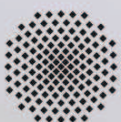
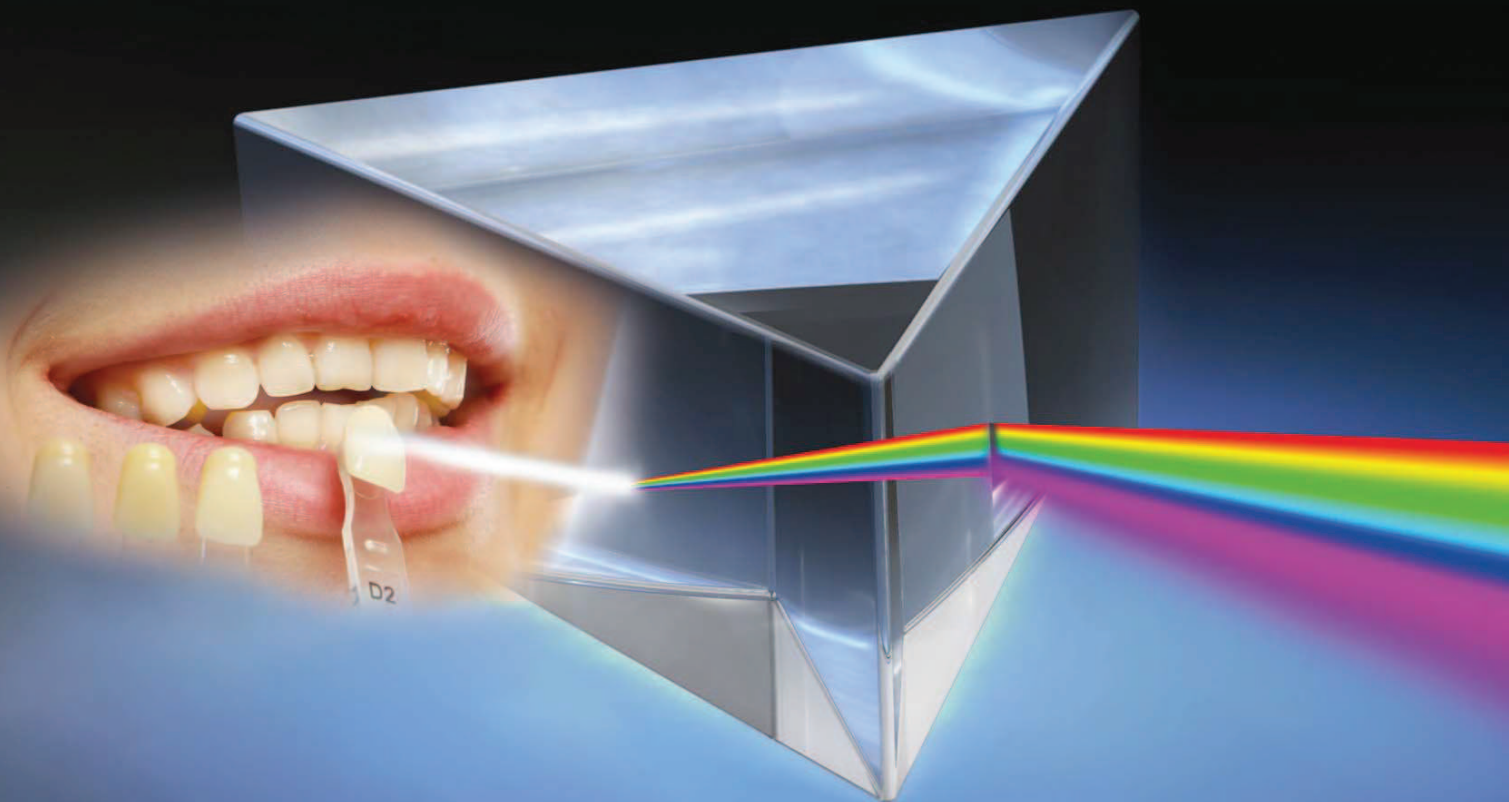


MARIUS PFLÜGER

Entwicklung eines robusten Verfahrens zur Farbbestimmung von Zähnen auf Basis spektraler Daten



Herausgeber:

Univ.-Prof. Dr.-Ing. Thomas Bauernhansl

Univ.-Prof. Dr.-Ing. Dr. h.c. mult. Alexander Verl

Univ.-Prof. a. D. Dr.-Ing. Prof. E.h. Dr.-Ing. E.h. Dr. h.c. mult. Engelbert Westkämper

Marius Pflüger

**Entwicklung eines robusten Verfahrens
zur Farbbestimmung von Zähnen auf Basis
spektraler Daten**

Kontaktadresse:

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA, Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon 0711 970-00, Telefax 0711 970-1399
info@ipa.fraunhofer.de, www.ipa.fraunhofer.de

STUTTGARTER BEITRÄGE ZUR PRODUKTIONSFORSCHUNG**Herausgeber:**

Univ.-Prof. Dr.-Ing. Thomas Bauernhansl
Univ.-Prof. Dr.-Ing. Dr. h.c. mult. Alexander Verl
Univ.-Prof. a. D. Dr.-Ing. Prof. E.h. Dr.-Ing. E.h. Dr. h.c. mult. Engelbert Westkämper

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA, Stuttgart
Institut für Industrielle Fertigung und Fabrikbetrieb (IFF) der Universität Stuttgart
Institut für Steuerungstechnik der Werkzeugmaschinen und Fertigungseinrichtungen (ISW)
der Universität Stuttgart

Titelbild: © Fotolia

Bibliografische Information der Deutschen Nationalbibliothek

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

ISSN: 2195-2892

ISBN (Print): 978-3-8396-0698-8

D 93

Zugl.: Stuttgart, Univ., Diss., 2014

Druck: Mediendienstleistungen des Fraunhofer-Informationszentrum Raum und Bau IRB, Stuttgart
Für den Druck des Buches wurde chlor- und säurefreies Papier verwendet.

© by **FRAUNHOFER VERLAG**, 2014

Fraunhofer-Informationszentrum Raum und Bau IRB
Postfach 80 04 69, 70504 Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon 0711 970-25 00
Telefax 0711 970-25 08
E-Mail verlag@fraunhofer.de
URL <http://verlag.fraunhofer.de>

Alle Rechte vorbehalten

Dieses Werk ist einschließlich aller seiner Teile urheberrechtlich geschützt. Jede Verwertung, die über die engen Grenzen des Urheberrechtsgesetzes hinausgeht, ist ohne schriftliche Zustimmung des Verlages unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Speicherung in elektronischen Systemen.

Die Wiedergabe von Warenbezeichnungen und Handelsnamen in diesem Buch berechtigt nicht zu der Annahme, dass solche Bezeichnungen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und deshalb von jedermann benutzt werden dürften. Soweit in diesem Werk direkt oder indirekt auf Gesetze, Vorschriften oder Richtlinien (z.B. DIN, VDI) Bezug genommen oder aus ihnen zitiert worden ist, kann der Verlag keine Gewähr für Richtigkeit, Vollständigkeit oder Aktualität übernehmen.

GELEITWORT DER HERAUSGEBER

Produktionswissenschaftliche Forschungsfragen entstehen in der Regel im Anwendungszusammenhang, die Produktionsforschung ist also weitgehend erfahrungsbasiert. Der wissenschaftliche Anspruch der „Stuttgarter Beiträge zur Produktionsforschung“ liegt unter anderem darin, Dissertation für Dissertation ein übergreifendes ganzheitliches Theoriegebäude der Produktion zu erstellen.

Die Herausgeber dieser Dissertations-Reihe leiten gemeinsam das Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA und jeweils ein Institut der Fakultät für Konstruktions-, Produktions- und Fahrzeugtechnik an der Universität Stuttgart.

Die von ihnen betreuten Dissertationen sind der marktorientierten Nachhaltigkeit verpflichtet, ihr Ansatz ist systemisch und interdisziplinär. Die Autoren bearbeiten anspruchsvolle Forschungsfragen im Spannungsfeld zwischen theoretischen Grundlagen und industrieller Anwendung.

Die „Stuttgarter Beiträge zur Produktionsforschung“ ersetzt die Reihen „IPA-IAO Forschung und Praxis“ (Hrsg. H.J. Warnecke / H.-J. Bullinger / E. Westkämper / D. Spath) bzw. ISW Forschung und Praxis (Hrsg. G. Stute / G. Pritschow / A. Verl). In den vergangenen Jahrzehnten sind darin über 800 Dissertationen erschienen.

Der Strukturwandel in den Industrien unseres Landes muss auch in der Forschung in einen globalen Zusammenhang gestellt werden. Der reine Fokus auf Erkenntnisgewinn ist zu eindimensional. Die „Stuttgarter Beiträge zur Produktionsforschung“ zielen also darauf ab, mittelfristig Lösungen für den Markt anzubieten. Daher konzentrieren sich die Stuttgarter produktionstechnischen Institute auf das Thema ganzheitliche Produktion in den Kernindustrien Deutschlands. Die leitende Forschungsfrage der Arbeiten ist: Wie können wir nachhaltig mit einem hohen Wertschöpfungsanteil in Deutschland für einen globalen Markt produzieren?

Wir wünschen den Autoren, dass ihre „Stuttgarter Beiträge zur Produktionsforschung“ in der breiten Fachwelt als substanziell wahrgenommen werden und so die Produktionsforschung weltweit voranbringen.

Alexander Verl

Thomas Bauernhansl

Engelbert Westkämper

Entwicklung eines robusten Verfahrens zur Farbbestimmung von Zähnen auf Basis spektraler Daten

Von der Fakultät Konstruktions-, Produktions- und Fahrzeugtechnik
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

Vorgelegt von
Marius Pflüger
aus Albstadt

Hauptberichter: Univ.-Prof. Dr.-Ing. Dr. h. c. mult. Alexander Verl
Mitberichter: Prof. Dr. med. dent. Alexander Hassel

Tag der mündlichen Prüfung 06.02.2014

Institut für Steuerungstechnik der Werkzeugmaschinen und Fertigungseinrichtungen (ISW)
der Universität Stuttgart

2014

Vorwort des Autors

Die vorliegende Arbeit entstand während meiner Tätigkeit am Fraunhofer-Institut für Produktionstechnik und Automatisierung (IPA) in Stuttgart. Herrn Univ.-Prof. Dr.-Ing. Dr. h.c. mult. Alexander Verl danke ich für seine wohlwollende Unterstützung und Förderung meiner Arbeit. Herrn Prof. Dr. med. dent. Alexander Hassel danke ich für die Durchsicht der Arbeit und der Übernahme des Mitberichts.

Ein besonderes Dankeschön gebührt Dr. Wolfgang Rauh für die Grundidee der Arbeit und der Ermunterung, dieses Thema zu bearbeiten. Seine große fachliche Unterstützung in allen Phasen der Entwicklungsarbeit und die vielen produktiven Diskussionen haben maßgeblich zum Gelingen der Arbeit beigetragen.

Besonders erwähnen möchte ich auch Ernst Schmidberger, welcher mich besonders in der Anfangsphase der Arbeit durch viele gute Ideen unterstützt hat. Er war für mich am Anfang meiner beruflichen Laufbahn eine Art Mentor und hat meinen Werdegang nachhaltig geprägt.

Ein großer Dank geht an meinen ehemaligen Kollegen Dr. Joachim Neher. In unserer gemeinsamen Zeit stand er mir immer mit Rat und Tat zur Seite. Viele meiner Entwicklungen basieren auf seinen Vorarbeiten und die produktiven Diskussionen waren häufig der Grundstein für neue Ideen und Lösungen.

Nicht unerwähnt lassen möchte ich auch meine anderen Kollegen aus der Abteilung Bild- und Signalverarbeitung unter der Leitung von Markus Hüttel für die offene und inspirierende Arbeitsatmosphäre. Besonders zu erwähnen sind hier meine Kollegen Julia Kroll und Markus Kaupp mit denen ich mich zu jederzeit austauschen konnte.

Zu guter Letzt möchte ich mich auch noch bei meiner Verlobten Melanie bedanken, die mir viel Rückhalt und Unterstützung gegeben und mich immer ertragen hat, auch wenn es mit meiner Motivation mal nicht zum Besten bestellt war.

Stuttgart, im März 2014

Marius Pflüger

Kurzinhalt

In der heutigen Gesellschaft hat das Aussehen einen immer größeren Stellenwert. Entsprechend kritisch werden deshalb unpassende Farbausprägungen beim Zahnersatz gesehen. In der Praxis wird die Zahnfarbe hierfür meist visuell mithilfe von Farbringen bestimmt, was jedoch fehleranfällig ist.

Diese Arbeit befasst sich deshalb mit der Entwicklung eines Verfahrens zur automatischen Bestimmung der Zahnfarbe mit Hilfe spektraler Messdaten sowie der Integration in den spektralen Smart-Sensor VITA Easyshade. Kernelement ist die Entwicklung eines mehrstufigen Prognosemodells für die Zahnfarbe auf Basis von künstlichen Neuronalen Netzen, welches anhand realer Messdaten trainiert wird. Als vorgelagertes Verfahren wurde außerdem eine Methode entwickelt, um die realen Messdaten hinsichtlich Ihrer Tauglichkeit für das Modelltraining automatisch zu bewerten und inkonsistente Daten ausschleusen zu können. Zur Steigerung der Ergebnisqualität wurde für die Prognoseergebnisse zudem eine Plausibilitätsprüfung durch Nachbarschaften im Farbraum konzipiert und umgesetzt.

Zur Evaluierung der ausgewählten Lösungsansätze konnte im Rahmen der Arbeit auf fast 4000 spektrale Messungen an Zähnen von Probanden zurückgegriffen werden. Die Validierung des Verfahrens zeigte, dass anhand definierter Kriterien die Prognosegüte im Vergleich zu herkömmlichen Verfahren um knapp 40% gesteigert werden konnte.

Short summary

In contemporary society the appearance becomes more and more important. Accordingly, inappropriate color characteristics of dentures are critical. In practice, shade is usually determined by using color rings. But this is highly error-prone.

This work deals with the development of a method for automatic determination of tooth color using spectral data and its integration in the spectral Smart Sensor VITA Easyshade. The core element is the implementation of a multi-stage predictive model for the tooth color on the basis of artificial neural networks. The model is trained by means of real measurement data. As preliminary process a method has been developed to evaluate the real data also regarding their suitability for training the model. Moreover an automatic eject rejection of inconsistent data can be done. To increase the prediction quality a method was designed and implemented to perform a plausibility check through neighborhoods in the color space.

For the evaluation of the selected solutions almost 4000 spectral measurements of teeth from volunteers have been acquired. The validation of the method showed that the prediction accuracy could be raised by almost 40%, compared to conventional methods.

Inhaltsverzeichnis

Abbildungsverzeichnis	vii
1 Einleitung	1
1.1 Problemstellung	1
1.2 Zielsetzung und Vorgehensweise	2
2 Ausgangssituation	4
2.1 Grundlagen der Farbwahrnehmung	4
2.2 Fehleranalyse bei der manuellen Farbbestimmung von natürlichen Zähnen	8
2.3 Anforderungen an ein optimiertes Verfahren zur Bestimmung der Zahnfarbe	13
3 Stand der Technik	18
3.1 Farbmesstechnik	18
3.1.1 Farbmessgeräte im industriellen Einsatz	20
3.1.2 Handliche Farbmessgeräte	22
3.1.3 Prinzipien für die Auswahl von Farbmessgeräten	24
3.2 Datenauswertung und Klassifikation	26
3.2.1 Modellierungsmethoden	26
3.2.2 Klassifikationsverfahren mit maschinellem Lernen	29
3.3 Problemspezifische Ableitung aus dem Stand der Technik	39
4 Detaillierter Lösungsansatz	44
4.1 Visualisierung hochdimensionaler Vektoren	44
4.2 Bestimmung der Zahnfarbe durch Klassifikation von Farbspektren	46
4.2.1 Analyse der spektralen Messdaten	46
4.2.2 Konzeption einer geeigneten Klassifikationsstrategie zur Ermittlung der Zahnfarben	54
4.3 Methoden zur automatischen Bewertung von Trainingsdaten	55
4.3.1 Hüllkurven-Schwellwert-Methode	57
4.3.2 Clustering-Methoden	58
4.3.3 Majority-Filtering-Methode	60

4.4	Auswahl geeigneter Klassifikationsmethoden	62
4.5	Konzeption unterschiedlicher Ansätze zur Klassifikation	67
4.5.1	Ein Modell für alle Klassen	67
4.5.2	Klassenspezifisches Modell	68
4.5.3	Mehrstufiges Modell	69
4.6	Bewertungskriterien und Plausibilitätsprüfungen der Ergebnisse	72
5	Verfahrenstechnische Umsetzung und Validierung	76
5.1	Ermittlung von Trainingsdaten	76
5.2	Anwendung der ausgewählten Verfahren	81
5.3	Bewertung der Ergebnisse und Fehleranalyse	85
5.4	Überprüfung der Anforderungen und Schlußfolgerungen	88
6	Systemseitige Realisierung der Funktionsmuster	92
6.1	Modul zum Training der Klassifikationsverfahren	92
6.1.1	Beschreibung des Trainingsmoduls	92
6.1.2	Datenformate	94
6.1.3	Parametrierung der Trainingsalgorithmen	97
6.2	Modul zur automatischen Klassifikation	99
6.2.1	Beschreibung des Klassifikationsmoduls	99
6.2.2	Datenformate	101
6.2.3	Verbesserung der Ergebnisqualität	102
6.3	Ergebnisse zur realisierten Systemtechnik	103
7	Zusammenfassung und Ausblick	105
7.1	Zusammenfassung	105
7.2	Ausblick	106

Abbildungsverzeichnis

2.1	Farbkreis nach Newton	5
2.2	Additive Farbmischung	5
2.3	CIE Normfarbtafel	8
2.4	Lab Farbraum	9
2.5	Vita Farbskalen	11
2.6	Prinzip 3D-Master	12
2.7	Vita Farbsysteme im Lab-Farbraum	13
3.1	Industrielle Farbmessung	21
3.2	Handliche Sensoren zur Ermittlung der Zahnfarbe	22
3.3	Entscheidungsbaum für die Klassifikation	32
3.4	Aufbau eines MLP-Netz	34
3.5	Aktivierungsfunktionen für Neuronen	35
3.6	Blockschaltbild GRNFN	38
4.1	Prinzip des Sammons-Mapping	45
4.2	Messfehler bedingt durch die Sensorkomponenten	48
4.3	Messfehlern bedingt durch Sensorverkippen	48
4.4	Messfehlern bedingt durch den Nutzer	49
4.5	Messfehlern bedingt durch Verfärbungen im Zahn	49
4.6	Relevanz der Wellenlängenbereiche	51
4.7	Exemplarische Farbverteilung bei realen Zähnen	52
4.8	Vergleich von Sensordaten und Farbwahrnehmung	53
4.9	Labels der Referenzdaten	56
4.10	Sammons-Mapping für Messungen an realen Zähnen	56
4.11	Sammons-Mapping nach Hüllkurven-Schwellwert.Methode	58
4.12	Sammons-Mapping nach der Hierarchischen-Clustering-Methode	60
4.13	Prinzipieller Ablauf des Majority-Filterings	61
4.14	Sammons-Mapping nach der Majority-Filtering-Methode	62
4.15	Einfaches Klassifikationsmodell	67

4.16	Farbspezifisches Klassifikationsmodell	69
4.17	Verkleinerung des Lösungsraums	70
4.18	Mehrstufiges Klassifikationsmodell	71
4.19	Histogrammverteilung	74
5.1	Verteilung der Trainingsdaten	77
5.2	Labels der ausgewählten Trainingsdaten	77
5.3	Sammons-Mapping der 3D-Master Trainingsdaten	78
5.4	Sammons-Mapping der classical Trainingsdaten	79
5.5	Sammons-Mapping der reduzierten 3D-Master Trainingsdaten	80
5.6	Sammons-Mapping der reduzierten classical Trainingsdaten	80
6.1	Flowchart Trainingsmodul	95
6.2	Dialog zur Konfiguration des Netztrainings	98
6.3	Dialog zur Modellkonfiguration	99
6.4	Flowchart Klassifikationsmodul	100
6.5	Klassifikation der Zahnfarbe mit dem VITA Easyshade	104

Kapitel 1

Einleitung

1.1 Problemstellung

In der heutigen informationsgeprägten Gesellschaft haben Aussehen und Image einen immer größeren Stellenwert, da einem optischen Schönheitsideal bestimmte positive Eigenschaften zugeordnet werden. So gelten beispielsweise Menschen mit einem sportlich proportionierten Körper und einem jugendlichen Aussehen in der Regel als fit, zielstrebig und erfolgreich. Neben dem Körper spielt besonders das Gesicht in der Ästhetik eine sehr große Rolle. Ein optisch ansprechendes Gesicht hängt hierbei neben den Gesichtszügen maßgeblich von den Zähnen ab. Vor allem die Werbung verbindet schöne und weiße Zähne häufig mit Attraktivität, Kraft und Erfolg. Entsprechend kritisch werden Auffälligkeiten im Gebiss gesehen, welche zum Beispiel durch Zahnlücken oder optisch unpassende Zähne vorliegen. Bei einer notwendigen Restauration von Zähnen kommt deshalb neben den funktionellen Aspekten des Zahnersatzes den ästhetischen Gesichtspunkten eine immer größere Bedeutung zu. Hier ist neben der räumlichen Anordnung der Zähne vor allem die Bestimmung der Zahnfarbe als ein essentielles Qualitätskriterium zu nennen. Ein Zahnersatz soll sich nahtlos in die natürlichen Zähne einfügen, das Gesamtbild nicht stören und am Besten nicht von den natürlichen Zähnen unterscheidbar sein. Zahlreiche Untersuchungen und wissenschaftliche Arbeiten untermauern deutlich, dass die passende Zahnfarbe für den Patienten eine große ästhetische Bedeutung hat [Wünnemann 2009, Biermann 2009, N.N. 2008a, Müller 2004, Baltzer u. a. 2004, Ross 2002]. Dies ist darin zu begründen, dass nahezu jeder Mensch in der Lage ist, zu beurteilen, ob die Zahnfarbe im Prinzip richtig ausgewählt wurde oder störend auffällt. Selbst Personen mit Farbsehschwächen können in der Regel unpassende oder abweichende Farben im direkten Vergleich erkennen. Passt die gewählte Zahnfarbe also nicht zu den anliegenden Zahnnachbarn oder anderen sichtbaren Zähnen im Mundbereich, so ist die Gefahr groß, dass der Patient unzufrieden ist, verunsichert wird oder den Zahnersatz sogar ablehnt.

In der Praxis wird die Zahnfarbe meist visuell mithilfe von Farbringen oder Farbskalen bestimmt, welche die in der Natur vorkommenden Farben möglichst gut repräsentieren. Hierbei werden die Farbmuster nach vorgegebenen Regeln mit den natürlichen Zähnen im Mundraum visuell vom Farbnehmer verglichen. Diese Methode ist jedoch prinzipiell stark fehleranfällig. Störfaktoren wie beispielsweise die Lichtverhältnisse bei der Farbnahme, die Farbe der Kleidung des Patienten oder des Raumes sowie Sehschwächen beim Farbnehmer sind potenzielle Fehlerquellen bei der visuellen Farbermittlung. Eine weitere Schwierigkeit ist in der Anzahl der Musterfarben zu sehen. Eine steigende Anzahl an Vergleichsfarbmuster bildet den möglichen Farbraum der natürlichen Zähne immer besser ab, wodurch die Möglichkeit, die passende Farbe zu finden, im Prinzip größer ist als bei wenigen Referenzfarben. Eine große Anzahl von Farbmustern führt jedoch auch zu einer steigenden Anzahl von Farbvergleichen bei der Zahnfarbermittlung und dazu, dass das menschliche Auge schnell ermüdet. Dies wiederum erschwert eine zuverlässige und reproduzierbare Farbbestimmung [Ross 2002]. Abhilfe schafft hier eine automatisierte Farbnahme auf Basis optischer Sensoren. Verfügbare Sensoren sind jedoch meist zu teuer, zu unhandlich oder zu ungenau, weshalb sich eine automatische Farbnahme bisher noch nicht flächendeckend durchgesetzt hat [Faber 2006].

Vor diesem Hintergrund ist die exakte Farbbestimmung bei Zähnen als stark fehleranfällig anzusehen, was sich direkt auf die Zufriedenheit der Patienten auswirken kann. Ein enormes Verbesserungspotenzial liegt in der Farbnahme mit optischen Sensoren inklusive einer automatischen Auswertung. Eine zuverlässige und gleichzeitig wirtschaftliche Lösung würde stark dazu beitragen, sowohl die Ergebnisqualität als auch die Reproduzierbarkeit der Farbbestimmung zu erhöhen.

1.2 Zielsetzung und Vorgehensweise

Ziel der Arbeit ist die Entwicklung eines neuen Verfahrens zur detaillierten und hochgenauen Farbklassifikation an natürlichen Zähnen. Das Verfahren muss dem Farbnehmer in einer realen Einsatzumgebung sowohl zuverlässige als auch reproduzierbare Ergebnisse in kürzester Zeit liefern. Hierfür wird in dieser Arbeit zunächst die allgemeine Ausgangssituation bei der manuellen und automatischen Farbbestimmung dargestellt und analysiert. Hieraus werden Anforderungen abgeleitet und prinzipiell geeignete Methoden zur automatischen Klassifikation hinsichtlich der Rahmenbedingungen bei der Zahnfarbermittlung ausgewählt. Daraufhin wird der Stand der Wissenschaft und Technik für geeignete Sensoren und Klassifikationsverfahren analysiert. Anhand einer Bewertung mit den ermittelten Anforderungen werden daraus die Defizite bei den aktuellen Verfahren zur Farbbestimmung abgeleitet. Als Folge werden für die ermittelten Defizite neue Lösungen zur automatischen Farbklas-

sifikation konzipiert, umgesetzt und getestet. Anschließend wird ein Funktionsmuster mit speziellen Auswerte- und Klassifikationsalgorithmen für eine vorgegebene Sensorkomponente realisiert. Anhand ermittelter Messdaten von natürlichen Zähnen aus Feldversuchen werden die realisierten Funktionsmuster validiert und bewertet. Die Arbeit schließt mit einer Zusammenfassung der bearbeiteten Teilthemen, der Ergebnisse und einem Ausblick.

Kapitel 2

Ausgangssituation

Im folgenden Kapitel wird zunächst die Ausgangssituation bezüglich Farbwahrnehmung beim Menschen und der Zahnfarbbestimmung dargestellt. Die Bestimmung der Zahnfarbe ist vor allem dann relevant, wenn für Patienten ein Zahnersatz benötigt wird. Es handelt sich hierbei um ein typisches Verfahren, welches bei Zahnärzten und in Zahntechniklaboren allein in Deutschland jährlich tausendfach durchgeführt wird.

2.1 Grundlagen der Farbwahrnehmung

Der Begriff der Farbe wird laut DIN-Norm definiert als „diejenige Gesichtsempfindung eines dem Auge des Menschen strukturlos erscheinenden Teiles des Gesichtsfeldes, durch die sich dieser Teil bei einäugiger Beobachtung mit unbewegtem Auge von einem gleichzeitig gesehene, ebenfalls strukturlosen angrenzenden Bezirk allein unterscheiden kann.“ [N.N. 1979]. Die Farbe ist also eine Empfindung, durch die der Mensch zwischen zwei strukturlosen Objekten unterscheiden kann. In der Geschichte der Farbtheorie gilt Isaak Newton (1643-1727) als der Begründer der modernen Farbtheorie. Er entdeckte mit Hilfe einer Versuchsanordnung zur Lichtbrechung die Existenz einzelner Spektralfarben. Daraus entwickelte er einen Farbkreis mit den Farben Rot, Orange, Gelb, Grün, Blau, Indigo und Violett (Abbildung 2.1), welcher als erstes Farbmodell und wissenschaftliche Erklärung für die Farbwahrnehmung gilt [Newton 1672].

Als weiterer Begründer in der Farbenlehre ist James Clerk Maxwell (1831-1879) zu nennen, welcher die physikalische Intensitätsmessung des Lichts bei verschiedenen Wellenlängen erforschte. Weiterhin gelang ihm durch Projektion der Nachweis der additiven Farbmischung (Abbildung 2.2), welche besagt, dass alle Farben durch die Grundfarben Rot, Blau und Grün synthetisiert werden können. Dieser Nachweis kann als allgemeine Basis der Farbmetrik des RGB-Farbraums angesehen werden [Küppers 2005, Marchesi 2005]. Eine Weiterentwicklung dieser Theorie gelang Hermann Ludwig Ferdinand von Helmholtz (1821-1849), der die

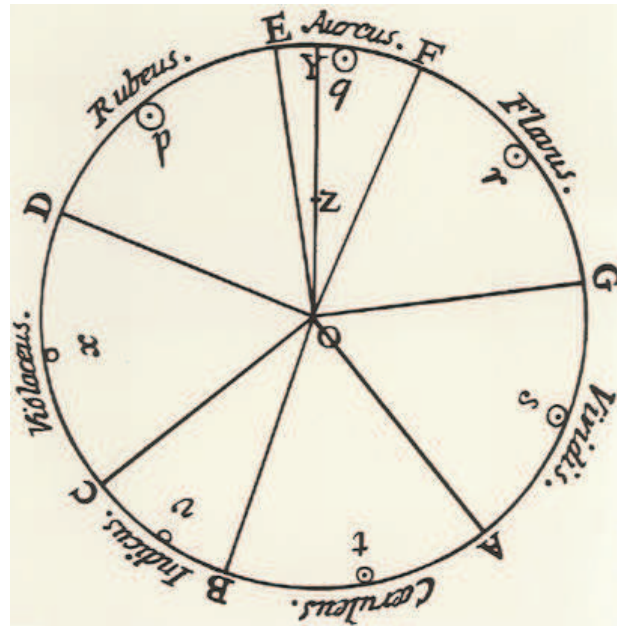


Abbildung 2.1: Der Farbkreis nach Newton mit sieben Grundfarben (Quelle: [Küppers 2005])

Farbwahrnehmung auf verschiedene Rezeptoren im Auge mit jeweils unterschiedlichen Pigmenten für Rot, Blau und Grün zurückführte. Er erkannte auch, dass bei der Farbe Schwarz keine Rezeptoren angeregt werden, während bei Weiß entsprechend alle Rezeptoren angeregt sind. Auch geht auf Helmholtz die Beschreibung von Farben mit den physikalischen Größen Helligkeit, Sättigung und Farbton zurück. Diese Beschreibung stellt die Grundlage für wahrnehmungsorientierte Farbräume (beispielsweise HSV und HSL) dar [Küppers 1999].



Abbildung 2.2: Additive Farbmischung mit den Grundfarben (Quelle: [Heintz u. a. 2010])

Um Farben wahrzunehmen, ist physikalisch betrachtet zuerst Licht notwendig, welches als elektromagnetisch Schwingung unterschiedlicher Wellenlängen definiert werden kann. Das "sichtbare Licht" liegt im Wellenlängenbereich etwa zwischen 400 und 750 Nanometer. Weiterhin ist für die Farbwahrnehmung ein Objekt notwendig, welches von dem Licht angestrahlt wird. Das Licht wird je nach Objekteigenschaften absorbiert oder reflektiert, wodurch sich die Intensität der einzelnen Wellenlängen verändert. Da das Licht an sich "farblos" ist, entsteht die Farbe schließlich durch den Betrachter, welcher durch sein Auge die Farbe wahrnimmt. Das menschliche Auge besitzt hierfür lichtempfindliche Sinneszellen auf der Netzhaut, die sogenannten Zapfen und Stäbchen. Die Intensität der verschiedenen Wellenlängen des Lichtes wird von den Stäbchen auf der Netzhaut wahrgenommen. Die Zapfen hingegen rufen durch Ihre spektrale Empfindlichkeit in Abhängigkeit von der jeweiligen Wellenlänge die Farb- oder Buntempfindung hervor. Dadurch entsteht im Gehirn der Farbeindruck. Es können drei Arten von Zapfen unterschieden werden. Bei Menschen ohne Farbsehschwäche sind ungefähr 10% der Zapfen für Blau empfindlich, 48% für Grün und 42% für Rot [Richter 1981].

Die drei Einflussfaktoren der Farbbeurteilung (Licht, Objekt und Betrachter) bewirken zudem, dass gleiche Objekte bei unterschiedlichen Beleuchtungen verschieden farbig wahrgenommen werden können. Dies lässt sich beispielsweise an einem Kleidungsstück bei Tageslicht oder Raumbeleuchtung gut beobachten. Ebenso können gleiche Objekte eine unterschiedliche Farbe aufweisen, obwohl Sie mit dem gleichen Licht beleuchtet werden. In der freien Natur lässt sich dies sehr anschaulich beobachten. Auf einer Wiese können bei Tageslicht unterschiedlichste Grüntöne wahrgenommen werden, obwohl alle Objekte vom "gleichen" Licht der Sonne beleuchtet werden. Somit beeinflusst einerseits die Art des Lichtes und andererseits die Absorptions- und Reflexionseigenschaften des Objektes unsere Farbwahrnehmung [Berger-Schuhn 1991].

Zusammenfassend lässt sich sagen, dass die Farbe weder eine Stoffeigenschaft noch ein Zustand, sondern eine subjektive Sinneswahrnehmung ist. Zudem ist die wahrgenommene Farbe abhängig von der Beleuchtung, den physikalischen Eigenschaften des Objekts sowie den biologischen Gegebenheiten des Betrachters hinsichtlich seiner Sehfähigkeiten [Heintz u. a. 2010]. Als Sinneswahrnehmung lässt sich die Farbe zwar einerseits durch unterschiedliche physikalische Größen beschreiben, wodurch eine Farbe im Prinzip messbar ist. Andererseits ist eine eindeutige Zuordnung zwischen Physik und menschlicher Wahrnehmung nicht möglich, denn je nach vorliegender Umgebung können abweichende Wahrnehmungen oder Farbtäuschungen auftreten [Franke 2007]. Somit ist die Farbwahrnehmung ständigen Veränderungen unterworfen. Aus diesem Grund eignet sich die als subjektiv einzu-stufende Farbwahrnehmung des Menschen nicht als eindeutiges Messverfahren zur Bestim-

mung und zum Vergleich von Farben. Deshalb wurde bereits Anfang des 20. Jahrhunderts nach objektiven Wegen der Farbmessung gesucht. Ziel war ein genormtes Farbsystem, welches unabhängig von der Wahrnehmung des menschlichen Auges ist und eine eindeutige mathematische Beschreibungsmöglichkeit für unterschiedliche Farben liefert, beispielsweise durch Relationen der Primärfarben. Dadurch sollten objektive Farbvergleiche ermöglicht werden, die keine subjektiven Vergleiche mit Farbmustern benötigen [Baltzer u. a. 2004].

Die internationale Beleuchtungskommission “Commission internationale de l’éclairage” entwickelte daraufhin 1931 die erste CIE-Normfarbtafel. Basierend auf der additiven Farbmischung von Maxwell wurden die vom Menschen wahrnehmbaren Farben experimentell erfasst. Mit Hilfe von genormten Lichtquellen wurde für jede Farbe der jeweilige Anteil des Wellenlängenbereichs für die definierten Grundfarben Rot, Grün und Blau ermittelt. Im englischsprachigen Raum werden diese Grundfarben als “Tristimulus” bezeichnet und mit x , y und z abgekürzt. In diesem System kann jede Farbe durch ein Trippel ihrer Tristimuluswerte beschrieben werden, wobei durch die Normierung der der Farbwerte gilt:

$$x + y + z = 1$$

Diese von der CIE entwickelte Normtafel, welche prinzipbedingt eine hufeisenförmige Darstellung aufweist (Abbildung: 2.3), wird zweidimensional mit den xy -Werten dargestellt (der z -Wert ergibt sich aus der Differenz zu 1) [Hunt 1998]. Das System hat jedoch einige Nachteile. Die zugrundeliegenden Werte wurden lediglich experimentell ermittelt, was sich nachteilig in der Genauigkeit auswirkt. Zudem ist die Farbverteilung ungünstig gewählt, da der Grünbereich verglichen mit Rot und Blau von seiner Fläche in der Farbtafel stark überrepräsentiert ist. Dies führt dazu, dass sich die Abstände der Farben je nach Position in der Farbtafel nicht mit der empfindungsgemäßen Farbwahrnehmung des Menschen decken. Der geometrische Abstand entspricht also nicht dem visuellen Abstand, weshalb es nicht ohne Weiteres möglich ist, reale Farbunterschiede als einfachen Abstand auf der Farbtafel zu beschreiben [Heintz u. a. 2010].

Wegen den genannten Nachteilen wurde das CIE-Farbsystem 1976 zum Lab-System erweitert. Hierbei werden die farbmatischen Koordinaten x , y und z , welche die Grundfarben repräsentieren, logarithmisch in die physikalischen Größen Helligkeit, Farbton und Sättigung umgeformt. Der Farbraum bildet dadurch eine dreidimensionale Kugel. Die senkrechte L -Achse repräsentiert die Helligkeit, an deren Ende sich Schwarz und Weiß befinden. Die waagrechte a -Achse beschreibt den Grün- und Rotanteil, während die ebenfalls waagrechte b -Achse den Blau- und Gelbanteil der darzustellenden Farbe beschreibt (Abbildung 2.4).

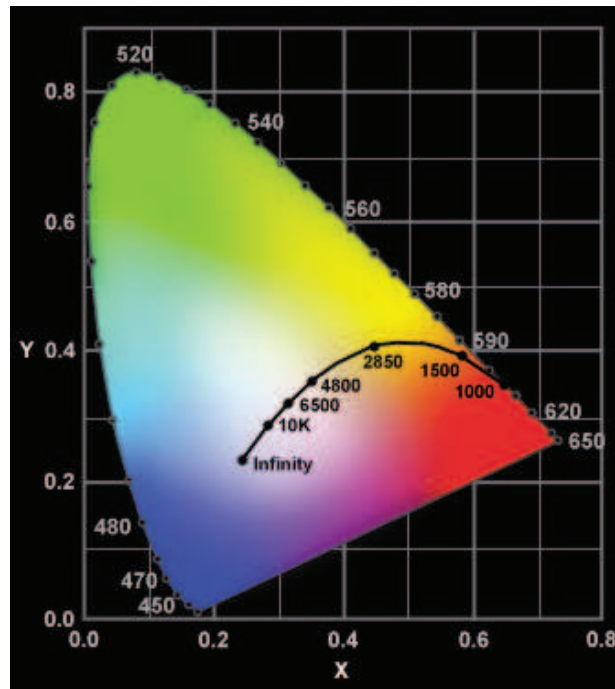


Abbildung 2.3: Die CIE Normfarbtafel von 1931, (Quelle: [N.N. 2011a])

Eine weitere gängige Möglichkeit der Farbdarstellung ist der LCH-Farbraum. Hierbei wird der Lab-Farbraum in die physikalischen Größen Helligkeit (L), Chroma oder Farbintensität (C) sowie den Winkel der Achse zum Farbort (H) umgerechnet. Der Farbraum ist hierbei identisch, lediglich die Lokalisation der Farbe wird durch Zylinderkoordinaten und nicht wie beim Lab-Farbraum mit kartesischen Koordinaten durchgeführt [Baltzer u. a. 2004].

Im Lab und LCH-Farbraum kann wie im CIE-Farbsystem von 1976 jede Farbe durch drei Koordinaten beschrieben werden. Im Gegensatz zum CIE-Farbsystem ist die Anordnung der Farben jedoch so gewählt, dass Unterschiede zwischen den Farben möglichst gut den euklidischen Abständen zwischen den Farbwerten entsprechen. Somit ist die zugrundeliegende Farbmetrik weitestgehend geräte- und wahrnehmungsunabhängig, weshalb viele technische Methoden der Farbmessung auf diesem System basieren [Berger-Schuhn 1991].

2.2 Fehleranalyse bei der manuellen Farbbestimmung von natürlichen Zähnen

Zur Herstellung von künstlichem Zahnersatz ist die Bestimmung der Zahnfarbe ein elementarer Baustein. Für die manuelle Bestimmung wird in der Regel folgendes Vorgehen favorisiert: Die Farbnahme erfolgt nach der Reinigung der Zahnoberfläche aber noch vor einer Präparation der Zähne, da die Zähne schnell dehydrieren und sich aufhellen können. Der Farbnehmer sitzt in einer Entfernung von ca. 60 cm bei einem seitlichen Lichteinfall

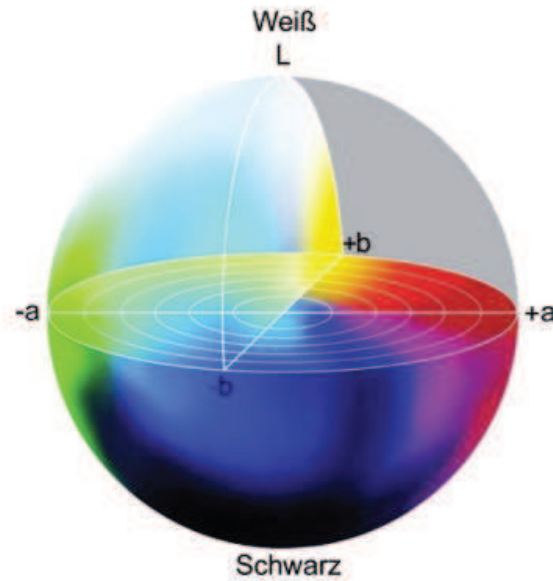


Abbildung 2.4: Der Lab Farbraum von 1976 (Quelle: [Altmann 2008])

und vergleicht vorgegebene Farbmuster mit der Zahnfarbe des Patienten. Die Auswahl der Übereinstimmung muss hierbei zügig getroffen werden, da das Auge nach kurzer Zeit ermüdet. Um das Auge wieder zu entspannen, sollte der Farbnehmer kurz auf eine graue, neutrale Fläche sehen. Es wird generell empfohlen die Farbnahme nicht alleine durchzuführen. Im Idealfall assistiert der Zahntechniker, welcher den Zahnersatz herstellt, oder ein entsprechend geschulter Arzthelfer [N.N. 2010c]. Bei der Bestimmung der Zahnfarbe geht es nicht um ein funktionelles, sondern hauptsächlich um ein ästhetisches Qualitätskriterium, welches für den Patienten von sehr große Bedeutung hinsichtlich seiner Kundenzufriedenheit ist [N.N. 2008a]. Statistische Untersuchungen haben jedoch ergeben, dass ca. 80% aller Patienten einen Farbunterschied zwischen ihren natürlichen Zähnen und dem Zahnersatz erkennen können [Devigus 2003]. Aufgrund von Fehlern bei der Farbbestimmung ist deshalb bei bis zu 65% des künstlichen Zahnersatzes eine Nacharbeit oder eine komplette Neufertigung notwendig [Bücking 2002]. Dies zeigt deutlich, dass Verfahren zur manuellen Bestimmung der Zahnfarbe stark fehleranfällig sind und entsprechendes Optimierungspotenzial bieten. Erschwerend kommt hinzu, dass die Farbdifferenzierung, die durch Schulung und Training verbessert werden kann, im zahnmedizinischen Studium nur eine untergeordnete Rolle spielt. Sie hängt somit stark von kognitiven Fähigkeiten, Talent und Erfahrung des jeweiligen Farbnehmers ab [Wünnemann 2009].

Fehlklassifikationen bei Zahnfarben können unterschiedliche Ursachen haben. Bei manueller, vergleichender Farbbestimmung durch den Zahnarzt oder dessen Helfer ist die subjektive Farbwahrnehmung des Menschen eine potenzielle Fehlerquelle. Derartige Fehlerquellen können zudem durch verschiedene Arten erblicher Farbsehschwächen verstärkt werden, welche

bei gut 4% der Bevölkerung auftreten und oft unerkannt bleiben. Als weiterer starker Einflussfaktor ist das vorhandene Licht zu nennen. So kann ein dunkler Zahn mit schwacher Farbausprägung im Licht einer OP-Lampe in der Zahnarztpraxis nahezu gleich aussehen wie ein hellerer und farbintensiverer Zahn bei Tageslicht. Dies ist darin zu begründen, dass derartige Lichtquellen einen dunkleren Zahn wesentlich heller erscheinen lassen, als er eigentlich ist. Außerdem können störende Hintergrundfarben im Umfeld, beispielsweise auffällige Tapeten, Vorhänge oder Möbel die Farbwahrnehmung negativ beeinflussen. Auch grelle Kleidung, Schmuck oder Schminke beim Patienten sowie Sehhilfen mit getönten Gläsern beim Farbnehmer gelten als negative Einflussfaktoren [Baltzer u. a. 2004]. Wahrnehmungswissenschaftlich betrachtet lassen sich diese Fehler in drei Kategorien der Farbtäuschung einteilen [Viénot u. a. 2005]. Diese sind Farbinduktion (die Farbe wird durch Umgebungsfarben verändert), Farbkonflikte (unterschiedliche Sehinformationen stehen in Konkurrenz zueinander) sowie subjektives Farbempfinden (es werden Farben wahrgenommen, welche nicht vorliegen). Zur größtmöglichen Reduzierung von Farbtäuschungen sollten Farbnehmer deshalb direkte Sonneneinstrahlung auf den Patienten weitestgehend vermeiden. Auch auf starkes künstliches Licht wie das einer OP-Lampe sollte verzichtet werden. Zudem kann ein gleichmäßig ausgeleuchteter Raum, eine farbneutrale Raumgestaltung sowie das Tragen von grauen Überziehkitteln beim Patienten das Ergebnis erheblich verbessern [N.N. 2010c].

Fehler bei der Farbnahme von Zähnen sind nicht nur auf die Umgebungsbedingungen oder menschliche Sehschwächen zurückzuführen. Auch in den physikalischen Eigenschaften der Zähne liegen weitere potenzielle Fehlerquellen hinsichtlich der Farbermittlung. Der natürliche Zahn besitzt für Lichtquellen aus den sichtbaren Wellenlängenbereichen reflektierende, absorbierende und streuende Bereiche. Diese können bei den Schneidezähnen und Eckzähnen unterschiedlich verteilt sein. Die direkte Oberfläche des Zahnschmelzes, welche die oberste Zahnschicht bildet, sowie dessen Form sind hierbei für einen großen Teil der Reflexion des Lichts am Zahn verantwortlich. Der diffundierte Lichtanteil wird in den unteren Dentinschichten des Zahns durch dessen Kristalle und Prismen gebrochen und gestreut. Der Zahnschmelz erzeugt hierbei einen gräulichen Anteil, während die Dentinschicht eher gelblich schimmert. Da der Zahnschmelz nicht überall gleich dick ist, weisen die Zähne somit in der Regel keine homogene Farbe über den ganzen Zahn auf. Weitere Einflussfaktoren hinsichtlich der Farbermittlung sind Verfärbungen, welche in der Regel nicht homogen über den Zahn verteilt sind. Beispielsweise können die Zähne durch Dehydrierung oder durch Ablagerungen von Partikeln tendenziell dunkler werden. Weitere Verfärbungsfaktoren sind in der Hygiene und den Ernährungsgewohnheiten des Patienten zu sehen (z.B. Konsum von Nikotin, Kaffee oder Rotwein)[Krueger-Janson 2006].



Abbildung 2.5: Farbskalen Vita classical (links) und 3D-Master (rechts) (Quelle: [N.N. 2010b])

Neben den physikalisch bedingten Schwierigkeiten bei der Farbnahme können auch handwerkliche Fehler das Ergebnis negativ beeinflussen. Hierbei ist zuerst die Wahl einer geeigneten Farbskala eine wichtige Voraussetzung für eine erfolgreiche vergleichende Farbermittlungen. In der Praxis werden hierfür sogenannte Farbringe verwendet, welche eine bestimmte Anzahl künstlicher Zähne in unterschiedlichen Referenzfarben als Vergleichsmuster beinhalten und in der Regel von den Herstellern künstlicher Zähne angeboten werden (Abbildung 2.5). So bietet die Firma Vita Zahnfabrik H. Rauter GmbH & Co. KG (im Folgenden “Vita” genannt) beispielsweise die Farbsysteme VITAPAN® classical (im Folgenden “classical” genannt) und VITA Toothguide 3D-Master (im Folgenden 3D-Master genannt) an [N.N. 2010b].

Das classical System wurde 1956 durch empirische Untersuchungen entwickelt und wird nach wie vor in Zahnarztpraxen eingesetzt. Es unterteilt die Farben in die vier Gruppen A1-A4 (rötlich-bräunlich), B1-B4 (rötlich-gelblich), C1-C4 (gräulich), D2-D4 (rötlich-grau) [N.N. 2010b]. Die empirische Ermittlung der Farben hat jedoch den Nachteil, dass keine gleichmäßige Verteilung der Musterfarben im Farbraum vorliegt. Dies lässt sich sehr deutlich im standardisierten LCH-Farbraum zeigen (Abbildung 2.7). Das classical Farbsystem enthält nur vergleichsweise wenige Farben, welche zudem keine definierten Abstände haben. Auch decken die classical Farben den Raum realer Zahnfarben nicht optimal ab, denn es existieren sowohl Häufungen als auch unterrepräsentierte Farbteilräume mit über großen Farbdistanzen, was sich negativ auf die manuelle Farbbestimmung von Zähnen auswirkt. [Paravina 2009].

Aufgrund der Nachteile im classical Farbsystem wurde anhand neuerer physikalischer und wissenschaftlicher Erkenntnisse das 3D-Master Farbsystem von Vita entwickelt. Dessen zugehörigen Musterfarben verfügen im Lab-Farbraum untereinander über weitestgehend gleiche Abstände und decken den möglichen Bereich der Zahnfarben wesentlich homogener

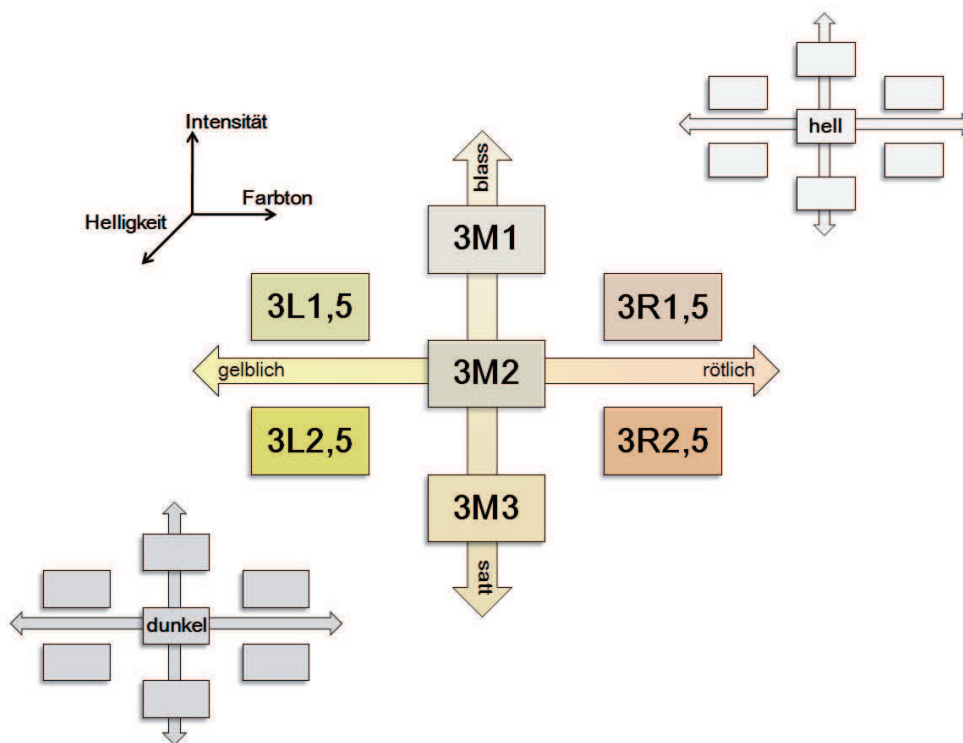


Abbildung 2.6: Der prinzipielle Aufbau des 3D-Master Farbsystem

ab [Baltzer u. a. 2004]. Die Gruppierung der einzelnen Farbmuster erfolgt beim 3D-Master System analog zum Lab-Farbraum nach Helligkeit (0..5 für hell bis dunkel), Farbton (L, M und R für gelblich bis rötlich) und Farbintensität (1..3 für blass bis satt) . Die Nomenklatur der Farbmuster setzt sich entsprechend aus drei Werten zusammen, von 1M1 (0M1 wenn gebleichte Zähne mit abgebildet werden sollen) bis 5M3 (Abbildung 2.6). Die beim 3D-Master vorhandene strukturelle Untergliederung des Farbraums kann zudem für die Farbermittlung genutzt werden. Es empfiehlt sich hierbei, jede Koordinate der Farbe nacheinander zu ermitteln. Hierfür wird zuerst mit den jeweils mittleren Farbmustern (1M1, 2M1, 3M1, 4M1 und 5M1) durch eine Ja/Nein-Entscheidung die passende Helligkeit bestimmt. Anschließend ermittelt der Farbnehmer für die Helligkeitsstufe die Farbintensität und zuletzt wird der passende Farbton ausgewählt [N.N. 2006].

Ein am Lab-Normfarbraum orientiertes Farbsystem kann folglich mögliche Farbwarnungsfehler gegenüber einem experimentell ermittelten Farbsystem deutlich reduzieren. Wichtig zu erwähnen ist auch, dass der Zahnarzt und der Zahntechniker dasselbe Farbsystem zur Kommunikation benutzen müssen, da sich unterschiedliche Systeme in der Regel nicht 1:1 ineinander umrechnen lassen. Weiterhin spielt für die Gestaltung von Farbringen die Anzahl der zur Verfügung stehenden Farbmuster eine große Rolle (beispielsweise 29 beim 3D-Master System gegenüber 16 beim classical-System). Hierbei muss einerseits beachtet werden, dass genug Farbmuster vorhanden sind, um den Farbraum möglichst op-

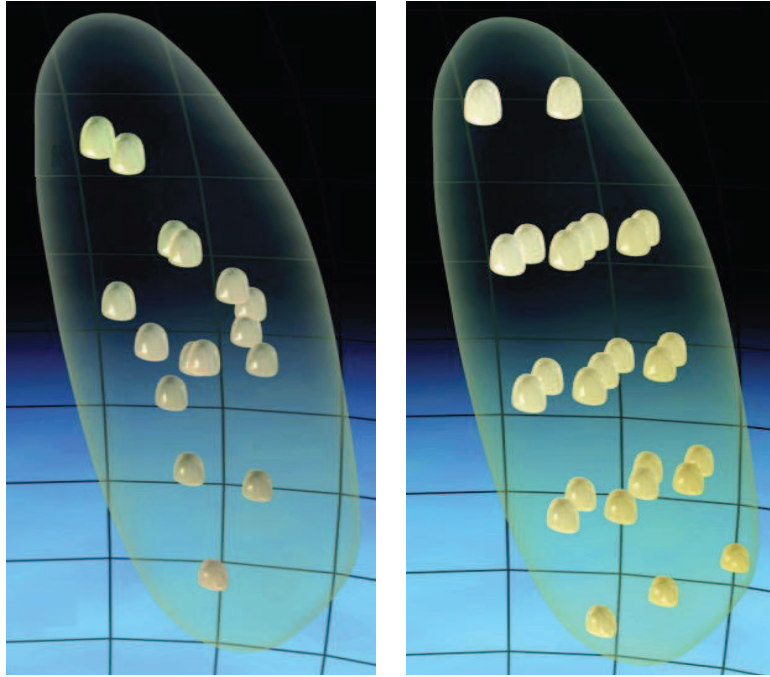


Abbildung 2.7: Die Farbsysteme der Firma Vita classical (links) und 3D-Master (rechts, ohne Helligkeit 0) und deren Lage im LCH-Farbraum (Quelle: Vita Mediendatenbank)

timal abzudecken. Andererseits kann es den Farbnehmer jedoch verwirren, wenn zu viele Farbmuster vorhanden sind, welche sich in der menschlichen Wahrnehmung ähneln. Es besteht die Gefahr, dass dann vom Farbnehmer nur bestimmte Standardfarben verwendet werden und der Rest der Muster ignoriert wird [Müller 2004]. Der Lösungsansatz beim 3D-Master ist, dem Farbnehmer einerseits mit 29 Farben eine vergleichsweise große Menge an möglichen Vergleichsfarbmustern zu Verfügung zu stellen. Andererseits wird durch das mehrstufige Auswahlverfahren diese Menge jedoch deutlich reduziert. Nachdem beispielsweise die Helligkeit ausgewählt wurde, stehen nur noch zwischen drei und sieben potenzielle Referenzfarben zur Auswahl zur Verfügung [N.N. 2006].

Die praktische Erfahrung zeigt jedoch, dass trotz der systematischen Optimierung der Farbmuster, Anleitungen zur Farbnahme und großer Erfahrung bei der Farbermittlung die menschliche Wahrnehmung nach wie vor die entscheidende Fehlerquelle bei der Zahnfarbermittlung bleibt.

2.3 Anforderungen an ein optimiertes Verfahren zur Bestimmung der Zahnfarbe

Wie oben beschrieben kommt der Farbermittlung beim Zahnersatz eine wichtige Rolle zu. Falsch ermittelte Farben, mangelhafte Reproduzierbarkeit und damit einhergehende

Kundenunzufriedenheit bei Patienten sowie hohe Kosten für anstehende Nacharbeit sind die Hauptgründe, die für die Entwicklung einer optimierten Methoden zur Farbermittlung sprechen [Wünnemann 2009, Ross 2002, Bücking 2002, Devigus 2003]. Da ein Großteil der Farbnahme noch immer manuell und visuell durchgeführt wird, ist ein großes Potenzial für den Einsatz automatischer Verfahren gegeben. Hierbei spielt das Preis-/Leistungsverhältnis eine entscheidende Rolle. Ausgehend von der Analyse möglicher Fehlerquellen bei der manuellen Farbermittlung ergeben sich für ein automatisches Verfahren zur Bestimmung der Zahnfarbe und deren zugrundeliegenden messtechnischen Komponenten nachfolgende Anforderungen. Da in der Regel nicht alle Anforderungen zu 100% erfüllt werden können, werden die Anforderungen in “zwingend notwendig” und “wünschenswert” unterschieden und soweit möglich quantifiziert. Tabelle 2.1 gibt einen zusammenfassenden Überblick über die ermittelten Anforderungen.

Plausible Ermittlung der Zahnfarbe (notwendige Anforderung)

Das Verfahren muss zuverlässig und plausibel die “richtige” Zahnfarbe ermitteln. Es muss hierbei die Güte der Ergebnisse bisheriger Methoden überbieten. Wie oben bereits beschrieben, ist die Erfolgsquote bei der Ermittlung der Zahnfarbe schwer quantifizierbar. Selbst wenn mehrere erfahrene Zahntechniker und Zahnärzte den selben Zahn beurteilen, klassifizieren Sie in den seltensten Fällen die gleiche Zahnfarbe (Quelle: Eigene Gespräche mit Zahntechnikern der Firma Vita). Deshalb kann als einziger sicherer Erfolgsfaktor die Zufriedenheit des Patienten genannt werden, welcher jedoch nicht objektiv messbar ist. Aus diesem Grund werden hier folgende Kriterien definiert.

- Die mit dem neuen Verfahren ermittelte Farbe muss bei mindestens 85% mit der von Experten ermittelte Farbe übereinstimmen oder einem direkten Nachbar dieser ermittelten Farbe im Farbraum zugeordnet werden.
- Die mit dem neuen Verfahren ermittelte wahrscheinlichste und zweitwahrscheinlichste Farbe muss bei mindestens 85% direkte Nachbarn im Farbraum sein .

Robuste Ermittlung der Zahnfarbe (notwendige Anforderung)

Das Verfahren muss möglichst robust gegen äußere Einflüsse sein, wie sie in typischen Zahnarztpraxen vorkommen. Dies beinhaltet hauptsächlich Unempfindlichkeit gegen den Einfluss von Fremdlicht wie starke Sonneneinstrahlung oder künstliches Licht von Lampen. Weiterhin soll die unterschiedliche Oberflächenbeschaffenheit beim Zahn des Patienten das Verfahren möglichst wenig beeinflussen.

Reproduzierbare Ergebnisse (notwendige Anforderung)

Das Verfahren muss reproduzierbare Ergebnisse liefern. Dies bedeutet, dass eine Wiederholmessung an der gleichen Stelle eines Zahns auch das gleiche Ergebnis liefern muss, also

auf gleiche Eingangsdaten auch die gleichen Systembewertungen folgen. Wünschenswert wäre zudem, dass unterschiedliche Färbungen an verschiedenen Stellen der Zahnoberfläche ausgeglichen werden können.

Mobil einsetzbares Verfahren (notwendige Anforderung)

Die eingesetzte Messtechnik muss möglichst mobil sein, damit das Verfahren direkt im Praxisraum des Zahnarztes oder im Zahnlabor angewendet werden kann. Eine größere vorbereitende Präparation der Zähne, welche über das normale Reinigen hinausgeht, eine Entnahme von Proben oder eine größere stationäre Ausrüstung muss so weit wie möglich vermieden werden.

Sterilisierbarkeit (notwendige Anforderung)

Die eingesetzte Messtechnik wird direkt im Mund des Patienten angewendet, einschließlich direkter Berührungen im Mundraum. Die gewählte Messtechnik muss deshalb auf sterilisierbaren Komponenten basieren.

Einklang mit Medizinproduktegesetz (notwendige Anforderung)

Medizinische Gerätschaften in der Zahnarztpraxis fallen in der Regel unter das Medizinproduktegesetz (MPG) [N.N. 2012]. Das Verfahren und die zugrunde liegende Messtechnik darf deshalb keine prinzipielle Unverträglichkeit mit dem MPG aufweisen und muss somit dessen Anforderungen genügen.

Kostengünstige Realisierung des Verfahrens (wünschenswerte Anforderung)

Die Zahnfarbe ist zwar ein wichtiges Qualitätskriterium für künstliche Zähne, jedoch muss bei einer automatisierten Lösung zur Bestimmung der Zahnfarbe die Wirtschaftlichkeit im Blick behalten werden. Seit 2005 zahlt die Krankenkasse nur noch einen festen Zuschuss beim Zahnersatz und selbst durch private Zusatzversicherungen oder Bonusmodelle bleibt die Gewinnmarge begrenzt [Braun 2011]. Aus diesem Grund muss das Verfahren und die zugrunde liegende Messtechnik so kostengünstig realisiert werden können, dass eine spätere Vermarktung an Zahnärzte und Zahntechniker wirtschaftlich möglich ist. Für die Arbeit wird deshalb als Anforderung festgelegt, dass die Vermarktung des späteren Produkts für ca. 2000 Euro möglich sein muss.

Einfache und intuitive Bedienung (wünschenswerte Anforderung)

Ein Sensor zur Bestimmung der Zahnfarbe muss intuitiv zu bedienen sein. Dies gilt sowohl für die eigentliche Messung, als auch für eine eventuell benötigte Systemparametrierungen. Es soll möglichst kein Vorwissen bezüglich Hardware, Software oder Grundlagen der Farbmessung erforderlich sein. Im Idealfall wird die Zahnfarbe durch einen einfachen Knopfdruck ermittelt.

Kurze Berechnungsdauer (wünschenswerte Anforderung)

Das Verfahren soll möglichst schnell Ergebnisse liefern. Die Zeit vom Start der Farbermittlung bis zur Ergebnisausgabe sollte innerhalb weniger Sekunden erfolgen.

Funktion für sämtliche Farbräume und Farbsysteme (wünschenswerte Anforderung)

Je nach Hersteller gibt es unterschiedliche Farbsysteme hinsichtlich der Zahnfarbe. So nutzt die Firma Vita beispielsweise das Farbsystem classical und 3D-Master, welche zwar beide den Lab-Farbraum nutzen, dort aber völlig anders verteilt sind. Auch kann natürlich nicht ausgeschlossen werden, dass in Zukunft ein neues Farbsystem entwickelt wird. Aus diesem Grund muss das neue Verfahren unabhängig vom Farbraum und adaptierbar auf neue Farbsysteme sein

Erweiterbar für neue Farben (wünschenswerte Anforderung)

In der heutigen Zeit nimmt der Trend zum Bleichen der Zähne aus ästhetischen Gründen immer weiter zu [Biermann 2009]. Im Farbsystem 3D-Master wurden deshalb eine neue Helligkeitsstufe und drei neue Farben speziell für gebleichte Zähne hinzugefügt (0M1 bis 0M3) [N.N. 2010b]. Aus diesem Grund muss das Verfahren zur Bestimmung der Zahnfarbe neben der prinzipiellen Unabhängigkeit von Farbräumen und Farbsystemen auch erweiterbar für neue Farben sein.

Anforderung	notwendig	wünschenswert	Quantifizierbar
Plausible Ermittlung der Zahnfarbe	✓		Plausibilität bei 85% der Ergebnisse im Vergleich mit Expertenurteil
Robuste Ermittlung der Zahnfarbe	✓		
Reproduzierbare Ermittlung der Zahnfarbe	✓		Reproduzierbare Ergebnisse bei 99%, wenn gleiche Eingangsdaten vorliegen
Mobil einsetzbares Verfahren	✓		Autonomes Verfahren Keine Probenpräparation zulässig Kann in jeder Umgebung eingesetzt werden Gewicht < 1 Kilogramm
Sterilisierbarkeit	✓		
Einklang mit dem Medizinproduktegesetz	✓		
Kostengünstige Realisierung		✓	Produktkosten ca. 2.000 Euro
Einfache und intuitive Bedienung		✓	
Kurze Berechnungsdauer		✓	Berechnungsdauer + Anzeige des Ergebnisses in < 2 Sekunden
Funktion für sämtliche Farbräume und Farbsysteme		✓	
Erweiterbar für neue Farben		✓	

Tabelle 2.1: Anforderungen an ein Verfahren und der zugrunde liegenden Messtechnik zur Bestimmung der Zahnfarbe

Kapitel 3

Stand der Technik

Das Ziel der Arbeit ist die Entwicklung eines robusten und effizienten Verfahrens zur Bestimmung der Zahnfarbe. Aufgrund der in Abschnitt 2.2 beschriebenen Ausgangssituation wird deutlich, dass die manuellen und visuelle Farbbestimmung mit Hilfe von Referenzfarbtafeln aufgrund der menschlichen Wahrnehmung stark fehleranfällig ist. Zwar kann die manuelle Farbbestimmung durch verschiedene Faktoren optimiert werden, beispielsweise mit einer systematischen Verteilung der Referenzfarben im Farbraum, einer mehrstufigen Farbbestimmung sowie durch Übung und Erfahrungsaufbau der Farbnehmer. Jedoch stoßen all diese Maßnahmen schnell an natürliche Grenzen. Ein manuelles Verfahren hat prinzipbedingt starke Defizite hinsichtlich richtiger Klassifikation und Reproduzierbarkeit, weshalb automatisierte Verfahren angestrebt werden. Um ein automatisiertes und robustes Verfahren zu entwickeln, muss zum Einen ein geeigneter Sensor zur Messung physikalisch optischer Spektralwerte ausgewählt werden. Zum Anderen wird eine Systemtechnik und Methode benötigt, welche die Messdaten zuverlässig, automatisch und reproduzierbar auswertet und der entsprechenden Referenzfarbe zuordnet. Das Gesamtsystem muss dem Zahnarzt oder dem Farbnehmer direkt in der Praxis zur Verfügung stehen und in einem akzeptablen Kosten-Nutzen-Verhältnis stehen. Im Folgenden wird deshalb zunächst der Stand der Technik für optische Sensoren zur Farbmessung sowie Methoden zur Datenauswertung vorgestellt. Im weiteren Verlauf der Arbeit werden diese hinsichtlich der gestellten Anforderungen zur Bestimmung der Zahnfarbe analysiert und bewertet.

3.1 Farbmesstechnik

Zur Bestimmung einer Farbe mit Hilfe eines Sensors werden im Prinzip drei Komponenten benötigt. Eine Lichtquelle, welche das zu messende Objekt beleuchtet. Weiterhin eine Komponente, welche das reemittierte Licht empfängt sowie eine Sensoreinheit, welche anhand des reemittierten Lichtes die Farbe bestimmt. Hierfür werden die Messwerte mit den entsprechenden Werten vorgegebener Referenzfarben verglichen. Sensoren zur Bestimmung

der Farbe arbeiten also in der Regel mit einer vergleichenden Messung [Hendrych 2011]. Es gibt zwei grundlegende Messprinzipien für die Farbbestimmung [N.N.]:

Dreibereichsmessung (Kolorimeter)

Die Dreibereichsmessung orientiert sich stark am menschlichen Auge. Hierbei wird ein Objekt mit einer Lichtquelle beleuchtet und die Lichtreflexion mit optischen Filtern in den Rot-, Grün- und Blauanteil zerlegt. Die Filter entsprechen hierbei den menschlichen Zapfenarten. Durch eine entsprechende Intensitätsnormierung können hierdurch die drei Normfarbwerte unmittelbar gemessen und die Farbe im entsprechenden RGB-Farbraum ermittelt werden. Die Messgenauigkeit hängt hier maßgeblich von den Farbfiltern ab.

Spektralphotometrie

Die Spektralphotometrie ermittelt die Farbe anhand des Wellenlängenspektrums des Lichts, welches vom Objekt remittiert wird. Von dem Licht wird der sichtbare Wellenlängenbereich (sichtbares Licht von 400 bis 700nm) durch ein Spektrometer oder Monochromatoren in definierte Wellenlängen aufgeteilt. Aus diesen diskretisierten spektralen Kurven kann auf die Farbe geschlossen werden. Je mehr echte Stützstellen des vorliegenden Spektrums verwendet werden, umso genauer kann die Farbe ermittelt werden.

Sowohl für die Dreibereichsmessung als auch für die Spektralphotometrie existieren unterschiedlichste Messgerätetypen, welche meist auf branchenspezifische Anforderungen ausgelegt sind. Bezüglich der Zahnfarbmessung weist die Dreibereichsmessung starke Nachteile auf. Aufgrund der Farbwahrnehmung des Menschen können verschieden zusammengesetzte Lichtspektren beim Menschen den gleichen Farbeindruck hervorrufen, also hinsichtlich der Farbvalenz ein mehrdeutiges Ergebnis liefern. Dies wird allgemein als Metamerie bezeichnet [N.N.]. Die Folge davon ist, dass bezogen auf die Farbwahrnehmung die Ergebnisse der Dreibereichsmessung nicht exakt genug auf den LCH-Farbraum übertragen werden können. Gleiche oder sehr ähnliche Messergebnisse bei Sensoren nach dem Messprinzip der Dreibereichsmessung können für den menschlichen Betrachter unter einer anderen Lichtquelle einen erheblich anderen Farbeindruck hervorrufen. Diese Messart kann für die hier vorliegende Aufgabe der Zahnfarbmessung also zu mehrdeutigen Ergebnissen führen [N.N. 2005b]. Im Gegensatz dazu decken Spektrometer das gesamte Spektrum des sichtbaren Lichts ab, wodurch eine eindeutige Zuordnung zum LCH-Farbraum gegeben ist. Somit ist die Methode prinzipbedingt zur Bestimmung der Zahnfarbe zu bevorzugen.

Bei den vorkommenden Messgeräten kann man für beide Messmethoden zwischen folgenden Arten der Messgeräte unterscheiden: Es gibt einerseits die große Gruppe der hochwertigen, meist stationären industriellen Farbmessgeräten und andererseits vergleichsweise einfache,

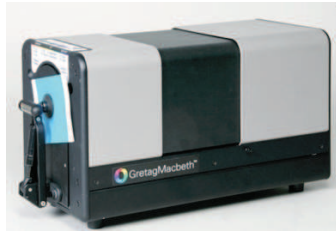
mobile und handliche Sensorsysteme. Die mobilen, einfache Sensorik ist meist einfach zu bedienen, preiswert und flexibel bezüglich der Messstelle. Jedoch werden diese Vorteile oft auf Kosten der Genauigkeit erkaufte. Die hochwertigen und stationären Messgeräte sind entsprechend teurer und unflexibel bei der Probenhandhabung, ermöglichen jedoch in der Regel eine deutlich höhere Messgenauigkeit. Weiterhin verfügen derartige Messgeräte oft noch über zusätzliche Messfunktionen, beispielsweise eine Messung der Transmission, wodurch die Ergebnisgüte weiter gesteigert werden kann [N.N.]. Im Folgenden werden beispielhaft am Markt verfügbare industrielle Farbmessgeräte sowie handliche Sensoren hinsichtlich deren Eignung zur Bestimmung der Zahnfarbe beschrieben.

3.1.1 Farbmessgeräte im industriellen Einsatz

Es existieren eine Vielzahl von hochgenauen Farbmessgeräten, sowohl nach dem Prinzip der Dreibereichsmessung als auch der Spektralphotometrie, wobei die Spektralphotometrie aufgrund der genannten Vorteile deutlich verbreiteter ist. Neben stationären Labormessgeräten werden immer mehr flexible Farbmessgeräte bezüglich der Anforderungen an zu messende Proben entwickelt [N.N.]. Deutlich wird dies auch durch den Einsatz von Farbmessgeräten in unterschiedlichsten Industrieapplikationen, beispielsweise zur Farbmessung an Stahl [Pryor 2011], zur Erkennung von Verunreinigungen in der Lebensmittelindustrie oder im Recycling [Aderhold u. a. 2010]. Industrielle Farbmessgeräte lassen sich je nach Einsatzgebiet grob in die folgenden drei Kategorien gliedern:

1. **Labormessgeräte**, welche in der Regel unter Laborbedingungen eingesetzt werden und einen festen abgeschirmten Probenplatz benötigen (Abbildung 3.1 a). Wenn mobile Flexibilität oder eine Messung direkt im Prozess keine zwingende Anforderung ist, sind diese Geräte in der Regel erste Wahl, da sie höchstmögliche Genauigkeit und Reproduzierbarkeit liefern.
2. **Inline-Prozessmessgeräte**, welche in der Regel stationär im laufenden Prozess berührungslos an bewegten Gütern oder Flüssigkeiten eingesetzt werden (Abbildung 3.1 b und c). Gängige Vertreter dieser Klasse sind Farmkameras oder berührungslose Spektrometer [Pflüger 2010, N.N.].
3. **Messwagen mit flexiblen Messkopf**, welche meist für die mobile Messung von Stichproben eingesetzt werden (Abbildung 3.1 d).

Bezüglich der Messung von Zahnfarben konnte bisher entsprechend den oben genannten Kategorien noch kein hochgenaues industrielles Messgerät konzipiert werden. Ein Hinderungsgrund ist, dass ein Messgeräte für die Zahnfarbmessung zu einer intraoralen Messung fähig sein muss. Aus diesem Grund können Sensoren aus der Klasse der Labormessgeräte



(a) Stationäres Laborspektrometer Color i5; Quelle: www.farbkarten-shop.de



(b) Berührungsloses Spektrometer VS 450; Quelle: www.farbkarten-shop.de



(c) Kamera mit Spektrometeroptik LuxVIS; Quelle: www.zeutec.com



(d) Laborspektrometer sph-IS mit flexibler Messsonde; Quelle: www.colorlite.de

Abbildung 3.1: Verschiedene Geräte zur industriellen Farbmessung

mit einem fixen, abgeschirmten Probenplatz als ungeeignet für die Bestimmung der Zahnfarbe angesehen werden.

Eine weitere Besonderheit für die Farbmessung ist in der Beschaffenheit der Zähne zu sehen. Da Zähne Transluzenzen und unterschiedliche Schichten aufweisen, muss das Licht bis in den Dentalbereich vordringen. Nur so wird eine ausreichend genaue Messung der Zahnfarbe zu gewährleisten. Da dieser Bereich stark für die Lichtabsorption und entsprechenden Farbwahrnehmung verantwortlich ist. Deshalb sollte sowohl das Anregungslicht als auch die Messeinheit möglichst nahe am Zahn positioniert sein. Bei größeren Abständen von Sensor und Zahn kann sonst Fremdlicht die absorbierten Wellenlängen durch den Zahn überlagern. Dieser Effekt wird durch die diffuse Reflektion an der Zahnoberfläche weiter verstärkt. Dadurch wird die Messung ungenau und kann kaum reproduziert werden [N.N. 2010b]. Hochwertige Komponenten können diese Nachteile zwar teilweise kompensieren, wodurch die Kosten für das Messgerät jedoch stark steigen. Aus diesen Gründen können Sensoren aus der Klasse der berührungslosen inline-Messgeräte ebenfalls als ungeeignet für die vorliegende Aufgabe angesehen werden.

Zusammenfassend kann festgestellt werden, dass Laborgeräte mit abgeschirmten Messbereichen oder mit Probenvorbereitung (3.1 a) sowie berührungslose Messverfahren (3.1 b)

und c) aufgrund der genannten Defizite als ungeeignet für die Bestimmung der Zahnfarbe angesehen werden können. Spektrale Messgeräte mit einem flexiblen Messkopf (3.1 d) kommen, sofern sie sterilisierbar sind, für die hier vorgestellte Anwendung prinzipiell in Frage. Jedoch bewegen diese sich immer noch in einem wesentlich höheren Preissegment als einfache handliche Farbmessgeräte.

3.1.2 Handliche Farbmessgeräte

Es existieren bereits mehrere handliche Messgeräte, welche speziell auf die Ermittlung der Zahnfarbe ausgelegt sind. Die gängigsten Geräte basieren aufgrund der oben genannten Metamerie auf dem Messprinzip der Spektralphotometrie. Hierbei gibt es unterschiedliche Schwerpunkte beim Design und den verwendeten Komponenten (Abbildung 3.2). So besitzen beispielsweise das VITA Easyshade und Crystaleye kompakte Gehäuse mit direkt integriertem Messkopf, während der Sensor Digital Shade Guide von der Firma Rieth aus einem größeren Gehäuse mit integrierter Auswerteeinheit sowie einen stiftförmigen Messkopf besteht [N.N. 2010a, Kourtis u. a. 2009, N.N. 2008b].



Abbildung 3.2: Verschiedene handliche Sensoren zur Ermittlung der Zahnfarbe, VITA Easyshade (links), Crystaleye (mitte), DSG (rechts) (Quellen: [N.N. 2010b, N.N. 2010a, N.N. 2008c])

Im Folgenden wird der Sensor VITA Easyshade exemplarisch für die Klasse der handlichen Sensoren zur Ermittlung der Zahnfarbe hinsichtlich technischer Details sowie der Vor- und Nachteile im praktischen Einsatz untersucht:

VITA Easyshade

Das VITA Easyshade ist ein tragbarer faseroptischer Sensor, welcher speziell zur intraoralen Messung von Zahnfarben optimiert wurde. Das System besteht aus einem Handstück mit Messsonde sowie einer Basisstation mit Halterung und Kalibrierblock. In dem Handstück ist neben der Beleuchtung auch die spektrale Auswerteeinheit und ein Mikroprozessor in-

tegiert. Das Messverfahren nutzt hierbei ein Lichtleitersystem zur Beleuchtung des Zahns und zur Detektion des von der Dentinschicht des Zahns reflektiertem Licht. Als spektrale Auswerteeinheit kommt eine mechanisch stabile Konstruktion aus Interferenzfiltern und Wandlern zum Einsatz. Damit wird eine spektrale Auflösung von ca. 25nm im sichtbaren Lichtbereich erzielt. Zur Wahrung der Hygiene bei intraoralen Messungen kann die Messsonde mit einer speziell angefertigten Einweg-Schutzfolie überzogen werden.

Jede Farbmessung ist ein vergleichender Vorgang, weshalb der Sensor vor dem eigentlichen Messvorgang kalibriert werden muss. Dies ist nötig, da einerseits die Auswirkungen der Schutzfolie auf die Lichtquelle kompensiert werden müssen und andererseits die Messkomponenten gewissen alterungs- und umgebungsbedingten Schwankungen unterworfen sind. Für die Kalibrierung wird zuerst ein Keramikblock mit bekannter Farbe (welcher sich in der Halterung des Sensors befindet) als Referenzmaterial gemessen. Die Kalibrierung erfolgt daraufhin automatisch auf Knopfdruck und die entsprechenden Korrekturfaktoren für die spektralen Messwerte werden berechnet.

Nach dem Kalibrieren ist der Sensor bereit für Messungen an Zähnen oder Dentalmaterialien. Mit Hilfe eines Displays können verschiedene Einstellungen vorgenommen werden, beispielsweise die Auswahl des gewünschten Farbraums. Anschließend wird die Messsonde mit der zu messende Oberfläche in Kontakt gebracht und per Knopfdruck der Messvorgang ausgelöst. Beim Messvorgang wird zuerst bei geschlossener Sonde nur das Anregungslicht gemessen, um bei der eigentlichen Reflexionsmessung Fremdlichteinflüsse kompensieren zu können [Kourtis u. a. 2009].

Für eine direkte Auswertung der Messdaten im VITA Easyshade auf Basis von classical und 3D-Master Farbräumen ist der Mikroprozessor H8SX/1653 von Renesas [N.N. 2005a] in das Handstück des Sensors integriert. Somit kann das VITA Easyshade als Smart-Sensor bezeichnet werden. Sämtliche benötigten Vorgänge beim Messprozess (Kalibrieren, Messen, Korrekturfunktionen, Messdatenauswertung, Ergebnisverwaltung und Visualisierung) werden hierbei automatisiert durchgeführt, ohne dass ein externer Computer benötigt wird. Zur Dokumentation und Weiterverarbeitung der Messdaten auf externen PCs ist eine USB-Schnittstelle vorhanden.

Das VITA Easyshade ist folglich ein auf die Zahnmessung angepasster Sensor, der in Aufbau und Funktionsweise viele spezifische Probleme berücksichtigt. Nachteilig ist jedoch zu nennen, dass einige potenzielle Fehler zwar technisch korrigiert aber nicht komplett beseitigt werden können. So ist das Messergebnis nach wie vor beeinflusst von Bedienerfehlern wie Verkippen des Sensors. Auch muss prinzipbedingt Kontakt zwischen der Messsonde

und der Oberfläche des Zahns bestehen. Nichthomogene Zahnoberflächen und die Zahnschmelzdicke beeinflussen das Ergebnis ebenfalls.

3.1.3 Prinzipien für die Auswahl von Farbmessgeräten

Bei der Auswahl geeigneter Farbmessgeräte bieten viele Hersteller Checklisten an, um Messprinzipien oder geeignete Sensoren zu ermitteln. Ein fragenbasierter Leitfaden nach [Schröder 2004] sowie dessen Beantwortung hinsichtlich der spezifischen Probleme bei der Bestimmung der Zahnfarbe ist im folgenden dargestellt:

1. Frage: Was und welche Art von Proben sollen hauptsächlich gemessen werden?
 - Primär reale Zähne von Patienten ohne Probenvorbereitung
 - Sekundär künstlicher Zahnersatz, Probenvorbereitung möglichFolgerung: Es muss ein Messgerät gewählt werden, welches am Menschen intraoral genutzt werden kann und mit dem Medizinproduktegesetz in Einklang steht.
2. Frage: Ist die Probe homogen und gibt es besondere Herausforderungen bezüglich der Geometrie?
 - Natürliche Zähne haben in der Regel keine komplett homogene Farbe, jedoch meist auch keine gravierenden Farbunterschiede.
 - Natürliche Zähne weisen leichte Rundungen an der Oberfläche auf.Folgerung: Der Messfleck muss mindestens 5mm Durchmesser aufweisen, um die Farbunterschiede im Zahn und Geometrieunebenheiten auf der Oberfläche ausreichend abzubilden.
3. Frage: Wo soll gemessen werden?
 - Messungen müssen in der Zahnarztpraxis durchgeführt werden könnenFolgerung: Das Messgerät darf keine besonderen Anforderungen an die Umgebung aufweisen, muss jedoch auch nicht resistent gegenüber extremen Umweltbedingungen sein.
4. Frage: Müssen die Daten unternehmensübergreifend vergleichbar sein?
 - Die Messungen müssen unternehmensübergreifend vergleichbar sein. Der Zahnarzt, welcher die Farbe misst und der Zahntechniker sind in der Regel in unterschiedlichen Unternehmensorganisationen und räumlich getrennt.Folgerung: Das Messgerät sollte eine gute Instrument zu Instrument Vergleichbarkeit besitzen. Dadurch wird gewährleistet, dass der Zahntechniker den Zahnersatz nach der Farbvorgabe des Zahnarztes fertigen und Kontrollieren kann.
5. Frage: Welche Auswertungen sind erforderlich?
 - Die Farbe muss exakt bestimmt werden.

Folgerung: Eine reine Differenzauswertung ist nicht ausreichend, die Farbe muss möglichst exakt durch eine vergleichende Messung bestimmt werden.

6. Frage: Wer soll die Messung durchführen?

→ Zahnärzte und Zahntechniker.

Folgerung: Das Gerät muss einfach zu bedienen sein und muss die Messdaten direkt auswerten. Expertenwissen zur Farbmessung, Auswertung oder Probenvorbereitung kann nicht vorausgesetzt werden.

7. Frage: Wird eine Qualitätskontrolle oder Qualitätssicherung durchgeführt?

→ Qualitätssicherung und Qualitätskontrolle.

Folgerung: Der primäre Anwendungsfall ist der Einsatz beim Zahnarzt. Dieser liefert durch die Farbbestimmung die Basisinformation für den Zahntechniker. Dies kann im erweiterten Sinne als qualitätssichernde Maßnahme angesehen werden. Eine Kontrolle der Farbe des Zahnersatzes findet beim Zahntechniker während und nach der Fertigung des Zahnersatzes statt. Dies kann als Qualitätskontrolle angesehen werden.

Anhand der vorgestellten Kriterien zur Auswahl von Messgeräten und den aufgestellten Anforderungen bietet ein handlicher Sensor mit spektralen Messdaten das größte Potenzial [Faber 2006, Müller 2004]. Prinzipiell gibt es somit zwei Ansätze für eine automatische Bestimmung der Zahnfarbe. Die erste liegt in der Nutzung hochwertigerer Messgeräte als die bisherigen handlichen Sensoren. Hierfür kann ein hochwertiges Messgerät mit flexiblen Messkopf auf die Bestimmung der Zahnfarbe hin optimiert werden. Weiterhin besteht die Möglichkeit, handliche Sensoren mit verbesserten Komponenten auszustatten, um deren Messqualität zu erhöhen. Da dies die Hardware jedoch verteuern würde ist dieser Ansatz hinsichtlich des Marktpotenzials als nachteilig einzustufen. Ein weiterer Ansatz liegt in der Verbesserung der Auswertelgorithmen. Nahezu alle handlichen Messgeräte nutzen einfache lineare Klassifikatoren auf Basis von Abstandsmaßen im Farbraum. Intelligente Auswerteverfahren und Modellierungsmethoden bieten ein großes Potenzial, um die Genauigkeit der Farbprognose erheblich hinsichtlich der Farbwahrnehmung zu verbessern. Da hierfür nur die Software der bestehenden Geräte erweitert werden muss, sind keine oder nur sehr geringe zusätzliche Hardwarekosten, beispielsweise hinsichtlich Rechenleistung und Speicherplatz, notwendig. Dieser Ansatz kann somit als vielversprechend angesehen werden, da einem preiswertes Messgerät mit guter Farbbestimmung ein großes Marktpotenzial eingeräumt werden kann. Voraussetzung ist allerdings, dass die Klassifikationsgüte durch entsprechende Auswertelgorithmen auch entsprechend gesteigert werden kann. Mögliche Klassifikationsansätze werden im Folgenden genauer untersucht.

3.2 Datenauswertung und Klassifikation

Aufgrund der oben genannten Rahmenbedingungen zur Bestimmung der Zahnfarbe sind auf Spektroskopie basierende handliche Sensoren oder Tischgeräte mit flexiblen Messkopf für die vorliegende Aufgabe geeignet. Speziell für handliche Messgeräte besteht die Herausforderung darin, ein geeignetes Auswerteverfahren zu entwickeln, welches das gemessene Wellenlängenspektrum robust und reproduzierbar einer definierten Farbe im LCH-Farbraum zuordnet. Wegen der spektralen Natur der Messdaten ist im Gegensatz zur Dreibereichsmessung eine klare Zuordnung von Spektren zu Musterfarben nur bedingt intuitiv darstellbar. Ebenso wenig können mit dem bloßen Auge die verschiedenen Spektren den Musterfarben eindeutig zugeordnet werden, da es sich um mehrfache Superpositionen von Spektren der einzelnen Stoffwechselprodukte der unterschiedlichen Zahnschichten handelt. Diese unterscheiden sich oft nur in einzelnen Bereichen und sind zudem abhängig von der jeweiligen Schichtdicke. Dies ist auch der Grund, weshalb die Betrachtung einzelner Wellenlängen nicht ausreichend ist, um eine eindeutige Zuordnung zu den Farben vornehmen zu können. Daher muss ein entsprechendes Verfahren auf jeden Fall in der Lage sein, das gesamte relevante Spektrum bei der Klassifikation zu berücksichtigen. Für die Auswertung und Klassifikation kommt erschwerend hinzu, dass die gemessenen Spektren häufig stärkeren Schwankungen unterworfen sind. Gründe hierfür sind in Fehlern bei den Messkomponenten, Bedienerfehler sowie der inhomogenen Zahnoberfläche zu sehen. Aufgrund dieser Fehlerquellen muss eine leistungsfähige Methode gewählt werden, welche derartige Schwankungen ausreichend kompensieren und weitestgehend fehlertolerant verarbeiten kann.

Im Folgenden werden ausgewählte Methoden vorgestellt, welche sich prinzipiell zur Erstellung eines Modells zur Klassifikation eignen. Mit solchen Methoden ist es möglich, anhand der spektralen Eingangsdaten eine eindeutige, robuste und reproduzierbare Zuordnung zu entsprechenden diskreten Referenzfarben zu bilden.

3.2.1 Modellierungsmethoden

In der Wissenschaft und Technik werden oft Modelle verwendet, um komplexe Sachverhalte darzustellen. Die gewählten Modelle sollen dabei auftretende Phänomene sinnvoll beschreiben und somit auch Vorhersagen und Prognosen erlauben. Die Modelle werden hierbei typischerweise wie folgt eingesetzt: Gegeben ist eine Menge von X Eingangs- sowie Y Ausgangswerten. Die Eingangswerte werden Kenngrößen genannt und sind in der Regel skalare Werte, die zu einem n -Tupel $x = (x_1 \dots x_n) \in X$ zusammengefasst werden. Der Wertebereich der Ausgangswerte kann diskret oder kontinuierlich sein. Ziel ist es nun ein Modell zu finden, welches bei allen möglichen Kombinationen von Eingangswerten die zugehörigen Ausgangswerte zuordnet. Für die Prognose wird also ein Algorithmus benötigt, welcher ein

geeignetes Modell repräsentiert. Dieses Modell ist dabei eine Funktion $f : X \rightarrow Y$, die für jeden Eingangswert $x \in X$ einen möglichst guten Näherungswert des erwarteten Ausgangswertes $y \simeq f(x)$ liefert [Hastie u. a. 2009].

In der vorliegenden Aufgabenstellung sind die Kenngrößen die Werte eines spektralen Sensors, die Ausgangswerte die zugehörigen Referenzfarben als diskrete Bezeichnung. Je nach Farbraum sind auch Teilgrößen der Farbe wie Helligkeit, Intensität oder Farbton als Ausgangswert denkbar. Somit liegt eine Klassifikationsaufgabe vor mit dem Ziel, jeder Kombination von Eingangsgrößen eine zugehörige Farbklasse zuzuordnen. Zur Lösung einer Klassifikationsaufgabe kann im Prinzip zwischen zwei Arten von Modellierungsmethoden unterschieden werden. Einerseits kann ein mathematisches Modell erstellt werden, um den Sachverhalt exakt abbilden zu können. Auf der anderen Seite können datenbasierte Ansätze verwendet werden, um ein Modell durch maschinelles Lernen anhand von Beispieldatensätzen zu generieren.

Mathematische Modelle

Generell liefern exakte mathematische Beschreibungen von physikalischen Zusammenhängen bei entsprechender Modellgüte die genauesten Ergebnisse hinsichtlich eines Prognosemodells. Der zu modellierende Sachverhalt wird durch mechanische, chemische, elektrische und physikalischen Gesetzmäßigkeiten wie beispielsweise der Lichtbrechung, Reflexionsgesetzen oder Absorptionen von Wellenlängen beschrieben. Eine solche Modellierung ist zwar meist hochgenau, jedoch schon für einfachste physikalische Zusammenhänge hoch komplex. In der Regel werden für eine ausreichend genaue Modellierung mehrdimensionale gekoppelte Differenzialgleichungen benötigt. Diese Gleichungen weisen dazu noch oft spezifische Konstanten auf, welche ebenfalls von Zustandsgrößen abhängen können, beispielsweise umgebungsabhängige Materialkonstanten. Vor allem im Bereich der optischen Sensoren werden in der Praxis nahezu keine mathematische Modelle zur Ergebnisprognose von Sensordaten eingesetzt. Zwar sind die Grundlegenden physikalischen Zusammenhänge sehr umfassend erforscht, jedoch ist die Menge an möglichen Einflüssen sehr groß. Für eine Modellierung der Beleuchtung müssten beispielsweise neben diversen Lichtquellen wie ambientes Licht (Grundausleuchtung der Umgebung), diffuses Licht (positionsabhängiges Licht) oder spekulares Licht (Glanzlicht) auch deren Wechselwirkungen mit der Oberfläche bei unterschiedlicher Beschaffenheit, Geometrie und Beleuchtungswinkeln modelliert werden [Peters 2010, Dana u. a. 2007]. Zudem müsste eine Möglichkeit geschaffen werden, sämtliche individuellen Eigenschaften von Zähnen zu erfassen und zu modellieren. Aus diesen Gründen war es bisher nicht möglich, ein entsprechendes praxistaugliches mathematisches Modell zu erstellen.

Weiterhin besteht bei mathematischen Modellen generell das Problem der Störgrößenanfälligkeit. Derartige Modelle gehen meist von einem idealisierten Zustand aus, welcher in der Praxis nur selten vorkommt. Da jedoch nur tatsächlich modellierte Störgrößen und Randbedingungen bei der Auswertung berücksichtigt werden können, besteht hier eine große Fehlerquelle für Prognosewerte hinsichtlich realer Sensordaten. Ein handlicher spektraler Sensor weist viele mögliche Störgrößen wie beispielsweise ein Verkippen des Sensorkopfes oder uneben Zahnoberflächen auf. Außerdem sind die Sensordaten wie oben dargestellt nahezu immer mit mehr oder weniger starken Störungen behaftet, welche hier in der Regel nicht ausreichend abgeschätzt werden können [Filip u. a. 2009].

Zusammenfassend lässt sich feststellen, dass eine praxistaugliche mathematische Modellierung für eine Farbmessung mit einem optischen Sensor erhebliche Anforderungen an die Abbildung der physikalischen Zusammenhänge sowie der Berücksichtigung möglicher Störgrößen stellt. Die mathematische Modellierung kann deshalb für die gegebene Aufgabenstellung als ungeeignet angesehen werden.

Datenbasierte Modelle auf Basis maschinellen Lernens

Aufgrund des häufig unverhältnismäßig großen Aufwands, welcher zur Erstellung von mathematischen Modellen nötig ist, werden für die Modellierung auch empirische, also datenbasierte Ansätze des maschinellen Lernens verwendet. Diese Modelle werden auf der Basis von real ermittelten Daten (auch Trainingsdaten oder Beobachtungen genannt) berechnet. Diese Art der Modellierung und entsprechende Lernverfahren sind in Wissenschaft und Technik weit verbreitet und kommt in vielen praktischen Anwendungen wie beispielsweise der Prozessüberwachung zum Einsatz [Teti u. a. 2010, Schmidberger u. a. 2007, Rauh u. a. 2001].

Zur Erstellung des Modells wird dem Algorithmus eine Menge $D \in X \times Y$ an Trainingsdatensätzen vorgegeben. Jeder dieser Datensätze $(x, y) \in D$ besteht aus einem Tupel von Eingangskenngrößen x und dem zugehörigen Ausgangstupel y . Auf Basis der Trainingsdaten wird nun ein geeignetes Modell f erzeugt, indem die Modellparameter auf die vorhandenen Trainingsdaten angepasst werden [Jankowski u. a. 2006]. Den zum Training verwendeten Daten kommt hierfür eine entscheidende Bedeutung zu. Üblicherweise kann nur ein kleiner Teil der tatsächlich möglichen Kombinationen an Eingangsdaten ermittelt und eingesetzt werden, weshalb die Trainingsdaten einen möglichst repräsentativen Bereich abdecken müssen. Hinsichtlich des Modells genügt es nicht, nur die Trainingsdaten quasi auswendig zu lernen. Das Modell muss auch für unbekannte Eingangstupel sinnvolle Ergebnisse liefern, also eine ausreichende Fähigkeit zur Abstraktion aufweisen [Hastie u. a. 2009]. Die Verallgemeinerungsfähigkeit oder Generalisierung eines Modells ist somit auch ein wesentlicher Indikator für dessen Güte. Zu deren Berechnung wird üblicherweise eine Abweichungsfunk-

tion L festgelegt. Diese vergleicht eine vom Modell ermittelte Ist-Prognose $f(x)$ mit der zu x gehörigen Ausgangsgröße y . Der Fehlerwert $L(y, f(x))$ ist hierbei umso größer, je weiter die Ist-Prognose vom Soll-Wert abweicht. Für Modelle mit hoher Güte muss der Fehlerwert L folglich möglichst minimiert werden.

In der Praxis ist die exakte Berechnung des Fehlerwertes der Modellgüte nicht möglich, da hierfür sämtliche Kombinationen an Eingangstupel nötig wären. Deshalb wird lediglich eine Abschätzung vorgenommen. Hierzu werden in der Regel die Trainingsdaten in Lern- und Verifikationsdaten aufgeteilt. Aus der Menge der Trainingsdaten werden also m Datensätze $(x_1, y_1), \dots, (x_m, y_m) \in D$ als Verifikationsdaten ausgewählt. Diese werden, im Gegensatz zu den übrigen Lerndaten, nicht für das eigentliche Training des Modells verwendet. Stattdessen dienen diese Daten der Modellverifikation, also zur Abschätzung des Mittelwerts der Fehlerwerte für die ausgewählten Verifikationsdaten. Der Fehlerwert wird wie folgt berechnet [Bishop 2006]:

$$R[f] = \frac{1}{m} \sum_{i=1}^m L(y_i, f(x_i))$$

Mit datenbasierten Modellen können also komplexe Zusammenhänge zwischen Eingangsdaten und Zielgrößen näherungsweise abgebildet werden. Der große Nachteil der datenbasierten Modelle liegt darin, dass nur Zusammenhänge erkannt und modelliert werden können, welche sich zumindest indirekt in den Trainingsdaten abbilden. Damit die Modelle entsprechend aussagekräftig sind, ist somit oft ein relativ großer Aufwand nötig, um repräsentative Trainingsdaten zu ermitteln. Entsprechend hängt die Güte des Modells von der Qualität der Trainingsdaten ab. Idealerweise wird für die Ermittlung der Trainingsdaten ein statistischer Versuchsplan verwendet. Dadurch kann sichergestellt werden, dass Wechselwirkungen und Abstraktionen möglichst gut abgebildet werden [Montgomery 2009, Kleppmann 2003].

Zusammenfassend lässt sich feststellen, dass mit einer genügend großen Menge an repräsentativen Trainingsdaten das maschinelle Lernen eine sehr vielversprechende Methode darstellt, um ein geeignetes Modell zur Farbklassifikation zu entwickeln.

3.2.2 Klassifikationsverfahren mit maschinellem Lernen

Künstliche Neuronale Netze sind eines der ältesten und bekanntesten Verfahren für eine datenbasierte Modellbildung durch maschinelles Lernen und kommen in vielen technischen Verfahren zum Einsatz. In den letzten Jahren wurden weitere neue Verfahren entwickelt, welche in allgemeinen Leistungsvergleichen sehr gut abschneiden [Guyon 2009, Guyon u. a. 2008, Guyon u. a. 2007, Lutz 2006, Guyon u. a. 2006]. Neben Verfahren aus der

Klasse der Neuronalen Netzen und Neuro Fuzzy werden deshalb beispielhaft im Folgenden kernel-basierende Verfahren, Bayes-Klassifikatoren und Boosting-Verfahren untersucht. Als Randbedingung wird bei der Betrachtung der Verfahren davon ausgegangen, dass diskrete Ausgangsgrößen vorliegen, da dies keine Einschränkung bei der gestellten Aufgabe darstellt.

Support Vector Machines (SVM)

Support Vector Machines sind ein Kernel-basiertes Verfahren mit einem reell wertigen Eingangsraum $X \in \mathbb{R}^d$. Hierbei wird versucht die einzelnen Klassen durch eingepasste Hyperebenen zu trennen. Hierfür werden die Daten in einen geeigneten hochdimensionalen Raum F überführt, in dem die Klassen möglichst linear separierbar sind. Als Maß für die Trenngüte der Hyperebene werden in der Regel die Abstände der Datenvektoren zur Ebene verwendet.

Der gesuchte Raum F kann in der Regel nicht direkt definiert werden, weshalb man sogenannte Kernels nutzt. Dies sind Funktionen, welche F durch ein Skalarprodukt der Form $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow F$ induzieren. Dadurch können implizite Berechnungen in F durchgeführt werden. Voraussetzung dafür ist, dass F ein Hilbertraum ist und eine Abbildung $\Phi : X \rightarrow F$ existiert, welche die Eingangswerte in F projiziert [Cawley u. a. 2010, Cristianini u. a. 2000].

Gängige Kernel Funktionen sind beispielsweise der Gauß-Kernel, der polynomiale Kernel oder ein linearer Kernel, welche wie folgt definiert sind [Guyon u. a. 2009]:

$$K_{Gau\beta}(x_i, x_j) = e^{-(x_i - x_j)^2 \cdot \sigma^{-2}}$$

$$K_{Poly}(x_i, x_j) = (1 + x_i^T \cdot x_j)^q$$

$$K_{Lin}(x_i, x_j) = x_i^T \cdot x_j$$

Der Parameter σ ist hierbei die Empfindlichkeit des Gauß-Kernels, welcher der Breite der Gauß-Glocke entspricht. Der Parameter q repräsentiert den Grad des Polynoms.

Eine prinzipielle Voraussetzung für einen erfolgreichen Einsatz des Verfahrens ist die Auswahl eines geeigneten Kernels. Dieser muss so angepasst werden, dass eine möglichst gute Separation der Eingangsdaten ermöglicht wird. Optimierungsverfahren bieten hier die Möglichkeit eine Kernelfunktion empirisch direkt auf die Trainingsdaten anzupassen. Um zu verhindern, dass dadurch die Generalisierungsfähigkeit des Modells sinkt, nutzt man sogenannten Schlupfvariabel. Diese erlauben es dem Klassifikator, eine gewisse Anzahl an Datensätzen "falsch" zu klassifizieren, wodurch eine Überanpassung des Modells verhindert werden kann. Auf diese Weise erreicht man auch mit Standard-Kernels häufig sehr gute Ergebnisse [Hüllermeier u. a. 2010].

Naive Bayes Klassifikator (NBK)

Bayes-Klassifikatoren basieren auf dem Bayestheorem, welches jeder Eingangsgröße $x \in X$ die Ausgangsklasse mit der größten Wahrscheinlichkeit zuordnet:

$$P(c_i | x) = \frac{P(x | c_i) P(c_i)}{P(x)}.$$

Dabei ist $P(x)$ die Wahrscheinlichkeit, welche das Eintreten des Ereignisses x repräsentiert. $P(c_i)$ ist die sogenannte A-Priori-Wahrscheinlichkeit. Diese gibt an, ob eine beliebige Eingangsgröße zur Klasse c_i gehört. $P(x | c_i)$ beschreibt die Wahrscheinlichkeit, mit der x in der Klasse c_i vorkommt, was als klassenbedingte Wahrscheinlichkeit bezeichnet wird. Aus diesen drei Wahrscheinlichkeiten lässt sich die sogenannte A-Posteriori-Wahrscheinlichkeit $P(c_i | x)$ berechnen, welche angibt, wie wahrscheinlich x zur Klasse c_i gehört.

Mit Hilfe des Bayestheorems könnte nun theoretisch der optimale Bayes-Klassifikator berechnet werden, welcher für jede Klassifikationsaufgabe das garantiert optimale Ergebnis liefert [Jankowski u. a. 2006]. Da jedoch in der Praxis die benötigten Wahrscheinlichkeiten in der Regel nicht ermittelt werden können, muss auf vereinfachte Annahmen zurückgegriffen werden. Eine Möglichkeit der Vereinfachung liefert der Naive Bayes Klassifikator. Hierbei wird die Annahme getroffen, dass die Eingangskenngrößen linear unabhängig sind. Dies trifft zwar nur selten in der Praxis zu, jedoch vereinfacht sich dadurch die Ermittlung der Zugehörigkeitswahrscheinlichkeiten enorm. Trotz dieser Vereinfachung werden in der Praxis mit dem NBK häufig gute Ergebnisse erzielt, besonders dann, wenn nur wenig Trainingsdaten zur Verfügung stehen [Hand u. a. 2001].

Zur Parametrierung des NBK werden mit dem Trainingsalgorithmus anhand der vorgegebenen Trainingsdaten und den vorkommenden Klassen im Ereignisraum Y die bedingten Wahrscheinlichkeiten $P(x_j | y)$ ermittelt. Bei einem kontinuierlichen Wertebereich für die Eingangsgrößen x_j werden hierfür zunächst Schwellwerte gebildet, um den Wertebereich künstlich zu diskretisieren. Alternativ kann man versuchen, die Wahrscheinlichkeiten durch eine Standardverteilung anzunähern, beispielsweise durch die Normalverteilung [Jankowski u. a. 2006]. Als wesentliche Vorteile des NBK sind dessen Einfachheit und geringe Komplexität zu nennen. Sollten zwischen den Eingangskenngrößen starke Abhängigkeiten auftreten, existieren weitere Abwandlungen, um diese teilweise zu berücksichtigen. Diese höhere Leistungsfähigkeit zur automatischen Kompensation von Abhängigkeiten durch einen komplexeren Aufbau des Modells bedingt jedoch in der Regel wesentlich schlechtere Laufzeiteigenschaften [Hand u. a. 2001].

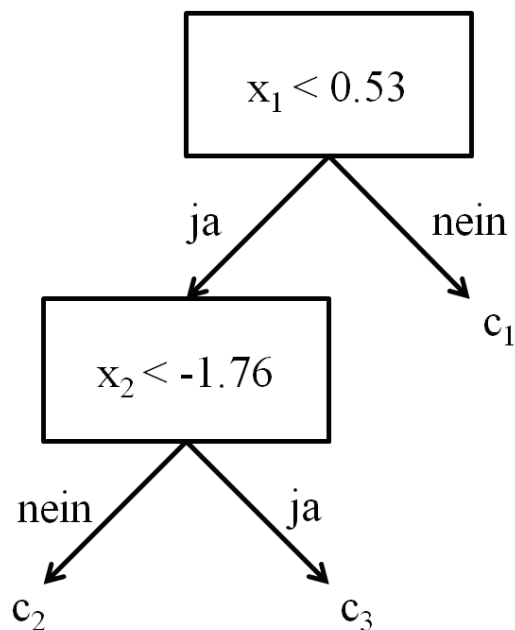


Abbildung 3.3: Exemplarischer Aufbau eines Entscheidungsbaums für die Eingangsgrößen x_1 und x_2 sowie den Klassen c_1 , c_2 und c_3

Logit Boost (LB)

Das Logit Boost Verfahren kommt aus der Gruppe der Boosting Verfahren. Dies sind Algorithmen, welche aus einer Menge aus m einfachen, gewichtet Basisklassifikatoren bestehen, die meist sequenziell zu einem mächtigeren Klassifikator fusioniert werden. Vom Prinzip her ähnelt die Klassifikation dem Naive Bayes Klassifikator, da auch beim LB anhand der A-Posteriori-Wahrscheinlichkeit der Eingangswerte $x \in X$ die wahrscheinlichste Klasse $c_i \in Y$ zugeordnet wird. Jedoch werden beim LB die gesuchten Wahrscheinlichkeiten nicht auf Basis des Bayestheorem sondern anhand von Basisklassifikatoren ermittelt. Diese Basisklassifikatoren sind meist sehr einfach aufgebaut und berücksichtigen beispielsweise nur ein einziges Merkmal, was dazu führt, dass sie schnell ausgewertet werden können.

Typische Basisklassifikatoren können beispielsweise einfache Entscheidungsbäume sein. Jeder Entscheidungsbaum besteht aus sogenannten Knoten und Blättern, wobei die Knoten eine Entscheidung über eine Kenngröße definieren, während die Blätter die zugeordnete Klasse darstellen (Abbildung 3.3). Indem man sequenziell entlang des Baumes durch die Knoten geht und deren Regeln abprüft, können so die einzelnen Kenngrößen klassifiziert werden. Abhängig von der Prüfung in den einzelnen Knoten wird der nächste Knoten ermittelt, bis man schließlich ein Blatt erreicht, welches das Ergebnis der Basisklassifikation darstellt. Entscheidungsbäume in Boosting Verfahren werden in der Regel automatisch anhand der Trainingsdaten erzeugt [Guyon u. a. 2009, Cai u. a. 2006, Jankowski u. a. 2006].

Mit Hilfe der einzelnen Basisklassifikatoren kann nun ein umfassendes Modell zur Gesamtklassifikation erstellt werden. Hierbei sei $b : X \rightarrow \{0, 1\}$ das Klassifikationsmodell und $\vartheta_1 \dots \vartheta_m$ frei konfigurierbare einstellbare Parameter des Modells. $b(x; \vartheta_i)$ ist der Wert, welcher von dem mit ϑ_i parametrisierten Modell für die Eingangsgröße $x \in X$ prognostiziert wird. Aus diesen Einzelwerten bildet der Algorithmus n Funktionen, welche die Zugehörigkeit F_i zur jeweiligen Klasse c_i repräsentieren.

$$\begin{aligned} F_1(x) &= \sum_{j=1}^m v_{1,j} b(x, \vartheta_j) \\ &\vdots \\ F_n(x) &= \sum_{j=1}^m v_{n,j} b(x, \vartheta_j) \end{aligned}$$

Je größer $F_i(x)$ ist, desto höher ist also die Wahrscheinlichkeit, dass x zur Klasse c_i gehört. Aus diesen Kennwerten kann mit dem Modell die gesuchte A-Posteriori-Wahrscheinlichkeit p_i berechnet werden. Diese gibt an, wie wahrscheinlich ein Eingangswert $x \in X$ zur Klasse c_i gehört und wird wie folgt berechnet [Li 2010, Friedman u. a. 2000]:

$$p_i(x) = \frac{e^{F_i(x)}}{\sum_{k=1}^n e^{F_k(x)}}$$

Ein auf Entscheidungsbäumen basierendes Logit Boost Verfahren gilt als ein sehr leistungsfähiges Klassifikationsverfahren, welches in eingesetzten Anwendungen oft gute Ergebnisse liefert und auch bei Vergleichstest häufig gut abschneidet [Lutz 2006, Dettling u. a. 2003, Schapire u. a. 1999].

Künstliche Neuronale Netze (KNN)

Einer der bekanntesten Ansätze in Wissenschaft und Technik zur datenbasierten Modellierung sind die künstlichen Neuronalen Netze. So basieren beispielsweise im Bereich der Prozessüberwachung mehr als 50% der “intelligenten Maschinenmodelle” auf KNNs [Abellan-Nebot u. a. 2010].

Der Begriff der KNN steht für eine große Menge an unterschiedlichen konnektionistischen Modellen. Diese Modelle haben in der Regel gemein, dass die Modellierung auf einer großen Anzahl an einfach aufgebauten Elementen (den sogenannten Neuronen) basiert. Diese Neuronen werden durch unterschiedlichste Topologien miteinander verbunden. Hierbei dient das Gehirn als Vorbild, in dem die Synapsen ebenfalls miteinander vernetzt sind, wenn auch ungleich komplexer [Knieling 2007].

Die wesentliche Eigenschaft von KNNs ist ihre Fähigkeit, nichtlineare Zusammenhänge anhand vorliegender Daten zu ermitteln, indem die Modellparameter meist iterativ an die

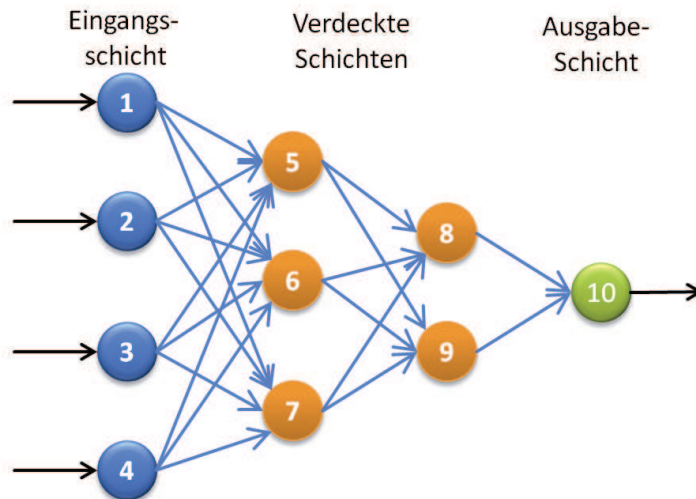


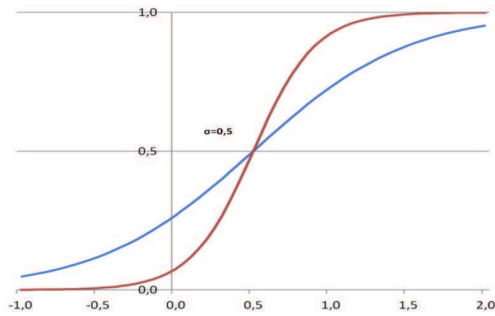
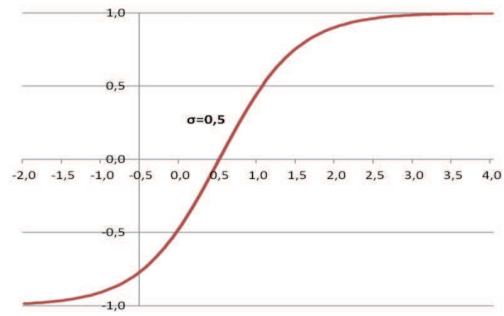
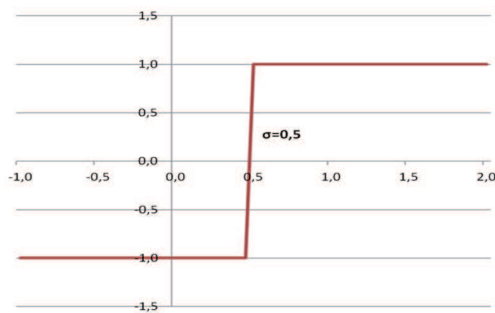
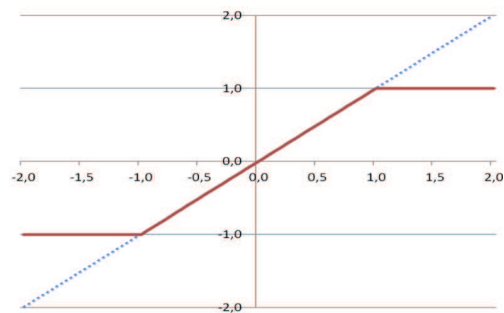
Abbildung 3.4: Exemplarischer Aufbau eines MLP-Netztes mit zwei verdeckten Schichten.

Trainingsdaten angepasst werden. KNNs weisen hierbei oft vergleichsweise gute Abstraktionsfähigkeiten auf, sofern geeignete Abbruchkriterien eine Überanpassung des Modells während des Trainings verhindern. Dadurch können zu einem gewissen Grad auch Eingangsdaten interpretiert werden, welche sich zwischen den Trainingsmustern befinden. Unterscheidungsmerkmale für die Neuronalen Netze sind die Anordnung und Verknüpfung der Neuronen sowie deren Aktivierungsfunktion. Mathematisch ist die Topologie eines Netztes eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, wobei n und m jeweils die Dimension der Eingangsdaten und der Ausgangsdaten darstellt [Bishop 1995].

Einer der gängigsten Verknüpfungsansätze bei den KNN ist das multi-layer perceptron (MLP), welches bereits in den 60er Jahren entstanden ist [Rosenblatt 1958]. Ein MLP-Netz besteht aus einer Eingangsschicht, welche die Dimension der Eingangsdaten repräsentiert. Es folgen eine oder mehrere verdeckte Schichten und eine Ausgabeschicht mit der Dimension der Ausgangsdaten. Der schematische Aufbau eines mehrschichtigen Perzeptrons ist in Abbildung 3.4 zu sehen.

Die wesentlichen Kennzeichen der Neuronen sind deren Eingangswert, eine Aktivierungsfunktion und eine Ausgabefunktion. Die Eingangswerte der ersten Schicht sind in der Regel die normierten Eingangsgrößen. Für die Neuronen der verdeckten Schichten und der Ausgabeschicht sind die Eingangswerte jeweils die Summen aus den gewichteten Ausgabewerten der verknüpften Vorgängerneuronen. Die Funktion zur Berechnung der Netzeingabe net_{pj} wird auch Propagierungsfunktion genannt und ist wie folgt berechnet:

$$net_{pj} = \sum_{i=1}^t out_{pi} w_{i,j}$$

(a) logistische Aktivierungsfunktion mit Schwellwert σ und Steigung T (b) Tangens hyperbolicus mit Schwellwert σ (c) binäre Aktivierungsfunktion mit Schwellwert σ 

(d) Lineare Aktivierungsfunktion mit Sättigung

Abbildung 3.5: Verschiedene Aktivierungsfunktionen für Neuronen der Neuronalen Netze

Hierbei ist t die Anzahl der Vorgängerneuronen, $out_{p,i}$ Ausgabe des Vorgängerneurons zum aktuellen Neuron und $w_{i,j}$ das Verbindungsgewicht zwischen den Neuronen i und j .

Der Ausgabewert der Neuronen wird anhand der Eingangswerte und der Aktivierungsfunktion berechnet, wobei diese Berechnungsvorschrift in der Regel ein wählbarer Konfigurationsparameter ist. Meist besteht die Aktivierungsfunktion aus einem begrenzten Wertebereich in Kombination mit einem Schwellwert σ . Gängige Aktivierungsfunktionen sind in Abbildung 3.5 dargestellt. Eine häufig verwendete Aktivierungsfunktion bei kontinuierlichen Eingangsparametern ist die logistische, welche sehr sensibel auf eine Netzeingabe im Bereich des Schwellwerts reagiert. Sie wird wie folgt berechnet:

$$f(in) = \frac{1}{1 + e^{-(in-\sigma)/T}}$$

T beschreibt hierbei einen Parameter für die Steigung und stellt somit die Sensibilität der Funktion dar.

Als Ausgabefunktion der Neuronen $out_{p,i}$ wird in der Regel die Identität der Aktivierungsfunktion verwendet, aber auch komplexere Berechnungsvorschriften sind je nach Aufgabenstellung möglich. Eine Besonderheit der MLP-Netze im Vergleich zu anderen Model-

lierungsverfahren wie SVM oder LB ist, dass nicht nur die Modellierung diskreter Klassen sondern auch von kontinuierlichen Merkmalen möglich ist. Dadurch sind MLP-Netze auch für Regressionsmodellierungen geeignet. Die Modellerstellung bei KNNs kann als vergleichsweise aufwendig bezeichnet werden, die Prognose des Netzes besteht jedoch aus einfachen Rechenoperationen und kann sehr schnell durchgeführt werden [Zell 1994].

Learning Vector Quantization (LVQ)

LVQ ist ein maschinelles Lernverfahren, welches der großen Klasse der künstlichen Neuronalen Netzen zugeordnet wird. Beim LVQ handelt es sich um ein Klassifikationsverfahren, welches als Modell eine Anzahl an Referenzvektoren oder sogenannte Codebookvektoren berechnet, mit deren Hilfe die möglichen Klassen abgebildet werden. Unbekannte Eingangsdaten werden im Gegensatz zu MLP-Netzen nun anhand eines Abstandsmaßes den möglichen Referenzklassen zugeordnet. Hierbei sei c die Anzahl der Codebookvektoren F_i und n die Anzahl der möglichen Ausgangsklassen, wobei $c \geq n$ gilt und F_i die gleiche Dimension wie die Eingangsdaten $x \in X$ besitzt.

Beim Training des Modells werden die Codebookvektoren zunächst initialisiert. Während des Trainings werden dann die Codebook-Vektoren durch iterative Verschiebung so angeordnet, dass die Dichteverteilung der Eingangsdaten möglichst gut repräsentiert wird. Trainingsdaten und Codebookvektoren der gleichen Klassen wirken dabei “anziehend”, während unterschiedlich Klassen “abstoßend” wirken. Die Größe der Verschiebung wird mathematisch durch einen von den Trainingsdaten abhängigen reellen Faktor α bestimmt. Dieser wird mit den Differenzvektoren multipliziert, welche zu den Codebookvektoren hinzuaddiert oder abgezogen werden.

In der Klassifikation wird für die Eingangswerte der ähnlichste Codebookvektor gesucht, was über ein definiertes Abstandsmaß geschieht. Es handelt sich somit praktisch um eine nächste-Nachbar-Klassifikation. Ein häufig verwendetes Abstandsmaß ist der euklidische Abstand d :

$$d_i(x, F_i) = \sqrt{\sum_{j=1}^m (x_j - F_{ij})^2}$$

Der Ausgabewert des Modells ist die Klasse des ermittelten Codebookvektors mit dem kleinsten Abstand [Kohonen 2001].

Besonders bei nicht optimal getrennten Eingangsdaten kann das Klassifikationsergebnis verbessert werden, indem nicht nur der nächste Nachbar sondern die k-nächsten Codebookvektoren nach entsprechenden Homogenitätsregeln zur Klassifikation verwendet werden.

Auch kann das Verfahren durch ein intelligentes Initialisierungsverfahren weiter verbessert werden.

Das LVQ-Verfahren ist prinzipbedingt anfällig gegenüber ungleich verteilten Trainingsdaten oder überlappenden Clustern innerhalb der Klassen. Dies kann sich dahingehend auswirken, dass Klassen mit vielen Datensätzen die Verschiebung der Codebookvektoren während des Trainings überproportional stark beeinflussen. Jedoch ist das Verfahren schnell und einfach zu berechnen und bietet durch das Hinzufügen neuer Codebookvektoren auch eine einfache Möglichkeit der Modellerweiterung ohne komplettes Neutraining. Weiterhin kann das trainierte Modell durch entsprechende Projektionsmethoden (z.B. Sammons-Mapping) der Trainingsdaten und Codebookvektoren vom Anwender leicht und intuitiv visualisiert und nachvollzogen werden, was bei vielen anderen Modellverfahren nicht der Fall ist.

General Regression Neural-Fuzzy Networks (GRNFN)

GRNFN ist ein interpolierendes Modellierungsverfahren basierend auf KNNs und Fuzzy-Logik. Fuzzy Logik ist eine Theorie, welche zur Modellierung von unscharfen Zusammenhängen verwendet wird und eine immer größere Verbreitung in Wissenschaft und Technik aufweisen. So kommen laut einer Studie Systeme mit Fuzzy-Komponenten in etwa 15% der angewendeten intelligenten Überwachungssysteme zum Einsatz [Abellan-Nebot u. a. 2010]. Fuzzy-Systeme gehen auf die Arbeiten von Zadeh im Jahre 1965 zurück [Zadeh 1965]. Hierbei wird versucht, ein System anhand von Expertenwissen mit unscharfen Wertezuordnungen zu modellieren. Dieses Expertenwissen wird in einer Regelbasis codiert und mit geeigneten Zugehörigkeitsfunktionen, Abbildungsvorschriften und Operatoren verknüpft [Hanss 2005]. Während datenbasierte Modelle automatisch aus den ermittelten Trainingsdaten berechnet werden, basieren Fuzzy-Systeme auf Regeln, welche von Experten erstellt werden müssen. Es liegt nahe, die Vorteile beider Prinzipien zu vereinen, was allgemein als Neuro-Fuzzy-Modellierung bezeichnet wird. Derartigen Modellen liegt häufig ein mehrschichtiges KNN zugrunde, welches mit Fuzzy-Komponenten ergänzt wird. So können beispielsweise die Parameter der Neuronen als Fuzzymenge interpretiert und entsprechende Zugehörigkeitsfunktionen anhand von Trainingsdaten optimiert werden [Mitra u. a. 2000, Park u. a. 2000, Nauck u. a. 1997].

Eine weitere Anwendung sind interpolierende Neuro-Fuzzy Modelle. Interpolierende Modelle haben generell die Eigenschaft, dass das Modelltraining nicht anhand einer zielgrößenbezogenen Optimierung der Netzparameter erfolgt. Stattdessen wird eine direkte Verknüpfung von Eingangs- zu Ausgangsgrößen anhand von Stützpunkten verwendet, welche aus den Trainingsdaten abgeleitet werden. Das Modell wird somit nach einer festen Berechnungsvorschrift in einem Durchgang ermittelt [Hastie u. a. 2009]. Basierend auf dem

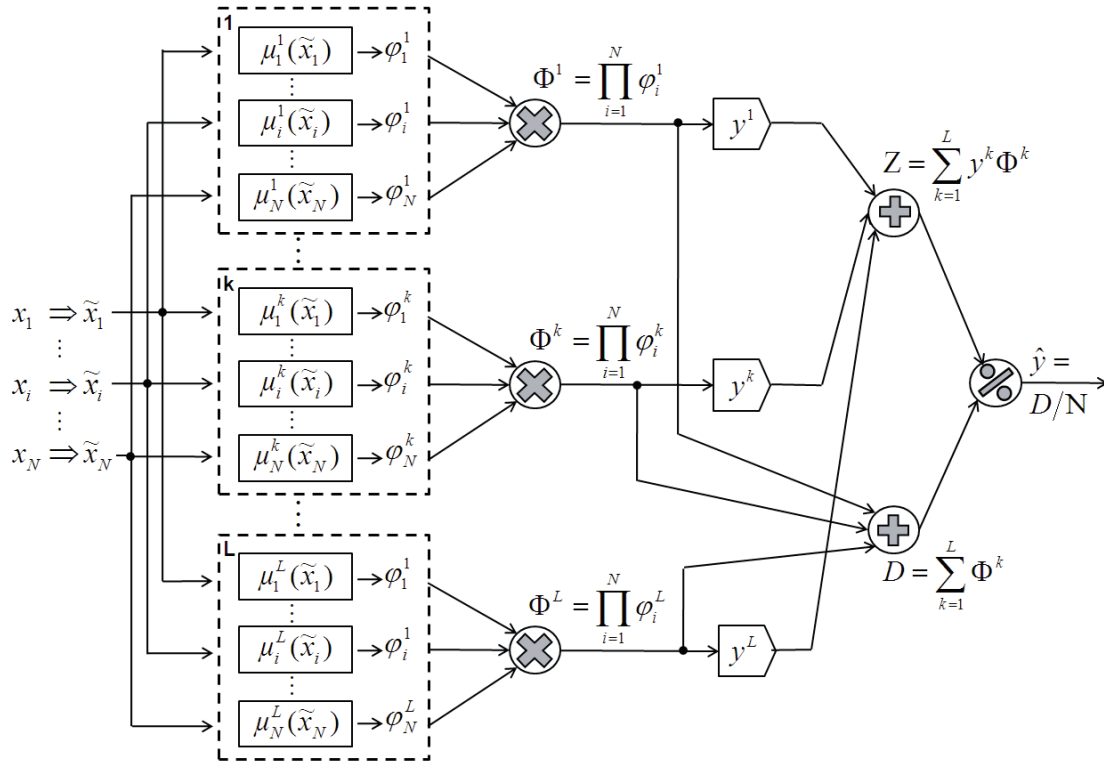


Abbildung 3.6: Prinzipielle Architektur von General Regression Neural-Fuzzy Networks (Quelle: [Bodyanskiy u. a. 2007])

Ansatz der interpolierenden Modellstruktur wurden die GRNFN entwickelt, deren Architektur in Abbildung 3.6 dargestellt ist. Es handelt sich um ein sechsschichtiges Netzmodell mit Fuzzy-Komponenten. In der ersten Schicht werden hierbei die Eingangswerte vor der eigentlichen Verarbeitung normiert und in der zweiten Schicht durch Zugehörigkeitsfunktionen in Fuzzy-Mengen umgewandelt (fuzzyfiziert). Aus den daraus resultierenden Einzelaktivierungswerten φ_i^k werden in Schicht drei die Gesamtaktivierungen Φ^k berechnet. Diese werden in Schicht vier durch entsprechende Gewichtungsfaktoren y^k zu sogenannten Inferenzen zusammengefasst. Abschließend wird in Schicht fünf die Netzausgabe \hat{y} durch Verknüpfung (Akkumulation) der einzelnen Inferenzen berechnet.

Die Berechnungsvorschrift der Netzausgabe ist somit:

$$\hat{y}(x) = \frac{\sum_{k=1}^L y^*(k) \varphi(D_k(x))}{\sum_{k=1}^L \varphi(D_k(x))}$$

wobei y^* der Ausgabewerts des Trainingsstützvektors, L die Anzahl der ermittelten Stützpunkte, D_k eine stützpunktspezifische Distanzfunktion und φ eine entsprechende Aktivierungsfunktion darstellt [Bodyanskiy u. a. 2008, Bodyanskiy u. a. 2007].

Die Vorteile von GRNFN-Modellen liegen in einer vergleichsweise geringen Anzahl an Modellparametern, aufgrund der nur eindimensionalen Aktivierungsfunktionen. Dies führt in der Regel zu einem sehr schnellen und deterministischen Training. Weiterhin kann das Netz bis zu einem gewissen Grad direkt und ohne Neutraining angepasst werden. Bei vorhandenem Prozesswissen können zudem manuelle Fuzzy-Regeln in das Modell integriert werden, wodurch die Modellgüte gesteigert werden kann [Neher 2011].

Weitere Verfahren zur Modellierung

Neben den genannten Verfahren gibt es eine Fülle weiterer Beschreibungsformen und Modellierungsmethoden (beispielsweise einfache Regressionsmethoden oder Hidden-Markov-Modelle), welche jedoch aufgrund von Randbedingungen und Zielen für die hier gestellte Aufgabe unpassend erscheinen und deshalb nicht näher untersucht werden. Für einen erweiterten Überblick über weitere Modellierungsmethoden wird deshalb auf entsprechende Literatur verwiesen, z.B. [Hartung u. a. 2009, Hastie u. a. 2009, Elliott u. a. 2008, Bishop 2006, Duda u. a. 2000, Berthold u. a. 1999, Mitchell 1997].

3.3 Problemspezifische Ableitung aus dem Stand der Technik

Spiegelt man die Anforderungen aus Abschnitt 2.3 mit dem Stand der Technik, so wird deutlich, dass zur Entwicklung eines robusten Verfahrens zur Bestimmung der Zahnfarbe zwei vielversprechende Möglichkeiten existieren: Einerseits die Nutzung genauerer Sensoren oder verbesserter Messkomponenten sowie andererseits die Entwicklung intelligenter Auswerteverfahren für die bestehenden handlichen Sensoren.

Es existieren zwar bereits handliche Smart-Sensoren (z.B. VITA Easyshade), welche zudem speziell auf die Bestimmung der Zahnfarbe zugeschnitten sind. Derartige Sensoren haben jedoch Defizite bei der Farbbestimmung aus den verfügbaren Sensordaten, weshalb sie von vielen anwendenden Zahnärzten noch kritisch gesehen werden. Weiterhin gibt es andere Sensorkonzepte für hochgenaue Farbmessungen (Farbkameras mit Spektrometern sowie stationäre oder berührungslose Laborgeräte), die jedoch nicht für intraorale Messungen an Zähnen taugen. Prinzipiell geeignete, hochgenaue Farbsensoren zur Bestimmung der Zahnfarbe mit mobilen Messsonden existieren zwar, jedoch sind diese deutlich teurer als bisher eingesetzte Sensoren. Deshalb steckt in der Verbesserung der Auswertung ein hohes Verbesserungspotenzial, beispielsweise durch Methoden des maschinellen Lernens.

Im Folgenden werden die Anforderungen an das Verfahren mit den beiden möglichen Lösungen abgeglichen und bewertet. Es werden Punkte zwischen “0” und “3” vergeben, wobei “0” für ungeeignet und “3” für sehr gut geeignet steht.

Eine der wichtigsten Anforderung an die Lösung ist, dass unter sämtlichen realen Bedingungen die richtige Farbe so gut wie möglich ermittelt wird. Die Ergebnisse müssen also plausibel, robust und reproduzierbar sein. Da die plausible Ermittlung der Zahnfarbe stark von der Datenqualität abhängt, ist hier ein Sensor mit genaueren Komponenten im Vorteil. Jedoch bieten intelligente Klassifikationsverfahren auch bei schlechteren Sensordaten das Potenzial, zufriedenstellende Ergebnisse hinsichtlich der Plausibilität erreichen zu können.

Bezüglich der Robustheit weisen beide Lösungsmöglichkeiten ähnliche Schwächen auf. Da die Probe nicht abgeschlossen in einer Laborumgebung gemessen werden kann, besteht bei beiden Ansätzen die Gefahr von negativen Umgebungseinflüssen, unebenen Zahnoberflächen sowie Bedienerfehlern wie eine Verkippung des Messkopfes.

Die Reproduzierbarkeit der Ergebnisse hinsichtlich der Auswertelgorithmen kann bei beiden Lösungen als identisch angesehen werden, da bei beiden Verfahren bei gleichen Eingangsdaten jeweils ihre gleichen Ergebnisse liefern. Da jedoch die Eingangsdaten bei den besseren Sensorkomponenten eine größere Wiederholgenauigkeit bei den Messung aufweisen, sind diese hier im Vorteil. Bei den handlichen Smart-Sensoren kann jedoch, je nach Leistungsfähigkeit der Auswertemethode, der beschriebene Nachteil verringert werden.

Anforderungen zur Klassifikationsqualität	Hochgenauer Sensor mit verbesserten Messkomponenten	Intelligente Auswertung für handliche Smart-Sensoren
Plausible Ermittlung der Zahnfarbe	3	2
Robuste Ermittlung der Zahnfarbe	1	1
Reproduzierbare Ermittlung der Zahnfarbe	2	1

Tabelle 3.1: Bewertung geeigneter Lösungen zur Entwicklung eines robusten Verfahrens zur Bestimmung der Zahnfarbe hinsichtlich der Ergebnisqualität

Eine weitere Anforderung ist die uneingeschränkte Einsetzbarkeit in Praxisräumen bei Zahnärzten. Dies beinhaltet zum Einen, dass das Verfahren möglichst mobil ist, also am besten direkt zum Patienten gebracht werden kann. Denkbar wäre auch ein stationäres Gerät in einem separaten Raum, ähnlich einem Röntgengerät. Ein hochgenauer Sensor ist in der Regel weniger mobil als ein einfacher handlicher Sensor, da er meist fest montiert oder aufgrund sperriger Komponenten z.B. auf einem rollenden Tisch transportiert werden

muss. Jedoch wären hochgenauen Laborgeräte teilweise noch prinzipiell tauglich für den Einsatz in Zahnarztpraxen. Die bestehenden, handlichen Smart-Sensor liefern alle Vorteile bezüglich Mobilität, da sie ohne Aufwand zum Patienten gebracht werden können.

Hinsichtlich der Sterilisierbarkeit können beide Sensorarten die Anforderungen erfüllen. Zwar wird mit einer Messsonde berührend und intraoral gemessen, jedoch kann die Messsonde bei beiden Lösungsansätzen einfach mit Schutzfolien entsprechend präpariert oder durch entsprechende Mittel desinfiziert werden.

Auch die Konformität mit dem Medizinproduktegesetz ist eine zwingend zu erfüllende Anforderung. Zwar ist die entsprechende Zertifizierung nicht Teil der Arbeit, jedoch darf keines der Verfahren potenzielle Risiken aufweisen, das Medizinproduktegesetz zu verletzen. Beide in Frage kommenden Lösungen weisen hierbei keine erkennbaren Schwierigkeiten auf, weshalb auch diese Anforderung als erfüllt angesehen werden kann.

Anforderungen zur Einsetzbarkeit in Zahnarztpraxen	Hochgenauer Sensor mit verbesserten Messkomponenten	Intelligente Auswertung für handliche Smart-Sensoren
Möglichst mobiles Verfahren	1	3
Sterilisierbarkeit	3	3
Einklang mit dem Medizinproduktegesetz	3	3

Tabelle 3.2: Bewertung geeigneter Lösungen zur Entwicklung eines robusten Verfahrens zur Bestimmung der Zahnfarbe hinsichtlich der Eignung im Praxisbetrieb

Neben der Einsetzbarkeit in der realen Einsatzumgebung in Zahnarztpraxen ist auch die Akzeptanz der Geräte durch die Nutzer sowie die Wirtschaftlichkeit eine wichtige Anforderung. Deshalb muss die gewählte Lösung über ein akzeptables Preis-/ Leistungsverhältnis verfügen, einfach und intuitiv bedienbar sein sowie das Ergebnis schnell und nachvollziehbar visualisieren. Hinsichtlich der kostengünstigen Realisierung liegen Lösungen auf Basis von hochgenauen Sensoren preislich oft im Bereich größer 10.000 Euro, während handliche Smart-Sensoren in der Größenordnung von 2000 Euro liegen. Da die intelligente Auswertesoftware keine zusätzlichen Materialkosten verursacht und eventuelle Entwicklungskosten auf eine große Anzahl von Geräten umgelegt werden können, ist dieser Lösungsansatz hier eindeutig im Vorteil.

Bezüglich der einfachen und intuitiven Bedienung verfügen beide Verfahren im Prinzip über die Möglichkeit, per Knopfdruck ein Messergebnis zu produzieren. Auch benötigte Kalibriervorgänge können weitestgehend automatisiert durchgeführt werden. Selbst kom-

plexe Auswertelgorithmen können völlig unabhängig vom Benutzer für die Auswertung genutzt werden. Lediglich Fehlbedienungen lassen sich bei beiden Ansätzen nicht komplett vermeiden.

Hinsichtlich der Berechnungsdauer sind ebenfalls keine Probleme zu erwarten. Hierbei ist anzumerken, dass ein Ergebnis lediglich innerhalb weniger Sekunden vorliegen muss und kleinere Verzögerungen akzeptabel sind. Bei hochgenauen Sensoren sind meist größere Auswerteeinheiten integriert, welche die nötigen Berechnungsschritte ohne Probleme in der benötigten Zeit durchführen können. Auch handliche Smart-Sensoren mit nur einer integrierten, leistungsschwächeren Auswerteeinheit können einfache mathematische Operationen in der benötigten Zeit durchführen. Komplexe Trainingsvorgänge von Verfahren des maschinellen Lernens lassen sich jedoch meist nicht auf einem Mikroprozessor realisieren. Das Modelltraining kann jedoch problemlos auf einem externen Computer durchgeführt werden. Somit muss nur das fertige Modell für die Farbklassifikation in den Smart-Sensor integriert werden. Somit kann diese Anforderung mit beiden Lösungsansätzen erfüllt werden.

Anforderungen zur Akzeptanz potenzieller Nutzer	Hochgenauer Sensor mit verbesserten Messkomponenten	Intelligente Auswertung für handliche Smart-Sensoren
Kostengünstige Realisierung	0	3
Einfache und intuitive Bedienung	2	2
Kurze Berechnungsdauer	3	2,5

Tabelle 3.3: Bewertung geeigneter Lösungen zur Entwicklung eines robusten Verfahrens zur Bestimmung der Zahnfarbe hinsichtlich Akzeptanz der Nutzer

Neben einer funktionierenden Lösung für die Ermittlung der Zahnfarbe, muss das System auch flexibel und erweiterbar sein. Das bedeutet, dass neue Farben in die verwendeten Farbräume integrierbar sein müssen. Weiterhin muss es auch möglich sein, die Klassifikation auf komplett neue Farbräume anzupassen.

Das Hinzufügen von komplett neuen Farbräumen anhand der Sensordaten ist bei beiden Lösungsansätzen mit größerem Aufwand verbunden, aber jeweils teilautomatisiert machbar. Hierbei muss lediglich darauf geachtet werden, dass neue Farben auch im Wellenlängenbereich der Sensoren liegen. Da die meisten Sensoren jedoch den gesamten sichtbaren Spektralbereich abdecken, sind hier keine Probleme zu erwarten.

Die Verwendung neuer Farben bei bestehenden Farbräumen ist ebenfalls bei beiden Lösungen möglich. Der Aufwand hierfür hängt teilweise von den verwendeten Algorithmen

ab. Intelligente Klassifikatoren benötigen in der Regel nur eine entsprechende Anzahl an Trainingsdaten, um das Modell durch zusätzliche Farben zu erweitern. Selbst wenn ein komplettes Neutraining nötig ist, kann dies in der Regel ohne Expertenwissen automatisiert durchgeführt werden.

Anforderungen zur Erweiterbarkeit	Hochgenauer Sensor mit verbesserten Messkomponenten	Intelligente Auswertung für handliche Smart-Sensoren
Funktion für sämtliche Farbräume und Farbsysteme	2	2
Erweiterbar für neue Farben	3	3

Tabelle 3.4: Bewertung geeigneter Lösungen zur Entwicklung eines robusten Verfahrens zur Bestimmung der Zahnfarbe hinsichtlich Erweiterbarkeit des Verfahrens

Gesamtbewertung der Lösungsansätze zur Erfüllung der Anforderungen

Zur Ermittlung der Gesamtpunktzahl werden die Punkte für die Erfüllung der notwendigen Anforderungen doppelt gewichtet, die Punkte für die optionalen Anforderungen einfach. Daraus ergibt sich folgende Gesamtbewertung:

Bewertung der Anforderungen	Hochgenauer Sensor mit verbesserten Messkomponenten	Intelligente Auswertung für handliche Smart-Sensoren
Gesamtpunktzahl	36	38,5

Tabelle 3.5: Gesamtbewertung der geeigneten Lösungen zur Entwicklung eines Verfahrens zur Bestimmung der Zahnfarbe

Speziell der hohe Preis und die schlechtere Mobilität geben den Ausschlag für einen handlichen Smart-Sensor mit intelligentem Auswerteverfahren. Der Nachteil bei dieser Lösung ist die schlechtere Qualität der Sensordaten. Hierfür müssen für die Umsetzung entsprechende Strategien entwickelt werden, um diese prinzipbedingte Schwäche bei der Datenaufnahme möglichst gut durch Auswertemethoden kompensieren zu können.

Als Beispielapplikation wird in dieser Arbeit der Smart-Sensor VITA Easyshade verwendet, welcher bereits über alle benötigte Komponenten zur Erfüllung der Entwicklungsanforderungen verfügt und auch bezüglich der Kosten Akzeptanz bei potenziellen Nutzern genießt. Zudem wurden alle benötigten Versuchs- und Validierungsdaten, welche für die nachfolgenden behandelten Entwicklungen genutzt werden, von der Firma Vita mit dem Easyshade Smart-Sensor ermittelt und aufgezeichnet.

Kapitel 4

Detaillierter Lösungsansatz

Aufgrund der in Abschnitt 3.3 durchgeführten Anforderungsanalyse wurde als geeignete Lösung ein handlicher Spektroskop mit intelligenter Auswertelogik auf Basis maschinellen Lernens ermittelt. Das Hauptdefizit für eine robuste Ermittlung der Zahnfarbe liegt in der Qualität der Messdaten, um hieraus mit den bisherigen Methoden die exakte Zahnfarbe für einen definierten Farbraum bestimmen zu können. In diesem Kapitel sollen daher basierend auf realen Messdaten Strategien und Methoden entwickelt werden, welche eine möglichst optimale Klassifikation der gesuchten Zahnfarbe ermöglichen. Dabei liegen die Schwerpunkte in der Entwicklung geeigneter Lösungen in den Teilgebieten Vorverarbeitung der Daten, Klassifikation der Daten sowie Plausibilitätsprüfung der Ergebnisse. Weiterhin wird die Möglichkeit der Integration der erarbeiteten Methoden in den ausgewählten Smart-Sensor VITA Easyshade betrachtet, welcher bereits über alle relevanten Hardwarevoraussetzungen zur Erfüllung der Systemanforderungen verfügt. Zur besseren Beurteilung der Methoden werden zudem Strategien zur Projektion hochdimensionaler Vektoren vorgestellt, um die Ergebnisse visualisieren zu können.

4.1 Visualisierung hochdimensionaler Vektoren

Hochdimensionale Vektoren sind schwer direkt zu visualisieren oder intuitiv zu verstehen. Deshalb ist für Analysen eine Projektion in den zweidimensionalen Raum in vielen Fällen sinnvoll. Die Abstände der Vektoren bezüglich einer definierten Metrik sind eine Möglichkeit um deren Ähnlichkeit charakterisieren zu können. Hierbei gilt generell, dass ähnliche Vektoren geringere Abstände aufweisen, unähnlichere Vektoren größere. Eine gängige Metrik ist hierbei der euklidische Abstand d welcher für zwei n -dimensionale Vektoren x und y wie folgt definiert ist:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

Projektionen ermöglichen es nun, Ähnlichkeiten im hochdimensionalen Raum auch im zweidimensionalen Raum sichtbar zu machen. Hierbei wird eine Abbildungsvorschrift benötigt, welche die Abstände der Vektoren verhältnismäßig in den zweidimensionalen Raum projiziert. Da der zweidimensionale Raum intuitiver verständlich ist, lassen sich somit Abhängigkeiten und Zusammenhänge leichter darstellen.

Eine nichtlineare Methode zur Projektion von hochdimensionalen Vektoren ist das Sammons-Mapping [Sammon 1969]. Diese iterative Projektionsmethode ist so angelegt, dass Abstandsverhältnisse zwischen zwei Vektoren möglichst gut im zweidimensionalen Raum beibehalten werden. Dadurch liegen die Projektionen von ähnlichen Vektoren auch im zweidimensionalen Raum nahe beieinander. Häufungen von ähnlichen Vektoren bilden dadurch auch in der Projektion Cluster, was dazu führt, dass die Verteilung der Daten visuell interpretiert werden kann. Durch eine Farbcodierung können zudem Zusammenhänge hinsichtlich einer entsprechenden Zielklasse (z.B. der Zahnfarbe) dargestellt werden. Das Prinzip des Sammons-Mapping ist an einem einfachen Beispiel in Abbildung 4.1 dargestellt.

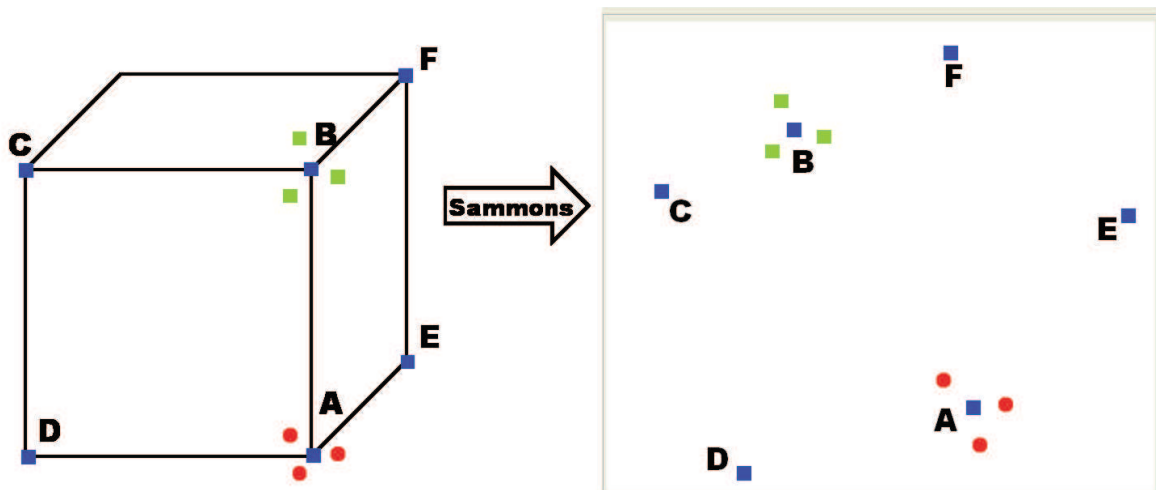


Abbildung 4.1: Projektion für sechs Eckpunkte eines Würfels mit Zusatzpunkten nahe der Ecken

Berechnung der Projektionsvorschrift

Die Berechnung der Projektion erfolgt iterativ für eine Menge an Datenvektoren durch folgende Zielfunktion:

$$E_s = \frac{1}{\sum_i \sum_{j>i} d_{ij}} \sum_i \sum_{j>i} \frac{(d_{ij} - \|r_i - r_j\|)^2}{d_{ij}}$$

Hierbei ist d_{ij} der Abstand von Vektor i zu Vektor j im N -dimensionalen Ursprungsraum bezüglich der definierten Abstandsmetrik. Weiterhin entsprechen r_i und r_j den zugeordneten Projektionen in der zweidimensionalen Bildebene, deren Abstand ebenfalls nach einer zu definierenden Metrik berechnet wird. Während der Berechnung der Projektion wird die Differenz der paarweisen Abstände im hochdimensionalen Vektorraum und im projizierten zweidimensionalen Raum quadriert und für alle Vektoren aufsummiert. Durch die iterative Minimierung des Gütemaßes E_s (beispielsweise durch ein Gradientenverfahren) kann nun eine Abbildung berechnet werden, welche die Ähnlichkeitsverteilung der Vektoren möglichst gut abbildet. Ein weiteres Merkmal der Berechnungsvorschrift ist die Normierung durch die Abstandsdifferenzen, wodurch kleinere Abstände stärker gewichtet werden. Dies führt dazu, dass sehr große Abstände bei den Vektoren die Projektion nicht überproportional stark dominieren können [Sammon 1969].

Eigenschaften von Sammons-Mapping

Die iterative Berechnung der Projektion bedingt, dass die genaue Lage der Punkte im zweidimensionalen Bildraum von der Startinitialisierung abhängt und keine exakte mathematisch darstellbare Abbildungsvorschrift existiert. Neue Daten können also nicht einfach zu bestehenden Projektionen hinzugefügt werden sondern bedingen eine komplette Neuberechnung der Projektion. Zudem können in der Projektion gewisse Verzerrungen auftreten, welche in der Regel von der Größe der Dimensionsreduktion abhängen. Trotz dieser Ungenauigkeiten ist die grobe Aufteilung der Daten in entsprechende Cluster in der Regel gut erkennbar, so dass die Struktur der Daten durch die Projektion hinsichtlich einer Klassenseparierbarkeit ausreichend abgeschätzt werden kann.

4.2 Bestimmung der Zahnfarbe durch Klassifikation von Farbspektren

Aufgrund der physikalischen Eigenschaften der spektralen Daten ist eine klare Zuordnung zu entsprechenden Referenzfarben nicht intuitiv machbar. Deshalb wird für die Klassifikation der Spektren eine automatische Methode auf Basis des maschinellen Lernens benötigt. Zur Auslegung und Realisierung geeigneter Klassifikationsstrategien wird zunächst eine generelle Analyse der spektralen Messdaten hinsichtlich ihrer Datenstruktur, potenzieller Fehler sowie möglicher aussagekräftiger Kenngrößen durchgeführt.

4.2.1 Analyse der spektralen Messdaten

Für die Auswertung liefert der Sensor VITA Easyshade elf Intensitätswerte aus dem sichtbaren Wellenlängenspektrum von 400 bis 700nm. Diese Messwerte stellen die Eingangsda-

ten und somit die Basis für ein intelligentes Auswerteverfahren dar. Das bedeutet, dass sich Fehlerquellen in den Eingangsdaten unmittelbar auf das Klassifikationsergebnis niederschlagen. Deshalb muss zuerst das Verhalten der Daten hinsichtlich Streuungen, Datenfehler, Messgenauigkeit, Unstetigkeiten sowie Zuordnungsmöglichkeiten zu den Farben im Farbraum untersucht und bewertet werden. Für eine intelligente Klassifikation muss außerdem geprüft werden, ob die gemessenen Werte vorverarbeitet werden müssen, das gesamte Spektrum zur Klassifikation genutzt werden kann, und ob durch Methoden der Signalverarbeitung aussagekräftige Kenngrößen gewonnen werden können. In einem letzten Schritt muss untersucht werden, ob potenzielle Trainingsdaten in ausreichender Menge und Qualität ermittelt werden können.

Abschätzung der Qualität der Messdaten und potenzieller Fehlerquellen

Zur Abschätzung der Qualität der Messdaten ist die Wiederholgenauigkeit der Messung ein essentieller Faktor. Hierbei müssen drei potenzielle Fehlerquellen untersucht werden. Dies sind einerseits Fehler, welche direkt durch die Funktionsweise des Sensors bedingt sind, Fehler, welche durch die Bedienung des Nutzers ausgelöst werden, sowie Fehler, die durch spezifische Beschaffenheit menschlicher Zähne bedingt sind.

Zur Abschätzung von Fehlern aufgrund der Sensorkomponenten wurde der Sensor fest verankert und unter definierter Umgebungsbeleuchtung verwendet. Dadurch reduzieren sich Bedienfehler wie Verkippungen oder äußere Einflüsse beim Messvorgang auf ein Minimum. Mit diesem Aufbau wurden Messungen an einer definierten Stelle eines künstlichen Zahns der Farbe 1M1 durchgeführt, wodurch potenzielle Fehlerquellen aufgrund unterschiedlicher Messpunkte auf der Zahnoberfläche ebenfalls minimiert werden. Die Ergebnisse sind beispielhaft in Abbildung 4.2 dargestellt. Die maximale Abweichung bei den einzelnen Messwerten liegt bei ca. 1,5%. Somit kann dem Sensor funktionell eine sehr gute Wiederholgenauigkeit bescheinigt werden. Fehler aufgrund der Messkomponenten haben somit keinen großen Einfluss auf das Ergebnis und können nahezu vernachlässigt werden.

Zur Abschätzung von Fehlern, welche durch den Bediener verursacht werden können, wurde der Sensor mit dem fest verankertem Aufbau bei Messungen bis zu 45° in unterschiedliche Richtungen verkippt. Hieraus resultierten sehr große Abweichungen von bis zu 60% (Abbildung 4.3). Derartig große Verkippungen sind zwar in der praktischen Anwendung unwahrscheinlich, es zeigt jedoch die Anfälligkeit der Messwerte aufgrund von Fehlbedienungen bei realen Messungen. Zur weiteren Abschätzung wurde ein farblich eindeutiger Zahn, welcher von mehreren Experten mit der Farbe 1M1 klassifiziert wurde, vom selben Bediener drei Mal gemessen. Hierbei traten Abweichungen von bis zu 10% auf (Abbildung 4.4). Aufgrund der farblich homogenen Oberflächenstruktur des Zahns, liegt die Vermutung nahe, dass

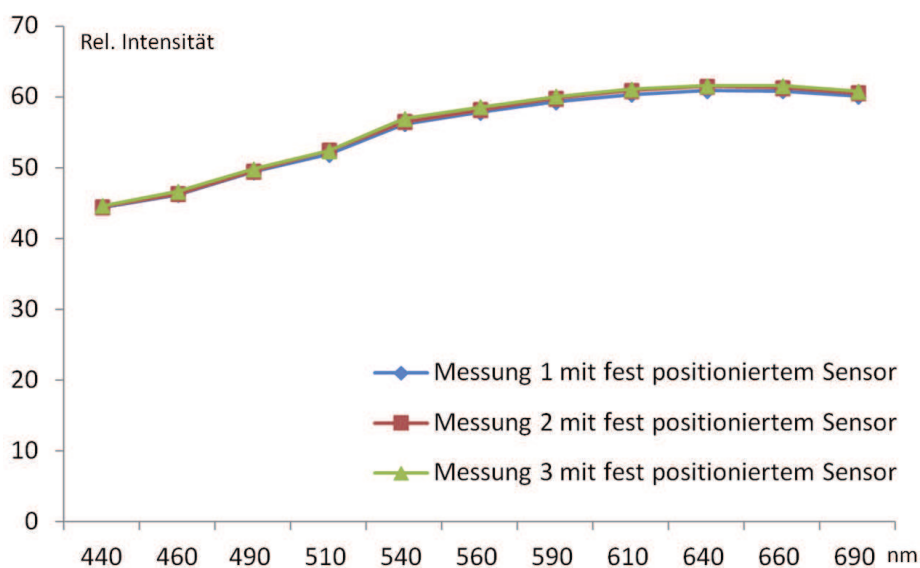


Abbildung 4.2: Bei Messungen mit fest positioniertem Sensor an künstlichen Zähnen ergibt sich eine sehr gute Wiederholgenauigkeit (maximale Abweichung 1,5%)

Verkippungen oder Unebenheiten in der Zahnoberfläche diese Abweichungen verursachen. Die Verkippung des Sensors bedingt durch den Bediener des Sensors oder eine inhomogene Zahnoberfläche beeinflussen das Ergebnis folglich stark.

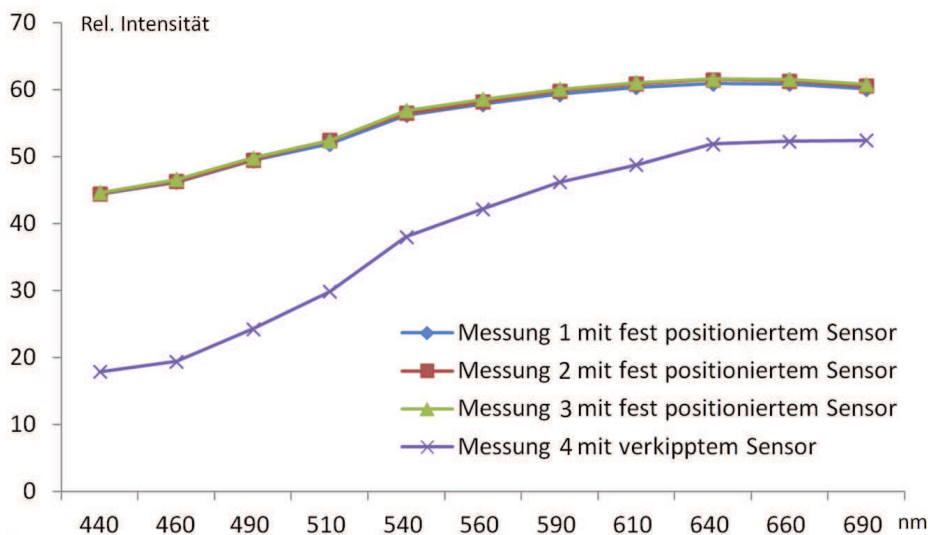


Abbildung 4.3: Bei Messungen mit einem stark verkippten Sensors ergeben sich Abweichungen von bis zu 60%.

Zur Abschätzung von Fehlern aufgrund zahnspezifischer Eigenschaften (beispielsweise eine inhomogene Farbverteilung auf der Zahnoberfläche) wurden am gleichen Zahn an zwei unterschiedlichen Stellen je drei Messungen durchgeführt. Es wurde darauf geachtet, dass

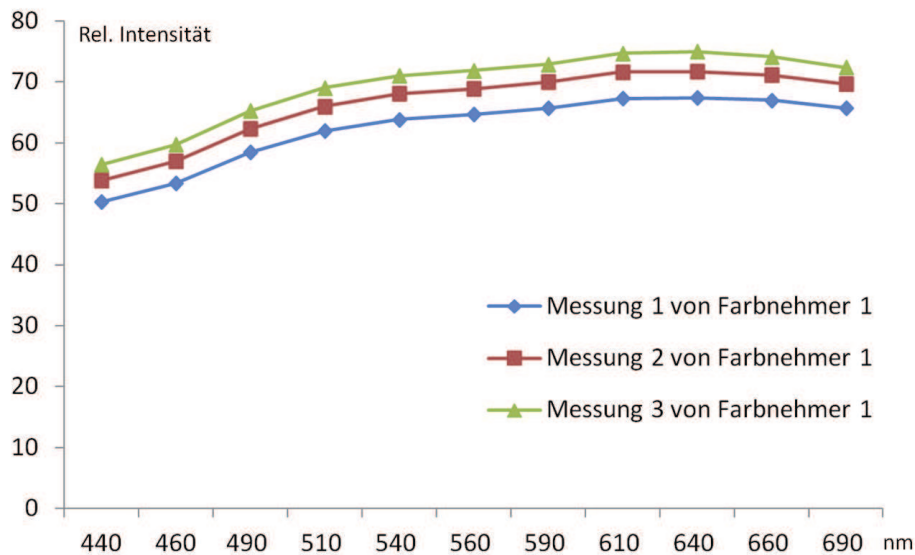


Abbildung 4.4: Bei Messungen am gleichen, realen Zahn ergeben sich Abweichungen von bis zu 10%.

der Sensor während der Messung möglichst nicht verkippt. Während die Schwankungen innerhalb der Messungen an der gleichen Stelle bei unter 5% liegen, treten zwischen den Messstellen Schwankungen von bis zu 12% auf (Abbildung 4.5). Hierbei muss festgestellt werden, dass nahezu jeder Zahn über Inhomogenitäten in der Verfärbung verfügt, welche je nach Person unterschiedlich stark ausgeprägt sind. Aus diesem Grund kann auch dieser Effekt als starker Einflußfaktor auf das Messergebnis und als mögliche Fehlerquelle angesehen werden.

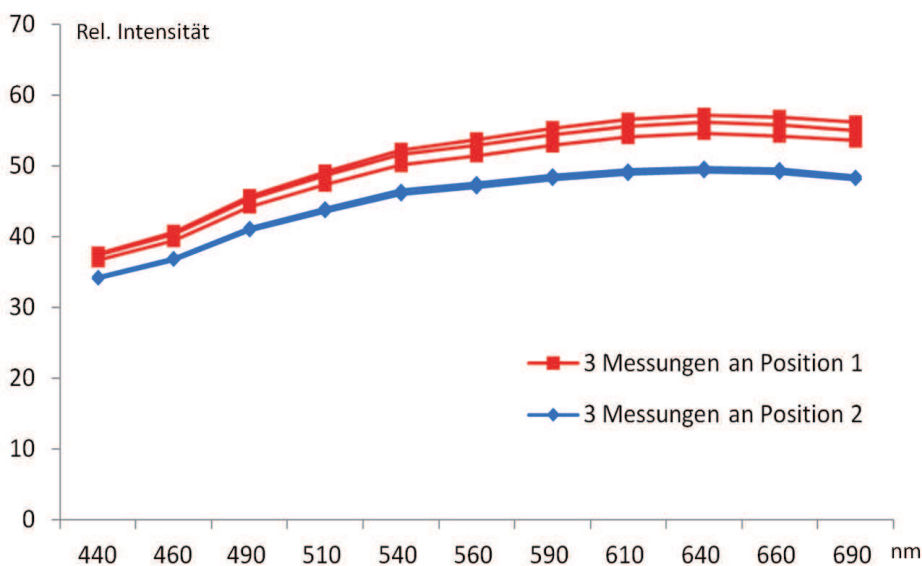


Abbildung 4.5: Bei Messungen an unterschiedlichen Stellen des realen Zahns ergeben sich Abweichungen von bis zu 12%.

Zusammenfassend kann festgestellt werden, dass der Sensor bei optimaler Bedienung und entsprechenden Randbedingungen eine sehr große Wiederholgenauigkeit aufweist. Jedoch sind Verkippungen sowie Inhomogenitäten des Zahns potenzielle Fehlerquellen, deren Einfluß soweit wie möglich bei der Entwicklung der Auswertalgorithmen berücksichtigt werden muss.

Abschätzung bezüglich relevanter Spektralbereiche, Strategien zur Vorverarbeitung und Feature Selection

Nach Untersuchung und Ursachenanalyse von Streuung und potenziellen Fehlerquellen bei Messungen werden im nächsten Schritt die Signalcharakteristika der Messdaten analysiert.

Betrachtet man die Signale hinsichtlich möglicher Vorverarbeitungsstrategien so wird deutlich, dass kaum Störungen anhand von Messrauschen oder Ausreißern an einzelnen Spektralbereichen vorliegen. Dies wird bereits bei der Datenaufnahme akzeptabel kompensiert, weshalb auf aufwändigere Filterverfahren verzichtet werden kann. Relevante Spektralbereiche oder eine besondere Aussagekraft einzelner Spektralbereiche, welche für die Klassifikation stärker gewichtet werden müssten, konnten ebenfalls nicht festgestellt werden. Abbildung 4.6 verdeutlicht dies. Für die jeweils sieben Farben der Helligkeitsstufen zwei und drei ist kein Bereich erkennbar, welcher für einzelne Farben oder zumindest hinsichtlich Helligkeit, Intensität oder Farbton ein besonders starkes Klassifikationsmerkmal darstellen würde. Somit scheint der Verlauf der gesamten Wellenlänge für die Farbbestimmung relevant zu sein.

Der Sensor liefert elf Werte über das Wellenlängenspektrum verteilt. Die digitalisierung des Signals ist damit zwar vergleichsweise grob, liefert aber genug Informationen, um die Verlaufsform des Spektrums repräsentativ wiederzugeben. Weitere Kenngrößen wie beispielsweise Bereichsintegrale, lokale Maxima oder eine Hauptachsentransformation lieferten in eigenen Voruntersuchungen keine signifikantere Informationen. Es konnte im Gegenteil ein tendenzieller Informationsverlust hinsichtlich der Klassifikation festgestellt werden. Eine Transformation in den LCH-Raum erscheint hier ebenfalls nicht zweckmäßig. Hierdurch gehen nachweislich relevante Informationen verloren, welche für die Farbwahrnehmung des Menschen relevant sind (Abschnitt 3.1). Hinsichtlich der Datenverarbeitung auf einem Mikroprozessor gibt es keinen Grund, die Datenmenge zu reduzieren. Elf Kenngrößen sind mit gängigen Klassifikationsalgorithmen auf einem Mikroprozessor problemlos in der gewünschten Zeit zu verarbeiten. Aus diesen Gründen erscheint es zweckmäßig, das gesamte abgetastete Spektrum als Eingangsgrößen zur Klassifikation heranzuziehen. Von einer Berechnung zusätzlicher Merkmale oder einer Reduktion der Messdaten ist keine Verbesserung der Klassifikation zu erwarten.

Bei Messdaten für Klassifikationsaufgaben und speziell bei der Arbeit mit Spektren wird typischerweise eine Normierung der Rohspektren durchgeführt (sofern dies nicht automatisch bei den Klassifikationsalgorithmen integriert ist, beispielsweise durch eine Mittelwertzentrierung) [Aderhold u. a. 2010, Pflüger 2010, Ebel u. a. 1988]. Hierbei bieten sich die Maximumnormierung ($x_{i\text{norm}} = \frac{x_i}{x_{\text{max}}}$) oder eine Normierung auf ein vorgegebenes Minimum und Maximum ($x_{i\text{norm}} = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$) an. Neben der ausgleichenden Wirkung auf diverse Störungen wie Fremdlicht oder Reflexionen dient die Normierung auch dazu, den Einfluss von Wellenlängenbereichen mit unterschiedlich großem Intensitätsbereich bei der Klassifikation zu beschränken. Bei den hier vorliegenden Farbspektren haben die höheren Wellenlängen in der Regel größere Intensitätsbereiche als die niedrigeren Wellenlängen (Abbildungen in Abschnitt 4.2.1). Da jedoch alle Wellenlängenbereiche im Wesentlichen gleich zur Farbe beitragen und es keinen Wellenlängenbereich gibt, der über mehr Aussagekraft bezüglich der Farbklassifikation verfügt (Abbildung 4.6), erscheint eine Normierung der Daten sinnvoll.

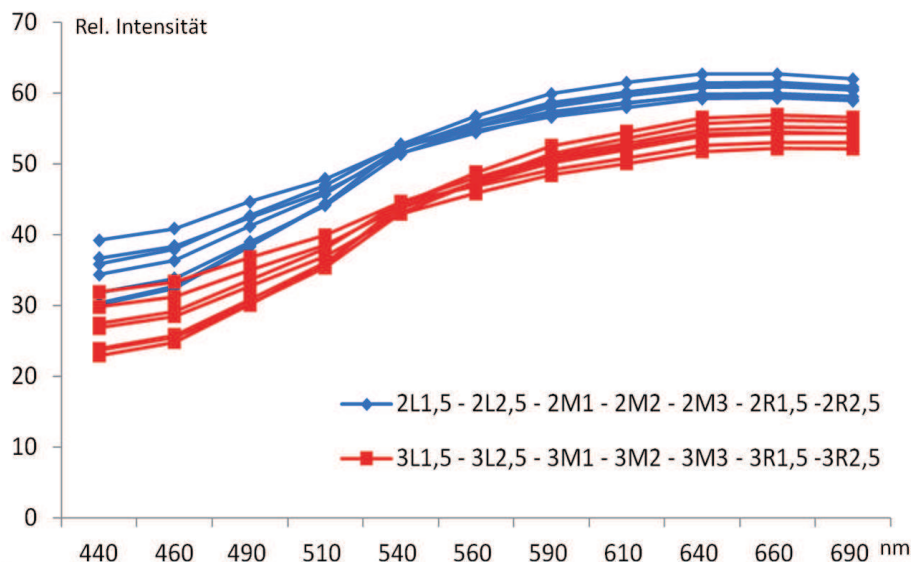


Abbildung 4.6: Beispielhafte Darstellung der Spektren aller Farben der Helligkeitsstufe 2 (rot) und Helligkeitsstufe 3 (blau).

Zusammenfassend bedeutet dies, dass das gesamte abgetastete Spektrum zur Klassifikation verwendet werden muss. Die Bildung von Kenngrößen oder eine Transformation der Spektren liefert keine Verbesserung der Aussagefähigkeit der Daten. Auch hinsichtlich der Verarbeitungsgeschwindigkeit der Algorithmen ist keine Kenngrößenbildung oder Datenfunktionsreduzierung zur Reduzierung der Datenmenge nötig. Damit alle Wellenlängenbereiche gleich gewichtet werden, wird für jede Wellenlänge eine Normierung auf Minimum und Maximum durchgeführt.

Abschätzung bezüglich Fehlerquellen bei der Ermittlung von Trainingsdaten

Bei intelligenten Klassifikatoren ist eine repräsentative und qualitativ hochwertige Auswahl von Trainingsdaten essentiell. Häufig wird hierfür extra ein statistischer Versuchsplan erstellt und durchgeführt, um sämtliche Einflüsse der Eingangskenngrößen optimal abschätzen zu können. Bei der hier vorliegenden Aufgabe ist die Ermittlung von Trainingsdaten problematisch. Zum Einen können die als Eingangsparameter fungierenden Intensitäten der Wellenlängen nicht nach einem Versuchsplan variiert werden, da es sich bei den Daten um reale Messungen an Probanden handelt. Die erste potenzielle Fehlerquelle bei der Ermittlung von Trainingsdaten ist somit, dass diese anhand realer Messungen an menschlichen Probanden gewonnen werden müssen. Somit lässt sich keine beliebige Menge an repräsentativen Trainingsdaten erzeugen. Ein weiterer Punkt ist, dass die Zahnfarben nicht gleichmäßig verteilt sind. Zudem sind die Trainingsdaten aus realen Messungen mit den oben beschriebenen Fehlern behaftet. In einem Großversuch der Firma Vita lagen bei über 1000 Messungen nur neun der Helligkeitsstufe 5 vor, während über 500 Messungen für Zähne der Helligkeitsstufe zwei durchgeführt werden konnten. Weiterhin wurden in diesem Versuchen überhaupt keine Probanden mit Zahnfarben 5M2 und 5M3 gefunden. Die sehr heterogene Verteilung der vorkommenden Farben in den Versuchsdaten ist in Abbildung 4.7 dargestellt. Da die Güte der lernenden Klassifikatoren stark von den Trainingsdaten abhängt, sind fehlende oder stark unterrepräsentierte Farben sowie Fehler in den Messwerten eine weitere potenzielle Fehlerquelle, welche für die Entwicklung eines robusten Verfahrens unbedingt berücksichtigt werden muss.

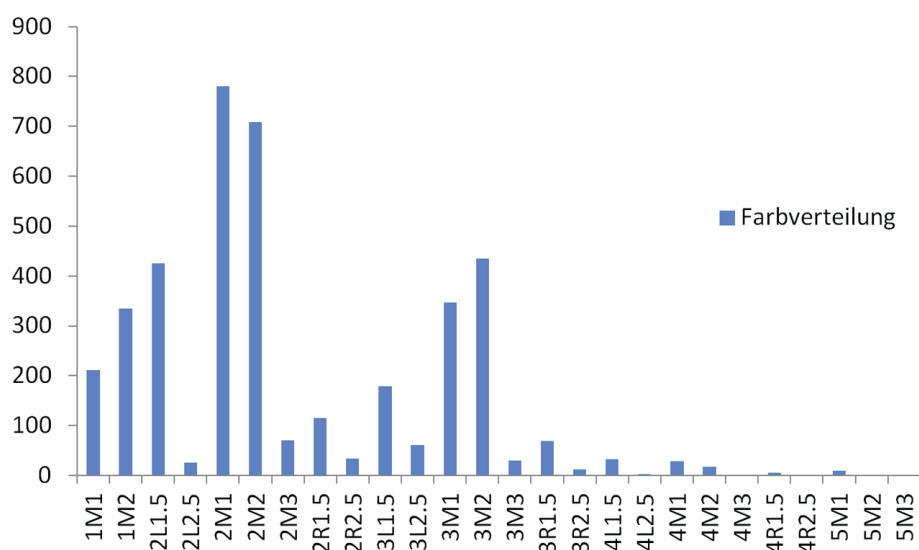


Abbildung 4.7: Exemplarische Farbverteilung bei mehreren Großversuchen von Vita mit knapp 4000 Messungen an realen Zähnen und manueller Klassifikation

Eine weitere Schwäche bei der Gewinnung der Trainingsdaten ist, dass die Messwerte nicht objektiv klassifiziert werden können. Da die zum Training der Klassifikationsverfahren verwendeten Messdaten direkt an Probanden gemessen werden müssen, ist eine Vergleichsmessung mit einem hochgenauen Messgerät in der Regel nicht möglich oder mit unverhältnismäßig hohem Aufwand verbunden (Abschnitt 3.1.1). Somit müssen die gemessenen Spektren für das Training durch Experten vor Ort einer entsprechenden Referenzfarbe zugewiesen werden. Da die eindeutige Farbbestimmung von Zähnen selbst Experten schwer fällt, entsteht dadurch ein weiteres Fehlerpotenzial für die Trainingsdaten. So wurden beispielsweise in dem oben genannten Großversuch der Zahn eines Probanden von drei Experten unabhängig gemessen und klassifiziert. Das Ergebnis ist in Abbildung 4.8 dargestellt. Während die Sensordaten nur eine kleine Schwankung von ca. 4% aufweisen, wurde der Zahn dreimal unterschiedlich klassifiziert (als 1M1, 2M1 und 3M1). Dieser Umstand muss bei der Entwicklung des Verfahrens ebenfalls berücksichtigt werden.

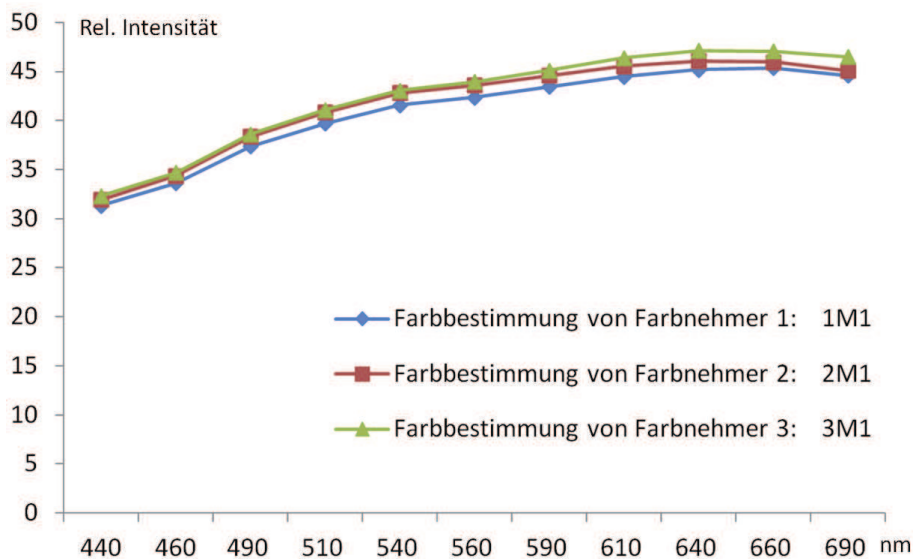


Abbildung 4.8: Drei Messungen am selben, realen Zahn mit unterschiedlicher manueller und visueller Farbbestimmung von drei verschiedenen Experten (1M1, 2M1 und 3M1)

Zusammenfassend kann festgestellt werden, dass mögliche Trainingsdaten für die Klassifikation an realen Zähnen stark heterogen und fehlerbehaftet sind. Dies liegt zum Einen daran, dass für einige Farben wenig bis keine geeigneten Probanden gefunden werden können und zum Anderen, dass die Messungen in der Regel manuell durchgeführt und von Experten subjektiv klassifiziert werden müssen.

4.2.2 Konzeption einer geeigneten Klassifikationsstrategie zur Ermittlung der Zahnfarben

Anhand der Analyse der spektralen Daten wurden mehrere Fehlerquellen identifiziert, welche das Ergebnis der Messung sowie der darauf aufbauenden Klassifikation stark beeinflussen können. Aus diesem Grund müssen Methoden entwickelt werden, welche gegenüber gewissen Schwankungen in den Daten robust sind. Weiterhin müssen zusätzliche Strategien für die Datenverarbeitung ausgearbeitet und umgesetzt werden, um fehlerhafte Daten zu erkennen und möglichst gut kompensieren zu können. Im Prinzip gibt es vier Bereiche, in welchen entsprechende Strategien zum Umgang mit fehlerhaften Daten zur Optimierung des gesamten Verfahrens umgesetzt werden können.

1. Bewertung und Auswahl geeigneter Trainingsdaten

Die Güte des gewählten Klassifikationsverfahrens hängt direkt von der Qualität der Trainingsdaten ab. Diese sind jedoch nicht optimal verteilt und können fehlerbehaftet sein. Deshalb empfiehlt es sich, die Trainingsdaten vorher zu bewerten und offensichtlich fehlerhafte Daten auszuschleusen. Durch das Aussortieren solcher Trainingsdatensätzen kann das Training optimiert und folglich das Modell stark verbessert werden. Hinsichtlich der ungleichen Farbverteilung der Trainingsdaten kann kurzfristig nur durch das Duplizieren von vorhandenen Trainingsdaten entgegengewirkt werden, sofern das gewählte Verfahren dies nötig macht. Langfristig sollten weitere Messungen an Probanden durchgeführt werden, um einen möglichst umfassenden Pool an Trainingsdaten akquirieren zu können. Dies kann in der durchgeführten Arbeit nicht geleistet werden. Es muss jedoch eine entsprechende Möglichkeit geschaffen werden, neu akquirierte Trainingsdaten in das Modell zu integrieren und ein entsprechendes Nachtraining durchführen zu können.

2. Optimierte Klassifikation

Nahezu alle intelligenten Klassifikationsverfahren liefern eine sehr gute Klassifikationsgüte, wenn charakteristische Kenngrößen und gut trennbare Cluster in den Trainingsdaten vorliegen. Dies ist hier nur begrenzt der Fall. Vielmehr unterliegen die Messdaten starken Schwankungen. Weiterhin sind die Unterschiede in den Spektren hinsichtlich der unterschiedlichen Referenzfarben vor allem bei Farben der gleichen Helligkeitsstufe zum Teil gering. Aus diesem Grund muss einerseits ein Klassifikationsverfahren entwickelt werden, welches mit derartigen Bedingungen möglichst gut zurecht kommt. Andererseits müssen neue und problemangepasste Klassifikationsstrategien entwickelt werden, welche die Klassifikationsgüte steigern. Beispielsweise bietet eine Verkleinerung des Lösungsraum Potenzial, um auch kleine Änderungen in den Spektren möglichst gut interpretieren zu können. Ansätze hierfür bietet eine Orientierung an der manuellen und visuellen Farbbestimmung, indem beispielsweise eine mehrstufige oder eine geteilte Klassifikation durchgeführt wird.

3. Plausibilitätsprüfung der Ergebnisse

Anhand der spezifischen Probleme bei der Akquise von Trainingsdaten sowie den Unsicherheiten bei der Messung, muss das Ergebnis einer Klassifikation auf Plausibilität geprüft werden. Hierfür müssen zwei aufgabenspezifische Bedingungen beachtet werden. Zum Einen wird der trainierte Klassifikator mit bekannten Referenzdaten validiert, um dessen Modellgüte zu bestimmen. Diese Validierungsdaten können aus genannten Gründen jedoch ebenfalls fehlerbehaftet sein. Dies muss bei dem zu entwickelnden Bewertungsverfahren berücksichtigt werden. Weiterhin müssen die Klassifikationsergebnisse im laufenden Betrieb auf Plausibilität überprüft werden. Ein möglicher Ansatz zur Erfüllung der Bedingungen ist eine erweiterte Analyse der wahrscheinlichsten Klassenzuweisung hinsichtlich derer Lage und Nachbarschaften im Farbraum. Zudem kann der Wertebereich der Messdaten vor der Klassifikation überprüft und Messungen in einem unwahrscheinlich erscheinenden Wertebereichen herausgefiltert werden.

4. Anwendungsoptimierung

Da der Sensor im Prinzip über eine gute Wiederholgenauigkeit verfügt, können Schulungen im Umgang mit dem Sensor oder Hinweise auf die Gefahr des Verkippens die Messergebnisse und die Robustheit des Klassifikationsverfahrens stark verbessern. Weiterhin bietet der Messkopf Optimierungspotenzial. Ein flexibel gelagerter Messkopf kann beispielsweise die Verkippungsgefahr kompensieren. Schulungskonzepte und Hardwareoptimierung sind jedoch nicht Teil der Arbeit, weshalb die ersten drei Punkte fokussiert betrachtet werden.

Im folgenden werden Strategien und Methoden analysiert und umgesetzt, um geeignete Lösungen für die Teilansätze 1. - 3. zu erarbeiten.

4.3 Methoden zur automatischen Bewertung von Trainingsdaten

Die für die gestellte Aufgabe verfügbaren Trainingsdaten sind aus oben genannten Gründen mit Fehlern behaftet und somit nicht optimal. Für ein Klassifikationsverfahren auf Basis von maschinellem Lernen sind gute Trainingsdaten jedoch ein elementarer Baustein, um eine hohe Klassifikationsgüte zu erreichen. Aus diesem Grund wird eine Methode gesucht, welche aus einer gegebenen Menge an Trainingsdaten die einzelnen Datensätze hinsichtlich Plausibilität bewertet und dazu fehlerhafte oder inkonsistente Trainingsdaten automatisch herausfiltert.

Die Notwendigkeit für die Bewertung der Trainingsdaten wird im Folgenden genauer verdeutlicht. Als Referenzdaten dienen knapp 700 Messungen an Zähnen mit der Helligkeitsstu-

fe zwei, welche von Experten der Firma Vita gemessen und klassifiziert wurden. Abbildung 4.9 zeigt, wie die einzelnen Farbmuster dieser Referenzdaten im Sammons-Mapping mit den folgenden Labels dargestellt werden.



Abbildung 4.9: Labels der Referenzdaten für die Darstellung im Sammons-Mapping

In Abbildung 4.10 werden alle Datensätze der Referenzdaten als zweidimensionale Projektion dargestellt. Es ist deutlich zu sehen, dass nahezu keine Cluster hinsichtlich der Referenzfarben vorhanden sind, wodurch die Qualität der Daten für Trainingszwecke als unzureichend einzustufen ist.

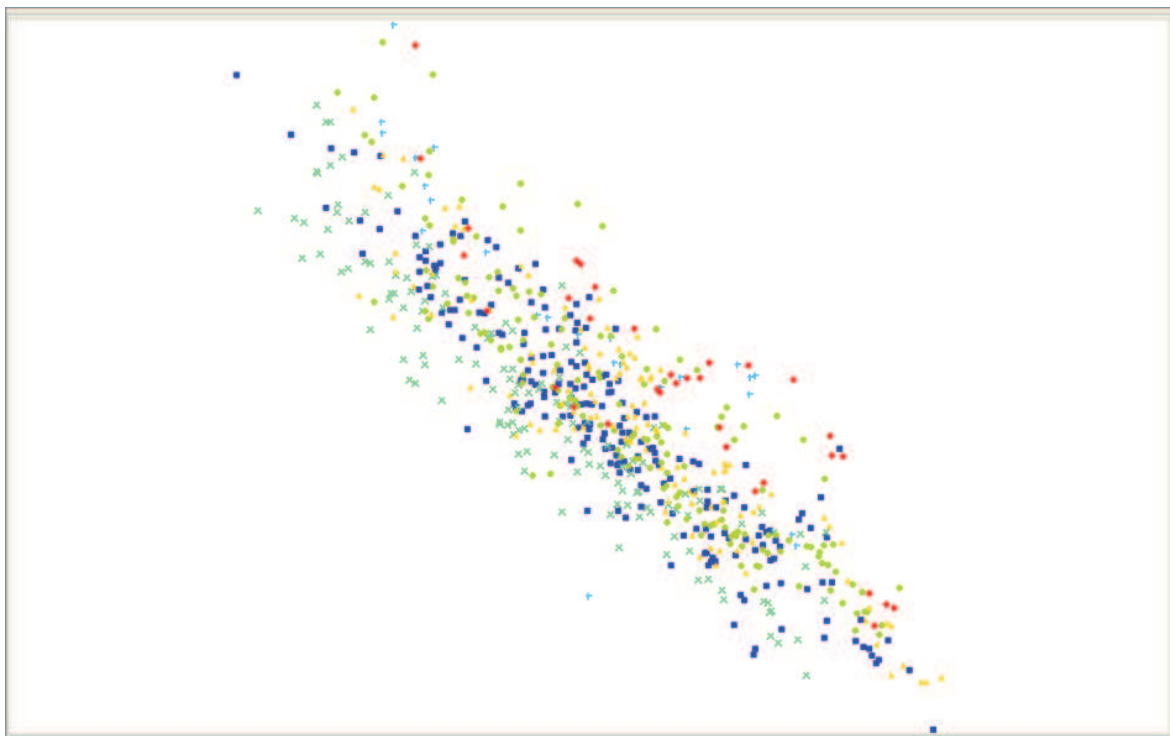


Abbildung 4.10: Sammons-Mapping von Messdaten an realen Zähnen der Helligkeit 2.

Auch durch die Vorverarbeitung der Daten oder die Bildung von Kenngrößen lässt sich dieser Zustand nicht verbessern. Das Herausfiltern von potentiell fehlerhaften Datensätzen bietet somit eine aussichtsreiche Möglichkeit, um die Qualität der Daten hinsichtlich Clusterbildung und Trainingstauglichkeit zu steigern.

4.3.1 Hüllkurven-Schwellwert-Methode

Bei der Hüllkurvenüberwachung werden die einzelnen Messwerte der Datensätze dahingehend bewertet, ob sie sich innerhalb definierter Grenzen über den Verlauf des gesamten Datensatzes befinden. Zur Definition dieser Grenzen kann ein Referenzdatensatz ausgewählt oder erzeugt werden, beispielsweise der statistische Mittelwert aus mehreren gemessenen Daten. Weiterhin muss für jede Kenngröße ein entsprechender Toleranzbereich festgelegt werden. Da im hier vorliegenden Fall keine relevanten Spektralbereiche identifiziert werden konnten, kann eine prozentuale Hüllkurvenabweichung für jede Wellenlänge als zweckmäßig angesehen werden. Spektren, welche außerhalb dieses Dynamikbereichs liegen, sind mit hoher Wahrscheinlichkeit Fehlmessungen oder falsch klassifizierte Datensätze. Wie groß die Toleranz der Hüllkurven ist, muss generell individuell und problemangepasst festgelegt werden.

Bei Versuchen mit den Referenzdaten wurde aus allen vorhandenen Messungen einer Referenzfarbe der statistische Mittelwert gebildet. Als Hüllkurve wurde eine Abweichung von 5% vom jeweiligen Mittelwert zugelassen. Hierbei blieben für die Helligkeitsstufe zwei noch 68 Messungen übrig, deren Verteilung in Abbildung 4.11 dargestellt ist. Die restlichen Datensätze wurde als fehlerhaft verworfen. Durch die deutliche Ausdünnung entsteht eine wesentlich bessere Clusterausprägung für die einzelnen Referenzfarben, wodurch ein effizienteres Training mit intelligenten Klassifikationsverfahren möglich ist.

Problematisch an dieser Methode ist, dass eine größere Menge an Daten für jede Referenzfarbe benötigt wird. Nur so ist bei den auftretenden Schwankungen in den Messdaten die Berechnung eines brauchbaren statistischen Mittelwerts möglich, welcher dem idealen Mittelwert für die jeweilige Referenzfarbe akzeptabel angenähert ist. Bei nur wenigen Daten pro Farbe ist diese Methode sehr fehleranfällig, da die Hüllkurven einen großen stochastischen Anteil aufweisen können. Zudem ist die Methode anfällig gegenüber von den Experten falsch klassifizierten Labels, da diese in entsprechender Anzahl den statistischen Mittelwert farbübergreifend ebenfalls verzerren können. Erschwerend kommt hinzu, dass aufgrund der stark schwankenden Messdaten die Hüllkurven-Methode eine sehr große Anzahl an Daten aussortiert. Somit stehen für das eigentliche Training eines Klassifikators nur noch relativ wenige Daten zur Verfügung, was sich wiederum negativ auf das Modelltraining auswirken kann.

Zusammenfassend kann festgestellt werden, dass durch die Hüllkurven-Schwellwert-Methode inkonsistente Datensätze und Ausreiser bei einzelnen Wellenlängen automatisch herausgefiltert werden können. Jedoch ist diese Methode sehr anfällig gegenüber ungleich verteilten Trainingsdaten und schleust in der Regel sehr viele Daten aus. In der vorliegenden Aufga-

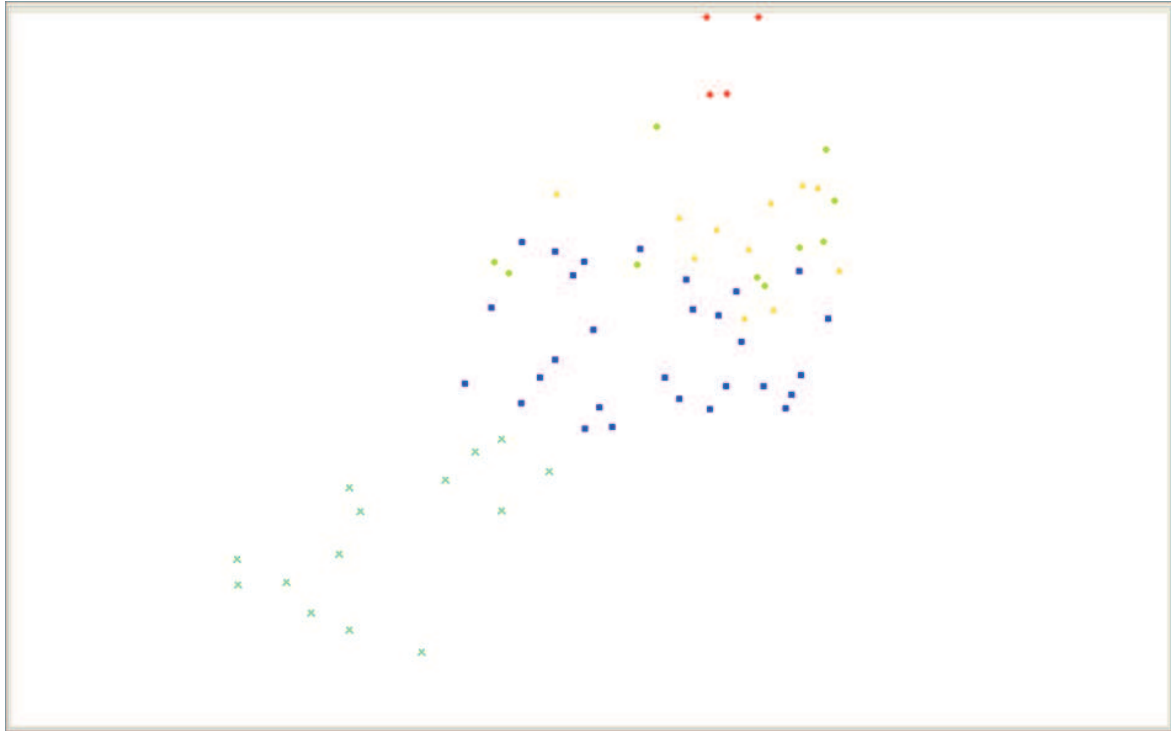


Abbildung 4.11: Sammons-Mapping von Messdaten an Zähnen mit der Helligkeit 2 nach Durchführung der Hüllkruven-Schwellwert-Methode.

be kann sie deshalb bestenfalls zur Vorauswahl oder zur nachträglichen Kontrolle genutzt werden.

4.3.2 Clustering-Methoden

Das Clustering beschreibt generell eine distanzbasierte Methode zur Findung von Clustern und Strukturen in unbekanntem Datenbeständen. Die Clusteranalyse hat hierbei im Gegensatz zu Klassifikationsalgorithmen nicht das Ziel, Daten bestehenden Klassen zuzuordnen. Vielmehr sollen ohne Vorwissen über Klassenzugehörigkeiten nur anhand der Struktur der Daten potenzielle Zugehörigkeiten identifiziert werden.

Eine verbreitete Methode zur Analyse unbekannter Daten ist das hierarchische Clustering. Beim hierarchischen Clustering unterscheidet man allgemein zwischen Top-Down- und Bottom-Up-Verfahren. Beim Top-Down-Verfahren werden alle Objekte zu einem Cluster zusammengefasst und dann schrittweise aufgeteilt. Beim Bottom-Up-Verfahren wird jeder Datensatz zuerst als ein eigenes Cluster interpretiert, welche dann schrittweise zu größeren Clustern zusammengelegt werden. Der Vorteil des hierarchischen Clustering ist, dass durch unterschiedliche Distanzmaße eine strukturierte Clusterhierarchie automatisch berechnet werden kann. Auch können Ausreißer in den Daten gut detektiert werden, da diese oft eigene Cluster mit wenigen Elementen bilden. Nachteilig ist der Analyseaufwand des

Endergebnisses, da die Cluster je nach Beschaffenheit der Daten oft eine unterschiedliche Integrationstiefe besitzen. Dies führt dazu, dass die richtigen Klassen händisch über mehrere Ebene gesucht werden müssen, wodurch diese Methode häufig nicht vollautomatisiert durchgeführt werden kann. [Bacher u. a. 2010].

Eine Möglichkeit die genannten Nachteile des hierarchischen Clusterings zu verbessern, liefert der k -Means-Algorithmus. Hierbei wird für das Clustering die Anzahl der zu suchenden Klassen fest vorgegeben, wodurch die unterschiedlichen Integrationstiefen des hierarchischen Clusterings kompensiert werden können. Nach der Festlegung der Anzahl der Clusterschwerpunkte k werden beim k -Means-Algorithmus initiale Schwerpunkte für jedes Cluster zufällig ermittelt und verteilt. Mit Hilfe einer Distanzfunktion (z.B. den euklidische Abstand) werden nun alle Daten einem der k Cluster zugeordnet. Basierend auf diesen neuen Clustern werden deren Schwerpunkte neu berechnet und die Daten wieder neu geclustert usw. Als Abbruchkriterium kann eine vorgegebene Integrationstiefe dienen oder eine vorgegebene Sättigung bei der Änderung der Clusterschwerpunkte [MacKay 2003]. Vorteilhaft ist hierbei, dass eine einheitliche Klassendichte automatisch und effizient berechnet werden kann. Nachteilig ist hier, dass Ausreiser zu den einzelnen Klassen nur über das Abbruchkriterium und Distanzmaß identifiziert werden können, nicht aber über ein separates Cluster wie beim hierarchischen Clustering. Dies führt dazu, dass der k -Means-Algorithmus zwar wesentlich bessere Ergebnisse hinsichtlich des Clusterings der Daten liefert. Ausreisser und inkonsistente Daten können jedoch kaum automatisch identifiziert werden, da generell jeder Datensatz einem Cluster zugeordnet wird und zudem die Anzahl der Cluster fest vorgegeben ist. Für die hier vorliegende Aufgabe eignet sich somit nur das hierarchische Clustering mit variabler Klassenanzahl.

Bei Versuchen mit den Referenzdaten wurden mit hierarchischem Clustering und manueller Nachbearbeitung inkonsistente Daten ermittelt. Hierbei zeigten Vorversuche, dass die beste Klassentrennung dann vorliegt, wenn die potenziellen Cluster mindestens zu 40% von einer Farbe dominiert werden und mindestens 15% der ursprünglichen Datensätze dieser Farbe enthalten. Das bedeutet, dass beispielsweise bei 215 Datensätze der Farbe 2L1.5 zu einem potenziellen Cluster mindestens 33 Datensätze dieser Farbe gehören müssen. Zudem dürfen nicht mehr als 82 andere Farblabels diesem Cluster zugeordnet sein, ansonsten wird es als potenzielles Farbcluster verworfen. Problematisch ist hierbei, dass manche Farben keine dieser Bedingungen erfüllten, andere wiederum in mehreren Clustern vertreten waren. Auch sind die Endcluster nur selten in einer Hierarchieebene, was eine manuelle Nacharbeitung nötig macht. Für die Helligkeitsstufe 2 bleiben nach obigen Vorgaben und Anwendung der Methode 391 Datensätze übrig, deren Verteilung in Abbildung 4.12 dargestellt ist. Es wird deutlich sichtbar, dass trotz des hierarchischen Clusterings kaum Verbesserungen

hinsichtlich potenzieller Cluster auszumachen sind.

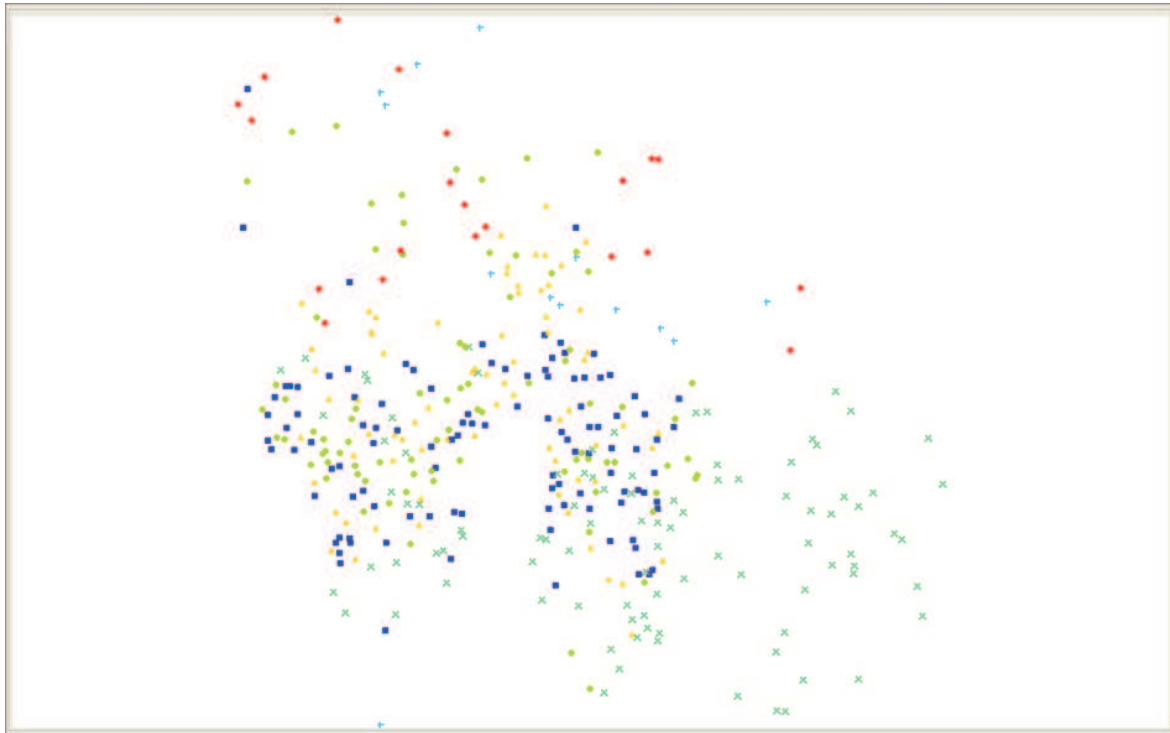


Abbildung 4.12: Sammons-Mapping von Messdaten an Zähnen mit der Helligkeit 2 nach Durchführung der Hierarchischen-Clustering-Methode.

Zusammenfassend kann festgestellt werden, dass das hierarchische Clustering hier nur mit manueller Nachbearbeitung verwendbare Ergebnisse liefert, welche in der Regel jedoch selten von ausreichender Qualität sind. Für die vorliegende Aufgabenstellung muss diese Methode somit als nur bedingt geeignet eingestuft werden.

4.3.3 Majority-Filtering-Methode

Die Methode des Majority-Filtering zur Identifikation von inkonsistenten oder fehlerhaften Trainingsdaten geht auf Brodley und Friedl zurück [Brodley u. a. 1996]. Die Identifikation von schlechten Trainingsdaten wird hierbei als Filter-Verfahren interpretiert. Als “Filter” dienen mehrere Klassifikationsverfahren, welche zuvor mit den zu analysierenden Trainingsdaten trainiert wurden. Im Anschluß werden diese Trainingsdaten wieder von den “Filtern” klassifiziert, um inkonsistente Datensätze zu erkennen. Von einem oder mehreren Filtern als falsch klassifizierte Datensätze sind in diesem Verfahren also potenziell fehlerhafte Datensätze und können nach problemangepassten Regeln herausgefiltert werden. Entscheidend für die Güte des Verfahrens sind hierbei also einerseits die Güte der gewählten Klassifikationsverfahren und andererseits die Schwellen sowie die Logik, wie Trainingsdaten herausgefiltert werden. In der Regel berechnet jeder Filter für den zu bewertenden Trainingsdatensatz eine

Wahrscheinlichkeit oder einen Prognosewert. Durch deren logische Kombinationen wird ein Schwellwert für das Ausschleusen ermittelt. Der prinzipielle Ablauf der Majority-Filtering Methode ist in Abbildung 4.13 dargestellt [Guan u. a. 2010].

Algorithm: MajorityFiltering (MF)

Input: E (training set)

Parameter: n (number of subjects), y (number of learning algorithms)

A_1, A_2, \dots, A_y (y kinds of learning algorithms)

Output: A (detected noisy subset of E)

```

(1) form  $n$  disjoint almost equally sized subsets of  $E_i$ , where  $\bigcup_i E_i = E$ 
(2)  $A \leftarrow \emptyset$ 
(3) for  $i = 1, \dots, n$  do
(4)   form  $E_t \leftarrow E \setminus E_i$ 
(5)   for  $j = 1, \dots, y$  do
(6)     induce  $H_j$  based on examples in  $E_t$  and  $A_j$ 
(7)   end for
(8)   for every  $e \in E_i$  do
(9)      $ErrorCounter \leftarrow 0$ 
(10)    for  $j = 1, \dots, y$  do
(11)      if  $H_j$  incorrectly classifies  $e$ 
(12)        then  $ErrorCounter \leftarrow ErrorCounter + 1$ 
(13)      end for
(14)      if  $ErrorCounter > \frac{y}{2}$ , then  $A \leftarrow A \cup \{e\}$ 
(15)    end for
(16) end for

```

Abbildung 4.13: Prinzipieller Ablauf des Majority-Filtering. Quelle: [Guan u. a. 2010]

Vorteilhaft an dieser Methode ist die relativ einfache und sehr effiziente Möglichkeit, automatisch potenziell fehlerhafte Datensätze zu identifizieren. Weiterhin besteht durch eine problemangepasste Auswahl von Klassifikationsverfahren als Filter sowie passende Ausschlusslogiken ein großes Optimierungspotenzial. Nachteilig ist hier die große Vielfalt an Einstellungsmöglichkeiten zu nennen, was zwangsläufig Expertenwissen voraussetzt. Jedoch hat die Methode selbst mit Standardlernverfahren und einfachen Verknüpfungen der Prognosewerte als Schwellwert ein großes Potenzial und liefert auch in praktischen Anwendungen häufig gute Ergebnisse [Guan u. a. 2010]. Ein weiterer Nachteil der Methode ist, dass sie bis zu einem gewissen Grad die Nachteile der als Filter verwendeten Klassifikationsverfahren aufweist (Abhängigkeit der Verteilung und Anzahl von Trainingsdaten, Generalisierung, Extrapolation, Feature Selection, Merkmalsabhängigkeit usw.). Durch die Kombination von mehreren Klassifikationsverfahren als Filter können diese Nachteile kompensiert werden. Jedoch besteht dann bei stark verrauschten Daten die Gefahr, dass durch die unterschiedlichen Filter eine vergleichsweise große Menge an Datensätzen ausgeschleust wird. Somit bietet diese Methode für die hier vorliegende Aufgabe ein großes Potenzial zur

Optimierung, vorausgesetzt es ist eine entsprechend große Menge an Trainingsdaten vorhanden.

Bei Versuchen mit den Referenzdaten wurde die Majority-Filtering Methode mit drei Klassifikationsverfahren als Filter eingesetzt (künstliche Neuronale Netze, Logit Boost sowie Support Vector Machines). Für jeden dieser Filter wurde ein Prognosewert zwischen Null und Eins berechnet. Ausgeschleusst wurden Datensätze, bei denen ein mittlerer Prognosewert aller drei Filter von kleiner als 0,4 vorliegt. Hierbei bleiben 107 Datensätze übrig, welche im Sammons-Mapping trennbarere Cluster mit deutlich weniger Überlappungen als bei den originalen Daten liefern (Abbildung 4.14).

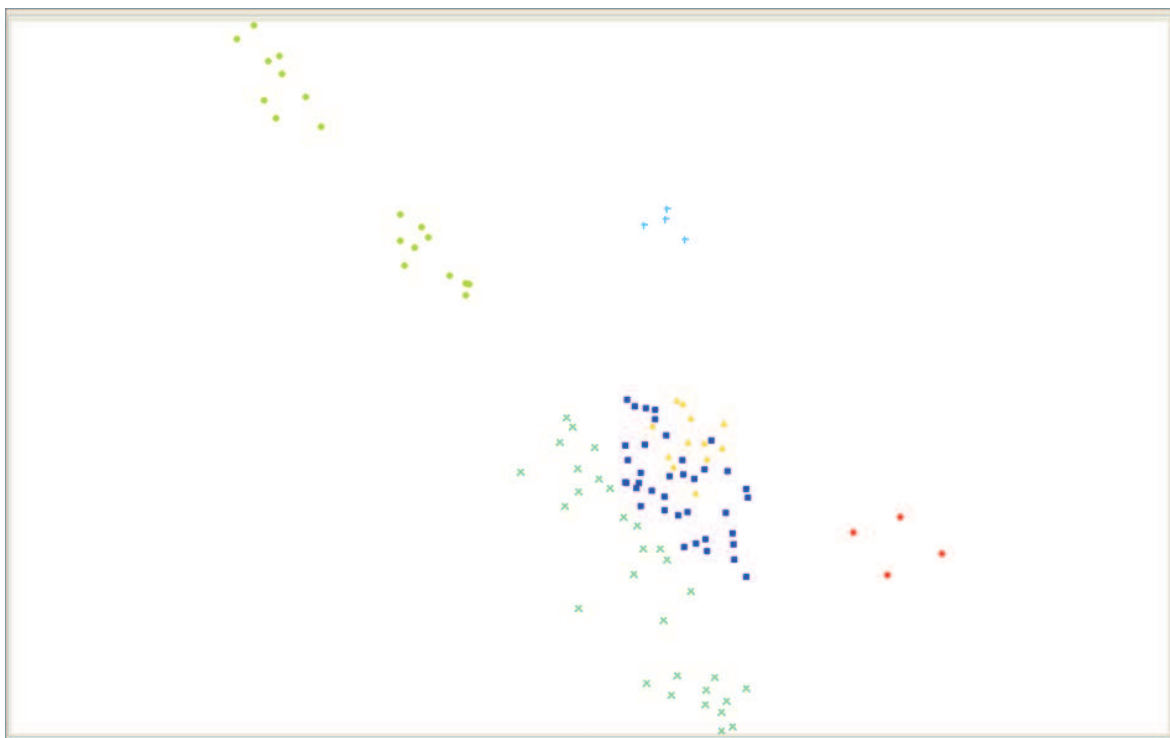


Abbildung 4.14: Sammons-Mapping von Messdaten an Zähnen mit der Helligkeit 2 nach Durchführung der Majority-Filtering-Methode.

Zusammenfassend kann festgestellt werden, dass das Majority-Filtering Cluster mit weniger Überlappungen liefert und somit für die vorliegende Klassifikationsaufgabe Optimierungspotenzial bietet.

4.4 Auswahl geeigneter Klassifikationsmethoden

Das Kernelement der automatischen Klassifikation ist ein geeignetes Modellierungsverfahren. Derartige Modelle müssen unterschiedlichste Anforderungen möglichst gut erfüllen.

Dies beinhaltet zunächst eine gute und robuste Prognose der aktuellen Zahnfarbe auf Basis spektraler Sensordaten. Hierbei sollte neben dem diskreten Prognosewert der Farbe auch ein kontinuierlicher Signifikanzwert berechenbar sein, um die Sicherheit des prognostizierten Ergebnisses hinsichtlich Plausibilität abschätzen zu können. Weiterhin müssen die Verfahren eine gewisse Robustheit hinsichtlich fehlerbehafteten Trainingsdaten aufweisen. Aufgrund praktischer Einschränkungen bei der experimentellen Ermittlung von Trainingsdaten sind diese nur in begrenzter Menge verfügbar und in der Regel stark ungleich verteilt. Dazu muss das gewählte Klassifikationsverfahren Fehler in den Datensätzen handhaben können, welche durch fehlerhafte Bedienung, äußere Einflüsse oder aus mangelnder Qualität der Hardware resultieren. Möglichkeiten zur Erweiterung des Modells oder zu einem Nachtraining bei veränderten Trainingsdaten sowie dem Hinzufügen neuer Farbklassen müssen ebenfalls gegeben sein.

Zur Auswahl eines geeigneten Modellierungsverfahrens auf Basis des maschinellen Lernens muss zuerst das sogenannte “No Free Lunch Theorem” (NFL) berücksichtigt werden [Wolpert u. a. 1997]. Das NFL-Theorem besagt, dass es für Optimierung-, Such- und Lernverfahren im Allgemeinen kein Verfahren gibt, welches generell besser ist, als ein anderes. Demzufolge kann ein gutes Modellierungsverfahren für das eine Problem für ein anderes Problem wesentlich schlechter ausfallen. Dies bedeutet ebenfalls, dass die unterschiedlichen Klassifikationsverfahren zwar theoretisch nach ihren mathematischen Eigenschaften bewertet und praktisch an unterschiedlichsten Benchmarking-Verfahren evaluiert werden können. Jedoch kann nicht garantiert werden, dass dadurch auch die beste Lösung gefunden wird. Speziell Lernverfahren, die von der Zusammensetzung der Trainingsdaten oder der Startinitialisierung abhängen, können große Schwankungen in der Modellgüte aufweisen.

Obwohl es also nicht möglich ist, den “besten” Klassifikator auszuwählen, weisen alle Klassifikatoren doch verschiedene Vor- und Nachteile auf, welche für die vorliegende Aufgabe als relevant erscheinen. Aus diesem Grund werden die gängigsten Klassifikatortypen (Abschnitt 3.2.2) hinsichtlich der Anforderungen der vorliegenden Aufgabe nach folgenden Punkten untersucht:

1. Rechenaufwand

Das entwickelte Verfahren muss auf einem Mikroprozessor direkt im Sensor ausgeführt werden können. Somit steht zumindest für die Prognose nur eine begrenzte Rechenleistung zur Verfügung (welche jedoch nicht echtzeitfähig sein muss). Das Training des Modells anhand von Trainingsdaten kann auch auf einem externen Rechner ausgeführt werden. Die Anzahl der auszuwertenden Merkmale bei der Prognose ist vergleichsweise gering und folglich irrelevant.

2. Einfluss der Trainingsdaten auf die Modellgüte

Die vorhandenen Trainingsdaten sind mit hoher Wahrscheinlichkeit fehlerbehaftet und bezüglich der Referenzklassen ungleich verteilt. Das gewählte Modellierungsverfahren muss also eine hohe Robustheit gegenüber fehlerhaften Daten aufweisen.

3. Komplexität des Modells und Nachtraining

Das Einsatzgebiet des Auswerteverfahrens liegt in der Ermittlung der Zahnfarbe anhand der Daten eines handlichen Smart-Sensors. Deshalb muss das Auswerteverfahren ohne Expertenwissen nutzbar sein. Dies beinhaltet sowohl eventuellen Parametrierungsaufwand des Modells hinsichtlich Training und Klassifikation als auch eine möglichst einfache Plausibilitätsprüfung der Ergebnisse. Das gewählte Modell muss zudem für neue Trainingsdatensätze sowie neue Farbklassen oder ganze Farbsysteme einfach erweiterbar sein. Die Möglichkeit einer Anpassung eines bestehenden Modells an neue Datensätze wäre wünschenswert, die automatische Möglichkeit eines kompletten Neutrainings mit geänderten Trainingsdaten ist eine Notwendigkeit. Wünschenswert wäre zudem ein Modell, welches für den Nutzer intuitiv verständlich ist und aussagefähig visualisiert werden kann.

4. Abstraktionsfähigkeit und Klassifikationsgüte

Aufgrund der Fehleranfälligkeit der Messdaten, muss das Modellierungsverfahren über eine hohe Abstraktionsfähigkeit verfügen. Modelle, welche die Trainingsdaten stark nachbilden und zu einer Überanpassung neigen, sind hier nur bedingt geeignet. Auch eine gewisse Robustheit hinsichtlich fehlerbehafteten Eingangsdaten oder einzelnen fehlerhaften Merkmalen wird benötigt. Eine hohe Klassifikationsgüte bei gleichzeitiger Robustheit gegen Schwankungen ist eine der wichtigsten Anforderungen.

Anforderungsanalyse Support Vector Machines:

SVM bieten eine schnelle und einfache Klassifizierung auch bei vielen Merkmalen. Der Einfluss der Verteilung der Trainingsdaten auf die Modellgüte kann als eher gering eingeschätzt werden. Die Modellgüte ist vom gewählten Kernel abhängig, jedoch erreicht man mit Standardkernen und einer Optimierung der Hyperebenen meist sehr gute Ergebnisse. Auch bei stark verrauschten und inkonsistenten Daten sowie teilweise überlappenden Clustern weisen SVM meist noch eine ordentliche Generalisierungsfähigkeit auf. SVM sind bei mehrdimensionalen Merkmalsvektoren nicht intuitiv verständlich und ein bestehendes Modell kann nicht erweitert werden. Jedoch ist ein Neutrainings mit weiteren Klassen problemlos möglich, auch wenn das Training verhältnismäßig viel Zeit in Anspruch nimmt.

Anforderungsanalyse Naive Bayes Klassifikator:

NBKs zeichnen sich durch eine nicht sehr rechenintensive und robuste Klassifikation aus.

Durch ein inkrementelles Lernen wird das Modell umso genauer, je mehr Trainingsdatensätze zur Verfügung stehen, ohne allzu anfällig für eine Überanpassung des Modells zu sein. Entsprechend jedoch benötigt der NBK auch eine vergleichsweise größere Anzahl an Trainingsdaten. Weiterhin berücksichtigt ein NBK keine Abhängigkeiten zwischen den einzelnen Merkmalen, wodurch vor allem hochdimensionale Vektoren oder sich beeinflussende Merkmale (wie sie bei den Wellenlängen zur Farbermittlung vorliegen) als potenzielle Fehlerquellen einzuschätzen sind. Das Modell des NBK arbeitet mit bedingten Wahrscheinlichkeiten. Dadurch ist sowohl ein Nachtraining des bestehenden Modells bei neuen Trainingsdaten als auch die Integration von mathematisch beschreibbarem Expertenwissen möglich. Weiterhin zeichnet sich der NBK durch ordentliche Abstraktionseigenschaften und Robustheit gegenüber Fehlern in der Datenstruktur aus, unter der Voraussetzung unabhängiger Merkmale und einer ausreichenden Anzahl an Trainingsdaten.

Anforderungsanalyse Logit Boost:

Der LB ist durch eine sehr schnelle und effiziente Klassifikation charakterisiert. Das Modell weist jedoch in der Regel größere Abhängigkeiten hinsichtlich der Verteilung der Trainingsdaten auf. So verzweigen die zugrunde liegenden Entscheidungsbäume besonders oft bei Klassen, denen viele Trainingsdatensätze zugrunde liegen. Die Modelle sind jedoch in der Regel auch für Laien sehr gut verständlich und können zum Teil sogar intuitiv auf Plausibilität bewertet werden. Das Hinzufügen neuer Klassen ist jedoch nicht möglich. Neue Farbklassen oder Trainingsdatensätzen bedingen somit ein komplettes Neutraining. Der LB weist eine ordentliche Abstraktionsfähigkeit und allgemeine Robustheit hinsichtlich Schwankungen in den Daten auf. Größere Abweichungen einzelner Merkmale können jedoch das Ergebnis aufgrund der Entscheidungsbaumstruktur überproportional stark beeinflussen.

Anforderungsanalyse künstlicher Neuronaler Netze:

KNN liefern eine schnelle und einfache Klassifizierung, welche auf einfachen Rechenoperationen basiert. Für ein effizientes Netztraining ist eine große Menge an Trainingsdaten nötig. Auch kann eine ungleiche Verteilung der Trainingsdaten zu Verzerrungen bei der Modellbildung führen. KNNs verfügen über eine große Menge an Parametern, liefern jedoch mit Standardparametern und kleineren problemspezifischen Optimierungen meist gute und robuste Modelle, auch für nicht lineare Zusammenhänge. Die KNN-Modelle sind klassische Blackbox-Modelle und somit in der Regel nicht intuitiv verständlich. Bei neuen Trainingsdaten können bestehende Modelle nachtrainiert werden. Weiterhin weisen KNNs mit entsprechenden Abbruchkriterien auch bei vergleichsweise schlechten Trainingsdaten eine gute Generalisierungsfähigkeit auf und liefern auch bei inkonsistenten Daten und starkem Rauschen meist noch gute Ergebnisse. Zudem sind KNNs in den mit Abstand meisten

industriellen Praxisanwendungen zu finden, was ein nicht zu unterschätzendes Akzeptanzkriterium bei potenziellen Nutzern darstellt.

Anforderungsanalyse Learning Vector Quantization:

LVQ ist im Prinzip eine Nächste-Nachbar-Klassifikation, weshalb sie schnell und einfach auch bei hochdimensionalen Vektoren durchgeführt werden kann. Vom Prinzip sind die Modelle relativ robust gegen ungleiche Verteilungen bei den Trainingsdaten, jedoch hängt die Modellgüte sehr von der Startinitialisierung der Stützvektoren ab. Bestehende Modelle sind jedoch sehr intuitiv verständlich und können vergleichsweise einfach erweitert werden. Bei linearen Zusammenhängen sind so sehr effiziente Modelle möglich. Schwächen haben die LVQ Modelle bei nichtlinearen Zusammenhängen, starkem Rauschen und ähnlichen Klassen, da diese durch die Stützvektoren nur schwer abgebildet werden können.

Anforderungsanalyse General Regression Neuro Fuzzy Networks:

GRNFN basieren auf einem neuronalen Netz in Kombination mit Elementen der Fuzzylogik. Die Modelle können jedoch vergleichsweise schnell berechnet werden. Für das Training wird keine allzu große Menge an Trainingsdaten benötigt. Bei gut trennbaren Klassen ist das Verfahren sehr robust gegenüber ungleich verteilten Trainingsdaten. Fehleranfällig ist das Modell bei überlappenden Clustern, besonders im Bereich der Stützstellen. Das Hinzufügen neuer Klassen ist dank der Stützstellenstruktur sehr einfach möglich, zudem kann fuzzifiziertes Expertenwissen dem Modell hinzugefügt werden. GRNFN-Modelle weisen eine ordentliche Generalisierungsfähigkeit auf und liefern bei der Klassifikation von kontinuierlichen Werten meist gute Ergebnisse. Jedoch hat das Modell in der Regel Probleme, Unstetigkeiten in den Daten effizient abzubilden.

Zusammenfassend kann festgestellt werden, dass alle Klassifikationsverfahren entscheidende Vor- und Nachteile aufweisen, was im Einklang mit dem NFL-Theorem zu erwarten war. Aufgrund ihrer Robustheit bei stark verrauschten und inkonsistenten Daten bieten KNNs ein sehr großes Potenzial zur Klassifikation spektraler Daten. Weiterhin erfüllt der LB viele der gestellten Anforderungen, wenn größere Ausreißer bei einzelnen Merkmalen vermieden werden können. Auch SVM erscheinen sehr geeignet. Im Gegensatz dazu haben der NBK durch die Merkmalsabhängigkeit und GRNFN durch die überlappenden Daten eher schlechtere Grundvoraussetzungen, eine entsprechende Klassifikationsgüte zu erreichen. Das LVQ-Verfahren muss als vergleichsweise ungeeignet eingestuft werden, da es aufgrund von Rauschen in den Daten sowie sehr ähnlichen Klassen fehleranfällig ist. Auch durch Vorteile hinsichtlich Nachtrainingsmöglichkeiten und Modellkomplexität können NBK, GRNFN und LVQ die genannten Nachteile nur schwer kompensieren.

Es muss jedoch angemerkt werden, dass keines der vorgestellten Klassifikationsverfahren als optimal angesehen werden kann. Aus diesem Grund werden im Folgenden weitere Ansätze untersucht, wie die Klassifikationsgüte der Modelle weiter verbessert werden kann.

4.5 Konzeption unterschiedlicher Ansätze zur Klassifikation

Die Klassifikationsgüte der vorgestellten Modelle ist für die Kundenzufriedenheit das entscheidende Kriterium. In den vorherigen Kapiteln wurden bereits verschiedenen Klassifikationsverfahren anhand der vorliegenden Rahmenbedingungen vorgestellt. Darauf aufbauend werden in diesem Kapitel Möglichkeiten untersucht, um die Klassifikationsgüte durch neue Modellstrukturen und Klassifikationsansätze zu steigern. Diese werden im Folgenden diskutiert und bewertet.

4.5.1 Ein Modell für alle Klassen

Der typische Standard-Modellierungsansatz ist, alle Klassen in einem Modell abzubilden. Hierbei werden sämtliche Modellparameter so eingestellt, dass diese Klassen möglichst gut abgebildet werden. Die prinzipielle Funktionsweise ist in Abbildung 4.15 dargestellt. In der vorliegenden Aufgabe liegen als Eingangsdaten die einzelnen spektralen Werte bei den gemessenen Wellenlängen vor. Als Ergebnis wird die entsprechende Farbe oder ein Prognosewert für die vorhandenen Referenzfarben berechnet.

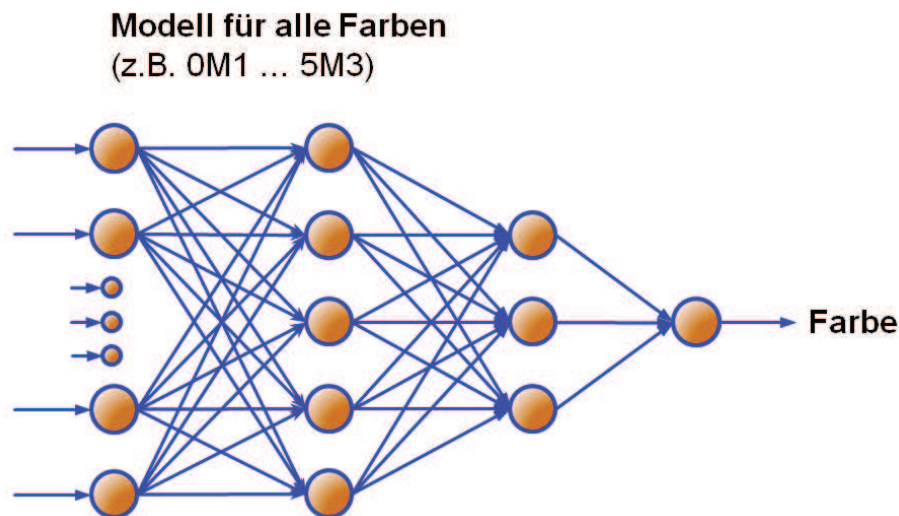


Abbildung 4.15: Klassifikationsansatz mit einem Modell für alle Klassen

Diese Form der Klassifikation hat den Vorteil, dass sie auch für Nutzer mit wenig Erfahrung im Bereich des maschinellen Lernens meist relativ einfach nutzbar ist. Die zugrunde liegenden Algorithmen sind meist ausführlich dokumentiert. Zudem existieren für viele Algorithmen zu diesem Modellierungsansatz eine Reihe an fertigen Softwarebibliotheken, welche mit wenig Aufwand genutzt werden können. In der Regel muss lediglich eine entsprechende Menge an Trainingsdaten mit zugehörigen Klassenlabels bereit gestellt werden, um ein Modell anhand von Standardparametern zu trainieren. Hinsichtlich des Klassifikationsablaufs ist nur ein Klassifikationsschritt nötig, um das Ergebnis zu berechnen. Dank leistungsfähiger Verfahren liefert auch bereits ein einzelnes Modell bei vielen Anwendungen gute Ergebnisse. Jedoch hat diese Form der Modellierung den Nachteil, dass bei einem großen Lösungsraum mit vielen ähnlichen Clustern die Klassifikationsgüte deutlich sinken kann. Da dies in der hier vorliegenden Aufgabenstellung der Fall ist, ist dieser Ansatz im Hinblick auf die Prognosegüte als nicht optimal einzustufen.

4.5.2 Klassenspezifisches Modell

Bei einem klassenspezifischen Modell wird für jede Klasse ein eigenes Modell berechnet. Diese Art der Modellierung wurde bereits in anderen Fachbereichen erfolgreich zur Verbesserung der Klassifikation eingesetzt [Schmidberger u. a. 2007]. Im Gegensatz zum vorherigen Ansatz liegen hier also N Modelle vor, wobei N der Anzahl an unterschiedliche Klassen in den Trainingsdaten entspricht. Jedes Modell $n \in N$ entspricht somit einem Zwei-Klassen-Problem, bei dem überprüft wird, mit welchem Prognosewert y ein Datensatz zur Klasse des jeweiligen Modells gehört. Bei der Klassifikation werden die Eingangsdaten zuerst mit jedem Modell n_i klassifiziert, um die Prognosewerte $y_1 \dots y_n$ zu ermitteln. Anschließend werden diese Prognosewerte mit einer zu definierenden Logik (beispielsweise das Maximum) verknüpft, um das Endergebnis zu ermitteln. Der prinzipielle Ansatz des klassenspezifischen Modells ist in Abbildung 4.16 dargestellt.

Das klassenspezifische Modell hat den Vorteil, dass der Lösungsraum in mehrere Zwei-Klassen-Probleme aufgeteilt und somit die Modelle zielgerichteter auf die jeweilige Klasse hin optimiert werden können. Hierdurch können ähnliche Klassen häufig effizienter getrennt werden. Auch für den Nutzer ist dieses Modell noch vergleichsweise einfach zu bedienen und umzusetzen. Es sind zwar softwaretechnische Erweiterungen bei verfügbaren Bibliotheken hinsichtlich des allgemeinen Ablaufs der N Klassifikationen sowie einer logischen Verknüpfung der Einzelergebnisse y_i nötig. Sind diese jedoch umgesetzt, so muss auch hier vom Nutzer nur ein Satz an Trainingsdatensätzen vorgegeben werden. Der Rechenaufwand bei diesem Klassifikationsansatz ist um den Faktor N höher als bei dem einfachen Modell, da für einen zu klassifizierenden Datensatz der Prognosewert y_i für jedes der N Modelle berechnet werden muss. Ein weiterer Nachteil ist, dass die Modelle zwar mit den gleichen Trainingsda-

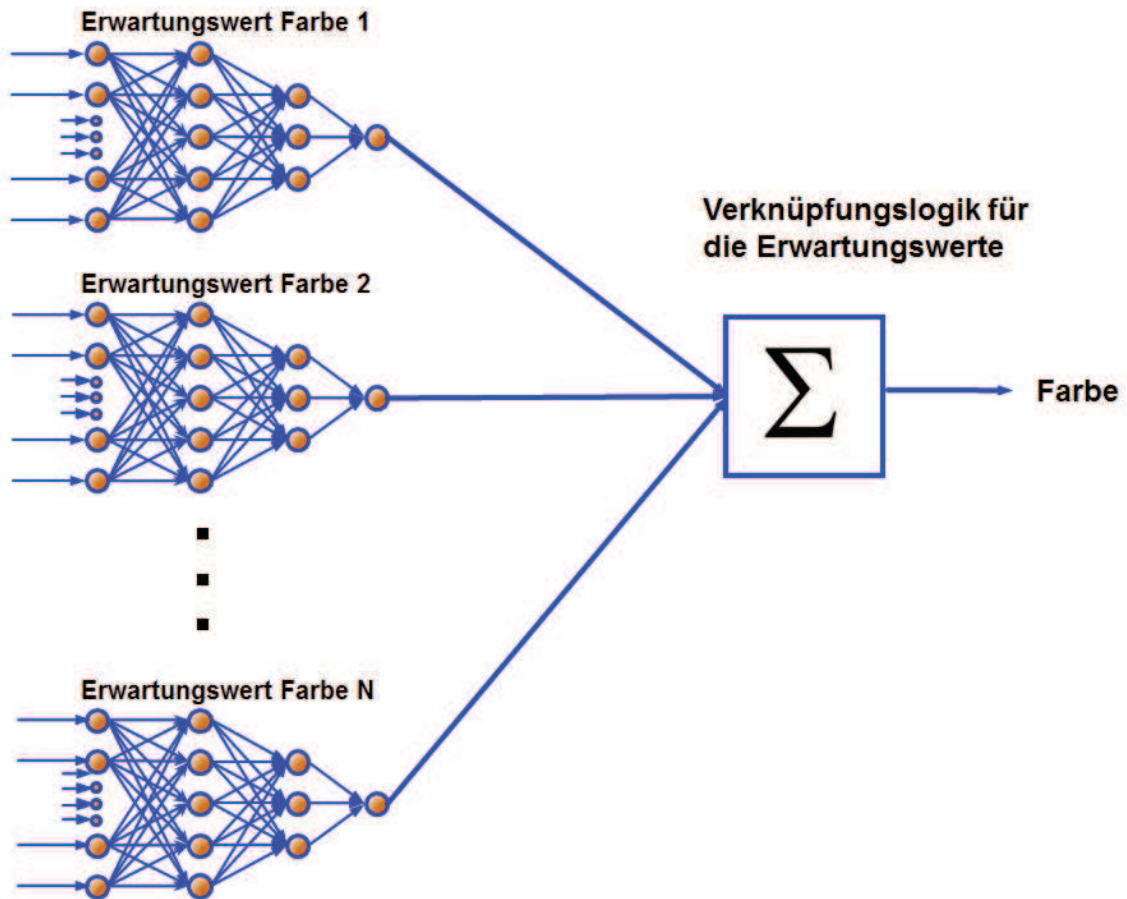


Abbildung 4.16: Klassifikationsansatz mit einem eigenen Modell für jede Klasse

ten trainiert, jedoch nicht abhängig hinsichtlich der Netzverknüpfungen voneinander sind. Je nach Verteilung der Trainingsdaten und Eigenschaften der gewählten Modellierung können dadurch Verzerrungen in den Verhältnissen der berechneten Prognosewerte auftreten. Zusammenfassend bietet dieses Verfahren eine verbesserte Klassentrennung. Jedoch besitzt es schlechtere Laufzeiteigenschaften und weist aufgrund der unabhängigen Modelle potenzielle Unsicherheiten hinsichtlich der Klassifikationsgüte auf. Somit bietet dieses Verfahren nur wenig Optimierungspotenzial.

4.5.3 Mehrstufiges Modell

Betrachtet man die spektralen Daten und deren zugehörigen Klassen im Farbraum, so liegen einige Klassen sehr eng beieinander und überlappen sich teilweise. Eine Verkleinerung des Lösungsraums, also eine Konzentration auf bestimmte Klassen, bietet somit einen möglichen Optimierungsansatz, welcher in Abbildung 4.17 verdeutlicht wird. Liegen mehrere ähnliche Klassen in einem Bereich des Lösungsraums (Abbildung 4.17, links), so können Verfahren des maschinellen Lernens an Effizienz hinsichtlich der Trennbarkeit der Klassen verlieren. Wird der Lösungsraum jedoch verkleinert oder Klassen zusammengefasst, so

kann in der Regel eine bessere Klassentrennbarkeit durch die Klassifikationsverfahren erreicht werden. (Abbildung 4.17, rechts). Eine Möglichkeit hierfür bietet ein mehrstufiges Klassifikationsmodell.

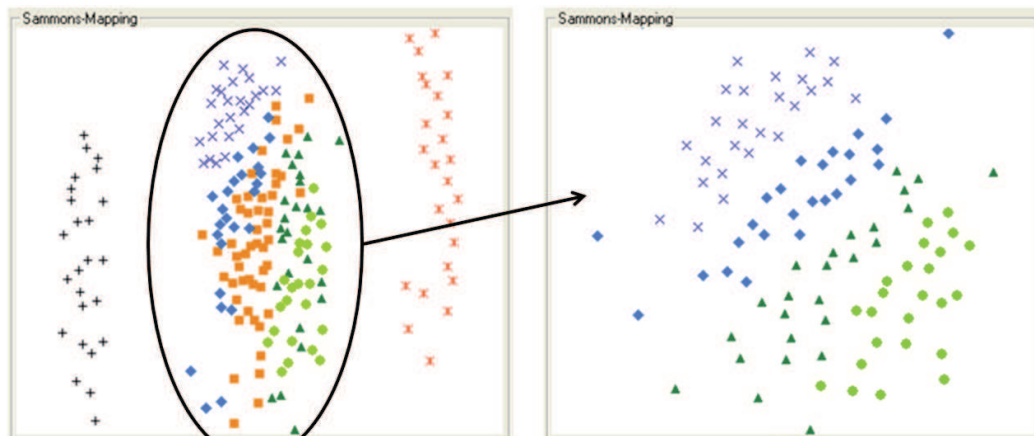


Abbildung 4.17: Beispielhafte Darstellung, wie eng aneinanderliegende Cluster in einem verkleinerten Lösungsraum besser getrennt werden können. Quelle: Eigene Darstellung, angelehnt an [Neher 2011]

Der mehrstufige Klassifikationsansatz besteht aus einer Vorklassifikation und einer oder mehreren Detailklassifikationen. Die Entscheidung, welche Detailklassifikation zur Anwendung kommt, hängt vom Ergebnis der Vorklassifikation ab. Hierbei werden für die Vorklassifikation zunächst ähnliche Klassen zusammengefasst, so dass M neue Klassen entstehen, wobei $M < N$ gilt. Hierdurch wird die ursprüngliche Anzahl der Klassen zu weniger aber besser trennbaren Clustern im Lösungsraum reduziert. Für jede neue Klasse M wird nun ein Modell erstellt, welches die zusammengefassten Klassen i bis j enthält, wobei $[i, j \in N]$ gilt. Somit besteht das gesamte Modell zur Klassifikation aus mehreren voneinander abhängigen Teilmodellen. Je nach Ergebnis der Vorklassifikation wird das entsprechend nächste Teilmodell ausgewählt und aus den noch möglichen Klassen die entsprechende Referenzfarbe klassifiziert. Übertragen auf die Zahnfarbe bedeutet dies, dass beispielsweise alle Farben mit der gleichen Helligkeit für die Vorklassifikation zu einer Klasse zusammengefasst werden. In dem Modell der Detailklassifikation zur Helligkeitsstufe 1 wären dann dementsprechend die Farben 1M1 und 1M2. Der allgemeine Ansatz zur mehrstufigen Klassifikation ist in Abbildung 4.18 dargestellt.

Der Vorteil des mehrstufigen Klassifikationsansatzes ist, dass der Lösungsraum effizienter aufgeteilt und somit eine Verbesserung der Klassentrennbarkeit zu erwarten ist. Zudem ist der Ansatz der manuellen Farbnahme mit dem 3D-Master System nachempfunden, bei der ebenfalls zuerst die Helligkeit und anschließend die detaillierte Farbausprägung ermittelt wird. Die Ergebnisse sind hinsichtlich der Netzverknüpfungen abhängig voneinander,

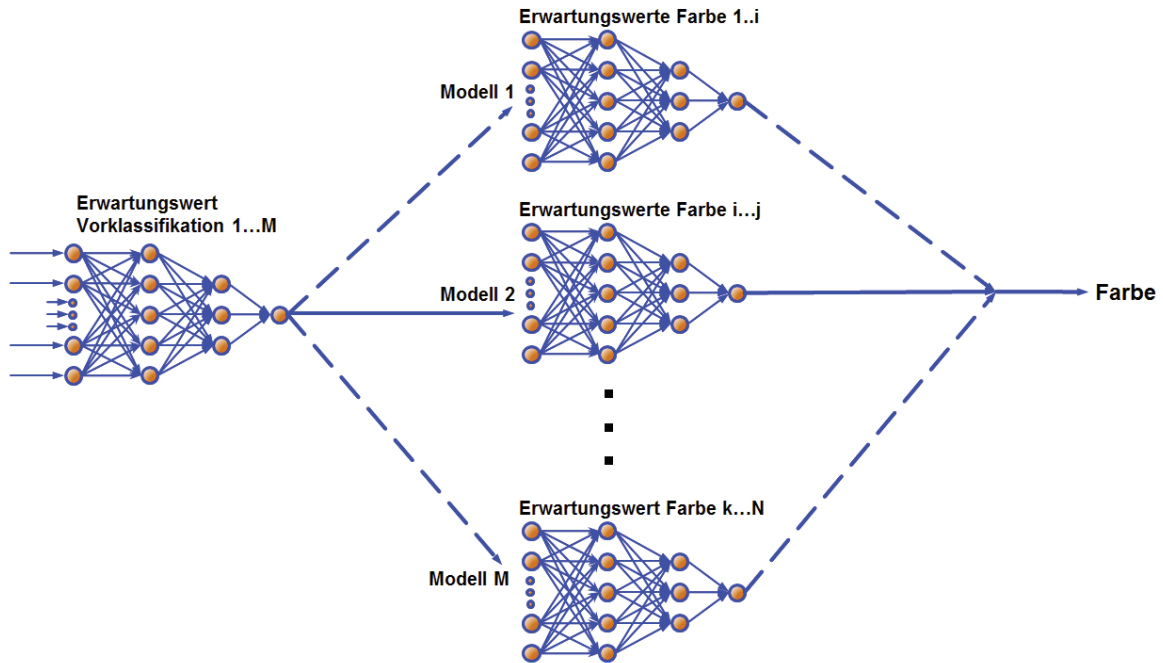


Abbildung 4.18: Mehrstufiger Klassifikationsansatz bestehend aus einer Vorklassifikation und einer abhängigen Detailklassifikation

wodurch keine Minderung der Klassifikationsgüte durch Verzerrungen in den Verhältnissen der Signifikanzwerte auftreten. Auch ist der Rechenaufwand mit $M+1$ Modelltrainings und zwei Klassifikationsberechnungen nicht sehr groß. Der Ansatz des mehrstufigen Modells liefert zudem weitere Möglichkeiten zur Optimierung. So kann der gesamte Klassifikationsprozess durch das Hinzufügen weiterer Klassifikationsschritte auf das jeweilige Problem sehr effizient angepasst werden. Erweiterungen, wie eine Kombination mit dem klassenspezifischen Ansatz, sind ebenfalls möglich, indem dieser beispielsweise für die Detailklassifikation verwendet wird. Als nachteilig ist der Aufwand für den Nutzer zu nennen, da die Definition der zusammengefassten Klassen für die Vorklassifikation je nach Farbsystem mehr oder weniger manuell erfolgen muss. Bei Farbsystemen wie dem 3D-Master ist dies relativ einfach und intuitiv durch die Helligkeit möglich. Beim classical System wird hierfür viel Expertenwissen benötigt. Dies wirkt sich weiterhin auch auf die Flexibilität hinsichtlich eines potenziellen Nachtrainings oder dem Hinzufügen neuer Farben aus.

Zusammenfassend liefert der mehrstufige Klassifikationsansatz die besten Klassifikationsergebnisse hinsichtlich der Klassentrennbarkeit, jedoch auf Kosten der einfachen Bedienbarkeit. Bei einer entsprechenden softwaretechnischen Assistenzfunktion, welche eine einfache Bedienbarkeit für den Nutzer sicherstellt, besitzt dieser Ansatz für die hier vorliegende Aufgabenstellung das größte Potenzial.

4.6 Bewertungskriterien und Plausibilitätsprüfungen der Ergebnisse

In Abschnitt 4.2.1 wurden bereits die Probleme anhand der Datenqualität untersucht. Zudem wurden Lösungen erarbeitet, um die vorhandenen Trainingsdaten zu bewerten und die Klassifikation zu optimieren. Aufgrund der Schwächen der menschlichen Wahrnehmung sowie der potenziellen Fehler des Messgerätes sind sämtliche Trainingsdaten und daraus resultierenden Modelle mit einer gewissen Unsicherheit behaftet. Deshalb sollten die Ergebnisse einer Plausibilitätsprüfung unterzogen werden. Diese Plausibilitätsprüfungen müssen im Lösungsansatz zwei Funktionen erfüllen. Zum einen dienen sie dazu, die vorgestellten Ansätze zu bewerten und das geeignetste Klassifikationsverfahren nach objektiven Kriterien auszuwählen. Zum anderen wird eine Funktion benötigt, um später im laufenden Betrieb die klassifizierten Ergebnisse online bewerten zu können und ggfs. weitere Maßnahmen wie eine Wiederholungsmessung einleiten zu können. Im Folgenden werden zuerst geeignete Kriterien hergeleitet, um die vorgestellten Klassifikationsalgorithmen und Klassifikationsansätze vergleichen und bewerten zu können. Daraufhin werden Strategien erarbeitet, um eine Inline-Plausibilitätsprüfung der Klassifikationsergebnisse direkt im Sensor umsetzen zu können.

Als Basis zur Entwicklung dieser Kriterien wird davon ausgegangen, dass zwei Blöcke an Datensätze zur Verfügung stehen, welche beide mit entsprechenden Farblabels versehen sind. Die einen Datensätze werden zum Training der zu bewertenden Klassifikationsverfahren genutzt (Trainingsdaten), die anderen zur Validierung des gesamten Verfahrens (Validierungsdaten). Angewendet werden die erstellten Bewertungskriterien ausschließlich auf die Validierungsdatensätze. Folgende Kriterien wurden hierfür erarbeitet:

G1: Richtig klassifizierte Datensätze:

Ein obligatorisches Kriterium bei Klassifikationsaufgaben ist die Anzahl der richtig klassifizierten Datensätze. Die Trainingsdaten sind jedoch zu einem großen Teil inkonsistent und fehlerbehaftet, weshalb das Kriterium in der hier vorliegenden Aufgabe stark an Aussagekraft verliert. Nichts desto trotz ist es ein wichtiger Indikator für die Klassifikationsgüte. Das Gütekriterium G1 zur Bewertung des Klassifikationsverfahrens ist somit die Rate der richtig klassifizierten Datensätze.

G2: Richtig klassifizierter Nachbar:

Natürliche Zähne im Mundraum weisen in den seltensten Fällen eine komplett homogene Farbausprägung auf. Auch lassen sich diese nicht immer exakt einer definierten Referenzfarbe in einem Farbsystem zuordnen. In Kombination mit der Farbwahrnehmung des Menschen

ist es somit nur schwer möglich, die Farbe eines natürlichen Zahns eindeutig zu klassifizieren. Aus diesen Gründen ist es meist schon ein ausreichend gutes Ergebnis, wenn die vom Klassifikator prognostizierte Farbe in unmittelbarer Nähe zur erwarteten Referenzfarbe im Farbraum liegt. Deshalb ist nicht nur die Rate der direkt richtig klassifizierten Farben ein Gütekriterium. Auch eine im Farbraum gelegene direkte Nachbarfarbe ist oft ausreichend. Das Gütekriterium G2 zur Bewertung des Klassifikationsverfahrens ist somit die Rate der richtig klassifizierten Farbe oder einer direkten Nachbarfarbe im Farbraum. Hierbei ist jedoch anzumerken, dass die Anzahl der Nachbarn einer Farbe je nach gewähltem Farbraum variieren kann. Im Farbraum 3D-Master hat die Farbe 1M1 beispielsweise die Nachbarn 0M1, 0M2, 1M2, 2M1, 2R1,5 und 2L1,5 während im classical Farbraum die Farbe A1 die Nachbarn A2, B1 und B2 aufweist (Abbildung 2.6 und Abbildung 2.7).

G3: Histogrammverteilung:

Wie bereits in Abschnitt 4.2.1 beschrieben, ist die Ermittlung von Trainingsdaten an realen Probanden nicht beliebig möglich. Folglich sind die Trainingsdaten in der Regel nicht gleich verteilt. Im Gegenteil dominieren einige Farben deutlich, beispielsweise die Farben 2L1,5 und 2M1 aus der Helligkeitsstufe zwei. Somit muss ein Klassifikationsverfahren in der Lage sein, diese Ungleichverteilung möglichst gut abzubilden. In der Untersuchung möglicher Klassifikationsverfahren wurde bereits festgestellt, dass einige Klassifikationsverfahren prinzipbedingt hier eine höhere Güte aufweisen als andere, vor allem in Kombination mit starken Schwankungen in den Datensätzen. Aus diesem Grund bietet sich eine Analyse der Histogrammverteilung durch einen Soll-Ist-Vergleich an. Wenn die Histogramme eine ähnliche Verteilung hinsichtlich der Klassenzugehörigkeit aufweisen, so deutet dies auf eine gute Qualität des Modells hinsichtlich der Generalisierungsfähigkeit hin.

Folgendes theoretisches Beispiel verdeutlicht das Problem. Validierungsdaten der Helligkeitsstufe zwei bestehen aus 20 Datensätzen, wobei auf die Farbe 2M1 14 Datensätze, auf die 6 übrigen je ein Datensatz entfallen. Klassifiziert das trainierte Modell nun jeden Datensatz gleich, also sämtliche 20 Datensätze zu 2M1, so ist das Gütekriterium G1 zu 70% und G2 zu 85% erfüllt. Von den absoluten Zahlen zur Klassifikationsgüte ist dies unter den vorliegenden Randbedingungen ein ordentliches Ergebnis. Jedoch ist offensichtlich, dass dieses Modell nicht funktioniert, da es immer dieselbe Farbe klassifiziert. Die Bewertung der Datenverteilung bietet somit großes Potenzial, um derartige "versteckte" Modellschwächen zu finden und zu bewerten.

Zur Berechnung eines Bewertungskriteriums bietet sich ein Soll-Ist-Vergleich für die Anzahl an Datensätzen in jeder Farbklasse mit anschließender Aufsummierung über alle Klassen an. Zum besseren Vergleich wird das Ergebnis auf einen Wertebereich von Null bis 100

normiert. Sei N die Menge aller vorkommenden Farbklassen in den Validierungsdaten. Dann berechnet sich das Gütekriterium G_{norm} der Histogrammverteilung wie folgt:

$$G_{norm} = \left(1 - \frac{\sum_{i=1}^N \sqrt{(n_{i\,soll} - n_{i\,ist})^2}}{G_{max}}\right) * 100,$$

wobei $n_{i\,soll}$ die Anzahl der Datensätze zur Farbe i und $n_{i\,ist}$ die Anzahl der vom Modell prognostizierten Datensätze zur Farbe i beschreibt. G_{Max} repräsentiert das theoretische Maximum des unnormierten Gütekriteriums, welches vorliegen würde, wenn die Histogramme von Soll- und Istwerten keine Übereinstimmung hätten. Das bedeutet

$$G_{max} = 2 * N.$$

Durch die Differenz zu Eins stimmen die Histogrammverteilungen der Soll- und Ist-Werte somit besser überein, je größer das Gütekriterium ausfällt. Die Qualität des zugrunde liegenden Klassifikationsmodells steigert sich also mit steigendem Gütekriterium G_{norm} . Abbildung 4.19 zeigt dies an einem Beispiel von Farben der Helligkeitsstufe zwei. Da im linken Histogramm die Farbe 2L1.5 überproportional oft prognostiziert wird, liegt hier eine relativ schlechte Verteilungsübereinstimmung von Soll- und Istwerten vor, während im rechten Histogramm die Verteilung der Farben besser übereinstimmt.

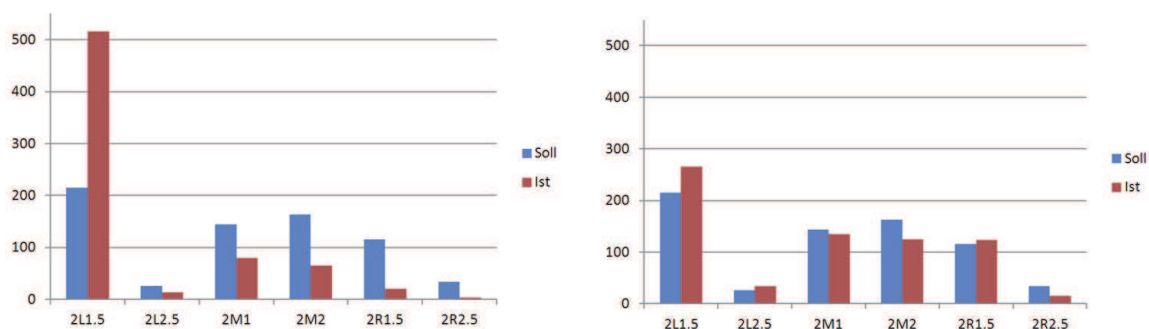


Abbildung 4.19: Schlechte Histogrammverteilung (links) und ordentliche Histogrammverteilung (rechts) zur Bewertung von Klassifikationsmodellen.

Das dritte Gütekriterium G_3 ist somit die Übereinstimmung in den Histogrammen der Farbverteilung von Soll- und Istwerten der prognostizierten Farben.

G4: Erst- und zweitbestes Klassifikationsergebnis:

Die Ergebnisse des Klassifikationsverfahrens müssen konsistent und logisch sein. Ein Hinweis darauf liefert eine genauere Untersuchung der Prognose eines Klassifikators. In der Regel liefern Klassifikationsverfahren für alle Labels einen kontinuierlichen Prognosewert oder einen diskreten Wert, welcher dem Label direkt zugeordnet werden (oder mit überschaubarem Aufwand dahingehend erweitert werden) kann. Dadurch kann für die Eingangsdaten eine wahrscheinlichste und eine zweitwahrscheinlichste Farbklasse bestimmt werden. Damit die Ergebnisse und damit auch das gesamte Klassifikationsverfahren an sich konsistent und logisch sind, müssen die prognostizierte wahrscheinlichste und zweitwahrscheinlichste Referenzfabre direkte Nachbarn im Farbraum sein. Ist dies nicht der Fall, so deutet einiges auf unerwünschte Scheininterpretationen im Modell oder eine schlechte Generalisierung hin. Aus diesem Grund ist eine Überprüfung, ob das beste und zweitbeste Klassifikationsergebnis eine direkte Nachbarschaft im Farbraum aufweisen, das vierte Gütekriterium G4 zur Bewertung des Klassifikationsverfahrens.

Das vierte Gütekriterium ist zudem im Vergleich zu den anderen unabhängig von den Validierungsdaten. Die ersten drei Kriterien benötigen eine Menge an Validierungsdaten mit bekannten Labels, um darauf basierend den Gütewert mit Hilfe von Soll-Ist-Vergleichen zu ermitteln. Das Gütekriterium des erst- und zweitbesten Klassifikationsergebnisses kann jedoch auch bei Messungen im laufenden Betrieb berechnet werden. Es wird kein Label für einen Soll-Ist-Vergleich benötigt. Liegen das beste und zweitbeste Klassifikationsergebnis nicht in direkter Nachbarschaft sondern weiter auseinander, so könnte beispielsweise eine Wiederholungsmessung ausgelöst werden. Dies kann als zusätzliche Erweiterung auch beispielsweise von den Absolutwerten der jeweiligen Prognosewerte abhängig gemacht werden.

Zusammenfassend kann festgestellt werden, dass alle vier Kriterien (richtig klassifizierte Datensätze, richtig klassifizierter Nachbar, Histogrammverteilung sowie erst- und zweitbestes Klassifikationsergebnis) geeignet sind, um den Klassifikationsansatz und das trainierte Modell zu bewerten. Das vierte Kriterium (erst- und zweitbestes Klassifikationsergebnis) ermöglicht zudem eine direkte Bewertung des aktuellen Klassifikationsergebnis im Smart-Sensor während des Betriebs.

Kapitel 5

Verfahrenstechnische Umsetzung und Validierung

Die im vorangegangenen Kapitel entwickelten Lösungskonzepte, speziell die automatische Bewertung von Trainingsdaten sowie die mehrstufige Klassifikation, werden nachfolgend im Rahmen der Validierung überprüft. Hierfür werden die Methoden auf reale spektrale Messdaten von Probanden angewendet und anschließend ausgewertet. Dabei wird die Klassifikationsgüte durch die im vorigen Abschnitt vorgestellten Kriterien ermittelt und diskutiert.

5.1 Ermittlung von Trainingsdaten

Für Klassifikationsverfahren auf Basis des maschinellen Lernens werden aussagekräftige Trainingsdaten benötigt. Da sich diese in der vorliegenden Aufgabenstellung nicht künstlich erzeugen lassen, muss eine entsprechende Menge an Trainingsdaten an Probanden akquiriert und entsprechend bewertet werden. Um ein möglichst allgemeingültiges Verfahren entwickeln zu können, sollten hierfür eine große Menge an Datensätzen zur Verfügung stehen.

Als Sensorsystem für die Akquisition von Trainingsdaten wurde der in Abschnitt 3.1.2 vorgestellte Smart-Sensor VITA EasyShade verwendet. Mit dessen Hilfe werden die Spektren zur Klassifikation der Zahnfarbe gemessen und gespeichert. Als potenzielle Trainingsdaten wird hierfür auf die Daten mehrerer Großversuche der Firma Vita zurückgegriffen. Die Daten wurden hierbei wie folgt ermittelt: Die Firma Vita und das Universitätsklinikum in Mainz führte an mehreren Probanden Messungen durch. Hierbei wurde an jedem Probanden spektrale Messungen am jeweils gleichen Zahn durchgeführt wurden. Die Farben der gemessenen Zähne wurden von Experten vor Ort manuell nach dem Farbsystem 3D-Master und classical bestimmt und dokumentiert. Somit stand ein Volumen von fast 4000

Datensätzen für jeweils beide Farbsysteme zur Verfügung, was für die hier vorliegende Untersuchung mit den bereits erwähnten Einschränkungen eine ausreichende Menge darstellt.

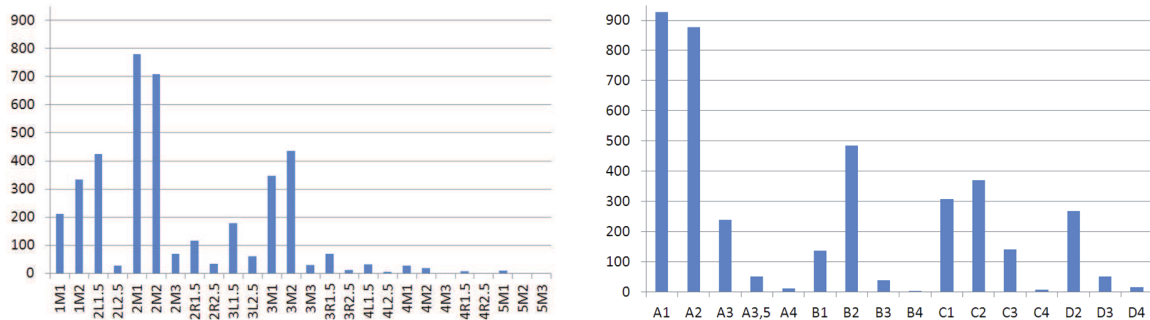


Abbildung 5.1: Anzahl der verfügbaren Trainingsdaten aus Messungen an Probanden von den Farbsystemen 3D-Master (links) und classical (rechts)

Es muss jedoch angemerkt werden, dass sämtliche in Abschnitt 4.2.1 hergeleiteten Fehlerpotenziale auch für diese Datensätze gelten. Weiterhin konnten trotz der Datenmenge im Farbsystem 3D-Master nicht zu allen möglichen Referenzfarben hinreichend viele Datensätze zur Verfügung gestellt werden. Die Verteilung der verfügbaren Daten zu den jeweiligen Farbsystemen ist in Abbildung 5.1 dargestellt. Auch hier ist, wie zu erwarten war, deutlich die Dominanz einzelner Farben zu sehen.

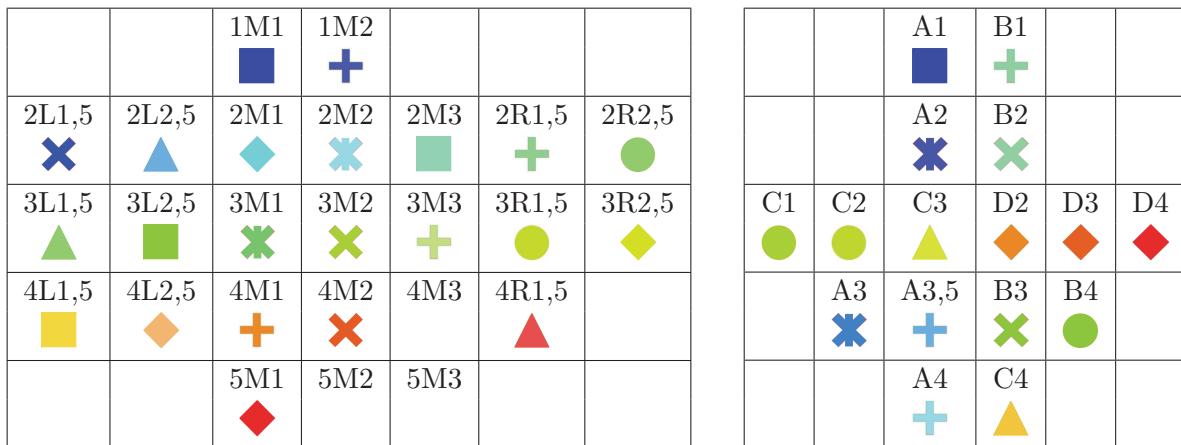


Abbildung 5.2: Labels der ausgewählten Trainingsdaten der Farbsysteme 3D-Master (links) und classical (rechts) für die Darstellung im Sammons-Mapping

Wie bereits in Abschnitt 4.6 beschrieben, werden die vorhandenen Datensätze in Trainingsdaten und Validierungsdaten aufgeteilt. Die Trainingsdaten werden hierbei für das eigentliche Training der Klassifikationsverfahren genutzt. Die Validierungsdaten hingegen werden zur Ermittlung der ebenfalls in Abschnitt 4.6 erarbeiteten Bewertungskriterien unabhängig von den Trainingsdaten verwendet. Zur Vergleichbarkeit mit bisherigen Versuchen

von Vita hinsichtlich der Klassifikationsgüte des VITA EasyShade wurde in Absprache mit Vita 1533 Datensätze als Trainingsdaten ausgewählt. Die restlichen Datensätze werden zur Validierung genutzt.

Zur Untersuchung der Cluster in den Trainingsdaten wird im Folgenden wieder auf die Projektionsdarstellung zurückgegriffen. Für die Darstellung der Trainingsdaten im Sammons-Mapping werden die Labels aus Abbildung 5.2 verwendet. Normfarben aus dem Farbsystem 3D-Master für die keine Datensätze akquiriert werden konnten, bleiben ohne Label. Für das Farbsystem classical liegen für jede Normfarbe des Farbsystems Trainingsdaten vor. Betrachtet man die zum Training ausgewählten Datensätze anhand der Projektion, so bestätigt sich die bereits beschriebene geringe Qualität der Trainingsdaten hinsichtlich der Trennbarkeit der Farbklassen. Sowohl im Farbsystem 3D-Master (Abbildung 5.3) als auch im Farbsystem classical (Abbildung 5.4) lassen sich so gut wie keine trennbaren Cluster finden. Hierbei ist anzumerken, dass die Darstellung im Sammons-Mapping teilweise etwas verzerrt sein kann, jedoch empfiehlt sich die Bewertung der Trainingsdaten und das Herausfiltern von offensichtlich falschen und inkonsistenten Datensätzen. Dadurch kann die Qualität der Trainingsdaten und der damit trainierten Klassifikationsverfahren optimiert werden.

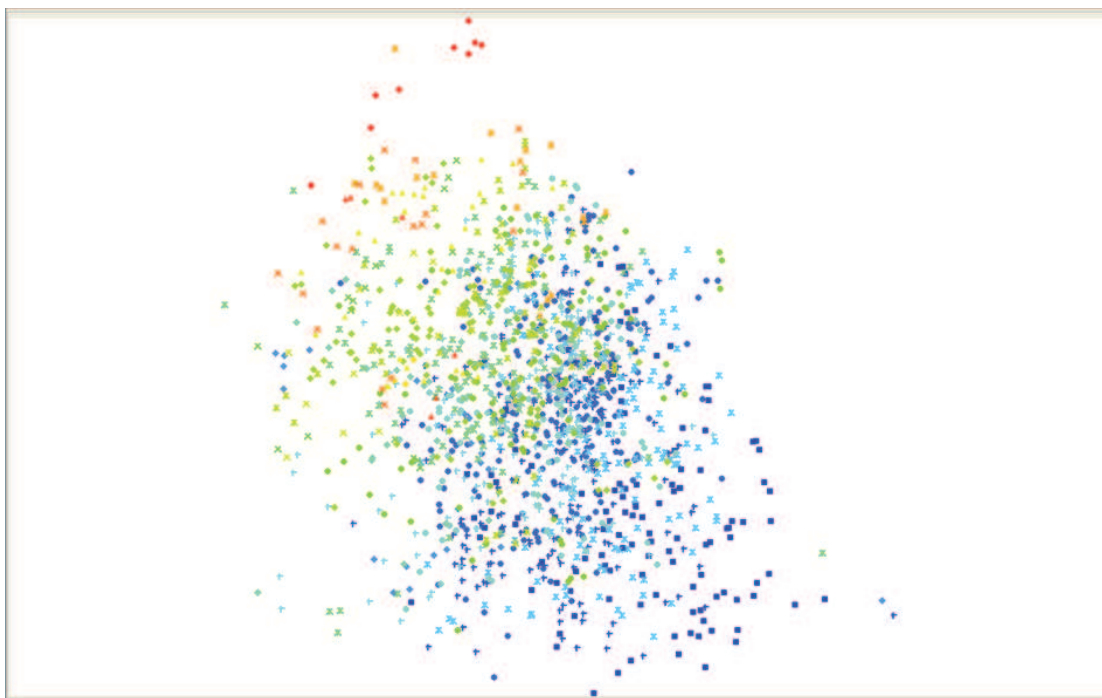


Abbildung 5.3: Sammons-Mapping der ausgewählten Trainingsdaten von dem Farbsystem 3D-Master

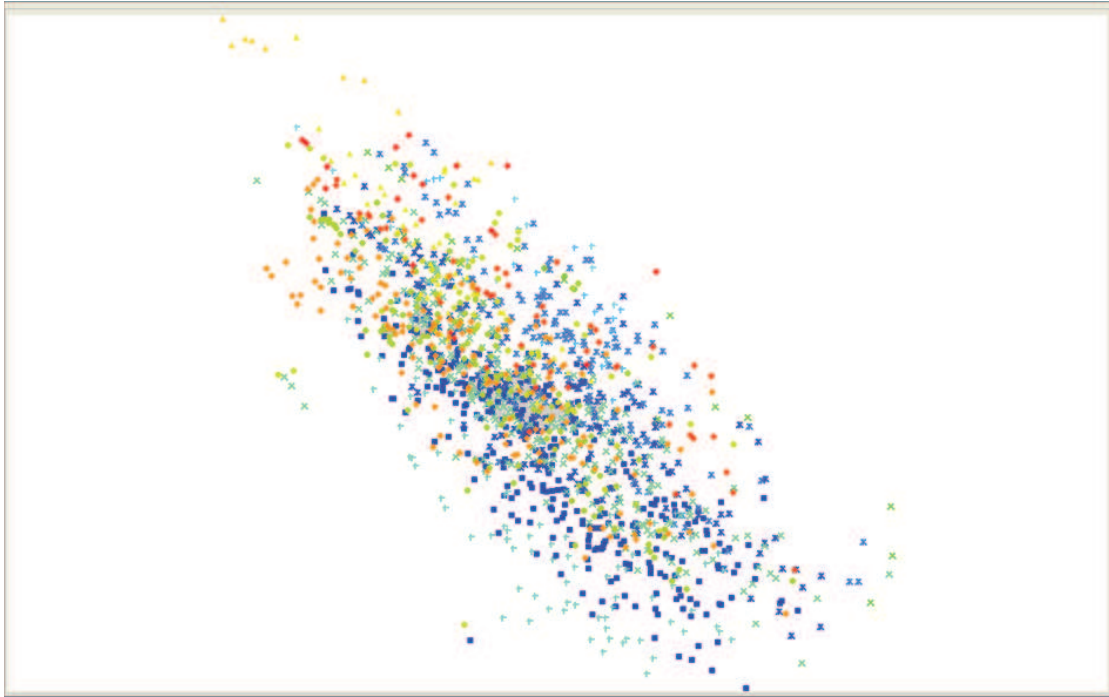


Abbildung 5.4: Sammons-Mapping der ausgewählten Trainingsdaten von dem Farbsystem classical

In den Voruntersuchungen hat sich die Methode des Majority-Filtering am geeignetsten erwiesen, um ungeeignete Trainingsdaten teilweise herauszufiltern. Aus diesem Grund wird das Majority-Filtering auf die ermittelten Trainingsdaten angewendet. Entsprechend den Voruntersuchungen wird die Majority-Filtering Methode mit den drei am geeignetsten erscheinenden Klassifikationsverfahren als Filter eingesetzt (künstliche Neuronale Netze, Logit Boost sowie Support Vector Machines). Für jeden dieser “Filter” wird ein auf Null und Eins normierter Prognosewert berechnet. Ausgeschleust werden alle Datensätze, bei denen der gemittelte Prognosewert kleiner als 0,4 beträgt. Nach Anwendung der Methode zeigt die Visualisierung im Sammons-Mapping eine deutlich verbesserte Klassentrennung (Abbildung 5.5 und 5.6).

Allerdings wird diese bessere Trennung auf Kosten des Ausschleusens einer großen Anzahl an Trainingsdaten erkauft. So bleiben bei den zur Verfügung stehenden Daten des 3D-Master Systems 415 Datensätze und bei den Farben vom System classical 412 Datensätze übrig. Dies entspricht nur knapp einem Drittel der ursprünglich 1533 Datensätze. Jedoch sind für das Training von Klassifikatoren in der Regel wenige gute und repräsentative Datensätze besser geeignet als mehrere inkonsistente Datensätze mit überlappenden Bereichen. Hierbei ist zudem anzumerken, dass die Schwelle für das Herausfiltern der Daten beim Majority-Filtering herabgesetzt werden kann, so dass bei Bedarf mehr Trainingsdaten zur Verfügung stehen.

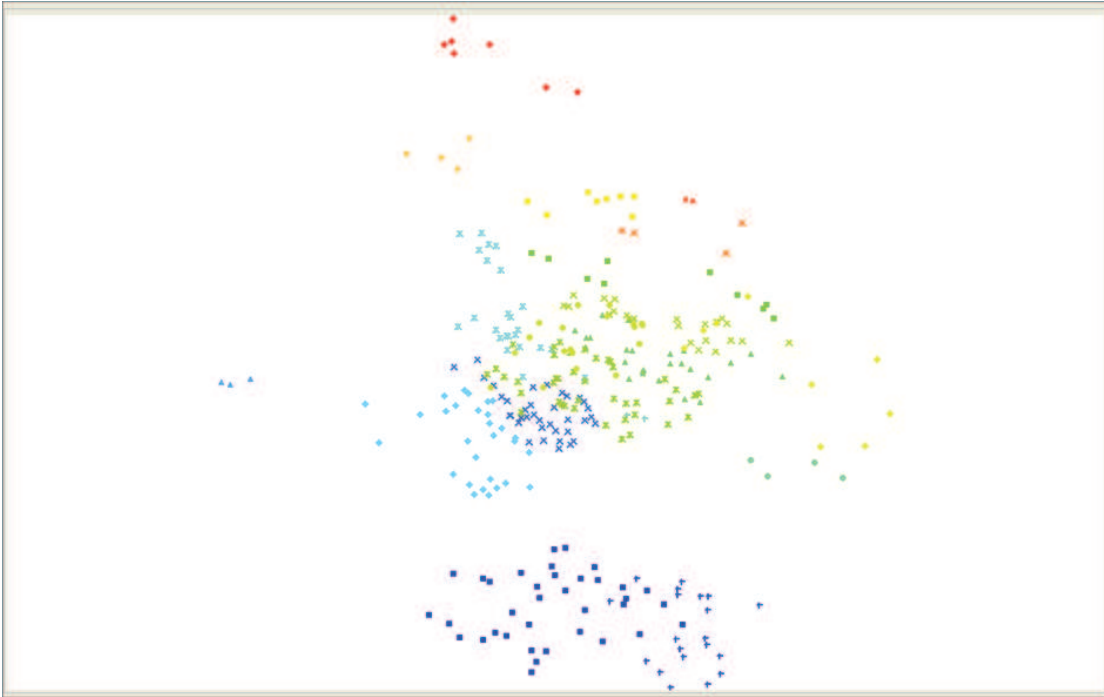


Abbildung 5.5: Sammons-Mapping der durch die Majority-Filtering reduzierten Trainingsdaten von dem Farbsystem 3D-Master

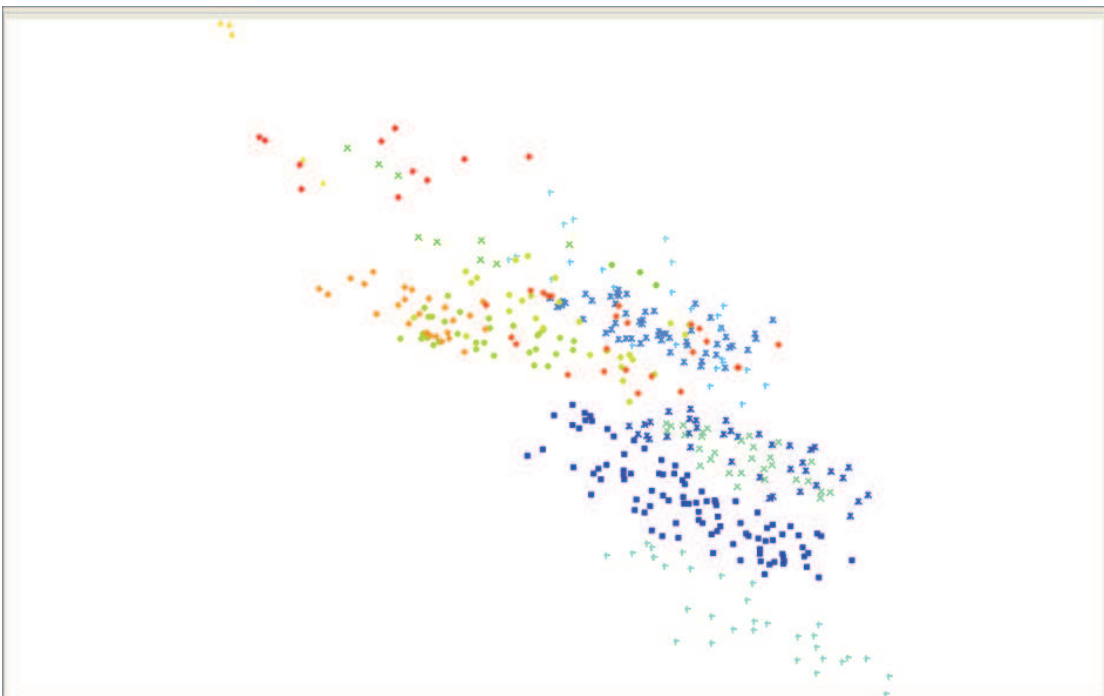


Abbildung 5.6: Sammons-Mapping der durch die Majority-Filtering reduzierten Trainingsdaten von dem Farbsystem classical

5.2 Anwendung der ausgewählten Verfahren

Nach der Optimierung der Trainingsdaten liegt der nächste Schritt in der Anwendung der am besten geeigneten Klassifikationsmethoden. Hierfür werden die diskutierten Verfahren mit den Trainingsdaten trainiert. Im Anschluss werden die Validierungsdatensätze mit den trainierten Modellen klassifiziert und die im vorigen Kapitel vorgestellten Kriterien zur Bewertung der Verfahren berechnet. Als mögliche Verfahren wurden sechs Klassifikationsmethoden sowie drei verschiedene Klassifikationsansätze vorgestellt und umgesetzt. Insgesamt werden somit 18 Verfahren hinsichtlich der vier Gütekriterien bewertet und verglichen.

Ergebnisse bisheriger Verfahren

Für die bisherige Klassifikation im VITA Easyshade wurden lineare Klassifikatoren und einfache Regressionsmodelle genutzt. Die Spektren wurden hierfür in deren LCH-Werte umgerechnet und anschließend mit gewichteten Schwellwerten klassifiziert. Mit dem bisher realisierten Verfahren ist keine Berechnung der zweitbesten Farbe möglich. Insofern wird eine vergleichende Auswertung mit den neuen Verfahren nur mit den Gütekriterien G1, G2 und G3 (Abschnitt 4.6) durchgeführt. Für diese Gütekriterien wurden folgende Ergebnisse für die gewählten Validierungsdaten erzielt:

bisherige Verfahren	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
3D-Master	12,5	88,4	13	-
classical	22,8	56,7	32	-

Tabelle 5.1: Ergebnisse bisheriger Methoden zur Klassifikation der Farbspektren

Ergebnisse der Validierung der Support Vector Machines-Methode bei unterschiedlichen Klassifikationsansätzen

Für die SVM-Methode wurde ein Gauß-Kernel mit Schlupfvariablen zur automatischen Anpassung verwendet. Damit wurden folgende Ergebnisse erzielt:

3D-Master	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	22,4	95,1	59	94,9
Klassenspezifische Modelle	16,0	85,7	55	72,8
Mehrstufige Klassifikation	21,9	92,8	64	95,1

Tabelle 5.2: Ergebnisse der SVM-Methode für die 3D-Master Datensätze

classical	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	24,7	74,5	70	71,0
Klassenspezifische Modelle	24,2	72,1	66	81,0
Mehrstufige Klassifikation	22,5	71,5	69	81,2

Tabelle 5.3: Ergebnisse der SVM-Methode für die classical Datensätze

Ergebnisse der Validierung der Naive Bayes-Methode bei unterschiedlichen Klassifikationsansätzen

Für die NBK-Methode wurden die bedingten Wahrscheinlichkeiten anhand der Trainingsdaten über Schätzverfahren ermittelt. Weitere Parameter müssen nicht vorgegeben werden. Die folgende Ergebnisse wurden erzielt:

3D-Master	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	12,1	90,2	45	97,4
Klassenspezifische Modelle	12,8	91,4	44	94,5
Mehrstufige Klassifikation	12,5	91,7	41	93,8

Tabelle 5.4: Ergebnisse der NBK-Methode für die 3D-Master Datensätze

classical	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	13,0	57,7	65	59,6
Klassenspezifische Modelle	22,1	59,3	47	74,6
Mehrstufige Klassifikation	31,8	69,5	53	81,4

Tabelle 5.5: Ergebnisse der NBK-Methode für die classical Datensätze

Ergebnisse der Validierung der Logit Boost-Methode bei unterschiedlichen Klassifikationsansätzen

Für die LB-Methode wurden über Regressionsverfahren die bedingten Wahrscheinlichkeiten der einzelnen Unterklassifikatoren ermittelt, weitere Parametrierungen sind nicht notwendig. Die folgende Ergebnisse wurden erzielt:

3D-Master	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	18,0	94,0	51	94,4
Klassenspezifische Modelle	17,2	88,9	42	78,7
Mehrstufige Klassifikation	24,2	94,4	55	88,8

Tabelle 5.6: Ergebnisse der LB-Methode für die 3D-Master Datensätze

classical	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	25,4	69,8	56	41,5
Klassenspezifische Modelle	26,6	76,2	63	79,3
Mehrstufige Klassifikation	31,6	79,4	73	84,9

Tabelle 5.7: Ergebnisse der LB-Methode für die classical Masters Datensätze

Ergebnisse der Validierung der künstlichen Neuronale Netze-Methode bei unterschiedlichen Klassifikationsansätzen

Für die KNN-Methode wurden Feed-Forward Netze mit einem Back-Propagation Training-Algorithmus verwendet. Das Modell besteht aus Eingangs- und Ausgangsschicht sowie zwei verdeckten Schichten. Die folgenden Ergebnisse wurden erzielt:

3D-Master	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	15,8	96,2	45	97,7
Klassenspezifische Modelle	21,9	96,2	54	96,0
Mehrstufige Klassifikation	24,3	96,6	71	93,8

Tabelle 5.8: Ergebnisse der KNN-Methode für die 3D-Master Datensätze

classical	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	25,9	73,8	51	88,1
Klassenspezifische Modelle	25,9	56,3	51	43,0
Mehrstufige Klassifikation	27,2	75,8	71	93,0

Tabelle 5.9: Ergebnisse der KNN-Methode für die classical Datensätze

Ergebnisse der Validierung der Learning Vector Quantization-Methode bei unterschiedlichen Klassifikationsansätzen

Für die LVQ-Methode wurden $5 * N$ Stützvektoren verwendet, wobei N die Anzahl der Klassen bzw. Labels darstellt. Die Stützvektoren wurden zufällig initialisiert. Für die Klassifikation mussten mindestens zwei Stützvektoren nächste Nachbarn zu dem zu klassifizierenden Datenvektor sein. Als Abstandsmaß wurde der Euklidische Abstand verwendet. Die folgenden Ergebnisse wurden erzielt:

3D-Master	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	15,7	91,2	62	87,6
Klassenspezifische Modelle	6,5	72,6	22	94,1
Mehrstufige Klassifikation	17,1	92,5	58	95,7

Tabelle 5.10: Ergebnisse der LVQ-Methode für die 3D-Master Datensätze

classical	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	19,8	71,9	69	74,4
Klassenspezifische Modelle	23,6	68,4	46	90,5
Mehrstufige Klassifikation	24,8	73,5	70	79,1

Tabelle 5.11: Ergebnisse der LVQ-Methode für die classical Datensätze

Ergebnisse der Validierung der General Regression Neuro Fuzzy Network-Methode bei unterschiedlichen Klassifikationsansätzen

Für die GRNFN-Methode wurde eine Standardglockenkurve als Aktivierungsfunktion ohne manuelle Fuzzy-Regeln verwendet, wobei fünf Stützpunkte als nächste Nachbarn gewählt wurden. Die folgenden Ergebnisse wurden erzielt:

3D-Master	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	13,5	82,5	56	68,3
Klassenspezifische Modelle	6,8	69,8	38	52,5
Mehrstufige Klassifikation	19,2	94,1	63	84,7

Tabelle 5.12: Ergebnisse der GRNFN-Methode für die 3D-Master Datensätze

classical	Gütekriterium G1 [%]	Gütekriterium G2 [%]	Gütekriterium G3 [Σ]	Gütekriterium G4 [%]
Ein Modell für alle Klassen	17,4	50,7	63	62,3
Klassenspezifische Modelle	19,1	69,2	68	62,0
Mehrstufige Klassifikation	20,3	71,0	72	86,6

Tabelle 5.13: Ergebnisse der GRNFN-Methode für die classical Datensätze

5.3 Bewertung der Ergebnisse und Fehleranalyse

Bei der Analyse der erzielten Ergebnisse fällt als erstes die vergleichsweise niedrige Klassifikationsquote bei den direkt richtig klassifizierten Datensätzen auf (G1). Dies kann auf die Probleme bei der Ermittlung der Trainings- und Validierungsdaten, wie in Abschnitt 4.2.1 beschrieben, zurückgeführt werden. Speziell kommt hier das Problem zum Tragen, dass die gemessenen Daten von Experten visuell klassifiziert werden müssen und dass hierbei häufig die Farbe nicht exakt bestimmt werden kann (Abbildung 4.8). Da dieser Fakt sowohl bei den Trainingsdaten als auch bei den Validierungsdaten der Fall ist, sind hohe Klassifikationsraten nur sehr begrenzt möglich.

Aus diesem Grund wurde das zweite Gütekriterium G2 eingeführt. Hierbei wird überprüft, ob der Datensatz mit der zu erwartenden Farbe oder einem direkten Nachbarn im Farbraum klassifiziert wurde. Die Klassifikation des richtigen Nachbarn ist für den praktischen

Einsatz in der Regel ein akzeptables Ergebnis. Hierbei liegen die Klassifikationsraten für die 3D-Master Farben deutlich höher, bei mehreren Verfahren deutlich über 90% .

Die Klassifikationsrate für G1 liegt bei den classical Datensätzen im Mittel höher als bei den 3D-Master Farben. Bei G2 liegt die Klassifikationsrate bei den classical Datensätzen jedoch deutlich niedriger. Dies kann wie folgt begründet werden: Die höhere Klassifikationsrate bei den direkt richtig klassifizierten Datensätzen ist darauf zurückzuführen, dass das classical Farbsystem nur über 16 Referenzfarben verfügt, während beim 3D-Master System 29 Referenzfarben vorhanden sind. Somit ist die Möglichkeit einer Fehlklassifikation durch die Experten beim 3D-Master prinzipbedingt höher, was sich auch in den Trainings- und Validierungsdaten widerspiegelt. Die schlechtere Klassifikationsrate bei G2 kann dadurch erklärt werden, dass die direkten Farbnachbarn im classical System nicht gleichmäßig verteilt sind (Abschnitt 2.2). Das 3D-Master System wurde auf Basis des LCH-Farbraums entwickelt und somit sind klare und definierte Nachbarschaftsstrukturen vorhanden. Das classical System entstand im Gegensatz dazu empirisch, wodurch gleichmäßig verteilte Klassen und Abstände im Farbraum nicht vorhanden sind. Dieser Umstand wirkt sich negativ auf G2 aus.

Bezüglich G3 (der Histogrammabweichung) und G4 (der Plausibilitätsprüfung hinsichtlich Nachbarschaften im Farbraum) können keine systematischen Besonderheiten oder Auffälligkeiten zwischen den einzelnen Klassifikationsverfahren oder Farbsystemen festgestellt werden. Aufgrund der unterschiedlichen Nachbarschaftsstrukturen zwischen den Farbsystemen sind die classical Farbsysteme im Allgemeinen bei Kriterium drei etwas schlechter, dafür bei Kriterium vier geringfügig besser. Für die Gesamtauswahl eines Klassifikationsverfahren können diese marginalen Unterschiede jedoch ignoriert werden.

Betrachtet man nun die Ergebnisse, welche die Verfahren hinsichtlich der Bewertungskriterien die besten Ergebnisse erreicht haben, so erzielen die mehrstufigen Verfahren die besten Ergebnisse, während die beiden anderen Verfahren im Mittel deutlich schlechter abschneiden. Hierbei muss jedoch auch festgestellt werden, dass die Ergebnisse teilweise recht eng zusammen liegen und auch potenziell bessere und schlechtere Verfahren bei einzelnen Kriterien Ausreisser nach unten oder oben aufweisen können. Mögliche Ursachen hierfür sind beispielsweise die Abhängigkeit der Startinitialisierungen einzelner Methoden oder die zufällige Aufteilung der Trainingsdaten (Abschnitt 5.2).

Hinsichtlich der Klassifikationsverfahren wird deutlich, dass die Methoden KNN, SVM und LB am besten abschneiden. LVQ, NBK und GRNF liefern nur vereinzelt entsprechend gute Ergebnisse und sind deshalb, trotz einiger Ausnahmen bei einzelnen Gütekriterien, generell

als schlechter zu bewerten. Die Gründe hierfür wurden bereits in Abschnitt 4.4 dargestellt und decken sich mit den Ergebnissen der Validierung.

Die Auswahl des am besten geeigneten Verfahrens erfolgt über ein Punktesystem für die ermittelten Gütekriterien. Zur Ermittlung der Punktezahlen für jedes Verfahren wird hierbei das Ergebnis für jedes Kriterium wie folgt auf Werte zwischen Null und Eins normiert:

$$P_{j\text{ Norm}} = \frac{P_j}{P_{Max}}$$

Zur Ermittlung eines geeigneten Normierungswertes P_{Max} wird das in den Klassifikationsergebnissen mögliche Maximum von P_j gewählt. Die Gesamtpunktzahl setzt sich anhand der Summe der Einzelpunktzahlen $P_{j\text{ Norm}}$ zusammen. Hierbei wird für jedes Verfahren der Punktwert für die Ergebnisse jeweils von classical und 3D-Master ermittelt, wobei insgesamt ein theoretisches Maximum von acht Punkten möglich ist:

Anhand der Bewertung zeigt sich deutlich, dass von den Klassifikationsansätzen die mehrstufige Klassifikation die besten Ergebnisse liefert und die ersten drei Plätze beinhaltet. Wie in Abschnitt 4.5 beschrieben ist dieser Ansatz im Prinzip die Nachbildung des Vorgehens bei der manuellen und visuellen Klassifikation im 3D-Master Farbsystem. Hinsichtlich der Klassifikationsmethoden bestätigen die erreichten Punktzahlen, dass KNN, LB und SVM für die Aufgabe die geeignetsten Methoden sind.

Das beste Klassifikationsverfahren für Zahnfarben auf Basis spektraler Daten sind nach den definierten Kriterien künstliche Neuronale Netze mit mehrstufigem Klassifikationsansatz. Das Verfahren liegt in allen Gütekriterien im oberen Wertungsbereich. Dies zeigt, dass das Verfahren eine gute Generalisierungsfähigkeit, ein plausibles Modell und eine gute Klassifikationsquote aufweist. Verglichen mit dem Verfahren, welches im VITA Easyshade bisher genutzt wurden, erreicht das KNN mit mehrstufigem Klassifikationsansatz deutlich bessere Ergebnisse. Die Gesamtpunktzahl liegt hierbei für die vergleichbaren Gütekriterien G1 bis G3 im Mittel bei knapp über 40%, verglichen mit der Punktzahl des bisherigen Ansatzes. Weiterhin soll hier noch einmal auf die besonders hohe Akzeptanz der KNNs in der Praxis hingewiesen werden, da es das mit Abstand am weitesten verbreitete Verfahren im industriellen Einsatz ist. Somit erfüllt das Verfahren alle Voraussetzungen zur Lösung der hier vorliegenden Aufgabe.

Ein Modell für alle Klassen	3D-Master				classical				Gesamt
	G1	G2	G3	G4	G1	G2	G3	G4	
SVM	0,64	0,90	0,77	0,88	0,71	0,70	0,91	0,66	6,16
NBK	0,35	0,85	0,57	0,91	0,37	0,54	0,84	0,55	4,98
LB	0,51	0,89	0,66	0,88	0,73	0,66	0,73	0,39	5,43
KNN	0,45	0,91	0,58	0,91	0,74	0,69	0,66	0,82	5,76
LVQ	0,45	0,86	0,79	0,81	0,57	0,68	0,89	0,69	5,74
GRNFN	0,39	0,78	0,73	0,64	0,50	0,48	0,82	0,58	4,90
bisheriges Verfahren	0,36	0,83	0,17	-	0,65	0,53	0,41	-	-

Farbspezifisches Modell	3D-Master				classical				Gesamt
	G1	G2	G3	G4	G1	G2	G3	G4	
SVM	0,46	0,81	0,71	0,68	0,69	0,68	0,85	0,75	5,62
NBK	0,37	0,86	0,57	0,88	0,63	0,56	0,61	0,69	5,17
LB	0,49	0,84	0,54	0,73	0,76	0,72	0,82	0,74	5,63
KNN	0,63	0,91	0,70	0,89	0,74	0,53	0,66	0,40	5,46
LVQ	0,19	0,68	0,28	0,88	0,68	0,64	0,59	0,84	4,78
GRNFN	0,20	0,66	0,49	0,49	0,55	0,65	0,87	0,58	4,48

Mehrstufige Klassifikation	3D-Master				classical				Gesamt
	G1	G2	G3	G4	G1	G2	G3	G4	
SVM	0,63	0,87	0,83	0,88	0,64	0,67	0,89	0,76	6,18
NBK	0,36	0,86	0,52	0,87	0,91	0,65	0,69	0,76	5,62
LB	0,69	0,89	0,71	0,83	0,91	0,75	0,95	0,79	6,51
KNN	0,70	0,91	0,91	0,87	0,78	0,71	0,91	0,86	6,65
LVQ	0,49	0,87	0,75	0,89	0,71	0,69	0,90	0,74	6,04
GRNFN	0,55	0,89	0,81	0,79	0,58	0,67	0,92	0,81	6,01

Tabelle 5.14: Gesamtbewertung der Klassifikationsverfahren.

5.4 Überprüfung der Anforderungen und Schlußfolgerungen

Für eine abschließende Bewertung des Verfahrens müssen die ermittelten Anforderungen aus Abschnitt 2.3 überprüft werden. Als Lösung wird für die Hardwarebasis mit dem VITA Easyshade ein handlicher Smart-Sensor zur Aufnahme spektraler Daten eingesetzt, welcher alle oben beschriebenen Voraussetzungen erfüllt und bereits von Vita vertrieben wird. Zur Auswertung der Daten kommt ein neues Verfahren basierend auf künstlichen Neuronalen Netzen zum Einsatz, welches eine mehrstufige Klassifikation zur Ermittlung der passenden Farbe durchführt. Unterstützt wird das Verfahren durch Module zur Bewertung der

Trainingsdaten sowie eine Plausibilitätsprüfung der Ergebnisse. Folgende Anforderungen wurden an das Verfahren gestellt:

Plausible Ermittlung der Zahnfarbe (notwendige Anforderung)

Es wurde gefordert, dass das System eine zuverlässige Ermittlung der Zahnfarbe ermöglicht, welche Patienten generell zufrieden stellt. Aus diesem Grund wurde als Bewertungskriterien festgelegt, dass das Verfahren für die Gütekriterien G2 und G4 wie oben festgelegt mindestens 85% erreichen muss. Das realisierte Verfahren liefert für die 3D-Master Farben für G2 Werte von deutlich über 90%. Bei den classical Farben liegt die Klassifikationsquote unter der geforderten Schwelle. Dies wurde bereits durch ungleiche Abstandsverteilung der Referenzfarben im Farbraum begründet und kann prinzipbedingt kaum verbessert werden. Somit kann festgestellt werden, dass diese Anforderung im Mittel über beide Farbräume erfüllt ist. Betrachtet man nur das Farbsystem 3D-Master, wird die Anforderung sehr gut erfüllt. Hinsichtlich G4 wird die Anforderung für beide Farbräume problemlos erfüllt.

Robuste Ermittlung der Zahnfarbe (notwendige Anforderung)

Der Sensor ist aufgrund seiner Möglichkeiten zur Messung spektraler Daten im Prinzip gut geeignet, um Spektren der benötigten Qualität zu messen. Jedoch zeigen sich in der praktischen Anwendung besonders bei inhomogenen Zahnoberflächen potenzielle Fehlerquellen, welche von den Algorithmen nur teilweise ausgeglichen werden. Ein weiteres Problem liegt in der Ermittlung von Trainingsdaten, da diese prinzipbedingt fehlerbehaftet sind. Durch die Softwareentwicklung der Module zur Bewertung von Trainingsdaten sowie der Plausibilitätsprüfung der Ergebnisse kann dies zu einem akzeptablen Grad aufgefangen werden. Zudem können Anwender durch Schulungen hinsichtlich potenzieller Messstörungen sensibilisiert werden. Aus diesem Grund kann diese Anforderung als erfüllt angesehen werden.

Reproduzierbare Ergebnisse (notwendige Anforderung)

Wie in Abschnitt 4.2.1 dargestellt, weist der Sensor bei einer entsprechenden Bedienung eine robuste Wiederholgenauigkeit auf. Die Auswertalgorithmen verfügen über ein fest codiertes Modell, so dass gleiche Eingangsdaten immer zu gleichen Ausgangsdaten führen. Somit wird die geforderte Quote von 99% Wiederholgenauigkeit auf Seiten des Auswerteverfahrens erreicht. Die Validierung in Abschnitt 5.3 zeigt zudem, dass größere Abweichungen in den Eingangsdaten in der Regel trotzdem zu plausiblen Ergebnissen hinsichtlich der Klassifikation führen.

Mobile einsetzbares Verfahren (notwendige Anforderung)

Das gewählte Sensorsystem ist durch den handlichen Smart-Sensor mobil, benötigt keine Probenvorbereitung und kann in jeder zahnärztlichen Umgebung eingesetzt werden. Auch

Messungen direkt am Patienten und Wiederholmessungen sind jederzeit möglich. Somit kann diese Anforderung als erfüllt angesehen werden.

Sterilisierbarkeit (notwendige Anforderung)

Zur Sterilisation wird für den Sensor eine Schutzfolie verwendet, welche ohne Expertenwissen einfach über den Messkopf gezogen werden kann. Da dies vor der Kalibrierung des Sensors durchgeführt wird, beeinträchtigt die Folie das Messergebnis nicht. Somit kann das Gerät steril gehalten werden, weshalb diese Anforderung erfüllt ist.

Einklang mit Medizinproduktegesetz (notwendige Anforderung)

Auch wenn das Medizinproduktegesetz im Rahmen der Arbeit nicht näher betrachtet wurde, besteht keine generelle Unverträglichkeit, welche gegen eine Zulassung des gewählten Verfahrens spricht. Aus diesem Grund kann die Anforderung als erfüllt angesehen werden.

Kostengünstige Realisierung des Verfahrens (wünschenswerte Anforderung)

Die gewählte Sensorplattform ist bereits kostendeckend im praktischen Einsatz. Die entwickelte Algorithmik stellt keine neuen Anforderungen an die bestehenden Komponenten, so dass keine zusätzlichen Hardwarekosten entstehen. Der Aufwand für die Auswertesoftware ist einmalig, wodurch auch hier keine signifikanten Kostensteigerungen beim gesamten System zu erwarten sind. Aus diesem Grund kann der Sensor weiterhin zu einem marktgerechten Preis vermarktet werden. Die Anforderung ist somit ebenfalls erfüllt.

Einfache und intuitive Bedienung (wünschenswerte Anforderung)

Bezüglich der Klassifizierung erlaubt das Verfahren eine sehr einfache Durchführung von Farbbestimmungen, welche vollautomatisch erfolgen. Der Anwender muss lediglich den Sensorkopf an den zu prüfenden Zahn halten und erhält auf Knopfdruck die klassifizierte Farbe. Für den Nutzer ist keine weitere Interpretation der Messergebnisse notwendig, wodurch kein Spielraum für Fehlinterpretationen gegeben ist. Auch hinsichtlich des Trainings ist das Verfahren nutzerfreundlich und kann mit einem Minimum an Expertenwissen einfach erweitert werden. Hierfür muss lediglich eine neue Datei mit Trainingsdaten angelegt werden. Auf Knopfdruck werden dann die KNN Modelle für die mehrstufige Klassifikation erstellt und als Netzstrukturdatei mit definierten Format gespeichert (Kapitel 6). Somit kann diese Anforderung als erfüllt angesehen werden.

Kurze Berechnungsdauer (wünschenswerte Anforderung)

Die Berechnung der Ergebnisse basiert bei dem entwickelten Verfahren nur auf einfachen mathematischen Standardoperatoren. Die Berechnung und Ergebnisanzeige ist somit in weniger als den geforderten zwei Sekunden möglich. Die Berechnung des zu generierenden Modells für die Klassifikation ist weit rechenintensiver und dauert mehrere Minuten, muss

aber nur einmal pro Systemadaption durchgeführt werden. Da dies in der Regel auf einem externen Rechner und vor der Auslieferung des Smart-Sensors geschieht, ist die Trainingszeit für diese Anforderung nicht relevant. Die Anforderung wird somit erfüllt.

Funktion für sämtliche Farbräume und Farbsysteme (wünschenswerte Anforderung)

Das entwickelte Gesamtsystem wurde an den gängigen Farbsystem classical und 3D-Master validiert, es funktioniert aber im Prinzip für jede Art von ähnlichen Farbsystemen. Voraussetzungen sind lediglich eine ausreichende Menge an Trainingsdaten sowie deren Einteilung für die Vorklassifikation. Diese Anforderung kann deshalb als erfüllt angesehen werden.

Erweiterbar für neue Farben (wünschenswerte Anforderung)

Das entwickelte Verfahren kann um eine beliebige Anzahl neuer Farbmuster erweitert werden. Auch hierfür müssen nur entsprechende Trainingsdaten bereit gestellt werden. Allerdings muss das gesamte Modell neu trainiert und validiert werden. Die alten Modelle können nicht mehr verwendet werden. Dies muss jedoch ebenfalls nur einmal auf einem externen Rechner durchgeführt werden. Zudem können die bereits ausgelieferten Sensoren durch ein einfaches Update mit den neuen Modelldaten ausgestattet werden. Somit ist die Anforderung erfüllt.

Zusammenfassend kann festgestellt werden, dass das gewählte und umgesetzte Verfahren zur mehrstufigen Klassifikation in Kombination mit dem bestehenden Smart Sensor VITA Easyshade alle notwendigen Anforderungen in ausreichendem Maße erfüllt und somit als uneingeschränkt geeignet angesehen wird.

Kapitel 6

Systemseitige Realisierung der Funktionsmuster

Dieses Kapitel beschreibt die Realisierung des ausgewählten Verfahrens in Form von Funktionsmustern. Die Realisierung besteht im Prinzip aus zwei Modulen. Das erste Modul liefert sämtliche Funktionalitäten für das Modelltraining sowie der Behandlung der Trainingsdaten. Das zweite Modul dient zur Klassifikation der Spektren sowie der Plausibilitätsprüfung. Zudem wurde die Integration des Moduls in den VITA Easyshade Sensor durchgeführt. Tabelle 6.1 gibt einen Überblick über die Teilfunktionalitäten, Definitionen und Schnittstellen, welche im Rahmen der Arbeit definiert und entwickelt wurden:

6.1 Modul zum Training der Klassifikationsverfahren

Das folgende Abschnitt beschreibt die Funktionsweise des Moduls zum Training der mehrstufigen Klassifikation.

6.1.1 Beschreibung des Trainingsmoduls

Das Trainingsmodul dient zur Bewertung und Auswahl der Trainingsdaten sowie dem Erstellen und Training des Modells auf Basis einer mehrstufigen Klassifikation mit künstlichen Neuronalen Netzen. Zuerst werden bei Bedarf die Trainingsdaten mit Hilfe des Majority-Filtering auf fehlerhafte und inkonsistente Daten hin untersucht werden. Basierend darauf werden fehlerhafte Daten ausgeschleust. Hierfür muss der Nutzer lediglich eine Schwelle für den mittleren Prognosewert der Filter vorgeben (Abschnitt 4.3.3). Die finalen Trainingsdaten werden automatisch als eigene Datei ausgegeben.

Bezeichnung	Erklärung
Trainingsmodul	Softwaremodul, welches sämtliche Funktionalitäten zur Bewertung von Trainingsdaten und zum Training der Modelle bereitstellt.
Klassifikationsmodul	Softwaremodul zur Klassifikation der Spektren anhand der zuvor trainierter Modelle und Ausgabe der Klassifikationsergebnisse.
Trainingsdaten	Datei, welche spektrale Datensätze mit bekannter Klassifikation zum Training der Modelle beinhaltet.
Lerndaten	Die Lerndaten sind ein automatisch ausgewählter Teil der Trainingsdaten, anhand derer das Modell trainiert wird.
Verifikationsdaten	Die Verifikationsdaten sind ein automatisch ausgewählter Teil der Trainingsdaten, anhand derer während des Modelltrainings die Generalisierungsfähigkeit und Leistung des aktuellen Modells überprüft wird.
Validierungsdaten	Datei, welche spektrale Datensätze mit bekannter Klassifikation beinhaltet, anhand derer nach abgeschlossenem Training die Modellgüte verifiziert wird.
Trainingsparameter	Einstellbare Parameter für das Modelltraining (Ausschlußkriterium für Trainingsdatensätze, Modellstruktur der KNNs, Lernverfahren, Abbruchkriterien usw.).
Modellparameter für die Teilmodelle	Datei mit Strukturdaten und Gewichtungsparametern für ein Teilmodell
Modellgüte Training	Protokolldatei mit der Güte des trainierten Modells anhand der Abweichungen von Lern- und Verifikationsdaten.
Modellgüte Validierung	Protokolldatei mit der Güte des trainierten Modells anhand der Klassifikation der Validierungsdaten.
Spektrale Daten	Aktuelle Messwerte des VITA Easyshade.
Ergebnis	Ausgabe des Klassifikationsergebnisses, bestehend aus den zwei wahrscheinlichsten Farben sowie deren Prognosewerten.
Nachbarschaftsmatrix	Nachbarschaften der Normfarben für die Plausibilitätsprüfung

Tabelle 6.1: Übersicht über Softwaremodule, Dateiformate und benötigte Parameter für die realisierte Gesamtfunktionalität

Anhand der ermittelten Trainingsdaten erfolgt das Modelltraining. Hierbei werden zuerst die Trainingsdaten automatisch normiert und anschließend automatisch in Lern- und Verifikationsdaten aufgeteilt. Während des iterativen Trainings werden die Gewichte und Übergangsfunktionen zwischen den Neuronen der KNNs ermittelt. Es werden hierbei sequenziell sämtliche benötigten freien Parameter für die KNNs für die Vor- und Detailklassifikation berechnet und gespeichert. Zudem können optional verschiedene Parameter für das Netztraining vorgegeben werden. Für die trainierten Modelle werden definierte Kriterien berechnet, welche die Güte des Trainings quantifizieren. Außerdem können die Modelle nach dem Training mit Validierungsdaten verifiziert werden, um die Modellgüte zu validieren.

Ist die Modellqualität hierbei nicht ausreichend, kann ein Neutraining aller Modelle oder ausgewählter Teilmodelle der Vor- und Detailklassifikation eingeleitet werden. Wichtig zu erwähnen ist, dass hinsichtlich der Nutzerfreundlichkeit sämtliche möglichen Operationen automatisch durchgeführt werden. Der Ablauf des gesamten Modelltrainings ist in Abbildung 6.1 dargestellt.

Folgende Teilfunktionalitäten und Definitionen wurden hierfür realisiert:

- Definierte Datenformate für Trainings- bzw. Validierungsdaten, Bewertungsparameter für das Modelltraining, die Parameter des trainierten Modells sowie die Ergebnisse der Validierung (sämtliche Dateien sollten möglichst flexibel gegenüber Änderungen und Erweiterungen sein)
- Ein Softwaredialog zur Einstellung der Trainingsparameter, aufgeteilt in notwendige und optionale Einstellungen.
- Funktionen zum Ein- und Auslesen der Dateien sowie zur automatischen Verwaltung sämtlicher Daten
- Funktionen zur automatischen Bewertung und Auswahl von Trainingsdaten
- Funktionen zur Aufteilung der Trainingsdaten in Lern- und Verifikationsdaten sowie Auswahl der Daten für Vor- und Detailklassifikation
- Funktion zur automatischen Durchführung des Modelltrainings, bei Bedarf unterteilbar in Einzeltrainings für Vor- und Detailklassifikation

6.1.2 Datenformate

Wie bereits im vorigen Abschnitt erwähnt, werden verschiedene Dateien und Datenformate benötigt. Diese sind wie folgt definiert:

Datei für Trainings- und Validierungsdaten

Diese Dateien werden im Ascii-Format angelegt und beinhalten Datensätze zum Training oder zur Validierung, welche bereits von Experten klassifiziert wurden. Um das Dateiformat möglichst flexibel zu halten, wurden vier Kopfzeilen definiert, in welchen die allgemeine Datenstruktur der Trainingsdaten definiert werden kann:

- Die erste Zeile enthält die Versionsnummer der Datei, um auf allgemeine Änderungen der Struktur reagieren zu können.
- Die zweite Zeile enthält die Anzahl der zu klassifizierenden Kenngrößen, welche in der aktuellen Sensorkonfiguration die spektralen Messwerte an elf Punkten im Wellenlängenspektrum sind.

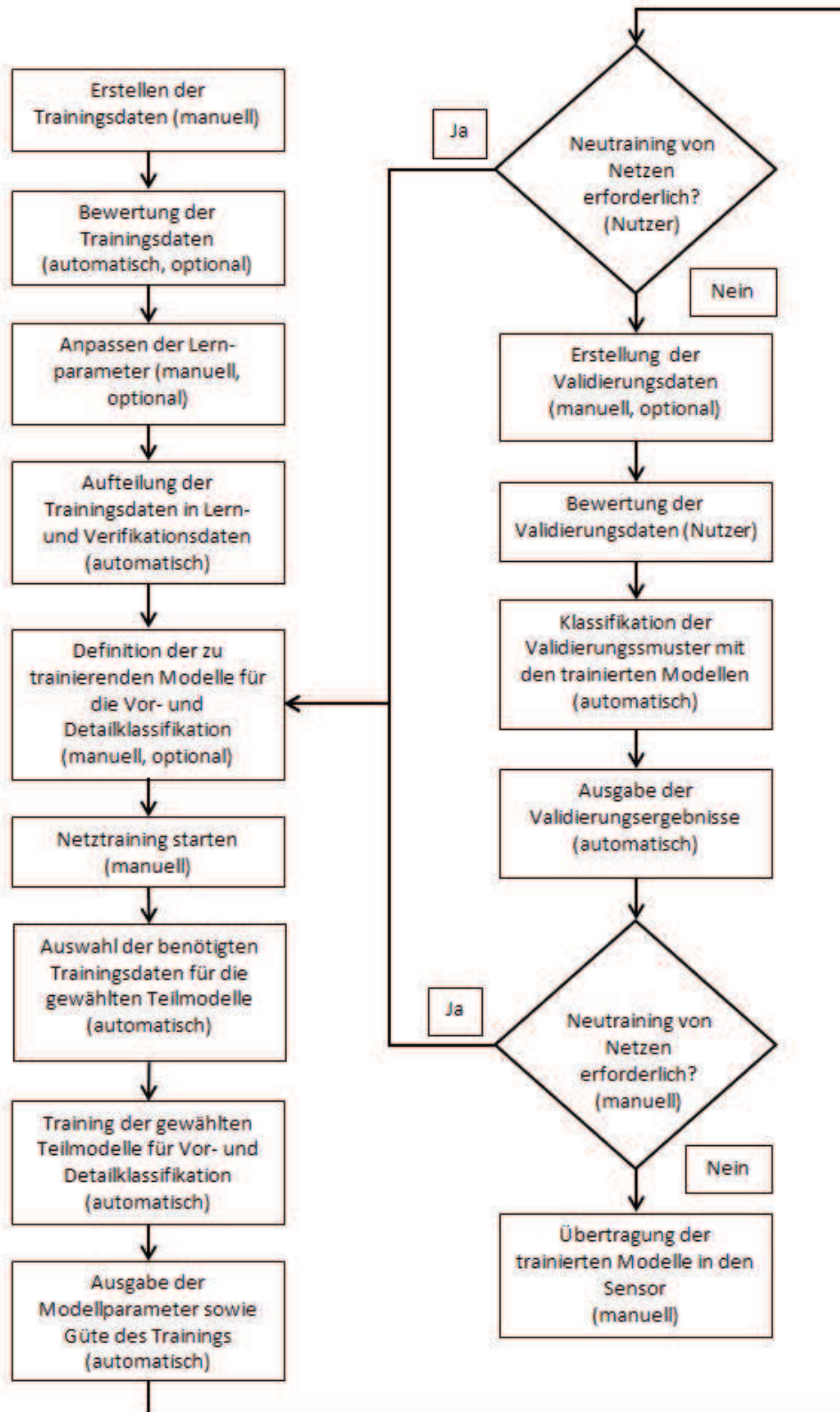


Abbildung 6.1: Ablauf des Netztrainings mit dem realisierten Trainingsmodul

- Die dritte Zeile enthält die Anzahl an Datensätzen, womit interne Kontrollabfragen hinsichtlich der Vollständigkeit der Datensätze möglich sind.
- Die vierte Zeile enthält eine Kennung für den zu nutzenden Farbraum, womit intern kontrolliert werden kann, dass Trainingsdaten, Modelle und Verifikationsdaten zueinander konsistent sind.
- Ab Zeile fünf werden die Kenngrößen und deren Label eingetragen, wobei jeder Wert durch einen Tabulator getrennt ist.
- Kommentare sind überall in der Datei möglich und werden durch das Zeichen “#” eingeleitet.

Durch die Struktur sind die Datendateien flexibel gegenüber Änderungen an der Datenstruktur, wie beispielsweise der Anzahl der Eingangsgrößen. Um dem Benutzer die Zuordnung zu unterschiedlichen Gruppen für die Vorklassifikation zu erleichtern wird für die Labels eine Struktur festgelegt. Vor jedes Label wird eine Zahl geschrieben, welche dessen Zugehörigkeit zu einer Gruppe für die Vorklassifikation angibt. Die Farben vom System 3D-Master sind somit schon in der richtigen Struktur, wenn die Vorklassifikation anhand der Helligkeit erfolgen soll (z.B. 1M1 gehört zu Gruppe 1 während 3L2.5 zur Gruppe 3 gehört). Vor die Bezeichnungen der classical Farben muss manuell eine Gruppenzugehörigkeit für die Vorklassifikation angegeben werden (z.B. bei 1A1 gehört A1 zur Gruppe eins und bei 4A3.5 gehört A3.5 zur Gruppe vier). Das Hinzufügen neuer Farben oder die Integration eines neuen Farbsystems für das Modelltraining kann dadurch von jedem Nutzer mit Hilfe entsprechender Trainingsdaten durchgeführt werden. Änderungen am Dateiformat oder der Software sind dabei nicht nötig.

Datei zur Bewertung des Modelltrainings

Diese Datei wird automatisch nach dem Training eines Teilmodells erstellt und beinhaltet das Gütekriterien des jeweiligen Teilmodell. Hierbei wird in der Datei für Identifikationszwecke das Datum des Trainings und die Kennung des aktuellen Teilmodells geschrieben (z.B. “Vorklassifikation” für das Modell zur Vorklassifikation oder “Detailklassifikation_2M1” für ein Teilmodell zur Detailklassifikation der zweiten Gruppe). Als Bewertungskriterien hinsichtlich der Trainingsgüte werden zum Einen der maximale und der mittlere Fehler beim Modelltraining hinsichtlich der Abweichung zwischen Lern- und Verifikationsdaten in die Datei geschrieben. Zum Anderen wird die Anzahl der Trainingszyklen für das Modell dokumentiert.

Dateien für die Modellparameter der Teilmodelle

Diese Dateien werden als Header-Datei erstellt, damit sie mittels eines Cross-Compilers direkt auf den Mikroprozessor des Sensors geladen werden können. Die Dateien beinhalten sämtliche Parameter zur vollständigen Beschreibung des jeweiligen Modells. Beim hier gewählten mehrstufigen Klassifikationsansatz auf Basis von KNNs müssen die Dateien zum Einen über eine Kennung verfügen, welches Teilmodell sie beinhalten. Dies wird anhand des Dateinamens realisiert (z.B. "Vorklassifikation.h" für das Modell zur Vorklassifikation oder "Detailklassifikation_2M1.h" für ein Teilmodell zur Detailklassifikation der zweiten Gruppe). Die Datei definiert die Netzstruktur sowie die einstellbaren Modellparameter der KNNs wie folgt:

- Die erste Zeile enthält eine ID der Farbklasse, welche in chronologischer Reihenfolge anhand der Labels der Trainingsdaten vergeben wird.
- Die zweite Zeile enthält die Anzahl der Eingangs- und Ausgabewerte für interne Kontrollabfragen.
- Die dritte Zeile enthält die Anzahl der Schichten der KNNs.
- Die vierte Zeile enthält die Verteilung der Neuronen der KNNs als Array (z.B. 11; 6; 3; 1 für ein Netz mit einer Eingangsschicht mit elf Elementen, zwei verdeckten Schichten mit sechs und drei Neuronen sowie einer Ausgabeschicht mit einem Ausgang).
- Die fünfte Zeile enthält die Anzahl der einstellbaren Parameter für die Gewichte der Netze (die Anzahl hängt hierbei von der gewählten Netzstruktur ab).
- Ab Zeile sechs werden die konkret ermittelten Werte für die ermittelten Netzparameter eingetragen.

Durch die Identifikationsmöglichkeit anhand des Dateinamens und der gespeicherten Struktur, welche automatisch aus den Trainingsdaten generiert wird, können auch neue Farben und Farbsysteme sowie fehlende Farben eines bestehenden Farbsystems automatisch behandelt werden, ohne dass Änderungen an der Software nötig sind.

Datei für die Ergebnisse der Validierung

Diese Datei wird automatisch erstellt und beinhaltet die Klassifikationsergebnisse. Hierbei werden für jeden Datensatz der beste und zweitbeste Prognosewert für den jeweiligen Datensatz in die entsprechende Zeile eingefügt.

6.1.3 Parametrierung der Trainingsalgorithmen

Zur Parametrierung der Trainingsalgorithmen wurde im Rahmen der Entwicklung des Trainingsmoduls ein Dialog erstellt, in dem der Nutzer das Modelltraining entsprechend kon-

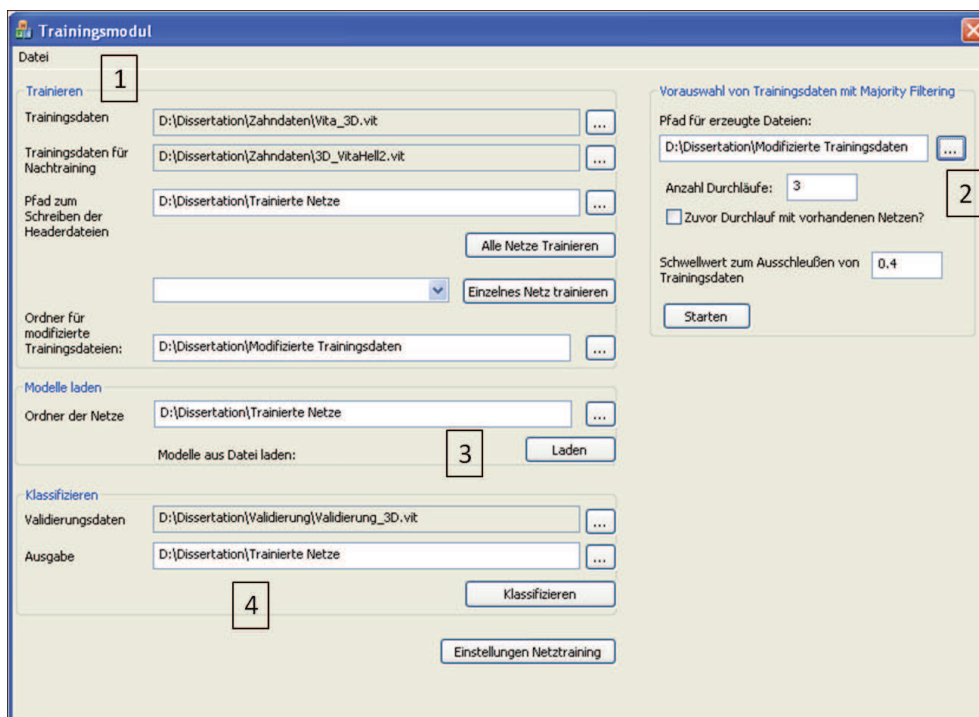


Abbildung 6.2: Dialog zur Konfiguration des Netztrainings

figurieren kann (Abbildung 6.2). Die einzige notwendige Eingabe für das Modelltraining ist der Pfad zu einer Datei mit Trainingsdaten. Auf Wunsch kann der Nutzer auch speziell einzelne Klassen zum Training auswählen oder ergänzende Daten für ein Nachtraining bestehender Modelle auswählen (Abbildung 6.2_Bereich 1). Weiterhin optional kann der Nutzer die Trainingsdaten mit dem Majority-Filtering bewerten lassen. Hier muss die Anzahl der Filterdurchgänge sowie ein Schwellwert zum Ausschleusen der Daten vorgegeben werden (Abbildung 6.2_Bereich 2). Für weiterführende Analysen und Versuche können bereits trainierte Modelle geladen werden (Abbildung 6.2_Bereich 3). Zudem können beliebige Datendateien ausgewählt werden, um die aktuellen Modelle zu validieren (Abbildung 6.2_Bereich 4).

Anzumerken bleibt, dass der Benutzer für ein effizientes Modelltraining lediglich Trainingsdaten auswählen muss. Weitere Einstellungen sind nicht zwingend erforderlich. Fortgeschrittene Nutzer können jedoch Modellparameter optimieren, um das Training auf die jeweilige Problemstellung anzupassen. Hierfür wurde für das Trainingsmodul ein erweiterter Dialog realisiert, welcher in Abbildung 6.3 dargestellt ist. Hier kann zum Einen die Grundstruktur der KNNs vorgegeben werden (Abbildung 6.3_Bereich 1). Zum Anderen können die Lernparameter des Modells eingestellt werden, beispielsweise Abbruchkriterien oder der Anteil an Lern- und Verifikationsdaten (Abbildung 6.3_Bereich 2).

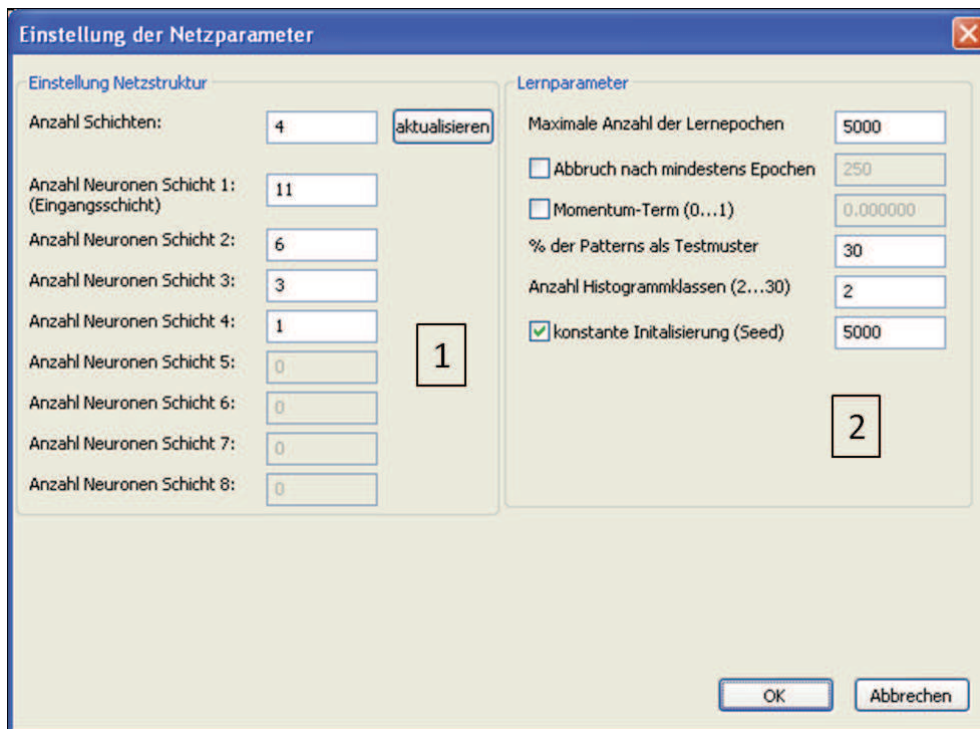


Abbildung 6.3: Dialog zur detaillierten Konfiguration der Modellstruktur und der Lernparameter

6.2 Modul zur automatischen Klassifikation

Das folgende Abschnitt beschreibt die Funktionsweise des Moduls zur mehrstufigen Klassifikation auf einem Mikroprozessor.

6.2.1 Beschreibung des Klassifikationsmoduls

Das Klassifikationsmodul dient zur Klassifikation der Zahnfarbe auf Basis der trainierten Modelle. Hierbei werden die vom Sensor gemessenen Spektren an das Klassifikationsmodul übergeben und dort zweistufig klassifiziert. Hierbei werden für die potenziellen Referenzfarben entsprechende Prognosewerte berechnet, anhand dessen die wahrscheinlichste Zahnfarbe ermittelt und ausgegeben wird. Zusätzlich kann optional die in Abschnitt 4.6 beschriebene Plausibilitätsprüfung durchgeführt werden, anhand derer die Ergebnisse der Klassifikation bewertet werden können. Der prinzipielle Ablauf ist in Abbildung 6.4 dargestellt. Wichtige Randbedingungen sind hierbei, dass sämtliche Funktionalitäten so konzipiert wurden, dass sie auf dem Mikroprozessor des VITA Easyshade genutzt werden können [N.N. 2005a].

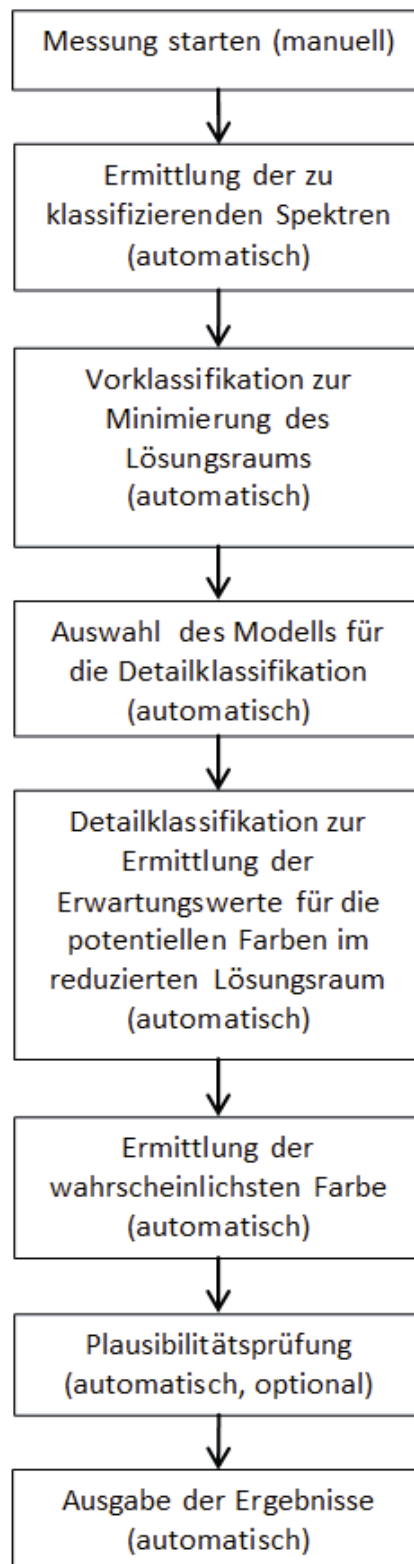


Abbildung 6.4: Ablauf der Klassifikation mit dem realisierten Klassifikationsmodul

Folgende Teilfunktionalitäten und Definitionen wurden hierfür realisiert:

- Definierte Datenformate für spektrale Daten, Modellparameter und Ergebnisausgabe
- Funktion zur modellbasierten Berechnung der benötigten Prognosewerte für die mehrstufige Klassifikation der Zahnfarbe
- Funktion zur optionalen Plausibilitätsprüfung

6.2.2 Datenformate

Wie bereits im vorigen Abschnitt erwähnt, werden auch für die Klassifikation verschiedene Datenformate benötigt. Diese sind wie folgt festgelegt:

Datenstruktur für die spektralen Daten

Diese Datenstruktur ist ein Array mit den gemessenen Intensitätswerten bei den definierten Stützpunkten aus dem Wellenlängenspektrum.

Datenstruktur für die Modellparameter der Teilmodelle

Diese Datenstruktur entspricht der Struktur der Datei für die Modellparameter der Teilmodelle. Da sich diese bereits im entsprechenden Format für den Mikroprozessor des Sensors befinden, werden diese nur noch in entsprechende Arrays für die Modellparameter eingelesen.

Datei für Nachbarschaftsmatrix (optional)

Diese Header-Datei wird manuell angelegt. Sie enthält ein doppeltes Array, welches eine quadratische Matrix repräsentiert. In der Matrix sind sämtliche mögliche Farbpaare des jeweiligen Farbraums abgebildet, indem die Referenzfarben jeweils in der ersten Spalte und Zeile eingetragen sind. Für jedes Farbenpaar muss der Nutzer eines der Elemente $[0, 1, -1]$ in die Matrix einfügen. Hierbei steht das Element $[1]$ für die gleiche Farbe, das Element $[0]$ für eine direkte Nachbarschaft und das Element $[-1]$ für keine direkte Nachbarschaft der jeweiligen Farben. Hierdurch werden für die Plausibilitätsprüfung sämtliche Nachbarschaften der Referenzfarben definiert. Falls der Nutzer keine Nachbarschaftsmatrix vorgibt, wird die Klassifikation ohne Plausibilitätsprüfung durchgeführt. Die entsprechenden Nachbarschaftsmatrizen der beiden gängigen Farbräume 3D-Master und classical wurden im Rahmen der Arbeit integriert.

Datenstruktur für den Ergebnisvektor

Diese Datenstruktur ist ein Array aus Strukturen, welche die Bezeichnung der Referenzfarbe und einen entsprechenden Prognosewert beinhalten. Der Ergebnisvektor enthält die

drei wahrscheinlichsten Normfarben und deren Prognosewerte und ist entsprechend der Wahrscheinlichkeiten sortiert. Anhand dieses Vektors wird das entsprechende Modell zur Detailklassifikation basierend auf der Vorklassifikation ausgewählt. Weiterhin wird die optionale Plausibilitätsprüfung durchgeführt und entweder das wahrscheinlichste Ergebnis der Klassifikation der Messdaten oder ein ungültiges Ergebnis ausgegeben, welches auf den Fehler hinweist.

6.2.3 Verbesserung der Ergebnisqualität

Um die Ergebnisqualität zu erhöhen und Fehlbedienungen oder Störungen zu vermeiden, wurden neben der bereits integrierten Plausibilitätsprüfung folgende Funktionen zur Sicherheit und zur erweiterten Auswertung der Ergebnisse realisiert:

Überprüfung der Modelldateien

Diese Funktion überprüft, ob sämtliche Modelldateien vorhanden sind und richtig geladen werden konnten. Sollten Modelldateien zu bestimmten Farben fehlen, beispielsweise weil für diese keine oder zu wenig Trainingsdaten vorliegen, so werden diese für die Klassifikation nicht berücksichtigt. Da bei den hier vorhandenen Trainingsdaten im Farbsystem 3D-Master nicht zu allen Referenzfarben entsprechende Datensätze vorhanden waren (beispielsweise für die Helligkeitsstufe 5), wurde im VITA Easyshade die Klassifikation durch ein Mischsystem erweitert. Für Farben ohne Referenzdaten wird auf die alten Klassifikationsmethoden zurückgegriffen. Sobald genügend Trainingsdaten für alle Referenzfarben verfügbar sind, kann die Klassifikation dann komplett mit der neuen Systemtechnik erfolgen.

Überprüfung des Wertebereichs der Spektren

Diese Funktion überprüft, ob die spektralen Eingangsdaten in einem plausiblen Wertebereich liegen. Hierbei werden anhand der Trainingsdaten entsprechende Schwellwerte für jede Wellenlänge berechnet. Hierfür werden die minimalen und maximalen Intensitätswerte in den Trainingsdaten um jeweils 15% erweitert. Liegt ein Messwert außerhalb dieses Bereichs, so deutet dies auf einen Fehler hin. Mögliche Ursache hierfür kann ein Defekt an der Hardware sein, beispielsweise an der Anregungsbeleuchtung, oder eine Sättigung des Detektors aufgrund stark reflektierender Materials. Weitere Fehler können durch Fehlbedienung des Nutzers bedingt sein, beispielsweise wenn der Sensorkopf nicht auf dem natürlichen Zahn aufliegt sondern mit großem Abstand gemessen wird. Wenn derartige Ereignisse zu unplausiblen Eingangswerten führen, klassifiziert das Modell in der Regel die Extremwerte der Referenzfarben. Vereinzelt können dadurch aber auch mehr oder weniger zufällige Ergebnisse auftreten. Aus diesem Grund werden derartige Messungen nicht klassifiziert und der Nutzer wird auf den Fehler hingewiesen.

Überprüfung der Ergebnissicherheit

Diese Funktion überprüft, wie belastbar die Prognosewerte des Klassifikationsergebnisses ist. Im Gegensatz zur bereits vorgestellten Plausibilitätsprüfung, welche die direkte Nachbarschaft der Ergebnisse überprüft, wird hier das Ergebnis anhand der Prognosewerte detaillierter untersucht. Zuerst sollte der beste Prognosewert, welcher einen Wertebereich zwischen Null und Eins aufweist, über einer bestimmten Schwelle liegen. Diese wird aufgrund der durchgeführten Untersuchungen auf 0.3 festgelegt. Liegt der Prognosewert unter dieser Schwelle, so ist die Klassifikation mit einem zu großen Unsicherheitsfaktor behaftet und sollte wiederholt werden. Die Folge daraus ist eine entsprechende Ausgabe, die den Benutzer darauf hinweist, eine Wiederholungsmessung durchzuführen. Da die Nutzerreaktion auf derartige Ausgaben jedoch noch untersucht werden muss, wurde diese Funktion in der bisherigen Auslieferung des VITA Easyshade noch nicht aktiviert, bis weitere Untersuchungen im praktischen Einsatz diesbezüglich abgeschlossen sind.

Überprüfung der Prognosewerte hinsichtlich Zwischenfarben

Diese Funktion wurde für den Farbraum 3D-Master entwickelt. Sie überprüft, wie eng die Prognosewerte des erst- und zweitbesten Klassifikationsergebnisses beieinander liegen. Ist die Differenz dieser beiden Prognosewerte kleiner als 0.1, so handelt es sich mit großer Wahrscheinlichkeit um eine Farbausprägung, welche zwischen den beiden Referenzfarben liegt. Diese sogenannten Zwischenfarben (beispielsweise liegt die Farbe 1,5M1 im Farbraum zwischen den Farben 1M1 und 2M1) können helfen, die exakte Farbausprägung genauer zu bestimmen. Liegt die Differenz der Prognosewerte unter der Schwelle, so bestimmt die Funktion die entsprechende Zwischenfarbe und gibt diese als optionales Ergebnis aus. Diese Funktion funktioniert nur bei symmetrischen und auf Koordinaten basierenden Farbraumbezeichnungen wie dem 3D-Master System. Die Funktion benötigt ebenfalls noch eine größere Testphase, bevor sie im ausgelieferten System aktiviert werden kann. Hierbei muss evaluiert werden, in wie weit die bereits erwähnten Fehlerquellen bei den Messergebnissen die Prognosewerte beeinflussen und in wie weit hierbei tatsächlich auf Zwischenfarben geschlossen werden kann.

6.3 Ergebnisse zur realisierten Systemtechnik

Mit dem Trainingsmodul und dem Klassifikationsmodul stehen zwei Funktionsmuster zur Verfügung, welche im Prinzip mit geringen Anpassungen für jedes Farbsystem in jede Art spektralen Sensor mit Mikroprozessor integriert werden können. Durch den modularen Aufbau können zudem einfach neue Klassifikationsverfahren hinzugefügt werden, ohne dass dies zu größeren Änderungen in der Datenstruktur oder der Behandlung von Trainingsdaten



Abbildung 6.5: Ergebnisanzeige der Klassifikation der Zahnfarbe mit dem VITA Easyshade unter Verwendung der neu entwickelten Klassifikationsmethoden.

führt.

In Zusammenarbeit mit Vita wurden die Ergebnisse der Arbeit in den VITA EasyShade Sensor integriert. Hierfür wurde zuerst wie oben beschrieben mit dem realisierten Trainingsmodul auf Basis der zur Verfügung stehenden gemessenen Daten von realen Zähnen ein mehrstufiges Modell zur Klassifikation der Zahnfarbe erstellt. Zudem wurde das Klassifikationsmodul auf den Mikroprozessor des Sensors installiert und die berechneten Modelle übertragen [N.N. 2011b]. Als Ergebnis der Klassifikation werden direkt die Bezeichnung der ermittelten Referenzfarben ausgegeben (Abbildung 6.5).

Die hier vorliegende Entwicklung beinhaltet somit ein neuartiges Verfahren zur Farbklassifikation von Zahnfarben auf der Basis von maschinellem Lernen. Durch die Integration der Ergebnisse in den VITA Easyshade steht dem Nutzer eine einfache und intuitive Möglichkeit zur Bestimmung der Zahnfarbe zur Verfügung. Die Klassifikation zeichnet sich in Bezug auf die vorhandenen preiswerten Hardwarekomponenten durch größtmögliche Robustheit und Fehlertoleranz aus. Das Verfahren liefert bei den vergleichbaren Gütekriterien eine um bis zu 40% bessere Klassifikationsgüte als das bisher eingesetzte Verfahren. Zudem kann ein Neutraining der Klassifikationsmodelle oder das Hinzufügen neuer Referenzfarben auch von nicht versierten Benutzern einfach durchgeführt werden, wodurch eine nachhaltige Entwicklung gegeben ist.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Zusammenfassung

In der vorliegenden Arbeit wurde zuerst die Problematik der menschlichen Wahrnehmung bei der manuellen und visuellen Farbermittlung an Zähnen mit Hilfe von Referenzfarbmustern vorgestellt. Anhand der ermittelten Defizite wurde das Entwicklungspotenzial für ein automatisiertes Verfahren zur Farbbestimmung abgeleitet sowie notwendige Anforderungen zur Bewertung möglicher Lösungen aufgestellt. Die Hauptanforderungen für eine automatische Bestimmung der Zahnfarbe sind ein robustes Verfahren mit möglichst hoher Klassifikationsgüte sowie einer kostengünstigen Realisierung. Dazu muss ein potenzieller Sensor möglichst mobil sein und Messungen direkt an den natürlichen Zähnen der Patienten ermöglichen. Daraufhin wurden bisher verwendete Systeme zur Farbmessung vorgestellt, wobei spektrale Sensoren prinzipbedingt hier am geeignetsten sind. Im Stand der Technik wurden daher bereits am Markt verfügbare spektrale Messgeräte auf deren Eignung zur Bestimmung der Zahnfarbe untersucht und mit den abgeleiteten Anforderungen verglichen. Dabei zeigte sich, dass preiswerte handliche Smart-Sensoren aufgrund der Wirtschaftlichkeit sowie der Möglichkeit zur intraoralen Messung über das größte Potenzial verfügen. Jedoch weist diese Klasse von Sensorsystemen eine vergleichsweise schlechte Klassifikationsgüte auf. Als Gründe hierfür sind die mäßige Qualität der Messdaten sowie Ineffizienz bei der Datenauswertung zu nennen. Für eine umfassende Akzeptanz bei potenziellen Nutzern ist somit die Verbesserung der Klassifikationsgüte unerlässlich. Der vielversprechendste Lösungsansatz lag in der Nutzung eines bestehenden spektralen Smart-Sensors - hier, der VITA Easyshade - sowie der Optimierung der Klassifikationsgüte mit einem neuen Verfahren auf Basis des maschinellen Lernens.

Ausgehend von den Ergebnissen der Analyse wurden sechs potentiell geeignete Methoden des maschinellen Lernens sowie drei vielversprechende Klassifikationsansätze analysiert und permutiert eingesetzt. Als vorgelagertes Verfahren wurden außerdem eine Methode auf Ba-

sis des Majority-Filtering entwickelt, um reale Messdaten hinsichtlich Ihrer Tauglichkeit für das Modelltraining automatisch zu bewerten und inkonsistente Daten ausschleusen zu können. Zur Steigerung der Ergebnisqualität wurde für die Klassifikationsergebnisse zudem eine Plausibilitätsprüfung auf Basis der Nachbarschaften im Farbraum konzipiert und umgesetzt.

Zur Evaluierung der ausgewählten Lösungsansätze konnte im Rahmen der Arbeit auf fast 4000 spektrale Messungen an Zähnen von Probanden zurückgegriffen werden. Mit diesen Daten konnten die Modelle der diskutierten Klassifikationsmethoden und Auswertansätze mit einer definierten Menge an Daten trainiert und mit dem System unbekanntem Daten validiert werden. Zur objektiven Bewertung der Klassifikationsgüte der potenziellen Verfahren wurden vier geeignete Qualitätskriterien definiert. Als bestes Verfahren wurde ein mehrstufiger Klassifikationsansatz auf Basis von künstlichen Neuronalen Netzen ausgewählt. Dessen Klassifikationsgüte liegt bei den vergleichbaren Gütekriterien im Mittel um knapp 40% höher als bei dem bisher eingesetzten Verfahren zur Ermittlung der Zahnfarbe. Die Validierung zeigte weiterhin, dass sämtliche gestellte Anforderungen erfüllt werden konnten, wodurch die Tauglichkeit des Verfahrens zur robusten Ermittlung der Zahnfarbe mit einem handlichen und preiswerten spektralen Sensor gegeben ist.

Anschließend folgte eine Realisierung in Form von Funktionsmustern. Hierbei wurde zuerst ein vom Sensor unabhängiges Modul zum Training der Klassifikationsmodelle realisiert. Hiermit können die Modelle optimiert sowie neue Referenzfarben für die Klassifikation hinzugefügt werden. Weiterhin wurde ein Klassifikationsmodul entwickelt, welches direkt auf den Mikroprozessor des Smart-Sensors integriert werden kann. Somit können die gemessenen spektralen Daten direkt ausgewertet und für den Nutzer visualisiert werden. Exemplarisch wurde das Gesamtsystem für die gängigen Farbräume 3D-Master und classical umgesetzt und in Zusammenarbeit mit der Firma Vita in den VITA Easyshade Sensor integriert.

7.2 Ausblick

Durch die systemtechnische Umsetzung kommt das entwickelte Verfahren in den realen Einsatz, wodurch weitere Erkenntnisse hinsichtlich der praktischen Eignung zu erwarten sind. Abgesehen davon bieten folgende Punkte aus dem Gebiet der Klassifikation von Zahnfarben Potenzial hinsichtlich weiterer Untersuchungen.

Im Rahmen der Arbeit wurde der Fokus auf die Entwicklung der Auswertalgorithmik und Klassifikationsansätze gelegt, da die Hardwarekosten des Sensors im Hinblick auf die

Kundenakzeptanz möglichst gering gehalten werden müssen. Jedoch bietet auch der Sensor Optimierungspotenzial. So könnten mit einem flexiblen Sensorkopf mögliche Messfehler aufgrund unebener Zahnoberflächen oder Verkippungen stark reduziert werden.

Eine weitere äußerst interessante Untersuchung bezüglich der Zahnfarbe wäre die Analyse der Farbverteilung einer Zahnoberfläche. Natürliche Zähne weisen in den seltensten Fällen eine komplett einheitliche Farbe auf der Oberfläche auf. Messungen an unterschiedlichen Punkten der Zahnoberfläche können somit das Ergebnis der bisherigen Messmethode mehr oder weniger stark beeinflussen. Durch die Kenntnis der relativen Farbverteilung eines Zahns könnte so anhand der genauen und absoluten Messung an einem Ort auf der Zahnoberfläche auf die entsprechende absolute Farbverteilung geschlossen werden. Die relative Farbverteilung könnte hierbei beispielsweise durch eine Farbkamera ermittelt werden. Durch die Klassifikation der Farbe am Messpunkt und der Histogrammverteilung des Bildes könnte dann in einem weiteren Schritt die absolute Farbverteilung berechnet oder zumindest ausreichend abgeschätzt werden. Dieses Vorgehen könnte die Klassifikationsgüte der Zahnfarbe bei entsprechender Auswertalgorithmik optimieren, allerdings auf Kosten des Hardwareaufwandes. Eine einfachere Möglichkeit liegt in der Durchführung von mehreren Messungen an unterschiedlichen Stellen des Zahns sowie einer entsprechenden Interpolation.

Hinsichtlich der Auswertung gibt es fortlaufend mathematische Neuentwicklungen sowie Optimierungen der bestehenden Algorithmen, welche unter Umständen auch für die Klassifikationen der Zahnfarbe geeignet sind. Durch den modularen Aufbau der Software können deshalb bei Bedarf neue Klassifikationsverfahren mit wenig Aufwand in die realisierten Funktionsmuster integriert werden. Auch eine Untersuchung hinsichtlich der Kombinationen von unterschiedlichen Auswertverfahren und deren Varianten wäre eine interessante Ergänzung und könnte eine Möglichkeit zur Verbesserung der Klassifikationsgüte darstellen.

Großes Potenzial zur Verbesserung des entwickelten Verfahrens ist prinzipbedingt in der Gewinnung weiterer repräsentativer Trainingsdaten zu sehen. Eine bereits genannte Problematik ist hierbei, dass die gemessenen Daten einer Referenzfarbe manuell und visuell durch Experten beurteilt werden müssen. Darin liegt aufgrund der menschlichen Wahrnehmung ein großes Fehlerpotenzial. Eine Verbesserung bei der Ermittlung von Trainingsdaten wäre, dass beispielsweise mehrere Experten die Zahnfarbe zusammen ermitteln und sich auf eine Farbe einigen. Dies würde den Einfluss der menschlichen Wahrnehmung reduzieren und die Qualität der Trainingsdaten signifikant erhöhen. Problematisch ist hierbei jedoch der dafür nötige sehr hohe Zeit- und Personalaufwand. Jedoch liegt trotz aller Aufwände in der Verbesserung der Trainingsdaten die effektivste Möglichkeit zur Steigerung der Klassifikationsgüte, da diese direkt von der Qualität der Trainingsdaten abhängt.

Literaturverzeichnis

- [Abellan-Nebot u. a. 2010] Abellan-Nebot, J.V. und F.R. Subirón, 2010. A review of machining monitoring systems based on artificial intelligence process models. *International Journal of Advanced Manufacturing Technology* **47**(1), S. 237–257. ISSN 0268-3768.
- [Aderhold u. a. 2010] Aderhold, J. und B. Plinke, 2010. Spektroskopische Charakterisierung von Oberflächen. In: Sackewitz, Michael (Leitung); Fraunhofer Allianz Vision: *Inspektion und Charakterisierung von Oberflächen mit Bildverarbeitung*. Seminar mit Praktikum, Handbuch, 9. und 10. Dezember 2010, Karlsruhe.
- [Altmann 2008] Altmann, R., 2008. *Mehr Bits für Farbe - Der richtige Farbraum*, 25.07.2011. Verfügbar: www.heise.de/foto/artikel/Mehr-Bits-fuer-Farbe-226838.html. Zugriff: 10.08.2012.
- [Bacher u. a. 2010] Bacher, J., Pöge, A. und K. Wenzig, 2010. *Clusteranalyse - Anwendungsorientierte Einführung in Klassifikationsverfahren*. 3. Auflage. München: Oldenburger Wissenschaftsverlag GmbH. ISBN 978-3486584578.
- [Baltzer u. a. 2004] Baltzer, A. und V. Kaufmann-Jinoian, 2004. Die Bestimmung der Zahnfarben. *QZ Quintessenz Zahntechnik* **30**(7), S. 726–740. ISSN 0340-4641.
- [Berger-Schuhn 1991] Berger-Schuhn, A., 1991. *Praktische Farbmessung*. Göttingen: Muster-Schmidt Verlag. ISBN 978-3788140496.
- [Berthold u. a. 1999] Berthold, M. und D.J. Hand, 1999. *Intelligent data analysis: An introduction*. Berlin: Springer. ISBN 978-3-540-48625-1.
- [Biermann 2009] Biermann, B.E., 2009. *Die Beeinflussung der Schmelzstruktur durch laseraktivierte Bleichverfahren. Eine In-vitro-Studie*. Bonn, Hohe Medizinischen Fakultät der Rheinischen Friedrich-Wilhelms-Universität, Diss. Verfügbar: hss.ulb.uni-bonn.de/2009/1897/1897.pdf. Zugriff: 10.08.2012.
- [Bishop 1995] Bishop, C.M., 1995. *Neural networks for pattern recognition*. New York: Oxford University Press. ISBN 978-0198538646.

- [Bishop 2006] Bishop, C.M., 2006. *Pattern recognition and machine learning*. New York: Springer. ISBN 978-0387310732.
- [Bodyanskiy u. a. 2007] Bodyanskiy, Y. u. a., 2007. Nonlinear process identification and modeling using general regression neuro-fuzzy network. In: *52nd IWK - Internationales Wissenschaftliches Kolloquium* 10. - 13. September 2007, Ilmenau. Ilmenau: Universitätsverlag Ilmenau, S. 25–34.
- [Bodyanskiy u. a. 2008] Bodyanskiy, Y. und N. Teslenko, 2008. General regression neuro-fuzzy network for identification of nonstationary plants. *International Journal Information Technologies and Knowledge* **2**, S. 136–142. ISSN 0973-4414.
- [Braun 2011] Braun, R., 2011. *Kosten für Zahnersatz*, 29.07.2011. Verfügbar: <http://www.stern.de/zaehne/ueberblick/kosten-fuer-zahnersatz-wie-sie-bei-bruecken-und-kronen-sparen-koennen-607045.html>. Zugriff: 10.08.2012.
- [Brodley u. a. 1996] Brodley, C. und M. Friedl, 1996. Identifying and eliminating mislabeled training instances. In: *13th national conference on artificial intelligence*. 4.–8. August 1996, Portland, Oregon. Cambridge: MIT Press Ltd, S. 799–805.
- [Bücking 2002] Bücking, W., 2002. Die fehlerfreie Zahnfarbenbestimmung – eine Herausforderung für Zahnarzt und Zahntechniker. *Quintessenz* **53**(2), S. 1199–1207. ISSN 0033-6580.
- [Cai u. a. 2006] Cai, Y.D. u. a., 2006. Using LogitBoost classifier to predict protein structural classes. *Journal of Theoretical Biology* **238**, S. 172–176. ISSN 0022-5193.
- [Cawley u. a. 2010] Cawley, G.C. und N.L.C. Talbot, 2010. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research* **11**(6), S. 2079–2107. ISSN 1532-4435.
- [Cristianini u. a. 2000] Cristianini, N. und J. Shawe-Taylor, 2000. *An Introduction to Support Vector Machines*. Cambridge: University Press. ISBN 978-0521780193.
- [Dana u. a. 2007] Dana, K.J., Cula, O.G. und J. Wang, 2007. Surface detail in computer models. *Image and Vision Computing* **25**(7), S. 1037–1049. ISSN 0262-8856.
- [Dettling u. a. 2003] Dettling, M. und P. Buhlmann, 2003. Boosting for tumor classification with gene expression data. *Bioinformatics* **19**(9), S. 1061–1069. ISSN 1471-2105.
- [Devigus 2003] Devigus, A., 2003. Die digitale Farbmessung in der Zahnmedizin. *Quintessenz* **54**(5), S. 495–500. ISSN 0340-4641.

- [Duda u. a. 2000] Duda, R.O., Hart, P.E. und D.G. Stork, 2000. *Pattern Classification*. 2nd Edition. Oxford: Wiley. ISBN 978-0471056690.
- [Ebel u. a. 1988] Ebel, S. und W. Mück, 1988. Algorithmen zum automatischen Vergleich von Spektren in der HPLC/UV-Kopplung. *Fresenius Zeitschrift für Analytische Chemie* **331**, S. 351–358. ISSN 1618-2642.
- [Elliott u. a. 2008] Elliott, R.J., Aggoun, L. und J.B. Moore, 2008. *Hidden Markov Models: Estimation and control*. 3rd Edition. Berlin: Springer. ISBN 978-0387943640.
- [Faber 2006] Faber, F. J., 2006. *Die Bestimmung der Zahnfarbe – Stellungnahme der DGZMK*, 18.11.2011. Verfügbar: www.dgzmk.de/uploads Zugriff: 10.08.2012.
- [Filip u. a. 2009] Filip, J. und M. Haindl, 2009. Bidirectional texture function modeling: A state of the art survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(11), S. 1921–1940. ISSN 0162-8828.
- [Franke 2007] Franke, K.H., 2007. Farb- und spektroskopische Bildverarbeitung. In: *35. Heidelberger Bildverarbeitungsforum*. Seminar, Handbuch, 9. Oktober 2007, Waldbronn.
- [Friedman u. a. 2000] Friedman, J., Hastie, T. und R. Tibshirani, 2000. Adaptive logistic regression: A statistical view of boosting. *Annals of Statistics* **28**(2), S. 337–407. ISSN 0090-5364.
- [Guan u. a. 2010] Guan, D. u. a., 2010. Identifying mislabeled training data with the aid of unlabeled data. *Applied Intelligence* **35**(3) S. 345–358. ISSN 0924-669X.
- [Guyon u. a. 2006] Guyon, I. u. a., 2006. Performance prediction challenge. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, 16.–21. Juli 2006, Vancouver, BC. Piscataway, N.J.: IEEE Xplore, S. 2958–2965.
- [Guyon u. a. 2007] Guyon, I. u. a., 2007. Agnostic learning vs. prior knowledge challenge. In: *IEEE, 2007 International Joint Conference on Neural Networks*, 12.–17. August 2007, Orlando, Florida. Piscataway, N.J.: IEEE Catalog Number: 07CH37922C, S. 829–834.
- [Guyon u. a. 2008] Guyon, I. u. a., 2008. Design and analysis of the causation and prediction challenge. In: *JMLR Workshop and Conference Proceedings: 3. Causation and Prediction Challenge WCCI*, 1.–6. Juli 2008, Hong Kong. , S. 1–33. Verfügbar: jmlr.org/proceedings/papers/v3/guyon08a/guyon08a.pdf. Zugriff: 12.08.2013

- [Guyon 2009] Guyon, I., 2009. *A practical guide to model selection*, 05.09.2013. Verfügbar: eprints.pascal-network.org/archive/00005768/01/guyon-mlss.pdf. Zugriff: 10.08.2012
- [Guyon u. a. 2009] Guyon, I. u. a., 2009. Analysis of the KDD Cup 2009: Fast scoring on a large orange customer database. In: *JMLR: Workshop and Conference Proceedings*: 3. , S.1–22. Verfügbar: jmlr.org/proceedings/papers/v7/guyon09/guyon09.pdf. Zugriff: 10.08.2013
- [Hand u. a. 2001] Hand, D.J. und K. Yu, 2001. Idiot’s Bayes - not so stupid after all?. *International Statistical Review* **69**(3), S. 385–398. ISSN 1751-5823.
- [Hanss 2005] Hanss, M., 2005. *Applied fuzzy arithmetic*. Berlin: Springer. ISBN 978-3540242017.
- [Hartung u. a. 2009] Hartung, J., Elpert, B. und K.H. Klösner, 2009. *Statistik: Lehr- und Handbuch der angewandten Statistitk*. München: Oldenburger Wissenschaftsverlag GmbH. ISBN 978-3486259056.
- [Hastie u. a. 2009] Hastie, T. u. a., 2009. *The elements of statistical learning - data mining, inference and prediction*. 2nd Edition. New York: Springer. ISBN 978-0387848570.
- [Heintz u. a. 2010] Heintz, R. und J. Beyerer, 2010. Theorie und Methoden der Farbmessung. In: Sackewitz, Michael (Leitung); Fraunhofer Allianz Vision: *Inspektion und Charakterisierung von Oberflächen mit Bildverarbeitung*. Seminar mit Praktikum, Handbuch, 9. und 10. Dezember 2010, Karlsruhe.
- [Hendrych 2011] Hendrych, B., 2011. Farberkennung mit Festoptik und Lichtleiter. *Journal für Oberflächentechnik* **9**, S. 68–71. ISSN 0940-8789.
- [Hunt 1998] Hunt, R.W.G., 1998. *Measuring colour*. 3rd Edition. Kingston-upon-Thames: Fountain Press. ISBN 978-1119975373.
- [Hüllermeier u. a. 2010] Hüllermeier, E. und S. Vanderlooy, 2010. Combining predictions in pairwise classification: an optimal adaptive voting strategy and its relation to weighted voting. *Pattern Recognition* **43**(1), S. 128–142. ISSN 0031-3203.
- [Jankowski u. a. 2006] Jankowski, N. und K. Grabczewski, 2006. *Feature Extraction: Foundations and Applications - Learning Machines*. Berlin: Springer. ISBN 978-3540354871.
- [Kleppmann 2003] Kleppmann, W., 2003. *Versuchsplanung - Produkte und Prozesse optimieren*. München: Hanser. ISBN 978-3446427747.

- [Knieling 2007] Knieling, S., 2007. *Einführung in die Modellierung künstlich neuronaler Netzwerke*. Duisburg: WiKu. ISBN 978-3865531926.
- [Kohonen 2001] Kohonen, T., 2001. *Self-organizing maps*. Berlin: Springer. ISBN 978-3540679219.
- [Kourtis u. a. 2009] Kourtis, S., Chronopoulos, V. und P. Papadimitriou, 2009. Die Verwendung von Easy shade und Spectroshade - Bewertung elektronischer Geräte zur Bestimmung der Zahnfarbe. *Journal of Continuing Dental Education* **12**(2), S. 233–240. ISSN 0362-9120.
- [Krueger-Janson 2006] Krueger-Janson, U., 2006. Analyse der Farb- und Strukturmerkmale eines Zahnes. *Journal of Multidisciplinary Collaboration of Restorative Dentistry*. **9**(3), S. 212–223. ISSN 1862-2879.
- [Küppers 1999] Küppers, H., 1999. *Harmonielehre der Farben*. Ostfildern: Dumont. ISBN 978-3770121922.
- [Küppers 2005] Küppers, H., 2005. *Schnellkurs Farbenlehre*. Ostfildern: Dumont. ISBN 978-3832176402.
- [Li 2010] Li, P., 2010. Robust LogitBoost and adaptive base class (ABC) LogitBoost. In: *Proceedings of the Twenty-Sixth Annual Conference on Uncertainty in Artificial Intelligence*. Poster, 8.–11. Juli 2010, Catalina Island, California. Verfügbar: arxiv.org/ftp/arxiv/papers/1203/1203.3491.pdf. Zugriff: 12.08.2013
- [Lutz 2006] Lutz, R.W., 2006. LogitBoost with trees applied to the WCCI 2006 performance prediction challenge datasets. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings* 16.–21. Juli 2006, Vancouver, BC. Piscataway, N.J.: IEEE Xplore, S. 1657–1660.
- [MacKay 2003] MacKay, D., 2003. *Information theory, inference and learning algorithms*. Cambridge: University Press. ISBN 978-0521642989.
- [Marchesi 2005] Marchesi, J.J., 2005. *Handbuch der Fotografie - Physikalische Farbtheorie und moderne Farbmetrik*. Gilching: Verlag Photographie. ISBN 978-3933131775.
- [Mitchell 1997] Mitchell, T.M., 1997. *Machine learning*. New York: McGraw-Hill. ISBN 978-0070428072.
- [Mitra u. a. 2000] Mitra, S. und Y. Hayashi, 2000. Neuro-fuzzy rule generation: Survey in soft computing framework. *IEEE Transactions on Neural Networks* **11**(3), S. 748–768. ISSN 1045-9227.

- [Montgomery 2009] Montgomery, D.C., 2009. *Design and analysis of experiments*. New York: Wiley. ISBN 978-1118146927.
- [Müller 2004] Müller, M., 2004. *Untersuchungen zur klinischen Anwendbarkeit des computergestützten Zahnfarberkennungssystems Digital Shade Guide 4*. Gießen, Fachbereich Humanmedizin der Justus- Liebig- Universität, Diss. Verfügbar: geb.uni-giessen.de/geb/volltexte/2004/1637/pdf/MuellerMarcus-2004-07-07.pdf. Zugriff: 10.08.2012
- [N.N.] N.N., *Farbmetrik - Farben sehen und messen: Eine Einführung und Übersicht*, 11.11.2011. Verfügbar: <http://www.farbkarten-shop.de/torso-data/download/FARBMETRIK.pdf>. Zugriff: 12.08.2012
- [N.N. 1979] DIN 5033-1 (1979-03-00). *Farbmessung; Grundbegriffe der Farbmetrik*
- [N.N. 2005a] N.N., 2005. *Renesas 32-Bit CISC Microcomputer Hardware Manual H8SX/1653 Group*. Sales Strategic Planning Div. Renesas Technology Corp..
- [N.N. 2005b] DIN 6172:1993-03. *Metamerie-Index von Probenpaaren bei Lichtartwechsel*
- [N.N. 2006] N.N., 2006. *Das VITA System 3D-Master*, 14.07.2011. Verfügbar: www.vita-zahnfabrik.com. Zugriff: 12.08.2012
- [N.N. 2008a] N.N., 2008. *Schöne Zähne liegen im Trend-Im Auftrag der Initiative proDente-Zusammenfassung Zahnärztekammer*, 13.07.2011. Verfügbar: www.zahnaerztekammer-sh.de/Presseservice/tipps/tip66.html. Zugriff: 12.08.2012
- [N.N. 2008b] N.N., 2008. *Handbuch digital shade guide DSG 4 plus Version 2.3*, 15.11.2011. Verfügbar: www.rieth-dentalprodukte.de. Zugriff: 12.08.2012
- [N.N. 2008c] N.N., 2008. *Zahnfarbmessung digital shade guide*, 15.11.2011. Verfügbar: www.rieth-dentalprodukte.de. Zugriff: 12.08.2012
- [N.N. 2010a] N.N., 2010. *TECHNIK crystaleye*, 15.11.2011. Verfügbar: <http://www.dentalimpulse.de/download/pdfs/ditechnik/ditechnik-crystaleye-2010-1.pdf>. Zugriff: 12.08.2012..
- [N.N. 2010b] N.N., 2010. *VITA Kompendium*, 14.07.2011. Verfügbar: <https://www.vita-zahnfabrik.com/de/Kompendium-5240.html>. Zugriff: 12.08.2013
- [N.N. 2010c] N.N., 2010. *Tipps zur Zahnfarbnahme*, 14.07.2011. Verfügbar: <https://www.vita-zahnfabrik.com/de/Farbnahme-27.html>. Zugriff: 12.08.2012

- [N.N. 2011a] N.N., 2011. *CIE Farbraum*, 25.07.2011. Verfügbar: <http://messtechnik.harzoptics.de/farbort.gif>. Zugriff: 10.08.2012
- [N.N. 2011b] N.N., 2011. *Evolution in der Farbmessung: VITA Easyshade Advance*, 25.06.2012. Verfügbar: <http://www.offenes-presseportal.de/gesundheitsmedizin/evolution-in-der-farbmessung-vita-easyshade-advance-248400.htm>. Zugriff: 11.08.2012
- [N.N. 2012] Bundesrepublik Deutschland, Bundesministerium der Justiz, 2012. Gesetz über Medizinprodukte (Medizinproduktegesetz – MPG; BGBl. I S. 2192).
- [Nauck u. a. 1997] Nauck, D. und R. Kruse, 1997. A neuro-fuzzy method to learn fuzzy classification rules from data. *Fuzzy Sets and Systems* **89**(3), S. 277–288. ISSN 0973-421X.
- [Neher 2011] Neher, J., 2011. *Neuro-Fuzzy-Modellierung zur umfassenden Prozessüberwachung am Beispiel des Ultraschallschweißens von Kunststoffteilen*. Stuttgart, Univ., Diss. Verfügbar: publica.fraunhofer.de/eprints/urn:nbn:de:bsz:93-opus-75874.pdf. Zugriff: 12.08.2012
- [Newton 1672] Newton, I., 1672. A new theory about light and colours. *Philosophical Transactions of the royal society* **6**, S. 3075–3087. ISSN 1364-503X.
- [Paravina 2009] Paravina, N., 2009. Der beste Laufschuh für eine einfache Farbbestimmung im Dentallabor. *ZT Produkte* **1**(1), S. 9.
- [Park u. a. 2000] Park, S. und T. Han, 2000. Iterative inversion of fuzzified neural networks. *IEEE Transactions on Fuzzy Systems* **8**(3), S. 266–280. ISSN 1063-6706.
- [Peters 2010] Peters, S., 2010. *Automatischer Entwurf und Autokonfiguration von Bildverarbeitungssystemen für die industrielle Oberflächeninspektion*. Kaiserslautern, Technische Universität, Diss. Verfügbar: kluedo.ub.uni-kl.de. Zugriff: 12.08.2012
- [Pflüger 2010] Pflüger, M., 2010. Identifikation und Analyse von Kontaminationen und technischen Substanzen mittels IR-Spektroskopie. In: Sackewitz, Michael (Leitung); Fraunhofer Allianz Vision: *Inspektion und Charakterisierung von Oberflächen mit Bildverarbeitung*. Seminar mit Praktikum, Handbuch, 9. und 10. Dezember 2010, Karlsruhe.
- [Pryor 2011] Pryor, D., 2011. Online-Spektralmesssystem misst Farbe am Stahlband - Heißer Auftrag. *Qualität und Zuverlässigkeit (QZ)* **56**(4), S. 48–49. ISSN 0720-1214.

- [Rauh u. a. 2001] Rauh, W. und Schmidberger E., 2001. Die Messlatte liegt bei Null. *Spektrum der Wissenschaft* **5**, S. 88–89. ISSN 170-2971.
- [Richter 1981] Richter, M., 1981. *Einführung in die Farbmatrik*. 2. Auflage. Berlin: Walter de Gruyter. ISBN 978-3110082098.
- [Rosenblatt 1958] Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* **65**(6), S. 386–408. ISSN 0033-295X.
- [Ross 2002] Ross, F., 2002. *Untersuchung zur Reproduzierbarkeit der Zahnfarbbestimmung mit einem neuen Zahnfarbringsystem*. Hannover, Medizinische Hochschule, Diss. Verfügbar: nbn-resolving.de/urn:nbn:de:gbv:354-20060718484. Zugriff: 12.08.2012
- [Sammon 1969] Sammon, J.W., 1969. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers* **18**(5), S. 401–409. ISSN 0018-9340.
- [Schapire u. a. 1999] Schapire, R.E. und Y. Singer, 1999. Improved boosting algorithms using confidence-rated predictions. *Machine Learning* **37**, S. 297–336. ISSN 0885-6125.
- [Schmidberger u. a. 2007] Schmidberger, E. und J. Neher, 2007. Was das Signal weiss. *Qualität und Zuverlässigkeit (QZ)* **52**(12), S. 54–55. ISSN 0720-1214.
- [Schröder 2004] Schröder, U., 2004. Farbmeßgeräte richtig einsetzen und einkaufen - Leitfaden zur Erstellung eines Anwendungsprofils zur Ermittlung des richtigen Farbmeßgerätes. *Welt der Farben* **7**, S. 18–21.
- [Teti u. a. 2010] Teti, R. u. a., 2010. Advanced monitoring of machining operations. *CIRP Anals - Manufacturing Technology* **59**(2), S. 717–739. ISSN 0007-8506.
- [Viénot u. a. 2005] Viénot, F. und J. Le Rohellec, 2005. Psychologie Farbwahrnehmung - Bunte Hirngespinnste. *Gehirn & Geist, Magazin für Psychologie und Hirnforschung* **5**, S. 34–39. ISSN 1618-8519.
- [Wolpert u. a. 1997] Wolpert, D.H. und W.G. Macready, 1997. No Free Lunch Theorems for Optimization. *IEEE Transaction on Evolutionary Computation* **1**(1), S. 62–82. ISSN 1089-778X.
- [Wünnemann 2009] Wünnemann, P., 2009. *Untersuchung: Ist die Zahnfarbdifferenzierung lernbar?*. Berlin, Medizinische Fakultät Charité - Universitätsmedizin, Diss. Verfügbar: www.deutsche-digitale-bibliothek.de/item/LKPQX2VWDDRYK7LNJOBEEHX3DBWOPUYJ. Zugriff: 12.08.2012

[Zadeh 1965] Zadeh, L.A., 1965. Fuzzy sets. *Information and Control* **8**, S. 338–353. ISSN 0019-9958.

[Zell 1994] Zell, A., 1994. *Simulation Neuronaler Netze*. München: Oldenburger Wissenschaftsverlag GmbH. ISBN 978-3486243505.

In der heutigen Gesellschaft hat das Aussehen einen immer größeren Stellenwert. Entsprechend kritisch werden deshalb unpassende Farbausprägungen beim Zahnersatz gesehen. In der Praxis wird die Zahnfarbe hierfür meist visuell mithilfe von Farbringen bestimmt, was jedoch fehleranfällig ist.

Diese Arbeit befasst sich deshalb mit der Entwicklung eines Verfahrens zur automatischen Bestimmung der Zahnfarbe sowie die Integration in den Sensor VITA Easyshade. Kernelement ist die Entwicklung eines mehrstufigen Prognosemodells für die Zahnfarbe auf Basis von künstlichen Neuronalen Netzen.

Zur Evaluierung der ausgewählten Lösungsansätze konnte im Rahmen der Arbeit auf fast 4000 spektrale Messungen an Zähnen von Probanden zurückgegriffen werden. Die Validierung des Verfahrens zeigte, dass anhand definierter Kriterien die Prognosegüte im Vergleich zu herkömmlichen Verfahren um knapp 40% gesteigert werden konnte.

ISBN 978-3-8396-0698-8



FRAUNHOFER VERLAG