

3D-Gesichtsmodellierung mit bilddatenbasierter Modifikation eines generischen Gesichtsmodells

**Gesichtsbildanalyse und Bereitstellung der Daten für
die Modifikation**

Bachelor-Thesis

zur Erlangung des akademischen Grades B.Sc.

Evelyn Kreutz

2025884



Hochschule für Angewandte Wissenschaften Hamburg
Fakultät Design, Medien und Information
Department Medientechnik

Erstprüfer: Prof. Dr. Torsten Edeler

Zweitprüfer: Prof. Dr. Andreas Plaß

Hamburg, 15.08.2016

Zusammenfassung

Das Ziel der vorliegenden Bachelorarbeit ist die Entwicklung eines teilautomatisierten Verfahrens zur Gewinnung und Bereitstellung von Bildanalysedaten, die ein menschliches Gesicht in seinen Formeigenschaften beschreiben. Die Bedingungen für die Entwicklung bestehen darin, eine ähnlich gute Analysedatenqualität zu erreichen wie durch die manuelle Datenerfassung und gleichzeitig den Benutzeraufwand für die Datengewinnung gering zu halten. Die Analysedaten werden in der Bachelorarbeit von [Burmeister \(2016\)](#) dazu genutzt, ein 3D-Modell des Gesichts zu erstellen, indem mit den Daten ein generisches Gesichtsmodell modifiziert wird. Die Datengewinnung erfolgt durch die Analyse einer Frontalansicht des Gesichts. Dafür wird das Verfahren zur Konturlokalisierung der Gesichtsmerkmale von [Hähnel \(2008\)](#) (Modellbasierte posen- und mimikinvariante Gesichtserkennung) an die Anforderungen dieser Arbeit angepasst.

Der Analyseerfolg wurde empirisch anhand eines Tests mit zwei Gruppen von Gesichtsbildern bemessen. Die Versuchsergebnisse dienen als Grundlage für die Einschätzung der Qualität der Datengewinnung und des damit verbundenen Benutzeraufwands. Aufgrund dieser Einschätzung wird festgestellt, dass die Qualität der Gesichtsanalyse zu stark variiert oder der Benutzeraufwand die vorgegebene Grenze überschreitet. Deshalb ist das Datengewinnungsverfahren in der aktuellen Form noch nicht dazu geeignet, eine zuverlässig gute Datenqualität bei möglichst geringem Benutzeraufwand zu erreichen. Darum werden zum Schluss Möglichkeiten aufgeführt, mit denen die Ergebnisqualität der Daten verbessert und die Benutzerfreundlichkeit erhöht werden kann.

Abstract

The aim of this thesis is to develop an image analysis system that collects face specific data which describe its shape. The requirements for the developed system are to achieve a similar high quality of analysis data as a pure manual data collection and to keep the user effort for this data collection as minimal as possible at the same time. Based on this data a 3D model is created in the thesis of [Burmeister \(2016\)](#) by modifying a generic face model. The data collection is realised by analysing the frontal view of the face based on the method of Michael Hähnel (Modellbasierte posen- und mimikinvariante Gesichtserkennung, 2007). This method is adapted to the requirements of this thesis.

The success of the analysis was measured in an empirical way with a test that consisted of two groups of face images. The quality of the data collection and the required user effort is assessed based on the results of the test. The assessment shows that either the quality of the face analysis varies too much or that the user effort is not in line with the desired effort. Therefore the method of data collection is not yet suitable to achieve reliably high data quality while keeping the user effort as minimal

as possible. On that account options for optimising the quality of the resulting data and increasing the usability are mentioned.

Ich versichere, die vorliegende Arbeit selbstständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangaben eindeutig kenntlich gemacht.

Ort, Datum

Evelyn Kreutz

Inhaltsverzeichnis

1	Einleitung	1
1.1	Heranführung an das Thema	1
1.2	Konzept	2
2	Gesichtsanalyse in digitalen Bildern	4
2.1	Verfahren zur Gesichtsanalyse	4
2.2	Grundlagen der Bildverarbeitung	6
2.2.1	Beschaffenheit des Eingabematerials der Analyse	7
2.2.1.1	Eigenschaften von Gesichtern	7
2.2.1.2	Anforderungen an die Bildbeschaffenheit	12
2.2.2	Bildvorverarbeitung	19
2.2.2.1	Bildtransformationen	19
2.2.2.2	Filteroperationen	21
2.2.3	Bildanalyse	22
2.2.3.1	Segmentierung	22
2.2.3.2	Merkmalsextraktion	26
3	Implementierung der Gesichtsanalyse	27
3.1	Gesichtsmerkmalslokalisierung	27
3.1.1	Eingrenzung des Suchbereichs	27
3.1.2	Konturlokalisierung und -interpolation	29
3.1.2.1	Augen	29
3.1.2.2	Augenbrauen	33
3.1.2.3	Nase	35
3.1.2.4	Mund	38
3.1.2.5	Kinnlinie	40
3.2	Ausgabedaten der Analyse	42
3.2.1	Berechnung der Merkmalskorrespondenzen	42
3.2.2	Berechnung der Gesichtsverhältnisvektoren	44
4	Ergebnisse	46
4.1	Versuchsbeschreibung	46
4.2	Versuchsauswertung	48
4.3	Ergebnisreflexion	54
5	Fazit	57

Inhaltsverzeichnis

5.1	Erfolgsrate	57
5.2	Verbesserungsvorschläge	59
5.3	Schlussbemerkung	61
	Abbildungsverzeichnis	62
	Tabellenverzeichnis	63
	Literaturverzeichnis	64

1 Einleitung

Diese Arbeit gehört zur Entwicklung einer Software, die ein 3D-Modell aus der Frontalansicht eines beliebigen menschlichen Gesichts erstellt. Das Gesamtthema hat zwei Bereiche: 1. In dieser Arbeit wird ein teilautomatisiertes Verfahren zur Gesichtsbildanalyse entwickelt, das die Daten für die Erstellung des 3D-Modells bereitstellt. 2. In der Arbeit von [Burmeister \(2016\)](#) werden die Daten für die Modifikation eines generischen Gesichtsmodells genutzt, das dadurch an das analysierte Gesicht angepasst wird.

Nach der Heranführung an das Thema 3D-Gesichtsmodellierung und der sich dadurch ergebenden Problemstellung wird das Konzept zur Behandlung des Problems vorgestellt. Zum besseren Verständnis der Umsetzung dieses Konzepts werden dann alle dafür erforderlichen Informationen und Grundlagen beleuchtet. Im Anschluss daran werden die Realisierung des Konzepts und alle dabei durchgeführten Schritte im Detail beschrieben und erläutert. Zum Schluss erfolgt ein Test zur Bemessung der Erfolgsrate der Analyse und eine Ergebnisreflexion in Hinblick auf eine Prozessoptimierung.

1.1 Heranführung an das Thema

Im modernen Informationszeitalter steht die Speicherung, Verarbeitung und Übermittlung von Informationen in digitaler Form im Vordergrund. Wo bisher nur die Mittel der Fotografie, der Zeichnung und der plastischen Modellierung zur Darstellung und Veranschaulichung von Gegenständen und Personen herangezogen werden konnten, besteht heutzutage auch die Möglichkeit der softwarebasierten 3D-Modellierung. Der enorme Vorteil, der sich durch diese Art der Darstellung ergibt, sind die vielfältigen und verschiedenen Nutzungs- und Verarbeitungsmöglichkeiten, die sich durch die Eigenschaften des Mediums digitaler Daten eröffnen.

Der Einsatz von 3D-Modellen aus verschiedenen Gründen und zu verschiedenen Zwecken gewinnt immer mehr an Bedeutung. In einigen Bereichen spielt vor allem die möglichst realistische Modellierung von Menschen und menschlichen Gesichtern eine große Rolle. Das professionelle Umfeld der 3D-Branche liefert für die mehr oder minder exakte Nachbildung von Gesichtern verschiedene optimierte Möglichkeiten, sodass der Einsatz von Gesichtsmodellen im kommerziellen Bereich längst gang und gäbe ist.

Potentielle Benutzer außerhalb des professionellen Bereichs sind in ihren Möglichkeiten zur Erstellung eines 3D-Gesichtsmodells bei angemessener Zeit- und Kosteneff-

1 Einleitung

fizienz bislang jedoch noch sehr eingeschränkt. Zurzeit sind auf dem Markt nur wenige softwarebasierte Lösungen dazu erhältlich, die einen kostspieligen 3D-Scan oder eine zeitaufwendige 3D-Modellierung von Hand ablösen oder vereinfachen könnten und bei denen die Modellierung außerdem unabhängig von der Einbettung in anwendungsgebundene Programme stattfindet. Eine entsprechende kostenfreie Software ist im Internet unter dem Namen Faceworx zu finden. Sie bietet dem Anwender die Möglichkeit, automatisch ein 3D-Gesichtsmodell anhand der Vorlage eines frontalen und eines seitlichen Gesichtsbildes zu erstellen. Dazu notwendig sind mehrere vom Benutzer manuell ausgeführte, spezifische Markierungen auf den Eingabefotos. Derartige Programme können die praktikable Nutzung von 3D-Gesichtsmodellen für den Konsumerbereich attraktiv gestalten und zugänglich machen.

Die zuvor angeführten Informationen zeigen einen Mangel an alternativen Lösungen bei bestehendem Entwicklungspotenzial für den Bereich der Erstellung und Nutzung von Gesichtsmodellen auf. Deshalb birgt die Entwicklung eines neuen Verfahrens zur 3D-Gesichtsmodellierung das Potenzial für eine Optimierung bisheriger Softwarelösungen.

Eine Optimierung besteht vor allem in der Reduktion des Benutzeraufwands für die Erstellung eines Gesichtsmodells bei gleicher oder ähnlicher Ergebnisqualität. Ein reduzierter Benutzeraufwand wird durch die Umwandlung von manuellen in automatisierte Prozesse erreicht. Darum widmen sich diese und die daran anschließende Arbeit der Entwicklung eines Verfahrens, das den Modellierungsprozess soweit automatisiert, dass der einzige manuelle Anteil daran eine kurze Initialisierung der Software ist.

Bezüglich der Umsetzung dessen stellen sich zwei Fragen:

- Kann eine teilautomatisierte, zielgerichtete Analyse eines geeigneten Frontalbildes eines Gesichts eine ausschließlich manuelle Datenerfassung gleichwertig ersetzen, und dies in einer Art und Weise, die den Benutzeraufwand reduziert?
- Inwiefern kann die vollautomatische Nutzung von komplexen Funktionen einer schon existierenden Grafiksuite sowohl den Modellierungsprozess vereinfachen als auch ein zufriedenstellendes Ergebnis liefern?

Diese Arbeit behandelt die erste der Fragen. Das Ziel ist die Gewinnung und Bereitstellung der Bildanalysedaten einer Frontalansicht eines menschlichen Gesichts. Mittels dieser wird die vollautomatisierte Modifikation eines generischen 3D-Gesichtsmodell ermöglicht. Dadurch wird das Modell an das abgebildete Gesicht angepasst, sodass eine möglichst vorlagengetreuen Nachbildung dessen entsteht.

1.2 Konzept

Diese Arbeit behandelt einen Teilbereich der Entwicklung einer Software zur Erstellung eines 3D-Gesichtsmodells anhand der Vorlage eines Gesichtsbildes.

1 Einleitung

Um zu erläutern, an welcher Stelle des Gesamtkonzepts der Inhalt dieser Arbeit steht, wird im Folgenden der chronologische Ablauf des Modellierungsprozesses mittels der Software kurz umrissen:

Zuerst erfolgt eine manuelle Markierung der Suchbereiche für die Gesichtsmerkmale auf dem Eingabefoto. Diese initialisiert die anschließende vollautomatisierte Berechnung der spezifischen Formeigenschaften dieser Merkmale. Davon ausgehend werden verschiedene charakteristische Merkmalspunkte bestimmt, die das Gesicht in seinen spezifischen Eigenschaften gänzlich beschreiben. Nachfolgend werden diese gesichtsspezifischen Analysedaten zur Erstellung eines 3D-Modells genutzt, indem damit ein generisches 3D-Gesichtsmodell an das Gesicht angepasst wird. Hierzu werden festgelegte Korrespondenzpunkte mittels der Daten aus der Gesichtsbildanalyse verschoben, sodass sich die Gesichtsform des Modells an die des Eingabegesichts angleicht. Der Anpassungsvorgang wird in einem externen Programm für 3D-Grafik durchgeführt, indem die Modellierungsfunktionen des Programms automatisch per Skript ausgeführt werden.

Diese Arbeit beschäftigt sich ausschließlich mit der Erstellung und Bereitstellung der oben genannten gesichtsspezifischen Analysedaten in einer Form, die für deren programminterne Weiterverarbeitung zur Modifikation erforderlich ist. Die Datengewinnung erfolgt mittels der Bildanalyse der Frontalansicht eines Gesichts. Für die Umsetzung dieser Analyse werden die Hilfsmittel einer Programmbibliothek für Bildverarbeitungsalgorithmen verwendet.

Das Verfahren zur Gesichtsbildanalyse ist teilautomatisiert. Es setzt sich zusammen aus der manuellen Markierung der Suchbereiche auf dem Eingabebild und der anschließenden vollautomatisierten Merkmalsanalyse innerhalb der markierten Bereiche. Für die Gesichtsmerkmalsanalyse wird ein von [Hähnel \(2008\)](#) (Modellbasierte posen- und mimikinvariante Gesichtserkennung) entwickeltes Verfahren genutzt und an die Anforderungen dieser Arbeit angepasst. Die Datengewinnung und die Art und Weise der Bereitstellung der Daten richtet sich nach den Vorgaben, die durch die Arbeit von [Burmeister \(2016\)](#) festgelegt werden.

Die nach diesem Konzept entwickelte Software ist die prinzipielle Lösung der beschriebenen Problemstellung. Sie bildet also das Grundprinzip für eine weitgehende Automatisierung des Prozesses der Gesichtsmodellierung. Deshalb ist sie zunächst nur für die Verarbeitung stereotypischer Eingabebilder ausgelegt, sodass Eingabefotos von Gesichtern mit Anomalien nicht vollständig oder gar nicht verarbeitet werden können. Allerdings wird der Programmcode so aufgebaut, dass er offen für Erweiterungen und Verbesserungen zu einem späteren Zeitpunkt ist.

2 Gesichtsanalyse in digitalen Bildern

Die Gesichtsanalyse in den digitalen Eingabebildern der Software ist das Kernstück dieser Arbeit. Deshalb werden alle dazu notwendigen Grundlagen innerhalb dieses Kapitels aufgegriffen.

Das erste Unterkapitel beschäftigt sich mit Verfahren zur Gesichtsanalyse und beleuchtet die Vorgehensweise dafür innerhalb dieser Arbeit. Das nächste Unterkapitel stellt alle Grundlagen der Bildverarbeitung vor, die für die Implementierung der Gesichtsanalyse wichtig sind. Dazu gehört die Beschäftigung mit den Eigenschaften des Analysematerials sowie die einzelnen Aufgaben der Bildverarbeitungskette.

2.1 Verfahren zur Gesichtsanalyse

Die computergestützte Analyse eines Bildes wird dazu genutzt, Aussagen über Bildinhalte zu treffen. Sie ist dann erforderlich, wenn das Bildmaterial nicht durch Menschen eingesehen und beurteilt werden kann oder soll. Dies gilt für Daten einer Überwachungskamera, den Abgleich eines Bildes mit den Bildern einer Datenbank oder den Videodaten eines Roboters mit künstlicher Intelligenz.

In vielen Fällen dient die Bildanalyse der anschließenden Objekterkennung und -klassifizierung. Diese sind kein triviales Problem und erfordern sehr individuelle Herangehensweisen, die jedoch nicht immer stabil und zielführend sind. Die Herausforderung liegt in der unterschiedlichen Art und Weise, wie Menschen oder Computer Bildinformationen erfassen und auswerten. Grundsätzlich fallen dem Menschen die Aufgaben Zuordnung und Einschätzung leicht, weil das menschliche Gehirn darauf trainiert ist, nach Kriterien, die den menschlichen Sinnen angepasst sind, Zusammenhänge zu erkennen, Unwichtiges auszublenden und Ähnliches zu gruppieren. Der Computer hingegen ist gänzlich objektiv und logisch und kann das Einordnen von Informationen nach menschlichen Kriterien höchstens anhand von Algorithmen, die von Menschen entwickelt wurden, nachahmen. Die Stärken der rechnergestützten Bildverarbeitung liegen vielmehr in der quantitativen Verarbeitung von Informationen, wie dem Zählen von Objekten oder dem Bestimmen von Pixelwerten. Aus diesen Gründen ist es sinnvoll, den Computer, wenn möglich, nur für letztere Aufgaben einzusetzen und alle Erkennungs- und Klassifizierungsaufgaben durch manuelle Eingaben zu lösen.

Dies gilt auch für die Gesichtsanalyse in Bildern. Das menschliche Gehirn ist im Bereich der Gesichtserkennung so hoch entwickelt, dass es die Gesichter anderer Menschen innerhalb von Millisekunden erkennen, zuordnen und ihre Mimik deuten kann.

2 Gesichtsanalyse in digitalen Bildern

Ein computerbasiertes System zur vollautomatisierten Gesichtserkennung und/oder -analyse hingegen kann diesen Vorgang nur bedingt simulieren. Die Schwierigkeit dabei liegt vor allem in der Variabilität der Abbildung von Gesichtern. Damit gemeint ist die Posen- und Mimikvarianz des Gesichts, aber auch die Position des Gesichts innerhalb eines Bildes.

Für Anwendungsbereiche, die abhängig von einer Echtzeitanalyse von Gesichtern in diversem Bildmaterial sind (beispielsweise bei der Interaktion von Mensch und Roboter mit künstlicher Intelligenz), gibt es inzwischen verschiedene Verfahren, die bei unterschiedlichen Bedingungen optimale Analyseergebnisse erzielen. Dafür verwenden sie modellbasierte und ansichtenbasierte Ansätze. Beide basieren auf der Erstellung eines Klassifikators, der Gesichter analysiert und dadurch erkennt und zuordnet. Erstere verwenden dafür die Bilder einer Gesichtsdatenbank, mithilfe derer der Klassifikator trainiert wird. Letztere hingegen bedienen sich dazu eines Gesichtsmodells, dessen Parameter zur Erkennung und Zuordnung eines Gesichts an dessen spezifische Werte angepasst werden. Aus beiden Bereichen gibt es favorisierte Verfahren (wie die Active Appearance Models, das Elastic Graph Matching und die Independent Component Analysis), die aber die Problematik von Fehldetektionen oder schlechten Erkennungsraten nicht gänzlich ausräumen. Diese Problematik lässt sich nur unter kontrollierten Bedingungen minimieren, indem vorab Faktoren, die Probleme hervorrufen könnten, verringert werden. Derartige Einschränkungen sind für die Gesichtsanalyse innerhalb dieser Arbeit realisierbar, insbesondere weil keine Echtzeitanalyse erforderlich ist.

Es stellt sich die Frage, ob eine vollautomatische Gesichtsanalyse für die Datenerfassung des hier entwickelten Verfahrens überhaupt notwendig und sinnvoll ist oder inwieweit nur Teile der Analyse automatisiert werden sollten.

Die folgende Tabelle stellt die Vor- und Nachteile der vollautomatischen und der manuellen Gewinnung von gesichtsspezifischen Daten für die Zielsetzung der Vereinfachung des 3D-Modellierungsprozesses von Gesichtern gegenüber.

	Vollautomatisch	Manuelle
Anzahl an Fehldetektionen	hoch	niedrig
Erkennungsrate	variabel	hoch
Datenbank erforderlich	ja	nein
Benutzeraufwand	niedrig	hoch
Rechenaufwand	hoch	niedrig
Datenqualität (Genauigkeit)	variabel	variabel
Störanfälligkeit bei diversem Bildmaterial	hoch	niedrig

Tabelle 2.1: Automatische vs. Manuelle Datenerfassung - Legende: Grün = positiv zu bewerten, Rot = negativ zu bewerten, schwarz = variierende Bewertung

Aus der Bewertungstabelle wird deutlich, dass für die hier aufgeführten Kriterien, die relevant für die Entwicklung einer Modellierungssoftware sind, vieles für die manuelle Datenerfassung spricht. Deshalb wird sie für bereits erhältliche Software für die Gesichtsm modellierung (Faceworx, Faceshop, FaceGen) genutzt, obwohl damit ein hoher Benutzeraufwand verbunden ist. Diese Arbeit verfolgt jedoch das Ziel, den Benutzeraufwand zu minimieren, ohne die Ergebnisqualität zu mindern. Zu diesem Zweck kann die Kombination von wenigen manuell eingegebenen Benutzerdaten, die eine Initialisierung automatisierter Analyseprozesse zulässt, zielführend sein.

Durch eine halbautomatische Gesichtsanalyse würden die Vorteile beider Verfahren ausgenutzt und eine gute Kosten-Nutzen-Bilanz erreicht, in Bezug auf die Komplexität der Programmbedienung sowie auch der Implementierung und des Rechenaufwands im Verhältnis zu den daraus resultierenden Ergebnissen. Aus diesem Grund erfolgt die Gesichtsanalyse innerhalb dieser Arbeit als Kombination der manuellen Einschränkung des Suchbereichs durch den Benutzer und der anschließenden vollautomatisierten Gesichtsanalyse innerhalb der markierten Bereiche.

2.2 Grundlagen der Bildverarbeitung

Zum besseren Verständnis der in folgenden Kapiteln beschriebenen Verfahren zur Gesichtsanalyse und der darin vorgenommenen Gesichtsmerkmalslokalisierung werden in diesem Unterkapitel alle dafür nötigen Grundlagen der Bildverarbeitung aufgegriffen. Der Begriff Bildverarbeitung bezeichnet alle Arten der Speicherung und Aufbereitung visueller Informationen. Da Bilder als zweidimensionale Signale aufgefasst werden können, bedient sich die Bildverarbeitung der Mittel der Signalverarbeitung. Die Gliederung des Kapitels orientiert sich an der Bildverarbeitungskette, die in Abbildung 2.1 dargestellt ist.



Abbildung 2.1: Bildverarbeitungskette

Der letzte Schritt der Kette, die Objekterkennung, ist die Zuordnung der detektierten Objekte zu den entsprechenden Objektklassen. Dieser Schritt ist für diese Arbeit nicht notwendig, weil die Bildverarbeitung in dieser Arbeit der Merkmalsanalyse von bekanntem Bildinhalt dient, sodass die Objekte nicht zugeordnet werden müssen. Allerdings erfolgt eine Klassifikation von einzelnen Pixeln, auf die im [autorefBildanalyse](#) näher eingegangen wird.

2.2.1 Beschaffenheit des Eingabematerials der Analyse

Eine nahezu fehlerfreie, zuverlässige, automatische Erkennung von bestimmten Objekten und in Bezug auf die Gesichtsanalyse von bestimmten Gesichtsmerkmalen in diversem Bildmaterial hängt oftmals vom Einsatz eines Systems unter kontrollierten Bedingungen ab. Zum einen ist es deshalb von Bedeutung, Vorkenntnisse über die Eigenschaften der gesuchten Objekte zu erlangen und diese für den Erkennungsprozess zu nutzen. Zum anderen gibt es verschiedene Faktoren, die die Erkennungsleistung von automatischen Erkennungssystemen beeinflussen und die es zu reduzieren gilt, um die Treffsicherheit zu optimieren. Die Reduzierung einiger Einflussfaktoren kann nur algorithmisch angegangen werden, weil sie Bildeigenschaften betreffen, die im Vorwege nicht ausgeschlossen werden können. Die Minimierung anderer Faktoren lässt sich ermöglichen, indem Vorkehrungen getroffen werden wie der Einsatz der richtigen Hardware oder Einschränkungen für die Beschaffenheit des Analysematerials.

Für das in dieser Arbeit entwickelte Analyseverfahren beziehen sich diese Vorkehrungen auf den Einfluss, den der Benutzer auf die Beschaffenheit der Eingabedaten hat und inwiefern dieser Einfluss genutzt werden kann, um das Material an die Programmanforderungen anzupassen. Durch die Auswahl eines geeigneten Fotos hat der Nutzer damit maßgeblichen Einfluss auf eine erfolgreiche Datenanalyse und eine zielführende, anschließende Weiterverarbeitung dieser Daten. Das dafür notwendige Wissen um alle relevanten und allgemeinen Eigenschaften von Gesichtern sowie auch die genaue Spezifizierung der Anforderungen an das Analysematerial wird in folgenden Unterkapiteln behandelt.

2.2.1.1 Eigenschaften von Gesichtern

Für den automatisierten Teil der Gesichtsmerkmalsanalyse sind die äußerlich sichtbaren Eigenschaften von Gesichtern grundlegend. Deshalb befasst sich der folgende Text mit allen Merkmalen von Gesichtern, die das Erscheinungsbild betreffen. Alle Eigenschaften in Bezug auf physiologische und biologische Funktionen von Gesichtern werden außer Acht gelassen. Die Beschreibung und Erläuterung des Aussehens menschlicher Gesichter gliedert sich dabei in zwei Kategorien: allgemeine und spezifische Merkmale.

Die *allgemeinen Merkmale* umfassen alle Eigenschaften, die zwar auf das Gesicht zutreffen, aber nicht spezifisch dafür sind. Hierzu gehören alle Merkmale, die durch die Zugehörigkeit des Gesichts zu einem menschlichen Organismus begründet sind wie beispielsweise die Knochenstruktur des Kopfes und die Beschaffenheit von Haut und Haaren in Bezug auf die Farbgebung und Oberflächenstruktur des Gesichts.

Die *spezifischen Merkmale* umfassen alle nur das Gesicht charakterisierenden Eigenschaften wie den typischen Aufbau eines Gesichts mit der Ausprägung der Gesichtsmarkkmale und deren geometrischen Verhältnissen. In der Erläuterung beider Kategorien wird sowohl auf die generischen Eigenschaften als auch auf mögliche spezifische Abweichungen davon eingegangen.

Prägend für das aktuelle Erscheinungsbild eines Gesichts ist auch der jeweilige Gesichtsausdruck in einem Moment, wobei die Eigenschaft der Gesichtsmimik wie bereits angemerkt in Bezug auf die Zielsetzung dieser Arbeit gesondert zu betrachten ist und deshalb innerhalb dieses Unterkapitels nicht berücksichtigt wird. Ihr Einfluss auf die Bildanalyse wird im Abschnitt 2.2.1.2 über die menschlichen Einflussfaktoren auf das Analysematerial aufgegriffen.

Allgemeine Merkmale:

1. Die *Knochenstruktur des Kopfes* hat für das Gesicht eine formgebende Funktion und kann (abgesehen von Wachstumsprozessen über einen längeren Zeitraum) als unveränderlich und unbeweglich angenommen werden. Die Knochen des Schädelskeletts sind im Erwachsenenalter fest miteinander verbunden, mit Ausnahme der gelenkig verbundenen Ober- und Unterkieferknochen. Wie auch der restliche Körper weist der Schädel eine vertikale Achsensymmetrie auf, so dass beide Gesichtshälften annähernd als Spiegelung der anderen gesehen werden können.

Erleidet ein Mensch Verletzungen am Schädelskelett, können diese auch äußerlich sichtbare Auswirkungen haben, die insofern das Erscheinungsbild verändern, als dass sie vom typischen Aussehen einer Gesichtsförm abweichen. Dazu gehören zum Beispiel ein Nasenbeinbruch oder ein Gesichtsschädelbruch, die beide Einfluss auf die Gesichtssymmetrie haben können, aber auch seltenere, schwerere Verletzungen, bei denen ganze Teile des Gesichtsknochens fehlen.

2. Die *Beschaffenheit von Haut und Haaren* ist ein zentrales Merkmal für das äußere Erscheinungsbild eines Gesichts. Als Haut bezeichnet man alle Gewebeschichten, die die Schädelknochen bedecken, sichtbar ist jedoch nur die Epidermis als oberste Gewebeschicht. Sie hat verschiedene Eigenschaften, die das Aussehen maßgeblich beeinflussen. Die Stärke der Gesichtshaarung ist eine dieser Eigenschaften. Mit dem Bartwuchs prägt sie insbesondere das männliche Gesichtsbild. Zwischen dem erblich bedingten Farbton von Haut und Haaren besteht ein Zusammenhang, jedoch variiert beides von Mensch zu Mensch. Die Farbe der Haut wird in der Kunst als inkarnat bezeichnet und setzt sich bei subtrak-

tiver Farbmischung aus Rot, Weiß und verschiedenen Brauntönen zusammen. Die Anteile der Farben variieren je nach Hauttyp sowie auch stellenweise und zeitabhängig, je nach aktuellem Durchblutungs- und Bräunungsgrad. Abgesehen von der Hautfarbe wird das Hautbild jedes Menschen spezifisch auch durch regionale Verfärbungen bestimmt, beispielsweise durch Leberflecken, Sommersprossen oder Altersflecken. Eine weitere Eigenschaft von Haut ist die elastische Verformung während der Bewegung der darunterliegenden Muskelschichten. Sie führt aufgrund der unnachgiebigen Schädelknochen im Kopfinneren bei verzogener Mimik zu einer Faltenbildung der Gesichtshaut, die altersabhängig auch bei neutralem Gesichtsausdruck dauerhaft auftritt und ebenfalls das Aussehen des Gesichts beeinflusst.

Außer diesen natürlichen, äußeren Unterschieden des Gesichts verschiedener Menschen gibt es auch Unterschiede im Erscheinungsbild von Haut und Haar aufgrund krankheitsbedingter Veränderungen. Verletzungen, die Narben hinterlassen, Hautausschläge, Geschwülste, Haarausfall oder Albinismus sind Beispiele für solche Veränderungen und führen dazu, dass das Aussehen des Gesichts von seiner typischen Gestalt abweicht. Ebenfalls gesondert zu betrachten ist der Fall, dass die Gesichtshaut einer Person ungewöhnlich viele Fetteinlagerungen hat, wodurch die Gesichtsförmigkeit beeinflusst wird. Bei dickleibigen Menschen verliert dadurch die Knochenstruktur als formgebendes Kriterium seine Bedeutung, weil sich das über den Knochen liegende Fettgewebe formverändernd auswirkt. Auch eine Lähmung des Muskelgewebes im Gesicht kann das Erscheinungsbild beeinträchtigen, indem zum Beispiel eine Gesichtshälfte erschlafft, wodurch die Gesichtssymmetrie beeinträchtigt wird.

Spezifische Merkmale:

1. *Gesicht:* Jedes Gesicht wird charakterisiert durch seinen typischen Aufbau mit allen enthaltenen sichtbaren, anatomischen, gesichtsspezifischen Merkmalen und deren Proportionen. Dieser Abschnitt befasst sich mit dem Gesichtsaufbau und beschreibt und erläutert dessen spezifische Merkmale.

Als Gesicht wird die Vorderseite des menschlichen Kopfes bezeichnet. Sie wird an den Seiten jeweils von den Ohren und oben und unten von Haaransatz und Kinn begrenzt. Die Gesichtsförmigkeit entspricht der frontalen Gesichtskontur. Sie basiert auf fünf verschiedenen Grundformen: oval, rechteckig, rund, herzförmig und trapezförmig. Die Gesichtsförmigkeit eines Menschen verändert sich im Laufe seines Lebens aufgrund von Wachstums- und Alterungsprozessen. Männliche und weibliche Gesichter unterscheiden sich in ihrer Form und der Ausprägung der Gesichtsmarkierungen.

2 Gesichtsanalyse in digitalen Bildern

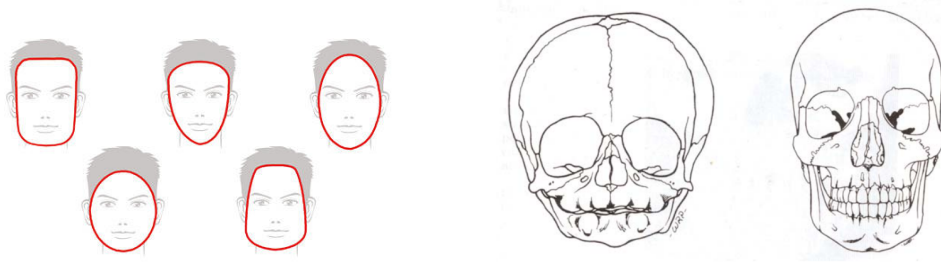


Abbildung 2.2: Grundformen von Gesichtern, Bitzer (2009), und Unterschiede zwischen Schädeln von Kindern und Erwachsenen, Enlow (1996)



Abbildung 2.3: Unterschiede in Gesichtern von Männern und Frauen, Grammer (2000)

Der grundsätzliche Gesichtsaufbau ist immer derselbe. In der Literatur werden keine allgemein gültigen Proportionen dafür aufgezeigt, häufig wird hingegen die hohe Variabilität der Ausprägung der Gesichtszüge genannt und das Schönheitsideal des goldenen Schnitts für dessen anatomische Verhältnisse angeführt. Der goldene Schnitt ist ein Verhältnis, das in der Natur häufig auftritt und allgemein hin als ästhetisch gilt. Nach dem goldenen Schnitt verhält sich das kürzere Stück einer unterteilten Strecke zum längeren wie das längere zur Gesamtstrecke.

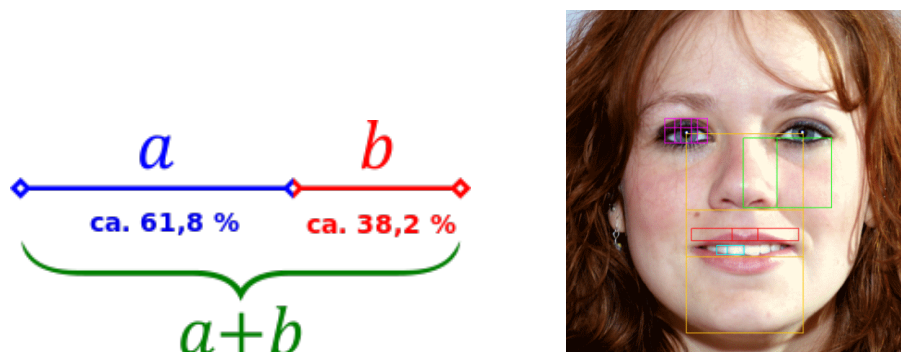


Abbildung 2.4: Goldener Schnitt, Stanner (2012), PhiPoint Solutions

2. *Augen:* Die Augen befinden sich im oberen Gesichtsteil und können als zentrales Merkmal des Gesichts bezeichnet werden, weil sie für die menschliche Kommunikation von enormer Bedeutung sind und deshalb für den Betrachter einen entsprechend hohen Stellenwert einnehmen. Geöffnet haben sie eine mandelförmige äußere Kontur, die durch die Wimpern an den Augenrändern hervorgehoben wird. Innerhalb der Augen befindet sich die runde Iris, in der zentrisch die ebenfalls kreisförmige Pupille liegt. Die Augen unterscheiden sich von den anderen Gesichtsmerkmalen vor allem durch ihren komplexen Aufbau und die besondere Farbgebung. Der gesamte Augapfel weist mit der weißen Lederhaut, der farbigen Iris und der dunklen Pupille im Unterschied zum restlichen Gesicht einen sehr geringen Rotanteil, aber einen großen Kontrastumfang auf. Die Farbe der Iris variiert dabei von Mensch zu Mensch und entstammt einem der Bereiche Braun, Blau oder Grün sowie zahlreichen Mischvarianten. Bei Menschen mit einer erblich bedingten Pigmentstörung erscheint die Iris jedoch rot.
3. *Nase:* Unterhalb der Augen und in der Mitte des Gesichts befindet sich die Nase. Sie vermittelt dem Betrachter durch ihre zentrale Position und starke räumliche Ausprägung einen Tiefeneindruck. Die Nase ist hautfarben und wird durch die Symmetrieachse des Gesichts, auf der sie liegt und die seitlich positionierten, dunkler erscheinenden Nasenlöcher charakterisiert. Die Form der Nase ist trapezförmig und nimmt von Mensch zu Mensch verschieden viel Raum ein bezüglich ihrer Breite und Höhe. Sie hat für das Erscheinungsbild des Gesichts eine strukturierende und formgebende Funktion, weil sie dem Betrachter Anhaltspunkte für eine geometrische Unterteilung des Gesichts gibt. In Bezug auf die Relevanz für die zwischenmenschliche Kommunikation ist sie jedoch zweitrangig und erhält bei der Gesichtsbetrachtung entsprechend weniger Aufmerksamkeit.
4. *Mund:* Der Mund liegt unterhalb der Nase und gilt neben den Augen als zweite wichtige Region im Gesicht, was in seiner Rolle für den Ausdruck von Emotionen und den Informationsaustausch zwischen Menschen begründet liegt. Charakteristisch für den Mund ist seine rote Farbe und seine horizontale Unterteilung in Ober- und Unterlippe. Die Grundform des Mundes ähnelt einem spitz zulaufenden Oval. Grundsätzlich unterscheidet sich die Form des Mundes bei Frauen und Männern stark. Der Mund ist durch seine hohe Beweglichkeit sehr formvariabel, sodass seine Form auch bei vermeintlich ausdruckslosem Gesicht geringfügig unterschiedlich erscheint.
5. *Augenbrauen:* Die Augenbrauen befinden sich oberhalb der Augen und spielen für den Ausdruck von Emotionen zwar eine erhebliche Rolle, treten durch ihre wenig hervorstechende Farbgebung im Betrachtungsprozess aber erst im Zusammenhang mit den Augen in den Vordergrund. Grundsätzlich kann von zwei getrennten Brauen gesprochen werden, deren Grundform annähernd einer horizontal liegenden und nach unten geöffneten Sichel entspricht, in einigen Fällen

sind die Augenbrauen aber auch oberhalb der Nase zusammengewachsen. Sie haben für das Erscheinungsbild eine form- und konturgebende Funktion, weil sie die obere Kante der Augenhöhlen bilden und das oberste der Gesichtsmerkmale sind. Ihre Farbe entspricht häufig der Farbe der Kopfbehaarung oder unterscheidet sich von dieser nur in Nuancen, es sei denn, die Haare einer Person ergrauen altersbedingt, wobei es sein kann, dass die Brauen zunächst noch ihre ursprüngliche Farbe haben.

6. *Kinnlinie*: Die Kinnlinie schließt das Gesicht nach unten hin ab und hat annähernd die Form eines nach oben geöffneten Polynoms sechsten Grades. Sie unterscheidet sich bei Männern und Frauen und von Mensch zu Mensch und tritt mehr oder weniger markant in Erscheinung. Es gibt Varianten von rund bis trapezförmig und von schmal bis breit. Grundsätzlich prägt sich die Kinnlinie erst innerhalb des Wachstums stärker aus und hat bei Kindern noch eine eher runde Form.

2.2.1.2 Anforderungen an die Bildbeschaffenheit

Der folgende Text widmet sich Spezifizierungen für das Eingabematerial, die für die Verarbeitung durch das Analyseprogramm und in Hinblick auf die Weiterverarbeitung der Analysedaten notwendig sind. Die Spezifizierungen ergeben sich aus allen Einflussfaktoren auf die Bildbeschaffenheit.

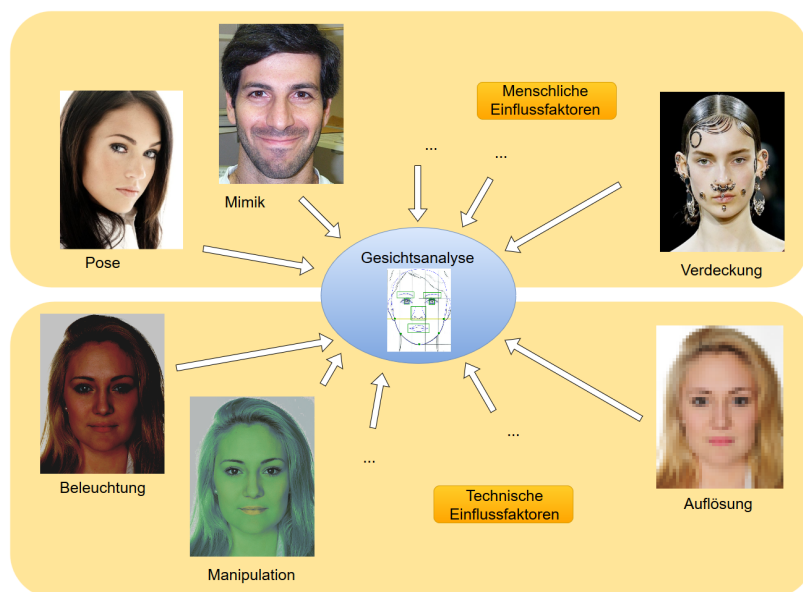


Abbildung 2.5: Einflussfaktoren auf die Gesichtsanalyse (Frau4, Frau gedreht, Frau Schmuck, Mann2 Mimik)

Deshalb erfolgt die Gliederung des Kapitels anhand der Unterteilung der Einflussfaktoren in die Kategorien Technische und Menschliche Einflussfaktoren, wobei die Menschlichen Einflussfaktoren nochmals unterteilt werden in biologische und verhaltensbedingte Einflussfaktoren. Das folgende Unterkapitel erläutert alle diese Faktoren, um dann auf die Art und Weise einzugehen wie mit dem Einfluss der Faktoren umgegangen werden muss, um den Erfolg der Bildanalyse zu gewährleisten.

Technische Einflussfaktoren:

Alle Einflussfaktoren auf die Erkennungsleistung eines Bildanalyse-Systems, die sich auf die technische Art und Weise der Entstehung oder Bereitstellung der Eingabedaten für das System beziehen, können als technische Faktoren zusammengefasst werden. Dazu gehören die eingesetzte Hardware und die Beleuchtungssituation während der Bildaufnahme sowie auch die anschließende Form der Bildbearbeitung und/oder Speicherung.

Für die Reduzierung ihres Einflusses auf die Erkennungsleistung ist es erforderlich, die verwendeten Geräte und den Aufnahmeort geeignet zu wählen und die Aufnahmedaten unverfälscht und unkomprimiert zu speichern. Dies erweist sich als schwierig, sobald auf bereits vorhandenes Analysematerial zurückgegriffen werden soll, sodass kein Einfluss mehr auf die Entstehungsbedingungen genommen werden kann. Die wichtigsten technischen Einflussfaktoren sind die Bildauflösung, das Bilddateiformat und die Bilddatenbeschaffenheit sowie die Beleuchtungssituation während der Bildaufnahme. Diese Einflüsse werden im folgenden Abschnitt kurz erläutert sowie die sich aus ihnen ergebenden Anforderungen an die Gesichtsbilder benannt, für die der im Zuge dieser Arbeit entwickelte Programmcode optimale Ergebnisse liefert.

1. *Bildauflösung:* Die Bildauflösung einer Grafik gibt an, wie viele Bildpunkte pro Länge existieren. Je höher also die Auflösung ist, desto größer ist die Unterteilung des Bildinhalts und desto genauer können einzelne Bildbereiche voneinander unterschieden werden. Ist die Bildauflösung sehr gering, wird die Erkennungsleistung eines Systems gemindert. Ist das Raster der Bildauflösung jedoch feiner als für die kleinste Einheit des Bildinhalts notwendig, so erhöht dies nur den erforderlichen Rechenaufwand, verbessert aber das Erkennungsergebnis nicht weiter. Es muss bei der Bildauflösung grundsätzlich immer zwischen verbesserter Erkennungsleistung und dem dafür notwendigem Rechenaufwand abgewogen werden. Aufgrund dieses Konflikts beschränkt/skaliert die innerhalb dieser Arbeit entwickelte Software die Bildhöhe jedes Eingabebildes auf eine Höhe von 800 Bildpunkten, wobei die zugehörige Bildbreite vom Seitenverhältnis des Originalbildes abhängt und anhand dessen berechnet wird. Das bedeutet, dass die Erkennungsleistung des Systems bei einer niedrigeren Auflösung des Quellbildes nicht voll ausgeschöpft werden kann, aber eine höhere Auflösung deshalb auch keine verbesserte Bildanalysequalität zur Folge hat. Für die Gewährleistung einer genügend genauen Auswahl der Gesichtsmerkmalsbe-

reiche durch den Benutzer darf die Höhe des Eingabebildes 200 Pixel nicht unterschreiten (wodurch sich die entsprechende spezifische Mindestauflösung ergibt).

2. *Bilddateiformat*: Das Bilddateiformat beschreibt den Aufbau der Datei, in der der Bildinhalt als digitale Daten vorliegt. Es gibt unzählige Dateiformate für Rastergrafiken. Jedoch haben sich nur einige davon durchgesetzt und werden gängigerweise von verschiedener Software unterstützt. Diese unterscheiden sich in der Art ihrer Kodierung, wodurch festgelegt ist, welche Eigenschaften das Dateiformat hat. Dazu gehören zum Beispiel, ob und welche Kompression verwendet wurde, die Anzahl der Kanäle, ob das Format Metadaten enthält und in welcher Farbtiefe die Bildinformation abgespeichert wurde. Die Wahl des Bilddateiformats ist also auch entscheidend dafür, in welcher Qualität die Bildinformation vorliegt. Wird beispielsweise eine verlustbehaftete Kodierung verwendet, sinkt dadurch die Bildqualität, wodurch auch die Chance auf eine erfolgreiche Bildverarbeitung durch ein Erkennungssystem gemindert wird. Das hier entwickelte Programm zur Gesichtsanalyse ist nur für die Verarbeitung von Farbbildern ausgelegt und arbeitet mit den verschiedenen Bilddateitypen, die durch die Verwendung der `imread`-Funktion von OpenCV vorgegeben sind. Dazu gehören alle gängigen Formate wie JPEG, Windows bitmaps, TIFF und Portable Network Graphics (PNG), aber auch die seltener verwendeten Formate wie JPEG2000, WebP, Portable Image Format und Sun rasters. Um die Erkennungsleistung des Systems nicht zu mindern, ist die Wahl eines Dateiformats, in dem die Bildinformation unkomprimiert oder verlustfrei komprimiert vorliegt sinnvoll. Dies sind zum Beispiel Windows Bitmaps-, PNG- und in der Regel auch TIFF-Dateien, JPEG-Dateien gehören nicht dazu.
3. *Bilddatenbeschaffenheit*: Die Beschaffenheit der Bilddaten bezeichnet in hiesigem Zusammenhang den Grad und die Art und Weise der vorherigen Anwendung von Bildbearbeitungsalgorithmen auf die Bilddaten, die als Analysematerial dienen sollen. Als Bildbearbeitung wird im Allgemeinen die Manipulation der Bilddaten zum Zwecke der visuellen Aufwertung oder der Erzeugung ästhetischer Effekte aus Sicht des Betrachters verstanden. Dieser Gesichtspunkt ist jedoch nicht deckungsgleich mit dem der Aufbereitung der Daten zwecks softwarebasierter Weiterverarbeitung. Die im Zuge dieser Arbeit angewandten Analysealgorithmen erwarten als initiale Eigenschaften des Eingabefotos für die Werte von Farbton, Sättigung, Helligkeit und ihrer statistischen Verteilung die typischen Bildwerte einer unveränderten Aufnahme eines ausreichend beleuchteten, unverrauschten und nach Weiß abgeglichenen Gesichts. Deshalb kann eine zuvor erfolgte Veränderung dieser Bilddaten die erfolgreiche Objekterkennung entweder fördern, einschränken oder sogar verhindern. Das bedeutet, grundsätzlich wirkt sich eine Verfälschung der Originalwerte eines Fotos negativ auf die Objekterkennung aus. Es sei denn, das Foto weicht vom erwarteten generischen

Typ ab und wird dann so verändert, dass es diesem Typ wieder gebracht wird. In diesem Fall begünstigt die Veränderung der Quellbilddaten den Analyseerfolg.

4. *Beleuchtungssituation*: Die Beleuchtungssituation während der Bildaufnahme hat entscheidenden Einfluss auf das Erscheinungsbild abgebildeter Objekte. Da die Bildinformation mithilfe des Lichts, das alle abzubildenden Objekte reflektieren, gewonnen wird, ist auch die Information über alle dabei existierenden Lichtquellen unausweichlich mit in der Bildinformation enthalten. Dies führt dazu, dass ein und dasselbe Objekt bei verschiedener Beleuchtung verschieden abgebildet wird. Die Schwierigkeit bei der Entwicklung von Objekterkennungsalgorithmen besteht darin, die Objektinformation von der Beleuchtungsinformation zu unterscheiden. Dafür müssen beispielsweise beleuchtungsunabhängige Merkmale und Klassifikatoren gewählt oder die Beleuchtungssituation modelliert und das Bild anhand dessen vorverarbeitet werden. Die Modellierung der Beleuchtungssituation ist aufwendig, weil sie die Kenntnis der Positionen der Lichtquellen oder die Rekonstruktion dieser Positionen aus mehreren verschiedenen Ansichten des jeweiligen Objekts erfordert. Deshalb basieren die Analysealgorithmen dieser Arbeit auf Grundlage beleuchtungsunabhängiger Merkmale. Diese Methode hat jedoch deutliche Grenzen, sodass sich zur Begünstigung der erfolgreichen Gesichtsanalyse ein neutral belichtetes, beleuchtungsinvariantes Foto empfiehlt.

Menschliche Einflussfaktoren:

Menschliche Einflussfaktoren auf die Gesichtsmerkmalserkennung eines Bildanalyse-Systems sind alle biologischen und spezifischen Unterschiede in Gesichtern sowie die verhaltensbedingten Veränderungen des Erscheinungsbilds des Gesichts.

Die *biologischen und spezifischen Unterschiede* sind für jede Person einzigartig, als gegeben hinzunehmen und können nicht verändert werden.

Die *verhaltensbedingten Veränderungen* des Erscheinungsbilds durch Bewegung und/oder dem Hinzufügen von Objekten sind dagegen temporär und deshalb veränderlich.

Im Folgenden wird erklärt, bei welchen menschlichen Faktoren für das in dieser Arbeit entwickelte Verfahren zur Gesichtsmerkmalserkennung Schwierigkeiten auftreten/aufreten können und wie eine Reduzierung aller verhaltensbedingten Faktoren erfolgen kann. Beides dient der Aufklärung des Benutzers im Umgang mit der Software. Die Reduzierung der verhaltensbedingten Faktoren ist dabei von besonderer Bedeutung, weil sie die Analyseleistung des Systems steigert und dadurch die Qualität der Analysedaten in Bezug auf deren Weiterverarbeitung beeinflusst und verbessert. Zum Teil bestehen auch Vorgaben für das Eingabematerial, die sich nur auf die Arbeit von [Burmeister \(2016\)](#) beziehen und deshalb nicht erläutert werden.

Biologische Einflussfaktoren sind alle für einen Menschen spezifischen Eigenschaften seines Gesichts, bedingt durch seine genetische Veranlagung, sein Alter und seine Gesundheitsgeschichte. Diese Faktoren können und sollen für die Zielsetzung dieser Arbeit nicht verändert werden, weil es gerade darum geht, ein bestimmtes menschliches Gesicht möglichst vorlagengetreu nachzubilden. Allerdings beeinflussen diese Eigenschaften durch ihre Unkontrollierbarkeit dennoch den Bildanalyseprozess, so dass einige Gesichter besser analysiert werden können als andere.

1. *Farbgebung*: Die Farbgebung des Gesichts und seiner einzelnen Merkmale kann Einfluss auf die Erkennungsleistung des Systems haben. Generell wirkt sich eine kontrastreiche Farbgebung positiv auf die Unterscheidbarkeit der Merkmale und damit auf ihre Lokalisierung aus. Augenfarbe, Lippenfarbe und Hautfarbe variieren von Mensch zu Mensch. Dabei heben sich die Lippenfarbe und die Haarfarbe mehr oder minder deutlich von der Hautfarbe der jeweiligen Person ab. Dies hat Einfluss auf die sichere Erkennung dieser Merkmale und kann innerhalb der in dieser Arbeit entwickelten Gesichtsanalyse zu Problemen führen. Die Analysealgorithmen greifen als Vorverarbeitungsschritt zum Teil auf Farbfilter zurück, um das Bildmaterial zu segmentieren und die Gesichtsmerkmale von der Haut zu separieren. Bei den Augenbrauen führt dies beispielsweise dazu, dass eine Erkennung erschwert ist oder gar nicht erfolgen kann, sobald Haut und Haarfarbe sich zu stark ähneln. Auch bei der Konturlokalisierung der Lippen treten Probleme in Bezug auf deren Genauigkeit auf, wenn sich die Lippen nicht ausreichend von der umgebenden Haut differenzieren. Die Software ist grundsätzlich nur für hellhäutige Gesichtsbilder ausgelegt, eine richtige Lokalisierung bei dunkelhäutigen Menschen ist unwahrscheinlich.
2. *Anatomie des Gesichts*: Der Aufbau jedes Gesichts ist grundsätzlich der selbe, dennoch unterscheidet sich bei jedem Gesicht die Lage der einzelnen Merkmale zueinander, ebenso wie deren Proportionen und Ausprägung. Auch diese unkontrollierbaren Voraussetzungen haben im Allgemeinen Einfluss auf die automatische Erkennung der Merkmale, jedoch wird dieser Einfluss bereits durch das Konzept der vorliegenden Arbeit auf ein Minimum beschränkt, weil der Nutzer der Software die Merkmalsbereiche markiert, wodurch die Eingabedaten des Programms den Suchbereich für jedes Merkmal vorgeben. Dennoch können auch die Proportionen und die Ausprägung innerhalb eines Merkmals vor- oder nachteilig für eine erfolgreiche Merkmalsextraktion sein, beispielsweise wenn die spezifischen Eigenschaften des Merkmals ungewöhnlich stark von durchschnittlichen Werten abweichen, die der Suchalgorithmus als Anhaltspunkte für die Suche verwendet (Entsprechendes Vorgehen wird im Kapitel Merkmalslokalisierung ausführlicher behandelt). Die Gesichtsmerkmalserkennung des in dieser Arbeit entwickelten Systems wird durch die Gesichtsanatomie nur in seltenen Fällen beeinflusst. Allerdings hat der Aufbau des Gesichts große Auswirkungen auf die Qualität der anschließenden Modifikation des generischen Kopfmodells.

Dieser wird jedoch nicht hier, sondern in [Burmeister \(2016\)](#) Bachelorarbeit behandelt.

3. *Geschlecht*: Das Geschlecht einer Person hat zwar Einfluss auf die Ausprägung der Gesichtsmarkmale, jedoch wird für das hier entwickelte Verfahren keine Beeinflussung der automatischen Erkennung der jeweiligen Merkmale erwartet, zumindest nicht in Bezug auf die Gesichtsanatomie. Allerdings beeinflusst das Geschlecht die Ausprägung der Gesichtshaarung, die signifikanten Einfluss auf die Lokalisierung der Kinnlinie oder der Lippen hat. Dieser Faktor kann indirekt auch als verhaltensbedingt eingestuft werden, da er durch eine Rasur reduzierbar ist. Außerdem hat das Geschlecht auch indirekten Einfluss auf die Merkmalslokalisierung, weil es ein geschlechtstypisches Styling zur Folge haben kann. Da dies jedoch sicher den verhaltensbedingten Faktoren zuzuordnen ist, wird erst im dazugehörigen Abschnitt darauf eingegangen.
4. *Verletzungen/Krankheiten*: Verletzungen und Krankheiten, die das Gesicht in Form und/oder Gestalt verändern können, sofern sie ein relevantes Gesichtsmerkmal betreffen, haben einen Einfluss auf die Erkennungsleistung der innerhalb dieser Arbeit entwickelten Software. Die Lokalisierungsalgorithmen für die Gesichtsmarkmale arbeiten anhand der Vorkenntnis der jeweilig als typisch angenommen spezifischen Eigenschaften dieser Merkmale in Bezug auf Form, Symmetrie, Farbe und Helligkeit. Treffen diese Eigenschaften aufgrund einer Verletzung oder Krankheit der abgebildeten Person nicht mehr zu, wird anhand falscher Annahmen nach ihnen gesucht. Dadurch kann die Suche nicht zielführend sein. Wie bereits erwähnt, ist die Erkennungsleistung des Systems durch die Kontrollierbarkeit der Eingabedaten beschränkt, sodass für Sonderfälle ein befriedigendes Ergebnis unwahrscheinlich ist.
5. *Alter*: Während des Alterungsprozesses eines Menschen verändert sich das Erscheinungsbild seines Gesichts fortlaufend. Zum einen verändert sich die Form des Schädels von der Kindheit bis zum Erwachsenenalter, aber auch Haut und Haar ändern ihr Aussehen in Abhängigkeit vom Alter einer Person. Die Veränderung der Schädelform sollte kaum oder keinen Einfluss auf die Erkennungsleistung des Systems haben, weil sowohl die kindliche als auch die erwachsene Gesichtsforn grundsätzlich den gleichen Aufbau hat und die gleichen Merkmale enthält, sodass das System in beiden Fällen mit den gleichen Annahmen arbeiten kann. Die altersbedingte Veränderung von Haut und Haar kann die Erkennungsleistung des Systems jedoch beeinflussen, weil zum Beispiel die faltige oder schlaffe Haut, das fleckige Hautbild oder das weiße Haar eines älteren Menschen andere Farb- und Kontrasteigenschaften hat als ein typisches Gesicht eines Menschen im mittleren Alter. Deshalb kann bei der Analyse von Gesichtsbildern älterer Menschen die Erkennungsleistung gemindert sein.

Verhaltensbedingte Einflussfaktoren beziehen sich auf den aktuellen Zustand des Gesichts bei dessen Abbildung, das heißt, sie bezeichnen alle temporären Eigenschaften eines menschlichen Gesichts, die durch das Verhalten des Menschen bedingt sind. Für die Erstellung von möglichst zielführendem Analysematerial gilt es, diese Faktoren auf ein Minimum zu reduzieren, um jegliche Störung des Erkennungsprozesses auszuklammern und eine optimale Ausgangssituation für die anschließende Gesichtsmodellierung zu gewährleisten. Dies ist natürlich nur möglich, wenn auf die Erzeugung der Analysefotos Einfluss genommen werden kann und nicht auf bereits vorhandenes Material zurückgegriffen werden muss.

1. *Mimik*: Die Mimik oder auch der Gesichtsausdruck verändert maßgeblich das aktuelle Erscheinungsbild eines Gesichts. Durch die Bewegung der Gesichtsmuskulatur werden einzelne Gesichtspartien und ihre Merkmale verschoben, verformt oder auch verdeckt sowie die Symmetrieeigenschaft des Gesichts beeinträchtigt. Hierdurch wird einerseits der Erkennungsprozess der Gesichtsmerkmale beeinflusst (aufgrund der bereits im Abschnitt über Verletzungen und Krankheiten genannten Eigenschaften der Lokalisierungsalgorithmen), andererseits bewirkt dies auch eine verfälschte Darstellung des Gesichts bei seiner Modellierung als 3D-Modell, weil die Berechnung der Korrespondenzen auf Grundlage der Gesichtsform während eines bestimmten Gesichtsausdrucks erfolgt. Werden diese Daten bei ihrer Weiterverarbeitung auf das generische Kopfmodell mit neutralem Gesichtsausdruck angewandt, entsteht bei der Modellierung eine Verzerrung, weil der mimikverzogene Zustand des Gesichts nicht mit dem ausdruckslosen Gesicht des Gesichtsprototypen übereinstimmt. Hierauf wird jedoch nicht näher eingegangen, weil dies nicht Teil dieser Arbeit ist. Zur Bereitstellung von optimal auf das hier entwickelte System abgestimmtem Eingabematerial muss die Abbildung des Gesichts bei neutralem Gesichtsausdruck und offenen Augen mit nach vorn gerichtetem Blick, mit weitgehend oder gänzlich geschlossenem Mund und unbewegter Mimik erfolgen, weil nur dann garantiert wird, dass sowohl der Erkennungsprozess fehlerfrei durchlaufen werden kann als auch die Ausgabedaten der Analyse zielführend sind.
2. *Ausrichtung*: Die Software erwartet als Eingabe die Frontalaufnahme eines Gesichts. Die Abbildung eines anders ausgerichteten Gesichts führt zwangsläufig zu Schwierigkeiten bei der Gesichtsanalyse und falsch berechneten Korrespondenzen für die nachfolgende Modellierung. Ist der Kopf im Extremfall seitlich abgebildet, können die Merkmale, die auf der anderen Kopfseite liegen nicht oder nicht vollständig erkannt werden.

Auch ein geringerer Rotationswinkel für die Aufnahme hat negative Auswirkungen auf die Merkmalerkennung und -lokalisierung. Die Algorithmen des Erkennungsprozesses arbeiten wie schon in vorherigen Abschnitten erwähnt, häufig mit der Vorkenntnis über sehr wahrscheinlich zutreffende Eigenschaften der gesuchten Merkmale. Die Veränderung des Rotationswinkels hat zur Folge, dass

erwartete Eigenschaften wie Symmetrie und Form ebenfalls verändert sind, wodurch die angenommenen Suchkriterien für die gedrehte Aufnahme nicht mehr zutreffen. Ein rotiertes Gesicht hat durch die 3D- zu 2D-Projektion bei der Aufnahme eine Formänderung der abgebildeten Merkmale zur Folge. Dadurch hat eine rotierte Ansicht die gleichen Auswirkungen auf die Weiterverarbeitung der Analysedaten wie auch ein durch Mimik verzerrtes Gesichtsbild. Wie bereits oben aufgeführt, führt eine nicht übereinstimmende Ausgangsposition von Gesicht und generischem 3D-Kopfmodell bei der Modifikation zu einer verzerrten Gesichtsdarstellung. Für die Begünstigung eines optimalen Ergebnisses muss eine möglichst frontale Aufnahme des zu modellierenden Gesichts gewählt werden.

3. *Verdeckung*: Es liegt auf der Hand, dass für die vollständige Erkennung und Lokalisierung aller Gesichtsmerkmale auch deren vollständige Abbildung auf dem Foto erforderlich ist.

Das Eingabefoto des Analysesystems muss deshalb den gesamten Gesichtsausschnitt ohne verdeckte Anteile abbilden. Gesichtsfremde Elemente, die im Gesichtsfeld liegen, wie Brillen, Schals, Schmuck, Kosmetik oder Tätowierungen, aber auch körpereigene Materialien wie Kopfbehaarung und Bärte oder andere Körperteile wie Arme und Hände verhindern eine korrekte Gesichtsmerkmalserkennung sowie die Berechnung der Lage der Korrespondenzpunkte.

2.2.2 Bildvorverarbeitung

Die Bildvorverarbeitung ist eine aufgabenbezogene Anpassung oder Korrektur der Bildinformation. Ihr Zweck ist die visuelle Bildverbesserung für den Betrachter oder die Aufbereitung der Daten für weitere Verarbeitungsstufen der rechnergestützten Mustererkennung. Innerhalb dieser Arbeit dient sie ausschließlich letzterem Anwendungszweck. Grundsätzlich kann zwischen zwei Verfahrenstypen unterschieden werden, solchen, anhand derer ein neues/verändertes Bild erzeugt wird, und denen, die Informationen über das Bild liefern.

Die Methoden der Bildvorverarbeitung umfassen zum einen verschiedene Transformationen, die die Daten in eine Form überführen, die sich besser für folgende Verarbeitungsschritte eignet, und zum anderen lineare und nicht lineare Filteroperationen, die relevante Bildinhalte hervorheben oder redundante Bildinhalte unterdrücken. Beide Vorverarbeitungsschritte werden in den folgenden Unterkapiteln kurz erläutert.

2.2.2.1 Bildtransformationen

Eine Bildtransformation überführt die Bildinformation in ein vom Originalraum verschiedenes Basissystem, wodurch die Daten anders repräsentiert werden und sich deshalb anders oder besser verarbeiten lassen. Viele Transformationen sind reine Umformungen, bewirken keinen Datenverlust und sind deshalb auch umkehrbar. Es

gibt jedoch auch Transformationen, die nicht invertierbar sind, weil sie verlustbehaftet sind. Gängige Transformationen, die in dieser Arbeit als Vorverarbeitungsschritt zum Einsatz kommen, sind Farbraum- und Koordinatentransformationen sowie die Berechnung von Histogrammen. Die Hough-Transformation wird gesondert betrachtet, weil sie kein Vorverarbeitungsschritt ist, sondern ein bestimmtes Verfahren der Bildsegmentierung darstellt und deshalb erst im dazugehörigen Unterkapitel erklärt wird.

Farbraumtransformationen rechnen die Farbmaßzahlen des Ausgangsfarbraums in die des Zielfarbraums um. Es gibt diverse Farbräume, deren Bildung teils durch technische Geräte und deren Signalverarbeitung und/oder durch eine sinnvolle, zweckgebundene systematische Beschreibung von Farben begründet ist. Die meisten Fotobilddaten liegen in der zur Darstellung auf Computer-Bildschirmen gängigen RGB-Form vor. Im Zuge dieser Arbeit erweist sich jedoch für einige Bildverarbeitungsschritte eine Konvertierung in andere Farbräume wie den HSV-, den YUV- oder den I1I2I3-Farbraum als notwendig, da die speziellen Eigenschaften des jeweiligen Farbraums für bestimmte Arbeitsschritte von Vorteil sind. Die jeweilige Bewandnis wird im entsprechenden Zusammenhang näher erläutert.

Koordinatentransformationen übertragen die Koordinaten von einem System in ein anderes. Sie verändern also nicht das Objekt, sondern das System, in dem es abgebildet wird, indem es zum Beispiel verschoben, gedreht, gespiegelt oder skaliert wird. In dieser Arbeit sind vor allem die Skalierung und die Verschiebung von Koordinaten von Bedeutung.

Die Berechnung eines *Histogramms* dient der statistischen Erfassung von Bildinformation und gibt Aufschluss über die Häufigkeitsverteilung von Pixelwerten eines Bildkanals. Histogramme von Farbbildern geben Auskunft über die Häufigkeitsverteilung der Helligkeitswerte innerhalb eines Bildes. Sie gehören zu den Punktoperator-Transformationen und modifizieren den Helligkeitswert jedes Pixels in Abhängigkeit des Helligkeitswertes selbst und teilweise auch in Abhängigkeit der Pixelposition im Bild.

Für diese Arbeit ist insbesondere der Histogrammausgleich von Bedeutung. Er bewirkt eine Kontrastverbesserung, die über eine Kontrasterhöhung hinaus geht. Dabei wird aus der Grauwertverteilung des Helligkeitskanals eines Bildes eine Gleichverteilung berechnet, was zur Folge hat, dass der gesamte, zur Verfügung stehende Wertebereich optimal ausgenutzt wird. Ein Histogrammausgleich ist also immer dann sinnvoll, wenn die Helligkeitswerte der Pixel sich möglichst gut voneinander unterscheiden und abheben sollen, weil anhand dessen in folgenden Verarbeitungsschritten Entscheidungen getroffen werden sollen.

2.2.2.2 Filteroperationen

Filteroperationen modifizieren Bilder dahingehend, dass informative Anteile hervorgehoben und störende reduziert werden. Filter sind typischerweise Nachbarschaftsoperationen. Der Unterschied zu den Punktoperationen liegt in der Größe der Region der Eingangsdaten, also darin, dass in die Berechnung eines Zielpixelwerts nicht nur sein Quellwert, sondern auch die benachbarten Pixelwerte einfließen.

Für die Bildanalyse in dieser Arbeit werden drei Filter angewandt, das Gaußfilter, das Medianfilter und der Sobeloperator. Die morphologischen Operationen Dilatation und Erosion sind keine Filter im eigentlichen Sinne, sondern eher der Bildmanipulation zuzuordnen. Sie sind deshalb gesondert zu betrachten, werden aber aufgrund ihrer Relevanz für spätere Verarbeitungsprozesse dennoch in diesem Kapitel mit aufgeführt.

Das *Gaußfilter* ist ein Frequenzfilter, das innerhalb der Bildverarbeitung als Weichzeichnungsfilter und zur Rauschreduzierung eingesetzt wird, das heißt, feine Bildstrukturen gehen verloren. Die spektrale Wirkung des Gaußfilters ist ein Tiefpass.

Das *Medianfilter* wird zur Rauschunterdrückung eingesetzt und gehört der Gruppe der Rangordnungsfilter an. Es ist ein nichtlineares Filter, das alle benachbarten Pixel einer Maske nach ihrer Größe sortiert und dann den aktuellen Pixelwert im Zentrum der Maske durch den mittleren Wert der Rangordnung ersetzt.

Das Medianfilter unterscheidet sich vom Mittelwertfilter dadurch, dass kein Mittelwert aus der Pixelumgebung berechnet wird, sondern lediglich der mittlere Wert der vorhandenen Pixel in dieser Umgebung ausgesucht wird. Dadurch hat das Medianfilter eine hohe Wirksamkeit und Stabilität gegenüber Ausreißern, ohne eine Kantenglättung, also eine Detailverfälschung der Bildinformation, zu bewirken. Das Medianfilter eliminiert binäres Rauschen, zeigt jedoch schlechte Wirksamkeit bei Gaußschem Rauschen.

Der *Sobeloperator* ist eine *Faltungsoperation*, die zur Kantendetektion in Bildern eingesetzt wird. Das Resultat des Algorithmus ist ein Gradientenbild, das hohe Bildfrequenzen, also ein starkes Helligkeitsgefälle, in Form von Grauwerten abbildet. Dazu werden im Originalbild die Richtungsableitungen der Helligkeitswerte in X- sowie in Y-Richtung mittels der jeweiligen Faltung mit einer 3x3-Matrix gebildet und deren Ergebnisse per Betragsbildung im Gradientenbild kombiniert.

Hohe Beträge bedeuten ein starkes Helligkeitsgefälle im Quellbild und verweisen deshalb auf darin abgebildete Kanten. Auf die Richtung des Gefälles kann geschlossen werden, indem der Steigungsquotient in die Arcustangens-Funktion eingesetzt und ein entsprechender Steigungswinkel berechnet wird. Der Sobeloperator ist ein Hochpassfilter, weshalb er einerseits sanfte Kanten nicht detektieren kann und was ihn andererseits anfällig für Bildrauschen macht. Deshalb dient er meist nur als Vorverarbeitungsstufe für komplexere Kantendetektionsalgorithmen.

Die *morphologischen Operationen Dilatation und Erosion* sind besondere Filter, weil der Bildinhalt durch sie strukturell verändert wird. Ihr Einsatz dient dem Zweck, die Form von Bildstrukturen zu analysieren oder zu verändern. Dazu wird ein Strukturelement verwendet, das einen Teilbereich des Ursprungsbildes darstellt. Dieser ist ein- oder mehrfach vollständig im Quellbild enthalten. Darin wird ein Bezugspixel gewählt und für den Vorgang des pixelweisen Verschiebens des Strukturelements über das Bild immer auf die Position des aktuell zu prüfenden Bildpunktes gelegt.

Im Fall der Erosion werden alle Pixel, an deren Position das Element jeweils vollständig mit dem Originalbild übereinstimmt, zur gesuchten Menge gezählt und alle Pixel, für die dies nicht zutrifft, auf einen festgelegten Wert gesetzt.

Im Fall der Dilatation hingegen wird die Umgebung jedes Pixels des Originalbildes, das mit dem Bezugspixel des Strukturelements übereinstimmt, um dessen Form erweitert. Morphologische Operatoren werden meist auf bereits segmentierte Binärbilder angewandt, um in diesen bestimmte Strukturen gegenüber anderen hervorzuheben.

2.2.3 Bildanalyse

Die Bildanalyse geht einen Schritt weiter als die Bildvorverarbeitung. Durch sie wird der Bildinhalt quantitativ klassifiziert und qualitativ bewertet, indem sie einen Übergang von der reinen Bilddarstellung zur Bildbeschreibung schafft. Dadurch können Aussagen über den Bildinhalt getroffen werden, zum Beispiel, ob ein bestimmtes Objekt abgebildet und wo es auf dem Foto zu finden ist.

Die Bildanalyse durchläuft die Schritte Segmentierung, und Merkmalsextraktion. Sie dienen normalerweise der anschließenden Objekterkennung und -klassifizierung. Innerhalb dieser Arbeit werden keine für zusammenhängenden Bildbereiche klassifiziert. Im Zuge der Konturlokalisierung der Gesichtsmerkmale erfolgt jedoch die Klassifizierung einzelner Pixel, indem für potentielle Pixel eine Zugehörigkeit zur Kontur entschieden wird oder nicht.

2.2.3.1 Segmentierung

Die Segmentierung ist ein Mittel zur Unterteilung der Bildpunkte in relevante und irrelevante Bereiche. Im Prinzip ist sie ein Verfahren zur Bildvereinfachung und zur Reduktion der Bildinformation auf das Wesentliche. Dafür wird für jedes Pixel festgestellt, welcher Gruppe es zuzuordnen ist. Eine vollständige Segmentierung des Bildes bedeutet eine überlappungsfreie Abgrenzung aller Objekte zum Hintergrund. Grundsätzlich ist das Ziel der Segmentierung die Erstellung einer Binärmaske, in der alle Objektpixel weiß und der Hintergrund schwarz erscheinen. Hierzu gibt es verschiedene Vorgehensweisen oder Methoden, von denen viele in dieser Arbeit Anwendung finden und in diesem Unterkapitel erläutert werden.

Pixelbasierte Segmentierung:

Ein pixelbasiertes Segmentierungsverfahren ordnet jedes Pixel anhand einer Entschei-

dungsregel einem Bildbereich zu. Solche Verfahren benötigen normalerweise nur geringen Rechenaufwand, sind deshalb schnell in der Umsetzung, liefern aber häufig keine vollständige und/oder fehlerfreie Segmentierung.

Schwellenwertverfahren arbeiten pixelorientiert, indem jeder Pixelwert mit einem Schwellenwert verglichen wird. Je nachdem, ob dieser unter- oder überschritten wird, erfolgt eine Zuordnung des Quellpixels zum Objekt oder zum Hintergrund. Dabei kann auch mit mehreren Schwellenwerten gearbeitet werden. Dadurch entstehen zulässige und unzulässige Bereiche der Pixelwerte. Schwellenwertverfahren arbeiten häufig mit den Helligkeitsgrauwerten eines Bildes, können aber auch auf andere Bildwerte angewandt werden.

Die Schwierigkeit von Schwellenwertverfahren liegt in der Bestimmung eines geeigneten Schwellenwertes und der Größe seines Anwendungsbereichs innerhalb des Bildes. Der Grund dafür ist die Kopplung einer erfolgreichen Schwellenwertberechnung mit den Eigenschaften des Bildinhalts und der Bildqualität.

Es gibt verschiedene Vorgehensweisen zur Berechnung von Schwellenwerten. Sie unterscheiden sich in ihrem Rechenaufwand und den Segmentierungsergebnissen, die damit erreicht werden können. Als Standard zur automatischen Berechnung hat sich das Verfahren von *Otsu* aus dem Jahr 1979 etabliert, welches auf den statistischen Grundlagen des Bildes basiert. Der Schwellenwert wird dabei so bestimmt, dass die Streuung der Pixel innerhalb der dadurch entstehenden Segmente möglichst gering, die Streuung zwischen den Segmentpixeln und den Hintergrundpixeln jedoch möglichst groß ist.

Globale Schwellenwerte beziehen sich auf das gesamte Bild und führen nur zu brauchbaren Segmentierungsergebnissen, wenn das Objekt sich stark vom Hintergrund abhebt, es also zumindest in einer der Eigenschaften Helligkeit, Farbe oder Sättigung keine Gemeinsamkeiten mit dem Hintergrund aufweist. Diese Bedingung wird für Abbildungen realer Szenen nur selten erfüllt, da meistens die Beleuchtung variiert und die Farbgebung durchwachsen ist.

Um dem Problem eines wechselnden Hintergrunds entgegenzuwirken, kann eine Schwellenwertbildung auch *lokal* angewandt werden. Das bedeutet, dass das Bild in Regionen unterteilt und für jede Region ein eigener Schwellenwert berechnet wird. Hierbei ist es absolut notwendig, dass in jeder Region auch ein zu segmentierender Objektbereich enthalten ist, da die Schwellenwertbildung sonst nur anhand von Hintergrundpixeln erfolgt, was zu fehlerhaften Segmentierungsentscheidungen führt. Zwar verringert das lokale Verfahren gegenüber dem globalen Verfahren eine fehlerhafte Segmentierung bei ähnlichem Rechenaufwand, jedoch kann es durch die unterschiedlichen, lokal angewandten Schwellenwerte zum Teil zu Versatz an den Grenzen der Regionen kommen.

Die Weiterentwicklung des lokalen Schwellenwertverfahrens sind *dynamische* Verfahren, die für jedes Pixel automatisch einen Schwellenwert anhand seiner Nachbar-

schaft berechnen. Der Rechenaufwand steigt bei dieser Variante erheblich, da für jedes Pixel eine Neuberechnung des Schwellenwertes erforderlich ist, dafür sind sie gegenüber lokalen Helligkeitsänderungen weitgehend stabil.

Kantenbasierte Segmentierung:

Kantenbasierte Segmentierungsverfahren arbeiten mit den Kanten und Übergängen im Bild. Sie nutzen die Eigenschaft, dass Kanten häufig als Konturen zwischen den zusammengehörigen Pixelbereichen im Bild liegen. Sie grenzen die Objekte voneinander und vom Hintergrund ab und bieten eine geeignete Segmentierungsgrundlage. Um diese nutzbar zu machen, müssen die Kanten im Bild zuerst detektiert, herausbildet und gegebenenfalls vervollständigt werden.

Ein kantenbasiertes Verfahren zur Bildsegmentierung ist die *Wasserscheidentransformation*. Dabei wird der Gradientenbetrag eines Bildes als Höhenfeld interpretiert, das nach und nach mit Wasser geflutet wird. Dadurch entstehen verschiedene Becken, die durch die Kanten, die im Gradientenbild Bereiche hoher Beträge darstellen, getrennt werden, die sogenannten Wasserscheiden. Die Becken sind äquivalent zu den Segmenten, in die das Bild unterteilt wird. Häufig führt dieses Verfahren zu einer Übersegmentierung, insbesondere bei verrauschtem Bildmaterial. Um Einfluss auf die Segmentierung nehmen zu können und eine zu hohe Segmentierung zu vermeiden, wurden Varianten der Wasserscheidentransformation entwickelt.

In dieser Arbeit wird die *markerbasierte Wasserscheidentransformation* verwendet. Diese Variante ermöglicht eine zielgerichtete Segmentierung, indem sie das Bildmaterial mit morphologischen Operatoren vorverarbeitet. Die Idee bei dabei ist, zunächst nur eine grobe, aber absolut sichere Zuordnung von Bildbereichen zum Vordergrund oder Hintergrund vorzunehmen und entsprechend mit Markern zu versehen. Die genaue und gänzliche Zuordnung der bis dahin nicht markierten Bereiche wird dann anhand der Wasserscheidentransformation vorgenommen. Sie weitet die schon sicheren Bereiche aus, bis alle Pixel einem bestimmten Bereich zugewiesen wurden. Diese Methode erweitert die einfache Wasserscheidentransformation um eine Vorentscheidung, die eine Übersegmentierung ausschließt, sodass ihr Ergebnis der tatsächlich vorliegenden Bildaufteilung sehr nahe kommt.

Die Schwierigkeit des Verfahrens liegt in der sinnvollen Positionierung der Marker und einer Vorsegmentierung, die sich als Entscheidungsgrundlage eignet. Hierzu kann ein Schwellenwertverfahren dienen, dessen Maskenresultat anschließend nacheinander mit den morphologischen Operatoren Dilatation und Erosion bearbeitet wird. Bei geeigneter Wahl des Schwellenwerts und der Größe des Strukturelements können auf diese Weise zwei Masken erstellt werden. Eine Vordergrundmaske, deren weiße Regionen der Binärmaske durch die Dilatation verkleinert wurden, und eine Hintergrundmaske, deren schwarze Regionen der Binärmaske durch Erosion verkleinert wurden. Sie enthalten die Information für Bereiche, die sicher zugeordnet werden können. Ihre Addition mit vorheriger Graufärbung des jeweils unbestimmten Bereichs führt deshalb zu einer Markermaske, in der drei Bildbereiche markiert wurden, Vordergrund,

Hintergrund und der dazwischenliegende unbestimmte Bereich.

Modellbasierte Segmentierung:

Modellbasierte Segmentierungsverfahren verwenden die modellierten Eigenschaften der gesuchten Objekte im Bild, um diese gezielt herauszubilden zu können. Dies erfordert die Kenntnis von Objekteigenschaften, die als Eingabeparameter für entsprechende Segmentierungsalgorithmen dienen.

Die *Houghtransformation* ist eines dieser Verfahren und wird zur Erkennung von geometrischen, parametrisierbaren Figuren in binären Kantenbildern eingesetzt. Üblicherweise werden Kreise oder Linien gesucht, deren Parameter vorher festgelegt oder eingegrenzt werden können. Die Houghtransformation unterscheidet sich grundsätzlich von den zur Vorverarbeitung eingesetzten Transformationsarten, da ihre Anwendung nicht auf eine verlustfreie Überführung der Bilddaten in eine andere Form abzielt, sondern den Bildinhalt einem bestimmten Schema nach bewertet. Dazu transformiert sie das Bild in ein niederdimensionales System, einen sogenannten Dual- oder auch Parameterraum, der speziell Houghraum genannt wird.

Die Houghtransformation dient innerhalb dieser Arbeit der Liniensuche, darum wird das Prinzip anhand dieses Beispiels erläutert: Es besteht darin, statt einer komplexen Merkmalsextraktion eine einfache Extremwertbestimmung durchzuführen. Dazu wird die binäre Kanteninformation eines Bildes den Houghraum transformiert, dessen Punktkoordinaten die Parameter jeder möglichen Geraden im Binärbild repräsentieren. Liegen im Kantenbild weiße Pixel auf einer Geraden, so erfüllen diese damit eine bestimmte Geradengleichung. Dadurch steigt im Houghraum der Zählerbetrag des Punktes am Ort der Geradenparameter, die diese Linie beschreiben, um die Anzahl dieser Pixel an. Die Orte mit den höchsten Wertigkeiten im Parameterraum zeigen auf die sehr wahrscheinlich existierenden Geraden im Binärbild. Sie liefern mit ihren Koordinaten die Geradenparameter der gesuchten Linien. Für das Aufspannen des Houghraumes werden Parameter mit endlichen Werten benötigt. Die Steigung und der Achsenabschnitt eignen sich nicht dazu, weil eine vertikale Gerade ins Unendliche steigt. Deshalb wird die hessesche Normalform einer Geraden verwendet, um den Houghraum mit deren Parameter Winkel α und euklidischer Abstand d aufzuspannen.

Ein weiteres modellbasiertes Verfahren ist das *Template Matching*. Mit ihm wird die Übereinstimmung eines Modells der gesuchten Bildinformation (Template) mit einem vorhandenen Pixelbereich im Quellbild festgestellt (Match). Hierzu wird das Template, das einen Bildausschnitt aus dem Original selbst oder aber eine gute Annäherung daran darstellt, mittels vollständiger Suche positionswise mit dem Quellbild verglichen. Dann wird anhand eines gewählten Maßstabs entschieden, an welchen Stellen eine gute bis perfekte Übereinstimmung von Template und Original zutrifft.

Für eine erfolgreiche Suche sollte das Template den gesuchten Pixelbereich eingebettet in seine Umgebung möglichst gut modellieren. Aufgabenabhängig wird nur

nach genau diesem Template oder aber auch nach allen möglichen Skalierungen und Ausrichtungen dessen gesucht. Als Maßstab für eine ausreichende Übereinstimmung von Modell und Bildobjekt können verschiedene Kriterien genutzt werden, zum Beispiel der durchschnittliche quadratische Abstand oder der Abstand der Beträge von Bild f und Template t . Beide sind relativ stabil gegenüber Bildrauschen, jedoch instabil gegenüber Intensitätsunterschieden von Template und Bild. Ein anderer Maßstab ist der Korrelationskoeffizient, mit dem auch mühelos alle im Bild enthaltenen Negierungen des Templates festgestellt werden können. Das Template Matching bedarf aufgrund der vollständigen Bildsuche eines hohen Rechenaufwands.

2.2.3.2 Merkmalsextraktion

Die Merkmalsextraktion ist die gezielte Suche nach Merkmalen in Bildern, die zu Objekteigenschaften korrelieren. Ihr Ziel ist die Gewinnung von objektspezifischen Informationen, die auch komplizierte Objekte möglichst eindeutig beschreiben können und anhand derer eine Erkennung und weitergehend auch eine Klassifizierung der abgebildeten Objekte ermöglicht wird. Die Merkmalsextraktion folgt dem Schritt der Bildsegmentierung und beschreibt die zuvor segmentierten Bereiche qualitativ und quantitativ. Welche Merkmale benötigt werden, hängt von der Aufgabenstellung des Mustererkennungssystems ab, also davon, zu welchem Zweck ein Bild analysiert werden soll.

Das Ziel der Merkmalerkennung in dieser Arbeit ist die Rekonstruktionsfähigkeit von Objektform und -position innerhalb des Bildes. Deshalb konzentrieren sich die hier angewandten Algorithmen auf das Extrahieren von Konturen und die Lokalisierung der Merkmale. Grundsätzlich werden die Merkmale jedes Objekts in einem Merkmalsvektor gespeichert, der als komprimierte Darstellung eines Bildobjekts aufgefasst werden kann. Das Problem der Merkmalerkennung ist für jede Aufgabenstellung und jeden Bildinhalt sehr spezifisch, weshalb es keine Generallösung, sondern maximal Teillösungen gibt. Zielführend ist in vielen Fällen der an das spezifische Problem angepasste Einsatz einer abgestimmten Kombination aus verschiedenen gängigen Bildverarbeitungsalgorithmen.

3 Implementierung der Gesichtsanalyse

Dieses Kapitel befasst sich mit der Implementierung des hier entwickelten Verfahrens zur Gesichtsanalyse sowie der Aufbereitung und Ausgabe der Analysedaten.

3.1 Gesichtsmerkmalslokalisierung

Dieses Kapitel befasst sich mit der Implementierung der Gesichtsanalyse in Form der Gesichtsmerkmalslokalisierung. Sie dient der Datenerfassung für die darauffolgende Berechnung der Korrespondenzpunkte für die Anpassung des 3D-Gesichtsmodells. Dafür werden beide Teilbereiche des hier entwickelten teilautomatisierten Verfahrens näher beleuchtet, die manuelle Benutzereingabe und die daran anschließende automatische Konturlokalisierung für jedes Gesichtsmerkmal.

Zunächst wird auf die Eingrenzung des Suchbereichs für die jeweiligen Merkmale eingegangen. Danach werden die einzelnen Verfahren erläutert, die zur genauen Lokalisierung der Merkmalskonturen innerhalb der eingegrenzten Merkmalsbereiche eingesetzt werden. Im Zuge dessen wird auch die eventuell erforderliche Interpolation von Daten, die nur angenähert werden können, beschrieben.

Das Ziel des Gesamtprozesses der Gesichtsanalyse ist die Bereitstellung einer geeigneten Berechnungsgrundlage für die erforderlichen Korrespondenzpunkte. Sollte dieses Ziel aus irgendwelchen Gründen nicht erreicht werden, ist ein wiederholter Versuch unter anderen/geeigneteren Bedingungen notwendig. Deshalb verfügt jeder der nachfolgend beschriebenen Arbeitsschritte über Kontrollmechanismen, die es erlauben, Metadaten über den Erfolg oder Misserfolg einzelner Prozesse zu liefern, damit eine Interaktion mit dem Benutzer möglich wird.

3.1.1 Eingrenzung des Suchbereichs

Es gibt zwei Gründe, weshalb die Eingrenzung des Suchbereichs für die nachfolgende Konturlokalisierung der Gesichtsmerkmale sinnvoll ist. Zum einen wird dadurch die Anzahl der Fehldetektionen stark eingegrenzt, weil der Anteil irrelevanter Bildinformation gegenüber der relevanten Information sinkt. Zum anderen erhöht sich dadurch die Verarbeitungsgeschwindigkeit, weil nur ein kleiner Pixelbereich durchsucht werden muss. In Kapitel 2.1 wurde erläutert, weshalb die Erstellung einer Suchmaske

manuell per Benutzerauswahl erfolgen wird. Dadurch ergeben sich bestimmte Anforderungen an das GUI, die nachfolgend aufgeführt werden und die Grundlage für die Entwicklung des GUI in der Arbeit von [Burmeister \(2016\)](#) sind.

Die Form der Bereichsauswahl, die für die Konturlokalisierung der Gesichtsmerkmale erfolgen muss, ergibt sich aus den für das jeweilige Lokalisierungsverfahren erforderlichen Eingabedaten und unterscheidet sich deshalb von Merkmal zu Merkmal. Für die Augenbrauen, die Nase und den Mund ist die Auswahl in Form eines eingrenzenden Rechtecks um das jeweilige Merkmal ausreichend, weil die Suchalgorithmen für diese Merkmalskonturen auf den gesamten Bereich der Vorauswahl angewandt werden.

Für die Augen und die Kinnlinie bedingen die Lokalisierungsverfahren jedoch eine abweichende Benutzerauswahl. Für die dort angewandten Suchalgorithmen werden Linien benötigt, von denen ausgehend die Suche durchgeführt wird, sodass der Benutzer dafür bestimmte Markierungen setzen muss, die zur Berechnung der Ausgangslinien erforderlich sind. Für die Augen muss der Benutzer deshalb die Möglichkeit erhalten, die Iris in Form eines quadratischen Auswahlfeldes eingrenzen zu können, damit anhand dessen der Radius und der Mittelpunkt der Pupille übergeben werden können, die der Berechnung der kreisförmigen Iriskontur dienen. Das Suchfenster für die Augenkontur entspricht dann einem Rechteck, das um den Mittelpunkt der Iris mit einer Kantenlänge des achtfachen Radius für die Breite und des dreifachen Radius für die Höhe aufgespannt wird. Die Werte orientieren sich an den typischen Größenverhältnissen innerhalb des Auges. Für die Suche nach der Kinnkontur eignet sich als Ausgangslinie eine angenäherte Polygonlinie aus insgesamt vier Teilstücken. Zu ihrer Berechnung ist die Benutzerauswahl von fünf markanten Kinnpunkten notwendig, die über das GUI realisiert werden muss.

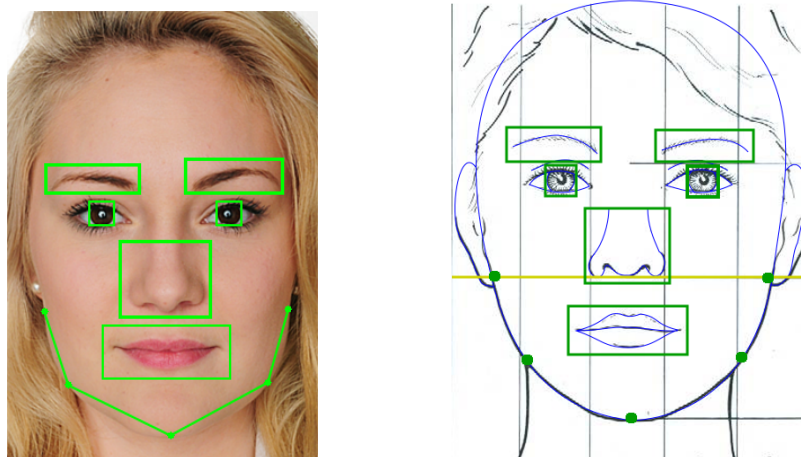


Abbildung 3.1: Auswahlbereiche des Benutzers für die lokale Merkmalsuche - links: Ansicht der anforderungsgerecht gewählten Suchfelder im GUI ([Frau4](#)), rechts: Hilfschema zur Erstellung der Benutzerauswahl

3.1.2 Konturlokalisierung und -interpolation

Dieses Kapitel stellt die auf die einzelnen Gesichtsmerkmale spezialisierten Verfahren zur genauen Lokalisierung der Kontur oder bestimmter Merkmalspunkte vor. Diese Verfahren kombinieren für die lokale Suche der Merkmalskonturen verschiedene, gängige Bildverarbeitungsalgorithmen und nutzen die Symmetrieeigenschaften von Frontalansichten sowie verschiedene spezifischen Merkmalseigenschaften (Unterunterabschnitt 2.2.1.1) aus.

Die hier verwendeten Verfahren gleichen in ihrem Grundprinzip den in der Dissertation "Modellbasierte posen- und mimikinvariante Gesichtserkennung" von Hähnel (2008) verwendeten Verfahren zur Konturlokalisierung von Gesichtsmerkmalen. Jedoch wurden sie insofern abgeändert, als dass die dort genutzten und vollautomatisch bestimmten Suchbereiche für die Merkmale durch die Eingabedaten des Benutzers ersetzt wurden. Außerdem wurde jeweils nur das dort erklärte Schema verwendet und falls notwendig erweitert sowie alle dafür verwendeten Schwellenwerte und Berechnungen den spezifischen Anforderungen dieser Analysesoftware angepasst.

Die Verfahren werden in folgenden Unterkapiteln zunächst grob erklärt und danach erfolgt die Erläuterung der einzeln durchgeführten Schritte. Für die Implementierung der Konturlokalisierung wird die Open Computer Vision (OpenCV) Programmibibliothek genutzt, die für die Programmiersprachen C, C++ und Python geschrieben wurde und alle gängigen Algorithmen der Bildverarbeitung enthält.

3.1.2.1 Augen

Die Lokalisierung der Augenkontur erweist sich aufgrund der Komplexität des Erscheinungsbilds der Augen als schwierig. Hähnel geht aufgrund dessen zu einer schrittweisen Detektion über, in der er zunächst die Iris mittels Houghtransformation detektiert und davon ausgehend erst die Augenkontur lokalisiert. Das hier verwendete Lokalisierungsverfahren verzichtet aufgrund des mäßigen Erkennungserfolgs in durchgeführten Tests auf den ersten Schritt und erhält stattdessen den Radius und den Mittelpunkt der Iris per Benutzereingabe. Die lokale Suche der Augenkontur erfolgt dann innerhalb des zuvor berechneten Bildausschnitts der Augensuchmaske in folgenden Schritten:

Schritt 1 Der geringe Rotanteil der Iris im Verhältnis zur umgebenden Haut kann für die Berechnung der Öffnungshöhe ausgenutzt werden, sodass im ersten Schritt der Rotkanal für die Augenausschnitte extrahiert wird.

Schritt 2 Dieser Schritt dient der Vorbereitung auf Schritt 4. Hierzu wird der Luminanzkanal für die Augenausschnitte berechnet und darin auf jede Zeile ein eindimensionaler, horizontaler Gaußfilter angewandt. Für eine maximale Kontrastinformation wird anschließend ein Histogrammausgleich für den Kanal berechnet. Die von Hähnel durchgeführte Rauschreduzierung für das gesamte Bild

3 Implementierung der Gesichtsanalyse

per Medianfilter entfällt an dieser Stelle, weil für das vorgesehene Eingabematerial keine verrauschten Bilder erwartet werden und sich der Filter innerhalb der Tests als entbehrlich erwiesen hat.

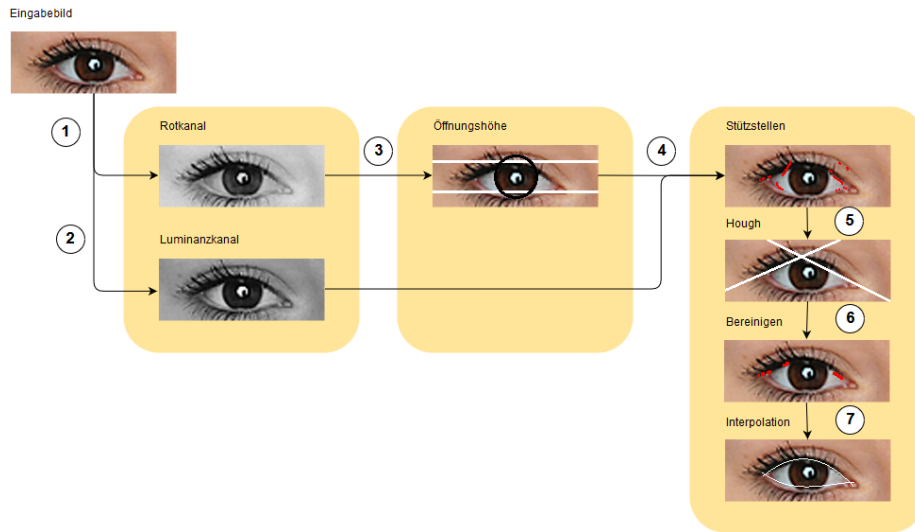


Abbildung 3.2: Augenlokalisierung

Schritt 3 Die Öffnungshöhe beschreibt den Abschnitt, der Iris, der nicht von den Augenlidern verdeckt wird und entspricht bei geöffneten Augen dem Abstand zwischen dem höchsten und dem niedrigsten Punkt des Auges. Wegen der zuvor festgelegten Einschränkungen für das Eingabematerial wird davon ausgegangen, dass die Augen der abgebildeten Person in jedem Fall geöffnet sind. Deshalb muss für die Lokalisierung der Öffnungshöhe nur in einem bestimmten Bereich am Rande der Iris gesucht werden. Darum wird die in Seite 30 dargestellte Suchmaske verwendet, die 40% der Mitte der Iris ausspart.



Abbildung 3.3: Iris-Suchmaske

Dies hat den Vorteil, dass Reflexionen im Innern der Iris weniger Einfluss auf das Verfahren zur Berechnung der Öffnungshöhe haben, das nach folgendem Prinzip durchgeführt wird:

Für die Fläche der unverdeckten Iris (ausgenommen von dem Bereich, der durch die Suchmaske ausgespart wird), wird zeilenweise der durchschnittliche Pixelwert des Rotkanals \bar{P}_r berechnet:

$$\bar{P}_r = \frac{1}{n} \sum_{i=0}^n$$

3 Implementierung der Gesichtsanalyse

mit \bar{P}_r = Durchschnittlicher Rotton der Zeile,

i = Pixelindex,

n = Anzahl der Pixel pro Zeile

Für die Ermittlung der Öffnungshöhe wird nun der Abschnitt innerhalb der Funktion gesucht, in dem die Rotwerte signifikant geringer sind als am Anfang und am Ende der Funktion. Hierzu werden die durchschnittlichen Rottöne von Zeile zu Zeile verglichen. Die Zeile, in der der Pixelwert im Vergleich zur darauffolgenden stark abfällt, beschreibt den Ort der oberen und die Zeile, in der der Pixelwert im Vergleich zur vorangegangenen stark ansteigt, beschreibt den Ort der unteren Lidkante, der Bereich dazwischen ist also die Öffnungshöhe des Auges.

Schritt 4 Der vierte Schritt dient der Detektion der Augenkontur und entspricht dem sechsten Schritt aus Hähnel's Vorgehensweise. Hierbei werden ausgehend von den Informationen über die Ausmaße der Iris und die vertikale Begrenzung des Augenbereichs durch die Öffnungshöhe sogenannte Stützstellen für die anschließende Interpolation der Augenkontur gesucht. Dazu wird der Luminanzkanal als Höhenfeld interpretiert und zeilenweise ausgehend vom Irisrand in beide Richtungen abgetastet. Pixel, die sehr wahrscheinlich auf der Augenkontur liegen, erfüllen eine der folgenden beiden Bedingungen:

1. Das Pixel entspricht einem Punkt im Höhenfeld, der den Anfang eines starken Anstiegs darstellt. Dort wo die Luminanzwerte stark ansteigen, befindet sich der Übergang vom Auge zur umliegenden Haut.
2. Der Pixelwert entspricht außerdem einem lokalen Luminanzminimum, sodass sein Betrag einen bestimmten Schwellenwert unterschreitet. Dieser Schwellenwert wird jeweils abhängig vom höchsten Luminanzwert der betrachteten Zeile gebildet und verhindert, dass geringfügige Luminanzschwankungen fälschlicherweise als Minima interpretiert werden.

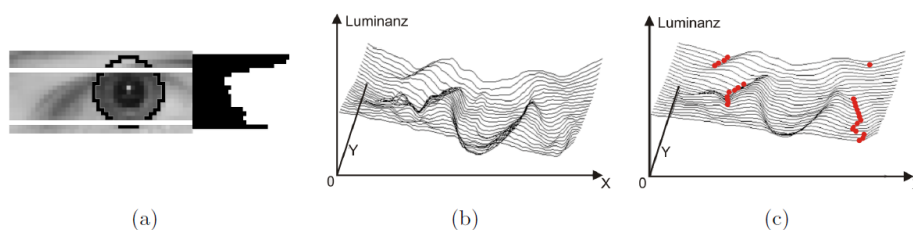


Abbildung 3.4: Interpretation des Luminanzkanals als Höhenfeld: (a) typischer Verlauf von $h(y)$ (b) Ausgangsprofil, (c) lokale Minima im geglätteten Profil, entnommen Hähnel (2008), Seite 101

Schritt 5 Die in Schritt 4 detektierte Pixelmenge setzt sich zusammen aus den tatsächlichen Konturpunkten und solchen Punkten, die zwar eine der Bedingungen erfüllen, aber außerhalb der Augenkontur liegen. Um die Ausreißer herauszufiltern, wird innerhalb dieses Schrittes für die Kontur der oberen Lidkante für jede Augenhälfte separat eine Houghtransformation auf die entsprechende Punktmenge angewandt. Das Ergebnis der Houghtransformation sind die Geraden, auf denen die meisten Punkte liegen. Da für das obere Augenlid ein Öffnungswinkel von 30° anzunehmen ist, wird der Parameterraum dementsprechend begrenzt. Von den möglichen Geraden wird dann die wahrscheinlichste gewählt, indem zunächst alle Geraden zu den Rändern des Bildausschnitts verlängert werden und anschließend diejenige ausgewählt wird, deren Mittelpunkt am nächsten am Irismittelpunkt liegt.

Die Problematik, dass zu wenige Stützstellen für die rechte oder linke Seite der oberen Augenkontur gefunden wurden, um brauchbare Geraden aus der Houghtransformation zu erhalten, wird versucht zu umgehen, indem auf die Punkte zuvor die morphologische Operation Dilatation angewandt wird. Dazu wird ein Strukturelement in Form eines kleinen Geradenstücks mit einer Steigung von 30° verwendet, das die Geradenstruktur der Punktmenge hervorheben soll.

Für den Fall, dass die Houghtransformation für eine oder sogar für beide Hälften des einen Auges erfolglos bleibt, wird die Symmetrieeigenschaft des Gesichts ausgenutzt und auf die Informationen des anderen Auges zurückgegriffen. In diesem Fall werden die entsprechenden fehlenden Geraden des anderen Auges auf die unvollständige Seite gespiegelt, um trotzdem eine Richtlinie für die Lage der Punkte zu gewinnen, die zur Augenkontur gehören. Schritt 6: Mit Hilfe der zuvor berechneten Geraden werden die Ausreißer aus der Punktmenge der Stützstellen identifiziert und eliminiert. Dazu wird der Abstand jedes Punktes zu der ermittelten Geraden berechnet und alle Punkte entfernt, deren Abstand einen empirisch gewählten Schwellenwert überschreiten.

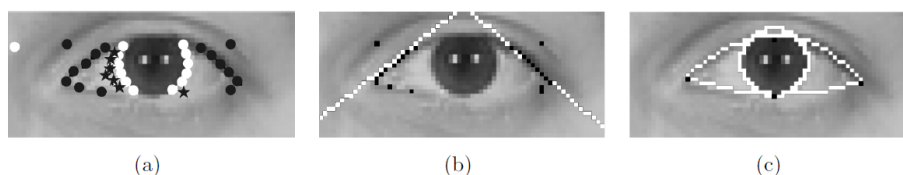


Abbildung 3.5: Lokale Suche der Augenkontur: (a) Suche pro Zeile ausgehend von der Iris (Stern: Luminanzmaxima, weiß: Startposition, schwarz: Ergebnis), (b) Entfernen von Ausreißern mittels Hough-Transformation, (c) Least-Squares-Ausgleichskurven, entnommen [Hähnel \(2008\)](#), Seite 102

Schritt 6 Mit Hilfe der zuvor berechneten Geraden werden die Ausreißer aus der Punktmenge der Stützstellen identifiziert und eliminiert. Dazu wird der Abstand jeden Punktes zu der ermittelten Geraden berechnet und alle Punkte entfernt, deren Abstand einen empirisch gewählten Schwellenwert überschreiten.

Schritt 7 Im letzten Schritt werden mit den Stützstellen der Augenkontur per Least-Squares-Anpassung für die untere und die obere Augenkontur kubische Polynome berechnet, die die tatsächliche Kontur interpolieren. Dazu werden zunächst die Augenwinkel, also die beiden äußersten Punkte der Augenkontur, bestimmt. Hähnel verwendet hierfür die zwei Punkte aus der Menge der Stützstellen, die den maximalen und den minimalen Wert für die X-Koordinate aufweisen. Davon abweichend werden die Augenwinkel innerhalb dieses Verfahrens anhand der typischen Längenverhältnisse für ein Auge mit

$$x = \text{Irismittelpunkt} + \text{Irisradius} \cdot 2,2 \quad \text{und} \quad y = mHough \cdot x + bHough$$

interpoliert.

Diese Lösung bedingt, dass die Iris tatsächlich mittig im Auge liegt und nimmt eine eventuell auftretende geringe Abweichung für die Position der Augenwinkel in Kauf. Dadurch wird eine gänzlich falsche Bestimmung der Position ausgeschlossen, die innerhalb von Tests mit Hähnels Vorgehensweise mehrfach festgestellt wurde, weil die äußersten Augenpunkte in vorangegangenen Schritten falsch detektiert und deshalb eliminiert wurden. Anhand der Horizontalen, die zwischen den Augenwinkeln verläuft, können die Punkte jeweils entweder zur unteren oder oberen Augenkontur zugeordnet werden. Letztendlich kann mittels der beiden Punktteilmengen der Außenpunkte und dem tiefsten und höchsten Irispunkt die Least Squares Anpassung durchgeführt und deren Lösungen für die obere und untere Augenkante als Augenkontur zusammengeführt werden.

3.1.2.2 Augenbrauen

Die Augenbrauenkontur hebt sich in ihrer Intensität nur wenig von der umliegenden Haut ab, sodass sich laut Hähnel die Verfahren, die allein auf der Schwellenwertsegmentierung von Intensitätskanal oder Gradientenbild beruhen, vor allem bei Personen mit lichten oder hellen Brauen als instabil erweisen. Deshalb schlägt Hähnel ein Verfahren vor, das für die Lokalisierung der Brauenkontur die Kombination aus Farb- und Gradienteninformation nutzt, nämlich in Form der Wasserscheidentransformation und einer anschließenden Approximation der Oberkante der Augenbrauen.

Schritt 1 Als erstes wird der Intensitätskanal des Bildausschnittes berechnet.

Schritt 2 Abweichend von Hähnels Vorgehensweise wird innerhalb dieses Verfahrens zunächst eine Maske erstellt, die alle hautfarbenen Bereiche schwarz und den

3 Implementierung der Gesichtsanalyse

Rest, also die Augenbrauen, weiß darstellt. Diese Maske ist für die Konturberechnung ungeeignet, weil die Abgrenzung zu ungenau ist. Wird sie allerdings per gewichteter Addition mit dem Intensitätskanal verrechnet, führt dies zu einer Verstärkung der Braue gegenüber der Haut. Bei hellhaarigen Menschen, bei denen sich die Braue nur sehr schwach von der Umgebung abhebt, hat sich dieser Vorverarbeitungsschritt als vorteilhaft erwiesen. Anhand des so bearbeiteten Bildes können anschließend auch in kritischen Fällen die Marker für die Wasserscheidentransformation sicher gesetzt werden, um eine gute Basis für die Transformation zu schaffen.

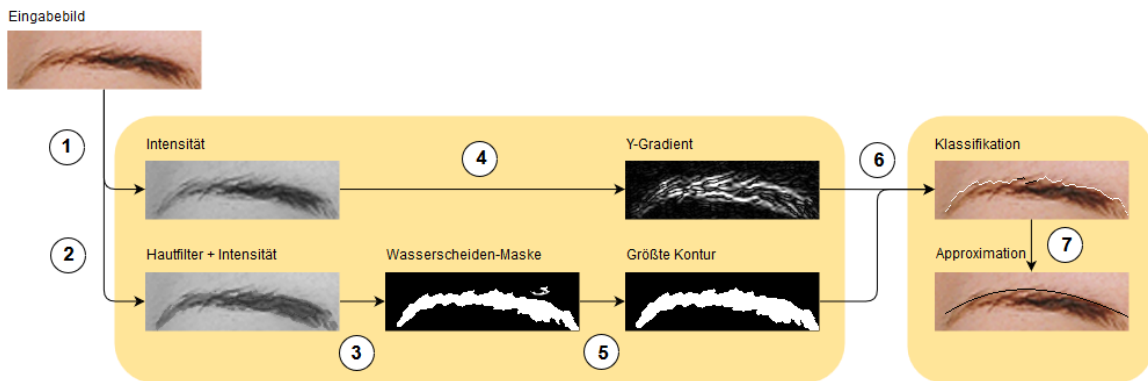


Abbildung 3.6: Augenbrauenlokalisierung

Schritt 3 Zur farblichen Segmentierung des Eingabebildes wird eine markerbasierte Wasserscheidentransformation angewendet, in deren Ergebnis Vordergrundelemente weiß und Hintergrundelemente schwarz dargestellt werden. Im Idealfall hebt sich die Augenbraue unter Verwendung der zuvor berechneten Marker gänzlich vom Hintergrund ab und bildet die größte Segmentfläche.

Schritt 4 Zur Berechnung des vertikalen Gradienten wird auf den Intensitätskanal des Bildausschnitts ein Sobel-Operator angewandt.

Schritt 5 In diesem Schritt wird das Segmentierungsergebnis der Wasserscheidentransformation bereinigt. Dafür wird die größte zusammenhängende Vordergründfläche extrahiert und die restlichen Segmente geschwärzt.

Schritt 6 Für die Lokalisierung der Oberkante der Augenbraue werden in diesem Schritt die Informationen aus dem Gradienten und der Binärmasken aus der Wasserscheidensegmentierung kombiniert. Die Kante ist durch einen Übergang von hell (Haut) nach dunkel (Braue) charakterisiert, sodass alle Konturpixel einen negativen Gradienten aufweisen. Eben solche Pixel, die zusätzlich im weißen Bereich der Augenbrauenmaske liegen, sind potenzielle Oberkantenpixel. Da diese Eigenschaft jedoch auch auf andere Punkte innerhalb der Braue zutreffen kann, wird jeweils immer nur das Pixel pro Spalte gewählt, das den

niedrigsten Wert für die Y-Koordinate aufweist.

Für die Bestimmung der Oberkante kommt nun ein Wachstumsalgorithmus zur Anwendung. Dieser unterteilt die potentiellen Konturpixel in zwei Mengen, eine Aktiv- und eine Passivmenge. Die Aktivmenge beschreibt die Zugehörigkeit zur Augenbrauenoberkante. Sie beinhaltet zu Beginn nur den mittleren Punkt aus der Kandidatenmenge. Er gehört mit großer Wahrscheinlichkeit zur Oberbrauenkante, weil der Bildausschnitt die Braue ausreichend genau erfasst. Danach werden nach und nach all jene Punkte aus der Passivmenge der Aktivmenge hinzugefügt, deren Distanz zu einem Punkt in der Aktivmenge unterhalb einer festgelegten Maximaldistanz liegt. Der Wachstumsprozess bricht ab, sobald die Passivmenge keine Punkte mehr enthält, die genügend nah an Punkten der Aktivmenge liegen.

Schritt 7 Der letzte Schritt umfasst die Approximation der Augenbrauenoberkante mittels der Least Squares Lösung für die zuvor bestimmten Konturpixel. Der Verlauf der Oberkante wird durch ein kubisches Polynom ausreichend genau angenähert, dessen Koeffizienten anhand der Punkte aus der Aktivmenge bestimmt werden. Dadurch werden auch eventuell enthaltene Ausreißer geglättet.

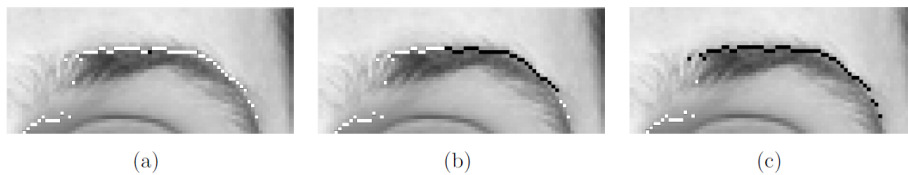


Abbildung 3.7: Bestimmung der Augenbrauenkontur: (a) Initialisierung, (b) Situation nach der Hälfte der Iterationen, (c) Endergebnis, entnommen [Hähnel \(2008\)](#), Seite 104

3.1.2.3 Nase

Weil die Nase sich farblich nur geringfügig von anderen Gesichtsräumen abhebt, ist die Kontur für ihre Lokalisierung ungeeignet. Hähnel beschreibt, dass die häufigsten Ansätze zur Nasendetektion anhand ihrer Struktur vorgenommen werden, weil die Nase in ihrer Form statisch ist. Als charakteristische Merkmale dienen dafür entweder die Nasenlöcher, die sich sehr gut anhand der Intensität bestimmen lassen, oder aber das Nasenmuschelprofil oder die Nasenspitze.

Das folgend beschriebene und für diese Arbeit verwendete Verfahren zur Nasenlokalisierung stützt sich auf die Symmetrieeigenschaft der Nase und die gute Detektierbarkeit der Nasenlöcher mittels Template Matching.

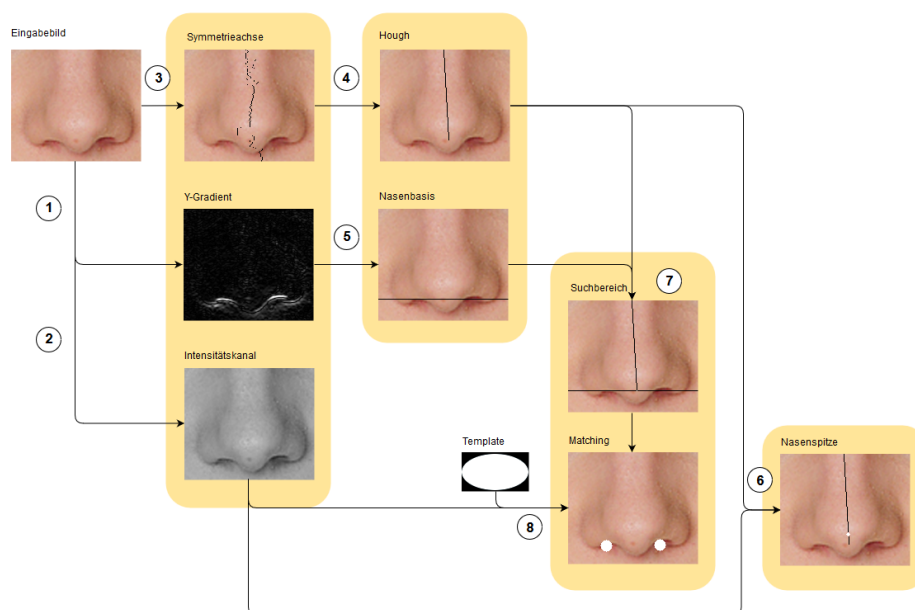


Abbildung 3.8: Nasenlokalisierung

Schritt 1 Als Vorbereitung für die Detektion der Nasenbasis werden die Beträge des vertikalen Gradienten des Nasenausschnitts berechnet.

Schritt 2 Anhand des Eingabebildausschnitts wird der Intensitätskanal berechnet, weil er die Kontrastinformation zur Detektion der Nasenlöcher enthält.

Schritt 3 Zur Lokalisierung der Nasenmittelachse wird für jeden RGB-Kanal des Nasenbilds einzeln pro Zeile das Pixel, das für diese Zeile den Symmetriepunkt bildet, bestimmt. Hierzu erhält jedes Pixel eine Bewertung anhand der Formel

$$score(x, y) = \sum_i \sum_{j \in \{r, b, g\}} |C_j(x + i, y) - C_j(x - i, y)|$$

Diese Summe schrumpft für ein Pixel dann auf ein Minimum, wenn die Pixel, die jeweils zu seiner Rechten und zur Linken dessen angeordnet sind, eine Symmetrie aufweisen, weil ihre Werte sich bei Subtraktion voneinander gegenseitig aufheben.

Schritt 4 Die Bestimmung der Nasenmittelachse erfolgt durch eine Houghtransformation der zuvor berechneten Stützstellen aus Schritt 3. Dadurch werden automatisch alle Ausreißer entfernt.

Schritt 5 Für die Information über die Höhe und Position der Nasenspitze und der Nasenlöcher wird in diesem Schritt die Nasenbasis in Form einer Hilfsgeraden für weitere Berechnungen bestimmt. Die Nasenbasis bezeichnet die horizontale

3 Implementierung der Gesichtsanalyse

Linie, unterhalb derer sich nahezu immer eine Region mit hohem Kontrast zu ihrer Umgebung abzeichnet. Dieser entsteht zum einen durch die dort befindlichen Nasenlöcher, aber auch den Schatten, den die vorstehende Nase wirft. Die Oberkante dieses Bereichs ist durch hohe vertikale Gradientenwerte gekennzeichnet. Deshalb erfolgt die Detektion der Nasenbasis über die Bestimmung der höchsten Zeilensumme über die Gradientenbeträge.

Schritt 6 Für die Detektion der Nasenspitze wird laut Hähnel's Recherchen in vielen Verfahren die hellste Stelle auf der Nasenmittellachse im Intensitätskanal gesucht. Allerdings weist Hähnel auf eine bestehende Unzuverlässigkeit dieser Methode während seiner Arbeit hin, sodass er als Merkmalspunkt den Schnittpunkt von Nasensymmetrieachse und Nasenbasislinie wählt. Für das hier entwickelte Verfahren hat sich die erste Methode jedoch bewährt und wird deshalb verwendet. Die unterschiedlichen Ergebnisse liegen wahrscheinlich in der Verwendung von anders beschaffenem Ausgangsmaterial für beide Arbeiten begründet.

Schritt 7 Der Suchbereich für die Nasenlöcher wird mittels der Kenntnis über die Position der Nasensymmetrieachse und der Nasenbasislinie eingegrenzt.

Schritt 8 Die Nasenlöcher entsprechen dunklen, kreisförmigen Strukturen. Die Lokalisierung erfolgt am besten mittels Template Matching im Intensitätskanal. Die Templategröße wird in Abhängigkeit der Breite des Nasenbildausschnitts gewählt:

$$\text{Breite} = 0.1 \cdot \text{Ausschnittsbreite}, \text{ Höhe} = 0.05 \cdot \text{Ausschnittsbreite}.$$

Das linke und das rechte Nasenloch wird dann separat auf der jeweiligen Seite der Symmetrieachse gesucht. Das Maß für die Übereinstimmung von Template und Bildregion bildet die Integration der Intensität über den Templatebereich an dieser Position. Weil die binäre Template Maske ein Modell des invertierten Nasenlochs darstellt (weiße Ellipse auf schwarzem Hintergrund), entspricht die Position der höchsten Übereinstimmung dem Pixel mit der niedrigsten Bewertung.

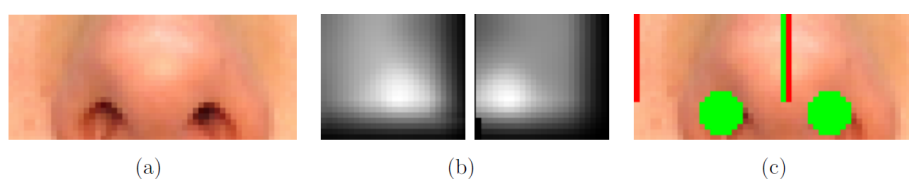


Abbildung 3.9: Template Matching: (a) Ausgangsbild, (b) Bewertungskarten für linkes und rechtes Nasenloch (helle Punkte repräsentieren eine gute Template-Bewertung), (c) Gewinnerpaar, entnommen Hähnel (2008), Seite 106

3.1.2.4 Mund

Der Mund ist sehr formvariabel und auch seine Farbe oder Intensität sind als Merkmale nicht spezifisch genug, um allein als Lokalisierungsgrundlage dienen zu können. Hähnel erstellt innerhalb seines Verfahrens deshalb mehrere Merkmalskarten, die verschiedene Eigenschaften herausbilden und später zur genauen Konturlokalisierung kombiniert werden. Zwar wird sein Verfahren innerhalb dieser Arbeit nur im Ansatz verwendet, aber auch die hier entwickelte Methode dient dazu, die Lippenfarbe gegenüber der umgebenden Hautfarbe mittels verschiedener Berechnungen stärker hervorzuheben als es im RGB-Farbraum der Fall ist. Letztendlich dient jedoch eine Wasserscheidentransformation mit anschließender Konturapproximation der Lokalisierung des Mundes.

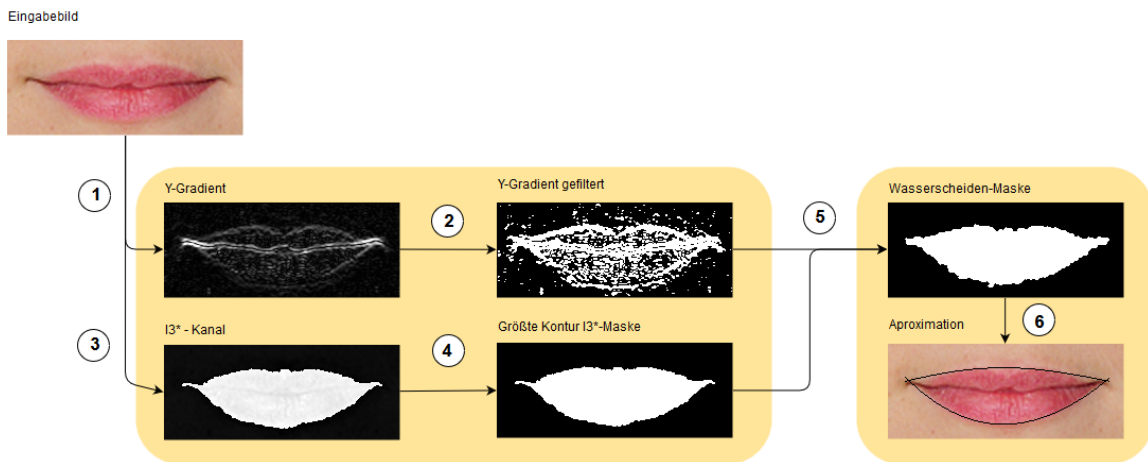


Abbildung 3.10: Mundlokalisierung

Schritt 1 Als Vorbereitung für die Extraktion der Kontureigenschaft des Mundes wird der Y-Gradient berechnet.

Schritt 2 Für die Herausbildung der Konturinformation wird auf den Y-Gradienten ein Schwellenwert angewendet, Das Resultat ist eine Binärmaske, in der Pixel mit einem Gradientenbetrag oberhalb 23 weiß und alle restlichen Pixel schwarz dargestellt werden.

Schritt 3 In diesem Schritt wird die farbliche Segmentierung vorbereitet. Hähnels Verfahren nutzt zur dazu sowohl den HSV- als auch den I1I2I3-Farbraum. Aus ersterem berechnet er den sogenannten logarithmischen Hue, eine abgewandelte Form des Hue-Kanals des HSV-Farbraums. Allerdings filtert wird dieser Kanal danach mit einer Funktion gefiltert, die mit durchschnittlichen Lippenfarbtönen arbeitet, die für diese Arbeit nicht zur Verfügung stehen, weil Hähnel sie anhand einer Bilddatenbank berechnet hat. Deshalb wird hier auf die Verwendung

3 Implementierung der Gesichtsanalyse

des HSV-Farbraums gänzlich verzichtet. Der I1I2I3-Farbraum ermöglicht laut Hähnel's Recherchen eine gute Separierbarkeit von farblich zusammenhängenden Bereichen. Die einzelnen Kanäle werden wie folgt aus den RGB-Kanälen berechnet:

$$I_1 = \frac{R+G+B}{3}$$

$$I_2 = \frac{R-B}{2}$$

$$I_3 = \frac{2G-R-B}{4}$$

Innerhalb dieses Verfahrens wird der I3 Kanal verwendet, weil in ihm der Kontrast von den Lippen zur Haut am stärksten ausgeprägt ist. Noch bessere Segmentierungsergebnisse wurden mit dem gewichteten I3* - Kanal erzielt. Für dessen Berechnung fließt der Blau-Kanal nur mit einem Gewichtungsfaktor 0,5 ein, weil er für die Lippenfarbe eine untergeordnete Rolle spielt.

$$I_3^* = \frac{2G-R-0.5 \cdot B}{4}$$

Schritt 4 In diesem Schritt erfolgt die farbliche Segmentierung, indem auf den I3*-Kanal ein empirisch gewählter Schwellenwert angewendet wird. Aus der Schwellenwertsegmentierung resultiert eine Binärmaske, in der nur die Lippenregion weiß erscheint. Diese Maske ist für eine genaue Konturlokalisierung zwar zu ungenau, zeigt jedoch sie mit Sicherheit die Region auf, in der sich die Lippen befinden.

Schritt 5 Eine eventuell auftretende Gesichtsbehaarung der abgebildeten Person kann das Segmentierungsergebnis verfälschen. Um dem entgegenzuwirken, wird in Anlehnung an das Verfahren zur Augenbrauendetektion auch hier eine markerbasierte Wasserscheidentransformation angewendet, deren Marker anhand der zuvor erstellten Masken berechnet werden. Für den Vordergrund wird die Markierung innerhalb der größten Kontur der Lippenmaske gesetzt, weil diese Region mit Sicherheit die Lippen repräsentiert. Der Hintergrund wird markiert, indem der Bereich gewählt wird, für den die Konturen innerhalb der Y-Gradienten-Maske keine Überschneidungen mit der Lippenregion aufweisen, da dies mit großer Wahrscheinlichkeit die fälschlich in der Maske enthaltenen Bereiche des Bartes sind.

Schritt 6 Für die Approximation der Mundkontur wird im letzten Schritt jeweils für die obere und untere Hälfte der größten Kontur der Segmente der Wasserscheidentransformation eine Least-Squares-Anpassung vorgenommen. Die Oberlippenkante wird durch ein Polynom 4. Grades und die Unterlippenkontur durch ein Polynom 3. Grades genügend genau angenähert.

3.1.2.5 Kinnlinie

Das Kinn stellt in Bezug zu den anderen Gesichtsmerkmalen eine Besonderheit dar, weil es weder geometrische noch farbliche Merkmale aufweist, anhand derer eine Detektion vorgenommen werden kann. Hähnel beschreibt, dass die einzige Möglichkeit für eine Lokalisierung der Kinnkontur in der Analyse von Kanteninformation liegt, wobei die Kenntnis der ungefähren Position der Kinnlinie erforderlich ist, um Anhaltspunkte für die Suche zu haben. Sein Verfahren beruht deshalb auf einer Analyse des Gradientenbildes und der Verwendung einer initialen Annäherung an die Kinnkontur, die als Ausgangspunkt für die lokale Suche genutzt wird. Dieses Verfahren findet auch innerhalb dieser Arbeit Anwendung, wobei die Initialkurve für die Kinnkontur anhand der vom Nutzer eingegebenen Kinnpunkte berechnet wird.

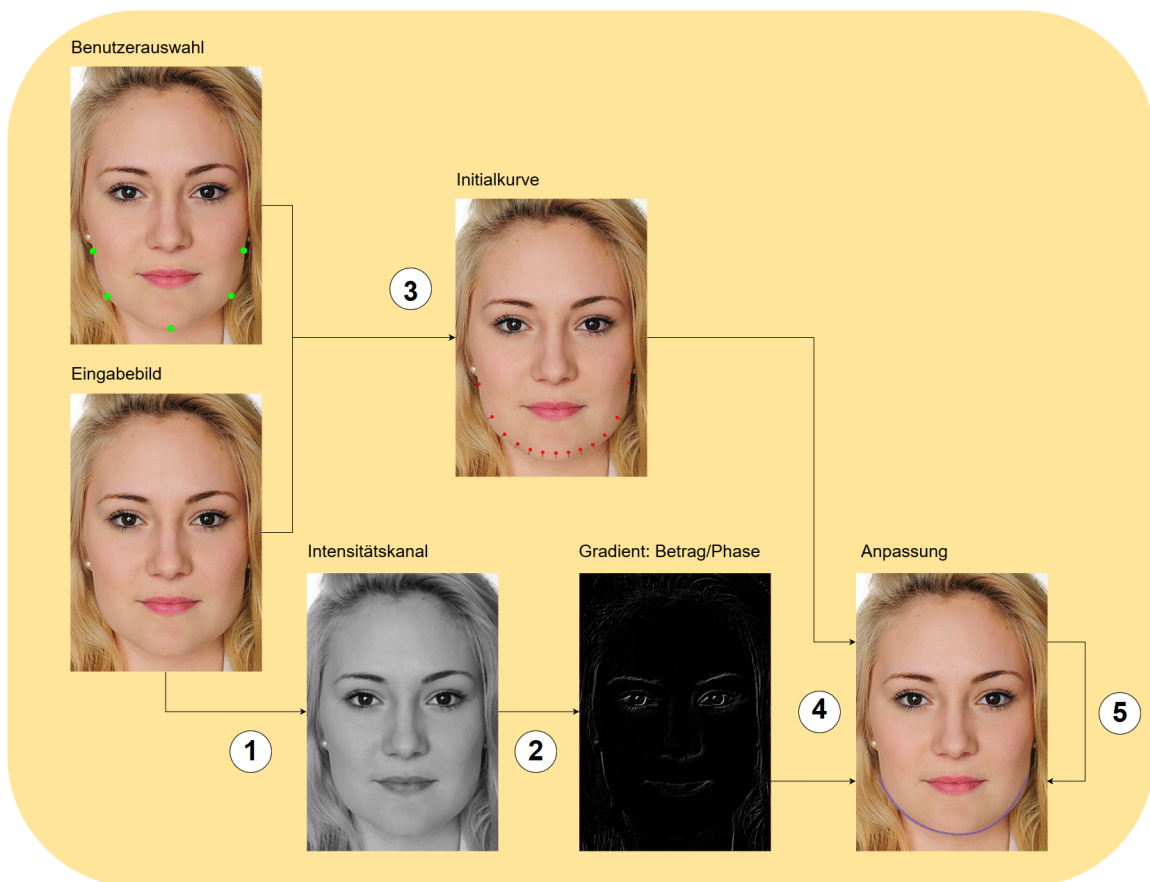


Abbildung 3.11: Kinnlokalisierung (Frau4)

Schritt 1 Als Grundlage für weitere Schritte wird der Intensitätskanal des Eingabebildes extrahiert und mittels Medianfilter entrauscht. Dieser entfernt Störeinflüsse, ohne wichtige Kanteninformation zu glätten.

Schritt 2 Als nächstes werden aus dem Intensitätskanal Betrag und Phase des Gradienten berechnet.

Schritt 3 Anhand der übergebenen Punkte aus der Benutzerauswahl wird als Ausgangskurve für die lokale Suche der Kinnkontur ein initiales Polynom berechnet. Hierzu wird die Polygonlinie, die durch die Merkmalspunkte des Benutzers aufgespannt wird, durch eine Least Squares Lösung angenähert. Das resultierende Polynom 4. Grades liegt immer vollständig innerhalb der tatsächlichen Kinnkontur, weil Verbindungsgeraden zwischen den Merkmalspunkten der Benutzerauswahl im Inneren des Kinns verlaufen. Anschließend werden auf der Initialkurve weitere Merkmalspunkte berechnet, die das Polynom in insgesamt 12 Kurvensegmente unterteilen.

Schritt 4 Die zuvor berechnete initiale Kinnkontur wird nun iterativ an die Kanteninformation aus dem Gradientenbild angeglichen. Dafür werden die Merkmalspunkte des Polynoms pixelweise auf ihrer Normalen zum Gesichtsrand hin verschoben und anschließend das Polynom nach jedem Durchgang neu berechnet. So nähern sich die Merkmalspunkte auf dem Initialpolynom Schritt für Schritt ihrer Position auf der tatsächlichen Kinnkontur an.

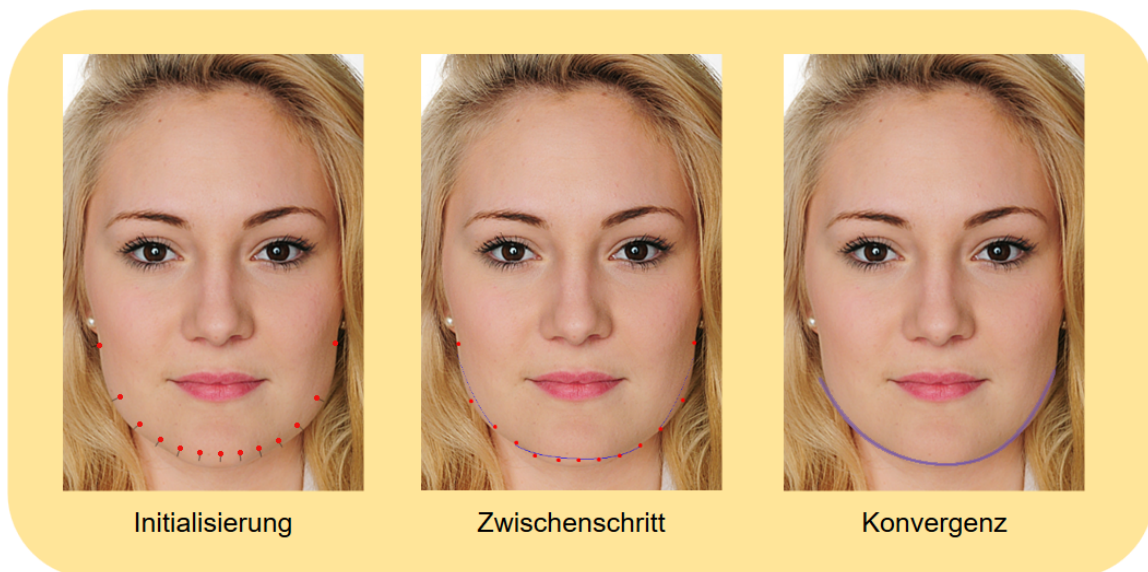


Abbildung 3.12: Iterationsweise Anpassung der Initialkurve an die Kinnkontur (Frau4)

Schritt 5 Die iterationsweise Verschiebung der Merkmalspunkte soll nur bis zu dem Schritt erfolgen, an dem die dazwischenliegenden Kurvensegmente die Zielkurve ausreichend genau annähern. Ob dies zutrifft, wird anhand der Betrags- und Phaseninformation des Gradientenbildes festgestellt. Für die Pixel der Kurvensegmente zwischen den Merkmalspunkten des Kinnpolynoms wird dazu in

jedem Iterationsschritt jeweils verglichen, ob die Ableitung der Kurve an dieser Stelle ausreichend genau mit der Orientierung des Gradienten übereinstimmt und ob außerdem entweder die betrachtete Kante hinreichend ausgeprägt ist oder im Intensitätskanal ein lokales Minimum erreicht wurde. Trifft dies auf mindestens die Hälfte der Pixel eines Kurvenabschnitts zu, werden die Merkmalspunkte dieses Teilstücks fixiert und nicht weiter verschoben. Alle bisher nicht fixierten Merkmalspunkte werden in folgenden Iterationsschritten weiter bewegt. Die Laufdauer des Verschiebungsprozesses wird durch eine maximale Anzahl an Iterationen begrenzt, sodass der Anpassungsversuch für einige Merkmalspunkte abgebrochen wird, wenn keine Position erreicht wurde, in der eine Fixierung erfolgte. Diese Punkte bleiben bei der finalen Berechnung der Approximation der Kinnkontur als Polynom 4. Grades unberücksichtigt.

3.2 Ausgabedaten der Analyse

Die Nutzung der Analysedaten für die Modifikation des generischen 3D-Gesichtsmodells bedingt die Bereitstellung dieser in einer dafür erforderlichen Form. Das heißt, es müssen sogenannte korrespondierende Punkte im Gesichtsbild berechnet werden, die an den Positionen liegen, die mit bestimmten vorher definierten Knotenpunkten des 3D-Gesichtsmodells übereinstimmen. Dazu werden innerhalb der Merkmalskonturen, die als Punktmengen vorliegen, für jedes Merkmal die entsprechenden Korrespondenzpunkte gesucht und für die Weiterverarbeitung in einem Vektor abgespeichert.

Anschließend muss für jeden Korrespondenzpunkt sein Verhältnis zur Position der Parallelprojektion des Gesichtsreferenzpunktes ermittelt werden. Dieser stellt im 3D-Gesichtsmodell quasi den „Ursprung“ für die Vektorrechnung bei der Modifikation dar, also den Ausgangspunkt für alle erfolgenden Berechnungen wie Skalierungen und Verschiebungen. Um die Lage aller Korrespondenzpunkte im Verhältnis zum Gesichtsreferenzpunkt übermitteln zu können, wird für jeden dieser Punkte der Richtungsvektor vom Referenzpunkt P_{ref} zum jeweiligen Korrespondenzpunkt P_{kor} berechnet und in einem Vektor übergeben. Dieser Vektor ist ein verschachtelter Vektor, welcher als Wörterbuch fungiert, in dem für die abgespeicherten Werte eine Zuordnungsvorschrift existiert, die für den Auslesevorgang der Informationen erforderlich ist.

Die Berechnung der Merkmalskorrespondenzen und der zugehörigen Gesichtsverhältnisvektoren wird in den folgenden Unterkapiteln aufgegriffen und beschrieben. Die Gesichtsverhältnisvektoren sind die Daten, die zur Weiterverarbeitung bereitgestellt werden, also die Ausgabedaten des Analysesystems.

3.2.1 Berechnung der Merkmalskorrespondenzen

Die Merkmalskorrespondenzpunkte sind die Punkte auf den Konturen der Gesichtsmerkmale, die charakteristisch für ein bestimmtes Merkmal sind und dieses in seiner

3 Implementierung der Gesichtsanalyse

Position und der Ausprägung seiner Form quantitativ beschreiben. Sie werden dazu genutzt, die Frontalansicht des 3D-Gesichtsmodells an das analysierte Gesichtsbild anzupassen, indem entsprechend zugewiesene Merkmalspunkte auf dem Modell an die aus dem Gesichtsbild berechnete Position verschoben werden. Jedes Merkmal wird durch eine unterschiedliche Anzahl an Korrespondenzpunkten ausreichend genau beschrieben.

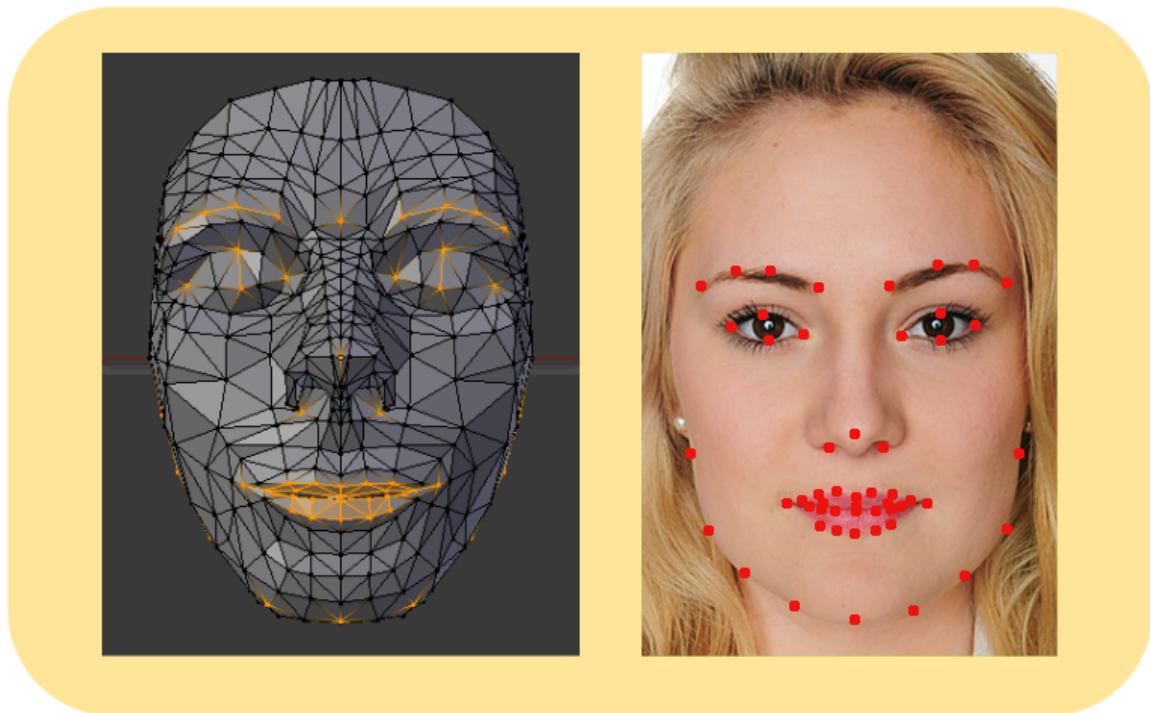


Abbildung 3.13: Korrespondenzpunkte - links: Markierte Korrespondenzpunkte im generischen Gesichtsmodell, rechts: Korrespondenzpunkte einer Frontalansicht (Frau4)

Abbildung 3.13 zeigt alle verwendeten Korrespondenzpunkte, für die die entsprechenden korrespondierenden Punkte auf den 2D-Gesichtsbildern berechnet werden sollen. Die Bestimmung ihrer Position auf dem Gesichtsbild anhand der Analysedaten wird nachfolgend beschrieben:

Die Informationen, die für die einzelnen Merkmale vorliegen, sind in den meisten Fällen Konturen und Merkmalspunkte, aus denen die Lage der jeweiligen Korrespondenzpunkte erst noch berechnet werden muss. Diese Berechnung erfolgt nach einem grundsätzlich geltenden Schema:

Zunächst werden alle Punkte bestimmt, deren Position eindeutig definiert ist, weil keine weiteren in Frage kommenden Punkte für das jeweilige Bestimmungskriterium existieren. Solche Punkte sind zum Beispiel die Augenwinkel, die die jeweils äußersten Punkte Augenkontur sind. Selbiges trifft auch auf die Mundwinkel innerhalb der

Mundkontur sowie die inneren und äußeren Endpunkte der Augenbrauen zu. Außerdem wird auch der oberste Punkt des Auges eindeutig durch die Öffnungshöhe festgelegt und auch die Nasenspitze ist klar definiert, weil sie zuvor schon als hellster Punkt auf der Nasenmittellachse bestimmt wurde. Ebenso können die beiden Korrespondenzpunkte, die die Gesichtsbreite definieren und die Kinnkontur nach oben begrenzen, in ihrer Position genau bestimmt werden, indem auf die Information der Nutzereingabe zurückgegriffen wird. Diese Punkte werden aus zwei Gründen nicht berechnet. Zum einen werden sie ausreichend genau vom Benutzer markiert und zum anderen entsteht bei der Bestimmung ihrer Y-Koordinate anhand des Approximationspolynoms für die Kinnkontur eine große Ungenauigkeit, die in der starken Steigung des Grafen an dieser Stelle begründet liegt.

Auf Grundlage dieser Merkmalspunkte mit eindeutig definierter Position werden die anderen benötigten Korrespondenzen berechnet, deren Lage nur durch ihr Verhältnis zu den bereits bestimmten Punkten definiert ist. Diese Punkte liegen zum Beispiel zwischen den äußersten Punkten auf einer Kontur und unterteilen diese in einem bestimmten Verhältnis. Hierzu wurden innerhalb des 3D-Gesichtsmodells für alle Korrespondenzpunkte horizontale Streckenverhältnisse berechnet, die nun zur Berechnung der restlichen Korrespondenzpunkte auf die Konturen der Merkmale angewendet werden. Dadurch wird die X-Koordinate der Punkte bestimmt, mittels derer die entsprechende y-Koordinate berechnet wird.

3.2.2 Berechnung der Gesichtsverhältnisvektoren

Die Gesichtsverhältnisvektoren sind die Daten, die letztlich für die Modifikation des 3D-Gesichtsmodells bereitgestellt werden. Sie werden in Form von Richtungsvektoren übergeben.

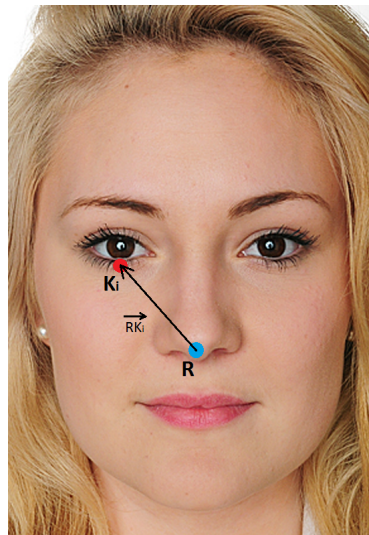


Abbildung 3.14: Gesichtsverhältnisvektor (Frau4)

3 Implementierung der Gesichtsanalyse

Die Berechnung der Gesichtsverhältnisvektoren erfolgt nach der Formel

$$\overrightarrow{RK_i} = \vec{K}_i - \vec{R}$$

mit $\overrightarrow{RK_i}$ = Gesichtsverhältnisvektor eines Korrespondenzpunkts i ,
 \vec{K}_i = Ortsvektor des Korrespondenzpunkts i ,
 \vec{R} = Ortsvektor des Gesichtsreferenzpunkts

Die Gesichtsverhältnisvektoren besitzen durch die anfängliche Normierung des Eingabebildes die Länge der Entfernung vom Gesichtsreferenzpunkt zu dem jeweiligen Merkmalspunkt auf dem normierten Eingabebild. Sie zeigen in die Richtung vom Gesichtsreferenzpunkt zu dem jeweiligen Merkmalskorrespondenzpunkt und dienen in der Arbeit von [Burmeister \(2016\)](#) als Grundlage für die Berechnung der Verschiebungsvektoren. Die Verschiebungsvektoren beschreiben, wie die Korrespondenzpunkte des generischen Gesichtsmodells bewegt werden müssen, um in der Frontalansicht des 3D-Modells an der gleichen Position zu liegen, wie die entsprechenden Merkmalspunkte auf dem Gesichtsbild es vorgeben. Mittels der Verschiebungsvektoren erfolgt die Anpassung des 3D-Modells an das Gesicht auf dem Eingabebild

4 Ergebnisse

Anfangs wurde die Frage gestellt, ob eine teilautomatisierte, zielgerichtete Analyse eines geeigneten Frontalbildes eines Gesichts eine ausschließlich manuelle Datenerfassung gleichwertig ersetzen kann, in einer Art und Weise, die den Benutzeraufwand reduziert. Für ihre Beantwortung wird eine Bewertung des Analyseverfahrens anhand eines Tests durchgeführt. Dafür werden verschiedene Ausgangssituationen erzeugt und die Reaktion des Systems darauf dokumentiert. Das Eingabematerial für den Test besteht aus zwei Gruppen. Eine davon enthält Bilder, die den Systemanforderungen für eine erfolgreiche Analyse gerecht werden. Die andere Gruppe ist eine Kontrollgruppe mit Bildern, für die aus bekannten Gründen keine erfolgreiche Analyse erwartet wird. Die folgenden beiden Unterkapitel befassen sich mit der genauen Versuchsbeschreibung und der anschließenden Auswertung der Testergebnisse.

4.1 Versuchsbeschreibung

Der folgend beschriebene Versuch gibt Aufschluss über das Systemverhalten unter verschiedenen Bedingungen und liefert die statistische Grundlage für die Bewertung der Leistungsfähigkeit des Analysesystems. Es wird insbesondere den folgenden zwei Fragen nachgegangen:

1. Wie reagiert das System bei gleichem Muster für die Benutzerauswahl (Bei Einhaltung der empfohlenen Vorgehensweise für die Auswahl) auf verschiedenes Eingabematerial?
2. Wie reagiert das System für auf unterschiedliche Arten der manuellen Merkmalsmarkierung bei ein und dem selben Eingabebild?

Für die Beantwortung der ersten Untersuchungsfrage wird das System mit zwei Gruppen von Eingabebildern getestet. Eine Gruppe besteht aus anforderungsgerechten (den Anforderungen des Programms entsprechenden) Eingabebildern, die für die Analyse zulässig sind, sich aber bezüglich ihrer Eigenschaften deutlich voneinander unterscheiden. Die Bilder beinhalten also verschiedene menschliche und technische Einflussfaktoren, die eine entsprechende Systemreaktion hervorrufen und später Rückschlüsse auf ein Muster zulassen.

Die andere Gruppe enthält für die Analyse unzulässiges, also den Anforderungen des Systems nicht entsprechendes Bildmaterial. Sie wird als Kontrollgruppe verwen-

det, für die aufgrund bekannter Eigenschaften der Fotos keine erfolgreiche Analyse zu erwarten ist.

Der Programmdurchlauf erfolgt für beide Gruppen mittels einer manuellen Merkmalsmarkierung, die den Vorgaben des Schemas entspricht, das das Programm als Hilfestellung oder Anhaltspunkt zur Verfügung stellt.

Die Gruppe der anforderungsgerechten Bilder enthält zwölf Bilder, davon sind sechs Fotos Abbildungen von Frauen, auf vier Fotos sind Männer abgebildet und zwei Fotos bilden Kinder ab, ein Mädchen und einen Jungen. Die Gruppe enthält Bildmaterial zum einen aus dem Internet und zum anderen aus der Datenbank Caltech Human face (Front) dataset. Alle Personen wurden frontal fotografiert, sind gut ausgeleuchtet, und die Abbildung hat eine ausreichende Auflösung.

Die Kontrollgruppe enthält elf Fotos, darin enthalten sind jeweils vier Abbildungen von Männern und Frauen und außerdem drei Abbildungen von Kindern, zwei Jungen und ein Mädchen. Diese Fotos weichen in ein oder mehreren Eigenschaften von den Vorgaben dafür ab: Getestet werden Fotos mit verdeckten Gesichtspartien, ungünstigen Gesichtsmerkmalen, unvorteilhaften Beleuchtungsverhältnissen, geringer Auflösung und eingeschränkter Farbinformation. Das entsprechend betrachtete Kriterium ist in der Bezeichnung der Fotos jeweils vermerkt.

Der Versuch zur Beantwortung der zweiten Untersuchungsfrage (Einfluss unterschiedlicher Arten der Merkmalsmarkierung), wird mit vier Fotos aus der anforderungsgerechten Gruppe durchgeführt. Diese setzen sich zusammen aus zwei Bildern, die innerhalb des ersten Tests gute Analyseergebnisse erzielen und zwei Bildern, für die die Analyse nur mäßige, also verbesserungswürdige Ergebnisse ergibt. Der Test besteht darin, den Analysevorgang für jedes dieser Bilder mehrfach zu durchlaufen, indem bei jedem Durchgang eine andere Merkmalsmarkierung durch den Benutzer erfolgt. So kann geprüft werden, inwieweit eine bestimmte Art der Benutzerauswahl den Erfolg der Analyse begünstigt oder hemmt.

Zunächst erfolgt ein Durchlauf mit vorgabengetreuer manueller Merkmalsmarkierung. Sie entspricht also den Empfehlungen des Schemas, das dem Benutzer als Hilfestellung angeboten wird. In drei weiteren Analyse-Durchgängen wird die Benutzerauswahl variiert, sodass sie zwar die Merkmale markiert, aber im Rahmen des Interpretationsspielraums des Benutzers von der eigentlichen Vorgabe abweicht:

1. In der ersten Variation werden die Suchfenster für die Merkmale groß gewählt, sodass im Verhältnis zum Merkmalsbereich auch viel Hintergrund eingeschlossen wird.
2. In der zweiten Variation wird der Bereich der Suchfenster wiederum sehr klein gewählt, sodass die Fenstergröße nahezu an den Merkmalskonturen abschließt.
3. Die dritte Variation der Benutzerauswahl bezieht sich nicht auf die Größe der Markierungsbereiche, sondern auf ihre Position. In diesem Durchlauf werden

4 Ergebnisse

die Markierungen unsymmetrisch vorgenommen, zum einen in Bezug auf die Position des Merkmals innerhalb des Suchfensters, zum anderen auch auf die Position der Suchfenster innerhalb des Gesichts. So ergibt sich für die gesamte Auswahl eine unsymmetrische Anordnung.

Die Markierung der Kinnlinie ist bei allen Testdurchläufen gesondert zu betrachten, weil sie nicht anhand eines Auswahlfensters vorgenommen wird, sondern mittels eines Polygonzugs. Dieser wird in den drei Analyse-Durchgängen variiert, indem die Unterteilungspunkte auf der Kinnlinie verschoben werden. Für die *große* Benutzerauswahl werden die beiden äußersten Teilstrecken auf dem Polygonzug kurz gewählt, sodass die Unterteilungspunkte dort dicht aneinander liegen. Im zweiten Durchlauf wird der Polygonzug annähernd spitz geformt, sodass die Teilstrecken nahe des Kinnmittelpunktes sich verkürzen. In der dritten Variation wird eine unsymmetrische Anordnung der Punkte auf der Kinnlinie erzeugt, das heißt, dass die Teilstrecken unterschiedlich lang gewählt werden und auch der Polygonmittelpunkt nicht auf dem Kinnmittelpunkt liegt.

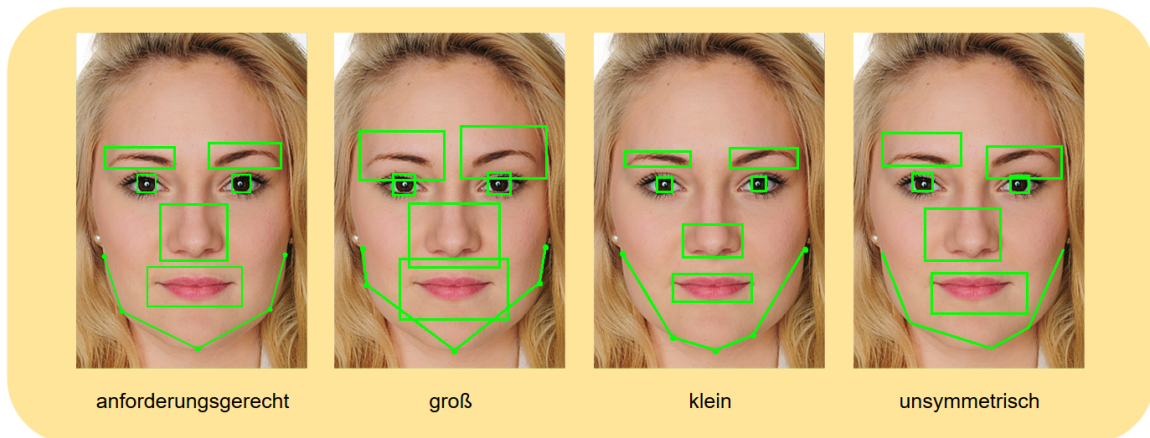


Abbildung 4.1: Verschiedene Formen der Markierung (Frau4)

4.2 Versuchsauswertung

Die Versuchsauswertung wird anhand der Wertetabellen der Analyseergebnisse für jede Versuchsreihe vorgenommen. Jede Tabelle wird zunächst dargestellt und anschließend kurz beschrieben und erläutert, um danach eine statistische Bewertung vorzunehmen.

Die Tabelle 4.1 zeigt die Testergebnisse der Eingabebilder, für welche eine erfolgreiche Analyse anzunehmen ist. Die ersten fünf Spalten sind mit den Merkmalen

4 Ergebnisse

bezeichnet, anhand derer die Analyse erfolgt. Dahinter ist jeweils die Anzahl der Korrespondenzpunkte, die das Merkmal beschreiben, vermerkt. Für jedes Testbild wird dann zeilenweise aufgeführt, wie viele der geforderten Korrespondenzpunkte vom Analyseprogramm ausreichend genau (nach subjektivem Ermessen) berechnet wurden. Die letzten beiden Spalten dienen der Gesamtbewertung der Gesichtsanalyse pro Testbild. Dafür wird einerseits eine subjektive, farbliche Bewertungsskala von Rot über Orange, Gelb, Hellgrün bis Grün angewandt, wobei Rot für schlechte und Grün für gute Analyseergebnisse steht.

Diese Bewertung bezieht sich auf den Gesamteindruck der berechneten Korrespondenzpunkte. Das heißt, sie bewertet die Lage und Übereinstimmung aller Punkte auch in Abhängigkeit zu den anderen Punkten und stellt ein qualitatives Maß nach subjektivem Empfinden dar. In der letzten Spalte ist für jedes Eingabebild aufgeführt, wie viel der prozentuale Anteil der richtig berechneten Korrespondenzpunkte beträgt. Dies ist ein rein quantitatives Maß zur Bewertung und berücksichtigt weder die menschliche Wahrnehmung noch die Abhängigkeit bestimmter Punkte zueinander und eignet sich deshalb nur Richtwert für eine Bewertung der Analyseleistung.

	Brauen (8)	Augen (8)	Nase (3)	Mund (12)	Kinn (7)	Summe	Bewertung	%
Barbie	8	6	1	9	6	30	1	78,95
Frau 1	8	8	3	12	7	38	2	100
Frau 2 1	8	7	2	12	7	36	2	94,74
Frau 2 2	7	8	3	12	7	37	2	97,37
Frau 3	5	4	3	11	6	29	1	76,32
Frau 4	8	8	3	12	7	38	2	100
Junge 1	4	4	3	9	5	25	-1	65,79
Mädchen 1	8	7	3	12	7	37	2	97,37
Mann 1	7	8	1	4	5	25	0	65,79
Mann 2	4	7	2	8	5	26	0	68,42
Mann 3	5	7	2	12	5	31	1	81,58
Mann 4	8	5	2	11	6	32	1	84,21
Summe	80	79	28	124	73			
Anzahl vollständig erkannt	6	4	6	6	5			
Anzahl Korresp. Insges.	96	96	36	144	84			
Erkannt in %	83	83	70	86	87			

Tabelle 4.1: Versuch Analysequalität - Legende: Rot = ungenügende Formbeschreibung (nicht enthalten), Orange = ungenaue Formbeschreibung, Gelb = teilweise ungenaue Formbeschreibung, Hellgrün = ausreichende Formbeschreibung, Grün = exakte Formbeschreibung

Die Tabelle zeigt, dass fünf von insgesamt 12 getesteten Eingabebildern zufriedenstellend analysiert werden konnten, das heißt, das Gesicht wird durch die Ergebnisse der Analyse richtig oder ausreichend genau beschrieben. Für vier Testbilder liefert die Analyse gute, aber teilweise ungenaue Ergebnisse, das heißt, die Gesichtsmerkmale werden durch die berechneten Korrespondenzpunkte ausreichend, aber nicht exakt beschrieben. Bei zwei Bildern der Gruppe liefert die Analyse nur mäßige Ergebnisse, weil zwar die überwiegende Anzahl an Korrespondenzpunkten richtig berechnet wurde, die restlichen Punkte jedoch zum Teil an stark abweichender Position liegen.

4 Ergebnisse

Für ein Foto ergibt die Analyse kein zufriedenstellendes Ergebnis, weil die Merkmale durch einige der berechneten Korrespondenzpunkte falsch beschrieben werden. Das Ergebnis des Tests ist positiv zu bewerten, weil für 75% der Eingabebilder ein gutes Analyseergebnis erzielt wurde. Es ist auffällig, dass der Analyseerfolg bei Eingabefotos von männlichen Gesichtern geringer ist als bei Fotos von Frauen.

Betrachtet man die Ergebnisse merkmalsbezogen, ist festzustellen, dass die Prozentsätze für die Gesamtanzahl der richtig berechneten Korrespondenzpunkte für alle Bilder oberhalb von 70% liegt. Außerdem weisen sie nur geringe Differenzen zueinander auf. Diese Tatsache zeigt, dass die unterschiedlichen Verfahren zur Lokalisierung der Merkmale ähnlich stabil und zuverlässig arbeiten. Zwar ist der Prozentsatz der Anzahl der richtig berechneten Korrespondenzpunkte für die Nase am geringsten, jedoch ist eine Bewertung der Nasenerkennung anhand des Prozentsatzes nicht aussagekräftig. Genau wie der Mund und die Augenbrauen wurde die Nase in 6 der 12 Fälle vollständig richtig erkannt. Ein Korrespondenzpunkt der Nase entspricht allerdings einem höheren Prozentwert als dem eines Korrespondenzpunktes eines der anderen Gesichtsmerkmale, das insgesamt durch eine höhere Anzahl von Korrespondenzpunkten beschrieben wird.

Für die Auswahl der Bilder dieses Versuchs wurde eine positive Tendenz für die Erkennungsleistung angenommen, da die Fotos augenscheinlich den Anforderungen des Systems entsprechen. Insgesamt fällt das Ergebnis des Tests erwartungsgemäß aus.

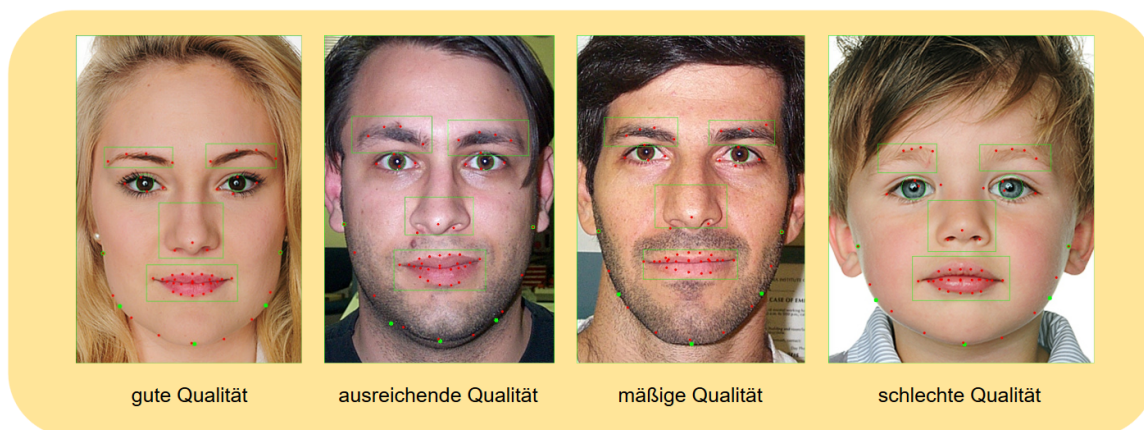


Abbildung 4.2: Beispiele der Bewertung der Analysequalität - von links nach rechts: Frau4, Mann3, Mann2 Mimik, Junge1

Die Tabelle 4.2 zeigt die Versuchswerte der Kontrollgruppe, also der Eingabebilder, für die keine erfolgreiche Analyse erwartet wurde. Auch hierbei sind die Gesichtsmerkmale einzeln spaltenweise vermerkt. Jedem Eingabebild wurden zwei Zeilen zu-

4 Ergebnisse

gewiesen.

Die obere enthält die Erwartungswerte der Anzahl an Korrespondenzpunkten, die für jedes Merkmal angenommen werden. Die Prognose basiert auf den jeweiligen Eigenschaften des Fotos. Ist beispielsweise im Eingabebild der Mund verdeckt, wird die Erkennung des Mundes beeinträchtigt und entsprechend eine geringe Anzahl an richtig berechneten Korrespondenzpunkten erwartet.

In der unteren Zeile stehen die tatsächlich vom Analysesystem zurückgegebenen Werte für die Merkmale. Die Bewertung des Versuchs erfolgt anhand der Skala *Erwartung unterschritten*, *Erwartung erfüllt* und *Erwartung übertroffen*, was in der letzten Tabellenspalte aufgeführt ist.

	Brauen (8)	Augen (8)	Nase (3)	Mund (12)	Kinn (7)	Summe	Bewertung
Frau gedreht							
	5	3	2	1	1	12	
	5	4	1	3	0	13	1
Frau Schmuck							
	8	8	1	10	7	34	
	8	5	1	7	6	27	-1
Frau Verdeckung							
	0	8	1	12	5	26	
	1	8	3	12	7	31	1
Junge Aufl Licht							
	8	4	3	2	3	20	
	8	7	3	9	2	29	1
Junge Brille							
	8	8	3	12	5	36	
	8	6	3	12	7	36	0
Mädchen Augenbrauen							
	1	8	3	12	5	29	
	2	8	3	7	6	26	-1
Mann Highkey							
	2	4	2	0	3	11	
	2	8	3	3	4	20	1
Mann 2 Mimik							
	8	8	1	0	5	22	
	8	7	1	9	7	32	1
Mann 3 Beleuchtung							
	5	8	2	0	2	17	
	6	5	3	0	7	21	1
Mann 5 Bart							
	8	6	1	0	2	17	
	5	4	1	7	5	22	1

Tabelle 4.2: Versuch Analysequalität mit Kontrollgruppe - Legende: Blau = Erwartete Anzahl der Korrespondenzpunkte, Magente = Erzielte Anzahl der Korrespondenzpunkte, Rot = geringere Anzahl als erwartet, Gelb = übereinstimmende Anzahl, Grün = höhere Anzahl

Die Gesichtsanalyse der Kontrollgruppe erreichte für sieben der insgesamt zehn Bilder bessere Ergebnisse als erwartet, weil die Anzahl der richtig berechneten Korrespon-

4 Ergebnisse

denzpunkte den zuvor angenommenen Wert dafür überschreitet. Bei zwei Testbildern liefert die Analyse schlechtere Ergebnisse als erwartet, weil die Anzahl der richtig berechneten Korrespondenzpunkte den zuvor angenommenen Wert unterschreitet. In einem Fall stimmt das erwartete mit dem erreichten Analyseergebnis überein und es wurden durchschnittlich so viele Korrespondenzpunkte richtig berechnet wie zuvor angenommen.

Tendenziell wurde die erwartete Erkennungsleistung des Systems für diesen Versuch überschritten oder bestätigt. Das heißt, das System wird von Störfaktoren weniger beeinflusst als erwartet, sodass das Gesamtergebnis des Versuchs positiv bewertet wird.

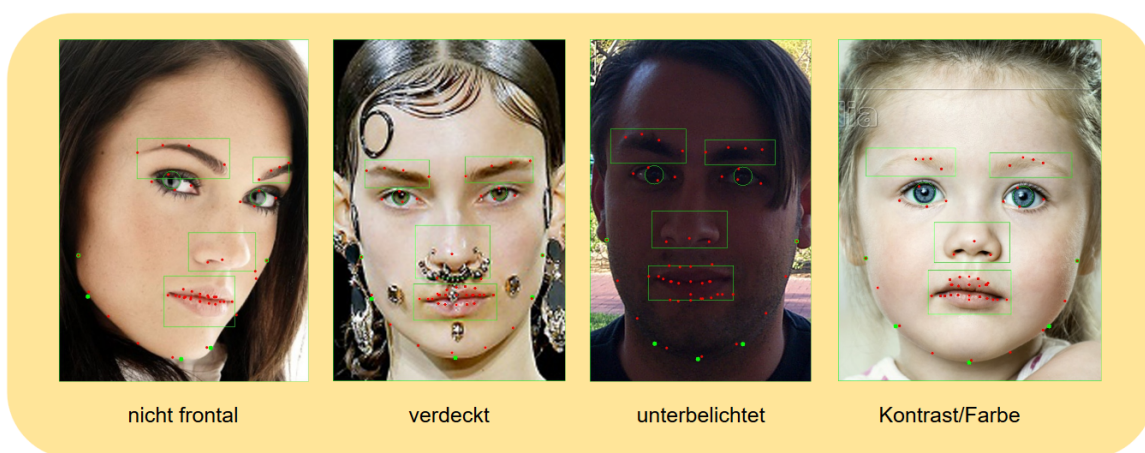


Abbildung 4.3: Beispiele der Ergebnisse der Kontrollgruppe - von links nach rechts: Frau gedreht, Frau Schmuck, Mann3 Beleuchtung, Mädchen Augenbrauen

Die Tabelle 4.3 zeigt die Ergebnisse des Versuchs mit variierender Merkmalsmarkierung durch den Benutzer. Jedem Testbild werden vier Zeilen zugeordnet, in denen die Werte der verschiedenen Markierungsvarianten dargestellt sind. Die Werte geben die Anzahl der richtig berechneten Korrespondenzpunkte für jedes Merkmal bei der jeweiligen Markierungsweise an.

Außerdem wurde eine farbliche Zuordnung vorgenommen, die die Werte ins Verhältnis setzt mit der Anzahl an Korrespondenzpunkten, die mittels der Markierungsvariante *nach Vorgabe* erreicht wurden. Blau hinterlegt sind alle Ergebnisse, deren Berechnung nach dem Markierungsschema der Vorgabe erfolgte. Rot hinterlegte Werte unterschreiten dieses Ergebnis. Eine gelbe Hinterlegung zeigt die Übereinstimmung dieser Werte mit dem nach anforderungsgerechter Markierung berechneten Ergebnis, und grün hinterlegte Werte überschreiten die Anzahl dieses Ergebnisses.

4 Ergebnisse

	Brauen (8)	Augen (8)	Nase (3)	Mund (12)	Kinn (7)	Summe	%
Frau 1							
anforderungsgerecht	8	8	3	12	7	38	100
groß	6	6	3	10	5	30	78,95
klein	8	6	3	12	4	33	86,84
unsymmetrisch	5	5	1	10	5	26	68,42
Frau 4							
anforderungsgerecht	8	8	3	12	7	38	100
groß	7	5	1	12	4	29	76,32
klein	7	3	3	12	7	32	84,21
unsymmetrisch	8	7	1	12	4	32	84,21
Mann 1							
anforderungsgerecht	8	8	1	5	5	27	100
groß	8	8	1	0	4	21	77,78
klein	7	8	3	11	5	34	125,93
unsymmetrisch	4	8	1	0	6	19	70,37
Mann 2							
anforderungsgerecht	5	8	1	10	7	31	100
groß	4	8	1	0	4	17	54,84
klein	4	4	2	10	6	26	83,87
unsymmetrisch	4	7	1	11	7	30	96,77
Mittelwert gesamt							82,38
Mittelwert groß							71,97
Mittelwert klein							95,21
Mittelwert unsymmetrisch							79,94

Tabelle 4.3: Versuch Markierungsvariation - Legende: Blau = Anzahl aus Versuch 1, Rot = Anzahl unterschritten, Gelb = Übereinstimmende Anzahl, Grün = Anzahl übertroffen

Die Bewertung der Versuchsreihe ergibt, dass insgesamt 55% der Fälle mittels einer vom empfohlenen Markierungsschema abweichenden Benutzerauswahl kein besseres Analyseergebnis erzielt werden konnte. In 35% der Fälle konnte mit der dabei angewandten Markierungsvariante ein übereinstimmendes Analyseergebnis erzielt werden. Nur 10% der Versuchswerte liefern ein besseres Ergebnis als das Ergebnis der Analyse mit anforderungsgerechter Merkmalsmarkierung.

Eine merkmalsbezogene Betrachtung der Werte ergibt: Bei den Augen und Augenbrauen führt keine der getesteten Markierungsvarianten zu einem besseren Ergebnis für die Berechnung der Merkmalspunkte, sondern maximal zu einer Übereinstimmung der Anzahl der richtig berechneten Korrespondenzpunkte mit den erwarteten Korrespondenzpunkten. Für die Nase und den Mund ist deutlich erkennbar, dass die Wahl eines kleinen Auswahl Fensters bei allen Fotos zu besseren oder gleichbleibenden Analyseergebnissen führt. Für die Berechnung der Korrespondenzpunkte des Kinns führte eine spitz zulaufende Wahl des Polygonzugs oder aber eine unsymmetrische

4 Ergebnisse

Form der Markierung zum Teil zu höheren oder gleichen Ergebnissen. Die breite Anordnung der Unterteilungspunkte führte jedoch in jedem Fall zu niedrigeren Werten.

Eine Bewertung der Ergebnisse pro Markierungsvariante bezogen auf die einzelnen Eingabebilder zeigt, dass mit einer kleinen Wahl des Suchbereichs tendenziell noch am ehesten gute Analyseergebnisse erzielt werden, für Mann 1 sogar deutlich bessere als mit der Markierung nach Vorgabe. Mittels des unsymmetrischen Markierungsschemas werden maximal ähnliche, aber insgesamt keine besseren Ergebnisse erzielt und die Wahl großer Suchfenster hat für keine der abgebildeten Personen positive Auswirkungen auf die Analyseergebnisse.

Insgesamt gilt, dass die Art und Weise der Markierung der Suchfenster für die Merkmale Einfluss auf die Analyseergebnisse hat. Es werden durchschnittlich mit keiner der abweichenden Auswahlmethoden mehr als 95% der Korrespondenzpunkte richtig berechnet, die mittels der Markierungsweise nach Vorgabe erzielt wurden.

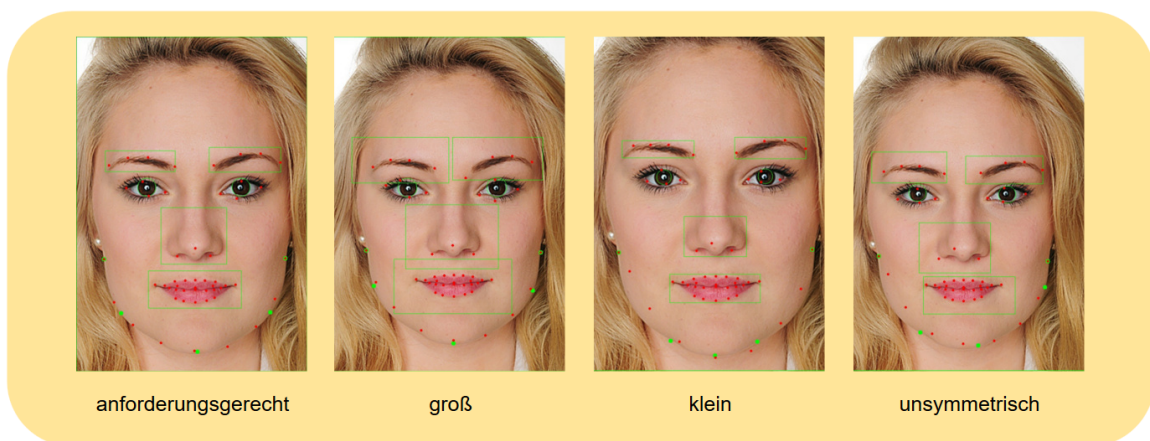


Abbildung 4.4: Beispielhaftes Ergebnis der Markierungsvariation (Frau4)

4.3 Ergebnisreflexion

Im folgenden Kapitel werden die Ergebnisse aus den durchgeführten Versuchen reflektiert. Das heißt, es werden Rückschlüsse auf das Systemverhalten gezogen und das Auftreten bestimmter Muster dabei erklärt. Das Kapitel dient der Vorbereitung auf das anschließende Fazit der Arbeit, weil darin auf Grundlage der hier aufgeführten Erkenntnisse argumentiert und geschlussfolgert wird.

Versuch zur Analysequalität:

Die Ergebnisse des ersten Versuchs zeigen deutlich, dass das Analysesystem für die meisten Fotos gute Leistungen erbringt. Für einige Eingabebilder wurde das ge-

4 Ergebnisse

wünschte Ergebnis jedoch nicht erreicht, weil ein oder mehrere Merkmale unzureichend beschrieben wurden. Die Analyse der Versuchsreihe der Kontrollgruppe liefert in den meisten Fällen sogar ein besseres Ergebnis als erwartet. Es gilt festzustellen, wie die Analysefehler für die Gruppe der als anforderungsgerecht eingestuften Fotos zustande kommen.

Mögliche Ursachen der Analysefehler bei den als anforderungsgerecht eingestuften Bildern werden nachfolgend erläutert:

Die Merkmalsanalyse-Algorithmen basieren auf der Nutzung unterschiedlicher Bildinformationen. Fehlt im Quellbild eine der grundlegenden Informationen für die erfolgreiche Lokalisierung eines oder mehrerer Merkmale oder wird die entsprechende Information durch Störeinflüsse unbrauchbar, wird die Analyse des jeweiligen Merkmals fehlschlagen. Für den Betrachter ist dies nicht immer sichtbar, weil die benötigte Bildinformation für die Analyse nicht zwangsläufig Auswirkungen auf die optische Darstellung des Bildes haben muss. Deshalb enthält die Gruppe der augenscheinlich anforderungsgerechten Bilder eben auch solche, deren nicht sichtbarer Informationsgehalt für eine erfolgreiche Analyse unzureichend ist.

Ein für den Benutzer nicht unmittelbar erkennbarer Einflussfaktor auf die Bildanalyse ist eine JPEG-Komprimierung des Analysebildes. Diese ist ein Datenreduktionsverfahren und wird deshalb nahezu immer bei Bildern durchgeführt, die im Internet veröffentlicht werden. JPEG wurde eigens dafür entwickelt, bei minimalen optischen Veränderungen eine maximale Einsparung von Bilddaten vornehmen zu können. Auf die Bildverarbeitung kann eine derart geminderte Bildqualität jedoch starken Einfluss haben, wie bereits in Unterunterabschnitt 2.2.1.2 genannt. Werden durch die Komprimierung Details geglättet, die für eine korrekte Segmentierung wichtig sind, resultiert daraus eine fehlerhafte Basis für die Konturdetektion. Da alle Bilder, die für den Test verwendet wurden, JPEG-komprimierte Bilder sind, ist ein Zusammenhang dieser Eigenschaft mit den abweichenden Positionen für einige Korrespondenzpunkte naheliegend.

Ein Großteil der Analysealgorithmen nutzt für die Konturlokalisierung die Gradienteninformation des Bildes. Auftretende Gesichtsbehaarung oder Lichtreflexe verändern die Gradienteninformation. Durch die zusätzlichen Kanten im Bild wird die Konturinformation der Merkmale gestört und kann nur schlecht davon separiert werden. Zum einen ist das der Grund dafür, dass die Analysewerte für männliche Gesichter im Bereich von Nase, Mund und Kinnlinie tendenziell niedriger sind als bei den Bildern, die Frauen abbilden. Zum anderen erklärt dies auch die teilweise falsche Detektion der Nasenmerkmalspunkte, auf die die Beleuchtungssituation den größten Einfluss hat.

Alle Lokalisierungsverfahren, die eine Interpolation der Merkmalskontur mittels der Least Squares Anpassung vornehmen, sind darauf angewiesen, dass einerseits genügend tatsächliche Konturpunkte gefunden werden und die Kontur andererseits nicht

zu stark von der Form abweicht, die durch das berechnete Polynom überhaupt darstellbar ist. Ansonsten ist der Spielraum für die Berechnung eines passenden Ausgleichspolynoms zu groß, oder das Ausgleichspolynom bildet die Form der Kontur trotz ausreichender Konturpunkte nur mangelhaft nach. Im schlimmsten Fall wird eine Konturkurve interpoliert, die zwar der Least Squares Bedingung genügt, auf der aber tatsächlich keiner der gefundenen Konturpunkte liegt. Dadurch entstehen für Merkmale, für die eine komplexe Form angenähert werden muss, häufig starke Abweichungen für die Lage der Korrespondenzpunkte, obwohl die gefundenen Konturpunkte richtig detektiert wurden. Die Versuchswerte ergeben, dass insbesondere für die Kinn und Mundkontur, aber auch teilweise für die Augenbrauen eine derartige Fehlanalyse stattgefunden hat.

Versuch zur Markierungsvariation:

Die Versuchsergebnisse des zweiten Versuchs bestätigen, dass die Markierungsmethode, die den Vorgaben entspricht, einen positiven Einfluss auf die anschließende Analyse hat und in vielen Fällen eine erfolgreiche Berechnung der Korrespondenzpunkte begünstigt. In den meisten Fällen führt eine davon abweichende Vormarkierung entweder zu niedrigeren oder zu übereinstimmenden Ergebnissen. Dies zeigt die Anfälligkeit des Systems auf benutzerspezifische Veränderungen der Markierungsmethode und ist negativ zu bewerten. Allerdings wurde der Rahmen der Abweichung für den Versuch sehr extrem gewählt, sodass eine derartige Auswahl im Realfall unwahrscheinlich ist. Diese Tatsache wertet das Gesamtergebnis des Versuchs auf.

In wenigen Fällen wurden mittels der kleinen oder unsymmetrischen Vormarkierung bessere Analyseergebnisse erzielt. Die Begründung dafür sind bildspezifische Störfaktoren, die durch ein kleines oder oder unsymmetrisches Auswahlfenster ausgespart oder ausgeglichen werden können. Gemeint ist zum Beispiel Bartwuchs, der den Analyseprozess stört und dessen Einfluss durch gezieltes Platzieren des Suchfensters vermindert werden kann. Gleiches gilt auch für Lichtreflexe innerhalb der Augen oder auf der Nase.

Ein anderes Beispiel ist eine geringfügig gedrehte Pose des abgebildeten Gesichts, die Einfluss auf die Symmetrieeigenschaften der Merkmale haben kann. In diesem Fall führt eine unsymmetrische Vormarkierung teilweise zu besseren Analyseigenschaften, weil sie die Drehung einbezieht. Insbesondere die Lokalisierung der Nase und des Kinns werden damit begünstigt.

Auf die Lokalisierung der Nasenlöcher hat vor allem die Größe des Suchfensters Einfluss, weil davon die Größe des Templates abhängt, mit dem die Nasenlöcher gesucht werden.

Zusammengefasst wird geschlussfolgert, dass die Merkmalsmarkierung nach Vorgabe zwar erfolversprechend ist, eine merkmalsbezogene, gezielte Abweichung davon jedoch unter Umständen die Analyse begünstigt. Deshalb kann das Ergebnis durch mehrmaliges Korrigieren der Vorauswahl verbessert werden.

5 Fazit

In diesem Kapitel wird die Fragestellung dieser Arbeit mittels der durch den Versuch gewonnenen Erkenntnisse beantwortet. Dafür wird die Erfolgsrate der Gesichtsanalyse des Systems betrachtet und mit dem Benutzeraufwand ins Verhältnis gesetzt. Es erfolgt die kritische Auseinandersetzung mit den Versuchsergebnissen. Weiterhin werden zum Schluss Vorschläge für Verbesserungen und Erweiterungen des Systems aufgeführt und diskutiert.

5.1 Erfolgsrate

Der Erfolg der Gesichtsanalyse nach dem hier entwickelten Verfahren wird mittels der Versuchsergebnisse des ersten Versuchs für die Gruppe der anforderungsgerechten Fotos bemessen. Eine Bemessung anhand der Versuchsreihe mit der Kontrollgruppe oder dem Versuch zur Variation der Benutzermarkierung ist aus den folgenden Gründen nicht sinnvoll:

1. Die Kontrollgruppe besteht aus Bildmaterial, für das keine erfolgreiche Analyse angenommen wird, weil die Fotos nicht den Vorgaben für geeignetes Eingabematerial entsprechen.
2. Der Versuch der Variation der Benutzerauswahl gibt Aufschluss über die Abhängigkeit des Systems von benutzerspezifischen Auswahlmöglichkeiten bei Eingabematerial, das innerhalb des ersten Versuchs schon analysiert wurde.

Außerdem wird die Analyse nur als erfolgreich eingestuft, wenn sie für das gesamte Gesicht gute Ergebnisse erzielt, weil nur dann eine Verwertbarkeit der Ausgabedaten gewährleistet ist. Hierzu wird der Gesamteindruck der berechneten Korrespondenzpunkte betrachtet. Das bedeutet, ein Analyseergebnis wird auch als Erfolg eingestuft, wenn bei der berechneten Position vieler Korrespondenzpunkte zwar von ihrer exakten Position abweicht, aber durch die Lage aller Punkte zueinander die Gesichtsform trotzdem gut annähert wird.

Anhand dieser Bewertung der Versuchsergebnisse wird die Erfolgsrate des Analyseverfahrens wie folgt berechnet:

$$Erfolgsrate = \frac{\text{erfolgreiche Analysefälle}}{\text{Gesamtanzahl der Bilder}} * 100 = \frac{7}{12} * 100 = 58\%$$

Die Erfolgsrate der Gesichtsanalyse für den Versuch liegt bei 58%. Das heißt, für mehr als die Hälfte der Eingabebilder konnten die benötigten Korrespondenzpunkte so genau berechnet werden, dass die Daten als Grundlage für die Modifikation des 3D-Modells brauchbar sind. Allerdings reicht diese Rate nicht aus, um die Analyseleistung des Systems als zuverlässig bewerten zu können.

Diese Rate wurde erreicht, indem die Vormarkierung der Gesichtsmerkmalsbereiche innerhalb eines Durchlaufs erfolgte. Das heißt, es erfolgte nach Überprüfung der Analyseergebnisse keine nachfolgende Korrektur für einen weiteren Systemdurchlauf. Es wurde seitens des Benutzers also minimaler Aufwand betrieben. Da für einige Bilder eine spezifisch auf den Bildinhalt angepasste Markierung für die Analyse vorteilhaft sein kann, könnte die Erfolgsrate durch mehrfaches Korrigieren der Vormarkierung gegebenenfalls angehoben werden. Damit verbunden ist jedoch ein höherer Benutzeraufwand.

Die Fragestellung der Arbeit (Abschnitt 1.1) kann mittels dieser Erkenntnisse wie folgt beantwortet werden:

Der *erste Teil* der Frage bezieht sich auf das Ergebnis der Analyse. Es wird gefragt, ob mit dem Analyseverfahren gleichwertige Ergebnisse erzielt werden können wie mittels einer ausschließlich manuellen Datenerfassung. Gleichwertig bedeutet in diesem Zusammenhang, genauso umfangreich und genau, wie es per manueller Eingabe der Fall wäre. Für die Beantwortung der Frage wird die Erfolgsrate herangezogen. Diese liegt bei 58% ohne Korrekturdurchläufe. Die manuelle Datenerfassung weist eine hundertprozentige Erfolgsrate auf, was dadurch begründet ist, dass die Analyseleistung des menschlichen Gehirns den Maßstab für die Bewertung der Gesichtsanalyse darstellt. Somit wird die Abweichung der Systemrate als zu hoch eingeschätzt, sodass das Verfahren in dieser Form keinen gleichwertigen Ersatz für die manuelle Datenerfassung darstellt.

Der *zweite Teil* der Fragestellung bezieht sich auf den mit dem Verfahren verbundenen Benutzeraufwand. Es wird gefragt, ob das Verfahren in seiner Art und Weise den Benutzeraufwand gegenüber der ausschließlich manuellen Datenerfassung reduziert. Dies trifft zumindest für die Analyse, die ohne Korrekturdurchläufe vorgenommen wird/auskommt, zu.

Für die Beantwortung der Fragestellung als Ganzes müssen beide oben betrachteten Aussagen gegeneinander abgewogen werden. Die Erfolgsrate der Analyse ohne Korrekturdurchläufe ist gegenüber der manuellen Datenerfassung unzureichend. Allerdings gibt es zwei Wege, wie die Analyseergebnisse verbessert werden können:

1. Die optimal an die Systemanforderungen angepasste Erstellung des Eingabematerials
2. Eine ein- oder mehrfach vorgenommene Korrektur der manuell erfassten Bereiche bei wiederholten Analysedurchläufen

Beide Möglichkeiten für die Aufwertung der Analyseergebnisse erfordern einen höheren Benutzeraufwand. Entweder findet eine Verlagerung des Aufwands hin zur aufwendigeren Erzeugung/Auswahl des Eingabematerials oder aber zur aufwendigeren Bedienung des Programms hin statt.

Aus diesen Gründen treffen die in der Frage genannten Bedingungen auf die hier entwickelte Form des Verfahrens zur Gesichtsanalyse nicht zu, weil das hier betrachtete (entwickelte) Verfahren zur teilautomatisierten, zielgerichtete Analyse eines geeigneten Frontalbildes eines Gesichts eine ausschließlich manuelle Datenerfassung nicht gleichwertig ersetzt, in einer Art und Weise, die den Benutzeraufwand reduziert.

5.2 Verbesserungsvorschläge

Aus dem vorangegangenen Kapitel geht hervor, dass das in dieser Arbeit entwickelte Verfahren zur Gesichtsanalyse, die gestellten Bedingungen bezüglich der Datenqualität und des Benutzeraufwands nicht erfüllt, zumindest nicht mit der derzeitig vorliegenden Form des Analysesystems. Allerdings ist eine stabile und überwiegend erfolgreiche Analysefunktion gegeben. Das heißt, es besteht die Möglichkeit, das Analyseprogramm dahingehend abzuwandeln/zu erweitern, dass es die Kriterien für die Erreichung der Zielsetzung erfüllt. Vorschläge und Ideen für derartige Verbesserungen und Erweiterungen werden im folgendem Text aufgeführt und erläutert.

1. Mimik und Poseninvarianz:

Das System ist darauf ausgelegt, Frontalansichten mimikneutraler Gesichter zu analysieren. Die Konturlokalisierung der Gesichtsmerkmale funktioniert zwar mit Einschränkungen auch für Fotos von Gesichtern mit abweichender Mimik oder Pose. Allerdings kann für derartige Daten keine Berechnung der Korrespondenzpunkte vorgenommen werden, weil deren Position anhand der Verhältnisse in der Frontalansicht des generischen Modells ermittelt werden. Aufgrund dieser Abhängigkeit wird für die Berechnung von falschen Annahmen ausgegangen und die Modifikation anhand derartig gewonnener Daten muss fehlschlagen. Der Benutzer muss jedoch weniger Aufwand betreiben, wenn es weniger Einschränkungen bei der Auswahl geeigneten Eingabematerials gibt. Deshalb ist eine Erweiterung der Programmfunktion für posen- und mimikvariantes Eingabematerial sinnvoll.

2. Manuelle Korrektur einzelner Korrespondenzpunkte:

Im ersten Versuch zur Bewertung der Analyseleistung konnte festgestellt werden, dass die überwiegende Anzahl der Korrespondenzpunkte zwar richtig berechnet wurde, aber in vielen Fällen keine erfolgreiche Gesamtanalyse erzielt werden konnte, wofür die wenigen falsch berechneten Korrespondenzpunkte ausschlaggebend waren. Der Benutzeraufwand für die Einzelkorrektur der wenigen falsch positionierten Punkte ist nur gering, aber das Analyseergebnis würde dadurch perfektioniert. Eine derartige

Einflussmöglichkeit für die Korrektur der automatisch berechneten Analyseergebnisse verbessert die Grundlage für die anschließende Modifikation erheblich. Außerdem liegt es im Ermessen des Benutzers, ob der Aufwand dafür erforderlich ist und wie viel Aufwand er dafür betreiben möchte. Dadurch wird die Benutzerfreundlichkeit erhöht.

3. Kubisches Spline für Mundkontur:

Die Mundkontur wird in der aktuellen Version des Analysesystems per Least Squares Anpassung interpoliert. Die Oberlippenkontur wird dabei durch ein Polynom vierten Grades angenähert. In den Tests ist aufgefallen, dass diese Annäherung oft nicht ausreicht, um den Amorbogen, also die kleine Vertiefung in der Mitte der Oberlippenkante nachzubilden. Hähnel verwendet für die Mundkontur deshalb ein kubisches Spline, also eine Kurve, die sich aus drei Teilkurven zusammensetzt. Dadurch besteht die Oberlippenkontur aus zwei Teilstücken, wodurch der Amorbogen besser nachgebildet wird. Diese Verbesserung würde die Analysegenauigkeit erhöhen.

4. Funktionen von OpenCV ersetzen

Für die Implementierung gängiger Bildverarbeitungsalgorithmen, die für typische Problemstellungen der Bildanalyse Anwendung finden, wurden bei der Entwicklung des Gesichtsanalysesystems die Programmbibliothek OpenCV genutzt. Das heißt, die entsprechenden Funktionen wurden nicht selbst programmiert, sondern auf vorhandene Funktionen aus der Bibliothek zurückgegriffen. Dadurch fehlte jedoch teilweise die Möglichkeit, Einfluss zu nehmen. Für die spezifische Anpassung an die Bedingungen, unter denen die Algorithmen innerhalb der Analyse Anwendung finden, fehlte teilweise der Einfluss auf die vorhandene Form der Funktionen. Zum einen fehlte zur sinnvollen Anwendung der bestehenden Funktionen in einigen Fällen die genaue Kenntnis über die darin erfolgende Verarbeitung der Daten. Zum anderen war teilweise eine andere Art der Initialisierung erforderlich als die, die per vorhandenem Konstruktor vorgegeben wurde. Deshalb kann das Analysesystem verbessert werden, indem entweder eine Anpassung oder Einsicht in die verwendeten Funktionen vorgenommen wird oder die entsprechenden Funktionen durch eine selbst entwickelte Variante ersetzt werden.

5. Zusätzliche Analyse einer Profilsicht

Die exakte Nachbildung eines Gesichts in Form eines 3D-Modells erfordert zusätzlich zu den Daten über die Maße und Formeigenschaften der Frontalansicht des Gesichts ebenfalls Informationen über die Tiefenmaße, also die dritte Dimension. Diese Daten können aus der Frontalansicht nicht gewonnen werden, sondern erfordern die Analyse von ein oder mehr Fotos, die das Gesicht aus einer anderen Perspektive abbilden. Diese Erweiterung des Programms unterstützt einerseits den Erfolg und die Genauigkeit der Gesichtsanalyse, weil insgesamt mehr Informationen über das Gesicht vorliegen. Andererseits wird dadurch die Qualität des 3D-Modells erheblich verbessert, weil die Daten, die für die Modifikation des generischen Modells verwendet werden, das Gesicht umfangreicher beschreiben.

6. Erweiterung der Analysefunktion in Bezug auf Anomalien

Zurzeit ist das Analysesystem unfähig, Anomalien in Gesichtern zu erkennen und/oder zu analysieren. Die Funktion der Analyse ist beschränkt auf Gesichter, die den Eigenschaften entsprechen, die im Kapitel zu den Anforderungen an das Analysesystem genau festgelegt werden. Das bedeutet, dass nur stereotypische Gesichter durch das System untersucht und beschrieben werden können. Eine sinnvolle Erweiterung des Programms besteht darin, die Analysefunktion auch für Gesichter, die vom Stereotyp abweichen, zu ermöglichen.

5.3 Schlussbemerkung

Der Wunsch des Menschen danach, diverse Aufgaben und Arbeiten durch Computertechnologie vereinfachen und ersetzen zu können, treibt die Forschung und Entwicklung in diesem Gebiet seit Jahrzehnten voran. Die Frage, die sich stellt, ist also niemals, ob eine Lösung für die jeweilige Problemstellung gefunden und entwickelt werden kann, sondern viel mehr wann die Lösung auf welchem Weg gefunden wird. Jede Idee und jeder Ansatz wird so lange verfolgt, erweitert und perfektioniert, bis aus der einmal befremdlichen Vorstellung, die jeweilige Technologie jemals nutzen zu können, längst ein alltäglich befriedigtes Bedürfnis geworden ist.

Dies gilt auch für den Bereich der Gesichtserkennung und -analyse. Wie bereits am Anfang der Arbeit erwähnt und festgestellt wurde, ist diese Aufgabe kein triviales Problem, und den Beweis dafür liefert diese Arbeit, weil darin die entsprechenden Tücken und Schwierigkeiten dazu geführt haben, dass letztendlich noch keine zufriedenstellende Analysequalität erreicht werden konnte. Die Ansätze sind gut, aber noch nicht ausgereift, und es konnten viele Erfolge erzielt, aber nicht alle Einzelheiten bedacht und behandelt werden.

Sehr wahrscheinlich kann das hier entwickelte Verfahren zur Datenerfassung mittels Gesichtsbildanalyse noch soweit verbessert und weiterentwickelt werden, dass es zuverlässig und effektiv arbeitet und qualitativ gute Ergebnisse liefert. Vielleicht ist es aber auch gut, in einer Welt, die sich immer abhängiger von verschiedenen Technologien macht, immer wieder daran erinnert zu werden, wie leicht dem Menschen komplexe alltägliche Aufgaben wie die Gesichtserkennung und -analyse fallen und wie kompliziert es ist, eine Maschine diese Aufgaben lösen zu lassen und, dass diese Imitation wohl niemals an das Original heranreichen wird.

Abbildungsverzeichnis

2.1	Bildverarbeitungskette	6
2.2	Verschiedene Gesichtsformen	10
2.3	Unterschiede in Gesichtern von Männern und Frauen	10
2.4	Goldener Schnitt und Gesichtsverhältnisse	10
2.5	Einflussfaktoren auf die Gesichtsanalyse	12
3.1	Auswahlbereiche des Benutzers für die lokale Merkmalsuche	28
3.2	Augenlokalisierung	30
3.3	Iris-Suchmaske	30
3.4	Luminanzkanal als Höhenfeld interpretiert	31
3.5	Lokale Suche der Augenkontur	32
3.6	Augenbrauenlokalisierung	34
3.7	Bestimmung der Augenbrauenkontur	35
3.8	Nasenlokalisierung	36
3.9	Template Matching	37
3.10	Mundlokalisierung	38
3.11	Kinnlokalisierung	40
3.12	Iterationsweise Anpassung der Initialkurve an die Kinnkontur	41
3.13	Korrespondenzpunkte	43
3.14	Gesichtsverhältnisvektor	44
4.1	Verschiedene Formen der Markierung	48
4.2	Ergebnisbeispiele Versuch Analysequalität	50
4.3	Ergebnisbeispiele Kontrollgruppe	52
4.4	Ergebnisbeispiel Versuch Markierungsvariation	54

Tabellenverzeichnis

2.1	Automatische vs. Manuelle Datenerfassung	5
4.1	Versuch Analysequalität	49
4.2	Versuch Analysequalität mit Kontrollgruppe	51
4.3	Versuch Markierungsvariation	53

Literaturverzeichnis

- Aurich, Volker: *Bildverarbeitung*, http://www.acs.uni-duesseldorf.de/~aurich/Skripte/Bildverarbeitung_1.pdf, 2011, Letzter Zugriff: 12. 08. 2016
- Baur, Dominikus: *Automatische Gesichtserkennung: Methoden und Anwendungen*, http://www.medien.ifi.lmu.de/fileadmin/mimuc/hs_ws0506/papers/Automatische_Gesichtserkennung.pdf, Letzter Zugriff: 12. 08. 2016
- Becker, Markus: *Gewichtsform und Attraktivität*, http://scidok.sulb.uni-saarland.de/volltexte/2003/102/pdf/markus_becker.pdf, 2003, Letzter Zugriff: 12. 08. 2016
- Bitzer, Meike: *Gesichtsformen: Die richtige Brille für jeden Mann*, <http://www.menshealth.de/artikel/die-richtige-brille-fuer-jedes-gesicht.115522.html>, 2009, Letzter Zugriff: 12. 08. 2016
- Borghoff, Höll: *Bildverarbeitung und Algorithmen - Vortrag Segmentierung*, www.gm.fh-koeln.de/~konen/WPF-BV/Segmentierung-SS05.doc, 2005, Letzter Zugriff: 12. 08. 2016
- Burmeister, Gunnar: *3D-Gesichtsmodellierung mit bilddatenbasierter Modifikation eines generischen Gesichtsmodells - Entwicklung einer Benutzeroberfläche und Modifikation des generischen Gesichtsmodells anhand bereitgestellter Daten*, 1. Aufl., BA 2016
- Enlow, Hans: *Essentials of facial growth*, Saunders 1996
- Geisendrees, Bärbel: *Methoden zur Erkennung von Hautfarbe in Farbbildern*, http://www.uni-weimar.de/medien/webis/teaching/theses/geisendrees_2001.pdf, 2001, Letzter Zugriff: 12. 08. 2016
- Grammer, Karl: *Signale der Liebe. Die biologischen Gesetze der Partnerschaft.*, Deutscher Taschenbuchverlag 2000
- Haasdonk, Dr. Bernhard: *Digitale Bildverarbeitung - Einheit 8 - Lineare Filterung*, http://lmb.informatik.uni-freiburg.de/people/haasdonk/DBV_FHO/DBV_FHO_SS07_E08_handout.pdf, 2007, Letzter Zugriff: 12. 08. 2016
- Hähnel, Michael: *Modellbasierte posen- und mimikinvariante Gesichtserkennung*, 1. Aufl., 2008

- Heine, Anne-Kathrin: *Individuelle dreidimensionale Simulation von Weichteilveränderungen im Gesichtsbereich ? eine Finite-Elemente-Studie*, https://edoc.ub.uni-muenchen.de/3662/1/Heine_Anne-Kathrin.pdf, 2005, Letzter Zugriff: 12. 08. 2016
- Karakoc, Gürcan: *Morphologische Bildoperationen*, <http://www9.in.tum.de/seminare/ps.SS06.gdbv/ausarbeitungen/psbv-ss06-morphologie-karakoc.pdf>, 2006, Letzter Zugriff: 12. 08. 2016
- Klaus, Peschke, Burkhardt: *Praktikumsversuch Bildvorverarbeitung*, http://lmb.informatik.uni-freiburg.de/lectures/praktika_brox/bvpraktikum/BVAnl_bildvor.pdf, Letzter Zugriff: 12. 08. 2016
- Kosmala, Andreas: *HMM-basierte Online Handschrifterkennung - ein integrierter Ansatz zur Text- und Formelerkennung, Kapitel 3*, https://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-5152/thesis_kosmala_ch3.pdf, 2000, Letzter Zugriff: 12. 08. 2016
- Kramer, Pricking: *Rauschunterdrückung in der Theorie & Praxis*, http://www.mathematik.uni-ulm.de/stochastik/lehre/ws03_04/seminarws03_04/kramer.pdf, 2004, Letzter Zugriff: 12. 08. 2016
- Kreutz, Verhoff: , <http://geb.uni-giessen.de/geb/volltexte/2004/1879/pdf/KreutzVerhoff-gu37-2004.pdf>, 2004, Letzter Zugriff: 12. 08. 2016
- Lordemann, Lambers: *Objekterkennung in Bilddaten*, <https://ivv5hpp.uni-muenster.de/u/lammers/EDU/ws03/Landminen/Abgaben/Gruppe9/Thema09-ObjekterkennungInBilddaten-ChristianGrosseLordemann-MartinLambers.pdf>, 2004, Letzter Zugriff: 12. 08. 2016
- Miehlke, Alexander: *Entwicklung eines Motion Tracking-Systems zur Positionssteuerung per Handbewegung*, <http://people.f4.htw-berlin.de/~tj/da/miehlke.pdf>, 2007, Letzter Zugriff: 12. 08. 2016
bibitem[OpenCV()]opencv OpenCV: *Open Source Computer Vision*, <http://opencv.org/>, Letzter Zugriff: 12. 08. 2016
- Peterwitz, Julia: *Grundlagen: Bildverarbeitung / Objekterkennung*, http://www9.in.tum.de/seminare/hs.SS06.EAMA/material/01_ausarbeitung.pdf, 2006, Letzter Zugriff: 12. 08. 2016
- PhiPoint Solutions: *The Human Face and the Golden Ratio*, <http://www.goldennumber.net/face/>, Letzter Zugriff: 11. 08. 2016
- Schöley, Paula: *Kantendetektoren*, <https://www.inf.tu-dresden.de/content/institutes/ki/is/VORTRAG/Kantendetektoren.pdf>, 2011, Letzter Zugriff: 12. 08. 2016

Stanner, Ed: *Golden ratio line percentages*, https://commons.wikimedia.org/wiki/File:Golden_ratio_line_percentages.svg, 2012, Letzter Zugriff: 12. 08. 2016

Syllwasschy, Robin: *Erstellung eines generischen Kopfmodells für die computer-gestützte Simulation und Animation menschlicher Gesichter*, http://www.cgvtugraz.at/PubArc/dipl/syllwasschy97_dipl.pdf, 1997, Letzter Zugriff: 12. 08. 2016

Wikipedia: <https://de.wikipedia.org/wiki/Wikipedia:Hauptseite>, Letzter Aufruf: 12. 08. 2016

Zhao, Yuanyuan: *Modellbasierte Segmentierung*, <http://www9.in.tum.de/seminare/ps.WS05.gdbv/ausarbeitungen/ps-gdbv-ws05-07.pdf>, 2005, Letzter Zugriff: 12. 08. 2016

Zipser, Bartusch, Steiner: *Komplexpraktikum Automatisierungstechnik - Versuch : Bildverarbeitung*, <https://www.htw-dresden.de/fileadmin/userfiles/et/Labore/Prozessmesstechnik/Bildverarbeitung.pdf>, 2013, Letzter Zugriff: 12. 08. 2016

Für den Versuch verwendete Bilder:

Barbie: <http://www.salubia.de/bilder/puppen/bianca01g.jpg>, Letzter Zugriff: 19. 07. 2016

Frau1: <https://www.whitezine.com/fr/photography/portraits-by-martin-schoeller.html/attachment/martin-schoeller-angelina-jolie-portrait>, Letzter Zugriff: 14. 07. 2016

Frau2_1: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016

Frau2_2: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016

Frau3: <http://www.polodigital.net/images/ErikaH/ErikaH010f.jpg>, Letzter Zugriff: 19. 07. 2016

Frau4: <http://www.fotostyle-in.de/file/2015/02/passfoto-frau.jpg>, Letzter Zugriff: 19. 07. 2016

Jungel1: <http://www.fotostyle-in.de>, Letzter Zugriff: 19. 07. 2016

Literaturverzeichnis

- Mädchen1: <http://www.fotostudio-mohl.de/blog/passfoto/passfoto-passbilder-bei-fotostudio-mohl-in-reutlingen-tel-07121-55294.html>,
Letzter Zugriff: 19. 07. 2016
- Mann1: <https://inst.eecs.berkeley.edu/~cs194-26/fa14/upload/files/proj5/cs194-db/>, Letzter Zugriff: 14. 07. 2016
- Mann2: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016
- Mann3: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016
- Mann4: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016
- Frau gedreht: <http://www.wallpapersimages.co.uk/master/NEW/1024x768/6883940-portrait.jpg>, Letzter Zugriff: 19. 07. 2016
- Frau Schmuck: http://bilder.desired.de/b/75/14/53/56/id_75145356/610/tid_da/bindis.jpg, Letzter Zugriff: 19. 07. 2016
- Frau Verdeckung: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016
- Junge Aufl Licht: <https://www.fotolia.com/p/201402472>, Letzter Zugriff: 14. 07. 2016
- Junge Brille: <http://jugendgemeinderat-tuttlingen.de/images/1256068628Passbild.JPG>, Letzter Zugriff: 19. 07. 2016
- Mädchen Augenbrauen: <http://bildagentur.panthermedia.net/m/lizenzfreie-bilder/12994826/makro-grossaufnahme-macro-makroaufnahme-macroaufnahme-nahaufnahme>, Letzter Zugriff: 14. 07. 2016
- Mann2 Mimik: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016
- Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016
- Mann5 Bart: *Caltech Human face (Front) dataset* <http://www.robots.ox.ac.uk/~vgg/data3.html>, Letzter Zugriff: 14. 07. 2016
- Mann Highkey: <http://www.readtheface.com/wp-content/uploads/2010/12/oldmancfaceread.jpg>, Letzter Zugriff: 19. 07. 2016