curac2018
IN LEIPZIG
VERNETZT IN DIE ZUKUNFT
13.-15.09.

17. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V.

Tagungsband

Herausgeber:
Thomas Neumuth
Andreas Melzer
Claire Chalopin

CURAC 2018

Tagungsband

17. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V.

13.–15. September 2018, Leipzig

Impressum

Herausgeber:
Thomas Neumuth
Andreas Melzer
Claire Chalopin

Universität Leipzig
Medizinische Fakultät
Innovation Center Computer Assisted Surgery (ICCAS)
Semmelweisstraße 14
04103 Leipzig

Redaktion: Kathrin Scholz
Grafik: Jürgen Auge

ISBN: 978-3-00-060786-8

Gesellschaft

Deutsche Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V.
CURAC Geschäftsstelle
Albstraße 45
70597 Stuttgart

Ansprechpartnerin: Frau Gabriele Schäfer
Tel.: +49 711 76 54 219
E-Mail: geschaeftsstelle@curac.org
www.curac.org

Vorstand

Tagungspräsidenten

Thomas Neumuth
Andreas Melzer

Vorsitzende des Programmkomitees

Thomas Neumuth
Andreas Melzer
Claire Chalopin

Programmkomitee

Andreas Boehm, Leipzig
Oliver Burgert, Reutlingen
Jessica Burgner-Kahrs, Hannover
Claire Chalopin, Leipzig
Georg Eggers, Weinheim
Rudolf Fahlbusch, Hannover
Stefan Franke, Leipzig
Wolfgang Freysinger, Innsbruck
Hubertus Feußner, München
Michael Gessat, Lindau
Christian Hansen, Magdeburg
Horst Karl Hahn, Bremen
Florian Imkamp, Hannover
Lüder Kahrs, Hannover
Ron Kikinis, Bremen
Thomas Klenzner, Düsseldorf
Wolfram Lamadé, Überlingen
Omid Majdani, Wolfsburg
Jürgen Meixensberger, Leipzig
Andreas Melzer, Leipzig
Arya Nabavi, Hannover
Thomas Neumuth, Leipzig
Tobias Ortmaier, Hannover
Bernhard Preim, Magdeburg
Jörg Raczkowsky, Karlsruhe
Jörg Schipper, Düsseldorf
Alexander Schlaefer, Hamburg
Patrick Schuler, Ulm
Stefanie Speidel, Dresden
Simon Sündermann, Berlin
Jürgen Wahrburg, Siegen
Stefan Weber, Bern
Thomas Wittenberg, Erlangen
Stefan Zachow, Berlin

Grußwort

Liebe Teilnehmerinnen und Teilnehmer,

auch in diesem Jahr setzt die CURAC mit ihrer Jahrestagung ein Zeichen für den vertrauensvollen und belastbaren Austausch zwischen Klinikern und Medizitechnik-Experten.

Unsere Gesellschaft hat sich damit über die Grenzen des deutschsprachigen Raumes hinweg einen ausgezeichneten internationalen Ruf erarbeitet. Regelmäßig gewinnen wir führende Technik-Forscher und hochkarätige Kliniker aus aller Welt zur Teilnahme. In diesem Jahr freuen wir uns ganz besonders auf die Gastvorträge von Prof. Guang-Zhong Yang, Direktor des Hamlyn Robotic Centers London sowie Prof. Jens Rassweiler, Gründungsmitglied der Section of Uro-Technology (ESUT) der European Association of Urology (EAU), die mit ihrem Auftritt das Netzwerk der CURAC nachhaltig erweitern.

Ein großer technischer Fortschritt ist insbesondere in der Informatik, bei den Verfahren der künstlichen Intelligenz, in den Materialwissenschaften, bei Halbleitertechnologien, additiven Fertigungstechnologien und auch bei medizinischen Sensortechnologien spürbar. Unsere Aufgabe ist es, diesen Fortschritt für neue computer- und roboterassistierte Behandlungsverfahren zu nutzen und im Sinne des Patienten in klinisch effektive Lösungen umzusetzen. Dazu ist die Einbindung von Industriepartnern, etablierten Firmen und jungen Start-ups unerlässlich.

So passt auch das Motto „Vernetzt in die Zukunft“ der 17. Jahrestagung der CURAC-Gesellschaft zum Thema der technischen und klinischen Zusammenarbeit. Beinhaltet es doch zwei wichtige Aspekte, die uns als diesjährigem Veranstalter am Herzen liegen: Zum einen die intensive Verflechtung der Forschungsarbeiten zwischen Informatikern, Ingenieuren und Medizinern, so wie sie Kern neuer medizintechnischer Entwicklungen für die klinische Anwendung in unserem Forschungszentrum Innovation Center Computer Assisted Surgery (ICCAS) Leipzig ist; zum anderen die zukünftige, technologische Herausforderung der offenen Vernetzung von Medizingeräten und IT-Systemen im Operationsbereich – gemeinhin auch „Internet of Things“ genannt. Wir unterstützen durch die örtliche und thematische Nähe der Forschenden im Bereich der computer- und roboterassistierten Medizin auf dem Universitätsmedizincampus Leipzig den aktiven Austausch interdisziplinärer Forschungsimpulse. Unser integrierter und voll vernetzter Operationssaal ist ein praktisches Beispiel, wie vernetzte Strukturen bei uns in Leipzig zu erfolgreichen Medizintechnik-Innovationen führen können. Den Intelligenten Operationssaal präsentieren wir neben anderen hoch innovativen Forschungsprojekten beim Get Together am ersten Veranstaltungsabend im ICCAS.

Wir freuen uns darauf, unsere Erfahrungen mit den Technologien einer vernetzten Zukunft gemeinsam während der Tagung vorzustellen. 47 spannende Vorträge zu den Themenbereichen Bildverarbeitung, Quantifizierung und Visualisierung, Chirurgische Planung und Risikoanalyse, Navigation, Workflowanalyse und Prädiktion, Chirurgische Assistenzsysteme, Bildgestützte Systeme, Wissensbasierte Systeme und Entscheidungsunterstützung sowie Modellierung und Simulation bilden das wissenschaftliche Programm der ersten beiden Veranstaltungstage. Am dritten Tag findet im Rahmen der CURAC der beliebte Clinical Day statt, mit dem die Sektion für minimal-invasive Computer- und Telematik-assistierte Chirurgie der Deutschen Gesellschaft für Chirurgie (CTAC) zusammen mit dem CURAC e. V. besonders die Mediziner ansprechen möchte. Wir danken hier besonders Prof. Hubertus Feußner für seine Unterstützung.

Der CURAC-Festabend am zweiten Veranstaltungstag wird im Da Capo, der Leipziger Eventhalle mit einer der größten Oldtimersammlungen Mitteldeutschlands stattfinden. Umgeben von Glanz und Chrom aus der Zeit, als die Technik erst fahren lernte, werden wir einen fantastischen Abend verbringen, in dessen Rahmen die Überreichung der Best Paper Awards stattfindet.

Wir hoffen, die Veranstaltung bringt allen Teilnehmern viele Kommunikationsmöglichkeiten und Inspiration für das weitere Wirken und bedanken uns herzlich bei allen Autoren und Co-Autoren, den Referenten, dem Programmkomitee, dem Organisationsteam und den Sponsoren für deren Engagement.

Prof. Dr. Thomas Neumuth
Prof. Dr. Andreas Melzer
Tagungspräsidenten CURAC 2018

Prof. Dr. Stefan Weber
Präsident CURAC

Wir bedanken uns bei den Sponsoren der 17. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e. V.

Platinsponsor

Hill-Rom

Sibersponsoren

STORZ KARL STORZ — ENDOSKOPE

atracsys

AXIOS 3D SERVICES

Weitere Aussteller und Sponsoren

KUKA

NDI

PHACON 3D LIFE MODELING & SIMULATION

LOCALITE

Inhaltsverzeichnis

Session VI – Workflowanalyse und Prädiktion

Session VII – Navigation

Session VIII – Chirurgische Assistenzsysteme

Session IX – Bildgestützte Systeme

Session X – Wissensbasierte Systeme und Entscheidungsunterstützung

Vorträge CURAC

Streaming of 3D Data from Ultrasound Systems to Augmented Reality Glasses (HoloLens)

Felix von Haxthausen[1], Floris Ernst[1], Ralf Bruder[1], Verónica García-Vázquez[1]

[1]Institute for Robotics and Cognitive Systems, University of Lübeck, Lübeck, Germany

Contact: vonhaxthausen@rob.uni-luebeck.de

Abstract

Even though two-dimensional (2D) ultrasound (US) is one of the most common tools for diagnostic procedures, this imaging modality requires highly experienced and skilled operators to mentally reconstruct three-dimensional (3D) anatomy from these images and the physician's gaze is focused on the US system. Real-time 3D US in combination with augmented reality has the potential to overcome these problems. We propose an approach that allows streaming of volumetric data from the US system to the HoloLens and that renders the volume next to the US probe. Raw data access to the US system was provided by an in-house modification and a multi-language remote procedure call was used to transmit data to HoloLens. The volumetric data was rendered with a time interval of 72±55 ms and an end-to-end latency of 259±85 ms. Streaming 3D US data to HoloLens gives the possibility of a more intuitive visualization than using a standard screen.

Keywords: Real-time three-dimensional ultrasound, Mixed reality, AR marker

1 Problem

Two-dimensional (2D) ultrasound (US) imaging is a method extensively used for diagnostic and interventional tasks [1]. Nevertheless, this technique suffers from several disadvantages. First, the user mentally needs to fuse several 2D images in order to get an impression of three-dimensional (3D) anatomical structures. Second, out-of-plane motion results in imaging of different anatomical slices. Third, the relocation of the probe at the exact same position, as needed for monitoring the progression of pathology, is difficult [1]. Also, the physician's gaze is focused on the US system, missing the spatial relation between the position of the probe and the region of interest. Real-time 3D US (also referred as four-dimensional (4D) US) volumes displayed as a virtual object next to the US probe has the potential to overcome these problems. The Microsoft HoloLens is an optical see-through head-mounted computer that displays stereoscopic projections in the user's natural environment (www.microsoft.com/de-de/hololens). We present an approach that allows streaming of volumetric US data from the system to the HoloLens which in turn displays the volume next to the US probe.

2 Materials and Methods

A 4D cardiac US system (Vivid7, GE Healthcare) with a 3V matrix array probe and an in-house modification, that allows real-time access to the raw data, was utilized for volume data acquisition [2]. Volumes were sent from the US system to a computer via Ethernet and from there to HoloLens via Wi-Fi. An open-source multi-language remote procedure call (gRPC) was used for the second transmission path. A conversion from beam space to Cartesian space took place on the computer. An augmented reality (AR) marker was attached to the US probe to display the rendered volume relative to the transducer. Lossless data encoding was used to reduce the amount of required bandwidth. On HoloLens, the US volumes had an isotropic voxel size of 1.5 mm and were interpolated using the nearest-neighborhood method. A Universal Windows Platform (UWP) application was developed in Unity (Version 2018.1.0b9) for displaying the US volumes on the HoloLens. The application consists of the following parts: the gRPC receiving module, the decoding module, the volume rendering and the module for tracking the AR marker based on the open-source HoloLensARTToolkit [3].

This approach was evaluated regarding the frame rate of displayed volumes on HoloLens and the end-to-end latency. US volumes (depth of 15 cm, matrix size in beam space of 495 x 72 x 26) were acquired at a frame rate of 13.8 Hz. On HoloLens the volume had a matrix size of 103 x 74 x 134. The US system and the HoloLens were first synchronized using a Network Time Protocol (NTP) server. Hereafter, the timestamp of each sent volume from the US system was used to calculate the end-to-end latency on the HoloLens after volume rendering.

3 Results

During the evaluation, all volumes were successfully transferred from the US system to HoloLens. Figure 1 shows the physician's point of view where the US volume was displayed relative to the US probe. Successive volumes were displayed with an interval of 72 ± 55 ms (mean ± standard deviation). The end-to-end latency was 259 ± 86 ms. The lossless data encoding allowed a decrease of 63 % in data size.

Figure 1*: The rendered US volume above the transducer, which was placed on the FAST Ultrasound Training Model (Blue Phantom). The HoloLens displayed the US data in the physician's field of view.*

4 Discussion

The results showed that the mean interval between displayed volumes on HoloLens corresponded to the frequency when sending US data, thus the system is capable of processing and rendering all acquired volumes without increasing latency. The end-to-end latency is still too high for an instantaneous perception of the displayed volumes compared to the suggested threshold of 100 ms mentioned in [4]. A performance improvement regarding the conversion from beam space to Cartesian space and the gRPC module can result into lower latencies. The AR marker on the US probe allowed to place the rendered volume relatively to the transducer such that a translation and rotation of the probe resulted into the according transformation of rendered data. Future work will also include the evaluation regarding the minimum size of the marker and the stability of the rendered data. Since HoloLens offers voice commands and hand gestures, a future research topic will target the user interaction with 3D US data in order to resize, rotate or change the transfer function of the rendered volume. The technology also supports other ultrasound stations using the same gRPC service specification. It is currently implemented for the Vivid 7 as described here and Philips' EPIQ 7G station with the X6-1 transducer.

5 Summary

Streaming of volumes from the US system to the HoloLens and rendering next to the US probe was feasible by using the approach presented in this study. There is a growing interest in using 4D US for clinical applications and HoloLens gives the possibility of a more intuitive visualization of volumetric data than using a standard screen.

Funding

This study was supported by the German Federal Ministry of Education and Research (grant number 13GW0228B), the Ministry of Economic Affairs, Employment, Transport and Technology of Schleswig-Holstein and the German Research Foundation (DFG) (grant number ER 817/1-1).

References

[1] Fenster A, Parraga G, Bax J, *Three-dimensional ultrasound scanning*, Interface Focus, 1(4) 503–19 (2011)

[2] Bruder R, Ernst F, Schläfer A, Schweikard A, *A Framework for Real-Time Target Tracking in IGRT using Three-dimensional Ultrasound*, Int J CARS, 6(Suppl 1) S306–7 (2011)

[3] Qian L, *ARToolKit on HoloLens*, http://longqian.me/2017/01/20/artoolkit-on-hololens/, [Accessed: 21-May-2018]

[4] Miller RB, *Response Time in Man-computer Conversational Transactions*, Western Joint Computer Conference, San Francisco (1968)

Liver surface reconstruction from navigated ultrasound during image-guided liver surgery

Luca Sahli [1], Stefan Weber[1], Iwan Paolucci [1]

[1]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

Contact: luca.sahli@students.unibe.ch

Abstract

During navigated liver resections, most systems rely on registration of a preoperative model to the patient using surface or ultrasound landmarks. However, by opening the abdominal wall and mobilizing the liver, the shape is altered and therefore registration of a preoperative model becomes very challenging and error prone. To overcome this issue, an intraoperatively acquired model for planning of a resection could be used instead. As navigated ultrasound is commonly used in image-guided surgery, these images could be used to create an anatomical model and a surgical plan. In this study we evaluated an approach to reconstruct the liver surface based on intra-operative ultrasound.

Keywords: Liver Surgery, Navigated Ultrasound, Surgical Planning

1 Problem

Surgical resection is the gold standard for curative care for primary and secondary hepatic tumors. This procedure usually involves removing the segment of the liver where the tumor is located. In this treatment, it is important to spare enough healthy parenchyma to preserve the function of the liver after surgery. Therefore, non-anatomical resection approaches are becoming more popular, as they try to spare as much healthy tissue as possible. This way, only the tumor and a safety margin of 5 – 10 mm are removed which allows multiple resections and re-treatments in case of recurrence [1]. However, especially in these non-anatomical resections, maintaining the safety margin is challenging as the tumor is removed by cutting around the tumor in a conical or wedge shape rather than a plane along anatomical landmarks.

Therefore, image-guidance systems have been introduced to guide the surgeon to precisely follow a planned resection path. These systems rely on tracking devices (optical or electromagnetic) to measure the pose of the surgical instruments and use a registration process to align a preoperative model with the patient intraoperatively [2, 3]. However, the setup and use of such systems is time consuming, complex and requires extensive training, which is a major reason why they are not widely used [4]. Additionally, the registration process introduces errors due to organ deformation between the image acquisition and the surgery.

During conventional, non-anatomical resections a resection plan is drawn onto the liver before the start of the resection. Therefore, an important part of the resection plan is an accurate reconstruction of the liver surface. This surface is then used to project the outline of the tumor and a safety margin onto the surface. This is where the surgeon would start with the transection of the parenchyma. Previous work used such surface reconstructions based on laser scanners [5] for intraoperative registration, which requires additional equipment. In this study, we evaluated an ultrasound (US) based method to automatically reconstruct the liver surface intraoperatively.

2 Material and Methods

The image processing pipeline consists of three steps, the acquisition, the contact detection and the surface reconstruction (Fig 1). First the data from an ultrasound scanner (Flex focus 800, BK Medical, Denmark) and a tracking camera (Polaris, NDI, Canada) is acquired and fused on a navigation system (CAS-One Vario, CAScination AG, Switzerland) for liver surgery. Then each image is classified whether it has contact to the liver surface or not. The position of the images with contact to the liver surface are then further processed in the surface reconstruction step to create a model of the liver surface. The result is then visualized in a 3D viewer.

Fig 1: The data processing pipeline

2.1 Acquisition

During the data acquisition phase, the ultrasound image and the corresponding 6D pose are recorded using the navigation system. The ultrasound was calibrated using a Z-wire phantom [6] and is tracked by an optical tracking system. To simulate a liver surgery, a multimodal liver phantom (Fig 2) and an intraoperative ultrasound was used to get the ultrasound images. During the simulation the ultrasound device had a trackable and calibrated marker attached.

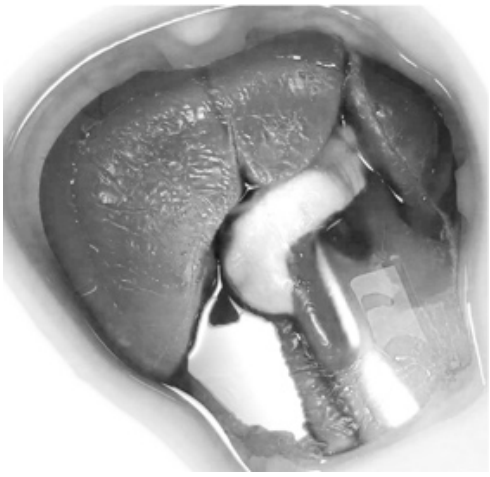


Fig 2: The US liver phantom used for the experiments in this study

To find an optimal sampling method, the liver was scanned with six different techniques (Fig 3: Different sampling movements of the ultrasound device over the surface of the liver.Fig 3). The two spiral techniques represent recordings of moving the US device to draw a spiral onto the liver. Either from the center to the peripheral part (spiral out) or vice versa (spiral in). The two sweep techniques represent recordings of moving the US device left and right (Sweep LR) or forward and backward (Sweep FB). The flower technique represents a recording of moving the US-device to draw a flower onto the liver. Additionally, a point grid was acquired as reference points for evaluation of the other reconstructions.

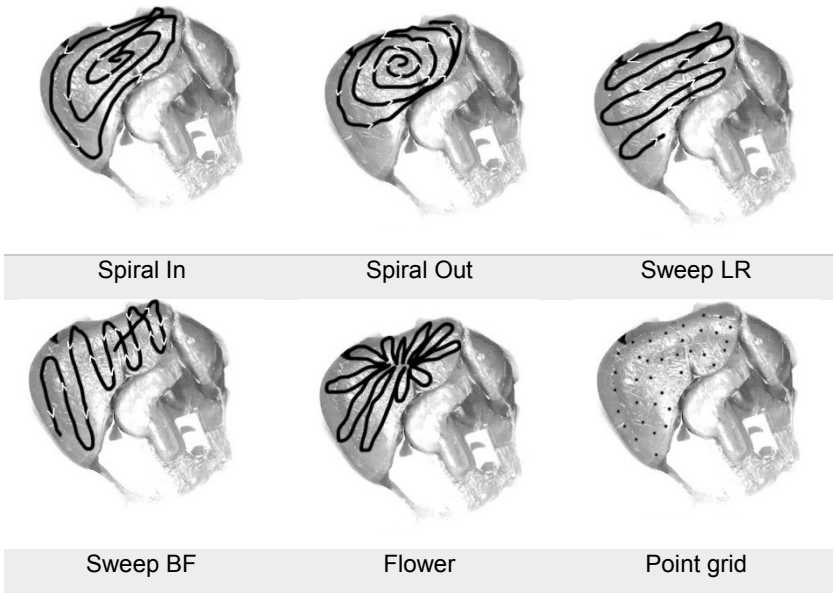


Fig 3: Different sampling movements of the ultrasound device over the surface of the liver.

2.2 Surface Contact Detection

In the surface contact detection step, a classifier detects whether the US probe has contact to the liver or not. Therefore, a support vector machine (SVM) was trained with US images from the phantom and from previous navigated liver surgeries. The SVM was trained to classify the image into "no surface contact" (left) and "surface contact" (middle and right). The images were labelled as "surface contact" if at least 50% and the center had contact to the surface (Fig 4 middle). The classifier takes into account that US waves are reflected at the US probe-air interface when the US probe has no contact to the liver and therefore no image is formed.

Fig 4: Three ultrasound images from left to right: No contact with the liver, difficult to decide (In this case it would be contact because the middle part of the image shows contact), contact with the liver.

The features for the classifier were: mean, median, minimum, maximum, variance, skewness and kurtosis of the pixel values. All features are calculated on the upper half of the image. For training, a set of 2'311 images (1'056 with contact, 1'255 without contact) were used. The training data was composed of images from a phantom (88%) and images from previous navigated liver surgeries (12%). All computations were performed using the SciPy software package.

2.3 Surface Reconstruction

To reconstruct the surface of the liver from the sampled points, the surface reconstruction algorithm by Hoppe et al. [7] was used. The algorithm consists of three phases. From an unorganized set of points, phase 1 constructs an initial dense mesh. Starting with the dense mesh created in phase 1, phase 2 reduces the number of faces and improves the fit to the data points. In phase 3, the surface representation is changed from a piecewise linear one (meshes) to a piecewise smooth one. For the computations the implementation in VTK (*SurfaceReconstructionFilter)* was used (neighborhood size of 50 and sample spacing of 10). Due to the different latencies of the US and the tracking system (with the US being slower), a delay of 4 frames (0.2 seconds) is applied to the tracking data.

2.4 Experimental Evaluation

For evaluation of the surface detector the data was split into training (80%) and test data (20%). The precision and recall were calculated for performance analysis on the test set. To quantitatively evaluate the reconstructed surfaces, the points of the point grid measurement (69 points) were used as a reference. These reference points represent points on the surface of the liver in an undeformed state. For each of these reference points, the error is computed as the shortest distance to the reconstructed surface. All computations were performed using SciPy.

3 Results

Overall, the surface contact detector was evaluated on a test set with 2414 images. Additionally, 14 scans of the liver surface were evaluated against the reference points to measure the accuracy of the surface reconstruction.

3.1 Surface Contact Detection

To evaluate the contact detector, a test set of 2414 images with 50% contact and 50% no contact was used. The detector has a sensitivity of 0.95 and a specificity of 0.98. Out of all negative samples, 1.9% were detected as having contact. The prediction of one image takes 15 ms where most of the time (approx. 99%) is spent for feature extraction.

3.2 Surface Reconstruction

Visual Assessment

The reconstruction of the liver surface created lead to a smoothed version of the measured surface part. The measured surface corresponds to the surface of the liver (Fig 2). However, the reconstructed surface area is larger than the sampled part of the surface. This is a property of the algorithm, as it estimates a rectangular grid.

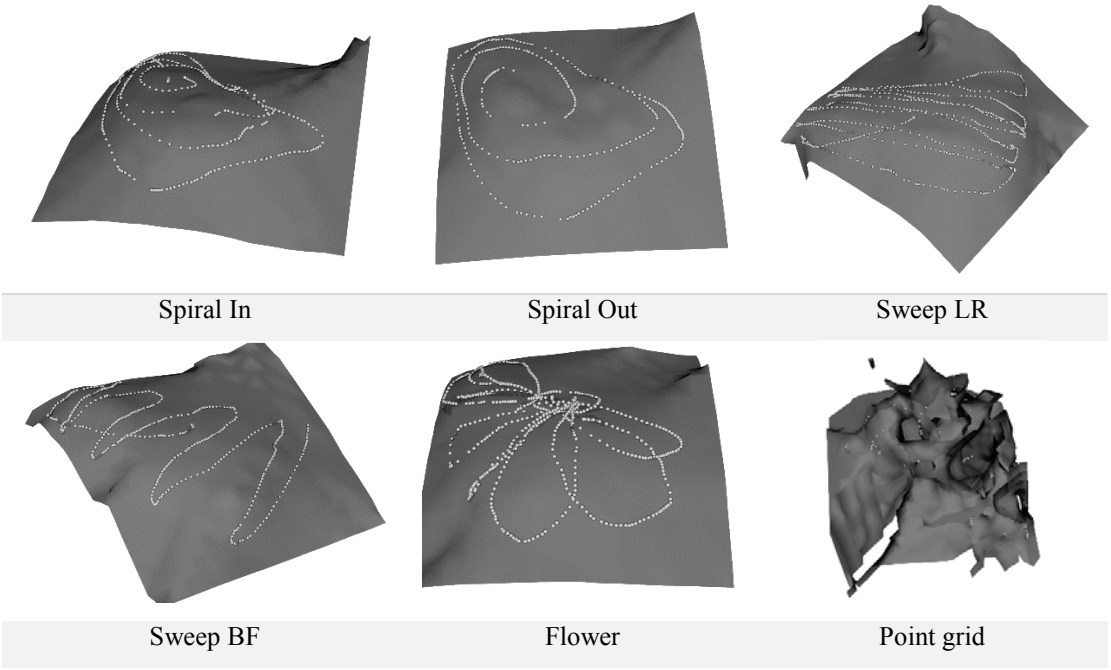


Fig 5: Reconstructed surfaces from different movements of the ultrasound device over the surface of the liver model. The names correspond to the movement drawings (Fig 3).

Quantitative Analysis

To evaluate the accuracy of the reconstructed surfaces, the shortest distance of the reference points to the surface were computed. The overall median error for all the measurements is 2,5mm with an interquartile range of 1 mm – 5 mm. By projecting these errors corresponding to each reference point onto the liver surface, one can see that the highest errors are in segments 2 and 3 And the lowest in segments 4, 5, 6 and 8 (Fig 6).

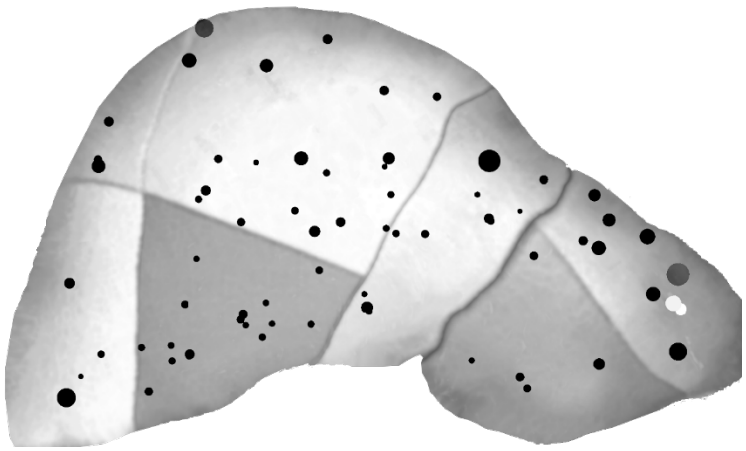


Fig 6: The errors at the reference points projected onto the liver surface. The size of the points corresponds to the median error at this reference point.

4 Discussion

The surface contact detector is correctly classifying in 96% of the cases, with a very low false positive rate of 1.9%. This is especially important, as false positives lead to artifacts in the reconstructed surface. Furthermore, the processing time of 15 ms per image makes it suitable for real-time processing of the images as the ultrasound scanner runs at 20 Hz (50 ms / frame). When the US probe is removed from the liver surface there are 3-5 images which are wrongly classified as having a signal. This would cause artifacts in the surface reconstruction, and therefore they are filtered out later for surface reconstruction. This is mainly, because of the latency of the US scanner itself compared to the tracking system. The images are slightly blurrier (Fig 7), but from the tracking positions one can clearly see that they are not on the surface.

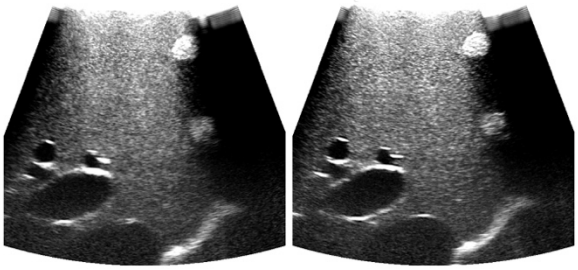


Fig 7: Wrong and correct classified images at the end of the measurement.

4.1 Surface Reconstruction

From a visual point of view the reconstructed surfaces of the US liver phantom look similar to the surface of the liver model. However, the spiral and the flower movement, led to a more accurate reconstruction.

From a quantitative point of view, it turned out that the largest average errors are in segment 2 and 3. This is likely because these liver segments are the softest on the phantom used for the measurements. Due to that, this segment was pressed down during the measurement which lead to a surface with a large distance to the reference points. Additionally, one can see that the average distance at the boundary of the liver is large as well. This could be because the US device was held between the wall of the tank and the liver model. Because of the small space between the wall and the liver, the pressure applied on the liver was larger than in other areas and the consequent distance between the reference points and the deformed surface became larger. However, this might also be the case in a clinical setting, as these regions are more difficult to reach with the US probe. Overall, the best accuracy, could be achieved in segments 4,5,6 and 8, which were the easiest to access in this setup. In a next step, this would also be analyzed on the human liver, to see in which segments this technique can be applied accurately.

To conclude, we presented a surface reconstruction technique, which can be used to intraoperatively acquire a surface model of the liver using navigated US. This can then be further used for intraoperative resection planning or surface-based registration.

References

1. Aghayan DL, Pelanis E, Avdem Fretland Å, Kazaryan AM, Sahakyan MA, Røsok BI, Barkhatov L, Bjørnbeth BA, Jakob Elle O, Edwin B (2018) Laparoscopic parenchyma-sparing liver resection for colorectal metastases. Radiology and Oncology 52: . doi: 10.1515/raon-2017-0046

2. Langø T, Vijayan S, Rethy A, V\a apenstad C, Solberg OV, M\a arvik R, Johnsen G, Hernes TN (2012) Navigated laparoscopic ultrasound in abdominal soft tissue surgery: technological overview and perspectives. International journal of computer assisted radiology and surgery 7:585–599

3. Banz VM, Müller PC, Tinguely P, Inderbitzin D, Ribes D, Peterhans M, Candinas D, Weber S (2016) Intraoperative image-guided navigation system: development and applicability in 65 patients undergoing liver surgery. Langenbeck's Archives of Surgery 401:495–502 . doi: 10.1007/s00423-016-1417-0

References

1. Aghayan DL, Pelanis E, Avdem Fretland Å, Kazaryan AM, Sahakyan MA, Røsok BI, Barkhatov L, Bjørnbeth BA, Jakob Elle O, Edwin B (2018) Laparoscopic parenchyma-sparing liver resection for colorectal metastases. Radiology and Oncology 52: . doi: 10.1515/raon-2017-0046

2. Langø T, Vijayan S, Rethy A, V\a apenstad C, Solberg OV, M\a arvik R, Johnsen G, Hernes TN (2012) Navigated laparoscopic ultrasound in abdominal soft tissue surgery: technological overview and perspectives. International journal of computer assisted radiology and surgery 7:585–599

3. Banz VM, Müller PC, Tinguely P, Inderbitzin D, Ribes D, Peterhans M, Candinas D, Weber S (2016) Intraoperative image-guided navigation system: development and applicability in 65 patients undergoing liver surgery. Langenbeck's Archives of Surgery 401:495–502 . doi: 10.1007/s00423-016-1417-0

4. Kingham TP, Jayaraman S, Clements LW, Scherer MA, Stefansic JD, Jarnagin WR (2013) Evolution of Image-Guided Liver Surgery: Transition from Open to Laparoscopic Procedures. Journal of Gastrointestinal Surgery 17:1274–1282 . doi: 10.1007/s11605-013-2214-5

5. Simpson AL, Kingham TP (2016) Current Evidence in Image-Guided Liver Surgery. Journal of Gastrointestinal Surgery 20:1265–1269 . doi: 10.1007/s11605-016-3101-7

6. Peterhans M, Anderegg S, Gaillard P, Oliveira-Santos T, Weber S (2010) A fully automatic calibration framework for navigated ultrasound imaging. IEEE, pp 1242–1245

7. Hoppe H (1995) Surface Reconstruction from Unorganized Points. PhD Thesis, University of Washington

Panoramic Endoscopy of the Stomach: First results from Phantom and Patient Data

R. Hackner[1], A. Hann[2], D. Franz[1], A. Meining[2], M. Raithel[3], T. Wittenberg[1,4]

[1]Fraunhofer-Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Universtiätsklinikum Ulm, Ulm, Germany
[3]Malteser Waldkrankenhaus St. Marien, Erlangen, Germany
[4]Chair of Computer Graphics, University Erlangen, Germany

Contact: ralf.hackner@iis.fraunhofer.de

Abstract

For gastroscopic examinations of the stomach, the limited field of vision related to the keyhole view of the endoscope can be a problem. A panoramic view of the gastroscopic images can enlarge the field of vision and may ensure that the performing physician has examined all of the stomach walls. Nevertheless, the complex geometry the stomach renders does not allow a straight forward stitching or mosaicking of a panorama.

Therefore, for primary experiments in the field of panoramic endoscopy of the stomach we subdivided the stomach in four areas (fundus, antrum, lesser and greater curvature of the body of the stomach). We adapted methods for panoramic bladder-mosaicking in order to stitch endoscopic image-sequences from two different phantoms of the stomach and a real gastroscopic examination. With the obtained "panoramic map", the physician can ensure that the entire stomach was examined, document and annotate findings and communicate these to other physicians.

Keywords: stomach, panorama, endoscopy, stitching, mosaicking, registration, view enhancement

1 Problem

One core problem in diagnostic and interventional endoscopic interventions including gastroscopy (endoscopic examination of the upper GI-tract, including esophagus, stomach and duodenum) is the limited "field of view" through an endoscope, also known as "keyhole view".

At the beginning of a gastroscopic examination, the endoscope is advanced through the patient's mouth into the esophagus, entering the stomach through the cardia and is moved through the stomach until the duodenum is reached. During this process, the physician does not look at closer details of the surface, but concentrates on the mechanical insertion of the endoscope into the lumen. The examination happens in a second step, when the endoscope is slowly withdrawn through the pylorus, passing the pyloric antrum and the body of the stomach (cf. Figure 1). In the stomach body, the physician bends the endoscope for a closer examination of the surface. In the end of the examination, the endoscope tip is retroflexed in order to allow an examination of the fundus and the cardia.

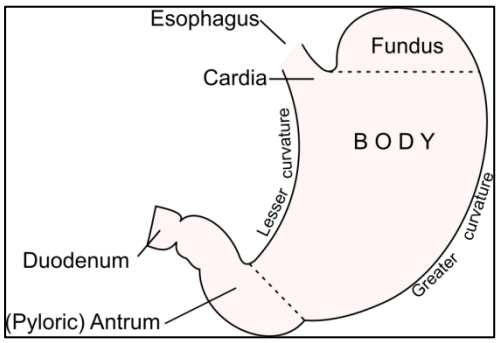


Figure 1*: Overview to the regions of the stomach. Based on [1], modified*

For the clinical documentation of an endoscopic examination of the upper (or lower) gastro-intestinal tract, a number of still-images photos or video-recordings are usually made. On one hand, still photos are not capable to show the entire surface of the organ lumen and the depicted areas are not comparable. On the other hand, it usually takes much time to find a certain area of interest in an endoscopic video sequence.

One approach to solve the above mentioned problem of the "key-hole-view" is to create a so-called "panorama image" of the examined body lumen, such as the stomach. Therefore, particular images captured by the endoscope are combined to a single panorama image by an image processing system. The resulting image can hence be regarded as "a map-like structure" of the surface of the examined hollow or at least significant parts of it. For hollow organs with a more simple inner geometry such as the spherical bladder, solutions already have been proposed, which are able create such image panoramas in real time [2] [3].

State of the Art

For the computation of panoramic images of the stomach from endoscopy image sequences, hardly any work is known.

Liu et al. [4] have recently introduced an approach of creating a complete gastric panorama. To address the problem of the 6-DOF freedom of the endoscope tip in gastroscopy, they mounted an additional electric magnet (EM Tracker) at the tip of the endoscope, in order to track the motion and orientation of the endoscope tip inside the gastric hollow with external sensors placed around the patient. Moreover, they suggested a method to project the stomach geometry onto a plane.

Ali [3] suggested the computation of a panorama of the region around the pyloric antrum (cf. Figure 1) with an endoscopic system, that does not need additional tracking sensors, but uses the optical flow between successive image frames from the input sequence for tracking. This approach has originally been developed for panoramic cystoscopy, and has only been transferred exemplarily to one gastroscopic sequence.

Capsule endoscopes (so called "PillCams") are an alternative approach for the assessment and examination of the gastrointestinal tract. Such acquired "PillCam"-sequences have been used by varying research groups to create image panoramas [5] [6] [7] of the stomach. The control over the image recording process is limited, since the currently commercially available capsule endoscopes are not able to move actively. Nevertheless, there are some attempts to control their motion in the stomach by magnetic fields [8].

2 Methods and Material

Methods

Our research group has recently developed a software system that allows creating a large-scale image panorama of the bladder during a cystoscopy in real time called, referred to as "Endorama" [2]. This system can capture a live video stream from an endoscopy system, to fuse the acquired images to an image panorama, and to display the resulting panorama to a control screen, that the physician gets a real-time feedback during the examination, if the entire surface of the bladder is captured in sufficient quality for stitching (cf. Figure. 2).

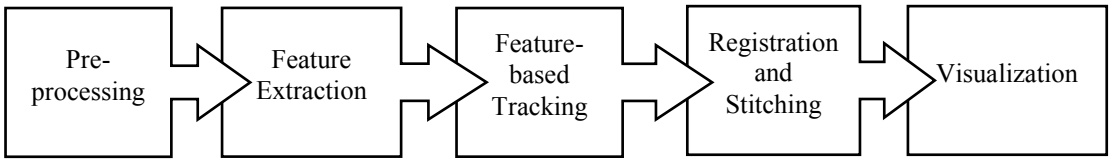


Figure 2*: Panorama generation overview*

The "Endorama" system captures images from the endoscope system with a frequency of 25 to 60 Hz, depending on the data source. A number of preprocessing steps is applied on every frame [9] (cf. Figure 2): At first a circular aperture mask typical for endoscopic images is applied. A color shading correction algorithm [10] is applied to handle inhomogeneous illumination, caused by the point light source used in endoscopy. Furthermore, lens distortion is corrected based on a previously recorded checker board pattern.

The preprocessing step is followed by a feature extraction step, where the system extracts a number of robust features based on the DOP algorithm [2] [9], which can be used as landmarks.

These key-points are handed over to a tracking module that estimates correspondence between the features from one frame to another by using a RANSAC scheme [11]. Based on this information, the distance between two

successive frames is estimated, so that the position of the endoscope tip can be tracked relative to the surface of the body lumen and the starting position.

Based on this tracking and correspondence information images are registered against each other in the next step. Only frames with a sufficient distance are registered, while frames that do not add additional information (e.g. are located close to another frame) are be discarded. While this tracking based registration step is performed in real time, a second registration process runs in parallel to perform a global optimization, which also includes a loop detection method [2].

The system has currently been optimized for a spherical geometry, a fully illuminated small field of vision, no abrupt changes in the angle of view, and no occlusion of parts of the surface by other structures.

"Endorama" is written in C++, using OpenCV and requires CUDA support to achieve real-time processing of a live stream or video.

Materials

To apply this approach to endoscopic image data obtained from the stomach, we had to mitigate the problem, that the geometry of the stomach is more complex than the bladder. Instead of stitching the entire stomach surface at once, we defined a number of sub-regions in the stomach with geometries that fit the current limitations of our system better and respect the anatomy of the stomach (cf. also Fig. 1): the fundus (and the cardia), the pyloric antrum, and the greater and the lesser curvature with the adjacent parts of the stomach wall. These regions are enlisted in the order of similarity between their geometry and the geometry of the bladder.

- The **fundus** and the adjacent cardia is a widely spherical region and the orientation of the endoscope tip is roughly perpendicular to the surface, so there are similar geometric conditions as in the bladder. The organ itself does not occlude any part of the fundus, but parts of the cardia are hidden behind the endoscope, that enters the stomach here.
- The **pyloric antrum** is delimited by sphere-like region, but is slightly stretched. Under certain conditions parts of the antrum might be slightly occluded by a notch. The surface texture of this region is not as distinct as in the fundus and wide parts of the stomach body; therefore finding landmarks for image registration is harder.
- The **greater curvature** has a concave shape, but the angle of vision alters.
- The **lesser curvature** is the only convex region and the angle of vision alters massively.

To test our stitching approach, we recorded image sequences of each of the described regions (fundus, antrum, curvatures) in a silicon phantom of the stomach (Phantom A: Tübingen silicon stomach phantom [12] inserted into a CLA torso, cf. Fig. 3, left) with a Storz 13924 PKS gastroscope. The previously described "Endorama" software processed the image sequence during the recording and gave direct visual feedback to the examiner about the quality of the recording as well as the panorama. A second stomach phantom (Phantom B: Erler-Zimmer Phantom, see Fig. 3, right) was recorded with an Olympus GIF HQ190 endoscope without visual feedback from the system. Both phantoms differ in their size and the texture of the inner surface.

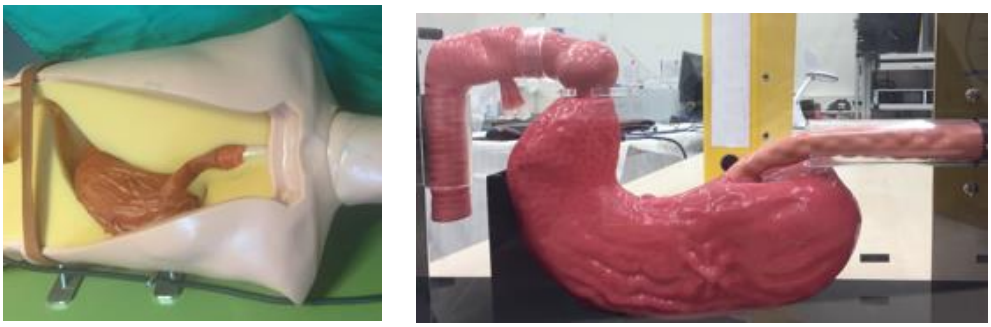


Figure 3*: left side: "Phantom B" (Tübingen silicon stomach phantom in a CLA-torso), right side: "Phantom A" (Erler-Zimmer-Phantom)*

Moreover, we used a collection of video sequences of one routine gastroscopic examination, recorded with an Olympus GIF HQ180 endoscope. Here the physician had no visual feedback, whether the recorded material would be sufficient for the generation of a panorama.

For the phantoms every region was recorded separately. To create panorama images for the real patient data, we manually selected and cropped subsequences of the video recording that show the required parts of the stomach.

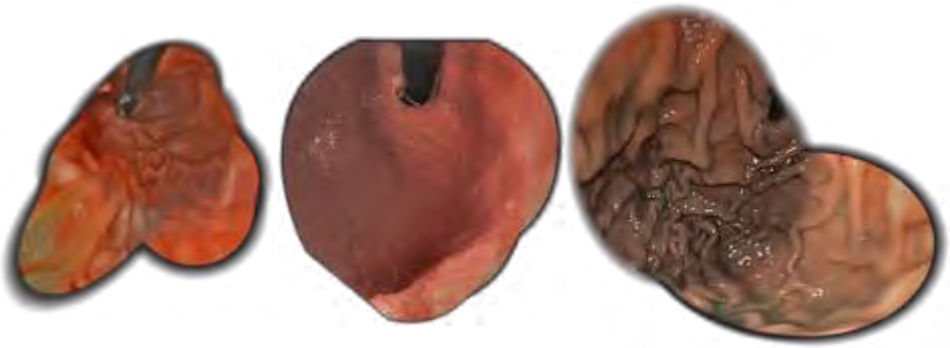


Figure 4*: Panoramic maps of the fundus region of phantom A (left), phantom B (middle), and patient (right) consisting of 21 tiles, 19 tiles, and 10 tiles, respectively.*

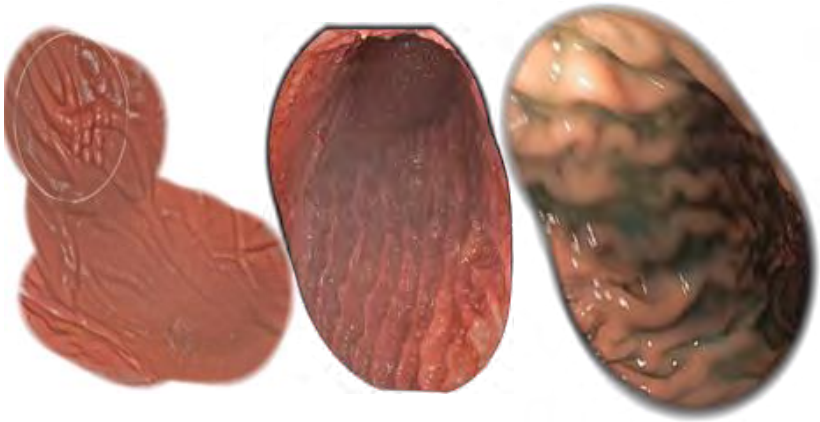


Figure 5*: Panoramic maps of the pyloric antrum of phantom A (left), phantom B (middle), and patient (right) consisting of 23 tiles, 68 tiles, and 22 tiles, respectively*

Figure 6*: Panoramic fragments of the body of the stomach along the greater curvature of phantom A (left), phantom B (middle), and patient (right) consisting of 32 tiles, 68 tiles, and 6 tiles.*

3 Results

Computed panoramic images of the fundus, the pyloric antrum and parts of the greater curvature obtained from endoscopic image sequences of the two phantoms and the patient data are depicted in Figures 4 – 6. For the **fundus region** we managed to create a panorama that shows nearly the entire region. Also, was possible for both the phantom video sequences as well as the recording of the patient. The panorama images are composed of 10 - 21 single image tiles each, neglecting all frames which do not add new information to the panorama. The

endoscope entering through the cardiac antrum can be seen in the top of the image due to the retroflexed endoscope tip (Figure 4).

Creating a panorama of the **pyloric antrum** was also possible for all three cases, even though the fraction of the panoramic region that we were able to compute, differ in the three cases. In the case of the two phantoms nearly the entire region is covered by the panorama, while we could only stitch a fraction of it in the recording of the real patient (Figure 5).

Creating a partial panorama of the body of the stomach along the **greater curvature** was also possible. In case of the two phantoms, the panorama images consist of up to 32 single tiles for the first phantom and up to 68 tiles for the second phantom. Stitching the real patient data was much more complicated, as there were challenges with fast movements of the endoscope tip (Figure 6). Moreover, the tissue surface often left the focal range of the endoscope. Currently, we were not able to create panorama images of the region around the **lesser curvature** with sufficient quality.

All panoramas were generated by our "Endorama" software. The calculations were performed in real-time, so the calculation time of a panorama corresponds to the playback time of the video or live stream used.

4 Discussion

For the **fundus** region we managed to create panoramas that depict nearly the entire region, based on the phantom video sequences as well as the real world patient data. As expected the current tracking and stitching algorithm fit quite well on this hemispheric regions. One disruptive factor is the problem that the endoscope itself is visible due to the retroflexed tip and occludes part of panorama, so that the stitching process is partially interrupted at the moment. The border regions are partially missing, since the angle of view becomes too low for the system.

In the case of the **antrum**, the enlargement of the field of vision was lower than originally expected. Again the shrinking angle of view and – at least in the real patient data – a change of perspective due to the notch in the antrum caused the tracking and registration algorithm to lose the correspondence after a while. In the case of real patient's data, there is also the special issue that the pylorus will automatically close, as soon the endoscope passes through. Thus, this automated sphincter muscle contraction may cause problems in the stitching process. In the phantoms the notch in the antrum was less distinctive as in the real world patient data. So it might be necessary to evaluate additional phantoms for our future research, to cover additional variations of the stomach geometry.

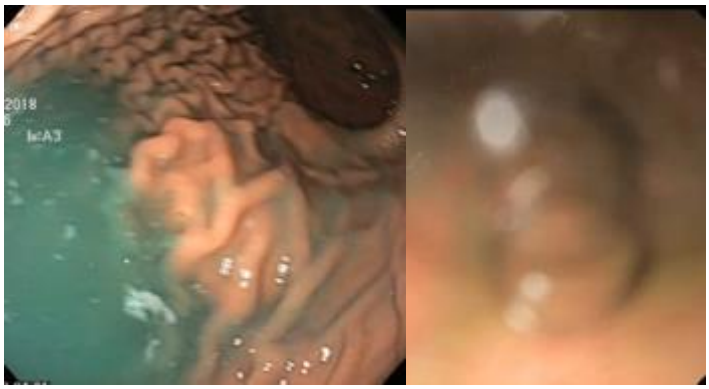


Figure 7*: Disruptive factors: remains of gastric juice (left), bubbles during cleaning (right)*

As depicted in Fig. 7 (left), a general problem using real world patient data are disruptive factors such as the remains of gastric juice hiding parts of the tissue surface as well as bubbles that appear during the cleaning procedure, when the physician is removing these remains.

Another problem is the blur (see Fig. 7, right) due to the low focusing range of the endoscope and fast motions of the endoscope tips. Under these conditions, it was not possible to extract enough stable und robust image features in order to track the endoscope tip yet and to register the acquired images with each other afterwards. Most likely, this problem will be reduced, when it is possible to provide real-time visual stitching feedback to the endoscopist during the endoscopic examination

5 Conclusion

We aim to reduce the problem of the limited field of view during examinations in the field of gastroscopy and thus try to improve the documentation of the results by creating panorama images of the stomach. To mitigate the complex geometry of the stomach we have subdivided the stomach in four distinct areas, regions or fractions of them. It was possible to create (partial) panorama images for most regions by using common endoscopes in the examined test cases. Parts of the body of the stomach are still missing in the panorama. The tracking of the endoscope tip was done by using only visual information delivered by the endoscope. Due to the complex geometry and various disruptive factors, the tracking and image registration process need improvements to handle such interruptions and some geometric conditions. The robustness of the methods must be tested with additional real world data.

Acknowledgement

This work has partially been supported by the BMBF grant “Interactive visualization of 2D/3D panorama endoscopy data of the bladder (03ZZ0444E).

The authors would like to thank Prof. Dr. med. Karl-Ernst Grund and his team, who provided us with a copy of the Tübingen silicon stomach phantom.

References

[1] *https://commons.wikimedia.org/wiki/File:Gray1046.svg*

[2] T. Bergen: *Real-time Endoscopic Image Stitching for Cystoscopy*. PhD Thesis, Univ. Koblenz-Landau. (2017).

[3] S. Ali: *Total variational optical flow for robust and accurate bladder image mosaicing*. PhD Thesis, Université de Lorraine (2016)

[4] J. Liu, B. Wang, W. Hu, *Global and Local Panoramic Views for Gastroscopy: An Assisted Method of Gastroscopic Lesion Surveillance*, IEEE Transactions on Biomedical Engineering, 62(9) 2296-2307 (2015)

[5] Y. Fan, M. Meng, B. Li, *3D reconstruction of wireless capsule endoscopy images*, Procs 2010 IEEE Ann. Int. Conf. Eng. in Medicine and Biology Society (EMBC), Buenos Aires (2010)

[6] M. Turan, Y. Pilavci, I. Ganiyusufoglu, H. Araujo, E. Konukoglu, M. Sitt, *Sparse-then-dense alignment-based 3D map reconstruction method for endoscopic capsule robot.*, Machine Vision and Applications, 29(2) *345-359* (2018)

[7] L. Maciura, J. G. Bazan, *Granular computing in mosaicing of images from capsule endoscopy*, Natural Computing, 14(4) 569-577 (2015)

[8] J.-F. Rey, H. Ogata, N. Hosoe, K. Ohtsuka, N. Ogata, K. Ikeda, H. Aihara, I. Pangtay, T. Hibi, S.-E. Kudo, H. Tajiri, *Blinded nonrandomized comparative study of gastric examination with a magnetically guided capsule endoscope and standard videoendoscope,* Gastrointestinal Endoscopy, Bd. 75(2) 373-381 (2012)

[9] M. Bergler, M. Weiherer, T. Bergen, M. Avenhaus, D. Rauber und T. Wittenberg, *Stitching pathological tissue using DOP feature tracking.* Bildverarbeitung für die Medizin, 322 -327 (2018)

[10] C. Münzenmayer, F. Naujokat, S. Mühldorfer und T. Wittenberg, *Enhancing Texture Analysis by Color Shading Correction*, 9. Workshop Farbbildverarbeitung (2003)

[11] T. Bergen, P. Hastreiter, C. Münzenmayer, M. Buchfelder und T. Wittenberg, *Image stitching of sphenoid sinuses from monocular endoscopic views*, Proc’s 12. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie (2013).

[12] K.-E. Grund, *Tübinger Endo-Trainer „SUSI”,* Perspektiven der Chirurgie im 21. Jahrhundert. Deutsche Gesellschaft für Chirurgie (2000)

Automatic Detection of Commissures in Mitral Valve Geometry

Lennart Tautz[1,2], *Mathias Neugebauer*[1], *Isaac Wamala*[3], *Simon Sündermann*[2,3], *Volkmar Falk*[2,3], *Anja Hennemuth*[1,2]

[1] *Fraunhofer MEVIS, Bremen, Germany*
[2] *Charité - Universitätsmedizin Berlin, Berlin, Germany*
[3] *German Heart Institute Berlin - DHZB, Berlin, Germany*

Contact: lennart.tautz@mevis.fraunhofer.de

Abstract Mitral valve diseases are among the major causes of heart failure. For surgical or interventional treatment of an impaired valve, detailed measurements of the valve geometry are necessary to support the surgeon in the therapy decision. We propose a labeling and quantification method that separates a valve surface model into anterior and posterior leaflet as well as leaflet scallops. We evaluated the clinical utility of our approach by comparing commissure positions and derived parameters between automatically labeled valve geometry and an expert definition of commissure points in 10 clinical image volumes. The differences in commissure position and derived parameters are comparable to inter-observer variation.
Keywords: mitral valve, commissures, quantification, geometry

1 Problem

Mitral valve regurgitation is the second most common heart valve disease and is associated with shortness of breath, impaired physical performance and consecutive heart failure.

The mitral valve is located between the left atrium (LA) and the left ventricle (LV). The valve apparatus consists of the mitral annulus, to which the anterior and posterior leaflets are attached, and the chordae tendineae which connect the leaflets to the papillary muscles (Fig. 1 (a)). During the diastolic phase blood flows through the open valve into the LV, while during the systolic phase the closed valve prevents blood from flowing back. There are different mechanisms that are the underlying cause for mitral valve regurgitation, for example chordal elongation or rupture, defects in the leaflets or annular dilatation. Blood then regurgitates into the atrium during the systolic phase, impairing heart function and increasing preload on the LA. In patients with mitral valve stenosis, the valve opening is narrowed, increasing pressure in the LA and the pulmonary vessels during the diastolic phase.

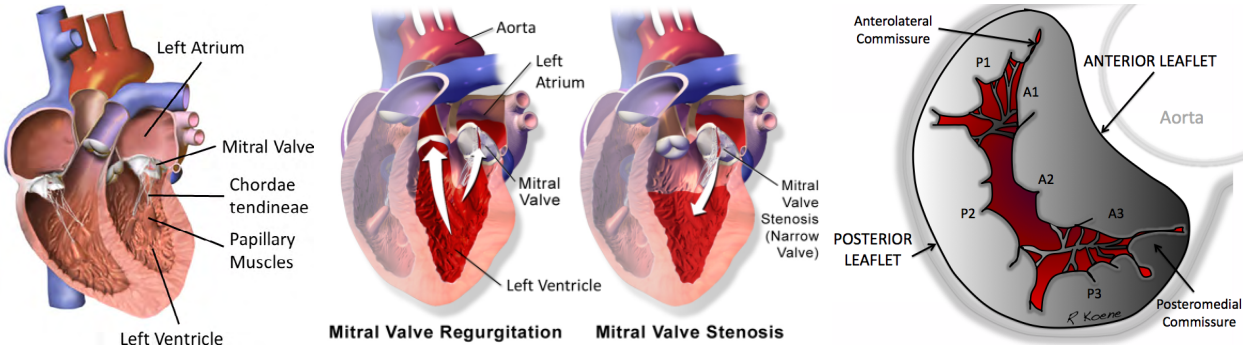


Figure 1: *Illustration of mitral valve anatomy (a)*[1]*, mitral valve pathologies (b)*[1]*, and mitral valve commissures and subdivisions (b)*[2]*.*

For mitral valve regurgitation, surgical mitral valve repair is the gold standard therapy resulting in better long-term survival compared to mitral valve replacement. Options to repair the valve are leaflet reconstruction, chordae replacement and annuloplasty. The choice of therapy is made based upon pre-operative imaging, as well as on the surgeon's personal experience and preference. Clinical imaging modalities such as echocardiography (US) and computed tomography (CT) are used to visually assess the valve, and to perform measurements such

[1]Images adapted from: Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". WikiJournal of Medicine 1 (2): 10. doi:10.15347/wjm/2014.010. ISSN 2002-4436)
[2]Image from public domain

as quantification of the annulus diameter. A detailed model of the valve apparatus can be extracted from such image data, enabling quantification of parameters like coaptation height and length, leaflet angulation in relation to the mitral valve annulus, and valve opening area. These parameters can support therapy planning by providing a more detailed view on the patient's mechanism of pathology. The goal of this work is the objective and reproducible commissure detection and quantification of the detailed geometry of the valve apparatus from an image-based, modality-independent valve model and context information. As medical text books as well as clinical analysis software that experts are familiar with commonly employ the labeling scheme shown in Fig. 1 (c), it is beneficial to use the same scheme for documentation and reporting of the quantitative assessment.

1.1 Related Work

Approaches for the detection of commissures can be grouped into methods with explicit commissure point detection, methods which derive the points from model and context information, and manual methods.

Explicit detection methods include multi-layered learning-based shape models that contain models of the anterior and posterior leaflets as well as of the commissures [3, 7]. A related approach used manual positioning of the commissures due to the lack of training data [2]. These models were evaluated extensively against manual annotations of the contained landmarks. A commercial solution offered by the Philips Ultrasounds Mitral Valve Quantification software has been noted to be accurate in general, but lacking in more complex or unusual geometry [1].

Implicit detection methods include a multi-atlas segmentation approach with separate atlas labels for anterior and posterior leaflets [9], and fitting of a shape model with components for both leaflets [8]. The atlas approach was cross-validated against the manually segmented data sets used for atlas creation.

There have been a number of studies using manual labeling or custom software for commissure detection and leaflet labeling [10, 13, 14, 4].

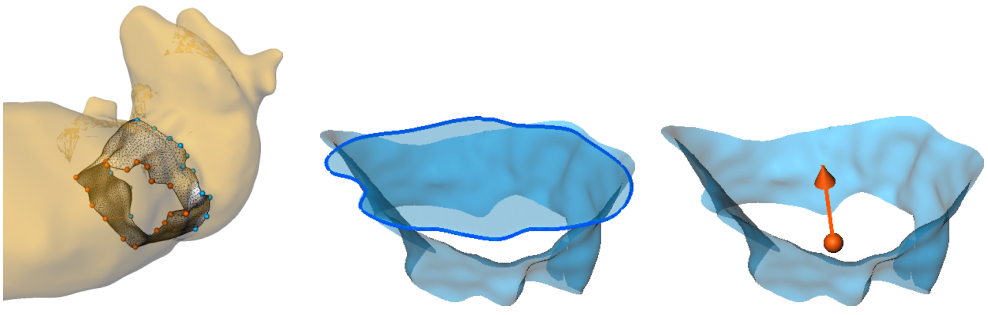
Several publications are concerned with the detection of commissures and segmentation of two separate mitral valve leaflets. All automatic approaches require extensive training on specific image data, as well as substantial context information in the data to locate the commissures. There is few evaluation data on the detection accuracy of the commissures [3]. We aim to detect the commissures in a valve surface model, supported by minimal image-based context information, and to evaluate the accuracy against expert observers.

2 Material and Methods

2.1 Data and Preprocessing

We used ten contrast-enhanced 4D computed tomography angiography (CTA) acquisitions (Siemens SOMATOM Definition Flash, Siemens, Germany). The images were acquired pre-surgery of adult patients diagnosed with different degrees of mitral valve insufficiency and scheduled for surgical or transcatheter cardiac intervention. The 4D volumes were reconstructed using retrospective ECG gating and a B30f kernel, and show the heart and surrounding structures.

A prototypical application based on the MeVisLab platform [11] was used for segmentation. Two clinical experts extracted LV, LA and aorta masks and an open-state, mesh-based mitral valve surface model for each volume using the approach proposed by Tautz et al. [12] (Fig.2 (a)).



(a) (b) (c)

Figure 2: *LV and LA mask and valve surface model extracted using the approach proposed in [12] (a). Annulus plane (b). Valvular axis defined by the annulus and orifice centers (c).*

2.2 Labeling of Valve Leaflets

Our labeling approach is based on the automatic detection of commissures to subdivide a valve surface model into the anterior and posterior leaflets and leaflet components.

The annulus plane is defined by the shortest eigenvector of the principal component analysis (PCA) of all annulus points, and annulus and orifice centers as the centers of gravity of the PCA of the respective points. The connection of these centers is the valvular axis (Figs. 2 (b) and (c)).

The orientation of the mitral valve surface model with respect to the patient coordinate system is derived from the aorta mask. The boundary points of the surface model closest to the aorta mask form a first estimate of anterior annulus points (A_C). The medial point of A_C is labeled as *anterior anchor point* ($\mathbf{p}_{\text{anchor}}$) (Fig. 3 (b)). Annulus and leaflets are bisected by $\mathbf{p}_{\text{anchor}}$ and it's opposite point on the annulus into two paths A_1 and A_2 of equal length L (Fig. 3 (c)).

Starting from the anterior anchor point, the commissures are searched along both paths in a subrange assuming a typical ratio between anterior and posterior annulus circumference of 1:2 [6, 5] (Fig. 3 (d)). Each subrange is defined by a start and an end point as

$$\mathbf{p}_{i,\text{start}} = \mathbf{p} \in A_i \cap A_C \;\mid\; d(\mathbf{p}) = \min\left(d_{i,max}, 0.30 \cdot L\right), i \in \{1,2\} \tag{1}$$

$$\mathbf{p}_{i,\text{end}} = \mathbf{p} \in A_i \cap A_C \;\mid\; d(\mathbf{p}) = \max\left(d_{i,max}, 0.35 \cdot L\right), i \in \{1,2\} \tag{2}$$

with

$$d_{i,max} = \max_{\mathbf{p} \in A_i \cap A_C} d(\mathbf{p}), i \in \{1,2\} \tag{3}$$

and d defining the path length between the anchor point $\mathbf{p}_{\text{anchor}}$ and a point $\mathbf{p}$ along A_i as

$$d(\mathbf{p}) = \int_{\mathbf{p}_{\text{anchor}}}^{\mathbf{p}} ds \tag{4}$$

For each annulus point, the corresponding orifice point is found by locating the closest point on the orifice along the valvular axis. The shortest path between the annulus point and the corresponding orifice point along the surface defines the leaflet height for one section. The shortest section is considered the commissure (Fig. 3 (e)). The commissures are labeled as anterolateral and posteromedial based on the valve orientation derived from the anterior anchor point.

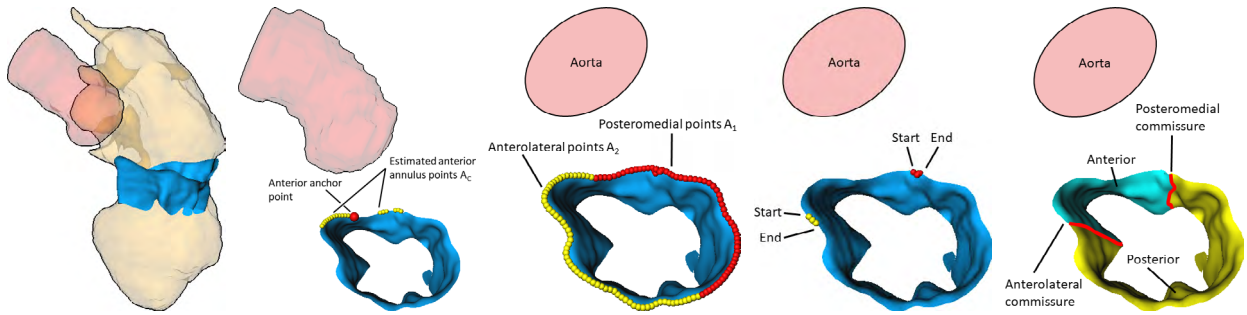


Figure 3: *Valve surface model (blue) from segmentation (view orientation towards posterior), with LA and LV (orange) and aorta (red) surfaces (a). Initially estimated anterior annulus and anchor points (b), paths of the bisected annulus (c), subranges for commissure search along the annulus as defined by Eqs. 1 and 2 (d), commissures and leaflets labeled as anterior and posterior (e).*

Both leaflets are then trisected into parts with equal annulus path length to allow analysis on leaflet component level, in particular for the clinically relevant A2 and P2 components [6] (cf. Fig. 1 (c)).

3 Results

For evaluation, two experienced cardiac surgeons marked commissure points on all 10 surface models. During annotation, they had access to the underlying image data.

First, we inspected the cases visually. Then differences between the commissures positions and derived leaflet parameters such as anterior circumference and leaflet height were assessed quantitatively.

3.1 Visual Comparison

All cases were analyzed qualitatively by the author and the respective observer during the annotation. According to expert feedback, the detection of the commissures worked well in most cases (Figs. 4 (a) and (b)), but in general underestimated the extent of the anterior leaflet (Figs. 4 (c) and (d)). In some cases, the estimated commissure points differed considerably from the expert definition (Figs. 5 (b) and (c)).

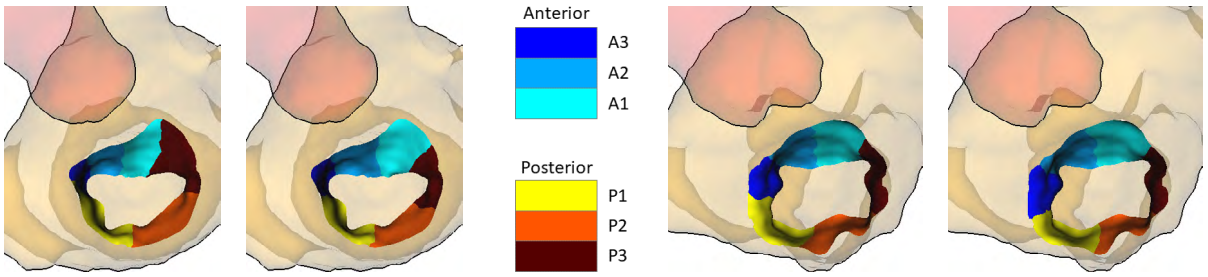


Figure 4: *Leaflet segmentations where the automatic approach (a, c) closely matches the commissures defined by the user (b, d). View orientation towards the apex. Note the general underestimation of the anterior leaflet (a, b), and the minor differences at both commissures (c, d).*

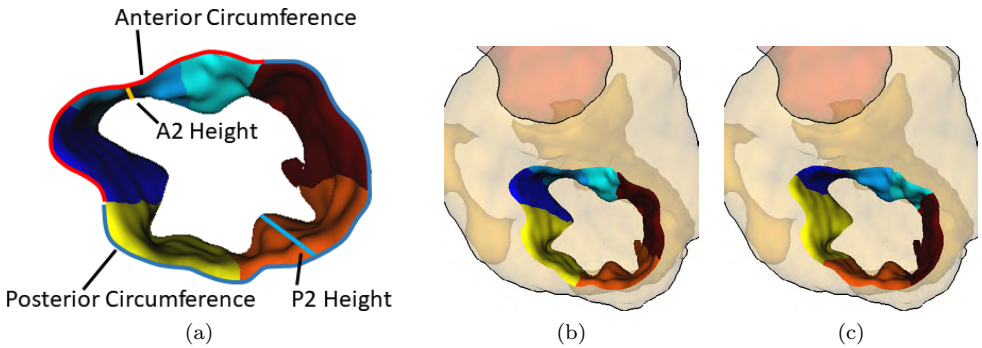


Figure 5: *Measurements on the labeled surface: anterior and posterior annulus circumference, and A2 and P2 heights (a). Leaflet segmentation where the automatic approach (b) differs considerably from the commissures defined by the user (c). View orientation towards the apex. Note the differences in commissure positions, in particular at the anterolateral commissure.*

3.2 Quantitative Assessment of Valve Leaflet Labeling

We calculated the differences between automatic and user-defined commissure positions to evaluate the agreement with the reference positions (Table 1). The variation of the commissure positions defined by the observers was comparable to the variation between user definition and automatic detection. We also derived clinically relevant measurements (Fig. 5 (a)): the anterior and posterior annulus circumference, and the maximum height of A2 and P2 leaflet components (Table 2). For calculation of the anterior and annulus circumferences, the annulus positions were projected to the annulus plane.

Differences in the commissure position are strongly visible in the annulus circumferences, while the A2 and P2 heights are more robust with respect to the commissure definition. The automatic detection in general underestimated the anterior leaflet when compared to the expert definitions. User 2 positioned the posteromedial commissure systematically closer to the aorta than user 1.

4 Discussion

The variation in the commissure point detection and the derived parameters between users and automatic leaflet labeling is comparable to the respective inter-user variation, with a mean error of 9.11 ± 6.15 mm between automatic and user-defined commissure and of 8.19 ± 8.56 mm between users. The inter-user agreement is

Signed Distance on Annulus [mm]		**Anterolateral Commissure**			**Posteromedial Commissure**			**Commissures Combined**		
Method 1	Method 2	min	median	max	min	median	max	min	median	max
Automatic	User 1	4.66	9.51	22.62	0.00	12.47	19.07	-4.66	11.12	22.62
Automatic	User 2	1.50	11.34	24.20	-10.81	0.72	5.95	-10.81	4.51	24.20
User 1	User 2	-12.53	0.00	14.94	-26.14	-11.00	0.00	-26.14	-0.43	14.94
Absolute Distance on Annulus [mm]										
Automatic	User 1	2.62	9.51	22.62	0.00	12.47	19.07	0.00	11.12	**22.62**
Automatic	User 2	1.50	11.34	24.20	0.00	**4.51**	**10.81**	0.00	7.33	24.20
User 1	User 2	**0.00**	**2.96**	**14.94**	0.00	11.00	26.14	0.00	**4.21**	26.14

Table 1: *Signed and absolute distances between commissures from automatic detection and user input along the annulus. Negative values indicate that method 2 detected the commissure closer to the anterior anchor point than method 1. Minimal absolute distances have been highlighted.*

Measurement [mm]		**Anterior Circumference**			**Posterior Circumference**			**A2 Height**			**P2 Height**		
		min	median	max	min	median	max	min	median	max	min	median	max
User 1		48.28	63.50	80.58	63.14	72.70	79.31	14.51	22.68	27.35	13.97	21.34	31.74
User 2		45.79	53.70	66.65	64.96	78.98	103.16	17.99	20.86	25.87	13.97	21.22	31.66
Automatic		38.10	43.33	51.67	71.96	92.19	108.37	11.70	21.03	27.35	13.97	23.17	31.74
Difference [mm]													
Method 1	Method 2												
Automatic	User 1	9.28	22.91	36.05	8.82	22.90	36.41	0.00	**0.19**	**5.16**	0.00	**0.05**	**6.01**
Automatic	User 2	1.08	9.82	**22.78**	1.19	10.10	**23.02**	0.00	1.48	10.45	0.00	0.41	6.07
User 1	User 2	**0.00**	**6.13**	27.42	**0.00**	**5.95**	27.36	0.00	1.39	7.64	0.00	0.34	6.83

Table 2: *Anterior and posterior annulus circumferences and A2 and P2 leaflet heights from automatic and user-defined commissures pooled over all cases. Minimal differences have been highlighted.*

higher than the agreement of either user with the automatic detection (Table 1). Grbic et al. reported a mean error of 2.75 ± 1.19 mm between expert ground truth and automatically fitted landmarks, and a mean inter-user error of 1.97 ± 1.4 mm. These errors have been averaged over seven mitral valve landmarks, including both commissures.

The users expected more pronounced commissures in the valve model. We assume that the underlying segmentation method is not able to detect this as accurate as desired, and that the results of both manual and automatic analysis would improve with more accurate imaging and segmentation approaches.

A limitation of the automatic approach is that it only considered the annulus shape and position relative to the aorta as anatomical information, whereas the clinical experts also interpreted the available image information during annotation. This suggests that using an extended anatomical model, including structures such as the coronary arteries or other valves as suggested in [2], could further improve our method.

5 Conclusion

We have presented an approach to automatically detect the mitral valve commissures in a surface model supported by an aorta segmentation. The valve model is separated into anterior and posterior leaflets, as well as into sub-leaflet components. The automatic quantitative assessment of the anterior and posterior geometry has been compared with annotations by clinical experts. To our knowledge, this is the first detailed comparison of automatically detected and manually defined commissure points. The results demonstrate principal clinical utility of the leaflet quantification based on the automatic commissure detection. The presented labeling scheme will be used in the automatic reporting of the location of valve pathologies such as calcification and leaflet prolapse.

Further work will focus on the improvement of the underlying anatomy detection in order to integrate more context information in the labeling process.

6 References

[1] Patric Biaggi, Christiane Gruner, Sean Jedrzkiewicz, Jacek Karski, Massimiliano Meineri, Annette Vegas, Tirone E. David, Anna Woo, and Harry Rakowski. Assessment of Mitral Valve Prolapse by 3D TEE. JACC Cardiovasc Imaging, 4(1):94–97, January 2011.

[2] Sasa Grbic, Thomas F. Easley, Tommaso Mansi, Charles H. Bloodworth, Eric L. Pierce, Ingmar Voigt, Dominik Neumann, Julian Krebs, David D. Yuh, Morten O. Jensen, Dorin Comaniciu, and Ajit P. Yoganathan. Personalized mitral valve closure computation and uncertainty analysis from 3D echocardiography. Med Image Anal, 35:238–249, January 2017.

[3] Sasa Grbic, Razvan Ionasec, Dime Vitanovski, Ingmar Voigt, Yang Wang, Bogdan Georgescu, Nassir Navab, and Dorin Comaniciu. Complete valvular heart apparatus model from 4D cardiac CT. Med Image Anal, 16(5):1003–1014, July 2012.

[4] Amir H. Khalighi, Andrew Drach, Robert C. Gorman, Joseph H. Gorman, and Michael S. Sacks. Multi-resolution geometric modeling of the mitral heart valve leaflets. Biomech Model Mechanobiol, October 2017.

[5] Karyn S. Kunzelman, Richard P. Cochran, Edward D. Verrier, and Robert C. Eberhart. Anatomic basis for mitral valve modelling. J Heart Valve Dis, 3(5):491–496, 1994.

[6] Patrizio Lancellotti, Luis Moura, Luc A. Pierard, Eustachio Agricola, Bogdan A. Popescu, Christophe Tribouilloy, Andreas Hagendorff, Jean-Luc Monin, Luigi Badano, and Jose L. Zamorano. European Association of Echocardiography recommendations for the assessment of valvular regurgitation. Part 2: mitral and tricuspid regurgitation (native valve disease). Eur J Echocardiogr, 11:307–332, 2010.

[7] Tommaso Mansi, Ingmar Voigt, Bogdan Georgescu, Xudong Zheng, Etienne Assoumou Mengue, Michael Hackl, Razvan I. Ionasec, Thilo Noack, Joerg Seeburger, and Dorin Comaniciu. An integrated framework for finite-element modeling of mitral valve biomechanics from medical images: Application to MitralClip intervention planning. Med Image Anal, 16(7):1330–1346, October 2012.

[8] Alison M. Pouch, Chun Xu, Paul A. Yushkevich, Arminder S. Jassar, Mathieu Vergnat, Joseph H. Gorman, Robert C. Gorman, Chandra M. Sehgal, and Benjamin M. Jackson. Semi-automated mitral valve morphometry and computational stress analysis using 3D ultrasound. J Biomech, 45(5):903–907, March 2012.

[9] A.M. Pouch, H. Wang, M. Takabe, B.M. Jackson, J.H. Gorman, R.C. Gorman, P.A. Yushkevich, and C.M. Sehgal. Fully automatic segmentation of the mitral leaflets in 3D transesophageal echocardiographic images using multi-atlas joint label fusion and deformable medial modeling. Med Image Anal, 18(1):118–129, January 2014.

[10] Yonghoon Rim, Ahnryul Choi, David D. McPherson, and Hyunggun Kim. Personalized computational modeling of mitral valve prolapse: Virtual leaflet resection. PloS one, 10(6):e0130906, 2015.

[11] F. Ritter, T. Boskamp, A. Homeyer, H. Laue, M. Schwier, F. Link, and H. Peitgen. Medical Image Analysis. IEEE Pulse, 2(6):60–70, November 2011.

[12] Lennart Tautz, Mathias Neugebauer, Markus Hüllebrand, Katharina Vellguth, Franziska Degener, Simon Sündermann, Titus Kuehne, Volkmar Falk, and Anja Hennemuth. Extraction of Open-State Mitral Valve Geometry from CT Volumes. In Proceedings of CARS 2018, 2018. (Accepted).

[13] Milan Toma, Morten Ø. Jensen, Daniel R. Einstein, Ajit P. Yoganathan, Richard P. Cochran, and Karyn S. Kunzelman. Fluid-Structure Interaction Analysis of Papillary Muscle Forces Using a Comprehensive Mitral Valve Model with 3D Chordal Structure. Ann Biomed Eng, 44(4):942–953, April 2016.

[14] Qian Wang and Wei Sun. Finite Element Modeling of Mitral Valve Dynamic Deformation Using Patient-Specific Multi-Slices Computed Tomography Scans. Ann Biomed Eng, 41(1):142–153, January 2013.

Intraoperative tissue characterization and classification in hyperspectral imaging: first results

H. Köhler[1], M. Maktabi[1], S.M. Rabe[2], M. Barberio[2,3], J.P. Takoh[2], N. Rayes[2], M. Diana[3], T. Neumuth[1], B. Jansen-Winkeln[2], I. Gockel[2], C. Chalopin[1]

[1]Universität Leipzig, Innovation Center Computer Assisted Surgery (ICCAS), Leipzig, Germany
[2]University Hospital Leipzig, Department of visceral, thorax, transplantation and vascular surgery, Leipzig, Germany
[3]Institute of Image-Guided Surgery (IHU), Strasbourg, France

Kontakt: Hannes.Koehler2@medizin.uni-leipzig.de

Abstract

Hyperspectral imaging (HSI) is a new imaging method in medicine. It combines the principles of spectroscopy with imaging. The approach is non-invasive and the acquisition is performed contactless. HSI systems record 3D hypercubes of data including the absorption spectra corresponding to the tissue surface. No image data is directly available. Tissue characterization can be performed based on the spectra, in order to perform an automatic classification and identification of target structures. Therefore, the goal of this work is to study the feasibility of tissue characterization and classification based on HSI data on three medical applications: the discrimination of the thyroid and parathyroid, the classification of intestine perfusion during ischemia and the identification of organs of the abdomen. First results showed that organs and different perfusion states have specific spectral features. These features could be used in computer-assisted algorithms for automatic tissue and perfusion classification.

Keywords: visceral surgery, operating room, tissue spectroscopy

1 Problem

Hyperspectral imaging (HSI) is a relative new optical approach which combines a visual camera with a spectrometer. It is a contactless and non-invasive method which requires no contrast agent. The measurement principle consists in illuminating the object under examination and in measuring the reflected light. The spectrum of the emitted light varies according to the application. The reflected light is split and analyzed according to the different wavelengths of the light spectrum. Therefore, a 3D cube of data, also called 3D hypercube, is recorded. The X- and Y-axis represent the spatial dimensions while the Z-axis is the spectral dimension. This approach is described as imaging technique because it measures a 2D spectral cartography of the object surface although it does not provide any images directly.

HSI is being evaluated in the medical research field since the beginning of the 2010's [1]. Commercial camera systems are transportable, easy to use and therefore suitable for medical applications. The main applications are the diagnosis and monitoring of scares [2], the assessment of tissue perfusion [3], also in the case of peripheral vascular diseases [4], the identification of tumors [5], especially brain tumors [6], or other structures like the nerves and ureter [7].

Some approaches developed for the discrimination of tissue consist in computing spectral parameters from the measured absorption spectra. Their values are depicted in false color images for visual interpretation. Examples of functional parameters are the oxygenation StO2, the NIR Perfusion- , the Tissue Hemoglobin- and the Tissue Water Index [3]. Also, the discrimination can be performed automatically using computer-assisted algorithms. For example classification methods based on machine learning approaches like support vector machines, artificial neural networks, random forest or K-nearest neighbors were implemented to automatically discriminate brain tumor from healthy tissue [8].

In this work, we present an approach for the characterization and automatic discrimination of tissue of the gastrointestinal tract. Preliminary results are presented on three different applications, which are 1. discrimination of the thyroid and parathyroid glands, 2. classification of perfusion states of the intestine and 3. identification of organs of the abdomen.

The parathyroid gland is a small endocrine gland in the neck area. Humans usually have four parathyroid glands which are located on the back of the thyroid gland. During thyroid and parathyroid surgery, the discrimination of both glands can be challenging. There is so far no standard imaging method to identify the parathyroid gland. In the context of research, low dose methylene blue together with NIR (near-infrared) cameras was tested to identify the parathyroids during neck surgery [9]. The limitation of this technique is the use of a contrast agent, administered prior to surgery and which is potentially neurotoxic at higher doses. Recently, the autofluorescence property of the parathyroid was evaluated. This approach consists in exciting the fluorophores included in the parathyroid with a light source from 750-785 nm. The gland emits as response a fluorescence peak around 820 nm which is larger than for the thyroid and background [10]. Autofluorescence imaging is a promising technique for the discrimination of the parathyroid. However, the measured signal is still too low to reach a visual discrimination of tissue intraoperatively [11].

Acute mesenteric ischemia is a frequent disease in old and ill patients. Consequences are the injury of the small intestine due to insufficient blood supply. Therapy is often performed in urgency and consists in the revascularization of tissue and the surgical resection of necrosis. The visual assessment of tissue perfusion is however subjective.

2 Material and Methods

The HSI system TIVITA™ Tissue (Diaspective Vision GmbH, Am Salzhaff, Germany) was used for the acquisition of the patient data sets. The system includes a CMOS camera providing images with resolution of 640 x 480 pixels, a spectrometer measuring light spectra in the range of 500 to 1000 nm with 100 channels (resolution of 5 nm) and an illumination component of 6 halogen lamps of 20 W each. The camera lens of the system works with a fixed focus at 30 cm distance. The camera system is attached to a side arm of a medical cart carrying the PC and a monitor to display the images intraoperatively.

The HSI measurements were performed at the University Hospital of Leipzig. The study protocol was approved from the ethics committee of the Leipzig University. During the acquisitions, the window blinds of the operating room were closed, the headlamps were switched off and the theatre lights turned away, in order to avoid disturbances from ambient light. Right after the acquisition of the 3D hypercubes (approximately 20 seconds), target structures were manually annotated as regions of interests (ROI) by a surgeon in the RGB images provided by the camera system and using the custom analysis software panel.

1. In this work HSI is used to record spectral data of seven patients during thyroid and parathyroid surgeries. The absorbance spectra of both glands were averaged on the set of patients and the standard deviations were computed. They were visually analyzed to extract specific spectral signatures for both glands with the objective to obtain an automatic classification. In a first approach these specific spectral signatures were used to build up a decision tree with the classes thyroid, parathyroid and background.

2. HSI data of one patient with acute mesenteric ischemia was acquired during surgery. The absorption spectra corresponding to healthy tissue and necrosis areas were plotted and analyzed. Thresholds of the absorbance values were defined from the visual inspection of the spectral signatures. These thresholds are integrated in a decision tree which classifies each pixel as necrotic tissue, critical tissue or background. The first two classes were overlaid as color map over the RGB images.

3. In HSI measurements of one patient depicting different organs of the abdomen, different target structures, liver, stomach, colon, ileum and peritoneum, were annotated. The corresponding absorbance spectra were visually analyzed to point out specific spectral signatures. Classification of organs was done using a principle component analysis (PCA) and k-means clustering.

3 Results

3.1 Discrimination of the thyroid and parathyroid

Figure 1 shows the averaged normalized absorbance spectra of the in vivo thyroid and parathyroid with standard deviations. The first observation is that the water and deoxygenated hemoglobin peaks are clearly visible at 960 and 760 nm. Secondly, the averaged spectra of the thyroid show a lower absorbance in the range 600 to 750 nm and a larger absorbance in the range 800 to 1000 nm in comparison with the parathyroid. This means that the thyroid is in average better oxygenated than the parathyroid and that the water content is larger for the thyroid. These differences can be used as spectral signature to characterize and discriminate both glands.

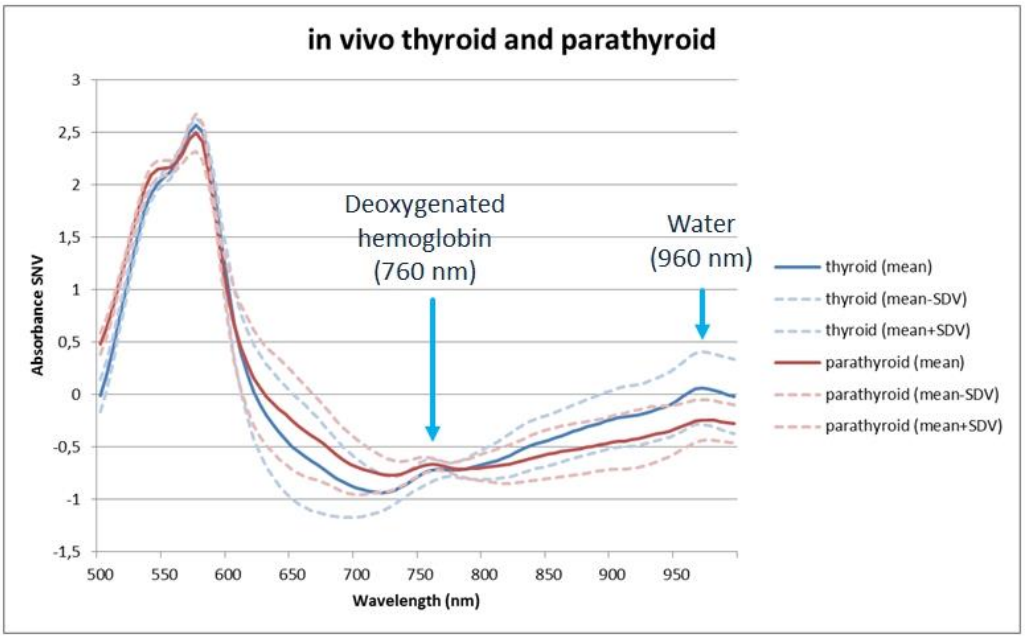


Figure 1: Averaged standard normal variate (SVN) transformed spectra with standard deviation (dotted lines) corresponding to the in vivo thyroid (blue) and parathyroid (red).

The feasibility of the development of classification algorithms based on these spectral signatures was studied. For that, the second derivatives of the absorbance spectra of one patient datasets depicting clearly both structures were plotted to enhance the minor spectral features independently from the baseline (Figure 2). Thresholds in four spectral regions, at 610 nm, 660 nm, 760 nm and 965 nm, build up a decision tree to classify each absorbance spectrum. These thresholds correspond also to the differences noticed in the mean absorbance spectra: 610 and 660 nm for the oxygenation state, 760 and 965 nm for the deoxygenated hemoglobin and water content. The result of the classification was visualized with two colors, green for the thyroid and red for the parathyroid, and overlapped on the RGB images (Figure 2)

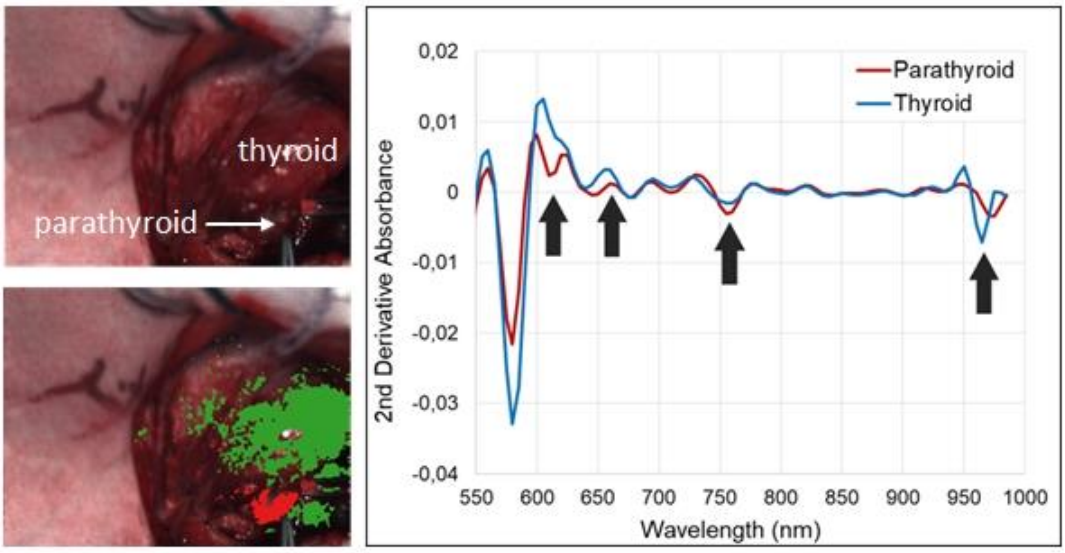


Figure 2: Right: second derivative of the mean absorbance spectra of the thyroid (blue) and parathyroid (red) of one patient. The arrows indicate the wavelengths of interest included in the classification algorithm. Left: RGB-image (above) and result of the classification (bellow).

3.2 Classification of intestine perfusion during ischemia

The mean absorbance spectra of the intestine depicted in Figure 3 show that the content of hemoglobin (500 to 600 nm) is lower in necrosis than in healthy tissue. A local peak at 630 nm, assumed typical for necrosis, is clearly visible on the blue and green curves. The larger absorbance of the necrosis in the 600 to 700 nm range means a lower oxygenation of tissue in comparison with healthy tissue.

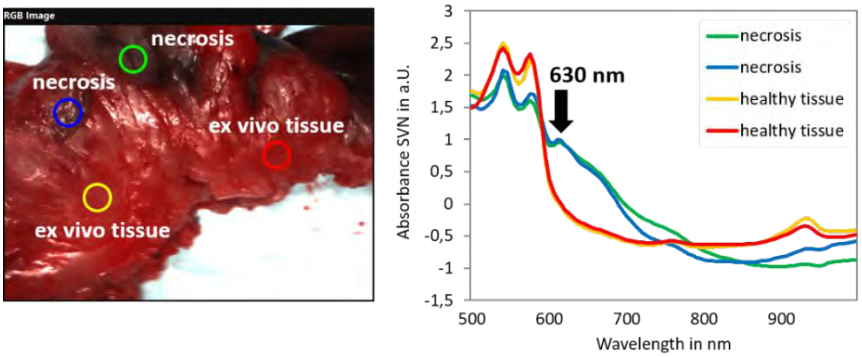


Figure 3: Mean absorbance spectra of four regions of interest selected in healthy and necrotic areas of ex vivo intestine part. The spectra of the necrosis show a typical local peak at 630 nm.

Features were defined to characterize necrosis and critical tissue. The detection of the peak at 630 nm was performed based on the first derivative of the absorbance spectra. A second parameter was computed as the average of the StO2 and NIR Perfusion indices. A detailed description and validation of these parameters can be found in Holmer et al. [12]. If the parameter value is smaller than 30% the perfusion in tissue is considered as critical. A decision tree based on the previous features was implemented to classify ex vivo intestine tissue in necrosis, critical and healthy tissue. Figure 4 shows the result of the classification on one data set.

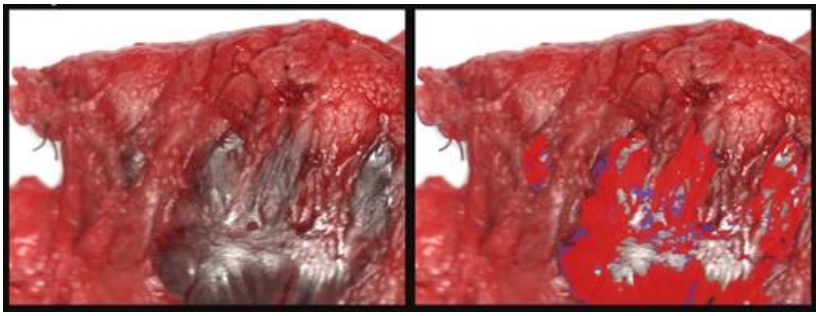


Figure 4: Automatic classification of the perfusion state of ex vivo intestine during mesenteric ischemia. Left: the necrosis is visible as dark tissue in the RGB image. Right: classified necrosis and critical tissue are depicted in red and purple. The remaining part of tissue was classified as normally perfused or ignored because of reflection artefacts.

3.3 Identification of organs of the abdomen

The absorbance spectra of the HSI data showed different behaviors according to the organs. Especially, the liver shows the characteristic absorption spectrum of bile (600 – 750 nm), which is produced in that organ [13]. The fat peak (930 nm) is well depicted in the absorbance curve of the ileum and the water peak (960 nm) is low, likewise to the peritoneum. The stomach and colon curves are very similar. Their water content is, like for the skin, larger compared to the other structures as shown in Figure 5. Using a principle component analysis (PCA) and k-means clustering it was possible to discriminate the liver, ileum, peritoneum and skin, which is illustrated in Figure 5. Stomach and colon belong to a separated class but could not be differentiated.

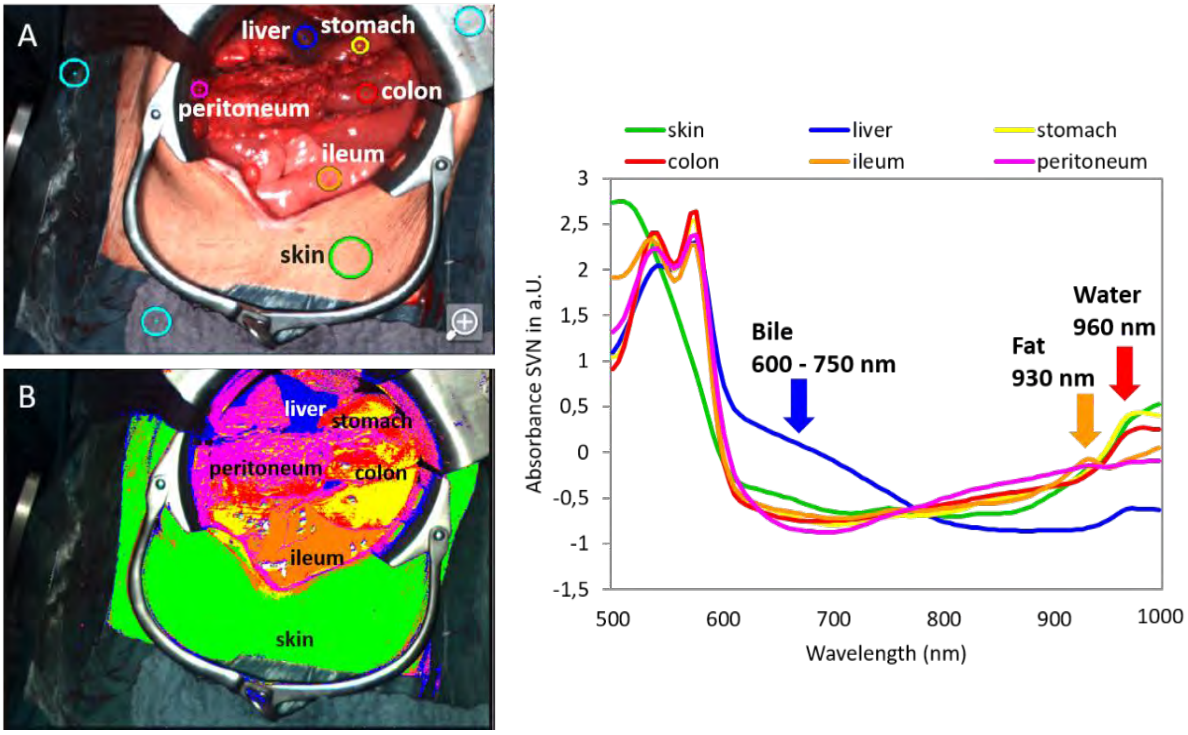


Figure 5: Mean absorbance spectra of regions of interest selected in different in vivo organs of the abdomen (A). Automatic identification of different organs in the abdomen area based on a PCA (B).

4 Discussion

In this study we demonstrate the feasibility of automatically discriminating organs and different tissue perfusion states based on HSI data with high spectral und spatial resolution. The described HSI-system is designed for use in wound diagnostics. For applications in open surgery, improvements are under development. This includes focused illumination to light up tissue deep in the body and correct for surrounding light during surgery. The automatic estimation of the object distance will in future decrease the time needed to get a focused image. A main limitation of this work is the restricted number of patient data, which makes it difficult to compare the results with other work. The analysis of further patient data has to be performed to confirm these preliminary results. Afterwards machine learning methods can be implemented and evaluated for intraoperative classification of tissue.

5 Conclusion

The intraoperative discrimination of tissue is of great interest for different surgical applications. This work demonstrates, that HSI with high spectral and spatial resolution is suitable for contact-free, non-invasive and rapid intraoperative assessment of different tissues. Organs and perfusion states of the intestine showed specific spectral features in HSI. These signatures were used in computer-assisted algorithms to automatically discriminate target structures and assess tissue perfusion. Such tools could assist the surgeon for intraoperative decision making process.

References

[1] Lu G, Fei B, *Medical hyperspectral imaging. A review*, Journal of biomedical optics, 19(1) 10901 (2014)

[2] Paul DW, Ghassemi P, Ramella-Roman JC, Prindeze N, *Noninvasive imaging technologies for cutaneous wound assessment. A review*, Wound repair and regeneration : official publication of the Wound Healing Society [and] the European Tissue Repair Society, 23(2) 149–162 (2015)

[3] Holmer A, Tetschke F, Marotz J, Malberg H, Markgraf W, *Oxygenation and perfusion monitoring with a hyperspectral camera system for chemical based tissue analysis of skin and organs*, Physiological measurement, 37(11) 2064–2078 (2016)

[4] Chiang N, Jain JK, Sleigh J, Vasudevan T, *Evaluation of hyperspectral imaging technology in patients with peripheral vascular disease*, Journal of Vascular Surgery, 66(4) 1192–1201 (2017)

[5] Goto A, Nishikawa J, Kiyotoki S, Nakamura M, Nishimura J, *Use of hyperspectral imaging technology to develop a diagnostic support system for gastric cancer*, Journal of biomedical optics, 20(1) 16017 (2015)

[6] Ortega S, Fabelo H, Camacho R, La Luz Plaza M, Callicó G, *Detecting brain tumor in pathological slides using hyperspectral imaging*, Biomedical optics express, 9(2) 818 (2018)

[7] Nouri D, Lucas Y, Treuillet S, *Hyperspectral interventional imaging for enhanced tissue visualization and discrimination combining band selection methods*, International journal of computer assisted radiology and surgery, 11(12) 2185–2197 (2016)

[8] Fabelo H, Ortega S, Ravi D, Kiran BR, Sosa C, Bulters D, Callicó G, *Spatio-spectral classification of hyperspectral images for brain cancer detection during surgical operations*, PloS one, 13(3) e0193721 (2018)

[9] Tummers QRJG, Schepers A, Hamming JF, Kievit J, Frangioni J, *Intraoperative guidance in parathyroid surgery using near-infrared fluorescence imaging and low-dose Methylene Blue*, Surgery, 158(5) 1323–1330 (2015)

[10] McWade MA, Sanders ME, Broome JT, Solórzano CC, Mahadevan-Jansen A, *Establishing the clinical utility of autofluorescence spectroscopy for parathyroid detection*, Surgery, 159(1) 193–202 (2016)

[11] Falco J, Dip F, Quadri P, La Fuente M, Prunello M, Rosenthal RJ, *Increased identification of parathyroid glands using near infrared light during thyroid and parathyroid surgery*, Surgical endoscopy, 31(9) 3737–3742 (2017)

[12] Holmer A, Tetschke F, Marotz J, Malberg H, Markgraf W, Thiele C, et al. *Oxygenation and perfusion monitoring with a hyperspectral camera system for chemical based tissue analysis of skin and organs.* Physiol Meas. 37(11) 2064–78 (2016)

[13] Nachabé R, Evers DJ, Hendriks BHW, Lucassen GW, van der Voort M, Wesseling J, et al. *Effect of bile absorption coefficients on the estimation of liver tissue optical properties and related implications in discriminating healthy and tumorous samples.* Biomedical Optics Express. 2(3) 600 (2011)

Deep Learning Based 3D Pose Estimation of Surgical Tools Using a RGB-D Camera at the Example of a Catheter for Ventricular Puncture

Fabian Reister[1], *Christian Kunz*[1], *Max Schneider*[2], *Torsten Kröger*[1]

[1] *Karlsruhe Institute of Technology, Institute for Anthropomatics and Robotics — Intelligent Process Automation and Robotics, Karlsruhe, Germany*
[2] *Department of Neurosurgery, University of Ulm, Günzburg, Germany*

Contact: christian.kunz@kit.edu

Abstract

Assistance systems improve the quality of treatment by showing the surgeon appropriate aids in the individual phases of a surgery. Often, the surgeon has only partial knowledge about the situation due to occlusion or the fact that the patient cannot be looked into. To increase the chances of success in surgery with a high demand on precision, a concept and implementation for providing missing data directly using augmented reality is introduced. This work addresses pose estimation of surgical tools, which is a required component for such an assistance system. First, we segment the tool in the color image based on a deep neural network and fuse this gathered information with the depth image. We then perform a model fitting to estimate the tool model parameters and its pose in 3D. The implementation and evaluation is based upon a catheter for ventricular puncture.

Keywords: Deep Learning, Semantic Segmentation, 3D Pose Estimation, Ventricular Puncture, PCA

1 Introduction

Computer vision for imaging methods like magnetic resonance imaging (MRI) and computer tomography (CT) already found its way into routine clinical practice. Simple tasks like segmentation of organs, but also more complex challenges like detecting malignant tissue can be achieved by using machine learning techniques. [4]. Augmented reality (AR) enables to visualize the gathered data during surgery. For example, it can be used in computer-aided neurosurgery [9] to minimize excision of benign tissue by displaying tumorous areas.

Procedures like a ventricular puncture, in which a catheter penetrates several centimeters through brain tissue before it reaches the cerebrospinal fluid (CSF), require a precise manipulation of surgical tools [15]. AR can be utilized to display information such as the position of the venctricular system, optimal penetration orientation of the catheter as well as the position of the catheter's tip. In addition, the status of the procedure can be monitored by tracking the catheter. When the catheter enters the head, it can be followed up in the AR scene.

2 Material and Methods

2.1 Experimental Setup

We used a RGB-D camera, an Intel Realsense SR300, that provides texture and direct depth data. In our experiments, the camera was mounted statically on a tripod. A phantom head from Synbone (www.synbone.com) was placed on a patient table. The head was partially covered by surgical covers. Different light situations were produced using a surgical lamp. In our setup we used a machine learning workstation with an Intel Core i7-8700K (3,7GHz), 64GB RAM and a Nvidia GeForce 1080 Ti GPU.

2.2 Segmentation of Color Images

To segment the surgical tool in the image, we employ a Full-Resolution Residual Network based on the FRRN-A architecture proposed in [8]. By using a pooling and a residual stream, the network is able to achieve good classification results as well as pixel accuracy.

The input image is first processed by residual units and then split into two streams. The pooling stream undergoes multiple convolution and pooling operations to gather information on object level. Alongside, the residual stream is enriched by this information but on full image resolution which requires an up-sampling.

The described computation of pooling and residual stream is summed up as Full-Resolution-Residual-Unit (FRRU). Multiple FRRU's in sequence will yield the Full-Resolution-Residual-Network (FRRN). In contrast to the original FRRN-A architecture we changed the last layer to consist of 2 output units to be able to perform a two class classification. Although, it would be possible, to use only a single output unit in this case, we highlight the ease of extensibility to segment multiple classes of tools.

As a cost function, we use the categorical cross entropy with a slight adaption by only considering the worst k pixels as favored in [8] and perform backpropagation using the Adam optimizer [5] and augment the dataset by horizontally mirroring the image and varying gamma. We are therefore able to perform an accurate segmentation for fine image structures.

2.3 Generation of Training Data

As there was no dataset available for this specific task, it had to be generated in a preferably (semi-)automatic fashion. In [14], the authors showed, that a neural network that was trained on segmenting single objects is able to segment complex scenes with multiple objects. Furthermore, they generated pixel-accurate foreground masks for known scenes and camera viewpoints.

Based on this idea, we put the camera statically into the scene and labeled it initially as completely being background. Afterwards, we added a single object to the scene which then was segmented by means of color and depth image segmentation.

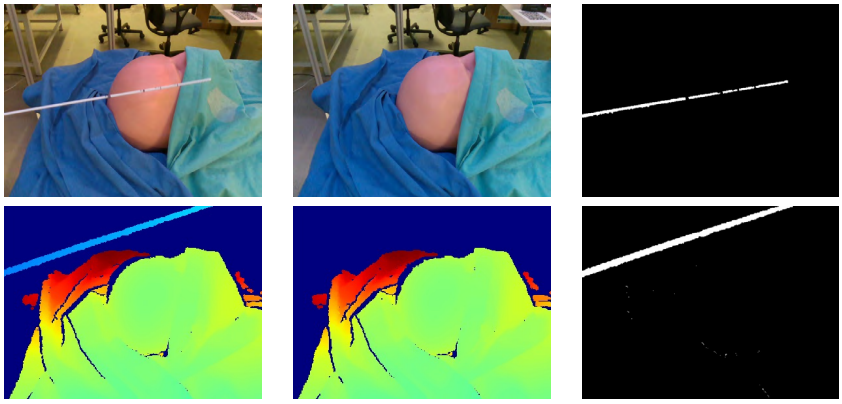


Figure 1: Segmentation based on color image (top) and depth image (bottom). From left to right: Image with catheter, background image and segmentation. The camera position is similar for all images.

Fig. 1 shows an exemplary segmentation based on color and depth image. The main challenges regarding the color segmentation are the robust handling of tool shadows and general illumination changes. As a consequence, we choose a Gaussian mixture based approach [16] that is able to tackle the issues by using a dynamic background model. When performing depth image based segmentation, a pixel precise registration of color and depth image is required. We create a background depth map and perform a distance based segmentation. To train our network, we created 3 datasets with the extend shown in table 1.

Dataset	1	2	3
Scene	Desk	Surgery	Surgery with additional items
Training images	5000	5000	4000
Validation images	2000	1500	1500

Table 1: Datasets.

We focused on generating different scenes by means of different objects, lighting conditions and viewpoint. An exemplary image for each dataset is given in figure 2.

2.4 3D Pose Estimation

We combine the segmented part of the image and the depth image to generate a point cloud of the tool. Due to inaccurate segmentation, a robust clustering is required. We apply a density based clustering algorithm [3] which suppresses falsely detected points and makes it possible to identify multiple objects of the same type. These clusters are individually registered to a reference model based on their type. Here, the catheter is modeled as a

Figure 2: Exemplary images of the generated datasets.

cylinder with known diameter. We use a principal component based approach which has shown to be robust. The eigenvector corresponding to the greatest eigenvalue indicates the direction of the cylinder. The length might vary due to occlusion.

2.5 Detection and Tracking of the Catheter Markings

When the tool is inserted into the brain, the tip is no longer visible. To estimate its position, we track the markings on the catheter indicating the depth of penetration. We use a convolutional neural network (CNN) to extract and match features from hidden layers of the network. As a similarity measure, we choose the cosine similarity that has been widely used in computer vision as a similarity measure for HOG descriptors [12] and features extracted from a CNN to perform place recognition [7]. Exploiting the activation of convolutional feature maps was also used for tracking [6], [11], [13].

Therefore, we employ the popular VGG16 [10], a CNN that has been trained on ImageNet [1]. Although it has been designed to perform an ambitious classification task, we focus on the activation of hidden cells but not on the final classification layer. We first take a reference image and apply the semantic segmentation described earlier in this paper to localize the tool in the image. We then perform the forward pass for the reference image using VGG16 and select and rank the features belonging to the tool as depicted in 3. Then in the same fashion, we proceed with the second image and so on to search for the best match in the extracted feature vectors.

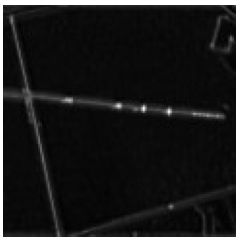


Figure 3: Tracked markings of the catheter.

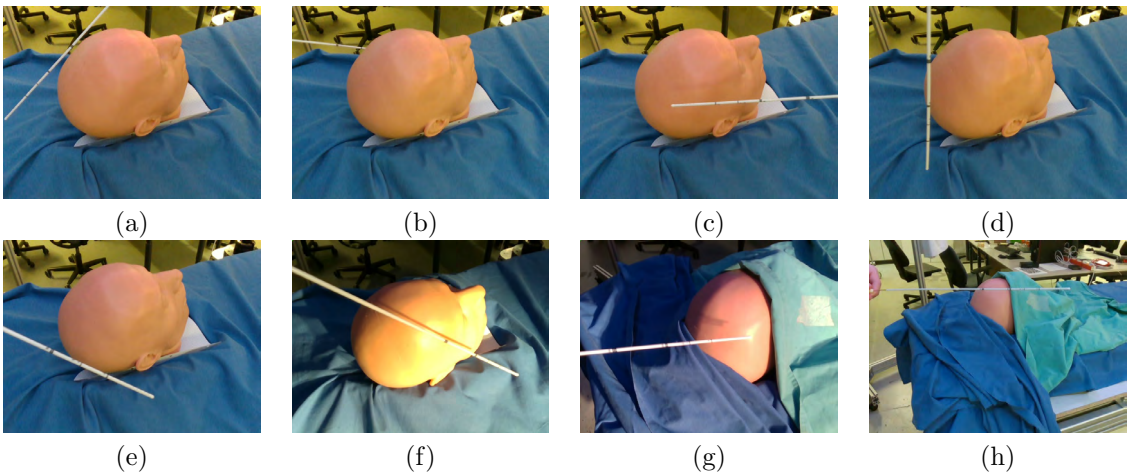
3 Results

3.1 Segmentation Evaluation

To evaluate the performance of the neural network and its ability to generalize, we have created several validation data sets. This includes variations of the tool itself in the form of different orientations in the image, because the CNN is not invariant against rotations and variations of the scene. The camera can also be mounted at different positions in the scene such that the distance to the patient and the tool might vary as well as the lighting conditions. Each dataset consists of approximately 1500 images in which the catheter is partially or fully visible. We use precision and recall as well as their harmonic mean, the F1-Score as metrics [2]. Accuracy in this case is not meaningful due to the imbalance between background and foreground pixels.

The results shown in table 2 highlight that the prediction is not invariant against different rotations of the instrument. If the catheter is placed vertically in the scene, precision and recall are dropping significantly. If we change the lighting conditions recall is dropping but precision is still high. This indicates that for the network it is becoming harder to detect the tool if the contrast to the phantom head is decreased. If we cover the head partially the contrast is enhanced and prediction improved. By increasing the distance to the patient, the amount of pixels belonging to the tool decreases which makes precision become more affected by outliers.

In figure 5 there are some representative training results shown that indicate a good handling of noise during training (first and second column). In some situations the prediction yields better results than the training data



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 4: Validation datasets. Images corresponding to the datasets listed in Table 2.

Dataset	Surgery 1 (a)	Surgery 2 (b)	Catheter horizontal (c)	Catheter vertical (d)	Catheter diagonal (e)	Bright (f)	Bright, patient covered (g)	Increased distance (h)
Precision	84.1%	64.0%	77.1%	47.7%	81.0%	59.0%	72.9%	11.5%
Recall	58.7%	61.9%	51.2%	19.2%	31.0%	15.4%	43.1%	55.8%
F1-Score	69.2%	62.9%	61.6%	27.4%	44.9%	24.4%	54.1%	19.0%

Table 2: Evaluation results for semantic segmentation.

(third column) because the network has learned to generalize. On the other hand, insufficient good training samples will lead to poor prediction (last column).

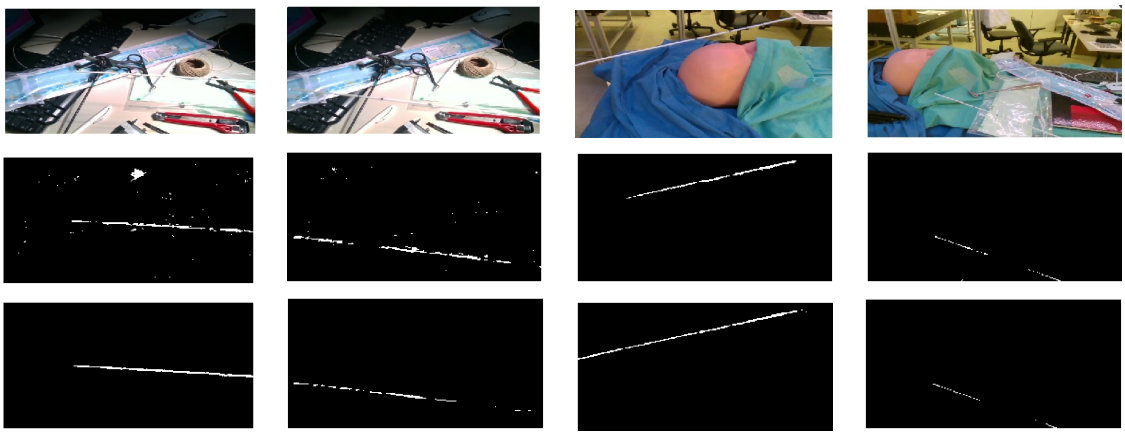


Figure 5: Selected training samples: color image (top), training data (center) and prediction (bottom).

3.2 Evaluation of Model Fitting Performance

To evaluate the model fitting, we recorded about 500 frames in which the catheter was fully visible. The model length estimation should yield 34.5 cm which is the known length of the catheter.

First, we placed the catheter in front of the camera (frames 1-100). As shown in fig. 6, the length prediction is fairly stable but biased. We then moved the catheter's tip towards and away from the camera (frames 100-250) such that the length estimation became a bit worse. Afterwards, we moved closer to the image border (frames 250-350). Under this condition, the prediction is limited due to the depth prediction capabilities of the camera. This can be identified in the right part of fig. 6 where there are no points visible at the shaft close to the hand. We then brought the catheter back into the image center (frames 350-450) and the estimated length became more stable again. In the last frames, the catheter was rotated away from the camera and was therefore hard to be recognized (frames 450-500).

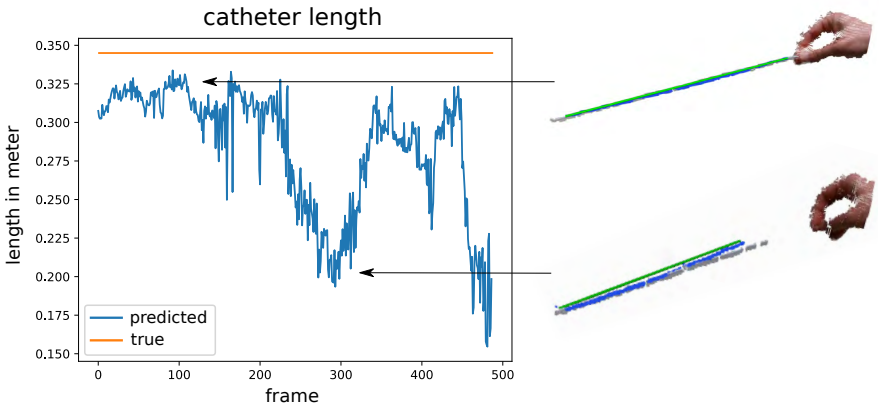


Figure 6: Left: predicted catheter length, right: RGB point cloud with overlaid semantic point cloud (blue) and predicted catheter model (green).

4 Discussion

As shown in the evaluation, the semantic segmentation is capable of segmenting fine image structures like the catheter for ventricular puncture. The performance issues can be explained by systematic errors in the creation of the training datasets and the coverage of different situations as well as the accuracy of the labels. To compensate these shortcomings, it seems reasonable to extend the training dataset. One way to automate this could be the use of two industrial robot arms: One to manipulate the position of the camera and a second to do the same with the tool. With the known pose of the tool and camera the segmentation of the training data can be enhanced and simplified.

Furthermore, it has been shown that the quality of depth estimation has a huge impact on the overall system performance. The employed consumer camera seems to have a limited utilizability for the described task. In early experiments, we also noticed that the depth prediction fails for dark or reflective surfaces like scalpels or scissors due to the projection technique for depth computation. To cope with reflective tools, time-of-flight sensors in combination with a color camera might be a prospective alternative to the Intel RealSense. The FRRN currently runs with 11.7 Hz for an image size of 512x256 pixels. For a real-time application at least 25 Hz are required. Cropping the image based on the previous position of the tool would immensely increase the segmented frames per second (FPS) as we already reach 38 FPS at half scale. Remarkably, the input size of the FRRN is not fixed as it lacks fully connected layers. Therefore, for a combination of the detection of the catheter's markings and the pose estimation, a fusion algorithm like the Kalman filter seems promising. The aim is to expand the current system to segment the whole scene and to determine the involved entities including the surgeon, the patient, the operating table and different tools. This would provide additional information that can be used, e.g. for surgical workflow analysis. This can be achieved without any algorithmic change to the applied training algorithm due to the fact that the approach can easily be adopted to multiple classes.

5 Conclusion

We have shown that segmentation of a surgical tool in the color image and subsequent combination with the depth image allows to perform an accurate pose estimation of parameterized models. The segmentation has already shown reasonably good performance for different situations that we investigated in section 3, although the amount of training data is quite limited. We carved out influencing factors diminishing the overall system performance that will be addressed in future work.

We showed that a highly accurate semantic segmentation is feasible in complex situations. Further studies should address clinical applicability, e.g. in an operating room. In the beginning, the ventricular puncture scenario would be well suited, as the surgical scene is clearly visible and the medical procedure is defined precisely. In ventricular puncture, the surgery can be followed by an estimation of the tool pose. This is an advantage for the patient, as the angle and penetration depth can be determined. This minimizes errors. It is of interest to combine this approach with an augmented reality based surgical navigation system. This offers the possibility to increase the surgeon's perception in order to get better prepared and accessible information about the surgery and the status of the patient. This can be of particular advantage in dynamic and highly complex scenarios such as the operating room.

6 Acknowledgement

This work was funded by the Federal Ministry of Education and Research within the project 'HoloMed - (Context-sensitive support of a surgeon in the operating room by Augmented Reality)'. It started on June 1st, 2017 and runs for two and a half years.

7 References

[1] J. Deng, W. Dong, R. Socher, L. J. Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pages 248–255, June 2009.

[2] Siddharth Dinesh and Tirtharaj Dash. Reliable Evaluation of Neural Network for Multiclass Classification of Real-world Data. arXiv preprint arXiv:1612.00671, 2016.

[3] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In Kdd, volume 96, pages 226–231, 1996.

[4] Afsaneh Jalalian, Syamsiah B.T. Mashohor, Hajjah Rozi Mahmud, M. Iqbal B. Saripan, Abdul Rahman B. Ramli, and Babak Karasfi. Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. Clinical imaging, 37(3):420–426, 2013.

[5] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980, 2014.

[6] Chao Ma, Jia-Bin Huang, Xiaokang Yang, and Ming-Hsuan Yang. Hierarchical Convolutional Features for Visual Tracking. In Proceedings of the IEEE International Conference on Computer Vision, pages 3074–3082, 2015.

[7] Tayyab Naseer and Wolfram Burgard. Deep Regression for Monocular Camera-based 6-DoF Global Localization in Outdoor Environments. In Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on, pages 1525–1530. IEEE, 2017.

[8] T. Pohlen, A. Hermans, M. Mathias, and B. Leibe. Full-Resolution Residual Networks for Semantic Segmentation in Street Scenes. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3309–3318, July 2017.

[9] Jeffrey H Shuhaiber. Augmented reality in surgery. Archives of surgery, 139(2):170–174, 2004.

[10] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[11] Jack Valmadre, Luca Bertinetto, João Henriques, Andrea Vedaldi, and Philip HS Torr. End-To-End Representation Learning for Correlation Filter Based Tracking. In Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on, pages 5000–5008. IEEE, 2017.

[12] Olga Vysotska, Tayyab Naseer, Luciano Spinello, Wolfram Burgard, and Cyrill Stachniss. Efficient and Effective Matching of Image Sequences Under Substantial Appearance Changes Exploiting GPS Priors. In Robotics and Automation (ICRA), 2015 IEEE International Conference on, pages 2774–2779. IEEE, 2015.

[13] Lijun Wang, Wanli Ouyang, Xiaogang Wang, and Huchuan Lu. Visual Tracking With Fully Convolutional Networks. In Proceedings of the IEEE International Conference on Computer Vision, pages 3119–3127, 2015.

[14] Andy Zeng, Kuan-Ting Yu, Shuran Song, Daniel Suo, Ed Walker, Alberto Rodriguez, and Jianxiong Xiao. Multi-view self-supervised deep learning for 6D pose estimation in the Amazon Picking Challenge. In Robotics and Automation (ICRA), 2017 IEEE International Conference on, pages 1386–1383. IEEE, 2017.

[15] Uwe K Zettl, Reinhard Lehmitz, and Eilhard Mix. Klinische Liquordiagnostik. Walter de Gruyter, 2005.

[16] Zoran Zivkovic. Improved adaptive Gaussian mixture model for background subtraction. In Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, volume 2, pages 28–31. IEEE, 2004.

Generierung und Verwendung künstlicher Trainingsdaten für CNNs zur Erkennung von Operationsinstrumenten

N. Graebling[1], *T. Neumuth*[1], *S. Franke*[1]

[1] *ICCAS, Universität Leipzig, Leipzig, Deutschland*

Kontakt: nico.graebling@medizin.uni-leipzig.de

Abstract

Diese Arbeit evaluiert das Potential künstlich erzeugter Trainingsdaten für die automatische Erkennung von Instrumenten anhand eines Convolutional Neural Networks (CNN). Die automatisierte Erkennung bildet eine zentrale Informationsquelle für die Workflowerkennung und Kontrollsysteme. Diese Systeme können das OP-Personal entlasten und die Patientensicherheit erhöhen. Es wurde einerseits eine Infrastruktur für die Generierung von Trainingsbildern für diesen Anwendungsfall geschaffen und andererseits wurden die Parameter und Klassifikationsgenauigkeiten des Trainings evaluiert. Hierfür wurden Modelle der Instrumente für das Rendering einer künstlichen Szene erstellt. Anhand der generierten Trainingsbilder wurde ein Finetuning des AlexNets durchgeführt, das dann anhand von echten Bildern aus dem Demo-OP des ICCAS evaluiert wurde. Während die Erkennung auf Bildern einzelner Instrumente gute Klassifikationsgenauigkeiten für Instrumentenklassen erzielte führte die Anwendung auf Bildern mit mehreren Intrumenten lediglich zu unzureichenden Ergebnissen und muss daher in zukünftigen Arbeiten weiter untersucht werden.

Keywords: Instrumentenerkennung, Maschinelles Lernen, Künstliche Trainingsdaten, Operationsassistenz

1 Problemstellung

Der moderne Operationssaal ist geprägt von komplexen und teils hochgradig individualisierten Abläufen, einer hohen Personaldichte, ausgeprägten Interaktionen verschiedenster Akteure und technischer Systeme. In dieser Arbeitsumgebung können digitale Unterstützungen Prozesse vereinfachen, eine automatisierte Dokumentation ermöglichen und die Patientensicherheit erhöhen [1]. Solche Unterstützungs- und Kontrollsysteme basieren häufig auf der Erkennung des momentan verwendeten Instrumentes als Indikator der aktuellen Operationsphase [1, 2]. Bisherige Arbeiten zur Instrumentenerkennung erfordern jedoch häufig eine Modifikation der Instrumente oder benötigen eine weiteres in den Operationssaal zu integrierendes System. Ein genereller Nachteil der Nutzung optischer Systeme zum intraoperativen Instrumententracking, wie beispielsweise die Verwendung von Barcodes [3], farblichen Markierungen [4] oder dem Abgleich von Kamerabildern mit Ontologiewissen [5], ergibt sich aus der Einschränkung des Bewegungsablaufs beim Greifen nach und Ablegen der Instrumente aufgrund der Notwendigkeit, optische Verdeckungen zu vermeiden. Die Notwendigkeit der Modifikation der Instrumente bei momentan existierenden Ansätzen, führt zudem zu einem erhöhten finanziellen und zeitlichen Aufwand. Des Weiteren wird in einigen Fällen dadurch das Sterilisieren der Instrumente erschwert. Auch die Notwendigkeit weitere Systeme in die OP-Umgebung zu integrieren und Wechselwirkungen auszuschließen, wie beispielsweise bei der Verwendung von RFID-Sensoren [6], stellt einen zentralen Nachteil bisheriger Systeme dar.
Diese Nachteile adressiert der Ansatz lediglich eine Kamera und unmodifizierte Instrumente zu verwenden und die Instrumentenerkennung anhand eines CNNs durchzuführen. CNN sind eine auf die Bildklassifikation ausgerichtete Spezialform Künstlicher Neuronaler Netze und finden auch im medizinischen Kontext zunehmend Anwendung. Es handelt sich hierbei um einen Ansatz aus dem Bereich des maschinellen Lernens. Eine zentrale Herausforderung bei der Anwendung von CNNs stellt der große Datenumfang, der für das Training benötigt wird, dar. Der Trainingskorpus muss nicht nur möglichst viele Variationen des Anwendungsfalles abdecken, sondern auch fehlerfrei annotiert sein. Da die Annotation momentan meist noch manuell durchgeführt wird, enthält sie häufig Fehler und ist mit einem hohen zeitlichen Aufwand verbunden. Bisherige Arbeiten haben gezeigt, dass künstlich erzeugte Trainingsdaten diese Probleme lösen könnten [7]. Sie ermöglichen eine absolut fehlerfreie Annotation und sind, abgesehen von einem initialen Aufwand zur Erzeugung der Bildgenerierungsinfrastruktur, ohne großen manuellen Aufwand leicht skalierbar.

2 Material und Methoden

Die Arbeit umfasst die Generierung der Instrumentenmodelle, die Erzeugung der künstlichen Trainingsdaten, die Auswahl einer Teilmenge dieser und das Training des CNNs. Abbildung 1 zeigt den generellen Ablauf

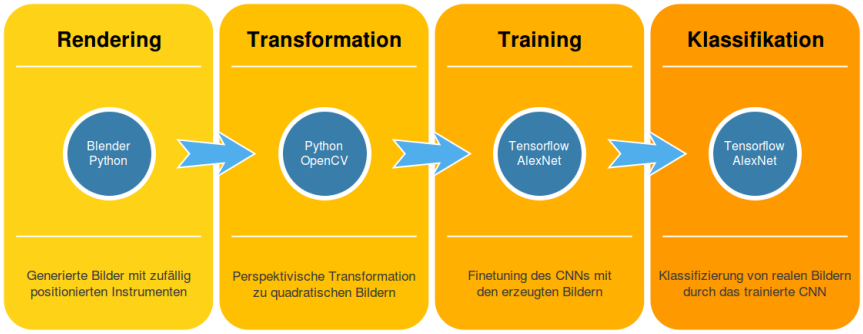


Abbildung 1: *Ablauf des Vorgehens vom Rendering bis zur Klassifikation*

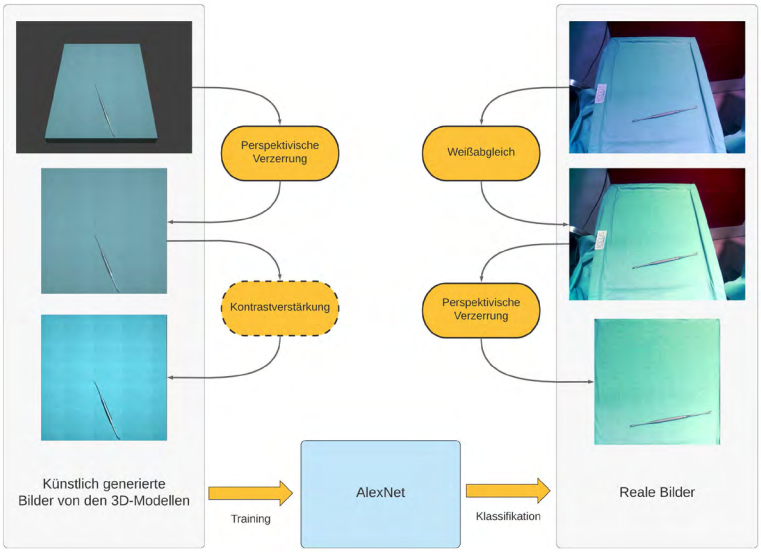


Abbildung 2: *Übersicht über das Vorgehen*

des Vorgehens vom Rendering bis zur Klassifikation und die verwendeten Technologien. Abbildung 2 zeigt schematisch die Verwendung der generierten Bilder für das Training und die Klassifikation.

2.1 Modellgenerierung

Aufgrund der unzureichenden Qualität der von einem 3D-Scanner erzeugten Modelle, die auch durch verschiedene Oberflächenmodifikationen der Instrumente nicht signifikant verbessert werden konnte, wurden die Modelle der Instrumente manuell mit Hilfe der Software *Blender* erstellt. Die Modellierung erfolgte anhand von Referenzbildern durch die Komposition geometrischer Grundformen und komplexerer Freihandmodellierungen. Abbildung 3 zeigt die Modelle der 11 ausgewählten Instrumente der Nasennebenhöhlenchirugie. Um die Komlexität der Modelle zu beschränken wurden die Instrumente bezüglich des Öffnungswinkels nur in einem Zustand modelliert. Diese Einschränkung wurde bei der Aufnahme der realen Bilder entsprechend übernommen.
Um die Validität der Modelle für den Anwendungsfall zu prüfen wurden verschiedene Computervision-Verfahren zum Bildvergleich in Erwägung gezogen. Da bei diesen jedoch die Ähnlichkeit eines gerenderten Bildes und eines realen Fotos von der Ablage des Instruments und dem Blickwinkel sowie der Lichtsituation abhängen, genügen sie für die Beurteilung der Ähnlichkeit der echten Instrumente und der Instrumentenmodelle im Rahmen der Validierung nicht. Die Abschätzung der Validität bleibt daher der menschlichen Kognition überlassen.
Die minimalen Abweichungen in der Modellierung wirken sich nicht stark auf die globale Form des Instruments aus, sondern haben lediglich lokale Effekte zur Folge. Durch die geringe Auflösung der Trainingsbilder für das AlexNet von 224*224 Pixeln und die perspektivische Verzerrung beim Vorverarbeiten der Bilder ist zudem damit zu rechnen, dass die geringfügigen Artefakte und minimalen Abweichungen, die im ursprünglichen, gerenderten Bild auftreten, später in den Trainingsdaten nicht ins Gewicht fallen. Daher sind die Modelle trotz im Detail erkennbarer Unterschiede für diesen Anwendungsfall valide.

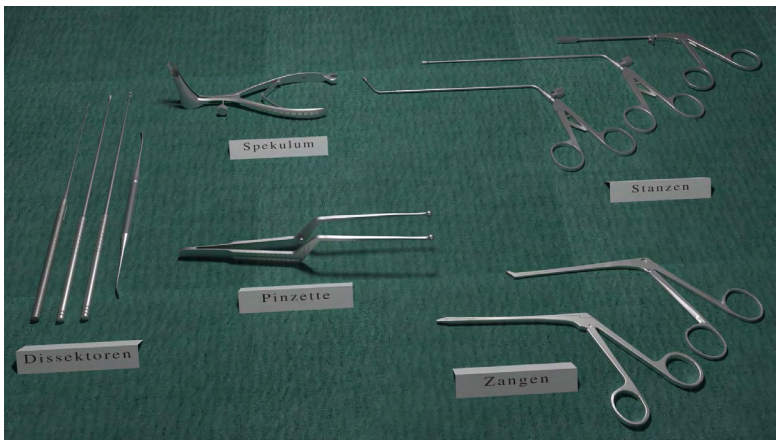


Abbildung 3: *Rendering der Modelle der ausgewählten Instrumente*

2.2 Training des CNNs

CNNs sind ein Ansatz für die Bildklassifikation aus dem Bereich des maschinellen Lernens, die, wie auch die allgemeinere Form der Künstliche Neuronale Netze, die Prozesse des menschlichen Gehirns durch eine Modellierung der Wechselwirkungen von Neuronen simulieren. Hierbei ermöglichen die Anordnungen der Schichten aus Neuronen sowie deren unterschiedlichen Varianten verschiedene Netzarchitekturen. Wie schon Sarkar, Varanasi und Stricker beschrieben, bietet das AlexNet die Kombination aus hinreichender Mächtigkeit für die Bildkassifizierung und einer einfachen Hierarchie [7]. Basierend auf diesen Vorteilen sowie der Möglichkeit des einfachen Finetunings und der guten Dokumentation des AlexNets wurde es anderen Netzarchitekturen vorgezogen.
Da die ersten Schichten eines CNN ausschließlich für eine allgemeine Kanten- und Feature-Erkennung zuständig sind und lediglich die hinteren Schichten die Zuordnung dieser zu den Klassen des Klassifikationsproblems übernehmen, ist es möglich, die auf einem anderen Datensatz vortrainierten Gewichte der vorderen Schichten zu übernehmen [7]. Hierfür wurden Gewichte verwendet, die aus dem Training mit dem ImageNet-Datensatz resultieren. Im Rahmen des Finetunings werden lediglich die letzte, die beiden letzten oder die drei letzten vollständig verbundenen Schichten neu traininiert. Die Beschränkung auf das Finetuning im Vergleich zum Training des gesamten Netzes, führt zu einer beachtlichen Zeitersparnis beim Trainingsprozess und ermöglicht es somit, mehr Trainingskonfigurationen zu erforschen.

3 Evaluation

Um die Komplexität der Klassifizierungsaufgabe zu reduzieren, wurden in den folgenden Untersuchungen jeweils auf ein Instrument beziehungsweise eine Instrumentenklasse spezialisierte Netze trainiert anstatt ein Netz zu trainieren, das die gesamte Klassifikation durchführt. Somit mussten die Netze lediglich erkennen, ob das entsprechende Instrument oder die Instrumentenklasse im zu klassifizierenden Bild enthalten ist. Der Anteil der korrekt klassifizierten Bilder bezüglich aller Testbilder wird im Folgenden als Klassifikationsgenauigkeit bezeichnet.

3.1 Auswahl der Trainingsdaten

Aus den erzeugten und den aufgenommenen Bildern wurden Teilmengen verschiedener Größe für das Training und das Testen der Genauigkeit der Netze erstellt. Die Auswahl der einzelnen Bilder für die Trainingskorpora erfolgte hierbei zufallsbasiert. Für alle Trainingsdatenmengen gilt, dass sie keine Überschneidungen mit den für die Auswertung verwendeten Bildmengen haben. Dies ist nötig, um zu garantieren, dass das Netz tatsächlich das Konzept des entsprechenden Instruments und nicht die Einzelausprägungen auf bestimmten Bildern gelernt hat. In allen Trainings- und Testkorpora sind die Instrumente beziehungsweise Instrumentenklassen in gleichen Anteilen enthalten.

3.2 Parameteroptimierung

Um die optimalen Parameter des Trainings zu bestimmen, wurden Tests ausschließlich für das Elevatorium und anhand von Bildern lediglich eines Instrumentes durchgeführt. Das Elevatorium weist keine zu markante Form

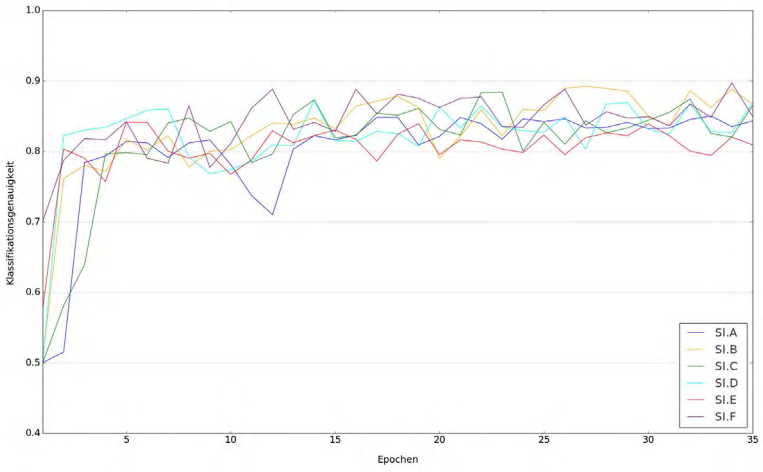


Abbildung 4: *Klassifikationsgenauigkeit im Verlauf des Trainings bei verschiedenem Datenumfang*

auf und ist zu anderen Instrumenten trotzdem auch nicht zu ähnlich geformt. Die Auswahl dieses Instruments ermöglicht daher eine repräsentative Parameterbestimmung für alle Instrumente. Die Untersuchungen haben gezeigt, dass die Verwendung des Weißabgleichs für die zu klassifizierenden, realen Bilder bessere Ergebnisse erzielt als die Klassifikation ohne vorherigen Weißabgleich. Ein signifikanter, positiver Einfluss durch eine Kontrastverstärkung auf den künstlich erzeugten Trainingsdaten ließ sich nicht feststellen. Das Finetuning sollte auf die letzten zwei Schichten angewandt werden, um optimale Ergebnisse zu erzielen. Für die Lernrate hat sich ein Wert von 0.004 als optimal herausgestellt. Das Training sollte mindestens 35 Epochen umfassen, um sicherzugehen, dass Schwankungen nachgelassen haben und ein stabiler Wert der Klassifizierungsgenauigkeit erreicht wurde. Nach der 35. Epoche sind keine signifikanten Verbesserungen des Trainings mehr zu erwarten.

3.3 Variationen des Datenumfangs

Wie bereits beschrieben, stellt die Generierung großer Trainingskorpora ohne hohen manuellen Aufwand einen der zentralen Vorteile der Nutzung künstlich erzeugter Trainingsdaten dar. Daher wurde in den Untersuchungen SI.A bis SI.F der Einfluss des Trainingsdatenumfangs auf die Klassifizierungsgenauigkeit untersucht. Tabelle 1 zeigt den Umfang der verschiedenen Datensätze, die sich ansonsten in ihrer Form nicht unterscheiden. Für alle Netze wurden die im vorigen Abschnitt ermittelten, optimierten Parametern verwendet. Der Vergleich wurde, wie auch die Untersuchungen im letzten Abschnitt, anhand des Elevatoriums durchgeführt. Es lässt sich bei dem Training mit künstlich erzeugten Trainingsdaten in diesem Anwendungsfall kein signifikanter, positiver Einfluss der Steigerung des Datenumfangs auf die Klassifikationsgenauigkeit feststellen da die Klassifikationsgenauigkeiten für alle Datensätze während des Großteils des Trainings in einem Bereich zwischen 80% und 90% deutlich schwanken (Abbildung 4). Eine möglicher Erklärung hierfür ist, dass die Ähnlichkeit der Art der Bilder des Trainingskorpus zu den zu klassifizierenden Bildern weitaus entscheidender ist als der Umfang der Trainingsdaten.

Tabelle 1: *Übersicht der verwendeten Variationen des Trainingsdatenumfangs*

Datensatz	Trainingsbilder	Validierungsbilder	Testbilder
SI.A	400	180	1000
SI.B	600	260	1000
SI.C	800	340	1000
SI.D	1000	420	1000
SI.E	1200	500	1000
SI.F	1400	600	1000

3.4 Training für Bilder mit einem Instrument

In dem folgenden Teil der Evaluation wurden wie zuvor Netze mit Bildern einzelner Instrumente sowohl für das Training als auch das Testen verwendet. Vergleichende Untersuchungen ergaben sehr große Unterschiede in der Qualität und der Stabilität der Klassifikationsgenauigkeit für die verschiedenen Instrumente. Während die Netze, die auf das Nasenspekulum oder die Tumorpinzette spezialisiert waren am Ende des Trainings in mehr als 95% bzw. circa 90% der Fälle korrekt erkannten, ob das jeweilige Instrument auf dem Bild zu sehen ist, blieben die Klassifikationsgenauigkeiten beispielsweise für die Zangen stark schwankend deutlich unter 80%. Die guten Ergebnisse für das Nasenspekulum und die Tumorpinzette lassen sich durch ihre sehr markante Form erklären und spiegeln sich auch in den Ergebnissen des zu Vergleichszwecken durchgeführten Trainings mit realen Daten wider.
Aufgrund dieser stark variierenden Ergebnisse wurde versucht die Komplexität der Klassifikationsaufgabe zu reduzieren, indem funktional ähnliche Instrumente gruppiert wurden. Somit ergaben sich vier Klassen: Die Klasse der Zangen & Stanzen als Gewebe abtragende Instrumente, die Klasse der Dissektoren sowie das Spekulum und die Pinzette. Die Reduzierung der Aussagekraft der Erkennung bei der Verwendung der Klassen statt der einzelnen Instrumente ist gering, da die Unterscheidung sehr ähnlicher Ausprägungen einzelner Instrumente einer Klasse nur wenig Nutzen bringt. Für das Training der auf die Instrumentenklassen spezialisierten Netze, wurden aus den generierten Datensätzen 700 Trainings- und 300 Validierungsbilder je Klasse ausgewählt. Für Klassen, die mehrere Instrumente enthalten, wurden Datensätze erzeugt, die die einzelnen Instrumente zu gleichen Anteilen beinhalten. Dies gilt auch für die je 1000 reale Bilder umfassenden Testkorpora. Durch die Verwendung von Instrumentenklassen konnten signifikante Verbesserungen der Ergebnisse erreicht werden. Für alle Klassen liegt die Klassifikationsgenauigkeit über 90% und verhält sich schon in frühen Epochen des Trainings stabil. Diese sehr guten Ergebnisse bleiben jedoch unter den Ergebnissen des Trainings mit realen Bildern im gleichen Anwendungsfall zurück, da diese Klassifikationsgenauigkeiten von über 95% erreichen. Die deutlich gestiegene Qualität der Klassifikationsergebnisse und die nur geringe Einschränkung der gewonnenen Informationen führen zu der Schlussfolgerung, dass die Verwendung der Instrumentenklassen sinnvoll ist.
Versuche die Vorteile künstlicher und realer Trainingsdaten durch eine Kombination der Datensätze auszunutzen führten nicht zu einer Steigerung der Klassifikationsgenauigkeit über die Ergebnisse des Trainings mit realen Daten hinaus. Ansätze die Varianz des künstlichen Datensatzes zu vergrößern, indem dieser durch zufällige Hintergrundbilder erweitert wurde, führten sogar zu signifikanten Verschlechterungen. Dies ist mit der gesteigerten Komplexität der Aufgabe, das Instrument im Vordergrund von dem zufälligen Hintergrund zu abstrahieren, zu erklären.

3.5 Training für Bilder mit mehreren Instrumenten

In dem folgenden Teil der Untersuchungen wurde das Training weiter an den Anwendungsfall im Operationssaal angepasst. Statt der Klassifikation von Bildern einzelner Instrumente wie bisher, wurde die Erkennung von Bildern mit bis zu vier Instrumenten untersucht. Aufgrund der Erkenntnisse der vorangegangenen Untersuchung wurden auch hier die Erkennung von Instrumentenklassen statt einzelner Instrumente trainiert und evaluiert. Das Ziel der Netze in der folgenden Untersuchung ist es demnach, zu erkennen, welche der vier Instrumentenklassen auf dem zu klassifizierenden Bild zu sehen sind. Untersuchungen zur Erkennung der einzelnen Instrumente wurden im folgenden Abschnitt nicht durchgeführt.
In einem ersten Ansatz wurden die auf Bildern einzelner Instrumente trainierten Netze verwendet. Die Ergebnisse schwankten je nach Klasse hierbei stark. Während die Klasse der Stanzen und Zangen mit Klassifikationsgenauigkeiten von circa 95% erkannt wurden, lagen die Ergebnisse für die Dissektoren lediglich bei circa 55%. Die schwachen Ergebnisse für die Erkennung der Dissektoren lassen sich durch die geringe Größe der Instrumente erklären. Einfache und gerade Strukturen ohne weitere Besonderheiten fallen in einem Bild mehrerer Instrumente mit Überlagerungen weniger stark auf. Daher ist die Klassifikation für diese Klasse schwieriger als beispielsweise für die Klasse der Stanzen und Zangen. Die auffällige Struktur der Griffe ermöglicht für diese sehr hohe Klassifikationsgenauigkeiten.
Auch für Bilder mehrerer Instrumente zeigte sich, wie zu erwarten war, eine Verschlechterung der Ergebnisse bei einer Expansion des Trainingskorpus durch die Komposition der Bilder einzelner Instrumente mit zufälligen Hintergründen. Auch erneute Untersuchungen zur Vergrößerung der Trainingsdatenmengen steigerten die Klassifikationsgenauigkeit nicht. Dies ist auf die bereits zuvor erwähnten Gründe zurückzuführen.

4 Diskussion

Es wurden gute Ergebnisse für die Erkennung von Instrumentenklassen auf Bildern einzelner Instrumente erzielt. Es zeigte sich jedoch wider Erwarten in den durchgeführten Untersuchungen nicht, dass durch die Verwendung von Trainingskorpora aus generierten Daten die Klassifikationsgenauigkeit der Netze, die mit kleineren Trainingsdatensätzen aus realen Bildern trainiert wurden, übertroffen werden können. Dies lässt vermuten, dass die Ähnlichkeit der Trainingsdaten zu den zu klassifizierenden Daten für die Funktionalität des Netzes entscheidender ist als der Umfang des Trainingskorpus. Diese These sollte in zukünftigen Untersuchungen mit noch größeren künstlichen Datensätzen überprüft werden.
Eine mögliche Erklärung dafür, dass die guten Ergebnisse der Einzelbilder nicht auf Bilder mehrerer Instrumente übertragen werden konnten, ist die hohe Komplexität der Klassifizierungsaufgabe, die sich durch eine große Ähnlichkeit der Instrumente und die Überlagerungen dieser begründet.
Zukünftige Ansätze um die Klassifikationsergebnisse für Bilder mehrerer Instrumente zu verbessern sind beispielsweise die Segmentierung des zu klassifizierenden Bildes, die Evaluation anderer Netzarchitekturen und die strukturisierte statt randomisierte Erzeugung der künstlichen Trainingsdaten. Auch eine Kombination des bestehenden Systems mit einer weiteren Sensorik wie beispielsweise einer Waage oder das Verwenden eines Meta-Klassifikators für die Gewichtung der Ergebnisse der schwächeren Klassifikatoren bieten vielversprechende Möglichkeiten weiterer Untersuchungen.

5 Zusammenfassung

Es wurde eine valide Infrastruktur für die Generierung künstlicher Trainingsdaten für den Anwendungsfall geschaffen und der Nutzen der somit erzeugbaren künstlichen Trainingsdaten evaluiert. Es ließen sich gute Klassifikationsgenauigkeiten von über 90% für alle Instrumentenklassen auf Bildern einzelner Instrumente erzielen. Diese Ergebnisse konnten nicht auf die Klassifikation von Bildern mehrerer Instrumente übertragen werden. Weitere Ansätze um das bestehende System zu verbessern bieten jedoch vielversprechende Möglichkeiten und sollten, auch wegen des hohen Nutzens für die automatisierte Workflowerkennung und somit eine mögliche Steigerung der Patientensicherheit im OP, in zukünftigen Arbeiten weiter untersucht werden.

6 Referenzen

[1] S Franke, M Rockstroh, E Schreiber, J Neumann, T Neumuth, “Towards the intelligent OR - Implementation of distributed, context-aware automation in an integrated surgical working environment”, *MICCAI, Athens*, 2016

[2] P Liebmann, J Meixensberger, P Wiedemann, T Neumuth, “The impact of missing sensor information on surgical workflow management”, *International Journal of Computer Assisted Radiology and Surgery* 85S. 867–875, 2013

[3] M Kranzfelder, A Schneider, G Blahusch, H Schaaf, H Feussner, “Feasibility of opto-electronic surgical instrument identification”, *Minimally Invasive Therapy & Allied Technologies* 185S. 253–258, 2009

[4] G Toti, M Garbey, V Sherman, B L Bass, B J Dunkin, “A Smart Trocar for Automatic Tool Recognition in Laparoscopic Surgery”, *Surgical Innovation* 221S. 77–82, 2015

[5] H Nakawala, E De Momi, A Morelli, C Tomasina, G Ferrigno, “Ontology-based surgical assistance system for instruments recognition”, *CRAS Joint Workshop On New Technologies For Computer/Robot Assisted Surgery*, 2015

[6] C Meißner, T Neumuth, “RFID-based surgical instrument detection using Hidden Markov models”, *Biomedizinische Technik. Biomedical engineering* 57, 2012

[7] K Sarkar, K Varanasi, D Stricker, “Trained 3D Models for CNN based Object Recognition”, *VISIGRAPP* 5S. 130–137, 2017

Electrical Impedance Tomography Lung Imaging with partial access to the thorax: A simulation study

J. Mrongowius[1], A. W. Reske[1,2], T. Neumuth[1], P. Salz[1]

[1]University of Leipzig, Innovation Center Computer Assisted Surgery, Leipzig, Germany
[2]Heinrich-Braun-Hospital, Department of Anesthesiology, Intensive Care Medicine, Emergency Medicine and Pain Management, Zwickau, Germany

Contact: julia.mrongowius@medizin.uni-leipzig.de

Abstract

For the potential use of electrical impedance tomography (EIT) in prehospital emergency scenarios, situations of poor access to the patient must be taken into account. In this case, it is not possible to place all electrodes on the thorax. To still allow EIT measurements in this situation, the effect of a reduced set of electrodes needs to be examined. In this simulation study, EIT measurement sets are manually modified and reconstructed to simulate the situation of a reduced electrode set. The resulting images are compared to the original image reconstructed with the full electrode set. These first simulation results show a severe degradation of image quality and interpretability even for a small number of missing electrodes.

Keywords: Electrical Impedance Tomography, Simulation Study, Electrode Configuration

1 Problem

The initial assessment of the patient's ventilation and vital condition in a prehospital environment such as situations of emergencies is a challenging task. Electrical impedance tomography (EIT) is a noninvasive imaging method for the monitoring of lung ventilation [3] with the potential of a mobile application in prehospital situations. However, the existing commercially available EIT devices require a full enclosement of the patient's thorax with electrodes. Regarding hard-to-reach patients, e.g. trapped patients after car crashes, it must be considered that only parts of the patient's thorax are accessible for the initial examination of the vital state. In such cases, only a reduced number of electrodes can be attached to the patient's thorax. In this simulation study, an examination of a successive reduction of electrodes on one side of a measurement target is conducted. The results are compared to the image with a complete electrode set and the effect on the EIT-image of the reduction is described.

2 Material and Methods

In order to examine the effect of EIT imaging with a reduced electrode set, a simulation study regarding a simplified phantom is conducted. Evaluating the findings on real data, the same simulation study for real measurement data is performed successively. Both studies are performed with MATLAB[1] and EIDORS [1], an open source toolbox providing functions for the forward and inverse modeling of EIT.

For the forward simulation of EIT-measurements, a simple cylindrical 2.5-D phantom is generated virtually with the help of the EIDORS toolbox. In the center of the phantom a cylinder of different conductivity than the background is added. The 16 electrodes are placed on one plane in the middle of the phantom's outer surface. A set of measurements is calculated. As reference data, a set of measurements without the cylinder of different conductivity is computed. To make the simulated voltages more realistic, noise is added. Subsequently, the simulated data is reconstructed with the GREIT algorithm [2]. In the two measurement sets, the voltages corresponding to one electrode are deleted and the remaining voltages are reconstructed with the help of an adapted inverse model. The measurements sets are reduced successively as described above until seven electrodes remain and in each step an image is reconstructed. The electrodes are reduced on one (the dorsal) side of the object to simulate the situation of a trapped patient with no access to his back.

[1] MATLAB R2013b (The MathWorks, Inc., Natick, Massachusetts, United States).

In order to evaluate the effect of a reduced set of measurements in a real setting, the same approach is conducted with real measured data of a pig, which is recorded in a pre-clinical study using the commercially available EIT-device PulmoVista 500[2]. The inverse model for the reconstruction is generated based on a computer tomography-based segmentation of the animal's thorax. Again, the measured data is reduced stepwise until seven of the 16 electrodes remain. For the evaluation, a mean tidal image over the signal of one minute is calculated for each electrode set.

3 Results

In Fig. 1, the reconstructed images for the simulation study on the cylindrical phantom (upper row) and the real data of a ventilated pig (lower row) are depicted. In order to illustrate the increasing impact on the image of a progressive reduction of electrode, selected results of decreasing number of electrodes are visualized. In the first column, the reconstruction of the complete electrode set is shown. The following columns show the reconstruction results for decreased number of electrodes, which are given above each image. In the images of the second column in Fig. 1, the original object is still recognizable. In the third column, the objects are smeared towards the borders, while the last column shows image results that are not visually similar anymore to the original ones. Regarding here the simulated image in the upper row, the image seems to belong to a phantom of two objects, since the original single object is smeared to the boundaries and one blob on each side occurs. A similar effect can also be seen in the animal data for 7 electrodes. Conversely, here the two blobs on the left and right side of the thorax result from the two lungs, but the original significant difference in the amount of ventilation (the right lung on the left side of the image is more ventilated in the original image), cannot be distinguished anymore.

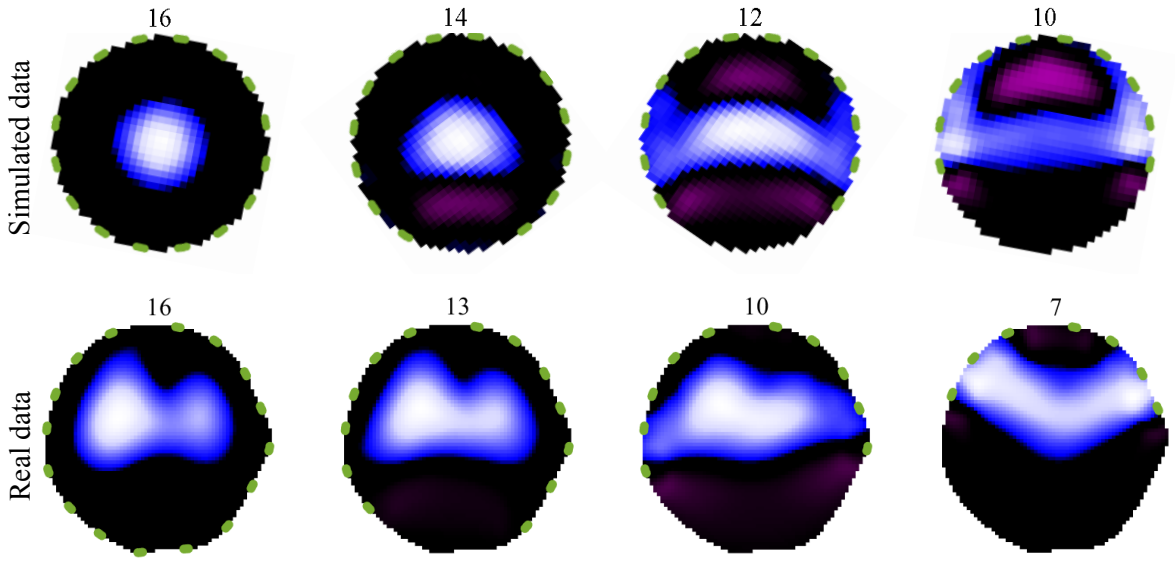


Figure 1*: Simulation results of a successively reduced number of electrodes on one side (dorsal) of the simulated EIT-voltages with the cylindrical phantom (upper row) and for a tidal image in real pig data (lower row). White pixels indicate positive impedance changes, purple pixels indicate negative changes. The number of electrodes is given above each image. The available electrodes are marked as green rectangles.*

4 Discussion

The results show a distinct distortion of the imaged object when reducing the measuring electrodes. Especially regarding the simulated data reconstructed with ten electrodes, the image interpretation of two objects in the phantom is problematic, since such a misled interpretation in an emergency situation could imply seeing two ventilated lungs, instead of one.

[2] Drägerwerk AG & Co. KGaA, Lübeck, Germany.

5 Conclusion

In this study, the effect of a reduced set of electrodes for Electrical Impedance Tomography was examined on two data sets. It was observed that a small number of reduced electrodes already have a strong effect on the resulting image and the depiction of the imaged object. Thus, it can be concluded, that the simple reduction of electrodes may result in misleading image interpretations and more elaborate methods have to be developed to handle incomplete access to a patient's thorax.

Acknowledgement

This project is sponsored by the German Federal Ministry of Education and Research (BMBF) as part of the KMU-Innovative-Program (project number: 13GW0179C).

References

[1] A. Adler and W. R. B. Lionheart, *Uses and abuses of EIDORS: An extensible software base for EIT*, Physiological Measurement, 27(5) 25-42 (2006)

[2] A. Adler, J. H. Arnold, R. Bayford, A. Borsic, B. Brown, P. Dixon, T. J. C. Faes, et al., *GREIT: a unified approach to 2D linear EIT reconstruction of lung images,* Physiological Measurement, 30(6) 35–55 (2009)

[3] I. Frerichs, M. B. P. Amato, A. H. van Kaam, D. G. Tingay, Z. Zhao, B. Grychtol, M. Bodenstein, et al., Chest Electrical Impedance Tomography Examination, Data Analysis, Terminology, Clinical Use and Recommendations: Consensus Statement of the TRanslational EIT DevelopmeNt StuDy Group, Thorax, 72(1) 83-93 (2017)

Risikomodell zur Abschätzung des intraoperativen Verletzungsrisikos an der lateralen Schädelbasis

C. Voigtmann[1], J. Bredemann[1], A. Knott[1], R. H. Schmitt[1], I. Stenin[2], J. Schipper[2],J. Kristin[2]

[1]RWTH Aachen, Lehrstuhl für Fertigungsmesstechnik und Qualitätsmanagement, Aachen, Deutschland

[2] Hals-Nasen-Ohrenklinik, Universitätsklinikum, Düsseldorf, Deutschland

Kontakt: c.voigtmann@wzl.rwth-aachen.de

Abstract

Neue, minimalinvasive Operationsverfahren können erst dann eingeführt werden, wenn das mit der Operation einhergehende Risiko für den Patienten vertretbar ist. Das gilt vor allem für hochkomplexe Operationsverfahren, bei denen die Gefahr besteht Risikostrukturen irreversibel zu verletzen. Trotz der Forderungen nach einheitlichen, objektiven Bewertungsverfahren für die Erprobung und Anwendung von Operationsverfahren, existieren derzeit keine einheitlichen Methoden. Speziell für die Risikobewertung minimalinvasiver Operationen an der lateralen Schädelbasis bestehen verschiedene Risikomodelle. Bisher ist es jedoch nicht möglich ein Gesamtverletzungsrisiko anzugeben. Im Folgenden wird daher zum ersten Mal ein Risikomodell vorgestellt, mit dem das intraoperative Gesamtrisiko unbeabsichtigter Verletzungen für minimalinvasive Eingriffe an der lateralen Schädelbasis präoperativ und patientenindividuell abgeschätzt werden kann.

Keywords: Medizinische Risikoanalyse, Verletzungsrisiko, Unsicherheit

1 Problemstellung

Das Konzept neuer, minimalinvasiver Zugangswege zur lateralen Schädelbasis beruht auf dem Anlegen von mindestens einem Bohrkanal zu einem definierten Zielpunkt im Operationsgebiet [1–4]. Durch diese Methode soll ein geringeres Trauma verbunden mit einer schnelleren Genesungszeiten des Patienten bestehen. Im Vergleich zur konventionellen Chirurgie an der lateralen Schädelbasis, die in der Regel als Grundlage eine Mastoidektomie beinhaltet, welche entsprechend der Indikation erweitert wird, mangelt es dem Chirurgen jedoch bei den neuen Verfahren sowohl an einem optischen als auch haptischem Feedback. Dadurch besteht die Gefahr Risikostrukturen wie beispielsweise den Gesichtsnerv unbeabsichtigt zu verletzen. Neben mechanischen Einwirkungen durch den Bohrer sind in diesem Zusammenhang auch thermische Einflüsse durch die Wärmeentwicklung am Bohrer von Bedeutung [5]. Aus diesen Gründen liegt der Fokus derzeitiger technischer Entwicklungen vor allem auf der sensorgestützten Absicherung des Bohrprozesses [1, 6].

Durch erste klinische Studien konnte die Durchführbarkeit minimalinvasiver Eingriffe im Bereich der lateralen Schädelbasis am Beispiel von Cochlea-Implantationen nachgewiesen werden [1, 5]. Neue, innovative Operationsverfahren können jedoch erst dann in der Breite implementiert werden, wenn das mit der Operation einhergehende Risiko für den Patienten vertretbar ist. Trotz der Forderung nach objektiven Ansätzen zur Bewertung innovativer Operationsverfahren mangelt es bisher an einheitlichen Methoden [7]. Im Folgenden wird daher ein Risikomodell vorgestellt, welches eine Abschätzung des Verletzungsrisikos bei minimalinvasiven Operationen im Bereich der lateralen Schädelbasis präoperativ und patientenindividuell ermöglicht.

2 Bestehende Risikomodelle

Forderungen nach einem regulatorischen Rahmen für die Erprobung und Anwendung neuer Operationsverfahren wurden im Jahr 2009 von BARKUN et al. [8] formuliert. In diesem Zusammenhang ist das für den Patienten resultierende Risiko von besonderer Bedeutung. Die Forderungen nach einem regulatorischen Rahmen wird in Form der Beschreibung der IDEAL-Empfehlungen durch MCCULLOCH et al. [9] aufgegriffen und detailliert. Die beschriebene Vorgehensweise bietet einen groben Rahmen für die Entwicklung und Erprobung neuer

Operationstechniken. Der Fokus der Arbeiten liegt jedoch nicht auf der patientenindiviuellen Ermittlung des Verletzungsrisikos [7].

Ein Ansatz zur Bestimmung der Sicherheit von minimalinvasiven Cochlea-Implantationen wird von NOBLE et al. [10] vorgestellt. Das entsprechende Risikomodell variiert den erreichten Zielpunkt durch einen Störterm und ermittelt auf Basis einer Monte-Carlo-Simulation in wie viel Prozent der Operationen mechanische Verletzungen von Risikostrukturen auftreten. Relevante Einflüsse wie die Positionierunsicherheit des Bohrers und die Temperaturentwicklung am Bohrer werden jedoch nicht berücksichtigt [7].

Ein weiterer Ansatz für die Risikoanalyse minimalinvasiver Eingriffe im Bereich der lateralen Schädelbasis beschreiben NAU et al [11]. Dieser basiert auf der Übertragung eines etablierten Fähigkeitskennwertes zur Beurteilung von Produktionsprozessen in den medizinischen Kontext. Das beschriebene Therapeutische Sicherheitsmaß (Therapeutic Risk Index) ermöglicht die präoperative Abschätzung des mechanischen Verletzungsrisikos auf Basis der Unsicherheiten der präoperativen CT-Bildgebung und der Unsicherheit der Bohrerposition und –ausrichtung. In Kombination mit den patientenindividuell ermittelten Abständen zu den Risikostrukturen wird die „Fähigkeit“ des Operationsprozesses ermittelt. Wenn der ermittelte Kennwert einen definierten Grenzwert überschreitet, wird das Verletzungsrisiko als ausreichend gering angesehen. Neben den Unsicherheiten, die aus dem Bohrprozess resultieren wird bei dieser Art der Modellbildung auch das thermische Verletzungsrisiko vernachlässigt.

Weitergehende Ansätze zur Abschätzung des Risikos von minimalinvasiven Eingriffen an der lateralen Schädelbasis wurden von den Autoren dieses Papers bereits 2016 [12] vorgestellt. Neben einem Risikomodell, das zur Abschätzung des mechanischen Verletzungsrisikos herangezogen werden kann, wurde auch ein Risikomodell für die Abschätzung thermischer Verletzungen entwickelt. Um präoperativ ein Gesamtrisiko angeben zu können, fehlt jedoch noch ein einheitliches Risikomodell, das die Auftretenswahrscheinlichkeit von mechanischen und thermischen Verletzungen kombiniert. Im Folgenden wird daher zum ersten Mal ein Risikomodell vorgestellt, mit dem das Gesamtrisiko unbeabsichtigter Verletzungen für minimalinvasive Eingriffe an der lateralen Schädelbasis präoperativ und patientenindividuell abgeschätzt werden kann. Dafür werden zunächst die beiden Risikomodelle erläutert, auf denen das Gesamtmodell basiert.

2.1 Risikomodell für mechanische Verletzungen

Das Modell zur Abschätzung des Risikos für mechanische Verletzungen basiert auf der Annahme, dass die kritische Distanz zwischen dem geplanten Bohrkanal und den verschiedenen Risikostrukturen groß genug sein muss, um die Unsicherheiten entlang des Operationsprozesses zu kompensieren. Die kritische Distanz zur sensitiven Struktur ist der minimale Abstand d zwischen dem Rand des Bohrpfades (mit Punkt P auf der Mittelachse und Radius r) und der sensitiven Struktur (Punkt S) (vgl. Abbildung 1) :

$$d = ||\mathbf{P} - \mathbf{S}|| - r \qquad (1)$$

Die Prozessunsicherheiten e, die dazu führen können, dass dieser Abstand in der Realität kleiner sein kann, werden als Zufallsvariablen mit Erwartungswert b und Varianz u^2 modelliert. b bezeichnet systematische Abweichungen (bias) und u die zufälligen Prozessschwankungen (uncertainty). Folgende Unsicherheiten werden für minimalinvasive Operationen der lateralen Schädelbasis als kritisch bewertet [12, 13]

- Positionier- und Ausrichtungsfehler des Bohrers, genannt Navigationsfehler e_{nav},
- Fehler in der Erkennung der sensitiven Strukturen durch das Bildgebungsverfahren e_{imag},
- Unsicherheit des Bohrprozesses e_{drill}, z. B. das Verlaufen des Bohrers.

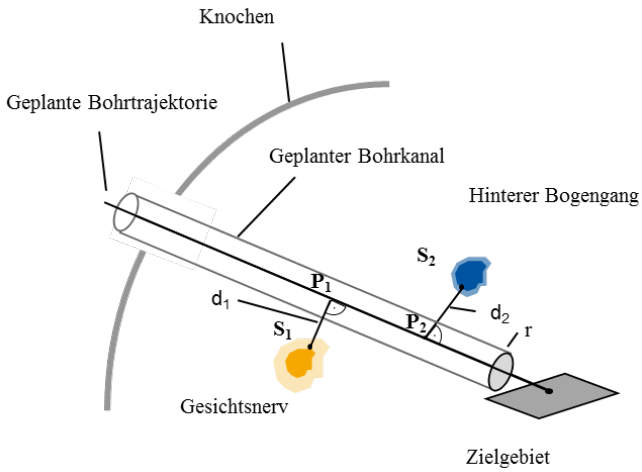


Abbildung 1: Planungsszenario schematisch

Daraus wurde bereits von den Autoren dieses Beitrages 2016 [12] die Wahrscheinlichkeit angegeben, eine sensitive Struktur mechanisch zu verletzen. Die Wahrscheinlichkeit wird berechnet durch:

$$P(\text{"mech. Treffen min. einer Struktur"}) = 1 - P(\text{"keine Struktur wird getroffen"})$$
$$= 1 - \prod_{i=1}^{m} P(\text{"Struktur } \mathrm{S_i} \text{ wird nicht getroffen"}) = 1 - \prod_{i=1}^{m} \mathrm{F_{hit,i}}(d_i) \qquad (2)$$

Dabei ist $\mathrm{F_{hit,i}}$ die Verteilungsfunktion der Zufallsvariable, die die oben genannte Prozessunsicherheit entlang des minimalen Abstandes d_i zur i-ten kritischen Struktur modelliert. Insgesamt sind m sensitive Strukturen entlang des Bohrpfades vorhanden, die in die Risikobetrachtung einbezogen werden müssen.

2.2 Risikomodell für thermische Verletzungen

Das zweite große Risiko bei minimalinvasiven Bohrprozessen ist die Temperaturentwicklung, die während des Bohrvorgangs zu Schädigungen des umliegenden Gewebes führen kann. Diese Nekrose des Gewebes tritt ab ca. 43°C auf, welche bei üblichen Bohrparametern schnell überschritten werden kann [14, 15]. Die Höhe der Temperatur alleine ist jedoch nicht entscheidend für eine Schädigung, sondern der gesamte Energieeintrag in das Gewebe über die Zeit. Dies wird durch den Kennwert der kumulativen Äquivalenzminuten bei 43°C (CEM43) ausgedrückt [16–18]. Der Kennwert gibt an, wie lange ein Gewebe einer Temperatur von 43°C ausgesetzt sein müsste, um denselben Energieeintrag zu erhalten. Der CEM43-Wert wird in Minuten angegeben und ist definiert als:

$$CEM43(T) = \int_0^{t=t_{final}} R\big(T(t)\big)^{43-T(t)} dt\,, \qquad R\big(T(t)\big) = \begin{cases} 0{,}5\,, & T(t) > 43°C \\ 0{,}25\,, & T(t) \leq 43°C \end{cases} \qquad (3)$$

$T(t)$ bezeichnet hierbei den Temperaturverlauf an der sensitiven Struktur über die Zeit $t \in [0, t_{final}]$ und t_{final} die Dauer der Temperatureinwirkung (in Minuten). In einem Übersichtsbeitrag von AUGUSTIN ET AL. [19] werden Material und anwendungsabhängige Grenzwerte diskutiert. Innerhalb dieser Risikobetrachtung wird angenommen, dass eine permanente thermische Schädigung ab einem Wert von $CEM43 = 60$ min auftritt [17–19].

Ein Risiko für den Patienten besteht nur dann, wenn dieser Wert in der Planung unsicher abgeschätzt werden kann oder während der Durchführung nur unsicher bestimmt werden kann. Wären der CEM43-Wert und sein Grenzwert exakt bekannt, könnte mit Sicherheit entschieden werden, ob eine Nekrose auftritt oder nicht. Die unabdingbare Unsicherheit von Temperaturmessungen oder anderen Arten der Abschätzung der Temperatur führt daher immer zu einer Unsicherheit des CEM43-Wertes und damit zu einem Risiko für den Patienten. Trotz dieses Zusammenhanges wurde jedoch nach Informationen der Autoren noch keine Unsicherheitsbetrachtung für den CEM43-Wert durchgeführt. Innerhalb dieses Beitrages wird vor allem auf die Auswirkung dieser Unsicherheit auf das Patientenrisiko eingegangen. Zur Kompensation dieses Umstandes werden zurzeit noch Bohrstrategien mit langen Abkühlpausen und „Angsttoleranzen“ verwendet [1, 16, 20], welche zum einen die Eingriffszeit erhöhen und zum anderen eine Optimierung des Prozesses auf Basis eines gesamtheitlichen Prozessverständnisses erschweren.

Für die Berechnung des Risikos einer thermischen Verletzung ist das Verhältnis aus Unsicherheit des CEM43-Wertes zum verbleibenden Abstand zum kritischen Wert entscheidend. Bezeichne e_{CEM43} nun diese Unsicherheit

des CEM43-Wertes in Form der Standardabweichung einer Zufallsvariablen und F_{CEM43} die zugehörige Verteilungsfunktion der Zufallsvariablen. Der Erwartungswert der Zufallsvariablen entspricht dabei dem ermittelten oder prognostizierten mittleren CEM43-Wert. Damit ergibt sich für die Wahrscheinlichkeit einer thermischen Verletzung für eine einzelne sensitive Struktur für den Grenzwert 60 min zu:

$$P(\text{"Thermische Verletzung"}) = 1 - \text{P("Keine thermische Verletzung")}$$
$$= 1 - P(e_{CEM43} < 60) = 1 - F_{CEM43}(60) \qquad (4)$$

Daraus kann nun wieder, wie bei der Wahrscheinlichkeit für mechanische Verletzungen in Gleichung (2), die Wahrscheinlichkeit für die thermische Verletzung bei mehreren sensitiven Strukturen abgeleitet werden. Es wird dazu angenommen, dass die Wahrscheinlichkeiten von thermischen Verletzungen an verschiedenen sensitiven Strukturen unabhängig voneinander sind. Damit ergibt sich die Wahrscheinlichkeit einer thermischen Verletzung an mindestens einer der *m* sensitiven Strukturen (mit minimalen Abständen d_i) als

$$P(\text{"Therm. Schädigung min. einer sens. Struktur"}) = 1 - \text{P(Keine sens. Struktur wird therm. geschädigt")}$$
$$= 1 - \prod_{i=1}^{m} P(\text{"Keine therm. Schädigung der Struktur } S_i\text{"})$$
$$= 1 - \prod_{i=1}^{m} F_{CEM43,d_i}(60) \qquad (5)$$

mit F_{CEM43,d_i} ist die Verteilungsfunktion des CEM43-Wertes für die i-te sensitive Struktur.

Damit kann nun erstmalig das Risiko für eine thermische Verletzung des Patienten auf Basis einer theoretischen Risikomodellierung angegeben werden, sofern die Unsicherheit des CEM43-Wertes bekannt ist.

3 Modellierung des Gesamtrisikos

Beide Teilmodelle werden nun zu einem Gesamtmodell für das Patientenrisiko kombiniert. Es wird in erster Näherung davon ausgegangen, dass das thermische und das mechanische Risiko unabhängig voneinander sind. Damit ergibt sich die gesamte Verletzungswahrscheinlichkeit als Kombination aus Gleichung (2) und (5):

$$P(\text{"Patient wird geschädigt"}) = 1 - P(\text{"Keine Schädigung des Patienten"})$$
$$= 1 - P(\text{"Keine sens. Struktur wird therm. geschädigt"})$$
$$\cdot P(\text{"Keine sens. Struktur wird getroffen"})$$
$$= 1 - \left(\prod_{i=1}^{m} F_{CEM43,d_i}(60) \cdot \prod_{i=1}^{m} F_{hit,i}(d_i)\right) \qquad (6)$$

Zusammenfassend lässt sich festhalten, dass das Patientenrisiko für minimalinvasive Bohrprozesse durch folgende Wahrscheinlichkeit beschrieben wird:

$$P(\text{"Patient wird geschädigt"}) = 1 - \left(\prod_{i=1}^{m} F_{CEM43,d_i}(60) \cdot \prod_{i=1}^{m} F_{hit,i}(d_i)\right)$$

mit

F_{CEM43,d_i}	Verteilungsfunktion des CEM43-Wertes für die *i*-te sens. Struktur (siehe Gleichung (4))
$F_{hit,i}$	Gemeinsame Verteilungsfunktion aller Unsicherheitskomponenten entlang des minimalen Abstandes zur *i*-ten sens. Struktur (siehe Gleichung (2))

Mit diesem Modell lässt sich direkt die Wahrscheinlichkeit für eine Verletzung des Patienten berechnen.

4 Anwendungsbeispiel

Basis für die Zahlenwerte in diesem Anwendungsbeispiel ist das Operationsszenario für eine Cochlea-Implantation (SCHIPPER ET AL. [21, 22], SCHMITT ET AL. [23], NAU [7] und POLLMANNS [24, 25]). Es wird in diesem Beispiel davon ausgegangen, dass systematische Abweichungen bekannt und kompensiert sind und alle Unsicherheiten normalverteilt sind. Das Risiko für eine mechanische Schädigung einer sensitiven Struktur wurde bereits in [6] berechnet und angegeben. Die Zahlenwerte werden von dort übernommen. Die

Unsicherheitsbeiträge von wurden von NAU und POLLMANNS anhand von Phantomen bestimmt. Neu ist die Berechnung des thermischen Risikos und die Kombination zu einem Gesamtrisiko.

4.1 Risikos eines mechanischen Treffers

Für einen durchschnittlichen Patienten müssen ca. 5 sensitive Strukturen in der Risikobetrachtung beachtet werden [22, 26]. Diese liegen meist in einem mittleren Abstand von 1 mm von der Bohrtrajektorie entfernt. Im Rahmen dieses Beispiels wird der minimale Abstand für alle sensitiven Strukturen als 1 mm angenommen und diese Abstände werden in den Bohrtiefen 1,2,3,4 und 5 mm erreicht. Die Auswirkung des initialen Ausrichtungsfehlers nimmt mit zunehmender Bohrtiefe zu, wodurch auch das Risiko tiefenabhängig ansteigt. Die von Nau bestimmte Unsicherheit der Navigation wurde für die maximale Tiefe von *l=5* mm bestimmt und muss für geringere Bohrtiefen linear verringert werden. Damit ergeben sich die in Tabelle 1 dargestellten Wahrscheinlichkeiten.

Tabelle 1 - Tiefenabhängige Wahrscheinlichkeit eine sensitive Struktur mechanisch zu verletzen

Bohrtiefe *l [in mm]*	1	2	3	4	5
Abstand sens. Struktur *[d in mm]*			1		
Unsicherheit Navigation $u_{nav,i}$	0,0634	0,1268	0,1902	0,2536	0,3170
Unsicherheit Bohrprozess u_{drill}			0,049		
Unsicherheit Bildgebung u_{imag}			0,382		
Wahrscheinlichkeit eines Treffers	$5{,}2 \cdot 10^{-3}$	$6{,}8 \cdot 10^{-3}$	$1{,}0 \cdot 10^{-2}$	$1{,}5 \cdot 10^{-2}$	$2{,}2 \cdot 10^{-2}$

Mit Gleichung (2) ergibt sich daraus eine Gesamtwahrscheinlichkeit mindestens eine sensitive Struktur direkt zu treffen von ca. 5,8%. Das bedeutet, dass im Mittel bei einer von 17 Operationen eine sensitive Struktur verletzt wird. Grund für die hohe Verletzungswahrscheinlichkeit ist die hohe Unsicherheit in der CT-Bildgebung sowie die Unsicherheit in der initialen Positionierung des Bohrers.

4.2 Risiko einer thermischen Verletzung

Typische Bereiche für die Cochlea-Implantation liegen nach Feldmann et al. bei CEM43-Werten von 0,1 min bis 25 min [16, 20]. Die Bohrstrategie und Bohrparameter sollten so gewählt werden, dass ein Sicherheitsabstand zum Grenzwert für eine thermische Verletzung eingehalten wird. Da noch keine Vorgaben oder Erfahrungswerte existieren, wird in diesem Beispiel ein CEM43-Wert von 6 min angenommen, was 1/10 des Grenzwertes von 60 min entspricht. Die Abschätzung der Unsicherheit des CEM43-Wertes wird aktuell noch erforscht. Erste Abschätzungen liegen aber in Bereichen von 10-50 min für die bei Feldmann et al. angegeben Temperaturverläufe [16, 20]. Um einen ersten Eindruck über die Größenordnung des Risikos zu gewinnen, wird die Unsicherheit des CEM43-Wertes für dieses Beispiel mit $u_{CEM43} = 32{,}39$ min für alle sensitiven Strukturen angenommen. Die Wahrscheinlichkeit, eine sensitive Struktur thermisch zu verletzen, liegt damit unter Verwendung von Gleichung (4) bei ca. 4,8% und die Wahrscheinlichkeit keine der 5 sensitiven Strukturen thermisch zu verletzen mit Gleichung (5) bei ca. 21,7%.

4.3 Bestimmung des Gesamtrisikos

Für das gesamte Patientenrisiko wird die thermische Verletzungswahrscheinlichkeit und die Wahrscheinlichkeit eines direkten Treffers nach Gleichung (6) kombiniert. Es ergibt sich eine Gesamtwahrscheinlichkeit für die Verletzung des Patienten im angenommen Operationsszenario von 26,2%. Eine Verletzung des Patienten bei jedem vierten Eingriff ist weit von einer vertretbaren Auftretenswahrscheinlichkeit entfernt. Dies liegt hauptsächlich an der hohen angenommen Unsicherheit des CEM43-Wertes. Weitere Untersuchungen zur Unsicherheit des CEM43-Wertes sind daher ratsam.

5 Diskussion

Das vorgestellte Risikomodell ist durch die Verwendung von Zufallsvariablen allgemein gehalten. Die Unsicherheitskomponenten können beliebig getauscht oder durch weitere Unsicherheitskomponenten ergänzt werden. Da die einzelnen Unsicherheitsbeiträge für die verschiedene Risikostrukturen variieren, muss jede der

Unsicherheitskomponenten spezifisch für jede Risikostruktur mit einer Wahrscheinlichkeitsdichtefunktion beschrieben werden. Darüber hinaus ist das Risikomodell unabhängig von der Art der chirurgischen Bohrung und kann auch ohne weiteres auf andere Bildgebungsverfahren (z.B. MRT), weitere Risikostrukturen und andere Positionierverfahren angewandt werden. Es müssen lediglich die Abstände von dem geplanten Bohrkanal zur den Risikostrukturen angegeben werden können. Die Modelle für die Bestimmung des Risikos für thermische und mechanische Verletzungen sind bisher vor allem die Planungsphase von Operationen ausgerichtet. Zusätzliche Informationen wie beispielsweise die Bestimmung der Bohrerposition oder der gemessenen Bohrerposition können in zukünftige Modelle integriert werden.

Das errechnete Gesamtrisiko ist mit 26,2% noch zu hoch, um das zugrunde gelegte Operationsverfahren [3] zulassen zu können. Durch die Bestimmung der Unsicherheitsbeiträge der einzelnen Komponenten ist es jedoch möglich die Komponenten/Modalitäten zu identifizieren, die weiterentwickelt werden müssen. Darüber hinaus wurde der Unsicherheitsbeitrag der Bildgebung, der von der jeweiligen Risikostruktur abhängig ist, zur sicheren Seite hin abgeschätzt [27]. Zukünftig ist die Unsicherheit in Abhängigkeit von der jeweiligen Messaufgabe zu berücksichtigen. Zudem ist zu beachten, dass das Risikomodell nur das Prozessrisiko abdeckt. Es wird davon ausgegangen, dass die Planung und Durchführung des Prozesses fehlerfrei erfolgt. Zudem werden Vitalparameter des Patienten noch nicht berücksichtigt.

6 Zusammenfassung und Ausblick

Mit dem entwickelten Risikomodell kann erstmalig das Patientenrisiko für minimalinvasive Bohrprozesse in der Planungsphase angegeben und damit entschieden werden, ob der Eingriff durchgeführt werden soll. Das Risikomodell kann als einheitliche Grundlage für die Bewertung der Verfahren, insbesondere für die Zulassung, genutzt werden. Weiterhin bildet es erstmalig ein Beschreibungsmodell, mit dem eine detaillierte Analyse der Risikostruktur möglich ist. Damit kann eine zielgerichtete Optimierung der wesentlichen Unsicherheitskomponenten hin zu einem für die Anwendung der Eingriffe akzeptablen Risiko durchgeführt werden. Das dies notwendig ist, zeigt die Anwendung auf das konkrete Beispiel der Cochlea-Implantation, bei dem das Patientenrisiko mit 26,2% abgeschätzt wurde.
Aktuell sind alle Modelle noch auf die Planungsphase der Operation ausgerichtet. Zusätzliche Informationen, die während der Operation hinzukommen, wie bspw. reale Positionen oder gemessene Temperaturwerte können in zukünftige Modelle integriert werden. Zum Einsatz des Risikomodells ist es notwendig, die Unsicherheit des CEM43-Wertes zu kennen. Dies ist zum aktuellen Zeitpunkt größtenteils unbekannt und sollte weiter erforscht werden. Zur Senkung des Risikos sollte vor allem an der Beherrschung der Temperatur und des Energieeintrages geforscht werden, da dies den größten Unsicherheitsbeitrag liefert. Das Risiko eines direkten Treffers alleine ist ebenfalls zu hoch und kann dadurch gesenkt werden, dass die Systeme zur Bildgebung, zur Segmentierung sowie zur Referenzierung und Positionierung des Bohrers hinsichtlich ihrer Unsicherheit verbessert werden.

Danksagung

Wir danken der Deutschen Forschungsgemeinschaft DFG für die Unterstützung und die Finanzierung des Forschungsprojektes „Messtechnische Absicherung eines Bohrprozesses bei bildgestützten minimalinvasiven Eingriffen am Beispiel der Otobasis“.

Referenzen

[1] Caversaccio, M.; Gavaghan, K.; Wimmer, W. et al.: Robotic cochlear implantation: surgical procedure and first clinical experience. Acta oto-laryngologica 137 (2017), Nr. 4, S. 447–454.

[2] Labadie, R. F.; Noble, J. H.; Dawant, B. M. et al.: Clinical validation of percutaneous cochlear implant surgery: initial report. The Laryngoscope 118 (2008), Nr. 6, S. 1031–1039.

[3] Stenin, I.; Hansen, S.; Nau-Hermes, M. et al.: Minimally invasive, multi-port approach to the lateral skull base. A first in vitro evaluation. International journal of computer assisted radiology and surgery 12 (2017), Nr. 5, S. 889–895.

[4] Majdani, O.; Rau, T. S.; Baron, S. et al.: A robot-guided minimally invasive approach for cochlear implant surgery: preliminary results of a temporal bone study. International journal of computer assisted radiology and surgery 4 (2009), Nr. 5, S. 475–486.

[5] Labadie, R. F.; Balachandran, R.; Noble, J. H. et al.: Minimally invasive image-guided cochlear implantation surgery: first report of clinical implementation. The Laryngoscope 124 (2014), Nr. 8, S. 1915–1922.

[6] Bredemann, J.; Voigtmann, C.; Jantzen, M. et al.: Bohrkopfintegrierte Sensorik zur Absicherung medizinischer Bohrprozesse. In: Burgner-Kahrs, J.; Kahrs, L. A.; Modes, V. et al. (Hrsg.): CURAC 2017 - Tagungsband. 16. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie (CURAC) : 5.-7. Oktober 2017, Hannover. Garbsen: PZH Verlag 2017.

[7] Nau-Hermes, M.: Ein messtechnischer Beurteilungsansatz für das Verletzungsrisiko anatomischer Strukturen bei bildgestützten minimalinvasiven Eingriffen. Aachen: Hochschulbibliothek der Rheinisch-Westfälischen Technischen Hochschule Aachen 2014.

[8] Barkun, J. S.; Aronson, J. K.; Feldman, L. S. et al.: Evaluation and stages of surgical innovations. The Lancet 374 (2009), Nr. 9695, S. 1089–1096.

[9] McCulloch, P.; Altman, D. G.; Campbell, W. B. et al.: No surgical innovation without evaluation. The IDEAL recommendations. The Lancet 374 (2009), Nr. 9695, S. 1105–1112.

[10] Noble, J. H.; Majdani, O.; Labadie, R. F. et al.: Automatic determination of optimal linear drilling trajectories for cochlear access accounting for drill-positioning error. The international journal of medical robotics + computer assisted surgery : MRCAS 6 (2010), Nr. 3, S. 281–290.

[11] Nau, M.; Pollmanns, S.; Schmitt, R. et al.: Assessing the risk of minimally-invasive surgery: a metrological approach. In: 16th International Congress of Metrology, Paris, France, October 7-10, 2013, S. 7002 2013.

[12] Bredemann, J.; Voigtmann, C.; Schmitt, R. et al.: Determining the patient's risk in minimally invasive surgery to the lateral skull base. 15.CURAC Annal Conference- Tagungsband der Jahrestagung der Deutschen Gesellschaft für computer- und roboterassistierte Chirurgie e.V.; 29.09- 01.10.2016, Bern/ Swiss (2016), S. 155–161.

[13] Nau-Hermes, M.; Schmitt, R.; Becker, M. et al.: Quality assurance of multiport image-guided minimally invasive surgery at the lateral skull base. BioMed research international 2014 (2014), S. 904803.

[14] Augustin, G.; Davila, S.; Mihoci, K. et al.: Thermal osteonecrosis and bone drilling parameters revisited. Archives of orthopaedic and trauma surgery 128 (2008), Nr. 1, S. 71–77.

[15] Davidson, S. R. H.: Drilling in Bone. Modeling Heat Generation and Temperature Distribution. Journal of Biomechanical Engineering 125 (2003), Nr. 3, S. 305.

[16] Feldmann, A.; Anso, J.; Bell, B. et al.: Temperature Prediction Model for Bone Drilling Based on Density Distribution and In Vivo Experiments for Minimally Invasive Robotic Cochlear Implantation. Annals of biomedical engineering 44 (2016), Nr. 5, S. 1576–1586.

[17] Sapareto, S. A.; Dewey, W. C.: Thermal dose determination in cancer therapy. International Journal of Radiation Oncology*Biology*Physics 10 (1984), Nr. 6, S. 787–800.

[18] Lee, J.; Rabin, Y.; Ozdoganlar, O. B.: A new thermal model for bone drilling with applications to orthopaedic surgery. Medical engineering & physics 33 (2011), Nr. 10, S. 1234–1244.

[19] Augustin, G.; Zigman, T.; Davila, S. et al.: Cortical bone drilling and thermal osteonecrosis. Clinical biomechanics (Bristol, Avon) 27 (2012), Nr. 4, S. 313–325.

[20] Feldmann, A.; Gavaghan, K.; Stebinger, M. et al.: Real-Time Prediction of Temperature Elevation During Robotic Bone Drilling Using the Torque Signal. Annals of biomedical engineering (2017).

[21] Schipper, J.; Klenzner, T.; Aschendorff, A. et al.: Navigiert-kontrollierte Kochleostomie. Ist eine Verbesserung der Ergebnisqualität in der Kochleaimplantatchirurgie möglich? HNO 52 (2004), Nr. 4, S. 329–335.

[22] Stenin, I.; Hansen, S.; Becker, M. et al.: MUKNO - Multi-Port-Knochenchirurgie am Beispiel der Otobasis. Virtuelle Planung und Machbarkeitsanalysemultiangulärer Bohrkanäle. In: Deutsche Gesellschaft für Computer- und Roboterassistierte Chirurgie (Hrsg.): CURAC2012@MEDICA. 11. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V. (CURAC), Düsseldorf, 15.-16.11., S. 181–184. Stuttgart 2012.

[23] Schmitt, R.; Nau, M.; Pollmanns, S. et al.: Bestimmung relevanter Einflussfaktoren auf die Unsicherheit minimalinvasiver Operationen am Beispiel der Otobasis. In: Deutsche Gesellschaft für Computer- und Roboterassistierte Chirurgie (Hrsg.): CURAC2012@MEDICA. 11. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V. (CURAC), Düsseldorf, 15.-16.11., S. 2–18. Stuttgart 2012.

[24] Pollmanns, S.: Bestimmung von Unsicherheitsbeiträgen bei medizinischen Computertomografiemessungen für die bildbasierte navigierte Chirurgie. Aachen: Apprimus Verlag 2014.

[25] Nau-Hermes, M.; Pollmanns, S.; Schmitt, R.: Uncertainty Evaluation for Surgical Processes. Key Engineering Materials 613 (2014), S. 317–326.

[26] Stenin, I.; Hansen, S.; Becker, M. et al.: Minimally invasive multiport surgery of the lateral skull base. BioMed research international 2014 (2014), S. 379295.

[27] Bredemann, J.; Voigtmann, C.; Schmitt, R. H.: Unsicherheitsbetrachtung für bildbasierte minimalinvasive Eingriffe an der Otobasis. tm - Technisches Messen 84 (2017), Nr. 5.

Präoperative Planung roboterassistierter Eingriffe mithilfe von augmentierter Realität und Robotik-Simulation

M. Stoljar[1], *O. Burgert*[1]

[1] *Hochschule Reutlingen, CaMed, Reutlingen, Deutschland*

Kontakt: maxim.stoljar@student.reutlingen-university.de

Abstract

Die minimal-invasive Chirurgie (MIC) entwickelt sich durch den Einsatz von medizinischen Robotern wie dem da Vinci-System von Intuitive Surgical stetig weiter. Hierdurch kann eine bessere oder gleichwertige Operation bei deutlich geringerer körperlicher Belastung des Operateurs erreicht werden. Dabei entstehen jedoch neue Problemstellungen wie beispielsweise Kollision zwischen Roboterarmen und die benötigte Zeit zum Einrichten einer geeigneten Roboterkonfiguration. Daher ist eine effiziente Vorbereitung und Planung der Interventionen erforderlich. Diese Arbeit präsentiert einen Ansatz für eine verbesserte Planung mit Augmented Reality (AR) und einer Robotik-Simulationssoftware (RS). die Robotik-Simulation dient zur Berechnung einer Roboterkonfiguration unter Vorgabe der Port-Positionen. Augmented Reality wird verwendet, um die berechneten Pose in der realen Umgebung zu visualisieren und somit leichter in den Operationssaal zu übertragen.

Keywords: Minimal-invasive Chirurgie, Augmented Reality, Robotik-Simulation

1 Problemstellung

Die größte Barriere für den Einsatz von roboterassistierte MIC ist neben dem enormen Anschaffungspreis, die benötigte Zeit zum Aufbau des Systems. Dabei ist der Aufbau entscheidend für den Ausgang der Operation, aber auch kostenintensiv, weil in dieser Zeit nicht operiert werden kann um Einnahmen mit dem System zu erzielen. Deshalb muss das Personal einen perfekten Spagat zwischen optimaler Operationsvorbereitung und minimalem Zeitaufwand zur Inbetriebnahme schaffen. [1]

Die Problemstellung kann dabei in zwei Aspekte untergliedert werden. Zum einen muss eine optimierte Platzierung ermittelt werden und zum anderen müssen diese Ergebnisse auf die reale Umgebung übertragen werden. Die Schwierigkeit bei der Platzierung besteht in der Auswahl der Port-Positionen, sowie der Positionierung der Roboterarme am Patienten. Wird dies suboptimal durchgeführt, können eine Reihe von Problemen auftreten. Die Zielanatomie ist aufgrund der ungeeigneten relativen Orientierung der Instrumente zum Zielbereich nicht erreichbar. Ein anderes Problem ist die Kollision von Instrumenten innerhalb des menschlichen Körpers oder Kollisionen der Roboterarme außerhalb des Patienten. Ebenso müssen Robotersingularitäten und die Robotergelenkanschläge für ein positives Operationsergebnis berücksichtigt werden. Diese Schwierigkeiten können zu einer Veränderung der Port-Positionen oder zur Neuausrichtung des Systems führen und in schweren Fällen zum Abbruch der Operation führen. [2] Um diese negativen Auswirkungen zu vermeiden, ist eine effiziente Vorbereitung der Operation erforderlich [3]. Der zweite Aspekt der Problemstellung adressiert die Übertragung der errechneten Roboterkonfiguration und Port-Position in den realen Operationssaal. Die Verwendung der präoperativen Planung ist nur dann möglich, wenn diese visualisiert oder in einer anderen Form auf das reale Setup angewendet werden kann.

2 Material und Methoden

2.1 Stand der Technik

Es gibt eine Reihe von Arbeiten die sich mit der automatischen Optimierung der Port-Platzierung und Roboterkonfiguration beschäftigen, wobei die Identifizierung der Port-Position als mathematisches Problem beschrieben wurde und die Visualisierung in einer virtuellen Welt dargestellt wurde [3–8]. Durch den Einsatz der virtuellen Realität war es nur schwer möglich die Ergebnisse in die realen Umgebungen zu übertragen.

Ein etwas anderer Ansatz zur Lösung des Planungsproblems ist der wissensbasierte Ansatz, wobei nur die Ergebnisse der Planung durch eine Mixed-Reality-Umgebung auf den menschlichen Körper projiziert wird [9,10]. Diese Arbeit zeigt an einfacheren Transfer der Planungsergebnisse, währenddessen nur die Port-Platzierung angezeigt wird und keine Darstellung der Roboterkonfiguration erfolgt.

Im Rahmen der Recherche wurde nur die Visualisierung der Port-Position mithilfe von Mixed Reality auf der Hautoberfläche des Patienten ermittelt. Dabei gibt es nach aktueller Recherche kein System, das die errechnete

Roboterkonfiguration visualisiert, die für die Platzierung der Roboterarme am Patienten notwendig wäre. Zudem gibt es kein System, dass die Visualisierung mit Hilfe von AR durchführt. Daher wird eine Planungsumgebung vorgeschlagen, um die Roboterkonfiguration zu optimieren und für die Inbetriebnahme in AR darstellt.

2.2 Planungsumgebung

Die Planungsumgebung besteht aus den zwei Hauptkomponenten Robotik-Simulation (RS) und Augmented Reality (AR). Dabei wird ein extern bezogenes Robotermodell [1] von der RS zur Berechnung der Roboterkonfiguration und von der AR zur Visualisierung in der realen Welt verwendet. Bei dem verwendeten Robotermodell handelt es sich um den da Vinci S von Inuitiv Surgical, der mit drei Armen ausgestattet ist, weil dieses System für eine mögliche Evaluierung zur Verfügung steht. Die RS wurde mit der Robotiksoftware V-Rep umgesetzt, wobei darin die Aufbereitung des Modells und die Modellierung der kinematischen Kette vorgenommen wurde. Die AR ist die zentrale Komponente, die die benötigten Informationen für den Benutzer visualisiert und Interaktionen wie Sprachsteuerung, Gestensteuerung und Feedback ermöglicht. Dabei interagiert der Benutzer nur über die AR mit der Planungsumgebung. Dazu wurde eine Applikation in der Spieleentwicklungs-Engine Unity3D erstellt, die auf die Microsoft HoloLens portiert wird, die anhand von eingebauter Rechenleistung, Sensorik sowie Optik in der Lage ist virtuelle Objekt innerhalb der realen Welt einzublenden. Die HoloLens ist eine kabelloses Head Mounted Display (HMD), das autark betrieben wird.

Im Falle dieser Arbeit wird die RS auf einem separaten Rechner ausgeführt und über WLAN mit der AR synchronisiert. Die synchronisierten Daten bestehen aus der Port-Position und Kommandos zum Starten der Berechnungen, die von der AR an die RS übertragen wird. Im Gegenzug wird die Gelenkkonfiguration an die AR übermittelt. Abbildung 1 zeigt den beschriebenen Datenaustausch. In unseren Ansatz kommuniziert der Benutzer nur mit der AR, während dessen die RS im Hintergrund läuft.

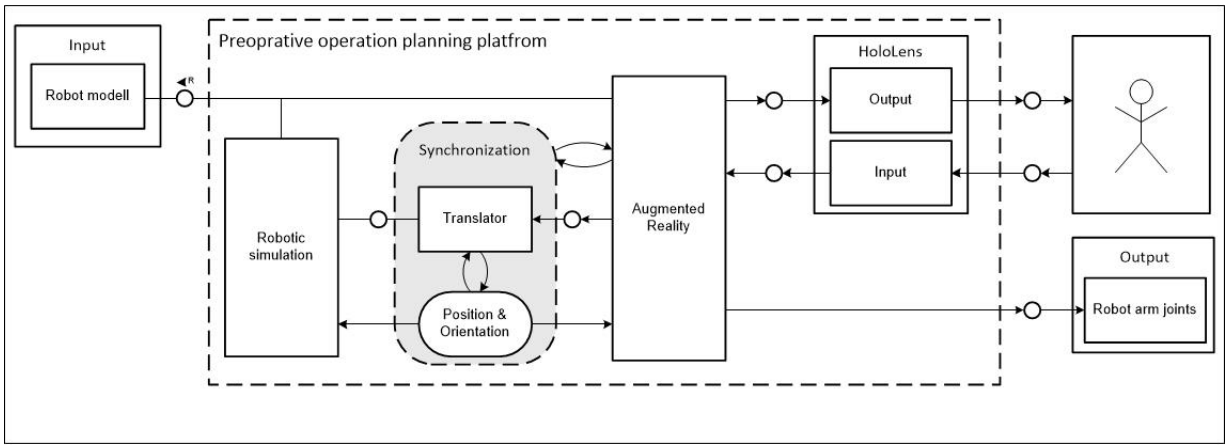


Abbildung 1: *Datenaustausch zwischen AR und RS*

2.2.1 Robotik-Simulation (RS)

Die Berechnung der Roboterkonfiguration wird anhand von Abbildung 2 verdeutlicht, wobei im ersten Teil des Bildes das Gesamtsystem, im zweiten Teil die Gelenke des mittleren Arms und im dritten Teil des Bildes die Gelenk der beiden äußeren Arme dargestellt sind. Die Ermittlung einer Konfiguration beinhaltet auf drei Einschränkungen. Als Erstes muss eine Konfiguration gefunden werden, bei der der Instrumenten-Rotationspunkt mit der Pose des entsprechenden Ports übereinstimmt. Als Zweites darf keine Kollision zwischen unterschiedlichen Roboterarmen und innerhalb eines Roboterarmes vorliegen. Als letztes müssen die Gelenke G3 der äußeren Roboterarme einen maximal großen Abstand zu dem Gelenk G2 des mittleren Roboterarms besitzen, um nun eine optimierte Konfiguration zu erhalten . Dabei ist der mittlere Arm immer in Richtung des rechten Arms ausgerichtet. Die Implementierung der Konfigurationsermittlung basiert auf der Berechnung von 15 unterschiedlichen Konfiguration, wobei die Erreichung der Port-Position ohne jegliche Kollision berechnet wird. Anschließend wird die Roboterarmkombination gewählt, die den größten Abstand der Gelenk G2 des mittleren Arms und G3 der beiden äußeren Arme besitzt.

[1] https://www.turbosquid.com/3d-models/max-da-vinci-robot/788008

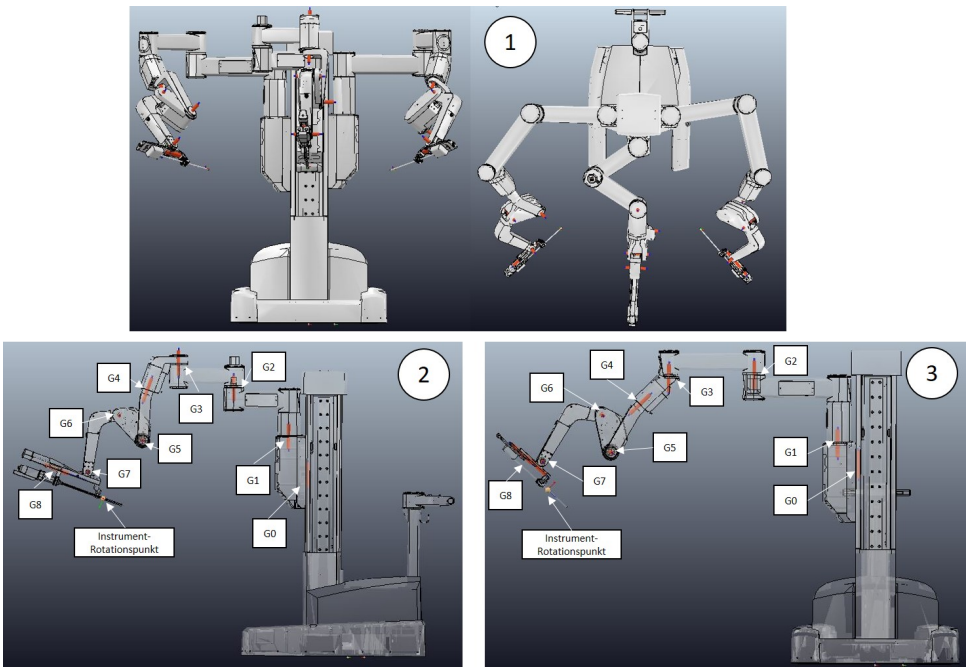


Abbildung 2: *1: Darstellung des Gesamtsystem aus Vorderansicht und Draufsicht. 2: Darstellung der Gelenkpositionen des mittleren Arms. 3: Darstellung der Gelenkpositionen der äußeren Arme*

2.2.2 Augmented Reality (AR)

Diese Komponente ist verantwortlich für die Hologramm-Platzierung in der realen Welt und die Synchronisation. Die Hologramm-Platzierung ist ein wichtiger Aspekt, um die Ergebnisse der Konfigurationsberechnung in die reale Welt zu übertragen. Daher sollte die Hologramm-Roboterbasis identisch über der realen Roboterbasis des chirurgischen Roboters liegen. Dabei kann die Hardware keine feste Pose für den Ursprung einrichten. Nach jedem Neustart der Anwendung wird der Ursprung des relativen Koordinatensystems auf die aktuelle Position des HMD gesetzt. Um diese Einschränkung zu lösen, wurden zwei Methoden entwickelt. Die erste Methode verwendet Marker, die die absolute Pose im Raum bestimmen. Dazu wird ein Marker auf der Roboterbasis verwendet, der von den eingebauten Kameras erkannt und durch Bildverarbeitung verarbeitet wird[2]. Zur Realisierung der Markererkennung wurde das Toolkit von Vuforia verwendet. Dabei wird ein auditives Feedback implementiert sobald ein Marker erkannt wird. Zusätzlich wurden die Sprachbefehle „Move Robot“, „Place Robot“, „Start Calculation“ und „Pause Calculation“ eingeführt, die ebenfalls mit auditiven Feedback versehen wurden. „Move Robot“ löst die starre Kopplung des Markers mit dem virtuellen Model, wobei sich das virtuelle Modell immer mit dem Marker bewegt. Wird der Befehl „Place Robot“ ausgelöst, wird die Kopplung entfernt und das Modell wird fest im Raum platziert. Die zweite Methode zur Platzierung des virtuellen Roboters über dem realen ist die Freihandplatzierung durch den Benutzer, die mit dem HoloToolkit von Microsoft realisiert wurde. Dabei wird die Blickrichtung sowie die Hand des Benutzers erkannt und Greifbewegungen registriert, die eine dreidimensionale Positionierung im Raum ermöglichen.

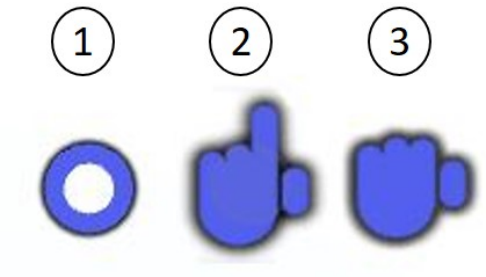


Abbildung 3: *Visualisierung der Blickrichtung (1), der Handerkennung (2),der Greifbewegung(3)*

Damit die Blickrichtung sowie die Handerkennung für den Benutzer erkenntlich wird wurde ein Feedback-

[2]https://library.vuforia.com/articles/Training/Image-Target-Guide

System realisiert, dass in Abbildung 3 dargestellt ist. Das erste Symbol visualisiert die Blickrichtung, wobei diese zu jedem Zeitpunkt sichtbar ist oder durch eines der folgenden Symbole ersetzt wird. Mit dem zweiten Symbol wird die Handerkennung visualisiert. Dabei wird die Visualisierung der Blickrichtung durch das zweite Symbol ersetzt. Das dritte Symbol zeigt die Registrierung einer Greifbewegung, wobei das vorhergehende Symbol ersetzt wird.

Wurde ein greifbares Objekt anvisiert und eine Greifbewegung durchgeführt wird das auditiv zurückgemeldet und kann anschließend dreidimensional im Raum bewegt werden. Im Rahmen der Arbeit haben wir eine Freihandplatzierung für die Ports, allen Gelenken und der Roboterbasis ermöglicht. Damit soll eine flexible Platzierung und benutzerspezifische Konfigurationsanpassung geschaffen werden. Während einer manuellen Anpassung der Roboterkonfiguration wird die dritte Einschränkung mit der Berechnung der des größten Abstands zwischen dem Gelenk G2 des mittleren und G3 der äußeren Arme außer kraft gesetzt. Der Ablauf der Interaktion wird in Abbildung 4 und die AR-Sicht durch das HMD in Abbildung 5 dargestellt.

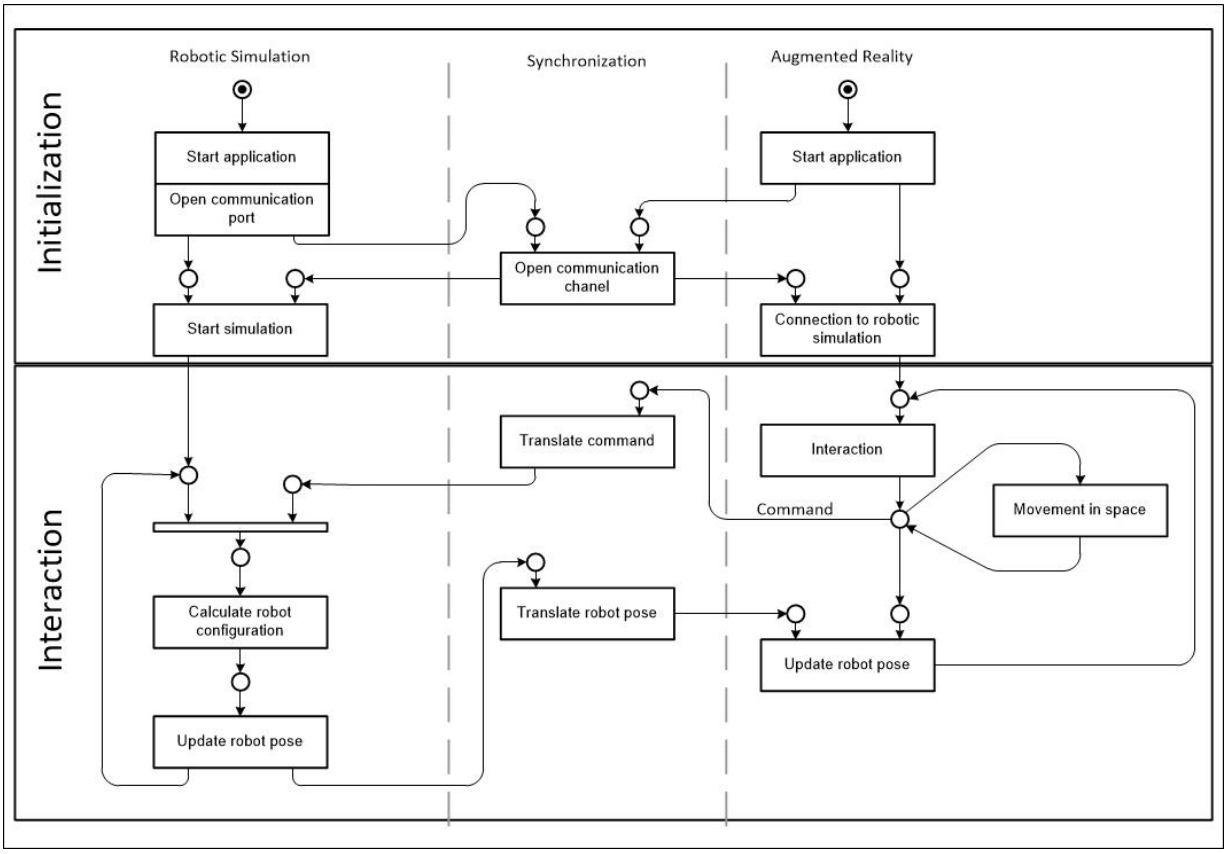


Abbildung 4: *Ablaufdiagramm des Gesamtsystems*

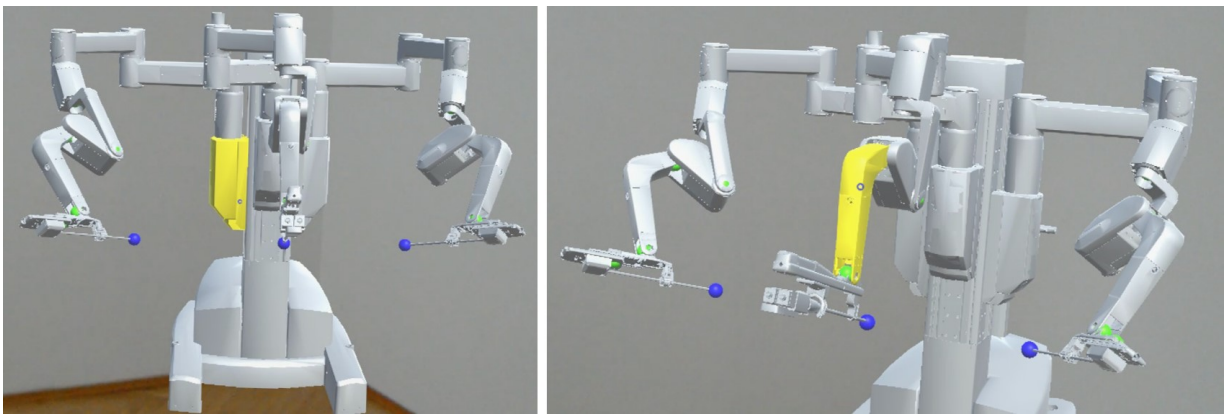


Abbildung 5: *AR-Sicht durch die Microsoft HoloLens auf das virtuelle Planungsmodell*

3 Ergebnisse

Das Ergebnis dieser Arbeit ist eine prototypische Planungsumgebung, in der die präoperative Positionierung der Roboterarme unter Vorgabe der Port-Position durchgeführt und trainiert werden kann. Dabei wurde eine Möglichkeit zur Ergebnisübertragung geschaffen, die AR verwendet um die errechnete Roboterkonfiguration in einer realen Umgebung immersiv zu visualisieren. Das Mapping zwischen virtuellen Objekt und realer Umgebung kann durch das integrierte Marker-Tracking einfach angewendet werden, wobei jederzeit manuelle Verbesserungen vorgenommen werden können.

Des weiteren wurde ein Ansatz zur Konfigurationsberechnung vorgestellt, der in der Praxis ähnliche Anwendung findet. Dabei wird anhand von Markierungen auf den Gelenken G3 der äußeren Arme versucht diese möglichst weit nach außen innerhalb des markierten Bereichs zu bewegen, während die Port-Position eingehalten wird. Darauf aufbauend wurde die vorgestellte Konfigurationsberechnung entwickelt, die diese manuelle Verfahren automatisieren kann, wobei manuelle Modifizierungen durch die Erfahrung des Chirurgen und somit benutzerspezifische Präferenzen einbezogen werden können.

Um die virtuelle Konfiguration zu modifizieren, wurde ein intuitives Interaktionskonzept benötigt, das durch Sprachbefehle, Gestenbefehle, visuelle und auditive Rückmeldungen realisiert wird. Wobei die Verwendung von Blick- und Handerkennung der Art und Weise ähnelt, wie wir mit alltäglichen Dingen in der realen Welt interagieren, wie etwas betrachten und danach greifen. Zusätzlich helfen die Feedback-Methoden dem Benutzer, die Aktionen auf Hologrammen mit dem zu verbinden, was man sehen und hören kann.

4 Diskussion

Die vorgestellte präoperative Planungsumgebung befindet sich noch in einem prototypischen Stadium, indem noch eine Evaluierung der leichteren Ergebnisübertragung von der virtuellen Planung in die reale Umgebung durchgeführt werden muss. Außerdem ist die aktuelle Konfigurationsoptimierung noch nicht auf dem Stand der Technik hinsichtlich anderer Optimierungsverfahren. Nichtsdestotrotz ist das Potential durch die Verwendung einer Visualisierung in AR deutlich erkennbar. Werden Planungsumgebungen, die ihre Ergebnisse in virtueller Realität darstellen, mit der in dieser Arbeit vorgestellten Ergebnisvisualisierung in AR verglichen, so liegt eine leichtere Überführung in die reale Umgebung mit unserem Ansatz nahe. Ebenso kann davon ausgegangen werden, dass die immersive Darstellung der Roboterkonfiguration einen Vorteil bei der Platzierung des Medizinroboter am Patienten bietet.

5 Zusammenfassung

Es wurde eine präoperative Planungsumgebung für roboterassistierte Eingriffe mithilfe von augmentierter Realität und Robotik-Simulation vorgestellt, die sich noch in einem prototypsichen Stadium befindet. Dabei wurde das Potential zur Optimierung in der Ergebnisübertragung aus der Planung in die reale Umgebung aufgezeigt. Als Ausblick wird die Weiterentwicklung des Optimierungsverfahrens bei der Konfigurationsermittlung sowie eine Evaluierung der getroffenen Annahmen zum Ziel gesetzt.

6 Referenzen

[1] D. B. Camarillo, T. M. Krummel, and J. K. Salisbury, JR, "Robotic technology in surgery: past, present, and future," American journal of surgery, vol. 188, no. 4A Suppl, pp. 2S–15S, 2004.

[2] J. W. Cannon, J. A. Stoll, S. D. Selha, P. E. Dupont, R. D. Howe, and D. F. Torchiana, "Port placement planning in robot-assisted coronary artery bypass," IEEE transactions on robotics and automation : a publication of the IEEE Robotics and Automation Society, vol. 19, no. 5, pp. 912–917, 2003.

[3] L. Adhami and E. Coste-Manirei, "Optimal planning for minimally invasive surgical robots," IEEE Transactions on Robotics and Automation, vol. 19, no. 5, pp. 854–863, 2003.

[4] L. Adhami, È. Coste-Manière, and J.-D. Boissonnat, "Planning and simulation of robotically assisted minimal invasive surgery," in Medical image computing and computer assisted intervention - MICCAI 2000 (S. L. Delp, ed.), vol. 1935 of Lecture Notes in Computer Science, pp. 624–633, Berlin u.a.: Springer, 2000.

[5] R. Konietschke, H. Weiss, T. Ortmaier, and G. Hirzinger, "A preoperative planning procedure for robotically assisted minimally invasive interventions," in CURAC 2004 3. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V., München, 8.-9. Okt. 2004, 2004.

[6] M. Hayashibe, N. Suzuki, M. Hashizume, Y. Kakeji, K. Konishi, S. Suzuki, and A. Hattori, "Preoperative planning system for surgical robotics setup with kinematics and haptics," The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 1, no. 2, pp. 76–85, 2005.

[7] A. L. Trejos, R. V. Patel, I. Ross, and B. Kiaii, "Optimizing port placement for robot-assisted minimally invasive cardiac surgery," The international journal of medical robotics + computer assisted surgery : MRCAS, vol. 3, no. 4, pp. 355–364, 2007.

[8] L. W. Sun, F. van Meer, J. Schmid, Y. Bailly, A. A. Thakre, and C. K. Yeung, "Advanced da vinci surgical system simulator for surgeon training and operation planning," The international journal of medical robotics + computer assisted surgery : MRCAS, vol. 3, no. 3, pp. 245–251, 2007.

[9] O. Weede, J. Wunscher, H. Kenngott, B.-P. Muller-Stich, and H. Worn, "Knowledge-based planning of port positions for minimally invasive surgery," in IEEE Conference on Cybernetics and Intelligent Systems (CIS), (Piscataway, NJ), pp. 12–17, IEEE, 2013.

[10] J. Hutzl and H. Worn, "Spatial probability distribution for port planning in minimal invasive robotic surgery (mirs)," in 2015 6th International Conference on Automation, Robotics and Applications (ICARA) (D. Bailey, ed.), (Piscataway, NJ), pp. 204–210, IEEE, 2015.

VR-basierte Interaktion mit 3D-Organmodellen zur Planung und Simulation laparoskopischer Eingriffe

F. Heinrich[1], S. Rohde[1], T. Huber[2], M. Paschold[2], W. Kneist[2], H. Lang[2], B. Preim[1], C. Hansen[1]

[1] *Fakultät für Informatik und Forschungscampus STIMULATE, Otto-von-Guericke-Universität Magdeburg*
[2] *Klinik für Allgemein-, Viszeral und Transplantationschirurgie, Universitätsmedizin der Johannes-Gutenberg-Universität Mainz*

Kontakt: florian.heinrich@ovgu.de

Abstract

Die Planung chirurgischer Eingriffe stellt einen entscheidenden Schritt bei der Behandlung von Erkrankungen wie Lebermetastasen dar. Zur Unterstützung von Chirurgen wurden daher zahlreiche Anwendungen zur Planung und abstrakten Simulation solcher Eingriffe entwickelt. Diese stoßen jedoch schnell an verschiedene Grenzen, wie unzureichende Interaktions- und Visualisierungsmöglichkeiten. In dieser Arbeit wird ein auf aktueller Virtual-Reality-Technologie basierender Softwareprototyp vorgestellt, der einen neuen Lösungsansatz in diesem Bereich darstellen soll und sich zunächst auf laparoskopische Leberresektionen spezialisiert. Durch den Einsatz geeigneter Eingabegeräte und der Möglichkeit patientenspezifische Datenmodelle in die virtuelle Umgebung einzufügen, liefert der Prototyp neuartige Interaktionsmöglichkeiten zur Eingriffsplanung sowie individuelle Trainingsmöglichkeiten. Eine erste qualitative Nutzerstudie zeigt die Nützlichkeit und Benutzbarkeit der Anwendung und bestätigt, dass der entwickelte Prototyp eine sinnvolle Basis für zukünftige Arbeiten bildet. Die Weiterentwicklung des hier vorgestellten Ansatzes könnte zukünftig dazu beitragen das Potenzial bisheriger Planungs- und Simulationsmethoden besser auszuschöpfen.

Keywords: Virtual Reality, Laparoskopie, Simulation

1 Problemstellung

Die Planung eines chirurgischen Eingriffs sowie die mentale Vorbereitung auf diesen stellen für Chirurgen eine wichtige Aufgabe dar. Durch bildgebende Verfahren kann die Lage und Position von im Körperinneren liegenden Strukturen genauer eingeschätzt werden. Zur Planung muss sich der ausführende Chirurg anhand der vorliegenden Bilder ein mentales 3D-Modell aufbauen. Moderne Operationsplanungssoftware ermöglicht es 3D-Organmodelle auf Basis tomographischer Bilder zu erstellen und in medizinische Dokumente einzubetten, z.B. in PDF-Dateien [1]. Weitere Anwendungen ermöglichen die Interaktion und chirurgische Planung mittels solcher Modelle [2]. In der Regel werden diese Daten jedoch auf einem zweidimensionalen Monitor betrachtet, wodurch Tiefeninformationen verloren gehen. Ein weiteres Verfahren wird durch das physische Drucken eines 3D-Modells dargestellt. Je nachdem, welche Komponenten zum Drucken verwendet werden, können dabei haptische Begebenheiten realistisch wiedergegeben werden, was für Simulations- und Trainingszwecke vorteilhaft ist. Jedoch müssen dabei für jeden Patienten neue Modelle gedruckt werden, um alle spezifischen Merkmale berücksichtigen zu können, was zurzeit kosten- und zeitintensiv ist [3]. Aus diesem Grund werden gedruckte 3D-Modelle zum aktuellen Zeitpunkt nur in besonderen Fällen verwendet.

Neben der Planung eines operativen Eingriffes sind vor allem eine vorangegangene chirurgische Ausbildung und zahlreiche Trainingseinheiten notwendig, um das Eingriffsrisiko zu minimieren. Um eine ausreichende Vorbereitung zu ermöglichen, werden daher verschiedenste Simulatoren eingesetzt, an denen angehende Chirurgen innerhalb virtueller Realitäten bestimmte Operationstechniken und psychomotorische Fähigkeiten trainieren können [4]. Chirurgen können dabei unterschiedliche laparoskopische Trainingsszenarien wählen, welche sich durch über die Zeit steigende Qualität und Komplexität auszeichnen. Die dabei verwendeten Szenarien sind jedoch im Regelfall nicht patientenspezifisch [5]. Die Möglichkeit einen Eingriff bereits im Vorfeld virtuell anhand der vorliegenden Daten zu simulieren, würde einen deutlichen Mehrwert bei der Vorbereitung auf einen Eingriff bieten. Zudem ist der Anteil virtueller Realität bei diesen Simulatoren stark begrenzt, da lediglich das Körperinnere simuliert und über einen konventionellen Monitor wiedergegeben wird. Dies begrenzt die wahrgenommene Immersion der Nutzer, welche sich stetig bewusst sind, lediglich eine Trainingssimulation zu erleben. So konnten Huber et al. bereits zeigen, dass die Verwendung eines aktuellen Head-Mounted-Displays (HMD) in Verbindung mit einem laparoskopischen Simulator ein hohes Maß an Immersion und Präsenz hervorrufen kann [6].

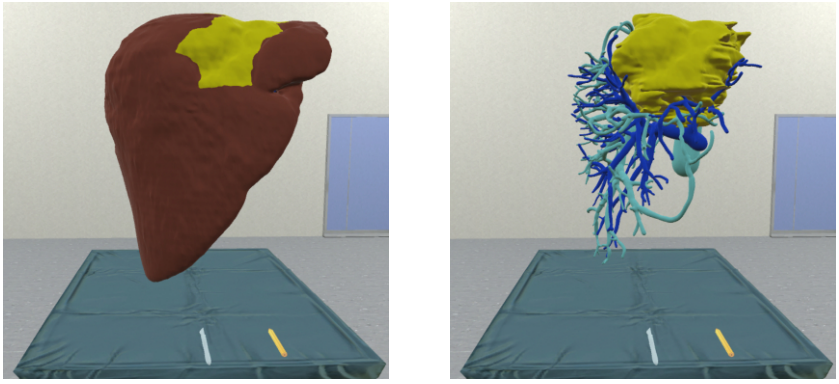
In dieser Arbeit wird deshalb eine Alternative zu bestehenden Planungs- und Simulationsmöglichkeiten vorgestellt, die es dem Nutzer unter Verwendung von aktuellen HMD-Technologien ermöglicht, patientenindividuelle

Daten zu explorieren und anhand dessen einen laparoskopischen Eingriff innerhalb hoch-immersiver Virtueller Realität (VR) zu planen und zu simulieren. Im Folgenden werden zunächst die technischen Komponenten des Systems beschrieben und anschließend auf eine erste qualitative Evaluierung des Prototyps eingegangen.

2 Material und Methoden

Um den in dieser Arbeit vorgestellten Prototyp umzusetzen, wurde die kommerziell verfügbare HTC Vive als VR-HMD-Lösung ausgewählt. Das relativ große Sichtfeld zusammen mit Head-Tracking und einer auf Stereoskopie basierenden Anzeigemethode ermöglichen eine hoch-immersive Erfahrung für den Nutzer. Neben der zur HTC Vive zugehörigen Controller finden dabei die für die Simulation laparoskopischer Eingriffe entwickelten Eingabegeräte Simball-Joysticks von G-Coder Systems Anwendung. Aufgrund ihres Designs lassen sich natürliche Interaktionsmöglichkeiten realisieren. Als Laufzeit- und Entwicklungsumgebung wurde die Computerspielegine Unity verwendet. Diese bietet unter anderem eine native Unterstützung der HTC Vive sowie mehrere Werkzeuge, die das Entwickeln einer VR-Anwendung unterstützen und beschleunigen können. So liefert das Virtual Reality Toolkit (VRTK) bereits vorgefertigte Lösungen zur grundlegenden Interaktion mit Objekten in VR. Das SOFA Framework [7], welches mittels zweier Schnittstellen in Unity integriert wurde, dient der physikalischen Simulation des Verhaltens von Weichgewebe. Unter Verwendung der Simulationsschleife des SOFA Frameworks kann so die visuelle Darstellung von Objekten innerhalb Unitys entsprechend berechnet und aktualisiert werden. Da die Anwendung zwei verschiedene Ziele, nämlich Planung und Simulation, verfolgt, fand eine Unterteilung in zwei verschiedene Modi statt.

2.1 Explorationsmodus – Eingriffsplanung



(a) Darstellung der Explorationsszene mit Lebermodell. Virtuelle Skalpell und Stift befinden sich auf dem OP-Tisch.

(b) Visualisierung der Lebergefäße und des Tumors mit ausgeblendeter Leberoberfläche.

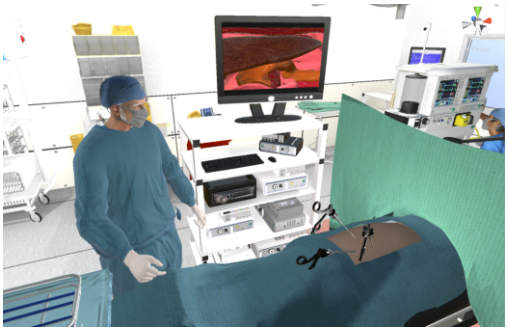
Abbildung 1: Explorationsmodus. Visualisierung eines Lebermodells mit den Komponenten: Leberoberfläche (braun), Portalvene (blau), Lebervene (türkis) und Tumor (gelb).

Durch die vergrößerte Darstellung eines interaktiven 3D-Organmodells basierend auf patientenspezifischen Daten kann eine Basis zur Planung chirurgischer Eingriffe geboten werden. Exemplarisch wurde zunächst ein Lebermodell basierend auf einer Fraunhofer MEVIS Leberoperationsplanung verwendet [8]. Mit Hilfe eines Vive-Controllers wird die Translation und Rotation des 3D-Objekts durch direkte Interaktion ermöglicht. Bei Kontakt des Controllers mit dem zuvor definierten interaktiven Modell wird dieses zunächst durch eine Farbänderung hervorgehoben, um dem Nutzer eine mögliche Aktion zu vermitteln. Durch Nutzen der Trigger-Taste kann das Objekt gegriffen, d.h. an den Controller gebunden, und durch einfache und natürliche Handbewegungen in der Szene verschoben und gedreht werden. Zusätzlich dazu wurde eine Rotation über das Trackpad des Vive-Controllers implementiert, um eine präzisere Kontrolle zu ermöglichen.

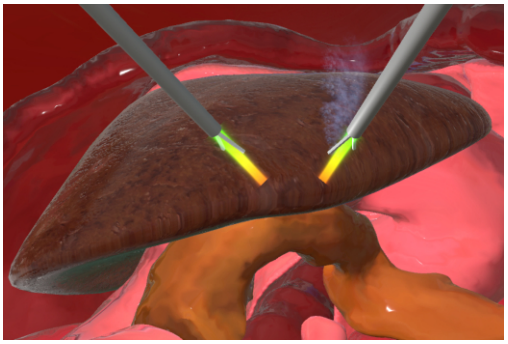
Entsprechend Abbildung 1(b) können zudem verschiedene segmentierte Organstrukturen ein- oder ausgeblendet werden, um den Planungsprozess zu unterstützen. Die Strukturen werden durch eine geeignete Farbgebung klar voneinander abgegrenzt. Dafür kann sowohl eine bestimmte Tastenkombination mit dem Vive Controller als auch ein speziell für diesen Modus entwickeltes User-Interface in Form eines VR-Menüs verwendet werden. Außerdem können durch das Interface weitere Funktionen wie das Laden oder Löschen von Objekten, das Wechseln der Szene oder auch das Beenden der Anwendung ausgewählt werden. Die Navigation durch das Menü erfolgt

mit Hilfe eines vom Controller ausgehenden virtuellen Laserpointers. Dieser kann per Tastendruck aktiviert und als Alternative zur direkten Interaktion durch den Controller eingesetzt werden. Um weitere Planungsmethoden bereitstellen zu können, wurden virtuelle Modelle eines Skalpells und eines Stifts in die Szene eingefügt. Dadurch soll dem Nutzer die Option gegeben werden, Resektionslinien einzuzeichnen sowie daraus entstehende Schnittresultate besser bewerten zu können.

2.2 Laparoskopiemodus – Eingriffssimulation



(a) Darstellung des virtuellen OP-Saals. Endoskopische Sicht in den Bauchraum kann über einen virtuellen Monitor betrachtet werden.



(b) Vergrößerung des virtuellen Monitors. Darstellung des Abdomens zur Eingriffssimulation.

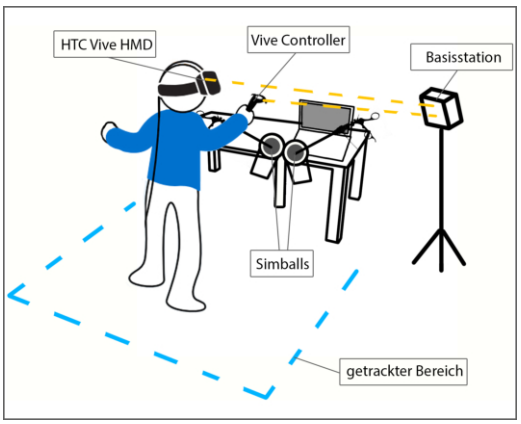
Abbildung 2: Laparoskopiemodus. Sicht eines Nutzers vor einem Patienten in einem virtuellen OP-Saal. Laparoskopische Werkzeuge sowie ein virtueller Monitor befinden sich vor ihm.

Bei der Darstellung des Laparoskopieszenarios sollte vor allem ein hoher Immersionsgrad erreicht werden. Dies konnte unter anderem durch Nutzung des von Huber et al. [9] verwendeten virtuellen Operationssaals und der Repräsentation der angebundenen Simball-Joysticks sowie derer Bewegungen innerhalb der virtuellen Realität erreicht werden. Zusätzlich konnte durch Verwendung des SOFA Frameworks [7] die physikalische Simulation von Gewebeverhalten beim Schneiden innerhalb von Unity umgesetzt werden. Im Inneren des virtuellen Patienten wurde ein anatomisch realistisches 3D-Abdomenmodell auf Basis des OpenHELP-Datensatzes [10] positioniert. Die dazugehörigen Leberdaten wurden durch ein anderes 3D-Lebermodell ersetzt, welches zur Verwendung mittels des SOFA Frameworks vorbereitet und verarbeitet wurde.

Da zunächst nur von der Nutzung durch eine Person ausgegangen wurde, wurde zudem ein drittes virtuelles Laparoskop als Kamera starr im Körperinneren platziert und das davon ausgehende Kamerabild auf den virtuellen Bildschirm des Operationssaals übertragen. Durch Verwendung der zwei Simball-Joysticks kann der Nutzer die virtuellen Instrumente frei bewegen. Über eine Art Laser ausgehend von den Endpunkten der Geräte wird eine verlängerte Darstellung dieser gegeben. Bei Kontakt mit dem Lebermodell kann dieses sowohl gegriffen als auch reseziert werden. Hierzu muss der zum Simball gehörende Griff vom Anwender zusammengedrückt werden. Das entsprechende Werkzeug wird über einen Schwellwert aktiviert, was wiederum durch eine Umfärbung des zuvor genannten Lasers verdeutlicht wird. Die Resektion entspricht hierbei dem Entfernen einzelner Dreiecke beziehungsweise dazugehöriger Tetraeder des Modellnetzes, sodass die daraus resultierenden Schnitte und Deformationen entsprechend der Modellauflösung mitunter unpräzise und grob wirken. Realistische Resultate könner daher aktuell nur bedingt erzielt werden.

2.3 Pilotstudie

Für eine erste Evaluierung der entwickelten Anwendung wurde eine qualitative Studie durchgeführt. Ziel war die Nützlichkeit und Benutzbarkeit des Systems zu bestimmen sowie Feedback zur weiteren Verbesserung der einzelnen Komponenten zu erhalten. Dazu wurden drei Chirurgen mit Erfahrung in VR und laparoskopischer Leberresektion gebeten, den Planungs- und Simulationsmodus der Anwendung zu benutzen und dabei die *Think-Aloud-Methode* anzuwenden . Der Prototyp wurde dabei zusammen mit den zuvor beschriebenen Hardware-Komponenten HTC Vive, Vive-Controller und Simball-Joysticks auf einem Rechner mit Intel Core i7-Prozessor einer NVIDIA Geforce GTX 1070 Grafikkarte und 16GB RAM Arbeitsspeicher getestet. Der Versuchsaufbau kann schematisch in Abbildung 3(a) betrachtet werden. Analog zum Prototyp wurde auch die Evaluierung in zwei Teilabschnitte gegliedert. Versuch 1 galt der Interaktion im Explorationsmodus während Versuch 2 der Interaktion im Laparoskopiemodus gewidmet war. Zuvor wurden für beide Teilversuche Aufgaben definiert, die



(a) Schematische Darstellung des Versuchsaufbaus.

(b) Nutzerinteraktion mit Simball-Joysticks im VR-Laparoskopiemodus.

Abbildung 3: Studienablauf. Proband befindet sich vor Simball-Joysticks und betrachtet VR-Inhalte über eine HTC-Vive. Basisstation wird zum tracken des Vive-Controllers und der Kopfposition des Nutzers verwendet.

von jedem Studienteilnehmer absolviert werden sollten.

Im Explorationsmodus stand dabei vor allem die Navigation durch das VR-Menü und die allgemeine Steuerung in VR mittels des Vive Controllers im Vordergrund. Nach Erklärung der grundlegende Steuerung mittels des HTC Vive Controllers, sollte zunächst das Interface des Menüs geöffnet und damit interagiert werden. Insbesondere sollten die Nutzer ein Objekt über das Menü laden und bestimmte Ebenen des dargestellten Objekts deaktivieren. Weitere Interaktionsaufgaben stellten die Translation und Rotation des geladenen Objekts sowie das Setzen von Markierungen mittels Marker-Tool dar.

Im Laparoskopiemodus sollten die Nutzer evaluieren, wie sicher mit den Simballs-Laparoskopiewerkzeugen in VR interagiert werden kann und ob die prototypische Anwendung zur Simulation bestimmter laparoskopischer Eingriffe geeignet ist. Dazu wurden die Anwender in der virtuellen Laparoskopieszene platziert und gebeten mit Hilfe der Instrumente ein sich im virtuellen Abdomen befindliches Lebermodell zu greifen und einen grob festgelegten Bereich des Lebermodells zu entfernen (Resektion der Lebersegmente 2/3). Abbildung 3(b) zeigt exemplarisch die Nutzung des Simulators im verwendeten Versuchsaufbau.

Entsprechend der *Think-Aloud-Methode* wurden die Nutzer gebeten alle Handlungen und Interaktionen mit dem System zu kommentieren und die jeweilige Situation zu beschreiben. Zudem wurden Fragen der Dimensionen Benutzbarkeit und Nützlichkeit des Fragebogens meCUE verwendet, um eine erste quantifizierbare Aussage bezüglich dieser Aspekte treffen zu können [11]. Dabei wurden die Probanden nach jedem Teilversuch gebeten, Antworten bezüglich der meCUE-Fragen auf einer siebenstufigen Likert-Skala von "lehne völlig ab"bis ßtimme völlig zußu geben. Anschließend wurden entsprechend des Fragebogen-Designs die Ergebnisse innerhalb der einzelnen Dimensionen und danach über die Probanden gemittelt.

3 Ergebnisse

Nach der Pilotstudie wurden die Kommentare der Probanden ausgewertet und zusammengefasst. Bezüglich des Explorationsmodus stellte vor allem die Möglichkeit, das Modell frei begehen zu können, eine neuartige Erfahrung dar und wurde als hilfreiche Ergänzung bei der Planung beurteilt. Dabei wurde die allgemeine Interaktion mit dem dargestellten 3D-Organmodell mittels des verwendeten Vive Controllers positiv wahrgenommen. Auch die Interaktion mittels Laserpointer wurde als positiv und intuitiv bewertet. Allerdings stellte sich der Umgang mit den virtuellen Werkzeugen, wie dem Marker-Tool, als vergleichsweise mühsam heraus. Außerdem ergab die Auswertung, dass das User-Interface durch in der Szene befindliche Objekte verdeckt werden konnte, wodurch die Interaktion mit dem Laserpointer erschwert wurde. Schließlich gaben die Probanden an, sich einige zusätzliche Funktionen zu wünschen. So sollte es die Möglichkeit geben, die Transparenz einzelner Ebenen des 3D-Modells modifizieren zu können. Außerdem wurde das Fehlen zusätzlicher Informationen wie dem Volumen der Leber oder einzelner Lebersegmente bemängelt, da auf diese Informationen für eine vollständige Planung nicht verzichtet werden kann.

Bei der Verwendung des Laparoskopiemodus bemerkten die Teilnehmer ein vergleichsweise unrealistisches physikalisches Verhalten des Lebermodells beim Greifen und Schneiden mit den genutzten Instrumenten. Zudem

erwies sich die Positionierung der virtuellen Trokare sowie der Umsetzung der Bewegung zwischen realem und virtuellem Instrument noch als unausgereift. So waren bestimmte Bereiche des Lebermodells nur schwer oder gar nicht erreichbar. Außerdem haben sich die Probanden weitere Interaktionsmöglichkeiten gewünscht. So ergab die Auswertung, dass die Steuerung der laparoskopischen Kamera im Abdomen sinnvoll wäre und, dass durch Anbindung eines Fußpedals zur Steuerung der Elektroden am Schneidewerkzeug ein höherer Grad an Realitätstreue erzielt werden könnte. Bezüglich des Realitätsgrades wurde zudem angemerkt, dass weitere, im OP häufig durchgeführte Arbeitsschritte – wie das Verschließen offener Bereiche mittels Clips – simuliert werden sollten. Die Darstellung zusätzlicher Strukturen wie bestimmter Ligamente am verwendeten Modell sei dafür ebenso relevant und fördere zudem die präzise Orientierung während des Eingriffs. Schließlich wurden fehlende Feedback-Elemente, wie die Darstellung der eigenen Hände sowie Haptik, bemerkt.

Zusammenfassend wurde der entwickelte Prototyp von den evaluierenden Chirurgen als „gute Basis“ für zukünftige Entwicklungen bewertet. Vor allem die Möglichkeit, patientenspezifische Daten verwenden zu können, wurde als Vorteil im Vergleich zu bekannten Simulatoren angesehen. Die Verwendung von VR zeigte speziell im Explorationsmodus einen Mehrwert und konnte mit den gegebenen Interaktionsmöglichkeiten überzeugen.

Die Auswertung der meCUE-Fragebögen ergab für den Explorationsmodus eine durchschnittliche Bewertung von $5,33 \pm 0,33$ hinsichtlich der Benutzbarkeit und $4,67 \pm 1,45$ hinsichtlich der Nützlichkeit. Für die Verwendung des Laparoskopiemodus wurde ein Score von rund $4,33 \pm 1,73$ bezüglich der Benutzbarkeit und $4,33 \pm 1,73$ bezüglich der Nützlichkeit erreicht.

4 Diskussion

Beide Komponenten der Anwendung wurden hinsichtlich ihrer Benutzbarkeit und Nützlichkeit positiv bewertet. Dies deutet auf Potenzial für weitere Entwicklungsschritte hin und deckt sich mit der qualitativen Bewertung der Anwendung als „gute Basis“. Aufgrund der kleinen Versuchsgruppe der Nutzerstudie müssen die Ergebnisse entsprechend vorsichtig bewertet werden. Dies trifft vor allem auf die quantitativen meCUE-Daten zu und spiegelt sich in vergleichsweise hohen Standardabweichungen von bis zu $1,73$ wieder. Deshalb ist eine Bewertung des Prototyps aufgrund dieser Ergebnisse nur begrenzt aussagekräftig. Dennoch lässt sich ein Trend erkennen.

Es konnten wertvolle Erkenntnisse aus den Ergebnissen des *Think-Aloud-Versuchs* gewonnen werden. Die direkte Interaktion mit virtuellen Objekten via Vive-Controller und Laserpointer wurde positiv bewertet während die Handhabung von virtuellen Werkzeugen durch Greifen und Führen mittels Controller als mühsam betrachtet wurde. Der Vive-Controller sollte daher stets direkt verwendet werden, um zusätzliche mentale Beanspruchungen durch die indirekte Übertragung der Controller-Bewegung auf die gehaltenen Werkzeuge zu vermeiden. Außerdem gaben die Probanden an, von der Darstellung weiterer Informationen zur Eingriffsplanung profitieren zu können. Ansätze zur Darstellung von Risokoanalysedaten wären hierbei von Nutzen [12], [13]. Außerdem sollten in zukünftigen Entwicklungen verschiedene Ansätze zum Messen von Distanzen und Volumina bedacht werden. Das IMHOTEP VR Framework stellt dabei bereits grundlegende Funktionen bereit und kann beliebig um eigene Ansätze erweitert werden [14]. Eine Integration des hier vorgestellten Prototyps innerhalb des Frameworks würde somit verschiedene Vorteile bieten. Bezüglich der laparoskopischen Simulation wurde vor allem mangelnde Realitätsnähe angemerkt. Da es sich hierbei um einen zentralen Aspekt der Simulation handelt, müssen dahingehend weitere Arbeitsschritte vorgenommen werden. Eine Kombination aus verbesserter Parametrisierung der Modelle sowie Anbindung weiterer Eingabegeräte zur Steuerung des Endoskops und der Elektroden kann dabei wertvolle Verbesserungen bieten.

Die Handhabung laparoskopischer Instrumente unterliegt dem Fulcrum-Effekt – die Endpunkte der Instrumente bewegen sich stets entgegengesetzt der eigenen Hände. Diese nicht-intuitive motorische Aufgabe wird aufgrund fehlenden Feedbacks auf Seiten des entwickelten Prototyps weiter erschwert. Deshalb sollten weitere Feedbackmodalitäten erschlossen werden. So kann sich die Visualisierung der eigenen Hände als hilfreich erweisen und die Immersivität der Anwendung steigern [15].

Die Verwendbarkeit patientenspezifischer Daten wurde von den Probanden als Vorteil herausgestellt. Der hier beschriebende Prototyp setzt jedoch aufbereitete 3D-Modelle voraus. Aufgrund aufwändiger Bildanalysen medizinischer Datensätze zur Generation dieser Modelle muss auch in anderen Bereichen weitere Forschungsarbeit betrieben werden.

Aktuell wurde der Prototyp auf den Anwendungsfall Leberresektion ausgelegt. In weiteren Arbeiten könnte der Simulator jedoch um weitere laparoskopische Eingriffe ergänzt werden. Dabei müsste die virtuelle OP-Szene auf die jeweiligen Anforderungen des Anwendungsfalls angepasst und segmentierte 3D-Modelle der betreffenden Organe verwendet werden. Durch eine Implementierung spezifischer Komplikationen könnten die Eingriffe somit in VR trainiert werden.

5 Zusammenfassung

In dieser Arbeit wurde ein Softwareprototyp vorgestellt, der eine neue Methode zur Planung und Simulation laparoskopischer Leberresektionen mittels VR-Technologie und intuitiver Interaktionsmöglichkeiten bietet. Die Anwendung wurde in einer Nutzerstudie evaluiert, wodurch Stärken und Verbesserungspotenziale aufgezeigt werden konnten. Besonders die Exploration patientenspezifischer Datenmodelle (Abschnitt 2.1) sowie einhergehende Interaktionsmöglichkeiten wurden von den Nutzern sehr positiv aufgenommen. Die virtuelle Eingriffsplanung (Abschnitt 2.2) könne jedoch in Zukunft von einer Visualisierung für den Chirurgen wichtiger Informationen profitieren. Die Eingriffssimulation wurde insgesamt eher schlechter als die Datenexploration bewertet, erhielt aber dennoch Zuspruch als sinnvolle Grundlage für weitere Entwicklungen in diesem Bereich. Vor allem die Schnittsimulation erwies sich als zu unpräzise und unrealistisch. Zudem stelle die Anbindung weiterer Eingabegeräte eine sinnvolle Möglichkeit dar, Eingriffe realitätsnäher zu simulieren und kollaborative Tätigkeiten zu trainieren. Insgesamt bietet der entwickelte Prototyp großes Potenzial als Grundlage für zukünftige Arbeiten die Möglichkeiten bisheriger Planungs- und Simulationsmethoden besser auszuschöpfen und zeigt auf, dass hochimmersive VR eine sinnvolle Ergänzung in der Medizin darstellen kann.

Danksagung Die vorliegende Arbeit wurde von der EU und dem Land Sachsen-Anhalt im Rahmen der Initiative „Sachsen-Anhalt WISSENSCHAFT Schwerpunkte“ gefördert (Förderkennzeichen ZS/2016/04/78123).

6 Referenzen

[1] A. Newe, L. Becker und A. Schenk, “Application and evaluation of interactive 3D PDF for presenting and sharing planning results for liver surgery in clinical routine”, *PLoS One*, Jg. 9, Nr. 12, e115697, 2014.

[2] S. Birr, J. Mönch, D. Sommerfeld et al., “The LiverAnatomyExplorer: A WebGL-Based Surgical Teaching Tool”, *IEEE Computer Graphics and Applications*, Jg. 33, Nr. 5, S. 48–58, 2013.

[3] J. T. Lambrecht, D. Berndt, R. Schumacher et al., “Generation of three-dimensional prototype models based on cone beam computed tomography”, *International journal of computer assisted radiology and surgery*, Jg. 4, Nr. 2, S. 175–180, 2009.

[4] R. Aggarwal, T. P. Grantcharov, J. R. Eriksen et al., “An evidence-based virtual reality training program for novice laparoscopic surgeons.”, *Annals of surgery*, Jg. 244, S. 310–314, 2 2006.

[5] T. Huber, A. Kirschniak und J. Johannink, “Survey of Training in Laparoscopic Skills in Germany”, *Zentralblatt fur Chirurgie*, Jg. 142, S. 67–71, 1 2017.

[6] T. Huber, M. Paschold, C. Hansen et al., “New dimensions in surgical training: immersive virtual reality laparoscopic simulation exhilarates surgical staff.”, *Surgical endoscopy*, Jg. 31, S. 4472–4477, 11 2017.

[7] J. Allard, S. Cotin, F. Faure et al., “SOFA - an open source framework for medical simulation”, S. 7, 2007.

[8] A. Schenk, D. Haemmerich und T. Preusser, “Planning of image-guided interventions in the liver”, *IEEE pulse*, Jg. 2, Nr. 5, S. 48–55, 2011.

[9] T. Huber, T. Wunderling, M. Paschold et al., “Highly immersive virtual reality laparoscopy simulation: Development and future aspects”, *International Journal of Computer Assisted Radiology and Surgery*, Jg. 13, Nr. 2, S. 281–290, 2018.

[10] H. G. Kenngott, J. J. Wünscher, M. Wagner et al., “OpenHELP (heidelberg laparoscopy phantom): Development of an open-source surgical evaluation and training tool”, *Surgical Endoscopy*, Jg. 29, Nr. 11, S. 3338–3347, 2015.

[11] M. Minge und L. Riedel, “meCUE-Ein modularer Fragebogen zur Erfassung des Nutzungserlebens.”, in *Mensch & Computer*, 2013, S. 89–98.

[12] C. Hansen, S. Zidowitz, B. Preim et al., “Impact of model-based risk analysis for liver surgery planning.”, *International journal of computer assisted radiology and surgery*, Jg. 9, S. 473–480, 3 2014.

[13] C. Hansen, S. Zidowitz, F. Ritter et al., “Risk maps for liver surgery”, *International Journal of Computer Assisted Radiology and Surgery*, Jg. 8, Nr. 3, S. 419–428, 2013.

[14] M. Pfeiffer, H. Kenngott, A. Preukschas et al., “IMHOTEP: virtual reality framework for surgical applications.”, *International journal of computer assisted radiology and surgery*, Jg. 13, S. 741–748, 5 2018.

[15] T. Wunderling, T. Huber, M. Paschold et al., “Immersives Laparoskopie-Training mit Hand-Tracking im virtuellen Operationssaal”, *Proceedings of the Annual Meeting of the German Society of Computer- and Robot-Assisted Surgery*, S. 67–72, 2017.

Feasibility of Robotic Multiport Cochlear Implantation – Evaluation in an Ex-Vivo Model

Daniel Schneider[1], Igor Stenin[2], Juan Ansó[1], Jan Hermann[1], Fabian Müller[1], Gabriela Pereira Bom Braga[1], Stefan Weber[1], Lukas Anschütz[3], Thomas Klenzner[2]

[1] ARTORG Center for Biomedical Engineering Research, University of Bern, Switzerland
[2] Department of Otorhinolaryngology, University Hospital Düsseldorf, Germany
[3] Department of ENT, Head and Neck Surgery, Inselspital, University Hospital Bern, Switzerland

Kontakt: daniel.schneider@artorg.unibe.ch

Abstract:
Objective: Robotic cochlear implantation requires visualization and potentially manipulative assistance during the electrode-array insertion process. This study investigates the feasibility of cochlear implant electrode-array insertion through three robotically drilled middle ear access tunnels.
Methods: Three tunnels targeting the round window were planned and drilled in four human cadaveric ear specimens using a surgical robotic system. Two tunnels passed through the facial recess, whereas one tunnel was planned supra-meatal. The electrode-array was inserted into the cochlea on four heads through the superior facial recess tunnel. Endoscopic visualization of the insertion process was provided through the inferior facial recess tunnel. Manipulative assistance was provided through the suprameatal tunnel.
Results: All tunnels could successfully be planned and drilled. Successful insertion was achieved in all specimens in under 15 minutes.
Conclusions: Electrode-array insertion with a multi-port approach is feasible.

Keywords: Cochlear implantation, Robotic surgery, Multi-port surgery

1 Problem Statement

Robotic cochlear implantation includes cochlear access planning, middle ear access drilling, inner ear access milling and electrode-array insertion. Clinically tested robotic systems so far only provide functionality to plan and create a tunnel access to the middle ear targeting the cochlea. The subsequent access of the inner ear and electrode-array insertion are of major importance for atraumatic and hearing preserving cochlear implantation. Visualization and potentially manipulative assistance is required to assure correct electrode-array insertion. The mentioned robotic approaches to cochlear implantation therefore involve a manually created tympanomeatal flap providing access to the middle ear through the external auditory canal. The inner ear access and the insertion procedure are then assisted manually under microscopic visualization through the external auditory canal [1], [2]. This involves the manipulation of the ear drum and ossicular chain, therefore additional trauma and potentially related loss of residual hearing. To tackle this challenge, an approach to electrode-insertion based on multiple tunnels targeting the round window was proposed. The tunnels serve as ports for the electrode-array, visualization and manipulation. The planning multiple tunnels with adequate safety clearance to critical structures was found to be feasible also in consideration of anatomical variations [3]. Furthermore, the drilling of tunnels with the adequate accuracy was shown to be feasible [4]. The feasibility of inserting a cochlear implant electrode-array through a multi-port approach into the cochlea however was never investigated. In this article an ex-vivo study is presented investigating the feasibility of inserting a cochlear implant electrode-array into the cochlea using three robotically drilled tunnels: one for the electrode-array, the second for endoscopic visualization of the insertion procedure and the third for manipulative assistance during the insertion procedure.

2 Material und Methods

The electrode insertion procedure via a multi-port approach was assessed in four human cadaveric ear specimens. First, a c-shaped retroauricular incision was created and four fiducial screws were inserted. Subsequently, a computed tomography (CT) scan was acquired (resolution: 0.15 x 0.15 x 0.2 mm). Three tunnels targeting the round window were planned for each specimen with a minimum clearance of 0.2 mm to the facial nerve, the chorda tympani, the ossicles, and the external ear canal (Fig. 1A) using a software dedicated to otologic surgery planning (OTOPLAN™, CAScina-

tion). Two tunnels through the facial recess were planned (FR_{inf}, FR_{sup}, inferior and superior as in standard anatomical position) and the third tunnel was planned suprameatal (SM). Subsequently, the tunnels were drilled (Fig. 1B) using a surgical robot (HEARO®, CAScination). The bony overhang was milled away robotically through the FR_{sup} tunnel under endoscopic visualization (diameter: 1 mm, angle: 0 deg, semi-rigid fiber optic, Karl Storz) and permanent monitoring of the milling forces. The milling of the bony overhang was manually stopped by the surgeon as soon as the milling force dropped sharply. Subsequently, the endoscopic view was adjusted to obtain a good view on the round window region and in particular of the round window membrane. The endoscope was fixed with a custom passive endoscope holder to free the hands for manipulative assistance and handling of the electrode-array. A guide tube was inserted in the FR_{sup} tunnel. The guide tube was positioned in contact with the remaining bony overhang to prevent the electrode array from bearing away into air cells the mastoid or the middle ear cavity. Finally, the electrode-array was guided to the round window through the guide tube (Fig. 1C). Manipulative assistance was provided through the SM tunnel using a surgical curved pick. The electrode-array was manually pushed through the guide tube into the cochlea. Finally, the round window was sealed with fat through the SM tunnel. Time for the insertion was recorded. Recording was started with all instrumentation away from the surgical site and stopped when the round window was sealed. The insertion depth was measured by counting the number of electrodes inserted into the cochlea.

In a second experiment, after the bony overhang milling, each tunnel was analysed concerning the potential for visualization of and manipulative assistance during the insertion procedure. For this purpose, a set of 13 anatomical landmarks and regions was defined (round window membrane, round window niche, promontory, oval window, stapes, chorda tympani, mastoidal segment of facial nerve, sinus tympani, fustis, facial recess, incus, pyramidal eminence, tympanic segment of the facial nerve). The round window membrane was defined as indispensable for cochlear implantation. Therefore, each tunnel, was analysed concerning the visibility and reachability of the round window membrane. Furthermore, through each tunnel, the percentage of visible and reachable anatomical structures was assessed.

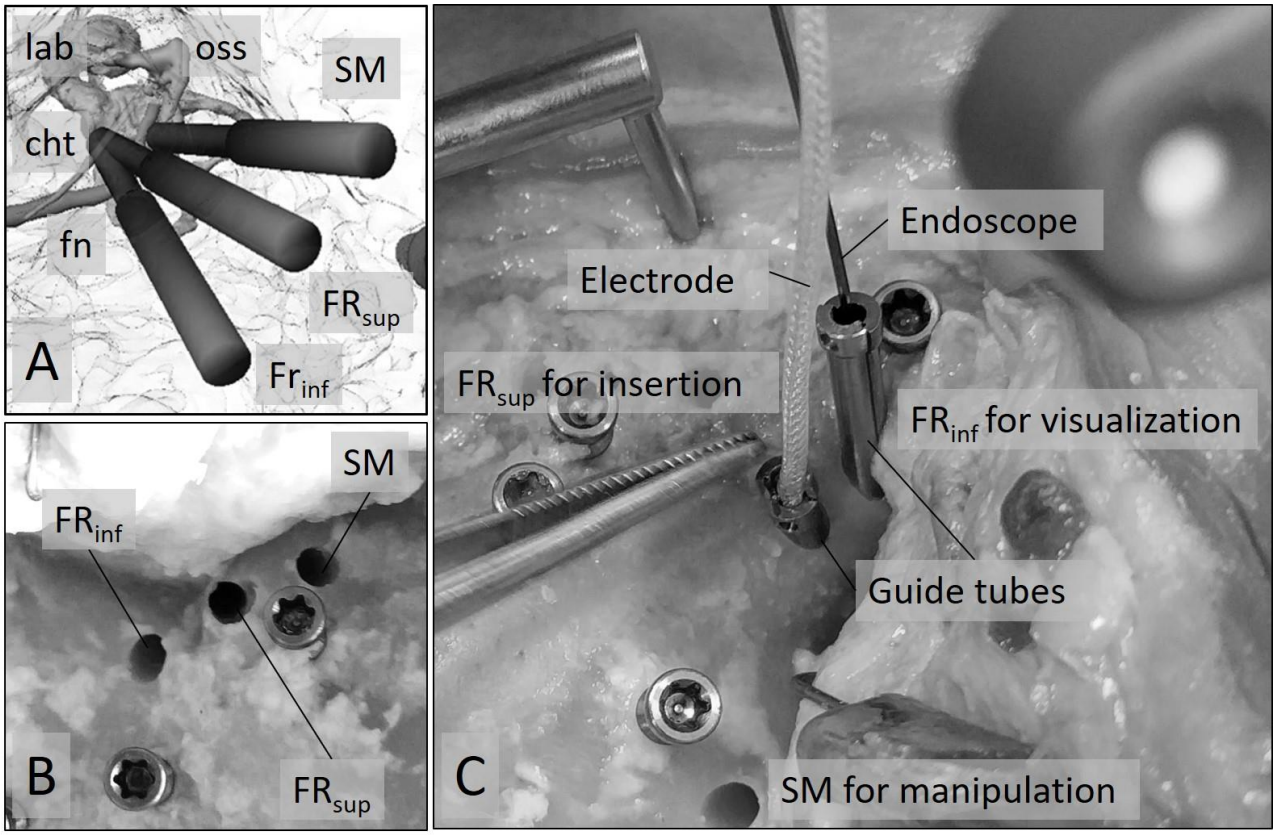


Fig. 1: Planning, drillings, and insertion procedure: Three tunnels were planned and drilled to the round window (A & B) with a minimum clearance of 0.2 mm to critical structures of the temporal bone. The electrode-array was guided to the round window through the FR_{sup} tunnel using a guide tube. Endoscopic visualization and manipulative assistance were enabled through the FR_{inf} and SM tunnel respectively (C).

3 Results

All three tunnels could successfully be planned and drilled in all four specimens. The minimum planned distance of a tunnel to a critical anatomical structure was 0.2 mm (FR_{inf} and FR_{sup} to chorda tympani in one of the four specimens). The electrode-array could fully be inserted (12/12 electrodes) in all specimens (Fig. 2). The time to insert the four electrode-arrays was 9.7 +/- 2.4 min. The endoscopic video enabled visualization of the round window region during the whole insertion procedure. Manipulative assistance of the insertion procedure through the SM tunnel was used only in one specimen.
The round window membrane was visible and reachable through the FR_{inf} and FR_{sup} tunnel all four specimens. Through the SM trajectory, the round window membrane was partially visible in two of four specimens and reachable in one of four specimens. The most anatomical landmarks could be visualized and reached through FR_{inf} (62.5%).

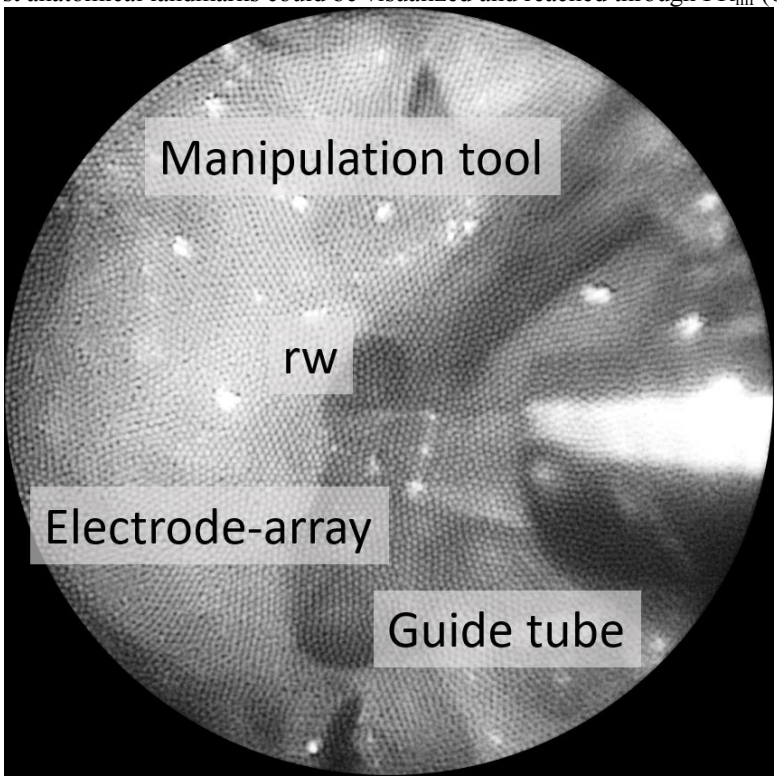


Fig. 2: Electrode-array insertion: Electrode-array inserted in the cochlea through the round window. Endoscopic visualization and manipulative assistance through the FR_{inf} and SM tunnel respectively.
rw: round window

4 Discussion

This study investigated the feasibility of inserting a cochlear implant electrode-array into the cochlea via a multi-port approach. Two robotically drilled tunnels joining in the facial recess served as ports for the electrode-array and the endoscope. A third robotically drilled tunnel served as port for a surgical instrument.
The electrode-array could be inserted fully in all specimens. The visualization with the endoscope was sufficient. Ideally, the guide tube restricts the electrode-array to go into the cochlea via the round window without manipulative assistance. Positioning the guide tube in contact with the promontorial bone and orienting the guide tube allowing to visualize the insertion procedure is of major importance for a successful insertion procedure. In one specimen, the electrode-array eluded to the middle ear cavity. Manipulative assistance and twisting of the electrode-array was necessary to finish the insertion. To preserve residual hearing during the insertion procedure, the jerky movements due to the twisting and manipulation are inadequate. Furthermore, assistance through the SM tunnel is a coordinative and motoric challenge for the surgeon due to the proximity of the ossicles to the surgical instrument. Also, the surgical instrument must be blindly guided through the epitympanum. To tackle these challenges, a transparent guide tube restricting any deviation of the electrode-array into the middle ear cavity while allowing to visualize the insertion procedure is currently being investigated. Furthermore, a construction in front of the endoscope to assist the insertion procedure allows to provide visualization and manipulation through one tunnel and is due to further work. A prerequisite to successful electrode-array insertion via a multi-port approach is robotic or manual bony overhang milling through one of the ports. Straight access to the

scala tympani is required to insert the electrode into the cochlea without the need for complex manipulation tasks with the surgical pick.
The anatomical landmarks visible and reachable through the different tunnels indicate the adequateness of a specific tunnel for a particular surgical application. The FR_{sup} and FR_{inf} trajectory both allowed the visualization and the manipulation of the anatomy in the round window region in all specimens and hence are suitable as ports to visualize and assist the insertion procedure. Depending on the patient's anatomy, the round window membrane is not visible through the SM tunnel even after milling of the bony overhang. Hence the tunnel is not adequate for visualization of the insertion procedure since electrode-array insertion into the scala tympani cannot be confirmed. However, the ossicles and the tympanic part of the facial nerve were visible through the SM tunnel in all four specimens indicating its adequateness for a potential minimally invasive access to the epitympanum.

5 Summary

This ex-vivo study (n=4) investigated the feasibility of a multi-port approach to insert a cochlear implant electrode-array into the cochlea. Three robotically drilled tunnels allowed to guide the electrode to the round window, visualizing the insertion procedure with an endoscope and assisting the procedure with a surgical curved pick. Insertion depth and time were used to quantify the feasibility of the multi-port approach. Furthermore, each tunnel was investigated concerning the potential to visualize and manipulate anatomical landmarks within the middle ear. The potential was quantified by the percentage of visible and reachable anatomical landmarks of a predefined set of anatomical landmarks.
The electrode-array could be fully inserted all specimens in under 15 minutes. While endoscopic visualization through the FR_{inf} tunnel enables adequate monitoring of the insertion procedure, manipulative assistance through the SM tunnel is cumbersome. The FR_{inf} and FR_{sup} tunnel allowed to reach and visualize the most anatomical landmarks respectively. Furthermore, both tunnels allowed to visualize and touch the round window membrane.

7 References

[1] S. Weber *et al.*, "Instrument flight to the inner ear," *Sci. Robot.*, vol. 2, no. 4, p. eaal4916, 2017.
[2] R. F. Labadie *et al.*, "Minimally invasive image-guided cochlear implantation surgery: First report of clinical implementation," *Laryngoscope*, vol. 124, no. 8, pp. 1915–1922, 2014.
[3] I. Stenin *et al.*, "Minimally invasive multiport surgery of the lateral skull base.," *Biomed Res. Int.*, vol. 2014, p. 379295, 2014.
[4] B. Bell *et al.*, "In Vitro Accuracy Evaluation of Image-Guided Robot System for Direct Cochlear Access.," *Otol. Neurotol.*, vol. 34, pp. 1284–1290, 2013.

Mastoid and ear model mimicking anatomical and physiological properties for development and training of robotic cochlear implantation

J. Hermann[1], F. Müller[1], D. Schneider[1], Y. Jegge[1], L. Müller[1], M. Caversaccio[1-2], S. Weber[1], J. Ansó[1]

[1]ARTORG Center for Biomedical Engineering Research, University of Bern
[2]Department of Head and Neck Surgery, University Hospital, Inselspital, Bern

Contact: jan.hermann@artorg.unibe.ch

Abstract

Mechanistic replications of the human anatomy are an imperative element of preclinical development and validation of medical devices. Such phantoms provide for reproducible and repetitive test benches to investigate a given device's performance parameters. Validation and verification of the elements of the robotic cochlear implantation approach requires phantoms that appropriately replicate anatomical, electrophysiological and mechanical properties. So far, different types of phantoms are used to investigate the various elements independent from each other. Hence, the objective of this work was to develop and validate an integrated phantom capable of emulating robotic drilling, functional electrical stimulation of the facial nerve, monitoring of the connected muscles and subsequently cochlear electrode insertion. Additionally, the phantom is integrated in a whole head replica to provide for a realistic intraoperative setup during testing activities.

Keywords: Technical model, ear anatomy, surgical training, facial nerve monitoring, robotic surgery

1 Problem

Cochlear implantation is an otologic microsurgical procedure for people with profound sensorineural hearing loss where an electronic device is implanted in the lateral skull base with an electrode array placed in the scala tympani of the cochlea. We have proposed a robotic treatment model (Robotic Cochlear Implantation) that can eventually lead to improved consistency of residual hearing preservation and better audiological outcome [1]. The robotic cochlear implantation model incorporates steps of computer assisted planning, robotic middle ear access (RMA), robotic inner ear access (RIA) and robotic electrode insertion (REI).

During the procedure, a tunnel is drilled from the mastoid surface through the facial recess towards the round window and the basal turn of the cochlea, passing through a narrow gap between the facial nerve (FN) and the chorda tympani (ChT) with submillimeter margin (typically, < 0.5 mm). Insertion is executed through the tunnel under visual inspection using a microscope.

During robotic passage of the facial nerve and the chorda tympani, nerve monitoring is used as a safety mechanism to prevent damage. Facial nerve activity is monitored passively throughout the complete procedure and is actively stimulated when the drill passes. Muscle reaction is assessed and classified through electrodes in the facial muscles (orbicularis oculi and oris) and classified as safe (> 0.4 mm) or potentially unsafe (< 0.1 mm) [2]. Such stimulation is executed at 5 positions axially spaced along the facial recess at predefined distances below 1 mm from the facial nerve center, as projected on the planned drill axis [3].

Validation and verification of the underlying technology utilizes phantoms for accuracy assessment [4], [5], animal models for investigation of electrophysiological properties [2], [6] and ultimately human specimen for general feasibility and workflow requirements [7], [8] as well as for cochlear electrode insertion [9]–[11]. Commercially available temporal bone models exist (e.g. PHACON GmbH, Leipzig, Germany) accurately mimicking anatomy and provide optical monitoring of the cochlea insertion process through built-in cameras. Resident training is being enhanced by computer-based surgery simulators, where the whole workflow of temporal bone surgeries can be executed on virtual 3d models [12], [13]. To date, no setup is available that combines all of the above elements in one technical replica and allow for holistic, realistic and subsequent testing of all elements of the robotic workflow.

The aim of this work was to create a technical model of human ear anatomy for robotic cochlear implantation surgery that contains all relevant anatomical structures (mastoid bone, facial nerve, chorda tympani, middle ear cavity, bony overhang and scala tympani of the cochlea). Furthermore, the model should mimic the similar

electrical characteristics of the mastoid bone and the embedded nerves and allow for visual observation of the surgery through conscious choice of material transparency, windows to the anatomy and built-in cameras. Such a model makes it possible to conduct investigations on the performance of: i) neuromonitoring electrodes at an early stage of development, ii) detection algorithms of a breakthrough of the round window membrane after milling away the bony overhang and, iii) electrode array buckling during insertion. In addition to that, demonstrations could be held for users and other technical personnel to better understand the above mentioned aspects of the procedure.

2 Material and Methods

An artificial mastoid box was created via additive manufacturing. The mastoid box contains fiducial screws that are used to register a preoperative surgical plan (computed tomography data) of the patient's (or phantom's) anatomical structures and a predefined RCI trajectory [14] to the physical model (Figure 1). The developed box is screwed into a holder on a modified whole head replica that includes a small-size camera directed at the cochlea such that the drilling process, the milling of the bony overhang and the electrode insertion can be observed.

The facial nerve and the chorda tympani are represented as tracks on a printed circuit board (PCB). Various shapes and distances can be designed to represent patient variability. Fitted into the head interior is a microcontroller that electronically simulates facial nerve activity and responds to facial nerve stimulation, which is used during robotic cochlear implantation for facial nerve monitoring [2]. The simulated muscle response is physically transmitted to connectors on the face of the whole head replica for easy connection of the facial muscle electrodes (Figure 2).

Drilling an access tunnel through the facial recess to the round window of the cochlea requires in some cases depending on the patient's anatomy the milling of a bony overhang over the round window ([10], [11]). The surrounding middle ear cavity, the bony overhang and the scala tympani of the cochlea are 3d-printed in two parts with a transparent material using micro-computed tomography data of human anatomy. In between the two parts a latex membrane is fitted to model the round window membrane.

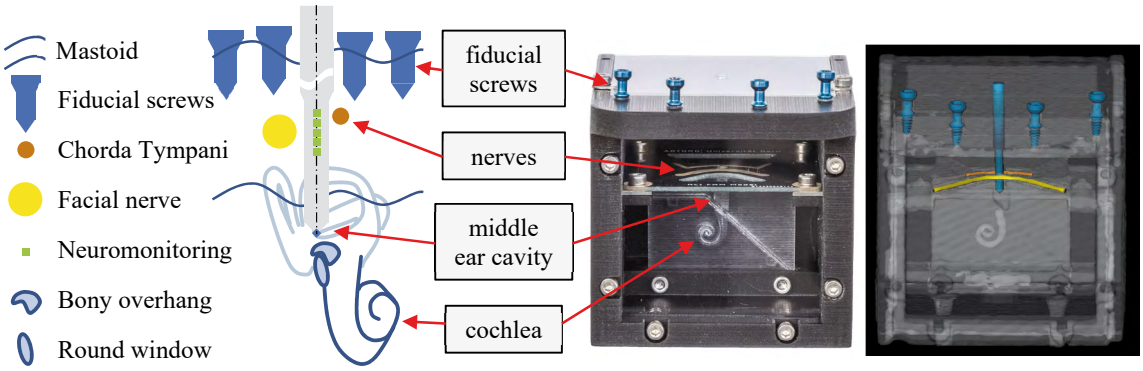


Figure 1: *On the left a schematic of the model-involved anatomy during RCI surgery. On the right a picture and a 3d surgery plan of the artificial mastoid box that contains the models of the represented anatomical structures.*

The usability of this technical model was assessed during an experimental robotic cochlear implant surgery experiment, which consisted of: i) drilling of the access tunnel, ii) neuromonitoring along the facial recess (near the facial nerve), iii) milling of the bony overhang, iv) round window membrane perforation and v) manual electrode array insertion. The model is comprised of the following:

- The **Mastoid Bone** is created by filling the artificial mastoid box from the bottom with a solidifying material. For this purpose the top is closed with a lid that can be removed before use. For electrical stimulation and demonstrations a gelatin, water and salt mixture is chosen because of its similar electrical properties to human tissue and its transparency [15]. For milling purposes a harder material such as artificial bone substrate could be used, preferably recreating the cancellous structure as well. For the purpose of testing the neuromonitoring system the material used as mastoid model needs to have similar electrical properties as living mastoid bone, which likely is not the case with a simple gelatin-based model. To identify realistic electrical properties of the gelatin mixture, electrical impedance spectroscopy (EIS) was conducted and compared to measured electrical properties of mastoid bone from in vivo trials (sheep model) [16].
- **Facial Nerve and Chorda Tympani Model** An electronics circuit model of the nerve-to-muscle interface was developed to emulate electromyogram responses based on stimulating intensity, integrated via a capacitive facial nerve model. The EMG model is composed of a signal conditioning amplifier, a microcontroller and a digital to analog converter, which translate electrical impulses from the stimulating

probe to a predefined (shape) electromyogram (EMG) response. Monophasic impulses (0.2 mA < intensity < 2 mA, 250 μs) are converted into an idealized compound muscle potential (sinewave period, 50 μV < amplitude < 8 mV, 10 ms). To enable the detection of a stimulation event based on the input impulse, an integration (or convolution) operation is carried out in the microcontroller. If the integrated signal is above a predefined minimum threshold (membrane potential, V_m = 20 mV), an EMG signal response is generated as a function of the measured input voltage. The intensity-to-EMG function follows a linear model below a predefined level of saturation (V_{sat} = 2 mV), simulating a scenario where the electrical field from the stimulation source covers the entire nerve's cross section, i.e. all nerve bundles and its internal axons are stimulated. Under this circumstance, the summation of action potentials in the innervated muscle fibers (EMG signal) stays equal with increasing stimulation intensity. The nerve-to-muscle model will be calibrated to enable discrimination of safety clearances from the RCI trajectory to the facial nerve (< 0.1 mm, > 0.4 mm) based on results from in vivo trials [2]. For monopolar stimulation the stimulating probe tip is the cathode and the far-field reference electrode is used as anode. Bipolar stimulation can be achieved by placing the anode on the electrode rings on the stimulation probe (Figure 4).

- **Middle Ear Cavity and Bony Overhang Model** The middle ear cavity is an air-filled cavity between the mastoid bone and the inner ear. A bony structure called the bony overhang located in front of the round window niche partially covers the entrance to the round window. The shape and size of the bony overhang can vary considerably between different subjects. In most subjects, the bony overhang has to be removed in order to expose the round window and get full access to the inner ear for a successful electrode insertion. The goal of robotic inner ear access is to remove the bony overhang while conserving the round window membrane. This prevents extensive trauma to the inner ear structures and maximizes residual hearing. In this model, an air filled rectangular cavity with a length of 3.5 mm and a width of 2.5 mm forms the middle ear cavity. The bony overhang is reconstructed from micro-CT data of a human temporal bone. For the purpose of testing the inner ear access the data set used originates from a subject with a dominant bony overhang. The anatomical models are created via additive manufacturing (Stratasys Objet Eden260VS, Ultimaker 3) using materials from *Vero* product line.
- **Inner Ear Model** The relevant structure of the inner ear for the purpose of testing tympanic electrode insertion is the scala tympani, whose shape is reconstructed from micro-CT data. After removal of the bony overhang and the round window membrane the scala tympani is accessible through the round window, which varies in shape and size and is seldom round. The round window used in this model is ovoid and nonplanar with a maximum diameter of 1.4 mm and an area of 1.5 mm^2. During the milling process the inner ear structures are protected from entering particles by the round window membrane, which in the model is formed by a latex foil with a thickness of 60 µm. The membrane separates the anatomical model of the middle ear and inner ear in two parts.

3 Results

The model for RCI surgery (Figure 2) was manufactured via additive manufacturing and manual handcraft. A verification drilling in the technical head showed its usability. Despite its elasticity, the gelatin mixture as mastoid bone could be drilled without the drilling process impacting its overall structure and visibility. However, the forces during drilling in the direction of the drill axis were below 1 N, where in patients the forces can exceed 8 N.

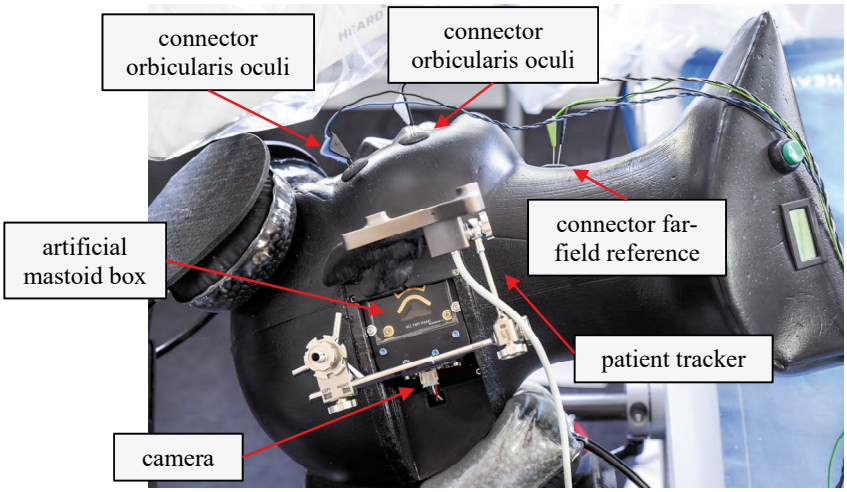


Figure 2: *The technical model of the human ear anatomy*

Verification of Electrical Characteristics of the Mastoid Bone Model

Electrical impedance spectroscopy was measured 39 times in: a) gelatin and b) saline solution. The same number of measurements was taken from existing *in vivo* trial data. Figure 3 shows the mean of the measured EIS data for each material. The standard deviation of the measurement at each frequency is depicted for each of the curves.

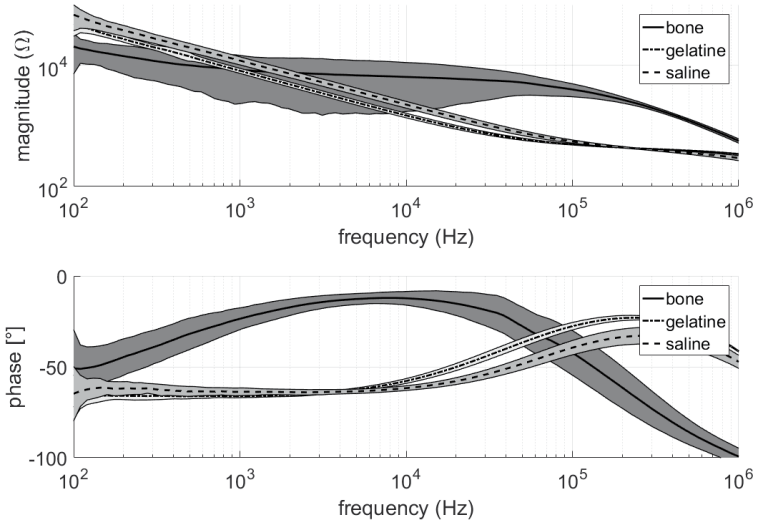


Figure 3: *Electrical impedance spectroscopy of the measured gelatine model as compared to the saline solution and to bone measurement from existing in vivo trials. The lines show the mean of all 39 measurements in each material, and the area depicts the standard deviation of the measurements.*

Facial nerve-to-muscle Model

The stimulation of the facial nerve-to-muscle model proved successful and a compound muscle action potential was measured in the connectors modelling the facial muscles. With the monopolar configuration, where the stimulation return electrode is situated far from the stimulating tip (Figure 4), an EMG response was recorded at a stimulating intensity of 2 mA, indicating sufficient clearance to the facial nerve. The bipolar configuration B3 generated a similar response to the monopolar, whereas B1 and B2 failed to produce a response, suggesting similar functional electrical field pattern of the configuration B3 and the monopolar.

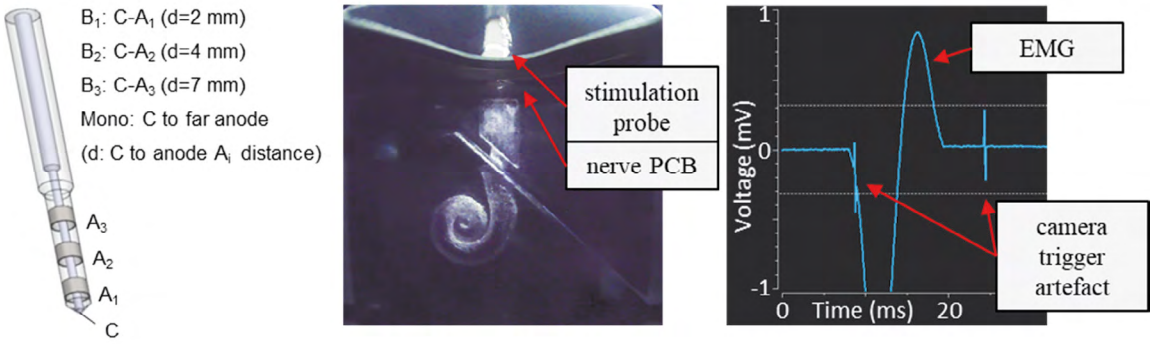


Figure 4: On the left the stimulating electrodes of the probe depicting each of the four cathode-to-anode combinations ($B_{1...3}$) and monopolar configuration. In the middle a picture from the built-in camera taken during facial nerve stimulation. On the right the simulated compound muscle response can be seen that was generated by the facial nerve-to-muscle model electronics at the detection of stimulation.

Inner Ear Access and Insertion

The bony overhang was milled away robotically with a 1 mm diameter diamond-coated milling bit. Guided by an insertion tube, full electrode array insertion could be performed in comparable time to surgery on real patients (< 5 min). From the built-in camera images the middle and inner ear structures could be visualized during inner ear access and electrode insertion (Figure 5).

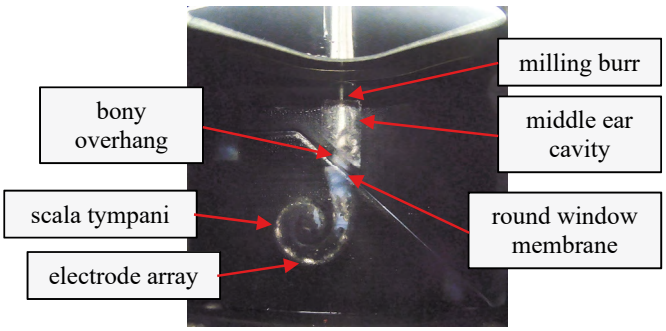


Figure 5: *A picture from the built-in camera taken after the electrode array insertion. Parts of the perforated round window membrane (latex in this model) can be seen in the scala tympani, which should be avoided for minimal trauma.*

4 Discussion

In this work, a technical replica of the human head comprising a mastoid bone region and inner ear anatomy containing all the relevant structures for a robotic cochlear implantation surgery has been developed. The model allows the drilling of an access tunnel to the round window of the cochlea. Facial nerve monitoring can be used during the drilling as the facial nerve is physically and electrically emulated using built-in specifically designed electronics. A small-size camera and material transparency allow for observation of the robotic surgery.

The mastoid model can be created with a solidifying material. For transparent material properties a gelatin mixture was chosen to enable direct inspection, which is of value for further development and validation of stimulating electrodes and neuromonitoring protocols, while being suitable for demonstrations and training. The gelatin model enables forces acting on the drill bit along the tunnel trajectory of below 1 N, which is an order of magnitude lower than forces in human mastoid. Furthermore, the electrical properties of the gelatin mixture are closer to saline solution than to living bone. While the current state enables initial tests for assessment of neuromonitoring protocols during RCI, the properties of the gelatin-bone model should become closer to the anisotropic morphology and properties of the mastoid bone. In the future reinforced gelatin models with additive manufactured support scaffolds could increase the hardness while preserving electrically representative conductivity properties.

Visualization of the presented RCI phantom could be improved by installing a second camera showing the passage of the drill bit between the facial nerve and chorda tympani, which would allow for observation of the stimulation probe during facial nerve monitoring.

In the presented model a computed tomography scan is required to provide a surgical plan with the location of the relevant anatomical structures. This step, although representative of the RCI approach, may not be necessary in the future, which could minimize preparation time and effort. The artificial mastoid box could be milled with a computer numerical control (CNC) machine with sufficient accuracy out of an adequately radiolucent material (e.g. PEEK).

The facial nerve and the chorda tympani are reproduced as tracks on a PCB with size and shape similar to real anatomy, even though they are contained within the PCB plane. The facial nerve model capacitance is designed to be in the same range as the capacitance of human myelinated nerves. The response of the electronics circuit model of the nerve-to-muscle interface to stimulation was functioning in monopolar and all bipolar ($B_{1...3}$) configurations. Stimulation level detection and corresponding response amplitude will be calibrated based on results from in-vivo studies.

The air-filled middle ear cavity was recreated with a bony overhang via additive manufacturing. This allows for the fabrication of various shapes and sizes of bony overhangs, which can be used for training of bony overhang removal without inflicting damage to the round window membrane for minimal trauma. The material used in additive manufacturing will have to be validated in the future to represent bony overhang hardness accurately.

The scala tympani is represented in transparent material recreated from human micro-computed tomography data. Same as with the bony overhang, this enables surgeons and engineers to train insertion with various shapes and locations of the round window, as well as different insertion angles. Representative electrode insertion forces can be achieved by modifying printed model wall roughness and potentially by filling the scala tympani with a lubricating liquid. Visualization of the middle ear cavity and the scala tympani was successful with the embedded camera, but could benefit in clarity and details from integrated light sources.

5 Conclusion

The developed model provides an initial test bench of the lateral skull base specifically for robotic cochlear implantation. It contains electrically functioning models of the mastoid bone and the facial nerve-to-muscle interface and physical models of the chorda tympani, the middle ear cavity, the bony overhang over the round window and the scala tympani. Further development of electrical, mechanical and anatomical properties are currently under investigation to provide a closer representation of a human mastoid and inner ear.

References

[1] S. Weber *et al.*, "Instrument flight to the inner ear," *Sci. Robot.*, vol. 2, no. 4, p. eaal4916, 2017.

[2] J. Ansó *et al.*, "A Neuromonitoring Approach to Facial Nerve Preservation During Image-guided Robotic Cochlear Implantation," *Otol. Neurotol.*, vol. 37, no. 1, pp. 89–98, 2016.

[3] J. Ansó *et al.*, "Neuromonitoring during Robotic Cochlear Implantation: Initial Clinical Experience," *Ann. Biomed. Eng.*, no. Special issue-Medical Robotics, p. Submitted, 2018.

[4] B. Bell *et al.*, "In Vitro Accuracy Evaluation of Image-Guided Robot System for Direct Cochlear Access.," *Otol. Neurotol.*, vol. 34, pp. 1284–1290, Aug. 2013.

[5] N. Gerber *et al.*, "High Accuracy Patient-to-Image Registration for the Facilitation of Image Guided Robotic Microsurgery on the Head.," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 4, pp. 960–8, Jan. 2013.

[6] A. Feldmann *et al.*, "Temperature Prediction Model for Bone Drilling Based on Density Distribution and In Vivo Experiments for Minimally Invasive Robotic Cochlear Implantation.," *Ann. Biomed. Eng.*, vol. 44, no. 5, pp. 1576–86, May 2016.

[7] B. Bell *et al.*, "A self-developed and constructed robot for minimally invasive cochlear implantation," *Acta Otolaryngol.*, vol. 132, no. 4, pp. 355–360, Apr. 2012.

[8] T. M. Williamson *et al.*, "Estimation of Tool Pose Based on Force–Density Correlation During Robotic Drilling," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 4, pp. 969–976, Apr. 2013.

[9] W. Wimmer *et al.*, "Cone beam and micro-computed tomography validation of manual array insertion for minimally invasive cochlear implantation.," *Audiol. Neurootol.*, vol. 19, no. 1, pp. 22–30, Jan. 2014.

[10] F. Venail *et al.*, "Manual Electrode Array Insertion Through a Robot-Assisted Minimal Invasive Cochleostomy: Feasibility and Comparison of Two Different Electrode Array Subtypes.," *Otol. Neurotol.*, vol. 36, no. 6, pp. 1015–22, Jul. 2015.

[11] W. Wimmer *et al.*, "Semiautomatic cochleostomy target and insertion trajectory planning for minimally invasive cochlear implantation.," *Biomed Res. Int.*, vol. 2014, p. 596498, Jan. 2014.

[12] R. Nash, R. Sykes, A. Majithia, A. Arora, A. Singh, and S. Khemani, "Objective assessment of learning curves for the Voxel-Man TempoSurg temporal bone surgery computer simulator," *J. Laryngol. Otol.*, vol. 126, no. 07, pp. 663–669, Jul. 2012.

[13] D. Franz *et al.*, "Haptisches Lernen für Cochlea Implantationen Konzept – HaptiVisT Projekt," *CURAC Dtsch. Gesellschaft für Comput. und Robot. Chir.*, vol. 15th, pp. 40–45, 2016.

[14] N. Gerber, B. Bell, K. Gavaghan, C. Weisstanner, M. D. Caversaccio, and S. Weber, "Surgical planning tool for robotically assisted hearing aid implantation.," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 9, no. 1, pp. 11–20, Jan. 2014.

[15] C. Marchal, M. Nadi, A. J. Tosser, C. Roussey, and M. L. Gaulard, "Dielectric properties of gelatine phantoms used for simulations of biological tissues between 10 and 50 MHz.," *Int. J. Hyperthermia*, vol. 5, no. 6, pp. 725–32.

[16] Y. Jegge *et al.*, "Feasibility of Nerve Proximity Detection using Tissue-Impedance Spectroscopy during Robotic Cochlear Implantation," in *CARS:Computer Assisted Radiology and Surgery Proceedings of the 31th International Congress and Exhibition*, 2018.

Feasibility of Pediatric Robotic Cochlear Implantation in Phantoms

Gabriela Pereira Bom Braga[1], *Wilhelm Wimmer*[1], *Kate Gerber*[1], *Daniel Schneider*[1], *Stefan Weber*[1], *Marco Caversaccio*[1]

[1] *University of Bern, Center of Biomedical Engineering Research (ARTORG), Bern, Switzerland*

Contact: gabriela.braga@artorg.unibe.ch

Abstract

Objective: To demonstrate the feasibility of robotic cochlear implant surgery in phantoms by anatomical study and drilling experimentation, followed by electrode insertion, drilling-accuracy study and patient positioning study.

Study design: Pilot study.

Materials and Methods: Computed tomographic preoperative encrypted data of twelve pediatric patients (total of 24 sides) between eight months and six years of age, who underwent cochlear implant surgery were studied, eight were used to create modified 3D printed replicas of the mastoid region and three other full skull replicas were used for accuracy and patient positioning evaluation. A robotic mastoidectomy was carried out.

Results: From the original 24 cases, on 17 planning a safe trajectory to the cochlea was possible. Drilling was successfully performed on 8 cases and insertion on 2. On 7 cases (29.1%) planning the trajectory was not possible, on 4 that was due to image quality and on 3 due to the small size of the facial recess.

Conclusion: To validate the proposed procedure in pediatric population in a more realistic clinical setting, optimization of the current surgical workflow, further investigations about cochlear angle behavior, mastoid development and response to trauma relative to electrode positioning and electrode insertion, will be accessed in future studies.

With this work we have demonstrated the feasibility to perform a robotic middle ear access in phantom models of pediatric cases. More data needs to be derived to address relevant aspects including bone growth after surgery and effects on electrode array position and migration.

Comparison of Haptic and Auditory Feedback Methods for the Teleoperation of Concentric Tube Continuum Robots

S. Lilge[1], *D. Black*[2], *A. Nabavi*[3], *J. Burgner-Kahrs*[1]

[1] *Leibniz Universität Hannover, Laboratory for Continuum Robotics, Hanover, Germany*
[2] *University of Bremen: Jacobs University, Fraunhofer MEVIS, Medical Image Computing, Bremen, Germany*
[3] *International Neuroscience Institute, Hanover, Germany*

Contact: lilge@lkr.uni-hannover.de

Abstract

Due to their high dexterity and good miniturizability, concentric tube continuum robots are well suited for minimally invasive surgeries. In such scenarios, the manipulator is often directly teleoperated by a surgeon, while the feedback is limited to a camera feed from an endoscope which is guided to follow the robot's movements. To further assist the operating surgeon, additional feedback for the teleoperation in medical scenarios is investigated.

This work focuses on the comparison of haptic and auditory feedback, generated using the method of virtual fixtures, for the teleoperation of concentric tube continuum robots. For evaluation, a user study using a simulated medical application is conducted. Results show an increase in satisfaction and usability scores, as well as a decrease in task load and execution time when applying haptic feedback. The use of auditory feedback does not show quantitative benefits and received diverse ratings from the participants.

1 Problem

Continuum robots are a growing research area in the context of medical robotics. Many different designs exist for continuum robots. One widely investigated design is the concentric tube continuum robot (CTCR) [1]. This manipulator consists of precurved Nitinol tubes that are nested inside each other. The independent axial translation and rotation of each tube constitutes the actuation of the CTCR. Compared to conventional robots with rigid links and discrete joints, this robot is composed of a compliant backbone and offers a high dexterity at the end-effector. Due to the jointless structure and extrinsic actuation, CTCR are miniturizable down to a millimeter in diameter. Furthermore, with their ability to follow curvilinear paths and to navigate in confined spaces, CTCR are well suited for applications in the context of minimally invasive surgery [2].

During such a surgery, a manipulator can either be automatically controlled or directly teleoperated by a surgeon. When controlled by a surgeon, an input device that is physically decoupled from the robot manipulator is used to manipulate the end-effector. Thus, the feedback received by the teleoperating user is often limited to a camera feed of an endoscope that is guided along the manipulator's end-effector. This visual feedback is often not sufficient to accurately steer the robot and avoid collision with the sensitive patient anatomy. To further assist the surgeon during robot teleoperation, various feedback methods for medical and continuum robots have been proposed. For instance, virtual fixtures in the form of haptic feedback were applied to teleoperated CTCR navigation scenarios in [3]. A similar approach was implemented in [4] and evaluated the benefit of haptic feedback in a user study of a simulated medical application. Furthermore, auditory feedback, or so-called auditory displays, have been investigated in the context of surgeries and medical robotics to increase the surgeon's awareness of surrounding structures or path deviations. Such methods have been evaluated for different medical applications [5, 6, 7]. Results have shown that the accuracy could be improved while the cognitive work load could be lowered when using auditory displays as an additional feedback. A review of auditory displays in image-guided interventions can be found in [8]. So far, the mentioned feedback methods in the context of robot teleoperation have only been compared to visual feedback. Hence, the comparison between the different feedback methods themselves is still an open question which is tackled throughout this work.

We compare different types of feedback generated based on the approach of virtual fixtures. We consider haptic and auditory feedback in comparison with visual feedback only. The evaluation consists of a user study that involves the teleroperation of a CTCR within a simulated medical application. Thus, the main contribution of this paper lies within the outright comparison of haptic and auditory feedback in the context of CTCR teleoperation.

2 Material and Methods

In this section, we describe the different components of the implemented system. This includes a CTCR in a simulation environment that allows teleoperation within a medical application where CTCR can be suitable

candidates. Furthermore, the generation of the additional haptic as well as auditory feedback on the basis of virtual fixtures is explained. Finally, the user study to evaluate the different feedback methods is described.

2.1 Concentric Tube Continuum Robot and Teleoperation Algorithm

A CTCR offers two degrees of freedom at its end-effector per tube (translation and rotation). Thus, three tubes are needed to freely translate and rotate the manipulator in every direction. We define the configuration of a three-tube CTCR as $\boldsymbol{q} = [\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3]^T$, where α_j is the rotation and β_j the translation of each tube $j \in \{1, 2, 3\}$. In this work, the kinematics of the CTCR, describing the mapping between robot actuation $\boldsymbol{q}$ and end-effector pose $\boldsymbol{x}_{\text{robot}}$ are calculated based on the Cosserat rod theory [9].

To teleoperate the manipulator (process shown in Fig. 1), the desired end-effector velocity $\dot{\boldsymbol{x}}_{des}$ is calculated from the deviation $\boldsymbol{\varepsilon}$ between the current input device position $\boldsymbol{x}_{des}$ and end-effector position $\boldsymbol{x}_{des}$ using a proportional gain λ. Afterwards, the damped-least squares inverse of the manipulator Jacobian matrix $\boldsymbol{J}^+$ is used to calculate the needed joint velocities $\dot{\boldsymbol{q}}$ [10]. The absolute desired joint values $\boldsymbol{q}$ are obtained by integration and applied to the robot. Finally, the estimated end-effector position is updated using the kinematic model.

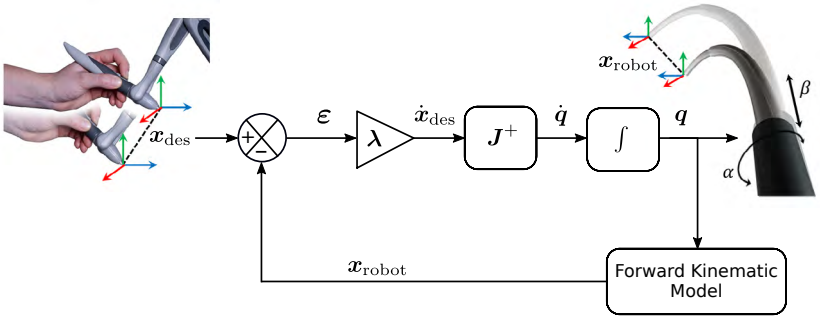


Figure 1: Overview of the teleoperation process to manipulate the CTCR end-effector using an input device.

The input device used in this work (Geomagic Touch, 3D Systems, Rock Hill, South Carolina, United States) is an impedance-controlled input device that allows manipulation of 6 degrees of freedom using a hand-held pen stylus. The current position of the pen tip is calculated using the device's geometry. Due to the different workspaces of the input device and of the robot, a so-called 'clutch-in' mechanism is applied by pressing a button on the stylus to move the end effector of the robot. By releasing this button, the stylus can be repositioned without end-effector movement. Furthermore, the used input device allows applying force feedback in the three translational degrees of freedom.

2.2 Simulation Environment

The implemented simulations environment allows the teleoperation of a CTCR in the context of different medical applications. Different anatomical data sets can be imported to serve as the work space for the manipulator. The imported data is then stored in kd-trees, which are space-partitioning data structures. These trees are used to query range and nearest-neighbor searches. This allows us to perform distance checks between the robot's end-effector as well as its backbone and the constraining patient anatomy online.

In addition, different navigation paths can be imported to conduct teleoperated navigation tasks. The simulation environment is then augmented with a visualization of the chosen path. Furthermore, different camera views are implemented for the teleoperating user. A first camera gives a brief overview of the scenery including the planned insertion path, while a second camera view simulated the camera feed of an endoscope guided along the CTCR end-effector (see Fig. 2, left). The simulation environment is perceived by the teleoperating user using a stereoscopic 3D monitor to enable depth perception.

2.3 Feedback Generation Based on Virtual Fixtures

The term virtual fixtures (VF) in the context of teleoperated robots was first introduced by [11]. In general, the method of virtual fixtures is divided into two categories: forbidden region (FR) and guidance (G). The method of forbidden region should prevent the penetration of a manipulator into specific regions, whereas the guidance method assists the teleoperating user to follow a predefined path. In this work, a suggested motion direction for the manipulator's end-effector is calculated utilizing both virtual fixtures methods, whether to avoid collisions with a forbidden region, or to follow a specific trajectory. The suggested motion direction is

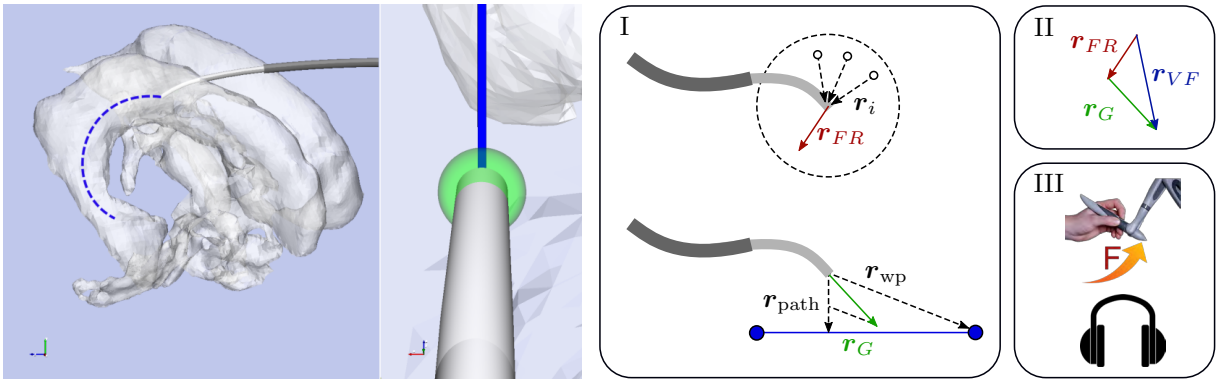


Figure 2: Left: Visualization of the simulation environment with two camera views; Right: Workflow of the feedback generation: (I) The suggested motion directions according to the guidance and forbidden region methods are calculated; (II) The vectors are combined to result in a total suggested motion direction for the methods of virtual fixtures; (III) The resulting vector is mapped to the haptic and auditory feedback methods.

calculated according to [4], which is briefly described in the following. Fig. 2 (right) shows an overview of the feedback generation based on virtual fixtures.

For the forbidden region method, a radius search around the robot's end-effector is performed in the work space to find n nearby data points of possible obstacles. A vector $\boldsymbol{r}_i$ pointing from a data point $\boldsymbol{p}_i$ to the robot's end-effector is calculated for each point with $i \in \{0, \ldots, n\}$. Afterwards, all vectors are weighted by the distance of corresponding data point to the end-effector, and a combined direction vector $\boldsymbol{r}_{\mathrm{FR}}$ is determined. This resulting vector $\boldsymbol{r}_{\mathrm{FR}}$ is then considered as the suggested motion direction of the end-effector to increase the distance between robot manipulator and obstacles in the work space.

For the guidance method, predefined paths consisting of multiple waypoints, which should be passed successively, and a linear interpolation between them are considered. These paths can be obtained by discretizing the planned insertion path. During the navigation of the manipulator, the vector between the next waypoint and the robot's end-effector $\boldsymbol{r}_{\mathrm{wp}}$, as well as the vector between the closest point on the current linear path segment and the robot's end-effector $\boldsymbol{r}_{\mathrm{path}}$ are calculated. Depending on the distance between the manipulator and the path to be followed, both vectors are weighted and combined into a resulting vector $\boldsymbol{r}_{\mathrm{G}}$. Eventually, both suggested motion directions $\boldsymbol{r}_{\mathrm{FR}}$ and $\boldsymbol{r}_{\mathrm{G}}$ are averaged in a final virtual fixtures vector $\boldsymbol{r}_{\mathrm{VF}}$. Note that the vector $\boldsymbol{r}_{\mathrm{VF}}$ is defined in the robot's end-effector frame, which is the same reference frame for the teleoperation process.

Using this suggested motion direction $\boldsymbol{r}_{\mathrm{VF}}$, feedback is generated to help the teleoperating user navigating the robot, either in the form of haptic feedback on the input device or an auditory feedback using headphones. For the generation of the haptic feedback, we use the suggested motion direction $\boldsymbol{r}_{\mathrm{VF}}$ and scale it to receive force vectors in the range of 0 to 1 N. Then, we directly apply every dimension of the vector to the translational degrees of freedom of the input device. For the auditory feedback, an auditory display is implemented that uses a set of recorded samples of vocal singing. The suggested motion direction $\boldsymbol{r}_{\mathrm{VF}}$ is then mapped on this 'vocal syllable choir'. The y-component of $\boldsymbol{r}_{\mathrm{VF}}$ is mapped to the pitch, syllable type and gender of the sung syllables. The used syllables included [o:] to indicate that the end-effector is *above* the suggested motion direction (German: *oben*), [u:] to indicate that the end-effector is *below* the suggested motion direction (German: *unten*) and a hummed [m] to indicate that the elevation of the end-effector does not need to be adjusted. The x-component of $\boldsymbol{r}_{\mathrm{VF}}$ is mapped similar, using a stereo auditory signal. If the sound is heard on the left channel, the robot end-effector is too far left compared to the suggested motion direction. The same is true for the right channel, respectively. If the azimuth of the manipulator does not need to be adjusted, the signal will be heard on both channels. Finally, the z-component of $\boldsymbol{r}_{\mathrm{VF}}$ is mapped to the amount of reverberation and inter-onset between the syllables. In addition to the mapping of $\boldsymbol{r}_{\mathrm{VF}}$, the auditory display also provides a notification sound when a waypoint has been reached by the CTCR end-effector.

2.4 User Study

To evaluate the effectiveness of the different feedback methods, a user study is conducted in a laboratory environment. In total, 13 participants are recruited. During the user study, the participants teleoperate a simulated CTCR within the simulation environment outlined in Sec. 2.2.

The participants are confronted with a navigation task in the context of minimal invasive surgery. As a medical application, the choroid plexus cauterization to treat hydrocephalus is chosen. The choroid plexus

produces watery cerebrospinal fluid in the ventricles of the brain. If a patient suffers from hydrocephalus, more cerebrospinal fluid is produced than absorbed, which leads to high intracranial pressures and compression of brain tissue. This can result in neurologic dysfunction and even death [12]. In [12], the insertion of CTCR into the brain ventricles was proposed to cauterize the choroid plexus to reduce the cerebrospinal fluid production as a treatment for hydrocephalus. The main advantage of using a CTCR in this application, compared to conventional endoscopes, is its ability to also follow the original insertion path, when retracting it. For the conducted user study, we choose a navigation task within a simulated model of the brain ventricles along insertion paths that lead to the proposed cauterization sites. The model is obtained from the 'National Institutes of Health Print Exchange' website [16]. This specific neurosurgical intervention serves as an example application for the use of CTCR manipulation in anatomical cavities like the lateral ventricles. Furthermore, it is perceivable for the participants without detailed explanations and allows them to focus on the navigation task.

First, each participant performs a training period to become accustomed to the robot teleoperation and the different feedback methods. Afterwards, the proposed navigation task is executed four times by every participant: Without any additional feedback (NF), with auditory feedback (AF), with haptic feedback (HF) and with a combination of haptic and auditory feedback (HAF). Note that visual feedback is present for every task execution and feedback modality. To counter learning effects, we use a within-subjects study design with an alternating order of task executions using the latin square method [13]. Furthermore, two different insertion paths were used during the study. For the first and third task execution, the first path is chosen, while for the second and fourth task execution, the second path is used.

In order to quantify the effectiveness of the different feedback methods, we record the task completion times and the average path deviations. In addition, after every finished task, each participant is asked to complete a questionnaire. The questionnaire contains questions to measure the satisfaction and usability according to the van der Laan's system acceptance scale [14], as well as the task load according to the NASA-TLX multidimensional assessment tool [15]. Furthermore, we ask five custom questions regarding the user's confidence during the task, the ease of use, their satisfaction with the time they needed, as well as the helpfulness and intuitiveness of the generated feedback. Finally, the participants rank the feedback methods.

3 Results

The quantitative results from the conducted user study are depicted in Fig. 3 (left and middle). It can be seen that the execution time is considerably lower with the use of haptic feedback and the combination of haptic and auditory feedback. When using auditory feedback only, the execution time stays in the same range compared to no feedback usage. The average path deviations did not alter using different feedback methods and mainly lie between 1 and 2 mm.

Fig. 3 (right) shows the results from the ranking of the feedback methods. The users rated the feedback methods from 'most favorite' to 'least favorite' which results in a score of 3 to 0 points for every method. It can be seen that on average, the users preferred the haptic feedback method closely followed by the combined feedback. On average, the participants liked the navigation without any feedback the least.

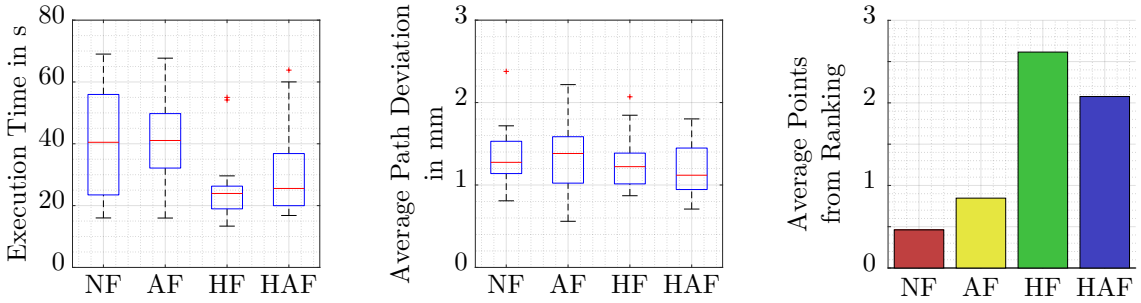


Figure 3: Box plots for execution times (left), average path deviations (middle), and average points from ranking (right) using different feedback methods.

The determined satisfaction, usability, and task load scores using the questionnaires are plotted in Fig. 4. Compared to the usage of no feedback, the usability is considerably higher when the participants were supported by additional haptic feedback. While the usability does not benefit from the usage of auditory feedback alone, the combined auditory and haptic feedback shows an improvement in the usability of the system. Similar results can be shown for the satisfaction of the generated feedback. The task load remains nearly the same for all feedback methods with a marginal better result for haptic feedback and the combined feedback.

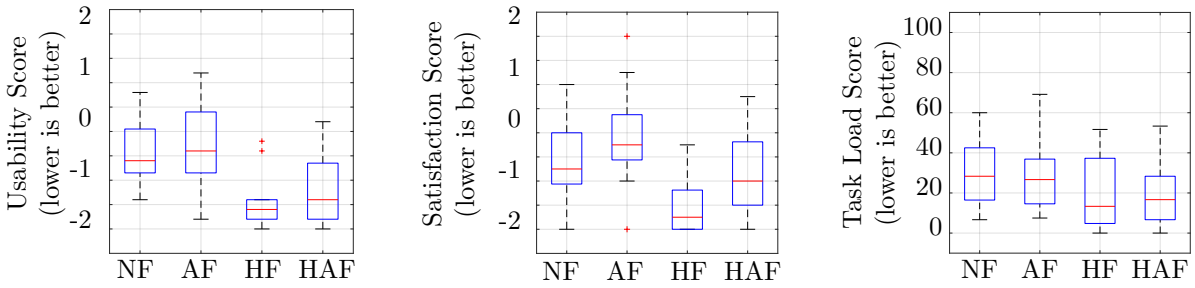


Figure 4: Box plots for the usability score (left), the satisfaction score (middle), and the task load score (right) of the system using different feedback methods.

Finally, the evaluation of the custom questions shows that the participants were more confident in the task execution and more satisfied with the time they needed when haptic feedback and the combined feedback was used (see Fig. 5). In addition, the haptic and combined feedback could improve the ease of use of the system and was described as helpful and intuitive. The results regarding the auditory feedback only show high variations between the participants. Some stated that the auditory feedback was very helpful while others did not find it useful at all. The same conclusion is valid for the satisfaction regarding the needed time. Other than that, the results of auditory feedback are comparable to the teleoperation without any feedback.

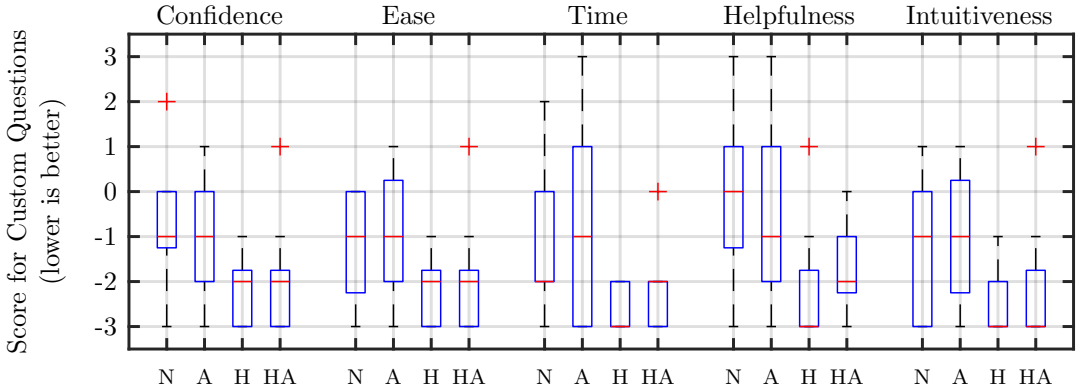


Figure 5: Box plots for the score of the custom questions using different feedback methods.

4 Discussion

On the one hand, the results presented in the previous section show a clear improvement in both usability, satisfaction, and task load of the teleoperation system, as well as in the execution time of the task when haptic feedback or a combination of haptic and auditory feedback is applied. Thus, we can confirm the results presented in [4], where the use of haptic feedback in the context of CTCR teleoperation was investigated as well. On the other hand, the auditory feedback alone does not show a clear benefit in quantitative and qualitative results but received polarizing evaluations, which can be especially seen in Fig. 5. Some of the users stated that the auditory feedback was helpful during the teleoperation and did improve their confidence regarding the task completion and the needed time, while others did not share this opinion. Compared to the haptic feedback, which addresses our sense of touch and directly leads to a change in input device motion, the audio feedback must be first interpreted by the teleoperating user. We believe a longer training phase could improve the helpfulness and satisfaction regarding the auditory feedback as the teleoperating user could become accustomed to it. Furthermore, we think that the strengths of auditory feedback might not have been fully utilized yet and other medical applications could benefit from the usage of auditory feedback, e.g., where complex anatomies obstruct the field of view. We would also like to stress that a user study with physicians could lead to different results as most of the recruited participants in this work were engineering students. Finally, we could envision to implement the auditory display as an addition to the haptic feedback instead of a possible replacement. Thus, the auditory feedback could be used to inform the teleoperating user about certain circumstances during a medical application. As our results indicate, haptic feedback is effective to support the user in navigation intuitively, such that this task alone does not necessarily require further feedback. The strong potential for

auditory signals is to inform the user about irregular events such as workspace limitations or sudden interactions with sensitive anatomy. In these cases, the auditory signal would not need to be permanently enabled, which could lead to a lower cognitive load for the teleoperating user.

5 Conclusion

In this work, the application of haptic as well as auditory feedback methods in the context of CTCR teleoperation is investigated. The conducted user study reveals that the satisfaction and usability of the teleoperation improves when haptic feedback and a combination of haptic and auditory feedback is applied. Furthermore, the task load for these feedback methods can be decreased and the given task can be executed faster on average. The auditory feedback does not show quantitative benefits and got diverse ratings in the questionnaires. We believe, that the full potential of the auditory feedback has not yet been utilized and its use might be beneficial in other scenarios, especially in situations where visual occlusion is present. In future work, we plan to investigate different application areas for auditory feedback in the context of CTCR manipulation and conduct a user study with physicians.

6 References

[1] Gilbert H.B., Rucker D.C., Webster III R.J., *Concentric Tube Robots: The State of the Art and Future Directions*, Robotics Research, pp. 253-269, 978-3-319-28872-7, Springer, (2016)

[2] Burgner-Kahrs J., Rucker D.C., Choset H.,*Continuum Robots for Medical Applications: A Survey*, IEEE Transactions on Robotics Vol. 31(6), pp. 1261-1280, (2015)

[3] Leibrandt K., Bergeles C., Yang G.-Z., *On-line collision-free inverse kinematics with frictional active constraints for effective control of unstable concentric tube robots*, IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3797-3804, (2015)

[4] Kreuer L.J., Lilge S., Burgner-Kahrs J., *Entwicklung von Kontrollalgorithmen für die Teleoperation von tubulären Kontinuumsrobotern*, 16. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie, pp. 79-84, (2017)

[5] Bork F., Fuerst B., Schneider A., Pinto F., Graumann C., Navab N., *Auditory and Visio-Temporal Distance Coding for 3-Dimensional perception in Medical Augmented Reality*, IEEE International Symposium on Mixed and Augmented Reality, pp. 7-12, (2015)

[6] Dixon B., Daly M., Chan H., Vescan A., Witterick I., Irish J., *Augmented real-time navigation with critical structure proximity alerts for endoscopic skull base surgery*, Laryngoscope 124, pp. 853-859, (2014)

[7] Kitagawa M., Dokko D., Okamura A., Yuh D., *Effect of sensory substitution on structure-manipulation forces for robotic surgical systems*, Thoracic and Cardiovascular Surgery Vol. 129(1), pp. 151-158, (2005)

[8] Black D., Hansen C., Nabavi A., Kikinis R., Hahn H., *A Survey of Auditory Display in Image-Guided Interventions*, International Journal of Computer Assisted Radiology and Surgery Vol. 12(10), pp. 1665-1676, (2017)

[9] Rucker D.C., Jones B.A., Webster III R.J., *A Geometrically Exact Model for Externally Loaded Concentric-Tube Continuum Robots*, IEEE Transactions on Robotics Vol. 26(5), pp. 769-780, (2010)

[10] Rucker D.C., Webster III R.J., *Computing Jacobians and Compliance Matrices for Externally Loaded Continuum Robots*, IEEE Conference on Robotics and Automation, pp. 945-950, (2011)

[11] Rosenberg L.B., *Virtual fixtures: Perceptual tools for telerobotic manipulation* , Virtual Reality Annual International Symposium, (1993)

[12] Bergeles C., Gosline A.H., Vasiyev N.V., Codd P.J., Del Nido P.J., Dupont P.E., *Concentric Tube Robot Design and Optimization Based on Task and Anatomical Constraints*, IEEE Transactions on Robotics Vol. 31(1), pp. 67-84, (2015)

[13] MacKenzie I.S., *Human-Computer Interaction*, 978-0124058651, Morgan Kaufmann, (2013)

[14] Van Der Laan J.D., Heino A., De Waard D., *A simple procedure for the assessment of acceptance of advanced transport telematics*, Transportation Research Part C: Emerging Technologies Vol. 5(1), pp. 1-10, (1997)

[15] Hart S.G., Staveland L.E., *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research*, Advances in Psychology Vol. 52, pp. 139-183, (1988)

[16] *National Institutes of Health Print Exchange Website; Lateral Ventricles*: https://3dprint.nih.gov/discover/3dpx-003132, accessed: June 2018

Segmentierung und Tracking von minimal-invasiven robotergeführten Instrumenten

S. Hanisch[1], O. Burgert[1], C. Curio[1], F. A. Fröhlich[2]

[1]Hochschule Reutlingen, Informatik Fakultät, Deutschland
[2]DLR, Robotik und Mechatronik, Oberpfaffenhofen, Deutschland

Kontakt: Simone.Hanisch@Student.reutlingen-university.de

Abstract

Die Segmentierung und das Tracking von minimal-invasiven robotergeführten Instrumenten ist ein wesentlicher Bestandteil für verschiedene computer-assistierte Eingriffe. Allerdings treten in der minimal-invasiven Chirurgie, die das Anwendungsfeld für den hier beschriebenen Ansatz darstellt, häufig Schwierigkeiten durch Reflexionen, Schatten oder visuelle Verdeckungen durch Rauch und Organe auf und erschweren die Segmentierung und das Tracking der Instrumente.

Dieser Beitrag stellt einen Deep Learning Ansatz für ein markerloses Tracking von minimal-invasiven Instrumenten vor und wird sowohl auf simulierten als auch realen Daten getestet. Es wird ein simulierter als auch realer Datensatz mit Ground Truth Kennzeichnung für die binäre Segmentierung von Instrument und Hintergrund erstellt. Für den simulierten Datensatz werden Bilder aus einem simulierten Instrument und realem Hintergrund zusammengesetzt. Im Falle des realen Datensatzes spricht man von der Zusammensetzung der Bilder aus einem realen Instrument und Hintergrund. Insgesamt wird auf den simulierten Daten eine Pixelgenauigkeit von 94.70 Prozent und auf den realen Daten eine Pixelgenauigkeit von 87.30 Prozent erreicht.

Keywords: minimally invasive, segmentation, tracking, neural network

1 Problemstellung

Im Gegensatz zur traditionellen offenen Chirurgie, bei der der Chirurg meist ein gutes Sichtfeld des zu operierenden Gebietes und haptisches Feedback durch die direkte Übertragung der Bewegung des Instruments auf den Patienten hat, treten bei der minimal-invasiven Chirurgie einige Nachteile auf. In den meisten Fällen verfügt der Chirurg über ein eingeschränktes Sichtfeld, das zusätzlich von der Qualität und der Führung der Kamera abhängig ist. Werden robotische Assistenzsysteme genutzt, erfolgt keine direkte Kraftübertragung der Instrumente auf den Patienten, sondern eine Kraftübertragung durch eine manipulative Einheit. Dies hat zur Folge, dass der Arzt kein haptisches Feedback mit der Bewegung der Instrumente erhält und Schwierigkeiten für das Kontextbewusstsein auftreten. [1,2,3,4,5]

Um dem Chirurgen Assistenzfunktionen, welche die genannten Limitationen abmildern können, anbieten zu können, ist es häufig erforderlich die genaue Position und räumliche Relation der OP-Instrumente zu erfassen. Weiterhin kann das Bild mit prä- oder intraoperativen Bildern aus bildgebenden Verfahren überlagert werden und dem Chirurgen zusätzliche Informationen bereitstellen, die sein Kontextbewusstsein erweitern. Folgend ist eine komplette Situationserkennung und darauf aufbauende Assistenzfunktionen möglich. Die markerlose Eigenschaft ist deshalb so wichtig, da eine Anbringung von komplexen Markern auf dem kleinen Raum der Instrumente nur schwer möglich ist. Weiterhin ist der Sterilisationsprozess kein Problem, da das Produkt nicht geändert wird und die Biokompatibilität weiterhin gegeben ist. Darüber hinaus wird bei einem markerlosen Ansatz der Arbeitsablauf des Chirurgen nicht gestört. [1,2,3,4,5]

Eine Lösung zur markerlosen Segmentierung liegt in der semantischen Segmentierung. In aktuellen Ansätzen wird die semantische Segmentierung von modernen Deep Learning Architekturen gelöst. Der Ursprung der Segmentierung liegt in der Klassifizierung, die daraus besteht, eine Vorhersage für das komplette Bild zu treffen. Die Objektlokalisierung ist der nächste Schritt in der Schlussfolgerung, bei dem nicht nur die Klassen, sondern zusätzliche Informationen bezüglich der räumlichen Lage der Klassen erschlossen werden. Bei der semantischen Segmentierung wird ein Schritt weitergegangen und es werden dichte Vorhersagen, bei der jedem Pixel ein Label mit der Klasse des beinhaltenden Objektes zugewiesen wird, getroffen. [6]

2 Material und Methoden

Um einen möglichst vielseitigen und großen Datensatz automatisch zu erzeugen, wird zunächst die Möglichkeit der Simulation gewählt. Da die Visualisierungssoftware MediView bereits ein System zur Visualisierung der MiroSurge Plattform des am DLR entwickelten Operationsroboters ist, wird diese gewählt, um einen Datensatz zu erzeugen. Für die automatische Erzeugung der Ground Truth Daten werden verschiedene Bildverarbeitungstechniken aus der Standardbibliothek Open Source Computer Vision Library (OpenCV) verwendet, um das Instrument mit einem Hintergrund aus einem öffentlichen Datensatz zusammenzusetzen. Zusätzlich wird mit OpenCV eine Ground Truth Kennzeichnung für die binäre Segmentierung von Instrument und Hintergrund erstellt. Im fortlaufenden Text wird diese Kennzeichnung als segmentierte Maske bezeichnet. Neben einem simulierten Datensatz eines robotischen Instruments wird ein weiterer realer Datensatz des robotischen Instruments in einem Phantomumfeld erzeugt. Für die Umsetzung eines Deep Learning Ansatzes wird ein passendes Netzwerk für Segmentierungsaufgaben gewählt und in TensorFlow [7] implementiert. Für das Training des Netzwerkes werden für verschiedene Experimente unterschiedliche Parameter gewählt. Dabei werden Experimente sowohl mit simulierten Daten als auch Echtdaten durchgeführt, um diese miteinander vergleichen zu können. Zusätzlich dienen die Experimente mit simulierten Daten dazu, um zu überprüfen ob der Deep Learning Ansatz für die Lösung des Problems zur semantischen Segmentierung herangezogen werden kann.

2.1 Erzeugung eines Datensatzes

Für die Erzeugung eines simulierten Datensatzes müssen verschiedene Schritte realisiert werden:

- Erstellung der segmentierten Maske als Ground Truth
- Zusammensetzung Instrument mit beliebigem Hintergrund
- Data Augmentation des zusammengesetzten Bildes

Dabei werden für die Erzeugung eines simulierten Datensatzes möglichst viele unterschiedliche Konfigurationen der Gelenkwinkel der Instrumente vor einem grünen Hintergrund und der Kameraposition unterschiedliche Werte übergeben. Die Ground Truth Daten werden dem Neuronalen Netz als Vorhersage, die gelernt werden soll, gegeben. Insgesamt umfasst der simulierte Datensatz ungefähr 2300 Bilder. Sowohl für die Erstellung der segmentierten Maske als auch die Zusammensetzung des Instruments mit einem beliebigen Hintergrund werden verschiedene Algorithmen aus OpenCV verwendet. Die Bandbreite der Algorithmen reicht dabei von Skalierungsalgorithmen, Filtern bis hin zu Algorithmen, die zur Extrahierung einer Farbe dienen. Als segmentierte Maske versteht man die binäre Zuordnung der Farbe Schwarz $RGB_{Schwarz} = (255; 255; 255)$ für den Hintergrund und die Farbe Weiß $RGB_{Weiß} = (0; 0; 0)$ für das Instrument. Dies kann durch die Extrahierung eines Schwellenwertes für die grüne Farbe des Hintergrundes erfolgen. Bei der Erzeugung eines realen Datensatzes, bei dem die realen Instrumente vor einem grünen Hintergrund bewegt werden, muss der untere und oberer Schwellenwert für die Extrahierung entsprechend angepasst werden. Der Datensatz mit realem Instrument umfasst 2200 Bilder.

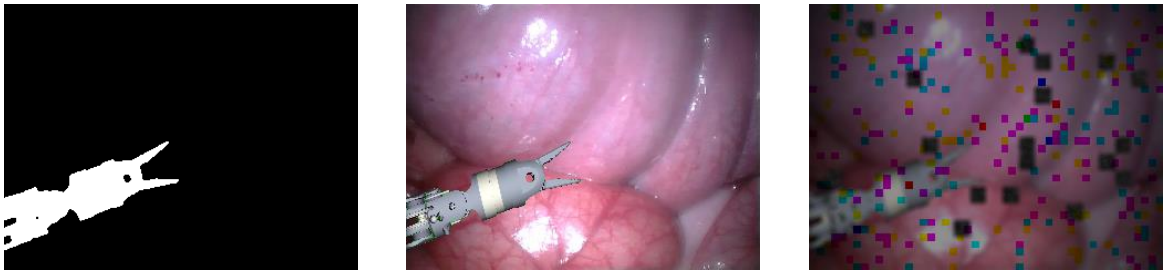


Abbildung 1: *Beispiel aus dem simulierten Datensatz, von links nach rechts: segmentierte Maske, Zusammensetzung Instrument mit Hintergrund und Data Augmentation*

Abbildung 2: *Beispiel aus dem realen Datensatz, von links nach rechts: segmentierte Maske, Zusammensetzung Instrument mit Hintergrund und Data Augmentation*

Für die Zusammensetzung des Instrumentes mit einem beliebigen Bild werden Datensätze des Imperial London College [8] herangezogen. Um einen größeren Datensatz mit leicht veränderten Bildern zu erzeugen, wird eine Data Augmentation auf das zusammengesetzte Bild angewandt. Dies ist nötig, da Deep Learning Netzwerke eine große Menge an Trainingsdaten benötigen, um eine gute Leistung zu erbringen. Zusätzlich führt die Data Augmentation dazu, dass die Kostenfunktion schneller minimiert wird und ein Overfitting des Netzwerkes durch das Handeln als Regularizer verhindert werden kann. Typischerweise besteht die Data Augmentation aus dem Einsetzen einer Menge an Transformationen wie Translation, Rotation, Verzerrung, Skalierung oder Farbverschiebungen [6]. Zusätzlich besteht die Möglichkeit verschiedene Transformationen zu kombinieren. Techniken, die von der imgaug-Python Bibliothek [9] eingesetzt werden sind unter anderen: GaussianBlur, Add (per channel), Multiply, ContrastNormalization (per channel), CoarsePepper, CoarseSalt, EdgeDetect, CoarseDropout oder Elastic Transformation.

Abbildung 1 und Abbildung 2 zeigen jeweils ein Beispiel aus dem simulierten und realen Datensatz.

2.2 Anwendung eines Neuronales Netzes

Für die Implementierung eines Netzwerkes wird die Architektur des U-Net, einem *Convolutional Neural Network* für die Segmentierung von biomedizinischen Bildern, gewählt. Für das Training des Neuronalen Netzes wird eine Kostenfunktion definiert, die minimiert wird, um das Trainingsergebnis zu optimieren. Außerdem wird der Datensatz in Trainings- und Testbilder aufgeteilt, die schubweise (*batch*) verarbeitet werden. Da bei biomedizinischen Anwendungen eine große Anzahl an Trainingsbildern außer Reichweite liegt, kann mit der Architektur des U-Net und einer starken Data Augmentation auch mit sehr wenigen Trainingsbildern eine genaue Segmentierung erreicht werden. Es wird die Architektur eines *Fully Convolutional Neural Network* weiter ausgebaut. Der Grundgedanke der Architektur ist, dass der komprimierende Zweig, der den Kontext erfasst, von fortlaufenden Schichten ergänzt wird, bei denen *pooling* Operationen mit *upsampling* Operationen sich abwechseln. Somit basiert die Architektur des U-Net auf einem komprimierenden Zweig, der den Kontext erfasst und einem symmetrisch expandierenden Zweig, der die präzise Lokalisierung ermöglicht. Daher stellt die Architektur eine U-Form dar. [10]

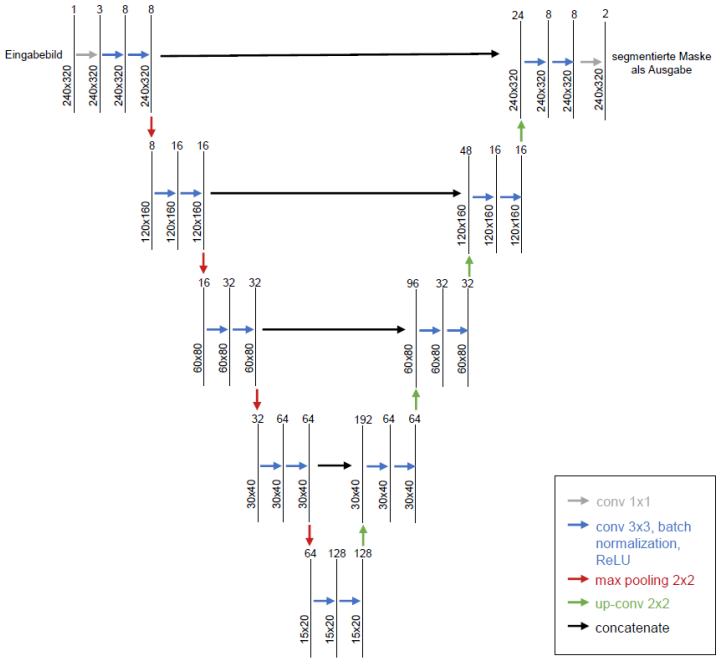


Abbildung 3: *Architektur des benutzten Neuronalen Netzes mit Parametrierung der jeweiligen Operationen.*

In Abbildung 3 ist die implementierte Architektur des U-Net abgebildet. Das Netzwerk startet mit einem Eingabebild der Auflösung 240x320 Pixel. Jeder senkrechte Strich in der Abbildung entspricht einer Merkmalsabbildung. Die Auflösung der Merkmalsabbildung steht links unten und die Anzahl der Merkmale steht rechts oben. In den meisten Fällen wird eine 3x3 *convolution* mit einer unterschiedlichen Anzahl an Merkmalen angewandt. Da das Netzwerk nur *same convolutions* mit *zero padding* benutzt, enthält die Segmentierungsausgabe nur Pixel, für die der volle Kontext im Eingabebild verfügbar ist. Dies erlaubt eine nahtlose Segmentierung von beliebig großen Eingabebildern und gewährleistet, dass die Ausgabe der segmentierten Maske dieselbe Auflösung wie das Eingabebild besitzt. Jeder *convolution* folgt ein *batch normalization* Schritt, eine *ReLU* (Recitified Linear

Unit) Aktivierungsfunktion und eine *max pooling* Operation in der Nachbarschaft von 2 Pixeln. Bei jedem *downsampling* Schritt wird die Anzahl der Merkmale verdoppelt. Dies führt dazu, dass das Eingabebild bis auf eine Auflösung von 15x20 Pixeln mit 128 Merkmalen reduziert wird. Für die Segmentierungsaufgabe folgt nun der expandierende Zweig. Jeder Schritt in dem expandierenden Zweig besteht aus einer 2x2 *up-convolution* und einer Verkettung mit der korrespondierenden zugeschnittenen Merkmalsabbildung des komprimierenden Pfades. Danach folgen jeweils zwei 3x3 *convolutions* mit einer *batch normalization* und einer *ReLU* Aktivierungsfunktion. Somit wird am Ende des *upsamplings* eine Merkmalsabbildung von 240x320x8 erreicht. Die finale Schicht wird mit einer 1x1 *convolution* abgeschlossen, bei dem jeder 8-Einheiten Merkmalsvektor auf die gewünschte Anzahl an Klassen, in diesem Fall 2 Klassen, eine für das Instrument selber und eine für den Hintergrund, abgebildet wird.

Für das Training des Neuronalen Netzes wird für die pixelweise Segmentierung die *softmax cross-entropy* Kostenfunktion gewählt. Als Optimierungsalgorithmus wird der *Adam* Algorithmus, der eine individuelle *learning rate* benutzt, um die Gewichte zu aktualisieren, gewählt. Da das Training in einer Schleife durchlaufen wird, muss TensorFlow mitgeteilt werden, wie viele Iterationen durchlaufen werden. Als *batch* bezeichnet man die Anzahl der Trainingsbeispiele, die dem Neuronalen Netz in einem Trainingsdurchlauf zugeführt werden. Daher muss vor dem Training des Neuronalen Netzwerkes die *learning rate*, *batch* Größe und Anzahl der Iterationen festgelegt werden.

2.3 Experimente

Insgesamt werden drei verschiedene Modelle trainiert. Im ersten Modell *Modell Sim* werden ausschließlich simulierte Daten (Trainingsdatensatz 2257 Bilder, Testdatensatz 103 Bilder) für das Training verwendet, im zweiten Modell *Modell Mix* werden simulierte und reale Daten (Trainingsdatensatz 4538 Bilder, Testdatensatz 243 Bilder) herangezogen und im dritten Modell *Modell Real* werden für das Training ausschließlich reale Daten (Trainingsdatensatz 2281 Bilder, Testdatensatz 140 Bilder) eingesetzt. Für die Hyperparameter *learning rate*, *batch* Größe und Anzahl der Iteration für das Training des Netzwerkes werden für die unterschiedlichen Modelle dieselben Werte eingesetzt, da diese die besten Ergebnisse geliefert haben. Ebenfalls die Data Augmentation unterscheidet sich zwischen den Modellen nicht. Für die *learning rate* und *batch* Größe werden die Werte 1e-4 bzw. 2 gewählt. Die Anzahl der Epochen wird so angepasst, dass bei gegebener Anzahl der Trainingsbilder ungefähr 60 000 Iterationen erreicht werden. Das Training wird auf einer GeForce Titan X GPU durchgeführt.

3 Ergebnisse

Um festzustellen, ob ein erfolgreiches Training des Netzwerkes stattgefunden hat, müssen die trainierten Modelle auf Testdaten evaluiert werden. Das Netzwerk wurde nicht mit diesen Daten trainiert. Schließlich werden die Experimente in einem letzten Schritt mit Standardmetriken evaluiert, um einen gerechten Vergleich mit bestehenden Methoden ziehen zu können. Um die Genauigkeit des Modells bei der semantischen Segmentierung zu bestimmen, kann eine Vielzahl an Evaluierungsmetriken verwendet werden. Meist sind die Evaluierungsmetriken bezüglich der Genauigkeit Varianten der Pixelgenauigkeit und des Intersection Over Union (IOU). Die Pixelgenauigkeit ist die einfachste Metrik und wird mit dem Verhältnis von richtig klassifizierten Pixeln durch die Gesamtmenge aller klassifizierten Pixel berechnet. Als zweite Evaluierungsmetrik wird das IOU untersucht. Das IOU berechnet das Verhältnis zwischen der Schnittmenge und der Vereinigungsmenge. [6, 11]

3.1 Evaluierung mit simulierten Daten

Abbildung 4 zeigt die Evaluierungsergebnisse der *Modelle Sim*, *Mix*, *Real* und die Ground Truth für die simulierten Daten. In Tabelle 1 sind die Ergebnisse des Beispiels für die Pixelgenauigkeit und das IOU aufgezeigt. Die jeweils besten Ergebnisse sind in fett gedruckt. Die Pixelgenauigkeit wird mit PA für Pixelaccuracy abgekürzt. Für den Mittelwert und die Standardabweichung der Pixelgenauigkeit und des IOU wurden insgesamt zum Vergleich der Evaluierung mit realen Daten, bei welcher 20 Beispiele ausgewertet werden, ebenfalls 20 Beispiele herangezogen.

Abbildung 4: Beispiel eines *Ergebnisses mit simulierten Daten, von links nach rechts: Original, Ground Truth, Modell Sim, Modell Mix und Modell Real*

Tabelle 1*: Ergebnisse mit simulierten Daten*

	IOU [%] *Modell Sim*	PA [%] *Modell Sim*	IOU [%] *Modell Mix*	PA [%] *Modell Mix*	IOU [%] *Modell Real*	PA [%] *Modell Real*
Beispiel	**98.28**	**99.13**	96.15	98.03	88.97	94.16
Mittelwert μ	**90.28**	**94.70**	81.03	88.50	71.42	80.99
Standardabweichung *S*	**7.76**	**4.83**	16.59	11.72	23.25	18.21

3.2 Evaluierung mit realen Daten

Abbildung 5 zeigt die Evaluierungsergebnisse der Modelle Sim, Mix, Real und die Ground Truth für die realen Daten. In diesem Fall wird für die Evaluierung ein neuer Datensatz in einem Phantomumfeld erzeugt. Die realen Instrumente werden direkt vor einem künstlichen Gewebe bewegt. Für den Mittelwert und die Standardabweichung der Pixelgenauigkeit und des IOU wurden insgesamt zur Erleichterung der Auswertung 20 Beispiele, da jedes Beispiel manuell gekennzeichnet werden muss, herangezogen. In Tabelle 2 sind die Ergebnisse des Beispiels für die Pixelgenauigkeit und das IOU aufgezeigt. Die jeweils besten Ergebnisse sind in fett gedruckt.

Abbildung 5: Beispiel eines *Ergebnisses mit realen Daten, von links nach rechts: Original, Ground Truth, Modell Sim, Modell Mix und Modell Real*

Tabelle 2*: Ergebnisse mit realen Daten*

	IOU [%] *Modell Sim*	PA [%] *Modell Sim*	IOU [%] *Modell Mix*	PA [%] *Modell Mix*	IOU [%] *Modell Real*	PA [%] *Modell Real*
Beispiel	52.39	68.76	74.96	85.68	**82.36**	**90.32**
Mittelwert μ	42.14	58.47	70.25	82.18	**77.76**	**87.30**
Standardabweichung *S*	10.92	11.34	9.14	6.64	**7.19**	**4.74**

4 Diskussion

Bei der Evaluierung mit simulierten Daten wird deutlich, dass *Modell Sim* in den meisten Beispielen *Modell Mix* und *Real* übertrifft. Mit einem Mittelwert der Pixelgenauigkeit von 94.70 % und dem IOU von 90.28 % zeigt dieses Modell die besten Ergebnisse für simulierte Daten. Ebenfalls die Standardabweichung zeigt bei *Modell Sim* einen geringeren Wert von 7.76 % des IOU und von 4.83 % der Pixelgenauigkeit an. Es wird deutlich, dass sowohl *Modell Mix* als auch *Real* Hintergrundpixeln die Instrumentenklasse zuweisen. Dies liegt vermutlich daran, dass in den realen Daten, die sowohl im Datensatz von *Modell Mix* als auch Modell Real vorliegen, vermehrt Reflexionen im Instrument auftreten, und daher Reflexionen im Hintergrund dem Instrument zugewiesen werden. Bei der Evaluierung mit realen Daten wird deutlich, dass *Modell Real* in den meisten Beispielen *Modell Sim* und *Mix* übertrifft. Mit einem Mittelwert der Pixelgenauigkeit von 87.30 % und dem IOU von 77.76 % zeigt dieses Modell die besten Ergebnisse für reale Daten. Ebenfalls die Standardabweichung zeigt bei *Modell Real* einen geringeren Wert von 7.19 % des IOU und von 4.73 % der Pixelgenauigkeit an. Damit sind die Ergebnisse der realen Daten zwar nicht so gut wie die der simulierten Daten, zeigen aber dennoch gute Segmentierungsergebnisse. Es ist zu erkennen, dass in *Modell Sim* weder in den Beispielen mit dem simulierten als auch realen Datensatz keine *false positives* auftreten. Diese können im laufenden Prozess allerdings umgangen werden, wenn lediglich die größte zusammenhängende Region als Instrument erkannt wird. Die Ergebnisse der durchgeführten Experimente mit Datensätzen aus simulierten als auch realen Instrumenten lassen darauf schließen, dass die Erweiterung eines Datensatzes mit simulierten Instrumenten nicht hilfreich ist, um reale Instrumente zu erkennen.

5 Zusammenfassung

Für die Segmentierung der Instrumente wurde das U-Net aus dem Bereich der Segmentierung von biomedizinischen Bildern gewählt. Diese Architektur entspricht einer bekannten und bewährten Vorgehensweise für Segmentierungsaufgaben und einer Weiterführung der erfolgreich etablierten *Fully Convolutional Neural Networks*. Da in der medizinischen Bildgebung vermehrt Schwierigkeiten bei der Verfügbarkeit von annotierten Datensätzen auftreten und Deep Learning auf einer großen Menge von Daten basieren, wurde in dieser Arbeit zunächst die Möglichkeit der Simulation gewählt. Bei der Evaluierung des Deep Learning Ansatzes steht die Qualität der Segmentierung gegenüber anderen Metriken im Vordergrund. Bei der Evaluierung dieses Ansatzes mit simulierten Daten kann eine sehr gute Pixelgenauigkeit von 94.70 % und einem IOU von 90.28 % erreicht werden. Allerdings zeigt sich bei der Evaluierung für reale Daten, dass diese nicht befriedigend segmentiert werden können. Dies liegt unter anderem an den starken Reflexionen, die bei realen Instrumenten auftreten, jedoch in der Simulation nicht ausreichend simuliert werden können. Es wird vermutet, dass die simulierten Instrumente gegenüber den realen Instrumenten zu unterschiedlich sind. Aufgrund dessen werden neben den simulierten Daten ebenfalls reale Daten in den Trainingsdatensatz für das Modell aufgenommen. Dabei werden Verbesserungen der Pixelgenauigkeit von 58.47 % auf 82.18 % und des IOU von 42.14 % auf 70.25 % erreicht. Wird ein Modell jedoch alleinig mit realen Daten trainiert, ohne dass simulierte Daten zum Einsatz kommen, so wird eine Steigerung auf 87.30 % und 77.76 % erreicht. Dennoch hängt die Verbesserung des Modells von vielen Faktoren ab und lässt offen, ob das Modell durch simulativ erzeugt Faktoren verbessert werden kann, zum Beispiel durch die Optimierung der Simulationsqualität. Ebenfalls eine Erhöhung der Trainingsepochen des *Modell Mix* kann zu einer Verbesserung führen, da bei diesem Experiment durch die Erhöhung der Trainingsbilder die Anzahl der Epochen zu den anderen beiden Epochen unterscheidet. Somit lässt sich das Ergebnis, dass die Evaluierung von Modellen basierend auf realen Daten mit einem realen Trainingsdatensatz am besten geeignet ist, nicht ohne weitere Untersuchungen auf andere Anwendungsszenarien übertragen.

Referenzen

[1] Garcia-Peraza-Herrera, Luis & Li,Wenqi & Fidon, Lucas & Gruijthuijsen, Caspar & Devreker, Alain & Attilakos, George & Deprest, Jan & Vander Poorten,x Emmanuel & Stoyanov, Danail & Vercauteren, Tom & Ourselin, Sebastien, *ToolNet: Holistically-Nested Real-Time Segmentation of Robotic Surgical Tools*, 2017

[2] Bouget, David & Allan, Max & Stoyanov, Danail & Jannin, Pierre, *Vision-Based and Marker-Less Surgical Tool Detection and Tracking: a Review of the Literature. Medical Image Analysis*, 2016

[3] Garcia-Peraza-Herrera, Luis & Li, Wenqi & Gruijthuijsen, Caspar & Devreker, Alain & Attilakos, George & Deprest, Jan & Vander Poorten, Emmanuel & Stoyanov, Danail & Vercauteren, Tom & Ourselin, Sébastien, *Real-Time Segmentation of Non-Rigid Surgical Tools based on Deep Learning and Tracking*, 2016

[4] Pakhomov, Daniil & Premachandran, Vittal & Allan, Max & Azizian, Mahdi & Navab, Nassir, *Deep Residual Learning for Instrument Segmentation in Robotic Surgery*, 2017

[5] Sarikaya, Duygu & J. Corso, Jason & A. Guru, Khurshid, *Detection and Localization of Robotic Tools in Robot-Assisted Surgery Videos Using Deep Neural Networks for Region Proposal and Detection*, IEEE Transactions on Medical Imaging, 2017

[6] Alberto Garcia-Garcia & Sergio Orts-Escolano & Sergiu Oprea & Victor Villena-Martinez & Jose Garcia-Rodriguez, *A Review on Deep Learning Techniques Applied to Semantic Segmentation*, Transactions on Pattern Analysis and Machine Intelligence TPAMI on Apr. 22, 2017, 2017

[7] TensorFlow, Open-Source Machine Learning Framework, https://www.tensorflow.org/, aufgerufen am 10.06.2018

[8] Imperial College London, *Hamlyn Centre Laparoscopic / Endoscopic Video Datasets*, http://hamlyn.doc.ic.ac.uk/vision/, aufgerufen am 10.06.2018

[9] Data Augmentation, imgaug - Python Bibliothek, https://github.com/aleju/imgaug, aufgerufen am 10.06.2018

[10] Olaf Ronneberger & Philipp Fischer & Thomas Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer. LNCS, Vol.9351: 234241, 2015

[11] Md Atiqur Rahman & Yang Wang, *Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation*, Advances in Visual Computing: 12th International Symposium, ISVC 2016, Las Vegas, NV, USA, 2016

Kollaborative Interaktion für die roboterassistierte ultraschallgeführte Biopsie

J. Berger[1], M. Unger[1], J. Keller[1], R. Bieck[1], L. Landgraf[1], T. Neumuth[1], A. Melzer[1]

[1]*Innovation Center Computer Assisted Surgery (ICCAS), Medizinische Fakultät, Leipzig, Deutschland*

Kontakt: johann.berger@medizin.uni-leipzig.de

Abstract

Für die Evaluation neuer Interaktionskonzepte in der kollaborativen Robotik wurde ein mobiler Demonstrator zur Nadelführung bei ultraschallgestützten Biopsien entwickelt. Unter Verwendung eines KUKA LBR iiwa 7 Roboterarms wurden verschiedene Touch-Gesten implementiert und in den Ablauf der Biopsie integriert. Ein mobiles Clarius L7 Ultraschallgerät wurde um eine flexible Nadelführung ergänzt und am Endeffektor des Roboters montiert. In System-Tests führten Probanden mit technischem und medizinischem Hintergrund sowohl manuelle als auch vom Roboter assistierte Biopsien an einem Abdomen Phantom durch.

Keywords: Robotik, ultraschallgestützt, Biopsie, Nadelführung, Gesten

1 Problemstellung

Robotersysteme verfügen über sehr hohe Präzision und können die Leistung des Chirurgen beim Durchführen von Biopsien enorm steigern. So wurden bereits Systeme vorgestellt, die beim Vorschub der Biopsienadel Assistenz leisten. Komplizierte Interaktion und fehlende Echtzeit-Bildgebung haben jedoch oft einen schlechten Einfluss auf die Akzeptanz. Kollaborative Konzepte für die Interaktion mit Robotern bieten hierfür vielversprechende Ansätze [1]-[2]. Diese Arbeit baut auf diesen Prinzipien auf, um neuartige Konzepte für Interaktionen mittels *Touch*-Gesten und die Handführung an einem kollaborativen Setup zu validieren.

2 Material und Methoden

2.1 Systemdesign

Ein KUKA LBR iiwa 7 R800 Roboter-Arm (Kuka AG, Augsburg, Deutschland) wurde in einem mobilen Aufbau für die Nadelführung bei ultraschallgestützten Biopsien integriert. Die Software für die *Touch*-Gesten Erkennung und Bewegungsplanung ist separat im Robot Operating System (ROS) implementiert und steuert den Roboter über die Standard Sunrise Cabinet Konsole an. Ein mobiles Clarius L7 (Clarius Mobile Health Corp, Kanada) Ultraschall (US) Gerät wurde mittels einer Halterung aus dem 3D Druck am Roboter montiert. Die Halterung umfasst ebenfalls eine einstellbare Nadelführung und einen Handgriff mit Druckknopf zur Freigabe des Handführmodus. Die Planung der Ziel-Trajektorien in präoperativen Bilddaten ist durch eine *Augmented-Reality* basierte Tablet-Anwendung realisiert. Patient und Tablet werden über ein optisches Tracking-System referenziert.

2.2 Studiendesign

Das System wurde an einem Abdomen-Phantom (Triple Modality 3D Abdominal Phantom, CIRS Inc., USA) getestet. Neun Probanden unterschiedlicher Expertisen führten Versuche an zwei verschiedenen Defekten durch. Zunächst verwendeten die Probanden das Clarius L7, um manuell beide Läsionen jeweils zweimal mit einer Nadel zu punktieren. Anschließend wurden die Interventionen mit Roboter-Assistenz wiederholt. Die Trajektorien wurden von den Probanden mit einem optischen *Pointer* geplant und der Winkel für die Nadelführung berechnet. Durch einen *Touch*-Befehl am Roboter wurde das Clarius L7 platziert. Die Genauigkeit der Gestenerkennung wurde in einer früheren Arbeit evaluiert [3]. Die Probanden führten den US-Kopf per Hand entlang der geplanten Trajektorie zum Phantom. Anschließend stellten sie den Winkel ein und schoben die Nadel durch die Führung ins Ziel. Nach einer Treffer-Bestätigung wurde der Roboter vom Phantom entfernt.

Die Zeiten und Anzahl der Versuche bis zum Treffen des Ziels wurden festgehalten. Ein Fragebogen diente der Erfassung der Nutzerfreundlichkeit des Systems.

3 Ergebnisse

Bei der manuellen Biopsie benötigten die Probanden im Durchschnitt 22 Sekunden, um die Läsionen im Phantom zu finden. Die Zeit das Ziel mit der Nadel zu punktieren lag im Schnitt bei 50 Sekunden. Die Planungszeit für die assistierte Biopsie betrug dagegen durchschnittlich 28 Sekunden und die Zeit zum Einstellen der Führung und Treffen des Ziels 64 Sekunden. Ohne Assistenz benötigten die Probanden bis zu fünf Versuche das Ziel erfolgreich zu punktieren. Im assistierten Falle wurde das Ziel immer beim ersten Versuch getroffen. In einem Ausnahmefall schob eine Probandin aufgrund eines Bildartefaktes die Nadel zur Sicherheit ein zweites Mal im Phantom vor. Die Nadelführung wurde von den Probanden als Verbesserung und das Interaktionskonzept als intuitiv empfunden (Abbildung. 1).

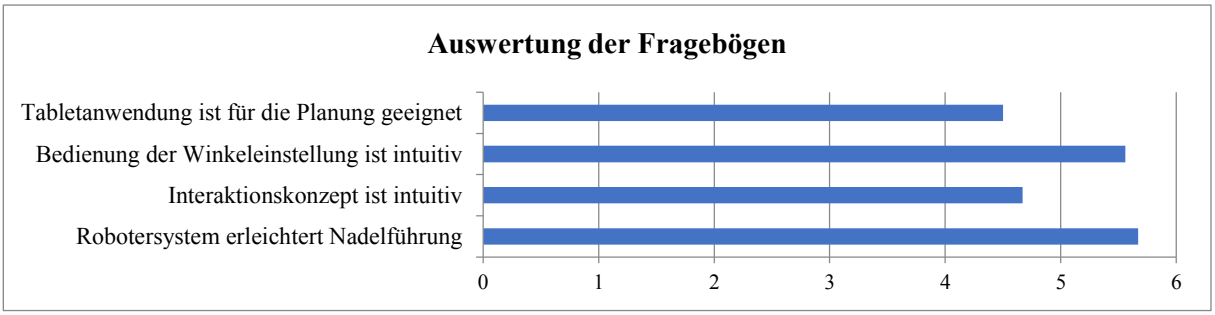


Abbildung 1: Auswertung der Fragebögen zum Assistenzsystem und Interaktionskonzept (6 - sehr gut, 1 - sehr schlecht). Es sind pro Frage die gemittelten Bewertungen aller 9 Probanden angegeben.

4 Diskussion

In dieser Arbeit konnten neuartige Interaktionsmöglichkeiten in ein kollaboratives Roboter-System für die interventionelle Positionierung von Biopsienadeln integriert werden. Das System wurde im Vergleich zur nicht assistierten US-gestützten Biopsie evaluiert. Ein direkter Vergleich der Eingriffszeiten ist aufgrund der unterschiedlichen Abläufe zwar nicht möglich, jedoch zeigen analoge Arbeitsschritte keine wesentliche Verzögerung im Eingriff. Von Probanden gewünschte Verbesserungen in der Planungsanwendung können die benötigte Zeit weiter reduzieren. Zudem konnte durch die Arbeit mit dem Robotersystem die Anzahl der Versuche zum Treffen des Ziels deutlich reduziert werden. Dies verbessert unter anderem auch die Einschätzung der Interaktionsqualität bei der assistierten Nadelführung.

5 Zusammenfassung

Die roboterassistierte Positionierung von Biopsienadeln reduziert die Anzahl der Punktionen. Sie bietet somit eine Verbesserung des Eingriffes und schont Arzt und Patient. Zur Steigerung der Nutzerfreundlichkeit kann mittels *Touch*-Gesten mit dem System interagiert werden. Diese natürlichere Art der Interaktion wurde von den Probanden als positives Interaktionskonzeptes bewertet.

Referenzen

[1] A. Simorov, R. S. Otte, C. M. Kopietz, and D. Oleynikov, “Review of surgical robotics user interface: what is the best way to control robotic surgery?,” Surg Endosc, vol. 26, no. 8, pp. 2117–2125, Aug. 2012.

[2] E. Magrini, F. Flacco, and A. D. Luca, “Control of generalized contact motion and force in physical human-robot interaction,” in 2015 IEEE International Conference on Robotics and Automation (ICRA), 2015, pp. 2298–2304.

[3] J. Berger, M. Unger, L. Landgraf, R. Bieck, T. Neumuth, and A. Melzer, “Assessment of Natural User Interactions for Robot-Assisted Interventions,” in Current Directions in Biomedical Engineering, Submitted.

Restdauerprädiktion bei laparoskopischen Eingriffen mittels nicht annotierten, multimodalen Daten

Sebastian Bodenstedt [1], *Martin Wagner* [2], *Lars Mündermann* [3], *Hannes Kenngott* [2], *Beat Müller-Stich* [2], *Sören Torge Mees* [4], *Jürgen Weitz* [4], *Stefanie Speidel* [1]

[1] *Translationale Chirurgische Onkologie, Nationales Centrum für Tumorerkrankungen (NCT), Dresden, Deutschland*
[2] *Allgemein-, Viszeral- und Transplantationschirurgie, Universitätsklinikum Heidelberg, Heidelberg, Deutschland*
[3] *Karl Storz SE & Co. KG – Tuttlingen, Tuttlingen, Deutschland*
[4] *Viszeral-, Thorax- und Gefäßchirurgie, Universitätsklinikum Dresden, Dresden, Deutschland*

Kontakt: sebastian.bodenstedt@nct-dresden.de

Abstract

Der Verlauf von chirurgischen Eingriffen ist oft nicht vorhersehbar, so dass es schwierig ist, die Dauer eines Eingriffes korrekt abzuschätzen. Kontextsensitive Methoden, die den Arbeitsablauf, anhand Informationen zum aktuellen Stand des Eingriffes, online analysieren und automatisch die verbleibende Dauer vorhersagen, würden diese Probleme lindern. Während laparoskopischer Eingriffe ist der endoskopische Videostrom hierfür eine ideale Informationsquelle. Das Extrahieren von quantitativen Informationen aus dem Video ist jedoch aufgrund der hohen Dimensionalität eine Herausforderung. Andere chirurgische Vorrichtungen (z. B. Insufflator, Lichter usw.) stellen Datenströme bereit die, im Gegensatz zu dem Videostrom, kompakter und leichter zu quantifizieren sind. Ob derartige Ströme jedoch ausreichende Informationen zur Schätzung der Operationsdauer liefern, ist ungewiss. In dieser Arbeiten stellen wir Methoden vor um die Dauer von laparoskopischen Interventionen basierend auf nicht annotierten Daten kontinuierlich vorherzusagen. Die Methoden werden an 80 Aufnahmen von laparoskopischen Eingriffen unterschiedlicher Art ausgewertet, für die Daten von chirurgischen Geräten und endoskopische Videoströme zur Verfügung stehen.

Keywords: Restdauerprädiktion, Workflow Analyse, Laparoskopie

1 Problemstellung

Die Zeit im Operationssaal (OP) und die Zeit des chirurgischen Personals sind kostenintensive Krankenhausressourcen und müssen genau eingeteilt werden. Die Planung der Verwendung des OP kann nicht statisch sein, da ein Vorgang, der länger dauert als zuvor geschätzt, dazu führen kann, dass die folgenden Operationen zurückgeschoben oder sogar abgebrochen werden, was sowohl dem Patienten als auch dem OP-Personal Unannehmlichkeiten bereitet. Auf der anderen Seite, wenn ein Eingriff vorzeitig beendet wird, bleibt das OP ungenutzt, was unnötige Leerlaufzeit für das chirurgische Personal zur Folge hat. Um dies zu verhindern, müssen durch die OP-Planung Zeitpläne dynamisch aktualisiert werden. Dies wird durch die Unvorhersehbarkeit des chirurgischen Arbeitsablaufs erschwert, welche es schwierig macht die Dauer der Eingriffe vorher genau abzuschätzen. Daher muss die OP-Planung ständig über den Fortschritt der laufenden Eingriffe informiert sein. Zu diesem Zweck muss sie periodisch den Status von Eingriffen abfragen, was zu sehr subjektiven Vorhersagen und vermeidbaren Störungen führt.

Integrierte OPs werden in Krankenhäusern immer häufiger eingesetzt und ermöglichen den Zugriff auf Datenströme von chirurgischen Geräten wie Kameras, Insufflatoren, Lichter usw. während eines Eingriffes. Solche Datenströme können Informationen bereitstellen, die eine kontextbezogene Assistenz ermöglichen, wie beispielsweise automatisch und kontinuierlich den Fortschritt einer laufenden Operation vorherzusagen. Insbesondere das endoskopische Video, über welches laparoskopische Eingriffe durchgeführt werden, bietet eine riesige Menge an Informationen. Methoden zur Workflow Analyse können verwendet werden, um Eingriffe in chirurgische Phasen zu unterteilen. Häufig wird die Instrumentennutzung, die einen Indikator für einen fortschreitende Operationsstatus darstellt, verwendet um die aktuelle chirurgische Phase zu bestimmen [1, 2, 3, 4]. Convolutional Neural Networks (CNNs) werden ebenfalls verwendet, um die chirurgische Phase direkt aus dem endoskopischen Videostrom herzuleiten [5, 6, 7].

Verfahren zur chirurgische Phasenerkennung können verwendet werden, um die Dauer von chirurgischen Eingriffen zu approximieren, aber diese Verfahren erfordern im Allgemeinen eine große Menge an annotierten Trainingsbeispielen als Eingabe [9]. Da Phasenmodelle im Allgemeinen für einen bestimmten Interventionstyp

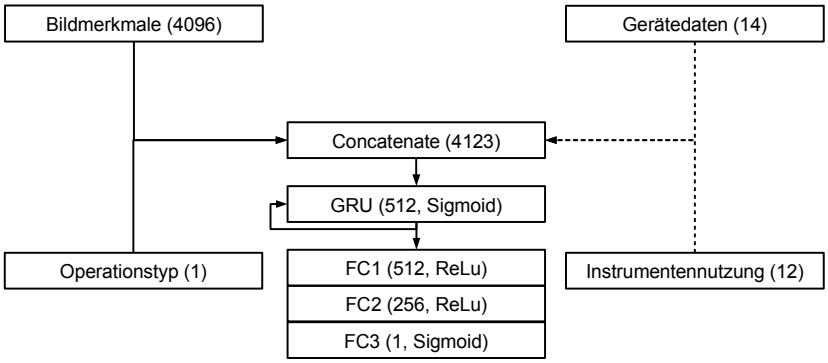


Abbildung 1: Die CNN-RNN Architektur für die Vorhersage der Restdauer von Videosignal, Instrumentennutzung und Gerätedaten. Die gestrichelten Linien symbolisieren optionale Eingaben. *FC* sind vollständig vernetzte Schichten und *GRU* steht für Gated Recurrent Unit. Die Zahlen in den Klammern stehen für die Anzahl an Neuronen und für die genutzte Nichtlinearität.

spezialisiert sind, müssten mehrere Detektoren trainiert werden. Daher würde die Verwendung eines phasenbasierten Verfahrens als allgemeine Lösung zur Bestimmung der verbleibenden Operationsdauer eine große Menge an annotierten Trainingsdaten erfordern. In [8] schlagen die Autoren ein System vor, das die verbleibende Operationszeit während laparoskopischer Cholezystektomien ohne Operationsphasen, aber direkt aus der Verwendung des elektrochirurgischen Geräts bestimmt.

In diesem Beitrag stellen wir drei auf CNN basierende Methoden zur direkten Vorhersage der Dauer laparoskopischer Interventionen vor. Diese Verfahren erfordern keine manuell annotierten Trainingsdaten und funktionieren für verschiedene Arten von laparoskopischen Verfahren. Die Auswertung der Methoden erfolgt an einem Datensatz mit 80 laparoskopischen Operationen unterschiedlicher Art und auf dem Cholec80-Datensatz [7]. Nach unserem Wissen ist unser Ansatz der erste, der die Dauer von laparoskopischen Eingriffen basierend auf einer Kombination aus nicht annotierten Video- und chirurgischen Gerätedaten vorhersagt.

2 Material und Methoden

Voraussetzung für eine Vorhersage der Restdauer von laparoskopischen Eingriffen sind Informationen über den aktuellen Zustand der Operation. Da der endoskopische Videosignal als Grundlage für die Handlungen des Chirurgen dient, gehen wir davon aus, dass er ausreichende Informationen über den Zustand des Verfahrens enthält, obwohl die Extraktion quantitativer Informationen aufgrund der hohen Dimensionalität des Datenstroms eine Herausforderung darstellt. Hier stellen wir eine Kombination aus CNN und einem rekurrenten neuronalen Netzwerk (RNN) vor, die es ermöglicht, die Dauer einer Operation aus einem endoskopischen Videostrom vorherzusagen. Außerdem schlagen wir eine Variante dieses neuronalen Netzwerk vor, das explizit die Erkennung von Instrumentennutzung durchführt und diese Information zur weiteren Verbesserung ihrer Vorhersagen verwendet.

Weiterhin hat die Verfügbarkeit von integrierte OPs in Krankenhäusern eine steigende Tendenz. Diese OPs ermöglichen den Zugriff auf Datenströme von chirurgischen Geräten in Form von Zeitreihen. Diese Zeitreihen sind, im Gegensatz zum Videosignal, kompakter und leichter zu quantifizieren, enthalten jedoch eine geringere Menge an Information. Wir stellen die Hypothese auf, dass sowohl Datenströme als auch Video- und Gerätedaten komplementäre Informationen enthalten, was bedeutet, dass eine Kombination der beiden die Genauigkeit der Vorhersagen erhöhen sollte. Um diesen Aussage zu evaluieren, schlagen wir eine Fusion der beiden Ströme in einem neuronalen Netzwerk (Abbildung 1) vor.

2.1 V-Net: Bildbasierte Schätzung

Die Topologie unseres CNN-RNN zur Vorhersage der Dauer der Operation aus dem endoskopischen Video, V-Net, ist identisch zu dem Netzwerk für die Phasenerkennung vorgeschlagen in [5], die aus einem Alexnet [10] Stil Netzwerk (Bildmerkmale in Abbildung 1) kombiniert mit einer Gated Recurrent Unit [11] (GRU) um zeitliche Informationen zu integrieren. Wir trainieren die der GRU vorausgehenden Schichten mit der in [5] vorgeschlagenen Methode vor. Als Input für V-Net tasteten wir Videos mit einer Rate von einem Bild pro Sekunde ab und reduzierten jedes Bild auf eine Auflösung von 320×240 Pixel. Dies wurde durchgeführt um die Datengröße und die Trainingszeit zu reduzieren.

Darüber hinaus geben wir V-Net Informationen über den Typ der durchgeführten laparoskopischen Operation (Operationstyp in Abbildung 1). Zu diesem Zweck kategorisieren wir unsere Daten in 5 allgemeine Typen von laparoskopischen Eingriffen (Tabelle 1). Wir weisen jedem Videoframe eine Zahl zwischen 0 und 1 als Label zu, d.H. die Bezeichnung des Videoframes i von einem Video, das aus N Videoframes besteht, ist $l_i = \frac{i}{N}$.

Operationstype	ID
Colorektal	1
Oberer Gastrointestinal Trakt und bariatrisch	2
Hepato Pankreato Biliäre	3
Allgemeine Laparoskopie	4
Einzelfälle	5

Tabelle 1: Allgemeinen Typen der verwendeten laparoskopischen Eingriffe

Da die der GRU vorausgehenden Schichten vortrainiert sind, optimieren wir nur die Gewichte der neu hinzugefügten Schichten. Dazu verwenden wir den Adam-Optimizer [12] mit einer Lernrate von 10^{-6} und einem sigmoid cross-entropy loss. Das Netzwerk wird für 50 Iterationen trainiert.

2.2 VI-Net: Schätzung basierend auf Bild- und Instrumentennutzung

Da das Vorhandensein eines bestimmten chirurgischen Instrumententypes im endoskopischen Video oft ein wichtiger Hinweis für den Fortschritt der Operation ist [7], gehen wir davon aus, dass die explizite Einbeziehung von Informationen über die derzeit sichtbaren Arten von Instrumenten die Leistung der Vorhersage der Prozedurdauer erhöhen würde.

Um die derzeit im endoskopischen Video verwendeten Instrumente zu identifizieren, modifizieren wir ein vortrainiertes ResNet-152 [13] für die Erkennung der Instrumentennutzung. Hier ersetzen wir die vortrainierte vollvernetzte Ausgangsschicht durch eine andere vollvernetzte Schicht, die aus 12 Knoten (einer pro Instrumententyp) mit sigmoiden Nichtlinearitäten besteht. Die Gewichte der Schichten *conv5_x* und der neuen vollständig vernetzten Schicht werden mit Hilfe eines Datensatzes von 24 kolorektalen Laparoskopien trainiert, in dem die Anwesenheit von 12 möglichen chirurgischen Instrumenten in einem Bild pro Minute annotiert wurde. Zum Trainieren wird der Adam-Optimizer [12] mit einer Lernrate von 10^{-5} und dem binäre cross-entropy loss verwendet.

Um Informationen über die Instrumentennutzung zur Vorhersage der Operationsdauer zu nutzen, erweitern wir die Architektur von V-Net auf VI-Net, um eine zusätzliche Eingabeschicht (Instrumentennutzung in Abbildung 1) zu akzeptieren. VI-Net wird auf die gleiche Weise trainiert wie V-Net.

2.3 VID-Net: Integration von Gerätedaten

Um den bereitgestellten Datenstrom von chirurgischen Geräten des integrierten OPs zu integrieren, erweitern wir VI-Net um eine weiteren Eingabeschicht (Gerätedaten in Abbildung 1). Die Eingabe besteht aus 14 Werten, die jeweils ein unterschiedliches Signal eines chirurgischen Gerät darstellen (eine Liste der verwendeten Geräte und Signale befindet sich in der Tabelle 2). Analog zum Video wählen wir eine Sekunde als Größe für jeden Zeitschritt aus. Bei Signalen mit einer Datenrate über 1 Hz werden alle Werte, mit Ausnahme des aktuellsten, ignoriert. Für Signale mit einer niedrigeren Datenrate verwenden wir den aktuellsten Wert, auch wenn er älter als eine Sekunde ist.

VID-Net wird auf die gleiche Weise wie die vorherigen zwei Netzwerke trainiert.

3 Evaluation und Ergebnisse

Grundlage unserer Auswertung ist ein Datensatz bestehend aus Aufnahmen von 80 laparoskopischen Eingriffen von unterschiedlichen Operationstypen (**MultiType**). Für jede Operation enthält der Datensatz das endoskopischen Video und die Datenströme der verschiedenen chirurgischen Geräten (Tabelle 2). Die Prozeduren wurden alle im selben OP des Universitätsklinikums Heidelberg mit dem integrierten OP-System OR1 ™(Karl Storz SE & Co. KG, Tuttlingen, Deutschland) aufgezeichnet. Die durchschnittliche Prozedurlänge im Dataset beträgt 105,6 Minuten. Die Datensätze, die für das Vortrainieren von V-Net und für das Trainieren des ResNet-152 für die Erkennung der Instrumentennutzung verwendet wurden, überschneiden sich nicht mit diesem Datensatz.

Geräte	Signal
Insufflator	Ist Gasflussrate
	Soll Gasflussrate
	Ist Gasdruck
	Soll Gasdruck
	Verbrauchtes Gasvolumen
	Druck Gasversorgung
	Gerät an?
OP Lichter	Alle Lichter aus?
	Helligkeit Licht 1
	Helligkeit Licht 2
Endoskopische Lichtquelle	Helligkeit
Endoskopische Kamera	Weißabgleich
	Gains
	Empfindlichkeitsindex

Tabelle 2: Chirurgischen Geräte und genutzte Signale

Um die drei vorgeschlagenen Methoden zu evaluieren teilen wir den Datensatz in vier Sets gleicher Größe auf und führen für jede Methode vier Leave-One-Set-Out-Bewertungen durch. Bei der Aufteilung des Datensatzes sorgen wir dafür, dass die Verteilung der verschiedenen Operationstypen in jedem Set ausgeglichen ist.

Während des Evaluation berechnen wir die Restdauerprädiktion $\tilde{N}_i$ für jeden Videoframe i: $\tilde{N}_i = \frac{i}{y_i}$, wobei $y_i \in [0, 1]$ der vorhergesagte Fortschritt der Prozedur ist. Mit $\tilde{N}_i$ können wir den Vorhersagefehler in Sekunden, $|\tilde{N}_i - N|$, und den Vorhersagefehler relativ zur Länge N jeder Prozedur berechnen, $\frac{|\tilde{N}_i - N|}{N}$. Für jede Methode stellen wir sowohl den absoluten als auch den relativen Vorhersagefehler bereit (siehe Tabellen 3a und 3b). Um zu messen, wie sich die Fehler im Verlauf einer Prozedur entwickeln, berechnen wir die durchschnittlichen Fehler während jedes zeitlichen Viertels einer gegebenen Prozedur (Q1-Q4). Als eine **naive** Baseline geben wir den Vorhersagefehler an, der auftreten würde, wenn die durchschnittliche Prozedurdauer über den Trainingsdaten als Wert für $\tilde{N}_i$ verwendet würde. Wir stellen auch eine auf den Operationstyp basierende Baseline (**Type**) zur Verfügung, wobei wir stattdessen die durchschnittliche Prozedurdauer für jeden Operationstyp in Tabelle 1 separat berechnen. Zur weiteren Auswertung werden V-Net und VI-Net auf den öffentlich zugänglichen

Method	Q1	Q2	Q3	Q4	Mean
V-Net	3818 ± 486	2496 ± 258	1611 ± 230	1353 ± 200	2320 ± 846
VI-Net	4271 ± 433	2421 ± 318	1052 ± 285	1457 ± 322	2300 ± 886
VID-Net	4143 ± 449	2289 ± 312	1071 ± 279	1313 ± 312	2204 ± 875
Baseline (Naive)	3208 ± 2085	3208 ± 2085	3208 ± 2085	3208 ± 2085	3208 ± 2085
Baseline (Type)	2093 ± 1787	2093 ± 1787	2093 ± 1787	2093 ± 1787	2093 ± 1787

(a) Absoluter Fehler (in Sekunden)

Method	Q1	Q2	Q3	Q4	Mean
V-Net	$53\% \pm 9\%$	$42\% \pm 6\%$	$42\% \pm 5\%$	$44\% \pm 4\%$	$45\% \pm 16\%$
VI-Net	$61\% \pm 8\%$	$35\% \pm 6\%$	$25\% \pm 5\%$	$37\% \pm 6\%$	$39\% \pm 15\%$
VID-Net	$59\% \pm 8\%$	$32\% \pm 5\%$	$24\% \pm 5\%$	$35\% \pm 6\%$	$37\% \pm 14\%$
Baseline (Naive)	$124\% \pm 245\%$	$124\% \pm 245\%$	$124\% \pm 245\%$	$124\% \pm 245\%$	$124\% \pm 245\%$
Baseline (Type)	$57\% \pm 99\%$	$57\% \pm 99\%$	$57\% \pm 99\%$	$57\% \pm 99\%$	$57\% \pm 99\%$

(b) Relativer Fehler

Tabelle 3: Die durchschnittlichen Vorhersagefehler auf den **MultiType** Datensatz. (a) zeigt den durchschnittlichen absoluten Fehler aller Methoden und (b) den durchschnittlichen relativen Fehler.

Cholec80-Datensatz [7] angewendet, der aus 80 Videos von laparoskopischen Cholezystektomien besteht. Wir unterteilen den Datensatz in vier Sets gleicher Größe und ähnlicher durchschnittlicher Eingriffslänge und führen für jede Methode vier Leave-One-Set-Out-Bewertungen durch. Der absolute und der relative Vorhersagefehler können in den Tabellen 4a und 4b gefunden werden. Es ist nur eine Baseline verfügbar, da das Dataset aus einem einzigen Prozedurentyp besteht. Im Datensatz sind keine Daten zu chirurgischen Geräten enthalten, sodass

VID-Net nicht verwendet werden kann.

Method	Q1	Q2	Q3	Q4	Mean
V-Net	934 ± 238	804 ± 135	581 ± 119	353 ± 92	668 ± 233
VI-Net	920 ± 147	746 ± 104	434 ± 93	236 ± 80	584 ± 214
Baseline	768 ± 669	768 ± 669	768 ± 669	768 ± 669	768 ± 669

(a) Absoluter Fehler (in Sekunden)

Method	Q1	Q2	Q3	Q4	Mean
V-Net	$42\% \pm 12\%$	$38\% \pm 7\%$	$31\% \pm 6\%$	$22\% \pm 5\%$	$33\% \pm 11\%$
VI-Net	$34\% \pm 7\%$	$28\% \pm 5\%$	$17\% \pm 4\%$	$12\% \pm 4\%$	$23\% \pm 9\%$
Baseline	$66\% \pm 27\%$	$66\% \pm 27\%$	$66\% \pm 27\%$	$66\% \pm 27\%$	$66\% \pm 27\%$

(b) Relativer Fehler

Tabelle 4: Die durchschnittlichen Vorhersagefehler auf den **Cholec80** Datensatz. (a) zeigt den durchschnittlichen absoluten Fehler aller Methoden und (b) den durchschnittlichen relativen Fehler.

Die Ergebnisse von **MultiType** zeigen, dass alle drei Methoden die Baseline übertreffen. Darüber hinaus liefert VID-Net genauere Ergebnisse als die anderen beiden Netzwerke. Dies zeigt, dass Daten aus chirurgischen Geräten tatsächlich komplementäre Informationen enthalten. Tabelle 5 zeigt, dass VID-Net in den ersten drei Verfahrenskategorien konsistent arbeitet. Teil der "Allgemeine Laparoskopie" Kategorie sind diagnostische Laparoskopien, die signifikant kürzer sind (< 15 min) als die durchschnittliche Prozedur im Datensatz, was diesen Leistungsabfall erklärt. Die letzte Kategorie enthält nur Einzelfälle, die sich von den anderen Kategorien stark unterscheiden, was Vorhersagen schwierig macht.

Bei **Cholec80** übertreffen beide Methoden die Baseline, während VI-Net auch V-Net übertrifft. Ein direkter Vergleich mit den Ergebnissen in [9] ist nicht möglich, da die Autoren einen privaten Datensatz zum Testen verwenden.

Type	Q1	Q2	Q3	Q4	Mean
1	6020 ± 961	3452 ± 733	1350 ± 652	1381 ± 710	3050 ± 2136
2	4102 ± 863	2096 ± 541	852 ± 474	1226 ± 527	2069 ± 1460
3	4070 ± 822	2090 ± 510	764 ± 376	1053 ± 475	1994 ± 1483
4	1098 ± 457	496 ± 241	706 ± 246	1270 ± 292	893 ± 572
5	1419 ± 522	707 ± 308	927 ± 256	1518 ± 382	1143 ± 746

(a) Absoluter Fehler (in Sekunden)

Type	Q1	Q2	Q3	Q4	Mean
1	$67\% \pm 12\%$	$37\% \pm 8\%$	$15\% \pm 7\%$	$18\% \pm 8\%$	$34\% \pm 24\%$
2	$62\% \pm 14\%$	$30\% \pm 9\%$	$12\% \pm 7\%$	$22\% \pm 9\%$	$32\% \pm 22\%$
3	$65\% \pm 14\%$	$32\% \pm 8\%$	$12\% \pm 6\%$	$18\% \pm 8\%$	$32\% \pm 24\%$
4	$43\% \pm 20\%$	$25\% \pm 12\%$	$45\% \pm 13\%$	$72\% \pm 14\%$	$46\% \pm 28\%$
5	$45\% \pm 21\%$	$26\% \pm 13\%$	$46\% \pm 11\%$	$71\% \pm 15\%$	$47\% \pm 30\%$

(b) Relativer Fehler

Tabelle 5: Die durchschnittlichen Vorhersagefehler von VID-Net auf den **MultiType** Datensatz, aufgeteilt nach Operationstyp. (a) zeigt den durchschnittlichen absoluten Fehler und (b) den durchschnittlichen relativen Fehler.

4 Diskussion

In dieser Arbeit haben wir, nach unserem Kenntnisstand, den ersten Ansatz zur Online Vorhersage der Restdauer unter Verwendung nicht annotierter, endoskopischer Videodaten und chirurgischer Gerätedaten in einer laparoskopischen Umgebung vorgestellt. Bei **MultiType** zeigte VID-Net einen durchschnittlichen Vorhersagefehler von 37% und einen Halbzeitfehler von etwa 28%, welche niedriger sind als die Ergebnisse der Baselines. Darüber hinaus haben wir gezeigt, dass eine Methode mit Bild- und Gerätedaten besser abschneidet als nur mit Bilddaten alleine. Die Auswertung zeigt, dass die vorgestellten Methoden bei Prozeduren mit kurzer Länge

(kürzer als 15 min) und in Einzelfällen derzeit überdurchschnittlich große Fehler aufweisen. Wir gehen davon aus, dass dies auf mangelnde Trainingsdaten zurückzuführen ist, da unser Datensatz hauptsächlich längere Operationen umfasst.

Auf **Cholec80** erzielte VI-Net einen durchschnittlichen Vorhersagefehler sowie einen Halbzeitfehler von 23% und übertraf damit die Baseline. Dieser Leistungsunterschied in den beiden Datensätzen kann dadurch erklärt werden, dass **MultiType** mehrere Prozedurtypen enthielt und die Eingriffe auch eine größere Varianz in der Dauer aufwiesen.

Da unsere Ergebnisse zeigen, dass die Kombination von Bild- und Gerätedaten mehr Informationen über den Fortschritt der Operation liefert, gehen wir davon aus, dass Daten, die dem Anästhesisten zur Verfügung stehen, wie Herzfrequenz, Blutdruck und Medikamentendosierungen, zusätzliche Verbesserungen bringen könnten.

Literatur

[1] T. Blum, H. Feußner, and N. Navab, Modeling and Segmentation of Surgical Workflow from Laparoscopic Video, pp. 400–407. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[2] O. Dergachyova, D. Bouget, A. Huaulmé, X. Morandi, and P. Jannin, "Automatic data-driven real-time segmentation and recognition of surgical workflow," International Journal of Computer Assisted Radiology and Surgery, vol. 11, no. 6, pp. 1081–1089, 2016.

[3] D. Katić, A.-L. Wekerle, F. Gärtner, H. Kenngott, B. P. Müller-Stich, R. Dillmann, and S. Speidel, Knowledge-Driven Formalization of Laparoscopic Surgeries for Rule-Based Intraoperative Context-Aware Assistance, pp. 158–167. Cham: Springer International Publishing, 2014.

[4] N. Padoy, T. Blum, S.-A. Ahmadi, H. Feussner, M.-O. Berger, and N. Navab, "Statistical modeling and recognition of surgical workflow," Medical Image Analysis, vol. 16, no. 3, pp. 632 – 641, 2012. Computer Assisted Interventions.

[5] S. Bodenstedt, M. Wagner, D. Katić, P. Mietkowski, B. Mayer, H. Kenngott, B. Müller-Stich, R. Dillmann, and S. Speidel, "Unsupervised temporal context learning using convolutional neural networks for laparoscopic workflow analysis," ArXiv e-prints, Feb. 2017.

[6] C. Lea, J. H. Choi, A. Reiter, and G. Hager, "Surgical phase recognition: From instrumented ors to hospitals around the world," M2CAI 2016, 2016.

[7] A. P. Twinanda, S. Shehata, D. Mutter, J. Marescaux, M. de Mathelin, and N. Padoy, "Endonet: A deep architecture for recognition tasks on laparoscopic videos," IEEE transactions on medical imaging, vol. 36, no. 1, pp. 86–97, 2017.

[8] A. C. P. Guédon, M. Paalvast, F. C. Meeuwsen, D. M. J. Tax, A. P. van Dijke, L. S. G. L. Wauben, M. van der Elst, J. Dankelman, and J. J. van den Dobbelsteen, "'it is time to prepare the next patient' real-time prediction of procedure duration in laparoscopic cholecystectomies," Journal of Medical Systems, vol. 40, no. 12, p. 271, 2016.

[9] I. Aksamentov, A. P. Twinanda, D. Mutter, J. Marescaux, and N. Padoy, "Deep neural networks predict remaining surgery duration from cholecystectomy videos," in International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 586–593, Springer, 2017.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, pp. 1097–1105, 2012.

[11] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," CoRR, vol. abs/1409.1259, 2014.

[12] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.

HeiKo – der offene Heidelberger Datensatz kolorektaler Operationen als neuer Benchmark für die chirurgische Workflow-Analyse

M. Wagner[1], S. Bodenstedt[2], L.B. Raedeker[1], M. Apitz[1], L. Maier-Hein[3], S. Speidel[2], B.P. Müller[1], H.G. Kenngott[1]

[1]Universitätsklinikum Heidelberg, Klinik für Allgemein-, Viszeral- und Transplantationschirurgie, Heidelberg, Deutschland
[2]Nationales Centrum für Tumorerkrankungen, Translationale Chirurgische Onkologie, Dresden, Deutschland
[3]Deutsches Krebsforschungszentrum, Computerassistierte Medizinische Interventionen, Heidelberg, Deutschland

Kontakt: hannes.kenngott@med.uni-heidelberg.de

Abstract

Kontext-Sensitivität bildet die Grundlage neuartiger Assistenzsysteme in der Chirurgie, beispielsweise (teil)autonomer Chirurgieroboter. Die Forschung an chirurgischer Workflowanalyse ist allerdings limitiert durch die geringe Zahl öffentlich verfügbarer Datensätze. In dieser Arbeit stellen wir daher einen neuen Datensatz von 30 kolorektalen Operationen vor, die in jeweils bis zu dreizehn Phasen eingeteilt wurden. Insgesamt wurden 96 Stunden Videodaten und fast 3 Millionen Sensordatenpunkte chirurgischer Medizingeräte (z.B. Insufflator) aufgenommen und mit 626 Phasenübergängen annotiert. Dieser Datensatz wurde zudem in einem internationalen Programmierwettbewerb genutzt. Die Ergebnisse mit einer Genauigkeit der Phasenerkennung zwischen 21 und 61% zeigen auf, dass chirurgische Workflowanalyse noch nicht gelöst ist. Unser Datensatz bildet damit einen neuen Benchmark, an dem sich Algorithmen für die Workflowanalyse messen können.

Keywords: Kontext-Sensitive Assistenzsysteme, Chirurgische Workflowanalyse, Laparoskopische Chirurgie, Kolorektalchirurgie, Open Data

1 Problemstellung

Kontext-Sensitivität und Workflowanalyse sind entscheidende Grundlage für die nächste Generation chirurgischer Roboter. Kognitive chirurgische Roboter werden aus Erfahrung lernen und komplexe Aufgaben (teil-)autonom durchführen. Allerdings ist die Forschung zu chirurgischer Prozessanalyse aktuell limitiert durch den Mangel an qualitativen klinischen Datensätzen. Um dieses Problem zu lösen, haben wir einen Datensatz laparoskopischer, kolorektaler Operationen erstellt und getestet, um einen neuen Benchmark für Algorithmen zur Workflowanalyse zu erstellen.

2 Material und Methoden

In einem ersten Schritt wurden am Universitätsklinikum Heidelberg laparoskopische Videos zusammen mit Daten chirurgischer Medizingeräte (Insufflator, OP-Licht, Kamera-Einstellungen) aufgezeichnet. Die Daten wurden in einem integrierten Operationssaal der Firma KarlStorz (Tuttlingen, Deutschland) aufgezeichnet. Bei der laparoskopischen Kamera handelte es sich um eine 2D-Kamera mit 30° Optik. Die Videos wurden mit einer Auflösung von 1920x1080 Bildpunkten und 25 Bildern pro Sekunde aufgenommen. Video- und Gerätedaten wurden durch Zeitstempel im Aufnahmerechner synchronisiert. Insgesamt 30 Operationen wurden ausgewählt, die sich aus jeweils zehn Eingriffen der Sigmaresektion (SR), Rektumresektion (RR) und Proktokolektomie (PK) zusammensetzten. Aus Datenschutzgründen erfolgte keine Auswertung der Operationen unterschieden nach Operateuren.

Anschließend annotierte ein chirurgischer Experte jedes Video, indem er jedem Einzelbild eine von dreizehn chirurgischen Phasen zuordnete (Vorbereitung, Mobilisation des Sigma, Dissektion des Rektums etc.).

In einem dritten Schritt wurde der gesamte Datensatz im Rahmen der internationalen Endoscopic Vision Challenge [1] von vier unabhängigen Forschungsgruppen getestet. Hier wurden 24 Operationen als Trainingsdatensätze und 6 Operationen als Testdatensätze genutzt, um die chirurgischen Phasen zu erkennen. Die Forschungsgruppen verwendeten verschiedene Kombinationen neuronaler Netze (u.a. Convolutional Neural Networks, Recurrent Neural Networks). Als Metriken wurden die Genauigkeit (Anteil korrekter Klassifikationen) und der Jaccard Index (JI, Überlappung zwischen Algorithmischer Klassifikation und chirurgischer Annotation geteilt durch ihre Vereinigung) gewählt.

3 Ergebnisse

Insgesamt 96 Stunden Video und ca. 2.97 Millionen Sensordatenpunkte wurden gesammelt und mit 626 Phasenübergängen annotiert. Jede Phase dauerte durchschnittlich 18.9±16.8 min. Die längste Phase in SR war „Mobilisation des Sigma“ mit 42.8±24.1 min. Dissektion des Rektums war die längste Phase in RR (45.4±21.3 min) und PK (51.2±9.7 min). Die zweitlängste Phase war in allen Operationstypen die extraabdominelle Vorbereitung der Anastomose mit 36.0±11.3 min für SR, 36.9±15.0 min für RR und 44.2±7.18 min für PK. Die Algorithmen zur Workflowanalyse erreichten zwischen 21% Genauigkeit mit einem durchschnittlichen JI von 8% (schwächster Algorithmus) und 61% Genauigkeit mit einem durchschnittlichen JI von 40% (stärkster Algorithmus). Allerdings schwankte der JI selbst beim stärksten Algorithmus zwischen 18% und 78% je nach Test-Operation.

4 Diskussion und Zusammenfassung

Der vorgelegte Datensatz stellt offensichtlich noch eine große Herausforderung für aktuelle Algorithmen zur Workflowanalyse dar. Mit dreißig Datensätzen handelt es sich um einen vergleichsweise großen Datensatz für die Chirurgie. Allerdings werden größere Datensätze für weitergehende Forschung benötigt. Weiterhin ist die Standardisierung und Reproduzierbarkeit der Annotationen notwendig, um standortübergreifende Zusammenarbeit zu ermöglichen. Mögliche Ansatzpunkte bilden hier Ontologien für die Annotation [2]. Kritisch ist zu bemerkten, dass unsere Annotationen nur von einem einzigen Chirurgen durchgeführt wurden und die intra- und inter-Observer-Variabilität nicht untersucht wurde.

Zusammenfassend zeigen unsere Ergebnisse, dass das Problem der Workflowanalyse durch aktuelle Algorithmen adressiert werden kann, aber noch nicht gelöst ist.

Referenzen

[1] EndoVisSub2017-Workflow - Home (o. J.)., URL: https://endovissub2017-workflow.grand-challenge.org/ [as of June 2018].

[2] Katić, D., Julliard, C., Wekerle, A.-L., Kenngott, H., Müller-Stich, B. P., Dillmann, R., Speidel, S., Jannin, P. und Gibaud, B. *LapOntoSPM: an ontology for laparoscopic surgeries and its application to surgical phase recognition*. Int J Comput Assist Radiol Surg *10*, 1427–1434 (2015)

Simulating the Patient-Individual Intervention Process for Staged Segmental Artery Occlusion - A Modeling Approach

R.Bieck[1], E. Schreiber[1], T. Neumuth[1]

[1]Innovation Center Computer Assisted Surgery, Universität Leipzig, Germany

Contact: richard.bieck@iccas.de

Abstract

Patients suffering from large thoracoabdominal aortic aneurysms (TAAA) are confronted with the need for substantial aortic repair. The risks of such an intervention are neurological pathologies, e. g. Paraplegia. The minimally invasive, selective segmental artery coil embolisation (MISACE) is a procedure employed to preemptively close supplying segmental arteries of the aorta to promote collateral perfusion network arteriogenesis before such a conservative procedure. Since the MISACE procedure is still in an early application phase there exist no general guidelines for artery closing patterns, the impact on the paraspinal collateral network perfusion compensation mechanics or the corresponding risk of ischemia and paraplegia. We aim to establish an understanding of the coiling patterns and their impact on spinal perfusion and the clinical outcome from a computational perspective. A modelling approach combining the patient-specific situation with an interventional process representation is introduced. A prediction method to estimate patient-individual consequences of coiling patterns on the overall procedure risk and the minimal recuperation time follows the model development.

Keywords: TAAA, Decision Support, Patient Modeling, Process Modeling

1 Problem

Large thoracoabdominal aortic aneurysms (TAAA) are substantial and complex forms of a permanent dilation of the aortic vessel wall mainly caused by atherosclerosis-based processes as well as autoimmune and infectious conditions. The extent of the TAAA is identified using the Crawford classification (I to IV). A repair of such a widespread pathology is preferably done using an endovascular minimally-invasive approach with the application of a stent graft. However, a successful intervention still poses a substantial risk to the patient in the form of ischaemic reactions in the spinal cord leading to paraplegia or death [1]. The minimally invasive, selective segmental artery coil embolization (MISACE) is a procedure employed to preemptively close supplying segmental arteries (Fig. 1) of the aorta to promote collateral perfusion network arteriogenesis before

Figure 1: Example for segmental arterie coilin (left) and different coiling patterns (right)

such a conservative procedure [2]. The procedure was then used to form the concept of “staged repair”, where segmental arteries are deliberately occluded in subsequent steps with sufficient time intervals to allow for structural changes within the arterial collateral network. This staging approach aims to secure spinal cord perfusion and, more importantly, eliminate ischaemic reactions following aortic repair. Since the MISACE procedure is still in an early application phase there, exist no general guidelines for segmental artery closing

patterns and their respective influence on the clinical outcome [3]. Furthermore, the impact on the paraspinal collateral network perfusion compensation mechanics and the corresponding risk of ischemia and paraplegia are unknown. The effectiveness of the MISACE procedure is currently being investigated in a multi-centric clinical trial in the *PAPAartis* EU project [4]. In the project, patients with Crawford type II and III will be pre-treated with the MISACE followed by TAAA (open or endovascular). Our goal is to develop a patient and an intervention process hybrid model to investigate the influence of varying coiling patterns and their impact on the convalescence of spinal perfusion and the clinical outcome using computational modelling strategies.

2 Material and Methods

Initially, an intervention modelling approach had to be identified that corresponds with the overall project trial setup and also covers all relevant aspects of the intended treatment process. Multimodal patient-specific information is acquired in various stages including before, during and after the MISACE staging as well as one month and twelve months post-operatively after aneurysm repair. In response to the increasing amount of patient-individual data and its consideration in the treatment process, the introduction of a software systems design for intelligent decision-support in the clinical workflow of MISACE planning is intended. An overall limitation is set by the machine-readability of such highly heterogeneous patient data, as well as the inference of relevant information. The extraction of expert knowledge from the respective medical domains is laborious and often counter-productive to the overall concept of assistance systems. Therefore, we chose the approach of digitally modelling the patient as well as the perioperative therapy processes. Using this approach, a software-internal representation of the individual patient state is generated at different time points during the treatment process. Investigating the temporal changes of specific patient data (vital, pathological, and procedure parameters) leads to the abstraction of more generalised representation of the treatment process and, furthermore, enables the comparison of new patients with this treatment representation. In a first step, the model components and relevant parameters were identified. With the start of the intervention phase in the PAPAartis project, the model components will be instantiated and key patient and therapy process model parameters identified which correspond to the MISACE staging in a meaningful way.

3 Results

The overall modeling approach is divided into 5 stages (Fig. 2). We initially identified the components of a patient model representation including parameters, e.g. demographic background, imaging data, anamnesis, and diagnosis. The individual patient data gathered from the trial is stored in a knowledge base. Patient-specific parameters are extracted, categorised and stored in an overall patient representation. From the knowledge base, potential treatment-relevant parameter candidates are identified by investigating patient clusters and corresponding intervention trajectories. For this, previously recorded treatment data and patient-specific characteristics from the trial are mapped for each patient and subsequently analysed by utilizing tailored similarity metrics for patient characteristics and treatment processes. The coherence revealed as a result of this analysis are potential causal connections and can be used to predict treatment processes, risks, recuperation time,

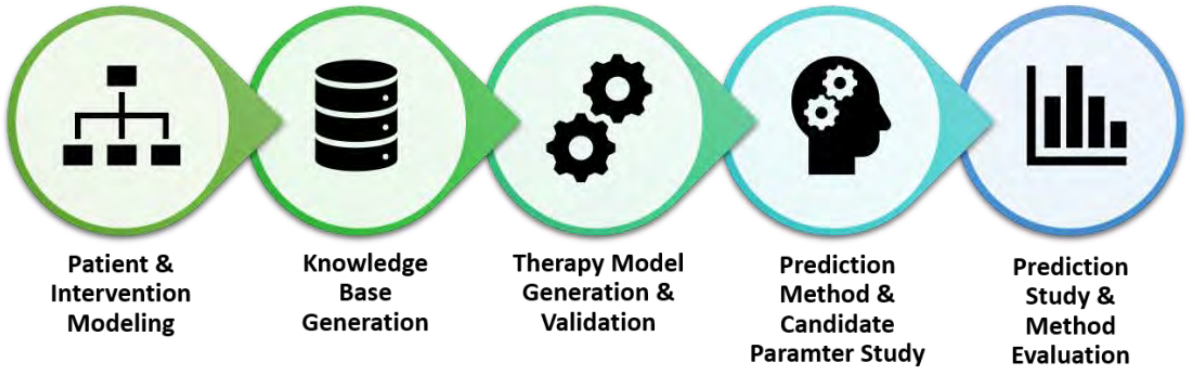


Figure 2: Stages of the modeling approach

best treatment as well as a patient outcome for each patient. Subsequently, the MISACE procedure is analysed using modelling approaches from previous works [4-7]. Workflow and process analysis steps then allow the definition of a representative process model including all relevant surgical actions executed for the various coiling configurations. Both the patient and the process model are then used to form a therapy model that is

instantiated at each time point of the trail, where patient-specific diagnostics are performed. The development of an estimation method to identify the ideal MISACE stage configuration is based on the highest correlation of parameters between patient models and process models at these time points. For example, patient characteristics that have been observed to likely cause high risk or poor outcome (e.g. for paralysis, extraordinary long recuperation time etc.) in combination with specific stage configurations, will cause an adaption of process model and treatment to minimize the risk for the patient, maximize patient outcome and thus contribute in choosing the optimal treatment plan for each individual patient.

4 Discussion

In this work, we introduced our modelling approach for the simulation of the patient-individual intervention process for staged segmental artery occlusion as performed in a staged MISACE procedure. The goal is to establish a model representation that enables the investigation of different coiling strategies and their direct influence on the paraspinal collateral network perfusion and the ischaemic reaction of the patient. Taking into consideration the multi-modal data, e.g. imaging & procedure parameters as well as cardiovascular and pathology parameters are challenging not only regarding complexity and resources needed. The modelling approaches for the patient situation and the MISACE process necessitate abstraction and simplification steps which could potentially lower the overall expressiveness of the overall model. Furthermore, the inter-patient range of parameter values and responses to the treatment could negatively impact computational methods optimised for convergence. Close attention must be paid to the meaningfulness of the selected model parameters and their validation in potential decision-support functionality.

5 Conclusion

The introduction of software systems for intelligent decision-support in the clinical workflow of interventional procedures is primarily a response to the increasing amount of patient-individual data and its consideration in the treatment process. However, the machine-readability of such patient data, as well as the inference of relevant information, is still restricted. The extraction of expert knowledge from the respective medical domains is still resource-consuming. The approach of digitally modelling the patient as well as the perioperative therapy processes are offering a way, to generate a software-internal representation of the individual patient state at different time points during the treatment process based on the currently specified domain knowledge and model aspects. Investigating the temporal changes of specific patient data (vital, pathological, and procedure parameters) could lead to the abstraction of a more generalised representation of the treatment process and, furthermore, enables the comparison of new patients with this treatment representation.

References

[1] Etz CD, Luehr M, Kari FA, Bodian CA, Smego D, Plestis KA, Griepp RB. J Thorac Cardiovasc Surg. 2008a Feb;135(2):324-30. Paraplegia after extensive thoracic and thoracoabdominal aortic aneurysm repair: does critical spinal cord ischemia occur postoperatively?

[2] Etz, Christian D.; Zoli, Stefano; Mueller, Christoph S.; Bodian, Carol A.; Di Luozzo, Gabriele; Lazala, Ricardo et al. (2010): Staged repair significantly reduces paraplegia rate after extensive thoracoabdominal aortic aneurysm repair. In The Journal of thoracic and cardiovascular surgery 139 (6), pp. 1464–1472. DOI: 10.1016/j.jtcvs.2010.02.037.

[3] Etz, Christian D.; Debus, E. Sebastian; Mohr, Friedrich-Wilhelm; Kolbel, Tilo (2015a): First in-man endovascular preconditioning of the paraspinal collateral network by segmental artery coil embolization to prevent ischemic spinal cord injury. In The Journal of thoracic and cardiovascular surgery 149 (4), pp. 1074–1079. DOI: 10.1016/j.jtcvs.2014.12.025.

[4] https://www.papa-artis.eu/

[5] Neumuth T, , P. Jannin, J. Schlomberg, J. Meixensberger, P. Wiedemann, und O. Burgert, „Analysis of surgical intervention populations using generic surgical process models“, Int J Comput Assist Radiol Surg, Bd. 6, Nr. 1, S. 59–71, Jan. 2011. (a)

[6] Neumuth T, P. Jannin, G. Strauss, J. Meixensberger, und O. Burgert, „Validation of Knowledge Acquisition for Surgical Process Models“, J Am Med Inform Assoc, Bd. 16, Nr. 1, S. 72–80, 2009.

[7] Neumuth D, F. Loebe, H. Herre, und T. Neumuth, „Modeling surgical processes: a four-level translational approach“, Artif Intell Med, Bd. 51, Nr. 3, S. 147–161, März 2011. (b)

[8] Cypko MA, Stoehr M, Kozniewski M, Druzdzel M, Dietz A, Lemke HU. Validation workflow for a clinical Bayesian network model in multidisciplinary decision making in head and neck oncology treatment. Int J Comput Assist Radiol Surg [Internet]. 2017 Feb;

Modellierung und Ausführung klinischer Prozesse in der molekularen Intervention

U. Thiel, S. Knoch, S. Schöning, J. Stallkamp

Fraunhofer-Projektgruppe für Automatisierung in Medizin und Biotechnologie (Fraunhofer PAMB), Mannheim, Deutschland

Kontakt: ulrich.thiel@ipa.fraunhofer.de

Abstract

Im Projekt M²PROCIT (Mannheim Molecular Process Information Technology) wird eine Prozessmanagementplattform realisiert, mit der die Prozesse vor, während und nach einer molekularen Intervention und der zugehörige Ressourceneinsatz geplant und gesteuert werden können. Die einheitliche Erfassung der Patientendaten und der Prozessparameter im Interventionsraum ermöglicht eine Qualitätskontrolle und Optimierung der Prozessstrukturen. Die Ansätze zur Modellierung und Ausführung der klinischen Prozesse werden skizziert und diskutiert.

Keywords: klinische Prozesse, molekulare Intervention, Prozessmanagement

1 Problemstellung

Im Forschungscampus „Mannheim Molecular Intervention Environment (M²OLIE)“ wird in interdisziplinärer Zusammenarbeit von Medizinern, Naturwissenschaftlern, Ingenieuren und Betriebswirtschaftlern aus fünf akademischen (Hochschulen und nicht-universitären Forschungseinrichtungen) und neun industriellen Partnern (Großunternehmen und kleine und mittlere Unternehmen) im Mannheimer Universitätsklinikum eine medizinische Interventionsumgebung für die Weiterentwicklung von minimal-invasiven Krebstherapien entwickelt. Im Laufe dieses Entwicklungsprozesses sollen Therapieverfahren aus den Bereichen Radiologie, Nuklearmedizin und Strahlentherapie durch modernisierte Bildgebungsverfahren, patientenindividuelle Radiopharmaka, roboterbasierende Interventionsassistenten und viele weitere personalisierte Therapieverfahren verbessert und ergänzt werden.

Am Behandlungsfall oligometastasierter Patienten werden minimalinvasive Diagnose und Therapie in einem geregelten Prozess, kurz Closed-Loop-Prozess, mit sehr kurzen Durchlaufzeiten als One-Stop-Shop umgesetzt. Das folgende Diagramm zeigt die wesentlichen Stationen des Prozesses und die zugeordneten Teilprojekte des Forschungscampus.

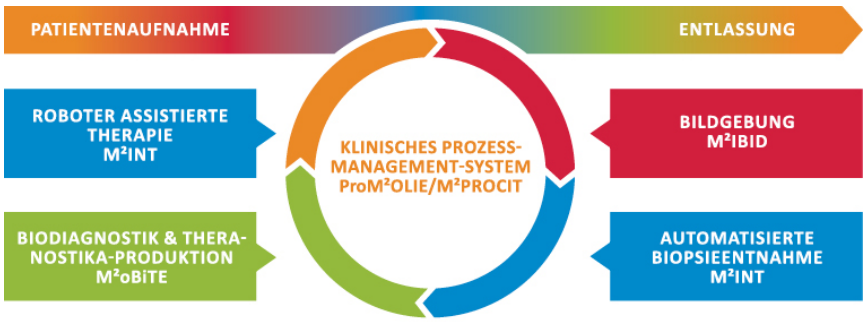


Abbildung 1*: Visualisierung des Closed-Loop-Ansatz und dazugehöriger Teilprojekte im Forschungscampus M²OLIE [1].*

Die Kombination verschiedener bildgebender Verfahren (3-Tesla-MR-Tomographen, PET-Scanner, CT-Scanner SOMATOM Force) ermöglicht die Planung der teilautonomen Navigation der robotischen Assistenzsysteme (Artis ZEEGO® System und Leichtbauroboter) im Interventionsraum und der Biopsieentnahme. Die Proben gelangen direkt zur Bioanalyse mit Verfahren der Massenspektrometrie. Die Resultate dienen dem Tumorboard neben den Patientendaten als weitere Entscheidungsgrundlage und stellen die Basis der radio-chemischen

Produktion von fallspezifischen Diagnostika und Therapeutika bereit, die unter Verwendung der Assistenzsysteme zeitnah minimalinvasiv appliziert werden.

Eine effiziente, patientenorientierte Behandlung in einem One-Stop-Shop erfordert neben der Bereitstellung der technischen, personellen und räumlichen Ressourcen auch eine möglichst exakte Planung der Schritte und deren Abfolge während der Behandlung. Das medizinisch notwendige Vorgehen – üblicherweise als Entscheidungsstruktur in Form eines "klinischen Pfades" erfasst – wird ergänzt durch situative Entscheidungen, die sich sowohl aus den spezifischen Patientendaten als auch den technischen und organisatorischen Voraussetzungen der Behandlung ergeben. Das Resultat sind Spezifikationen "klinischer Prozesse", die in einer formalen Notation vorliegen, die von einem Prozessmanagementsystem verarbeitet und zur Steuerung der Behandlung im Interventionsraum dienen, wobei den beteiligten Akteuren (medizinisches und technisches Personal) für ihre aktuellen Aufgaben zugeschnittene Informationsangebote zusammengestellt werden. Die Modellierung der klinischen Prozesse ist darüber hinaus für eine optimale Ressourcennutzung erforderlich, die eine Vorbedingung für den Einsatz der aufwändigen molekularen Techniken in der Praxis ist. Daher wird das Prozessmananagementsystem um Module für das Ressourcenmanagement, die optimale Planung der Reihenfolge von anstehenden Behandlungen ("Scheduling") und den Zugriff auf diverse Informationsquellen (elektronische Patientenakte, Radiologie, Labor etc) ergänzt. Weiterhin werden alle Prozessparameter während der Durchführung erfasst und in einem uniformen Format archiviert. Die Integration dieser Komponenten resultiert in einer Prozessmanagementplattform für das Management klinischer Prozesse in der molekularen Intervention, mit der die erforderlichen Ressourcen der diagnostischen und therapeutischen Verfahren dynamisch koordiniert werden können.

Im Folgenden stellen wir die Ansätze und Methoden zur Modellierung der klinischen Prozesse vor, die im Projekt M²ProcIT (Mannheim Molecular Process Information Technology) eingesetzt werden. Das Projekt ist als Querschnittsprojekt im Forschungscampus M²OLIE zuständig für die Bereitstellung einer Prozessplattform für die molekulare Intervention.

2 Material und Methoden

Prozesse sind in unserem Kontext Abläufe, die als Folgen von seriellen und z.T. parallelen Arbeitsschritten definiert werden. Diese Abläufe können in verschiedenen Aktivitätsbereichen erfolgen, so sind administrative Prozesse Modelle der Tätigkeiten der Krankenhausverwaltung, die z.B. die Patientenverwaltung von der Aufnahme bis zur Entlassung und Nachsorge umfasst. Weitere administrative Prozesse betreffen die Logistik der Ressourcen, wie z.B. Personalplanung, Raumplanung und Materialbeschaffung und -bevorratung.

Biologisch-medizinische Prozesse betreffen Aktivitäten zur Untersuchung und Behandlung des Patienten, von der Anamnese und Diagnose zur Behandlung. Ein solcher Prozess kann als (teilweise) Ausführung eines der medizinischen Schritte des administrativen Prozesses angesehen werden, der den Krankenhausaufenthalt eines Patienten strukturiert. So ist z.B. ein Biopsie-Prozess ein möglicher Teil der Ausführung des Diagnoseschrittes. Technische Prozesse betreffen mit (semi-) autonomen Geräten durchgeführte Eingriffe, die wiederum als Teilprozesse der biologisch-medizinischen Prozesse durchgeführt werden.

Eine **exakte Beschreibung der Prozesse** ist die Voraussetzung der weitergehenden informationstechnischen Nutzung und Bearbeitung der Prozesse. Zur Modellierung von Prozessen gibt es eine Reihe von (semi-) formalen Beschreibungssprachen. In Abhängigkeit vom gewählten Abstraktionsgrad der Modellierung und der ihrer Zielsetzung stehen verschiedene Methoden zur Auswahl. Die Abhängigkeiten zwischen Entscheidungen während der Durchführung des Prozesses werden mittels Entscheidungsbäumen oder Klinischen Pfaden [2] dargestellt. Die Dynamik der Prozessausführung kann durch Petrinetz-Modelle[3] analysiert werden, in denen verschiedenen Zuständen des Prozesses Knoten eines Netzwerks zugeordnet sind. Aktive Zustände sind durch zugeordnete „Token" gekennzeichnet, ein Folgezustand wird erreicht, wenn alle Voraussetzungen erfüllt sind, d.h. die Vorgängerzustände alle mit Token belegt sind. Durch die Analyse dieser Struktur (oder einer Simulation der Tokenverschiebungen) können Defizite in der Modellierung, z.B. Engpässe, oder wegen inkonsistenten Eingangsbedingungen nicht erreichbare Teilprozesse, bereits in der Modellierungsphase erkannt werden.

Zur **computergestützten Durchführung** benötigt man neben einem formalen Modell des Prozesses, seiner Teilaktivitäten und ihrer Beziehungen zueinander eine Plattform, die das Modell interpretiert, die Schritte der Prozessausführung vorgibt, den Beteiligten Alternativen zur Auswahl stellt, die dafür benötigten Daten bereitstellt, den Prozess gemäß den Entscheidungen der Akteure fortführt und die Prozessparameter laufend erfasst und protokolliert. Solche Systeme sind seit langer Zeit im Einsatz, um Geschäftsprozesse zu strukturieren und deren Ablauf zu steuern. In diesem Bereich gibt es eine normierte

Prozessmodellierungsprache, die Business-Process-Management-Notation (BPMN), für die eine Reihe von Ausführungsplattformen am Markt verfügbar ist[4].

In der letzten Dekade wurde der Prozessansatz zunächst in der amerikanischen Literatur diskutiert (vgl. z.B. [5]). Erste Entwicklungen in Deutschland legten den Schwerpunkt auf Vereinheitlichung und Integration der Abläufe[6]. Spätere Projekte erweiterten den Fokus um eine situationsadäquate Präsentation von Patientendaten [7]. Aktuelle Projekte streben die Kombination relevanter Daten aus der Patientenakte mit während der Intervention erhobenen Bilddaten an, um das OP-Team optimal zu unterstützen [8]. In M²OLIE dient die dynamische Bildgebung darüber hinaus zur Steuerung und Kontrolle der roboterbasierten Nadelplatzierung.

Ein wichtiger Aspekt der klinischen Prozesse sind die benötigten **Ressourcen**. Dies sind sowohl dem Prozess zugeordnete Personen und Räume als auch die technische Ausstattung des Interventionsraumes. Diese Ressourcen werden in einem der Plattform zugehörigen Ressourcenmanager verwaltet, der zu jeder Ressourcenart die vorhandenen, die bereits eingesetzten oder verplanten und die noch verfügbaren Quantitäten ermittelt. Die Repräsentation der Ressourcen erfolgt als Klassen in einer Ressourcenontologie (vgl. [9]), die aus den initialen Modellen der klinischen Prozesse der molekularen Intervention abgeleitet wurde. Verschiedenen Ressourcenarten werden durch ihre unterschiedlichen Attribute differenziert, sie werden aber in einem einheitlichen Format darstellt. So sind begriffliche Beziehungen zwischen Ressourcenklassen aus dem Modell ableitbar, und das Ressourcenmodell kann durch Hinzufügen neuer Klassen mit differenzierten Attributen ergänzt werden (vgl. Abb.2).

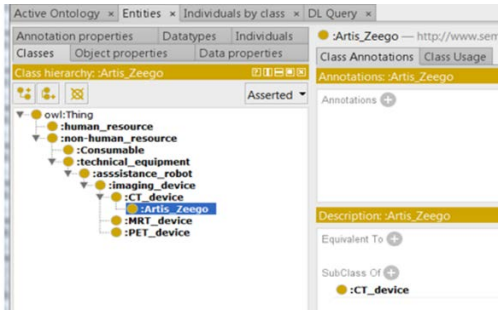


Abbildung 2*: Darstellung des Artis Zeego in der Ressourcenontologie*

Die Verfügbarkeit von Ressourcen bestimmt die Ausführbarkeit von klinischen Prozessen, so kann mit dem Ressourcenmanagement festgestellt werden, ob für einen geplanten Prozess in einem Zeitfenster die Ressourcen gegeben sind (Ressourcenmanagement) und in welcher Reihenfolge die Prozesse (teilweise parallel) auszuführen sind, um eine gute Auslastung der Ressourcen zu gewährleisten.

Ein weiterer Schwerpunkt des Projektes M²ProcIT liegt in der **Informationsbereitstellung** für die Akteure in klinischen Prozessen. Hierzu werden die Informationsbedarfe der Beteiligten in jeder Phase des Prozesses erfasst und entsprechende Schnittstellen in HTML5 definiert, der Datenabruf erfolgt über HL7 und DICOM Schnittstellen zu den angeschlossenen klinischen Informationssystemen.

Die **Architektur der Prozessplattform** ist in der Abb. 3 dargestellt. Die Basis des Prozessmanagers ist ein Business Process Manager der Fa. IBM. Dieses System wird durch den M²OLIE-Prozessmanager erweitert, der einerseits vorhandene Datenhaltungssysteme des Krankenhauses einbindet (bspw. Krankenhausinformationssysteme KIS) und andererseits die Prozessinteraktion über graphische Bedienschnittstellen („Benutzerclients“) realisiert. Weiterhin sind dem Prozessmanager Komponenten zur Ressourcenverwaltung sowie der Ablaufplanung beiseite gestellt.

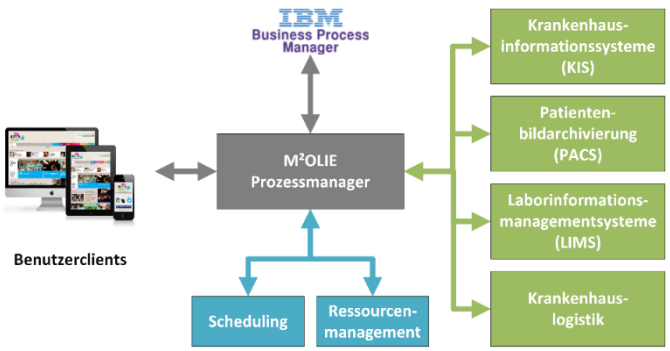


Abbildung 3: Softwarearchitektur der Prozessplattform und ihrer Komponenten.

Bei der Ausführung der Prozesse insbesondere im Interventionsraum können vielfältige **Prozessparameter** erhoben werden, die den Ablauf dokumentieren und somit eine Basis für eine Qualitätskontrolle darstellen, aber auch zur späteren Prozessoptimierung herangezogen werden können. Um die Archivfunktion zu gewährleisten und die nachfolgende Analyse zu ermöglichen, ist es notwendig, die vielfältigen Datentypen zu integrieren. Im Projekt M²ProcIT nutzen wir dazu einen Ansatz, der in einem weiteren Projekt gemeinsam mit der QIT GmbH entwickelt wurde.

Das Verfahren erfasst verschiedene Datenströme aus dem Interventionsraum – von den Eingaben der Prozessbeteiligten (Mediziner, MTAs, Techniker usw.), den bildgebenden Verfahren, bis zu Sensordaten und Verlaufsdaten der technischen Geräte – in Logfiles, aus denen DICOM-Objekte erzeugt werden, die in einem geeigneten Archivsystem gespeichert werden können. Dazu wurden im Projekt neue DICOM-Klassen spezifiziert, die z.B. den zeitlichen Verlauf von Sensordaten der Roboter repräsentieren, und eine Speicherung in einem PACS der QIT GmbH experimentell durchgeführt, das den erweiterten DICOM-Standard unterstützt (vgl. Abb. 4).

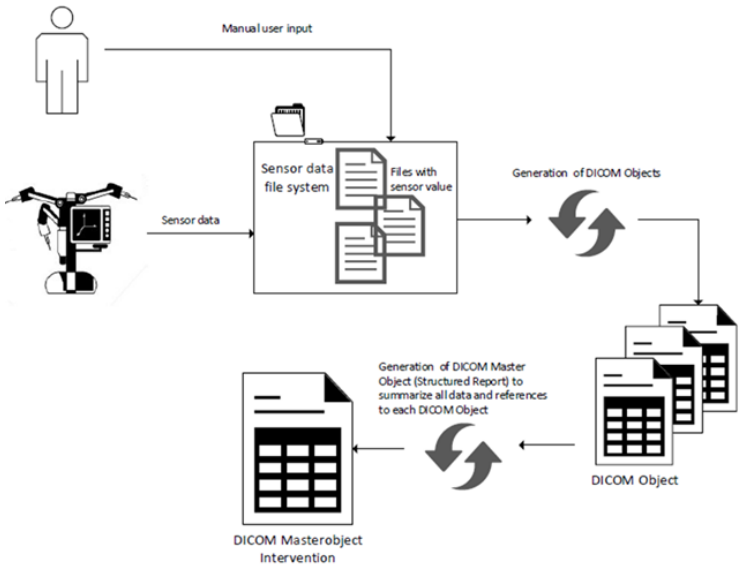


Abbildung 4: *Erfassung von Prozessparametern in DICOM-Objekten*

3 Ergebnisse

Erste Modelle der Teilprozesse des Molekularen Interventionsprozesses wurden vom Projektpartner Unversität Mannheim in Kooperation mit Onkologen der Universitätsmedizin Mannheim (UMM) konzipiert und in Petrinetzmodelle überführt[3]. Für die Ausführung im Prozessmanager wurden hieraus BPMN-Modelle abgeleitet, die auch die benötigten Ressourcen erfassen und parallel ausführbare Schritte koordinieren, so dass seine Fortführung des Prozesses nach Abschluss aller dieser Teilaktivitäten erfolgen kann. Insgesamt liegen ca.15 Prozessmodelle vor, die unterschiedliche Phasen der Closed-Loop spezifizieren, von der Aufnahme des Patienten und der Diagnostik bis zur Therapie (z.B. IORT).

Die Komponenten der Prozessplattform wurden prototypisch entwickelt und werden zur Zeit für präklinische Tests im Interventionsraum installiert. Dabei werden zunächst robotergestützte Verfahren für die Biopsieentnahme und radiologische Interventionen implementiert. Abb. 5 illustriert das Szenario der Biopsieentnahme: Die Nadel ist auf dem Roboterarm montiert, der Roboter platziert die Nadel autonom in der geplanten Position. In einer präklinischen Studie für 13 verschiedene Nadelpositionen wurden Winkelabweichungen von unter 1° und Zielpräzisionswerte von weniger als 4mm mit einer 15cm_Nadel gemessen[10]. Der Operateur führt dann die Nadel zum Tumorgewebe, wobei die Navigation durch Aufnahmen mit dem Artis Zeego unterstützt werden kann, um die aktuelle Lage der Nadelspitze im Gewebe zu überprüfen.

Erste Experimente zeigten, dass Cone-beam Computed Tomographie im Vergleich zu einem Multi-Detector-Computed-Tomograph-Ansatz bei vergleichbar Präzision Geschwindigkeitsvorteile bietet [11].

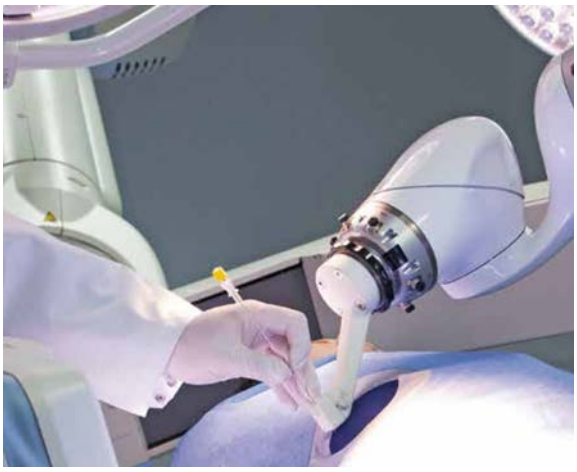


Abbildung 5: *Positionierung der Biopsienadel durch den Roboterarm*

Der Biopsie-Prozess ist als ausführbares BPMN-Modell spezifiziert (Abb.6 zeigt ein kleines Fragment der Struktur). Elementare Aktionen werden als Knoten (Rechtecke) im Graph dargestellt, die mit Verzweigungs- und Entscheidungsknoten durch Kanten verbunden sind. Alternative Aktionen (z.B. Initiierung mit/ohne Anästhesie) und parallele Aktionen (Entnahme Bioptat, Unterstützung der Entnahme) werden auf dezidierten Ebenen (sog. Swim lines) angeordnet, die jeweils einem Akteur (oder einer Gruppe von Akteuren) zugeordnet sind.

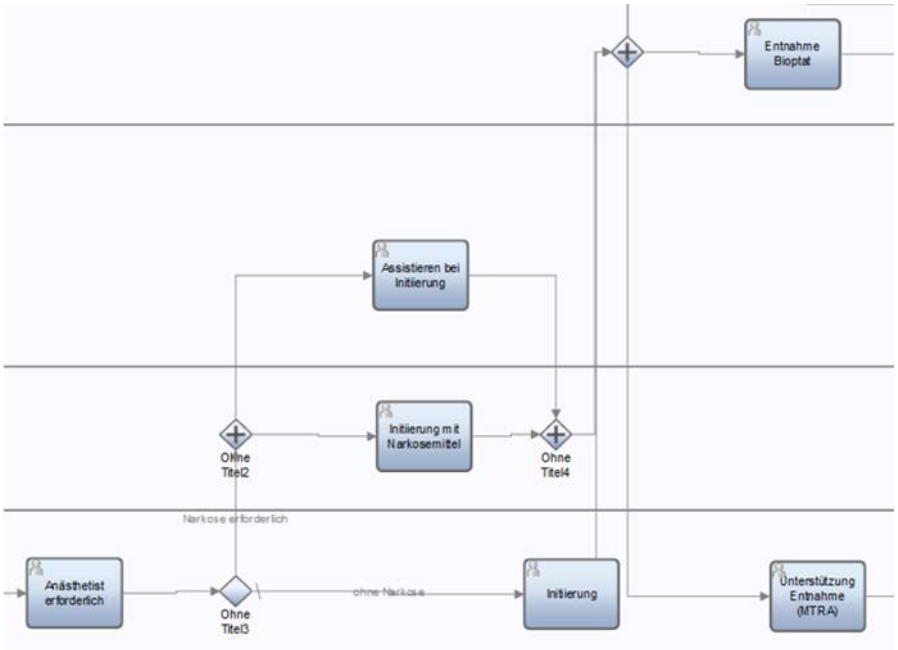


Abbildung 6: *Fragment des BPMN-Modells des Biopsieprozesses*

Der Prozessmanager nutzt dieses Modell zur Bereitstellung der notwendigen Informationen, die die Akteure zur Barbeitung ihrer aktuellen Aufgabe benötigen, und stellt sicher, dass alle Teilaktionen ordnungsgemäss abgeschlossen sind, bevor der nächste Schritt erfolgt. Im medizinischen Kontext ist immer die Möglichkeit gegeben, von dem geplanten Ablauf abzuweichen (etwa infolge einer Verschlechterung der Vitalwerte oder einer Blutung), so dass die Schnittstelle des Interventionsraumportals auch die Option bietet, alternative Schritte durchzuführen. Der so ad-hoc entstehende Prozess wird vom System archiviert und kann später in einer Analysephase zu neuen Erkenntnissen führen und die Prozessoptimierung unterstützen.

4 Diskussion

Das Konzept der Modellierung klinischer Prozesse ermöglicht die Koordination der komplexen Techniken der molekularen Intervention in flexibler Weise. Eine dynamische Zuordnung von Ressourcen leistet einen entscheidenden Beitrag zur Wirtschaftlichkeit des Ansatzes.

Eine Änderung oder Anpassung der Modelle aufgrund der Analyse realer Ausführungen und der Auswertung der dabei angefallenen Prozessdaten ist ein weiterer Vorteil des Ansatzes. Es ist geplant, die Datenbasis systematisch auf alle Bereiche der molekularen Intervention auszudehnen, um neben den konkreten Verlaufsdaten der Ad-hoc Prozesse auch die zugehörigen Ergebnisse aus Teilprozessen der Laboranalyse und der Bildgebung als Basis einer Prozessoptimierung zu nutzen. Hierzu sollen teilautomatisierte Verfahren des Process Mining genutzt werden.

Auch die technische Weiterentwicklung der molekularen Intervention ist durch die modulare Struktur der Prozessplattform antizipiert, die normierten Schnittstellen erlauben die Integration weiterer Diagnose- und Therapietechniken, Informationsquellen und Endgeräte.

5 Zusammenfassung

Dieser Beitrag skizziert die informationstechnischen Ansätze, die dazu beitragen, das im Forschungscampus M²OLIE angestrebte Ziel einer hocheffizienten minimal-invasiven molekularen Intervention bei oligo-metastasierten Patienten zu erreichen. Die klinischen Prozesse werden als ausführbare Konstrukte explizit modelliert und von einer Prozessmanagementplattform gesteuert, die mit den beteiligten Akteuren im Interventiosraum interagiert, situationsadäquate Informationen bereitstellt und alle relevanten Prozessparameter erhebt und archiviert.

Danksagung

Das Projekt M²ProcIT wird als Querschnittsprojekt im Forschungscampus M²OLIE unter dem Förderkennzeichen 13GW0093A gefördert, die Kooperation mit der QIT GmbH vom EU Network EIT Health mit einem Startup-Projekt unterstützt.

Referenzen

[1] *"Forschungscampus M²OLIE."* [Online]. Link: http://www.m2olie.de. [Zugriff: 30.05.2018].

[2] Schwarzbach, Matthias; Ronellenfitsch, Ulrich, *Klinikpfade in der Chirurgie: Ein Instrument für den Routinebetrieb?,* Dtsch Arztebl, 105(47): A-2512/ B 2135 /C 2059 (2008)

[3] Hubert, Simon, *Modeling and Dynamic Analysis of a Molecular Intervention Process for the Treatment of Oligometastases*, Masterthesis, Mannheim University (2013)

[4] Allweyer, T., *BPMN 2.0 – Introduction to the Standard for Business Process Modeling*, Norderstedt Books on Demand, 2. überarbeitete Auflage, 2016.

[5] Walker, J. M.; Carayon, P., *From Tasks to Processes: The Case for Changing Health Information Technology To Improve Health Care*. Health Affairs 28, No.2 (2009): 467-477

[6] Scharnetzke, Marco, *Clinical BPM – Prozessmanagement zur Steigerung der Wirtschaftlichkeit in Krankenhäusern*. INUBIUM Das Kundenmagazin der Inubit AG, 2010, No. 07:4-7

[7] Meier, J. et al., *Design and Evaluation of a Multimedia Electronic Patient Record "Oncoflow" with Clinical Workflow Assistance for Head and Neck Tumor Therapy*. Int J CARS (2014) 9:949-965

[8] "*OP 4.1*" [Online]. Link: https://www.op41.de/ [Zugriff: 25.07.2018].

[9] Fadel, F.G. et al., *A Generic Enterprise Resource Ontology*, Proceedings of the third IEEE Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, April 1994, Morgantown, West Virginia (WET ICE '94)

[10] Paetz T. et al., *An Ex-vivo Study on Needle Placement Accuracy of a Pre-clinical Robotic Guidance System*. JVIR Vol. 28, Issue 2, Suppl., Feb. 2017, p. S197

[11] Kostrzewa, M. et al., *Accuracy of Percutaneous Soft-tissue Interventios Using a Multi-axis, C-arm CT System and 3D Laser Guidance*. Eur J Radiol 84 (2015):1970-1975

PROMIS - Standardized Modelling of Surgical Procedures in Minimally Invasive Surgery for Open Datasets in Surgical Data Science

M. Wagner[1], L.B. Raedeker[1], S. Bodenstedt[2], S. Speidel[2], B.P Müller[1], H.G. Kenngott[1]

[1]Department of General, Visceral and Transplant Surgery, Heidelberg University Hospital, Heidelberg, Germany
[2]National Center for Tumor Diseases, Translational Surgical Oncology, Dresden, Germany

contact: hannes.kenngott@med.uni-heidelberg.de

Abstract

Availability of Open datasets is crucial for the development of research in machine learning and artificial intelligence. However, in surgery there is a lack in both, high quality data sets as well as standardization of procedures concerning the annotation of the data. Thus the very few existing open datasets cannot be processed mutually. To address this need, we are developing a model for standardized annotation of common surgical workflows in laparoscopic operations.

Keywords: Surgical Data Science, Context-aware information systems, Laparoscopic surgery, Open Science, Machine learning

1 Problem

Artificial intelligence and machine learning brought numerous technological advances to consumers such as voice recognition or autonomous driving. Here, vast amounts of data are created and processed. In surgery, surgical data can be used for the analysis and prediction of workflows for intelligent assisting systems [1]. In the near future, learning cognitive systems could be developed based upon this data that drive autonomous surgical robots or predict the surgeons need for specific assistance at any specific time during an operation [2].
Up until now, there has been the attempt to create a standardized set of terms, but only to descripe a limited number of workflows [3]. Thus the few existing data-sets [4, 5] cannot be annotated and therefore not processed in a standardized manner.
Here we introduce our attempt of creating a standardized modelling for annotation applicable for a wide range of surgical procedures as a foundation for the creation of open datasets to advance surgical data science.

2 Material and Methods

As a first step, a list of the 13 most relevant surgical procedures was defined. Laparoscopic videos of these operations were then captured within the Heidelberg University hospital. A workflow template was designed to ensure the best possible comprehensibility of provided data.
On this basis, a greath bandwith of recommended procedures for performing a laparoscopic surgery was researched using surgical textbooks, online databases such as “Touch Surgery” [6] as well as published research articles. The knowledge sources were researched for the standard workflow of the proposed procedure as well as possible deviations and complications. In addition we collected hints for the inexperienced surgeons.
An experienced clinician who did not participate in the surgery of our videos then oversaw the list and was involved in creating an optimized and standardized workflow. Similar tasks were arranged in “bags of tasks”, other steps were listed in the necessary order.
In the second part of our aproach, we developed a set ot annotation rules upon the defined workflow to allow for reproducability of workflow annotations. The respective laparoscopic video for each procedure was then revied upon these and annotated accordingly.
Third, using the workflow and the gathered laparoscopic videos, we then created an educational video in German and English language with the purpose of establishing our workflow annotation rules and educating future surgeons.

3 Results

Up until now we were able to produce an in-depth comparison of existing procedures for laparoscopic cholecystectomy, comparing 9 different workflows. Using this, we've established a standardized workflow on our own, specifically designed for the use in open datasets.
The standardized workflow was produced as an educational video subdivided into the following sections: (1) Ports, (2) Patient positioning and (3) Positioning of surgeon, assistant, nurse and monitors. The operations itself comprised (4) defined steps, (5) phases/bag of tasks (6) important activities (7) additional hints.
Additionally, we added information about (8) instruments used, (9) definition of start and end of phase, (10) anatomical structures involved and (11) important complications
We then applied this set of annotation rules to an example video of a laparoscopic cholecystectomy.
Beside the cholecystectomy, we are currently working on further surgical procedures such as sigmoidectomy for diverticulitis, rectal resection for cancer and proctocolectomy in chronic inflammatory bowel disease.

4 Discussion

The goal of our work was to define universal models of surgical workflows for the use with open Dataset, especially for annotation of workflows in laparoscopic videos. We succeded in creating a model template and accomplished applying it for a first standard procedure. The usability was ensured by annotating a demonstration video. Furthermore, we were able to create an corresponding educational video and a set of standardized annotation rules. In the future we will integrate our models into LapOntoSPM and the standards of the OntoSPM-collaboration [7]. Also, we will have to apply the workflow annotation to a larger set of laparoscopic videos to analyse inter- and intra-observer variance of our annotation models.
At this point of work, we would highly value the input and needs of the CURAC-community in order to advance the research in open and shared surgical workflow models.

5 Conclusion and References

Our approach of creating standardized surgical workflows for the annotation of datasets is a foundation for the further development of enhanced machine learning used during surgical operations.

References

[1] R Maier-Hein, L., Vedula, S. S., Speidel, S., Navab, N., Kikinis, R., Park, A., Eisenmann, M., Feussner, H., Forestier, G., Giannarou, S., Hashizume, M., Katic, D., Kenngott, H., Kranzfelder, M., Malpani, A., März, K., Neumuth, T., Padoy, N., Pugh, C., Schoch, N., Stoyanov, D., Taylor, R., Wagner, M., Hager, G. D. und Jannin, P. (2017). *Surgical data science for next-generation interventions.* Nature Biomedical Engineering *1*, 691–696.

[2] Kenngott, H.G., Wagner, M., Nickel, F., Wekerle, A.L., Preukschas, A., Apitz, M., Schulte, T., Rempel, R., Mietkowski, P., Wagner, F., Termer, A., Muller-Stich, B.P.: *Computer-assisted abdominal surgery: new technologies.* Langenbecks Arch Surg (2015)

[3] Katić, D., Julliard, C., Wekerle, A.-L., Kenngott, H., Müller-Stich, B. P., Dillmann, R., Speidel, S., Jannin, P. und Gibaud, B. *LapOntoSPM: an ontology for laparoscopic surgeries and its application to surgical phase recognition*. Int J Comput Assist Radiol Surg *10*, 1427–1434 (2015)

[4] Twinanda, A. P., Shehata, S., Mutter, D., Marescaux, J., de Mathelin, M. und Padoy, N. (2017). *EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos*. IEEE Transactions on Medical Imaging *36*, 86–97.

[5] Stauder, R., Ostler, D., Kranzfelder, M., Koller, S., Feußner, H. und Navab, N. (2016). *The TUM LapChole dataset for the M2CAI 2016 workflow challenge*. arXiv:161009278 [cs], URL: http://arxiv.org/abs/1610.09278 [as of June 2018].

[6] URL: https://www.touchsurgery.com/ [as of June 2018].

[7] start [OntoSPM Collaborative Action] (o. J.)., URL: https://ontospm.univ-rennes1.fr/doku.php?id=start [as of June 2018].

Ontology-based Annotation of Endoscopic Images

A. Uciteli[1], S. Siemoleit[1,2], K. Tahar[1], R. Bieck[3] and H. Herre[1]

[1]Institute of Medical Informatics, Statistics and Epidemiology, University of Leipzig, Germany
[2]Stoneball GmbH, Leipzig, Germany
[2]Innovation Center for Computer Assisted Surgery, University of Leipzig, Germany

Contact: auciteli@imise.uni-leipzig.de

Abstract

The established optical surgical navigation systems require training and experience for the registration process and the overall optical marker setup. The use of these complex technologies is time-consuming and often error-prone.

The BIOPASS project develops a novel localization approach for markerless navigation systems based on process and image databases of learned surgical procedures to intra-operatively identify anatomical landmarks and risk structures. In order to recognize anatomical structures visible in the endoscopic images, several classification algorithms were integrated into the system's architecture.

We present the Annotation Configuration Ontology (ACO) and the software tool SWEConfigGen, which allow the annotation of endoscopic images using ontological concepts. Two external ontologies FMA and VCO were successfully integrated and used for annotating 30 video recordings of ENT procedures. This domain knowledge was utilized to train the system's classifiers developed in the BIOPASS project.

Keywords: Minimally-invasive surgery, Endoscopic images, Annotation, Ontology, Classification

1 Problem

Minimally-invasive surgical procedures offer numerous advantages for the patients, e.g., reduced access trauma, faster recovery and shorter periods for in-patient stays and rehabilitation. The special challenges for the surgeon are orientation and navigation without a direct view of the surgical area as well as the guidance of instruments with limited hand-eye coordination by looking at images provided by the endoscope. The established optical surgical navigation systems require training and experience for the registration process and the overall optical marker setup. The use of these complex technologies is time-consuming and often error-prone.

The BIOPASS project [1] develops a novel localization approach for markerless navigation systems based on process and image databases of learned surgical procedures to intra-operatively identify anatomical landmarks and risk structures. In order to recognize anatomical structures visible in the endoscopic images, several classification algorithms were integrated into the system's architecture. Additional information is provided by sensors developed in this project to further enrich the classifier data. An introduction of the project is given in [2]. Further ontology-based components developed in the project are presented in [3,4].

This work focusses on the creation of training data for the system's classifiers by the annotation of existing video recordings of surgical interventions. We present the Annotation Configuration Ontology (ACO) and the software tool SWEConfigGen, which allow the use of ontology concepts for the annotation.

2 Material and Methods

The necessary domain knowledge is specified by experts in form of annotations of existing videos and functions as training data for the system's classifiers. Not only the anatomical structures that are visible in the videos, but also descriptions of the operating procedures (e.g., activities of the surgeons or instruments) had to be annotated. Same goes for the various types of image properties (e.g., the brightness and hue values). These annotations needed to be created using a controlled vocabulary (ontology), which enables a unified internal communication within the system.

To allow the annotation of endoscopic images with ontology concepts, the Surgical Workflow Editor [5] was used. The Surgical Workflow Editor is ideal to create and visualize a taxonomy of annotation concepts. However, the tool does not support the integration of external ontologies or terminologies without additional manual effort. To solve this problem, we developed the Annotation Configuration Ontology (ACO) and the SWEConfigGen, an ACO-based plug-in for the ontology editor Protégé [6]. The tool enables domain experts to load arbitrary ontologies into the Surgical Workflow Editor and use them for annotations.

When configuring the Surgical Workflow Editor (including the specification of annotation concepts), a configuration file is created in XML format. Our idea was to generate such configuration files including the integration of ontologies using an ontology-based software tool.

2.1 Annotation Configuration Ontology (ACO)

The Annotation Configuration Ontology is based on the XML structure of the configuration files of Surgical Workflow Editor and was implemented in the Web Ontology Language (OWL) [7]. The ACO contains concepts representing certain XML tags, e.g., <*itemType*>, <*actors*> or <*relations*>. The concept **itemType** plays a central role in the integration of concepts contained in ontologies that have to be used for annotation. These concepts have to be set as subclasses of **itemType**. If not the entire ontology, but only its subtree has to be utilized for the annotation, it is possible to specify the desired taxonomy depth with the annotation property **taxonomy_depth** as shown in Figure 1.

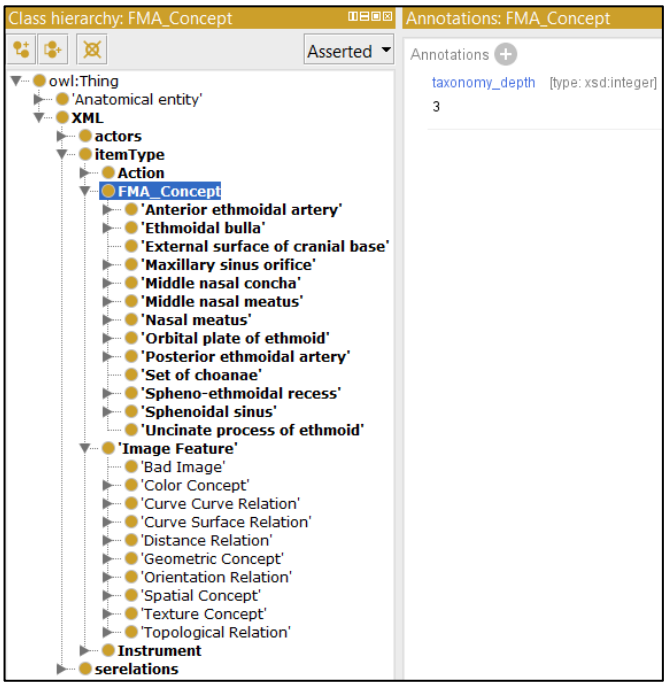


Figure 1: *Screenshot of Protégé showing ACO with integrated FMA and VCO concepts*

2.2 SWEConfigGen: The Tool for Generating Configuration Files of Surgical Workflow Editor

The tool for generating configuration files of the Surgical Workflow Editor, and integrating external ontologies (SWEConfigGen) is designed and developed as a Protégé plug-in according to the three ontology method [8]. This method for developing software is based on the interactions of three different kinds of ontologies: a task ontology (TO), a domain ontology (DO) and a top-level ontology (TLO). The TO is an ontology for the general problem that the software is intended to solve. The DO provides the domain-specific knowledge, whereas the TLO integrates the TO and the DO and is used as the foundation of them. The TLO also provides means for integrating data from different domains.

In our case, the ACO functions as a TO. It describes the XML structure of the configuration files of Surgical Workflow Editor and enables to integrate external ontologies. Thus, the ACO can help to solve the problem of generating such configuration files containing concepts from desired ontologies that have to be used for annotation. To describe the required domain-specific knowledge, we utilized, e.g., the Foundational Model of Anatomy (FMA) [9] and the Visual Concept Ontology (VCO) [10] as DOs. For integrating the TO and the DOs we used the General

Formal Ontology (GFO) [11,12] as TLO (for the sake of simplicity, the GFO is not shown in the figure and is not described in this paper).

The SWEConfigGen implements the access to classes and properties of the ACO but it does not know the DOs embedded in the ACO. The tool uses the TO as an interface to access the DO entities. For example, the SWEConfigGen obtains access to the root concepts of integrated domain ontologies (e.g., FMA Concept and Image Feature) by querying the subclasses of **itemType**. Then, the tool navigates through the taxonomy downwards and fetches the desired concepts.

3 Results

To allow an easy adaption to other systems, we decided to use external ontologies as our controlled vocabulary. Moreover, we profited from the soundness of their taxonomies and improved the quality of our created data streams. Up to now, two ontologies were successfully integrated, namely the Foundational Model of Anatomy (FMA) and the Visual Concept Ontology (VCO). Both ontologies were used to annotate video recordings of surgical interventions.

FMA models anatomical structures of biological organism and describes the relations between them. We specified FMA concepts that are able to function as anatomical landmarks and as surgical risk structures in Functional Endoscopic Sinus Surgery (FESS). VCO is used to visually describe the processed video images. The ontology is subdivided into three partitions, i.e., texture concepts, color concepts and spatial concepts. Both ontologies are used in the internal communication of our system, such that a web service will receive the notification of an identified anatomical structure represented by its corresponding concept in FMA. Analogously, the image properties of an endoscopic image are encoded with the help of VCO. Moreover, if FMA and VCO turn out to be not expressive enough for the use case of minimally-invasive surgery, we can augment ACO with additional ontologies.

The annotation process was performed on 30 video recordings of FESS procedures. The concepts of external ontologies were successfully integrated and annotated in the videos (Figure 2).

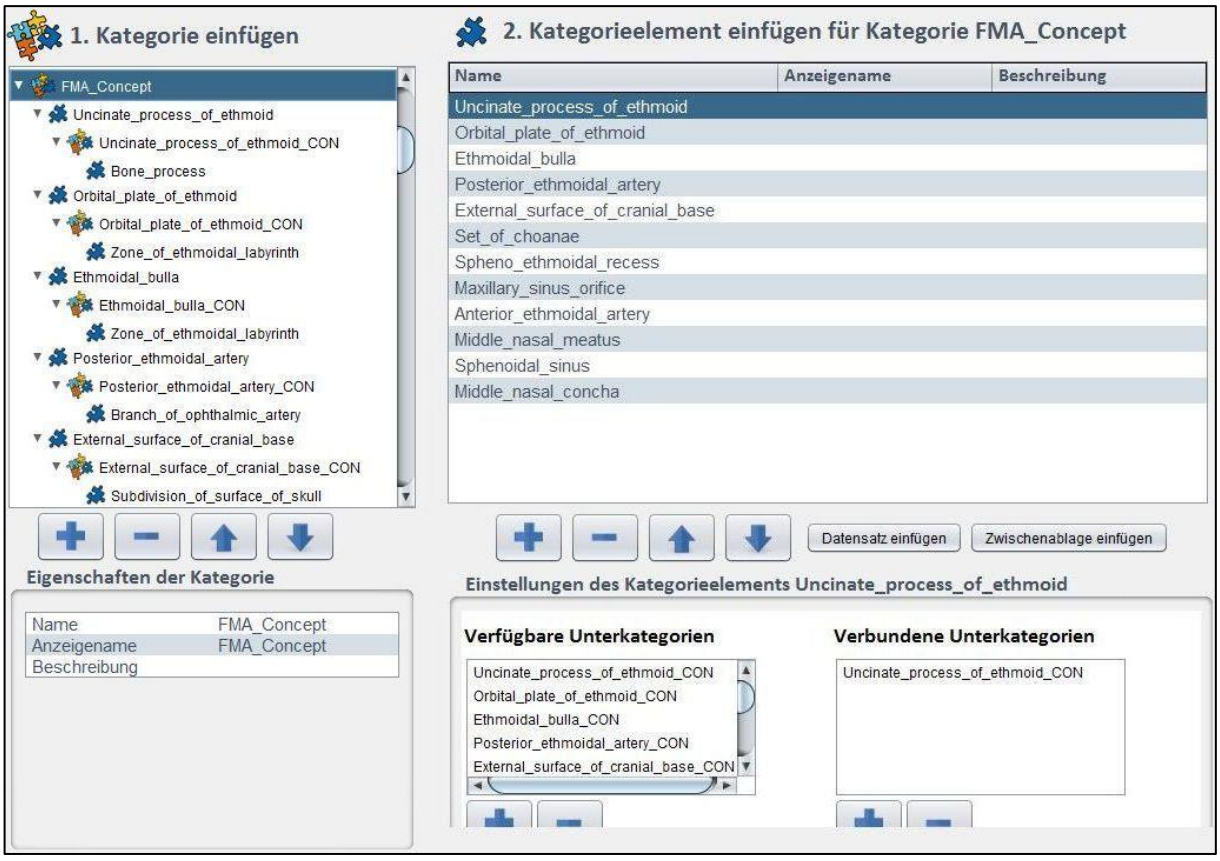


Figure 2: *Surgical Workflow Editor with integrated FMA concepts*

4 Discussion

Ontologies are widely used in biological and biomedical research. According to [13], their success is based on the combination of four main features found in almost all ontologies:

1. Providing standard identifiers for classes and relations
2. Providing a vocabulary for a domain
3. Providing metadata describing the meaning of the classes and relations
4. Providing machine-readable axioms and definitions

Each of these features enables various applications in the fields of data integration, data access, and analysis, but the combination of these four features can additionally support the integrative analysis and interpretation of multimodal data [13].

The use of ontologies can also support the incorporation of various types of features for multimodal learning and classification [13,14]. The combination of information from text, images, videos, or structured data to improve classification can be facilitated using ontologies. Relevant features from each type of information have to be extracted and the results have to be represented by a single ontology combining the information used for training of a classifier [13].

Similar to this approach, videos were annotated using ACO in the BIOPASS project. Hence, relevant concepts from FMA and VCO were integrated with ACO to generate training data. This training data, including additional sensor data, was used to train the classifiers developed during the project. In other words, the classifier's algorithms were trained based on surgical video recordings annotated with concepts from FMA and VCO. In live operation, the BIOPASS system analyzes the current endoscopic images and suggests, inter alia, the recognized anatomical structures in the terminology of FMA.

With relatively little effort, further standardized ontologies can be integrated with ACO in order to describe additional aspects of the surgical interventions. Our solution allows the usage of a controlled vocabulary within the overall system, a unified internal communication between its components, as well as an easy adaption to other systems.

5 Conclusion

In this paper, we presented the Annotation Configuration Ontology (ACO) and the software tool SWEConfigGen, which allow the annotation of endoscopic images using ontological concepts. Two external ontologies FMA and VCO were successfully integrated and used for annotating 30 video recordings of FESS procedures. This domain knowledge was utilized to train the system's classifiers developed in the BIOPASS project. The classifiers' results should help to intra-operatively identify anatomical landmarks and risk structures in the endoscopic images.

Acknowledgment

This work was funded by the German Federal Ministry of Education and Research (BMBF) under reference number 16SV7254K as part of the BIOPASS project [1].

References

[1] BIOPASS Project, (n.d.). https://www.iccas.de/biopass/.
[2] R. Bieck, K. Heuermann, M. Schmidt, A. Schmitgen, S. Arnold, A. Dietz, and T. Neumuth, Towards an Information Presentation Model of a Situation-Aware Navigation System in Functional Endoscopic Sinus Surgery, in: 15th Conference of the German Society of Computer and Robotic Assisted Surgery, Der Andere Verlag, Bern, 2016: pp. 27–32.
[3] S. Siemoleit, A. Uciteli, R. Bieck, and H. Herre, Ontological Modelling of Situational Awareness in Surgical Interventions, in: Proceedings of the Joint Ontology Workshops (JOWO) 2017, Episode 3: The Tyrolean Autumn of Ontology, Bozen-Bolzano, Italy, 2017, CEUR Workshop Proceedings, Vol. 2050, Aachen, 2018.
[4] S. Siemoleit, A. Uciteli, R. Bieck, and H. Herre, Processual Reasoning over Sequences of Situations in Endoscopic Surgery, *Studies in Health Technology and Informatics*. **243** (2017) 222–226.

[5] T. Neumuth, B. Kaschek, D. Neumuth, M. Ceschia, J. Meixensberger, G. Strauss, and O. Burgert, An observation support system with an adaptive ontology-driven user interface for the modeling of complex behaviors during surgical interventions, *Behavior Research Methods*. **42** (2010) 1049–1058.
[6] Protégé: A free, open-source ontology editor and framework for building intelligent systems, (n.d.). http://protege.stanford.edu/.
[7] OWL 2 Web Ontology Language, (n.d.). https://www.w3.org/TR/owl2-overview/.
[8] R. Hoehndorf, A.-C.N. Ngomo, and H. Herre, Developing Consistent and Modular Software Models with Ontologies, in: H. Fujita, and V. Marik (Eds.), New Trends in Software Methodologies, Tools and Techniques: Proceedings of the Eighth SoMeT_09, IOS Press, 2009: pp. 399–412.
[9] C. Rosse, and J.L.V. Mejino, The Foundational Model of Anatomy Ontology, in: Anatomy Ontologies for Bioinformatics, Springer, London, 2008: pp. 59–117.
[10] N. Maillot, M. Thonnat, and A. Boucher, Towards ontology-based cognitive vision, *Machine Vision and Applications*. **16** (2004) 33–40.
[11] H. Herre, B. Heller, P. Burek, R. Hoehndorf, F. Loebe, and H. Michalek, General Formal Ontology (GFO): A Foundational Ontology Integrating Objects and Processes. Part I: Basic Principles (Version 1.0), Research Group Ontologies in Medicine (Onto-Med), University of Leipzig, 2006.
[12] H. Herre, General Formal Ontology (GFO): A Foundational Ontology for Conceptual Modelling, in: R. Poli, M. Healy, and A. Kameas (Eds.), Theory and Applications of Ontology: Computer Applications, Springer, Netherlands, 2010: pp. 297–345.
[13] R. Hoehndorf, P.N. Schofield, and G.V. Gkoutos, The role of ontologies in biological and biomedical research: a functional perspective, *Brief Bioinform*. **16** (2015) 1069–1080.
[14] A. Skowron, H. Wang, A. Wojna, and J. Bazan, Multimodal Classification: Case Studies, in: Transactions on Rough Sets V, Springer, Berlin, Heidelberg, 2006: pp. 224–239.

Optimizing External Surface Sensor Locations for Respiratory Tumor Motion Prediction

Yusuf Özbek, Zoltan Bardosi, Srdjan Milosavljevic, Wolfgang Freysinger

4D Visualization Research Group, Univ. ENT Clinic,
Medical University of Innsbruck, Innsbruck, Austria

Kontakt: Yusuf.Oezbek@Student.i-med.ac.at

Abstract

An optimization method is presented for sensor locations in respiratory motion prediction using external surface fiducials/sensors to be used with the optical and electromagnetic tracking systems. The purpose of this work is to propose a procedure that automatically identifies a sparse set of location combinations on the patient's surface, based on the placed 3D printed sensor-holders with the optimal prediction power for the internal tumor motion. The method can automatically find optimal location combinations on the patient's surface, based on automatically determined external sensor-holder locations and serves a recommended list to place sensors for optimal prediction of the single or multiple internal tumor motions. The method is validated on an artificial respiratory phantom model. The external sensor-holders within a CT marker (fiducial) and internal markers(tumors) are placed on/in the built phantom before treatment and a 4D-CT image with 11 different respiratory phases is acquired. The locations of the surface markers and tumor/s are determined by a morphology-based algorithm in the CT images. The preoperative recommendation procedure for determining the optimal sensor locations to be used in the real-time prediction phase is performed with the machine learning Gaussian Processes Regression algorithm. The method is evaluated using RMSE that gives deviations between measured and predicted internal marker trajectories. The average RMSE value is between 0.07 and 0.33 mm.

Keywords: Tumor Tracking, Respiratory Motion, Prediction, Optimization

1 Problem

The localization of the internal tumors or structures and the detection of respiratory organ or tumor movement under certain therapies (radiation therapy, radiofrequency ablation, etc.) and real-time visualization of these movements are an important concern in the safe and effective provision of precision radiotherapy, computer-assisted tumor surgery and biopsy [1, 5]. The information on the movements and tracking of the current positions of targeted tumors, e.g. in prostate, liver, lung or soft tissue tumors are traditionally performed by using abdominal compression, breath hold, respiratory gating, implanted radiation-impermeable markers and real-time motion tracking of these markers with optical (OP)/electromagnetic(EM) tracker or over real-time image processing of interventional 4D-CT, -MRI, and 3D/4D-ultrasound [6], [7]. Because of the difficulties with respect to intraoperative complexity of these methods, there is currently no clinically established solution for the reliable determination of respiratory movements of internal organs, tumors or soft tissues.

Specifically, in the real-time motion tracking of external sensors with OP or EM tracking systems, it is difficult to determine the optimal amount of sensors to be used and the exact placement locations of these sensors on the patient's surface in preoperative phase. The used surface markers are randomly distributed and placed on the patient in the near region of a surgical area and all of them are used for respiratory motion prediction, which can increase the error rate in the real-time prediction step.

2 Material and Methods

2.1 Phantom Respiratory System Model

To replicate the human respiratory system and to predict the motion based on the respiratory simulation, a custom phantom model is built (Fig. 1). The main components consist of a standard rubber hot-water bottle (replicates eg. abdominal region of human), a spherical rubber balloon (replicates a moving organ), flexible silicone tube, water cannon, sensor-holders within CT skin markers (external) and skin markers or retro-reflective balls (internal).

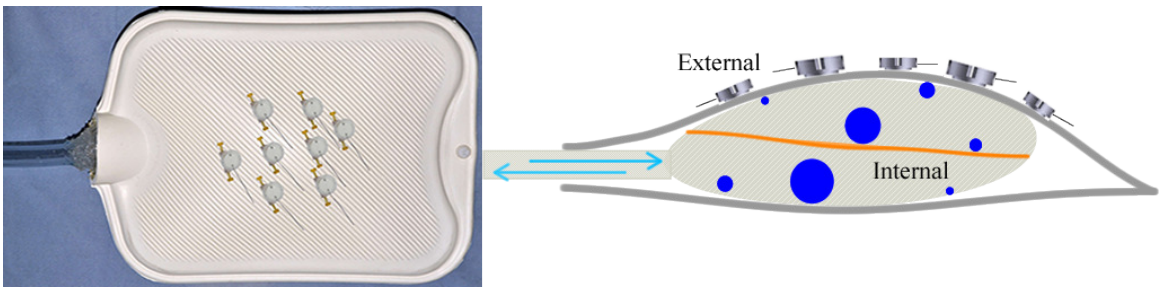


Figure 1: *Left: Respiratory model with the fixed sensor-holders on it. Right: Interior view of the model during an inhalation. The inside of the balloon (max. inflation ⊘ 120 cm) is brought out and a rubber band from both ends at the body-neck transition and at the drip point of the balloon is glued, so that it stays in the middle of the balloon if the outside brought in and inflated. The XSpot skin markers(⊘ 1.5 mm, on the rubber band and on the inner wall) and sphere-shaped retro-reflective markers (⊘ 12 mm, on the rubber band and on the inner wall), which symbolize some tumors, are placed inside of the balloon. A flexible silicone tube (L 200 cm , ⊘ 20 mm) is used for inflating the balloon with the water cannon (82x5x15 cm) to perform an inhalation and exhalation simulations of the respiratory model. The balloon is inserted in the hot-water bottle and fixed on the neck of bottle.*

2.2 Sensor Holders

Custom made sensor-holders are developed (Fig. 2), which helps to perform an offline prediction preoperatively, without using real sensors or tracking system for the prediction and optimize the determination of sensor locations. Based on the offline prediction results, the operator gets best possible sensor count and locations to use in real-time prediction as recommended. In case of single or multiple tumor motion prediction, the operator can change the locations of the sensors by taking out the sensor from one sensor-holder and replacing into the other, according to the given recommendation list and perform intervention for the next selected tumor.
The sensor-holder consist of two parts:
Optical-Sensor-Holder: Main part. Contains within a XSpot CT marker in the center, which position is determined in CT image space and holds the active IRED sensor (11x7x5 mm) of NDI Optotrak Certus tracker during real-time prediction in intraoperative phase concentrically. They are attached to the phantom patient preoperatively without sensors and used during interventional imaging.
Magnetic-Sensor-Holder: Contains one 5DOF sensor (L 8 mm , ⊘ 0.6 mm) of the NDI Aurora tracking system and fits in main part concentrically.

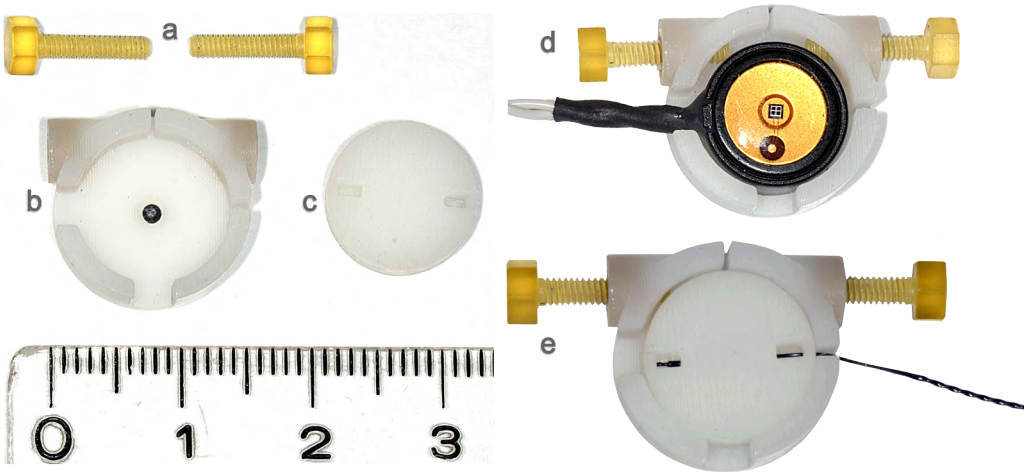


Figure 2: *a) Two M2 screws that are placed from both directions to establish a rigid setup once an OP or EM sensor is placed in the main part. b) The main part within a XSpot marker. It used in offline prediction step. c) The Magnetic-Sensor-Holder to fix the EM sensor in it. d) View of sensor-holder within an OP marker. e) View of sensor-holder within an electromagnetic sensor. The parts a,c,d and e are intended for the real-time prediction.*

2.3 Data Acquisition

To validate the optimization method with realistic clinical cases a 4D-CT scan of the phantom is acquired in the following manner: The phantom model with 10 main sensor-holders on the surface within external markers and 7 internal markers in the balloon is placed in CT device. A person has performed the respiration (inhalation/exhalation) manually while standing behind the device and using a water cannon, which is physically connected to the phantom model over flexible tube. For the 4D-CT a scanner in Univ. Clinic for Radiology (Siemens healthcare Austria) in Medical Univ. of Innsbruck is used. The 4D-CT scan consists of 11 discrete time steps of a breathing cycle. Each axial CT slice (512x512 px) has a thickness of 1.0 mm and 11 discrete CT phases consist of 261 images with 0.488x0.488x0.488 mm pixel spacing.

2.4 Automatic Marker Detection

In order to learn the respiratory cycle of the patient from the observed CT images, where the optimal sensor locations for prediction are determined, a regressor is trained to predict internal motion given this data. The precise locations of external and internal fiducials in the CT image space are detected by a GPU accelerated volumetric detection method [8].
For this purpose, each of 11 CT phases are thresholded and binarized to determine the centroid of the fiducials. On the resulting image the 3D fiducials are filtered for spherical structures using morphological opening with a spherical structuring ball element of the appropriate scale given the voxel size of the dataset and the physical dimensions of the markers. Using a geometry filter on the resulting spherical blobs best candidates are selected based on ellipticity and size. From the best candidates the blob centroids are calculated and stored to use in prediction step (Fig. 3).

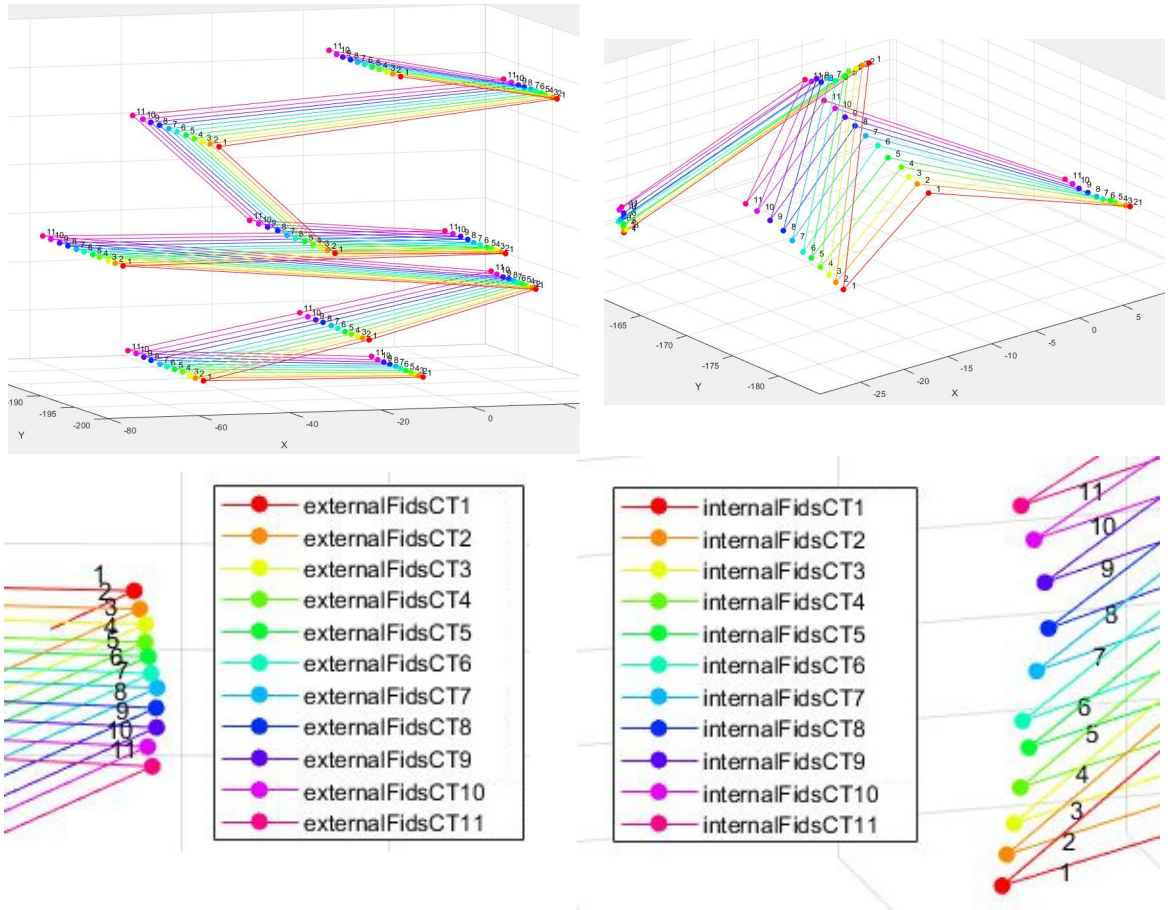


Figure 3: *Top Left: Discrete chronological 3D movement positions of all external surface fiducials during in-/exhalation over 11 time-steps (Time-step 1: Fully inhaled, Time-step 11: Fully exhaled). The location coordinates are obtained by automatic marker detection. Bottom left: The positions of one external fiducial in 4D-CT image space. Movement positions of internal fiducials in top right and bottom right.*

2.5 Respiratory Motion Prediction with GPR

For the offline prediciton, a machine learning approach (Gaussian Process Regression) [9] with six different kernel (covariance) functions are implemented and tested statistically. RBF, RBF + W, RQ, C * RBF, ESS + W and M, where RBF is radial basis function, W is white kernel, RQ is rational quadratic, C is constant kernel, ESS is exp sine squared and M is matern kernel.
In the automatic marker detection prestored external and internal fiducial coordinates (each detected fiducial contains 11 different positions in 3D Cartesian coordinate system in CT image space) for the 11 CT image dataset are used respectively. The recommendation procedure to specify best sensor combinations and numbers in our GPR implementation, first it reads every surface fiducial coordinates as input and internal fiducials as target data. Secondly, based on the existing targets, the procedure creates a combined group with a user defined tuple of input coordinates and tries to predict over those combination tuple the selected target with every kernel function in an iterative loop. Thirdly, each combination member (input) is mapped to the target and this mapping is used for the prediction.

To validate the prediction, the Leave One Out cross validation procedure (LOO) on the mapped data is applied. The parameter optimization is performed automatically on 10 of the 11 3D data and then the performance of the tuned algorithm is tested on the 11th input/target data. So, the 11th pair is the test set and the other 10 pairs are the training data for optimizing the free parameters of the GPR algorithm. This process is repeated 11 times, each time leaving out a different pair to use as the single test case.
As an outcome, the prediction results of each combination with given target are sorted and best sensor combinations for all targets are found. Finally the operator becomes a recommendation sensor list for every target to use for real-time prediction.

Since the respiration rate and regulation of breathing can be difficult in time, an over-fitting problem by the prediction can come across. It occurs when GPR captures the noise in the data. Specifically, if the algorithm shows low bias, but high variance. Over-fitting is often a result of an excessively complicated model, and it can be prevented by fitting multiple additional data (Data Augmentation) and using validation or cross-validation to compare their predictive accuracies on test data to get a reliably predictions in future.
To overcome this problem, we add random noisy data (Gaussian distributed) in the original input and target data (raw data) by analyzing every three columns of 11 different positions of an input/target marker automatically. The production of noisy data is based on the default raw data for every different marker and for every different individual column of that marker in 3D coordinate space. The noisy data was not fixed and for each run of algorithm 50 different noise in given limits (between min. and max. of individual columns) are generated. When generating random data the mean of distribution μ= 0.2 and standard deviation σ= 0.200 is used.

3 Results

Table 1 represents the prediction results for the first 5 target markers in the balloon with automatically recommended 3 surface marker combinations. Each input marker in the recommended triple tuple is processed with the listed target respectively. The best RMSE is obtained from the RBF + W kernel with min. 0.07 mm for target 1 and max. 0.33 mm for target 4 using recommended input fiducial combinations. All GPR predictions are performed with a raw and noisy data for each input and target marker respectively.
The worst prediction RMSE is obtained from the ESS + W kernel covariance function for the same targets using recommended input fiducials with min. 0.48 mm and max. 3.45 mm RMS after prediction using raw data.

Internal Target Marker	Recomm. Input Comb.	Used Kernel	# Raw Data	# Noisy Data	# Raw + Noisy Data	μ Raw Pred. RMS	μ Noisy Pred. RMS	μ Raw + Noisy Pred. RMS
1	8	RBF+WK	11x3	50x3	61x3	0.07	0.05	0.06
	7					0.07	0.04	0.05
	2					0.07	0.08	0.06
2	8					0.11	0.23	0.21
	2					0.11	0.17	0.22
	3					0.12	0.16	0.19
3	5					0.14	0.28	0.14
	3					0.14	0.28	0.15
	7					0.14	0.25	0.14
4	8					0.32	0.10	0.12
	2					0.32	0.08	0.09
	5					0.33	0.09	0.10
5	8					0.15	0.06	0.06
	5					0.16	0.07	0.07
	10					0.16	0.05	0.06

Table 1: *Overview of the offline prediction results. The used kernel is configured with the length scale: 100.0, length scale bounds: 1e-2, 1e3, noise level: 1 and noise level bounds: 1e-10, 1e+1. The Gaussian Process Regressor is configured with normalized target value and the number of restarts of the optimizer for finding the kernel parameters is set to 10. All RMSE values are in mm.*

4 Discussion and Conclusion

In many treatments, where respiratory motion prediction and tracking is a necessary approach to apply, the success of a treatment strongly depends on the accuracy of prediction, which is related to the detection and placement of the external surface marker locations on the patient.
Therefore, the location and distribution of the sensors in e.g. external radiotherapy, needle biopsy, etc. is important and challenging during real-time the tumor motion prediction. In the most clinical approaches, the location of external markers is chosen empirically; that is, operator-dependent [10]. Within this work, we examined that the distribution and selection of the best placement locations automatically and intelligently, gives better prediction results with using less numbers of external sensors (at least 3) to use e.g. in the thorax or abdominal regions. The proposed method is tested through a custom built respiratory phantom model to provide the required dataset, which surrogates same breathing circle of a real patient [11] (Table 2).

To find the optimum placement locations of external markers onto phantom patient, the GPR was proposed, which predicts the value of certain variables, based on measurements of other variables that parameterized by a mean and covariance or kernel functions. Based on the results from the GPR prediction (Table 1), the regions on the patient as optimum marker placement are determined.

Internal Target Marker	1	2	3	4	5
3D Movements in mm					
x	0.69	6.52	6.21	-4.05	1.8
y	6.44	17.6	16.57	3.1	1.97
z	1.51	2.39	1.4	-9.47	6.58

Table 2: *The 3D positional movement variations of the internal target markers (marker centroids) from fully exhaled to fully inhaled time-steps after 4D-CT.*

In this work, a method is presented to automatically determine the optimal locations of the external OP/EM sensors on the patient's surface to predict internal tumor motions. The preopertively determined optimal sensor locations can be used in real-time tumor motion tracking using the sensors of NDI Optotrak Certus or NDI Aurora tracking systems.

The experiments and evaluations on the built realistic respiratory phantom model have shown that by using our method, 3 surface sensors can be enough for a reliable prediction, based on the selected tumor inside

the body. The experiments give also information the distributing of recommended sensors locations has a high correlation between the surface motion and the internal tumor motion. With this procedure, EM or OP 3D measurement technologies can be used for real-time prediction and it is suitable for use in the medical environment.
All processes in the preoperative phase respectively (placing sensor holders on the patient, performing CT scan, visualizing patient data, performing automatic marker detection, performing offline GPR, removing unnecessary sensor holders and placing sensors in sensor holders) will be evaluated in the future on the phantom model and real-patients to validate the presented approach. For this, the OP or EM tracker will serve real-time sensor data to be used as input for the real-time prediction in the intraoperative phase.

5 References

[1] I. Buzurovic, K. Huang, Y. Yu, T.K. Podder, *A robotic approach to 4D real-time tumor tracking for radiotherapy*, Phys. Med. Biol. 56(5), 12991318 (2011)

[2] J.R. Wong, L. Grimm, M. Uematsu, R. Oren, C.W. Cheng, S. Merrick, P. Schiff, *Image-guided radiotherapy for prostate cancer by CT-linear accelerator combination: prostate movements and dosimetric considerations*, Int. J. Radiat. Oncol. Biol. Phys. 61(2), 561569, (2005)

[3] A. Sawant, R.L. Smith, R.B. Venkat, L. Santanam, B. Cho, P. Poulsen, H. Cattell, L.J. Newell, P. Parikh, P.J. Keall, *Toward submillimeter accuracy in the management of intrafraction motion: the integration of real-time internal position monitoring and multileaf collimator target tracking*, Int. J. Radiat. Oncol. Biol. Phys. 74(2), 575582 (2009)

[4] W.D. DSouza, S.A. Naqvi, C.X. Yu, *Real-time intra-fraction-motion tracking using the treatment couch: a feasibility study*, Phys. Med. Biol. 50(17), 40214033 (2005)

[5] I. Buzurovic, T.K. Podder, K. Huang, Y. Yu, *Tumor motion prediction and tracking in adaptive radiotherapy. IEEE International Conference on Bioinformatics and Bioengineering*, pp.273278, (2010).

[6] Suk Jin Lee, Yuichi Motai, *Prediction and Classification of Respiratory Motion*, ISBN:978-3-642-41508-1, Springer-Verlag, Berlin Heidelberg, (2014)

[7] AAPM Task Group 76, *The Management of Respiratory Motion in Radiation Oncology*, ISBN:1-888340-61-4, American Association of Physicists in Medicine One Physics Ellipse, College Park, MD, (2006)

[8] Bardosi Z., *OpenCL accelerated GPU binary morphology image filters for ITK*, The Insight Journal, (2015)

[9] C. Rasmussen, C. Williams, *Gaussian Processes for Machine Learning*, ISBN:026218253X, Mit University Press Group Ltd, (2005)

[10] Hui Yan, Guopei Zhu, James Yang, Mei Lu, Munther Ajlouni, Jae Ho Kim, Fang-Fang Yin, *Investigation of the location effect of external markers in respiratory-gated radiotherapy*, Journal of Applied Clinical Medical Physics, Vol.9 Nr.2, (2008)

[11] Weiss E., Wijesooriya K., Dill SV., Keall PJ., *Tumor and normal tissue motion in the thorax during respiration Analysis of volumetric and positional variations using 4D CT*, Int. J. Radiation Oncology Biol. Phys., Vol. 67, No. 1, pp. 296307, (2007)

Peer-to-Peer-Navigation in der computerassistierten Chirurgie

S. Strzeletz[1], *S. Hazubski*[1], *J. L. Moctezuma*[2], *H. Hoppe*[1]

[1] *Hochschule Offenburg, Labor für Computerassistierte Medizin, Offenburg, Deutschland*
[2] *Stryker Robotics, Freiburg, Deutschland*

Kontakt: simon.strzeletz@hs-offenburg.de

Abstract

Optische Navigationssysteme weisen bisher eine eindeutige Trennung zwischen nachverfolgendem Gerät (Tool Tracker) und nachverfolgten Geräten (Tracked Tools) auf. In dieser Arbeit wird ein neues Konzept vorgestellt, dass diese Trennung aufhebt: Jedes Tracked Tool ist gleichzeitig auch Tool Tracker und besteht aus Marker-LEDs sowie mindestens einer Kamera, mit deren Hilfe andere Tracker in Lage und Orientierung nachverfolgt werden können. Bei Verwendung von nur einer Kamera geschieht dies mittels Pose Estimation, ab zwei Kameras werden die Marker-LEDs trianguliert. Diese Arbeit beinhaltet die Vorstellung des neuen Peer-To-Peer-Tracking-Konzepts, einen sehr schnellen Pose-Estimation-Algorithmus für beliebig viele Marker sowie die Klärung der Frage, ob die mit Pose Estimation erreichbare Genauigkeit vergleichbar mit der eines Stereo-Kamera-Systems ist und den Anforderungen an die chirurgische Navigation gerecht wird.

Keywords: peer-to-peer navigation, pose estimation, tracking

1 Problemstellung

Optische Navigationssysteme sind bislang unterteilt in das nachverfolgende Gerät (Tool Tracker) und die nachverfolgten Geräte (Tracked Tools). Das nachverfolgende Gerät verwendet in der Regel wahlweise Zeilenkameras (z. B. Stryker FP 6000) oder Flächenkameras (z. B. NDI Polaris), die nachverfolgten Geräte beinhalten Marker (LEDs, retroreflektierende Kugeln, Schwarz-Weiß-Marker), die mit den Kameras leicht nachverfolgt werden können.

Das nachverfolgende Gerät nimmt dabei eine zentrale Rolle ein, benötigt freie Sicht auf alle an den Objekten befestigten Markern und muss daher in aller Regel aus größerer Entfernung auf die komplette Szene blicken. Dies ist insbesondere im Operationssaal oft problematisch, da der Operationssitus vom OP-Personal umstellt ist. Außerdem resultiert aus der zentralisierten Grundstruktur, dass das System keinerlei Redundanz beinhaltet. Die Relativtransformation zwischen Tracked Tool 1 und Tracked Tool 2 hängt von der freien Sicht des zentralen Navigationsgeräts auf die beiden Objekte ab. Fällt eine der beiden Strecken aus, kann die Relativtransformation nicht bestimmt werden.

Das hier vorgestellte neue Konzept eines Peer-to-Peer-Trackers hebt die grundsätzliche Trennung zwischen nachverfolgendem und nachverfolgtem Gerät auf. Dieser Tracker ist so ausgeprägt, dass er von anderen Peers leicht nachverfolgt werden kann, beinhaltet aber auch mindestens eine Kamera, mit deren Hilfe die Lage und Orientierung weiterer Tracker bestimmt wird. Bei Verwendung einer Kamera erfolgt die Bestimmung der Transformation durch Pose Estimation.

Fischler und Bolles [1] stellten 1981 einen Pose-Estimation-Algorithmus vor, der robust und iterativ die Transformation eines bekannten Modells in einen Satz von Geraden berechnet. Heutzutage werden verbesserte Algorithmen verwendet, die die Transformation von komplexeren Modellen mit mehr als drei Punkten berechnen können. Dabei spielen Geschwindigkeit, Robustheit und Genauigkeit eine entscheidende Rolle. In dieser Arbeit wird ein neuer, robuster, iterativer und dennoch sehr schneller Pose-Estimation-Algorithmus vorgestellt, der ohne initiale Schätzung von Lage und Orientierung das globale Optimum findet.

In vielen Bereichen werden Verfahren zur Bestimmung der Lage und Orientierung von Objekten benötigt, die kompakt, preiswert und schnell sind, jedoch keine zu großen Einbußen in der Genauigkeit aufweisen. Einige dieser Bereiche sind die Satellitennavigation, die Berechnung der Relativtransformation zwischen unbemannten Luftfahrzeugen wie Quadrocoptern oder die Bestimmung der Relativtransformation zwischen chirurgischen Instrumenten und der Anatomie des Patienten. Auch für Augmented-Reality-Anwendungen ist die Kenntnis der Relativtransformation zwischen realen Objekten und der einblendenden Brille von entscheidender Bedeutung.

Teixeira et. al. [2] stellen in ihrer Arbeit eine Problemstellung vor, bei der Pose-Estimation mit LEDs verwendet wird, um die Lage und Orientierung eines Quadrocopters im Raum zu ermitteln. Dabei wird darauf verwiesen, dass auch in diesem Bereich eine kompakte und kostengünstige Lösung benötigt wird. Faessler et. al. [3] nutzen ebenfalls Pose Estimation mit einer Kamera, wobei ein iterativer Algorithmus zum Einsatz kommt, welcher als Startwert die Position der letzten Berechnung verwendet. Dies kann Folgefehler erzeugen, falls die

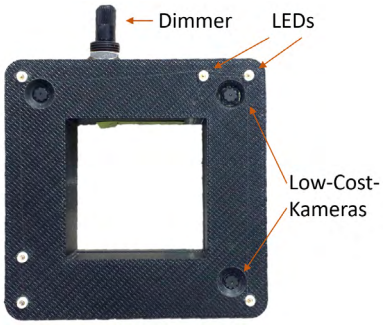


Abbildung 1: *Der entwickelte Peer-to-Peer-Tracker mit sechs LEDs und drei Low-Cost-USB2-Kameras.*

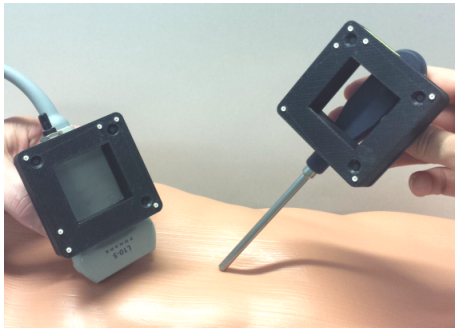


Abbildung 2: *Anwendungsfall für Biopsie mit Sonografie und zwei Peer-to-Peer-Trackern.*

letzte Berechnung ein fehlerhaftes Ergebnis geliefert hat. Auch könnte die Berechnungszeit bei einer schlecht gewählten Startposition steigen oder nicht zum gewünschten Ergebnis führen.

Die in dieser Arbeit vorgestellten Peer-to-Peer-Tracker bestehen aus mindestens einer Kamera und mindestens vier Markern, wobei zur eindeutigen Identifizierung eines bestimmten Trackers mehr als die mindestens vier notwendigen Marker-LEDs verbaut sind. Bei Verwendung einer Kamera erfolgt die Berechnung der Transformation mittels Pose-Estimation-Algorithmus, bei zwei oder mehr Kameras kann trianguliert werden. Dies führt mittels Point-to-Point-Matching zu noch genaueren Ergebnissen. Abbildung 1 zeigt den in dieser Arbeit verwendeten Tracker mit drei kostengünstigen Kameras und sechs Marker-LEDs. Abbildung 2 zeigt einen möglichen Anwendungsfall zum genauen Positionieren einer Biopsienadel relativ zu einer Ultraschallsonde.

Die Verwendung von Peer-to-Peer-Trackern hat mehrere Vorteile:

- Die Relativtransformation zwischen zwei Trackern wird in der Regel zwei Mal bestimmt. ⇒ Redundanz und größere Genauigkeit
- Ist die direkte Sicht eines Trackers auf den anderen versperrt, kann die Relativtransformation über einen oder mehrere dritte Tracker bestimmt werden. ⇒ höhere Ausfallsicherheit
- Die Relativtransformation kann je nach Anzahl beteiligter Tracker zusätzlich über Umwege bestimmt werden: 1 sieht 2 direkt, 1 sieht aber auch 3 und 3 sieht 2. ⇒ Transformation von 1 nach 2 redundant bestimmt
- Über den Aufbau von Transformationsketten kann auch die Relativtransformation zwischen Trackern bestimmt werden, die sich grundsätzlich nicht sehen können.

Von entscheidender Bedeutung ist die Tatsache, dass sehr günstige, sehr einfache Kameras verwendet werden können. Wichtig ist lediglich deren präzise Kalibrierung, die durch ein Verfahren gewährleistet ist, das ebenfalls am Labor für Computerassistierte Medizin der Hochschule Offenburg entwickelt wurde [4].

Die Kosten eines Peer-to-Peer-Trackers sind bei Verwendung von Low-Cost-Kameras gering und eröffnen insbesondere in der computergestützten Chirurgie die Möglichkeit, Einwegprodukte zu konzipieren, wodurch die Anforderung entfällt, dass die Tracker autoklavierbar bzw. sterilisierbar sind.

2 Material und Methoden

Im Zentrum der vorliegenden Arbeit steht die Frage, ob die Verwendung einer Kamera pro Beobachtungsrichtung eines Peer-to-Peer-Trackers ausreichend ist, um die Genauigkeitsanforderung an ein chirurgisches Navigationssystem zu erfüllen. Die mittlere Abweichung am Tool Center Point (TCP) soll dabei nicht größer als 1,5 mm sein, wobei eine maximale Distanz von 200 mm zwischen TCP und Schwerpunkt der Marker festgelegt wurde. Bei Verwendung einer Kamera ergibt sich die gesuchte Transformation mittels Pose Estimation bzw. Point-to-Line-Matching. Hierfür werden im Kamerabild die Zentren der LEDs bestimmt und über die Kamerakalibrierung die zugehörigen Weltgeraden berechnet. Auf diese müssen nun die im Koordinatensystem des Trackers vermessenen LED-Positionen abgebildet werden. In Kapitel 2.1 wird ein sehr schneller, iterativer Algorithmus präsentiert, der ohne Kenntnis eines Startwertes robust das gesuchte globale Minimum findet.

2.1 Pose Estimation mittels Point-to-Line-Matching

Gegeben ist eine Menge von Geraden $\vec{a}_i + s_i \cdot \hat{d}_i$ mit normierten Richtungsvektoren $\hat{d}_i$ sowie eine Menge korrespondierender Punkte $\vec{p}_i$ mit $i = 1 \dots N$. Gesucht sind die Rotationsmatrix R und der Translationsvektor $\vec{T}$ so, dass die Summe der quadrierten Abstände der transformierten Punkte $R \cdot \vec{p}_i + \vec{T}$ zu den zugehörigen Geraden minimal wird. Der Abstand c eines Punktes $\vec{p}$ zu einer Geraden $\vec{a} + s \cdot \hat{d}$ ist

$$c = \left| \vec{h} - \left(\hat{d}^T \cdot \vec{h} \right) \cdot \hat{d} \right| = \left| \vec{h} - \left(\hat{d} \cdot \hat{d}^T \right) \cdot \vec{h} \right| = \left| \left(E - \hat{d} \cdot \hat{d}^T \right) \cdot \vec{h} \right| \tag{1}$$

mit $\vec{h} = \vec{p} - \vec{a}$. Somit muss folgende Funktion minimiert werden:

$$f(R, \vec{T}) = \sum_{i=1}^{N} \left| \left(E - \hat{d}_i \cdot \hat{d}_i^T \right) \cdot \left(R \cdot \vec{p}_i + \vec{T} - \vec{a}_i \right) \right|^2 \tag{2}$$

Mit den Abkürzungen $\vec{c}_i = R \cdot \vec{p}_i - \vec{a}_i$ und $W_i = E - \hat{d}_i \cdot \hat{d}_i^T$ erhält man

$$f(R, \vec{T}) = \sum_{i=1}^{N} \left(\vec{c}_i^T + \vec{T}^T \right) \cdot W_i^T \cdot W_i \cdot \left(\vec{c}_i + \vec{T} \right) \tag{3}$$

Im Gegensatz zu R unterliegt $\vec{T}$ keinen Randbedingungen, weswegen die partielle Ableitung $\frac{\partial f}{\partial \vec{T}}$ unmittelbar verschwinden muss:

$$\frac{\partial f}{\partial \vec{T}}(R, \vec{T}) = \sum_{i=1}^{N} \left(\vec{c}_i^T + \vec{T}^T \right) \cdot W_i^T + \left(\vec{c}_i^T + \vec{T}^T \right) \cdot W_i = \sum_{i=1}^{N} 2 \cdot \left(\vec{c}_i^T + \vec{T}^T \right) \cdot W_i^T \stackrel{!}{=} \vec{0}^T \tag{4}$$

Transponieren und auflösen nach $\vec{T}$ ergibt

$$\sum_{i=1}^{N} W_i \cdot \vec{T} = -\sum_{i=1}^{N} W_i \cdot \vec{c}_i \quad \Leftrightarrow \quad \vec{T} = W^{-1} \cdot \sum_{i=1}^{N} W_i \cdot (\vec{a}_i - R \cdot \vec{x}_i) \tag{5}$$

wobei die Abkürzung $W = \sum_{i=1}^{N} W_i$ eingeführt wurde. Mit $\vec{u}_i = R \cdot \vec{p}_i + \vec{T} - \vec{a}_i$ und der einfach zu beweisenden Identität $W_i^T \cdot W_i = W_i$ lässt sich die zu minimierende Funktion $f(R, \vec{T})$ auch folgendermaßen schreiben:

$$f(R, \vec{T}) = \sum_{i=1}^{N} |W_i \cdot \vec{u}_i|^2 = \sum_{i=1}^{N} \vec{u}_i^T \cdot W_i^T \cdot W_i \cdot \vec{u}_i = \sum_{i=1}^{N} \vec{u}_i^T \cdot W_i \cdot \vec{u}_i \tag{6}$$

Es seien $\vec{w}_{i1}$, $\vec{w}_{i2}$ und $\vec{w}_{i3}$ die Spaltenvektoren der Matrizen W_i. Unter Ausnutzen von $d_{ix}^2 + d_{iy}^2 + d_{iz}^2 = 1$ lässt sich leicht zeigen, dass die nützliche Identität $W_i = \sum_{j=1}^{3} \vec{w}_{ij} \cdot \vec{w}_{ij}^T$ gilt. Eingesetzt in Gleichung (6) ergibt sich

$$f(R, \vec{T}) = \sum_{i=1}^{N} \vec{u}_i^T \cdot \left(\sum_{j=1}^{3} \vec{w}_{ij} \cdot \vec{w}_{ij}^T \right) \cdot \vec{u}_i = \sum_{i=1}^{N} \sum_{j=1}^{3} \vec{u}_i^T \cdot \vec{w}_{ij} \cdot \vec{w}_{ij}^T \cdot \vec{u}_i = \sum_{i=1}^{N} \sum_{j=1}^{3} \left(\vec{w}_{ij}^T \cdot \vec{u}_i \right)^2 \tag{7}$$

Nun kann $\vec{T}$ aus Gleichung (5) in Gleichung (7) eingesetzt werden und man erhält, die zu optimierende Funktion f, die nun nur noch von R abhängt:

$$f(R) = \sum_{i=1}^{N} \sum_{j=1}^{3} \left(\vec{w}_{ij}^T \cdot \left(R \cdot \vec{x}_i - \vec{a}_i + W^{-1} \cdot \sum_{k=1}^{N} W_k \cdot (\vec{a}_k - R \cdot \vec{x}_k) \right) \right)^2 \tag{8}$$

Mit den Abkürzungen $\vec{v}_{ijk}^T = \vec{w}_{ij}^T \cdot W^{-1} \cdot W_k$ und $k_{ij} = \vec{w}_{ij}^T \cdot \left(W^{-1} \cdot \sum_{k=1}^{N} W_k \cdot \vec{a}_k - \vec{a}_i \right)$ lässt sich die in folgender Form schreiben:

$$f(R) = \sum_{i=1}^{N} \sum_{j=1}^{3} \left(\vec{w}_{ij}^T \cdot R \cdot \vec{x}_i - \sum_{k=1}^{N} \vec{v}_{ijk}^T \cdot R \cdot \vec{x}_k + k_{ij} \right)^2 \tag{9}$$

Es sei $\vec{q}^T = \begin{pmatrix} q0 & q1 & q2 & q3 \end{pmatrix}$ das zur Rotationsmatrix R gehörende Einheitsquaternion mit

$$R = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2 \cdot (q_1 q_2 - q_0 q_3) & 2 \cdot (q_1 q_3 + q_0 q_2) \\ 2 \cdot (q_1 q_2 + q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2 \cdot (q_2 q_3 - q_0 q_1) \\ 2 \cdot (q_1 q_3 - q_0 q_2) & 2 \cdot (q_2 q_3 + q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \tag{10}$$

Ein beliebiger Ausdruck der Form $\vec{a}^T \cdot R \cdot \vec{c}$ lässt sich immer in der Form $\vec{q}^T \cdot B(\vec{a},\vec{c}) \cdot \vec{q}$ schreiben mit

$$B(\vec{a},\vec{c}) = \begin{pmatrix} a_x c_x + a_y c_y + a_z c_z & a_z c_y - a_y c_z & a_x c_z - a_z c_x & a_y c_x - a_x c_y \\ a_z c_y - a_y c_z & a_x c_x - a_y c_y - a_z c_z & a_x c_y + a_y c_x & a_x c_z + a_z c_x \\ a_x c_z - a_z c_x & a_x c_y + a_y c_x & a_y c_y - a_x c_x - a_z c_z & a_y c_z + a_z c_y \\ a_y c_x - a_x c_y & a_x c_z + a_z c_x & a_y c_z + a_z c_y & a_z c_z - a_x c_x - a_y c_y \end{pmatrix} \tag{11}$$

Eingesetzt in Gleichung (9) ergibt sich somit die finale Version der zu optimierenden Funktion:

$$f(\vec{q}) = \sum_{i=1}^{N} \sum_{j=1}^{3} \left(\vec{q}^T \cdot B_{ij} \cdot \vec{q} + k_{ij}\right)^2 \tag{12}$$

Hier wurden die symmetrischen Matrizen $B_{ij} = B(\vec{w}_{ij}^T, \vec{x}_i) - \sum_{k=1}^{N} B(\vec{v}_{ijk}^T, \vec{x}_k)$ eingeführt.

Optimierung

Gleichung (12) muss unter der Nebenbedingung $\vec{q}^T \cdot \vec{q} = 1$ optimiert werden, da $\vec{q}$ normiert ist. Der Lagrange-Formalismus liefert die zu optimierende Funktion

$$F(\vec{q},\lambda) = \sum_{i=1}^{N} \sum_{j=1}^{3} \left(\vec{q}^T \cdot B_{ij} \cdot \vec{q} + k_{ij}\right)^2 + \lambda \cdot \left(\vec{q}^T \cdot \vec{q} - 1\right) \tag{13}$$

wobei die ersten Ableitungen von $F(\vec{q},\lambda)$ verschwinden müssen:

$$\frac{\partial F}{\partial \vec{q}}(\vec{q},\lambda) = 2\lambda \cdot \vec{q} + \sum_{i=1}^{N} \sum_{j=1}^{3} 4 \cdot \underbrace{\left(\vec{q}^T \cdot B_{ij} \cdot \vec{q} + k_{ij}\right)}_{=:v_{ij}} \cdot B_{ij} \cdot \vec{q} \stackrel{!}{=} \vec{0} \tag{14}$$

$$\frac{\partial F}{\partial \lambda}(\vec{q},\lambda) = \vec{q}^T \cdot \vec{q} - 1 \stackrel{!}{=} 0 \tag{15}$$

Dieses nichtlineare Gleichungssystem kann mittels Lagrange-Newton-Iteration gelöst werden:

$$\begin{pmatrix} \vec{q}_{\text{new}} \\ \lambda_{\text{new}} \end{pmatrix} = \begin{pmatrix} \vec{q}_{\text{old}} \\ \lambda_{\text{old}} \end{pmatrix} - \begin{pmatrix} \Delta\vec{q} \\ \Delta\lambda \end{pmatrix} \quad \text{mit} \quad H_F(\vec{q}_{\text{old}}, \lambda_{\text{old}}) \cdot \begin{pmatrix} \Delta\vec{q} \\ \Delta\lambda \end{pmatrix} = \vec{\nabla} F(\vec{q}_{\text{old}}, \lambda_{\text{old}}) \tag{16}$$

Die ersten Ableitungen von $F(\vec{q},\lambda)$ sind schon bekannt (siehe Gleichung (14) und (15)) und man erhält

$$\vec{\nabla} F(\vec{q},\lambda) = \begin{pmatrix} 2\lambda \cdot \vec{q} + \sum_i \sum_j 4 \cdot v_{ij} \cdot \vec{h}_{ij} \\ q_0^2 + q_1^2 + q_2^2 + q_3^2 - 1 \end{pmatrix} \tag{17}$$

mit der für die Implementierung nützlichen Abkürzung $\vec{h}_{ij} = \begin{pmatrix} h_{ij0} & h_{ij1} & h_{ij2} & h_{ij3} \end{pmatrix}^T = B_{ij} \cdot \vec{q}$. Die Hesse-Matrix $H_F(\vec{q},\lambda)$ der zweiten Ableitungen von $F(\vec{q},\lambda)$ ist symmetrisch und man erhält

$$H_F(\vec{q},\lambda) = \begin{pmatrix} H_{00} & * & * & * & * \\ H_{10} & H_{11} & * & * & * \\ H_{20} & H_{21} & H_{22} & * & * \\ H_{30} & H_{31} & H_{32} & H_{33} & * \\ 2q_0 & 2q_1 & 2q_2 & 2q_3 & 0.0 \end{pmatrix} \tag{18}$$

mit

$$H_{kl} = \frac{\partial^2 F}{\partial q_k \partial q_l} = 2\lambda \cdot \delta_{kl} + \sum_{i=1}^{N} \sum_{j=1}^{3} 4 \cdot (2 \cdot h_{ijl} \cdot h_{ijk} + v_{ij} \cdot B_{ijkl}) \tag{19}$$

wobei es sich bei δ_{kl} um das Kronecker-Delta handelt und B_{ijkl} das Element in der k-ten Zeile und j-ten Spalte der Matrix B_{ij} ist.

Startwert der Iteration

Von entscheidender Bedeutung für die oben beschriebene Optimierung ist ein geeigneter Startwert für die Rotationsmatrix R bzw. das Quaternion $\vec{q}$, um tatsächlich das globale Minimum und nicht nur ein lokales zu finden. Den Autoren war es wichtig, hier nicht auf vorherige Berechnungen der Rotationsmatrix als Schätzwert für die neue zurückgreifen zu müssen, sondern grundsätzlich so viele verschiedene Start-Rotationen zu wählen, dass ausgehend von mindestens einem dieser Startwerte das globale Minimum gefunden wird.

Eine möglichst uniforme Verteilung dieser Startrotationen führt auf die Frage nach einer uniformen Verteilung von Einheitsquaternionen auf einer vierdimensionalen Hyperkugel. Hier bietet es sich an, die Eckpunkte eines vierdimensionalen platonischen Polychoras zu verwenden, von denen sechs zur Auswahl stehen. Bestens geeignet ist hier das vierdimensionale Pendant eines Ikosaeders: der 600-Zeller oder Hexacosichora. Seine dreidimensionale Oberfläche besteht aus 600 Tetraedern von denen sich jeweils fünf eine Kante teilen (in Analogie zu den fünf gleichseitigen Dreiecken, die sich beim Ikosaeder einen Ecke teilen). Die 120 Ecken des 600-Zellers sind $(\pm 1, 0, 0, 0)$ mit allen Permutationen, $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$ mit allen Permutationen und $(\pm\frac{\tau}{2}, \pm\frac{1}{2}, \pm\frac{1}{2\tau}, 0)$ mit allen geraden Permutationen, wobei $\tau = (1 + \sqrt{5})/2$ ist. Da $\vec{q}$ und $-\vec{q}$ jeweils die gleiche Rotation definieren, ergeben sich auf diese Weise 60 uniform verteilte Rotationen.

Es hat sich gezeigt, dass das globale Minimum bei Verwendung der so definierten 60 Startrotationen immer gefunden wird. Mit dem oben beschriebenen Point-to-Line-Matching-Algorithmus lassen sich die hieraus iterativ resultierenden 60 Endrotationen in weniger als zwei Millisekunden berechnen (Intel Core i5-6500 mit 3,2 GHz und 4 Kernen).

2.2 Genauigkeitsuntersuchung

Zur Bestimmung der Genauigkeit des Pose-Estimation-Algorithmus wird ein Stereo-Kamera-System mit Hilfe der in [4] vorgestellten Methode sehr präzise kalibriert und in ca. 300 mm Entfernung vom nachzuverfolgenden Tracker mit einer LED-Seitenlänge von 63 mm platziert. Dessen Lage und Orientierung wird nun auf drei Arten bestimmt: Pose Estimation linke Kamera, Pose Estimation rechte Kamera sowie Triangulation der Marker mit beiden Kameras und anschließendem Point-to-Point-Matching. Dies ergibt drei Transformationen F_{CT} vom Tracker-Koordinatensystem T ins Kamera-Koordinatensystem C, die identisch sein müssten.

Für jede dieser drei Transformationen kann nun die mittlere und maximale Abweichung der transformierten LED-Postionen zu den triangulierten LED-Positionen bestimmt werden. Darüber hinaus werden im Tracker-Koordinatensystem sechs Tool Center Points definiert, die jeweils 200 mm entlang der Koordinatenachsen vom Zentrum der LEDs entfernt sind. Diese werden ebenfalls transformiert und die mittlere bzw. maximale Abweichung der durch Pose Estimation bestimmten TCPs mit denen durch Triangulation entstandenen TCPs berechnet.

3 Ergebnisse

Als Ground Truth werden die mittels Triangulation bestimmten Markerpositionen verwendet. Zur Bestimmung der Genauigkeit des Ground Truth werden alle durch Triangulation entstandenen Marker-Frames gemittelt und nach Point-to-Point-Matching (Least Square Fit) mit den Messungen einer optischen Koordinatenmessmaschine verglichen:

- Die mittlere Abweichung der triangulierten LED-Positionen zu denen des gemittelten Frames beträgt 0,05 mm. Die Wiederholgenauigkeit kann somit als sehr gut bewertet werden.
- Die mittlere Abweichung der LED-Positionen des gemittelten Frames zu denen der optischen Messmaschine liegt bei 0,04 mm. Somit ist auch die Absolutgenauigkeit gewährleistet.

Die Genauigkeit des Ground Truth ist somit gesichert. Die Abweichungen der transfomierten LED-Positionen zu den triangulierten LED-Positionen sind in Tabelle 1 zusammengefasst.

Tabelle 2 zeigt die Abweichungen der TCPs, die sich mit den aus Pose Estimation berechneten Transformationen ergeben, von denen, die mit Hilfe der Transformation berechnet werden, die aus dem Point-to-Point-Matching (Least Square Fit) folgt.

4 Diskussion

Die Ergebnisse aus Kapitel 3 legen nahe, dass die mittels Pose Estimation bzw. Point-to-Line-Matching mit einer Kamera bestimmten Transformationen genau genug sind, um das in Kapitel 2 gesteckte Ziel einer mitt-

Tabelle 1: *Abweichungen der transformierten LED-Positionen von den triangulierten LED-Positionen*

	Mittlere Abweichung	Maximale Abweichung
Point-to-Point Kamera 1 und 2	0,06 mm	0,19 mm
Pose Estimation Kamera 1	0,45 mm	0,92 mm
Pose Estimation Kamera 2	0,45 mm	0,72 mm

Tabelle 2: *Abweichungen der mittels Pose Estimation transformierten TCPs von den mittels Point-to-Point-Matching bestimmten TCPs*

	Mittlere Abweichung	Maximale Abweichung
Pose Estimation Kamera 1	0,69 mm	1,62 mm
Pose Estimation Kamera 2	0,54 mm	1,59 mm

leren Abweichung von 1,5 mm am TCP (in 200 mm Entfernung vom Zentrum der sechs LEDs) erreichen zu können. Die bisherigen Ergebnisse beziehen sich auf eine durchaus realistische Entfernung von 300 mm zwischen Kamera und LEDs, wurden jedoch mit einer Industriekamera (Ximea MQ013MG-E2 mit 60 fps und 1280 x 1024 Pixeln) erzeugt, um einen verlässlichen Ground Truth zu erhalten. Weitere Untersuchungen werden zeigen, ob die im Tracker selbst verbauten Kameras (Vimicro Venus USB2.0, siehe Abbildung 1) vergleichbare Ergebnisse liefern und somit die Verwendung einer Kamera pro Tracker und Beobachtungsrichtung ausreichend ist. Dies wäre insbesondere für Tracker wünschenswert, die in mehreren Richtungen nach „Peers“Ausschau halten. Der vorgestellte Algorithmus für Pose Estimation mittels Point-to-Line-Matching funktioniert insbesondere im Hinblick auf das Auffinden des globalen Minimums sehr robust. Die Verwendung der 60 vorgestellten Startrotationen ist zeitlich unkritisch, hat sich bewährt und sollte auch für weitere iterative Probleme mit unbekannter Startrotation in Betracht gezogen werden.

5 Zusammenfassung

In dieser Arbeit wurde ein neues Peer-to-Peer-Tracking-Konzept für die chirurgische Navigation vorgestellt. Die Trennung zwischen nachverfolgendem und nachverfolgtem Gerät wurde aufgehoben und die Frage erörtert, ob eine Kamera pro Beobachtungsrichtung eines Peer-to-Peer-Trackers ausreichend ist, um die mittels Pose Estimation berechneten Transformationen mit ausreichender Genauigkeit bestimmen zu können. Das vorgestellte Konzept löst das in der Problemstellung beschriebene Line-of-Sight-Problem und erhöht die Ausfallsicherheit durch mehr Redundanz. Die geringen Herstellungskosten eines solchen Peer-To-Peer-Trackers legen nahe, diesen als Einwegprodukt zu konzipieren.

Darüber hinaus wurde ein sehr robuster (durch Verwendung von 60 geeigneten Startrotationen) und schneller (weniger als 2 ms pro Transformation auf einem Intel Core i5-6500 mit 3,2 GHz und 4 Kernen) Point-to-Line-Matching-Algorithmus vorgestellt, der trotz seines iterativen Ansatzes ohne initiale Schätzung auskommt und dennoch das globale Minimum findet.

6 Danksagung

Diese Arbeit wurde unterstützt von der Stryker Leibinger GmbH & Co. KG, Freiburg, Deutschland.

7 Referenzen

[1] Fischler, Bolles, *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*, Communications of the ACM Volume 24 Number 6 (1981)

[2] Teixeira, Maffra, *VI-RPE: Visual-Inertial Relative Pose Estimation for Aerial Vehicles*, IEEE Robotics and Automation Letters (Early Access), (2018)

[3] Faessler, Mueggler, *A Monocular Pose Estimation System based on Infrared LEDs*, IEEE International Conference on Robotics & Automation (ICRA), (2014)

[4] Hoppe H., Seebacher F., Klemm M., *Nicht modellbasierte Kalibration von Kameras mit Monitoren*, Proceedings Bildverarbeitung für die Medizin (BVM) 2016, S. 50-55, Berlin, 13.-15-03. (2016)

Comparison of Auditory Display Methods for Real-time Catheter Navigation

D. Black[1,2], Ewald Schmidt[2,3], J. Strehlow[1,2], T. Pätz[2], R. Rascher-Friesenhausen[2,3],

[1] Jacobs University, Bremen, Germany

[2] Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen, Germany

[2] University of Applied Sciences, Bremerhaven, Germany

Contact: david.black@mevis.fraunhofer.de

Abstract:

During endovascular navigation, fiber-optic curvature sensors can help inform clinicians about the position of a catheter in a body. This helps reduce radiation exposure and limits the use of X-ray contrast agents compared to existing methods. We propose three auditory feedback methods to transmit the distance to the next vascular bifurcation and angle during navigation in the hopes of increasing usability of novel wire-guided catheter systems and reducing time spent viewing screens. The results of a user evaluation show that such methods can be used for screen-free guidance during a task involving correctly navigation paths and angles of rotation.

Keywords: auditory display, catheter navigation, computer-assisted interventions, sonification, vascular system, human-computer interaction, user experience.

1 Problem

During catheter interventions, a patient's vascular system is used as an access path to reach target structures. Catheter or guide wires are inserted at a suitable entry point and navigated through the vascular system. The position and surroundings of the catheter are usually displayed using X-ray projections so that the navigation can be monitored. The vessels through which the catheter will be navigated are, however, not usually seen in these images without further support. To temporarily make the vessels visible, the clinician can inject a small amount of X-ray contrast agent into the patient. However, the complexity of the vascular system and its many branches provide difficulties for navigation to the target lesion, as depicted in Figure 1. The cognitive

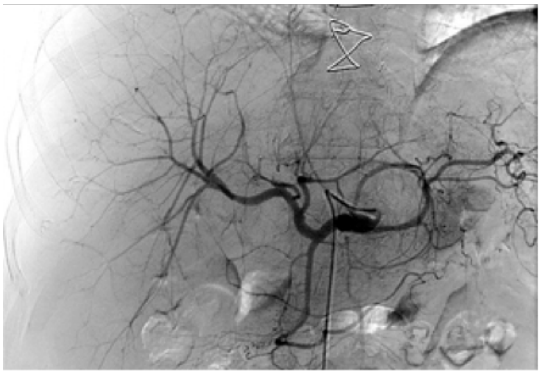


Figure 1: *Contrast-enhanced angiography of liver arteries*

fusion of a 2D vessel tree projection to a 3D vessel model and the projection of the catheter present a high burden for the clinician, and in addition, is not always clear. For the patient, this approach also entails the side effects of exposure to radiation as well as contrast agents [1, 2].

Approaches exist, however, that simplify navigation for clinicians and reduce patient exposure to radiation and contrast agents. Using preoperative or intraoperative images, a spatial model of the patient's vascular system can be generated. In addition, a catheter can be equipped with fiber-optic curvature sensors. During an intervention, comparing sensor data with the spatial model of the vascular system can determine the position of the catheter in the spatial model [3, 4, 5], see Figure 2. Thanks to this calculation, novel navigation techniques can be used by the clinician. This type of navigation is especially suitable for scenarios in which preinterventional imaging is a part of the clinical workflow and for which navigation is especially demanding. This could include, for instance, radioembolization of liver tumors, where the clinician must find a segment of the arterial vascular tree, which almost entirely supplies the tumor with blood. In these tedious operations, time could be saved and interventions could be made easier to plan.

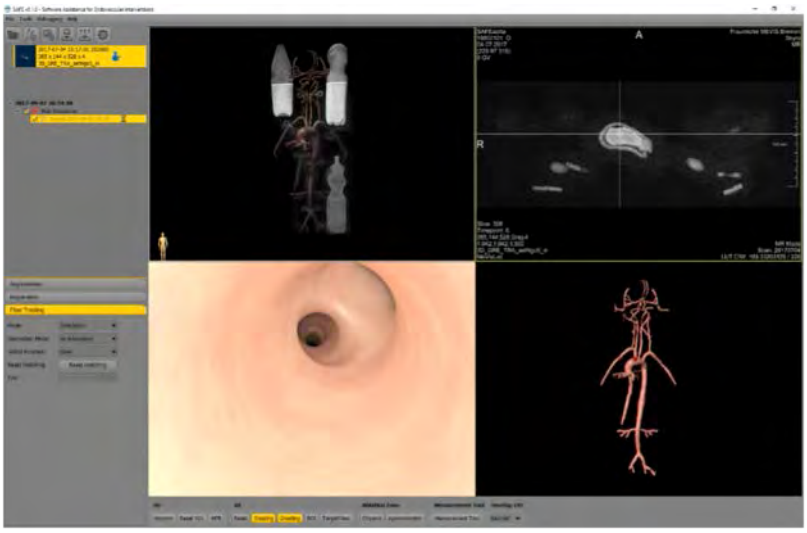


Figure 2: *Prototypical integration of intelligent catheter position using wavefront curvature sensor, showing sample 3D view in bottom left and vasculature in bottom right*

In this work, we describe an approach that attempts to improve the shortcomings of standard catheter navigation by implementing a so-called auditory display to aid the clinician in novel guide-wire catheter navigation. An auditory display is one that uses sound (as opposed to visualizations) to transmit information. Auditory displays in medical image-guided navigation are scarce, but attempts described in the literature have shown benefits including enhanced recognition of the distance to structures or targets, improved placement accuracy, reduced cognitive workload, and reduced rates of clinical complication. Drawbacks reported in the literature include increased task time and increased non-target tissue removed during resection. For a thorough review of the state of research in auditory displays for medical image-guided navigation, see [6]. We hypothesize that an auditory display for guide-wire-based catheter navigation could help improve inter-

action with these novel systems by transmitting the distance of the catheter tip to the next bifurcation in the vascular system tree as well as the correct angle so that the catheter may be properly guided into sequential branches of the vasculature. In such a scenario, the clinician could be informed when nearing the next bifurcation as well as avoiding incorrect branches. This could reduce time, increase accuracy, reduce cognitive load of the clinician, and help protect the patient by minimizing harmful radiation and contrast-agent exposure.

2 Materials and Methods

This work compares three different auditory display methods for navigating points on a path and angle of rotation and describes the results of a laboratory evaluation in terms of task completion time, accuracy, system usability, and cognitive load. The test setup consisted of a computer gaming joystick (Logitech Attack 3) as an input device used to simulate catheter navigation, with x- and y-ranges of 0 to 255. This provides virtual navigation such catheter paths such as those depicted in Figure 3, where green indicated a correct path and red indicates false paths to be avoided. For both path length and angle of rotation, values of 0 to 255 were mapped to a set of discrete auditory display parameters to provide the participants the three real-time synthesis methods that could be 'played' so that changes in path distance and angle of rotation of the catheter could be heard. The three evaluated auditory display methods cover a wide range of sound-design techniques:

FM: The first auditory display synthesis method consisted of a mapping of path length and angle of rotation to fundamental frequency of a sine wave and subsequent frequency modulation (FM) of the target tone. The base tone consisted of a sine wave with a fundamental frequency mapped to the difference of the position of the joystick to the difference of the target point, between 0 and 255 Hz. Inside of a margin of -20 to 20 between base and target, the base tone was turned off and a target signal was played which consisted of a frequency-modulated [7] tone with a modulation frequency of 12 Hz and a carrier frequency of 56 Hz.

SUB: The second auditory display synthesis method utilized subtractive synthesis (SUB). In this method, the base tone consisted of a sawtooth wave with a fundamental frequency of 100 Hz. When approaching the target position, the center frequency of two sequential band-pass filters with a Q factor of 2 were mapped from 0 to 255 Hz. In contrast to the method employing FM, this method does not use a separate target tone, but rather relies only on changes in fundamental frequency of the sawtooth wave as well as center frequency of the band pass filters.

MSS: Finally, the third auditory display method consisted of a multiple sound synthesis (MSS). The base tone, similarly to FM, consisted of a single sine wave mapped to the range of the input device (0 to 255). The target tone was created using two alternating sine waves with differing fundamental frequencies. Between path length and angle of rotation distances between -20 and 20, the target tone was played. The two sine waves, with fundamental frequencies of 200 Hz and 100 Hz were alternated each 500 ms.

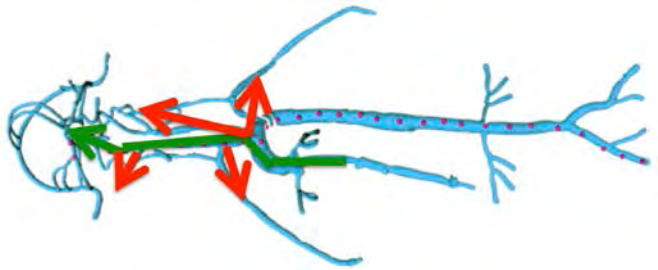


Figure 3: *Example vascular system navigation, where green indicates correct path and red indicates incorrect branches*

The task of the participant was to position the joystick, first in the y-direction (forwards and backwards) to locate a randomized point along the path length, and then to move the joystick in the x-direction (left and right) to locate the correct angle of rotation. After a brief (<5 minute) training period and introduction to each of the three auditory display synthesis methods (FM, SUB, and MSS), the participants completed the set of tests. The task of finding a correct path placement and angle of rotation was repeated twice for each auditory display method, thus resulting in a total of 12 data points for each participant. The trigger on the joystick was used to activate and deactivate the recording of the path length and angle of rotation. The studies took place in an enclosed laboratory. A pair of over-ear headphones were provided for each participant. Participants were seated during the entire evaluation procedure. Participants were permitted to adjust chair and joystick position on the testing table. The evaluation was completed without the use of a screen. After the set of tasks were completed, for each auditory display synthesis method, each participant completed both the van der Laan system acceptance questionnaire [8] as well as the NASA-TLX multidimensional assessment questionnaire [9] that rates perceived workload. No compensation was given.

3 Results

The evaluation was completed by 10 (7 male, 3 female) novice participants with an average age of 27.9 years recruited from a local research institute and all had backgrounds in fields of medical engineering. For each of the auditory display synthesis methods (FM, SUB, and MSS) and each path length and angle of rotation task, the percentage of reached point (RP) as well as task completion time were recorded. For the average results of all participants for each of these measures as well as the results of the qualitative questionnaires, please see Figure 4.

To help determine the accuracy of each method, we recorded the number of reached target points for both path and angle of rotation. X- and y-position for each dimension was recorded at intervals of 100ms. If all of the last 5 samples were in the range of -20 to 20 of the target point or angle of rotation, when the participant released the joystick trigger this was considered as having found the target. The participants located the desired path length in 85% using FM, 100% using SUB, and 75% using MSS. In contrast, for angle of rotation, 95% were found using FM, only 20% using SUB, and 75% using MSS. Thus, the overall accuracy in terms of number of reached points was highest using FM (90%), and lowest using SUB (60%). The task completion time was calculated as the elapsed time between joystick trigger presses. Participants needed an average

of 26.76 seconds to find each path point using FM, 17.29 using SUB, and 22.84 using MSS. Results were similar for angle of rotation: Participants needed an average of 30.39 seconds to find each path point using FM, 12.52 using SUB, and 21.61 using MSS. Thus, the FM method required the most amount of time, but provided more accuracy. SUB provided more speed and high accuracy for path point location, but poor accuracy for finding angle of rotation.

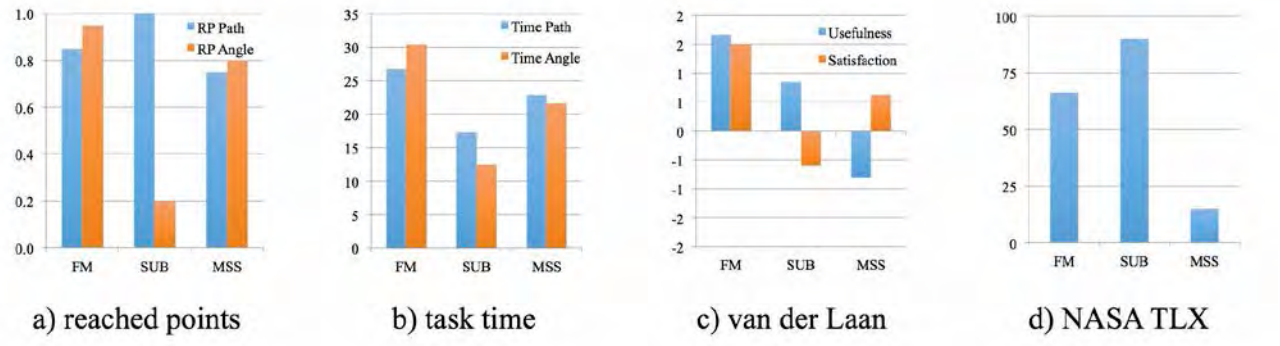


Figure 4: *Graphical display of average results over all participants. Quantitative results show average percentage of reached points (a), task completion time (b) in seconds for path and angle location. Qualitative results show Van der Laan system usability scale (c) with subfactors of usefulness and satisfaction ranging from -2 (poor) to 2 (best) and NASA Raw-TLX questionnaires (d) from 0 (low cognitive load) to 100 (high cognitive load).*

The results of the qualitative questionnaires were given as scores of the van der Laan system acceptance scale and NASA TLX. The van der Laan scale includes 9 pairs of adjectives including “undesirable / desirable” or “nice / annoying” to generate composite ratings for usefulness and satisfaction. The NASA-TLX multidimensional assessment questionnaire queries the participants’ self-reported cognitive workload experienced while completing the task. For system acceptance, ranging from -2 (worst) to 2 (best), the FM method was rated highest, with average scores of 1.6 for usefulness and 1.50 for satisfaction. SUB had an average of 0.85 for usefulness and -0.60 for satisfaction. MSS had an average of -0.80 for usefulness and 0.63 for satisfaction. The results of the NASA-TLX multidimensional assessment questionnaire, which ranges from 0 (low cognitive load) to 100 (high cognitive load), show that MSS was rated with the lowest load (15.00). FM was rated at 66.25, and SUB at 90.00.

4 Discussion

We performed an evaluation of three auditory display synthesis methods to gauge whether participants could locate both points along a path as well as angles of rotation without using a computer screen as a means of navigating a catheter for novel guide-wire applications. The results of the quantitative measures as well as the qualitative van der Laan and NASA-TLX questionnaires reveal great differences in the performance, usability, and perceived cognitive load of each of the three auditory display synthesis methods.

In terms of quantitative performance measures, participants found the highest percentage of targets using the FM method. This was accompanied by a substantially higher task completion time, the highest of all methods. Results using SUB showed a large difference in task time for path (100%) and angle of rotation (20%), showing that the method might be useful for path target locating but less so for angle of rotation. However, participants exhibited the lowest task completion times when using SUB. Finally, average results of MSS were between those of FM and SUB in terms of reached points and task completion time. In terms of qualitative measures gathered from the van der Laan and NASA-TLX questionnaires, the methods were also widely different. FM received the best usefulness and satisfaction scores, even though the results of the perceived cognitive load were much higher than using MSS. SUB was rated as somewhat useful, but less satisfying, whereas MSS was rated as somewhat satisfying but less useful. SUB received the worst NASA-TLX scores, showing that the method is not well suited for a task that should be completed by clinicians in an already highly cognitively demanding scenario. Overall, it is clear that the FM synthesis method provides the best tradeoff between accuracy, task completion time, and qualitative results.

We believe that with an extended amount of training, the FM method could be adapted to use in a clinical scenario: with extended training, the perceived cognitive load and task completion time would most likely be reduced. In addition, the combination of the auditory display method with existing novel navigation aids, such as those seen in Figure 2, require careful attention in terms of integration and synchronization. The auditory display must be adapted to other aids in such a way that neither detracts from the other; a hybrid audiovisual feedback would be ideal. The interaction with such hybrid audiovisual displays would also exhibit different results when used with a physical wire-guided catheter instead of the computer joystick as input device features in this evaluation.

5 Conclusion

In this work, we have shown that auditory display can be a useful modality of feedback for a virtual novel wire-guided catheter application. The evaluation shows that path points and angles of rotation of a virtual catheter scenario can be found without the use of the screen, using a joystick as a virtual input device for a wire-guided catheter. The developed methods could assist, for instance, the clinician in avoiding incorrect paths and by informing when an upcoming bifurcation is about to be reached. Although the results presented in this work have not been combined with existing visual feedback methods, we believe that such a combination could provide even higher benefit to the clinician by harnessing the information of the wire-guided catheter system without having to keep the view on the system screen during an entire intervention. Future work will include enhancing the employed methods to further reduce cognitive load, increase accuracy, and reduce task completion time. This should help clinicians better understand the planned path, help execute the navigation while reducing difficulty and increasing confidence, as well as limiting patient exposure to radiation or contrast agents.

References

[1] M. Ingwersen, A. Drabik, U. Kulka, U. Oestreicher, S. Fricke, H. Krankenberg, C. Schwencke, D Mathey, *Physicians' Radiation Exposure in the Catheterization Lab: Does the Type of Procedure Matter?*, JACC: Cardiovascular Interventions 6(10) 1095-102 (2013)

[2] M. Andreassi, E. Piccaluga, G. Guagliumi, M. Del Greco, F. Gaita, E. Picano, *Occupational Health Risks in Cardiac Catheterization Laboratory Workers*. Circulation: Cardiovascular Interventions, *9*(4) e003273 (2016)

[3] J. Strehlow, T. Pätz, N. Papenpberg, J. Georgii, A. Schreiber, H. Hahn. *Catheter Tracking based on Catheter Curvature and Vessel Anatomy,* Proceedings of the 3rd Conference on Image-Guided Interventions & Fokus Neuroradiologie, Magdeburg, Germany (2017)

[4] R. Trivisonne, E. Kerrien, S. Cotin. *Augmented 3D Catheter Navigation using Constrained Shape from Template.* Hamlyn Symposium, London, United Kingdom (2017)

[5] K. Mandal, F. Parent, S. Martel, R. Kashyap, S. Kadoury. *Vessel-based registration of an optical shape sensing catheter for MR navigation*. International Journal of Computer Assisted Radiology and Surgery, 11(6) 1025-34 (2016)

[6] D. Black, C. Hansen, A. Nabavi, R. Kikinis, H. Hahn, *A Survey of Auditory Display in Image-Guided Interventions*, International Journal of Computer Assisted Radiology and Surgery, 12(10) 1665-1676 (2017)

[7] J. Chowning, *The Synthesis of Complex Audio Spectra by Means of Frequency Modulation*, Journal of the Audio Engineering Society, 21(7) 526-534 (1973)

[8] J. Van Der Laan, A. Heino, D. De Waard, *A simple procedure for the assessment of acceptance of advanced transport telematics*, Transportation Research Part C: Emerging Technologies 5(1) 1-10 (1997)

[9] S. Hart, L. Staveland, *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,* Advances in Psychology 52 139-183 (1988)

A System for Augmented Reality Guided Ventricular Puncture Using a HoloLens: Design, Implementation and Initial Evaluation

Christian Kunz[1], *Michal Hlavâc*[2], *Max Schneider*[2], *Björn Hein*[1], *David Puljiz*[1], *Steffen Peikert*[1]

[1] *Karlsruhe Institute of Technology, Institute for Anthropomatics and Robotics — Intelligent Process Automation and Robotics, Karlsruhe, Germany*
[2] *Department of Neurosurgery, University of Ulm, Günzburg, Germany*

Contact: christian.kunz@kit.edu

Abstract

In the treatment of patients surgical assistance systems are becoming increasingly established worldwide. The idea is to provide the surgeon with suitable assistance at the right time during the ongoing surgical procedure. We propose a surgical assistance system that supports the surgeon during ventricular puncture by visualizing the target area using a Microsoft HoloLens^{TM}. Preoperatively an individual patient head model is reconstructed from computer tomography (CT)-data and a marker is attached to the patient, which is tracked during the procedure. The visualized target region must be positioned as accurately as possible during the surgical procedure to provide safe guidance for the surgeon. To ensure this we register the marker system to the patient's head by using an Iterative Closes Point (ICP) based surface matching algorithm. The presented system will be examined in detail with regard to the render performance of the HoloLens. Furthermore, a qualitative online analysis is performed. The results show that the Augmented Reality (AR)-headset can be successfully used to visualize medical data and that the patient can be successfully tracked at runtime.

Keywords: Computer Assisted Surgery, Surgical Assistance, Augmented Reality, Ventricular Puncture, Surface Matching, Marker-based Tracking

1 Problem

Computer-assisted procedures in medical technology should support the surgeon in the treatment of patients in order to improve the outcome for the patient. It is important that such systems are seamlessly integrated into the medical workflow through user-centered development and that a high degree of usability is ensured. If this is not the case, their use can fail due to the lack of acceptance or even the safety of the procedure can be compromised. The ventricular puncture of the human brain for the drainage of cerebrospinal fluid is one of the routine procedures in neurosurgery. Across Germany, about 20 000 punctures are performed either as an independent procedure or as part of a more complex surgery per year. The puncture is carried out via a burr hole trepanation at the Kocher's point. The direction of the catheter is determined using external anatomical landmarks. In addition, the anatomical and pathological characteristics of the patient must be taken into account by the surgeon. This procedure is often indicated in emergency situations, such as brain hemorrhages, traumatic brain injury or strokes. The ventricle is about 4 cm deep, has an elongated shape with a width of about 1 cm, which can vary greatly due to the underlying disease. Optimal catheter placement is achieved in only about 60% of cases [1], sometimes requiring multiple puncture attempts. The difficulty of the intervention is that the determination of the penetration angle often allows only a few degrees of tolerance and the target cannot be visualized directly.

Currently there are only limited options for providing assistance for the procedure. The commercially available navigation systems require complex preparation usually with rigid head fixation, manual planning and registration as well as additional navigation instruments. This significantly increases the amount of time and resources needed for the procedure. Other supportive tools like the Thomale guide also require presurgical planning, increase complexity and potential risks. They only achieve a slightly higher optimal catheter position in approximately 70% of cases [5]. The success rate is to be significantly improved by displaying the target area using a 3-D overlay.

2 Material and Methods

It is the basic idea of our system to provide the surgeon with an easy-to-use, exact and robust interface for surgical assistance. The system should be integrated as seamlessly as possible into the surgical workflow. This

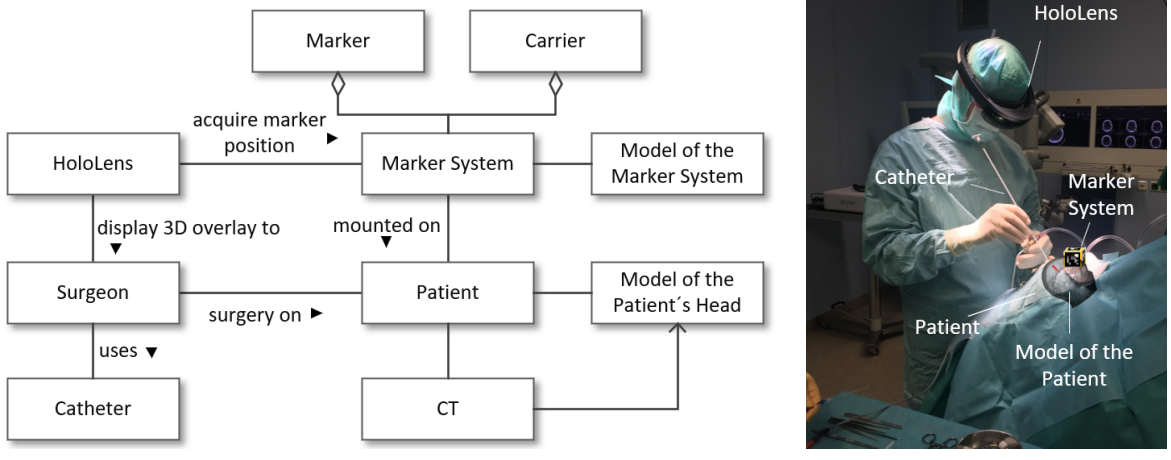
contributes significantly to user acceptance. The assistance system should only run on and depend on the AR-headset. Expensive external sensors should not be used for environmental sensing.

The surgeon puts on the AR-headset before the surgery (see figure 1 b) and carries out all necessary steps with it. Before covering the patient's head in sterile drapes a fully automated registration of the marker system to the patient is carried out. This means that the position of the marker system on the patient's head is determined precisely. The patient can then be covered for the operation. From this point on, it is only necessary to track the marker to overlay the 3D patient model with the real world correctly.

2.1 Experimental Setup

We use a phantom head from Synbone (www.synbone.com) mounted on a stewart platform from PI (www.pi-usa.us) on a solid measuring table made of stone (the head is displayed in figure 3 a). A Microsoft HoloLens is used as AR-platform to display the 3D-overlays and to track the environment. The HoloLens employs the Unity 3D graphics engine to visualize the holographic scene. The computer vision library OpenCV and the ARUCO library is used for marker tracking. In addition, a RGB-D camera, an Intel Realsense SR300, was used as a sensor for pre-testing. It was positioned firmly on a tripod. When Microsoft released the research mode for the HoloLens, access to the raw data streams of the different sensors, e.g. depth stream and environmental cameras, became available. Before that, access to the data streams were very limited. It was only possible to access the camera in the center of the headset and a spatial map.

In figure 1 (a) and (b) the system setup is depicted. The marker system consists of a carrier and the actual marker. Initially ARUCO markers [2] are used in our system. The marker system is attached to the patient's head and fixed with an adhesive pad. Usually, neurosurgical head fixation and navigation is an invasive and time consuming procedure, in which metal pins are pierced through the skin into the cranial bone. Since ventricular punctures are often carried out in time critical emergency situations, this can be undesirable. The human head is relatively stiff and immobile at its surface. It will be further examined how well a fixing method using adhesive pads works and whether this is sufficient for the present case. Initial tests were carried out with a marker adapted to the patient's head by using its original surface as the base of the carrier. This allows a very accurate fit on the patient's head. Since the head does not move much during the surgery the marker system is very promising.



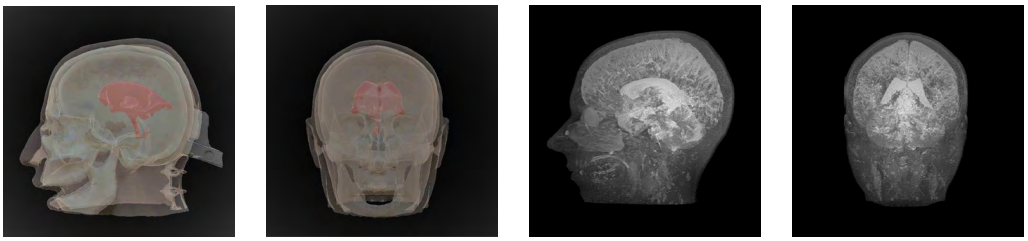
(a) (b)

Figure 1: Setup of the system.

2.2 The Surgical Assistance Workflow

Preoperatively a CT-scan of the patient is carried out and individual 3D-models of the patient's head are generated: A volume model and a surface model (figure 2). Each voxel in the volume model has an individual tissue value. During rendering each type of tissue is rendered with a different color value, usually as a grayscale value. This allows the volume model to be segmented, e.g. to display the ventricle system. The volume model can be used for preoperative analysis of the patient's anatomy and pathology. It can be viewed within the HoloLens, although we do not achieve 60 fps. Therefore it is not used during the surgery as an overlay due to the limited resources of the HoloLens. The surface model consists of three single surface models:

The ventricular system, the cranial bone and the skin. During the surgery the surface model is overlayed with the patient's head. The surface models need much less performance to be rendered and all the needed informations are also presented.



(a) (b) (c) (d)

Figure 2: Surface models of the Synbone phantom head (a, b) and volume models of a real patient (c, d).

At the beginning of the surgery preparation the marker is attached to the patient's head. We use an ARUCO based tracking. The markers are generated as described in [3]. We use a small marker dictionary consisting of 50 markers. It is planned in the future to use more than two markers to stabilize tracking. This allows the tracking to be averaged over several values to compensate for fluctuating detections. In our system we have multiple observable coordinate frames (see figure 3) that must be registered to each other: the HoloLens frame T_H, the marker system frame T_{MS} and the patient model frame T_{P_m}. The patient frame T_P is synchronized with T_{P_m} to create a coordinate system in which the patient is located. Additionally there is the catheter tool frame T_C. The catheter is not tracked at the moment. The surgeon is responsible for the insertion.

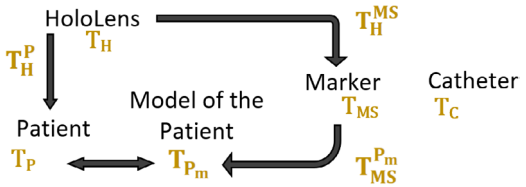


Figure 3: Registration of the coordinate frames of the system.

To be able to overlay the virtual patient head precisely during the surgery, the marker system must be correctly registered to the real patient head. The process is depicted in figure 4.

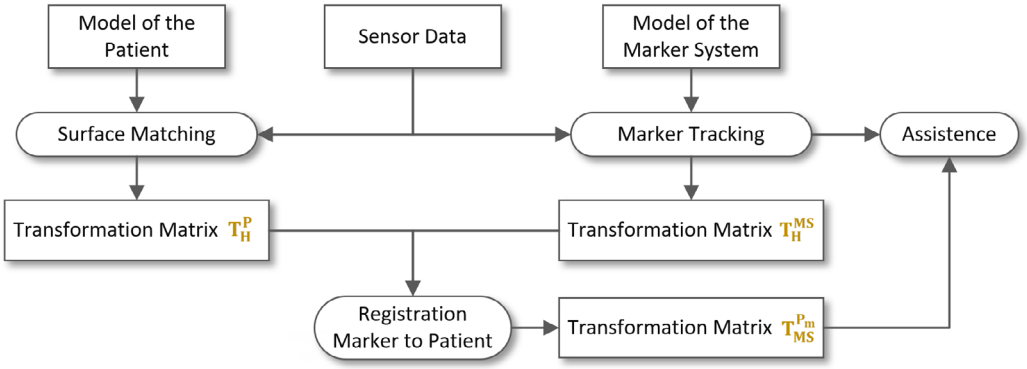


Figure 4: Activity diagram of the AR-based surgical assistance workflow.

This requires to determine the 3D-pose of the marker system and the patient's head in the scene. The marker is detected directly through ARUCO tracking. The result is a transformation matrix T_H^{MS} in the coordinate frame of the HoloLens camera. Next the translation and rotation of the patient's head is determined. We do so by using a Surface matching algorithm. We have investigated various variants and their implementations: the OpenCV implementation of Picky Iterative Closest Point [7], the Point Cloud Library and ICP [4] and 'Fast global registration' [6]. All three approaches yielded good results. We get the transformation matrix T_H^P in the coordinate frame T_H. When the localization of the marker system and the patient's head are found in the scene

the relative transformation from marker system to the patient's head can be calculated using formula 1.

$$T_{MS}^{P_m} = (T_H^{MS})^{-1} * T_H^P \tag{1}$$

During the surgery the patient is tracked only through the marker multiplying the transformation matrix $T_{MS}^{P_m}$ with the position of the marker to determine the position of the patient's head.

2.3 The HoloLens Deployment Process

In order to visualize the two different model types on the HoloLens they need to be imported into a Unity 3D project. The surface models can be visualized using the default shader. Volume models cannot be displayed by default, so a Ray Trace Shader has been programmed to do this. Volume files are provided in the '.nrrd' data format. These files are read and converted to an internal binary data format as a three dimensional array. Once the volume has been converted, a memory dump is created and used as an asset in Unity 3D projects. The finished project can then be transferred to the HoloLens.

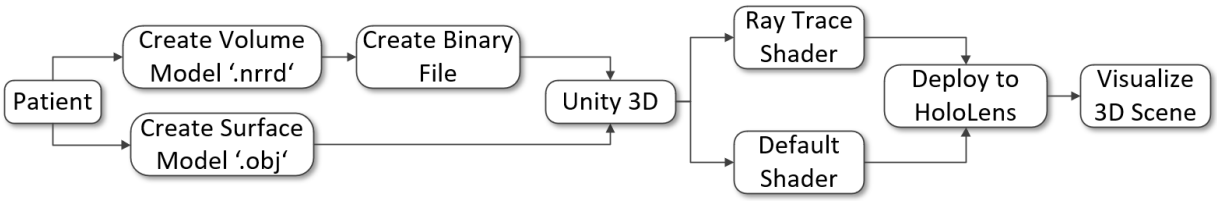


Figure 5: The process to deploy and render 3D-models on the HoloLens.

3 Results

We prove the feasibility of the proposed design through initial implementations and evaluations. We carried out detailed performance measurements with the HoloLens to show which possibilities the headset provides for the presentation of medical data. We used different models to evaluate the render performance. Tables 1 and 2 list the different models. Surface models of the patient were provided in the '.obj' file format. Blender (v.2.79b) was used to reduce the vertex and face count.

	Model 1 (real patient data)				Model 2 (Synbone phantom head)			
	Skin	Brain	Ventricle	Total	Skin	Brain	Ventricle	Total
Vertex count	4 595	8 103	4 068	16 766	17 633	42 310	5 977	65 950
Triangle count	8 942	16 255	8 149	33 346	34 994	83 805	11 922	130 721
Size (in kB)	280	500	248	1 028	1 154	2 839	369	4 362

Table 1: For evaluation 2 combined surface models were used (skin, brain and ventricle). Model 1 was generated from real patient data. Model 2 was generated from a CT-scan of the Synbone phantom.

Volume files were provided in the '.nrrd' file format. Because applications running on the Hololens have a memory limit of 900 MB, we were not able to use the original head model. Although our app with the full-scale model requires only 235 MB of memory, Unity 3D allocates up to 950 MB of RAM causing our application to crash. We then reduced the size of the volume models. Resampling was done using the 'unu' resample tool with default parameters.

	Head 100%	Head 75%	Head 50%
Dimension (in voxel)	512 x 160 x 512	384 x 120 x 384	256 x 80 x 256
Voxel size (in mm)	0.488 x 1.000 x 0.488	0.653 x 1.333 x 0.653	0.980 x 2.000 x 0.980
Binary file size	160 MB	67.5 MB	20 MB

Table 2: Volume models: The original volume model was generated from a CT-scan of a real patient. The 75% and 50% reduced models were generated out of the original one.

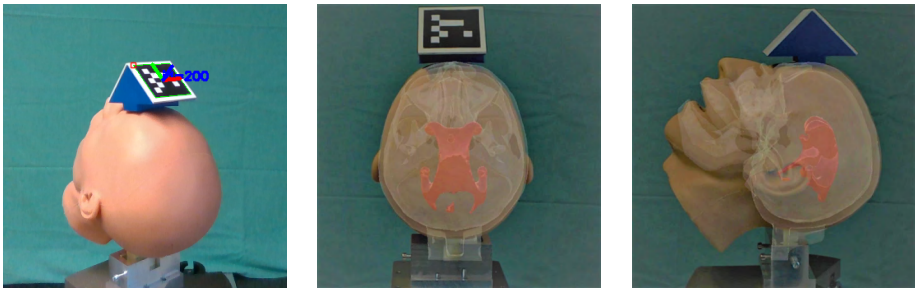
All measurements were recorded with the built-in performance tracker of the HoloLens. For the performance analysis a side view of the displayed head was chosen. All values (with the exception of RAM) are the average

of a one minute measurement interval once the scene was loaded. Since no memory allocation is necessary once the scene is active, the given values for RAM utilization were taken as soon as the scene was loaded.

	Volume models					Surface models	
	Head 100%	Head 75%		Head 50%		Model 1	Model 2
Samples per ray	-	16	96	16	96	-	-
FPS	-	54	13.5	52	13	60	46
CPU utilization (in %)	-	11	3	10	2.5	12	9
RAM (commit size) (in MB)	-	507.2	509.4	271.3	271.1	162.8	164.7
SoC power utilization (in %)	-	100	95	61	61	97	116

Table 3: Evaluation of the render performance of the volume models.

The results give an intial overview of what can be displayed on the HoloLens. The 50% reduced volume model runs at a lower fps as the 75% volume model. The fps number depends more on the render pipeline than on the actual size of the models. We will further improve the pipeline in future work. However, the visualization of volumetric data requires a lot of performance and we achieve good results with 54 fps. The surface models can be visualized with 60 fps. In figure 6 a first qualitative analysis is shown. The results are promising.



(a) (b) (c)

Figure 6: (a) ARUCO based marker tracking, (b) 3D-Overlay in cranio-caudal direction, (c) 3D-overlay from lateral direction.

Further evaluations will be carried out in the future: The registration of the marker system to the patient's head, the accuracy of the ARUCO marker tracking, as well as the stability of the holographic overlay.

4 Discussion

We have described an optimized process for surgical assistance in the scenario of a ventricular puncture. Our results show that the concept of using the HoloLens is capable of displaying medical data and overlay it with a real patient during an operation. However, several process steps (see figure 4) need to be revised and need to be made more robust.

It has been shown that the preprocessing steps are error-prone. The model files need to be in the right format and size and must not have any errors. This requires a lot of manual intervention. An automated solution is targeted. The marker system we have initially used must be further improved and adapted to the requirements of the surgical procedure. Adhesive marker systems are already in use with commercial navigation systems and achieve a very good precision (e.g. the Stryker Cranial Map system (www.stryker.com)). These systems use external camera systems, which can be adjusted, to achieve optimal tracking. Our hardware consists only of the HoloLens, which means the marker system has to be adapted to allow tracking from behind the patient, where the surgeon is positioned. We achieve this by using a marker system that comprises multiple 2D markers that are visible from any point of view. In addition, a solution must be found for transferring the marker system from the non-sterile environment of the initial registration to the sterile environment of the surgical field. This is also the usual process in commercial navigation systems requiring a mounting system which can guarantee that the sterile marker will be fitted in exactly the same position as the non sterile version. In the current state this is not taken into account in the system. At the moment the marker tracking is subject to fluctuations. Smaller tracking inaccuracy of the marker lead to larger fluctuations the further away from the marker the hologram is located. This needs to be further improved, e.g. through Extended Kalman Filtering. Our system

is completely mobile and does not depend on external sensors. This leads to challenging tasks. The marker tracking is more complicated compared to a fixed reference system. In addition, the HoloLens still has strong limitations due to the underlying hardware. Initial tests on wearing the headset during a simulated surgery were positive. It benefits from the fact that it is light and the field of vision is not severely restricted. With our design, tracking of the patient with a set of 2D markers reaches high accuracy allowing sufficient precision of the virtual projection. In summary, the approach presented is very promising and may increase the success rate of surgical procedures in the future.

5 Conclusion

We have demonstrated the feasibility of the proposed approach by registering a marker system to the patient's head and then overlaying the virtual patient model by only tracking the marker. The models were generated automatically from CT-data and a process was implemented to visualize them on the HoloLens with a sufficient frame rate. Providing good surgical assistance based only on an AR-headset is largely dependent on the underlying hardware. The release of the HoloLens enabled many scenarios that were previously impossible to implement. Thanks to the Research mode and the associated opening of the platform, all sensors of the headset can now be used. In this paper we have described, implemented and initially evaluated the design of an AR-based navigation system for surgical procedures. A good basis was achieved to further improve the underlying system. The big advantage of our system is the independence from external sensors and expensive additional hardware. In the future we will work on the robustness and accuracy of the system to further improve the stability. This can lead to a much higher success rate for ventricular puncture.

6 Acknowledgement

This work was funded by the Federal Ministry of Education and Research within the project 'HoloMed - (Context-sensitive support of a surgeon in the operating room by Augmented Reality)'. It started on June 1st, 2017 and runs for two and a half years.

7 References

[1] Paul M Foreman, Philipp Hendrix, Christoph J Griessenauer, Philip GR Schmalz, and Mark R Harrigan. External ventricular drain placement in the intensive care unit versus operating room: evaluation of complications and accuracy. Clinical neurology and neurosurgery, 128:94–100, 2015.

[2] Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco José Madrid-Cuevas, and Manuel Jesús Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognition, 47(6):2280–2292, 2014.

[3] Sergio Garrido-Jurado, Rafael Munoz-Salinas, Francisco José Madrid-Cuevas, and Rafael Medina-Carnicer. Generation of fiducial marker dictionaries using mixed integer linear programming. Pattern Recognition, 51:481–491, 2016.

[4] Radu Bogdan Rusu and Steve Cousins. 3d is here: Point cloud library (pcl). In Robotics and automation (ICRA), 2011 IEEE International Conference on, pages 1–4. IEEE, 2011.

[5] Ulrich-Wilhelm Thomale, Andreas Schaumann, Florian Stockhammer, Henrik Giese, Dhani Schuster, Stefanie Kästner, Alexander Sebastian Ahmadi, Manolis Polemikos, Hans-Christoph Bock, Leonie Gölz, Johannes Lemcke, Elvis Hermann, Martin U Schuhmann, Thomas Beez, Michael Fritsch, Berk Orakcioglu, Peter Vajkoczy, Veit Rohde, and Georg Bohner. GAVCA Study: Randomized, Multicenter Trial to Evaluate the Quality of Ventricular Catheter Placement with a Mobile Health Assisted Guidance Technique. Neurosurgery, 1(1):40, August 2017.

[6] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast global registration. In European Conference on Computer Vision, pages 766–782. Springer, 2016.

[7] Timo Zinßer, Jochen Schmidt, and Heinrich Niemann. A refined icp algorithm for robust 3-d correspondence estimation. In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, volume 2, pages II–695. IEEE, 2003.

Towards Head Motion Compensation Using Multi-Scale Convolutional Neural Networks

O. Rajput[1], N. Gessert[1]*, M. Gromniak[1], L. Matthäus[2], A. Schlaefer[1]*

[1] *Institute of Medical Technology, Hamburg University of Technology, Hamburg, Germany*
[2] *eemagine Medical Imaging Solutions GmbH, Berlin, Germany*
* *Both authors contributed equally.*

Contact: omer.rajput@tuhh.de, nils.gessert@tuhh.de

Abstract

Head pose estimation and tracking is useful in variety of medical applications. With the advent of RGBD cameras like Kinect, it has become feasible to do markerless tracking by estimating the head pose directly from the point clouds. One specific medical application is robot assisted transcranial magnetic stimulation (TMS) where any patient motion is compensated with the help of a robot. For increased patient comfort, it is important to track the head without markers. In this regard, we address the head pose estimation problem using two different approaches. In the first approach, we build upon the more traditional approach of model based head tracking, where a head model is morphed according to the particular head to be tracked and the morphed model is used to track the head in the point cloud streams. In the second approach, we propose a new multi-scale convolutional neural network architecture for more accurate pose regression. Additionally, we outline a systematic data set acquisition strategy using a head phantom mounted on the robot and ground-truth labels generated using a highly accurate tracking system.

Keywords: Head Pose Estimation, CNN, Deep Learning

1 Problem

Motion compensation is necessary in various medical procedures including transcranial magnetic stimulation (TMS). TMS is a method used to stimulate cortical regions of the brain, e.g., to treat tinnitus. The necessity of motion compensation arises from the fact that patients are typically not static due to, e.g., breathing and other involuntary movements. Without the motion compensation, the patients mobility has to be restricted with unpleasant aides like plaster casts [1]. For motion compensation in TMS, the patient is tracked with a tracking system and the robot compensates the head motion. Previously, TMS with motion compensating robots have been proposed using marker-based tracking which involves placing/sticking retro-reflective markers on a patient's head [2]. The marker placement typically causes irritation to the treated patients and their placement is likely to be affected when the patients scratch or touch the markers, thus compromising the overall accuracy of the compensation.

Therefore, markerless tracking approaches based on adjusted models have been proposed where a subject-specific template is tracked with an Iterative Closest Point (ICP) algorithm [3]. The model-based approach splits the problem into two major subproblems: An offline fitting step to create a subject-specific face template [4] from noisy sensor data and its registration to the sensor frames using a variant of the Iterative Closest Point algorithm [5] during the actual pose estimation phase. Recently, CNNs have also been proposed for head pose estimation "in the wild" [6] and for automotive driver head pose estimation [7]. Often, these approaches use coarse, discrete poses which is insufficient for TMS application.

Thus, we propose CNN-based head pose estimation from RGBD+IR images for TMS-like applications using a data acquisition setup with a highly accurate tracking camera for ground-truth annotation. A robot moves a head phantom for automatic data acquisition which allows for generation of large datasets to study head pose estimation approaches. Based on the acquired datasets, we propose a multi-scale CNN architecture for accurate pose regression. We compare this method to a single-path CNN and a classical head model-based approach.

2 Materials and Methods

2.1 Head Phantom Dataset Generation

A dataset of the head phantom is generated to systematically cover predefined distances and orientations towards the Kinect sensor. Universal Robot's UR5 robot is used to systematically move the head in the Kinect's FoV.

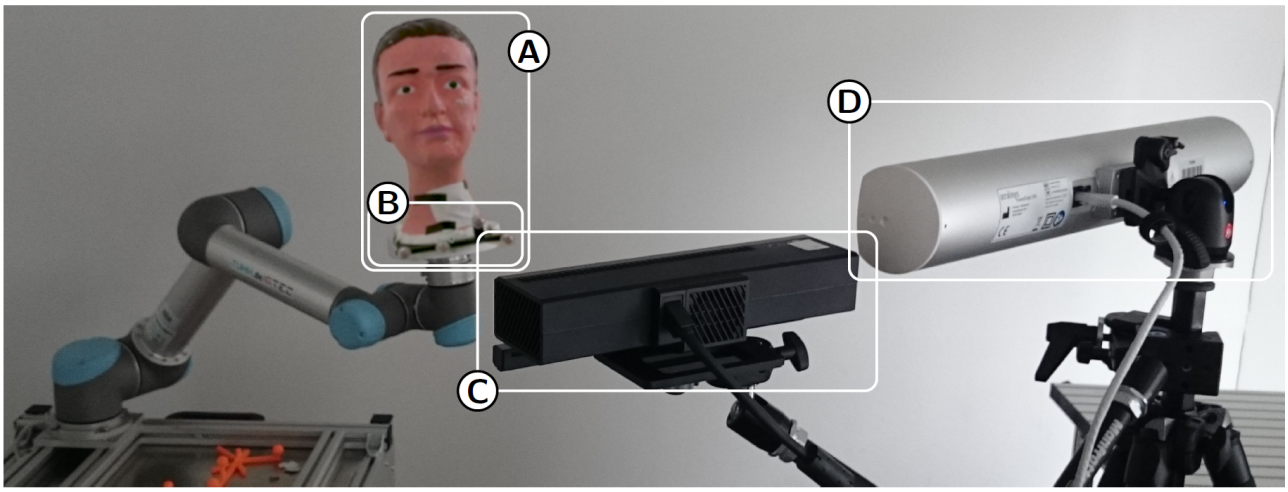


Figure 1: Acquisition setup. Head phantom (A), head marker geometry (B), Microsoft Kinect v2 (C), optical tracking system (D)

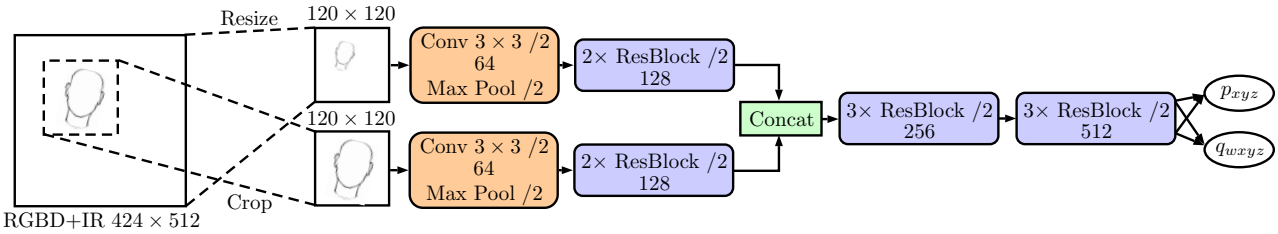


Figure 2: The proposed multi-scale CNN architecture. One path receives a downsampled version of the entire image, the other path receives a cropped version, containing mostly the head itself. Cropping is performed through a face detecting classifier. *ResBlock* refers to residual blocks with bottlenecks, following [8]. In each block, the number of output feature maps is denoted. /2 denotes spatial downsampling with a convolutional stride of 2. p_{xyz} denotes the position and q_{wxyz} denotes the orientation, expressed as a quaternion.

In addition to Kinect, a highly accurate optical tracking system 'Atracsys fusionTrack 500' is used to track the head as well to record the corresponding ground truth head poses. The tracking system tracks a retroreflective markers based rigid geometries, therefore, one of such marker geometries is fixed to the phantoms neck. This was done such that the markers do not interfere/occlude the head/face of the phantom as shown in Figure 1. The robot's trajectory is planned such that the overlap between robot workspace and Kinect FoV is discretized into a grid. At each position in the grid, the robot rotates the head phantom covering combinations of rotations in a range from +/-20 degrees for yaw, pitch and roll axes, respectively. Also, the order of rotations are

	Model-Based	Single-Path	Multi-Path
Position Error	4.16 ± 2.21 mm	2.23 ± 1.38 mm	1.86 ± 1.29 mm
Orientation Error	$4.63 \pm 1.42°$	$1.27 \pm 0.78°$	$0.676 \pm 0.350°$

Table 1: Mean and standard deviation of the errors for the different approaches. Errors are calculated by considering $T_P^{-1} T_{GT} = I_E \stackrel{!}{=} I$ where T_P is the predicted pose and T_{GT} the ground-truth. The result I_E should be approximately equal to the identity matrix I. Position error is the norm of the translational part of I_E and orientation error is the angle of the rotational part of I_E in axis-angle representation.

	Model-Based	Single-Path	Multi-Path
Setup Time	15 min	6 h	14 h
Processing Time	27.1 ± 8.6 ms	5.1 ± 0.3 ms	6.4 ± 0.3 ms

Table 2: Setup and processing time for the three approaches. For the model-based approach, setup time refers to the time required to generate the head template. For the CNN models, setup time refers to the training time needed. Processing time denotes the time that is required for processing a single image for tracking.

A 120x120 section around the ROI's center is cropped from the original image to get the cropped image for the CNN.

For training, we consider all five image channels (RGB+Depth+IR) as channels at the model input. The RGB images were registered to the IR-based representations. The network performs regression by outputting seven continuous values which represent position and orientation (quaternions). Images and target poses are scaled to a range of $[0, 1]$. We minimize the mean squared error between predictions and targets using the Adam algorithm [10]. We implement our model in Tensorflow [11].

We compare this approach to a standard single-path CNN that only takes the entire Kinect image as its input. We do not consider a single-path CNN with the cropped images, as they do not contain global position information. Furthermore, we consider a classical head model based approach for tracking, in which the head template is expressed in terms of a statistical face model [4]. The face model is deformed using a head scan to create a head template corresponding to the specific head shape to be tracked. An optical tracking system and a pointer probe are used to sample a subject's facial surface. The tracking algorithm is initialized by an RGB face detector and uses the Iterative Closest Point algorithm with subject-specific template for pose refinement.

Since a marker geometry is rigidly attached to the head, the head does not need to be static during the scan. The transformations from tracking camera to head marker ${}^{C}T_{M_i}$ and the transformations from tracking camera to pointer probe ${}^{C}T_{P_i}$ are synchronously recorded for n samples. The translational part of the relative transformations ${}^{M}T_{P_i} = ({}^{C}T_{M_i})^{-1} {}^{C}T_{P_i}, i \in 1, \cdots, n$ from head marker to pointer probe can be compiled into a three-dimensional head scan.

3 Results

The results are shown in Table 1. The CNN-based models clearly outperform the classical model-based approach. Moreover, the two-path architecture achieves higher rotational accuracy than the single-path model.

Additionally, we provide results for the training and runtime of each approach. The results are shown in Table 2. The training time for model based approach is required to setup the subject-specific head template as described in the previous section. The CNN requires significantly more setup time due to model training, compared to the classic approach. However, for application, the processing time for CNN inference for each image is four times faster than the model-based tracker.

4 Discussion

We propose a new multi-path CNN architecture for head pose estimation, that takes both global and local context into account. Our results in Table 1 show, that the method outperforms a classic model-based approach. This indicates that directly learning features from color and depth images is advantageous compared to point-cloud-based models that only rely on depth. Note, that these results are limited as they are shown for one head

phantom. Although generalization in between heads does not appear to be problematic for similar problems [7], this should be addressed in future research. Moreover, we show that the multi-scale approach outperforms a single-path model that only takes the entire image as the input. The higher resolution crop appears to provide more accurate information for orientation learning. The setup and processing times shown in Table 2 show that the CNN models need significantly longer preparation time, as the model has to be trained on a large dataset. However, when being applied for head pose estimation, the CNNs are significantly faster than the classic approach in terms of processing time. Thus, our new CNN approach is well suited for practical applications, as fast feedback on head pose changes during a TMS scenario can allow for better motion compensation strategies.

5 Conclusion

We address the problem of head pose estimation for motion compensation in TMS scenarios. For this purpose, we introduce an automatic data acquisition setup with highly accurate ground-truth annotation in order to obtain large datasets to study pose estimation methods. We introduce a new multi-scale CNN architecture for accurate pose regression. We show that our method outperforms a classic approach. Moreover, our new multi-scale CNN design improves performance over a simple single-path model. In this way, we show the successful integration of prior, coarse pose estimation methods in our new deep learning-based setup. Additionally, processing duration improves significantly, which implies that real-time estimation is feasible. For future work, this approach can be extended to multiple head phantoms and potential trials with human subjects.

Acknowledgements

This work was partially supported by grant number ZF 4026301KW5.

References

[1] Weltens, C., Kesteloot, K., Vandevelde, G., Van den Bogaert, W. (1995) Comparison of plastic and orfit® masks for patient head fixation during radiotherapy: precision and costs. International Journal of Radiation Oncology Biology Physics **33**(2), 499–507

[2] Richter, L., Trillenberg, P., Schweikard, A., Schlaefer, A. (2013) Stimulus intensity for hand held and robotic transcranial magnetic stimulation. Brain stimulation **6**(3), 315–321

[3] Breitenstein, M.D., Kuettel, D., Weise, T., Van Gool, L., Pfister, H. (2008) Real-time face pose estimation from single range images. In: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, pp. 1–8. IEEE

[4] Paysan, P., Knothe, R., Amberg, B., Romdhani, S., Vetter, T. (2009) A 3D face model for pose and illumination invariant face recognition. In: Advanced video and signal based surveillance, 2009. AVSS'09. Sixth IEEE International Conference on, pp. 296–301. Ieee

[5] Rusinkiewicz, S., Levoy, M. (2001) Efficient variants of the icp algorithm. In: 3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on, pp. 145–152. IEEE

[6] Patacchiola, M., Cangelosi, A. (2017) Head pose estimation in the wild using convolutional neural networks and adaptive gradient methods. Pattern Recognition **71**, 132–143

[7] Schwarz, A., Haurilet, M., Martinez, M., Stiefelhagen, R. (2017) DriveAHead – A large-scale driver head pose dataset. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on, pp. 1165–1174. IEEE

[8] He, K., Zhang, X., Ren, S., Sun, J. (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE CVPR, pp. 770–778

[9] Déniz, O., Bueno, G., Salido, J., De la Torre, F. (2011) Face recognition using histograms of oriented gradients. Pattern Recognition Letters **32**(12), 1598–1603

[10] Kingma, D., Ba, J. (2014) Adam: A method for stochastic optimization. In: ICLR

[11] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. (2016) Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467

Camera-based Instrument Navigation in ENT-Surgery

M. Katanacho[1], E. Uhlmann[1,2]

[1] Fraunhofer Institute for Production Systems and Design Technology IPK, Berlin, Germany
[2] TU Berlin, Institute for Machine Tools and Factory Management IWF, Berlin, Germany

Contact: {manuel.katanacho, eckart.uhlmann}@ipk.fraunhofer.de

Abstract

Algorithms for a camera-based surgical navigation system, with a localization based on the surface structure of the patients face, are developed and evaluated. The measurement technology is based on a single HD-camera that is integrated on a surgical instrument. Sophisticated and expensive state of the art systems are thus replaced by a software solution, which reduces the hardware costs and the space requirements in the operation room. A functional prototype is implemented and tested in a laboratory and an operating room. The challenge is to determine the camera's pose from a camera live stream with a feature based approach, without additional markers on the patient. The core software modules are mapping, registration and localization. The root-mean-square error of the system varies between 2.4 mm and 6.4 mm dependent on the lighting conditions. By removing a systematic error from the registration process, the accuracy is improved to 1.3 mm to 2.0 mm.

Keywords: Surgical Navigation, Visual Navigation, Computer Vision.

1 Problem

1.1 Background

Minimal-invasive surgery helps to improve surgical outcomes and a rapid recovery of patients. As a result, hospital stays can be reduced and this can minimize costs in the health care sector [6, 11]. The trend towards minimal-invasive surgery also increases the demand in surgical navigation systems [5, 6]. Established systems are used in different medical disciplines, yet they are disadvantageous in some aspects and are therefore not popular among all surgeons. Surgeons demand a stronger integration of medical systems into the operating room workflow [4]. Furthermore, it must be considered that all the computer-aided systems for surgery must physically fit into the operating room [9]. The aim of this work is to avoid complex and cost-intensive measurement technology by replacing it with a software solution. Within this work a function prototype of a camera-based navigation system for ENT-surgery is developed. The measuring technique is a single HD-camera that can be integrated to surgical instruments. The idea is to navigate based on the surface structure of the patient's face with an image processing solution based on the camera live stream.

1.2 State of the Art in Surgical Navigation

Stereo-optical tracking is the most popular measurement technology for surgical navigation. A stereo camera in the infrared spectrum acquires the position of circular markers based on triangulation, and with a set of markers the orientation respectively. Main disadvantages of optical tracking systems are the required line-of-sight between the camera and the markers on the instrument and patient, and the required space of the equipment in the in the operating room [3]. Advantageous is the large accuracy below 0.5 mm [13]. An alternative are electromagnetic tracking systems. They are used to localize small sensors in a direct current or alternating current magnetic field of known geometry [9]. The sensors are made of one (for 5 degrees of freedom) or more (for 6 degrees of freedom) small coils in which the magnetic field induces a voltage dependent on its location in the field [3]. Electromagnetic tracking systems can be used without visual contact between field generator and sensors (no line-of-sight problem) but are less accurate compared to optical tracking systems [4] and susceptible to conductive and ferromagnetic materials [9]. Another option, especially in neurosurgery, are stereotactic systems [10]. They are mechanically affixed to the patient. A reference system defines a coordinates system and is used for guiding surgical instruments. Landmarks that can be recognized in the image data are used for registration [9].

Besides the established systems mentioned above, there is several innovative measurement technology for surgical navigation. Camera-based systems, relevant to this work, are described next. Profess is a surgical navigation system for paranasal sinus and intranasal surgery. A flat patient tracker with a black and white pixel-pattern is

affixed at the eye region of the patient. Moreover, stickers with a similar pattern are fixed to the forehead, cheek and nose of the patient for automatic registration. The location of a camera, integrated in an instrument, is computed based on image processing algorithms. The accuracy in the nasal cavity is ±2 mm according to the manufacturer [12]. Intellijoint HIP is a camera-based surgical navigation system for intraoperative measurement in total knee arthroplasty. The systems measures the location between a camera, which is mounted on the iliac crest, and a tracker, which is mounted on the fermur or on the acetabulum implant, respectively. The system supplies real-time data on leg length, leg relocation and location of the acetabulum implant. A measurement of the leg length can be done with an accuracy below 1 mm and the location of the acetabulum implant below 1 ° [8]. Naviswiss is a camera-based surgical navigation system for total knee arthroplasty. The navigation camera is a handheld optical 3D-scanner. Markers, which are mounted on the iliac crest and the femoral bone, are used for determination of the relative localization. The prosthetic cup and the prosthesis stem are positioned based on the markers [7].

2 Material and Methods

2.1 Concept

The proposed camera-based navigation computes the transformation of the navigation instrument to the patient continuously from camera live images without using additional measurement devices or trackers (Figure 1, a). The camera, that is mounted on the instrument and connected to a computer, is the measurement system. The software that is installed on that computer serves as the navigation software. The transformation ${}^{\mathbf{TCP}}\mathbf{T}_{\mathbf{CAM}}$ from the optical center of the camera to the tool center point (TCP) of the instrument is known from fabrication or can be determined in a calibration process. The software workflow is structured into the single steps mapping, point cloud registration and localization. The mapping determines the patient surface $\mathbf{O}_{\mathbf{MAP}}$ from an image sequence (see Figure 1, b). The point cloud registration relates the mapping to the preoperative acquired computer tomography (CT) patient data by determining the registration transformation ${}^{\mathbf{CT}}\mathbf{T}_{\mathbf{MAP}}$ (Figure 1, c). The localization computes the transformation ${}^{\mathbf{MAP}}\mathbf{T}_{\mathbf{CAM}}$ from the camera to the mapping by continuously processing of the image data (see Figure 1, d). The final navigation transformation ${}^{\mathbf{CT}}\mathbf{T}_{\mathbf{TCP}}$ from the instrument to the CT patient data is given in equation 1.

$$ {}^{\mathbf{CT}}\mathbf{T}_{\mathbf{TCP}} = {}^{\mathbf{CT}}\mathbf{T}_{\mathbf{CAM}}\left({}^{\mathbf{TCP}}\mathbf{T}_{\mathbf{CAM}}\right)^{-1} = {}^{\mathbf{CT}}\mathbf{T}_{\mathbf{MAP}}\,{}^{\mathbf{MAP}}\mathbf{T}_{\mathbf{CAM}}\left({}^{\mathbf{TCP}}\mathbf{T}_{\mathbf{CAM}}\right)^{-1} \tag{1} $$

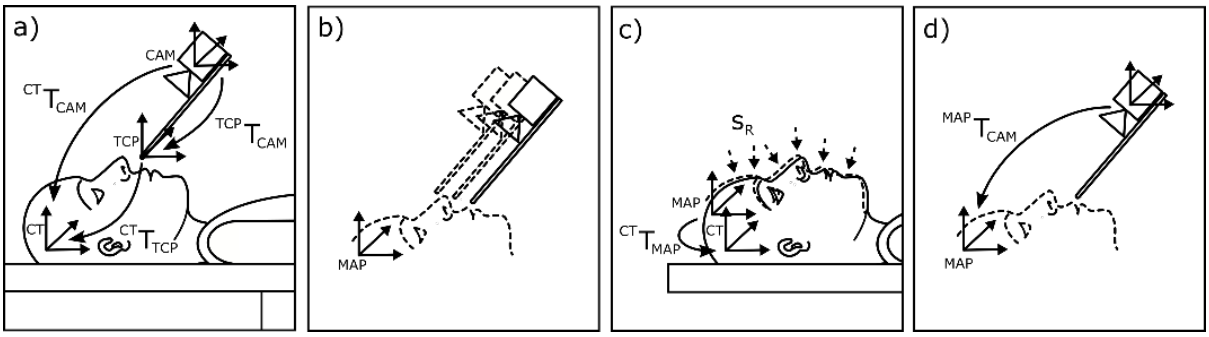


Figure 1: *Concept of the camera-based surgical navigation system: a) the navigation transformation ${}^{CT}T_{TCP}$ is defined by the camera to the patient data transformation ${}^{CT}T_{CAM}$ and camera to instrument transformation ${}^{TCP}T_{CAM}$, b) mapping from an image sequence, c) point cloud registration of the mapping to the CT data, d) camera pose estimation (localization) relative to the mapping.*

2.2 Implementation

Figure 2 shows the software workflow for the camera-based navigation system. The workflow is structured into the modules preoperative steps, image grabbing, patient registration, localization and visualization. The preoperative steps module is responsible for loading of the DICOM patient data and the camera calibration for computation of the internal camera parameters. Furthermore, an instrument calibration can be performed in order to compute the transformation of the camera's optical center to the tool center point. The image grabbing module acquires the camera image stream and signals it to the patient registration and localization module respectively. The patient registration module consists of the steps CT surface extraction, mapping and point cloud registration. The CT surface extraction computes the surface of the CT patient data and converts it to a point cloud. The mapping

module computes a depth map from of the patient's face based on an acquired camera image stream. Output data of the module are a disparity and reference image that is used in the navigation module as well as a point cloud that is used for point cloud registration. The point cloud registration is done in order to align the computed point clouds of the mapping and CT surface extraction. The process computes the required transformation and scaling of the mapping. The navigation module is started after the patient registration. The input of the module are the camera live images, the disparity and reference image from the mapping process and the registration transformation and scaling. The output of the module is the transformation ${}^{CT}\mathbf{T}_{TCP}$ of the tool center point to the coordinate system of the CT-DICOM data. The visualization module displays the navigation data and process information in a graphical user interface (GUI). The navigation data, i.e. the position and orientation of the navigated surgical instrument, is visualized in a slice (axial, coronal, sagittal) and rendered view of the CT patient data. Moreover, a visualization of selected processes, e.g. mapping, registration, tracking, is realized in the GUI for a live feedback for the user. A more detailed description of the CT surface extraction, mapping, registration and localization module is given in the following sections.

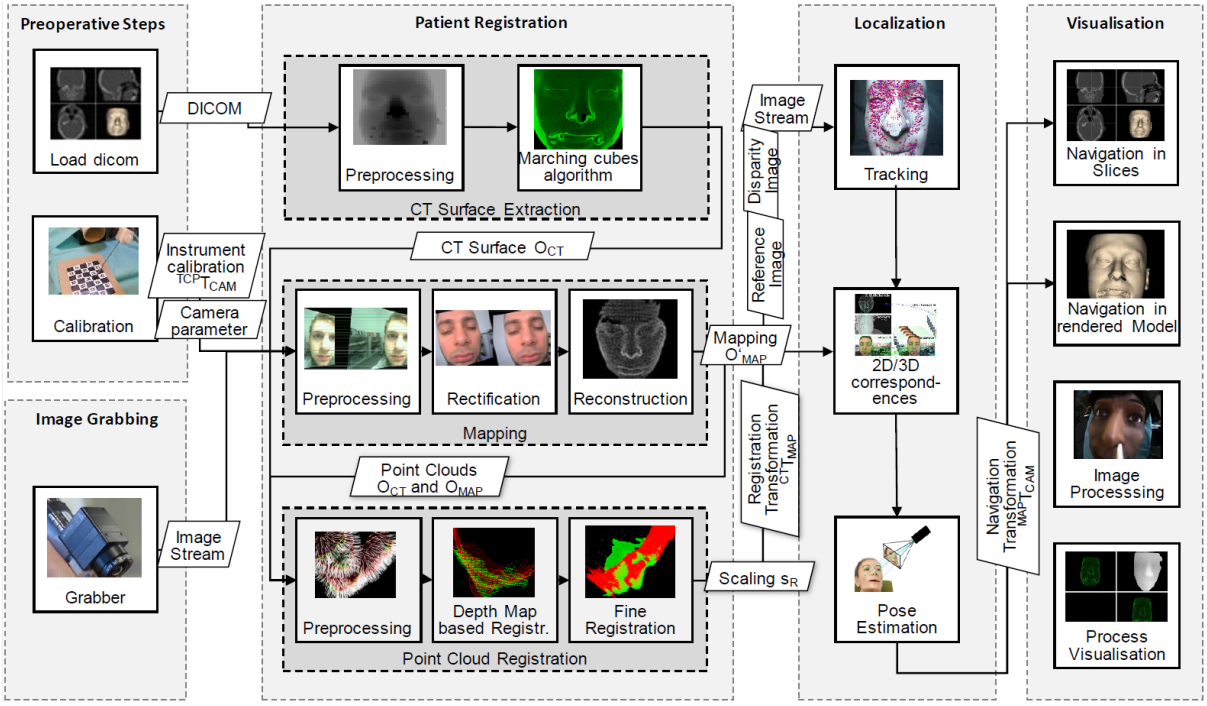


Figure 2: *Software workflow for the camera-based navigation system.*

CT Surface Extraction

The CT surface extraction module computes the surface point cloud of the patients face from the CT-DICOM data. The module is divided into the steps preprocessing and Marching-Cubes algorithm. The preprocessing filter computes a depth map of the frontal view on the patient face by thresholding the voxel data. The depth map is divided into partial maps of 27 x 51 pixels. Each partial map is filtered based on an interval defined by the mean value and standard deviation with a threshold of 2.28 σ. The Marching-Cubes filter processes and returns the point cloud of the patient face surface.

Mapping

The mapping module computes a depth map of the surface of the patient face from an acquired image sequence during the registration process. The mapping is divided into the steps preprocessing, rectification and reconstruction. The preprocessing is used for selection and preparation of appropriate image pairs, i.e. image pairs with a minimum disparity of 60 and maximum disparity of 100 pixels. The rectification filter computes the relative transformation, i.e. rotation and normalized translation, of a selected image pair based on a feature matching with SIFT and SURF. Based on this relative transformation the rectification is done with the Bouguet algorithm [1]. The reconstruction of the depth map is done with semi-global block matching [1]. Subsequently, the depth map is postprocessed to remove outliers and error pixels. The depth map is converted according the reprojection matrix Q to a point cloud. The output of the module is a point cloud, depth map, mask and reference image, i.e. left image of the stereo image pair. The resolution of the depth map is defined by the resolution of the camera and its field of view. Dependent on the perspective the face covers approximately 50 % of the sensor (1280 x 1024 px), which

results in approximately 650,000 points in the point cloud. Yet, the depth resolution is limited and depends on the camera translation, the object distance, the focal length and disparity increment.

Point Cloud Registration

The point cloud registration module computes the scaling s_R and transformation ${}^{CT}\mathbf{T}_{MAP}$ of the mapping $\mathbf{O}_{MAP}$ to the CT surface model $\mathbf{O}_{CT}$. The module is divided into the steps preprocessing, depth map based registration and fine registration. In the first step, the point clouds are both downsampled to a number of 4,000 points for faster processing. The preprocessing filter performs an outlier filtering and prescaling according to the point cloud dimensions. Furthermore, a coarse alignment is done by aligning the centers of mass and applying a known rotation of 90 °. The depth map based registration computes corresponding 3D points based on a feature matching process on the 2D depth maps O_{MAP} and O_{CT} (using BRISK features and SIFT descriptors) [2]. The corresponding features are transformed to 3D points, which are used for fine scaling and computation of the registration transformation. The fine registration is done with a standard implementation the ICP-algorithm. The results of the single registration steps are shown in Figure 3.

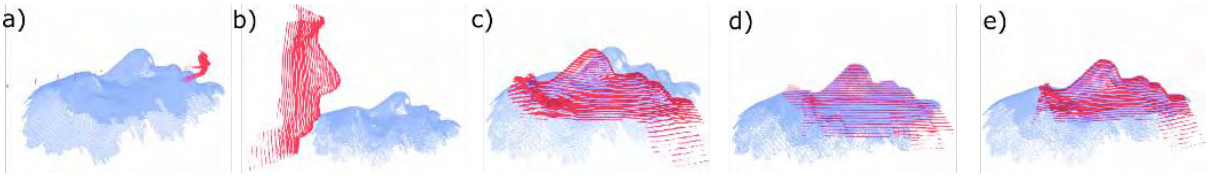


Figure 3: *Interim and final results of the registration process: a) input data O_{CT} and O_{MAP} (small scaled), b) after outlier filtering and prescaling, c) after coarse alignment, d) after depth map based registration, e) after fine registration.*

Localization

The localization module determines the position and orientation (pose) of the camera relative to the mapping $\mathbf{O}_{MAP}$ for every camera image. This is done with a pose estimation based on point correspondences between selected points of the mapping (3D) and tracked points on the camera image (2D). The module is divided into the filters tracking, 2D-3D-correspondences and pose estimation. The tracking filter performs a feature matching between the camera live image and the reference image that is kept from the mapping process. The matching process is replaced by an optical flow tracking, as far as a minimum number of 300 matched points are available. The 2D-3D-correspondences filter computes for every tracked point the corresponding 3D point from the correspondence to reference image (given by the matching) and the correspondence to the mapping $\mathbf{O}_{MAP}$ (see Figure 4). The pose estimation is done with a Levenberg-Marquardt optimization to estimate the pose by minimizing the reprojection error. For every tenth input image the pose estimation is performed with a RANSAC based computation in order to eliminate outliers. The matching, tracking and 2D-3D-corresponcences filters are executed in parallel on two threads with two different image resolutions of 640 x 512 pixels and 512 x 410 pixels for an improved robustness. The pose estimation is done with the output of both of the parallel computed correspondences.

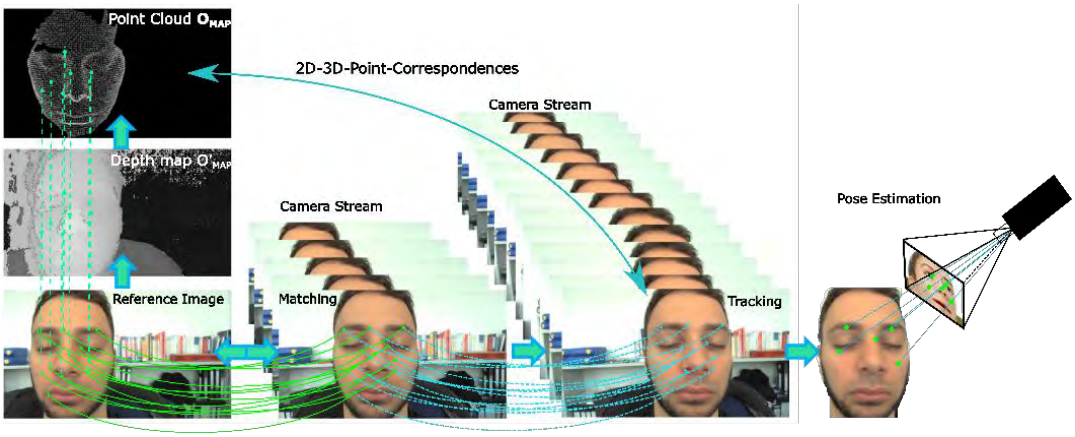


Figure 4: *Operating principle of the localization module. A matching between the reference image and the camera stream allows a computation of the 2D-3D-point-correspondences. By tracking the 2D points in the camera stream, the pose can be computed for every input image.*

3 Results

The system accuracy results from the accuracy of the transformations for instrument calibration ${}^{TCP}\mathbf{T}_{CAM}$, patient registration ${}^{CT}\mathbf{T}_{MAP}$, including scaling s_R, and localization ${}^{MAP}\mathbf{T}_{CAM}$. For a reliable evaluation, a measurement setup with real test persons was developed for testing in the laboratory and an operating room in the Charité Berlin (see Figure 5). Initially, 20 measurement points are marked with an erasable pen on the test person. Additionally, 20 torus-shaped targets, which can be detected on a laser scan, are fixed on each mark (see Figure 5 a). A laser scan of the patient surface is then created and converted to DICOM data. The targets, which are removed after the laser scan, define the position that is touched during measurement (see Figure 5 c - e).

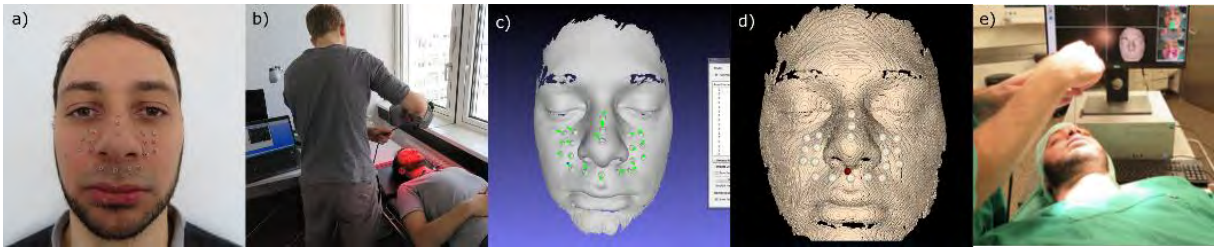


Figure 5: *Measurement setup for the system accuracy: a) Marking of measurement points and targets, b) laser scan of patient surface, c) detection of targets in 3D data, d) conversion to DICOM data, e) performing measurement in different lighting conditions (laboratory and OR).*

The measurements are done in three different lighting conditions, i.e. daylight, fluorescent (in the OR) and operating light (in the OR). The root-mean-square (RMS) error varies from 2.4 mm to 6.4 mm (see Table 1). The main error is present in z-direction (along the camera's optical axis). Either this results from a bad resolution in z-direction due to the single camera approach, or a wrong scaling s_R in the registration process. The system operation is tested for a measurement area of +5 cm and -10 cm around a defined target point and a +/-45 ° rotation around the horizontal axis and a +/- 30 ° rotation around the vertical rotation axis. The processing time of the modules are for CT surface extraction 6 s, mapping 25 s and point cloud registration 20 s. The current implementation of the localization module processes 8 frames per second, but can be accelerated with a graphics processing unit.

Table 1*: Root-mean-square error of the system accuracy for different lighting conditions.*

Root-mean-square error	Illumination		
	Daylight	Fluorescent lamp	Operating light
x_{rms} [mm]	1.7	2.4	2.4
y_{rms} [mm]	0.8	1.3	1.4
z_{rms} [mm]	1.4	5.7	5.7
r_{rms} [mm]	2.4	6.3	6.4

4 Discussion

In a further evaluation, the systematic error from the registration process is eliminated. This is done by correcting the registration transformation and scaling so that the sum of the squared error between the target points $\mathbf{z}_i$ and the

Table 2: *Root-mean-square error of the system accuracy for different lighting conditions with a correction of the registration.*

Root-mean-square error	Illumination		
	Daylight	Fluorescent lamp	Operating light
x_{rms} [mm]	0.7	0.5	0.8
y_{rms} [mm]	0.7	1.2	1.5
z_{rms} [mm]	0.7	1.4	1.1
r_{rms} [mm]	1.3	1.9	2.0

measurement points $\mathbf{m_i}$ is minimized. The scaling is varied between $s_k = 0.8$ and $s_k = 1.2$ with increments of 0.001. This way the root-mean-square error of the system accuracy can be reduced to 1.3 mm to 2.0 mm (see Table 2). The correction factor for the scaling s_R varies between 0.99 (daylight measurement), 0.92 (fluorescent lamp) and 0.94 (operating light). This correlates with the error r_{rms} in z-direction for the varying lighting conditions and allows to conclude that it results mainly from a scaling error in the registration. From this results can be concluded that the measurement results for the system accuracy include a systematic error that is mainly present due to the registration process. Especially a large deviation in the scaling factor leads to larger errors in z direction and a decline in the overall measurement precision.

5 Conclusion

In this work methods for camera-based surgical navigation, with a navigation based on the surface structure of the patients face is developed and evaluated. The measurement technology of the system is a HD-camera that is integrated on a surgical instrument. This software solution can replace or supplement sophisticated and expensive state of the art systems, which reduces hardware costs and space requirement in the OR.

A functional prototype was implemented and tested in the laboratory and an operating room. The root-mean-square error of the system between 2.4 mm and 6.4 mm varies with different lighting conditions. The main error results from an incorrect scaling of the mapping in the registration process, which is subject to future work. With a corrected registration, the RMS error of the system lies between 1.3 mm and 2.0 mm.

Acknowledgement

We would like to thank the Federal Ministry of Education and Research (BMBF) for their support of this research project (13GW0039).

References

[1] Bradski, G., Kaehler, A.: *Learning OpenCV. Computer Vision with the OpenCV Library*. O'Reilly Media, Inc., Sebastopol, California (2008)

[2] Katanacho, M., Lack, J.-C., Hoffmann, J., Uhlmann, E. (eds.): *Depth Map based Point Cloud Registration of CT Patient Data for Surgical Navigation*. CURAC 2017, Hannover (2017)

[3] Kleemann, M., Martens, V., Hildebrand, P., Besirevic, A., Schlichting, S., Mirow, L., Roblick, U.J., Bürk, C., Bruch, H.P.: *Die Erfassung des Raumes – Grundlagen der Trackingtechnologie*. FOCUS MUL(4) (2006)

[4] Knobel, B.: *Surgical measuring systems*, 10 November 2015

[5] Mickel, P.: *European Markets for Surgical Navigation Systems 2013*

[6] Mirota, D.J., Ishii, M., Hager, G.D.: *Vision-based navigation in image-guided interventions*. Annual review of biomedical engineering **13**, 297–319 (2011)

[7] Naviswiss AG: *Naviswiss - New Navigated Surgery*. https://naviswiss.eu/de/ (2018). Accessed 21 February 2018

[8] Paprosky, W.G., Muir, J.M.: Intellijoint HIP(R): *a 3D mini-optical navigation tool for improving intraoperative accuracy during total hip arthroplasty*. Medical devices (Auckland, N.Z.) **9**, 401–408 (2016). doi: 10.2147/MDER.S119161

[9] Peters, T.M.: *Image-guidance for surgical procedures*. Physics in medicine and biology **51**(14), R505-40 (2006). doi: 10.1088/0031-9155/51/14/R01

[10] Rahman, M., Murad, G.J.A., Mocco, J.: *Early history of the stereotactic apparatus in neurosurgery*. Neurosurgical focus **27**(3), E12 (2009). doi: 10.3171/2009.7.FOCUS09118

[11] Shannon, R.: *European Markets for Surgical Navigation Systems 2011*

[12] Stryker Navigation: *PROFESS Navigation*. http://www.stryker.com/stellent/groups/public/documents/web_prod/166358.pdf (2014)

[13] Wiles, A.D., Thompson, D.G., Frantz, D.D.: *Accuracy assessment and interpretation for optical tracking systems*. In: Medical Imaging 2004, p. 421

First experience with computer-assisted laparoscopic needle guidance for pancreatic tumour ablation in a pig

Benjamin Eigl[1,2], I. Paolucci[1], T. Rudolph[2], A. Garcia[3], E. Serra[3], S Weber[1], Mariano Gimenez[3]

[1]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland
[2]CAScination AG, Bern, Switzerland
[3]Institut de Chirurgie Guidée par l'Image, IHU, Strasbourg, France

Contact: benjamin.eigl@artorg.unibe.ch
mariano.gimenez@ihu-strasbourg.eu

Abstract

Pancreatic cancer is the deadliest of the gastrointestinal tract with 5-year survival rates of less than 5%. Irreversible-electroporation (IRE) is an ablation technique that spares vessels and is therefore suitable for treatment of locally advanced and unresectable pancreatic cancer. Placing these needles resembles a major challenge especially if the tumour is located nearby structures at risk such as the aorta, vena cava and portal vein. Hereby we present the first experience from an animal trial using our computer-assisted planning and navigation solution for laparoscopic needle guidance to support the surgeon during this difficult procedure.

Keywords: Pancreas, computer-assistance, needle guidance, minimally-invasive surgery

1 Problem

Local ablation is a treatment alternative for patients with unresectable pancreatic cancer. Irreversible-electroporation (IRE) is a non-thermal ablation technique that spares vessels and thus is predisposed for appropriate treatment of locally advanced pancreatic cancer [1]. Correct needle positioning and parallel placement is a prerequisite for an efficient and complication-free IRE treatment. In our previous work we proposed a 2D/3D planning and visualization solution to support the surgeon during the preoperative procedure and investigated a large potential for intraoperative needle guidance during an evaluation with 8 medical doctors with an average of 70 pancreatic surgeries per year [2].

A minimally invasive laparoscopic approach has the potential to combine the advantages of both, the procedure through open surgery and through CT-guidance. Navigated computer-assistance could support for tumour localization as well as precise and safe needle placement according to the pre-planned strategy, thus avoiding injury to vascular structures.

Stillström et al. have shown the feasibility of using a computer assisted navigation approach in placing the IRE needles in a minimally-invasive setting [3]. Their work serves as a basis for our research to further develop the existing solutions to assist the surgeon during the preoperative and intraoperative procedure.

Herein, we present the result of an animal study where we evaluated the applicability of computer-assisted planning and navigation for pancreatic tumour ablation using a single needle in a laparoscopic setting. Furthermore, we conducted the CT imaging with and without pneumoperitoneum and compared the registration outcome. This work serves as a basis to investigate the applicability of our proposed workflow for the future IRE use cases with multiple needles.

2 Material and Methods

The test site (IHU Strasbourg, France) consisted of a CT Scanner (Somatom Force, Siemens), a navigation system (Cas-One, CAScination), and a laparoscopy system (Karl Storz). The animal was under general anesthesia and for the study two artificial tumours were inserted percutaneously into the liver and pancreas body with the guidance of the interventional radiology system (CAS-One IR). Next, two CT scans were performed with and without pneumoperitoneum to evaluate the fiducial registration error (FRE) during the intraoperative navigation which resembles the quality of the point-based registration. Our proposed workflow consists of two steps which are described as follows.

2.1 Pre-operative planning

The CT scans were performed under apnea with 30 mm/Hg pressure and segmented using a commercially available software (Myrian, Intrasense), specifically targeted for segmentation of abdominal structures like the liver and pancreas parenchyma with tumours, aorta, vena cava and portal vein. The data were imported into the navigation system and the planning of a needle for each tumor was conducted using the developed planning solution (see Figure 1). The software gives the surgeon the ability to pre-operatively define targeting strategies using the CT image data and 3D reconstructions. In our use case a single needle was planned to evaluate our proposed workflow by means of targeting accuracy and overall feasibility. In previous work we have shown the potential of the 3D planning for multiple needles in pancreatic IRE treatments as it provides the surgeon the possibility to better prepare for the upcoming procedure [2].

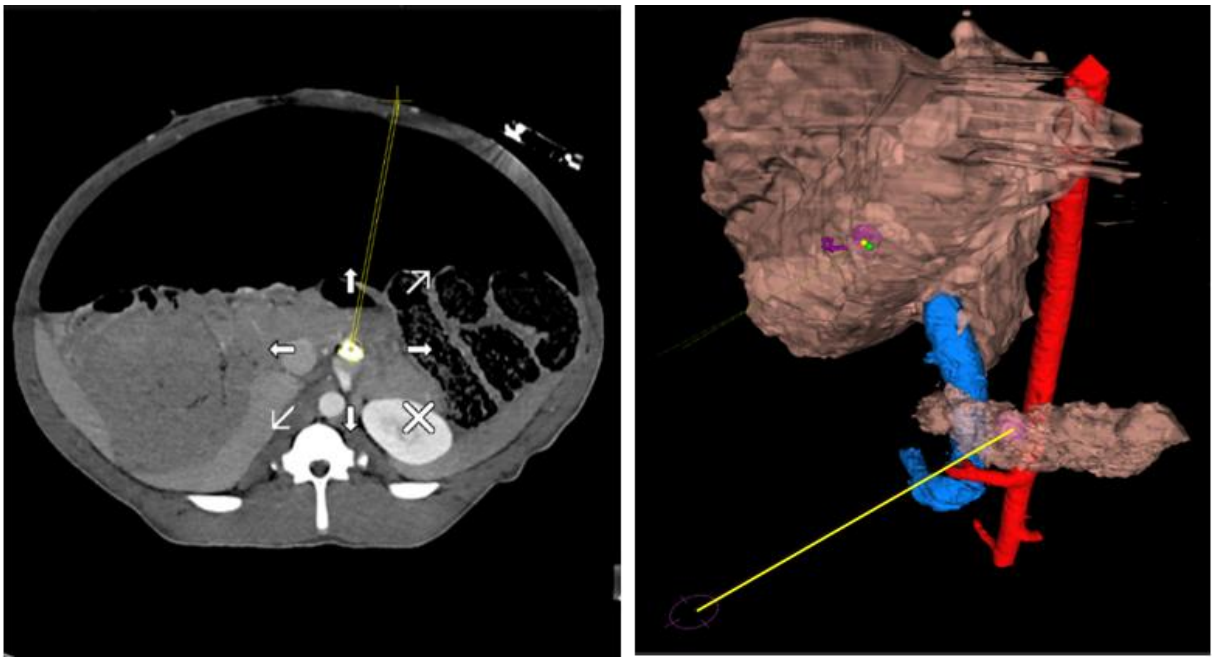


Figure 1: CT scan (pneumoperitoneum) and 3D reconstruction with planned trajectory

2.2 Intra-operative navigation

Using the planning data, the surgeon can perform a rigid registration using surface landmarks acquired by an optically tracked instrument. A rigid pointer with the length of 350mm was attached with a marker shield and calibrated. For the evaluation four corresponding point pairs were acquired on the organ surface of the liver (2), pancreas (1), and spleen (1). The landmarks were chosen after the surgical exploration (see Figure 2). The developed software allows the surgeon to define points on the CT image data as well as on the 3D reconstructions. The software tracks relevant data related to the acquisition attempts and re-calculates the FRE once a new point pair is acquired. The outcome is used to compare the conventional and pneumoperitoneum approach. Once a rigid transformation is calculated, the defined and acquired points are visualized in the 3D scene to give the surgeon a visualization of possible outliers.

For the needle placement the CT scan with pneumoperitoneum was used under live endoscopic ultrasound control and the animal was put under apnea during the insertion process. The needles were placed according to the planned 3D trajectory and the distance to the target was evaluated using the endoscopic ultrasound.

Figure 2: Intraoperative setup

3 Results

The registration workflow was conducted twice (CT scan with and without pneumoperitoneum) with three registration attempts each. When registering the non-insufflated scan, the mean FRE was located at 15.6mm with an average of 4 minutes per registration attempt. These numbers decreased for the pneumoperitoneum approach with a mean FRE of 8.7mm and an average time of 3 minutes per registration attempt.

The targeting of the liver tumour was conducted using the pneumoperitoneum approach and resulted in a displacement of the needle to the tumour center by < 20mm while the second attempt was considered as successful (< 10mm). Between these attempts the point pair acquired at the spleen was filtered out of the registration which resulted in a better registration around the region of interest. The same misplacement was also observed during the targeting attempt of the pancreatic tumour where the misplacement during the first attempt was even larger with < 30mm. The second attempt resulted in a slightly decreased misplacement, yet the accuracy was not as good as the targeting approach of the liver tumour.

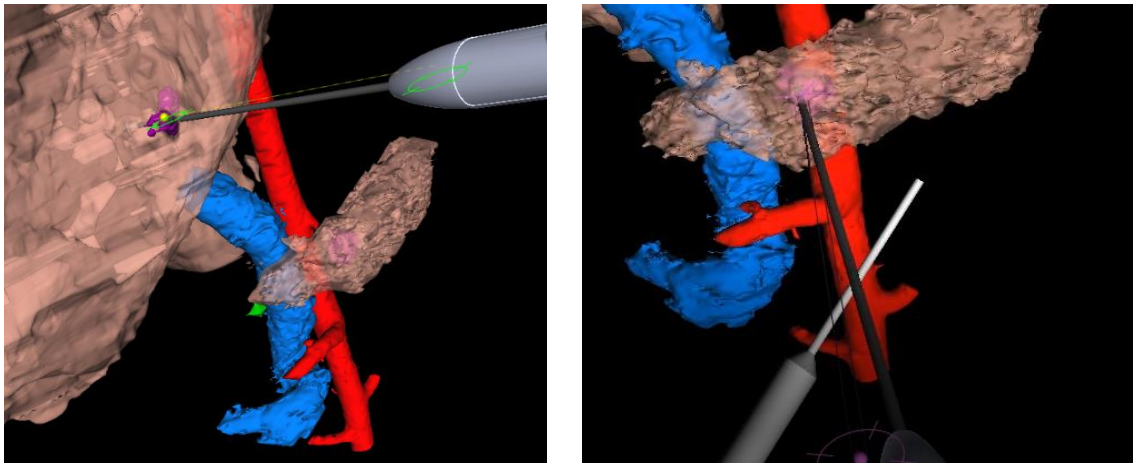


Figure 3: Targeting attempts visualizing the liver (left) and pancreas approach (right).

4 Discussion & Conclusion

In this work, we presented the feasibility of a computer-assisted planning and navigation approach for the placement of ablation needles in the liver and the pancreas. The planning enables the surgeon to pre-operatively define targeting strategies based on the 2D and 3D image data and thus better prepare for the upcoming procedure. Intraoperatively, we investigated that the system assists the surgeon to localize the tumour during the surgical exploration. It can be stated that for an accurate targeting according to the pre-operatively defined plan multiple factors must be considered. First, the CT scan with the pneumoperitoneum enabled better registration which was an expected result but is not easy to accomplish due to the small number of hybrid operating rooms. The conventional method still provides a sufficient registration result for the localization of the tumours. It needs to be said that one part of the improvement between the conventional and pneumoperitoneum approach can be explained that the surgeon became more familiar with the workflow. Furthermore, landmarks near the region of interest result in a better registration. The larger misplacement in the pancreatic targeting approach can be explained by the location of the registration points closer to the region of the tumour in the liver. Therefore, our the navigation solution enables the surgeon to acquire as many point pairs as necessary and provides the possibility to filter out certain pairs. This allows to easily adjust the registration for each tumor in case they are located at different spots. Beside the surface landmarks an ultrasound-based approach to acquire landmarks on internal structures need to be considered, especially in the pancreatic use case. We are currently working on the solution to track the laparoscopic ultrasound using electromagnetic tracking. One of the major challenges resembles the freehand navigation to follow the planned trajectory. Therefore, an approach with a tracked aiming device, commonly used in interventional radiology, will be investigated in the future.

In respect of the usage of multiple needles used for the IRE procedure we see a large potential for this workflow to, on the one hand allow the surgeon to pre-operatively define the IRE strategy, and on the other hand to assist during the needle placement while considering different constraints such as parallelism, spacing between the needles and anatomical structures at risk. Finally, this solution shall enable more minimally-invasive IRE treatments for patients with locally advanced, unresectable pancreatic cancer

References

[1] R. C. G. Martin, “Use of irreversible electroporation in unresectable pancreatic cancer,” *Hepatobiliary Surg. Nutr.*, vol. 4, no. 3, pp. 211–215, Jun. 2015.

[2] B. Eigl, M. Peterhans, S. Weber, D. Candinas, B. Gloor, and M. Worni, “3D planning of irreversible electroporation treatment in pancreatic carcinoma : a use case,” in *IHPBA*, 2018, p. 1.

[3] D. Stillström, H. Nilsson, M. Jesse, M. Peterhans, E. Jonas, and J. Freedman, “A new technique for minimally invasive irreversible electroporation of tumors in the head and body of the pancreas,” *Surg. Endosc.*, vol. 31, no. 4, pp. 1–4, Apr. 2016.

Stereotactic image-guidance for percutaneous transgastric irreversible electroporation of pancreatic tumors – an animal case

Iwan Paolucci[1], Benjamin Eigl[1,2], Alain Garcia[3], Edgardo Serra[3], Stefan Weber[1], Mariano Gimenez[3]

[1]*ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland*
[2]*CAScination AG, Bern, Switzerland*
[3]*Institut de Chirurgie Guidée par l'Image, IHU, Strasbourg, France*

Contact: iwan.paolucci@artorg.unibe.ch, mariano.gimenez@ihu-strasbourg.eu

Abstract

Pancreatic cancer has one of the highest mortality rates with a 5-year survival of less than 5%. Only about 10% of pancreatic cancers are curatively treatable with surgery. Irreversible-electroporation (IRE) represents a treatment method for advanced and unresectable pancreatic tumors. Due to the complex anatomy around the pancreas, this procedure is mostly performed in open surgery under intra-operative ultrasound guidance. Image-guidance could potentially allow the safe treatment in a percutaneous approach, which would lead to lower complication rates and shorter hospital stays.

Keywords: Pancreas, minimally-invasive surgery, stereotactic image-guidance

1 Introduction

Only about 10% of patients with pancreatic cancer are treatable with resection, which is considered the only curative treatment option. Local ablation is an alternative for advanced and nonresectable pancreatic cancer. These tumors have usually grown around the pancreatic duct and other blood vessels and therefore a non-thermal ablation technique must be used to preserve these vessel structures. Irreversible electroporation (IRE) is such a non-thermal ablation method, which induces apoptosis of cells by applying an electric field but does not the affect the connective tissue of vessels [1]. Due to the proximity of the pancreas to critical structures, this treatment is mostly performed using an open surgical access, which comes with high complication rates and long hospital stays.

The percutaneous access represents a minimally invasive approach for IRE treatment in the pancreas. However, due to the proximity of large blood vessels (aorta, vena cava, portal vein) it is a very challenging and risky procedure. Due to the location of the pancreas, a transgastric trajectory would be the shortest if targeted from them the anterior abdominal wall [2, 3]. As the stomach is usually filled with air it is not possible to use real-time ultrasound imaging for needle guidance. Relying on pure CT guidance exposes the patient and the interventionalist to a large amount of radiation. Additionally, as the needle must pass through multiple air-tissue interfaces, the needles are prone to bending. Therefore, stereotactic image-guidance might help to overcome these technical challenges and therefore making the percutaneous approach safer and more efficient. This would come with the advantages of percutaneous over open surgery such as less complications and faster recovery.

In this study, we present an initial animal case of a percutaneous transgastric IRE needle placement for pancreatic tumors in an animal trial.

2 Material and Methods

For planning and placement of the IRE needles (Nanoknife, Angiodynamics) a navigation system for interventional radiology (CAS-One IR, CAScination AG) was used. The animal was under general anesthesia during the whole treatment. First, an artificial tumor consisting of Arginate was percutaneously injected into the head of the pancreas using the navigation system. Then, a contrast enhanced CT (CECT) was acquired for planning of the needle trajectories. Each needle was placed through the skin and the stomach using the tracked aiming device of the navigation system. After all needles were placed, a control CT scan was acquired to validate the accurate needle positioning with respect to the plan. During all CT acquisitions and during the needle placement, the animal was put on apnea with 30 mm/Hg pressure.

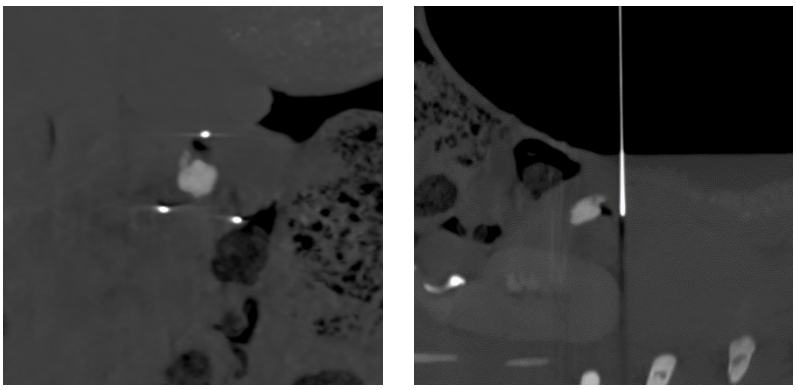


Figure 1: A top view and lateral view of the needle with respect to the tumor

3 Results

All three IRE needles could be safely placed around the tumor through the stomach (Figure 1) with lateral errors of *0.2, 7.1, and 4.5 mm*. Due to the vicinity of the needles to the portal vein, they were placed shorter than indicated by the navigation system, with a longitudinal distance to the tumor of *6.1, 4.5, and 1.0 mm*, and adjusted after the control scan. In the volume rendering (Figure 2), one can see that one needle was heavily bent, which led to a lateral error of 7.1 mm.

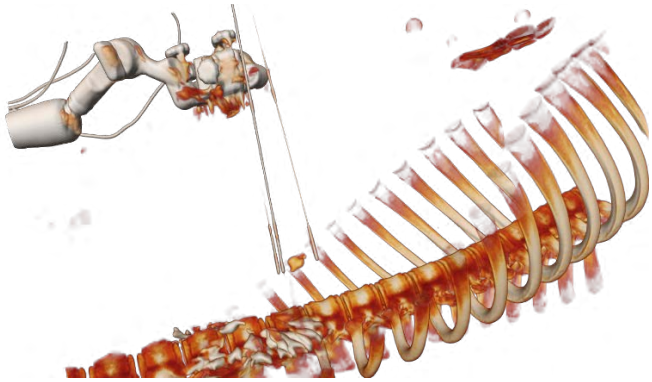


Figure 2: Volume rendering

4 Discussion

In this study, an IRE needle placement for a pancreatic tumor was successfully performed using stereotactic CT based image-guidance. The CT based planning was especially useful to plan safe trajectories with a sufficient distance to critical anatomical structures like the portal vein or the aorta, while meeting the technical constraints of the IRE system (needle spacing and parallelism).

To conclude, this image-guided approach might represent a safe and efficient approach for percutaneous IRE treatment of unresectable pancreatic cancer as an alternative to the open surgical approach.

References

1. Martin RCG (2015) Use of irreversible electroporation in unresectable pancreatic cancer. Hepatobiliary Surg Nutr 4:211–215 . doi: 10.3978/j.issn.2304-3881.2015.01.10
2. Trueba-Arguiñarena FJ, de Prado-Otero DS, Poves-Alvarez R (2015) Pancreatic Adenocarcinoma Treated With Irreversible Electroporation Case Report. Medicine (Baltimore) 94: . doi: 10.1097/MD.0000000000000946
3. Will U, Reichel A, Fueldner F, Meyer F (2015) Endoscopic ultrasonography-guided drainage for patients with symptomatic obstruction and enlargement of the pancreatic duct. World Journal of Gastroenterology 21:13140–13151 . doi: 10.3748/wjg.v21.i46.13140

Preclinical Evaluation of a Micro-Stereotactic Surgical Targeting System for Minimally Invasive Cochlear Implant Surgery.

Th. S. Rau[1], M. E. Timm[1], M. Kluge[2], S. John[2], J. Stieghorst[2], M. Fröhlich[4], L. Kahrs[3], T. Ortmaier[3], O. Majdani[1,5], Th. Lenarz[1],

[1]Department of Otolaryngology and Cluster of Excellence EXC 1077/1 „Hearing4All“, Hannover Medical School, Hannover, Germany
[2]OtoJig GmbH, Hannover, Germany
[3]Institute of Mechatronic Systems, Leibniz Universität Hannover, Hannover, Germany
[4]MED-EL Medical Electronics GmbH, Hannover Research Center, Hannover, Germany
[5]Department of Otolaryngology, Klinikum Wolfsburg, Wolfsburg, Germany

Kontakt: rau.thomas@mh-hannover.de

Abstract

In the past we have developed a micro-stereotactic targeting system to enable an image-guided, minimally invasive access to the inner ear in order to reduce intraoperative trauma to the patient in cochlear implant (CI) surgery. The system consists of a reusable, bone-anchored frame and a customized drilling jig. Recently, a profound preclinical evaluation of the system has been started regarding accuracy, handling, and safety issues. Full head cadaveric specimens are used in order to test the whole surgical procedure. In the following, preliminary results after the first five specimens are presented.

Keywords: image-guided surgery, micro-stereotactic frame, drill guide, drilling accuracy, direct cochlear access.

1 Problem

A minimally invasive approach to the inner ear requires drilling of a single bore hole from the surface of the skull down to the cochlea. The straight path typically passes the facial recess—a narrow region between the facial nerve and its branch chorda tympani. Therefore, high accuracy is mandatory. As the average diameter of the facial recess is only approx. 2.5 mm [1] the safety margins to the nerval structures as well as the allowable diameter of the bore hole and are limited. In addition, the size of the surgical access to the inner ear needs to be large enough to perform the insertion of the electrode array of a cochlear implant (CI). It is known from previous studies by Labadie et al. [2,3], that sufficient accuracy can be achieved using customized micro-stereotactic frames, which are rigidly fixed to the bony skull. Such a system serves as a drill guide, constraining the drilling direction to a predetermined trajectory, which is planned using patient-specific image data. A major drawback of the existing concepts is that neither 3D-printing of the STarFix [2] nor milling of the Microtable [3] allow for rapid customization under sterile conditions.

In order to overcome these limitations, we developed a two-part micro-stereotactic concept [4] that is driven by the aim to enable patient-specific fabrication during the surgery under sterile conditions. Thus, the customization procedure is simplified: Now, only a single counterbored hole needs to be drilled into a blank in order to finish a patient-individual jig. This drilling jig is deployed onto a bone-anchored carrier frame to establish a rigid fixation with the patient's head. A standardized mechanical coupling interface links both parts of the micro-stereotactic system. The carrier frame serves also for image-to-patient registration and, therefore, is equipped with four spherical markers. Up to now, only the accuracy of fabricating the individual drilling jig was investigated and was found to be 0.11 mm ± 0.04 mm [5]. However, determination of the overall drilling accuracy including all clinically relevant steps of the procedure (see Figure 1) is still missing, which was aim of the presented study.

In the following, preliminary results of the first evaluation study in a preclinical setting using full body donors are presented. This study was motivated by the following questions and issues: Is it possible—using that system—to drill through the facial recess without unexpected damage to vital structures? Which accuracy can be achieved when drilling is performed in the inhomogeneous bone material of the mastoid? Furthermore, we wanted to collect practical experience and surgical feedback for the further development of both hard- and

software components of the system; especially regarding the bone anchoring, imaging quality, and path planning. A minimum of 12 trials are planned. Preliminary results of the first five trials are reported in the following.

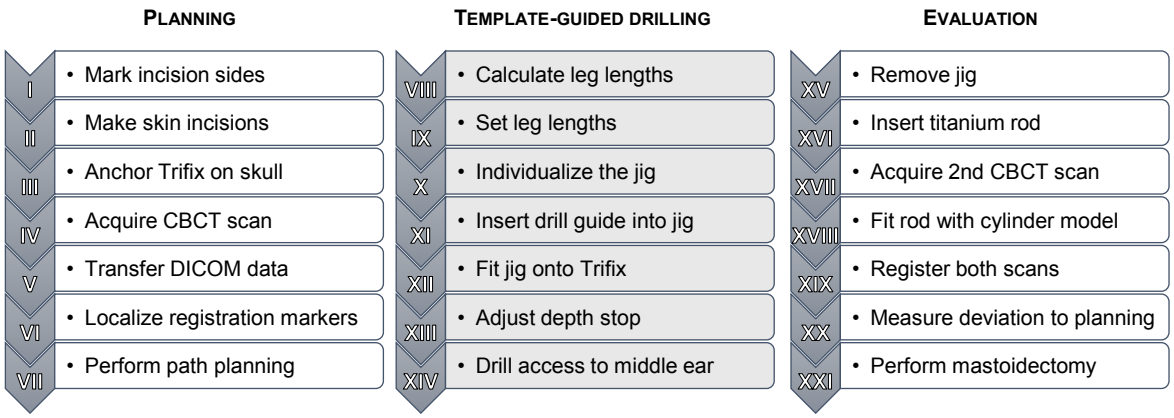


Figure 1: *Flowchart of steps involved for preclinical evaluation.*

2 Material and Methods

For preclinical evaluation five Thiel-fixed [6], anonymized, donated human bodies were used under approval of the authors' institutional review board. Only the left side of each head was used within this study. At the beginning of the procedure, the reusable reference frame, referred to as 'Trifix' in the following, had to be fixed to the skull using three self-drilling and self-tapping bone screws (see Fig. 2a, including steps I-III described in Fig. 1). The position of the Trifix was chosen by the surgeon just behind the pinna with respect to the expected drill path. Afterwards, imaging was performed (Fig. 2b, step IV) using a mobile, intraoperative cone beam computed tomograph (CBCT, xCAT, Xoran, Ann Arbor, MI, USA) with an isotropic voxel size of 300 µm. All scans were checked for proper bone fixation of the screws and visibility of the complete reference frame including all four spherical registration markers.

Approved image data were transferred (step V) to MacBook Pro and loaded into 3DSlicer 4.8.1 with custom-made extensions for semi-automated sphere detection (VI, Fig. 2d)) as well as path planning (VII, Fig. 2c). Using these tools, a safe trajectory was planned manually passing through the facial recess without violating surrounding vital structures of the temporal bone. Special focus was set on preservation of the facial nerve and a drill path which keeps the bony wall of the outer ear canal intact. In one case—due to a narrow facial recess—chorda tympani was sacrificed in order to keep a sufficient safety margin to the facial nerve. Drilling was planned to stop after reaching the middle ear. Start and entry point of each trajectory, relative to the carrier frame, were used to compute the inverse kinematics of the 'Jig Maker' (see below).

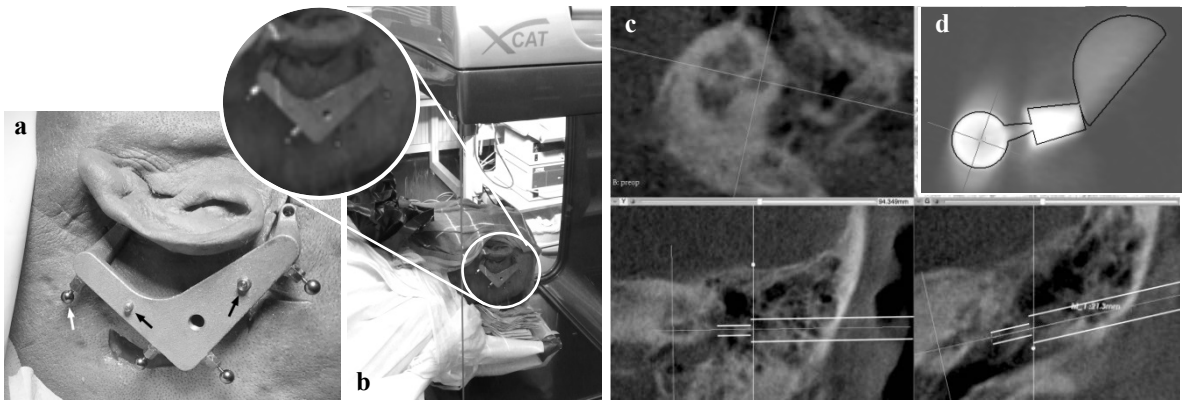


Figure 2: ***(a)*** *Trifix attached to the patient's skull just behind the pinna. Two dowel pins (black arrows) on its surface allows connecting the jig to the reference frame in a free-of-play manner. Four spherical markers (white arrow) serve for image-to-patient registration.* ***(b)*** *Patient's head with mounted reference frame (white circle) inside the xCAT CBCT scanner.* ***(c)*** *Screenshot of the trajectory planning using a custom-made 3DSlicer plug-in.* ***(d)*** *Detected position of the registration marker indicating low fiducial registration error.*

The individual part of the micro-stereotactic system was fabricated out of a blank (ULTEM 1000 resin) using a self-constructed device, referred to as Jig Maker in the following (Figure 3a, [5]). In that device, the blank is temporarily fixed using the same mechanical coupling interface, which is identical to the one on the surface of the Trifix. Adjustment of position and orientation in a patient-specific manner is conducted by use of a pose setting mechanism, which is a Stewart-Gough-Platform (Hexapod) with 6 degrees of freedom. The lengths of all six legs of the hexapod were manually adjusted (step IX) using values provided by the planning software (VIII) written in Python using the libraries NumPy [7] and Plotly. After length setting the main axis of the drill bit is in coaxial alignment with the planned trajectory. A special polymer drill bit was designed which enables drilling a counterbored hole with (12.5 and 15.0 mm in diameter, Fig. 3b) in a single-shot process. The drilling machine was manually advanced (step X).

Afterwards the jig was removed from the Jig Maker and equipped with a drill guide (step XI) suiting custom-made twist step drill bits with a diameter of 4.0 mm at the base and 1.8 mm at the tip. The whole assembly was mounted onto the Trifix (XII, Fig. 3c). The twist drill bit was equipped with a set collar for mechanical limitation of the drilling depth (XIII, Fig. 3d). The distance from the drill's tip to the set collar was measured during trajectory planning. Using the guidance of the jig, the access to the middle ear was drilled by the surgeon (XIV, Fig. 3d). Drilling was performed slowly applying only slight thrust and under continuous irrigation with water. The drilling stopped when the set collar, attached to the drill bit, reaches the upper drill bushing.

After drilling, the surgical jig was unplugged from the reference frame (step XV) in order to get access to the bore hole. It was flushed to remove bone dust. Following, a titanium rod with the same two-tier diameter like the drill bit was inserted (XVI, Fig. 4a) to improve visualization of the bore hole in radiological imaging. Afterwards, a second CBCT scan was acquired (XVII) to assess the actual drill path represented by clearly visible titanium rod. Therefore, the titanium rod was manually fitted to a cylinder model with the correct diameter of 4 mm (XVIII). Planning and drilling were compared by registration of both scans (XIX) using the spherical registration markers of the Trifix. The deviation between the planned and drilled trajectory was measured (XX, Fig. 4b) as the distance between the two center lines at the depth of the planned target point. Finally, preservation of facial nerve and chorda tympani was assessed under the microscope after conventional mastoidectomy (XXI).

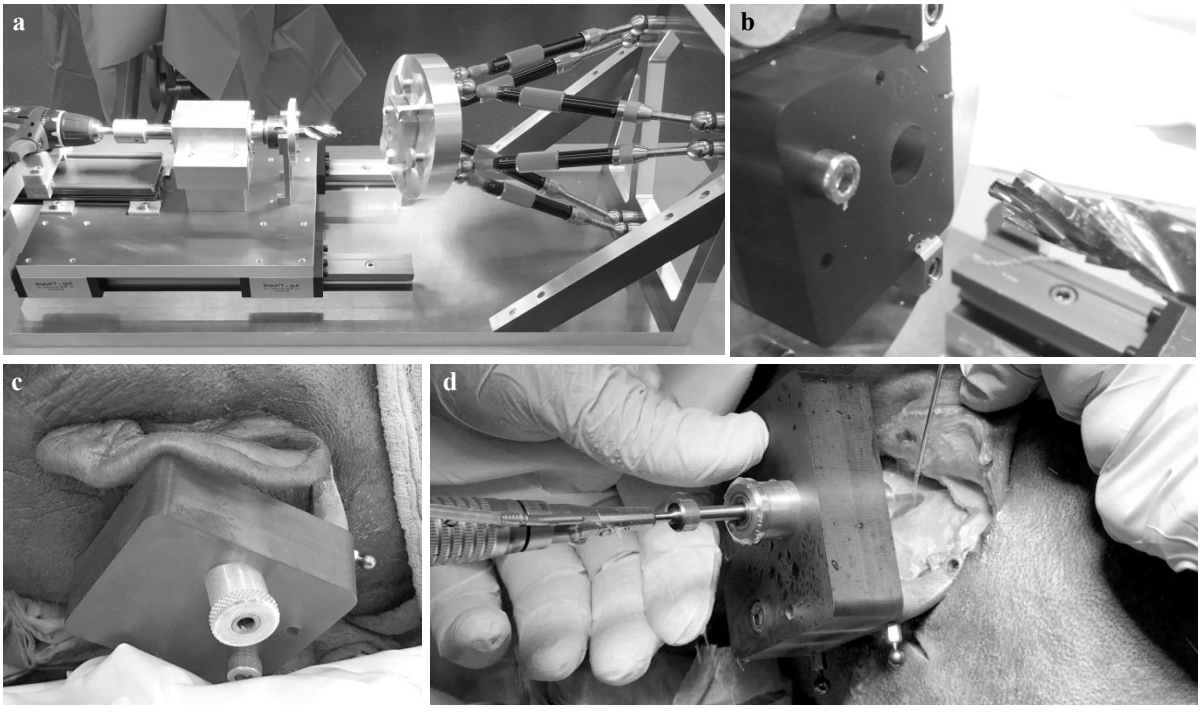


Figure 3: ***(a)*** *Jig Maker with the manual operated drilling machine on the left side and an opposite hexapod-based mechanism for pose setting of the surgical jig.* ***(b)*** *Finalization of the blank by drilling a patient-specific counterbored hole whose axis matches the planned trajectory. The hole serves as mounting port for surgical instruments.* ***(c)*** *Jig mounted on top of the Trifix with inserted drill guide.* ***(d)*** *Drilling the minimally invasive canal down to the middle ear under permanent irrigation.*

In the last specimen, feasibility of inserting a commercially available electrode array was evaluated. In order to perform manual insertion using only common otological instruments, a tympanomeatal flap was created through

a retroauricular incision. This enables access to the promontory and allows for visual control using a conventional microscope. A Flex[20] electrode array (MED-EL corporation, Innsbruck, Austria) was threaded through the drill canal and advanced into the cochlea.

3 Results

Minimally invasive mastoidectomy approach could be successfully performed in all five patients demonstrating the feasibility of this approach (Figure 4). The facial nerve was preserved in all cases. Chorda tympani was sacrificed in one case as accepted during planning. The difference between the planned and the drilled trajectory was found to be 0.40 mm ± 0.30 mm including one case with a deviation of approx. 0.94 mm. During the trials, we could immediately identify a human error in setting the length of one leg of the hexapod as the reason for this unexpected difference. After calculating the target error without the human component in adjusting the Jig Maker, the drilling accuracy in that specimen would be 0.09 mm. Without that outlier the average accuracy in this study was 0.27 mm ± 0.05 mm.

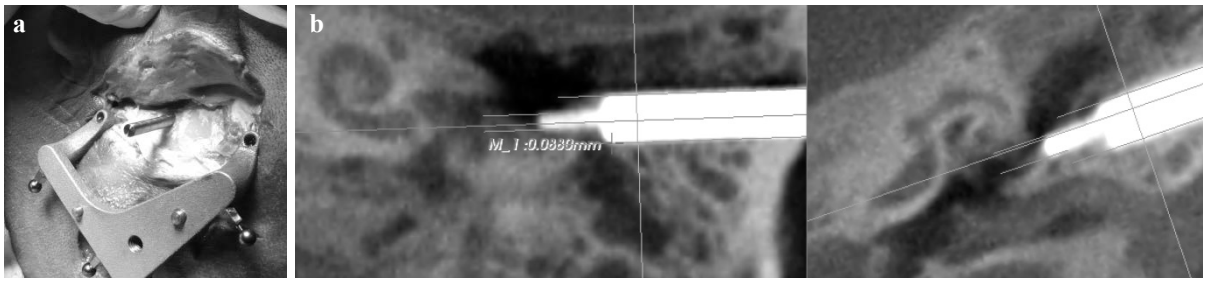


Figure 4: ***(a)*** *Bore hole with inserted titanium rod after removing the surgical jig.* ***(b)*** *Images from the second CBCT scan showing good compliance between the drilled canal and the originally planned trajectory.*

Table 1: *Results of comparing the planned trajectory with the drilled one. Target error was measured as the distance between the two center lines at the planned target point. Average value marked with asterisk (*) is calculated without specimen #03. FRE: fiducial registration error; SD: standard deviation.*

specimen	FRE [mm]	drill depth [mm]	target error [mm]
01	0.077	18.7	0.24
02	0.068	22.4	0.31
03	0.030	22.8	(0.94)
04	0.047	25.9	0.22
05	0.069	23.7	0.31
mean ± SD		22.7 ± 2.6	0.40 ± 0.30 (0.27 ± 0.05)*

One specimen showed an unusual thick skin flap. Due to the limited length of the Trifix's legs bone fixation was not possible through small skin incisions. Instead, a large retroauricular incision was conducted in order to exposure the bone behind the pillar for sufficient fixation of the reference frame.

It was possible to push the electrode array through the minimally invasive drill canal into the middle ear and further through the round window membrane into the inner ear using standard instruments. Post-experimental CBCT imaging confirmed correct location of the electrode inside scala tympani (Figure 5).

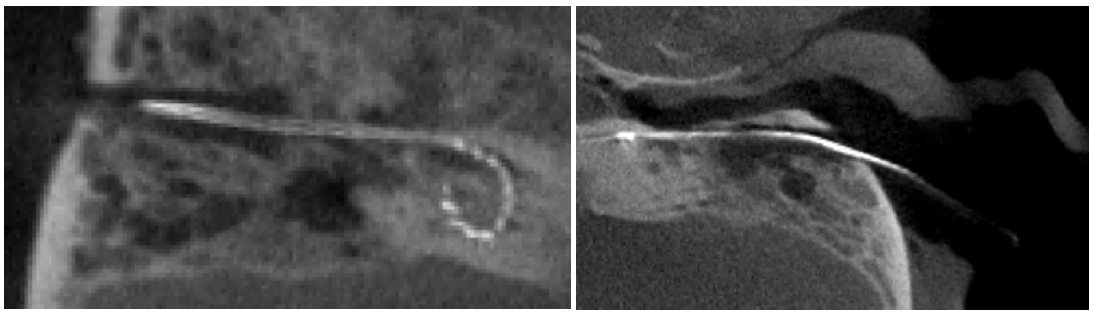


Figure 5: *Radiographic images showing the intra-cochlear position of the lateral wall electrode array as well as its lead running through the drill canal.*

4 Discussion

The presented trials are part of a first preclinical evaluation of the micro-stereotactic surgical targeting system in order to investigate its usability in human cadaveric temporal bones. This study covers the whole surgical workflow including preoperative imaging and planning, manufacturing of the surgical jig, drilling the minimally invasive tunnel through the mastoidectomy posterior tympanotomy approach (MPTA), and finally cochlear implant electrode insertion.

The complete workflow was feasible for all specimens (excluding electrode insertion which was performed in only one case). Facial nerve could be preserved in all cases, which was evaluated both, in post-experimental CBCT imaging, as well as histological evaluation using a conventional mastoidectomy. Apart from one outlier, the deviation between planning and drilled trajectory was sufficient good for minimally invasive cochlear implantation using the MPTA. Accuracy is in the same range as of the Microtable (0.31 ± 0.10 mm, [8])—another, already clinically evaluated micro-stereotactic system [3].

More detailed analysis of the outlier discloses a human error during manual length setting of the legs of the Jig Maker. Further quality measures, like multi user readings, as they are common in the clinical routine, can be implemented easily to limit such human errors. However, the other trials showed that manual length setting is feasible with high precision.

Overall, the whole process seems to be very robust. Cases where a leg of the frame showed insufficient contact to the skull surface (i.e. due to soft tissue in between the base and the skull surface) were clearly identified in radiological images. Based on these findings it was decided to revise bone fixation in two cases to ensure maximal stability of the bone anchorage. The critical steps of manufacturing the surgical jig using the Jig Maker, and drilling inside the inhomogeneous temporal bone contributed very little to the total error. The average error being smaller than the resolution of the CBCT images showed that deflection of the thin drill bit does not seem to be a problem in the hard cortical and mastoid bone.

At the current stage the most time-consuming steps of the whole procedure are planning and electrode insertion. For clinical implementation, a user-friendly planning software, providing tools for (semi-)automated segmentation and trajectory planning, is under development. In addition, we are working on insertion tools [9] to further simplify the process by waiving the tympanomeatal flap.

5 Conclusion

Preliminary results demonstrate the feasibility of a minimally invasive mastoidectomy approach using the developed micro-stereotactic frame. One crucial step is bone anchoring of the carrier frame as one has to ensure that all three bone screws are sufficiently fastened in the skull. This seems to require some amount of experience and practice. Another security relevant issue is manual adjustment of the hexapod's legs, which is a potential source of individual errors. However, this can be overcome either by automatization of the hexapod or by additional security measures of the jig.

Acknowledgements

The presented work was funded in part by the Federal Ministry of Education and Research of Germany (BMBF, grant numbers 13GW0019A/C/E).

Conflict of Interests

The authors TR, MK, SJ, TO, TL and OM declare being limited partners of HörSys IP GmbH & Co. KG that holds a financial stake in OtoJig GmbH, a German company that owns and further develops the described technology.

References

[1] Bielamowicz SA, Cocker NJ, Jenkins HA, Igarashi M. Surgical Dimensions of the Facial Recess in Adults and Childrens. Arch Otolaryngol Head Neck Surg 1988;114:534–7.

[2] Labadie RF, Noble JH, Dawant BM, Balachandran R, Majdani O, Fitzpatrick JM. Clinical Validation of Percutaneous Cochlear Implant Surgery : Initial Report. The Laryngoscope 2008;118:1031–9. doi:10.1097/MLG.0b013e31816b309e.

[3] Labadie RF, Balachandran R, Noble JH, Blachon GS, Mitchell JE, Reda F a., et al. Minimally invasive image-guided cochlear implantation surgery: First report of clinical implementation. Laryngoscope 2014;124:1915–22. doi:10.1002/lary.24520.

[4] Majdani O, Kluge M, Lexow GJ, John S, Kahrs LA, Ortmaier T, et al. Minimally invasive, mini-stereotactic approach for Cochlear Implant Surgery (CI): RoboJig. In: Proc. of 16th Annual Meeting of the German Society for Computer and Robot Assisted Surgery (CURAC), 05.-07.10.2017, Hannover, 2017, p. 37–8.

[5] Kluge M, Lexow GJ, Lenarz T, Majdani O, Rau TS. Investigation on the accuracy of the intra-operative fabrication device of the RoboJig (in German). In: Proc. of 16th Annual Meeting of the German Society for Computer and Robot Assisted Surgery (CURAC), 05.-07.10.2017, Hannover, 2017, p. 262–5.

[6] Thiel W. The preservation of the whole corpse with natural color (in German). Ann Anat 1992;174:185–95.

[7] Oliphant TE. A guide to NumPy. Trelgol Publishing; 2006.

[8] Balachandran R, Mitchell JE, Blachon G, Noble JH, Dawant BM, Fitzpatrick JM, et al. Percutaneous cochlear implant drilling via customized frames: An in vitro study. Otolaryngology - Head and Neck Surgery 2010;142:421–6. doi:10.1016/j.otohns.2009.11.029.

[9] Rau TS, Hügl S, Salcher R, Lenarz T, Majdani O. Hydraulically automated electrode insertion (in German). In: Proc. of 16th Annual Meeting of the German Society for Computer and Robot Assisted Surgery (CURAC), 05.-07.10.2017 Hannover, 2017, p. 131–5.

Awake surgery and intraoperative language mapping: setup and pre-clinical evaluation of a standardized time-synchronized documentation system

A.L. Roethe[1,2], M. Schmidt[3], M.R. Münch[1], T. Picht[1,2]

[1]Department of Neurosurgery, Charité-Universitätsmedizin Berlin, Berlin, Germany
[2]Project ›Image Guidance‹, Image Knowledge Gestaltung – An Interdisciplinary Laboratory, Humboldt-Universität zu Berlin, Berlin, Germany
[3]IT Division, Image Knowledge Gestaltung – An Interdisciplinary Laboratory, Humboldt-Universität zu Berlin, Berlin, Germany

contact: anna.roethe@charite.de

Abstract

Despite a multitude of available neuropsychological test batteries and computer-assisted systems for image-guided interventions, language-eloquent tumors in neurosurgery are still subject to heterogeneous best practice models in both performance and evaluation. Different established mapping modalities and standards for routine preoperative and intraoperative language assessment complicate the development of advanced, safe, patient-tailored protocols for awake language testing. To validate our institutional approach and facilitate the integration of additional patient-specific functional data for surgical guidance, we set up a multi-source intraoperative documentation system and conducted a pre-clinical feasibility study. The overarching goal was to guarantee a reproducible and interoperable solution that can be implemented into surgical routine workflows providing comprehensive high-quality data on patient performance during language testing.

Keywords: awake surgery, language mapping, intraoperative documentation system

1 Problem

In eloquent brain tumors, careful balancing of tissue removal and functional integrity is a major part of the surgical strategy. Non-invasive brain mapping helps to assess the patient's functional profile before the operation. During awake surgery, additional mapping and monitoring data are acquired through both active task performance and passive recordings with and without electrical stimulation of brain tissue. Awake surgery remains the gold standard for treatment of language-eloquent tumors to date [1]. However, interpreting and classifying those language mapping results beyond ad-hoc intraoperative decision-making remains a logistical challenge. Several data sources support the surgical removal of tumor tissue in language-eloquent areas: neuronavigation including preoperative language mapping data (e.g., navigated transcranial magnetic stimulation (nTMS) and white matter tractography), direct electrical brain stimulation during microsurgery, intraoperative neuromonitoring (IONM) (including surveillance of epileptic activity and motor function), and patient performance in language tasks (Fig. 1). It takes a team of specialists to combine and apply that intraoperative information to the individual case. Still, postoperative evaluation is complex due to different data formats, single-source data recording streams and lack of sufficient temporo-spatial information. This also affects the validation and further development of both non-invasive and invasive brain mapping techniques. Thus, the aim of this project was to find a clinically applicable solution to a) document all information sources time-synchronously for subsequent in-depth analysis and correlation with clinical outcome, and b) prepare the comparison of language performance across individual patients with respect to stimulation site, language task and electrical brain activity. Device interoperability had to be ensured, as the recording setting would comprise five different intraoperative sources from different manufacturers.

2 Material and Methods

Our standard workflow includes preoperative language examination by clinical tests and nTMS, followed by presurgical planning using nTMS-positive and negative spots and tractography of major language fascicles

(iPlan Net 3.6, Fibertracking Element 1.0, Brainlab AG, Munich). Anatomical and functional data is then being transferred to the neuronavigation system (see below). Awake language mapping tasks (e.g., picture naming) are shown to the patient on a tablet computer (see below) during cortical (bipolar, using a fork probe) and subcortical (monopolar, using the Mapping Suction Probe by Raabe, Inomed, Emmendingen) direct electrical stimulation (DES). Patient performance is reported to the surgeon and video-recorded for subsequent in-detail analysis of error types. Areas identified as language-eloquent during stimulation are saved as spatial coordinates in the neuronavigation software as well as microscope screenshots. Rather than for purposes of documentation and scientific evaluation, existing custom-made single-center solutions such as the intraoperative testing platform [2] or the intraoperative examination monitor for awake surgery [3] have been designed to support the task performance in patients and/or share information on the patient's performance directly with the surgeon on a dedicated screen. After an extensive market research, we decided to use a multi-component system consisting of commercially available devices: a laptop (MacBook Pro, Apple Inc., Cupertino), a screen splitter (MultiView 16, BlackMagic Design, Port Melbourne) and a video recording adapter (Intensity Shuttle, BlackMagic Design, Port Melbourne), interconnected to sources via converters where necessary (HDMI/SDI Mini Converter, BlackMagic Design, Port Melbourne). The recorded stream includes the surgical microscope (OPMI Kinevo 900, Carl Zeiss Meditec AG, Oberkochen)[1], a patient face camera with clip microphone (HD Camcorder, Panasonic Corporation, Kadoma)[2], a tablet computer (LifeTab P10505 64GB 10,1" with Android 5.1, Medion AG, Essen)[3] with the intraoperative language task set in PowerPoint format, the neuronavigation system (Curve with Cranial Navigation 3.1, Brainlab AG, Munich)[4] and the neuromonitoring cart (ISIS IOM System, Inomed, Emmendingen)[5] (Fig. 2). As preset recording layout grids are currently limited to 2x2, 3x3 or 4x4 split screens, we created a shell script using FFmpeg (FFmpeg project, GNU LGPL 2.1/GPL 2.0) that removes all redundant screens creating a 2x2 grid layout with the fifth source added using the picture-in-picture feature (4+1 grid). This customization is done in post-processing. The complete setup has been tested for feasibility and consistence during a number of laboratory and pre-clinical tests in the operating room (OR).

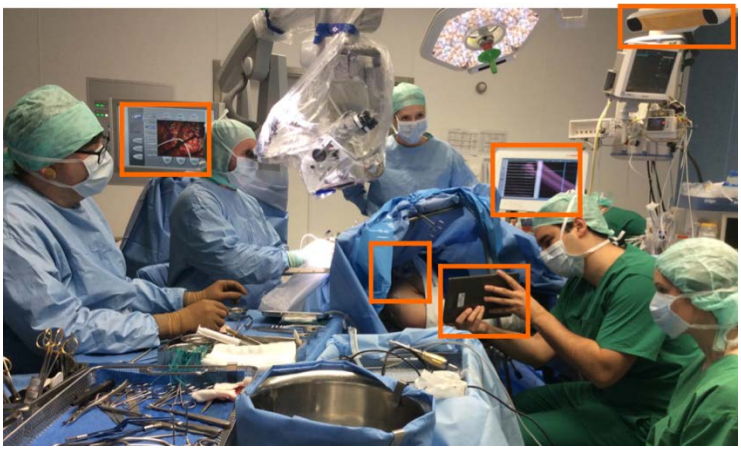


Fig. 1*: Current intraoperative setup of awake mapping technology (highlighted from left to right: surgical microscope video, patient face, tablet computer, neuromonitoring, navigation camera)*

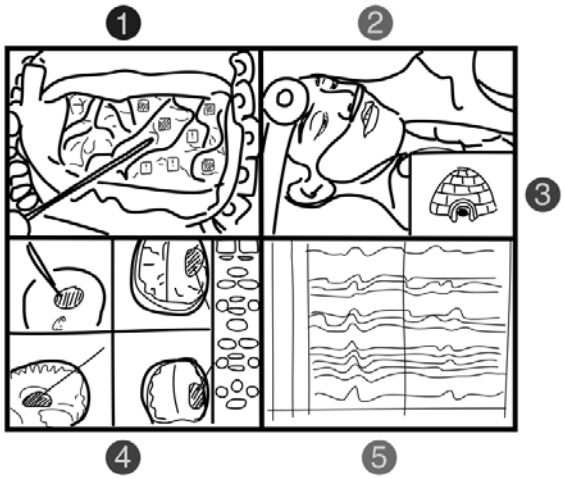


Fig. 2: *Schematic view of intraoperative documentation sources*

3 Results

For pre-clinical simulation of intraoperative routine settings, the system has been positioned on a mobile equipment cart on the right side of the OR table beside the anesthesia team. During tests, all data sources have been successfully connected with the documentation system and recording software (Blackmagic Media Express, BlackMagic Design, Port Melbourne), which confirmed the desired interoperability of its components (Fig. 3). In comparison with other customized solutions, the system's complexity proved compatible with clinical conditions so far (e.g., required space, time needed for setup, access to and wiring of devices). The video signal of all sources has been robust, high-quality video post-processing being fast and reliable. Time synchronization through the recording software proved sufficiently precise for comparative analysis of stimulation site and response during proband testing. Due to general availability and affordability of all components used, the setup can be easily reproduced for future work.

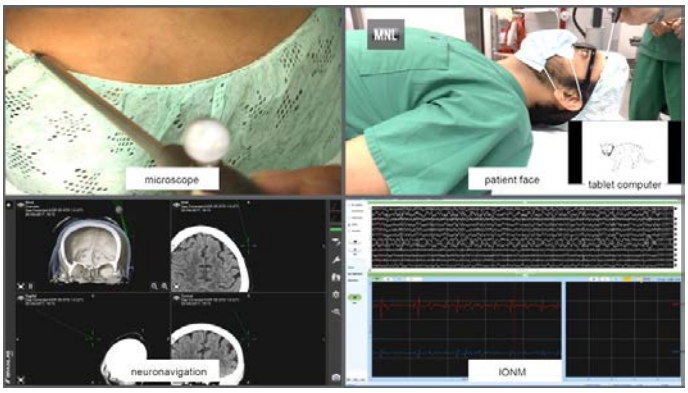


Fig. 3: *Pre-clinical test of recording setup with proband in the OR setting*

4 Discussion

While electrical mapping of brain functions during awake surgery constitutes one of the most insightful neuroscientific setups to enhance our understanding of brain function, the documentation of intraoperative interventions and resulting behavioral effects is rudimentary. In the process of updating our awake surgery protocol, we decided to standardize its documentation and evaluation in preparation of future adjustments and efforts of mapping validation. Instead of indirect filming of screens, all sources are recorded directly. Particularly the integration of IONM in such a documentation setup has, to our knowledge, not been done before, facilitating a better linguistic error categorization by using a more advanced analysis of electroencephalography and stimulation data to differentiate between causes of task errors. The video signal use of the neuromonitoring cart allows for situative reconstruction of the IONM examination by showing what was visibly measured at a certain point during surgery. It does not, however, enable the analyst to alter display parameters such as amplification or write speed, limiting the analysis to the intraoperative status of data utilization. Further exploration and quantification of this restriction as well as of any potential signal delay will follow with more extensive live recordings during upcoming clinical tests.

5 Conclusion

In despite of most surgical systems' integrated functions for video recording and connector outputs, there is no plug-and-play solution for surgical documentation in the OR. We were able to integrate devices from multiple vendors in a single time-synchronous documentation system adjusted to actual clinical and scientific demands. By replacing previous standards of observer-based paper documentation, we solved the preexisting dilemma of post-hoc analysis of comprehensive awake language mapping at our institution. In the next phase, we will evaluate a series of 10 consecutive awake surgeries using this documentation setup. Particularly the patient-tailored selection of cognitive and executive tasks is expected to greatly benefit from standardized assisted documentation and evaluation.

References

[1] Ferracci FX, Duffau H, *Improving surgical outcome for gliomas with intraoperative mapping*, Expert Rev Neurother, 18(4), 333-341 (2018).

[2] Morrison MA, Tam F, Garavaglia MM, Golestanirad L, Hare GM, Cusimano MD, Schweizer TA, Das S, Graham SJ, *A novel tablet computer platform for advanced language mapping during awake craniotomy procedures*, J Neurosurg, 124(4), 938-44 (2016).

[3] Yoshimitsu K, Maruyama T, Muragaki Y, Suzuki T, Saito T, Nitta M, Tanaka M, Chernov M, Tamura M, Ikuta S, Okamoto J, Okada Y, Iseki H, *Wireless modification of the intraoperative examination monitor for awake surgery*, Neurol Med Chir (Tokyo), 51(6), 472-6 (2011).

Eine Beurteilung der Anwendbarkeit von hyperspektralbasierter Bildgebungstechnologie bei viszeralonkologischen Eingriffen

M. Maktabi[1], H. Köhler[1], I. Gockel[2], B. Jansen-Winkeln[2], J.P. Takoh[2], S.M. Rabe[2], T. Neumuth[1], C. Chalopin[1]

[1]Universität Leipzig, Innovation Center Computer Assisted Surgery (ICCAS), Deutschland
[2]Universitätsklinikum Leipzig, Klinik und Poliklinik fürfür Viszeral-, Transplantations-, Thorax- und Gefäßchirurgie, Leipzig, Deutschland

Kontakt: marianne.maktabi@medizin.uni-leipzig.de

Abstract

Hyperspektrale Bildgebung stellt eine neue Möglichkeit dar um Gewebeperfusion bei viszeralchirurgischen Eingriffen präziser einschätzen zu können. Im Rahmen dieser Arbeit wurde ein derzeitig verfügbares Hyperspektralkamerasystem bei verschiedenen viszeralchirurgischen Anwendungsszenarien getestet. Danach wurden funktionale und nichtfunktionale Anforderungen an ein solches System zur optimalen klinischen Nutzung abgeleitet. In weiteren Arbeiten gilt es diese umzusetzen, um ein HSI-System erfolgreich in den klinischen Alltag zu integrieren.

Keywords: Hyperspektrale Bildgebungstechnologie, Viszeralchirurgie, Operationssaal

1 Problemstellung

Jede dritte Krebserkrankung in Deutschland betrifft die Hohlorgane wie beispielsweise den Darm (2014: 83.370 Fälle) [1]. Bei Operationen des Magen-Darm-Trakts spielt eine genaue Beurteilung der Gewebedurchblutung eine wichtige Rolle für den Erfolg der Eingriffe. Somit erfolgt standardweise die herkömmliche intraoperative Beurteilung der Durchblutungssituation durch den tastbaren Puls der vorsorgenden Gefäße und die Rosigkeit bzw. Vitalität der gastrointestinalen Strukturen. Das komplexe System der Durchblutung des Magen-Darm-Trakts ist für das bloße Auge des Chirurgen im operativen Situs nicht mit absoluter Genauigkeit und objektiv-reproduzierbar zu evaluieren. In den letzten Jahren wurde die Near-InfraRed (NIR)-Fluoreszenzbildgebung mit IndoCyanine-Grün (ICG) im Rahmen von Forschungsprojekten ausgewertet. Die Nutzung der Bildgebungsmodalität während der Operation hat beigetragen, die Komplikationen zu reduzieren [2, 3]. Dennoch stellen hierbei die subjektive Evaluation der Fluoreszenz-Intensität durch die visuelle Beurteilung des Chirurgen sowie die Invasivität der Methodik mittels in vivo-Verabreichung des Farbstoffs gravierende Limitationen der Technologie dar. Zudem stellt das Verfahren nur die Durchblutung von Gewebe dar, es ist aber keine funktionelle Bildgebungsmodalität. Somit sind objektive und präzise, quantitative und möglichst nicht-invasive Messmethoden der Perfusion (Durchblutung und Sauerstoffsättigung) erforderlich, um klare Grenzwerte zwischen normaler und kritischer Perfusion aufzuzeigen und konsekutiv Komplikationen und deren Folgen zu vermeiden.

Hyperspectral-Imaging (HSI) ist ein relativ neues Bildgebungsverfahren, das für die Erkennung von Strukturen und für die Auswertung der Gewebedurchblutung im Viszeral- und Thorax-Bereich im Forschungsfeld vielversprechende Ergebnisse gezeigt hat. Das Verfahren kombiniert das Prinzip der Spektroskopie mit der Bildgebung, die man von herkömmlichen visuellen Kameras kennt und erfordert keine Gabe eines Kontrastmittels. Zudem können Messdaten über größere flächenhafte Messbereiche aufgenommen werden, wodurch die Reduktion lokaler Variationen und damit stabilere Resultate ermöglicht werden. Das System generiert Absorptionsspektren, die für die Analyse der Daten weiter verarbeitet werden sollen. In der Literatur vorgestellte experimentelle und auch klinische Untersuchungen belegen die grundsätzliche Nutzbarkeit der HSI Technologie und das hohe Potential zur objektiven Bestimmung relevanter Durchblutungsparameter [4,5].

Im Rahmen dieser Arbeit wurde basierend auf ersten viszeralchirurgischen Untersuchungsreihen die intraoperative Anwendbarkeit hyperspektralbasierter Technologie zur Beurteilung von Gewebedurchblutung analysiert.

2 Material und Methoden

2.1 HSI-System

Die HSI-Aufnahmen während der Eingriffe wurden mit dem TIVITA™ Tissue Kamerasystem von der Firma Diaspective Vision (Am Salzhaff, Deutschland) durchgeführt (Abbildung 1). Die Kamera nimmt den Spekralbereich von 500 bis 1000 nm auf (Auflösung: 5nm). Die Aufnahmefläche kann bis zu 13x10 cm groß sein. Das Kamerasystem besteht aus einen Push-Broom-Spektrographen, welcher für die spektrale Datenaufnahme entlang der x-Achse zuständig ist, und einem CMOS-Kamerasensor, welcher mit Hilfe einer USB3-Schnittstelle und einen Schrittmotor die Daten des Spektrographen in y-Richtung aufnimmt. So entsteht ein sogenannter 3D-Datenwürfel. Die x- und y-Achse bilden die räumlichen Dimensionen ab, wobei die z-Achse die Wellenlängen repräsentiert. Als Beleuchtungseinheit wurden acht 20 W Halogenlampen verwendet. Im Abstand von circa 30 cm zum Objekt wurden mit Hilfe der Kamera die Aufnahmen durchgeführt. Eine detaillierte Beschreibung des technischen Prinzips ist in der Arbeit von Holmer et al. [6] beschrieben.

Die Analysesoftware verarbeitet die Daten und generiert ein RGB-Bild (640 x 480 Bildpunkte) sowie je ein Falschfarbbild für die Parameter Gewebeoxygenierung (StO2), Gewebe-Hämoglobin- (THI), NIR-Perfusion- und Gewebe-Wasser-Index (TWI). Die Berechnung dieser Parameter erfolgte mit Hilfe eines kameraspezifischen Softwarepakets namens TIVITA™ Suite. Diese Software ermöglichte es die Aufnahmen des Patienten näher zu untersuchen.

2.2 Klinische Anwendungsszenarien

Es wurden drei verschiedene intraoperative Szenarien untersucht.

In der ersten Untersuchungsreihe wurde eine Schlauchmagenbildung bedingt durch Speiseröhrenkrebs untersucht. Dazu wurde der gebildete Schlauchmagen direkt vor Verbindung der Gewebe mit einer HSI-Kamera aufgenommen. Ziel der Aufnahme war die Analyse des Schlauchmagens mittels Hyperspektralbildgebung (HSI) bei Patienten ohne vs. mit ischämischer Konditionierung um kritische Durchblutungszustände frühzeitig zu diagnostizieren. Aus den Aufnahmen mit hoher räumlicher und spektraler Auflösung wurde je ein Falschfarbbild für die Parameter StO2 und TWI und für einen angepassten THI, Organ-Hämoglobin-Index (OHI) genannt, und die Mittelwerte deren zur Gegenüberstellung der verschiedenen Patientengruppen berechnet. Die Lokalisation der späteren Anastomose ist durch eine Pinzette auf den HSI-Bildern markiert wurden.

In einer weiteren Untersuchungsreihe wurde bei Patienten nach zentraler Gefäßabsetzung und vor Durchtrennung bei Kolonresektionen die Perfusion des Darmes gemessen. Ziel dessen war es die subjektive Einschätzung des Chirurgen mit den objektiven Messungen der Hyperspektral-Bildgebung hinsichtlich der Festlegung der Resektionsgrenzen zu untersuchen. Am Patientensitus wurde dazu nach der Durchtrennung jede Minute eine HSI-Aufnahme durchgeführt (n = 15). Auf dem berechneten Gewebeoxygenierungsbild konnte dann die Resektionsgrenze identifiziert werden.

Um objektive Kriterien zur Beurteilung der Darmperfusion zu erhalten, wurde ein Patient mit Darmischämie hyperspektralbasiert untersucht. Dazu wurde mittels HSI-Kamera der Dünndarm im Situs und das Operationspräparat untersucht und folgende Parameter erhoben: Gewebeoxygenierung, Nahinfrarot-Perfusions-Index und Gewebe-Wasser-Index. Danach wurde das gesamte Spektrum von 500 bis 1000 nm auf Auffälligkeiten hin untersucht. Auch die Gewebeoxygenierung, Perfusion und der Gewebewasser-Index des Darms wurden näher betrachtet.

Abbildung 1*: Verwendetes HSI-System von der Firma Diaspective Vision GmbH (Am Salzhaff, Deutschland) zur Aufnahme während chirurgischer Eingriffe mit einer Standfläche von ca. 0,25 m² ohne Schwenkarm(rot: Beleuchtungseinheit mit Objektiv und Spektralkamerasystem an Schwenkarm, blau: Softwareeinheit mit Stativ, grün: Stromversorgung),*

3 Ergebnisse

Die klinischen Ergebnisse der drei verschiedenen Untersuchungsreihen waren vielversprechend. Das HSI-System konnte zeigen, dass die Gewebeoxygenierung, der OHI und der TWI im markierten Bildbereich bei den zwei verschiedenen Patientengruppen ohne und mit Präkonditionierung in der ersten Untersuchungsreihe signifikant verschieden waren. Somit kann man die hyperspektrale Bildgebungstechnologie grundsätzlich bei offenen viszeralchirurgischen Eingriffen, wie z.B. bei der Ösphagusresektion, einsetzen. Auch zur Darstellung der Perfusionsgrenze im Darm nach Resektionen in der zweiten Untersuchungsreihe zeigte das Verfahren eine gute Anwendbarkeit. Die HSI-Kamera bietet eine objektivere Entscheidungshilfe zum Finden der Resektionsgrenze. In der dritten Untersuchung zeigten sich die Indices für Gewebeoxygenierung und Perfusion von nekrotischen Darmabschnitten erniedrigt (Abbildung 2). Der Gewebe-Wasser-Index war in den minderperfundierten Dünndarmabschnitten erhöht. Auch zur Erkennung von Nekrose, zeigte sich eine typische Bande des Absorptionsspektrums bei einer Wellenlänge von 630 nm (Abbildung 3). Somit ist auch mittels intraoperativer Hyperspektralbildgebung, die Perfusion grenzwertig perfundierter Darmabschnitte bei akuter Ischämie im Darm objektivierbar.

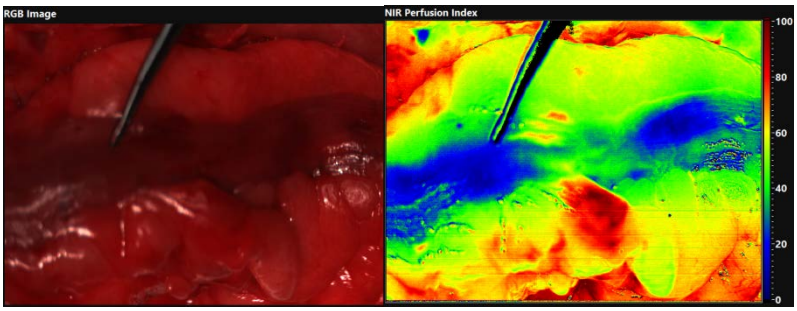


Abbildung 2 Ein Beispiel für ein intraoperatives Bild eines schlecht durchbluteteten Gewebes (mit Pinzette gekennzeichnet) (links RGB-Bild, rechts Nahinfrarot-Perfusion-Index)

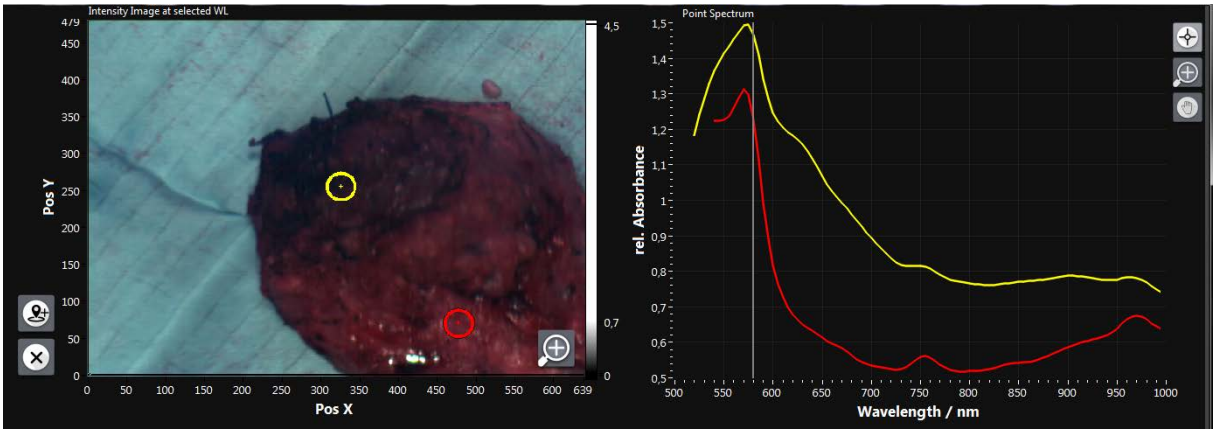


Abbildung 3 Ein Beispiel eines Präperats mit Absorptionsbande bei einer Wellenlänge von 630 nm für leicht nekrotische Abschnitte (gelb: nekrotisches Gewebe; rot: nicht-nekrotisches Gewebe)

Jedoch bietet bisher das HSI-System noch keine intraoperative Standardnutzbarkeit an. Es treten Verzögerungen bei Verwendung des Systems im OP-Ablauf auf. Zudem können die erhaltenen HSI-Daten noch nicht automatisch viszeralmedizinisch interpretiert werden.

Somit können Anforderungen an ein hyperspektralbasiertes OP-Kamerasystem definiert werden. Als nichtfunktionale Anforderungen sind besonders eine optimale Standfläche für das Kamerasystem mit Schwenkarm und eine einfache Strukturierung der Bedienelemente zur benutzerfreundlichen Anwendung durch das klinisches Personal hervorzuheben. Zudem muss eine schnelle Berechnung der jeweils interessanten Parameter in unter 5 Sekunden geschehen. Die Standfläche für das Kamerasystem mit Schwenkarm muss so dimensioniert werden, das man es auch in Operationssälen mit geringer Größe einfach verwenden kann. Als funktionale Anforderungen sind eine kabellose Energieversorgung, eine robuste und schwenkbare Kameravorrichtung mit automatischer Fixierung und Ausrichtung zum Situs, eine Datenintegration in das Krankenhausinformationssystem, eine Berücksichtigung vom bewegten Situs bei der Aufnahme (z.B. Atmung) und eine präzise medizinische Interpretation der HSI-Daten erforderlich.

4 Diskussion

Im Rahmen dieser Arbeit wurde die klinische Anwendbarkeit des HSI-Kamerasystems zur Beurteilung von Gewebeperfusion im OP untersucht und Anforderungen an ein solches System aufgestellt. Diese Technologie könnte perspektivisch in klinischer Umgebung dem Behandler eine schnelle Kategorisierung des aktuellen Zustands der Gewebsperfusion und eine präzise Verlaufsbeurteilung nach revaskularisierenden Prozessen ermöglichen.

Wir konnten Spezifikationen eines intraoperativen viszeralchirurgischen HSI-Systems aufstellen. Diese Spezifikationen können teilweise auch auf andere klinische Anwendungen wie beispielsweise auf HNO-Eingriffen übertragen werden. Bei neurochirurgischen Eingriffen wird meist bei Operationen mit einen Mikroskop gearbeitet. Deshalb müssen andere Spezifikationen berücksichtigt werden. Beispielsweise ist eine Miniaturisierung nicht zwingend erforderlich, da die HSI-Kameraeinheit direkt an ein Mikroskop ankoppelbar ist.

Die in dieser Arbeit untersuchten Operationen waren offene chirurgische Eingriffe. Bei minimalinvasiven chirurgischen Eingriffen müssten spezielle Anforderungen hinsichtlich eines endoskopischen Systems erfüllt werden. Beispielsweise ist die einfache Handhabbarkeit eines HSI-Systems dabei dringend erforderlich, da dieses manuell während Operationen gehalten werden muss. Zudem muss die Auswertungssoftware noch genauer dem Anwendungsfall angepasst werden, um beispielsweise die Perfusionsgrenze bei viszeralchirurgischen Eingriffen schnell und präzise visuell darstellen zu können.

5 Zusammenfassung

Im Rahmen dieser Arbeit wurde die Anwendbarkeit des HSI-Kamerasystems im OP untersucht und Anforderungen an das System aufgestellt. Diese Technologie könnte perspektivisch in klinischer Umgebung dem Behandler eine schnelle Kategorisierung des aktuellen Zustands der Gewebsperfusion und eine präzise Verlaufsbeurteilung nach revaskularisierenden Prozessen ermöglichen.

Wir konnten mit Hilfe unserer klinischen Anwendungsszenarien Anforderungen an das System definieren. Bisher ist kein HSI-System mit diesen Eigenschaften auf den Markt erhältlich. In zukünftigen Arbeiten müssen diese erarbeitenden Anforderungen umgesetzt werden, um ein HSI-System erfolgreich in den klinischen Alltag integrieren zu können.

Referenzen

[1] T. Delko, P. Kirchhoff, R. Rosenthal, J. Schäfer, M. Kraljević, and C. Kettelhack, "Colon Perfusion Patterns During Colorectal Resection Using Visible Light Spectroscopy," *World Journal of Surgery*, vol. 41, no. 11, pp. 2923–2932, http://link.springer.com/10.1007/s00268-017-4100-x, 2017.

[2] T. Wada, K. Kawada, R. Takahashi, M. Yoshitomi, K. Hida, S. Hasegawa, and Y. Sakai, "ICG fluorescence imaging for quantitative evaluation of colonic perfusion in laparoscopic colorectal surgery," *Surg Endosc*, vol. 31, no. 10, pp. 4184–4193, http://link.springer.com/10.1007/s00464-017-5475-3, 2017.

[3] M. D. Jafari, S. D. Wexner, J. E. Martz, E. C. McLemore, D. A. Margolin, D. A. Sherwinter, S. W. Lee, A. J. Senagore, M. J. Phelan, and M. J. Stamos, "Perfusion assessment in laparoscopic left-sided/anterior resection (PILLAR II): A multi-institutional study," *Journal of the American College of Surgeons*, vol. 220, no. 1, 82-92.e1, 2015.

[4] M. S. Chin, B. B. Freniere, Y.-C. Lo, J. H. Saleeby, S. P. Baker, H. M. Strom, R. A. Ignotz, J. F. Lalikos, and T. J. Fitzgerald, "Hyperspectral imaging for early detection of oxygenation and perfusion changes in irradiated skin," *Journal of biomedical optics*, vol. 17, no. 2, p. 26010–26010, 2012.

[5] M. Denstedt, B. S. Pukstad, L. A. Paluchowski, J. E. Hernandez-Palacios, and L. L. Randeberg, "Hyperspectral imaging as a diagnostic tool for chronic skin ulcers," in *SPIE BiOS*: SPIE, 2013, 85650N.

[6] A. Holmer, F. Tetschke, J. Marotz, H. Malberg, W. Markgraf, C. Thiele, and A. Kulcke, "Oxygenation and perfusion monitoring with a hyperspectral camera system for chemical based tissue analysis of skin and organs," Physiological measurement, vol. 37, no. 11, pp. 2064–2078, 2016.

Development of an improved ultrasound navigation system for neurosurgical procedures

JGC Aragon[1], D Lindner[2], S Arnold[3], A Schmitgen[3], C Chalopin[1]

[1]Innovation Center Computed Assisted Surgery, University Hospital Leipzig, Leipzig, Germany
[2] Department of Neurosurgery, University Hospital Leipzig, Leipzig, Germany
[3] Localite GmbH, Sankt Augustin, Germany

JesusGuillermo.CabalAragon@medizin.uni-leipzig.de

Abstract

The use of intraoperative ultrasound (iUS) imaging supports the neurosurgeon to identify tumorous tissue during brain tumor operations. 3D image data provides a complete overview of the surgical field. The commercial SonoNavigator product (Localite GmbH, Sankt Augustin, Germany) includes a navigation system connected to an ultrasound device and an optical tracking system. Since there is still high user-machine interaction, this work presents a package of three different tools to provide support to neurosurgeons into the operating room. First module is a tool called image based connector, for the automatic identification in real time of the parameter values in the US image data. Second one is a tool, which provides to neurosurgeons the visualization of the brain tumor, as well as other cerebral structures in the preoperative MR data. Third one is a registration tool, which performs the overlapping process between a pre-operative brain tumor model and the live 3D iUS image data.

Keywords: Automation systems, MR, Neuronavigation system, Template matching, US images, US parameters.

1 Problem

The surgical navigation system has experienced vast development over the past two decades for minimizing the risks of collateral damage and improving the precision of the surgery. Navigation in surgery is a high-up example of today's technological capabilities being applied into medicine. It has arisen as one of the most reliable developments of technology as it continues to turn surgical interventions into safer and well controlled invasive procedures. In neurosurgery, navigation has encourage technical progress, enabled more daring procedures. The merge of technology and surgery is stronger when faced with the challenge to operate one of the most delicate organ of the human body, as the brain is. The brain is confined in a close-fitting space together with other vital structures, like blood vessels, which can cause major functional deficits if damaged. Due to the affluence of risky structures surgical invasion has to be limited.

Furthermore, intraoperative ultrasound (iUS) imaging is standardly performed in the neurosurgery to support the surgeon in the identification of tumorous tissue during brain tumor operations [1]. Therefore, navigation systems integrate ultrasound devices for neurosurgical purposes, also called ultrasound navigation systems; perform an overlapping process of the iUS image data with preoperative MR or CT data. This supports the surgeon for a better understanding and analysis of the iUS images [2].

Nowadays a huge variety of navigation systems and ultrasound integration are offered. For example the models of the company Brainlab GmbH (München, Germany) include one US device into the complete system providing the advantage to take control of all variables and parameters involved. However, it is not so flexible to couple with other devices in the same manner. Moreover, segmentation tools are helpful to delineate target or risk structures, like tumor and blood vessels, for visualization purposes [3]. Indeed tumor margins can be smoothly defined in the image data and the monitor of the navigation system is sometimes positioned far away from the surgeon. Most models of navigation systems include only basic segmentation tools. More powerful software tools are often problematic because of the lack of user interface between the segmentation tool and the navigation system.

The aim of this work is the further development of a ultrasound navigation system to provide optimized support to neurosurgeons into the operating room (OR). This research project offers the advantage of having embedded software tools which includes automatically set the visualization and 3D reconstruction of the iUS data.

Moreover, a software tool for the segmentation and generation of a preoperative patient model and for the enhancement of the brain tumor contours in the 3D iUS image data was developed [4].

2 Material und Methods

The components of the overall system are the commercial SonoNavigator® (Localite GmbH, Sankt Augustin, Germany) including an optical tracking system, an US device and a research platform. The research platform is a laptop and includes: 1) a tool for the automatic setting of visualization and reconstruction of the iUS data, 2) an embedded segmentation tool for the generation of a patient model based on preoperative MR data, and 3) brain tumor contour enhancement tool in the iUS data. The navigation system and the research platform are connected using a local network in order to exchange information and data.

2.1 Automatic setting of visualization and reconstruction

Some ultrasound navigation systems integrate a complete system including the US device and it has therefore access to the values of the US parameters. These parameters, as depth, flip image, current probe and US device model, are necessary for the correct visualization and reconstruction of 3D US volume. For that, the calibration data corresponding to the currently used US parameters is automatically selected. On the other hand, the SonoNavigator® product is a neuronavigation system which can be connected to any US device model. This solution is more flexible and practical. However, the US parameters are not directly accessible anymore. During surgical interventions in the OR, user has to manually select the calibration file which corresponds to the parameter values used during the acquisition to activate the visualization and/or the reconstruction steps. The selection of the wrong file leads up to incorrect visualization and 3D reconstruction. Therefore, the automation of the selection of the correct calibration data makes more safety and reliable the use of iUS imaging during surgery. For that, a solution to identify the US parameter values which are only available through the US image data visualized on the monitor was developed.

The automatic identification of the US parameters in the iUS images is performed by the research platform and recognize them as annotations, these parameters are US model, currently probe (in this case three different prober were used, one linear 5.0L45, one curve C8-5 and one phase P4-2), depth (four depths were tested such as 4, 6, 8 and 10 cm) and flip. The tool for this task, called image based connector, includes image processing algorithms. The identified current US parameters are then automatically communicated to the navigation system. A new tool was implemented in the navigation system to automatically select the corresponding calibration data. Therefore, the implemented image based connector includes two parts. The first part is a tool for the creation of a database of template images including the parameter values that are commonly used during brain tumor operations. The second part is the tool for the automatic identification of the parameter values in the live iUS images by performing a matching between the iUS image and the templates of the database. Figure 1 shows a description of the tool.

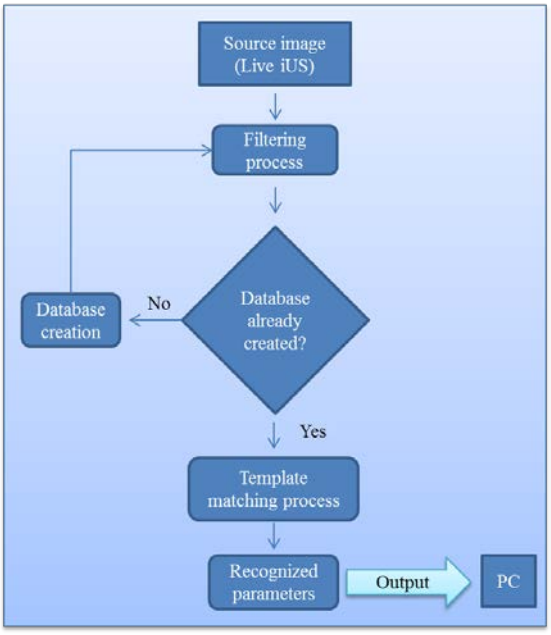


Figure 1: *Image based connector internal process.*

2.2 Generation of a patient model based on preoperative MR data

The delineation of target and risk structures and therefore the generation of a preoperative patient model based on MR data are helpful for visualization purposes [5]. To generate a patient model based on preoperative MR data it is necessary to load DICOM files. DICOM source files are loaded directly from the CD patient data. DICOM files are imported into the tool as a 3D volume to visualize through all slices. This research tool provides neurosurgeons to visualize in three different views (Axial, Sagittal and Coronal) to have a better focus of the brain tumor. Neurosurgeons can freely choose the more convenient view for the segmentation. Then, a manual delineation is performed by the neurosurgeon through a couple slices, and then a model of the tumor is generated by interpolating the slices. The lateral ventricles and blood vessels are possible to segment too. A region growing method is applied. All these extracted structures are saved as VRML format into a directory shared with the computer including the navigation system. They are automatically visualized on the monitor of the navigation system when pressing the button save the generated model in the research platform. A brief description of the internal process is shown in figure 2.

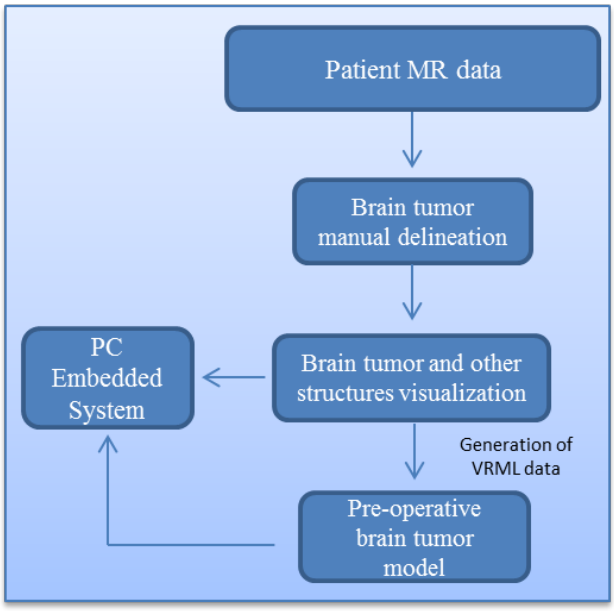


Figure 2*: Internal process of the Automatic setting of visualization and reconstruction tool. As a first step MR data are acquired, and then the delineation process by the user is performed, as well the identification of structures such as lateral ventricles and blood vessels. Finally a brain tumor model is obtained.*.

2.3 Brain tumor contour enhancement in the iUS data

To provide a better understanding of the iUS images data, the enhancement of the brain tumor contour is performed. First the segmentation of the brain tumor in the MR data is performed with the previous tool. Then with the coordinates of the brain tumor model, a ROI is established. Then a thresholding process is applied to the MR and iUS images to extract the hyperechogenic structures considered as features in the registration process. The segmented ROIs are automatically aligned to provide the parameters of a transformation function T. The brain tumor is finally registered using T. This provides the tumor contour enhancement in the iUS image. Within the registration/overlapping tool a variety of techniques are possible to apply. Rigid, Affine, Translation and Similarity methods are available on the tool. Figure 3 shows the internal process of the tool.

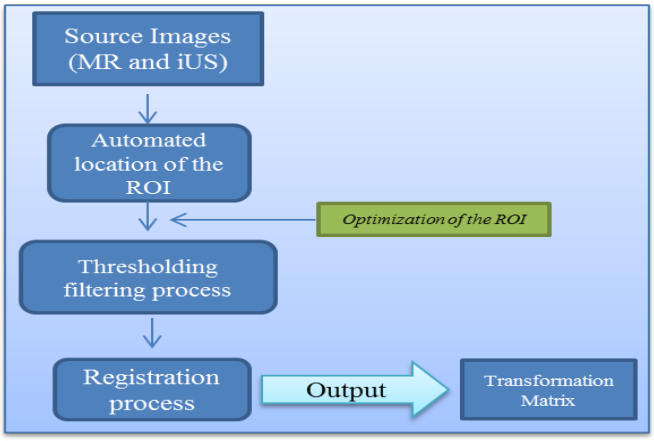


Figure 3*: Brain Tumor Contour Enhancement Tool.*

3 Results

A demonstrator including the navigation system, a Siemens Sonoline G60S US device, the research platform and a local network was developed at our institute. The overall system was used to test the implemented tools.

3.1 Automatic setting of visualization and reconstruction

The following parameters were identified: US device model, name of the probe currently used (3 different US probes, one linear 5.0L45, one curve C8-5 and one phase P4-2), image depth (4, 6, 8 and 10 cm) and orientation (standard view or mirrored). Once the correct parameters have been successfully recognized, they are sent to the PC through the socket connection. Then the navigation system processes them and fit with the correct configuration the platform. Figure 4 shows the correct configuration of the system when receiving US parameters.

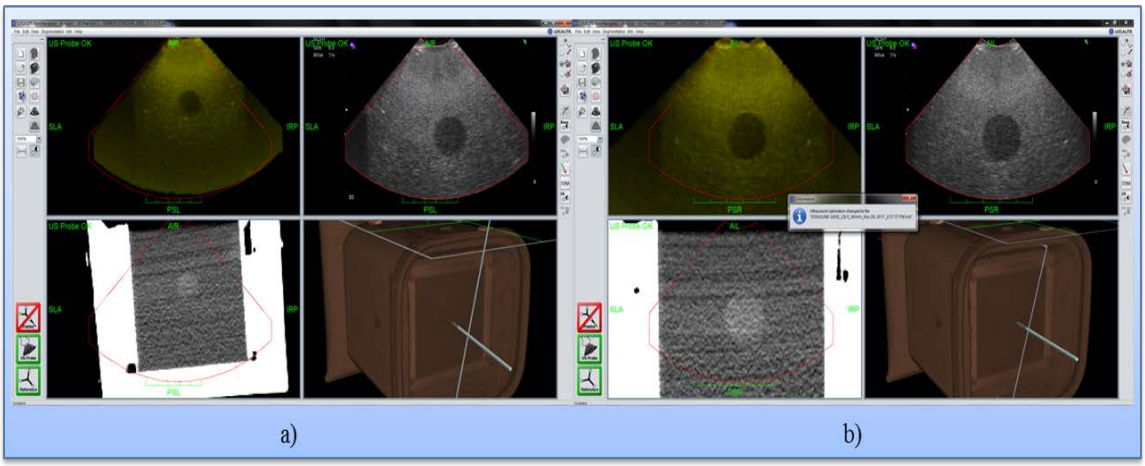


Figure 4*: Successful receiving of the US parameters through the research platform.*

3.2 Generation of a patient model based on preoperative MR data

Figure 5 shows the visualization of the segmented brain tumor, lateral ventricles and blood vessels as a rendering volume. With this tool it is easy to visualize and analyze possible connections between the represented structures. The tool was included into the demonstrator.

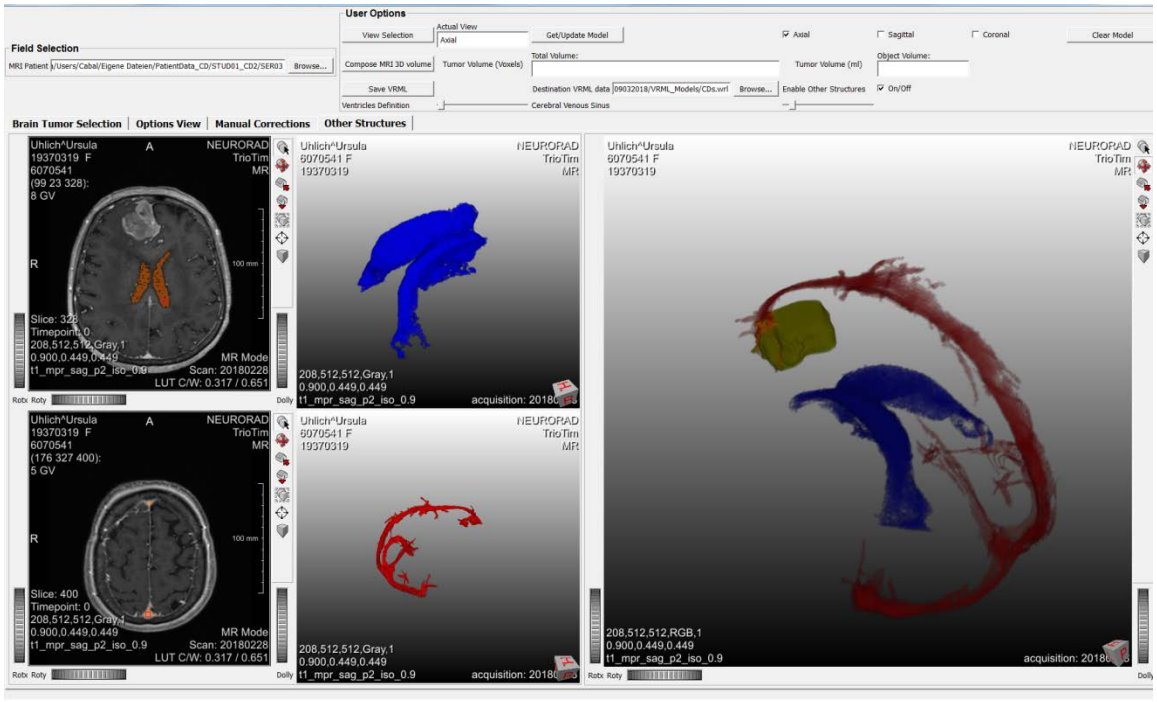


Figure 5: *Visualization and reconstruction Tool, on the top left corner is the MR data where user has to identify, just with a marker on the area of lateral ventricles, and then the rendering volume is shown on the top middle window. Same process on the bottom left corner for the blood vessels, as well the rendering volume is shown on the bottom middle window. On the right window is shown the three volumes to visualize possible connections between structures.*

3.3 Brain tumor contour enhancement in the iUS data

Figure 6 shows the registration process applying rigid method. On the left side MR image is located showing the brain tumor and the contour of it after rigid registration has been applied. On the right side US image is shown. Then it is possible to identify in a better way the contour of the brain tumor after rigid registration has been applied. This tool is not included into the demonstrator yet.

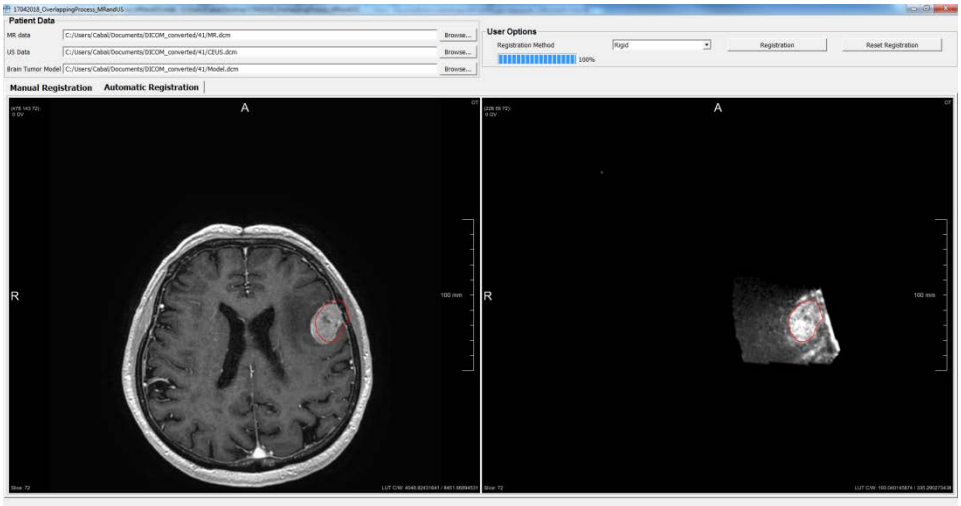


Figure 6*: Brain tumor contour enhancement in iUS tool.*

4 Discussion

The objective of this work was to present an overview of several tools developed to improve a commercial navigation system. First with the automation of the selection of calibration data for purposes of visualization and 3D reconstruction of iUS volumes. The main advantage of this system in comparison with products available on the market is its ability to work with any US devices. This solution is more convenient for the surgeon who can use different US device models. However, the main drawback is the creation of a database for each different US device. Second, a semi-automatic segmentation tool was developed. This interface provides to neurosurgeons the possibility to manually delineate the contour of the brain tumor, the ventricles and the cerebral blood vessels. The segmentation is performed on the research platform. When the task is finished, the contours of the structures are automatically visualized in the navigation system thanks to the socket connection. The tool was originally developed for the processing of MR data. It could be used for the extraction of relevant structures in the iUS data in the OR. A technician or a nurse could perform the task since it requires manual interaction. Third, an automatic method to enhance the brain tumor margin in the iUS image was presented. This tool should be soon included in the demonstrator.

5 Conclusion

Three different tools were developed to improve a commercial navigation system. They aim at automating or facilitating some processes, like the selection of calibration data and the extraction of target and risk structures. Two of these tools were already included into a demonstrator and were successfully tested. Next steps will be the integration of the last tool and the evaluation of the overall system in the OR. This project should increase the safety and the facility of the use of US imaging and therefore support this imaging modality in neurosurgery.

Acknowledgment

Authors thank the Mexican Council on Science and Technology (CONACyT) (Grant Number 302481) and Federal Ministry for Economic Affairs and Energy of Germany (Grant number ZF4148705 AW6) for the provided financial support that allows this research. Additionally, we would like to thank the department of neurosurgery, University Hospital Leipzig, for the clinical study and data collection in the context of an earlier research project funded by the German Research Society (Deutsche Forschungsgemeinschaft).

References

[1] Selbekk T, Jakola AS, Solheim O. Ultrasound imaging in neurosurgery: approaches to minimize surgically induced image artefacts for improved resection control. Acta Neurochir; 155: 973-980; 2013

[2] Unsgaard G, Rygh O, Selbekk T. Intra-operative 3D ultrasound in neurosurgery. Acta Neurochir; 148: 235-253; 2006

[3] Coupé P, Manjón JV, Fonov V, Pruessner J, Robles M, Collins DL. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. NeuroImage, 54(2): 940-954; 2011

[4] Angelini ED, Clatz O, Mandonnet E, Konukoglu E, Capelle L, Duffau H. Glioma dynamics and computational models: A review of segmentation registration and in silico growth algorithms and their clinical applications, Curr. Med. Imag. Rev., 3: 262-276; 2007

[5] Walter AH, Haiying L, Alastair JM, Christopher HP, Robert EM, Charles LT; Safety, Efficacy, and Functionality of High-field Strength Interventional Magnetic Resonance Imaging for Neurosurgery, Neurosurgery,46: 632–642; 2000

Hardware-Setup für die Erfassung von optischen und thermografischen Bilddaten im Rahmen von neurochirurgischen Eingriffen

J. Müller[1], M. Oelschlägel[1], C. Schnabel[1], N. Hoffmann[2], S.B. Sobottka[3], M. Kirsch[3], G. Schackert[3], U. Morgenstern[4], E. Koch[1] und G. Steiner[1]

[1]Technische Universität Dresden, Medizinische Fakultät Carl Gustav Carus, Klinik für Anästhesiologie und Intensivtherapie, Klinisches Sensoring und Monitoring, Dresden, Deutschland
[2]Technische Universität Dresden, Fakultät Informatik, Professur für Computergraphik und Visualisierung, Dresden, Deutschland
[3]Universitätsklinikum Carl Gustav Carus, Klinik und Poliklinik für Neurochirurgie, Dresden, Deutschland
[4]Technische Universität Dresden, Fakultät Elektrotechnik und Informationstechnik, Institut für Biomedizinische Technik, Dresden, Deutschland

Kontakt: Juliane.Mueller3@tu-dresden.de

Abstract

Im Rahmen von neurochirurgischen Interventionen existiert eine Reihe von klinisch relevanten Fragestellungen, die sich mit Hilfe von modernen bildgebenden Verfahren beantworten lassen. Zu diesen Verfahren zählen u. a. die Intraoperative Optische Bildgebung sowie die thermografische Bildgebung. Um das Outcome für den Patienten zu verbessern, bietet es sich an, beide Verfahren entsprechend ihrer Stärken kombiniert anzuwenden. Für diesen Einsatzzweck wurde, basierend auf einem kommerziell verfügbaren Operationsmikroskop, ein modulares Gerätesystem entwickelt, das die intraoperative Erfassung sowie Auswertung von Bilddaten beider Modalitäten ermöglicht. Der Fokus bei der Entwicklung lag auf einer bestmöglichen Integration in die operativen Abläufe und der Anpassung an vorhandene Gerätetechnik. Zur Evaluation wurden anschließend Messungen an Patienten synchron mit beiden Verfahren im Rahmen von Tumorresektionen und vaskulären Eingriffen durchgeführt. Die Resultate zeigen, dass sich das System einfach in den operativen Ablauf integrieren lässt und durch die Anwendung beider Modalitäten die Sicherheit der Ergebnisse erhöht sowie der Zeitaufwand reduziert werden kann.

Keywords: Thermografie, Optical Imaging, Neurochirurgie, Gerätesystem

1 Problemstellung

Trotz der immensen technischen Fortschritte in den letzten Jahren, zählen neurochirurgische Eingriffe am offenen Hirn auf Grund der Irreversibilität von entstandenen Schädigungen und der speziellen Rolle des Organs für den menschlichen Organismus immer noch zu den herausforderndsten Interventionen. In der Vergangenheit konnte gezeigt werden, dass sowohl die Intraoperative Optische Bildgebung (IOI), als auch die thermografische Bildgebung im Rahmen von spezifischen operativen Fragestellungen hilfreiche Informationen für den Chirurgen zur Verfügung stellen können. Zu diesen Informationen zählen u. a. die flächenhafte Lokalisation von funktionellen Arealen, die Abgrenzung von pathologisch veränderten Arealen (z. B. Tumoren) sowie Informationen zur zerebralen Perfusion (z. B. Blutfluss in Aneurysmen). Jede der genannten Modalitäten hat bezüglich bestimmter Anwendungsgebiete Stärken und Schwächen. Die IOI kann primär für die Darstellung von Änderungen im zerebralen Blutvolumen bzw. der zerebralen Oxygenierung angewendet werden und eignet sich somit für die Lokalisation funktionell aktiver Hirnareale [1], da auf Grund der neurovaskulären Kopplung neuronale Aktivität auch mit Änderungen im Metabolismus einhergeht. Im Rahmen von vaskulären Fragestellungen eignet sich insbesondere die Thermografie, da minderdurchblutete Areale sich durch eine geringere Temperatur von physiologisch perfundierten Arealen unterscheiden [2,3]. Die Abgrenzung von pathologisch veränderten Hirnarealen gegenüber gesundem Gewebe ist Forschungsgegenstand beider Methoden [4,5]. Durch die Kombination der Verfahren und der daraus gewonnen Erkenntnisse, kann auf diesem Gebiet u. U. der intraoperative Informationsgewinn gesteigert und die Genauigkeit der Ergebnisse erhöht werden. Zur kombinierten Erfassung sowohl von thermografischen als auch optischen Bilddaten wurde daher im Rahmen

dieser Arbeit ein Gerätesystem entwickelt, das speziell auf die Anwendung im neurochirurgischen Umfeld angepasst ist und einfach in bestehende Arbeitsabläufe integriert werden kann.

2 Material und Methoden

Zur Durchführung von IOI Messungen im sichtbaren Spektralbereich (VIS-Bereich, ca. 400…800 nm) befindet sich am entwickelten Gerätesystem (siehe Abbildung 1) eine CCD-Kamera (AxioCam MRm, Carl Zeiss MicroImaging, Jena, Deutschland), die über einen Strahlteiler (2 Ports, 50/50, Carl Zeiss Meditec AG, Oberkochen, Deutschland) in den optischen Weg des Mikroskops eingekoppelt wird. In den Strahlengang der Kamera kann zusätzlich ein optischer Filter (z. B. Bandpassinterferenzfilter) eingebracht werden, um den Fokus bei den Messungen auf die Auswertung unterschiedlicher physiologischer Phänomene, entsprechend des Absorptionspektrums des signalgebenden Hämoglobins (z. B. Blutvolumenänderung bei 568 nm oder Oxygenierungsänderung bei 600 nm) zu legen. Der zweite Port des Strahlteilers wird für die Aufnahme von Weißlichtbildern mit einer zusätzlichen VIS-Kamera (hier Basler acA1920 mit CMOS-Sensor, Basler AG, Ahrensburg, Deutschland) genutzt. Diese Daten werden zusammen mit den thermografischen Bildern des Operationssitus dargestellt und unterstützen die Orientierung des Chirurgen bei der korrekten Positionierung der Optik über dem exponierten Kortex des Patienten.

Für die thermografische Bildgebung wird ein ungekühltes Thermografie-System (VarioCAM HD head, InfraTec GmbH, Dresden, Deutschland) auf Basis von Mikrobolometerdetektoren verwendet. Es erfasst Infrarotstrahlung im Spektralbereich von 7,5…14 µm. Damit können Temperaturänderungen von bis zu 30 mK mit einer räumlichen Auflösung von 125 µm (bei 30 cm Objektabstand) und bis zu 60 Bildern pro Sekunde aufgenommen werden. Das Thermografie-System ist mittels einer eigens konstruierten Halterung mit Schwalbenschwanz-verbindung am Strahlteiler befestigt.

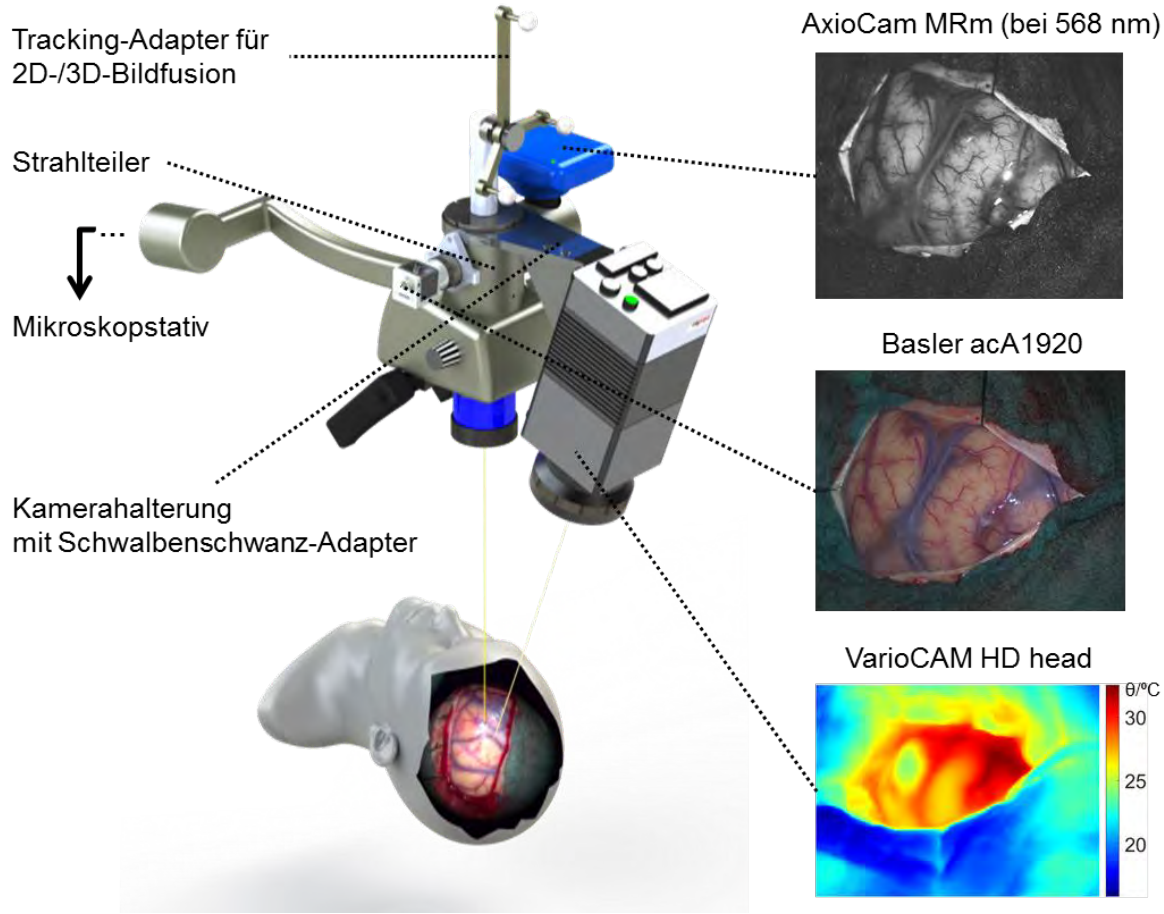


Abbildung 1*: Gerätesystem zur Datenerfassung von optischer und thermografischer Bildgebung bestehend aus einer CCD-Kamera (AxioCam MRm), einem ungekühlten Thermografie-System (VarioCAM HD head), einer RGB-Kamera (Basler acA1920) sowie einem Tracking-Adapter zur Verfolgung der Systemposition während der Intervention durch die Neuronavigation.*

Bildgeführte Interventionen in der Neurochirurgie basieren auf Neuronavigationssystemen, welche die Ortung (Tracking) von Position und Orientierung des Gerätesystems bezüglich des Patienten ermöglichen. Zur Fusionierung der aufgezeichneten thermografischen und optischen Bilddaten mit den für die Neuronavigation präoperativ aufgezeichneten Bilddatensätzen, gehört daher eine Vorrichtung zur Befestigung eines Instrumententracking-Adapters (Ultraschallreferenzeinheit, BrainLab AG, München, Deutschland) zum Gerätesystem. Das im Operationssaal verwendete Navigationssystem VectorVision 2 (BrainLab AG, München, Deutschland) erlaubt die kontinuierliche Verfolgung der Position des Instrumententracking-Adapters im Welt- und Patientenkoordinatensystem. Das Bildkoordinatensystem der Kameras kann wie in [6] beschrieben mit dem Weltkoordinatensystem registriert werden. Die Projektion der registrierten Bilddaten auf die Kortexoberfläche wird mit einem projektiven Textur-Mapping durchgeführt. Die Grauwerte der MRT-Aufnahme werden mit den 2D-Bilddaten durch gewichtete Mittelung fusioniert (vgl. Abbildungen 2C und 3C). Alle Ergebnisse können mittels der Visualisierungssoftware 3DSlicer dargestellt werden.

Die Einzelkomponenten des Systems sind an einem kommerziell erhältlichen Operationsmikroskop (OPMI Pico, Carl Zeiss Meditec AG, Oberkochen, Deutschland) befestigt. Die intraoperative Analyse und Auswertung der Daten erfolgt für jede Modalität auf einem separaten Laptop mit den eigens entwickelten Programmpaketen für MATLAB (The MathWorks, Massachusetts, Vereinigte Staaten) und LabVIEW (National Instruments, Texas, Vereinigte Staaten). Eine genaue Beschreibung der Methodik zur Stimulation funktioneller Areale sowie der dazugehörigen Datenaufnahme und -auswertung findet sich u. a. für die funktionell optische Bildgebung in [1] und für die funktionell thermografische in [7]. Weitere Informationen zur thermografischen Perfusions-bildgebung werden in [2] und [3] beschrieben.

3 Ergebnisse

Das beschriebene Gerätesystem wurde erfolgreich bei neurochirurgischen Interventionen an Patienten mit unterschiedlichen Pathologien eingesetzt. Primär wurden Eingriffe zum Zweck der Tumorresektion untersucht, bei denen die Läsion zusätzlich in der Nähe von funktionell testbaren Arealen lag. Weiterhin wurde das Setup bei vaskulären Eingriffen, wie dem Clipping von Aneurysmen, der Resektion von arteriovenösen Malformationen oder der Darstellung zerebraler Ischämien genutzt. Bei diesen Fällen wurde vor allem die thermografische Bildgebung eingesetzt, die über die generierten Perfusionskarten eine Evaluation des regionalen Blutflusses ermöglichte.

Exemplarisch sind in Abbildung 2 die Ergebnisse der optischen und thermografischen Bildgebung eines Patienten mit einer Metastase (Primärtumor: Melanom) in der postzentralen Kortexregion dargestellt. Mit Hilfe der IOI wurde der sensorische Kortex intraoperativ identifiziert und das Ergebnis mit Hilfe der Messung der Phasenumkehr des elektrophysiologischen Monitorings erfolgreich verifiziert. Auf diese Weise konnten funktionelle Areale während der Operation zur Vermeidung postoperativer funktioneller Defizite geschont werden. Die subkortikal liegende Läsion grenzte sich im thermografischen Bild durch einen Temperaturgradienten vom umliegenden gesunden Hirngewebe ab. Mittels der Fusion von präoperativen MRT-Daten und intraoperativen IOI-Auswertungen konnten die Ergebnisse anatomisch verifiziert und visualisiert werden.

Abbildung 3 zeigt das Beispiel eines Patienten mit einem niedriggradigen Astrozytom (WHO°II, hirneigener Tumor) im frontalen Kortex, der ebenfalls in der thermografischen Bildgebung durch eine hypertherme Region abgegrenzt werden konnte. Zusätzlich zur hier ebenfalls dargestellten Bildfusion von prä- und intraoperativ aufgenommenen Bilddaten, wurde bei diesem Patienten eine zerebrale Perfusionskarte erstellt, die die Evaluation des regionalen kortikalen Blutflusses möglich machte. Dargestellt wird in der Perfusionskarte der Zeitpunkt, zu dem ein intravenös injizierter Bolus von kalter physiologischer Natriumchloridlösung die jeweilige Region erreicht hat.

Bei der Integration des Gerätesystems in die operativen Abläufe ergaben sich für die Durchführung der Messungen keine Schwierigkeiten. Da das Hardware-Setup auf kommerziell erhältlichen Komponenten aufbaut, konnten die für das Mikroskop üblicherweise vorgesehenen Bezüge ebenfalls für die Verwendung im sterilen Operationsbereich genutzt werden. Die Ausrichtung des Bildaufnahmesystems über dem Kortex des Patienten erfolgte nach der Eröffnung des Schädels durch den Operateur, der zur exakten Positionierung die Bildausschnitte der Kameras in einer Live-Darstellung betrachten konnte. Diese wurde auf dem jeweiligen Messlaptop angezeigt, der sich außerhalb des sterilen Bereiches im Operationssaal befand. Für Messungen mit der IOI wurde nach der Positionierung die Raumbeleuchtung ausgeschaltet und der Kortex lediglich durch die im System integrierte Beleuchtung illuminiert (180 W, Xenon). Die thermografische Bildgebung besitzt prinzipbedingt keine speziellen Anforderungen an die Beleuchtung. Nach Durchführung der Messungen erfolgte ein Wechsel zurück auf das übliche Operationsmikroskop, sofern für den Eingriff eines verwendet wurde.

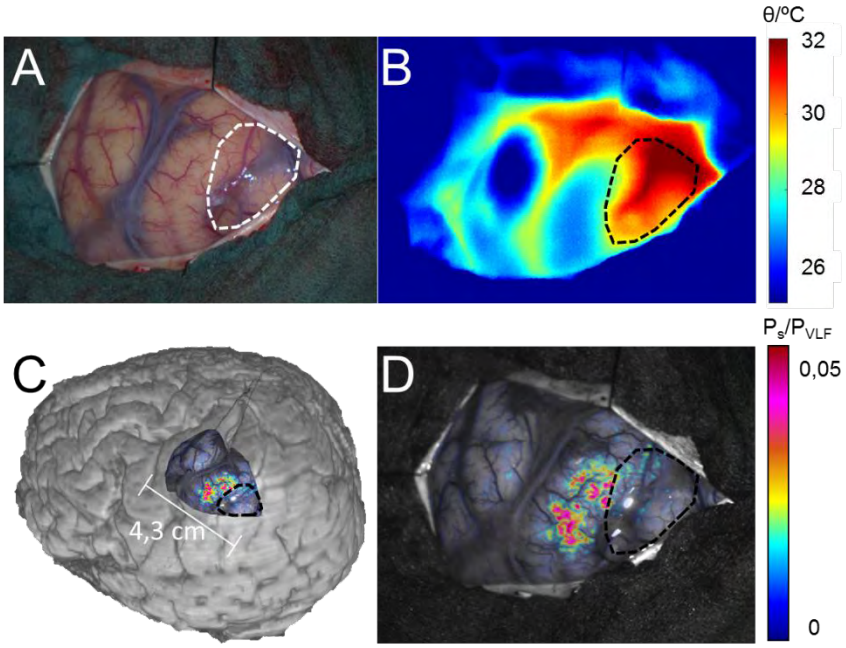


Abbildung 2*: Ergebnisse der optischen und thermografischen Bildgebung eines Patienten mit Metastase im postzentralen Kortex.* ***A*** *RGB-Bild des Operationssitus* ***B*** *thermografische Aufnahme korrespondierend zu A* ***C*** *2D-/3D-Fusion der präoperativen MRT-Aufnahme und der IOI-Aktivierungskarte mit 3DSlicer nach [6]* ***D*** *IOI-Aktivierungskarte nach [1]. Die gestrichelte weiße bzw. schwarze Linie repräsentiert die Grenzen der subkortikal liegenden Metastase.*

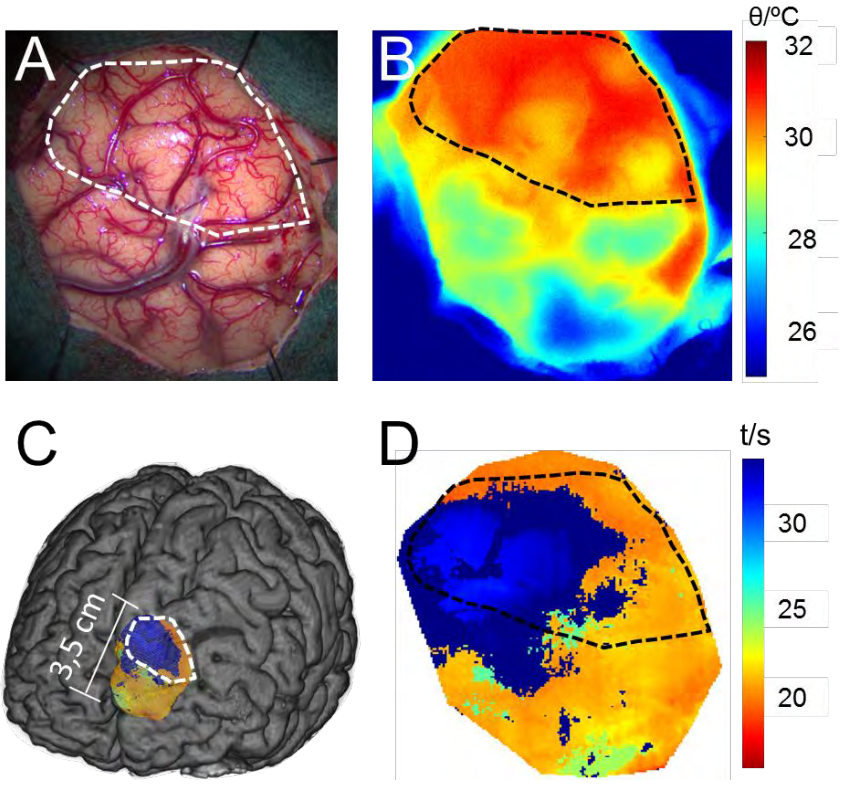


Abbildung 3*: Ergebnisse der thermografischen Bildgebung eines Patienten mit niedriggradigem hirneigenen Tumor im frontalen Kortex.* ***A*** *RGB-Bild des Operationssitus* ***B*** *thermografische Aufnahme korrespondierend zu A* ***C*** *2D-/3D-Fusion der präoperativen MRT-Aufnahme und des thermografischen Bildes mit 3DSlicer nach [6]* ***D*** *zerebrale Perfusionskarte der Kortexoberfläche nach [3]. Die gestrichelte weiße bzw. schwarze Linie repräsentiert die Grenzen des Tumors.*

4 Diskussion

Während der Evaluation, der mit dem entwickelten Gerätesystem durchgeführten Patientenmessungen, wurde festgestellt, dass sich die Kombination beider Bildgebungsmodalitäten besonders für Eingriffe zur Resektion von pathologisch veränderten Geweberegionen (sowohl hirneigene Tumore als auch Metastasen) eignet und gerade hier die gemeinsame Anwendung beider Verfahren einen großen Vorteil gegenüber der jeweiligen Einzelanwendung hat. Die parallele Messung mit beiden Modalitäten bei gleichen Aufnahmebedingungen verbessert die Vergleichbarkeit der Aufnahmen und erhöht die Sicherheit der Ergebnisse, sowohl in Bezug auf die Gewebedifferenzierung als auch auf die Identifizierung funktioneller Areale. Zudem reduziert sich der Zeitaufwand für die intraoperative Messung, da kein Wechsel des Aufnahmesystems notwendig ist. Für die Darstellung von funktionellen Arealen kann die Auswertung der optischen Bilddaten intraoperativ durchgeführt werden, die Berechnungsdauer auf einem mittelmäßig ausgestatteten Laptop beläuft sich auf ca. 2 Minuten, sodass sich inklusive der Aufnahmezeit von 9 Minuten eine Gesamtdauer für die Bildgebung von 11 Minuten ergibt. Das thermografische Rohbild zur Abgrenzung von hypo- bzw. hyperthermen Regionen steht nahezu instantan intraoperativ zur Verfügung, wohingegen die Erstellung der zerebralen Perfusionskarten derzeit auf Grund einer erhöhten Rechendauer postoperativ erfolgt. Perspektivisch wird die Perfusionsanalyse durch die Implementierung des Algorithmus auf einem Grafikprozessor ebenfalls intraoperativ zur Verfügung stehen. Ein weiterer Vorteil ist, dass zu jedem Zeitpunkt die IOI-Ergebnisse durch vaskuläre Analysen der Thermografie ergänzt werden können. Das Gesamtsystem lässt sich einfach in das neurochirurgische Umfeld integrieren und erfüllt die Anforderungen für die Integration in die operativen Abläufe.

In der bisherigen Ausführung muss das Gerätesystem ohne Beobachtertubus verwendet werden. Die Ausrichtung der Kameras erfolgt ausschließlich über den jeweiligen Laptop-Bildschirm, dessen für den Chirurgen gut sichtbare Positionierung mitunter schwierig sein kann, abhängig von der Größe des Operationssaals und der Anzahl zusätzlich vorhandener Geräte. Dies lässt sich jedoch durch die zukünftige Anbindung an das bereits im Operationssaal vorhandene Videosystem lösen. Zudem erfolgt die Kamerasteuerung und Datenerfassung auf getrennten Messlaptops, was momentan einen zusätzlichen Synchronisationsschritt im Messablauf erforderlich macht. Eine Softwareintegration der IOI und thermografischen Bildgebung wird daher zukünftig angestrebt. Bei einer Weiterentwicklung des Gerätesystems empfiehlt es sich zudem, zwei baugleiche Kameras im visuellen Spektralbereich zu verwenden. Die synchrone Erfassung des Operationssitus macht eine spätere 3D-Rekonstruktion der Hirnoberfläche möglich, die für stereoskopische oder augmentierte Visualisierungen genutzt werden kann.

5 Zusammenfassung

Das entwickelte Gerätesystem bietet ein effizientes Framework für die kombinierte Erfassung von thermografischen und optischen Bilddaten inklusive der Möglichkeit, die intraoperativ erfassten Daten mit präoperativen Patientendaten zu fusionieren und visualisieren. Die Kombination ermöglicht intraoperativ die Einbeziehung der Informationen beider Bildgebungsverfahren und unterstützt den Chirurgen optimal entsprechend der jeweiligen klinischen Fragestellung. Auf diese Art kann zukünftig u. U. das Risiko für postoperative funktionelle Defizite vermindert, sowie die Überlebensdauer der Patienten nach einer Tumorresektion erhöht werden.

Danksagung

Das Projekt wurde vom Europäischen Sozialfond (ESF) und dem Freistaat Sachsen sowie dem Sächsischen Staatministerium für Wissenschaft und Kunst im Rahmen der ESF-Nachwuchsforschergruppe "Optische Technologien in der Medizin" (Projektnr. 100270108) und dem ESF-Projekt „NeuroFusion“ (Projektnr. 100312752) gefördert.

Referenzen

[1] Sobottka S. B., Meyer T., Kirsch M., Koch E., Steinmeier R., Morgenstern U., Schackert G., *Intraoperative optical imaging of intrinsic signals: a reliable method for visualizing stimulated functional brain areas during surgery,* Journal of neurosurgery, 119(4), 853-863, (2013)

[2] Steiner G., Sobottka S. B., Koch E., Schackert G., Kirsch M., *Intraoperative imaging of cortical cerebral perfusion by timeresolved thermography and multivariate data analysis*, JBO, 16(1):016001-1-016001-6, (2011)

[3] Hollmach J., Radev Y., Hoffmann N., Schnabel C., Sobottka S. B., Kirsch M., Schackert G., Koch E., Steiner G., *Intraoperative perfusion imaging of the cerebral cortex by time-resolved thermography*, Biomedical Engineering / Biomedizinische Technik, 59(1), (2014)

[4] Oelschlägel M., Meyer T., Schackert G., Kirsch M., Sobottka S. B., Morgenstern U., *Intraoperative optical imaging of metabolic changes after direct cortical stimulation–a clinical tool for guidance during tumor resection?*, Biomedical Engineering/Biomedizinische Technik, (2018)

[5] Gorbach A.M., Heiss J.D., Kopylev L., Oldfield E.H., *Intraoperative infrared imaging of brain tumors*, Journal of Neurosurgery, 101:960-969, (2004)

[6] Hoffmann N., Weidner F., Urban P., Meyer T., Schnabel C., Radev Y., Schackert G., Petersohn U., Koch E., Gumhold S., Steiner G., Kirsch M., *Framework for 2D-3D image fusion of infrared thermography with preoperative MRI*, Biomedical Engineering / Biomedizinische Technik, 62(6), 599–607, (2017)

[7] Hoffmann N., Petersohn U., Schackert G., Koch E., Gumhold S., Kirsch M., *Fast mapping of the eloquent cortex by learning L2 penalties*, Proceedings of the 21th International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI), Granada, Spain, (2018), Accepted for publication

Interventionelle MR-geführte Herzbiopsie: Entwicklung einer innovativen MR-tauglichen Biopsiezange

C. Martin Reich[1], Annekatrin Pfahl[1], Moritz Lehnhardt[1], Bernhard Uihlein[2], Sandra Petershans[2], Christina Unterberg-Buchwald[3], Martin Uecker[4], Andreas Melzer[1]

[1]*Innovation Center Computer Assisted Surgery, Universität Leipzig, Leipzig, Deutschland*
[2]*EPflex Feinwerktechnik GmbH, Dettingen an der Erms, Deutschland*
[3]*Universitätsmedizin Göttingen, Klinik für Kardiologie und Pneumologie, Göttingen, Deutschland*
[4]*Institut für Diagnostische und Interventionelle Radiologie, Universitätsmedizin Göttingen, Göttingen, Deutschland*

Kontakt: ChristophMartin.Reich@medizin.uni-leipzig.de

Abstract

Die Bedeutung der Magnetresonanztomographie (MRT) nahm im Bereich minimal-invasiver Interventionen in den vergangenen Jahren stark zu. Die Vorteile einer MR-geführten Intervention sind ein hoher Weichgewebekontrast, der Verzicht auf nierenschädigende Kontrastmittel, das Nichtvorhandensein ionisierender Strahlung sowie eine hohe räumliche Auflösung. Dennoch dienen für Myokardbiopsien die Röntgen-Fluoroskopie und Echokardiografie als Standardverfahren der bildgestützten Echtzeitüberwachung des Eingriffs. Ursache dafür ist nicht zuletzt die begrenzte Verfügbarkeit MR-geeigneter Instrumente. Im Fokus der Studie stehen daher die Entwicklung und Evaluation einer MR-sicheren Biopsiezange und die Etablierung eines angepassten klinischen Workflows. Intensive ex-vivo-Experimente in einem realistischen Gefäßmodell werden dabei konzipiert und dienen als Grundlage der Analysen.

Keywords: interventionelle MRT, Herzbiopsie, flexible MR-taugliche Biopsiezange, MRT-Biopsie

1 Problemstellung

Nach den Bestimmungen des ASTM werden Materialen den Klassen *MR-sicher*, *MR-kompatibel* oder *MR-unsicher* zugeordnet. Maßgeblich sind dabei die Eigenschaften dieser Materialien in der Umgebung starker Magnetfelder. Herkömmliche Instrumente bestehen zumeist aus biokompatiblen Metallen. Aufgrund deren Erhitzung durch Radiowellen [1] sind diese für den Einsatz innerhalb MR-geführter Interventionen nur bedingt geeignet. Für den klinischen Gebrauch existieren lediglich wenige *MR-kompatible* Instrumente wie Biopsie-Punktionskanülen (*Somatex*, *Invivo*). Neben materialbedingter Sicherheitsvorschriften spielt weiterhin die Visualisierung der Instrumente in der MRT-Bildgebung eine entscheidende Rolle. Diese lässt sich prinzipiell mittels passiver, semi-aktiver oder aktiver Techniken verbessern [2]. Vorteil aktiver Methoden ist vor allem die genaue räumliche Darstellung - allerdings benötigen diese Verfahren spezielle Setups. Passive und semi-aktive Methoden hingegen können in allen verfügbaren MRT-Systemen zum Einsatz kommen und stellen somit eine flexiblere Alternative dar. Mithilfe integrierter passiver Marker konnte der weltweit erste *CE-zertifizierte MR-kompatible* Führungsdraht (*EPflex*) auf den Markt gebracht werden. Ziel des Verbundprojektes *MRT-Biopsie* ist daher die Entwicklung und Evaluierung einer flexiblen *MR-sicheren* Zange für die Anwendung innerhalb MR-geführter Herzbiopsien.

2 Material und Methoden

Zunächst wurde ein Design des Zangenkopfes entwickelt, das eine Produktion aus nicht-magnetischen, nicht-metallischen und nicht-leitfähigen Werkstoffen ermöglicht. Die Auswahl der Materialien erfolgte in Hinblick auf die Zielgabe, der Einstufung *MR-sicher* zu genügen. Da Polymere und Keramiken im Allgemeinen geringe Artefakte in der MRT erzeugen, muss deren Sichtbarkeit erhöht werden. Die Einzelteile werden dazu mit passiven und semi-aktiven Markern versehen und in verschiedenen MRT-Systemen (1,5 und 3 Tesla, *GE, Siemens, Philips*) mit Standard- und Echtzeit-Sequenzen getestet. In Folge werden diejenigen Sequenzen identifiziert und nötigenfalls angepasst, die die geeignetste Visualisierung für den Anwendungsfall erzeugen. Für

weitere ausführliche Studien in einem möglichst realistischen Setup werden ex-vivo-Experimente durchgeführt. Hierzu wurden aus anonymen medizinischen Bilddaten Gefäße (Arterien/Venen) segmentiert und in ein 3D-Modell überführt, welches wiederum als Vorlage für einen 3D-Druck dient (siehe Abb. 1). Um dabei natürliches Gewebe bioptieren zu können, werden jeweils frische Schweineherzen an das Gefäßmodell angeschlossen. Mit diesem Phantom und dem Zangensystem ist es anschließend möglich, realitätsnahe Evaluationen im MRT durchzuführen. Nach erfolgreichen ex-vivo Tests sollen später in-vivo Erprobungen folgen.

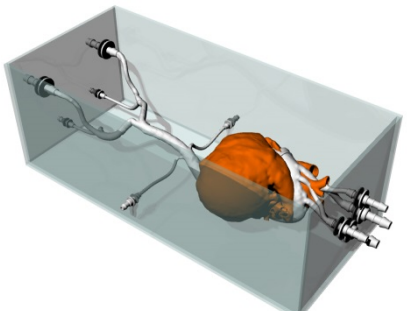


Abbildung 1:3D-Modell des Aorten-Phantoms (Phacon).

3 Ergebnisse

Erste Muster des Zangenkopfes aus nicht-ferromagnetischen Verbundwerkstoffen wurden bereits hergestellt. Besonderes Augenmerk lag dabei auf der Schärfbarkeit des Materials, einem möglichst geringen Reibungskoeffizienten und der Korrosionsbeständigkeit. Die Geometrie der Zange wurde an eine Herzbiopsie mit formalerem Zugang mit Führungskatheter angepasst. So wurde im Vergleich zu herkömmlichen metallischen Biopsiezangen der Zangenkopf um etwa 50% verkürzt und der Öffnungswinkel um ca. 50% vergrößert.

4 Diskussion

Nach erfolgreicher Produktion von Musterstücken müssen weiterhin intensive Untersuchungen zur Sichtbarkeit im MRT durchgeführt werden. Gleichzeitig muss die Funktionalität der Zange an tierischem Gewebe getestet werden, um eine möglichst hohe Qualität des Bioptats zu gewährleisten. Zuletzt erfolgt die Formulierung eines adäquaten klinischen Workflows zur MR-geführten Herzbiopsie. Detaillierte Recherche, Beobachtung und Protokollierung von Standardprozeduren und MR-Interventionen werden hierbei analysiert und ausgewertet. Problemfelder sind dabei vor allem die Optimierung der MRT-Sequenzen, die MRT-Steuerung und die Kommunikation zwischen dem Interventionalisten und den Assistenten im Kontrollraum. Zudem muss direkter Zugang zum Patienten für den Notfall, bspw. eine Reanimation, gewährleistet sein.

5 Zusammenfassung

Um eine neuartige, flexible und *MR-sichere* Zange für Myokardbiopsien und einen korrespondierenden klinischen Workflow zu etablieren, wurden intensive Untersuchungen durchgeführt und bereits ein erster Prototyp entwickelt. Weiterführende in-vitro-Experimente und Recherchen werden derzeit durchgeführt. Das Ziel ist der Nachweis der Anwendbarkeit in Form von reproduzierbarer Bioptat-Qualität und reliabler MR-geführter Intervention.

Referenzen

[1] Wolska-Krawczyk M, Rube MA, Immel E, Melzer A, Buecker A. Heating and safety of a new MR-compatible guidewire prototype versus a standard nitinol guidewire. Radiol Phys Technol. 1. Januar 2014;7(1):95–101.

[2] Saikus CE, Lederman RJ. Interventional Cardiovascular Magnetic Resonance Imaging: A New Opportunity for Image-Guided Interventions. JACC: Cardiovascular Imaging. 1. November 2009;2(11):1321–31.

[3] Unterberg-Buchwald C, Ritter CO, Reupke V, Wilke RN, Stadelmann C, Steinmetz M, u. a. Targeted endomyocardial biopsy guided by real-time cardiovascular magnetic resonance. Journal of Cardiovascular Magnetic Resonance [Internet]. Dezember 2017 [zitiert 3. April 2018];19(1). Verfügbar unter: http://jcmr-online.biomedcentral.com/articles/10.1186/s12968-017-0357-3

Optical Coherence Tomography – Towards New Medical Applications

N. König[1], M. Riediger[1], D. Wilhelm[2], M. Witte[1,3], K. Masamune[4], R. Schmitt[1,5]

[1]Fraunhofer Institute for Production Technology IPT, Aachen, Germany
[2]Klinikum rechts der Isar, Technical University of Munich, Munich, Germany
[3]Fraunhofer Center for International Management and Knowledge Economy IMW, Leipzig, Germany
[4]Institute of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo, Japan
[5]Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Aachen, Germany

Contact: niels.koenig@ipt.fraunhofer.de

Abstract

The non-invasive imaging technology OCT (Optical Coherence Tomography) is based on low-coherence interferometry and can enable a three-dimensional representation of organic tissue with close to microscopy resolution. With penetration depths of several millimeters, resolutions of a few microns, real-time scanning rates, and miniaturization opportunities, imaging is a frequently sought-after clinical field. Clinical experts from various disciplines described concrete needs as part of the "Research Alliance OCT for Medicine" FOMed, which can be optimally addressed through the future use of OCT technology. Improved diagnoses, gentler therapies for the patient or the simultaneous acceleration of the surgical workflow are the expected benefits - the potential advantages and the applications in question show the great potential for using OCT in medicine.

Keywords: OCT, Optical Coherence Tomography, Intraoperative Imaging, Oncology, Minimal Invasive Surgery

1 Problem

Optical Coherence Tomography (OCT) is an imaging technology based on low-coherence interferometry, that can penetrate transparent and semi-transparent materials and make its internal structures visible. This applies in particular for the analysis of human tissue. OCT works non-invasively or without additional invasion. Due to its optical working principles, it causes no radiation or ionization damage. Furthermore, OCT provides high resolutions at the micrometer level and acquires tomographic cross-sectional and volumetric images in real-time. Because of the fiber-optical system concepts, there is a high potential for miniaturization of OCT for use in minimally invasive medicine.

However, the effective use of OCT technology is highly dependent on the optical properties of the particular tissue and requires customized system design as well as carefully selecting the OCT mode and signal post-processing. In this respect, OCT is not a universal technology, which can be used equally for all possible applications, but has to be adopted in an application-specific way. With the functionality mentioned above, OCT is able to characterize tissue morphologies, many other application perspectives arise. An almost universal application of OCT is the detection of tissue malformations and especially cancer [1], since OCT offers a resolution superior to ultrasound imaging with larger penetration depths than confocal microscopy.

Today, OCT has become the gold standard as diagnostic tool for retinal and corneal diseases in ophthalmology [2]. Although a lot of feasibility studies have been reported [3] and more than 4.000 journal papers are published annually, only a few products outside ophthalmology have found their way to the market. Therefore, we have initiated the innovation forum FOMed [4], which serves to open up new medical application perspectives.

2 Material and Methods

The development of new medical applications for OCT requires the concerted interaction of all necessary contributors. For this purpose, we gathered clinicians and companies as well as academic and institutional partners

for the generation of contributions for systemic solutions in an appropriate network structure. The network structure largely covers the value chain for the creation of new medical OCT devices, as shown in Fig. 1.

Figure 1*: Value Chain OCT for New Medical Devices*

An essential principle in the methodical procedure was to collect the clinical requirements as the decisive determinant of future system designs for suitable OCT solutions and to base this on the further technical details. The initial start of the project therefore was a clinical workshop in which representatives of different clinical disciplines highlighted current field specific limitations. On that basis, application examples were selected in which the use of OCT would contribute to a benefit. The necessary technical solutions were coordinated within the framework of theme workshops. The results were presented and discussed in a 2-day final conference on May 16, 2018 in Berlin.

The clinical contributions were supplemented by several clinical workshops at the Tokyo Women Medical University (TWMU) under the direction of Prof. Ken Masamune in January 2018 in Tokyo. These workshops contributed esp. the use perspectives in the fields of Neurosurgery, Cardiology and Interventional Pulmonology.

In summary, we collected possible OCT applications from the following clinical disciplines: dermatology, gastroenterology, visceral surgery, vascular surgery, urology, gynecology, ENT, orthopedics, neurosurgery, cardiology, pulmonology and pathology.

3 Results

The identified clinical applications of OCT were elaborated in a series of workshops according to Fig. 1. In the following we describe the subject-specific requirements using the example of visceral surgery.

According to the clinical experts, imaging such as OCT, which can be used in vivo, will not deliver the image quality as the first choice for histology. Therefore, OCT as quasi-microscopic imaging must not use histology as a benchmark or score on it and not be intended as a replacement method. Instead, the focus must be placed on the investigation of tissue properties that can be detected in vivo and that allow for valid, usable information. Since operatively active physicians (users) are not pathologists, work should not be carried out in the direction of a realistic imaging, but rather a decision support should be provided, which carries out a qualitative assessment. This can be done, for example, in the form of a software-assisted tissue analysis that provides information on the type of tissue abnormality (benign / malignant / inflammatory). The compensation of patient and user movements is a particular challenge which can be addressed by appropriate image processing tools.

It is often said that the OCT basically covers an interesting resolution range, but the scan area is too small for an effective and ergonomic application. The intraoperative OCT should therefore not be regarded as a search modality, but as a confirmation modality. A hybrid imaging with OCT as a micro-imaging combined with an additional modality as pre- and intra-operative macro-imaging for navigation and detection of conspicuous areas was considered useful by all experts. The preferred modality is based on the field of application, for example OCT could be supplemented by imaging with a higher scan range but lower resolution (e.g. US or endoscopy) or with even higher resolution (e.g. OCT / LSM) for maximum efficiency.

Furthermore, image-processing that gives the therapist an automated diagnosis recommendation or highlights conspicuous areas, is addressed across all disciplines. For this purpose, the above mentioned approach of hybrid imaging can be supplemented by automated image analysis and interpretation, or automated provision of diagnostic criteria.

In some clinical settings, images of relatively low depth of penetration can be used to delineate neoplasia and adjust the therapy. Here, the assessment of the basement membrane (in about 250 - 500µm depth) is of primary importance, esp. in the field of gynecology [5] and gastroenterology [6], but also in ENT [7], as the invasiveness of squamous cell carcinomas (e.g. esophagus, trachea, cervix, etc.) can be tested. It represents the boundary between the epithelial layer and the stroma and is localized in body cavities, which can be well visualized endoscopically. The OCT is thus suitable for endoscopic imaging of the tomography and topography of external tissue layers such as organ surfaces or body cavity walls (mucosa, serosa) or for imaging of deeper structures when used in resectoscopes.

3.1 Use Case: OCT in Visceral Surgery

In literature, OCT has been judged as a tool for tissue differentiation in visceral surgery especially for the detection of cancer sufficiently positive. This is also confirmed by the preliminary investigations by Fraunhofer IPT together with the TU Munich [8], see fig. 2.

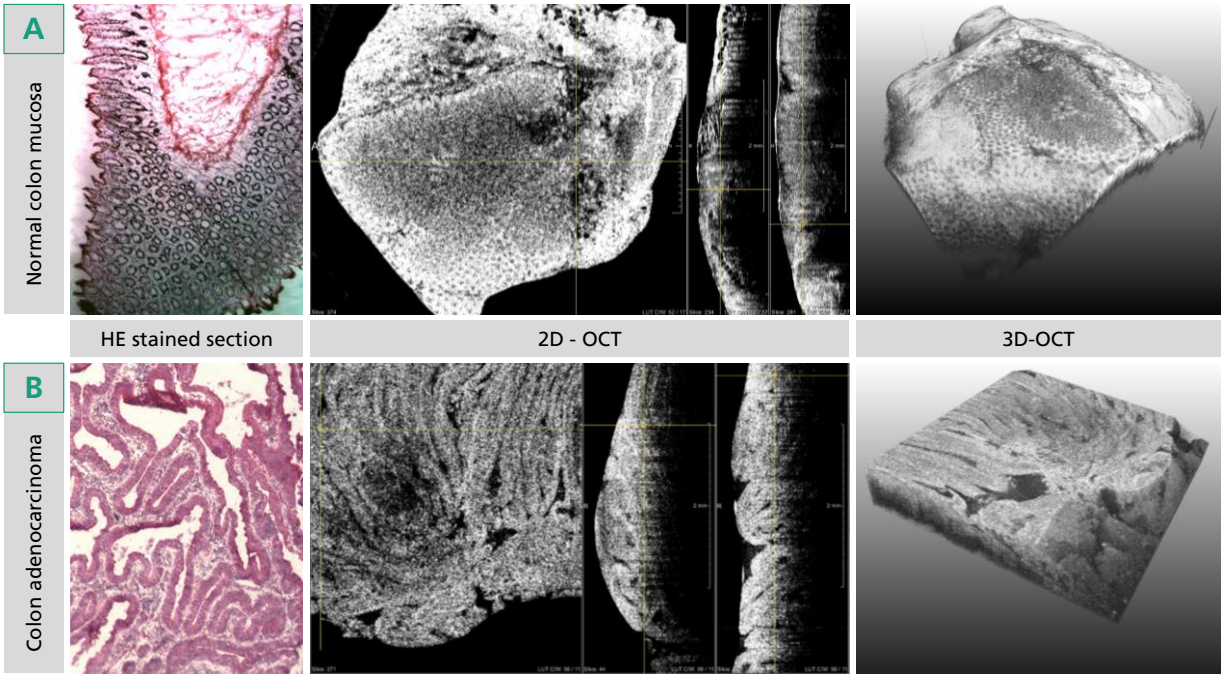


Figure 2: OCT imaging on normal colon mucosa (A) and colon adenocarcinoma (B) with addition of stained sections. The OCT workup included a 3D reconstruction as well as multi-planar visualization including orthogonal oriented planes [8]

As a need for endoscopic use of OCT in Visceral Surgery the radicality of therapies is mentioned: Visceral Surgery often operates with excessive and sometimes unnecessary radicality (excessive removal of surrounding healthy tissue) to ensure the complete removal of the malignant tissue. The quality of the surgery is evaluated histologically, but relies on macroscopic imaging (CT / MRI), so that. there is a continuing conflict between therapeutic decisions to operate too radically and microscopically too close.

The frozen section investigation as a current solution has many disadvantages, such as delaying the operation, labor and cost-intensity, availability and delay regarding the workflow.

When using minimally invasive methods, the tactility is lacking, so that it is not available as sensory information and must be visually replaced. The preoperative imaging is too inaccurate and the transformation to the intraoperative situs is not or only insufficiently possible. Therefore, there is currently the need for "over-radical" therapy. In the long term, partially autonomous approaches involving on-line therapy control are needed, i.e. the system only resects what is really infested. Approaches that are not based on existing concepts and allow for better implementation of necessary therapies are preferred.

All areas of visceral surgery would benefit from OCT-based in vivo imaging. In the combination of OCT with other modalities, clinical assessment is as follows: for visceral surgery, the combination with US is interesting to allow for a rough navigation in parenchymatous organs. E.g. liver lesions could be targeted and then examined micro-optically. The combination is also useful in the pancreas and possibly the lungs. The scan area and depth must cover a larger area. Most visceral tumors are not superficially recognizable, so the probe must penetrate the tissue and must be directed into suspicious regions. Target applications are from the today´s view liver, pancreas and lymph node biopsy for oncological surgery.

3.2 Sensor Concepts

As explained above, a medical OCT system always consists of several components, which have to be selected and developed as explained in the value chain shown in fig. 1. The largest customization towards the clinical application affects the sensor, which we understand as the frontend interacting with the patient. Less affected is the actual OCT evaluation system containing the opto-electronic set-up and the spectrometer. For the clinical disciplines listed in section 2, several applications have been identified. In joint workshops with companies of the

FOMed network, we developed 10 OCT sensor concepts, which can be subdivided in the following groups: (1) flexible endoscopes, (2) rigid endoscopes, (3) OCT-integrated microscopes, (4) handheld probes and (5) autonomous therapeutic systems. In the following, the details of a probe concept suitable for applications in visceral surgery described in section 3.1 are presented.

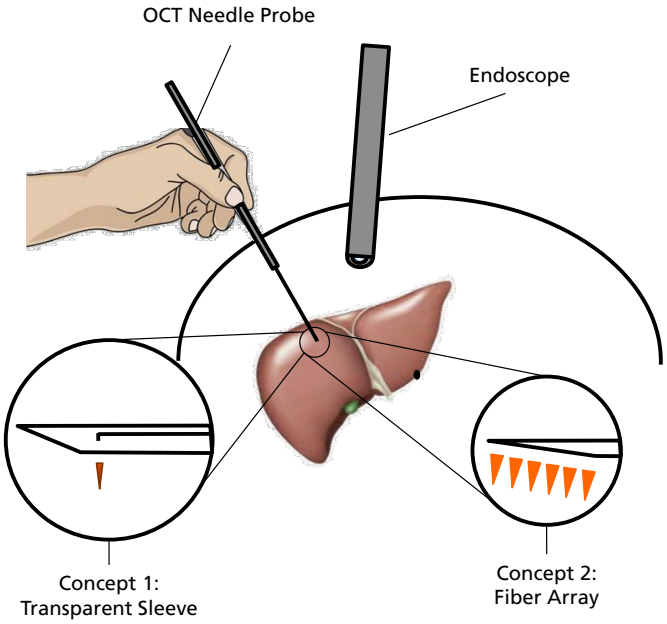


Figure 3: Needle type OCT probe for locating tumorous tissue in visceral surgery, illustrated with the liver as use case scenario. The probe diameter is reduced below 1 mm for minimal trauma. A visual endoscope or ultrasound imaging may be added for guiding purposes. Depending on the diameter requirements, either a single channel (1) or multi-channel (2) concept can be realized.

Fig. 3 shows the concept of a miniaturized needle type OCT probe for the localization of tumorous regions in soft tissues such as the liver, making it attractive for intraoperative application in laparoscopic surgery. In case of concept 1, the result of each measurement would be an A-scan containing the tissue density profile, which can be an indicator for tissue inhomogeneity. Concept 2 would deliver a B-scan through an array of fibers, which are addressed by an external fiber switch or scanning device. The navigation of the probe can be supported by tracking markers and conventional navigation systems. For decreasing the trauma associated with the insertion of the probe, the diameter can be reduced below 1 mm. Diameters down to 0,31 mm have already been shown for concept 1 [9]. A scanning set-up using an array of 100 singlemode fibers according to concept 2 has been reported in [10].

3.3 Results of the Utility Value Analysis

For the clinical disciplines mentioned above, a number of OCT applications have been identified and combined with the five groups of sensor concepts, resulting in 70 potential medical applications of OCT. In order to assess and prioritize these applications with respect to their clinical and economical relevance, we developed the following list of evaluation criteria together with our network partners: (1) utility values for clinical application (specifity and sensitivity, duration of procedure, period of reconvalescence, recurrence rate); (2) utility values for clinical workflow (handling, training requirements, material costs, field of view in relation to area of interest, required human resources); (3) number of cases and cost savings; (4) utility values from industrial perspective (degree of innovation, perceived potential, reimbursement possibilities, certification efforts, market exploitation efforts, competitive pressure from alternatives, feasibility). Utility values in (1)-(3) were evaluated by the clinical experts and (4) by the companies respectively. Each of the criteria listed in brackets have a value range from 1 to 10, resulting in a maximum score of 190. Medical applications of OCT, which reached the Top 4 scorings, were: 1. Handheld OCT for tumor and eczema recognition in dermatology. 2. Rigid endoscope for diagnostics of the epithelium in the upper aero-digestive tract. 3. Rigid endoscope for diagnostics re. tumor and cartilage damages (using PS-OCT) in orthopedics. 4. Needle type OCT probe for tumor diagnostics in visceral surgery (see sec. 3.2).

4 Discussion

Clinical treatment decisions in surgery oncology rely on accurate data about the histological and cellular prerequisites of the suspicious tissue area.

Unfortunately, histological examination and the following tissue-based diagnoses delays the performance of surgery, whether it is an intraoperative frozen section or a routine procedure. Methodological approaches and suitable devices that enable real-time tissue analysis are therefore of great relevance and clinical interest. OCT can also contribute to paradigm change in medical imaging: The future role of OCT as a confirmation tool would lead to new approaches in imaging routines.

According to the findings to date, there is almost no clinical discipline that has not indicated the need for action that can be significantly supported by the use of OCT, see fig. 4.

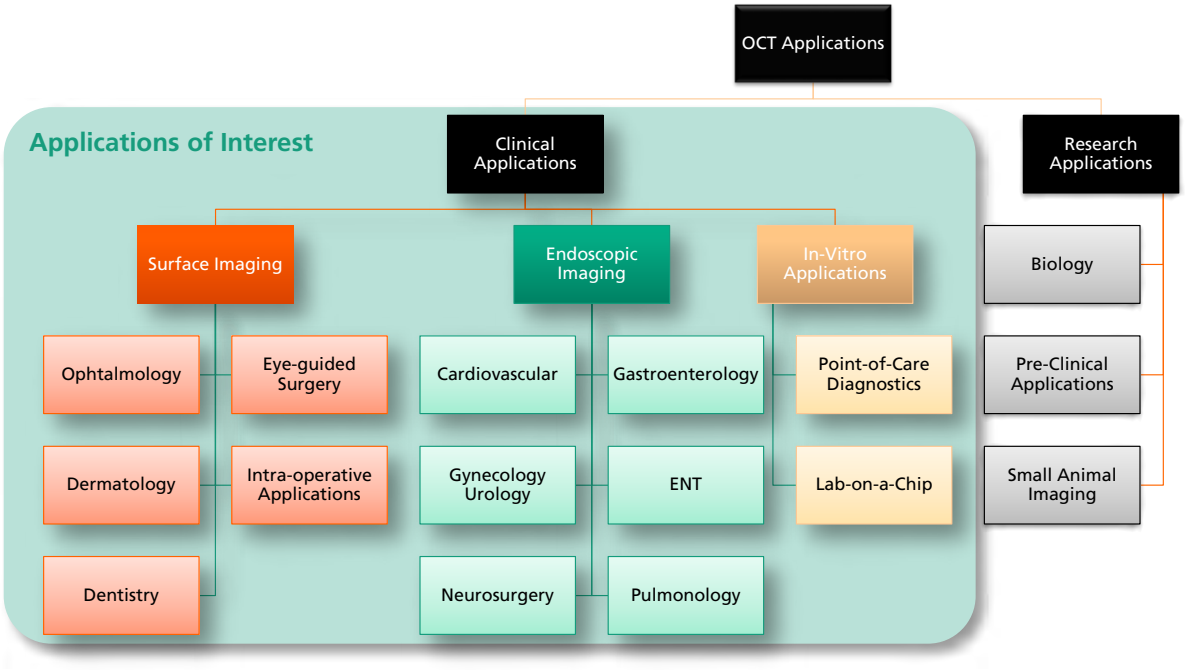


Figure 4*: Overview of OCT Applications*

5 Conclusion

The work carried out with the OCT research alliance FOMed was dedicated to the goal of bringing together suitable network partners for systemic solutions and working on these together with clinically relevant solution concepts. These should be as substantial that implementation projects can be directly defined and initiated.

From today's point of view, the goal has been to merge contributions from different partners to identify suitable system solutions and then to deepen the defined development steps for further development steps as a suitable approach. Within the FOMed alliance the first goal was thoroughly met.

The evolution of the OCT technology towards new devices for clinical applications is at the very beginning. The fact that OCT has not yet established itself as a mass-attractive imaging beyond the application in ophthalmology and in cardiovascular fields of application has, above all, to do with the fact that a single actor alone can not succeed in implementing a new OCT device. Therefore, the continuation of the research alliance OCT is of particular interest.

Over the past year, the FOMed Innovation Forum, an alliance of doctors and industry and research partners, has devoted its attention to the subject and laid the groundwork for further collaboration to promote this optical technology for new medical applications. In a joint initiative with SPECTARIS, the German Industry Association for Optical, Medical and Mechatronic Technologies, the Fraunhofer Institute for Production Technology IPT will establish and expand the innovation network "Innovation Hub FOMed: OCT for new medical applications". The participants in this innovation platform include all actors involved in the field of OCT - from medical applications to research and development of systems and components. In addition to the regular exchange of actors, the development of high-impact consortia and the implementation of R&D projects will be promoted.

Acknowledgement

The results presented within this paper have arisen from two projects, which are kindly supported by the German Ministry of Education and Research (BMBF) and the DLR Project Management Agency: »FOMEd« (FKZ 01MI1709) was carried out in cooperation with Klinikum rechts der Isar of Technical University of Munich and several clinical partners and companies, which are listed on the project homepage www.forschungsallianz-oct.de. »OCTmapp« (FKZ 01DR17028) is carried out in cooperation with Tokyo Women´s Medical University TWMU. The following clinical experts have contributed to the discussion and definition of OCT applications (in alphabetical order): Rafael Beck, University of Leipzig; Christian Betz, University Hospital Tübingen; Rainer Burgkart, Technical University of Munich; David Ellebrecht, University Medical Centre Schleswig-Holstein; Georgios Gakis, University Hospital of Würzburg; Julia Gallwas, Ludwig Maximilians University Munich; Carolin Knebel, Technical University of Munich; Takako Matsumoto, Tokyo Women´s Medical University; Alexander Meining, Ulm University; Yoshihiro Muragaki, Tokyo Women´s Medical University; Arya Nabavi, International Neuroscience Institute Hannover; Björn Rath, Universal Hospital RWTH Aachen; Kenta Uto, Tokyo Women´s Medical University; Dirk Wilhelm, Technical University of Munich; Linn Wölber, University Medical Center Hamburg-Eppendorf; Junichi Yamaguchi, Tokyo Women´s Medical University; Kenji Yasui, Tokyo Women´s Medical University

References

[1] Benjamin J. Vakoc, Dai Fukumura, Rakesh K. Jain, Brett E. Bouma, *Cancer imaging by optical coherence tomography: preclinical progress and clinical potential*, Nat. Rev. Cancer, 12(1) 363–368 (2012)

[2] Johannes F. de Boer, Rainer Leitgeb, Maciej Wojtkowski, *Twenty-five years of optical coherence tomography: the paradigm shift in sensitivity and speed provided by Fourier domain OCT [Invited]*, Biomed. Opt. Express 8, 3248-3280 (2017)

[3] Oscar M. Carrasco-Zevallos, Christian Viehland, Brenton Keller, Mark Draelos, Anthony N. Kuo, Cynthia A. Toth, Joseph A. Izatt, *Review of intraoperative optical coherence tomography: technology and applications [Invited]*, Biomed. Opt. Express 8, 1607-1637 (2017)

[4] Fraunhofer IPT, FOMed Project Homepage: www.forschungsallianz-oct.de

[5] Julia Gallwas, Aydan Jalilova, Roland Ladurner, Theresa Maria Kolben, Thomas Kolben, Nina Ditsch, Christian Homann, Eva Lankenau, Christian Dannecker, *Detection of cervical intraepithelial neoplasia by using optical coherence tomography in combination with microscopy*, J. Biomed. Opt. 22(1) 016013 (2017)

[6] Anne-Fré Swager, Guillermo J. Tearney, Cadman L. Leggett, Martijn G.H. van Oijen, Sybren L. Meijer, Bas L. Weusten, Wouter L. Curvers, Jacques J.G.H.M. Bergman, *Identification of volumetric laser endomicroscopy features predictive for early neoplasia in Barrett's esophagus using high-quality histological correlation*, Gastrointest. Endosc. 85 (5), 918-926.e7 (2017)

[7] Veronika Volgger, Herbert Stepp, Stephan Ihrler, Marcel Kraft, Andreas Leunig, Parag M. Patel, Malavika Susarla, Kathleen Jackson, Christian S. Betz, *Evaluation of optical coherence tomography to discriminate lesions of the upper aerodigestive tract*, Head Neck 35 (11), 1558-1566 (2013)

[8] Matthias Rehberger, Dirk Wilhelm, Klaus-Peter Janssen, Matthias Günther, Michael Witte, Niels König, Niels, Robert Schmitt, *Optical coherence tomography (OCT): The missing link in gastrointestinal imaging?*, Int J CARS 11 (Suppl 1), 32-33 (2016)

[9] D. Lorenser, X. Yang, R. W. Kirk, B. C. Quirk, R. A. McLaughlin, D. D. Sampson, *Ultrathin side-viewing needle probe for optical coherence tomography*, Opt. Lett. 36 (19), 3894-3896 (2011)

[10] Robert Schmitt, Klaus Eder, *A Two Dimensional Scanning Setup for Precise Addressing of Fibers in a Fiber Bundle*, CIRP Annals 56 (1), 505-508 (2007)

One-lung ventilation and Pneumothorax detection using Electrical Impedance Tomography: a preliminary study report

Landeck T[1, 3], Prägler S[1], Salz P[1], Mrongowius J[1], Reske AW[1, 3], Neumuth T[1], Girrbach FF[1, 2]

[1]*University of Leipzig, Innovation Center Computer Assisted Surgery, Leipzig, Germany*

[2]*University of Leipzig, Department of Anesthesiology and Intensive Care Medicine, Leipzig, Germany*

[3]*Heinrich-Braun-Hospital, Department of Anesthesiology, Intensive Care Medicine, Emergency Medicine and Pain Management, Zwickau, Germany*

Kontakt: tobias.landeck@iccas.de

Abstract

Dislocation of the endotracheal tube into the main bronchus and pneumothorax (PTX) are differential diagnoses of one lung ventilation (OLV) which are difficult to distinguish in clinical practice. However, Electrical Impedance Tomography (EIT) might be a future diagnostic tool for automatized diagnosis of a PTX. Therefore, characterizing the appearance of single lung ventilation with open and closed chest during thoracic surgery as a model for different entities of PTX might help to develop new diagnostic algorithms for automatized PTX detection using EIT. We report preliminary findings of 12 patients who were monitored with EIT during thoracic surgery. Our results show that EIT is able to correctly identify the side of OLV. Additionally, we detected an out of phase ventilation like signal in the collapsed lung during open chest. This pattern might help to detect PTX in EIT without a reference image.

Keywords: Pneumothorax, Electrical impedance tomography, out of phase ventilation like signal

1 Problem

Electrical Impedance Tomography is a non-invasive, radiation free, dynamic, bedside monitoring technique for lung ventilation[1]. Steinmann et al[2] demonstrated that OLV with dual lumen tube can be detected with EIT. A clinical scenario where the recognition of OLV is of high importance is the dislocation of the tracheal tube in the bronchial main stem. A crucial differential diagnosis of one sided ventilation is an unilateral pneumothorax (PTX), which can be detected with EIT as well if a reference image was recorded before[3]. The distinction between both with auscultation is difficult, especially when ambient noise is high as in pre-hospital trauma patients. Our hypothesis is that EIT is able to distinct OLV with closed (mimicking endobronchial intubation) and open (mimicking PTX) chest, without a reference image. To test this hypothesis we conducted a clinical trial with thoracic surgery patients, were OLV is necessary for the procedure and thoracotomy is leading to PTX.

2 Material and methods

EIT is a radiation-free, noninvasive monitoring technique. A belt with 16-32 electrodes is placed around the chest of the patient. Through the surface electrodes, small currents are injected in the chest and resulting potential differences between pairs of passive electrodes are measured. Using different reconstruction algorithms, global and local impedance curves can be calculated and displayed as a real-time transversal image of the chest depicting pulmonary ventilation.

After approval by the local ethics committee (No. EK-BR-67/17-1), we conducted an observational study in a total of 12 patients scheduled for open or minimally-invasive thoracic surgery requiring OLV. After informed consent, an EIT belt was placed around the patient's chest preceding induction of anaesthesia and placement of a double lumen tube. EIT files were recorded at different, predefined time-points with the Timpel Enlight 500 EIT monitoring device (Timpel, Sao Paulo, Brazil).

Events recorded on EIT were induction of anesthesia and placement of the double lumen tube, check for correct placement of the tube by subsequent clamping of one of the two lumina of the tube and ausculation, confirmative bronchoscopy, one-lung ventilation (OLV) after opening of the chest (mimicking open pneumothorax), and re-inflation of the collapsed lung at the end of the operation.

3 Results

Representative results of one patient are shown in Figure one. In the shown patient during OLV of the right lung, 100% of the ventilated pixels were detected in the right region of interest (1). Correct tube placement was confirmed or denied by EIT in all patients and verified via bronchoscopy and by auscultation. After thoracotomy during OLV of the right lung (2A) we detected an out of phase ventilation like signal of the collapsed, left lung that persisted throughout the procedure (2B). The absence of ventilation in the collapsed lung was assured by a water trap, the correctly placed double lumen tube and the surgeon who did not observe any inflation of the lung. Impedance changes of the ventilated lung was 23.6 (±0.1) AU and 7.2 (±0.1) AU of the collapsed lung over the five shown cycles. The out of phase signal stopped when the thorax was closed and chest drains reestablished negative pleural pressure at the end of the operation.

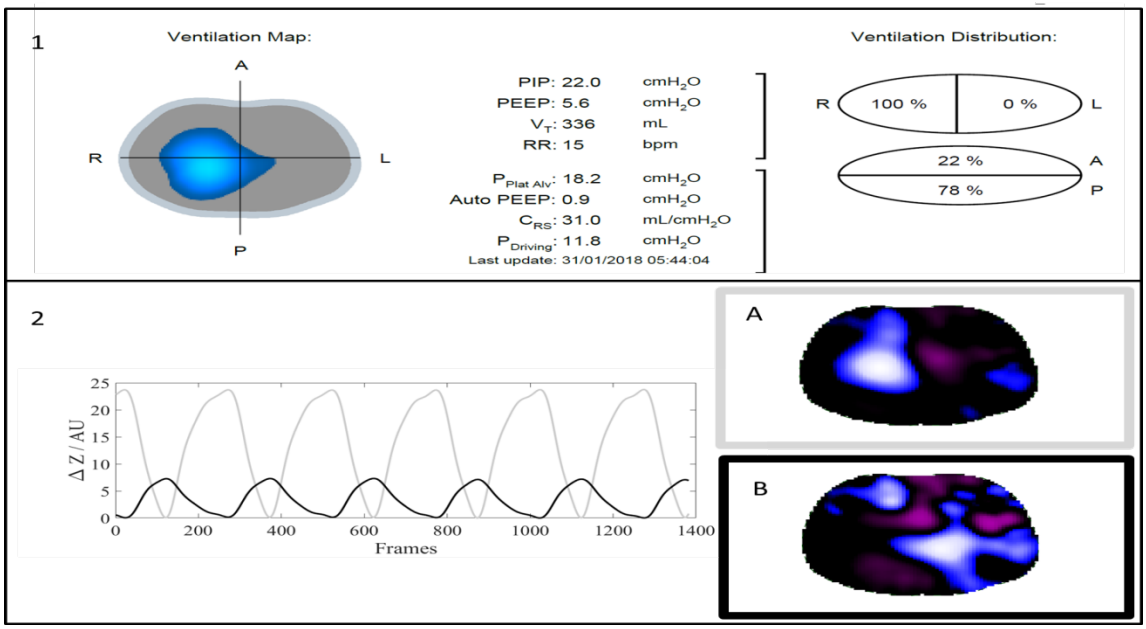


Figure 1: Example of EIT imaging during single lung ventilation of the right lung with closed (1) and open chest (2). *In the closed chest, the EIT ventilation map and the calculation of ventilation distribution clearly identify isolation of the left lung. 2) Impedance curves (left side) during single lung ventilation in the open chest. The impedance curves of the right (light grey) and left (black) lung reveal a delayed (out of phase) ventilation signal of the left lung. The corresponding tidal ventilation images of the EIT are depicted on the lower right side. 2A) Tidal images at the peak of the light grey impedance curve and 2B) at the peak of the black impedance curve.*

4 Discussion

We demonstrated that EIT is able to confirm correct placement of a double lumen tube for OLV. We also found that unilateral lung collapse with open chest generates an ipsilateral out of phase ventilation like signal whose origin needs to be further investigated. One possible explanation is that loss of contact of the lung to the inner chest wall alters the classical path of electrical current and therefore generates artificial, out of phase ventilation like signals. Another explanation might be a passive movement of the collapsed lung in the thorax, following the moving mediastinum with inspiration and expiration of the ventilated lung in lateral position. To our knowledge this is the first description of an out of phase ventilation like pattern in EIT in open total PTX. The signal seems to be specific for an open PTX with a completely collapsed lung and may be valuable for the development of algorithms for automatized PTX detection using EIT. Further experimental studies are required to determine if this phenomenon can also be identified in smaller PTX without total collapse of the affected lung.

5 Conclusion

This case indicates that the EIT is a feasible tool to detect endobronchial intubation. The out of phase ventilation like signal might be characteristic for open PTX in EIT and thus might help to automatically identify a life threatening condition in an emergency scenario, where no reference image is available. Further research is necessary to validate our results.

Acknowledgement

This project is sponsored by the German Federal Ministry of Education and Research (BMBF) as part of the KMU-Innovative-Program (project number: 13GW0179C).

References

1. Frerichs I, Amato MBP, van Kaam AH, et al. Chest electrical impedance tomography examination, data analysis, terminology, clinical use and recommendations: consensus statement of the TRanslational EIT developmeNt stuDy group. *Thorax*. 2016:thoraxjnl-2016-208357. doi:10.1136/thoraxjnl-2016-208357
2. Steinmann D, Stahl CA, Minner J, et al. Electrical impedance tomography to confirm correct placement of double-lumen tube: A feasibility study. *Br J Anaesth*. 2008;101(3):411-418. doi:10.1093/bja/aen166
3. Costa EL, Chaves CN, Gomes S, et al. Real-time detection of pneumothorax using electrical impedance tomography. *Crit Care Med*. 2008;36(4):1230-1238. doi:10.1097/CCM.0b013e31816a0380

Electrical Impedance Tomography for ventilation delay analysis

R. Fuchs[1], J. Mrongowius[1], A. W. Reske[1,2], T. Neumuth[1], P. Salz[1]

[1]University of Leipzig, Innovation Center Computer Assisted Surgery, Leipzig, Germany
[2]Heinrich-Braun-Hospital, Department of Anesthesiology, Intensive Care Medicine, Emergency Medicine and Pain Management, Zwickau, Germany

Contact: reinhard.fuchs@medizin.uni-leipzig.de

Abstract

We propose the usage of ventilation images, obtained through Electrical Impedance Tomography to compute local ventilation delays in regions or single pixels. With this method, medical users will have a tool for identifying unilateral diaphragmatic paresis and estimating severity of inhomogeneous ventilation. Furthermore, the method might be useful for the detection of pneumothoraces.

Keywords: Electrical Impedance Tomography, ventilation delay, diaphragmatic paresis

1 Problem

The interscalene brachial plexus block (ISB) is a common method for postoperative analgesia to reduce pain or paresthesia caused by a shoulder surgery [1]. However, inducing anesthesia can cause more intense side effects. Phrenic nerve blocks are a common by-product, causing unilateral diaphragmatic paresis and reducing respiratory function even after the patient's discharge from the hospital [2].

Diagnosing the impaired function of the diaphragm is commonly done by sonography, while the reduced respiratory capabilities are quantified by spirometry [3]. Electrical Impedance Tomography (EIT) presents an alternative, visualizing pulmonary ventilation through a combination of small-voltage stimulations and measurements, performed by a multitude of electrodes around the patient's thorax [4]. This way, it is possible to achieve real-time monitoring of the lung's ventilation and diagnose the hypoventilation caused by a unilateral phrenic nerve blockage [3].

With a real-time detection method of the ventilation and its delay in one side of the lung, medical personnel would have a tool which allows for an immediate recognition of the symptoms caused by unilateral diaphragmatic paresis. An alternative approach would be the analysis of the relative impedance change ratio between the left and the right half of a reconstructed tidal image. However, since the left-to-right ventilation ratio can be unobtrusive or misleading, this work will not focus on the distribution of pixel values, but rather the delay with which regions reach their peak value.

2 Materials and Methods

Measured voltages were filtered to reduce cardiac signals. EIT-reconstruction was done with EIDORS, an open-source software employing the GREIT-algorithm for ventilation reconstruction by a trained 2.5D thorax-lung-model [5].

The reconstruction returns a series of images at a resolution of 64x64 pixels. The image values represent the relative impedance change $\Delta Z_{ij}(t)$ of each pixel $P[\,i,j\,]$ of the reconstructed image .

Following the accumulation of every single pixel trend, the global relative impedance change $\Delta Z(t)$ was calculated with

$$\Delta Z(t) = \frac{1}{64^2} \sum_{i=1}^{64} \sum_{j=1}^{64} \Delta Z_{ij}(t)\,.$$

Alternatively, the following calculations can be executed with averaged signals for predefined regions of interest. This way, the obtained delay value describes the behavior of a lung region, reducing the amount of contributing signals.

End-inspiratory and end-expiratory times were calculated for the global impedance signal $\Delta Z(t)$ (t_{Ik} - end-inspiratory, t_{Ek} - end-expiratory) and every local signal $\Delta Z_{ij}(t)$ (t_{ijIk} - end-inspiratory, t_{ijEk} - end-expiratory) with I and E defining the type of moment and k denoting the breath number (with $k = 1, 2, \ldots K$). Thus, it was possible to calculate the average time delay between the global and the local extreme points for every pixel / region with

$$t_{ijD} = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{2}\left((t_{ijIk} - t_{Ik}) + (t_{ijEk} - t_{Ek}) \right) .$$

The resulting delay values are drawn on a 64x64 matrix, allowing the visualization through a colormap, whose boundaries are based on the extreme values of the delay matrix. Finally a mask was applied to subdue pixel values with low variance.

3 Results

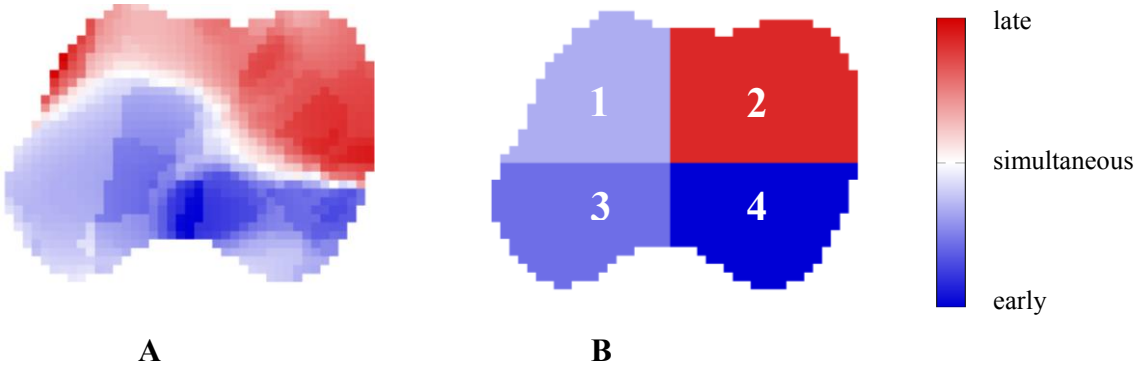


Figure 1: *Delay-Map of a patient with phrenic nerve blockage after surgery; A – delay calculated for each pixel, B – delay calculated for regional averaged impedance signals; red areas mark a positive delay, blue areas mark a negative delay*

The resulting delay maps are images with 64x64 pixels and identify pixels / regions with a phase shift. Figure 1 displays two delay maps, one containing the calculated delay values for every pixel (A), while the other is based on a delay calculation with four averaged pixel signals (B). The sign and magnitude of the phase shift is distinguishable through the color and its intensity (positive - red, negative - blue).

The first map (A) shows the delay for each pixel signal, resulting in a plot showing most of the delay in the ventral regions of the lung and especially on the left lung side (the right map side). The dorsal part of the map has mostly negative delay values, indicating that local signals reach their extreme points well before the global impedance signal.

The second map (B) simplifies the analysis by averaging the pixel impedance signals into 4 distinctive quadrants, thereby reducing the calculations and dividing the final values into ventral right (1), ventral left (2), dorsal right (3) and dorsal left (4). Quadrant 1, 3 and 4 show that the averaged signals reach their respective extreme points earlier than the global impedance signals, while quadrant 2 shows a strong delay.

In summary, the delay map is in accordance with the diagnosed phrenic nerve block on the left side (the right image side), caused by an ISB during shoulder surgery on the left shoulder.

4 Discussion

The proposed method provides a visual feedback for the localized delay of ventilation in corresponding lung regions. This allows for a fast identification of regions affected by an impaired respiratory system, something that may stem from an unilateral activation of the diaphragm. Furthermore, it is possible to divide the maps into regions of interest, which allows for a faster diagnosis and an easier differentiation between affected lung areas and the magnitude of their delay in regard to each other.

When compared to alternative approaches for detecting asynchronous ventilation, i.e. the left-to-right ratio, this method enables the user to perform a better diagnosis based on dynamic information. Instead of focusing on the relative impedance change values and their distribution, the map delivers a feedback which depends on the regional temporal progress in regard to local air flow.

5 Conclusion

We proposed a method for calculating the ventilation delay, occurring when nerve blocks caused by a postoperative analgesia result in an impaired function of the diaphragm and unilateral ventilation. Through imaging using EIT, it is possible to compare local and global times of end-inspiratory and end-expiratory moments. The resulting plot of the delay calculation allows for a fast assertion whether or not significant delays exist and which lung regions are affected by it. Additionally, in contrast to other works [6], our method does not require a slow-inflation maneuver.

Acknowledgements

This project is sponsored by the German Federal Ministry of Education and Research (BMBF) as part of the KMU-Innovative-Program (project number: 13GW0179C).

References

[1] M. J. Fredrickson, S. Krishnan, and C. Y. Chen, 'Postoperative analgesia for shoulder surgery: a critical appraisal and review of current techniques: Postoperative analgesia for shoulder surgery', Anaesthesia, vol. 65, no. 6, pp. 608–624, May 2010.

[2] W. F. Urmey, K. H. Talts, and N. E. Sharrock, 'One Hundred Percent Incidence of Hemidiaphragmatic Paresis Associated With Interscalene Brachial Plexus Anesthesia as Diagnosed by Ultrasonography', Anesthesia & Analgesia, vol. 72, no. 4, p. 498–503, Apr. 1991.

[3] M. Wiegel, S. Hammermüller, H. Wrigge, and A. W. Reske, 'Electrical Impedance Tomography Visualizes Impaired Ventilation Due to Hemidiaphragmatic Paresis after Interscalene Brachial Plexus Block', Anesthesiology, vol. 125, no. 4, p. 807, Oct. 2016.

[4] Frerichs et al., 'Chest electrical impedance tomography examination, data analysis, terminology, clinical use and recommendations: consensus statement of the TRanslational EIT developmeNt stuDy group', Thorax, vol. 72, no. 1, pp. 83–93, Jan. 2017.

[5] Adler et al., 'GREIT: a unified approach to 2D linear EIT reconstruction of lung images', Physiological Measurement, vol. 30, no. 6, pp. S35–S55, Jun. 2009.

[6] T. Muders et al., 'Tidal recruitment assessed by electrical impedance tomography and computed tomography in a porcine model of lung injury', Critical Care Medicine, vol. 40, no. 3, pp. 903–911, Mar. 2012.

Automatic Liver and Tumor Segmentation in Late-Phase MRI Using Fully Convolutional Neural Networks

G. Chlebus[1], *H. Meine*[1,2], *N. Abolmaali*[3], *and A. Schenk*[1]

[1] *Fraunhofer Institute for Medical Image Computing MEVIS, Bremen, Germany*
[2] *Medical Image Computing Group, University of Bremen, Bremen, Germany*
[3] *Department of Radiology, Städtisches Klinikum Dresden, Dresden, Germany*

Contact: grzegorz.chlebus@mevis.fraunhofer.de

Abstract

Liver and tumor segmentation plays an important role for many liver interventions. Automation of segmentation steps will bring a substantial improvement to clinical workflows by reducing the planning time and decreasing the inter-observer variability. While liver and tumor segmentation in CT data is a well-studied subject, only few publications are available for MRI data. We present an automatic segmentation approach based on 2D fully convolutional neural networks, which can be applied to both segmentation tasks not only in CT, but also in MR images. For the challenging MRI data, we evaluated two algorithms, one using an axial network and a second one based on three orthogonal models. The latter and better method resulted in a mean Dice index of 0.95 and 0.65 for liver and tumor segmentation, respectively. The liver segmentation quality of the method matches that of experienced clinical users and is higher compared to other automatic approaches and results in the literature. The tumor segmentation requires further improvements to make it comparable to the results of human experts.

Keywords: deep learning, liver, tumor, segmentation

1 Problem

Many liver interventions such as radiofrequency ablation or selective internal radiation therapy (SIRT) require liver and tumor segmentation for procedure planning. Robust and fully automatic segmentation algorithms could speed up the planning process and reduce the inter-observer variability of manual segmentations. Liver and tumors exhibit large variation in shape and appearance, with the latter depending additionally on the used imaging modality and protocol. These factors make the development of automatic segmentation algorithms very challenging.

Several approaches have been proposed for CT images, e.g., within the recent Liver Tumor Segmentation Challenge (LiTS) organized in 2017, where organ and tumor segmentation was evaluated [1]. The best five methods submitted to the ISBI and MICCAI rounds used fully convolutional neural networks (FCN). These methods included approaches using one FCN for segmentation of both liver and tumors [2], cascaded FCNs with separate models for liver and tumors [3], and 3D networks combining original image data with features derived from 2D models [4].

Although MRI has proven to show a significantly higher sensitivity for primary liver cancer in most studies [5], automatic segmentation in MRI has received less attention in the last five years compared to CT. To our best knowledge, there is currently only one publication applying FCNs to MRI data for liver and tumor segmentation [6], while other groups used established approaches for this task. Examples are Suzuki et al., who employed a fast-marching method for rough liver segmentation followed by a geodesic active contour model coupled with a level-set algorithm for mask refinement [7] and Le et al., who proposed a histogram-based liver segmentation method, where the histogram analysis was done by a neural network [8]. The final liver shape was obtained using geodesic active contours. A liver probability map was computed by Bereciartua et al., which was subsequently used to guide the evolution of a 3D liver model [9]. Huynh et al. computed a rough liver mask using a watershed segmentation and statistical properties of the liver [10]. The subsequent refinement step was done with an active contour model. A method based on k-Means clustering and active contours was proposed by Ivashchenko et al. to automatically segment the liver, hepatic artery and biliary tree [11].

In this work, we present our segmentation method based on 2D fully convolutional neural networks for liver and tumor segmentation in MRI and its evaluation. We provide results for two different network designs and show the influence of the training set size on the resulting liver model performance. A summary of segmentation results in MRI data from the literature is given and discussed in comparison to the presented method and data used. This paper extends our previous work on liver segmentation in MRI, where a first evaluation on a smaller dataset was presented and tumor segmentation was not included [12].

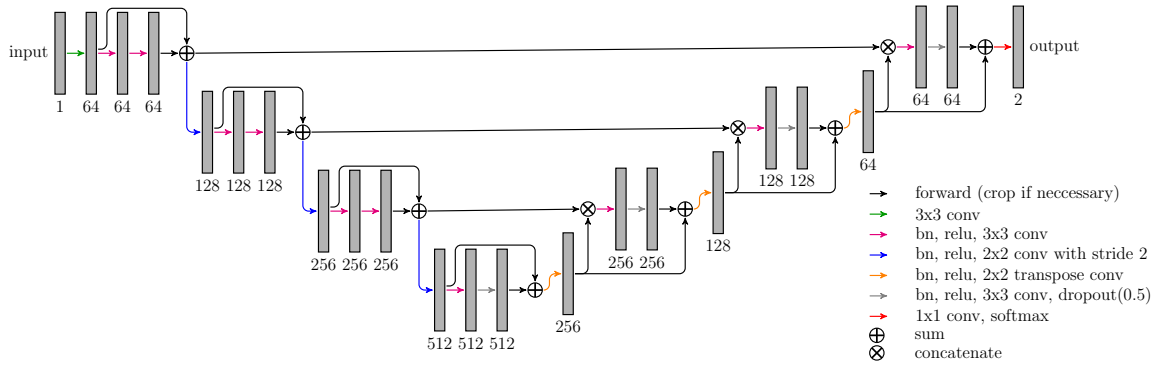


Figure 1: Overview of the neural network architecture. The numbers denote the feature map count.

2 Material and Methods

2.1 Data

For our study we used DCE-MRI images of 90 patients with primary liver cancer and/or liver metastases (76 scheduled for SIRT) acquired at Städtisches Klinikum Dresden, Germany, on a 3T Discovery MRI scanner (GE Healthcare Systems, USA). The image in-plane resolution ranged from 0.74 to 1.76 mm and the slice thickness from 2 to 5 mm. For segmentations we selected the late hepatocellular phase acquired 15 min after iv-injection of contrast agent (Primovist®, Bayer Healthcare, Germany), because it provides a high liver/tumor and liver/surrounding organs contrast. The reference masks used for training and evaluation of automatic methods were defined by medical radiology assistants with over ten years of experience with a semi-automatic contouring tool [13]. The resulting dataset consisted of 90 and 85 images with reference liver and tumor segmentations, respectively. The tumor dataset was smaller, because five patients had no tumors. We randomly selected 57 (60) training, 5 (5) validation and 28 (20) test cases for training and evaluation of liver (liver tumor) deep learning models.

Imaging data for this study was evaluated after approval by the local ethics committee (EK-BR-79/16-1 / 118834).

2.2 Neural network

Architecture The employed neural network is a U-net like architecture [14] (Fig. 1). The model works on four resolution levels, which allows for accumulation of local and global features. Downscaling (upscaling) is done using strided (transposed) 2×2 convolutions. Features computed in the downscaling path are passed to the upscaling path via long skip connections. Short skip connections were added to improve the gradient flow throughout the network [15]. Each convolution block has 2 convolutional layers using 3×3 filter size. Each convolution is followed by a batch normalization layer and ReLU activation function. To avoid overfitting we used spatial dropout in the upscaling path with drop probability 0.5.

Training We trained the networks using 2D image patches (size 228×228) and mini-batch size of 16. Because all convolutions are computed in the so-called valid mode, the input was reflectively padded with 45 pixels on each side. We employed the soft dice as a loss function, which was computed on the pixelwise softmax of the network output. We trained the models using the Adam optimizer with 10^{-4} learning rate. During validation (every 500 training steps) the model quality was evaluated using the Jaccard coefficient. The training was stopped if the model quality had not improved for 20 validations.

Data preprocessing To normalize the MR images, we mapped the intensities between the 2nd and 98th percentile of each 3D volume to the [0, 1] range. Values outside of the target range were clipped. Images were resampled to 2 mm isotropic voxel size. We augmented the training dataset by random rotations $\alpha[\text{deg}] \sim \mathcal{N}(0, 10)$ and random intensity shifts $x \sim \mathcal{N}(0, 0.1)$.

2.3 Segmentation pipeline

We trained separate models for liver and tumor segmentation. For both segmentation tasks, we consider two types of segmentation pipelines: the first one (*Axial*) employs one model trained with axial patches; the second one (*OrthoMean*) is a triplanar approach using three networks trained on axial, coronal and sagittal patches [16]. The output of *OrthoMean* is computed as an average of masks produced by the three models.

The final segmentation is obtained by applying a threshold of 0.5. For the liver segmentation only the biggest connected component is retained, whereas the tumor segmentation output is masked with a liver region to remove false positives found outside of the organ.

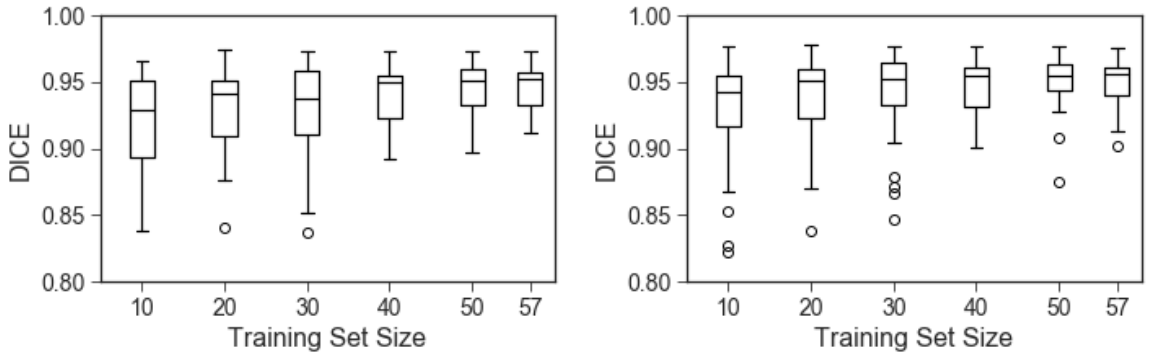
2.4 Segmentation evaluation

We evaluated the segmentation quality by comparing the automatic results with the reference on the test cases left out for evaluation. The quality was assessed using the Dice index (DICE) and relative volume error (RVE). Additionally, we reported processing times on a desktop PC (Intel Core i7-4770K, 32GB RAM, NVIDIA Titan Xp). All significance tests were done using the Wilcoxon signed-rank test at 0.05 level.

3 Results

3.1 Liver segmentation

We computed the liver segmentation for 28 test cases. Box plots showing DICE distribution for *Axial* and *OrthoMean* when varying the number of training cases are shown in Fig. 2. All increases of the training size led to a significantly better DICE except for transitions 20-30, 50-57 for *Axial* and 40-50, 50-57 for *OrthoMean*. The DICE for *Axial* and *OrthoMean* trained on all training cases was 0.946 and 0.951, respectively (details are given in Tab. 1). *OrthoMean* showed significantly better liver segmentations than *Axial* according to DICE, while there was no significant difference for the RVE metric. Example cases and results of both approaches are shown in Fig. 3. Large liver tumors located near the organ boundary led in one case to incorrect exclusion of liver region from the segmentation mask (Fig. 3b). The processing time for *Axial* was faster than for *OrthoMean* (factor of 3).



(a) (b)

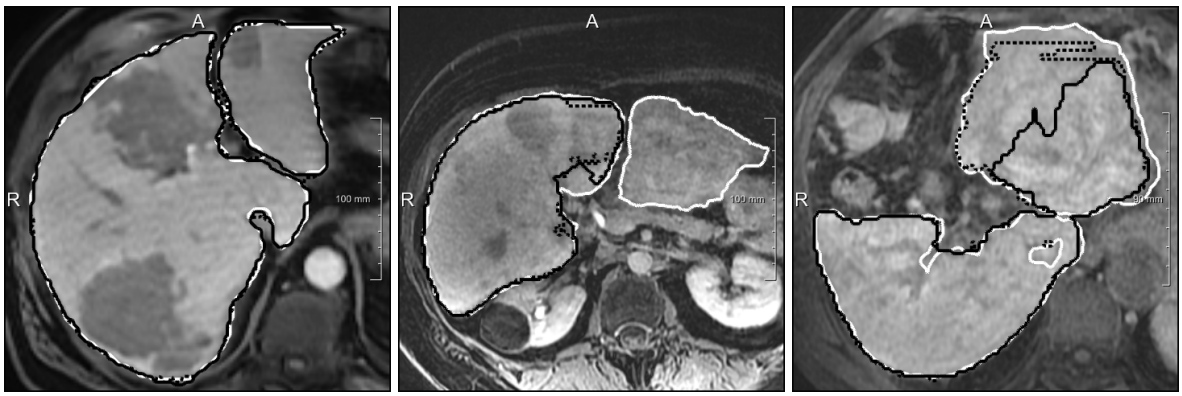
Figure 2: Box plots showing the relationship between the training set size and DICE computed on test cases for *Axial* (a) and *OrthoMean* (b).

Table 1: Evaluation results (mean ± stddev) for liver and tumor segmentation.

	DICE	RVE[%]	t[s]
Liver			
Axial	0.946 ± 0.018	4.20 ± 3.34	2.05 ± 0.34
OrthoMean	0.951 ± 0.018	4.20 ± 3.65	7.32 ± 0.36
Tumor			
Axial	0.627 ± 0.241	48.9 ± 53.3	1.73 ± 0.73
OrthoMean	0.647 ± 0.210	35.9 ± 28.2	7.63 ± 2.23

3.2 Liver tumor segmentation

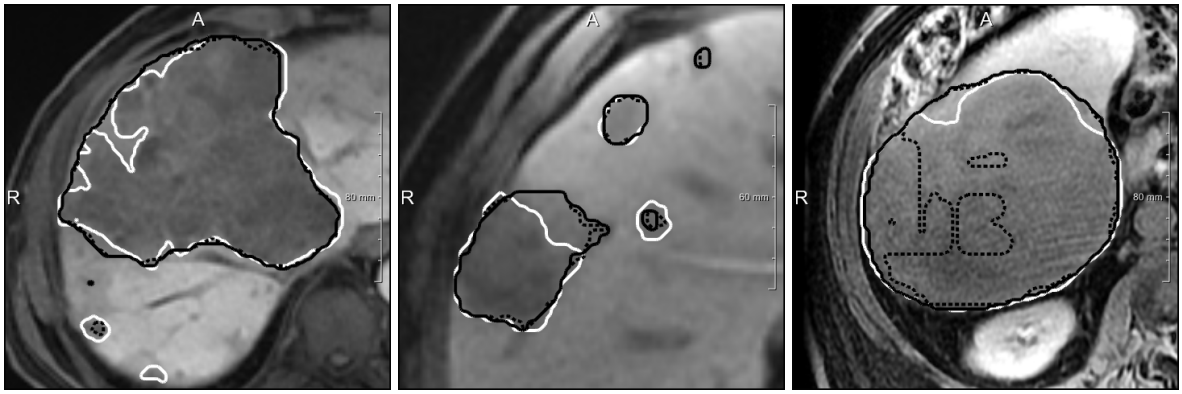
The evaluation results for 20 test cases are summarized in Tab. 1. The output of tumor segmentation models was limited to the reference liver masks to avoid influence of liver segmentation quality on the tumor segmentation results. No significant differences were found between *OrthoMean* and *Axial* for DICE and RVE. Fig. 4 shows



(a) (b) (c)

Figure 3: Example liver segmentation results of *Axial* (solid) and *OrthoMean* (dashed) compared with reference (white). (a) Case where both approaches produced accurate segmentations (0.97 and 0.98 DICE for *Axial* and *OrthoMean*, respectively). (b) Case with a big tumor in the left liver lobe causing segmentation errors (0.93 and 0.91). (c) Case where the additional orientations used by *OrthoMean* improved the results in the left liver lobe compared to *Axial* (0.91 and 0.95).

exemplary cases and outputs of both methods. Both approaches were able to produce results close to the reference in many cases (Fig. 4a, 4b). In some cases, additional orientations used by *OrthoMean* helped to detect small tumors (Fig. 4a), in others they caused errors (Fig. 4c). Similarly to liver segmentation, the processing time of *OrthoMean* was approximately 3 times that of *Axial*.



(a) (b) (c)

Figure 4: Example tumor segmentation results of *Axial* (solid) and *OrthoMean* (dashed) compared with reference (white). (a) Case with the highest DICE score (0.84 and 0.87 for *Axial* and *OrthoMean*, respectively). (b) Case with a low RVE 3% and 7% for *Axial* and *OrthoMean*, respectively (0.84 and 0.85). (c) Case where substantial part of a very big tumor was missed by *OrthoMean* (0.87 and 0.83).

4 Discussion

Liver segmentation When comparing the two approaches of *Axial* and *OrthoMean*, the additional image orientations used by *OrthoMean* helped to increase the segmentation quality in particular in cases of resected livers (Fig. 3c). The liver segmentation quality of both investigated approaches was comparable to that of routine segmentations done by one radiologist and two residents as reported in [12]: 0.94±0.02, 0.95±0.02, 0.95±0.02 and 5.37±4.38, 3.63±4.57, 4.33±4.62 for DICE and RVE[%], respectively. Comparing the results of our approaches to the accuracy measures of other automatic methods given in the literature, the achieved DICE was higher for our method (Tab. 2). This is particularly remarkable, since we trained and applied our algorithms to a highly challenging dataset with patients mostly scheduled for radioembolisation and thus including a large number of tumors and diseased livers. It is important to note, that a comparison based on

literature remains difficult due to patient and image differences (scanner type, acquisition protocol, contrast agent, liver pathologies, etc.).

Tumor segmentation The DICE for our two approaches was 0.627 and 0.647 comparing the results to the reference defined by the experts. The DICE of 0.697 reported by Christ et al. was a bit higher, but a direct comparison seems not to be applicable due to differences in datasets used for evaluation with respect to image resolution and acquisition protocols for example (late-phase of contrast-enhanced MRI vs. diffusion weighted images, see Tab. 2 for details). Nevertheless, both approaches are far from matching human performance, which according to an inter-reader study can achieve a DICE of 0.90-0.93 [17]. One of the reasons for such a big quality difference may be the fact, that clinicians in contrast to our method use more than one contrast phase or MRI sequence for segmentation purposes.

Liver vs. tumor Although we applied the same method to both segmentation tasks, the liver models produced segmentations of a substantially higher quality than the tumor ones. This difference could be explained by the fact, that tumors exhibit a bigger variability in shape and appearance compared to the liver. Thus, collecting more training data with different tumor types would help to improve the tumor segmentation performance.

Table 2: Comparison of our approaches to results from the literature on automatic liver and tumor segmentation in MRI.

			Dataset			
	DICE	#	Sequence	Scanner	Resolution [mm]	Slice thickness [mm]
Liver						
Axial, *OrthoMean*	0.946 ± 0.018, 0.951 ± 0.018	90	LAVA	3T Discovery(GE)	0.74-1.76	2.0-5.0
Christ et al. [6]	0.870	31	n/a	1.5T Avanto(Siemens)	n/a	5.0
Suzuki et al. [7]	0.936 ± 0.017*	23	LAVA, THRIVE	1.5T Signa HDx/HDxt(GE), 1.5T Achieva(Philips)	1.17-1.72	4.0-5.0
Le et al. [8]	0.910 ± 0.028	10	VIBE	1.5T Avanto(Siemens)	1.18-1.40	3.5-4.0
Bereciartua et al. [9]	0.902 ± 0.086	18	VIBE	1.5T Avanto(Siemens)	n/a	n/a
Huynh et al. [10]	0.911 ± 0.019	27	VIBE, LAVA, THRIVE	1.5T Signa HDx/HDxt(GE), 1.5T Achieva(Philips), 1.5T Avanto(Siemens)	1.17-1.72	3.5-5.0
Ivashchenko et al. [11]	n/a	5	mDIXON	n/a	1.00	1.5
Tumor						
Axial, *OrthoMean*	0.627 ± 0.241, 0.647 ± 0.210	85	LAVA	3T Discovery(GE)	0.74-1.76	2.0-5.0
Christ et al. [6]	0.697	31	n/a	1.5T Avanto(Siemens)	n/a	5.0

* Average for CT and MRI dataset.

5 Conclusion

In this work, we presented two deep learning based approaches for automatic liver and liver tumor segmentation in MR images. We evaluated both methods on data including mostly patients scheduled for SIRT, which is particularly challenging due to numerous liver pathologies. Our triplanar segmentation algorithm *OrthoMean* using three image orientations resulted in segmentations of higher quality for liver than *Axial*. The liver segmentations showed a similar accuracy compared to routine segmentations done by clinicians and achieved higher mean Dice score than other automatic methods given in the literature. The quality of tumor segmentation was relatively low compared to that of the liver segmentation. One option to increase the performance of tumor segmentation models may be increasing the size of the training data set. Moreover, other architectures as 3D networks, which we are currently evaluating, or using additional contrast phases and MRI sequences (future work) could further increase the segmentation accuracy. To allow for direct comparisons of automatic methods a benchmark for liver and tumor segmentation in MRI similar to the LiTS challenge would be helpful.

6 Acknowledgments

We would like to thank Christiane Engel, Andrea Koller, Smita Thoduka and Claudia Böhme for providing

feedback and manual segmentations. The Titan Xp used for this research was donated by the NVIDIA Corporation.

References

[1] Christ P, Ettlinger F, Gruen F, Lipkova J, Kaissis G, *LiTS - Liver Tumor Segmentation Challenge*, www.lits-challenge.com (2017)

[2] Han X, *Automatic Liver Lesion Segmentation Using A Deep Convolutional Neural Network Method*, arXiv preprint arXiv:170407239 (2017)

[3] Vorontsov E, Tang A, Pal C, Kadoury S, *Liver lesion segmentation informed by joint liver segmentation*, Proceedings of the IEEE 15th International Symposium on Biomedical Imaging (2018)

[4] Li X, Chen H, Qi X, Dou Q, Fu CW, Heng PA, *H-DenseUNet: Hybrid densely connected UNet for liver and liver tumor segmentation from CT volumes*, arXiv preprint arXiv:170907330 (2017)

[5] Roberts LR et al., *Imaging for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis*, Hepathology, **67(1)** 401–421 (2018)

[6] Christ PF et al., *Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks*, arXiv preprint arXiv:170205970 (2017)

[7] Suzuki K et al., *Computerized segmentation of liver in hepatic CT and MRI by means of level-set geodesic active contouring*, Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2013)

[8] Le TN, Bao PT, Huynh HT, *Fully automatic scheme for measuring liver volume in 3D MR images*, Bio-Medical Materials and Engineering, **26(s1)** 1361–1369 (2015)

[9] Bereciartua A, Picon A, Galdran A, Iriondo P, *Automatic 3D model-based method for liver segmentation in MRI based on active contours and total variation minimization*, Biomedical Signal Processing and Control, Biomedical Signal Processing and Control, **20** 71–77 (2015)

[10] Huynh HT, Le-Trong N, Bao PT, Oto A, Suzuki K, *Fully automated MR liver volumetry using watershed segmentation coupled with active contouring*, International Journal of Computer Assisted Radiology and Surgery, **12(2)** 235–243 (2017)

[11] Ivashchenko O et al., *Automated segmentation of complete 3D liver model including arterial hepatic vasculature and biliary treee anatomy from a single multiphase MR scan*, International Journal of Computer Assisted Radiology and Surgery, **12(s1)** 25–26 (2017)

[12] Schenk A, Chlebus G, Meine H, Thoduka S, Abolmaali N, *Deep learning for liver segmentation and volumetry in late phase MRI*, 2018 European Congress of Radiology, Vienna, Feb 28–Mar 4 (2018)

[13] Schenk A, Prause G, Peitgen HO, *Efficient semiautomatic segmentation of 3D objects in medical images*, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000, 186–195, 978-3-540-40899-4, Springer, Berlin, Heidelberg (2000)

[14] Ronneberger O, Fischer P, Brox T, *U-net: Convolutional networks for biomedical image segmentation*, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 234–241, 978-3-319-24574-4, Springer, Cham (2015)

[15] Drozdzal M, Vorontsov E, Chartrand G, Kadoury S, Pal C, *The importance of skip connections in biomedical image segmentation*, Deep Learning and Data Labeling for Medical Applications, 179–187, 978-3-319-46976-8, Springer, Cham (2016)

[16] Prasoon A et al., *Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network*, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, 246–253, 978-3-642-40763-5, Springer, Berlin, Heidelberg (2013)

[17] Budjan J at al., *Semi-automatic Volumetric Measurement of Treatment Response in Hepatocellular Carcinoma After Trans-arterial Chemoembolization*, Anticancer Research, **36(8)** 4353–4358 (2016)

A Framework For The Quantitative Assessment of Image-guided Percutaneous Ablation of Hepatic Lesions

R. Hrabuska[1], R. Sandu[1], I. Paulocci[1], K. Gerber[1], S. Weber[1]

[1]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

Contact: radek.hrabuska@artorg.unibe.ch

Abstract

In recent years, percutaneous ablations of hepatic lesions are becoming a viable alternative to liver resection. Currently, radiologists' subjective evaluation of the ablation success is considered as gold standard. However, further improvements in ablation technique require quantitative evaluation.

To compute such evaluation autonomously, it is necessary to annotate both lesions and resulting ablation zones. This responsibility has been recently transferred from physician to automatic methods based on computer vision and machine learning algorithms. Nevertheless, latter method requires currently unavailable large amounts of data. A semi-automatic method for lesion annotation could be a solution to this issue.

During intervention, metrics computed from annotated data will be used to support outcome evaluation. Metrics can be devised on basis of volume or distance measurement. Introduced metrics include coverage ratios and distances, in order to confirm sufficient size of ablation margin. Resulting set of evaluation criteria will provide tools for future studies of treatment results.

Keywords: Thermal Ablation, Image Segmentation, Quantitative Measurement, Image Processing

1 Problem

Percutaneous microwave ablation is a minimally invasive, image-guided therapy for the local treatment of primary and secondary hepatic lesions in selected patients. Microwave ablation is an alternative therapy if surgical resection, which is considered to be the gold standard, is contraindicated. This therapy aims to encompass the lesion and a 5–10 mm safety margin of tissue around it for complete destruction of the cancerous cells. The outcome of the ablation therapy is evaluated in clinical practice intra-operatively. During a CT-guided intervention, the physician visualizes the resulted necrosis area post-ablation on a CT image. If the physician decides that the lesion was not completely ablated, re-ablation of the lesion is performed. Generally, it is assumed that risk of local recurrence increases in case the minimal safety margin of 5-10mm is not achieved. However, a visual assessment of such small range of distances between two 3D objects could be not only challenging, but also subjective and thus error prone. To address this issue a quantitative way to assess the outcome of the ablation therapy is necessary. To the best of our knowledge, most quantitative assessment tools for ablation outcome were only used post-operatively and focused on retrospective analysis. Moreover, their end goal were registration benefits and accuracy, and not treatment efficacy evaluation [1], [2].

In order to obtain quantitative metrics, in the first step, the lesion and the ablation must be delineated accordingly. The physicians defines the margins of the lesion and ablation and approximates the distance between the two mentally. An annotation tool that provides fast and accurate segmentations of liver structures intra-operatively would enable not only the computation of quantitative metrics, but also the creation of valuable ground-truth data. Nonetheless, creating such database proves to be a time-consuming task to be performed mundanely during or outside the intervention [3], [4]. With development of machine learning and image processing algorithms, this process can be accelerated [3].

In recent years, machine learning and expert systems were developed capable of handling image segmentation tasks for lesions [5]–[7]. However, there are no sufficiently large datasets to cover all peculiarities of lesion and specifically, ablation segmentation. Specific tissue changes in different types of lesions and different ablation modalities have many degrees of freedom, resulting in unsatisfactory automatic segmentation, consequently increasing the demand for user input. One way to minimize user interaction is to employ semi-automatic methods to initialize segmentation and manual correction tool. For such initialization, there is a large spectrum of algorithms that allow for efficient segmentation initialization [8].

2 Material and Methods

The following workflow is already established in commercial systems for computer-assisted surgery. Before the intervention, a CT scan is acquired with contrast agent, followed by identification of lesions for planning purposes. After planning, the ablation probe is placed precisely into the lesion using image guided navigation. In the next step, the probe position is verified and ablation is performed. After intervention, post interventional contrast CT scan is used to identify the ablation zone. The physician visually evaluates the result of ablation with overlaid pre- and post- interventional CT scan. In case of insufficiently ablated lesion, intervention is repeated. As an extension of current workflow in commercial system (CAS-One IR), image annotation tool and quantitative metrics are proposed.

2.1 Image Segmentation

The image annotation tool integrated into the computer-assisted system employs a semi-automatic algorithm. The Fast Marching algorithm [9] was selected based on speed, sufficiently accurate segmentation and minimal user input. The only user interaction required is initialization of the algorithm by setting the seed point.

In principle, fast marching algorithm considers image as topological map, on which curve is evolved. Curve parameters can be obtained by solving boundary conditions problem. For each voxel of segmented region of interest, speed function $F|\nabla\phi|$ is calculated producing time of arrival t from selected initial seed point x to every voxel in volume, as stated in following equations [9]:

$$\Phi_t + F|\nabla_\Phi| = 0, \tag{1}$$

$$\Phi(x, t = 0) = given \tag{2}$$

Resulting image with times of arrival is thresholded to obtain binary segmentation mask.

In case the initial annotation of the liver structure is not deemed satisfactory by the physician, adjustment can be performed with a manual annotation tool.

2.2 Ablation efficacy metrics

To quantitatively measure the outcome of the ablation treatment multiple metrics are proposed. These are defined by volume and distance metrics.

Euclidian distances between individual surface voxels of ablation and lesion were calculated using Maurer's algorithm[10]. To visualize the relative frequency of distances from lesion surface to ablation surface the distances were displayed into a histogram. Moreover, a traffic light color coding scheme was applied to distinguish the safety margin convention as shown in Figure 1.

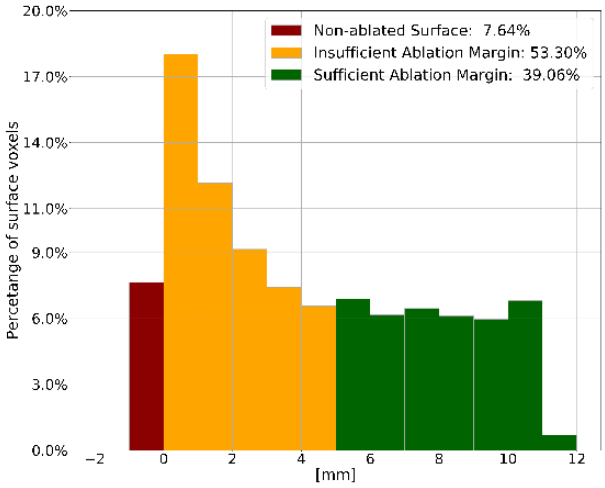


Figure 1*: Surface-to-surface histogram created on the synthetic example. Colors represent adequacy of achieved margin (red – up to 0mm, ablation does not cover lesion volume completely; yellow – up to 5mm, ablation does not cover ablation margin completely; green – over 5mm, sufficient margin of ablation)*

To complement information from the histogram, volume based metrics are implemented to provide overall status of ablation efficacy and create a parameter of comparison for future studies. Therefore, we propose to assess completeness of lesion destruction by Volume Coverage Ratio (VCR):

$$VCR = 1 - \frac{|V_A \cap V_I|}{|V_I|}, \quad (3)$$

where V_A is volume of ablation, V_I is target investigated volume. VCR can be applied to lesion volume resulting in Lesion Coverage Ratio (LCR), or lesion volume extended to demanded ablation margin, resulting in Margin Coverage Ratio (MCR). The residual lesion volume (V_t) was also computed.

All of these metrics can be intraoperatively displayed to performing staff in order to quantitatively assess the outcome of ablation treatment.

2.3 Experiment

As CT scanners vary from model to model, spatial resolution is one the key factors in precision of proposed metrics. Therefore, three settings of spatial resolution are tested, namely 0.5mm, 1mm and 1.5mm, as values used by common CT scanners.

To assess precision of metrics under such conditions, synthetic examples were crafted. First example consists of concentric spherical masks, with ablation zone with radius of 15mm and lesion with radius of 10mm, second one of identical sphere masks (radius 10mm) and third one of spheres with same radius (10mm) and offset of 2mm.

Precision of three metrics is evaluated. For volume measurement, difference in volume of spheres is measured. Distance based metrics are evaluated on spheres with different diameters. Precision of VCR is evaluated on identical spheres with shifted centers.

3 Results

In Tables 1-3 results of precision measurements are presented. Table 1. contains results for volume measurements, in which real volume is calculated by empirical formula for sphere and measured volume is calculated from volume of voxels in binary mask.

Table 1*: Volume measurement accuracy*

Spacing [mm]	Real Volume [ml]	Measured Volume [ml]	Volume Difference [%]
0,5	4,18	4,18	0
1	4,18	4,17	0,24
1,5	4,18	4,17	0,24

Table 2. contains results for surface-to-surface distance measurements from ablation to lesion surface.

Table 2*: Distance measurement accuracy*

Spacing [mm]	Maximal Surface-to-surface Distance [mm]	Minimal Surface-to-surface Distance [mm]	Median Surface-to-surface Distance [mm]	Real Distance [mm]
0,5	9,07	4,3	5,1	5
1	5,1	3,46	4,36	5
1,5	4,85	2,5	3,35	5

Table 3. contains results for volume coverage measurements, in which VCR real was calculated by using empirical formula to evaluate volume of overlapping spheres and VCR measured was calculated using equation 3.

Table 3*: Volume coverage ratio accuracy*

Spacing [mm]	VCR Real [-]	VCR Measured [-]	VCR Difference [%]
0,5	0,78	0,78	0
1	0,78	0,78	0
1,5	0,78	0,78	0

4 Discussion

The results obtained in Tables 1 and 3 support the fact that the volume metrics are suitable for a quantitative assessment of the ablation therapy.

Table 2 introduces results of surface-to-surface distances. The difference between real and measured distances can be explained by surface calculation algorithm`s imperfections. Difference can be further decreased by increasing mask resolution.

For both types of metrics, it's necessary to take into account registration error between images. Results of measurements are more likely to be affected more by registration error between pre- and post- interventional image, than by inaccuracies of segmentation, voxel partitioning and low resolution.

5 Conclusion

Improvements in quantifying MWA results and procedure efficacy require further improvements in methods of evaluation. Such incremental improvements require modern image processing methods and their application specific optimization, in order to increase efficacy of intervention and decrease duration of intervention. In addition introduced quantitative measures improve reproducibility of intervention. .

6 **The presented framework is the first step towards a quantitative intra-operative evaluation, given we found no previous established ground work. From clinicians' point of view, added value of the framework lies in precise measurements available to aid expert decision. By incorporating such solution into commercial systems, large amount of data can be produced to study efficacy of image guided percutaneous ablation. Future work can also be oriented on retrospective analysis of percutaneous interventions.**

7 References

[1] C. Rieder *et al.*, "Software-assisted post-interventional assessment of radiofrequency ablation," 2014, p. 903604.

[2] "Local incompressible registration for liver ablation surgery assessment - Fu - 2017 - Medical Physics - Wiley Online Library." [Online]. Available: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.12535. [Accessed: 30-May-2018].

[3] B. Kerbl *et al.*, "Intervention Planning of Hepatocellular Carcinoma Radio-Frequency Ablations," in *Clinical Image-Based Procedures. From Planning to Intervention*, vol. 7761, K. Drechsler, M. Erdt, M. G.

Linguraru, C. Oyarzun Laura, K. Sharma, R. Shekhar, and S. Wesarg, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 9–16.

[4] "Interactive Volumetry Of Liver Ablation Zones | Scientific Reports." [Online]. Available: https://www.nature.com/articles/srep15373.epdf. [Accessed: 30-May-2018].

[5] M. Moghbel, S. Mashohor, R. Mahmud, and M. I. B. Saripan, "Automatic liver tumor segmentation on computed tomography for patient treatment planning and monitoring," *EXCLI J. 15Doc406 ISSN 1611-2156*, 2016.

[6] P. F. Christ *et al.*, "Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, vol. 9901, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds. Cham: Springer International Publishing, 2016, pp. 415–423.

[7] W. Wu, S. Wu, Z. Zhou, R. Zhang, and Y. Zhang, "3D Liver Tumor Segmentation in CT Images Using Improved Fuzzy *C* -Means and Graph Cuts," *BioMed Res. Int.*, vol. 2017, pp. 1–11, 2017.

[8] R. Rajagopal and P. Subbaiah, "A SURVEY ON LIVER TUMOR DETECTION AND SEGMENTATION METHODS," vol. 10, no. 6, p. 5, 2015.

[9] J. A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, 1999.

[10] C. R. Maurer, Rensheng Qi, and V. Raghavan, "A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 265–270, Feb. 2003.

Model-Based Imitation of Patient Scenarios for Oncological Decision Support

J. Gaebel[1], *J. Müller*[1], *M. Stoehr*[2], *S. Oeltze-Jafra*[1]

[1] *Innovation Center Computer Assisted Surgery (ICCAS), University Leipzig, Leipzig, Germany*
[2] *University Hospital Leipzig, Dept. of Otolaryngology, Head and Neck Surgery, Leipzig, Germany*

Contact: jan.gaebel@medizin.uni-leipzig.de

Abstract

Finding an individualized therapy decision for a cancer patient remains a challenge for interdisciplinary teams. The large amount of treatment options, their possible combinations, and the respective medical knowledge must be digested to finally decide for the most promising treatment. In this work, we propose a concept for the imitation of patient characteristics for treatment decision support based on a digital patient model. This will allow the interdisciplinary team to evaluate the possible outcomes that are produced by different clinical scenarios. We conclude that with the aid of such a system, the experts can weigh the results of different treatments more efficiently and thereby increase treatment success and patient safety.

Keywords: Clinical Decision Support, Digital Patient Model, Head and Neck Cancer

1 Introduction

Model-based decision support systems can assist interdisciplinary teams in finding suitable and individualized treatments. Using Bayesian Networks (BN), decisions can be modeled based on guidelines, studies, and clinicians expertise [1]. Utilizing patient-specific data from the hospital information system (HIS), the resulting patient-specific model may support therapeutic decision making [2]. It can calculate unobserved data (e.g. therapy side effects) and propose suitable treatments with the most desirable outcomes. Our approach expands the previous work on model-based clinical decision support by Cypko et al. [3] with an exemplaric model for laryngeal cancer.

Oftentimes, physicians discuss possible different scenarios that might occur after the treatment was performed to weigh the suitability of the treatment. A digital model of the patient that integrates knowledge about all potential outcomes might assist in the recognition of the relation between the patient's situation, the chosen therapy and the desired outcome.

Simulation and imitation has always been used as a method for testing new scenarios, related to a potentially dangerous procedure, in a manner that eliminated the hazards of potentially unintended consequences for patients and harmful therapy procedures or side effects (e.g. in surgery or chemotherapy) [4]. Especially in complex diseases and therapy procedures with harmful features, like chemotherapy, we believe that the imitation of patients' characteristics and processes can increase the safety of treatment decisions.

2 Materials and methods

The prototypical model-based decision support system by Cypko et al. relies on the integration of patient-specific data. These data are transferred from the HIS into the model. For every patient, a specific individualized model is instantiated. It proposes the inferred therapy options and provides ways for interaction. It is enhanced with the proper medical knowledge that the inference algorithm is based on. However, it can only depict the calculations based on the real-life scenario. Additional functions for a comprehensive analysis of a case and safe, experienced decision making are needed [5, 6]. Useful features would be:

- presentation of inference results of most suitable treatment options
- imitation of certain patient entities (e.g. selecting different treatment options or quality of life-factors)
- triggering of the inference algorithm based on imitated data
- storing different imitation-scenarios to be discussed and evaluated by the interdisciplinary team

Based on these factors, we propose a concept of integrated patient data and simulated patient scenarios to properly evaluate an oncological case.

3 Results

We propose a two-layered approach. On the first layer, the existing patient-specific data is integrated into the decision model. Using the exemplaric decision model for laryngeal cancer, these would comprise of diagnostic data from radiology, pathology, clinical examinations and so on. Based on these data, *unobservable* entities will be inferred (e.g. suitable treatment options). On the second layer, to be able to imitate a patient's situation, it must be possible to enter different types of data. Not only the diagnostic (*observable*) entities must be entered to depict different scenarios. Also, the *unobservable* entities influence the inference algorithm.

When creating a simulated patient scenario for decision support, at first, all patient-specific data from the patient record are integrated into the model. Secondly, different entities can be changed according to the scenario that is created. This process comprises of the added (unobservable) values as well as changed values to observed variables that may have already been set by the data from the patient file. Hence, simulated entities override already set values from observed data. Simulated patient scenarios would comprise of the following data:

procedural data all observed data from electronic patient record

simulated data all manual entered entries, possibly overriding the procedural data

For a comprehensive understanding of a patient's situation, several different imitated scenarios would be necessary, next to the original patient-specific model, to compare possible procedures and outcomes. For example, in preparation of a upcoming tumorboard meeting, the attending physician imitates different treatment options by setting the information on the specific options and comparing the respective outcomes to assess the causal relations. He or she would then take the conclusions to present the case to the interdisciplinary team.

4 Discussion

Imitating patient data or clinical scenarios respectively does not replace the thorough studying of the clinical case. Decisions must be made based on clinical knowledge, evidences and experiences. But given complex oncological cases, where the variety of correlations and possible procedures are vast, the assessment of the next steps in the clinical pathway can be supported by providing information on different possible outcomes. However, it must be made clear that the scenario that is examined is a manually entered, imitated scenario. All reasoning steps must be provided with the respective medical knowledge that underlie the calculation. Data storage of simulated entities must be separated from the regular electronic patient file, so that no simulated data is wrongfully used in the clinical daily routine.

5 Conclusion

Imitation of clinical scenarios seems to be a promising method so support decision making. We will implement the proposed approach and integrate it into an existing decision support architecture. The system will be installed for study purposes in the University hospital Leipzig and evaluated with the medical experts from the ENT department.

6 References

[1] Cypko MA, Stoehr M, Oeltze-Jafra S, Dietz A, Lemke HU. *A Guide for Constructing Bayesian Network Graphs of Cancer Treatment Decisions.* Studies in Health Technology and Informatics 2017:1355-1355. doi:10.3233/978-1-61499-830-3-1355.

[2] Gaebel J, Cypko MA, Lemke HU. *Accessing Patient Information for Probabilistic Patient Models Using Existing Standards.* Stud Health Technol Inform. 2016;223:107-12.

[3] Cypko MA, Stoehr M, Kozniewski M, Druzdzel MJ, Dietz A, Berliner L, Lemke HU. *Validation workflow for a clinical Bayesian network model in multidisciplinary decision making in head and neck oncology treatment.* International Journal of Computer Assisted Radiology and Surgery. 2017:1–12.

[4] Moss J, Berner E. *Evaluating clinical decision support tools for medication administration safety in a simulated environment.* International Journal of Medical Informatics, 2015(84:5):3081–318,ISSN 1386–15056.

[5] Kneebone RL. *Practice, Rehearsal, and Performance:An Approach for Simulation-Based Surgical and Procedure Training.* JAMA. 2009;302(12):1336-1338. doi:10.1001/jama.2009.1392

[6] Doudican NA, Kumar A, Singh NK et al. *Personalization of cancer treatment using predictive simulation.* J Transl Med. 2015 Feb 1;13:43. doi: 10.1186/s12967-015-0399-y.

Towards the optimization of generator modulation and applicator positions for radiofrequency ablation of spine metastases

Caroline v. Dresky[1], *Hanne Ballhausen*[1], *Mathias Becker*[2], *Steffen Serowy*[2], *Martin Skalej*[2], *Christian Rieder*[1]

[1] *Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen, Germany*
[2] *Institute of Neuroradiology, Otto von Guericke University Magdeburg, Germany*

Contact: caroline.von.dresky@mevis.fraunhofer.de

Abstract We suggest a method for advanced generator modulation and electrode positioning for radiofrequency ablation. This method is based on a patient individual numerical simulation of the therapy. The treatment is simulated for various different applicator positions while ensuring that risk structures are protected by deactivating electrodes that cause too high temperatures. The optimal treatment is chosen by means of an evaluation formula weighting the therapy goals in a user preferred manner. This method shall support the forward planning of spinal metastases ablations.

Keywords: radiofrequency ablation, therapy planning, optimization of treatment parameters

1 Problem

Radiofrequency ablation (RFA) is a procedure option for the palliative treatment of painful spinal metastases based on local tissue heating. The goal is the complete coagulation of the metastases while protecting surrounding risk structures such as the spinal canal from overheating. To achieve this, electrode positions and ablation parameters have to be chosen carefully. Hence, a method for the determination of applicator position and generator modulation resulting in an optimal temperature distribution in the tissue is desirable. However, as this is a complex procedure, we present a method that improves the result of the standard therapy, but not necessarily yields the optimal solution. To this end, we developed a numerical simulation method that improves the standard generator modulation as well as applicator positions and predicts the treatment outcome.

In the last years, numerical methods to simulate the thermal necrosis including cooling effects (1, 2, 3) as well as fast approximations based on such simulations (4) have been developed to estimate the therapy outcome in particular for ablation of liver tumors. Recently, numerical simulations have been also presented for spine metastases (5, 6). To automate the access path determination, methods have been proposed which aim at the optimization of the percentage of destroyed tumor volume either by approximating the coagulation necrosis with ellipsoids (7) or by utilizing simulation methods (8). Also, multiple additional clinically relevant criteria such as distance to risk structures, path length, and angulation have been considered (9, 10, 11). A comprehensive discussion of related work is presented in (12).

2 Material and Methods

The proposed procedure is based on a numerical method developed for the patient individual simulation of the temperature distribution of RFA therapies. A screenshot of the simulation platform is shown in figure 1. The standard modulation for a multipolar RFA Celon generator is the switch between all possible electrode configurations considering the impedance of the tissue. In the case of two applicators comprising two electrodes each, the generator switches between six combinations. Due to the anatomical situation, the two applicators can not be placed in parallel on both sides of the spinal canal, for which reason this procedure does not necessarily result in the optimal necrosis shape. Therefore, we consider variations in the generator modulation. Moreover, we allow shiftings of the applicators. From the clinical perspective, there are only marginal variation possibilities in the access paths. Hence, we restrict the method to small variations of the applicator tips while the skin penetration point remains fixed.
The indicated variation options result in four different methods for the choice of generator modulation and applicator position, which will be discussed in some more detail in the following.

2.1 Standard generator modulation

To simulate the therapy outcome using the standard generator modulation, we consult data of real RFA therapies. We consider the applicator positions chosen by the clinician and switch the generator every two seconds to the next of the six possible electrode combinations.

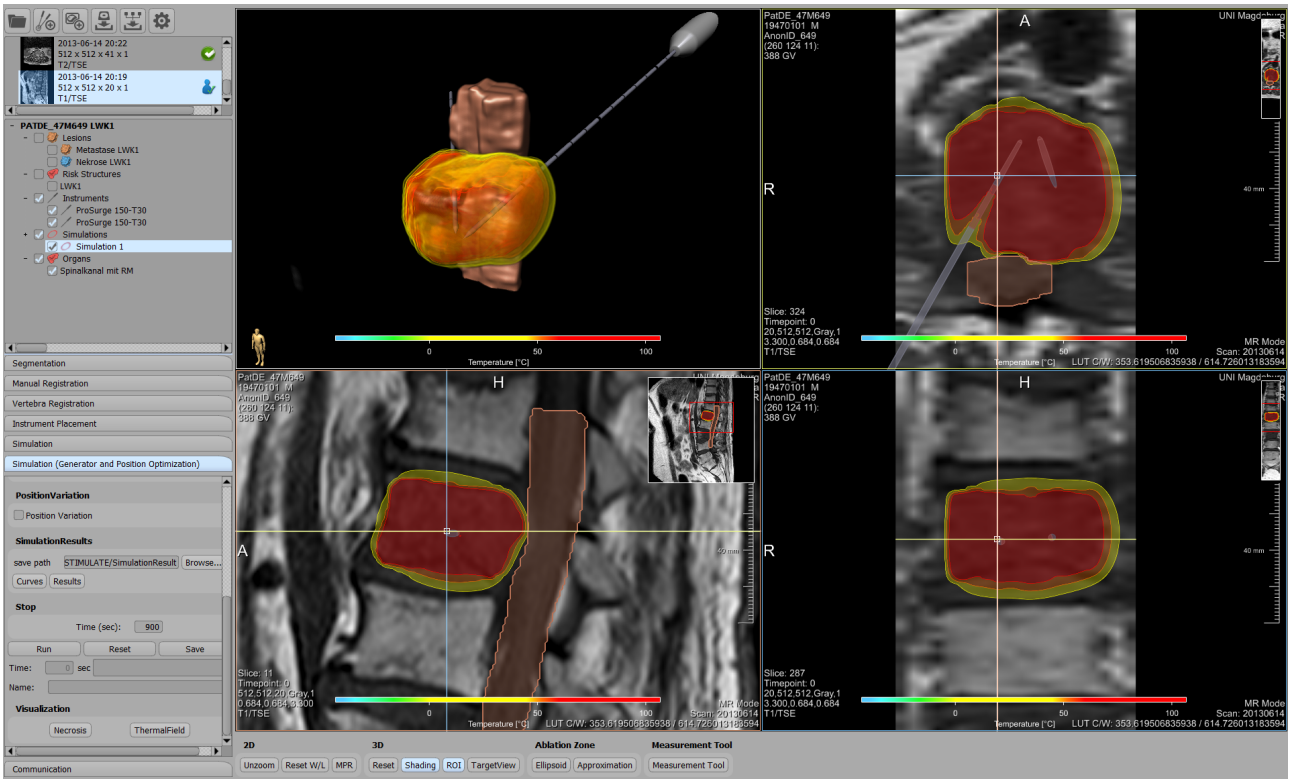


Figure 1: *Platform for the patient individual simulation and advanced parameter determination for radiofrequency ablation of spine metastases. The current viewer setting reveals the segmented spinal canal and the temperature field resulting from the simulation in 3D as well as in three orthogonal slices underlaid with the corresponding magnetic resonance images.*

2.2 Advanced generator modulation

We start the simulation using the standard generator modulation. In order to prevent risk structures from overheating, we check the maximum temperature in the respective tissue after each two seconds of simulation. In this context we focus only on the spinal canal as most important risk structure. If the temperature exceeds a predefined threshold temperature, the electrode, which is closest to the position of the maximum temperature, is turned off and with this all combinations involving the respective electrode will not be used anymore. The generator switches then between the remaining electrode combinations until the temperature in the risk structure rises further or the therapy duration is reached. In case the electrode next to the overheated tissue is already switched off, the second closest electrode is deactivated, and so on. If a predefined critical temperature is reached, the therapy is terminated immediately. This procedure, which is sketched in figure 2, ensures that risk structures are not harmed and moreover, it allows for the continuation of the treatment – far enough from the tissue at risk – although high temperatures are reached. The result of this advanced generator modulation procedure is a therapy plan consisting of a specific order of the six electrode combinations. The resulting treatment duration might be shorter than the aimed one.

2.3 Advanced applicator positioning

In order to allow for small variations in the electrode positions, we place a cube of appropriate size over each applicator so that the target position, i.e. the position in the middle of both electrodes, is in the center, and the lateral faces of the cube are all perpendicular or parallel to the shaft (see figure 3). Additional target positions are the center points of the six faces of the cube, so that we get seven positions for each applicator and hence 49 positions in combination.
Subsequently, we simulate the therapy for each of the 49 applicator positions using the standard generator modulation. For each simulation result, we compute an evaluation value, so that the highest score defines the best therapy outcome and hence the best applicator positions. The evaluation formula respects the following goals:

(i) The fraction of the metastasis, which is ablated (V_{met}^{abl}/V_{met}), shall be maximized.

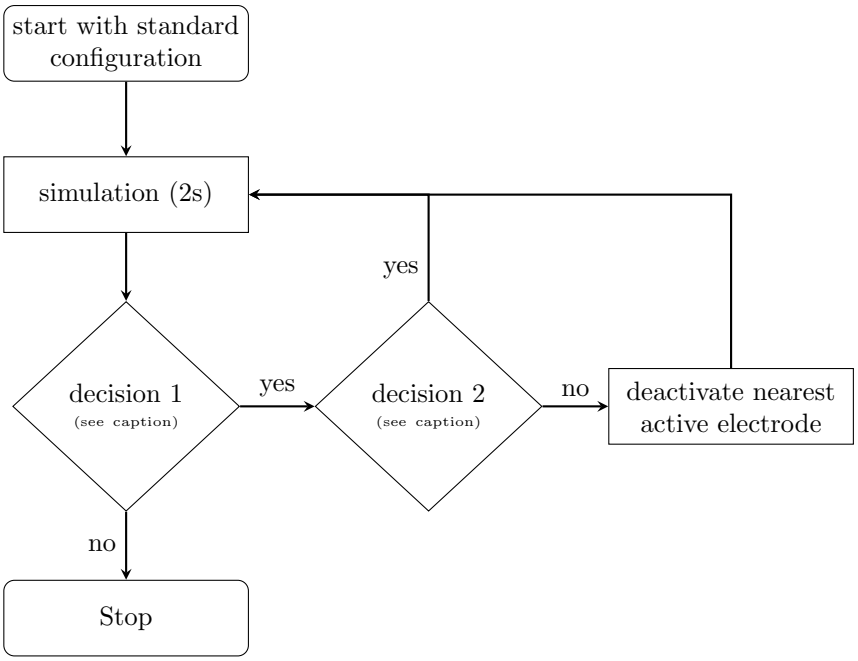


Figure 2: *Advanced generator modulation: In decision 1 it is tested whether the maximum temperature in the risk tissue is below the predefined critical temperature, whether at least one generator configuration is active and whether the simulation time already performed is less than the predefined ablation time. In decision 2 it is checked whether the maximum temperature in the risk material is below a predefined threshold temperature.*

(ii) The fraction of the total ablation, which is inside the metastasis ($V_{met}^{abl}/V_{total}^{abl}$), shall me maximized.

(iii) Risk structures shall be protected by ensuring that a predefined critical temperature is not exceeded.

The formula used in this context is a weighting of goals (i) and (ii), while applicator positions contradicting goal (iii) are not incorporated in the optimization step:

$$g(V_{met}, V_{met}^{abl}, V_{total}^{abl}) = f \cdot \frac{V_{met}^{abl}}{V_{met}} + (1-f) \cdot \frac{V_{met}^{abl}}{V_{total}^{abl}}, \qquad f \in [0,1]. \tag{1}$$

As can easily be seen, the evaluation value g is in the interval $[0,1]$, and the higher the score for a fixed factor f the better.

2.4 Advanced generator modulation and applicator positioning

The final method suggested here combines both approaches, namely the advanced applicator positioning and the advanced generator modulation. To this end we simulate the therapy once more for all 49 position combinations, as described in 2.3. However, in contrast to the simple applicator positioning we use the advanced generator modulation instead of the standard one for each therapy simulation. Hence, all 49 simulations meet goal (iii) and will be incorporated in the optimization step, which also uses function (1).

3 Results

We compared the simulation results of the four methods discussed above applied to real clinical data sets. More precisely, we used the planning MR images along with the segmentation of metastasis, vertebra, and spinal canal as well as the information on the applicator positions extracted from intraoperative images. Then, we performed patient individual simulations with all four methods for 900 s with a generator power of 50 W. We considered only one risk structure, namely the spinal canal, with a threshold temperature of 40 °C and a critical temperature of 40.1 °C. For the position variation we chose the side length of the cube to be 3 mm, and the evaluation factor in formula (1) is $f = 0.6$. The results for two exemplary patients are listed in table 1,

Figure 3: *Advanced applicator positioning: While the cross on the applicator marks the initial target position, the center points of the six lateral cube faces are the additional target positions used in the advanced applicator positioning procedure.*

where capital letters denote the rows for both patients and numbers indicate columns used for references in the following discussion. The columns comprise the outcomes of the four methods, namely

- the result of the standard procedure, supplemented by the intermediate result at reaching the threshold temperature in the spinal canal,
- the outcome of the advanced generator modulation, which is computed for the same applicator positions as the standard procedure, as indicated by the numbers in row A,
- the best output of the 49 simulations performed in the advanced applicator positioning procedure, and
- the result linked to the largest of the 49 evaluation values achieved in the advanced generator modulation and applicator positioning method.

4 Discussion

The standard generator configuration can yield high temperature in risk structures, as seen in cells B1 in table 1. In contrast, all three suggested methods restrict this temperature, which is evident in rows B. Instead of just terminating the therapy as soon as the threshold temperature is reached, which is shown in column 2, the advanced generator modulation method continues the treatment far enough from the tissue at risk. This can be understood by a comparison of G2 and G3. In this way we can reach larger ablation volumes and higher fractions of the metastasis that are ablated, which is seen by comparing C2 and C3 as well as D2 and D3. The reason that the evaluation value of the advanced generator modulation result can be worse than for the terminated standard procedure, as seen in H2 and H3 for patient 2, is the following: Due to the fact that we only consider the spinal canal as risk structure, the advanced generator modulation does not prohibit that even a large part of the ablation is created outside the metastasis in the surrounding tissue. Although the protection of this tissue is not part of the simulation, the evaluation formula takes the fragmentation of the ablation (inside/outside the metastasis) into account and hence penalizes the ablation outside the metastasis. However, in order to account for goal (ii) already in the simulation, future analyses could also define threshold temperatures for vertebra and vertebra surrounding tissue. The advanced applicator positioning process considers various different applicator positions in the neighborhood of the standard position and hence provides the opportunity to find a position yielding better therapy results than the standard modulation, as is evident from H1 and H4 for patient 2. However, as it is based on the standard generator modulation for the respective position, which does not protect risk structures, it is not guaranteed that this procedure yields an acceptable result at all, as is the case in column 4 for patient 1. In order to assure that we end up with an acceptable therapy plan, we combine the advanced applicator positioning with the advanced generator modulation, which is presented in column 5. As expected, this method yields the best results. It is related to the largest evaluation value, reported in rows H, and a therapy duration shorter than the standard one, which is seen in rows F.

5 Conclusion

The results indicate that the presented advanced generator modulation and applicator positioning method has the potential to create a patient individual therapy plan, which yields better results than the standard

treatment procedure. The risk tissue is protected, the ablation result might be more satisfying and in some cases the treatment duration can be shortened. In order to strengthen this indication, further validation based on phantom data as well as patient data is needed. Moreover, not only the spinal canal but also other tissue should be protected. If we defined threshold and critical temperatures for vertebra or surrounding tissue, the fraction of the ablation, which is outside the metastasis, could be further limited. So far, we concentrated only on the procedures but not on computing times, which are currently in the order of hours for the simulation of 49 therapies of 900 s. Therefore, it will be necessary to shorten this duration by means of code optimization as well as a sophisticated decision on the number of simulations and applicator positions to be taken into account in the clinical environment. Finally, it should be emphasized that the proposed method is a simplified optimization, which is based on the selection of the best result out of a limited number of cases. In further analyses, this procedure should be compared to a profound optimization method with respect to quality of the results as well as time required for the computation of a therapy plan.

Table 1: *Results of patient individual treatment simulations using the standard generator modulation, the advanced generator modulation, the advanced applicator positioning and the advanced generator modulation as well as applicator positioning method.*

		1	2	3	4	5
	characteristics	standard modulation	termination at threshold temperature	advanced generator modulation	advanced applicator positioning	advanced generator mod./applicator pos.
	automatic protection of risk structures		+	+	(+)[1]	+
	consideration of multiple applicator positions				+	+
	PATIENT 1					
A	applicator position	1	1	1	-	2
B	maximum temperature in spinal canal (in °C)	42.9	40.0	40.0	-	40.0
C	ablation volume (in cm^3)	29.5	2.9	22.9	-	23.3
D	fraction of metastasis, which is ablated (in %)	93.7	11.9	88.8	-	91.8
E	fraction of total ablation, which is in metastasis (in %)	43.0	55.7	52.4	-	53.6
F	therapy duration (in s)	900	126	606	-	606
G	therapy duration per configuration (in s)	150,150,150, 150,150,150	22,22,22, 20,20,20	182,22,22, 178,20,182	-	190,12,12, 190,12,190
H	evaluation value	0.734	0.458	0.742	-	0.765
	PATIENT 2					
A	applicator position	1	1	1	2	3
B	maximum temperature in spinal canal (in °C)	40.9	40.0	40.0	39.7	40.0
C	ablation volume (in cm^3)	34.8	28.7	34.6	35.0	29.4
D	fraction of metastasis, which is ablated (in %)	87.6	83.6	87.8	90.2	85.3
E	fraction of total ablation, which is in metastasis (in %)	66.1	76.4	66.7	67.9	75.9
F	therapy duration (in s)	900	534	900	900	518
G	therapy duration per configuration (in s)	150,150,150, 150,150,150	90,90,90, 88,88,88	212,90,88, 212,88,210	150,150,150, 150,150,150	86,88,86, 86,84,88
H	evaluation value	0.790	0.807	0.794	0.813	0.815

[1]The therapy simulation for each applicator position does not respect the protection of risk structures as it is the standard generator configuration for the respective position. However, the determination of the best simulation result by means of the evaluation formula (1) only considers simulations that do not yield too high temperatures in risk structures.

6 Acknowledgement

This work was funded by the Federal Ministry of Education and Research within the Forschungscampus STIMULATE under grant number 13GW0095A.

7 References

[1] T. Kröger, I. Altrogge, T. Preusser, P. L. Pereira, D. Schmidt, A. Weihusen, H.-O. Peitgen, *Numerical simulation of radio frequency ablation with state dependent material parameters in three space dimensions*, Proc. of MICCAI 2006, Springer, 380–388 (2006)

[2] C. Rossmann, F. Rattay, D. Haemmerich., *Platform for patient-specific finite-element modeling and application for radiofrequency ablation*, Visualization, Image Processing and Computation in Biomedicine, **1(1)** (2012)

[3] P. Weir, D. Reuter, R. Ellerweg, T. Alhonnoro, M. Pollari, P. Voglreiter, P. Mariappan, R. Flanagan, C. S. Park, S. Payne, E. Staerk, P. Voigt, M. Moche, and M. Kolesnik, *Go-Smart: Web-based computational modeling of minimally invasive cancer treatments*, E-Health and Bioengineering Conference (EHB), 1-4 (2015)

[4] C. Rieder, T. Kröger, C. Schumann, H. K. Hahn, *GPU-based real-time approximation of the ablation zone for radiofrequency ablation*, IEEE Trans Vis Comput Graphics, **17(12)** 1812–1821 (2011)

[5] J. Matschek, E. Bullinger, F. von Haeseler, M. Skalej, R. Findeisen, *Mathematical 3D modelling and sensitivity analysis of multipolar radiofrequency ablation in the spine*, Math Biosci, **284** 51-60 (2017)

[6] Z. Tian, T. Dong, Y. Cheng, J. Hu, Q. Nan, *A treatment planning of radio frequency ablation for spinal tumor*, Int J Comput Methods, Online Ready (2018)

[7] K. Trovato, S. Dalal, J. Krücker, A. Venkatesan, B. J. Wood, *Automated RFA planning for complete coverage of large tumors*, Proc. of SPIE 2009, 72610D-72610D-7 (2009)

[8] C.-C.R. Chen, M.I. Miga, and R.L. Galloway, *Optimizing electrode placement using finite-element models in radiofrequency ablation treatment planning*, IEEE Trans Biomed Eng, **56(2)** 237-245 (2009)

[9] C. Villard, C. Baegert, P. Schreck, L. Soler, A. Gangi, *Optimal trajectories computation within regions of interest for hepatic RFA planning*, Proc. of MICCAI 2005, Springer, 49–56 (2005)

[10] C. Baegert, C. Villard, P. Schreck, L. Soler, *Multi-criteria trajectory planning for hepatic radiofrequency ablation*, Proc. of MICCAI 2007, Springer, 676–684 (2007)

[11] C. Schumann, J. Bieberstein, C. Trumm, D. Schmidt, P. Bruners, M. Niethammer, R. T. Hoffmann, A. H. Mahnken, P. L. Pereira, H.-O. Peitgen, *Fast automatic path proposal computation for hepatic needle placement*, Proc. of SPIE, 76251J-76251J-10 (2010)

[12] C. Schumann, *Visualization and heuristic modeling for planning of minimally- and non-invasive tissue ablation*, PhD thesis, Jacobs University, Bremen (2007)

A haptic model for virtual petrosal bone milling

T. Eixelberger[1], T. Wittenberg[1], J. Perret[2], U. Katzky[3], M. Simon[1], S. Schmitt-Rüth[1], M. Hofer[4], M. Sorge[4], R. Jacob[5], F. Engel[5], A. Gostian[6], C. Palm[7], D. Franz[1]

[1] *Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany*
[2] *Haption GmbH, Aachen, Germany*
[3] *szenaris GmbH, Bremen, Germany*
[4] *Universitätsklinikum Leipzig, Leipzig, Germany*
[5] *Bundeswehrzentralkrankenhaus Koblenz, Koblenz, Germany*
[6] *Universitätsklinikum Erlangen, Erlangen, Germany*
[7] *Regensburg Medical Image Computing (ReMIC), Ostbayerische Technische Hochschule Regensburg, Regensburg, Germany*

Contact: thomas.eixelberger@iis.fraunhofer.de

Abstract

Virtual training of bone milling requires real-time and realistic haptics of the interaction between the "virtual mill" and a "virtual bone". We propose an exponential abrasion model between a virtual bone and the mill bit and combine it with a coarse representation of the virtual bone and the mill shaft for collision detection using the Bullet Physics Engine. We compare our exponential abrasion model to a widely used linear abrasion model and evaluate it quantitatively and qualitatively. The evaluation results show, that we can provide virtual milling in real-time, with an abrasion behavior similar to that proposed in the literature and with a realistic feeling of five different surgeons.

Keywords: virtual milling, virtual reality, physics, surgerical training, OR simulation, haptics

1 Problem

The surgical access to the middle ear through the petrosal bone (e.g. for CI insertion) requires considerable theoretical and practical skills of the surgeons. During their education, surgeons should be exposed to real patients to train the required skills, but the safety and the patient's well-being should nevertheless be ensured at all times. As an alternative method for education surgeons can use virtual surgery simulators.

The HaptiVisT project aims at immerse haptical and visual virtual surgery training for manipulating bones in the sense of virtual hand surgery [18] and petrosal bone milling [6]. With this system, surgeons can potentially train virtual petrosal bone milling as often as they need and with an increasing level of difficulty. For realistic virtual petrosal bone milling both visualization and haptics need to be real-time and realistic. Dreher et al. [9] compared different approaches for data visualization in the milling system. The framerate of the visualization has to be above 30 FPS and the haptic thread must run at least at 1000 Hz to be real-time. An interaction with the virtual bone (abrasion, collision, slip,...) that behaves like real bone milling, with respect to the correct amount of abrasion and a correct milling behavior for experienced surgeons, to be considered.

State of the Art. Agus et al. [11] have developed a model of bone erosion analytically. The erosion is a function of the applied force and the velocity. Petersik et al. [13] have used the Voxmap-Pointshell algorithm to calculate an appropriate force at each sampled point. Others use linear models for modeling the abrasion [14, 15]. The realism of virtual milling has been evaluated by Wong et al. [16], who focus on the differences between virtual milling and milling on an animal petrosal bone. Bryan et al. [12] give an overview over the necessary aspects of a surgical training simulator. We compare the amount of abrasion with Federspil et al. which have evaluated the best abrasion in the case of skullbase surgery [17].

2 Material and Methods

The HaptiVisT system for pertrosal bone milling has been designed for a sensorial learning experience (Figure 1, left). Haptic feedback (forces and torques) is delivered by the Virtuose 6D Desktop haptic device from Haption [1] (Figure 1, right). At its tip, we extended it with an aluminum bone mill model (Figure 2, left) for a realistic grasp on the instrument. Visual feedback is provided by an auto-stereoscopic 3D monitor from SeeFront [19], that has been mounted on top of a mirror. The mirror serves, to better mimic, the intraoperative use of an

OR microscope for real petrosal bone milling. Like the OR microscope the mirror occludes the hand of the practicing surgeon during the intervention and it prevents the eyes of the surgeon to converge to a specific point. The mirror also gives the possibility to extend the system to other scenarios. A touch screen (LG Electronics) provides the possibility to the practicing surgeon to change the position of the virtual petrosal bone, the viewing angle, and the type and size of the virtual bone mill. The complete system runs on a mobile workstation with an Intel i7 processor, 16GB RAM and a NVIDIA Quadro graphics card.

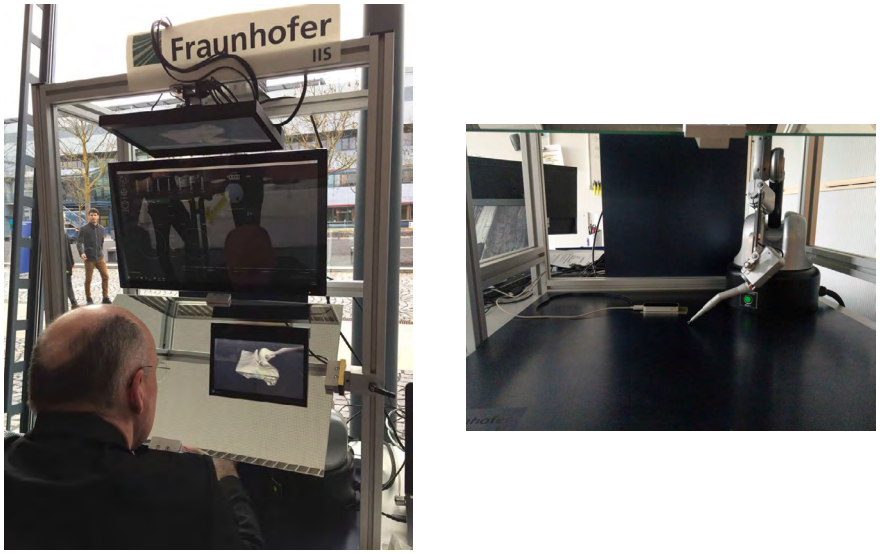


Figure 1: *Over- and detail view of the demonstrator, see text for details*

Because of performance reasons the visualization and the haptic simulation are running in separate threads [9]. Minimal requirement for the visualization is a frame rate of 30 FPS and for the haptic rendering is a haptic rate of 1000 Hz [4]. OpenGL and QT are used for visualization and the graphical user interface, respectively. chai3D [2], an open source C++ interface for haptics is used to communicate with the haptics device.

The virtual petrosal bone is based on computed tomography (CT) data and rendered as an iso-surface model via chai3d [5]. The DICOM files of the CT are loaded via gdcm [8] and the spacing is equal to 0.3 mm for the used test data. The virtual mill consists of two parts: the virtual mill bit and the virtual mill shaft (see Fig.2, right). With chai3D, we detect collisions between the virtual mill bit and the high resolution virtual petrosal bone and apply precomputed forces to the virtual mill bit. With the Bullet Physics Engine [3], collisions between the virtual mill shaft and a coarse model of the virtual petrosal bone are detected. The subdivision in virtual mill bit and shaft is necessary, because (1) in the Bullet Engine, the interaction of the mill bit with the virtual petrosal bone is computationally too intensive and (2) because chai3d does only allow spherical haptic points, like the mill bit, and does not allow for freeform models, like the mill shaft.

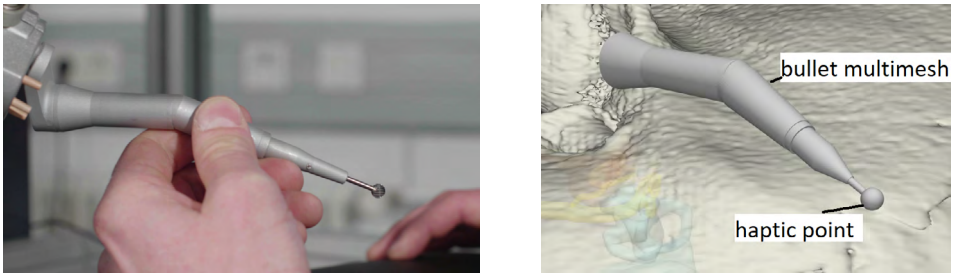


Figure 2: *Physical and virtual model of the mill.*

2.1 Mill bit: Abrasion simulation

We implemented two selectable types of bits. "blades" and "diamond", with seven eligible sizes each: 1.4 mm, 1.8 mm, 2.1 mm, 2.7 mm, 3.1 mm, 3.5 mm and 4 mm. The virtual abrasion is a function of the drillspeed and

combined with the force at the haptic point:

$$Abrasion = \begin{cases} 0 & \text{if } Speed = 0 \\ \exp\left(\frac{Speed}{k_a}\right) \cdot \frac{k_b}{10^7} \cdot Force & \text{else} \end{cases} \quad (1)$$

The force at the haptic point in Newton is computed via chai3d. It depends on the angle between the virtual mill and the surface normal of the petrosal bone at the contact point, the impact velocity, and the stiffness and friction of the virtual bone. *Speed* is the rotational speed of the mill bit in rpm (revolutions per minute). The modifiable parameter $k_a = 2875$ controls the influence of the *Speed* parameter, k_b equals 2 for the blades bit and 0.0002 for the diamond bit and biases the abrasion at low driving speed. We set the modifiable parameters heuristically with the feedback of an experienced surgeon (surgeon 1). The value of the abrasion is multiplied by a value between 0 and 1 for the sharpness and then subtracted from the alpha channel a of all interacting voxels (changing with the mill diameter). If $a \leq 0$ the regarding voxel will be deleted from the virtual petrosal bone and with that it has been virtually milled away.

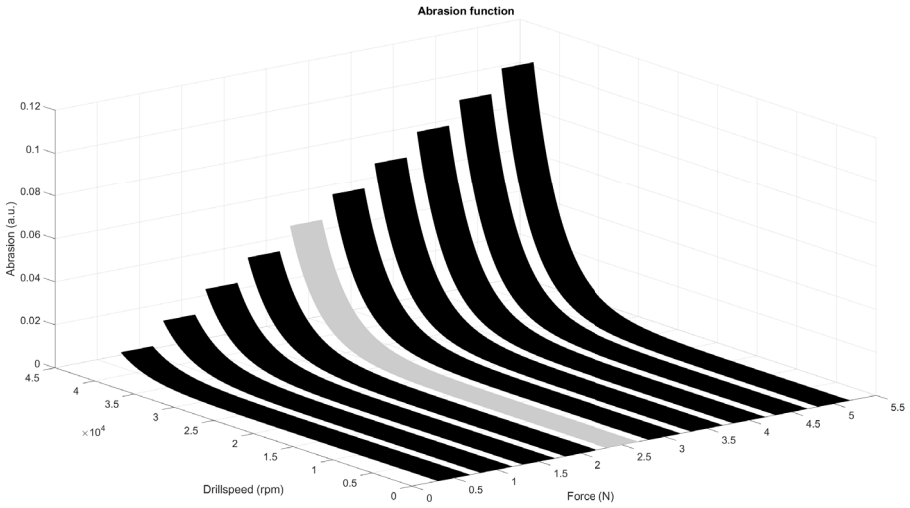


Figure 3: *Exponential abrasion model with* $Force = 0...5\ N$, $Speed = 0...40000\ rpm$, $k_a = 2875$ *and* $k_b = 2$.

2.2 Mill shaft: Collision detection

To ensure that the mill shaft does not clip into the petrosal bone during the milling process, a collision detection between shaft and bone is implemented using the Bullet Physics Engine. In the background of the visible petrosal bone the physics model is generated. This 3D model is a low resolution Bullet rigid body mesh [10] with an opacity level of zero, due to this it is a transparent colliding model at the same position as the visualized petrosal bone. If the mill shaft touches the Bullet model, a force and a torque are computed. The torque and force depend on the position of the mill, direction of the virtual mill, surface normal of the virtual petrosal bone, distance of the point of contact to the mill bit and stiffness and friction of the bone.

2.3 Evaluation of the virtual milling model

We evaluated our virtual milling model qualitatively regarding its real-time capability and the amount of abrasion and qualitatively regarding the realism of the haptics.

Quantitative evaluation. For an evaluation of the real-time capability of the model, we logged whether the haptic rate dropped below 1000 Hz. For an evaluation of the amount of abrasion, we observed the typical milling behavior of five surgeons. We determined typical *Force* and *Speed* values and compared the resulting abrasion to the work of Federspil et al. [17]. They stated an optimal abrasion for milling in the mastoid with a blades mill of 1 mm/s.

Qualitative evaluation. We did a small initial user study with five ENT surgeons (four experienced surgeons and one assistant doctor) which milled with the system for about half an hour each. They did tested the model with the focus on the haptics and visualization while milling into the petrosal bone. We presented the linear abrasion model of the literature to surgeon 1 to get a reference statement and asked for feedback about the exponential model until it met its requirements. After that we presented the exponential model to four additional surgeons and asked them to critically think aloud while they milled with the system. If they did not tell us something about their experience regarding one of the following stated features of the virtual milling model, we asked them explicitly.

- **Abrasion:** The amount of the abrasion.
- **Force:** The difference in haptic feedback for different forces applied to the mill bit.
- **Speed:** The difference in haptic feedback for different speeds of the mill bit.
- **Bit type:** The difference in haptic feedback for different types (blades, diamond) of the mill bit.
- **Diameter:** The difference in haptic feedback for different diameters of the mill bit.
- **Fall:** The haptic feedback, when the mill bit falls in a bone hole.
- **Slip:** The haptic feedback, when the mill bit slips of the bone.
- **Caught:** The haptic feedback, when the mill bit is caught behind a bone edge.

We organized the surgeons answers in categories ranging from - - (very negative) over - (negative), o (neutral), and + (positive) to + + (very positive). During both parts of the evaluation we locked the visual frame rate at 60 FPS. The subjects were able to adjust the drillspeed between 0 and 40000 rpm and the milling force ranges from 0 to 10 N.

3 Results

Quantitative evaluation. Figure 4 shows an example of the measured drillspeed, force, the abrasion of a linear model and of our exponential over the frames. Without bone contact the haptic rate is at about 30 kHz, with bone contact it never drops below 10 kHz.

Federspil et al. conclude that the optimal abrasion in the mastoid is 1 mm/s. We figured out that the mean force is by 2.5 N and the mostly used drillspeed is 40000 rpm. With this experience the abrasion in figure 3 is about 0.055. At a locked framerate of 60 FPS the decrementation of the alpha channel is 3.3 per second, so 3.3 voxels with an alpha channel of 1 are deleted per second. The spacing leads us to an abrasion of approximately 1 mm/s.

Qualitative evaluation Organized answers of the five surgeons are depicted in Table 1.

Table 1: *Evaluation results. - - very bad, - bad, o neutral, + positive, + + very positive*

Model	Surgeon	Abrasion	Force	Speed	Bit type	Diameter	Fall	Slip	Caught
Linear	Surgeon 1	- -[1]	o	- -	- -[2]	- -	-	-	+
Exponential	Surgeon 1	+ +[3]	+	+	+	+ +[5]	+[4]	+	+
	Surgeon 2	+ +	+	+ +	+	+ +	+	+	+ +
	Surgeon 3	+	+	+	+	+ +	+ +	+	+ +[6]
	Surgeon 4	+[7]	o	+	+	+ +	o[8]	+ +	+ +
	Surgeon 5	+ +	+ +	+[9]	+ +	+ +	+ +	+	+ +

[1] the abrasion is much too low [2] the diamond mill does more abrasion than the blades and that is wrong [3] the abrasion is correct [4] the feeling of falling into a bone hole is good [5] the difference between a small and a big diameter is very good [6] it is perfect. The feeling of being caught behind an edge is great [7] a bit too low [8] in the one direction it is good but in the other bad [9] good difference between the drillspeeds

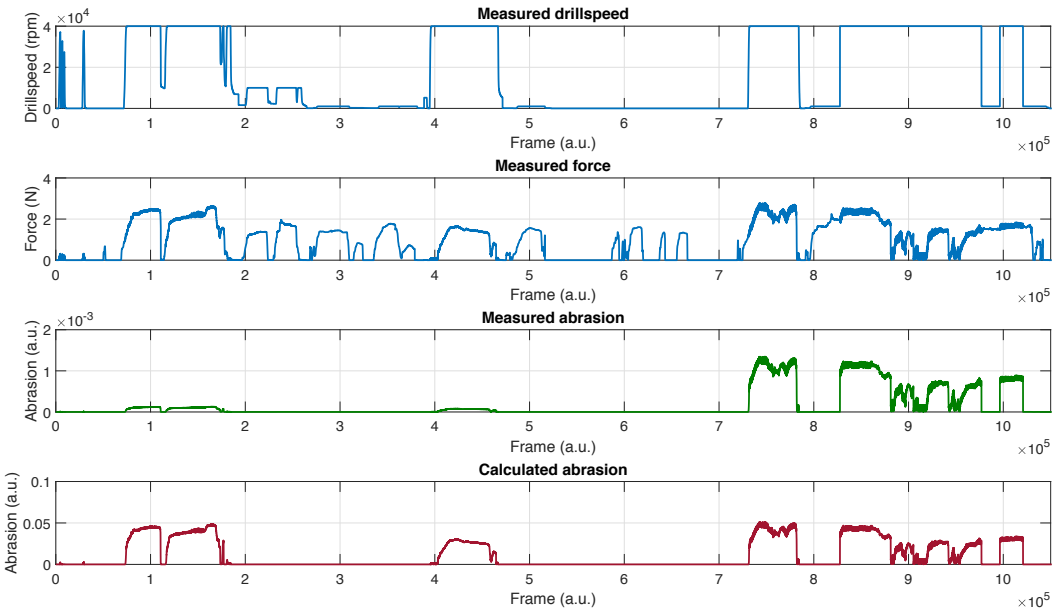


Figure 4: *From top to bottom: Measured drillspeed in rpm over frames, Measured force in N over frames, Abrasion calculated with a linear model over frames and Abrasion calculated with our exponential model over frames.*

4 Discussion

With a minimal haptic rate of 1000 Hz, our abrasion model met the requirement of real-time haptics. Also the abrasion at an experimentally determined mean $force = 2.5\ N$ and with $speed = 40000\ rpm$ (the preferred one of the surgeons) met the abrasion stated by Federspil et al. [17]. The comparison of linear and exponential model (Figure 4) shows that through the exponential weighting, the influence of the drillspeed on the abrasion is larger with our model than with the linear one. The five different surgeons stated that the abrasion model is close to reality, but not completely realistic. The most prominent features regard the collision detection: the haptics of falling into a bone hole or slipping over the surface of the bone needs further improvement. Also, the difference between the two types of the bit requires further work. While the interaction with bone of both types is not so different, the interaction with soft tissue is different for both types of mill bits. The blades mill damages soft tissue much faster and more severely than the diamond mill. In a further step we will integrate haptic interaction of soft tissue and mill bit in the system.

5 Conclusion

We describe a model for virtual milling of the petrosal bone that is both real-time and realistic. A quantitative evaluation of the amount of abrasion shows, that it is comparable to abrasion values given by the literature. Modifiable parameters of our abrasion models have been determined heuristically to fit virtual milling of the petrosal bone. Nevertheless, our model can be adapted quite easily to other types of tissue and mill bits. Our next step will be to include a haptic interaction between mill bit and soft tissue in the petrosal bone. With that we will model the risk structures, like facial nerve, dura and sigmoid sinus and than the practicing surgeon can experience haptically and visually the risk of damaging them.

6 Acknowledgement

This work was supported by the German Federal Ministry of Education and Research under the grant number 16SV7559.

7 References

[1] Haption GmbH, *Virtuose 6D Desktop Documentation*, last visited 06/2018, https://downloads.haption.com/marketing/datasheet/Datasheet-Virtuose6DDesktop-2018.pdf

[2] chai3d, *chai3d Documentation*, last visited 06/2018, http://www.chai3d.org/documentation/getting-started

[3] Bullet Physics Engine, *Bullet Physics Engine*, last visited 06/2018, http://bulletphysics.org/Bullet/BulletFull/

[4] Zhengyi, Y. and Y. Chen, *Haptic rendering of milling*, Mechanical Engineering Department, The University of Hong Kong, Hong Kong China, 2003

[5] Hansen, C. and C. Johnson, *Visualization handbook*, p. 7-11, Academic Press, 2004

[6] Franz, D., U. Katzky, S. Neuman, J. Peret, M. Hofer, M. Huber, S. Schmitt-Rüth, S. Haug, K. Weber, M. Prinzen, C. Palm, T. Wittenberg, *Haptisches Lernen für Cochlea Implantationen, Konzept - HaptiVisT Projekt*, p. 21 - 26, Proc. of CURAC, 2016

[7] Eriksson, M., H. Flemmer, and J. Wikander, *A haptic and virtual reality skull bone surgery simulator*, Proceedings of World Haptics, 2005

[8] gdcm, *gdcm Documentation*, last visited 06/2018, http://gdcm.sourceforge.net/wiki/index.php/Main_Page

[9] Dreher, M., D. Franz, M. Prinzen, M. Teßmann, C. Palm, U. Katzky, J. Perret, M. Hofer, T. Wittenberg, *CT-basiertes virtuelles Fräsen am Felsenbein. Bild- und haptische Wiederholfrequenzen bei unterschiedlichen Rendering Methoden.*, p. 176-181, Proc. of Bildverarbeitung für die Medizin, 2018

[10] Bender, J., K. Erleben, and J. Trinkle, *Interactive simulation of rigid body dynamics in computer graphics.*, Computer Graphics Forum, Vol. 33 No. 1, 2014

[11] Agus, M., Giachetti, A., Gobbetti, E., Zanetti, G., John, N.W., Stone, R.J, *Mastoidectomy simulation with combined visual and haptic feedback*, Proc. of Medicine Meets Virtual Reality, Newport Beach, 2002

[12] Bryan, J., D. Stredney, G. Wiet, D. Sessanna, *Virtual temporal bone dissection: A case study*, Proc. of IEEE Visualizaton, 2001

[13] Petersik, A., B. Pflesser, U. Tiede, K.H. Hohne, R. Leuwer, *Haptic volume interaction with anatomic models at sub-voxel resolution*, Proc. of IEEE Virtual Reality, 2002

[14] Agus, M., A. Giachetti, E. Gobbetti, G. Zanetti, A. Zorcolo, *Real-time haptic and visual simulation of bone dissection*, p. 110-122, Presence: Teleoperators & Virtual Environments 12.1, 2003

[15] Arbabtafti, M., M. Moghaddam, A. Nahvi, M. Mahvash, A. Rahimi, *Haptic and visual rendering of virtual bone surgery: A physically realistic voxel-based approach*, p. 30-35, IEEE Haptic Audio Visual Environments and their Applications, 2008

[16] Wong, D., B. Unger, J. Kraut, J. Pisa, C. Rhodes, J. Hochman, *Comparison of cadaveric and isomorphic virtual haptic simulation in temporal bone training*, Journal of Otolaryngology-Head & Neck Surgery 43:31, 2014

[17] Federspil, P., B. Plinkert and P. Plinkert, *Experimental robotic milling in skull-base surgery.*, Computer Aided Surgery 8.1, p. 42-48, 2003

[18] Maier, J., M. Huber, U. Katzky, J. Perret, T. Wittenberg, C. Palm, *Force-feedback assisted bone drilling simulation based on CT data*, p. 291 - 296, Proc. of Bildverarbeitung für die Medizin, 2018

[19] SeeFront GmbH, *SeeFront SF3D-133CR Documentation*, last visited 06/2018, https://www.seefront.com/fileadmin/content/downloads/SF3D-133CR_0.96.pdf

Vorträge CURAC / CTAC

Universal surgical display

D. Wilhelm[1,3], R. Stauder[2], D. Ostler[3], H. Feussner[1,3]
[1]*Klinikum rechts der Isar, Technische Universität, München, Deutschland*
[2]*CAMP, Technische Universität München,* ***München, Deutschland***
[3]*Arbeitsgruppe MITI, Technische Universität München, München, Deutschland*

Kontakt: dirk.wilhelm@de

Abstract

Aufgrund der hohen Komplexität des chirurgischen Operationssaals und der zunehmenden Informationsflut im klinischen Alltag ist eine selektive, situativ angepasste Darstellung medizinischer Inhalte im hohen Maße wünschenswert. Das „Universal Surgical Display" wird in der Bestrebung entwickelt, dem Chirurgen während einer Operation nur die Informationen zu visualisieren, die für den aktuellen operativen Teilschritt erforderlich sind, bzw. die eine Auswirkung auf den unmittelbaren Arbeitsablauf besitzen.

Keywords: Workflow, Display, Autonomie, OP Integration

1 Problemstellung

Der klinische Alltag ist u.a. gekennzeichnet durch eine fortwährende Informationsflut, die sich vor allem im chirurgischen OP negativ bemerkbar macht. Hier muss der Chirurg, während er eine komplexe Operation durchführt, zeitgleich eine Vielzahl von Problemen bearbeiten (Weiteres Vorgehen bei der Operation, Kontrolle der Systemparameter, OP Planung, Koordination von Anrufen, Terminplanung, etc.). Durch diese Mehrfachbelastung sind der Erfolg der Operation und der Patient gefährdet. Ein Instrument, welches selektiv und an die aktuelle OP Situation verfügbare Informationen filtert und in aufbereiteter Form dem Chirurgen präsentiert, wäre in hohem Maße wünschenswert.

2 Material und Methoden

Im Rahmen der DFG Forschergruppe Plafokon entwickeln wir derzeit ein universelles chirurgisches Display, welches an das Workflow-Modul des Experimental OP angebunden über den aktuellen OP Ablauf informiert ist und situativ angepasst verfügbare, aber auch erforderliche Informationen visualisiert. Das Display ist an das OP Modul (Laparoskopie) angebunden, kann aber auch Bildgebungen, Daten des KIS, Workflow-Daten und context-unabhängige Informationen (z. B. Telefonie) darstellen.

3 Ergebnisse

Das Universal Surgical Display ist aktuell noch in Entwicklung. Erste Funktionsmuster werden derzeit evaluiert, um die Selektion der verfügbaren Informationen anzupassen und die grafische Darstellung zu optimieren. Die Funktionalität umfasst neben dem OP Feld die Darstellung von System- und Workflowparametern, weitere Module sind in Vorbereitung.

4 Diskussion

Die Arbeitsumgebung, nicht nur im medizinischen Umfeld, ist überaus belastend und gekennzeichnet durch eine kontinuierliche Überladung mit ungefilterten Informationen. Um den Chirurgen wieder die Ruhe zu geben, die er für die Durchführung eines operativen Eigriffes benötigt, zeitgleich aber diese Informationen bereit zu stellen, die für die Ausführung seiner medizinischen Tätigkeit erforderlich sind, verlangt die Einbindung des operativen Workflows und eine intelligente Selektion auf Basis derselben.

5 Zusammenfassung

Das universelle chirurgische Display ist ein erster Ansatz für die selektive, situativ angepasste Informationsvisualisierung im chirurgischen OP und ein erfolgsversprechendes Instrument für die Erleichterung des chirurgischen Alltags.

Referenzen

Keine

Markerless Endoluminal Navigation (Project BIOPASS)

Deep Learning Based Detection of Intestinal Segments for Colorectal Endoscopic Investigations

N. Kohn[1], J. Fuchtmann[1], J. Morandell[1], D. Ostler[1], D. Wilhelm[1,2], H. Feußner[1,2]

[1]Research Group for Minimally Invasive Interdisciplinary Therapeutical Intervention at the University Hospital rechts der Isar of the TU Munich, Trogerstr. 26, 81675 München

[2] Department for Visceral Surgery at the University Hospital rechts der Isar of the TU Munich, Ismaninger Str. 22, 81675 München

Contact: nils.kohn@tum.de

Abstract

Position estimation within the large intestine during colonoscopies still depends significantly on the physician's capabilities and experience. To provide assistive information, a location prediction framework, relying on informative frame filtering and a deep learning approach, was developed. Also, for live utilization during clinical procedures, a setup comprising an embedded platform for real-time artificial intelligence applications as well as temporal smoothing and visual prediction feedback was implemented. The project indicates that image pre-processing and previous filtering of uninformative video frames is of greater importance than the architecture of the neural net itself. While the overall accuracy of the model has yet to be improved, major differences between the detectability of different colon sections have been noticed.

Keywords: Colonoscopy, Position Detection, Deep Learning, Large Intestine

1 Problem

Colonoscopic procedures require physicians to get familiar with uncommon perspectives from the inside of the human body. A monoscopic camera restricts the view and most importantly impedes orientation via anatomic landmarks. Knowing the exact position of the endoscope's camera-head is however crucial for executing an accurate colonoscopy and to ensure the patient's safety [1]. Unfortunately, the inner structures of the large intestine show great similarities across different areas and therefor position estimation is currently only based upon the experience of the physician. A system to provide assistive location information without an additional tracking device can only rely on the video stream. To overcome the obstacle of finding representative structures the project uses informative frame filtering and deep learning in order to automatically detect meaningful features for determining the current position of the tip of the endoscope within the large intestine.

2 Material und Methods

Twenty individual colonoscopies have been recorded at Klinikum rechts der Isar, Munich, and serve as a data set for training the deep neural network. Videos were captured in 1080i with either 25 or 50 fps. Since the original image comprises patient information as well as signals of an electromagnetic shape tracking device, the image was cropped to the relevant area showing only the camera signal. In all videos 11 predefined anatomic landmarks were identified by an experienced endoscopist. To ensure high quality of the annotation, each colonoscopy was annotated two times. Extracting the frames from the videos resulted in a data set with more than 900000 images composed of 11 different classes.

Because foreign substances (i.e. stool, cleansing agent), out of focus or motion blurred frames often occur, only a fraction of colonoscopic images can provide information relevant to the physician [2]. Ground truth data quality constraints get even more pivotal when machine learning techniques should be applied. To evaluate the informativeness of images within the generated data set, the informative frame filtering approach of Armin M. A. et al. [3] was implemented. Each frame is therefore represented in HSV color space and methods like edge detection or the Kanade Lucas Tomasi tracker are used to extract features (Fig. 1). Finally, a random forest classifier that was trained by a manually selected set of 300 images, divides the data into informative and non-informative frames.

For the position estimation of the endoscope-tip itself, different commonly used architectures of convolutional neural networks such as VGG16 [4] have been tested for classifying individual images into 9 predefined classes of colon sections (anus external, rectum, sigmoid colon, descending colon, left flexure, transverse colon, right flexure, ascending colon and caecum) and one additional class for non-informative frames. Available videos were split for reliable assessment of the model into training-, validation- and test- sets in the ration of 12:4:4. Transfer learning and finetuning using image augmentation and various optimizer were assessed, while starting with initial weights from ImageNet [5].

To use the framework in a clinical environment and present the final position estimation in an ergonomic way, the neural network was deployed on a Nvidia Jetson TX2 - an embedded platform for real-time artificial intelligence applications. Colonoscopic video signals are captured by a frame-grabber and feet into the trained deep learning model. For increasing the reliability of the system, temporal smoothing and an output logic, that increases weights on results which align with the anatomical sequence of colon sections, were added as a postprocessing step. Non-informative frames were rejected for position calculation. The estimated position is then graphically displayed next to the camera image (Fig. 2 d).

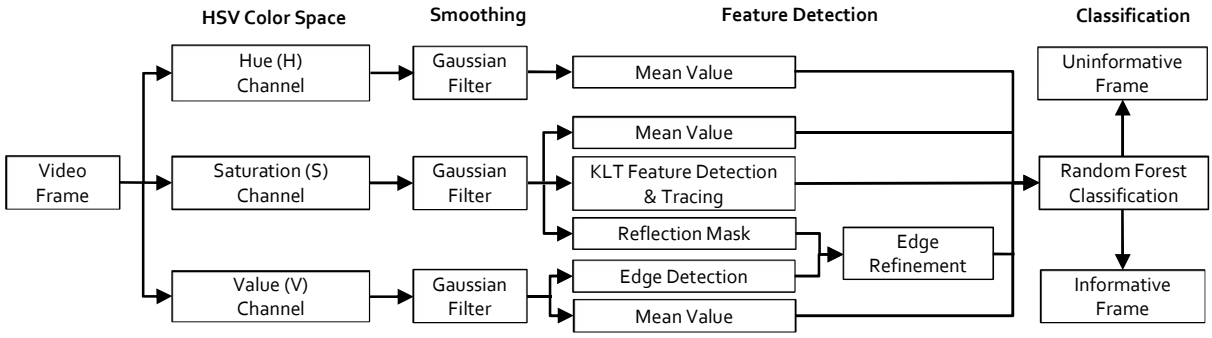


Figure 1: *Informative frame filtering as a data preprocessing step*

3 Results

Using informative frame filtering, the number of 920000 extracted frames was reduced and divided into two subsets: subset A containing 195000 frames and subset B containing 103000 frames (Fig. 2 a, b, c).

Subset A contains images of the intestine wall and lumen while subset B mainly consists of pictures showing only lumina. Even though subset B reduces the number of images of subset A by nearly 50%, the overall accuracy of the VGG16 network was increased by 21%. While sections like the caecum, anus external, rectum or ascending colon were identified reliably, others such as the terminal ileum, left/ right flexure or transverse colon turned out to be more challenging to classify.

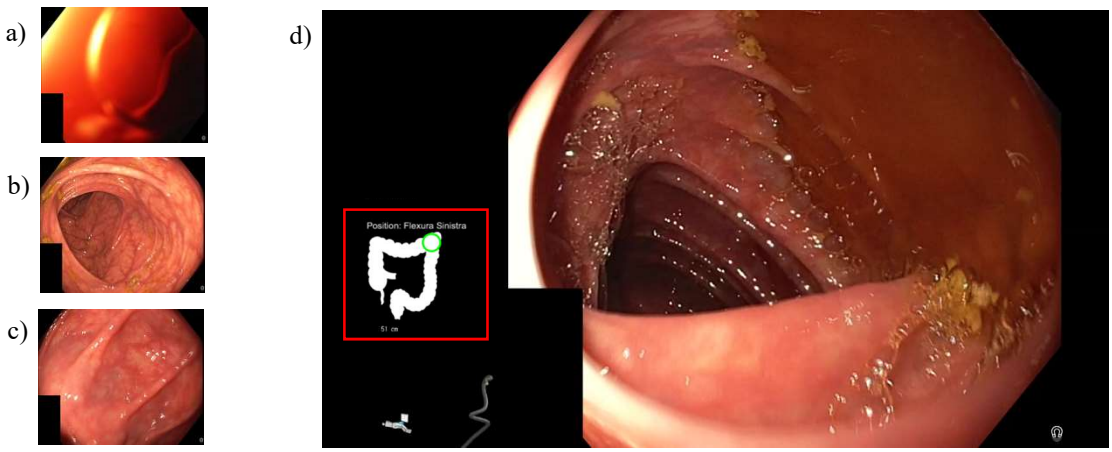


Figure 2: *a) uninformative frame, b) informative frame included in dataset A and B, c) informative frame included only in dataset A, d) visual presentation of the prediction*

4 Discussion

The significant improvement in the overall prediction accuracy between the two data subsets indicates that training data selection based on informativeness and preprocessing steps are essential in developing solid prediction models for colonoscopies. Yet, a quantitative definition of “informative frame” in the meaning of being beneficial for training a deep-learning based classifier remains elusive. Misclassification in certain classes may also be explained by the lack of training data due to the frame filtering and noise in the ground truth data, as clear transition points of colon sections are even for medical experts not uniquely identifiable. Image pre-processing and previous filtering of uninformative video frames seems to be of greater importance than the architecture of the neural net itself. However, results can be improved in the postprocessing by averaging predictions over multiple frames and restricting the output to anatomically valid correlations.

5 Conclusion

The work shows that providing assistive position information using live classification during colonoscopies is feasible even though the algorithm still has to be refined. Although data set quality shows strong influence, overall classification accuracy varies strongly between different colon sections. With most of the captured fames being uninformative for the location detection, deep-learning based models can only use a scarce set of images while making online predictions during colonoscopies.

6 Acknowledgement

This work was done with grant support by the Federal Ministry of Education and Research (BMBF).

References

[1] Kirschniak A., Granderath F.A., et al., *Endoskopisch-Chirurgische Anatomie*, Introduction / 2-6, 978-3 642 04732-9, Springer, Berlin, Heidelberg (2010)

[2] An Y. H., Hwang S., et al., *Informative-frame filtering in endoscopy videos*, Medical Imaging 2005, San Diego (2005)

[3] Armin M. A., Chetty G., et al., *Uninformative frame detection in colonoscopy through motion, edge and color features*. International Workshop on Computer-Assisted and Robotic Endoscopy, Munich (2015)

[4] Simonyan K., Andrew Z, *Very deep convolutional networks for large-scale image recognition*, arXiv preprint arXiv:1409.1556 (2014)

[5] Russakovsky, O., Deng, J., et al., *Imagenet large scale visual recognition challenge*, International Journal of Computer Vision, 115(3), 211-252. (2015).

Tumor- und Gewebeklassifikation mittels hyperspektralem Imaging

René Thieme[1], Claire Chalopin[2], Hannes Köhler[2], Marianne Maktabi[2], Nada Rayes[1], Rafael Beck[3], Andreas Dietz[3], Stefan Niebisch[1], Nico Holfert[1], Jonathan Philip Takoh[1], Sebastian Murad Rabe[1], Matthias Mehdorn[1], Thomas Neumuth[1], Andreas Melzer[1], Boris Jansen-Winkeln[1], Ines Gockel[1]

[1] Klinik für Viszeral-, Transplantations-, Thorax- und Gefäßchirurgie, Universitätsklinikum Leipzig, AöR
[2] ICCAS (Innovation Center Computer Assisted Surgery), Universität Leipzig
[3] Klinik für Hals-, Nasen- und Ohrenheilkunde, Universitätsklinikum Leipzig, AöR

Kontakt: rene.thieme@medizin.uni-leipzig.de

Keywords: Tumorklassifikation, Gewebeklassifikation, Hyperspektralem Imaging (HSI)

1 Problemstellung

Das hypersektrale Imaging (HSI) ist ein innovatives Bildgebungs-Verfahren. Durch die Anwendung einer speziellen HSI-Kamera, werden die Spektroskopie und die Bildgebung technisch kombiniert. Dieses Verfahren soll dazu genutzt werden, um zum einen Gewebe zu charakterisieren und dessen spektrale Eigenschaften zu erfassen und zum anderen soll es Tumore vom umliegenden gesunden Gewebe abgrenzen. Diese Technik ist nicht nur auf vitale Gewebe und Tumore begrenzt, sondern kann auch dazu dienen, Zellen in histologischen Präparaten zu erkennen und zu klassifizieren. Ziel wird es sein, dass mathematische Algorithmen erstellt werden die eine automatisierte Gewebeerkennung zulassen.

2 Material und Methoden

Mittels einer HSI-Kamera werden Spektren zwischen den Bereichen des visuellen Lichts und nahen-infrarot Bereich an histologischen Präparaten aufgezeichnet (500-995nm). Für die hier gezeigte Machbarkeitsstudie wurden ein Kollektiv von gesundem ösophagealen Plattenepithel (n=22) und ösophagealen Adenokarzinomen (n=45) aufgenommen. Bei den histologischen Präparaten handelt es sich um Resektate von ösophagealen Adenokarzinomen, die mittels Formaldehyd fixiert, in Paraffin eingebettet, in 3µm dünne Schnitte geschnitten wurden und mittels Eosin und Hematoxilin angefärbt wurden.

3 Ergebnisse

In allen n=45 untersuchten Gewebeschnitten konnten Zellen eines ösophagealen Adenokarzinoms detektiert und mittels HSI-Kamera aufgezeichnet werden. In n=22 der untersuchten Präparate befanden sich auch Anschnitte von gesundem unverhorntem Plattenepithel. Beim Vergleich des spektralen Verhaltens des Plattenepithels und der Adenokarzinomzellen konnten Unterschiede in der Absorption im Wellenlängenbereich von 500-700nm aufgezeigt werden. Die Intrvarianz der untersuchten Präparate hinsichtlich der Eigenschaften von Plattenepithel und Adenokarzinomzellen war gering und wies eine geringe Streuung auf.

4 Diskussion

Areale mir unverhorntem Plattenepithel und Zellen eines ösophagealen Adenokarzinoms weisen spezifische spektrale Charakteristika auf. Diese Merkmale können dazu dienen mittels Computer-assentierter Algorithmen eine automatische Diskriminierung auf das Vorhandensein eines Adenokarzinoms in histologischen Präparaten zu genieren, um eine Entscheidungshilfe bei der histopathologischen Diagnose zu liefern.

Intraoperative Beurteilung grenzwertig perfundierten Darms im Rahmen der akuten Mesenterialschämie mittels Hyperspektralbildgebung: Eine Falldarstellung

M.Mehdorn[1] , S.Murad Rabe[1], H.Köhler[2], M.Maktabi[2], J.Takoh[1], T.Neumuth[2], A.Melzer[2], Y.Moulla[1], B. jansen-Winkeln[1], C.Chalopin[2], I.Gockel[1]

[1]*Klinik undPoliklinik für Viszeral-, Transplantations-, Thorax- und Gefäßchirurgie, Universitätsklinikum Leipzig, Leipzig, Deutschland*
[2]*Innovation Center Computer Assisted Surgery (ICCAS), Universität Leipzig, Leipzig, Deutschland*

Kontakt: SebastianMurad.Rabe@medizin.uni-leipzig.de

1 Einleitung

Die akute Mesenterialischämie ist ein häufiges Krankheitsbild in alten und kritisch kranken Patienten, welche mit einer hohen Morbidität und Mortalität einhergeht (1). Therapieziele sind die Revaskularisation und die Resektion der ischämischen Darmabschnitte. Dabei ist die intraoperative Beurteilung der Perfusion in der Notfallsituation nicht immer eindeutig.
Hyperspektral-Imaging (HSI) ist dabei ein neues, intraoperatives Bildgebungsverfahren, welches nicht-invasiv die Gewebeperfusion objektivieren kann. Das untersuchte Gewebe wird mit Licht im visuellen und nah-infraroten Spektrum beleuchtet und das reflektierte Licht gemessen.

2 Ziele

Objektive Kriterien zur Beurteilung der Darmperfusion sind hilfreich, um das notwendige Resektionsausmaß zu bestimmen und das Outcome zu verbessen. Unser Ziel war dahingehend, HSI auf Machbarkeit zu untersuchen.

3 Methodik

Wir berichten den Fall einer 52-jährigen Patientin mit generalisierter Arteriosklerose im Multiorganversagen mit hohem Katecholaminbedarf, bei der sich eine multisegmentale Darmischämie bei chronischem Verschluss des Truncus coeliacus zeigte. Die bereits gangränösen Abschnitte wurden mittels Dünndarmteilresektion sowie erweiterter Hemikolektomie links mit anteriorer Rektumresektion unter Blindverschluss der Darmenden versorgt. In der programmierten Re-Laparotomie zeigte sich das gesamte Intestinum nekrotisch. Die Patientin verstarb im Multiorganversagen. Mittels HSI-Kamera wurde der intraoperative Dünndarm und das Operationspräparat untersucht und folgende Parameter erhoben: Gewebeoxygenierung, Nahinfrarot-Perfusions-Index und Gewebe-Wasser-Index.

4 Ergebnis

Nekrotische Darmabschnitte im Vergleich zu nicht-nekrotischen Abschnitten zeigten eine typische Bande des Absorptionsspektrum bei einer Wellenlänge von 630 nm (Abb.1). Erwartungsgemäß zeigten sich die Indices für Gewebeoxygenierung und Perfusion erniedrigt (Abb.2). Der Gewebe-Wasser-Index war in den minderperfundierten Dünndarmabschnitten erhöht. Korrespondierend zeigte sich eine Absorptionsbande bei 960 nm. Auch die Werte der Gewebeoxygenierungs- und Nahinfrarot-Perfusions-Indexen sind niedriger als im gesunden Darm.

5 Schlußfolgerung

Mittels intraoperativer Hyperspektralbildgebung ist es möglich, die Perfusion grenzwertig perfundierter Darmabschnitte bei akuter Mesenterialischämie zu objektivieren. Die Methode ist einfach, schnell und nicht-invasiv. Zukünftig könnten computer-assistierte Verfahren nekrotischer und kritischer Darm aus den beschriebenen spektralen Merkmalen automatisch erkennen, um den Chirurg bei der Resektion zu unterstützen.

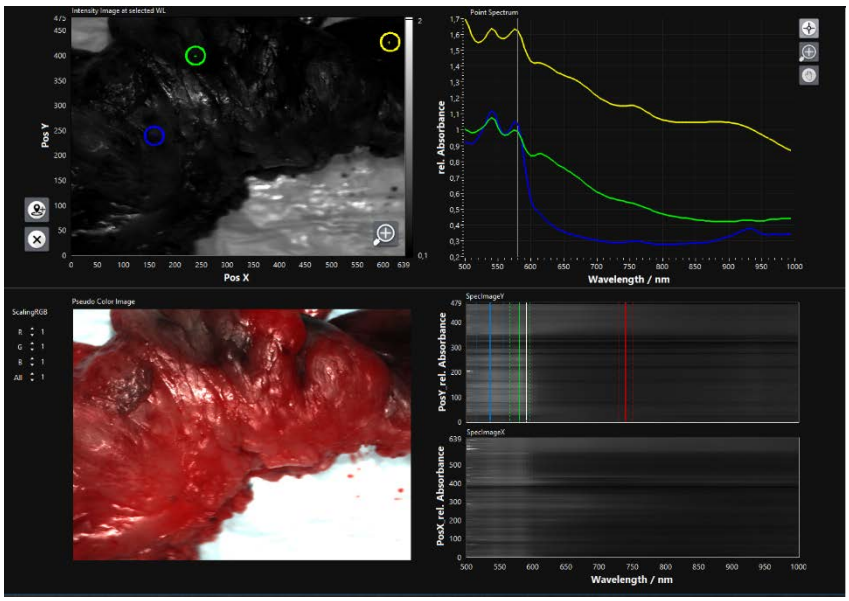


Abb 1.:*Dickdarmresektat mit Absorptionsbande bei einer Wellenlänge von 630 nm für nekrotische Abschnitte (grün: nekrotisches Gewebe; blau: nicht-nekrotisches Gewebe)*

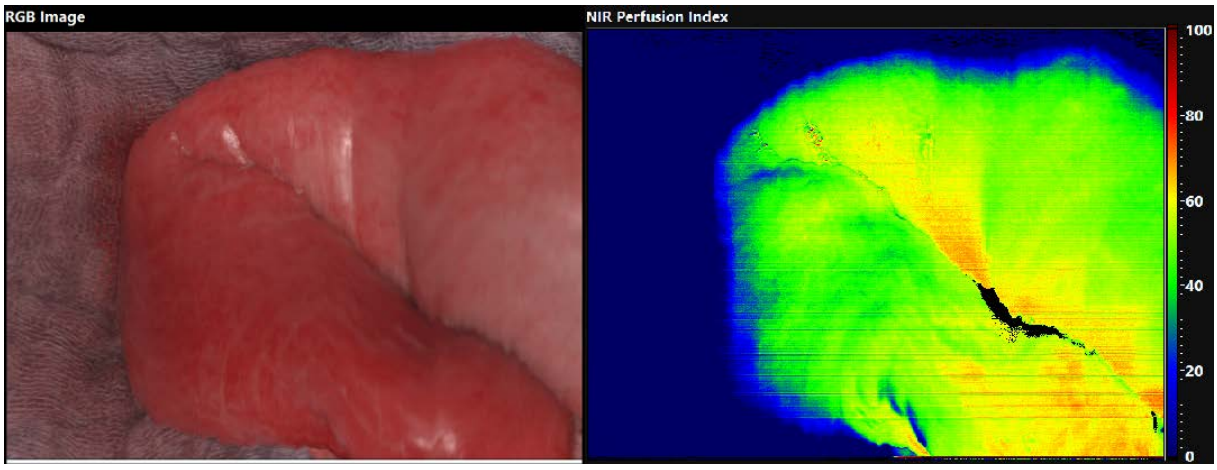


Abb 2.: *Intraoperatives Bild des minderperfundierten Dünndarmabschnittes (links RGB-Bild, rechts Nahinfrarot-Perfusion-Index)*

Referenzen

1) Acute mesenteric ischemia (part I) - Incidence, etiologies, and how to improve early diagnosis. Kärkkäinen JM, Acosta S. Best Pract Res Clin Gastroenterol. 2017 Feb;31(1):15-25

3D-Druck und Navigation - Auf dem Weg zur strahlungsfreien Stentgraftimplantation in der Aorta - Das Nav EVAR-Projekt

3 D-Printing and Navigation – The Way to a X-ray-free implantation of Aortic Stent grafts – The NAV EVAR Project.

Markus Kleemann

[1]*Division of Vascular and Endovascular Surgery, Department of Surgery, UKSH Campus Lübeck, Germany*

Kontakt: Markus.Kleemann@uksh.de

1 Introduction

Due to the continuous medical development of minimal invasive treatments, the endovascular aortic aneurysm repair (EVAR) has become an established procedure in the clinical routine. This procedure still requires the application of X-rays and contrast agent. The Nav EVAR project is a new approach with a promising potential for treatment of abdominal aortic aneurysms without using harmful ionizing radiation and contrast agent.

2 Material und Methods

The Nav EVAR procedure uses a real-time navigation. Based on the preoperative CT-scan a 3-dimensional endoluminal view of the patient vascular system during the treatment is enabled (Fig. 1). The navigation of the catheter through the patient body can be realised with a Bragg fibre which is implemented into the stent system. The change of the reflection spectrum through the fibre is used to determine the exact position and orientation of catheters and stentgrafts inside the aorta. Furthermore, optical coherence tomography can provide a continuous local feedback to verify the position of the catheter inside the vessel. During the video demonstration the catheter position is displayed with a mixed reality headset as a 3-dimensional projection inside a virtual angioscopy of the vascular system. Additionally, desired information like the preoperative CT-scan or the OCT data can be visualized on a virtual screen.

3 Results

The presented navigation system of the Nav EVAR project enables the virtual 3-dimensional real-time navigation for endovascular treatments. First 3-dimensional navigation inside a vascular phantom and visualization with a mixed reality headset has been realised.

4 Conclusion

Nav EVAR is a disruptive method for treatment of abdominal aneurysms. Due to the possibilities for visualization and navigation it may reduce the usage of harmful radiation and contrast agent. Further experimental studies on a phantom with patient specific 3D-printed abdominal aneurysm are planned to verify the accuracy of the navigation system.

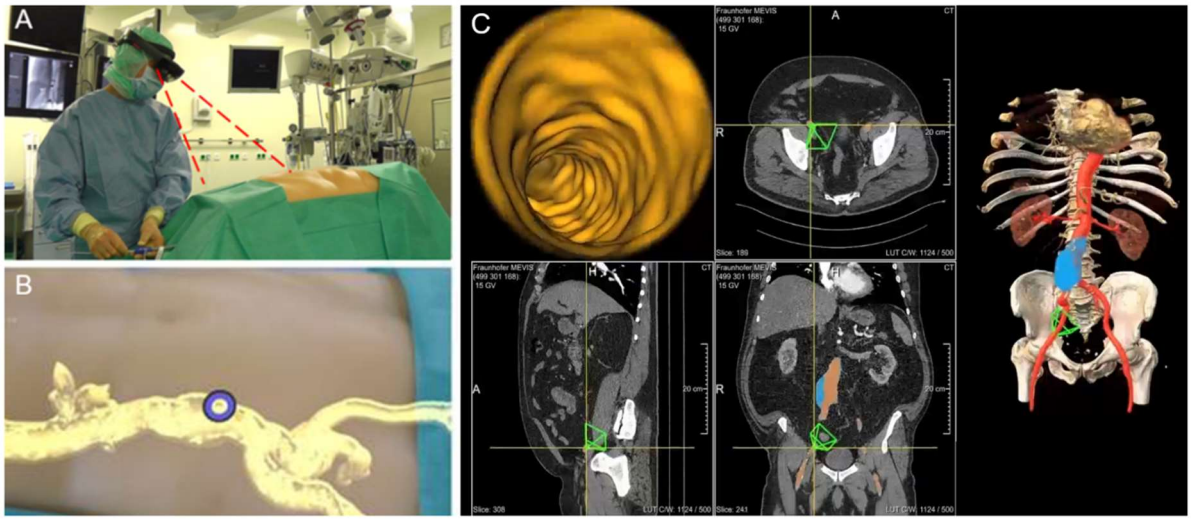


Fig. 1. View of the Surgeon via HoloLens (Microsoft inc.); B Augmented Reality of preoperative CT-Angiography on experimental Model; C Virtual Angioscopy, with courtesy Fraunhofer MeVis, Lübeck Nils Papenberg

Kontinuierliches Neuromonitoring CIONM – Status quo

E.Lauzana[1], T.Tuncer[1], S.Klein[2], W. Lamadé[1]

[1]Helios Spital Überlingen,Überlingen, Deutschland
[2]Bundesforschungsinstitut für Tiergesundheit, Neustadt, Deutschland

Kontakt: wolfram.lamade@helios-gesundheit.de

1 Einleitung

23 Jahre nach der Erstbeschreibung des kontinuierlichen Neuromonitorings cIONM bei Schilddrüsenoperationen wird dieses nun zunehmend eingesetzt und Signalverluste von 50% von vielen als Kriterium für den Abbruch der OP gesehen. Postoperative Recurrensfunktionsstörungen sind noch immer gefürchtete und oft völlig unerwartete Komplikationen in der Schilddrüsenchirurgie. Feinste mechanische Traumata führen manchmal zu prolongierten Recurrensparesen, während in anderen Fällen starke mechanische Belastungen ohne Konsequenz bleiben.

Ziel dieser Arbeit war es, die Reaktionsmuster des cIONM auf mechanische Zugkräfte zu korrelieren mit individuellen morphologischen Eigenheiten des N. laryngeus recurrens „RLN“.

2 Material und Methoden

An 20 Schweinen der Deutschen Landrasse wurden die Nervenleitfähigkeiten des NLR beidseits (n=40) durch kontinuierliches Neuromonitoring mit der Saxophonelektrode mittels Schwellstromkurven unter Zugbelastung analysiert. 30 NLR der Gruppe I wurden einer Zugbelastung von 1,2 N ausgesetzt bis nur noch 15 % des Ausgangssignals mittels Schwellstromkurve messbar waren. 10 NLR der Gruppe II wurden zeitkonstant über 10 und 20 min belastet. Messgröße war der prozentuale Abfall des Signals. Daraufhin folgte die Erholungsphase unter völliger Entlastung des Nervs, bis sich das Ausgangssignal wieder einstellte. Darauf erfolgte bei 11 Tieren die chirurgische Nervenentnahme und nach Aufbereitung folgt die Untersuchung unter dem konfokalen Lasermikroskop.

3 Ergebnisse

Die Ausgangssignale (0,08 – 1,50 mV) sind intra- und interindividuell sehr variabel. Die Dauer bis zum Eintritt eines Signalverlusts von 85 % reichen von 0,3 bis 171 min, Median 36 min. (li 0,3-171 min, re 0,3-168 min). Es fällt auf, dass immer eine Seite eines Tieres besonders schnell mit einem Signalverlust reagiert (0,3-98,9 min), wohingegen die andere Seite 26,8-171 min benötigt, um auf 15 % des Ausgangssignals abzufallen. Der empfindlichere Nerv stellt damit seine Funktion 1,4-146,4 mal (Median 4,3 mal) schneller ein, als der gegenseitige ($p<0,003$). Dieser überempfindliche Nerv wurde als „Weepy Nerve“ bezeichnet. Nach 20 min und konstantem Zug in der Gruppe II fanden sich Signalabfälle von 27-96% (Median 38%) beim empfindlicheren, Weepy Nerve, und 0-27% (Median 7%) Signalabfall auf der Gegenseite. Die empfindlichere Seite zeigte einen Signalabfall, welcher die Gegenseite im Median um 31% (26-68%) übersteigt ($p=0,008$). Morphologisch konnte keine Korrelation mit dem Nervendurchmesser oder der Faszikelanzahl gefunden werden. Die gesamte Fläche der Faszikel innerhalb des Nerven war jedoch hoch signifikant korreliert mit dem Weepy Nerve. Das heißt, der "Weepy Nerve" hatte eine kleinere Faszikelfläche als die Gegenseite (p-value: 0.014; Weepy Nerve – 0.0887mm2, Gegenseite – 0.1150mm2). Keine Seite war signifikant bevorzugt ($p = 0.19$).

4 Schlussfolgerung

Eines der wichtigsten Kriterien für das Reaktionsmuster des cIONM auf Zugkräfte stellt eine intrinsische, nicht symmetrische Eigenschaft des Nerven dar. Als "Weepy Nerve" wurde die mechanisch überempfindliche Seite bezeichnet, um den 1,4-146 mal empfindlicheren RLN zu beschreiben. Eine einfache Handlungsanweisung an den Chirurgen bei einem pauschalen Signalverlust im EMG scheint nur unter Kenntnis des „Weepy Nerve“ möglich.

Bonseyes - The open AI Marketplace supporting Surgical Data Science

D. Ostler[1], N. Marahrens[1], N. Kohn[1], D. Wilhelm [1,2], H. Feussner[1,2]

[1] Research Group for Minimally Invasive Interdisciplinary Therapeutical Intervention at the University Hospital rechts der Isar of the TU Munich, Trogerstr. 26, 81675 München, Germany

[2] Department for Visceral Surgery at the University Hospital rechts der Isar of the TU Munich, Ismaninger Str. 22, 81675 München, Germany

Contact: daniel.ostler@tum.de

Abstract

Recent advancements in digitalization and information technology influence healthcare and the goal of its optimization. Within surgery, the interdisciplinary field of surgical data science emerges and grows. Hereby machine learning in particular deep learning plays a key role in data-driven development. The Horizon 2020 funded project "Bonseyes" aims on the creation of an open platform, which significantly facilitates and improves the development process of such applications. In this work, we present our utilization of the Bonseyes platform demonstrated for the task of surgical tool detection.

Keywords: Surgical Data Science, Open Platform, Deep Learning

1 Problem

Driven by the goal of optimization of the "magic" triangle of the contradicting aims of quality of outcome, expenses and therapeutic trauma [1] for surgical health care, advancements in digitalization and information technology is having an increasing impact on interventional medicine. Due to the strive towards evidence based medicine, interdisciplinary data science becomes an emerging field, whose manifestation in surgery is called "Surgical Data Science". Surgical data science will pave the way from artisanal to data-driven interventional healthcare, with concomitant improvements in quality and efficiency of care [2]. Since "artificial intelligence" (AI) and especially its emerging branch of deep learning, plays a key role in the data-driven development, the demand for tool-chains and infrastructure that facilitate the development of such AI-systems. The Bonseyes collaborative project, funded by the European Commission, aims to create a platform for open development of systems of AI, consisting of a Data Marketplace, a Deep Learning Toolbox, and Developer Reference Platforms for organizations wanting to adopt AI. To evaluate the effectiveness, technical feasibility, and to quantify the real-world improvements, complementary demonstrators from different industries are currently built [3]. We present our work on the health care demonstrator.

2 Material and Methods

2.1 Healthcare Demonstrator

Many applications of computer assisted surgery depend on the analysis of the surgical workflow. This challenge is often tried to be solved by analyzing the laparoscopic video, in particular by the detection and tracking of surgical tools. As a sub-task within the Bonseyes Healthcare Demonstrator, we are focusing on object detection by leveraging the Bonseyes platform in order to facilitate the development process.

2.2 Bonseyes Development Platform

The concept of the Bonseyes platform is depicted in Figure 1. The Marketplace will facilitate the exchange of artefacts of AI systems: data, expert knowledge in the form of annotations, models, benchmarks and applications that are optimized for embedded edge computing. While the different components of the platform are under continuous development, we focused in this work particularly on the end-to-end development of a deep learning application based on a containerized development pipeline. Starting from the data set (laparoscopic images with

bounding box annotations of surgical instruments), the consecutive application of different process steps (i.e. pre-processing, model training, model benchmarking and deployment) results in a model that is optimized for the mobile embedded application.

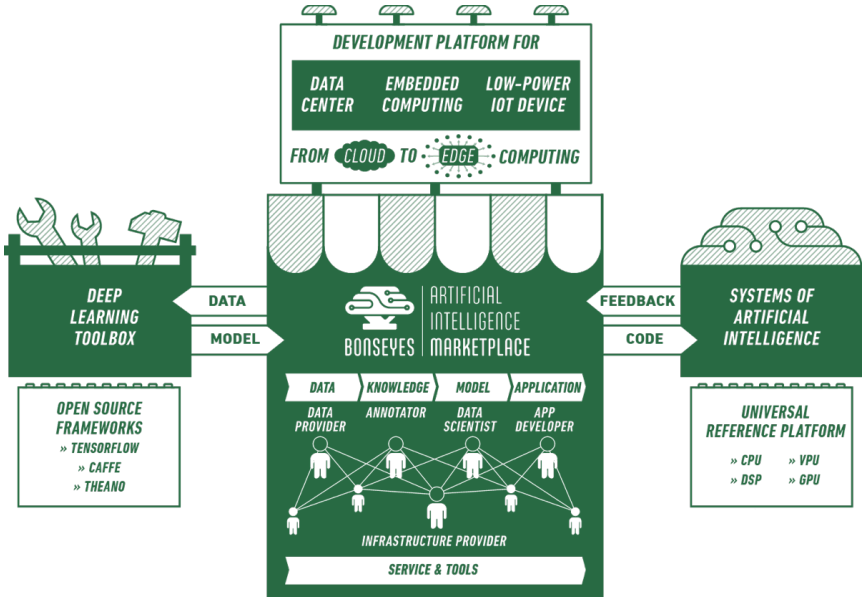


Figure 1: Concept of the Bonseyes platform with its central marketplace.

3 Results

Based on the collaborative work and support of consortium partners (in particular for this use case experts of nviso[1], UCLM[2], HES-SO[3], ICCS[4], FHNW[5]), an easy to use development pipeline was created, which allows the successfully training and deployment of the model *Single Shot Multibox Detector (SSD)* [4] with *MobileNets*. On our custom data set, the benchmarking resulted in a mean average precision (mAP) 86.4% across 11 classes. Figure 2 shows an exemplary image of the detection result.

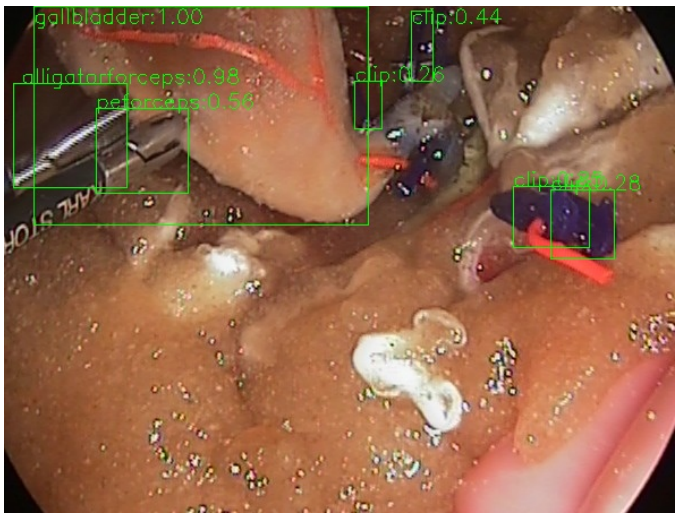


Figure 2: Shown is an image of a lap. Cholecystectomy performed in the ELITE-Phantom [5] after being processed by the SSD network. The network proposes bounding boxes around detected objects.

[1] nViso SA, Lausanne, Switzerland
[2] La Universidad de Castilla-La Mancha, Ciudad Real, Spain
[3] University of Applied Sciences Western Switzerland, Délémont, Switzerland
[4] Institute of Communication and Computer Systems, Athens, Greece
[5] Fachhochschule Nordwestschweiz, Windisch, Switzerland

Due to the Bonseyes Deep Learning Toolbox, the trained model can be easily transformed to platform-specific code and deployed on an embedded system. We utilized the automated code generation for a deployment on the NVIDIA Jetson TX2.

4 Discussion

While the presented application case of device detection can be considered as a standard object detection task, the novelty lies in the utilized toolchain. The Bonseyes toolchain enables a significantly acceleration of the end-to-end development of AI applications through its containerized design concept. Hereby it facilitates reutilization and sharing of data and models, while eliminating the common problem of dependency issues. Especially in an interdisciplinary field like SDS, this has great advantages in the adaptation of state-of-the-art machine learning techniques in terms of development time and cost of ownership. The empiric measurements of those Key Performance Indicators (KPIs) is foreseen as future work.

5 Conclusion

In this work, we presented an end-to-end development process of a machine learning application for surgical instrument detection, utilizing the Bonseyes platform. Its containerized modularity and ease of use allow fast development even for non-experts in the field of machine learning and hence is well suited for interdisciplinary research as the field of surgical data science. Since the deployment on embedded hardware is facilitated, the application in real world scenarios is enabled. The Bonseyes project and its different components are still under development and further applications in the field of SDS (e.g. end-to-end surgical workflow phase recognition) will follow.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732204 (Bonseyes). This work is supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0159. The opinions expressed and arguments employed herein do not necessarily reflect the official views of these funding bodies.

References

[1] Feussner, H., Ostler, D., Kranzfelder, M., Kohn, N., Koller, S., Wilhelm, D., ... & Schneider, A. (2017). Surgery 4.0. In Health 4.0: How Virtualization and Big Data are Revolutionizing Healthcare (pp. 91-107). Springer, Cham.

[2] Maier-Hein, L., Vedula, S. S., Speidel, S., Navab, N., Kikinis, R., Park, A., ... & Hashizume, M. (2017). Surgical data science for next-generation interventions. Nature Biomedical Engineering, 1(9), 691.

[3] Llewellynn, T., Fernández-Carrobles, M., Deniz, O., Fricker, S., Storkey, A., Pazos, N., ... & Goumas, G. (2017, May). BONSEYES: platform for open development of systems of artificial intelligence. In Proceedings of the Computing Frontiers Conference (pp. 299-304). ACM.

[4] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.

[5] Gillen, S., Wilhelm, D., Meining, A., Fiolka, A., Doundoulakis, E., Schneider, A., ... & Feussner, H. (2009). The" ELITE" model: construct validation of a new training system for natural orifice transluminal endoscopic surgery. Endoscopy, 41(05), 395-399.

Workflow Adaptive Cooperative Support System: Telephone Management

N. Samm1, N. Kohn2, D. Ostler2, N. Marahrens2, M. Kranzfelder1,2, T. Vogel1,2, D. Wilhelm1,2, H. Feußner1,2

[1] *Research Group for Minimally Invasive Interdisciplinary Therapeutical Intervention at the University Hospital rechts der Isar of the TU Munich, Trogerstr. 26, 81675 München*

[2] *Department for Visceral Surgery at the University Hospital rechts der Isar of the TU Munich, Ismaninger Str. 22, 81675 München*

Contact: nicole.samm@tum.de

Abstract

As much as mobile phone in the hospital increase communication and the exchange of crucial information about patients and their treatment, they are also a major source of interruption during surgical procedure and may negatively impact medical outcomes. This work presents the results of a user study for a prototypical telephone call management system. In conclusion the system may vastly minimize the number of non-critical calls.

Keywords: Cognitive OR, Telephone Call Management, Reduction of Surgical Workflow Interruption

1 Problem

Mobile phones increase the reachability of a surgeon and help to improve patient's care by stimulating communication [1]. Conversely, constant interruptions of the workflow in the surgical procedure due to incoming calls have an influence on the surgeon's concentration and may therefore negatively impact patient safety and treatment outcome [2,3]. Incoming telephone calls account for as much as every fourth disturbing intraoperative event [4]. To alleviate the effects of interrupting phone calls during surgery, we devised a system that is capable of automatically answering incoming phone calls with respect to the current situation in the OR (see Figure 1 – Left). By stimulating the caller to scrutinize the importance of their calls, the amount of phone calls that actually reaches the surgeon is supposed to be minimized.

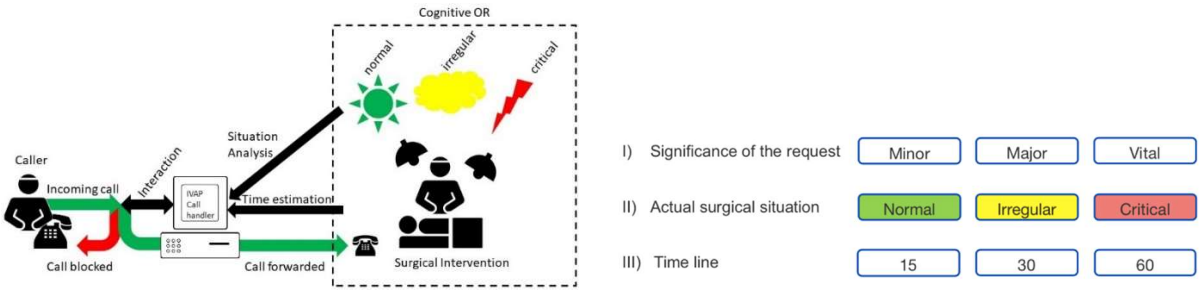


Figure 1*: Left: Schematic drawing of the proposed system. Right: Scenario parameter variations*

2 Material and Methods

To evaluate the effectiveness of the system and get further insides into the different factors that may help limit the amount of phone calls, we performed a user questionnaire study. In the questionnaire, participants were placed in the role of a caller, whose call was answered by the phone call management system. Therefore, participants were played different automatic voice recordings of the phone management system for different situations during surgery and for the case of different call types with respect to importance. In total, 46 volunteers from different functional areas of the University Hospital rechts der Isar were presented with three call type scenarios of different degrees of urgency in combination with different OR situations and time lines (see Figure 1 - Right). Class III scenarios were of minor importance, class II were of major importance and class I scenarios of vital importance for patient care.

For each call type scenario they were informed about the approximate remaining operating time (60, 30 or 15 minutes) and depending on the current situation in the OR (normal, irregular, critical), they were asked with increasing severity if their call was really necessary and crucial for the care of the patient, resulting in a total of 27 different overall scenarios (Figure 1 – Right). Subsequently, participants were asked to decide whether they still felt the need that their call should be forwarded to the operating surgeon.

3 Results

For class III calls (minor inquiries) only about 1% (4 out of 414 calls) of all participants confirmed to forward the call, for class II calls (major inquiries), about 7% (29 out of 414 calls) of all calls were routed to the surgeon and class I (vital inquiries) 78% (324 out of 414 calls) of all test persons opted for a forwarding to the surgeon. With a remaining operating time of 15 minutes, only 37% (57 out of 90 blocked calls) of the participants decided to be forwarded in case of a vital inquiry and after being informed by the system. Furthermore, results did not show a statistically significant correlation between increasing message severity for different OR situations (normal, irregular, critical) and the percentage of calls being forwarded. No significant difference with regard to the gender of the caller was observed.

4 Discussion

The results of the questionnaire indicate a statistically significant correlation between the importance of the call and the number of calls being forwarded. Whereas on average only every hundredth call for inquiries of minor and less than every fourteenth call for major importance was confirmed to be forwarded, more than two thirds of all vital calls were forwarded. Assuming that the majority of calls falls into classes II and III and are therefore no emergency calls, we can conclude that the vast majority of potentially unnecessary calls can be prevented from disturbing the surgery by means of the proposed system. Important emergency calls of class I in turn are less effected by the system or may even be postponed if the end of the surgery is near. Lastly, the different messages for different OR situations had no clear influence on the percentage of forwarded calls. This finding may be due to the messages themselves that might not have completely conveyed the level of severity of the OR situation.

5 Conclusion

In this paper we presented a concept of a telephone call management system as part of the cognitive OR with the aim of reducing intraoperative stress due to incoming phone calls without blocking important calls. Our survey showed that by merely stating that the surgeon is currently performing surgery, a vast majority of non-critical calls (class II and class III) were not forwarded and did, as desired, not disturb the surgery. Future work will be devoted into integrating more information into the automatic replies and evaluate their effect on the percentage of calls being forwarded to the OR. Furthermore, we may attempt to increase the number of participants and make a clearer distinction between different interest groups (leading surgeons, assistant surgeons, OR staff).

Acknowledgements

This work was done with grant support by the Bavarian Research Foundation (BFS).

References

[1] Mentis, H. M., Chellali, A., Manser, K., Cao, C. G., Schwaitzberg, S. D. (2016). A systematic review of the effect of distraction on surgeon performance: directions for operating room policy and surgical training. Surgical endoscopy, 30 (5), 1713-1724.

[2] Weigl, M., Stefan, P., Abhari, K., Wucherer, P., Fallavollita, P., Lazarovici, M., Catchpole, K. (2016). Intra-operative disruptions, surgeon's mental workload, and technical performance in a full-scale simulated procedure. Surgical endoscopy, 30 (2), 559-566.

[3] Arora, S., Sevdalis, N., Nestel, D., Woloshynowych, M., Darzi, A., Kneebone, R. (2010). The impact of stress on surgical performance: a systematic review of the literature. Surgery, 147 (3), 318-330.

[4] Weigl, M., Antoniadis, S., Chiapponi, C., Bruns, C., Sevdalis, N. (2015). The impact of intra-operative interruptions on surgeons' perceived workload: an observational study in elective general and orthopedic surgery. Surgical endoscopy, 29 (1), 145-153.

Elektrostimulation im GI-Trakt; Aktueller Stand und Ausblick

Peter Wilhelm, Universitätsklinik für Allgemeine, Viszeral- und Transplantationschirurgie, Tübingen

Die elektrische Modulation von Sakralnerven ist seit Ihrer Erstbeschreibung durch *Matzel et al.* 1995 heute fester Bestandteil des Armamentariums in der Therapie von Kontinenzstörungen und - mit Einschränkungen – von therapierefraktärer Obstipation [1, 2]. Die Modulation erfolgt mittels Implantation von „tined leads" in die Sakralforamina 2 - 4 und direkter Stimulation der angrenzenden Nervenstrukturen. Es handelt sich somit um eine rein efferente Technik, Afferenzen sind nicht ableitbar. Der Wirkungsmechanismus im Viszerum ist in klinischen und experimentellen Studien kontrovers diskutiert, bis heute aber nicht hinreichend verstanden.
Die Beeinflussung des oberen GI-Traktes ist über das System nicht möglich [3]. Die Datenlage hinsichtlich des Einflusses der Technik auf die Kolonmotilität ist widersprüchlich und bedarf weiterführender Studien [4]. Ein Effekt auf die Motorneurone, Wandspannung oder Compliance des Rektums ist ebenfalls nicht nachgewiesen [5]. Teilaspekte des Wirkmechanismus könnten dagegen in der Sensitivitätserhöhung für die Rektumdehnung und der arteriellen Blutflusserhöhung unter Stimulation im Endorgan liegen, jedoch erlaubt die derzeitige Studienlage auch hier keine validen Aussagen [4, 6, 7].
Trotz der sehr hohen Heterogenität der veröffentlichten Studien ist der Nachweis eines Effektes auf die Sphinkterfunktion am besten untersucht. Die positiven Effekte auf den willkürlichen maximalen Kneif- und den Ruhedruck bleiben in ihrem Mechanismus aber aufgrund der Studiendesigns weiterhin ungeklärt [4]. Nur wenige Studien haben bislang die zentralnervösen Effekte der Therapie adressiert. Mit ihrem Nachweis einer Perfusionszunahme des zur Implantationsseite kontralateralen frontalen Cortex unterstrichen *Lundby et al.* den multifaktoriellen Wirkmechanismus der Therapie [8]. *Hirabayashi et al.* wiesen in ihrer tierexperimentellen Arbeit über die selektive Stimulation der Sakralnerven Propulsionswellen vom proximalen Kolon bis zum Anus nach und zogen den Schluss, dass die grundsätzliche Machbarkeit für ein „colonic pacing" gegeben ist, sofern die Stimulation auf Höhe der Sakralnerven erfolgt [9]. Eine Modulation mittels viszeraler Elektrodenplatzierung untersuchten Sobocki et al. im porcinen Tiermodell [10]. Über eine laparoskopisch implantierte dreiarmige Stimulationssonde mit insgesamt 9 Elektroden erfolgte die Stimulation des Plexus hypogastricus. Dabei konnte eine 1:1 Stimulations:Defäkations und eine 1:3 Stimulations:Miktions Überleitung erzielt werden.
Die hohe Wirksamkeit und die unmittelbare Stimulationslokalität am Plexus hypogastricus machen diesen Ansatz besonders interessant. Rückschlüsse auf den zugrundeliegenden Mechanismus oder die Messung von gastrointestinalen Afferenzen lassen sich jedoch nicht gewinnen.
Multielektroden-Arrays (MEA) können mittels 3-dimensionaler Elektrodenanordnung sowohl Efferenzen als auch Afferenzen in der Fläche vermitteln. Ihr Einsatzgebiet ist insbesondere in der neurophysiologischen Forschung etabliert [11], viszerale Applikationen sind jedoch bislang auf experimentelle ex-vivo Ansätze beschränkt [12]. Im Rahmen eines BMBF-geförderten experimentellen Forschungsprojektes wird in einem gänzlich neuen Ansatz ab Herbst 2018 die MEA-vermittelte Neurostimulation viszeraler Nervenplexus und die Ableitung gastrointestinaler Afferenzen an der Universitätsklinik für Allgemeine, Viszeral- und Transplantationschirurgie in Tübingen untersucht.

Bibliografie

1. Matzel, K.E., et al., Electrical stimulation of sacral spinal nerves for treatment of faecal incontinence. Lancet, 1995. **346**(8983): p. 1124-7.
2. Holzer, B., et al., Sacral nerve stimulation in patients with severe constipation. Dis Colon Rectum, 2008. **51**(5): p. 524-29; discussion 529-30.
3. Worsoe, J., et al., Turning off sacral nerve stimulation does not affect gastric and small intestinal motility in patients treated for faecal incontinence. Colorectal Dis, 2012. **14**(10): p. e713-20.
4. Carrington, E.V., et al., A systematic review of sacral nerve stimulation mechanisms in the treatment of fecal incontinence and constipation. Neurogastroenterol Motil, 2014. **26**(9): p. 1222-37.
5. Michelsen, H.B., et al., Rectal motility after sacral nerve stimulation for faecal incontinence. Neurogastroenterol Motil, 2010. **22**(1): p. 36-41, e6.
6. Kenefick, N.J., et al., Effect of sacral nerve stimulation on autonomic nerve function. Br J Surg, 2003. **90**(10): p. 1256-60.
7. Abdel-Halim, M.R., et al., Temporary sacral nerve stimulation alters rectal sensory function: a physiological study. Dis Colon Rectum, 2011. 54(9): p. 1134-40.
8. Lundby, L., et al., Relief of fecal incontinence by sacral nerve stimulation linked to focal brain activation. Dis Colon Rectum, 2011. **54**(3): p. 318-23.
9. Hirabayashi, T., et al., Colorectal motility induction by sacral nerve electrostimulation in a canine model: implications for colonic pacing. Dis Colon Rectum, 2003. **46**(6): p. 809-17.
10. Sobocki, J., et al., Laparoscopically implanted system for stimulation of the hypogastric plexus induces colonic motility, defecation, and micturition: experimental study. Surg Innov, 2015. **22**(1): p. 70-6.
11. Eickenscheidt, M., et al., Electrical stimulation of retinal neurons in epiretinal and subretinal configuration using a multicapacitor array. J Neurophysiol, 2012. **107**(10): p. 2742-55.
12. Lutton, E.J., et al., Identification of uterine pacemaker regions at the myometrial-placental interface in the rat. J Physiol, 2018. **596**(14): p. 2841-2852.

What is new in non invasive Surgery with Focused Ultrasound?

Andreas Melzer [1]..

[1]*ICCAS, University Leipzig, Germany*
[2]*IMSaT, University Dundee, Scotland*

General Secretary EUFUS European charitable Society of Focused Ultrasound

www.EUFUS.org office@eufus.org

General Secretary iSMIT international Society for Medical Innovation and Technology

www.iSMIT.org office@ismit.org

Contact: andreas.melzer@iccas.de

Abstract

Focused ultrasound (FUS/HIFU) relies on ablation of pathological tissues, by delivering sufficiently high level of acoustic energy in situ into the human body. Magnetic Resonance guided FUS (MRgFUS/HIFU) and Ultrasound guided (USgFUS/HIFU) are image guided techniques combined with therapeutic FUS for monitoring purposes. Clinical applications FUS/HIFU gained CE and FDA approvals for the treatment of various benign and few malignant lesions in the last two decades. Current technical limitations of ultrasound guided and MRI guided Focused Ultrasound are related to challenges of ablating moving organs (Liver and Kidney). An outlook to possible new applications will be provided. Among those are real-time MRI guided focused ultrasound of the liver through a novel controlling system (Fraunhofer MEVIS funded by the European grant TransFusimo) which allows combining for 4D MRI tracking of liver lesions with beamforming of a novel phased array transducer (InSightec, Israel). Ultrasound guided focused ultrasound supported by robotic positioning has been realized ffor the first time funded by the European grant FUTURA. Two ABB arms are coordinated through specially developed software algorithms and positioning control is provided by Kinect and optical navigation. New applications using to KUKA robot arms including wireless ultrasound devices are currently being developed at ICCAS in Leipzig.

Support of radiation therapy through FUS and HiFU is currently being developed in Leipzig and Dresden (D) funded by a large grant from the German Ministry of science. Preliminary results of sonication of glioblastoma cells will be presented. The work includes the development of PET MRI compatible focused ultrasound (Fraunhofer IBMT) positioned by a PET MR compatible robot arm.

Acknowledgement

The presenter gratefully acknowledges funding by the European Community FP7 program for TransFusimo and FUTURA as well as funding from the German Ministry of Science (BMBF) for the SONO-RAY project.

Autorenverzeichnis

Die 17. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie in Leipzig ist das deutschsprachige Forum für den Austausch aktueller Methoden, Ergebnisse und Erkenntnisse in der computer- und roboterassistierten Chirurgie.

Dieser Band umfasst die wissenschaftlichen Beiträge von Ingenieur_innen, Informatiker_innen und Chirurg_innen in den Themenfeldern: Bildverarbeitung, Quantifizierung und Visualisierung, Chirurgische Planung und Risikoanalyse, Navigation, Robotik, Workflowanalyse und Prädiktion, Chirurgische Assistenzsysteme, Bildgestützte Systeme, Wissensbasierte Systeme und Entscheidungsunterstützung sowie Modellierung und Simulation.