



h_da

HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES

fbi

FACHBEREICH INFORMATIK

Hochschule Darmstadt
- Fachbereich Informatik -

**DETEKTION VON OHREN AUF 3D-PROFILBILDERN UNTER
VERWENDUNG DER FLÄCHENKRÜMMUNGEN**

Abschlussarbeit zur Erlangung des akademischen Grades
Bachelor of Science (B.Sc.)

vorgelegt von
Adrian Winterstein

Referent: Prof. Dr. Wolf-Dieter Groch
Korreferent: Prof. Dr. Christoph Busch

Durchgeführt am:



Center for Advanced Security Research Darmstadt

ERKLÄRUNG

Ich versichere hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die im Literaturverzeichnis angegebenen Quellen benutzt habe.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder noch nicht veröffentlichten Quellen entnommen sind, sind als solche kenntlich gemacht.

Die Zeichnungen oder Abbildungen in dieser Arbeit sind von mir selbst erstellt worden oder mit einem entsprechenden Quellennachweis versehen.

Diese Arbeit ist in gleicher oder ähnlicher Form noch bei keiner anderen Prüfungsbehörde eingereicht worden.

Darmstadt den 25. Januar 2012

ABSTRAKT

Das Ohr erhält als biometrisches Merkmal wenig Aufmerksamkeit im Vergleich zu Fingerabdrücken oder dem Gesicht. Dennoch ist die Ohrmuschel auf Grund ihrer vielseitigen anatomischen Struktur gut für die Unterscheidung von Personen geeignet und wird im Bereich der Forensik seit langem verwendet. Das Ohr wird zu diesem Zweck in Form von zwei- oder dreidimensionalen Bildern erfasst, wobei dreidimensionale Aufnahmen durch die fortschreitende Entwicklung entsprechender Sensoren an Bedeutung gewinnen.

Damit die aufgenommenen Bilder in einem biometrischen System verwendet werden können muss eine Detektion der Ohrregion vorgenommen werden. In dieser Arbeit wird ein neues Konzept zur automatischen Detektion von Ohren auf dreidimensionalen Profilbildern vorgestellt, bei dem nur Tiefen-, aber keine Texturinformationen benötigt werden. Hierbei werden die für das Ohr spezifischen starken Krümmungen in der Oberfläche der Helix des Ohres genutzt.

Beim hier vorgestellten Ansatz wird der Verlauf der Helix über die, in ein zweidimensionales Binärbild transformierten, maximalen Krümmungen der Oberfläche bestimmt. Es werden die Zusammenhangskomponenten des Binärbildes ermittelt, über kurze Distanzen verbunden und zu Kandidaten für den Ohrumriss kombiniert. Schließlich werden Bewertungskriterien definiert, die eine Unterscheidung von Ohrenkandidaten und damit die Auswahl des Kandidaten ermöglichen, der die Ohrregion am besten darstellt.

Der entwickelte Algorithmus wurde auf der öffentlich zugänglichen UND-J2-Datenbank getestet, die über 2000 dreidimensionale Profilaufnahmen von 404 Personen enthält. Dabei konnte die Ohrregion auf 94% der Bilder erfolgreich ermittelt werden. Die Detektion funktionierte sowohl für linke, als auch rechte Ohren problemlos und erwies sich als robust gegen Rotationen der Eingabebilder in der Bildebene.

INHALTSVERZEICHNIS

ABBILDUNGSVERZEICHNIS	7
1 EINLEITUNG	9
1.1 Das Ohr als biometrisches Charakteristikum	9
1.1.1 Kriterien zur Bewertung von Charakteristiken	10
1.1.2 Identifikation von Personen auf Fotografien	12
1.2 Grundlagen eines biometrischen Systems	12
1.2.1 Detektion von Ohren	13
1.3 Das Gesamtprojekt	13
1.4 Aufbau der Arbeit	14
2 GRUNDLAGEN	15
2.1 Flächenkrümmungen	15
2.1.1 Shape-Index	17
2.2 Morphologische Operationen	18
2.2.1 Hit-Miss-Operator	18
2.3 Bresenham-Algorithmus	20
2.4 Krümmung zweidimensionaler Komponenten	21
2.5 Eckendetektion	23
2.6 Datenbanken biometrischer Testdaten	24
3 STAND DER TECHNIK	25
3.1 Ohrendetektion auf zweidimensionalen Aufnahmen	25
3.1.1 AdaBoost-Klassifikator	26
3.1.2 Verwendung der Helix	26
3.1.3 Weitere kantenbasierte Ansätze	27
3.2 Ohrendetektion auf dreidimensionalen Aufnahmen	27
3.2.1 Active-Contour-Algorithmus	28
3.2.2 Template- und Modell-basierte Ansätze	28
3.2.3 Klassifikator von Shape-Index-Histogrammen	29
3.2.4 Zusammenfassung	30
3.3 Anforderungen	30
4 KONZEPT ZUR OHRENDETEKTION	31
4.1 Ablauf der Detektion	32

Inhaltsverzeichnis

4.2	Bildsegmentierung	34
4.2.1	Glättung	35
4.2.2	Binarisieren	35
4.2.3	Finden der Zusammenhangskomponenten	37
4.3	Verbinden über kurze Distanzen	40
4.3.1	Eigenschaften von Shapes	41
4.3.2	Suchen möglicher Verbindungen	43
4.3.3	Metriken zur Bewertung einer Verbindung	45
4.3.4	Zusammenfassen zu einem gemeinsamen Shape	45
4.3.5	Limitierungen und Probleme beim Verbinden	46
4.4	Auswahl einiger prominenter Shapes	47
4.5	Untersuchung von Parallelität	48
4.5.1	Überprüfung paralleler Abschnitte	49
4.5.2	Alternatives Suchen paralleler Shapes	51
4.6	Erweiterung des Verbindens	52
4.6.1	Verbinden über größere Distanzen	53
4.6.2	Der Shape-Index als weiteres Flächenmaß	53
4.7	Kategorisierung von Kandidaten	55
4.8	Kombination von Ohrenkandidaten	56
4.8.1	Kombination paralleler Shapes	57
4.8.2	Weitere Kombinationen von Shapes	57
4.9	Bewertung der Ohrähnlichkeit	60
4.9.1	Absolute Bewertungskriterien	60
4.9.2	Relative Bewertungskriterien	63
4.9.3	Kombination zu einer gemeinsamen Bewertung	64
4.9.4	Auswahl des besten Kandidaten	65
4.10	Markieren der Ohrregion	65
4.10.1	Achsenparalleles Rechteck	66
4.10.2	Ellipse	66
5	REALISIERUNG	67
5.1	Framework	67
5.2	Modul zur Ohrendetektion	68
5.2.1	Datenstruktur	69
5.2.2	Evaluators	70
5.2.3	Konfiguration	71
6	ERGEBNISSE & DISKUSSION	72
6.1	Messmethode	72
6.1.1	Kennzahlen des Überlappungsgrades	73
6.1.2	Erkennungsrate	74
6.2	Messergebnisse	75
6.2.1	Erkennungsrate	75
6.2.2	Überlappungsgrade	76

Inhaltsverzeichnis

6.2.3	Geometrische Transformationen	78
6.2.4	Positive Ergebnisse	79
6.2.5	Verbleibende Probleme	81
7	FAZIT	87
7.1	Ausblick	87
7.2	Zusammenfassung	88
8	ANHANG	90
8.1	Vergleich verschiedener Schwellwerte	90
8.2	Vergleich verschiedener Strategien zur Auswahl prominenter Shapes	91
	LITERATURVERZEICHNIS	93

ABBILDUNGSVERZEICHNIS

1.1	Merkmale der Ohrmuschel	10
1.2	Grundlegender Aufbau eines biometrischen Systems	12
1.3	Aufnahmeperspektive am Geldautomat	14
2.1	Hauptkrümmungen einer Oberfläche	16
2.2	Darstellung von fünf Kategorien des Shape-Indexes	18
2.3	Beispiel einer Hit- und einer Miss-Maske	19
2.4	Hit- und Miss-Masken zum Ausdünnen von Objekten	19
2.5	Hit- und Miss-Masken zur Detektion von Verzweigungspunkten	20
2.6	Hit- und Miss-Masken zur Detektion von Endpunkten	20
2.7	Rasterung einer Linie mit dem Bresenham-Algorithmus	21
2.8	Skizze der Ableitungen nach x und y	22
2.9	Gegenüberstellung von runder, stumpfer und falscher Ecke	23
3.1	Beispielbilder aus der „IIT Kanpur Ear Database“	27
4.1	Darstellung der Oberfläche eines Ohres	31
4.2	Schritte des Konzeptes	33
4.3	Beispielbild zur Binarisierung	34
4.4	Auswirkungen der Glättung	35
4.5	Vergleich verschiedener Schwellwertverfahren	36
4.6	Spezialfall der Zusammenhangskomponenten	37
4.7	Beispiel des Entfernens kleiner Komponenten	38
4.8	Extraktion einer Zusammenhangskomponente	39
4.9	Boundary-Trace-Algorithmus	40
4.10	Boundary-Trace-Algorithmus auf Linien	40
4.11	Ermittlung der Ausgangsvektoren eines Shapes	42
4.12	Ungünstige Verbindung	43
4.13	Definition des Suchbereichs für Verbindungen	44
4.14	Metriken beim Verbinden von Shapes	45
4.15	Drehen von Komponenten beim Verbinden	46
4.16	Beispielbilder gewählter prominenter Shapes	48
4.17	Konvexe Minima und konkave Maxima der Flächenkrümmungen	49
4.18	Achsenparalleles Rechteck zur Einschränkung des Suchbereichs	50
4.19	Bestimmung potentiell paralleler Abschnitte	50
4.20	Bestimmung des orthogonales Abstands	50

Abbildungsverzeichnis

4.21	Krümmungsrichtungen abschnittsweise paralleler Shapes	51
4.22	Suchen paralleler Shapes	52
4.23	Probleme beim Verbinden von Shapes	53
4.24	Beispiele für ungünstige Shape-Index-Regionen	54
4.25	Die vier Kategorien der Ohrenkandidaten	55
4.26	Kombinieren des Shapes S_1 mit S_2 und S_3 zum gemeinsamen Shape S_k	57
4.27	Vergleich der möglichen Positionierungen von Shapes	58
4.28	Überprüfen zweier Shapes auf eine nebeneinander liegende Anordnung	58
4.29	Kombinieren nebeneinander angeordneter Shapes	59
4.30	Verschieben eines Shapes zum Kombinieren mit einem anderen	60
4.31	Bewertungskriterien für Ohrenkandidaten	61
4.32	Beispiele für Krümmungshäufung in Ecken	62
4.33	Vergleich von achsenparallelem Rechteck und Ellipse	66
5.1	Klassendiagramm des Testframeworks	68
5.2	Klassendiagramm der Datenstruktur	69
5.3	Vererbungshierarchie der Evaluators	71
6.1	Beispiele der manuell erzeugten Ground-Truth	73
6.2	Kennzahlen des Überlappungsgrades	74
6.3	Beispiele für einen Überlappungsgrad von etwa 60%	75
6.4	Erkennungsrate in Abhängigkeit des Überlappungsgrades	76
6.5	Überlappungsgrade für die detektierten Regionen	77
6.6	Übergrößen und Untergrößen der detektierten Ohrregionen	77
6.7	Überlappungsgrade bei verschiedenen Rotationen der Eingabebilder	78
6.8	Überlappungsgrade bei der horizontalen Spiegelung der Eingabebilder	79
6.9	Beispiele erfolgreicher Detektionen des Ohrumrisses	80
6.10	Beispiel mit Ohrringen als Störfaktoren	82
6.11	Löcher in der aufgenommenen Oberfläche des Ohres	82
6.12	Ausgangsbedingungen für eine durch Haare bedingte Fehlklassifikation	83
6.13	Beispiele fehlerhafter Detektionen des Ohrumrisses	84
6.14	Angepasster Shape und Original-Shapes von zwei Ohrenkandidaten	85
8.1	Testergebnisse der Auswahl prominenter Shapes	92

*Das Ohr der Neugier liegt nur an
den Türen des Glückes und der Lei-
denschaft.*

Friedrich Schiller

1

EINLEITUNG

Die Identifikation von Personen anhand biometrischer Merkmale ist ein aktuelles Thema, sei es in der Privatwirtschaft oder bei der Aufklärung von Straftaten. Biometrische Charakteristika, die dazu Verwendung finden, sind beispielsweise Fingerabdrücke, die Iris, das Gesicht oder auch der Gang einer Person. Vergleichsweise wenig Aufmerksamkeit erhält das Ohr. Dennoch wird es beispielsweise von Strafverfolgungsbehörden als wichtiges Merkmal für die Identifikation geschätzt. So spielt das Iannarelli-System, ein System zur manuellen Klassifikation von Ohren, seit über 40 Jahren eine wichtige Rolle in der US-amerikanischen Forensik [JAIN 2007].

1.1 DAS OHR ALS BIOMETRISCHES CHARAKTERISTIKUM

Ein „biometrisches Charakteristikum“ ist definiert als ein messbares „biologisches oder verhaltensabhängiges Charakteristikum eines Individuum“ [3D F 2009], das zur Unterscheidung von Personen verwendet werden kann. Die Identifikation unter Verwendung des Ohres nutzt zur Unterscheidung die einzigartige Struktur der Ohrmuschel. Einige dazu geeignete Merkmale sind die Helix und die Anti-Helix, der Tragus, das Cavum Conchae und das Ohrläppchen. In Abbildung 1.1 sind ihre Positionen in der Ohrmuschel gekennzeichnet.

1 Einleitung

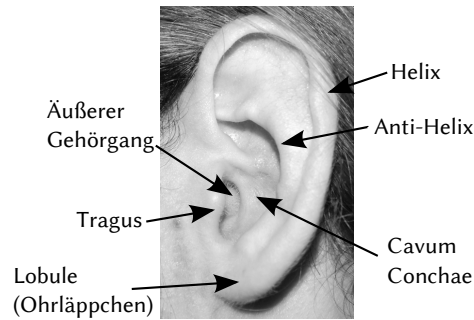


Abbildung 1.1: Darstellung einiger Merkmale der Ohrmuschel.

1.1.1 KRITERIEN ZUR BEWERTUNG VON CHARAKTERISTIKEN

Allgemein kann die Eignung einer Charakteristik für die Nutzung in einem biometrischen System nach Jain et al. anhand von sieben formalen Kriterien bewertet werden [JAIN 2007]:

1. **Universality:** Jedes Individuum der Zielgruppe muss das Merkmal besitzen.
2. **Uniqueness:** Es muss innerhalb der betrachteten Population einzigartig sein.
3. **Permanence:** Über den Anwendungszeitraum des biometrischen Systems, darf sich das Merkmal nicht signifikant verändern.
4. **Measurability:** Die Eigenschaften des Merkmals müssen ohne übermäßige Belästigung des Individuums gemessen und verarbeitet werden können. Geeignete Sensoren müssen technisch möglich und umsetzbar sein.
5. **Performance:** Die Genauigkeit und die nötigen Ressourcen zum Erreichen dieser Genauigkeit müssen die Randbedingungen des Anwendungsfalls erfüllen.
6. **Acceptability:** Die Zielgruppe muss bereit sein, das Merkmal von dem biometrischen System erfassen zu lassen.
7. **Circumvention:** Die Merkmale sollten sich nicht zu leicht imitieren lassen. Das biometrische System darf nicht umgangen werden können, beispielsweise durch Manipulationen am System oder den gespeicherten und zu verarbeitenden Daten.

Kein biometrisches Charakteristikum kann alle Anforderungen für alle Anwendungsfälle erfüllen, d.h. es gibt kein ideales Charakteristikum, aber einige akzeptable [JAIN 2007]. Bezogen auf das Ohr ergibt sich folgendes Bild: Die Einzigartigkeit (Uniqueness) wird vollständig erfüllt, denn kein Ohr gleicht dem anderen. Das gilt auch für die Ohren eineiiger Zwillinge

1 Einleitung

und sogar für das linke und rechte Ohr jedes Menschen [IANN 1989]. Iannarelli konnte auch zeigen, dass die Ohrmuschel sich im Laufe des Lebens kaum in ihrer Form verändert. Bereits wenige Monate nach der Geburt ist die Form des Ohres stabil und Änderungen finden nur noch in der Größe statt. Damit ist das Kriterium der Beständigkeit (Permanence) ebenfalls gut erfüllt. Im Gegensatz zum Gesicht ist die Ohrmuschel auch keinen Verzerrungen durch Mimik unterworfen [JAIN 2007].

„Next to fingerprints, the external ear constitutes the most unique design, characteristic features, and peculiarities for the purpose of identification. On no other part of the human body do we have flesh lines with such a unique design“ [IANN 1989].

Jeder Mensch besitzt eine Ohrmuschel, wenn sie auch in seltenen Fällen missgebildet oder nur in Ansätzen vorhanden sein kann. Iannarelli bescheinigt dem Ohr auch eine hohe Universalität (Universality) gerade im Bezug auf forensische Anwendungsfälle, da beispielsweise von einer vermissten Person meist Fotografien vorhanden sind, auf denen auch die Ohren zu erkennen sind, während nicht unbedingt die Fingerabdrücke erfasst wurden.

Die Akzeptanz (Acceptability) der Erfassung des Charakteristikums kann als hoch eingeschätzt werden, da sie auch aus größerer Distanz erfolgen kann. Das Freilegen des Ohres kann jedoch je nach Kultur als Eingriff in die Persönlichkeitsrechte aufgefasst werden, da beispielsweise Schleier abgenommen werden müssten. Bei der Fälschungssicherheit (Circumvention) ergeben sich für biometrische Systeme, die das Ohr als Charakteristikum verwenden die gleichen Probleme, wie beispielsweise bei der Verwendung des Gesichts. Es wird nicht überprüft, ob das aufgenommene Charakteristikum zu einer lebenden Person gehört (Lebenderkennung). Deshalb können Systeme, die mit zweidimensionalen Eingabedaten arbeiten, durch Fotos getäuscht werden. Dreidimensionale Modelle von Ohren, wie beispielsweise Gipsabdrücke oder Latex-Ohren, können Systeme täuschen, die 3D-Daten erfassen.

Je nachdem, wie die Merkmale der Ohrmuschel erfasst werden, also ob z.B. unkontrollierte Bedingungen mit variabler Distanz zur Kamera und nicht definierten Kopfposen vorherrschen, sind die Kriterien der Messbarkeit (Measurability) und Performanz (Performance) mehr oder weniger gut erfüllt. So kann die Ohrmuschel beispielsweise teilweise oder vollständig durch Haare oder eine Mütze verdeckt sein. Die Sicherstellung einer hohen Genauigkeit bei der automatisierten Messung der Merkmale ist eine Herausforderung in der Ohrbiometrie, gerade bei unkontrollierten Aufnahmebedingungen.

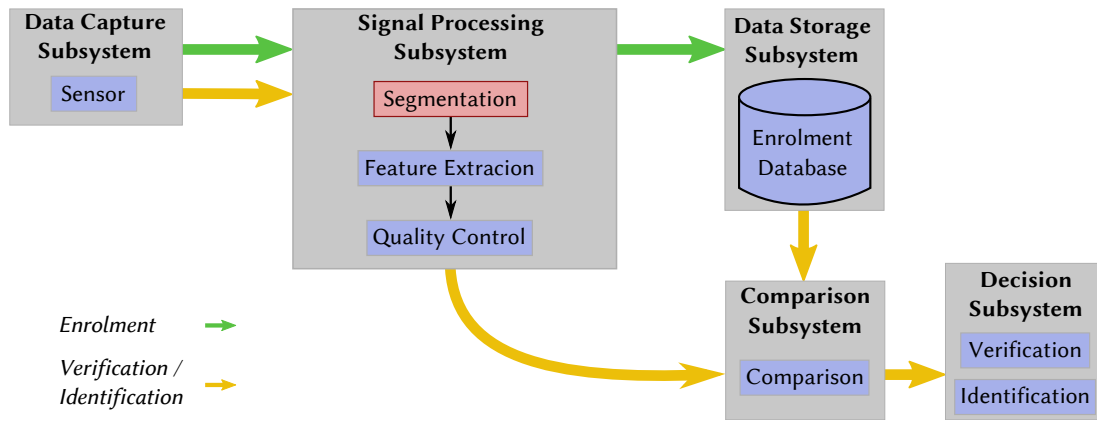


Abbildung 1.2: Grundlegender Aufbau eines biometrischen Systems. Der Bestandteil „Segmentation“ ist es, um den sich dieses Projekt dreht.

1.1.2 IDENTIFIKATION VON PERSONEN AUF FOTOGRAFIEN

Wie Iannarelli zeigte, lassen sich die Merkmale des Ohres problemlos aus Fotografien ermitteln. Es bietet sich daher an, Aufnahmen von Ohren, beispielsweise aus Überwachungsvideos, zur Identifikation mutmaßlicher Straftäter zu verwenden. Das Ohr kann dabei in Ergänzung zum Gesicht genutzt werden.

Der Stand der Technik bei 3D-Scannern ermöglicht es dabei, die Beschränkung auf zweidimensionale Fotos aufzuheben und zusätzlich dreidimensionale Aufnahmen zu verwenden. Dabei können beispielsweise 3D-Scans zur Erfassung der biometrischen Charakteristik Ohr verwendet und zur Identifikation einer Person mit zweidimensionalen Aufnahmen verglichen werden. Dieses Projekt hat zum Ziel, Ohren aus dreidimensionalen Aufnahmen zu detektieren und damit die automatische Verarbeitung dieser Aufnahmen zu ermöglichen.

1.2 GRUNDLAGEN EINES BIOMETRISCHEN SYSTEMS

Ein System zur automatischen Identifizierung oder Verifizierung von Personen anhand biometrischer Charakteristiken, wird als biometrisches System bezeichnet. Es besteht nach Definition der ISO [BIOM 2010] aus fünf Subsystemen, die in Abbildung 1.2 skizziert sind.

Daraus ergibt sich folgender Ablauf: Zunächst wird vom Sensor des *Data Capture Subsystems* eine „analoge oder digitale Repräsentation biometrischer Charakteristika“ [3D F 2009], das biometrische Sample, erfasst und an das *Signal Processing Subsystem* übermittelt. Dort findet eine Segmentierung des Samples, die Extraktion biometrischer Features und eine Qualitätskontrolle statt. Wird ein Enrolment durchgeführt, d.h. soll die biometrische Referenz in eine Datenbank eingefügt werden, wird die vom *Signal Processing Subsystem* er-

1 Einleitung

zeugte Referenz vom *Data Storage Subsystem* gespeichert. Sind bereits biometrische Referenzen vorhanden, kann die Verifikation oder Identifikation einer Person durchgeführt werden. Dazu wird die aus dem erfassten biometrischen Sample erzeugte Probe mit einer (bei der Verifikation) oder mehreren (bei der Identifikation) Referenzen aus der *Enrolment Database* verglichen, was die Aufgabe des *Comparison Subsystems* ist. Probe und Referenzen werden bezüglich ihrer Übereinstimmung bewertet. Schließlich wird auf Basis dieser Bewertungen unter Verwendung eines Grenzwertes im *Decision Subsystem* eine Entscheidung getroffen, ob die Probe zur selben Person gehört wie die Referenz.

1.2.1 DETEKTION VON OHREN

Im *Signal Processing Subsystem* ist zunächst eine Segmentierung nötig, damit später über die Extraktion von Features biometrische Referenzen erzeugt werden können. Bezogen auf den in diesem Projekt zugrunde gelegten Fall der Identifikation von Personen auf Fotografien unter Verwendung des Ohres, muss das Ohr zunächst aus den Bilddaten extrahiert werden. Manuelles Segmentieren der Ohrregion würde beim Enrolment einer großen Datenbank viel Zeit in Anspruch nehmen. Ein mit Hilfe von Bildverarbeitung automatisierter Prozess ist daher praktikabler.

Aus Eingabebildern, im Falle dieses Projekts dreidimensionale Aufnahmen, muss also in einem automatisierten Verfahren das Ohr extrahiert, d.h. auf diesen Bildern detektiert werden. Die genauen Anforderungen dafür sind in Abschnitt 3.3 spezifiziert.

1.3 DAS GESAMTPROJEKT

Praxisrelevant sind die Ergebnisse dieser Arbeit durch die Einbettung in ein Projekt zum Aufbau eines multimodalen biometrischen Erkennungssystems, d.h. eines biometrischen Systems in dem mehrere biometrische Charakteristiken genutzt werden. Es handelt sich hierbei um das Gesicht und das Ohr. Durchgeführt wird das Projekt in Zusammenarbeit mit dem BKA, mit dem Ziel die Identifikationsrate auf tatrelevanten Bildern zu erhöhen [Ver 2010].

Derzeit sind in den Datenbanken des BKA zweidimensionale Gesichtsbilder gespeichert. Ein automatisierter Vergleich dieser Bilder ist nur mit Frontalaufnahmen möglich. Notwendig ist die Nutzung von Bildern zur Identifikation möglicher Täter meist bei Taten im öffentlichen Raum, wo keine verwertbaren Fingerabdrücke oder DNA-Spuren gefunden werden können, aber Aufnahmen von Überwachungskameras existieren. Diese Aufnahmen sind häufig keine Frontalaufnahmen, sondern variieren in der Pose oder zeigen durch eine un-

1 Einleitung



Abbildung 1.3: Beispielhafte Darstellung der Aufnahmeperspektive am Geldautomat. Dieses Foto stammt aus einer Präsentation des BKA-Mitarbeiters Harald Sieber.

günstige Kameraposition nur einen Ausschnitt des Kopfes. Zusätzlich können die Beleuchtungsverhältnisse schlecht sein. Starkes Rauschen und niedrige Auflösung erschweren die Identifikation. Ein Beispiel für diese Faktoren ist in Abbildung 1.3 zu sehen. Die Charakteristik Ohr wird bisher nicht automatisiert ausgewertet, womit wichtige Unterscheidungsmerkmale nicht beachtet werden.

Mit der Verwendung dreidimensionaler Referenzen und dem Entwickeln von Verfahren zur 3D-Gesichtserkennung und 3D-Ohrenerkennung soll die Erkennungsleistung deutlich verbessert werden. Die in dieser Arbeit beschriebene Ohrendetektion ist ein wichtiger Schritt zur Entwicklung eines automatisierten Verfahrens zur Ohrenerkennung mit dreidimensionalen Bildern.

Diese Arbeit wurde am Center for Advanced Security Research Darmstadt (CASED) erstellt. Das CASED ist ein Cluster der TU-Darmstadt, des Fraunhofer SIT und der Hochschule Darmstadt, der „wichtiges Grundlagenwissen und IT-Sicherheitslösungen für die Praxis“ [CASE 2010] erarbeitet.

1.4 AUFBAU DER ARBEIT

Diese Arbeit gliedert sich in sieben Kapitel. Auf die Einleitung folgen zunächst einige mathematische Grundlagen und Konzepte der Bildverarbeitung. Dann wird der Stand der Technik in den Bereichen Identifikation und Ohrenerkennung dargestellt, bevor in Kapitel 4 das der entwickelten Ohrendetektion zugrundeliegende Konzept ausführlich beschrieben wird. Die programmiertechnische Realisierung des Algorithmus wird im darauf folgenden Kapitel kurz dargestellt. Im nächsten Kapitel werden die erzielten Ergebnisse vorgestellt und erläutert. Das abschließende Kapitel enthält einen kurzen Ausblick und eine Zusammenfassung.

*Wir glauben dem Auge
mehr als dem Ohr.*

Aus Ungarn

2

GRUNDLAGEN

In diesem Kapitel werden die Grundlagen beschrieben, auf denen das in Kapitel 4 dargestellte Konzept basiert. Ein wichtiger Bestandteil sind die in Abschnitt 2.1 erläuterten Flächenkrümmungsmaße. In den darauf folgenden Abschnitten werden zwei Konzepte der Bildverarbeitung erläutert, der morphologische Hit-Miss-Operator (Abschnitt 2.2), sowie der Bresenham-Algorithmus zur Rasterung von Linien (Abschnitt 2.3). Darauf folgt die Vorstellung eines Konzeptes zur Berechnung der Krümmungen zweidimensionaler Komponenten (Abschnitt 2.4), sowie zur Detektion von Ecken (Abschnitt 2.5). In Abschnitt 2.6 wird zuletzt die zum Testen des Algorithmus verwendete Bilddatenbank kurz beschrieben.

2.1 FLÄCHENKRÜMMUNGEN

Für Flächen im dreidimensionalen Raum \mathbb{R}^3 existieren zwei wichtige Krümmungsbegriffe: die Mittlere Krümmung H und die Gaußsche Krümmung K . Sie lassen sich über die beiden Hauptkrümmungen k_1 und k_2 der Oberfläche wie folgt definieren [BESL 1988]:

$$H = \frac{1}{2}(k_1 + k_2) \quad K = k_1 k_2 \quad (2.1)$$

2 Grundlagen

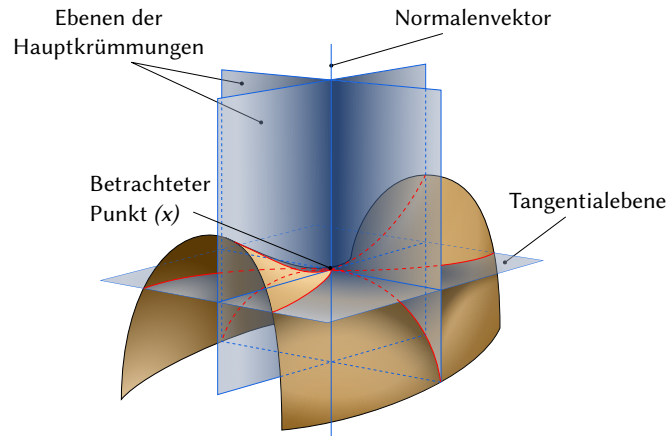


Abbildung 2.1: Darstellung der Hauptkrümmungen einer Oberfläche. Diese Grafik ist eine Übersetzung und geringfügige Erweiterung der Abbildung [Minimal surface curvature planes](#) von Eric Gaba, [CC BY-SA 3.0](#).

In [\[DORA 1995\]](#) wird zusätzlich noch die „Curvedness“ als ein Maß für die Stärke der Krümmung einer Oberfläche definiert:

$$C = \sqrt{\frac{k_1^2 + k_2^2}{2}} \quad (2.2)$$

Die Hauptkrümmungen einer Oberfläche im Punkt x sind die maximale und minimale Krümmung aller Flächenkurven in x . Flächenkurven sind dabei Kurven, die auf der Fläche verlaufen. In [Abbildung 2.1](#) sind einige Flächenkurven als rote Linien dargestellt. Alle Tangenten der Flächenkurven liegen in der ebenfalls in [Abbildung 2.1](#) dargestellten Tangentialebene. Senkrecht auf der Tangentialebene steht der Normalenvektor der Fläche für den Punkt x . Berechnen lassen sich die Hauptkrümmungen k_1, k_2 über folgenden Zusammenhang bei gegebenem H und K [\[BESL 1988\]](#):

$$k^2 - 2Hk + K = 0 \quad (2.3)$$

Da die Hauptkrümmungen entsprechend nur ermittelt werden können, wenn mittlere und gaußsche Krümmung bekannt sind, müssen H und K auf eine andere Art als über den in den [Gleichungen 2.1](#) beschriebenen Zusammenhang bestimmt werden. Möglich ist dies über die Koeffizienten E, F und G der ersten und die Koeffizienten L, M und N der zweiten quadratischen Fundamentalform der Fläche [\[BESL 1988\]](#):

$$H = \frac{EN + GL - 2FM}{2(EG - F^2)} \quad K = \frac{LN - M^2}{EG - F^2} \quad (2.4)$$

2 Grundlagen

Die Koeffizienten der ersten Fundamentalform für den Punkt x sind dabei folgende:

$$E = |x_u|^2 \quad F = x_u x_v \quad G = |x_v|^2 \quad (2.5)$$

Und die der zweiten Fundamentalform:

$$L = \frac{x_u \times x_v}{|x_u \times x_v|} x_{uu} \quad M = \frac{x_u \times x_v}{|x_u \times x_v|} x_{uv} \quad N = \frac{x_u \times x_v}{|x_u \times x_v|} x_{vv} \quad (2.6)$$

Hierbei sind x_u und x_v jeweils die ersten partiellen Ableitungen nach u und v für den betrachteten Punkt x und x_{uu} , x_{uv} und x_{vv} analog die zweiten partiellen Ableitungen. Diese Ableitungen lassen sich, wenn nur diskrete Punktwolken statt einer definierten Oberfläche vorhanden sind, über die Gradienten von x , beziehungsweise die Gradienten der ersten partiellen Ableitungen von x ermitteln.

In dieser Arbeit werden zwei verschiedene Arten von Krümmungen unterschieden. Zum einen die in diesem Abschnitt beschriebenen Krümmungen einer dreidimensionalen Oberfläche, für die im folgenden immer der Begriff *Flächenkrümmungen* verwendet wird und zum anderen die Krümmungen von Linien im zweidimensionalen Raum, welche nur als *Krümmungen* bezeichnet werden.

2.1.1 SHAPE-INDEX

Der Shape-Index bietet eine Möglichkeit der quantitativen Messung der Form einer Oberfläche für einen bestimmten Punkt i dieser Oberfläche [ZHOU 2010]. Er ist nach [DORA 1995] definiert durch:

$$S_i = \frac{1}{2} - \frac{1}{\pi} \arctan \left(\frac{k_{1_i} + k_{2_i}}{k_{1_i} - k_{2_i}} \right) \quad S_i \in [0,1] \quad (2.7)$$

Dabei sind k_{1_i} und k_{2_i} die beiden Hauptkrümmungen für Punkt i , mit $k_{1_i} \geq k_{2_i}$.

„Every distinct shape corresponds to a unique value of $[S_i]$, except the planar shape. Points on a planar surface have an indeterminate shape index since $[k_{1_i} = k_{2_i} = 0]$ “ [DORA 1995].

Die Werte des Shape-Index können dabei nach [ZHOU 2010] in neun Kategorien festgelegter Intervalle unterschieden werden, dargestellt in Tabelle 2.1. Abbildung 2.2 zeigt eine grafische Darstellung der Shape-Index-Kategorien.

Tabelle 2.1: Die neun Kategorien des Shape-Index [ZHOU 2010].

Shape-Kategorie	$S_i \in$	Shape-Kategorie	$S_i \in$
Spherical cup	$(0, 1/16)$	Saddle Ridge	$(9/16, 11/16)$
Through	$(1/16, 3/16)$	Ridge	$(11/16, 13/16)$
Rut	$(3/16, 5/16)$	Dome	$(13/16, 15/16)$
Saddle Rut	$(5/16, 7/16)$	Spherical cap	$(15/16, 1)$
Saddle	$(7/16, 9/16)$		

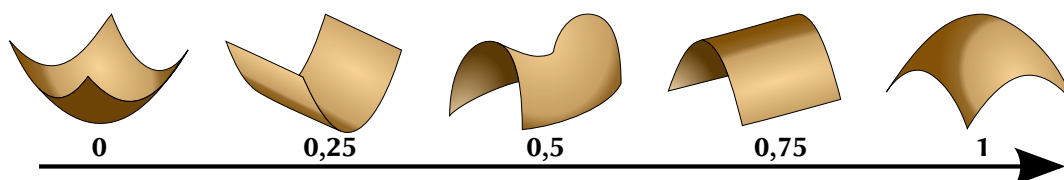


Abbildung 2.2: Darstellung von fünf Kategorien des Shape-Indexes: Spherical cup, Rut, Saddle, Ridge und Spherical cap. Die restlichen vier Kategorien sind Zwischenformen im Übergang von einer abgebildeten Kategorie zur nächsten. Diese Abbildung basiert auf „Illustration of the relationship between the values of the shape index and the shapes“ aus [YOSH 2002] und der bereits für Abbildung 2.1 verwendete Grafik „Minimal surface curvature planes“.

2.2 MORPHOLOGISCHE OPERATIONEN

Morphologische Operationen werden in der digitalen Bildverarbeitung auf Binärbilder angewandt und sind „Nachbarschaftsoperationen, [...] mit denen sich die Form von Objekten modifizieren und analysieren lässt“ [JAHN 2005]. Beispielsweise kann die Erosion oder Dilation eines Bildes mit einem morphologischen Operator statt mit einem Rangfolgeoperator durchgeführt werden, wenn ein Binärbild vorliegt. Definiert sind morphologische Operationen als Mengenoperationen auf der Menge der Bildpunkte eines Binärbildes B , die ungleich null sind.

2.2.1 HIT-MISS-OPERATOR

Mit dem Hit-Miss-Operator können Objekte detektiert werden, die eine spezifische Form aufweisen [JAHN 2005]. Zu diesem Zweck müssen nacheinander zwei morphologische Erosionen durchgeführt werden.

Zunächst wird das Binärbild B mit der sogenannten Hit-Maske erodiert. Dazu wird die Hit-Maske an jede Position des Eingabebildes B gelegt. Die entsprechende Position im Ausgabebild B_a erhält nur dann eine 1, wenn unter jeder 1 der Hit-Maske auch eine 1 im binären Eingabebild vorhanden ist. In allen anderen Fällen ergibt sich der Wert 0 für die Position in

2 Grundlagen

B_a . Dadurch werden alle Objekte entfernt, die kleiner als die durch die Hit-Maske definierte Struktur sind. Nach Anwendung der Hit-Maske wird die Miss-Maske analog auf das invertierte Binärbild \overline{B} angewendet. Dadurch werden alle Objekte entfernt, die größer als die gesuchte Struktur sind.

In Abbildung 2.3 sind eine Hit- und die zugehörige Miss-Maske zur Detektion von Objekten abgebildet, die aus zwei oder drei waagrecht nebeneinander angeordneten Pixeln bestehen. Positionen an denen sowohl die Hit- als auch die Miss-Maske eine 0 aufweisen spielen bei beiden Erosionen keine Rolle und können somit für eine gewisse Flexibilität sorgen, da an diesen Positionen ein Objektpixel vorhanden sein darf aber nicht vorhanden sein muss.

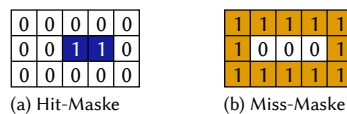


Abbildung 2.3: Beispiel einer Hit- und einer Miss-Maske zur Detektion eines zwei bis drei Pixel langen Objektes.

In den folgenden Abschnitten werden drei in dieser Arbeit verwendete Operatoren mit ihren Hit- und ihren Miss-Masken vorgestellt.

THINNING

Der Thinning-Operator ist ein Hit-Miss-Operator, mit dem sich Objekte verdünnen lassen, ohne dass sie aufgetrennt werden. Abbildung 2.4 zeigt die vier Hit- und die vier Miss-Masken, die notwendig sind um ein Objekt von allen Seiten aus verdünnen zu können.

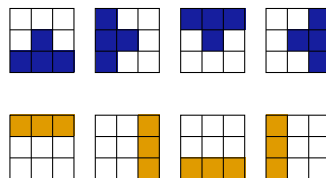


Abbildung 2.4: Hit-Masken (obere Reihe) und die zugehörigen Miss-Masken (untere Reihe) zum Ausdünnen von Objekten (vgl. [JAHN 1997]).

Die Masken liefern die äußerste Pixelschicht an allen Stellen, die breiter als ein Pixel sind. Wird das Ergebnisbild, das sich durch die Anwendung dieser Masken ergibt, vom Ausgangsbild subtrahiert, werden alle Objekte um einen Pixel verdünnt. Um das Verdünnen eines Objektes auf eine Breite von einem Pixel zu ermöglichen, kann der Thinning-Operator so oft angewendet werden, bis keine Änderungen mehr stattfinden [JAHN 1997].

DETEKTION VON VERZWEIGUNGSPUNKTEN

Verzweigungspunkte sind all jene Punkte in einem Binärbild, die genau drei Nachbarn besitzen. Diese beiden Nachbarn sind selbst nicht benachbart. Die Detektion von Verzweigungspunkten kann mit den in Abbildung 2.5 dargestellten Masken aus [BANS 2010] erfolgen.

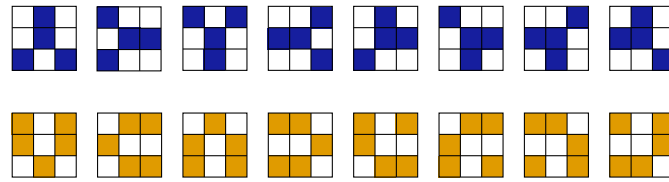


Abbildung 2.5: Hit-Masken (obere Reihe) und die zugehörigen Miss-Masken (untere Reihe) zur Detektion von Verzweigungspunkten. Basierend auf „Figure 12“ aus [BANS 2010].

DETEKTION VON ENDPUNKTEN

Endpunkte ausgedünnter Linien können ebenfalls mit einem Hit-Miss-Operator ermittelt werden. Dazu können die in Abbildung 2.6 dargestellten Hit- und Miss-Masken verwendet werden, die in [BANS 2010] eingeführt wurden.

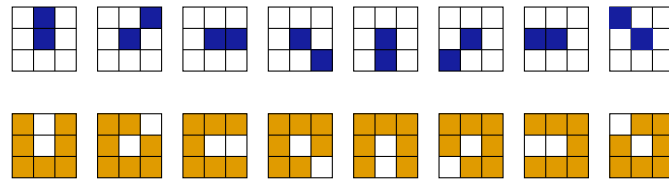


Abbildung 2.6: Hit-Masken (obere Reihe) und die zugehörigen Miss-Masken (untere Reihe) zur Detektion von Endpunkten. Basierend auf „Figure 11“ aus [BANS 2010].

2.3 BRESENHAM-ALGORITHMUS

Der Bresenham-Algorithmus ist ein Algorithmus zur Rasterung von geraden Linien, deren Anfangs- und Endpunkte gegeben sind. Diese Rasterung ist notwendig, um stetige Linien in diskrete Bildkoordinaten zu übertragen. Der große Vorteil des Bresenham-Algorithmus ist, dass er „die Berechnung der einzelnen Rasterpunkte auf die einfachen Rechenoperationen Addition und Subtraktion reduzier[t]“ [ZEPP 2003] und dadurch eine hohe Geschwindigkeit erreichen kann. Wie in [ZEPP 2003] beschrieben ist, wird dazu folgendes ausgenutzt: Es werden nur Linien betrachtet, die eine Steigung zwischen 0° und 45° aufweisen, da sich

alle anderen Fälle darauf abbilden lassen. In Abbildung 2.7 ist zu erkennen, dass sich die x -Koordinate bei der Beschränkung auf diesen Bereich der Steigungen in jedem Schritt um genau 1 erhöht und die y -Koordinate nur jeweils gleich bleiben oder sich um 1 erhöhen kann.

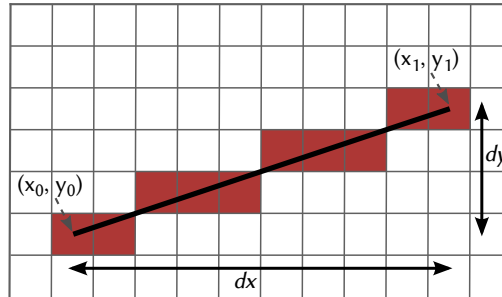


Abbildung 2.7: Rasterung einer Linie von der Koordinate (x_0, y_0) nach (x_1, y_1) mit dem Bresenham-Algorithmus.

Der Ablauf des Algorithmus ist wie folgt, wobei dx und dy wie in Abbildung 2.7 definiert sind und bei (x_0, y_0) begonnen wird. Es wird eine Fehlervariable θ definiert und mit $2dy - dx$ initialisiert. In jedem Schritt wird zunächst der Pixel an den aktuellen Koordinaten gezeichnet. Dann wird die aktuelle x -Koordinate um eins erhöht und $2dy$ auf den Wert von θ addiert. Die y -Koordinate wird nur dann um eins erhöht, wenn $\theta > 0$ ist. In diesem Fall wird von dem Wert von θ nach der Inkrementierung der y -Koordinate $2dx$ subtrahiert. Auf diese Art können stetige Linien, so wie in Abbildung 2.7 beispielhaft dargestellt, gerastert werden.

2.4 KRÜMMUNG ZWEIDIMENSIONALER KOMPONENTEN

Die Krümmungen zweidimensionaler Komponenten lassen sich nach [HE 2008] auf die im Folgenden beschriebene Art in vier Schritten bestimmen. Zunächst wird ein Binärbild erzeugt, um aus diesem die Zusammenhangskomponenten zu extrahieren. Die Koordinaten der Bildpunkte der Zusammenhangskomponenten werden in der Reihenfolge ihrer Nachbarschaft in je einer Liste für jede Zusammenhangskomponente abgelegt. Mit einem Gauß-Filter findet dann eine Glättung der x -Koordinaten und eine Glättung der y -Koordinaten statt. Nach der Glättung entsprechen die Koordinaten nicht mehr ganzzahligen Pixelkoordinaten, sondern es sind rationale Zahlen entstanden. Schließlich werden die Krümmungen für jeden Punkt der Zusammenhangskomponenten, also für jede Position in den Listen mit den geglätteten Koordinaten, berechnet. Die Krümmungen sind in diesem Fall Krümmun-

2 Grundlagen

gen im zweidimensionalen Raum und keine Flächenkrümmungen, wie sie in Abschnitt 2.1 eingeführt wurden.

Voraussetzung dazu ist, dass die Zusammenhangskomponenten eine Breite von genau einem Pixel an jeder Position aufweisen. Bei dem von [HE 2008] als Grundlage genommenen CSS-Algorithmus (Curvature Scale-Space), der zuerst in [MOKH 1998] eingeführt wurde, geschieht dies durch die Verwendung des Canny-Kantendetektors zur Binarisierung. Verzweigungen werden beim Extrahieren der Zusammenhangskomponenten ignoriert, indem jeweils nur ein Nachbarpixel gewählt wird.

Die extrahierten Zusammenhangskomponenten A werden durch eine Faltung ($*$) mit der Gaußfunktion g geglättet. Die Stärke der Glättung kann dabei durch σ der Gaußfunktion kontrolliert werden:

$$A_{smooth} = A * g \quad (2.8)$$

$$g = e^{-\frac{t^2}{2\sigma^2}} \quad (2.9)$$

Um die Glättung durchführen zu können, ohne die Komponente dabei zu kürzen, muss die Komponente vorher an ihren beiden Enden symmetrisch erweitert werden.

Nach dem Glätten, kann die Krümmung K_i für jeden Pixel i aller Pixel einer Komponente berechnet werden:

$$K_i = \frac{\Delta x_i \Delta^2 y_i - \Delta^2 x_i \Delta y_i}{[(\Delta x_i)^2 + (\Delta y_i)^2]^{1,5}} \quad (2.10)$$

Die Werte Δx_i und Δy_i sind dabei über die Differenz zu i benachbarter x und y wie folgt definiert:

$$\Delta x_i = \frac{1}{2}(x_{i+1} - x_{i-1}) \quad \Delta y_i = \frac{1}{2}(y_{i+1} - y_{i-1}) \quad (2.11)$$

Veranschaulichen lässt sich die Berechnung von Δx und Δy anhand von Abbildung 2.8. Der Wert Δx ist dabei die Differenz der x -Werte der beiden Nachbarpixel (x_{i-1}, y_{i-1}) und (x_{i+1}, y_{i+1}) des Pixels (x_i, y_i) . Analog dazu ist Δy über die y -Werte der Nachbarpixel definiert.

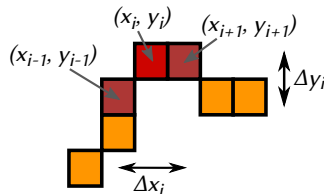


Abbildung 2.8: Berechnung von Δx_i und Δy_i durch die Differenz der x -, bzw. y -Werte der beiden Nachbarn von x_i . Zur besseren Verdeutlichung ist in dieser Skizze keine Glättung vorgenommen worden.

Analog zu Δx und Δy sind $\Delta^2 x$ und $\Delta^2 y$ als die Differenzen zu i benachbarter Δx und Δy definiert:

$$\Delta^2 x_i = \frac{1}{2}(\Delta x_{i+1} - \Delta x_{i-1}) \quad \Delta^2 y_i = \frac{1}{2}(\Delta y_{i+1} - \Delta y_{i-1}) \quad (2.12)$$

2.5 ECKENDETEKTION

Die Eckendetektion nach [HE 2008] läuft in mehreren Schritten ab. Zunächst werden alle lokalen Maxima der Beträge der Krümmungen in eine Liste von Eckenkandidaten aufgenommen. Die Berechnung der Krümmungen ist in Abschnitt 2.4 beschrieben. Aus dieser Liste werden dann nach den im folgenden beschriebenen Kriterien alle sogenannten runden und falschen Ecken entfernt.

Runde Ecken sollen aus der Liste der Kandidaten entfernt werden, während stumpfe Ecken beibehalten werden. Eine Darstellung dieser beiden Typen von Ecken ist in Abbildung 2.9 zu finden. Dazu wird für jeden Eckenkandidaten die sogenannte Region of Support bestimmt, die durch die beiden nächsten Krümmungsminima begrenzt ist. Innerhalb dieser Region wird über die durchschnittliche Krümmung ein lokaler, adaptiver Schwellwert bestimmt. Ist die Krümmung des Eckenkandidaten größer als der berechnete Schwellwert, so ist er tatsächlich eine Ecke, im anderen Fall wird er als runde Ecke aus der Liste entfernt. Dies funktioniert, da die Krümmungen bei runden Ecken sich im Bereich um die Ecke vergleichsweise langsam verringern.

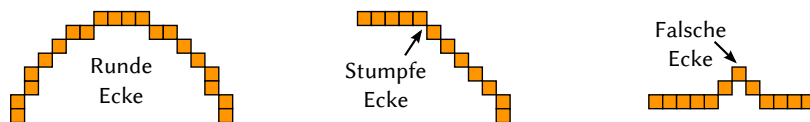


Abbildung 2.9: Gegenüberstellung von runder, stumpfer und falscher Ecke. Runde und falsche Ecken sollen ausgeschlossen werden, während stumpfe Ecken als Ecke detektiert werden sollen. Diese Abbildung basiert auf „Examples of round corner and obtuse corner“, sowie „Feature of different sizes“ [HE 2008].

Zur Entfernung falscher Eckenkandidaten wird wieder eine Region of Support definiert, diesmal aber mit den beiden nächsten Eckenkandidaten, die nach dem Entfernen runder Ecken übrig sind, als Grenzen. Für die beiden Seiten (vor und hinter dem untersuchten Eckenkandidaten) der Region of Support wird jeweils der Schwerpunkt berechnet, sodass zwei Geraden vom Eckenkandidaten aus durch die Schwerpunkte definiert werden können. Der Winkel zwischen diesen beiden Geraden kann als Kriterium zur Erkennung falscher Ecken dienen. Je kleiner der Winkel ist, desto wahrscheinlicher handelt es sich um eine

Ecke. Durch die Festlegung eines Schwellwertes können falsche Ecken daher ausgeschlossen werden. Nach dem Entfernen runder und falscher Ecken aus der Liste der Eckenkandidaten, enthält diese nur noch die gesuchten Ecken des Eingabebildes.

2.6 DATENBANKEN BIOMETRISCHER TESTDATEN

Die Tests von Algorithmen zur Ohrendetektion werden auf biometrischen Bilddatenbanken durchgeführt. Die umfangreichste derzeit verfügbare Datenbank mit dreidimensionalen Profilaufnahmen ist die Bilddatenbank der University of Notre Dame (UND) [YAN 2007]. Die Teilmenge J2 dieser Datenbank enthält linksseitige Profilaufnahmen von 404 Personen, wobei von jeder Person Aufnahmen in bis zu 22 Sitzungen über 2 Jahre erstellt wurden. Nicht jede Person war bei jeder Sitzung anwesend, so dass die Anzahl der Aufnahmen pro Person variiert. Insgesamt sind es 2414 verschiedene Aufnahmen, wobei jeweils ein zweidimensionales Foto und ein exakt registriertes Tiefenbild vorliegen. Beleuchtung und Hintergründe unterscheiden sich bei verschiedenen Bildern. Auf die UND-J2 wurde für das Testen des im Rahmen dieser Arbeit entwickelten Konzeptes zurückgegriffen. Die in dieser Arbeit abgebildeten Ergebnisse und Zwischenergebnisse wurden aus Bildern dieser Datenbank erzeugt.

Chen und Bhanu [CHEN 2007] verwenden eine eigene Bilddatenbank, das „UCR data set“. Auch in dieser Datenbank sind Tiefeninformationen und die zugehörigen Farbwerte vorhanden. Zusätzlich zu den Profilaufnahmen enthält die Datenbank von jeder Person zwei Aufnahmen mit um etwa 35° nach links und rechts gedrehtem Kopf. Die UCR-Datenbank ist bisher nicht öffentlich verfügbar, auch wenn laut der Internetseite des „Visualization and Intelligent Systems Laboratory“ der „University of California Riverside“ geplant ist, die Datenbank zugänglich zu machen [VISL 2009]. Da der Hinweis, dass dies bald geschehen solle, bereits seit 2009 auf der Internetseite zu finden ist, kann davon ausgegangen werden, dass eine Veröffentlichung nicht mehr stattfindet.

*Das menschliche Ohr ist besonders
geeignet, an die Wand gelegt zu
werden.*

Bruno Ziegler

3

STAND DER TECHNIK

Dieser Abschnitt gibt einen Überblick über den aktuellen Stand bei der Detektion von Ohren auf zwei- und dreidimensionalen Daten. Eine gebräuchliche Lösung sowohl im Zwei-, als auch im Dreidimensionalen ist die Verwendung eines Templates, also einer Schablone des Ohres. Die Ohrregion wird dabei über die am besten passendste Position des Templates im Bild definiert. Auch Klassifikatoren, die nach einer umfangreichen Trainingsphase die Regionen eines gegebenen Bildes in Ohrregionen und Regionen ohne Ohr unterteilen können, werden in verschiedenen Variationen eingesetzt.

3.1 OHRENDETEKTION AUF ZWEIDIMENSIONALEN AUFNAHMEN

Zur Detektion von Ohren auf zweidimensionalen Aufnahmen existieren viele Veröffentlichungen, von denen im Folgenden eine kleine Auswahl vorgestellt wird. Dabei wird sich auf kantenbasierte Ansätze und Klassifikatoren beschränkt, da nur diese relevant für das im Rahmen dieser Arbeit entwickelte Konzept sind, beziehungsweise als Vergleich bei der Präsentation der Messergebnisse herangezogen werden. Das eigentliche Thema dieser Arbeit ist die Detektion im dreidimensionalen Raum.

3.1.1 ADABOOST-KLASSIFIKATOR

Ein AdaBoost-Klassifikator, wie er beispielsweise bei [ISLA 2008] oder [SHIH 2009] zum Einsatz kommt, ist eine weit verbreitete Technik zur Detektion von Objekten, die auch für Gesichter und Ohren zum Einsatz kommt. Die Implementierung einer Objektdetektion mit einem AdaBoost-Klassifikator ist auch in der Bildverarbeitungsbibliothek OpenCV vorhanden und wird von [SHIH 2009] verwendet. Der Klassifikator selbst wird oft als Viola-Jones-Klassifikator bezeichnet und nur das Trainingsverfahren, bei dem die im folgenden Absatz beschriebene Gewichtung erfolgt, als AdaBoost.

Ein Klassifikator untersucht ein Bild mit Hilfe eines Fensters fester Größe, das an jede mögliche Position des Bildes verschoben wird. Die untersuchten Ausschnitte werden von dem Klassifikator unterteilt in jene, die aus einem Objekt der gesuchten Art bestehen und jene ohne diesen Objekttyp. Die Klassifikation erfolgt dabei auf Basis einfacher Features [VIOL 2001], den Basisfeatures, meist definiert als rechteckige Bereiche, bestehend aus weißen und schwarzen Rechtecken. Die Basisfeatures symbolisieren lokale Helligkeitsübergänge. Während des Trainingsprozesses wird ermittelt, wie die Regionen, die als positive und negative Beispiele verwendet werden, aus den Basisfeatures zusammengesetzt sind. Je nach Häufigkeit des Vorkommens bestimmter Zusammensetzungen werden diese gewichtet. Die einfachen Features, auch als schwacher Klassifikator bezeichnet, werden dann über mehrere Stufen, die sogenannte Kaskade, zu einem starken Klassifikator kombiniert. Dadurch kann in der Kombination eine hohe Genauigkeit erreicht werden.

Bei der Klassifikation wird dann nach den durch den Trainingsprozess bestimmten Zusammensetzungen der Hell-Dunkel-Übergänge, also der Basisfeatures, gesucht. Der mehrstufige Aufbau des AdaBoost macht ihn sehr zeiteffizient, da große Bereiche des Bildes bereits mit den ersten Stufen ausgeschlossen werden können und nur einige Bereiche genau untersucht werden müssen. Daher kann sogar Echtzeitfähigkeit bei der Detektion von Ohren erreicht werden [ABAZ 2010].

3.1.2 VERWENDUNG DER HELIX

Ansari und Gupta [ANSA 2007] nutzen die parallel verlaufenden Kanten der äußeren Helix zur Detektion der Ohrregion. Sie verwenden den Canny-Kantendetektor auf zweidimensionalen Grauwertbildern und unterteilen die ermittelten Kanten in konvexe und konkave. Von diesen Kanten werden, als mögliche Bestandteile der Helix, die Kanten ausgewählt, die abschnittsweise parallel zu anderen Kanten verlaufen. Schließlich werden aus diesen gewählten

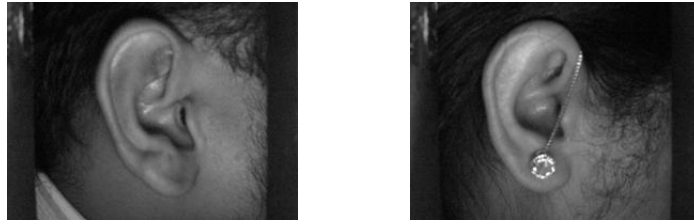


Abbildung 3.1: Beispielbilder aus der „IIT Kanpur Ear Database“ [KUMA 2011].

Kanten alle möglichen Ohrumrisse gebildet. Als korrekter Ohrumriss gilt dann jener, der aus den meisten Pixeln von Kanten aus dem Grauwertbild besteht.

In [ANSA 2007] wird mit diesem Ansatz eine Detektionsrate von 93,34% auf der „IIT Kanpur Ear Database“ erreicht. In Abbildung 3.1 ist zu sehen, dass die Bilder dieser Datenbank nur einen Ausschnitt der Region um das Ohr zeigen. Auf vollständigen Profilbildern (eventuell noch mit Hintergrund) ist eine deutlich geringere Performanz dieses Ansatzes zu erwarten, wie gerade auch die Erfahrungen mit dem in dieser Arbeit beschriebenen Konzept zeigen.

3.1.3 WEITERE KANTENBASIERTE ANSÄTZE

Den Canny-Kantendetektor verwenden auch [ARBA 2007]. Sie suchen danach die Ellipse, die der Ohrumriss bildet, mit Hilfe einer reduzierten Hough-Transformation. Bei der Anwendung auf Bildern, auf denen relativ viel Hintergrund zu sehen ist, muss für diesen Ansatz zunächst eine Hautdetektion als Vorsegmentierung vorgenommen werden.

Jeges und Máté [JEJE 2007] detektieren das Ohr ebenfalls auf einem Kantenbild. Das Kantenbild ist in diesem Fall allerdings kein Binärbild, sondern stellt die Intensitäten der Kanten im Bereich von 0 bis 255 dar. Aus den Kantenintensitäten berechnen sie für jeden Bildpunkt einer Kante die Orientierung der Kante. Ein Template-Matching wird dann auf einem Bild der Kantenorientierungen durchgeführt.

3.2 OHRENDETEKTION AUF DREIDIMENSIONALEN AUFNAHMEN

Zum Zeitpunkt der Erstellung dieser Arbeit existieren im Wesentlichen drei Ansätze zur Detektion von Ohren auf dreidimensionalen Aufnahmen. Diese Ansätze werden in den folgenden Absätzen vorgestellt.

3.2.1 ACTIVE-CONTOUR-ALGORITHMUS

Yan und Bowyer [YAN 2007] verwenden zur Ermittlung des Ohrumrisses einen Active-Contour-Algorithmus und arbeiten zu diesem Zweck auf dreidimensionalen Tiefenbildern und zweidimensionalen Farbbildern. Active-Contours, auch als Snakes bezeichnet, sind dabei Mengen von Punkten, die das jeweilige Zielobjekt möglichst genau umschließen sollen [NIXO 2002]. Eine Kontur als Initialwert ist für diese Algorithmen notwendig und wird bei [YAN 2007] durch eine Ellipse mit dem äußeren Gehörgang des Ohres als Mittelpunkt definiert. Für die Position des äußeren Gehörgangs in der Ohrmuschel siehe Abbildung 1.1. Um diesen Initialwert zu ermitteln, werden mehrere Schritte angewandt: Zunächst wird die Nasenspitze in dem Tiefenbild als der am weitesten links befindliche Bereich ermittelt, für den der Sensor des Aufnahmegerätes Informationen erhalten hat. Von ihr aus wird ein Kreissegment bestimmt, außerhalb dessen das Ohr nicht liegen kann. Im nächsten Schritt werden Farbinformationen benötigt, um innerhalb des Kreissegments eine weitere Segmentierung anhand der Hautfarbe vornehmen zu können. Schließlich wird der äußere Gehörgang über die Oberflächenkrümmungen der 3D-Aufnahme ermittelt. Der Active-Contour-Algorithmus arbeitet zuletzt mit einer Kombination der Farb- und Tiefeninformationen, um den Ohrumriss anzunähern.

Ein großer Nachteil dieses Ansatzes ist die Notwendigkeit der Verwendung von Farbinformationen, sowohl für die Hautsegmentierung, als auch die Active-Contour. Farben können je nach Beleuchtung deutlich variieren. Auch ist der äußere Gehörgang bereits bei geringen Drehungen des Kopfes vom Tragus überdeckt, so dass der Startpunkt für den Active-Contour-Algorithmus fehlt.

3.2.2 TEMPLATE- UND MODELL-BASIERTE ANSÄTZE

Chen und Bhanu stellen in [BHANU 2008] drei Methoden zur Detektion von Ohren auf 3D-Profilaufnahmen vor.

Bei der Template-basierten Methode werden Regionen maximaler Flächenkrümmungen über das Kriterium „Step Edge Magnitude“ ermittelt. Dies ist die maximalen Gradienten innerhalb eines Fensters vorgegebener Größe. Bei Chen und Bhanu ist die Fenstergröße 3×3 . Auf dem durch das Verschieben des Fensters über das Bild entstehenden „Step Edge Magnitue Image“ wird eine Binarisierung mit anschließender Dilatation zum Füllen von Löchern durchgeführt. Die Zusammenhangskomponenten werden dann ermittelt und Bounding-Boxes für jede Region definiert. Innerhalb der Bounding-Boxes werden Histogramme der Shape-Index-Kategorien (siehe Abschnitt 4.6.2) ermittelt und auf Ähnlichkeit

zu einem vorher erstellten Template untersucht. Das Template ist eine Mittlung der Shape-Index-Histogramme manuell extrahierter Ohrregionen.

Bei der zweiten Methode wird ein Modell für die Koordinaten der Punkte von Helix und Anti-Helix erstellt und die Position im Bild gesucht, an der dieses Modell mit dem geringsten Fehler registriert werden kann. Dazu wird ebenfalls ein „Step Edge Magnitude Image“ berechnet und aus diesem die Zusammenhangskomponenten ermittelt. In diesem Fall werden die Regionen aber nach der Dilatation skelettiert. Nahe beieinander liegende Regionen werden zu Clustern zusammengefasst, innerhalb derer die Registration des Modells dann stattfindet.

In [CHEN 2007] beschreiben Chen und Bhanu die dritte Methode ausführlich. Sie segmentieren das Farbbild und das zugehörige Tiefenbild zunächst nach der Hautfarbe, da Fehlklassifikationen bei den beiden zunächst entwickelten Algorithmen insbesondere von Pixeln in Regionen herrührten, die nicht zu den Hautbereichen zählen. Die Hautregionen werden zu einem Grauwertbild konvertiert, auf dem dann die Kanten detektiert werden. In den vorsegmentierten Bereichen des Tiefenbildes wird wie bei den ersten beiden Methoden die „Step Edge Magnitude“ bestimmt. Die dann aus dem Tiefenbild ermittelten Zusammenhangskomponenten werden mit den durch die Kantendetektion auf dem Farbbild bestimmten Kanten kombiniert. Auf diesem kombinierten Bild kann das Ohr dann mit Hilfe der Registrierung eines Modells detektiert werden.

3.2.3 KLASSIFIKATOR VON SHAPE-INDEX-HISTOGRAMMEN

Ausschließlich 3D-Daten (in Form von Tiefenbildern) verwenden Zhou et al. [ZHOU 2010]. Sie nutzen zur Detektion die Krümmungen der dreidimensionalen Oberfläche. Die ermittelten Krümmungsmaße sind die Curvedness (siehe Abschnitt 2.1), sowie der Shape-Index, der die Art der lokalen Krümmung angibt und eine Kategorisierung ermöglicht (siehe Abschnitt 2.1.1). Jedem Pixel des Bildes werden eine Shape-Kategorie und ein Wert zugeordnet, basierend auf dem berechneten Shape-Index und den Krümmungswerten. Das Bild wird mit einem Sliding-Window untersucht und für jede Fensterposition wird ein sogenannter Feature-Vektor angelegt, der aus konkatenierten Histogrammen besteht, deren Klassen den Shape-Kategorien entsprechen. Zur Ermittlung dieser Histogramme wird das Fenster in überlappende Blöcke verschiedener, festgelegter Größen unterteilt. Für jeden Block wird iterativ ein Histogramm berechnet, indem der Wert der Curvedness jedes Punktes innerhalb des Blockes auf den bisherigen Wert der Klasse des Histogramms addiert wird, die der Shape-Kategorie des Punktes entspricht.

Ein Klassifikator, der auf ebensolchen konkatenierten Histogrammen von Ohrregionen als Positiv- und anderen Bildbereichen als Negativbeispielen trainiert wurde, kann dann anhand der für das Ohr spezifischen Häufigkeit bestimmter Flächenkrümmungen, die Position des Ohres ermitteln. Dazu wird für jede Fensterposition ein Score berechnet, über den Vergleich des jeweiligen Histogramms mit dem aus dem Trainingsprozess resultierten.

3.2.4 ZUSAMMENFASSUNG

Die bisher existierenden Methoden zur Detektion von Ohren auf dreidimensionalen Bildern weisen deutliche Nachteile auf. Teilweise werden zusätzlich exakt registrierte zweidimensionale Aufnahmen benötigt oder aber es wird ein Klassifikator verwendet, der zunächst trainiert werden muss. Klassifikatoren besitzen, neben der notwendigen Trainingsphase, den Nachteil meist nur auf den Daten gute Ergebnisse zu liefern, die ähnlich zu jenen sind, auf denen das Training stattgefunden hat. Daher sind sie nur bedingt rotations- und skalierungsinvariant. Um transformierte Objekte zu detektieren, müssen die Bilder gezielt mit verschiedenen Fenstergrößen und -rotationen untersucht werden.

3.3 ANFORDERUNGEN

Aus den Limitierungen aktueller Ansätze zur Detektion von Ohren auf dreidimensionalen Aufnahmen ergeben sich folgende Ziele für ein neues Konzept: Der zu implementierende Algorithmus zur Ohrendetektion soll verschiedene Vorteile gegenüber anderen aktuellen Verfahren aufweisen. Als Eingabedaten sollen ausschließlich 3D-Informationen dienen, Farbinformationen aus Texturen sollen also nicht benötigt werden. Die Erkennung der Ohrregion soll ohne die Notwendigkeit einer Trainingsphase des Algorithmus erfolgen können. Damit ist die Verwendung eines Klassifikators ausgeschlossen. Ebenso sollen keine Annahmen über die Skalierung des Ohres nötig sein. Eine größtmögliche Unabhängigkeit von der Pose des abgebildeten Gesichts wird ebenfalls angestrebt. Ein weiteres Ziel ist auch die leichte Verständlichkeit des Algorithmus.

*Jedem Sprecher fehlt die Sprache,
fehlt dem Hörenden das Ohr.*

Franz Grillparzer

4

KONZEPT ZUR OHRENDETEKTION

In diesem Kapitel wird das, zur Detektion von Ohren auf dreidimensionalen Profilaufnahmen entwickelte, Konzept erläutert. Zunächst erfolgt ein kurzer Überblick der Konzeptes, dann werden die Details beschrieben, aufgeteilt in die sequentiell zu verarbeitenden Teilschritte und einige alternative Vorgehensweisen.

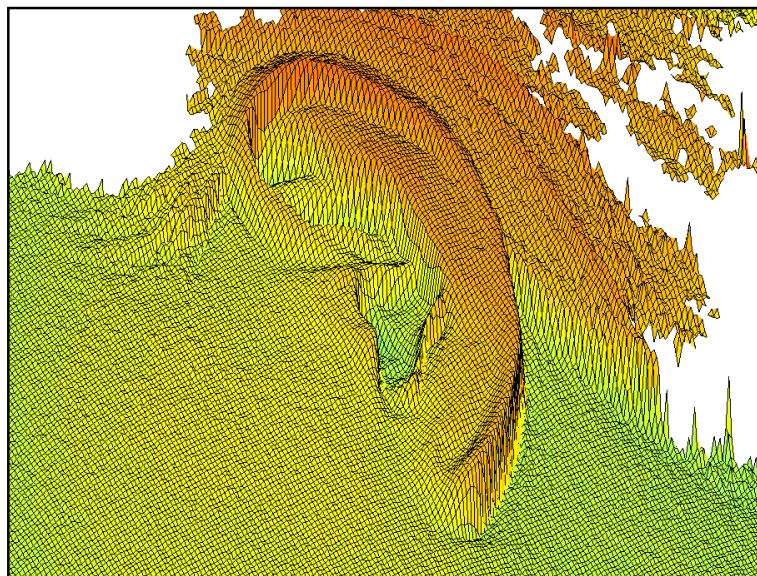


Abbildung 4.1: Darstellung der Oberfläche eines Ohres. Grundlage ist eine dreidimensionale Punktwolke aus der UND-J2-Datenbank.

Bei Betrachtung der Oberfläche des Gesichts auf einer dreidimensionalen Profilaufnahme fallen im Bereich des Ohres im Vergleich zu den restlichen Strukturen starke Flächenkrümmungen auf, insbesondere an der Helix. In [Abbildung 4.1](#) ist ein Beispiel für die Oberfläche eines Ohres zu sehen, die sich deutlich von ihrer Umgebung abhebt. Die Verwendung der Flächenkrümmungen, wie in [\[ZHOU 2010\]](#), bildet daher die Grundlage des im Folgenden beschriebenen Konzeptes zur Detektion des Ohres auf einer dreidimensionalen Aufnahme.

Die Grundidee des Konzeptes ist es, Zusammenhangskomponenten maximaler Flächenkrümmungen der dreidimensionalen Daten zu erzeugen und diese zu möglichen Ohrumrissen zu kombinieren. Für jede (x,y) -Koordinate des 3D-Eingabebildes wird über den zugehörigen z -Wert und die z -Werte der lokalen Umgebung die Flächenkrümmung berechnet. Die berechnete Flächenkrümmung wird dann an die entsprechende (x,y) -Koordinate einer zweidimensionalen Ausgabematrix geschrieben, so dass alle weiteren Bestandteile des Algorithmus nur auf zweidimensionalen Daten arbeiten müssen. Anhand von Bewertungskriterien, welche die Ohrähnlichkeit eines vermeintlichen Ohrumrisses quantifizieren, wird der beste Umriss als Ohr gewählt. Das Kombinieren von Komponenten zu möglichst vollständigen Ohrumrissen erfolgt in Anlehnung an den Ansatz aus [\[ANSA 2007\]](#), beschrieben in [Abschnitt 3.1.2](#), bei dem Kanten eines zweidimensionalen Bildes zu einem vollständigen Ohrumriss kombiniert werden.

4.1 ABLAUF DER DETEKTION

Die Detektion der Ohrregion läuft in mehreren sequentiell aufeinander folgenden Teilschritten ab, deren Zusammenhang im unteren Bereich von [Abbildung 4.2](#) dargestellt ist. Vom Berechnen der Flächenkrümmungen ausgehend, werden die einfachen Strukturen der Zusammenhangskomponenten über das Verbinden und Kombinieren mit zusätzlicher Semantik versehen, bis schließlich die Ohrregion markiert werden kann. Die beiden Verzweigungen stellen alternative Konzepte dar, die im sequentiellen Ablauf keine Rolle spielen. Logisch lässt sich die Detektion in drei Abstraktionsebenen gliedern, die die Elemente Komponenten, Shapes und Ohrenkandidaten enthalten. Den Ebenen werden in [Abbildung 4.2](#) die zugehörigen Schritte des Algorithmus zugeordnet.

Ohrenkandidaten stellen dabei die Elemente der obersten Abstraktionsebene dar. Sie beschreiben Strukturen des Bildes, die jeweils als semantisch zusammengehörig angesehen werden, ohne dabei zwangsläufig vollständige Ohrumrisse zu sein. Ohrenkandidaten können damit auch Teile von Ohren oder Strukturen aus anderen Bildregionen sein. Sie werden auf die in [Abschnitt 4.9](#) beschriebene Art auf ihre Ohrähnlichkeit untersucht, so dass eine

4 Konzept zur Ohrendetektion

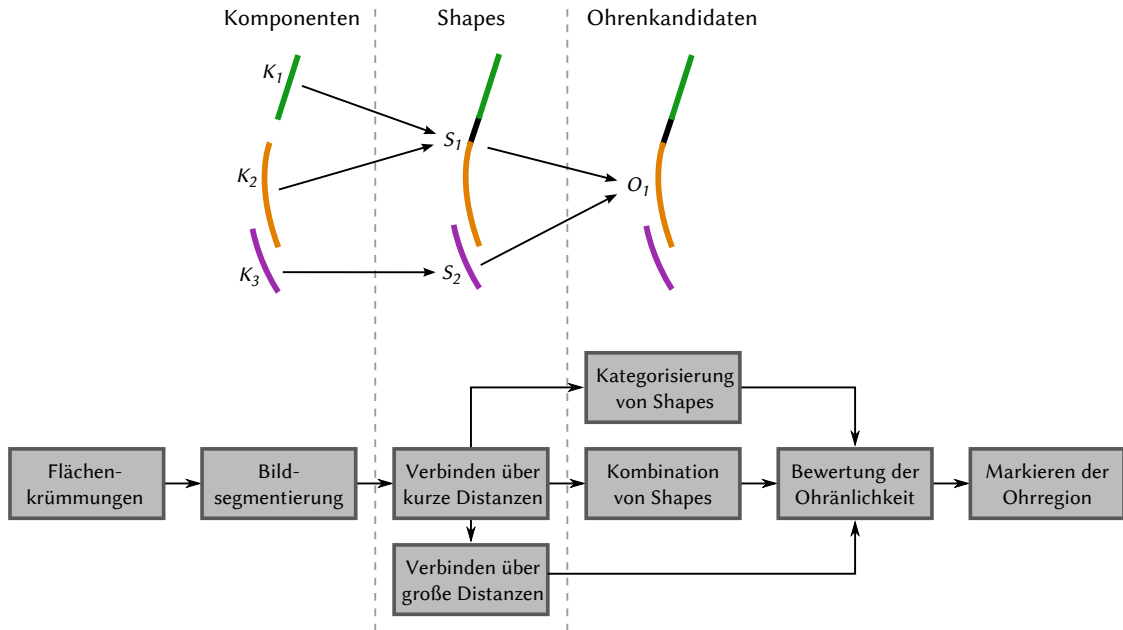


Abbildung 4.2: Darstellung der Einzelschritte im sequentiellen Ablauf der Detektion. Die beiden Verzweigungen stellen alternative Konzepte dar, die im Verlauf dieses Kapitels ebenfalls beschrieben werden. Die drei Ebenen Komponenten, Shapes und Ohrenkandidaten sind den Schritten zugeordnet, bei denen sie erstmals eine wichtige Rolle spielen.

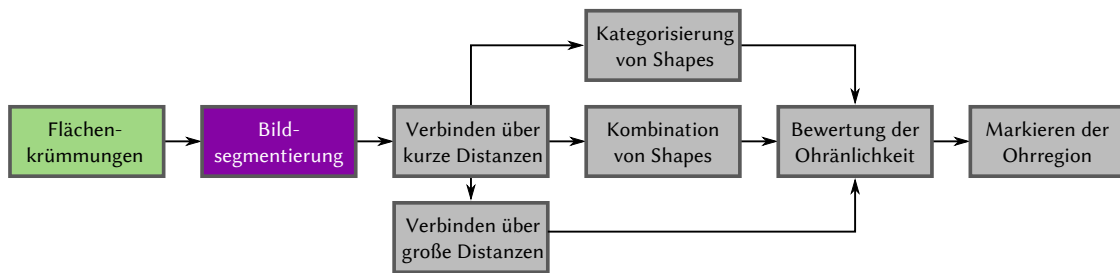
Rangliste der Ähnlichkeiten erstellt werden kann und der Umriss des am besten bewerteten Kandidaten im Ergebnis die Ohrregion bildet (siehe Abschnitt 4.10). Ohrenkandidaten können, wie in Abschnitt 4.8 dargestellt, miteinander kombiniert werden. Dies ist notwendig, da einzelne Kandidaten in der Regel nicht den vollständigen Ohrumriss darstellen, sondern nur Teile davon. Ohrenkandidaten bestehen aus Shapes, die zusammenhängende Regionen sind, also nur direkt benachbarte Punkte enthalten. Shapes wiederum können miteinander verbunden werden, so dass größere zusammenhängende Regionen entstehen, als es mit den Komponenten der Fall wäre, welche die unterste Ebene darstellen. Das Verbinden von Shapes ist in Abschnitt 4.3 beschrieben. Komponenten entsprechen entweder genau den Zusammenhangskomponenten der Bildmatrix oder sind sogenannte Verbindungskomponenten, die zwei Komponenten zu einer größeren Region verbinden. Die Ermittlung der Zusammenhangskomponenten aus der Bildmatrix der Flächenkrümmungen wird in Abschnitt 4.2 erläutert.

Die Elemente, die auf einer Ebene erzeugt wurden, werden von anderen Ebenen aus nicht mehr verändert. So werden beim Verbinden von Shapes nur neue Komponenten erzeugt, während die bereits vorhandenen als Bestandteil der Shapes bestehen bleiben. Ebenso ver-

hält es sich mit den Shapes, wenn Kandidaten, die aus den Shapes bestehen, kombiniert werden. Somit kann beispielsweise von einem Kandidaten ausgehend jederzeit nachvollzogen werden, welche Komponenten aus dem Bild der Flächenkrümmungen stammen und welche im Zuge der Detektion erzeugt wurden.

In den weiteren Abschnitten dieses Kapitels werden die einzelnen Schritte des Ablaufs der Detektion detailliert erläutert. Zu Beginn jedes Abschnittes ist die aktuelle Position im Ablauf des Detektionsalgorithmus zur leichteren Orientierung grafisch dargestellt. Der jeweils aktuelle Schritt im Ablauf wird lila hervorgehoben, während gelbe Hervorhebungen bei Vorverarbeitungsschritten andeuten, für welche Teile des Algorithmus sie notwendig sind.

4.2 BILDSEGMENTIERUNG



Nach dem Berechnen der Flächenkrümmung für jede Koordinate des dreidimensionalen Eingabebildes, durchgeführt nach der Beschreibung in Abschnitt 2.1, können auf diesen Werten einige Bildverarbeitungsschritte durchgeführt werden: die Glättung, zwei Binarisierungen, deren Ergebnisse in Abbildung 4.3 durch zwei verschiedene Farben beispielhaft dargestellt sind, und das anschließende Extrahieren von Zusammenhangskomponenten, auf denen die Ohrendetektion im weiteren Verlauf des Algorithmus arbeiten soll.

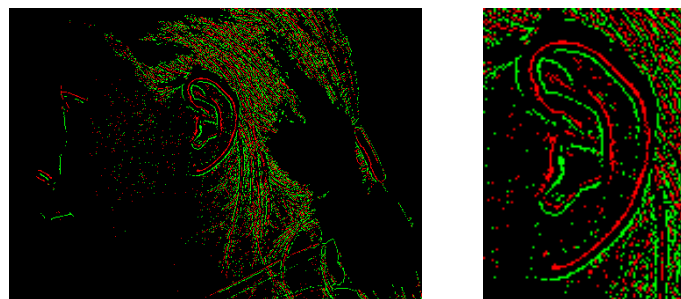


Abbildung 4.3: Binarisierte Bilder der berechneten Flächenkrümmungen in einem Bild dargestellt. Grün entspricht dem Binärbild der Maxima und rot dem der Minima der Flächenkrümmungen.

4.2.1 GLÄTTUNG

Auf die Berechnung der Flächenkrümmungen folgt zunächst eine Glättung der Krümmungswerte durch die Anwendung einer Tiefpass-Filterung. Die Glättung erfolgt durch eine Faltung der Flächenkrümmungswerte mit der Gauß-Funktion. Sie verringert das Rauschen im Bild und damit auch das Auftreten von Artefakten an den Enden der Zusammenhangskomponenten der danach aus den maximalen und minimalen Werten erzeugten Binärbildern, wie in den Ausschnitten (1) und (2) der Abbildung 4.4 zu sehen ist. Ausschnitt (2) zeigt, dass die Glättung außerdem die Anzahl der Lücken verringern kann. Zwar können auch negative Effekte, wie eine Verstärkung von Fragmenten oder das Verschwinden von Objekten auftreten (siehe Ausschnitt (3)), die positiven Effekte überwiegen jedoch deutlich, wie Tests zeigen konnten.

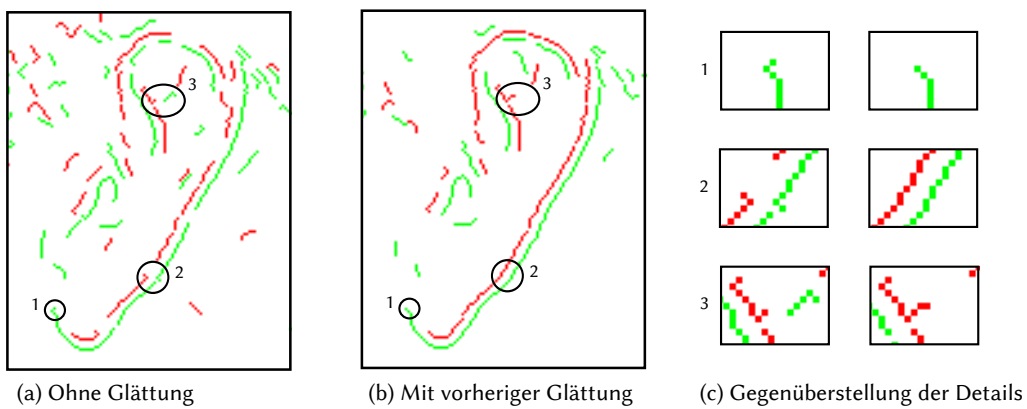


Abbildung 4.4: Auswirkungen der Glättung anhand des Vergleichs eines Binärbildes (a), das ohne vorige Glättung erzeugt wurde und (b), vor dessen Erzeugung ein Tiefpass-Filter angewandt wurde. In (c) sind die in (a) und (b) markierten Ausschnitte gegenübergestellt.

4.2.2 BINARISIEREN

Bei der Binarisierung, die auf die Glättung folgt, gilt folgender Zusammenhang, um den Wert b_{ij} des Pixels an der Position ij des Ausgabebildes bestimmen zu können:

$$b_{ij} = \begin{cases} 1 & k_{ij} \geq S \\ 0 & k_{ij} < S \end{cases} \quad (4.1)$$

Dabei enthält k die Flächenkrümmungen für jede Position ij , d.h. die Ergebnisse der Abbildung der dreidimensionalen Punktwolken auf ein zweidimensionales Bild. Die Flächenkrüm-

mungen nehmen positive und negative Werte an, die getrennt behandelt werden sollen. Deshalb werden sie zunächst in positive und negative Werte aufgeteilt. Die Binarisierung findet dann jeweils auf den Absolutwerten statt, so dass auch die Minima über den Zusammenhang 4.1 ermittelt werden können. Notwendig zur Durchführung der Binarisierung ist die Festlegung eines Schwellwertes S . Dieser kann, wie in Formel 4.1, global für alle Positionen auf den gleichen Wert festgelegt oder adaptiv lokal berechnet werden. Lokal verschiedene Schwellwerte haben den Vorteil selbst bei geringer Ausprägung der Krümmungen Maximalwerte zu finden. Dies kann aber auch von Nachteil sein, da dann selbst in großen Bereichen mit nur sehr geringen Krümmungen viele Bildpunkte den Wert 1 erhalten. Bei der Detektion der Ohren trifft das beispielsweise auf die Wangen des Individuums zu. Lokale Schwellwerte erwiesen sich daher als unbrauchbar, wie auch in Abbildung 4.5(c) zu sehen ist.

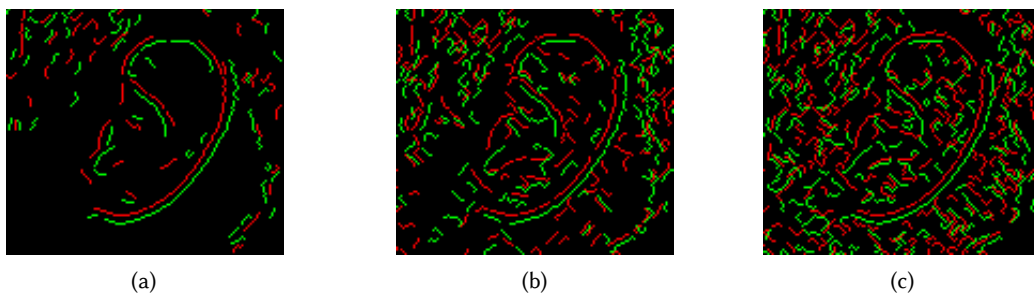


Abbildung 4.5: Vergleich verschiedener Schwellwertverfahren an der Region um ein Ohr: (a) fester globaler Schwellwert von 0,5, (b) Schwellwertverfahren nach Otsu, (c) adaptiver lokaler Schwellwert. Nach der Binarisierung sind Komponenten kleiner als fünf Pixel entfernt und eine Skelettierung durchgeführt worden.

Es bleibt daher nur die Verwendung eines globalen Schwellwertes, der entweder einen festen Wert erhält, der für jedes Bild gleich ist, oder jeweils aus den Eigenschaften der vorhandenen Krümmungswerte berechnet wird. Wie in Abbildung 4.5(b) zu sehen ist, entstehen auch bei der Verwendung eines variablen globalen Schwellwertes deutlich mehr Zusammenhangskomponenten, ohne dass der Ohrumriss vollständiger ist, als bei dem manuell für alle Bilder festgelegten Schwellwert. Daher wird ein fester Schwellwert zum Binarisieren verwendet. Besser als ein einzelner Schwellwert könnte dabei noch ein Hysterese-Verfahren sein, dass mit zwei Schwellwerten arbeitet: einem oberen Schwellwert S_h und einem unteren Schwellwert S_l [NIXO 2002]. Der Pixel an der Position ij wird auf 1 gesetzt, wenn $k_{ij} > S_h$ ist und auf 0 bei $k_{ij} < S_l$. Von allen $k_{ij} > S_h$ aus wird dann nach benachbarten Pixeln gesucht, für die gilt: $S_l \leq k_{ij} \leq S_h$. Diese werden ebenfalls auf 1 gesetzt.

Tests mit verschiedenen Schwellwerten für das Hysterese-Verfahren ergaben die besten Ergebnisse bei $S_h = 0,6$ und $S_l = 0,4$. Diese waren geringfügig besser, als die besten Ergeb-

nisse bei der Verwendung eines einzelnen Schwellwerts, die mit $S = 0,5$ erreicht werden konnten (siehe Abschnitt 8.1 für die Testergebnisse).

Nach der Binarisierung können Lücken an Positionen der Zusammenhangskomponenten des Binärbildes vorhanden sein, die semantisch nicht korrekt sind, wie beispielsweise in Abbildung 4.5(a) im oberen Bereich der Helix zu erkennen ist. Der Umgang mit kleinen Lücken wird in Abschnitt 4.3 beschrieben und große Lücken werden schließlich in Abschnitt 4.8 behandelt.

4.2.3 FINDEN DER ZUSAMMENHANGSKOMPONENTEN

Die Zusammenhangskomponenten der erzeugten Binärbilder werden im nächsten Schritt ermittelt. In dieser Arbeit wird der Begriff Zusammenhangskomponente eingeschränkt definiert, als dies in der Bildverarbeitung normalerweise der Fall ist: Jeder Bildpunkt einer Zusammenhangskomponente, mit Ausnahme der beiden Endpunkte hat exakt zwei Nachbarn. Endpunkte sind Punkte mit nur einem Nachbarn. Die Zusammenhangskomponenten haben also an jeder Position eine Breite von einem Pixel, sie liegen in einer skelettierten Form vor. Ein Vergleich von Zusammenhangskomponenten, die nicht skelettiert sind, und der obigen Definition, ist in Abbildung 4.6 dargestellt.

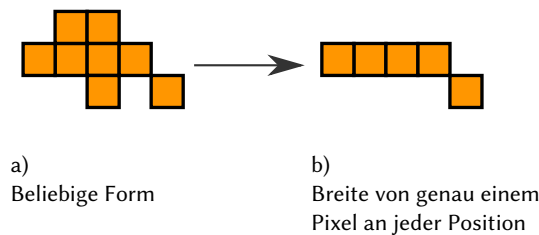


Abbildung 4.6: Zusammenhangskomponenten von a) beliebiger Form und b) dem durch für diese Arbeit definierten Spezialfall.

VORVERARBEITUNG

Damit sichergestellt ist, dass jeder Bildpunkt maximal zwei Nachbarn hat, müssen einige Vorverarbeitungsschritte auf den Binärbildern durchgeführt werden. Das Ausdünnen aller zusammenhängender Regionen zu Linien, also die Durchführung einer Skelettierung, ist zunächst sinnvoll. Dies kann durch eine morphologische Thinning-Operation (siehe Abschnitt 2.2.1) erfolgen. Ein skelettiertes Bild kann Verzweigungen, also Punkte mit mehr als zwei Nachbarn, in den zusammenhängenden Regionen enthalten. Der Verlauf der Krümmungsmaxima der Helix ist grundsätzlich frei von Verzweigungen. Sie können in den Helixre-

gionen daher nur auf Grund von Störfaktoren, wie beispielsweise Haaren auftreten. Alle Verzweigungspunkte werden entfernt, was morphologisch nach der Beschreibung in Abschnitt 2.2.1 erfolgen kann, so dass nach dem Ermitteln der jetzt getrennten Zusammenhangskomponenten, die semantisch zusammengehörenden wieder verbunden werden können. Ein letzter Vorverarbeitungsschritt ist schließlich das Entfernen sehr kleiner Regionen, damit die Laufzeit des in Abschnitt 4.3 beschriebenen Verbindens der Komponenten nicht zu hoch ist. Sehr kleine Regionen werden als Bildrauschen angesehen und können deshalb als nicht zum Ohr gehörend ausgeschlossen werden. Der Unterschied von vor und nach dem Entfernen ist in Abbildung 4.7 beispielhaft zu sehen.

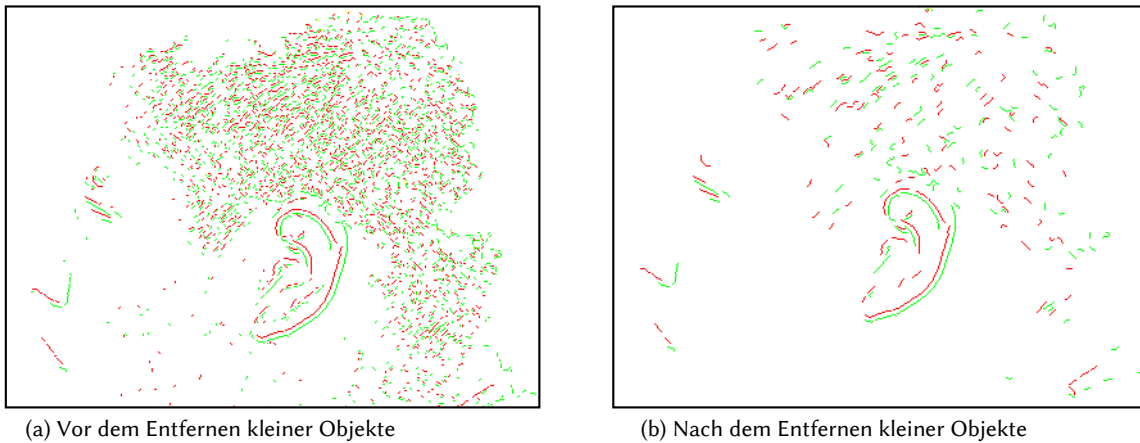


Abbildung 4.7: Beispiel des Entfernens kleiner Komponenten.

KOMPONENTENEXTRAKTION

Für spätere Schritte, wie das Verbinden über kurze Distanzen (siehe Abschnitt 4.3), ist es nötig die Bildkoordinaten jeder Zusammenhangskomponente in der Reihenfolge ihrer Nachbarschaften zu speichern. Nur so können die Krümmungen der Zusammenhangskomponenten berechnet werden, wie es in Abschnitt 2.4 beschrieben wird. Auch kann es nur bei einer definierten Reihenfolge einen Anfang und ein Ende einer Zusammenhangskomponente geben.

Flood-Fill-Algorithmen, zum Extrahieren von Zusammenhangskomponenten, durchlaufen das Bild zeilen- oder spaltenweise, bis sie einen Objektpixel finden, von dem aus das Objekt dann rekursiv oder iterativ geflutet wird, die Objektpixel also markiert werden. Nach dieser Flutung wird die zeilen-, bzw. spaltenweise Suche nach noch nicht markierten Objektpixeln fortgesetzt, bis das Bild vollständig abgesucht wurde. Das Fluten der Zusammenhangskomponenten beginnt dabei entsprechend nicht am jeweiligen Endpunkt, so dass im

Ergebnis die Reihenfolge der Objektpixel nicht von einem Endpunkt zum anderen gegeben ist.

Ein mögliches Verfahren, um die Objektpixel der Zusammenhangskomponenten in ihrer Reihenfolge aus einem Binärbild zu extrahieren, ist in Abbildung 4.8 skizziert. Dabei müssen zunächst die Endpunkte der Zusammenhangskomponenten ermittelt werden. Dies kann durch die Anwendung eines morphologischen Hit-Miss-Operators erfolgen (siehe Abschnitt 2.2.1). Von den Endpunkten aus können die Komponenten dann extrahiert werden. Dazu wird folgende Matrix benutzt, die auch in Abbildung 4.8 als Suchmaske skizziert ist:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (4.2)$$

Die Maske wird, beim Endpunkt beginnend, jeweils zum nächsten durch die Maske gefundenen Punkt verschoben, bis kein weiterer Punkt mehr gefunden werden kann. Nach jedem Schritt wird der zuletzt gefundene Punkt aus dem Bild entfernt. Der letzte gefundene Punkt einer Komponente, der ebenfalls ein Endpunkt ist, wird zuletzt noch aus der Liste zu untersuchender Endpunkte entfernt, damit von ihm aus keine weitere Suche durchgeführt wird.

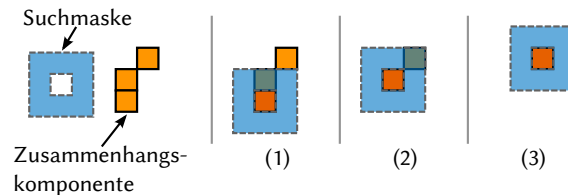


Abbildung 4.8: Extraktion einer Zusammenhangskomponente durch die Verwendung einer Suchmaske für Nachbarpixel.

Eine andere Möglichkeit ist die Verwendung von Boundary-Trace-Algorithmen, die wie in Abbildung 4.9 zu sehen ist, die Pixel am Rand einer Region in ihrer Reihenfolge ermitteln. Zusammenhangskomponenten der für diese Arbeit definierten Art bestehen nur aus dem Objektrand, weshalb ein solcher Algorithmus grundsätzlich sinnvolle Ergebnisse zur Extraktion solcher Zusammenhangskomponenten liefern sollte.

Bei der Verwendung einer solchen Implementierung, ohne Änderungen daran vorzunehmen, ist folgendes zu beachten: Auf Zusammenhangskomponenten mit einer Breite von einem Pixel werden alle Pixel außer den beiden Endpunkten doppelt aufgenommen (siehe Abbildung 4.10), da Border-Trace-Algorithmen terminieren, wenn der nächste Pixel der Startpunkt ist. Nach Entfernen der doppelt gewählten Pixel entspricht das Ergebnis genau dem

4 Konzept zur Ohrendetektion

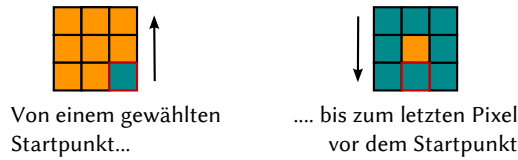


Abbildung 4.9: Boundary-Trace-Algorithmus auf einer Zusammenhangskomponente mit einer Breite von mehr als einem Pixel.

des zunächst erläuterten Verfahrens. Die Verwendung eines Boundary-Trace-Algorithmus hat den Vorteil, dass auf fertige, optimierte Implementierungen zurückgegriffen werden kann und daher keine Optimierungen zum Erreichen einer guten Zeiteffizienz vorgenommen werden müssen.

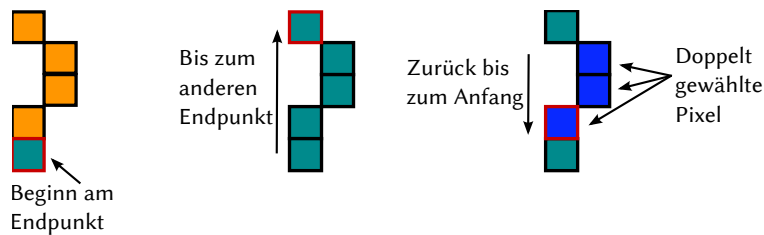
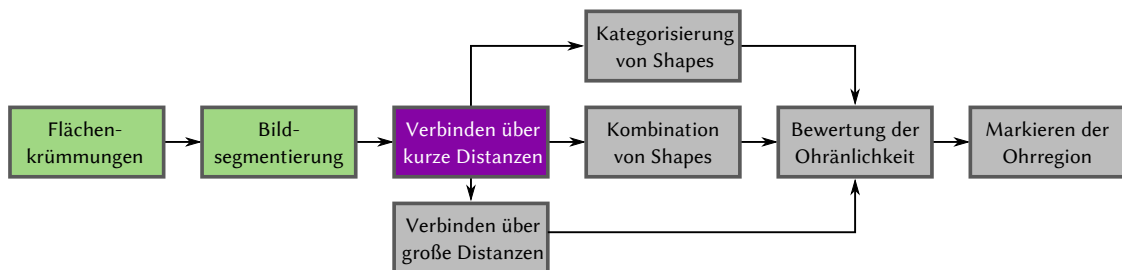


Abbildung 4.10: Boundary-Trace-Algorithmus auf Linien.

4.3 VERBINDEN ÜBER KURZE DISTANZEN



Durch das Binarisieren (siehe Abschnitt 4.2.2) sind nur die Maximal- und Minimalwerte der Krümmungen in den entsprechenden Binärbildern abgebildet. Da die Oberflächenkrümmungen aber beispielsweise durch Störeinflüsse schwanken, weisen die extrahierten Zusammenhangskomponenten Lücken auf. Der nächste Schritt ist daher das Verbinden der Zusammenhangskomponenten über kurze Distanzen. Dabei werden mehrere Zusammenhangskomponenten zu jeweils einem Shape verbunden (siehe Abschnitt 4.1 für die Einführung des Begriffes). Im folgenden wird nur noch von Komponenten gesprochen, wenn

explizit auf die unveränderten Zusammenhangskomponenten des Eingabebildes Bezug genommen werden soll. In allen anderen Fällen wird die Bezeichnung Shape verwendet, auch wenn dieser nur aus einer einzigen Zusammenhangskomponente besteht.

Das Verbinden läuft folgendermaßen ab: Zunächst müssen zwei Eigenschaften von Shapes ermittelt (Abschnitt 4.3.1) und dann mögliche Verbindungen gesucht (Abschnitt 4.3.2) und bewertet (Abschnitt 4.3.3) werden. Zuletzt werden die zu verbindenden Shapes zu einem gemeinsamen Shape zusammengefasst (Abschnitt 4.3.4). Das entwickelte Verfahren zum Verbinden von Shapes über kurze Distanzen ist ein rein lokales Verfahren. Verbindungen werden immer von dem jeweiligen Shape aus in seiner unmittelbaren Nachbarschaft gesucht.

4.3.1 EIGENSCHAFTEN VON SHAPES

Damit das Verbinden von Shapes über kurze Distanzen nach sinnvollen Kriterien erfolgen kann, müssen zunächst zwei Eigenschaften der Shapes ermittelt werden: ihre Ausgangsvektoren und ihre Krümmungen.

AUSGANGSVEKTOREN

Für jeden Shape S ist die Bestimmung seiner Ausgangsvektoren notwendig. Als Ausgangsvektoren werden in dieser Arbeit jene Vektoren bezeichnet, die von den Endpunkten aus die Richtung des Verlaufs des Shapes angeben, wie es in Abbildung 4.11 dargestellt ist, indem der Verlauf des Shapes in der Umgebung der Endpunkte abgeschätzt wird. Die Ermittlung der Ausgangsvektoren erfolgt über die Differenzen der x - und y -Werte des ersten und k -ten Pixels von S , sowie die Differenzen des $(n-k)$ -ten und n -ten Pixels, bei einer Länge des Shapes von n Pixeln:

$$\vec{z}_1 = \begin{pmatrix} x_1 - x_k \\ y_1 - y_k \end{pmatrix} \quad \vec{z}_2 = \begin{pmatrix} x_n - x_{n-k} \\ y_n - y_{n-k} \end{pmatrix} \quad (4.3)$$

Dabei sollte k größer als 3 gewählt werden, da sonst die Abstufungen in den Richtungen unterschiedlicher Ausgangsvektoren auf Grund der ganzzahligen x - und y -Werte sehr grob sind. Je größer k gewählt wird, desto größer wird der Bereich des Shapes, der zur Berechnung herangezogen wird und desto weniger wird entsprechend der Verlauf explizit im Endbereich des Shapes angegeben. Ein Wert von 5 erwies sich als sinnvoll zur Berechnung der Ausgangsvektoren. Anschließend werden die Vektoren \vec{z}_1 und \vec{z}_2 noch normiert:

$$\vec{a}_1 = \frac{\vec{z}_1}{\|\vec{z}_1\|} \quad \vec{a}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} \quad (4.4)$$

4 Konzept zur Ohrendetektion

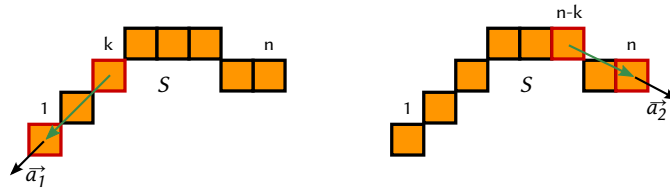


Abbildung 4.11: Ermittlung der beiden Ausgangsvektoren \vec{a}_1 und \vec{a}_2 eines Shapes S über die Differenz der x - und y -Werte der hervorgehobenen Pixel.

Es ist auch möglich die Ausgangsvektoren über die Ermittlung einer Ausgleichsgeraden durch die letzten k Punkte von S zu ermitteln. Es zeigte sich aber, dass das einfache Verfahren des Bildens der Differenzen der Pixelkoordinaten eine ausreichende Genauigkeit für den Einsatz beim Suchen von Verbindungen bietet. Die Berechnung einer Ausgleichsgeraden wird insbesondere auch deshalb nicht durchgeführt, da sie eine Sonderbehandlung des Falls parallel zur y -Achse verlaufender Pixel benötigt und eine Gerade keine Richtung besitzt, diese also noch aus der Steigung und den ursprünglichen Punkten berechnet werden muss.

KRÜMMUNG

Die Krümmungen in jedem Punkt der Shapes werden über das in Abschnitt 2.4 beschriebene Verfahren bestimmt. Anhand der Krümmungswerte kann eine Klassifizierung der Shapes in drei Kategorien erfolgen: nicht gekrümmte Shapes, im mathematisch positiven Sinn gekrümmte Shapes und im mathematisch negativen Sinn gekrümmte Shapes. Die Krümmungsrichtung bezieht sich dabei relativ auf die Betrachtung der Shapes von ihrem ersten zu ihrem zweiten Endpunkt.

Als nicht gekrümmt gilt ein Shape, wenn die Summe seiner positiven Krümmungen und der Betrag der Summe seiner negativen Krümmungen jeweils kleiner als ein definierter Schwellwert sind.

Im mathematisch positiven Sinn gekrümmt ist ein Shape hingegen, wenn die Summe seiner positiven Krümmungen größer und der Betrag der Summe seiner negativen Krümmungen kleiner als der Schwellwert sind. Analog gilt das für negativ gekrümmte Kurven. Sind die positiven Krümmungen und der Betrag der negativen Krümmungen größer als der Schwellwert, so gilt der Shape als gekrümmt in beide Richtungen und wird im weiteren Verlauf nicht mehr berücksichtigt, da Shapes mit solcher Krümmung im Bereich des Ohres nicht auftreten.

4.3.2 SUCHEN MÖGLICHER VERBINDUNGEN

Verbindungen zwischen Shapes können grundsätzlich auf zwei Arten untersucht werden: Zum einem, indem jede mögliche Verbindung zwischen den Shapes betrachtet wird, zum anderen, indem von jedem Shape aus jeweils ein Suchbereich definiert wird und nur Verbindungen zu anderen Shapes betrachtet werden, wenn diese innerhalb des Suchbereichs liegen. Die Untersuchung aller möglichen Verbindungen ist zwar konzeptuell sehr einfach, insbesondere auch, da eine Bewertung von Verbindungen bei der Verwendung eines Suchbereichs ebenfalls erfolgen muss. Die Zeitkomplexität ist mit $O(n^2)$ für n Komponenten jedoch sehr hoch. Sinnvoller ist daher die Definition eines Suchbereichs, um viele mögliche Verbindungen direkt ausschließen zu können.

Dabei werden die Shapes in der Reihenfolge ihrer Längen untersucht, beginnend mit dem längsten Shape, da kurze Shapes wahrscheinlicher durch Rauschen bedingt sind, als dies bei längeren der Fall ist. Durch dieses Vorgehen wird das in Abbildung 4.12 skizzierte Problem umgangen, bei dem eine Verbindung von einem kurzen Shape S_1 ausgehend eine sinnvolle Verbindung von S_3 aus verhindert.

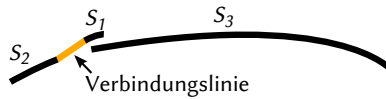


Abbildung 4.12: Ungünstige Verbindung durch Verbinden von S_1 aus. Besser wäre zunächst von S_3 aus eine Verbindung erzeugt worden.

Mit den in Abschnitt 4.3.1 definierten Eigenschaften eines Shapes ist es möglich einen Bereich zu definieren, in dem von dem Endpunkt eines Shapes S_u aus, nach Endpunkten weiterer Shapes gesucht wird. Dieser Suchbereich wird, wie in Abbildung 4.13 dargestellt, durch den Ausgangsvektor \vec{a}_u von S_u und zwei Rotationen dieses Vektors \vec{p}_u und \vec{n}_u um den Endpunkt E_{u_1} definiert. Entgegen der für S_u berechneten Krümmungsrichtung wird dabei deutlich weniger rotiert als in Richtung der Krümmung, um dem angenommenen weiteren Verlauf von S_u Rechnung zu tragen. Die beiden Vektoren \vec{p}_u und \vec{n}_u ergeben sich dabei durch folgenden Zusammenhang:

$$\vec{p}_u = \begin{bmatrix} \cos \frac{9}{10}\theta & v_u \sin \frac{9}{10}\theta \\ -v_u \sin \frac{9}{10}\theta & \cos \frac{9}{10}\theta \end{bmatrix} \cdot \vec{a}_u \quad \vec{n}_u = \begin{bmatrix} \cos \frac{1}{10}\theta & -v_u \sin \frac{1}{10}\theta \\ v_u \sin \frac{1}{10}\theta & \cos \frac{1}{10}\theta \end{bmatrix} \cdot \vec{a}_u \quad (4.5)$$

Der Koeffizient v_u ist dabei aus $\{-1,1\}$ und gibt die Krümmungsrichtung des Shapes S_u an. Mit θ kann der Öffnungswinkel des Suchbereichs global für alle Shapes definiert wer-

4 Konzept zur Ohrendetektion

den. Bei einem nicht gekrümmten Shape ($v_u = 0$), werden stattdessen folgende Rotationen durchgeführt, so dass der Suchbereich in beide Richtungen die gleiche Ausdehnung besitzt:

$$\vec{p}_u = \begin{bmatrix} \cos \frac{1}{2}\theta & \sin \frac{1}{2}\theta \\ -\sin \frac{1}{2}\theta & \cos \frac{1}{2}\theta \end{bmatrix} \cdot \vec{a}_u \quad \vec{n}_u = \begin{bmatrix} \cos \frac{1}{2}\theta & -\sin \frac{1}{2}\theta \\ \sin \frac{1}{2}\theta & \cos \frac{1}{2}\theta \end{bmatrix} \cdot \vec{a}_u \quad (4.6)$$

Der Suchbereich kann als ein Polygon P durch fünf Punkte $\{a_u, p_u, n_u, l_u, r_u\}$ definiert werden. Drei werden durch \vec{a}_u , \vec{p}_u und \vec{n}_u bestimmt, während zwei weitere durch die beiden orthogonal zu \vec{a}_u liegenden Vektoren \vec{l}_u und \vec{r}_u bestimmt werden.

Die beiden orthogonal zum Ausgangsvektor liegenden Punkte werden statt E_{u_1} zur Definition des Polygons herangezogen, damit auch in nächster Nachbarschaft des Endpunktes leicht versetzt liegende Endpunkte anderer Shapes gefunden werden können. Durch die Diskretisierung des Suchpolygons könnten dort sonst Probleme auftreten. Auch ermöglicht der über fünf Punkte definierte Suchbereich im Nahbereich von E_{u_1} das Finden von Endpunkten die etwas versetzt liegen, ohne dass p_u und n_u sehr weit von a_u entfernt sein müssen.

Mit einer Konstanten d kann die Ausdehnung des Suchbereichs festgelegt werden. Die Länge der Vektoren \vec{a}_u , \vec{p}_u und \vec{n}_u beträgt d , während die Vektoren \vec{l}_u und \vec{r}_u eine Länge von $d/2$ aufweisen. Damit ist sichergestellt, dass die Suche mehr in Richtung des Ausgangsvektors von K_u erfolgt, als orthogonal zu diesem.

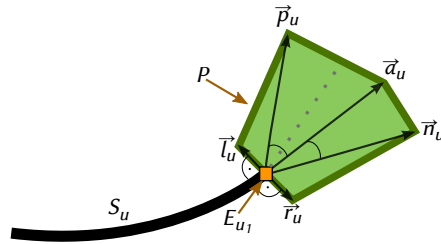


Abbildung 4.13: Definition des Suchbereichs P für Verbindungen über die fünf Vektoren \vec{a}_u , \vec{p}_u , \vec{n}_u , \vec{l}_u und \vec{r}_u . Die gestrichelte Linie gibt den angenommenen weiteren Verlauf von K_u an. Diese Annahme beruht auf der Krümmungsrichtung, die für die Komponente K_u ermittelt wurde.

Für jeden Endpunkt eines anderen Shapes, der innerhalb des Suchpolygons P liegt, muss dann untersucht werden, ob eine Verbindung zwischen diesem und dem Endpunkt von S_u sinnvoll ist. Deshalb werden Metriken zur Bewertung festgelegt.

4 Konzept zur Ohrendetektion

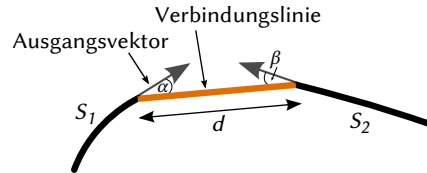


Abbildung 4.14: Metriken beim Verbinden der Shapes S_1 und S_2 zu S_g , der neben den Komponenten der beiden Shapes aus der Verbindungslinie besteht. Dabei sind α und β die Winkel zwischen Verbindungslinie und jeweiligem Ausgangsvektor und d ist die Länge der Verbindung.

4.3.3 METRIKEN ZUR BEWERTUNG EINER VERBINDUNG

Eine sinnvolle Verbindung wird wie folgt definiert: Der durch die Verbindung entstandene Shape S_g soll an den Verbindungspunkten keine Knicke aufweisen, die stärker als ein festgelegter Wert sind. Die Stärke der Knicke ergibt sich dabei aus den in Abbildung 4.14 dargestellten Winkeln α und β . Durch den Suchbereich sind bereits sinnvolle Maxima für den Winkel α und die Distanz d festgelegt, so dass diese nicht mehr untersucht werden müssen. Der Suchbereich ist, wie in Abschnitt 4.3.2 beschrieben, über die Länge und die Rotationswinkel von fünf Vektoren auf sinnvolle Werte begrenzt. Der Winkel β wird berechnet und die Verbindung wird zugelassen, wenn β kleiner oder gleich dem maximal möglichen Wert für α ist. Gilt dies für mehrere mögliche Verbindungen von E_{u_1} aus, so wird eine Rangliste mit Hilfe der Bewertungen B erstellt. B wird für jede mögliche Verbindung nach folgendem Zusammenhang ermittelt:

$$B = d(\alpha + \beta) \quad (4.7)$$

Von allen Verbindungsmöglichkeiten, wird die mit dem kleinsten Wert für B als beste Verbindung gewählt.

Neben der maximalen Distanz d_{max} über die zwei Shapes verbunden werden können, die über die Größe des Suchbereichs festgelegt ist, müssen beim Verbinden zwei weitere harte Kriterien eingehalten werden. Ein Shape darf sich nach der Verbindung mit einem anderen nicht selbst überschneiden und er darf keine geschlossene Form bilden.

4.3.4 ZUSAMMENFASSEN ZU EINEM GEMEINSAMEN SHAPE

Die beiden verbundenen Shapes S_1 und S_2 sollen zu einem gemeinsamen Shape S_g zusammengefasst werden, damit von diesem aus weitere Verbindungen gesucht werden können und nach Abschluss des Verbindens möglichst große Shapes entstanden sind.

Zu diesem Zweck muss zunächst die Verbindungslinie zwischen den Endpunkte von S_1 und S_2 gerastert werden, d.h. die diskreten Pixelkoordinaten müssen berechnet werden. Eine Möglichkeit dazu bietet der Bresenham-Algorithmus (siehe Abschnitt 2.3). Nach dem Rastern können die drei Einzelteile S_1 , S_2 und die Verbindungslinie zu S_g zusammengefasst werden. Dabei ist zu beachten, dass die Reihenfolge der Punkte des Shapes S_2 eventuell umgekehrt werden muss, damit die Pixel von S_g weiterhin in der Reihenfolge ihrer Nachbarschaft betrachtet werden können. Abbildung 4.15 zeigt einen solchen Fall.

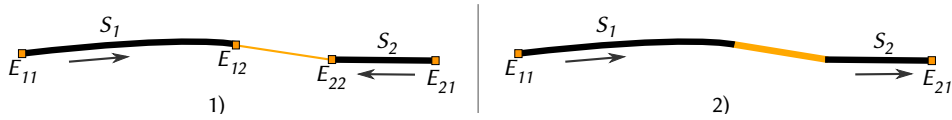


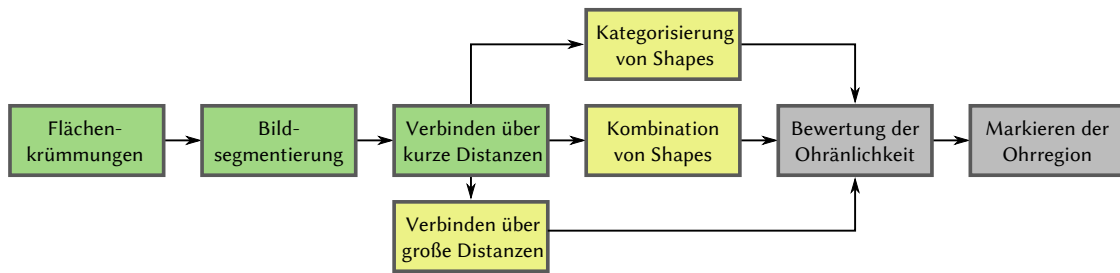
Abbildung 4.15: Beim Verbinden des Endpunktes E_{12} mit E_{21} muss die Pixelliste einer der beteiligten Komponenten umgekehrt werden. Dasselbe gilt analog beim Verbinden von E_{11} und E_{21} .

4.3.5 LIMITIERUNGEN UND PROBLEME BEIM VERBINDEN

Mit Hilfe des Verbindens von Shapes über kurze Distanzen kann der Ohrumriss über kleine Lücken hinweg vervollständigt werden. Das Verbinden ist aber nicht ausreichend, wenn größere Teile der Helix im Bild der Flächenkrümmungen fehlen. Dies kann von einer Überdeckung durch Haare verursacht sein oder auch von direkt neben dem Ohr liegenden Haaren, durch welche die Flächenkrümmungen abschnittsweise geringer sind. Abbildung 4.16(a) zeigt einen solchen Fall.

Drei mögliche Konzepte zur Behebung der Probleme beim Verbinden werden in den folgenden Abschnitten beschrieben und ihre Tauglichkeit wird untersucht. Zunächst sind aber noch zwei Zwischenschritte notwendig, um einige Voraussetzungen für diese Konzepte und nachfolgende Schritte zu schaffen: die in Abschnitt 4.4 beschriebene Auswahl von Kandidaten und die im darauf folgenden Abschnitt 4.5 beschriebene Untersuchung von Parallelität.

4.4 AUSWAHL EINIGER PROMINENTER SHAPES



Die Auswahl einiger Shapes aus der Menge aller Shapes ist ein notwendiger Vorbereitungsschritt für alle drei alternativen Konzepte für das weitere Vorgehen. Die auf diesen Konzepten basierenden Algorithmen weisen eine hohe Zeitkomplexität auf, weshalb eine Beschränkung auf wenige Shapes sinnvoll ist.

Dazu muss zunächst festgelegt werden, welche Shapes für die Auswahl in Frage kommen, was also einen Shape, der einen Teil des Ohres repräsentiert, von anderen Shapes unterscheidet. Bei der Betrachtung des Verlaufs von Helix und Anti-Helix fällt auf, dass diese Strukturen grundsätzlich nur in eine Richtung gekrümmt sind, sieht man von leichten Krümmungsänderungen im Übergang von der Helix zum Ohrläppchen ab. Gerade Shapes und Shapes, die starke Krümmungen in beide Richtungen aufweisen, können also bereits ausgeschlossen werden. Grundannahme des gesamten Konzeptes ist außerdem, dass durch die Flächenkrümmungen des Ohrs prominente, lange Shapes entstehen. Nach der Durchführung des in Abschnitt 4.3 beschriebenen Verbindens sollten die Shapes des Ohres demnach zu den längsten Shapes im Bild gehören.

Deshalb werden die zehn längsten, in eine Richtung gekrümmten Shapes ausgewählt. Sie werden im weiteren Verlauf als prominente Shapes bezeichnet. Die im folgenden beschriebenen Konzepte werden nur noch ausgehend von den prominenten Shapes ausgeführt.

Bei der Auswahl der prominenten Shapes könnte auch das Auftreten von Parallelität (siehe nächsten Abschnitt 4.5) als ein weiteres Kriterium berücksichtigt werden. In Tests (siehe Abschnitt 8.2) ergab sich dadurch aber eine Verschlechterung der Detektionsergebnisse. Eine andere untersuchte Erweiterung stellt das Aufspalten der prominenten Shapes an detektierten Ecken dar. Starke Ecken sollten im Ohrumriss nicht auftreten, sie können aber beispielsweise durch Haare entstehen, deren Flächenkrümmungen mit einem Teil der Helix eine gemeinsame Zusammenhangskomponente bilden. Das Aufspalten an Ecken sorgte weder für eine Verbesserung noch für eine Verschlechterung der Detektionsergebnisse. Das Aufspalten an Wendepunkten im Verlauf der Krümmungen, durchgeführt nur bei starken gegensätzlichen Krümmungen, ist eine weitere Möglichkeit Shapes vor der Auswahl der

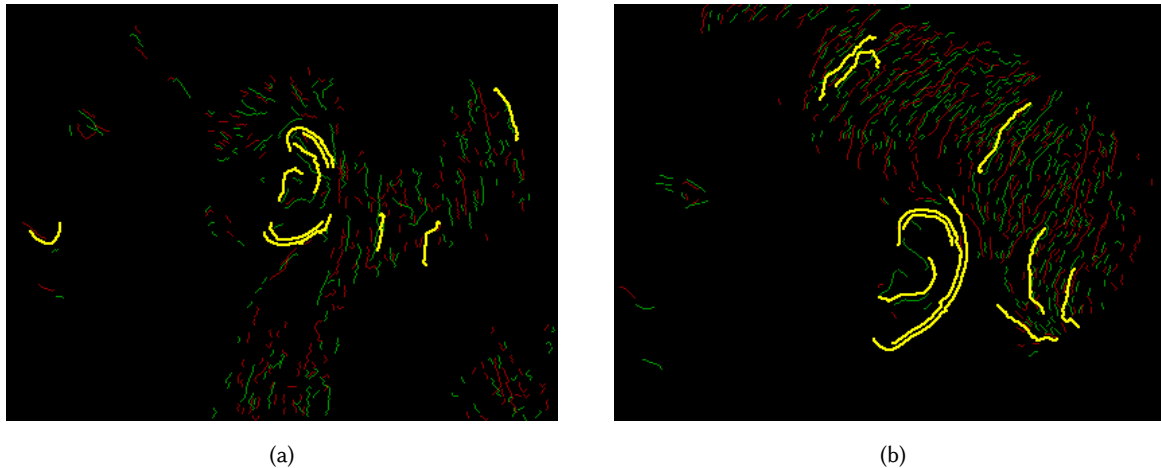
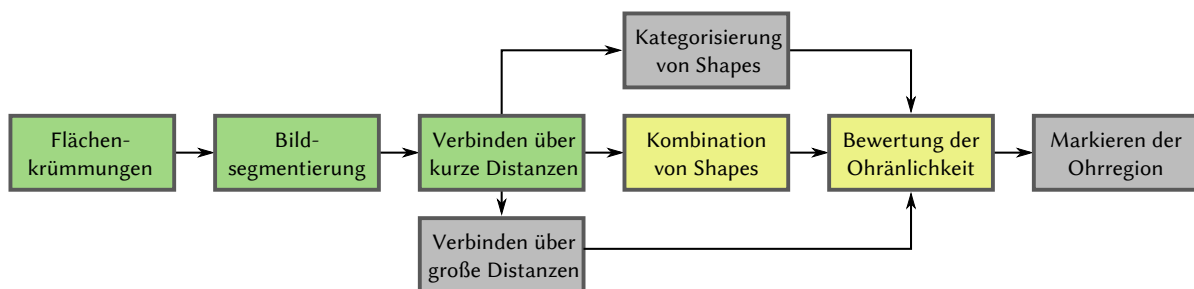


Abbildung 4.16: Beispielbilder gewählter prominenter Shapes (dargestellt in gelb). Im Hintergrund sind die verbundenen Shapes der negativen und positiven Oberflächenkrümmungen zu sehen.

prominenten Shapes zu bearbeiten, da Krümmungsänderungen nicht im Verlauf des Ohrumrisses auftreten sollten, sondern dort nur durch Störeinflüsse entstehen können. Hiermit konnte eine Verbesserung erreicht werden, die allerdings sehr gering war. Die Testergebnisse (siehe Abschnitt 8.2) zeigen, dass die Erweiterungen bei der Auswahl prominenter Shapes keinen großen Einfluss auf die Detektionsrate haben.

4.5 UNTERSUCHUNG VON PARALLELITÄT

Das Untersuchen der Parallelität von Shapes ist kein separater Schritt im sequentiellen Ablauf des Algorithmus. Parallelität wird aber bei der Kombination von Shapes und der Bewertung der Ohrähnlichkeit, die in den Abschnitten 4.8 und 4.9 beschrieben werden, als wichtiges Kriterium verwendet.



Die anatomische Struktur der Ohrmuschel weist parallel verlaufende Maxima und Minima der Flächenkrümmungen insbesondere im Bereich der Helix auf, wie in Abbildung 4.17

zu sehen ist. Auch verlaufen Helix und Anti-Helix tendenziell parallel [JAIN 2007]. Beim Suchen des Ohrumrisses sollte die Parallelität daher berücksichtigt werden, insbesondere auch zur Unterscheidung der Shapes des Ohres von anderen Shapes, beispielsweise von Haaren.

Im Folgenden werden zwei Möglichkeiten zur Untersuchung der Parallelität vorgestellt. Bei der Überprüfung paralleler Abschnitte werden je zwei Shapes auf gegenseitige Parallelität untersucht, während beim Suchen paralleler Shapes von je einem Shape aus nach parallelen Shapes gesucht wird.

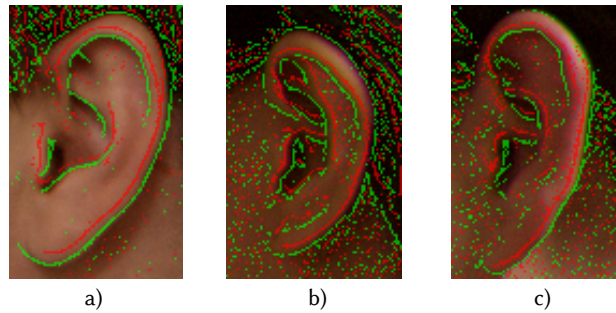


Abbildung 4.17: Darstellung der konvexen Minima (rot) und der konkaven Maxima (grün) der Flächenkrümmungen. Zur besseren Veranschaulichung sind sie auf den zugehörigen zweidimensionalen Aufnahmen dargestellt, die vom Algorithmus nicht verwendet werden.

4.5.1 ÜBERPRÜFUNG PARALLELER ABSCHNITTE

Die Suche paralleler Shapes wird für jeden Shape S_i aus der Menge der Shapes aller Kandidaten S durchgeführt, indem die Parallelität zwischen $S_i \in S$ und jeweils einem anderen Shape $S_p \in S$ untersucht wird. Auch hier ist es aber, wie beim Suchen von Verbindungsmöglichkeiten in Abschnitt 4.3, nur sinnvoll Shapes in unmittelbarer Nachbarschaft zu untersuchen. Dies vermeidet die Untersuchung unnötig vieler Kombinationen von Shapes. Deshalb wird ein achsenparalleles Rechteck um S_i definiert, dessen Ränder von jedem Punkt von S_i mindestens D Pixel entfernt sind, wobei D die maximale Distanz zwischen zwei Punkten eines Shapes bezeichnet, bis zu der auf Parallelität untersucht wird. Nur Shapes innerhalb dieses Rechtecks werden zur Suche paralleler Shapes von S_i herangezogen.

Die Untersuchung von S_i und einem potentiell parallelen Shape S_p erfolgt dabei nach folgendem Schema: Zunächst muss herausgefunden werden ob die beiden Shapes überhaupt nebeneinander verlaufen und wenn ja, in welchem Bereich dies der Fall ist. Dazu wird von jedem Endpunkt von S_i und S_p eine Untersuchung auf einen Schnittpunkt mit dem jeweils anderen Shape durchgeführt. Beispiele dafür sind in Abbildung 4.19 dargestellt. Zwei Shapes

4 Konzept zur Ohrendetektion

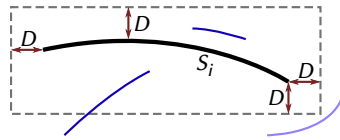


Abbildung 4.18: Achsenparalleles Rechteck (gestrichelt dargestellt) zur Einschränkung des Suchbereichs. Die minimale Distanz zu jedem Punkt von S_1 beträgt D . Nur die dunkelblau dargestellten Shapes werden auf Parallelität zu S_1 untersucht.

können nur dann parallel sein, wenn Schnittpunkte von mindestens zwei Endpunkten aus gefunden wurden. Schnittpunkte bei mehr als zwei Endpunkten sind nur bei dem Spezialfall möglich, dass zwei Endpunkte sich exakt orthogonal zueinander befinden.

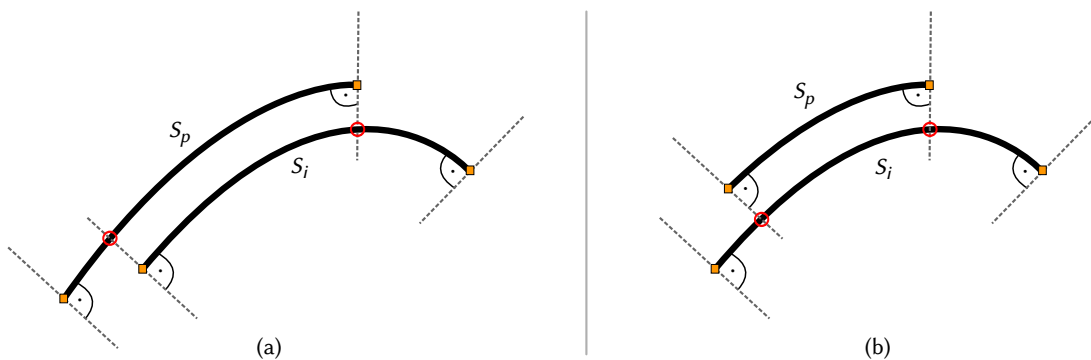


Abbildung 4.19: Bestimmung potentiell paralleler Abschnitte der Shapes S_p und S_i über das Suchen von Schnittpunkten in orthogonaler Richtung der Endpunkte.

Wurden die potentiell parallelen Abschnitte der beiden Shapes ermittelt, kann für jeden Pixel k aus der Menge aller Pixel in dem Abschnitt von S_i der Abstand d zu seinem orthogonalen Nachbarn bestimmt werden. Am Beispiel eines Pixels k ist dies in [Abbildung 4.20](#) dargestellt.

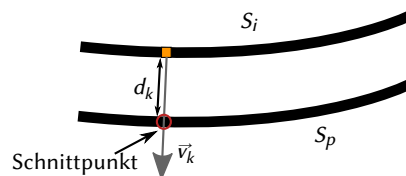


Abbildung 4.20: Bestimmung des orthogonalen Abstands d_k zu einem Shape S_p von einem Pixel k des Shapes S_i aus.

Über alle Abstände wird dann mit Hilfe des Mittelwerts μ der Abstände die Standardabweichung σ gebildet. Die Shapes S_i und S_p gelten als parallel bei einer Standardabweichung unter einem festgelegten Schwellwert G .

$$\sigma_{ip} = \sqrt{\sum (d_k - \mu)^2} \quad (4.8)$$

Sind zwei abschnittsweise parallele Shapes in unterschiedliche Richtungen gekrümmt, wie es in Abbildung 4.21(b) dargestellt ist, können sie nicht beide zum Ohrumriss gehören, da für diesen die Beschränkung auf eine Krümmungsrichtung definiert ist (siehe Abschnitt 4.4). Deshalb müssen die Krümmungsrichtungen von S_i und S_p untersucht werden. Die Krümmungen von Shapes können nur relativ zur Reihenfolge der Pixel bestimmt werden und die Leserichtung ist nicht global festgelegt. Da ihre Leserichtung entgegengesetzt sein kann, können zwei Shapes nicht direkt hinsichtlich ihrer Krümmungsrichtungen verglichen werden, wie auch Abbildung 4.21 verdeutlicht. Für die Untersuchung der Parallelität wurde aber bereits für beide Shapes das Intervall bestimmt, in denen S_i und S_p nebeneinander verlaufen (siehe Abbildung 4.19). Anhand der Intervallgrenzen lässt sich die Leserichtung der beteiligten Shapes leicht angleichen, indem die Reihenfolge der Pixel beider Shapes von einer Grenze zur anderen verglichen wird. Werden dabei die Pixel beider Shapes in der gleichen Reihenfolge durchlaufen, ist die Leserichtung gleich. Im anderen Fall ist die Leserichtung von S_i und S_p unterschiedlich und die Krümmungsrichtung von S_p wird für den Vergleich invertiert.

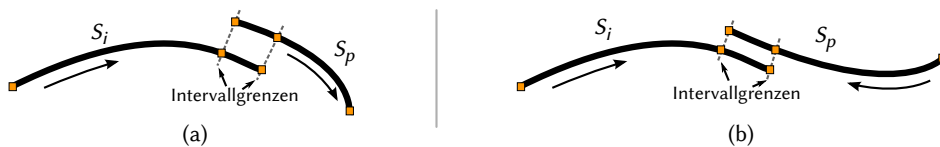


Abbildung 4.21: Abschnittsweise parallele Shapes, die bezüglich der Reihenfolge ihrer Pixel jeweils im mathematisch negativen Sinn, global betrachtet in (b) aber entgegengesetzt gekrümmt sind. Die Reihenfolge der Pixel ist für jeden Shape durch einen Pfeil angedeutet. Zusätzlich sind die Grenzen der nebeneinander verlaufenden Bereiche der Shapes eingezeichnet.

4.5.2 ALTERNATIVES SUCHEN PARALLELER SHAPES

Eine andere Möglichkeit zu der im vorigen Abschnitt beschriebenen Untersuchung der Parallelität durch den Vergleich von jeweils zwei Shapes, besteht in der Suche nach parallelen Shapes ausgehend von einem Shape S_i , wie es in Abbildung 4.22 dargestellt ist. Dazu wer-

4 Konzept zur Ohrendetektion

den von S_i , der aus dem Binärbild B_1 stammt, Nachbarpixel im Binärbild B_2 gesucht, ohne dass diese als Teile von Shapes oder Komponenten angesehen werden. Für jeden gefundenen Nachbarpixel wird die Distanz zu S_i berechnet. Über alle ermittelten Distanzen kann dann die Standardabweichung gebildet werden, um zu überprüfen, ob die gefundenen Nachbarpixel parallel zu S_i verlaufen. Dies ließe sich effizienter implementieren, da nicht in einem Bereich explizit jeder Shape mit jedem verglichen werden müsste, sondern von jedem Shape ausgehend nur eine Suche ausgeführt würde. Problematisch ist dabei der in Abbildung 4.22 skizzierte Fall nicht zusammenhängender Nachbarpixel, da diese zu großen Unterschieden in den ermittelten Distanzen und damit einer hohen Standardabweichung führen können, auch wenn die beiden Shapes, zu denen die Pixel gehören, jeder für sich parallel zu S_i wäre. Deshalb wird die in Abschnitt 4.5.1 beschriebene Methode bevorzugt, die außerdem direkt den zugehörigen parallelen Shape ermittelt. Bei der in diesem Abschnitt beschriebenen Methode hingegen müssen nach dem Feststellen von Parallelität noch die beteiligten Shapes aus den gefundenen Nachbarpixeln ermittelt werden.

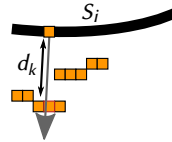
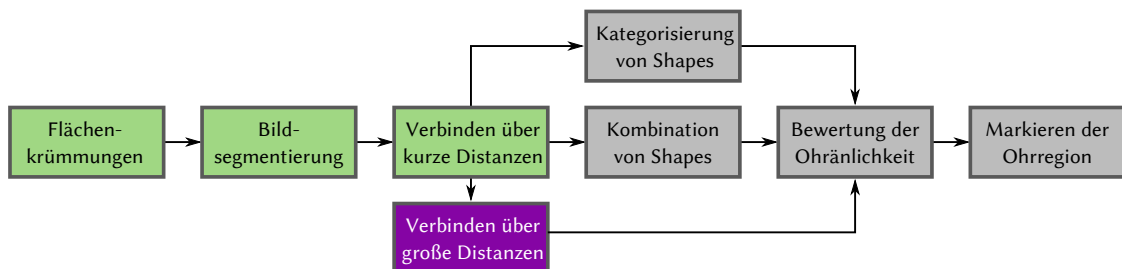


Abbildung 4.22: Suchen paralleler Shapes.

4.6 ERWEITERUNG DES VERBINDENS



Als ein einfacher Ansatz zur Behebung der in Abschnitt 4.3.5 beschriebenen Limitierungen beim Verbinden über kurze Distanzen erscheint die Erweiterung des Verbindens auf größere Distanzen, die in diesem Abschnitt beschrieben wird.

4.6.1 VERBINDEN ÜBER GRÖßERE DISTANZEN

Größere Lücken in den Zusammenhangskomponenten der Helix können überbrückt werden, indem die Begrenzung auf kurze Distanzen bei dem in Abschnitt 4.3.2 beschriebenen Verbinden von Shapes aufgehoben wird. Der lokale Ansatz des Verbindens wird dabei beibehalten, die Nachbarschaft ist nur deutlich größer definiert. Das erweiterte Verbinden wird dabei nicht für alle Shapes durchgeführt, sondern nur von den in Abschnitt 4.4 gewählten prominenten Shapes aus. Ansonsten würden zu viele mögliche Verbindungen existieren, um das Verbinden mit einer vertretbaren Laufzeit durchführen zu können.

Die beim Verbinden über kurze Distanzen verwendeten Kriterien sind bei der Erweiterung auf größere Distanzen nicht ausreichend, um jeweils die sinnvollste Verbindung zu wählen, weshalb eine Anpassung notwendig ist. Da das Ziel des Verbindens die Konstruktion eines vollständigen Ohrumrisses ist, sollte sich die Gesamtkrümmung des Shapes bis ungefähr 2π erhöhen. Vor dem Verbinden zweier Shapes wird deshalb zusätzlich untersucht, ob sich die Gesamtkrümmung durch die Verbindung erhöht, ohne dabei 2π deutlich zu überschreiten.

Das Verbinden über größere Distanzen teilt mit dem Verbinden über kurze Distanzen zwei Limitierungen, die in Abbildung 4.23 skizziert sind. Stark versetzt oder abschnittsweise nebeneinander verlaufende Shapes können nicht zu einem Ohrumriss verbunden werden und Artefakte an den Enden der Shapes können das Verbinden verhindern.

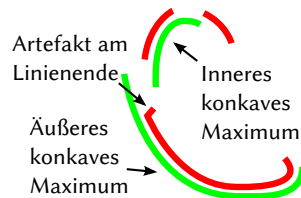


Abbildung 4.23: Probleme beim Verbinden von Shapes. Ein Artefakt verhindert das Verbinden der konvexen Flächenkrümmungsmaxima, während die konkaven abschnittsweise parallel verlaufende Shapes ebenfalls nicht verbunden werden können.

4.6.2 DER SHAPE-INDEX ALS WEITERES FLÄCHENMAß

Der Shape-Index ist ein Maß zur Kategorisierung von Krümmungen, das in Abschnitt 2.1.1 kurz beschrieben und auch in [ZHOU 2010] benutzt wird. Er kann ein sinnvolles Maß zur Vervollständigung des Ohrumrisses sein, da die Oberfläche der Helix durch genau zwei Ka-

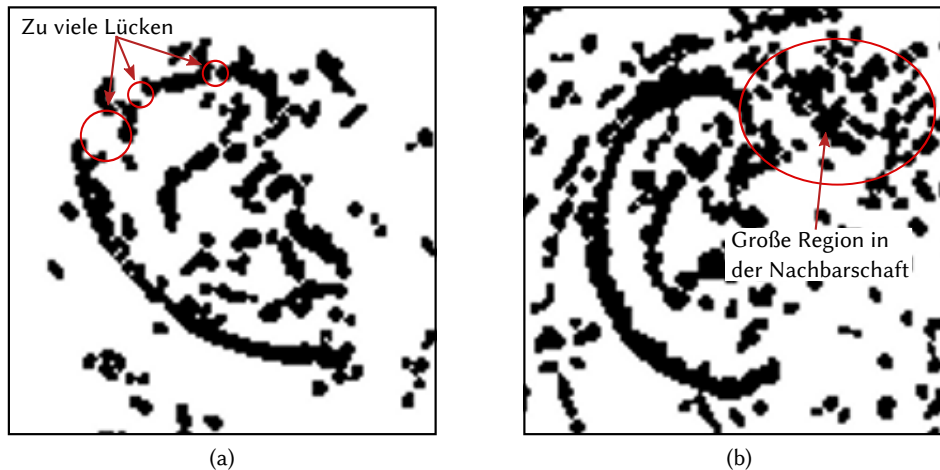


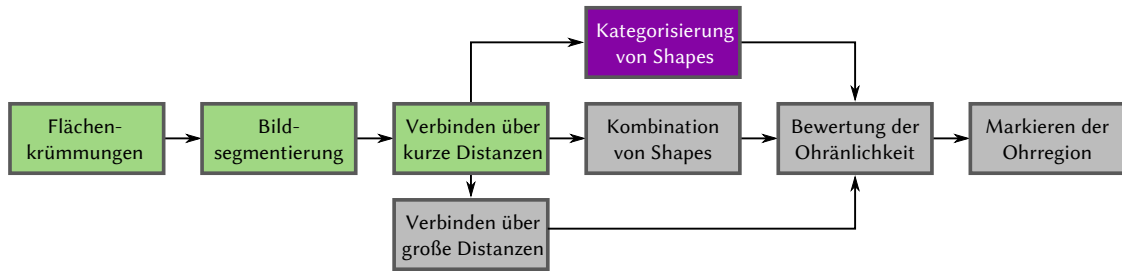
Abbildung 4.24: Beispiele für ungünstige Shape-Index-Regionen der Kategorie „Ridge“. In Bild (a) kann das Ohr nicht vollständig rekonstruiert werden, da zu viele Lücken existieren, während in (b) möglicherweise zu viele Komponenten zum Ohr gezählt werden.

tegorien des Shape-Index gebildet wird: Ridge und Rut. D.h. das Auftreten dieser beiden Kategorien kann neben den Krümmungsmaxima als weiteres Indiz für die Helix gelten.

Es könnten also Regionen bestimmter Shape-Index-Werte zur Vervollständigung der Helix verwendet werden. Beispielsweise indem Komponenten zum Ohr gezählt werden, wenn sie innerhalb der Region liegen, in der auch der prominente Shape zu finden ist, der als Teil der Helix ausgewählt wurde. Da auch die Shape-Index-Regionen Lücken aufweisen können werden die direkt benachbarten Regionen ebenfalls verwendet. Dabei treten aber Probleme auf, die sich gut an zwei Beispielen verdeutlichen lassen, die in [Abbildung 4.24](#) dargestellt sind. Es gibt grundsätzlich zwei ungünstige Fälle des Auftretens von Shape-Index-Regionen: zu große Lücken im Bereich der Helix und zu große Shape-Index-Regionen. Im ersten Fall kann trotz Verwendung der Shape-Index-Regionen nicht der vollständige Ohrumriss rekonstruiert werden, während im zweiten Fall zu viele Shapes zum Ohrumriss gerechnet werden.

Die Verwendung des Shape-Indexes kann demnach keine sinnvollen Informationen zur Vervollständigung des Ohrumrisses beitragen. Das Verbinden über größere Distanzen gemeinsam mit der Verwendung von Shape-Index-Regionen genügt daher in vielen Fällen nicht zur erfolgreichen Detektion der Ohrregion. Die Probleme des Konzeptes des erweiteren Verbindens zeigen, dass eine Bewertung gefundener und verbundener Strukturen notwendig ist, um sinnvolle Ergebnisse erhalten zu können.

4.7 KATEGORISIERUNG VON KANDIDATEN



Damit die prominenten Shapes sinnvoll mit weiteren Shapes zu vollständigen Umrissen von Ohren erweitert werden können, ist es nützlich wenn bekannt ist, welchen Teil des Ohres der jeweilige prominente Shape darstellt. Das heißt wenn der prominente Shape einer bestimmten Position eines zuvor erstellten, möglichst allgemeingültigen Modells des Ohres zugeordnet werden kann. Mit dieser Information kann der Bereich bestimmt werden, in dem nach weiteren Shapes gesucht wird. Wenn keine weiteren Shapes gefunden werden, kann die Größe des Ohres auf diese Weise zumindest abgeschätzt werden. Zunächst wird dazu die Ebene der Ohrenkandidaten eingeführt. Ohrenkandidaten bestehen aus einem oder mehreren Shapes. Diese Ohrenkandidaten sollen dann einer von vier Kategorien zugewiesen werden, die mit ihren Kriterien in Abbildung 4.25 dargestellt sind.

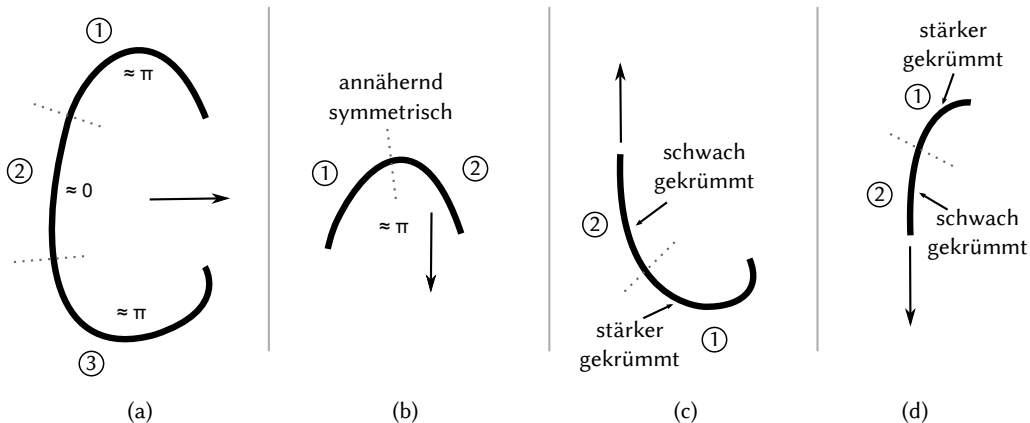


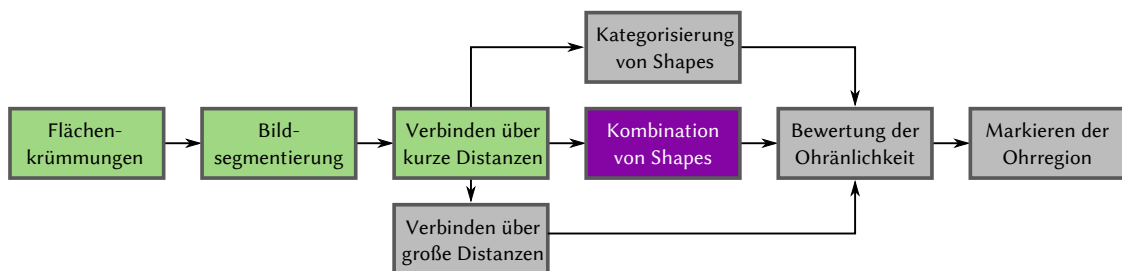
Abbildung 4.25: Die vier Kategorien der Ohrenkandidaten. Die jeweils vertikal gespiegelten Strukturen gehören in die selben Kategorien. Sie müssen aber nicht extra betrachtet werden, da bei der Definition der Kategorien keine Unterscheidung zwischen rechts und links, sondern nur zwischen oben und unten vorgenommen wird. Der Pfeil in Bild (a) symbolisiert nur eine horizontale Öffnungsrichtung.

Kategorie (a) soll dabei vollständige Ohren beschreiben. Die Gesamtkrümmung von ungefähr 2π unterteilt sich bei der Aufteilung in drei gleich große Abschnitte folgendermaßen: in

den Abschnitten (1) und (3) jeweils ungefähr π , während in Abschnitt (2) fast keine Krümmungen auftreten. Die Öffnung zeigt zur Seite, wobei nicht zwischen links oder rechts unterschieden wird. Die Kategorien unvollständiger Ohren unterteilen den Kandidaten jeweils in zwei gleich große Abschnitte. Bei der Kategorie (b), dem oberen Teil des Ohres, sollten die Krümmungen annähernd symmetrisch, sowie gleichmäßig verteilt und die Öffnung nach unten gerichtet sein. Während die Krümmungen der beiden Abschnitte der Kategorien (c) und (d) jeweils nicht symmetrisch sein sollten, bei entgegengesetzter Richtung des Ausgangsvektors des geringer gekrümmten Bereichs.

Problematisch ist bei dieser Herangehensweise insbesondere, dass Umrisse die in Kategorie (c) fallen müssten sehr unterschiedlich aussehen können. Beispielsweise kann das Ohr läppchen frei oder angewachsen sein, was in deutlich unterschiedlichen Shapes für diesen Bereich resultiert. Insgesamt stellte sich heraus, dass die Kategorisierung auf Grund der starken Variationen in der Form der Ohren verschiedener Personen, nicht zuverlässig funktionieren kann. Krümmungen und Proportionen von Ohrenkandidaten eignen sich dennoch gut als Kriterien zur Unterscheidung von Ohren und anderen Strukturen, wenn sie Bestandteil umfangreicherer Bewertungskriterien sind, wie in Abschnitt 4.9 gezeigt wird.

4.8 KOMBINATION VON OHRENKANDIDATEN



Die Kombination von Ohrenkandidaten ist das letzte der drei in dieser Arbeit vorgestellten Konzepte zur Vervollständigung des Ohrumrisses. Sie arbeitet wie auch bereits die Kategorisierung von Shapes (siehe Abschnitt 4.7) auf der Ebene der Ohrenkandidaten. Das Kombinieren hebt die am Ende von Abschnitt 4.3 beschriebenen Limitierungen beim Verbinden über kurze Distanzen auf, indem das Verschieben und Kürzen von Shapes ermöglicht wird. Ein Beispiel ist in Abbildung 4.26 skizziert. Außerdem wird die Kombination als globales Verfahren entwickelt, d.h. es wird nicht nur in der Nachbarschaft eines Shapes, sondern bezogen auf das gesamte Bild nach Kombinationen gesucht. Durch das Kombinieren von Shapes entstandene Shapes S_k werden als angepasste Shapes bezeichnet, um sie von Shapes zu unterscheiden, die nur durch Verbinden entstanden sind. Im Zusammenhang mit Ohren-

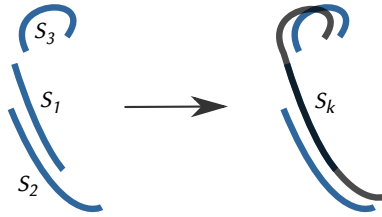


Abbildung 4.26: Kombinieren des Shapes S_1 mit S_2 und S_3 zum gemeinsamen Shape S_k . Die ursprünglichen Shapes bleiben dabei als Original-Shapes unverändert vorhanden, während S_k den neuen angepassten Shape des Ohrenkandidaten darstellt.

kandidaten wird der Begriff Original-Shapes des Kandidaten verwendet, wenn die Shapes gemeint sind, die zu S_k kombiniert wurden.

4.8.1 KOMBINATION PARALLELER SHAPES

Über die in Abschnitt 4.5 beschriebene Parallelitätsuntersuchung können die parallelen Shapes eines Kandidaten ermittelt werden. Diese Shapes werden dann zur weiteren Vervollständigung des Ohrumrisses genutzt. Die Kombination eines zu einem Original-Shape des Ohrenkandidaten parallelen Shapes kann zu einem vollständigeren Ohrumriss führen, als die Verwendung nur eines Shapes, da sich der parallele Verlauf häufig nur über einen Teil der Shapes erstreckt. In den anderen Bereichen können zusätzliche Informationen über den Verlauf der Helix gewonnen werden.

Parallele Shapes werden über das gleiche Verfahren zu einem gemeinsamen Shape kombiniert, das bei dem im nächsten Abschnitt beschriebenen Kombinieren nebeneinander angeordneter Shapes zum Einsatz kommt. Parallele Shapes sind nur ein Spezialfall nebeneinander angeordneter Shapes, der getrennt betrachtet wird, da bei parallelen Shapes in jedem Fall davon ausgegangen werden kann, dass sie semantisch zusammengehören.

4.8.2 WEITERE KOMBINATIONEN VON SHAPES

Bei der Kombination von Shapes, die nicht parallel zueinander verlaufen, sind drei Fälle zu unterscheiden, die jeweils getrennt behandelt werden müssen (siehe Abbildung 4.27): voreinander angeordnete Shapes, nebeneinander angeordnete, sowie Shapes die nicht auf eine der beiden Arten zueinander liegen und für die kein Versuch der Kombination unternommen werden soll, da der Ohrumriss durch sie nicht sinnvoll erweitert werden kann.

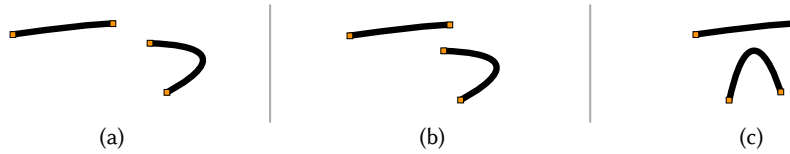


Abbildung 4.27: Vergleich der möglichen Positionierungen von Shapes. (a) Voreinander, (b) Nebeneinander, (c) keine sinnvolle Positionierung

Die Unterscheidung voreinander angeordneter Shapes S_1 und S_2 (siehe Abbildung 4.27(a)) von den anderen beiden Möglichkeiten kann über den in Abbildung 4.28(1) skizzierten Winkel α erfolgen. Dieser Winkel entspricht dem in Abschnitt 4.8.1 beschriebenen Winkel α , er wird hier aber zu einem anderen Zweck verwendet und nur für ein Endpunktpaar ermittelt. Es werden die beiden Endpunkte der beteiligten Shapes gewählt, welche die geringste Distanz zueinander aufweisen. Der Winkel α kann dann zwischen dem Ausgangsvektor \vec{a} von S_2 und dem Verbindungsvektor \vec{v} zwischen den beiden Endpunkten berechnet werden (siehe Abbildung 4.28(1)). Ein Winkel von $\alpha \leq \pi/2$ bedeutet eine voreinander liegende Anordnung, während $\alpha > \pi/2$ für die anderen beiden Fälle steht. Zur weitergehenden Unterscheidung der Positionierung von Shapes nebeneinander und dem Fall, dass sie nicht sinnvoll zueinander positioniert sind, wird das in Abbildung 4.28(2) skizzierte Verfahren angewandt. Ausgehend vom Endpunkt des Shapes S_1 wird orthogonal nach einem Schnittpunkt mit S_2 gesucht. Kann ein Schnittpunkt ermittelt werden, so sind die Shapes nebeneinander angeordnet. Wenn kein Schnittpunkt existiert, sind die Shapes nicht sinnvoll zueinander positioniert.



Abbildung 4.28: Überprüfen zweier Shapes auf eine nebeneinander liegende Anordnung.

KOMBINATION NEBENEINANDER ANGEORDNETER SHAPES

Die Kombination nebeneinander angeordneter Shapes S_1 und S_2 erfolgt, indem der entlang von S_1 verlaufende Teil von S_2 entfernt und der verbleibende Rest an den zugehörigen Endpunkt von S_1 angehängt wird. Dieser Vorgang ist in Abbildung 4.29 dargestellt. Zunächst muss dazu ermittelt werden, welche Enden der beiden Shapes zusammengehören. Dies kann ähnlich zu dem in Abschnitt 4.5.1 beschriebenen Ermitteln der parallelen Abschnitte erfolgen, indem von den Endpunkten aus Pixel des anderen Shapes in orthogonaler Richtung gesucht werden.

4 Konzept zur Ohrendetektion

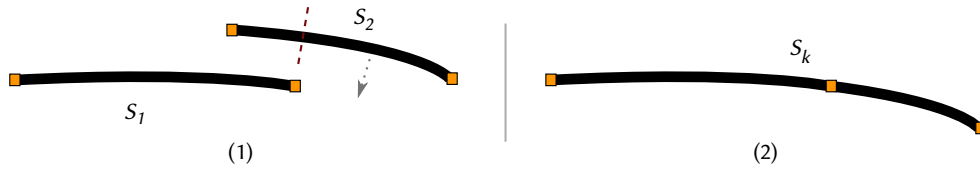


Abbildung 4.29: Kombinieren der nebeneinander angeordneten Shapes S_1 und S_2 zum Shape S_k . Zu diesem Zweck wird S_2 gekürzt und an den Endpunkt von S_1 verschoben.

KOMBINATION VOREINANDER ANGEORDNETER SHAPES

Voreinander angeordnete Shapes können aus zusammengehörigen Bereichen maximaler Flächenkrümmungen entstanden sein, dies muss aber nicht zwangsläufig der Fall sein. Ein häufig auftretender Fall ist, dass ein Teil des äußeren Ohrumrisses und ein Teil der inneren Flächenkrümmungen der Helix gefunden wurde. Unter Ausnutzung der Ähnlichkeit des Verlaufs der äußeren und inneren Flächenkrümmungsmaxima können aber auch diese zu einem sinnvollen Ohrumriss kombiniert werden. Dabei ist wie beim Verbinden über kurze Distanzen in Abschnitt 4.3 das Erzeugen einer neuen Komponente zwischen den beiden Shapes S_1 und S_2 nötig. Zunächst müssen die beiden beteiligten Shapes aber, wie in Abbildung 4.30 dargestellt, zueinander ausgerichtet werden. Als ausgerichtet gelten S_1 und S_2 , wenn folgender Zusammenhang für die Winkel α' und β' (siehe Abbildung 4.30(2)) zwischen der direkten Verbindung der beiden Shapes und dem jeweiligen Ausgangsvektor gilt: $\alpha' = -\beta'$. Es findet zum Erreichen dieses Zusammenhangs aus den ursprünglichen Winkeln α und β (siehe Abbildung 4.30(1)) lediglich eine Translation des Shapes S_2 statt. Eine Rotation wird nicht durchgeführt.

Die Position P , an die der Endpunkt von S_2 verschoben werden muss, kann durch den Vektor \vec{v}_r ermittelt werden, der eine Rotation des Verbindungsvektors \vec{v} zwischen S_1 und S_2 um den Winkel φ darstellt:

$$\vec{v}_r = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \cdot \vec{v} \quad (4.9)$$

Dabei wird φ nach folgendem Zusammenhang bestimmt, wobei $\text{sgn}(x)$ das Vorzeichen von x liefert:

$$\varphi = \begin{cases} \frac{\alpha + \beta}{2} & \text{sgn}(\alpha) = \text{sgn}(\beta) \\ \frac{\text{sgn}(\alpha) \cdot (|\alpha| - |\beta|)}{2} & \text{sonst} \end{cases} \quad (4.10)$$

Zuletzt kann die Position P über den Vektor \vec{v}_r vom Endpunkt E des Shapes S_1 aus bestimmt werden:

$$P = E + \vec{v}_r \quad (4.11)$$

4 Konzept zur Ohrendetektion

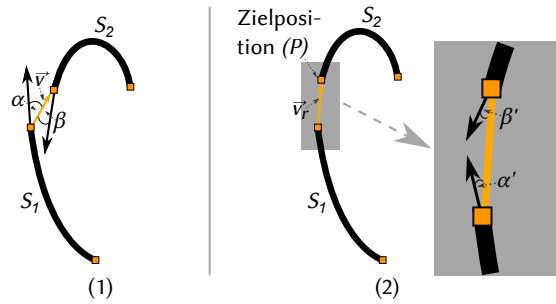
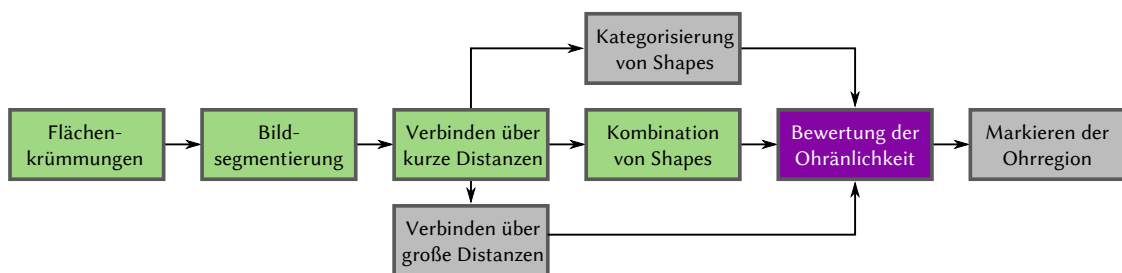


Abbildung 4.30: Verschieben von S_2 , so dass die Beträge von α und β gleich sind. Die Orientierung der Winkel ist dabei entgegengesetzt.

4.9 BEWERTUNG DER OHRÄHNLICHKEIT



Zur Unterscheidung des Ohrumrisses von anderen langen, eventuell ebenfalls kombinierten Shapes, ist es notwendig Kriterien zur Bewertung der Ohrähnlichkeit zu definieren. Die Kriterien werden in diesem Abschnitt beschrieben, wobei unterschieden wird zwischen absoluten Bewertungskriterien, die unabhängig für jeden Kandidaten berechnet werden, und relativen Bewertungskriterien, die nur im Verhältnis der Kandidaten zueinander definiert sind.

4.9.1 ABSOLUTE BEWERTUNGSKRITERIEN

Für jeden Kandidaten kann eine absolute Bewertung seiner Ohrähnlichkeit anhand von vier Kriterien bestimmt werden, die in Abbildung 4.31 dargestellt sind: die Proportion, die Gesamtkrümmung, das Vorhandensein paralleler Shapes und das Auftreten von Ecken. Diese Kriterien beziehen sich jeweils auf den angepassten Shape des Kandidaten K_i aus der Menge aller Kandidaten K . Die Bewertungszahl und alle Kriterien sind dabei auf den Bereich $[0, 1]$ normiert, wobei 0 die schlechteste und 1 die beste Bewertung darstellt.

4 Konzept zur Ohrendetektion

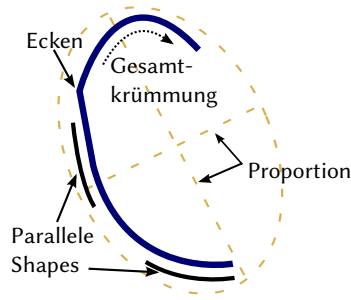


Abbildung 4.31: Bewertungskriterien für Ohrenkandidaten

SUMME DER KRÜMMUNGEN

Ein kompletter Ohrumriss sollte eine Gesamtkrümmung nahe an 2π aufweisen, da 2π einen vollständigen Kreis beschreibt. Werte deutlich unterhalb von 2π lassen auf einen unvollständigen Umriss schließen, während Werte deutlich darüber nicht von ohrähnlichen Strukturen erreicht werden können.

Zunächst wird die Gesamtkrümmung des angepassten Shapes von K_i durch Aufsummierung aller Krümmungswerte k_j von K_i berechnet:

$$Ksumme_i = \sum_{j=1}^n |k_j| \quad (4.12)$$

Dann kann aus der Gesamtkrümmung ein Wert ermittelt werden, der umso größer ist, je näher die Gesamtkrümmung an 2π ist:

$$B_{K_i} = \begin{cases} 1 - \frac{|2\pi - Ksumme_i|}{1,5\pi} & 0,5\pi < Ksumme_i < 3,5\pi \\ 0 & sonst \end{cases} \quad (4.13)$$

Die Konstanten sind dabei so definiert, dass die Bewertung ab einer Abweichung von $1,5\pi$ den Wert 0 annimmt, ohne dass eine Lücke in den Werten auftritt.

PROPORTION

Die Proportion eines Ohres liegt ungefähr zwischen $2/1$ und $3/1$, wenn sie als Verhältnis der Haupt- zur Nebenachse einer das Ohr umgebenden Ellipse definiert wird. Dies kann anhand von Fotografien von Ohren verifiziert werden. Auch bei der Proportion werden Abweichungen von dem definierten Bereich wie bei der Krümmungssumme nicht-linear gewichtet, um Ohrenkandidaten mit größeren Abweichung stärker abzuwerten.

$$B_{P_i} = \begin{cases} 1 - \frac{(2 - \text{proportion}_i)^2}{4} & 0 < \text{proportion}_i < 2 \\ 1 - \frac{(\text{proportion}_i - 3)^2}{4} & 3 < \text{proportion}_i < 5 \\ 1 & 2 \leq \text{proportion}_i \leq 3 \\ 0 & \text{sonst} \end{cases} \quad (4.14)$$

PARALLELE SHAPES

Das auftreten paralleler Shapes bei einem Kandidaten erhöht die Wahrscheinlichkeit, dass es sich um ein Ohr handelt, da insbesondere im Bereich der Helix parallel verlaufende Maxima und Minima der Flächenkrümmungen existieren, wie in Abschnitt 4.5 dargelegt wird. Die Bewertung ist daher um so besser, je größer die Länge der Abschnitte der Original-Shapes ist, in denen parallele Abschnitte anderer Shapes gefunden wurden, im Vergleich zu der Länge aller Abschnitte der Original-Shapes.

$$B_{R_i} = \frac{\text{LängeParallelerAbschnitte}_i}{\text{LängeAllerAbschnitte}_i} \quad (4.15)$$

ECKEN

Ecken treten weniger bei korrekten Ohrumrissen auf, als vielmehr bei Shapes, die durch Haare oder andere Störeinflüsse entstanden sind. Im Extremfall ist fast die gesamte Krümmung eines Shapes in Ecken gehäuft, während der Rest fast gerade ist. Der Extremfall mit einer Krümmung von 2π bei einer Häufung sämtlicher Krümmungen in Ecken ist in Abbildung 4.32(a) skizziert.

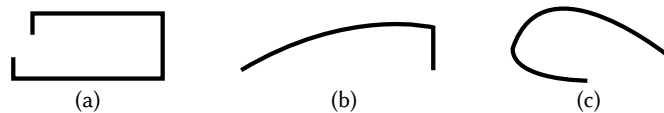


Abbildung 4.32: Von (a) nach (c) abnehmende Häufung von Krümmungen in Ecken.

Die Krümmungshäufung in einer Ecke wird dabei durch die Summierung aller Krümmungswerte $k_j > 0,1$ in dem Bereich um die Ecke bestimmt. Der Bereich um die Ecke ist durch die nächsten Minima der Krümmungswerte begrenzt. Es werden nur die großen Krümmungen summiert, um zu verhindern, dass ein großer Bereich um eine Ecke mit monoton fallenden Krümmungen zu einem zu großen Wert für die Krümmungshäufung führt. Je geringer die Krümmungshäufungen in Ecken ausgeprägt sind, desto wahrscheinlicher stellt ein Kandidat ein Ohr dar. Die Bewertung der Ecken B_{C_i} bildet dies ab, indem das Verhältnis

4 Konzept zur Ohrendetektion

der Summe aller Krümmungen in Bereichen der Ecken zu der Gesamtkrümmung des Shapes als Malus von der besten möglichen Bewertung subtrahiert wird.

$$B_{C_i} = 1 - \left(\frac{\text{KrümmungenInEcken}_i}{\text{Gesamtkrümmung}_i} \right)^2 \quad (4.16)$$

Die Ecken können mit dem in Abschnitt 2.5 beschriebenen Eckendetektor aus [HE 2008] ermittelt werden. Ein Vorteil dieses Ansatzes ist, dass zur Bestimmung der Ecken bereits eine Region um die jeweilige Ecke definiert wird, deren Krümmungswerte zur Berechnung der Bewertung herangezogen werden können.

ZUSAMMENFÜHREN ZU EINER BEWERTUNG

Die Teilbewertungen können zu einer gemeinsamen absoluten Bewertung zusammengeführt werden. Sie werden alle mit der gleichen Gewichtung versehen, da keines der Kriterien das Ohr im Vergleich zu den anderen Kriterien besonders gut oder besonders schlecht beschreibt.

$$A_i = \frac{1}{4}(B_{K_i} + B_{P_i} + B_{R_i} + B_{C_i}) \quad (4.17)$$

Die Teilbewertungen sind dabei auf 150% des Wertes der Teilbewertung mit dem kleinsten Wert begrenzt. Der Effekt der Obergrenze ist umso größer, je kleiner der minimale Wert der Teilbewertungen ist. Damit wird vermieden, dass ein Kandidat eine mittlere bis gute Bewertung erhalten kann, obwohl beispielsweise seine Proportion sehr stark vom Idealwert abweicht. Die Obergrenze auf 150% festzulegen bietet sich an, da sie bei guten Teilbewertungen ($> 0,67$) keine Auswirkungen auf die Gesamtbewertung hat.

4.9.2 RELATIVE BEWERTUNGSKRITERIEN

Neben den absoluten Bewertungskriterien werden auch relative Kriterien benötigt, da die absolute Bewertung nicht in jedem Fall ausreicht, um vollständige von unvollständigen Kandidaten zu unterscheiden. Die Ähnlichkeit von Helix und Anti-Helix kann dazu führen, dass die Anti-Helix allein die definierten Kriterien der Ohrähnlichkeit erfüllt. Ein entsprechendes Beispiel wird in Kapitel 6 vorgestellt. Aus diesem Grund sollten größere Kandidaten im Vergleich zu kleineren bevorzugt werden. Durch die Betrachtung dieses Kriteriums in Relation der Kandidaten zueinander kann die Festlegung auf eine absolute Skalierung vermieden werden.

4 Konzept zur Ohrendetektion

Drei Maße werden zur Ermittlung der relativen Bewertung herangezogen: die Anzahl der Pixel aller Shapes des Ohrenkandidaten l_i , die Summe aller Entfernungen g_i , die beim Kombinieren überbrückt wurden und die Summe aller Distanzen m_i , um die Shapes beim Kombinieren verschoben wurden. Aus diesen Werten ergibt sich für jeden Kandidaten i aus der Menge der Kandidaten eine Bewertungszahl N_i über nachfolgenden Zusammenhang. Die Werte g_i und m_i gehen dabei nicht-linear als Malus ein, da relativ große Werte darauf schließen lassen, dass Shapes kombiniert wurden, die nicht zusammengehören. Die Stärke der nicht-Linearität kann über den Exponent λ festgelegt werden. Ein Wert von $\lambda = 1,2$ erwies sich für die verwendeten Testdaten als sinnvoll.

$$N_i = l_i - g_i^\lambda - m_i^\lambda \quad (4.18)$$

Wurde für jedes i aus der Menge der Kandidaten die Bewertungszahl N_i berechnet, können alle N auf den Bereich $[0 \ 1]$ normiert werden. Dies geschieht durch die Betrachtung der Bewertungszahlen bezogen auf die größte ermittelte Zahl, also relativ zueinander.

$$R_i = \begin{cases} \frac{N_i}{\max(N)} & N_i > 0 \\ 0 & N_i \leq 0 \end{cases} \quad (4.19)$$

Diese Normierung ist notwendig, um die relative Bewertung mit der ebenfalls auf den Bereich $[0 \ 1]$ normierten absoluten Bewertung kombinieren zu können.

4.9.3 KOMBINATION ZU EINER GEMEINSAMEN BEWERTUNG

Die absolute Bewertung A_i und die relative Bewertung R_i jedes Kandidaten werden zu einer gemeinsamen Bewertung der Ohrähnlichkeit S_i zusammengefasst, anhand derer eine Rangliste der Kandidaten erstellt und der beste gewählt werden kann. Die beiden Konstanten κ und ω sind Gewichtungsfaktoren.

$$S_i = \kappa A_i + \omega R_i \quad (4.20)$$

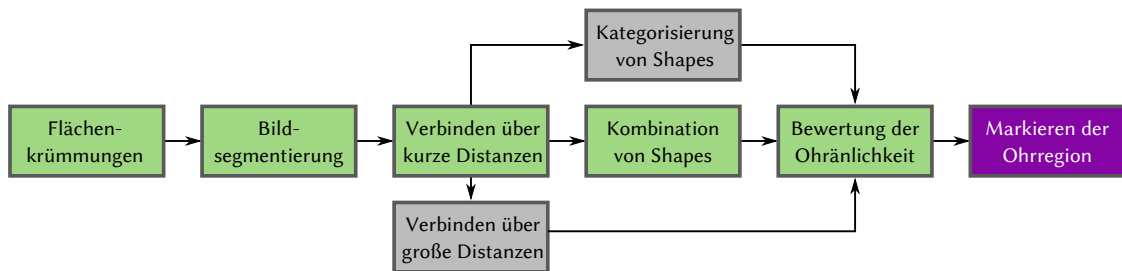
Die absolute Bewertung sollte stärker gewichtet werden, da sie die Ohrähnlichkeit beschreibt und damit die Abgrenzung zu anderen Strukturen leistet, während die relative Bewertung nur die Bevorzugung größerer und vollständigerer Kandidaten ermöglicht. Eine Normierung von S_i ist nicht mehr notwendig kann, aber durch die Einhaltung von $\kappa + \omega = 1$ bei der Wahl der Gewichtungsfaktoren erfolgen. Eine sinnvolle Abschätzung der Gewichtungsfaktoren konnte mit $\kappa = 3/4$ und $\omega = 1/4$ gefunden werden.

4.9.4 AUSWAHL DES BESTEN KANDIDATEN

Anhand der Gesamtbewertung der Ohrähnlichkeit kann der Ohrenkandidat mit der besten Bewertung als detektierter Ohrumriss gewählt werden. Dabei ist es sinnvoll eine Untergrenze zu definieren, so dass ein Kandidat mit einer Bewertung unter diesem Wert kein Ohrumriss sein kann, auch wenn er den besten gefundenen Ohrenkandidaten darstellt. Zwar kann in diesem Fall dann trotzdem keine Ohrregion detektiert werden, es wird aber auch kein fehlerhafter Bereich markiert.

Ein Wert von 0,15 als minimal zu erreichende absolute Bewertung sorgt dafür, dass Kandidaten nicht als Ohrumriss ausgewählt werden können, wenn sie die Kriterien der Ohrähnlichkeit so gut wie gar nicht erfüllen. Gleichzeitig besteht aber noch eine hohe Toleranz gegenüber Ohrumrissen, deren Bewertung weit vom Optimum entfernt ist. Deshalb eignet sich ein Unterschreiten dieses Werts gut als Ausschlusskriterium.

4.10 MARKIEREN DER OHRREGION



Die Ohrregion kann durch einen Hüllkörper um die gefundenen Shapes angenähert werden. Mögliche Körper sind die konvexe Hülle, die Ellipse und das achsenparallele Rechteck. Die Verwendung der konvexen Hülle scheidet aus, da ihre Form nicht mathematisch definiert werden kann, wie das bei einer Ellipse oder einem Rechteck der Fall ist. Auch lägen zu leicht zum Ohr gehörende Bereiche außerhalb der ermittelten Region, da die Detektion auf dem Finden der Helix (inklusive Ohrläppchen) beruht und die gegenüberliegende Seite des Ohres, auf der sich der Tragus befindet, nicht in den Ohrumriss einbezogen wird. Zur Wahl stehen daher nur das Rechteck und die Ellipse. In [Abbildung 4.33](#) sind diese beiden Hüllkörper auf einem Beispielfeld dargestellt. Unabhängig von der gewählten Form kann ein Toleranzbereich definiert werden, um den der Hüllkörper vergrößert wird, so dass er nicht direkt an die Shapes anschließt. Dies ist sinnvoll, da die Shapes nicht aus den Kanten, also den Rändern von Objekte erzeugt wurden, sondern aus den maximalen Krümmungen,

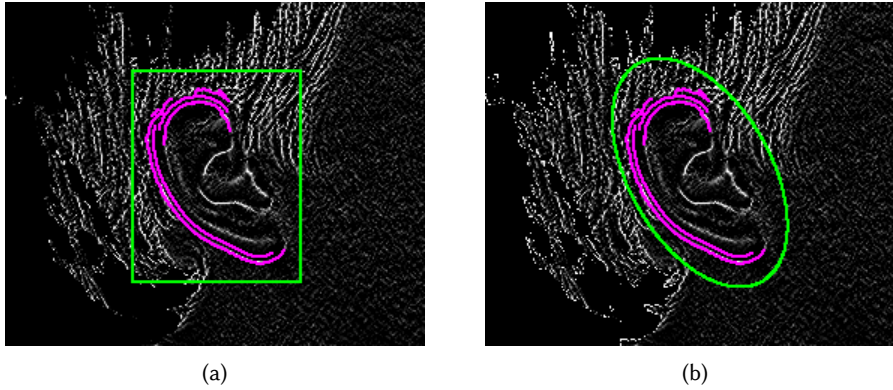


Abbildung 4.33: Vergleich eines achsenparallelen Rechtecks (a) mit einer Ellipse (b) als Hüllkörper.

die nicht direkt an der Kante auftreten. Eine Vergrößerung um 10%, wie sie auch für Abbildung 4.33 angewandt wurde, erwies sich in Tests als passender Wert.

4.10.1 ACHSENPARALLELES RECHTECK

Das achsenparallele Rechteck lässt sich mathematisch besonders einfach definieren. Die Berechnung des Rechtecks und die Segmentierung eines Bildes anhand eines achsenparallelen Rechtecks lässt sich entsprechend effizient implementieren.

4.10.2 ELLIPSE

Die Verwendung einer Ellipse bietet Vorteile gegenüber dem achsenparallelen Rechteck. Zwar ist sie in ihrer Definition und Berechnung komplexer, doch nähert sie den Umriss eines Ohres besser an, wie in Abbildung 4.33 zu erkennen ist. Außerdem kann aus ihrer Rotation die Rotation des Ohres auf dem Bild abgelesen werden. Die Ellipse ist daher für die Segmentierung des Bildes für die weiteren Module des biometrischen Systems vorzuziehen.

5

REALISIERUNG

Das in Kapitel 4 beschriebene Konzept ist objektorientiert in Matlab realisiert. In diesem Kapitel wird zunächst das Testframework eines biometrischen Systems beschrieben, in das der Algorithmus als Modul eingebettet ist. Darauf folgen einige Grundlagen der Realisierung des Konzeptes zur Detektion von Ohren.

5.1 FRAMEWORK

Im Rahmen des Gesamtprojekts wird am CASED ein Framework zum Testen der einzelnen Module und der kompletten Pipeline eines biometrischen Systems entwickelt. Die Module sind dabei in die drei Teilbereiche „Preprocessing“, „Feature Extraction“ und „Comparison“ unterteilt. Eine Übersicht über dieses Framework ist in Abbildung 5.1 dargestellt.

Ergebnisse der Teilbereiche können in dem zugehörigen Report gespeichert werden. Diese Ergebnisse sind im Fall des Preprocessing beispielsweise die detektierte Region und die Zeitdauer der Detektion pro Bild. Es können auch Statistiken, bezogen auf eine vorher erzeugte Ground-Truth (siehe Abschnitt 6.1), erstellt werden, wie beispielsweise die Detektionsrate.

Das zur Detektion von Ohren entwickelte Modul implementiert das Preprocessing-Interface und ist damit Teil des PreprocessingSubsystem des Frameworks. Um dieses

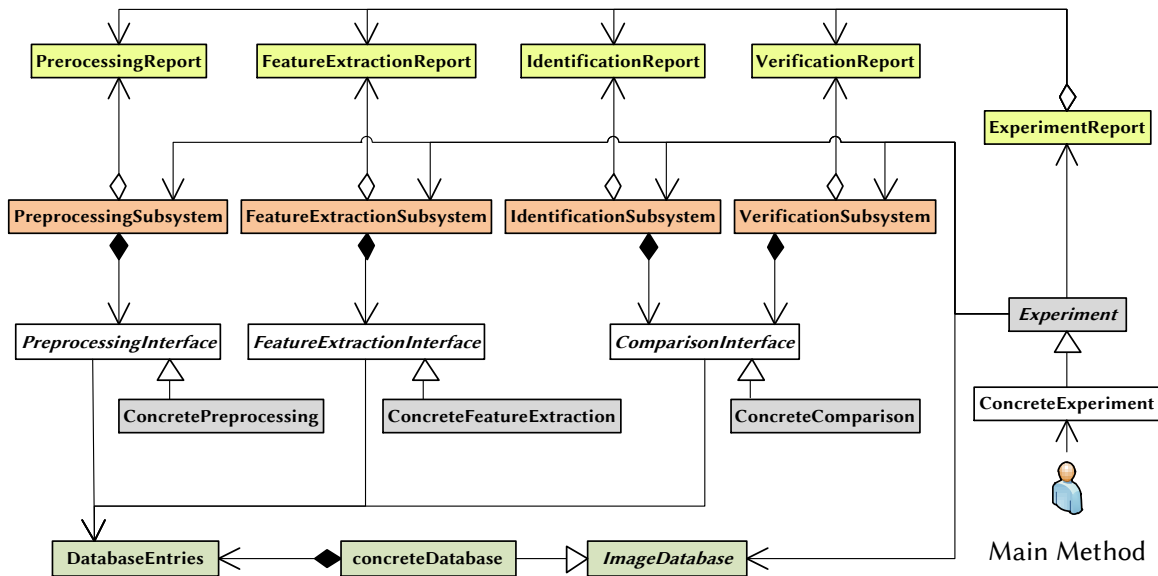


Abbildung 5.1: Klassendiagramm des Testframeworks, erstellt von Anika Pflug.

Modul zu testen kann ein Experiment erstellt werden, dessen *PreprocessingSubsystem* die Ohrendetektion enthält.

5.2 MODUL ZUR OHRENDETEKTION

Das Ohrendetektionsmodul findet die Ohrregion auf einem gegebenen Eingabebild, also einer Punktwolke mit drei Koordinaten für jeden Punkt. Als Ausgabewert wird eine Maske geliefert, die der detektierten Region entspricht, sowie alle Shapes, die als dem Ohr zugehörig erkannt wurden. Der Ablauf des Detektionsalgorithmus ist sequentiell, entsprechend der Beschreibung des Konzeptes in Kapitel 4.

Zunächst werden die Flächenkrümmungen des Eingabebildes berechnet und aus den maximalen und minimalen Werten dieser Krümmungen je ein Binärbild erzeugt. Die Binärbilder werden dann entsprechend der Beschreibungen von Abschnitt 4.2.3 bearbeitet, so dass die Zusammenhangskomponenten extrahiert und über kurze Distanzen verbunden werden können. Dann erfolgt die Auswahl von Ohrenkandidaten nach den in Abschnitt 4.4 beschriebenen Kriterien. Die Kandidaten werden dann zunächst mit ihren parallelen Shapes kombiniert und schließlich mit den anderen Kandidaten. Der beste Ohrenkandidat nach diesen beiden Kombinationsschritten wird als Ohr gewählt und schließlich wird mit Hilfe der originalen Shapes des Kandidaten die Ohrregion bestimmt.

Der Ablauf des Algorithmus wird dabei von der Klasse `SurfatureEarDetection` gesteuert, die das `PreprocessingInterface` des Testframeworks implementiert und damit auch die Schnittstelle des Moduls zum Framework darstellt. Zur Durchführung der Detektion greift `SurfatureEarDetection` auf die in den folgenden Abschnitten beschriebene Datenstruktur (5.2.1) und sogenannte Evaluators (5.2.2) zurück, welche die Evaluierungsfunktionalität von Shapes und Kandidaten in Bezug auf Verbindungen, Parallelität oder Kombinationsmöglichkeiten beinhalten.

5.2.1 DATENSTRUKTUR

Die in Abschnitt 4.1 beschriebenen Ebenen werden in der Klassenstruktur der implementierten Anwendung abgebildet. Abbildung 5.2 zeigt den Zusammenhang von Komponenten, Shapes und Ohrenkandidaten, die als die Klassen `Components`, `Shapes` und `EarCandidates` implementiert sind.

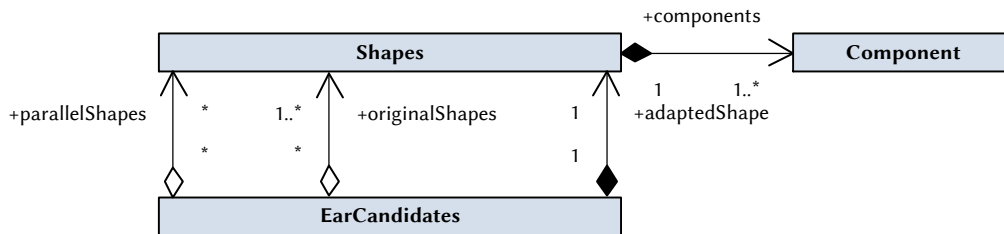


Abbildung 5.2: Klassendiagramm der Datenstruktur, bestehend aus Komponenten, Shapes und Ohrenkandidaten.

Ein Ohrenkandidat, der in der obersten Ebene der Datenstruktur angesiedelt ist, besteht aus genau einem angepassten Shape (`adaptedShape`), der den Ohrenumriss beschreibt. Dieser Shape wird für die Bewertung des Ohrenkandidaten verwendet und seine Komponenten müssen nicht unbedingt Zusammenhangskomponenten des Bildes entsprechen, da sie durch Kombinieren verschoben und gekürzt sein können (siehe Abschnitt 4.8). Die `originalShapes` eines Ohrenkandidaten wiederum sind jene Shapes, auf deren Grundlage der `adaptedShape` erzeugt wurde. Nur diese definieren am Ende die Ohrregion und werden als Features an weitere Module des biometrischen Systems weitergereicht. Die Liste der `parallelShapes` wird benötigt, um die parallelen Shapes des Kandidaten mit diesem zu kombinieren. Da das Hinzufügen jedes parallelen Shapes nicht zwangsläufig zu dem besten möglichen Kandidaten führt, muss es möglich sein Kandidaten mit verschiedenen Kombinationen mit ihren parallelen Shapes zu erzeugen.

Ein Objekt der Klasse `Shapes` wiederum besteht aus einem oder mehreren Objekten der Klasse `Component`, wobei jede `Component` nur genau einem `Shape` zugeordnet sein kann. `Shapes` sind damit auf der mittleren Ebene angeordnet. Für `Components`, welche die unterste Ebene der Datenstruktur bilden, wird dabei unterschieden, ob sie aus den Maxima der Flächenkrümmungen extrahiert wurden und damit tatsächliche Zusammenhangskomponenten aus den Binärbildern sind oder ob sie eine Verbindungslinie des in Abschnitt 4.3 beschriebenen Verbindens sind. Diese Unterscheidung ist für die Ermittlung der Ohrregion nicht notwendig, sie bietet aber die Möglichkeit die extrahierten Komponenten in einem weiteren Modul des biometrischen Systems zu nutzen.

5.2.2 EVALUATORS

Die im Konzept beschriebenen Überprüfungen auf Verbindungen, Parallelität oder Ecken werden in `Evaluators` vorgenommen, die über die in Abbildung 5.3 dargestellte Klassenhierarchie strukturiert sind. Die abstrakte Superklasse `Evaluator` enthält einige Funktionalitäten, die zur Vorbereitung der zu untersuchenden `Shapes` für die Subklassen benötigt werden. Diese werden über die Funktion `prepareShape` zur Verfügung gestellt, die einen gegebenen `Shape` so vorbereitet, dass er beispielsweise auf Verbindungen oder Parallelität untersucht werden kann. Dazu werden die Krümmungen in jedem Punkt des gegebenen `Shapes` berechnet (`calculateCurvature`) und auf Basis dessen in der Methode `examineCurvature` eine Krümmungsrichtung ermittelt (siehe Abschnitt 4.3.1). Außerdem werden mit `calculateDirectionVectors` die Ausgangsvektoren des `Shapes` (siehe ebenfalls Abschnitt 4.3.1) bestimmt. Die Klasse `Evaluator` ermöglicht auch das Aufspalten von `Shapes` an definierten Positionen (siehe Abschnitt 4.4) über die Methode `splitShapeAtPositions`. Diese Positionen können zum Beispiel Ecken sein.

Die Detektion von Ecken (`getCorner`) wird vom `CornerEvaluator` bereitgestellt. Der `ConnectionEvaluator` bietet die Möglichkeit `Shapes` nach dem in Abschnitt 4.3 beschriebenen Konzept zu verbinden. Dafür liefert die Methode `checkShapeForConnections` eine nach der Erfüllung der festgelegten Bewertungskriterien sortierte Liste aller möglichen Verbindungen des gegebenen `Shapes`. Die Untersuchung der Parallelität (siehe Abschnitt 4.5) wird vom `ParallelismEvaluator` durchgeführt. Dazu ist es möglich sich mit `getParallelShapes` alle `Shapes` zu erhalten, die zu einem gegebenen `Shape` parallel sind. Der `CompletenessEvaluator` ermöglicht es schließlich, die in Abschnitt 4.8 beschriebenen Kombinationen von `Shapes` durchzuführen und Ohrenkandidaten aus kombinierten `Shapes` zu bewerten.

5 Realisierung

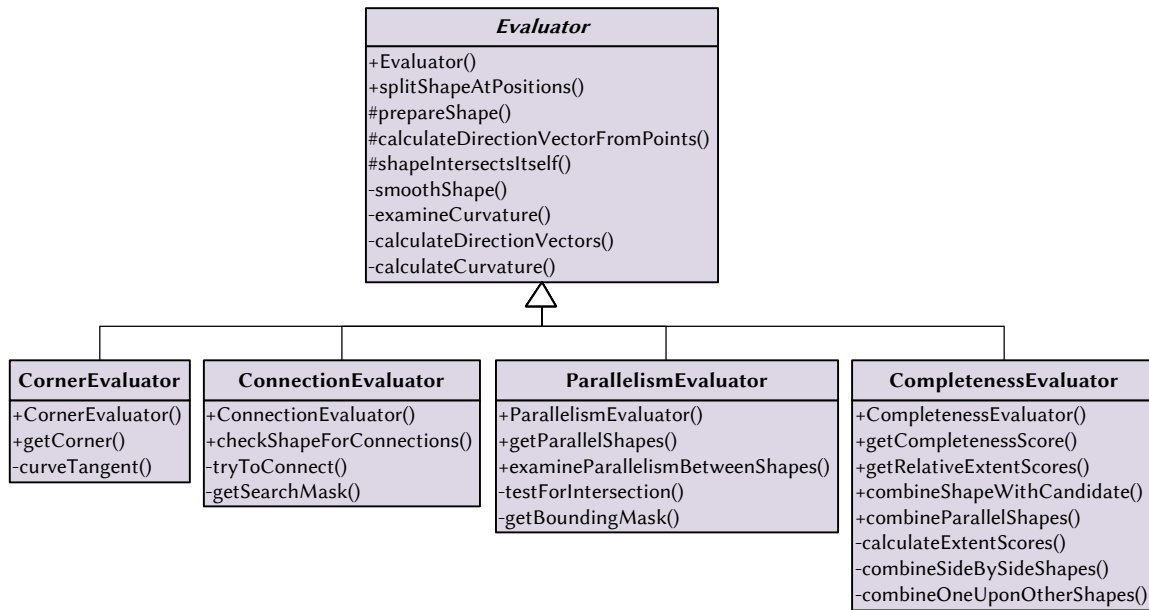


Abbildung 5.3: Vererbungshierarchie der Evaluators. Die Attribute der dargestellten Klassen, sowie die Parameter der Methoden sind in diesem Diagramm nicht enthalten.

5.2.3 KONFIGURATION

Das Detektionsmodul lässt sich umfangreich konfigurieren. Das bedeutet insbesondere, dass alle im Konzept beschriebenen Parameter, wie beispielsweise der Öffnungswinkel des Suchbereichs beim Verbinden von Komponenten (siehe Abschnitt 4.3.2), flexibel angepasst werden können. Damit ist es einfach möglich mit verschiedenen Werten zu experimentieren. Auch lassen sich geometrische Transformationen der Eingabedaten simulieren, ohne dass die festgelegten Vergleichsdaten (siehe Kapitel 6) verändert werden müssen, indem die Eingabedaten innerhalb des Moduls auf die gewünschte Art transformiert und am Ende die Transformationen umgekehrt auf die Ergebnisse angewendet werden.

Über das in Matlab vorhandene Konzept von Function-Handles ist es auch möglich einzelne Funktionen, wie das Binarisieren oder die Skelettierung auszutauschen, ohne Änderungen am Quellcode des Moduls vornehmen zu müssen. Dazu muss nur der Name einer entsprechenden Funktion, welche die erwartete Anzahl ein Ein- und Ausgabeparametern aufweist, als Option übergeben werden.

*Die Erschütterung der Luft wird
erst Schall, wo ein Ohr ist.*

Georg Christoph Lichtenberg

6

ERGEBNISSE & DISKUSSION

In diesem Kapitel werden in Abschnitt 6.1 die Metriken zum Testen des implementierten Algorithmus erläutert und darauf folgend in Abschnitt 6.2 die Testergebnisse präsentiert. Außerdem werden anhand von Beispielen erfolgreicher und fehlgeschlagener Detektionen die Fähigkeiten und verbliebenen Probleme des Algorithmus diskutiert und bewertet.

6.1 MESSMETHODE

Das Konzept zur Detektion von Ohren wurde auf der in Abschnitt 2.6 beschriebenen Bild-datenbank UND-J2 der University of Notre Dame [YAN 2007] getestet. Dabei wurden nur die dreidimensionalen Aufnahmen verwendet, da der entwickelte Algorithmus keine Texturinformationen benötigt.

Die Vergleichsdaten, im folgenden auch Ground-Truth genannt, wurden manuell aus den 2D-Bildern erzeugt, deren Koordinaten mit denen der 3D-Bilder übereinstimmen. Die Vergleichsdaten sind tendenziell etwas größer als die tatsächliche Ohrregion. Abbildung 6.1 zeigt zwei Beispiele auf Basis der Ground-Truth zugeschnittener 2D-Aufnahmen mit relativ großer Region um das Ohr herum. Dies muss beim Auswerten der Messergebnisse beachtet werden, da die Region in den Ergebnissen des Algorithmus genauer an die tatsächliche Größe der Ohren angepasst ist. Da die Ground-Truth in Form achsenparalleler Rechtecke vor-



Abbildung 6.1: Beispiele der manuell erzeugten Ground-Truth, dargestellt anhand zweidimensionaler Aufnahmen.

liegt, wurde der Algorithmus zum Testen so konfiguriert, dass er als Ergebnisregion ebenfalls Rechtecke lieferte.

Um eine bessere Einordnung der Ergebnisse zu ermöglichen, werden diese mit den Ergebnissen eines auf den zweidimensionalen Bildern der UND-J2-Datenbank trainierten und ausgeführten AdaBoost-Klassifikators (siehe 3.1.1) aus dem OpenCV-Framework verglichen.

6.1.1 KENNZAHLEN DES ÜBERLAPPUNGSRADES

Damit eine quantitative Aussage über die Güte der Ohrendetektion auf einem Bild getroffen werden kann, werden drei Kennzahlen ermittelt, welche die Überlappung zwischen Ergebnisregion und Ground-Truth-Region beschreiben. Die Regionen sind jeweils eine rechteckige binäre Maske. Die Kennzahlen sind als Mengenoperationen auf der Ergebnismenge E_i und der Ground-Truth-Menge G_i für ein untersuchtes Bild i definiert. Dabei ist E_i die Menge aller Pixel ungleich null der Ergebnismaske und G_i analog die Menge aller Pixel ungleich null der Ground-Truth-Maske. Mit Hilfe dieser beiden Mengen werden die wie folgt definierten Werte für jedes untersuchte Bild i berechnet. Abbildung 6.2 zeigt eine Skizze der Zusammenhänge.

Die Überlappung der beiden Regionen wird über die Schnittmenge im Verhältnis zu den Mächtigkeiten der beiden Mengen bestimmt:

$$L_i = 2 \frac{|E_i \cap G_i|}{|E_i| + |G_i|} \quad L_i \in [0,1] \quad (6.1)$$

Ein Maß für die Übergröße der Ergebnisregion, d.h. welche Bereiche der Ergebnisregion nicht in der Ground-Truth-Region enthalten sind:

$$O_i = 1 - \frac{|E_i \setminus G_i|}{|E_i|} \quad O_i \in [0,1] \quad (6.2)$$

Und ein Maß für die Untergröße der Ergebnisregion, d.h. welche Bereiche der Ground-Truth-Region nicht in der Ergebnisregion enthalten sind:

$$U_i = 1 - \frac{|G_i \setminus E_i|}{|G_i|} \quad U_i \in [0,1] \quad (6.3)$$

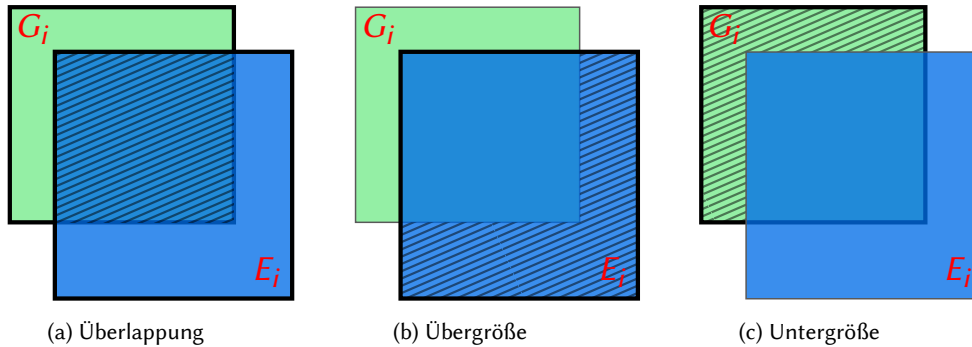


Abbildung 6.2: Kennzahlen des Überlappungsgrades. Grün dargestellt ist dabei die Ground-Truth-Region G_i , während die Ergebnisregion E_i blau ist. Die Kennzahl ergibt sich jeweils über die Division der Mächtigkeit des schraffierten Bereichs, durch die Mächtigkeit des eingerahmten Bereichs, bzw. der eingerahmten Bereiche.

Alle drei Kennzahlen sind auf das Intervall $[0,1]$ normiert. L_i nimmt bei vollständiger Überlappung zweier gleich großer Masken den Maximalwert 1 an, also beispielsweise bei der Verwendung derselben Maske für Ergebnis und Ground-Truth. Sind die beiden verglichenen Masken disjunkt, so wird der Wert 0 angenommen. O_i und U_i sind genau entgegengesetzt definiert, d.h. sie nehmen bei vollständig überlappenden Masken den Wert 0 an, da sie die Fehler in der Größe und Positionierung der Ergebnisregion angeben sollen. Die Plausibilität der Kennzahlen wurde durch einen Vergleich der Ground-Truth mit sich selbst überprüft: L_i nahm wie erwartet den Wert 1 an und O_i und U_i jeweils den Wert 0.

6.1.2 ERKENNUNGSRATE

Die Erkennungsrate gibt den Anteil der Bilder an, bei denen die Ohrregion erfolgreich detektiert werden konnte. Sie kann mit Hilfe des Überlappungsgrades L_i von Ergebnis- und Ground-Truth-Region bestimmt werden, indem eine minimale Überlappung festgelegt wird, die für eine erfolgreiche Detektion erreicht sein muss. Eine Überlappung von 60% erwies sich bei der vorliegenden Ground-Truth als eine sinnvolle Untergrenze, da das Ohr bei einer solchen Überlappung in den meisten Fällen annähernd vollständig in der detektierten Region enthalten ist. In Abbildung 6.3 sind die Ergebnis- und Ground-Truth-Regionen von drei

Bildern mit ungefähr 60% Überlappung dargestellt. Bei (a) und (b) ist das Ohr annähernd vollständig innerhalb der detektierten Region. Wenn die Ground-Truth-Region sehr knapp definiert wurde, wie es bei (c) der Fall ist, wurde bei 60% Überlappung dennoch zumindest ein großer Teil des Ohres korrekt erkannt.

In dem Bereich von 50% bis 60% Überlappung kommen ebenfalls noch fast vollständig detektierte Ohrregionen vor, die Wahrscheinlichkeit, dass bei einer Überlappung in diesem Intervall wesentliche Teile des Ohres außerhalb der detektierten Region liegen ist aber sehr hoch. Deshalb wurde 60% als Untergrenze festgelegt.

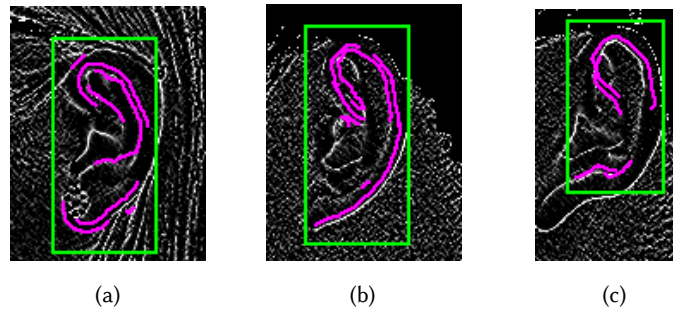


Abbildung 6.3: Beispiele für einen Überlappungsgrad von etwa 60%. Der Bildausschnitt entspricht der Ground-Truth, während die detektierte Region durch ein grünes Rechteck dargestellt ist.

6.2 MESSERGEBNISSE

In diesem Abschnitt werden die Messergebnisse der systematischen Tests des implementierten Algorithmus präsentiert und ausgewertet. Es werden Statistiken zu den Überlappungsgraden von Ergebnis- und Ground-Truth-Regionen gezeigt und die Erkennungsrate des Algorithmus in Abhängigkeit der Überlappung dargestellt. Auch ein Vergleich mit den Ergebnissen eines AdaBoost-Klassifikators auf der selben Datenbank wird gezeigt. Erfolgreiche und nicht erfolgreiche Detektionen des Algorithmus werden diskutiert.

6.2.1 ERKENNUNGSRATE

Die Erkennungsrate des Algorithmus liegt bei 94%, wenn die Definition aus Abschnitt 6.1.2 und damit ein Überlappungsgrad von 60% als Grundlage genommen wird. Für unterschiedliche Überlappungsgrade sind die Detektionsraten in Abbildung 6.4 dargestellt. Dabei ist zu erkennen, dass die Detektionsraten zwischen 90% und 70% Überlappung stark ansteigen

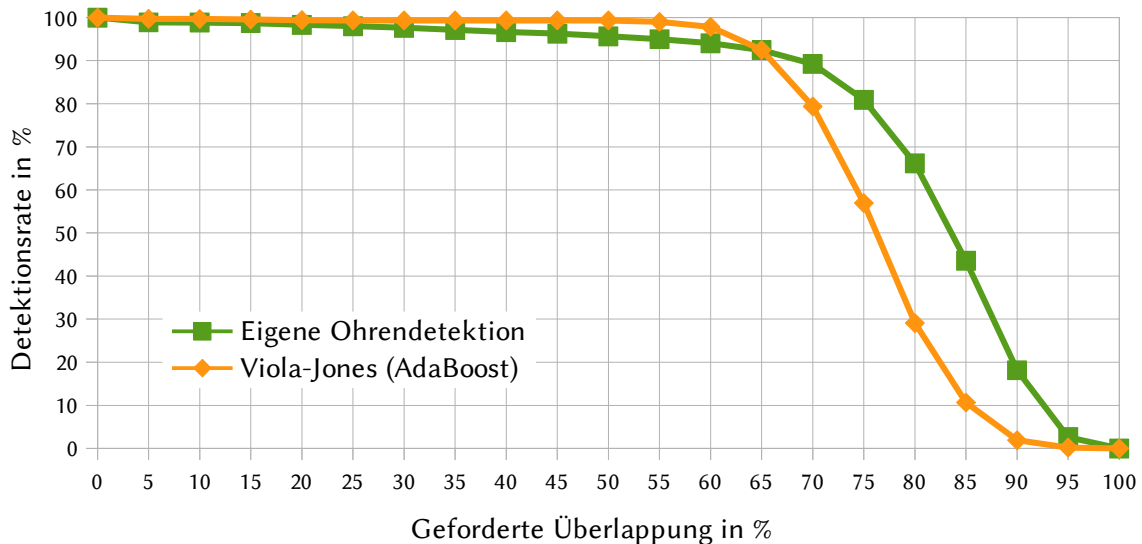


Abbildung 6.4: Erkennungsrate in Abhängigkeit des geforderten Überlappungsgrades.

und die Kurve danach abflacht. Selbst wenn 70% Überlappung als Untergrenze festgelegt würde, könnte daher eine Detektionsrate von immerhin 89% erreicht werden.

Betrachtet man die Erkennungsrate eines auf den zweidimensionalen Bildern der UND-J2-Datenbank trainierten und ausgeführten AdaBoost-Klassifikators, die zum Vergleich ebenfalls in Abbildung 6.4 dargestellt ist, fällt zunächst auf, dass die Erkennungsrate erwartungsgemäß sehr gut ist. Bei der Festlegung des minimalen Überlappungsgrades für eine erfolgreiche Detektion auf 60%, ergibt sich eine Detektionsrate von 97,8%. Bei einer Überlappung von 55% gelten bereits 99% der Ohren detektiert, so dass in dem restlichen Intervall bis 0% Überlappung kaum noch eine Änderung an der Erkennungsrate stattfindet.

Der in dieser Arbeit vorgestellte Algorithmus weist zwar eine etwas geringere Detektionsrate auf, als der AdaBoost-Klassifikator, er kann bei einer erfolgreichen Detektion die Ohrregion aber genauer bestimmen, wie der Verlauf der beiden Kurven im Bereich von 65% bis 95% Überlappung zeigt.

6.2.2 ÜBERLAPPUNGSRADE

Die Detektionsergebnisse des Algorithmus können in Form von Histogrammen dargestellt werden, indem die Wertebereiche der drei in Abschnitt 6.1.1 beschriebenen Kennzahlen in Klassen mit einer Breite von fünf Prozentpunkten unterteilt werden. Für die Überlappung von Ergebnis- und Ground-Truth-Regionen ergab sich bei den durchgeführten Tests das in Abbildung 6.5 dargestellte Histogramm. Dieses Histogramm korreliert dabei direkt mit den

6 Ergebnisse & Diskussion

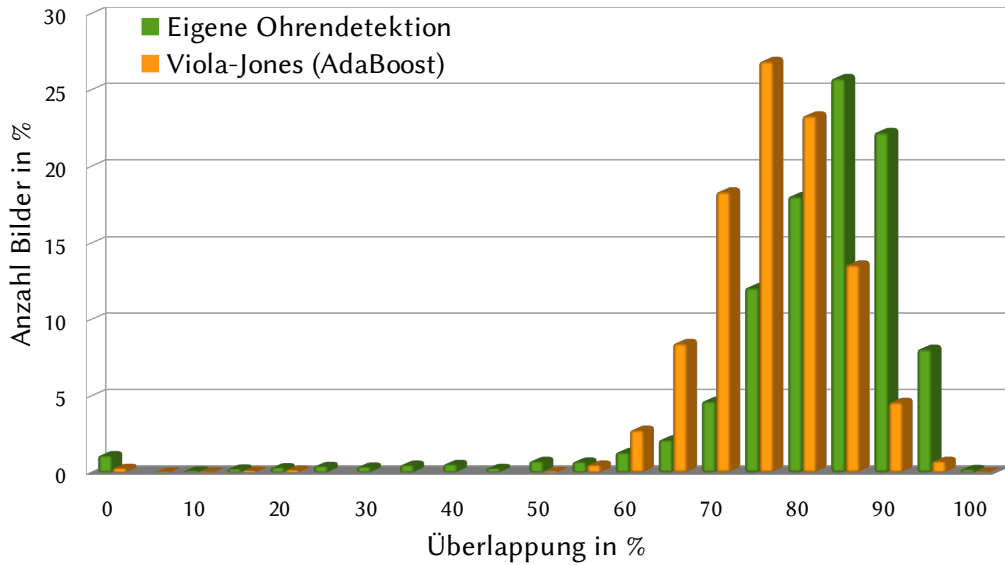


Abbildung 6.5: Überlappungsgrade für die detektierten Regionen. Dieses Diagramm zeigt die Überlappungen von Ergebnis- und Ground-Truth-Regionen für den in dieser Arbeit vorgestellten Detektionsalgorithmus und AdaBoost im Vergleich.

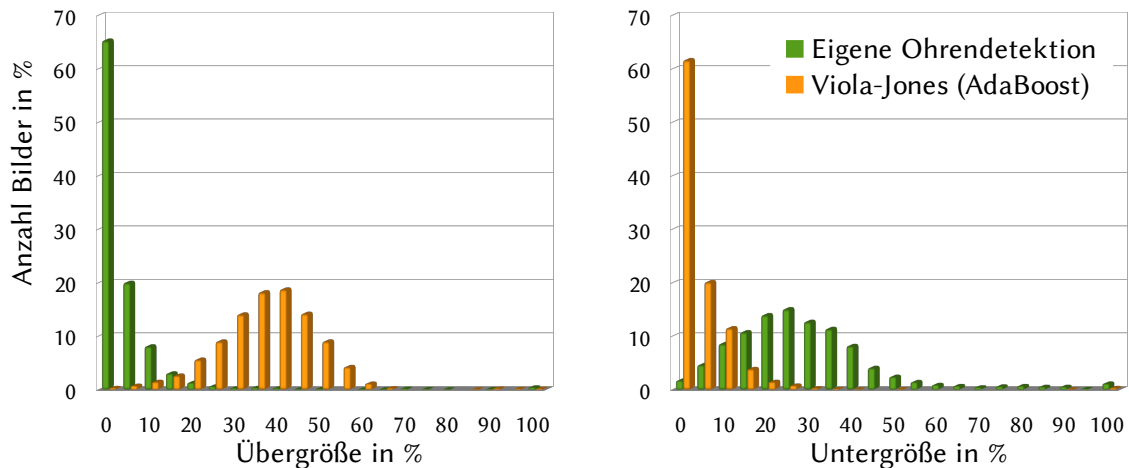


Abbildung 6.6: Übergrößen und Untergrößen der detektierten Ohrregionen. Die von AdaBoost detektierten Regionen sind eher größer als die der Ground-Truth, während die eigene Implementierung eher etwas kleinere Regionen detektiert.

in Abbildung 6.4 dargestellten Detektionsraten, da diese gerade eine summierte Darstellung des Überlappungshistogramms sind.

Zusätzliche Informationen zur Analyse der Detektionsergebnisse lassen sich aus den beiden Kennzahlen der Übergröße und der Untergröße der detektierten Region erhalten. Die zugehörigen Histogramme sind in Abbildung 6.6 dargestellt. Dabei ist zu erkennen, dass der

im Rahmen dieser Arbeit entwickelte Detektionsalgorithmus die Ohrregion tendenziell kleiner detektiert, als die jeweilige Ground-Truth-Region definiert ist und nur geringe Bereiche außerhalb als Bestandteil der Ohrregion ansieht. Da die Ground-Truth-Region für die meisten Bilder etwas größer als die tatsächliche Ohrregion ist, zeigt sich in diesem Histogrammen das positive Ergebnis der besseren Anpassung der Ergebnisregion an die tatsächliche Ohrregion.

Der AdaBoost-Klassifikator zeigt hier ein genau entgegengesetztes Bild: Die detektierte Region ist eher etwas größer als die Ground-Truth-Region, dafür ist die Untergröße sehr gering ausgeprägt, so dass eine zu klein detektierte Ohrregion sehr unwahrscheinlich ist.

6.2.3 GEOMETRISCHE TRANSFORMATIONEN

Der entwickelte Algorithmus ist auf Grund seiner Konzeption robust gegenüber den geometrischen Transformationen Rotation und Spiegelung, sowie invariant gegenüber Translationen. Da das Ohr bereits auf den Bildern der verwendeten Testdatenbank jeweils an unterschiedlicher Position zu finden ist, wird die Translation nicht gesondert überprüft.

ROTATION

Die Ohrendetektion ist robust gegen Rotationen des Eingabebildes, wie die Ergebnisse mehrerer Testdurchläufe zeigten, wobei die Eingabebilder für jeden Durchlauf jeweils um 90° gedreht wurden. In Abbildung 6.7 ist die Detektionsrate in Abhängigkeit der Drehung des Bildes dargestellt.

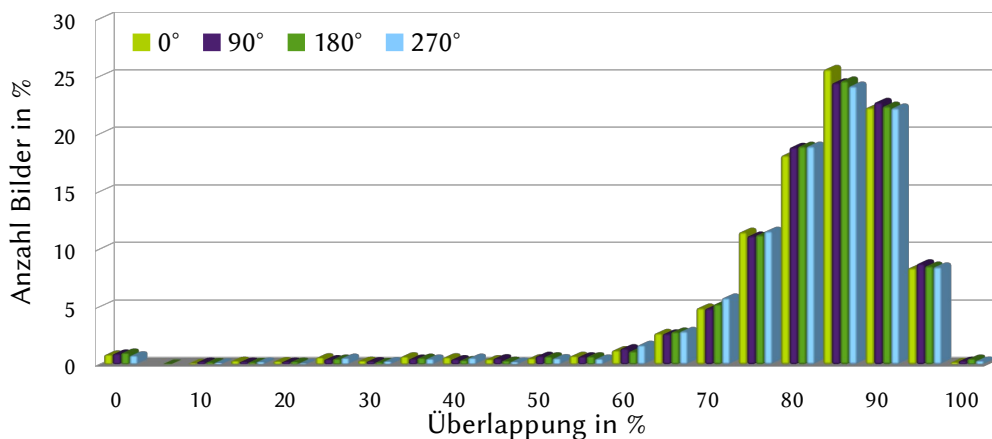


Abbildung 6.7: Histogramm der Überlappungsgrade bei verschiedenen Rotationen der Eingabebilder. Der Grad der Rotation ist dabei jeweils auf die ursprüngliche Orientierung der Eingabebilder bezogen.

SPIEGELUNG

In den Ergebnissen von Detektionen auf horizontal gespiegelten Eingabebildern, verglichen mit nicht gespiegelten Bildern, dargestellt durch ein Histogramm der Überlappungsgrade von Ergebnis- und Ground-Truth-Regionen in Abbildung 6.8, sind nur geringe Unterschiede zu erkennen. Der Algorithmus ist dementsprechend robust gegen die geometrische Transformation der Spiegelung, woraus sich ableiten lässt, dass sowohl linke als auch rechte Ohren ohne Probleme detektiert werden können.

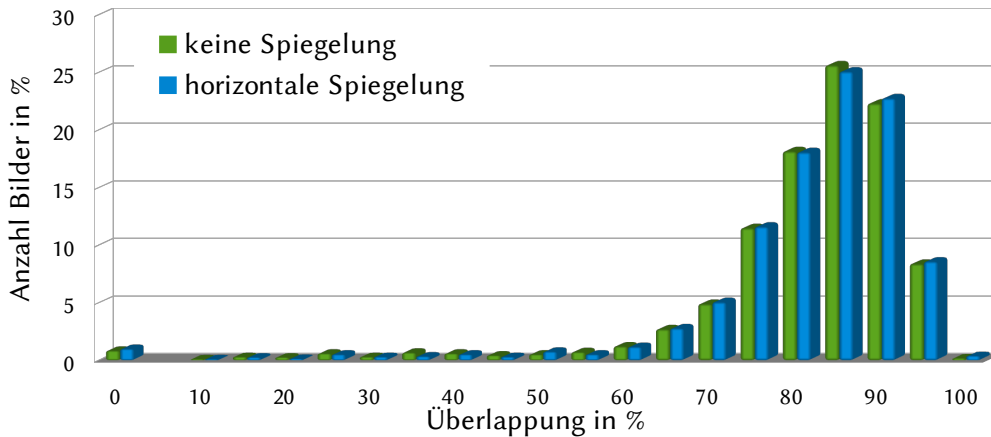


Abbildung 6.8: Überlappungsgrade für die detektierten Regionen bei der horizontalen Spiegelung der Eingabebilder im Vergleich zu den nicht gespiegelten Bildern.

6.2.4 POSITIVE ERGEBNISSE

Die Detektion des Ohres funktionierte problemlos, wenn die Helix annähernd vollständig durch das in Abschnitt 4.3 beschriebene Verbinden von Shapes rekonstruiert werden konnte. Wird die Helix so vollständig repräsentiert, macht es für die Detektion kaum einen Unterschied ob zusätzlich nur sehr wenige oder viele Komponenten großer Flächenkrümmungen vorhanden sind, die nicht zum Ohr gehören. Zwei entsprechende Detektionsbeispiele sind in den Bildern (a) und (b) der Abbildung 6.9 zu sehen. In diesen Fällen wäre eine erfolgreiche Detektion auch ohne das in Abschnitt 4.8 beschriebene Kombinieren von Ohrenkandidaten möglich. Der beste Ohrenkandidat in (b) hat dabei eine geringere Bewertung als der beste Ohrenkandidat des in (a) dargestellten Detektionsergebnisses, da die Gesamtkrümmung geringer ist. Außerdem konnte keine Parallelität zwischen den konvexen und konkaven Flächenkrümmungen erkannt werden, da die beiden entsprechenden Shapes sich im oberen Bereich des Ohres auseinander bewegen. Dennoch konnte auch in diesem Fall eine relativ

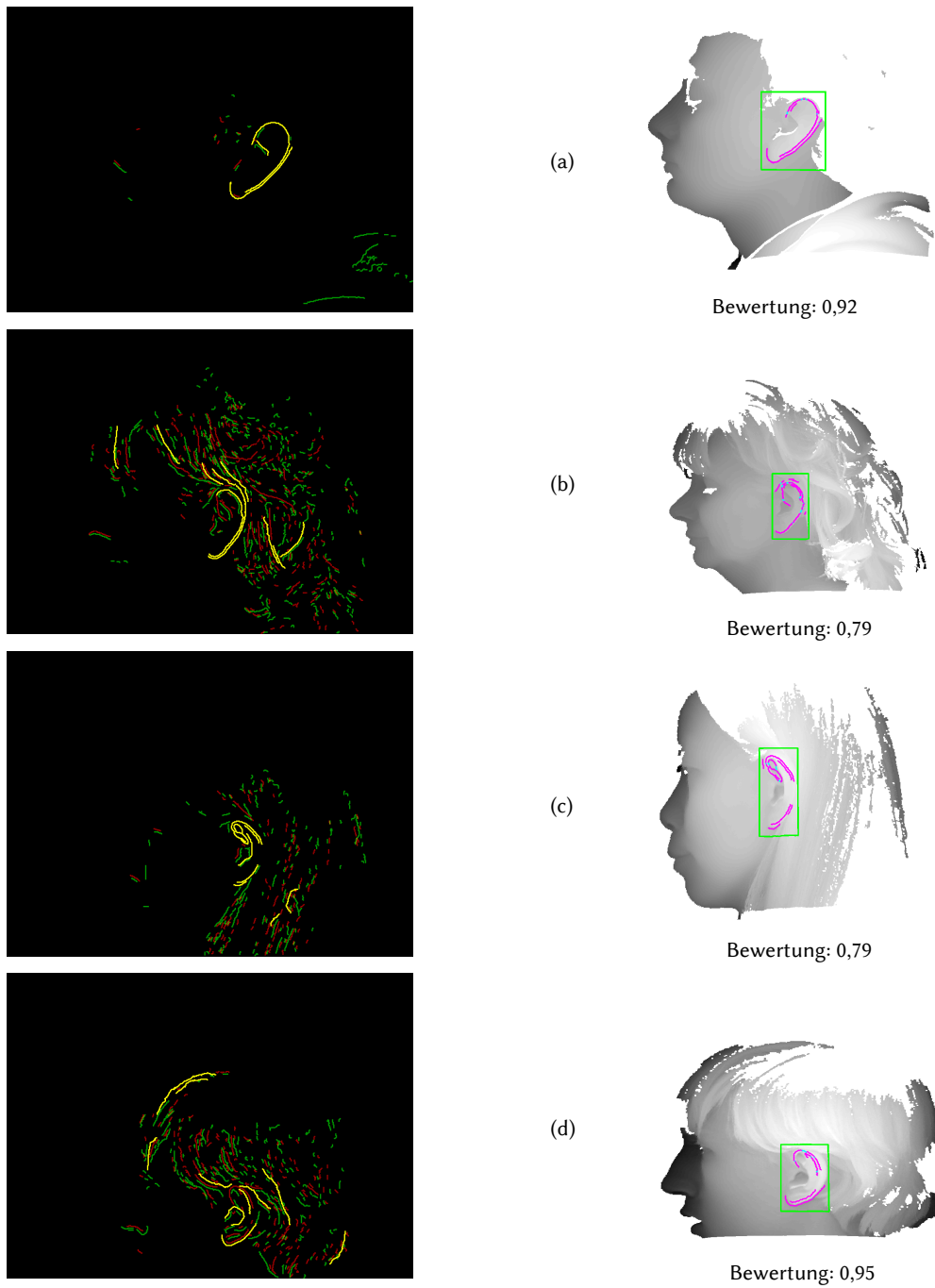


Abbildung 6.9: Beispiele erfolgreicher Detektionen des Ohrumrisses. Bei den oberen beiden Bildern konnte die Helix vollständig durch das Verbinden konstruiert werden, während bei den unteren beiden größere Lücken vorhanden sind. Die Bewertung des besten Ohrenkandidaten wird für jedes Bild dargestellt.

hohe Bewertung erreicht werden, die für eine gute Erfüllung der definierten Kriterien der Ohrähnlichkeit spricht.

Auch mit größeren Lücken im Bereich der Helix konnte der Algorithmus gut umgehen, solange prominente Shapes sowohl aus dem oberen als auch dem unteren Bereich des Ohres gefunden werden konnten, wie beispielhaft in den Abbildungen 6.9(c) und 6.9(d) zu sehen ist. In diesen Fällen konnte durch die Kombination der prominenten Shapes jeweils ein guter Ohrenkandidat gebildet werden. Das Verbinden über kurze Distanzen hätte alleine nicht zu einem vollständigen Ohrumriss führen können. Auch hierbei können, wie in (d) zu erkennen ist, neben den Shapes des Ohrumrisses weitere prominente Shapes vorhanden sein, ohne dass die Detektion fehlschlägt, da über die Bewertung der Ohrenkandidaten der ohrähnlichste Kandidat gewählt werden kann. Der beste Kandidat in (d) konnte eine sehr gute Bewertung durch die in weiten Teilen auftretende Parallelität und eine hohe Gesamtkrümmung erreichen, die sich nicht auf Ecken konzentriert. Auch die Proportion ist innerhalb des optimalen Bereichs. Bei (c) erhielt der beste Ohrenkandidat eine etwas geringere Bewertung, da die Eckendetektion an dem angepassten Shape des Kandidaten eine Ecke erkannte, die dann zu einer Abwertung führte.

Ohringe haben auf die Komponenten der Flächenkrümmungsmaxima denselben Effekt, der auch durch Schwankungen in der Ausprägung der Flächenkrümmungen auftritt: Die Komponenten sind unterbrochen. Solange die einzelnen Komponenten des Ohres dennoch zu den längsten des Bildes zählen, verhindern auch Ohringe die Detektion nicht. In Abbildung 6.10 ist ein Beispiel mit zwei Ohringen zu sehen, die die Detektion des Ohrumrisses nicht behinderten. Im schlimmsten Fall kann allerdings auch ein Teil des Ohres fehlen, wenn durch den Ohring beispielsweise im Bereich des Ohrläppchens keine großen Flächenkrümmungen des Ohrumrisses auftreten.

6.2.5 VERBLEIBENDE PROBLEME

Nicht in jedem Fall funktionierte die Detektion fehlerfrei, wie an der in Abschnitt 6.2.1 dargestellten Erkennungsrate zu sehen ist. In diesem Abschnitt werden daher einige Beispiele vorgestellt, die dem Algorithmus Probleme bereiten.

FEHLKLASSIFIKATIONEN

Es konnte vorkommen, dass das Ohr nicht gefunden wurde oder sogar ein anderer Bereich des Bildes als vermeintliche Ohrregion erkannt wurde. Das betraf 2% der Bilder, wenn man einen Überdeckungsgrad von bis zu 20% als eine solche Fehlklassifikation ansieht.

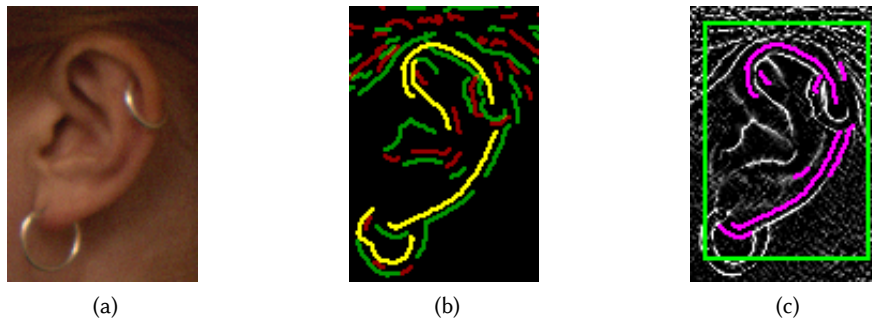


Abbildung 6.10: Beispiel einer korrekten Detektion trotz Ohringen als Störfaktoren. In (a) ist die Farbaufnahme des Ohres zu sehen, anhand der die Ohringe gut zu erkennen sind, (b) zeigt die extrahierten Komponenten und die gewählten Kandidaten, während in (c) das von den Ohringen unbeeinflusste Detektionsergebnis dargestellt ist.

Keine Möglichkeit das Ohr zu detektieren bestand für den Algorithmus, wenn der Kamerasensor die Oberfläche des Ohres, beziehungsweise den Bereich direkt außerhalb der Helix nur lückenhaft erfasst hat, wie es in Abbildung 6.11(a) zu sehen ist. Zur Berechnung der Flächenkrümmungen konnten diese fehlenden Bereiche nicht verwendet werden, so dass dort keine Flächenkrümmungsmaxima ermittelt werden konnten. Es gab entsprechend keine Komponenten, die zu einem Ohrumriss kombiniert werden konnten (siehe Abbildung 6.11(b)).

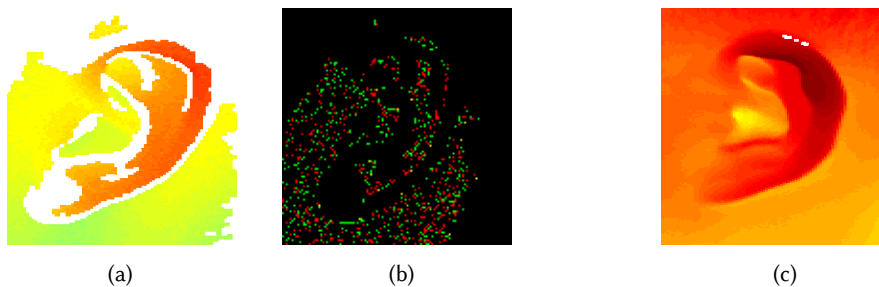


Abbildung 6.11: Stark mit Löchern durchsetzte aufgenommene Oberfläche eines Ohres (a) und die daraus resultierenden Flächenkrümmungsmaxima (b), bei denen keine zusammenhängenden Bereiche vorhanden sind. Alle Positionen, für die der Sensor des Aufnahmegerätes keine Informationen erhalten hat, sind weiß dargestellt. Die Farben geben die Entfernung zum Sensor an, wobei rot für die nächsten Bereiche steht. In (c) ist zum Vergleich eine Oberfläche mit sehr viel weniger Lücken dargestellt.

Die Haare der aufgenommenen Personen erwiesen sich neben den Lücken in der Oberfläche als weitere Ursache für Fehlklassifikationen. Problematisch waren Haare, wenn sie selbst starke Flächenkrümmungen zeigten und der Ohrumriss sich kaum von den umliegen-

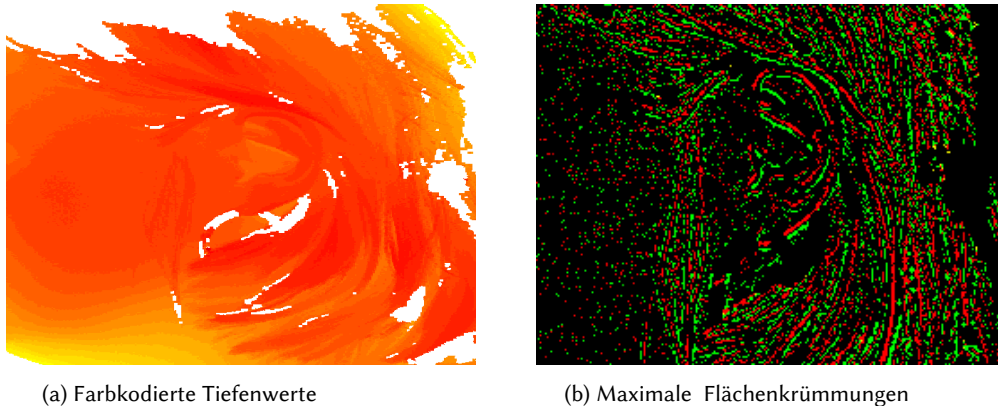


Abbildung 6.12: Beispiel für die Ausgangsbedingungen einer durch Haare bedingten Fehlklassifikation: große Flächenkrümmungen im Bereich der Haare und geringe Unterschiede in den Tiefenwerten der Haare und des Ohrumrisses.

den Haaren abhob, wie auch das Beispiel in Abbildung 6.12 zeigt. In diesem Fall war es dem Algorithmus nicht möglich, die Ohrregion zu finden, da die Komponenten in diesem Bereich durch die auf gleicher Höhe liegenden Haare sehr stark unterbrochen waren. Es konnten an vielen Stellen des Ohrumrisses keine großen Flächenkrümmungen ermittelt werden. Das zugehörige Detektionsergebnis ist in Abbildung 6.13(a) dargestellt. Der beste Kandidat erhielt eine absolute Bewertung von 0,01, woraus sich mit der relativen Bewertung eine Gesamtbewertung von 0,26 ergab, da er der einzige Kandidat mit einer Bewertung größer 0 und damit der längste betrachtete Kandidat war. Auf Grund der sehr niedrigen absoluten Bewertung wurde ausgeschlossen, dass es sich um einen tatsächlichen Ohrumriss handeln könnte (siehe Abschnitt 4.9.4), so dass keine Ohrregion detektiert wurde.

ZU GROßE UND ZU KLEINE OHRREGIONEN

Bei weiteren fehlerhaft detektierten Ohrregionen (Überlappungsgrad kleiner als 60%) war dennoch das ganze Ohr zusammen mit weiteren nicht zum Ohr gehörenden Bereichen in der Ergebnisregion vorhanden oder es wurde ein Teil des Ohres detektiert, während andere Teile fehlten. Die Klassifikation war also nicht vollständig fehlgeschlagen, da die richtige Struktur als Ohr erkannt wurde. Beispiele solcher Detektionen werden in diesem Abschnitt vorgestellt.

Haare in direkter Umgebung des Ohres führten unter bestimmten Umständen zu einer detektierten Ohrregion, die zu groß war. Dabei traten verschiedene Fälle auf: Zum einen wurden Shapes des Ohrumrisses mit nicht zum Ohr gehörenden Shapes (meistens Haaren) verbunden, da die Endpunkte nahe beieinander lagen und die Richtungen zueinander pas-

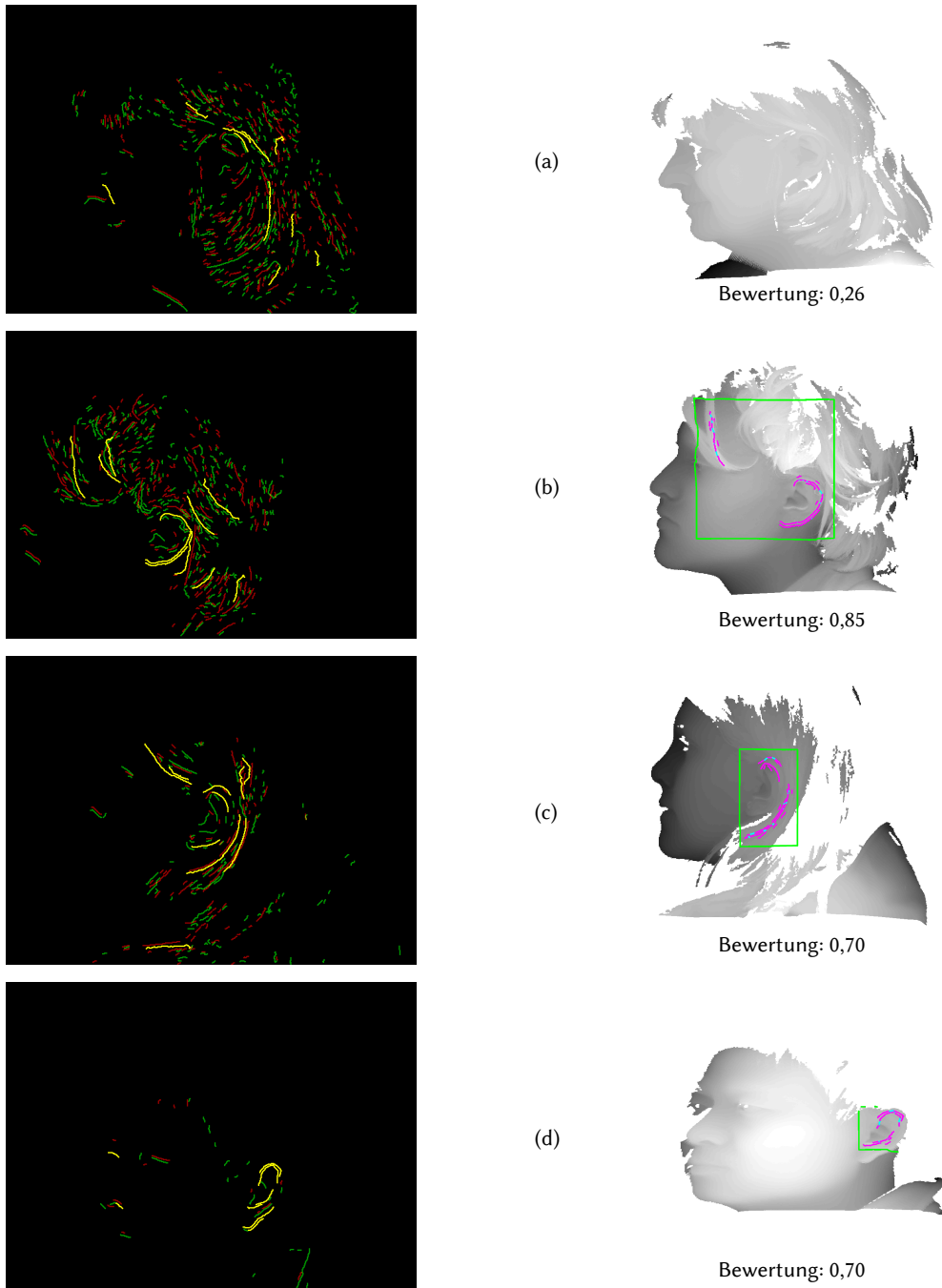


Abbildung 6.13: Beispiele fehlerhafter Detektionen des Ohrumrisses: (a) kein Ohr gefunden, (b) & (c) Haare zum Ohrumriss gezählt, (d) Teil des Ohres als vollständiges Ohr klassifiziert.

6 Ergebnisse & Diskussion

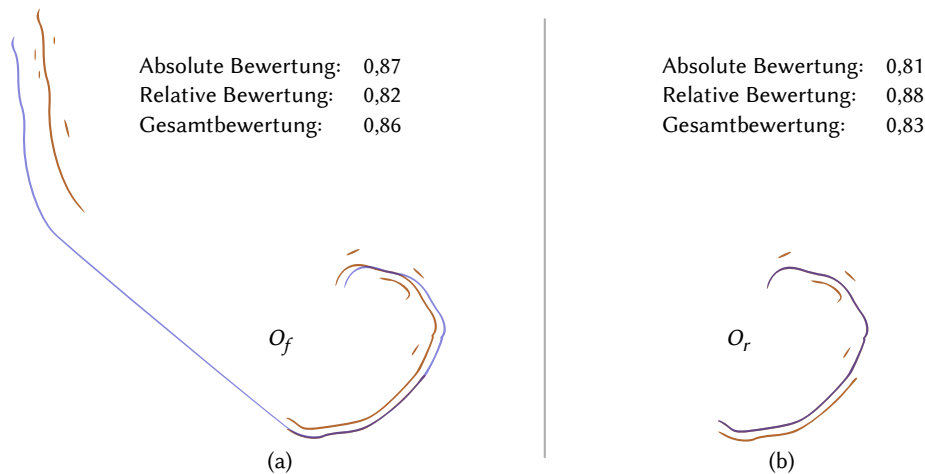


Abbildung 6.14: Darstellung des angepassten Shapes (blau) und der Original-Shapes (braun) von zwei Ohrenkandidaten des selben Eingabebildes. An den dunklen Positionen sind sowohl Pixel des angepassten Shapes, als auch der Original-Shapes vorhanden.

sten, zum anderen wurden abschnittsweise parallele Shapes oder neben anderen Shapes liegende Shapes kombiniert, obwohl sie ebenfalls nicht zum Ohr gehörten. In Abbildung 6.13(b) ist ein Detektionsergebnis zu sehen, bei dem ein durch Haare entstandener Shape zur Ohrregion gezählt wird, so dass diese deutlich zu groß ist. Der Grund für dieses Detektionsergebnis und die sehr gute Bewertung des Ohrenkandidaten, der keinen Ohrumriss darstellt, zeigt sich in der Darstellung des Kandidaten O_f in Abbildung 6.14(a). Der fälschlicherweise dazugenommene Shape konnte mit geringer Verschiebung und ohne das Auftreten von Ecken mit den anderen Shapes kombiniert werden, so dass der dadurch erzeugte angepasste Shape, der in der Abbildung blau dargestellt ist, die Kriterien der Ohrähnlichkeit sehr gut erfüllen kann. Im Vergleich mit einem den tatsächlichen Ohrumriss darstellenden Kandidaten O_r (siehe Abbildung 6.14(b)), weist O_f eine höhere Gesamtkrümmung auf, was zu einer leicht besseren Bewertung führte, da die Gesamtkrümmung von O_r noch unterhalb des optimalen Bereichs war.

Bei der Detektion, deren Ergebnis in Abbildung 6.13(c) dargestellt ist, wählte der Algorithmus die rechts neben dem Ohr liegenden Haare statt dem unteren Bereich der Helix als Bestandteil des Ohrumrisses aus. Da der entsprechende Shape sehr nah neben der Helix angeordnet war, eine ähnlich zur Helix verlaufende Krümmung aufwies und in weiten Teilen parallel verlaufende Shapes gefunden werden konnten, bestand keine Möglichkeit diese Struktur von einem tatsächlichen Bestandteil des Ohrumrisses zu unterscheiden.

Neben dem Fall zu großer detektierter Regionen trat das Fehlen einiger Teile des Ohres in der Ergebnisregion auf. Es wurden dann zwar jeweils nur Komponenten als Bestandteile

des Ohres klassifiziert, die tatsächlich zum Ohr gehören, aber es konnten nicht alle Bestandteile gefunden werden. Zu kleine Regionen wurden beispielsweise dann detektiert, wenn die Oberflächenkrümmungen am oberen oder unteren Teil der Helix so schwach ausgeprägt waren, dass an diesen Stellen keine Komponenten ermittelt werden konnten.

In Abbildung 6.13(d) ist zu erkennen, dass bereits ein Teil des Ohres die definierten Bewertungskriterien gut erfüllen kann. Die absolute Bewertung des besten Kandidaten lag in diesem Fall bei 0,61. Die Bevorzugung größerer Kandidaten, durch das in Abschnitt 4.9.2 eingeführte relative Bewertungskriterium war nicht ausreichend, für eine erfolgreiche Detektion, da der Verlauf der äußeren Flächenkrümmungen der Helix im oberen Bereich des Ohres vollständig fehlte, so dass kein sinnvoller größerer Kandidat konstruiert werden konnte. Das Fehlen von Maxima der Flächenkrümmungen in weiten Teilen des äußeren Bereichs des Ohres zeigt hier ein Problem des Algorithmus auf, das bei Drehungen des abgebildeten Kopfes auftritt. Wenn hinter dem Ohr keine Haut mehr ist, kann der Kamerasensor dort keine Informationen aufnehmen und die Flächenkrümmungen können für die Randbereiche des Ohres keine Maximalwerte ergeben. Kopfposen, die nicht einem Vollprofil entsprechen, bereiten demnach eher Probleme durch fehlende Daten hinter dem Ohr, als durch die perspektivische Verzerrung des Ohrumrisses, wie auch die gute Bewertung des besten gefundenen Kandidaten zeigt.

ZUSAMMENFASSUNG

Detektionsprobleme lassen sich, wie die Darstellung der Ergebnisse in den vorigen Abschnitten zeigt, auf zwei grundsätzliche Ursachen zurückführen. Zum einen das Fehlen von Oberflächeninformationen im Bereich des Ohres, beziehungsweise in direkter Umgebung der Helix. Da dann an diesen Stellen keine großen Flächenkrümmungen berechnet werden können, fehlen Komponenten, die möglicherweise für eine erfolgreiche Detektion nötig sind. Zum anderen können insbesondere Haare große Flächenkrümmungen verursachen, was mitunter zu einer großen Anzahl durchaus langer Komponenten führt. Je stärker diese beiden Dinge ausgeprägt waren, desto eher schlug die Detektion der Ohrregion fehl.

*Gibt es ein Ohr so fein, dass es die
Seufzer der welkenden Rose zu hö-
ren vermöchte?*

Arthur Schnitzler

7

FAZIT

In diesem abschließenden Kapitel wird ein Ausblick auf mögliche Aktivitäten zur Verbesserung des vorgestellten Algorithmus gegeben und die erzielten Ergebnisse werden kurz zusammengefasst.

7.1 AUSBLICK

Der implementierte Algorithmus liefert zwar gute Detektionsergebnisse (siehe Kapitel 6), es bleibt aber noch Raum für Verbesserungen. Mögliche Erweiterungen werden in diesem Abschnitt erläutert.

Fehler bei der Detektion der Ohrregion werden, wie in Abschnitt 6.2.5 dargestellt ist, häufig durch Lücken in der aufgenommenen Oberfläche verursacht. Damit die Detektion auch bei lückenhaften Aufnahmen funktioniert, wäre eine Vorverarbeitung der Bilder denkbar, bei der beispielsweise schmale Lücken überbrückt oder schmale Stege, wie sie im oberen Bereich des Ohres bei Abbildung 6.11(a) auftreten, als stark gekrümmte Bereiche behandelt werden. Es wäre auch möglich ein weiteres Binärbild aller Übergänge zu Bereichen fehlender Daten zu erzeugen und die Komponenten dieses Bildes ebenfalls zum Vervollständigen des Ohrumrisses zu benutzen. Da der gesamte Umriss des Kopfes einen Rand der aufgenommenen Daten darstellt, müssten weitere Kriterien zum Behandeln dieser Komponenten definiert werden.

Die absolute Bewertung der Ohrenkandidaten ließe sich möglicherweise durch die Betrachtung des Krümmungsverlaufs, zusätzlich zur Gesamtkrümmung, verbessern. Der Verlauf der Krümmungen dürfte sich dabei nicht sehr gut zur Unterscheidung ohrähnlicher Strukturen von anderen Strukturen eignen, da der Verlauf der Helix und damit der Verlauf der Krümmungen insbesondere im Bereich des Ohrläppchens sehr unterschiedlich sein kann (siehe auch Abschnitt 4.7). Es könnte aber besser abgeschätzt werden, wann die ermittelte Gesamtkrümmung für ein vollständiges Ohr steht. Auch die relative Bewertung bietet noch Verbesserungspotential, da durch die Wahl der beiden Exponenten (siehe Abschnitt 4.9.2) eine Abhängigkeit von der Skalierung der Eingabedaten besteht. Als Richtwert für den erlaubten Umfang der Verschiebung von Shapes könnte der Abstand gelten, der zwischen parallelen Shapes des Ohrumrisses besteht.

Bisher wird das Ohr nur über den Verlauf der Helix detektiert, eine Betrachtung der gegenüberliegenden Seite der Ohrmuschel findet nicht statt. Mit dem Tragus existiert dort aber ebenfalls eine bezüglich ihrer Oberfläche prominente Struktur, die eine starke Flächenkrümmung aufweist. Zusätzlich sollten in direkter Umgebung des Tragus lokal größte Entfernungen der Oberfläche vom Sensor auftreten, je nach Perspektive verursacht durch den äußeren Gehörgang oder das Cavum Conchae. Die Existenz einer Tragus-ähnlichen Struktur in einem durch den detektierten Ohrumriss definierten Bereich, könnte als weiteres Indiz für eine korrekte Ermittlung des Ohrumrisses gewertet werden und damit für eine bessere Unterscheidbarkeit korrekter und fehlerhafter Ohrenkandidaten sorgen.

Das Verhalten des Algorithmus bei Abbildungen im Dreiviertel- oder Halbprofil konnte auf Grund fehlender Testdaten bisher nicht untersucht werden. Dies zu untersuchen und gegebenenfalls die Robustheit der Detektion gegen Posenänderungen zu verbessern wäre ein sinnvoller weiterer Schritt.

7.2 ZUSAMMENFASSUNG

Das in dieser Arbeit entwickelte Konzept zur Detektion von Ohren auf dreidimensionalen Profilbildern lieferte auf der UND-J2-Bilddatenbank mit einer Detektionsrate von 94% sehr gute Ergebnisse. Damit konnte gezeigt werden, dass eine erfolgreiche Detektion bei ausschließlicher Verwendung der Tiefeninformationen möglich ist, ohne dass ein Klassifikator trainiert werden muss und ohne dass zusätzlich Farbinformationen benötigt werden. Durch die in Abschnitt 7.1 beschriebenen Änderungen sind weitere Verbesserungen des Detektionsalgorithmus möglich, so dass eine noch höhere Detektionsrate erreicht werden könnte.

7 Fazit

Die mittleren Flächenkrümmungen erwiesen sich damit als ein für die Ohrendetektion geeignetes Merkmal, wenn sinnvolle Kriterien zur Bewertung der aus den Flächenkrümmungen ermittelten Zusammenhangskomponenten definiert werden. Dabei reichen, wie in Kapitel 6 gezeigt wurde, bereits sehr einfache Kriterien wie die Proportion oder die Gesamtkrümmung aus, um die wenigen stark gekrümmten Strukturen unterscheiden zu können, die auf dreidimensionalen Profilbildern zu finden sind.

Ein auf dieser Arbeit basierendes Paper wurde bei der Konferenz „IHMSP 2012“ [III 2011] eingereicht.

8

ANHANG

8.1 VERGLEICH VERSCHIEDENER SCHWELLWERTE

Die Testergebnisse für verschiedene Schwellwerte S bei zwei in Kapitel 4 vorgestellten Binarisierungsverfahren (siehe Abschnitt 4.2.2) werden in diesem Abschnitt aufgeführt. Dies ist zum einen die Binarisierung mit einem globalen Schwellwert (siehe Tabelle 8.1) und zum anderen die Binarisierung mit einem Hysterese-Schwellwertverfahren (siehe Tabelle 8.2).

Tabelle 8.1: Detektionsrate bei der Binarisierung mit verschiedenen globalen Schwellwerten S . Die Detektionsrate ist jeweils für 60% Überlappung zwischen Ergebnis- und Ground-Truth-Region (siehe Abschnitt 6.1) angegeben. Die Zeile des besten Ergebnisses ist hervorgehoben.

S	Detektionsrate
0,3	90,60%
0,4	93,16%
0,5	94,03%
0,6	91,80%
0,7	80,07%

Das beste Ergebnis für einen Schwellwert wird bei $S = 0,5$ erreicht, wobei eine Erhöhung des Schwellwertes für deutlich schlechtere Ergebnisse sorgt, als die Verwendung eines etwas niedrigeren Wertes.

Tabelle 8.2: Detektionsrate bei der Binarisierung mit einem Hysterese-Schwellwertverfahren mit zwei Schwellwerten S_h und S_l . Auch hier bezieht sich die Detektionsrate auf 60% Überlappung und die Zeile des besten Ergebnisses ist hervorgehoben.

S_h	S_l	Detektionsrate
0,5	0,2	92,05%
0,5	0,3	93,12%
0,5	0,4	94,28%
0,6	0,4	94,41%
0,6	0,5	93,74%
0,7	0,4	94,28%
0,7	0,5	92,75%

Bei der Verwendung des Hysterese-Schwellwertverfahrens konnte das beste Ergebnis für $S_h = 0,6$ und $S_l = 0,4$ erreicht werden. Vergleicht man das beste Ergebnis der Binarisierung mit einem Schwellwert mit dem besten Ergebnis des Hysterese-Schwellwertverfahrens, erweist sich das Hysterese-Verfahren als etwas besser.

8.2 VERGLEICH VERSCHIEDENER STRATEGIEN ZUR AUSWAHL PROMINENTER SHAPES

In diesem Abschnitt werden die Ergebnisse von Tests verschiedener Möglichkeiten bei der Auswahl prominenter Shapes (siehe Abschnitt 4.4) dargestellt. Die Tests wurden auf einem Ausschnitt der Bilddatenbank ausgeführt, der von jeder Person genau eine zufällig gewählte Aufnahme und damit insgesamt 404 Aufnahmen enthält.

Getestet wurde die Auswahl prominenter Shapes mit der Berücksichtigung des Auftretens von Parallelität als Merkmal eines prominenten Shapes, mit dem Aufspalten der Shapes an detektierten Ecken und dem Aufspalten bei starken Krümmungsänderungen an den Wendepunkten im Verlauf der Krümmungen. Zum Vergleich ist in Abbildung 8.1 auch die Detektionsrate bei der Auswahl prominenter Shapes ohne Berücksichtigung von Parallelität oder dem Aufspalten von Shapes dargestellt.

Das beste Ergebnis konnte durch das Aufspalten an Wendepunkten erreicht werden. Die Unterschiede der ermittelten Detektionsraten sind, mit Ausnahme der Verwendung von Parallelität, aber sehr gering. Das Aufspalten von Shapes nach den beiden angewandten Kriterien, hat demnach kaum einen Einfluss auf das Detektionsergebnis.

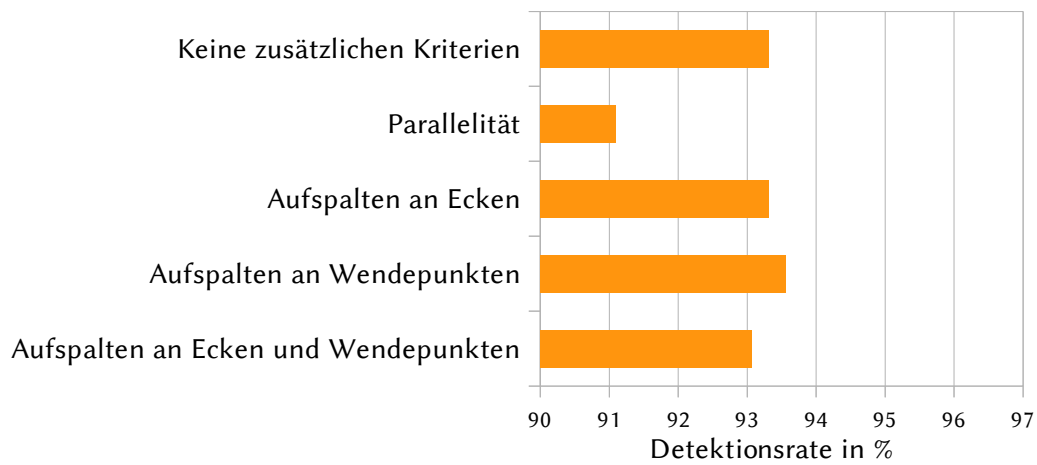


Abbildung 8.1: Testergebnisse verschiedener Variationen bei der Auswahl prominenter Shapes. Die Tests wurden auf einem 404 Bilder umfassenden Ausschnitt der UND-J2-Datenbank durchgeführt. Die Detektionsrate bezieht sich auf 60% Überlappung zwischen Ergebnis- und Ground-Truth-Region (siehe Abschnitt 6.1).

LITERATURVERZEICHNIS

- [Ver 2010] *Verbundskizze zur Bekanntmachung Biometrie*. Mai 2010
- [IIH 2011] *The Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. <http://iihmsp12.unipi.gr>. 2011, zuletzt besucht am 23.01.2012
- [3D F 2009] 3D Face Consortium: *Harmonized Biometric Vocabulary*. <http://www.3dface.org/media/vocabulary.html>. Oktober 2009, zuletzt besucht am 19.01.2012
- [ABAZ 2010] Abaza, A.; Hebert, C. ; Harrison, M.A.F.: Fast Learning Ear Detection for Real-time Surveillance. In: *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS 2010)*, 2010, S. 1–6
- [ANSA 2007] Ansari, S.; Gupta, P.: Localization of Ear Using Outer Helix Curve of the Ear. In: *International Conference on Computing: Theory and Applications, 2007. ICCTA '07*, 2007, S. 688–692
- [ARBA 2007] Arbab-Zavar, B.; Nixon, M.: On Shape-Mediated Enrolment in Ear Biometrics. 2007. http://dx.doi.org/10.1007/978-3-540-76856-2_54. In: *Advances in Visual Computing* Bd. 4842. Berlin / Heidelberg : Springer, 2007, 549–558
- [BANS 2010] Bansal, Roli; Sehgal, Priti ; Bedi, Punam: Effective Morphological Extraction of True Fingerprint Minutiae based on the Hit or Miss Transform. In: *International Journal of Biometrics and Bioinformatics* 4 (2010), Nr. 2, S. 71–85
- [BESL 1988] Besl, Paul J.; Jain, Ramesh (Hrsg.): *Surfaces in Range Image Understanding*. New York : Springer, 1988 (Springer Series in Perception Engineering)
- [BHAN 2008] Bhanu, Bir; Chen, Hui; Singh, Sameer (Hrsg.): *Human Ear Recognition by Computer*. London : Springer, 2008 (Advances in Computer Vision and Pattern Recognition). <http://dx.doi.org/10.1007/978-1-84800-129-9>. <http://dx.doi.org/10.1007/978-1-84800-129-9>

- [BIOM 2010] Biometrics, ISO/IEC JTC1 S.: *ISO/IEC SC37 SD11 Overview Standards Harmonization Document*. International Organization for Standardization, 08 2010. – N 3972
- [CASE 2010] CASED: *IT-Sicherheitsforschung und -Entwicklung in Darmstadt*. <http://www.cased.de/ueber.html>. September 2010, zuletzt besucht am 24.11.2011
- [CHEN 2007] Chen, Hui; Bhanu, B.: Human Ear Recognition in 3D. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (2007), April, Nr. 4, S. 718–737. <http://dx.doi.org/10.1109/TPAMI.2007.1005>
- [DORA 1995] Dorai, Chitra; Jain, Anil K.: COSMOS-a representation scheme for free-form surfaces. In: *Fifth International Conference on Computer Vision*, 1995, S. 1024–1029
- [HE 2008] He, Xiao C.; Yung, Nelson H. C.: Corner detector based on global and local curvature properties. In: *Optical Engineering* 47 (2008), Mai, Nr. 5
- [IANN 1989] Iannarelli, Alfred V.: *Ear identification*. Paramount Publishing Company, 1989
- [ISLA 2008] Islam, S.M.S.; Bennamoun, M. ; Davies, R.: Fast and Fully Automatic Ear Detection Using Cascaded AdaBoost. In: *Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on*, 2008, S. 1–6
- [JAHN 1997] Jähne, Bernd: *Digitale Bildverarbeitung*. Berlin Heidelberg : Springer-Verlag, 1997
- [JAHN 2005] Jähne, Bernd: *Digitale Bildverarbeitung*. Berlin Heidelberg : Springer-Verlag, 2005
- [JAIN 2007] Jain, A. K. (Hrsg.); Flynn, P. (Hrsg.) ; Ross, A. A. (Hrsg.): *Handbook of Biometrics*. Springer, 2007
- [JEGE 2007] Jeges, Ernő; Máté, László: Model-Based Human Ear Localization and Feature Extraction. In: *IC-MED 1* (2007), Nr. 2, S. 101–112
- [KUMA 2011] Kumar, Ajay; Wu, Chenye: *Automated human identification using ear imaging*. <http://dx.doi.org/10.1016/j.patcog.2011.06.005>. Juni 2011
- [MOKH 1998] Mokhtarian, Farzin; Suomela, Riku: Robust Image Corner Detection Through Curvature Scale Space. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (1998), Dezember, Nr. 12, S. 1376–1381

- [NIXO 2002] Nixon, Mark S.; Aguado, Alberto S.: *Feature Extraction and Image Processing*. Oxford : Newnes, 2002
- [SHIH 2009] Shih, Hao-Chung; Ho, C.C.; Chang, H.T. ; Wu, Chin-Song: Ear Detection Based on Arc-Masking Extraction and AdaBoost Polling Verification. In: *Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2009)* Bd. ear detection, 2009, S. 669–672
- [VIOL 2001] Viola, Paul; Jones, Michael: Rapid Object Detection using a Boosted Cascade of Simple Features. In: *Conference on Computer Vision and Pattern Recognition*, 2001
- [VISL 2009] VISLab: *Visualization and Intelligent Systems Laboratory*. <http://www.vislab.ucr.edu>. 2009, zuletzt besucht am 02.01.2012
- [YAN 2007] Yan, Ping; Bowyer, K.W.: Biometric Recognition Using 3D Ear Shape. In: *Pattern Analysis and Machine Intelligence* 29 (2007), August, S. 1297–1308. <http://dx.doi.org/10.1109/TPAMI.2007.1067>
- [YOSH 2002] Yoshida, Hiroyuki; Masutani, Yoshitaka; MacEneaney, Peter; Rubin, David T. ; Dachman, Abraham H.: Computerized Detection of Colonic Polyps at CT Colonography on the Basis of Volumetric Features: Pilot Study1. In: *Radiology* 222 (2002), Nr. 2, 327-336. <http://dx.doi.org/10.1148/radiol.2222010506>
- [ZEPP 2003] Zeppenfeld, Klaus: *Lehrbuch der Grafikprogrammierung: Grundlagen, Programmierung, Anwendung*. Spektrum Akademischer Verlag, 2003
- [ZHOU 2010] Zhou, Jindan; Cadavid, S. ; Abdel-Mottaleb, M.: Histograms of Categorized Shapes for 3D ear detection. In: *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, 2010