

Non-invasives Wachstums-Monitoring
von Steinbutten bei hoher
Populationsdichte in der Aquakultur
mittels eines Stereokamerasystems

Dissertation

zur Erlangung des akademischen Grades
Doktor der Agrarwissenschaften
(Dr. agr.)

der Agrar- und Ernährungswissenschaftlichen
Fakultät
der Christian-Albrechts-Universität zu Kiel

Dipl.-Inf. Christian BAUER

Kiel

2013

Gedruckt mit Genehmigung der Agrar- und Ernährungswissenschaftlichen Fakultät der Christian-Albrechts-Universität zu Kiel

1. Gutachter: Professor Dr. Eberhard HARTUNG, CAU Kiel
2. Gutachter: Professor Dr. Carsten SCHULZ, CAU Kiel

Datum der mündlichen Prüfung: 24.04.2013

Danksagung

Hiermit möchte ich meinen Dank an meinen Doktorvater Prof. Dr. Eberhard Hartung und meinen Betreuer Dr. Eiko Thiessen für ihre Unterstützung und Beratung in vielen Fragen richten, wodurch die Fertigstellung dieser Arbeit überhaupt erst möglich wurde.

Weiterer Dank gilt dem Institut für landwirtschaftliche Verfahrenstechnik und der Christian-Albrechts-Universität zu Kiel, die mir im Rahmen meiner Anstellung als wissenschaftlicher Mitarbeiter die Möglichkeit zur Realisierung dieser Doktorarbeit gaben. Weiter danken möchte ich noch der GMA Büsum für die Unterstützung beim Einsatz des Prototypen.

Neben allen Kollegen für das angenehme Arbeitsklima möchte ich explizit den Technikern und der Werkstatt des Instituts für die technische Realisierung des Prototypen danken.

Ich möchte meinen Eltern für all ihre Liebe und Unterstützung danken, dass sie mir das Studium ermöglichten und mich in kein vorgeplantes Bild zu zwängen versuchten.

Besonderer Dank gilt Nine, die mich in allen schweren Phasen stets unterstützt und aufgebaut hat, meine Launen ertrug, mir mit Rat und Tat beiseite stand und trotz all dem bereit war, den nächsten Schritt zu gehen und meine Frau zu werden.

Inhaltsverzeichnis

1	Ziel der Arbeit	1
1.1	Motivation	1
1.2	Fokus der Arbeit	2
2	Wissensstand	5
2.1	Steinbutt (Scophthalmus maximus)	5
2.1.1	Steinbutt in der Aquakultur	8
2.1.1.1	Aquakultur	8
2.1.1.2	Zahlen und Fakten	9
2.1.1.3	Produktion von Steinbutt	10
2.2	Grundlagen	12
2.2.1	Bildentstehung	12
2.2.1.1	Kameramodell	12
2.2.1.2	Projektive Geometrie	13
2.2.1.3	Epipolargeometrie	15
2.2.1.4	Stereopsis	18
2.2.1.4.1	Rektifizierung	18
2.2.1.4.2	Disparität	19
2.2.1.5	Robuste Schätzer	20
2.2.2	Bildverarbeitung	21
2.2.3	Mustererkennung	25
2.2.3.1	Klassifikation	26
2.2.3.2	Formerkennung	28
2.3	Aktueller Stand der Technik	30
2.3.1	Bildverarbeitung im Kontext der Fischdetektion	30
2.3.2	Unterwasser-Kalibrierung von Stereokamerasystemen	34
2.3.3	Stereosysteme	35
2.3.4	Konturbasierte Formerkennung	37
3	Aquakultur-Monitoring von Fischwachstum	39
3.1	Anforderungen an das Stereokamerasystem	39
3.2	Material: Hardwaresystem	42
3.2.1	Planung des Stereokamerasystems	42
3.2.2	Resultierendes Stereokamerasystem	48
3.2.2.1	Unterwasser-Stereokamera	48
3.2.2.2	Beleuchtungssystem	49
3.2.2.3	Stromversorgung	52

3.2.3	Systemaufbau am Versuchsstandort	54
3.3	Methoden: Parameter des Beleuchtungssystems und Unterwasser-Kalibrierung	56
3.3.1	Versuchsaufbau zur Farbwahl des Beleuchtungssystems	57
3.3.2	Versuchsaufbau zur Ausrichtung des Beleuchtungssystems	60
3.3.3	Kalibrierung des Stereokamerasystems für Unterwasser-Einsatz	62
3.4	Ergebnisse: Parameter des Beleuchtungssystems und Unterwasser-Kalibrierung	66
3.4.1	Ergebnisse zur Farbwahl des Beleuchtungssystems . . .	66
3.4.2	Ergebnisse zur Ausrichtung des Beleuchtungssystems .	77
3.4.3	Evaluation der Kalibrierung	78
4	Optische Fischvermessung	83
4.1	Ausgangssituation und Überblick	83
4.2	Vorverarbeitung	87
4.2.1	Separierter nichtlinearer Gaußfilter	87
4.2.2	Histogrammäqualisierung	88
4.2.3	Canny Edge Detector	89
4.3	Fischerkennung	91
4.3.1	Training	91
4.3.1.1	Trainingsdaten	91
4.3.1.2	Kostenberechnung	93
4.3.1.3	Ablauf des Trainings	94
4.3.2	Clustering	95
4.3.3	Matching	96
4.3.4	Validierung	97
4.3.5	3D-Vermessung	99
4.4	Evaluation der Parameter	100
4.4.1	Ermittlung des Schwellwerts zur Bestimmung repräsentativer Konturen - Versuch	100
4.4.2	Ermittlung des Schwellwertes zur Bestimmung repräsentativer Konturen - Ergebnisse	103
4.4.3	Erkennungsraten des Konturen-Matchings - Versuch . .	105
4.4.4	Erkennungsraten des Konturen-Matchings - Ergebnisse	107
5	Diskussion	109
5.1	Entwicklung Hardwaresystem	109
5.2	Algorithmus zur Steinbuttvermessung	113

Inhaltsverzeichnis	vii
<hr/>	
6 Ausblick	121
6.1 Hardwaresystem	121
6.2 Algorithmus zur Steinbuttvermessung	122
7 Zusammenfassung	125
7.1 Zusammenfassung	125
7.2 Summary	126
A Anhang	129
A.1 Ergebnistabellen Versuche zur Farbwahl	129
Literaturverzeichnis	137

Abbildungsverzeichnis

2.1	Schemadarstellung des Steinbutts aus [41].	6
2.2	Steinbutt aus Aquakultur.	6
2.3	Links: Netzkäfigsystem mit Lachsen. Rechts: Teichwirtschafts- system. ¹	8
2.4	Kreislaufsystem mit Wasseraufbereitungsanlage. ²	9
2.5	Entwicklung der Steinbuttproduktion in der Aquakultur. Dia- gramm aus [41].	10
2.6	Produktionskreislauf des Steinbutts von der Brut bis zum ver- kaufsfertigen Fisch. Illustration aus [41].	11
2.7	Schematische Darstellung des Lochkameramodells.	13
2.8	Schematische Darstellung der Epipolargeometrie.	16
2.9	Stereobildpaar, bei dem der Holzquader aufgrund der unter- schiedlichen Perspektiven an unterschiedlichen Positionen er- scheint.	18
2.10	Beispiel für die Rektifizierung. Links: Originalbild. Rechts: Rek- tifiziertes Bild.	19
2.11	Schematische Darstellung der Disparität, mit dl, dr als linke und rechte Disparität, b als Länge der Grundlinie, f als Brenn- weite und r als Distanz zum Objekt.	20
2.12	Beispiele für verschiedene Punktoperatoren. Oben links: Ori- ginal. Oben rechts: Invertierung. Unten links: Binarisierung. Un- ten rechts: Histogrammäqualisierung.	22
2.13	Beispiele für verschiedene lokale Operatoren. Oben links: Ori- ginal. Oben rechts: Gaußfilter. Unten links: Sobelfilter. Unten rechts: Canny Edge Detector.	24
2.14	Beispiele für verschiedene morphologische Operatoren unter Ver- wendung einer quadratischen Maske. Von links nach rechts: Ori- ginal, Erosion, Dilatation, Öffnung, Schließung.	24
3.1	Winkeländerung aufgrund Lichtbrechung zum Beispiel beim Wechsel von Luft in Wasser [26].	43
3.2	Durch das Gehäuse verdeckte Randbereiche bei Fisheye-Objek- tiven verkleinern den effektiven Bildausschnitt.	44
3.3	Schema zur Berechnung der Baseline b	45
3.4	Beide Kameras vom Typ GigE uEye 5240SE-M mit Objektiven der Sorte Pentax H416 (KP) montiert auf Metallschiene für variable Positionierung.	48

3.5	Oberseite des Gehäuses mit Kabelzuführung und Halterung für das Fischbecken.	49
3.6	Links: Dichtung zum wasserdichten Verschließen des Deckels. Rechts: Komplett geschlossenes Gehäuse.	50
3.7	Konstruktionszeichnung der LED-Halterung.	50
3.8	Konstruktionszeichnung des einzelnen LED-Gehäuses. A: Flüssigkeitsschauglas. B: Kühlkörper aus Aluminium. C: Schlauchverschraubung POM, Blau.	51
3.9	Links: Kameragehäuse in der Mitte. Außen LED-Halterung mit 8 LED-Gehäusen. Rechts: Einzelnes LED-Gehäuse mit einstellbarem Winkel.	52
3.10	Netzgerät und Präzisions-Stromregelungseinheit im Edelstahlgehäuse.	53
3.11	Links: Aufgebautes Stereokamerasystem über Fischbecken. Rechts: Recheneinheit im Serverraum verbunden mit dem Stereokamerasystem über das Firmennetzwerk.	54
3.12	Links: Fest montiertes Edelstahlgehäuse mit abgehenden Stromleitungen. Rechts: Verteilerkasten mit Anschluss an Stromnetz und Firmennetzwerk.	55
3.13	Halterung für Kameragehäuse und Beleuchtungssystem montiert an Trägerstange über Fischbecken.	55
3.14	Stereokamerasystem im Unterwasser-Einsatz bei der GMA in Büsum.	56
3.15	LED-Träger für Farblicht-Versuch. 6 LEDs kreisförmig auf Basisplatte aus Aluminium, die als Kühlkörper dient. Scheibe aus Plexiglas.	58
3.16	Die beim Versuch verwendeten Steinbutte. Links: Paar mit heller Musterung. Rechts: Paar mit dunkler Musterung.	59
3.17	Aufnahmen von Steinbutten bei farbigem Licht. Oben links: Blau. Oben mitte: Gelb. Oben rechts: Ultraviolett. Unten links: Weiß. Unten mitte: Rot. Unten rechts: Grün.	60
3.18	Aufnahmen bei unterschiedlichen Einstellungswinkeln der LEDs. Links: 20° nach innen gedreht. Rechts: senkrecht ausgerichtet.	61
3.19	Aufnahmen bei unterschiedlichen Einstellungswinkeln der LEDs. Links: 20° nach außen gedreht. Rechts: 40° nach außen gedreht.	61
3.20	Vignettierung des Pentax-Objektivs H416 (KP) aus [52].	62
3.21	4 Beispiele für Aufnahmen des Kalibrierusters mit unterschiedlichen Positionierungen und Ausrichtungen.	64
3.22	Ein Beispiel für die Rückprojektion. Die Position der gemessenen Punkte des Kalibrierusters werden durch grüne Kreise und die rückprojizierten Punkte durch rote Kreise dargestellt.	65

3.23	Mittlere Grauwerte eines Steinbutts bei einer Belichtungszeit von 10 ms. Ergebnisse sind die Mittelwerte aus 3 Aufnahmen je Variante.	67
3.24	Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei einer Belichtungszeit von 10 ms.	69
3.25	Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.	70
3.26	Empirischer Variationskoeffizient innerhalb eines Steinbutts bei einer Belichtungszeit von 10 ms.	72
3.27	Empirischer Variationskoeffizient innerhalb eines Steinbutts bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.	73
3.28	Überbelichtung der Randpartien eines Steinbutts bei weißem Untergrund.	74
3.29	Mittlere Gradientenstärke S_m im betrachteten Bildausschnitt am Übergang zwischen den Steinbutten bei einer Belichtungszeit von 10 ms.	76
3.30	Mittlere Gradientenstärke S_m im betrachteten Bildausschnitt am Übergang zwischen den Steinbutten bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.	76
3.31	Mittelwerte der Grauwerte mit ansteigendem Abstand vom Zentrum.	78
3.32	Beispiel für 2 mögliche berechnete Distanzen unter Verwendung des Kalibrierungsmusters.	79
3.33	Ergebnisse von 10 Testmessungen im Bereich der Bildmitte. Sollwert der Distanz beträgt 2,4 cm. Die Abweichung beträgt im Mittel 1,7 %.	80
3.34	Ergebnisse von 10 Testmessungen im Bereich des Bildrandes. Sollwert der Distanz beträgt 2,4 cm. Die Abweichung beträgt im Mittel 2 %.	80
4.1	Beispiele unterschiedlichen Wasserqualitäten. Links: Wasser klar. Mitte: Viele Schwebeteilchen behindern die Sicht. Rechts: Trübes Wasser mit schlechter Sicht.	84
4.2	Typische Aufnahme mit einem Sichtfeld komplett gefüllt mit Steinbutten.	85
4.3	Schematische Darstellung des Softwaresystems zur Erkennung und Vermessung von Steinbutten.	86
4.4	Beispiel für Histogrammäqualisierung. Links: Originalbild. Rechts: Nach Histogrammäqualisierung.	89

4.5	Zwei Beispiele für die Schwellwerte t_{low} und t_{high} beim Canny Edge Detector. Links: Original. Mitte: Niedrige Schwellwerte mit $t_{low} = 30$, $t_{high} = 60$. Rechts: Hohe Schwellwerte mit $t_{low} = 110$, $t_{high} = 150$	90
4.6	Schematische Darstellung der Trainingsphase des Softwaresystems.	92
4.7	Beispiele für 3 Trainingssets. Links: Trainingsbild mit Boundingbox, Objekt-Mittelpunkt, Maul- und Flossenspitze. Mitte: Trainingsset Positiv mit Objekt-Mittelpunkt. Rechts: Trainingsset Negativ ohne Steinbutt aber anderen Strukturen.	93
4.8	Funktionsweise Chamfer distance. Das Suchbild mit eingetragenen kürzesten Distanzen zur nächsten Kontur sowie zu matchende Kontur. Illustration aus [16].	94
4.9	Beispiele für 2 Cluster repräsentativer Kanten beim Steinbutt.	96
4.10	Nachbarschaftssuche von Mittelpunkten repräsentativer Konturen an einem einzelnen Beispiel.	98
4.11	Beispiele vom Trainingsset Negativ. Oben links: Leeres Becken mit Equipment. Oben rechts: Schwebeteilchen. Unten links: Vervielfachte Kaustik. Unten rechts: Vervielfachte Oberflächenstruktur eines Steinbutts.	101
4.12	Beispiele für Trainingsdaten bestehend aus Laboraufnahmen von Steinbutten.	102
4.13	Beispiele für Trainingsdaten von Aufnahmen in Büsum. Für eine bessere Erkennbarkeit wurden die Bilder nachbearbeitet.	102
4.14	Testvariante 1: Training von Laboraufnahmen. Testvariante 2: Training von Aufnahmen in realer Umgebung, Variante 1. Testvariante 3: Training von Aufnahmen in realer Umgebung, Variante 2. Testvariante 4: Training von Aufnahmen in realer Umgebung, Auswahl bester Trainingsbilder. Testvariante 5: Training von negativen Strukturen.	103
4.15	Zwei Beispiele der synthetischen Steinbuttkonturen. Die Konturen wurden für diese Darstellung verstärkt.	106
4.16	Ergebnisbilder des Konturen-Matchings. Links: Bild aus Validierungsset Labor. Rechts: Bild aus Validierungsset Büsum. Gematchten Konturen werden angezeigt und farbcodiert. Konturen des Validierungsbildes: weiß. Niedrige Matching-Kosten ($c_{match} < 0,3$): rot. Mittlere Ergebnisse ($c_{match} < 0,4$): grün. Hohe Ergebnisse ($c_{match} \geq 0,4$): blau. Konturen für Darstellung verstärkt.	108

5.1	Beispiele für Aufnahmen für Verfahren zur Fischerkennung. A: Lachs in Becken aus [140]. B: Thunfische in offener See aus [32]. C: Fisch auf Fließband aus [134]. D: Steinbutte in Becken der GMA Büsum.	116
5.2	Beispiel für den Einsatz von Differenzbildern. Oben: Zwei aufeinander folgende Aufnahmen aus Büsum. Unten links: Differenzbild. Unten rechts: Sobelfilter zur Kantendetektion auf erstem Bild.	118
5.3	Beispiel für den Einsatz von Differenzbildern. Oben: Zwei Aufnahmen aus Büsum mit circa 2 Sekunden Abstand, Steinbutt wechselt Position vollständig. Unten links: Differenzbild. Unten rechts: Sobelfilter zur Kantendetektion auf erstem Bild.	119

Tabellenverzeichnis

2.1	Anzahl x der Iterationen, die benötigt werden, damit mit einer Wahrscheinlichkeit von $p = 0,99$ eine Teilmenge aus s Elementen und entsprechendem Anteil von Ausreißern mindestens einmal keine Ausreißer enthält.	22
3.1	Farbtemperatur für neutral weiße LED beziehungsweise dominante Wellenlänge λ_D für andere LEDs inklusive Minimal- und Maximal-Wert.	57
3.2	Belichtungszeiten der verschiedenen Farbkombinationen in ms im automatischen Belichtungsmodus. Bei maximaler Belichtungszeit von 99,8 ms wird die zusätzlich benötigte Verstärkung angegeben, um einen mittleren Grauwert von 127 zu erreichen.	66
3.3	Anzahl der im 23×77 Pixel großen Bereich durch den Canny Edge Detector als Konturpixel erkannten Pixel.	75
4.1	Quantile der Varianten 1 bis 5 bei entsprechenden $p_1 \dots p_5$	104
4.2	Anzahl korrekt gematchter und skaliertes Konturen von 66 relevanten Konturen. Validierung mit Laborbildern. Labor: Trainingsdaten mit Laborbildern. Büsum: Trainingsdaten mit Bildern aus Büsum. Synt1: Synthetische Bilddaten trainiert mit Laborbildern. Synt2: Synthetische Bilddaten trainiert mit Bildern aus Büsum.	108
4.3	Anzahl korrekt gematchter und skaliertes Konturen von 72 relevanten Konturen. Validierung mit Bildern aus Büsum. Labor: Trainingsdaten mit Laborbildern. Büsum: Trainingsdaten mit Bildern aus Büsum. Synt1: Synthetische Bilddaten trainiert mit Laborbildern. Synt2: Synthetische Bilddaten trainiert mit Bildern aus Büsum.	108
A.1	Mittlerer Grauwert eines Steinbutts bei einer Belichtungszeit von 10 ms. Berechnung des Mittelwerts aus drei Aufnahmen pro Konfiguration von Steinbutt, Boden- und Lichtfarbe.	130
A.2	Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei einer Belichtungszeit von 10 ms.	131
A.3	Standardabweichung vom relativen Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei einer Belichtungszeit von 10 ms.	132

A.4	Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.	133
A.5	Standardabweichung vom relativen Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.	134
A.6	Empirischer Variationskoeffizient innerhalb eines Steinbutts bei einer Belichtungszeit von 10 ms.	135
A.7	Empirischer Variationskoeffizient innerhalb eines Steinbutts bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.	136

KAPITEL 1

Ziel der Arbeit

Inhalt

1.1 Motivation	1
1.2 Fokus der Arbeit	2

1.1 Motivation

Die Aquakultur entwickelt sich zu einem immer wichtiger werdenden Aspekt im Hinblick auf die Versorgung der Weltbevölkerung mit Eiweiß durch Speisefisch und Schalentieren. Während der Bedarf an Fisch ständig steigt, ist die Ertragskraft aus der Fangfischerei am stagnieren, überwiegend verursacht durch Fangquoten und Überfischung. Die Aquakultur belegt heute bereits ein Drittel der Weltproduktion an Fisch und ist weiterhin ein stark wachsender Industriezweig mit glänzenden Zukunftsaussichten, auch gerade in Schleswig-Holstein.

Aufgrund der relativen Jugend dieses Wirtschaftszweiges sind die Möglichkeiten für technologische Fortschritte enorm, speziell auch im Bereich der Anlagentechnik oder der Optimierung von Arbeitsabläufen. So liegt bei Durchfluss- und Kreislaufsystemen ein großer Arbeitsanteil darin, ein regelmäßiges Monitoring und Vermessen des Bestandes durchzuführen, um Größendiskrepanzen innerhalb eines Zuchtbeckens zu vermeiden. Das liegt darin begründet, dass größere Fische das Wachstum von kleineren Fischen negativ beeinflussen, was entsprechende Auswirkungen auf die Produktivität der Fischfarm hat.

Deswegen muss im Allgemeinen eine Sortierung der Fische anhand des Größenmaßes durchgeführt werden, um so die Varianz der Fischgrößen pro Becken möglichst gering zu halten. Dafür wird jeder Fisch einzeln aus dem Becken genommen und in ein neues Becken mit gleich großen Fischen sortiert. Da die Sortierung heutzutage noch von Hand durchgeführt werden muss, stellt dies einen großen Zeit- und somit Kostenfaktor dar. Des Weiteren stresst dieser Vorgang die Fische immens, was deren Größenzuwachs negativ beeinflussen kann.

Ein sehr wichtiger Aspekt ist hierbei, den besten Zeitpunkt für eine Sortierung herauszufinden. Ist der Zeitpunkt zu früh gewählt, ist der große Aufwand

einer Sortierung von wenig Nutzen geprägt, da die Größendiskrepanz noch zu gering ist. Ist der Zeitpunkt zu spät gewählt, kann sich der Größenunterschied bereits negativ auf das Wachstum der kleineren Fische ausgewirkt haben. Dieses Problem wird hier aufgegriffen und ein System zum automatisierten Monitoring des Wachstums der Steinbutte vorgestellt.

1.2 Fokus der Arbeit

Diese Arbeit behandelt die Entwicklung eines Systemkonzepts zur non-invasiven Vermessung von Steinbutten in Aquakulturanlagen mittels eines Stereokamerasystems unter realen Arbeitsbedingungen. Hierbei existieren 2 Themenschwerpunkte, in die die Arbeit aufgeteilt wurde. Erstens die Entwicklung eines Hardwaresystems zur Aufnahme von Bildmaterial unter den gegebenen Umständen und zweitens die Umsetzung des Konzeptes als Implementierung eines softwarebasierten Verfahrens zur automatischen Erkennung von Steinbutten im Bildmaterial und deren Vermessung basierend auf Stereometrie.

Im ersten Teil der Arbeit (Kapitel 2) wird der aktuelle Wissensstand betrachtet. Grundlegende Informationen über den Steinbutt und dessen Aufzucht in der Aquakultur (Kapitel 2.1), die mathematischen Grundlagen der Stereopsis sowie eine Einführung in die Bildverarbeitung und Mustererkennung (Kapitel 2.2) als auch der aktuelle Stand der Technik in den Bereichen Fischerkennung mittels Bildverarbeitung, Unterwasser-Stereokalibrierung, Stereosysteme sowie konturbasierte Formerkennung (Kapitel 2.3) werden hier thematisiert.

Kapitel 3 beschreibt die Umsetzung eines robusten Stereokamerasystems für den Unterwasser-Einsatz im Arbeitsumfeld einer Aquakulturanlage und den damit verbundenen Anforderungen. Diese beinhalten Eigenschaften wie nicht leitend, spritzwassergeschützt und korrosionsbeständigkeit für alle Komponenten wie Kameragehäuse, Stromversorgung und Beleuchtung. Im Rahmen der Entwicklung des Beleuchtungssystems wird evaluiert, welche Lichtfarbe für die spätere Erkennung von Steinbutten das beste Bildmaterial liefert und welches Arrangement der Lichtquellen eine möglichst homogene Ausleuchtung des Sichtfeldes der Kameras im vorliegenden Arbeitsumfeld bereitstellt. Abschließend erfolgt die Genauigkeitsüberprüfung der durchgeführten Unterwasser-Kalibrierung des Stereokamerasystems.

Das Verfahren zur Erkennung und Vermessung von Steinbutten (Kapitel 4) stellt sich mehreren Herausforderungen. Das Hauptproblem ist die extrem hohe Besatzdichte der Steinbutte, wodurch Vorder- als auch Hintergrund im Bildmaterial nur aus Steinbutten besteht. Andere Verfahren der Fischvermessung (siehe hierzu auch Kapitel 2.3.1) profitieren aufgrund niedriger Besatz-

dichte von einer Abhebung der zu erkennenden Fische vom Hintergrund, sei es Wasser oder eine Beckenwand. Prinzipiell gehen die meisten Verfahren zur Segmentierung und Erkennung von Objekten grundlegend von einer farblichen oder textuellen Differenz zwischen dem Objekt und dem Hintergrund aus, wodurch die vollständige Erfassung einer Objektkontur möglich ist (siehe hierzu auch Kapitel 2.3.4), was in dieser Arbeit nicht der Fall ist.

Erschwerend kommt hinzu, dass aufgrund der Einsatzbedingungen und der Anforderungen lebender Steinbutte die Aufnahmebedingungen nicht optimal für die Bildverarbeitung sind, während im industriellen Einsatz von optischer Vermessung die Bedingungen normalerweise optimal an das Verfahren angepasst werden können.

Das hier vorgestellte Verfahren (Kapitel 4) vergleicht gefundene Kontursegmente mit einem Set zuvor gelernter repräsentativer Konturfragmente. Anschließend werden passende Kontursegmente anhand geometrischer Zusammenhänge gruppiert, um so einzelne Objekte im Bild zu identifizieren. Die Vermessung erfolgt dann über eine Erfassung der Maul- und Flossenbereiche und der Berechnung von deren 3D-Koordinaten.

KAPITEL 2

Wissensstand

Inhalt

2.1 Steinbutt (<i>Scophthalmus maximus</i>)	5
2.1.1 Steinbutt in der Aquakultur	8
2.2 Grundlagen	12
2.2.1 Bildentstehung	12
2.2.2 Bildverarbeitung	21
2.2.3 Mustererkennung	25
2.3 Aktueller Stand der Technik	30
2.3.1 Bildverarbeitung im Kontext der Fischdetektion	30
2.3.2 Unterwasser-Kalibrierung von Stereokamerasystemen	34
2.3.3 Stereosysteme	35
2.3.4 Konturbasierte Formerkennung	37

2.1 Steinbutt (*Scophthalmus maximus*)

Der Steinbutt, mit lateinischem Namen *Scophthalmus maximus*, ist eingeordnet in die Unterklasse der *Strahlenflosser (Actinopterygii)* sowie die Teilklasse der *echten Knochenfische (Teleostei)*. In der Ordnung der *Plattfische (Pleuronectiformes)* gehört er von den 15 vorhandenen Familien zur Familie der *Steinbutte (Scophthalmidae)*. Der Steinbutt selbst gehört zur Gattung *Scophthalmus* und stellt die Art *Scophthalmus maximus*. Der Steinbutt ist auch unter dem Namen *Psetta Maxima* bekannt [48].

Die *Scophthalmidae* sind hierbei eine sehr kleine Familie bestehend aus 4 Gattungen mit insgesamt 9 Arten. Die Größe der Arten variiert dabei von kleinen Exemplaren bis zu Vertretern, die eine Größe von bis zu 100 cm erreichen können. Des Weiteren ist die Familie der Steinbutte charakterisiert durch einen fast runden und asymmetrischen Körper, bei dem sich das Augenpaar auf der linken Seite des Körpers befindet. Das Maul und die Augen sind relativ groß im Vergleich zu anderen Fischarten.

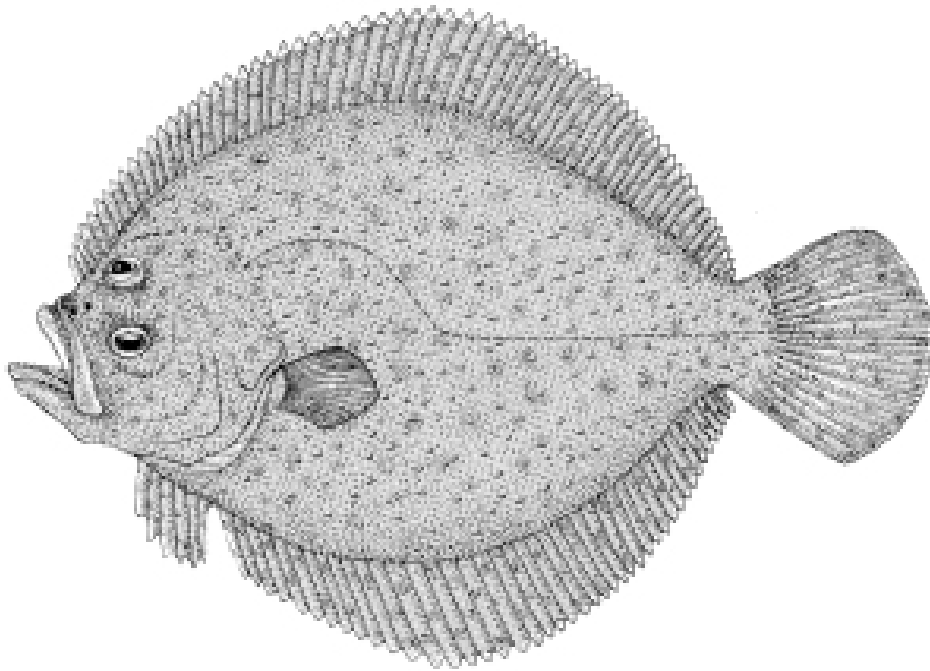


Abbildung 2.1: Schemadarstellung des Steinbutts aus [41].

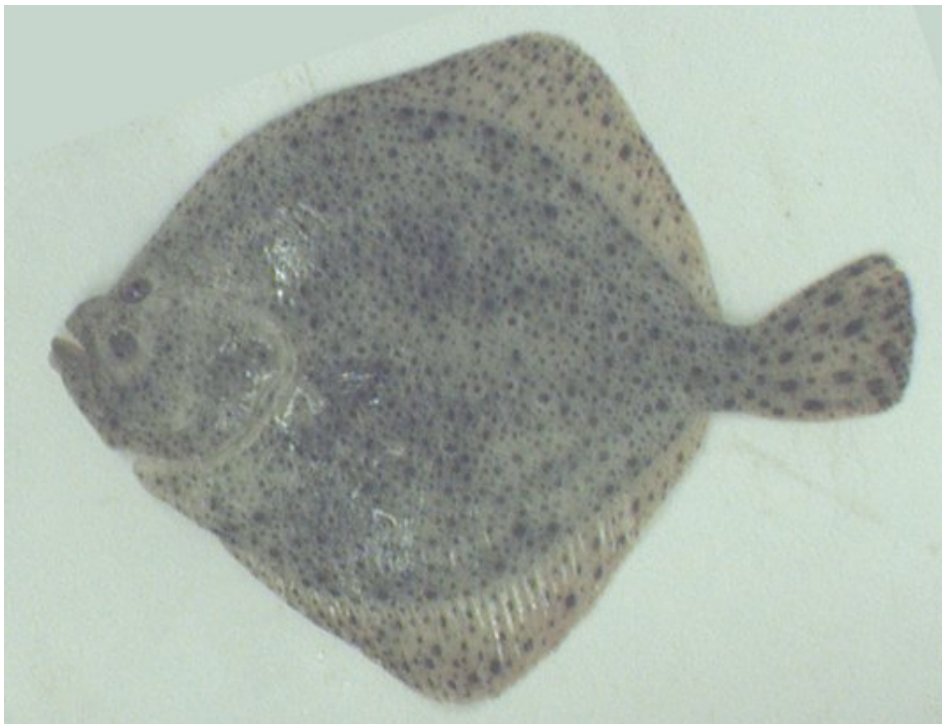


Abbildung 2.2: Steinbutt aus Aquakultur.

Die blinde rechte Körperseite entspricht der Unterseite des Steinbutts und weist eine hell-weißliche Färbung auf. Die linke Seite ist schuppenlos und von variabler Färbung, meistens im Farbbereich grau-braun mit dunklen und hellen Flecken, und dient der Tarnung. Darüber hinaus besitzt der Steinbutt unregelmäßig verteilte Knochenvorsprünge, denen er seinen Namen verdankt. Eine Schemadarstellung sowie ein reales Bild des Steinbutts sind in Abbildung 2.1 und Abbildung 2.2 zu sehen [48] [41] [44]. Der Steinbutt wird normalerweise 50 bis 70 cm lang. Einzelne Exemplare können aber auch Größen von bis zu 100 cm mit einem Maximalgewicht von 25 kg erreichen. Dabei erreichen sie mit die höchste Wachstumsrate unter den Plattfischen mit circa 30 cm alle 3 Jahre. Der älteste bisher gemeldete Steinbutt erreichte ein Alter von 25 Jahren. Bei Fischen besteht ein Zusammenhang zwischen der Größe beziehungsweise Länge und dem Gewicht, welches näherungsweise mit folgender Formel berechnet werden kann [121]:

$$W = a * L^b \quad (2.1)$$

mit Gewicht W in Gramm und Länge L in cm. b steht für einen Exponent, der das Wachstum des Fisches beschreibt. $b \approx 3$ beschreibt ein isometrisches Wachstum, während $b < 3$ ein negatives allometrisches Wachstum und $b > 3$ ein positives allometrisches Wachstum darstellen. Für $b \approx 3$ ist a ein Formfaktor mit den Richtwerten $a = 0.001$ für aalähnliche, $a = 0.01$ für spindelförmige und $a = 0.1$ für kugelförmige Fische. Für den Steinbutt können näherungsweise die Werte $b = 3.1$ und $a = 0.01$ basierend auf mehreren Studien angenommen werden [118] [67] [117] .

Das natürliche Habitat des Steinbutts befindet sich im Nordostatlantik, entlang der Küsten vom mediterranen Raum bis hin zum Polarkreis, sowie der Ostsee und dem Mittelmeer. Er ist ein Raubfisch, der auf sandigem, steinigem oder gemischtem Untergrund in einer Tiefe von bis zu 100 m lebt. Während sich die Jungtiere überwiegend von Krustentieren ernähren, jagen die ausgewachsenen Exemplare hauptsächlich andere in der Nähe des Grundes lebende Fische.

Die Plattfische haben die besondere Fähigkeit, ihre Färbung und Musterung zu verändern, um sich ihrer Umgebung anzupassen und somit einen Camouflage-Effekt zu erzeugen, der ihnen Vorteile bei der Jagd verschafft. Der Effekt wird dabei durch visuelle Stimuli ausgelöst, die der Untergrund verursacht. Durch das Ausdehnen beziehungsweise Zusammenziehen von 3 verschiedenen Arten von *Chromatophoren* (pigmenthaltige Zellen) erfolgt die Veränderung von Muster und Farbe und kann innerhalb weniger Minuten abgeschlossen sein [48].

2.1.1 Steinbutt in der Aquakultur

2.1.1.1 Aquakultur

Die Aquakultur entwickelt sich immer weiter zu einem Grundpfeiler für die Versorgung der Weltbevölkerung mit Fischen und Meeresfrüchten. Während deren Bedarf stetig ansteigt, stagniert die Produktion der Fischereindustrie aufgrund von Überfischung oder vorhandener Fangquoten [41]. Die Aquakultur erwirtschaftet inzwischen ein Drittel der Weltproduktion an Fisch und Schalentieren und ist weiterhin eine schnell wachsende Industrie [39].

Es gibt 4 typische Grundformen der Fischeaufzucht in der Aquakultur. Für die sogenannten *Netzkäfigsysteme* sind die *Offshore*-Anlagen in küstennahen Gewässern, in denen Lachse gezüchtet werden, ein typisches Beispiel. Bei *Teichwirtschaftssystemen* werden die Fische *Onshore* beziehungsweise an Land in überwachten Teichanlagen gehalten, während bei *Kreislauf-* sowie *Durchflusssystemen* Beckenanlagen verwendet werden. Bei Kreislaufsystemen wird das verwendete Wasser aufbereitet, gereinigt und wiederverwendet, während bei Durchflusssystemen das Wasser einer natürlichen Quelle, wie zum Beispiel einem Fluss oder dem Meer, entnommen wird und nach einmaliger Nutzung wieder zurückgeführt wird. Beispiele für die verschiedenen Varianten kann man in den Abbildungen 2.3 sowie 2.4 sehen.



Abbildung 2.3: Links: Netzkäfigsystem mit Lachsen. Rechts: Teichwirtschaftssystem. ¹

Der Steinbutt wird fast ausschließlich in Kreislauf- oder Durchflusssystemen gezüchtet. Dabei werden runde oder rechteckige Becken mit unterschiedlichen Größen von 25 bis 100 m³ verwendet. Die Wassertiefe beträgt hierbei meistens zwischen 50 cm und 70 cm. Eine andere Möglichkeit besteht darin,

¹Bilder mit freundlicher Genehmigung von Mahmoud Gargouri.

²Bilder mit freundlicher Genehmigung von Carsten Schulz, GMA Büsum.



Abbildung 2.4: Kreislaufsystem mit Wasseraufbereitungsanlage. ²

sogenannte Raceways zu verwenden, sehr flache Becken, bei denen die Wassertiefe nur 25 cm beträgt und die in mehreren Bahnen übereinander aufgebaut sind. Im Offshore-Bereich können für den Steinbutt dem Netzkäfigsystem ähnlich flache Käfige mit ebenem Grund ins Wasser gesetzt werden. Diese Methode ist im Vergleich zu Onshore-Anlagen kosteneffektiver, befindet sich aktuell aber noch in der experimentellen Phase und Standorte mit optimalen Bedingungen sind nicht sehr zahlreich, so dass Onshore-Anlagen vorerst Standard bleiben werden [41].

2.1.1.2 Zahlen und Fakten

Plattfische gehören mit zu den beliebtesten und teuersten Fischarten auf dem Markt, was sie für die Aquakultur besonders attraktiv macht. In Europa gehören Steinbutt, Atlantischer Heilbutt und Seezunge mit zu den teuersten Meeresfischen, in Asien ist es vor allem die Japanische Flunder oder auch Hirame genannt, die sich großer Beliebtheit erfreut. Kommerziell erfolgreich kommen von den Plattfischen zurzeit nur der Steinbutt und die Japanische Flunder in der Aquakultur zum Einsatz [48][39].

In Europa wird der Steinbutt in Spanien, Portugal, Frankreich, Deutschland, Dänemark, Island, Norwegen, Wales und Irland in nennenswerten Mengen produziert. Hierbei nimmt Spanien vor Frankreich und Portugal mit Abstand die Führungsrolle ein. In Spanien wurden im Jahr 2002 insgesamt 3.847 t Steinbutt produziert, und somit circa 75 % der damaligen Weltproduktion [41]. Wieviel Wachstumspotenzial in diesem Wirtschaftszweig steckt, kann man gut daran erkennen, dass das spanische Unternehmen Pescanova 2010 im portugiesischen Mira eine der größten Fischzuchtanlagen überhaupt eröffnet hat, die aktuell circa 7.000 t Steinbutt pro Jahr produziert [108], doppelt soviel wie im Jahr 2002 in Spanien insgesamt produziert wurde.

Um 2005 begannen auch Unternehmen in China mit der Produktion von Steinbutt und erreichen inzwischen Produktionsmengen von 50.000 bis 60.000 t pro Jahr, was ungefähr der siebenfachen Menge aller europäischer Produzenten zusammengefaßt entspricht.

Die Entwicklung der Produktionsmenge und das extreme Wachstum mit dem Eintritt Chinas in die Produktion zeigt die Abbildung 2.5.

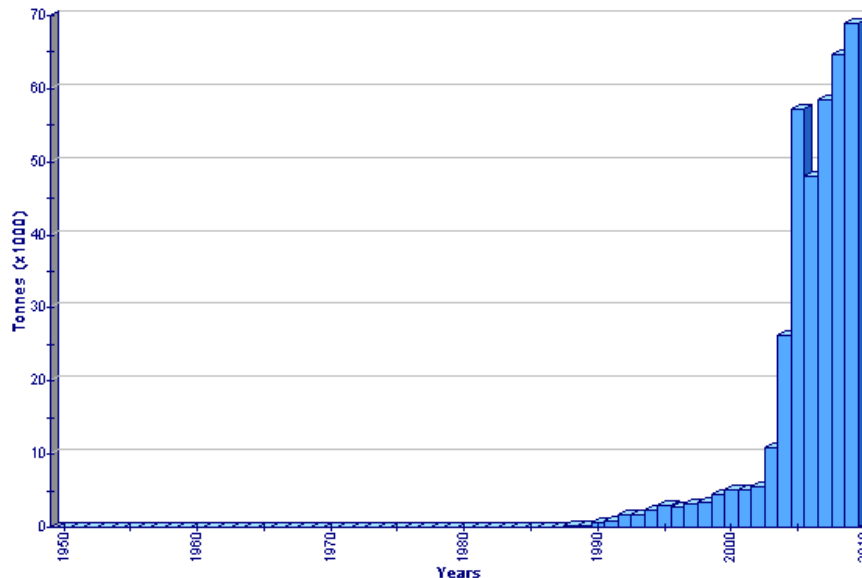


Abbildung 2.5: Entwicklung der Steinbuttproduktion in der Aquakultur. Diagramm aus [41].

2.1.1.3 Produktion von Steinbutt

Die Aufzucht von Steinbutt erfolgt in unterschiedlichen Phasen: der gezielten Reproduktion, der Brut, der Larvenaufzucht, der Juvenilenaufzucht, sowie der eigentlichen Mast. Dabei wird die Vermehrung, Brut und Aufzucht bis zu einer bestimmten Größe oft von darauf spezialisierten Unternehmen, sogenannten Hatcheries, durchgeführt und an die Aquakulturunternehmen verkauft, die für die Mast ausgelegt sind. Abbildung 2.6 zeigt eine schematische Übersicht über den Produktionszyklus des Steinbutts.

Die Elterntiere bestehen in der Regel aus 2 bis 4 Jahre alten Steinbutten, die sowohl aus dem Wildfang als auch aus der Aquakultur stammen, um die genetische Variabilität zu erhöhen. Sie werden in Betontanks mit einem Volumen von 20 bis 40 m³ und einer Besatzdichte von 3 bis 6 kg/m³ gehalten. Die Produktion kann das ganze Jahr über aufrecht gehalten werden.

Dafür werden die Wassertemperatur und die künstlichen Tag-Nacht-Zyklen entsprechend angepasst, um eine Fortpflanzungsperiode bei den Steinbutten auszulösen.

Da der Steinbutt in Gefangenschaft nicht freiwillig laicht, müssen die Gameten von Hand gewonnen werden. Von den Weibchen werden bei einem Laichvorgang 5 bis 10 Millionen pelagische, rund geformte Eier mit einem Durchmesser von circa 1 mm produziert.

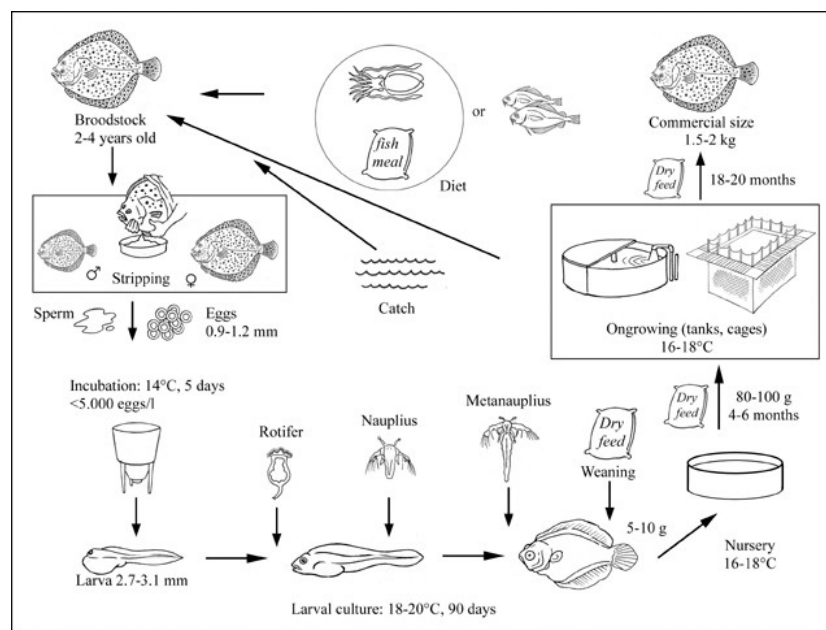


Abbildung 2.6: Produktionskreislauf des Steinbutts von der Brut bis zum verkaufsfertigen Fisch. Illustration aus [41].

Bis zum Schlüpfen der Larven dauert die embryonale Phase 4 bis 6 Tage in Abhängigkeit von der Wassertemperatur. Die Steinbutt-Larven haben dann eine Größe von 2,7 bis 3,1 mm. Die Larven werden entweder intensiv mit einer Besatzdichte von 15 bis 20 Exemplaren je Liter oder semi-intensiv bei einer Besatzdichte von 2 bis 5 Exemplaren je Liter in kleinen Becken gehalten. Dabei handelt es sich um Kreislaufsysteme mit Meerwasser, welche bei einer Wassertemperatur von 18 bis 20 °C gehalten werden. Anfangs ernähren sich die Larven noch von ihrem Dottersack bis sich das Maul nach 3 Tagen öffnet und mit der Fütterung begonnen wird. Hierfür werden *Brachiomus plicatilis*, *Artemia nauplii* und *Phytoplankton* verwendet. Nach 40 bis 50 Tagen erreichen die Larven eine Länge von circa 25 mm und die Metamorphose ist abgeschlossen.

Nach insgesamt 2 bis 3 Monaten und mit einem Gewicht von 5 bis 10 g

wechseln die Juvenile in Aufzuchtbecken. Hierbei handelt es sich ebenfalls um Kreislaufsysteme mit Meerwasser. Hinzu kommen Belüftungssysteme, um den Sauerstoffgehalt des Wassers gesättigt zu halten. Hier werden die Juvenile mit Trockenpellets gefüttert, bis sie eine Größe von 80 bis 100 g haben und in die großen Zuchtbecken umgesetzt werden. Diese Phase des Vormästens dauert 4 bis 6 Monate. Abschließend werden die Steinbutte für die Mast zumeist in Kreislaufsystemen mit Meerwasser gehalten. Die Wassertemperatur wird konstant zwischen 14 und 18 °C gehalten und der Sauerstoffgehalt mittels Sauerstoffzufuhr gesättigt. Als Futter dienen Trockenpellets. Die hohe Toleranz im Bezug auf die Besatzdichte ist aus Sicht der Aquakultur ein Vorteil des Steinbutts, da so Dichten bis zu 100 kg/m² Grundfläche des Beckens erreicht werden können. Im Allgemeinen erreichen die Steinbutte ihr finales Verkaufsgewicht von 1,5 bis 2 kg nach 18 bis 20 Monaten [41] [48].

2.2 Grundlagen

2.2.1 Bildentstehung

Um bei der Arbeit mit Kameras Informationen aus digitalen Bildern zu erhalten, ist ein Verständnis von Kamera- und Optik-Grundlagen unabdingbar. Dieses Kapitel beschreibt die Themen Kameramodelle (Kapitel 2.2.1.1), Projektive Geometrie (Kapitel 2.2.1.2) und Epipolargeometrie (Kapitel 2.2.1.3), die grundlegend sind, wenn Informationen wie die Größe von Objekten und Distanzen zwischen ihnen ermittelt werden sollen.

2.2.1.1 Kameramodell

Möchte man verstehen, wie die Objekte in der Welt auf ein Bild abgebildet werden, wird ein theoretisches Kameramodell benötigt. Das einfachste Modell ist das Lochkameramodell beziehungsweise die sogenannte *Zentralprojektion*. Es ist das allgemeinste Projektionsmodell, liefert aber in den meisten Fällen zufriedenstellende Approximationen des Bildgebungsverfahrens [42].

Das Prinzip des Lochkameramodells ist eine Zentralprojektion von Punkten im euklidischen 3-dimensionalen Raum \mathbb{R}^3 auf eine 2-dimensionale Ebene im euklidischen 2-dimensionalen Raum \mathbb{R}^2 . Eine schematische Darstellung findet sich in Abbildung 2.7. Dafür wird das *Projektionszentrum* beziehungsweise *optische Zentrum* als der Ursprung eines euklidischen Koordinatensystems definiert. Die Ebene Z_f , aufgespannt durch beliebige Koordinaten x und y und einer fixen Koordinate $z = f$, mit f der Brennweite der Kamera, repräsentiert die Bildebene [60]. Ein Punkt $p = (x, y, z)^T$ wird dabei auf den Punkt $p' = (x', y')^T$ der Bildebene abgebildet, an dem sich die Linie von p zu dem

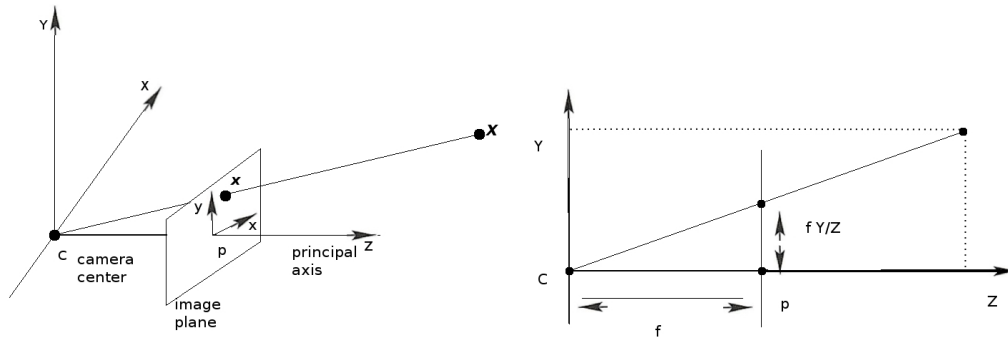


Abbildung 2.7: Schematische Darstellung des Lochkameramodells.

Projektionszentrum mit der Ebene schneidet:

$$(x, y, z)^T \mapsto (fx/z, fy/z)^T. \quad (2.2)$$

Die vom optischen Zentrum ausgehende Linie, die senkrecht zur Bildebene steht und diese am *Bildmittelpunkt* schneidet, ist die sogenannte *Hauptachse*. Andere weit verbreitete Kameramodelle sind die *orthographische Projektion* oder *Parallelprojektion* und die *schwach-perspektivische Projektion*, beide klassifiziert als sogenannte *affine Projektionen*. Bei der orthographischen Projektion sind alle Strahlen parallel zur *Z*-Achse und ein Punkt $p = (x, y, z)^T$ wird auf den Bildpunkt $p' = (x, y)^T$ abgebildet, indem die *Z*-Koordinate ignoriert wird. Bei der schwach-perspektivischen Projektion wird eine Ebene bei $Z = z_0$ definiert. Punkte im Raum werden zuerst orthographisch auf diese Ebene abgebildet, bevor von dieser Ebene ausgehend, die Zentralprojektion angewendet wird. Die Abbildung eines Punktes $p = (x, y, z)^T$ auf den Bildpunkt $p' = (x', y')^T$ erfolgt somit durch

$$\begin{cases} x' = f \frac{x}{z_0} \\ y' = f \frac{y}{z_0} \end{cases}. \quad (2.3)$$

Diese beiden Kameramodelle können dann eingesetzt werden, wenn die abgebildeten Objekte in Relation zu ihrem Abstand zur Kamera sehr klein sind.

2.2.1.2 Projektive Geometrie

Um einen 3-dimensionalen Punkt p in der realen Welt zu beschreiben, wird ein sogenanntes *Welt-Koordinaten-System* (WKS) verwendet. Die Abbildung dieses Punktes auf ein 2-dimensionales Bild, das durch ein *Bild-Koordinaten-System* (BKS) beschrieben wird, erfolgt durch die Anwendung perspektivischer Projektion unter Verwendung von *homogenen Koordinaten*. Bei den ho-

mogenen Koordinaten werden die Vektoren um einen weiteren Wert am Ende erweitert, so dass ein 3-dimensionaler Punkt $p = (x, y, z)^T$ durch einen 4-dimensionalen Vektor $(x, y, z, 1)^T$ repräsentiert wird. Durch diese Änderung wird die nicht-lineare Berechnung der Projektion (Kapitel 2.2.1.1) zu einer linearen Abbildung zwischen WKS und BKS, in Form einer Matrix-Multiplikation:

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fx \\ fy \\ z \end{pmatrix} = \begin{bmatrix} f & & 0 \\ & f & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}. \quad (2.4)$$

Die Abbildung von 4 Dimensionen auf 3 Dimensionen erfolgt mittels einer 3×4 -Matrix, der sogenannten *Projektionsmatrix* P :

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (2.5)$$

Die fehlende Brennweite im Vergleich zu der Matrix aus 2.4 ist Teil der *Kalibriermatrix* K . Außerdem geht das Lochkameramodell davon aus, dass sich der Ursprung der Bildkoordinaten am Hauptpunkt befindet. In Wirklichkeit existiert meistens ein Offset, so dass $(p_x, p_y)^T$ die korrigierten Koordinaten des Hauptpunktes repräsentieren. Im diskreten Fall einer digitalen Kamera können die Pixel des Bildsensors über unterschiedliche Ausmaße in x - und y -Richtung verfügen, weswegen die Skalierungsfaktoren m_x und m_y beachtet werden müssen. Ein weiterer zu beachtender Parameter ist die mögliche Schräglage oder auch *Skew* s der Pixel des Bildsensors. Werden die genannten Parameter beachtet, so erhält man die folgende K -Matrix:

$$K = \begin{bmatrix} fm_x & s & p_x \\ 0 & fm_y & p_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.6)$$

Alle in K enthaltenen Parameter bilden zusammen die sogenannten *intrinsischen Kameraparameter*, die die Eigenschaften einer Kamera beschreiben. Die K -Matrix beschreibt die Abbildung vom BKS auf die diskreten Werte des sogenannten *Pixel-Koordinaten-Systems* (PKS).

Da man nicht davon ausgehen kann, dass sich die Kamera stets im Ursprung des WKS befindet, wird eine weitere Transformation benötigt. Die D -Matrix enthält die sogenannten *extrinsischen Kameraparameter*, die die Orientierung und Position der Kamera im WKS beschreiben:

$$D = \begin{bmatrix} R & t \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (2.7)$$

mit R als 3×3 -Rotationsmatrix und t als 3-dimensionalen Translationsvektor. Die D -Matrix beschreibt die Abbildung eines Punktes im WKS in das *Kamera-Koordinaten-System* (KKS). Um die Abbildung von 3-dimensionalen Punkten im WKS auf 2-dimensionale Pixel-Koordinaten im PKS zusammenzufassen, kann die Gesamtgleichung wie folgt geschrieben werden:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = K * P * D * \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}. \quad (2.8)$$

Um die intrinsischen und extrinsischen Kameraparameter für eine möglichst akkurate Modellierung der Projektion zu erhalten, muss jede einzelne Kamera, die verwendet wird, kalibriert werden. Für diese Aufgabe existiert eine Anzahl weithin akzeptierter Softwarelösungen³. Die angewendeten Verfahren basieren oft auf der Arbeit von Tsai [142] [143], welche als Standardmethode angesehen werden kann. Dabei werden Kalibriermuster wie zum Beispiel Schachbrettmuster mit bekannten Abmessungen verwendet, um anhand der Verzerrungen in den Bildern die intrinsischen und extrinsischen Kameraparameter zu berechnen. Weitere Beispiele für Kalibriermethoden sind [152] und [40].

2.2.1.3 Epipolargeometrie

Die *Epipolargeometrie* beschreibt die intrinsische Geometrie zwischen Kameras und hängt nur von den internen Parametern der Kameras sowie der relativen Pose dazwischen ab. Mit ihrer Hilfe können die Parameter zur Rekonstruktion von 3-dimensionalen Koordinaten im WKS aus korrespondierenden Punkten in verschiedenen Ansichten mehrerer PKS bestimmt werden. Die Abbildung 2.8 illustriert das generelle Konzept der Epipolargeometrie am Beispiel von 2 Kameras. Die 2 Kamerazentren C und C' aus verschiedenen Ansichten von Punkt X sind mit der sogenannten Grundlinie verbunden. Die Epipole e und e' bezeichnen die Punkte, in denen die Grundlinie die Bildebenen durchstößt. Die Grundlinie spannt damit zusammen mit X die Epipolarebene auf. Die Schnittkanten dieser Ebene mit den Bildebenen ergeben die Epipolarlinien l und l' . Auf diesen Linien liegen auch x und x' , welche der Projektion des Punktes X auf die Bildebenen entsprechen. Für beliebige Punkte im WKS wird schnell ersichtlich, dass sie jeweils eine Epipolarebene aufspannen, deren gemeinsamer Schnittpunkt immer der Epipol des zugehörigen Bildes ist (Ebenenbündel). Einer der bedeutendsten Zusammenhänge der Epipolargeometrie

³zum Beispiel OpenCV oder die Camera Calibration Toolbox für Matlab

ist die Möglichkeit, die Suche nach Punktkorrespondenzen in den verschiedenen Ansichten auf die Epipolarlinie einschränken zu können.

Ein weiterer Zusammenhang besteht darin, dass die Reprojektion von gefundenen Punktkorrespondenzen im Bild zu Strahlen im Raum führt, die sich unter idealen Bedingungen treffen müssen, da sie den gleichen Raumpunkt beschreiben.

Mit Hilfe der Triangulation kann hiermit der tatsächliche 3-dimensionale Punkt im WKS gefunden werden. Die Zusammenfassung der mathematischen Bezüge erfolgt im Allgemeinen in der sogenannten E-Matrix (*Essentialmatrix*) oder der F-Matrix (*Fundamentalmatrix*) [60][42].

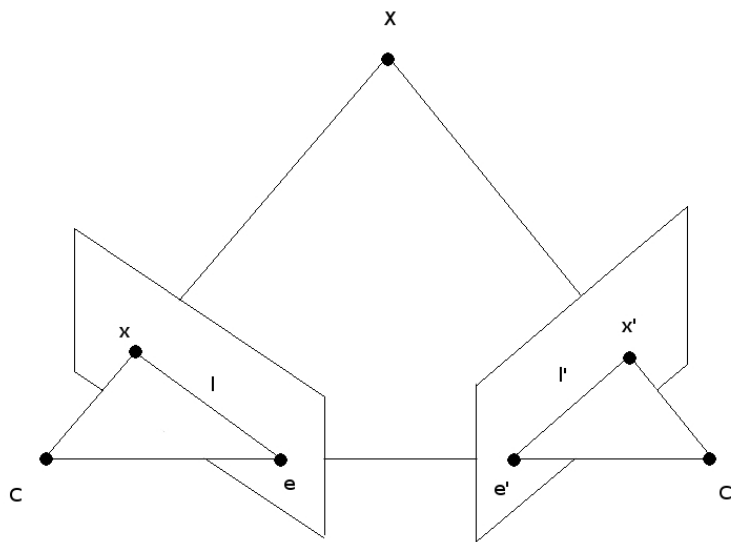


Abbildung 2.8: Schematische Darstellung der Epipolargeometrie.

Die Korrelation zwischen 2 Kameras kann durch eine Rotationsmatrix R und eine Translation t ausgedrückt werden und beschreibt dabei die Abbildung vom KKS der ersten Kamera (im Allgemeinen Master- oder Hauptkamera genannt) zum KKS der zweiten Kamera (Slave- oder Nebenkamera):

$$C = \begin{bmatrix} R & t \\ 0 & 0 & 0 & 1 \end{bmatrix} C'. \quad (2.9)$$

Im Bezug zu den epipolaren Bedingungen, die aussagen, dass die Vektoren zwischen C , C' und X koplanar sind, folgt

$$\mathbf{x} \cdot [t \times (R\mathbf{x}')] = 0, \quad (2.10)$$

mit $\mathbf{x} = (u, v, 1)^T$ und $\mathbf{x}' = (u', v', 1)^T$ als homogene Koordinaten der Bildpunkte x und x' . Das kann wie folgt formuliert werden:

$$\mathbf{x}^T E \mathbf{x}' = 0 \quad (2.11)$$

mit

$$E = R \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}. \quad (2.12)$$

Die E-Matrix besitzt 5 Freiheitsgrade, ist damit bis auf einen Skalierungsfaktor definiert und bildet den Punkt \mathbf{x}' auf seine korrespondierende Epipolarlinie l ab mit:

$$l = E \mathbf{x}'. \quad (2.13)$$

Dies gilt auch für die Partnerkamera, da das *Dualitätsprinzip* der projektiven Geometrie aussagt, dass der Austausch der Begriffe Punkt und Linie in den Axiomen wieder zu einer wahren Aussage führt.

Die E-Matrix setzt kalibrierte Kameras voraus und berücksichtigt damit nicht die intrinsischen Parameter der Kameras. Für den Fall, dass die intrinsischen Parameter nicht bekannt sind, kann mit K und K' als Kalibriermatrizen und $\hat{\mathbf{x}}$ und $\hat{\mathbf{x}}'$ als normalisierte Bildkoordinaten die F-Matrix aufgestellt werden. Hier gilt

$$\mathbf{x} = K \hat{\mathbf{x}} \quad (2.14)$$

und

$$\mathbf{x}' = K' \hat{\mathbf{x}}', \quad (2.15)$$

was zu

$$F = K^{-1T} E K'^{-1}, \quad (2.16)$$

mit F als Fundamentalmatrix führt. Auch die F-Matrix beschreibt die algebraische Repräsentation der Epipolargeometrie durch die Beziehungen zwischen den Kameras. Mit einem Rang von 2 und 7 Freiheitsgraden gilt für die F-Matrix, dass für jeden Punkt \mathbf{x} die korrespondierende Epipolarlinie $l' = F \mathbf{x}$, beziehungsweise $l = F^T \mathbf{x}'$ für die entgegengesetzte Richtung, definiert werden kann. Die F-Matrix hilft dabei, die Suche nach den Korrespondenzen \mathbf{x} und \mathbf{x}' auf eine einzige Linie im Bild zu reduzieren, anstatt das gesamte Bild testen zu müssen. Da die F-Matrix 7 Freiheitsgrade besitzt, reichen 7 Punktkorrespondenzen aus, um die F-Matrix mit Hilfe von intrinsischen Parametern zu berechnen [60][42].

2.2.1.4 Stereopsis

Die Prinzipien der Epipolargeometrie wurden im letzten Abschnitt erläutert und werden nun in der Anwendung eines *Stereosystems* vorgestellt, das heißt für diesen Kontext, ein auf einer Halterung installiertes Set von 2 Kameras, dem sogenannten Rig. Dies führt zur Anwendung der *Stereopsis*, welche die Möglichkeit beschreibt, eine komplette Tiefenkarte aus der Ansicht der Kameras zu erstellen. Solche Systeme haben eine breite Anwendung in der Roboternavigation, der Kartographie oder der Photogrammetrie gefunden. Da es sich um 2 verschiedene Perspektiven der Kameras handelt, erscheinen die Objekte im Bild je nach Entfernung zur Kamera an verschiedenen Positionen im Bild. Diese Beziehung wird durch die bestehende *Parallaxe* ausgelöst und mit dem Maß der *Disparität* beschrieben. Mit Hilfe der intrinsischen Parameter der Kameras kann das Maß der Disparität eine Aussage über die Entfernung, also die Tiefe von Objekten im Raum, geben. Die Abbildung 2.9 zeigt ein Beispiel hierzu. Die Berechnung für eine 3-dimensionale Rekonstruktion von Raumpunkten kann signifikant vereinfacht werden, wenn die Bilder in der Vorverarbeitung *rektifiziert* werden [91][71].

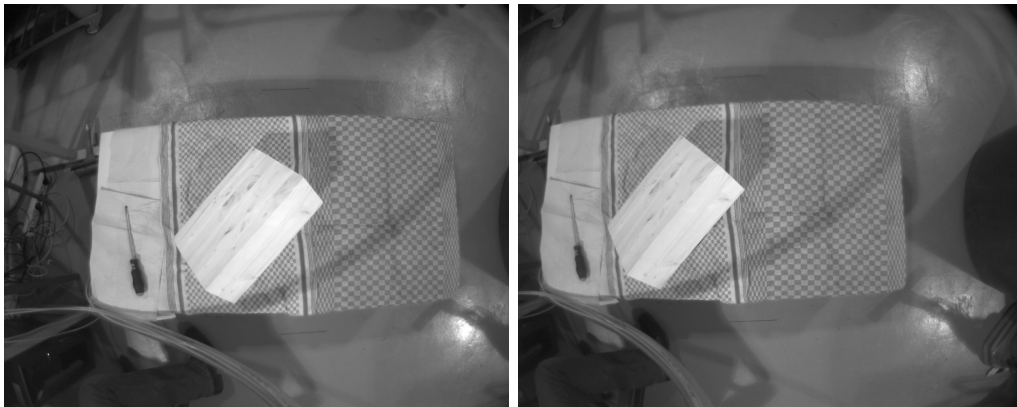


Abbildung 2.9: Stereobildpaar, bei dem der Holzquader aufgrund der unterschiedlichen Perspektiven an unterschiedlichen Positionen erscheint.

2.2.1.4.1 Rektifizierung Die Rektifizierung von Stereobildern ist eine Überführung der beiden Ansichten in die sogenannte *epipolare Standard-Konfiguration*. Damit ist eine rotationsfreie, ideale Anordnung der Kameras parallel zueinander beschrieben. Dies kann bewerkstelligt werden, indem die Bilder auf eine neue Bildebene projiziert werden. Danach ergeben sich neue Bedingungen für die Ansichten. Sie haben eine gemeinsame Bildebene, die parallel zur Grundlinie liegt und die optischen Zentren miteinander verbindet. Daraus

folgend sind die optischen Achsen ebenso parallel. Des Weiteren ist die Distanz vom optischen Zentrum zur Bildebene identisch. Die extrinsischen Parameter sind justiert, so dass die einzige Transformation zwischen den optischen Zentren eine Translation in x -Richtung ist. Ein Beispiel gibt die Abbildung 2.10. Die Voraussetzung ist eine Kalibrierung der Ausrichtung der Kameras zueinander. Der Vorteil daran ist, dass die Epipolarlinien nicht länger verzerrte

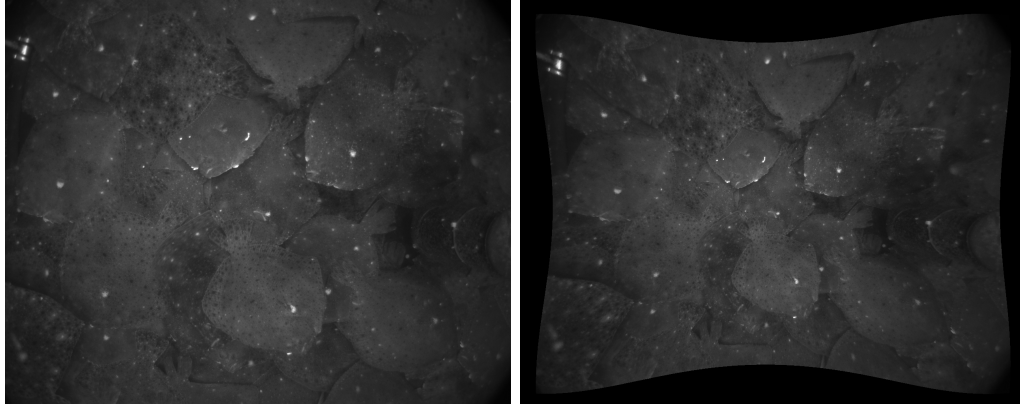


Abbildung 2.10: Beispiel für die Rektifizierung. Links: Originalbild. Rechts: Rektifiziertes Bild.

Kurven im korrespondierenden Bild sind, sondern nun parallel zu den Pixelreihen verlaufen, das heißt mit einer konstanten y -Koordinate, was die Suche extrem vereinfacht. Des Weiteren wird auch die Berechnung der Disparität erheblich vereinfacht [42].

2.2.1.4.2 Disparität Unter der Voraussetzung von rektifizierten Bildpaaren kann die Disparität d von der x -Koordinate der Punkte $p = (x, y)^T$ und $p' = (x', y)^T$ des linken und rechten Abbildes des Punktes P mit $d = x' - x$ berechnet werden. Mit homogenen Koordinaten ergibt sich die Tiefe von P mit $z = -b/d$ mit b als Länge der Grundlinie. Der Koordinatenvektor von P bezüglich des Koordinatensystems der linken Kamera ist dann

$$\mathbf{P} = -\left(\frac{b}{d}\right)\mathbf{p} \quad (2.17)$$

mit $\mathbf{p} = (x, y, 1)^T$ als normalisierte Bildkoordinaten von p . Für die Illustration der Abhängigkeiten siehe Abbildung 2.11. Eine weitere Methode für die Berechnung der Koordinaten von P ist eine direkte Triangulation des Schnittpunktes der Strahlen $R = Cp$ und $R' = C'p'$, wobei C und C' die Kamerazentren sind. Aus verschiedenen Gründen, z.B. Kalibrierungenauigkeiten oder -fehler, sind die Umstände nicht ideal und die Strahlen werden

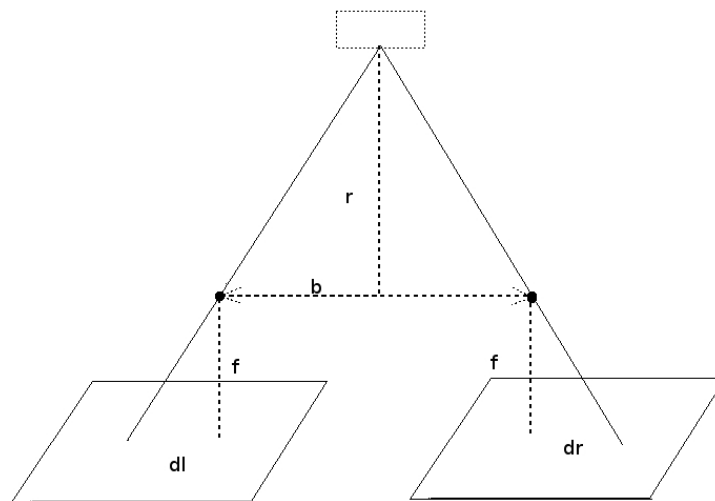


Abbildung 2.11: Schematische Darstellung der Disparität, mit dl , dr als linke und rechte Disparität, b als Länge der Grundlinie, f als Brennweite und r als Distanz zum Objekt.

sich höchstwahrscheinlich nicht schneiden. Um dieses Problem zu lösen, wurde eine Vielzahl an geometrischen (zum Beispiel die Berechnung eines Lots zu beiden Strahlen) sowie numerischen Schätzverfahren entwickelt [91] [42].

2.2.1.5 Robuste Schätzer

Bei der Rekonstruktion von 3-dimensionalen Modellen aus Bildpaaren oder der dafür benötigten Kalibrierung des Stereosystems wird eine große Anzahl an korrespondierenden Bildpunkten verwendet. Ein großes Problem sind hierbei Ausreißer, da diese das Ergebnis stark verfälschen können und durch Messungenauigkeiten oder fehlerhaftes Zuweisen von Korrespondenzen entstehen. Um dieses Problem zu minimieren, werden sogenannte *robuste Schätzer* eingesetzt, das sind Verfahren, die sich durch ihre Unempfindlichkeit gegen Ausreißer auszeichnen. Typische Verfahren sind hier die *M-Schätzer*, eine Gruppe von Schätzverfahren, die basierend auf der *Maximum-Likelihood*-Methode versuchen, eine bestimmte Fehlerfunktion zu minimieren. Beispiele hierfür sind die *Methode der kleinsten Quadrate* oder der *Hubert-k-Schätzer*.

Eine andere Variante, die Ausreißer zu behandeln, stellt der *RANSAC-Algorithmus* dar. RANSAC steht für *RAN*dom *SAM*ple *CON*sensus, also übereinstimmende zufällige Stichprobe. Ziel des Verfahrens ist es, aus einem Daten-

satz D mit vielen Messwerten, die Ausreißer zu ermitteln und auszuschließen, um für die Modellberechnung ein sogenanntes *Consensus Set* zu verwenden. Dabei hat sich der RANSAC-Algorithmus als ein sehr robustes Verfahren herausgestellt und ist heutzutage Standard. Er soll hier stellvertretend genauer vorgestellt werden.

Der Ablauf des Algorithmus ist wie folgt [106]:

1. Zufällige Auswahl von n Messwerten d aus D , n sei hierbei die minimale Menge an Messdaten, die benötigt wird, um das Modell zu berechnen.
2. Berechnen des Modells aus dieser Teilmenge.
3. Für alle anderen Messwerte ist die Distanz zum Modell zu ermitteln. Ist diese unterhalb eines Schwellenwertes t , wird der Messwert in den Consensus-Datensatz D_i übernommen.
4. Erreicht D_i eine Mindestgröße T , wird angenommen, dass der Consensus-Datensatz ein gutes Modell erzeugt und gespeichert.
5. Wiederholen der Schritte 1 bis 4 für x Iterationen. Der größte Consensus-Datensatz D_i wird gewählt und das Modell wird mit allen Messwerten daraus erneut berechnet.

Entscheidend sind die Parameter t für die Distanz zum Modell, der Schwellwert T für die Größe des Consensus-Datensatzes sowie die Anzahl der Iterationen x . Der Schwellwert t für die Distanz zum Modell ist entscheidend für den Erfolg des Verfahrens. Hierbei gibt es verschiedene Verfahren zur Ermittlung eines guten Schwellwertes [60]. T kann aber auch genutzt werden, um den Algorithmus vorzeitig abzubrechen, wenn ein Consensus-Datensatz T überschreitet und somit ein gutes Ergebnis erwartet wird. Die empfohlene Anzahl der Iterationen x ist abhängig von der benötigten Größe der Teilmenge und der gewünschten Wahrscheinlichkeit p , dass mindestens eine Teilmenge keine Ausreißer beinhaltet, in Abhängigkeit der zu erwartenden Ausreißer. Ein Beispiel aus [106] ist in der Tabelle 2.1 dargestellt.

2.2.2 Bildverarbeitung

Mit Hilfe der Bildverarbeitung ist es möglich, die Eigenschaften von Bildern zu verändern. Zumeist dient dies der Verbesserung oder der Vorverarbeitung des Bildmaterials, um somit spätere Verfahren, wie zum Beispiel die Erkennung und Segmentierung sowie Klassifizierung von Objekten zu ermöglichen. So ist es möglich, die Helligkeit oder den Kontrast des Bildes zu optimieren, Rauschen zu unterdrücken, das Bild weicher zu zeichnen, zu schärfen oder

Tabelle 2.1: Anzahl x der Iterationen, die benötigt werden, damit mit einer Wahrscheinlichkeit von $p = 0,99$ eine Teilmenge aus s Elementen und entsprechendem Anteil von Ausreißern mindestens einmal keine Ausreißer enthält.

Größe Teilmenge s	Menge der Ausreißer			
	5%	20%	30%	50%
3	3	7	11	35
5	4	12	26	146
7	4	20	54	588

Merkmale wie Ecken oder Kanten zu verstärken. Dabei werden stets die vorhandenen Grau- beziehungsweise Farbwerte des Bildes zur Berechnung der Ergebnisse herangezogen. Hier sollen die grundlegenden Verfahren der Bildanalyse kurz erläutert werden, wobei aufgrund des großen Themenfeldes die Auswahl auf solche Varianten beschränkt wird, die in dieser Arbeit Anwendung finden.

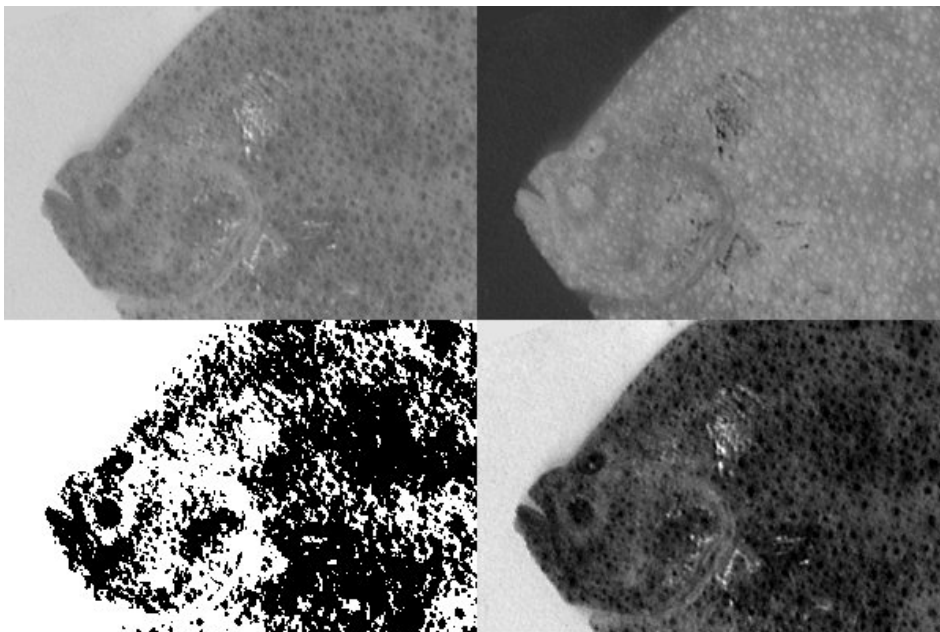


Abbildung 2.12: Beispiele für verschiedene Punktoperatoren. Oben links: Original. Oben rechts: Invertierung. Unten links: Binarisierung. Unten rechts: Histogrammqualisierung.

Punktoperatoren verwenden für die Berechnung des neuen Pixelwertes nur den aktuell betrachteten Pixel. Dabei können globale Parameter, die zum

Beispiel auf dem Histogramm des Bildes basieren, verwendet werden. So kann durch gleichmäßiges Verändern der Grauwerte das Bild aufgehellt oder abgedunkelt werden. Durch das einfache Strecken des Histogramms kann der Kontrast eines Bildes verstärkt werden. Eine weitere häufig genutzte Methode zur Bildverbesserung ist die Histogrammäqualisierung, bei der die Grauwerte aller Pixel gleichmäßig auf das vorhandene Spektrum verteilt werden.

Weitere Möglichkeiten sind die Invertierung des Bildes sowie die Binarisierung durch einen Grenzwert, der auf unterschiedlichste Art ermittelt werden kann [125]. Für Beispiele siehe Abbildung 2.12.

Eine sehr häufig eingesetzte Methode der Bildanalyse ist der *lokale Operator*, bei dem zur Berechnung des neuen Pixelwertes eine Menge benachbarter Pixel, die sogenannte Filtermaske, herangezogen wird. Eine Variante ist die *diskrete Faltung*, bei der anhand einer diskreten Funktion der neue Wert berechnet wird. Je nach Berechnung werden die Filter in linear und nichtlinear unterschieden und es ergeben sich unterschiedliche Filterergebnisse. *Mittelwert-, Gauß- oder Medianfilter* sind Glättungsfilter und dienen zur Rauschunterdrückung und Kantenglättung. Ein *nichtlinearer Gaußfilter* [146] ist ein Beispiel für einen Glättungsfilter, der in der Lage ist, Rauschen zu unterdrücken und die Weichzeichnung von Kanten zu minimieren. Mittels *Sobel- und Laplacefiltern* können Kanten verstärkt und detektiert werden [71] [149]. Der *Canny Edge Detector* [24] ist ein oft eingesetzter Filter zur Kantenerkennung in Graubildern. Beispiele für verschiedene Filter sind in der Abbildung 2.13 dargestellt.

Eine besondere Form der lokalen Operatoren sind die *morphologischen Operatoren*, die auf Binärbildern angewendet werden. Hier wird eine binäre *Maske*, auch *Strukturelement* genannt, verwendet. Bei der *Erosion* bleiben nur die Pixel aus dem Ursprungsbild weiß, bei denen das Strukturelement komplett von weißen Pixeln belegt ist, während bei der *Dilatation* alle Pixel weiß werden, bei denen das Strukturelement mindestens einen weißen Pixel enthält. Bei der *Öffnung* wird zuerst eine Erosion und anschließend eine Dilatation durchgeführt, was zum Aufbrechen schwacher Verbindungen und dem Löschen kleiner Strukturen führt. Bei der *Schließung* erfolgt zuerst die Dilatation mit anschließender Erosion, was zum Schließen kleiner Lücken und dem Verstärken schwacher Strukturen führt [91] [125].

Zuletzt sollen noch die *globalen Operatoren* genannt werden, bei denen sich der neue Pixelwert aus der Gesamtheit des Bildes berechnet. Eines der gängigsten Verfahren ist die Bildglättung oder -schärfung im Frequenzraum. Dazu wird mittels der *Fourier-Transformation* das Bild in den Frequenzraum abgebildet und entsprechende hohe oder niedrige Frequenzen entfernt beziehungsweise bearbeitet. Abschließend erfolgt eine inverse Fouriertransformation, die das bearbeitete Frequenzbild zurück in seine Ausgangsform bringt

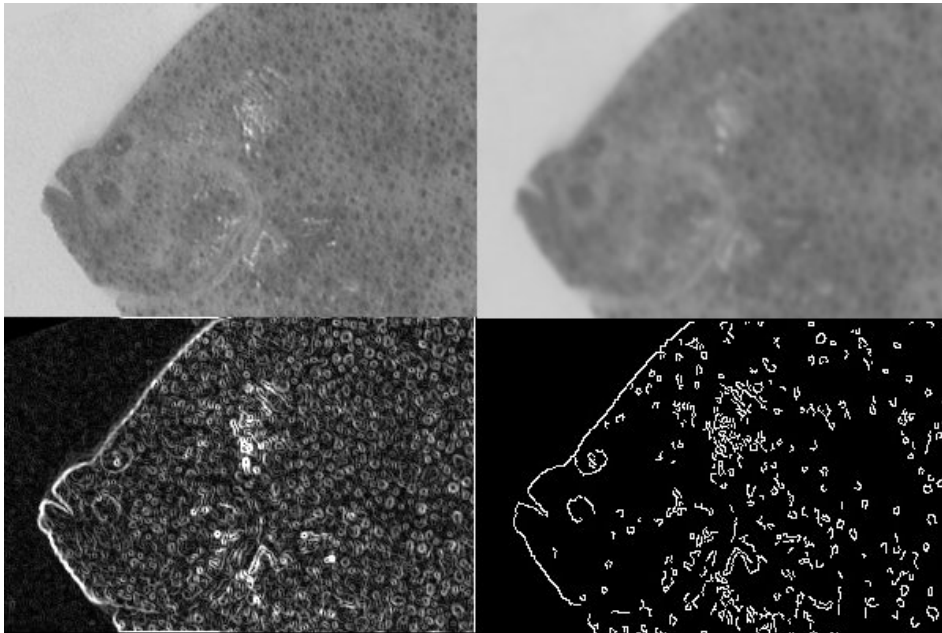


Abbildung 2.13: Beispiele für verschiedene lokale Operatoren. Oben links: Original. Oben rechts: Gaußfilter. Unten links: Sobelfilter. Unten rechts: Canny Edge Detector.

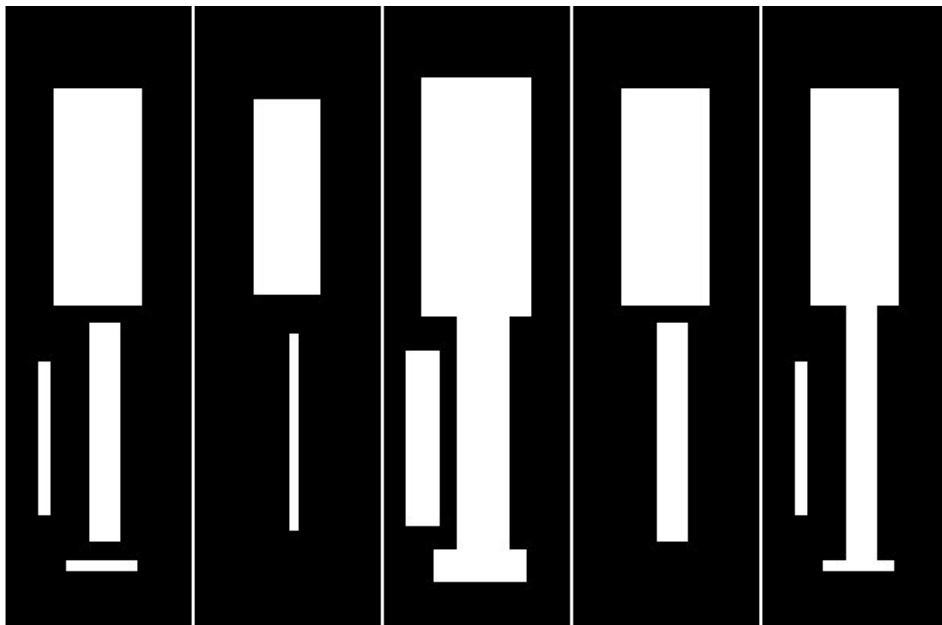


Abbildung 2.14: Beispiele für verschiedene morphologische Operatoren unter Verwendung einer quadratischen Maske. Von links nach rechts: Original, Erosion, Dilatation, Öffnung, Schließung.

[71] [149].

Für eine tiefergehende Behandlung der Möglichkeiten und der Vielzahl an Verfahren sei auf die entsprechende Literatur verwiesen [42] [125] [71] [91] [111] [114].

2.2.3 Mustererkennung

Mustererkennung ist das wissenschaftliche Fachgebiet, welches sich mit der Frage der Objektbeschreibung anhand von Eigenschaften und deren Klassifizierung in unterschiedliche Gruppen beschäftigt. Dabei kann es sich um verschiedenste Arten von Daten handeln, die ausgewertet werden, um bestimmte Muster zu erkennen. Einige Hauptthemen mit Beipielverfahren für die verschiedenartigen Einsatzgebiete der Mustererkennung sind nach [92]:

- Spracherkennung [49],
- Verkehrsanalyse und Ampelsteuerung [76],
- Genanalysen [9],
- Textanalyse zum Beispiel auf Webseiten [94],
- Analyse von Nah-Infrarot-Spektroskopie [85] sowie
- Gesichtserkennung [84].

Ein Verfahren zur Mustererkennung besteht grundlegend aus den 3 folgenden Phasen [37] [87]:

1. *Vorverarbeitung*. Hier erfolgt die Bearbeitung der Eingangsdaten, so dass sie für die anschließenden Schritte verwertbar sind. Dementsprechend muss der erste Schritt an die Anforderungen des zweiten Schrittes angepasst sein. Bei dieser Arbeit handelt es sich bei den Eingangsdaten um Bilder. Typische Arbeitsschritte der Vorverarbeitung von Bilddaten sind Rauschminimierung, Detektion von Kanten oder anderen auffälligen Strukturen und die Segmentierung von ähnlichen Bildbereichen basierend auf Farbe oder Struktur, wie teilweise in Kapitel 2.2.2 beschrieben.
2. *Extraktion von Features*. Als Feature wird ein Parameter bezeichnet, der aus den vorverarbeiteten Informationen gewonnen wird, um diese anschließend zu bewerten. Alle verwendeten Features zusammen bilden den sogenannten *Feature-Vektor*. Die Auswahl der passenden Features erfolgt unter den Anforderungen der späteren Klassifikation und sollte

möglichst zielführend sein. Allein im Bereich der Mustererkennung auf Bildmaterial ist die Auswahl an möglichen Features und deren Berechnung extrem vielseitig. Siehe hierzu auch [96]. Auf die für diese Arbeit wichtige Gruppe der Konturen und deren Erkennung wird ausführlicher in Kapitel 2.2.3.2 eingegangen.

3. *Klassifikation*. Bei der Klassifikation wird der zuvor gesammelte Feature-Vektor anhand des gewählten Klassifikations-Verfahrens bewertet und innerhalb der vorhandenen Objektklassen zugeordnet. Hierbei ist die Auswahl an möglichen Verfahren sehr vielschichtig. Ein kurzer Überblick wird in Kapitel 2.2.3.1 gegeben.

Der Übergang zwischen Extraktion von Features und Klassifikation ist dabei eher fließend. Nach der Klassifikation folgt dann das sogenannte *Post-Processing*, bei dem die durch die Klassifikation getroffenen Zuteilungen in die vorhandenen Klassen weiter verwendet werden. In dieser Arbeit entspricht dies der Vermessung des Fisches.

2.2.3.1 Klassifikation

Für die bei der Mustererkennung durchgeführte Klassifikation existieren sehr viele unterschiedliche Ansätze und Methoden. In diesem Kapitel soll ein kurzer Überblick der gängigsten der sogenannten *Klassifikatoren*, gegeben werden, wobei eine eindeutige Kategorisierung aufgrund vieler Überschneidungen schwierig ist.

Eine grundlegende Unterteilung der verschiedenen Methoden ist die Unterscheidung zwischen *überwachten* und *unüberwachten* Verfahren. Bei den überwachten Verfahren werden Informationen über die Klassen zum Beispiel aus Trainingsdaten verwendet, um damit verschiedene Klassen zu definieren. Bei unüberwachten Verfahren sind solche Vorgaben nicht gegeben und die Klassifikation hat das Ziel, selbstständig eine Unterteilung in Klassen vorzunehmen, basierend auf den Feature-Vektoren und dem gewählten Verfahren [15].

Außerdem kann man die Klassifikatoren nach ihren Hauptcharakteristiken unterscheiden, wobei hier überwachte und unüberwachte Verfahren einem Charakteristikum zugeordnet werden können [92].

Daten-Clustering verfolgt das Ziel, ohne Vorwissen über Klassen, Objekte in sich ähnelnde Gruppen zu unterteilen. Damit gehört es zu den unüberwachten Verfahren. Dies erfolgt beim Clustering über entsprechend zu wählende Verfahren zur Ähnlichkeitsberechnung.

Beispiele für typische Vertreter sind das *k-Means*-, *Agglomerative*- und *Nearest-Neighbour-Clustering* [37] [138]. *Support Vector Machines (SVM)* gehören ebenfalls zu den Clustering-Verfahren. Dabei teilen sie die Klassen derart auf, dass sich in einem möglichst großen Bereich zwischen den Klassen keine Objekte befinden, auch *Maximum Margin* genannt. Dazu erzeugt die SVM zwischen den Objekten im Vektorraum eine Hyperebene. Sind die Objekte nicht durch eine Ebene trennbar, so wird der Vektorraum um so viele Dimensionen erhöht, bis dies möglich ist. Hierfür werden *Kernelfunktionen* benötigt, um diesen Vorgang berechenbar zu halten [15] [35] [93]. Da es beim Clustering Verfahren gibt, die auf die Berechnung von Wahrscheinlichkeiten zurückgreifen, insbesondere auch bei den SVMs, entsteht hier eine Überschneidung und somit auch eine mögliche Einordnung in die nächste Gruppe.

Statistische Mustererkennung ist eine der klassischen und am häufigsten eingesetzten Varianten der Mustererkennung. Sie basiert auf der Verwendung von Wahrscheinlichkeitsmodellen, um anhand der Verteilung der Feature-Vektoren abzuschätzen, in welche Klasse ein Objekt gehört. Dafür wird ein Trainingsset benötigt, bei dem eine Zuordnung der Objekte zu den unterschiedlichen Klassen erfolgt ist, weswegen die statistische Mustererkennung den überwachten Verfahren zuzuordnen ist [92].

Eine Untergruppe sind die *Fuzzy-Sets*, die darauf basieren, dass eine exakte Einordnung eines Objektes nicht immer möglich ist. So ermöglicht ein Fuzzy-Set, dass ein Objekt mehreren Objektklassen angehören kann, wenn eine klare Trennung nicht möglich ist, wodurch es zu Überschneidungen kommt [28] [55].

Weitere Beispiele für statistische Mustererkennung sind der *Bayes-Klassifikator*, *Maximum-Likelihood* oder die Verwendung von Verfahren, die auf Bäumen basieren [145].

Strukturelle Mustererkennung wird eingesetzt, wenn die verwendeten Features in Relation zueinander stehen und sich ihre Auswertung durch das Ergebnis eines anderen Features beeinflussen lässt. Hierfür werden oft unterschiedliche statistische Klassifikatoren in Netzwerken miteinander verbunden. Durch diese Kombination ist es möglich, komplexe Expertensysteme zu schaffen, wie zum Beispiel zur Diagnoseunterstützung von Ärzten. Ein typischer Vertreter für die strukturelle Mustererkennung sind die *Bayes-Netzwerke* [70] [34]. Eine spezielle Untergruppe der strukturellen Mustererkennung ist die *Syntaktische Mustererkennung*. Sie konzentriert sich auf die zugrunde liegende Struktur der zu klassifizierenden Objekte und wie sie aus den verschiedenen Features zusammengesetzt sind. So entstehen Grammatiken aus Regeln, anhand derer die Zuordnung zu einer Klasse erfolgt. Aufgrund der benötigten Grammatik werden diese Klassifikatoren zumeist bei möglichst präzise defi-

nierbaren Problemstellungen eingesetzt [46].

Neuronale Netzwerke basieren auf dem Prinzip der Funktionsweise des menschlichen Gehirns. Dabei wird ein Netzwerk aus *Neuronen* erstellt, vergleichbar mit Knoten in anderen Netzwerken. In jedem Neuron werden Entscheidungen getroffen und entsprechende Signale an die verbundenen Neuronen weiter geleitet. Mit Hilfe von Gewichtungen kann die Auswirkung eines Signals auf ein Neuron beeinflusst werden. Der Aufbau eines Neuronalen Netzes muss entsprechend der vorgesehenen Aufgaben entwickelt werden und kann je nach Einsatz ein überwachtes oder unüberwachtes Verfahren darstellen. Die Neuronalen Netze haben sich inzwischen zu einem eigenen sehr großen Forschungszweig entwickelt und für weitere Informationen sei hier auf die Literatur verwiesen [105] [14] [113].

Für jede hier erwähnte Art der Klassifikatoren existiert eine sehr große Anzahl an verschiedenen Verfahren mit unterschiedlichen Herangehensweisen und Möglichkeiten an Ergebnissen. Die Auswahl sollte stets anhand der vorliegenden Aufgabe, der vorhandenen Daten und der gewünschten Art an Ergebnissen ausgewählt werden.

2.2.3.2 Formerkennung

Bei der Mustererkennung in Bilddaten ist zur Beschreibung von Objekten die Analyse von deren Form eine sehr weit verbreitete Vorgehensweise. Dabei unterscheidet man Verfahren zur Beschreibung des Objekts in solche, die nur auf die Objekt-Umrisse zurückgreifen, die *konturbasierten Verfahren*, und solche, die die gesamte Objekt-Fläche einbeziehen, die *regionbasierten Verfahren*. Jede dieser Gruppen kann erneut in *globale* und *strukturelle* Verfahren unterteilt werden [151].

Konturbasierte globale Verfahren verwenden die Kontur als Ganzes und erzeugen je nach Methode einen für das Objekt repräsentativen Feature-Vektor. Die Ähnlichkeit zwischen 2 Objekten wird dann zumeist mittels eines Distanzmaßes über die Feature-Vektoren berechnet. Zur Bestimmung des Feature-Vektors stehen eine Vielzahl an verschiedenen Verfahren zur Auswahl. So gibt es simple Ansätze, die auf grundlegenden geometrischen Eigenschaften der Kontur basieren wie Fläche, Kreisförmigkeit oder Orientierung der Hauptachse. Diese Verfahren beschreiben Objekte nur sehr grob und werden eher in Kombination mit anderen Verfahren genutzt, um den Feature-Vektor zu erweitern. Komplexere und aussagekräftigere Verfahren sind zum Beispiel:

- *Korrespondenz-Verfahren* basieren auf dem direkten Vergleich von Pixeln, um so einen Feature-Vektor zu erzeugen.

- *Kontur-Signaturen*, die anhand von Kontur-Eigenschaften, wie dem Abstand zum Mittelpunkt, einen Feature-Vektor berechnen [36].
- Das *Elastische Matching* versucht ein Template an die Kontur anzupassen. Die benötigten Verformungen und weitere Parameter dienen dann zur Konturbeschreibung [13].
- *Stochastische Methoden* wenden zur Beschreibung der Konturen stochastische Modelle auf Feature-Vektoren an, die mittels Kontur-Signaturen gewonnen wurden [75].
- *Spektrale Transformationen* verwenden mit Kontur-Signaturen gewonnene Feature-Vektoren, um diese im Spektralraum weiter auszuwerten. Die Verwendung von *Fourier-Deskriptoren* [27] gehört zu den am häufigsten genutzten Verfahren der Formerkennung.

Konturbasierte strukturelle Verfahren zerlegen die Kontur in *Primitive* und organisieren diese zu einer Kontur-Beschreibung. Die Verfahren unterscheiden sich darin, wie die Primitive erzeugt werden und die Art der Organisation zur Objektbeschreibung. Beispiele hierfür sind:

- *Chain-Codes* verwenden Konturabschnitte mit Einheitslänge und die generalisierte Orientierung über den Konturabschnitt. Die Elemente werden wie an einer Kette aneinander gereiht [43].
- Die *Polygon-Zerlegung* beschreibt die Kontur anhand von Polygonen, von denen nur die Auffälligsten verwendet werden. Die Organisation der Polygone erfolgt zum Beispiel in Baumstrukturen [56].
- *Smooth-Curve-Zerlegung* glättet die Kontur mittels Gauß-Verfahren und unterteilt sie in Kurven-Elemente, von denen die stärkste Krümmung und deren Orientierung zur Beschreibung dient. Zur Organisation werden erneut häufig Baumstrukturen verwendet [12].
- Die *Syntaktische Analyse* verwendet einen Satz nummerierter und vorher definierter Primitive. Die Kontur wird dann in solche Segmente zerlegt, die den vorgegebenen Primitiven entsprechen. Die Organisation erfolgt durch die Aneinanderreihung der entsprechenden Nummern [46].

Regionbasierte globale Verfahren betrachten sämtliche Pixel des Objekts in ihrer Gesamtheit, um daraus eine Objektbeschreibung in Form eines Feature-Vektors zu ermitteln. Die Ähnlichkeit zwischen 2 Objekten wird meist über die Distanz zwischen deren Feature-Vektoren bestimmt. Einige Beispiele für solche Verfahren sind:

- *Geometrisch invariante Momente* werden aus den gemittelten Grauwerten der Pixel gebildet und können gegenüber Rotation, Translation und Skalierung invariant sein. Ein Set bestehend aus solchen invarianten Momenten wird genutzt, um einen Feature-Vektor des Objekts zu erstellen [66]. Darüber hinaus gibt es noch Verfahren, die auf anderen Varianten der Momente basieren.
- Der *Generische Fourier-Deskriptor* verwendet eine 2D-Fourier-Transformation auf die in kartesische Koordinaten umgewandelten Polarkoordinaten der Region [150].
- *Gitter-basierende Methoden* legen ein Gitternetz bestimmter Größe über die Region. Zellen, die das Objekt enthalten, werden auf 1 gesetzt, ansonsten auf 0. Dadurch entsteht ein binärer Feature-Vektor. Auch hier gibt es Varianten, die auf anderen Formen des Netzes basieren.

Regionbasierte strukturelle Verfahren zerlegen wie bei den konturbasierten strukturellen Verfahren die Region in einzelne Bereiche, die dann zur Beschreibung des Objektes verwendet werden. Beispiele für solche Verfahren sind:

- Die *Konvexe Hülle* einer Region wird erzeugt. Die Bereiche, die nicht zur ursprünglichen Region gehören, werden genutzt, um die Region zu beschreiben [36].
- Bei der *Medialen Achse* wird die Region *skelettiert* und die entstandenen Linien werden in Segmente zerlegt und als Graph dargestellt, um die Region zu beschreiben.

2.3 Aktueller Stand der Technik

2.3.1 Bildverarbeitung im Kontext der Fischdetektion

Bei der Vermessung von Fischen mit Hilfe von Kamerasystemen müssen 2 Varianten unterschieden werden. Zum einen Verfahren, bei denen die Fische im Trockenen vermessen werden und zum anderen Verfahren, bei denen die Vermessung unterwasser erfolgt.

Zunächst werden die Verfahren betrachtet, bei denen die Fische im Trockenen vermessen werden. Dies erfolgt zumeist auf Fließbändern mit der Zielsetzung für die Industrie eine Sortierung zu ermöglichen. Bereits 1991 wurde in [132] ein Verfahren zur Gewichtsschätzung von Plattfischen präsentiert, wobei hier zusätzlich strukturiertes Licht eingesetzt wurde. Zwei weitere Varianten zur Längenmessung von Fischen auf Fließbändern werden in [134] und

[147] vorgestellt. Bei [147] befinden sich die Fische auf einem halbdurchsichtigen Fließband, das von unten beleuchtet wird, um die Kontur des Fisches zu verstärken. Sobald ein Fisch sich im Kamerabild befindet, wird ein Bild aufgenommen. Aufgrund der Hintergrundbeleuchtung kann die Kontur des Fisches leicht ermittelt werden. Anschließend wird die Orientierung des Fisches anhand der 2 am weitesten voneinander entfernten Punkte der Kontur ermittelt. Für die Ermittlung der Länge des Fisches wird, um eine Mittellinie zu erzeugen, eine Skelettierung der Kontur vorgenommen. Aufgrund der bekannten Abmessungen der Anlage kann daraus die Länge in cm berechnet werden. Darüber hinaus wird anhand der Kontur zwischen 7 unterschiedlichen Fischarten unterschieden.

Da bei den Fließbandverfahren wegen der fest definierten Systemkonfiguration die Abstände zwischen Kamera und Fließband bekannt sind, wird hier meist nur eine einzelne Kamera eingesetzt. Solche Vermessungssysteme werden inzwischen in der Industrie zum Beispiel auch auf Fischtrawlern erfolgreich eingesetzt. Das Unternehmen VAKI Aquaculture Systems Ltd. [88] hat das Prinzip des Fließbandes für die Aquakultur in der Art weiterentwickelt, dass die Fische anstatt über ein Fließband über einen flachen Wasserkanal transportiert werden und sich deswegen während der Vermessung weiterhin im Wasser befinden. Darüber hinaus bietet das Unternehmen Produkte zur Vermessung von Fischen an, bei denen die Fische eine Schleuse zur Vermessung passieren müssen.

Aufgrund der Anforderung dieser Arbeit, dass die Steinbutte non-invasiv in ihrem Lebensraum vermessen werden sollen, können Verfahren, die auf Fließbändern, beziehungsweise trockener Vermessung basieren, nicht als Grundlage dienen, da sie entweder Stress bei den Fischen verursachen oder tote Fische verwenden. Wasserkanäle und Schleusen reduzieren zwar den Stress im Vergleich zum Fließband, erfüllen aber noch nicht die Anforderung an ein non-invasives Vermessungssystem.

Die Alternative hierzu ist die Vermessung der Fische im Wasser. Im Gegensatz zu den Fließbandverfahren wird im Bereich der Unterwasser-Vermessung von Fischen zumeist auf Stereokamerasysteme zurückgegriffen, da hier der Abstand, in dem sich die Fische befinden, nicht bekannt ist. Das Forschungsgebiet der Fischvermessung ist im Vergleich zu vielen anderen Bereichen relativ jung und die Menge an vorhandener Literatur somit vergleichsweise gering.

Die Themengebiete erstrecken sich dagegen auf einer sehr großen Spannweite. Arbeiten, die sich mit den grundlegenden Aspekten der Fischvermessung unter Wasser befassen, sind zum Beispiel [2], [110] und [116]. Es gibt Lösungen, die mit Hilfe manueller Interaktion eine semi-automatische Vermessung bewerkstelligen. Die Anwendungsbereiche erstrecken sich hierbei von der Vermessung von Thunfischen und der Evaluation der Genauigkeit [61], bis hin

zu der Verwendung von ferngesteuerten Unterwasser-Kamerastationen, um eine Beeinflussung der Fische durch Taucher zu verhindern [38] [72]. In den meisten Fällen handelt es sich um Speziallösungen für explizite Aufgabenstellungen.

Während bei den bisher genannten Verfahren die zu vermessenden Stellen der Fische noch von Hand ausgewählt wurden, ist es von großem Interesse, diesen Vorgang zu automatisieren. Um Fische automatisiert zu vermessen, müssen diese zuvor von einem System eigenständig erkannt werden. Hier gibt es mehrere Verfahren, die sich mit der Klassifikation von mehreren meistens gezielt ausgewählten Fischarten befassen. Bereits 1990 haben Strachan et al. in [135] Verfahren vorgestellt, um 7 Fischarten voneinander zu unterscheiden. Auch [154] präsentiert ein Verfahren, um zwischen 3 Fischarten, die zusammen in Teichanlagen gehalten werden, zu differenzieren. Ein ähnliches Verfahren präsentieren Sze et al. in [136]. Sie stellen ein auf Form, Farbe oder anderen Features basierendes Suchverfahren für eine Fischdatenbank vor. Leider basieren die Verfahren entweder auf geschlossenen Konturen, siehe hierzu Kapitel 2.3.4, oder auf eine trockene Vermessung.

Auch morphometrische Methoden [22] können zur Beschreibung der Fischform verwendet werden, um eine Erkennung von Fischarten zu ermöglichen.

Zur Erkennung von Fischarten werden auch Neuronale Netzwerke eingesetzt, zum Beispiel in [5], wo zwischen 20 Fischfamilien unterschieden werden kann. Bei [133] wird eine Laserschranke und dessen Verformung genutzt, um die Form von Fischen auf einem Fließband zu erfassen und mittels eines Neuronalen Netzwerks einer von 6 Sorten zuzuordnen. Aufgrund dessen, dass die Versuchsbedingungen durch Fließbänder und definierte Hintergründe stark reglementiert sind, scheint die Robustheit für den Anwendungsfall in der hier vorliegenden Arbeit nicht gegeben zu sein.

Eine weitere Möglichkeit zur Detektion von Fischarten ist die Nutzung von Template-Matching-Methoden, deren Anwendung in [115] beschrieben wird.

Ein Framework zur automatischen Fischvermessung wird in [1] vorgestellt. Das Framework verwendet anstatt eines Stereokamerasystems nur eine einzelne Kamera und ist davon abhängig, dass die Distanz zum Fisch bekannt ist.

Ein Verfahren, bei dem frei schwimmende Lachse in einem Becken durch ein Stereokamerasystem vermessen werden, wird in [140] und [86] präsentiert. Zuerst werden potentielle Lachse anhand einer in Differenzbildern erkennbaren Sichelform des Fischkopfes identifiziert. Anschließend wird ein 3-dimensionales Punkt-Verteilungsmodell eines Lachses verwendet, das an die im Bildmaterial gefundenen Konturen der Fische angepasst wird. Aus den Positionen von Maul und Flosse wird dann anhand der Stereobilddaten die Länge berechnet. Dabei wurden in 322 Bildern mit mindestens einem Lachs und einer schwarzen Beckenwand für maximalen Kontrast in 86 Bildern ein Fischkopf

identifiziert, davon 55 korrekt identifizierte Lachsköpfe und 31 falsch positiv identifizierte Objekte. Bei Optimierung der falsch positiv Rate auf 2 fehlerhafte Erkennungen, wurden insgesamt 47 Lachsköpfe korrekt identifiziert, was circa 14,5 % Erkennungsrate entspricht, ausgehend von nur einem zu identifizierenden Lachskopf pro Bild. Das anschließende Matching des Lachsmodells erfolgte bei circa 63 % der Testbilder korrekt. Zusammenfassend ergibt das eine Rate von ungefähr 10 % korrekt erkannter Lachsköpfe mit erfolgreichem Matching des Modells. Die Ergebnisse der Arbeit zeigen jedoch, dass die Konturen von sich in der Nähe befindlichen anderen Lachsen das Ergebnis der Anpassung stören kann. Ebenso führt eine zu starke Differenz zwischen der Form und der Größe des Modells und dem tatsächlichen Lachs im Bild zu Problemen bei der Anpassung des Modells.

Ein weiteres Verfahren, bei dem frei schwimmende Thunfische bei der Überführung von Fangnetzen in Zuchtkäfige mittels eines von Tauchern bedienten Stereokamerasystems vermessen werden, wurde von Costa et al. in [32] und [33] vorgestellt. Zuerst filmen die Taucher die Thunfische gegen die Wasseroberfläche. Aufgrund der kontrastreichen Bilder kann nach Anwendung eines Sobelfilters, einer Binarisierung und mehrerer morphologischer Filter (siehe hierzu auch Kapitel 2.2.2) die Kontur der Thunfische segmentiert werden. Relevante Landmarken der Thunfischkontur werden ermittelt und deren 3-dimensionale Koordinaten mithilfe eines Neuronalen Netzwerkes ermittelt. Aus diesen Daten werden abschließend Größe und Biomasse der Thunfische berechnet. Im Testeinsatz konnten so 992 von insgesamt 1014 Thunfischen erkannt und deren Masse bestimmt werden. Die grobe Unterteilung in große (13 %) und kleine (87 %) Thunfische war dabei sehr nah an der Realität (15 % groß, 85 % klein) und damit präziser als konventionelle Videoverfahren (54 % groß, 46 % klein). Die berechnete Biomasse wies zwar einen Fehler von circa 50 % auf, war damit aber erheblich präziser als die Schätzung basierend auf manueller Videoanalyse, deren Fehleinschätzung der Biomasse bei circa 350 % lag.

Ein junges Forschungsgebiet beschäftigt sich mit der Analyse von Bewegungsmustern von Fischen zur Erkennung von gesunden und kranken Verhaltensweisen. Während in [78] primär Transponder verwendet werden, um das Auftauchverhalten von Heilbutten bei unterschiedlicher Besatzdichte zu ermitteln, existieren verschiedenste Ansätze, die auf Kamerasysteme zurückgreifen.

Zum Beispiel können die Becken mit Markierungen versehen werden, um die relative Positionierung des gesamten Bestandes zu evaluieren. So verwendet [131] ein Kamerasystem unterwasser, das die Sichtbarkeit von Markierungen auf der gegenüberliegenden Beckenwand auswertet, um die Tiefenverteilung von Lachsen in den Becken zu ermitteln.

Der klassische Ansatz ist allerdings ein Motion-basiertes Verfahren, das eine 3D-Trajektorie über die Zeit aufbaut. Um Stressfaktoren, wie Verunreinigungen des Wassers und die daraus resultierenden Verhaltensänderungen zu erfassen, wurde in [74] ein System entwickelt, mit dem mehrere Killifische in separaten Becken automatisiert überwacht werden. [103] nutzt ein Stereokamerasystem, um die Auswirkungen von unterschiedlichen Hintergrundgeräuschen auf das Verhalten von Kugelfisch, Blaubarsch und Seebrasse und deren Bewegung im Becken zu analysieren. Des Weiteren stellt [124] ein System zur Analyse der Bewegungsmuster von Fischen vor, um daraus auf Giftstoffe im Wasser zu schließen. In [153] werden Spiegel verwendet, um mehrere Blickwinkel auf ein Aquarium mit einem einzelnen Kamerabild zu erhalten und daraus 3-dimensionale Positionen der Fische zu berechnen.

Um Muster interpretieren zu können, sind andere Methoden notwendig. Ein Beispiel hierfür sind Markov-Modelle wie zum Beispiel das *Hidden-Markov-Modell* [139] [107]. Mit ihrer Hilfe können gelernte Verhaltensmuster auch wiedererkannt werden. Die Analyse von Bewegungsmustern spielt in dieser Arbeit nur eine untergeordnete Rolle, da sie in der weiteren Anwendung des Systems liegt.

Im Rahmen der Recherche konnte kein Verfahren gefunden werden, dessen Anwendung auf den in dieser Arbeit beschriebenen Anwendungsfall direkt übertragbar war. Dies liegt darin begründet, dass bei keinem der beschriebenen Verfahren eine Besatzdichte vorlag, bei der kein sich vom zu segmentierenden Fisch unterscheidbarer Hintergrund zu sehen ist.

2.3.2 Unterwasser-Kalibrierung von Stereokamerasystemen

Die in früheren Arbeiten verwendeten Konstruktionen für die Kalibrierung von Kamerasystemen unterwasser sind relativ groß und wurden mit Markern versehen [45] [62] [63] [127]. Als Beispiel dient der Aluminiumrahmen in [83] mit den Maßen 1,4 m \times 1,4 m \times 0,7 m. Eine solche Konstruktion ist für die Nutzung in offenen Gewässern oder in großen Becken ausgelegt und kann aufgrund der Größe nicht im vorliegenden Szenario eingesetzt werden.

Für die Kalibrierung unterwasser sind besondere physikalische Effekte wie Lichtbrechung an Trennflächen, Strahlenausbreitung im Wasser oder Reflexionen zu modellieren. Verfahren, die sich mit diesen Herausforderungen befassen haben, sind in [130] und [129] vorgestellt. Eine Analyse der Brechungseffekte und die Vorstellung eines Modells, das diese Effekte ebenso berücksichtigt, wird in [137] präsentiert. Ein Kalibrierverfahren basierend auf einer geometrischen Prozedur zur Brechkorrektur wird in [80] erläutert.

Keines der physikalischen oder geometrischen Modelle konnte die komplexe

Realität bisher perfekt abbilden. Gegensätzlich dazu gibt es die Möglichkeit, Unterwasser-Kameras zu kalibrieren, ohne dass die Kameras sich unterwasser befinden. Dies wird in [81] und [82] mit dem Ergebnis untersucht, dass eine Approximation durch Anpassung der Kameraparameter möglich ist. In [122] wird ein Kameramodell angewandt, das explizit von einer Lichtbrechung ausgeht, mit dem gleichen Ziel, die Kameraparameter zu ermitteln, ohne eine Kalibrierung unterwasser vornehmen zu müssen.

Treibitz stellt in [141] ein Kalibrierverfahren vor, das entgegen typischer Methoden von mehreren Standpunkten der Kamera ausgeht. Dies geschieht, um einem Effekt entgegenzuwirken, der durch die Lichtbrechung an den Trennflächen entsteht und mit steigendem Abstand der Kamera verstärkt wird und zu größer werdenden Fehlmessungen führt. Ist der Abstand der Kamera jedoch so gering, dass der Faktor zwischen der Brennweite einer Luft- und einer Unterwasser-Kalibrierung circa 1,33 entspricht, kann ein einzelner Standpunkt der Kamera angenommen werden.

Da der Hardwareaufbau Inhalt der eigenen Konzeption ist, besteht die Möglichkeit, Rücksicht auf diesen Sachverhalt zu nehmen. Dabei stellt sich die Frage nach der Form des Kameragehäuses. In [79] wird der Unterschied zwischen planen und hemisphärischen Trennflächen zwischen Kamera und Wasser untersucht. Die hemisphärische Trennfläche bietet die Möglichkeit, keine weiteren Verzerrungen durch die Lichtbrechung den intrinsischen Kameraparametern hinzuzufügen, sofern die Rundung einer solchen Kuppel genau einer Kugel entspricht und die Positionierung der Kamera exakt in der Mitte der Kuppel liegt. Für die Kalibrierung beider Varianten wird ein physik-basiertes Modell vorgestellt, aber auch darauf hingewiesen, dass typische Verfahren eine zumeist ausreichende Genauigkeit liefern.

Für den Fall, dass in der Kalibrierphase die Wasserqualität schlecht ist, wird in [20] ein Verfahren zur Kalibrierung von Unterwasser-Kameras vorgestellt, dass nicht, wie sonst üblich, sämtliche Markierungen des Kalibrierobjekts benötigt.

Einen interessanten Ansatz bietet das Verfahren zur Selbstkalibrierung einer Unterwasser-Kamera in [109]. Dabei wird kein Kalibriermuster verwendet, sondern die Veränderungen in der Szene, die durch die Kamerabewegung entstehen.

2.3.3 Stereosysteme

Stereokamerasysteme sind heutzutage weit verbreitet und werden in vielen unterschiedlichen Bereichen eingesetzt. Die Standardliteratur ist in diesem Bereich vielfältig [42] [60].

In den letzten Jahren erreichten Stereokamerasysteme den massentaugli-

chen Markt der Unterhaltungsbranche. So bietet Microsofts Kinect [31] [11] die Möglichkeit, mittels Gesten Videospiele oder das Windows Media Center zu steuern. Für die Produktion von 3D-Kinofilmen werden professionelle Stereokameras verwendet [23]. Durch die Verbreitung von 3D-Fernsehern ist auch im Privatbereich der Einsatz von Stereokameras immer beliebter. So gibt es inzwischen Camcorder [29], Digicams [30] und Smartphones [51] mit Stereokameras.

Im industriellen Sektor entwickelten zum Beispiel Daimler [4] und Continental [3] Stereokamerasysteme zur Unterstützung ihrer Fahrerassistenzsysteme und für eine Verbesserung der Verkehrssicherheit. Ziel ist es, Fußgänger und andere Hindernisse zu erfassen und entsprechende Reaktionen einzuleiten. Ein weiteres Einsatzgebiet ist die robotergestützte minimal-invasive Chirurgie. Hier werden endoskopische Stereokameras eingesetzt, um für die Chirurgen ein 3D-Bild zu erzeugen, wodurch die Steuerung der Roboterarme vereinfacht wird [65].

Ein weiteres Einsatzgebiet von Stereokamerasystemen ist die Vermessung und Rekonstruktion von Objekten, Oberflächen und Landschaften. So befindet sich die *High Resolution Stereo Camera* (HRSC) im Orbit des Planeten Mars, um dessen Oberfläche zu erfassen [69], [64]. Aber auch mit Stereokameras ausgerüstete Flugzeuge werden eingesetzt, um Landschaften und Städte zu rekonstruieren [73].

Ein großes Forschungsgebiet ist die Robotik, bei der für die Navigation häufig Stereokamerasysteme eingesetzt werden. [58] präsentiert ein Verfahren zur Rekonstruktion von Innenräumen aus Stereobildern, das für die Roboternavigation geeignet ist. Bei der Robotik werden oft viele unterschiedliche Sensorvarianten kombiniert. So werden in [104] Stereokameras in Verbindung mit polarisierten Bildern verwendet, um Wasserhindernisse auszumachen. Ein anderes Beispiel ist [148], wo neben einem Stereokamerasystem ein Laserscanner für die Navigation durch unebenes Gelände zum Einsatz kommt.

Diese große Auswahl an unterschiedlichsten Einsatzmöglichkeiten zeigt, dass die allgemeine Verwendung von Stereokamerasystemen inzwischen das Stadium der Forschung verlassen hat und die zugrunde liegenden Verfahren eine entsprechende Robustheit erreicht haben, um im Massenmarkt eingesetzt zu werden.

Aus Sicht dieser Arbeit ist die Rekonstruktion von Strukturen unter Wasser am interessantesten. Die Arbeit von [8] verwendet eine große Baseline zur Rekonstruktion von Korallen. Manche Verfahren verwenden zusätzliche Sensoren, um das Stereokamerasystem zu unterstützen. So kommt bei [68] ein Fächerecholot zum Einsatz und in [19] werden Muster strukturierter Lichter verwendet. Diese dienen dazu, die Oberflächenstruktur der Objekte hervorzuheben und damit die Rekonstruktion zu verbessern. In dieser Arbeit sollte auf

solche zusätzlichen Hilfsmittel verzichtet werden, da diese für die Steinbutte Stress bedeuten würden. Darüber hinaus liegen die Steinbutte überwiegend auf dem Grund des Beckens und bilden dabei eine sehr gleichmäßige Oberfläche, bei der die Übergänge zwischen den Fischen nicht hervorgehoben werden.

Zur Analyse von Stereobildpaaren kann inzwischen auf frei verfügbare Frameworks zurückgegriffen werden, wie zum Beispiel das in dieser Arbeit verwendete OpenCV [102].

2.3.4 Konturbasierte Formerkennung

Die Segmentierung in Bildern ist ein eigener Forschungszweig. Betrachtet man die möglichen Beschreibungskriterien Farbe, Form und Textur, so muss man im Kontext dieser Arbeit feststellen, dass weder Farbe (Kapitel 3.3.1), noch Textur (Kapitel 2.1) für eine Analyse in Frage kommen. Für eine Übersicht zu grundlegenden Verfahren der konturbasierten Formerkennung siehe Kapitel 2.2.3.2 und die dort genannten Beispiele.

Die meisten Verfahren zur konturbasierten Formerkennung benötigen eine vollständige Kontur, die das Objekt umschließt. Kann für das gesuchte Objekt eine geschlossene Kontur segmentiert werden, so steht eine große Anzahl an Verfahren zur Verfügung, die unter Verwendung unterschiedlichster Parameter ein Objekt einer bestimmten Objektklasse zuweisen. Als mögliche Beispiele seien die folgenden Verfahren genannt. In [21] basiert die Objektbeschreibung auf der Position der Konturpunkte und deren Distanz zum Mittelpunkt. Das Matching der zu vergleichenden Konturen wird von einem Split-Merge-Verfahren unterstützt und anhand von Werkzeug-Konturen überprüft. [10] verwendet Punktkorrespondenzen der Konturen, um die Ähnlichkeit zweier Konturen zu überprüfen. Für die Korrespondenzfindung wird für jeden Punkt die relative Verteilung der anderen Konturpunkte berechnet und verglichen. In [123] werden Konturen, die Schleifen enthalten, derart behandelt, dass von 2 Konturen mit denselben Endpunkten ausgegangen wird, die bei Verzweigungen und Schleifen die unterschiedlichen Wege ablaufen. Eingesetzt wurde dieses Verfahren an Röntgenaufnahmen von Knien, mit vermehrter Schleifenbildung am Knochenende. Bei [126] wird die Kontur des Objekts in gleichlange Kurven unterteilt. Im Anschluß werden daraus halb so viele eindimensionale Funktionen ermittelt, deren Mittelwerte und Varianzen als Featurevektor dienen. Getestet wird dieses Verfahren an einem Datenset bestehend aus Baumblättern und einem Datenset mit diversen Objekten. [47] verwendet dynamische Programmierung, um 2 Konturen zu matchen. Dabei werden zuerst die beiden Konturen gematcht und anschließend in 2 Symbolstränge überführt, für die erneut die optimale Ausrichtung mit maximaler Ähnlichkeit ermittelt wird. Im Kontext dieser Arbeit muss jedoch aufgrund der hohen Besatzdichte

davon ausgegangen werden, dass ein Fisch in seiner gesamten Form nicht segmentiert werden kann, wodurch diese Verfahren hier nicht eingesetzt werden können und eine Alternative gefunden werden muss.

Bereits 1988 beschäftigte sich Gorman [54][53] mit dem Matching von unvollständigen Konturen, in diesem Fall von Flugzeugen, wobei in der Kontur keine Lücken vorhanden sind, sondern fehlerhafte Passagen. Insgesamt 4 unterschiedliche Kontur-Signaturen werden in [95] vorgestellt und auf ihre Qualität bei unvollständigen Konturen hin überprüft.

Die bisher genannten konturbasierten Verfahren zur Formerkennung benötigen eine vollständige Kontur des zu erkennenden Objekts, wie in [25] erwähnt. Bei einer sehr hohen Besatzdichte und nicht optimalen Lichtbedingungen muss davon ausgegangen werden, dass von Einzelobjekten nur Bruchstücke der Kontur erfasst werden können. Für dieses spezielle Anwendungsgebiet scheinen 2 Verfahren sehr interessant. Ein Verfahren wird in [97] und [98] vorgestellt, bei dem die Objekterkennung über die Verwendung von Konturfragmenten und deren relative Position zum Objekt-Mittelpunkt erreicht werden kann. Die Detektoren für die Konturfragmente werden dabei in einem Postprocessing-Schritt in einem sogenannten Boosting verbessert, dabei werden mehrere schwache Detektoren zu einem starken Detektor zusammengefasst. Auf diese Weise ist es möglich auch Objekte zu erfassen, von denen aufgrund von schlechten Sichtverhältnissen oder teilweisen Verdeckungen keine vollständig zusammenhängende Kontur erzeugt werden kann. Zu Testzwecken werden Objekte aus 17 unterschiedlichen Kategorien, wie zum Beispiel Autos, Fahrräder, Gesichter, Kühe und Pferde, erkannt. Eine weitere Methode in [128] ist ähnlich aufgebaut, verzichtet aber auf den Boosting-Ansatz.

Aquakultur-Monitoring von Fischwachstum

Inhalt

3.1	Anforderungen an das Stereokamerasystem	39
3.2	Material: Hardwaresystem	42
3.2.1	Planung des Stereokamerasystems	42
3.2.2	Resultierendes Stereokamerasystem	48
3.2.3	Systemaufbau am Versuchsstandort	54
3.3	Methoden: Parameter des Beleuchtungssystems und Unterwasser-Kalibrierung	56
3.3.1	Versuchsaufbau zur Farbwahl des Beleuchtungssystems	57
3.3.2	Versuchsaufbau zur Ausrichtung des Beleuchtungssystems	60
3.3.3	Kalibrierung des Stereokamerasystems für Unterwasser-Einsatz	62
3.4	Ergebnisse: Parameter des Beleuchtungssystems und Unterwasser-Kalibrierung	66
3.4.1	Ergebnisse zur Farbwahl des Beleuchtungssystems . .	66
3.4.2	Ergebnisse zur Ausrichtung des Beleuchtungssystems .	77
3.4.3	Evaluation der Kalibrierung	78

3.1 Anforderungen an das Stereokamerasystem

Im Folgenden sollen die Anforderungen an ein System zur non-invasiven Vermessung von Steinbutten in der Aquakultur benannt werden. Neben Anforderungen, die die Arbeitssicherheit und Robustheit in entsprechenden Arbeitsumgebungen betreffen, können die im Bereich der Beleuchtung getroffenen Entscheidungen negative Auswirkungen auf die Entwicklung der Fische ausüben. Ziel des non-invasiven Verfahrens ist es, den auf die Steinbutte ausgeübten Stress möglichst gering zu halten, um das Wachstum nach Möglichkeit

nicht zum Negativen zu beeinflussen. Aufgrund dessen entfallen auch Varianten, die ein Entnehmen des Fisches aus dem Becken beinhalten oder bei denen die Fische durch eine Schleuse oder eine spezielle Konstruktion getrieben werden, um so Einzelmessungen zu ermöglichen.

Kameras, Objektive, Kameraanordnung Ein Stereokamerasystem soll für die non-invasive Vermessung der Steinbutte eingesetzt werden. Dafür müssen die Kameras sowie die Optik entsprechend ausgewählt werden. Entscheidende Faktoren sind hier Geschwindigkeit, Lichtempfindlichkeit, Bildgröße und Bildqualität sowie der überwachbare Bereich des Beckens. Des Weiteren ist für ein Stereokamerasystem die Anordnung der Kameras für die Genauigkeit der Vermessung von Bedeutung.

Umweltbedingungen Ein entscheidender Faktor ist, dass das System unterwasser funktionieren muss. Ein auf Kameras basierendes System, das von oberhalb der Wasseroberfläche auf Fische im Wasser schaut, hätte mehrere Probleme zu überwinden. Erstens entstehen Reflexionen auf der Wasseroberfläche sowie Kaustiken unterwasser von diversen Lichtquellen, seien es Sonneneinfall, Deckenbeleuchtung oder eine etwaige Beleuchtung für das System. Zweitens ist die Wasseroberfläche nicht eben. Dadurch werden zum einen die eben erwähnten Reflexionen auf der gesamten Wasseroberfläche verteilt und zum anderen kommt es durch die an der Wasseroberfläche stattfindende Lichtbrechung bei unregelmäßiger Oberfläche zu nicht nachvollziehbaren Verzerrungen, so dass eine adäquate Größenmessung der Fische nicht mehr möglich ist.

Die Robustheit des Systems ist eine weitere essentielle Anforderung. Da es das Ziel ist, eine kontinuierliche Wachstumsüberwachung durchzuführen, müssen Hardware und Recheneinheit über einen langen Zeitraum im Betrieb sein, anstatt nur kurzzeitig eingesetzt zu werden. Demnach muss das System so konstruiert sein, dass es den Umweltbedingungen einer Aquakulturanlage, wie hohe Luftfeuchtigkeit mit hohem Salzgehalt sowie das Eintauchen in Salzwasser, dauerhaft ausgesetzt werden kann.

Elektrik Aus Sicherheitsgründen wurde für das Gesamtsystem festgelegt, dass eine Niederspannung von 48 Volt nicht überschritten werden soll. Da das System ins Wasser getaucht wird, ergibt sich eine weitere Anforderung basierend auf der Tatsache, dass Wasser Strom leitet. In der Aquakultur werden unterschiedliche Sensoren zur Überwachung der Anlage verwendet, die zum Beispiel die Wasserqualität oder den Sauerstoffgehalt messen. Die Sensoren, die den Sauerstoffgehalt messen, reagieren bereits auf Spannungen im Bereich von Millivolt. Außerdem befinden sich im Wasser lebende Tiere, die keinem

steten Stromfluss ausgesetzt werden sollen. Deswegen ist es für das System unumgänglich, dass es vollständig abgedichtet ist und keinerlei Kriechströme in das Wasser gelangen können.

Verschmutzung Die Arbeit mit einem Kamerasystem unterwasser birgt noch weitere Anforderungen. Schwebeteilchen, die im Wasser umher treiben, befinden sich neben den Fischen auch im Becken. Werden diese von einer Lichtquelle in unmittelbarer Kameranähe direkt angestrahlt, reflektieren sie sehr stark und verursachen überbelichtete Spotlights in den Aufnahmen. Hinzu kommt, dass sich Schwebeteilchen mit der Zeit am System festsetzen und es dadurch verschmutzen. Dies trifft auch auf die Kameras beziehungsweise vorhandene Sichtfenster für die Kameras zu, was die Sicht und somit die Bildqualität stört. Neben den Schmutzpartikeln können sich ebenfalls Luftblasen festsetzen und die Bildqualität ebenfalls beeinträchtigen. Aufgrund der Anreicherung des Wassers mit Sauerstoff muss im Vergleich zu unangereichertem Wasser mit einem vermehrten Aufkommen an Luftblasen gerechnet werden.

Beleuchtung Die Aquakulturanlage befindet sich im Allgemeinen und auch in diesem Fall in einer Halle mit relativ geringer Beleuchtung. Des Weiteren bestehen in der Versuchsanlage die Mastbecken aus schwarzem Kunststoff, welcher einfallendes Licht großteilig absorbiert. Die Becken sind bei einem Durchmesser von circa 1,8 m und einer Wandhöhe von circa 1,1 m relativ schmal und hochwandig, wodurch nur wenig Umgebungslicht in die Becken gelangt. Es ist eine zusätzliche Beleuchtung nötig, um bei der Aufnahme von Bildmaterial eine ausreichende Bildqualität für die weitere Verarbeitung zu gewährleisten. Umso mehr Licht vorhanden ist, desto geringer fällt das Rauschen im Bild aus und die Belichtungszeit kann kürzer gewählt werden, wodurch die Bewegungsunschärfe reduziert wird. Eine Erhöhung der Lichtmenge, bis das Bildmaterial eine optimale Qualität erreicht, ist jedoch nicht möglich. Dies liegt am Steinbutt begründet, da diese Fischart die Dunkelheit bevorzugt. Eine zu intensive Beleuchtung würde bei den Tieren zu Stress führen. Das Ziel der non-invasiven Vermessung ist es, den auf die Tiere ausgeübten Stress möglichst gering zu halten. Daher ist es notwendig, hier einen Kompromiss zwischen Bildqualität und Beleuchtungsstärke zu finden.

Fischerkennung Der Steinbutt beinhaltet aufgrund seiner Eigenschaften weitere Anforderungen an das System. Durch die hohe Besatzdichte und die Tatsache, dass der Steinbutt in mehreren Lagen auf dem Grund des Beckens liegt, muss das Verfahren zur Fischerkennung in der Lage sein, ein einzelnes Exemplar vor einem Hintergrund bestehend aus anderen Steinbutten zu erkennen und durch Überlagerungen unvollständige Steinbutte auszulassen.

sen. Erschwerend kommt in diesem Fall noch hinzu, dass der Steinbutt ein Tarnmuster besitzt, um sich von seiner Umgebung optisch nicht abzuheben. Darüber hinaus besitzt er die Fähigkeit, dieses Muster den Umgebungsbedingungen anzupassen. Eine ausführliche Betrachtung der Anforderungen an das Verfahren zur Erkennung von Steinbutten erfolgt in Kapitel 4.1.

3.2 Material: Hardwaresystem

3.2.1 Planung des Stereokamerasystems

Ziel dieser Arbeit ist die Konzeption eines Stereokamerasystems mit einer spezialisierten Aufgabe. Deswegen müssen die Komponenten für die Aufgabe entsprechend ausgewählt werden. Im Speziellen sei hier die Auswahl der Kameras und der Objektive genannt. Weitere entscheidende Punkte, die es zu beachten gibt, sind die Anordnung der Kameras zueinander, der Aufbau des Gehäuses, die zu verwendenden Materialien, das Beleuchtungssystem, die Stromversorgung sowie die Datenübertragung von den Kameras zur Recheneinheit. Dabei dienen die in Kapitel 3.1 genannten Anforderungen als einzuhaltende Rahmenbedingungen für den Prototypen.

Kameras Bei der Wahl der Kameras sind folgende Eigenschaften von Interesse: Auflösung, Größe und Lichtempfindlichkeit der Bildsensoren sowie Aufnahmegeschwindigkeit.

Aufgrund der dunklen Lichtverhältnisse müssen die Kameras eine hohe Lichtempfindlichkeit aufweisen. Diese ist abhängig von der Größe des Bildsensors in Relation zur Auflösung der Kamera sowie der Art des Bildsensors. Umso größer der Bildsensor ist, desto mehr Licht kann aufgenommen werden. Mit ansteigender Auflösung reduziert sich die Lichtmenge pro Pixel und somit die allgemeine Lichtempfindlichkeit der Kamera. Prinzipiell ist ein größerer Bildsensor von Vorteil, bei der Auflösung muss aber ein Kompromiss gefunden werden zwischen Lichtempfindlichkeit und dem möglichen Detailgrad einer höheren Auflösung.

Bei der Art des Bildsensors gibt es die Wahl zwischen *CCD-Sensoren* und *CMOS-Sensoren*. Ursprünglich waren die CCD-Sensoren den CMOS-Sensoren bei der Lichtempfindlichkeit überlegen, in der Bildauslesung jedoch langsamer. Inzwischen sind jedoch CMOS-Sensoren mit einer höheren Lichtempfindlichkeit als CCD-Sensoren erhältlich, bei gleich bleibendem Geschwindigkeitsvorteil, und sollten, sofern verfügbar, gegenüber den CCD-Sensoren vorgezogen werden.

Objektive Für die Wahl der Objektive ist das Arbeitsumfeld genauer zu betrachten. Die Becken haben einen Durchmesser von circa 1,80 m und der Wasserpegel liegt bei circa 80 cm. Vom Beckenboden soll ein möglichst großer Bereich überwacht werden. Dies kann mit einem großen *Field of view* (FoV) beziehungsweise Öffnungswinkel erreicht werden. Dabei muss beachtet werden, dass sich das FoV von Objektiven unterwasser aufgrund der Lichtbrechung beim Übergang von Luft ins Wasser verengt. Hier gilt das *Snelliussche Brechungsgesetz*, mit dem das neue FoV des Objektivs berechnet werden kann:

$$\sigma_2 = \arcsin\left(\frac{\sin(\sigma_1) n_1}{n_2}\right) \quad (3.1)$$

mit σ_1 als Einfallswinkel (halber FoV), σ_2 als Brechungswinkel sowie n_1 und n_2 als Brechungsindex der jeweiligen Medien. Für den vorliegenden Fall sind das für Luft $n_1 = 1,000292$ und für Wasser $n_2 = 1,33$. In der Abbildung 3.1 kann man sehen, dass bei steigendem Einfallswinkel der resultierende Brechungswinkel in Relation immer kleiner wird.

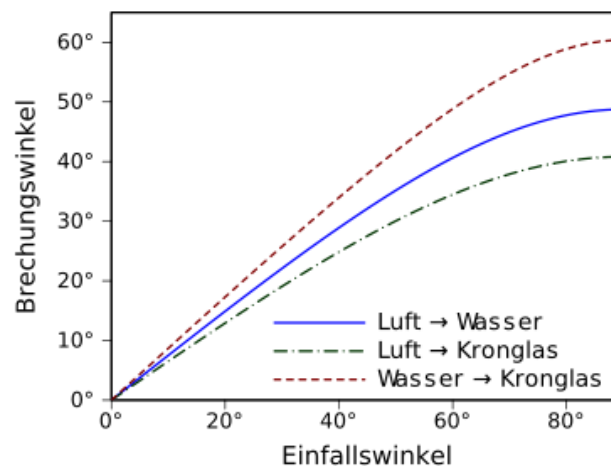


Abbildung 3.1: Winkeländerung aufgrund Lichtbrechung zum Beispiel beim Wechsel von Luft in Wasser [26].

Fisheye-Objektive bieten FoVs von über 180° und würden unter normalen Bedingungen ermöglichen, das gesamte Becken zu überwachen. Im Wasser relativiert sich der Eintrittswinkel von 90° auf circa 48° und der FoV des Fisheye-Objektivs verengt sich auf 96° . Bei einem FoV von mehr als 180° und dem konstruktionsbedingten Hervorstehen der Linse würde außerdem der Effekt eintreten, dass vor dem Wechsel in das Medium Wasser im Randbereich des Bildes das Innere des Gehäuses aufgenommen wird, was auch in der Abbildung 3.2 zu sehen ist. Dies führt zu einem Verlust von nutzbarem

Bildbereich als auch zu möglichen Komplikationen bei der Kalibrierung des Stereosystems. Um das Sichtfeld von Fisheye-Objektiven vollständig nutzen zu können, müssten die Objektive in Kuppeln untergebracht sein, was erstens einen großen Mehraufwand bei der Konstruktion bedeuten würde und zweitens durch die Krümmung des Materials bei der Kalibrierung des Stereosystems extreme Probleme erwarten lässt. Bei Objektiven mit einem FoV von circa

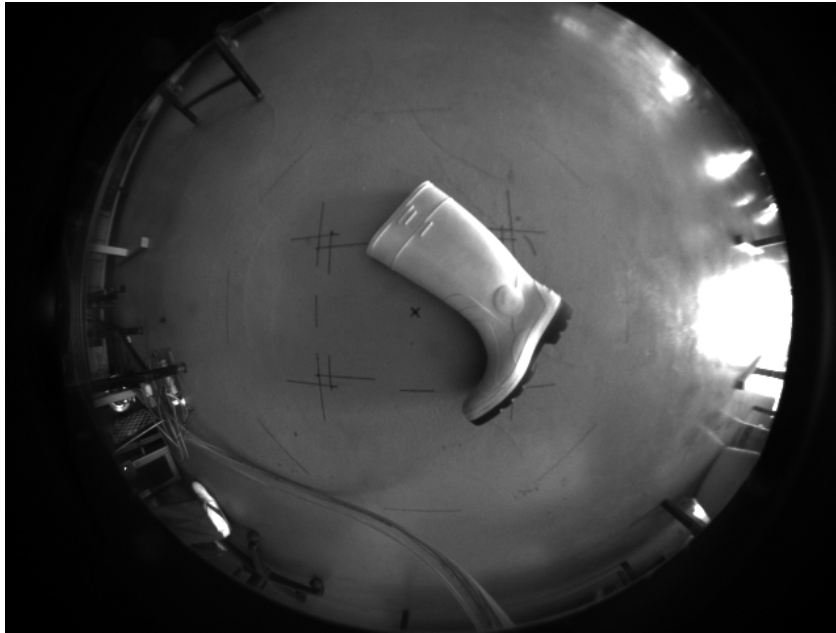


Abbildung 3.2: Durch das Gehäuse verdeckte Randbereiche bei Fisheye-Objektiven verkleinern den effektiven Bildausschnitt.

90° reduziert sich dieser durch die Lichtbrechung auf ungefähr 64°. Bei einer Wassertiefe von 80 cm wäre mit einem solchen Objektiv die Überwachung von einem circa 1 m durchmessenden Bodenbereich möglich.

Kameraanordnung Ein charakterisierender Faktor für ein Stereosystem ist die sogenannte *Baseline* oder Grundlinie b , die den Abstand der beiden Kameras voneinander beschreibt. Umso größer der Abstand der beiden Kameras zueinander, desto stärker ist die Disparität (Kapitel 2.2.1.4.2) zwischen korrespondierenden Pixeln und eine Tiefenbestimmung wird dadurch, insbesondere auf größere Distanz, exakter [71]. Die Arbeitsumgebung für das Stereosystem ist mit einer maximalen Distanz von circa 1 m stark begrenzt, weswegen die Wahl einer möglichst großen Grundlinie hier nicht benötigt wird. Entscheidender für die Wahl der Grundlinie ist in diesem Fall der Nahbereich des Stereosystems. Umso größer die Grundlinie gewählt wird, desto größer ist die

Distanz p_1 , in der sich die Sichtbereiche der beiden Kameras zu überschneiden beginnen:

$$p_1 = \frac{b}{2} * \cot \frac{\alpha}{2}, \quad (3.2)$$

mit α als dem FoV des Objektivs. Soll ein Objekt einer bestimmten Größe x in beiden Bildern vollständig vorhanden sein, muss der Bereich der Überschneidung mindestens diese Größe besitzen, was bei einem FoV von α in der Distanz p_2 erreicht ist:

$$p_2 = \frac{x}{2} * \cot \frac{\alpha}{2} + p_1. \quad (3.3)$$

Somit ist die Distanz p_2 der Mindestabstand vom Stereosystem, der benötigt wird, um ein Objekt der Größe x in beiden Kameras zu erfassen. Die Abbildung 3.3 zeigt eine Schemadarstellung der Zusammenhänge. Ein Stereosystem mit einer Baseline $b = 15$ cm und einem FoV in Wasser von $\alpha = 64^\circ$ ergibt für einen Steinbutt der Größe $x = 20$ cm einen Mindestabstand von circa $p_2 = 28$ cm. Die Verringerung der Baseline auf $b = 12$ cm reduziert den Mindestabstand auf circa $p_2 = 25,5$ cm. In Anbetracht der Tatsache, dass Objektive mit fixen Brennweiten und einem FoV um die 90° meistens eine *Mindest-Objekt-Distanz* (M.O.D.) von 20 - 30 cm aufweisen¹, macht eine kürzere Baseline wenig Sinn. Auf der anderen Seite sollte die Baseline auch nicht viel größer gewählt werden, damit ein möglichst großer Bereich überwacht werden kann.

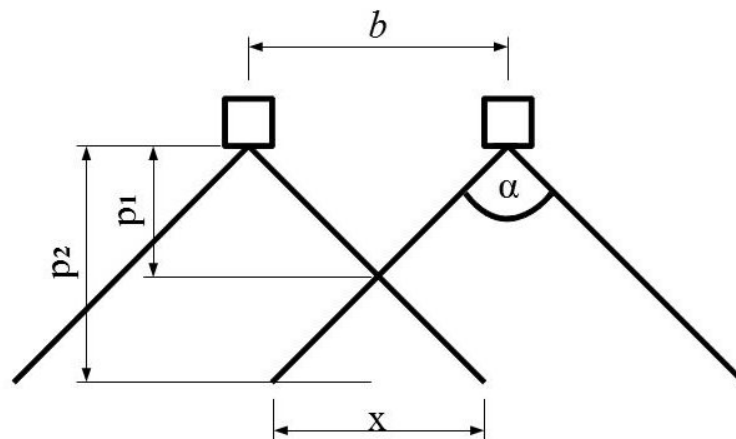


Abbildung 3.3: Schema zur Berechnung der Baseline b .

¹zum Beispiel <http://www.lensation.de> oder <http://security-systems.pentax.de/>

Umweltbedingungen Um mit dem Stereokamerasystem Aufnahmen unter Wasser durchführen zu können, wird ein wasserdichtes Gehäuse benötigt. Aufgrund der angedachten langen Einsatzzeit in diesem Medium, ist es unabdingbar, dass das verwendete Material verwitterungs- und korrosionsbeständig ist. Hier bieten sich Kunststoffe wie PVC an. Ein solches Gehäuse muss darüber hinaus Sichtfenster für die Kameras besitzen. Hier erscheint die Verwendung einer möglichst großen Sichtscheibe am praktikabelsten, da hierdurch die Möglichkeit gegeben wird, Änderungen an der Baseline oder bei den verwendeten Objektiven vorzunehmen. Dies wäre bei vorher in Größe und Position festgelegten Sichtlöchern nicht gegeben. Als Material für die Sichtscheibe erscheint Acrylglas aufgrund seiner Robustheit und guten Verarbeitbarkeit als die richtige Wahl. Darüber hinaus benötigt ein solches Gehäuse eine Kabelzuführung für Strom- und Datenkabel. Hier muss ebenfalls darauf geachtet werden, dass kein Wasser in das Gehäuse eindringen kann.

Das Stereokamerasystem wird in einer Halle mit hoher Luftfeuchtigkeit und der Möglichkeit von Spritzwasser montiert und wird dementsprechend konstruiert werden müssen. Um die Kosten gering zu halten dient ein handelsüblicher PC als Recheneinheit, der üblicherweise für diese Umweltbedingungen nicht geeignet ist. Die Recheneinheit wird aus diesem Grund nicht in der Halle, sondern in einer anderen Räumlichkeit aufgebaut werden. Typische Kameraschnittstellen wie USB oder Firewire (IEEE 1394) sind für eine Übertragung über so lange Strecken nicht ausgelegt. Daher empfiehlt es sich, für die Datenübertragung auf *Gigabit-Ethernet* (GigE) zurück zu greifen. Neben einer ausreichend hohen Datenübertragungsrate kann auch die große Distanz zwischen Kameras und Recheneinheit überbrückt werden. Im Idealfall ist die Nutzung des unternehmensinternen Netzwerks möglich.

Elektronik Es ist für einen ausreichenden Spritzwasserschutz bei sämtlichen elektronischen Komponenten des Systems wie zum Beispiel dem Netzteil zur Stromversorgung der Komponenten, etwaiger anderer Steuerungskomponenten oder den Steckverbindungen von Kabeln zu sorgen. Mit diesem Vorgehen wird vor allem die Sicherheit der Mitarbeiter gewährleistet, wie auch die Funktionalität des Kamerasystems sichergestellt.

Verschmutzung Um die Sichtscheibe von Verunreinigungen und Luftblasen frei zu halten, erscheinen mehrere Vorgehensweisen erfolgversprechend. Die einfachste Variante ist es, die Scheibe regelmäßig mit einem Abzieher manuell zu reinigen. Eine andere Möglichkeit wäre es, eine motorisierte automatische Vorrichtung an dem Gehäuse zu befestigen. Der Reinigungsvorgang könnte dann in gleichmäßigen Zeitabständen aktiviert werden. Hier wäre auch vorstellbar, dass der Verschmutzungsgrad über das Stereosystem ermittelt

werden könnte und die Reinigung softwareseitig ausgelöst wird. Eine andere Idee, um die Sichtscheibe frei zu halten, ist die Verwendung eines kontinuierlichen Wasserstrahls auf die Scheibe, wodurch Luftblasen und Schmutz durch den Strom fortgespült würden. Eine Möglichkeit den Wasserstrahl zu betreiben, wäre über eine separate Wasserpumpe, die das Wasser aus dem Becken ansaugt. Eine andere vorstellbare Möglichkeit wäre es, Wasser aus dem Zufluss abzuzweigen und zu verwenden. Sofern der dabei entstehende Druck ausreicht, um die Scheibe frei zu halten, könnte bei dieser Variante auf eine zusätzliche Pumpe verzichtet werden.

Beleuchtung Das Beleuchtungssystem für das Stereosystem bedarf einiger vorhergehender Überlegungen. Bei der Wahl der Leuchtmittel erscheinen LEDs aufgrund des guten Verhältnisses von Stromverbrauch zu Leistung sowie der kompakten Größe und somit guter Verarbeitungsmöglichkeiten als die beste Wahl. Die LEDs sollten dabei von ihrer Positionierung her maximal auf gleicher Höhe mit der Sichtscheibe sein, bevorzugt sogar etwas dahinter. Hierdurch wird verhindert, dass eine LED direkt in die Kameras strahlt. Um das Problem der direkten Bestrahlung von Schwebeteilchen zu reduzieren, dürfen die LEDs nicht in unmittelbarer Nähe der Kameras montiert werden. Hier kann sich an der -Fotografie orientiert werden, wo sich die Blitzlichter an verstellbaren Armen befinden. Der Abstand zwischen LEDs und den Kameras sollte möglichst groß sein, um so eine indirekte Beleuchtung zu erzeugen. Bei der Konstruktion der Halterung für die LEDs sollte darauf geachtet werden, dass ein großer Abstand nicht durch schlechte Handhabbarkeit oder geringe Stabilität erkauft wird. Neben der Möglichkeit einer dauerhaften Beleuchtung sollten die LEDs auch als Blitzlicht verwendet werden können, um die Gesamtmenge des ausgestrahlten Lichtes zu reduzieren und dadurch eventuell die Steinbutte weniger zu belasten.

Um eine annähernd homogene Ausleuchtung der gesamten Szenerie zu ermöglichen, sollte nicht eine einzelne Lichtquelle verwendet werden. Statt dessen ist es sinnvoller, mehrere Lichtquellen gleichmäßig um das Stereosystem herum zu verteilen, um eine homogene Ausleuchtung zu ermöglichen. Bei der Konstruktion sollte die Möglichkeit zur Erweiterung oder Reduzierung der Lichtquellen gegeben sein. Darüber hinaus sollten die Halterungen für die Lichtquellen in ihrer Ausrichtung einstellbar sein, um eine variable Ausleuchtung zu ermöglichen. Eine annähernd homogene Ausleuchtung soll für diesen Fall durch eine maximale Helligkeitsdifferenz von 50 % zwischen dem hellsten und dem dunkelsten Bereich des Bildes definiert sein.

3.2.2 Resultierendes Stereokamerasystem

3.2.2.1 Unterwasser-Stereokamera

Basierend auf den Überlegungen in Kapitel 3.2.1 wurde folgendes Stereokamerasystem entwickelt. Als Kameras wurden 2 GigE uEye 5240SE-M der Firma IDS gewählt [50]. Diese Kameras verwenden einen 1/1.8“ großen CMOS-Sensor mit hoher Lichtempfindlichkeit und einer Auflösung von 1280×1024 Pixeln. Die Datenübertragung erfolgt über die GigE-Schnittstelle. Die Objektive sind vom Typ Pentax H416 (KP) [112] mit einer Brennweite von 4,2 mm und einem FoV von $86,8^\circ$ an der Luft. Die M.O.D. beträgt bei diesen Objektiven 20 cm.

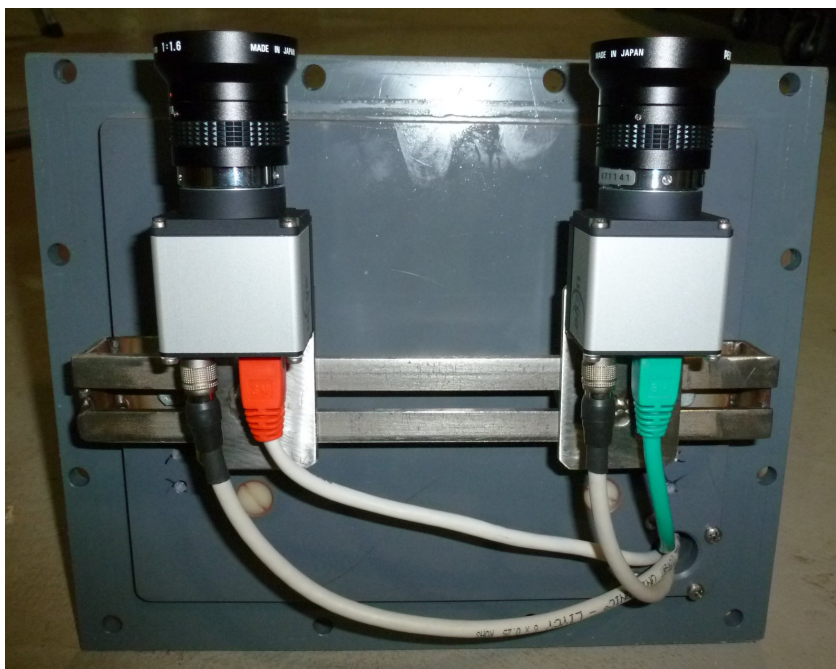


Abbildung 3.4: Beide Kameras vom Typ GigE uEye 5240SE-M mit Objektiven der Sorte Pentax H416 (KP) montiert auf Metallschiene für variable Positionierung.

Die Baseline beträgt circa 13 cm, wodurch ab einer Entfernung von ungefähr 27 cm ein 20 cm großes Objekt von beiden Kameras vollständig erfasst werden kann. Die Kamerahalterung befindet sich am Deckel des wasserdichten Gehäuses zusammen mit der Kabelzuführung, bestehend aus einem 1,5 m langen und 4 cm durchmessenden klarsichtigem PVC-Schlauch auf der Außenseite. Hier befindet sich zusätzlich die Halterung für die Montage über dem Fischbecken. Siehe hierzu die Abbildungen 3.4 und 3.5 zur Veranschaulichung.

Für das wasserdichte Gehäuse als auch für den Deckel wurde PVC-Hart

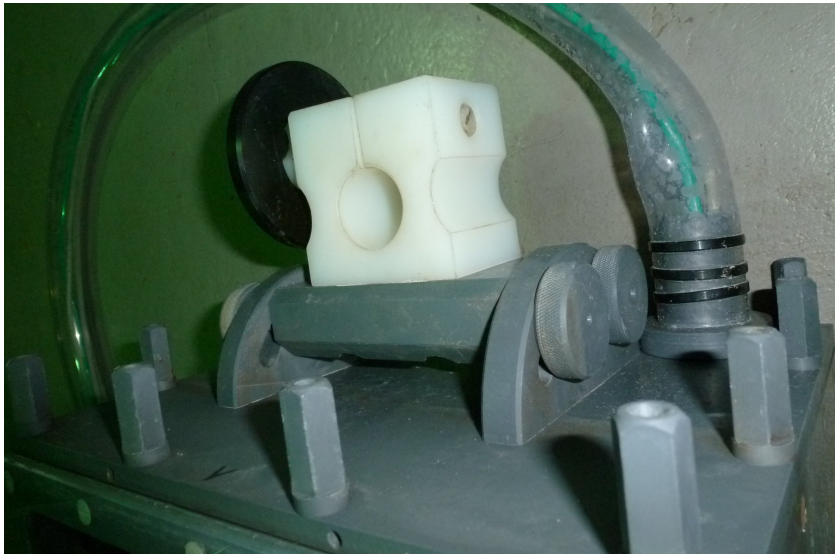


Abbildung 3.5: Oberseite des Gehäuses mit Kabelzuführung und Halterung für das Fischbecken.

Grau verwendet, das mit Tangitkleber verklebt und anschließend von innen zusätzlich mit flüssigem Silikon-N abgedichtet wurde. Die Sichtscheibe besteht aus PMMA (Polymethylmethacrylat oder auch Acrylglas) und wurde ebenfalls wasserdicht mit dem restlichen Gehäuse verklebt. Das Gehäuse hat die Maße 312,5 mm×140 mm×260,5 mm. Die wasserdichte Abdichtung von Deckel und Gehäuse erfolgt über 10 Kunststoffschrauben und einer extra gegossenen selbstnivelierenden Dichtung aus Silikon-N. Das Gehäuse sowie die Dichtung sind in der Abbildung 3.6 zu sehen.

Erste Versuche für die Entfernung von Luftblasen auf der Sichtscheibe verliefen nicht zufriedenstellend, weswegen für diese Arbeit vorerst auf eine manuelle Säuberung zurückgegriffen wurde.

3.2.2.2 Beleuchtungssystem

Für einen großen Abstand der Lichtquellen von den Kameras wurde für das Beleuchtungssystem eine kreisförmige Halterung aus PVC-Hart Grau hergestellt. Die Halterung hat einen Durchmesser von 512 mm, an die die einzelnen LED-Halterungen montiert werden, wodurch für den Ring aus Lichtquellen ein Durchmesser von 600 mm entsteht. Siehe hierzu auch die Konstruktionszeichnungen in Abbildung 3.7. Für diese Arbeit wurden 8 LED-Gehäuse an der Halterung montiert, wobei eine Ausbaureserve für 8 weitere LED-Gehäuse gegeben ist.

Für die LEDs wurden wasserdichte Gehäuse entwickelt, die mit einer dreh-

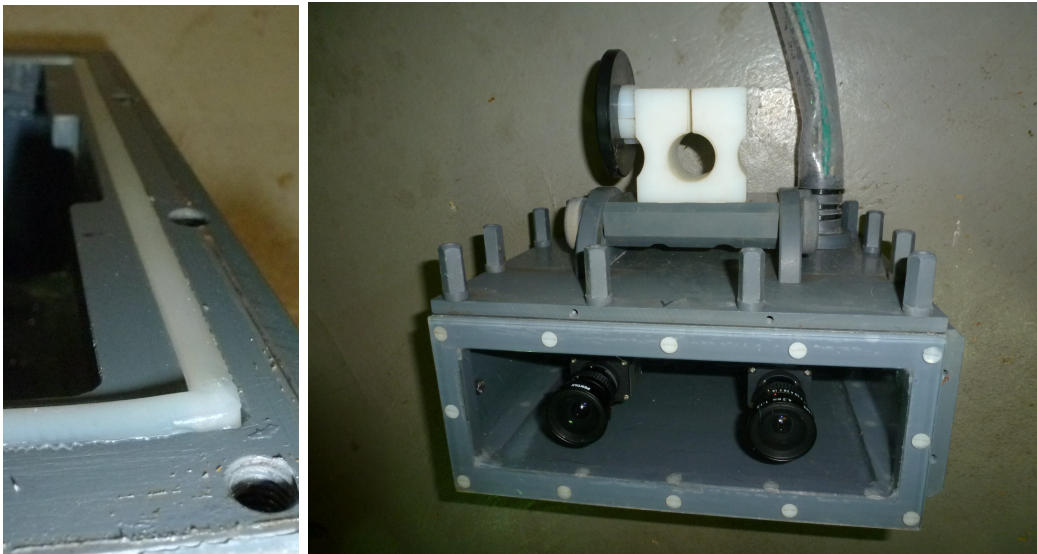


Abbildung 3.6: Links: Dichtung zum wasserdichten Verschließen des Deckels. Rechts: Komplett geschlossenes Gehäuse.

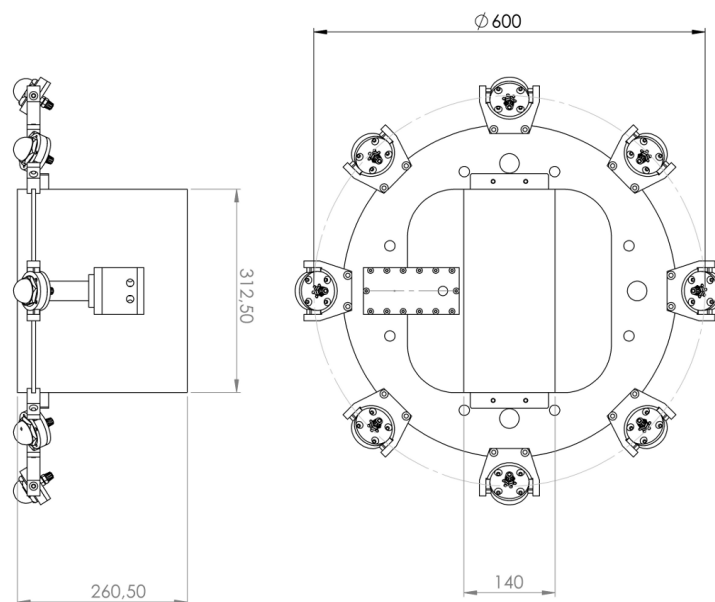


Abbildung 3.7: Konstruktionszeichnung der LED-Halterung.

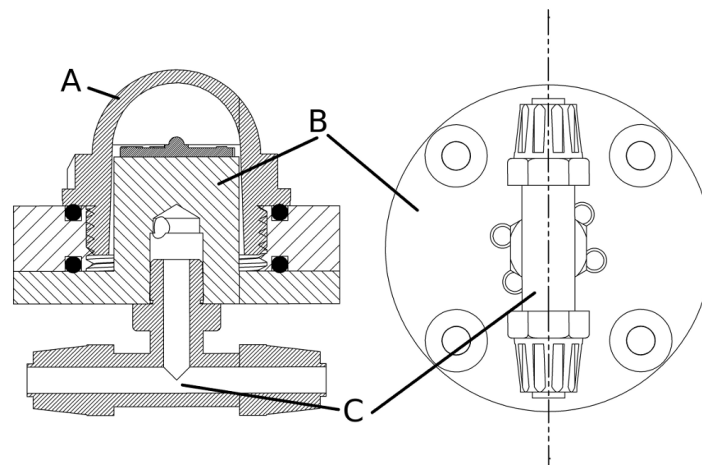


Abbildung 3.8: Konstruktionszeichnung des einzelnen LED-Gehäuses. A: Flüssigkeitsschauglas. B: Kühlkörper aus Aluminium. C: Schlauchverschraubung POM, Blau.

baren Halterung versehen wurden, um so einen variabel einstellbaren Abstrahlwinkel zu ermöglichen. Das LED-Gehäuse besitzt folgende erwähnenswerte Bauteile, siehe hierzu auch die Konstruktionszeichnung in Abbildung 3.8:

- A) Flüssigkeitsschauglas, dessen Form einer Halbkugel einen maximalen Abstrahlwinkel der LEDs ermöglicht.
- B) Kühlkörper aus Aluminium AW2007, der durch eine Kunststoffbeschichtung gegenüber Wasserkontakt isoliert wurde. Durch den Kontakt mit Wasser wird eine sehr starke Wärmeableitung erreicht, was den Einsatz leistungsstarker LEDs ohne die Gefahr einer Überhitzung ermöglicht.
- C) Schlauchverschraubungen der Sorte POM, Blau, die aus der Pneumatik stammen, um eine wasserdichte Kabeleinführung zu gewährleisten. Darüber hinaus wurden für die Stromversorgung der LEDs ebenfalls klar-sichtige PVC-Schläuche verwendet, die aufgrund ihrer Stabilität eine wasserdichte Verschraubung mit den POMs ermöglichen. Silikon-schläuche zum Beispiel sind dafür zu weich.

Abbildung 3.9 zeigt 2 Beispielbilder für das fertige Beleuchtungssystem.

Welche Ausrichtung der Lichtquellen für diese Arbeit eine annähernd homogene Ausleuchtung erzielt, wird in Kapitel 3.3.2 genauer untersucht und die Ergebnisse in Kapitel 3.4.2 präsentiert.

Ob die Verwendung von farbigem Licht Vorteile bei der Erkennung von Steinbutten bringt, wird in Kapitel 3.3.1 untersucht und die Ergebnisse in Kapitel 3.4.1 präsentiert. Dementsprechend erfolgt dann die Wahl der LEDs für das Beleuchtungssystem.

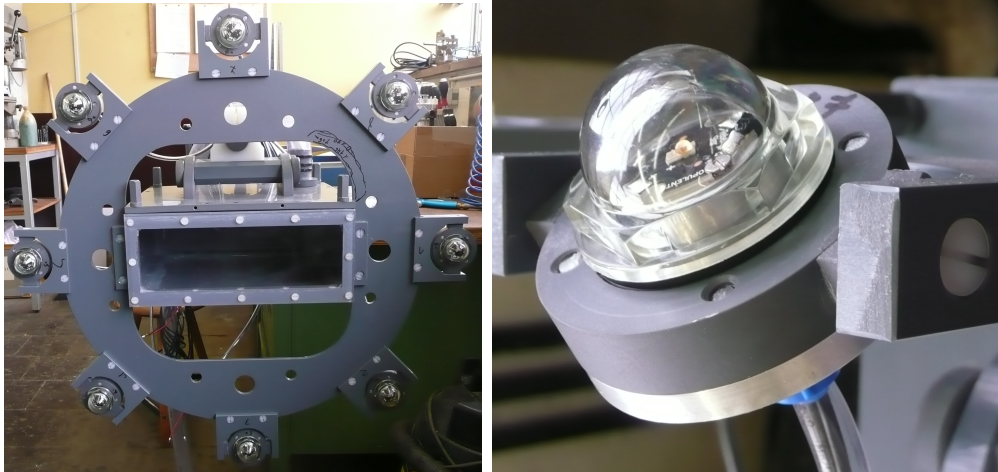


Abbildung 3.9: Links: Kameragehäuse in der Mitte. Außen LED-Halterung mit 8 LED-Gehäusen. Rechts: Einzelnes LED-Gehäuse mit einstellbarem Winkel.

3.2.2.3 Stromversorgung

Für die Stromversorgung der LEDs und der Kameras wurde ein 3-kanaliger LED- und Kamera-Controller entwickelt. Dieser ermöglicht die präzise Stromsteuerung von bis zu 3 LED-Ausgängen und den Betrieb von bis zu 3 Kameras der Baureihe GigE uEye SE der Firma IDS. Das Gerät besteht im Wesentlichen aus einem Basis-Netzgerät für die grundlegende Spannungs- und Stromversorgung sowie einer für jeden LED-Kanal nachgeschalteten, ultraschnellen Präzisions-Stromregelungseinheit als LED-Treiberstufe. Sie erlauben Ein- und Ausschaltzeiten im Bereich von circa 25 ns, was 40 KHz entspricht, bedingt durch die Anstiegs- und Abfallzeiten ($<10/\lt;15$ ns) der elektronischen Ansteuerung bis DC = Dauerstrom. Die 3 LED-Ausgänge lassen sich stufenlos im Bereich 0 bis 1.000 mA einstellen.

Ein 12 VDC Netzgerät zum Betrieb der Kameras und ein Multimeter für Prüf- und Abgleicharbeiten sind ebenfalls integriert. Zum Schutz gegen Umwelteinflüsse sind die Geräte in ein Edelstahlgehäuse eingebaut, wie in Abbildung Kapitel 3.10 zu sehen ist. Für sämtliche Ein-/Ausgänge wurden spritzwassergeschützte Stecker verwendet.

Die Kanäle 1 und 2 werden mit den zugehörigen Potentiometern des LED- und Kamera- Controllers auf jeweils 3 Stromwerte (Low, Medium, High) vor-

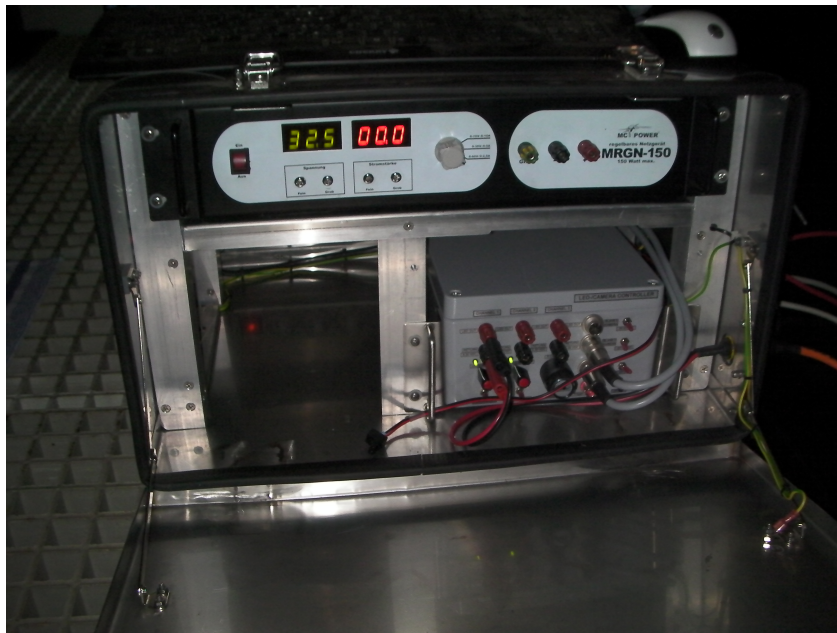


Abbildung 3.10: Netzgerät und Präzisions-Stromregelungseinheit im Edelstahlgehäuse.

eingestellt und können dann über Stufenschalter abgerufen werden. In den Schaltstufen für geringsten und mittleren Strom lassen sich die Ausgänge durch Kippschalter zwischen den Varianten *Ein* (Dauerbetrieb), *Aus* und *Impulsbetrieb* (nur getriggert) umschalten. Die Schaltstufe für den höchsten Strom lässt ausschließlich den getriggerten Impulsbetrieb zu, um angeschlossene LEDs nicht zu überlasten. Die Ansteuerung für den Impulsbetrieb erfolgt wahlweise intern über die angeschlossene Master-Kamera, oder mittels Umschalter über einen externen Triggereingang (BNC-Buchse).

Der Kanal 3 ist voll variabel und wird durch ein Digitalpotentiometer auf den erforderlichen Stromwert eingestellt. Im Gegensatz zu den Kanälen 1 und 2 kann dieser mit dem Ein/Aus-Impulsbetrieb Kippschalter in allen Funktionen betrieben werden. Hierbei muss der Anwender dafür Sorge tragen, dass die LEDs nicht überlastet werden.

3 Kamera-Eingänge und Kamera-Ausgänge dienen dem Anschluss einer Master- und bis zu 2 Slave-Kameras. Die Master-Kamera generiert bei entsprechender Softwareeinstellung einen digitalen Ausgangsimpuls. Dieser dient als interner Trigger zur Ansteuerung der entsprechend zugeschalteten LED-Kanäle und Slave-Kameras.

3.2.3 Systemaufbau am Versuchsstandort

Der Aufbau des Stereokamerasystems bei der GMA in Büsum soll in diesem Abschnitt beschrieben werden. Die Abbildung 3.11 zeigt im linken Bild den Aufbau des Stereokamerasystems über einem Fischbecken. Im rechten Bild befindet sich die Recheneinheit im Serverraum. Der Datentransfer erfolgt dabei über das Firmennetzwerk der GMA. Zusätzlich wurden das Netzwerk und die Recheneinheit so konfiguriert, dass eine ferngesteuerte Nutzung per Remote-Funktion aus der Universität Kiel möglich ist.

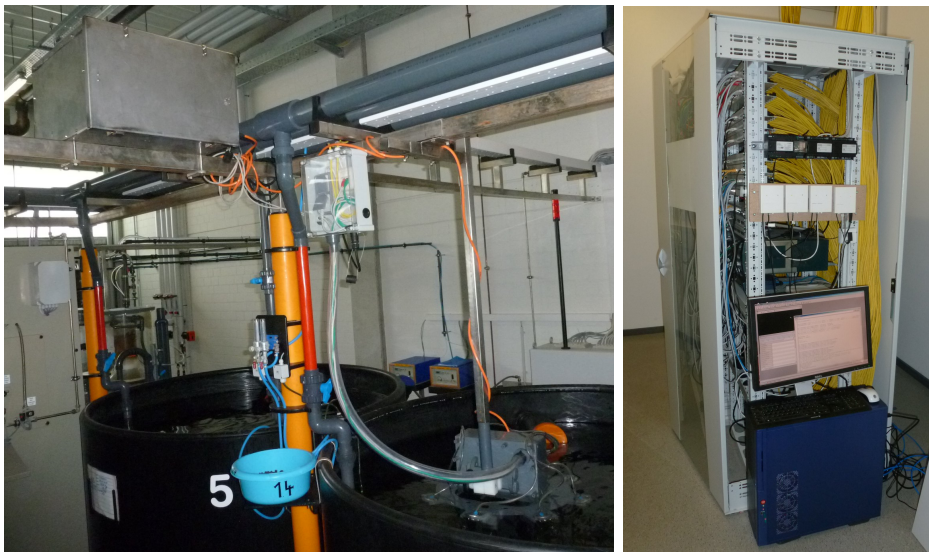


Abbildung 3.11: Links: Aufgebautes Stereokamerasystem über Fischbecken. Rechts: Recheneinheit im Serverraum verbunden mit dem Stereokamerasystem über das Firmennetzwerk.

Das Edelstahlgehäuse mit der Präzisions-Stromregelungseinheit und dem Netzgerät wird auf einem Querträger über dem Fischbecken fest verschraubt, siehe hierzu Abbildung 3.12 im linken Bild. Von hier aus gehen die Stromleitungen für LEDs und Kameras zu den entsprechenden Komponenten. Im Falle der Kameras erfolgt dies über einen Verteilerkasten, der am Fischbecken montiert ist. In diesem befindet sich außerdem der Anschluss für das Netzgerät an das Stromnetz sowie die Anbindung an das Firmennetzwerk für die Datenübertragung, wie in Abbildung 3.12 im rechten Bild gut zu sehen ist.

Die Stereokameras im wasserdichten Gehäuse werden mit einer extra konstruierten Halterung an einer Trägerstange über dem Fischbecken befestigt, wie in Abbildung 3.13 zu sehen ist. Aufgrund der Konstruktion, bestehend aus Führungsschienen, Stangen und Feststellschrauben, ist so eine variable Positionierung der Stereokameras möglich.

Die Stereokameras in ihrem wasserdichten Gehäuse werden dann für ihren Einsatz soweit in das Fischbecken getaucht, bis sich die Sichtscheibe und sämtliche LEDs unterhalb der Wasseroberfläche befinden, wie in Abbildung 3.14 zu sehen ist.

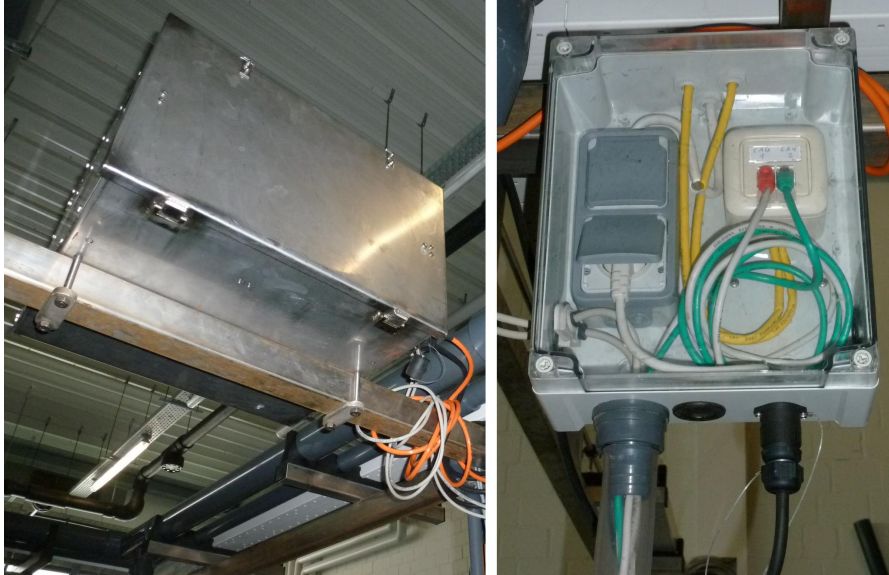


Abbildung 3.12: Links: Fest montiertes Edelstahlgehäuse mit abgehenden Stromleitungen. Rechts: Verteilerkasten mit Anschluss an Stromnetz und Firmennetzwerk.



Abbildung 3.13: Halterung für Kameragehäuse und Beleuchtungssystem montiert an Trägerstange über Fischbecken.

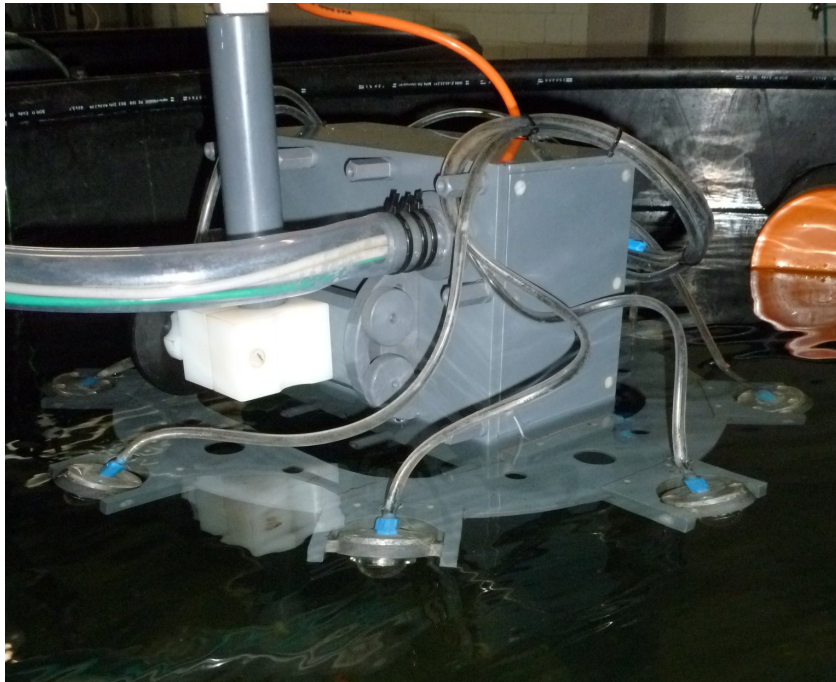


Abbildung 3.14: Stereokamerasystem im Unterwasser-Einsatz bei der GMA in Büsum.

3.3 Methoden: Parameter des Beleuchtungssystems und Unterwasser-Kalibrierung

Für die Entwicklung des Beleuchtungssystems mussten einige Vorüberlegungen und Versuche durchgeführt werden. Die zu klärenden Fragen betrafen die zu verwendende Farbe für die Beleuchtung als auch für das Material des Beckens sowie das Arrangement der Beleuchtung. Dabei musste bedacht werden, dass aufgrund der Lichtempfindlichkeit der Steinbutte die verwendete Lichtmenge nicht beliebig hoch gewählt werden sollte, und zum anderen, dass vorhandene Schwebeteilchen nicht zu stark beleuchtet werden sollten. Die Farbwahl wurde mittels des in Kapitel 3.3.1 beschriebenen Versuchs durchgeführt. Das entwickelte Arrangement für die Beleuchtung sowie der durchgeführte Test zur Ausleuchtung des Sichtbereiches werden in Kapitel 3.3.2 vorgestellt. Die Ergebnisse der Versuche werden in Kapitel 3.4.1 und Kapitel 3.4.2 präsentiert.

Die durchgeführte Unterwasser-Kalibrierung des Stereokamerasystems wird in Kapitel 3.3.3 genauer erläutert. Die Evaluation der resultierenden Messgenauigkeit erfolgt in Kapitel 3.4.3.

3.3.1 Versuchsaufbau zur Farbwahl des Beleuchtungssystems

Das Ziel des ersten Versuches war herauszufinden, welche Art der Beleuchtung die besten Resultate bei der Bildaufnahme für die folgende Erfassung von Steinbutten erwarten läßt. Dafür werden die folgenden Anforderungen an die Bilder gestellt, um für die Bildverarbeitung eine gute Ausgangssituation zu erlangen:

1. Kontrast: Ein möglichst hoher Kontrast innerhalb des Bildes bedeutet eine bessere Detektion von auffälligen Strukturen, wie zum Beispiel den Umrissen der Steinbutte.
2. Helligkeit: Die Bilder dürfen nicht zu dunkel sein. Zum einen reduziert dies den möglichen Kontrast, zum anderen wird durch eine mögliche elektronische Korrektur der Helligkeit zumeist das Rauschen im Bild verstärkt. Helligkeit läßt sich durch längere Belichtungszeiten erreichen.

Zur Ermittlung der Parameter Kontrast und Helligkeit wurde der folgende Versuch durchgeführt. Es wurden 5 LEDs mit vergleichbarer Leistungsaufnahme in den Farben weiß, gelb, blau, grün sowie rot ausgewählt, um somit das Lichtspektrum ansatzweise abzudecken. Tabelle 3.1 zeigt für die verwendeten LEDs die Farbtemperatur (weiß) beziehungsweise die dominante Wellenlänge λ_D sowie deren Minimum- und Maximum-Werte. Siehe hierzu auch die Datenblätter [90] und [89].

Tabelle 3.1: Farbtemperatur für neutral weiße LED beziehungsweise dominante Wellenlänge λ_D für andere LEDs inklusive Minimal- und Maximal-Wert.

Farbe	Min.	Farbtemperatur bzw. λ_D	Max.
Weiß	3500 K	4100 K	4500 K
Gelb	584,5 nm	590 nm	597 nm
Blau	460 nm	470 nm	490 nm
Grün	520 nm	530 nm	550 nm
Rot	620,5 nm	627 nm	645 nm

Des Weiteren wurden versuchsweise eine UV-LED sowie eine IR-LED mit unbefriedigenden Ergebnissen ausprobiert, so dass diese LEDs nicht in die weitere Betrachtung mit einbezogen wurden. Die farbigen LEDs wurden kreisförmig und mit geringem Abstand auf einer angepassten Träger-Platte aus Aluminium montiert, die gleichzeitig als Kühlkörper für die LEDs dient. Abschließend wurden die LEDs mit einer Plexiglasscheibe wasserdicht versiegelt,

siehe hierzu Abbildung 3.15. Durch den geringen Abstand zwischen den LEDs sollte erreicht werden, dass die Szenerie von den einzelnen LEDs möglichst vergleichbar ausgeleuchtet wird. Bei einem zu großen Abstand zwischen den einzelnen LEDs wären die Bereiche der intensivsten Ausleuchtung an unterschiedlichen Positionen und könnten so die Auswertung verfälschen.

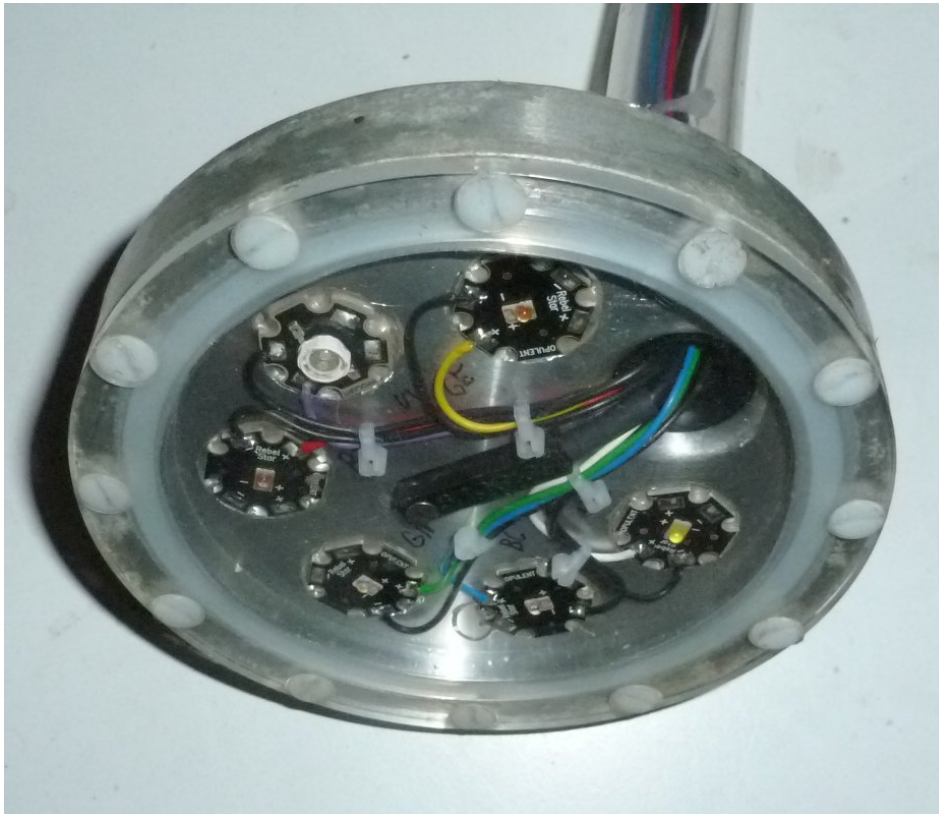


Abbildung 3.15: LED-Träger für Farblicht-Versuch. 6 LEDs kreisförmig auf Basis-Platte aus Aluminium, die als Kühlkörper dient. Scheibe aus Plexiglas.

Da neben den Steinbutten auch die Umgebung Einfluss auf die Aufnahmen ausüben kann, wurden austauschbare Bodenplatten verwendet, um so die unterschiedlichen Varianten an Beckenmaterialien zu simulieren. Die Platten bestehen aus PVC und wurden entsprechend den in der Aquakultur typischen Farben der Becken gewählt: weiß, grau, schwarz und grün.

Aufgrund des breiten Spektrums von hellen bis dunklen Musterungen, die Steinbutte aufweisen können, wurden für diesen Versuch Steinbutte mit unterschiedlichen Musterungen verwendet, ein Paar mit heller Musterung und ein Paar mit dunkler Musterung, zum Vergleich siehe Abbildung 3.16. Für diesen Versuch wurden tote, noch unverwaste, Steinbutte verwendet, um eine gleich bleibende und vergleichbare Position zu ermöglichen.

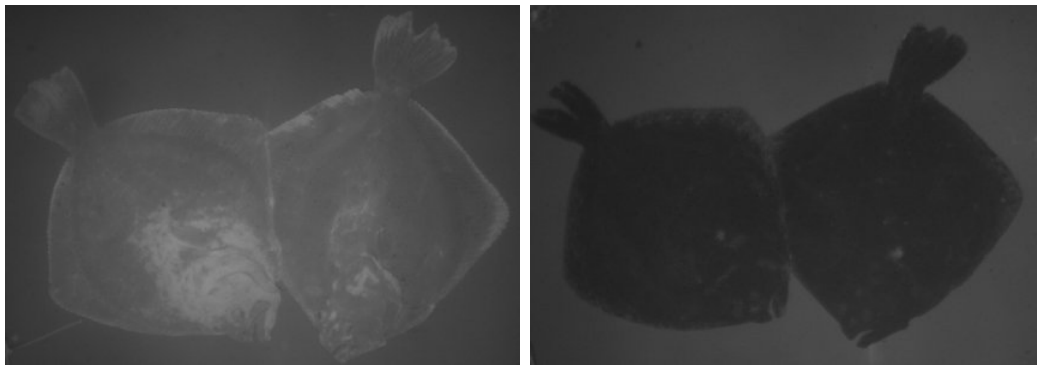


Abbildung 3.16: Die beim Versuch verwendeten Steinbutte. Links: Paar mit heller Musterung. Rechts: Paar mit dunkler Musterung.

Der Versuch wurde in einem Becken aus schwarzem PVC mit einem Bodendurchmesser von circa 1 m durchgeföhrt. Der Wasserpegel wurde vergleichbar zu den realen Bedingungen bei der GMA in Büsum auf 80 cm festgelegt. Der Boden wurde mit jeweils einer Bodenplatte ausgelegt und eines der Steinbuttpaare wurde in einer vorher bestimmten Position auf die Bodenplatte gelegt. Die LEDs wurden am Kameragehäuse befestigt und zusammen in das Becken getaucht, bis die Sichtscheiben kurz unter der Wasseroberfläche waren.

Für jede Farbvariante wurden mehrere Aufnahmen mit 2 unterschiedlichen Belichtungszeiten durchgeföhrt. Die erste Belichtungszeit betrug 10 ms und alle weiteren automatischen Optionen zur Bildverbesserung wurden deaktiviert. Die zweite Belichtungszeit wurde automatisch von der Kamerasoftware im automatischen Belichtungsmodus bestimmt, wo auch die sonstigen Optionen zur Bildverbesserung aktiviert waren. Dieser Modus stellt die Kamera derart ein, dass ein mittlerer Grauwert von 127 über das gesamte Bild erreicht wird.

Für die Aufnahmen wurde das vom Kamerahersteller bereitgestellte Softwarepaket² verwendet. Diese Prozedur wurde für jede Kombination der 5 Beleuchtungsfarben, 4 Bodenplatten und 2 Steinbuttmusterungen wiederholt. Das Ergebnis sind 40 Kombinationen, von denen jeweils mit 2 Belichtungszeiten Aufnahmen durchgeföhrt wurden. Zur Veranschaulichung der verschiedenen Farben wurden auch Aufnahmen mit einer Farbkamera vorgenommen, siehe hierzu Abbildung 3.17.

²<http://www.ids-imaging.de> - uEye Softwarepaket

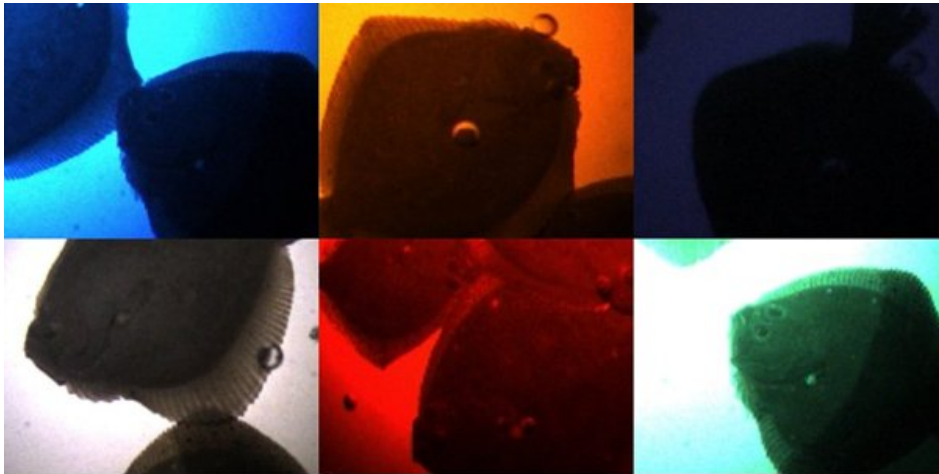


Abbildung 3.17: Aufnahmen von Steinbutten bei farbigem Licht. Oben links: Blau. Oben mitte: Gelb. Oben rechts: Ultraviolett. Unten links: Weiß. Unten mitte: Rot. Unten rechts: Grün.

3.3.2 Versuchsaufbau zur Ausrichtung des Beleuchtungssystems

Der zweite Versuch zum Beleuchtungssystem sollte klären, welche Grundeinstellung der drehbar verbauten LEDs einer möglichst gleichmäßigen Ausleuchtung des Sichtfeldes dient. Dabei ist es wichtig, die *Vignettierung* zu beachten. Vignettierung ist ein Effekt, der vor allem Objektiv mit großem FoV betrifft und die Helligkeitsabnahme des Bildes zum Bildrand hin beschreibt.

Für das verwendete Pentax-Objektiv H416 (KP) ist der Verlauf der Vignettierung in Abbildung 3.20 dargestellt. Ziel dieses Versuches war es, für eine möglichst homogene Ausleuchtung der Szene zu sorgen und dabei die Vignettierung zu einem gewissen Maß auszugleichen. Dabei sollte die Differenz zwischen den Grauwerten im Bildzentrum und dem Randbereich relativ gering sein. Wie bereits in Kapitel 3.2.1 erwähnt, wurde für eine annähernd homogene Ausleuchtung eine maximale Helligkeitsdifferenz von 50 % zwischen dem hellsten und dunkelsten Bereich des Bildes bei einem homogenen Objekt festgelegt. Der mittlere Bildbereich darf wiederum nicht zu schwach ausgeleuchtet werden, da in diesem Bereich die höchste Genauigkeit erreicht werden kann.

Dafür wurde die resultierende Beleuchtung des Beleuchtungsrings auf einen schwarzen PVC-Hintergrund aus einer Distanz von 60 cm aufgenommen und ausgewertet. Die getesteten Winkel für die LEDs reichten von 20° nach innen gedreht bis 40° nach außen gedreht. Dabei wurden Schritte von 20° durchgeführt. Beispiele für die entstandenen Beleuchtungen sind in Abbildung 3.18 zu sehen.

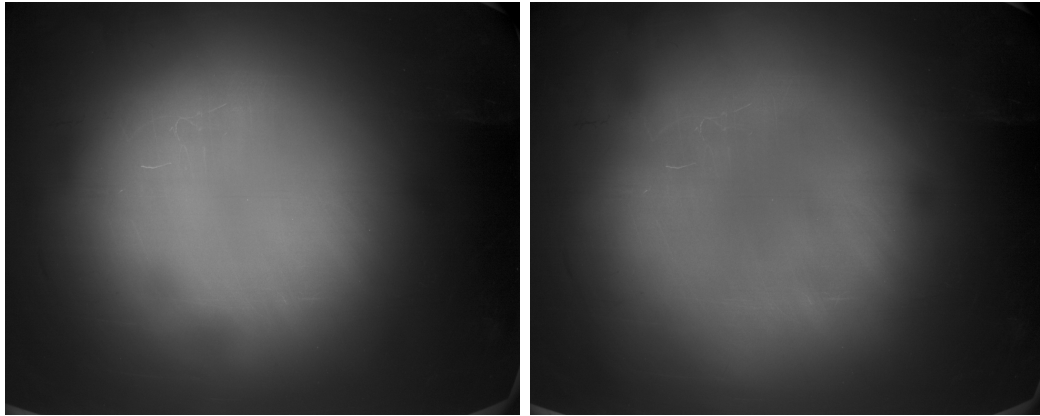


Abbildung 3.18: Aufnahmen bei unterschiedlichen Einstellungswinkeln der LEDs. Links: 20° nach innen gedreht. Rechts: senkrecht ausgerichtet.

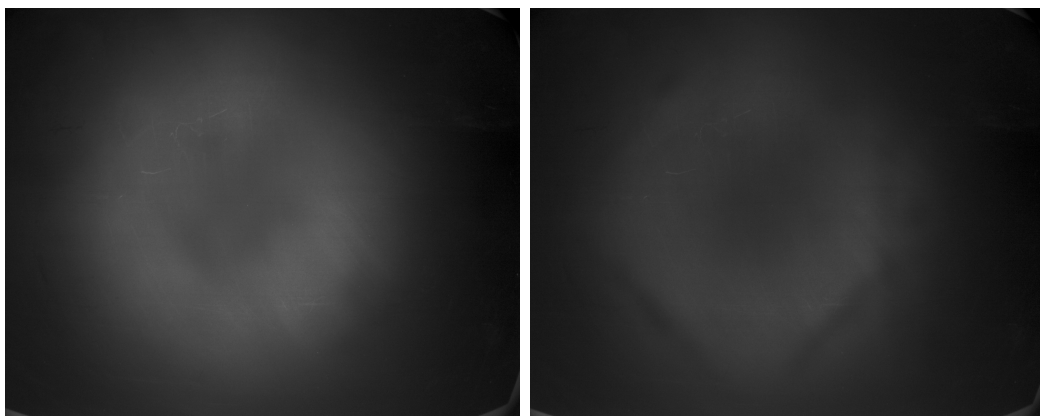


Abbildung 3.19: Aufnahmen bei unterschiedlichen Einstellungswinkeln der LEDs. Links: 20° nach außen gedreht. Rechts: 40° nach außen gedreht.

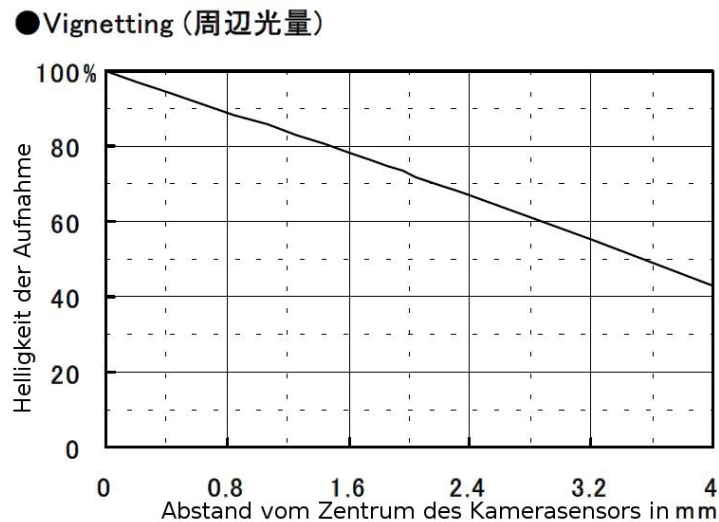


Abbildung 3.20: Vignettierung des Pentax-Objektivs H416 (KP) aus [52].

3.3.3 Kalibrierung des Stereokamerasystems für Unterwasser-Einsatz

Die Kalibrierung von Stereokamerasystemen ist inzwischen ein standardisiertes Verfahren, für das ein Objekt mit bekannten Abmessungen und Markierungen eingesetzt wird. In den meisten Fällen handelt es sich dabei um Schachbrettmuster, wo die Ecken zwischen den einzelnen Feldern als Markierungen dienen. Von diesem Objekt wird dann eine Vielzahl von Aufnahmen aus unterschiedlichen Perspektiven und Positionen erstellt. Wie bereits im Kapitel 2.2.1.2 kurz erwähnt, können anhand der bekannten Abstände zwischen den Markierungen des Objekts und vorhandener Verzerrungen in den Bildern sowohl die intrinsischen Kameraparameter als auch die extrinsischen Kameraparameter des Stereosystems berechnet werden. Für weiterreichende Informationen seien hier die Arbeiten von Tsai [142] [143] empfohlen.

Für die Kalibrierung eines Stereokamerasystems existieren mehrere, frei erhältliche Softwaretools wie zum Beispiel die *Camera Calibration Toolbox for Matlab* [18] oder die *MIP - MultiCameraCalibration* [119]. In dieser Arbeit wird auf die MIP - MultiCameraCalibration zurückgegriffen. Für weitere Informationen hierzu siehe [120].

Wie bereits in Kapitel 3.2.1 dargestellt, verändert die Brechung des Lichts beim Übergang in ein anderes Medium den Öffnungswinkel der Objektive und somit auch die intrinsischen Kameraparameter des Stereosystems, vor allem die Brennweite. Da sich die Kamera in einem wasserdichten Gehäuse befinden, existiert immer ein Wechsel zwischen Luft, wo sich die Kamera befindet, und Wasser, getrennt durch eine Scheibe. Es muss geklärt werden, inwiefern

die Brechung des Lichts Einfluss auf eine Kalibrierung des Kamerasystems ausübt. Dabei ist zu beachten, dass die Sichtstrahlen der Kameras zwar durch die Lichtbrechung abgelenkt werden, es im weiteren Verlauf jedoch zu keiner Krümmung der Strahlen kommt und die für die Kalibrierung zugrunde liegende Epipolareometrie aus Kapitel 2.2.1.3 weiterhin gültig bleibt.

Anhand der Aussage von Treibitz in [141] kann überprüft werden, ob ein auf Standardverfahren basierendes Softwaretool für die Unterwasser-Kalibrierung eingesetzt werden kann. Demnach können Standardverfahren mit ausreichender Genauigkeit unterwasser verwendet werden, sofern das Verhältnis zwischen den Brennweiten einer Luft- und einer Unterwasser-Kalibrierung circa 1:1,33 beträgt. Ob dieser Faktor erreicht wird, hängt zum großen Teil von der Hardwarekonfiguration und dem Abstand zwischen Objektiv und der Sichtscheibe ab. Aufgrund dieses Vorwissens wurden in dieser Arbeit die Kameras derart im Kameragehäuse montiert, dass ein minimal möglicher Abstand vorliegt.

Für die Überprüfung wurde jeweils eine Luft- und eine Unterwasser-Kalibrierung mit dem Softwaretool durchgeführt und das Verhältnis zwischen den ermittelten Brennweiten berechnet. Das Ergebnis war mit einem Verhältnis von 1:1,31 sehr nah am Optimum und bestätigte somit den möglichen Einsatz der MIP - MultiCameraCalibration.

Für die Kalibrierung des Stereokamerasystems wurde als Kalibriermuster ein Schachbrett der Größe 9×7 verwendet, wobei jedes Feld eine Kantenlänge von 24 mm aufwies. Von diesem Kalibriermuster wurden Stereoaufnahmen angefertigt, für die folgende Punkte beachtet werden mussten:

- Das Kalibriermuster muss in beiden Stereobildern vollständig abgebildet sein.
- Das gesamte Sichtfeld der Kameras soll möglichst vollständig durch die Aufnahmen abgedeckt sein, um die Verzerrungsparameter modellieren zu können.
- Die Aufnahmen sollen in unterschiedlichen Distanzen zu den Kameras durchgeführt werden.
- Die Ausrichtung des Kalibriermusters zu den Kameras soll verschiedene Winkel aufweisen.
- Die extrahierten Punkte müssen sich in jedem Bild entsprechen. Das heißt insbesondere bei automatischen Detektoren, dass wegen des symmetrischen Musters die Ausrichtung der Oberkante des Kalibriermusters gleich bleiben muss.

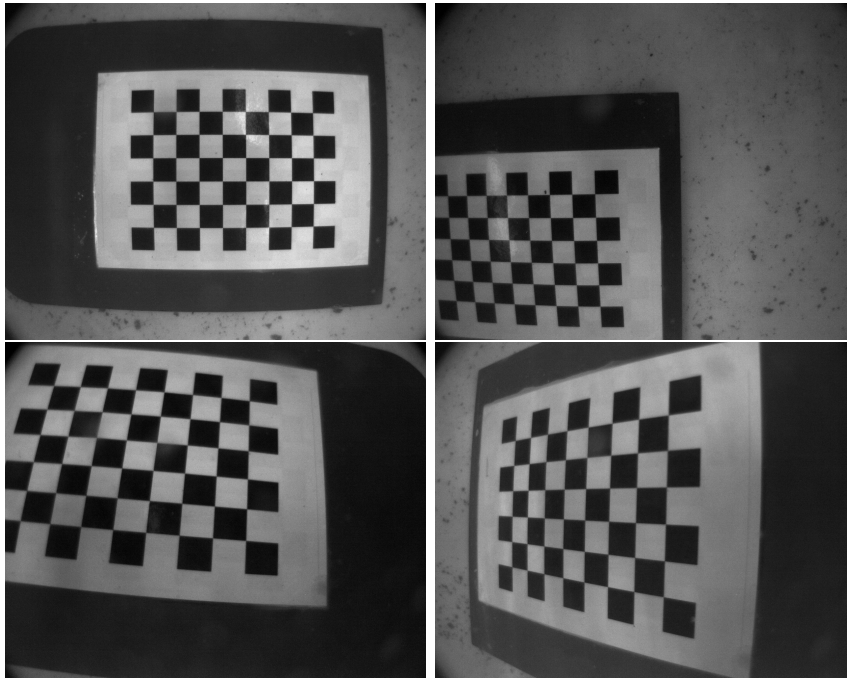


Abbildung 3.21: 4 Beispiele für Aufnahmen des Kalibrieramusters mit unterschiedlichen Positionierungen und Ausrichtungen.

4 Beispiele für solche Kalibrieraufnahmen werden in Abbildung 3.21 gezeigt.

Anhand der Markierungen in den Kalibrieraufnahmen wurden dann vom Softwaretool die Kameraparameter, basierend auf den in Kapitel 2.2 vorgestellten Grundlagen, berechnet. Dabei wird ein Polynom approximiert, dass die real auftretenden Verzerrungen in allen Bildbereichen möglichst genau widerspiegeln soll.

Für die Berechnung eines solchen nichtlinearen Modells werden im Allgemeinen 4 Stützstellen als ausreichend empfohlen. Damit diese möglichst genau berechnet werden können, ist eine große Anzahl an Kalibrieraufnahmen notwendig. Hierfür wird als Grundlage eine fünffache Redundanz zu den unbestimmten Parametern im zu berechnenden Gleichungssystem empfohlen [77]. Dies bedeutet für diesen Kalibrierprozess, dass aufgrund der 11 Unbekannten für die intrinsischen und extrinsischen Parameter ein Datensatz von mindestens 55 möglichst unterschiedlichen Kalibrieraufnahmen ausreichen sollten [17]. Für die Approximation ist hierbei auch der Randbereich wichtig, da bei der Verwendung von kleinen Brennweiten hier die größten Verzerrungen auftreten, die es zu modellieren gilt [144]. Anderenfalls gilt die Kalibrierung nur für das während der Kalibrierphase abgedeckte Kamera- Sichtfeld. Das Ergeb-

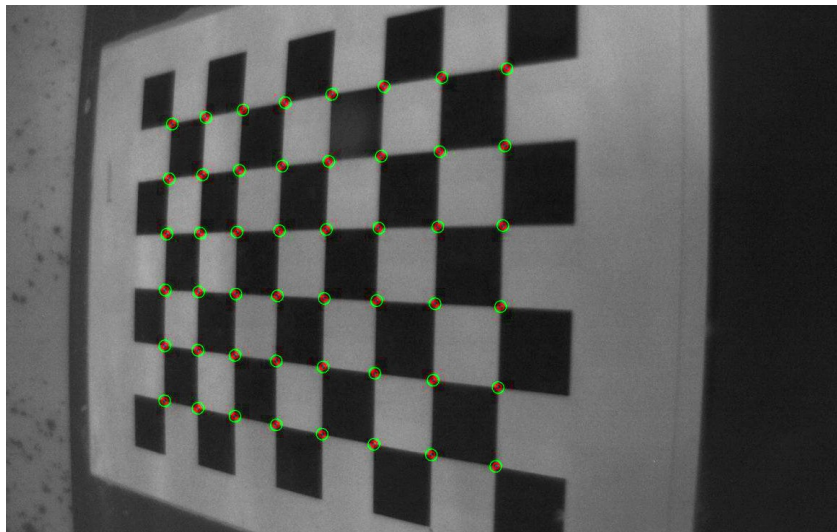


Abbildung 3.22: Ein Beispiel für die Rückprojektion. Die Position der gemessenen Punkte des Kalibrieremusters werden durch grüne Kreise und die rückprojizierten Punkte durch rote Kreise dargestellt.

nis der Kalibrierung des Stereokamerasystems enthält folgende Parameter:

- Brennweite f beider Kameras.
- Hauptpunkt x_0, y_0 beider Kameras.
- Verzerrungskoeffizienten beider Kameras.
- Translation t zwischen den Kameras.
- Rotation R zwischen den Kameras.

Ein Qualitätsmaß für die Kalibrierung ist der sogenannte *Mean Reprojection Error*. Dabei ist der Reprojection Error die Ungenauigkeit zwischen dem berechneten Kameramodell und der Realität. Dies wird ermittelt durch die Rückprojektion von Punkten des Kalibrieremusters zurück ins Bild. Die Beträge der Distanzen zwischen den gemessenen und rückprojizierten Messpunkten werden gemittelt und als Mean Reprojection Error zurückgegeben. Umso geringer dieser Wert ist, desto genauer ist die Kalibrierung ausgefallen. Im Idealfall wäre der Mean Reprojection Error $e = 0$ Pixel. Da es sich aber immer um Approximationen handelt, kann dies nicht erreicht werden. Ein Fehler von 0,5 Pixel beschreibt im Allgemeinen eine genaue Kalibrierung. Ein Beispiel für die Rückprojektion der Kalibrierenaufnahmen in [Abbildung 3.21](#) wird in [Abbildung 3.22](#) gezeigt.

3.4 Ergebnisse: Parameter des Beleuchtungssystems und Unterwasser-Kalibrierung

3.4.1 Ergebnisse zur Farbwahl des Beleuchtungssystems

Für die Auswertung jeder der 80 Kombinationen von Lichtfarbe, Untergrundfarbe, Fischmusterung und Belichtungszeit wurden für die Bildbereiche der beiden Steinbutte sowie des Untergrundes die Mittelwerte sowie die Standardabweichungen ermittelt.

Die Tabelle 3.2 zeigt für alle Kombinationen die Belichtungszeit, die benötigt wird, um im automatischen Belichtungsmodus einen mittleren Grauwert von 127 über das gesamte Bild zu erreichen. Reicht die maximal mögliche Belichtungszeit von 99,8 ms hierfür nicht aus, wird der Wert der in dem Fall aktivierten Verstärkung ebenfalls angegeben. Die Ergebnisse zeigen für alle Varianten von Fischmuster und Bodenplatte eine erheblich kürzere Belichtungszeit für weißes Licht im Vergleich zu allen anderen Farbtönen. Das blaue und grüne Licht weisen die nächst kürzeren Belichtungszeiten auf, in Abhängigkeit von der Bodenplatte. Rotes Licht hat im Vergleich zu blauem und grünem Licht etwas längere Belichtungszeiten, die näherungsweise bei einem Faktor 2 gegenüber dem weißen Licht liegen. Das gelbe Licht weist so lange Belichtungszeiten auf, dass bei jeder Kombination eine zusätzliche Bildverstärkung benötigt wurde.

Tabelle 3.2: Belichtungszeiten der verschiedenen Farbkombinationen in ms im automatischen Belichtungsmodus. Bei maximaler Belichtungszeit von 99,8 ms wird die zusätzlich benötigte Verstärkung angegeben, um einen mittleren Grauwert von 127 zu erreichen.

Fisch	Boden	Weiß	Grün	Blau	Rot	Gelb
Dunkel	Weiß	9,80	18,00	15,70	19,60	99,8/1,3
Dunkel	Grau	44,10	75,70	64,30	92,20	99,8/6
Dunkel	Schwarz	90,30	99,8/1,67	99,8/1,5	99,8/1,9	99,8/12,7
Dunkel	Grün	41,60	52,40	78,00	99,8/1,4	99,8/8
Hell	Weiß	9,50	17,60	17,50	18,90	99,8/1,2
Hell	Grau	27,90	51,70	46,80	56,80	99,8/3,6
Hell	Schwarz	40,90	79,20	72,70	82,80	99,8/5,4
Hell	Grün	29,40	44,70	53,60	79,60	99,8/4,5

Vergleichbar sehen die Ergebnisse in der Abbildung 3.23 aus. Hier werden die mittleren Grauwerte innerhalb eines Fisches für die unterschiedlichen

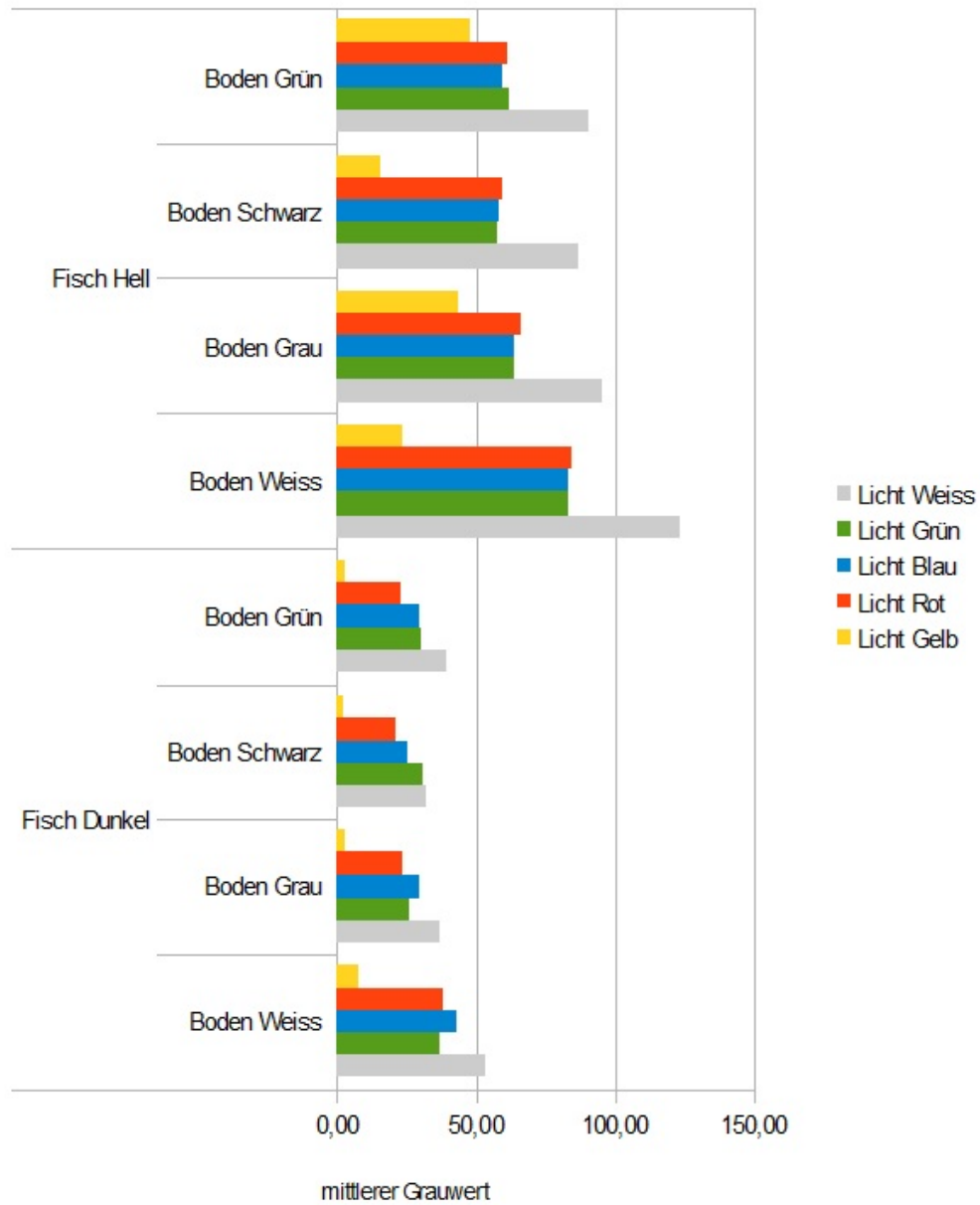


Abbildung 3.23: Mittlere Grauwerte eines Steinbutts bei einer Belichtungszeit von 10 ms. Ergebnisse sind die Mittelwerte aus 3 Aufnahmen je Variante.

Kombinationen bei fixen 10 ms Belichtungszeit dargestellt. Dafür wird innerhalb des Fisches ein abgestecktes Areal ausgewertet und die Grauwerte der einzelnen Pixel gemittelt. Bei den Aufnahmen der hell gemusterten Steinbutte werden so insgesamt zwischen 37.000 und 42.000 Messwerte pro Bodenfarbe in die Berechnung aufgenommen. Bei den Aufnahmen der kleineren dunkel gemusterten Steinbutte werden pro Bodenfarbe zwischen 23.000 und 26.000 Messwerte verarbeitet. Für jede Variante werden 3 Aufnahmen ausgewertet und deren Ergebnisse erneut gemittelt. Zwischen den einzelnen Aufnahmen treten nur sehr geringe Unterschiede, bedingt durch das Kamerarauschen, auf.

Weißes Licht weist hier ebenfalls die besten Resultate in Form von hohen mittleren Grauwerten auf, gefolgt von grün, blau und dann rot. Gelbes Licht resultiert in eindeutig niedrigeren mittleren Grauwerten. Obwohl es sich bei den Ergebnissen nur um die mittleren Grauwerte innerhalb eines Fisches handelt, variieren die Ergebnisse mit wechselnder Bodenfarbe. Dies liegt darin begründet, dass aufgrund eines helleren Hintergrundes das gesamte Bild stärker ausgeleuchtet wird und somit auch die Fische einen höheren mittleren Grauwert erreichen.

Die Ergebnisse bei gelbem Licht sind für keine Variante der Belichtungszeit zufriedenstellend. Bei automatisch gewählter Belichtungszeit verursacht die zusätzlich benötigte Verstärkung unerwünschtes Rauschen, während bei 10 ms Belichtungszeit die Ergebnisse in mitunter derart niedrigen Grauwertebereichen liegen, dass schon kleine Schwankungen starke Auswirkungen auf das Ergebnis haben. Dadurch entstehen für die folgenden Auswertungen scheinbar gute Resultate für gelbes Licht, die nicht der praktischen Realität entsprechen.

In den Abbildungen 3.24 und 3.25 sind die relativen Kontraste zwischen der Bodenplatte sowie den beiden Steinbutten dargestellt, in Abbildung 3.24 bei 10 ms Belichtungszeit und in Abbildung 3.25 bei automatisch festgelegter Belichtungszeit.

Dafür wird jeweils ein mittlerer Grauwert für einen Steinbutt sowie für die sie umgebende Hintergrundplatte errechnet, indem die Grauwerte der Pixel der entsprechend abgesteckten Areale verrechnet werden. Bei den hell gemusterten Steinbutten entstehen so zwischen 73.000 und 82.000 Messwerte für die Steinbutte und zwischen 94.000 und 102.000 Messwerte für die Bodenplatte je nach Bodenfarbe. Bei den dunkel gemusterten Steinbutten sind es zwischen 47.000 und 52.000 Messwerte für die Steinbutte und zwischen 71.000 und 78.000 Messwerte für die Bodenplatte, ebenfalls abhängig von der Bodenfarbe.

Für den relativen Kontrast wird die folgende Formel verwendet:

$$K = \left| \frac{x_1 - x_2}{x_1 + x_2} \right| \quad (3.4)$$

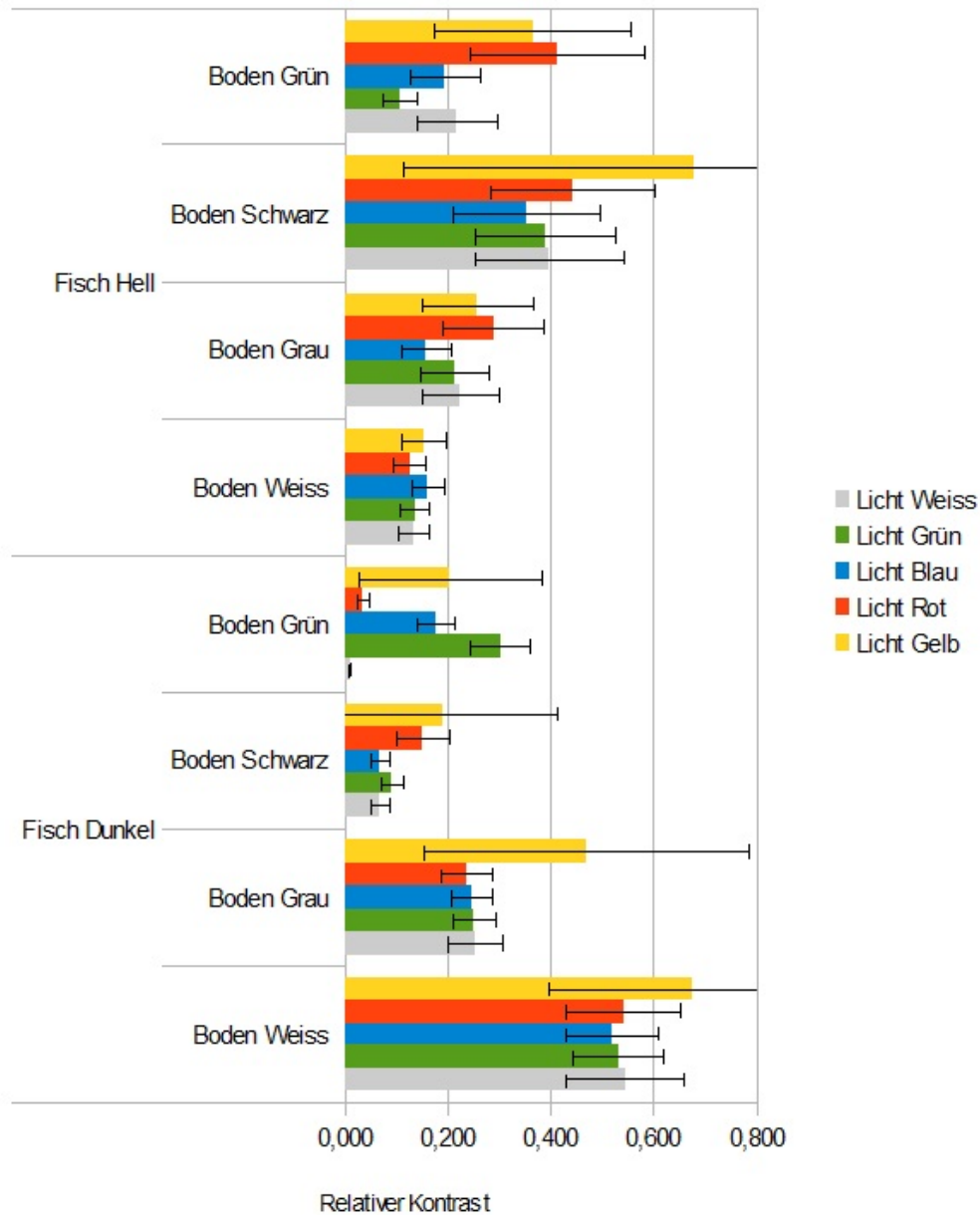


Abbildung 3.24: Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei einer Belichtungszeit von 10 ms.

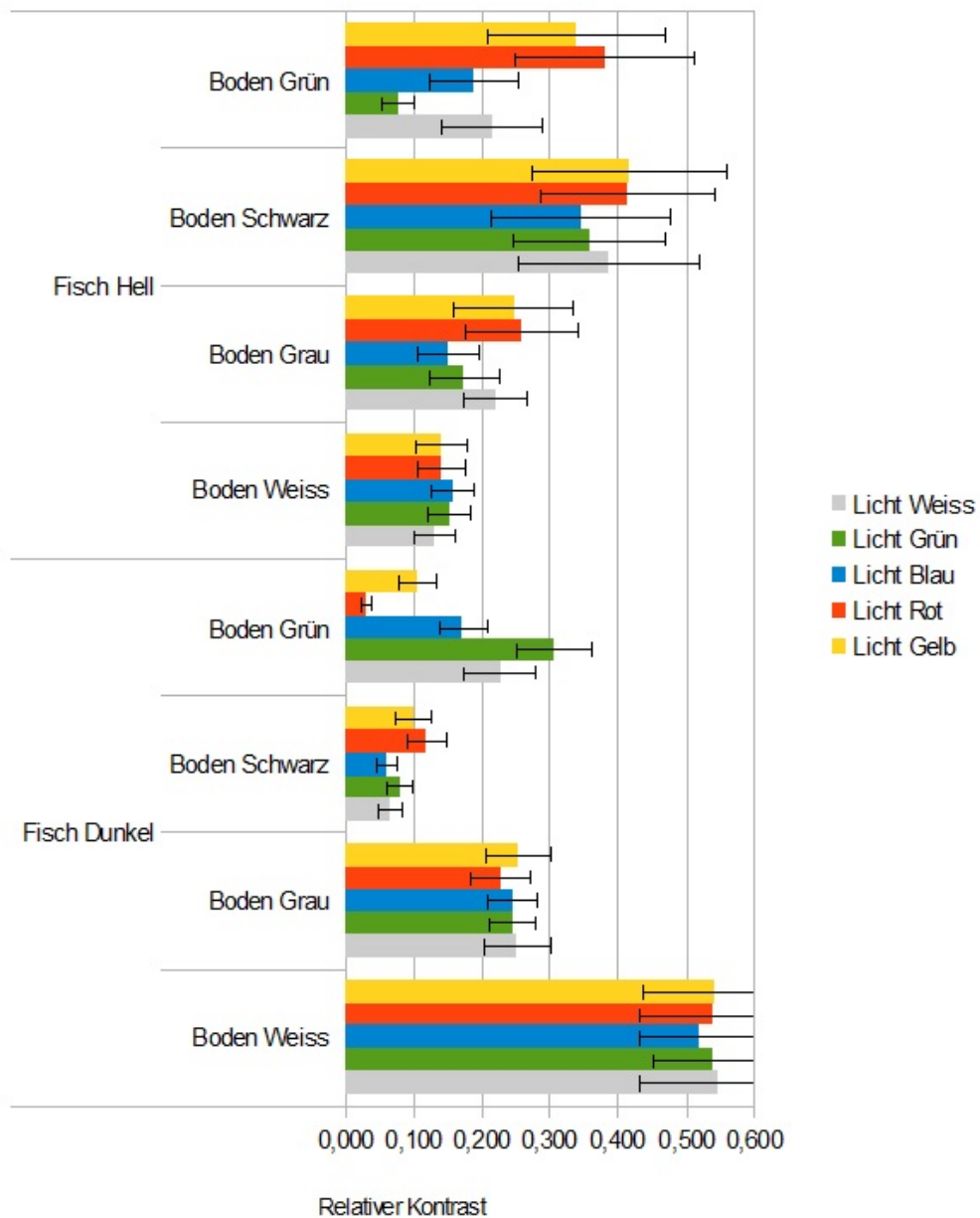


Abbildung 3.25: Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.

mit x_1 als Mittelwert der Bodenplatte und x_2 als Mittelwert der Steinbutte. Die Standardabweichung des relativen Kontrast σ_K beträgt:

$$\sigma_K = \left| \frac{\delta K}{\delta x_1} \right| \sigma_{x_1} + \left| \frac{\delta K}{\delta x_2} \right| \sigma_{x_2} \quad (3.5)$$

mit σ_{x_1} als Standardabweichung der Bodenplatte und σ_{x_2} als Standardabweichung der Steinbutte.

Der Kontrast zeigt den Helligkeitsunterschied zwischen den betrachteten Bereichen. Umso größer der Kontrast ist, desto stärker ist der Unterschied in den vorhandenen Grauwerten, was für eine Segmentierung der unterschiedlichen Bereiche in der Bildverarbeitung von Vorteil ist.

Die Kontraste zwischen Bodenplatte und Steinbutt zeigen keine besonderen Vorteile für eine bestimmte Kombination. Helle Untergründe mit dunklen Steinbutten sowie dunkle Untergründe mit hellen Steinbutten resultieren in hohen Kontrasten, während gleichartige Paarungen relativ niedrige Kontraste ergeben. Innerhalb der Farbvarianten des Lichts sind in den meisten Fällen nur geringe Unterschiede festzustellen. Einzig das Verhalten von rotem und grünem Licht bei einer grünen Bodenplatte ist auffällig. Da der grüne Boden das grüne Lichtspektrum fast vollständig reflektiert, entspricht das Ergebnis eher einem hellen Untergrund. Das rote Licht wiederum wird fast vollständig absorbiert und das Ergebnis entspricht dadurch dem eines dunklen Untergrunds. Das hat den Effekt, dass sich die Ergebnisse der beiden Lichtfarben bei grünem Untergrund bei jeweils hellem und dunklem Steinbutt konträr zueinander verhalten.

Die Abbildungen 3.26 und 3.27 zeigen den *empirischen Variationskoeffizienten* V der Grauwerte innerhalb der Steinbutte, in Abbildung 3.26 bei 10 ms Belichtungszeit und in Abbildung 3.27 bei automatisch festgelegter Belichtungszeit. Dabei werden auf gleichen Areale und Messwerte innerhalb der Steinbutte verwendet, wie bei der Berechnung der Kontraste zwischen den Steinbutten und dem Untergrund.

Der empirische Variationskoeffizient

$$V = \frac{\text{Standardabweichung}}{\text{Mittelwert}} \quad (3.6)$$

ist eine Normierung der Varianz, wodurch die Ergebnisse für die verschiedenen Farbkombinationen von Steinbutt, Untergrund und Licht vergleichbar werden. Ein hoher empirischer Variationskoeffizient deutet auf eine hohe Varianz der Grauwerte innerhalb eines Steinbutes hin, was auf das Vorhandensein von Texturen und Strukturen hindeutet, die mittels Bildverarbeitung erkannt und ausgewertet werden können. Starkes Rauschen kann die Ergebnisse extrem verfälschen, was speziell an den Ergebnissen für das gelbe Licht in Abbildung 3.26 sichtbar ist.

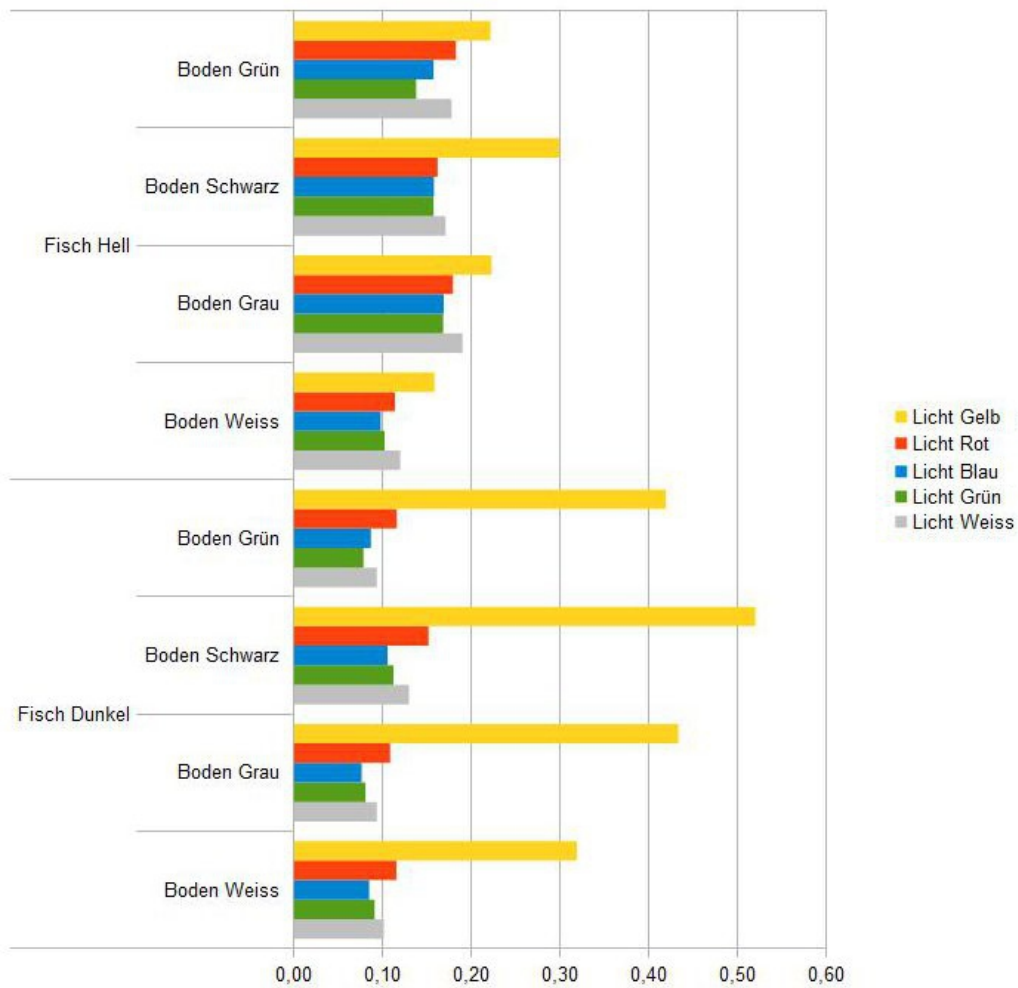


Abbildung 3.26: Empirischer Variationskoeffizient innerhalb eines Steinbutts bei einer Belichtungszeit von 10 ms.

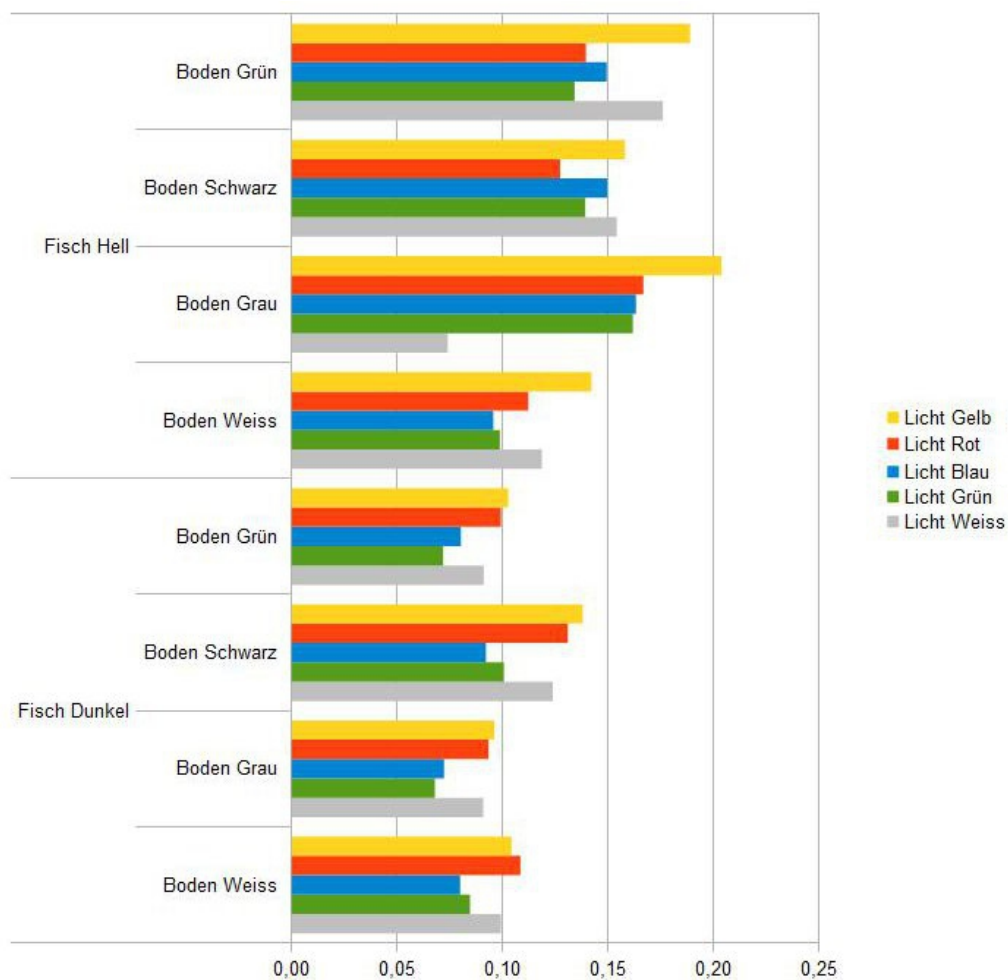


Abbildung 3.27: Empirischer Variationskoeffizient innerhalb eines Steinbutts bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.

Die Ergebnisse bei einer fixen Belichtungszeit von 10 ms zeigen durchgehend eine übereinstimmende Tendenz bei alle Farbvarianten des Lichts. Rotes und weißes Licht weisen im geringen Maße höhere empirische Variationskoeffizienten auf als blaues und grünes Licht, wobei das rote Licht in den meisten Fällen das höchste Ergebnis erzielen konnte. Bei der automatisch ermittelten Belichtungszeit bestätigt sich dieser Trend im Allgemeinen, beim Versuch mit den hellen Steinbutten ergaben sich bei manchen Kombinationen von Boden- und Lichtfarbe Schwankungen.

Es zeigt sich, dass der empirische Variationskoeffizient bei dunklen Böden oft höhere Werte erreicht als bei einem hellen Boden. Dies lässt sich darauf zurückführen, dass durch den relativ stark beleuchteten Boden eine Überbelichtung erfolgt, was die Varianz der Grauwerte innerhalb der Steinbutte reduziert.



Abbildung 3.28: Überbelichtung der Randpartien eines Steinbutts bei weißem Untergrund.

Die Abbildungen 3.29 und 3.30 zeigen die mittlere Gradientenstärke S_m der Konturpixel am Übergang zwischen den beiden Steinbutten, in Abbildung 3.26 bei 10 ms Belichtungszeit und in Abbildung 3.27 bei automatisch festgelegter Belichtungszeit. Mittels des Gradienten kann die Stärke einer Kontur beschrieben werden, umso höher der Gradient ist, desto ausgeprägter ist die Differenz zwischen den Grauwerten der Kontur. Für die Ermittlung der Gradientenstärke S der Konturpixel wurde auf das im Canny Edge Detector angewendete Verfahren zurückgegriffen, der in Kapitel 4.2.3 beschrieben wird.

Für die Bestimmung der mittleren Gradientenstärke S_m der Konturpixel am Übergang zwischen 2 Steinbutten wird in den Bildern ein Bereich von 23×77 Pixeln festgelegt in dem der Übergang von einem Steinbutt zum anderen liegt. Auf diesen Bereich wird der Canny Edge Detector angewendet und die dabei ermittelten Gradientenstärke S pro Konturpixel aufsummiert

gemittelt.

Dies wird für die beiden Belichtungszeiten, die 5 zu prüfenden Lichtfarben sowie die helle und dunkle Musterung der Steinbutte durchgeführt. Auf eine Variation der Bodenfarbe wird in diesem Fall verzichtet und nur die schwarze Bodenplatte verwendet.

Der untersuchte Bereich mit einer Größe von 23×77 Pixeln besteht aus insgesamt 1771 Pixeln, die als Konturpixel hätten klassifiziert werden können. Die Anzahl der durch den Canny Edge Detector als Konturpixel klassifizierten Pixel in Abhängigkeit von Fischmuster, Belichtungszeit und Lichtfarbe findet sich in Tabelle 3.3. Diese Anzahl Pixel wurde auch für die Berechnung der jeweiligen mittleren Gradientenstärke S_m herangezogen.

Tabelle 3.3: Anzahl der im 23×77 Pixel großen Bereich durch den Canny Edge Detector als Konturpixel erkannten Pixel.

Fisch	Bel.Zeit	Weiß	Grün	Blau	Rot	Gelb
Dunkel	10 ms	593	633	629	608	625
Dunkel	Auto	574	577	577	604	621
Hell	10 ms	597	639	616	581	605
Hell	Auto	592	583	578	607	645

Anhand der Schwankungen in der Anzahl der Konturpixel zwischen den unterschiedlichen Lichtfarben läßt sich nicht aussagen, ob eine Lichtfarbe besser für die Konturenfindung geeignet ist, da sich hier schon durch kleine Änderungen sehr schwache Gradienten bilden und mit in die Auswertung kommen. Auffällig sind erneut die hohen Werte bei gelbem Licht und automatisch festgelegter Belichtungszeit. Dies liegt begründet am hohen Rauschen in dieser Beleuchtungsvariante, da auch einzelne durch das Rauschen entstandene Fehlpixel als Kontur klassifiziert werden können.

Die mittleren Gradientenstärken S_m bei 10 ms Belichtungszeit zeigen für beide Mustervarianten der Steinbutte für weißes Licht die höchsten Durchschnittswerte. Rotes und grünes Licht weisen ebenfalls relativ hohe Werte auf, während blaues und gelbes Licht in den niedrigsten Werten resultierten. Bei den dunklen Fischen ist der Unterschied zwischen weiß, rot und grün auf der einen Seite und blau und gelb auf der anderen Seite besonders auffällig.

Die mittleren Gradientenstärken S_m bei automatisch festgelegter Belichtungszeit zeigen bei heller Musterung der Steinbutte für alle Varianten der Lichtfarbe relativ gleiche Werte. Die höchsten Ergebnisse erreichten hier rotes und gelbes Licht, wobei beim Ergebnis für das gelbe Licht das Rauschen in Betracht gezogen werden muss.

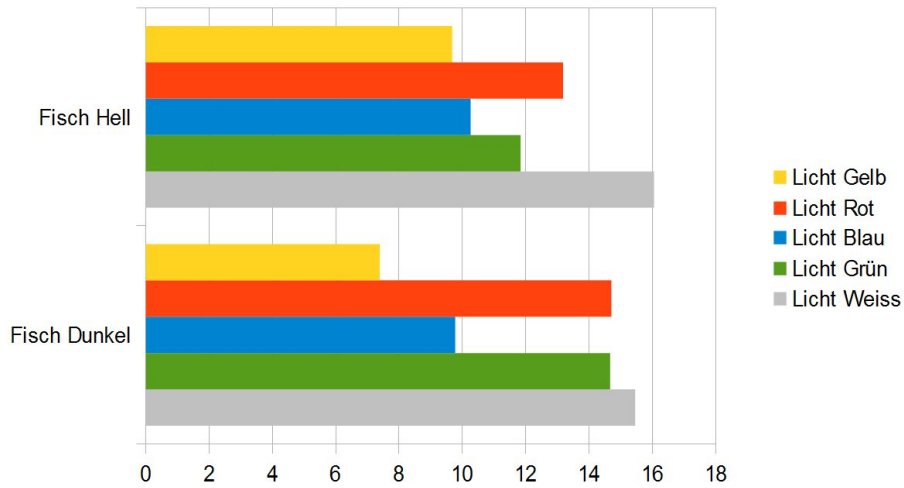


Abbildung 3.29: Mittlere Gradientenstärke S_m im betrachteten Bildausschnitt am Übergang zwischen den Steinbutten bei einer Belichtungszeit von 10 ms.

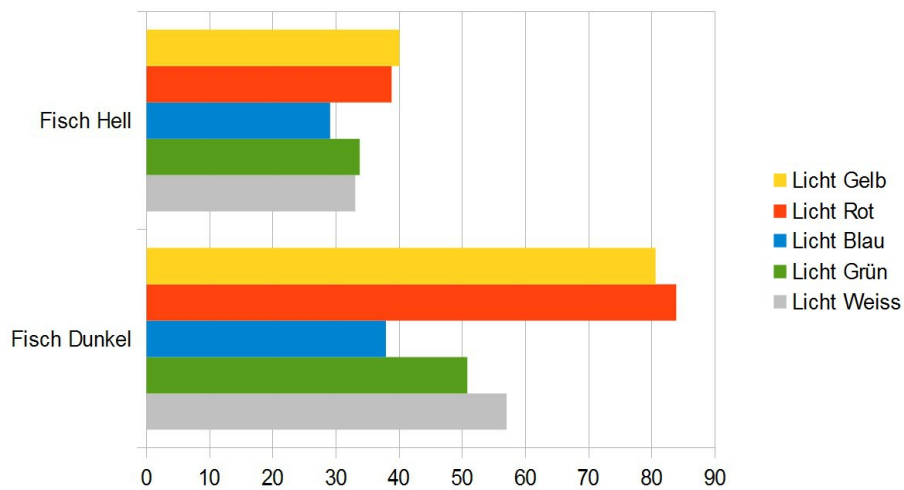


Abbildung 3.30: Mittlere Gradientenstärke S_m im betrachteten Bildausschnitt am Übergang zwischen den Steinbutten bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.

Bei dunkler Musterung der Steinbutte weisen rotes und gelbes Licht erheblich höhere mittlere Gradientenstärken S_m als die anderen Farbvarianten auf. Beim Ergebnis für gelbes Licht ist auch hier aufgrund des hohen Rauschens von einer starken Verfälschung auszugehen. Bei den anderen 3 Lichtfarben ist das Ergebnis vergleichbar mit dem Resultat bei 10 ms Belichtungszeit. Den höchsten Wert erreicht das weiße Licht, gefolgt von grünem und blauem Licht.

Es zeigt sich in den Versuchsbildern, dass helle Untergründe aufgrund ihrer starken Lichtreflexion zwar große Kontraste zu den Steinbutten aufweisen, aber neben niedrigeren empirischen Variationskoeffizienten ebenfalls die Randbereiche und somit die Konturen der Steinbutte überbelichtet und verfälscht werden können, wie man in der Abbildung 3.28 sehen kann. Aufgrund der hohen Besatzdichte ist davon auszugehen, dass nur in seltenen Fällen der Untergrund zu sehen sein wird. Für diesen Fall sollte für den Untergrund eine Farbe ohne großen Einfluss auf die Bildverarbeitung gewählt werden. Dies trifft am ehesten auf einen schwarzen Untergrund zu, weswegen die Entscheidung auf diese Farbe fiel.

In den meisten Kombinationen erzielen rotes und weißes Licht in Form eines höheren empirischen Variationskoeffizienten V die besten Ergebnisse. Hier erzielt das rote Licht oft sogar höhere Ergebnisse als das weiße Licht.

Bei der Bestimmung der mittleren Gradientenstärken S_m überzeugen ebenfalls rotes und weißes Licht mit durchgängig guten Ergebnissen für beide Belichtungszeiten und Mustervarianten der Steinbutte.

Daraus lässt sich folgern, dass für die Erkennung von Kanten und Strukturen bei Steinbutten sowie von Übergängen zwischen Steinbutten rotes und weißes Licht die am besten geeigneten Farben sind. Aufgrund der Tatsache, dass das rote Licht gegenüber dem weißen Licht in den meisten Fällen nur geringfügig höhere Ergebnisse aufweist, weißes Licht aber erheblich bessere Ergebnisse für die Belichtungszeiten aufweist, wird in dieser Arbeit weißes Licht verwendet.

3.4.2 Ergebnisse zur Ausrichtung des Beleuchtungssystems

Für die entstandenen Beleuchtungsbilder werden ausgehend vom Zentrum des Lichtfeldes bis zu den Bildrändern kreissektorförmige *Regions of Interest* (ROI) mit einer Breite von 20 Pixeln gesetzt. Dabei wird der Abstand vom Zentrum in Schritten von 1 Pixel vergrößert. Für diese ROIs wird der mittlere Grauwert berechnet. Abbildung 3.31 zeigt das resultierende Diagramm. Die Kurve für 20° nach innen gedrehte LEDs zeigt einen Bereich hoher Helligkeit in der Mitte und einen starken Abfall zum Rand hin so dass ein Verhältnis von 3,5:1 entsteht. Da durch diese Einstellung auch Schwebeteilchen erneut

sehr direkt bestrahlt werden, erscheint diese Einstellung nicht praktikabel. Bei parallel zur Blickrichtung der Kameras eingestellten LEDs ist das Verhältnis von Zentrum zu Rand mit 2,5:1 immer noch zu hoch. Bei 20° nach außen gedrehten LEDs liegt das Verhältnis knapp unter 2:1 und liegt damit im akzeptablen Rahmen mit einer Differenz von unter 50 % zwischen hellstem und dunkelstem Bereich. Bei 40° nach außen gedrehten LEDs ist das Verhältnis von Zentrum zu Rand unter 1,5:1 und bietet somit die homogenste Beleuchtung. Auffällig ist, dass die Helligkeit im Zentrum zwar wie erwartet abnimmt

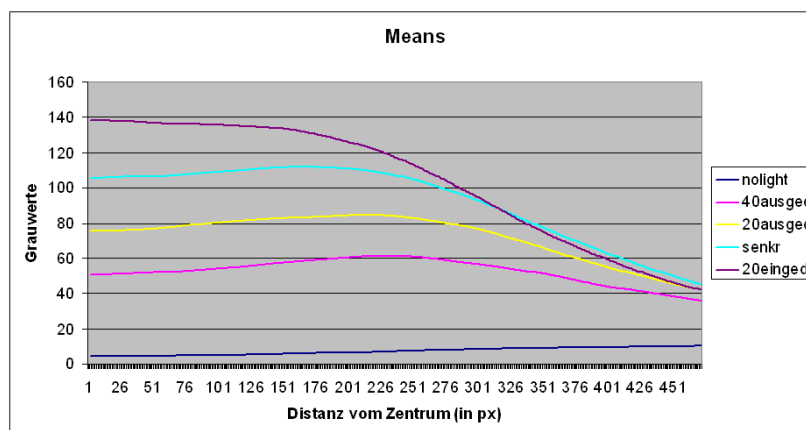


Abbildung 3.31: Mittelwerte der Grauwerte mit ansteigendem Abstand vom Zentrum.

umso weiter die LEDs nach außen gedreht werden, im Randbereich kann aber aufgrund der Vignettierung kein Ansteigen der Helligkeit registriert werden. Dadurch ist die Variante mit 40° nach außen gedrehten LEDs insgesamt sehr dunkel. Hinzu kommt, dass bei dieser Einstellung konstruktionsbedingte Artefakte durch Schattenwurf auftreten. Aufgrund dieser Tatsachen erscheint die Einstellung von 20° nach außen gedrehten LEDs als die beste Alternative. Sie bietet ein akzeptables Verhältnis der Helligkeit von Zentrum zu Rand, ist insgesamt nicht zu dunkel, hat keine Schattenwurf-Artefakte und reduziert durch die indirekte Beleuchtung die direkte Bestrahlung von Schwebeteilchen.

3.4.3 Evaluation der Kalibrierung

Die Messgenauigkeit des kalibrierten Stereokamerasystems wird anhand von Testmessungen überprüft. Dafür wird auf Stereoaufnahmen des zur Kalibrierung verwendeten Kalibrieremusters zurückgegriffen, da hierfür die Abmessungen exakt bekannt sind und somit eine hohe Genauigkeit bei der Überprüfung erreicht werden kann. Es werden in unterschiedlichen Aufnahmen des Kalibrieremusters korrespondierende Punktpaare ausgewählt und deren Koordi-

naten im Raum und der Abstand zueinander berechnet, wie es in Abbildung 3.32 beispielhaft dargestellt wird.

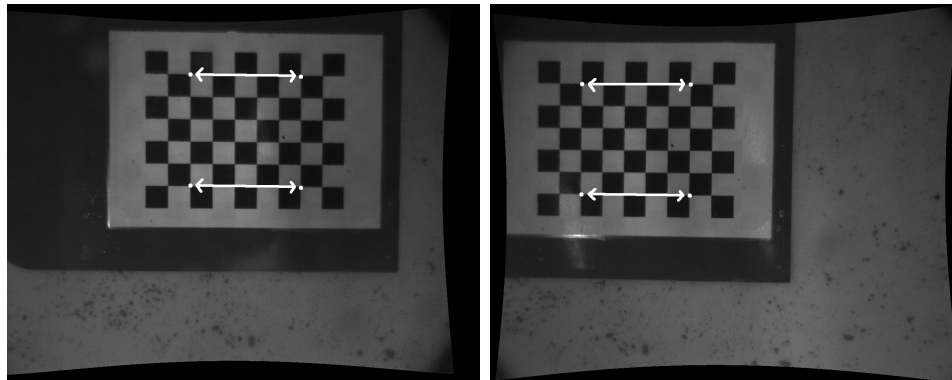


Abbildung 3.32: Beispiel für 2 mögliche berechnete Distanzen unter Verwendung des Kalibrierungsmusters.

Zur Überprüfung der Genauigkeit der Kalibrierung werden jeweils 10 Messungen im Bereich der Bildmitte und des Bildrandes durchgeführt. Der Abstand zum Kamerasystem beträgt dabei maximal 60 cm, bedingt durch die Beckentiefe, in der die Kalibrierung durchgeführt wurde. Die Ergebnisse sind in den Abbildungen 3.33 und 3.34 dargestellt und zeigen die Abweichungen der Distanzmessungen gegenüber dem Sollwert.

Abbildung 3.33 zeigt die Ergebnisse für die Testmessungen im Bereich der Bildmitte. Bei 10 Messungen und einem Sollwert für die Distanz von 2,4 cm, erreicht die präziseste Messung einen Wert von 2,389 cm und somit eine Abweichung von 0,46 %, während die Messung mit der größten Abweichung einen Wert von 2,326 cm und somit eine Abweichung von 3,1 % erzielt. Die im Mittel der Beträge erzielte Abweichung vom Sollwert über die 10 Messungen beträgt 1,7 % oder 1,7 mm auf 10 cm Messlänge.

Abbildung 3.34 zeigt die Ergebnisse für die Testmessungen im Bereich des Bildrandes. Bei 10 Messungen und einem Sollwert für die Distanz von 2,4 cm, erreicht die präziseste Messung einen Wert von 2,386 cm und somit eine Abweichung von 0,57 %, während die Messung mit der größten Abweichung einen Wert von 2,49 cm und somit eine Abweichung von 3,8 % erzielt. Die im Mittel der Beträge erzielte Abweichung vom Sollwert über die 10 Messungen beträgt 2,0 % oder 2,0 mm auf 10 cm Messlänge.

Zur weiteren Überprüfung der Genauigkeit werden exemplarisch weitere Messungen über die Längen 7,2 cm, 9,6 cm sowie 14,4 cm durchgeführt. Die Abweichungen überschreiten dabei nicht die 2,4 %-Marke und bestätigten damit die Ergebnisse aus den Testmessungen auch über größere Distanzen.

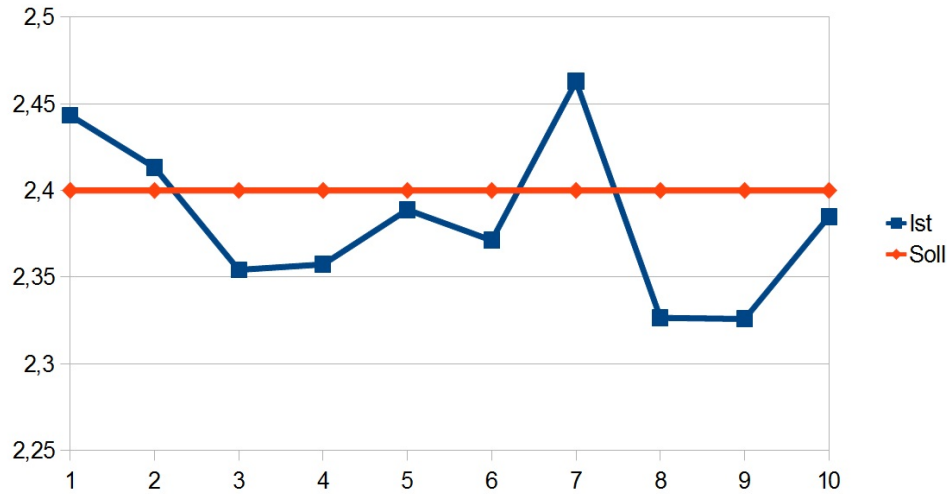


Abbildung 3.33: Ergebnisse von 10 Testmessungen im Bereich der Bildmitte. Sollwert der Distanz beträgt 2,4 cm. Die Abweichung beträgt im Mittel 1,7 %.

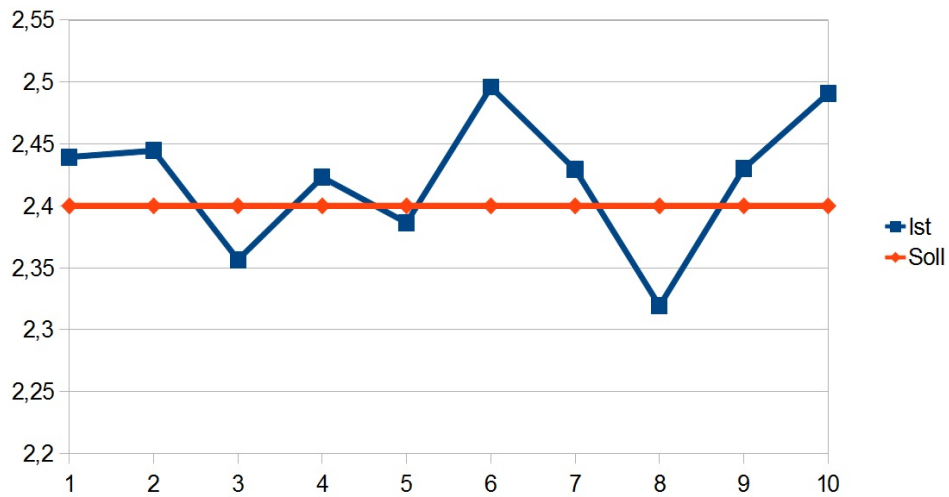


Abbildung 3.34: Ergebnisse von 10 Testmessungen im Bereich des Bildrandes. Sollwert der Distanz beträgt 2,4 cm. Die Abweichung beträgt im Mittel 2 %.

Die Ergebnisse zeigen eine ausreichend hohe Genauigkeit, um für das Wachstums-Monitoring von Steinbutten in der Aquakultur eingesetzt zu werden, da Abweichungen von 1-2 % das Endergebnis bestehend aus sehr vielen gemittelten Messungen nicht negativ beeinflusst. Die Hauptaufgaben des Monitorings, das generelle Wachstum der Steinbutte zu verfolgen sowie die Ermittlung einer erforderlichen Sortierung der Steinbutte, aufgrund einer zu großen Disparität in der Fischgröße, sind bei einer Messungenauigkeit dieser Größenordnung realisierbar.

Der Messfehler bei der Längenbestimmung eines Steinbutts setzt sich aus 2 Fehlerwerten zusammen. Zur Messungenauigkeit der Kalibrierung kommt die Ungenauigkeit bei der Bestimmung der zu berechnenden Positionen von Maul und Flosse hinzu. Bei der Objekterkennung werden jeweils für die Maulspitze und die Flossenspitze ein Punkt festgelegt, für den die Koordinaten im 3-dimensionalen Raum berechnet werden. Umso weiter dieser Punkt von der optimalen Position entfernt ist, desto stärker kann die Abweichung von der korrekten Länge sein. Zur Veranschaulichung sei hier angefügt, dass bei dem Stereokamerasystem in dieser Arbeit (siehe Kapitel 3.2.1 und Kapitel 3.2.2.1) ein Positionsänderung von einem Pixel einem Abstand von 100 cm in einer Messungenauigkeit von bis zu 1 mm resultiert, je nachdem wie die Punkte von Maulspitze und Flossenspitze zueinander gelagert sind. Im Abstand von 50 cm beträgt die Messungenauigkeit noch bis zu 0,5 mm. Bei einem Steinbutt von 30 cm Länge, der in 50 cm Entfernung in der Bildmitte aufgenommen wurde, entspricht die mittlere Ungenauigkeit durch die Kalibrierung circa 5 mm. Bei der Ungenauigkeit durch die Positionsbestimmung von Maulspitze und Flossenspitze ist ein Fehler von circa 5 mm bereits bei einer Abweichung von 10 Pixeln vom Optimum möglich.

Optische Fischvermessung

Inhalt

4.1	Ausgangssituation und Überblick	83
4.2	Vorverarbeitung	87
4.2.1	Separierter nichtlinearer Gaußfilter	87
4.2.2	Histogrammäqualisierung	88
4.2.3	Canny Edge Detector	89
4.3	Fischerkennung	91
4.3.1	Training	91
4.3.2	Clustering	95
4.3.3	Matching	96
4.3.4	Validierung	97
4.3.5	3D-Vermessung	99
4.4	Evaluation der Parameter	100
4.4.1	Ermittlung des Schwellwerts zur Bestimmung repräsentativer Konturen - Versuch	100
4.4.2	Ermittlung des Schwellwertes zur Bestimmung repräsentativer Konturen - Ergebnisse	103
4.4.3	Erkennungsraten des Konturen-Matchings - Versuch	105
4.4.4	Erkennungsraten des Konturen-Matchings - Ergebnisse	107

4.1 Ausgangssituation und Überblick

Das Ziel des Verfahrens, das in dieser Arbeit entwickelt wurde, ist die Längenmessung von Steinbutten in der Aquakultur anhand von Stereobildern, die mit dem in Kapitel 3 vorgestellten Stereokamerasystem erstellt wurden. Für die Längenmessung eines Fisches werden hierfür die Position des Mauls und der Schwanzflosse benötigt und es muss sichergestellt werden, dass beide zu einem Fisch gehören. Um dies zu gewährleisten, muss der gesamte Steinbutt im Bild als solcher erkannt werden.

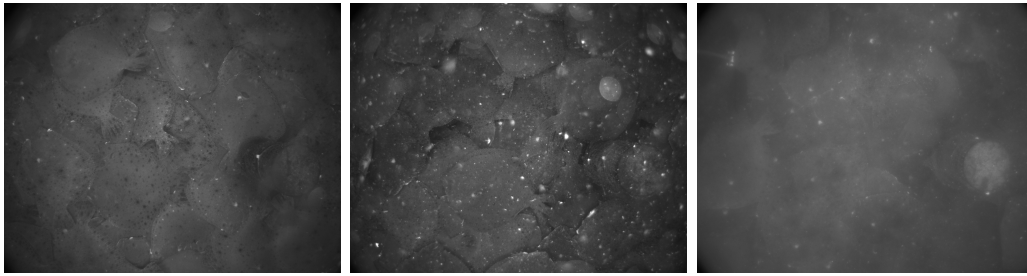


Abbildung 4.1: Beispiele unterschiedlichen Wasserqualitäten. Links: Wasser klar. Mitte: Viele Schwebeteilchen behindern die Sicht. Rechts: Trübes Wasser mit schlechter Sicht.

Die besondere Herausforderung an das Verfahren wird durch die realen Arbeitsbedingungen, in denen die Stereoaufnahmen erfolgen, gegeben. Die Lichtverhältnisse sind durch die Ansprüche der Steinbutte nicht optimal. Da es sich um Unterwasser-Aufnahmen handelt, können Schwebeteilchen als Störfaktoren auftreten und die allgemeine Verunreinigung des Wassers die Sicht mitunter extrem verschlechtern. Deswegen wird für das Verfahren in dieser Arbeit das Vorhandensein von möglichst klarem Wasser als Grundvoraussetzung gestellt. In Abbildung 4.1 sind Beispiele für Wasserqualitäten dargestellt. Das linke Beispiel zeigt relativ klares Wasser mit einer akzeptablen Menge an Schwebeteilchen und kann für das Verfahren in dieser Arbeit genutzt werden. Die anderen beiden Beispiele können auf keinen Fall verwendet werden, in der Mitte aufgrund des sehr hohen Anteils an Schwebeteilchen, rechts aufgrund des trüben Wassers. Aber bereits weniger stark gestörte Aufnahmen behindern die Erkennung massiv.

Die größte Herausforderung an das System ist die hohe Besatzdichte der Becken mit Steinbutten, wodurch das gesamte Sichtfeld der Aufnahme praktisch vollständig mit Steinbutten ausgefüllt ist, was in Abbildung 4.2 gut zu erkennen ist. Hier wird der grundlegende Unterschied zu vielen anderen Einsatzgebieten deutlich.

Im Normalfall hebt sich das gesuchte Objekt von einem Hintergrund ab, der sich in Farbe, Form oder Textur vom Objekt unterscheidet. In diesem Fall jedoch bestehen Vorder- als auch Hintergrund aus den gleichen zu identifizierenden Objekten. Die Unterscheidung zwischen einem Objekt im Vordergrund und Objekten im Hintergrund ist hier lediglich dadurch definiert, dass ein Steinbutt im Vordergrund vollständig erkennbar ist.

Dieser Umstand führt dazu, dass dieselben Strukturen, die das gesuchte Objekt beschreiben, ebenfalls in großer Zahl im Hintergrund vorgefunden werden. Demnach steht das Verfahren vor der Aufgabe, Strukturen von Stein-

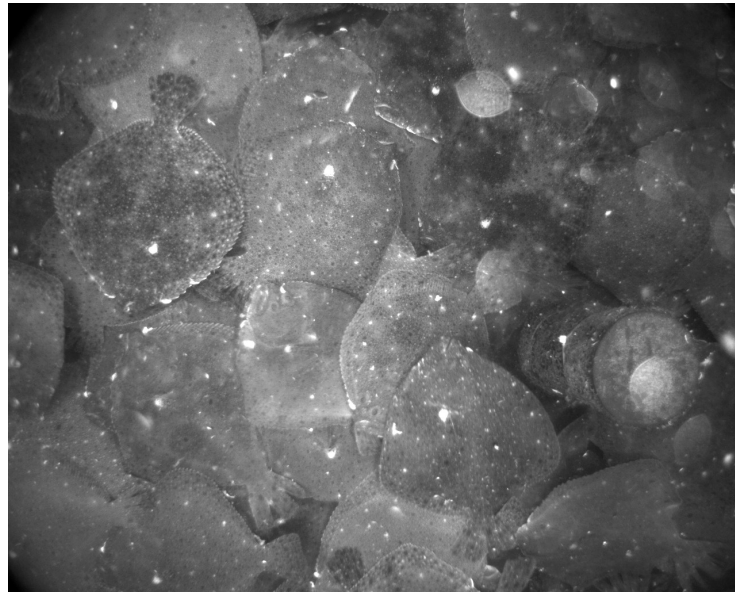


Abbildung 4.2: Typische Aufnahme mit einem Sichtfeld komplett gefüllt mit Steinbutten.

butten im Vordergrund von Strukturen von Steinbutten im Hintergrund zu unterscheiden.

Die in Kapitel 2.1 beschriebenen Tarnmuster und Camouflage-Fähigkeiten erschweren die Erkennung von einzelnen Steinbutten auf extreme Weise, da ihre Funktion darin besteht, mit dem Hintergrund zu verschmelzen und sich somit einer Erkennung auf visueller Basis zu entziehen.

Aufgrund dieser Rahmenbedingungen kann nicht davon ausgegangen werden, dass eine vollständige Kontur eines Steinbuttes ermittelt werden kann, wie sie für die meisten Verfahren zur konturbasierten Formerkennung benötigt werden, siehe hierzu auch die Kapitel 2.2.3.2 und 2.3. Das in dieser Arbeit entwickelte Verfahren orientiert sich an den Arbeiten von Opelt et al. [97] [98] [99], wo die Formerkennung anhand von Konturfragmenten und einer räumlichen Zuordnung erfolgt. Diese Methode wurde für diese Arbeit angepasst und erweitert. Einen Überblick über den Aufbau des Verfahrens gibt die schematische Darstellung in Abbildung 4.3.

Das Verfahren besteht aus 3 Hauptphasen. In der ersten Phase wird anhand von Trainingsdaten ein *Set geordneter repräsentativer Konturfragmente* für das Objekt Steinbutt erstellt, mit deren Hilfe die spätere Erkennung von Steinbutten erfolgt. Die dafür benötigten Arbeitsschritte werden in den Kapiteln 4.3.1 bis 4.3.2 beschrieben. Die zweite Phase beinhaltet die Erkennung und Lokalisierung von Steinbutten in unbekanntem Bildmaterial. Dafür erfolgt

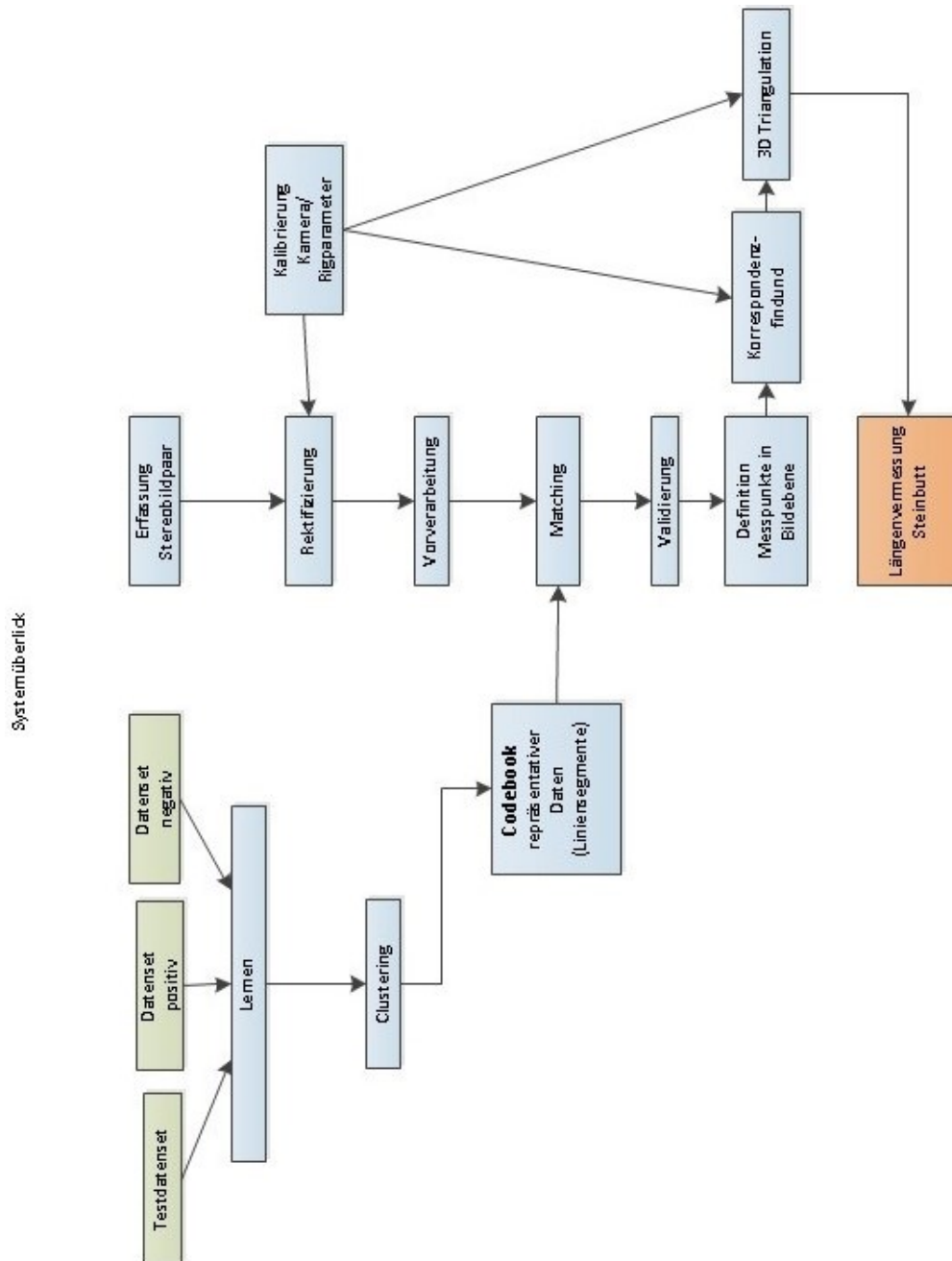


Abbildung 4.3: Schematische Darstellung des Softwaresystems zur Erkennung und Vermessung von Steinbutten.

zuerst die Vorverarbeitung des Bildmaterials, um dies für das Verfahren zu optimieren. Dies wird in Kapitel 4.2 erläutert. Die folgenden Arbeitsschritte zum Matching der repräsentativen Konturen und der anschließenden Ergebnisvalidierung erfolgt in den Kapiteln 4.3.3 und 4.3.4. Die dritte Phase ist die Längenvermessung des gefundenen Steinbutts anhand von Stereobilddaten. Der Ablauf der Korrespondenzfindung und der Triangulierung der Punkte im Raum wird in Kapitel 4.3.5 beschrieben. Abschließend erfolgt die Evaluierung des Verfahrens in Kapitel 4.4.

Zur Realisierung des Verfahrens wird auf einen handelsüblichen PC mit Windows XP als Betriebssystem und Microsoft Visual Studio 2005 als Entwicklungsumgebung zurückgegriffen. Die verwendete Hochsprache ist C++. Zusätzlich kommt die freie Programmbibliothek OpenCV [102] zum Einsatz.

4.2 Vorverarbeitung

Das Ziel der Vorverarbeitung ist die Aufbereitung der empfangenen Bilddaten derart, dass sie für die nachfolgende Objekterkennung optimiert wurden. Dafür werden Verfahren eingesetzt, wie sie in Kapitel 2.2.2 vorgestellt wurden. Um Konturen besser erkennbar zu machen, erfolgt in dieser Arbeit der Einsatz einer Rauschreduktion sowie eine Verstärkung der Kontraste. Abschließend wird für die weitere Verarbeitung der Bilddaten ein Kantenbild erstellt.

4.2.1 Separierter nichtlinearer Gaußfilter

Bei den gängigsten Verfahren zur Rauschunterdrückung erfolgt durch Mittelung ein Verwischen der Konturen, was die folgende Konturenfindung entsprechend verschlechtert. Deswegen wird aufgrund seiner kantenerhaltenden Eigenschaften für die Rauschunterdrückung zu Beginn der Vorverarbeitung ein *separierter nichtlinearer Gaußfilter* (SNLG) [7] [6] verwendet.

Ein linearer Gaußfilter G_λ berechnet den neuen Wert für Grauwert f an Position p aus den benachbarten Positionen q durch:

$$G_\lambda f(p) = \frac{1}{N} \sum_q g_\lambda(\|q - p\|) f(q) \quad (4.1)$$

$$g_\lambda(t) = \exp\left(-\frac{t^2}{2\lambda^2}\right) \quad (4.2)$$

$$N = \sum_q g_\lambda(\|q - p\|) \quad (4.3)$$

mit $g_\lambda(t)$ als unnormierte Gaußfunktion mit Streuung λ und N zur Normierung des Gesamtgewichtes auf 1. Dadurch reduziert sich die Gewichtung mit steigender Distanz von q zu p .

Der Gaußfilter ist nicht kantenerhaltend. Um dies zu erreichen, wird eine zusätzliche Gewichtung

$$g_{\zeta}(|f(q) - f(p)|) \quad (4.4)$$

hinzugefügt, wodurch mit steigender Differenz der Grauwerte die Gewichtung reduziert wird. Der resultierende nichtlineare Gaußfilter ist dann wie folgt definiert:

$$NLG_{\lambda,\zeta}f(p) = \frac{1}{N} \sum_q g_{\sigma}(\|q - p\|) g_{\zeta}(|f(q) - f(p)|) f(q) \quad (4.5)$$

$$N = \sum_q g_{\lambda}(\|q - p\|) g_{\zeta}(|f(q) - f(p)|) \quad (4.6)$$

Der Wert für ζ muss für eine gute Rauschunterdrückung groß und für eine optimale Kantenerhaltung klein gewählt werden. Dieses Problem kann durch Verkettung mehrerer Filter mit unterschiedlichen Parametern umgangen werden. Hierbei beginnt man mit großen Werten für ζ und kleinen Werten für λ . In den darauf folgenden Filterungen werden die Werte für ζ reduziert und für λ erhöht. Dabei sind Verkettungen von 3 bis 5 Filtern in den meisten Fällen ausreichend.

Die Separierung des nichtlinearen Gaußfilters erfolgt durch die Verwendung von 2 eindimensionalen Filtern in x - sowie y -Ausrichtung, die hintereinander ausgeführt werden. Dies hat vor allem eine Verkürzung der Berechnungszeit zur Folge.

4.2.2 Histogrammäqualisierung

Die Histogrammäqualisierung dient der Verbesserung des Kontrastes von Bildern. Dabei wird, wie bereits in Kapitel 2.2.2 erwähnt, das vorhandene Grauwertspektrum auf das gesamte Grauwertspektrum gestreckt und nach Möglichkeit soweit ausgeglichen, dass alle Grauwerte in ihrer Verteilung gleichmäßiger auftreten. Dafür werden die folgenden Schritte ausgeführt [101]:

1. Erstellen der *Kumulativen Verteilungsfunktion* $H'(i)$ für ein Histogramm $H(i)$ durch:

$$H'(i) = \sum_{0 \leq j < i} H(j) \quad (4.7)$$

2. $H'(i)$ in normalisierte Form $H''(i)$ bringen, so dass der Maximalwert dem maximalen Grauwert entspricht.
3. Remapping der Grauwerte durch:

$$\text{Grauwert}_{neu}(x, y) = H''(\text{Grauwert}_{alt}(x, y)) \quad (4.8)$$

von Pixel (x, y) im Bild.

Den stärksten Effekt hat die Histogrammäqualisierung auf Bilder, die nur ein geringes Grauwertspektrum aufweisen. Ein Beispiel hierfür ist in Abbildung 4.4 dargestellt.

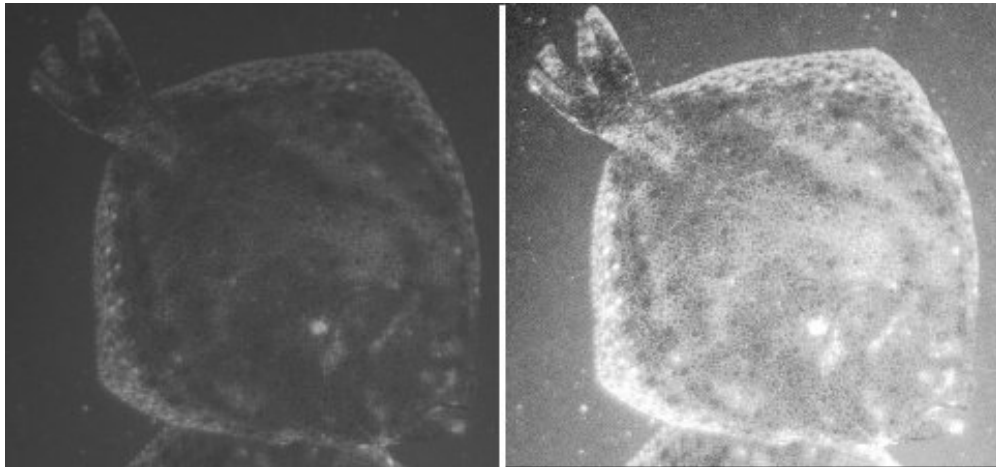


Abbildung 4.4: Beispiel für Histogrammäqualisierung. Links: Originalbild. Rechts: Nach Histogrammäqualisierung.

4.2.3 Canny Edge Detector

Für die Erstellung des Kantenbildes wird der Canny Edge Detector [24] verwendet. Dieser basiert auf 4 Arbeitsschritten [100]:

1. Rauschunterdrückung durch Anwendung eines Gaußfilters.
2. Anschließend wird der Intensitätsgradient für das Bild berechnet. Dafür werden 2 Sobelfilter S_x und S_y auf den aktuell betrachteten Pixel p_i angewendet

$$S_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 1 & +2 \\ -1 & 0 & +1 \end{bmatrix}, S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \quad (4.9)$$

Anschließend wird die Gradientenstärke S und die Orientierung des Gradienten θ , reduziert auf 4 Richtungen (waagrecht, senkrecht und 2 diagonal), berechnet

$$S = \sqrt{(S_x(p_i))^2 + (S_y(p_i))^2} \quad (4.10)$$

$$\theta = \arctan \frac{S_y}{S_x} \quad (4.11)$$

3. Bei der Unterdrückung von Nicht-Maxima wird anhand von Stärke und Orientierung des Gradienten ermittelt, ob ein Pixel zu einer Kante gehört. Dadurch entsteht ein Binärbild von dünnen Linien.
4. Abschließend werden mittels Hysterisis-Schwellwert schwache Kanten eliminiert. Hierfür werden ein oberer Schwellwert t_{high} und ein unterer Schwellwert t_{low} benötigt. Kantenpixel p mit $S_p > t_{high}$ gehören garantiert zu einer Kante. Bei $t_{low} < S_p < t_{high}$ muss der Kantenpixel mit einem Kantenpixel mit $S_p > t_{high}$ verbunden sein. Ist $S_p < t_{low}$ wird der Kantenpixel direkt entfernt.

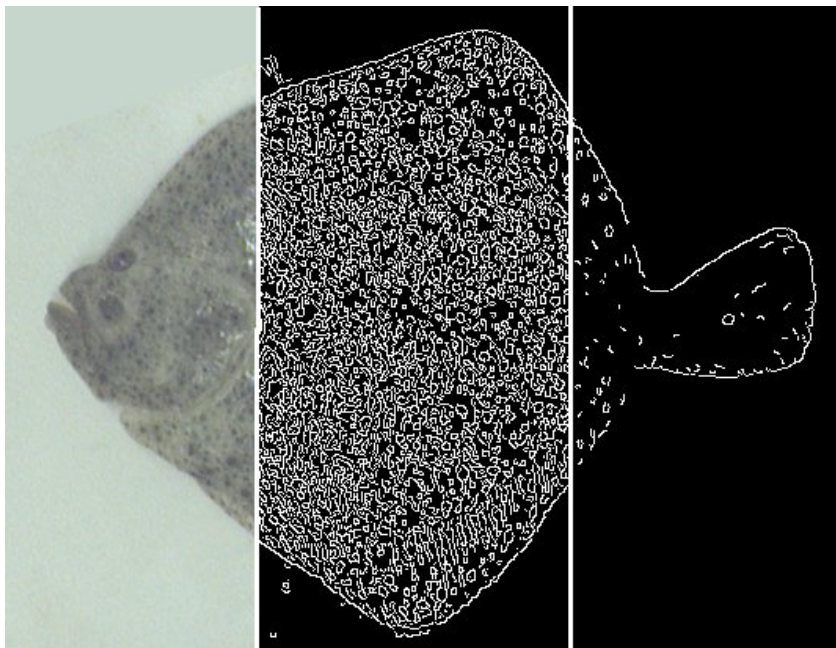


Abbildung 4.5: Zwei Beispiele für die Schwellwerte t_{low} und t_{high} beim Canny Edge Detector. Links: Original. Mitte: Niedrige Schwellwerte mit $t_{low} = 30$, $t_{high} = 60$. Rechts: Hohe Schwellwerte mit $t_{low} = 110$, $t_{high} = 150$.

Das resultierende Kantenbild wird maßgeblich von den beiden Schwellwerten t_{low} und t_{high} bestimmt. Bei niedrigen Werten werden selbst schwache Strukturen mit nur geringen Gradienten als Kanten erkannt. Bei hohen Schwellwerten werden nur sehr ausgeprägte Kanten erkannt. Diese sind dafür oft in kleinere Teilstücke aufgeteilt. Beispiele für unterschiedliche Schwellwerte werden in Abbildung 4.5 illustriert.

Die Wahl der Schwellwerte ist ebenfalls abhängig von den Eigenschaften des Ausgangsbildes. So werden bei kontrastarmen Bildern bereits bei niedrigen

Schwellwerten viele Kanten aussortiert. Es gilt, für den Canny Edge Detector je nach gewünschtem Ergebnis und Eigenschaften des Ausgangsbildes die passenden Parameter zu finden.

4.3 Fischerkennung

4.3.1 Training

Das Ziel des Trainings ist es, ein Set repräsentativer Konturfragmente für das Objekt Steinbutt zu ermitteln, das die Basis bildet, um Steinbutte in unbekanntem Bildmaterial zu identifizieren und die Position im Bild zu ermitteln. Für eine Übersicht wird der prinzipielle Ablauf des Trainings in Abbildung 4.6 schematisch dargestellt. Die Trainingsdaten (Kapitel 4.3.1.1) werden zu Beginn mittels Vorverarbeitung (Kapitel 4.2) in Kantenbilder überführt. Anschließend erfolgt über mehrere Schleifendurchläufe die Ermittlung von repräsentativen Konturfragmenten (Kapitel 4.3.1.3) anhand einer Kostenberechnung basierend auf der *Chamfer distance* (Kapitel 4.3.1.2)

4.3.1.1 Trainingsdaten

Für das Training werden 3 Trainingssets aus Bildern mit zusätzlichen Informationen zusammengestellt:

1. Trainingsbilder: Enthält das zu trainierende Objekt, das in allen Trainingsbildern gleich ausgerichtet und auf die gleiche Größe skaliert ist. Benötigte Zusatzinformationen für das Training sind die Boundingbox um das Objekt sowie der Objekt-Mittelpunkt. Für diese Arbeit wurden für die spätere Verwendung im Verfahren noch die Positionen der Maul- sowie Flossenspitze hinzugefügt.
2. Trainingsset Positiv: Enthält das zu trainierende Objekt mit der gleichen Skalierung und Ausrichtung wie in den Trainingsbildern. Eine benötigte Zusatzinformation für das Training ist der Objekt-Mittelpunkt.
3. Trainingsset Negativ: Enthält weder das zu trainierende Objekt noch weitere Zusatzinformationen. Hier empfehlen sich Aufnahmen von anderen zu erwartenden Strukturen wie Schwebeteilchen oder Gegenstände im Becken.

Die Daten über Maul- und Schwanzflossenpunkte werden für das Training an sich nicht benötigt, aber bereits jetzt wie die Objekt-Mittelpunkte bei den Trainingsbildern gespeichert. Beispiele für alle 3 Trainingssets werden in Abbildung 4.7 dargestellt.

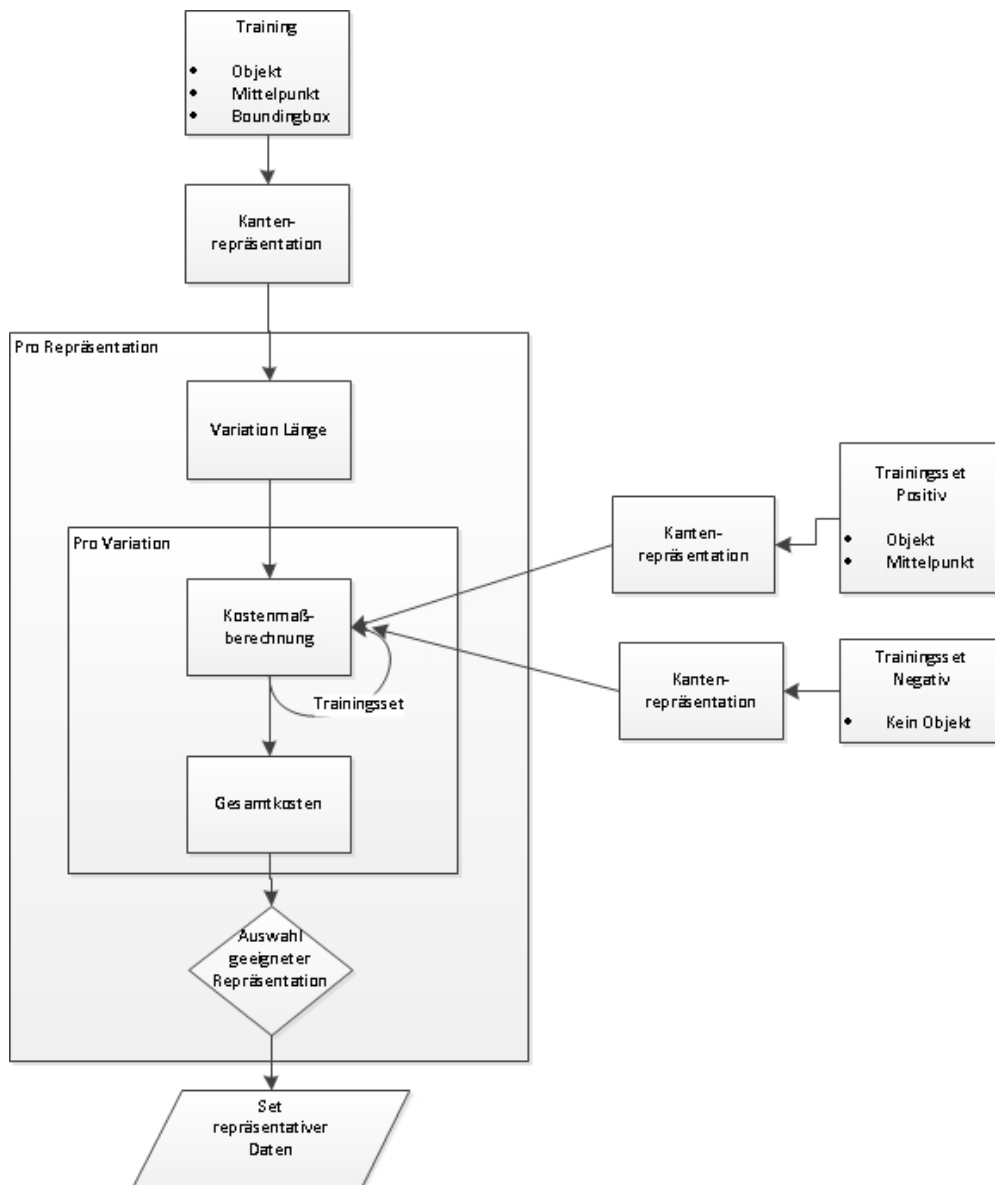


Abbildung 4.6: Schematische Darstellung der Trainingsphase des Softwaresystems.

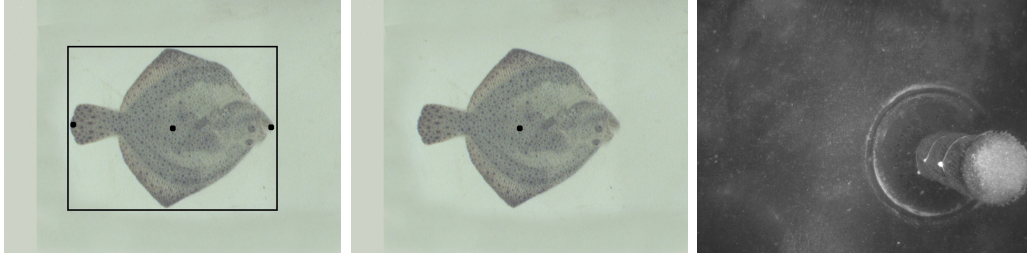


Abbildung 4.7: Beispiele für 3 Trainingssets. Links: Trainingsbild mit Boundingbox, Objekt-Mittelpunkt, Maul- und Flossenspitze. Mitte: Trainingsset Positiv mit Objekt-Mittelpunkt. Rechts: Trainingsset Negativ ohne Steinbutt aber anderen Strukturen.

4.3.1.2 Kostenberechnung

Die Gesamtkosten $C(\gamma_i)$ für eine Kontur γ_i aus den Trainingsbildern basieren auf [97] und sind definiert durch:

$$C(\gamma_i) = c_{chamf}(\gamma_i)c_{loc}(\gamma_i). \quad (4.12)$$

c_{loc} ist die über die Anzahl der Bilder des Trainingsset Positiv gemittelte euklidische Distanz zwischen den beiden Objekt-Mittelpunkten vom Bild des Trainingssets Positiv und der gematchten Kontur aus den Trainingsbildern:

$$c_{loc}(\gamma_i) = \sum_{i=1}^{L^+} pixdist(\gamma_i, P_{v_i})/L^+ \quad (4.13)$$

mit L^+ als die Anzahl der Bilder im Trainingsset Positiv P_{v_i} .

$c_{chamf}(\gamma_i)$ sind die Matching-Kosten der Kontur γ_i zu den Konturen in den Trainingssets Positiv und Negativ:

$$c_{chamf}(\gamma_i) = \frac{\sum_{i=1}^{L^+} dist(\gamma_i, P_{v_i})/L^+}{\sum_{i=1}^{L^-} dist(\gamma_i, N_{v_i})/L^-} \quad (4.14)$$

mit L^- als die Anzahl der Bilder im Trainingsset Negativ N_{v_i} .

$dist(\gamma_i, I_{v_i})$ beschreibt die Distanz zwischen der Kontur γ_i und der am besten übereinstimmenden Kontur in einem Bild I_{v_i} und basiert auf dem *Chamfer Matching* nach [16]. Das verwendete Distanzmaß DT ist die *Chamfer distance*.

Bei der Chamfer distance wird im Suchbild jedem Pixel ein Wert gleich der kürzesten Distanz zu nächsten Kontur in dem Bild zugewiesen und Konturpixel erhalten den Wert 0. Von der zu matchenden Kontur wird eine Auswahl von Pixeln in das Suchbild übertragen und die gespeicherten Distanzwerte

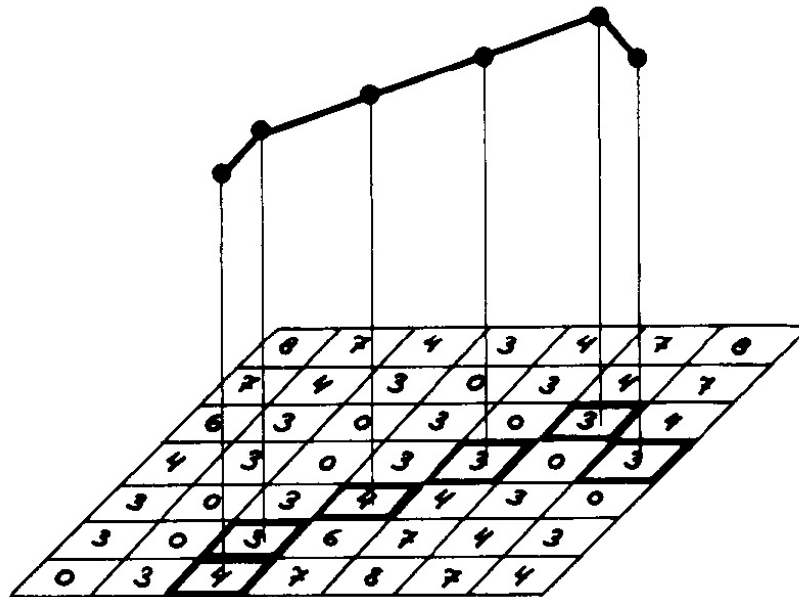


Abbildung 4.8: Funktionsweise Chamfer distance. Das Suchbild mit eingetragenen kürzesten Distanzen zur nächsten Kontur sowie zu matchende Kontur. Illustration aus [16].

der überlagerten Pixel werden gemittelt. Liegt die zu matchende Kontur genau auf einer Kontur im Suchbild resultiert dies in einer Distanz von 0 [16]. Zur Veranschaulichung dient Abbildung 4.8.

Anstatt, wie beim Chamfer Matching üblich, das niedrigste Distanzmaß zu verwenden, werden beim Training der repräsentativen Konturen zuerst die 10 niedrigsten Ergebnisse ermittelt. Hiervon wird das Ergebnis verwendet, bei dem die euklidische Distanz zwischen den Objekt-Mittelpunkten vom Bild des Trainingssets Positiv und der gematchten Kontur des Trainingsbildes am geringsten ist. Abschließend wird die Distanz über die Länge der Kontur normalisiert:

$$dist(\gamma_i, I_{v_i}) = \frac{1}{|\gamma_i|} \min_{\gamma_i \subset I_{v_i}} \sum_{t \in \gamma_i} DT_{I_{v_i}}(t). \quad (4.15)$$

4.3.1.3 Ablauf des Trainings

Zu Beginn des Trainings werden zunächst alle Bilder aus den 3 Trainingssets vorverarbeitet, wie in Kapitel 4.2 beschrieben. Die dabei verwendeten Parameter wurden systematisch variiert und für die Trainingsdaten einzeln angepaßt, um optimierte Ergebnisse zu erhalten.

Jede Kontur aus einem Trainingsbild, die sich komplett innerhalb der

Boundingbox befindet, ist ein Kandidat für ein repräsentatives Konturfragment. Von jedem Kandidaten werden mehrere Varianten unterschiedlicher Länge ausgehend von zufallsgenerierten Saatpunkten erstellt.

Für jede Längenvariante γ_{i_v} wird nun für jedes Bild in Trainingsset Positiv und Trainingsset Negativ die Distanz

$$\text{dist}(\gamma_{i_v}, I_{v_i}) \quad (4.16)$$

berechnet und anschließend die Gesamtkosten $C(\gamma_{i_v})$. Von allen Längenvarianten γ_{i_v} für einen Kandidaten wird diejenige mit den geringsten Gesamtkosten als repräsentatives Konturfragment k_i für das Objekt Steinbutt gewertet und gespeichert, sofern ein Schwellwert t_{train} nicht überschritten wird, um ein Mindestmaß an Ähnlichkeit zwischen den gematchten Konturen zu gewährleisten. Zur Ermittlung des Schwellwertes t_{train} wurde ein Versuch durchgeführt, dessen Ablauf und Ergebnisse in den Kapiteln 4.4.1 und 4.4.2 behandelt werden.

4.3.2 Clustering

Als Resultat des Trainings liegt ein Set von repräsentativen Konturfragmenten für das Objekt Steinbutt vor. Diese entstanden aus mehreren Trainingsbildern, wodurch Linien auftreten können, die einen näherungsweise gleichen Bereich des Objektes beschreiben und dadurch sehr ähnlich sein können.

Um sehr ähnlich verlaufende Konturfragmente zu gruppieren, wird das Set mittels *agglomerativem Clustering* über den *Medoiden* sortiert. Das agglomerative Clustering ist ein hierarchisches Verfahren und gehört zu den unüberwachten Klassifikationsmethoden (Kapitel 2.2.3.1).

Der Ablauf für das Clustering ist der folgende:

1. Jedes der n Konturfragmente k_1, \dots, k_n wird einem eigenen Cluster K_1, \dots, K_n zugewiesen.
2. Berechnung des Distanzmaßes $d(K_i, K_j)$ mit $i, j \in n$ zwischen allen Clustern. Als Distanzmaß dient 4.15 aus Kapitel 4.3.1.
3. Die 2 Cluster mit dem geringsten Distanzmaß zueinander werden zu einem Cluster zusammengefasst. Dabei wird ein Schwellwert von t_{clust} angewendet, bei dessen Überschreiten kein weiteres Cluster mehr gebildet wird.
4. Medoid des Clusters ermitteln. Als Medoid wird das mittlere Element eines Clusters beschrieben und über die kleinste Summe der Distanzmaße zwischen den Elementen eines Clusters definiert. Für die weiteren Distanzberechnungen des Clusters wird der Medoid verwendet. Bei nur 2 Elementen wird die längere Kontur als Medoid festgelegt.

5. Wiederholen der Schritte 2 bis 5, bis kein weiteres Cluster mehr gebildet werden kann.

Das Ergebnis ist ein Set geordneter repräsentativer Konturfragmente. Ähnliche Konturfragmente sind zu Gruppen zusammengefasst und werden durch einen Medoiden vertreten. Abbildung 4.9 zeigt 2 Beispiele für solche Cluster.

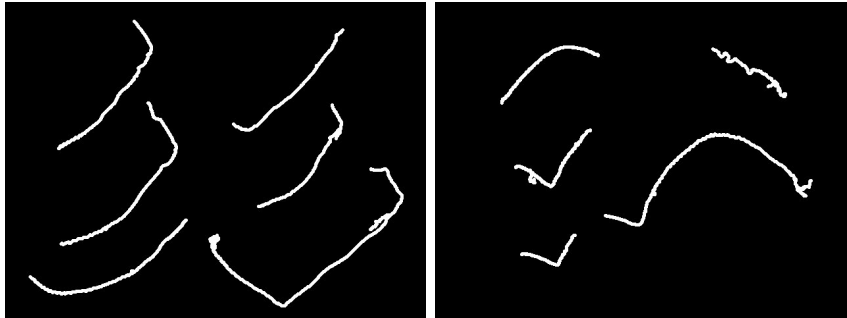


Abbildung 4.9: Beispiele für 2 Cluster repräsentativer Kanten beim Steinbutt.

4.3.3 Matching

Das in den Kapiteln 4.3.1 bis 4.3.2 erstellte Set geordneter repräsentativer Konturfragmente wird jetzt angewendet, um in Bilddaten mit unbekanntem Inhalt nach ähnlichen Strukturen zu suchen. Dafür durchlaufen die neuen Bilder zuerst die in Kapitel 4.2 vorgestellte Vorverarbeitung. Die dafür verwendeten Parameter wurden für diese Arbeit systematisch variiert und für die entsprechenden Situationen jeweils angepasst.

Das nach der Vorverarbeitung vorliegende Kantenbild wird in einem weiteren Arbeitsschritt von kurzen Kantenelementen und vollständig abgeschlossenen Konturen befreit. Kurze Kantenelemente sind in der Praxis überwiegend generisch und bieten Möglichkeiten für fehlerhaftes Matching. Die vollständig abgeschlossenen Konturen stellen unter den gegebenen Bedingungen zumeist Schwebeteilchen oder weiße Flecken auf der Haut des Steinbuttes dar, die für das in dieser Arbeit verwendete Verfahren keinen Mehrwert bieten.

Von dem jetzt vorliegenden Kantenbild wird jede Kante mit dem Set geordneter repräsentativer Konturfragmente abgeglichen. Dafür wird zwischen der Kante und jedem Medoiden eines Clusters das aus Kapitel 4.3.1 bekannte Distanzmaß 4.15 berechnet, jedoch ohne die abschließende Normalisierung über die Länge der Kontur. Das Cluster mit der niedrigsten Distanz wird gewählt. Enthält das Cluster mehr als ein Konturfragment, wird für diese ebenfalls das Distanzmaß zur Kante berechnet und das Konturfragment mit dem

niedrigsten Ergebnis c_{match} wird verwendet, sofern es einen zuvor festgelegten Schwellwert t_{match} unterschreitet.

Dabei werden auch Skalierungen und Positionsveränderungen des Konturfragments berücksichtigt und ebenfalls auf den zum Konturfragment gehörenden Objekt-Mittelpunkt sowie die Maul- und Schwanzflossenpunkte übertragen. Dieses resultierende Konturfragment wird für die anschließende Validierung in einem Ergebnisvektor gespeichert.

4.3.4 Validierung

Die in Kapitel 4.3.3 gefundenen Konturfragmente können aufgrund verschiedener Störquellen fehlerbehaftet sein oder nicht zu einem Fisch gehören und können aufgrund von vorhandenem Vorwissen validiert werden. Die folgenden Abschnitte zeigen die durchgeführten Validierungsschritte.

Punktmengen Für das hier verwendete Verfahren war es notwendig, zu jeder Kontur den Mittelpunkt des zugehörigen Objektes zu erfassen. Anhand dieses Kriteriums kann ausgewählt werden, welche der Vielzahl an Linien aus einem Bild zu einem Objekt gehören, wenn die einzelnen Linien einen ähnlichen Mittelpunkt aufweisen.

Idealerweise würden die Objekt-Mittelpunkte der repräsentativen Konturen, die zu einem Objekt gematcht wurden, perfekt aufeinander liegen. Aufgrund von prinzipiellen Unterschieden in Größe und Form der Steinbutte und den dadurch entstehenden Variationen der erfassten Konturen ist in der Realität davon auszugehen, dass es zu Abweichungen kommt. Diesbezüglich kann nur eine Nachbarschaft erfasst werden, was im Nachfolgenden beschrieben wird. Für das Clustern der Kontursegmente wird eine klassische Distanzsuche verwendet. Dazu wird für jeden Mittelpunkt p_c ein Cluster erstellt, dem all jene Mittelpunkte p_i zugeordnet werden, deren Distanz den Schwellwert t_{dist} unterschreitet:

$$t_{dist} > \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2}. \quad (4.17)$$

Die Abbildung 4.10 zeigt ein Beispiel, in dem mehrere Mittelpunkte einem Cluster in Form eines Kreises zugeordnet wurden. In der Abbildung sind die Konturen des Objekts weiß dargestellt, repräsentative Konturen sowie deren zugehörige Objekt-Mittelpunkte werden je nach ihren Matchingkosten rot ($c_{match} < t_{match} = 0,3$) beziehungsweise grün ($c_{match} < t_{match} = 0,4$) dargestellt. Zu weit entfernte Mittelpunkte gehören zu Konturen, die fehlerhaft gematcht wurden. Nicht dargestellt ist die Möglichkeit, dass ein zu weit entfernter Mittelpunkt einem anderen Steinbutt zugehörig ist. Anschließend werden redundante Cluster entfernt. Weitere Validierungsschritte sind die Überprüfung auf redundante Cluster sowie die Erfüllung der Mindestanzahl an Kon-

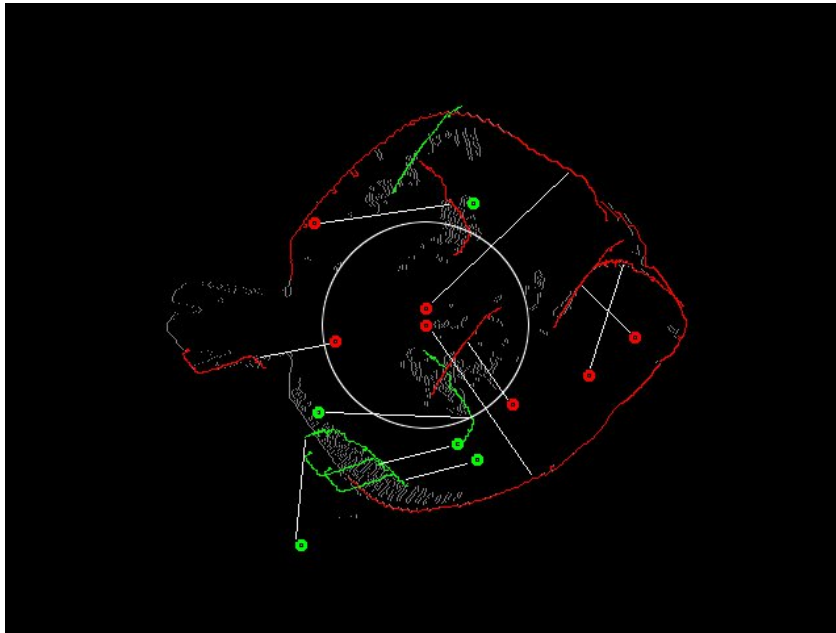


Abbildung 4.10: Nachbarschaftssuche von Mittelpunkten repräsentativer Konturen an einem einzelnen Beispiel.

turen für die Erkennung eines Objektes. Sollte anhand des zunächst niedrig gesetzten Schwellwert t_{dist} keine korrespondierenden Objekte gefunden werden, wird t_{dist} stufenweise bis zu einem Maximalwert erhöht. Die Auswahl des Objektes erfolgt dann über die höchste Anzahl an Clusterelementen.

Abdeckung Ein weiterer Validierungsschritt ist die Überprüfung der prozentualen Abdeckung eines Objektes. Diese Überprüfung ist notwendig, da sich eine sichere Erfassung eines Objektes erst mit einer gewissen Abdeckung der Objekt-Kontur ergibt. Dadurch soll sichergestellt werden, dass sich nicht ein Großteil der gematchten Konturen auf denselben kleinen Bereich des Objekts bezieht.

Vorhandensein von Maul- und Flossenbereich Dem Trainingsdatensatz wurden äquivalent zu den Objekt-Mittelpunkten ebenfalls Informationen über die Position von Maul und dem Endpunkt der Schwanzflosse in Form von Punktkoordinaten hinzugefügt. Dies soll als Grundlage für die Längenmessung des Steinbuttes dienen, da dies die aktuell einzige verwendete Methode zur Gewichtsapproximation ist. Im Kontext dieser Arbeit wird ebenfalls evaluiert, ob sich trotz nicht vorhandener Liniensegmente von Maul und Schwanzflosse trotzdem eine Längenschätzung durchführen lässt. Anhand der

zuvor durchgeführten Überprüfung des Abdeckungsbereiches der gefundenen Konturen lässt sich ebenfalls bestimmen, ob explizit Konturen im Bereich des Mauls und der Schwanzflosse gematcht wurden. Ist dies der Fall, wird von einem erfassten Maul- und Flossenbereich ausgegangen (siehe Kapitel 4.3.4). Wurden im Maul- und Flossenbereich keine Konturen gematcht, so muss dieser Bereich approximiert werden (siehe Kapitel 4.3.4).

Erfasster Maul- und Flossenbereich Für den Fall, dass Maul und Schwanzflosse erfasst wurden, müssen 2 Varianten unterschieden werden. Wurde für jeden der beiden Bereiche nur eine repräsentative Kontur gematcht, kann davon ausgegangen werden, dass der dazugehörige Maul- beziehungsweise Schwanzflossenpunkt an der optimalen Position des gematchten Steinbutts liegt. Dies setzt voraus, dass die Linien perfekt gematcht wurden, was zu einer hohen Fehleranfälligkeit führt. Die zweite Variante basiert darauf, dass die Maul- und Schwanzflossenpunkte aller gematchten Konturen einen Bereich eng beieinander liegender Punkte bilden, genau wie die Objekt-Mittelpunkte der Konturen. Aufgrund der gematchten Konturen in den beiden Bereichen ist davon auszugehen, dass diese Bildbereiche zum erfassten Steinbutt gehören. Von den beiden Punktgruppen wird jeweils der Schwerpunkt ermittelt und für die weitere Berechnung als Maul- beziehungsweise Schwanzflossenpunkt verwendet.

Approximierter Maul- und Flossenbereich Grundlegend wird das zuvor erläuterte Verfahren für erfasste Maul- und Flossenbereiche verwendet. Wenn für die entsprechenden Bereiche jedoch keine Konturen gematcht werden konnten, muss die Möglichkeit einer Verdeckung durch einen anderen Steinbutt angenommen werden. Da es sich hierbei um ein über dem gematchten Steinbutt schwimmendes Exemplar handeln kann, ändert sich für den approximierten Maul- beziehungsweise Flossenpunkt die Höhenposition im Raum, wodurch es zu Fehlmessungen kommen kann.

Das Ergebnis beider Berechnungen ist jeweils ein Punkt im Sinne von einem definiertem Pixel für das Maul und für die Schwanzflosse des Steinbuttes, für die bei der 3D-Vermessung eine Korrespondenz im Stereobild gefunden werden muss.

4.3.5 3D-Vermessung

Im Kalibrierprozess wurde eine F-Matrix (Kapitel 2.2.1.3) ermittelt, die die Grundlage für die weiteren Berechnungen bildet. Um für die ermittelten Maul- sowie Schwanzflossenpunkte deren 3-dimensionalen Positionen zu ermitteln, müssen zuerst deren korrespondierende Punkte im Stereobild der zweiten Ka-

mera lokalisiert werden. Dafür werden für die beiden Punkte im ersten Bild die entsprechenden Epipolarlinien im zweiten Bild berechnet. Aufgrund einer zuvor durchgeführten Rektifizierung (Kapitel 2.2.1.4.1) verlaufen die Linien parallel zur x-Achse.

Die korrespondierenden Punkte können dann mittels *Template Matching*-Verfahren entlang der Epipolarlinien ermittelt werden. Beim Template Matching wird für ein Bildausschnitt bestimmter Größe, dem Template, anhand eines Differenzmaßes in einem anderen Bild nach dem Bildausschnitt gesucht, das die größte Übereinstimmung mit dem Template aufweist. In diesem Fall wird als Template der Bildbereich mit dem gesuchten Pixel als Zentrum gewählt. Vom gefundenen Bildausschnitt wird wiederum der zentrale Pixel als korrespondierender Pixel verwendet.

Als letztes folgt die Triangulierung, eine Reprojektion der gefundenen Pixel in den 3-dimensionalen Raum. Hierbei entstehen zunächst Strahlen im Raum. Wären ideale Voraussetzungen gegeben, müssten sich diese Strahlen aufgrund der mathematischen Bedingungen schneiden, wie in Kapitel 2.2.1.2 dargestellt. Da dies in der Realität selten gegeben ist, muss hier der Schnittpunkt approximiert werden. Hierfür kann ein Standardverfahren nach [59] genutzt werden.

Abschließend kann anhand dieser 3-dimensionalen Koordinaten jetzt die Länge des Steinbuttes im Raum berechnet werden und eine Abschätzung des Gewichts anhand der in Kapitel 2.1 vorgestellten Formel erfolgen.

4.4 Evaluation der Parameter

4.4.1 Ermittlung des Schwellwerts zur Bestimmung repräsentativer Konturen - Versuch

Beim Training eines Sets repräsentativer Konturfragmente (Kapitel 4.3.1) wird die Entscheidung darüber, ob eine Kontur repräsentativ für das Objekt Steinbutt ist, über den Schwellwert t_{train} geregelt. Eine repräsentative Kontur ist in dieser Arbeit durch folgende Punkte charakterisiert:

- Sie gehört zum äußeren Rand des Steinbutts und somit zur Beschreibung von dessen Form.
- Sie gehört nicht zur Musterung oder anderer visuell sichtbarer Strukturen auf der Oberfläche des Steinbutts.
- Sie gehört nicht zu einem Objekt, das sich neben den Steinbutten im Becken befindet, wie Schwebeteilchen oder auch Gerätschaften.

Sind die Trainingssets entsprechend den Anforderungen gewählt, sollten die Gesamtkosten $C(\gamma_{i_v})$ für repräsentative Konturen niedriger ausfallen, als für andere Konturen. Mit dem Schwellwert t_{train} wird jetzt festgelegt, bis zu welchem Wert eine Kontur als repräsentativ gilt. Ziel ist es, einen Schwellwert zu finden, der die Konturen in den Trainingsbildern effektiv in repräsentative Konturen und solche, die den Anforderungen nicht mehr entsprechen, aufteilt.

Dafür werden mehrere unterschiedliche Varianten von Trainingsdaten erstellt, wie in Kapitel 4.3.1.1 beschrieben wurde. Bei allen Varianten gleich bleibend ist das Trainingsset Negativ, das aus Bildern besteht, die keine Steinbutte enthalten. Dafür enthalten sie Elemente, die ebenfalls in den Aufnahmen auftreten können, wie Schwebeteilchen oder Hardwarekomponenten, die zum Becken gehören. Außerdem werden Bilder erzeugt, die zu vermeidende Strukturen enthalten, wie Kaustiken oder Musterungen der Steinbutte, da für die repräsentativen Konturen nur deren Außenlinie Verwendung finden soll. Vier Beispiele für Bilder des Trainingssets Negativ finden sich in Abbildung 4.11

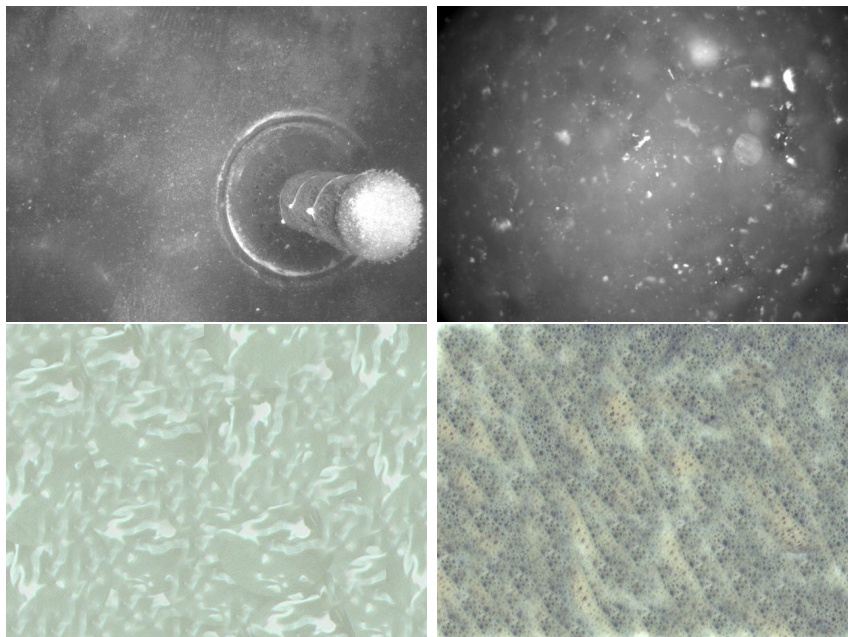


Abbildung 4.11: Beispiele vom Trainingsset Negativ. Oben links: Leeres Becken mit Equipment. Oben rechts: Schwebeteilchen. Unten links: Vervielfachte Kaustik. Unten rechts: Vervielfachte Oberflächenstruktur eines Steinbutts.

Für die erste Variante der Trainingsdaten werden verschiedene Aufnahmen von einzelnen Steinbutten unter Laborbedingungen für die Trainingsbilder sowie das Trainingsset Positiv verwendet. Die Steinbutte befinden sich dabei im Trockenem oder in einer dünnen Wasserschicht in einer hellen Schale und

werden für die Aufnahmen gleichmäßig ausgeleuchtet. Beispiele hierfür sind in Abbildung 4.12 zu finden.

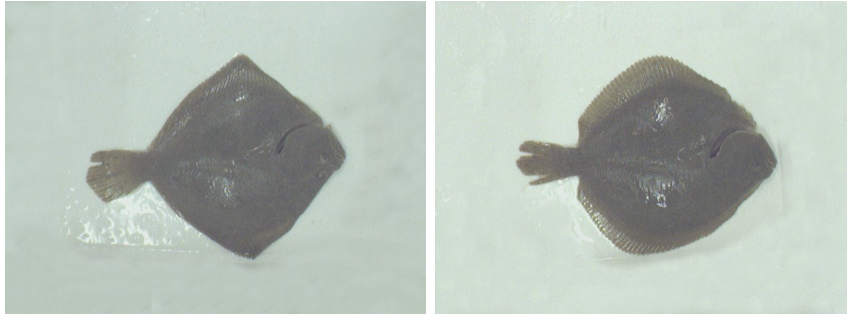


Abbildung 4.12: Beispiele für Trainingsdaten bestehend aus Laboraufnahmen von Steinbutten.

Die Varianten 2 bis 4 der Trainingsdaten verwenden für Trainingsbilder sowie das Trainingsset Positiv Aufnahmen aus Büsum unter realen Arbeitsbedingungen, siehe Abbildung 4.13 für Beispiele. Für die zweite und dritte Variante werden unterschiedliche Trainingsdaten erstellt. Die Trainingsbilder der vierten Variante stellen eine Kombination der Trainingsbilder aus der zweiten und dritten Variante dar, die die markantesten repräsentativen Konturen erzeugen.

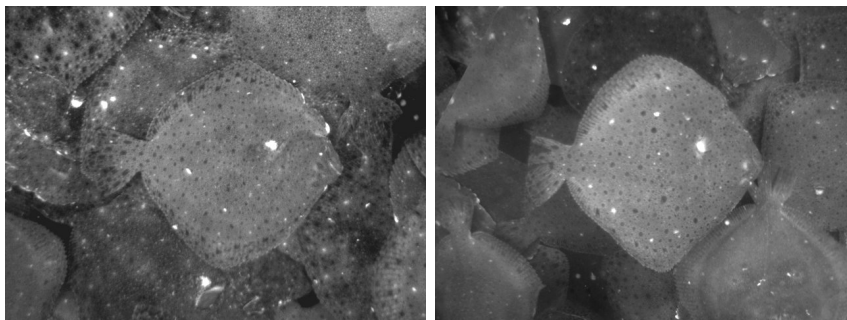


Abbildung 4.13: Beispiele für Trainingsdaten von Aufnahmen in Büsum. Für eine bessere Erkennbarkeit wurden die Bilder nachbearbeitet.

Als fünfte Variante der Trainingsdaten wird für die Trainingsbilder ein Set von Bildern ohne Steinbutte ähnlich dem Trainingsset Negativ erstellt. Das Trainingsset Positiv besteht wie bei der ersten Variante aus Steinbutt Bildern unter Laborbedingungen. Der Sinn hinter dieser Variante liegt darin, zu ermitteln, welche Gesamtkosten $C(\gamma_{i_v})$ für Konturen entstehen, die garantiert nicht zu einer Außenlinie eines Steinbutts gehören.

Für jede dieser 5 Varianten wird das Training zur Ermittlung eines Sets repräsentativer Konturen durchgeführt, wobei der Schwellwert t_{train} auf unendlich gesetzt wird. Jede einzelne Kontur sowie die für sie berechneten Gesamtkosten $C(\gamma_{i_v})$ werden gespeichert und anschließend ausgewertet.

4.4.2 Ermittlung des Schwellwertes zur Bestimmung repräsentativer Konturen - Ergebnisse

Zur Veranschaulichung der Ergebnisse dient Abbildung 4.14. Das Training von Variante 1 mit Laborbildern ergibt insgesamt 32 Konturen, die bewertet werden. Die niedrigsten Gesamtkosten liegen bei 5,5, die höchsten erreichen einen Wert von 72,6. Der Mittelwert beträgt 28,5 mit einer Standardabweichung von 20,5.

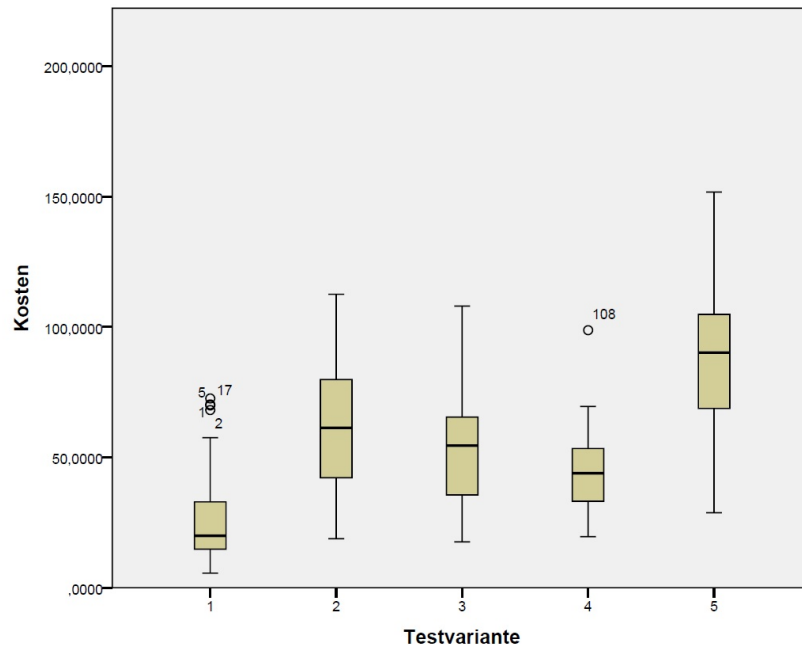


Abbildung 4.14: Testvariante 1: Training von Laboraufnahmen. Testvariante 2: Training von Aufnahmen in realer Umgebung, Variante 1. Testvariante 3: Training von Aufnahmen in realer Umgebung, Variante 2. Testvariante 4: Training von Aufnahmen in realer Umgebung, Auswahl bester Trainingsbilder. Testvariante 5: Training von negativen Strukturen.

Bei Variante 2 resultiert das Training in 30 Konturen. Das Minimum der Gesamtkosten beträgt 18,8, das Maximum liegt bei 112,5 bei einem Mittelwert von 63,1 und einer Standardabweichung von 26. Die Variante 3 erreicht

vergleichbare Resultate zu Variante 2. Insgesamt 30 Konturen für deren Gesamtkosten der Mittelwert 56,3 beträgt mit einer Standardabweichung von 25,2. Das erreichte Minimum liegt bei 17,6 und das Maximum bei 108.

Variante 4, bestehend aus Bildern mit den markantesten Konturen von Variante 2 und 3, enthält nur 20 Konturen, weist aber auch eine Verbesserung der Resultate auf. Der Mittelwert beträgt nur 45,2 und die Standardabweichung liegt bei 19. Das Minimum der Gesamtkosten beträgt 19,5, das Maximum ist 98,7.

Besonderes Interesse galt im Vergleich zu den 4 anderen Varianten der Variante 5, da in deren Trainingsbildern keine repräsentativen Konturen vorhanden sind. Somit stellen die bewerteten 41 Konturen Gegenbeispiele für repräsentative Konturen dar. Die Gesamtkosten resultieren in einem Mittelwert von 86,1 mit einer Standardabweichung von 28,7. Der niedrigste Wert liegt bei 28,7 und die höchsten Gesamtkosten werden mit einem Wert von 151,7 erreicht. Wie erwartet liegen die Werte für diese Variante höher als die Resultate der anderen Varianten und bestätigen, dass repräsentative Konturen niedrigere Gesamtkosten aufweisen als andere Konturen.

Tabelle 4.1: Quantile der Varianten 1 bis 5 bei entsprechenden $p_1 \dots p_5$.

$p_{1..4}/p_5$	Var 1	Var 2	Var 3	Var 4	Var 5
0,90/0,10	67,0	105,0	94,3	62,8	43,9
0,88/0,12	60,5	99,3	90,0	62,1	45,8
0,86/0,14	56,8	92,9	86,4	62,0	50,9
0,84/0,16	55,4	90,4	83,3	61,7	55,4
0,82/0,18	53,9	87,3	79,9	58,6	58,7
0,80/0,20	49,0	83,5	76,0	55,4	63,5
0,78/0,22	37,4	81,3	71,1	53,6	66,2
0,76/0,24	33,0	80,0	65,8	53,3	68,0
0,74/0,26	31,7	78,9	65,3	53,1	70,3
0,72/0,28	31,0	77,8	65,1	52,8	73,1
0,70/0,30	30,6	76,2	65,0	52,6	74,8
0,68/0,32	30,6	75,0	63,7	51,9	75,4

Für die Ermittlung eines Schwellwertes werden die Ergebnisse miteinander verglichen. Da die Trainingsbilder der Varianten 1 bis 4 auch Konturen enthalten können, die nicht repräsentativ sind, muss der Schwellwert unterhalb des Maximums der jeweiligen Variante liegen. Variante 5 enthält keine repräsentativen Konturen, jedoch ist es möglich, dass beim Training der Konturen niedrige Gesamtkosten entstehen, weswegen der Schwellwert nicht unterhalb

des Minimums dieser Variante liegen sollte. Die hier gewählte Lösung ist eine gleichmäßige Annäherung zwischen den niedrigen Gesamtkosten der Variante 5 und den hohen Gesamtkosten der jeweiligen Variante 1 bis 4.

Hierfür wird in dieser Arbeit das sogenannte *Quantil* Q_p mit $0 \leq p \leq 1$ verwendet, das jenes Element einer Ergebnismenge beschreibt, für das ein Anteil der Ergebnisse von p unterhalb dieses Elements liegt. So beschreibt das Quantil $Q_{0,25}$ diejenige Kontur und deren Gesamtkosten, für die 25 % der Ergebnisse niedrigere Gesamtkosten haben. Beginnend mit einem niedrigen Quantil und einer geringen Gesamtkost für Variante 5 und einem hohen Quantil und entsprechend hohen Gesamtkosten für die anderen Varianten, werden die Quantile schrittweise angenähert, bis sich die Gesamtkosten überschneiden. In diesem Wertebereich wird der Schwellwert t_{train} festgelegt. Die Ergebnisse hierfür werden in Tabelle 4.1 aufgeführt. Auf diese Weise erhält man für Variante 1 den Schwellwert $t_{train1} = 56$, für Variante 2 $t_{train2} = 75$, für Variante 3 $t_{train3} = 68$ und für Variante 4 $t_{train4} = 58$.

Mit steigenden Schwellwerten wachsen auch entsprechend die Chancen, dass nicht repräsentative Konturen als solche eingestuft werden. Da die Varianten 2 und 3 auffällig höhere Schwellwerte als Variante 4 aufweisen, die eine Kombination der beiden darstellt, ist Variante 4 gegenüber den anderen beiden Varianten vorzuziehen und wurde im weiteren Verlauf dieser Arbeit als Trainingsdatensatz eingesetzt.

Abschließend wird für die Varianten 1 und 4 eine visuelle Sichtung der Konturen und deren Gesamtkosten durchgeführt. Ziel ist es, zu überprüfen, ob die auf diese Weise ermittelten Schwellwerte für die verschiedenen Varianten von Trainingsdaten eine effektive Beurteilung von repräsentativen Konturen ermöglicht. Die Schwellwerte und deren Zuweisung von repräsentativen Konturen kann dabei bestätigt werden und werden für die weitere Verwendung in dieser Arbeit eingesetzt.

4.4.3 Erkennungsraten des Konturen-Matchings - Versuch

Nach dem Training eines Sets von repräsentativen Konturen werden diese verwendet, um in unbekanntem Bildmaterial an vergleichbare Konturen gematcht zu werden, wodurch eine Erkennung von Steinbutten erfolgen soll. Zur Überprüfung der Korrektheit der gematchten Konturen wird die folgende Evaluation durchgeführt.

Hierfür werden 4 unterschiedliche Sets repräsentativer Konturen verwendet, um zu ermitteln, ob sich eine der Varianten beim Matching gegenüber den anderen Varianten als vorteilhaft herausstellt indem sie in einer höheren Rate an korrekt erkannten Konturen resultiert.

Für die ersten beiden Sets werden die aus 4.4.2 bekannten Varianten 1, bestehend aus Laborbildern, und 4, bestehend aus ausgewählten Bildern aus Büsum, verwendet.

Für die weiteren Sets erfolgt die Erstellung von einem Datensatz aus 6 Bildern, die pro Bild nur die synthetischen von Hand erzeugten Konturen eines Steinbutts enthalten, ohne weitere nicht verwendbare Konturen. Dafür werden die Konturen von Steinbutten per Hand nachgezeichnet und in mehrere Abschnitte unterteilt. Zwei Beispiele hierfür finden sich in Abbildung 4.15. Mit diesem synthetischen Datensatz wird das Training zweimal durchgeführt, einmal mit dem Trainingsset Positiv mit Laborbildern und einmal mit dem Trainingsset Positiv mit Bildern aus Büsum. Dabei wird auf die Verwendung eines Schwellwertes t_{train} verzichtet, da sämtliche Konturen in diesen Datensätzen zu Steinbutten gehören und somit keine Trennung von repräsentativen und nicht repräsentativen Konturen erfolgen muss. Die 2 dabei entstehenden Sets repräsentativer Konturen sind die Sets 3 und 4 für diesen Versuch.

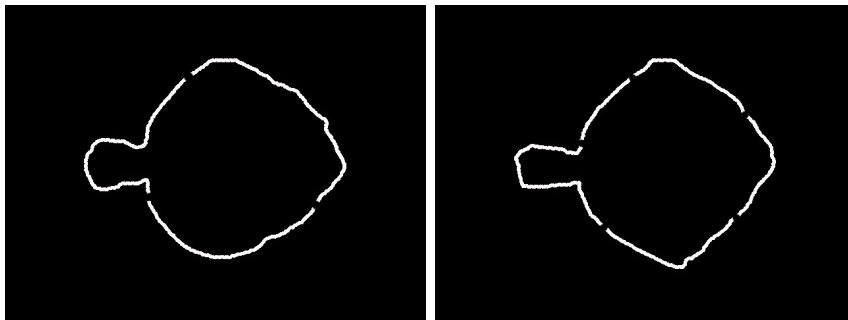


Abbildung 4.15: Zwei Beispiele der synthetischen Steinbuttkonturen. Die Konturen wurden für diese Darstellung verstärkt.

Alle 4 Sets repräsentativer Konturen werden mit 2 Datensätzen von je 14 positiven Validierungsbildern getestet. Der erste Datensatz besteht aus Laborbildern, wie sie auch für das Training verwendet wurden. Der zweite Datensatz besteht aus ausgewählten Bildern aus Büsum, die eine möglichst hohe Qualität und einen nach menschlichen Maßstäben gut erkennbaren Steinbutt aufweisen, um die Erkennung nicht zusätzlich zu erschweren.

Das in 4.3.3 beschriebene Matching der Konturen wird für jedes der Validierungsbilder durchgeführt. Für jede Kontur werden die Matching-Kosten erfasst. Darüber hinaus wird die Distanz zwischen dem gematchten Mittelpunkt der repräsentativen Kontur und dem hierfür zuvor ermittelten Mittelpunkt des Steinbutts im Validierungsbild berechnet. Diese Distanz wird als Anhaltspunkt verwendet, um die korrekte Skalierung der gematchten Konturen zu überprüfen. Bei korrekter Positionierung und Skalierung ist die Distanz zwi-

schen den Mittelpunkten minimal. Umso fehlerhafter die Skalierung ist, desto größer wird die Distanz zwischen den Mittelpunkten. Unter der Annahme einer korrekten Positionierung wird für diesen Versuch als Grenzwert für eine ausreichend genaue Skalierung eine Distanz von maximal 50 Pixeln festgelegt.

Die Überprüfung auf die korrekte Auswahl der gematchten Kontur durch das Matching-Verfahren erfolgt anhand visueller Sichtung der Ergebnisse. Wird die repräsentative Kontur an eine Stelle des Steinbutts gematcht, der sie selbst entstammt, wird dies als korrekte Positionierung gewertet. Sobald das Ergebnis für die Position positiv ausfällt, wird anschließend die Skalierung anhand der Distanz zwischen den Mittelpunkten geprüft.

4.4.4 Erkennungsraten des Konturen-Matchings - Ergebnisse

Abbildung 4.16 zeigt 2 Ergebnisse für das Konturen-Matching. Die Laborbilder weisen im Vergleich zu den Bildern aus Büsum erheblich weniger Konturen auf, für die ein Matching durchgeführt wird. Die 14 Validierungsbilder aus Laborbildern bestehen aus insgesamt 2433 Konturen, von denen nach Aussortierung kurzer und geschlossener Konturen 134 übrig bleiben, die überprüft werden müssen. Bei den Validierungsbildern aus Büsum bleiben von insgesamt 10126 Konturen 436 zu überprüfende Konturen übrig.

Diese Konturen gehören nicht alle zu Steinbutten und werden bei der visuellen Sichtung dahingehend sortiert, ob sie zu einer Kontur des zu erkennenden Steinbutts gehören und somit für die Fischerkennung relevant sind, oder ob sie Konturen aus der Umgebung sind. Bei den Laborbildern sind dies 66 und bei den Bildern aus Büsum 72 relevante Konturen.

Tabelle 4.2 zeigt, dass bei der Validierung mit Laborbildern nur circa 40 bis 54 % der relevanten Konturen korrekt gematcht werden. Eine ausreichend gute Skalierung weisen nur circa 7 bis 14 % der relevanten Konturen auf. Ausgehend von 14 Validierungsbildern ist dies weniger als eine korrekt gematchte und skalierte Kontur pro Bild. Tabelle 4.3 zeigt für die Validierung mit Bildern aus Büsum noch niedrigere Ergebnisse. Während noch circa 48 bis 56 % der relevanten Konturen gematcht werden, liegt die Quote für die korrekte Skalierung bei circa 1- bis 6 %.

Aufgrund der unzureichenden Ergebnisse beim Matching der repräsentativen Konturen an die Konturen der Validierungsbilder ist eine weiterführende Bearbeitung zur Erkennung von Steinbutten unter realen Arbeitsbedingungen in der Aquakultur nicht möglich, weswegen für die in den Kapiteln 4.3.4 und 4.3.5 beschriebenen Folgeschritte vorerst keine weiteren Untersuchungen durchgeführt werden.

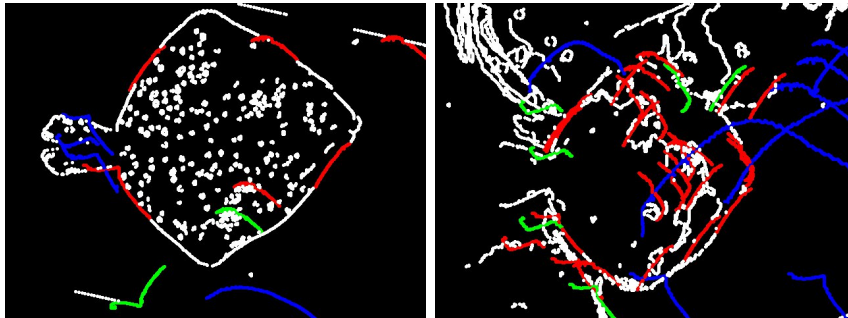


Abbildung 4.16: Ergebnisbilder des Konturen-Matchings. Links: Bild aus Validierungsset Labor. Rechts: Bild aus Validierungsset Büsum. Gematchten Konturen werden angezeigt und farbcodiert. Konturen des Validierungsbildes: weiß. Niedrige Matching-Kosten ($c_{match} < 0,3$): rot. Mittlere Ergebnisse ($c_{match} < 0,4$): grün. Hohe Ergebnisse ($c_{match} \geq 0,4$): blau. Konturen für Darstellung verstärkt.

Tabelle 4.2: Anzahl korrekt gematchter und skaliertes Konturen von 66 relevanten Konturen. Validierung mit Laborbildern. Labor: Trainingsdaten mit Laborbildern. Büsum: Trainingsdaten mit Bildern aus Büsum. Synt1: Synthetische Bilddaten trainiert mit Laborbildern. Synt2: Synthetische Bilddaten trainiert mit Bildern aus Büsum.

	Labor	Büsum	Synt1	Synt2
Korrekt gematchte Konturen	36	28	34	27
Korrekt skalierte Konturen	9	5	8	9

Tabelle 4.3: Anzahl korrekt gematchter und skaliertes Konturen von 72 relevanten Konturen. Validierung mit Bildern aus Büsum. Labor: Trainingsdaten mit Laborbildern. Büsum: Trainingsdaten mit Bildern aus Büsum. Synt1: Synthetische Bilddaten trainiert mit Laborbildern. Synt2: Synthetische Bilddaten trainiert mit Bildern aus Büsum.

	Labor	Büsum	Synt1	Synt2
Korrekt gematchte Konturen	35	40	38	37
Korrekt skalierte Konturen	3	4	1	1

KAPITEL 5

Diskussion

Inhalt

5.1	Entwicklung Hardwaresystem	109
5.2	Algorithmus zur Steinbuttvermessung	113

5.1 Entwicklung Hardwaresystem

Zu Beginn der Entwicklung wurde eine Anforderungsanalyse erstellt und anhand dieser ein flexibel konfigurierbares Hardwarekonzept erstellt. Nach dessen Realisierung konnten nach durchgeführten Probeaufnahmen noch weitere Verbesserungen am Hardwaresystem vorgenommen werden. Nach Auswertung der Untersuchungen zur Farbe der Beleuchtung erfolgte die Fertigstellung der Endfassung des Hardwaresystems. Mit dieser Version wurde die Untersuchung zur Ausrichtung der LEDs für eine möglichst homogene Ausleuchtung der Szenerie durchgeführt. Parallel zu den Untersuchungen zum Beleuchtungssystem konnte die Kalibrierung des Stereokamerasystems durchgeführt werden, da diese unabhängig vom Beleuchtungssystem ist.

Das entwickelte Hardwaresystem, bestehend aus dem Stereokamerasystem in einem wasserdichten Gehäuse, dem wasserdichten Beleuchtungssystem, der Steuerelektronik und der Recheneinheit, erwies sich in der praktischen Anwendung als robust, siehe hierzu Kapitel 3.2.2. Das Hardwaresystem ist in der Lage, über längere Zeiträume den Umweltbedingungen einer Aquakulturanlage ausgesetzt zu sein und wäre somit auch für dauerhafte Installationen geeignet. Wichtig ist hierbei jedoch eine regelmäßige Reinigung des Gehäuses von Salzablagerungen sowie Verschmutzungen.

Am Hardwaresystem positiv zu werten sind die Möglichkeiten zur Anpassung des Systems. Die Montage der Stereokameras ist variabel und ermöglicht eine Veränderung von Baseline sowie Blickrichtung der Kameras, während die Gehäusegröße unterschiedliche Objektive zulässt, wodurch eine Anpassung an andere Einsatzgebiete in der Aquakultur möglich ist. Die Steuerelektronik ermöglicht den Einsatz einer dritten Kamera. Diese kann zum Beispiel dafür genutzt werden, um die Stereoerfassung auszubauen, einen weiteren Bereich

zu überwachen oder die Szenerie aus einer anderen Perspektive zu überwachen. Das Beleuchtungssystem ist derart modular gestaltet, dass die Anzahl der LED-Gehäuse variabel den Bedürfnissen angepasst werden kann, während jedes einzelne LED-Gehäuse im Abstrahlwinkel einstellbar ist. Darüber hinaus lassen sich die LEDs mit relativ geringem Arbeitsaufwand gegen andere Varianten austauschen.

Das Kameragehäuse mit montiertem Beleuchtungssystem sowie die Steuerelektronik im Edelstahlgehäuse sind für eine einzelne Person relativ sperrig zu handhaben, nach der Montage über dem Fischbecken ist dies jedoch kein weiterer Störfaktor. Die stabile Bauweise von Kameragehäuse und Beleuchtungssystem hat sich sogar als vorteilhaft herausgestellt. Wenn, zum Beispiel während der Fütterung, die Steinbutte in Aufregung versetzt werden, kann es zu Zusammenstößen mit dem Stereokamerasystem kommen, die aufgrund der robusten Konstruktion keine Auswirkungen auf die Systemeinstellungen und die Kalibrierung hatten.

Bei der Entwicklung des Hardwaresystems wurde anhand der Untersuchung von vier Parametern überprüft, ob eine bestimmte Farbe des Lichts Vorteile bei der Objekterkennung bietet, zum Beispiel in Form einer besseren Hervorhebung von Konturen und Strukturen. Neben der Farbe des Lichts wurde auch untersucht, ob eine bestimmte Färbung des Materials der Beckenuntergründe Vorteile haben kann.

Für die zu überprüfenden Farben des Lichts wurden die Grundfarben rot, grün, blau und gelb neben weiß gewählt, um so die verschiedenen Bereiche des sichtbaren Lichtspektrums abzudecken. Dafür wurde eine spezielle LED-Halterung entwickelt, bei der der Abstand zwischen den farbigen LEDs sehr gering ist, um eine möglichst gleichartige Ausleuchtung der Szene zu gewährleisten. Aufgrund dieser Konstruktion konnte bei den Untersuchungen sehr schnell zwischen den einzelnen Farben gewechselt werden, wodurch es, abgesehen von der Beleuchtung, zu keiner Veränderung der Szene kam und eine Vergleichbarkeit der Ergebnisbilder gewährleistet wurde.

Anfangs wurden versuchsweise noch ultraviolettes und infrarotes Licht verwendet, deren Resultate aber sehr unbefriedigend ausfielen. Die schlechten Ergebnisse liegen hauptsächlich darin begründet, dass Wasser hohe Absorptionsraten für die Frequenzbereiche von ultraviolettem und im besonderen Maße infrarotem Licht aufweist [57]. Um für die verschieden farbigen LEDs eine Vergleichsbasis zu schaffen, wurden für die Untersuchungen LEDs mit gleicher Leistungsaufnahme verwendet. Bei den Farben für die Becken wurden die im Bereich der Aquakultur typischen Materialfarben weiß, grau, schwarz und grün gewählt.

Die Untersuchung zum relativen Kontrast zwischen den Steinbutten und dem Untergrund zeigte für keine Kombination von Licht, Untergrund und

Fischmusterung signifikante Vorteile gegenüber den anderen Kombinationen. Es hat sich jedoch gezeigt, dass bei der vorliegenden Besatzdichte der Untergrund praktisch nicht zu sehen ist, wodurch die Überprüfung des relativen Kontrastes zwischen Steinbutt und Untergrund im Nachhinein als nicht optimal gewählt erscheint.

Die Untersuchung zeigte jedoch Ergebnisse, die für den Ausblick mögliche Optionen offenbarten. So sind die Resultate von grünem Untergrund bei rotem und grünem Licht abweichend von den anderen Ergebnissen. Aufgrund der fast vollständigen Absorption von rotem Licht sind die Ergebnisse vergleichbar mit denen eines schwarzen Untergrundes. Die starke Reflexion von grünem Licht wiederum bewirkt, dass die Ergebnisse vergleichbar mit denen eines hellen Untergrundes sind.

Die durch die Ergebnisse der Untersuchungen zu erwartenden starken Kontraste zwischen hellem Untergrund und dunklen Steinbutten beziehungsweise der umgekehrten Kombination werden in der Realität praktisch nicht erreichbar sein, da sich die Steinbutte der Umgebung anpassen. Das heißt, dass sie in einem hellen Becken ihre Musterung aufhellen und in einem dunklen Becken abdunkeln werden. Auch hier zeigt sich eine Besonderheit des grünen Untergrunds, da sich die Steinbutte an diesen Farbton nicht anpassen können, abgesehen von einer eventuellen Anpassung der Steinbutte innerhalb ihres grau-braunen Farbspektrums an die Helligkeit des Grüntons.

Unter dem Gesichtspunkt der Anpassungsfähigkeit der Steinbutte, wäre bei einem Monitoring-System, bei dem der Kontrast zwischen Boden und Steinbutt für die Ermittlung einer Kontur effektiv genutzt werden könnte, die Verwendung von einem grünen Untergrund zu empfehlen. Zum einen wegen der fehlenden Anpassung an den Farbton und zum anderen erscheint es realistisch, dass durch den Einsatz von rotem oder grünem Licht gut verwendbare Kontraste zwischen den Steinbutten und dem Untergrund erzeugt werden können. Denkbar wäre auch der Einsatz von rotem und grünem Licht, wo bei minimal möglichem Zeitabstand zwei Aufnahmen hintereinander, einmal mit rot und einmal mit grün beleuchtet, werden. So erhält man zwei Bilder, bei denen der Hintergrund einmal dunkel und einmal hell ist. Wenn Belichtungszeit sowie zeitlicher Abstand zwischen den Aufnahmen gering genug sind, sollte die Differenz der Bilder durch Bewegung der Steinbutte minimal gehalten werden können und ein starker Kontrast zwischen Steinbutt und Untergrund kann errechnet werden.

Für die Überprüfung, ob eine der Kombinationen von Licht, Untergrund und Fischmusterung die vorhandenen Strukturen und Konturen stärker hervorhebt als die anderen Kombinationen, wurden als Parameter der empirische Variationskoeffizient sowie die Gradientenstärke gewählt. Der empirische Variationskoeffizient diente dabei als Maßstab, wie sehr die Grauwerte im Be-

reich der Steinbutte insgesamt variieren. Da aufgrund des Versuchsaufbaus ein gleichmäßiger Verlauf vom Hellen ins Dunkle innerhalb des untersuchten Bereichs ausgeschlossen werden konnte, was ansonsten auch zu einem hohen empirischen Variationskoeffizienten führt, konnten die Ergebnisse deshalb als Indiz für das Vorhandensein von auffindbaren Strukturen herangezogen werden. Die Gradientenstärke diente als Anhaltspunkt dafür, wie stark ausgeprägt die gefundenen Konturen sind, wodurch maßgeblich das Vorhandensein von Kanten und Linien, die gefunden werden sollten, beschrieben wird. Bei den Untersuchungen stellte sich heraus, dass keine der Farben einen signifikanten Vorteil gegenüber den anderen aufweisen konnte. Aus diesem Grund wurde für das Beleuchtungssystem auf weiße LEDs zurückgegriffen. Diese erzielten bei der Überprüfung der Relation von Belichtungszeit und mittlerem Grauwert des Bildes das beste Verhältnis zwischen Leistungsaufnahme und Helligkeit, unabhängig von der Farbe des Untergrundes.

Um bei den Untersuchungen zum Beleuchtungssystem die Anpassungsfähigkeit der Steinbutte ansatzweise zu simulieren, wurden Exemplare mit sehr heller und sehr dunkler Musterung verwendet. Erste Probeaufnahmen mit lebenden Steinbutten stellten sich aufgrund deren Bewegung als problematisch heraus, weswegen für die durchgeführten Versuche auf tote Exemplare zurückgegriffen wurde, um möglichst vergleichbare Ergebnisse zu erzielen. Da die gesamten Aufnahmen innerhalb eines kurzen Zeitraums durchgeführt wurden, kam es zu keinen Verwesungsprozessen bei den toten Steinbutten, die die Vergleichbarkeit der Ergebnisse negativ beeinflussen konnten.

Für eine funktionierende Bildverarbeitung muss ein Mindestmaß an Helligkeit im Bildmaterial vorhanden sein. Sind die Bilder dunkel, gehen Details verloren, die bei der Erkennung von Steinbutten hilfreich sein könnten. Dies kann durch eine höhere Lichtmenge von der Beleuchtung oder eine längere Belichtungszeit erreicht werden. Um die Bewegungsunschärfe der Steinbutte zu minimieren, muss jedoch eine möglichst kurze Belichtungszeit gewählt werden. Bei der in dieser Arbeit maximal verwendeten Belichtungszeit von circa 20 ms waren bereits mit bloßem Auge Unschärfen zu erkennen, falls es zu schnellen Bewegungen der Steinbutte kam. Eine Belichtungszeit von maximal 10 ms wäre erstrebenswert, um die Bewegungsunschärfen so gering wie möglich zu halten. Entsprechend hoch muss dann die Lichtmenge gewählt werden. Die natürliche Grenze sollte der Steinbutt mit seinen Ansprüchen an die Umgebung sein, worauf im Ausblick in Kapitel 6.1 weiter eingegangen wird.

Es war bereits während der Planung des Hardwaresystems bekannt, dass Steinbutte lichtscheue Tiere sind, die ein Minimum an zusätzlicher Beleuchtung bevorzugen. Aus diesem Grund wurde die Steuerelektronik sowie das Beleuchtungssystem derart entwickelt, dass die Möglichkeit bestand, das Licht durch die Kameras zu triggern. Dabei sollte das Beleuchtungssystem wie ein

Blitzlicht für jedes Einzelbild über die Dauer der Belichtungszeit aktiviert werden, um so die Gesamtmenge des Lichts zu reduzieren. Beim versuchsweisen Einsatz des getriggerten Lichts entstand bei den für den späteren Einsatz zu erwartenden 25 bis 30 Bildern pro Sekunde und einer Belichtungszeit zwischen 10 und 20 ms ein Stroboskop-Effekt. Dieser führte bei den Steinbutten nach einiger Zeit zu einer Stressreaktion, so dass sie das aktive Beleuchtungssystem angriffen, es rammten und nach den LED-Gehäusen bisßen. Daraufhin musste für den weiteren Verlauf der Arbeit auf getriggertes Licht verzichtet werden.

Die Kalibrierung des Stereokamerasystems (Kapitel 3.3.3) zeigte eine ausreichende Genauigkeit für die Vermessung von Steinbutten mit einer Messungenauigkeit von unter 2 % (Kapitel 3.4.3). Ein solcher Fehler bedeutet bei einer Fischlänge von zum Beispiel 30 cm eine Ungenauigkeit von bis zu 0,6 cm. Eine normale Vermessung von Hand ist aufgrund der Beweglichkeit der Steinbutte und der flexiblen Schwanzflosse ähnlich präzise. Zum Vergleich sei hier aufgeführt, dass Testmessungen des Verfahrens in [141] ebenfalls eine Abweichung von circa 1 - 2 % der realen Objektängen aufwiesen, während das Vergleichsverfahren Ungenauigkeiten von 10 - 15 % aufwies.

Zu beachten ist hier die Hardwarekonfiguration mit paralleler Ausrichtung der Kameras und minimiertem Abstand zu einer ebenen Sichtscheibe. Durch diesen Aufbau wurde zwischen einer Kalibrierung an der Luft und einer unterwasser ein Verhältnis der errechneten Brennweiten von annähernd 1,33 erreicht. Nach [141] ist in einem solchen Fall die Anwendung von Standardverfahren zur Stereokalibrierung möglich. So konnte eine frei erhältliche Software zur Berechnung der Kalibrierung verwendet werden, anstatt das Verfahren zur Berechnung selbst implementieren zu müssen. Für eine andere Konfiguration, bei der die Voraussetzungen mit parallel ausgerichteten Kameras und minimalem Abstand zur Sichtscheibe nicht erfüllt werden können, muss davon ausgegangen werden, dass ein anderes Kalibrierverfahren benötigt wird, um unter diesen Bedingungen ausreichend genaue Messergebnisse zu erhalten. Die Verfahren aus [141] und [122] erscheinen für diesen Fall bestens geeignet.

5.2 Algorithmus zur Steinbuttvermessung

Der Algorithmus zur Erkennung von Steinbutten und deren Vermessung wurde mit C++ sowie OpenCV, einer frei erhältlichen Bildverarbeitungs-Bibliothek, entwickelt. Die Wahl wurde aufgrund von Vorkenntnissen in der Programmiersprache sowie der vorhandenen Flexibilität bei der Entwicklung getroffen.

Die verwendete Vorverarbeitung, bestehend aus kantenerhaltender Rauschunterdrückung durch einen separierten nichtlinearen Gaußfilter, einer Histogrammäqualisierung zur Verstärkung der Kontraste und dem Einsatz des Can-

ny Edge Detectors zur Extraktion der Konturen, zeigte sich in der Anwendung sehr abhängig von den vorliegenden Lichtbedingungen und der Verunreinigung des Wassers. Abhängig von diesen Faktoren war eine jeweilige Anpassung der Parameter für die Rauschunterdrückung und die Konturextraktion nötig. So gab es keine allgemein anwendbaren Einstellungen, die für die unterschiedlichen auftretenden Aufnahmebedingungen durchgängig verwendbare Ergebnisse lieferten.

Als Lösung für die benötigte Anpassung der Filter-Parameter ist denkbar, ein zusätzliches Verfahren zu entwickeln, das das Bildmaterial zuerst analysiert und anhand der Ergebnisse automatisiert die Parameter für den separierten nichtlinearen Gaußfilter sowie die Schwellwerte für den Canny Edge Detector setzt.

Bei Betrachtung der erzielten Ergebnisse aus dem Konturen-Matching fällt auf, dass die vorliegenden Konturen sich in vielen Fällen als generische Formen herausstellen, oft bestehend aus nur einem Bogen oder Geraden mit leichten Kurven. Dies trifft auf die Bilder zu, in denen die Steinbutte gefunden werden sollen, den Validierungsbildern, als auch auf den überwiegenden Teil der Bilder, die zum Training der repräsentativen Konturen für den Algorithmus zur Erkennung von Steinbutten verwendet wurden, seien es Laborbilder von Steinbutten, Bilder unter realen Bedingungen aus Büsum oder synthetisch erzeugte Konturen. Ausnahmen existieren bei den repräsentativen Konturen, die mit Laborbildern trainiert wurden. Hier sind teilweise Konturen vorhanden, die bis zu drei Viertel eines Steinbutts abdecken. Solche Konturen haben aber den Nachteil, dass sie bei Bildern unter realen Bedingungen praktisch nicht auftreten und kürzere Konturen bessere Ergebnisse beim Matching erzielen.

Die generische Formgebung der Konturen ist ebenfalls die Ursache für die fehlerhafte Skalierung im Matching-Prozess, da eine signifikante Formgebung fehlt, an der die Skalierung angepasst werden kann. So wird in diesem Fall eine relativ gerade Linie an eine andere relativ gerade Linie angepasst, bei der für die Skalierung keine Anhaltspunkte über die reale Größe des Konturfragments vorliegen.

Die fehlerhafte Skalierung resultiert in einer ungenauen Positionierung der potentiellen Steinbutt-Mittelpunkte. Um diese Mittelpunkte weiterhin gruppieren zu können, um somit deren Zugehörigkeit zu einem Steinbutt zu bestimmen, würde ein sehr großer Suchradius benötigt werden. Hierdurch wird jedoch das Kriterium, wonach dicht beieinander liegende potentielle Mittelpunkte einen Steinbutt repräsentieren, ignoriert und Konturen, die nicht zum Steinbutt gehören, der Gruppierung hinzugefügt.

Die Notwendigkeit bei der Erkennung von Steinbutten ein Verfahren basierend auf Konturfragmenten anzuwenden, entstand aus den gegebenen Anforderungen an das System. Fast sämtliche in der Recherche untersuchten

Verfahren zur Erkennung von Objekten basieren darauf, dass eine das Objekt vollständig umschließende Kontur vorliegt, was bei den in dieser Arbeit vorherrschenden Arbeitsbedingungen praktisch nicht erreicht werden konnte, wie die Ergebnisse des Konturen-Matchings bestätigen.

Die Ergebnisse aus der Evaluation des entwickelten Verfahrens zur Erkennung von Steinbutten sowie die Erkenntnisse aus der anschließenden Analyse führen zu dem Schluss, dass eine robuste Erkennung und Vermessung von Steinbutten unter den gegebenen Bedingungen nicht durchgeführt werden kann.

Im Vergleich mit den Arbeiten [97] und [98], auf denen das für diese Arbeit entwickelte Verfahren grundlegend basiert, können für die erfolgreiche Erkennung entscheidende Unterschiede festgemacht werden. So bestehen die verwendeten Objekte wie Kühe, Pferde oder Motorräder aus komplexeren Konturen als ein Steinbutt, wodurch auch die Konturfragmente, solange sie eine bestimmte Länge nicht unterschreiten, komplexere Formen annehmen. Des Weiteren befinden sich die gesuchten Objekte im Regelfall vor einem anders gearteten Hintergrund, wodurch sich das Objekt stärker von seiner Umgebung abhebt als ein getarnter Steinbutt von anderen getarnten Steinbutten.

Neben den beeinträchtigten Sichtbedingungen, entstanden durch die Anforderung der Steinbutte an geringe Helligkeit sowie der generellen Verschmutzung des Wassers, war die hohe Besatzdichte der Becken das Hauptproblem, da bei Besatzdichten von bis zu 100 kg/m^2 die Steinbutte in bis zu 5 Lagen übereinander liegen und das gesamte Sichtfeld der Aufnahme vollständig mit Steinbutten ausgefüllt ist.

Hier wird der grundlegende Unterschied zwischen dem Ziel dieser Arbeit, der Erkennung von Steinbutten vor einem Hintergrund bestehend aus Steinbutten, und Einsatzgebieten der Objekterkennung in Forschung und Industrie sichtbar. In Forschung und Industrie hebt sich das gesuchte Objekt von einem Hintergrund ab, der sich in Farbe, Form oder Textur vom Objekt unterscheidet. In dieser Arbeit bestehen Vorder- als auch Hintergrund jedoch aus den gleichen zu identifizierenden Objekten. Die Unterscheidung zwischen einem Objekt im Vordergrund und Objekten im Hintergrund ist hier lediglich dadurch definiert, dass ein Steinbutt im Vordergrund vollständig erkennbar ist. Das Tarnmuster und die Camouflage-Fähigkeiten der Steinbutte erschweren die Erkennung von einzelnen Steinbutten zusätzlich zu der hohen Besatzdichte, da ihre Funktion darin besteht, mit dem Hintergrund zu verschmelzen und sich somit einer Erkennung auf visueller Basis zu entziehen.

Verfahren zur Erkennung von Fischen und in den meisten Fällen auch deren Vermessung, die ihre Aufgabe erfüllten, wurden in Kapitel 2.3.1 vorgestellt. Bei diesen Verfahren ermöglichten die Rahmenbedingungen die Erkennung einer vollständigen Kontur und somit die Anwendung von entsprechen-

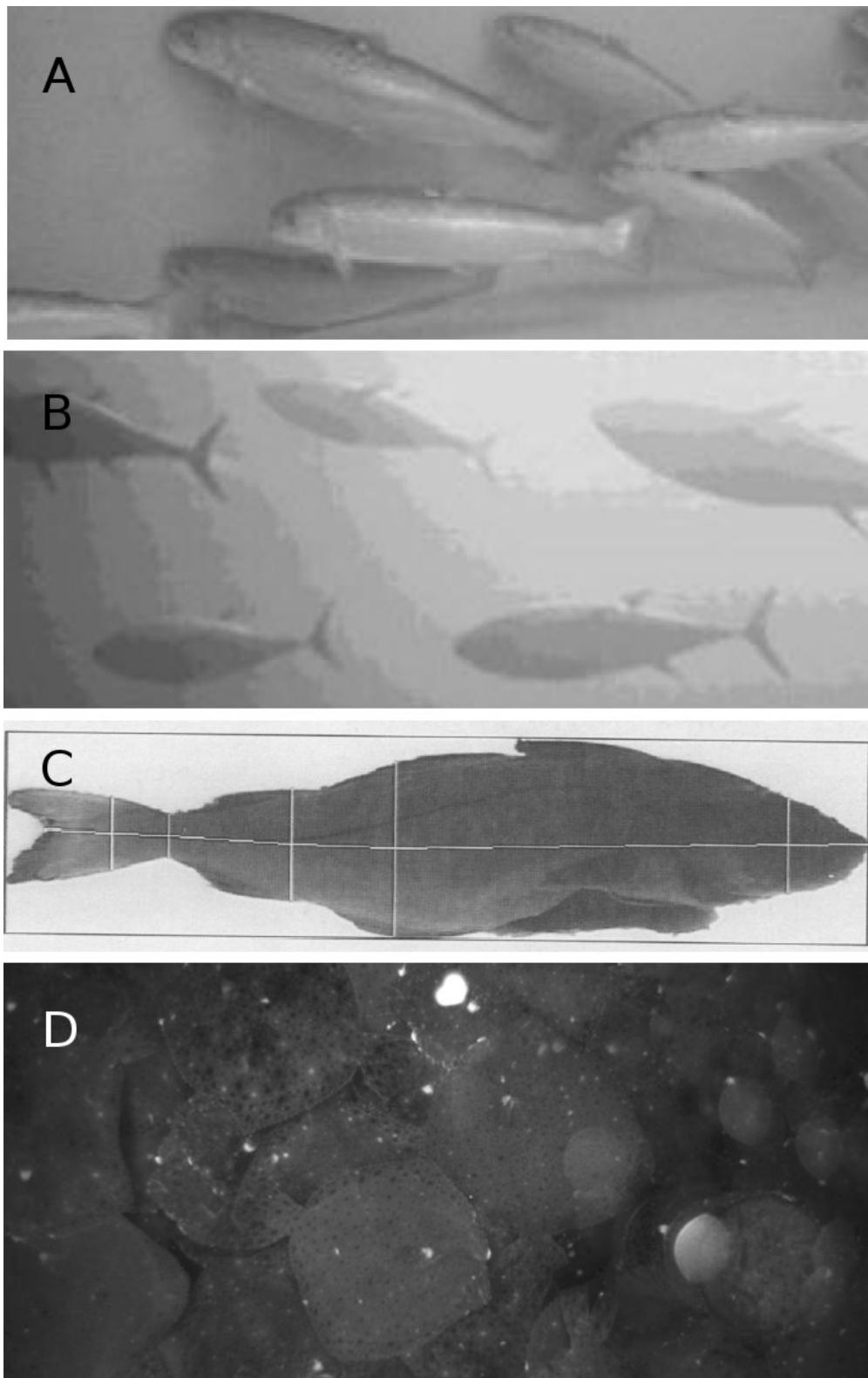


Abbildung 5.1: Beispiele für Aufnahmen für Verfahren zur Fischerkennung. A: Lachs in Becken aus [140]. B: Thunfische in offener See aus [32]. C: Fisch auf Fließband aus [134]. D: Steinbutte in Becken der GMA Büsum.

den Verfahren zur Objekterkennung. Zum Vergleich wurden mehrere Beispielfelder aus verschiedenen Arbeiten in Abbildung 5.1 zusammengefasst.

In [140] und [86] wird ein Verfahren zur Erkennung und Vermessung von frei schwimmenden Lachsen in einem Becken vorgestellt. Wie in Abbildung 5.1 A gut zu erkennen, heben sich die Lachse gut von der Rückwand des Beckens ab. Bei der Anpassung des in dem Verfahren verwendeten Fisch-Modells traten Probleme auf, sobald sich weitere Lachse in unmittelbarer Nähe befanden. Die in Kapitel 2.3.1 beschriebenen Ergebnisse lassen auf eine Erkennungsrate von circa 10 % schließen, ausgehend von nur einem Lachs pro Bild. Da das veröffentlichte Bildmaterial stets mehr als einen Lachs aufwies, muss von einer wesentlich geringeren Erkennungsrate ausgegangen werden. In Anbetracht der Tatsache, dass die Untersuchung in einem Becken mit schwarzem Hintergrund zur Optimierung des Kontrastes durchgeführt wurde und das reale Einsatzgebiet ein Käfigsystem in offenem Gewässer ohne Rückwand wäre, muss sich die Frage gestellt werden, wie effektiv das Verfahren unter solchen Bedingungen funktionieren würde.

Ein Verfahren, das bei der Vermessung von Thunfischen bei der Überführung von Fangnetzen in Zuchtkäfige unter realen Bedingungen arbeiten musste, wird in [32] und [33] vorgestellt. Hier wurden die Thunfische durch die Taucher jedoch gegen die Wasseroberfläche gefilmt, wodurch ein gut zu erkennender Schattenumriss der Fische entstand, was in Abbildung 5.1 B gut zu sehen ist. Außerdem war die Dichte der Fische derart niedrig, dass Bereiche mit der Wasseroberfläche zwischen den Fischen zu erkennen sind. Bei diesem Verfahren wurden fast sämtliche Thunfische erkannt, jedoch wies die Vermessung und daraus kalkulierte Biomasse eine Ungenauigkeit von circa 50 % auf.

Drei Verfahren, bei denen Fischarten erkannt, unterschieden und vermessen wurden, werden in [132], [134] und [147] vorgestellt. Sie haben im Vergleich mit den zuvor erwähnten Verfahren den Charakter von industriellen Bildverarbeitungsverfahren, da hier die Fische auf Fließbändern laufen, wo die Umgebungsparameter für die Erkennung der Fische optimiert wurden. Die Fische liegen einzeln auf dem farblich konträrem Fließband, werden eventuell von unten beleuchtet, um einen Schattenumriss zu erhalten und die Abstände zwischen Fischen und Kameras und weiterer Sensoren sind bekannt. Eine solche Aufnahme zeigt Abbildung 5.1 C. All diese Faktoren erleichtern die Erkennung einer vollständigen Kontur des zu erkennenden Fisches und sind bei dieser Arbeit nicht gegeben.

Abbildung 5.1 D zeigt abschließend ein Bild von Steinbutten von der GMA in Büsum. Die Besatzdichte ist derart hoch, dass sich nur Steinbutte im Bild befinden und sich Steinbutte im Vordergrund kaum von denen im Hintergrund abheben. Eine Beleuchtung von hinten, um Schattenrisse zu erzeugen,

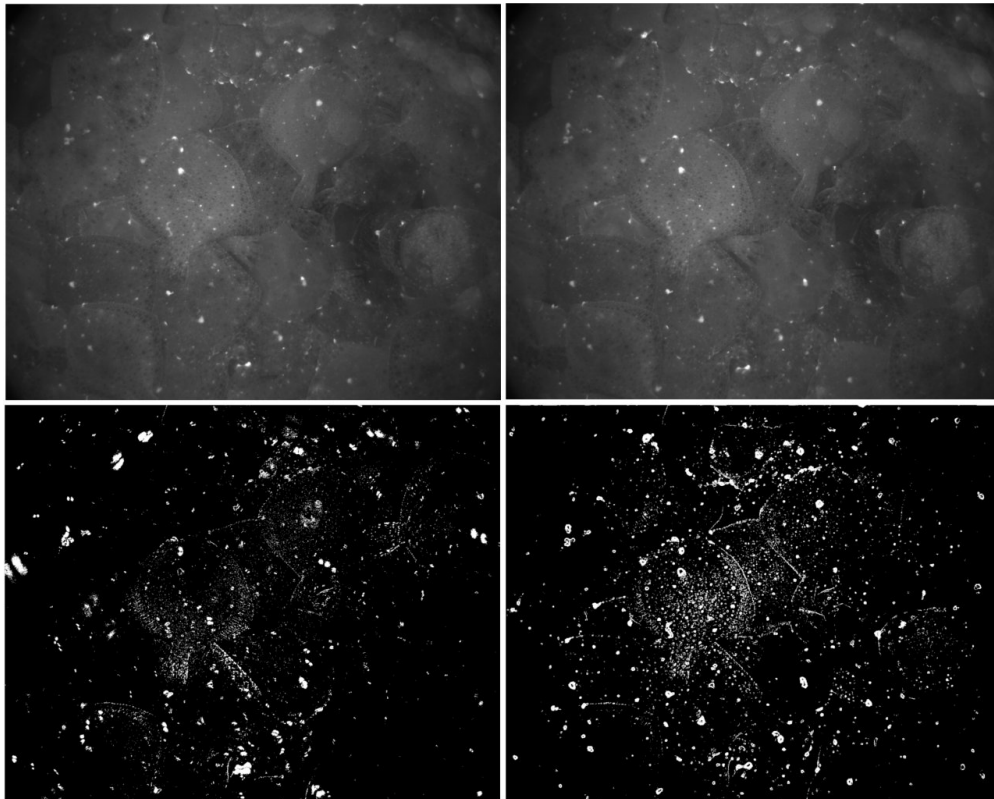


Abbildung 5.2: Beispiel für den Einsatz von Differenzbildern. Oben: Zwei aufeinander folgende Aufnahmen aus Büsum. Unten links: Differenzbild. Unten rechts: Sobelfilter zur Kantendetektion auf erstem Bild.

ist ebenfalls nicht möglich und die in [140] und [86] problematische Nähe zu weiteren Fischen ist hier prinzipiell vorhanden.

Zwei Ansätze zur Erkennung von Steinbutten unter realen Arbeitsbedingungen müssen noch genauer betrachtet werden. Der Einsatz von Shapes hat zum Ziel, Modelle von vollständigen Steinbutten an die im Bild gefundenen Konturen zu matchen, um so einen Steinbutt zu lokalisieren. Das Problem hierbei ist die Tatsache, dass in dem vorliegenden Bildmaterial fast sämtliche Konturen zu einem von vielen Steinbutten gehören, was das Matching des Modells an Konturen, die nur zu einem einzelnen Steinbutt gehören, extrem erschwert. [140] und [86] verwenden deformierbare Modelle und berichten von Problemen beim Matching, wenn Fische im Bild in unmittelbarer Nähe zueinander liegen.

Der zweite Ansatz ist die Verwendung von Differenzbildern. Eine Möglichkeit ist die Verwendung von aufeinander folgender Aufnahmen. Ein Beispiel hierfür zeigt Abbildung 5.2. Aus den beiden oberen Aufnahmen entstand das Differenzbild unten links. Unten rechts ist zum Vergleich ein einfacher Sobel-

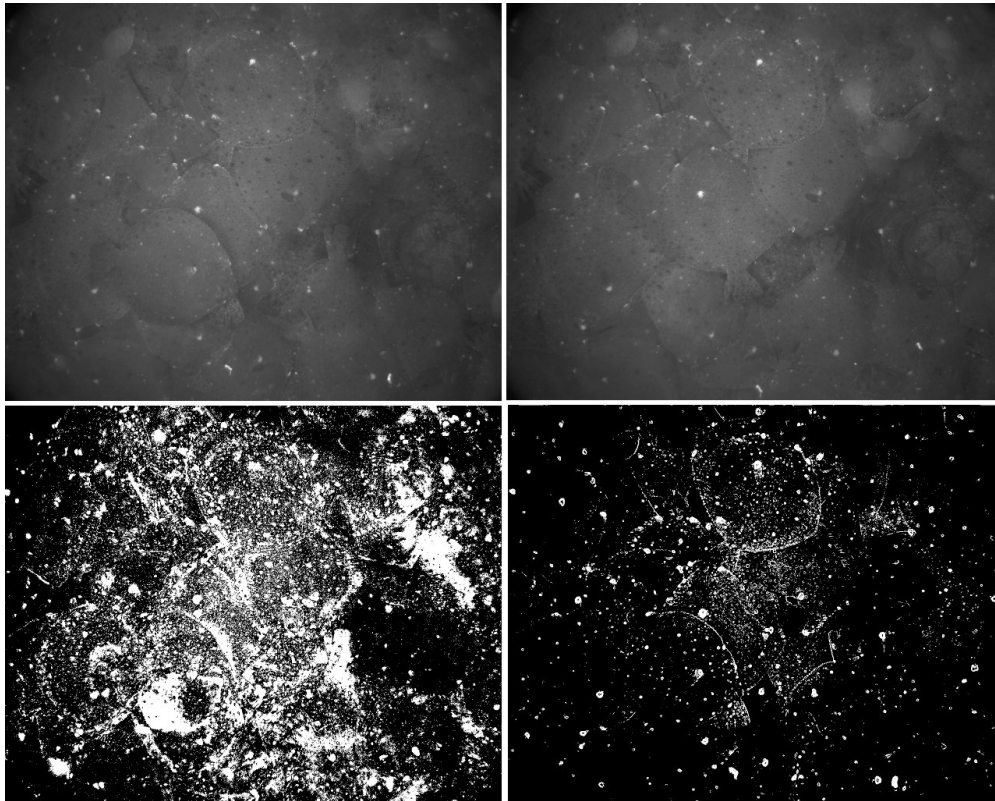


Abbildung 5.3: Beispiel für den Einsatz von Differenzbildern. Oben: Zwei Aufnahmen aus Büsum mit circa 2 Sekunden Abstand, Steinbutt wechselt Position vollständig. Unten links: Differenzbild. Unten rechts: Sobelfilter zur Kantendetektion auf erstem Bild.

filter auf ein Bild der oberen Reihe angewendet worden. Es werden in beiden Varianten sehr ähnliche Strukturen hervorgehoben. Das liegt darin begründet, dass bei einem Differenzbild, eine Bewegung nur dann angezeigt werden kann, wenn eine hervorstechende Struktur, wie zum Beispiel eine Kante, die Position ändert. Solche Strukturen werden auch von Kantendetektoren erkannt, die sogar den Vorteil besitzen, Strukturen zu finden, die sich nicht bewegen.

Eine andere Möglichkeit beim Einsatz von Differenzbildern ist es, den Zeitabstand zwischen den zwei Bildern derart auszudehnen, dass ein Steinbutt seine Position komplett ändert und aufgrund seines Fehlens dieser Bereich durch das Differenzbild entsprechend hervorgehoben wird. Ein Beispiel hierfür zeigt Abbildung 5.3. Aus den beiden oberen Aufnahmen, bei denen ein Steinbutt im unteren Bildbereich seine Position vollständig ändert, entstand das Differenzbild unten links. Unten rechts ist zum Vergleich ein einfacher Sobelfilter auf das linke Bild der oberen Reihe angewendet worden. Das Differenzbild zeigt in fast allen Bereichen des Bildes Veränderungen an, gerade auch in Be-

reichen, in denen sich die Steinbutte kaum vom Ort bewegt haben, aber kleine Positionsänderungen vornahmen. Der Bereich, in dem der Steinbutt seine Position vollständig änderte, wird nicht im besonderen Maße hervorgehoben. Die Anwendung des Sobelfilters zeigt hier deutlichere Konturen von Steinbutten.

Die Idee, das Differenzbild dazu zu nutzen, um den Suchbereich auf schwimmende Steinbutte einzugrenzen, in der Annahme, dass nur solche Bewegung erzeugen, musste verworfen werden, da auch von verdeckten Steinbutten miteinander starke Bewegungen erzeugt werden, und somit kein Vorteil durch das Differenzbild entsteht.

Aufgrund der zuvor genannten Faktoren, wie eine sehr hohe Besatzdichte, die eine Abgrenzung von Steinbutten zu einem sich differenzierenden Hintergrund verhindert sowie die Camouflage des Steinbutts, die eine Konturfindung zusätzlich erschwert, muss ausgehend von den in dieser Arbeit gewonnenen Kenntnissen davon ausgegangen werden, dass die Erkennung von Steinbutten unter den bei dieser Arbeit gegebenen Arbeitsbedingungen zum jetzigen Zeitpunkt nicht realisierbar ist. Ideen zur Lösung dieser Problematik werden im folgenden Kapitel aufgezeigt.

KAPITEL 6

Ausblick

Inhalt

6.1	Hardwaresystem	121
6.2	Algorithmus zur Steinbuttvermessung	122

6.1 Hardwaresystem

Die wichtigsten Verbesserungsmöglichkeiten am Hardwaresystem bietet das Beleuchtungssystem, um so eine Balance zwischen den Ansprüchen der Steinbutte und der bestmöglichen Bildqualität für die Objekterkennung von Steinbutten zu erreichen.

So wäre für den Dauereinsatz eine Optimierung der Beleuchtung des Systems im Sinne der Steinbutte sinnvoll. Für diese Arbeit wurde untersucht, ob eine bestimmte Art der Lichtfarbe Vorteile bei der Fischerkennung aufweist, siehe Kapitel 3.3.1. Da Steinbutte von Natur aus lichtscheu sind, besteht normalerweise ein Bedarf an möglichst geringen Lichtmengen durch die Beleuchtung. Für einen dauerhaften Einsatz beim Monitoring von Steinbutten wäre zu ermitteln, welche Lichtmengen und zu welchen Zeitpunkten diese eingesetzt werden können, ohne negative Einflüsse auf das Wachstum der Steinbutte auszuüben. Ist diese als für Steinbutte optimal zu bezeichnende Beleuchtung gefunden, kann die restliche Hardware, wie die Belichtungszeit der Kameras oder die Blende der Objektive, darauf abgestimmt werden, um so das bestmögliche Bildmaterial zu erzeugen.

Um die Lichtmenge für die Steinbutte niedrig zu halten und für die Bildaufnahme zu maximieren, wäre eine Überarbeitung des in Kapitel 5.1 diskutierten getriggerten Lichts denkbar. Dafür müssten zuerst Untersuchungen erfolgen, bei welcher Frequenz der Lichtblitze die Steinbutte das getriggerte Licht als konstant empfinden. Daraufhin wäre eine Anpassung des Hardwaresystems nötig. Zum Beispiel der Einsatz von Kameras deren Framerate zur benötigten Frequenz passt, oder die Modifizierung der Steuerelektronik, so dass Lichtblitze außerhalb einer Bildaufnahme durchgeführt werden können, um so die benötigte Frequenz zu erreichen.

Die Verwendung von Licht aus dem nicht sichtbaren Spektrum wie Infrarot oder Ultraviolett kann aufgrund der hohen Absorptionsraten von Wasser nicht effektiv umgesetzt werden.

In Anbetracht der folgenden Lösungsansätze für das Verfahren zur Erkennung von Steinbutten sollte überprüft werden, inwiefern die in Kapitel 5.1 diskutierte Kombination von grünem Untergrund und rotem beziehungsweise grünem Licht die Erkennung einer vollständigen Kontur eines einzeln liegenden Steinbutts begünstigt. Bei positiven Resultaten wäre eine entsprechende Anpassung des Beckens sowie des Beleuchtungssystems vorzunehmen.

Für einen dauerhaften Einsatz erscheint der Bedarf einer automatisierten Möglichkeit, die Sichtscheibe frei von Verschmutzungen und Luftblasen zu halten, unumgänglich. Hier gilt es eine technische Lösung zu finden, die regelmäßig beziehungsweise beständig die Sichtscheibe reinigt und eine klare Sicht durch die Scheibe ermöglicht. Gedankenansätze hierzu sind bereits in Kapitel 3.2.1 zu finden, zum Beispiel ein Abzieher oder ein Schwamm. Eine andere Idee basiert darauf mit Hilfe eines kontinuierlichen Wasserstrahls die Sichtscheibe frei von Schmutz und Luftbläschen zu halten.

6.2 Algorithmus zur Steinbuttvermessung

Die Ergebnisse dieser Arbeit zeigen, dass unter den vorliegenden realen Arbeitsbedingungen eine Erkennung von Steinbutten nicht realisierbar erscheint.

Um die Erkennung von Steinbutten mittels eines non-invasives Verfahrens basierend auf einem Stereokamerasystem und damit deren Vermessung sowie ein Wachstums-Monitoring zu ermöglichen, müssen die Arbeitsbedingungen angepasst werden. Das Hauptproblem liegt in der hohen Besatzdichte auf engem Raum und der dadurch praktisch nicht vorhandenen Unterscheidbarkeit von einem Steinbutt im Vordergrund, der vermessen werden soll, und dem Hintergrund, der ebenfalls aus Steinbutten besteht.

Die naheliegendste Änderung der Arbeitsbedingungen ist die Verwendung anderer Anlagensysteme für die Haltung der Steinbutte. Anstatt Becken mit relativ kleiner Grundfläche können Steinbutte auch in Becken mit großer Grundfläche von beispielsweise 8×8 m gehalten werden. Hier kommt es selbst bei einer hohen Besatzdichte zu Bereichen, wo Steinbutte vereinzelt auf dem Untergrund liegen oder durch das Sichtfeld der Kameras schwimmen. Da bei einem derart großen Becken nur ein relativ kleiner Bereich von einem Stereokamerasystem überwacht werden kann und der Tatsache, dass die Bereiche mit geringer Steinbuttverteilung nicht vorherbestimmt werden können, wäre hier die Installation mehrerer Systeme eine Option.

Für eine Reduzierung der Besatzdichte in dem kleinen Becken der GMA

in BÜsum, sodass vereinzelt liegende Steinbutte auftreten, soll hier eine beispielhafte Rechnung herangezogen werden. Ausgehend von kreisrunden Steinbutten und der Gewichtsrechnung aus Kapitel 2.1, hat ein Steinbutt von 30 cm Länge ein Gewicht von circa 380 g und eine Fläche von circa 700 cm². Das Becken von 180 cm Durchmesser kommt auf eine Grundfläche von circa 2,5 m². Das ergibt 36 Steinbutte, um den Boden vollständig zu bedecken bei einer Biomasse von circa 14 kg. Daraus folgt eine Besatzdichte von nur 5,5 kg/m² im Vergleich zu den in der Mast üblichen 100 kg/m². Selbst wenn eine Verdopplung der Steinbutte ausreicht, um regelmäßig vereinzelt liegende Steinbutte zu erhalten, müsste die Besatzdichte auf 10 % der üblichen Menge reduziert werden.

Eine weitere Möglichkeit ist die Anpassung der Becken derart, dass Bereiche entstehen, wo sich Steinbutte nur vereinzelt aufhalten können, die dann vom Monitoring-System explizit überwacht werden. Podeste, auf denen nur einzelne Steinbutte Platz finden, wären eine Möglichkeit. Die Verwendung von einer überwachten Schleuse, durch die die Steinbutte schwimmen müssen, beziehungsweise sogar dazu gezwungen werden, würde jedoch dem Anspruch eines non-invasiven Systems widersprechen.

Ist so eine Trennung von Objekt und Hintergrund realisiert, kann die Hardware an die neuen Bedingungen entsprechend angepasst werden, um die Bildfassung zu optimieren. Unter diesen Bedingungen kann eine vollständige Kontur eines Steinbutts erzeugt und ein entsprechendes Verfahren zur konturbasierten Objekterkennung angewendet werden. Basierend auf diesen Daten erfolgt dann eine Vermessung des Steinbutts anhand der Stereobilder.

Wie in Kapitel 5.2 angesprochen, ist die Entwicklung eines Verfahrens zur automatisierten Anpassung der Parameter in der Vorverarbeitung eine sinnvolle Erweiterung des Funktionsumfangs und könnte die Bildverarbeitung unabhängiger von Umweltbedingungen gestalten.

Die Anwendung der in dieser Arbeit gewonnenen Erkenntnisse sowie die Einbeziehung der genannten Überlegungen sollten die Realisierung eines non-invasiven Wachstums-Monitorings von Steinbutten bei hoher Populationsdichte in der Aquakultur mittels eines Stereokamerasystems ermöglichen.

Zusammenfassung

Inhalt

7.1 Zusammenfassung	125
7.2 Summary	126

7.1 Zusammenfassung

Ziel dieser Arbeit ist die Entwicklung eines Prototypen zur non-invasiven Vermessung von Steinbutten in Aquakulturanlagen mittels eines Stereokamerasystems unter realen Arbeitsbedingungen. Die zwei Themenschwerpunkte sind die Entwicklung eines Konzeptes für ein Hardwaresystem zur Aufnahme von Bildmaterial unter den gegebenen Umständen sowie die Entwicklung eines Algorithmus zur automatischen Erkennung von Steinbutten im Bildmaterial und deren Vermessung basierend auf Stereometrie. Die bei der Entwicklung von Hardwaresystem und Algorithmus durchgeführten Arbeitsschritte und Untersuchungen werden in dieser Arbeit vorgestellt und deren Ergebnisse präsentiert.

Für die Entwicklung des Hardwaresystems wurde zuerst eine Anforderungsanalyse erstellt, basierend auf den Bedingungen, die ein derartiges Arbeitsumfeld an ein solches System stellt.

Im Rahmen dieser Arbeit wurde anschließend ermittelt, ob die Verwendung von farbigem Licht und einer bestimmten Farbe des Beckenuntergrundes die softwareseitige Objekterkennung unterstützen kann. Dies wurde anhand der Parameter Belichtungszeit, relativer Kontrast zwischen Untergrund und einem Steinbutt, empirischer Variationskoeffizient der Grauwerte innerhalb eines Steinbuttes, sowie Gradientenstärke im Übergangsbereich zwischen 2 Steinbutten unter Verwendung von 5 unterschiedlichen Lichtfarben, 4 unterschiedlichen Beckenfarben sowie 2 unterschiedlich heller Steinbuttmusterungen untersucht. Bei der Hervorhebung von Strukturen wurden keine signifikanten Unterschiede zwischen den einzelnen Farbkombinationen festgestellt. Weißes Licht zeigte im Vergleich zu den anderen Lichtfarben eindeutig kürzere Belichtungszeiten, und wurde aus diesem Grund für die weitere Verwendung in dieser Arbeit gewählt.

Um eine möglichst homogene Ausleuchtung des beobachteten Bereichs zu erreichen, wurden mehrere Grundeinstellungen für die Abstrahlwinkel des Beleuchtungssystems überprüft und ausgewertet. Die Ausrichtung der Leuchtkörper um 20° nach außen gedreht erwies sich als die beste Grundeinstellung der überprüften Varianten.

Die Prüfung der Genauigkeit der Kalibrierung des Stereokamerasystems wurde anhand von Testmessungen durchgeführt. Dabei zeigte sich eine Messungengenauigkeit von circa 1,7 % im mittleren Bildbereich und von circa 2 % im Randbereich.

Der entwickelte und vorgestellte Algorithmus stellt sich der Aufgabe, einzelne Steinbutte in Bildmaterial zu erkennen, dessen Vorder- als auch Hintergrund vollständig aus Steinbutten besteht und die Aufnahme nicht unter optimalen Bedingungen erfolgen kann. Das Verfahren basiert auf der Erkennung von repräsentativen Konturfragmenten und deren Zuordnung anhand geometrischer Zusammenhänge, die zuvor automatisiert erlernt wurden.

Der Schwellwert zur Unterscheidung zwischen repräsentativen und nicht repräsentativer Konturen wird bestimmt, basierend auf den Ergebnissen von mehreren Trainings mit unterschiedlichen Sets an Trainingsdaten. Dabei konnte eindeutig ein Wertebereich identifiziert werden, in dem ein Übergang von repräsentativen zu nicht repräsentativen Konturen stattfindet.

Für mehrere unterschiedliche Sets repräsentativer Konturen wurden anhand zweier Sets von Validierungsbildern das Konturen-Matching durchgeführt. Die Ergebnisse zeigten bei allen Varianten sehr geringe Raten an korrekt gematchten und skalierten repräsentativen Konturen, weswegen eine Fortführung des Verfahrens nicht erfolgen konnte.

Das Hardwaresystem erwies sich als robust in der Anwendung und von der Konstruktion als flexibel genug, um in verschiedenen Bereichen der Aquakultur als Monitoring-System eingesetzt zu werden.

Der entwickelte Algorithmus zur Erkennung von Steinbutten war nicht in der Lage, Steinbutte zuverlässig unter realen Arbeitsbedingungen zu erkennen. In der Diskussion und dem Ausblick werden Gründe hierfür genannt und Möglichkeiten zur Behebung der vorliegenden Probleme diskutiert.

7.2 Summary

Goal of this work is the development of a prototype for the non-invasive measuring of turbots in aquacultural facilities via a stereo camera system under realistic working conditions. The two key subjects are the development of a concept for a hardware system for the recording of image data under the given conditions as well as the development of an algorithm for the automatic re-

cognition of turbots in image data and their measuring based on stereometry. The working steps and researches done in the development of the hardware system and the algorithm are shown and the results are presented in this work.

For the development of the hardware system, first a requirement analysis was done, based on the conditions of the actual working area.

In this work it was tested, if the usage of colored light and colored material for the basin can support the object recognition algorithm. This was tested based on the parameters exposure time, relative contrast, empirical coefficient of variation and the gradient intensity using 5 different colors of light, 4 different colors of the basin and 2 differently bright patterned turbots. In emphasising the structures of the turbots no significant differences between the combinations of color were found. White light showed the shortest exposure times in comparison with the other colors of light and was therefore chosen for further usage.

To reach the most homogenous illumination of the monitored area possible, multiple basic settings for the radiation angles of the lighting system were evaluated. The orientation of the lamps 20° degrees to the outer side proved to be the best basic setting of the examined variations.

The testing of the precision of the calibration of the stereo camera system was done using exemplary measurements. A measuring inaccuracy of approximately 1.7 % in the image center and of approximately 2 % at the image border was reached.

The developed and presented algorithm meets the challenge to recognize single turbots in image data, where foreground and background consist completely out of turbots and the shot can not be done under optimized conditions. The algorithm bases on the recognition of representative contour fragments and their assignment based on geometric correlations, that are learned automatically.

The threshold to differentiate between representative and non representative contours is determined based on the results of multiple trainings of the algorithm using different sets of training data. A distinct range of values could be identified, where a transition from representative to non representative contours occurs.

For multiple sets of representative contours the contour matching was executed using 2 sets of validation images. The results showed for all variations very low rates of correctly matched and scaled representative contours, wherefore a continuation of the algorithm could not be done.

The hardware system proved itself as robust in its usage and is constructed flexible enough to be used in different areas of aquaculture as a monitoring system.

The developed algorithm to recognise turbots was not able to detect tur-

bots under realistic working conditions reliably. In the discussion and the outlook reasons are mentioned and possible solutions are discussed.

ANHANG A
Anhang

A.1 Ergebnistabellen Versuche zur Farbwahl

Tabelle A.1: Mittlerer Grauwert eines Steinbutts bei einer Belichtungszeit von 10 ms. Berechnung des Mittelwerts aus drei Aufnahmen pro Konfiguration von Steinbutt, Boden- und Lichtfarbe.

Fisch	Boden	Rot	Weiß	Grün	Blau	Gelb
Dunkel	Weiß	35,60	53,30	37,40	42,40	6,90
	Weiß	40,00	53,30	37,40	42,40	8,40
	Weiß	39,90	53,30	37,40	42,40	8,40
	Mittelw.	38,50	53,30	37,40	42,40	7,90
	Grau	24,00	37,50	26,00	30,10	3,00
	Grau	23,90	37,40	26,00	30,10	3,10
	Grau	23,90	37,40	26,00	30,10	3,00
	Mittelw.	23,93	37,43	26,00	30,10	3,03
	Schwarz	21,10	32,40	31,20	25,60	2,50
	Schwarz	21,10	32,40	31,10	25,60	2,40
	Schwarz	21,10	32,40	31,20	25,70	2,50
	Mittelw.	21,10	32,40	31,17	25,63	2,47
	Grün	23,30	39,60	30,60	30,00	3,10
	Grün	23,30	39,60	30,60	30,10	3,10
	Grün	23,30	39,60	30,60	30,10	3,10
	Mittelw.	23,30	39,60	30,60	30,07	3,10
Hell	Weiß	84,30	123,50	83,30	82,80	24,00
	Weiß	84,30	123,60	83,30	82,80	24,00
	Weiß	84,30	123,70	83,40	82,90	24,00
	Mittelw.	84,30	123,60	83,33	82,83	24,00
	Grau	65,70	95,20	63,70	63,40	43,10
	Grau	65,70	95,20	63,70	63,40	43,20
	Grau	65,70	95,20	63,70	63,60	43,20
	Mittelw.	65,70	95,20	63,70	63,47	43,17
	Schwarz	59,30	86,60	57,20	57,70	15,70
	Schwarz	59,30	86,60	57,20	57,70	15,70
	Schwarz	59,30	86,60	57,10	57,70	15,70
	Mittelw.	59,30	86,60	57,17	57,70	15,70
	Grün	60,80	90,10	61,70	59,10	47,40
	Grün	60,80	90,00	61,70	59,20	47,40
	Grün	60,80	90,10	61,60	59,10	47,50
	Mittelw.	60,80	90,07	61,67	59,13	47,43

Tabelle A.2: Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei einer Belichtungszeit von 10 ms.

	Fisch Dunkel			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,546	0,254	0,066	0,010
Licht Grün	0,531	0,251	0,089	0,302
Licht Blau	0,519	0,247	0,067	0,177
Licht Rot	0,540	0,237	0,150	0,033
Licht Gelb	0,675	0,469	0,190	0,205
	Fisch Hell			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,132	0,224	0,397	0,218
Licht Grün	0,135	0,213	0,390	0,107
Licht Blau	0,160	0,157	0,351	0,194
Licht Rot	0,125	0,289	0,443	0,414
Licht Gelb	0,152	0,258	0,679	0,364

Tabelle A.3: Standardabweichung vom relativen Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei einer Belichtungszeit von 10 ms.

	Fisch Dunkel			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,115	0,052	0,019	0,002
Licht Grün	0,089	0,040	0,022	0,057
Licht Blau	0,089	0,039	0,017	0,038
Licht Rot	0,111	0,049	0,053	0,010
Licht Gelb	0,278	0,316	0,222	0,178
	Fisch Hell			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,031	0,074	0,144	0,077
Licht Grün	0,029	0,065	0,136	0,034
Licht Blau	0,032	0,050	0,142	0,069
Licht Rot	0,031	0,098	0,161	0,169
Licht Gelb	0,044	0,108	0,566	0,189

Tabelle A.4: Relativer Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.

	Fisch Dunkel			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,547	0,252	0,065	0,227
Licht Grün	0,539	0,245	0,080	0,306
Licht Blau	0,520	0,245	0,061	0,173
Licht Rot	0,540	0,228	0,118	0,031
Licht Gelb	0,542	0,254	0,099	0,105
	Fisch Hell			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,130	0,220	0,386	0,214
Licht Grün	0,153	0,174	0,357	0,076
Licht Blau	0,158	0,152	0,344	0,189
Licht Rot	0,142	0,258	0,414	0,380
Licht Gelb	0,141	0,247	0,417	0,338

Tabelle A.5: Standardabweichung vom relativen Kontrast zwischen farbiger Bodenplatte und einem Steinbutt bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.

	Fisch Dunkel			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,115	0,050	0,017	0,052
Licht Grün	0,087	0,035	0,018	0,055
Licht Blau	0,087	0,037	0,014	0,036
Licht Rot	0,106	0,043	0,029	0,008
Licht Gelb	0,104	0,048	0,026	0,028
	Fisch Hell			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,030	0,047	0,133	0,074
Licht Grün	0,032	0,051	0,113	0,023
Licht Blau	0,032	0,046	0,132	0,065
Licht Rot	0,035	0,083	0,127	0,132
Licht Gelb	0,038	0,087	0,144	0,132

Tabelle A.6: Empirischer Variationskoeffizient innerhalb eines Steinbutts bei einer Belichtungszeit von 10 ms.

	Fisch Dunkel			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,101	0,093	0,130	0,093
Licht Grün	0,091	0,081	0,112	0,078
Licht Blau	0,085	0,076	0,105	0,087
Licht Rot	0,115	0,108	0,152	0,116
Licht Gelb	0,319	0,433	0,520	0,419
	Fisch Hell			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,120	0,190	0,171	0,178
Licht Grün	0,102	0,168	0,157	0,138
Licht Blau	0,098	0,169	0,158	0,157
Licht Rot	0,114	0,179	0,162	0,183
Licht Gelb	0,158	0,223	0,299	0,222

Tabelle A.7: Empirischer Variationskoeffizient innerhalb eines Steinbutts bei automatisch festgelegter Belichtungszeit für einen mittleren Grauwert von 127.

	Fisch Dunkel			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,099	0,091	0,124	0,091
Licht Grün	0,085	0,068	0,101	0,072
Licht Blau	0,080	0,072	0,092	0,080
Licht Rot	0,108	0,093	0,131	0,099
Licht Gelb	0,104	0,096	0,138	0,103
	Fisch Hell			
	Boden Weiß	Boden Grau	Boden Schwarz	Boden Grün
Licht Weiß	0,119	0,074	0,154	0,176
Licht Grün	0,099	0,162	0,139	0,134
Licht Blau	0,096	0,163	0,150	0,149
Licht Rot	0,112	0,167	0,127	0,140
Licht Gelb	0,142	0,204	0,158	0,189

Literaturverzeichnis

- [1] N. Abdullah, M. Shafry, M. Rahim, and I. M. Amin. Measuring fish length from digital images (filedi). In *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, ICIS '09, pages 38–43, New York, NY, USA, 2009. ACM. 32
- [2] Hedin A.E., Fleming E.L., Manson A.H., Schmidlin F.J., Avery S.K., Clark R.R., Franke S.J., Fraser G.J., Tsuda T., Vial F., Vincent R.A., Abovsky M., Lanir Y., Nevo E., Dissanayake C., Sigman M., Kasari C., van Rooij J.M., and Videler J.J. A simple field method for stereo-photographic length measurement of free-swimming fish: merits and constraints. *Journal of Experimental Marine Biology and Ecology*, 195(2):237–249, 1996. 31
- [3] Continental AG. Two eyes are better than one the stereo camera. http://www.conti-online.com/generator/www/com/en/continental/pressportal/themes/press_releases/3_automotive_group/chassis_safety/press_releases/pr_20110504_stereo_camera_en.html. 05.04.2011. 36
- [4] Daimler AG. 6d-vision - meilenstein der fahrzeugsicherheit. <http://www.daimler.com/technologie-und-innovation/einblicke-in-forschung-und-entwicklung/6d-vision>. 19.04.2012. 36
- [5] M. K. Alsmadi, K. B. Omar, S. A. Noah, and I. A. Almarashdeh. Fish recognition based on robust features extraction from color texture measurements using back-propagation classifier. *Journal of Theoretical and Applied Information Technology*, 18(1):11–18, 2010. 32
- [6] V. Aurich, E. Mühlhaus, and S. Grundmann. Kantenerhaltende glättung von volumendaten bei sehr geringem signal-rausch-verhältnis. In *Bildverarbeitung für die Medizin'98*, pages –1–1, 1998. 87
- [7] V. Aurich and J. Weule. *Non-Linear Gaussian Filters Performing Edge Preserving Diffusion*, volume 17, pages 538–545. Springer-Verlag, 1995. 87
- [8] C. Beall, B. J. Lawrence, V. Ila, and F. Dellaert. 3d reconstruction of underwater structures. *Atlantic*, 0448111:4418–4423, 2010. 36

- [9] J. Bedo, M. Chetty, A. Ngom, and S. Ahmad. *Pattern Recognition in Bioinformatics*, volume 4774. Springer Berlin Heidelberg, 2007. 25
- [10] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:509–522, 2001. 37
- [11] T. Berg. Kinect and 3d. http://www.tamaraberg.com/teaching/Spring_11/CSE364/slides/kinect.pdf, 2011. 08.02.1013. 36
- [12] S. Berretti, A. D. Bimbo, and P. Pala. Retrieval by shape similarity with perceptual distance and effective indexing. *IEEE Transactions on Multimedia*, 2(4):225–239, 2000. 29
- [13] A. D. Bimbo and P. Pala. Visual image retrieval by elastic matching of user sketches. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):121–132, 1997. 29
- [14] C. M. Bishop. *Neural Networks for Pattern Recognition*, volume 92. Oxford University Press, 1995. 28
- [15] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer Science+Business Media, LLC, 2006. 26, 27
- [16] G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):849–865, 1988. xii, 93, 94
- [17] J. Y. Bouguet. 3d transformations & camera calibration. www.multires.caltech.edu/teaching/courses/3DP/notes/calib_notes.ps. 01.05.2012. 64
- [18] J. Y. Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/. 01.05.2012. 62
- [19] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. V. Rationale. Experimentation of structured light and stereo vision for underwater 3d reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4):508–518, 2011. 36
- [20] M. Bryant, D. Wettergreen, S. Abdallah, and A. Zelinsky. Robust camera calibration for an autonomous underwater vehicle. In *In Proceedings of the 2000 Australian Conference of Robotics and Automation*, 2000. 35

- [21] D. J. Buehrer, C. S. Lu, and C. C. Chang. A shape recognition scheme based on real-valued boundary points. *Tamkang Journal Of Science And Engineering*, 3(1):5574, 2000. 37
- [22] S. X. Cadrin and K. D. Friedland. The utility of image processing techniques for morphometric analysis and stock identification. *Fisheries Research*, 43(13):129 – 139, 1999. 32
- [23] J. Cameron and V. Pace. Cameron | pace group (cpg). <http://www.cameronpace.com/v2/>. 18.04.2012. 36
- [24] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, 1986. 23, 89
- [25] M. E. Celebi and Y. A. Aslandogan. A comparative study of three moment-based shape descriptors. *International Conference on Information Technology Coding and Computing ITCC05 Volume II*, 1:788–793 Vol. 1, 2005. 38
- [26] Cepheiden. Brechungsgesetz-winkel. <http://de.wikipedia.org/wiki/Datei:Brechungsgesetz-Winkel.svg>, 2010. 08.02.2013. ix, 43
- [27] R. Chellappa and R. Bagdazian. Fourier coding of image boundaries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(1):102–105, 1984. 29
- [28] Z. Chi, H. Yan, and T. Pham. *Fuzzy algorithms: with applications to image processing and pattern recognition*. Advances in fuzzy systems. World Scientific, 1996. 27
- [29] JVC Americas Corp. Gs-td1b full hd 3d everio camcorder. <http://resources.jvc.com/Resources/00/01/48/99.PDF>. 20.04.2012. 36
- [30] FUJIFILM Corporation. Fujifilm finepix real 3d w3. http://www.fujifilm.com/products/3d/camera/finepix_real3dw3/. 20.04.2012. 36
- [31] Microsoft Corporation. Kinect for xbox360. <http://www.xbox.com/kinect>. 22.12.2012. 36
- [32] C. Costa, A. Loy, S. Cataudella, D. Davis, and M. Scardi. Extracting fish size using dual underwater cameras. *Aquacultural Engineering*, 35(3):218 – 227, 2006. xiii, 33, 116, 117

- [33] C. Costa, M. Scardi, V. Vitalini, and S. Cataudella. A dual camera system for counting and sizing northern bluefin tuna (*thunnus thynnus*; linnaeus, 1758) stock, during transfer to aquaculture cages, with a semi automatic artificial neural network tool. *Aquaculture*, 291(34):161 – 167, 2009. 33, 117
- [34] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems*, volume 43. Springer, 1999. 27
- [35] J. Cui, Z. Li, R. Lv, X. Xu, and J. Gao. The application of support vector machine in pattern recognition. *2007 IEEE International Conference on Control and Automation*, 00(3):3135–3138, 2007. 27
- [36] E. R. Davies. *Machine Vision: Theory, Algorithms, Practicalities*. Morgan Kaufmann Publishers Inc., 2004. 29, 30
- [37] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley & Sons, second edition, 2001. 25, 27
- [38] R. L. Dunbrack. In situ measurement of fish body length using perspective-based remote stereo-video. *Fisheries Research*, 82(13):327 – 331, 2006. 32
- [39] FAO. The state of world fisheries and aquaculture 2010, 2010. 8, 9
- [40] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. The MIT Press, 1993. 15
- [41] Food and Agriculture Organization of the United Nations. Fisheries and aquaculture department. <http://www.fao.org>. 08.02.2013. ix, 6, 7, 8, 9, 10, 11, 12
- [42] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002. 12, 16, 17, 19, 20, 25, 35
- [43] H. Freeman. On the encoding of arbitrary geometric configurations. *Electronic Computers IRE Transactions on*, 10(2):260–268, 1961. 29
- [44] R. Froese and D. Pauly. Fishbase. World Wide Web electronic publication <http://www.fishbase.org>. 02.05.2012. 7
- [45] J. G. Fryer and C. S. Fraser. On the calibration of underwater cameras. *Photogrammetric Record*, (12(67)):73–85, 1986. 34

- [46] K. S. Fu. *Syntactic Methods in Pattern Recognition*, volume 112. Academic Press, 1974. 28, 29
- [47] N. Gherabi. A new algorithm for shape matching and pattern recognition using dynamic programming. *Context*, 2010. 37
- [48] R. N. Gibson. *Flatfishes. Biology and Exploitation*. Blackwell Publishing, Oxford, 2005. 5, 7, 9, 12
- [49] J. Glass. A brief introduction to automatic speech recognition. *Artificial Intelligence*, pages 1–22, 2007. 25
- [50] IDS Imaging Development Systems GmbH. Gige ui-5240se datenblatt. <http://www.ids-imaging.de/>. 06.04.2012. 48
- [51] LG Electronics Deutschland GmbH. Lg p920 optimus 3d. www.lg.com/de/. 27.01.2013. 36
- [52] Pentax Ricoh Imaging Deutschland GmbH. C60402kp - h416 (kp) datasheet. 06.04.2012. x, 62
- [53] J. W. Gorman. Recognition of incomplete polygonal objects. In *Twenty-First Southeastern Symposium on System Theory*, pages 518–522, 1989. 38
- [54] J. W. Gorman, O. R. Mitchell, and F. P. Kuhl. Partial shape recognition using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(2):257–266, 1988. 38
- [55] R. Grekova. Methods of fuzzy pattern recognition. In *Scientific proceedings of Riga Technical University*, 2002. 27
- [56] W. I. Grosky and R. Mehrotra. Index-based object recognition in pictorial data management. *Computer Vision Graphics and Image Processing*, 52(3):416–436, 1990. 29
- [57] G. M. Hale and M. Q. Querry. Optical constants of water in the 200-nm to 200- μm wavelength region. *Applied Optics*, 12:555–563, 1973. 110
- [58] C. Hane, C. Zach, J. Lim, A. Ranganathan, and M. Pollefeys. Stereo depth map fusion for robot navigation. *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1618–1625, 2011. 36
- [59] R. I. Hartley and P. Sturm. Triangulation. *CVIU - Computer Vision and Image Understanding*, 68:957–966, 1994. 100

- [60] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004. 12, 16, 17, 21, 35
- [61] E. Harvey, M. Cappelletti, M. Shortis, S. Robson, J. Buchanan, and P. Speare. The accuracy and precision of underwater measurements of length and maximum body depth of southern bluefin tuna (*thunnus maccoyii*) with a stereo-video camera system. *FISH RES*, 63 (3):315 – 326, 2003. 31
- [62] E. Harvey and M. R. Shortis. A system for stereo-video measurement of sub-tidal organisms. *Marine Technology Society Journal*, (29(4)):10–22, 1996. 34
- [63] E. Harvey and M. R. Shortis. Calibration stability of an underwater stereo-video system: Implications for measurement accuracy and precision. *Marine Technology Society Journal*, (32(2)):3–17, 1998. 34
- [64] H. Hirschmüller and R. Liu. Terrain and city modeling from aerial images. http://www.dlr.de/rm/desktopdefault.aspx/tabid-4847/8044_read-13028/. 20.04.2011. 36
- [65] G. Hirzinger and U. Hagn. Mirosurge - ein innovatives robotik-system für die minimal invasive chirurgie. *forschung*, pages 4–7, 2009. 36
- [66] M. K. Hu. Visual pattern recognition by moment invariants. *Information Theory IRE Transactions on*, 8(2):179–187, 1962. 30
- [67] J. N. Ianelli, T. K. Wilderbuer, and D. Nichol. 5. assessment of greenland turbot in the eastern bering sea and aleutian islands. http://www.afsc.noaa.gov/REFM/docs/2010/BSAI_turbot.pdf, 2010. 7
- [68] G. Inglis and C. Roman. Terrain constrained stereo correspondence. In *OCEANS 2009 MTSIEEE Biloxi Marine Technology for Our Future Global and Local Challenges*, pages 1–9, 2009. 36
- [69] R. Jaumann, G. Neukum, T. Behnke, T. Duxbury, K. Eichentopf, J. Flohrer, S. Gasselt, B. Giese, K. Gwinner, and E. Hauber. The high-resolution stereo camera (hrsc) experiment on mars express: Instrument aspects and experiment conduct from interplanetary cruise through the nominal mission. *Planetary and Space Science*, 55(7-8):928–952, 2007. 36
- [70] F. V. Jensen. *An Introduction to Bayesian Networks*, volume 39. UCL Press, 1996. 27

- [71] B. Jähne. *Digitale Bildverarbeitung*. Springer Verlag, 2005. 18, 23, 25, 44
- [72] C. Johansson, M. Stowar, and M. Cappelletti. The use of stereo images for measuring fish size. Technical report, Report to the Marine and Tropical Sciences Research Facility. Reef and Rainforest Research Centre Limited, Cairns, 2008. Australian Governments Marine and Tropical Sciences Research Facility. 32
- [73] I. K. Jung and S. Lacroix. High resolution terrain mapping using low altitude aerial stereo imagery. *Proceedings Ninth IEEE International Conference on Computer Vision*, 2(Iccv):946–951 vol.2, 2003. 36
- [74] A. S. Kane, J. D. Salierno, G. T. Gipson, T. C. A. Molteno, and C. Hunter. A video-based movement analysis system to quantify behavioral stress responses. *Water Research, Elsevier*, 38:39934001, 2004. 34
- [75] R. Kashyap and R. Chellappa. Stochastic models for closed boundary analysis: Representation and reconstruction. *IEEE Transactions on Information Theory*, 27(5):627–637, 1981. 29
- [76] T. Kim, H. Kim, C. Oh, and B. Son. Traffic flow forecasting based on pattern recognition to overcome memoryless property. *2007 International Conference on Multimedia and Ubiquitous Engineering MUE07*, pages 1181–1186, 2007. 25
- [77] A. Kläser. Kamerakalibrierung und stereo vision. www.multires.caltech.edu/teaching/courses/3DP/notes/calib_notes.ps, 2005. 08.02.2013. 64
- [78] T. S. Kristiansen, A. Fernö, J. C. Holm, L. Privitera, S. Bakke, and J. E. Fosseidengen. Swimming behaviour as an indicator of low growth rate and impaired welfare in atlantic halibut (*hippoglossus hippoglossus* l.) reared at three stocking densities. *Aquaculture*, 230(14):137 – 151, 2004. 33
- [79] C. Kunz and H. Singh. Hemispherical refraction and camera calibration in underwater vision. *OCEANS 2008*, 2008. 35
- [80] Y. H. Kwon, R. H. Sanders, and B. J. Gibson. *A camera calibration algorithm for the underwater motion analysis*, pages 257–260. School of Biomedical and Sports Science, Edith Cowan University, Perth, WA, 1999. 34

- [81] J. M. Lavest, G. Rives, and J. T. Lapresté. Underwater camera calibration. *Computer Vision ECCV 2000*, 1843:654–668, 2000. 35
- [82] J. M. Lavest, G. Rives, and J. T. Lapresté. Dry camera calibration for underwater applications. *Machine Vision and Applications*, 13(5-6):245–253, 2003. 35
- [83] R. Li, H. Li, W. Zou, R.G. Smith, and T.A. Curran. Quantitative photogrammetric analysis of digital underwater video imagery. *Oceanic Engineering, IEEE Journal of*, (22):364 – 375, April 1997. 34
- [84] S. Z. Li and A. K. Jain. *Handbook of Face Recognition*, volume 54. Springer, 2005. 25
- [85] Y. Z. Li, S. G. Min, and X. Liu. Study on the methods and applications of near-infrared spectroscopy chemical pattern recognition. *Spectroscopy and Spectral Analysis*, 27(7):1299–1303, 2007. 25
- [86] J.A. Lines, R.D. Tillett, L.G. Ross, D. Chan, S. Hockaday, and N.J.B. McFarlane. An automatic image-based system for estimating the mass of free-swimming fish. *Computers and Electronics in Agriculture*, 31(2):151–168, 2001. 32, 117, 118
- [87] J. Liu, J. Sun, and S. Wang. Pattern recognition: An overview. *IJCSNS International Journal of Computer Science and Network Security*, 6(6):57–61, 2006. 25
- [88] VAKI Aquaculture Systems Ltd. Vaki is leading company in development, manufacture and marketing of hi-tech equipment for fish farming and the environmental market. <http://www.vaki.is/>. 08.02.2013. 31
- [89] Luxeon. power light source luxeon star - technical datasheet ds23, 2006. 57
- [90] Luxeon. power light source luxeon rebel - technical datasheet ds56, 2007. 57
- [91] H. A. Mallot. *Sehen und die Verarbeitung visueller Informationen*. Vieweg Verlag, 2000. 18, 20, 23, 25
- [92] J. P. Marques de S. *Pattern Recognition Concepts, Methods and Applications*. Springer, 2001. 25, 26, 27
- [93] H. Men, Y. Wu, Y. Gao, X. Li, and S. Yang. Application of support vector machine to pattern classification. *2008 9th International Conference on Signal Processing*, pages 1612–1615, 2008. 27

- [94] R. J. Mooney and U. Y. Nahm. Text mining with information extraction. *Language*, 7(September 2003):1–2, 2005. 25
- [95] A. Mustafa. Matching incomplete objects using boundary signatures. In Carlo Arcelli, Luigi Cordella, and Gabriella di Baja, editors, *Visual Form 2001*, volume 2059 of *Lecture Notes in Computer Science*, pages 563–572. Springer Berlin / Heidelberg, 2001. 10.1007/3-540-45129-3_52. 38
- [96] M. Nixon and A. Aguado. *Feature extraction image processing*. Academic Press, 2002. 26
- [97] A. Opelt, A. Pinz, and A. Zisserman. A boundary-fragment-model for object detection. *Computer Vision ECCV 2006*, 3952:575–588, 2006. 38, 85, 93, 115
- [98] A. Opelt, A. Pinz, and A. Zisserman. Incremental learning of object detectors using a visual shape alphabet. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Volume 1 CVPR06*, 1(c):3–10, 2006. 38, 85, 115
- [99] A. Opelt, A. Pinz, and A. Zisserman. Learning an alphabet of shape and appearance for multi-class object detection. *International Journal of Computer Vision*, 80(1):16–44, 2008. 85
- [100] OpenCV. Image processing - canny edge detector. http://opencv.itseez.com/doc/tutorials/imgproc/imgtrans/canny_detector/canny_detector.html#canny-detector. 22.04.2012. 89
- [101] OpenCV. Image processing - histogram equalization. http://opencv.itseez.com/doc/tutorials/imgproc/histograms/histogram_equalization/histogram_equalization.html#histogram-equalization. 22.04.2012. 88
- [102] OpenCV. Opencv (open source computer vision). <http://opencv.willowgarage.com/wiki/>. 30.04.2012. 37, 87
- [103] Y. Oya and K. Kawasue. Three dimensional measurement of fish movement using stereo vision. *Artificial Life and Robotics*, 13:69–72, 2008. 10.1007/s10015-008-0516-6. 34
- [104] A. Pandian. Robot navigation using stereo vision and polarization imaging. *Image Rochester NY*, (November):64, 2008. 36
- [105] A. S. Pandya and R. B. Macy. *Pattern Recognition with Neural Networks in C++*. CRC Press, 1996. 28

- [106] D. Paulus. Structure from motion. <http://www.uni-koblenz-landau.de/koblenz/fb4/institute/icv/agpaulus/agas-team/paulus/paulus-en>, 2003. 08.02.2013. 21
- [107] M. W. Pedersen. *Hidden Markov modelling of movement data from fish*. PhD thesis, Technical University of Denmark, Department of Informatics and Mathematical Modeling, Mathematical Statistics, 2010. 34
- [108] Pescanova. Grupo pescanova investors presentation, january 2010. http://www.pescanova.com/i/pdf/agenda_inversor/2010/HR_210110_InvPres_EN.pdf, 2010. 08.02.2013. 9
- [109] N. Pessel, J. Opderbecke, and M. J. Aldon. Camera self-calibration in underwater environment. In *WSCG*, 2003. 35
- [110] R. J. Petrell, X. Shi, R. K. Ward, A. Naiberg, and C. R. Savage. Determining fish size and swimming speed in cages and tanks using simple video techniques. *Aquacultural Engineering*, 16(12):63 – 84, 1997. 31
- [111] W. K. Pratt. *Digital Image Processing*. John Wiley & Sons, New York, second edition, 1991. 25
- [112] Pentax Ricoh Imaging Deutschland GmbH. C60402kp - h416 (kp) spezifikationen. <http://security-systems.pentax.de>. 06.04.2012. 48
- [113] B. D. Ripley. *Pattern Recognition and Neural Networks*, volume 40. Cambridge University Press, 1996. 28
- [114] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*. Academic Press, second edition, 1982. 25
- [115] A. Rova, G. Mori, and L. M. Dill. One fish, two fish, butterfly, trumpeter: Recognizing fish in underwater video. In *In IAPR Conference on Machine Vision Applications*, 2007. 32
- [116] B. P. Ruff, J. A. Marchant, and A. R. Frost. Fish sizing and monitoring using a stereo image analysis system applied to fish farming. *Aquacultural Engineering*, 14(2):155 – 173, 1995. 31
- [117] T. Sahin and E. Günes. A preliminary study on population characteristics of turbot, *psetta maxima*, in the eastern black sea, turkey. *Turkish Journal of Science & Technology*, 6:1–9, 2011. 7

- [118] N. Samsun, F. Kalayci, and O. Samsun. Seasonal variation in length, weight, and sex distribution of turbot (*scophthalmus maeoticus pallas*, 1811) in the sinop region (black sea) of turkey. *Turkish Journal of Zoology*, 31:371–378, 2007. 7
- [119] I. Schiller. Mip - multicameracalibration. <http://www.mip.informatik.uni-kiel.de/tiki-index.php?page=Calibration>. 02.05.2012. 62
- [120] I. Schiller, C. Beder, and R. Koch. Calibration of a pmd camera using a planar calibration object together with a multi-camera setup. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVII. Part B3a, pages 297–302, Beijing, China, 2008. XXI. ISPRS Congress. 62
- [121] J. C. Schneider. *Manual of fisheries survey methods II: with periodic updates*. Michigan Department of Natural Resources, Fisheries Special, 2000. 7
- [122] A. Sedlazeck and R. Koch. Calibration of housing parameters for underwater stereo-camera rigs. In *BMVC 2011, Proceedings of the British Machine Vision Conference*, 2011. 35, 113
- [123] M. Seise, S. J. McKenna, I. W. Ricketts, and C. A. Wigderowitz. Double contour active shape models. *British Machine Vision Conference*, 2:159168, 2005. 37
- [124] C. Serra-Toro, R. Montoliu, V. J. Traver, I. M. Hurtado-Melgar, M. Nunez-Redo, and P. Cascales. Assessing water quality by video monitoring fish swimming behavior. *2010 20th International Conference on Pattern Recognition*, 0:428–431, 2010. 34
- [125] L. Shapiro and G. Stockman. *Computer Vision*. Prentice Hall, 2001. 23, 25
- [126] C. Shi and B. Wang. Shape retrieval using statistical chord-length features. In L. W. Chang and W. N. Lie, editors, *Advances in Image and Video Technology*, volume 4319 of *Lecture Notes in Computer Science*, pages 403–410. Springer Berlin / Heidelberg, 2006. 10.1007/11949534_0.37
- [127] M. R. Shortis and E. Harvey. Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. *International Archives Photogrammetry and Remote Sensing*, (32(5)):792–799, 1998. 34

- [128] J. Shotton, A. Blake, and R. Cipolla. Contour-based learning for object detection. In *ICCV*, pages 503–510, 2005. 38
- [129] A. Silvatti, M. Rossi, F. Dias, N. Leite, and R. Barros. Non-linear camera calibration for 3d reconstruction using straightline plane object. *27 International Conference on Biomechanics in Sports*, 2009. 34
- [130] A. Silvatti, M. Rossi, F. Dias, T. Telles, and R. Barros. Underwater non-linear camera calibration: An accuracy analysis. *28 International Conference on Biomechanics in Sports*, 2010. 34
- [131] L. H. Stien, S. Bratland, I. Austevoll, F. Oppedal, and T. S. Kristiansen. A video analysis procedure for assessing vertical fish distribution in aquaculture tanks. *Aquacultural Engineering*, 37(2):115 – 124, 2007. 33
- [132] F. Storbeck and B. Daan. Weight estimation of flatfish by means of structured light and image analysis. *Fisheries Research*, 11(2):99 – 108, 1991. 30, 117
- [133] F. Storbeck and B. Daan. Fish species recognition using computer vision and a neural network. *Fisheries Research*, 51(1):11 – 15, 2001. 32
- [134] N. J. C. Strachan. Length measurement of fish by computer vision. *Computers and Electronics in Agriculture*, 8(2):93 – 104, 1993. xiii, 30, 116, 117
- [135] N. J. C. Strachan, P. Nesvadba, and A. R. Allen. Fish species recognition by shape analysis of images. *Pattern Recogn.*, 23(5):539–544, June 1990. 32
- [136] C. J. Sze, H. R. Tyan, H. Y. M. Liao, C. S. Lu, and S. K. Huang. Shape-based retrieval on a fish database of taiwan. *Tamkang Journal of Science and Engineering*, 2(3):163–173, 1999. 32
- [137] G. Telem and S. Filin. Photogrammetric modeling of underwater environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5):433 – 444, 2010. 34
- [138] S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. Third edition, 2006. 27
- [139] U. H. Thygesen, M. W. Pedersen, and H. Madsen. Geolocating fish using hidden markov models and data storage tags. In Jennifer L. Nielsen, Nuno Fragoso, Molly Lutcavage, Haritz Arrizabalaga, Alistair Hobday,

- and John Sibert, editors, *Tagging and Tracking of Marine Animals with Electronic Devices*, volume 9 of *Reviews: Methods and Technologies in Fish Biology and Fisheries*, pages 277–293. Springer Netherlands. 34
- [140] R. D. Tillett, N. J. B. Mcfarlane, and J. Lines. Estimating dimensions of free-swimming fish using 3d point distribution models. *CVIU - Computer Vision and Image Understanding*, 79:2000, 2000. xiii, 32, 116, 117, 118
- [141] T. Treibitz, Y. Y. Schechner, and H. Singh. Flat refractive geometry. *IEEE Conference on Computer Vision and Pattern Recognition (2008)*, 34(Cvpr):1–8, 2008. 35, 63, 113
- [142] R. Y. Tsai. An efficient and accurate camera calibration technique for 3d machine vision. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 364–374, 1986. 15, 62
- [143] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *RA-3(4):323–344*, 1987. 15, 62
- [144] W. Wang, B. Guo, X. Li, and J. Cao. Influence factors evaluation on high-precision planar calibration of non-metric digital camera. *ISPRS - International Society for Photogrammetry and Remote Sensing*, 37:1071–1076, 2008. 64
- [145] A. R. Webb. *Statistical Pattern Recognition*, volume 9. Wiley, 2002. 27
- [146] J. Weule. *Iteration nichtlinearer Gauß-Filter in der Bildverarbeitung*. PhD thesis, Heinrich-Heine-Universität Düsseldorf, 1994. 23
- [147] D. J. White, C. Svellingen, and N. J. C. Strachan. Automated measurement of species and length of fish by computer vision. *Fisheries Research*, 80(23):203 – 210, 2006. 31, 117
- [148] D. Wooden, M. Malchano, K. Blankespoor, A. Howard, A. A. Rizzi, and M. Raibert. Autonomous navigation for bigdog. *Jet Propulsion*, 23(11/12):4736–4741, 2010. 36
- [149] A. Young. *Handbook of Pattern Recognition and Image Processing*. Academic Press Inc., 1986. 23, 25
- [150] D. Zhang and G. Lu. Generic fourier descriptor for shape-based image retrieval. *Proceedings IEEE International Conference on Multimedia and Expo*, 1(1):425–428, 2002. 30

- [151] D. Zhang and G. Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37, 2004. 28
- [152] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000. 15
- [153] L. Zhu and W. Weng. Catadioptric stereo-vision system for the real-time monitoring of 3d behavior in aquatic animals. *Physiology and Behavior*, 91(1):106 – 119, 2007. 34
- [154] B. Zion, A. Shklyar, and I. Karplus. Sorting fish by computer vision. *Computers and Electronics in Agriculture*, 23(3):175 – 187, 1999. 32

Lebenslauf Christian Bauer

- 17.01.1980 geboren in Bremen
- 1998 Abitur
- 11.98–09.99 Bundeswehr
Grundwehrdienst
- 10.99–09.00 Studium der Informatik
an der Christian-Albrechts-Universität zu Kiel
- 10.00–08.05 Studium der Computervisualistik
an der Universität Koblenz-Landau, Campus Koblenz
Abschluss: Diplom Informatik
- 02.06–01.09 Wissenschaftlicher Mitarbeiter
an der Universität Flensburg, Berufsbildungsinstitut
Arbeit und Technik
- 03.09–06.12 Wissenschaftlicher Mitarbeiter / Doktorand
an der Christian-Albrechts-Universität zu Kiel,
Institut für landwirtschaftliche Verfahrenstechnik

