

Klassifikationsverfahren zur Materialerkennung

Grenzschichterkennung mittels laserinduzierter Fluoreszenz in mineralischen Lagerstätten am Beispiel der Braunkohlegewinnung

Von der Fakultät für Georessourcen und Materialtechnik der
Rheinisch - Westfälischen Technischen Hochschule Aachen

zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

genehmigte Dissertation

vorgelegt von Diplom-Ingenieur

Arne Kristoffer Bayer

aus Lüneburg

Berichter: Univ.-Prof. Dr.-Ing. Karl Nienhaus

Univ.-Prof. Dr.-Ing. Andreas Seeliger

Tag der mündlichen Prüfung: 17. Dezember 2004

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

Arne Kristoffer Bayer - Klassifikation zur Materialerkennung
Grenzschichterkennung mittels laserinduzierter Fluoreszenz in mineralischen Lagerstätten am
Beispiel der Braunkohlengewinnung

ISBN: 3-86130-757-X

1. Auflage 2005

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Das Werk einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwendung ist ohne die Zustimmung des Herausgebers außerhalb der engen Grenzen des Urhebergesetzes unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Vertrieb:

1. Auflage 2005
© Verlagshaus Mainz GmbH Aachen
Süsterfeldstr. 83, 52072 Aachen
Tel. 0241/87 34 34
Fax 0241/87 55 77
www.Verlag-Mainz.de

Herstellung:

Druck und Verlagshaus Mainz GmbH Aachen
Süsterfeldstraße 83
52072 Aachen
Tel. 0241/87 34 34
Fax 0241/87 55 77
www.DruckereiMainz.de
www.Druckservice-Aachen.de

Satz: nach Druckvorlage des Autors
Umschlaggestaltung: Druckerei Mainz

printed in Germany

Danksagung:

Ohne die Neugierde, den Wissensdurst und den Drang, unbekannten Dingen auf den Grund zu gehen und stets das Bekannte zu hinterfragen, hätte diese Arbeit nie das Licht der Welt erblickt. Den Grundstein der vorliegenden Dissertation haben damit bereits sehr früh meine Eltern gelegt: Danke dafür und die immerwährende Unterstützung. Neben meiner Familie danke ich besonders Ina für das Verständnis und den Beistand auch in Momenten, in denen es mal nicht reibungslos lief. Nur mit dieser Motivationsquelle war diese Arbeit möglich.

Mein Dank gilt auch allen Mitstreitern und Freunden, die mich über all die Jahre begleitet haben: in Kuhstedt, in Aachen, an der RWTH und besonders am BGMR.

Allen voran Univ.-Prof. Dr.-Ing. Karl Nienhaus, der mich durch seine unkonventionelle und sehr pragmatische Herangehensweise an Herausforderungen aller Art stets motiviert hat. Diese Zeit war prägend. Aber auch die Zusammenarbeit mit allen Kollegen, Diplomanden, Studien- und Seminararbeitern sowie studentischen Hilfskräften war stets ein Grund, mit Freude und Spaß die anstehenden Probleme anzugehen.

Ohne die Unterstützung des Projektes „Laserinduzierte Fluoreszenz in der mineralischen Rohstoffgewinnung“ durch die Deutsche Forschungsgemeinschaft (DFG) und die RWE Power AG wäre diese Arbeit nicht entstanden. Besonderer Dank gebührt hier der Mannschaft des Tagebau Hambachs und vor allem Markus Poths. Auch die Diskussionen mit Daniel Keyzers vom Lehrstuhl für Informatik VI an der RWTH Aachen waren eine wichtige Hilfe.

Darüber hinaus danke ich Herrn Univ.-Prof. Dr.-Ing. Andreas Seeliger für die Übernahme des Korreferats.

Inhaltsverzeichnis

Inhaltsverzeichnis	I
1 Motivation	1
2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz.....	4
2.1 Entstehung der Photolumineszenz	4
2.1.1 Stimulierte Absorption.....	5
2.1.2 Stimulierte Emission	5
2.2 Eigenschaften der Fluoreszenz	6
2.2.1 Absorption.....	7
2.2.2 Emission und Fluoreszenzintensität.....	7
2.2.3 Quantenausbeute	8
2.2.4 Abklingverhalten	8
2.3 Materialerkennung mittels LIF	9
2.3.1 Grundlagen der Materialerkennung mittels LIF.....	10
2.3.2 Auswertung und Interpretation von LIF-Signalen.....	11
2.4 Genereller Aufbau eines LIF Systems	12
2.4.1 Lichtquelle.....	13
2.4.2 Strahlengang.....	13
2.4.3 Empfangseinheit	14
2.4.4 Weitere Bestandteile.....	15
2.5 LIF-Geräte für den Einsatz in der Rohstoffindustrie	16
2.5.1 Stationäre Anlagen	16
2.5.2 LIF Scanner	18
3 Praktische Untersuchungen in der Braunkohlegewinnung.....	21
3.1 Braunkohlentagebau Hambach	23
3.1.1 Eingesetzte Gewinnungstechnik.....	24
3.1.2 Lagerstättensituation und Qualitätsmanagement.....	24
3.2 Messkampagnen im Tagebau Hambach	27
3.2.1 Messkampagne Liegendes Frimmersdorf a.....	28
3.2.2 Messkampagne Zwischenmittel Frimmersdorf a/b.....	32
3.2.3 Diskussion der Fehlerquellen.....	34
3.2.4 Fazit der Feldversuche	36
4 Mustererkennung: Grundlagen, Algorithmen und Programmierung.....	38
4.1 Grundlagen der Mustererkennung.....	38
4.2 Klassifikationsverfahren	41
4.2.1 Gaußverfahren.....	42
4.2.2 k-nächster Nachbar Verfahren (k-NN)	44
4.2.3 Lineares Regressionsverfahren	47
4.3 Verfahren zur Merkmalsselektion	50
4.3.1 Sequential Forward Selection (SFS).....	52
4.3.2 Sequential Backward Selection (SBS)	53
4.3.3 Sequential Forward Floating Selection (SFFS).....	54
4.3.4 Wrapper2-Ansatz	55
5 Parametereinstellungen zur Merkmalsselektion	58
5.1 Ermittlung der optimalen Durchlaufanzahl	60
5.2 Ermittlung der optimalen Merkmalsanzahl.....	64
5.3 Ermittlung der optimalen Methode zur Merkmalsselektion	67
6 Durchführung der Merkmalsselektion und Materialklassifikation	69
6.1 Originäre Merkmalsselektion und Materialklassifikation.....	71
6.2 Analyse der Überkreuzkorrelation	73
6.3 Optimierung des reduzierten Merkmalssets	74
6.4 Optimierung der Materialklassifikation mittels eines selektierenden gewichteten Mittelwerts (SGM).....	79

7	Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation	83
7.1	Einfluss der Anzahl an Materialklassen	84
7.1.1	Liegendes, 9 Materialien	84
7.1.2	Zwischenmittel, 6 Materialien, 355 nm	87
7.1.3	Fazit.....	89
7.2	Einfluss der Merkmalsselektion und des Klassifikationsverfahrens	90
7.2.1	Abschnitt nördlich Liegendrücken	90
7.2.2	Abschnitt Liegendes Tiefstufe	93
7.2.3	Abschnitt Zwischenmittel Flöz Frimmersdorf a/b.....	96
7.3	Einfluss der Anregungswellenlänge des Lasers	102
7.4	Verwendung von übergreifenden und universellen Merkmalen für das Liegende.....	102
8	Diskussion der Ergebnisse.....	106
9	Ausblick	111
10	Zusammenfassung.....	117
	Literaturverzeichnis	121
	Abbildungsverzeichnis	124
	Tabellenverzeichnis	129
	Anhang A: LIF Scanner / Analysers, Typ Sc 2-1-2	131
	A 1: Technische Daten und Anschlüsse	131
	A 2: Technische Zeichnungen	132
	Anhang B: Tagebau Hambach: Daten, Profile und Messeinstellungen der Messkampagne	134
	Anhang C: Algorithmen zur Merkmalsselektion und Materialklassifikation (Flowcharts und Programmcode).....	140
	Anhang D: Ergebnisse der Merkmalsselektion und Mustererkennung	167

1 Motivation

Der Qualitätsgedanke rückt in der Rohstoffindustrie immer mehr in den Vordergrund des unternehmerischen Handelns, da sich die Produkte nicht mehr ausschließlich über Preis und Menge, sondern in zunehmendem Maße über die Garantie einer konstanten stofflichen Zusammensetzung absetzen bzw. weiterverarbeiten und nutzen lassen. In jüngster Vergangenheit haben in diesem Zusammenhang Online-Messmethoden zur Bemusterung des gesamten Förderstromes bereits auf dem Fördermittel oder direkt am Stoß an Bedeutung gewonnen, um das Material entsprechend der Qualität effizienter lenken zu können. Die Online-Analyse bildet dabei das Bindeglied zwischen der Rohstoffgewinnung und den entsprechenden Lagerstättenmodellen sowie der Weiterverarbeitung, in der die Analyse von Teilströmen seit langem Stand der Technik ist. Dabei wird zudem die zeitbedingte Informationslücke zwischen Gewinnung und Auf- bzw. Weiterverarbeitung geschlossen, die konventionelle, nasschemische Untersuchungen aufgrund der längeren Analysedauer mit sich bringen würden.

Neben dem Abgleich zwischen der erwarteten und der tatsächlichen stofflichen Zusammensetzung des im bergmännischen Gewinnungsprozess hereingewonnenen Materials ermöglicht der Einsatz von intelligenten Analysemethoden darüber hinaus eine effektivere Lagerstättenausnutzung und eine Minimierung des mitgeschnittenen Nebengesteins. Die beiden letzteren Aspekte resultieren aus der Möglichkeit, unerwünschte Materialien entlang von Grenzschichten mittels der Daten der Online-Analyse aushalten zu können und so unerwünschte Vermischungen bereits beim Gewinnungsvorgang zu verhindern.

Der Einsatz der optischen Analyseverfahren der laserinduzierten Fluoreszenz (LIF) bietet eine Möglichkeit, unterschiedliche Materialien zu erkennen und Trennflächen auch über Entfernungen zwischen Analysegerät und Materialoberfläche zu detektieren, die mit anderen Verfahren wie etwa Röntgenstrahlung oder radioaktiven Methoden nicht zu realisieren wären. Am Lehr- und Forschungsgebiet „Betriebsmittel für die Gewinnung mineralischer Rohstoffe (BGMR)“ der RWTH Aachen wird seit einigen Jahren untersucht, wie sich diese Methode für die Materialerkennung in der Rohstoffgewinnung nutzen lässt, um den Gewinnungsvorgang hinsichtlich der Qualitätssteuerung zu optimieren. Dabei kommt einer systematischen Auswertung der Fluoreszenzsignale für eine zuverlässige Materialklassifikation besondere Bedeutung zu. Die Grundlagen für die Grenzschichterkennung bilden stabile Klassifikationsalgorithmen, die eine unkomplizierte Zuordnung der Messwerte in für die Qualitätssteuerung relevante Materialklassen ermöglichen.

Im Rahmen eines Forschungsvorhabens mit der RWE Power AG wurden im Zeitraum zwischen Juni 2002 und April 2003 die Einsatzmöglichkeiten des innovativen optischen Verfahrens der laserinduzierten Fluoreszenz (LIF) untersucht. Es wurde versucht, mit Hilfe charakteristischer Fluores-

1 Motivation

zenzsignale qualitative Unterschiede zwischen signifikanten Lagerstättenpartien im Braunkohletagebau Hambach zu erkennen, die nicht unmittelbar auf einer chemischen Analyse der stofflichen Zusammensetzung des Materials, sondern eher auf einem pragmatischen, phänomenologischen Ansatz beruhen.

Es sollte geprüft werden, ob auf der Basis dieser Technik eine Steuerung entwickelt werden kann, die den Fahrer des Schaufelradbaggers in die Lage versetzt, während des Gewinnungsvorganges in Echtzeit die Stratigraphie qualitativ und unabhängig von für das menschliche Auge sichtbaren Qualitätsmerkmalen zu beurteilen. Schichten, die farblich gleich erscheinen, wie etwa Braunkohle und dunkle Tone, könnten so unterschieden oder Zwischenmittel innerhalb des Flözes besser ausgehalten werden. Durch eine auf der LIF-Technik basierenden Warnfunktion könnten ungeplante und unkontrollierte Verunreinigungen während des Gewinnungsprozesses verhindert werden. Dies ginge mit einer Optimierung des Lagerstättenausbringens durch ein gezieltes Heranfahren an die jeweilige Trennfläche einher.

Für die Bestimmung von materialspezifischen Fluoreszenzcharakteristika, die auch als „Fingerabdrücke“ bezeichnet werden, lässt sich bei der LIF auf eine Vielzahl von Merkmalen zurückgreifen, die aus dem Emissionsspektrum gewonnen werden. Diese Datenfülle erschwert jedoch in erheblichem Maße eine Automatisierung der Materialerkennung wie auch die Dimensionierung und Auslegung von praxistauglichen Analysegeräten.

In dieser Arbeit sollen daher Wege aufgezeigt werden, wie aus LIF-Messdaten systematisch signifikante Merkmale extrahiert werden können, die eine optimale Zuordnung und Klassifikation von unterschiedlichen Materialien erlauben. Dazu werden drei Klassifikations- und Selektionsalgorithmen hinsichtlich ihrer Eignung zur Materialerkennung verglichen und für die beiden untersuchten Lagerstättenabschnitte zur Mustererkennung eingesetzt. Dies stellt einen generell neuen Ansatz bei der Materialklassifikation in der mineralischen Rohstoffgewinnung mittels LIF dar, da bis dato lediglich Mittelwert basierte Algorithmen zum Einsatz kamen. Auch die Selektion der für eine Materialerkennung geeigneten Wellenlängenbereiche des Emissionsspektrums fand bisher nicht im Rahmen eines systematischen Vorgehens statt. Neben der grundsätzlichen Methodik werden die Schwierigkeiten bei der Nutzung der LIF-Technologie und die damit verbundenen besonderen Anforderungen an die Merkmalsselektion und anschließende Materialklassifikation ausführlich diskutiert.

Zunächst wird im Kapitel 2 auf die physikalischen und technischen Aspekte der laserinduzierten Fluoreszenz eingegangen, bevor im nachfolgenden Abschnitt die Besonderheiten der Qualitätssteuerung bei der Braunkohlegewinnung aufgeführt und die beiden Messkampagnen im Tagebau Hambach vorgestellt werden. Auch die Probleme und Fehlerquellen, die mit dem Einsatz des LIF Scanner / Analysers, einem Prototyp, verbunden sind, werden in diesem Abschnitt angesprochen.

Die generelle Vorgehensweise bei der Durchführung einer Mustererkennung wird im Abschnitt 4 dokumentiert. Die drei betrachteten Klassifikationsverfahren und die unterschiedlichen Ansätze zur Merkmalsselektion werden ausführlich erläutert, wobei auch auf die programmiertechnische Umsetzung der Algorithmen eingegangen wird. Bei den verwendeten Verfahren handelt es sich um Ansätze, die in anderen Bereichen der Muster- und Bilderkennung eingesetzt werden und sich somit in der Praxis bewährt haben. Für die speziellen Anforderungen bei der Zuordnung von LIF-Messdaten und dem Einsatz in der mineralischen Rohstoffgewinnung mussten die Algorithmen modifiziert und teilweise erheblich überarbeitet werden, um die gestellte Aufgabe erfolgreich zu bewältigen.

Nach der Ermittlung der Parametereinstellungen für die Mustererkennung erfolgt die eigentliche Durchführung der Merkmalsselektion und Materialklassifikation in Kapitel 6. Hier wird exemplarisch das Vorgehen am Beispiel eines Lagerstättenabschnitts veranschaulicht. Die Ergebnisse der Materialerkennung für die betrachteten Materialübergänge im Tagebau Hambach werden im folgenden Kapitel vorgestellt und hinsichtlich ihrer praktischen Verwendung zur Trennflächenerkennung im Gewinnungsprozess bewertet. Dabei wird neben dem Einfluss der Anzahl der zu unterscheidenden Materialklassen auf den Klassifikationserfolg die Eignung der verwendeten Algorithmen besprochen, sowie Ansätze erörtert, mit welchen LIF-Merkmalen eine universelle Nutzung der optischen Materialerkennung in der Braunkohle möglich wäre. Die Einsatzmöglichkeiten - aber auch das Verbesserungspotential - dieser neuen Technologie werden in Kapitel 9 diskutiert.

2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Die Klassifizierung und Erkennung mineralischer Rohstoffe, wie sie in dieser Arbeit durchgeführt wird, basiert auf dem physikalischen Phänomen der Photolumineszenz. Darunter wird das „kalte Leuchten“ von Atomen, Molekülen oder Kristallen verstanden, d.h. die Aussendung von Photonen nach einer Anregung, die nicht thermischer Natur ist. In Abhängigkeit der Dauer der Emission nach Beendigung der Anregung unterscheidet man zwischen Phosphoreszenz, die Sekunden bis Minuten nachleuchtet, und der Fluoreszenz, die in der Regel eine Lebensdauer im Bereich von 10^{-8} s aufweist.

Bei der hier genutzten laserinduzierten Fluoreszenz (LIF) erfolgt die Anregung des zu analysierenden Materials mittels monochromatischen Lichts. Die Wechselwirkungen zwischen elektromagnetischer Strahlung und der Materie sowie die Möglichkeiten der Nutzung dieser zur Unterscheidung unterschiedlicher Stoffe sollen in diesem Kapitel erläutert werden.

Das Prinzip der LIF hat seit langem einen festen Platz in der Biochemie und der Medizin sowie in Bereichen der Labor- und Umweltanalytik. Auch bei der Optimierung von Verbrennungsprozessen in Kraftmaschinen bedient man sich mittlerweile standardmäßig der laserinduzierten Fluoreszenz als Analysewerkzeug [HOF, 2002]¹.

Phosphoreszierende und fluoreszierende Eigenschaften von Mineralen werden seit langem in der Mineralogie intensiv laboranalytisch genutzt. Großtechnische Feldanwendungen in der mineralischen Rohstoffgewinnung sind jedoch erst seit Mitte der 1970er Jahre bekannt, diese kamen über Testeinsätze nicht hinaus. Die derzeitigen Anwendungen zur Materialerkennung in diesem Bereich, die gerätetechnischen Anforderungen sowie die mineralogischen Grundlagen bei der Nutzung der LIF werden in diesem Abschnitt vorgestellt.

2.1 Entstehung der Photolumineszenz

Die quantenmechanischen Wechselwirkungen zwischen Molekülen bzw. Atomen und elektromagnetischer Strahlung, zu denen auch die Photolumineszenz zählt, lassen sich mit Hilfe des Zwei-Niveau-Modells darstellen. Dieses basiert auf der vereinfachten Annahme, dass das Atom bzw. Molekül nur über zwei diskrete quantenmechanische Energiezustände definiert

¹ Eckige Klammern enthalten einen Verweis auf das Literaturverzeichnis am Ende der Arbeit. Die Großbuchstaben stehen für das Kürzel des Autors, die Jahreszahl kennzeichnet das Jahr der Veröffentlichung.

ist. Dabei handelt es sich um den so genannten Grundzustand und den angeregten Zustand.

Durch den Beschuss von Materie mit Laserphotonen definierter Intensität und Wellenlänge können, wie im Fall der laserinduzierten Fluoreszenz, unterschiedliche Reaktionen im quantenmechanischen System ausgelöst werden. Dies sind im Wesentlichen die Prozesse der stimulierten Absorption bzw. Emission, die für den Einsatz der LIF in der Materialerkennung praktische Bedeutung haben und die Basis für eine Klassifikation bilden.

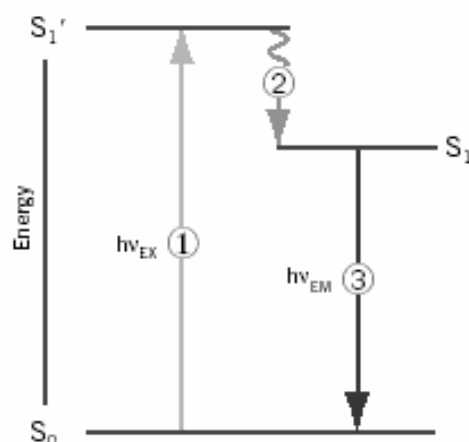
Darüber hinaus sind weitere Reaktionsmechanismen eines angeregten Systems bekannt, die ohne zusätzliche Energiezufuhr von außen ablaufen. Im Einzelnen sind dies die spontane Emission, die auch als Quenching bezeichnete Stoßauslöschung sowie die Prädissoziation und die Photoionisation [FUC, 2003].

2.1.1 Stimulierte Absorption

Die stimulierte Absorption beschreibt den Übergang vom Grundzustand in den angeregten Zustand durch den Beschuss mit einem Photon. Unterschiedliche Materialien können dabei unter Umständen gänzlich unterschiedliche Absorptionsspektren aufweisen, d.h. aus dem Spektrum der anregenden Strahlung werden nicht alle Wellenlängen in gleichem Maße absorbiert. Die daraus abgeleitete Absorptionsfähigkeit diskreter Wellenlängen ist somit eine materialspezifische Eigenschaft.

2.1.2 Stimulierte Emission

Unter der stimulierten Emission wird der energetische Wechsel des Systems von dem angeregten Zustand zurück in den ursprünglichen Grundzustand unter einer gleichzeitigen Abgabe von Photonen verstanden. Die quantenmechanischen Reaktionsprozesse sind in Abbildung 2-1 im erweiterten Zwei-Niveau-Modell dargestellt.



Durch die Energiezufuhr in Form eines Photons einer definierten Wellenlänge wird das Atom aus dem stabilen Zustand S_0 auf das Niveau des in-

2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Abbildung 2-1: Darstellung des quantenmechanischen Reaktionen bei der Fluoreszenz: Anregung (1), vibrationale Entspannung (2) und Emission (3) [APB, 2000]

2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

stabilen, angeregten Zustands S_1' angehoben. Dieser Zustand weist nur eine extrem kurze Halbwertszeit auf, in der Regel im Bereich von Nanosekunden. Das angeregte Molekül „entspannt“ sich in dieser Zeit in Richtung des niedrigsten Energieniveaus des angeregten Zustandes (S_1) unter Abgabe von Wärmeenergie. Beim Rückfall in den stabilen Grundzustand S_0 wird die Energie eines Photons Wellenlänge λ_{EM} der anregenden elektromagnetischen Strahlung.

Die Wellenlänge des emittierten Lichts λ_{EM} stellt die Energiedifferenz Δ_{EM} zwischen dem Grundzustand S_0 und dem angeregten Zustand S_1 dar und errechnet sich wie folgt:

$$\lambda_{EM} = \frac{h \cdot c}{\Delta_{EM}},$$

mit der Planck'schen Konstante h und der Lichtgeschwindigkeit c .

Ein fluoreszierendes Atom bzw. Molekül kann nun Photonen unterschiedlicher Wellenlängen emittieren. Abbildung 2-2 zeigt, dass die Energiedifferenz zwischen angeregtem Zustand und dem Grundniveau variieren und damit unterschiedliche Wellenlängen auftreten können. Bei der Anregung des Materials können unterschiedliche Energieniveaus mit dem Beschuss angeregt werden.

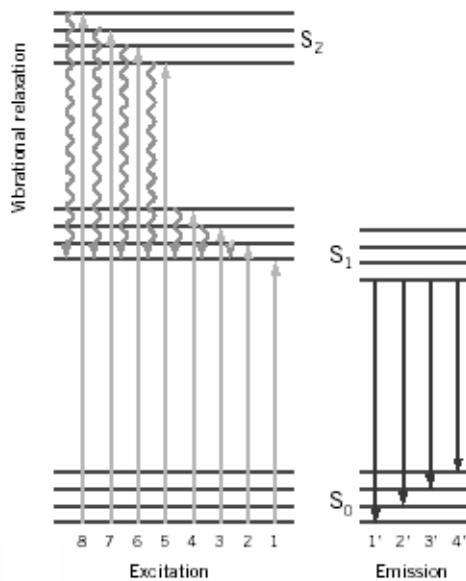


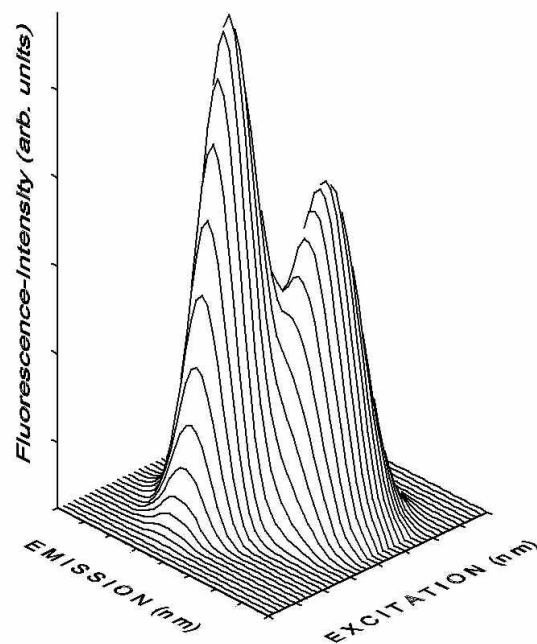
Abbildung 2-2: Darstellung der Energieniveaus eines fluoreszierenden Moleküls [APB, 2000]

2.2 Eigenschaften der Fluoreszenz

Zur Beschreibung von Fluoreszenzerscheinungen lassen sich vier Parameter heranziehen, um Aussagen über materialspezifische Charakteristika zu treffen. Dies sind neben den bereits vorgestellten Absorptions- und Emissionsspektren und deren Intensitäten die Quantenausbeute und das Abklingverhalten, also der zeitliche Verlauf des Fluoreszenzereignisses.

2.2.1 Absorption

Das Absorptions- oder auch Anregungsspektrum eines Moleküls oder Materials beschreibt die Wahrscheinlichkeit der Anregung durch ausgewählte Wellenlängen. Exemplarisch zeigt Abbildung 2-3 dieses Verhalten. Es ist zu erkennen, dass zwei Wellenlängenbereiche ein ausgeprägtes Absorptionsverhalten zeigen. Dies bedeutet, dass mit diesen Wellenlängen das Material am Effektivsten zum Leuchten angeregt werden kann. Aus diesem Grunde fällt der Analyse des Absorptionsspektrums besondere Bedeutung bei der Auswahl der Anregungswellenlänge des Lasers zu, um bei der späteren Materialerkennung optimale Ergebnisse zu erzielen.



2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Abbildung 2-3:
Absorptions- und Emissionsspektren [BRO, 2000]

2.2.2 Emission und Fluoreszenzintensität

Das Emissionsspektrum kennzeichnet für eine definierte Anregungswellenlänge die Intensität der Fluoreszenz. Wie bereits beschrieben, lässt sich die Fluoreszenz auf die Energiedifferenzen zwischen den unterschiedlichen Energiezuständen zurückführen, daher beschreibt das Emissionsspektrum die stoffliche Zusammensetzung des Materials. In den meisten Fällen tritt nicht ein einzelnes Maximum im Emissionsspektrum auf, sondern mehrere - so genannte Leuchtzentren.

Die Fluoreszenzintensität ist linear proportional zur einfallenden Laserintensität [FUC, 2003]. Die absoluten Werte des Fluoreszenzsignals sind jedoch von internen Geräteparametern abhängig. Zu diesen gerätespezifischen Parametern gehört neben der Regelung der Empfängersensitivität durch Spannungsverstellung und der Nutzung von Filtern z.B. auch eine Kalibrierung der Empfangseinheit auf variierende Entfernungen.

2.2.3 Quantenausbeute

Die Quantenausbeute wird aus der Emissions- und der Anregungsenergie berechnet.

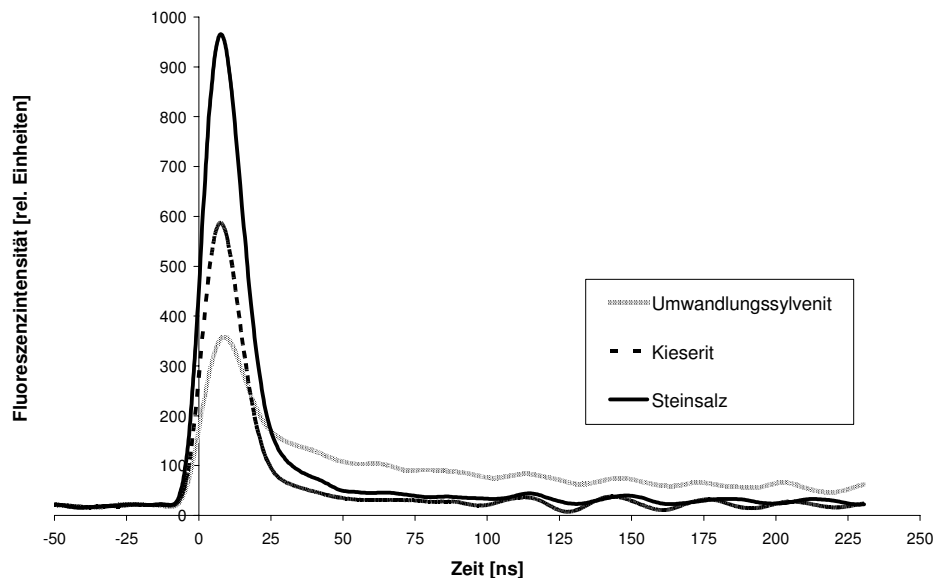
$$\text{Quantenausbeute} = \frac{\sum \text{emitierte Quanten}}{\sum \text{absorbierte Quanten}}$$

Somit stellt der Wert der Quantenausbeute einen spezifischen Wirkungsgrad dar, der durch die Absorptionsfähigkeit des Materials und die Fähigkeit, die absorbierte Energie in emittierte Fluoreszenzsignale zu konvertieren, beeinflusst wird.

2.2.4 Abklingverhalten

Die Intensität eines Fluoreszenzereignisses sinkt im zeitlichen Verlauf nach der Beendigung der Anregung. Diese exponentielle Abnahme der Emissionsintensität wird durch Abklingkurven beschrieben (Abbildung 2-4). Unter Normalbedingungen ist die Lebensdauer der Emission nur von der Anregungsintensität und dem angeregten Material abhängig und somit ebenfalls eine materialspezifische Eigenschaft. Aufgrund von gerätetechnischen Einstellungen des verwendeten LIF-Systems erscheint in der Abbildung bereits ein Fluoreszenzsignal vor dem Nulldurchgang, dies wird bei der Ermittlung des absoluten Signals berücksichtigt.

Abbildung 2-4:
Abklingverhalten verschiedener
Evaporite



Als Beispiel zeigt Abbildung 2-4 das Abklingverhalten verschiedener Salze. Die Evaporite wurden mit einem Laser der Wellenlänge 266 nm angeregt, die zeitaufgelöste Aufnahme des Fluoreszenzsignals erfolgte bei einer Emissionswellenlänge von 400 nm. Neben dem exponentiellen Abklingverhalten, das allen Proben gemeinsam ist, fällt die zeitliche Verzögerung des Abklingens des Umwandlungssylvenits auf. Dieser wurde radioaktiv beein-

flusst und weist daher ein charakteristisches materialspezifisches Abklingverhalten auf.

Mathematisch beschreibt die nachfolgende Formel das Abklingverhalten:

$$I_t = I_0 \times e^{-t/\tau}$$

, wobei τ den Lebensdauerfaktor repräsentiert.

2.3 Materialerkennung mittels LIF

Die Materialerkennung in der Rohstoffgewinnung ist ein relativ junges Einsatzgebiet der laserinduzierten Fluoreszenz. Im Bergbausektor wurde die LIF erstmals in den 1970er Jahren bei der Prospektierung mit Helikoptern in Kanada eingesetzt. Die deutsche Preussag AG entwickelte ein Jahrzehnt später einen auf einem Unimog montierten Analyser, der ebenso wie der Kohlesensor der australischen CSIRO aus finanziellen und geräte-technischen Gründen keinen dauerhaften Erfolg hatte [BRO, 1999]. Wie in Kapitel 2.5 ausgeführt wird, wurden in den letzten Jahren jedoch erfolgreich Geräte bei der Materialerkennung in der Rohstoffgewinnung eingesetzt.

Verglichen mit anderen optischen Analyseverfahren wie der Röntgenfluoreszenz (RFA) oder der Prompt Gamma Neutronenaktivierung (PGNA) bemustert die LIF lediglich die Oberfläche und kann nicht wie andere Verfahren das Material durchleuchten. Andererseits stellen variierende Entfernungen und weitere Strecken zwischen Material und Analyser für die LIF keine Restriktion dar.

Die Leuchtzentren weisen im Falle der LIF im Gegensatz zur RFA relativ breite Banden auf, bei der diese Peaks sehr schmal sind und direkt mit spezifischen Elementen korrelieren. Da die meisten mineralischen Rohstoffe inhomogene Gemische verschiedener Minerale sind, kommt es bei der LIF zu einer Aufsummierung aller Fluoreszenzereignisse. Aus diesen Gründen muss beim Einsatz der LIF berücksichtigt werden, dass nur unter idealen Bedingungen (homogenes Material, wenige Leuchtzentren) das Resultat der Messungen quantitativer Natur ist.

Im Bergbauumfeld findet eine Materialerkennung in der Regel unter klar definierten geologischen und mineralogischen Rahmenbedingungen mit einer begrenzten Anzahl von Wertmineral- und Nebengesteinsklassen statt. In diesem Falle hat sich der Abgleich von Fluoreszenzcharakteristika der Feldmessungen mit denen bekannter Gesteinsklassen nach diesem Prinzip des „Fingerabdrucks“ als sehr effektive Methode der Materialklassifizierung herausgestellt. LIF-Analysegeräte wurden so erfolgreich im Eisenerzbergbau und der Kalk- und Zementindustrie getestet [NIE, 2002, NIE, 2003c].

2.3.1 Grundlagen der Materialerkennung mittels LIF

Bei der Betrachtung der Fluoreszenz von Mineralen und Mineralgemischen liegt der Fokus besonders auf den fluoreszierenden Eigenschaften spezifischer Moleküle und Elemente. Durch den Mehrstoffcharakter von Mineralen ist es zumeist nicht möglich, die Interaktionen zwischen den verschiedenen Elementen und die resultierenden Spektren vorherzusagen. Lediglich für ausgewählte Einzelminerale lassen sich entsprechende Zusammenhänge theoretisch und experimentell eindeutig herstellen [MAR, 1979].

Im Allgemeinen lassen sich die Fluoreszenzerscheinungen von Mineralen auf zwei Hauptursachen zurückführen:

- Intrinsische Fluoreszenzeigenschaften eines Minerals und
- Verunreinigungen (aktivierende Substanzen), die in die Kristallstruktur des Minerals eingebaut sind.

Beispiele für Minerale mit intrinsischen Fluoreszenzeigenschaften sind Scheelit und uranhaltige Minerale. In der Vielzahl der Fälle verursachen jedoch aktivierende Substanzen die Fluoreszenz. Besonders die Elemente der Seltenen Erden wie etwa Scandium oder Yttrium rufen solche Erscheinungen hervor. Versuche mit künstlich erzeugten Mineralen zeigten, dass auch Elemente wie Wolfram, Antimon oder Chrom, Kupfer, Mangan und Titan Aktivatoren sein können [RIE, 2002].

Abhängig von der geologischen Historie eines Minerals können unterschiedliche Verunreinigungen in das Kristallgitter eingebaut werden, so dass keine generellen Aussagen hinsichtlich der Fluoreszenzeigenschaften einzelner Minerale getroffen werden können. Dies bedeutet, dass die speziellen Lagerstättenbedingungen berücksichtigt werden müssen und die Methode der LIF nur im Kontext der lokalen Gegebenheiten eine qualitative Analyse und Klassifikation ermöglicht.

Gewisse Minerale sind jedoch für ihre starke Fluoreszenz bekannt, so dass eine Detektion bereits mit UV-Handlampen geringer Energieleistung möglich ist. Zu diesen Mineralen, die eine durch Verunreinigungen verursachte natürliche Fluoreszenz aufweisen, zählen unter anderem Apatit, Kalzit, Diamanten, Fluorit, Zinkblende und Zinkit.

Neben den Elementen, die Fluoreszenzerscheinungen aktivieren, sind auch solche bekannt, die die Fluoreszenz dämpfen („Quenching“) wie etwa Eisen, Kobalt und Nickel.

An dieser Stelle soll kurz auf das allgemeine Vorgehen bei der Durchführung einer LIF Messkampagne eingegangen werden. Um eine Kalibrierung und Klassifikation von Gesteinen einer Lagerstätte vornehmen zu können, bedarf es des Wissens um die materialspezifischen Fluoreszenzeigenschaften, also den Fingerabdruck. Zunächst werden in Voruntersuchungen in einem Fluorospektrometer, einem Laborgerät, die Absorptions- und Emissionsspektren der relevanten Gesteine untersucht, um die optimalen

Anregungswellenlängen zu ermitteln sowie Erfolg versprechende Bereiche des Emissionsspektrums einzugrenzen. Basierend auf diesen Ergebnissen wird das Analysegerät konzipiert und anschließend mit repräsentativem Probenmaterial kalibriert. Es werden also die klassen- und gesteinspezifischen Referenzmerkmale ermittelt.

2.3.2 Auswertung und Interpretation von LIF-Signalen

Die Klassifikation von Gesteinen bzw. Mineralen basierend auf der LIF beruht wie bereits beschrieben auf dem Vergleich der aktuellen Messwerte mit materialspezifischen Referenzwerten, also dem Prinzip des Fingerabdrucks. Dieses Ziel sollte mit einer minimalen Merkmalsmenge erreicht werden, die zudem möglichst unabhängig von äußeren Störgrößen wie etwa Staub oder Feuchtigkeit im Strahlengang ist.

Unterschiedliche Ansätze lassen sich bei der Merkmalsextraktion bzw. der Bildung von Merkmalen aus den Fluoreszenzsignalen verfolgen. In einem 2-Komponentenmix lässt sich dabei die einfachste Form der Unterscheidung realisieren. In Abbildung 2-5 wäre dies z.B. der Fall bei der Differenzierung zwischen den beiden Roherzklassen. So fluoresziert das Material mit hohen Apatitgehalten sehr stark, während das reine Magnetit Erz (schwarze Linie) nicht durch die Bestrahlung zum Leuchten angeregt wird. Bei dieser Konstellation ist eine Emissionswellenlänge theoretisch ausreichend, um die Materialklasse zu ermitteln.

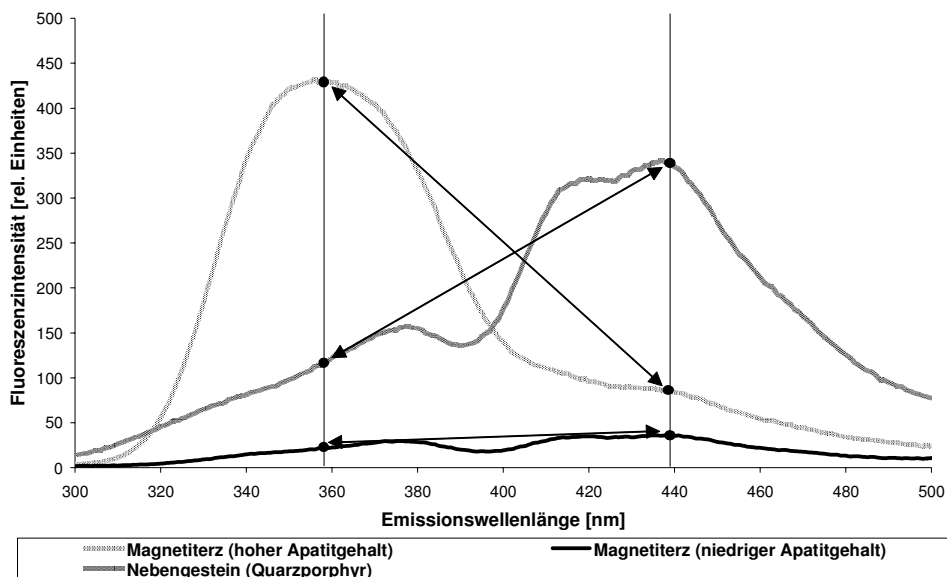


Abbildung 2-5: Merkmalsbildung zur Materialerkennung mittels LIF

Daneben besteht die Möglichkeit, eine qualitative Klassifikation basierend auf ausgewählten Quotienten verschiedener Abschnitte des Spektrums durchzuführen. Im gewählten Beispiel wären dies die beiden signifikanten Peaks des apatithaltigen Erzes und des Nebengesteins bei 360 nm bzw. 440 nm. Die Steigung dieser Quotienten erlaubt eine eindeutige Unterscheidung zwischen allen drei Gesteinsklassen.

2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

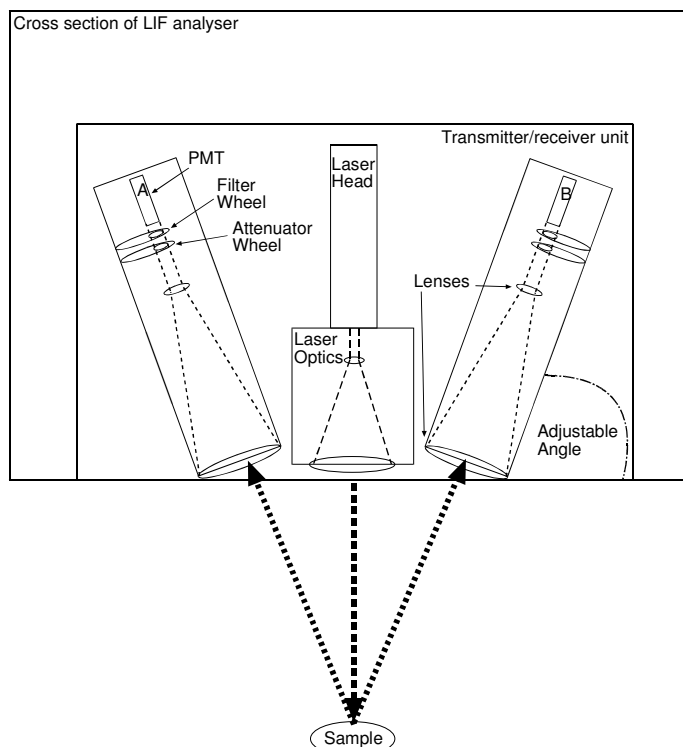
Ein wesentlicher Vorteil der Nutzung von Quotienten als Merkmale ist die Unabhängigkeit von äußeren Einflüssen wie Staub oder Feuchtigkeit. Diese Störgrößen wirken sich dämpfend auf die Fluoreszenzintensität aus, dies jedoch relativ zur Signalhöhe, so dass bei der Quotientenbildung die äußeren Effekte keine Ergebnisverfälschung bewirken [BAY, 2002].

Bei ähnlichen Emissionsspektren bietet sich unter Umständen der Einsatz mehrerer Anregungswellenlängen zur Materialunterscheidung an. Daneben bietet auch die Betrachtung des zeitlichen Verlaufs des Fluoreszenzereignisses eine Möglichkeit der Differenzierung von Materialklassen, da im Allgemeinen organische und anorganische Verbindungen durch ein ungleiches Abklingverhalten charakterisiert werden.

2.4 Genereller Aufbau eines LIF Systems

Zu den Hauptkomponenten eines LIF Analysers zur Materialerkennung zählt neben der Lichtquelle die Empfangseinheit zur Detektierung der induzierten Fluoreszenz. Beispielhaft ist hier der Aufbau der Detektoreinheit mit Photomultipliern dargestellt (Abbildung 2-6). Um mineralische Rohstoffe anzuregen bedarf es einer energiereichen Strahlung im ultravioletten Spektralbereich unter 400 nm, die zudem einen monochromatischen Charakter aufweist, um eindeutige Korrelationen zwischen Anregungs- und Emissionsspektrum zu gewährleisten.

Abbildung 2-6:
Schematische Darstellung
des Messprinzips und gene-
rellen Aufbaus eines LIF
Analysegerätes [RIE,2002]

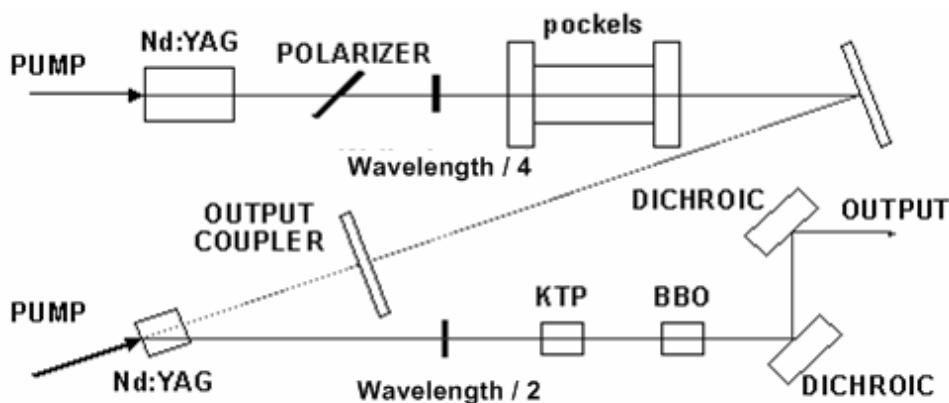


Darüber hinaus gehören eine Vielzahl von optischen Bauteilen wie Spiegel, Prismen und Blenden sowie elektronische Komponenten zur Steuerung und Datenverarbeitung zum komplexen System eines LIF Analysers.

2.4.1 Lichtquelle

Zur Erzeugung des ultravioletten Lichtes zur Anregung des Materials über Entfernungen bis zu mehreren Metern stehen grundsätzlich Gasentladungslampen und Laser zur Verfügung. Aufgrund der monochromatischen Eigenschaften des Laserlichtes, den hohen Ausgangsleistungen und der Reduktion des apparativen Aufwandes kommen in der Rohstoffgewinnung ausschließlich gepulste Festkörperlaser zum Einsatz. Nachfolgend wird auf die Funktionsweise des Nd:YAG-Lasers eingegangen, der weit verbreitet ist und bei dem in dieser Arbeit genutzten Gerät verwendet wurde.

Bei diesem Laser wird das Licht von Nd^{3+} -Ionen erzeugt. Das optisch aktive Nd^{3+} -Ion ist in YAG-Kristalle (Yttrium-Aluminium-Granat – $\text{Y}_3\text{Al}_5\text{O}_{12}$) eingebettet, die eine hohe Lichtverstärkung sowie gute mechanische und thermische Eigenschaften aufweisen. Die Anregung erfolgt durch optisches Pumpen mit Blitzlampen oder Diodenbänken. Unter normaler Betriebstemperatur emittieren Nd:YAG-Laser Licht der Wellenlänge 1.064 nm.



2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Abbildung 2-7:
Optisches System: Laser –
harmonische Generatoren
[THA, 2002]

Nichtlineare Kristalle (harmonische Generatoren) wandeln das infrarote Licht in die harmonischen Wellenlängen 532 nm, 355 nm und 266 nm um (siehe Abbildung 2-7). Da die nachgeschalteten harmonischen Generatoren auch in abgeschwächter Form Licht der vorangegangenen durchlassen, beobachtet man erhöhte Fluoreszenzintensitäten um 532 nm, der Reflexion der zweiten harmonischen Wellenlänge.

Der Wirkungsgrad eines Nd:YAG-Lasers liegt zwischen vier und zehn Prozent. Festkörperlaser erreichen Lebensdauern von bis zu 50 Millionen Pulsen. Bei der LIF in der Rohstoffindustrie werden in Abhängigkeit der A/D-Wandler Laserpulsfrequenzen zwischen 15 und 40 Hz genutzt.

2.4.2 Strahlengang

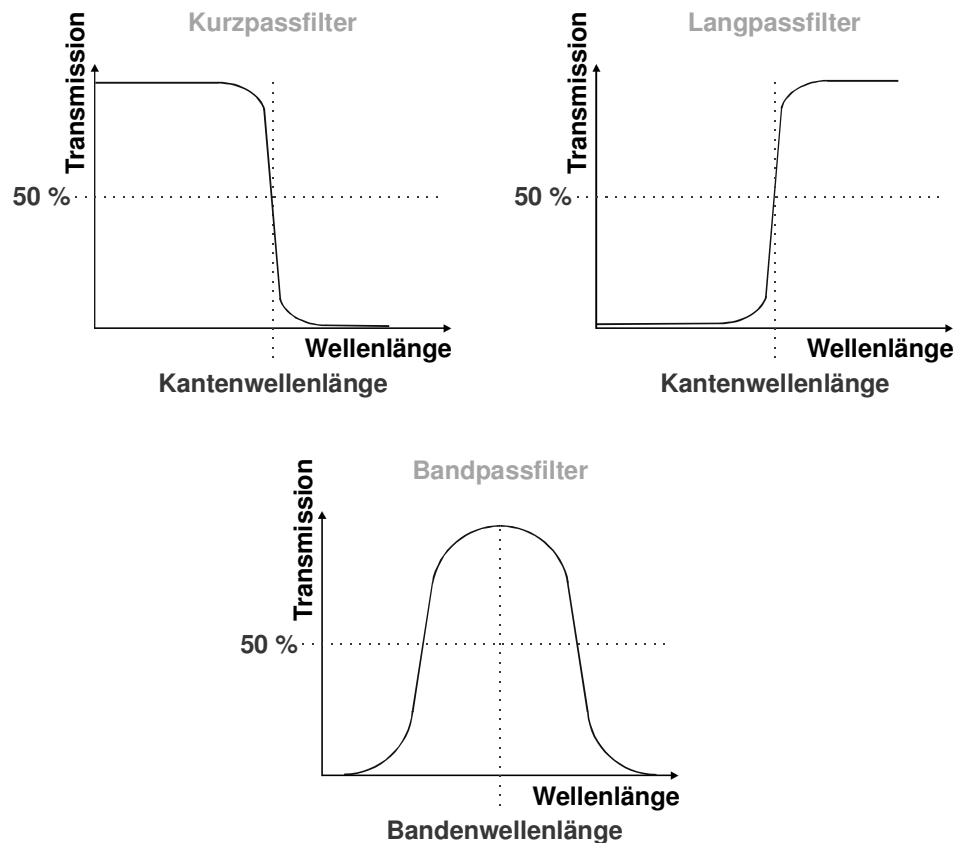
Neben der Aufweitung und Bündelung des Lichtes mittels Prismen und Linsen kommt der Selektion der Wellenlängen bei der Materialanalyse mittels LIF-Technologie eine große Bedeutung zu. Auf der Sendeseite des Systems erfolgt zunächst eine Aufweitung des Laserstrahls, um eine größere

2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Fläche bemustern zu können. Bevor das Fluoreszenzsignal dann in die Empfangseinheit eintritt, wird das Licht wieder gebündelt und passiert eine Reihe von Blenden und Filtern, die eine Selektion der Wellenlänge vornehmen. Ausnahme bilden Messinstrumente, die parallel das gesamte Spektrum analysieren können, hier ist eine Selektion, also eine Vorauswahl der Wellenlänge nicht erforderlich.

Beispiele für Filter sind in Abbildung 2-8 dargestellt [FUC, 2003]. Die genannten Filter sind in ihrem spezifischen Arbeitsbereich keine idealen Lichtleiter. Das bedeutet, dass die Lichtintensität des passierenden Lichtes bis zu 70% abgeschwächt werden kann.

Abbildung 2-8:
Kurz-, Lang- und Bandpassfilter



Kurzpassfilter sind Kantenfilter, die Licht transmittieren, welches kurzwelliger ist als ihre eigene Kantenlänge. Licht mit einer längeren Wellenlänge wird hingegen gesperrt. Die so genannte Kantenwellenlänge ist die Wellenlänge, bei der die Transmission 50% des Maximums erreicht. Nach dem umgekehrten Prinzip arbeiten Langpassfilter. Sie transmittieren Licht, das länger ist als ihre eigene Kantenlänge. Bandpassfilter arbeiten nach dem Prinzip eines Bandenfilters. Ihr Transmissionsbereich ist zu beiden Seiten begrenzt.

2.4.3 Empfangseinheit

Prinzipiell kann jedes photosensitive Messgerät zur Erfassung und Quantifizierung der einfallenden Fluoreszenzsignale verwendet werden. Photo-

multiplier (PMT) erfassen auch die generell sehr niedrige Intensität bei der Nutzung der LIF über weite Entfernungen. Es können auch Photodioden, Diodenarrays oder CCD-Kameras („Charge-coupled detector arrays“) zum Einsatz kommen, deren Empfindlichkeit jedoch bis zu mehrere Zehnerpotenzen geringer ist. Im letzten Fall ist jedoch das Vorschalten einer Wellenlängenselektion nicht erforderlich.

Die Funktionsweise des PMT, auch Sekundärelektronenvervielfältiger genannt, beruht auf dem physikalischen Prinzip des Photoeffektes. Die auf die Photokathode treffenden Photonen lösen Photoelektronen aus, die durch ein elektrisches Feld in Richtung der ersten Dynode beschleunigt werden. Dort werden pro Elektron bis zu 100 Sekundärelektronen ausgelöst, die ihrerseits zur nächsten Dynode driften. Gemessen wird das elektrische Messsignal an der Anode.

Die Empfindlichkeit eines Photomultipliers ist von der eingehenden Lichtwellenlänge und von seiner Bauart abhängig. Auf der langwelligen Seite wird die Grenze der Empfindlichkeit von den Materialparametern der Photokathode und auf der kurzwelligen Seite vom Fenstermaterial festgelegt. Am effizientesten arbeiten PMT's unter stabilen Umgebungsbedingungen und niedrigen Temperaturen, eine Kühleinrichtung optimiert daher in der Regel die Arbeit der Empfangseinheit.

2.4.4 Weitere Bestandteile

Zum Betrieb eines LIF-Analysesystems sind zu den bereits genannten optischen Bauteilen Komponenten zur Datenverarbeitung und Steuerung erforderlich. Kühlaggregate und Übergangsstromversorgungselemente stellen den reibungslosen Betrieb der Geräte sicher.

U.a. gehören Umlenkspiegel, dichroitische Spiegel, Lichtwellenleiter, Teleskopoptiken in die Kategorie der optischen Bauteile. Ihre Rolle für das Analysegerät darf nicht unterschätzt werden, da sie in einigen Fällen aufgrund der hohen Energiedichte der Laserstrahlung erheblichen Belastungen unterliegen. Besonders Lichtwellenleiter weisen beim Betrieb im kurzwelligen UV-Bereich nur sehr kurze Lebensdauern auf, da sie im Laufe ihres Lebens zunehmend „erblinden“, d.h. der Transmissionsgrad beständig sinkt [THI, 2003].

Das elektrische Analogsignal des PMT wird im A/D-Wandler in ein digitales gewandelt, damit es weiter verarbeitet werden kann. Um dabei neben dem integrierten Intensitätssignal auch eine zeitaufgelöste Analyse der Fluoreszenz gewährleisten zu können, bedarf es einer hohen Samplerate mit Auflösungen von bis zu 5 ns. Spezielle Softwaretools geben die Ergebnisse direkt graphisch aus bzw. stellen sie für weitere Analysen bereit.

2.5 LIF-Geräte für den Einsatz in der Rohstoffindustrie

Bei den derzeit in der Rohstoffindustrie eingesetzten LIF-Analysegeräten lässt sich grundsätzlich zwischen zwei unterschiedlichen Arbeitsweisen differenzieren. Die Aufgabenstellung beider Systeme ist jedoch identisch: Die online Ermittlung der stofflichen Merkmale des gewonnenen oder noch zu gewinnenden Materials, um die Materialströme entsprechend ihrer Qualitäten zu steuern.

Stationäre Anlagen bemustern Stoffströme über eine konstante Entfernung und werden zur Klassifikation über Bandanlagen und Bahnentladeeinrichtungen oder Schaufelfüllungen von Radladern eingesetzt. Dieser Bereich der Qualitätssicherung in der Rohstoffgewinnung ist bereits kommerziell mit Geräten in Kleinstserienfertigung erschlossen. In der Entwicklungs- und Prototypentestphase befinden sich hingegen noch die LIF Scanner / Analyser, die nicht nur punktuell Informationen über die Fluoreszenzeigenschaften des Materials ermitteln, sondern Profile bemustern und damit flächhaft Daten erheben.

Daneben wird zurzeit im Rahmen eines InnoNet - Forschungsvorhabens ein LIF-System entwickelt, das die stoffliche Zusammensetzung von Einzelkörnern analysieren soll. Das Projekt wird unter anderem vom Lehr- und Forschungsgebiet „Aufbereitung mineralischer Rohstoffe (AMR)“ an der RWTH Aachen und dem Aachener Fraunhofer Institut für Lasertechnik (ILT) bearbeitet und firmiert unter dem Namen „EIGER“ (Einzelkornanalyse zur Mineraltrennung mit hohem Massendurchsatz bei der Gewinnung von Primärrohstoffen).

2.5.1 Stationäre Anlagen

Stationäre Anlagen kamen erstmalig Anfang 2000 im schwedischen Eisenerzbergwerk Kiruna zum Einsatz. Die Problematik der Unterscheidung des apatithaltigen Erzes von reinem Magnetit Erz auf der einen und Nebengestein auf der anderen Seite wurde bereits in Abbildung 2-5 verdeutlicht. Seit diesem ersten Testeinsatz im Eisenerz folgte eine Reihe von Betriebsstudien in unterschiedlichen Bereichen der mineralischen Rohstoffgewinnung. Unter anderem wurden stationäre LIF-Anlagen im Kalkstein, Dolomit und Magnesit zur Kontrolle der Gehalte an MgO , Fe_2O_3 and Al_2O_3 eingesetzt (Abbildung 2-9). Die Unterscheidung zwischen Nebengestein und Wertmineral stand bei Messkampagnen im Diamantenbergbau sowie Gold-, Kupfer- und Blei-Zink-Bergbau im Vordergrund [WEB, 2001; RIE, 2003].



2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Abbildung 2-9:
Stationärer LIF-Analyser im Einsatz in einem Kalksteinbruch

Derzeit gibt es drei Anbieter von stationären Systemen zur Qualitätssteuerung mittels laserinduzierter Fluoreszenz in der mineralischen Rohstoffgewinnung. Die Scan & Sort GmbH, Wedel, konnte 2003 ein Gerät für den Einsatz im österreichischen Magnesitbergbau verkaufen. Die Sommer GmbH, Wedel, bietet LIF-Systeme an, wie sie bei den Testkampagnen in Kiruna 2000 / 2001 zum Einsatz kamen. Die finnische INA Oy vertreibt ebenfalls ähnliche Analysegeräte für den Einsatz über Bandanlagen und zur Bemusterung der Schaufelfüllungen von Radladern.

Die generelle Funktionsweise eines stationären Geräts zur Qualitätssteuerung von Schüttgütern verdeutlicht Abbildung 2-10. Nach der Kalibrierung des Gerätes auf die charakteristischen Fingerabdrücke der auftretenden Gesteinsklassen erfolgt die online Zuordnung des Materials in eine dieser Klassen beim Durchfahren der Messstrecke. Anschließend wird der Materialstrom entsprechend der Qualität entweder direkt der Aufbereitung zugeführt oder vorbehandelt. Auf diese Weise wird der Gesamtwertstoffanteil in der Aufbereitungsanlage bei gleichzeitiger Volumenreduktion und Maximierung des Durchsatzes der Anlage erhöht, da die eigentliche Aufbereitung weniger Material behandeln muss, jedoch insgesamt mehr Rohmaterial auf den Vorbrecher aufgegeben werden kann.

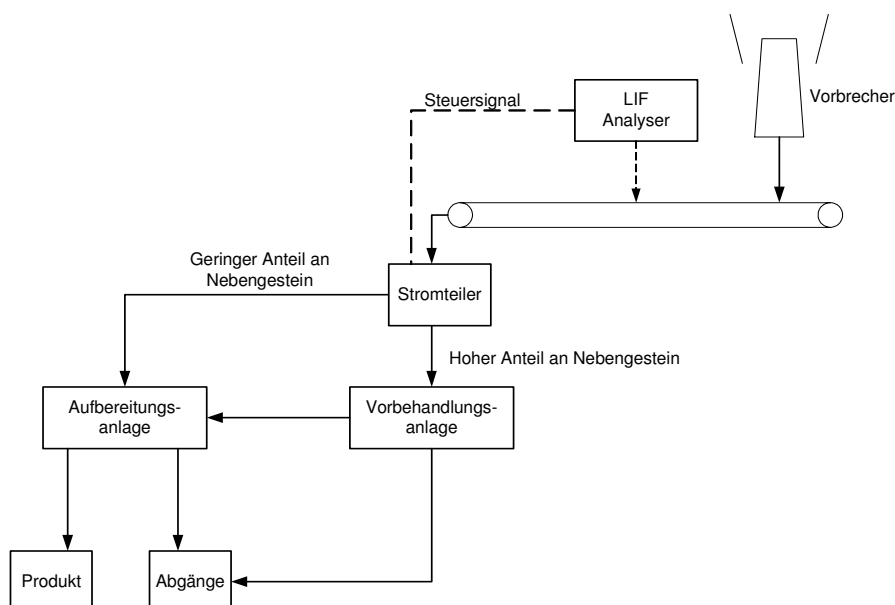


Abbildung 2-10:
Schematische Darstellung einer LIF-Analysers zur Steuerung eines Förderstroms [NIE, 2003c]

2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Kombinationen von stationären LIF-Geräten mit anderen Klassifikations- und Analyseeinheiten wie optischen Erkennungssystemen erscheinen denkbar.

2.5.2 LIF Scanner

Seit Mitte 2002 ist auf dem Markt neben den oben beschriebenen stationären Anlagen auch die LIF Scannertechnologie erhältlich. Die Besonderheit dieses Anlagentyps ist es, dass komplette Bruchwände und Stöße oder auch geschossenes Haufwerk in situ hinsichtlich der Zusammensetzung mit der laserinduzierten Fluoreszenz untersucht werden können. Die Idee dieses Ansatzes liegt daran begründet, dass die Identifizierung von Trenn- und Grenzflächen zwischen unterschiedlichen Materialien entscheidend für die Steuerung von Gewinnungsmaschinen ist.

Das Lehr- und Forschungsgebiet Betriebsmittel für die Gewinnung mineralischer Rohstoffe (BGMR) an der RWTH Aachen konnte den LIF Scanner / Analyser, Typ LA SC 2-1-2, im September 2002 mit Mitteln des Hochschulbauförderungsgesetzes (HBFUG) anschaffen. Das Gerät wurde von der LIF GmbH, Wedel, gefertigt. Im Rahmen des durch die Deutsche Forschungsgemeinschaft (DFG) geförderten Forschungsvorhabens „Bergmännische Nutzbarkeit der Grenzschichterkennung in mineralischen Lagerstätten basierend auf dem Prinzip der laserinduzierten Fluoreszenz“ wurden im Zeitraum zwischen Mai 2002 und April 2004 zahlreiche Rohstoffe und Lagerstätten mit dem LIF Scanner / Analyser (Abbildung 2-11) bemustert.

Abbildung 2-11:
LIF Scanner / Analyser im
Technikum des BGMR

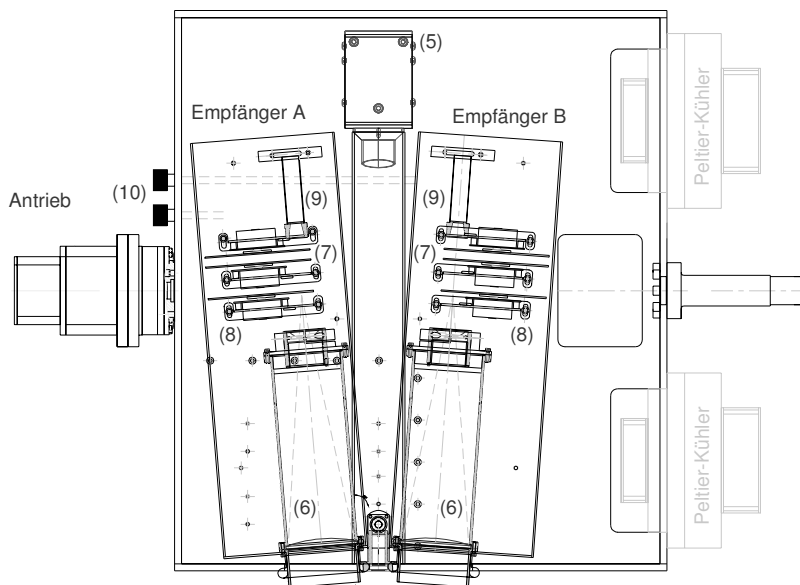


Hauptkomponenten sind zwei diodengepulste Nd:YAG-Laser vom Typ DIVA II der Firma Thales (Frankreich) mit Anregungswellenlängen im tiefen UV bei 266 nm und bei 355 nm, die beiden PMT basierten Empfangseinheiten und der schwenkbare Scannkopf. Die technischen Daten des LIF Scanner / Analysers sind im Anhang A beigefügt.

Die von den Lasern mit einer maximalen Pulsfrequenz von 20 Hz emittierten Strahlen werden von einem Durchmesser von 1,5 mm auf ca. 10 mm aufgeweitet, bevor sie den Scannkopf verlassen. Entsprechend sinkt die Energiedichte auf 0,1775 mJ/cm² an der Austrittsoptik, die jedoch ausreichend ist, um Minerale in Entfernungen über 10 m zum Fluoreszieren anzuregen. Die Laser können alternierend oder einzeln arbeiten, jedoch nicht gleichzeitig.

Ein Servoantrieb mit Absolutencoder kann den Scanner-Messkopf um 180° drehen: Von der horizontalen Ausgangsstellung um plus 39° nach oben und minus 141° nach unten. Die Steuerung erfolgt mittels Software durch den Bediener am Computer.

Abbildung 2-12 zeigt die Empfangsseite des LIF Scanner / Analysers. Der vom pulsierenden Laser angestrahlte Bereich wird mit zwei lichtstarken Empfangsoptiken (6) betrachtet und das induzierte Fluoreszenzsignal gelangt gebündelt durch die selektiven Bandpassfilter (7) sowie einer Blendeinrichtung (8) zu den PMT's (9). Zur Messung der Distanz zwischen Analysegerät und dem angestrahlten Material wird ein im infraroten Spektrum arbeitender Entfernungsmesser (5) eingesetzt, der gleichzeitig den maximalen Sicherheitsbereich überwacht. Entsprechend der Entfernung können die Empfangsoptiken mit Stellschrauben (10) auf den optimalen Brennpunkt fokussiert werden.



Der als PMT A bezeichnete Photomultiplier erfasst das Spektrum zwischen 265 nm und 640 nm, während PMT B die Fluoreszenzintensitäten zwischen 365 nm und 720 nm erfasst. Die Bandpassfilter können die entsprechenden Emissionswellenlängen selektieren. Mit den Blenden kann die Intensität des einfallenden Lichts geregelt werden, um eine Übersteuerung der photosensitiven Empfangseinheiten zu verhindern. Des Weiteren lässt sich die an den PMT anliegende Spannung verändern, um eine für die Auswertung

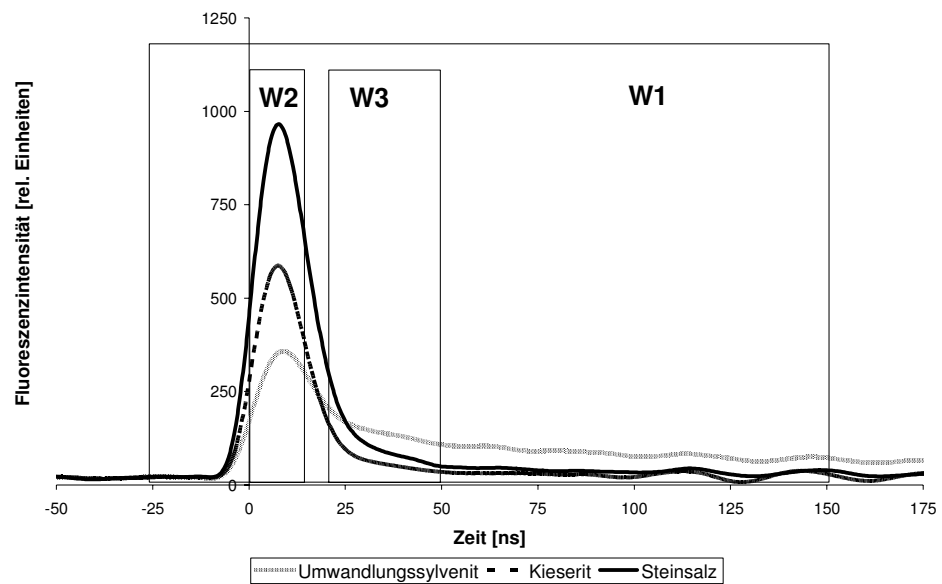
2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Abbildung 2-12: Empfangsseite des LIF Scanner / Analysers [LIF; 2002]

2 Physikalische und technische Aspekte der laserinduzierten Fluoreszenz

Abbildung 2-13:
Zeitfenster W1, W2 und W3
zur Ermittlung der Intensitätsintegrale

optimale Signalhöhe zu erhalten. Die maximal mögliche Spannung beträgt 1.200 V.



Die im separaten Schaltschrank untergebrachte Datenerfassung digitalisiert die Fluoreszenzsignale mit einer Frequenz von 200 MHz, also einer zeitlichen Auflösung von 5 ns. Die Messauswertung der Fluoreszenzsignale beinhaltet neben den Werten der Entfernungsmessung zwischen Gerät und Objekt und der Winkelstellung des Scanner-Messkopfes auch Informationen über Bandpassfilter und Blende. Diese ASCII-Daten werden in txt-Format ausgegeben. Zur zeitaufgelösten Analyse der Signale kann die Abklingkurve detailgenau aufgenommen (vergleiche Abbildung 2-4) oder über interpolierte Intensitätsintegrale (W1, W2 und W3) ausgewertet werden, wie Abbildung 2-13 exemplarisch für drei Zeitspannen zeigt.

Die LIF-Software ermöglicht eine Zusammenfassung von Fluoreszenz-messdaten und Primärdaten sowie deren Darstellung auf dem Bildschirm des Bediener-Computers in Echtzeit. Eine weiterführende Datenauswertung, wie z.B. Ermittlung von Materialgruppen oder Qualitätsklassen, wird dem Bediener durch Zugriff auf die Primärdaten ermöglicht.

Der Schaltschrank enthält darüber hinaus Bedienelemente und Zustandsanzeigen für den LIF-Analyser wie auch die Schnittstellen für die Ethernet-Verbindung zu dem Bediener-Computer, die Betriebszustandsanzeige, die fernbedienbare Sicherheitsverriegelung sowie Anschlüsse für verschiedene Leitungen zum Scanner-Messkopf.

3 Praktische Untersuchungen in der Braunkohlegewinnung

Durch die veränderten Rahmenbedingungen auf dem liberalisierten Strommarkt haben sich für die Braunkohlegewinnung im rheinischen Revier ebenfalls die Prioritäten verschoben. Generell definiert die Optimierung der Stromerzeugung den Rahmen für sämtliche Aktivitäten der RWE Power AG; die nun als integrierter Bergbau- und Energieerzeugungskonzern am Markt agiert. Damit ergeben sich drei wesentliche Ziele: eine Minimierung der Stromerzeugungskosten, eine Maximierung der Wirkungsgrade bei der Verstromung und eine Optimierung der Verfügbarkeit der Braunkohlekraftwerke. Für den Tagebaubetrieb bedeutet der letzte Punkt, dass die Braunkohle bedarfsgerecht in den geforderten Mengen und Qualitäten bereitgestellt wird. Das selektive Hereingewinnen der Lagerstätte, das die Anforderungen der Kraftwerke und Veredlungsbetriebe bereits während der Gewinnung berücksichtigt, ist daher entscheidend für den wirtschaftlichen Erfolg.

Die Einführung moderner Geoinformationssysteme und computergestützter Lagerstättenmodelle, die sensorische Erfassung der Motorströme der Schaufelradbagger zur Unterstützung der Fahrer beim Einhalten der Schneidhorizonte, die tagesscharfe Disposition von Gewinnungsgeräten und die Optimierung der Bunkerbewirtschaftung in qualitativer und quantitativer Hinsicht lassen jedoch weiterhin die entscheidende Frage offen, ob die hereingewonnene Kohle auch tatsächlich mit der prognostizierten stofflichen Zusammensetzung übereinstimmt.

Seit 2003 setzt daher die RWE Power z.B. im Tagebau Hambach an der Übergabe des Strossenbandes im Bereich des Bandsammelpunktes ein Analysegerät ein, das die Qualitätsmerkmale der Rohkohle online mit Hilfe der Röntgenfluoreszenzanalyse (RFA) bestimmt, um so mögliche Diskrepanzen zur erwarteten Kohlequalität zu erkennen. Nachteilig bei dieser Installation ist zum einen die zeitlich verzögerte Information für den Baggerführer über mögliche Verunreinigungen, da die Kohle in Abhängigkeit vom Standort des Gewinnungsgeräts bis zu mehreren Minuten zum Analysator benötigt. Zum anderen bedarf es beim Einsatz der RFA erheblicher Anstrengungen an die Probenahme und Teilung, so dass nur kleine Stoffströme aus dem Gesamtmaterialfluss bemustert werden können, was die Repräsentativität beeinflusst [LIN, 2003].

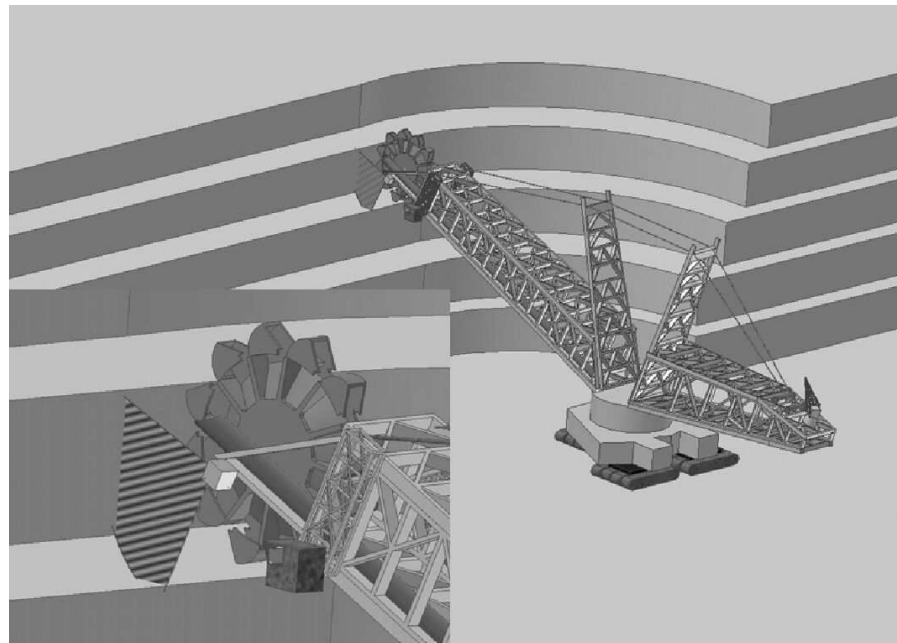
Das Erkennen der kritischen Schichten innerhalb der Braunkohlenflöze etwa in Form von Zwischenmitteln sowie die trennflächenscharfe Detektierung von Übergängen zum Nebengestein am Hangenden und Liegenden der Lagerstätte direkt vor Ort am Stoß bzw. in der Schneidspur würde neben einer besseren Schneidhorizontsteuerung auch eine Optimierung der Kohlequalität bedeuten.

3 Praktische Untersuchungen in der Braunkohlegewinnung

Im Rahmen des gemeinsamen Forschungsvorhabens „Trennflächenerkennung mittels laserinduzierter Fluoreszenz (LIF)“ untersuchte das Lehr- und Forschungsgebiet Betriebsmittel für die Gewinnung mineralischer Rohstoffe (BGMR) an der RWTH Aachen zusammen mit der RWE Power AG zwischen Juni 2002 und Juli 2003 die Möglichkeiten des Einsatzes des innovativen optischen Verfahrens der laserinduzierten Fluoreszenz zur Qualitätssteuerung direkt im Gewinnungsprozess.

Ziel des Projektes war die Untersuchung und Bewertung des Potenzials der LIF für ein einsetztaugliches Sensorsystem zur Unterscheidung von Kohle und Nebengestein unmittelbar am Schaufelrad (Abbildung 3-1). Es wurde untersucht, ob auf Basis dieser Technik eine Steuerung entwickelt werden kann, die den Führer des Schaufelradbaggers in die Lage versetzt, während des Gewinnungsvorganges in Echtzeit die Stratigraphie unabhängig von für das menschliche Auge sichtbaren Qualitätsmerkmalen zu beurteilen. Die Vermeidung von unerwünschten Verunreinigungen während des Gewinnungsprozesses durch eine auf der LIF-Technologie basierenden Warnfunktion würde eine Optimierung des Lagerstättenausbringens durch ein näheres Heranfahren an die jeweilige Grenze zum Nebengestein ermöglichen.

Abbildung 3-1:
Mögliche Einsatzkonfiguration eines LIF Scanners im Braunkohlebergbau



Vorteile eines solchen Systems wären neben der Erhöhung des Lagerstättenausbringens auch eine verbesserte Qualitätssteuerung. Zudem bestünde die Möglichkeit, das Lagerstättenmodell mit den Ergebnissen der Materialklassifikation an der Schaufel abzugleichen und kontinuierlich zu aktualisieren.

In diesem Kapitel soll kurz auf den Tagebau Hambach der RWE Power AG unter bergtechnischen und geologischen Aspekten eingegangen werden. Daneben werden ebenfalls die heutige Qualitätssteuerung und die Anstrengungen bei der selektiven Gewinnung sowie der auf einzelne Kraftwerksblöcke ausgerichteten Kohlebereitstellung angerissen.

Anschließend erfolgt eine Darstellung der Messkampagnen, die im November 2002 und April 2003 im Tagebau Hambach mit dem LIF Scanner / Analyser an kritischen Trennflächen durchgeführt wurden. Auch auf die wesentlichen technischen Fehlerquellen, die die Durchführung der Messungen teilweise erheblich beeinträchtigten, sowie deren Einfluss auf die Aussagen über die Eignung der LIF Technologie soll eingegangen werden. Das Kapitel schließt mit einem Fazit hinsichtlich einer Nutzung der LIF-Signale für eine Materialklassifikation und Trennflächenenerkennung ab.

3 Praktische Untersuchungen in der Braunkohlegewinnung

3.1 Braunkohlentagebau Hambach

Der Tagebau Hambach liegt im Zentrum des rheinischen Reviers und versorgt über die Bahnverbindungen der RWE Power AG die anliegenden Kraftwerke und Veredlungsanlagen mit hochwertiger Braunkohle (Abbildung 3-2). Insgesamt fördert der Tagebau jährlich 40 Mio. t verwertbare Kohle und hält damit ca. einen Anteil von 20% an der deutschen Gesamtbraunkohleförderung.

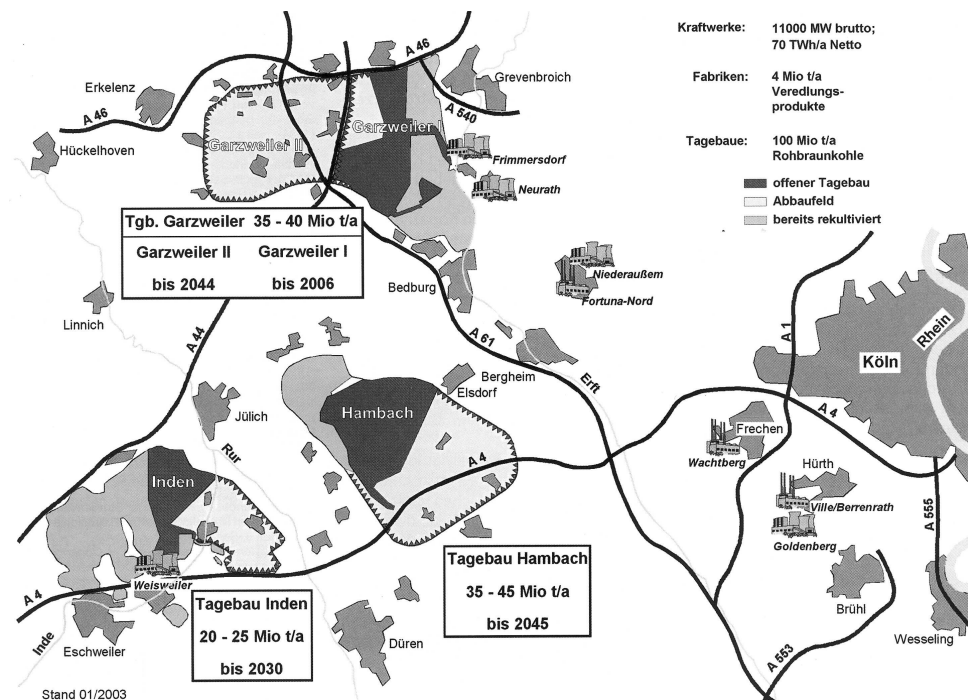


Abbildung 3-2: Einbettung des Tagebaus Hambach ins rheinische Braunkohlerevier [HEM, 2003]

Das genehmigte Abbauegebiet umfasst eine Fläche von ca. 8.500 ha östlich der Städte Düren und Jülich mit einem ursprünglichen Kohleinhalt von annähernd 2,5 Mrd. t. Das Abraum zu Kohle - Verhältnis des in der Erftscholle liegenden Tagebaus beträgt 6,2:1 ($m^3:t$), d.h. es werden bis ins Jahr 2045 mehr als 15,4 Mrd. m^3 Abraum gefördert. Zusätzlich werden jährlich 400 Mio. m^3 Wasser gehoben. Die Personalstärke wurde in den letzten Jahren durch Rationalisierungsmaßnahmen und Outsourcing von Nichtkernaufgaben kontinuierlich reduziert, für Ende 2004 wird eine Belegschaft von 1.480 Mitarbeitern angestrebt [GÄR, 2003a; HEM, 2003].

3 Praktische Untersuchungen in der Braunkohlegewinnung

3.1.1 Eingesetzte Gewinnungstechnik

Die Gewinnung der Braunkohle im Tagebau Hambach erfolgt mit der im rheinischen Revier bewährten Betriebsmittelkette der Großgeräte Schaufelradbagger, Bandanlagen und Absetzer. Auf der Gewinnungsseite kommen sieben Schaufelradbagger zum Einsatz (1 x 110.000 (m³+t)/d, 1 x 200.000 (m³+t)/d und 5 x 240.000 (m³+t)/d), die mit sieben Absetzern auf der Innen- und Außenkippenenseite verbunden sind.

In der Vergangenheit erfolgte die Kohlegewinnung auf der 7. Sohle im Wesentlichen im so genannten 1-Punkt-Betrieb, d.h. mehr als 80% der geförderten Kohle wurde von einem Bagger gefördert. Dies schränkte die Leistungsfähigkeit bei der selektiven Gewinnung der Lagerstätte aufgrund einer geringen Scheibenmächtigkeit, wie sie etwa im Bereich von Zwischenmitteln sowie im Hangenden und Liegenden gegeben ist, deutlich ein. Aus diesem Grunde wurde der Abbau in den letzten Jahren derart umgestaltet, dass ein Mehrpunktbetrieb mit der Beteiligung von mindestens zwei Schaufelradbaggern an der Kohleförderung in Zukunft möglich ist. Planerisch bedeutet dies, dass die 6. Sohle abgesenkt wurde, damit die Kohle auch vom dortigen Großgerät gefördert werden kann. Auch soll mittelfristig ein Bagger der 100.000er Klasse aus dem Tagebau Garzweiler nach Hambach überstellt werden, um eine selektive Gewinnung ballastreicher Kohlepartien wirtschaftlicher zu gestalten. Zwischenzeitlich werden Sonderbetriebe mit mobiler Hydraulikbagger-Dumper-Technologie eingesetzt. Dies hauptsächlich in Bereichen, in denen die Leistungsfähigkeit der 240.000er Schaufelradbagger drastisch abgesunken wäre, etwa in der Nähe des Schwenkpunktes des Tagebaus.

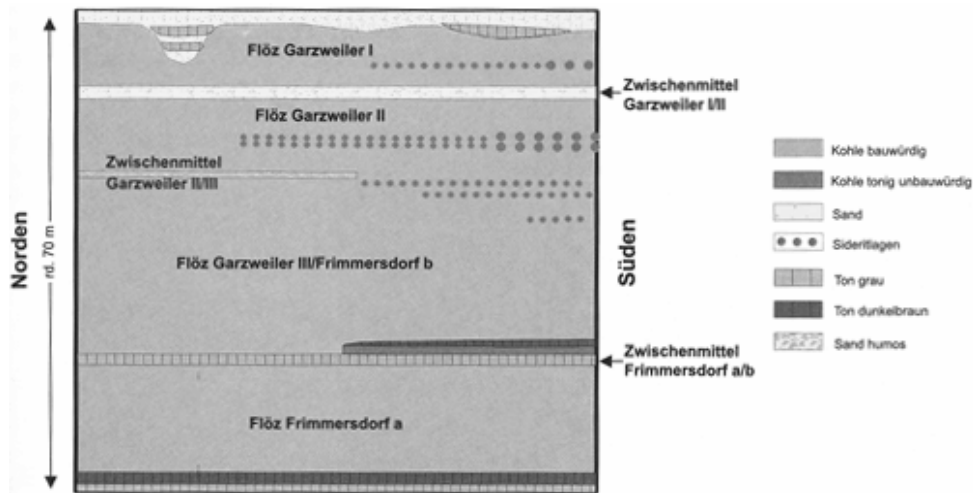
Über zwei Bunkergräben mit einem Gesamtvolumen von 800.000 t gelangt die Braunkohle zur Zugbeladung. Die Züge mit jeweils 14 Wagen zu je 100 Tonnen bringen die Kohle über die Hambachbahn und die Nord-Süd-Bahn je nach Disposition zu den Kraftwerken und Veredlungsbetrieben.

3.1.2 Lagerstättensituation und Qualitätsmanagement

Geologisch lässt sich die Entstehung der Braunkohle des rheinischen Reviers den Zeitaltern des Oligozäns und Miozäns zuordnen. Vor etwa 35 bis sechs Millionen Jahren dehnten sich umfangreiche Moore im Deltabereich der rheinischen Bucht aus, aus denen sich im mittleren und oberen Miozän die heutigen Braunkohleflöze entwickelten.

Die Lagerstätte Hambach fällt mit 3° bis 11° nach N ordosten ein und wird im Osten durch die Rurrand Ost Verwerfung begrenzt. Das Deckgebirge setzt sich im Wesentlichen aus Schluffen und Tonen sowie fluviatilen und marinen Sanden zusammen (Abbildung 3-3). Abgebaut werden die Hauptflöze Frimmersdorf und Garzweiler, die als kompaktes Flöz mit einer Gesamtmächtigkeit von bis zu 70 m anstehen. Es sind drei geringmächtige Zwischenmittel mit tonig-sandigem Charakter eingeschaltet, so dass sich die Lagerstätte in bis zu fünf Teilflöze aufteilt. Die beiden Zwischenmittel Garzweiler I/II und Frimmersdorf a/b sind über den gesamten Abbaubereich

ausgedehnt. Im Liegenden steht ein schluffiger, humoser Ton an [GÄR, 2003b].



3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-3: Vereinfachter geologischer Schnitt des Tagebaus Hambach [GÄR, 2003b]

Entsprechend des Einfallens der Flöze wird die Mächtigkeit der überdeckenden Abraumschichten von derzeit 200 m im südlichen bzw. 300 m im nördlichen Bereich des Abbaus auf voraussichtlich bis zu 400 m Deckgebirge im östlichen Randbereich ansteigen. Mit dieser Teufenzunahme der Braunkohle geht eine erhöhte Verdichtung der Kohle und Reduktion des Wassergehaltes mit einem Anstieg des Heizwertes einher. Im Süden beträgt der Wassergehalt derzeit 55% bei einem Heizwert von 9.500 kJ/kg gegenüber 52.5% und 10.300 kJ/kg im nördlichen Bereich des Abbaubereichs.

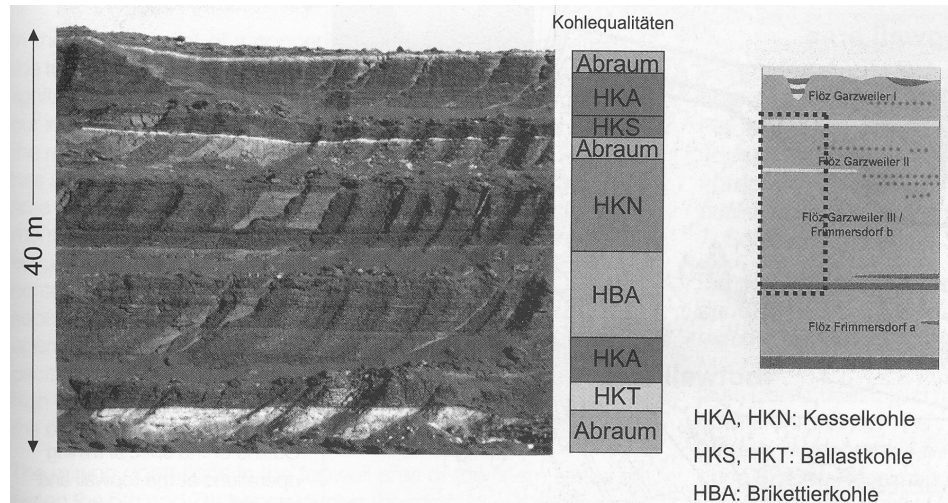
Neben dem Heizwert steigt auch der Natriumgehalt von Süden nach Norden sowie auch in Abbaurichtung an. Eisenhorizonte werden ebenfalls vereinzelt angefahren. Der Gesamtaschegehalt liegt in der Regel unter 3,5% mit Ausnahmen bei den Übergängen im Liegenden und Hangenden der Lagerstätte. Die Hambachkohle zeichnet sich durch einen relativ geringen Schwefelanteil von durchschnittlich 0,25% aus.

Die Ansprache der Kohle im Sinne der Qualitätssteuerung erfolgt bei erhöhten Werten von Verunreinigungen wie etwa Natrium, Kalium, Schwefel, Eisen, Silizium oder Aschegehalt nach einem einheitlichen Schlüssel. Diese ballastreichen Bereiche müssen selektiv in Scheiben gewonnen werden, um später mit qualitativ höherwertiger Kohle gemischt oder in ausgewählten Kraftwerksblöcken verstromt zu werden (Abbildung 3-4).

Die Qualitätssteuerung wird dadurch erschwert, dass sowohl in vertikaler als auch horizontaler Richtung erhebliche kleinräumige Schwankungen der Kohlequalität auftreten. Insgesamt sind die einzelnen Kohlen zu neun Kohlesorten zusammengefasst, die wiederum durch weitere Parameter bei Bedarf unterteilt werden können [FUC, 2003].

3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-4: Scheibeneinteilung am Kohlestoß gemäß Kohlequalitätsschlüssel [GÄR, 2003b]



Kraftwerksseitig werden für die einzelnen Blöcke geeignete Qualitäten festgelegt und blockspezifische Qualitätsspannen fixiert. Die Sortendefinition und die chemisch-physikalischen Qualitätsangaben sind in allen Datenverarbeitungssystemen zur Planung und Steuerung abgelegt und gestatten eine einheitliche, in der gesamten Prozesskette verfügbare Terminologie für alle Beteiligten. Angebot und Nachfrage werden über Datenbankmodule zur Planung und Steuerung von Kraftwerken und Tagebauen abgeglichen. Eine Vielzahl von Steuerungssystemen und Datenbanken soll das reibungslose Zusammenspiel der unterschiedlichen Akteure auf Tagebau- und Kraftwerks- bzw. Veredlungsseite gewährleisten [FUC, 2003]. Dafür stehen derzeit die folgenden Systeme zur Verfügung:

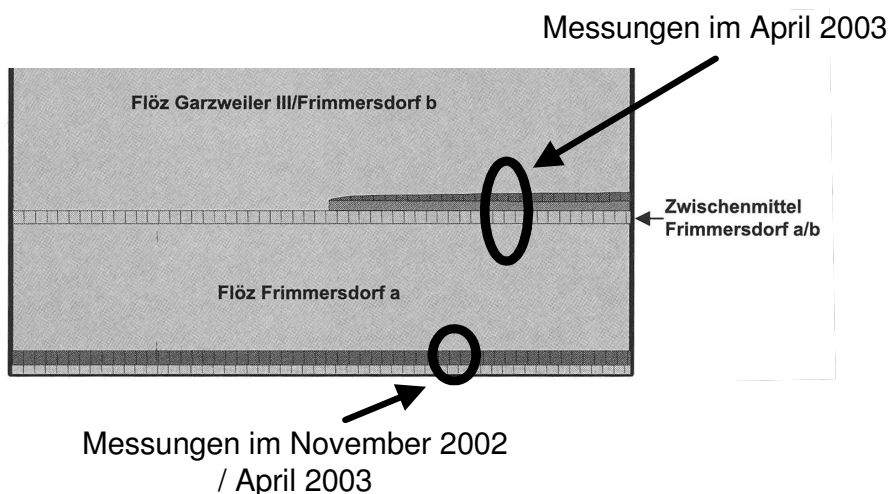
- IBUKOM (Integriertes bunkergestütztes Kohle-Management)
- IKOLA (Interaktive Bestimmung der Kohlequalität auf der Basis von Lagerstätten-dateien)
- QUABUS (Quadermodellbasiertes Bunkerbewirtschaftungssystem)
- KOFIS (Kohleflussinformationssystem)
- KUKIS (Kraftwerks- und Kohleinformationssystem)
- SPANKO (Sortenorientierte Planung von Angebot und Nachfrage Kohle-förderung)

Trotz all dieser Tools muss beachtet werden, dass alle Systeme auf eine gemeinsame Datenbasis zurückgreifen, nämlich die der Information über die geförderte Kohlequalität. Entsprechend können alle Systeme nur dann optimal eingesetzt werden, wenn auch die im Lagerstättenmodell prognostizierte Qualität mit der tatsächlichen übereinstimmt und der Baggerführer auch selektiv die Kohle hereingewinnt, also nicht gewünschte Partien konsequent aushält. In diesem Zusammenhang nimmt die online Qualitätsbestimmung eine bedeutende Rolle ein. Ein automatisches Klassifikationsverfahren zur Materialerkennung stellt hier eine weitere tagesbauseitige Maßnahme zur Sicherung der hohen Qualitätsansprüche dar.

3.2 Messkampagnen im Tagebau Hambach

Den mehrtägigen Feldkampagnen im November 2002 und April 2003 im Rahmen des gemeinsamen Projekts „Trennflächenerkennung mittels laser-induzierter Fluoreszenz (LIF)“ gingen intensive Voruntersuchungen voraus. So wurden im März 2002 zwei markante Übergänge in der Lagerstätte Hambach mit dem Vorläufergerät LA1 über eigens angelegte Rampen vermessen und die lagerstättenspezifische Datenbasis in einer stationären Messreihe durch Handstückvermessungen im Juni 2002 erweitert. Auch der Einfluss von betrieblichen Störgrößen wie Staub und Feuchtigkeit wurde im Rahmen dieser Voruntersuchungen bewertet [BAY, 2002].

Die Feldkampagnen bemusterten Stoßpartien über mehrere Meter bankrecht zur Schichtung. Es wurden Bereiche im Liegenden der Lagerstätte – also dem Übergang zwischen dem Flöz Frimmersdorf a und dem Liegendton – sowie am Zwischenmittel Frimmersdorf a/b untersucht (Abbildung 3-5). Diese Partien stellen im Gewinnungsbetrieb die kritischen Übergangsbereiche dar, bei denen eine selektive Scheibeneinteilung aufgrund der ballastreichen Kohle unumgänglich ist.



3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-5:
Messorte im Tagebau
Hambach

Die Messungen erfolgten am frisch freigelegten Stoß in Schürfgräben, die zuvor durch einen Hydraulikbagger angelegt worden waren (Abbildung 3-6). Die Ausrichtung des LIF Scanner / Analysers wurde so gewählt, dass ausschließlich die Spur zwischen den Zähnen des Hydraulikbaggers abgescannt wurde, um das Bemustern von verschobenem Material und damit eine Verfälschung der Signale zu verhindern. Der LIF Scanner / Analyser wurde vor der Messkampagne im Technikum des BGMR auf eine maximale Entfernung von 15 m kalibriert, um wechselnde Intensitäten aufgrund schwankender Entfernungen zwischen Laser und dem untersuchten Material mathematisch zu korrigieren und so zu einem entfernungsunabhängigen Fluoreszenzsignal zu gelangen.

3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-6:
Versuchsaufbau des LIF
Scanner / Analysers



Jede Scannspur wurde vermessen, dokumentiert und geologisch angesprochen, um eine möglichst genaue Korrelation der Fluoreszenzsignale mit der tatsächlichen Schichtung und den damit verbundenen Informationen über die stoffliche Zusammensetzung zu gewährleisten. Jeder Einzelmessung des LIF Scanner / Analysers wurde eine eindeutige Position aufgrund der Winkelangabe zugewiesen, die eine spätere Umrechnung bzw. Zuordnung zur entsprechenden Stelle am Stoß erlaubt.

Pro Spur wurden abhängig von der Anregungswellenlänge bis zu 15 Filterradkombinationen für die PMT's gewählt, d.h. eine Spur wurde bis zu fünfzehnmal abgescannt. Grund für diese langwierige Versuchsdurchführung war die Tatsache, dass ein möglichst vollständiges Emissionsspektrum aufgenommen werden sollte, um in späteren Einsätzen die Anzahl der Spektralbanden reduzieren zu können.

Die Profile und Ansprachen der einzelnen Scanns sowie die Einstellungen auf der Senderseite (Laserpulsfrequenz, Schwenkgeschwindigkeit) und der Empfängerseite (PMT-Spannung, Filterrad- und Blendeneinstellungen) sind dem Anhang B zu entnehmen. Ebenso ist dort die resultierende Anzahl an Merkmalen dargestellt, die sich aus Abklingzeiten für die einzelnen Filterradeneinstellungen je PMT, den Quotienten aus den Einstellungen der jeweiligen PMT's sowie die Quotienten aus den Einstellungen von PMT A zu PMT B ergibt. Letztere Quotienten ergeben nur im Falle von nicht baugleichen PMT's einen Zugewinn an Informationen.

3.2.1 Messkampagne Liegendes Frimmersdorf a

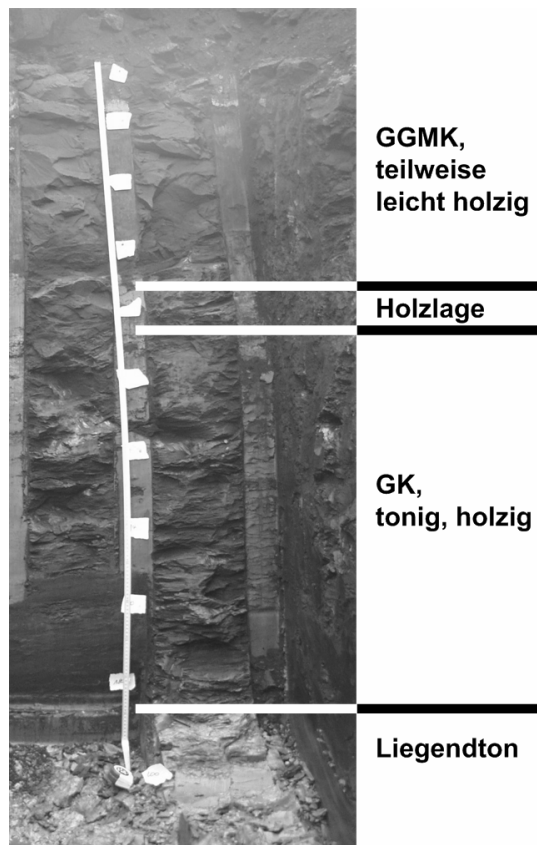
Bei den Messungen im Bereich des Übergangs vom Flöz Frimmersdorf a zum Liegendton wurden zwei unterschiedliche Stratigraphien an insgesamt drei Lokalitäten bemustert. Im November 2002 wurde im nördlichen Bereich des Tagebaus Hambach zunächst ein Schürfgraben untersucht, bei dem die Stratigraphie durch eine charakteristische Holzlage oberhalb der be-

dingt bauwürdigen Gewebekohle (GK) mit hohem Holz- bzw. Tonanteil und dem Liegendton gekennzeichnet ist. Die Holzlage wird von einer als Gewebegrundmassenkohle (GGMK) bezeichneten Braunkohle überlagert. Dieser Scann wird nachfolgend als „Nördlich Liegendrücken“ bezeichnet. Ebenfalls wurde in dieser Messkampagne ein Bereich untersucht, in dem eine charakteristische sandige Kohle sowie eine bedingt bauwürdige Gewebekohle mit hohem Holz- bzw. Tonanteil oberhalb des Übergangs zum Liegendton anstehen („Tiefstufe“).

Auch im April 2003 wurde dieser Bereich der Lagerstätte mit den leicht sandigen Kohlepartien im Liegendbereich des Flöz Frimmersdorf a untersucht. Bei den ersten beiden Messungen wurde lediglich der obere Bereich des hellen Liegendtons (ca. 25 cm) angeschnitten, während im April 2003 mehrere Meter mittels LIF untersucht wurden. Diese Messungen am Liegenden im April 2003 werden in dieser Arbeit jedoch nicht berücksichtigt, da aufgrund von äußeren Einflüssen (extreme Temperaturen) und Gerätefehlern keine verwertbaren Ergebnisse generiert wurden.

Nördlich Liegendrücken

Über einen Winkel von $15,5^\circ$ wurden mit Laser 1 (266 nm) zunächst etwa 4.500 Messpunkte erfasst (Scann #1) und anschließend knapp 1.500 Werte (Scann #2). Die Scanngeschwindigkeit betrug im ersten Fall 0,07%/s, bei der zweiten Bemusterung 0,2%/s. Bei beiden Vorgängen wurde ohne Blenden bei 1.000 V PMT Spannung auf beiden PMT's gearbeitet.



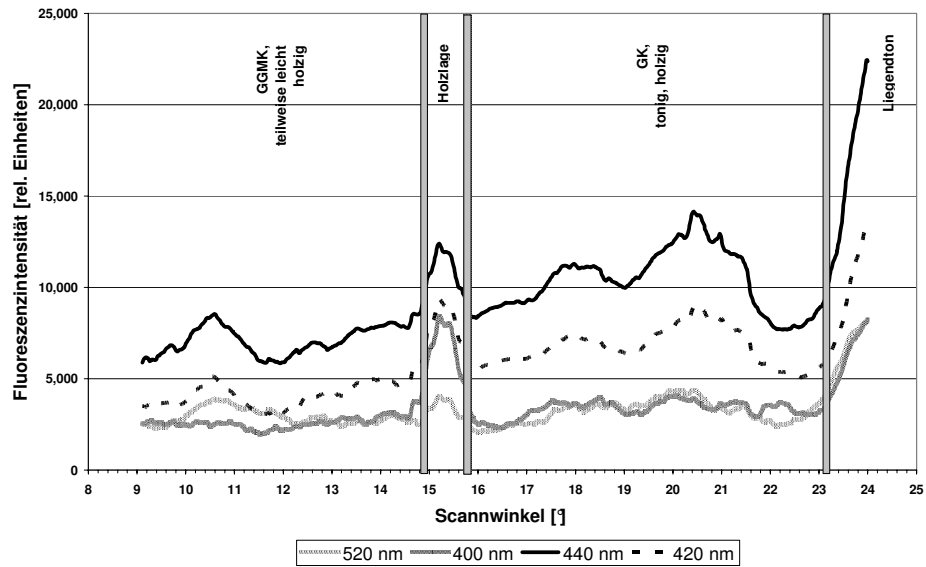
3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-7:
Profil des Scannbereichs
am Liegendübergang, nörd-
lich Tiefstufe, 12.11.2002

3 Praktische
Untersuchungen in der
Braunkohlegewinnung

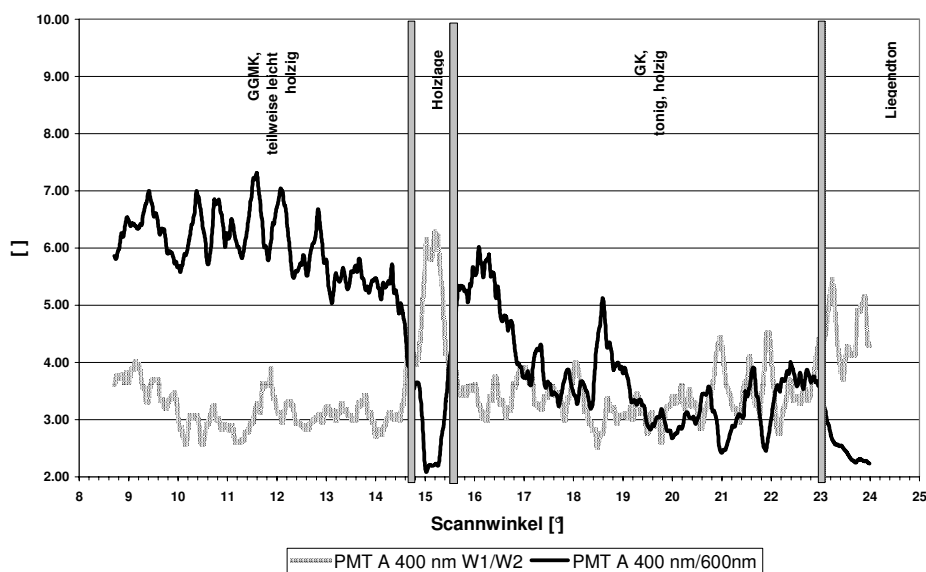
Abbildung 3-8:
 Verlauf der LIF-Intensitäten
 am Liegendübergang, nördlich
 Liegendrücken,
 12.11.2002

Die nachfolgende Abbildung 3-8 zeigt die Intensitätsverteilung über den gescannten Bereich des Profils im Bereich nördlich des Liegendrückens. Deutlich zu erkennen ist der Intensitätspeak auf Höhe der Holzlage sowie der abrupte Anstieg beim Übergang im Bereich zum Liegendton.



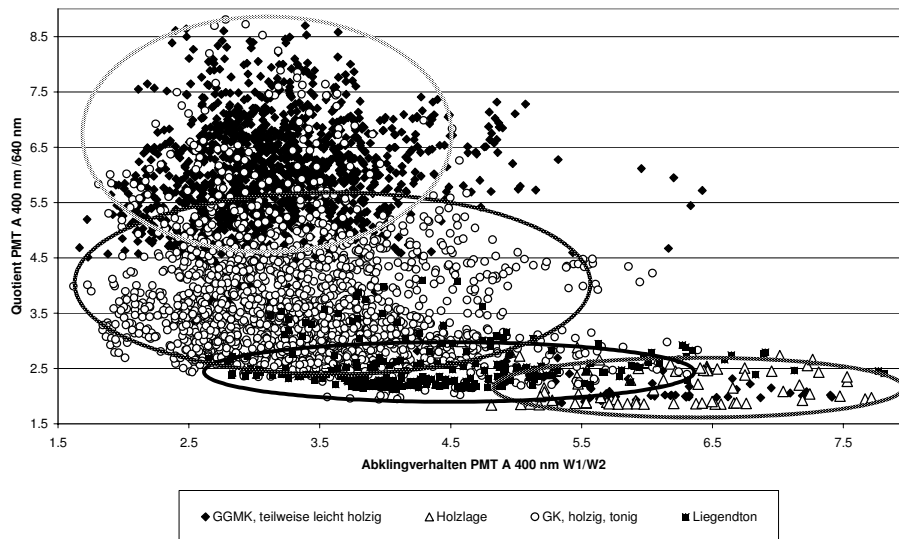
Die Abbildung zeigt auch die grundsätzliche Schwierigkeit des Einsatzes der LIF zur Materialunterscheidung und Trennflächenerkennung in der Braunkohlegewinnung. Die reinen Intensitäten sind auf keinen Fall aussagekräftig, da sich ähnliche Signalhöhen für unterschiedliche Schichtpakete einstellen. Auch beim Rückgriff auf Quotienten bzw. Abklingkurven lassen sich mit der Auswahl nur eines einzigen Merkmals keine eindeutigen Zuordnungen durchführen, wie Abbildung 3-9 verdeutlicht. Die Materialübergänge lassen sich zwar recht scharf abgrenzen, doch wäre z.B. anhand des Quotienten aus PMT A 400 nm/ 600 nm die Holzlage nicht vom Liegendton zu differenzieren. Für beide Schichtpakete sind Werte unter 3 zu verzeichnen.

Abbildung 3-9:
 Verlauf der Quotienten und
 Abklingkurven am Liegend-
 übergang, nördlich Liegen-
 drücken, 12.11.2002



Aus diesem Grunde ist nur eine mehrdimensionale Berücksichtigung von Merkmalen zielführend. Graphisch veranschaulicht Abbildung 3-10 diese Aussage für den Bereich nördlich des Liegendrückens. Mit zwei Merkmalen

ist es möglich, eine grobe Einordnung der Materialien mittels Clustern vorzunehmen. Die Streuungen respektive Überlappungen sind jedoch erheblich, so dass weitere Merkmale herangezogen werden müssen, um eine für den betrieblichen Einsatz befriedigende Klassifikation zu erreichen.



3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-10: Liegendübergang, nördlich Liegendrücken, Clusterbildung, 12.11.2002

Tiefstufe

Über einen Winkel von 18° wurden mit Laser 1 in beiden Scans jeweils etwa 1.500 Messpunkte erfasst. Um die Richtungsunabhängigkeit des LIF Scanner / Analysers zu untersuchen, wurde zunächst dabei von 0° auf 18° von oben nach unten bemustert und anschließend von 18° auf 0°. Die Scangeschwindigkeit betrug in beiden Fällen 0,2 9/s. Bei beiden Vorgängen wurde ohne Blenden bei 1.000 V PMT Spannung auf beiden Kanälen gearbeitet. Eine Richtungsabhängigkeit wurde nicht festgestellt. Nachfolgend zeigt Abbildung 3-11 den Verlauf der Fluoreszenzintensität im Bereich der Tiefstufe.

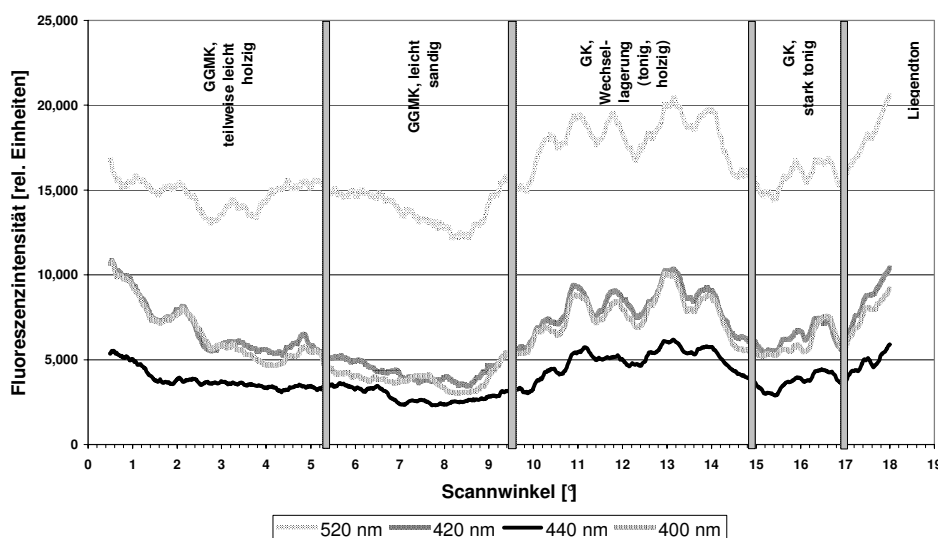


Abbildung 3-11: Verlauf der LIF-Intensitäten am Liegendübergang, Tiefstufe 12.11.2002

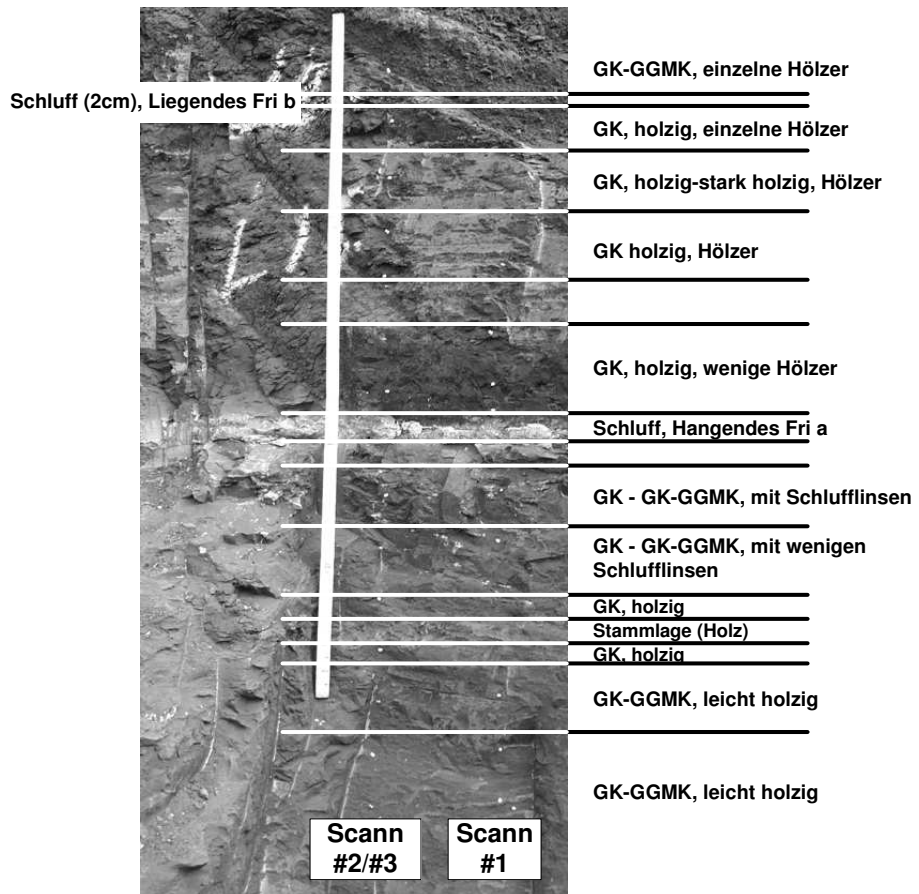
3 Praktische
 Untersuchungen in der
 Braunkohlegewinnung

3.2.2 Messkampagne Zwischenmittel Frimmersdorf
 a/b

Das Zwischenmittel Frimmersdorf a/b ist in dem betrachteten Bereich der Lagerstätte durch eine Zweiteilung gekennzeichnet. Das Liegende des Flözes Frimmersdorf b, ein dünnes Schluffpaket, befindet sich etwa 150 cm oberhalb des Hangenden des Flözes Frimmersdorf a, einer mehrere Dezimeter mächtigen Schlufflage. Unterhalb dieser schließt sich ein Bereich mit stark schluffig verunreinigter GK bzw. GK mit Schlufflinsen an. Es folgen holzige GK sowie eine Stammlage, die im Weiteren sehr deutlich auszuhalten sein wird. Darunter schließen bauwürdige Partien mit GK bis GK-GGMK aus dem Flöz Frimmersdorf a an (Abbildung 3-12).

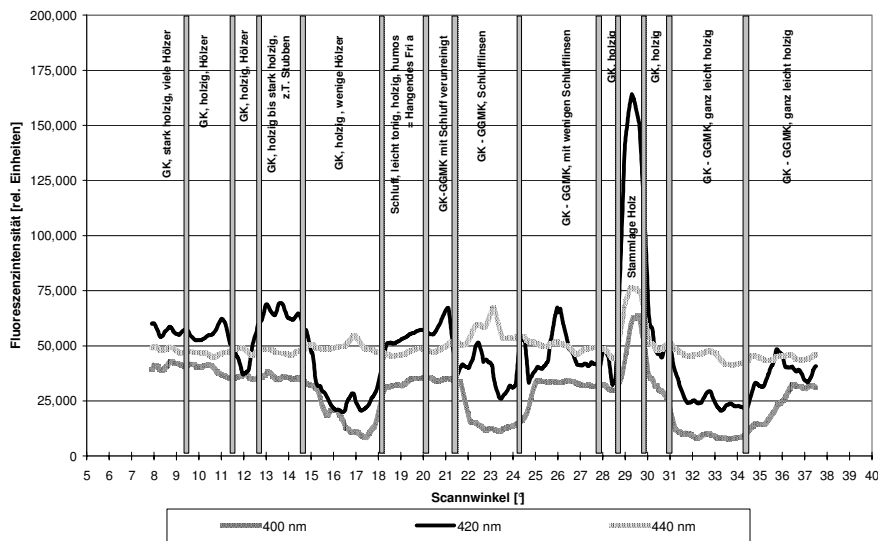
Das freigelegte Profil wurde in drei Scanns mit beiden Lasern auf seine Fluoreszenzcharakteristik hin untersucht. Die Scanngeschwindigkeit lag bei 0,16%/s. Aufgrund schwankender Frequenzen wurden pro Scannvorgang zwischen 825 und 2.700 Messwerte erfasst. Die Gründe dafür sind in Abschnitt 3.2.3 dokumentiert. Insgesamt wurden 68 Scannvorgänge durchgeführt. Beide Photomultiplier (PMT's) arbeiteten bei 1.100 V ohne vorge-schaltete Blenden.

Abbildung 3-12:
 Profil des Scannbereichs
 am Zwischenmittel Frim-
 mersdorf a/b, 09.04.2003



Eine Abgrenzung einzelner Schichten ist mittels der Intensität in ausge-wählten Spektralbereichen durchaus möglich. So kennzeichnet ein deutli-cher Intensitätsabfall bei 400 nm von ca. 35.000 zu ca. 12.000 den Über-

gang zur GK – GGMK mit wenigen Schlufflinsen. Sehr deutlich sticht auch die Holzlage Stammlage bei 28° - 30° mit Intensität n bis zu 150.000 heraus.



3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-13:
Verlauf der LIF-Intensitäten am Zwischenmittel Frimmersdorf a/b 09.04.2003

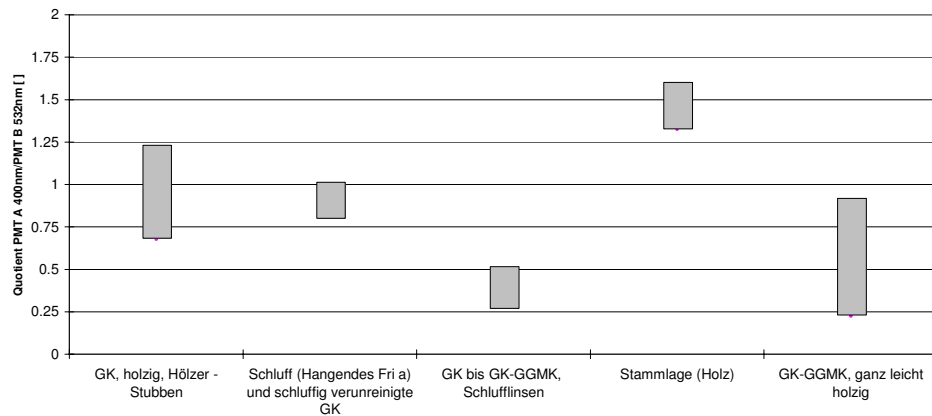
Gleichzeitig lässt sich auch hier wieder das grundlegende Problem der nur auf Intensitäten beruhenden Materialunterscheidung ausmachen: Das Hangende Frimmersdorf a und die Holzige GK zeigen bei 400 nm ungefähr gleich hohe Intensitäten. Würden also in einem anderen Profil diese beiden Schichten direkt aufeinander folgen, wäre eine auf Intensitäten beruhende Unterscheidung schwierig. Auch zeigen gewisse Spektralbereiche wie etwa 440 nm nur sehr geringe Differenzen zwischen materialspezifischen Signalcharakteristika.

Daneben führen die erheblichen Streuungen innerhalb der definierten Materialklassen zu Unzulänglichkeiten bei der Zuordnung mittels nur eines Merkmals. Neben den gerätetechnischen Unzulänglichkeiten, die im Folgenden angesprochen werden, sind dafür die Inhomogenitäten verantwortlich. Vereinzelt Holzstücke in ansonsten reinen Kohlepartien oder partiell tonige Abschnitte beeinflussen die LIF-Messungen erheblich. Dabei handelt es sich um das grundsätzliche Problem, dass gewachsenes Material bemustert wurde, das eben kein einheitliches Material darstellt, auch wenn die Ansprache dies vortäuscht.

Die Streuung der Werte veranschaulicht exemplarisch Abbildung 3-14 für den Quotienten aus den Spektralbereichen des PMT A 400 nm und PMT B 532 nm. Die Holzlage lässt sich eindeutig von den anderen vier Materialien aushalten. Doch eine Unterscheidung zwischen dem nichtbauwürdigen Material mit Schlufflinsen und den bauwürdigen Kohlen im Hangenden des Flözes Frimmersdorf a unterhalb der Holzlage ist aufgrund der erheblichen Streuung nicht einwandfrei möglich.

3 Praktische Untersuchungen in der Braunkohlegewinnung

Abbildung 3-14: Streuung der Merkmale für ausgewählte Materialien, Zwischenmittel Frimmersdorf a/b, 09.04.2003



3.2.3 Diskussion der Fehlerquellen

Beim Einsatz einer neuen Technologie, um die es sich beim LIF Scanner / Analyser ohne Zweifel handelt, sind unerwartete Störungen und Ausfälle unvermeidlich, zumal das Gerät zwar für den robusten Technikumseinsatz, jedoch nur bedingt für den direkten Einsatz im Feld ausgelegt wurde. Aus diesem Grund ist es nicht verwunderlich, dass eine Vielzahl von Unterbrechungen die Feldeinsätze begleitete. Einige dieser Fehler wurden zeitnah korrigiert, so dass die Messkampagne im April 2003 bei der Durchführung der Versuche und dem Versuchsaufbau optimiert werden konnte.

Einige Fehlerquellen sollen hier exemplarisch aufgeführt werden, die zugleich als Vorgriff auf die Bewertung der Tauglichkeit der LIF-Technik als Sensor für die automatisierte Trennflächenerkennung bei der Braunkohlegewinnung zu sehen sind. Insgesamt konnten von den 24 Scans, die während der Messkampagnen im November 2002 und April 2003 vor Ort im Tagebau Hambach durchgeführt wurden, nur 13 für eine sinnvolle Mustererkennung herangezogen werden. Entweder wurde die Abfolge des Scans bereits nach einer statistisch zu geringen Anzahl von Filterradeneinstellungen abgebrochen, oder es lag ein Ausfall nach wenigen Laserimpulsen vor. Konsequenz beider Ausfallursachen war jeweils ein erneutes Booten des LIF Scanner / Analysers, was zu erheblichen zeitlichen Verzögerungen führte. Folgende Fehler werden an dieser Stelle besonders angesprochen:

- Bei den Messungen am Zwischenmittel Frimmersdorf a/b im April 2003 betrug z.B. bei einer eingestellten Mess- bzw. Schussfrequenz der Laser von 16 Hz die Scangeschwindigkeit 0,16 %/s. Aufgrund der extrem niedrigen Temperaturen zu Beginn der Messungen – gegen 10.00 Uhr lagen die Werte deutlich unter 0° Celsius – konnten zunächst nur niedrigere Messfrequenzen realisiert werden. Gegen Mittag wurde die eingestellte Frequenz annähernd erreicht. Allgemein wurde die Versuchsdurchführung durch die niedrigen Temperaturen negativ beeinflusst, da die elektronischen Bauteile des LIF Scanners / Analysers nicht für diese extremen Bereiche ausgelegt sind.

- Schwankungen in der Spannungsversorgung, die durch Generatoren sichergestellt wurde, sorgten dafür, dass sich der LIF Scanner / Analyser kontrolliert terminierte. Begonnene Messungen wurden also abrupt abgebrochen und konnten daher nur bedingt für die Auswertung genutzt werden, abhängig von der Anzahl der durchgeführten Scans und Filtereinstellungen. Dieses Problem konnte mit der Verwendung größerer Aggregate gelöst werden.
- Direkt in die Empfangsoptiken einstrahlendes Sonnenlicht führte im November 2002 zu erheblichen Messstörungen, da anstelle der induzierten Fluoreszenz des Stoßes direkt das Spektrum der Sonne vermessen wurde. Aus diesem Grunde wurde in der zweiten Messkampagne mittels auf die Empfangsoptiken aufgesetzter Tuben sowie dem Einsatz von Sonnensegeln der Einfluss der direkten Sonneneinstrahlung auf die Messungen minimiert.
- Die PMT's reagieren auf eine zu hohe Signalintensität mit kurzfristigem „Erblinden“. Dieses Phänomen liegt im Aufbau und physikalischen Prinzip der Detektoren begründet, so dass man die Ausgangsspannung der maximalen Signalstärke der jeweiligen Fluoreszenzintensität anpassen muss. Zudem kommt es bei einer so genannten Übersteuerung der PMT's zu einem nicht-linearen Zusammenhang zwischen Spannungssignal und tatsächlicher Intensität des Fluoreszenzsignals. Eine Justierung der Empfangseinheit mittels Anpassung der anliegenden PMT-Spannung oder ein Einbringen von Blenden in den Strahlengang ist in diesem Fall erforderlich, um die Übersteuerung zu vermeiden.

Im Falle der Messungen im Tagebau Hambach stellte sich heraus, dass vor allem der Übergang vom Liegenden der Lagerstätte zum Liegendton mit einem erheblichen Intensitätsanstieg einhergeht. Eine manuelle Justierung der Einstellungen auf der Empfangsseite war also bei dem eingesetzten LIF Scanner / Analyser unumgänglich. Eine Vergleichbarkeit zwischen einzelnen Scans ist damit nicht mehr gegeben.

- Bei vielen Messungen wurde eine erhebliche Streuung der Messwerte festgestellt, die in vielen Fällen über ein vertretbares Grundrauschen hinausging. Gründe für diese Schwankungen sind im internen Geräteaufbau, aber auch in externen Quellen zu suchen.

Jeder Laser unterliegt einer internen Puls-zu-Pulsschwankung des Energieinhaltes eines Laserimpulses. Diese variiert nach Herstellerangaben bei einer Schussfrequenz von 16 Hz zwischen 3,5% (355 nm) und 4,5% (266 nm). Doch zeigte sich bei weiterführenden Untersuchungen im Technikum des BGMR, dass diese Energieschwankungen bei der Anregung vor allem bei Laser 1 erheblich höher lagen. Bereits der punktuelle Beschuss eines homogenen Materials ergab je nach Spektralbereich Streuungen von bis zu 25%.

Darüber hinaus stellt die Entfernungskalibrierung eine weitere Fehlerquelle dar, die zur Streuung der Messwerte führt. Die Kalibrierung ist erforderlich, um die sich ändernden Entfernungen zwischen Stoß und Analyser zu normieren, so dass sich vergleichbare Intensitäten ergeben.

3 Praktische Untersuchungen in der Braunkohlegewinnung

Hierzu wurden im Technikum des BGMR umfangreiche Kalibrierungsmessreihen durchgeführt. Die Ursachen und Zusammenhänge, die zu den entfernungsbedingten Streuungen führen - abhängig von der absoluten Signalintensität betragen sie bis zu 10% - sind noch nicht vollständig verstanden, so dass hier ebenfalls ein Aufklärungsbedarf besteht.

Die Wahl der Integrationszeiten barg ebenfalls ein erhebliches Fehlerpotenzial, welches jedoch nach der Änderung der Integrale in der zweiten Feldkampagne nicht mehr zum Tragen kam. Dieser Fehler konnte somit eliminiert werden.

- Gerätetechnische Mängel bei der Auswahl bzw. bei Verarbeitung, Einbau und Anordnung einzelner Komponenten durch den Hersteller des LIF Scanner / Analysers sowie die Tatsache, dass es sich um einen Prototypen handelt, führten zu kompletten Ausfällen des Systems. In einigen Fällen konnten diese erst nach langwierigen Reparaturen behoben werden, was zum Abbruch der Messkampagnen führte. Dies war bei den ersten Versuchen im Tagebau Hambach im Oktober 2002 der Fall, als es zu einem Ausfall des Schrittmotors des Scannkopfes sowie zu einer Fehlfunktion der Motoren beim Wechseln der Bandpassfilter- bzw. Blendenräder kam. Die Messreihe im April 2003 wurde erheblich gestört durch einen Fabrikationsfehler der A/D-Wandlerkarte, die für die Umwandlung der analogen Spannungssignale der PMT's in digitale Signale verantwortlich ist, bzw. war die mangelhafte Herstellung verantwortlich für den Abbruch der Messungen und einen Ausfall des LIF Scanner / Analysers für mehr als einen Monat. Es wird also ersichtlich, dass es derzeit noch erheblicher Optimierungsanstrengungen an der verwendeten Technik bedarf, bevor diese in einem automatisierten Sensorsystem genutzt werden kann.

3.2.4 Fazit der Feldversuche

Die Messkampagnen im Tagebau Hambach der RWE Power AG zeigten, dass eine grundsätzliche Unterscheidung von Materialien mit der laserinduzierten Fluoreszenz unter realen Einsatzbedingungen möglich ist. Technische Mängel am LIF Scanner / Analyser sowie äußere Umstände wie schlechte Witterung oder schwankende Stromversorgung erschwerten die Durchführung der Messungen, doch zeigt die Auswertung der Daten, dass eine Klassifikation ausgehend von den vorhandenen Ergebnissen bereits möglich ist [BAY, 2003; FUC, 2003].

Die erfolgreiche Klassifikation und die Übertragung der LIF-Technologie in den Gewinnungsbetrieb ist im Wesentlichen von zwei Faktoren abhängig: Eine eindeutige Differenzierung und Zuordnung der Materialien und die Erreichung dieses Ziels mit Hilfe einer minimalen Anzahl von Merkmalen. Der letztere Punkt begründet sich in den Restriktionen an die baulichen Dimensionen eines Analysers, der mit PMT-Technologie arbeitet, und den Kosten für ein solches Gerät.

Die vorhandenen Daten wurden, wie gezeigt, bereits teilweise vor dem Hintergrund der Merkmalsreduktion bearbeitet und analysiert. Im Rahmen von Voruntersuchungen zu dieser Arbeit wurde versucht, mittels materialspezifischer Mittelwerte unter Berücksichtigung der Standardabweichung geeignete Merkmale bei gleichzeitiger Maximierung des Klassifikationserfolges zu selektieren. So liegen als Ergebnis einer Diplomarbeit eine Access-Datenbank sowie ein Algorithmus vor, der eine Selektion der Merkmale für eine definierte Anzahl von Materialklassen vornimmt [FUC, 2003].

Die Voruntersuchungen zeigen ein erhebliches Verbesserungspotential hinsichtlich des Klassifikationserfolges auf. Dieses soll im Rahmen der vorliegenden Arbeit mit der Verwendung von alternativen Selektionsansätzen und Klassifikationsverfahren zur Trennflächen- und Grenzschichterkennung nutzbar gemacht werden.

4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

In diesem Kapitel soll zunächst kurz auf die theoretischen Grundlagen der Muster- und damit auch der Materialerkennung eingegangen werden, bevor die für diese Arbeit entwickelten und genutzten Algorithmen im Detail vorgestellt werden. Bei den verwendeten Algorithmen handelt es sich um in anderen Bereichen der Mustererkennung etablierte Verfahren, die jedoch für die spezielle Applikation in der mineralischen Rohstofferkennung modifiziert und für die neue Problemstellung in der Grenzschichterkennung weiterentwickelt wurden. Dies ermöglicht erstmals eine systematische Analyse und Selektion von LIF-Messwerten sowie erweiterte Materialerkennung.

Die Darstellung der programmiertechnischen Umsetzung der gewählten Verfahren erfolgt in der Form von Flussdiagrammen (Flow Charts). Der Code, der in dem mathematischen Programm „MATLAB“, Version 6.5, umgesetzt wurde, ist im Anhang C aufgeführt.

4.1 Grundlagen der Mustererkennung

Bei der Mustererkennung handelt es sich allgemein gesprochen um den Prozess der Identifikation von Objekten anhand bereits gelernter Beispiele. Nach Niemann „beschäftigt sich die Mustererkennung mit den mathematisch-technischen Aspekten der automatischen Verarbeitung und Auswertung von Mustern. Dazu gehören sowohl die Klassifikation einfacher Muster als auch die Analyse komplexer Muster [NIE, 1983].“

Ein Muster ist in diesem Sinne eine Anzahl von Merkmalen (Features), d.h. eine Beschreibung bzw. Eigenschaft eines Objektes: Dies kann eine Textur oder Struktur, aber auch ein Verhaltensmuster oder eine elektromagnetische Welle wie etwa ein Ton oder Licht sein. Klassen fassen viele unterschiedliche Objekte unter gemeinsamen Eigenschaften zusammen. Dies kann eine ausgewählte Menge an Eigenschaften, oder aber die Gesamtheit an Eigenschaften sein. Jedes Objekt wird durch einen mehrdimensionalen Merkmalsvektor, auch als Merkmalsset bezeichnet, beschrieben, der alle zur Beschreibung erforderlichen Merkmale enthält.

Die Identifikation beschreibt den Akt der Zuordnung eines Objektes zu einer Klasse von Objekten, zu denen es gehört. Zu unterscheiden ist in diesem Kontext zwischen Erkennen und Identifizieren, da Erkennen nicht zwingend mit einer Zuordnung zu einer Klasse einhergeht. In der Literatur wird daher im deutschsprachigen Raum auch von Musteridentifikation gesprochen – im Gegensatz zum angloamerikanischen „Pattern recognition“.

Die Klassifikation oder auch Klassifizierung ist nun der Vorgang der Sortierung von Objekten unterschiedlicher Klassen in Gruppen, so dass jede Gruppe nur noch Objekte einer Klasse enthält. Exemplarisch wäre hier für

die Taxonomie in der Biologie z.B. die Klassifikation von Laubbäumen anhand von Blattform und -farbe zu nennen.

Systeme zur Mustererkennung lassen sich grundsätzlich in drei Hauptkomponenten zerlegen. Die Vorverarbeitung der Daten (Pre-processing), die Merkmalsselektion (Feature extraction) und die eigentliche Klassifizierung (Classification) [DUD, 2000]. In Abbildung 4-1 wird der systematische Aufbau einer Mustererkennung detaillierter dargestellt.

Im Einzelnen soll an dieser Stelle kurz auf die Anforderungen und speziellen Problemstellungen der Unterpunkte eingegangen werden, die bei der praktischen Durchführung einer Mustererkennung relevant sein können.

Die Datenerfassung wird durch limitierende Faktoren und individuelle Charakteristika des Messsystems stark beeinflusst. Im Fall der LIF-Messungen mittels LIF Scanner / Analyser ist dies z.B. die durch die PMT-Filterräder vorgenommene Auswahl an Spektralbändern. Auch die Segmentierung und Auflösung des Messpunktes bzw. des Objektes sind bei der Datenerfassung von Bedeutung. So findet beim LIF Scanner z.B. eine Mittelwertbildung der Fluoreszenzcharakteristika über den Durchmesser des anregenden Laserpunktes statt – die Frage der Auflösung bzw. Vereinzelung der Messung ist hier relevant.

4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

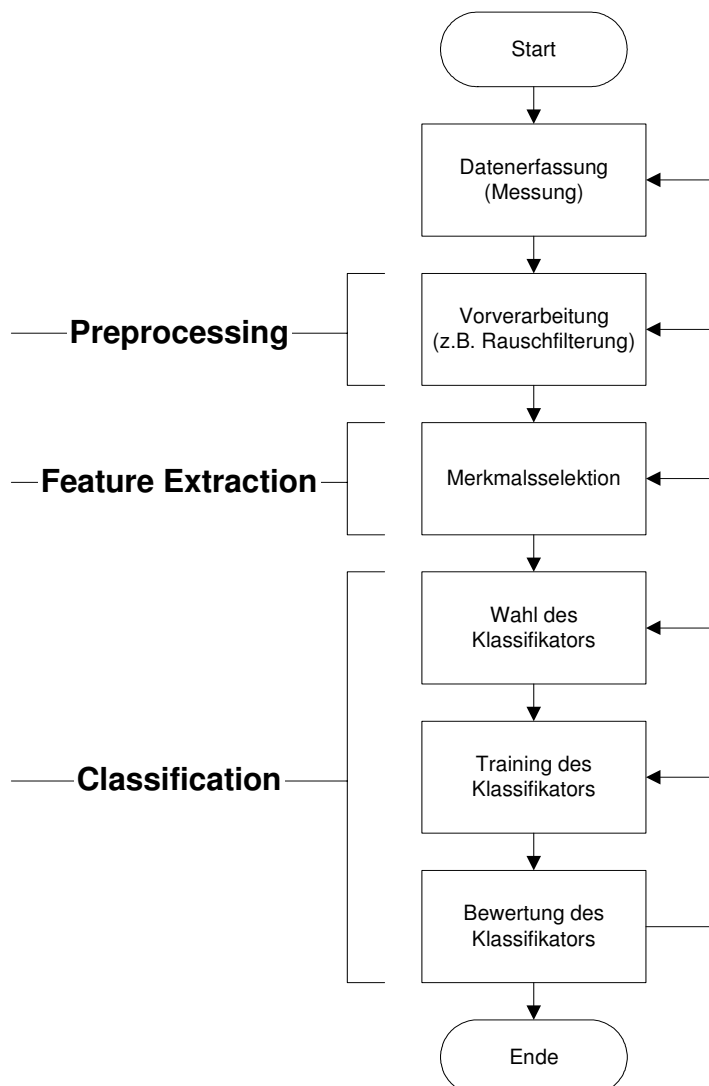


Abbildung 4-1:
Aufbau eines Systems zur
Mustererkennung
[nach DUD, 2000]

4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

Die Vorverarbeitung dient neben der Datenbereinigung auch der Datenintegration. Zunächst bedarf es in der Regel einer Bearbeitung von Fehlwerten und der Bestimmung des Rauschens. Auch die Detektion von Ausreißern und der Umgang mit inkonsistenten Messdaten fallen in diesen Bereich. Bei der Datenintegration steht das Erkennen und das Eliminieren von redundanten Datensätzen und eine eventuelle Normierung der Daten im Vordergrund.

Im Falle der LIF-Basisdaten aus den Feldversuchen im Tagebau Hambach wurde hier auf das Preprocessing - Konzept der vorhandenen Access-Datenbank zurückgegriffen [FUC, 2003]. Neben der Möglichkeit, negative Werte und Nullwerte zu eliminieren, können über eine Blendenfunktion mit wählbaren Ober- und Untergrenzen Ausreißer detektiert werden. Zusätzlich kann die Blendenöffnung, also die Anzahl der betrachteten Werte, reguliert werden. Das Ersetzen der als Ausreißer bzw. Fehlmessungen identifizierten Werte durch gemittelte Werte der benachbarten, korrekten Datensätze erfolgt bei der als Datenmodulation bezeichneten Funktion ebenfalls auf Wunsch. Redundanzen können dann auftreten, wenn der Laser aufgrund schwankender Schussfrequenzen für einen Winkel (die Ermittlung erfolgt auf 0.01° genau) mehrere Messwerte erzeugt. In diesem Falle kommt es ebenfalls zu einer Mittlung und Reduktion auf einen Messwert pro Winkelabschnitt.

Merkmalsselektion und eigentliche Klassifikation sind voneinander konzeptionell teilweise schwer zu unterscheiden, in einigen Fällen geschieht beides auch in engem Zusammenspiel. Ein optimaler Ansatz zur Merkmalsselektion würde bereits eine Zuordnung ermöglichen, die den Klassifikator überflüssig erscheinen lassen würde. Umgekehrt würde ein optimales Klassifikationsverfahren keine Merkmalsselektion mehr erfordern, da es mit allen vorhandenen Merkmalen eine eindeutige Klassifikation vornähme.

Aufgabe der Merkmalsselektion, die in Abschnitt 7 ausführlich diskutiert wird, ist es, die minimale Anzahl an Merkmalen zu finden, die eine Unterscheidung zwischen den Klassen erlaubt. Die Eigenschaften von Objekten derselben Klasse sollten sehr ähnlich sein und gleichzeitig eine eindeutige Differenzierung zu Objekten anderer Klassen erlauben.

Gründe für die Merkmalsreduktion sind zum einen, dass die Messwertaufassung vereinfacht werden kann, da zur einwandfreien Klassifikation nicht alle ursprünglichen Merkmale, sondern nur die selektierten hinreichend sind. Des Weiteren sinkt die Komplexität des Systems mit einer Abnahme der Merkmalsdimension, was mit einer Reduktion des Rechenaufwandes einhergeht. Generell zeigt sich, dass mit einer geringeren Anzahl an Merkmalen eine geringere Fehlerrate bei der Klassifikation mit gleichzeitiger Zunahme der Zuordnungsgeschwindigkeit erreicht werden kann.

Die Klassifikation bzw. die Wahl des geeigneten Klassifikationsverfahrens hat zur Aufgabe, anhand der selektierten Merkmale ein Objekt einer Klasse zuzuweisen. Dabei besteht im Wesentlichen die Schwierigkeit der fehlerfreien Zuordnung, da durch das Rauschen innerhalb von Klassen bzw. zu

geringer Unterschiede zwischen einzelnen Klassen eine eindeutige Zuweisung erschwert wird. Im folgenden Abschnitt werden unterschiedliche Ansätze der Klassifikation vorgestellt, die in parametrische und nicht-parametrische Verfahren unterschieden werden.

Für den Fall, dass die Klasseneinteilung bzw. die Klassenzugehörigkeit von Messwerten bereits vor der Klassifikation bekannt ist, spricht man von einer überwachten Klassifikation. Dies bedeutet, dass das Mustererkennungssystem mit bekannten Datensätzen auf ein optimales Ergebnis hin trainiert werden kann. Dabei spielt für den Einsatz der Mustererkennung später eine erhebliche Rolle, wie repräsentativ die Trainingsdaten für die Grundgesamtheit sind.

4.2 Klassifikationsverfahren

Die generelle Aufgabe aller Klassifikationsverfahren ist die Überführung des Merkmalsvektors einer Beobachtung in eine Klassenbezeichnung, also die Zuordnung eines Messergebnisses zu einem bestimmten Material. Dabei muss die Entscheidungsfunktion, nach der der Klassifikator die Klassifikation durchführt, anhand geeigneter Datensätze trainiert werden, um zu lernen, wie sie zu entscheiden hat.

Bei den Lernmethoden kann man, wie bereits angesprochen, zwischen überwachten und unüberwachten Verfahren unterscheiden. Bei letzterem Vorgehen findet das System selber geeignete Klassen aus dem vorhandenen Datenmaterial – es führt ein Clustering im Merkmalsraum durch. Im Falle der Material- und Trennschichterkennung in der mineralischen Rohstoffgewinnung wird jedoch die Methode des überwachten Lernens angewandt, da hier eine Korrektur von außen erfolgen kann, denn die Klasse jedes Trainingsobjektes ist durch die geologische Ansprache am Stoß bekannt. Den generellen Aufbau eines Klassifikators zeigt die nachfolgende Abbildung 4-2.

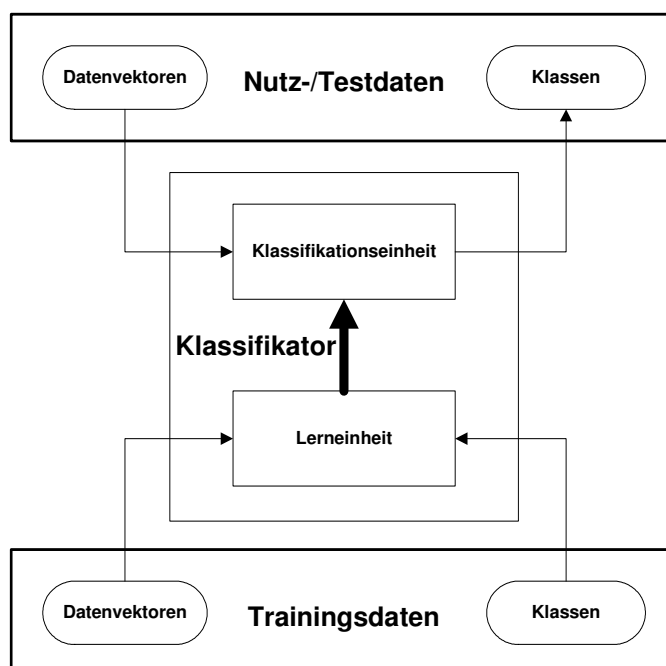


Abbildung 4-2:
Aufbau eines Klassifikators

Bevor im Einzelnen auf die hier betrachteten drei Ansätze zur Klassifikation eingegangen wird, muss zunächst definiert werden, wie sich der Erfolg einer Mustererkennung bewerten lässt. Die Qualitätsbeurteilung erfolgt mittels des Fehlerkriteriums, d.h. mittels des Verhältnisses zwischen den falsch klassifizierten Objekten und der Gesamtmenge der klassifizierten Objekte.

In dieser Arbeit sollen drei generelle Ansätze zur Klassifikation hinsichtlich ihrer Eignung für die Materialerkennung mittels der laserinduzierten Fluoreszenz gewonnenen Messdaten verglichen werden. Der Fokus wurde auf die Gauß'schen Verfahren, die Methodik der nächsten Nachbarn sowie einen Algorithmus, der sich der linearen Regression bedient, gelegt. Diese in der Literatur bekannten Verfahren mussten entsprechend der Aufgabenstellung und vor allem vor dem Hintergrund der unterschiedlichen Ansätze zur Merkmalsselektion überarbeitet bzw. neu entwickelt werden [DUD, 2000; KEY, 2002, KEY 2004].

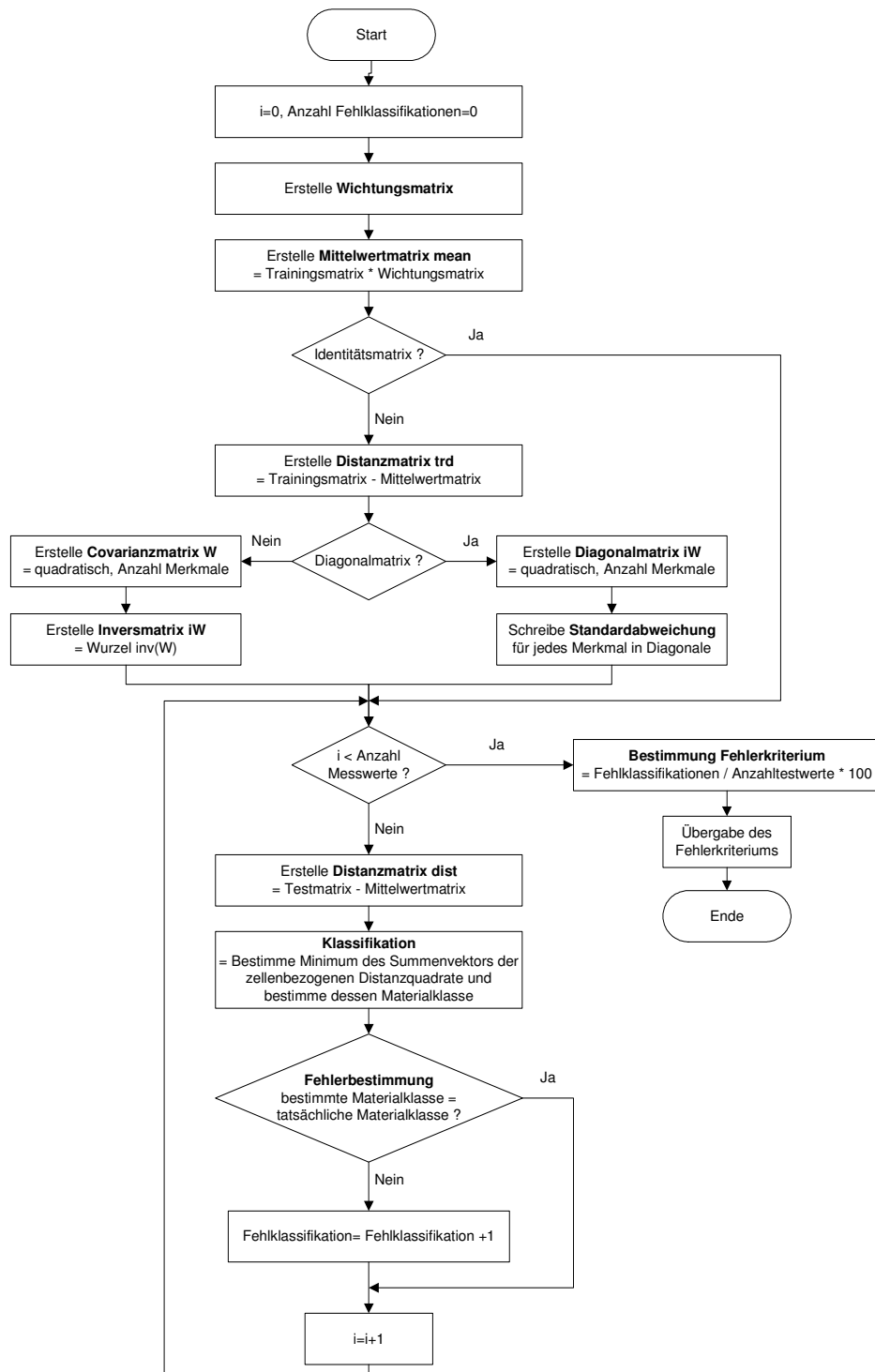
4.2.1 Gaußverfahren

Die Klassifikation nach dem Gaußverfahren arbeitet mit Mittelwerten, Varianzen und Kovarianzen. Es zählt damit zu den parametrischen Verfahren. Die Umsetzung erfolgt dergestalt, dass entweder ausschließlich mit den Mittelwerten der einzelnen Klassen die Klassifikation durchgeführt wird, oder aber mit den Varianzen der Mittelwerte, also der Streuung, gearbeitet wird. Darüber hinaus ist es möglich, auch die Beziehungen der Merkmale zueinander zu berücksichtigen, wenn nämlich die Kovarianzen mit einbezogen werden. Abbildung 4-3 zeigt den Ablauf der Klassifikation nach dem Gaußverfahren.

Zunächst wird die Mittelwertmatrix aus den Trainingsdaten erstellt, d.h. der Mittelwert jedes Merkmals für die unterschiedlichen Klassen errechnet. Entsprechend der Selektion (reine Mittelwerte, Varianzen oder Kovarianzen) wird im nächsten Schritt die Distanzmatrix für jede Messung der Testdatenmenge bestimmt. Diese Matrix besteht aus den Differenzen sämtlicher Merkmale der i -ten Testmessung zu den Werten der Mittelwertmatrix. Bildlich wird die Entfernung des einzelnen Wertes zu den Mittelwerten aller Trainingsklassen errechnet, um die Übereinstimmung oder Ähnlichkeit zu den jeweiligen Klassen zu bestimmen. Diese Distanzmatrix wird nun elementweise quadriert und klassenweise aufsummiert. Der sich ergebende Zeilenvektor stellt nun die Ähnlichkeit zwischen dem Testwert und den Klassen der Trainingsdatenmenge dar. Die Klassifikation des unbekanntes Testwertes erfolgt durch die Zuordnung in die Klasse, die die geringste Entfernung zum Testwert aufweist.

Der Erfolg der Klassifikation wird direkt durch den Vergleich der zugewiesenen mit der tatsächlichen Klasse überprüft. Das Fehlerkriterium für die gesamte Testdatenmenge kann bestimmt werden, indem der beschriebene Ablauf für alle i -Messdaten vollzogen wird.

Klassifikationsverfahren zur Materialerkennung



4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

Abbildung 4-3:
Flowchart des Ablaufs der
Klassifikation nach dem
Gaußverfahren

Die Schwierigkeit, die beim Einsatz des parametrischen Verfahrens auftritt, betrifft die Annahme der Normalverteilung der Merkmale. Ist dies jedoch nicht der Fall, sind komplizierte Berechnungen erforderlich und eine nachträgliche Modifikation des Algorithmus bzw. der Datensätze schwer durchführbar.

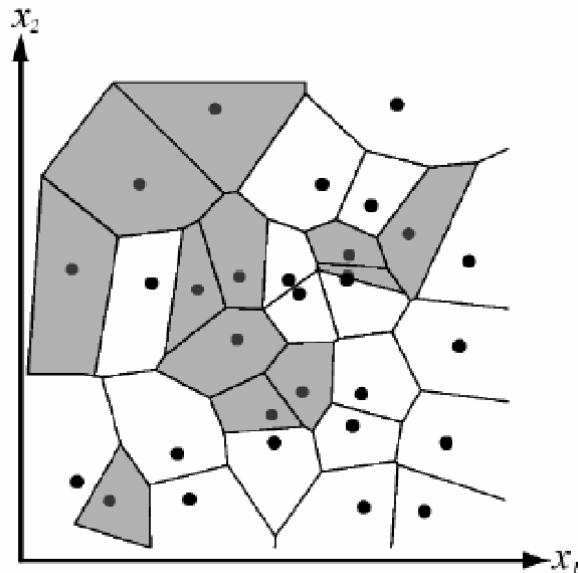
4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

4.2.2 k-nächster Nachbar Verfahren (k-NN)

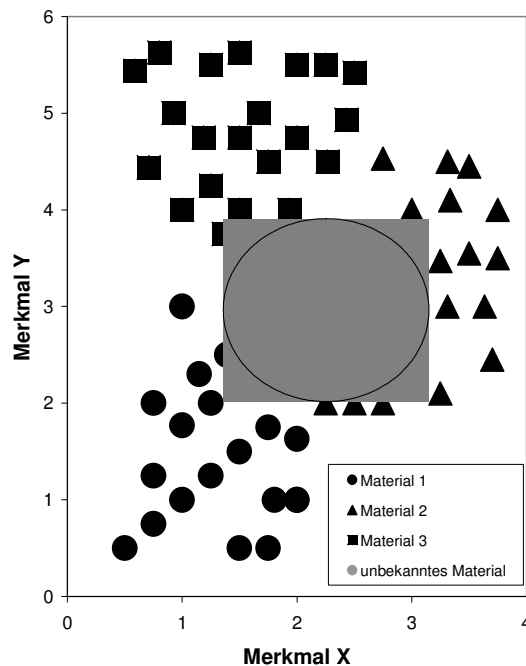
Der k-nächster Nachbar-Algorithmus zählt zu den nicht-parametrischen Klassifikationsverfahren und verzichtet auf die Nutzung von multivariaten Normalverteilungen. Vielmehr wird direkt aus den Trainingsdaten die Abschätzung der Zugehörigkeit eines Objektes zu einer Klasse vorgenommen. Der k-NN-Klassifikator ist in der Trainingsphase sehr schnell, da alle Trainingsdaten ausgewählt und gespeichert werden. Die Klassifikation ist jedoch mit erheblichem Rechenaufwand verbunden, da ein Vergleich des zu klassifizierenden Objektes mit allen Prototypen der Trainingsmenge durchgeführt wird. Dies bedeutet, dass z.B. beim 1-NN das unbekannte Objekt der Klasse des Trainingsprototypen zugewiesen wird, zu dem es den geringsten Abstand im Merkmalsraum aufweist. Zumeist wird hier der euklidische Abstand errechnet, aber auch andere Distanzmaße wie etwa der City-Block-Ansatz können verwendet werden, wenn dadurch Besonderheiten des Merkmalsraumes optimaler ausgenutzt werden können.

Anschaulich zeigt Abbildung 4-4 für einen zweidimensionalen Merkmalsraum die Vorgehensweise in der Trainingsphase. Der Merkmalsraum wird geometrisch durch so genannte Voronoi-Polygone aufgeteilt. Das einzelne Polygon ist durch die Mittelsenkrechten der Verbindungsstrecken zwischen einem Punkt und seinen Nachbarpunkten definiert. Die Polygone am Rand sind offen, da sie keine Nachbarpunkte in dieser Richtung haben.

Abbildung 4-4:
Voronoi-Polygone für 1-NN-
Verfahren [BEI; 2003]



Die Klassifikation eines unbekanntes Objektes erfolgt beim 1-NN nun in die Klasse des Trainingswerts, in dessen Polygon das zu klassifizierende Objekt fällt. Im Falle des k-NN wird nicht nur die Klasse des nächsten Trainingsprototypen in den Klassifikationsprozess eingezogen, sondern auch die k benachbarten Trainingsdaten. Abbildung 4-5 zeigt die Klassifikation für einen 10-NN-Klassifikator bei drei unterschiedlichen Materialklassen in einem zweidimensionalen Merkmalsraum.



4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

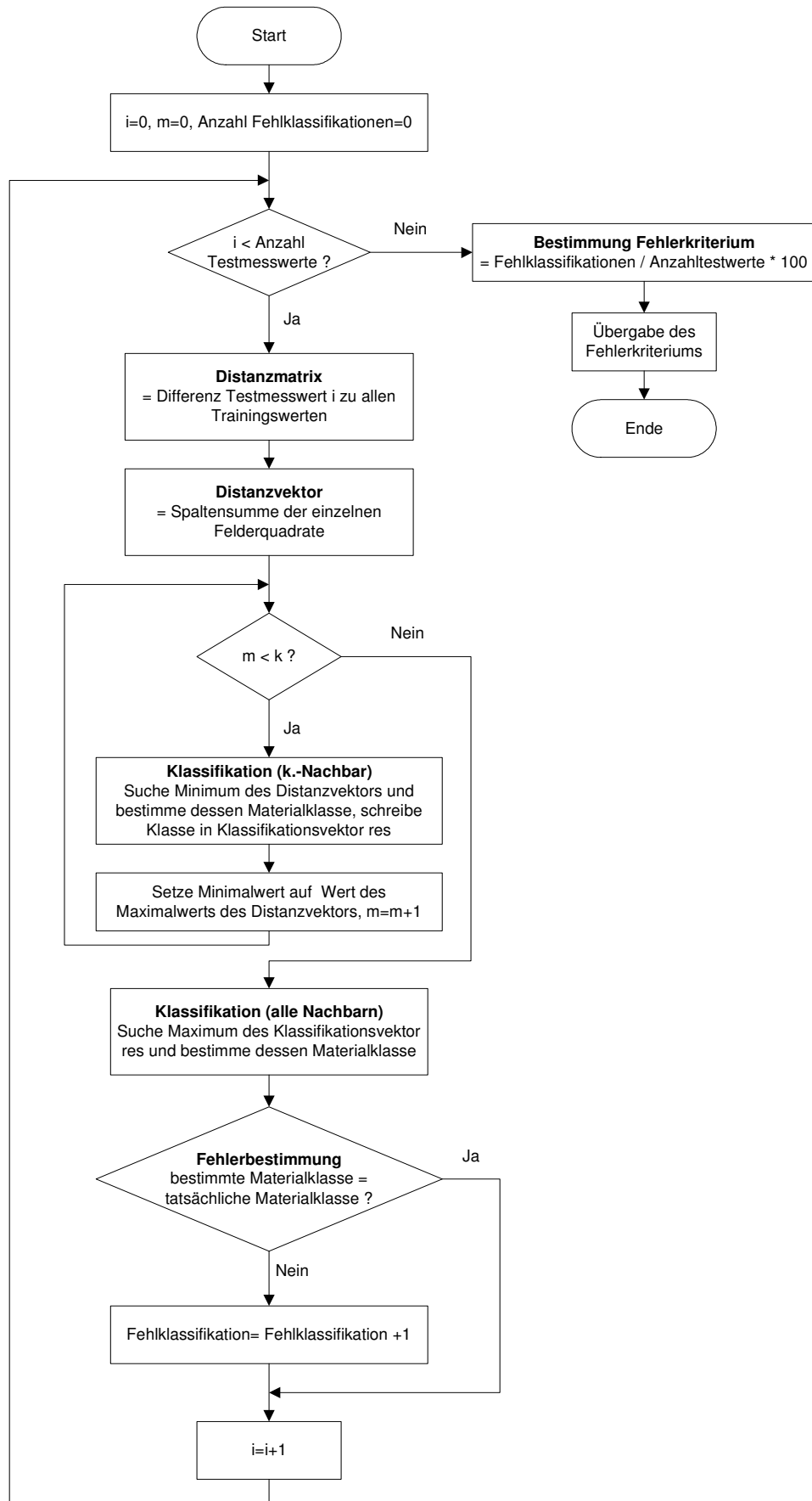
Abbildung 4-5:
Klassifikation bei einem 10-
NN-Verfahren

Das unbekanntes Messobjekt würde in dem dargestellten Fall der Klasse 2 zugeordnet werden, da diese mit sechs Nachbarn das Maximum darstellt.

Abbildung 4-6 zeigt den Ablauf der Klassifikation nach dem k-NN-Verfahren.

**4 Mustererkennung:
Grundlagen,
Algorithmen und
Programmierung**

Abbildung 4-6:
Flowchart des Ablaufs der
Klassifikation nach dem k-
Nächster Nachbar



Für jeden Messwert i wird zunächst die Differenz der erfassten Merkmale zu allen Trainingsprototypen in einer Distanzmatrix ermittelt und dann

durch Aufsummierung der elementweisen Quadrate ein Distanzvektor gebildet. Dieser stellt die Entfernung des unbekanntes Objektes zu jedem einzelnen Trainingsobjekt im Merkmalsraum dar. Nun werden entsprechend der Vorgabe von k die benachbarten Objekte ermittelt und deren Klasse in einen Klassifikationsvektor geschrieben. Dabei wird zunächst das jeweilige Minimum des Distanzvektors ermittelt und die Klassenzugehörigkeit dieses Objektes gespeichert. Anschließend wird der Distanzwert dieses Objektes auf das Maximum des Distanzvektors gesetzt und in einer weiteren Iteration die Klasse des nunmehr am dichtesten zum zu klassifizierenden Testwert gelegenen Trainingsobjektes ermittelt.

Um nun die Klasse des Testobjektes zu bestimmen, erfolgt die Auswertung des Klassifikationsvektors, indem das Maximum der Nennungen bestimmt wird. Dieses Vorgehen wird für jedes Messobjekt der Testdatenmenge wiederholt und dann entsprechend des Zuordnungserfolges das Fehlerkriterium ermittelt.

Um die Rechenzeit und den Speicherbedarf zu optimieren, kann man Abwandlungen des Algorithmus nutzen, die die Komplexität reduzieren. Eines dieser Verfahren ist das Entfernen von redundanten Datenpunkten, die keinerlei Informationszugewinn für die Klassifikation erbringen. In diesem als „Editing“ oder auch „Condensing“ bekannten Verfahren werden nur die Merkmale der Trainingsmenge genutzt, die sich entlang von Materialgrenzen im Merkmalsraum befinden. Diese Optimierung wurde in der Arbeit nicht weiterverfolgt. Grund dafür ist die erhebliche Erhöhung der Rechenzeit bei der Merkmalsselektion, da das Condensing für jeden Schritt der Selektion jeweils neu durchgeführt werden muss. Im Falle einer späteren Anwendung würde man jedoch sehr wohl auf diese Optimierungsmöglichkeit zurückgreifen. In diesem Falle wäre dann das relevante Merkmalsset des Klassifikators bekannt, so dass eine Modifikation der Trainingsdaten basierend auf diesen Merkmalen erfolgen und eine schnellere Klassifikation ermöglicht werden würde.

4.2.3 Lineares Regressionsverfahren

Lineare Klassifikationsverfahren trennen den Merkmalsraum je Merkmalsachse höchstens an einem Punkt, dies bedeutet, dass sich die Trennfunktionen als Geraden darstellen lassen. Im Gegensatz zum k -NN-Verfahren, das z.B. einen zweidimensionalen Merkmalsraum in mehr als zwei Klassen unterteilen kann, lässt sich dies bei linearen Klassifikatoren nicht bewerkstelligen.

Grundsätzlich reduziert sich die Aufgabenstellung bei allen linearen Regressionsverfahren auf die Lösung eines linearen Gleichungssystems. Die Darstellung des Systems erfolgt mittels Matrizen. Das Problem stellt sich dann als $Ya = b$ dar. Die rechteckige Matrix Y stellt in diesem Kontext die Trainingsmenge und b den Begrenzungsvektor dar, während a der Wichtigkeitsvektor ist, der für die Klassifikation entscheidend ist. Die Lösung basiert auf der Ermittlung dieses Vektors unter der Prämisse, den Fehler der

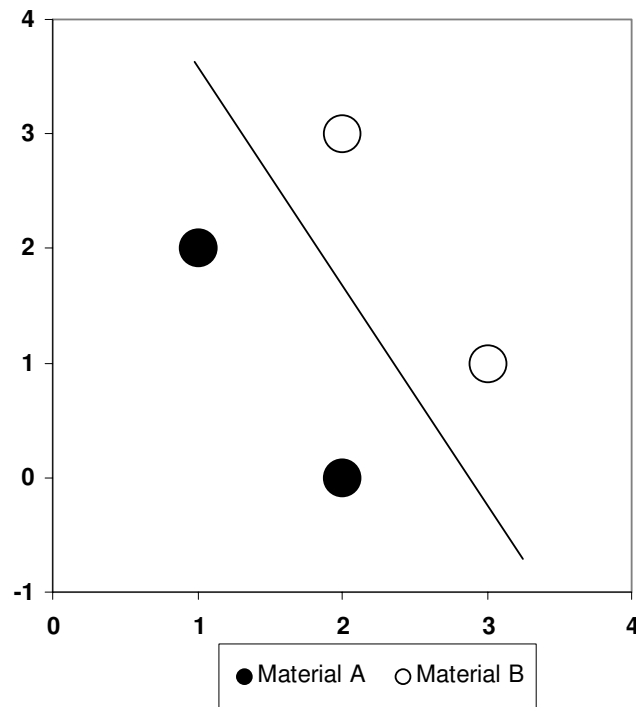
**4 Mustererkennung:
Grundlagen,
Algorithmen und
Programmierung**

Trennfunktionen zu minimieren. Dies lässt sich mit Hilfe der Pseudoinversmatrix Y^\dagger erreichen, die sich aus $(Y^t Y)^{-1} Y^t$ errechnet. Die Lösung des Problems mit dem Wichtungsvektor a ergibt sich dann aus:

$$a = (Y^t Y)^{-1} Y^t b = Y^\dagger b .$$

Das folgende Beispiel (Abbildung 4-7) soll für einen zweidimensionalen Raum das Vorgehen für ein lineares Klassifikationsverfahren erläutern [DUD, 2000]. Die beiden Materialien A und B werden durch ω_1 und ω_2 bezeichnet.

Abbildung 4-7:
Beispiel für ein lineares
Klassifikationsverfahren



Für die Trainingsdatenmatrix Y und die Pseudoinversmatrix Y^\dagger gilt dann:

$$Y = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 0 \\ -1 & -3 & -1 \\ -1 & -2 & -3 \end{pmatrix} \quad Y^\dagger = (Y^t Y)^{-1} Y^t = \begin{pmatrix} 4/5 & 13/12 & 3/4 & 7/12 \\ -1/2 & -1/6 & -1/2 & -1/6 \\ 0 & -1/3 & 0 & -1/3 \end{pmatrix} .$$

Bei der Verwendung einer gleich verteilten Begrenzungsmatrix $b=(1,1,1,1)^t$ ergibt sich die Lösung des Problems in Form der Trennfunktion zum Wichtungsvektor $a=(11/3,-4/3,-2/3)^t$.

Das generalisierte Problem mit mehr als nur zwei Klassen lässt sich auf den Grundfall zurückführen, in dem man es als eine Reihe von mehreren Zweiklassenproblemen auffasst. Es wird in diesem Fall der Wichtungsvektor a_i gesucht, der die Lösung des folgenden Problems darstellt:

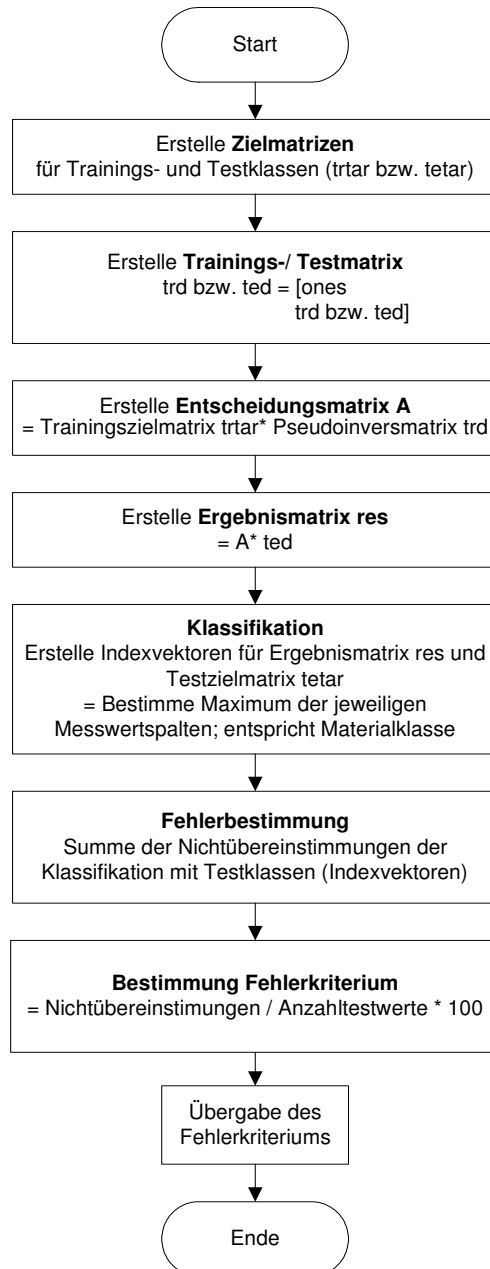
$$a_i^t y = 1 \text{ für alle } y \in Y_i$$

$$a_i^t y = 0 \text{ für alle } y \notin Y_i$$

Entsprechend erfolgt die Entscheidung über die Klassifikation dahingehend, dass das Testobjekt y der Klasse ω_i zugeordnet wird, wenn gilt:

$$a_i^t y \geq a_j^t y \text{ für } j \neq i.$$

Abbildung 4-8 zeigt den Ablauf der Klassifikation nach dem linearen Regressionsverfahren, der die oben beschriebene Systematik umsetzt.



Die Entscheidungsmatrix A wird aus der Zielmatrix, die der Begrenzungs-
matrix b entspricht, und der Pseudoinversmatrix der Trainingsdaten ermit-
telt. Anschließend wird die Ergebnismatrix aus der Entscheidungsmatrix
und den Testdaten gewonnen. Die Klassifikation erfolgt nun derart, dass

4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

Abbildung 4-8:
Flowchart des Ablaufs der
Klassifikation nach dem
Verfahren der linearen Reg-
ression

jedes Testobjekt i der Klasse zugewiesen wird, in deren Zeile das Spaltenmaximum des Objektes i der Ergebnismatrix vorliegt. Anschließend wird durch den Abgleich der Übereinstimmungen zwischen der Objektklassifikation und den tatsächlichen Testklassen das Fehlerkriterium ermittelt.

4.3 Verfahren zur Merkmalsselektion

In diesem Abschnitt sollen unterschiedliche Ansätze zur automatisierten Selektion und Reduktion von Merkmalen aus einem gegebenen Merkmalsraum vorgestellt werden. Ziel ist die Reduktion der Dimensionalität des Merkmalsvektors dergestalt, dass eine Unterscheidung zwischen den Klassen beibehalten wird, ohne dass dies mit einem Verlust an Information einhergeht.

Die Generierung der optimalen Merkmale kann über reine Kombinatorik erfolgen, doch stößt diese Methode sehr schnell an Grenzen. Bei einer Originalmerkmalsdimension von D und der gewünschten Größe d des zur Klassifikation eingesetzten Merkmalsvektors ergeben sich bei $D=20$ und $d=5$ aus der nachfolgenden Formel so bereits 15.504 Kombinationen, die überprüft werden müssen.

$$\text{Kombinationen}(D,d) = \frac{D!}{d!(D-d)!}$$

Bei den LIF-Messungen im Tagebau Hambach ergeben sich aus den Quotienten und Abklingkurven bei einigen Scans bis zu 405 Merkmale. Bereits bei einer Messung von nur acht Spektralbanden auf beiden PMT's ergibt sich eine Gesamtanzahl von 152 Merkmalen. Eine Reduktion auf fünf Merkmale geht in diesem Falle mit der Überprüfung von 632.671.880 möglichen Kombinationen einher.

Heuristiken liefern nicht zwangsweise die optimale Methode, stellen aber eine im Allgemeinen akzeptable Möglichkeit dar, zwischen den beiden Seiten - Rechenaufwand und zufrieden stellende Erfolgsquote der Klassifikation basierend auf dem selektierten Merkmalsvektor - abzuwägen. In diesem Abschnitt werden zwei grundsätzliche Ansätze solcher „nicht-optimaler“ sequenzieller Verfahren sowie eine optimierte Erweiterung vorgestellt [XIO, 2001]. Dabei handelt es sich im Einzelnen um die „Sequential Forward Selection (SFS)“ und die „Sequential Backward Selection (SBS)“ sowie beim optimierten Verfahren um den Ansatz der „Sequential Forward Floating Selection (SFFS)“.

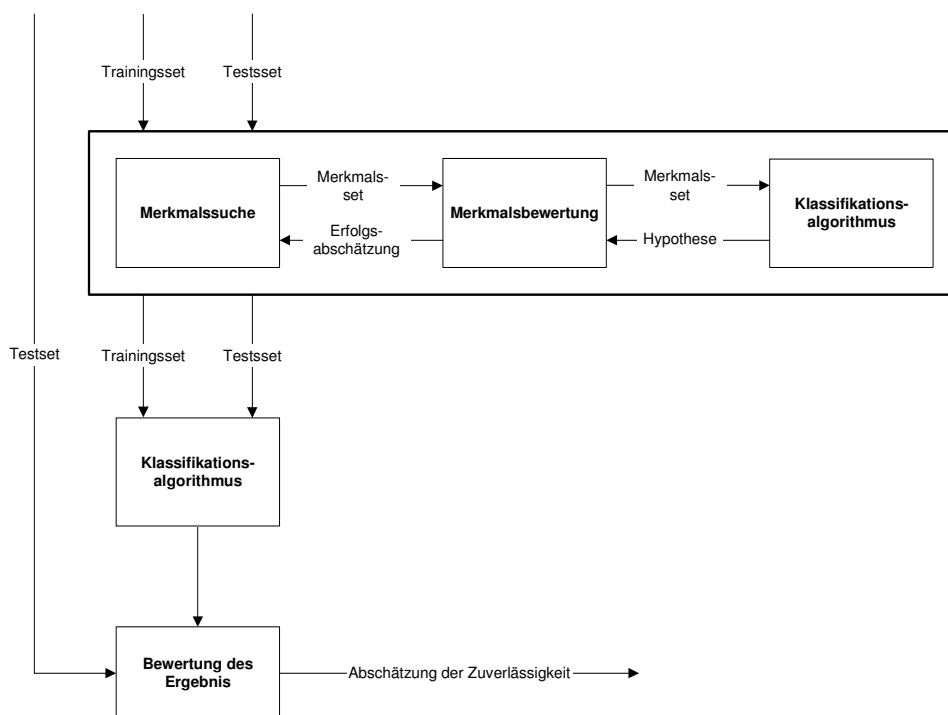
Im Gegensatz zu den erforderlichen mehr als 600 Millionen Kombinationen beim erschöpfenden Vergleich, der mit dem Verfahren des „Branch & Bound (B&B)“ zur optimalen Lösung führt, ergeben sich bei der SFS mit $D=152$ und $d=5$ insgesamt 750 Kombinationen. Bei der Verwendung von SBS würde es zu einer Anzahl von 11.613 zu überprüfenden Kombinationen führen. Bei einer Erhöhung der Dimension des Merkmalsvektors für die Klassifikation steigt die Anzahl der Kombinationen für das B&B sowie SFS,

während die Kombinationen für SBS sinken, wie im Folgenden detailliert ausgeführt wird.

Alle hier betrachteten Verfahren zur Merkmalsselektion werden zu den so genannten „Wrapper“-Strategien gezählt. Dies bedeutet, dass sich die Merkmalsselektion und die eigentliche Klassifikation „umarmen“. Das Prinzip, das allen den in dieser Arbeit betrachteten Selektionsstrategien gemein ist, zeigt nachfolgend Abbildung 4-9. Dieses Verfahren ermittelt quasi in einer Blackbox den besten Merkmalsvektor, der mit der eingesetzten Strategie der Merkmalsuche sowie dem gewählten Klassifikationsalgorithmus zu erreichen ist. Dabei wird auf einen Trainings- und Testdatensatz zurückgegriffen. Dies bedeutet, dass bei jedem Schritt der Merkmalsselektion eine Klassifikation durchgeführt wird, um das Subset an Merkmalen zu ermitteln, das das für diese Merkmalsanzahl geringste Fehlerkriterium aufweist.

**4 Mustererkennung:
Grundlagen,
Algorithmen und
Programmierung**

Abbildung 4-9:
Prinzip des Wrapper Ansatzes [nach KOV,1997]



Mit dem resultierenden Merkmalsvektor kann dann die eigentliche Klassifikation sowie eine Bewertung und Analyse des Ergebnisses der Mustererkennung durchgeführt werden.

Die Vorteile dieses Wrapper-Ansatzes liegen zum einen darin, dass die Merkmale im Zusammenhang bewertet werden, d.h. im Kontext mit anderen Merkmalen, so dass Abhängigkeiten und Korrelationen berücksichtigt werden. Zum anderen werden mögliche systematische Fehler des Klassifikators bereits in den Prozess der Merkmalsselektion einbezogen, da der Wrapper-Ansatz direkt auf die Minimierung des Fehlerkriteriums des Klassifikationsverfahrens abzielt.

4 Mustererkennung: Grundlagen, Algorithmen und Programmierung

4.3.1 Sequential Forward Selection (SFS)

Die „Sequential Forward Selection“ zählt zu den so genannten „Bottom up“-Suchstrategien, d.h. Ausgangspunkt ist die Initialisierung mit einem leeren Merkmalsset. In jedem Schritt der Suche wird nun ein Merkmal dem Merkmalsset hinzugefügt, bis die erwünschte Dimension des Merkmalsvektors erreicht ist. Es handelt sich bei SFS um eine suboptimale Strategie, da einmal gewählte Merkmale nicht wieder aus dem Merkmalsset entfernt werden können. In diesem Zusammenhang wird vom so genannten „Nesting effect“ gesprochen, da sich bildlich suboptimale Merkmale im reduzierten Merkmalsvektor „einnisten“.

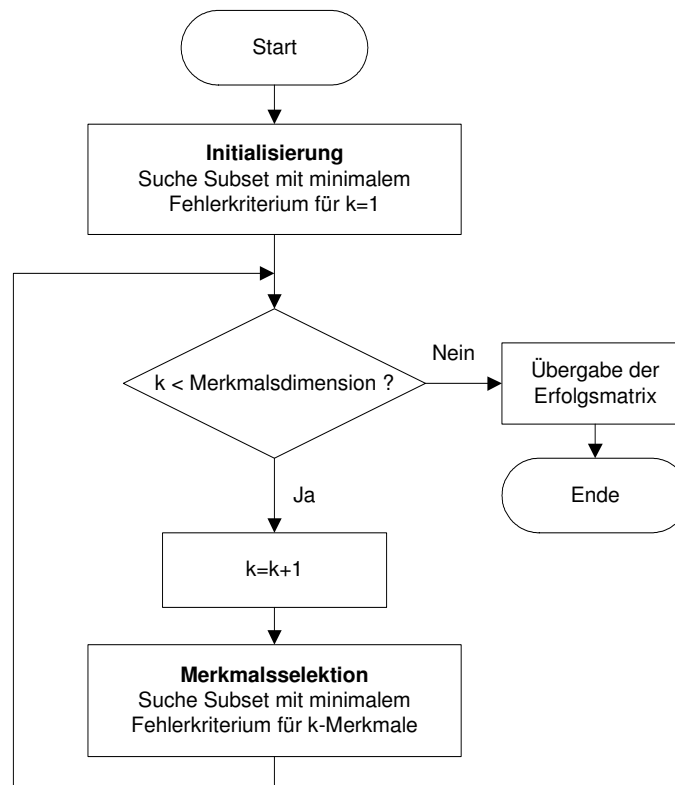
Die Anzahl der Kombinationen, die zum Ergebnis führen, ergibt sich aus der nachfolgenden Formel, mit D als ursprünglicher Merkmalsdimension und d als der gewünschten Größe des reduzierten Merkmalssets:

$$\text{Kombinationen}(D,d) = \sum_{i=1}^d (D+1) - i$$

Mit wachsender Größe des reduzierten Merkmalssets steigt auch die Anzahl der zu prüfenden Kombinationen bei unverändertem D .

Die schematische Darstellung der Merkmalsselektion mittels SFS ist als Flowchart in Abbildung 4-10 gezeigt. Die Auswahl der einzelnen Merkmale je Suchschritt erfolgt anhand des angewandten Fehlerkriteriums – es findet also eine Überprüfung aller möglichen Kombinationen eines Suchschrittes unter Anwendung des Klassifikationsverfahrens statt. Das Subset mit dem minimalen Fehlerkriterium wird als Folge ausgewählt und dient als Ausgangspunkt für den nächsten Suchschritt.

Abbildung 4-10:
Flowchart des Ablaufs der
Merkmalsselektion nach
SFS



4.3.2 Sequential Backward Selection (SBS)

Die „Sequential Backward Selection“ ist eine „Top down“-Strategie, die als eine Umkehrung der SFS verstanden werden kann. Ausgangspunkt ist die Gesamtmenge der Merkmale D . In jedem Schritt wird nun das „schlechteste“ Merkmal entfernt, bis die gewünschte Merkmalsanzahl d erreicht worden ist. In diesem Fall tritt der entgegengesetzte „Nesting effect“ ein, da einmal verworfene Merkmale nicht mehr eingefügt werden können, so dass es sich ebenfalls um eine suboptimale Suchstrategie handelt.

Die Anzahl der erforderlichen Kombinationen ergibt sich aus:

$$\text{Kombinationen}(D, d) = \sum_{i=d+1}^D i$$

Mit wachsender Größe des reduzierten Merkmalssets sinkt folglich die Anzahl der zu prüfenden Kombinationen bei unverändertem D , da die Suche entsprechend früher abgebrochen wird.

Abbildung 4-11 verdeutlicht als Flowchart die Merkmalsselektion mittels SBS. Ebenso wie bei der SFS erfolgt die Auswahl des zu eliminierenden Merkmals über eine Klassifikation aller vorhandenen Merkmalssets und der Wahl des Sets mit dem minimalen Fehlerkriterium. Dies bedeutet, dass jeweils das Merkmal, das den geringsten Einfluss auf das bestmögliche Ergebnis einer Klassifikation hat, entfernt wird.

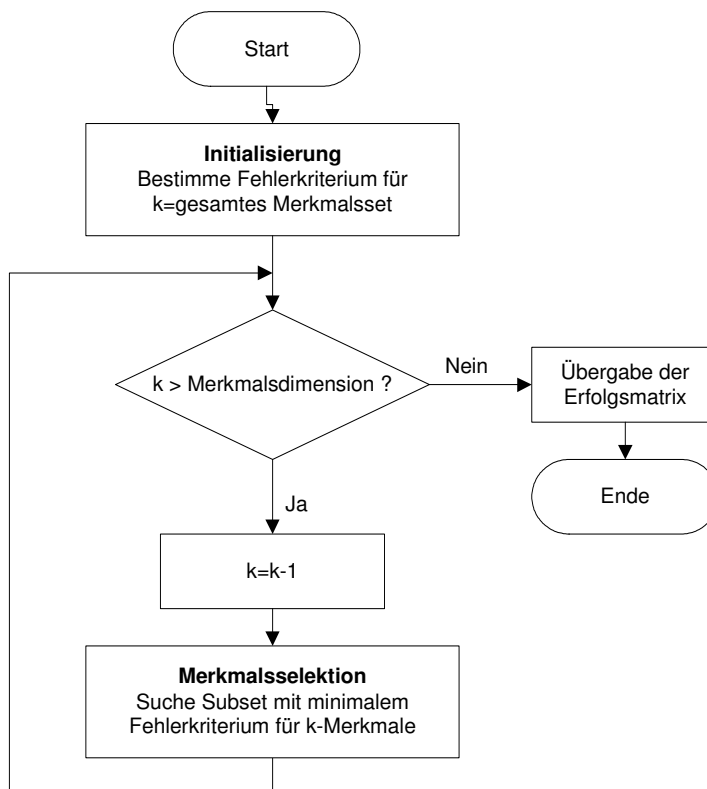


Abbildung 4-11:
Flowchart des Ablaufs der
Merkmalsselektion nach
SBS

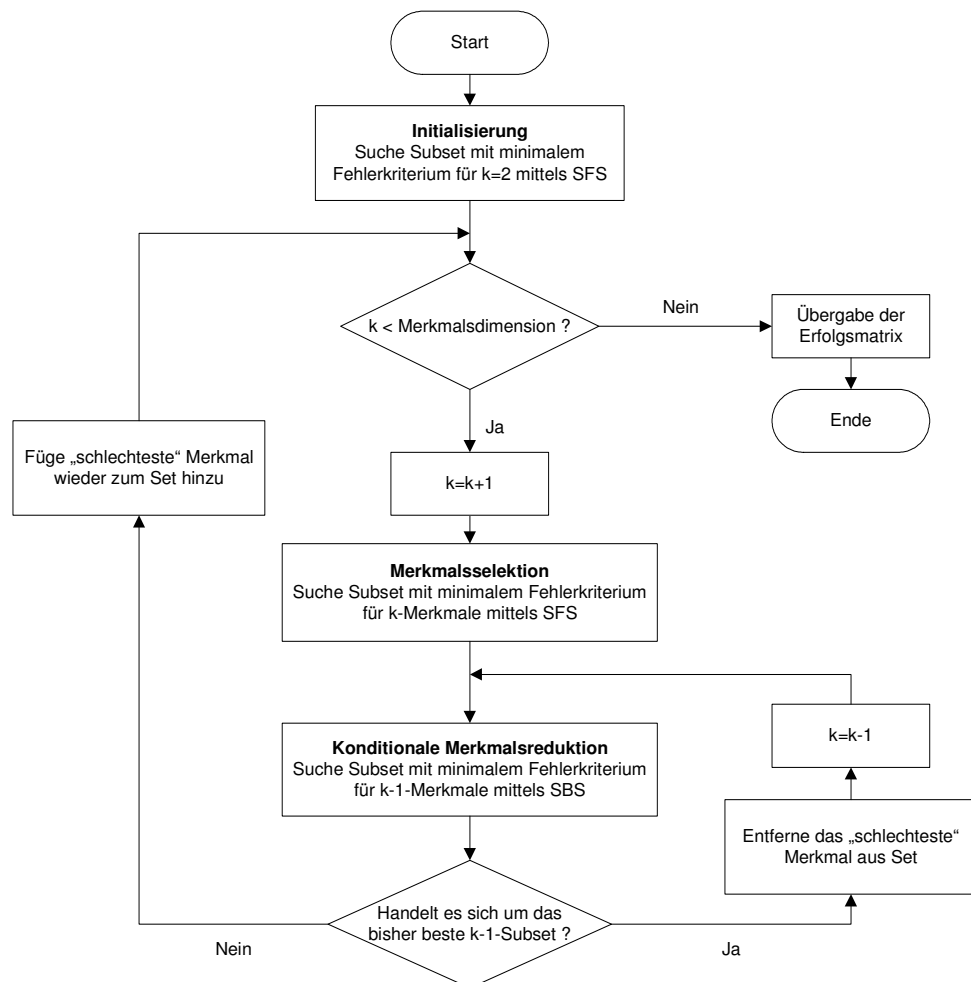
4 Mustererkennung:
Grundlagen,
Algorithmen und
Programmierung

4.3.3 Sequential Forward Floating Selection (SFFS)

Um zu verhindern, dass sich der „Nesting effect“ einstellt und bei der Merkmalsselektion einmal entfernte Merkmale unwiderruflich aus dem zur Verfügung stehenden Set entfernt werden, bietet sich die Verwendung von dynamischen Suchstrategien an. Zu diesen „Floating Search Strategies“ zählt unter anderem die „Sequential Forward Floating Selection (SFFS)“ [SOM, 1999]. Diese Verfahren lassen sich in gewissem Sinne im Vergleich zu den bereits oben vorgestellten „Sequential Selection“ Ansätzen auch als Algorithmen der zweiten Generation für die Merkmalsselektion bezeichnen.

Bei diesem flexiblen Ansatz werden die Vorteile der beiden bereits genannten Strategien, SFS und SBS, kombiniert. Da es sich um einen „Bottom Up“-Ansatz handelt, erfolgt die Initialisierung mit einem leeren Merkmalsset, das mittels SFS auf zwei Merkmale erweitert wird. Anschließend ist der SFFS – Algorithmus durch drei Subroutinen definiert, die im Ablaufdiagramm in Abbildung 4-12 graphisch verdeutlicht werden.

Abbildung 4-12:
Flowchart des Ablaufs der
Merkmalsselektion nach
SFFS



Zunächst fügt der Algorithmus das signifikanteste Merkmal aus den verbleibenden Merkmalen hinzu, d.h. ein weiterer Schritt mittels SFS. Dann folgt der Schritt der konditionalen Merkmalsreduktion, in der überprüft wird, ob sich durch die Entfernung eines Merkmals aus dem aktuellen Subset eine Minimierung des Fehlers ergibt. Wenn es ein Subset gibt, das ein bes-

seres Ergebnis als in dem bis dato besten $k-1$ -Set liefert, wird das entsprechende Merkmal entfernt, d.h. ein SBS-Schritt wird vorgenommen, und ein weiterer Durchlauf der konditionalen Merkmalsreduktion folgt. Weist das reduzierte $k-1$ -Subset jedoch kein kleineres Fehlerkriterium auf, so wird das entfernte Merkmal wieder hinzugefügt und der Algorithmus kehrt zum SFS-Teil zurück, falls die erwünschte Dimension des Merkmalsvektors noch nicht erreicht ist.

Hinsichtlich der Anzahl der erforderlichen Kombinationen zur Ermittlung des besten mittels SFFS zu erlangenden Merkmalssets lässt sich daher keine allgemeingültige Aussage treffen. Der Algorithmus erhöht oder reduziert dynamisch in Abhängigkeit vom besten Fehlerkriterium die Anzahl der im Subset vorliegenden Merkmale, bis die erwünschte Anzahl erreicht ist. Bei progressiver Verschlechterung der Ergebnisse mit steigender Merkmalsanzahl, d.h. die konditionale Merkmalsreduktion führt zu keinen besseren Ergebnissen, gleicht die SFFS dem Vorgehen bei der SFS.

In dieser Arbeit wird neben den Basisverfahren zur Merkmalsreduktion ausschließlich der SFFS betrachtet. Die rückwärtsgerichtete „Sequential Backward Floating Selection (SBFS)“ wird nicht genutzt, um die LIF-Ergebnisse auszuwerten. Grund hierfür ist, dass die Ergebnisse im Allgemeinen keine wesentliche Verbesserung des Klassifikationsergebnisses zur Folge haben. Auch die in der Literatur bekannten Strategien der Nachfolgegeneration, der „Adaptive Floating Search Methods (ASFM)“, werden an dieser Stelle nicht weiterverfolgt. Entscheidend für die Beschränkung auf die SFFS war, dass diese im Vergleich zur Selektion mittels SBFS oder ASFM erhebliche Reduktionen in der Rechengeschwindigkeit zur Folge hat.

4.3.4 Wrapper2-Ansatz

Bei den drei auf dem Wrapper-Ansatz basierenden Ansätzen zur Merkmalsselektion stellt sich die Frage nach der Stabilität der Ergebnisse. Unter Stabilität wird in diesem Zusammenhang verstanden, ob die verwendete Strategie unabhängig von einer geänderten Datenreihenfolge auch stets denselben reduzierten Merkmalsvektor generiert. Bei der im Rahmen dieser Arbeit durchgeführten Mustererkennung sind eine solche Variation der Ergebnisse und die damit verbundene Instabilität der Selektionsstrategie ebenfalls nicht auszuschließen. Um diese, für die Klassifikation der Materialien mit erheblichen Folgen verbundene Fehlerquelle bereits im Vorfeld zu eliminieren, wird auf den Ansatz des „Wrapper2“ oder der zweifachen „Umarmung“ zurückgegriffen [DUN, 2002].

Grundsätzlich bedient sich der „Wrapper2“ einer statistischen Betrachtung der selektierten Merkmale über eine gewisse Anzahl von Wiederholungen des Selektionsprozesses bei leicht veränderten Trainingsdaten. Im Einzelnen gliedert sich die Merkmalsselektion nach „Wrapper2“ für alle Selektionsstrategien in die in Abbildung 4-13 und Abbildung 4-14 dargestellten Abläufe.

4 Mustererkennung:
Grundlagen,
Algorithmen und
Programmierung

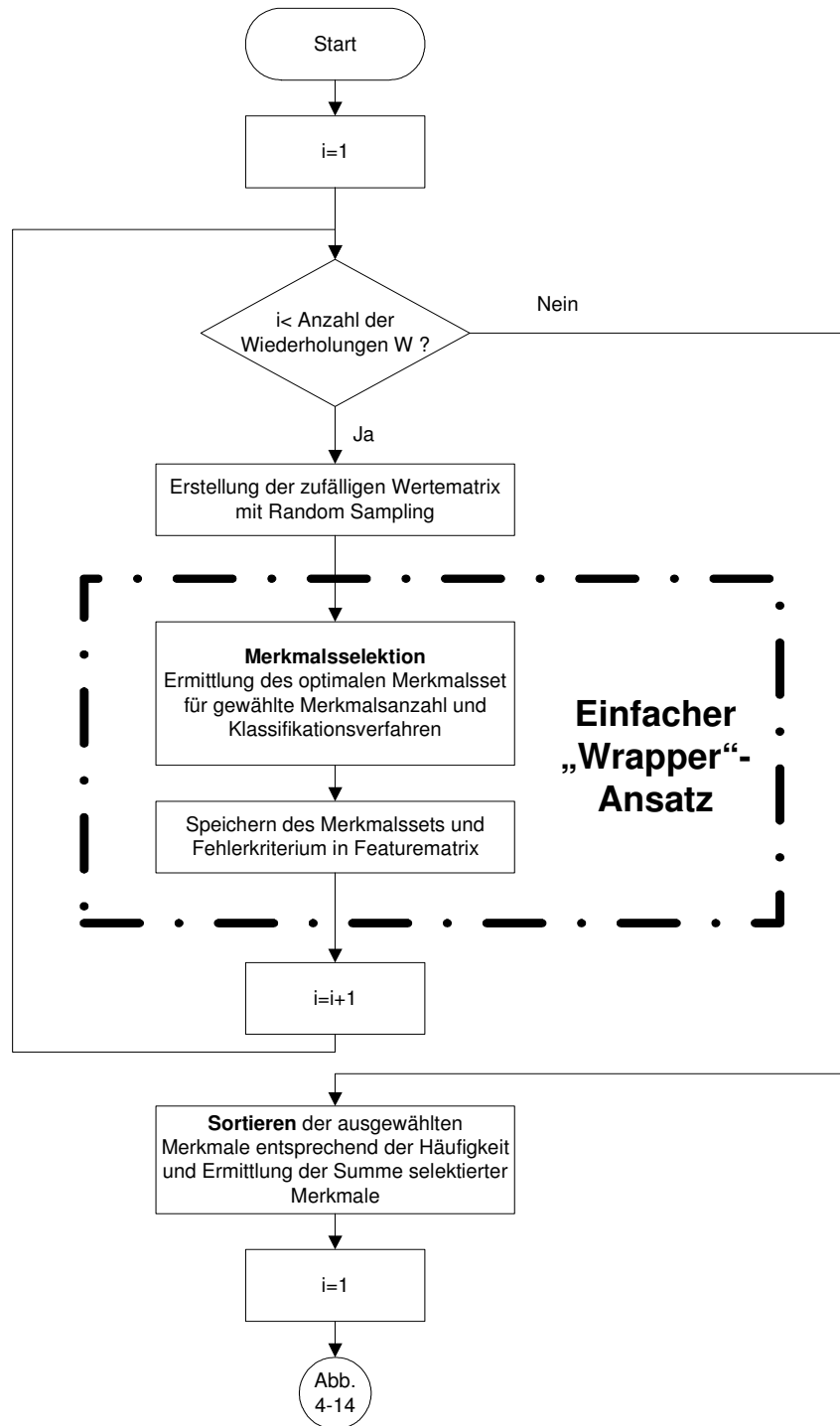


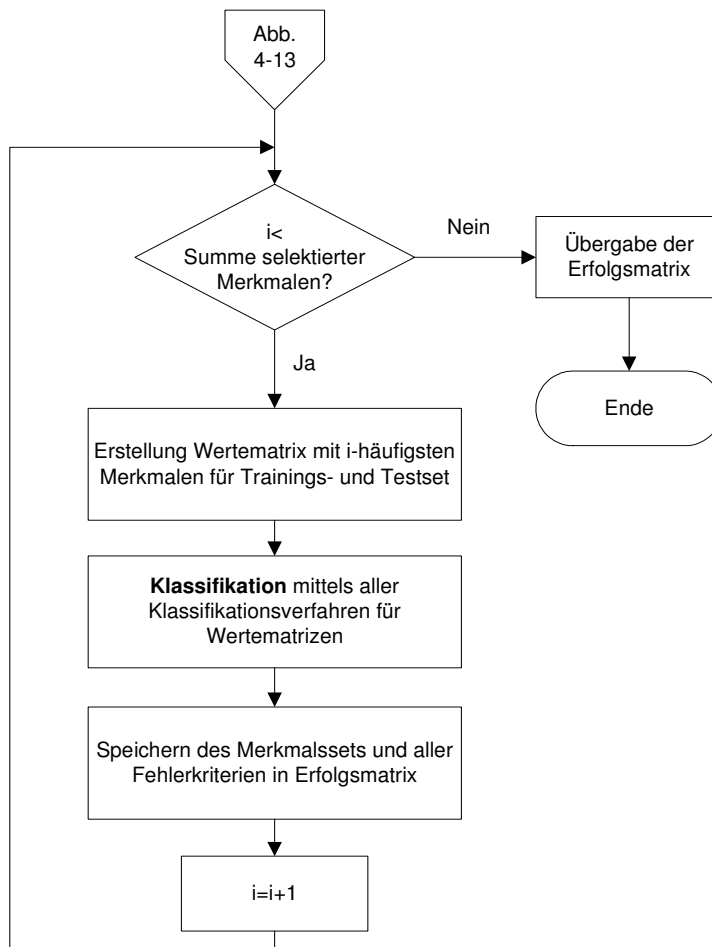
Abbildung 4-13:
Flowchart des Ablaufs der
Merkmalsselektion nach
Wrapper-2 (a)

Die Anzahl der Durchläufe oder Wiederholungen wird mit W bezeichnet. Bei $W=1$ entspricht die Merkmalsselektion dem ursprünglichen Verfahren. Die Variation der Trainingsdaten wird über die „Resampling rate“ R gesteuert, d.h. $R=80$ bedeutet, dass 80% der ursprünglichen Datenmenge in dem jeweiligen Durchlauf zufallsverteilt verwendet werden. Die Darstellung des Vorgangs des Random Sampling ist im Anhang angegeben. Die Testdatenmenge bleibt von der Variation unberührt.

Mit diesen neuen Trainingsdaten wird nun die Merkmalsselektion durchgeführt und der reduzierte Merkmalsvektor in einer Featurematrix gespeichert. Diese Matrix sammelt über die Anzahl der Wiederholungen alle generierten

Merkmalssets, aus denen dann ein Histogramm der Häufigkeitsverteilung aller Merkmale gewonnen wird.

Ausgehend von der Rangfolge der sortierten Merkmale wird nun die abschließende Klassifikation in Abhängigkeit von der Merkmalsanzahl durchgeführt. Zunächst erfolgt also die Klassifikation ausschließlich mit dem am häufigsten selektierten Merkmal und dann in absteigender Reihenfolge mit den n-Häufigsten. Auf diese Weise lässt sich der Einfluss der Merkmalsdimension auf das Fehlerkriterium bewerten. Die abschließende Klassifikation wird unabhängig vom dem im Basis-Wrapper verwendeten Algorithmus für alle in Kapitel 4.2 vorgestellten Klassifikatoren durchgeführt. Dies ermöglicht eine Analyse des Einflusses des Klassifikationsverfahrens auf das Fehlerkriterium bei gleichen Merkmalssets.



**4 Mustererkennung:
Grundlagen,
Algorithmen und
Programmierung**

Abbildung 4-14:
Flowchart des Ablaufs der
Merkmalsselektion nach
Wrapper-2 (b)

5 Parametereinstellungen zur Merkmalsselektion

Bevor die Merkmalsselektion und Klassifikation für alle bemusterten Lagerstättenpartien durchgeführt werden kann, bedarf es zunächst einer detaillierten Betrachtung der relevanten Parameter, um diese optimal auf die Aufgabenstellung hin abstimmen zu können. Zur Minimierung der durch äußere Fehlerquellen verursachten Instabilitäten bei der Merkmalsauswahl wurde die statistische Vorgehensweise des Wrapper2-Ansatzes zur Merkmalsselektion und Klassifikation gewählt. Der Grund dafür stellt das Problem der erheblichen Streuungen der Messwerte dar, die über ein vertretbares Grundrauschen hinausgehen. Verantwortlich sind dafür, wie in Kapitel 3 ausgeführt, systematische Fehler wie bei der Entfernungskalibrierung oder stochastische Schwankungen etwa bei der Laseranregungsenergie. Ebenso führt die Tatsache, dass die beprobten Materialien am Stoß selbst in den definierten Klassen sehr inhomogene Eigenschaften etwa durch eingelagerte Holzstücke etc. aufweisen, zu hohen Varianzen innerhalb der Klassen.

Durch die Variation der Testdatenmenge mittels der zufallsverteilten Selektion einer Teilmenge wird versucht, die Einflüsse der Streuungen auf die selektierten Merkmale zu minimieren. Der Anteil der berücksichtigten Teilmenge zur Gesamtmenge, als „Resampling Rate“ R bezeichnet, wird in dieser Arbeit mit 80% definiert. In der Literatur wurde in umfangreichen Testreihen festgestellt, dass dieser Wert ausreichend ist, um Varianzen innerhalb der Messdaten zurückzuführende Instabilitäten ausgleichen zu können [DUN, 2002]. Aus diesem Grunde wurden für die Grenzschiefererkennung mittels LIF in dieser Ausarbeitung keine gesonderten Untersuchungen durchgeführt.

Bei den durchgeführten Klassifikationen wurde dabei stets berücksichtigt, dass die Materialklassen nicht durch zu wenige Messwerte gebildet werden. Um stets statistisch relevante Stichprobengrößen zu gewährleisten, bedarf es aufgrund der Innerklassenvarianzen einer ausreichenden Anzahl an Messwerten, um die klassentypischen Eigenschaften abbilden zu können. Der minimale Stichprobenumfang ergibt sich bei den Untersuchungen im Tagebau Hambach für den Abschnitt nördlich des Liegendrückens für die Holzlage mit 71 Messwerten.

Bei den Voruntersuchungen zur Parameterbestimmung bei der Merkmalsselektion und Materialklassifikation sind über die „Resampling Rate“ hinaus drei weitere Stellgrößen zu betrachten.

Der erste Parameter ist der Einfluss der Durchlaufanzahl auf die Merkmalsselektion, da die Rechendauer des Algorithmus proportional zur Anzahl der Wiederholungen der Selektion ist.

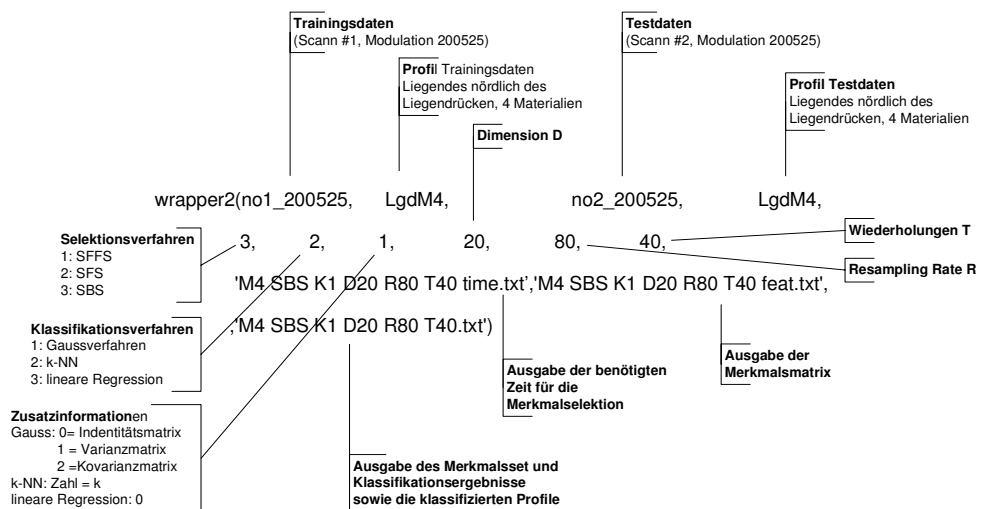
Ist die optimale Anzahl der als „Trials“ T benannten Durchläufe bekannt, stellt sich die Frage nach der optimalen Anzahl der zu ermittelnden Merkmale selber. Diese Größe wird im Weiteren als Dimension D bezeichnet. Wie bereits erläutert, steigt bei den „Bottom up“-Ansätzen wie SFFS und SFS die Anzahl der möglichen Kombinationen mit wachsender Größe des reduzierten Merkmalssets. Bei den „Top down“-Verfahren hingegen sinkt die Anzahl der zu untersuchenden Kombinationen mit steigender Größe des reduzierten Sets. Aufgabe ist es nun zu untersuchen, welchen Einfluss T auf die Qualität der selektierten Merkmale hat.

Der dritte Parameter, der in den Voruntersuchungen betrachtet werden muss, ist das für die Materialerkennung der beprobten Lagerstätteprofile bzw. Merkmalsselektion eingesetzte Selektionsverfahren. Der erhebliche Rechenaufwand ist ein Grund dafür, dass nicht alle drei vorgestellten Verfahren auch für die endgültige Aufgabe der Trennflächenerkennung eingesetzt werden können. Das Fehlerkriterium der Klassifikation für die jeweiligen Selektionsverfahren ist darüber hinaus ebenfalls ein entscheidendes Kriterium und Qualitätsmerkmal.

Die Voruntersuchungen wurden mit dem in Matlab programmierten Modul wrapper2.m durchgeführt. Die Nomenklatur ist in Abbildung 5-1 dargestellt. Der Aufruf des Moduls erfolgt beispielhaft für eine Merkmalsselektion über folgenden Befehl:

```
wrapper2 (no1_200525,LgdM4,no2_200525,LgdM4,3,2,1,20,80,40, 'M4 SBS K1 D20 R80 T40 time.txt', 'M4 SBS K1 D20 R80 T40 feat.txt', 'M4 SBS K1 D20 R80 T40.txt')
```

Abbildung 5-1: Nomenklatur des Aufrufs von wrapper2.m



Die Ausgabe erfolgt über txt-Dateien. Die Hauptdatei enthält sowohl die Häufigkeitsverteilung der selektierten Merkmale als auch die Ergebnisse der Klassifikation. Dabei wird die Materialzuordnung für die selektierten Merkmale in ansteigender Rangfolge für alle Klassifikationsverfahren durchgeführt, obwohl die Selektion nur auf der Basis eines Verfahrens erfolgt. Darüber hinaus werden für alle Klassifikationen auch die entsprechenden Profile protokolliert. Ein Vergleich zwischen dem tatsächlichen

5 Parametereinstellungen zur Merkmalsselektion

Profil und der Zuordnung der Messwerte bei der Klassifikation ist also möglich. Dies erlaubt es, die für die Klassifikation kritischen Bereiche des Profils eindeutig zu bestimmen und entsprechend zu analysieren.

Die Modulation der Daten für die Bestimmung der Parametereinstellungen erfolgte mit einer Lochweite der Blende von 20, d.h. es wurden jeweils die zehn Werte vor und die zehn Werte nach der entsprechenden Messung betrachtet. Werte unter dem 0,5fachen bzw. über dem 2,5fachen des 20er Mittels bezogen auf den betrachteten Wert wurden moduliert. Im Folgenden wird für eine solche Modulation von „20 0K5 2K5“ gesprochen.

Für die Voruntersuchungen wurde die Messung am Liegenden des Flözes Frimmersdorf a nördlich des Liegendrückens gewählt (12.11.2002). Grund hierfür waren die signifikanten Materialwechsel im Profil mit vier Klassen (LgdM4: GK-GGMK, teilweise leicht holzig; Holzlage; GK, tonig, holzig; Liegendton) und die hohe Anzahl von Messwerten je Klasse. Die Anzahl der Originalmerkmale liegt bei 79. Dabei wurden die Abklingparameter W1/W2 und W2/W3 für beide PMT's getrennt sowie die Quotienten $PMT A \frac{xy}{nm} W1$ bzw. $PMT B \frac{yz}{nm} W1$ herangezogen. Detaillierte Angaben über die verwendeten Einstellungen sind Anhang B zu entnehmen.

5.1 Ermittlung der optimalen Durchlaufanzahl

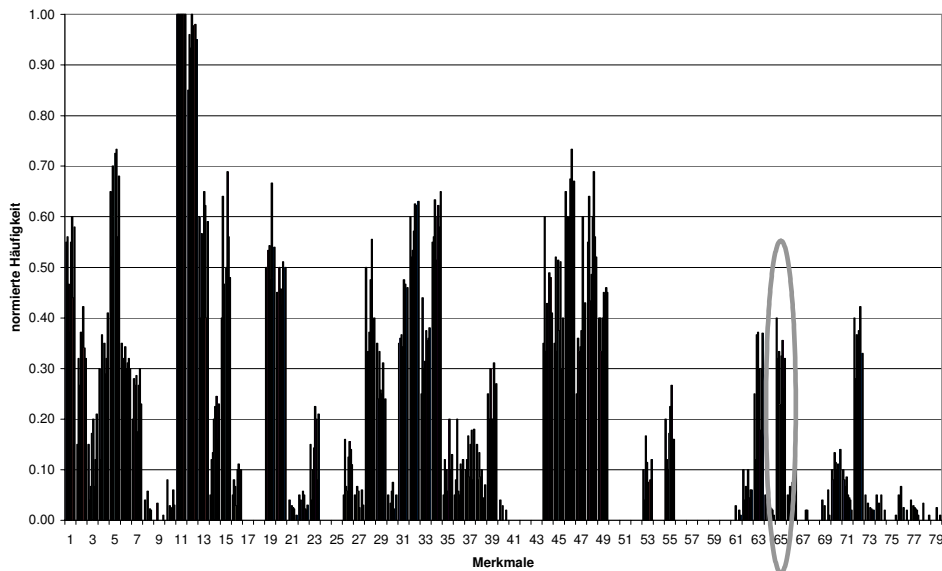
Um die optimale Durchlaufanzahl für die Merkmalsselektion bestimmen zu können, werden bis zu zehn unterschiedliche Einstellungen für T berechnet und analysiert. Für alle drei Klassifikationsverfahren wird dabei auf die Methode der „Sequential Forward Selection“ zur Reduktion der Merkmale zurückgegriffen, um eine gemeinsame Vergleichsbasis sowie auch ein einheitliches Vorgehen zu gewährleisten. Die Größe des reduzierten Merkmalssets wird für die Ermittlung des optimalen T auf 25 Merkmale (D) festgesetzt. R beträgt wie auch in den folgenden Berechnungen 80%.

Zur Entscheidungsfindung bezüglich eines definierten T für die Durchführung der Merkmalsselektion und Klassifikation sämtlicher Messdaten erfolgt eine Gegenüberstellung der selektierten Merkmale der unterschiedlichen Durchlaufvarianten gegen die von T100. Bei 100 Durchläufen wird unterstellt, dass sich eine stabile statistische Verteilung der Merkmale einstellt. Dass diese Stabilisierung der Rangfolge der selektierten Merkmale mit steigender Anzahl an Durchläufen erfolgt, zeigten schon Dunne et alii [DUN, 2002].

Wesentlich für die Auswahl von T ist darüber hinaus die erforderliche Rechendauer für die einzelnen Klassifikationsverfahren beim vorgegebenen SFS-Ansatz. Die optimale Durchlaufanzahl ergibt sich daher aus der Abwägung zwischen Rechenzeit und der minimalen Abweichung zur Rangfolge bei T100.

Abbildung 5-2 zeigt die normierte Häufigkeit der selektierten Merkmale in Form eines Histogramms. Die Durchlaufanzahl wurde zwischen T20 und

T50 in Fünferschritten variiert, darüber hinaus stellt der jeweils rechte Wert T100 dar. Als Klassifikationsverfahren kam in diesem Falle 1-NN zum Einsatz. Es zeigt sich in dieser Abbildung, dass eine Instabilität in der Selektion der Merkmale bei unterschiedlichen T vorliegt.



5 Parametereinstellungen zur Merkmalsselektion

Abbildung 5-2:
Histogramm (normierte Häufigkeit) T20-T100 für SFS 1-NN, D25

Zum besseren Verständnis des Histogramms verdeutlicht Abbildung 5-3 die normierte Häufigkeit für das Merkmal PMT A 600 nm W1/W3, das in der obigen Abbildung markiert ist. Die Durchläufe T25 und T45 liegen unter 20%, während bei den anderen Voruntersuchungen die Häufigkeiten zwischen 30 und 40% schwanken.

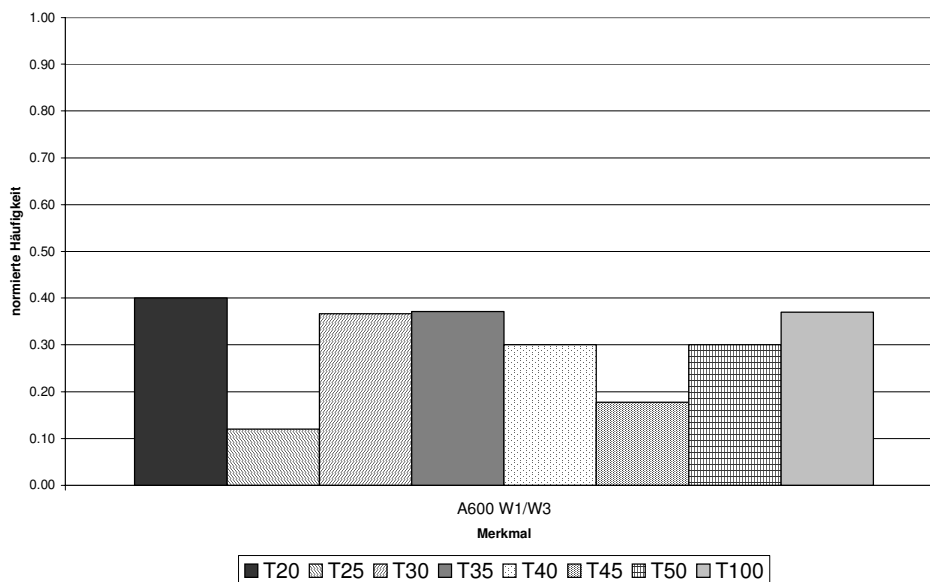


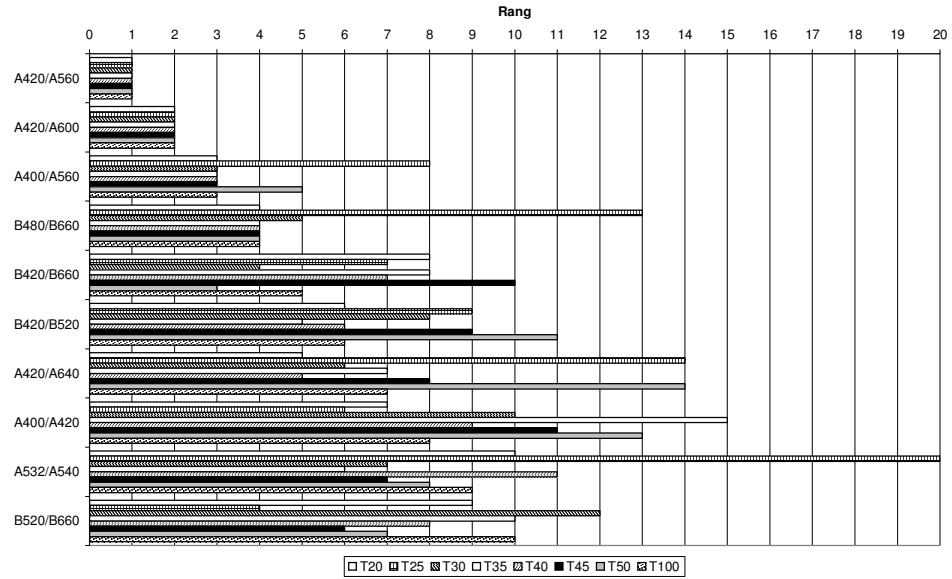
Abbildung 5-3:
Histogramm (normierte Häufigkeit) T20-T100 für SFS 1-NN, Merkmal A600 W1/W3, D25

Ausgehend von der vorherigen Darstellung lässt sich anhand Abbildung 5-4 ablesen, welchen Einfluss T auf die Merkmalsselektion, aber auch auf den Rang eines Merkmals im Sinne der Häufigkeit der Selektion in den T Durchläufen hat. Absteigend vom wichtigsten Merkmal, sprich am häufigsten ausgewählten Merkmal in den T Durchläufen, werden die Top 10 sortiert nach T100 dargestellt. Es zeigt sich, dass z.B. die beiden ersten Merkmale (A420 nm/ A 560 nm und A420 nm/ A 600 nm) unabhängig von

5 Parametereinstellungen
zur Merkmalsselektion

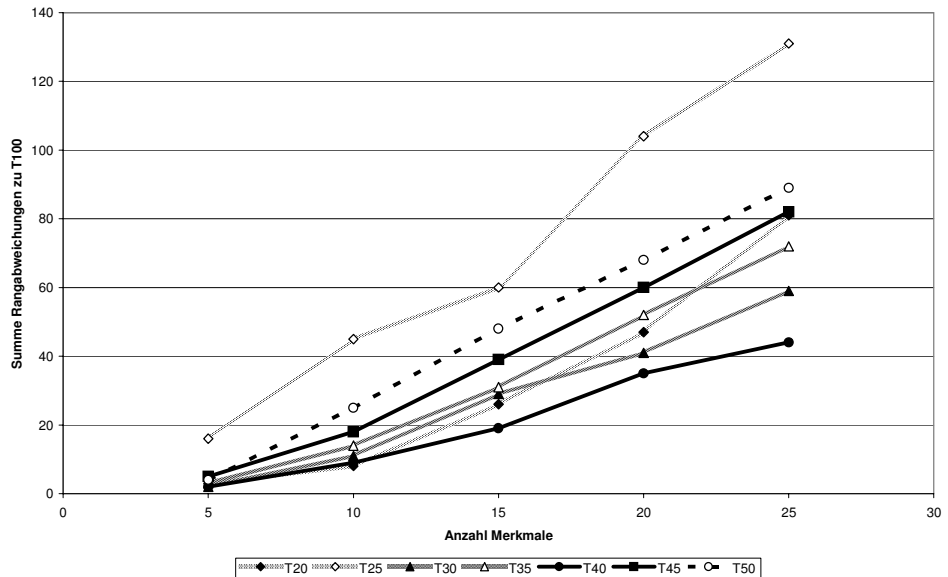
Abbildung 5-4:
Rangfolge nach T100 für
SFS 1-NN, D25

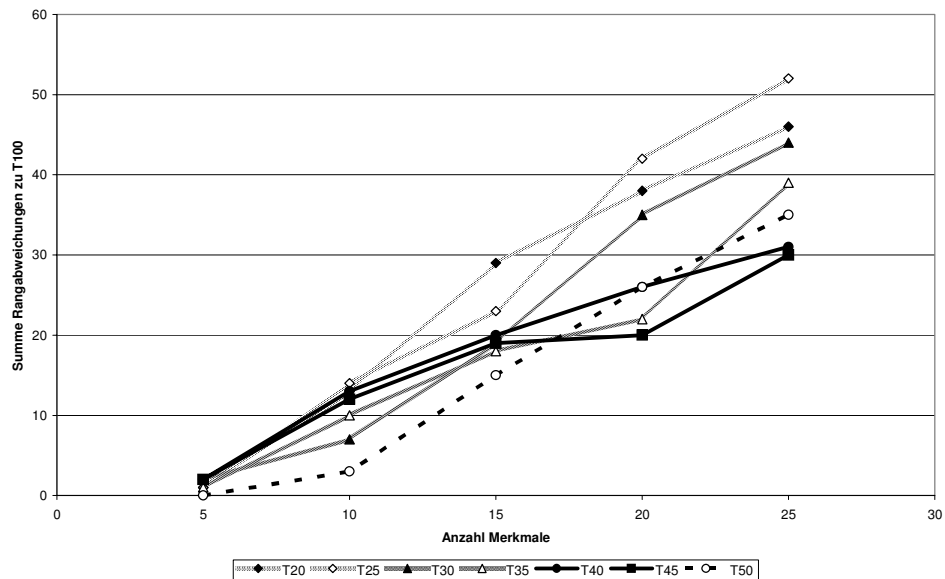
T am besten zur Klassifikation des Materials nach 1-NN geeignet sind. Mit sinkender Bedeutung des Merkmals stellen sich Abweichungen der Rangfolge für die kleineren Durchlaufzeiten im Vergleich zu T100 ein.



Zur Auswahl der optimalen Anzahl der Durchläufe bedarf es daher einer Bewertung der Abweichungen. Diese Analyse erfolgt mittels der Aufsummierung der absoluten Rangabweichungen der einzelnen Durchlaufvarianten gegen T100 für die ersten fünf, zehn, 15, 20 und 25 selektierten Merkmale. Damit soll der Einfluss von T auf die Zuverlässigkeit hinsichtlich der Übereinstimmung zu T100 bestimmt werden. Abbildung 5-5 und Abbildung 5-6 zeigen dies für 1-NN bzw. das lineare Regressionsverfahren.

Abbildung 5-5:
Summe der Rangabweichungen T20-T50 gegen
T100 für SFS 1-NN, D25





5 Parametereinstellungen zur Merkmalsselektion

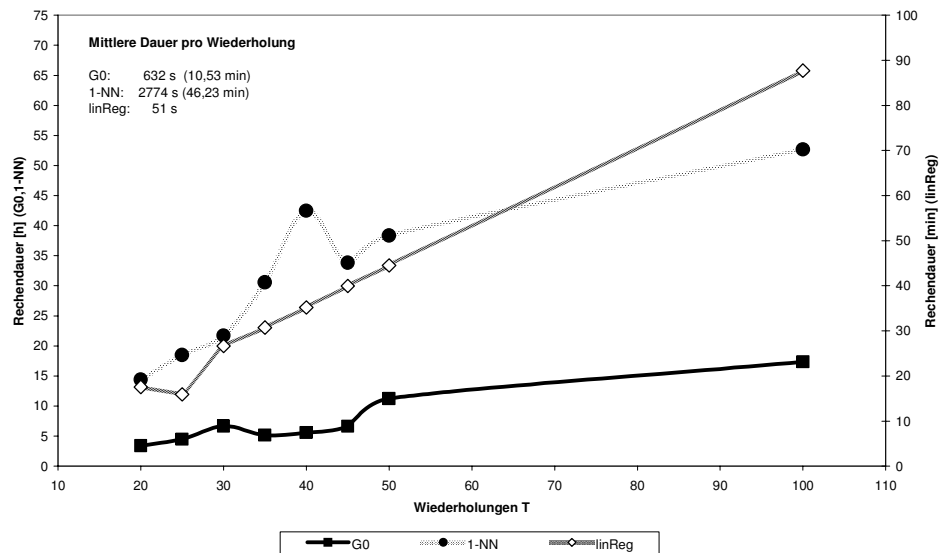
Abbildung 5-6:
Summe der Rangabweichungen T20-T50 gegen T100 für SFS, linReg, D25

Für beide Klassifikationsverfahren zeigt sich die bereits angesprochene Tendenz, dass mit wachsender Größe des reduzierten Merkmalssets die Abweichungen zu T100 steigen. Für kleine Merkmalssets stellt sich die Summe der Abweichungen für alle T sehr ähnlich dar. Die Bestimmung des optimalen Durchlaufparameters bei $d=15$ erweist sich als schwierigere Aufgabe, da dieser abhängig vom gewählten Klassifikationsalgorithmus schwankt. Der Bereich für T zwischen 30 und 40 zeigt tendenziell die geringsten Abweichungen zu T100 für steigende Sets. Bei der Reduktion auf die besten 10 bis 15 Merkmale sind jedoch auch 20 Durchläufe zu vertreten, da sich die Abweichungen von T100 erst bei einer Erhöhung des Merkmalssets verstärkt negativ auswirken.

Hinsichtlich der erforderlichen Rechendauer zeigt Abbildung 5-7 für die drei Klassifikationsverfahren den zeitlichen Aufwand für SFS. Zu beachten ist, dass die Skaleneinteilung der Graphen für G0 und 1-NN nicht der Skalierung der linearen Regression entspricht. Die Schwankungen, besonders für 1-NN, haben ihre Ursache in der Nutzung unterschiedlicher Prozessgenerationen mit entsprechenden Leistungseinbußen. Für alle Verfahren ist der Trend jedoch annähernd linear. Die mittlere Dauer für einen Durchlauf ist, wie die Abbildung zeigt, stark abhängig vom verwendeten Algorithmus. Das Verfahren der linearen Regression benötigt pro T lediglich 51 s, während für das Gaußverfahren mit der Identitätsmatrix bereits 10 min und 1-NN sogar mehr als 45 min pro Iterationsschritt erforderlich sind. Der Grund dafür liegt bei 1-NN im komplexen Abgleich aller Trainingsdaten mit den Testdaten.

5 Parametereinstellungen zur Merkmalsselektion

Abbildung 5-7: Rechendauer T20-T100 für SFS, D25



Um nun ein optimales Selektionsergebnis unter Berücksichtigung der Qualität der Merkmalsauswahl sowie der dafür erforderlichen Rechendauer zu garantieren, wird im folgenden mit einem Wert für T von 40 gearbeitet.

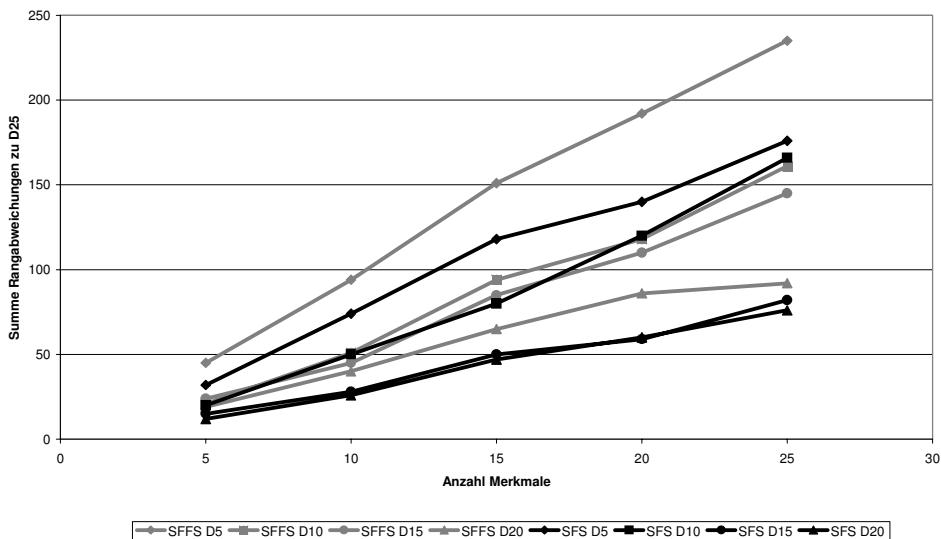
5.2 Ermittlung der optimalen Merkmalsanzahl

Um die optimale Merkmalsanzahl D für die Selektion bzw. Klassifikation bestimmen zu können, wurden fünf unterschiedliche Einstellungen für D berechnet und analysiert. Für alle drei Klassifikationsverfahren wurde dabei auf die drei Selektionsverfahren zur Reduktion der Merkmale zurückgegriffen. Ausgehend von den Ergebnissen aus dem vorherigen Unterkapitel 5.1 wurde die Anzahl der Durchläufe T auf 40 festgesetzt. R betrug 80%.

Zur Auswahl der optimalen Größe des reduzierten Merkmalssets bedarf es einer Bewertung der Abweichungen der selektierten Merkmale. Diese Analyse erfolgt mittels der Aufsummierung der absoluten Rangabweichungen der einzelnen Durchlaufvarianten gegen D25 für die ersten fünf, zehn, 15 und 20 selektierten Merkmale. Damit soll der Einfluss von D auf die Zuverlässigkeit hinsichtlich der Übereinstimmung zu D25 bestimmt werden. Abbildung 5-8 und Abbildung 5-9 zeigen dies für SFFS und SFS beim Gaußverfahren sowie 1-NN.

Bei beiden Verfahren sinkt die Übereinstimmung mit Erweiterung des Merkmalssets. Dabei weisen die Varianten mit größeren D geringere Abweichungen von D=25 auf, sowohl für SFFS als auch SFS.

Klassifikationsverfahren zur Materialerkennung



5 Parametereinstellungen zur Merkmalssektion

Abbildung 5-8: Summe der Rangabweichungen D5-D20 gegen D25 für SFFS und SFS, Gaußverfahren, T40

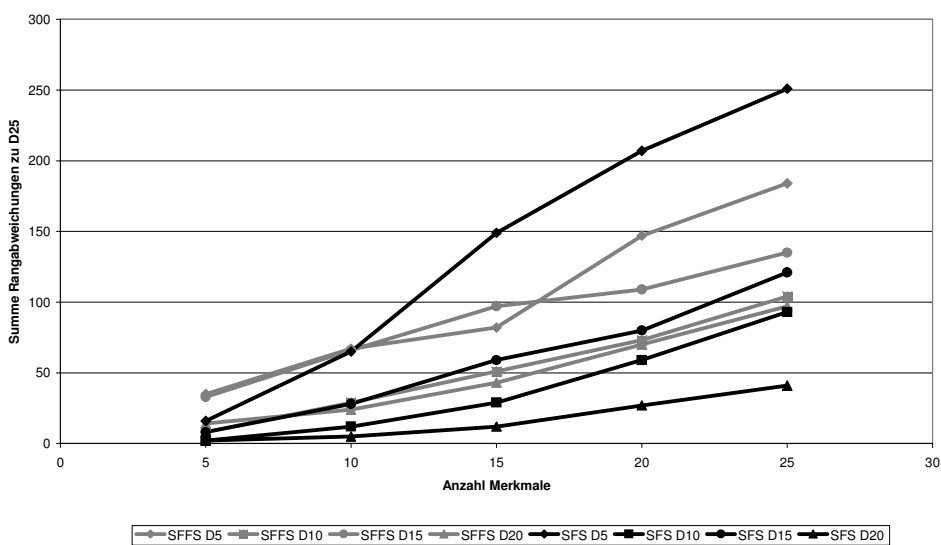


Abbildung 5-9: Summe der Rangabweichungen D5-D20 gegen D25 für SFFS und SFS, 1-NN, T40

Bei Einbeziehung der Ergebnisse der „Sequential Backward Selection“ zeigt sich, dass die Abweichungen im Gegensatz zu den „Bottom up“ - Verfahren größer sind. Aber auch hier zeigt z.B. D20 eine höhere Übereinstimmung zu D25 als D15 (Abbildung 5-10).

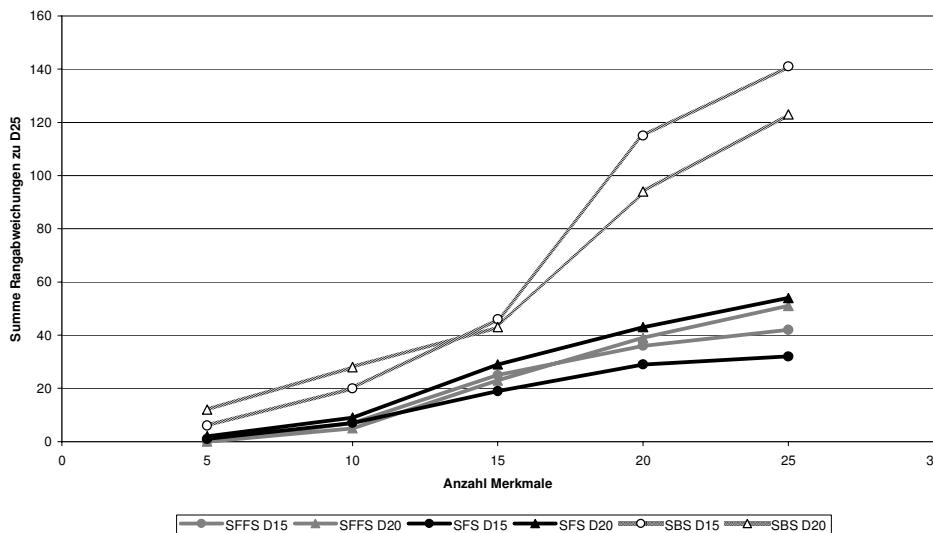


Abbildung 5-10: Summe der Rangabweichungen D15 und D20 gegen D25 für SFFS, SFS und SBS, linReg, T40

5 Parametereinstellungen zur Merkmalsselektion

Der erforderliche Rechenaufwand steigt für die beiden „Forward“-Verfahren mit steigender Größe des Merkmalssets, wie bereits formelmäßig in Kapitel 4.3 bei der Vorstellung der Verfahren zur Merkmalsselektion dargelegt. Dies zeigt sich in Abbildung 5-11. Die sinkende Rechendauer mit steigendem D des Merkmalssets lässt sich ebenfalls anhand der Ergebnisse der Voruntersuchungen graphisch darstellen (Abbildung 5-12). Aufgrund der sehr langen Rechendauer wurde auf die systematische Analyse des 1-NN-Verfahrens verzichtet. Alleine für die Ermittlung des reduzierten Merkmalssets D25 wurde bereits eine Rechendauer von 111 h (etwa 4,5 Tage) benötigt. Im Vergleich zwischen SFFS und SBS erhöht sich der durchschnittliche Rechenaufwand pro reduzierter bzw. erhöhter Dimension je nach eingesetztem Klassifikationsansatz um den Faktor 17 bis 25.

Abbildung 5-11: Rechendauer D5 bis D25 für SFFS und SFS, T40

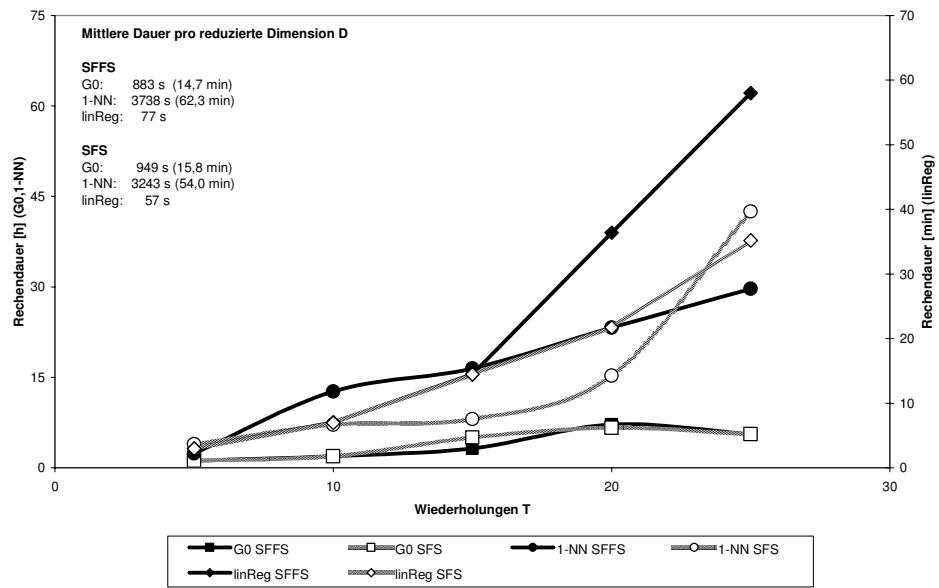
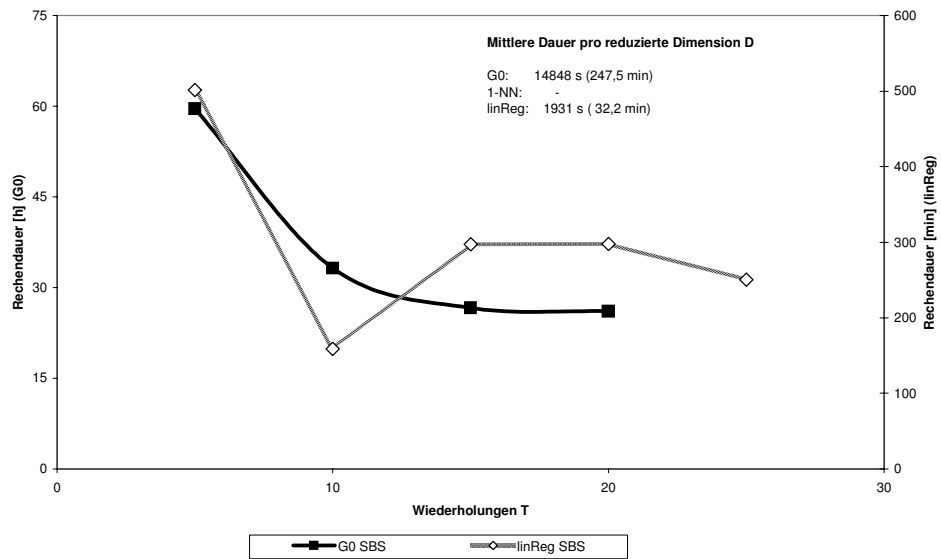


Abbildung 5-12: Rechendauer D5 bis D25 für SBS, T40



Um nun ein optimales Selektionsergebnis unter Berücksichtigung der Qualität der Merkmalsauswahl sowie der dafür erforderlichen Rechendauer zu garantieren, wird im folgenden mit einem Wert für D von 20 gearbeitet.

5.3 Ermittlung der optimalen Methode zur Merkmalsselektion

Bei der Durchführung der Parameterauswahl im Rahmen der Voruntersuchungen zeigte sich, dass die unterschiedlichen Verfahren zur Merkmalsselektion neben erheblichen Differenzen hinsichtlich des Rechenaufwandes auch qualitativ andere Resultate liefern. Aus diesem Grunde soll in diesem Abschnitt eine Auswahl getroffen werden, welches bzw. welche der vorgestellten Selektionsverfahren des Wrapper2-Ansatzes in dieser Arbeit verworfen bzw. zur Materialklassifikation genutzt werden sollen.

Ausgangspunkt dieser Betrachtungen sind die Einstellungen bei einer Durchlaufanzahl T40 sowie einem reduzierten Merkmalsset bei der Selektion der Dimension 20. Bei der Analyse des 1-NN-Klassifikators wird jedoch auf D25 zurückgegriffen, da die D20 Berechnung aufgrund des zeitlichen Aufwandes nicht durchgeführt wurde.

Abbildung 5-13 zeigt für D20 und T40 den Verlauf des Fehlerkriteriums für die ausgewählte Lagerstättensituation im Bereich des Liegenden nördlich des Liegendrückens. Mit wachsendem reduziertem Merkmalsset sinkt das Fehlerkriterium der Materialklassifikation für alle drei Selektionsverfahren. Dabei liegt dieser für die beiden „Bottom Up“-Verfahren unter 30% gegenüber SBS mit einem Fehlerkriterium größer als 30%. SFFS und SFS nähern sich bei der Verwendung eines größeren Merkmalssets zur Zuordnung der Testdaten zu den vier Klassen stark an und sind für $d > 13$ sogar deckungsgleich.

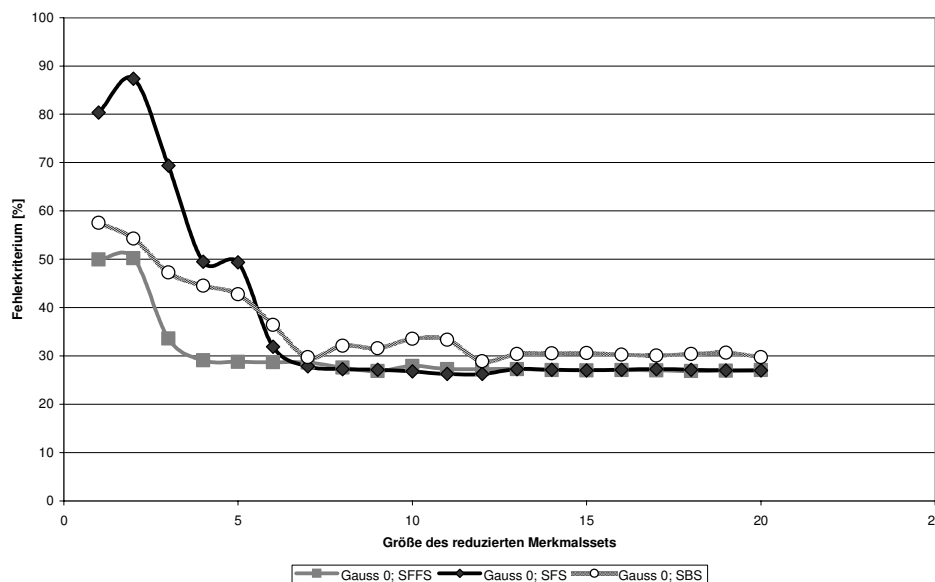
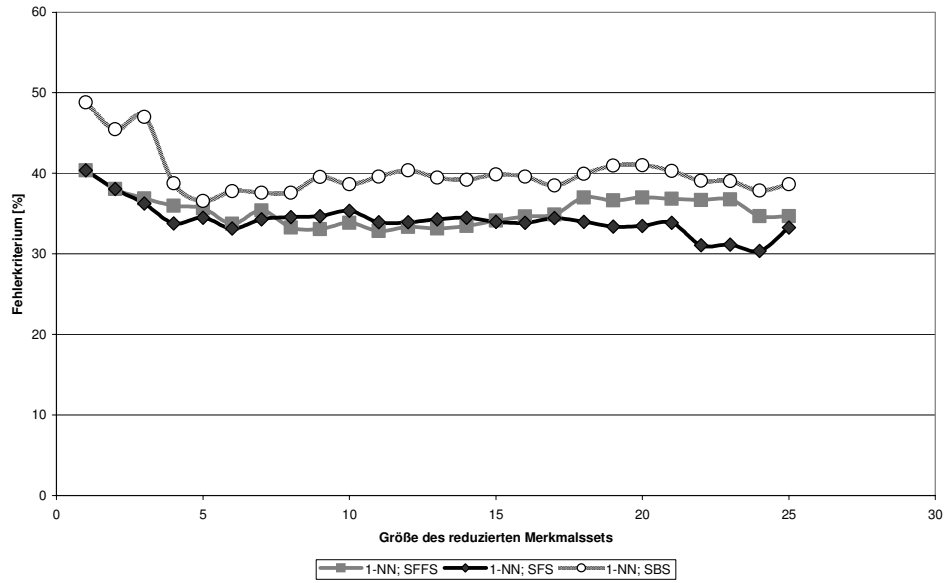


Abbildung 5-13: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, Gaußverfahren 0, T40

Diese Parallelität ergibt sich auch bei der Verwendung des 1-NN-Verfahrens als Klassifikator. In Abbildung 5-14 verläuft dabei die Kurve des Fehlerkriteriums für das SBS oberhalb der Forward-Verfahren. In diesem Fall divergieren deren Kurven jedoch ab $d=17$. Mit SFS wird unter den gegebenen Berechnungsparametern ein geringeres Fehlerkriterium mit wachsendem reduziertem Merkmalsset in der Klassifikation erreicht.

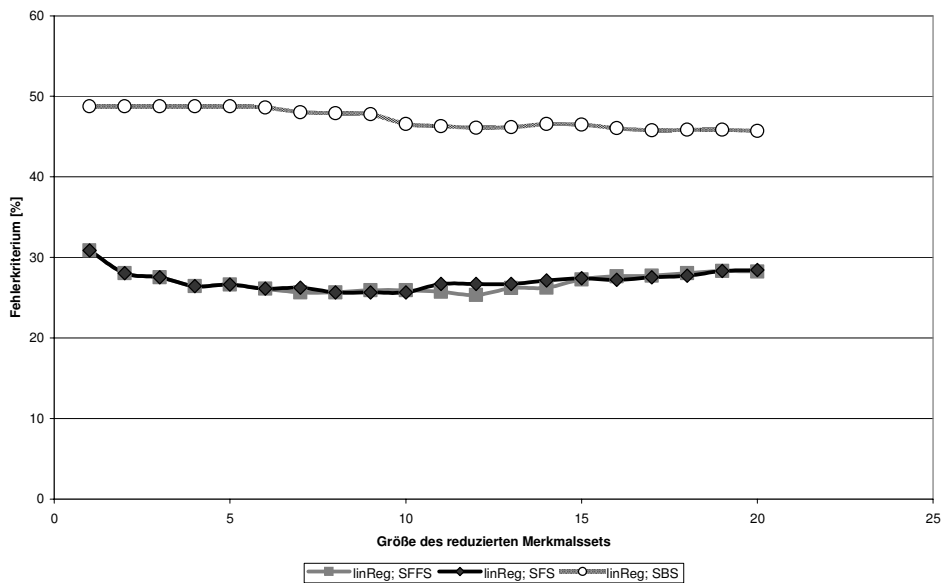
5 Parametereinstellungen zur Merkmalsselektion

Abbildung 5-14: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, 1-NN-Verfahren, T40



Beim linearen Regressionsverfahren zur Reduktion der Fluoreszenzmerkmale aus dem Trainingsdatenset erzielen die Forward-Verfahren erheblich bessere Ergebnisse als mit SBS erreicht werden können. Die Zuordnungsgenauigkeit bei SBS beträgt knapp 50%. Mit steigendem reduziertem Merkmalsset verschlechtert sich das Klassifikationsergebnis für die beiden Forward-Verfahren von einem Minimum bei 25% (Abbildung 5-15).

Abbildung 5-15: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, lineares Regressionsverfahren, T40



Das Ergebnis der Voruntersuchungen zur Bestimmung der Parametereinstellungen bei der Merkmalsselektion ist die alleinige Nutzung des Verfahrens der „Sequential Forward Floating Selection“. Trotz des geringfügig höheren Rechenaufwandes wird dieses Verfahren dem SFS vorgezogen, da aufgrund der Möglichkeit des konditionalen Ausschlusses der „Nesting effect“ verhindert wird. Das SBS kommt nicht zum Zuge, da neben der massiv erhöhten Rechenzeit zudem die Qualität der Klassifikation nicht der der Forward-Verfahren entspricht.

6 Durchführung der Merkmalsselektion und Materialklassifikation

Die Durchführung der Merkmalsselektion und Materialklassifikation für die Braunkohlelagerstätte Hambach erfolgt über den klassischen Weg eines Systems zur Mustererkennung, wie er in Abbildung 4-1 bereits vorgestellt wurde. Nach der Vorverarbeitung der Rohdaten zur Filterung des Rauschens schließen sich die Merkmalsselektion mittels Wrapper2-Ansatz und das Training des Klassifikators an. Die Bewertung des Klassifikators schließt den Vorgang ab, dabei stellt die Fehlerrate im vorliegenden Fall das wesentliche Entscheidungskriterium für die Beurteilung eines Verfahrens dar.

Da jedoch neben der einfachen Betrachtung der Fehlerrate auch die Praxistauglichkeit des Systems zur Materialerkennung berücksichtigt wurde, findet eine mehrstufige Modifikation und Optimierung des aus der Merkmalsselektion hervorgegangenen Merkmalssets statt. Ein System zur Grenzschiefererkennung sollte so z.B. für den Praxiseinsatz in der Braunkohle mit einer möglichst geringen Anzahl von Detektoreinheiten ausgestattet sein, die eine zuverlässige Materialklassifikation in allen relevanten Lagerstättenpartien gewährleistet. Ein optimales Merkmalsset vereint somit eine geringe Fehlerrate mit einer geringen Anzahl an erforderlichen PMT's. Das Motto „Keep it as simple as possible“ ist also die Leitlinie der Merkmalsselektion und Materialklassifikation, die mit einer möglichst geringen Anzahl von universell verwendbaren Filterradeneinstellungen erreicht werden sollte.

Insgesamt wurden für die drei untersuchten Lagerstättenabschnitte (Liegendes nördlich des Liegendrückens, Liegendes Tiefstufe sowie das Zwischenmittel Frimmersdorf a/b) 86 Merkmalsselektionen mit dem SFFS - Verfahren durchgeführt. Aufgrund des erheblichen Rechenaufwandes wurde für den Bereich des Zwischenmittels die Durchlaufanzahl von $T=40$ auf $T=20$ reduziert. Die Ergebnisse aus Kapitel 5.1 zeigen, dass sich durch diese Maßnahme die Repräsentativität der Ergebnisse nicht wesentlich verschlechtert. Die Gesamtrechendauer betrug 11.725610 s, was ca. 136 Tagen bzw. etwa 19,5 Wochen entspricht. Zeitweilig wurden bis zu sechs Rechner parallel eingesetzt.

Die Übersicht in Tabelle 6-1 zeigt die Selektionsverfahren, die zur Anwendung kamen. Bei drei Untersuchungsbereichen wurde auf das 5-NN bzw. bei Zweien auf 2-NN verzichtet: Aufgrund erster Ergebnisse wurde deutlich, dass die erforderliche Rechendauer in keinem Zusammenhang zum möglichen Erkenntnisgewinn steht, da keine fundamentale Verbesserung der Ergebnisse zu erwarten war. Zudem liegt der Fokus dieser Arbeit auf einer generellen Betrachtung des Potentials von Klassifikationsverfahren zur Materialerkennung und soll Tendenzen und Wege für eine pragmatische Lösung des Problems der Trennflächenerkennung aufzeigen.

6 Durchführung der Merkmalsselektion und Materialklassifikation

Bei der Durchführung der Klassifikation wurde stets eine Blindklassifikation durchgeführt. Bei zwei vorliegenden Datensätzen je Messung ergibt sich dabei der reduzierte Merkmalsvektor jeweils durch die Nutzung der Messung 1 (#1) als Trainingsset für den Wrapper2, das als Grundlage der Klassifikation des Testsets der zweiten Messung (#2) herangezogen wird. Umgekehrt erfolgt die zweite Blindklassifikation mit #2 als Trainingsset und #1 als Testset (#2 - #1).

Tabelle 6-1:
Übersicht der durchgeführten Merkmalsselektionen und Materialklassifikationen im Tagebau Hambach

	G0	G1	G2	1-NN	2-NN	5-NN	Lineare Regression
Liegendes nördlich Liegendrücken	X	X	X	X	X	X	X
Liegendes Tiefstufe	X	X	X	X	X		X
Zwischenmittel Laser 1 (266 nm)	X	X	X	X			X
Zwischenmittel Laser 2 (355 nm)	X	X	X	X			X

Die Rohdaten wurden wie schon bei den Voruntersuchungen mit dem Modul „Datenmodulation“ des auf Microsoft ACCESS™ basierenden Mustererkennungsprogramms vorverarbeitet [FUC, 2003]. Die Glättung der Rohdaten wurde mit einer Blendenweite von 20 Messwerten sowie einer Beschränkung der oberen Grenze auf das 5fache des gleitenden Mittelwerts bzw. der unteren Grenze auf das 0,2fache durchgeführt (20 0.2 5). Die Analyse der Rohdaten zeigte, dass mit diesen Einstellungen Ausreißer mit extremen Werten zuverlässig korrigiert werden, ohne die Aussagekraft der Messungen aufgrund zu starker Mittlung zu verringern. Mit dieser Vorgehensweise können kurzzeitige Ausfälle der PMT's („0“-Werte bzw. negative Messergebnisse) aus der Datenbasis entfernt werden. Auch Ausreißer nach oben, wie etwa durch Einlagerungen von anorganischem Material (Pyrit, Kalk, Siderit etc.) hervorgerufen, werden so eliminiert, da diese die Inhomogenitäten in der bereits in sich nicht homogenen Lagerstätte noch verstärken.

Dass diese Glättung keine Reduktion des Informationsinhaltes der Messungen zur Folge hat, lässt sich z.B. am Fall der Messungen mit Laser 2 am Zwischenmittel zeigen (vergleiche Abbildung 3-12). Die maximalen Sprünge in der Intensität der induzierten Fluoreszenz treten hier bei den Übergängen zur Holzlage sowie beim Wechsel von der GK mit Stubben zur GK ohne Holzeinlagerungen im Hangendbereich der Messungen auf. Die maximalen Wechsel in der Intensität liegen hier bei 30.000 zu 150.000 rel. Einheiten bzw. 65.000 zu 20.000 rel. Einheiten. Die materialspezifischen Charakteristika werden durch die Glättung also nicht beeinflusst, vielmehr werden lediglich verzerrende Extremwerte entfernt.

Darüber hinaus fand vor der eigentlichen Durchführung der Merkmalsreduktion die Eliminierung einiger theoretisch zur Verfügung stehender Datensätze statt. Aufgrund der in Kapitel 3.2 bereits beschriebenen Fehlerquellen sind nicht alle Messreihen für die Materialklassifikation geeignet, da

sie den Verlauf der Materialwechsel nicht zuverlässig darstellen. Zudem konnten einige Datensätze nicht verwendet werden, da einige Bereiche des Fluoreszenzspektrums lediglich bei Messung #1, aber aus unterschiedlichen Gründen nicht bei Messung #2 mit den PMT's erfasst wurden oder umgekehrt. Wo Wellenlängenbereiche doppelt gemessen wurden, fand eine winkelabhängige Mittelwertbildung statt. Die für die Merkmalsreduktion und Materialklassifikation herangezogenen Filtereinstellungen sind im Anhang dargestellt.

In diesem Kapitel wird die Durchführung der Merkmalsselektion und Materialerkennung exemplarisch am Beispiel des Liegenden nördlich der Tiefstufe erläutert. Die scharfen Materialübergänge sowie die Beschränkung auf vier relativ homogene Materialien sprechen für diese Wahl. Die originäre Merkmalsselektion wurde für den Bereich des Liegenden nördlich der Tiefstufe mit sieben Algorithmen durchgeführt.

Bei der Durchführung zeigte sich nach der originären Materialklassifikation mit den reduzierten Merkmalssets, dass die Aufgabe in keiner Hinsicht unter der Vorgabe der Verwendung einer minimalen Merkmalsanzahl, die für beide Scannvorgänge gültig sind, zufrieden stellend gelöst werden kann. Wie im Folgenden ausgeführt wird, lässt sich dieses Ziel nur mit der Optimierung der reduzierten Merkmalssets erreichen. Eine weitere Verbesserung der Resultate kann durch die Verwendung eines selektierenden gewichteten Mittelwerts (SGM) erzielt werden, der geologisch nicht sinnvolle Materialzuordnungen in den Profilen der Klassifikation glättet.

6.1 Originäre Merkmalsselektion und Materialklassifikation

Bei der originären Merkmalsselektion wurden mittels des Wrapper2 aus den 135 Merkmalen des Gesamtsets nördlich des Liegendrückens die für eine Klassifikation am besten geeigneten ausgewählt. Bei der anschließenden Materialklassifikation wurden dann jeweils bei der Materialzuordnung die besten Ergebnisse in Form des minimalen Fehlerkriteriums erzielt, wenn auch hier der Algorithmus der Merkmalsselektion verwendet wurde. Dies bedeutet, dass bei einer Selektion des optimalen Merkmalssets mit dem 1-NN-Verfahren dieses auch bei der Materialklassifikation den besten Ansatz für eine erfolgreiche Zuordnung der Messwerte zu einer Materialklasse darstellt.

Im Folgenden soll die Merkmalsselektion exemplarisch für den 5-NN-Algorithmus beschrieben werden, um damit die Vorgehensweise so übersichtlich wie möglich darzustellen.

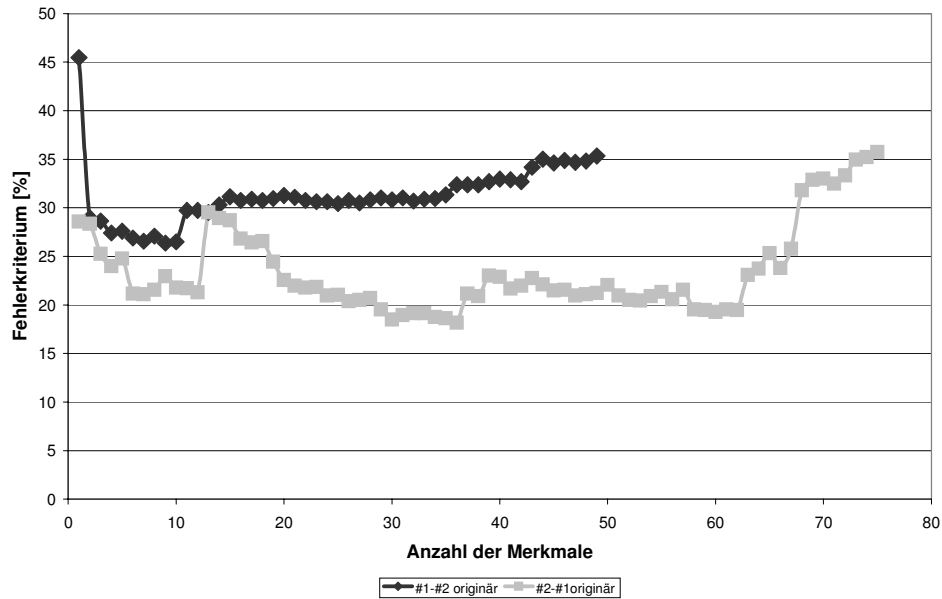
Als Maß für den Erfolg der Materialklassifikation wird das Fehlerkriterium herangezogen. Abbildung 6-1 zeigt die Abhängigkeit der korrekten Zuordnung der Messwerte in der Blindüberprüfung von der Anzahl der Merkmale im reduzierten Merkmalsset, dem Ergebnis der Merkmalsselektion. Generell verbessert sich zunächst die Materialklassifikation mit steigender

6 Durchführung der Merkmalsselektion und Materialklassifikation

Abbildung 6-1:
Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, 5-NN, T40, nördlich Liegendrücken

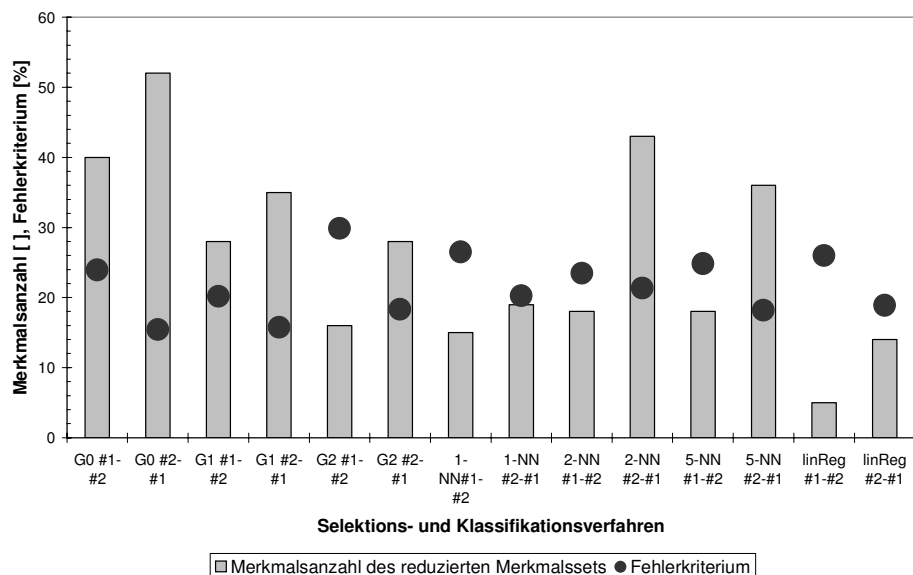
Merkmalsanzahl, verschlechtert sich dann jedoch ab einer gewissen Größe des Merkmalssets zunehmend.

Auffällig ist, dass bei den beiden Blindklassifikationen eine unterschiedliche Anzahl von Merkmalen aus der ursprünglichen Gesamtmerkmalsanzahl von 135 selektiert wurde. Dies weist auf inhomogene Datensätze der Trainingssets hin. Bei den unterschiedlichen Messungen werden also offensichtlich gleiche Emissionswellenlängen nicht durch gleiche Charakteristika definiert, so dass bei den zwei Blindselektionen gleiche Merkmale entsprechend nicht gleich gut für die Zuordnung geeignet sind. Auf diese Thematik wird im folgenden Unterkapitel näher eingegangen.



Darüber weisen beide Klassifikationen abweichende Ergebnisse beim minimalen Fehlerkriterium sowohl hinsichtlich des absoluten Wertes für die Güte der Zuordnung als auch bei der Anzahl der dabei verwendeten Merkmale auf. Abbildung 6-2 verdeutlicht dies für die Messungen im Bereich des Liegenden nördlich des Liegendrückens.

Abbildung 6-2:
Minimales Fehlerkriterium und korrelierende Größe des reduzierten Merkmalssets für #1-#2 und #2-#1, T40, nördlich Liegendrücken



Aufgrund der Abweichungen in der Größe des optimalen reduzierten Merkmalssets für die beiden Blindklassifikationen bedarf es individueller Anpassungen der für die Klassifikation herausgezogenen Merkmalsvektoren. Dieses Vorgehen wird im folgenden Unterkapitel detaillierter nachgezeichnet.

6 Durchführung der Merkmalsselektion und Materialklassifikation

6.2 Analyse der Überkreuzkorrelation

Zur Überprüfung der im vorhergehenden Abschnitt beobachteten Unstimmigkeiten zwischen den Ergebnissen der beiden Blindselektionen und -klassifikationen wird das reduzierte Merkmalsset aus #1-#2 für die Materialklassifikation von #2-#1 herangezogen. Mit diesem Verfahren wird die Überkreuzkorrelation überprüft, d.h. ob die aus dem Trainingsset #1 gewonnenen Merkmale allgemeingültig für beide Messreihen an diesem Profil einzusetzen sind. Beispielhaft lässt sich in Abbildung 6-3 erkennen, dass die als „gg“ bezeichnete Klassifikation über Kreuz schlechtere Zuordnungen liefert.

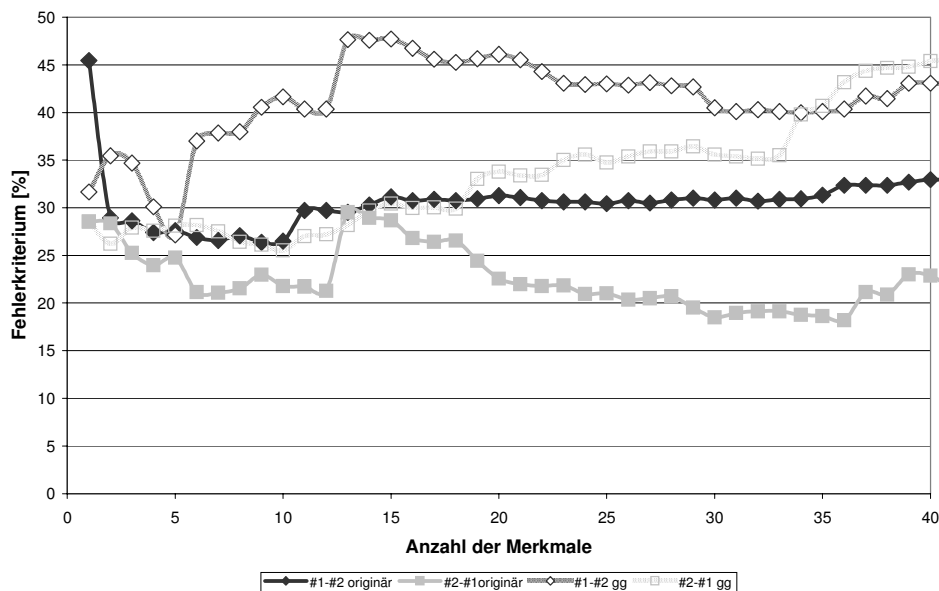


Abbildung 6-3: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, 5-NN, T40, nördlich Liegendrücken, Überkreuzkorrelation

Da auch die minimalen Fehlerraten nicht bei übereinstimmender Größe des reduzierten Merkmalssets erreicht werden und generell die Anzahl der Merkmale bei optimaler Klassifikation für eine praktische Nutzung nicht geeignet ist, bedarf es einer Optimierung des reduzierten Merkmalssets.

Gründe für diese unzureichende Korrelation der beiden Blindklassifikationen und der Nichtübereinstimmung der Merkmale in den optimalen reduzierten Merkmalssets sind in den Basisdaten zu suchen. Um jedoch auch bei wechselnden Daten eine Klassifikation zu erreichen, muss versucht werden, mit einer möglichst geringen Anzahl von allgemeingültigen Merkmalen ein optimales Fehlerkriterium für beide Blindklassifikationen zu realisieren.

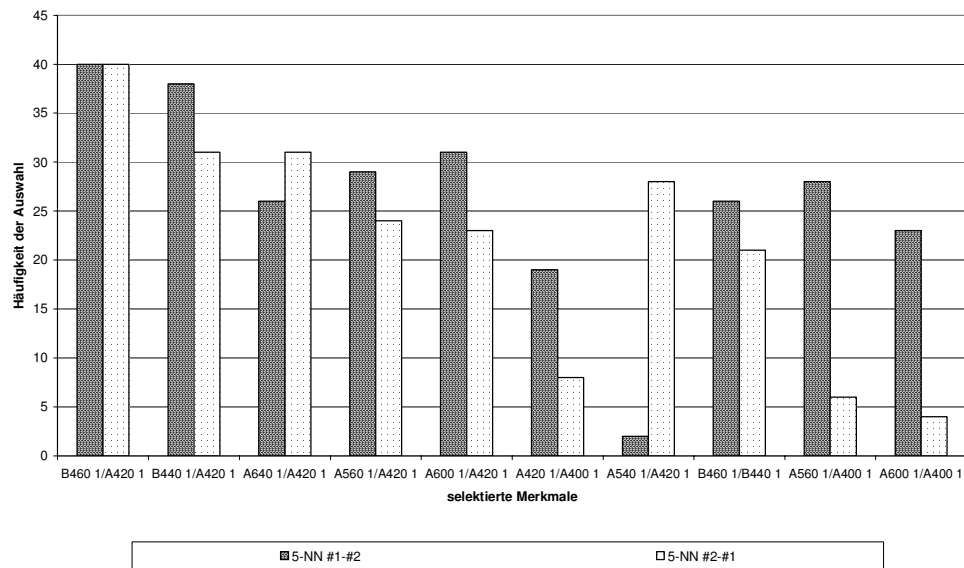
6 Durchführung der Merkmalsselektion und Materialklassifikation

6.3 Optimierung des reduzierten Merkmalssets

Die Diskrepanzen zwischen den reduzierten Merkmalssets und der Materialklassifikation für die beiden durchgeführten Untersuchungen machen es erforderlich, die Einbeziehung sowie Rangfolge der einzelnen Merkmale in das reduzierte Merkmalsset gesondert zu analysieren. Das Fehlerkriterium als alleiniges Maß ist somit nicht hinreichend für die Materialklassifikation.

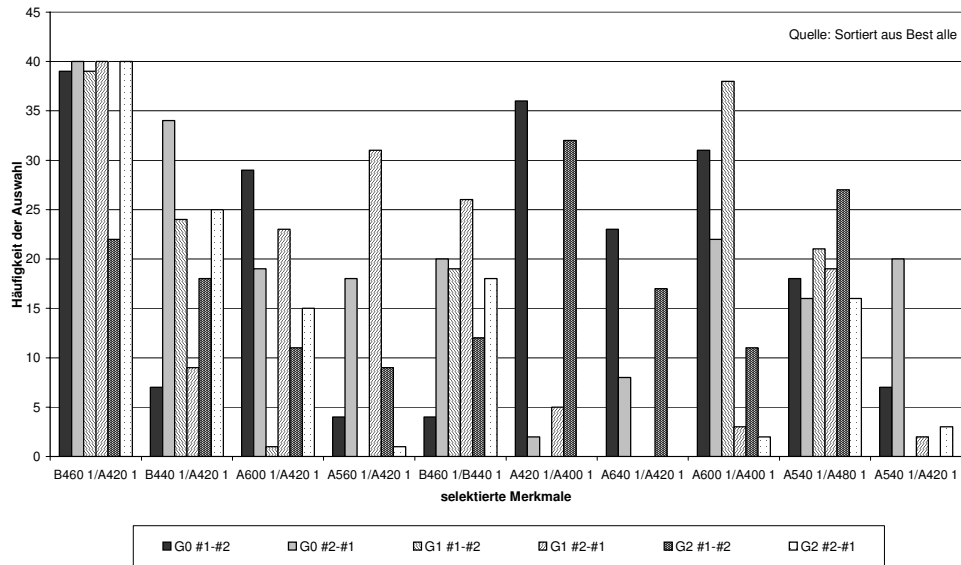
In der nachfolgenden Abbildung 6-4 werden für beide Selektionsvorgänge mit dem 5-NN-Algorithmus die zehn häufigsten ausgewählten Merkmale gegenübergestellt. 40 Nennungen bedeuten bei einer Durchlaufanzahl von T40, dass sich das entsprechende Merkmal bei allen Durchläufen als ein für die Klassifikation geeignetes erwiesen hat.

Abbildung 6-4: Häufigkeit der Wahl einzelner Merkmale, Top10 für #1-#2 sowie #2-#1 5-NN, T40, nördlich Liegendrücken



Neben Merkmalen, die bei beiden Blindklassifikationen annähernd gleich häufig ausgewählt wurden, wie etwa die Quotienten B460 / A420 und B440 / A420, die also als universelle Merkmale betrachtet werden können, fallen sehr einseitig verteilte Merkmale unter den Top 10 auf. Dies ist z.B. beim Quotienten aus A540 und A420 der Fall. Während dieser Parameter bei der Selektion #1-#2 mit nur zwei Nennungen annähernd irrelevant ist, wurde er bei #2-#1 mit 28 Nennungen in 70% aller Fälle für die Materialklassifikation herangezogen.

Dieses Phänomen ist auch bei den anderen Algorithmen zur Merkmalsselektion zu beobachten, wie Abbildung 6-5 für die drei Gaußverfahren zeigt. Neben Merkmalen, die bei allen drei Verfahren und jeweils entsprechend auch bei beiden Blindklassifikationen verwendet wurden, treten Merkmale auf, die nur bei einem Verfahren oder nur einer der beiden Blindklassifikationen eingesetzt wurden.



6 Durchführung der Merkmalsselektion und Materialklassifikation

Abbildung 6-5: Häufigkeit der Wahl einzelner Merkmale, Top10 alle Algorithmen für #1-#2 sowie #2-#1, Gaußverfahren, T40, nördlich Liegendrücken

Ausgehend von der Gegenüberstellung aller Merkmale für die Merkmalsselektion wurden sowohl für alle Selektionsalgorithmen zusammen als auch separat für die Gauß- und die k-NN-Verfahren sowie für die lineare Regression optimierte Merkmalssets, im Weiteren als „BEST“ bezeichnet, definiert (Abbildung 6-6).

Die Kriterien, nach denen die Zusammenstellung dieser Merkmalssets erfolgte, berücksichtigen neben der absoluten Häufigkeit der Merkmalsnennung die Gleichverteilung zwischen #1-#2 und #2-#1. Bei einer Durchlaufanzahl T40 werden Merkmale, die mehrfach weniger als 15 Mal in Abhängigkeit von der Anzahl der zur Optimierung herangezogenen Algorithmen genannt werden, nicht mehr berücksichtigt. Bei Nutzung aller sieben Algorithmen (14 Blindklassifikationen) im Falle des Liegenden nördlich des Liegendrückens findet ein Merkmal nur beim mehr als achtfachen Überschreiten der Grenze von 15 Nennungen Eingang in das „BEST alle“-Set. Die Grenze liegt bei den Gauß- und k-NN-Verfahren bei zwei sowie bei der linearen Regression bei null. Ausnahmen werden von dieser Regel dann gemacht, wenn eine sehr gleichmäßige Verteilung der Nennungen für alle Teilverfahren vorliegt und es bei keiner Klassifikation zu Nicht-Nennung des entsprechenden Merkmals kommt.

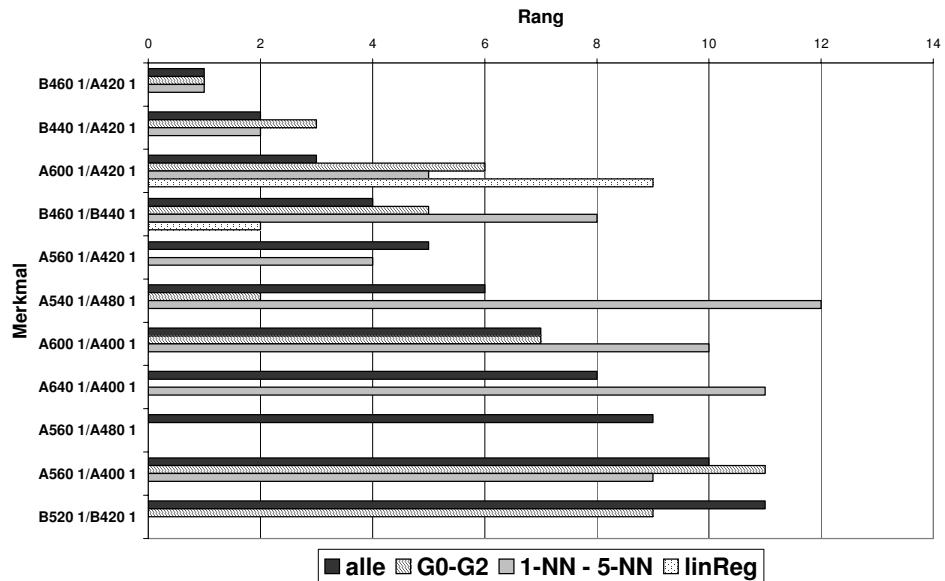
Tabelle 6-2 zeigt einen Ausschnitt aus der Selektionstabelle für das k-NN-Verfahren nach der oben beschriebenen Vorgehensweise.

Merkmal	1-NN		2-NN		5-NN		Sum	# > 15	# = 0	#1 = 0	#2 = 0
	#1-#2	#2-#1	#1-#2	#2-#1	#1-#2	#2-#1					
B460 1/A420 1	89	40	40	40	40	40	240	6	0	0	0
B440 1/A420 1	88	37	5	35	23	38	169	5	0	0	0
A640 1/A420 1	43	26	22	30	29	26	164	6	0	0	0
A560 1/A420 1	41	29	17	34	25	29	158	6	0	0	0
A600 1/A420 1	42	24	22	33	23	31	156	6	0	0	0
A420 1/A400 1	31	36	24	40	26	19	153	5	0	0	0
A540 1/A420 1	40	31	14	37	14	2	126	3	0	0	0
B460 1/B440 1	65	9	10	24	23	26	113	4	0	0	0
A560 1/A400 1	35	21	10	32	8	28	105	3	0	0	0
A600 1/A400 1	36	20	12	30	10	23	99	3	0	0	0

Tabelle 6-2: Ausschnitt der Selektionstabelle zur Ermittlung von „BEST k-NN“, T40, nördlich Liegendrückens

6 Durchführung der Merkmalsselektion und Materialklassifikation

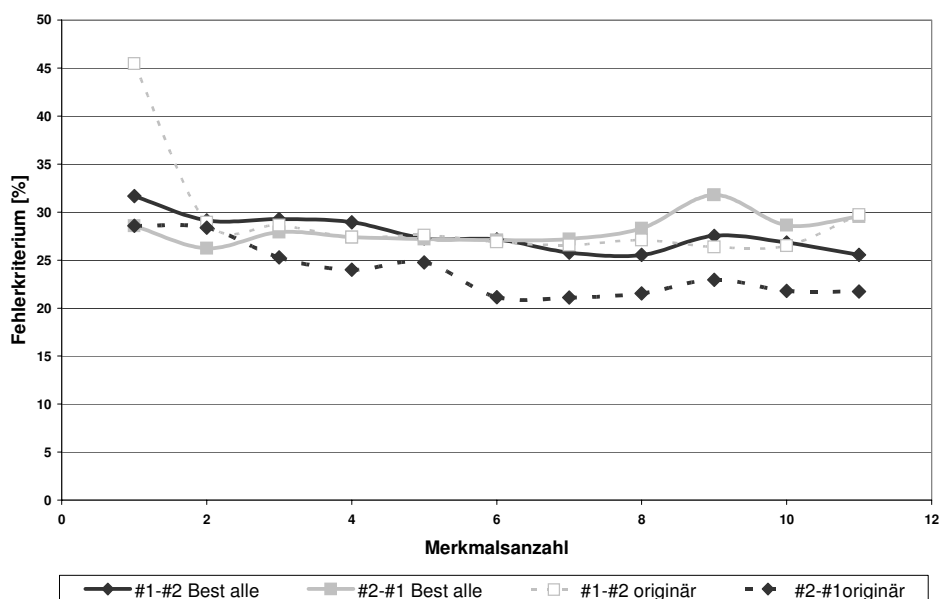
Abbildung 6-6: Darstellung der Rangfolge der optimierten Merkmalssets „BEST“, sortiert nach „BEST alle“, T40, nördlich Liegendrücken



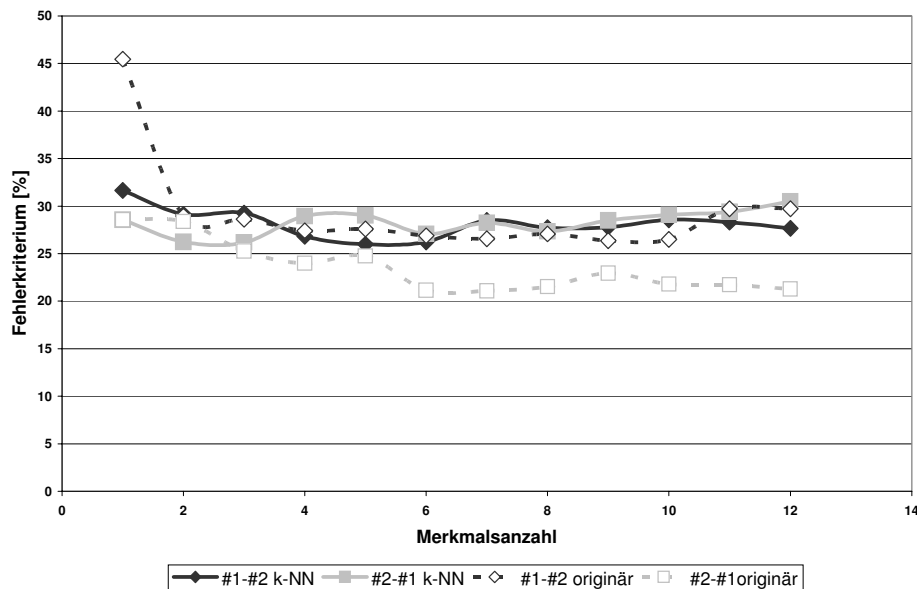
Wie aus der Rangfolge der optimierten Merkmalssets „BEST“ für die unterschiedlichen Algorithmen zu sehen ist, stimmen für die besten fünf Merkmale die Gauß- und k-NN-Verfahren annähernd mit dem „BEST alle“-Set überein. Die lineare Regression arbeitet in den meisten Fällen mit unterschiedlichen Merkmalen, so dass sich hier kaum Übereinstimmungen ergeben.

Wesentlich für die praktische Nutzung der laserinduzierten Fluoreszenz zur Materialerkennung ist neben der zuverlässigen Zuordnung der Materialklassen die Verwendung einer minimalen Anzahl von Merkmalen, die für unterschiedliche Lagerstättenbereiche einsetzbar sind. Die nachfolgenden Abbildungen zeigen vor diesem Hintergrund die Relevanz der Optimierung des reduzierten Merkmalssets (Abbildung 6-7 und Abbildung 6-8). Es werden die Fehlerkriterien sowohl für „BEST alle“ als auch „BEST kNN“ den reduzierten Merkmalssets für beide Blindklassifikationen gegenüber gestellt.

Abbildung 6-7: Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST alle“, 5-NN, T40, nördlich Liegendrücken



Die Graphen laufen für die „BEST“-Merkmalssets mit wenigen Ausnahmen über denen der originären Sets. Es sind jedoch zwei wesentliche Unterschiede ausschlaggebend, die für die Verwendung der optimierten Merkmalsauswahl sprechen. Zum einen sind bei der Verwendung der „BEST“-Merkmale die Differenzen hinsichtlich des Fehlerkriteriums zwischen den Blindklassifikationen erheblich geringer, d.h. es werden unabhängig vom Trainingsset übereinstimmende Materialzuordnungen vorgenommen. Der zweite Vorteil liegt in der Verwendung von gemeinsamen Merkmalen, so dass nur eine einzige Empfangskonstellation verwendet werden kann – eine Voraussetzung für den Einsatz in der Braunkohlegewinnung.



6 Durchführung der Merkmalsselektion und Materialklassifikation

Abbildung 6-8: Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST k-NN“, 5-NN, T40, nördlich Liegendrücken

Neben der Analyse der Klassifikation mittels des Fehlerkriteriums kommt dem resultierenden Profil eine nicht zu vernachlässigende Bedeutung zu. Erst im Kontext mit der Zuordnung des Materials in Abhängigkeit von der Lage am Stoß erlangt die Materialerkennung für mögliche Bediener praktische Relevanz. Eine Warnfunktion kann nur auf dem Profil basieren, das durch den Abgleich der aktuellen Messwerte mit der Trainingsdatenbasis generiert wird. Einen Vergleich zwischen dem Sollprofil, also der geologischen Ansprache vor Ort, und den durch die Verwendung des „BEST“-Merkmalsets erzeugten Profilen erlauben Abbildung 6-9 und Abbildung 6-10.

Unter Berücksichtigung von elf Merkmalen ergeben sich Fehlerkriterien von unter 30%. Die Bereiche der Holzlage und des Liegendton werden beinahe eindeutig identifiziert. Bei Bereichen, die durch die GGMK bzw. GK im Sollprofil gekennzeichnet sind, liegen jedoch Überlappungen vor. Die Zahlenwerte in der Legende stellen das jeweils realisierte Fehlerkriterium dar.

An dieser Stelle bedarf es einer Anmerkung zur Darstellung von Fehlzuordnungen in den Profilauswertungen der Klassifikation. Pro Profil werden bis zu 2.500 Einzelpunkte klassifiziert. Die Größe eines Punktes in der Darstellung ist stark überhöht, so kommt es zu einer scheinbaren optischen Verstärkung der Fehlzuordnungen. Aus diesem Grunde sind die Profile

6 Durchführung der Merkmalsselektion und Materialklassifikation

Abbildung 6-9:
 Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST alle“, 11 Merkmale, 5-NN, T40, nördlich Liegendrücken

stets in Kombination mit den in der Legende eingefügten realisierten Fehlerkriterien zu bewerten.

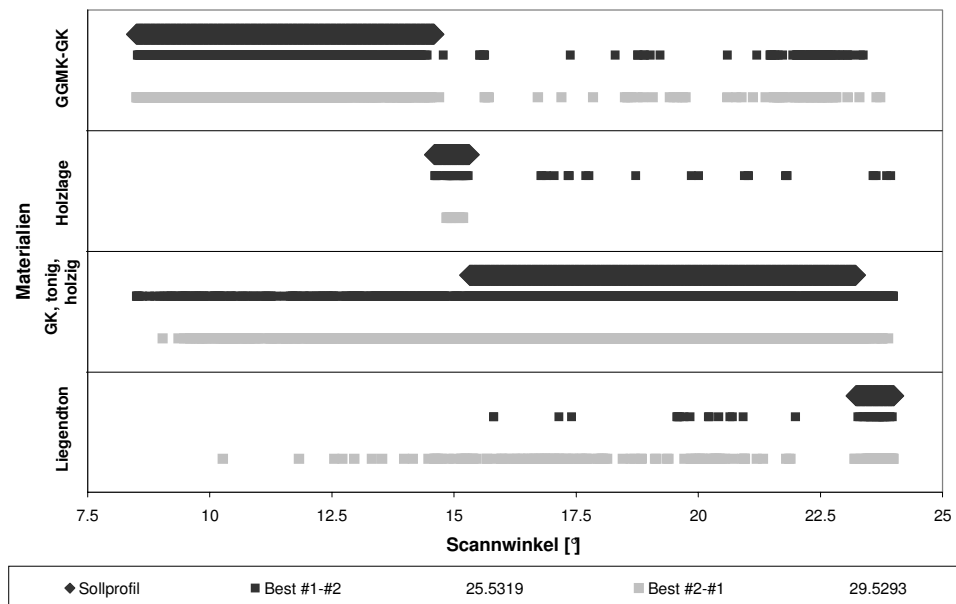
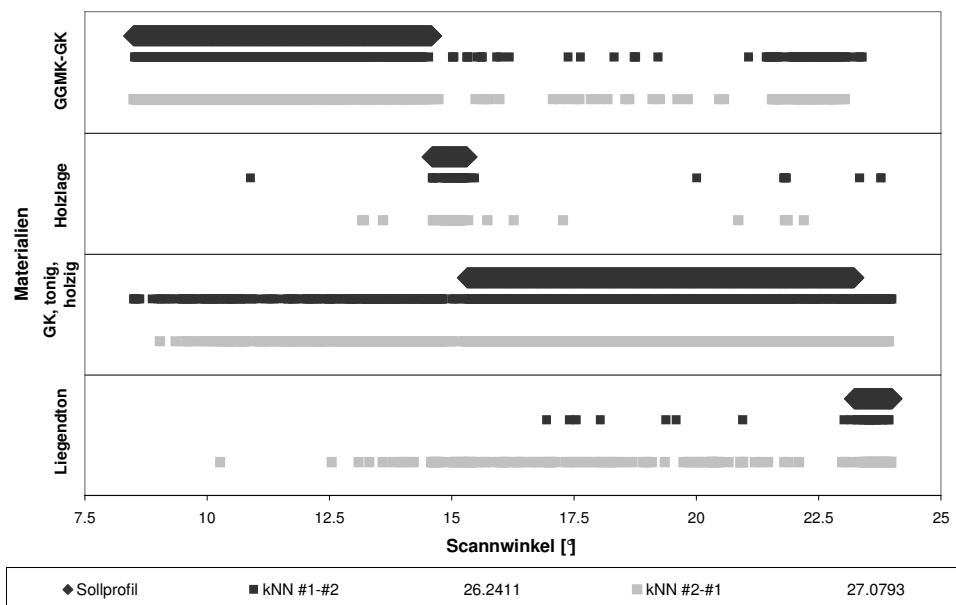


Abbildung 6-10:
 Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST k-NN“, 6 Merkmale, 5-NN, T40, nördlich Liegendrücken



Diese nicht eindeutigen Zuordnungen erschweren einen Einsatz im Tagebau zur Steuerung des Gewinnungsgerätes. Eine Interpretation für dieses Ergebnis ist das Auftreten von sowohl kohligen als auch tonigen Partien in der GK, die dann entsprechend der reineren GGMK bzw. dem Liegendton zugeordnet werden.

Die Darstellung der Profile der Materialklassifikation erscheint aufgrund der hohen Anzahl an Datenpunkten verzerrt. Insgesamt werden über die 15,5° des Stoßes 1.551 Messwerte visualisiert. Trotzdem besteht die Notwendigkeit, diese unter geologischen Gesichtspunkten unrealistischen Ausreißer zu eliminieren, um ein ebeneres Profil zu erhalten. Aus diesem Grund wird im folgenden Unterkapitel eine mögliche Alternative als Lösung dieses Problems vorgestellt.

6.4 Optimierung der Materialklassifikation mittels eines selektierenden gewichteten Mittelwerts (SGM)

Die Optimierung der Materialerkennung stellt eine Nachbearbeitung der eigentlichen Klassifikation dar. Sie überprüft in gewisser Hinsicht die Ergebnisse auf geologische Plausibilität. Dies kann jedoch nur in einem Rahmen erfolgen, der auch die betrieblichen Anforderungen mit einbezieht. Im vollen Eingriff schneidet ein Schaufelradbagger der 240.000er Klasse etwa 150 bis 200 t Material pro Minute. Die Schwenkdauer für einen sichelförmigen Schnitt dauert je nach Blockzuschnitt zwischen zwei bis drei Minuten. Das Signal an den Baggerführer sollte möglichst schnell erfolgen, damit dieser entsprechend reagieren kann. Dies bedeutet, der Baggerfahrer muss neben dem eigentlichen Signal dieses zusätzlich noch mit dem geschnittenen Horizont korrelieren können, um im Sinne der Grenzflächenerkennung spätestens beim darauf folgenden Schwenk den Schneidhorizont anzupassen.

Zur Glättung des aus der Materialklassifikation resultierenden Profils wird in dieser Arbeit der selektierende gewichtete Mittelwert (SGM) herangezogen. Dieser wurde bereits bei einer ersten Mustererkennung durch Fuchs eingesetzt und konnte wesentlich zur Verbesserung der qualitativen Aussage der Klassifikation beitragen [FUC, 2003]. Für das Liegende nördlich der Tiefstufe mit vier Materialien bei der Klassifikation #2-#1 verdeutlicht Tabelle 6-3 das Prinzip dieser Glättung für den Messpunkt 10,27°, der nach Sollprofil im Bereich der GGMK (Materialnummer 0) liegt.

Winkel [°]	Originäres Profil nach der Materialklassifikation	Profil nach SGM	Sollprofil
10,19	0	0	0
10,20	0	0	0
10,21	0	0	0
10,22	0	0	0
10,23	2	0	0
10,24	0	0	0
10,25	2	0	0
10,26	0	0	0
10,27	3	0	0
10,28	2	0	0
10,29	0	0	0
10,30	2	0	0
10,31	0	0	0
10,32	0	0	0
10,33	0	0	0
10,34	2	0	0
10,35	0	0	0
Einfaches Mittel	0,62		
Material 0	11x	65 %	
Material 1	0x	0 %	
Material 2	5x	29 %	
Material 3	1x	6 %	

6 Durchführung der Merkmalsselektion und Materialklassifikation

Tabelle 6-3:
Beispielhafter Einsatz des SGM zur Optimierung der Materialklassifikation, 4 Materialien, Liegend nördlich Tiefstufe

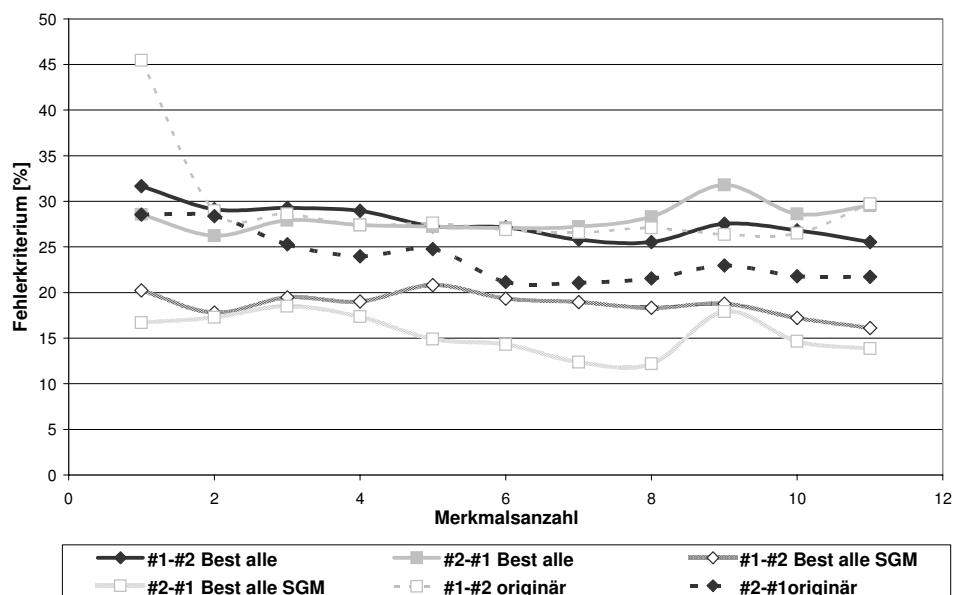
6 Durchführung der Merkmalsselektion und Materialklassifikation

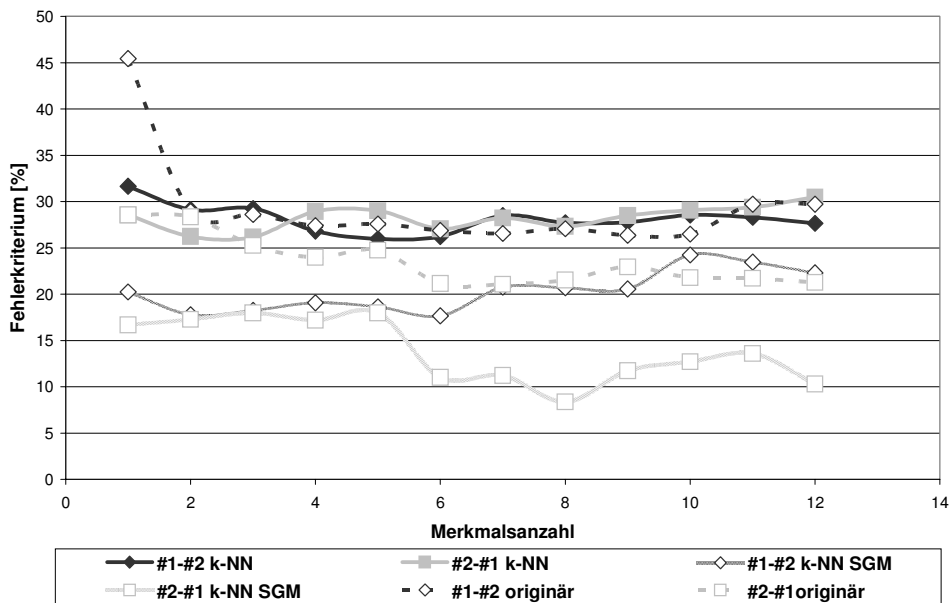
Bei einer Blendenweite von 16 werden die benachbarten acht Werte oberhalb sowie unterhalb des betrachteten Messwertes berücksichtigt. Die Blendenweite wurde so gewählt, dass bei einer Schussfrequenz des Lasers von 16 Hz die Glättung unter Einbeziehung der zeitlich nachfolgenden Messwerte über einen Zeitraum von 500 ms erfolgt. Das Ergebnis der Klassifikation kann also eine halbe Sekunde nach der Messung für Steuerungszwecke genutzt werden, dies ist eine vertretbare Abweichung von einer Echtzeitinformation. Es entspricht unter Vollast einer maximalen Kohlenmenge von unter 2 t.

Die Nutzung des einfachen Mittels führt bei dieser Aufgabenstellung zu keine Lösung. Der Algorithmus des SGM summiert vielmehr die Häufigkeit der innerhalb der Blendenweite auftretenden Materialien und modifiziert beim Überschreiten einer zuvor definierten Grenze die Materialklasse des entsprechenden Wertes bzw. belässt die Materialzuordnung. In dieser Arbeit wird der Entscheidungsfaktor auf 0,6 gesetzt, d.h. eine Änderung der Materialklasse der originären Klassifikation findet nur dann statt, wenn 60% der angrenzenden Messwerte des originären Profils einer abweichenden Klasse zugeordnet wurde. Dieser Wert hat sich als für eine Materialklassifikation am besten geeignet erwiesen [FUC, 2003]. Der Source Code zur SGM in Matlab ist im Anhang zu finden.

Die Optimierung der Materialerkennung spiegelt sich in einer Reduktion des Fehlerkriteriums wieder. Im Falle der Klassifikation mit dem Merkmalsset „BEST alle“ werden durchgehend Fehler unter 20% erreicht (Abbildung 6-11). Im Durchschnitt wird eine Verbesserung gegenüber der ungeglätteten Klassifikation von etwa 10% und gegenüber der originären Mustererkennung um 5% erzielt.

Abbildung 6-11:
Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST alle“ inklusive SGM, 5-NN, T40, nördlich Liegendrücken





6 Durchführung der Merkmalsselektion und Materialklassifikation

Abbildung 6-12: Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST k-NN“ inklusive SGM, 5-NN, T40, nördlich Liegendrücken

Bei der Klassifikation mit dem Merkmalsset der optimierten k-NN-Verfahren ergibt sich ebenfalls eine deutliche Verbesserung des Fehlerkriteriums. Mit steigender Merkmalsanzahl ist jedoch für #1-#2 ein Anstieg der Fehlklassifikationen zu verzeichnen, während bei #2-#1 mit der SGM bei acht Merkmalen eine Fehlerrate von unter 10% realisiert wird.

Die Reduktion der Fehlerrate der Klassifikation rührt im Wesentlichen durch die Modifikation von vereinzelt Fehlzuordnungen her, wie in Tabelle 6-3 beispielhaft demonstriert wurde. Abbildung 6-13 und Abbildung 6-14 stellen die Profile der „BEST“ denen der mittels SGM modifizierten Materialerkennung gegenüber. Besonders in den bereits angesprochenen Bereichen in den kohligten Partien werden mit dem selektierenden gewichteten Mittelwert Fehlansprachen durch die Klassifikation korrigiert.

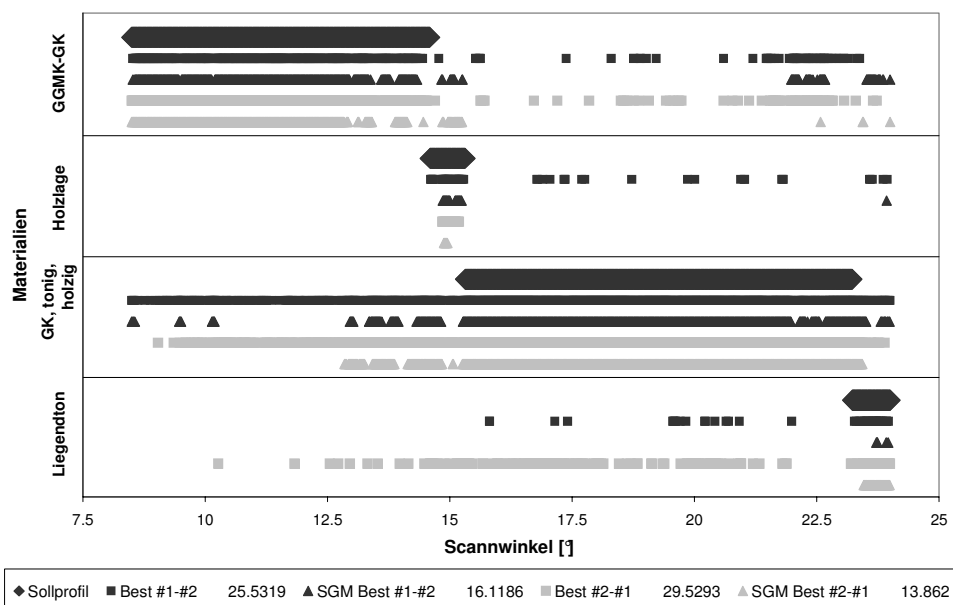
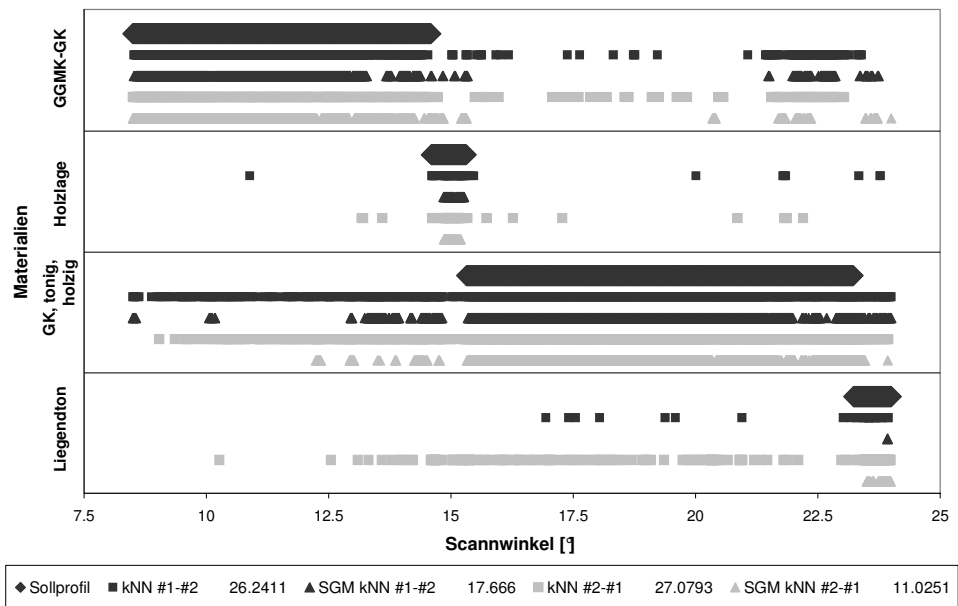


Abbildung 6-13: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST alle“ inklusive SGM, 11 Merkmale, 5-NN, T40, nördlich Liegendrücken

6 Durchführung der Merkmalsselektion und Materialklassifikation

Abbildung 6-14:
 Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST k-NN“ inklusive SGM, 6 Merkmale, 5-NN, T40, nördlich Liegendrücken



Durch die Glättung kommt es jedoch dazu, dass Übergangsbereiche zwischen einzelnen Materialien wie etwa die Holzlage oder der Liegendton aufgrund der Blende erst später erkannt werden. In den Übergangsbereichen erfolgt in der Regel die Klassifikation für die ersten Messpunkte uneinheitlich, da es sich im Falle der untersuchten Braunkohlelagerstätte um gewachsenes Material handelt. Die Übergangsbereiche sind daher durch eine sehr inhomogene Materialzusammensetzung charakterisiert. Die Mittelwertbildung über die durch den Laser bestrahlte Fläche von 10 mm Durchmesser ist hier anscheinend für eine Homogenisierung noch nicht ausreichend. Kleinere Blendenöffnungen des Glättungsalgorithmus führen unter diesen Umständen nicht zu schärferen Materialübergängen bei der Klassifikation.

7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Die Eignung der laserinduzierten Fluoreszenz als möglicher Bestandteil der zukünftigen Qualitätssteuerung bei der Gewinnung von Braunkohle soll in diesem Kapitel durch eine Bewertung der Merkmalsselektion und Materialklassifikation erörtert werden. Aufbauend auf der im vorherigen Kapitel vorgestellten Methodik wird für die drei Lagerstättenabschnitte des Untersuchungsbereichs der Erfolg der Materialerkennung dargestellt und analysiert. Neben der Bedeutung der Anzahl der Materialklassen eines Profils wird der Einfluss der unterschiedlichen Algorithmen zur Merkmalsselektion und Materialklassifikation auf das Zuordnungsergebnis untersucht. Es soll geprüft werden, ob sich eine Methode als optimal erweist bzw. welches Verfahren in Zukunft bei der Klassifikation unter Berücksichtigung des Rechenaufwandes und der Zuverlässigkeit der Zuordnung zu präferieren ist.

Aufgrund der zahlreichen Störungen bei der Durchführung der Messreihen kann ein Vergleich der Vorteilhaftigkeit einer bestimmten Anregungswellenlänge ausschließlich für den Bereich des Zwischenmittels Frimmersdorf a/b erfolgen und auch dies nur eingeschränkt. Im Liegenden wurde das anstehende Material ausschließlich mit Laserlicht der Wellenlänge 266 nm bemustert. Die Frage, ob es Übereinstimmungen zwischen den selektierten Merkmalen der optimalen Merkmalssets für die unterschiedlichen Lagerstättenpartien gibt, wird in diesem Kapitel ebenfalls diskutiert. Zu diesem Zweck wird ein universeller Merkmalsvektor, der Messstellen übergreifend für das Liegende gilt, aus dem Gesamtset extrahiert und auf seine Einsatztauglichkeit hin überprüft.

Bevor im Detail die angesprochenen Teilaspekte diskutiert werden, wird vorgehend auf die Problematik der Materialeinteilung im Bereich des Zwischenmittels hingewiesen. Die Vielzahl an unterschiedlichen Materialien im untersuchten Stoßbereich erschwert neben der nicht einfachen geologischen Ansprache auch die Kombination von Materialien zu gemeinsamen, allgemeingültigen Klassen. Es ist daher damit zu rechnen, dass die Ergebnisse in diesem Bereich der Lagerstätte Hambach nicht an die Qualität der Materialzuordnung am Übergang zum Liegenden heranreichen werden.

7.1 Einfluss der Anzahl an Materialklassen

Die Scheibeneinteilung für den Gewinnungsprozess erfolgt im Tagebau Hambach, wie im Kapitel 3.1.2 erläutert, anhand der geologischen Ansprache und der stofflichen Zusammensetzung des Materials hinsichtlich des Brennwertes und der Verunreinigungen (Aschegehalt, Beimengungen an Kalium, Natrium usw.). Es findet neben der Unterscheidung zwischen bauwürdigem und nichtbauwürdigem Material also auch eine Differenzierung innerhalb der Braunkohle statt. Bei den durchgeführten LIF Messkampagnen wurden die bemusterten Profile in Materialklassen unterteilt, die den geologisch-stratigraphischen Kriterien entsprechen. Neben der Einteilung in eine minimale Klassengröße, die bei allen drei Profilen bei vier Materialien liegt, wurden auch Unterteilungen der Profile in fünf, sechs und neun Klassen vorgenommen.

In diesem Unterkapitel soll der Einfluss der Anzahl an Materialklassen auf die Materialerkennung vorgestellt werden. Dazu erfolgt im Bereich des Liegenden nördlich der Tiefstufe der Vergleich zwischen einem Profil mit vier und neun Klassen. Die Unterschiede in der Merkmalsselektion und die Materialklassifikation für das Zwischenmittel zeigt die Gegenüberstellung der Zuordnungen für vier und sechs Materialklassen. Um die Besonderheiten der beiden Untersuchungsbereiche zu dokumentieren, erfolgt die Darstellung nicht ausschließlich an einem Stoß, sondern für beide Lagerstättenabschnitte. Die Analyse basiert auf der Auswertung der jeweiligen „BEST“-Merkmalssets sowie der entsprechenden Fehlerkriterien und Zuordnungsprofile, die mit dem SGM generiert wurden.

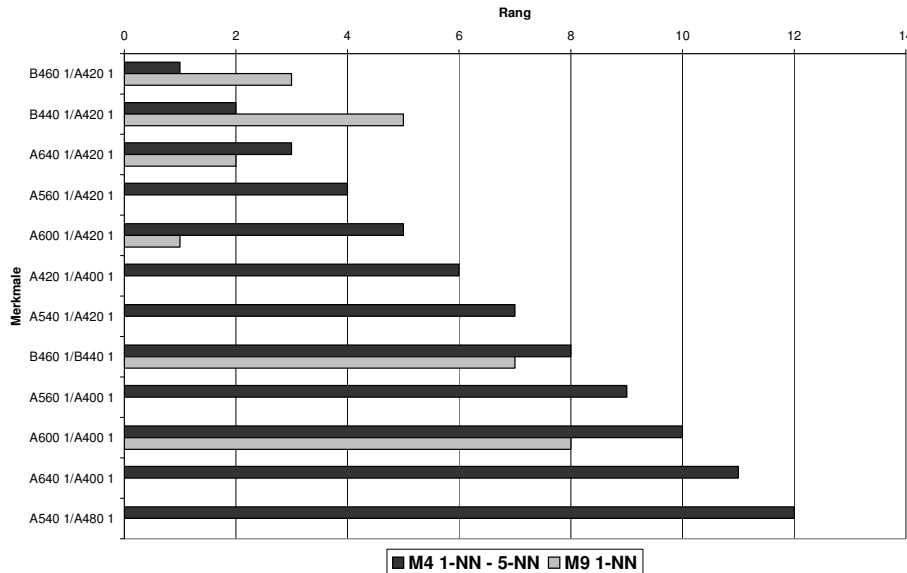
7.1.1 Liegendes, 9 Materialien

Der Abbau erfolgt im Bereich des Liegenden nördlich des Liegendrückens mit dem Aushalten der unterhalb der Holzlage befindlichen ballastreichereren Kohle, die dann separat hereingewonnen wird. Der Liegendton darf dabei unter keinen Umständen mit der tonig holzigen GK abgebaggert werden. Damit entspricht die Einteilung in vier Materialklassen (M4) der derzeitigen Praxis, bei der die Holzlage die optisch erkennbare Grenzschicht zwischen den beiden Kohlesorten darstellt. Die Einteilung in neun Klassen (M9) geht daher über das derzeit angewendete System hinaus und würde bei einer positiven Materialerkennung eine bessere Warnfunktion zur Folge haben. Der Baggerfahrer erhielte bereits vor dem Anschneiden der jeweilig kritischen Schicht detailliertere Informationen über seine aktuelle Position in der Lagerstätte.

Der Vergleich der selektierten optimierten Merkmale der „BEST“-Merkmalssets zeigt, dass abhängig von der Klassenanzahl unterschiedliche Merkmale für die Klassifikation herangezogen werden. Dies veranschaulicht für „BEST k-NN“ Abbildung 7-1.

Die Rangfolge ist nach der Eignung der Merkmale für vier Materialien sortiert. Es lässt sich erkennen, dass von den zehn Merkmalen des „BEST k-NN“ für neun Materialien, sechs auch für die Klassifikation für M4 verwen-

det werden. Entsprechend der Klassenanzahl lassen sich also gewisse Merkmale weiterhin für die Materialerkennung nutzen. Aufgrund der sich mit der Klassenanzahl ändernden klassenspezifischen Merkmalsmittelwerte sind jedoch bestimmte Merkmale nur für eine definierte Klassenanzahl geeignet und damit nicht universell einsetzbar.



7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-1: Vergleich der selektierten „BEST k-NN“ Merkmale für 4 und 9 Materialien, nördlich Liegendrücken

Ohne dem nachfolgenden Unterkapitel vorzugreifen, verdeutlicht die nachfolgende Abbildung 7-2, dass die Zuordnung bei neun Materialien im Durchschnitt einer Verschlechterung im Vergleich zu M4 um 35% bis zu 40% entspricht. Keines der vier „BEST“-Merkmalssets erreicht dabei eine Genauigkeit der Materialzuordnung von besser als 50% wie Abbildung 7-3 zeigt. Bei vier Klassen werden Fehlerkriterien unter 20% realisiert.

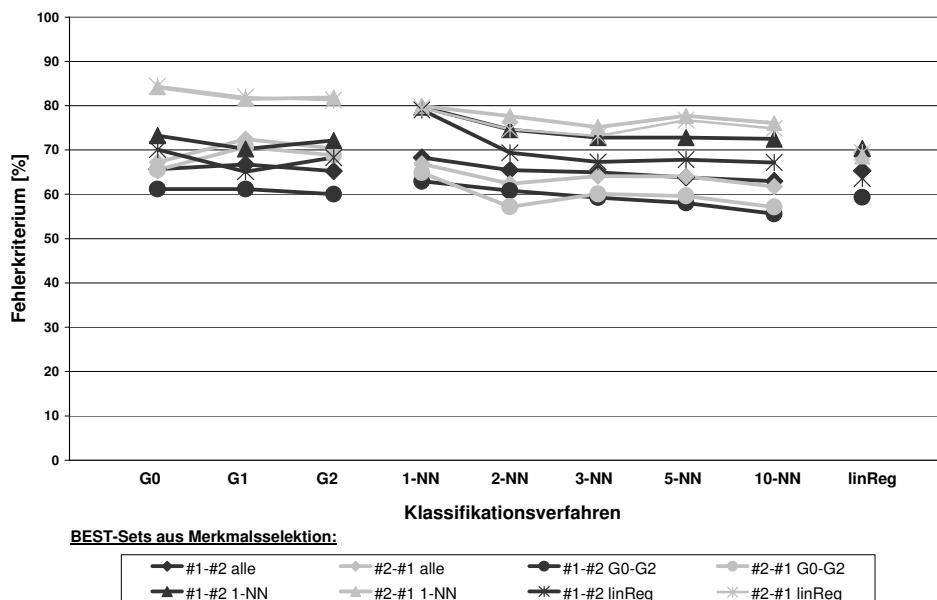
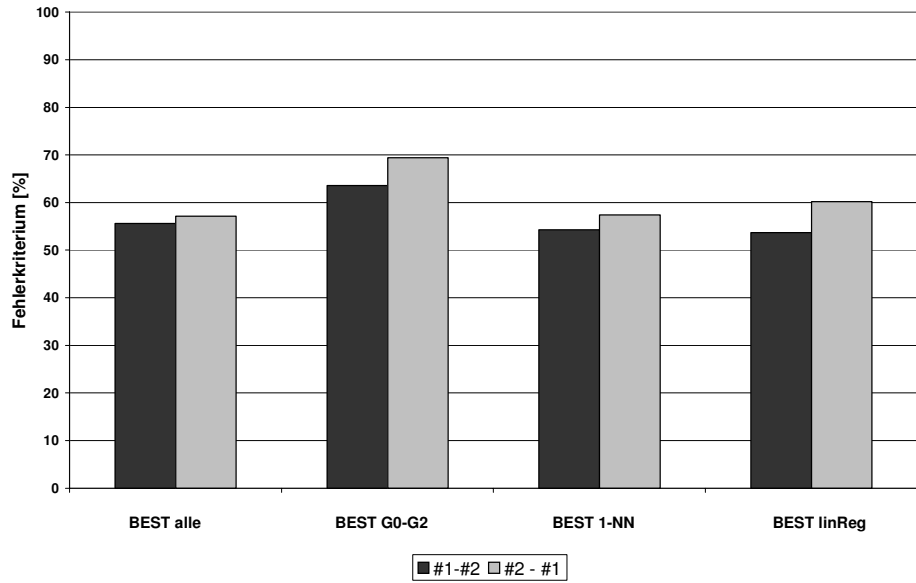


Abbildung 7-2: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 9 Materialien, nördlich Liegendrücken

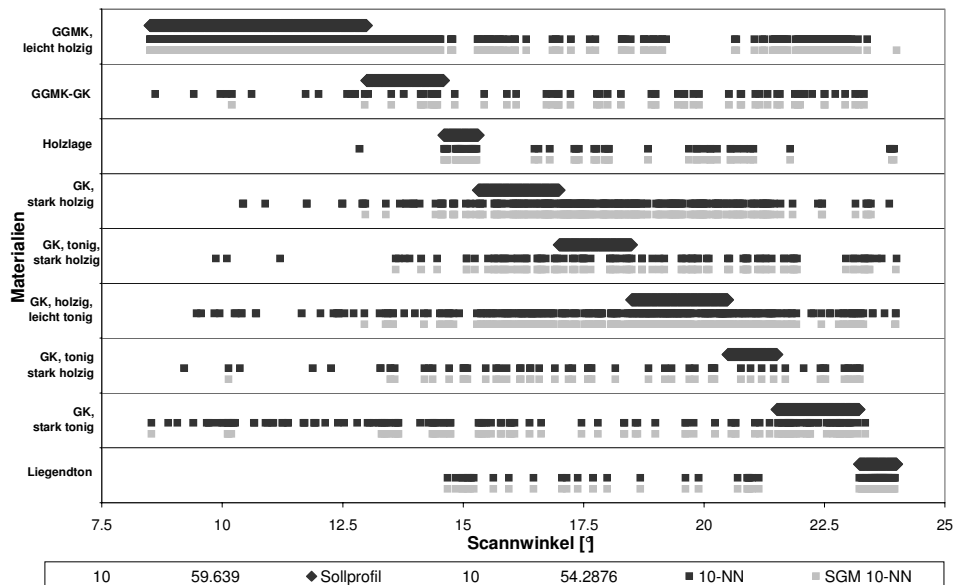
7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-3: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 9 Materialien, nördlich Liegendrücken



Die Ursache für diese mangelhafte Materialerkennung lässt sich aus der nachfolgenden Abbildung 7-4 ableiten. Während die leicht holzige GGМК im oberen Bereich des gescannten Abschnitts und der Liegendton ab 23,2° nach der Glättung mittels SGM beinahe eindeutig durch die Materialklassifikation zugeordnet werden konnten, treten bei weniger homogenen Klassen massive Fehlzuordnungen auf. Die minimalen Fehlerkriterien mit und ohne Nutzung der SGM bei der Klassifikation wurden mit zehn Merkmalen erreicht. Die erforderliche Merkmalsanzahl ist in der Abbildung jeweils dem Fehlerkriterium in der Legende vorangestellt.

Abbildung 7-4: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST 1-NN“ inklusive SGM, 10-NN, 9 Materialien, nördlich Liegendrücken



Die Klasse zwischen der GGМК und der Holzlage wird z.B. im ihr zugewiesenen Bereich fast überhaupt nicht erkannt, dafür aber in tiefer gelegenen Abschnitten. Große Schwierigkeiten bereitet die Klassifikation in den Klassen, die lediglich durch wechselnde Anteile von holzigen und tonigen Beimengungen bei einer GK als Hauptbestandteil gekennzeichnet sind. Hier sind die klassenübergreifenden Differenzen bei den zur Zuordnung heran-

gezogenen Merkmalen so gering, dass eine eindeutige Materialerkennung nicht erfolgt.

Die durch die Glättung mittels SGM erzielte Optimierung des Profils mit M9 fällt im Vergleich zu nur vier Klassen gering aus, da es sich bei den Fehlzugeordnungen nicht um einzelne Ausreißer, sondern in den Bereichen der inhomogenen Kohlen um massive Fehlklassifikationen handelt. Auch die Änderung der Entscheidungsgrenze auf 0,5 führt zu keiner substantiellen Verbesserung des Fehlerkriteriums.

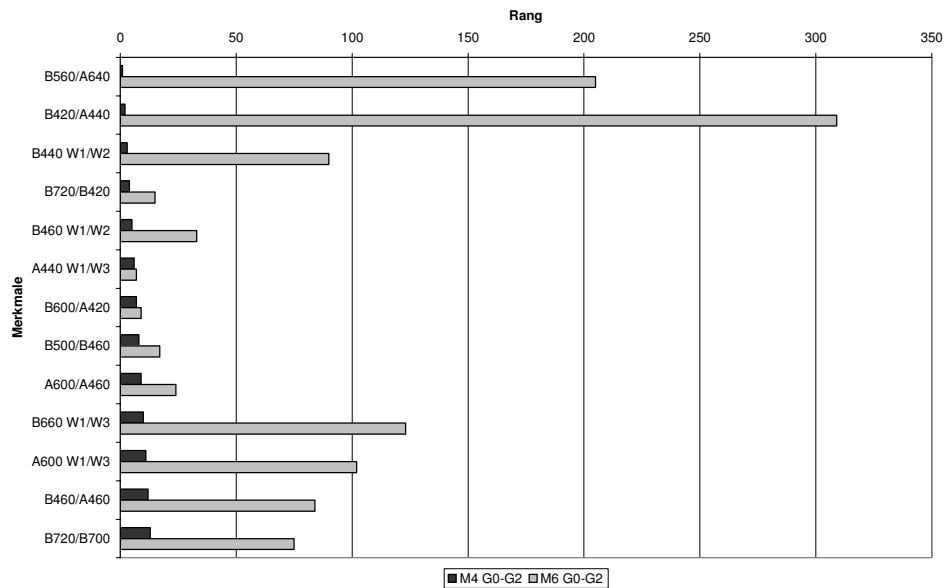
7.1.2 Zwischenmittel, 6 Materialien, 355 nm

Das betrachtete Profil im Bereich des Zwischenmittels Frimmersdorf a/b ist im Gegensatz zum oben beschriebenen Liegenden durch einen häufigen Wechsel unterschiedlicher Materialien gekennzeichnet. Aus diesem Grunde ist eine verallgemeinernde Klasseneinteilung sehr schwierig, da bei der geologischen Ansprache elf Materialklassen unterschieden wurden. Ähnliche Materialien wurden in gemeinsamen Klassen zusammengefasst, so dass in dieser Arbeit neben der Einteilung in vier Materialien ebenfalls eine Merkmalsselektion und Materialklassifikation für sechs Klassen durchgeführt wurde. Dieses Profil unterteilt im Gegensatz zum allgemeinen Fall zusätzlich im oberen Abschnitt und unteren Abschnitt des Scans in holzigere Bereiche und Materialien, die nur leicht holzig sind (siehe auch Anhang B).

Der Vergleich zwischen den optimierten selektierten Merkmalen M4 und M6 in Form der „BEST“-Merkmalssets für eine Anregungswellenlänge von 355 nm zeigt im Gegensatz zum Liegenden der Lagerstätte keinerlei Übereinstimmung. Abbildung 7-5 stellt daher lediglich die Rangfolge für beide Klassenanzahlen basierend auf der nicht bearbeiteten Nennungshäufigkeit sortiert nach M4 dar. Für einige Merkmale ergeben sich ansatzweise Übereinstimmungen der Häufigkeit der Selektion wie etwa bei den Abklingzeiten A 440 nm W1/W3 sowie beim Quotienten B 600 nm / A 420 nm. Jedoch kennzeichnet diesen Quotienten z.B. bei M6 eine extreme Ungleichmäßigkeit der Selektion, da dieser ausschließlich bei der Blindklassifikation #2-#1, aber nicht bei #1-#2 ausgewählt und daher nicht in das „BEST“-Merkmalsset übernommen wurde. Auch bei vermeintlichen anderen Übereinstimmungen treten diese Unregelmäßigkeiten bei der Merkmalsselektion auf.

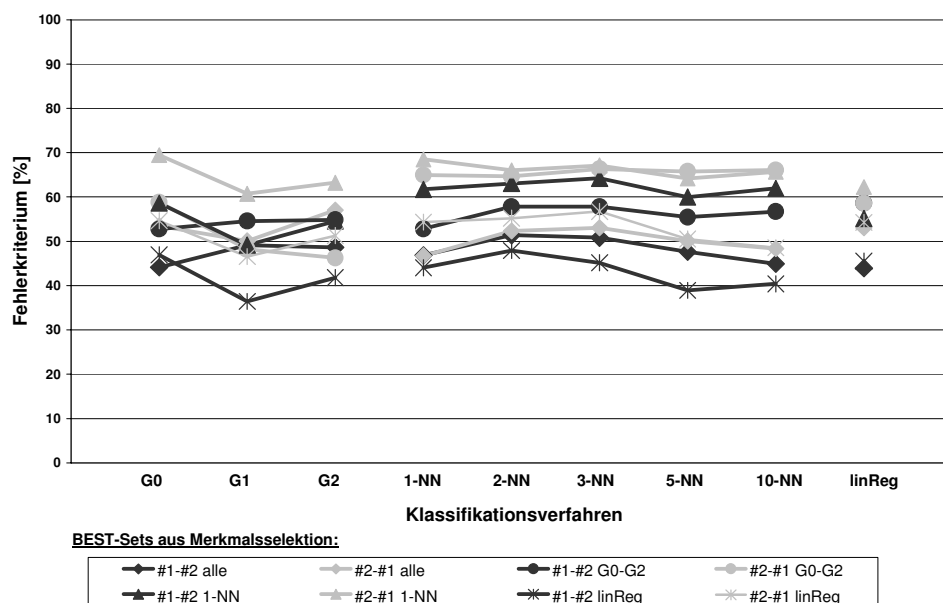
7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-5: Vergleich der Merkmale für 4 und 6 Materialien G0-G2, Zwischenmittel, 355 nm



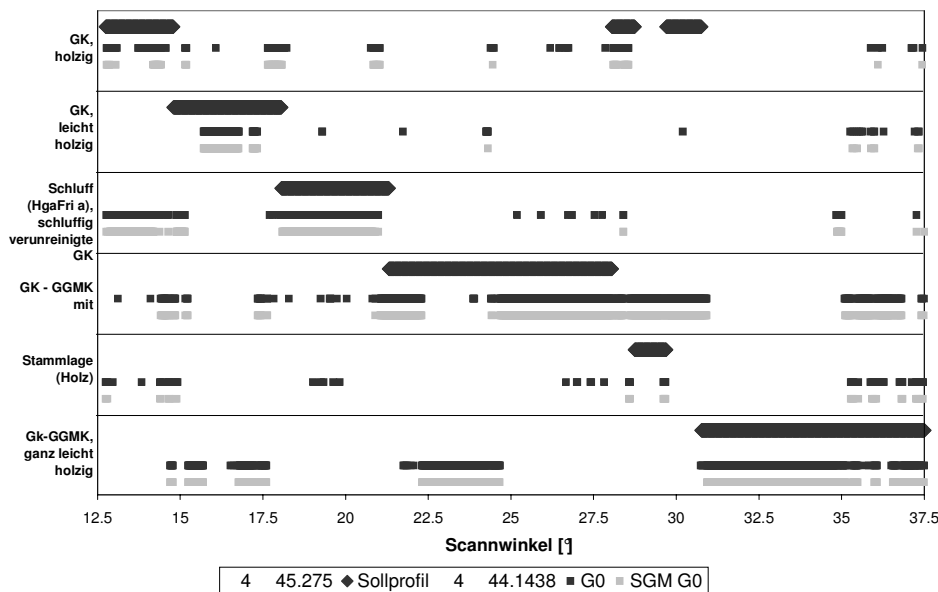
Aufgrund dieser Unterschiede in den resultierenden „BEST“-Sets ließe sich eine Materialklassifikation für M4 und M6 nicht mit denselben PMT-Einstellungen in der Praxis realisieren. Nichtsdestotrotz lässt sich auch für sechs Materialien eine Klassifikation durchführen. Die Fehlerkriterien in Abhängigkeit der bei Selektion und Klassifikation verwendeten Algorithmen zeigt Abbildung 7-6. Auch bei M6 im Zwischenmittel bei 355 nm dominiert kein Verfahren eindeutig, obwohl tendenziell mit der linearen Regression die besseren Ergebnisse erzielt werden. Die mit dem SGM erzielten minimalen Fehlzuordnungen liegen im Falle der linearen Regression bei der Blindklassifikation #1-#2 teilweise unter 40%, in der Regel jedoch zwischen 40% und 50%.

Abbildung 7-6: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 6 Materialien, Zwischenmittel, 355 nm



Das Profil der Materialklassifikation mit dem Merkmalsset „BEST alle“ ist für die einfache Materialerkennung sowie für die Glättung mittels SGM in Abbildung 7-7 gezeigt. Die Holzlage wird nur unzureichend klassifiziert, während die bauwürdige GK des Flöz Frimmersdorf a und das schluffige

Hangende zufrieden stellend zugeordnet werden. Aber wie schon bei M9 im Liegenden bleibt festzustellen, dass die Klassifikation für M6 nicht den betrieblichen Ansprüchen der RWE Power AG entspricht.



7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-7:
Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST alle“ inklusive SGM, G0, 6 Materialien, #1-#2, Zwischenmittel, 355 nm

7.1.3 Fazit

Als Ergebnis der Analyse des Einflusses der Klassenanzahl auf den Erfolg der Materialerkennung lässt sich festhalten, dass eine zu detaillierte Klasseneinteilung nicht zu praktisch nutzbaren Ergebnissen führt, da die zu erreichenden Fehlerquoten im Vergleich zu allgemeiner gehaltenen Materialeinteilungen unbefriedigend sind. Fehlerkriterien unter 40% wurden bei den durchgeführten Messreihen nur in Ausnahmefällen erreicht. Vielmehr wurde im Regelfall nur jeder zweite Messwert korrekt der zugehörigen Materialklasse zugeordnet.

Im Falle des Liegenden war es grundsätzlich möglich, die Klassifikation M9 mit denselben Merkmalen durchzuführen, mit denen auch die Materialerkennung für vier Klassen erfolgt. Aufgrund erheblicher Unterschiede in der Komposition der resultierenden „BEST“-Sets ließe sich jedoch für die Emissionswellenlänge 355 nm im Zwischenmittel eine Materialklassifikation für M4 und M6 in der Praxis nicht mit denselben PMT-Einstellungen realisieren.

Aus diesen Gründen findet in den folgenden Abschnitten dieses Kapitels ausschließlich eine Merkmalsselektion und Materialklassifikation für die Einteilung der gescannten Lagerstättenpartien in vier Klassen statt.

7.2 Einfluss der Merkmalsselektion und des Klassifikationsverfahrens

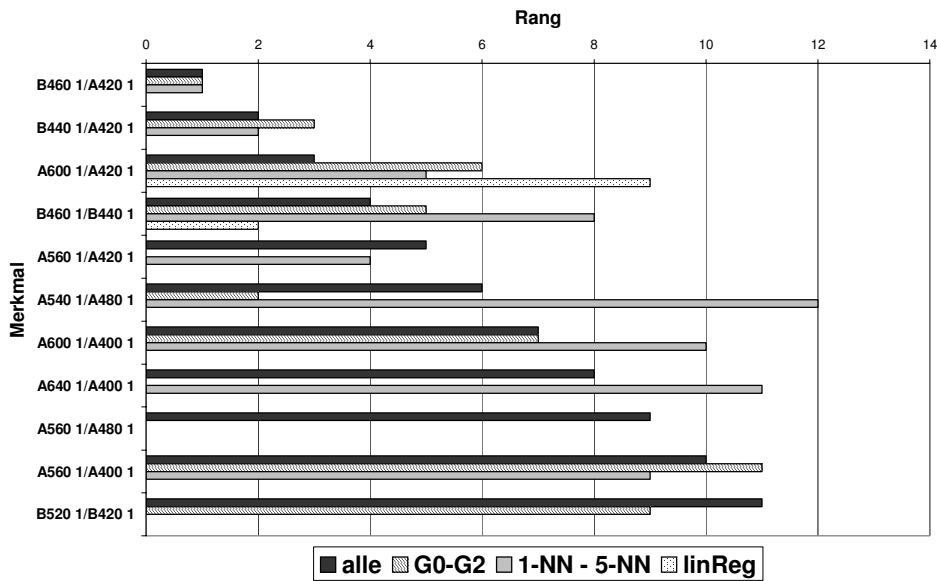
Im Rahmen der vorliegenden Arbeit fand die Merkmalsselektion und anschließende Materialklassifikation mit drei grundsätzlich unterschiedlichen Algorithmen statt. Ziel ist es, ausgehend von dieser allgemeinen Betrachtung zu prüfen, ob sich eine der verwendeten Methoden als optimal für alle Lagerstättenbereiche erweist. Dies würde für zukünftige Messreihen mit optimierter Gerätetechnik die Arbeit erheblich vereinfachen. Dabei zählen bei der Beurteilung der einzelnen Verfahren neben der Minimierung des Fehlerkriteriums auch die Aspekte des erforderlichen Rechenaufwandes sowie der Zuverlässigkeit der selektierten Merkmale und die Minimierung der erforderlichen Merkmalsanzahl zu den relevanten Entscheidungskriterien.

Im Einzelnen wird in diesem Unterkapitel für die vier untersuchten Lagerstättenbereiche bzw. Anregungswellenlängen 266 nm und 355 nm der Einfluss der jeweiligen Merkmalsselektion sowie des angewandten Klassifikationsverfahrens untersucht. Bei allen Profilen wurde mittels SGM eine Glättung von geologisch nicht plausiblen Ausreißern bei der Materialklassifikation durchgeführt.

7.2.1 Abschnitt nördlich Liegendrücken

Die Einteilung des Liegendabschnitts nördlich des Liegendrückens in vier Materialklassen folgt, wie im Unterkapitel 7.1.1 bereits ausgeführt wurde, im Wesentlichen der betrieblichen Vorgehensweise bei der Gewinnung der Kohle in diesem Lagerstättenbereich.

Das originäre Merkmalsset der beiden Messreihen im Bereich des Liegenden nördlich des Liegendrückens umfasst 135 Merkmale. Nach der Merkmalsreduktion lagen je nach Selektionsalgorithmus zwischen 35 und 81 Merkmale vor, aus denen die „BEST“-Merkmalssets generiert wurden. Deren Rangfolge, sortiert nach „BEST alle“, zeigt Abbildung 7-8. Mit Ausnahme der linearen Regression ergeben sich ähnliche Merkmalssets.



7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-8: Vergleich der Rangfolge der selektierten „BEST“ Merkmale für 4 Materialien, nördlich Liegendrücken

Im Gegensatz zur Auswertung der beiden Klassifikationen für mehr als vier Materialklassen lässt sich für M4 eine eindeutige Aussage zugunsten eines Klassifikationsverfahrens machen (Abbildung 7-9). Die k-NN-Verfahren führen für die Verwendung aller Merkmalssets mit Ausnahme von „BEST linReg“ zur besten Materialerkennung. Bei der Zuordnung der 1.551 Messpunkte lassen sich z.B. durch die Verwendung des Merkmalsvektors „BEST k-NN“ bei der Klassifikation mit 5- bzw. 10-NN weniger als 10% Fehlzugeordnungen realisieren.

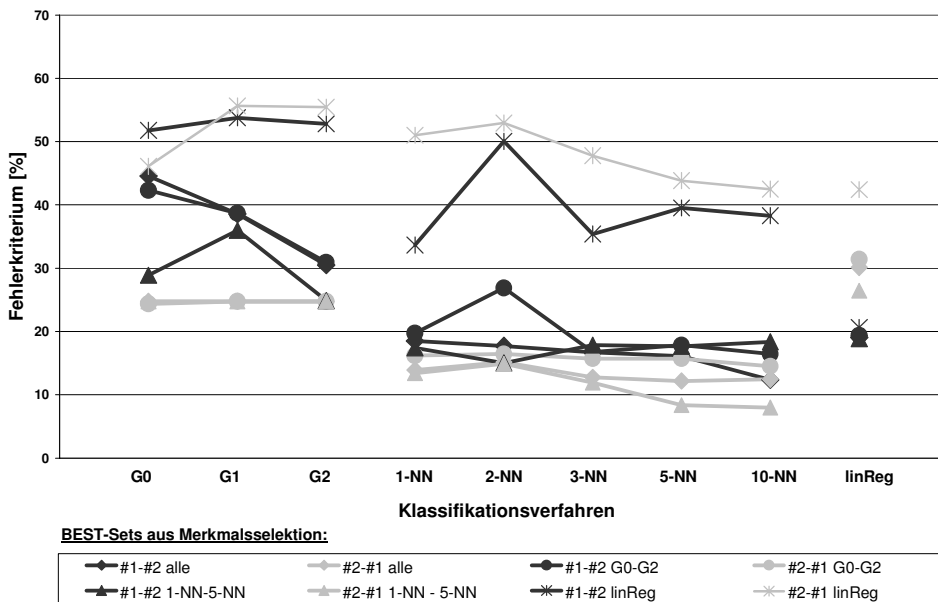


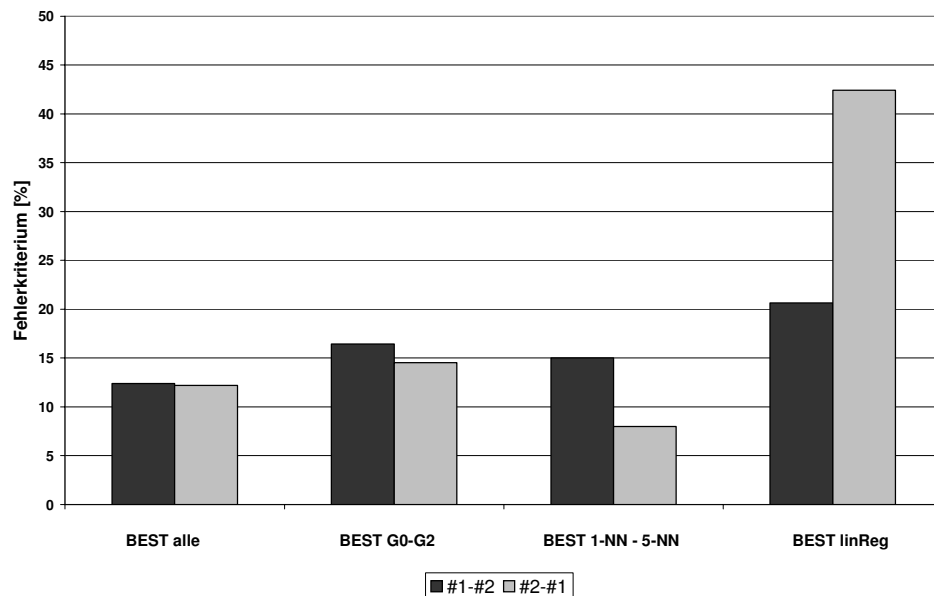
Abbildung 7-9: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, nördlich Liegendrücken

Die Frage, welches Merkmalsset und damit welches Selektionsverfahren optimal für die Reduktion der ursprünglichen 135 Merkmale ist, lässt sich jedoch nicht eindeutig beantworten. Die nachfolgende Abbildung 7-10 schließt lediglich das aus der linearen Regression gewonnene Merkmalsset als für diesen Lagerstättenabschnitt nicht geeignet aus. Tendenziell erweist sich das aus k-NN abgeleitete „BEST“-Set als am zweckdienlichsten für die

7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

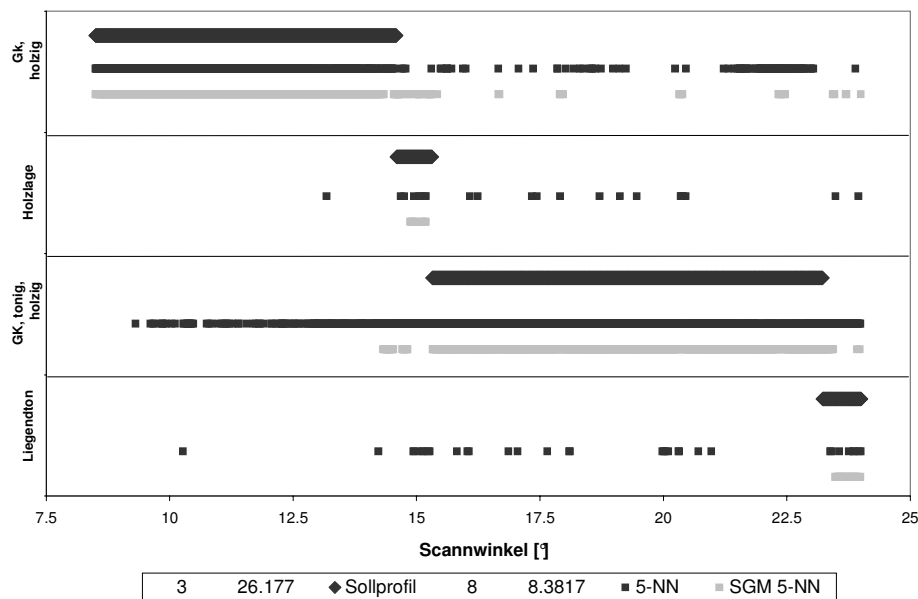
Abbildung 7-10:
Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, nördlich Liegendrücken

Aufgabe der Materialerkennung. Im Gegensatz zu den Gauß-Merkmalen besteht jedoch ein Unterschied zwischen den Ergebnissen der beiden Blindproben von mehr als 5 %, während der aus allen Selektionsalgorithmen gewonnene Merkmalsvektor „BEST alle“ in beiden Fällen ein annähernd identisches Fehlerkriterium von 12,5% aufweist. Dieser umfasst, wie im vorherigen Kapitel exemplarisch vorgestellt, die am häufigsten in den 14 Selektionsverfahren mittels Wrapper2 ermittelten Merkmale und ist aufgrund dieser Vorgehensweise am stabilsten.



Die Verbesserung des Zuordnungserfolgs durch die Verwendung des SGM beträgt je nach Merkmalsset und Klassifikation bis zu 20%. Für die Klassifikation der LIF-Messwerte mittels 5-NN, basierend auf dem Merkmalsvektor „BEST k-NN“, ergibt sich so bei der Verwendung von acht Merkmalen eine Fehlzurordnung von lediglich 8,3% (Abbildung 7-11).

Der Liegendton kann sehr klar von den beiden kohligen Bereichen und der Holzlage unterschieden werden. Beim Aushalten der bedingt bauwürdigen tonig-holzigen Partien der GK treten die bereits diskutierten Fehlzurordnungen aufgrund der Inhomogenitäten und resultierenden Ähnlichkeit zur GGМК im oberen Lagerstättenbereich auf. Die Holzlage kann auch nicht in allen Fällen eindeutig identifiziert werden, hier treten ebenfalls Überlappungen mit der oberhalb gelegenen GGМК auf. Bei anderen Profilen kommt es ebenfalls zu Fehlzurordnungen in diesen beiden kritischen Materialklassen, während die GGМК und der Liegendton in den meisten Fällen sehr deutlich und über die gesamte Klasse einheitlich zuverlässig mittels der Materialklassifikation ausgehalten werden können.



7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-11: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST k-NN“ inklusive SGM, 5-NN, 4 Materialien, #2-#1, nördlich Liegendrücken

Eine Klassifikation der bauwürdigen Kohle und des Liegendtons im Bereich nördlich des Liegendrückens lässt sich mit hoher Zuverlässigkeit mittels der mit laserinduzierter Fluoreszenz ermittelten Spektren durchführen. Eine Differenzierung zwischen den drei Klassen oberhalb des Liegendtons konnte mit den vorliegenden Basisdaten zwar für die Praxis zufrieden stellend, aber nicht eindeutig vorgenommen werden. Die Inhomogenitäten innerhalb der Klassen, besonders der ballastreichen GK mit tonigen und holzigen Beimengungen, sind im Falle der mit einer Anregungswelle von 266 nm untersuchten Profile die Ursache der Fehlzuordnungen. Mit Hilfe der Glättung durch den selektierenden gewichteten Mittelwert kann jedoch eine Vielzahl der Fehlklassifikationen im Nachgang ausgeglichen werden.

7.2.2 Abschnitt Liegendes Tiefstufe

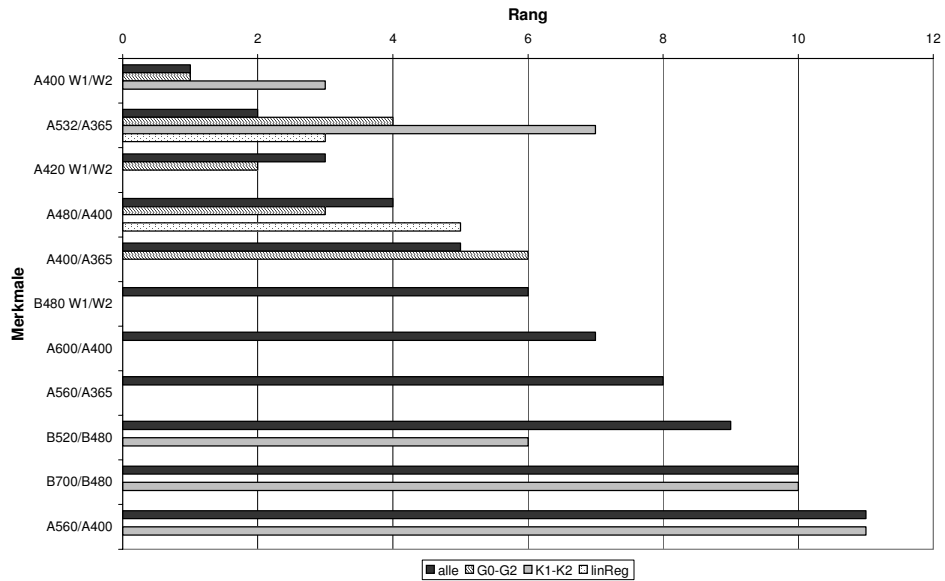
Im Bereich des Überganges zwischen Lagerstätte und dem Liegenden in der Tiefstufe ist anstelle der Holzlage eine leicht sandige Kohle anzutreffen, die den Übergang zur stärker tonig verunreinigten GK charakterisiert. Aus diesem Grunde korreliert die Einteilung des bemusterten Stoßes in vier Klassen auch an dieser Messstelle mit der betrieblichen Scheibenzuordnung des Gewinnungsbetriebes.

Das originäre Merkmalsset der beiden Messreihen in der Tiefstufe umfasst 275 Merkmale. Nach der Merkmalsreduktion lagen je nach Selektionsalgorithmus zwischen 91 und 164 Merkmale vor, aus denen die „BEST“-Merkmalssets generiert wurden. Deren Rangfolge, sortiert nach „BEST alle“, zeigt Abbildung 7-12.

Im Gegensatz zum Lagerstättenabschnitt nördlich des Liegendrückens ergeben sich nur für die ersten fünf optimierten Merkmale der Selektion Übereinstimmungen. Im allgemeinen Merkmalsset sind die Merkmale des „k-NN“-Sets sowie die des aus den Gauß-Verfahren resultierenden jeweils fünfmal vertreten. Die Differenzen bei der Merkmalsselektion sind also auffällig größer als im anderen untersuchten Bereich am Liegendübergang.

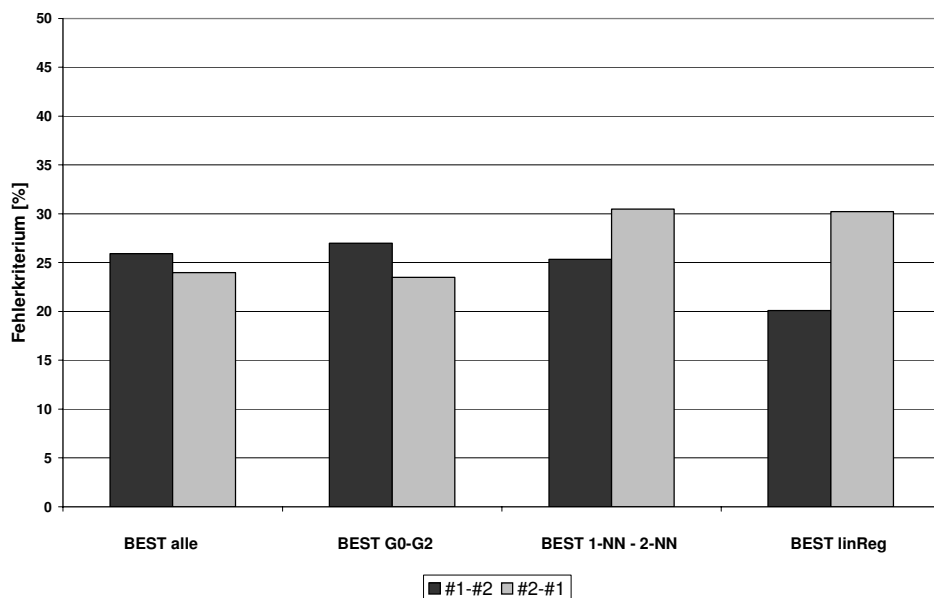
7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-12: Vergleich der Rangfolge der selektierten „BEST“ Merkmale für 4 Materialien, Liegendes Tiefstufe



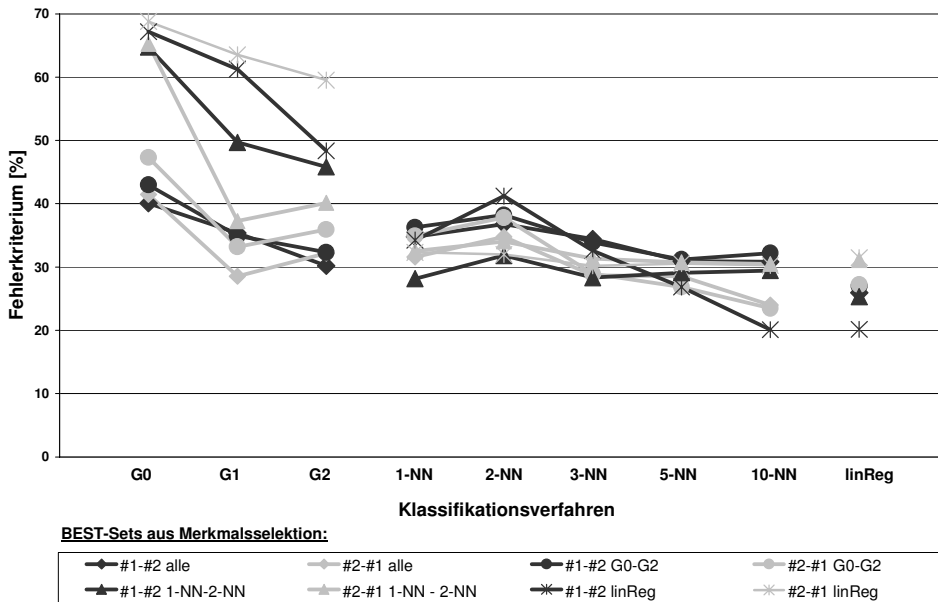
Beim Vergleich der minimalen Fehlerkriterien zeigen sich jedoch in Abhängigkeit von den verwendeten „BEST“-Merkmalssets nur geringe Unterschiede zwischen den Selektionsalgorithmen. Je nach Blindklassifikation treten Fehler in der Zuordnung der Messwerte im Vergleich zur tatsächlichen Materialkasse zwischen 20% und 30% auf. Die höchste Stabilität hinsichtlich beider Blindklassifikationen weist auch im Bereich der Tiefstufe „BEST alle“ auf. Dieser Vektor basiert hier Fall auf elf Merkmalsselektionen.

Abbildung 7-13: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, Liegendes Tiefstufe



Eine eindeutige Präferenz für einen speziellen Selektionsalgorithmus erlaubt die Bewertung der Ergebnisse auch für diesen abgescannten Schürfgaben nicht. Es lässt sich jedoch festhalten, dass die Schwankungen zwischen den beiden Blindproben mit der Einbindung einer höheren Anzahl an Selektionsverfahren abnehmen. Der optimierte Merkmalsvektor für die lineare Regression, der auf lediglich zwei unterschiedlichen Selektionen basiert, weist mit 10% aus diesem Grunde die extremste Differenz auf.

Am besten für die Materialerkennung geeignet ist wiederum das Verfahren der k-Nächsten Nachbarn. Wie Abbildung 7-14 verdeutlicht, werden dabei mit steigender Anzahl der bei der Klassifikation berücksichtigten Messwerte der Trainingsdaten, also Nachbarn des zu klassifizierenden unbekanntes Messpunktes, bessere Zuordnungen erreicht. Fehlerraten unter 30% sind in diesem Lagerstättenabschnitt zu realisieren.



7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-14: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, Liegendes Tiefstufe

Die allgemeine Problematik, die mit der Verwendung des Algorithmus der linearen Regression einhergeht, verdeutlicht exemplarisch das Profil in Abbildung 7-15. Die Klassifikation mit diesem Verfahren erkennt in den meisten der durchgeführten Fälle ein Material überhaupt nicht oder nur sehr schlecht. Dies liegt zum einen an der Anzahl der verwendeten Merkmale und zum anderen aber auch an der vorliegenden Klassenanzahl. Da das Verfahren den Merkmalsraum mittels Geraden unterteilt, darf für eine erfolgreiche Klassifikation die Anzahl der Klassen nicht die der unabhängigen Merkmale übersteigen.

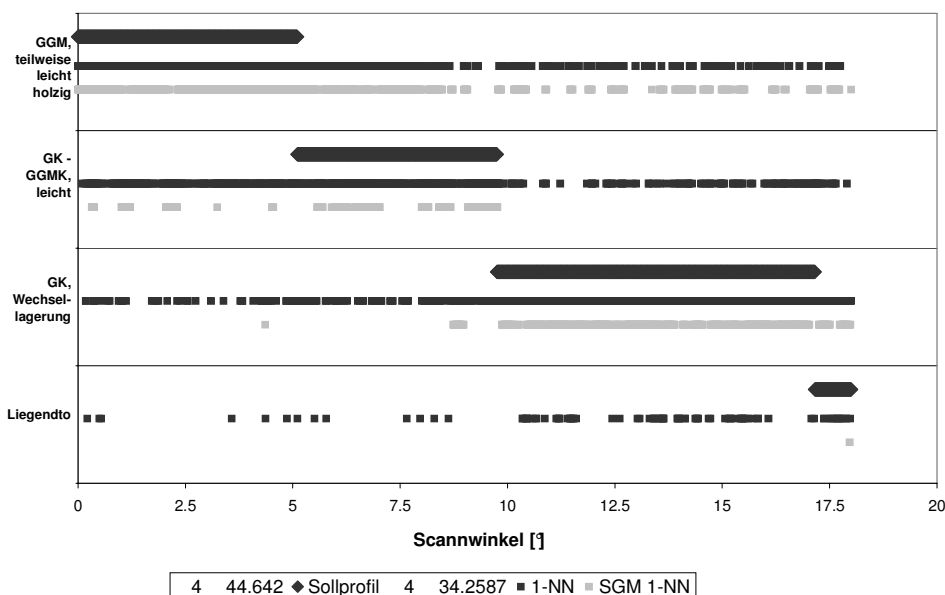


Abbildung 7-15: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST linReg“ inklusive SGM, 1-NN, 4 Materialien, #1-#2, Liegendes Tiefstufe

Trotz der stets ausreichenden Anzahl an Merkmalen erschwert ein anderer Umstand die Klassifikation bzw. Selektion mittels des Wrapper2-Ansatzes mit diesem Verfahren. Die Qualität der Merkmale entspricht aufgrund der hohen Varianzen innerhalb der Klassen nicht den Anforderungen, die dieses Selektionsverfahren an die Basisdaten stellt.

Aus diesen Gründen kann das Potential der linearen Regression aufgrund der teilweise unzureichenden Qualität der Messwerte nicht ausgeschöpft werden. Denn hinsichtlich der erforderlichen Rechendauer für die Selektion und Klassifikation benötigt dieses Verfahren nur einen Bruchteil der Zeit, wie bereits in Kapitel 5.3 graphisch gezeigt wurde.

Trotz dieser Unzulänglichkeiten ist es möglich, eine Materialzuordnung durchzuführen, die eine Unterscheidung zwischen den Klassen erlaubt, jedoch nicht den Ansprüchen der Qualitätssteuerung an eine solche Klassifikationsmethode genügt. Für den gesamten Bereich der Tiefstufe resultiert daher, basierend auf den vorliegenden LIF-Daten, nur eine eingeschränkte Materialerkennung.

7.2.3 Abschnitt Zwischenmittel Flöz Frimmersdorf a/b

Im Gegensatz zu den im Liegenden der Lagerstätte Hambach bemusterten Profilen ist das Zwischenmittel, das die Flöze Frimmersdorf a und b voneinander stratigraphisch trennt, durch eine Vielzahl von wechselnden Materialien gekennzeichnet. Diese Wechsel treten sowohl graduell auf, wie etwa unterhalb des Schluffs im Hangenden von Frimmersdorf a mit der Einlagerung von Schlufflinsen in GK, aber auch abrupt, wie beim Wechsel oberhalb des Schluffs oder auch bei der Holzlage, die die Grenze zum bauwürdigen oberen Bereich des Flözes Frimmersdorf a bildet.

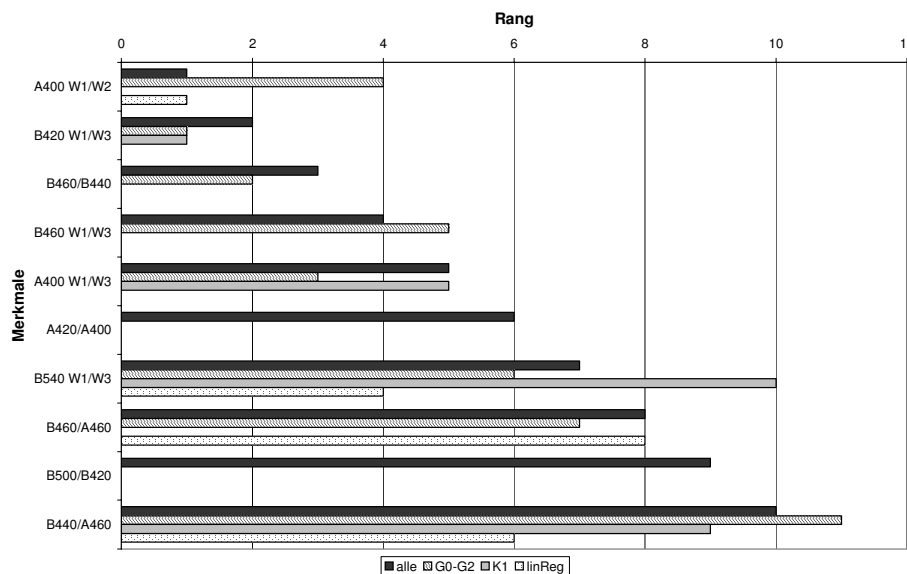
Wie bereits im vorherigen Unterkapitel bei der Bewertung der Materialerkennung anhand der Einteilung des Stoßes in neun Klassen diskutiert, bereitet die Zuordnung der angetroffenen Materialien in möglichst wenige Klassen erhebliche Schwierigkeiten. Es kommt zu Verallgemeinerungen, die zu einer Abschwächung der materialspezifischen Fluoreszenzcharakteristiken führen. Trotz dieser Unzulänglichkeiten, die zum Verlust von signifikanten Informationen führen, erfolgt die Bewertung der einzelnen Algorithmen zur Merkmalsselektion und Materialklassifikation für M4. Ein betrieblicher Einsatz würde unter Umständen mit fünf Klassen arbeiten, die den Bereich der GK oberhalb des schluffigen Zwischenmittels von der GK unterhalb differenzieren.

Die Schwierigkeiten durch verringerte Laserpulsfrequenzen und zwischenzeitliche Systemausfälle bei der Messkampagne am Zwischenmittel im April 2003 erschweren die Formulierung von Schlussfolgerungen allgemeingültiger Art für diesen Lagerstättenabschnitt jedoch sehr. Daher soll an dieser Stelle vor allem das Potential der Trennflächenerkennung mittels laserinduzierter Fluoreszenz für diesen Bereich des Tagebaus Hambach betrachtet werden.

Anregungswellenlänge 266 nm (Laser 1)

Die Messungen im kurzwelligen UV-Spektrum waren besonders von den Störungen aufgrund der kältebedingten geringen Messfrequenz betroffen. Aus diesem Grunde konnte die erste Messreihe nicht bei der Materialerkennung genutzt werden. Da der dritte Scann vorzeitig durch einen kompletten Systemabsturz unterbrochen wurde, besteht das ursprüngliche Merkmalsset lediglich aus 104 Parametern.

Durch die Reduktion wurden zwischen 34 und 70 Merkmale ausgewählt und auf ihre Eignung für die optimierten Merkmalssets „BEST“ hin überprüft. Die Gegenüberstellung der Rangfolge dieser selektierten Merkmale erfolgt in Abbildung 7-16. Es liegt eine hohe Übereinstimmung der Merkmale vor, die durch Gleichverteilung in den beiden Blindedurchläufen und darüber hinaus die häufigste Nennung bei der Merkmalsselektion gekennzeichnet sind.



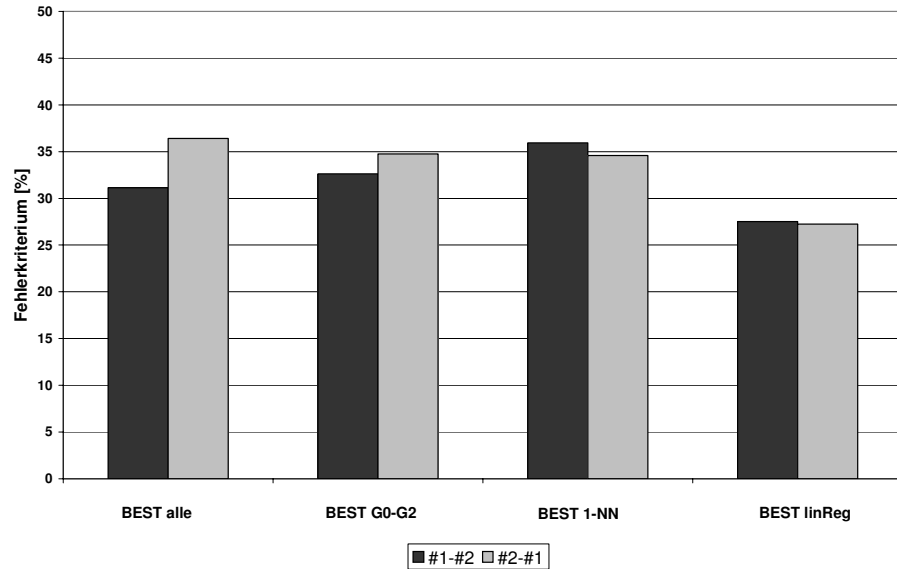
Hinsichtlich eines optimalen Selektionsalgorithmus spricht die Analyse der minimalen Fehlerkriterien für die „BEST“-Merkmalssets für das Verfahren der linearen Regression (Abbildung 7-17). Mit einer richtigen Zuordnung in mehr als 72% aller Fälle für beide Blindproben liegt diese um etwa 5% über den anderen Verfahren.

7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-16: Vergleich der Rangfolge der selektierten „BEST“ Merkmale für 4 Materialien, Zwischenmittel. 266 nm

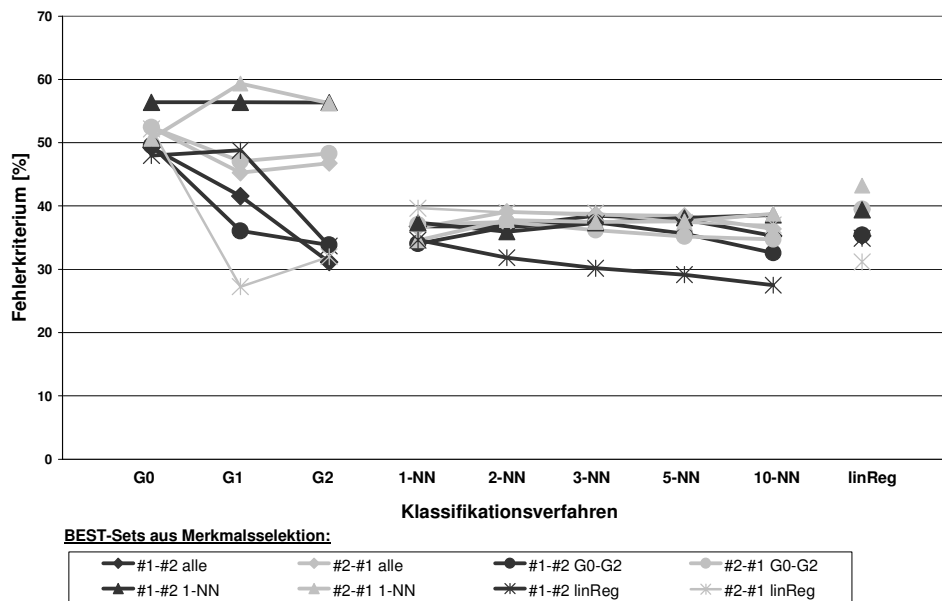
7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-17: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, Zwischenmittel, 266 nm

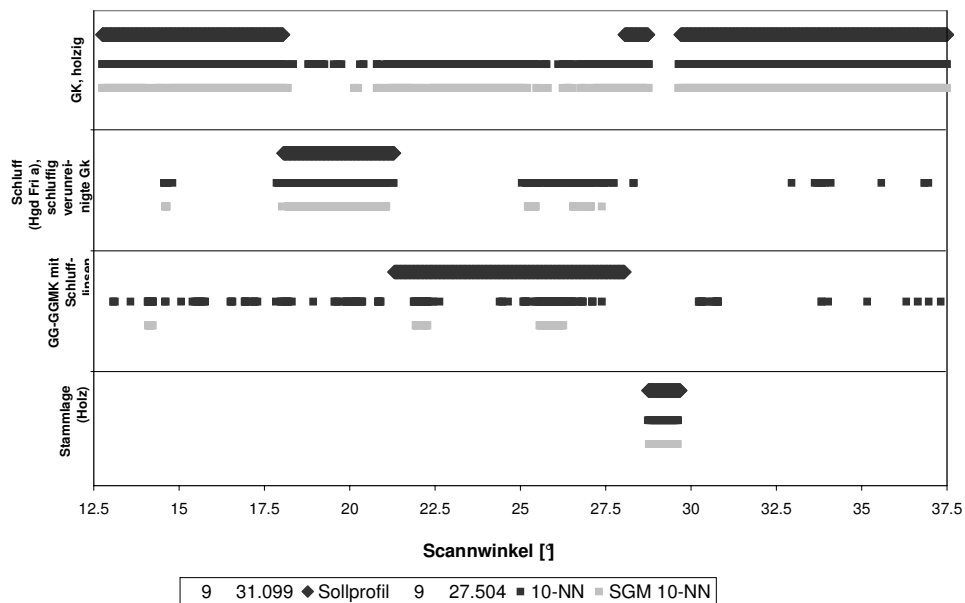


Wie bereits bei der Materialerkennung im Liegenden der Braunkohlelagerstätte erweist sich die Klassifikation mittels k-NN-Algorithmen als geeignete Wahl zur Minimierung des Fehlerkriteriums (Abbildung 7-18). Im Gegensatz zu den Messungen im Liegenden sind Fehlzuordnungen in der Größenordnung von 35% die Regel, was sich auf die Beeinträchtigung der Messungen und der damit verbundenen Qualität der Messwerte sowie die kompliziertere Klasseneinteilung zurückführen lässt.

Abbildung 7-18: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, Zwischenmittel, 266 nm



Der Einfluss der Klassendefinition innerhalb des inhomogenen Stoßabschnittes im Zwischenmittel spiegelt auch das Profil der Merkmalsklassifikation für das „BEST linReg“ Merkmalsset bei Verwendung des 10-NN-Algorithmus für die Materialerkennung wieder. Die charakteristische Stammlage und die unterhalb gelegene GK des bauwürdigen Hangenden des Flözes Frimmersdorf a werden scharf getrennt. Die Schlufflage sowie stark schluffige Partien werden ebenfalls richtig zugeordnet.



7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-19: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST linReg“ inklusive SGM, 10-NN, 4 Materialien, #1-#2, Zwischenmittel, 266 nm

Es treten hingegen wie bereits im Liegenden Schwierigkeiten bei der Zuordnung von inhomogenen bzw. mit Einlagerungen verunreinigten Materialklassen auf. So wird die mit Schlufflinsen verunreinigte GK häufig auch als holzige GK bzw. als Schlufflage selber erkannt, was als Ähnlichkeit der Messwerte zu den reineren Stoßpartien zu werten ist.

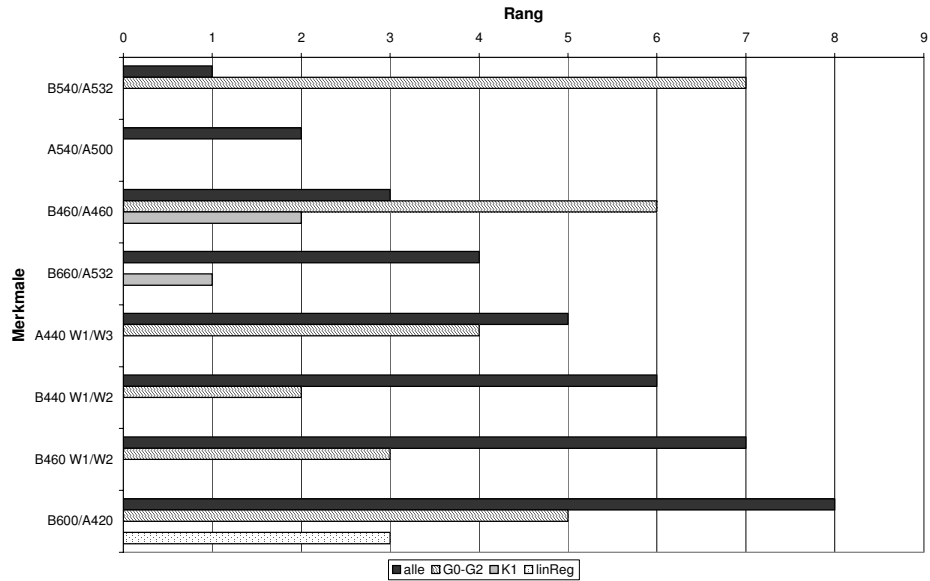
Anregungswellenlänge 355 nm (Laser 2)

Um überhaupt eine Materialerkennung für die Anregung im langwelligeren UV durchführen zu können, bedurfte es einer gesonderten Vorverarbeitung der Originaldatensätze. Aufgrund der hohen Frequenzschwankungen bei der ersten Bemusterung des Stoßes mit 355 nm wurden über einen Winkel von 25° nur zwischen 725 und 1.503 Messwerte erfasst (Scann #2: 1.968 – 2.148). Um durch den Einfluss der Mittelwertbildung bei der Datenmodulation („20 0,5 5“) die materialspezifischen Charakteristika nicht zu stark zu verzerren, wurde jeder Merkmalsatz, der nach der Glättung 2.551 winkelnormierte und gemittelte Werte je Filterradeneinstellung umfasst, auf die Hälfte reduziert, d.h. jeder zweite Werte wurde entfernt. Die Ergebnisse dieses Unterkapitels sind vor diesem Hintergrund der nicht unerheblichen Beeinflussung der Messungen durch äußere und Messgerät bedingte Störgrößen zu bewerten.

Aus dem originären Merkmalsset von 405 Parametern wurden zwischen 66 und 162 Merkmale mittels Reduktion als relevant ausgewählt. Aufgrund der erheblichen Unregelmäßigkeiten zwischen den beiden Blindklassifikationen aber auch der generell geringen Häufigkeit der Nennung einzelner Merkmale setzen sich die „BEST“-Merkmalsvektoren nur aus drei bis acht Merkmalen zusammen, die im Vergleich zu den vorherigen Sets die geringsten Übereinstimmungen aufweisen. Die Rangfolge, sortiert nach „BEST alle“, zeigt Abbildung 7-20.

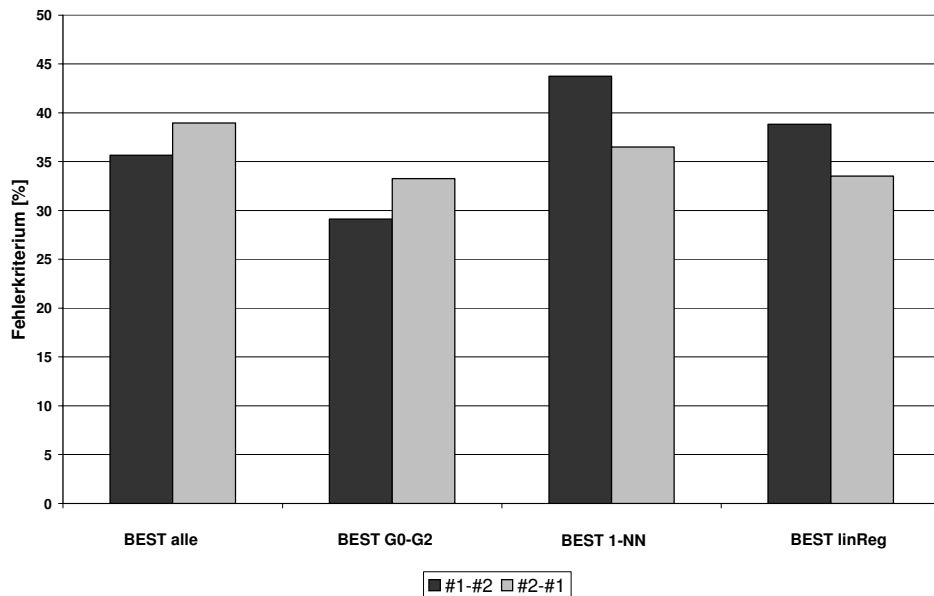
7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-20: Vergleich der Rangfolge der selektierten „BEST“ Merkmale für 4 Materialien, Zwischenmittel, 355 nm



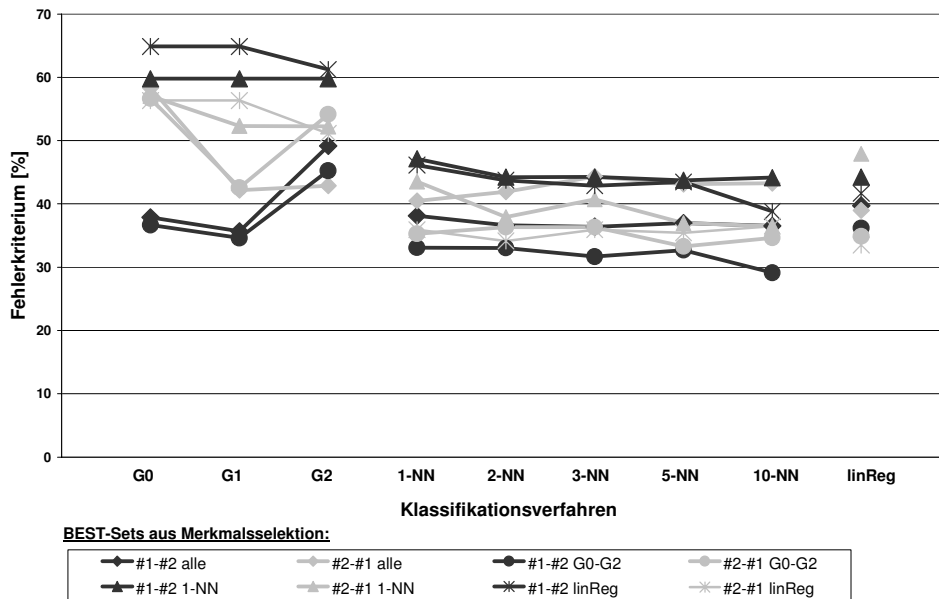
Die besten Klassifikationsergebnisse lassen sich bei der Nutzung der langwelligeren UV-Anregung mit dem optimierten Merkmalsset erzielen, das sich aus der Bearbeitung der Merkmalsselektion mittels der drei Gauß-Verfahren ergibt (Abbildung 7-21). Vergleichend ergeben sich ähnliche Größenordnungen hinsichtlich des Fehlerkriteriums wie bei 266 nm.

Abbildung 7-21: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, Zwischenmittel, 355 nm



Die Materialklassifikation mit Hilfe der k-NN-Algorithmen liefert auch beim Zwischenmittel mit Laser 2 die besten Ergebnisse, wobei Fehlzuordnungen im Bereich zwischen 30% und 40% die Regel sind (Abbildung 7-22).

Klassifikationsverfahren zur Materialerkennung



7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Abbildung 7-22: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, Zwischenmittel, 355 nm

Ebenso wie bereits bei Laser 1 gestaltet sich die Materialerkennung für die Bereiche inhomogener Materialien am schwierigsten. Die Stammlage sowie die unterhalb gelagerte GK sind ebenso wie die GK oberhalb des Schluffs und diese Lage selber sehr gut zu erkennen. Lediglich der Bereich der GK mit Schlufflinsen wird sehr schlecht mit der laserinduzierten Fluoreszenz der richtigen Materialklasse zugeordnet (Abbildung 7-23). Durch die Verwendung der Glättung mit Hilfe des SGM verringern sich die Fehlzurechnungen um maximal 5%.

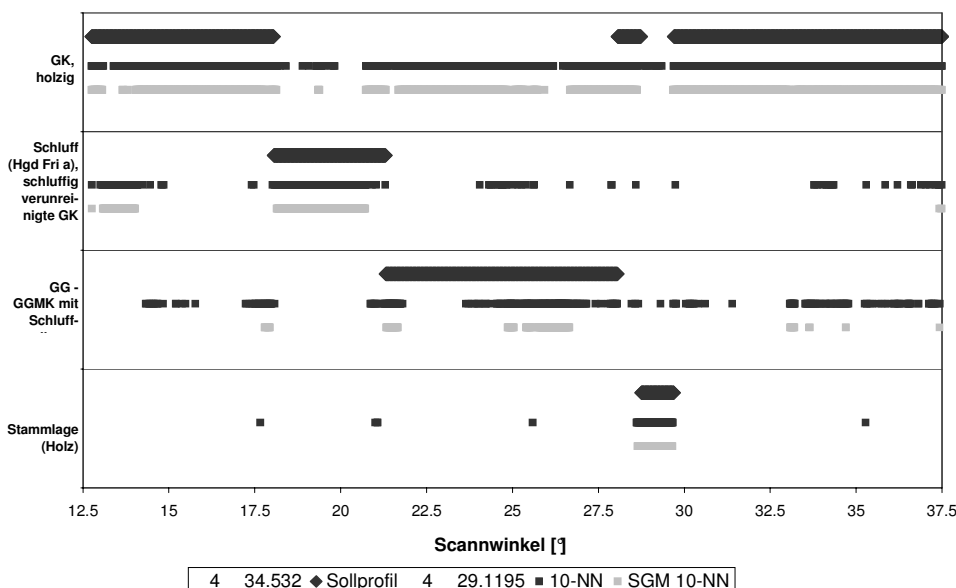


Abbildung 7-23: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST G0-G2“ inklusive SGM, 10-NN, 4 Materialien, #1-#2, Zwischenmittel, 355 nm

Bei der Materialerkennung mit dem Laser 2 handelt es sich um eine echte Blindprobe, da die beiden Scannspuren räumlich versetzt nebeneinander und nicht über einen identischen Abschnitt des Stoßes das Material bemusterten. Neben den Unzulänglichkeiten bei der Datenerfassung bestand hier zusätzlich die Problematik der Korrelation der einzelnen stratigraphischen Abschnitte zu gleichen Materialklassen.

Unter Berücksichtigung der Komplikationen bei der Durchführung der Messungen und der nachgelagerten Messwertvorverarbeitung sind die bei der Materialklassifikation im Bereich des Zwischenmittels erzielten Zuordnungen als Erfolg zu sehen. Bei störungsfreier Anwendung des LIF Scanner / Analysers unter idealen Bedingungen sollten die betrieblichen Vorgaben der Qualitätssteuerung an die optische Erkennung erfüllbar sein.

7.3 Einfluss der Anregungswellenlänge des Lasers

Der Einfluss der Anregungswellenlänge auf die Materialerkennung in der Braunkohle kann lediglich am Beispiel des Zwischenmittels bewertet werden, da im Liegenden der langwelligere Laser 2 nicht zum Einsatz kam. Wie im vorherigen Unterkapitel ausgeführt wurde, kennzeichnen eine aus vielen Gründen schlechte Datenqualität die Messungen in diesem Abschnitt der Lagerstätte, so dass keine konkreten Aussagen möglich sind. Bei beiden Anregungswellenlängen liegt das Fehlerkriterium über 25%, womit in keinem Falle betrieblichen Anforderungen entsprochen wird. In beiden Fällen scheitert die Materialerkennung im Wesentlichen in den Bereichen des bemusterten Stoßes, die durch Beimengungen sehr inhomogen sind, sei es durch Schlufflinsen oder holzige Bestandteile. Für die homogeneren Partien lassen sich mit beiden Anregungswellenlängen die einzelnen Materialklassen mit vergleichbarer Erfolgsquote identifizieren.

Die Erfahrungen mit denen bis dato am BGMR durchgeführten Untersuchungen mit der laserinduzierten Fluoreszenz an mineralischen Rohstoffen, lassen tendenziell eine bessere Materialklassifikation mit kurzwelligerem UV-Laserlicht zu. Dies wird durch die Messungen am Zwischenmittel der Lagerstätte Hambach nicht bestätigt. Aus diesem Grunde ist es erforderlich, dass diese Fragestellung in folgenden Versuchsreihen weitergehend untersucht wird. Die Reduktion der Gerätetechnik auf nur eine Anregungswellenlänge hätte über eine Vereinfachung von Aufbau und Steuerung eines LIF Scanner / Analysers hinaus den Vorteil der Verringerung der Investitionskosten.

7.4 Verwendung von übergreifenden und universellen Merkmalen für das Liegende

Eine Voraussetzung für den Einsatz der laserinduzierten Fluoreszenz zur Trennflächenerkennung in der mineralischen Rohstoffgewinnung ist die universelle Einsetzbarkeit innerhalb der Lagerstätte. Dies macht es erforderlich, dass mit einem definierten Set an Filterradeneinstellungen und damit ausgewählten Emissionsspektralbanden unterschiedliche Profile respektive Materialzusammensetzungen einheitlich erfolgreich klassifiziert werden können.

Im Rahmen der Messkampagnen im Tagebau Hambach stehen für die Verwendung von übergreifenden und universellen Merkmalen ausschließlich die Messungen mit der Anregungswellenlänge 266 nm zur Verfügung. Im Liegenden kam der Laser 2 mit 355 nm aufgrund der begrenzten Mess-

zeiten nicht zum Einsatz, so dass eine Betrachtung für diese Wellenlänge über verschiedene Lagerstättenteile nicht möglich ist.

Bei der Auswahl der geeigneten Parameter für alle Messungen mit Laser 1 können die Abklingzeiten nicht herangezogen werden, da die Torzeiten für die Integrale W2 und W3 im Herbst 2002 nicht mit denen im April 2003 übereinstimmen. Ohne Torzeiten stünden bei der Verwendung aller drei Profile nur 36 Merkmale zur Verfügung, da eine Reihe von Filterradeneinstellungen aus verschiedenen Gründen nicht bei allen Messungen zum Einsatz kam. Diese Tatsache und die generell schlechtere Datenqualität im Zwischenmittel begründen die alleinige Fokussierung bei der Bewertung von universellen Merkmalen auf die Messungen im Liegenden der Lagerstätte Hambach.

Für die Durchführung dieser Betrachtungen kann auf insgesamt 135 Merkmale als Ausgangsdatenbasis zurückgegriffen werden (siehe Anhang). Bei der Optimierung der gemeinsamen „BEST“-Merkmalsvektoren ließen sich keine Parameter ermitteln, die sowohl bei der Selektion mittels Gauß - Verfahren als auch bei den k-NN-Algorithmen den Auswahlkriterien entsprechen. Für die lineare Regression war es ebenfalls nicht möglich, gemeinsame Merkmale für den Bereich nördlich des Liegendrückens und der Tiefstufe zu generieren. Daher können lediglich zwei „BEST“-Merkmalssets zur Profil übergreifenden Untersuchung der Eignung der laserinduzierten Fluoreszenz herangezogen werden. Tabelle 7-1 zeigt die acht bzw. zwölf Merkmale, die den entsprechenden optimierten Merkmalsvektor bilden.

Rang	BEST G0-G2	BEST k-NN
1	B520/B480	A560/A420
2	A560/A400	A640/A420
3	A540/A532	A600/A420
4	A560/A532	A540/A420
5	B540/B480	A560/A400
6	A560/A540	A640/A400
7	B480/B420	A540/A480
8	A480/A400	A540/A400
9		B660/B420
10		B660/B520
11		B660/B480
12		B520/B420

Die Tabelle zeigt, dass bei den universellen Merkmalen die Verwendung der Abklingzeiten nicht zielführend zu sein scheint und auch die Quotientenbildung aus PMT A und PMT B nicht zur Materialerkennung herangezogen wird.

Anhand der Analyse der minimalen Fehlerkriterien für die beiden Selektionsverfahren kommt man zum Ergebnis, dass neben den tendenziell bes-

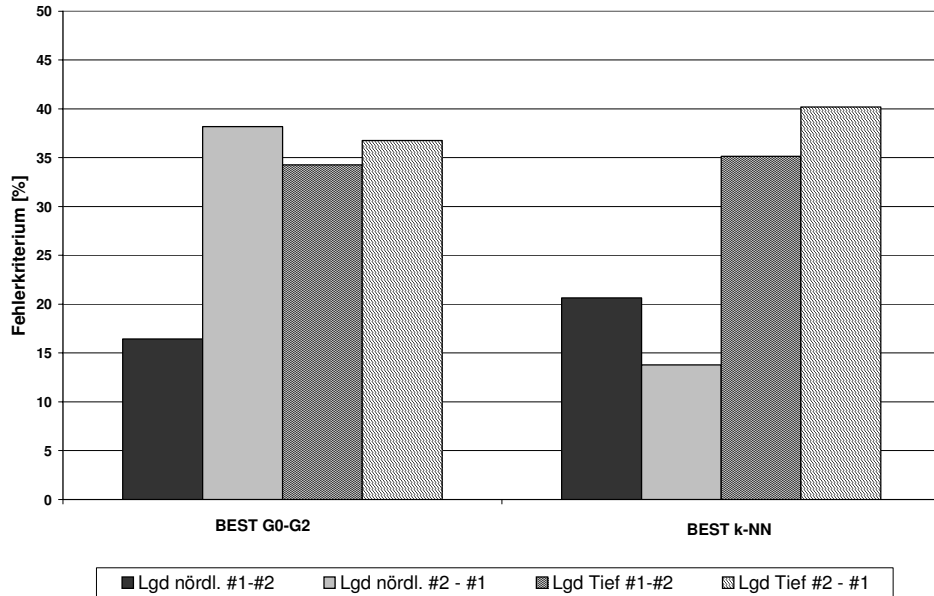
7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

Tabelle 7-1:
Rangfolge der universellen selektierten „BEST“ Merkmale für 4 Materialien, Liegendes, 266 nm

7 Darstellung und Bewertung der Merkmalsselektion und Materialklassifikation

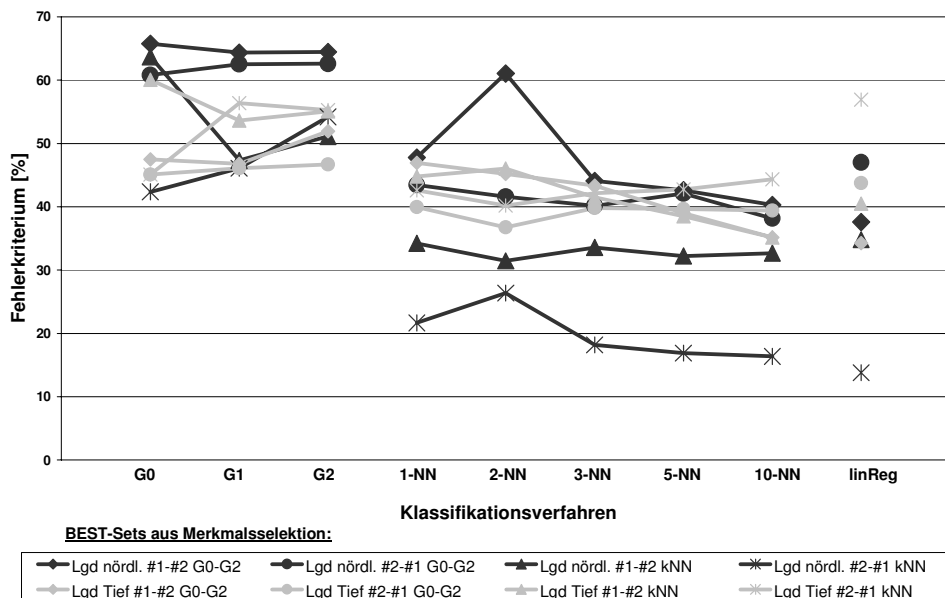
Abbildung 7-24: Vergleich der minimalen Fehlerkriterien für universelle BEST Merkmalssets, Liegendes

seren Ergebnissen im Lagerstättenabschnitt nördlich des Liegendrückens keine generelle Aussage zugunsten eines Reduktionsalgorithmus getroffen werden kann. Bei einer Mittlung der minimalen Fehlerkriterien über alle vier Scans schneidet „BEST k-NN“ tendenziell besser ab als die Gauß'sche Selektion (Abbildung 7-24).



Die Bewertung der Eignung der drei Klassifikationsalgorithmen zur Materialerkennung ergibt sich aus Abbildung 7-25. Die Gauß'schen Verfahren schneiden schlechter ab, was den Schluss nahe legt, dass aufgrund der hohen Varianzen die Bildung materialspezifischer Mittelwerte negativ beeinflusst wird.

Abbildung 7-25: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens für universelle Merkmalssets, 4 Materialien, Liegendes



Mit Ausnahme der beiden Blindklassifikationen des Profils Liegend nördlich mit dem „BEST k-NN“-Merkmalsset werden maximale Zuordnungsquoten von 60% erreicht. Die beiden angesprochenen Klassifikationen erreichen

Klassifikationsverfahren zur Materialerkennung

Werte darüber. Mit den „BEST k-NN“-Set bei #2-#1 wird bei der Materialerkennung mit 10-NN sogar ein Fehlerkriterium unter 20% realisiert.

Die Ergebnisse der Nutzung von übergreifenden und universell einsetzbaren Merkmalen im Bereich des Liegenden der Lagerstätte Hambach lassen nur den Schluss zu, dass ausgehend von der derzeitigen Datenbasis mit keinem Selektions- und Klassifikationsalgorithmus ein einheitlicher Merkmalsvektor generiert werden kann, der zu einer Erfolg versprechenden Materialerkennung führt. Die gezeigten Ansätze deuten jedoch darauf hin, dass prinzipiell eine universelle Vorgehensweise zu realisieren ist. Grundlage dafür sind neben der Güte der Messwerte ebenso die Erhöhung an universell einzusetzenden Merkmalen (Messung mit allen Filtradeinstellungen an allen Profilen) und die Größe der Datenbasis. Zur Minimierung der Varianzen innerhalb der einzelnen Materialklassen bedarf es vor allem der Umsetzung des letztgenannten Aspektes.

**7 Darstellung und
Bewertung der
Merkmalsselektion und
Materialklassifikation**

8 Diskussion der Ergebnisse

An dieser Stelle sollen die mittels der Merkmalsselektion und Materialklassifizierung gewonnenen Erkenntnisse kritisch hinterfragt und eine abschließende Bewertung vor dem Hintergrund der praktischen Nutzung der laserinduzierten Fluoreszenz zur Grenzschichterkennung in mineralischen Lagerstätten vorgenommen werden. Bevor die einzelnen Selektions- und Klassifikationsalgorithmen auf ihre Eignung zur Lösung der Problemstellung analysiert werden, sollen jedoch noch einmal die Schwierigkeiten bei der Erfassung der Daten und deren Auswirkungen auf die Resultate verdeutlicht werden.

Bei beiden Messkampagnen traten erhebliche Streuungen der Messwerte und unerklärliche Unstimmigkeiten der Fluoreszenzcharakteristiken zwischen parallelen Scans mit identischen Filterradeneinstellungen auf. Geräteausfälle bzw. Beeinträchtigungen durch Umwelteinflüsse stellen Gründe für diese schlechte Rohdatenqualität dar. Doch durch die Eliminierung von offensichtlich falschen Emissionsspektren und die Vorverarbeitung zur Reduktion des Rauschens mittels Datenmodulation sind die für die Klassifikation verwendeten Datensätze für die Trennflächenerkennung hinreichend signifikant. Allerdings muss bei der Bewertung berücksichtigt werden, dass in keinem Falle ideale Rahmenbedingungen im Hinblick auf die Datenqualität vorliegen und die Ergebnisse der Arbeit vielmehr nur das Potential der laserinduzierten Fluoreszenz und der ausgewählten Algorithmen aufzeigen. Vor dem Hintergrund der fehlerhaften Datenbasis sind die Ergebnisse eher als grundsätzliche Tendenzen zu verstehen. Es gelingt durch die Verwendung von anspruchsvollen, statistischen Verfahren wie dem Wrapper2-Ansatz bei der Materialklassifikation Ergebnisse mit guter Aussagequalität zu erreichen.

Die Abweichungen bei der Merkmalsselektion und auch der anschließenden Materialklassifikation zwischen den beiden Blindproben der Messungen sind als Ergebnis der oben geschilderten Unzulänglichkeiten bei der Durchführung der Messungen und der Datenerfassung zu werten. Die Nutzung des Wrapper2-Ansatzes trägt dieser Problematik der starken Streuung der Messwerte Rechnung und stabilisiert die Merkmalsselektion. In diesem Zusammenhang ist mit der Verwendung einer Resampling Rate von 80% die in der Literatur empfohlene Einstellung verwendet worden, die auch die teilweise kleine Klassengröße innerhalb bestimmter Stoßabschnitte vollständig berücksichtigt.

Eine inkorrekte Stoßansprache stellt eine systematische Fehlerquelle dar, die jedoch bis auf eine Ausnahme die beiden Blindproben gleichermaßen beeinflusst. Denn alle Scans wurden bis auf die auf die Materialklassifikation im Zwischenmittel mit dem Laser 2 jeweils in einer identischen Spur geführt. Falls eine Klasseneinteilung aufgrund der Ansprache falsch gewählt wurde, bestehen die Auswirkungen für jeweils beide Scans. Die teilweise nicht korrelierenden Fluoreszenzverläufe in einzelnen Abschnitten

könnten unter Umständen bei einer möglichen Fehleinteilung die klassen- und damit materialspezifischen Mittelwerte der Fluoreszenzcharakteristika unterschiedlich beeinflussen. Doch wird dieser Fehler als eher klein im Vergleich zu den Varianzen innerhalb der Messwerte einer Klasse eingeschätzt.

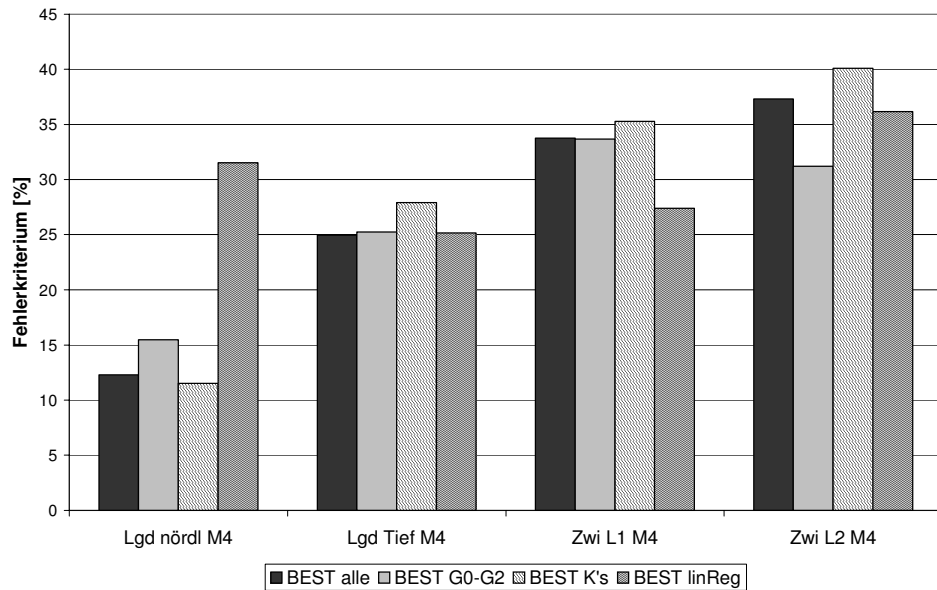
Als ein Ergebnis dieser Arbeit ist festzuhalten, dass eine zu detaillierte Einteilung der Lagerstätte in Materialklassen nicht im Sinne einer zuverlässigen Trennflächenerkennung zielführend ist. Zunächst sollte eine Materialerkennung darauf abzielen, dass eine eindeutige Erkennung von bauwürdigen und nichtbauwürdigen Partien der Lagerstätte möglich ist. Eine Aussage Schwarz, Weiß oder Grau ist für den Bediener des Gewinnungsgerätes zunächst ausreichend. Im Fall des Liegenden für den Abschnitt nördlich des Liegendrückens konnte dieses Ziel nur mit der Unterteilung des Stoßes in vier Materialklassen erreicht werden. Die Materialerkennung für einen neungeteilten Stoß scheiterte an der Ähnlichkeit der sich nur durch variierende Verunreinigungsgrade mit holzigen und tonigen Bestandteilen unterscheidenden Klassen. Auch für das Zwischenmittel konnten eindeutig bessere Zuordnungen bei einer geringeren Anzahl von Materialklassen erzielt werden.

Daher sollte eine Einteilung der Lagerstätte in Zukunft in solche Klassen erfolgen, die auch bei der Gewinnung für die Zuordnung der Baggerscheiben verwendet werden. Beim Vorliegen von qualitativ hochwertigen LIF-Messdaten sollte jedoch nochmals ein Vergleich mit einer höheren Klassenanzahl vorgenommen werden, da die Vorteile einer höher auflösenden Materialerkennung offensichtlich sind. Anstelle von sprunghaften Änderungen könnte das LIF-System dem Baggerführer graduelle Veränderungen des Materialverlaufs über die Schneidspur liefern. Es besteht dann nicht mehr der Notwendigkeit, nicht bauwürdige Schichten anzuschneiden, um die Trennfläche eindeutig detektieren zu können.

Die Auswertung der durchgeführten Merkmalsselektion für alle drei Lagerstättenbereiche kommt hinsichtlich einer Entscheidung für einen optimalen Algorithmus zu keinem eindeutigen Ergebnis. Eine Gegenüberstellung des Erfolges der optimierten selektierten „BEST“-Merkmalssets der drei grundsätzlichen Verfahren sowie des Merkmalsvektors, der aus allen Algorithmen generiert wurde, zeigt Abbildung 8-1. Hierbei ist zu berücksichtigen, dass die Überprüfung der zugeordneten Materialklassen im Sinne einer geologischen Plausibilität mittels des Verfahrens des gewichtenden selektierenden Mittelwerts (SGM) erheblich zur Verbesserung der Materialerkennung beiträgt. Durch diese Glättung stratigraphisch nicht sinnvoller Fehlklassifikationen kann das Fehlerkriterium um bis zu 20% verringert werden.

8 Diskussion der Ergebnisse

Abbildung 8-1:
Vergleich der minimalen Fehlerkriterien für BEST Merkmalssets in Abhängigkeit der Lokation, Tagebau Hambach



In Abhängigkeit vom betrachteten Lagerstättenabschnitt sind unterschiedliche Vorgehensweisen zu präferieren. Tendenziell erweisen sich Merkmalssets „BEST alle“ als die beste Wahl. Angesichts des erheblichen Rechenaufwandes zur Generierung der erforderlichen Merkmalsselektionen sollten jedoch andere Wege gegangen werden. Im Falle des untersuchten Bereichs nördlich des Liegendrückens stehen die Rechendauern für beide Blinndurchläufe im Verhältnis 1:3,73 oder 79,4 h:296 h beim Vergleich zwischen den drei Gaußverfahren und den k-NN-Algorithmen (1-NN, 2-NN sowie 5-NN). Die lineare Regression benötigte nur 1,2 h für die Durchführung der beiden Reduktionen.

Eine Kombination der Gauß-Verfahren sowie 1-NN und der linearen Regression könnte einen Kompromiss zwischen der Nutzung von „BEST alle“ und der alleinigen Beschränkung auf einen Algorithmus darstellen. Der Rechenaufwand läge in diesem Fall um etwa ein Drittel unter dem, der für die Ermittlung der entsprechenden Ergebnisse der kompletten Durchführung des Wrapper2 für alle sieben Algorithmen erforderlich wäre.

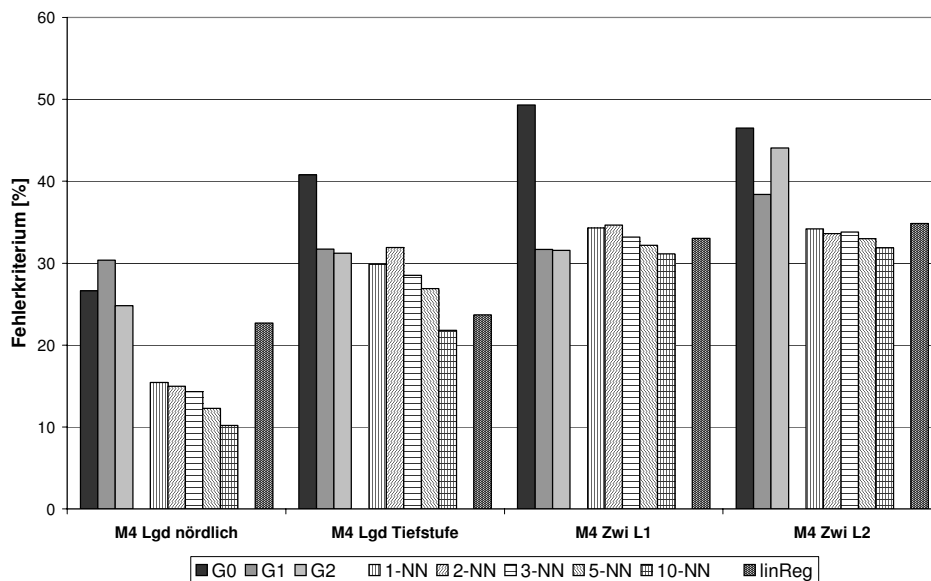
Die Qualität der Material- und damit Trennflächenerkennung offenbart sich ebenfalls in der obigen Abbildung. Die mittleren minimalen Fehlerkriterien für die jeweils besten Materialklassifikationen liegen im Übergangsbereich der Lagerstätte zum Liegenden erheblich unter denen für das Zwischenmittel Frimmersdorf a/b. Ein Grund für diese schlechtere Materialerkennung im Zwischenmittel liegt in der Einteilung der Klassen in eine minimale Anzahl, die unter geologisch-stratigraphischen und auch abbautechnischen Aspekten noch sinnvoll ist. Dadurch werden eigentlich unterschiedliche Fluoreszenzcharakteristika gemittelt und signifikante Unterschiede „verwässert“. Die geringere Materialanzahl bei den Untersuchungen nördlich des Liegendrückens erlaubt daher eine bessere Materialzuordnung.

Die Unterschiede zwischen den beiden Messungen im Liegenden können mit der generell geringeren Fluoreszenzintensität im Bereich der Tiefstufe begründet sein. Aufgrund der Varianzen treten daher materialspezifische Klasseneigenschaften nicht so deutlich hervor.

Die generelle Eignung der durchgeführten Materialerkennung für die Qualitätssteuerung in der Braunkohlegewinnung ist mit Ausnahme des Abschnitts im Liegenden der Lagerstätte nur bedingt gegeben. Auch wenn im Zwischenmittel die bauwürdigen Schichten eindeutig identifiziert werden und die hohen Fehlerkriterien hauptsächlich nur im Bereich einer Klasse auftreten, nämlich der GK-GGMK mit Schlufflinsen, lassen die Resultate nur dieses Fazit zu:

Die Materialien nördlich des Liegendrückens können mit den vorhandenen LIF Daten eindeutig identifiziert und im Sinne einer Echt-Zeit-Steuerung des Schaufelradbaggers genutzt werden.

Abbildung 8-2 zeigt für alle vier durchgeführten Materialerkennungen getrennt nach Klassifikationsalgorithmen die Fehlerkriterien. Diese stellen jeweils das für beide Blindproben gemittelte Minimum über die „BEST“-Merkmalssets dar. Es zeigt sich, dass die Verfahren nach dem Prinzip des nächsten Nachbarn über alle vier Bereiche die besten Resultate liefern. Mit steigender Anzahl der bei diesem Verfahren berücksichtigten Nachbarn als Referenzwerte aus den Trainingsdaten verbessert sich die Materialzuordnung. So treten Differenzen zwischen 1-NN und 10-NN von bis 10% auf. Mit steigender Anzahl an Nachbarn verringert sich der Einfluss der Messwertstreuungen auf das Klassifikationsergebnis.



8 Diskussion der Ergebnisse

Abbildung 8-2: Vergleich der minimalen Fehlerkriterien in Abhängigkeit des Klassifikationsverfahrens bei jeweils optimalem „BEST“-Merkmalsset, Tagebau Hambach

Als positiv für die Eignung der laserinduzierten Fluoreszenz zur Qualitätssteuerung im Gewinnungsprozess zu bewerten sind vor allem die im Bereich nördlich des Liegenden realisierten Erfolge bei der Materialerkennung von teilweise um 90 %. Die Fehlklassifikationen, die bei diesen Materialzuordnungen auftreten, betreffen dabei nicht die kritische Unterscheidung zwischen Wertmineral und Nebengestein, sondern lediglich verunreinigte Partien in der Braunkohle, die jedoch möglichst ebenfalls von reinerer Kohle ausgehalten werden sollte. Es bedarf aber auf jeden Fall noch erheblicher Verbesserungen bei der Qualität der Daten, um auch für die anderen Lagerstättenabschnitte ähnliche Ergebnisse zu erzielen. Zudem müssen die Scans je Schürfgraben gewährleisten, dass mehr als nur zwei Mess-

8 Diskussion der Ergebnisse

reihen je Übergangsbereich vorgenommen werden, dies dann außerdem für beide Anregungslaser. Nur mit einer größeren Datenbasis können repräsentative Aussagen hinsichtlich einer praktischen Nutzung der laserinduzierten Fluoreszenz getroffen werden.

Die Beantwortung der Frage, ob bereits auf Basis dieser Arbeit auf eine Anregungswellenlänge zur erfolgreichen Realisierung einer auf LIF basierenden Materialerkennung verzichtet werden kann, erlauben die Untersuchungen und die Ergebnisse dieser Arbeit jedoch nicht. Aufgrund der bereits angesprochenen Datenbasis waren dazu erforderliche Analysen nicht möglich, so dass es hier weiterer Messkampagnen bedarf.

Die Verwendung von allgemeingültigen Merkmalen zur Identifikation von Trennflächen, wie sie in Kapitel 7.4 zur Materialerkennung durchgeführt wurde, zeigte die Möglichkeiten, die mit einem solchen Vorgehen verbunden sind. Aufgrund der Datenlage konnte nur eine eingeschränkte originäre Merkmalsanzahl für die Extraktion der optimierten Sets herangezogen werden. Trotzdem liefern die Materialklassifikationen für beide Abschnitte des Liegenden Fehlerkriterien, die sich in den gleichen Größenordnungen wie für die alleinige Erkennung bewegen. Dies bedeutet, dass keine markante Verschlechterung der Zuordnung mit der Nutzung von universellen Merkmalssets einhergeht. Mit einer verbesserten Qualität der Rohmessdaten sollte es also möglich sein, nachzuweisen, dass eine definierte Anzahl an ausgewählten Filterradkombinationen zur Bestimmung der Trennflächen und damit auch der Materialerkennung in den kritischen Bereichen der Lagerstätte Hambach genutzt werden kann.

Die generelle Eignung der laserinduzierten Fluoreszenz als möglicher Bestandteil der zukünftigen Qualitätssteuerung bei der Gewinnung von Braunkohle konnte durch die in dieser Arbeit durchgeführte Merkmalsselektion und Materialklassifikation positiv bestätigt werden. Einschränkungen resultieren aus der zum einen unzureichenden Qualität der LIF-Messdaten und zum anderen aus der zu kleinen und damit nicht repräsentativen Größe der Datenbasis.

9 Ausblick

Obwohl eine Materialklassifikation und damit eine Trennflächenerkennung basierend auf der Messung der Fluoreszenzcharakteristika ausgewählter Bereiche der Lagerstätte Hambach mit denen in dieser Arbeit vorgestellten Algorithmen durchgeführt werden konnte, bedarf es einiger substanzieller Änderungen und Optimierungen, um dieses innovative optische Messverfahren zur Qualitätssteuerung erfolgreich in der Praxis bei der Rohstoffgewinnung einzusetzen.

Neben konstruktiven Änderungen am LIF Scanner / Analyser zur Minimierung der Messwertstreuungen bzw. der Nutzung komplett anderer Systemkomponenten zur Vermeidung von bisher aufgetretenen Störungen lässt sich auch mit der Optimierung der eingesetzten Algorithmen ein erhebliches Potential zur Verbesserung des Klassifikationserfolges freisetzen. Die Anforderungen an die Gerätetechnik werden am Ende dieses Abschnittes näher beleuchtet.

In den Bereich der Optimierung der eigentlichen Materialklassifikation gehört z.B. die kombinierte Betrachtung der Materialerkennung für unterschiedliche Klassifikationsalgorithmen. Die einzelnen Verfahren kennzeichnen individuelle Stärken und Schwächen, die weitestgehend durch die Datenqualität beeinflusst werden. So werden bei der Nutzung der linearen Regression teilweise komplette Klassen nicht erkannt, andere Stoßabschnitte jedoch sehr scharf ausgehalten. Eine Kombination von Mittelwertbasierten Verfahren und Ansätzen, die die nächsten Nachbarn zur Klassifikation von unbekanntem Messpunkten nutzen, soll an dieser Stelle exemplarisch vorgestellt werden (Abbildung 9-1).

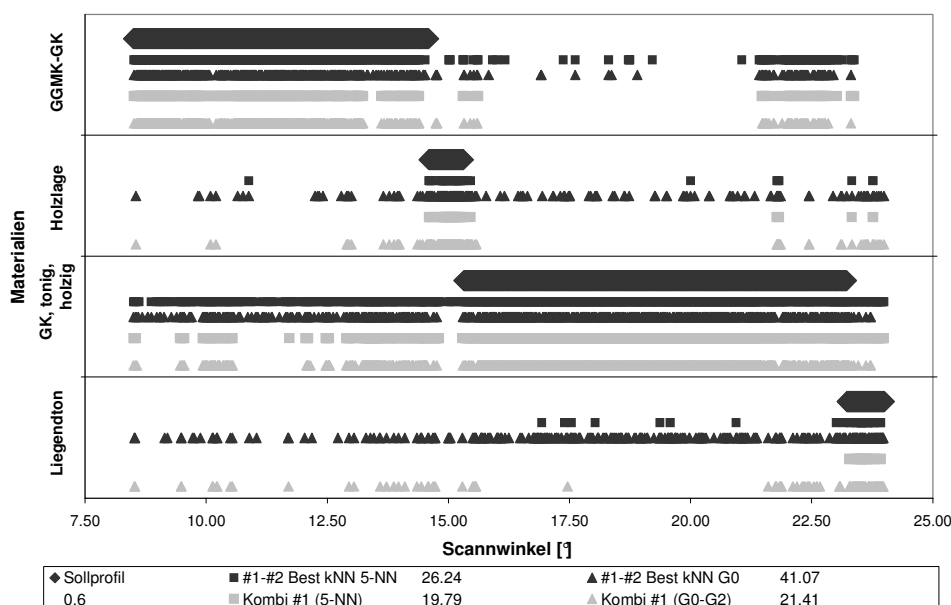


Abbildung 9-1: Profil der Merkmalsklassifikation, Kombination von Materialklassifikationen inklusive SGM, 5-NN und G0, 4 Materialien, #1-#2, nördlich Liegendrücken

In diesem Falle wurden die originären Profile für die Klassifikationsverfahren G0 und 5-NN des Scann #1-#2 basierend auf den sechs besten Merk-

9 Ausblick

malen des Sets „BEST k-NN“ mit einem gemeinsamen selektierenden Mittelwert geglättet. Die beiden als „Kombi“ bezeichneten Profile sind das Ergebnis der Glättung, dabei steht die Nomenklatur in der Klammer für das Profil, auf das im Falle der Nichterfüllung der SGM-Grenzbedingung verwiesen wird. Eine Korrektur der Materialklasse nach SGM bedarf eines Überschreitens des Entscheidungskriteriums, hier mehr als 60% (0,6) der angrenzenden Profilwerte innerhalb der Blende.

Mit der Kombination wird eine deutliche Verbesserung der Materialerkennung erreicht, so dass diese Möglichkeit der Reduktion des Fehlerkriteriums in Zukunft in Betracht gezogen werden sollte. Eine Kombination von bereits mittels SGM geglätteten Profilen hat sich als nicht sinnvoll erwiesen, da hier bereits die Fehlklassifikationen, die geologisch nicht plausibel sind, herausgefiltert wurden. Im geglätteten Profil liegen keine Ausreißer mehr vor, sondern nur massive Fehlzuordnungen, die mit diesem Glättungsalgorithmus nicht erkannt werden.

Um mit den vorhandenen Ergebnissen der Merkmalsselektion eine weitergehende Verbesserung der Materialklassifikation zu erreichen, besteht zudem die Möglichkeit, auf erweiterte Merkmalssets zurückzugreifen. Es handelt sich dabei im weiteren Sinne um eine Optimierung der „BEST“-Sets. Diese beschränken sich lediglich auf die in der Selektion am häufigsten und am gleichmäßigsten verteilten Merkmale. Dahinter stehen jedoch Fluoreszenzspektren, die durch die eben genannten Filtereinstellungen gemessen werden. Die im weiteren als „TOP“ bezeichneten Merkmalsvektoren umfassen hingegen neben den Merkmalen von „BEST“ auch alle durch die Kombination der Filtereinstellungen möglichen Merkmale, da diese durch die Messungen ebenfalls zur Verfügung stehen und unter Umständen eine Erhöhung des Informationsgehaltes bedeuten.

Entscheidend für den praktischen Einsatz der LIF mit der PMT-Technologie in der Rohstoffgewinnung ist eine Minimierung der Baugröße eines Analyserätes, die sich einfach über eine Einsparung an Empfangseinheiten erreichen lässt. Insofern kommt der Merkmalsreduktion eine erhebliche Bedeutung zu.

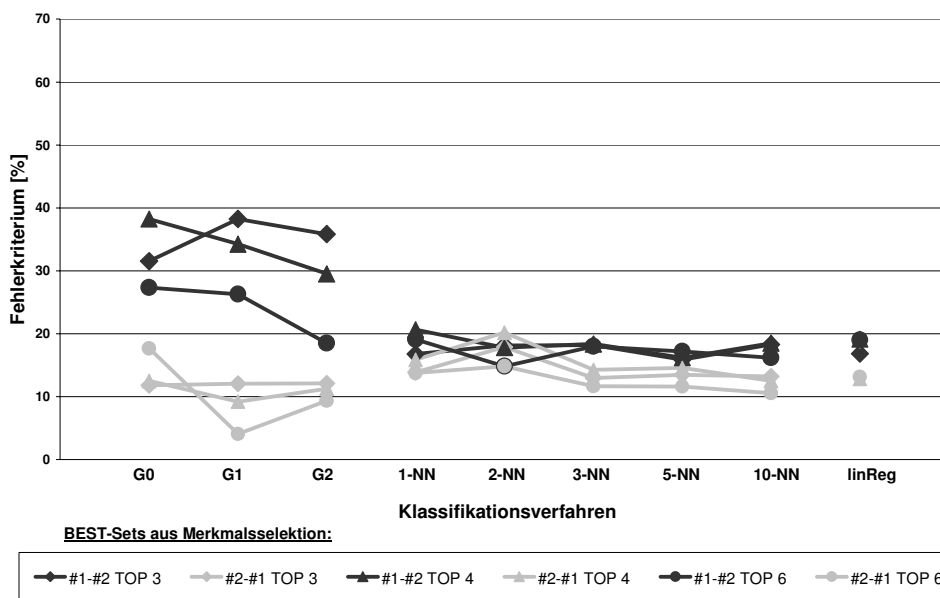
Das Potential der Erweiterung des Merkmalssets soll am Beispiel des Lagerstättenabschnitts nördlich des Liegendrückens für vier Materialklassen aufgezeigt werden. Die besten sechs Merkmale aus „BEST kNN“ und die Häufigkeit der entsprechenden Filtereinstellung zeigt Tabelle 9-1.

Rang	BEST k-NN	PMT-Einstellung	Häufigkeit
1	B 460 / A 420	A 400	1 x
2	B 440 / A 420	A 420	6 x
3	A 640 / A 420	A 560	1 x
4	A 560 / A 420	A 600	1 x
5	A 600 / A 420	A 640	1 x
6	A 420 / A 400	B 420	1 x
		B 440	1 x

Tabelle 9-1: Beste sechs Merkmale von „BEST k-NN“, nördlich Liegendrückens

Das Emissionsspektrum 420 nm auf PMT A ist in allen sechs Merkmalen vertreten, so dass die sechs Merkmale mit nur sieben gleichzeitig erfassten Emissionswellenlängen gebildet werden könnten. Die neuen Merkmalsvektoren wurden für die Kombinationen aus allen sechs, den besten vier sowie besten drei Merkmalen generiert. „TOP6“ besteht entsprechend der sieben Einstellungen aus 35 Merkmalen, „TOP4“ aus 20 und „TOP3“ aus 14 Merkmalen respektive fünf bzw. vier PMT Einstellungen. Die Reihenfolge der Merkmale bei der Erstellung des neuen Vektors erfolgt nach der gleichen Methodik wie bei den „BEST“-Vektoren, da auf die Daten der originären Merkmalsreduktion zurückgegriffen wird.

Mit diesen drei Merkmalsvektoren wurde die Materialklassifizierung für beide Blindproben durchgeführt. Die realisierten Fehlerkriterien der sechs Klassifikationen zeigt Abbildung 9-2. Die stabilsten Ergebnisse liefern die Verfahren mit k-NN sowie die lineare Regression, während die Gauß'schen Verfahren stark abweichende Ergebnisse für beide Blindproben aufweisen.



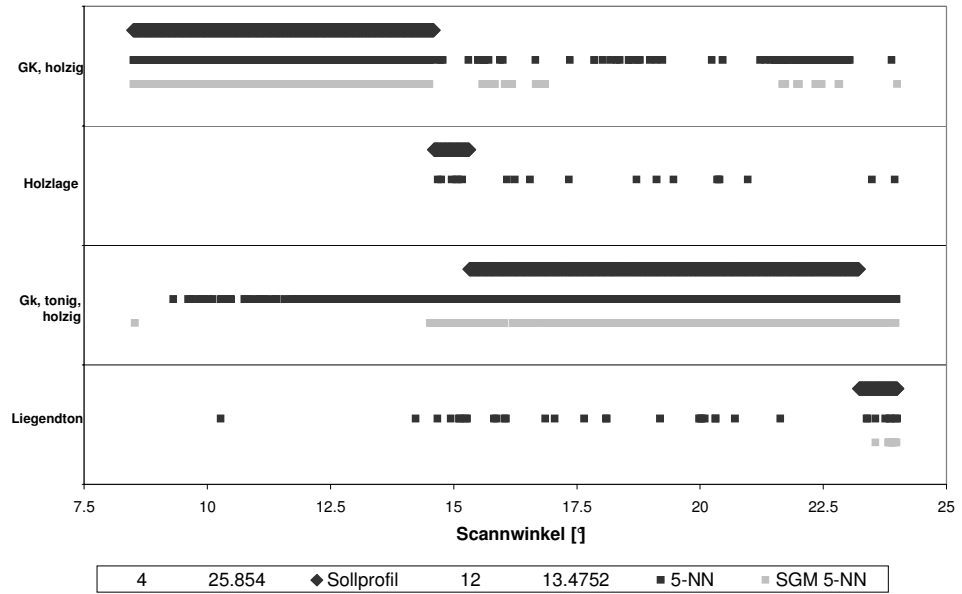
9 Ausblick

Abbildung 9-2: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens für erweiterte Merkmalssets TOP, 4 Materialien, nördlich Liegendrücken

Im Vergleich zur Materialerkennung mit „BEST“ ergibt sich auf den ersten Blick kaum eine Minimierung der Fehlzuordnungen: Entscheidend ist jedoch, dass bereits mit einer sehr geringen Anzahl an Filterradeneinstellungen Fehlerkriterien unter 15% realisiert werden (Abbildung 9-3). Bei der Nutzung von vier PMT Einstellungen kann die Holzlage zwischen den beiden Kohleklassen nach der SGM nicht identifiziert werden. Auch eine eindeutige Erkennung des Liegendtons wird auf diese Weise nicht erreicht.

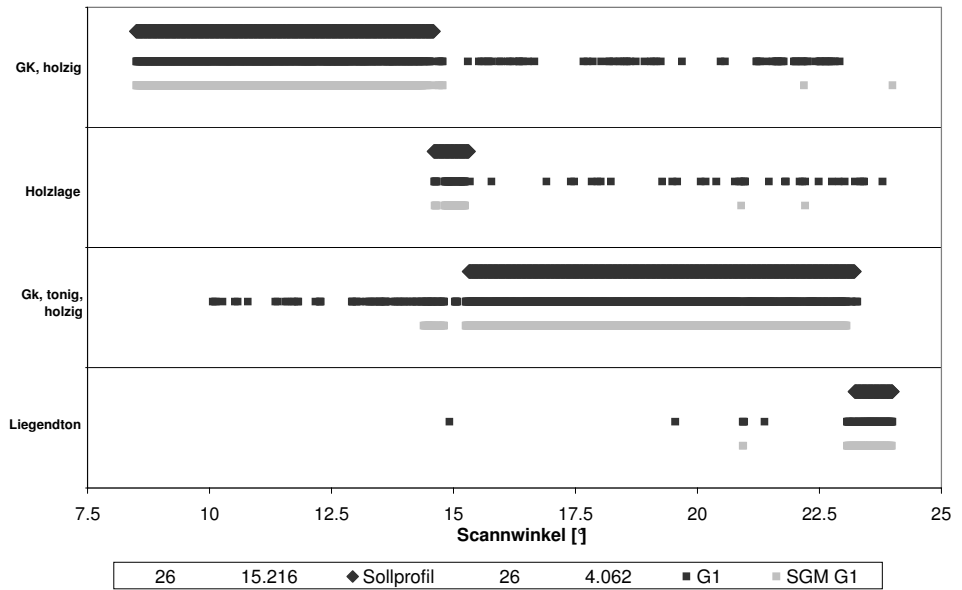
9 Ausblick

Abbildung 9-3:
 Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST TOP3“ inklusive SGM, 5-NN, 4 Materialien, #2-#1, nördlich Liegendrücken



Ein anderes Bild ergibt sich für „TOP6“, also sieben Filtereinstellungen, in Abbildung 9-4. Bereits das originäre Profil weist ein Fehlerkriterium von nur 15% auf, wobei 26 der 35 zu Verfügung stehenden Merkmale genutzt werden. Nach der Glättung kommt es lediglich bei 63 der 1.551 Messwerten zu einer fehlerhaften Materialerkennung, was das Potential der Erweiterung auf „TOP“-Merkmalssets widerspiegelt.

Abbildung 9-4:
 Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST TOP6“ inklusive SGM, G0, 4 Materialien, #2-#1, nördlich Liegendrücken



Auch eine Kombination der Fluoreszenzspektren beider Anregungswellenlängen könnte zu einer Verbesserung des Klassifikationsergebnisses beitragen. Aufgrund der Störungen der Messkampagnen liegen jedoch die erforderlichen Messreihen nicht vor, so dass zur Überprüfung des Potentials dieser Maßnahme weitere Feldversuche erforderlich sind.

Eine weitere Möglichkeit zur Optimierung bietet sich bei den k-NN-Verfahren durch die Vereinfachung der Trainingssets an, die im betrieblichen Einsatz als Grundlage für die Klassifikation der Messwerte in der Scannspur herangezogen werden. Durch das Editieren des Trainingssets werden redundante Trainingsprototypen aus der Datenbasis entfernt: Die-

ses auch als „Condensing“ bekannte Verfahren überprüft, ob Prototypen ausschließlich von Mitgliedern der eigenen Klasse umgeben sind. In diesem Fall enthält der entsprechende Prototyp keine neuen Informationen und wird eliminiert. So kann später die Geschwindigkeit der Klassifikation erheblich gesteigert werden. Bei diesem Vorgehen besteht allerdings das Problem, dass keine Trainingsdaten nachträglich hinzugefügt werden können, es bedarf daher aller Trainingsdaten vor der praktischen Durchführung der Materialerkennung.

Den zweiten Hebel zur Verbesserung der Qualität der Materialerkennung stellt die Optimierung der vorhandenen Gerätetechnik zur Reduktion der Messwertstreuungen dar. Dabei soll an dieser Stelle nur kurz auf diesen Aspekt eingegangen werden, da im Rahmen des laufenden DFG-Forschungsvorhabens „Bergmännische Nutzbarkeit der Grenzschieferkennung in mineralischen Lagerstätten basierend auf dem Prinzip der laserinduzierten Fluoreszenz“ bereits einige der Maßnahmen entweder umgesetzt oder konzeptionell erarbeitet wurden bzw. werden.

Einen wichtigen Punkt stellt die zuverlässige Implementierung der Entfernungskalibrierung des LIF Scanner / Analysers dar. Im Anschluss an die Messkampagnen bei der RWE Power AG wurde zusammen mit der Nachfolgegesellschaft des Messgeräteherstellers über Möglichkeiten der Verbesserung diskutiert. Neben der Vereinfachung und Automatisierung der Anpassung der Empfangsoptiken, um einen optimalen Fokus der PMT's bei unterschiedlichen Entfernungen zu garantieren, soll die für die mathematische Umrechnung des entfernungsabhängigen Intensitätsabfalls zuständige Splinefunktion zukünftig nicht mehr nur auf 50 sondern vielmehr 150 Kalibrierungswerte zurückgreifen. Auf diese Weise ließe sich eine bessere Interpolation realisieren, was die Schwankungen der Messwerte aufgrund wechselnder Entfernungen zwischen Stoß und LIF Scanner / Analyser abschwächen sollte.

Darüber hinaus wurde in späteren Messkampagnen an anderen mineralischen Rohstoffen im Technikum des BGMR festgestellt, dass die Ausgangsleistung der Laser weit über das vom Laserhersteller angegebene Maß hinaus schwankt. Anstelle von Puls-zu-Puls-Schwankungen unter 4,5% der Ausgangsleistung wurden vor allem bei 266 nm Anregungswellenlänge Streuungen von bis zu 25% beim punktuellen Beschuss homogenen Materials gemessen. In Folge dieser Messungen wurde der Laser vom Hersteller über einen Zeitraum von mehr als einem halben Jahr auf Produktionsfehler hin überprüft und nach einer Generalüberholung sowie einer intensiven Testphase im April 2004 wieder an das BGMR ausgeliefert. In Kombination mit einer in den LIF Scanner / Analyser integrierten automatischen Messung des Energieinhaltes eines jeden Laserpulses sollten in folgenden Messkampagnen sowohl frühzeitig ein Abfall bzw. nicht zulässige Schwankungen der Laserausgangsleistung erkannt, aber auch eine automatische Korrektur der Fluoreszenzsignale nach der A/D-Wandlung im LIF Scanner / Analyser umgesetzt werden können. Von der Umsetzung dieser Maßnahmen verspricht man sich eine erhebliche Reduktion der Varianz. Diese sollte zu einer Verbesserung der Datenqualität der Fluoreszenzmessungen um solche Größenordnungen beitragen, dass im Gegensatz zu den

9 Ausblick

Messungen im Tagebau Hambach 2002 / 2003 Zielgrößen des Fehlerkriteriums von unter 10% bei der Materialerkennung mittels LIF für alle Lagerstättenbereiche erreichbar sein sollten.

Darüber hinaus sollte, nachdem die Hauptursachen für die Messwertstreuungen mit den oben genannten Maßnahmen beseitigt worden sind, der Fokus neben der Ausweitung der Datenbasis ebenfalls wieder verstärkt auf den Einfluss von betrieblichen Störgrößen auf die laserinduzierte Fluoreszenz im bergmännischen Gewinnungsprozess gelegt werden, wie etwa Staub und Feuchtigkeit. In der am BGMR aufgebauten Klimakammer könnten die für die Auslegung eines praxistauglichen LIF Gerätes erforderlichen Parameter bestimmt und so die Anforderungen für ein mögliches Pflichtenheft formuliert werden.

Einen anderen Ansatz zur Umsetzung einer Trennflächenerkennung in der mineralischen Rohstoffgewinnung mittels laserinduzierter Fluoreszenz könnte ein völliger Systemwechsel bei der eingesetzten Empfangstechnologie mit sich bringen. Innerhalb der letzten Jahre ging die Entwicklung bei CCD-Chips und Diodenarrays hin zu immer lichtempfindlicheren, leistungstärkeren und kleineren Einheiten. Aufgrund der langsameren Zeitauflösung könnten dann jedoch nur eingeschränkt die Abklingcharakteristiken der untersuchten Materialien genutzt werden. Diese Empfangseinheiten würden jedoch das Problem der Beschränkung auf einzelne Bereiche des Spektrums, wie sie die PMT-Technologie mit sich bringt, ebenso lösen wie die Schwierigkeiten, die sich aus dem Platzbedarf der PMT's ergeben. Zur Bewältigung der resultierenden Datenflut kommt einer effektiven Merkmalsreduktion und schnellen Materialerkennung eine besondere Bedeutung zu, so dass die im Rahmen dieser Arbeit vorgestellten Algorithmen einen Lösungsansatz auch bei der Nutzung neuer Gerätetechnik bieten.

In diesem Zusammenhang wäre es unter Umständen sinnvoll, das ursprünglich für den LIF Scanner / Analyser vorgesehene Konzept des koaxialen Strahlengangs nochmals auf seine Eignung für die Materialerkennung hin zu untersuchen. Damals wurde vom Hersteller des Analysegerätes eine koaxiale Spiegelkonstruktion aufgrund von Steuerungsproblemen zugunsten des schwenkbaren Sende- und Empfangskopfes verworfen. Bei der Nutzung von CCD-Spektrographen hätte ein koaxiales Spiegelsystem aufgrund der einfacheren Entfernungskalibrierung der Empfangsoptiken und der erheblichen geringeren Abmaße einige Vorteile gegenüber der derzeit realisierten Ausführung vorzuweisen.

Die obigen Ausführungen zeigen, dass die Ergebnisse dieser Arbeit vor dem Hintergrund eines Prototypeneinsatzes einer in der Rohstoffgewinnung bis dato nur zu Testzwecken eingesetzten Technologie zu sehen sind und der laserinduzierten Fluoreszenz zur Materialerkennung im bergmännischen Gewinnungsprozess ein erhebliches Potential inne wohnt. Diese Arbeit legt damit den Grundstein für eine systematische Auswertung der Messdaten zur effektiven Nutzung dieser Online-Technik und liefert zudem Denkanstöße und Verbesserungspotentiale für zukünftige Arbeiten auf dem Gebiet der LIF-Analytik zur Materialerkennung.

10 Zusammenfassung

Die optische Analysemethode der laserinduzierten Fluoreszenz (LIF) bietet eine Möglichkeit, unterschiedliche Materialien zu erkennen und Trennflächen auch über weitere Entfernungen zwischen Analysegerät und Materialoberfläche zu detektieren, als es mit anderen Verfahren wie etwa Röntgenstrahlung oder radioaktiven Methoden zu realisieren wäre. Am Lehr- und Forschungsgebiet „Betriebsmittel für die Gewinnung mineralischer Rohstoffe (BGMR)“ der RWTH Aachen wird seit einigen Jahren untersucht, wie sich diese Methode für die Materialerkennung in der Rohstoffgewinnung nutzen lässt, um den Gewinnungsvorgang hinsichtlich der Qualitätssteuerung zu optimieren.

Die vorliegende Arbeit zeigt methodisch auf, wie signifikante Merkmale mit Hilfe von systematischen Suchalgorithmen aus dem Fluoreszenzspektrum extrahiert werden können. Im Gegensatz zu den bisherigen, eher willkürlichen Ansätzen, die vielfach nach dem „Trial & Error“-Prinzip arbeiteten, erlaubt diese Vorgehensweise eine bestmögliche Erkennung der Materialien. Durch die Anpassung und Optimierung von Klassifikations- und Selektionsalgorithmen, die in dieser Arbeit erstmals im Bereich der mineralischen Rohstoffe appliziert wurden, liegt nun ein Konzept vor, mit dem Fluoreszenzmesswerte erfolgreich zur Materialerkennung und Qualitätssteuerung eingesetzt werden können.

Eine Materialklassifikation und damit eine Trennflächenenerkennung basierend auf der Messung von Fluoreszenzcharakteristika der Lagerstätte Hambach konnte mit den in dieser Arbeit entwickelten Algorithmen durchgeführt werden.

Datengrundlage dieser Arbeit bilden LIF-Messwerte, die im Zuge eines Forschungsprojektes mit der RWE Power AG im Zeitraum zwischen Juni 2002 und April 2003 im Braunkohletagebau Hambach gewonnen wurden. Mit dem LIF Scanner / Analyser wurden signifikante Lagerstättenpartien am Übergang zum Liegenden sowie im Bereich des Zwischenmittels Frimmersdorf a/b untersucht. Es sollte geprüft werden, ob auf Basis dieser Technik eine Steuerung entwickelt werden kann, die den Fahrer des Schaufelradbaggers in die Lage versetzt, während des Gewinnungsvorganges in Echtzeit die Stratigraphie qualitativ unabhängig von für das menschliche Auge sichtbaren Qualitätsmerkmalen zu beurteilen.

Mittels materialspezifischen Fluoreszenzcharakteristika, den „Fingerabdrücken“, sollten Schichten, die farblich gleich erscheinen, wie etwa Braunkohle und dunkle Tone, unterschieden werden. Die Vermeidung von ungeplanten und unkontrollierten Verunreinigungen während des Gewinnungsprozesses durch eine auf der LIF-Technik basierenden Warnfunktion würde eine Optimierung des Lagerstättenausbringens durch das gezielte Heranfahren an die jeweilige Trennfläche bedeuten.

10 Zusammenfassung

Bei der LIF kann auf eine Vielzahl von Merkmalen zurückgegriffen werden, die aus dem Emissionsspektrum extrahiert werden. Im Falle des Zwischenmittels Frimmersdorf a/b liegen z.B. bei einer Anregungswellenlänge von 355 nm in den Rohdaten insgesamt 405 Merkmale vor. Diese Datenfülle erschwert jedoch erheblich die Automatisierung der Materialerkennung wie auch die Dimensionierung und Auslegung von praxistauglichen Analysegeräten. Der Selektion geeigneter Merkmale aus den zur Verfügung stehenden Rohdaten kommt daher eine besondere Bedeutung zu.

Bei der Extraktion der signifikantesten LIF-Merkmale wird der „Wrapper2“-Ansatz gewählt, um Varianzen bei der Merkmalsselektion aufgrund von Messwertstreuungen zu berücksichtigen und damit verbundene potentielle Instabilitäten der Ergebnisse bereits im Vorfeld auszuschließen. Grundsätzlich bedient sich dieser Ansatz einer statistischen Betrachtung der selektierten Parameter über eine definierte Anzahl von Wiederholungen des Selektionsprozesses bei leicht veränderten Trainingsdaten. Die eigentliche Merkmalsauswahl und Klassifikation erfolgt mit dem einfachen Wrapper-Ansatz, bei dem sich die Selektion und Materialzuordnung „umarmen“. Dieses Verfahren ermittelt in einem iterativen Vorgang die besten Kombinationen an Merkmalen, das Merkmalsset.

Während bisher in der mineralischen Rohstoffgewinnung ausschließlich Mittelwert basierte Verfahren zur Materialerkennung genutzt wurden, wird in dieser Arbeit neben diesen Gauß'schen Ansätzen auch das Verfahren des k-nächsten Nachbarn sowie die lineare Regression eingesetzt. Die Strategien, mit denen die optimalen Merkmale aus den Rohdaten extrahiert werden, erlauben effektive Suchvorgänge und Klassifikationen. Neben konventionellen sequentiellen Verfahren kommt auch die „Sequential Forward Floating Selection (SFFS)“ als optimierter Ansatz der zweiten Generation zum Einsatz. Für die eigentliche Durchführung der Merkmalsselektion und Materialklassifikation der Lagerstätte Hambach wird aufgrund der Resultate in Voruntersuchungen ausschließlich SFFS verwendet.

Neben der alleinigen Betrachtung der Fehlerrate fand auch die Praxistauglichkeit des Systems zur Materialerkennung in der Form Berücksichtigung, dass eine mehrstufige Modifikation und Optimierung des aus der Merkmalsselektion hervorgegangenen Merkmalssets durchgeführt wurde. Ein optimales Merkmalsset vereint eine geringe Fehlerrate mit einer minimalen Anzahl an erforderlichen PMT's, um eine praxistaugliche Dimensionierung eines LIF-Analysers zu gewährleisten. Das Motto „Keep it as simple as possible“ ist also die Leitlinie der Merkmalsselektion und Materialklassifikation, die mit einer möglichst geringen Anzahl von universell verwendbaren Filterradeneinstellungen erreicht werden sollte.

Aufgrund von Diskrepanzen zwischen den reduzierten Merkmalssets bei beiden Blindklassifikationen pro Lagerstättenabschnitt wurden die Merkmalsvektoren gesondert hinsichtlich ihrer Rangfolge analysiert und individuell in die Auswertung einbezogen. Die eigentliche Klassifikation wurde zusätzlich hinsichtlich der Optimierung der Materialerkennung nachbearbeitet, indem die Ergebnisse auf geologische Plausibilität hin geprüft wurden.

Zur Glättung des aus der Materialklassifikation resultierenden Profils wird in dieser Arbeit der selektierende gewichtete Mittelwert (SGM) herangezogen.

Neben der Bedeutung der Anzahl der Materialklassen wurde in dieser Arbeit der Einfluss der unterschiedlichen Algorithmen zur Merkmalsselektion und Materialklassifikation auf das Zuordnungsergebnis untersucht. Als ein Ergebnis kann hervorgehoben werden, dass sich keine Methode als optimal erwiesen hat bzw. basierend auf den vorliegenden Daten kein Verfahren bei der Klassifikation unter Berücksichtigung des Rechenaufwandes und der Zuordnungszuverlässigkeit eindeutig zu präferieren ist.

Als wesentliches Ergebnis der Analyse der Klassenanzahl auf den Erfolg der Materialerkennung lässt sich festhalten, dass eine zu detaillierte Klasseneinteilung nicht zu praktisch nutzbaren Ergebnissen führt, da die Fehlerquoten im Vergleich zu allgemeiner gehaltenen Materialeinteilungen unbefriedigend hoch sind. Fehlerkriterien unter 40% wurden bei den durchgeführten Messreihen nur in Ausnahmefällen erreicht. Konsequenterweise fand die Merkmalsselektion und Materialklassifikation ausschließlich für die Einteilung der gescannten Lagerstättenpartien in vier Klassen statt. Damit lässt sich eine Klassifikation der bauwürdigen Kohle und des Liegendtons im Bereich nördlich des Liegendrückens mit hoher Zuverlässigkeit mittels der mit laserinduzierter Fluoreszenz ermittelten Spektren durchführen. Bei der Klassifikation der 1.551 Messpunkte wurden so mit 5- bzw. 10-NN weniger als 10% Fehlzugeordnungen erreicht. Alle Trennflächen und Übergänge zwischen den bemusterten Materialien sind damit sehr gut zu unterscheiden und eine Nutzung der LIF zur Trennflächenenerkennung sinnvoll.

Für den zweiten Lagerstättenabschnitt am Liegendübergang lässt sich das positive Ergebnis der Materialerkennung nur bedingt bestätigen. Es ist zwar möglich, eine Materialzuordnung durchzuführen, die eine Unterscheidung zwischen den Klassen erlaubt, jedoch genügt sie mit minimalen Fehlerkriterien um 30%. nicht den Ansprüchen der Qualitätssteuerung an eine Klassifikationsmethode. Im Abschnitt des Zwischenmittels Frimmersdorf a/b sind die Zuordnungsergebnisse schlechter als im Bereich des Liegenden. Grundsätzlich eignen sich aber auch hier die verwendeten Algorithmen zur Klassifikation der gescannten Materialien.

Eine Voraussetzung für den Einsatz der laserinduzierten Fluoreszenz zur Trennflächenenerkennung in der mineralischen Rohstoffgewinnung ist die universelle Tauglichkeit eines Analysegerätes innerhalb der Lagerstätte. Dies macht es erforderlich, dass mit einem definierten Set an Filtereinstellungen und damit ausgewählten Emissionsspektralbanden unterschiedliche Profile respektive Materialzusammensetzungen einheitlich erfolgreich klassifiziert werden können.

Für beide Abschnitte im Liegenden der Lagerstätte Hambach wurde aus den vorliegenden LIF-Messwerten ein gemeinsamer Merkmalsvektor für die Materialklassifikation generiert. Das Fazit der Klassifikation mit diesen übergreifenden Merkmalen ist, dass sich prinzipiell eine universelle

10 Zusammenfassung

Vorgehensweise realisieren lässt. Grundlage für einen Praxiseinsatz sind jedoch neben der Güte der Messwerte ebenso die Erhöhung an universell einzusetzenden Merkmalen und die Größe der Datenbasis. Zur Minimierung der Varianzen innerhalb der einzelnen Materialklassen bedarf es vor allem der Umsetzung des letztgenannten Aspektes.

Diese Arbeit zeigt, dass mit der Verwendung einer optimierten Methodik bei der Durchführung der Merkmalsselektion und Materialklassifikation selbst bei der qualitativ nicht-optimalen LIF-Datenbasis eine Trennflächenerkennung in den untersuchten Lagerstättenabschnitten des Tagebaus Hambach möglich ist. Es konnte jedoch nicht nachgewiesen werden, dass sich ein Algorithmus besonders für diese Aufgabenstellung eignet. Bei weiterführenden Untersuchungen sollte eine Kombination der Verfahren zur Merkmalsselektion und Materialklassifikation Verwendung finden.

Die Grundlagen für eine erfolgreiche Grenzschichterkennung mittels LIF bilden die in dieser Arbeit entwickelten Klassifikationsalgorithmen, die eine unkomplizierte Zuordnung der Messwerte in für die Qualitätssteuerung relevante Materialklassen ermöglichen. Die optische Analyseverfahren der laserinduzierten Fluoreszenz ermöglicht dabei eine qualitative Materialerkennung und Trennflächendetektierung bereits am Stoß. Doch sind noch weitere Anstrengungen erforderlich, um einerseits die Messtechnik für den Praxiseinsatz bei der Online-Analyse im Gewinnungsprozess zu ertüchtigen, als auch andererseits eine endgültige Festlegung auf die für eine erfolgreiche Materialerkennung erforderlichen Emissionsspektren vornehmen zu können. Mit der vorliegenden Datenbasis kann eine solche Entscheidung zum jetzigen Zeitpunkt noch nicht erfolgen, da die Zuverlässigkeit der Klassifikation für die unterschiedlichen Lagerstättenbereiche noch nicht für den praktischen Einsatz in allen Lagerstättenpartien des Tagebaus Hambach ausreichend ist.

Ausgehend von der vorgestellten Methodik zur Merkmalsselektion und Materialklassifikation legt diese Arbeit den Grundstein für eine systematische Auswertung von Messdaten zur effektiven Nutzung der LIF-Technik. Darüber hinaus soll sie Denkanstöße und Verbesserungspotentiale für zukünftige Arbeiten auf dem Gebiet der LIF-Analytik zur Materialerkennung liefern. Die Ausführungen zeigen, dass die Ergebnisse dieser Arbeit vor dem Hintergrund eines Prototypeneinsatzes einer in der Rohstoffgewinnung bis dato nur zu Testzwecken eingesetzten Technologie zu sehen sind und der laserinduzierten Fluoreszenz zur Materialerkennung im bergmännischen Gewinnungsprozess ein erhebliches Potential inne wohnt.

Literaturverzeichnis

APB, 2000	AMERSHAM PHARMACIA BIOTECH	Fluorescence Imaging: Principles and Methods – Technical Manual, 2000, www.apbiotech.com
BAY, 2002	BAYER, A.K.	Messkampagne vom 25. bis 27.6.2002 im Tagebau Hambach der RWE Rheinbraun AG – Untersuchungen zur Ausweitung der lagerstättenbezogenen LIF-Datenbasis und Ermittlung des qualitativen Einflusses von Störgrößen, Juli 2002, unveröffentlichter Bericht
BAY, 2003	BAYER, A.K., VINZELBERG, G.	Trennflächenerkennung mittels laserinduzierter Fluoreszenz (LIF), Juli 2003, unveröffentlichter Abschlussbericht
BAY, 2004	BAYER, A.K., POTHS, M., VINZELBERG, G.	Innovative Approaches to Selective Lignite Mining by Bucket-Wheel Excavators - Boundary detection using laser-induced fluorescence (LIF), Surface Mining (Braunkohle & Other Minerals), voraussichtlich 2004, akzeptierter, noch nicht veröffentlichter Aufsatz
BEI, 2003	BEICHEL, R.	Bildverarbeitung und Mustererkennung, Vorlesungsunterlagen, WS 2003/2004, Graz, www.icg.tu-graz.at/~Education
BEI, 2002	BEISSNER, H., KULIK, L.	Planung und Steuerung von Braunkohletagebauen im liberalisierten Strommarkt, Surface Mining (Braunkohle & Other Minerals), 2002, Nummer 2, S. 150-160
BRO, 2000	BROICHER, H.	Bulk sorting by LIF: quality control of ores for bulk sorting and blending by laser induced fluorescence, Mai 2002, Mining Engineering
DUD, 2000	DUDA, R., HART, P., STORK, D.	Pattern Classification – Second Edition 2000, John Wiley & Sons, Inc., New York
DUN, 2002	DUNNE, K.; CUNNINGHAM, P. AZUAJE, F.	Solutions to Instability Problems with Sequential Wrapper-based Approaches to Feature Extraction, http://citeseer.nj.nec.com/cache/papers/cs/26841/http%3A%3Awww.cs.tcd.ie%2Fpublications%2Ftech-reports%2FzSzreports.02zSzTCD-CS-2002-28.pdf/dunne02solutions.pdf
FUC, 2003	FUCHS, S.	Entwicklung eines automatischen Klassifikationsverfahrens zur Materialerkennung bei der Gewinnung von Braunkohle auf der Basis von Fluoreszenzsignalen, Februar 2003, Diplomarbeit am BGMR
GÄR, 2003a	GÄRTNER, D., MARX, H.	Sonderbetrieb im Tagebau Hambach zur Ergänzung einer qualitätsorientierten Braunkohlegewinnung, Surface Mining (Braunkohle & other Minerals), 2003, Nummer 3 Juli/September, S. 256 ff.
GÄR, 2003b	GÄRTNER, D., GUDER, W.	25 Jahre Tagebau Hambach Von der Projektplanung zu einem leistungsstarken und umweltverträglichen Braunkohletagebau, Bergbau, 2003, Nummer 9, S. 388 ff.
HEM, 2003	HEMPEL, R.-J., KULIK, L.	Planung und Steuerung des Tagebau Hambachs, Bergbau, 2003, Nummer 9, S. 398 ff.
HOF, 2002	HOFFMANN, P.	Einsatzbereiche der laserinduzierten Fluoreszenz – Lichtblicke im Ultravioletten, Juni 2002, Seminararbeit

		am BGMR
KEY, 2002	KEYSERS, D., OCH, F.J.; NEY, H.	Maximum Entropy and Gaussian Models for Image Object Recognition. In DAGM 2002, Pattern Recognition, 24th DAGM Symposium, Zürich, Switzerland, Volume LNCS 2449 of Lecture Notes in Computer Science, Springer-Verlag, S. 498-506, September 2002
KEY, 2004	KEYSERS, D., MACHEREY, W.; NEY, H., DAHMEN, J.	Adaptation in Statistical Pattern Recognition Using Tangent Vectors. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 26, Number 2, S. 269-274, February 2004
KIN, 2003	KINDLICH, T.	Mit Glasfaser und CCD-Kameras zur Materialerkennung ?, Juni 2003, Seminararbeit am BGMR
KOV, 1997	KOHAVI, R., JOHN, G. H.	Wrappers for Feature Subset Selection, 20.05.1997, http://robotics.stanford.edu/~{ronnyk,gjohn}
LIF, 2002	LIF GMBH	Handbuch LIF Scanner / Analyser Typ 2-1-2
LIN, 2003	LINDENAU, J.	Qualitätsmanagement beim Materialtransport durch „Online Real time Analyse“, April 2003, Studienarbeit am BGMR
MAR, 1979	MARFUNIN, A.S.	Spectroscopy, Luminescence and Radiation centres in Minerals, 1979, Springer Verlag, New York
NIE, 1983	NIEMANN, H.	Klassifikation von Mustern, 1983, Springer Verlag, Berlin Heidelberg
NIE, 2002	NIENHAUS, K., BAYER, A.K., WEBER, S.	Beeinflussende Parameter bei der Anwendung der laserinduzierten Fluoreszenz – Betriebliche Studien in einem Eisenerzbergwerk, März 2003, Erzmetall, S. 158 ff.
NIE, 2003a	NIENHAUS, K., BAYER, A.K.,	Innovative Systems for Horizon Control of Mining Machines by means of Laser induced Fluorescence (LIF), April 2003, Konferenzbeitrag auf dem 12. internationalen Symposium MPES 2003 (Mine Planning and Equipment Selection), Kalgoorlie, Australien
NIE, 2003b	NIENHAUS, K., BAYER, A.K., RIEDEL, F.	Dilution Control by means of Laser-Induced Fluorescence (LIF) - Improving Run of Mine Ore Quality by innovative Rock Identification Methods, August 2003, African Mining Review, S. 44 - 45
NIE, 2003c	NIENHAUS, K., BAYER, A.K., RIEDEL, F.	Quality of LIF – Research and testing is opening up new frontiers for the application of laser-induced fluorescence in the mining industry, September 2003, Australian Mining Monthly, S. 68 ff.
RIE, 2002	RIEDEL, F.	Technical and Financial Evaluation of Dilution Control by means of laser induced Fluorescence (LIF) – A case study conducted at Premier Diamond Mine, South Africa, Dezember 2002, Diplomarbeit am BGMR
SOM, 1999	SOMOL, P.; PUDIL, P.; NOVOVICOVA, J. PACLIK, P.	Adaptive Floating Search Methods in Feature Selection, 1999, in Pattern Recognition Letters, Elsevier Science BV., Amsterdam
THA, 2002	THALES	Betriebsanweisung für den Laser DIVA II der Firma THALES LASER, 2002, Paris
WEB, 1998	WEBER, M.J..	Selected paper on photoluminescence of inorganic

Klassifikationsverfahren zur Materialerkennung

			solids, 1998, SPIE milestone series, Vol. MS 150
WEB, 2001	WEBER, S.		Quality control of iron ore by means of laser induced fluorescence (LIF) – conception, realisation and analysis of laboratory and operational field tests at Kiruna Mine, Sweden, September 2001, Diplomarbeit am BGMR
XIO, 2001	XIONG, M.; FANG, X.; ZHOA, J.		Biomarker Identification by Feature Wrapper, 2001, www.genome.org, S. 1178 ff.

Literaturverzeichnis

Abbildungsverzeichnis

Abbildung 2-1: Darstellung des quantenmechanischen Reaktionen bei der Fluoreszenz: Anregung (1), vibrationale Entspannung (2) und Emission (3) [APB, 2000].....	5
Abbildung 2-2: Darstellung der Energieniveaus eines fluoreszierenden Moleküls [APB, 2000].....	6
Abbildung 2-3: Absorptions- und Emissionsspektren [BRO, 2000]	7
Abbildung 2-4: Abklingverhalten verschiedener Evaporite.....	8
Abbildung 2-5: Merkmalsbildung zur Materialerkennung mittels LIF.....	11
Abbildung 2-6: Schematische Darstellung des Messprinzips und generellen Aufbaus eines LIF Analysegerätes [RIE,2002].....	12
Abbildung 2-7: Optisches System: Laser – harmonische Generatoren [THA, 2002]	13
Abbildung 2-8: Kurz-, Lang -und Bandpassfilter	14
Abbildung 2-9: Stationärer LIF-Analyser im Einsatz in einem Kalksteinbruch.....	17
Abbildung 2-10: Schematische Darstellung einer LIF-Analysers zur Steuerung eines Förderstroms [NIE, 2003c]	17
Abbildung 2-11: LIF Scanner / Analyser im Technikum des BGMR.....	18
Abbildung 2-12: Empfangsseite des LIF Scanner / Analysers [LIF; 2002].....	19
Abbildung 2-13: Zeitfenster W1, W2 und W3 zur Ermittlung der Intensitätsintegrale	20
Abbildung 3-1: Mögliche Einsatzkonfiguration eines LIF Scanners im Braunkohlebergbau	22
Abbildung 3-2: Einbettung des Tagebaus Hambach ins rheinische Braunkohlerevier [HEM, 2003]	23
Abbildung 3-3: Vereinfachter geologischer Schnitt des Tagebaus Hambach [GÄR, 2003b]	25
Abbildung 3-4: Scheibeneinteilung am Kohlestoß gemäß Kohlequalitätsschlüssel [GÄR, 2003b].....	26
Abbildung 3-5: Messorte im Tagebau Hambach.....	27
Abbildung 3-6: Versuchsaufbau des LIF Scanner / Analysers	28
Abbildung 3-7: Profil des Scannbereichs am Liegendübergang, nördlich Tiefstufe, 12.11.2002	29
Abbildung 3-8: Verlauf der LIF-Intensitäten am Liegendübergang, nördlich Liegendrücken, 12.11.2002.....	30
Abbildung 3-9: Verlauf der Quotienten und Abklingkurven am Liegendübergang, nördlich Liegendrücken, 12.11.2002	30
Abbildung 3-10: Liegendübergang, nördlich Liegendrücken, Clusterbildung, 12.11.2002.....	31
Abbildung 3-11: Verlauf der LIF-Intensitäten am Liegendübergang, Tiefstufe 12.11.2002	31
Abbildung 3-12: Profil des Scannbereichs am Zwischenmittel Frimmersdorf a/b, 09.04.2003	32
Abbildung 3-13: Verlauf der LIF-Intensitäten am Zwischenmittel Frimmersdorf a/b 09.04.2003	33

Abbildung 3-14: Streuung der Merkmale für ausgewählte Materialien, Zwischenmittel Frimmersdorf a/b, 09.04.2003	34
Abbildung 4-1: Aufbau eines Systems zur Mustererkennung [nach DUD, 2000].....	39
Abbildung 4-2: Aufbau eines Klassifikators.....	41
Abbildung 4-3: Flowchart des Ablaufs der Klassifikation nach dem Gaußverfahren	43
Abbildung 4-4: Voronoi-Polygone für 1-NN-Verfahren [BEI; 2003].....	44
Abbildung 4-5: Klassifikation bei einem 10-NN-Verfahren	45
Abbildung 4-6: Flowchart des Ablaufs der Klassifikation nach dem k-Nächster Nachbar.....	46
Abbildung 4-7: Beispiel für ein lineares Klassifikationsverfahren	48
Abbildung 4-8: Flowchart des Ablaufs der Klassifikation nach dem Verfahren der linearen Regression	49
Abbildung 4-9: Prinzip des Wrapper Ansatzes [nach KOV,1997].....	51
Abbildung 4-10: Flowchart des Ablaufs der Merkmalsselektion nach SFS	52
Abbildung 4-11: Flowchart des Ablaufs der Merkmalsselektion nach SBS	53
Abbildung 4-12: Flowchart des Ablaufs der Merkmalsselektion nach SFFS	54
Abbildung 4-13: Flowchart des Ablaufs der Merkmalsselektion nach Wrapper-2 (a)	56
Abbildung 4-14: Flowchart des Ablaufs der Merkmalsselektion nach Wrapper-2 (b)	57
Abbildung 5-1: Nomenklatur des Aufrufs von wrapper2.m	59
Abbildung 5-2: Histogramm (normierte Häufigkeit) T20-T100 für SFS 1-NN, D25	61
Abbildung 5-3: Histogramm (normierte Häufigkeit) T20-T100 für SFS 1-NN, Merkmal A600 W1/W3, D25	61
Abbildung 5-4: Rangfolge nach T100 für SFS 1-NN, D25.....	62
Abbildung 5-5: Summe der Rangabweichungen T20-T50 gegen T100 für SFS 1-NN, D25	62
Abbildung 5-6: Summe der Rangabweichungen T20-T50 gegen T100 für SFS, linReg, D25.....	63
Abbildung 5-7: Rechendauer T20-T100 für SFS, D25	64
Abbildung 5-8: Summe der Rangabweichungen D5-D20 gegen D25 für SFFS und SFS, Gaußverfahren, T40	65
Abbildung 5-9: Summe der Rangabweichungen D5-D20 gegen D25 für SFFS und SFS, 1-NN, T40	65
Abbildung 5-10: Summe der Rangabweichungen D15 und D20 gegen D25 für SFFS, SFS und SBS, linReg, T40.....	65
Abbildung 5-11: Rechendauer D5 bis D25 für SFFS und SFS, T40.....	66
Abbildung 5-12: Rechendauer D5 bis D25 für SBS, T40	66
Abbildung 5-13: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, Gaußverfahren 0, T40.....	67
Abbildung 5-14: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, 1-NN-Verfahren, T40.....	68
Abbildung 5-15: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, lineares Regressionsverfahren, T40	68

Abbildung 6-1: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, 5-NN, T40, nördlich Liegendrücken	72
Abbildung 6-2: Minimales Fehlerkriterium und korrelierende Größe des reduzierten Merkmalssets für #1-#2 und #2-#1, T40, nördlich Liegendrücken	72
Abbildung 6-3: Fehlerkriterium abhängig von der Größe des reduzierten Merkmalssets, 5-NN, T40, nördlich Liegendrücken, Überkreuzkorrelation.....	73
Abbildung 6-4: Häufigkeit der Wahl einzelner Merkmale, Top10 für #1-#2 sowie #2-#1 5-NN, T40, nördlich Liegendrücken.....	74
Abbildung 6-5: Häufigkeit der Wahl einzelner Merkmale, Top10 alle Algorithmen für #1-#2 sowie #2-#1, Gaußverfahren, T40, nördlich Liegendrücken.....	75
Abbildung 6-6: Darstellung der Rangfolge der optimierten Merkmalssets „BEST“, sortiert nach „BEST alle“, T40, nördlich Liegendrücken.....	76
Abbildung 6-7: Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST alle“, 5-NN, T40, nördlich Liegendrücken	76
Abbildung 6-8: Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST k-NN“, 5-NN, T40, nördlich Liegendrücken.....	77
Abbildung 6-9: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST alle“, 11 Merkmale, 5-NN, T40, nördlich Liegendrücken	78
Abbildung 6-10: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST k-NN“, 6 Merkmale, 5-NN, T40, nördlich Liegendrücken	78
Abbildung 6-11: Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST alle“ inklusive SGM, 5-NN, T40, nördlich Liegendrücken.....	80
Abbildung 6-12: Fehlerkriterium abhängig von der Größe des reduzierten optimierten Merkmalssets „BEST k-NN“ inklusive SGM, 5-NN, T40, nördlich Liegendrücken.....	81
Abbildung 6-13: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST alle“ inklusive SGM, 11 Merkmale, 5-NN, T40, nördlich Liegendrücken.....	81
Abbildung 6-14: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST k-NN“ inklusive SGM, 6 Merkmale, 5-NN, T40, nördlich Liegendrücken.....	82
Abbildung 7-1: Vergleich der selektierten „BEST k-NN“ Merkmale für 4 und 9 Materialien, nördlich Liegendrücken.....	85
Abbildung 7-2: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 9 Materialien, nördlich Liegendrücken	85
Abbildung 7-3: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 9 Materialien, nördlich Liegendrücken.....	86

Abbildung 7-4: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST 1-NN“ inklusive SGM, 10-NN, 9 Materialien, nördlich Liegendrücken	86
Abbildung 7-5: Vergleich der Merkmale für 4 und 6 Materialien G0-G2, Zwischenmittel, 355 nm	88
Abbildung 7-6: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 6 Materialien, Zwischenmittel, 355 nm	88
Abbildung 7-7: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST alle“ inklusive SGM, G0, 6 Materialien, #1-#2, Zwischenmittel, 355 nm	89
Abbildung 7-8: Vergleich der Rangfolge der selektierten „BEST “ Merkmale für 4 Materialien, nördlich Liegendrücken	91
Abbildung 7-9: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, nördlich Liegendrücken	91
Abbildung 7-10: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, nördlich Liegendrücken	92
Abbildung 7-11: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST k-NN“ inklusive SGM, 5-NN, 4 Materialien, #2-#1, nördlich Liegendrücken	93
Abbildung 7-12: Vergleich der Rangfolge der selektierten „BEST “ Merkmale für 4 Materialien, Liegendes Tiefstufe	94
Abbildung 7-13: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, Liegendes Tiefstufe	94
Abbildung 7-14: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, Liegendes Tiefstufe	95
Abbildung 7-15: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST linReg“ inklusive SGM, 1-NN, 4 Materialien, #1-#2, Liegendes Tiefstufe	95
Abbildung 7-16: Vergleich der Rangfolge der selektierten „BEST “ Merkmale für 4 Materialien, Zwischenmittel. 266 nm	97
Abbildung 7-17: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, Zwischenmittel, 266 nm	98
Abbildung 7-18: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, Zwischenmittel, 266 nm	98
Abbildung 7-19: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST linReg“ inklusive SGM, 10-NN, 4 Materialien, #1-#2, Zwischenmittel, 266 nm	99
Abbildung 7-20: Vergleich der Rangfolge der selektierten „BEST “ Merkmale für 4 Materialien, Zwischenmittel. 355 nm	100

Abbildung 7-21: Vergleich der minimalen Fehlerkriterien für „BEST“ Merkmalssets, 4 Materialien, Zwischenmittel, 355 nm.....	100
Abbildung 7-22: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens, 4 Materialien, Zwischenmittel, 355 nm.....	101
Abbildung 7-23: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST G0-G2“ inklusive SGM, 10-NN, 4 Materialien, #1-#2, Zwischenmittel, 355 nm.....	101
Abbildung 7-24: Vergleich der minimalen Fehlerkriterien für universelle BEST Merkmalssets, Liegendes.....	104
Abbildung 7-25: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens für universelle Merkmalssets, 4 Materialien, Liegendes.....	104
Abbildung 8-1: Vergleich der minimalen Fehlerkriterien für BEST Merkmalssets in Abhängigkeit der Lokation, Tagebau Hambach.....	108
Abbildung 8-2: Vergleich der minimalen Fehlerkriterien in Abhängigkeit des Klassifikationsverfahrens bei jeweils optimalem „BEST“ - Merkmalsset, Tagebau Hambach.....	109
Abbildung 9-1: Profil der Merkmalsklassifikation, Kombination von Materialklassifikationen inklusive SGM, 5-NN und G0, 4 Materialien, #1-#2, nördlich Liegendrücken.....	111
Abbildung 9-2: Fehlerkriterien in Abhängigkeit der Merkmalsselektion und des Klassifikationsverfahrens für erweiterte Merkmalssets TOP, 4 Materialien, nördlich Liegendrücken.....	113
Abbildung 9-3: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST TOP3“ inklusive SGM, 5-NN, 4 Materialien, #2-#1, nördlich Liegendrücken.....	114
Abbildung 9-4: Profil der Merkmalsklassifikation, optimiertes Merkmalsset „BEST TOP6“ inklusive SGM, G0, 4 Materialien, #2-#1, nördlich Liegendrücken.....	114

Tabellenverzeichnis

Tabelle 6-1: Übersicht der durchgeführten Merkmalsselektionen und Materialklassifikationen im Tagebau Hambach	70
Tabelle 6-2: Ausschnitt der Selektionstabelle zur Ermittlung von „BEST k-NN“, T40, nördlich Liegendrücken	75
Tabelle 6-3: Beispielhafter Einsatz des SGM zur Optimierung der Materialklassifikation, 4 Materialien, Liegend nördlich Tiefstufe	79
Tabelle 7-1: Rangfolge der universellen selektierten „BEST “ Merkmale für 4 Materialien, Liegendes, 266 nm	103
Tabelle 9-1: Beste sechs Merkmale von „BEST k-NN“, nördlich Liegendrücken	112

Anhang A: LIF Scanner / Analysers, Typ Sc 2-1-2

A 1: Technische Daten und Anschlüsse

LIF Analyser / Scanner

ohne Beine

Abmessungen B x H x T (mm)	1200 x 1100 x 1000
Raumbedarf B x H x T (mm)	1200 x 1100 x 1000
Gewicht (kg)	ca. 240 kg

mit Beinen

Abmessungen B x H x T (mm)	1330 x 2185 x 1000
Raumbedarf B x H x T (mm)	1330 x 2185 x 1000
Gewicht (kg)	ca. 320 kg

Werkstoff, Gehäuse und Beine	Stahl, rostfrei, blank
Werkstoff, Scanner-Messkopf	Stahl und Aluminium, elox.
Energieversorgung	über den Steuerschrank
Schutzart	IP 54
Umgebungstemperatur	+5 °C bis +40 °C
Luftfeuchtigkeit	10 % bis 75%

Schaltschrank

Hersteller	Rittal
Typ	Kompakt-Schaltschrank
Modell	1648.600
Abmessungen B x H x T (mm)	800 x 1200 x 400
Raumbedarf B x H x T (mm)	1200 x 1200
Werkstoff	Stahl lackiert
Gewicht	ca. 200 kg
Energieversorgung	1 x 230 V / N / PE / 50 Hz
Sicherung der Zuleitung	max. 16 A
Schutzart	IP 54

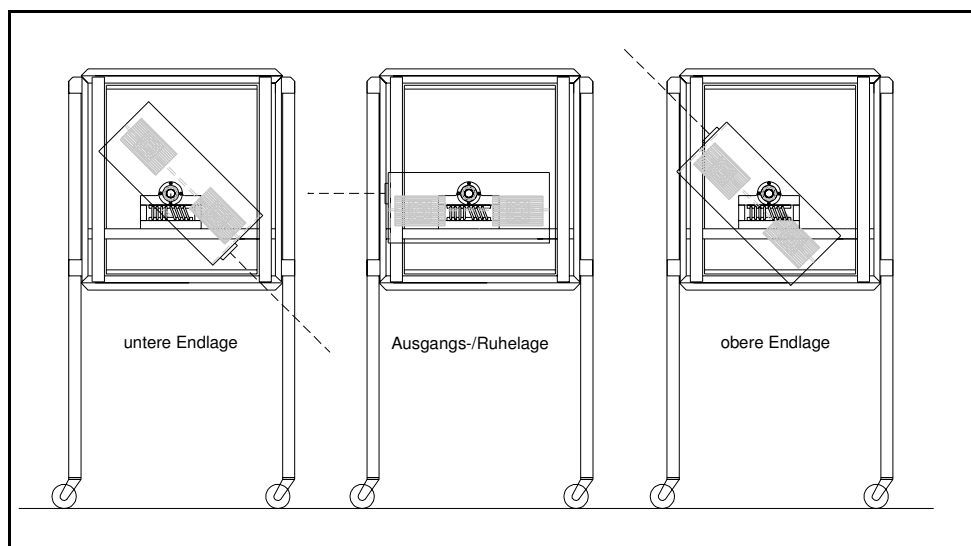
Klassifikationsverfahren zur Materialerkennung

Umgebungstemperatur	+5 °C bis +40 °C
Luftfeuchtigkeit	10 % bis 75%

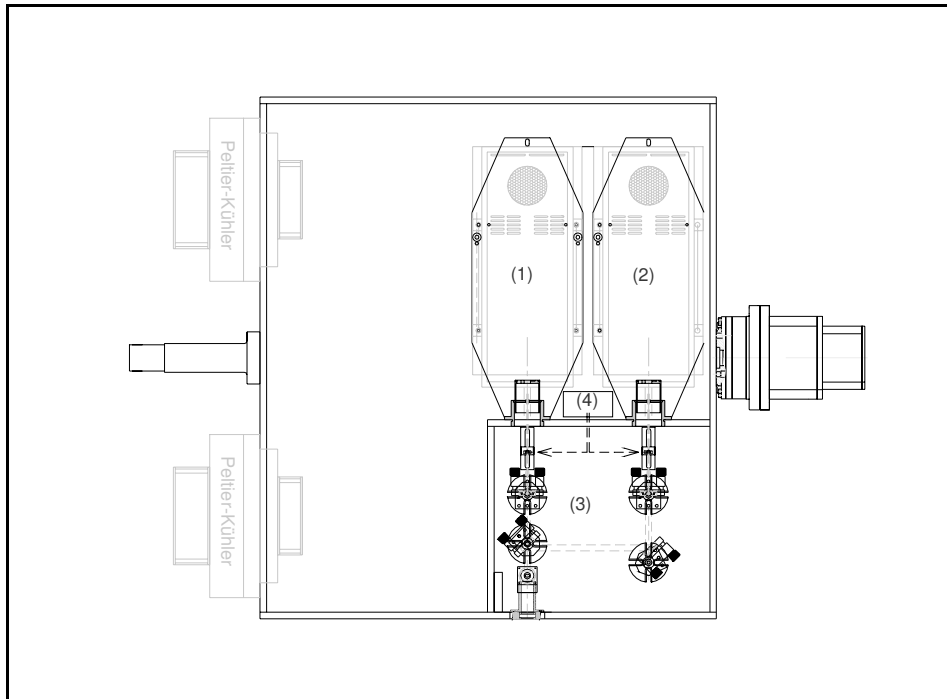
Kühlgerät

Hersteller	Rittal
Typ	Dachaufbau-Kühlgerät
Modell	SK-3290.100
Abmessungen B x H x T (mm)	600 x 400 x 430
Raumbedarf B x H x T (mm)	800 x 600 x 600
Werkstoff	Stahl lackiert
Gewicht	ca. 45 kg
Energieversorgung	1 x 230 V / N / PE / 50/60 Hz
Sicherung der Zuleitung	max. 6 A T
Kühlleistung	1060 W
Temperaturbereich	+20° bis +55° C
Schutzart	IP 54
Umgebungstemperatur	+5 °C bis +35 °C
Luftfeuchtigkeit	10 % bis 75%

A 2: Technische Zeichnungen

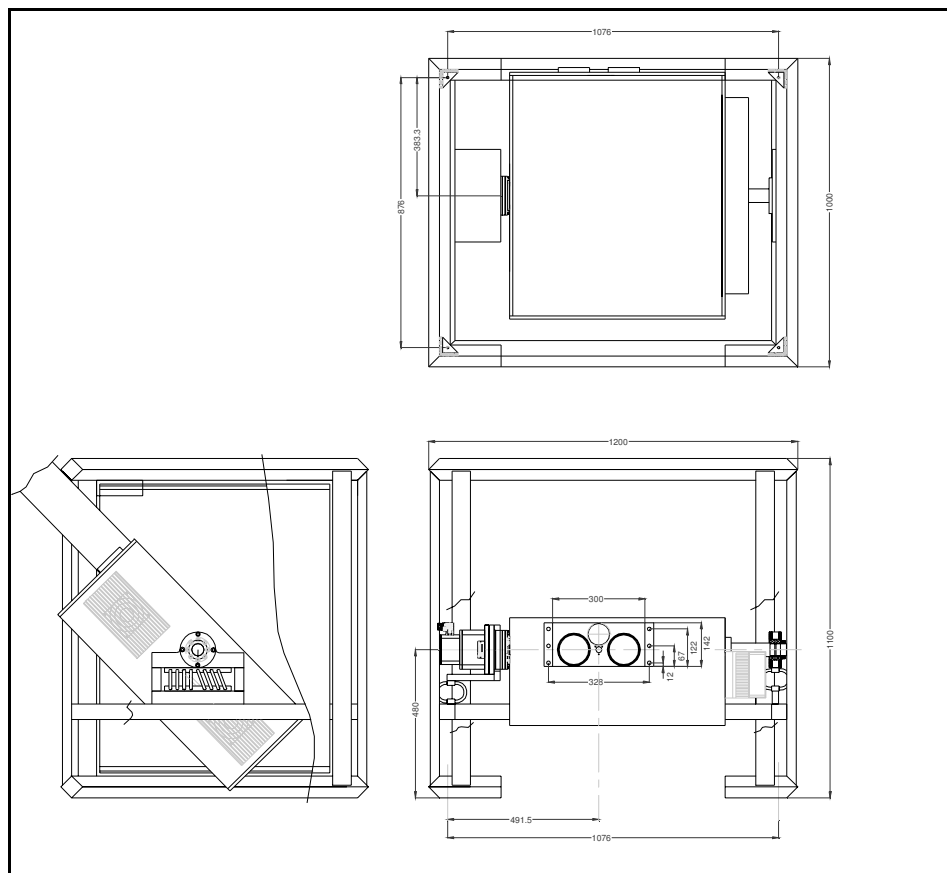


Klassifikationsverfahren zur Materialerkennung



(1) und (2) Laserköpfe

(3) und (4) optische Komponente



Anhang B: Tagebau Hambach: Daten, Profile und Messeinstellungen der Messkampagne

Kohlesorten des Tagebaus Hambach [BEI, 2002]

	Einsatzart	Sorte	Beschreibung	
Brikettierkohle	Kohle, die in den Veredlungsbetriebe eingesetzt werden kann.	HBA	Hambach B-Kohle für alle Abnehmer; auch "pur" in umgebauten Kesseln einsetzbar	Asche † 2,5%; Gel = 0 Schwefel † 0,24%; Eisen † 4000ppm Silizium † 1400ppm Kalium † 200ppm
		HBG	Hambach B-Kohle "gelorientiert" für Fortuna - Nord und Vile/Berrenrath; "pur" in umgebauten Kesseln einsetzbar	Asche † 2,5%; Gel = 0 oder 1 Schwefel † 0,24%; Eisen † 4000ppm Kalium † 200ppm
		HBg	siehe HBG	Asche 2,5 - 3%, sonst wie HBG Information an Kohlelastvert. u. Qualitätsbeauftragten Fabriken; im Bunkerrechner nachweisen
		HBS	Hambach B-Kohle schwefelreich	Asche † 2,5%; Schwefel > 0,24% Eisen † 4000ppm Silizium † 1400ppm + Gel = 0 => Einsatzzweck Brikettierung Silizium > 1400ppm + Gel = 0 oder 1 => Einsatzzweck Staub
Kesselkohle	Kesselkohle, die nicht ohne weiteres Mischen mit Garzweiler- oder Hambachkohle eingesetzt werden kann	HKR	Hambach Kesselkohle; reich an kritischen Bestandteilen (in der Regel sehr eisenreich)	Ausserhalb Dreistoffdiagramm Kalium † 300ppm
	Kesselkohle, die im Dreistoffdiagramm liegt und in Hambachblöcken eingesetzt werden kann	HKF	Hambach Kesselkohle eisenreich, alkaliarm	Dreistoffdiagramm und Eisen > 3000 bis 4000ppm Alkali < 2100ppm Asche < 4,0% Kalium < 200ppm
		HKN	Hambach Kesselkohle eisenarm, alkalireich	Dreistoffdiagramm und Eisen < 3000ppm Alkali > 2000 bis 2500ppm Asche < 4,0% Kalium < 200ppm
		HKA	Hambach Kesselkohle eisenarm, alkaliarm	Dreistoffdiagramm und Eisen < 3000ppm Alkali > 2000 Asche < 4,0% Kalium < 200ppm
	Kesselkohle, die in festgelegten Kraftwerksblöcken pur eingesetzt werden kann	HKE	Hambach Kesselkohle eisenreich, siliziumarm	Eisen < 10000ppm Silizium < 2000ppm Kalium < 200ppm

Ansprache und Profil Liegendes, nördlich des Liegendrückens, Messung am 12.11.2002 mit 266 nm

Winkel		9 Materialien		5 Materialien		4 Materialien		3 Materialien	
von	bis	Nr.	Ansprache	Nr.	Ansprache	Nr.	Ansprache	Nr.	Ansprache
8.50	12.99	6/0	GGMK, leicht holzig	0	GGMK, leicht holzig	1/0	Bauwürdige Schichten	1/0	Bauwürdige Schichten
13.00	14.59	7/1	GGMK - GK	1	GGMK - GK	2/1	Holzlage	2/1	Holzlage
14.60	14.70	2/2	Holzlage	2	Holzlage	0/1	0	0/1	0
14.71	14.79	0/2	0			2/1	Holzlage	2/1	Holzlage
14.80	15.31	2/2	Holzlage						
15.32	16.99	8/3	GK, stark holzig						
17.00	18.49	9/4	GK, tonig, stark holzig						
18.50	20.49	10/5	GK, holzig, leicht holzig						
20.50	21.51	11/6	GK, tonig, stark holzig						
21.52	23.00			3	Unbauwürdige Schichten (ohne Ton)	3/2	Unbauwürdige Schichten (ohne Ton)	5/2	Unbauwürdige Schichten (inklusive Ton)
23.01	23.22	12/7	GK, stark tonig, dkl., grau			0/2/3	0		
23.23	23.46			4	Ton	3	Ton		
23.47	24.00	4/8	Ton						

Ansprache und Profil Liegendes, Tiefstufe, Messung am 13.11.2002 mit 266 nm

Winkel		5 Materialien		4 Materialien		3 Materialien	
von	bis	Nr.	Ansprache	Nr.	Ansprache	Nr.	Ansprache
0.00	4.50	16/0	GGMK	0	GGMK, teilweise leicht holzig	13/0	GGMK, teilweise leicht holzig
4.51	5.10	0/0/1	0			14/1	GK-GGMK, sandig
5.11	8.90	17/1	GK-GGMK, leicht sandig	1	GK-GGMK		
8.91	9.49						
9.50	10.00	0/1/2	0	1/2			
10.01	14.80	18/2	GK, Wechsellagerung				
14.81	14.99	0/2/3	0				
15.00	16.99	19/3	GK, stark tonig	2	GK, Wechsellagerung	15/2	GK Wechsellagerung (unbauwürdig + Ton)
17.00	17.30	0/3/4	0	2/3			
17.31	18.00	4/4	Ton	3	Ton		

Ansprache und Profil Zwischenmittel Flöz Frimmersdorf a/b, Messung am 09.04.2003 mit 266 nm und 355 nm

#2 (3-37,5°) #3 (7,5-37,5°)	#1 (12-38°)		Ansprache	Profile (Materialnummern)									
	Winkel von	Winkel bis		11er	9er	7er	6er	4er					
3.00	4,3		GK-GGMK, einzelne Hölzer										
4,3	4,5		Schluff, tonig, humos (ca. 2cm) = Liegendes Fri b										
4.50	6.50		GK, holzig, einzelne Hölzer										
6.50	9.00		GK, stark holzig, viele Hölzer										
9.00	11.16		GK, holzig, Hölzer										
11.17	12.74		GK, holzig, Hölzer										
12.75	14.79	12.00	GK, sehr stark holzig, viele Hölzer, Stubben	0	0	0	0	0	0	0	0	0	0
14.80	18.05	15.51	GK, holzig, wenige Hölzer	1	1	1	1	1	1	1	1	1	1
18.06	20.15	18.57	Schluff, leicht tonig, holzig, humos = Hangendes Fri a	2	2	2	2	2	2	2	2	2	2
20.16	21.30	20.53	GK-GGMK mit Schluff verunreinigt	3	3	3	3	3	3	3	3	3	3
21.31	24.30	22.56	GK - Gk-GGMK, mit Schlufflinsen	4	4	4	4	4	4	4	4	4	4
24.31	28.05	24.51	GK - Gk-GGMK, mit wenigen Schlufflinsen	5	5	5	5	5	5	5	5	5	5
28.06	28.74	27.81	GK, holzig	6	6	6	6	6	6	6	6	6	6
28.75	29.69	29.04	Stammlage (Holz)	7	7	7	7	7	7	7	7	7	7
29.70	30.75	30.26	GK, holzig	8	8	8	8	8	8	8	8	8	8
30.76	34.23	31.31	GK-GGMK, ganz leicht holzig	9	9	9	9	9	9	9	9	9	9
34.24	37.50	35.56	GK-GGMK, ganz leicht holzig	10	10	10	10	10	10	10	10	10	10
		37.51	GK, holzig										
		38.00											

PMT A/B 1.000V, Blende 0

PMT Einstellungen

Liegendes

12.11.2002

#1 L1 0.07%_s, 3585 Schuß

Messung	Laser	PMT A	PMT B
1	1	365	420
2	1	400	440
3	1	420	460
4	1	440	460
5	1	480	480
6	1	500	500
7	1	532	520
8	1	540	540
9	1	560	660
10	1	600	680
11	1	640	720

Störeffekte

Störeffekte

8 7

#2 L1 0.25%_s, 1250 Schuß

Messung	Laser	PMT A	PMT B
12	1	400	400
13	1	420	420
14	1	440	440
15	1	460	460
16	1	480	480
17	1	500	500
18	1	520	520
19	1	532	540
20	1	540	560
21	1	560	600
22	1	600	640
23	1	640	660
24	1	640	700

PMT A:
Mittelwerte aus
640er

8 7

Gesamtmerkmalsanzahl	135
-----------------------------	------------

13.11.2002

#3 L1 0.2%_s

Messung	Laser	PMT A	PMT B
25	1	365	420
26	1	400	440
27	1	420	460
28	1	440	480
29	1	480	500
30	1	500	520
31	1	532	540
32	1	540	660
33	1	560	680
34	1	600	700
35	1	640	720

11 11

#4 L1 0.2%_s

Messung	Laser	PMT A	PMT B
36	1	365	420
37	1	400	440
38	1	420	460
39	1	440	480
40	1	480	500
41	1	500	520
42	1	532	540
43	1	540	660
44	1	560	680
45	1	600	700
46	1	640	720

11 11

Gesamtmerkmalsanzahl	275
-----------------------------	------------

Zwischenmittel

09.04.2003

#1 L1 0,16%_s

Messung	Laser	PMT A	PMT B
197	1	300	400
198	1	320	420
199	1	365	440
200	1	380	460
201	1	400	480
202	1	420	500
203	1	440	520
204	1	460	540
205	1	480	560
206	1	500	600
207	1	520	640
208	1	532	660
209	1	540	680
210	1	560	700
211	1	600	720
212	1	640	360

15 15

#3 L1 0,16%_s, korrigiert

Messung	Laser	PMT A	PMT B
261	1	0	400
262	1	0	420
263	1	0	440
264	1	400	460
265	1	420	480
266	1	440	500
267	1	460	520
268	1	480	540

5 8

#2 L1 0,16%_s

Messung	Laser	PMT A	PMT B
228	1	320	400
230	1	365	420
232	1	380	440
234	1	400	460
235	1	420	480
236	1	440	500
237	1	460	520
238	1	480	540
239	1	500	560
240	1	520	600
241	1	532	640
242	1	540	660
243	1	560	680
244	1	600	700
245	1	640	720

15 15

Gesamtmerkmalsanzahl mit #3	104
Gesamtmerkmalsanzahl mit #1 und #2	495

#1 L2 0,16%_s

Messung	Laser	PMT A	PMT B
213	2	400	400
214	2	420	420
215	2	440	440
216	2	460	460
217	2	480	480
218	2	500	500
219	2	520	520
220	2	532	540
221	2	540	560
222	2	560	600
223	2	600	640
224	2	640	660
225	2	640	680
226	2	600	700
227	2	560	720

12 15

PMT A: Mittelwerte
560,600,640

#2 L2 0,16%_s

Messung	Laser	PMT A	PMT B
246	2	400	400
247	2	420	420
248	2	440	440
249	2	460	460
250	2	480	480
251	2	500	500
252	2	520	520
253	2	532	540
254	2	540	560
255	2	560	600
256	2	560	720
257	2	600	640
258	2	600	700
259	2	640	660
260	2	640	680

12 15

PMT A: Mittelwerte
560,600,640

Gesamtmerkmalsanzahl	405
----------------------	-----

Anhang C: Algorithmen zur Merkmalsselektion und Materialklassifikation (Flowcharts und Programmcode)

Klassifikationsverfahren

Gaußverfahren

```
function [er,profil,result]=singleGaussneu(trd,trc,ted,tec,type,minvar)
```

```
if (nargin<6) minvar=0.001; end
if (nargin<5) type=1; end

% singleGauss(trd,trc,ted,tec[,covariance[,minvar]])
% covariance: 0:identity 1:diagonal(default) 2:full
% (pooling over all classes)
% returns er

trN=size(trd,2);
teN=size(ted,2);
trc=trc+ones(1,trN);
tec=tec+ones(1,teN);
K=max(trc);
D=size(trd,1);

classindex=zeros(trN,K);
for k=1:K
    classindex(:,k)=(trc==k)';
    classindexR(:,k)=(1.0/sum(classindex(:,k)))*classindex(:,k);
end

means=zeros(D,K);
means=trd*classindexR;

if(type>0)
    trd=trd-(means*classindex');
    if(type==1) %diag
        iW=sparse(D,D);
        for d=1:D
            iW(d,d)=var(trd(d,:),1);
            if(iW(d,d)==0.0) iW(d,d)=minvar; end
            iW(d,d)=1.0/sqrt(iW(d,d));
        end
    else %full
        W=cov(trd',1);
        iW=inv(W);
        iW=sqrtm(iW);
    end
    ted=iW*ted;
```

Klassifikationsverfahren zur Materialerkennung

```
means=iW*means;
end

er=0;

for n=1:teN
    dist=tcd(:,n)*ones(1,K)-means;
    dist=sum(dist.^2,1);
    [tmp,best]=min(dist);
    er=er+(best~=tec(n));
    profil(n,1)=best-1;
    dist=tcd(:,n)*ones(1,K)-means;
    dist=dist.^2;
    for m=1:D
        [tmp,best]=min(dist(m,:));
        res(m,1)=best;
    end
    for m=1:K
        resa(1,m)=size(find(res==m),1);
    end
    result(n,:)=resa;
end

er=100.0*er/teN;
profil;
result;
```

k-nächster Nachbar Verfahren (k-NN)

function [er,profil,result]=KnearestNeighbor(trd,trc,ted,tec,k)

% KnearestNeighbor(trd,trc,ted,tec,k)
% returns er and result of classification in results
% = classification per class + finally derived profile

```
trN=size(trd,2);
teN=size(ted,2);
```

```
K=max(trc)+1;
```

```
er=0;
```

```
for n=1:teN
    dist=tcd(:,n)*ones(1,trN)-trd;
    dist=sum(dist.^2,1);
    for m=1:k
        [tmp,best]=min(dist);
```

Klassifikationsverfahren zur Materialerkennung

```
    res(m,1)=trc(best);
    dist(1,best)=max(dist);
end
for m=1:K
    resa(1,m)=size(find(res==m-1),1);
end
result(n,:)=resa;
[tmp,best]=max(resa);
profil(n,1)=best-1;
er=er+((best-1)~=tec(n));
end

er=100.0*er/teN;
profil;
```

Lineares Regressionsverfahren

function [er,profil]=MyLinearRegression(trd,trc,ted,tec)

% linearRegression(trd,trc,ted,tec,dimension)
% returns er
% adds the constant offset feature automatically

```
Ntr=size(trd,2);
Nte=size(ted,2);
```

```
trd=[ones(1,Ntr);trd];
ted=[ones(1,Nte);ted];
```

```
D=size(trd,1);
```

```
trtar=maketargetvector(trc);
tetar=maketargetvector(tec);
```

```
A=trtar*pinv(trd);
```

```
res=A*tetar;
```

```
[dummy,indlab]=max(tetar);
[dummy,indres]=max(res);
```

```
er=sum((indlab-indres)~=0);
```

```
er=er/size(tetar,2)*100;
profil=indres'-ones(Nte,1);
```

```
function res=maketargetvector(classes)
```

```
K=max(classes)+1;
```

```
N=size(classes,2);
```

```
res=zeros(K,N);
```

```
for n=1:N
```

```
    res(classes(n)+1,n)=1;
```

```
end
```

Materialklassifikationsverfahren

Wrapper2 – Ansatz

```
function [final,count]=wrapper2(trd,trc,ted,tec,selection,classifier,  
type,dimension,resampling,trials,filenameetime,filenameefeature,filename)
```

```
% selection: 1=Sequential floating feature selection (SFFS)
```

```
%           2=sequential forward selection (SFS)
```

```
%           3=sequential backward selection (SBS)
```

```
% classifier: 1= Gauss (type: 0= identity, 1=diagonal, 2=full)
```

```
%           2= k-NN (type: k Neighbors)
```

```
%           3= linear Regression
```

```
% dimension: max. number of features returned
```

```
% resampling: percentage of data elements resampled for each trial
```

```
% trials: number of resampling trials performed
```

```
% filenameefeature.txt: name of txt-file with result of feature vector
```

```
% filename.txt: name of txt-file with result matrix
```

```
tic;
```

```
D=size(trd,1);
```

```
trN=size(trd,2);
```

```
teN=size(ted,2);
```

```
count=zeros(D+1,D);
```

```
for i=1:trials
```

```
    samples=randomsampling(trd,trc,resampling);
```

```
    trcneu=samples(1,:);
```

```
    trdneu=samples(2:D+1,:);
```

```
    if selection == 1
```

```
        h=floating(trdneu,trcneu,ted,tec,classifier,type,dimension);
```

```
    else
```

```
        if selection == 2
```

```
            h=forward(trdneu,trcneu,ted,tec,classifier,type,dimension);
```

```
        else
```

```
            h=backward(trdneu,trcneu,ted,tec,classifier,type,dimension);
```

Klassifikationsverfahren zur Materialerkennung

```

        end
    end
    [tmp,best]=min(h(D+1,:));
    count(:,i)=h(:,best);
    i
    c=clock;
    fix(c(1,4:6))
end

for i=1:D
    feature(i,:)=sum(count(i,:));
end

null=0;
for i=1:D
    if feature(i,1)==0 null=null+1;
    end;
end

durchlauf=D-null;

featureneu=[feature;zeros(11,1)];
dlmwrite(filenameefeature,feature,'\t');

% bar(feature);
% ylabel('frequency');
% xlabel('# of feature');

i=1;
test=zeros(D+12,durchlauf);
profile=zeros(teN*10,durchlauf+1);

while i < durchlauf+1
    [tmp,best]=max(feature);
    test(best,i)=1;
    feature(best)=-1;
    test(:,i+1)=test(:,i);
    i=i+1;
end

toc;
t=toc;

dlmwrite(filenameetime,t,'\t');

for o=1:durchlauf
    trdneu=zeros(o,trN);
    tedneu=zeros(o,teN);
    s=1;

```

Klassifikationsverfahren zur Materialerkennung

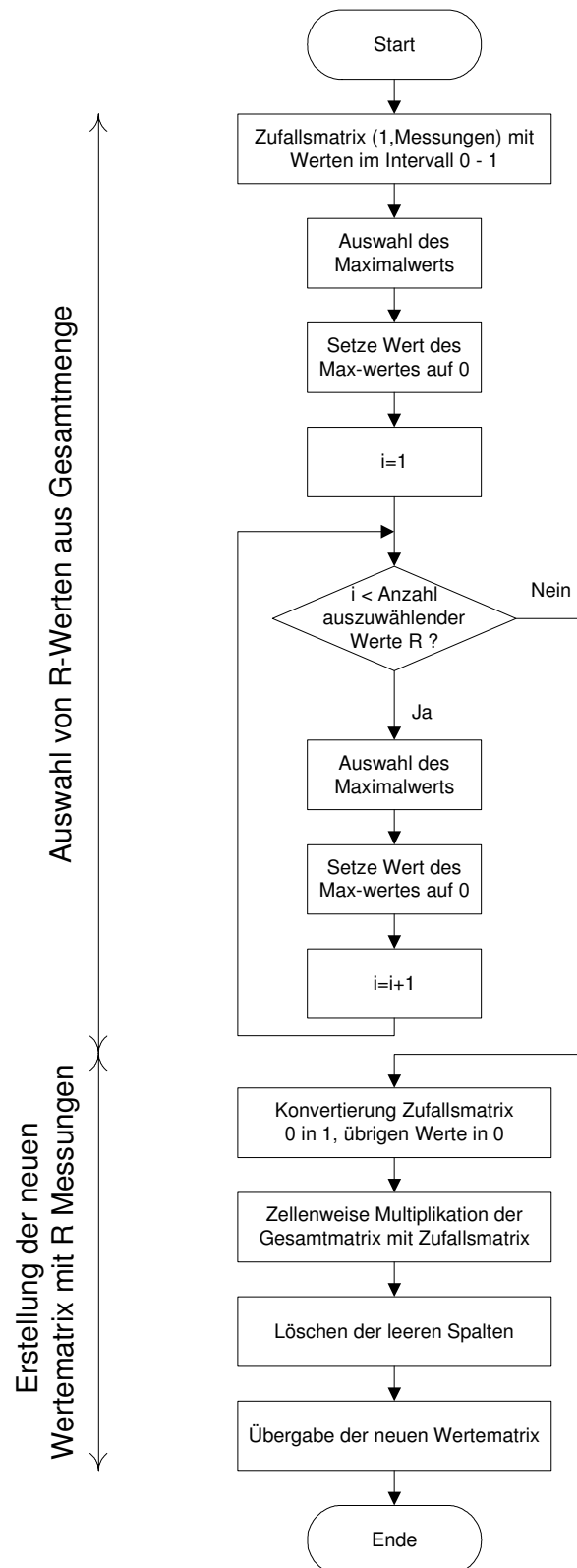
```
for n=1:D
    if test(n,o) == 1
        tedneu(s,:)=ted(n,:);
        trdneu(s,:)=trd(n,:);
        s=s+1;
    end
end
[a,b]=singleGaussneu(trdneu,trc,tedneu,tec,0);
test(D+1,o)=a; profile(1:1*teN,o+1)=b;
[a,b]=singleGaussneu(trdneu,trc,tedneu,tec,1);
test(D+2,o)=a; profile(1*teN+1:2*teN,o+1)=b;
[a,b]=singleGaussneu(trdneu,trc,tedneu,tec,2);
test(D+3,o)=a; profile(2*teN+1:3*teN,o+1)=b;
[a,b]=KnearestNeighbor(trdneu,trc,tedneu,tec,1);
test(D+4,o)=a; profile(3*teN+1:4*teN,o+1)=b;
[a,b]=KnearestNeighbor(trdneu,trc,tedneu,tec,2);
test(D+5,o)=a; profile(4*teN+1:5*teN,o+1)=b;
[a,b]=KnearestNeighbor(trdneu,trc,tedneu,tec,3);
test(D+6,o)=a; profile(5*teN+1:6*teN,o+1)=b;
[a,b]=KnearestNeighbor(trdneu,trc,tedneu,tec,5);
test(D+7,o)=a; profile(6*teN+1:7*teN,o+1)=b;
[a,b]=KnearestNeighbor(trdneu,trc,tedneu,tec,10);
test(D+8,o)=a; profile(7*teN+1:8*teN,o+1)=b;
[a,b]=MyLinearRegression(trdneu,trc,tedneu,tec);
test(D+9,o)=a; profile(8*teN+1:9*teN,o+1)=b;
[a,b]=mylda1(trdneu,trc,tedneu,tec);
test(D+10,o)=a; profile(9*teN+1:10*teN,o+1)=b;
end

featureneu=[featureneu;t];
test=[featureneu, test(:,1:durchlauf)];

for i=1:10
    profile((i-1)*teN+1:i*teN,1)=tec';
end

final=[test;profile]
dlmwrite(filename,final,'\t');
```

Random Sampling




```
function [result]=randomsampling(trd,trc,y)
```

```
% randomsampling returns y % of trd data samples in a combined matrix [trcneu;tedneu]
```

```
D=size(trd,2);
```

```
n=size(trd,1);
```

```
i=0;
```

```
Y=rand(1,D);
```

```
% selection of y% samples randomly
```

```
while ceil(D*y/100)>i
```

```
    [tmp,best]=max(Y);
```

```
    Y(1,best)=0;
```

```
    i=i+1;
```

```
end
```

```
for i=1:D
```

```
    if Y(1,i) == 0 Y(1,i)=1;
```

```
    else Y(1,i)=0; end
```

```
end
```

```
% creation of new samplematrix with ceil(y% x D) samples
```

```
results=zeros(n+1,D);
```

```
for i=1:n+1
```

```
    results(i,:)=Y;
```

```
end
```

```
results=[trc;trd].*results;
```

```
resultsneu=zeros(n+1,ceil(D*y/100));
```

```
m=1;
```

```
for i=1:D
```

```
    if sum(results(:,i)) > 0
```

```
        resultsneu(:,m)=results(:,i);
```

```
        m=m+1;
```

```
    end
```

```
end
```

```
result=resultsneu;
```

Sequential Forward Selection (SFS)

```

function [results]=forward(trd,trc,ted,tec,classifier,type,dimension)

% classifier = 1 : forwardSingleGauss
%               type 0:identity 1:diagonal(default) 2:full (pooling over all classes)
% classifier = 2 : forwardKnearestNeighbor
%               type = k
% classifier = 3 : forwardMyLinearRegression
% dimension: max. number of features returned
% returns er and feature selection in results

if classifier == 1
    results=forwardsingleGaussneu(trd,trc,ted,tec,type,dimension);
else
    if classifier == 2
        results=forwardKnearestNeighbor(trd,trc,ted,tec,type,dimension);
    else
        if classifier == 3
            results=forwardMyLinearRegression(trd,trc,ted,tec,dimension);
        else
            results=forwardMylda1(trd,trc,ted,tec,dimension);
        end
    end
end

results=results(:,1:dimension);

function [results]=forwardsingleGaussneu(trd,trc,ted,tec,type,dimension)

if (nargin<5) type=0; end

% singleGauss(trd,trc,ted,tec[,covariance[,minvar]])
% covariance: 0:identity 1:diagonal(default) 2:full
% (pooling over all classes)
% returns er and result of classification in profil

D=size(trd,1);

results=zeros(D+1,D);
result=zeros(D+1,D);
trN=size(trd,2);
teN=size(ted,2);

% initializing

for i=1:D
    result(i,i)=1;

```

Klassifikationsverfahren zur Materialerkennung

```
    result(D+1,i)=singleGaussneu(trd(i,:),trc,tcd(i,:),tec,type);
end

[tmp,best]=min(result(D+1,:));
results(:,1)=[result(:,best)];

% forward feature selection

for i=1:dimension-1
    for m=1:D
        result(:,m)=results(:,i);
        result(m,m)=1;
    end
    for o=1:D
        if sum(result(1:D,o)) > i
            trdneu=zeros(i+1,trN);
            tcdneu=zeros(i+1,tcN);
            s=1;
            for n=1:D
                if result(n,o)== 1
                    tcdneu(s,:)=tcd(n,:);
                    trdneu(s,:)=trd(n,:);
                    s=s+1;
                end
            end
            result(D+1,o)=singleGaussneu(trdneu,trc,tcdneu,tec,type);
            else result(D+1,o)=100; end
        end
        [tmp,best]=min(result(D+1,:));
        results(:,i+1)=[result(:,best)];
    end
end

results;
```

function [results]=forwardKnearestNeighbor(trd,trc,tcd,tec,k,dimension)

% returns er and feature selection in results

```
D=size(trd,1);
```

```
results=zeros(D+1,D);
```

```
result=zeros(D+1,D);
```

```
trN=size(trd,2);
```

```
tcN=size(tcd,2);
```

% initializing

```
for i=1:D
```

Klassifikationsverfahren zur Materialerkennung

```
    result(i,i)=1;
    result(D+1,i)=KnearestNeighbor(trd(i,:),trc,tcd(i,:),tec,k);
end

[tmp,best]=min(result(D+1,:));
results(:,1)=[result(:,best)];

% forward feature selection

for i=1:dimension-1
    for m=1:D
        result(:,m)=results(:,i);
        result(m,m)=1;
    end
    for o=1:D
        if sum(result(1:D,o)) > i
            trdneu=zeros(i+1,trN);
            tcdneu=zeros(i+1,tcN);
            s=1;
            for n=1:D
                if result(n,o)== 1
                    tcdneu(s,:)=tcd(n,:);
                    trdneu(s,:)=trd(n,:);
                    s=s+1;
                end
            end
            result(D+1,o)=KnearestNeighbor(trdneu,trc,tcdneu,tec,k);
        else result(D+1,o)=100; end
    end
    [tmp,best]=min(result(D+1,:));
    results(:,i+1)=[result(:,best)];
end

results;
```

```
function [results]=forwardMyLinearRegression(trd,trc,tcd,tec,dimension)
```

```
% returns er and feature selection in results
```

```
D=size(trd,1);
```

```
results=zeros(D+1,D);
```

```
result=zeros(D+1,D);
```

```
trN=size(trd,2);
```

```
tcN=size(tcd,2);
```

```
% initializing
```

Klassifikationsverfahren zur Materialerkennung

```
for i=1:D
    result(i,i)=1;
    result(D+1,i)=MyLinearRegression(trd(i,:),trc,tcd(i,:),tec);
end

[tmp,best]=min(result(D+1,:));
results(:,1)=[result(:,best)];

% forward feature selection

for i=1:dimension-1
    for m=1:D
        result(:,m)=results(:,i);
        result(m,m)=1;
    end
    for o=1:D
        if sum(result(1:D,o)) > i
            trdneu=zeros(i+1,trN);
            tcdneu=zeros(i+1,teN);
            s=1;
            for n=1:D
                if result(n,o)== 1
                    tcdneu(s,:)=tcd(n,:);
                    trdneu(s,:)=trd(n,:);
                    s=s+1;
                end
            end
            result(D+1,o)=MyLinearRegression(trdneu,trc,tcdneu,tec);
        else result(D+1,o)=100; end
    end
    [tmp,best]=min(result(D+1,:));
    results(:,i+1)=[result(:,best)];
end

results;
```

Sequential Backward Selection (SBS)

function [results]=backward(trd,trc,tcd,tec,classifier,type,dimension)

```
% classifier = 1 : backwardSingleGaussneu
%               type 0:identity 1:diagonal(default) 2:full (pooling over all classes)
% classifier = 2 : backwardKnearestNeighbor
%               type = k
% classifier = 3 : backwardMyLinearRegression
% dimension: max. number of features returned
% returns er and feature selection in results
```

Klassifikationsverfahren zur Materialerkennung

```
D=size(trd,1);

if classifier == 1
    results=backwardsingleGaussneu(trd,trc,ted,tec,type,dimension);
else
    if classifier == 2
        results=backwardKnearestNeighbor(trd,trc,ted,tec,type,dimension);
    else
        if classifier == 3
            results=backwardMyLinearRegression(trd,trc,ted,tec,dimension);
        else
            results=backwardMylda1(trd,trc,ted,tec,dimension);
        end
    end
end

results=results(:,1:D-dimension);
```

```
function [results]=backwardsingleGaussneu(trd,trc,ted,tec,type,dimension)
```

```
% singleGauss(trd,trc,ted,tec[,covariance[,minvar]])
% covariance: 0:identity 1:diagonal(default) 2:full
% (pooling over all classes)
% returns er and result of classification in profil
```

```
D=size(trd,1);
```

```
results=zeros(D+1,D);
result=zeros(D+1,D);
trN=size(trd,2);
teN=size(ted,2);
```

```
% initializing (all features included)
```

```
for i=1:D
    results(i,1)=1;
```

```
end
results(D+1,1)=singleGaussneu(trd,trc,ted,tec,type);
```

```
% backward feature selection
```

```
i=1;
while i < D-dimension+1
    for m=1:D
        result(:,m)=results(:,i);
        result(m,m)=0;
    end
```

Klassifikationsverfahren zur Materialerkennung

```

for o=1:D
    if sum(result(1:D,o)) == D-i
        trdneu=zeros(D-i,trN);
        tedneu=zeros(D-i,teN);
        s=1;
        for n=1:D
            if result(n,o)== 1
                tedneu(s,:)=ted(n,:);
                trdneu(s,:)=trd(n,:);
                s=s+1;
            end
        end
        result(D+1,o)=singleGaussneu(trdneu,trc,tedneu,tec,type);
    else result(D+1,o)=100; end
end
[tmp,best]=min(result(D+1,:));
results(:,i+1)=[result(:,best)];
i=i+1;
end

results;

```

function [results]=backwardKnearestNeighbor(trd,trc,ted,tec,k,dimension)

% returns er and feature selection in results

```
D=size(trd,1);
```

```
results=zeros(D+1,D);
```

```
result=zeros(D+1,D);
```

```
trN=size(trd,2);
```

```
teN=size(ted,2);
```

% initializing

```
for i=1:D
```

```
    results(i,1)=1;
```

```
end
```

```
results(D+1,1)=KnearestNeighbor(trd(i,:),trc,ted(i,:),tec,k);
```

% backward feature selection

```
i=1;
```

```
while i < D-dimension+1
```

```
    for m=1:D
```

```
        result(:,m)=results(:,i);
```

```
        result(m,m)=0;
```

```
    end
```

Klassifikationsverfahren zur Materialerkennung

```
for o=1:D
    if sum(result(1:D,o)) == D-i
        trdneu=zeros(D-i,trN);
        tedneu=zeros(D-i,teN);
        s=1;
        for n=1:D
            if result(n,o)== 1
                tedneu(s,:)=ted(n,:);
                trdneu(s,:)=trd(n,:);
                s=s+1;
            end
        end
        result(D+1,o)=KnearestNeighbor(trdneu,trc,tedneu,tec,k);
    else result(D+1,o)=100; end
end
[tmp,best]=min(result(D+1,:));
results(:,i+1)=[result(:,best)];
i=i+1;
end

results;
```

function [results]=backwardMyLinearRegression(trd,trc,ted,tec,dimension)

% returns er and feature selection in results

```
D=size(trd,1);
```

```
results=zeros(D+1,D);
```

```
result=zeros(D+1,D);
```

```
trN=size(trd,2);
```

```
teN=size(ted,2);
```

% initializing (all features included)

```
for i=1:D
```

```
    results(i,1)=1;
```

```
end
```

```
results(D+1,1)=MyLinearRegression(trd,trc,ted,tec);
```

% backward feature selection

```
i=1;
```

```
while i < D-dimension+1
```

```
    for m=1:D
```

```
        result(:,m)=results(:,i);
```

```
        result(m,m)=0;
```

```
    end
```



```

for o=1:D
    if sum(result(1:D,o)) == D-i
        trdneu=zeros(D-i,trN);
        tedneu=zeros(D-i,teN);
        s=1;
        for n=1:D
            if result(n,o)== 1
                tedneu(s,:)=ted(n,:);
                trdneu(s,:)=trd(n,:);
                s=s+1;
            end
        end
        result(D+1,o)=MyLinearRegression(trdneu,trc,tedneu,tec);
    else result(D+1,o)=100; end
end
[tmp,best]=min(result(D+1,:));
results(:,i+1)=[result(:,best)];
i=i+1;
end

results;

```

Sequential Forward Floating Selection (SFFS)

function [results]=floating(trd,trc,ted,tec,classifier,type,dimension)

```

% classifier = 1 : forwardSingleGauss
%             type 0:identity 1:diagonal(default) 2:full (pooling over all classes)
% classifier = 2 : forwardKnearestNeighbor
%             type = k
% classifier = 3 : forwardMyLinearRegression
% dimension: max. number of features returned
% returns er and feature selection in results

if classifier == 1
    results=floatingsingleGaussneu(trd,trc,ted,tec,type,dimension);
else
    if classifier == 2
        results=floatingKnearestNeighbor(trd,trc,ted,tec,type,dimension);
    else
        if classifier == 3
            results=floatingMyLinearRegression(trd,trc,ted,tec,dimension);
        else
            results=floatingmylda1(trd,trc,ted,tec,dimension);
        end
    end
end
end

```

Klassifikationsverfahren zur Materialerkennung

```
results=results(:,1:dimension);
```

```
function [results]=floatingsingleGaussneu(trd,trc,tcd,tec,type,dimension)
```

```
% floatforwardsingleGaussneu(trd,trc,tcd,tec[,covariance,dimension])  
% covariance: 0:identity 1:diagonal(default) 2:full (pooling over all classes)  
% dimension as maximum no. of features returned  
% returns er and result of classification in profil
```

```
D=size(trd,1);
```

```
results=zeros(D+1,D);  
result=zeros(D+1,D);  
trN=size(trd,2);  
teN=size(tcd,2);
```

```
% initializing
```

```
for i=1:D  
    result(i,i)=1;  
    result(D+1,i)=singleGaussneu(trd(i,:),trc,tcd(i,:),tec,type);  
end
```

```
[tmp,best]=min(result(D+1,:));  
results(:,1)=[result(:,best)];
```

```
for i=1:D  
    result(:,i)=results(:,1);  
    result(i,i)=1;  
    if sum(result(1:D,i)) == 2  
        trdneu=zeros(2,trN);  
        tcdneu=zeros(2,teN);  
        s=1;  
        for n=1:D  
            if result(n,i) == 1  
                tcdneu(s,:)=tcd(n,:);  
                trdneu(s,:)=trd(n,:);  
                s=s+1;  
            end  
        end  
        result(D+1,i)=singleGaussneu(trdneu,trc,tcdneu,tec,type);  
    else result(D+1,i)=100; end  
end
```

```
[tmp,best]=min(result(D+1,:));  
results(:,2)=[result(:,best)];
```

```
% sequential floating feature selection
```

Klassifikationsverfahren zur Materialerkennung

m=2;

h=0;

while m < dimension

 % adding best remaining feature

 for i=1:D

 result(:,i)=results(:,m);

 result(i,i)=1;

 end

 for o=1:D

 if sum(result(1:D,o)) == m+1

 trdneu=zeros(m+1,trN);

 tedneu=zeros(m+1,teN);

 s=1;

 for n=1:D

 if result(n,o)== 1

 tedneu(s,:)=ted(n,:);

 trdneu(s,:)=trd(n,:);

 s=s+1;

 end

 end

 result(D+1,o)=singleGaussneu(trdneu,trc,tedneu,tec,type);

 else result(D+1,o)=100; end

 end

 [tmp,best]=min(result(D+1,:));

 results(:,m+1)=[result(:,best)];

 m=m+1;

 % conditional exclusion

while h < 1

 for i=1:D

 result(:,i)=results(:,m);

 result(i,i)=0;

 end

 for o=1:D

 if sum(result(1:D,o)) == m-1

 trdneu=zeros(m-1,trN);

 tedneu=zeros(m-1,teN);

 s=1;

 for n=1:D

 if result(n,o)== 1

 tedneu(s,:)=ted(n,:);

 trdneu(s,:)=trd(n,:);

 s=s+1;

 end

 end

 end

Klassifikationsverfahren zur Materialerkennung

```
        result(D+1,o)=singleGaussneu(trdneu,trc,tedneu,tec,type);
        else result(D+1,o)=100; end
    end
    [tmp,best]=min(result(D+1,:));
    if tmp<results(D+1,m-1)
        results(:,m-1)=[result(:,best)];
        m=m-1;
        h=0;
    else h=1; end
end
end
results;
```

```
function [results]=floatingsingleGaussneu(trd,trc,ted,tec,type,dimension)
```

```
% floatingsingleGaussneu(trd,trc,ted,tec,k,dimension)
% type = k nearest neighbors included
% dimension as maximum no. of features returned
% returns er and result of classification in profil
```

```
D=size(trd,1);
```

```
results=zeros(D+1,D);
result=zeros(D+1,D);
trN=size(trd,2);
teN=size(ted,2);
```

```
% initializing
```

```
for i=1:D
    result(i,i)=1;
    result(D+1,i)=kNearestNeighbor(trd(i,:),trc,ted(i,:),tec,type);
end
```

```
[tmp,best]=min(result(D+1,:));
results(:,1)=[result(:,best)];
```

```
for i=1:D
    result(:,i)=results(:,1);
    result(i,i)=1;
    if sum(result(1:D,i)) == 2
        trdneu=zeros(2,trN);
        tedneu=zeros(2,teN);
        s=1;
        for n=1:D
            if result(n,i) == 1
                tedneu(s,:)=ted(n,:);
```

Klassifikationsverfahren zur Materialerkennung

```
        trdneu(s,:)=trd(n,:);
        s=s+1;
    end
end
    result(D+1,i)=kNearestNeighbor(trdneu,trc,tedneu,tec,type);
else result(D+1,i)=100; end
end

[tmp,best]=min(result(D+1,:));
results(:,2)=[result(:,best)];

% sequential floating feature selection

m=2;
h=0;

while m < dimension

    % adding best remaining feature

    for i=1:D
        result(:,i)=results(:,m);
        result(i,i)=1;
    end
    for o=1:D
        if sum(result(1:D,o)) == m+1
            trdneu=zeros(m+1,trN);
            tedneu=zeros(m+1,teN);
            s=1;
            for n=1:D
                if result(n,o)== 1
                    tedneu(s,:)=ted(n,:);
                    trdneu(s,:)=trd(n,:);
                    s=s+1;
                end
            end
            result(D+1,o)=kNearestNeighbor(trdneu,trc,tedneu,tec,type);
        else result(D+1,o)=100; end
    end
    [tmp,best]=min(result(D+1,:));
    results(:,m+1)=[result(:,best)];
    m=m+1;

    % conditional exclusion

    while h < 1
        for i=1:D
            result(:,i)=results(:,m);
            result(i,i)=0;
```

Klassifikationsverfahren zur Materialerkennung

```
end
for o=1:D
    if sum(result(1:D,o)) == m-1
        trdneu=zeros(m-1,trN);
        tedneu=zeros(m-1,teN);
        s=1;
        for n=1:D
            if result(n,o)== 1
                tedneu(s,:)=ted(n,:);
                trdneu(s,:)=trd(n,:);
                s=s+1;
            end
        end
        result(D+1,o)=kNearestNeighbor(trdneu,trc,tedneu,tec,type);
    else result(D+1,o)=100; end
end
[tmp,best]=min(result(D+1,:));
if tmp<results(D+1,m-1)
    results(:,m-1)=[result(:,best)];
    m=m-1;
    h=0;
else h=1; end
end
end
results;
```

function [results]=floatingMyLinearRegression(trd,trc,ted,tec,dimension)

% floatingMyLinearRegression(trd,trc,ted,tec,type,dimension)
% dimension as maximum no. of features returned
% returns er and result of classification in profil

D=size(trd,1);

results=zeros(D+1,D);
result=zeros(D+1,D);
trN=size(trd,2);
teN=size(ted,2);

% initializing

```
for i=1:D
    result(i,i)=1;
    result(D+1,i)=MyLinearRegression(trd(i,:),trc,ted(i,:),tec);
end
```

[tmp,best]=min(result(D+1,:));

Klassifikationsverfahren zur Materialerkennung

```
results(:,1)=[result(:,best)];
```

```
for i=1:D
```

```
    result(:,i)=results(:,1);
```

```
    result(i,i)=1;
```

```
    if sum(result(1:D,i)) == 2
```

```
        trdneu=zeros(2,trN);
```

```
        tedneu=zeros(2,teN);
```

```
        s=1;
```

```
        for n=1:D
```

```
            if result(n,i) == 1
```

```
                tedneu(s,:)=ted(n,:);
```

```
                trdneu(s,:)=trd(n,:);
```

```
                s=s+1;
```

```
            end
```

```
        end
```

```
        result(D+1,i)=MyLinearRegression(trdneu,trc,tedneu,tec);
```

```
    else result(D+1,i)=100; end
```

```
end
```

```
[tmp,best]=min(result(D+1,:));
```

```
results(:,2)=[result(:,best)];
```

```
% sequential floating feature selection
```

```
m=2;
```

```
h=0;
```

```
while m < dimension
```

```
    % adding best remaining feature
```

```
    for i=1:D
```

```
        result(:,i)=results(:,m);
```

```
        result(i,i)=1;
```

```
    end
```

```
    for o=1:D
```

```
        if sum(result(1:D,o)) == m+1
```

```
            trdneu=zeros(m+1,trN);
```

```
            tedneu=zeros(m+1,teN);
```

```
            s=1;
```

```
            for n=1:D
```

```
                if result(n,o) == 1
```

```
                    tedneu(s,:)=ted(n,:);
```

```
                    trdneu(s,:)=trd(n,:);
```

```
                    s=s+1;
```

```
                end
```

```
            end
```

```
            result(D+1,o)=MyLinearRegression(trdneu,trc,tedneu,tec);
```

Klassifikationsverfahren zur Materialerkennung

```

    else result(D+1,o)=100; end
end
[tmp,best]=min(result(D+1,:));
results(:,m+1)=[result(:,best)];
m=m+1;

% conditional exclusion

while h < 1
    for i=1:D
        result(:,i)=results(:,m);
        result(i,i)=0;
    end
    for o=1:D
        if sum(result(1:D,o)) == m-1
            trdneu=zeros(m-1,trN);
            tedneu=zeros(m-1,teN);
            s=1;
            for n=1:D
                if result(n,o)== 1
                    tedneu(s,:)=ted(n,:);
                    trdneu(s,:)=trd(n,:);
                    s=s+1;
                end
            end
            result(D+1,o)=MyLinearRegression(trdneu,trc,tedneu,tec);
        else result(D+1,o)=100; end
    end
    [tmp,best]=min(result(D+1,:));
    if tmp<results(D+1,m-1)
        results(:,m-1)=[result(:,best)];
        m=m-1;
        h=0;
    else h=1; end
end
end

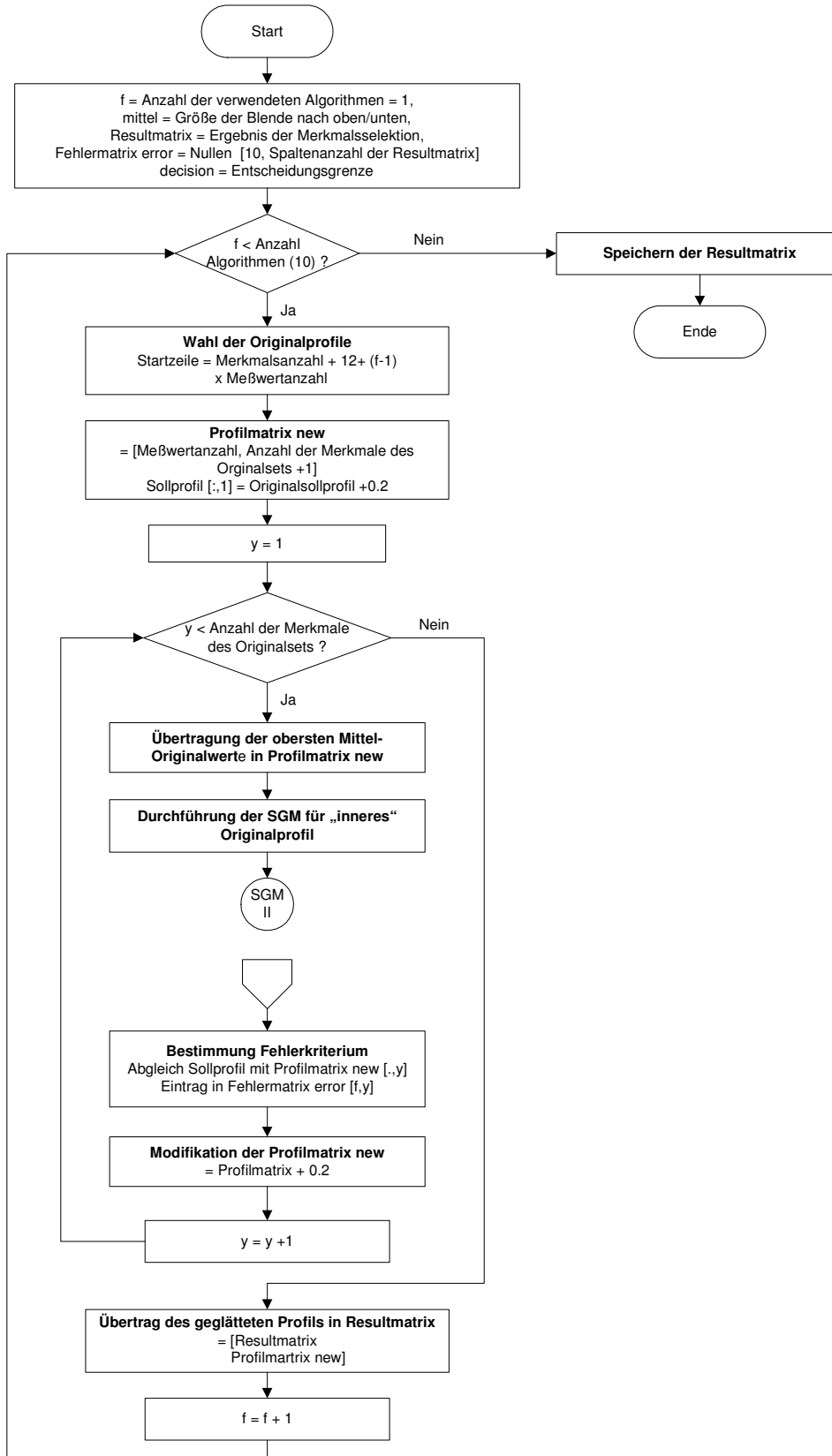
results;

```

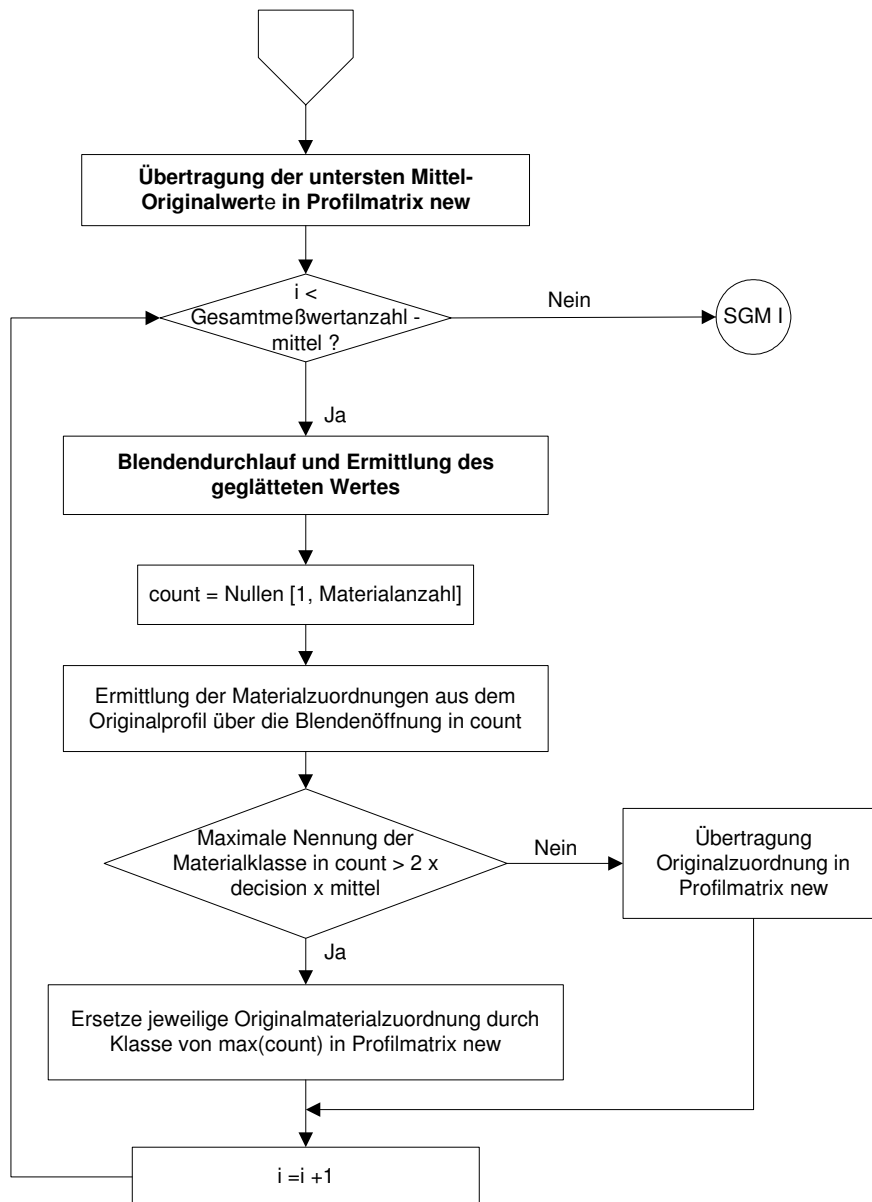

Optimierungsverfahren zur Materialklassifikation

Selektierender gewichtender Mittelwert (SGM)

SGM I



SGM II



function [error]=sgmall(results,ted,tec,mittel,decision,filename)

if (nargin<6) decision=0.6; end

% results: Eingabe des Ergebnis von wrapper2 bzw. wrapRes (Merkmalsvektor,Originalergebnisse, Originalmustererkennung

% mittel= Anzahl der mit einzubeziehenden Werte (+-mittel)

% decision: Entscheidungsgrenze (Standard = 0.6)

spalten=size(results,2);

error=zeros(10,spalten);

tec=tec';

% SGM für alle Spalten aller 10 Algorithmen

for f=1:10

 % SGM für alle Spalten einer der 10 Algorithmen

 start=size(ted,1)+12+(f-1)*size(ted,2);
 new=zeros(size(ted,2),spalten);

 for i=1:size(ted,2)
 new(i,1)=tec(i,1)+.2;
 i=i+1;
 end

 for y=1:spalten-1

 % Übertrag der untersten (Mittel) - Werte aus dem ungeglätteten Profil

 new(1:mittel,y+1)=results(start+1:start+mittel,y+1);

 % Durchlauf für alle Werte des zu glättenden Profils

 for i=mittel+1:size(ted,2)-mittel
 count=zeros(1,9);

 % Blendendurchlauf und Ermittlung des geglätteten Wertes

 for m=-mittel:mittel

 if results(start+i-m,y+1)==0
 count(1,1)=count(1,1)+1;

 else

 if results(start+i-m,y+1)==1
 count(1,2)=count(1,2)+1;

 else

 if results(start+i-m,y+1)==2
 count(1,3)=count(1,3)+1;

 else

 if results(start+i-m,y+1)==3
 count(1,4)=count(1,4)+1;

 else

 if results(start+i-m,y+1)==4
 count(1,5)=count(1,5)+1;

 else

 if results(start+i-m,y+1)==5
 count(1,6)=count(1,6)+1;

 else

 if results(start+i-m,y+1)==6
 count(1,7)=count(1,7)+1;

 else

 if results(start+i-m,y+1)==7
 count(1,8)=count(1,8)+1;

 else count(1,9)=count(1,9)+1; end

 end

```

        end
    end
end
end
end
end
end
m=m+1;
end

[tmp,best]=max(count);
if max(count)>decision*2*mittel
    new(i,y+1)=best-1;
else
    new(i,y+1)=results(start+i,y+1);
end
i=i+1;

end

% Übertrag der untersten (Mittel) - Werte aus dem ungeglätteten Profil

new(size(ted,2)-mittel:size(ted,2),y+1)=results(start+1+size(ted,2)-
mittel:start+1+size(ted,2),y+1);

% Ermittlung der Fehlerrate

hit=0;
for n=1:size(tec,1)
    if tec(n,1)==new(n,y+1) hit=hit+1; end
    n=n+1;
end
error(f,y+1)=(1-hit/size(tec,1))*100;

for z=1:size(ted,2)
    new(z,y+1)=new(z,y+1)+.2;
    z=z+1;
end

y=y+1;
end

results=[results;new];

f=f+1;
end

error;
results=[results;error];

dlmwrite(filename,results,'\t');
```

Anhang D: Ergebnisse der Merkmalsselektion und Mustererkennung

Liegendes (Laser 1, 266 nm)

Nördlich des Liegendrückens

M4 "BEST" Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

M4 „TOP3,4 und 6“ Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

M9 „BEST“ Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

Tiefstufe

M4 "BEST" Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

Liegendes Universell

M4 „BEST universell“ Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

Zwischenmittel

Laser 1 (266 nm)

M4 "BEST" Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

Laser 2 (355 nm)

M4 "BEST" Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

M6 „BEST“ Klassifikationsergebnisse, Merkmalsvektoren,
Profile der Merkmalsklassifikation

**Auswertung Mustererkennung
Lgd nördl, M4 Best (Diss1)**

M4		Auswahl der besten Merkmale aus Reduktion																	
		alle		G0-G2		1-NN - 5-NN		linReg		Min									
		normal <i>E_{rr}</i> #	SGM #	normal <i>E_{rr}</i> #	SGM #	normal <i>E_{rr}</i> #	SGM #	normal <i>E_{rr}</i> #	SGM #	normal <i>E_{rr}</i> #	SGM #	normal <i>E_{rr}</i> #	SGM #						
Mustererkennung																			
G0	G0 #1 #2 G0 #2 #1	48.0 38.9	8 1	44.6 24.8	1 1	44.3 38.9	10 1	42.3 24.4	10 2	41.1 38.9	6 1	28.9 24.8	6 1	62.4 54.5	9 8	51.8 46.1	8 8	41.1 38.9	28.9 24.4
G1	G1 #1 #2 G1 #2 #1	45.3 38.9	8 1	38.6 24.8	11 1	41.3 38.9	10 1	38.7 24.8	10 1	44.4 38.9	6 1	36.0 24.8	6 1	58.9 53.6	9 5	53.8 55.6	9 6	41.3 38.9	36.0 24.8
G2	G2 #1 #2 G2 #2 #1	40.6 38.9	8 1	30.5 24.8	8 1	35.8 38.9	10 1	30.9 24.8	10 1	43.8 38.9	9 1	24.9 24.8	11 1	53.0 51.6	9 5	52.8 55.4	9 7	35.8 38.9	24.9 24.8
1-NN	1-NN #1 #2 1-NN #2 #1	30.7 32.2	5 4	18.5 13.9	5 8	33.1 32.6	7 3	19.7 16.2	4 7	28.6 32.2	9 3	17.4 13.4	6 8	48.0 53.0	8 2	33.7 51.0	9 2	28.6 32.2	17.4 13.4
2-NN	2-NN #1 #2 2-NN #2 #1	28.7 28.2	8 7	17.7 15.2	8 11	34.6 29.8	3 11	26.9 16.4	4 11	27.3 28.0	9 8	15.0 14.9	12 8	51.2 51.8	9 8	50.0 52.9	9 9	27.3 28.0	15.0 14.9
3-NN	3-NN #1 #2 3-NN #2 #1	28.2 28.8	7 2	16.7 12.8	10 8	27.5 28.4	7 3	16.8 15.7	11 4	27.1 27.8	6 3	17.9 11.9	6 8	46.2 51.9	9 2	35.4 47.8	9 2	27.1 27.8	16.7 11.9
5-NN	5-NN #1 #2 5-NN #2 #1	25.5 26.2	8 2	16.1 12.2	11 8	27.4 26.2	11 3	17.9 15.7	3 11	26.0 26.2	4 3	17.7 8.4	6 8	46.0 50.3	9 2	39.5 43.8	9 2	25.5 26.2	16.1 8.4
10-NN	10-NN #1 #2 10-NN #2 #1	22.4 24.5	11 5	12.4 12.4	11 8	25.3 24.9	11 3	16.4 14.5	11 3	25.3 23.9	5 6	18.4 8.0	3 11	44.8 47.7	8 8	38.3 42.5	8 8	22.4 23.9	12.4 8.0
linReg	linReg #1 #2 linReg #2 #1	27.2 30.9	4 1	19.1 30.1	4 5	26.7 30.9	6 1	19.4 31.4	5 2	28.0 30.4	3 11	18.9 26.4	3 11	34.5 42.7	9 8	20.6 42.4	9 8	26.7 30.4	18.9 26.4
Min	#1-#2 #2 - #1	22.4 24.5		12.4 12.2		25.3 24.9		16.4 14.5		25.3 23.9		15.0 8.0		34.5 42.7		20.6 42.4		41.1 38.9	28.9 24.4

11 Merkmale

11 Merkmale

12 Merkmale

9 Merkmale

M4 Lgd nördl. neuer Merk-Vektor

Merkmal	#	Lgd nörd alle	Merkmal	#	Lgd nörd G0-G2	Merkmal	#	Lgd nörd K1-K5	Merkmal	#	Lgd nörd linReg
B460 1/A420 1	89	11	B460 1/A420	89	11	B460 1/A420	89	12	B460 W1/W3	22	9
B440 1/A420 1	88	10	A540 1/A480	45	10	B440 1/A420	88	11	B460 1/B440	65	8
A600 1/A420 1	42	9	B440 1/A420	88	9	A640 1/A420	43	10	A600 W1/W3	14	7
B460 1/B440 1	65	8	A540 1/A532	49	8	A560 1/A420	41	9	B480 W1/W3	24	6
A560 1/A420 1	41	7	B460 1/B440	65	7	A600 1/A420	42	8	A540 W1/W3	10	5
A540 1/A480 1	45	6	A600 1/A420	42	6	A420 1/A400	31	7	A560 1/A540	53	4
A600 1/A400 1	36	5	A600 1/A400	36	5	A540 1/A420	40	6	A640 1/A600	58	3
A640 1/A400 1	37	4	A560 1/A532	50	4	B460 1/B440	65	5	B460 1/A400	82	2
A560 1/A480 1	46	3	B620 1/B420	62	3	A560 1/A400	35	4	A600 1/A420	42	1
A560 1/A400 1	35	2	B640 1/B480	75	2	A600 1/A400	36	3	A400 W1/W2	1	
B620 1/B420 1	62	1	A560 1/A400	35	1	A640 1/A400	37	2	A400 W1/W3	2	
A400 W1/W2	1		A400 W1/W2	1		A540 1/A480	45	1	A420 W1/W2	3	
A400 W1/W3	2		A400 W1/W3	2		A400 W1/W2	1		A420 W1/W3	4	
A420 W1/W2	3		A420 W1/W2	3		A400 W1/W3	2		A480 W1/W2	5	
A420 W1/W3	4		A420 W1/W3	4		A420 W1/W2	3		A480 W1/W3	6	
A480 W1/W2	5		A480 W1/W2	5		A420 W1/W3	4		A532 W1/W2	7	
A480 W1/W3	6		A480 W1/W3	6		A480 W1/W2	5		A532 W1/W3	8	

M4			Auswahl der besten Merkmale aus Reduktion						
			TOP 3 SGM		TOP 4 SGM		TOP 6 SGM		Minimum SGM
			Err	#	Err	#	Err	#	
Mustererkennung	G0	G0 #1-#2	31.5	2	38.2	2	27.3	14	27.3
		G0 #2-#1	11.8	6	12.5	8	17.7	16	11.8
	G1	G1 #1-#2	38.2	1	34.2	6	26.3	11	26.3
		G1 #2-#1	12.1	7	9.2	13	4.1	26	4.1
	G2	G2 #1-#2	35.8	3	29.5	5	18.5	12	18.5
		G2 #2-#1	12.1	8	11.3	7	9.3	15	9.3
	1-NN	1-NN #1-#2	16.8	4	20.6	3	19.1	4	16.8
		1-NN #2-#1	13.9	12	15.9	7	13.7	15	13.7
	2-NN	2-NN #1-#2	18.1	4	17.8	5	14.8	16	14.8
		2-NN #2-#1	17.9	2	20.2	2	14.8	10	14.8
	3-NN	3-NN #1-#2	18.2	3	18.4	3	18.0	5	18.0
		3-NN #2-#1	13.0	11	14.2	6	11.7	8	11.7
5-NN	5-NN #1-#2	15.9	3	16.2	3	17.2	6	15.9	
	5-NN #2-#1	13.5	12	14.6	4	11.6	10	11.6	
10-NN	10-NN #1-#2	18.3	3	18.5	3	16.2	15	16.2	
	10-NN #2-#1	13.2	11	12.5	16	10.6	15	10.6	
linReg	linReg #1-#2	16.8	5	19.1	3	19.0	4	16.8	
	linReg #2-#1	19.0	6	12.8	8	13.2	21	12.8	
Min	#1-#2	15.9		16.2		14.8			
	#2-#1	11.8		9.2		4.1			

Rang	Merkmale		
	TOP 3	TOP 4	TOP 6
1	B460 1/A420 1	B460 1/A420 1	B460 1/A420 1
2	B460 1/B440 1	B460 1/B440 1	A600 1/A420 1
3	B440 1/A420 1	B440 1/A420 1	B460 1/B440 1
4	A640 1/A420 1	A560 1/A420 1	B440 1/A420 1
5	B440 W1/W2	A640 1/A560 1	A600 1/A400 1
6	B460 W1/W2	A640 1/A420 1	A560 1/A400 1
7	B440 W1/W3	B440 W1/W2	A560 1/A420 1
8	A640 W1/W2	B460 W1/W2	A420 1/A400 1
9	B460 W1/W3	A560 W1/W3	A640 1/A560 1
10	A420 W1/W2	A560 W1/W2	A640 1/A400 1
11	A420 W1/W3	B440 W1/W3	B460 1/A400 1
12	B460 1/A640 1	A640 W1/W2	A640 1/A600 1
13	B440 1/A640 1	B460 W1/W3	A640 1/A420 1
14	A640 W1/W3	A420 W1/W2	A600 1/A560 1
15		A420 W1/W3	B440 W1/W2
16		B460 1/A560 1	B460 W1/W2
17		B440 1/A560 1	A600 W1/W3
18		B460 1/A640 1	A560 W1/W3
19		B440 1/A640 1	B440 1/A400 1
20		A640 W1/W3	A600 W1/W2
21			A560 W1/W2
22			B440 W1/W3
23			A640 W1/W2
24			B460 W1/W3
25			A420 W1/W2
26			A400 W1/W3
27			B460 1/A600 1
28			A420 W1/W3
29			B460 1/A560 1
30			B440 1/A560 1
31			B460 1/A640 1
32			B440 1/A600 1
33			B440 1/A640 1
34			A640 W1/W3
35			A400 W1/W2

Auswertung Mustererkennung
Lgd nördl, M9 Best

M9	Auswahl der besten Merkmale aus Reduktion												
	alle		G0-G2		1-NN		linReg		SGM		Min		
	normal Err	#	normal Err	#	normal Err	#	normal Err	#	normal Err	#	normal Err	SGM	
G0	G0 #1-#2	65.6	7	73.2	8	68.2	10	65.1	2	70.5	8	65.2	8
	G0 #2-#1	67.2	6	84.1	1	69.6	3	67.9	10	76.6	6	74.6	3
G1	G1 #1-#2	66.7	8	70.2	7	65.6	8	59.6	8	63.2	10	56.0	10
	G1 #2-#1	72.4	4	81.5	4	76.5	9	74.8	9	73.1	4	70.0	4
G2	G2 #1-#2	65.2	8	72.1	7	63.0	10	58.0	10	64.5	3	58.8	3
	G2 #2-#1	70.3	4	81.8	4	79.9	5	79.8	9	70.9	4	68.2	4
1-NN	1-NN #1-#2	68.3	5	79.7	8	68.3	5	62.2	8	66.2	4	60.8	6
	1-NN #2-#1	66.9	6	79.9	4	68.0	5	64.4	4	70.3	7	68.1	7
2-NN	2-NN #1-#2	65.4	5	74.6	8	64.3	10	58.6	10	63.1	4	56.9	4
	2-NN #2-#1	62.3	6	77.6	4	66.0	6	61.3	6	68.0	7	63.3	7
3-NN	3-NN #1-#2	64.9	5	72.8	8	64.3	8	59.1	8	62.9	4	57.1	5
	3-NN #2-#1	64.1	6	75.2	4	64.7	10	60.2	10	67.3	4	61.1	4
5-NN	5-NN #1-#2	63.8	6	72.8	8	61.3	10	55.2	10	62.3	6	55.7	6
	5-NN #2-#1	64.0	6	77.7	8	65.0	10	60.2	10	65.6	4	61.3	4
10-NN	10-NN #1-#2	63.0	4	72.5	4	59.6	10	54.3	10	60.9	6	53.6	7
	10-NN #2-#1	61.7	6	76.1	4	62.7	5	57.4	5	64.1	7	60.2	7
linReg	linReg #1-#2	65.3	4	70.3	8	63.3	7	58.4	10	60.8	5	54.9	9
	linReg #2-#1	68.7	7	68.5	8	65.1	9	65.9	9	64.2	10	64.9	8
Min	#1-#2	63.0		70.2		59.6		54.3		60.8		53.6	
	#2-#1	61.7		68.5		62.7		57.4		64.1		60.2	

Mustererkennung

M9 Lgd nördl. neuer Merk-Vektor

Merkmal	#	Lgd nörd alle Rang	Merkmal	#	Lgd nörd G0-G2	Merkmal	#	Lgd nörd K1-K5	Merkmal	#	Lgd nörd linReg
B460 1/B440 1	65	8	A540 1/A480 1	45	8	A600 1/A420 1	42	10	A600 W1/W2	13	10
A540 1/A480 1	45	7	A560 1/A532 1	50	7	A640 1/A420 1	43	9	B460 1/A400 1	82	9
A600 1/A420 1	42	6	B460 1/B440 1	65	6	B460 1/A420 1	89	8	B440 1/A420 1	88	8
B460 1/A420 1	89	5	A420 1/A400 1	31	5	A540 1/A480 1	45	7	B460 1/B440 1	65	7
A420 1/A400 1	31	4	A540 1/A532 1	49	4	B440 1/A420 1	88	6	A560 1/A480 1	46	6
B660 1/B540 1	79	3	B660 1/B540 1	79	3	A560 1/A480 1	46	5	B660 1/B540 1	79	5
A560 1/A532 1	50	2	A600 1/A420 1	42	2	B460 1/B440 1	65	4	B440 1/A400 1	81	4
B520 1/B420 1	62	1	A600 1/A400 1	36	1	A600 1/A400 1	36	3	B480 W1/W3	24	3
A400 W1/W2	1		A400 W1/W2	1		A532 1/A480 1	44	2	B520 W1/W2	25	2
A400 W1/W3	2		A400 W1/W3	2		B640 1/B420 1	63	1	A600 1/A420 1	42	1

**Auswertung Mustererkennung
Lgd Tief, M4 Best**

M4	Auswahl der besten Merkmale aus Reduktion																
	alle			G0-G2			1-NN -2-NN			linReg			Min				
	normal Err	#	SGM #	normal Err	#	SGM #	normal Err	#	SGM #	normal Err	#	SGM #	normal	SGM			
G0	#1-#2	43.6	11	40.1	8	43.0	8	53.6	13	64.7	13	56.2	5	67.2	5	43.6	40.1
	#2 - #1	46.0	11	41.5	11	47.3	5	57.5	13	65.2	4	57.8	8	68.8	6	46.0	41.5
G1	#1-#2	43.4	4	35.4	5	43.4	4	47.7	12	49.7	11	54.8	5	61.2	5	43.4	34.9
	#2 - #1	40.5	11	28.5	9	43.9	4	41.2	11	37.3	11	53.4	10	63.5	11	40.5	28.5
G2	#1-#2	40.8	11	30.2	11	43.0	7	45.6	11	45.8	11	47.1	10	48.4	5	40.8	30.2
	#2 - #1	42.3	10	32.2	9	42.8	4	40.6	9	40.1	9	41.2	3	59.5	4	40.6	32.2
1-NN	#1-#2	44.3	6	34.8	11	46.3	6	42.0	13	28.2	13	44.6	4	34.3	4	42.0	28.2
	#2 - #1	43.8	11	31.6	11	45.3	6	40.5	13	32.5	3	44.3	5	32.3	5	40.5	31.6
2-NN	#1-#2	42.6	6	36.8	10	44.6	4	38.7	13	31.8	12	44.9	4	41.2	5	38.7	31.8
	#2 - #1	42.4	10	34.8	11	44.4	8	40.5	13	34.0	3	42.8	5	32.0	6	40.5	32.0
3-NN	#1-#2	41.4	10	34.4	10	42.6	4	39.5	13	28.3	12	43.4	4	32.6	4	39.5	28.3
	#2 - #1	40.3	11	28.7	10	41.3	6	39.1	12	31.3	3	41.2	4	30.2	6	39.1	28.7
5-NN	#1-#2	38.1	10	30.9	11	40.0	4	38.8	4	29.1	13	40.1	7	26.9	8	38.1	26.9
	#2 - #1	38.5	11	28.5	7	38.4	6	37.0	11	30.8	11	39.1	4	30.5	3	37.0	26.9
10-NN	#1-#2	37.8	11	30.8	11	38.7	6	36.6	13	29.4	12	44.8	4	20.1	5	36.6	20.1
	#2 - #1	36.0	11	24.0	10	35.5	6	36.3	12	30.5	10	36.9	5	30.3	3	36.5	23.5
linReg	#1-#2	34.3	10	25.9	10	34.6	4	34.6	9	25.3	7	34.5	10	20.2	10	34.3	20.2
	#2 - #1	36.7	7	27.2	4	37.4	5	36.1	7	31.1	5	39.2	3	31.5	2	36.1	27.2
Min	#1-#2	34.3		25.9		34.6		34.6		25.3		34.5		20.1			
	#2 - #1	36.0		24.0		35.5		36.1		30.5		36.9		30.2			

11 Merkmale

8 Merkmale

13 Merkmale

11 Merkmale

M4 Lgd Tief neuer Merk-Vektor

Merkmal	#	Tief alle	Merkmal	#	Tief G0-G2	Merkmal	#	Tief K1-K2	Merkmal	#	Tief linReg
A400 W1/W2	3	11	A400 W1/W2	3	8	B520/A400	171	13	B420/A400	166	11
A532/A365	50	10	A420 W1/W2	5	7	A560/A420	69	12	B480/A400	169	10
A420 W1/W2	5	9	A480/A400	57	6	A400 W1/W2	3	11	A532/A365	50	9
A480/A400	57	8	A532/A365	50	5	B520 W1/W2	33	10	B520/A400	171	8
A400/A365	45	7	A532/A440	74	4	A532 W1/W2	13	9	A480/A400	57	7
B480 W1/W2	29	6	A400/A365	45	3	B520/B480	128	8	A532 W1/W3	14	6
A600/A400	62	5	B480/B460	119	2	A532/A365	50	7	A600 W1/W2	19	5
A560/A365	52	4	B520/B460	121	1	A532/A420	67	6	B480 W1/W3	30	4
B520/B480	128	3	A365 W1/W2	1		A540/A420	68	5	B540 W1/W3	36	3
B700/B480	132	2	A365 W1/W3	2		B700/B480	132	4	B520/B500	134	2
A560/A400	61	1	A400 W1/W3	4		A560/A400	61	3	B680/B540	146	1
A365 W1/W2	1		A420 W1/W3	6		A500/A420	66	2	A365 W1/W2	1	
A365 W1/W3	2		A440 W1/W2	7		A540/A365	51	1	A365 W1/W3	2	

Liegendes universell

Merkmalsvektoren

K's	Rang	G's	Rang
A560/A420	1	B520 1/B480	1
A640/A420	2	A560 1/A400	2
A600/A420	3	A540 1/A532	3
A540/A420	4	A560 1/A532	4
A560/A400	5	B540 1/B480	5
A640/A400	6	A560 1/A540	6
A540/A480	7	B480 1/B420	7
A540/A400	8	A480 1/A400	8
B660/B420	9		
B660/B520	10		
B660/B480	11		
B520/B420	12		

M4	Mustererkennung											
	G0-G2				1-NN - 5-NN				Min			
	normal Err	#	SGM Err	#	normal Err	#	SGM Err	#	normal	SGM	Min	SGM
G0	#1-#2	66.8	5	65.8	5	63.8	12	63.7	12	63.8	63.7	
	#2-#1	63.1	2	60.8	2	47.8	10	42.4	10	47.8	42.4	
G1	#1-#2	65.2	5	64.3	5	49.6	12	47.3	12	49.6	47.3	
	#2-#1	64.4	2	62.5	2	50.6	10	46.1	10	50.6	46.1	
G2	#1-#2	64.7	8	64.5	8	54.6	12	51.1	12	54.6	51.1	
	#2-#1	64.4	2	62.6	2	57.3	1	54.2	1	57.3	54.2	
1-NN	#1-#2	50.2	2	47.8	2	38.0	9	34.2	3	38.0	34.2	
	#2-#1	51.8	7	43.5	7	38.0	6	21.7	6	38.0	21.7	
2-NN	#1-#2	51.7	4	61.1	4	37.1	9	31.5	9	37.1	31.5	
	#2-#1	50.3	7	41.6	7	38.9	6	26.4	12	38.9	26.4	
3-NN	#1-#2	47.4	2	44.1	2	36.6	2	33.6	2	36.6	33.6	
	#2-#1	48.2	7	40.2	2	35.7	6	18.2	9	35.7	18.2	
5-NN	#1-#2	46.4	3	42.6	2	34.7	1	32.2	1	34.7	32.2	
	#2-#1	48.9	8	42.1	4	32.4	6	16.9	7	32.4	16.9	
10-NN	#1-#2	44.7	2	40.3	2	33.6	1	32.7	1	33.6	32.7	
	#2-#1	47.8	2	38.2	2	31.4	5	16.4	7	31.4	16.4	
linReg	#1-#2	40.6	2	37.6	3	34.4	2	34.8	2	34.4	34.8	
	#2-#1	46.3	2	47.0	2	25.9	1	13.8	1	25.9	13.8	
Min	1 nördl. #1	40.6		37.6		33.6		31.5		33.6	31.5	
	nördl. #2	46.3		38.2		25.9		13.8		25.9	13.8	

M4		Mustererkennung											
		G0-G2				1-NN - 2-NN				Min			
		normal		SGM		normal		SGM		normal		SGM	
		Err	#	Err	#	Err	#	Err	#	Err	#	Err	#
G0	#1-#2	54.3	8	47.5	8	62.6	10	60.1	8	54.3	8	47.5	8
	#2-#1	51.2	8	45.1	8	54.2	12	45.1	5	51.2	8	45.1	5
G1	#1-#2	53.5	8	46.8	8	56.4	5	53.6	5	53.5	8	46.8	8
	#2-#1	52.9	8	46.1	8	60.0	1	56.4	12	52.9	8	46.1	12
G2	#1-#2	56.2	8	51.9	4	58.1	10	55.1	12	56.2	8	51.9	12
	#2-#1	52.2	8	46.7	8	59.9	5	55.3	6	52.2	8	46.7	6
1-NN	#1-#2	53.2	8	46.9	7	52.8	12	44.8	12	52.8	8	44.8	12
	#2-#1	48.8	8	40.0	8	50.5	10	42.6	10	48.8	8	40.0	10
2-NN	#1-#2	52.7	8	45.2	8	52.9	12	46.0	12	52.7	8	45.2	12
	#2-#1	46.8	8	36.8	8	48.5	10	40.2	10	46.8	8	36.8	10
3-NN	#1-#2	51.6	8	43.4	7	49.2	12	41.5	12	49.2	8	41.5	12
	#2-#1	47.6	8	39.8	8	49.3	9	42.1	9	47.6	8	39.8	9
5-NN	#1-#2	48.2	8	39.0	8	46.9	11	38.5	11	46.9	8	38.5	11
	#2-#1	45.3	8	39.6	5	46.8	10	42.8	8	45.3	8	39.6	8
10-NN	#1-#2	46.3	4	35.1	7	42.9	12	35.1	12	42.9	4	35.1	12
	#2-#1	43.4	8	39.4	8	45.5	8	44.4	6	43.4	8	39.4	6
linReg	#1-#2	42.0	8	34.3	8	44.1	10	40.5	1	42.0	8	34.3	1
	#2-#1	45.9	8	43.8	8	53.6	12	56.9	12	45.9	8	43.8	12
Min	id Tief #1	42.0		34.3		42.9		35.1		42.0		34.3	
	id Tief #2	43.4		36.8		45.5		40.2		43.4		36.8	

Auswertung Mustereerkennung
Zwi L1, M4 Best

M4	Auswahl der besten Merkmale aus Reduktion												Min							
	alle				G0-G2				1-NN						linReg		normal	SGM		
	Err	#	SGM	#	Err	#	SGM	#	Err	#	SGM	#	Err	#	Err	#				
Mustererkennung	G0	#2-#3	53.1	9	49.3	9	53.4	9	49.6	10	67.2	1	56.4	1	51.2	10	48.0	10	51.2	48.0
		#3 - #2	61.1	9	52.4	2	61.1	8	52.5	8	57.5	2	50.6	2	59.9	10	52.2	10	57.5	50.6
	G1	#2-#3	47.1	9	41.6	9	42.0	10	36.1	10	67.2	1	56.4	1	52.1	2	48.8	2	42.0	36.1
		#3 - #2	48.8	4	45.2	4	53.6	10	47.1	10	65.1	1	59.3	6	30.6	7	27.3	8	30.6	27.3
	G2	#2-#3	37.1	9	31.1	9	38.0	11	33.8	11	65.3	13	56.3	1	41.6	4	33.7	4	37.1	31.1
		#3 - #2	49.6	4	46.8	4	52.6	11	48.3	11	64.9	13	56.2	13	36.9	4	32.0	4	36.9	32.0
	1-NN	#2-#3	43.5	8	36.8	5	41.0	7	34.0	7	45.8	5	37.4	5	35.9	9	34.6	9	35.9	34.0
		#3 - #2	40.3	7	36.6	6	41.8	6	37.1	10	40.1	13	34.6	5	42.0	3	39.7	1	40.1	34.6
	2-NN	#2-#3	44.2	8	36.8	7	41.3	11	36.7	6	42.9	5	35.9	5	32.4	9	31.8	10	32.4	31.8
		#3 - #2	40.9	6	39.1	6	41.0	11	37.5	11	39.8	13	37.8	7	38.2	3	39.0	9	38.2	37.5
	3-NN	#2-#3	45.7	7	38.5	7	42.6	8	37.4	7	44.6	5	37.3	9	32.2	10	30.2	10	32.2	30.2
		#3 - #2	41.0	2	38.6	2	40.9	11	36.2	11	40.1	13	37.5	5	39.6	4	38.9	4	39.6	36.2
5-NN	#2-#3	44.7	8	37.9	7	41.6	11	35.6	8	44.7	5	38.1	3	32.8	10	29.2	10	32.8	29.2	
	#3 - #2	42.0	2	38.4	6	39.7	11	35.2	10	39.3	13	37.6	5	38.5	4	37.4	4	38.5	35.2	
10-NN	#2-#3	43.7	8	35.3	8	38.9	8	32.6	8	44.8	2	38.6	2	31.1	9	27.5	9	31.1	27.5	
	#3 - #2	40.5	4	36.4	6	39.2	11	34.7	10	40.9	13	38.8	11	37.8	4	37.0	3	37.8	34.7	
linReg	#2-#3	37.5	6	35.3	3	38.4	4	35.4	4	40.9	2	39.4	8	38.0	3	34.9	3	37.5	34.9	
	#3 - #2	40.6	1	39.5	8	40.8	10	39.5	8	44.2	1	43.2	13	34.2	10	31.2	10	34.2	31.2	
Min	#2-#3	37.1		31.1		38.0		32.6		40.9		35.9		31.1		27.5		31.1	27.5	
	#3 - #2	40.3		36.4		39.2		34.7		39.3		34.6		30.6		27.3		30.6	27.3	

11 Merkmale

13 Merkmale

10 Merkmale

M4 Zwi L1 neuer Merk-Vektor

Merkmal	#	Zwi L1 alle	Merkmal	#	Zwi L1 G0-G2	Merkmal	#	Zwi L1 K1	Merkmal	#	Zwi L1 linReg
A400 W1/W2	1	10	B420 W1/W3	14	11	B420 W1/W3	14	13	A400 W1/W2	1	10
B420 W1/W3	14	9	B460/B440	50	10	A460 W1/W3	8	12	B540/A480	104	9
B460/B440	50	8	A400 W1/W3	2	9	A420 W1/W3	4	11	B480/B460	55	8
B460 W1/W3	18	7	A400 W1/W2	1	8	B440 W1/W3	16	10	B540 W1/W3	26	7
A400 W1/W3	2	6	B460 W1/W3	18	7	A400 W1/W3	2	9	B420/A480	98	6
A420/A400	27	5	B540 W1/W3	26	6	A440 W1/W3	6	8	B440/A460	91	5
B540 W1/W3	26	4	B460/A460	92	5	A480 W1/W3	10	7	B480/B440	51	4
B460/A460	92	3	B500/A480	102	4	A460/A440	34	6	B460/A460	92	3
B500/B420	47	2	A460/A440	34	3	B440/A460	91	5	B440/B400	38	2
B440/A460	91	1	B500/B460	56	2	B540 W1/W3	26	4	A480 W1/W3	10	1
A420 W1/W2	3		B440/A460	91	1	B480 W1/W3	20	3	A400 W1/W3	2	
A420 W1/W3	4		A420 W1/W2	3		A480/A460	36	2	A420 W1/W2	3	
A440 W1/W2	5		A420 W1/W3	4		B500/B460	56	1	A420 W1/W3	4	

Auswertung Mustererkennung
Zwi L2, M4 Best

M4		Auswahl der besten Merkmale aus Reduktion												Min normal SGM		
		alle		G0-G2		1-NN		linReg		SGM						
Mustererkennung		normal	SGM	normal	SGM	normal	SGM	normal	SGM	normal	SGM	normal	SGM	normal	SGM	
		#	#	#	#	#	#	#	#	#	#	#	#	#	#	
G0	#1(50)-#2	39.1	37.9	5	36.7	5	71.3	3	59.8	1	64.9	1	64.9	1	38.2	36.7
	#2 - #1(50)	60.2	58.3	2	56.6	1	59.4	1	57.0	1	56.3	1	56.3	1	59.4	56.3
G1	#1(50)-#2	36.6	35.7	5	34.6	6	64.6	3	59.8	1	64.9	1	64.9	1	36.6	34.6
	#2 - #1(50)	47.8	42.2	8	42.6	7	59.4	1	52.3	2	56.3	1	56.3	1	45.8	42.2
G2	#1(50)-#2	49.4	49.2	8	45.2	3	63.3	3	59.8	1	61.2	4	61.2	4	49.4	45.2
	#2 - #1(50)	48.6	42.9	5	54.2	7	29.0	2	52.2	2	51.1	4	51.1	4	29.0	42.9
1-NN	#1(50)-#2	42.1	38.1	4	33.1	5	65.0	1	47.1	1	46.1	2	46.1	2	38.7	33.1
	#2 - #1(50)	50.6	40.5	4	35.3	4	47.4	2	43.5	3	35.9	4	35.9	4	46.0	35.3
2-NN	#1(50)-#2	40.2	36.6	4	33.0	5	54.4	1	44.2	1	43.7	4	43.7	4	37.9	33.0
	#2 - #1(50)	48.2	41.9	2	36.3	3	43.2	3	37.9	3	34.2	4	34.2	4	43.1	34.2
3-NN	#1(50)-#2	39.3	36.4	4	31.7	3	58.3	1	44.3	1	42.9	1	42.9	1	38.3	31.7
	#2 - #1(50)	50.5	44.3	7	36.3	3	45.6	2	40.7	2	43.1	3	43.1	3	43.1	35.9
5-NN	#1(50)-#2	39.6	37.0	3	32.7	3	55.3	1	43.7	1	43.5	4	43.5	4	37.6	32.7
	#2 - #1(50)	50.7	43.1	8	33.3	4	42.1	2	37.0	2	43.9	3	43.9	3	42.1	33.3
10-NN	#1(50)-#2	36.9	36.5	4	29.1	4	52.2	1	44.1	1	38.8	1	38.8	1	34.5	29.1
	#2 - #1(50)	48.4	43.3	5	34.6	2	41.5	2	36.5	2	43.8	3	43.8	3	40.6	34.6
linReg	#1(50)-#2	39.6	39.7	2	36.2	2	44.2	3	44.2	1	41.3	4	41.3	4	36.2	36.2
	#2 - #1(50)	40.2	39.0	3	34.9	4	49.0	1	47.9	1	38.0	2	38.0	2	38.0	33.5
Min	#1(50)-#2	36.6	35.7		29.1		44.2		43.7		40.6		38.8		38.0	36.6
	#2 - #1(50)	40.2	39.0		33.3		29.0		36.5		38.0		33.5		38.0	33.5

7 Merkmale

3 Merkmale

4 Merkmale

M4 Zwi L2 neuer Merk-Vektor

Merkmal	#	alle	Merkmal	#	G0-G2	Merkmal	#	K1	Merkmal	#	linReg
B600/A420	250	8	B720/B420	147	7	B660/A532	342	3	A460 W1.WG3	8	4
B460 W1.WG2	31	7	B440 W1.WG2	29	6	B460/A460	274	2	B420 W1.WG3	28	3
B440 W1.WG2	29	6	B460 W1.WG2	31	5	A560/A500	103	1	B600/A420	250	2
A440 W1.WG3	6	5	A440 W1.WG3	6	4	A400 W1.WG2	1		B720 W1.WG2	53	1
B660/A532	342	4	B600/A420	250	3	A400 W1.WG3	2		A400 W1.WG2	1	
B460/A460	274	3	B460/A460	274	2	A420 W1.WG2	3		A400 W1.WG3	2	
A540/A500	102	2	B540/A532	338	1	A420 W1.WG3	4		A420 W1.WG2	3	
B540/A532	338	1	A400 W1.WG2	1		A440 W1.WG2	5		A420 W1.WG3	4	

Auswertung Mustererkennung
Zwi L2, M6 Best

M4	Auswahl der besten Merkmale aus Reduktion												Min normal	SGM						
	alle				G0-G2				1-NN						linReg					
	normal	Er	#	SGM	normal	Er	#	SGM	normal	Er	#	SGM			normal	Er	#	SGM		
Mustererkennung	G0	#1(50)-#2	45.3	4	44.1	4	55.4	6	52.7	6	59.5	7	58.6	6	48.9	9	46.9	9	44.1	44.1
		#2 - #1(50)	56.8	5	53.4	6	58.7	7	58.8	7	69.8	3	89.4	3	57.9	9	54.6	7	56.8	53.4
	G1	#1(50)-#2	49.7	4	49.2	5	56.3	6	54.6	6	52.3	6	49.2	7	36.7	9	36.4	9	36.4	36.4
		#2 - #1(50)	56.6	6	50.1	6	52.6	7	48.2	7	61.8	10	60.7	10	49.2	9	46.6	9	49.2	46.6
	G2	#1(50)-#2	49.7	4	48.6	6	56.3	7	54.8	7	57.6	6	54.6	8	42.2	9	41.8	9	42.2	41.8
		#2 - #1(50)	60.3	5	57.0	5	49.7	6	46.2	6	62.8	10	63.2	10	55.9	9	51.3	9	49.7	46.2
	1-NN	#1(50)-#2	51.7	5	46.8	5	56.5	6	52.8	6	64.9	8	61.7	8	49.5	9	44.0	9	49.5	44.0
		#2 - #1(50)	53.5	6	46.7	6	66.7	7	65.0	5	69.3	4	68.5	4	57.5	7	54.3	6	53.5	46.7
	2-NN	#1(50)-#2	53.8	5	51.4	5	59.6	6	57.8	6	66.2	8	63.0	8	52.3	9	47.9	9	52.3	47.9
		#2 - #1(50)	56.7	6	52.3	6	66.2	6	64.6	6	68.2	4	66.0	3	59.3	7	55.2	9	56.7	52.3
	3-NN	#1(50)-#2	50.8	5	48.4	5	57.8	6	55.7	6	64.2	10	61.1	8	45.2	9	40.3	9	45.2	40.3
		#2 - #1(50)	53.0	6	48.4	6	66.3	5	64.6	5	67.1	4	65.6	4	56.7	7	51.8	6	53.0	48.4
	5-NN	#1(50)-#2	51.2	4	47.6	4	57.3	7	55.5	6	62.7	10	60.0	8	43.6	9	38.9	9	43.6	38.9
		#2 - #1(50)	54.0	6	50.1	6	66.2	6	65.8	5	66.5	2	64.2	2	56.1	6	50.5	6	54.0	50.1
	10-NN	#1(50)-#2	48.4	5	44.9	4	58.1	6	56.7	6	62.0	10	62.0	5	43.6	9	40.4	9	43.6	40.4
	#2 - #1(50)	53.0	6	48.3	6	65.8	6	66.1	6	66.7	2	65.7	4	55.3	6	48.6	6	53.0	48.3	
linReg	#1(50)-#2	47.1	6	43.9	6	59.9	7	58.6	7	55.9	7	55.1	8	46.6	9	45.5	9	46.6	43.9	
	#2 - #1(50)	56.7	3	53.0	3	60.2	7	58.6	7	61.8	10	62.2	10	55.0	9	54.2	9	55.0	53.0	
Min	#1(50)-#2	45.3		43.9		55.4		52.7		52.3		49.2		36.7		36.4		45.3		
	#2 - #1(50)	53.0		46.7		49.7		46.2		61.8		60.7		49.2		46.6		53.0		

M6 Zwi L2 neuer Merk-Vektor

Merkmal	#	alle Rang	Merkmal	#	G0-G2 Rang	Merkmal	#	K1 Rang	Merkmal	#	linReg Rang
A540/A460	89	6	A540/A460	89	7	A540/A460	89	10	E700/B560	209	9
B660/B540	201	5	A532/A520	106	6	A560 W1/W3	20	9	E600 W1/W2	35	8
B640/B600	211	4	B500/A532	336	5	A520/A460	87	8	E640/B600	211	7
B460 W1/W3	32	3	B720/B420	147	4	B660/B540	201	7	E600/A532	336	6
A520/A460	87	2	B660/B540	201	3	A500/A460	86	6	E720 W1/W3	54	5
B420/B400	121	1	B460 W1/W3	32	2	A440 W1/W2	5	5	B440/A532	333	4
A400 W1/W2	1		A520/A460	87	1	A540/A520	107	4	B420/B400	121	3
A400 W1/W3	2		A400 W1/W2	1		A540/A500	102	3	B440 W1/W2	29	2
A420 W1/W2	3		A400 W1/W3	2		A640/A540	117	2	E600/B400	129	1
A420 W1/W3	4		A420 W1/W2	3		A532/A460	88	1	A400 W1/W2	1	

Lebenslauf

Persönliche Daten:

Name: Arne Kristoffer Bayer
Geburtsdatum / -ort : 28.03.1975 / Lüneburg
Anschrift: Königstraße 30, 52064 Aachen
Familienstand: ledig
Staatsangehörigkeit: deutsch

Bildungsgang:

1981 – 1985 Grundschule in Kuhstedt
1985 – 1987 Orientierungsstufe in Gnarrenburg
1987 – 1994 Gymnasium in Bremervörde, Abitur
1994 – 1995 Wehrdienst
1995 – 2000 Studium der Fachrichtung Bergbau an der RWTH Aachen, Abschluss als Diplom-Ingenieur
1998 – 1999 Teilnahme am European Mining Course (EMC) an der Helsinki University of Technology (Finnland), Royal School of Mines, London (Großbritannien), TU Delft (Niederlande) und RWTH Aachen

Berufsprofil:

2000 – 2004 Wissenschaftlicher Mitarbeiter am Lehr- und Forschungsgebiet Betriebsmittel für die Gewinnung mineralischer Rohstoffe (BGMR) an der RWTH Aachen
seit 1.7.2004 Consultant bei Roland Berger Strategy Consultants in Düsseldorf im Competence Center „Utilities“