

Thema:

**Bildaufbereitung für die Objekterkennung**

**Ausarbeitung** im Rahmen des Seminars  
„Unterstützung von Landminendetektion  
durch Bildauswertungsverfahren und Robotereinsatz“

im Fachgebiet Informatik

Themensteller: Prof. Dr. X. Jiang

Betreuer: S. Wachenfeld, Dr. D. Lammers

vorgelegt von: Jana Gliet, Birgit Pieke, Mareike Kritzler und Timo Schreiber

Abgabetermin: 2004-01-05

## Inhaltsverzeichnis

Inhaltsverzeichnis .....	II
Abbildungsverzeichnis .....	III
1 Einleitung.....	1
1.1 Das Bild und seine Wahrnehmung .....	1
1.2 Bildvorverarbeitung .....	3
1.3 Ziele der Bildvorverarbeitung.....	3
2 Eigenschaften von Bildern.....	5
2.1 Bildarten .....	5
2.2 Charakteristika von Bildern.....	7
3 Bildverbesserung .....	9
3.1 Systematisierung der Verfahren .....	9
3.2 Punktbezogene Operatoren.....	9
3.3 Nachbarschaftsoperatoren.....	14
3.3.1 Glättungsfilter.....	16
Mittelwertfilter .....	16
Medianfilter .....	16
Zusammenhang zwischen Mittelwert- und Medianfilter .....	17
3.3.2 Korrektur von uneinheitlichem Hintergrund.....	18
3.4 Geometrische Transformationen .....	19
3.5 Bildverarbeitung im Frequenzraum.....	22
3.5.1 Einführung in die Fourier Analyse.....	22
3.5.2 Die Fourier Transformation und ihre Eigenschaften.....	24
3.5.3 Filtern im Frequenzraum.....	27
4 Bildrestaurierung .....	35
5 Luftbildauswertung.....	40
5.1 Luftbildauswertung am Beispiel der vorgestellten Methoden.....	40
5.2 Landminenerkennung auf Luftbildern.....	43
5.3 Probleme der Luftbildauswertung .....	44
Literaturverzeichnis .....	46
Anhang .....	47
A Einige Beispiele im Zusammenhang mit Filterung im Frequenzraum.....	47
A.a Das Spektrum und der Phasenwinkel einer DFT .....	47

## Abbildungsverzeichnis

Abb. 1.1:	Das elektromagnetische Spektrum .....	2
Abb. 1.2:	Einige bekannte optische Täuschungen.....	2
Abb. 1.3:	Beispiel zur Kontrastsimulation .....	3
Abb. 2.1:	RGB- und HSV-Farbmodelle .....	6
Abb. 3.1:	Kontraststreckung und Binärisierung .....	10
Abb. 3.2:	Grundlegende Grauwerttransformationen .....	11
Abb. 3.3:	Stückweise lineare Transformation .....	12
Abb. 3.4:	Grauwertdehnung .....	12
Abb. 3.5:	Original mit Histogramm.....	13
Abb. 3.6:	Grauwertdehnung mit Histogramm .....	13
Abb. 3.7:	Histogrammausgleich .....	14
Abb. 3.8:	Beispiel einer Nachbarschaftsoperation .....	15
Abb. 3.9:	Beispiel, gefiltert mit Mittelwert- und Medianfilter .....	18
Abb. 3.10:	Beispiel der Korrektur von uneinheitlichem Hintergrund (links: Original) ..	19
Abb. 3.11:	Beispiel der Korrektur von eineinheitlichem Farbton (links: Original) .....	19
Abb. 3.12:	Translation .....	20
Abb. 3.13:	Skalierung .....	20
Abb. 3.14:	Tiepoints .....	21
Abb. 3.15:	Nearest Neighbour Interpolation .....	22
Abb. 3.16:	Bilineare Interpolation .....	22
Abb. 3.17:	Zerlegung einer periodischen Funktion als Fourier-Reihe .....	24
Abb. 3.18:	Zentriertes Spektrum eines Bildes .....	27
Abb. 3.19:	Vorgang der Filterung im Frequenzraum .....	28
Abb. 3.20:	Der Notch-Filter .....	29
Abb. 3.21:	Gauß-Hoch- und Tiefpassfilter im Bild- und Frequenzraum .....	31
Abb. 3.22:	Der ideale Tiefpassfilter in der Frequenzdomäne.....	32
Abb. 3.23:	Anwendung des idealen Tiefpassfilters mit Ringeffekten.....	33
Abb. 3.24:	Der Butterworth-Tiefpassfilter .....	34
Abb. 4.1:	Modell der Bildverschlechterung und -restauration .....	36
Abb. 4.2:	Grundlagen des Wiener-Filters und seine Anwendung in der Bildanalyse..	37
Abb. 4.3:	Restaurierung einer horizontalen Bewegungsunschärfe.....	39
Abb. 4.4:	Restaurierung einer Defokussierung mit dem Wiener Filter .....	39
Abb. 5.1:	Reduzierung von Wolkenschatten durch „Unsharp Masking“ .....	41
Abb. 5.2:	Verzerrung eines Luftbildes bei geneigter Aufnahmerichtung .....	42
Abb. 5.3:	Schematisches Beispiel zur geometrischen Mosaikbildung .....	42
Abb. 5.4:	Schritte der Minenfeldererkennung auf Luftbildern: 1. Minenfeldregion, 2. Ergebnis der Minensuche, 3. Abgrenzung des Minenfeldes .....	44

## 1 Einleitung

Inhalt dieses Referates ist die Darstellung einiger Verfahren der Bildaufbereitung für die Objekterkennung. Diese Verfahren der digitalen Bildverarbeitung werden für die Aufbereitung der von einem Roboter mit Hilfe einer USB - Kamera gewonnenen Bilder benutzt. Es wird das Ziel verfolgt, eine bessere Darstellung der Bilder zu erreichen und es soll möglich werden, bestimmte Eigenschaften oder Strukturen aus den Bildern extrahieren zu können, die für die weitere Verarbeitung von Bedeutung sind. Die Auswahl der hierbei verwendeten Verarbeitungsschritte richtet sich dabei immer nach der vorhandenen Bildqualität.

Welches Interesse wir an der digitalen Bildverarbeitung haben, zeigt sich an den zwei Anwendungsbereichen, in die prinzipiell unterschieden wird. Zum einen gibt es die Verbesserung bildlicher Informationen für die menschliche Interpretation, zum anderen die Verarbeitung von Bilddaten für die Speicherung, Übertragung und Repräsentation für autonome Maschinenwahrnehmung.

Im Zusammenhang mit der Bildverarbeitung gibt es drei Typen von computerisierten Prozessen, die low-level, mid-level und die high-level Prozesse:

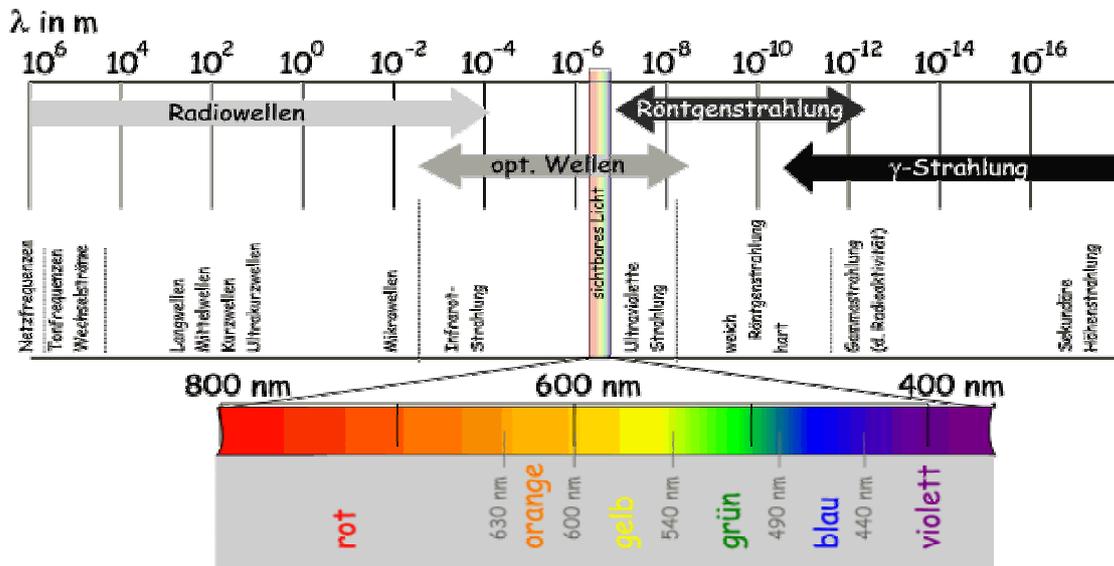
- Die low-level Prozesse umfassen primitive Operatoren wie Bildvorverarbeitung, z.B. um Rauschen zu reduzieren, Kontrastverbesserung vorzunehmen oder die Bildschärfe zu verstärken. Charakteristisch für diese Prozesse ist, dass sowohl Input als auch Output Bilder sind.
- Die mid-level Prozesse befassen sich unter anderem mit Segmentierung. Dabei wird das Bild in Regionen oder Objekte eingeteilt und eine Objektbeschreibung vorgenommen. Hier wird als Input ein Bild verwendet und als Output erhält man aus dem Bild extrahierte Attribute.
- Die high-level Prozesse analysieren ein Bild.

Im Folgenden werden einige low-level-Verfahren vorgestellt, die ein Bild hervorbringen, mit dem es anschließend möglich ist, Objekte zu erkennen.

### 1.1 Das Bild und seine Wahrnehmung

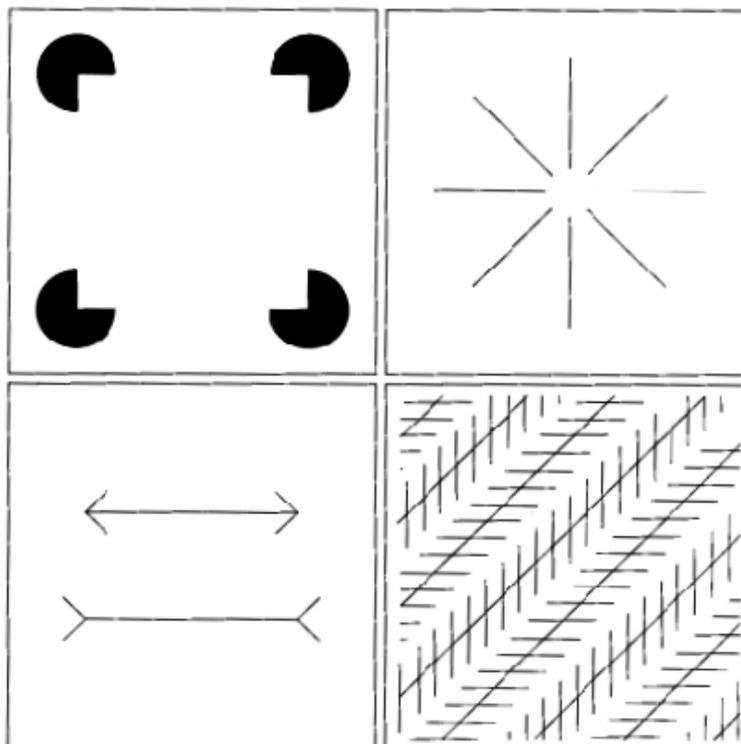
Ein digitales Bild ist als eine zweidimensionale Funktion  $f(x,y)$  definiert und besteht aus einer endlichen Anzahl von Elementen (Pixeln). Jedes von diesen Pixeln (Picture Element) hat eine bestimmte Position und einen Wert. So ist es möglich die verschiedenen Verfahren anzuwenden.

Bei uns Menschen ist der Sehsinn der ausgeprägteste unserer Sinne. Deshalb ist es nicht überraschend, dass Bilder die wichtigste Rolle in der menschlichen Wahrnehmung spielen, jedoch besitzen wir nur ein eingeschränktes visuelles Band von dem gesamten elektromagnetischen (EM) Spektrum (s. Abb. 1.1) und lassen uns leicht täuschen (s. Abb. 1.2).



Quelle: [http://www.roro-seiten.de/physik/lk12/emwellen/elektromagnetisches\\_spektrum.html](http://www.roro-seiten.de/physik/lk12/emwellen/elektromagnetisches_spektrum.html).

Abb. 1.1: Das elektromagnetische Spektrum



Quelle: Gonzales, Woods (2002), S. 42.

Abb. 1.2: Einige bekannte optische Täuschungen

## 1.2 Bildvorverarbeitung

Die digitale Bildvorverarbeitung gliedert sich in die Bereiche Bildverbesserung und Bildrestaurierung. Die Bildverbesserung ist das einfachste und verlockendste Feld der digitalen Bildvorverarbeitung. Hintergrundidee ist, nicht wahrnehmbare Details sichtbar zu machen oder einfach bestimmte Objekte, die in einem Bild interessant sind, deutlicher hervorzuheben. Die Bildverbesserung kann in zwei Kategorien unterteilt werden:

- Zum einen gibt es die **räumliche Domäne**. Hier findet die Realisierung auf der Ebene des Bildes statt, das heißt, die Manipulationen arbeiten direkt auf den Pixeln.
- Zum anderen existiert die **Frequenzdomäne**. Dabei wird ein Bild mit Hilfe der Fourier Transformation in Wellen verschiedener Frequenz zerlegt, die globales Filtern ermöglicht.

Ein bekanntestes Beispiel für eine Verbesserung ist, den Kontrast eines Bildes zu vergrößern, weil es „besser aussieht“. Bei der Bildverbesserung handelt es sich dabei um den subjektiven Teil der Bildvorverarbeitung.



Quelle: eigene Darstellung

**Abb. 1.3:** Beispiel zur Kontrastsimulation<sup>1</sup>

Die Techniken zur Bildrestaurierung haben zum Ziel, ein Bild in einem vorgefertigten Sinn zu verbessern, hier im Gegensatz zur Bildverbesserung als objektiver Prozess.

## 1.3 Ziele der Bildvorverarbeitung

Bildverbesserung und Bildrestaurierung werden also zur Aufbereitung von Bildsignalen benutzt, um wichtige Merkmale hervorzuheben. Folglich werden gering ausgeprägte Strukturen betont, während redundante oder gestörte Bildinformationen unterdrückt werden.

---

<sup>1</sup> Anm.: Die inneren Quadrate haben alle die gleiche Intensität, aber sie wirken dunkler, wenn der Hintergrund heller wird.

Das Ziel der Bildvorverarbeitung ist, ein Bild zu erstellen, in dem der Bildbereich hervorgehoben ist, der die später zu gewinnenden Daten enthält. Nach der Vorverarbeitung muss sich dieser Datenbereich in der Helligkeit von der Umgebung deutlich unterscheiden.

Es soll der subjektiv-visuelle Eindruck verbessert werden. Die Vorverarbeitung ist eine Bildtransformation, bei der als Ergebnis wieder ein Bild entsteht. Dabei wird der Informationsgehalt des Bildes in Anbetracht der späteren Auswertung nicht wesentlich verändert. Weiterhin sollen nachfolgende Verarbeitungsschritte vereinfacht werden.

## 2 Eigenschaften von Bildern

### 2.1 Bildarten

Es gibt viele Arten von Bildern, aber allgemein lassen sie sich in die drei Arten Intensitäts- bzw. Grauwertbilder, Binärbilder und Farbbilder unterteilen. Diese werden nachfolgend im Detail vorgestellt.

#### *Grauwertbilder*

Bei Grauwertbildern stellt jeder Pixel einen Intensitätswert dar, der bei einer 8-bit Speicherung Werte von 0 bis 255 annehmen kann. Es ergeben sich somit  $2^8 = 256$  verschiedene Intensitätsstufen, weshalb Grauwertbilder auch als Intensitätsbilder bezeichnet werden. Ein Intensitätswert von 0 entspricht dabei der Farbe schwarz, der Wert 255 der Farbe weiß. Mathematisch gesehen kann ein Grauwertbild als eine Funktion  $g(r,c)$  betrachtet werden. Der Funktionswert an der Stelle  $(r,c)$  entspricht dem Grauwert, bzw. der Helligkeit des Bildes.

Grauwertbilder haben mehrere Vorteile in der Bildverarbeitung. Zum einen haben sie gegenüber Farbbildern einen geringeren Speicherverbrauch und zum anderen lassen sich Bildverarbeitungsoperatoren recht einfach auf ihnen ausführen, da nur ein Kanal mit Werten berücksichtigt werden muss.<sup>2</sup>

#### *Binärbilder*

Aus Grauwertbildern lassen sich durch eine Schwellwertoperation Binärbilder erzeugen, was auch als Binärisierung bezeichnet wird. Dazu wird ein Intensitätswert als Schwellwert festgelegt, so dass jeweils die 0 den Werten unterhalb des Schwellwertes und die 1 denen oberhalb des Wertes zugewiesen wird. Die Visualisierung kann als Zweipegelebild geschehen, bei der den Werten 0 und 1 bestimmte Intensitäts- oder Farbwerte zugeordnet werden.<sup>3</sup> Durch die Schwellwertwahl wird das spätere Bild daher entscheidend beeinflusst. Eine Verwendung in der Bildvorverarbeitung und Objekterkennung finden Binärbilder u. a. bei der Trennung von einem Objekt und dem Hintergrund.

---

<sup>2</sup> Viele der folgenden Beispiele beziehen sich daher auf Grauwertbilder.

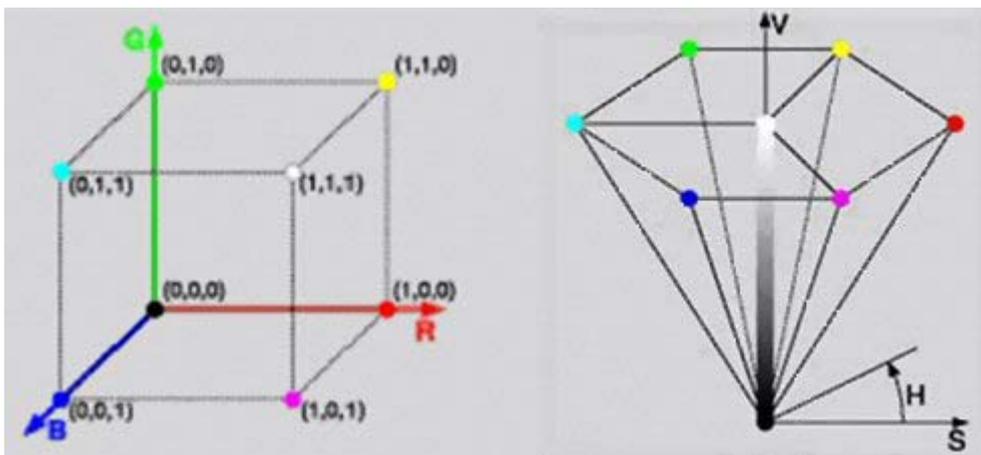
<sup>3</sup> Vgl. Jiang (2002), S. 26.

## Farbbilder

Zur Darstellung von Farbbildern gibt es zahlreiche Farbmodelle, von denen hier nur das RGB- und das HSV-Modell kurz vorgestellt werden sollen.

Das RGB-Modell beruht auf der additiven Überlagerung der drei Grundfarben rot, grün und blau. Die Farbe Weiß entsteht dabei durch die Überlagerung der drei Farben in voller Intensität.<sup>4</sup>

Das HSV-Modell hingegen beruht auf der intuitiven Zusammensetzung der Farben, so dass jede Farbe durch ein Tripel hue (Farbton), saturation (Sättigung) und value (Helligkeit) beschrieben wird. Der Vorteil dieses Farbmodells gegenüber dem RGB-Farbmodell besteht darin, dass bei Änderungen der Farbe oder der Helligkeit nur ein Kanal betrachtet werden muss gegenüber allen drei Kanälen im RGB-Modell. Das HSV-Modell ermöglicht daher eine unabhängige Betrachtungsweise von Farb- und Helligkeitsinformationen.<sup>5</sup>



Quelle: Vgl. Preim (2003), S. 14-16.

**Abb. 2.1:** RGB- und HSV-Farbmodelle

Operationen zur Bildverbesserung lassen sich auf Farbbildern ähnlich durchführen wie auf Grauwertbildern, im RGB-Modell muss dabei in den meisten Fällen die Berechnung auf allen 3 Kanälen durchgeführt werden. Eine Ausnahme bildet u. a. der Medianfilter<sup>6</sup> im RGB-Modell, da es keine sinnvolle lineare Ordnung „ $\leq$ “ auf Farbvektoren gibt und durch eine unabhängige Betrachtung der Farbkanäle Farben entstehen können, die nicht mehr mit den Ausgangsfarben zusammenhängen.

Die Konvertierung eines Farbbildes in ein Grauwertbild ist eindeutig und lässt sich relativ einfach durchführen (im RGB-Modell durch Mittelwertbildung der drei Kanäle). Eine

<sup>4</sup> Vgl. Abb. 2.1.

<sup>5</sup> Vgl. Pandit (2001), S. 164.

<sup>6</sup> Vgl. Jiang (2002), S. 25.

Pseudokolorierung eines Grauwertbildes ist im Gegensatz dazu nicht mehr eindeutig und somit wesentlich aufwendiger.

## 2.2 Charakteristika von Bildern

Charakteristische Eigenschaften von Bildern lassen sich durch Histogramme und statistische Kennwerte wie z. B. Mittelwert oder Varianz beschreiben. Die statistischen Kennwerte ermöglichen eine objektive Beurteilung des Bildes, die nicht durch subjektive Eindrücke beeinflusst wird. In der Bildvorverarbeitung helfen diese Kennwerte zu entscheiden, ob und in welchem Maße bestimmte Verbesserungsverfahren auf ein spezielles Bild angewendet werden können, um bestimmte Merkmale dieses Bildes besser herauszustellen und somit den objektiven Eindruck zu verbessern.

### *Histogramm*

Unter einem Histogramm eines Grauwertbildes  $g$  versteht man die relative oder absolute Häufigkeitsverteilung der im Bild vorkommenden Grauwerte  $k$ . Ein absolutes Grauwert-Histogramm  $H(k)$  ist die Summe aller Pixel mit gleichem Grauwert, ein relatives Histogramm  $h(k)$  beschreibt die Wahrscheinlichkeit, mit der ein bestimmter Grauwert im Bild auftritt.

$$H(k) = \left| \left\{ (r, c) \mid g(r, c) = k \right\} \right| \text{ absolutes Grauwert-Histogramm}$$

$$h(k) = \frac{H(k)}{M \cdot N} \text{ relatives Grauwert-Histogramm mit der Größe } M \times N$$

Ein Histogramm enthält Informationen über die Statistik erster Ordnung eines Bildes. Darunter versteht man Aussagen über das Auftreten von bestimmten Werten, die unabhängig von irgendwelchen anderen benachbarten Werten gemacht werden. Haben zwei Bilder das gleiche Histogramm, so sind sie deswegen nicht notwendigerweise identisch.<sup>7</sup> Auch für Farbbilder lassen sich Histogramme erstellen, doch diese sind entweder vierdimensional (drei Farbdimensionen und eine „Häufigkeitsdimension“), was sich schwer darstellen lässt oder beschreiben die drei einzelnen Dimensionen des RGB-Modells, denen der Zusammenhang fehlt.

---

<sup>7</sup> Vgl. Pandit (2001), S. 50.

### Mittlerer Grauwert

Der Mittlere Grauwert  $\bar{f}$  gibt Auskunft über die allgemeine Helligkeit eines Bildes, also darüber, ob bei einer subjektiven Betrachtung dieses Bild als zu hell oder zu dunkel empfunden wird. Je größer der Wert, desto heller erscheint das Bild. Eine Anwendung findet der Mittelwert auch im Frequenzraum im Zusammenhang mit der 2-dim DFT.<sup>8</sup> Berechnet wird der mittlere Grauwert aus den relativen Häufigkeiten wie folgt:<sup>9</sup>

$$\bar{f} = \frac{1}{N \cdot M} \sum_{r=0}^{N-1} \sum_{c=0}^{M-1} g(r, c) = \sum_0^{255} h(k)$$

### Varianz

Die Varianz ist ein Maß für die Abweichungen der Grauwerte aller Pixel vom mittleren Grauwert  $\bar{f}$ . Mit Hilfe der Varianz lässt sich der Kontrast eines Bildes beschreiben. Es wird als kontrastreich bezeichnet, wenn eine weite Streuung der Grauwerte im Histogramm vorliegt, da dann die Grauwertübergänge stärker hervorgehoben werden. Durch den direkten Zusammenhang mit dem mittleren Grauwert errechnet sich die Varianz aus den relativen Häufigkeiten durch folgende Formel:

$$\sigma^2 = \frac{1}{N \cdot M} \sum_{r=0}^{N-1} \sum_{c=0}^{M-1} (g(r, c) - \bar{f})^2 = \sum_0^{255} h(k) \cdot (k - \bar{f})^2$$

---

<sup>8</sup> Vgl. Kap. 3.5.

<sup>9</sup> Vgl. Jiang (2002), S. 17.

## 3 Bildverbesserung

### 3.1 Systematisierung der Verfahren

Um einen besseren Überblick über die große Anzahl von verschiedenen Verfahren zur Bildverbesserung zu bekommen soll hier eine Systematisierung der Verfahren vorgenommen werden. Verschiedene Einteilungskriterien sollen dabei helfen, die unterschiedlichen Verfahren besser in einen globalen Zusammenhang einordnen zu können.

Eine Möglichkeit der Gliederung bietet die Unterteilung nach der Anzahl der Pixel, die eine Verarbeitung beeinflussen. Dabei wird zwischen punktbezogenen, lokalen und globalen Operatoren unterschieden. Punktbezogene Operatoren wie z. B. die affine Transformationen beziehen sich bei der Verarbeitung ausschließlich auf einen Punkt, lokale Operatoren wie z. B. Filtermasken auf eine frei wählbare Nachbarschaft und globale Operatoren, wie beispielsweise die Fourier-Transformation, auf das gesamte Bild.<sup>10</sup>

Nach einer Einteilung in Bearbeitungsbereiche lassen sich homogene und inhomogene Verfahren unterscheiden. Ein homogenes Verfahren wie z. B. der Mittelwertfilter wird dabei für jede Region eines Bildes gleich benutzt und ist somit unabhängig von der Bildkoordinate. Im Gegensatz dazu steht das inhomogene Verfahren, in dem der Operator von der Bildkoordinate abhängig ist und aus der Umgebung berechnet werden kann.

Eine dritte Möglichkeit zur Systematisierung bietet die Einteilung nach der Linearität der Verarbeitung. Lineare Verfahren wie z. B. die Konvertierung in ein Negativbild können rückgängig gemacht werden und sind somit invertierbar. Nichtlineare Verfahren hingegen sind nicht invertierbar. Zu diesen Verfahren zählt auch die Binärisierung, da durch sie die Entropie, d.h. der Informationsgehalt eines Bildes sinkt und nicht wieder rückgängig zu machen ist.

### 3.2 Punktbezogene Operatoren

Alle punktbezogene Grauwerttransformationen basieren auf räumlichen Transformationen, so dass direkt die Pixel eines Bildes manipuliert werden. Die Anwendung dieser Transformationen ist in der Regel spezifisch und problemorientiert. So ist diese Form der Bildverarbeitung ein subjektiver Prozess und beruht auf der visuellen Auswertung der Bildqualität.

---

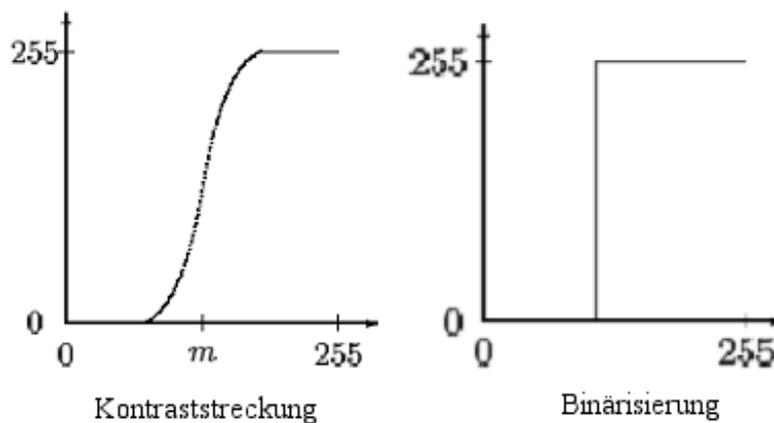
<sup>10</sup> Vgl. Jiang (2003), S. 20.

Grauwerttransformationen basieren auf der Gesamtmenge aller Pixel aus denen ein Bild besteht. Die allgemeine Formel lautet:

$$g(x, y) = T[f(x, y)]$$

Dabei ist  $f(x,y)$  der Grauwert des Eingangsbildes an der Stelle  $x, y$  und  $g(x,y)$  der Wert des bearbeiteten Bildes.  $T$  ist der Operator auf  $f$ , der in diesem Fall Grauwerttransformation-funktion ist:  $s = T(r)$

Hier sind  $s$  und  $r$  Variablen bzw. die Grauwerte von  $f(x,y)$  und  $g(x,y)$  von jedem Punkt  $x, y$ .  $T$  wird auch Transformationskennlinie genannt. Ein Beispiel in Abb. 3.1 dient zur Verdeutlichung.



Quelle: Vgl. Jiang(2003), S. 22-28.

**Abb. 3.1:** Kontraststreckung und Binarisierung

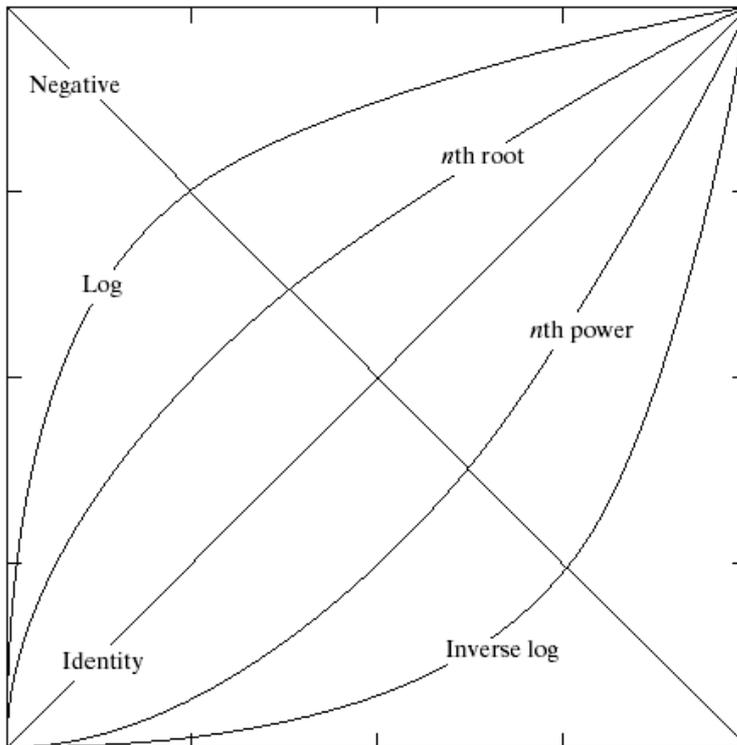
In Abb. 3.1 links wird ein höherer Kontrast erzeugt, da die dunklen Pixel im Bild verdunkelt und die hellen Anteile erhellt werden. So eine Transformationsfunktion nennt man auch Kontrastdehnung, oder Kontraststreckung.

Bei Abb. 3.1 rechts wird das Eingabebild in ein binäres Bild umgewandelt, da die Werte nur noch zwei Werte annehmen. Hier heißt  $T$  auch Schwellwertfunktion.

### *Grundlegende Grauwerttransformationen*

$r$  und  $s$  sind hier wieder die Werte der Pixel vor und nach der Bearbeitung.  $T$  ist die Transformation, die  $s$  aus  $r$  berechnet:  $s = T(r)$ .

Dann können folgende Grundtypen von Transformationen unterschieden werden:



Quelle: Gonzales, Woods (2002), S. 78.

**Abb. 3.2:** Grundlegende Grauwerttransformationen

1. *Identität und Negativ*. Am einfachsten ist die Identität zu berechnen:  $s = r$ .

Beim Negativbild wird der Wert  $r$  von dem höchst möglichen Grauwert abgezogen. Die Formel lautet:

$$s = 255 - r.$$

2. *Logarithmus Transformation* (kurz: log Transformation). Die Formel für diese Transformation lautet:  $s = 255 * \frac{\log(1+r)}{\log 256}$

mit  $r \geq 0$ . Diese Formel bewirkt eine Dehnung der Grauwertdynamik im unteren Grauwertbereich und gleichzeitig eine Stauchung im oberen Grauwertbereich. Also werden die dunklen Pixel im Bild expandiert und die hellen komprimiert. Dadurch werden Details im dunklen Bereich besser sichtbar gemacht. Die inverse log Transformation hat genau den gegenteiligen Effekt (als Beispiel das Bild aus dem Skript der Vorlesung Bildverarbeitung auf S. 23).

3. *Exponentielle Transformation* (power-law Transformation). Sie hat die allgemeine Form:  $s = 255 * (r/255)^\gamma$

mit  $\gamma$  als Konstante. Diese Funktion ist vielfältiger anwendbar als die vorherige. Für  $\gamma < 1$  ist die Funktion ähnlich der Log Transformation. Hier erhöht die exponentielle Transformation die Helligkeit und ermöglicht insbesondere eine bessere Sichtbarkeit

im dunklen Bereich. Für  $\gamma > 1$  ist das Verhalten genau umgekehrt. Das bearbeitete Bild wirkt dunkler. Das ermöglicht bei zu hellen Bildern eine bessere Betrachtung.

Allgemein werden die exponentiellen Transformationen gebraucht für Kontrastmanipulationen. Die Transformationskennlinien für die ersten drei Grundtypen sind in Abb. 3.2 zu erkennen.

4. *Stückweise-lineare Transformation.* Bei der Transformation müssen mehr Werte spezifiziert werden als bei den vorherigen. Diese Funktion kann verwendet werden um Bildinhalte aus einem Teilbereich  $[a,b]$  hervorzuheben (Abb. 3.3) oder um Grauwerte zu dehnen, z. B. bei kontrastarmen Bildern, bei denen die Grauwerte nur im Intervall  $[a,b]$  liegen und dann auf den Bereich  $[0,255]$  ausgedehnt werden (Abb. 3.4). Problem hierbei ist, dass man die Intervallgrenzen  $a$  und  $b$  bestimmen muss.

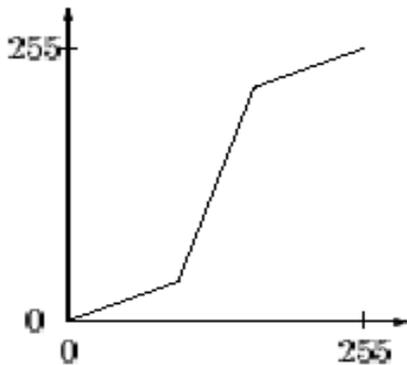


Abb. 3.3: Stückweise lineare Transformation

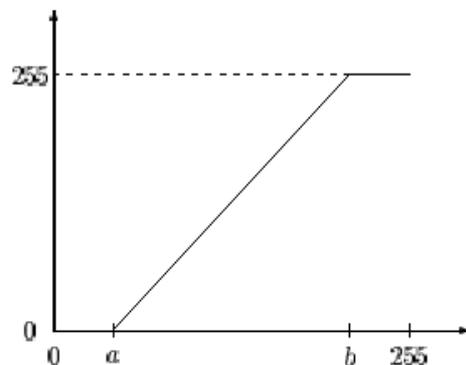


Abb. 3.4: Grauwertdehnung

Quelle: Jiang (2003), S. 26.

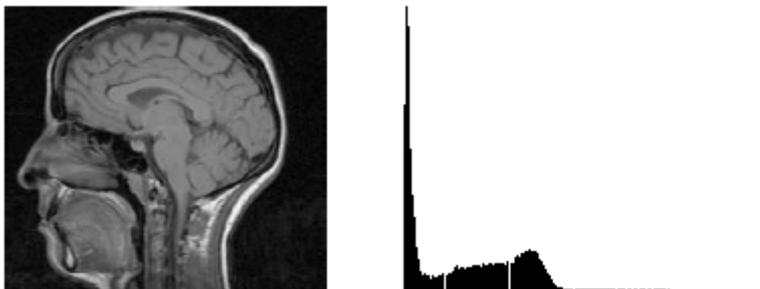
### Histogrammtransformationen

Histogramm:  $h(r_k) = n_k$

Das Histogramm eines Bildes beschreibt die Häufigkeit des Vorkommens eines Grauwertes. Ein Bild, das ein Histogramm mit großer Ausdehnung der Grauwerte und gleichmäßiger Verteilung hat, wirkt sehr kontrastreich und hat viele Grauwerte. Man kann anhand des Histogramms des Inputbildes eine Transformationsfunktion erstellen.

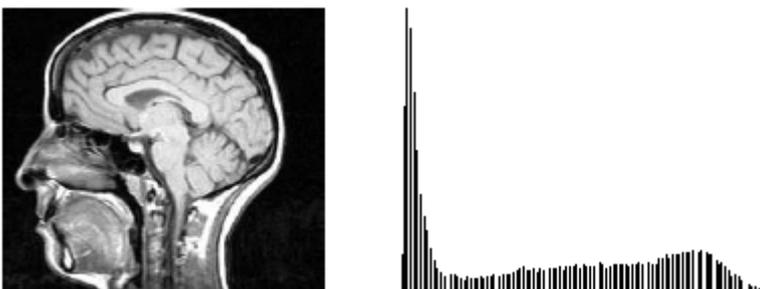
Bei den Histogrammtransformationen gibt es die Histogrammspreizung. Sie nutzt den Wertebereich zur Verbesserung der Darstellung aus. Die Zuordnungsvorschrift ergibt sich aus dem kleinsten und größten im Bild vorkommendem Grauwert. Diese Transformation ist ähnlich der stückweisen linearen Transformation. Mit Hilfe des Histogramms können hierbei die Intervallgrenzen einfach bestimmt werden. In Abb. 3.5 ist das Originalbild dargestellt, das in Abb. 3.6 durch eine Histogrammspreizung transformiert wurde.

$$g(x,y) = \frac{f(x,y) - a}{b - a} * 255$$



Quelle: Jiang (2003), S. 27.

**Abb. 3.5:** Original mit Histogramm



Quelle: Jiang (2003), S. 27.

**Abb. 3.6:** Grauwertdehnung mit Histogramm

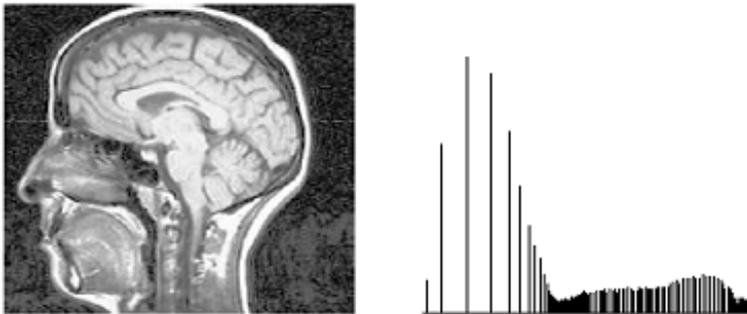
Eine weitere Transformation ist der Histogrammausgleich. Es wird das Ziel verfolgt, große Anhäufungen im Histogramm zu strecken oder auch weniger genutzte Grauwertbereiche

zu stauchen. Insgesamt soll das transformierte Bild damit ein gleichverteiltes Histogramm erhalten.

Gesucht ist deshalb eine monotone Abbildung  $T: [0,255] \rightarrow [0,255]$ , so dass  $g(x,y) = T(f(x,y))$  mit  $s = T(r)$  ein gleichverteiltes Histogramm  $h(r_k)$  hat. Jedem Grauwert  $r$  wird die Summe der bis zu  $r$  aufsummierten relativen Häufigkeiten zugeordnet.

$$s = 255 * \sum_{p=0}^r hf(r)$$

, mit  $hf(r)$  als relative Häufigkeit von  $r$ .



Quelle: Jiang (2003), S. 31.

**Abb. 3.7:** Histogrammausgleich

Im Bild ist das Resultat des Histogrammausgleichs deutlich zu erkennen. Es zeigt, dass der Histogrammausgleich ein gutes Verfahren zur Bildverbesserung ist.

### 3.3 Nachbarschaftsoperatoren

Bei den zuvor vorgestellten punktbezogenen Operatoren können die räumlichen Beziehungen der Grauwerte einer kleineren Umgebung nicht erkannt werden. Aus diesem Grund ist eine weitere Gruppe von Operatoren nötig, die die Bildpunkte einer Nachbarschaft in angemessener Weise kombinieren. Sie werden auch Nachbarschaftsoperatoren oder Filter genannt und arbeiten direkt auf den Pixeln des Bildes, wobei  $f$  das Eingabebild ist und durch die Operationen zu dem verbesserten Ausgangsbild  $g$  wird:

$$g(r,c) = \Omega(N(f(r,c)))$$

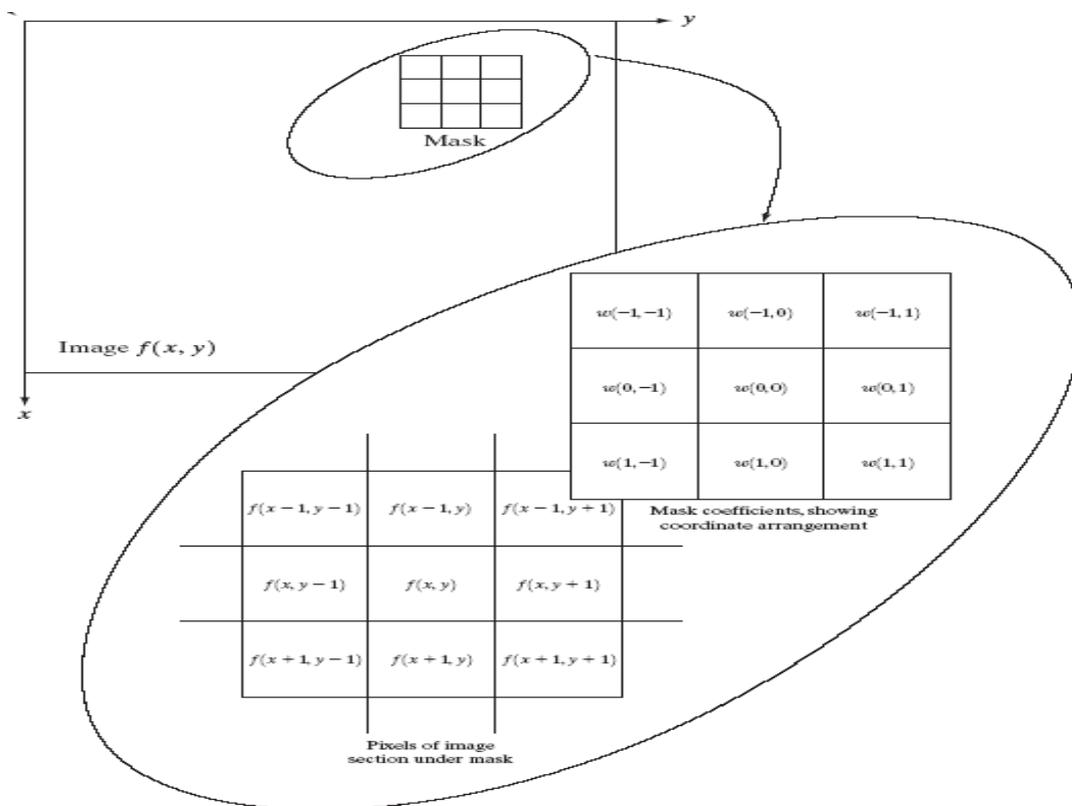
Eine grundlegende Kombinationsart von Bildpunkten in einer Nachbarschaft ist die lineare Operation mit einer Filtermaske. Eine Maske ist ein kleines (z. B.  $3 \times 3$ ) zweidimensionales Array, das die Größe der Nachbarschaft abdeckt und stellt sich wie folgt dar:

$$h(u,v), -k \leq u \leq k, -k \leq v \leq k \quad (\text{Annahme bei einer } n \times n \text{ Maske: } n = 2k + 1)$$

Bei einer Nachbarschaftsoperation wird jeder Bildpunkt, der von der Bildmaske bedeckt wird mit dem entsprechenden Gewicht in der Maske multipliziert. Anschließend werden alle Punkte addiert und die Summe wird an die Position des zentralen Bildpunktes gesetzt, was in der folgenden Formel sichtbar wird:

$$g(r,c) = \sum_{u=-k}^k \sum_{v=-k}^k h(u,v) f(r+u, c+v) \quad (k = (n - 1) / 2)$$

Dieser Prozess besteht einfach daraus, dass die Maske sich von Punkt zu Punkt im Bild bewegt und den neuen Punkt durch eine vordefinierte Nachbarschaft berechnet.



Quelle: Gonzales, Woods (2002), S. 117.

**Abb. 3.8:** Beispiel einer Nachbarschaftsoperation

Im Beispiel (s. Abb. 3.8) wird das Bild  $f$  mit einer 3x3 Maske  $h$  bearbeitet. Als Ergebnis erhält man ein verbessertes, jedoch verkleinertes Bild  $g$ :

$$h = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad f = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} \quad g = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & 18 & 27 & \cdot \\ \cdot & 18 & 27 & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Die Nichtberücksichtigung der Randpunkte führt dazu, dass das Bild bei wiederholter Anwendung der Maske immer kleiner wird. Um ein Ergebnisbild derselben Größe zu erhalten, müssen die Randpunkte gesetzt werden. Dazu gibt es mehrere Möglichkeiten, z.B. können die Randpunkte einfach unverändert beibehalten werden.

### 3.3.1 Glättungsfilter

Durch Glättungsfilter sollen Störungen rückgängig gemacht werden. Störungen in Bildern liegen dann vor, wenn benachbarte Bildpunkte stark unterschiedliche Grauwerte aufweisen, obwohl diese im Normalfall nur wenig schwanken sollten. Dabei wird angenommen, dass Störungen unabhängig von der Position und vom Grauwert des Bildpunktes sind. Das Ergebnis eines solchen Filters ist der Durchschnitt der Pixel, die in der Nachbarschaft der Filtermaske enthalten sind. Der Wert jedes Pixels wird durch den durchschnittlichen Grauwert der Nachbarschaft ersetzt, der durch die Filtermaske definiert wird.

#### *Mittelwertfilter*

Die Idee, die hinter dieser Glättung des Rauschens steht, ist, den arithmetischen Mittelwert der Grauwerte aus der Umgebung eines jeden Bildpunktes durch einen lokalen Filter zu bilden. Bei dem Mittelwertfilter handelt es sich um eine lineare Operation.

Die Rechenvorschrift für den Mittelwertfilter sieht wie folgt aus:

$$g(r,c) = \frac{1}{(2k+1)^2} \sum_{u=-k}^k \sum_{v=-k}^k f(r+u,c+v)$$

Die Maske für ein 3 X 3 großen Ausschnitt stellt sich dann so dar:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (k=1)$$

Der Nachteil des Mittelwertfilters besteht darin, dass die Kontraste sich vermindern und so eine Kantenverwischung entsteht.

#### *Medianfilter*

Das Ergebnisbild des Medianfilters ergibt sich als Median der Grauwerte im Bild f innerhalb eines bestimmten Ausschnitts der Größe n x n (wobei n= 2k+1).

Der Median wird berechnet, indem die  $(2k+1)^2$  Grauwerte der Größe nach sortiert werden und der  $(2k^2+2k+1)$ -te Wert der Median wird.

Im Beispiel für ein 3 x 3 Fenster in einem Bild wird nach der Formel für die Medianberechnung der 5. Wert der Median:

$$\left. \begin{array}{ccc} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 5 & 4 & 2 \end{array} \right\} \rightarrow 1, 1, 1, 1, 2, 2, 3, 4, 5$$

↑  
Median

Der Medianfilter ist ein nichtlinearer Operator und gehört zur Klasse der Rangordnungsoperatoren (rank order filtering), d.h. die Eingabedaten werden sortiert, und der mittlere Wert der aufsteigenden Folge wird als Median genommen.

Der Vorteil beim Medianfilter liegt darin, dass man als Ergebnis keine verwischten Kanten erhält. Aber Nachteil ist, dass die Ecken abgerundet und kleinere Objekte ausgelöscht werden. Zudem besitzt er eine höhere Rechenkomplexität als andere Operatoren, da die Werte immer der Größe nach sortiert werden müssen.

*Zusammenhang zwischen Mittelwert- und Medianfilter*

Der Mittelwert- und der Medianfilter können zusammen hergeleitet werden. Dabei wird für die Nachbarschaft:

$$f(r + u, c + v), -k \leq u, v \leq k$$

ein Grauwert  $f^*$  gesucht, der die  $(2k+1)^2$  Grauwerte am besten repräsentieren kann. Der Abstand zwischen den Grauwerten  $f^*$  und denen in der Nachbarschaft wird mit einer Abstandsfunktion  $d()$  definiert und so ergibt sich die Optimierungsaufgabe. Diese ermittelt den gemeinsamen kleinsten Abstand zwischen dem gesuchten Grauwert  $f^*$  und den Grauwerten aus der Umgebung. Mit verschiedenen Abstandsfunktionen erhält man nun die beiden verschiedenen Filter. Dabei zeigt sich, dass der Medianfilter gegenüber Ausreißern robuster ist als der Mittelwertfilter.



Quelle: Graber, Folie 76.

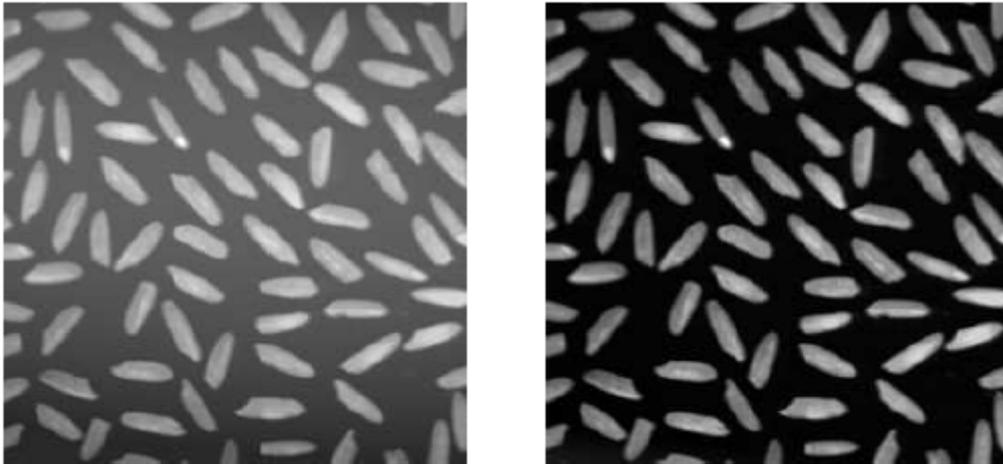
**Abb. 3.9:** Beispiel, gefiltert mit Mittelwert- und Medianfilter

### 3.3.2 Korrektur von uneinheitlichem Hintergrund

Ein uneinheitlicher Hintergrund kann unterschiedliche Ursachen haben, z.B. Beleuchtung oder Materialeigenschaft. Eine Verbesserungsmöglichkeit besteht in der Modellierung und der Subtraktion des Hintergrundes vom Bild. Dazu wird oft eine polynomiale Funktion der Form

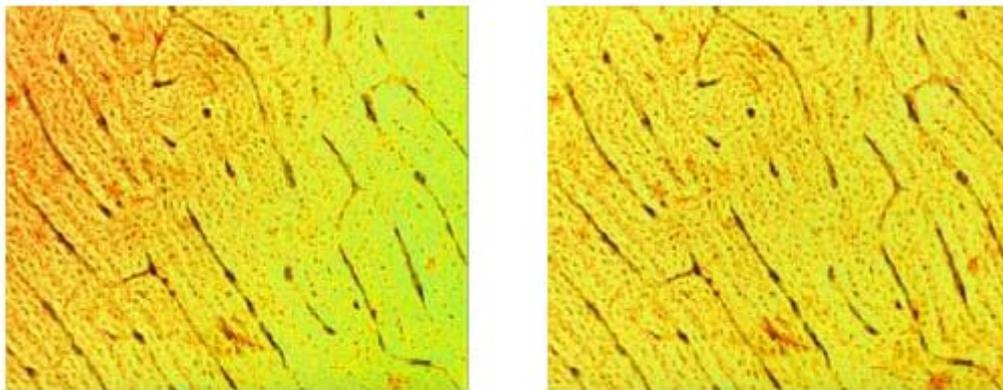
$$b(r,c)=a_{00} + a_{10}r + a_{01}c + a_{20}r^2 + a_{11}rc + a_{02}c^2$$

oder eine Funktion höherer Ordnung verwendet. Um die Parameter  $a_{ij}$  bestimmen zu können müssen ausreichend Hintergrundpunkte, mindestens aber 6 oder ein Vielfaches davon, manuell oder automatisch bestimmt werden. Eine Möglichkeit ist, das Originalbild in  $k \times k$  Subbilder aufzuteilen und für jedes Subbild den dunkelsten Punkt zu selektieren. Daraus ergeben sich  $k^2$  Hintergrundpunkte, für die anschließend die Funktionsapproximation durchgeführt wird.



Quelle: Jiang (2003), S. 54.

**Abb. 3.10:** Beispiel der Korrektur von uneinheitlichem Hintergrund (links: Original)



Quelle: Jiang (2003), S. 54.

**Abb. 3.11:** Beispiel der Korrektur von uneinheitlichem Farbton (links: Original)

### 3.4 Geometrische Transformationen

Bei geometrischen Transformationen nur die Position der Grauwerte verändert. Dabei verändert sich die räumliche Beziehung zwischen den Pixeln. Diese Funktion wird auch rubber-sheet-transformation genannt. Die geometrische Transformation bei der Bildverarbeitung besteht aus 2 Teilen. Erstens wird eine räumliche Transformation durchgeführt, zweitens eine Interpolation.

#### *Räumliche Transformation*

Diese Funktion wird verwendet für die Änderung, Neuordnung oder auch Umgestaltung der Pixel in der Bildebene. Beispiele sind:

### Translation

Verschiebung um gegebene x- und y-Distanzen  $t_x$  und  $t_y$ :

$$P = (x, y)$$

$$P' = (x', y') = (x + t_x, y + t_y)$$

$$x' = x + t_x$$

$$y' = y + t_y$$

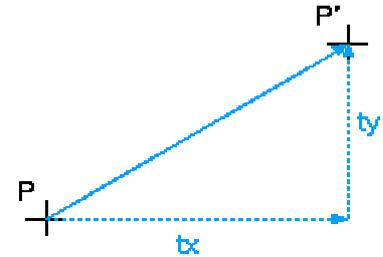


Abb. 3.12: Translation

### Skalierung

Ziel: Dehnung (bzw. Stauchung) um gegebenen Skalierungsfaktor  $s$  mit Zentrum = Ursprung

Je nach Skalierungsfaktor  $s$ :  $s < 1$ : Verkleinerung, Stauchung,  $s > 1$ : Vergrößerung, Dehnung,  $s = 1$ : keine Änderung

$$x' = x * s_x \quad \text{und} \quad y' = y * s_y.$$

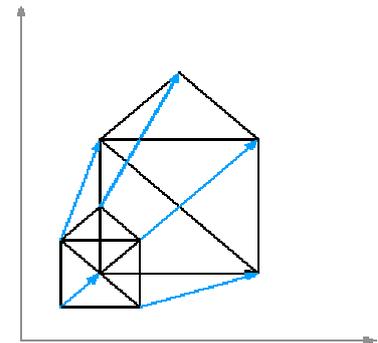


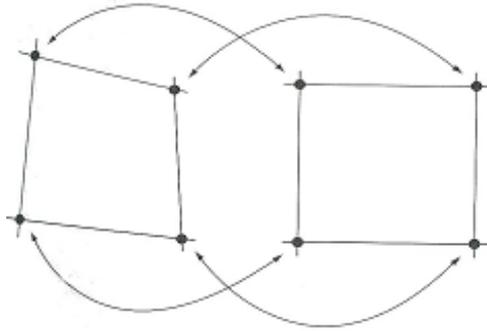
Abb. 3.13: Skalierung

Andere affine Transformationen sind z. B. Rotation, Scherung usw.

Ein Bild mit Pixel  $(x,y)$  durchlebt geometrische Verzerrungen. Die räumliche Transformation erzeugt ein neues Bild  $g(x',y')$  mit

$$x' = r(x,y) \quad \text{und} \quad y' = s(x,y).$$

Wenn  $r$  und  $s$  bekannt sind, kann man das frühere Bild wieder herstellen. Da man jedoch  $r$  und  $s$  meistens nicht kennt, benutzt man sog. tiepoints, Punkte die im Input- und Outputbild exakt bekannte Position haben.



Das Bild zeigt zwei korrespondierende Quadrate, das eine aus dem verzerrten Bild und das andere aus dem korrigierten Bild. Die Eckpunkte der Quadrate sind die korrespondierenden tiepoints.

Quelle: Gonzales, Woods (2002), S. 271.

**Abb. 3.14:** Tiepoints

$$x' = c_1x + c_2y + c_3xy + c_4$$

$$y' = c_5x + c_6y + c_7xy + c_8$$

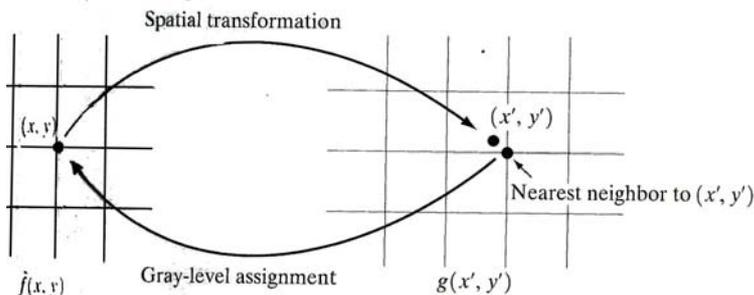
Mit acht tiepoints kann man nun die Koeffizienten berechnen. Diese bilden dann das geometrische Verzerrungsmodell und werden gebraucht um alle Pixel innerhalb des Quadrates zu berechnen. Die Auswahl der tiepoints erfolgt durch verschiedene Techniken, abhängig von der Anwendung.

Wenn man nun wissen will, welcher Wert im korrigiertem Bild an der Stelle  $(x_0, y_0)$  ist, muss man nur wissen, wo im verzerrtem Bild  $f(x_0, y_0)$  ist. Dafür muss man  $x_0$  und  $y_0$  in die Gleichungen einsetzen und bekommt dann die berechneten Werte  $x_0'$  und  $y_0'$ . Der Grauwert im verzerrtem Bild an der Stelle  $(x_0', y_0')$  ist dann  $g(x_0', y_0')$ . Dann ist  $f(x_0, y_0) = g(x_0', y_0')$ . Und so geht man dann Zeile für Zeile und Reihe für Reihe durch das Bild durch.

### *Interpolation*

Sind nun die Pixel berechnet, kommt es häufig vor, abhängig von den Koeffizienten, dass die Pixelwerte  $x'$  und  $y'$  keine Integer sind. Da die Grauwerte aber auf Pixelwerten basieren, die Integer sind, muss eine Grauwertinterpolation durchgeführt werden.

Die einfachste Variante ist dabei die Nearest Neighbour Interpolation.

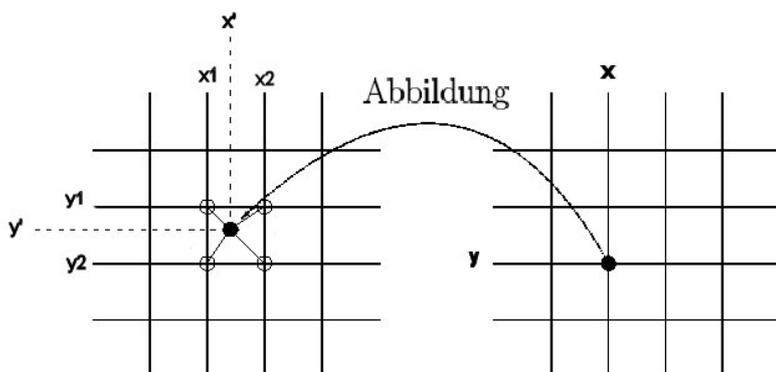


Quelle: Gonzales, Woods (2002), S. 273.

**Abb. 3.15:** Nearest Neighbour Interpolation

Hier wird der nächstgelegene Integer Nachbar zum berechneten Pixelwert ausgewählt und der Grauwert dieses Pixels wird dem Pixel zugewiesen. Diese Interpolation ist zwar einfach zu implementieren, aber sie erzeugt unerwünschte Nebeneffekte, z.B. die Verzerrung von glatten Ecken.

Eine weitere Interpolation ist die bilineare Interpolation. Hierbei werden die Grauwerte der vier nächsten Nachbarn mit in die Berechnung aufgenommen.



Quelle: Jiang (2003), S. 59.

**Abb. 3.16:** Bilineare Interpolation

Es gibt auch Interpolationen höherer Ordnung, d. mit Einbezug einer größeren Nachbarschaft, aber im allgemeinen hat die bilineare Interpolation ausreichende Qualität.

### 3.5 Bildverarbeitung im Frequenzraum

#### 3.5.1 Einführung in die Fourier Analyse

Neben der punktbasierten Betrachtungsweise existiert ebenfalls die Möglichkeit Bildverarbeitung im Frequenzraum durchzuführen. Dabei liegt die Idee zugrunde, ein Bild als Funk-

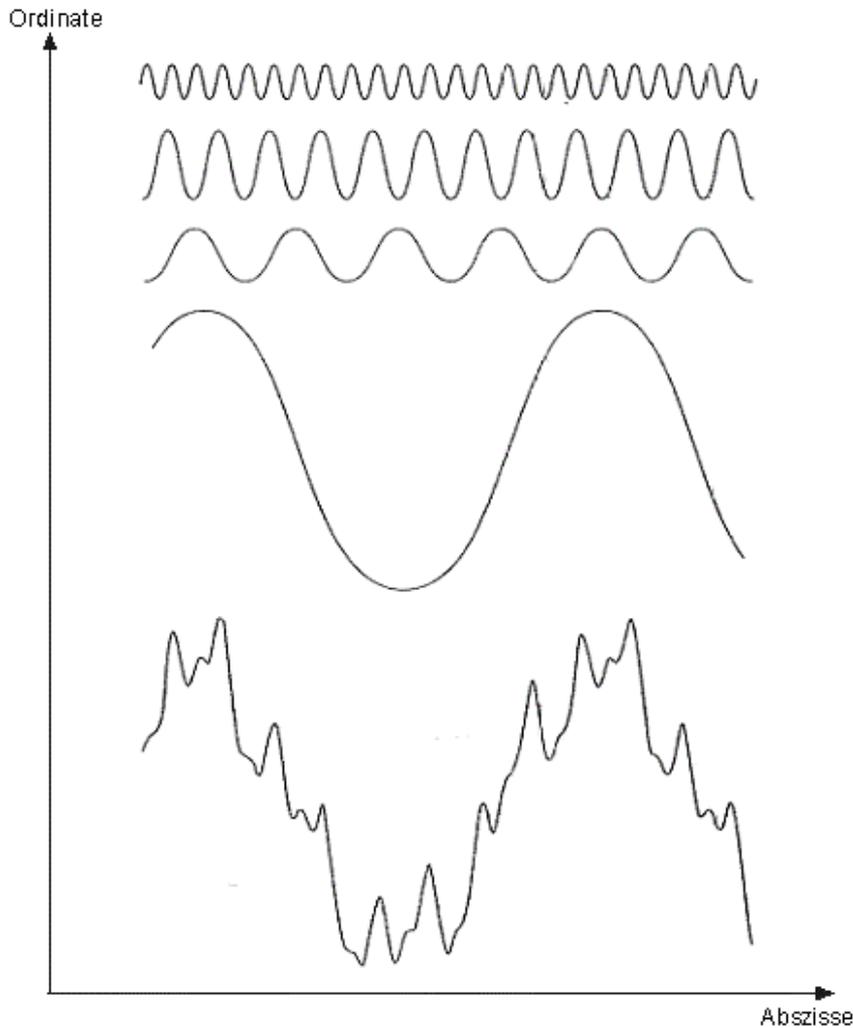
tion zu verstehen, die sich linear in periodische Funktionen separieren lässt – in sogenannte Schwingungen verschiedener Frequenzen. Ausgangspunkt ist das Buch „La Théorie Analytique de la Chaleur“ von Jean Baptiste Joseph Fourier. Dort entwarf er 1922 das revolutionäre Konzept, dass jede einigermaßen wohlgeformte periodische Funktion sich als gewichtete Summe von Sinuskurven verschiedener Frequenzen nachbilden lasse. Ursprünglich zur mathematischen Erklärung der Wärmeverbreitung eingesetzt, führte eine Erweiterung der Theorie zur Fourier Transformation, die insbesondere für die Bildverarbeitung von Interesse ist, da sie die Zerlegung finiter Funktionen unterstützt. Zuerst wird jedoch die ursprüngliche Definition Fouriers, die Zerlegung periodischer Funktionen, eingeführt. Sie wird auch als Fourier-Reihe bezeichnet:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[ a_k \cos\left(\frac{2\pi kx}{T}\right) + b_k \sin\left(\frac{2\pi kx}{T}\right) \right], \text{ mit } T \text{ als Periode der Funktion } f(x).$$

Die Funktion  $f(x)$  ist dabei definiert als unendliche Summe von Sinus- und Kosinusschwingungen, die jeweils mit eigenen Koeffizienten  $a_k$  und  $b_k$  gewichtet sind. Diese ergeben sich aus Integralen über die eigentliche Funktion und sollen hier nicht weiter vertieft werden. Entscheidend ist, dass sich aus sämtlichen periodischen Schwingungen ohne Informationsverlust die ursprüngliche Funktion herstellen lässt. Es kann dadurch erklärt werden, dass keine Interdependenzen zwischen den einzelnen Schwingungen bestehen, wie z. B. multiplikative Komponenten. Die Menge an Frequenzen (periodische Sinus-/ Kosinusschwingungen), aus denen die Funktion  $f(x)$  aufgebaut werden kann, spannen dabei den Frequenzraum von  $f(x)$  auf. Analog wird jede Schwingung als Frequenzkomponente des Frequenzraumes von  $f(x)$  bezeichnet. Wird nun ein Bild als zerlegbare Funktion verstanden, so ist es prinzipiell möglich, es aus dem Bildraum oder der Bilddomäne in den Frequenzraum zu konvertieren und nach Bearbeitung wieder in den Bildraum zurückzu-konvertieren.

Warum verdient aber der Frequenzraum überhaupt Interesse in der Bildverarbeitung? Das Beispiel in Abbildung 3.5 stellt eine einfache Zerlegung einer periodischen Funktion in vier Frequenzkomponenten dar und kann exemplarisch als eindimensionales Bild verstanden werden, wobei von der diskreten Bildbetrachtung zu abstrahieren ist. Interpretiert man die Ordinate als Farbintensitätswert, so ergibt sich eine erstaunliche Beobachtung. Die Schwingungskomponenten hoher Frequenz, d. h. einer häufigen Wiederholung der periodischen Schwingung innerhalb eines festen Abschnittes, tragen zu feinen Details von Farbübergängen bei, während die niedrigen Schwingungskomponenten grobe Farbverläufe skizzieren. Anders ausgedrückt: die hochfrequenten Komponenten beeinflussen kontrastreiche Regionen, während Tieffrequenzanteile Bereiche geringer Intensitätsvariation bestimmen. Jede Frequenzkomponente kann somit als eine globale Eigenschaft des Bildes verstanden werden, unabhängig von der Lokalität der betrachteten Kontrasteigenschaft.

Der Frequenzraum ermöglicht damit die globale Bearbeitung von Bildeigenschaften im Gegensatz zu den lokalen Nachbarschaftsoperationen in der Bilddomäne. Aus Anschauungsgründen wird nachfolgend mit Graubildern gearbeitet, da der Unterschied zu Farbbildern einzig darin besteht, die Filteroperation auf sämtliche Farbkanäle (z. B. beim RGB-Modell) anzuwenden.



Quelle: o.A.

**Abb. 3.17:** Zerlegung einer periodischen Funktion als Fourier-Reihe

### 3.5.2 Die Fourier Transformation und ihre Eigenschaften

Da in der Bildverarbeitung die Anforderung gestellt wird, mit finit diskreten Funktionen zu arbeiten, beschränken wir uns im weiteren auf die nützlichere Fourier-Transformation, die sich aus der Fourier-Reihe ergibt, wenn ein Grenzübergang  $\rightarrow \infty$  für die Periode  $T$  durchgeführt wird. Die Fourier-Transformation hebt damit die Beschränkung auf, für Funktionen den Nachweis der Periodizität zu erfüllen und fordert lediglich reelle Funktionen. Für den Fall einer eindimensionalen kontinuierlichen Variablen  $x$  besitzt sie die Form:

$F(u) = \int_{-\infty}^{\infty} f(x) \exp(-j2\pi ux) dx$ , mit  $j = \sqrt{-1}$ , und der Umkehrung:

$$f(x) = \int_{-\infty}^{\infty} F(u) \exp(j2\pi ux) du.$$

Anstelle der Sinus- und Kosinuscurven wird die komplexe Schreibweise verwendet, die auf die Euler-Identität  $\exp(j\theta) = \cos \theta + j \sin \theta$  zurückgeht. Die Sinuskomponenten sind dabei den imaginären Zahlanteilen zugeordnet, während die Kosinuskomponenten den Realzahlanteil darstellen. Für die Bildverarbeitung ist aber eine diskrete Summierung statt der Integration notwendig. Ein Bild der Auflösung  $M \times N$  stellt eine diskrete Funktion  $f(x, y)$  dar, mit  $x \in \{0, \dots, M-1\}$  und  $y \in \{0, \dots, N-1\}$ . Dies erfordert die diskrete Fourier-Transformation, wobei wir uns zur besseren Anschaulichkeit wieder auf den eindimensionalen Fall beschränken:

$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) \exp(-j2\pi ux / M)$ , und der Umkehrung:

$$f(x) = \sum_{u=0}^{M-1} F(u) \exp(j2\pi ux / M), \text{ für diskrete } u, x \text{ von } 0 \text{ bis } (M-1).$$

Die Berechnung von  $F(u)$  erfolgt für jedes  $u \in M$  getrennt. Damit besteht die diskrete Fourier-Transformation (DFT) aus  $M$  Termen, die den Frequenzraum aufspannen. Jede dieser Komponenten setzt sich wiederum aus einer Summe von  $M$  Sinus-/ Kosinustermen zusammen. Aus den  $M$  getrennten Komponenten wird dann bei der Inversion die ursprüngliche Funktion des Bildraumes wiederhergestellt, ohne Informationsverlust. Man sagt auch, dass für sämtliche  $u \in M$   $F(u)$  den **Frequenzraum** aufspannt, mit den einzelnen  $F(u)$  als **Frequenzkomponenten**. Die Fourier-Transformation kann auch als ein Glasprisma verstanden werden, welches Licht (analog der Bildfunktion) in seine Bestandteile (den Frequenzkomponenten) zerlegt.

Wie werden nun die Frequenzkomponenten zur Bildverarbeitung genutzt? Ausgangspunkt ist die Betrachtung der komplexen Frequenzkomponenten  $F(u) = R(u) + jI(u)$  (mit  $j = \sqrt{-1}$ ,  $R(u)$  als Realteil,  $I(u)$  als Imaginärteil) als Polarvektoren  $F(u) = |F(u)| \exp(-j\phi(u))$ , mit dem Betrag  $|F(u)| = [R^2(u) + I^2(u)]^{0.5}$  und dem (Phasen-)winkel  $\phi(u) = \tan^{-1}(I(u) / R(u))$ . Der Betrag der Polarvektoren wird auch als Spektrum der Fourier-Transformation bezeichnet. Für die folgenden Filter wird ausschließlich das Spektrum genutzt, da z. B. Kontrast durch Helligkeitsunterschiede und nicht durch Farbwerte bestimmt wird (s. Anhang A.a).

Vor der Behandlung einiger Filter wird nun die zweidimensionale DFT mit einigen weiteren Eigenschaften vorgestellt, wobei ein Bild der Auflösung  $M \times N$  zugrunde liegt:

$$(1) \quad F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp[-j2\pi(ux/M + vy/N)], \text{ mit der Umkehrung:}$$

$$(2) \quad f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp[j2\pi(ux/M + vy/N)].$$

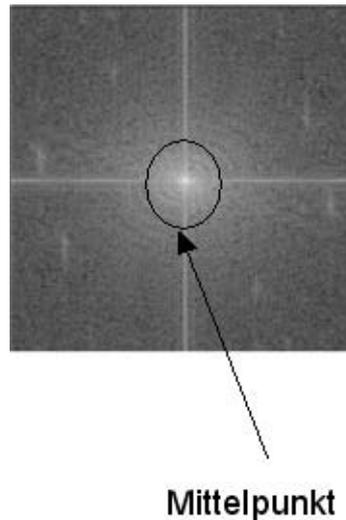
Dem Bild im Frequenzraum steht ein ebenso großes Bild im Frequenzraum gegenüber, mit  $u, v$  als Frequenzraumkoordinaten. Große Werte von  $u, v$  bezeichnen hohe Frequenzkomponenten, die ergo kontrastreiche Regionen bestimmen, wie z. B. Ecken und Kanten. Dagegen spezifizieren geringe  $u, v$  Werte geringe Intensitätsvariationen, wie beispielsweise einheitliche Zimmerwände. Der Wert am Ursprung  $(0, 0)$  entspricht dem mittleren Grauwert des Bildes, da:

$$F(0, 0) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp[-j2\pi(\underbrace{0x/M}_0 + \underbrace{0y/N}_0)]$$

$\underbrace{\hspace{10em}}_1$

$$F(0, 0) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \text{ (siehe mittlerer Grauwert in lokalen Bildoperationen).}$$

Vor der Bearbeitung eines Bildes in der Frequenzdomäne wird es jedoch so zentriert, dass die Koordinaten  $(M/2, N/2)$  dem Ursprung entsprechen, d.h. den mittleren Grauwert abbilden. Dafür wird  $f(x, y)$  im Bildraum transformiert mit  $f(x, y) \leftarrow f(x, y)(-1)^{(x+y)}$ . Eine Eigenschaft des Spektrums bei reellem  $f(x, y)$  ist die Symmetrie am Ursprung  $(M/2, N/2$  nach Transformation). Dies ergibt sich wegen der Gültigkeit von  $F(u, v) = F^*(-u, -v)$  und der Gleichheit des Betrages einer Konjugierten  $[F^*(u, v)]$  und seines Originals  $[F(u, v)]$ . Die Konjugierte stellt lediglich eine Spiegelung an der reellen Achse dar. Beides, sowohl die Zentriertheit, als auch die Punktsymmetrie, wird in Abb. 3.18 deutlich. Die Bereiche des Frequenzspektrums nahe des Mittelpunktes entsprechen hierbei den kontrastarmen Regionen niedriger Frequenz, mit dem mittleren Grauwert als schwingungsärmstes Zentrum. Radial vom Zentrum nach außen hin nehmen dann die Frequenzkomponenten an Kontrast zu. Diesen Zusammenhang machen sich Frequenzraumfilter zunutze, die kontrastreiche oder –arme Bildbereiche herausfiltern möchten.



**Abb. 3.18:** Zentriertes Spektrum eines Bildes

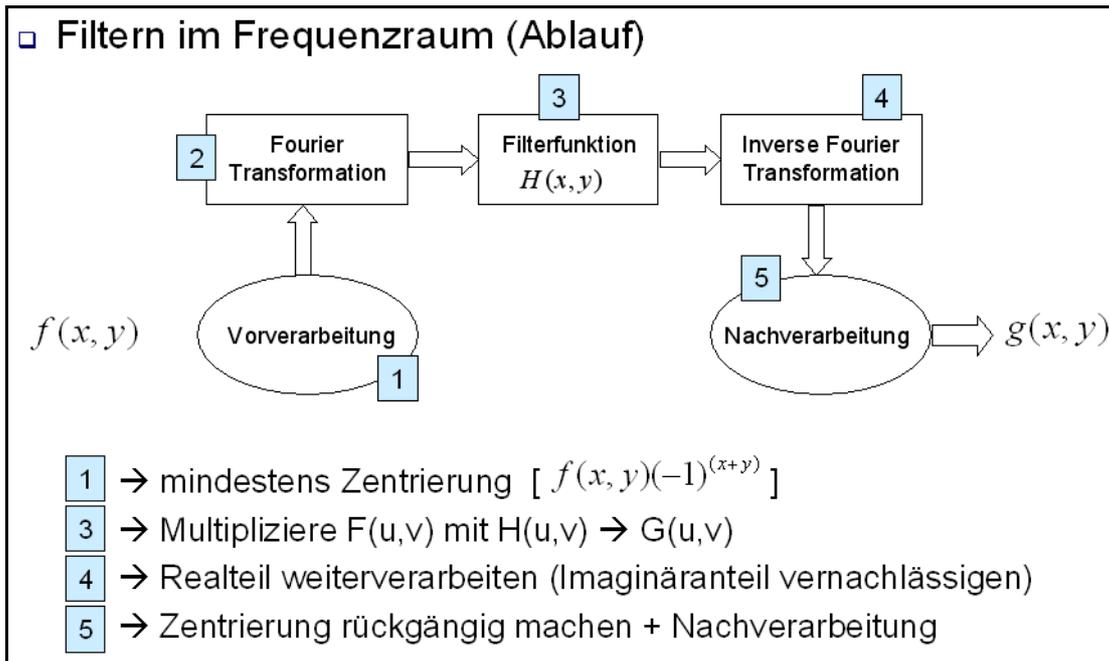
### 3.5.3 Filtern im Frequenzraum

Der Ablauf des Filterns im Frequenzraum ist in Abb. 3.19 dargestellt. Die bereits erwähnte Zentrierung ist Teil der Vorverarbeitung im Bildraum (s. Schritt 1). Hierzu gehört ebenfalls die Begrenzung von Grauwerten u. ä. Operationen.

Anschließend wird das Bild in den Frequenzraum konvertiert (s. Schritt 2). Dabei wird in der Praxis die Fast Fourier Transformation (FFT) angewandt, die mit  $O(M \log_2 M)$  Multiplikationen/ Additionen statt  $O(M^2)$  bei der konventionellen Summenbildung (s. Gleichung 1 und 2) auskommt.

Dann wird eine Filterfunktion  $H(x, y)$  mit  $F(u, v)$  elementweise multipliziert, d. h. das erste Element von  $H$  mit dem ersten Element von  $F$  usw. (s. Schritt 3). In diesem Aufsatz werden ausschließlich reelle Funktionen als Filter verwendet, da sie i) zu keiner Phasenverschiebung bei der Filterung führen, und ii) bei der Rückkonvertierung (s. Schritt 4) in den Bildraum ebenfalls zu einer reellen Funktion führen. Etwaige komplexe Zahlanteile werden ignoriert, da sie auf Rundungsfehlern basieren. Die fehlende Phasenverschiebung erklärt sich dadurch, dass der reelle Filter  $H(x, y)$  sowohl mit dem reellen, als auch mit dem imaginären Anteil von  $F(u, v)$  gleichartig multipliziert wird, d. h. das Verhältnis der Real- und Imaginärkomponente nach der Filterung bleibt gleich - ergo gleicher Phasenwinkel.

Als letzter Schritt muss die Zentrierung rückgängig gemacht werden. Anschließend kann eine Nachverarbeitung folgen (s. Schritt 5), wie beispielsweise eine Normierung von Grauwerten oder die Zusammenführung der Bildkanäle bei Farbbildern.



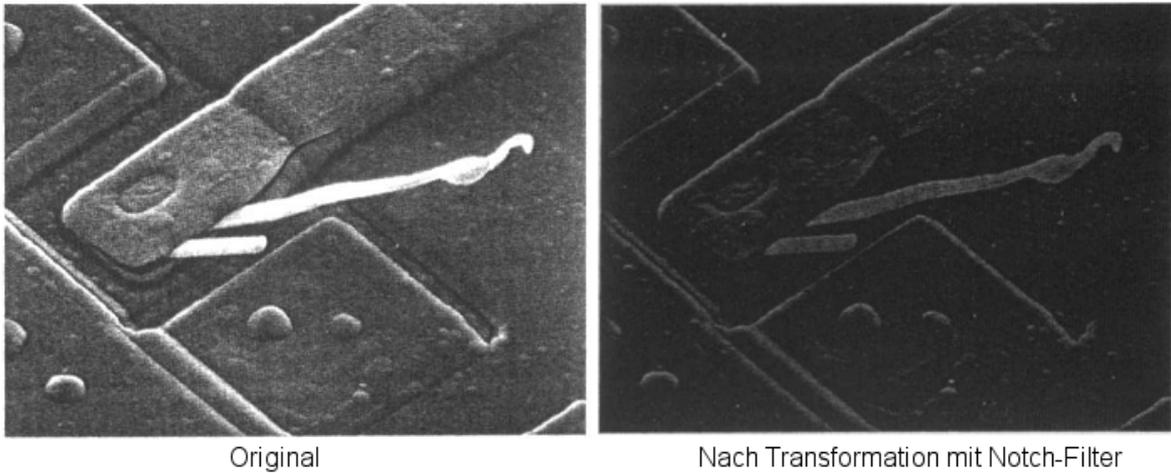
Quelle: eigene Darstellung nach Gonzales, Woods (2002), S. 159.

**Abb. 3.19:** Vorgang der Filterung im Frequenzraum

Die Funktion des ersten zu betrachtenden Filters, dem Notch-Filter, ist, ein begrenztes Intensitätsspektrum des Bildes hindurchzulassen. Im dargestellten Beispiel wird der mittlere Grauwert, d. h. das Zentrum des Frequenzraums, auf Null gesetzt. Der Filter hat dann die folgende Form:

$$H(u, v) = \begin{cases} 0 & \text{if } (u, v) = (M/2, N/2) \\ 1 & \text{otherwise} \end{cases}$$

In Abb. 3.20 wird deutlich, wie die Farbintensität des gesamten Bildes herabgesetzt wird, da der mittlere Grauwert, d. h.  $F(0, 0)$ , eine globale Eigenschaft des Bildes in der Bilddomäne darstellt.



Quelle: Gonzales, Woods (2002), S. 157-161.

**Abb. 3.20:** Der Notch-Filter

Da die Veränderung des mittleren Grauwertes ebenfalls als eine lokale Grauwerttransformation bekannt ist, stellt sich die Frage, ob und wie ein Zusammenhang zwischen Filtern im Bild- und Frequenzraum hergestellt werden kann. Eine Antwort liefert folgendes Faltungstheorem:

$$f(x, y) * h(x, y) \Leftrightarrow F(u, v)H(u, v)$$

$$\underbrace{f(x, y)h(x, y)}_{\text{Bildraum}} \Leftrightarrow \underbrace{F(u, v) * H(u, v)}_{\text{Frequenzraum}}, \text{ mit } f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n)h(x - m, y - n)$$

und  $\Leftrightarrow$  als Fourier/ inverse Fourier Transformation. Da Nachbarschaftsoperatoren [Bildraumfilter  $h(x, y)$ ] via Faltung [ $f(x, y) * h(x, y)$ ] auf die Bildfunktion [ $f(x, y)$ ] angewandt werden, lässt sich aus der ersten Zeile des Theorems folgern, dass zu jedem Bildraumfilter ein entsprechender Frequenzraumfilter existieren muss. Die Ergebnisse der Filterung im Frequenz-, als auch im Bildraum sind dabei über die Fourier-Transformation bzw. ihre Inverse ineinander überführbar. Dadurch wird deutlich, dass die Multiplikation im Frequenzraum der Faltung im Bildraum entspricht und umgekehrt.

Es kann darüber hinaus gezeigt werden, dass gilt:  $h(x, y) \Leftrightarrow H(u, v)$ .<sup>11</sup> Damit kann über die Fourier-Transformation aus einem Bildraumfilter der korrespondierende Frequenzraumfilter erzeugt werden und vice versa. Frequenzraumverarbeitung findet damit Verbindung mit der Bildraumverarbeitung, indem man statt getrennter Filter das **Fourier-Transformationspaar**  $h(x, y); H(u, v)$  betrachtet, dass dieselbe Filterfunktion in beiden Domänen erfüllt.

<sup>11</sup> Anm.: Der Beweis wird mit einer Impulsfunktion geführt, die den Einfluss der Filter auf einen Punkt beschränkt und kann ausführlich in Gonzales, Woods (2002) auf S. 163-164.

Wofür ist dann die Frequenzraumverarbeitung gut, wenn doch Filter aus dem Bildraum dieselbe Funktion erfüllen?

Erstens gibt es über die einfachen Filter hinaus Verfahren zur Bildrestaurierung, die allein im Frequenzraum durchgeführt werden können. Auf den Wiener Filter als ein Beispiel wird dabei im nächsten Abschnitt eingegangen.

Zweitens konnte nachgewiesen werden, dass mit der FFT Filterung im Frequenzraum effizienter ist als im Bildraum, wenn die Zahl der Punkte  $>32$  ist.<sup>12</sup> Die meisten Bilder liegen weit über dieser Grenze (z. B.  $800 \times 600$  Pixel eines normalen Bildes), weshalb die Frequenzraumfilterung meist schneller ist.

Nicht zuletzt sind die Filter im Frequenzraum eher der Intuition zugänglich. Dies soll am Beispiel des Gauß-Hochpassfilters deutlich gemacht werden (s. Abb. 3.21, rechte Seite). Hochpassfilter haben die Aufgabe, kontraststarke Regionen (hohe Frequenzen - weit entfernt vom Zentrum  $F(0,0)$ ) zu verstärken, während sie Regionen mit geringer Intensitätsvariation abschwächen (niedrige Frequenzen - in der Nähe vom Zentrum  $F(0,0)$ ). Es ist nicht intuitiv, dass der Bildraumfilter von einem positiven Ausschlag im Zentrum erst stark ins Negative verläuft, um sich dann langsam wieder der Abszisse anzunähern (rechts unten). Dagegen ist im Frequenzraum (rechts oben) eindeutig klar, dass die dem Zentrum nahe liegenden niederfrequenten Anteile abgeschwächt werden, während die außen liegenden kontrastreichen Gebiete vollständig durchgelassen werden.

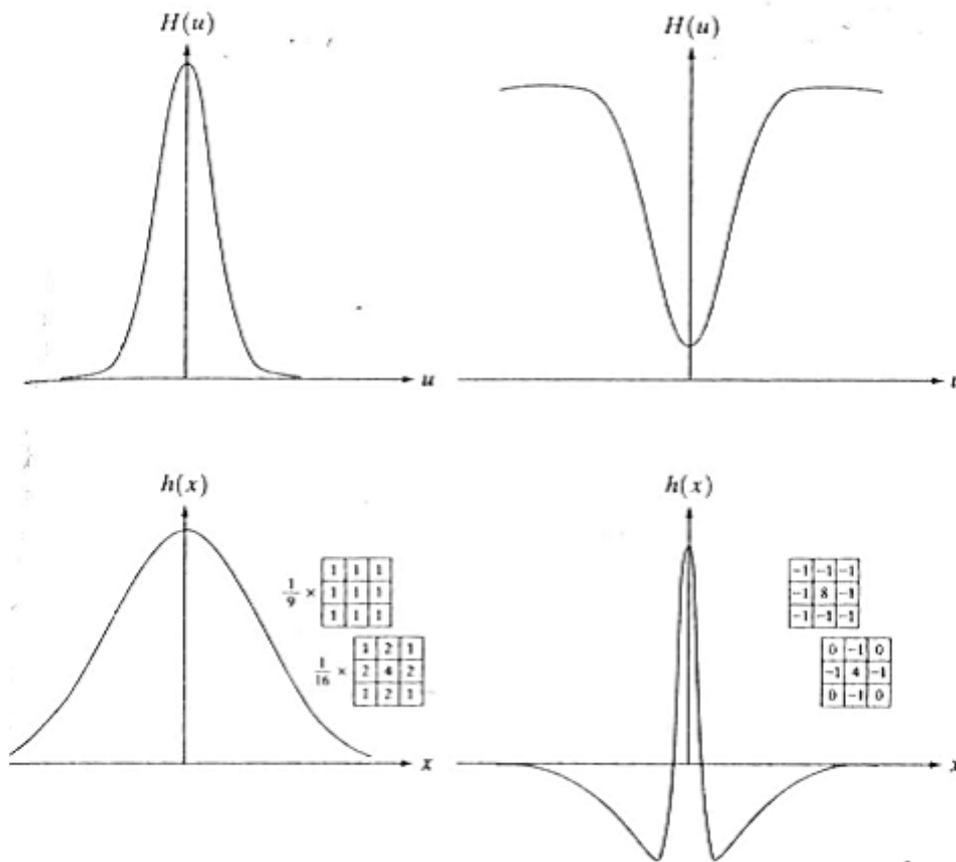
Neben dem Gauß-Hochpassfilter existiert ebenfalls der Gauß-Tiefpassfilter. Tiefpassfilter haben die umgekehrte Aufgabe von Hochpassfiltern. Sie lassen niederfrequente Bildanteile durch, während kontrastreiche Bildbereiche abgeschwächt werden (siehe ebenfalls Abb. 3.21). Am Gauß-Tiefpassfilter soll eine fundamentale Beziehung der Reziprozität des Transformationspaares deutlich gemacht werden. Der Gauß-Tiefpassfilter wird definiert als:

$$H(u) = A \exp(-u^2 / 2\sigma^2) \quad (\text{im Frequenzraum}),$$

$$h(x) = \sqrt{2\pi}\sigma A \exp(-2\pi^2\sigma^2 x^2) \quad (\text{im Bildraum}).$$

---

<sup>12</sup> Vgl. Gonzales, Woods (2002), S. 166.

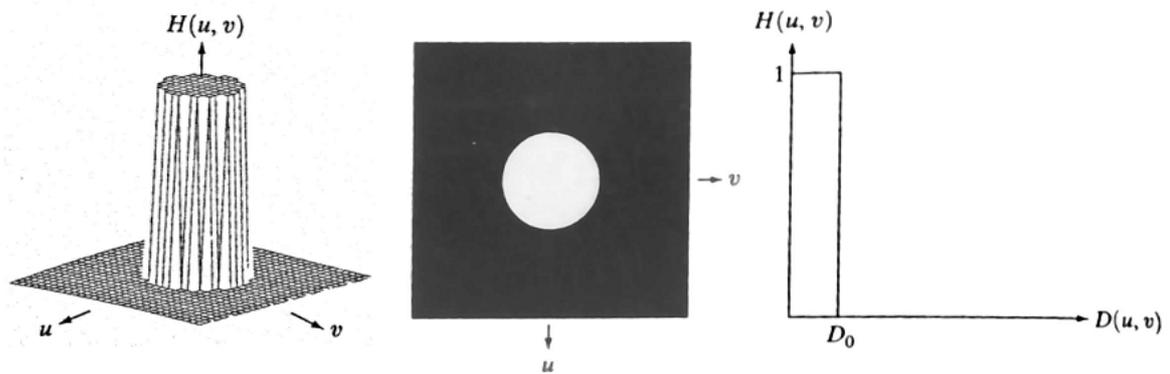


Quelle: Vgl. Gonzales, Woods (2002), S. 165.

**Abb. 3.21:** Gauß-Hoch- und Tiefpassfilter im Bild- und Frequenzraum

Die Breite der glockenähnlichen Kurvenform wird durch die Variabilität  $\sigma$  spezifiziert (diese kann mit dem später verwendeten Begriff der Cutoff-Frequenz gleichgesetzt werden). Auffällig ist, dass einer schmalen Glocke im Bildraum eine breite Kurve im Frequenzraum entspricht. Dies ist der Fall, da  $\sigma$  einerseits im Zähler liegt, im anderen Fall dagegen im Nenner. Wird der Gauß-Tiefpassfilter im Bildraum verbreitert ( $\sigma$  steigt), so verhält sich der Gauß-Tiefpassfilter im Frequenzraum **reziprok** und wird schmaler (steigendes  $\sigma$  im Nenner  $\rightarrow$  gegenläufiges Verhalten). Dieses reziproke Verhalten ist sämtlichen Filtern inhärent und eine wesentliche Eigenschaft des Fourier-Transformationspaars.

Neben dem Gauß-Filter existieren ideale Hoch- und Tiefpassfilter. Statt eines stetigen Übergangs von durchlässigen zu undurchlässigen Frequenzen bestimmt eine **Cutoff-Frequenz**  $D_0$  einen abrupten Übergang der durchzulassenden Frequenzen (siehe Abb. 3.22).

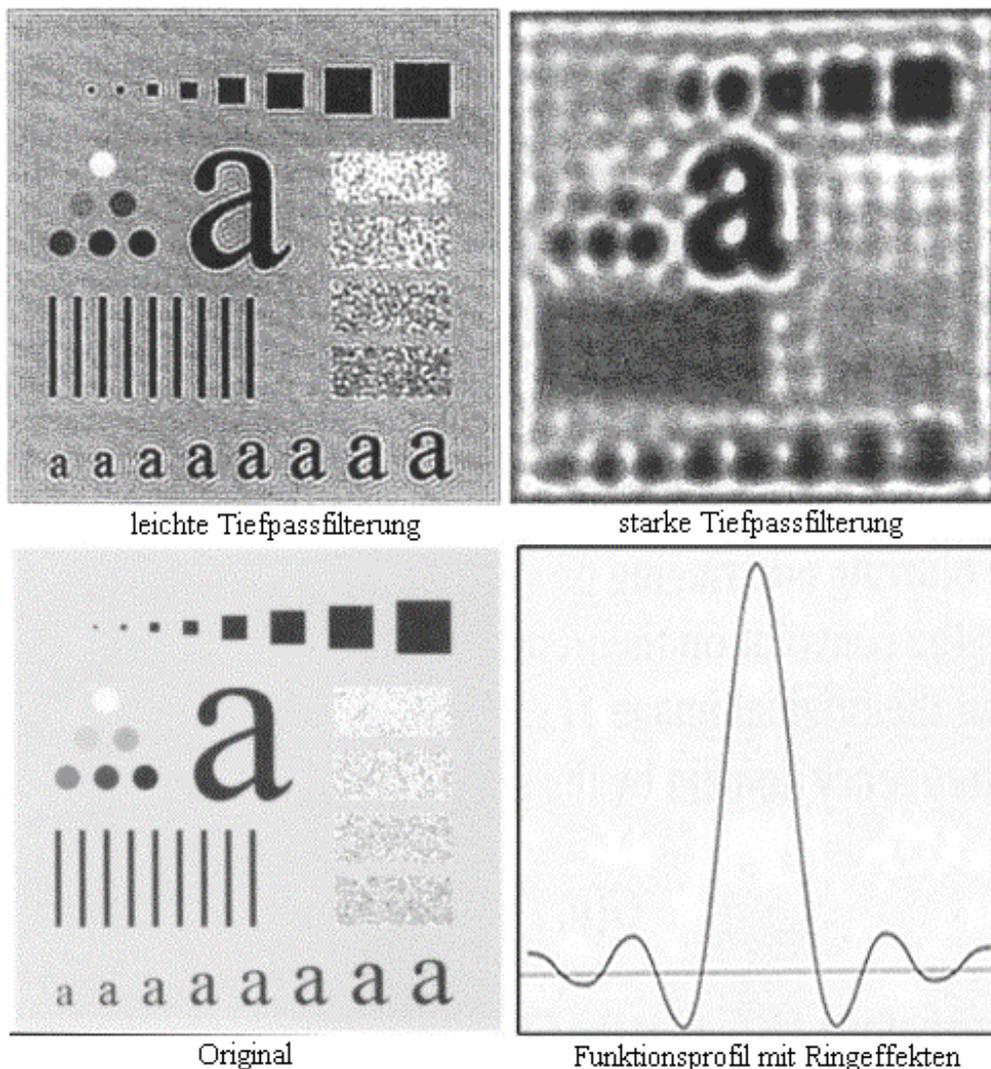


Quelle: Gonzales, Woods (2002), S. 168.

**Abb. 3.22:** Der ideale Tiefpassfilter in der Frequenzdomäne

Der ideale Tiefpassfilter hat die Form : 
$$H(u, v) = \begin{cases} 1 & \text{if } (D(u, v) \leq D_0) \\ 0 & \text{if } (D(u, v) > D_0) \end{cases}$$

Nur die Frequenzen mit der Entfernung  $D(u, v)$  zum Zentrum, die innerhalb der Cutoff-Frequenz  $D_0$  liegen, werden dabei durchgelassen. Allerdings bringt der ideale Tiefpassfilter Probleme mit sich, da bei abnehmender Cutoff-Frequenz, also stärkerem Herausfiltern der hohen Frequenzen, ringförmige Anomalien entstehen (siehe Abb. 3.11 oben). Dies ist leicht zu erklären, da der räumlich transformierte Filter eine glockenförmig abfallende Kurve darstellt, die in negativen und positiven Randschwingungen ausläuft, ähnlich den Wellen bei einem Stein, der ins Wasser geworfen wird. Die Randschwingungen treten dabei umso stärker auf, je schmaler der ideale Tiefpassfilter im Frequenzraum ist (Reziprozität). Dies ist im Funktionsverlauf in Abb. 3.23 rechts unten zu erkennen.



Quelle: Vgl. Gonzales, Woods (2002), S. 170ff.

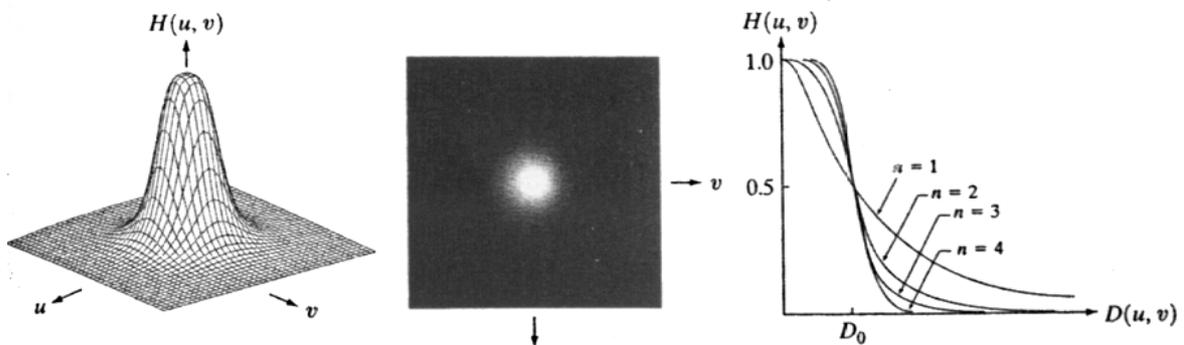
**Abb. 3.23:** Anwendung des idealen Tiefpassfilters mit Ringeffekten

Neben dem Gauß-Tiefpassfilter (siehe Abb. 3.21, links) gibt es einen weiteren Filter, der diese Nachteile vermeidet, den Butterworth-Tiefpassfilter (s. Abb. 3.24). Er wird definiert als:

$$H(u,v) = \frac{1}{1 + [D(u,v) / D_0]^{2n}}, \text{ mit } D(u,v), D_0 \text{ wie beim idealen Tiefpassfilter.}$$

Die Cutoff-Frequenz  $D_0$  kann im Gegensatz zum idealen Tiefpassfilter nicht als Grenzübergang verstanden werden, sondern wird als Ort definiert, bei dem die Funktion zu einem gewissen Grad abgefallen ist. Dies ist bei der vorigen Formel ein 50 % Abfall, unabhängig davon, welcher Wert für  $n$  gewählt wird, da mit  $D(u,v) = D_0$  der Exponentialterm immer 1 ergibt (s. Schnittpunkt im rechten Funktionsverlauf in Abb. 3.24). Der Butterworth-Tiefpassfilter kann durch die Wahl von  $n$  verschieden stark abfallen (siehe ebenfalls Abb. 3.24, rechts). Wie der Funktionsverlauf zeigt, nähert sich der Filter dem idealen Tiefpassfilter an, je größer die Werte für  $n$  als Parameter gewählt werden. Wird anderer-

seits  $n=1$  gesetzt, so weist der Butterworth-Tiefpassfilter Ähnlichkeit mit dem Gauß-Filter auf.



Quelle: Vgl. Gonzales, Woods (2002), S. 173.

**Abb. 3.24:** Der Butterworth-Tiefpassfilter

Sowohl der Gauß, als auch der ideale und der Butterworth-Tiefpassfilter können als Hochpassfilter aus den zugehörigen Tiefpassfiltern auf folgende Weise bestimmt werden:

$$H_{hp}(u, v) = 1 - H_{lp}(u, v).$$

$H_{hp}$  bezeichnet dabei den Hochpassfilter des zugehörigen Tiefpassfilters  $H_{lp}$ . Die Probleme des idealen Tiefpassfilters treten ebenso beim idealen Hochpassfilter auf, lediglich in invertierter Form. Es konnte ebenfalls gezeigt werden, dass der Butterworth-Filter das Bindeglied zwischen dem idealen und dem Gauß-Filter darstellt, indem der Parameter  $n$  die Steilheit des Funktionsverlaufs bestimmt und damit beide Filterarten annähern kann.

Abschließend werden einige andere Filter im Frequenzraum vorgestellt. Grundsätzliches Problem der Hoch-/ Tiefpassfilter ist, dass sie das Bild verdunkeln. Besonders beim Hochpassfilter wird der mittlere Grauwert nahezu weggefiltert (s. Abb. 3.21, rechts oben im Zentrum). Tiefpassfilter reduzieren zwar lediglich hochfrequente Anteile, der Effekt bleibt aber trotzdem beobachtbar. Lösung ist das **Unsharp Masking**. Ein hochpassgefiltertes Bild wird z. B. erzeugt, indem vom Original eine tiefpassgefilterte Version subtrahiert wird.

Eine Verallgemeinerung ist das **High-boost** Filtern, bei dem das Original vor Subtraktion gewichtet wird und damit der Beitrag des Originals in seiner Intensität erhöht werden kann. Im folgenden Abschnitt wird im Zusammenhang mit der Bildrestaurierung auf den Wiener Filter eingegangen, der ebenfalls auf der Frequenzraumverarbeitung basiert.

## 4 Bildrestaurierung

Die Bildrestaurierung hat die Aufgabe ein Bild, bei dem während der Bildaufnahme Störungen aufgetreten sind, zu rekonstruieren oder wiederzugewinnen. Bei den Restaurations-techniken wird die Verschlechterung modelliert und dann der umgekehrte Prozess angewendet, um das Originalbild wiederzugewinnen.

Die äußere Erscheinung des Bildes wird verbessert, jedoch objektiv, da die Restaurations-techniken auf mathematischen oder Wahrscheinlichkeitsmodellen basieren.

Bei der Bildrestaurierung besitzt man ein a-priori-Wissen über die Art der vorliegenden Störung. A-priori-Wissen über die Bildstörung wird entweder vor oder während der Aufnahme gewonnen (z. B. Messung der Richtung und Geschwindigkeit einer Objektbewegung). Je besser diese bekannt ist, desto genauer ist das Ergebnis der Restaurierung.

Mögliche Ursachen für Bildstörungen sind:

- Fehler bei der Bildaufnahme bzw. -übertragung
- Verschmierung aufgrund von Bewegungen oder Vibrationen der Aufnahmegeräte bzw. Objekte
- Defokussierung
- atmosphärische Turbulenzen (bei Luftbildern und in der Astronomie)
- additive Signale, welche die Helligkeits- und Kontrastwerte verändern

Die Störungen werden durch Messen oder Schätzen separat mathematisch modelliert. Um die Verschlechterungsfunktion zu schätzen, gibt es drei Wege: Bildbeobachtung, Schätzen durch Experimente und mathematisches Modellieren. Durch Anwendung der inversen Störfunktion wird versucht, das Originalbild ohne Störungen möglichst gut zu rekonstruieren. Die beste Restaurierung kann erreicht werden, wenn ein stochastisches Modell zu Grunde liegt, wie z.B. beim Wiener Filter.

### *Verschlechterungsprozess*

Es wird angenommen, dass man die Störung durch ein lineares System beschreiben kann. Der Verschlechterungsprozess wird als Verschlechterungsfunktion  $h$  modelliert, die zusammen mit einer additiven Störungsfunktion  $\eta(x,y)$  (die z.B. bei der Bildübertragung oder Bildgewinnung auftritt) auf einem Eingabebild  $f(x,y)$  arbeitet, um ein verschlechtertes Bild  $g(x,y)$  zu produzieren. Ist nun ein Bild  $g(x,y)$ , ein wenig Wissen über die Verschlechter-

rungsfunktion  $h$  und über die additive Störungsfunktion  $\eta(x,y)$  gegeben, ist die Aufgabe der Restauration eine Schätzung  $\hat{f}(x,y)$  des Originalbildes zu bekommen. Je mehr über die Störungen  $\eta$  und  $h$  bekannt ist, um so näher kommt die Schätzung von  $\hat{f}(x,y)$  an das Original  $f(x,y)$  heran.

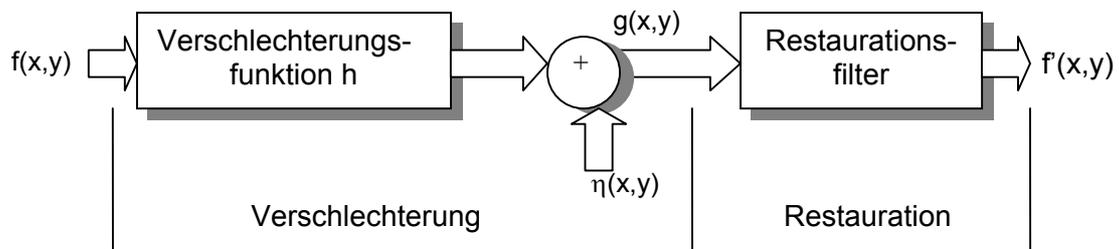
In der räumlichen Domäne sieht das so aus:

$$g(x,y) = h(x,y) * f(x,y) + \eta(x,y) (\eta(x,y) \text{ kann auch } = 0 \text{ sein})$$

In der Frequenzdomäne ergibt sich diese Formel:

$$G(u,v) = H(u,v) F(u,v) + N(u,v)$$

In der Frequenzdomäne entspricht der Faltung (\*) die Multiplikation.

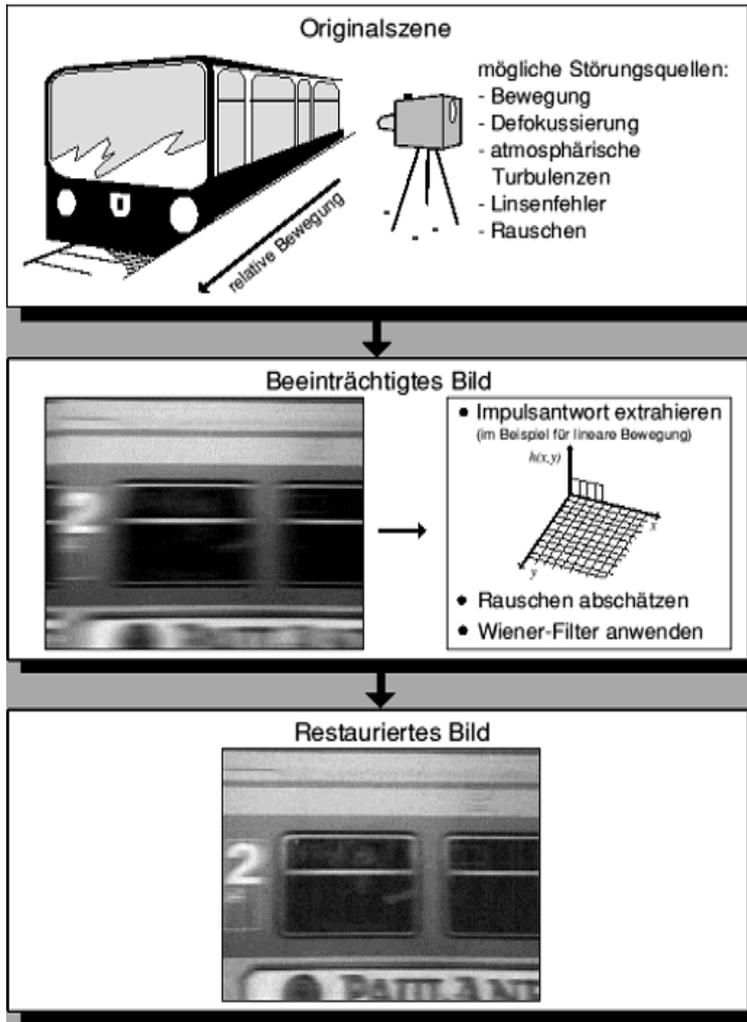


Quelle: eigene Darstellung nach Gonzales, Woods (2002), S. 221.

**Abb. 4.1:** Modell der Bildverschlechterung und -restauration

### *Wiener Filter*

Dieser Filter gehört zu den stochastischen Bildrestaurierungsverfahren und führt eine optimale Bildverbesserung im statistischen Sinne durch. Unter vielen Ansätzen hat sich der Wiener Filter als universelles Werkzeug für diesen Verfahrensweg herausgestellt.



Quelle: Lückenhaus, M. (1996).

**Abb. 4.2:** Grundlagen des Wiener-Filters und seine Anwendung in der Bildanalyse

Ziel des Wiener Filters ist, unter Ausnutzung des a-priori-Wissens über das Bildsignal und die Störung, ein neues Signal so zu berechnen, dass der quadratische Fehler  $\varepsilon^2$ , einer Schätzung des Originalbildes  $f$ , minimal wird.

$$\varepsilon^2 = E \{ ( f(i,j) - \hat{f}(i,j) )^2 \} \quad E\{\} \text{ ist der Erwartungswert.}$$

Es ist anzunehmen, dass die additive Störungsfunktion und das Bild nicht in Beziehung stehen, dass eins von beiden den Mittelwert null hat und dass die Graustufen in der Schätzung eine lineare Funktion der Stufen in dem verschlechterten Bild sind. Geht man von diesen Bedingungen aus, so erhält man folgende Gleichung:

$$\begin{aligned}
 \hat{F}(u,v) &= \left[ \frac{H^*(u,v)S_f(u,v)}{S_f(u,v) |H(u,v)|^2 + S_n(u,v)} \right] G(u,v) \\
 &= \left[ \frac{H^*(u,v)}{|H(u,v)|^2 + S_n(u,v)/ S_f(u,v)} \right] G(u,v)
 \end{aligned}$$

$$= \left[ \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)} \right] G(u,v)$$

Wir benutzen hier die Tatsache, dass das Produkt einer komplexen Größe mit ihrer konjugierten das Gleiche ist, wie die komplexe Größe zum Quadrat, so ist es möglich, die oben gezeigten Umformungsschritte durchzuführen, um eine Formel zu erhalten, die weiter vereinfacht werden kann, wie später gezeigt.

$H(u,v)$  = Verschlechterungsfunktion (Fouriertransformation der Bildstörung)

$H^*(u,v)$  = Komplexkonjugierte von  $H(u,v)$

$|H(u,v)|^2$  =  $H^*(u,v) H(u,v)$

$S_\eta(u,v) = |N(u,v)|^2$  = die spektrale Dichte des Rauschen

$S_f(u,v) = |F(u,v)|^2$  = die spektrale Dichte des gestörten Bildes

$G(u,v)$  = die Transformation des gestörten Bildes

Man kann  $S_\eta(u,v)$ , die spektrale Dichte des Rauschen, also die Stärke der Störung bei bestimmten additiven Störungsfunktionen als konstant betrachten. Dies ist ein Ansatz, der häufig benutzt wird, wenn diese Größe nicht bekannt ist oder nicht geschätzt werden kann, so ergibt sich die folgende Annäherung an unsere ursprüngliche Gleichung:

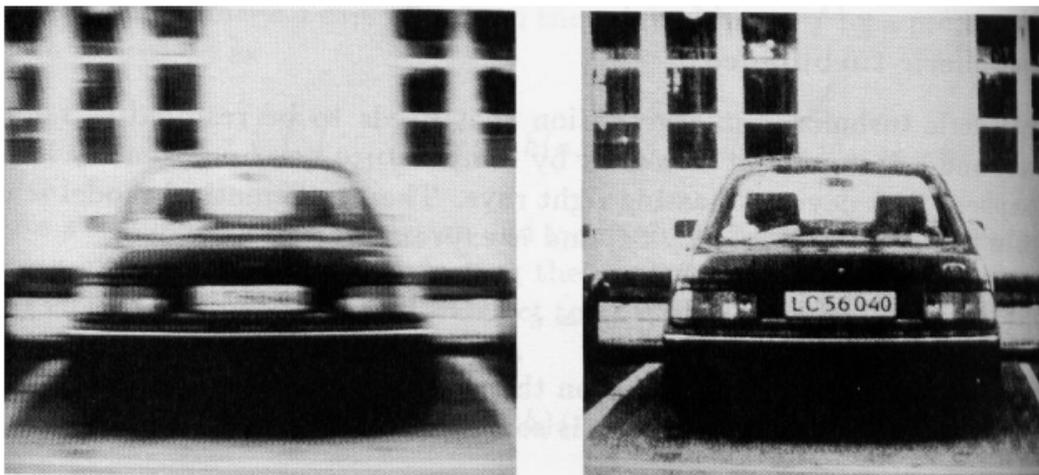
$$F'(u,v) = \left[ \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K} \right] G(u,v)$$

wobei  $K$  eine spezielle Konstante ist.

Das restaurierte Bild in der räumlichen Domäne ist durch die inverse Fouriertransformation von der Frequenzdomänen-Schätzung  $F'(u,v)$  gegeben.

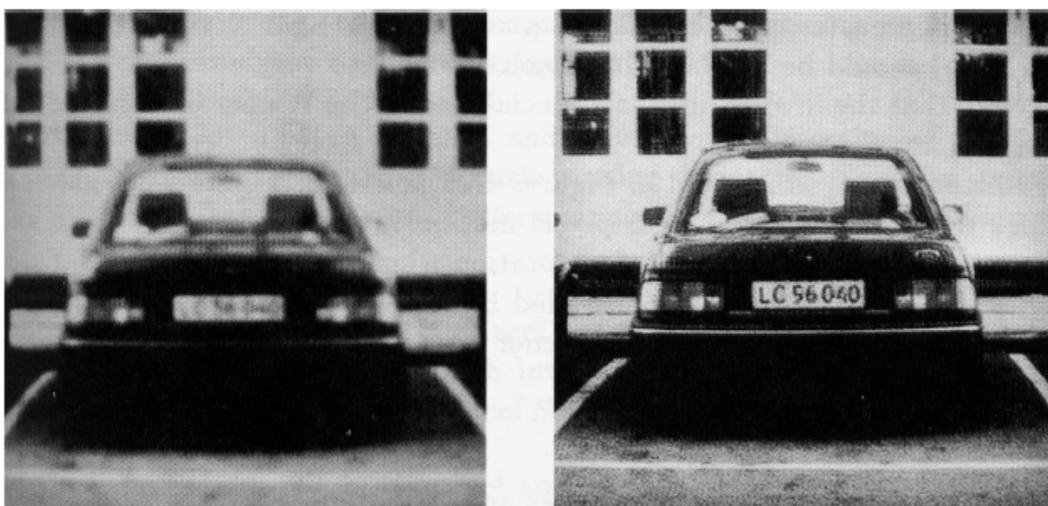
Um den Wiener Filter anwenden zu können müssen daher die Art der Störung  $H(u,v)$  und die statistischen Parameter des Rauschens bekannt sein.

Wird  $S_\eta(u,v)$  gleich null gesetzt, so wird aus dem Wiener Filter der inverse Filter, der hier nicht näher beschrieben wird.



Quelle: o. A.

**Abb. 4.3:** Restaurierung einer horizontalen Bewegungsunschärfe



Quelle: o. A.

**Abb. 4.4:** Restaurierung einer Defokussierung mit dem Wiener Filter

## 5 Luftbildauswertung

### 5.1 Luftbildauswertung am Beispiel der vorgestellten Methoden

Der letzte Themenbereich befasst sich im Speziellen mit der Aufbereitung von Luftbildern, die zusammen mit der Bildanalyse als ein Beispiel der Bildvorverarbeitung betrachtet werden sollen.

Viele der bisher vorgestellten Verfahren finden auch in der Bearbeitung und Auswertung von Luftbildern Verwendung. Ist es bei der Aufnahme der Bilder zu einer Über- oder Unterbelichtung gekommen, so lassen sich die zu geringen Grauwertunterschiede durch eine Kontrastverbesserung ausgleichen. Diese Verbesserung kann z. B. durch logarithmische oder exponentielle Transformationen, durch Dehnung des Histogramms oder durch einen Histogrammausgleich geschehen. Dabei handelt es sich um punktbezogene Operatoren, die nicht durch Nachbarschaftswerte beeinflusst werden.

Zur Reduzierung von gleichmäßigen Störungen, die z. B. durch Verwackelungen des Flugzeugs entstanden sind, eignen sich lokale Operatoren wie Glättungsfilter. Der Medianfilter ist dabei resistent gegenüber Ausreißern, dafür aber im Gegensatz zum Mittelwertfilter auch rechenintensiver.

Die Bearbeitung von Bildern im Frequenzraum lässt sich auch auf Luftbilder anwenden. Damit wird eine globale und effiziente Bearbeitung ermöglicht. Mögliche Filter, die im Frequenzraum angewendet werden können, sind z. B. der Hoch- und Tiefpassfilter und das „Unsharp Masking“.

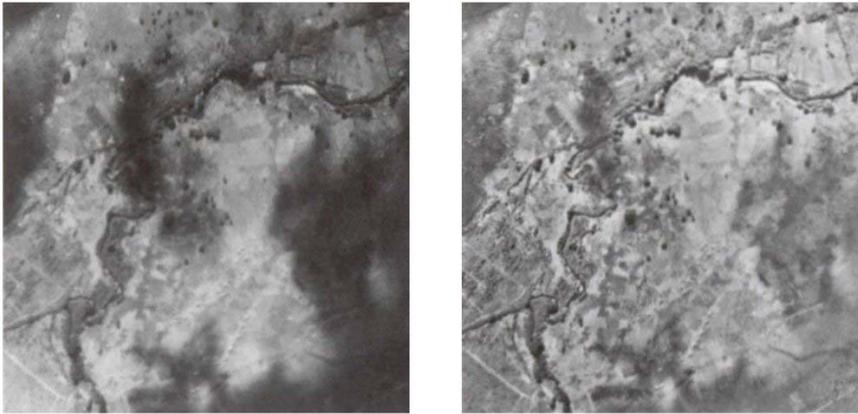
Ein Hochpassfilter dient dabei zur Verschärfung von Kanten und zur Hervorhebung von Details wie z. B. Minenobjekten, da durch ihn geringere Grauwertunterschiede betont werden. Allerdings werden durch diesen Filter auch unerwünschte Rauschanteile oder sonstige Störungen betont, was zu einer Qualitätsminderung im Ergebnisbild führen kann.<sup>13</sup>

Tiefpassfilter hingegen haben die entgegengesetzte Wirkung. Sie betonen große Grauwertunterschiede und führen somit zu einer Glättung des Bildes und einer Verwischung von Kanten. Dadurch gehen zwar Details verloren, für eine spätere Kantendetektion ist hingegen eine Glättung meist notwendig, um eine Reduktion auf die wesentlichen Strukturen eines Bildes – wie z. B. Kanten – vorzunehmen.

---

<sup>13</sup> Vgl. Albertz (2001), S. 108.

Ein weiteres für die Luftbildauswertung wichtiges Verfahren ist das „Unsharp Masking“. Dabei können unerwünschte großflächige Grauwertunterschiede, z. B. entstanden durch Wolkenschatten, reduziert werden, so dass ein in der Gesamthelligkeit ausgeglichenes Bild entsteht<sup>14</sup>.



Quelle: Vgl. Albertz (2001), S. 110.

**Abb. 5.1:** Reduzierung von Wolkenschatten durch „Unsharp Masking“

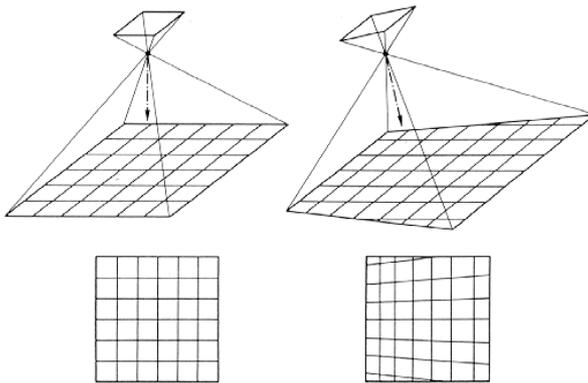
Es zeigt sich also, dass die Auswahl der Bearbeitungsmethoden zur Bildverbesserung stark von der vorliegenden Störung und der Art der Weiterverarbeitung durch die Objekterkennung abhängt. Es findet daher ein Abwägungsprozess zwischen den einzelnen Verfahren zur Bildverbesserung statt, der zu einer Auswahl an Verfahren führen sollte, die das Bild möglichst gut auf eine spätere Objekterkennung vorbereitet.

Neben diesen „generellen“ Verfahren für die Bildvorverarbeitung werden bei der Aufbereitung von Luftbildern noch weitere Verfahren benötigt, um Luftbilder z. B. zu entzerren und zu einem zusammenhängenden Bild zusammenzufügen. Eine Entzerrung ist notwendig, da es durch das Aufnahmesystem, das Geländere relief oder eine geneigte Aufnahmerichtung zu Verzerrungen im Bild kommen kann.<sup>15</sup>

---

<sup>14</sup> Vgl. Abb. 5.1.

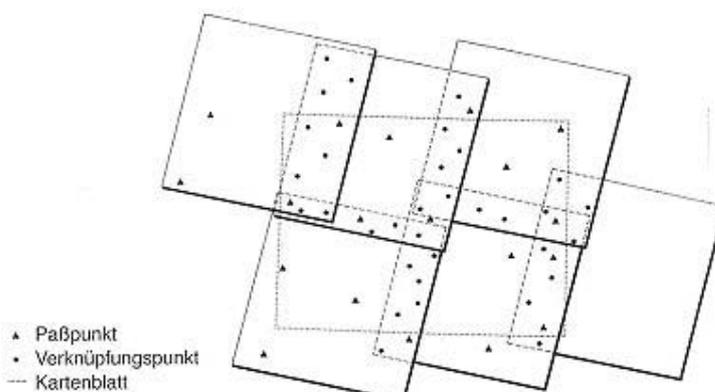
<sup>15</sup> Vgl. Abb. 5.2.



Quelle: Albertz (2001), S. 75.

**Abb. 5.2:** Verzerrung eines Luftbildes bei geneigter Aufnahmerichtung

Eine Entzerrung geschieht durch eine geometrische Transformation mit Hilfe von eindeutig bestimmten Passpunkten, aus denen dann die Transformationsfunktion berechnet wird. Diese Passpunkte dienen gleichzeitig auch dazu, das Luftbild zu rektifizieren, d. h. in ein bestimmtes geodätisches Koordinatensystem einzupassen. Dadurch ist es dann möglich, Punkten im Bild einen genauen Ort in der Wirklichkeit zuzuordnen, so dass mögliche Minenobjekte auch am Boden lokalisiert werden können. Um schließlich ein zusammenhängendes Bildmosaik eines Gebiets aus einzelnen Luftbildern erstellen zu können, werden Verknüpfungspunkte benötigt, die in den Überlappungsbereichen benachbarter Bilder liegen.<sup>16</sup> Die praktische Durchführung der Mosaikbildung ist dann im Prinzip wieder eine geometrische Transformation der Bilddaten.<sup>17</sup>



Quelle: Albertz (2001), S. 119.

**Abb. 5.3:** Schematisches Beispiel zur geometrischen Mosaikbildung.

<sup>16</sup> Vgl. Abb. 5.3.

<sup>17</sup> Albertz (2001), S. 118.

## 5.2 Landminenerkennung auf Luftbildern

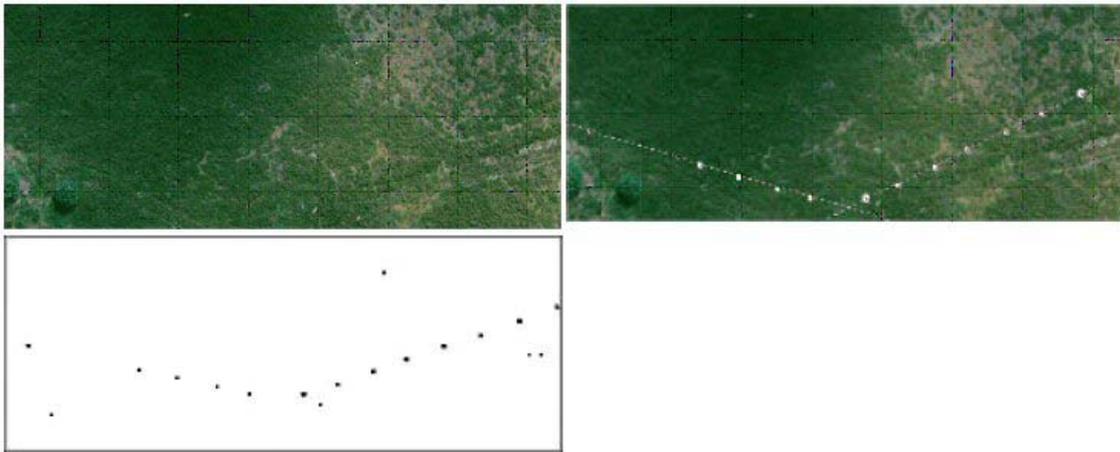
Eine weitere Möglichkeit die Auswertung von analogen oder digitalen Bildern zur Landminendetektion zu nutzen bietet die Fernerkundung, wenn auch nur in eingeschränktem Maße. Dabei geht es nicht voranging um die Erkennung einzelner Minen, sondern vielmehr um die Eingrenzung größerer Minenfelder. Dieses kann als ein Schritt auf dem Weg zur kompletten Minenräumung aufgefasst werden. In Gebieten, in denen noch keine Mittel zu Minenräumung zur Verfügung stehen, kann es daher schon hilfreich sein zu wissen, welche Gebiete potentiell gefährdet sind, um diese Gebiete speziell kennzeichnen zu können.

Mit der Erkennung von Landminen auf Luftbildern beschäftigt sich u. a. das europäische „Airborne Minefield Detection Pilot Project“. Dabei werden auf Testfeldern in Belgien und Mozambique auch optische Sensoren getestet, die Bilder von 1-32 cm Auflösung aus einer Höhe von ca. 200 m liefern können. Das generelle Modell dieser Erkennung beruht darauf, dass Minen durch eine Strukturveränderung der Vegetation oder des Bodens erkannt werden können.<sup>18</sup> Dazu wurde ein Vorgehensschema entwickelt, das durch vier Verarbeitungsschritte gekennzeichnet ist.

Zuerst werden Klassen für Vegetation, Bodenart usw. durch Merkmalsextraktion gebildet u. a. mit Hilfe von statistischen Kennwerten und der Fourier-Transformation. Dann werden die Merkmale ausgewählt, die sich am meisten unterscheiden, um eine Reduzierung der Merkmalsklassen zu erreichen. Der dritte Schritt ist die Aufstellung eines Modells zu jeder der vorher ausgewählten Merkmalsklassen. Jede Klasse kann dann durch einen Merkmalsvektor, eine Kovarianzmatrix und eine a priori Wahrscheinlichkeit beschrieben werden. Der Merkmalsvektor eines Objekts wird schließlich im vierten Schritt mit denen der verschiedenen Klassen verglichen, um Übereinstimmungen zu finden und somit Objekte klassifizieren zu können. Das Ergebnis einer solchen Segmentierung ist ein Binärbild, das mögliche Minenobjekte darstellt. Um falsch klassifizierte Pixel zu entfernen zu können ist allerdings zusätzliches Wissen über die Minenart und -größe nötig. Da mit Hilfe des vorgestellten Verfahrens einzelne Minen nur schwer auszumachen sind, wird versucht nach Minen in regelmäßigen Strukturen zu suchen, um somit Minenfelder besser abgrenzen zu können. Können solche regelmäßigen Strukturen wie z. B. eine Linie gefunden werden, so steigt dann auch gleichzeitig die Wahrscheinlichkeit, dass sich weitere Minen auf dieser Linie befinden, wodurch letztendlich die Suche nach Minen vereinfacht wird.

---

<sup>18</sup> Vgl. Pizurica et. al (1999), S. 2.



Quelle: Vgl. Pizurica et. al (1999), S. 2.

**Abb. 5.4:** Schritte der Minenfeldererkennung auf Luftbildern: 1. Minenfeldregion, 2. Ergebnis der Minensuche, 3. Abgrenzung des Minenfeldes

### 5.3 Probleme der Luftbildauswertung

Die generellen Probleme bei der Luftbildauswertung sind unterschiedlich. Durch eine zu geringe Auflösung und zu kleine Objektgröße wird die Chance geringer, die gesuchten Minen zu entdecken. Je kleiner die Minen und je älter das Minenfeld sind, desto kleiner die Wahrscheinlichkeit auf Erfolg. Das Alter ist insofern entscheidend, da mit zunehmendem Alter des Feldes die Vegetation zunimmt und somit die Minen überwachsen werden. Ein weiteres Problem besteht darin, dass es keine Möglichkeit gibt mit optischen Sensoren in der Form von Luftbildern Minen unter der Erdoberfläche zu orten. Die Erstellung von Luftbildern ist außerdem von den Wetterverhältnissen abhängig, so dass geeignete Wetterbedingungen wie z. B. eine klare Sicht herrschen müssen. Aufgrund diese Probleme war es bisher fast nur möglich Anti-Panzer-Minen zu erkennen, da Anti-Personen-Minen durch ihre geringere Größe keinen ausreichenden Kontrast auf den Bildern darstellten<sup>19</sup>.

Verbessert werden könnte die vorgestellte Methode durch weiteres *a priori* Wissen über das Gebiet und die Minen. Es können weiterer Layer hinzugenommen werden, die Informationen über Bodennutzung, z. B. durch Straßen, enthalten. Diese stelle eher sichere Bereiche dar.

Dennoch ist die Luftbildauswertung eine Technik mit dessen Hilfe relativ schnell große Gebiete auf Minenfelder überprüft werden können, sofern die äußeren Bedingungen dazu günstig sind. Auch wenn diese Möglichkeit der Minenerkennung bisher hauptsächlich in

<sup>19</sup> Vgl. Pizurica, Katartzis, Cornelis, Sahli (1999), S. 6.

Pilotprojekten eingesetzt und getestet wurde, könnte die Fernerkundung – sofern die Auflösungen der Bilder besser und die Suchalgorithmen genauer werden – zur Unterstützung der individuellen Minensuche am Boden beitragen.

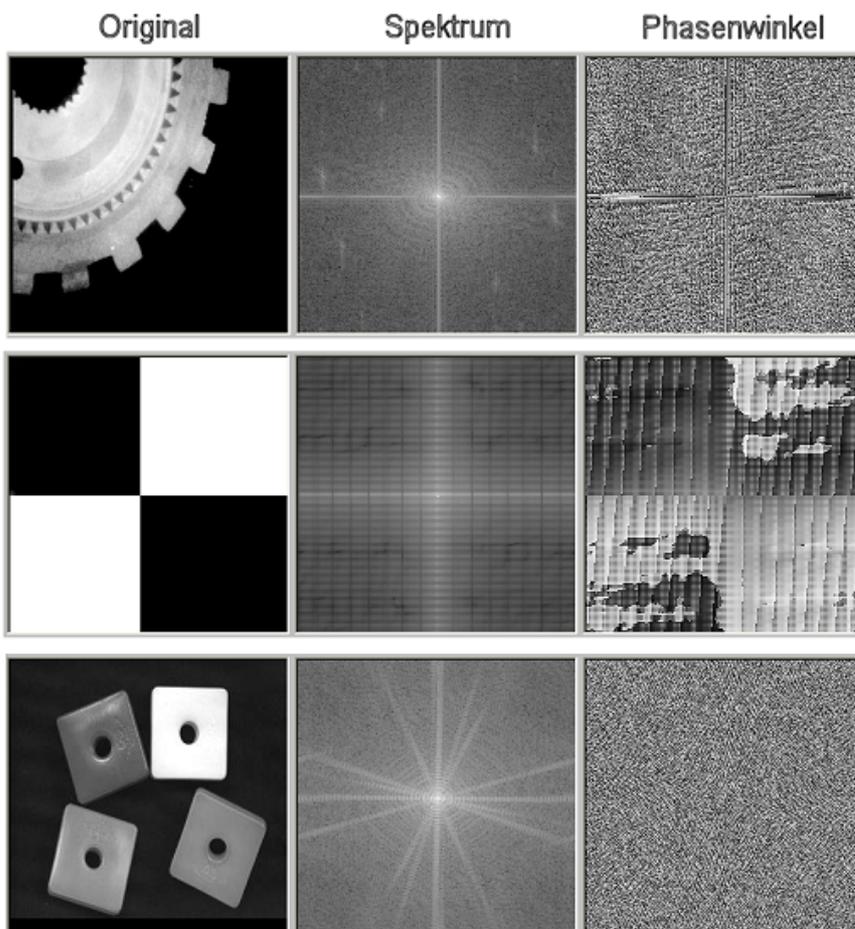
## Literaturverzeichnis

- Albertz, J.: Einführung in die Fernerkundung. Grundlagen der Interpretation von Luft- und Satellitenbildern. Wissenschaftliche Buchgesellschaft. 2. Aufl., Darmstadt 2001. S. 75-119.
- Gonzales R.C.; Woods, R.E.: Digital Image Processing. 2. Aufl., New Jersey 2002. S. 42-273.
- Gonzales, R.C.; Woods, E: Digital Image Processing. 2. Aufl., 2002.  
[http://www.cvip.uofl.edu/wwwcvip/education/ECE618\\_2002/download/Chapter02.pdf](http://www.cvip.uofl.edu/wwwcvip/education/ECE618_2002/download/Chapter02.pdf),  
[http://www.cvip.uofl.edu/wwwcvip/education/ECE618\\_2002/download/Chapter03.pdf](http://www.cvip.uofl.edu/wwwcvip/education/ECE618_2002/download/Chapter03.pdf). Abrufdatum 2004-01-02.
- Graber, G.: BACG Online. Graz. [http://fcggpc41.icg.tu-graz.ac.at/grabner/BACG\\_Online/Kapitel-07-Filterung/sld076.htm](http://fcggpc41.icg.tu-graz.ac.at/grabner/BACG_Online/Kapitel-07-Filterung/sld076.htm). Abrufdatum 2004-01-04.
- Jiang, X. : Vorlesung Bildverarbeitung. Universität Münster 2003. S. 22-59.  
<http://www.math.uni-muenster.de/~u/xjiang/lectures/SS03/SS03-BV.html>. Abrufdatum 2004-01-03.
- Jiang, X. : Vorlesung Bildanalyse. Universität Münster 2002. <http://www.math.uni-muenster.de/~u/xjiang/lectures/WS02/WS02-BA.html>. Abrufdatum 2004-01-03.
- Lückenhaus, M.: Grundlagen des Wiener-Filters und seine Anwendung in der Bildanalyse. Diplomarbeit. München 1996.
- O. V.: Das elektromagnetische Spektrum. [http://www.ro-ro-seiten.de/physik/lk12/emwellen/elektromagnetisches\\_spektrum.html](http://www.ro-ro-seiten.de/physik/lk12/emwellen/elektromagnetisches_spektrum.html). Abrufdatum 2004-01-02.
- Pandit, M.: Methoden der Digitalen Bildverarbeitung für die Anwendung. Fortschritt-Berichte VDI Reihe 10 Band 669. Düsseldorf 2001. S. 50-164.
- Pizurica, V.; Katartzis, A; Cornelis, J; Sahli, H: What can be expected from Computerised Image Analysis for Airborne Minefield Detection? Brussel 1999. S. 2-6.  
<http://citeseer.nj.nec.com/pizurica99what.html>. Abrufdatum 2004-01-02.
- Preim, B.: Vorlesung Computergrafik. Farbe (Wahrnehmung, Farbräume). Magdeburg 2003. S. 14-16. <http://isgwww.cs.uni-magdeburg.de/~bernhard/cg/vorl4Farbe.pdf>. Abrufdatum 2004-01-03.
- Rehrmann, V.: Foliensammlung Vorlesung digitale Bildverarbeitung. 1999.  
<http://www.uni-koblenz.de/~lb/publications/1999/bvscript.pdf>. Abrufdatum 2004-01-03.
- Weitkamp, A.: Untersuchungen zum Einsatz von Laserscanningverfahren beim Monitoring der Deichvorlandentwicklung. Diplomarbeit. Hannover 2001.

## Anhang

### A Einige Beispiele im Zusammenhang mit Filterung im Frequenzraum

#### A.a Das Spektrum und der Phasenwinkel einer DFT



Quelle: Bildausgaben eines Java-Programms

## **Abschließende Erklärung**

Wir versichern hiermit, dass wir unsere Seminararbeit „**Bildaufbereitung für die Objekterkennung**“ selbstständig und ohne fremde Hilfe angefertigt haben, und dass wir sämtliche von anderen Autoren wörtlich übernommenen Stellen wie auch die sich an die Gedankengänge anderer Autoren eng anlegenden Ausführungen unserer Arbeit besonders gekennzeichnet und die Quellen zitiert haben.

Münster, den 5. Januar 2004.

Jana Gliet, Mareike Kritzler, Birgit Pieke, Timo Schreiber.