



Universität Paderborn
Fachbereich 17 - Mathematik und Informatik
Arbeitsgruppe Wissensbasierte Systeme

Diplomarbeit
für den Studiengang Informatik

Methoden zur Erkennung von Hautfarbe in Farbbildern

BÄRBEL GEISENDREES
SEPTEMBER 2001

Eidesstattliche Erklärung

Hiermit versichere ich, daß ich die vorliegende Diplomarbeit selbstständig verfaßt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie Zitate kenntlich gemacht habe.

Paderborn, 11. Oktober 2001

Bärbel Geisendrees

Inhaltsverzeichnis

1	Einleitung	1
1.1	Allgemeines	1
1.1.1	Gesetzliche Grundlagen	2
1.2	Ablauf der Bilderkennung	3
1.3	Beschreibung der Aufgabe	4
1.4	Aufbau der Arbeit	5
2	Hautfarbe und Farbmodelle	6
2.1	Problem	6
2.2	Das RGB-Farbmodell	6
2.3	Das HSV-Farbmodell	7
3	Methoden zur Erkennung von Hautfarbe	9
3.1	Farbselektion	9
3.1.1	Farbselektion aus 16 Millionen Farben	10
3.1.2	Farbselektion aus 256 Farben	11
3.2	Lineare Klassifikation	12
3.2.1	Adaption der linearen Klassifikation	12
3.3	Regression	15
3.4	Neuronale Netze	15
3.5	Chromatischer Farbraum	16
3.6	Die Transformation in den HSV-Farbraum	19

4	Umsetzung	21
4.1	Selektion mit Gimp	21
4.2	Lineare Klassifikation in C++.	22
4.2.1	Hautfarbendefinition für die folgenden Methoden	23
4.2.2	Lineare Klassifikation	23
4.2.3	Octree Datenstruktur	23
4.3	Chromatischer Farbraum mit matlab	24
4.4	HSV in C++	25
4.5	Hilfsmittel für die Auswertung	26
4.5.1	Konverter für Bitmap-Dateien	26
5	Auswertung	29
5.1	Selektion	30
5.1.1	11 aus 256	30
5.1.2	8 aus 16 Millionen	31
5.2	Lineare Klassifikation	33
5.2.1	Schwellwertsuche	33
5.2.2	Ergebnisse	34
5.3	Regression und Neuronale Netze	37
5.4	Chromatischer Farbraum	37
5.5	Kopplung von zwei Verfahren	38
5.6	Das HSV-Farbsystem	39
5.7	Diskussion	40
5.8	Zusammenfassung und Ausblick	43
A	Diagramme	45
B	Ergebnisse zu einzelnen Auswertungen	49

Kapitel 1

Einleitung

1.1 Allgemeines

Das Internet ist ein sehr nützliches und hilfreiches Instrument, um sich in vielen Bereichen schnell und umfassend zu informieren. Aufgrund der leichten Bedienung ist es für Kinder selbstverständlich mit diesem Medium aufzuwachsen. Die Kehrseite ist aber die, daß man sich vor ungewünschten Inhalten nicht schützen kann.

Es gibt natürlich Möglichkeiten, bestimmte Seiten auszuschließen, das setzt aber voraus, das man diese Seiten kennt. Eine Methode wäre ein Programm, welches den Besuch einer Anzahl von angegebenen Internetseiten einfach unterbindet. Dazu muß das Programm erstens aber schon einen Grundstock von solchen Seiten enthalten und zweitens muß es ständig aktualisiert werden. Man kann nämlich zu keinem Zeitpunkt sagen, daß jetzt alle Seiten aufgenommen sind, da selbst bei guter Recherche immer unentdeckte Seiten übrig bleiben und außerdem ständig neue Seiten hinzukommen. Eine solche Methode kann aber immer nur im Nachhinein agieren, da die Seiten erst existieren müssen, um gefunden zu werden. Besser wäre daher ein Verfahren, das bei angeforderten Seiten immer aktuell eine Überprüfung dahingehend vornehmen kann, ob die Seite inhaltlich Gefährdungen birgt.

Speziell gewaltverherrlichende, rechtsradikale oder pornografische Inhalte stören die meisten Benutzer sehr. Da solche Seiten oft Internetadressen benutzen, die bekannten Adressen sehr ähnlich sind, wird man ohne eigenes Zutun auf diese Seiten verwiesen. Die Besitzer solcher Seiten sorgen durch geschicktes Plazieren von Suchwörtern dafür, daß ein Nutzer bei Eingabe von viel gesuchten Begriffen von einer Suchmaschine aus völlig zufällig auf ihrer Seite landet. Das führt dazu, daß Kinder eigentlich nicht unbeaufsichtigt das Angebot des Internets nutzen sollten, um solche Inhalte zu umgehen. Es wäre sehr wünschenswert, daß Kinder das Internet zu Informationszwecken nutzen können, ohne dieser Gefahr ausgesetzt zu sein.

Die Herausforderung für Entwickler liegt nun darin, Inhalte einer Internetseite schon vor dem Anschauen zu analysieren. Für Textseiten ist das in gewissem Umfang auch einigermaßen schnell möglich. Probleme bereitet aber das Erkennen von Bildern im Vorhinein. Den Inhalt eines Bildes vorauszusagen, ohne es gesehen zu haben, entpuppt sich als sehr komplizierter Prozeß.

1.1.1 Gesetzliche Grundlagen

Die Veröffentlichung von pornografischen Bildern im Internet wird immer mehr zum Problem. Zum einen gibt es keine tatsächliche rechtliche Handhabe gegen eine Veröffentlichung der Bilder. Zum anderen ist es für Eltern sehr schwierig, ihre Kinder im Internet vor solchen Inhalten zu schützen.

Das Jugendschutzgesetz [JÖSchG] weist in §7 darauf hin, daß

”Bildträger von der obersten Landesbehörde zu kennzeichnen sind, ob sie unter 18 Jahren freigegeben sind”.

Wenn Daten im Internet öffentlich zugänglich gemacht werden, muß man diesen Weg der Einstufung durch Dritte nicht gehen. Einem Benutzer jeden Alters ist es durch einen Internetanschluß möglich, alle Inhalte des Internets zu nutzen. Laut Jugendschutzgesetz [JÖSchG]

”handelt man ordnungswidrig, wenn man vorsätzlich oder fahrlässig nicht freigegebene Bildträger anbietet oder überläßt”.

D.h., jegliche Personen, die im Internet pornografisches Bildmaterial veröffentlichen, handeln im Sinne des Jugendschutzgesetzes lediglich ordnungswidrig. Sie müssen sich durch das Nichtvorhandensein von gesetzlichen Regelungen im Internet nicht strafrechtlich verantworten. Die Bereitstellung von pornografischem Material ist im gesetzessfreien Raum des Internets also ohne wesentliche Konsequenzen möglich.

In den USA ist 1998 ein Gesetz zum ”Schutz von Minderjährigen vor Internet-Pornografie” erlassen worden. Dieses umstrittene Gesetz sieht vor, daß Betreiber von Sex-Webseiten elektronische Systeme installieren müssen, um das Alter der Nutzer festzustellen. Bei erstmaligem Verstoß gegen das Gesetz sind bis zu sechs Monate Haft und eine Geldstrafe von 50.000 US-Dollar fällig. Die US-Bürgerrechtsorganisation American Civil Liberties Union (ACLU) hatte gegen das Gesetz geklagt. Die ACLU argumentierte, daß auch medizinische Beratungstellen, Diskussionsforen rund um Sexualkrankheiten, Online-Buchshops, Seiten mit homosexuellen Inhalten oder Kunstgalerien von dem Gesetz betroffen seien [USG].

Man sieht an den unterschiedlichen Handhabungen der Länder, daß es schwierig bis unmöglich ist, einheitliche Regelungen zu finden. Schon jedes einzelne Land hat genügend Probleme, überhaupt Gesetze zu definieren.

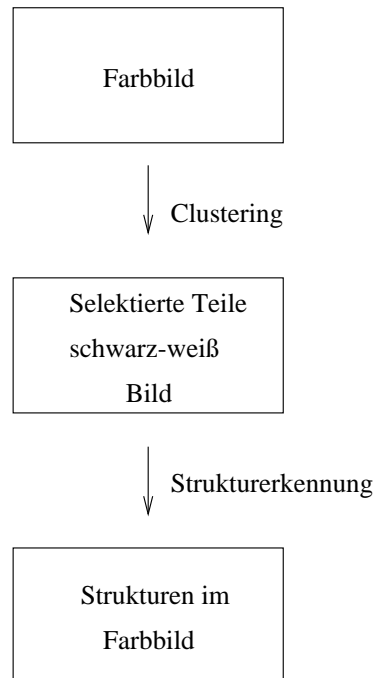


Abbildung 1.1: Ablauf der Bilderkennung

1.2 Ablauf der Bilderkennung

Den Ablauf des Erkennens von Bildern mit pornografischem Inhalt kann man wie in Bild 1.1 beschreiben. Zunächst werden die Bilder mit einer Haut-Erkennung vorsortiert. Dabei wird versucht, Hautfarbe in Bildern zu erkennen und zu selektieren und diese nach bestimmten Kriterien, wie Hautanteil im Bild im Verhältnis zum Gesamtbild, weiterzureichen. Ist das Kriterium erfüllt, wird das Bild weiter untersucht. Es findet eine sogenannte Strukturerkennung statt, die in dem Bereich, der als Haut erkannt wurde, weitere Prüfungen durchführt. In diesem Teil wird versucht eine bestimmte Struktur oder einen Menschen zu erkennen. Die Strukturerkennung ist ein sehr aufwendiges und kompliziertes Gebiet. Will man nur einfache Strukturen erkennen, ist der Aufwand noch abzuschätzen. Bei der Erkennung von Menschen allerdings ist es nötig, die einzelnen Glieder des Menschen zusammenzusetzen. Diese zunächst zu finden und dann in körpergerechter Weise zusammenzusetzen erfordert einen sehr hohen Aufwand.

Erst nach der Strukturerkennung kann man mit großer Sicherheit sagen, daß es sich im Bild um die Struktur handelt, die gesucht worden ist, hier die eines Menschen.

Meine Arbeit beschäftigt sich mit der Suche von Haut in Bildern. Es gibt unterschiedlichste Methoden, Hautfarbe in Bildern zu identifizieren. Ich möchte diese

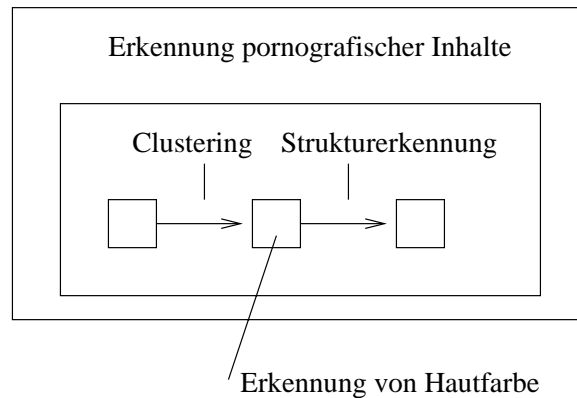


Abbildung 1.2: Einordnung der Arbeit in den Problemzusammenhang

verschiedenen Möglichkeiten erläutern und vergleichen. Sie sollen auf unterschiedliche Kriterien geprüft werden. An erster Stelle steht dabei natürlich die Güte der Erkennung von Haut. Nachdem man mit den Ergebnissen meiner Arbeit eine gute Lösung zur Erkennung von Haut gefunden hat, könnte daran anschließend eine Strukturerkennung im Bild durchgeführt werden.

1.3 Beschreibung der Aufgabe

Bei Tripod, dem Homepage-Building-Bereich der Lycos Europe GmbH, liegt das Problem an ähnlicher Stelle. Dort können sich Mitglieder eine eigene Homepage zusammenstellen und veröffentlichen. Diese Seiten sollen "sauber" sein, d.h. sie sollen unter anderem keine pornografischen Inhalte enthalten. Um diesem Anspruch gerecht zu werden, müßten die verantwortlichen Personen ständig die gesamte Datenmenge, die von Mitgliedern auf den Servern abgelegt wird, kontrollieren. Das ist bei der enormen Anzahl nicht mehr möglich. Außerdem ist ein Mitglied jederzeit in der Lage seine Daten zu ändern. Diese Problematik gilt natürlich nicht nur für pornografische, sondern ebenso für rechtsgerichtete oder gewaltverherrlichende Inhalte. Während man jedoch bei rechtsradikalen Inhalten noch das Instrument der Textsuche einsetzen kann, gestaltet sich das Finden pornografischer Seiten wesentlich schwieriger. Aus Erfahrung kann man davon ausgehen, daß die meisten Seiten mit pornografischem Inhalt kaum oder keinen Text enthalten. Die Seiten bestehen häufig nur aus einer Anzahl von Bildern oder Bildgalerien. Das Durchsuchen sämtlicher Bilder bereitet viel Arbeit. Um diesen Ablauf zu erleichtern, ist das Ziel meiner Arbeit, Bilder möglichst gut nach pornografischem oder nicht pornografischem Inhalt sortieren zu können.

Die Problematik des kontextbasierten Erkennens von Bildern mit pornografischem Inhalt liegt darin, daß es nur sehr wenige Anhaltspunkte gibt, um diese Bildin-

halte zu spezifizieren. Der Grundgedanke bei der Analyse von Bildinhalten ist die Hautfarbe. Auf Fotos und Bildern ist im allgemeinen wenig Hautfarbe zu finden, da Menschen kaum hautfarbene Kleidung tragen; andererseits sind im Hintergrund oft Holz oder Sand zu erkennen, die leicht als Hautfarbe identifiziert werden.

Um Hautfarbe in einem Bild zu erkennen, ist es zunächst einmal nötig Hautfarbe zu definieren. Dann gibt es unterschiedliche Methoden, um die Farben aus den Bildern herauszulösen. Ich werde verschiedene Methoden vorstellen, die unterschiedliche, farbige Bilder in schwarz-weiß Bilder umsetzen, wobei die Hautfarben dann in weißer und alle anderen Farben in schwarzer Farbe dargestellt werden. Im Rahmen dieser Arbeit soll deutlich werden, welches die besten Methoden sind, um diese Transformation durchzuführen.

1.4 Aufbau der Arbeit

Zunächst wird in Kapitel 2 das Problem Hautfarbe erläutert. Es wird erklärt, warum es sich so schwierig gestaltet, Hautfarbe zu definieren. Außerdem wird der Zusammenhang zwischen Bildern und Farbmodellen aufgezeigt und es wird dargelegt, aus welchem Grund zunächst das RGB-System und später auch das HSV-Farbsystem gewählt wurde. In diesem Rahmen werden beide Farbsysteme erklärt.

In Kapitel 3 werden verschiedene Methoden vorgestellt, mit denen Hautfarbe erkannt werden soll. Zunächst wird dabei die Selektion aus dem RGB-Farbraum vorgestellt, danach die sogenannte Regression und Neuronale Netze und schließlich werden Transformationen des Farbraums behandelt. Daran anschließend findet in Kapitel 4 die Umsetzung dieser verschiedenen Verfahren statt, in der der Verlauf der Implementierung der Methoden beschrieben ist. Nachfolgend werden einige Hilfsmittel erläutert, die für die Umsetzung wichtig sind. Zum Schluß folgt die Auswertung in Kapitel 5, in der die Ergebnisse dargestellt und diskutiert werden.

Kapitel 2

Hautfarbe und Farbmodelle

2.1 Problem

Es ist schwierig, herauszufinden, welche Farben Hautfarbe tatsächlich ausmachen. Die Farbwerte für Hautfarbe befinden sich nur in einem relativ kleinen Bereich des Farbraumes. Bei der Darstellung dieses Bereiches muß man sich jedoch mit einer Approximation zufrieden geben, da Hautfarbe sich als eine Punktwolke im Farbwürfel darstellt. Wie die Lage von Hautfarbe im Farbwürfel ist, kann man anschaulich in [SCMA] nachvollziehen. Hautfarbe setzt sich aus den Komponenten Blut (rot) und Melanin (gelb, braun) zusammen [FFNP]. Diese Farben variieren in den Anteilen der Zusammensetzung. Je nach Hautfarbe des Menschen wird etwas mehr rote, gelbe oder bräunliche Farbe dazugemischt. Diese Zusammensetzung läßt so asiatische, japanische oder afrikanische Hautfarben entstehen. Wenn man nun die Farben gefunden hat, die Hautfarbe ausmachen, ist das Problem immer noch nicht gelöst, da sich auf Bildern die Beleuchtung noch auswirkt. Bei starker Beleuchtung treten Bereiche auf, die eigentlich zur Hautfarbe gehören, sich aber fast weiß darstellen. Ebenso werden durch die starke Beleuchtung Schatten geworfen, woraufhin Teile eines Bildes, die eigentlich zur Hautfarbe gehören, fast schwarz erscheinen. Solche Bereiche zu erkennen, ist fast nicht möglich, es sei denn, sie liegen mitten in einer Hautpartie und sind somit von Haut umgeben.

2.2 Das RGB-Farbmodell

In Farbbildern wird das RGB-Farbsystem verwendet. Beim Auslesen der Bilder wird festgestellt, daß die Farben der Punkte im RGB-System gespeichert sind. Aus diesem Grund ist es naheliegend, das RGB-Farbmodell für die Erkennung von Hautfarbe zu verwenden.

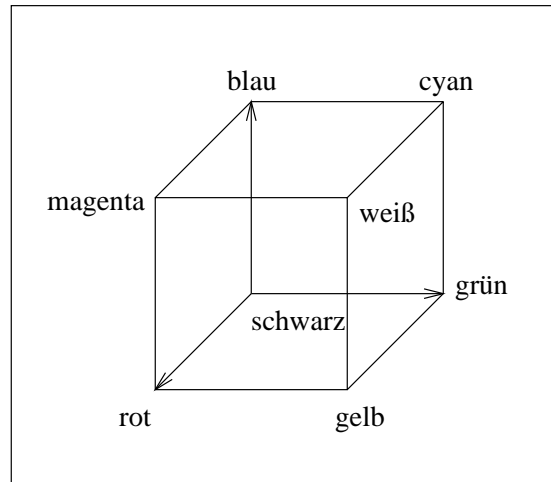


Abbildung 2.1: Der RGB Würfel

Das RGB-Farbsystem ist ein technisches, additives Farbsystem, das vor allem im Bereich Fernsehtechnik und Computergrafik Verwendung findet. Es basiert auf einem kartesischen Koordinatensystem. Die Betrachtung des RGB-Würfels beginnt in der schwarzen Ecke, wie in Abbildung 2.1 verdeutlicht. Zu der schwarzen Farbe werden die Farben rot, grün oder blau hinzuaddiert. Diese drei Komponenten können in Ausprägungen von 0 bis 1 angegeben werden, wobei dadurch alle Farben des Farbwürfels entstehen können. Sind alle Komponenten in voller Ausprägung dargestellt, so entsteht die Farbe weiß. In den anderen Ecken des Würfels, also bei jeweils einer vollen Ausprägung findet man die Farben rot, grün und blau, bei zwei vollen Ausprägungen die Farben gelb, cyan und magenta.

Das RGB-Farbsystem liegt der Darstellung von Farben auf dem Bildschirm zugrunde. Die Helligkeit der drei Hauptachsen (rgb) ist dabei für die Anzahl der darzustellenden Farben verantwortlich. Wenn die Stärke der Beleuchtung pro Phosphorprisma auf dem Bildschirm 256 beträgt, so befindet sich die Ausprägung jeder Achse im Bereich von 0 bis 255, statt 0 bis 1. Aus dieser Darstellung entsteht ein Farbraum von ungefähr 16 Millionen Farben [CG]. In dieser Arbeit wird der Farbraum von 0 bis 255 auf jeder Achse betrachtet.

2.3 Das HSV-Farbsystem

Auch das HSV-Farbsystem soll an dieser Stelle erklärt werden, da es in einer der Methoden zur Erkennung von Hautfarbe benutzt wird. Leider sind die Vorteile des Modells, auf die nach der Durchführung und Auswertung der dazugehörigen Methode noch eingegangen wird, erst deutlich geworden, als bereits alle anderen Me-

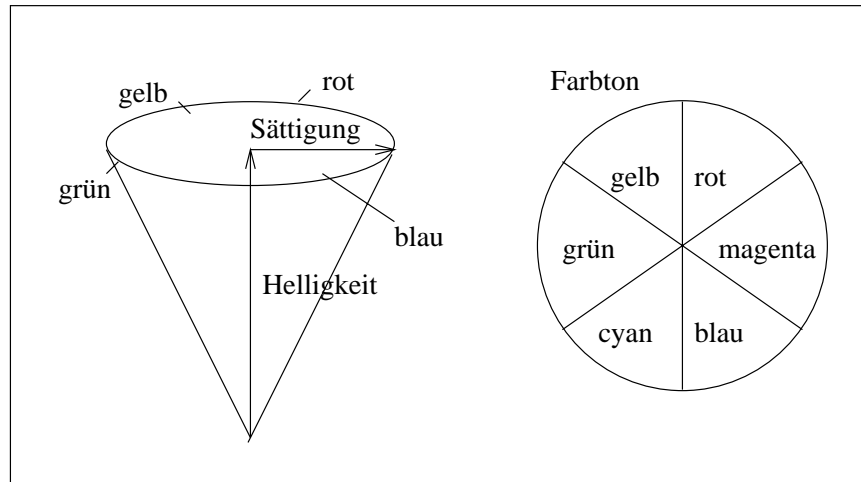


Abbildung 2.2: HSV-Farbsystem

thoden auf dem RGB-System aufsetzen.

Das HSV-Farbmodell (hue, saturation, value) findet Verwendung im Bereich Computergrafik zur interaktiven Spezifikation von Farbwerten, es gehört zu den perceptuellen Farbräumen. Diese Modelle werden durch Farbton, Helligkeit und Sättigung beschrieben. Man kann sich das Modell gut als Kegel vorstellen, wie in Abbildung 2.2 zu sehen ist. Die Grundfläche bildet ein Kreis, in dem die Farbtöne angeordnet sind. Die Anordnung der Farben beginnt mit der roten, die dabei einen Winkel von 0 und 360 Grad erhält. Die Winkel werden gegen den Uhrzeigersinn auf dem Kreis angeordnet. Dadurch erhält Grün den Winkel 120 Grad und Blau 240 Grad. Die Spitze des Kegels hat die Helligkeit 0, wobei diese auf dem Weg zur Grundfläche größer wird, bis sie 1 erreicht. Die Sättigung, welche sich ebenfalls im Bereich von 0 bis 1 bewegt, hat den Wert 0 auf der Mittelachse der Pyramide und nimmt mit dem Radius zu. Auf der Mantelfläche der Pyramide erreicht sie 1.

Das HSV-Farbmodell beschreibt die gleiche Anzahl von Farben, wie das RGB-System. Die Farben können von RGB in HSV umgewandelt werden und umgekehrt.

Kapitel 3

Methoden zur Erkennung von Hautfarbe

In diesem Kapitel sollen die Methoden erklärt werden, die benutzt werden sollen, um Hautfarbe in Bildern ausfindig zu machen. Die Methoden wurden von der Funktionsweise sehr unterschiedlich gewählt, um einen guten Querschnitt über verschiedene Herangehensweisen zu bekommen.

Begonnen wird mit einer Selektion, die Punkte aus einem Bild selektiert und diese einfärbt. Dazu werden zwei unterschiedliche Ausprägungen der Selektion vorgestellt. Bei der Selektion handelt es sich um ein intuitives Verfahren. Bei der linearen Klassifikation wurde versucht, um den Körper der Hautfarbe eine Hülle zu bilden, um diese Farben ebenfalls als Haut zu erkennen. Die nächste Methode wird mathematischer und heißt Regression. Ihr soll es gelingen, die Hautfarben in einer linearen Gleichung darzustellen. Das folgende Verfahren der Neuronalen Netze soll versuchen Funktionsweisen des Gehirns nachzubilden. Aus vorhandenen Forschungen zur Gesichtserkennung ist die nächste Methode des chromatischen Farbraums herausgelöst worden. Sie bildet einen Teil der Gesamtforschung zur Gesichtserkennung. Als letztes Verfahren wird das HSV-Farbsystem vorgestellt, da es erlaubt die Helligkeit bei der Farb-Erkennung auszuschalten.

3.1 Farbselektion

Die Möglichkeiten der Selektion, die im Folgenden vorgestellt werden, entsprechen beide dem gleichen Schema. Der Ablauf der Methoden wird in Abbildung 3.1 verdeutlicht. Die Unterschiede werden in der Beschreibung der Verfahren aufgezeigt.

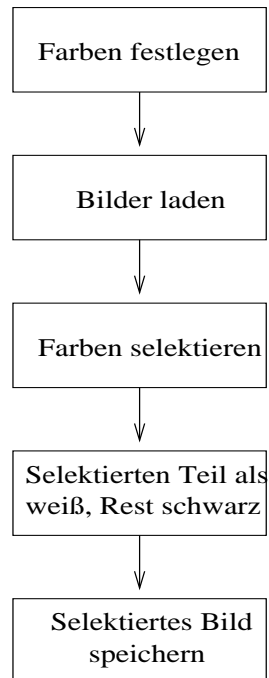


Abbildung 3.1: Ablauf der Farbselektion

3.1.1 Farbselektion aus 16 Millionen Farben

Um alle Bilder möglichst naturgetreu zu bearbeiten, bietet sich eine Selektion im Bereich der 16 Millionen Farben an. Fast alle Bilder werden in dieser Größenordnung dargestellt. Man hat damit eine sehr einfache, aber genaue Methode.

Um Hautfarbe aus einem Bild zu extrahieren, bietet es sich an, Farbwerte, die Haut entsprechen, aus dem Bild zu selektieren. Dazu ist es nötig, die Farben, die man selektieren möchte, zu kennen. Es ist bei einem Spektrum von 16 Millionen Farben nicht möglich, Hautfarben per Hand auszuwählen. Deshalb wäre es sinnvoll, Bereiche zu definieren, die Haut entsprechen. Es ist aber bei kaum einem bekannten Programm möglich, einen Farbbereich anzugeben, der herausgefiltert werden soll. Eine Möglichkeit in diesem Zusammenhang besteht darin, eine Reihe von Farben durch ihre RGB-Werte anzugeben. Automatisch werden dann die angegebenen Werte und ähnliche Farben aus dem Bild selektiert.

Diese Methode liefert, wie sich im folgenden zeigen wird, gute Ergebnisse, da hier wesentlich mehr Farben selektiert werden, als eigentlich angegeben sind.

3.1.2 Farbselektion aus 256 Farben

Der nächste Ansatz, den es im Bereich der einfachen Selektion gibt, ist die vorherige Transformation des Bildes. Diese Methode ist sinnvoll, da die Hautfarben durch Auswählen aus 256 Farben eingegrenzt werden können.

Im Regelfall liegt ein Bild im 16 Millionen-Farbbereich vor. Um die Hautfarben nicht aus 16 Millionen Farben spezifizieren zu müssen, ist die Idee, das gesamte Bild auf einen Farbbereich von 256 Farben herunter zu transformieren. Damit werden alle Farben aus 16 Millionen auf eine ähnliche Farbe aus 256 abgebildet.

Aus 256 können Farben, die Hautfarbe darstellen sollen, per Hand ausgesucht werden. Hier wurden 11 Farben aus 256 ausgewählt und als Hautfarbe definiert. Diese Farben werden in jedem Bild als Hautfarbe selektiert.

Bei dieser Methode erwartet man bessere Ergebnisse als bei der Selektion aus 16 Millionen Farben, da "alle" Hautfarben gewählt werden können. Dazu müssen aber einige Überlegungen einbezogen werden. Bei der Transformation vom 16 Millionen Farb-Bereich auf 256 Farben können einige Bildinformationen verloren gehen.

3.2 Lineare Klassifikation

Die Hautfarben sind in einer Menge definiert. Diese Menge kann man benutzen, um ähnliche Fälle zu finden. Eine neue Farbe kann daraufhin untersucht werden, ob sie in diese Menge hineinpaßt oder nicht.

Die lineare Klassifikation ist durch Überlegungen zur fallbasierten Klassifikation entstanden. Die grundsätzliche Idee der fallbasierten Klassifikation soll hier zunächst einmal erklärt werden. Die fallbasierte Klassifikation ist ein Ansatz zum Modellieren menschlichen Denkens. Dieser Ansatz kommt aus kognitionswissenschaftlichen Arbeiten, und auf diesem Gebiet wird in den USA schon seit 1977 und in Europa seit ca. 1988 geforscht. Die fallbasierte Klassifikation versucht, das menschliche Denken nachzubilden, indem es Probleme und deren Lösungen spezifiziert und aufbewahrt. Entsteht nun ein neues Problem, so wird versucht, ein ähnliches schon mal bearbeitetes Problem zu finden und dessen Lösung so anzupassen, das eine Lösung für das neue Problem entsteht. Der große Vorteil dieser Methode besteht darin, daß vorhandene Datenbestände wieder benutzt werden können. Zunächst muß natürlich ein Grundstock angelegt werden. Die Resultate zu diesen Problemen werden von Grund an neu entwickelt. Danach kann man aber all diese Lösungen benutzen, um andere Probleme effizient zu bearbeiten.

Der Zyklus der fallbasierten Klassifikation besteht aus 4 Zuständen, die auch in Bild 3.2 deutlich gemacht sind.

1. Das Finden eines ähnlichen Problems. Dieser Teil wird durch verschiedene Methoden realisiert. Zum Beispiel wird durch lineare Listen, Datenbanken, Bäume oder Diskriminationsnetze nach ähnlichen Problemen gesucht.
2. Hier wird die Adaption des Problems bearbeitet. Ist keine Anpassung nötig, wird die Lösung übernommen. Ansonsten findet in dieser Phase eine Anpassung der Lösung statt. Dies kann durch Benutzerinteraktion, regelbasiertes Schließen oder Constraintbefüllungsverfahren umgesetzt werden.
3. Die Überprüfung der Lösung ist eine wichtige Phase. Jedoch wird hier keine Verbesserung mehr vorgenommen. Die Ergebnisse werden durch Simulation überprüft und bewertet.
4. In der letzten Phase werden die Ergebnisse gespeichert, um sie später wieder verwenden zu können. Unter Gesichtspunkten von Ähnlichkeit und Merkmalen werden die neuen Lösungen in die vorhandenen Struktur aufgenommen [UKB97].

3.2.1 Adaption der linearen Klassifikation

Für diese Zwecke wird die Idee der fallbasierten Klassifikation genutzt und wie folgt in einer Methode beschrieben, die lineare Klassifikation heißen soll. Alle

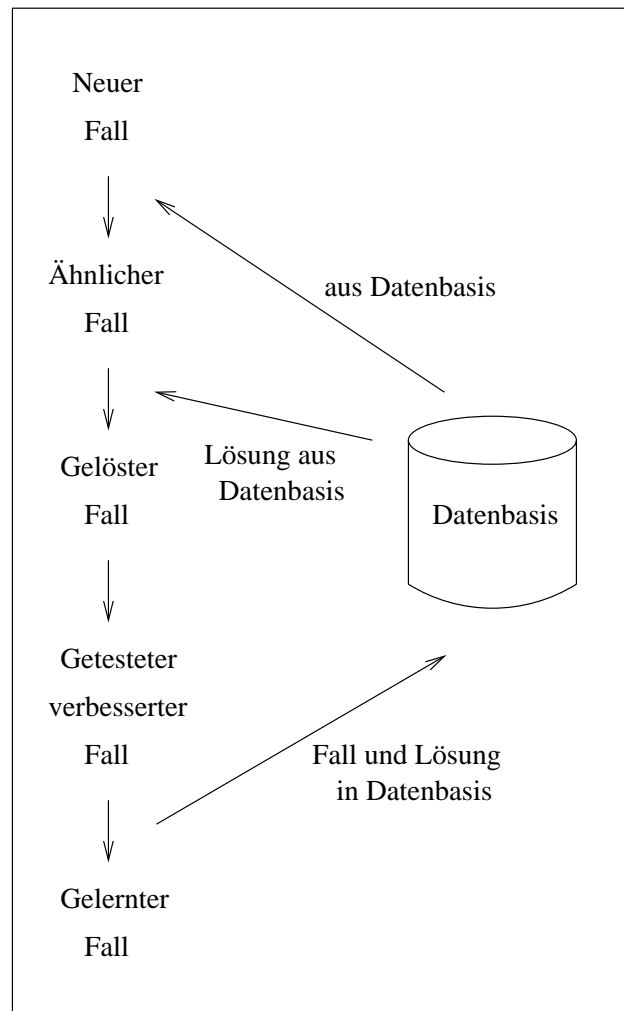


Abbildung 3.2: Fallbasierte Klassifikation

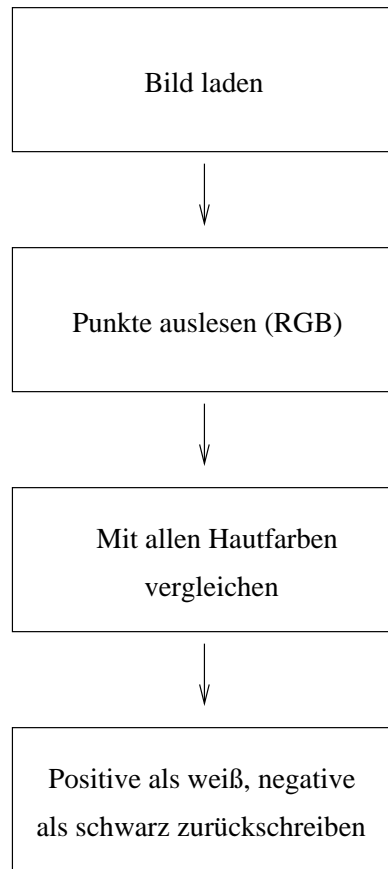


Abbildung 3.3: Ablauf der linearen Klassifikation

RGB-Punkte, die Hautfarbe darstellen, werden in einer Datei gehalten. Wenn jetzt ein neuer Punkt klassifiziert werden soll, wird er mit jedem Element aus der Hautfarbendatei verglichen. Liegt der Punkt in einer festgelegten Distanz zu einer der Hautfarben, dann wird er als Hautfarbe identifiziert, ansonsten nicht. Eine sinnvolle Distanz ist vorher festzulegen. Im Gegensatz zur Idee der fallbasierten Klassifikation werden neue Punkte nicht neu mit in die Datenbasis aufgenommen. Die Aufnahme der Punkte würde den Radius um das aus Hautfarben bestehende Gebilde zu groß werden lassen. Die Distanzen würden sich addieren und die Ergebnisse dadurch verfälscht werden.

Diese Methode weist Ähnlichkeiten zur Selektion aus 256 Farben auf. Wenn die beiden Verfahren genau betrachtet werden, wird festgestellt, daß sie umgekehrt arbeiten. Die Selektion schränkt die Hautfarben ein, indem sie ähnliche Farben, aus dem Umkreis, auf eine Farbe im 256 Farbraum abbildet. Danach werden nur die Farben selektiert, die sich im 256 Farbbereich noch als Hautfarbe darstellen. Bei der linearen Klassifikation wird umgekehrt vorgegangen. Als erstes werden sehr

viele Hautfarben aus dem 16 Millionen Farbbereich festgelegt. Anschließend wird um diesen Bereich ein Kreis, von Größe der unterschiedlichen Schwelle mitselektiert.

Die Ergebnisse der linearen Klassifikation werden deshalb den Ergebnissen der Selektion ähneln. Hier werden Farben mit einer bestimmten Spanne zu den definierten Hautfarben als Hautfarbe erkannt. Es muß nicht genau die definierte Farbe getroffen werden, sondern ein Bereich um die jeweilige Hautfarbe herum wird zur Hautfarbe gezählt. Der Bereich kann natürlich auch zu groß sein, und dadurch würden zu viele Farben als Haut erkannt.

3.3 Regression

Es wird vermutet, daß die Hautfarbe keinen komplexen Bereich, sondern einen zusammenhängenden Raum beschreibt. Aus diesem Grund bietet sich die Methode der Regression an. Die lineare Regression soll lernen, den Hautfarbenbereich durch eine Funktion zu beschreiben.

Die einfache lineare Regressionsanalyse sucht nach einer linearen Gleichung, die den Zusammenhang zwischen x und y zum Ausdruck bringt. x ist dabei die sogenannte exogene Variable (Einflußfaktor), während y die endogene, also Zielvariable ist [UBL0599]. In diesem Problem stellt der RGB-Wert den Einflußfaktor dar, und die Zielvariable ist eine positiv/negativ-Variable, die die Aussage beinhaltet, ob der dazugehörige RGB-Wert hautfarbig ist(1) oder nicht(0).

Die Regressionsrechnung verfolgt zwei Ziele:

- Die Gewinnung einer handlichen Beschreibung des Zusammenhangs zwischen x und y , der sich aus den Beobachtungsdaten herauslesen läßt.
- Die Erstellung von Prognosen der y -Werte, die für noch nicht beobachtete x -Werte zu erwarten ist.

Es wird mit der linearen Regression gearbeitet, deshalb könnte es Probleme dabei geben, einen Körper zu beschreiben. Außerdem könnten Ausbrecher im Hautfarbenbereich und Unebenheiten im Körper zu Problemen führen.

3.4 Neuronale Netze

Aus den gleichen Gründen wie bei der Regression werden auch die Neuronalen Netze als eine gute Methode gesehen, den Hautfarbenbereich zu lernen. Es handelt

sich um eine relativ kompakte Menge, sodaß das Lernen in einem Neuronalen Netz möglich sein sollte.

Das Prinzip der Neuronalen Netze basiert auf der Funktionsweise der Nervenzellen des Gehirns. Nach ihrem Vorbild wurden die Neuronalen Netze geschaffen. Ein künstliches Neuronales Netz besteht aus mehreren Neuronen, d.h. Zellkörper, Dendriten und Axon, die untereinander verbunden sind. Die Dendriten sind für den Empfang von Informationen zuständig, der Zellkörper für die Verarbeitung und schließlich das Axon zur Weitergabe eines Impulses. Es sind also mindestens drei Schichten von Nöten, die Eingabe-Schicht (input-layer), die verborgene Schicht (hidden-layer) und die Ausgabe-Schicht (output-layer). Das Axon eines Neurons verzweigt sich und tritt über Synapsen mit den Dendriten anderer Neuronen in Verbindung. Ein künstliches Neuron besteht aus $(\vec{x}, \vec{w}, f_a, f_o, o)$.

$\vec{x} = (x_1 \dots x_n)$ Eingabevektor
 $\vec{w} = (w_1 \dots w_n)$ Gewichtsvektor
 f_a Aktivierungsfunktion $f_a : IR \times IR \rightarrow IR$
 f_o Ausgabefunktion $f_o : IR \rightarrow IR$
 o Ausgabewert des Neurons $f_o(f_a(\vec{x}, \vec{w})) = o$

Häufig wird als Aktivierungsfunktion die gewichtete Summe verwendet:

$$f_a(\vec{x}, \vec{w}) = \sum_{l=1}^n x_l w_l$$

Verbindet man nun mehrere Neuronen miteinander, so können dadurch verschiedene Topologien von Netzwerken entstehen. Zu unterscheiden sind Netzwerke mit und ohne Rückkopplung. Bei Netzwerken ohne Rückkopplung handelt es sich um azyklische Graphen. Diejenigen mit Rückkopplung werden entweder direkt in ihrer Schicht oder in einer anderen verstärkt [WWU].

Die Probleme, die bei dieser Methode entstehen, sind die gleichen wie bei der Regression. Unebenheiten und Ausreißer können zu unklaren Ergebnissen führen.

3.5 Chromatischer Farbraum

Im Bereich der Gesichtserkennung wird viel geforscht. Häufig, wie auch in [DHF] und [FDCI] wird bei der Gesichtserkennung zunächst eine Erkennung von Haut durchgeführt. Viele der Forschungen gehen auf die Methode von Chang und Robles zurück, die im folgenden vorgestellt werden soll.

Bei der Haut-Erkennung in Gesichtern wird oft der Ansatz der Wahrscheinlichkeitsrechnung benutzt. Hier wird für jeden Punkt in einem Bild die Wahrscheinlichkeit dafür berechnet, daß es sich bei ihm um Hautfarbe handelt. Da die Hautfarbe in Gesichtern die gleiche ist wie die allgemeine Hautfarbe, bietet es sich an,

diese Methode zu untersuchen. Wie die Methode ganz allgemein abläuft, kann man in Abbildung 3.4 sehen.

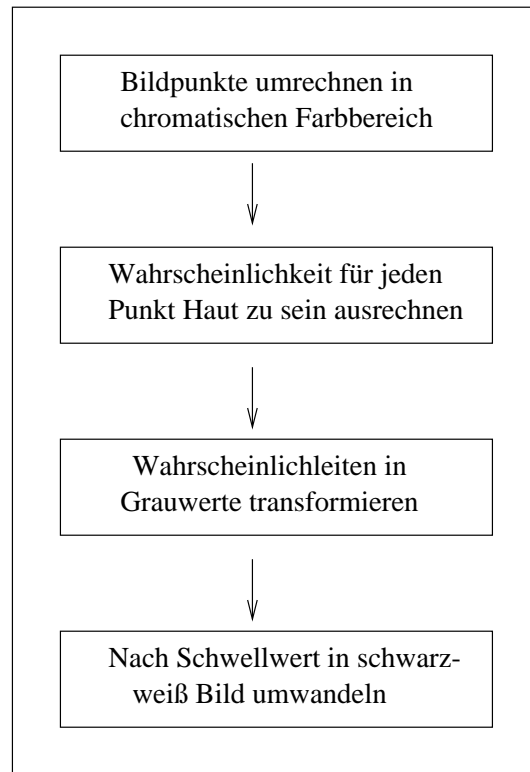


Abbildung 3.4: Haut-Erkennung als Teil der Gesichtserkennung

Henry Chang und Ulises Robles haben auf dem Gebiet der Gesichtserkennung hervorragende Erfahrungen gemacht. Sie haben sich speziell mit einer Hautfarben-Erkennung beschäftigt. Sie benutzen dabei nicht den RGB- sondern den chromatischen Farbraum, da im RGB von den drei Komponenten nicht nur Farbe sondern auch Helligkeit repräsentiert wird. Helligkeit variiert aber über einem Gesicht und ist deshalb keine zuverlässige Maßeinheit, in der man Haut von Nicht Haut unterscheiden kann. Im chromatischen Farbraum wird die Helligkeit aus der Farbrepräsentation entfernt. Chromatische Farben werden als "pure" Farben bezeichnet, bei denen die Helligkeit fehlt. Sie können mit einem Normalisierungsprozeß definiert werden:

$$r=R/(R+G+B)$$

$$b=B/(R+G+B)$$

grün ist redundant bei $r+g+b = 1$

Obwohl die Hautfarbe verschiedener Personen in einer großen Spanne variiert, unterscheidet sie sich viel weniger in der Farbe, als in der Helligkeit. Deshalb kommt

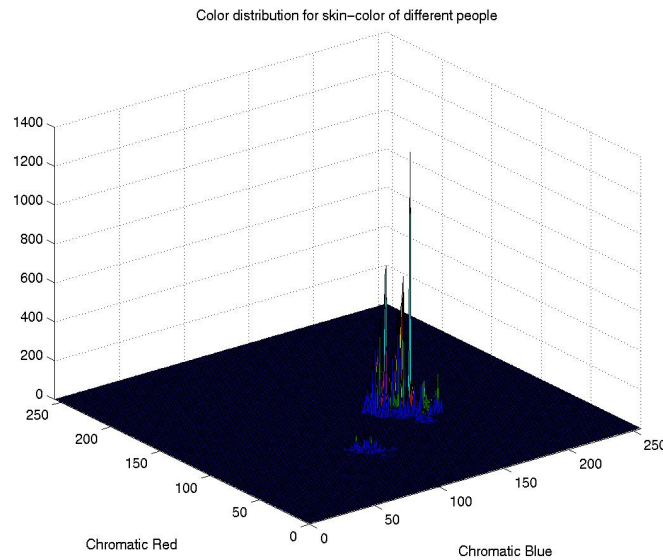


Abbildung 3.5: Hautverteilung im chromatischen Farbraum

die Hautfarbe verschiedener Personen dicht versammelt, wie in Abbildung 3.5, im chromatischen Farbraum vor.

Chang und Robles haben eine Gesamtheit von 32500 Hautstichproben aus 17 Farbbildern benutzt, um die Verteilung von menschlicher Haut im chromatischen Farbraum zu untersuchen. Die Hautstichproben wurden zunächst mit einem Tiefpaßfilter gefiltert, um die Rauscheffekte zu reduzieren.

Ein Farb-Histogramm zeigt, daß die Verteilung von Hautfarbe von verschiedenen Personen im chromatischen Farbraum sehr dicht zusammenhängt und mit einem Gauß-Modell $N(m,C)$ repräsentiert werden kann, wobei:

$$\begin{aligned} \text{Mittelwert: } m &= E\{x\} \text{ wobei } x = (rb)^T \\ \text{Kovarianz: } C &= E\{(x - m)(x - m)^T\} \end{aligned}$$

Mit diesem nach Gauß-Verteilung zugeschnittenen Farbmodell kann jetzt die Wahrscheinlichkeit von Haut für jeden Punkt eines Bildes ermittelt werden. Wenn ein Punkt vom RGB in den chromatischen Raum transformiert wurde, kann die Wahrscheinlichkeit von Haut für diesen Punkt wie folgt berechnet werden:

$$\begin{aligned} P(r, b) &= \exp[-0.5(x - m)^T C^{-1}(x - m)] \\ \text{wobei } x &= (r, b)^T \end{aligned}$$

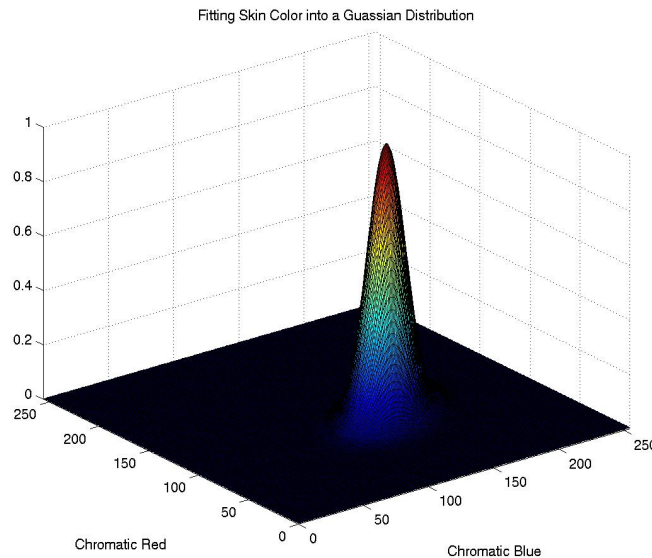


Abbildung 3.6: Gauß-Verteilung

Dieses Haut-Farbmodell kann ein Farbbild in ein grau skaliertes Bild transformieren, sodaß die Grauwerte jeden Punktes die Wahrscheinlichkeit für diesen Punkt "Haut zu sein" angibt. Mit einem passend gewählten Schwellwert kann das grau skalierte Bild in ein binäres Bild transformiert werden, das Haut- und Nicht-Haut-Regionen zeigt [ChRo0500].

Da die Untersuchungen von Chang und Robles in Gesichtern sehr gute Ergebnisse liefern, wird dieses auch für andere Körperteile erwartet. Allerdings hängen die Ergebnisse sehr stark von der Lernmenge der Hautfarben ab. Ist die Lernmenge zu einseitig oder zu klein, sind die Ergebnisse schlechter als bei guter Wahl der Lernmenge.

3.6 Die Transformation in den HSV-Farbraum

Da das HSV-Farbmodell eine Farbzusammenstellung besitzt, wie sie vom Menschen als natürlich empfunden wird, soll auch eine Anwendung aus diesem Bereich untersucht werden.

Die Methode die aus dem HSV-System entsteht, soll so vorgehen, daß alle Farben, die als Haut definiert wurden, in Werte des HSV-Farbmodells umgewandelt werden. Aus diesen Werten werden die Minima und Maxima von Farbton und Sättigung bestimmt. Bei der anschließenden Überprüfung einzelner RGB-Werte auf

Hautfarbe werden diese zunächst in einen HSV-Wert umgesetzt. Danach wird kontrolliert, ob diese Werte im Bereich von Farbton und Sättigung liegen. Ist das der Fall, so gelten sie in diesem Modell als Hautfarbe, anderenfalls nicht.

Farbton und Sättigung beinhalten Farbinformationen, während Helligkeit keine Farbinformationen enthält [CG]. Aus diesem Grund kann in der Betrachtung die Helligkeit unbeachtet bleiben und der Farbraum für Hautfarbe durch nur zwei Komponenten beschrieben werden. Da die Helligkeit in den anderen Verfahren Probleme bei der Erkennung von Hautfarbe bringt, wird dieses System, da es die Helligkeit nicht betrachtet, gute Ergebnisse produzieren.

Bei der Haut-Erkennung für die Gesichtserkennung gibt es Verfahren die auch mit den HSV-Werten arbeiten. Zum Beispiel werden in [FSCI] Farb-Histogramme benutzt, um Hautfarbe ausfindig zu machen. Einen ähnlichen Ansatz hat auch [STCM], wobei hier, wie bei der Methode im chromatischen Farbraum in dieser Arbeit, mit Wahrscheinlichkeiten gearbeitet wird.

Kapitel 4

Umsetzung

Die Umsetzung der in Kapitel 3 beschriebenen Methoden wurde auf ganz unterschiedliche Weise realisiert. So wurde die Selektion mit Hilfe eines Grafikbearbeitungsprogrammes namens Gimp durchgeführt. In Gimp ist es möglich, auf einem Bild verschiedene Befehle auszuführen und diese als Makro abzulegen. So kann der Prozeß automatisiert werden. Für die Methode im chromatischen Farbraum gibt es Quelltexte, die in matlab verarbeitet werden können. Dafür müssen vorher lediglich die Hautteile, mit denen das Programm lernen soll, angegeben werden. Die lineare Klassifikation und die Haut-Erkennung mit Hilfe des HSV-Farbsystems wurden in C++ realisiert. Dabei muß bei der linearen Klassifikation immer wieder die Hautfarbendatei mit zu überprüfenden Farbwerten verglichen werden, während beim HSV-System die Hautfarben nur einmalig gelesen werden und dann ein Bereich festgelegt wird, auf den andere Farben getestet werden.

4.1 Selektion mit Gimp

Diese beiden ersten Methoden wurden in Gimp realisiert. Gimp ist ein Bildbearbeitungsprogramm unter Linux/Unix. In Gimp kann man eigene Funktionalitäten definieren und automatisch ausführen lassen. Das Programm stellt wohl das Pendant zu PhotoShop unter Windows dar. Die Funktionen wurden unter Skript-Fu zusammengestellt, wie man mit diesem Modul umgeht ist in [GUM] beschrieben.

Für die erste Selektion sind 8 Farben aus 16 Millionen ausgewählt worden, die Haut darstellen sollen. Die RGB-Werte dieser 8 Farben sind in Tabelle 4.1 zu sehen. Der Vorgang der Selektion ist unter Script-Fu so aufgebaut worden, daß zunächst das jeweilige Bild geladen wird. Danach werden in dem Bild die 8 Farben aus Tabelle 4.1 selektiert und zusätzlich werden von Gimp intern ähnliche Farben selektiert. Die Ähnlichkeitsfunktion, die von Gimp benutzt wird, ist sehr mächtig. Sie stellt sehr gut fest welche Farben, den angegebenen ähnlich sind. Aus dieser Selektion

Nr.	Rot	Grün	Blau
1	201	154	134
2	166	119	101
3	243	180	149
4	156	113	96
5	182	95	75
6	230	179	122
7	202	126	92
8	253	197	182

Tabelle 4.1: Farbauswahl aus 16 Millionen

Nr.	Rot	Grün	Blau
1	153	0	0
2	204	0	0
3	153	51	0
4	153	102	0
5	204	102	0
6	204	102	51
7	204	102	102
8	204	153	0
9	204	153	102
10	204	153	153
11	255	204	102

Tabelle 4.2: Farbauswahl aus 256 Farben

entsteht ein neues Bild, in dem alle selektierten Bereiche, also die als Hautbereich definierten, in weißer Farbe dargestellt werden und der gesamte Rest in schwarz. Dieses schwarz-weiß Bild wird dann abgespeichert.

Für die zweite Selektion im 256-Farbbereich wurde eine Transformation des aktuellen Bildes vorgeschaltet. Das Originalbild liegt im 16 Millionen Farbbereich vor. Die Transformation ordnet jedem Punkt des Bildes einen Punkt im 256 Farbbereich zu, der möglichst nahe an dem ursprünglichen Punkt liegt. Danach wird wie bei der ersten Methode verfahren. Es werden diesmal die 11 Farben aus Tabelle 4.2 gewählt, die die Haut im 256-Farbbereich repräsentieren sollen. Diese Farben werden im 256-Farbbild selektiert und das neu generierte schwarz-weiß Bild gespeichert.

4.2 Lineare Klassifikation in C++

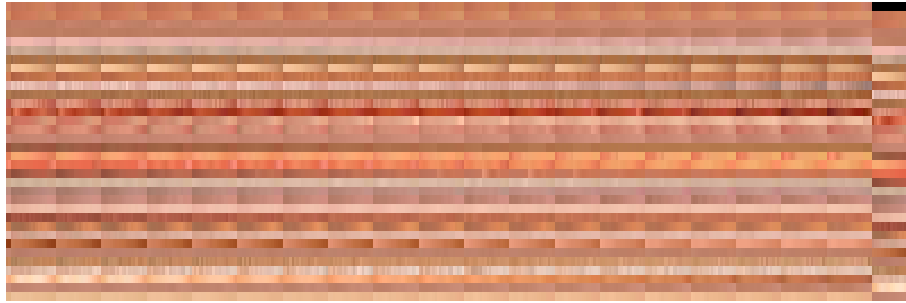


Abbildung 4.1: Auswahl der Hautfarben

4.2.1 Hautfarbendefinition für die folgenden Methoden

Für die lineare Klassifikation und das HSV-Farbsystem werden nicht nur einzelne Farben, sondern viele Farbbereiche benötigt. Um diese nicht einzeln per Hand auszuwählen, werden Teile von Bildern benutzt, die Hautfarbe darstellen. Aus diesen Teilen der Bilder werden die RGB-Werte herausgelesen und in eine Datei geschrieben. Auf diese Weise ist eine Datei mit ungefähr 14800 RGB-Einträgen entstanden. Die Farbdarstellung der gesamten Hautfarbendatei ist in Bild 4.1 abgebildet. Das Bild stellt die Farben dar, mit denen die einzelnen Bildpunkte bei der Klassifikation verglichen werden, und deren Farbton- und Sättigungs-Bereich nach Umrechnung in HSV festgestellt werden.

4.2.2 Lineare Klassifikation

Die lineare Klassifikation ist über eine Programmierung in C++ gelöst worden. Dabei existiert eine Datei, in der alle Hautfarben abgelegt sind. Außerdem gibt es eine Schwelle, die beliebig angegeben werden kann. Die RGB-Werte der zu untersuchenden Bilder werden ausgelesen und jeder einzelne Punkt des Bildes mit allen Hautfarben verglichen. Liegt der jeweilige Punkt höchstens im Bereich eines Hautpunktes plus Schwelle, so wird er als Hautfarbe identifiziert. Liegt er außerhalb des Bereichs, wird er nicht als Hautfarbe erkannt. Für jeden Punkt wird so ein weißer Punkt für Haut und ein schwarzer Punkt für Nicht-Haut zurückgegeben; diese lassen ein neues Bild entstehen. Wie die lineare Klassifikation umgesetzt worden ist, sieht man in Abbildung 4.2.

4.2.3 Octree Datenstruktur

Für die Wahl einer Datenstruktur gibt es mehrere Entscheidungskriterien, zum Beispiel die Menge und Art der Daten, die gespeichert werden sollen. Bei diesem

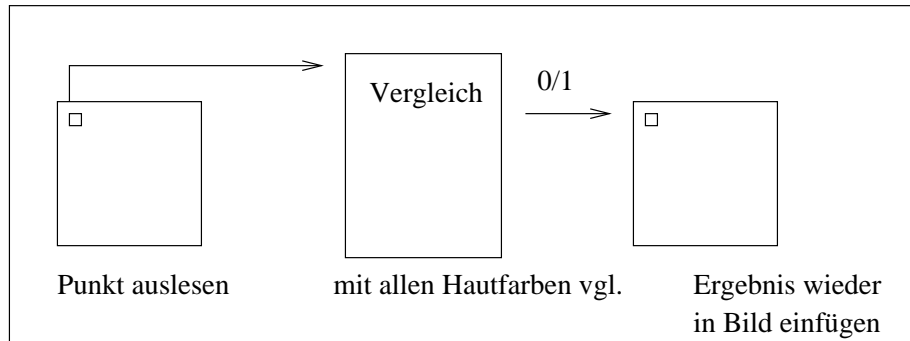


Abbildung 4.2: Umsetzung lineare Klassifikation

Problem wurde ein Octree gewählt, weil er ein Verfahren zur Speicherung von 3D-Objekten anbietet. Der 3D-Raum des Octree, in diesem Fall ein Würfel, teilt sich in 8 Subräume, d.h., in jeder Ausprägung wird der Würfel in der Mitte unterteilt. Jeder einzelne Subraum kann natürlich entsprechend weiter unterteilt werden. Der Aufbau einer Octree-Datenstruktur ist in Abbildung 4.3 zu sehen [Char].

Die Hautfarben sollen in die Datenstruktur eingefügt werden. Dabei stellt jede Farbe(RGB) eine Dimension dar. Der Octree wird auf jeder Ebene in gleich große Teile unterteilt. Die Ausdehnungen in xyz-Richtung sind die R-, G- und B-Werte.

Um bei der oben genannten Methode der linearen Klassifikation eine Laufzeitverbesserung zu erzielen, ist es sinnvoll die Daten geordnet in einer Struktur abzulegen, um so schneller an die gewünschten Werte zu kommen. Wenn ein Punkt auf Hautfarbe untersucht werden soll, wird im Octree nach dem nächsten Nachbarn zu diesem Punkt gesucht. Der Abstand zum nächsten Nachbarn wird festgestellt und es wird überprüft, ob dieser Abstand unter der angegebenen Schwelle liegt. Die Verfahrensweise ist also die gleiche wie bei der linearen Klassifikation, nur ist es wesentlich schneller, im Octree einen Nachbarn zu suchen, als die gesamte Hautfarbendatei zu überprüfen.

4.3 Chromatischer Farbraum mit matlab

Für die Gesichtserkennung von Chang und Robles gibt es matlab Quelltexte, die man benutzen kann. Diese Programme sind geringfügig zu verändern.

Zum Lernen werden Hautteile benötigt, die die Autoren nicht zur Verfügung stellen. Diese Hautteile sollen alle gleich groß sein und wurden von der Farbe möglichst unterschiedlich gewählt. D.h. helle und dunkle Hautpartien wurden gewählt und in die Programme integriert. Die Hautpartien werden verschiedenen Bildern entnommen. Das ist wichtig, damit das Programm mit den richtigen Farbwerten

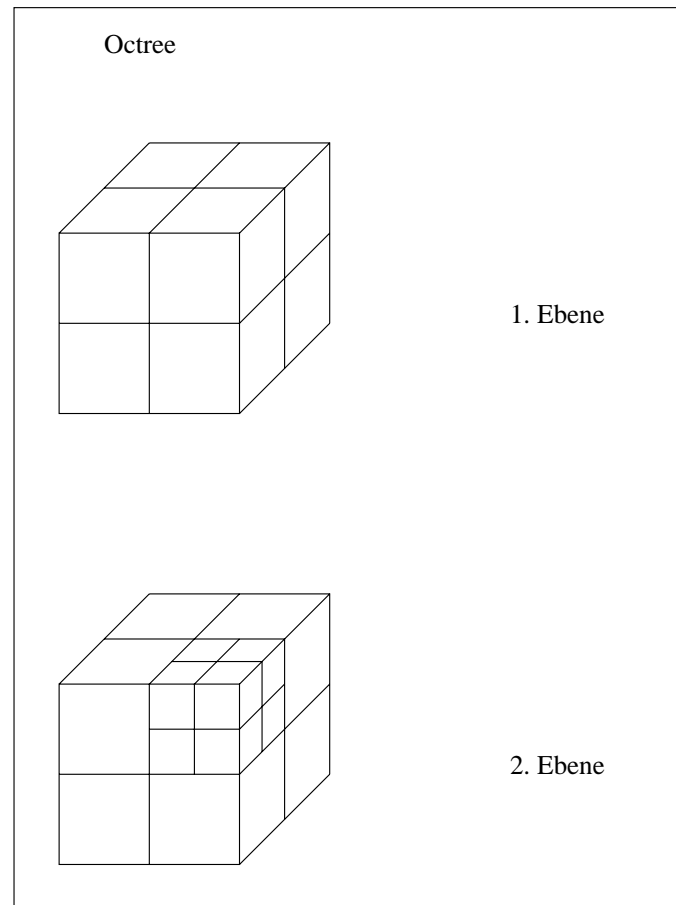


Abbildung 4.3: Aufbau einer Octree-Datenstruktur

angelernt wird. Wenn bei der Berechnung später Hauttypen verschiedenfarbiger Menschen auftauchen, können all diese erkannt werden. Nachdem die den Anforderungen entsprechenden Hautpartien gewählt wurden, werden auch aussagekräftige Ergebnisse erzielt.

4.4 HSV in C++

Die Umsetzung des HSV-Farbsystems fordert zunächst eine Umrechnung der RGB-Werte in das HSV-System. Dazu werden alle Hautwerte, die auch bei der linearen Klassifikation benutzt worden sind, in das HSV-System umgesetzt. Dies geschieht mit Hilfe des folgenden Algorithmus aus [CG].

Umrechnung RGB→HSV

```

min = min(R, G, B)
max = max(R, G, B)
V = max
if(max == min)
    S = 0
    H = 0
else
    S = (max - min)/max
    if(R == max)
        H = ((G - B)/(max - min)) * 60
    elseif(G == max)
        H = (2 + (B - R)/(max - min)) * 60
    elseif(B == max)
        H = (4 + (R - G)/(max - min)) * 60

```

Nachdem alle 14800 Hautwerte auf das HSV-Farbsystem umgesetzt sind, werden die Minima und Maxima der H- und S-Werte bestimmt. Die Resultate haben folgende Ergebnisse, wobei H den Farbton und S die Sättigung beschreibt:

```

H_MIN = 0
H_MAX = 50
S_MIN = 0.17
S_MAX = 1

```

Wenn nun ein Bild auf Hautfarben überprüft werden soll, so werden die ausgelesenen Farbwerte in HSV-Werte umgewandelt. Die einzelnen Werte werden dann daraufhin kontrolliert, ob H und S in den oben genannten Bereichen liegen. Ist dem so, dann werden sie als Haut erkannt, andernfalls als Nicht-Haut. Da die Ausgabe in ein schwarz-weiß Bild geleitet wird, ist keine Transformation von HSV in RGB nötig. Die RGB-Werte für schwarz und weiß werden je nach Resultat fest angegeben.

4.5 Hilfsmittel für die Auswertung

4.5.1 Konverter für Bitmap-Dateien

Um Bilder auszulesen und veränderte Farbwerte wieder in ein Bild umzuwandeln, war es nötig, einen Konverter zu schreiben. Dieser Konverter kann ausschließlich Bitmap-Dateien (.bmp) ein- und auslesen. Dieses Bildformat ist einfacher als andere zu lesen, da es nicht komprimiert ist. Andere Formate, wie jpg oder gif benutzen

Nr.	Bedeutung	Byte
1.	Zeichen 'B','M'	2
2.	Dateigröße	4
3.	Reserviert	4
4.	Offset (Adressabstand)	4
5.	Bild Kopfgröße	4
6.	Bildbreite	4
7.	Bildhöhe	4
8.	Anzahl der Ebenen = 1	2
9.	Kompression = 0	4
10.	Bildgröße in Byte	4
11.	Pixel per Meter x	4
12.	Pixel per Meter y	4
13.	Anzahl der benutzten Farben	4
14.	Anzahl der als wichtig befundenen Farben	4
15.	Einzelnen Farben jeweils in Reihenfolge bb gg rr	3

Tabelle 4.3: Struktur einer Bitmap-Datei

Komprimierungsalgorithmen, um die Größe der Datei zu verringern. Daher ist es nicht so leicht, Bilder in diesen Formaten auszulesen. Alle Bilder, die hier bearbeitet werden, wurden also zunächst in Bitmap-Dateien umgewandelt.

Struktur einer .bmp-Datei

Eine Bitmap-Datei ist in hexadezimaler Darstellung vorhanden. Dabei sind die einzelnen Optionen in 2 oder 4 Byte großen Integer dargestellt. Diese lassen sich umrechnen, um die Daten in der Dezimaldarstellung deuten zu können. Der Aufbau einer Bitmap-Datei ist in Abbildung 4.3 zu erkennen.

Konverter

Zunächst sollen die einzelnen RGB-Werte aus einem Bild extrahiert werden. Dazu werden alle in Tabelle 4.3 angegebenen Optionen mit `getLong` für 4 Byte und `getShort` für 2 Byte ausgelesen. Sind all diese Daten vorhanden, können die Pixel jeweils mit dreimal `getBytes` für blau, grün und rot ausgelesen werden. Es werden Breite multipliziert mit Höhe des Bildes viele Farbdaten ausgelesen. Wenn die Farbdaten jetzt vorliegen, werden sie je nach Verwendungszweck weiterverarbeitet. Entweder werden sie in eine Datei geschrieben, oder sie werden direkt

weiterverarbeitet. Beim Schreiben in eine Datei endet die Arbeit des Konverters hier, die Daten werden in einer Datei abgelegt und später anderweitig benutzt.

Bei der direkten Weiterverarbeitung werden die RGB-Werte, die aus dem Bild ausgelesen wurden, einer Bearbeitung unterzogen. Normalerweise in der Form, daß eine Methode für Haut-Erkennung angewandt wird. Die RGB-Werte werden untersucht und als Haut oder Nicht-Haut klassifiziert. Für jeden Punkt wird eins der Ergebnisse schwarz oder weiß in ein neues Ergebnisbild eingefügt. Optionen, wie Höhe und Breite, die das Ursprungsbild beinhaltet, können in das Resultatbild übernommen werden, lediglich die Farben sind schwarz oder weiß. Das neue Bild wird also wie in Tabelle 4.3 mit Werten und zum Schluß mit Farbwerten für schwarz oder weiss gefüllt. Dann wird es unter .bmp gespeichert und es wurde ein Bild im Bitmap Format geschaffen.

Kapitel 5

Auswertung

Zur Auswertung der einzelnen Methoden wurde ein Grundstamm von 70 verschiedenen Bildern angelegt. Diese Bilder sind vom Schwierigkeitsgrad unterschiedlich gut zu erkennen. Es gibt zum Beispiel Gesichter, die auf einem einheitlich schwarzen Hintergrund aufgenommen sind, bei denen es leichter sein sollte, die Haut zu erkennen, und dann gibt es Bilder mit Menschen und Holz oder Sand im Hintergrund, die ein größeres Problem darstellen. Von der Gesamtheit der Bilder wurden Optimalresultate erstellt. Durch Selektion wurde per Hand entschieden, welche Bereiche als Haut erkannt werden sollen. Die entstandenen Bilder sind die Grundlage der Auswertung.

Jedes aus einer der Methoden entstandene Bild wird mit seinem Optimalergebnis verglichen und danach festgestellt, wieviel Haut erkannt worden ist und wieviele Nicht-Hauptpunkte fälschlich als Haut identifiziert wurden. Zunächst wurde angenommen, daß es reicht zu wissen, wieviel Prozent der Haut eines Bildes erkannt worden sind, dem ist aber nicht so. So gibt es zum Beispiel Fälle, in denen fast das gesamte Bild weiß ist. Dabei wurde dann alles als Haut erkannt und demnach auch die gesamte Haut erkannt. Isoliert betrachtet ist die Hautfarben Erkennung kein guter Indikator dafür, wie gut Haut erkannt wird. Aus diesem Grund ist die Erkennung der Nicht-Hautfarbe als Attribut hinzugekommen. Es sind erst die Verfahren wirklich gut, die beide Seiten gut erkennen und deshalb wurden beide Prozentsätze gewählt und durch zwei dividiert. Da einige Verfahren bei der Nicht-Hautfarben-Erkennung sehr viel mehr Nicht-Hautfarben erkennen und damit auch viel Hautfarbe, liegen diese Prozentzahlen der erkannten Nicht-Hautfarben über 90%. Das führt zu einer Verfälschung der Ergebnisse. Aus diesem Grund wurde ein weiterer Indikator für die Güte der Verfahren hinzugenommen. Dieser Indikator ist der, wie gut die Bilder erkannt werden. Ein Bild gilt als gut erkannt, wenn beide Werte, der für Hautfarbe erkannt und der für Nicht-Hautfarbe erkannt, über 50% liegen. Eine jeweilige Endbewertung setzt sich also aus der Summe der Rangliste für den Durchschnittswert von erkannter Haut und erkannter Nicht-Haut und der Rangliste für die Anzahl der am besten erkannten Bilder zusammen. Dieser Wert wird durch



Tabelle 5.1: Originalbild

2 dividiert und das Ergebnis in die Endrangfolge gebracht. Zum Abschluß einer jeden Methode wird diese Auswertung durchgeführt.

Die Auswertungsdaten für alle Verfahren sind in Anhang A und Diagramme zur Verdeutlichung sind in Anhang B zu finden. Ein ausgewähltes Originalbild ist in Abbildung 5.1 zu sehen. Dieses Bild wird in den einzelnen Abschnitten mit entsprechendem Resultat der Methode wieder auftauchen.

5.1 Selektion

5.1.1 11 aus 256

Bei der Selektion aus 256 werden genau die 11 Farben selektiert, die vorher als Hautfarbe definiert wurden. Bei diesen Farben kommt es nur vereinzelt vor, daß eine Nicht-Hautfarbe als solches eingeordnet wird. Das passiert nur in den Fällen, in denen eine Nicht-Hautfarbe bei der vorherigen Transformation auf eine der 11 Farben abgebildet wurde. Da deshalb aber nur sehr wenige Nicht-Hautfarben falsch klassifiziert werden, sind die Prozentzahlen für erkannte Nicht-Hautfarbe sehr groß. In fast allen Fällen liegen sie zwischen 90 und 100%.

Die Hautfarbe an sich wird jedoch von dieser Methode am schlechtesten erkannt. Im Durchschnitt erreicht sie nur 27%. Dabei spielt natürlich die Transformation vom 16 Millionen Farbraum in 256 Farben eine große Rolle. Bei der Transformation aus 256 Farben können eine Reihe von Bildinformationen verloren gehen.



Tabelle 5.2: Optimales Ergebnis und Ergebnis der Methode aus 256 Farben

Viele Farben, die vorher nicht in den Hautfarbenbereich gefallen wären, werden nun auf eine der 11 gewählten Hautfarben abgebildet, da nur noch 256 Farben zur Verfügung stehen. Außerdem werden Farben, die bei 16 Millionen noch im Hautfarbenbereich lagen, auf eine Farbe außerhalb des Bereiches abgebildet.

Insgesamt werden von dem Verfahren nur 6 Bilder aus 70 gut erkannt, d.h. daß die Prozentzahlen für erkannte Haut und Nicht-Haut für diese Bilder jeweils über 50% liegen.

Daß die Selektion 11 aus 256 in der Durchschnittswertung auf Platz 3 liegt, läßt sich dadurch erklären, daß die "Nicht-Haut erkannt" Prozentzahlen stets sehr hoch sind. Sie heben das Gesamtergebnis an. Man muß die tatsächlich gut erkannten Bilder aber in die Gesamtbewertung mit einfließen lassen und kommt dann zu einem Gesamtergebnis, in dem die Selektion 11 aus 256 Platz fünf von sieben belegt.

Als Resultat kann weiter festgehalten werden, daß die gut erkannten Bilder dunkle Hintergründe haben und kaum über spiegelnde Flächen verfügen. Man kann von sehr einfachen Bildern sprechen, die hier erfolgreich sind.

5.1.2 8 aus 16 Millionen

Bei der Selektion von 8 Farben aus 16 Millionen werden mehr als diese 8 Farben selektiert. Die Anzahl von 8 Farben empfindet man als zu wenig, um in einem 16 Millionen Farbbereich ein Ergebnis zu erzielen. Wenn nur diese wenigen betrachtet würden, wären die Ergebnisse sicherlich sehr schlecht. Von Gimp, dem

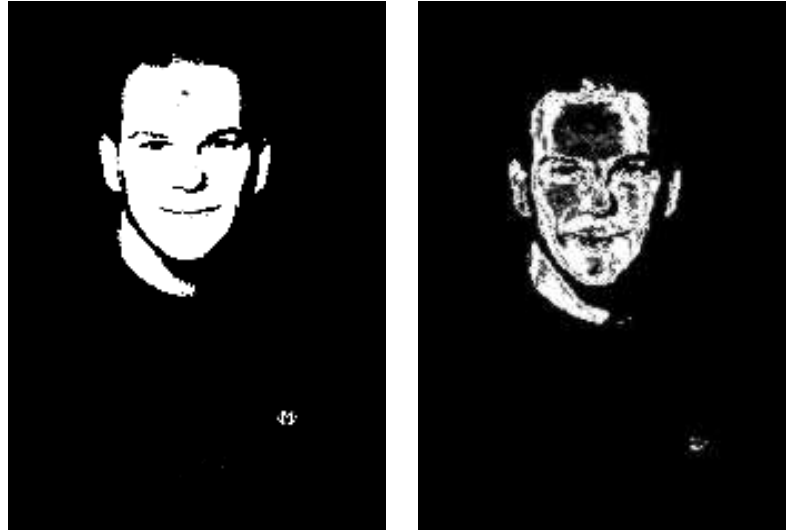


Tabelle 5.3: Optimales Ergebnis und Ergebnis der Methode aus 16 Millionen

Programm in dem diese Funktion umgesetzt wurde, wird intern eine Funktion benutzt, die ähnliche Farben mitschleift. Das ist grundsätzlich sehr gut, allerdings ist das Negative daran, daß man nicht sagen kann, wie diese Funktion arbeitet. Ob sie versucht in der Umgebung ähnliche Farben zu finden, oder das Spektrum der ausgewählten Farben vergrößert, ist unbekannt.

Die Hautfarbe wird von dieser Methode etwas besser erkannt, als von 11 aus 256. Das ist damit zu erklären, daß keine Transformation durchgeführt wird und damit auf diesem Wege keine Bildinformationen verloren gehen. Daß die Erkennung der Hautfarbe mit 31% nicht der herausragende Erfolg ist, ist mit der Aussage zu begründen, daß mit 8 Farben lange nicht alle Hautbereiche abgedeckt werden können, selbst wenn ähnliche Farben auch als Hautfarbe identifiziert werden.

Die Methode "aus 16 Millionen" belegt, wenn man nur den Durchschnittswert erkannter und nicht erkannter Hautfarbe betrachtet, Platz 2. Allerdings wird dieses Ergebnis auch, wie bei der vorherigen Selektion, dadurch angehoben, daß wenige Nicht-Hautfarben falsch identifiziert wurden. Untersucht man die gut erkannten Bilder, so sind sie mit einer Anzahl von 9 eher eine Enttäuschung und diese liegt in der Reihenfolge der am besten erkannten Bilder auf Platz 5. Deshalb sinkt die Methode in der Abschlußbewertung auf Platz 3 ab.

Sehr gut werden hier Bilder erkannt, die von einfacher Natur sind, klare Strukturen und einen dunklen Hintergrund aufweisen.

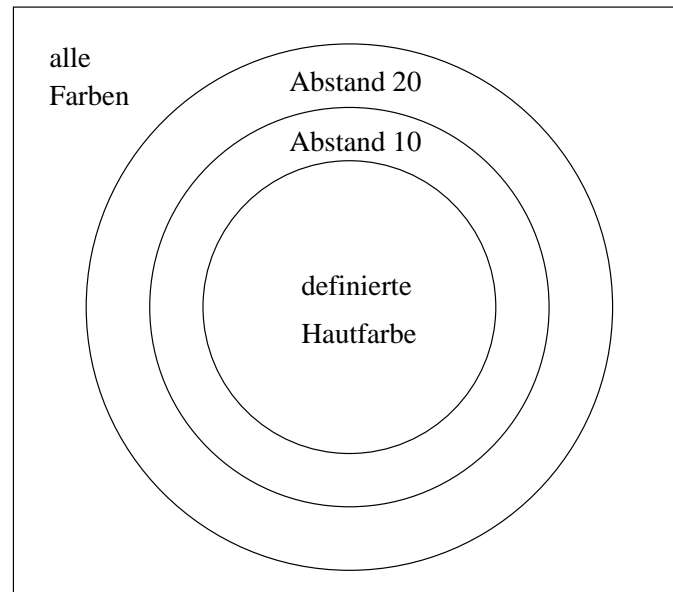


Abbildung 5.1: Problematik der Schwellwertsuche

5.2 Lineare Klassifikation

5.2.1 Schwellwertsuche

In dem Kapitel über die Umsetzung ist die Funktionsweise des Verfahrens bereits beschrieben worden. Es wurde jedoch noch nicht näher auf den Schwellwert eingegangen. Um eine ungefähre Vorstellung zu haben, wie groß ein solcher Schwellwert überhaupt sein soll, wurden einige Tests durchgeführt. Als Ausgangspunkt dienen zwei Dateien. Die eine beinhaltet die Hautfarben, wie sie bei der Methode benutzt werden. Die andere Datei enthält eine Testmenge, in der Haut- und Nicht-Hautfarben unsortiert abgelegt sind, allerdings als Haut und Nicht-Haut durch 0/1 markiert. Die Testmenge wird immer wieder neu aus einer Gesamtmenge zusammengestellt. Als Ergebnisse werden erkannte Haut von Gesamthaut und erkannte Nicht-Haut von der Gesamtheit Nicht-Haut dokumentiert. An den Resultaten, welche Tabelle 5.4 zu entnehmen sind, kann man erkennen, daß bei größer werdendem Abstand mehr Hautfarbe erkannt wird und gleichzeitig die erkannten Nicht-Hautfarben sinken. Eine solche Reaktion war zu erwarten, da bei größer werdender Schwelle Ausreißer mit identifiziert werden, aber auch immer mehr Nicht-Hautpunkte als Haut zurückgegeben werden. Bei sehr großer Schwelle wird das gesamte Bild weiß sein und damit jede Hautfarbe erkannt. Allerdings wird in diesem Fall Null Prozent der Nicht-Hautfarbe erkannt. Es ist also wichtig ein ausgeglichenes Verhältnis zwischen Erkennen der Haut- und der Nicht-Hautfarbe zu finden. Wenn jetzt noch ein Blick auf Tabelle 5.4 geworfen wird, stellt man fest, daß

Schwelle	HAUT			NICHT - HAUT		
	gesamt	richtig	Prozent	gesamt	richtig	Prozent
5	210	192	91	237	230	97
10	210	202	96	237	215	90
20	210	207	99	237	170	71
30	210	210	100	237	131	55
40	210	210	100	237	107	45
5	196	168	86	251	247	98
10	196	174	89	251	234	93
20	196	191	97	251	184	73
30	196	196	100	251	139	55
40	196	196	100	251	113	45
5	212	183	86	235	230	97
10	212	193	91	235	208	88
20	212	207	98	235	162	68
30	212	212	100	235	134	57
40	212	212	100	235	92	39

Tabelle 5.4: Ergebnisse zur Schwellwertsuche

bei einer Schwelle von 30 Punkten in unserem Fall immer 100 Prozent der Haut erkannt wird. Das ist zwar gut, aber nach der obigen Argumentation wird dabei wohl schon zu viel Nicht-Haut als Haut erkannt. Denn die Erkennung der Nicht-Haut sinkt beim Schwellwert von 30 schon rapide. Bei einem Abstand von 20 Punkten liefern beide Werte ganz gute Ergebnisse, man liegt bei der Haut-Erkennung noch über der Nicht-Haut-Erkennung, aber diese liegt immernoch klar über 50 Prozent.

Nach dieser Vorüberlegung wird 20 als gute Schwelle gewählt, und um die Auswertung zu überprüfen, werden außerdem die Schwellen 10 und 30 angewendet.

5.2.2 Ergebnisse

Bei der linearen Klassifikation wird versucht, den Bereich der Hautfarbe zu umschreiben. Durch die verschiedenen Abstände sollen noch nicht definierte Hautpunkte mit eingeschlossen werden. Dabei aber genau den besten Abstand zu finden, wie oben beschrieben, ist nicht so einfach. Um auch mit einer kleinen Schwelle möglichst gute Ergebnisse erzielen zu können, wurden als Hautfarben grundsätzlich sehr unterschiedliche Hautfarben ausgewählt.

Die Nachteile des Verfahrens sind:

- Es müssen sehr viele Vergleiche durchgeführt werden, da jeder Pixel mit

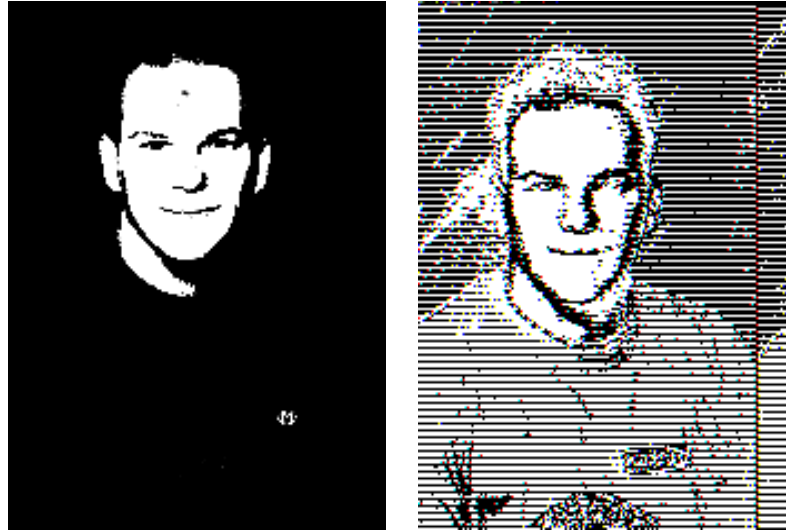


Tabelle 5.5: Optimales Ergebnis und Ergebnis der linearen Klassifikation mit Abstand 20

allen Hautfarben verglichen wird. Dieser Nachteil konnte nur durch die Verwendung der Octree-Datenstruktur wieder ausgeglichen werden.

- Durch die Definition von dunklen Hautfarben werden Farben mitselektiert, die nahe an schwarz grenzen, das kann Bilder mit helleren Hautfarben und dunklen Hintergründen stark verschlechtern, vor allem wenn der Abstand steigt, werden irgendwann schwarze Farben mit als Haut identifiziert. Die Resultate für die unterschiedlichen Abstände müssen gesondert behandelt werden.

Abstand 10: Bei einem Abstand von 10 werden nur 28% der Haut erkannt und 78% der Nicht-Haut erkannt. Das ist insgesamt für diesen Teil der zweitletzte Platz. Da hinzukommt, daß nur 2 Bilder gut erkannt wurden, belegt diese Methode für den Abstand 10 bei der Endwertung den letzten, also den 8ten Platz. Wenn der Abstand so gering ist, werden viele den Hautfarben ähnliche Farben noch nicht als Hautfarbe identifiziert. Der Anteil der nicht als Hautfarben erkannten Bereiche ist noch sehr groß, da viele Farben außerhalb der Hautfarbe nicht als Haut identifiziert werden, was auch so vorgesehen ist.

Abstand 20: Ein Abstand von 20 bringt bei der Hautfarben-Erkennung eine Prozentzahl von 54, also deutlich mehr als bei Abstand 10. Bei der Nicht-Hautfarben-Erkennung sinkt das Ergebnis auf 51%. Das war zu erwarten, da die Anteile an Nicht-Haut, die als Haut identifiziert werden, mit größerem Abstand zunehmen. Die Hautfarben, die als Ausreißer in dem Gebiet um



Tabelle 5.6: Ergebnisse für Abstand 10, 20 und 30 im direkten Vergleich

Abstand 20 liegen, können jetzt natürlich erkannt werden und dementsprechend steigt der Anteil an erkannter Haut. Da 15 Bilder gut erkannt wurden, liegt die Klassifikation mit Abstand 20 am Ende auf Platz 4.

Abstand 30: Liegt der Abstand bei 30, so entwickeln sich die Ergebnisse gleichbleibend. Die Prozentzahl der identifizierten Haut steigt weiter an und die der identifizierten Nicht-Haut wird weniger. Mit 72% erkannter Haut und 31% erkannter Nicht-Haut liegt die Methode dann auch auf dem letzten Platz. Auf den Bildern werden viel zu viele Farben der Hautfarbe zugesprochen, natürlich sind dabei noch einige, die auch Haut sind und dadurch richtig klassifiziert werden. Allerdings werden sehr viele Nicht-Hautfarben als solche klassifiziert und deshalb wird das Verfahren schlechter. Da es aber 14 Bilder gut erkennt, rettet es sich in der Gesamtwertung auf den zweitletzten, sechsten Platz.

An den drei gut erkannten Bildern in Abbildung 5.6 kann man den Verlauf der linearen Klassifikation sehr gut sehen. Das linke Bild ist eine lineare Klassifikation mit Abstand 10, das mittlere mit Abstand 20 und das rechte hat einen Abstand von 30 Punkten. Im linken Bild ist noch sehr vorsichtig selektiert, nämlich nur die Farben, die als Hautfarbe definiert sind, plus einen Abstand von 10 in jede Richtung. Beim Bild in der Mitte wirkt es so, als seien die Flächen, die in Bild 1 von Haut umgeben waren, gefüllt worden. Das ist erklärbar, da die Farben, die sich in der von Haut umgebenen Fläche befanden, ähnliche Farben besitzen. Bei einem Abstand von 20 zu den definierten Hautfarben wurden diese mitselektiert. Das Bild auf der rechten Seite hat einen sehr großen Teil des Bildes selektiert. Das ist ein Zeichen dafür, daß mit einem Abstand von 30 Punkten zu viele Farben als Hautfarbe erkannt werden. Hier sieht man auch die Schwellwertsuche aus 5.2.1 bestätigt, daß der Abstand von 20 Punkten gute Ergebnisse liefert.

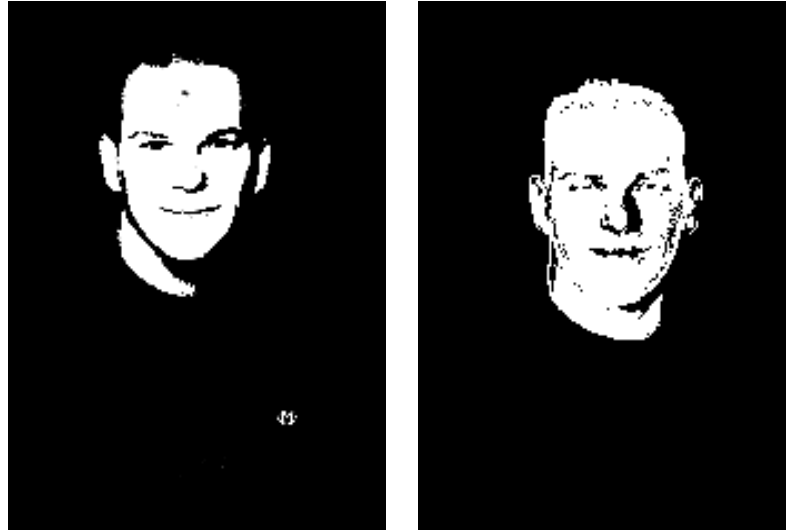


Tabelle 5.7: Optimales Ergebnis und Ergebnis aus chromatischem Farbraum

5.3 Regression und Neuronale Netze

Bei der Regression ebenso wie bei den Neuronalen Netzen sind die Vorüberlegungen sehr gut gewesen, die beiden Verfahren zu benutzen. Die Regressionsrechnung findet aber keine lineare Gleichung, die den Hautfarbenbereich beschreiben kann. Das liegt an der Tatsache, daß der Hautfarbenbereich nicht, wie angenommen, einen zusammenhängenden Raum beschreibt. Im RGB-Farbbereich beschreiben die Hautfarben eine Punktwolke im Raum. Diese ist nicht vollständig zusammenhängend, weshalb es der Regression nicht gelingt den Raum zu beschreiben. Im Hinblick auf die sehr schlechten Ergebnisse in der Vorbereitung zu diesen Methoden, sind sie nicht durchgeführt worden. Es existieren lediglich die Resultate aus den Vorbereitungen. Diese bestehen aus Lern- und Testmengen, die untersucht worden sind. Demnach ist es für den linearen Fall nicht möglich, gute Ergebnisse zu finden und die beiden Verfahren wurden ausgeschlossen.

5.4 Chromatischer Farbraum

Die Haut-Erkennung aus der Gesichtserkennung liefert die besten Ergebnisse. Da die Farbwerte stets in den chromatischen Farbraum umgerechnet werden, spielt bei dieser Methode die Helligkeit keine Rolle. Aus diesem Grund können viele Bilder, die schwierig sind, gut erkannt werden. Die einzige Problematik, die dieses System hat, ist die Auswahl der Hautbereiche, mit denen gelernt wird. Diese Bereiche müssen mit sehr viel Sorgfalt ausgewählt werden. Nach Rücksprache mit

aus 16 Mio.	Abst.30	Ergebnis
Haut	Haut	Haut
Haut	Nicht-Haut	Haut
Nicht-Haut	Haut	Haut
Nicht-Haut	Nicht-Haut	Nicht-Haut

Tabelle 5.8: Oder Verknüpfung bei Kopplung der Ergebnisse

den Autoren dieses Systems, kann man sagen, daß besonders drei Aspekte für die Wahl der Lernhautteile wichtig sind.

- Die "Farbteile" sollen die gleiche Größe haben.
- Die Farben sollen sehr unterschiedlich sein, um Haut verschiedener Rassen von Menschen zu finden.
- Es sollten möglichst gleiche Anteile von heller, dunkler, rötlich und gelber Hautfarbe vorhanden sein.

Mit der Güte der Lernmenge steigt die Güte des gesamten Verfahrens. Bei einer schlechten Lernmenge ergeben sich bei der Anwendung auch schlechte Ergebnisse. Die Lernmenge, die hier benutzt wurde, besteht aus 9 Hautteilen, die jeweils eine Größe von 20x20 Pixel haben. Das ergibt eine Lernmenge von 3600 Punkten. Wie die Ergebnisse zeigen, ist die Lernmenge gut gewählt.

Im chromatischen Farbraum zeigt sich ein gutes Verhältnis zwischen Haut- und Nicht-Haut-Erkennung. Haut wurde zu 52% und Nicht-Haut zu 84% erkannt. Der Durchschnittswert dieser beiden ist 68% und damit das beste Ergebnis. Beim Kriterium der möglichst gut erkannten Bilder liegt die Haut-Erkennung aus der Gesichtserkennung mit 37 Bildern an der absoluten Spitzenposition. In der Gesamtauswertung ist dieser Methode der erste Platz sicher.

5.5 Kopplung von zwei Verfahren

Um die Auswertung etwas kreativer zu gestalten, wurden die beiden Verfahren Selektion aus 16 Millionen Farben und lineare Klassifikation mit Abstand 30 ODER verknüpft. Bei der Auswertung wurden die jeweiligen Ergebnisbilder der beiden Verfahren mit dem Optimalbild verglichen. Die Resultate für die gekoppelte Auswertung setzen sich wie in Tabelle 5.8 gezeigt zusammen.

Wenn also eines der Verfahren einen Punkt als Haut identifiziert hat, wird er auch im Ergebnis der Kopplung Haut. Nur wenn beide Verfahren Nicht-Haut identifizieren ist das Ergebnis Nicht-Haut. 78% der Haut werden erkannt und 31% der

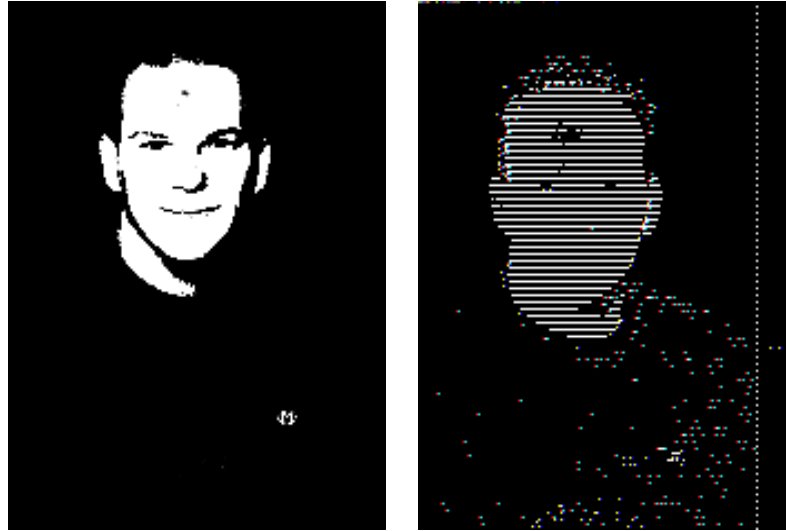


Tabelle 5.9: Optimales Ergebnis und Ergebnis aus dem HSV-Farbsystem

Nicht-Hautfarben. Das Ergebnis der "Haut erkannt Auswertung" liegt damit zwischen den beiden Verfahren, aus denen es entstanden ist. Es belegt Platz 5, wobei aus 16 Mio. auf Platz 2 und Abstand 30 auf Platz 7 liegt. Bei der Anzahl an gut erkannten Bildern steht die Kopplung beiden Methoden nach, indem nur 8 Bilder gut erkannt werden. Bei aus 16 Mio. sind es 9 Bilder und bei einem Abstand von 30 werden 14 Bilder gut erkannt. Insgesamt besetzt die Kopplung dann mit Abstand 30 zusammen den sechsten Platz. Die Methode wird also nicht besser als eine der Teilmethoden. Erstaunlich ist, daß nur eins der 8 Bilder, die gut erkannt werden, von beiden Teilmethoden gut erkannt wurden. 5 Bilder wurden nur von einer Methode gut erkannt und 2 wurden von keiner der Teilmethoden gut erkannt, sondern nur von der Kopplung. Das heißt, die Kopplung erkennt nicht automatisch die Bilder gut, die von einer der Teilmethoden gut erkannt wurden.

Allgemein liegt die Kopplung auf Platz 6. Im Anhang A kann man ein Diagramm finden, das die beiden Teilmethoden und die Kopplung gegenüberstellt.

5.6 Das HSV-Farbsystem

Das HSV-Farbsystem hat für den Zweck der Haut-Erkennung den Vorteil, daß es die Sicht des Menschen auf Farben widerspiegelt. Dieses intuitive System kann Hautfarbe in einem Bereich festlegen. Lediglich durch das Abprüfen des Farbtons und der Sättigung kann entschieden werden, ob die Farbe als Hautfarbe identifiziert wird. Helligkeit spielt keine Rolle, sie kann beliebige Werte annehmen.

Da fast alle Bilder das RGB-Farbsystem benutzen, ist eine Umwandlung von jedem Punkt in das HSV System nötig. Das ist der Nachteil, den dieses System liefert.

Mit der Anwendung des Systems erreicht man 36% der Haut, die erkannt wird und 81% der Nicht-Haut. Damit sind 59% ein vierter Platz. Das Verfahren erkennt, mit einer Anzahl von 21, am zweitbesten Bilder. Insgesamt ergibt das den 2ten Platz. Es werden ganz unterschiedliche Bilder gut erkannt. Sowohl dunkle Hintergründe und klare Strukturen, als auch unruhige Bilder mit schweren Übergängen. Verschiedene Beleuchtung spielt keine Rolle, da sie durch ledigliche Angabe des Farbtons und der Sättigung eliminiert wurden.

5.7 Diskussion

Methode	durchschn. erkannte Haut	durchschn. erkannte Nicht-Haut	Gesamt- durch- schnitt	# gut erkannter Bilder	Gesamt- be- wertung
Chromatisch	52%	84%	68% (1)	37 (1)	1
HSV	36%	81%	59% (4)	21 (2)	2
aus 16 Mio.	31%	95%	63% (2)	9 (5)	3
Abstand 20	54%	51%	53% (6)	15 (3)	4
11 aus 256	27%	95%	61% (3)	6 (7)	5
Abstand 30	72%	31%	52% (7)	14 (4)	6
Kopplung	78%	31%	54% (5)	8 (6)	6
Abstand 10	28%	78%	53% (6)	2 (8)	7

Tabelle 5.10: Übersicht gesammelter Endergebnisse

Die angegebene Endergebnistabelle 5.10 ist so aufgebaut, wie die Auswertungsdaten ausgelesen wurden. In der ersten Spalte sind die verschiedenen Methoden aufgetragen. In der zweiten die durchschnittlich erkannte Haut in Prozent. Dieser Wert setzt sich zusammen aus dem Produkt der Prozentzahlen der erkannten Haut von allen überprüften Bildern geteilt durch die Anzahl der Bilder. Die Einzelwerte können Anhang B entnommen werden. Die durchschnittlich erkannte Nicht-Haut ist von der Rechnung gleich aufgebaut und repräsentiert die Nicht-Hautfarben, die als solche erkannt wurden. Ein Durchschnitt dieser beiden Werte ist dann in Spalte vier zu finden. Dieser Durchschnitt über erkannte Farben wurde in eine Reihenfolge gebracht. Der höchste Wert wird auf Position 1 gesetzt und die anderen der Reihe nach darunter sortiert. Diese Reihenfolge reicht als Endergebnis nicht aus. Bei manchen Verfahren, die hier sehr gut abgeschnitten haben, stellte sich heraus, daß nur wenige Bilder von diesen Verfahren gut erkannt wurden. Als gut erkannt ist ein Bild definiert, wenn die beiden Werte von Haut und Nicht-Haut erkannt über 50% liegen. Die Anzahl der Bilder, die von dem jeweiligen Verfahren gut erkannt

wurden, ist als Auswertungsattribut hinzugekommen. Bei diesem Wert ist die Anzahl der gut erkannten Bilder angegeben worden. Auch hier hat der größte Wert in der Reihenfolge die Position 1 bekommen und alle anderen wurden darunter sortiert. Aus diesen beiden Auswertungsreihenfolgen, erkannte Haut und gut erkannte Bilder wurde eine Gesamtauswertung erstellt, die zeilenweise das Produkt aus den Reihenfolgen bildet und es durch zwei dividiert. Diese Ergebnisse wurden dann in eine Endreihenfolge gebracht, die in der letzten Spalte platziert wurde.

Jetzt wird noch auf die Beziehung unter den einzelnen Methoden eingegangen. Dabei wird an der untersten Stelle, der Klassifikation mit Abstand 10, begonnen.

Abstand 10 - Kopplung Die beiden "letzten" Verfahren sind etwas schwer zu vergleichen, da die Kopplung ja zum Teil aus einem ähnlichen Verfahren besteht. Dieser Tatsache ist auch zu verdanken, daß die Argumentation ähnlich ist, wie beim Vergleich der Einzelklassifikationen. Die Kopplung übernimmt ja nicht nur die Ergebnisse der Verfahren aus 16 Mio., sondern auch die von Abstand 30. Die Kopplung wird also auch nicht schlechter als eines der Teilverfahren. Damit ist es besser als die Klassifikation mit Abstand 10, denn bei dieser Methode werden zu wenige Hautfarben identifiziert. Im Gegensatz dazu werden bei einem Abstand von 30 zu viele Hautfarben und zu wenige Nicht-Hautfarben als solche identifiziert. Die beiden Methoden kehren sich gegenseitig genau um. Bei der Kopplung kann man das genauso sehen. Die Ergebnisse, die die Klassifikation mit Abstand 10 als schlecht liefert, nämlich wenig erkannte Hautfarbe und die es als gut liefert, gut erkannte Nicht-Hautfarbe, kehrt die Kopplung um und erkennt Hautfarbe gut und Nicht-Hautfarbe nicht gut. Beide erkennen recht wenige Bilder wirklich gut und landen deshalb auf den untersten Plätzen.

Kopplung - Abstand 30 Die Kopplung muß mit der Klassifikation mit Abstand 30 nicht verglichen werden, da die eine Methode die andere beinhaltet. Die Ergebnisse sind sich sehr ähnlich und sogar die erkannten Bilder stimmen in sechs Bildern überein.

Abstand 30 - 11 aus 256 Die Klassifikation mit Abstand 30 unterscheidet sich schon mehr zum Verfahren 11 aus 256. 11 aus 256 bietet ähnliche Ergebnisse wie die Klassifikation mit Abstand 10. Auch hier werden Hautfarben schlechter erkannt, dagegen Nicht-Hautfarben im Bereich um 90%. Dies läßt einen sehr großen Durchschnitt entstehen. Nur aus diesem Grund liegt 11 aus 256 vor Abstand 30, denn die Anzahl der gut erkannten Bilder beträgt nur 6. Bei der Methode 11 aus 256 werden, anders als bei Abstand 30, die Nicht-Hautfarben fast gar nicht mit eingeschlossen. Es kann höchstens bei der Transformation passieren, daß einige Farben falsch in den 256 Farbbereich abgebildet werden. Ansonsten wird eine Farbe, die nicht zu den 11 Hautfarben gehört, auch nicht als Hautfarbe identifiziert. Bei der Methode Abstand 30 sieht das anders aus. Der Abstand von 30 in alle drei, R-, G- und

B-Richtung schließt auch Nicht-Hautfarben mit ein, die als Hautfarbe identifiziert werden. Das ist der Grund dafür, daß 11 aus 256 insgesamt besser abschneidet, als Abstand 30.

11 aus 256 - Abstand 20 Bei der als nächsthöher liegenden Methode von Abstand 20 herrscht ein sehr ausgeglichenes Verhältnis. Sowohl Haut als auch Nicht-Haut werden gleichermaßen gut erkannt. Die Haut wird etwas weniger gut erkannt als noch bei Abstand 30, dafür wird die Nicht-Haut besser erkannt und beides ist dann eher ausgeglichen. Im Vergleich zum Verfahren 11 aus 256 erkennt es Haut besser, da es mehr Farben einschließt und es erkennt Nicht-Hautfarben schlechter, da diese teilweise als Hautfarbe mit eingeschlossen werden. Da von Abstand 20 aber mehr als die doppelte Anzahl von Bildern gut erkannt wird, ist die Methode in der Gesamtbewertung über 11 aus 256 anzusiedeln.

Abstand 20 - 16 Mio. Das Verfahren der Selektion aus 16 Millionen Farben wird, wie sein Vorgänger, die Selektion 11 aus 256, durch die Erkennung der Nicht-Hautfarbe gerettet. Da Farben, die nicht in den Bereich passen, automatisch keine Chance haben, als Hautfarbe identifiziert zu werden, werden die Nicht-Hautfarben gut erkannt. Das ist bei der darunterliegenden Methode von Abstand 20, wie oben beschrieben, etwas anders. Die Selektion aus 16 Millionen Farben erkennt wenige Bilder sehr gut. Aber aus der Tatsache heraus, daß die Nicht-Haut gut erkannt wird, besetzt es den dritten Platz.

16 Mio. - HSV Das HSV-System kommt dem Empfinden von Farben, wie der Mensch sie sieht, sehr nahe. Es erkennt im Durchschnitt zwar nur geringfügig mehr Haut als die Selektion aus 16 Mio. und bei der Nicht-Haut-Erkennung ist es sogar noch schlechter. Da von dem Verfahren, das auf HSV basiert aber 21 Bilder gut erkannt werden, ist es besser zu bewerten als aus 16 Millionen. Die Methoden sind schwieriger zu vergleichen, da sie auf unterschiedlichen Farbsystemen arbeiten. Es scheint so, als ob die Hautfarben einfach besser vom HSV-System erkannt werden. Es wird dabei nur der Farbton und die Sättigung überprüft und nicht die Helligkeit. Sie kann beliebige Werte annehmen. Bei der Methode aus 16 Mio. müssen schließlich alle Werte des RGB-Bereichs ungefähr übereinstimmen.

HSV - Chromatisch Die Gesichtserkennung basiert wie das HSV-System auf einer Transformation der Farbwerte, die selektiert werden sollen. Da die Verfahren der Gesichtserkennung wesentlich aufwendiger sind, als die des HSV-Systems, sind die Ergebnisse auch aussagekräftiger. Es erkennt in jeder Kategorie mehr als das HSV, mehr Haut wird erkannt, mehr Nicht-Haut wird erkannt und es werden mehr Bilder gut erkannt. Die Unterschiede in der Güte können natürlich zusätzlich mit der Wahl der Hautteile zusammenhängen, die zum Lernen in den beiden Systemen angegeben wurden. Die Hautteile sind in jedem System unterschiedlich.

5.8 Zusammenfassung und Ausblick

In dieser Arbeit werden einige Möglichkeiten untersucht, Hautfarbe in Farbbildern zu erkennen.

Dafür wurde zunächst eine Einordnung des Problems vorgenommen und der Gesamtablauf der Erkennung von pornografischen Bildern erklärt, in dem die Hautfarben-Erkennung einen Teil bildet.

Dann wurde erläutert, wie sich Hautfarbe zusammensetzt und welche Probleme bei der Wahl und Erkennung von Hautfarbe auftreten. Die Farbmodelle, die in dieser Arbeit verwendet wurden, werden erklärt und es wird begründet, warum ihre Benutzung sinnvoll ist.

Anschließend wurden die einzelnen Methoden, um die es hier geht, ganz allgemein erklärt. Basierend auf dem RGB-System sind das die Selektion und die lineare Klassifikation. Dann folgen die Regression und die Neuronalen Netze. Schließlich noch die beiden Verfahren, die Helligkeit außen vor lassen, der chromatische Farbraum und eine Methode, die auf dem HSV-System basiert.

Im darauffolgenden Kapitel wurde die Umsetzung der einzelnen Verfahren beschrieben, d.h., wie es von der Idee zu einem lauffähigen Programm gekommen ist. Die einzelnen Algorithmen werden erläutert und Hilfsmittel aufgeführt, die für die Bearbeitung der Bilder und die spätere Auswertung wichtig sind.

Schließlich wurden die Ergebnisse der einzelnen Methoden mit ihren Vor- und Nachteilen vorgestellt. In der anschließenden Diskussion wurden die Ergebnisse der verschiedenen Methoden verglichen und in eine Reihenfolge gebracht. Dabei stellte sich sehr klar heraus, daß die Methoden des chromatischen Farbraums und HSV, also jene in denen die Helligkeit im Bild vernachlässigt worden ist, die besten Ergebnisse produzieren.

Jetzt soll noch ein Ausblick darauf gegeben werden, wie dieses Projekt weitergeführt werden könnte.

Zunächst könnten natürlich noch eine Unzahl von Methoden zur Erkennung von Hautfarbe untersucht werden. Dabei sollte man den Ansatz von verschiedenen Farbsystemen weiterverfolgen, der hier in den Methoden von HSV und chromatischem Farbraum eingeführt wurde. Da gerade bei der Umrechnung von RGB in andere Systeme sehr gute Ergebnisse erzielt wurden, wäre das ein guter Ansatz, den man ausbauen sollte.

Um das Grundproblem, die Erkennung von pornografischen Bildern weiter voranzutreiben, sollten an die Hautfarben-Erkennung anschließend Strukturerkennungen durchgeführt werden. Dazu ist es, wie in [FFF] nötig, einzelne Glieder des Menschen zu erkennen und diese in den schwarz-weiß Bildern zusammenzufügen. Forschungen von M.Fleck und D.Forsyth haben auch schon versucht, wie in [FFI], nackte Menschen zu finden. Alternativ kann man natürlich auch mit einfacheren

Auswertungskriterien beginnen. Dazu bieten sich zum Beispiel statistische Methoden, wie Anteil von Hautfarbe im Verhältnis zur Gesamtanzahl von Punkten im Bild, an.

Diese Arbeit hat unterschiedliche Methoden zur Erkennung von Hautfarbe in Farbbildern aufgezeigt und analysiert. Durch die Erfahrungen, die im Zusammenhang mit den Methoden gemacht wurden, kann nach einer Haut-Erkennung durch ein Verfahren, das die Helligkeit vernachlässigt mit der Strukturerkennung begonnen werden.

Ein Teilschritt auf der Suche nach pornografischen Bildern wird durch diese Arbeit abgedeckt.

Anhang A

Diagramme

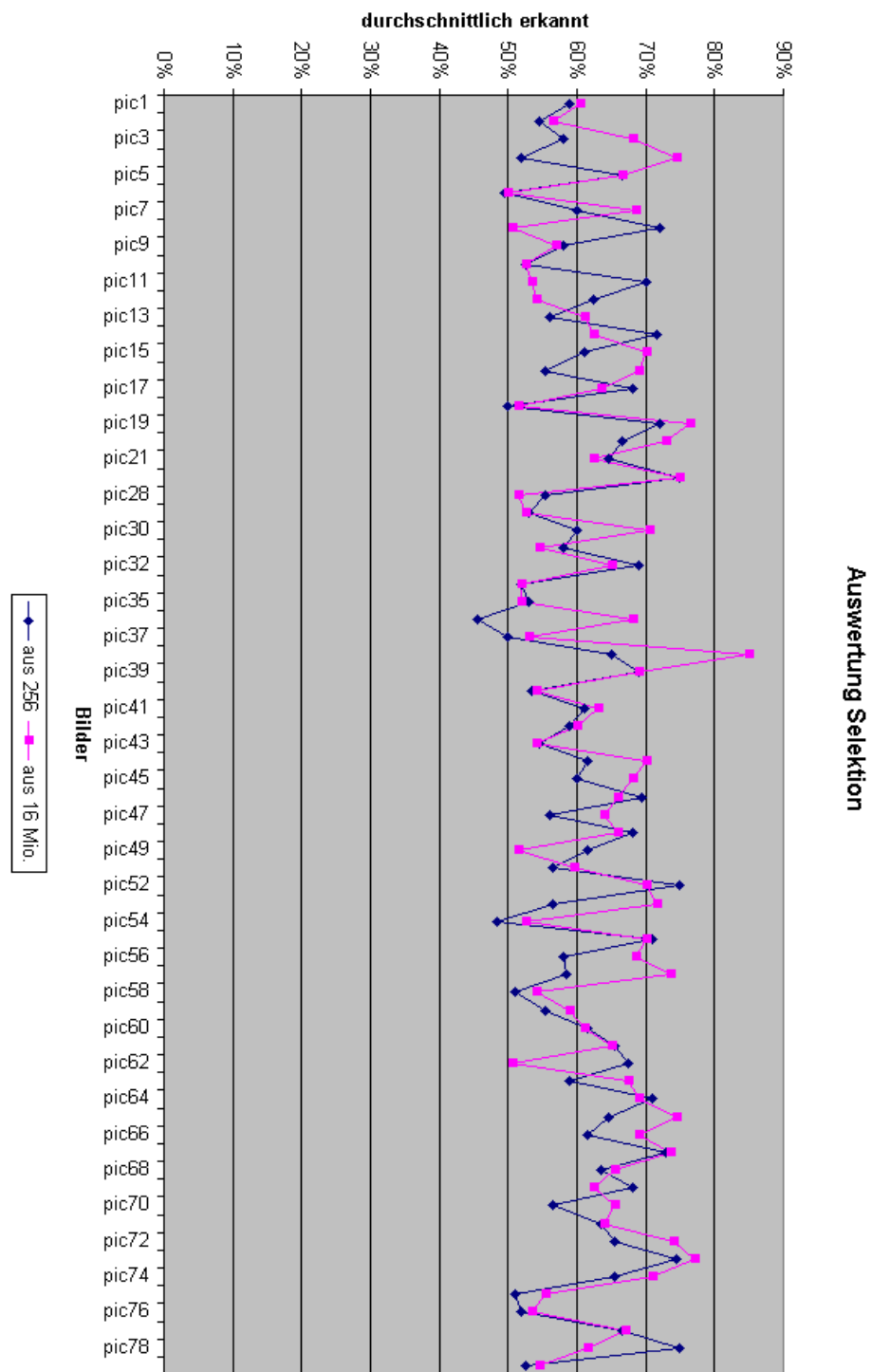


Abbildung A.1: Ergebnisse der verschiedenen Selektionen zusammengestellt

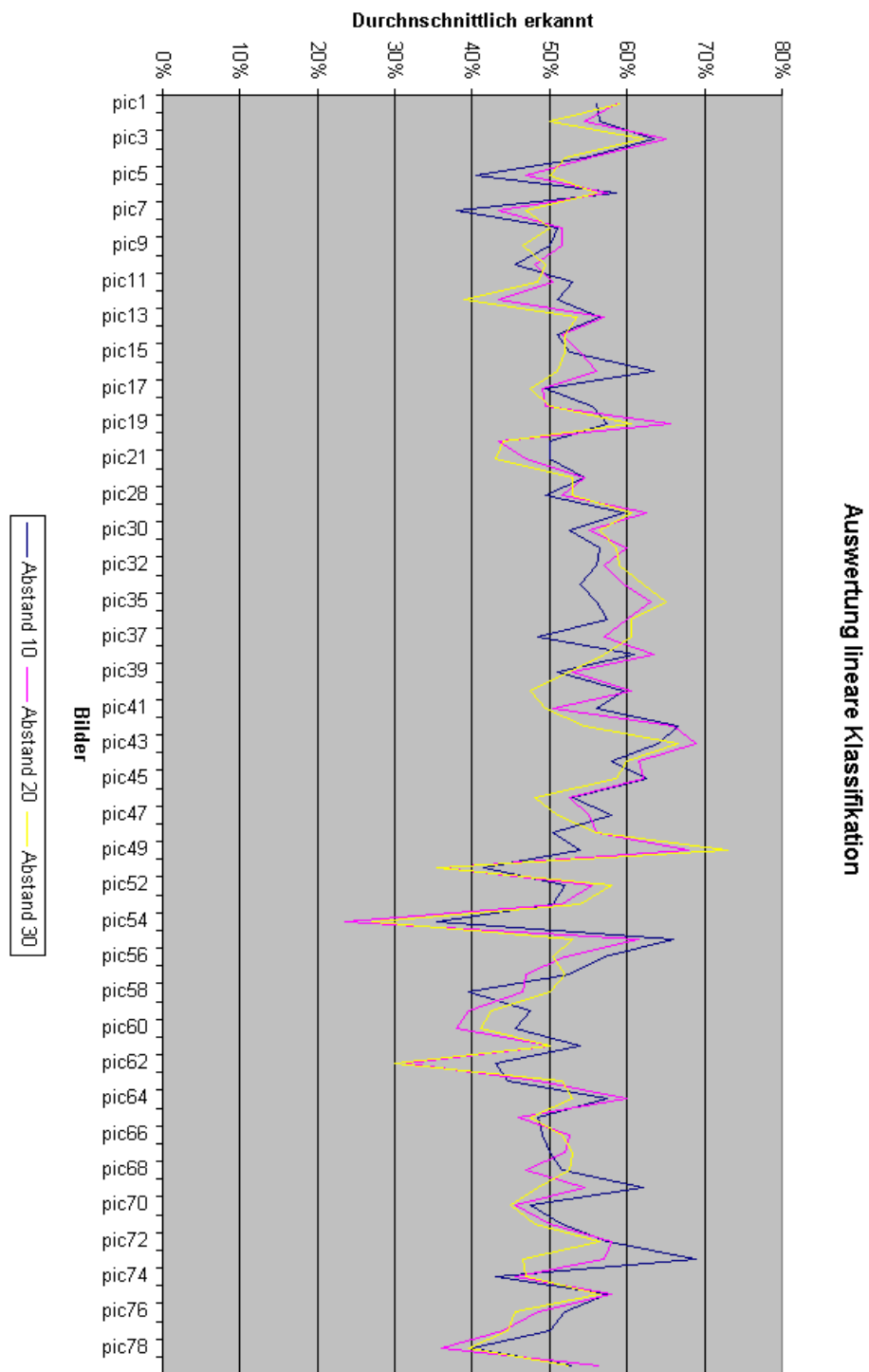


Abbildung A.2: Ergebnisse der unterschiedlichen Klassifikationen

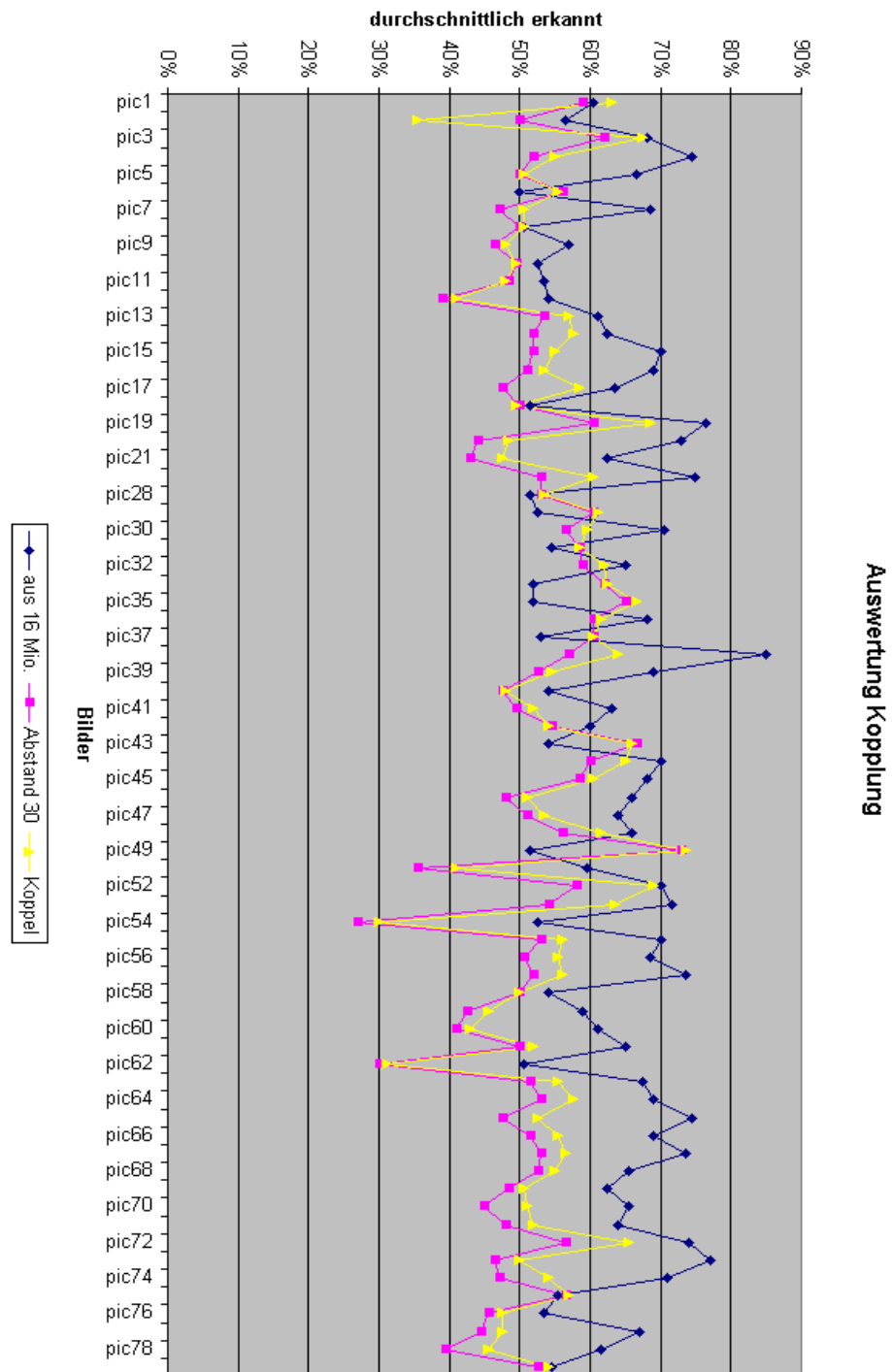


Abbildung A.3: Ergebnisse aus 16 Mio., Klassifikation mit Abstand 30 und Koppelung der Beiden

Anhang B

Ergebnisse zu einzelnen Auswertungen

Bilder	Selektion 11 aus 256			Selektion aus 16 Mio.		
	erkannt	n.erk.	Schnitt	erkannt	n.erk	Schnitt
pic1	23	95	59	22	99	61
pic2	12	97	55	16	97	57
pic3	21	95	58	37	99	68
pic4	8	96	52	55	94	75
pic5	37	96	67	40	93	67
pic6	7	92	50	5	95	50
pic7	21	99	60	42	95	69
pic8	50	94	72	1	100	51
pic9	18	98	58	22	92	57
pic10	6	99	53	7	98	53
pic11	49	91	70	16	91	54
pic12	28	97	63	10	98	54
pic13	25	87	56	27	95	61
pic14	56	87	72	32	93	63
pic15	26	96	61	43	97	70
pic16	18	93	56	44	94	69
pic17	40	96	68	33	94	64
pic18	1	99	50	5	98	52
pic19	44	100	72	54	99	77
pic20	33	100	67	47	99	73
pic21	31	98	65	26	99	63
pic27	52	98	75	53	97	75
pic28	17	94	56	5	98	52
pic29	6	100	53	6	99	53
pic30	21	99	60	43	98	71
pic31	17	99	58	11	98	55
pic32	48	90	69	46	84	65
pic34	6	98	52	5	99	52
pic35	8	98	53	6	98	52
pic36	7	84	46	52	84	68
pic37	13	87	50	9	97	53
pic38	33	97	65	76	94	85
pic39	41	97	69	40	98	69
pic40	8	99	54	9	99	54
pic41	23	99	61	28	98	63
pic42	24	94	59	28	92	60
pic43	15	94	55	13	95	54
pic44	32	91	62	47	93	70

Tabelle B.1: Ergebnisse für 11 aus 256 und aus 16 Mio. (1)
(alle Angaben in Prozent)

Bilder	Selektion 11 aus 256			Selektion aus 16 Mio.		
	erkannt	n.erk.	Schnitt	erkannt	n.erk	Schnitt
pic45	25	95	60	43	93	68
pic46	41	98	70	37	95	66
pic47	14	98	56	31	97	64
pic48	39	97	68	37	95	66
pic49	29	94	62	6	97	52
pic50	14	99	57	22	97	60
pic52	55	95	75	43	97	70
pic53	19	94	57	47	96	72
pic54	5	92	49	13	92	53
pic55	44	98	71	43	97	70
pic56	19	97	58	40	97	69
pic57	26	91	59	60	87	74
pic58	5	97	51	13	95	54
pic59	13	98	56	21	97	59
pic60	34	89	62	34	88	61
pic61	41	90	66	49	81	65
pic62	38	97	68	1	100	51
pic63	26	92	59	47	88	68
pic64	46	96	71	40	98	69
pic65	33	96	65	55	94	75
pic66	28	95	62	45	93	69
pic67	47	99	73	51	96	74
pic68	37	90	64	42	89	66
pic69	36	100	68	26	99	63
pic70	19	94	57	36	95	66
pic71	28	99	64	29	99	64
pic72	34	97	66	53	95	74
pic73	51	98	75	57	97	77
pic74	35	96	66	44	98	71
pic75	10	92	51	23	88	56
pic76	7	97	52	11	96	54
pic77	35	98	67	37	97	67
pic78	52	98	75	24	99	62
pic79	7	98	53	10	99	55
Schnitt	27	95	61	31	95	63

Tabelle B.2: Ergebnisse für 11 aus 256 und aus 16 Mio. (2)
(alle Angaben in Prozent)

Bilder	Klassifikation Abstand 10			Klassifikation Abstand 20		
	erkannt	n.erk.	Schnitt	erkannt	n.erk	Schnitt
pic1	24	88	56	41	76	59
pic2	24	89	57	51	58	55
pic3	38	89	64	56	74	65
pic4	37	72	55	73	37	55
pic5	40	41	41	81	13	47
pic6	91	26	59	95	20	58
pic7	29	47	38	61	26	44
pic8	3	99	51	8	95	52
pic9	25	75	50	60	43	52
pic10	27	64	46	53	43	48
pic11	45	61	53	74	27	51
pic12	5	97	51	17	70	44
pic13	22	91	57	44	70	57
pic14	18	84	51	37	66	52
pic15	43	62	53	69	39	54
pic16	44	83	64	73	39	56
pic17	13	86	50	42	56	49
pic18	21	90	56	51	48	50
pic19	29	86	58	55	76	66
pic20	39	61	50	66	21	44
pic21	12	88	50	39	55	47
pic27	20	89	55	45	64	55
pic28	3	96	50	16	87	52
pic29	31	88	60	54	71	63
pic30	24	81	53	58	52	55
pic31	32	81	57	69	51	60
pic32	24	88	56	46	68	57
pic34	18	90	54	41	78	60
pic35	20	92	56	40	86	63
pic36	35	80	58	72	48	60
pic37	36	61	49	73	41	57
pic38	34	88	61	56	71	64
pic39	34	68	51	65	41	53
pic40	25	94	60	47	74	61
pic41	31	81	56	61	40	51
pic42	45	88	67	71	61	66
pic43	40	88	64	69	69	69
pic44	24	92	58	51	72	62

Tabelle B.3: Tabelle mit allen Ergebnissen für Klassifikation mit Abstand 10 und 20 (alle Angaben in Prozent)

Bilder	Klassifikation Abstand 10			Klassifikation Abstand 20		
	erkannt	n.erk.	Schnitt	erkannt	n.erk.	Schnitt
pic45	32	93	63	69	55	62
pic46	13	93	53	41	64	53
pic47	55	61	58	75	35	55
pic48	3	98	51	35	77	56
pic49	14	94	54	58	78	68
pic50	21	62	42	43	29	36
pic52	10	94	52	27	84	56
pic53	15	86	51	38	65	52
pic54	4	67	36	18	29	24
pic55	47	85	66	74	49	62
pic56	22	92	57	53	51	52
pic57	25	80	53	59	35	47
pic58	52	27	40	77	16	47
pic59	26	69	48	69	10	40
pic60	10	81	46	28	48	38
pic61	45	63	54	65	35	50
pic62	6	80	43	18	45	32
pic63	26	63	45	56	41	49
pic64	22	93	58	60	60	60
pic65	20	77	49	64	28	46
pic66	9	89	49	65	40	53
pic67	39	61	50	65	39	52
pic68	20	83	52	51	43	47
pic69	40	84	62	73	36	55
pic70	11	84	48	30	61	46
pic71	21	82	52	45	55	50
pic72	31	84	58	50	66	58
pic73	56	82	69	77	37	57
pic74	26	60	43	51	40	46
pic75	47	68	58	75	41	58
pic76	19	85	52	42	55	49
pic77	43	57	50	74	14	44
pic78	20	60	40	29	43	36
pic79	32	74	53	57	56	57
Schnitt	28	78	53	54	51	53

Tabelle B.4: Tabelle mit allen Ergebnissen für Klassifikation mit Abstand 10 und 20 (2)
(alle Angaben in Prozent)

Bilder	Klassifikation Abstand 30			Chromatisch		
	erkannt	n.erk.	Schnitt	erkannt	n.erk	Schnitt
pic1	57	61	59	44	87	66
pic2	67	33	50	57	95	76
pic3	71	53	62	42	98	70
pic4	86	18	52	28	72	50
pic5	95	5	50	80	91	86
pic6	96	16	56	9	93	51
pic7	82	12	47	79	78	79
pic8	20	80	50	0	100	50
pic9	78	15	47	71	97	84
pic10	70	29	50	86	97	92
pic11	91	6	49	41	91	66
pic12	49	29	39	24	96	60
pic13	64	43	54	43	89	66
pic14	53	51	52	64	37	51
pic15	82	22	52	89	96	93
pic16	85	17	51	63	100	82
pic17	64	31	48	79	62	71
pic18	75	25	50	26	91	59
pic19	70	51	61	80	73	77
pic20	81	7	44	46	95	71
pic21	66	20	43	27	95	61
pic27	65	41	53	81	84	83
pic28	38	68	53	0	99	50
pic29	81	40	61	12	97	55
pic30	84	29	57	73	89	81
pic31	90	27	59	40	83	62
pic32	71	47	59	83	50	67
pic34	57	67	62	24	98	61
pic35	51	79	65	22	99	61
pic36	90	31	61	65	65	65
pic37	89	32	61	44	87	66
pic38	75	39	57	83	81	82
pic39	86	19	53	56	92	74
pic40	67	28	48	47	84	66
pic41	82	17	50	15	97	56
pic42	85	24	55	91	94	93
pic43	80	53	67	70	91	81
pic44	75	45	60	54	77	66

Tabelle B.5: Ergebnisse für Klassifikation mit Abstand 30 und Chromatisch (1)
(alle Angaben in Prozent)

Bilder	Klassifikation Abstand 30			Chromatisch		
	erkannt	n.erk.	Schnitt	erkannt	n.erk	Schnitt
pic45	89	28	59	76	80	78
pic46	62	34	48	28	87	58
pic47	89	13	51	52	85	69
pic48	68	44	56	82	50	66
pic49	87	59	73	9	45	27
pic50	60	11	36	65	90	78
pic52	40	76	58	37	92	65
pic53	60	48	54	35	76	56
pic54	39	15	27	68	75	72
pic55	87	19	53	67	92	80
pic56	72	29	51	60	77	69
pic57	82	22	52	69	60	65
pic58	91	9	50	70	98	84
pic59	81	4	43	52	93	73
pic60	46	36	41	44	78	61
pic61	83	17	50	87	62	75
pic62	28	32	30	0	98	49
pic63	73	30	52	56	85	71
pic64	79	27	53	72	78	75
pic65	83	12	48	77	87	82
pic66	78	25	52	40	89	65
pic67	82	24	53	47	90	69
pic68	75	30	53	55	84	70
pic69	88	9	49	83	80	82
pic70	48	42	45	55	89	72
pic71	67	29	48	16	86	51
pic72	65	48	57	64	75	70
pic73	86	7	47	71	75	73
pic74	68	26	47	66	93	80
pic75	88	25	57	41	59	50
pic76	63	28	46	94	77	86
pic77	85	4	45	35	98	67
pic78	50	29	40	29	99	64
pic79	70	35	53	30	90	60
Schnitt	72	31	52	52	84	68

Tabelle B.6: Ergebnisse für Klassifikation mit Abstand 30 und Chromatisch (2)
(alle Angaben in Prozent)

Bilder	Kopplung Ab.30/aus 16 Mio.			HSV		
	erkannt	n.erk.	Schnitt	erkannt	n.erk	Schnitt
pic1	66	60	63	28	78	53
pic2	68	3	36	39	91	65
pic3	83	52	68	23	69	46
pic4	93	17	55	43	73	58
pic5	96	5	51	20	94	57
pic6	96	15	56	5	93	49
pic7	90	11	51	24	89	57
pic8	21	80	51	21	83	52
pic9	82	14	48	25	80	53
pic10	71	28	50	11	97	54
pic11	91	5	48	40	88	64
pic12	54	28	41	62	67	65
pic13	73	41	57	41	82	62
pic14	19	96	58	25	78	52
pic15	88	22	55	27	91	59
pic16	91	16	54	24	95	60
pic17	21	96	59	18	86	52
pic18	75	24	50	23	89	56
pic19	86	51	69	61	87	74
pic20	90	7	49	68	97	83
pic21	75	20	48	14	93	54
pic27	81	40	61	72	72	72
pic28	40	67	54	16	85	51
pic29	82	40	61	16	94	55
pic30	90	29	60	24	89	57
pic31	91	26	59	10	90	50
pic32	81	43	62	80	50	65
pic34	58	67	63	22	92	57
pic35	55	78	67	24	91	58
pic36	96	27	62	25	78	52
pic37	90	31	61	25	73	49
pic38	91	37	64	81	58	70
pic39	91	18	55	15	93	54
pic40	69	27	48	19	88	54
pic41	87	17	52	22	74	48
pic42	87	21	54	28	91	60
pic43	82	50	66	35	90	63
pic44	87	43	65	30	72	51

Tabelle B.7: Ergebnisse für die Kopplung von zwei Verfahren und HSV (1)
(alle Angaben in Prozent)

Bilder	Kopplung Ab.30/aus 16 Mio.			HSV		
	erkannt	n.erk.	Schnitt	erkannt	n.erk.	Schnitt
pic45	94	27	61	26	87	57
pic46	70	32	51	23	81	52
pic47	94	13	54	60	80	70
pic48	81	42	62	29	76	53
pic49	89	58	74	19	45	32
pic50	71	11	41	64	88	76
pic52	64	74	69	62	74	68
pic53	80	47	64	64	76	70
pic54	46	14	30	32	88	60
pic55	94	18	56	28	96	62
pic56	82	29	56	54	58	56
pic57	93	19	56	31	73	52
pic58	92	8	50	19	92	56
pic59	87	4	46	55	88	72
pic60	56	30	43	95	62	79
pic61	91	13	52	76	56	66
pic62	30	32	31	65	68	67
pic63	83	28	56	63	83	73
pic64	89	26	58	70	71	71
pic65	94	11	53	41	75	58
pic66	88	23	56	28	71	50
pic67	90	23	57	20	91	56
pic68	84	26	55	89	61	75
pic69	92	9	51	18	92	55
pic70	62	40	51	24	84	54
pic71	75	29	52	20	84	52
pic72	84	47	66	21	82	52
pic73	93	7	50	32	91	62
pic74	82	26	54	18	89	54
pic75	91	23	57	13	84	49
pic76	69	26	48	23	90	57
pic77	91	4	48	37	91	64
pic78	62	29	46	66	71	69
pic79	74	34	54	46	86	66
Schnitt	78	31	54	36	81	59

Tabelle B.8: Ergebnisse für die Kopplung von zwei Verfahren und HSV (2)
(alle Angaben in Prozent)

Abbildungsverzeichnis

1.1	Ablauf der Bild-Erkennung	3
1.2	Einordnung der Arbeit in den Problemzusammenhang	4
2.1	Der RGB Würfel	7
2.2	HSV-Farbsystem	8
3.1	Ablauf der Farbselektion	10
3.2	Fallbasierte Klassifikation	13
3.3	Ablauf der linearen Klassifikation	14
3.4	Ablauf der Haut-Erkennung	17
3.5	Hautverteilung beim Modell im chromatischen Farbraum	18
3.6	Gauß-Verteilung	19
4.1	Hautfarben mit denen gearbeitet wird	23
4.2	Umsetzung lineare Klassifikation	24
4.3	Aufbau der Octree-Datenstruktur	25
5.1	Problematik der Schwellwertsuche	33
A.1	Ergebnisse der verschiedenen Selektionen zusammengestellt	46
A.2	Ergebnisse der unterschiedlichen Klassifikationen	47
A.3	Ergebnisse der Kopplung zweier Verfahren	48

Tabellenverzeichnis

4.1	Farbenauswahl aus 16 Millionen	22
4.2	Farbenauswahl aus 256 Farben	22
4.3	Struktur einer Bitmap-Datei	27
5.1	Originalbild	30
5.2	Ergebnisse: Optimal und 11 aus 256 Farben	31
5.3	Ergebnisse: Optimal und aus 16 Millionen	32
5.4	Ergebnisse zur Schwellwertsuche	34
5.5	Ergebnisse: Optimal, linearen Klassifikation mit Abstand 20 . . .	35
5.6	Ergebnisse für Abstand 10, 20 und 30 im direkten Vergleich . . .	36
5.7	Ergebnisse: Optimal und chromatischer Farbraum	37
5.8	Oder Verknüpfung bei Kopplung der Ergebnisse	38
5.9	Optimales Ergebnis und Ergebnis aus dem HSV-Farbsystem . . .	39
5.10	Übersicht gesammelter Endergebnisse	40
B.1	Ergebnisse für Selektion mit 11 aus 256 und aus 16 Mio. (1) . . .	50
B.2	Ergebnisse für Selektion 11 aus 256 und aus 16 Mio. (2)	51
B.3	Ergebnisse für Klassifikation mit Abstand 10 und 20 (1).	52
B.4	Ergebnisse für Klassifikation mit Abstand 10 und 20 (2).	53
B.5	Klassifikation mit Abstand 30 und Chromatisch (1)	54
B.6	Klassifikation mit Abstand 30 und Chromatisch (2)	55
B.7	Ergebnisse für die Kopplung von zwei Verfahren und HSV (1) . .	56
B.8	Ergebnisse für die Kopplung von zwei Verfahren und HSV (2) . .	57

Literaturverzeichnis

- [CG] Foley, van Dam, Feiner, Hughes *Computer Graphics: Principles and Practice* Chapter 13.3 Color Models for Raster Graphics
- [Char] Eric F.Charlton *The Octree* University of Michigan June 1997
<http://engin.umich.edu/CFD/users/charlton/Thesis/html/node29.html>
- [ChRo0500] Henry Chang, Ulises Robles: *Face Detection* Universität Stanford Mai 2000
- [DHF] J.Cai, A.Goshtasby, C.Yu *Detecting Human Faces in Color Images* Workshop on Multi-Media Database Management Systems, 1998
- [FFF] Margaret M.Fleck, David A.Forsyth *Finding Naked People* Department of Computer Science University of Iowa IA 52242
- [FDCI] Jay P.Kapur *Face Detection in Color Images* University of Washington Department of Electrical Engineering, Capstone Design Project, EE499, 1997 Computer Science Division, U.C.Berkeley, Berkeley CA 94720
- [FFI] Margaret M.Fleck, David A.Forsyth *Identifying nude pictures* Department of Computer Science, University of Iowa, Iowa City, IA 52240
Computer Science Division, U.C.Berkeley, Berkeley, CA 94720, 1996
- [FFNP] Margaret M.Fleck, David A.Forsyth *Naked People Skin Filter* Department of Computer Science, University of Iowa, Iowa City
Computer Science Division, U.C.Berkeley, 1996
- [FSCI] Mahbubul Islam Chowdhury *Finding Skin in Color Images* University of Newfoundland
- [GUM] Karin Kylander, Olof S.Kylander *Gimp: The Official Handbook* Library of Congress Cataloging-in-Publication Data, 1999
- [JöSchG] Gesetz zum Schutze der Jugend in der Öffentlichkeit *Jugendschutzgesetz JöSchG* vom 25.Februar 1985(BGBl.S.1227) mit allen Änderungen in der Fassung vom 28.Juni 1990 (BGBl.S.1227) zuletzt geändert am 28.Oktober 1994(BGBl.I.S.3186/3197)
www.kinderprojekte.de

- [SCMA] Jie Yang, Weier Lu, Alex Waibel *Skin-Color Modeling and Adaptation* School of Computer Science, Carnegie Mellon University Pittsburgh, PA 15213
- [STCM] Michael J.Jones, James M.Rehg *Statistical Color Models with Application to Skin Detection* Cambridge Research Laboratory, 1998
- [UBL0599] Dorthe Lübbert *Regressionsrechnung* Universität Bochum, Statistik C, Mai 1999
- [UKB97] R.Bergmann *Fallbasiertes Schließen* Universität Kaiserslautern, Zentrum für Lernende Systeme und Anwendungen (LSA), AG Prof. M. M. Richter, 1997
- [USG] Artikel aus heise online *Oberstes US-Gericht prüft Gesetz gegen Internet-Pornografie*
www.heise.de/newsticker/data/wst-22.05.01-001
- [UW] *Struktur einer .bmp Datei* Universität Wien, Advanced Computer Engineering
- [WWU] *Einführung in Neuronale Netze* Westfälische Wilhelms-Universität Münster, Neuro-Fuzzy-AG