



Forschungs- und Lehrereinheit Informatik VI: Robotics and Embedded Systems

Modellierung des Operationsstatus in der robotergestützten
minimalinvasiven Herzchirurgie

István Nagy

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen
Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. N. Navab, Ph.D.

Prüfer der Dissertation:

1. Univ.-Prof. Dr. A. Knoll
2. Univ.-Prof. Dr. B. Radig

Die Dissertation wurde am 14. März 2008 bei der Technischen Universität München
eingereicht und durch die Fakultät für Informatik am 7. April 2009 angenommen.

Zusammenfassung

Die moderne Chirurgie hat in den letzten Jahrzehnten die Lebenserwartung und die Lebensqualität nach Operationen beachtlich erhöht. Eine bedeutende Verbesserung, außer technischer und pharmazeutischer Weiterentwicklungen im Operationssaal, brachte die Einführung der minimalinvasiven Chirurgie. Die Vorteile für Patienten werden jedoch oft durch Nachteile für den operierenden Chirurgen aufgewogen. Die eingeschränkte Sicht erschwert die Orientierung, die Bedienung der langen Instrumente um die kleinen Eintrittsöffnungen erfordert eine angepasste Hand–Auge–Koordination und der Hebeleffekt verstärkt den Handtremor erheblich. Roboter-gestützte Telemanipulatoren beseitigen durch konstruktions- und regelungstechnische Maßnahmen diese Probleme. Doch die neue Technik brachte auch Nachteile, wobei die zwei eminentesten die fehlende taktile und kinästhetische Wahrnehmung und die deutlich verlängerte Operationsdauer sind. Sie motivierten die Entwicklung zweier Experimentierplattformen zur Evaluation von Krafrückkopplung und zur Integration teilautonomer Techniken.

Der Schwerpunkt dieser Dissertation liegt in der Verarbeitung von Bildern, die mit Hilfe medizinischer Stereoendoskope gewonnen wurden. Insbesondere wird die dreidimensionale Modellierung des Operationssitus im Kontext robotergestützter minimalinvasiver Eingriffe aus der Herzchirurgie behandelt. Die Arbeit bietet zunächst einen Überblick kommerziell erhältlicher Systeme und Forschungsaufbauten für die robotergestützte Chirurgie. Anhand noch vorhandener Unzulänglichkeiten verfügbarer Systeme wird die Vision eines teilautonomen chirurgischen Manipulators vorgestellt. Die primäre Zielsetzung ist, das System mit Fähigkeiten auszustatten, die ihm die Ausübung autonomer Assistenz Tätigkeiten ermöglichen. Um die Steuerungsschleife Aktorik/Sensorik zu schließen, wird das chirurgische Nahtmaterial in der Endoskopsicht segmentiert und anschließend zusammen mit dem Herzgewebe rekonstruiert, modelliert und dem Robotersystem zur Verfügung gestellt.

Das Ergebnis ist ein Subsystem zur Bildverarbeitung, das in die Steuerung der Roboterarme integriert wurde, und somit die Durchführung einfacher autonomer Greifvorgänge ausschließlich über den visuellen Kanal ermöglicht. Der rekonstruierte und modellierte Situs kann außerdem in eine Simulationsumgebung integriert, oder als Plattform für Augmentierungstechniken aufgrund aufbereiteter Texturen fotorealistic dargestellt werden. Die Anwendbarkeit konnte in einem, während dieser Arbeit entstandenen, realistischen Aufbau für robotergestützte minimalinvasive Chirurgie gezeigt werden.

Inhaltsverzeichnis

1	Einleitung	7
1.1	Motivation	7
1.2	Roboterassistierte Telemanipulatorchirurgie	9
1.3	Klinische Relevanz	11
1.4	Vision eines teilautonomen Chirurgesystems	13
1.5	Problemstellung und verwandte Arbeiten	14
1.6	Gliederung der Arbeit	15
2	Chirurgie und Robotik	17
2.1	Minimalinvasive Chirurgie	19
2.2	Chirurgieroboter	21
2.3	Roboterassistierte minimalinvasive Chirurgie	25
2.3.1	Vorteile	28
2.3.2	Nachteile	29
2.4	Stand der Forschung	30
2.5	Das Endo[PA]R System	35
2.6	Das ARAMIS ² System	39
2.7	Aufbau in der Großtier-OP am DHM	41
3	Kalibrierung	43
3.1	Motivation	45
3.2	Szenario	45
3.3	Begriffe und Notationen	46
3.3.1	3D-Koordinaten	47
3.3.2	Translationen (Verschiebungen)	47

3.3.3	Rotationen (Drehungen)	49
3.3.4	Rigide Transformationen	53
3.3.5	Homogene Transformationsmatrizen	53
3.4	Kameramodell und Parameter	56
3.4.1	Binokularer Aufbau	56
3.4.2	Bildgebende Sensoren und Linsen	56
3.4.3	Das Lochkameramodell	58
3.5	Kamerakalibrierung	67
3.5.1	Mathematische Grundlagen	67
3.5.2	Kalibriermethoden	69
3.5.3	Kalibrierergebnisse	70
3.6	Hand–Auge–Kalibrierung	72
3.6.1	Von Kamera– nach Weltkoordinaten	73
3.6.2	Richtungskosinus	75
3.6.3	Von Welt– nach Roboterkoordinaten	77
4	Bildvorverarbeitung	79
4.1	Spiegelreflexionen	81
4.1.1	Einführung	81
4.1.2	Vorarbeiten auf dem Gebiet	82
4.1.3	Erkennung von Spiegelreflexionen	83
4.1.4	Kettencodes zur Speicherung von Kontourdaten	90
4.1.5	Rekonstruktion mittels Linearinterpolation	91
4.1.6	Rekonstruktion mittels Collapsing Chains	97
4.1.7	Regularisierung vektorwertiger Daten mittels partiellen Differenzialgleichungen (RVWPDGL)	99
4.1.8	Evaluation der vorgestellten Verfahren	102
4.1.9	Visueller Vergleich anhand von Videosequenzen	107
4.2	Anwendung der Regularisierung von Bildern mittels PDGLen zur Rauschunterdrückung	110

5 Fadenerkennung	113
5.1 Motivation	115
5.2 Chirurgisches Nahtmaterial	116
5.3 Kantenfilter	118
5.3.1 Segmentierung von Linien nach Steger	122
5.3.2 Anwendungsbeispiele für chirurgisches Nahtmaterial	127
6 3D-Rekonstruktion des Situs	141
6.1 Motivation	143
6.2 Binokulare Stereobildverarbeitung	145
6.2.1 Abbildungsgeometrie	145
6.2.2 Genauigkeitsermittlung mittels 3D-Kalibrierkörper	146
6.2.3 Rekonstruktion von 3D-Punktwolken	150
6.3 Symbolische Repräsentation des Fadens	154
6.3.1 Splines	155
6.3.2 Evaluation der Genauigkeit der Splineinterpolation	159
6.4 Modellierung der Herzoberfläche	161
6.4.1 Bézier-Flächen	161
6.4.2 Texturierung	164
7 Zusammenfassung und Ausblick	167
7.1 Zusammenfassung	167
7.2 Ausblick	169

Kapitel 1

Einleitung

1.1 Motivation

Die Herzchirurgie vollzog seit ihren Anfängen in den fünfziger Jahren dieses Jahrhunderts bemerkenswerte Fortschritte in der Behandlung kardiovaskulärer Erkrankungen. Diese führen seit vielen Jahren die Mortalitätsstatistiken der meisten Industrienationen an, und sind auf demographische und ernährungsbedingte Ursachen zurückzuführen. Dabei stellen Erkrankungen der Herzkranzgefäße und Herzklappenfehler die häufigsten Indikationen für einen operativen Eingriff dar. Der Koronararterien-Bypass ist heute die effektivste Behandlungsmethode der koronaren Herzkrankheit und dadurch der häufigste Eingriff am Herzen. Wie alle Eingriffe in der konventionellen Herzchirurgie, benötigt auch der Bypass eine höchst invasive Technik, die Sternotomie, eine chirurgische Durchtrennung des Brustbeins um Zugang zum Operationsgebiet zu erlangen. Die Folge ist ein langwieriger Heilungsprozess, begleitet von einem hohen Infektionsrisiko bedingt durch das operative Trauma. Eine weitere Belastung für das Immunsystem stellt die Notwendigkeit einer Herz–Lungen–Maschine dar, die bei stillgelegtem Herzen die Pumpfunktion und die Sauerstoffversorgung übernimmt.

Neben dieser Art der operativen Behandlung der koronaren Herzkrankheit, die mittlerweile zu einer standardisierten Behandlungsform wurde, gibt es noch verschiedene medikamentöse Therapien, die Dilatationsbehandlung und die Stentimplantation. Motiviert durch den Wunsch der Traumaminimierung und bestärkt durch Erfolge in anderen chirurgischen Disziplinen wurden auch in der Herzchirurgie verschiedene minimalinvasive Verfahren eingeführt. Mangels eines einheitlichen Verfahrens kann man jedoch in der Herzchirurgie den Begriff *minimalinvasiv* nicht einer bestimmten Technik zuordnen. Es haben sich jedoch zwei Zielrichtungen herauskristallisiert, die den Patienten schonen sollen. Zum einen wird versucht, durch Minimierung des Zugangs zum Operationsgebiet das Knochen– und Gewebstrauma zu reduzieren. Zum anderen erlauben spezielle Techniken die Operation am schlagenden Herzen,

wodurch auf die extrakorporale Blutzirkulation verzichtet werden kann. Eine Kombination aus beiden, die total endoskopische Operation am schlagenden Herzen, konnte jedoch bis jetzt nicht minimalinvasiv durchgeführt werden.

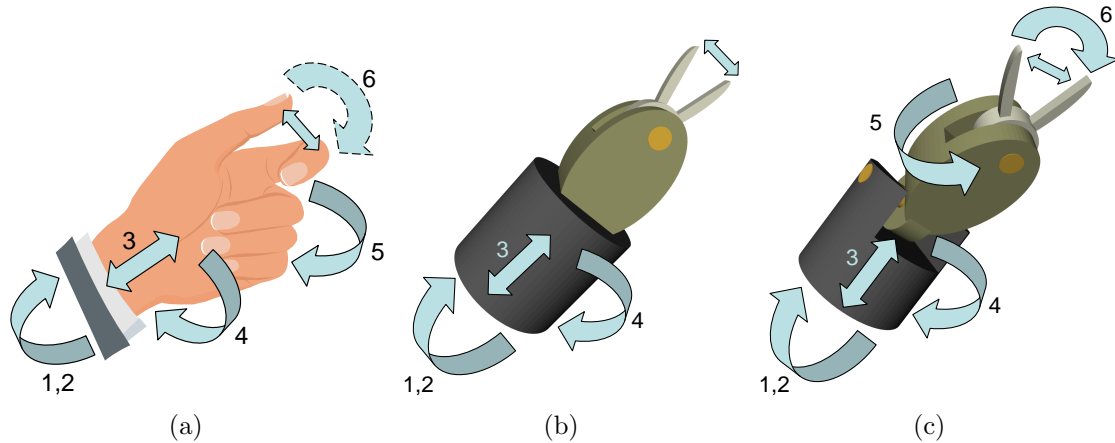


Abbildung 1.1: Freiheitsgrade der Hand und endoskopischer Instrumente

Der Hauptgrund hierfür ist das Fehlen von Freiheitsgraden (engl. *degrees of freedom*, DOF) im Körperinneren wegen mechanischer Einschränkungen klassischer endoskopischer Instrumente. Die menschliche Hand besitzt, wenn man nur Daumen und Zeigefinger als einfache Greifer berücksichtigt, sechs Freiheitsgrade, wobei der sechste nur beschränkt von Nutzen ist (siehe Abbildung 1.1(a)). Die ersten zwei erlauben eine Bewegung in der Ebene, wobei aus Gründen der Übersichtlichkeit nur ein Pfeil eingezeichnet wurde. Der dritte Freiheitsgrad ermöglicht den Vor- und Rücklauf längs der Armrichtung, der vierte rotatorische Freiheitsgrad eine Drehung um die Armachse. Der fünfte Freiheitsgrad entsteht durch die Abwinkelbarkeit des Handgelenks.

Bedingt durch die invarianten Eintrittsöffnungen (*Ports*) am Brustkorb und durch konstruktionsbedingte Einschränkungen endoskopischer Instrumente reduziert sich die Anzahl ihrer Freiheitsgrade auf vier. Abbildung 1.1(b) zeigt schematisch die Spitze eines klassischen endoskopischen Instruments. Der Chirurg hat hierbei lediglich die Möglichkeit die Backen des Greifers zu bedienen, wobei der Verlust des fünften Freiheitsgrades ihn bei der Ausrichtung erheblich einschränkt. Desweiteren muss er auf haptische Wahrnehmung gänzlich verzichten. Die langen Instrumente verstärken den Handtremor und erschweren die Hand-Auge-Koordination durch die Notwendigkeit inverser Handbewegungen. Diese fehlende Übereinstimmung von Hand- und Instrumentenbewegungen im Endoskopbild verhindert ergonomisches Arbeiten.

Abbildung 1.1(c) zeigt das Modell eines neuartigen im Körperinneren abwinkelbaren endoskopischen Instruments. Trotz der Wiedererlangung aller Freiheitsgrade einer menschlichen Hand durch konstruktionstechnische Maßnahmen, ist ein solches Instrument von einem Chirurgen ohne Hilfe nicht zu bedienen. Denn bei der üblichen

Handhabung endoskopischer Instrumente besteht keine Möglichkeit, einen zusätzlichen Freiheitsgrad manuell ergonomisch zu steuern. Betrachtet man jedoch ein solches Instrument als einen mehrgelenkigen Roboter, und bietet dem Chirurgen mit Hilfe spezieller Eingabegeräte eine intuitive Form der Fernsteuerung, so ergeben sich ganz neue Möglichkeiten, die wir im folgenden vorstellen.

1.2 Robotergestützte Telemanipulatorchirurgie

Das Modell aus Abbildung 1.1(c) stellt einen Greifer des seit Anfang der 1990er Jahre klinisch zugelassenen und kommerziell erhältlichen Telechirurgiesystems *da VinciTM* der Firma Intuitive Surgical Inc. dar. Abbildung 1.2 zeigt eine Installation des Systems am Deutschen Herzzentrum München (DHM).

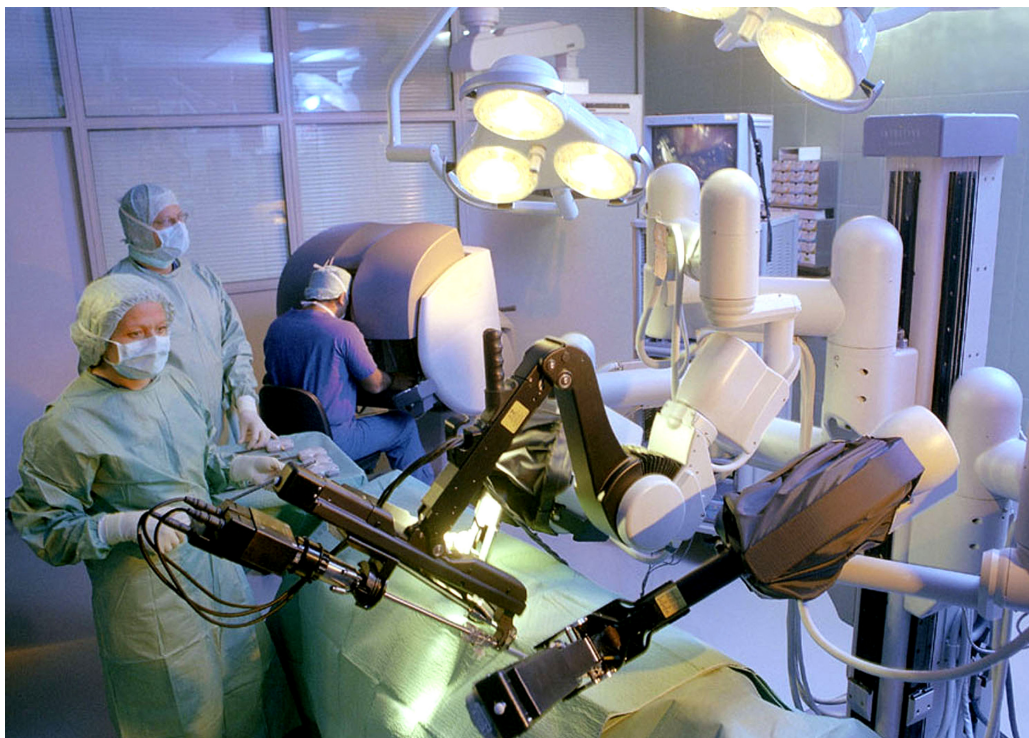


Abbildung 1.2: Das minimalinvasive Telechirurgiesystem *da VinciTM* am DHM

Um die Nachteile klassischer manuell bedienter endoskopischer Instrumente zu vermeiden, führen robotergestützte Telechirurgiesysteme eine räumliche Trennung zwischen Instrumenten und Operateur ein. Dadurch dass die Instrumente nicht mehr vom Chirurgen sondern von Roboterarmen gehalten und angetrieben werden, wird erst die Bedienung hinzugekommener Freiheitsgrade möglich. Spezielle Eingabegeräte sorgen für die intuitive Abbildung menschlicher Handbewegungen auf die der chirurgischen Instrumente. Außer den zusätzlichen Freiheitsgraden bieten sol-

che Systeme durch regelungstechnische Maßnahmen die Reduktion des Handtremors und mehrere Skalierungsstufen. Die Notwendigkeit inverser Handbewegungen entfällt ebenfalls. Nichtsdestotrotz bleibt die Problematik der fehlenden taktilen und kinästhetischen Wahrnehmung größtenteils bestehen, obwohl es bereits einige vielversprechende Entwicklungen auf dem Gebiet gibt.

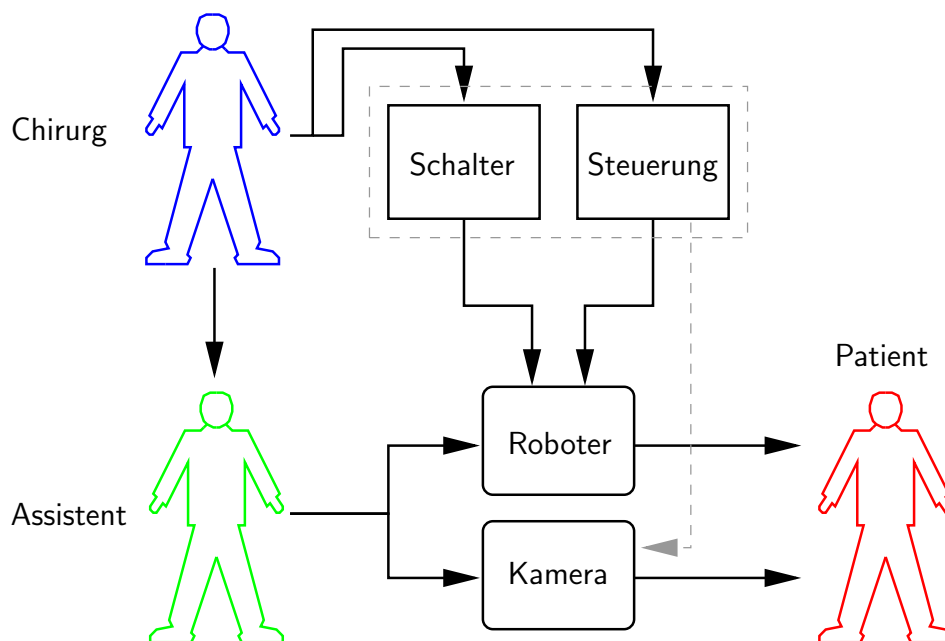


Abbildung 1.3: Klassischer Telechirurgieaufbau

Robotergestützte Chirurgesysteme sind prinzipbedingt typische Master–Slave–Architekturen, und bestehen aus zwei Hauptkomponenten: aus einer sogenannten Konsole (Master) und einem Operator (Slave). Abbildung 1.3 zeigt schematisch die Komponenten eines solchen Systems. Die Konsole (siehe auch Abbildung 1.2) dient als Ein- und Ausgabeeinheit für den Chirurgen, der patientenseitige Operator führt die an der Konsole ausgelösten Vorgänge durch, der Roboter operiert also nicht autonom. Die Einstellung etwaiger Systemparameter, wie beispielsweise die Skalierung der Bewegungen, oder das Maß der Tremorreduktion, werden vom Chirurgen ebenfalls an der Konsole vorgenommen. Dafür muss jedoch der Operateur seine Eingabegeräte “auskuppeln”, in eine Parkposition bringen, oder die Hilfe einer menschlichen Assistenz in Anspruch nehmen, deren Notwendigkeit auch bei solchen Systemen nicht entfällt. Sie übernimmt desweiteren den Tausch von Instrumenten, die Versorgung mit benötigten Implantaten oder die Repositionierung der Kamera (Abbildung 1.2).

Eine optimale Konfiguration wäre jedoch ein Arbeitsplatz, welcher die Bedienung aller Funktionalitäten eines solchen Systems ohne den Verlust der Immersion seitens des Chirurgen, und die Bedienung aller Komponenten ohne eine menschliche

Assistenz ermöglicht. Einige zeitaufwendige, aber immer wiederkehrende Aufgaben wie beispielsweise das Knotenlegen, haben ein hohes Potential für Automatisierung, und könnten unter Ausnutzung der ohnehin verfügbaren Robotertechnologie die Operationsdauer mancher Eingriffe verkürzen. Keines der kommerziell erhältlichen Telechirurgiesysteme nutzt die hochwertigen Stereoendoskope zur maschinellen Bildverarbeitung, oder die Stereosichtsysteme zur Augmentierung. Die Erstellung von Übersichtsmodellen zur einfacheren Navigation, die visuelle Überwachung des Arbeitsraums oder die Einblendung klinisch relevanter Daten sind nur einige der Möglichkeiten. In einem späteren Abschnitt werden wir die Vision eines teilautonomen robotergestützten Chirurgiesystems skizzieren, die für einige Unzulänglichkeiten heute verfügbarer Systeme Lösungsvorschläge bietet und weitere Techniken zur Verbesserung der Ergonomie, Effizienz und Sicherheit vorstellt.

1.3 Klinische Relevanz

Erst mit der Verfügbarkeit von Telechirurgiesystemen konnte erstmals die total endoskopische Bypassoperation (Totally Endoscopic Coronary Artery Bypass, TECAB) am schlagenden Herzen durchgeführt werden [1]. Abbildung 1.4 zeigt zwei Momentaufnahmen eines totalendoskopischen Eingriffs, durchgeführt mit dem *da Vinci*-System am Deutschen Herzzentrum München (DHM).

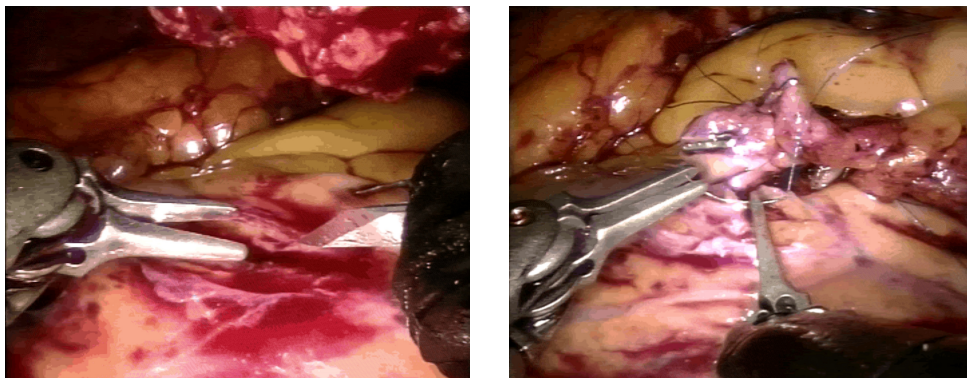


Abbildung 1.4: Schneiden und Nahtlegen mit *da Vinci*TM am DHM

Die große Bandbreite verfügbarer Instrumente und entsprechend angepasster Implantate ermöglicht auch die Durchführung anderer Eingriffsarten, wie beispielsweise den Verschluss des Vorhofscheidewanddefektes (Atrium Septum Defekt, ASD) oder die Rekonstruktion der Mitralk- und Aortenklappe. Wir beschränken uns in diesem Abschnitt auf Ergebnisse aus der Abdominal- und Thorakalchirurgie, und auf die zwei wichtigsten Vertreter robotergestützter Chirurgiesysteme *da Vinci* und das Konkurrenzprodukt *ZEUS* der Firma Computer Motion.

Die Studie aus [2] zeigt, dass sich das *da Vinci*-System prinzipiell zur Rekonstruktion von Herzklappen eignet und dass man nur wenige Einschränkungen in Kauf

nehmen muss. Die Autoren bemängeln jedoch die Verfügbarkeit einer integrieren Nähetechnologie. Bei einer großangelegten Studie [3] wurden bei 131 Patienten Koronararterien-Bypässe gelegt und bei 17 Patienten Mitralklappenrekonstruktionen durchgeführt. Als Ergebnis wurde festgehalten, dass für ausgewählte Patienten Eingriffe am stillgelegten Herzen mit dem *da Vinci*-System eine sichere Alternative darstellen, bieten jedoch keine nennenswerten Vorteile gegenüber der klassischen minimalinvasiven Technik. Der Grund ist der weiterhin notwendige Einsatz einer Herz-Lungen-Maschine. In [4] wird die Verwendbarkeit des Systems auch für Eingriffe im kleinen Becken gezeigt. Anhand von über 900 Fällen wird versucht, die Einsatzfelder robotergestützter Chirurgie für urologische Eingriffe zu definieren. Die meisten Erfolgsmeldungen werden bei der Behandlung der Prostata, der Gallenblase und der Niere verzeichnet.

Obwohl es instrumentenseitig mit einem Freiheitsgrad weniger ausgestattet als *da Vinci*, verbuchte auch das *ZEUS*-System zahlreiche klinische Erfolge. Insbesondere mit dem hauseigenen sprachgesteuerten System *AESOP* zur Nachführung des Endoskops erweist sich *ZEUS* als würdiger Konkurrent. In einer Studie [5] wurden bei einem Kollektiv bestehend aus 25 Patienten endoskopische Bypassoperationen durchgeführt. Davon fanden zehn mit Hilfe endoskopischer Stabilisatoren am schlagenden Herzen und ohne die Verwendung einer Herz-Lungen-Maschine statt. Bei der Verbindung der Blutgefäße (Anastomose) gab es interessanterweise keine nennenswerten Unterschiede in der Ausführungszeit der Eingriffe am stabilisierten oder stillgelegten Herzen. Ein grundlegend verschiedenes Anwendungsgebiet eines robotergestützten Telemanipulators wurde in [6] untersucht. Während offener mikrochirurgischer Experimente sollten die Vorteile der Bewegungsskalierung, der Tremorfilterung und des vergrößerten und hochaufgelösten Bildes des *ZEUS*-Systems evaluiert werden. Unter Hinzunahme eines Operationsmikroskops wurden verschiedene Konfigurationen zur Bilddarstellung ausprobiert. Erwartungsgemäß lieferte das Operationsmikroskop das beste visuelle Ergebnis, doch in Kombination mit den Roboter-Manipulatoren erwies sich *ZEUS* als eine sehr gute Alternative zu einer klassischen Mikrochirurgie-Arbeitsstation.

Wir werden in einem späteren Abschnitt weitere robotergestützte Chirurgesysteme auch aus anderen Bereichen der Chirurgie vorstellen und Referenzen auf klinische Resultate angeben. Trotz einiger Erfolge konnte sich diese Operationstechnik in der Abdominal- und Thorakalchirurgie aber nicht breitflächig durchsetzen können. Ein wichtiger Grund hierfür ist sicherlich die fehlende taktile und kinästhetische Wahrnehmung bei der Manipulation von Weichteilen. Auch die nicht nennenswert gesunkene Operationsdauer, im Vergleich mit herkömmlichen laparoskopischen Eingriffen, bei weitaus höheren Kosten, wirken sich negativ auf die generelle Akzeptanz aus. Eine Einschränkung auf nur bestimmte Eingriffsarten bedingt durch (noch) mangelnde Instrumentierung [7] und die fehlende Unterstützung beim Nähen erweisen sich ebenfalls als Schwachstellen.

1.4 Vision eines teilautonomen Chirurgesystems

Motiviert durch vorhandenes Verbesserungspotential kommerziell erhältlicher Systeme und bestärkt durch ihre bisherigen klinischen Erfolge, wurde am Lehrstuhl “Robotics and Embedded Systems” der Technischen Universität München in Zusammenarbeit mit dem Deutschen Herzzentrum München eine realitätsnahe Experimentierplattform für robotergestützte minimalinvasive Chirurgie aufgebaut. Die primären Ziele waren zunächst die Ausrüstung der chirurgischen Instrumente mit Kraftsensorik und die Integration teilautonomer Techniken zur Entlastung des Chirurgen und Ersetzung der menschlichen Assistenz. Unsere Vision eines teilautonomen Chirurgesystems lässt sich anhand der Darstellung aus Abbildung 1.5 erklären.

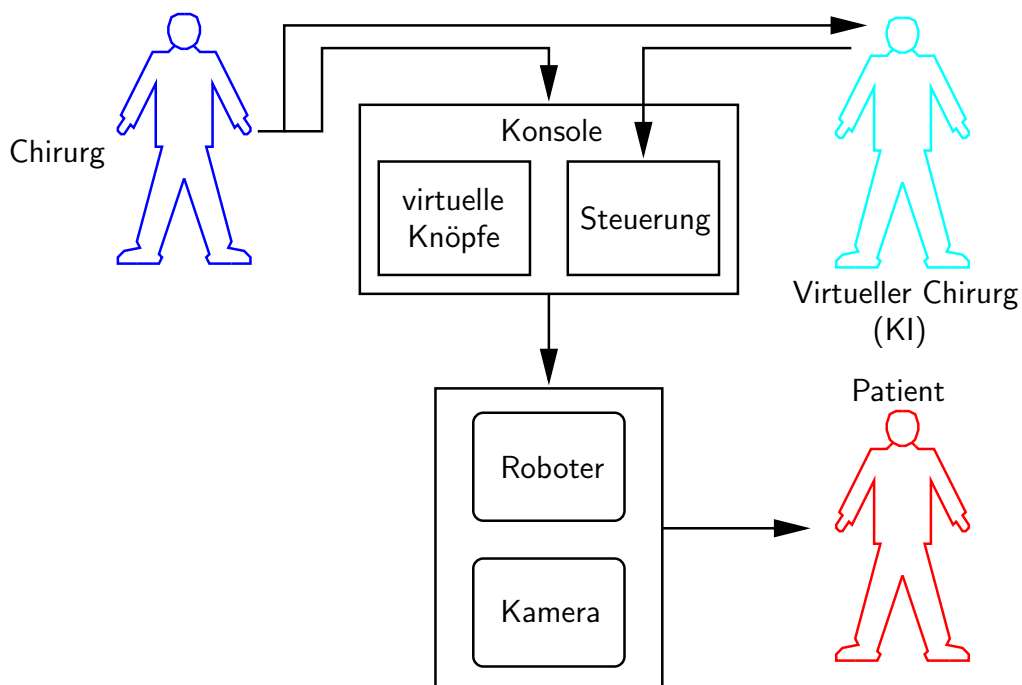


Abbildung 1.5: Schematischer Aufbau eines teilautonomen Telechirurgiesystems

Der menschliche Operateur bedient von einer mit Stereosicht und Krafrückkopplung ausgestatteten Konsole aus die gesamte Funktionalität des Teleoperators. Systemparameter werden mit Hilfe einer dreidimensionalen haptischen Benutzerschnittstelle eingestellt, die vollständig in die Bedienerkonsole integriert ist. Die Ausführung teilautonomer Aufgaben wird entweder per Fußpedal ausgelöst, oder situationsbedingt vom virtuellen Chirurgen initiiert. Sowohl Instrumentenarme als auch der Endoskoparm sind in die Steuerung integriert (vgl. Abb. 1.3) und können vom Operateur in ihre Initialposition gebracht oder repositioniert werden. Das Bildverarbeitungsmodul segmentiert das chirurgische Nahtmaterial und liefert eine 3D-Rekonstruktion des Operationssitus zur Unterstützung teilautonomer Ausführungen.

1.5 Problemstellung und verwandte Arbeiten

Die in dieser Dissertation vorgestellten Ergebnisse sind Teil der Arbeiten im Teilprojekt “Teilautonomie und multimodale Instruktion zum Transfer von Fertigkeiten” des DFG-geförderten [8] Sonderforschungsbereichs (SFB) 453 “Wirklichkeitsnahe Telepräsenz und Teleaktion” [9]. Eines der Ziele des Teilprojektes ist die Integration teilautonomer Techniken in robotergestützte minimalinvasive Chirurgieszenarien. In Zusammenarbeit mit dem Projektpartner von der Klinik für Herz- und Gefäßchirurgie am Deutschen Herzzentrum München [10] entstand der Aufbau einer realitätsnahen Experimentierplattform für herzchirurgische minimalinvasive und endoskopische Operationen, und dient der Verifikation neu entwickelter Techniken.

Als exemplarische automatisierbare Aufgabe wurde das Knotenlegen ausgewählt. Laparoskopische Instrumentalknoten benötigen wegen ihrer Komplexität einen erhöhten Zeitaufwand, und haben somit ein erhebliches Zeitersparnispotenzial, welches mittels robotergestützter Techniken ausgenutzt werden kann. Aus Sicherheitsgründen kann nicht der gesamte Knotenvorgang der Maschine überlassen werden, sodass sich der Ablauf wie folgt darstellt: i) der Chirurg initiiert an der gewünschten Stelle die Ausführung eines Knotens; ii) der Wickelvorgang wird von der Maschine kraftgeregelt und autonom ausgeführt; iii) bei der automatischen Komplettierung des Knotens wird das Fadenende gegriffen und kraftgeregelt festgezogen.

Für das autonome Greifen des Fadenendes ist dessen genaue Position im Koordinatensystem der Greifer notwendig. Diese kann aus visuellen Daten nur mit Hilfe eines Stereokamerasystems in ausreichender Genauigkeit gewonnen werden. Die Aufgabe der Bildverarbeitung ist demnach das Auffinden des Nahtmaterials im endoskopischen Bild, die 3D-Rekonstruktion des Fadenverlaufs aus dem Stereobildpaar, und die anschließende Modellbildung der Szene. Eine symbolische Repräsentation des Operationssitus kann nachfolgend dem Robotersystem zur Verfügung gestellt werden. Obwohl keine primäre Zielsetzung, ist die Rekonstruktion und Modellierung des Hintergrunds, beispielsweise der Herzoberfläche, ebenfalls möglich. Die Verfügbarkeit genauer Daten des Szenenmodells ermöglichen außerdem die Durchführung von Kollisionserkennung zwischen Instrumenten und Hintergrundgewebe, und dient als ein zusätzlicher Sicherheitsmechanismus.

Auf dem Gebiet der 3D-Rekonstruktion von anatomischen Strukturen mittels Bildverarbeitung gibt es bereits einige Vorarbeiten. Erwähnenswert sind Veröffentlichungen auf dem Gebiet der Neurochirurgie [11, 12], wo besondere Anforderungen an die Rekonstruktionsgenauigkeit gestellt werden. Einige Arbeiten [13, 14] auf dem Gebiet der Herzchirurgie waren lediglich durch den Wunsch motiviert, dem Chirurgen eine augmentierte Sicht vom Operationsgebiet zu präsentieren. Nach Kenntnisstand des Autors gibt es keine bisherige vergleichbare Arbeit, welche die Segmentierung und 3D-Rekonstruktion des Nahtmaterials zum Zweck der Manipulation in realistischen minimalinvasiven Chirurgieszenarien behandelt.

1.6 Gliederung der Arbeit

Der einführende Abschnitt 2 “Chirurgie und Robotik” bietet einen Einstieg in das Themengebiet der Roboterchirurgie. Es werden zunächst klassische offene chirurgische Eingriffe und minimalinvasive Verfahren gegenübergestellt. Anhand der Nachteile laparoskopischer Eingriffe wird die Integration robotergestützter Techniken motiviert. Nach der Vorstellung einiger kommerziell erhältlicher Systeme für Roboterchirurgie, werden die wichtigsten Forschungsaufbauten präsentiert. Anschließend werden die während dieser Arbeit entstandenen Forschungsplattformen für robotergestützte minimalinvasive Chirurgie Endo[PA]R und ARAMIS² vorgestellt.

Den Schwerpunkt des Kapitels 3 “Kalibrierung” bildet die *Hand-Auge-Kalibrierung* des Gesamtsystems bestehend aus mehreren Roboterarmen und dem Stereoendoskop. Ein neuartiger 3D-Kalibrierkörper wird vorgestellt, der sich sowohl für die Kalibrierung des Stereoendoskops, als auch des Gesamtsystems eignet. Er bietet außerdem Vorteile, die ihn für den Einsatz bei wechselnden Bedingungen und im sterilen Umfeld eines Operationssaals qualifizieren.

Der Abschnitt 4 “Bildvorverarbeitung” widmet sich einigen Schwierigkeiten der endoskopischen Bildverarbeitung, und stellt Lösungsansätze vor. Ziel ist es, das endoskopische Stereobildpaar für die nachfolgende Segmentierung und 3D-Rekonstruktion aufzubereiten, bzw. den operierenden Chirurgen zu entlasten. Zum einen werden verschiedene Verfahren zur Entfernung von Spiegelreflexionen vorgestellt und evaluiert, eine typische Erschwernis im chirurgischen Umfeld, bedingt durch das helle Licht und die nassen Organoberflächen. Zum anderen wird die Problematik des Bildrauschens analoger Endoskopiesysteme behandelt.

Im Abschnitt 5 “Fadenerkennung” wird der erste Schritt zur Rekonstruktion des Operationssitus behandelt. Eine Methode zur robusten Erkennung und Lokalisation von unterschiedlichem chirurgischem Nahtmaterial wird vorgestellt. Der mehrstufige Ansatz, dessen Schwerpunkt ein subpixelgenauer Kantenfilter darstellt, wertet sowohl Farbinformation als auch geometrische Merkmale aus, und liefert eine subpixelgenaue Repräsentation des Fadenverlaufs. Erst die Verfügbarkeit präziser Daten im 2D-Bildraum ermöglicht eine exakte Rekonstruktion im 3D-Raum.

Anschließend befasst sich Kapitel 6 “3D-Rekonstruktion des Situs” mit der Aufgabe der Modellbildung. Ziel ist es, während eines Eingriffs ein genaues Modell von einem Ausschnitt des Operationsgebietes zu erstellen und gegebenenfalls zu aktualisieren. Sowohl Nahtmaterial als auch Herzoberfläche werden rekonstruiert und modelliert. Die Daten schaffen zum einen die Grundlage für autonome Techniken, und ermöglichen zum anderen die Implementierung sicherheitsrelevanter Funktionen.

Der Abschnitt 7 “Zusammenfassung und Ausblick” schließt die Arbeit mit einem Resümee der erreichten Ziele und der verwendeten Methoden ab. Außerdem werden hier mögliche Verbesserungen vorgestellt und weitere potentielle Forschungsschwerpunkte in dieser Domäne skizziert.

Kapitel 2

Chirurgie und Robotik

Zusammenfassung Die Entwicklung fortgeschrittener chirurgischer Eingriffe hat in den letzten Jahrzehnten die Lebenserwartung, und die Lebensqualität nach Operationen beachtlich erhöht. Doch die Anfänge waren wegen den großen Schnitten von erheblichem Blutverlust und Infektionen begleitet, wobei viele Patienten als Folge des Eingriffs starben. Eine bedeutende Verbesserung, außer technischer und pharmazeutischer Weiterentwicklungen im Operationssaal, brachte die Einführung der minimalinvasiven Chirurgie um 1980. Im Gegensatz zu konventioneller offener Chirurgie wird das Operationsgebiet durch kleine Einschnitte mit Hilfe spezieller Instrumente und Endoskope erreicht. Die Vorteile gegenüber konventionellen Eingriffen sind offensichtlich: weniger Trauma und Schmerz wegen kleineren Schnitten, kürzere Rekonvaleszenz, und nicht zuletzt kosmetische Betrachtungen. Doch trotz aller Vorteile brachte die neue Technik entgegen der Erwartungen nicht den erhofften Durchbruch auf dem Gebiet der Chirurgie. Der Grund dafür ist, dass Vorteile auf Patientenseite durch Nachteile auf der Arztseite aufgewogen werden. Die eingeschränkte Sicht erschwert die Orientierung und das Auffinden anatomischer Strukturen. Die Instrumente werden um sogenannte *Trokarpunkte* am Abdomen oder Brustkorb des Patienten bewegt, wodurch die Freiheitsgrade des Operateurs reduziert werden, und umgekehrte Handbewegungen notwendig sind. Die langen Instrumente verstärken den Tremor und verschlechtern deutlich die haptische Rückkopplung.

Um einige dieser Probleme zu beseitigen, wurden Technologien aus der Robotik in die minimalinvasive Chirurgie integriert. Dabei werden spezielle Roboterarme als Instrumententräger eingesetzt, welche vom Chirurgen ferngesteuert werden. Auch die Sicht wurde mittels Stereoendoskopie verbessert, sodass beim Operieren ein realistischer räumlicher Eindruck entsteht. Kommerziell erhältliche und für den klinischen Einsatz zugelassene robotergestützte minimalinvasive Chirurgiesysteme sind Teleoperationssysteme, welche ausschließlich in striktem ferngesteuertem Modus betrieben werden. Das heißt, die einzige Informationsquelle für den operierenden Chirurgen ist der visuelle Kanal, und jede Instrumentenbewegung wird vom Chirurgen gesteuert. Diese Art vom teleoperativen Modus nennt man auch *visual servoing*. Die Verwen-

dung von Roboterarmen beseitigt mit Hilfe konstruktions- und regelungstechnischer Maßnahmen die Notwendigkeit umgekehrter Handbewegungen, und mindert den Handtremor erheblich. Die Armbewegungen des Chirurgen können außerdem nahezu beliebig fein skaliert werden. Nichtsdestotrotz bleibt die Problematik fehlender Krafrückkopplung und taktiler Eindrücke bestehen. Obwohl die Handhabung robotergestützter Chirurgesysteme im Vergleich mit traditionellen laparoskopischen Eingriffen einfacher wurde, ist der durchschnittliche Zeitbedarf für Operationen nicht nennenswert gesunken. Das ist zum einen auf den gestiegenen Planungs- und Vorbereitungsauwand des Eingriffs zurückzuführen. Zum anderen erschwert die fehlende Gesamtübersicht die Orientierung, und viele Standardaufgaben, wie beispielsweise Knotenlegen, gestalten sich im minimalinvasiven Fall aufwendiger.

Löst man sich jedoch von der Vorstellung eines ausschließlich teleoperierten Systems, und betrachtet man die Arme als autonomiefähige Roboter, so ergeben sich zahlreiche Verbesserungsmöglichkeiten hinsichtlich Ergonomie, Qualität und Effizienz. Eine erste naheliegende Erweiterung ist, die Systeme mit (teil-)autonomen Assistenzarmen auszustatten, zumal es auch bei konventionellen offenen Eingriffen fast immer einer menschlichen Assistenz bedarf. Einfache Positionier- und Haltetätigkeiten können vom System selbständig durchgeführt werden. Die dafür erforderliche Integration von Kraftsensorik- und Regelung ist seit Jahrzehnten Stand der Technik in der Industrierobotik und unter theoretischem Gesichtspunkt direkt anwendbar. Häufig wiederkehrende, aber zeitraubende Aufgaben, wie das schon erwähnte Knotenlegen, können ebenfalls automatisiert werden, was zu einer Verkürzung der Operationsdauer führt. Das dazu notwendige Modell des Operationssitus, um beispielsweise den Faden greifen und den Knoten festziehen zu können, liefert die Steuerung anhand der Daten des Stereoendoskops. Die eingeschränkte Sicht während eines Eingriffs kann dadurch kompensiert werden, dass mit Hilfe des Stereoendoskops zu Beginn ein 3D-Übersichtsmodell des Operationsgebietes erstellt, und um den jeweils aktuell beobachteten Ausschnitt eingeblendet wird. Die kinematische Einbindung des Kamerasystems in das Robotersystem ermöglicht eine kontinuierliche Anpassung des Modells bei nachfolgenden Kamerafahrten.

Die vorliegende Arbeit zeigt einige der oben beschriebenen Lösungsansätze, wobei der Schwerpunkt auf Bildverarbeitung und Modellierung liegt. Die Techniken wurden in einen realistischen Experimentalaufbau bestehend aus drei Instrumentenarmen und einem Endoskoparm integriert. Die verwendeten Instrumente und das Endoskop sind für den klinischen Einsatz zugelassen. Das System befindet sich zum Zeitpunkt der Erstellung dieser Arbeit am Deutschen Herzzentrum München, und wird dort im Großtierversuch unter klinischen Bedingungen von Chirurgen evaluiert.

2.1 Minimalinvasive Chirurgie

Die minimalinvasive Chirurgie (MIC) wurde Anfang der 1980er Jahre aus dem Wunsch heraus entwickelt, nach operativen Eingriffen eine schnellere Genesung mit verminderten Beschwerden zu gewährleisten. Sie steht im allgemeinen für Interventionen mit kleinstem Trauma, also mit kleinster Verletzung von Haut und Weichteilen. Zunächst nur auf Eingriffe in der Bauchhöhle begrenzt, wurden die Vorteile schnell erkannt und man begann auch auf anderen Gebieten, hauptsächlich im Bereich des Brustkorbs, mit der Entwicklung minimalinvasiver Operationsverfahren. Während minimalinvasiver Eingriffe werden die klinischen Ziele ausschließlich durch mehrere kleine Öffnungen an der Körperoberfläche erreicht. Der Chirurg operiert mit speziellen, langen Instrumenten, wobei er keinen direkten Zugang zum Operationsgebiet hat. Die nötige Sicht liefert ein Endoskop, welches meistens von einer Assistenz manuell nachgeführt wird. Sowohl Instrumente als auch Endoskop werden durch sog. Trokare¹ ins Körperinnere geleitet und um diese bewegt. Abbildung 2.1 zeigt schematisch Trokare, Instrumente und Endoskop bei einem minimalinvasiven Eingriff im Brustkorb. Bedingt durch die Art des Eingriffs kann eine zusätzliche Öffnung für die Aufdehnung des Brust- oder Bauchraumes mit CO_2 notwendig sein.

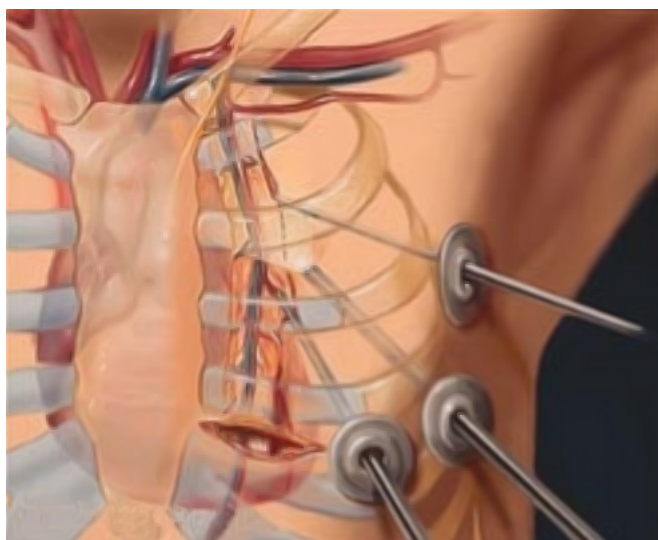
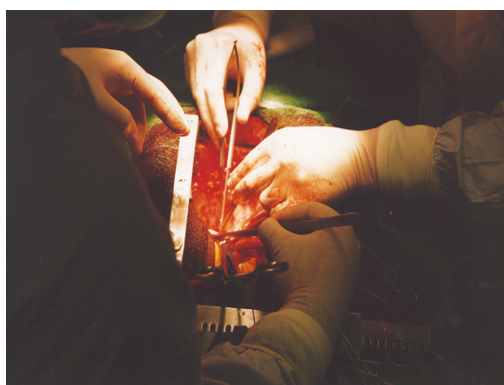


Abbildung 2.1: MIC (schematische Darstellung, Quelle unbekannt)

Bei manchen Eingriffen haben sich minimalinvasive Techniken gegenüber konventionellen, offenen Varianten durchgesetzt, und gelten nun als Standardverfahren. Die anfängliche Einschränkung auf Brust- oder Bauchraum schwindet immer mehr, doch entgegen der Prognosen, hat sich die minimalinvasive Chirurgie im allgemeinen nicht in dem erhofften Maße etabliert.

¹Trokare sind chirurgische Instrumente, welche eine Öffnung im Brust- oder Bauchraum schaffen und während eines Eingriffs offen halten

Offensichtliche Verbesserungen auf Patientenseite sind weniger Trauma, kürzere Rekonvaleszenz, und nicht zuletzt kosmetische Vorteile. Der verkürzte Krankenhausaufenthalt, oder sogar die Möglichkeit ambulanter Behandlungen stellen außerdem einen erheblichen Kostenfaktor dar. Abbildung 2.2 zeigt den Unterschied zwischen offenem und minimalinvasivem Eingriff unter dem Gesichtspunkt des verursachten Gewebstraumas. Auch bei unklarer Befundlage bietet die schonende Inspektion mittels einer Laparoskopie deutliche Vorteile, und kann unnötige Eingriffe vermeiden.



(a) Offener Eingriff



(b) Ports und Trokare

Abbildung 2.2: Äußere Unterschiede zwischen offenem und minimalinvasivem Eingriff (Quelle: Deutsches Herzzentrum München)

Doch die neue Technik brachte auch Nachteile mit sich, wovon zunächst nur der Operateur betroffen zu sein scheint. Die schwierige Handhabung der langen Instrumente erfordert spezielles Training, wobei bereits erlangte Fertigkeiten in konventioneller Chirurgie nicht übertragbar sind. Der Operateur hat im Körperinneren weniger Freiheitsgrade zur Verfügung, sodass angepasste Techniken entwickelt werden mussten. Bedingt durch den Trokar sind stets umgekehrte Handbewegungen notwendig, und weite Bewegungen behindern ergonomisches Arbeiten. Die Hebelwirkung der langen Schäfte minimalinvasiver Instrumente verstärken den Handtremor und der Reibwiderstand am Trokar verschlechtern das haptische Feedback; taktiler Feedback ist gänzlich unmöglich. Obwohl sich zunehmend stereoskopische Sichtsysteme durchsetzen, die eine räumliche Darstellung des Operationsgebietes ermöglichen, treten oft Orientierungsprobleme wegen der eingeschränkten Sicht auf. Ein weiterer Nachteil minimalinvasiver Eingriffe ist die verzögerte Zugriffsmöglichkeit bei einer bedrohlichen Komplikation, wie beispielsweise einer starken Blutung im Operationsfeld.

Inspiziert durch die Mechanik minimalinvasiver Instrumente begannen Forschungsgruppen Ende der 1980er, Anfang der 1990er Jahre mit der Anpassung und Integration bestimmter Techniken aus der Robotik in die Chirurgie. Die Hoffnung war, einige der hier aufgezählten Probleme zu beseitigen, bzw. zu mindern.

2.2 Chirurgeroboter

Bedingt durch die sehr unterschiedliche Art der Anforderungen in den verschiedenen Bereichen der Chirurgie, ist es nicht sinnvoll von einem “generischen” Chirurgeroboter zu sprechen. Doch es zeichnen sich drei Haupteinsatzgebiete ab, in denen immer mehr Roboterchirurgiesysteme kommerziell erhältlich und für den klinischen Einsatz zugelassen werden:

1. **(Kiefer-)Orthopädie** Typische Aufgaben sind in diesem Bereich das Fräsen und Bohren von Knochen mit hoher Präzision, wodurch beträchtliche Kräfte und störende Vibrationen auftreten können. Hierfür werden oft adäquat modifizierte Industrieroboter eingesetzt. Ein hoher Grad an Autonomie kann dank präoperativer Planung erreicht werden.
2. **Neurochirurgie** Wichtigste Anforderung in diesem Feld ist hohe Präzision in einem sehr beschränkten Arbeitsraum, wobei die aufzuwendenden Kräfte vernachlässigbar sind. Die größte Herausforderung stellt die genaue Navigation und Positionierung dar. Meistens nur im Zusammenhang mit präoperativer Planung anhand bildgebender Verfahren (CT, MRT) möglich.
3. **Abdomen- und Thoraxchirurgie** Der Bereich mit dem größten Arbeitsraum stellt besondere Anforderungen an Roboterarme und Endeffektoren. Die Behandlung deformierbarer Organe in einem online-teleoperierten Modus verlangt nach vielen Freiheitsgraden und haptischem Feedback.

Robotergestützte chirurgische Eingriffe entfalten erst im Zusammenhang mit leistungsfähigen bildgebenden Verfahren ihr Potenzial. In allen drei Bereichen ist eine sorgfältige präoperative Planung anhand von Computertomographie (CT) oder Magnetresonanztomographie (MRT) unerlässlich. Erwähnenswert ist noch ein Verfahren aus der Dentalrestauration, bei der mit Hilfe von stereoskopischen Kameraaufnahmen das notwendige 3D-Modell zur Anfertigung des Implantats erstellt wird.

Abbildung 2.3 zeigt die zwei prominentesten Vertreter aus der ersten Klasse. Das *ROBODOC*TM-System aus Abbildung 2.3(a) der Firma Integrated Surgical Systems [15] kam hauptsächlich bei der Implantation von Hüftgelenkprothesen zum Einsatz [16]. Die notwendigen Fräsungen im Knochen wurden von einem modifizierten Industrieroboter rechnergestützt und autonom durchgeführt. Fehler in der präoperativen Planung konnten somit während des Eingriffs nur schwer erkannt und behoben werden. Außerdem erfordert die *ROBODOC*-Methode zum Erreichen des Operationsgebiets eine größere Öffnung im Muskel als die herkömmliche Methode. Diese Umstände wurden *ROBODOC* zum Verhängnis, sodass aktuell die Benutzung bis auf weiteres eingestellt wurde. Zahlreiche Geschädigte haben Klagen gegen den Hersteller eingereicht und fordern die Zahlung von Schmerzensgeld.

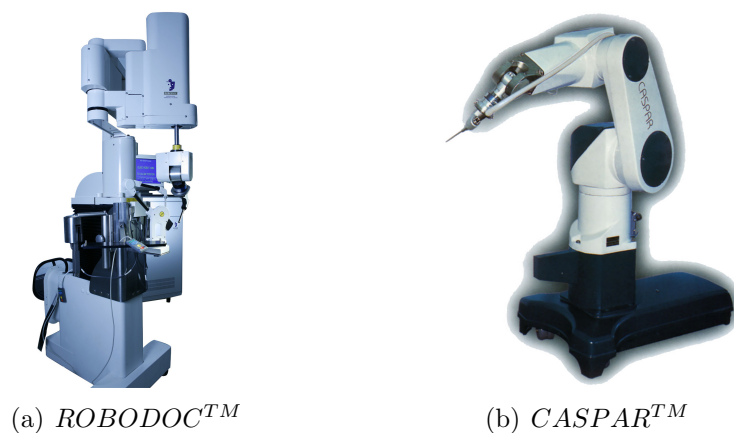


Abbildung 2.3: Orthopädische Systeme

Als direkter Konkurrent ist noch die Firma Universal Robotic Systems Ortho GmbH [17] mit ihrem *CASPARTM*-System zu erwähnen. Basierend auf einem *Stäubli RX90* Roboter (Abb. 2.3(b)) wurde das System hauptsächlich für Knie- und Hüftgelenkoperationen eingesetzt, aber auch die Behandlung von zervikaler Spondylose [18] war damit möglich. Doch auch *CASPAR* hat den durch *ROBODOC* verursachten Imageschaden nicht verkraftet, sodass zum aktuellen Zeitpunkt kein orthopädisches Chirurgesystem weltweit im Einsatz ist.

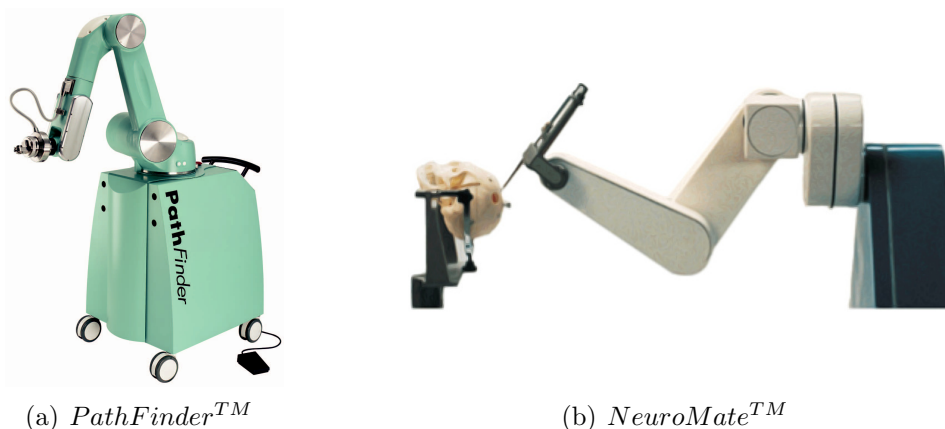


Abbildung 2.4: Neurochirurgische Systeme

Der erfolgreichste Vertreter eines robotergestützten Systems für Neurochirurgie ist *PathFinderTM* von Prosurgics [19] (Abb. 2.4(a)). Intracraniale Läsionen werden anhand von hochauflösenden CT- oder MRT-Scans lokalisiert und der Eingriff geplant. Es wird kein klassischer stereotaktischer Rahmen mehr benötigt, sondern am Schädel des Patienten befestigte optische Referenzmarkierungen, welche zum Scan registriert werden können [20]. Das Verfahren lässt einerseits viel kürzere OP-Zeiten zu, und ermöglicht andererseits eine gewisse Bewegungsfreiheit während der

OP. Ebenfalls von Integrated Surgical Systems wurde *NeuroMateTM* (Abb. 2.4(b)) entwickelt, welches zunächst auf die Verwendung eines stereotaktischen Rahmens angewiesen war. Zum Umfang des Systems gehört eine Workstation zur präoperativen Planung. In der zweiten Generation wurde das System dahingehend verbessert, dass ein für den Patienten schonender rahmenloser Betrieb möglich ist [21]. In einer vergleichenden Studie [22] wurde auch die Positioniergenauigkeit von *NeuroMate* evaluiert, und sie ist vergleichbar mit der herkömmlicher Systeme mit stereotaktischem Rahmen.

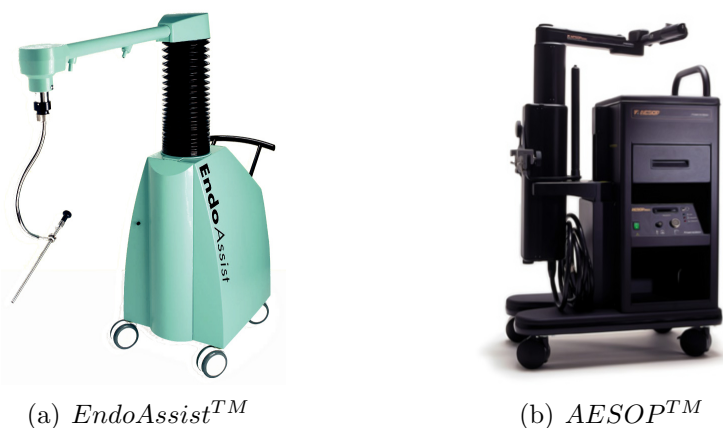


Abbildung 2.5: Systeme zur Positionierung des Endoskops

Auch ein weiteres Produkt von Prosurgics sollte hier erwähnt werden, und zwar ihr *EndoAssistTM*-System (Abb. 2.5(a)). Es handelt sich dabei um einen robotischen Manipulator für die Positionierung des Endoskops in der Abdominal- und Thorakalchirurgie. Die Besonderheit des Systems ist, dass es mittels Kopfbewegungen des Chirurgen gesteuert werden kann [23, 24]. Die Firma Computer Motion [25] bietet mit *AESOPTM* (Abb. 2.5(b)) einen sprachgesteuerten Konkurrenzprodukt an, der dank seiner sieben Freiheitsgrade [26] flexibler bei der Handhabung ist. Die Arbeit mit dem System beeinflusst nicht negativ die Operationsdauer [27], oft ermöglicht es sogar eine schnellere und effizientere Arbeitsweise als klassische Systeme mit Hand- und Fußsteuerung [28]. Der Vergleich beider Systeme ergab einen leichten Vorteil zugunsten von *EndoAssist*, sporadische Fehler in der Spracherkennung des *AESOP*-Systems führen nämlich zu leichten Verzögerungen [29, 30].

Obwohl es in der Kieferorthopädie keine Möglichkeit zum Einsatz eines Robotersystems im klassischen Sinne gibt, ist die ursprünglich von Siemens entwickelte und derzeit von Sirona [31] vertriebene *CEREC*-Technik (Chairside Economical Restorations of Esthetic Ceramic) vollständigshalber hier zu erwähnen, zumal sie prinzipiell große Ähnlichkeiten mit orthopädischen Systemen aufweist. Dabei werden zunächst anhand stereoskopischer Kameraaufnahmen dreidimensionale Modelle der zu restaurierenden Kieferpartien erstellt. Anschließend werden die Implantate mit hoher Genauigkeit aus einem Keramikblock ausgefräst. Der Hauptvorteil des

Systems ist, dass die Behandlung in einer Sitzung erfolgen kann. Abbildung 2.6(a) zeigt Behandlungs-, Planungs- und Fräseeinheit. Vorstellbar wäre jedoch auch ein robotergestütztes System, welches in der Schädel- und Gesichtschirurgie zum Einsatz käme. Beispiele hierfür sind in [32] und [33] bzw. das *RobaCKa*-System [34] aus Abbildung 2.6(b).



Abbildung 2.6: Kiefer-, Schädel- und Gesichtschirurgie

Bei den vorhin vorgestellten Systemen stand die Erhöhung der Präzision eines Eingriffs im Vordergrund, das Trauma kann oft ohnehin nicht reduziert werden. In der Abdominal- und Thorakalchirurgie, die auch den Schwerpunkt dieser Arbeit bilden, richtet sich das Hauptaugenmerk primär auf die Reduzierung des Knochen- und Gewebstraumas. Erst solche Systeme können als “minimalinvasiv” bezeichnet werden.

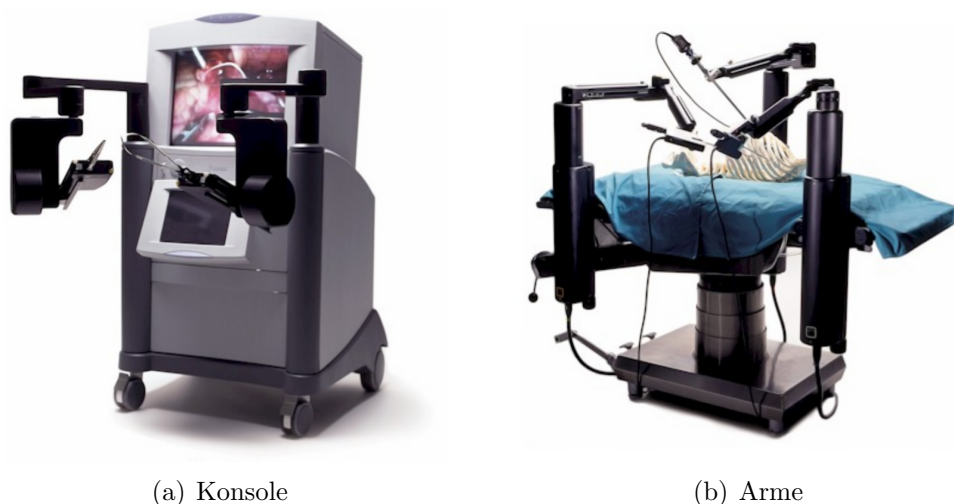


Abbildung 2.7: Das *ZEUS™*-System (©Computer Motion Inc.)

Die zwei ausgereiftesten, kommerziell erhältlichen und klinisch zugelassenen Systeme dieser Domäne sind *ZEUS™* (Abb. 2.7) von Computer Motion [25] und *da Vinci™* (Abb. 2.8) von Intuitive Surgical Inc. [35]. Die Firma Computer Motion gibt es zum Zeitpunkt der Erstellung dieser Arbeit nicht mehr, wodurch das

das *Vinci*-System zur einzigen Referenz eines abdominal-thorakalen minimalinvasiven Chirurgesystems wurde. Anhand dessen soll nun der prinzipielle Aufbau solcher Systeme erläutert werden.

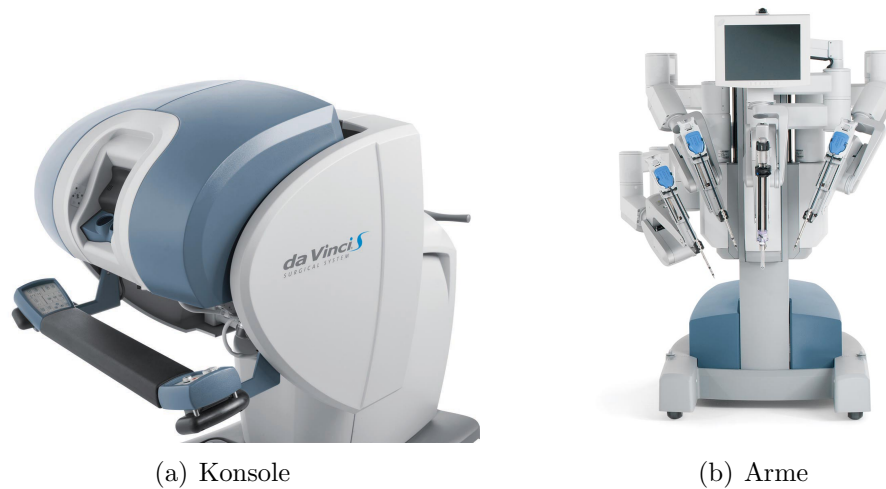


Abbildung 2.8: Das *da Vinci*TM-System (©Intuitive Surgical Inc.)

2.3 Robotergestützte minimalinvasive Chirurgie

Die Grundidee robotergestützter minimalinvasiver Chirurgie (RMIC) ist die Verwendung von Roboterarmen als Instrumententräger, welche vom Chirurgen mittels geeigneter Eingabegeräte ferngesteuert werden. Die Instrumente weisen große Ähnlichkeiten mit klassischen laparoskopischen Instrumenten auf, haben jedoch meistens mehr Freiheitsgrade. Somit sind robotergestützte Chirurgesysteme typische Master-Slave-Architekturen, und bestehen aus zwei Hauptkomponenten: aus einer sogenannten *Konsole (Master)* und einem *Operator (Slave)*. Die Konsole dient dabei als Ein- und Ausgabeeinheit für den Chirurgen, der patientenseitige Operationen führt die an der Konsole ausgelösten Vorgänge durch, der Roboter operiert also nicht autonom. Die Konsole hat demnach zwei Funktionen: einerseits dient sie als Eingabeinterface, und andererseits liefert sie visuellen Feedback vom Ort des operativen Eingriffs. Die Bewegungen des Chirurgen werden mit Hilfe von zwei Eingabemechanismen registriert und zum Operator übertragen. Der Operator ist im Prinzip ein in der Nähe des Operationstisches positionierter dreiarmer Roboter, wobei zwei der Arme die eigentliche Operationsaufgabe ausführen, der dritte Arm ist für die Positionierung einer endoskopischen Kamera zuständig. Die Kamera liefert ein dreidimensionales Bild, welches an die Konsole übertragen, und dem operierenden Chirurgen präsentiert wird. Die Instrumente können während eines Eingriffs ausgetauscht, und die Kamera repositioniert werden. Abbildung 2.9 zeigt schematisch

den typischen Aufbau eines Operationssaales ausgerüstet für robotergestützte minimalinvasive Chirurgie.

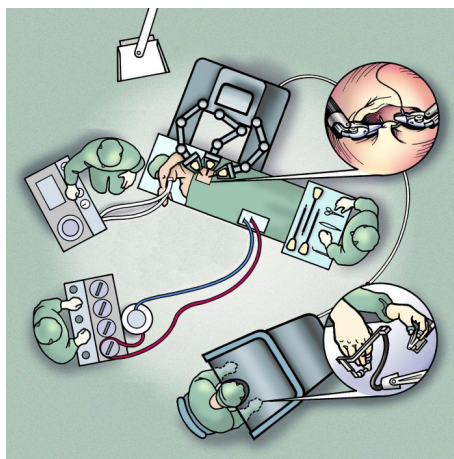


Abbildung 2.9: Schematische Darstellung RMIC (©Intuitive Surgical)

Der folgende Abschnitt gibt einen kurzen technologischen Überblick eines aus heutiger Sicht optimalen robotergestützten minimalinvasiven Chirurgesystems, betreffend sowohl Systemarchitektur, als auch Konfiguration. Wir beschränken uns dabei auf die Architektur klassischer Telepräsenzsysteme (online Mensch–Maschine–Interaktion), wie es bei den *da Vinci* und *ZEUS* Systemen der Fall ist. Wir erkennen drei Hauptkomponenten:

1. **Master** Dieses Subsystem (auch Eingabekonsolle genannt, siehe Abb. 2.7(a) und 2.8(a)) ist der Arbeitsplatz des Chirurgen. Hier werden die verschiedenen Modalitäten der visuellen und kinästhetischen Wahrnehmung dargestellt, welche auf der Slave–Seite von entsprechender Hardware generiert werden. Die Aktionen des Operators werden mit Hilfe spezieller (meist exoskeletaler) Eingabegeräte (siehe Abb. 2.10) erfasst. Sowohl die Skalierung der Bewegungen, als auch die Filterung des Handtremors sollten zum Leistungsumfang gehören. Die transparente Übertragung der Freiheitsgrade der menschlichen Hand ins Körperinnere ist technisch realisierbar, und sollte sowohl von Eingabemechanismen, als auch von den chirurgischen Instrumenten unterstützt werden. Eine hochauflösende Stereosicht ist unverzichtbar, denn eine fehlende Tiefeninformation ist nur schwer kompensierbar.
2. **Slave** Das auf der Patientenseite befindliche Subsystem (auch Operator oder Aktuator genannt, siehe Abb. 2.7(b) und 2.8(b)) besteht aus zwei Hauptkomponenten: den Roboterarmen und den minimalinvasiven chirurgischen Instrumenten (siehe Abb. 2.10). Die Trennung von Armen und Instrumenten ist durch die Anforderung motiviert, dass die Instrumente während eines Eingriffs

austauschbar sein müssen. Die Instrumente sollten zum einen die Freiheitsgrade der menschlichen Hand (man vergleiche die unterschiedliche Anzahl der Freiheitsgrade der *ZEUS*- und *da Vinci*-Instrumente) abbilden können, zum anderen mit haptischer Sensorik ausgestattet sein. Die Kinematik der Roboterarme muss in der Lage sein, die durch den Trokar bedingten Einschränkungen zu handhaben, ohne die Funktionalität einzuschränken.



Abbildung 2.10: Eingabemechanismen und Instrumente

3. **Kommunikationskanal** Mehrere Kanäle hoher Bandbreite sind für den Austausch von Haptik-, Positions- und Videodaten zwischen Master und Slave notwendig. Anforderungen wie garantierte Bandbreite, keine (oder sehr geringe) Verzögerung müssen für einen sicheren Betrieb gewährleistet sein. Das Kommunikationssystem muss flexibel genug sein, um die Verbindung mehrerer Master zum selben Slave, oder dynamische Master-Slave Neuordnung zu unterstützen.

Zusätzlich zu dem oben aufgeführten Leistungsumfang gibt es noch einige fortgeschrittene Techniken, welche an verschiedenen Forschungseinrichtungen erprobt werden, und potenziell in naher Zukunft eingesetzt werden könnten:

- **Automatische Kameranachführung** Aktuell verfügbare RMIC-Systeme überlassen die Kontrolle der Kamera dem Chirurgen. Immer wenn die Kamera neu positioniert werden muss, schaltet der Chirurg von der Steuerung der Instrumente auf die Steuerung der Kamera um, was zum einen zeitaufwändig ist, und zum anderen Gefahren birgt. Die Kenntnis der Lage der Instrumente, und infolgedessen des Arbeitsbereichs, ermöglicht einem Robotersystem diesen Bereich optimal zu erfassen.
- **Partielle Autonomie** Assistenz ist während konventionellen offenen Eingriffen nicht wegzudenken. Mögliche (partiell) autonome Aufgaben in minimalinvasiven Chirurgieszenarien wären: temporäres Halten (“Parken”) der Nadel

odes des Fadens, Knotenlegen und Straffung des Gewebes um den Schneidvorgang zu erleichtern.

- **Bewegungskompensation** Hauptsächlich auf dem Gebiet der abdominalen und thorakalen Chirurgie ist eine erhebliche Bewegung des Operationsgebietes wegen Herzschlag und Atmung des Patienten zu beobachten. Theoretisch möglich ist die synchrone, auf die Bewegungen des Chirurgen überlagerte Nachbewegung von Instrumenten und Kamera, um einen virtuell stehenden Situs zu präsentieren. Die niederfrequente Atembewegung könnte relativ leicht erfasst und kompensiert werden. Doch die Kompensation der meist unregelmäßigen (Rest-)Herzbewegung während eines Eingriffs gestaltet sich hingegen schwieriger.

Ursprünglich zur Entlastung des Chirurgen und zur Erhöhung der Präzision eingeführt brachte die neue robotergestützte Technik viele Vorteile mit sich. Doch erst im klinischen Betrieb stellten sich auch einige Nachteile heraus. Die folgenden Abschnitte stellen eine Zusammenfassung beider Seiten dar, und zeigen mögliche Lösungsvorschläge für noch bestehende Probleme.

2.3.1 Vorteile

Durch den Verzicht auf das komplette Öffnen der Bauchdecke bzw. des Brustkorbs wird das aus der Operation resultierende Trauma stark verringert, und es ergeben sich auf der Patientenseite die selben Vorteile, wie bei klassischen minimalinvasiven Eingriffen. Die Integration von Technologien aus der Robotik haben in erster Linie die Arbeit des Chirurgen erleichtert. Durch die Entwicklung angepasster Arm-Kinematiken und entsprechender Steuerung kann das durch den Trokar verursachte Problem umgekehrter Handbewegungen gelöst werden. Die Instrumente können nun intuitiv gehandhabt werden, alle Freiheitsgrade der menschlichen Hand sind auch ins Körperinnere transferierbar. Dank der beliebig feinen Bewegungsskalierung und der Möglichkeit den Handtremor regelungstechnisch zu eliminieren ist eine höhere Präzision erreichbar.

Die Verwendung stereoskopischer Optiken zusammen mit dem Einsatz entsprechender Wiedergabehardware ermöglichen eine realitätsnahe Darstellung des Operationsgebietes. Die dadurch hinzugekommene Tiefeninformation entlastet den Operateur und erhöht die Genauigkeit. Die kinematische Integration des Endoskops in das Robotersystem gestattet die Augmentierung mit Daten, die aus verschiedenen bildgebenden Verfahren, wie beispielsweise Angiographie, gewonnen wurden. Moderne Endoskopiesysteme bieten außerdem die Möglichkeit zur digitalen Bildaufbereitung, wie Vergrößerung, Kontrasterhöhung oder Hervorhebung bestimmter Farben und Strukturen. Die Aufzeichnung kompletter Eingriffe zu Lehr- und Dokumentationszwecken ist ebenfalls möglich.

Präoperative Planung anhand bildgebender Verfahren ist im robotergestützten Fall praktisch unverzichtbar. Um das Operationsgebiet optimal erreichen zu können, muss die Lage zur Platzierung der Trokare sorgfältig ermittelt werden, denn nachträgliche Änderungen sind nur schwer möglich. Dies trifft wegen der Einschränkung auf den Interkostalraum insbesondere auf die Thoraxchirurgie zu. Durch den Einsatz von speziellen Markern auf der Körperoberfläche des Patienten kann die Registrierung und Planung nahezu automatisch erfolgen.

2.3.2 Nachteile

Die fehlende taktile und kinästhetische Wahrnehmung erschwert die Arbeit des Chirurgen bei jeglicher Art minimalinvasiver Eingriffe, eine Roboterunterstützung bringt zunächst sogar Nachteile. Krankhafte Veränderungen, wie beispielsweise Kalkablagerungen in Blutgefäßen, können visuell oft nicht erkannt werden. Die Eigenschaften von Gewebe, insbesondere in Wechselwirkung mit dem Nahtmaterial, kann ohne Krafrückkopplung ebenfalls nur schwer geschätzt werden. Abhilfe schaffen hier sensorbestückte chirurgische Instrumente in Verbindung mit entsprechenden haptischen Mensch-Maschine-Schnittstellen. Das Endoskop bietet nur eine beschränkte Sicht des Operationsgebietes, sodass Orientierungsprobleme die Arbeit des Chirurgen erschweren. Eine geeignete Abhilfe kann das Robotersystem in Form eines Übersichtsmodells schaffen, welches um das real beobachtete Szene eingeblendet wird. Die Synchronisierung zwischen Übersicht und Situs bei nachfolgenden Kamerafahrten kann automatisch erfolgen.

Robotergestützte Eingriffe zeichnen sich typischerweise durch eine verlängerte Operationsdauer aus. Dieser Umstand resultiert zum einen aus dem höheren Vorbereitungsaufwand, wodurch sie sich meistens nur für geplante Interventionen eignen, Notoperationen sind kaum möglich. Zum anderen beanspruchen laparoskopische Ausführungen bestimmter Techniken, wie Nähen und Knoten, mehr Zeit für ihre Ausführung. Verschiedene organisatorische Aspekte, wie die hohen Anschaffungskosten und das notwendige Training des Personals, verhindern ebenfalls die Verbreitung solcher Systeme. Forschungseinrichtungen versuchen hauptsächlich die technischen Probleme zu lösen, und konzentrieren sich auf die Entwicklung chirurgischer Instrumente (oft mit Kraft/Moment-Sensorik), haptischer Eingabegeräte und angepasster Armkonstruktionen. Großer Handlungsbedarf besteht bei der Entwicklung einheitlicher Methoden und Prozesse, die den gesamten Ablauf eines robotergestützten Eingriffs abdecken. Der nächste Abschnitt bietet einen Überblick verschiedener Forschungssysteme, wir beschränken uns jedoch auf veröffentlichte Systeme, für die es funktionierende Prototypen gibt.

2.4 Stand der Forschung

Das ARTEMIS-System Entwickelt am Forschungszentrum Karlsruhe (FZK) war das ARTEMIS-System (Advanced Robotic and Telemanipulator System for Minimal Invasive Surgery) der erste Aufbau für Roboterchirurgie in Deutschland, und in dieser Komplexität einer der ersten weltweit. Obwohl mittlerweile nicht mehr weiterentwickelt, ist er auch nach heutigen Maßstäben auf sehr hohem technischen Niveau. Beim ARTEMIS wurde streng auf die logische Trennung der einzelnen Subsysteme und ihre Kommunikation geachtet, und wir beobachten folgende Komponenten: Mensch-Maschine-Schnittstelle, Ausführungseinheit und Steuerungssystem.



Abbildung 2.11: ARTEMIS am Forschungszentrum Karlsruhe

Die Mensch-Maschine-Schnittstelle (Abb. 2.11(b)) besteht aus mehreren Einheiten: zwei haptische Manipulatoren, graphische Benutzerschnittstelle, 3D-Sichtsystem, Spracheingabe (für die Steuerung des Laparoscops), Pedale und ein Trackball. Die Ausführungseinheit (Abb. 2.11(c)) besteht aus dem Instrumentenführungssystem TISKA, welches für die Positionierung der chirurgischen Instrumente zuständig ist, und dem computergesteuerten Kameraführungssystem ROBOX. Die multifunktionalen Endeffektoren sind am distalen Ende flexibel und verfügen über 6 Freiheitsgrade. Die Kenntnis der relativen Lage zwischen ROBOX und den TISKA Robotern ermöglicht eine automatische Kameranachführung. Jedes Eingabemaster kann jeden Slave steuern, die einzelnen Kinematiken müssen nicht identisch sein. Verschiedene Steuermodi (Weltkoordinaten, Bildkoordinaten) und Funktionen (Skalierung, Indexing) sind möglich. Der ethernetbasierte Kommunikationkanal MONSUN (Manipulator Control System Utilizing Network Technology) ermöglicht auch Teleoperationen über weite Entfernungen. Die Steuerung sorgt für die Einhaltung der Trokarkinematik und enthält zahlreiche Sicherheitsvorkehrungen. Die 3D-Simulationssoftware KISMET [36] ist ebenfalls Teil des Systems, das einzige Manko ist die fehlende Krafterückkopplung. Siehe [37], [38] und [39] für eine ausführliche Beschreibung der Systemarchitektur.

Das UCB/UCSF-System In einem Kooperationsprojekt zwischen dem Robotics and Intelligent Machines Laboratory der University of California, Berkeley

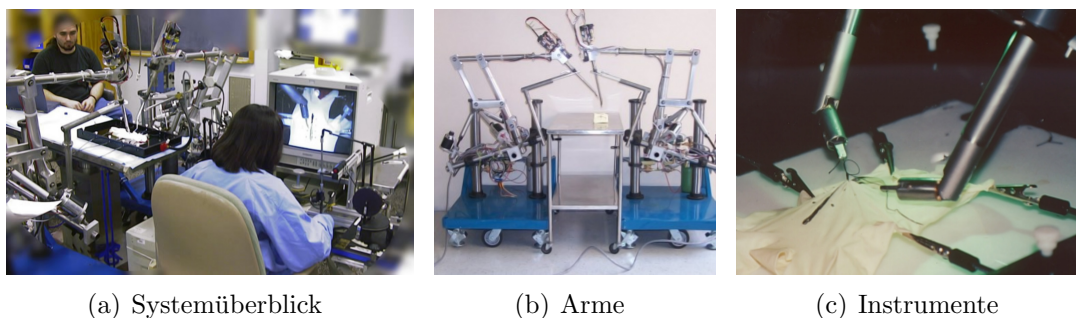


Abbildung 2.12: “Robotic Telesurgical Workstation for Laparoscopy”

(UCB) und dem Department of Surgery der University of California, San Francisco (UCSF) wurde ein robotergestützter, laparoskopischer Telechirurgieaufbau entwickelt (Abb. 2.12(a)). Das System ist ein bimanueller Teleoperator, welcher sowohl auf Master- als auch auf Slave-Seite über 6 Freiheitsgrade verfügt. Die Instrumente (Abb. 2.12(c)) sind mit Greifern bestückt und werden mit Hilfe der bekannten PHANToM-Eingabegeräte [40] gesteuert. Der Teleoperator verfügt weder über Krafterückkopplung, noch über dreidimensionale Sicht. Als Besonderheit ist es zu erwähnen, dass die Arme (Abb. 2.12(b)) hydraulisch angetrieben werden. Die zweite Generation des Systems wurde im Jahr 2000 fertiggestellt, und war in der Lage einfache Näh- und Knotenaufgaben zu bewerkstelligen, scheint seither aber nicht weiterentwickelt worden zu sein. Das System wird in [41], [42], [43] und [44] ausführlich beschrieben.

Das KAIST-System An der Korea Advanced Institute of Science and Technology (KAIST) wurde ein Teleoperationsroboter für mikrochirurgische Aufgaben entwickelt. Er besteht aus einem parallelen Mikromanipulator (basierend auf einer Stewart-Plattform, siehe Abb. 2.13(c)), welcher an einen herkömmlichen Industrieroboter (Abb. 2.13(b)) montiert ist. Der haptische Eingabemaster (Abb. 2.13(a)) kann sowohl Kräfte als auch Momente rückkoppeln und besteht aus DC-Servomotoren mit Harmonic Drive Getrieben. Laut [45], [46] und [47] verfügt der Aufbau über kein Sichtsystem, der Schwerpunkt liegt auf Kraftmessung und Rückkopplung. Die Besonderheit des Systems liegt in der Kraftregelung, welche Reibungs- und Gravitationskräfte des Eingabegerätes erkennt und kompensiert. Die Kommunikation zwischen Master und Slave erfolgt via Ethernet.

Die Univ. Tokio Systeme Eine ganze Reihe von Experimentalsystemen auf unterschiedlichen Gebieten der Roboterchirurgie wurden von mehreren Forschungsgruppen um den Wissenschaftler Mamoru Mitsuishi von der University of Tokyo vorgestellt. Abbildung 2.14 zeigt das in [48] und [49] beschriebene Teleoperatorsystem für die Mikrochirurgie. Der Aufbau wurde speziell entwickelt, um sehr feine

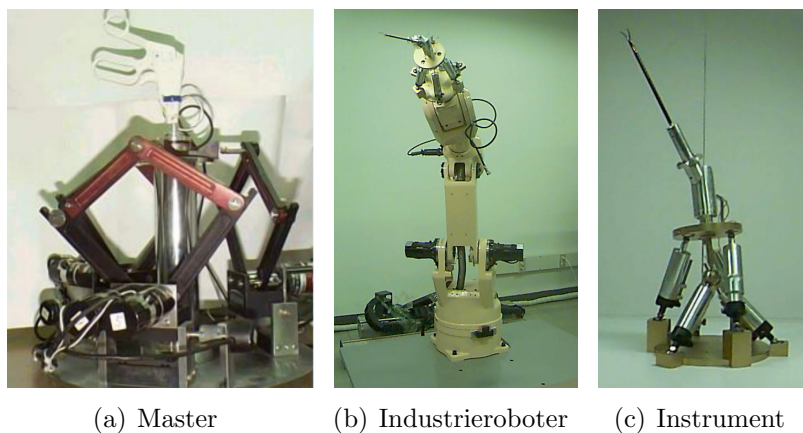


Abbildung 2.13: Das KAIST-System

anatomische Strukturen zu vernähen. Die Autoren berichten über das erfolgreiche Vernähen eines künstlichen Blutgefäßes von 1 mm Durchmesser aus einer Entfernung von 700 km. Eingabegeräte und Manipulatoren verfügen über sechs Freiheitsgrade und einem zusätzlichen Freiheitsgrad für das Öffnen und Schließen der Mikrozange.

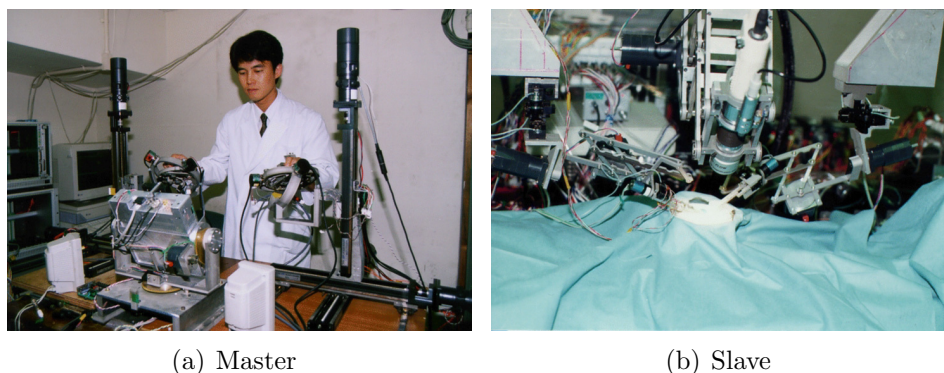


Abbildung 2.14: "Tele-micro-surgery system"

Die kreisförmig-konzentrische Konstruktionsweise des Eingabegerätes (Abb. 2.14(a)) ermöglicht eine hohe Positioniergenauigkeit. Die Spitze des Instruments verändert ihre Position nicht, wenn nur rotatorische Bewegungen stattfinden, was die Präzision weiter erhöht. Kräfte werden am Master rückgekoppelt und Bewegungen werden 1 : 20 skaliert ausgeführt. Das Mikroskop liefert nur eine zweidimensionale Sicht, doch mit Hilfe eines Laserpointers und dessen Schatten kann Tiefeninformation in ausreichendem Maße geliefert werden.

Der nächste Aufbau wurde für die Bauchraumchirurgie entwickelt, und ist ebenfalls ein bimanueller Master-Slave Manipulator bestehend aus Multi-Media-Cockpit, Arme und Kommunikationskanal. Das Multi-Media-Cockpit besteht aus kraftrückkoppelnden Eingabegeräten, Pedalen und Geräten zur audio-visuellen Wiedergabe des

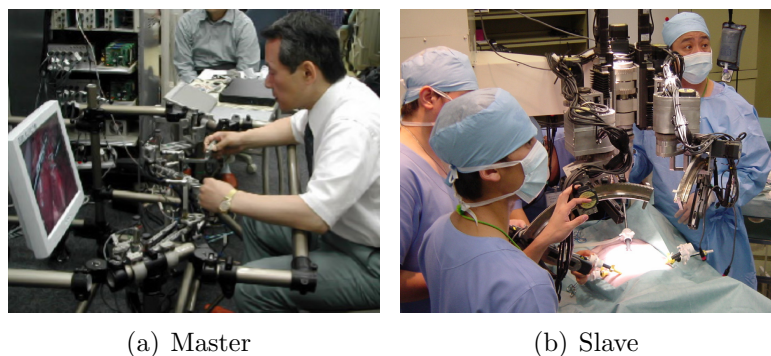


Abbildung 2.15: "Dexterous minimally-invasive surgical system"

Geschehens am Situs (Abb.2.15(a)). Die drei Arme vom SCARA-Typ (Abb.2.15(b)) halten Instrumente und Endoskop. Das Hauptaugenmerk liegt bei diesem System auf der Entwicklung abwinkelbarer Instrumente mit vielen Freiheitsgraden und Kraftrückkopplung [50]. Die patentierte Zangenkonstruktion hat vier Freiheitsgrade und ermöglicht die Messung der Greifkraft. Das System wurde in einem Experiment evaluiert, wobei aus 150 km Entfernung die Gallenblase eines Schweines erfolgreich entfernt wurde [51].

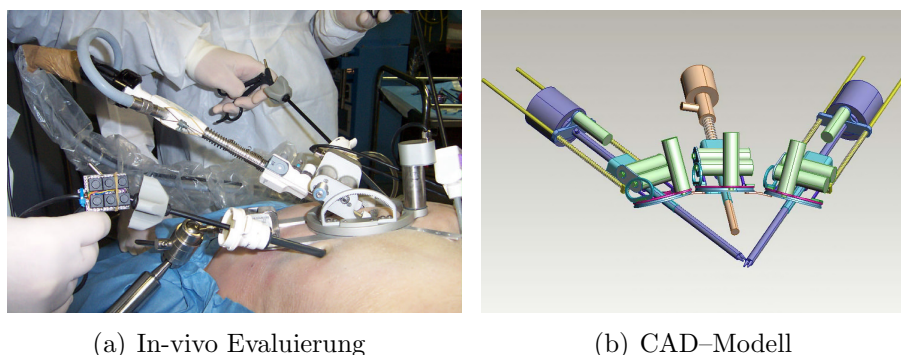


Abbildung 2.16: "Light Endoscope Robot"

Das LER-System Das in [52] und [53] erstmals beschriebene LER-System (Light Endoscope Robot) wurde ursprünglich als Endoskopführungssystem für die minimalinvasive Bauchraumchirurgie entwickelt. Das Gerät wird einfach auf dem Abdomen des Patienten mittels Gummibänder befestigt (Abb. 2.16(a)) und sorgt für die Positionierung des Endoskops. Es wurden verschiedene Benutzerschnittstellen entwickelt, sodass die Bedienung mittels Tasten, automatischer (optischer) Instrumentenverfolgung oder Sprachkontrolle erfolgen kann. Dabei operiert der Chirurg mit herkömmlichen laparoskopischen Instrumenten. Verglichen mit kommerziell erhältlichen Systemen (vgl. Abb. 2.5) bietet dieser Aufbau bei gleichem Leistungsumfang und bei erheblich kleinerem Preis und Platzbedarf die selbe Funktionalität.

Eine logische Weiterentwicklung sind chirurgische Instrumente nach dem selben Konstruktionsprinzip [54, 55]. Abbildung 2.16(b) zeigt das CAD-Modell eines bimanuellen Operators. Ein zusätzlicher rotatorischer Freiheitsgrad muss für die Instrumente vorgesehen werden, dieser war beim Endoskop nicht notwendig. Der hier vorgestellte Ansatz scheint eine kostengünstige Alternative zu den weit verbreiteten Konstruktionen der Form Roboterarm-Instrumente zu sein.

Das Remote Microsurgery System Ein neuartiges Konzept für Mikrochirurgie wurde vom Department of Micro System Engineering der Nagoya University vorgeschlagen. Anvisierte Einsatzgebiete sind schwer zugängliche und enge Körperregionen wie der Gehörgang oder die oberen Atemwege. Das System besteht aus einem flexiblen, katheter-ähnlichen Slave (Abb. 2.17(b)), an dessen Ende sich ein Mikromanipulator mit sieben Freiheitsgraden befindet. Er wird mittels DC-Servomotoren mit Harmonic Drive Getrieben angesteuert, die Bewegungen werden durch Drahtseile übertragen.

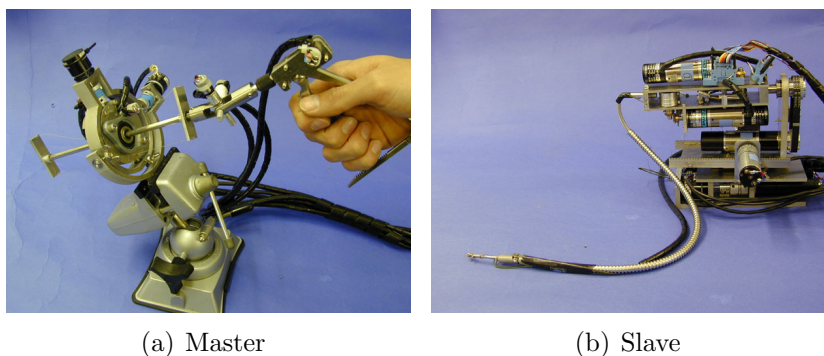


Abbildung 2.17: “Remote Microsurgery Robot for Deep and Narrow Space”

Der Master (Abb. 2.17(a)) ist dem Griff eines herkömmlichen laparoskopischen Instruments nachempfunden, und verfügt ebenfalls über sieben Freiheitsgrade, wobei das System auch das Problem umgekehrter Handbewegungen der klassischen Laparoskopie löst. Der Prototyp, mittlerweile in der dritten Generation, wurde laut [56] und [57] im Tierversuch erfolgreich getestet.

Das “Hyper Finger”-System Am Department of Micro System Engineering der Nagoya University wurde ein neuartiges Konzept eines Robotersystems für minimalinvasive Chirurgie in tiefegelegenen Organen entwickelt. Dies ist einer der kleinsten Master-Slave Roboter in der Medizin. Jeder Finger hat neun Freiheitsgrade und wird mit Hilfe von dünnen Drahtseilen betrieben. Ein prototypischer abnehmbarer Greifer wurde ebenfalls entwickelt und ein Mechanismus zur Kompensation der Längendehnung der Seile integriert. Man beachte, dass der Master (Abb.2.18(a)) nicht exoskeletal ist, wie bei den meisten Systemen, sondern von Benutzer wie ein

Stift gehalten wird. Die Konstruktion benötigt keine speziellen Roboterarme, der Manipulator kann einfach auf ein Stativ montiert werden. Das System bietet weder Stereosicht, noch Kraftrückkopplung, aber laut [58] und [59] wurde die Anwendbarkeit in in-vivo Experimenten bestätigt. Als Haupteinsatzgebiete kommen Eingriffe in schwer zugänglichen und engen Körperregionen (z.B. obere Atemwege) in Frage.

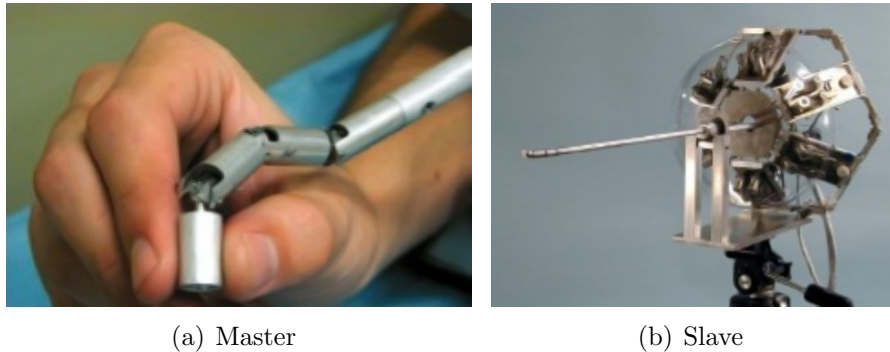


Abbildung 2.18: "Hyper Finger" von der Nagoya University

2.5 Das Endo[PA]R System



Abbildung 2.19: Endo[PA]R an der Technischen Universität München

Am Lehrstuhl "Robotics and Embedded Systems" der Technischen Universität München wurde in Zusammenarbeit mit dem Deutschen Herzzentrum München (DHM) eine realitätsnahe Experimentierplattform für robotergestützte minimalinvasive Chirurgie aufgebaut [60]. Das Akronym Endo[PA]R (Endoscopic Partial-Autonomous Robot) weist auf eines der ursprünglichen Ziele hin, nämlich die Integration teilautonomer Techniken in robotergestützte Chirurgieszenarien [61, 62].

Motiviert durch die fehlende Kraftmessung- und Rückkopplung kommerziell erhältlicher Chirurgesysteme war die Evaluation haptischer Möglichkeiten die zweite wichtige Zielsetzung [63].

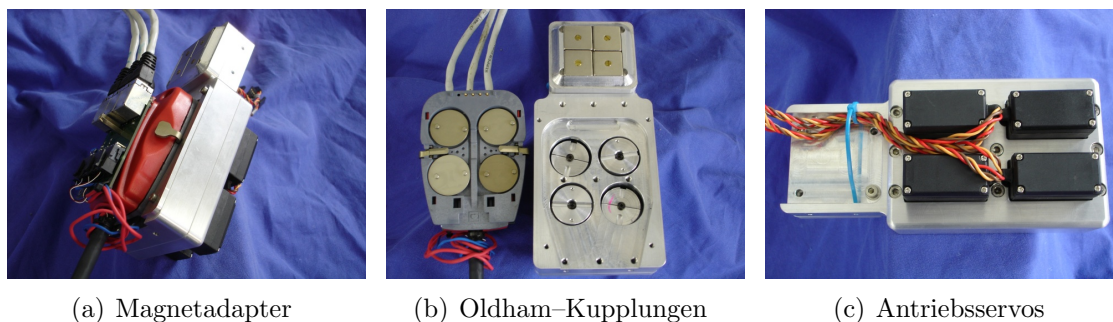


Abbildung 2.20: Instrumentenadapter und Servoeinheit

Die Ausführungseinheit bestand aus zwei Niedriglast-Robotern vom Typ Kuka KR6/2, die als Instrumententräger dienten. Speziell angefertigte Magnet-Adapter (Abb. 2.20(a)) ermöglichten die Verwendung der Originalinstrumente des *da Vinci*-Systems und erlaubten außerdem das schnelle Auswechseln der Instrumente. Der integrierte Servo-Block trieb mittels Oldham-Kupplungen (Abb. 2.20(b)) die vier Freiheitsgrade des angeschlossenen Instruments an. Die Einhaltung von invarianten Punkten (Trokarkinematik) wurde steuerungstechnisch realisiert, und nicht konstruktionstechnisch wie bei *da Vinci* [64, 65]. Der so entstandene Manipulator verfügte über alle notwendigen Freiheitsgrade im Körperinneren bei ausreichender Dynamik während der Manipulation.



Abbildung 2.21: 3D-Endoskop und Adapterflansch

Für die Stereosicht sorgte ein 3D-Endoskop der Firma Richard Wolf GmbH (Abb. 2.21), das mit Hilfe eines ähnlichen Magnetadapters an einem dritten Roboter vom Typ Stäubli RX 90 befestigt war. Die synchronen Bildströme wurden mittels einer Framegrabberkarte digitalisiert, und konnten in Realzeit dargestellt werden. Außerdem bot diese Konfiguration die Möglichkeit, parallel auch Bildverarbeitung zu betreiben. Die Bilddarstellung erfolgte am einem Head Mounted Display (HMD), das an der Bedienerkonsole (Abb. 2.22) befestigt war. An der Konsole waren auch

die Eingabegeräte vom Typ PHANToM angebracht, mit Hilfe derer die Fernsteuerung des Systems erfolgte. Sie dienten gleichzeitig der Rückkopplung der an den Instrumentenspitzen gemessenen Kräfte. Ein ebenfalls speziell angefertigter und am PHANToM–Stylus befestigter Greifmaster ermöglicht das Schließen und Öffnen der Greifbacken.

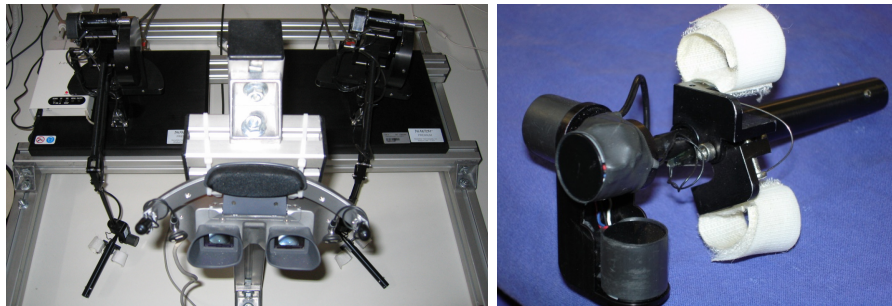
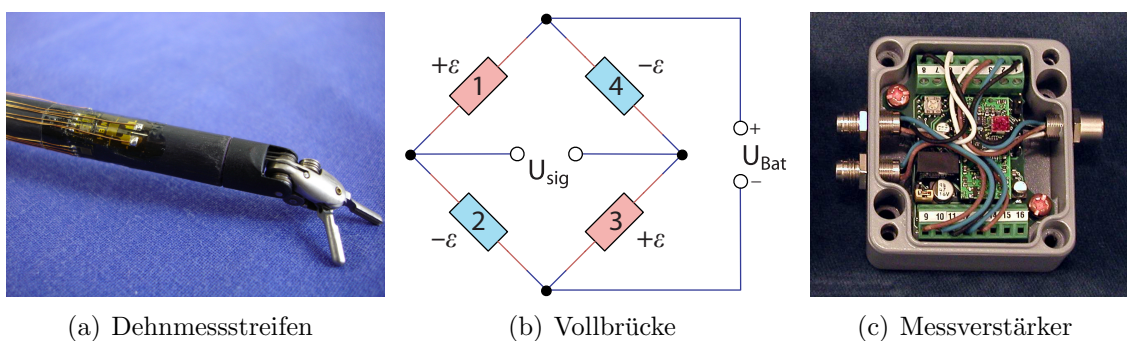


Abbildung 2.22: Bedienerkonsole und Greifmaster

Für die Kraftmessung wurden auf die Instrumente Dehnmessstreifen appliziert [66]. Diese sind passive Halbleiterbauelemente, die bei Verformung ihren elektrischen Widerstandswert ändern. Wirkt nämlich eine Kraft F auf einen elektrischen Leiter, wodurch sich seine Querschnittsfläche reduziert, so führt dies zu einer Änderung seines spezifischen Widerstandes ρ . Bei sehr kleinen Dehnungen ($\epsilon = \Delta l/l < 10^{-3}$) ist die Verformung elastisch, und es besteht ein linearer Zusammenhang zwischen mechanischer Spannung ($\sigma = F/A$), relativer Dehnung ($\epsilon = \Delta l/l$), Querkontraktion ($\Delta r = -\mu r \epsilon$) und Widerstandsänderung ($\Delta R/R$). Die Poissonzahl μ , auch Querdehnzahl genannt, ist eine Materialkonstante, die aus dem Verhältnis der relativen Dickenänderung zur relativen Längenänderung gebildet wird [67].



(a) Dehnmessstreifen

(b) Vollbrücke

(c) Messverstärker

Abbildung 2.23: Sensorbestückung der *da Vinci*-Instrumente

Abbildung 2.23(a) zeigt die applizierten Dehnmessstreifen, die mit Hilfe einer Wheatstoneschen Messbrücke in der Vollbrückenordnung (Abb. 2.23(b)) verschaltet werden. Um temperaturbedingte Widerstandsschwankungen auszufiltern wurden vier aktive Dehnmessstreifen im uniaxialen Spannungsfeld für die drei kartesischen

Richtungen X, Y, Z angeordnet. Aus der Widerstandsänderung kann die Dehnung des Dehnmessstreifens (oder des Festkörpers, mit dem er verbunden ist) bestimmt werden und schließlich die einwirkende Kraft:

$$R = \frac{l\sigma}{\pi r^2} \rightarrow \frac{\Delta R}{R} = \frac{\Delta l}{l} - \frac{\Delta r}{r} + \frac{\Delta \sigma}{\sigma} \quad (2.1)$$

$$= \left(1 + 2\mu + \frac{\Delta \sigma / \sigma}{\epsilon}\right) \epsilon \quad (2.2)$$

$$= K \cdot \epsilon \quad (2.3)$$

Für die Vollbrücke gilt $U_{sig}/U_{bat} = K \cdot \epsilon$, wobei U_{sig} die Signalspannung, U_{bat} die Speisespannung und K die Materialkonstante ist. Halbleiterdehnmessstreifen sind sehr reaktionsschnell, sodass mit Hilfe des Messverstärkers GSV3CAN der Firma ME-Messsysteme GmbH aus Abbildung 2.23(c) über 1000 Messwerte pro Sekunde erfasst werden konnten. Während Kräfte senkrecht auf den Instrumentenschaft sehr fein aufgelöst wurden, war die Messung von Kräften in longitudinaler Richtung mit starken Störungen behaftet. Dies ist darauf zurückzuführen, dass die Antriebsseile der Instrumente bei Betätigung Spannungen innerhalb der Schaftwände verursachen.

Der voll funktionsfähige Aufbau wurde von 25 Chirurgen des Deutschen Herzzentrums München bei klinisch relevanten Manipulationsaufgaben evaluiert. Das Postulat, nach dem die ungewöhnlich hohe Ermüdung bei robotergestützten Eingriffen durch die Notwendigkeit der visuellen Kompensation des haptischen Kanals zustandekommt, sollte überprüft werden. Außerdem sollte die Qualität und Performanz bei der Ausführung der Aufgaben in Abhängigkeit der Krafrückkopplung untersucht werden. Bei drei Abstufungen der Krafrückkopplung (keine, 1 : 1, 1 : 2) waren jeweils die Aufgaben Knotenlegen, Nahtmaterial an die Reißgrenze bringen und die Detektion von Stenosen zu bewältigen. Nach jedem Durchgang wurde mittels des Flimmerverschmelzungstests das Grad der Ermüdung der Probanden ermittelt. Als Ergebnisse wurden festgehalten: (i) eine Erhöhung der Kraftskalierung führt zu einer signifikanten Reduktion der aufgewendeten Kräfte, ohne dabei die Performanz zu beeinflussen; (ii) die Ermüdung wird bei zugeschalteter Krafrückkopplung ebenfalls signifikant reduziert [68, 69, 70].

Endo[PA]R verfügte bereits über autonome Fähigkeiten [71, 72, 73], sodass ein zuvor aufgezeichneter Knotenvorgang wieder abgespielt werden konnte. Dabei bestand die Möglichkeit, die Trajektorien der Instrumente zu glätten (Tremorfilterung) und mit einer höheren Geschwindigkeit abzuspielen. Der Regelkreis Sichtsystem \leftrightarrow Instrumentenarme war zu diesem Zeitpunkt jedoch noch nicht geschlossen, folglich konnte auf Fehlersituationen nicht reagiert werden. Der Aufbau erwies sich wegen den sperrigen Industrierobotern in der Bodenmontage für weitere klinische Versuche als nachteilig, und motivierte die Entwicklung des Nachfolgersystems ARAMIS².

2.6 Das ARAMIS² System

Das primäre Ziel bei der Entwicklung von ARAMIS² (Autonomous Robot-Assisted Minimally Invasive Surgery System) war ein Chirurgesystem zu entwickeln, das sich für die Anwendung in einem Operationsaal eignet. Die eminentesten Probleme des Endo[PA]R-Systems, die es für klinische Anwendungen disqualifizierten, waren die sperrige Bodenmontage, die unergonomische Konsole und die fehlende Möglichkeit einer robusten Realzeitsteuerung. Abbildung 2.24 zeigt den neuen Aufbau bestehend aus vier identischen deckenmontierten Robotern vom Typ Mitsubishi MELFA 6SL. Sie sind leichter und schlanker als die KUKA-Roboter und eignen sich sehr gut für die Deckenmontage. Die Ethernet-Schnittstelle der Roboter-Controller ermöglicht zudem die Steuerung in Realzeit mit Hilfe des UDP-Protokolls in 7ms Takt.

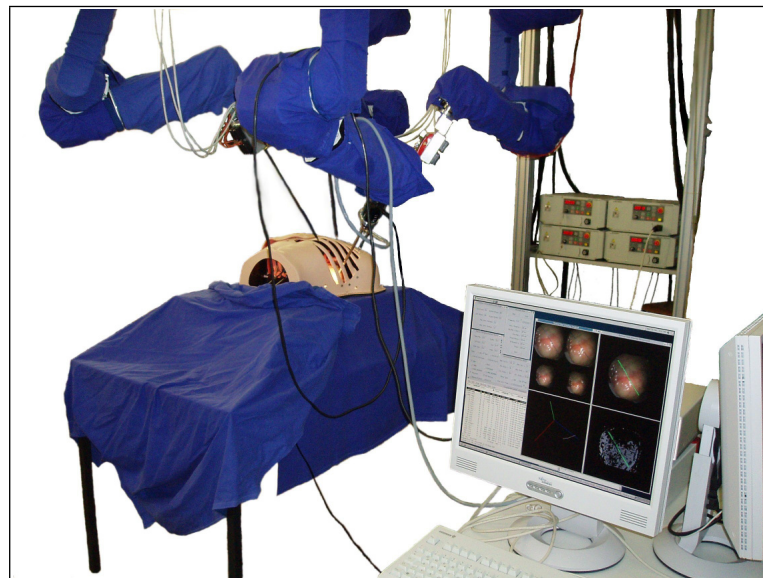


Abbildung 2.24: ARAMIS² an der Technischen Universität München

Die Hinzunahme eines dritten Assistenzarms war eine wichtige Voraussetzung für die Erprobung teilautonomer Techniken. Durch die Montage aller Roboter auf ein starres Aluminiumgerüst war die Vermessung und Kalibrierung des Gesamtsystems bestehend aus Instrumentenarmen und 3D-Endoskop möglich. Der Regelkreis Sichtsystem \leftrightarrow Instrumentenarme konnte nun geschlossen werden, wodurch das autonome Greifen des Fadenendes zum Komplettieren eines teilautonomen Knotens möglich war. Verschiedene Techniken des maschinellen Lernens wurden integriert und evaluiert [74]. Eine neuartige Methode basierend auf Analogien zur Fluidsimulation [75] erwies sich für dieses Einsatzgebiet als besonders geeignet, und erlaubt bereits nach einmaligem Vormachen eine autonome Ausführung.

ARAMIS² diente auch als Plattform zur Erprobung alternativer Bedienkonzepte eines Telechirurgiesystems. Dabei stand die Erhöhung der Ergonomie und Immersion

im Vordergrund. Ein grundsätzliches Problem bei der Bedienung ist der Umstand, dass die Eingabegeräte nur die Steuerung der Instrumentenarme ermöglichen. Um andere Aufgaben zu erledigen, wie beispielsweise die Anpassung des Endoskopbilds oder die Einstellung von Systemparametern, wird der Operateur gezwungen, auf sekundäre Eingabegeräte zurückzugreifen. Dies hat zum einen den Verlust der Immersion zur Folge, und birgt durch die Abwendung vom Operationsgebiet potentielle Gefahren für den Patienten. Um eine einheitliche Bedienung des Systems zu ermöglichen, wurde eine haptisch-graphische Benutzerschnittstelle (Haptical Graphical User Interface, HGUI) entwickelt, welche in das 3D-Sichtsystem der Konsole integriert werden kann [76].

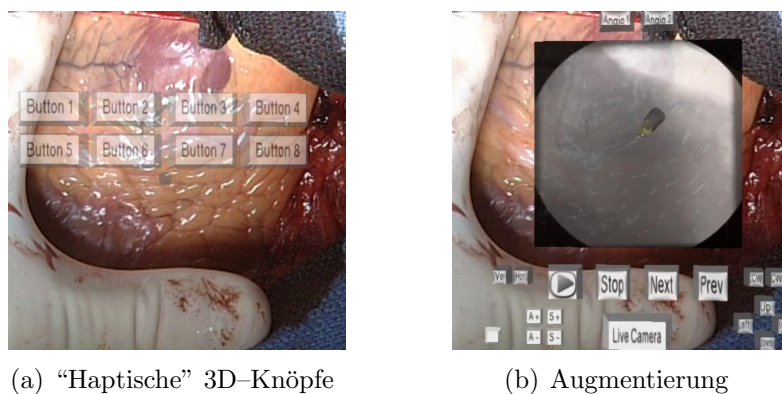


Abbildung 2.25: Dreidimensionale haptische Benutzerschnittstelle

Abbildung 2.25(a) zeigt das Funktionsprinzip: eine halbtransparente Bedienoberfläche mit dreidimensionalen virtuellen Knöpfen wird beim Betätigen eines Fußpedals über die Livebilder vom Operationsgebiet eingeblendet. Gleichzeitig verliert der Chirurg die Kontrolle über die Roboterarme, die vom System in ihrer letzten Position gehalten werden. Nun kann der Operateur die freigewordenen Eingabegeräte zur Bedienung der Oberfläche heranziehen, ohne dabei das Operationsgebiet aus den Augen zu verlieren. Als “Mauszeiger” dient ein 3D-Modell einer Instrumentenspitze, wobei die Interaktion mit den HGUI-Elementen durch simulierte Druckpunkte und Trägheiten haptisch unterstützt wird. Desweiteren ermöglicht die Benutzerschnittstelle die Einblendung von Vitalfunktionen oder präoperativ aufgenommenen Daten (CT-Scans, Angiogramme) in das Sichtfeld (Abb. 2.25(b)).

Für die Resynchronisation zwischen Eingabegerät(en) und Instrument(en) wurde ein sogenannter *Snap-Mode* integriert [77]. Abbildung 2.26 veranschaulicht die Vorgehensweise. Ein Eingabegerät, visualisiert durch das virtuelle 3D-Modell eines Instruments, wird in der Simulation mit einem realen Instrument in Deckung gebracht. Stimmen Position und Orientierung von Modell und Instrument überein, erlangt der Operateur die Kontrolle wieder und blendet das Benutzerinterface aus. Mit Hilfe des *Snap-Mode* kann auch die beliebige Abbildung zweier Eingabegeräte auf drei Instru-

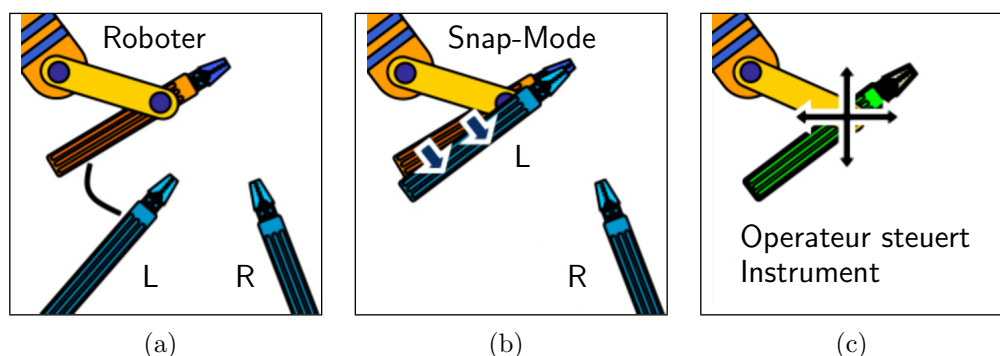


Abbildung 2.26: *Snap Mode* zur Synchronisation Eingabegerät \leftrightarrow Instrument

mentenarme erfolgen. Dadurch ist auch das kurzzeitige “Parken” eines Instruments möglich, oder das Initiieren einfacher teilautonomer Assistenzfähigkeiten.

Die minimalinvasiven Instrumente wurden ebenfalls für einen teilautonomen Betrieb vorbereitet. Sie wurden dahingehend weiterentwickelt, dass sie durch Plug&Play-fähigkeit ein automatisches Auswechseln im Betrieb ermöglichen [78, 67, 79].

2.7 Aufbau in der Großtier-OP am DHM

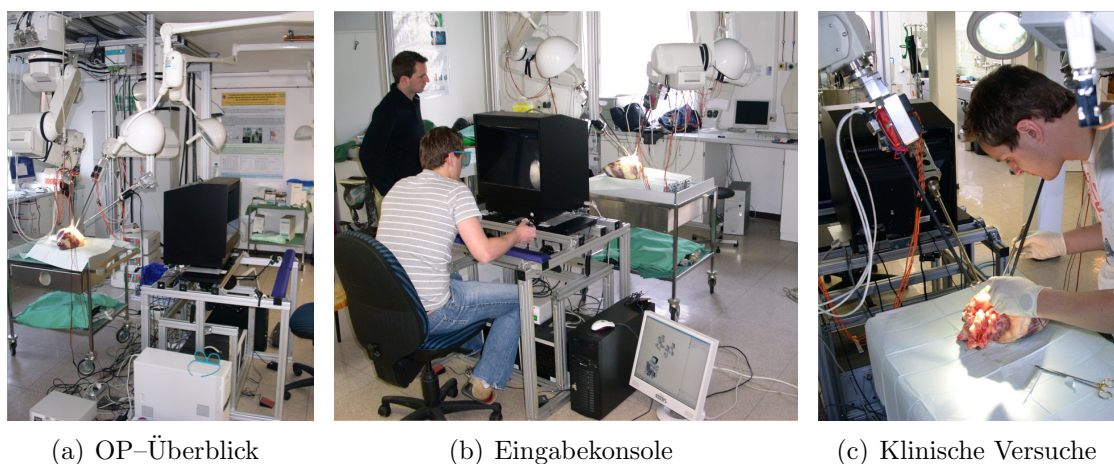


Abbildung 2.27: Aufbau am Deutschen Herzzentrum München

Für eine weitere Evaluation unter klinischen Bedingungen wurde ARAMIS² ins Deutsche Herzzentrum München umgezogen. In einem speziell ausgerüsteten Operationsaal, der für die Durchführung von Großtierexperimenten zugelassen ist, wurde der Aufbau ebenfalls in der Deckenmotage realisiert. Die Eingabekonsolle wurde dahingehend verbessert, dass die PHANToM-Eingabegeräte horizontal gespiegelt montiert wurden, wodurch ein größerer Arbeitsraum entstand [80]. Auch die 3D-Darstellung

wurde verbessert, statt des HMD kommt nun ein Polarisationsdisplay und entsprechende Brillen zum Einsatz, das einen wesentlich besseren Tiefeneindruck vermittelt. Die Montage des Displays über den Eingabegeräten bietet gegenüber dem Aufbau aus Abbildung 2.22 den Vorteil, dass der freigewordene Platz für eine Armauflagefläche verwendet werden kann. Der Arbeitsplatz gewinnt an Ergonomie und ermöglicht auch längere Eingriffe ermüdungsfrei durchzuführen. Desweiteren kann einer begrenzten Anzahl von Zuschauern ebenfalls eine 3D-Sicht vom Operationsgebiet geboten werden.

Auch dieses System wurde von einem Kollektiv bestehend aus 15 Studenten und 15 Chirurgen evaluiert. Im Gegensatz zur Evaluation des Endo[PA]R-Systems, wo die Auswirkungen der Krafrückkopplung untersucht wurden, stand hier die Benutzbarkeit bei der Ausführung realistischer herzchirurgischer Teilaufgaben im Vordergrund. Die drei Aufgaben der Versuchsreihe waren das Vernähen eines einfachen Oberflächenschnitts, das Schließen eines Vorhofscheidewanddefektes (engl. *atrial septum defect*, ASD) und die Ersetzung eines Papillarmuskels. Abbildung 2.28 zeigt Momentaufnahmen dieser Eingriffe.

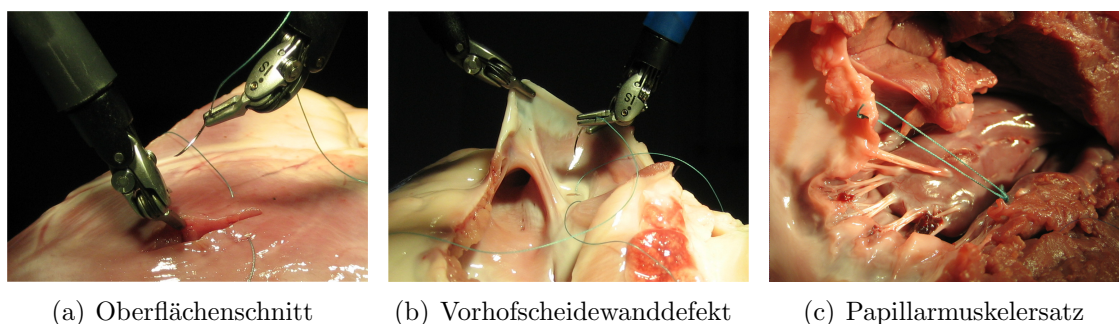


Abbildung 2.28: Evaluation unter klinischen Bedingungen

Sowohl Chirurgen als auch Studenten waren in der Lage, die spezifizierten Aufgaben zu erledigen, wobei es eine signifikante Differenz in der Ausführungszeit zugunsten der Chirurgen gab. Sie benötigten zwischen 45 und 110 Minuten (im Schnitt 75 Minuten), während Studenten zwischen 70 und 170 Minuten (im Schnitt 105 Minuten) für die Erledigung der Aufgaben beanspruchten. Die Ergebnisse dieser Evaluation wurden zum Zeitpunkt der Entstehung dieser Arbeit noch nicht genau ausgewertet, zeigen jedoch einen leichten Vorteil zugunsten erfahrener Chirurgen. Die Qualität der Krafrückkopplung schien bei diesen Tätigkeiten eher eine untergeordnete Rolle zu spielen, sie wurde sogar oft als störend empfunden. Das wichtigste Ergebnis ist jedoch, dass die Verwendbarkeit des ARAMIS²-Systems unter klinischen Bedingungen gezeigt werden konnte.

Kapitel 3

Kalibrierung

Zusammenfassung In einem Aufbau bestehend aus Roboter(n) und Kamerasystem(en) ist es möglich, den beobachteten Arbeitsraum mit den darin befindlichen Objekten genau zu vermessen und zu modellieren, um diese gegebenenfalls zu manipulieren. Hierfür muss die Voraussetzung erfüllt sein, dass das Robotersystem über eine interne Repräsentation der Szene verfügt, welche möglichst genau der Realität entspricht. Das erreicht man in zwei Schritten. Als erstes ermöglicht ein Stereokamerasystem aus der beobachteten Szene genaue 3D-Koordinaten zu extrahieren. Zweitens müssen die gewonnenen Raumpunkte zu einem Modell zusammengefasst, und dem Robotersystem adäquat aufbereitet zur Verfügung gestellt werden. Voraussetzung für beide Schritte ist eine entsprechende Kalibrierung. Zum einen muss das Stereokamerasystem photogrammetrisch vermessen werden. Dabei werden Parameter wie Bildmittelpunkt, Brennweiten der Linsen und ihre bauartbedingte Position zueinander ermittelt. Zum anderen benötigt man die relativen Lagen von Robotern und Kameras zueinander, sodass als Ergebnis alle Komponenten in das selbe kartesische Koordinatensystem transformiert werden können. Von einer meistens werkseitig korrekt durchgeführten Gelenkkalibrierung der Roboter wird im allgemeinen ausgegangen.

Unter Kalibrierung einer Kamera versteht man hier die Erstellung eines möglichst genauen Modells, welches 3D-Weltkoordinaten auf 2D-Bildkoordinaten abbildet (und umgekehrt). Die Idee dabei ist, dass man die zu ermittelnden Kameraparameter solange iterativ anpasst, bis die Abbildungen bekannter Raumpunkte mit den erwarteten übereinstimmen. Die erwähnten Raumpunkte sind gut wiedererkennbare Merkmale von Objekten, welche man im allgemeinen Kalibrierkörper nennt. Für die Kalibrierung von Stereokamerasystemen verwendet man überwiegend planare Kalibrierkörper mit regulären Mustern und maschinell einfach zu extrahierenden Merkmalen. Der Hauptgrund dafür ist, dass diese preiswert und relativ genau mit heutiger Laserdrucktechnik hergestellt werden können. Die Nachteile sind die relativ große Anzahl aufzunehmender Bilder und ungenaue Kalibrierergebnisse bei suboptimaler Anzahl und Auswahl von Aufnahmen des Kalibrierkörpers. Es finden sich

zwar Hinweise in der Literatur auf dreidimensionale Kalibrierkörper, doch diese bestehen meistens nur aus rechtwinklig zueinander angeordneten planaren Flächen. In der Domäne der minimalinvasiven Chirurgie ergeben sich außerdem noch zwei gravierende Nachteile solcher Kalibrierkörper: zum einen können sie nicht in die sterile Umgebung des Operationsgebietes gebracht werden. Zum anderen macht die typischerweise häufige Verstellung der Kameraparameter während eines Eingriffs eine zeitaufwendige Rekalibrierung notwendig.

In dieser Arbeit wird ein neuartiger 3D-Kalibrierkörper vorgestellt, wodurch einige obengenannte Nachteile (quasi-)planarer Kalibrierkörper beseitigt werden. Es handelt sich dabei um pyramidenartig angeordnete Balken, welche aus einem Aluminiumblock mittels CNC-Technik mit hoher Genauigkeit ausgefräst wurden. Dies gewährleistet eine ausreichende Anzahl von Punkten, deren 3D-Weltkoordinaten bekannt sind, und deren Position im Bild genau bestimmt werden kann. Dadurch ergeben sich mehrere Vorteile: der Zeitaufwand der Kalibrierung reduziert sich erheblich, und kann auch vom medizinischen Personal durchgeführt werden. Dank des Fertigungsmaterials ist der Kalibrierkörper sterilisierbar und damit tauglich für den Einsatz in Operationssälen, ohne dabei Abnutzungserscheinungen aufzuweisen. Für die Kalibrierung des Gesamtsystems (“Hand-Auge-Kalibrierung”), bestehend aus Stereoendoskop und mehreren, mit chirurgischen Instrumenten bestückten Roboterarmen, kann der Kalibrierkörper ebenfalls verwendet werden. Dafür muss der bedienende Arzt lediglich mit jedem der Instrumentenarme bestimmte Punkte des Kalibrierkörpers anfahren. Aus den bekannten 3D-Weltkoordinaten dieser Punkte, und aus den zu berechnenden 3D-Kamerakoordinaten werden die für die Umrechnung notwendigen Transformationen automatisch ermittelt. Ein weiterer Vorteil der vorgestellten Methode ist die anschließende Verfügbarkeit von genauen metrischen Daten, welche sowohl zur Kollisionserkennung, als auch zur Vermessung des Operationsitus herangezogen werden können.

Durch das Aufkommen neuer, auf binokulare Bildverarbeitung basierender Technologien in der robotergestützten minimalinvasiven Chirurgie ergibt sich die Notwendigkeit, auch an die neuen Aufgabenbereiche angepasste Rahmenbedingungen zu schaffen. Zum einen ergeben sich höhere Anforderungen an die Genauigkeit, zum anderen muss die klinische Tauglichkeit gewährleistet sein. Die beschriebene Methode kann mit wenig Aufwand in bestehende Chirurgesysteme integriert werden, und würde sich, bestimmte Anpassungen vorausgesetzt, prinzipiell als einheitliches Framework zur Kalibrierung robotergestützter Chirurgesysteme eignen.

3.1 Motivation

Die ursprüngliche Bedeutung des Wortes *Kalibrierung* bezieht sich auf die Überprüfung von Messgeräten, für die es keine gesetzlichen Vorgaben gibt. Demgegenüber steht die *Eichung*, die die vom Gesetzgeber vorgeschriebene Prüfung eines Messgerätes auf Einhaltung der zugrundeliegenden eichrechtlichen Vorschriften vornimmt. Das Wörterbuch der Metrologie (VIM)¹ definiert Kalibrierung:

Kalibrieren umfasst die Tätigkeiten zur Ermittlung des Zusammenhanges zwischen den ausgegebenen Werten eines Messmittels [...] und den bekannten Werten der Messgröße unter bekannten Bedingungen.

Man könnte also den Begriff Kalibrierung auch als einen Abgleich zwischen einem bestimmten (mathematischen) Modell und Realität definieren. In der Robotik umfasst die Kalibrierung zwei Vorgänge: zum einen gewährleistet sie die Korrektheit der Gelenkwinkel, um eine hohe Positionier- und Wiederholgenauigkeit zu erreichen. Zum anderen integriert sie das Robotersystem samt Werkzeugen in seinen Arbeitsraum. Im folgenden wird von einer (meistens werkseitig) korrekt durchgeführten Gelenkkalibrierung ausgegangen, und nur die Schritte zur Einbettung des Robotersystems in einen konkreten Aufbau betrachtet. Das primäre Ziel der in diesem Kapitel vorgestellten Methoden ist es, alle Komponenten eines Roboterchirurgieaufbaus in ein und das selbe kartesische Koordinatensystem zu bringen. Wir werden dieses gemeinsame Koordinatensystem, wie in der Literatur üblich, das *Weltkoordinatensystem* nennen.

3.2 Szenario

Die hier vorliegende Arbeit beschäftigt sich schwerpunktmäßig mit der Erprobung und Integration von Techniken aus der (Stereo-)Bildverarbeitung in minimalinvasive robotergestützte Chirurgiesysteme. Ein typisches Roboterchirurgieszenario besteht aus kalibriertechnischer Sicht aus mehreren mit Instrumenten bestückten Roboterarmen und einem (Stereo-)Kamerasystem. Abbildung 3.1 zeigt schematisch den Aufbau eines solchen Systems, wobei aus Gründen der Übersichtlichkeit nur eine Kamera und ein Instrument ohne den dazugehörigen Roboterarm dargestellt sind. Obwohl binokulare Sicht für einen menschlichen Operateur nicht zwingend notwendig ist [81, 82], benötigt ein (teil-)autonomes System genaue Positionsangaben vom Operationsgebiet. Dies ist nur mit Hilfe eines stereoskopischen Aufbaus von Kameras mit ausreichender Genauigkeit zu erreichen. Wegen der Übersichtlichkeit werden wir meistens nur eine Kamera (die linke) des Stereo-Kamerasystems darstellen. Die Position und Orientierung der anderen Kamera kann durch eine homogene Transformationsmatrix relativ zur ersten angegeben werden. Diese Transformationsmatrix

¹„Internationales Wörterbuch der Metrologie“, welches grundlegende und allgemeine Begriffe der Metrologie erklärt (Vocabulaire international des termes fondamentaux et généraux de métrologie)

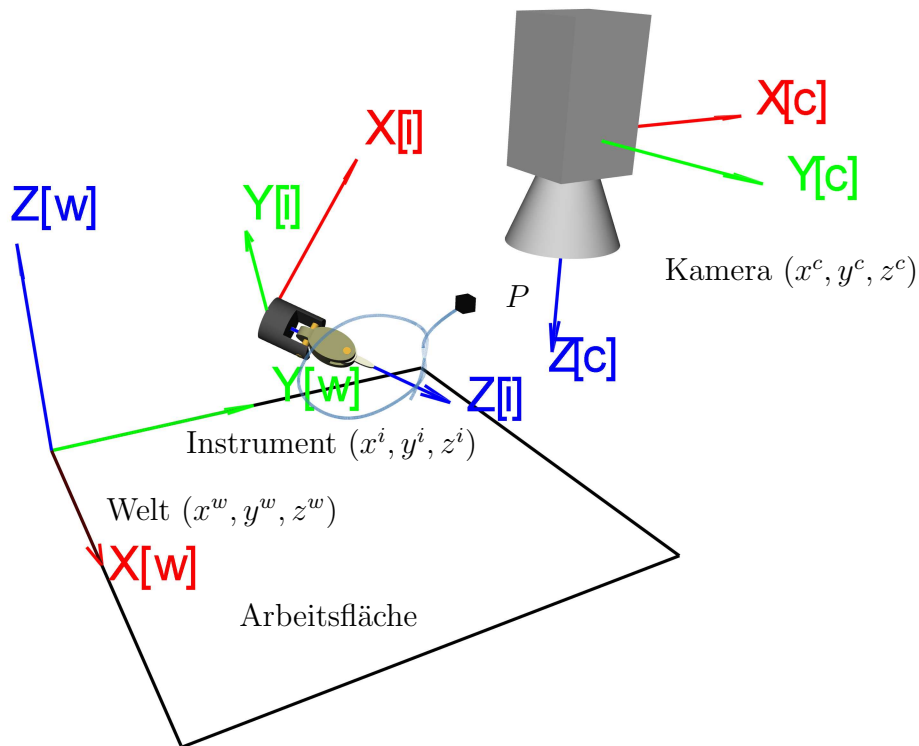


Abbildung 3.1: Schematische Darstellung eines Aufbaus für Roboterchirurgie

(*extrinsische Parameter*) ist eines der Ergebnisse der Stereo-Kamerakalibrierung. Zu den beteiligten Koordinatensystemen sind noch einige Bemerkungen notwendig. Im Koordinatensystem einer Kamera des Stereoendoskops werden die Objekte nach der 3D-Rekonstruktion angegeben. Desweiteren befinden sich alle zum Aufbau gehörenden Instrumente in einem anderen gemeinsamen Koordinatensystem. Um eine über die Bildverarbeitung geschlossene Steuerungsschleife Aktorik \leftrightarrow Sensorik zu gewährleisten, ist eine entsprechende Kalibrierung des Gesamtsystems, auch “Hand-Auge-Kalibrierung” genannt, notwendig.

3.3 Begriffe und Notationen

Bevor wir anfangen über Kalibrierung zu sprechen, müssen wir einige notwendige Begriffe und Notationen einführen, zumal diese in der Literatur oft recht abweichend dargestellt werden. Die grundlegendsten Fragen, die wir beantworten müssen, sind:

- Wie beschreibt man eine beliebige Transformation (Translation und Rotation) von Punkten und Koordinatensystemen,
- wie beschreibt man die Position und Orientierung eines Koordinatensystems relativ zu einem anderen, und

- wie bestimmt man die Koordinaten eines Punktes in verschiedenen Koordinatensystemen, genauer, wie transformiert man Koordinaten zwischen Koordinatensystemen.

In der Tat kann man alle oben genannten Aufgaben mit Hilfe ein und desselben mathematischen Werkzeugs lösen, und zwar mit homogenen Transformationsmatrizen. Schauen wir uns also zunächst kartesische Koordinatensysteme mit ihren Transformationsmöglichkeiten im dreidimensionalen Raum genauer an.

3.3.1 3D-Koordinaten

Die Position eines Punktes P im dreidimensionalen Raum beschreibt man mit Hilfe seiner 3D-Koordinaten (x_p, y_p, z_p) . Die Koordinaten werden üblicherweise als ein 3D-Vektor interpretiert und mit \mathbf{p} bezeichnet. Das Koordinatensystem, in welchem die Koordinaten angegeben werden, erscheint als hochgestellter Index des Vektors. Demzufolge hat ein Punkt P in einem Aufbau bestehend aus einer Kamera und einer Arbeitsfläche zwei verschiedene Koordinaten \mathbf{p}^c und \mathbf{p}^w bezogen jeweils auf die Kamera und auf die Arbeitsfläche (Welt):

$$\mathbf{p}^c = \begin{pmatrix} x_p^c \\ y_p^c \\ z_p^c \end{pmatrix} \quad \mathbf{p}^w = \begin{pmatrix} x_p^w \\ y_p^w \\ z_p^w \end{pmatrix} \quad (3.1)$$

Abbildung 3.2 zeigt den Punkt P im Kamera- und Weltkoordinatensystem. Die Koordinatensysteme müssen natürlich nicht rechtwinklig (wie im Bild) zueinander angeordnet sein. Vielmehr besteht die Schwierigkeit darin, die genaue räumliche Relation zweier beliebig zueinander angeordneten Koordinatensysteme zu ermitteln.

3.3.2 Translationen (Verschiebungen)

Translation von Punkten Die beliebige Verschiebung (Translation) eines Punktes in einem bestimmten Koordinatensystem kann mittels einer einfachen Vektoraddition beschrieben werden. Die Koordinate eines verschobenen Punktes P_2 berechnet sich aus der Addition seines ursprünglichen Koordinatenvektors \mathbf{p}_1 und dem Verschiebungsvektor \mathbf{t}_1 (Abbildung 3.3):

$$\mathbf{p}_2 = \mathbf{p}_1 + \mathbf{t}_1 = \begin{pmatrix} x_{p_1} + x_{t_1} \\ y_{p_1} + y_{t_1} \\ z_{p_1} + z_{t_1} \end{pmatrix} \quad (3.2)$$

Mehrere Translationen werden durch sukzessives Aufaddieren der Verschiebungsvektoren erreicht, wobei die Operation kommutativ ist, d.h. die Reihenfolge der Additionen beeinflusst das Ergebnis nicht.

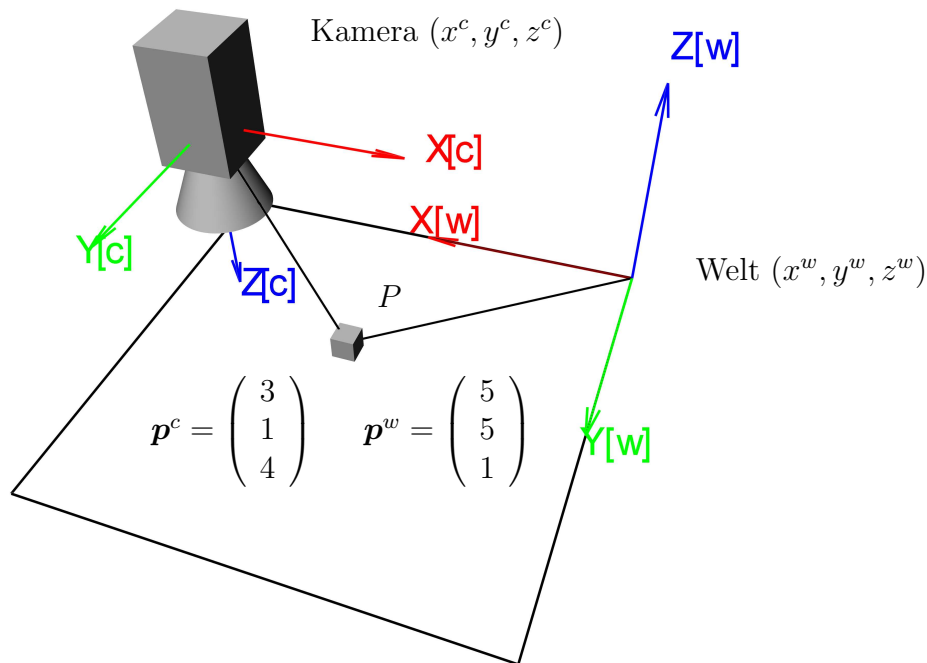


Abbildung 3.2: Koordinaten eines Punktes in den zwei verschiedenen Koordinatensystemen c (Kamera) und w (Welt)

Translation von Koordinatensystemen Analog werden Koordinatensysteme verschoben. Wenn beispielsweise ein Koordinatensystem c_1 verschoben wird um das neue Koordinatensystem c_2 zu bilden, dann ist die Position von c_2 in c_1 , also der Koordinatenvektor seines Ursprungs relativ zu c_1 ($\mathbf{o}_{c_2}^{c_1}$), identisch mit dem Verschiebungsvektor (Abbildung 3.3):

$$\mathbf{t}_2^{c_1} = \mathbf{o}_{c_2}^{c_1} \quad (3.3)$$

Koordinatentransformationen Desweiteren kann die Koordinatentransformation eines Punktes zwischen (verschobenen) Koordinatensystemen als eine Verschiebung samt seines lokalen Koordinatensystems aufgefasst werden. Verschieben wir beispielsweise das Koordinatensystem c_1 , zusammen mit dem Punkt Q_1 , um einen Vektor \mathbf{t}_2 , mit dem Ergebnis c_2 und Q_2 , dann haben die Punkte Q_1 und Q_2 die selben Koordinaten bezogen auf ihr lokales Koordinatensystem ($\mathbf{q}_1^{c_1} = \mathbf{q}_2^{c_2}$). Wenn also Koordinatensysteme relativ zueinander verschoben werden, dann können Koordinaten einfach durch Addition des Verschiebungsvektors transformiert werden:

$$\mathbf{q}_2^{c_1} = \mathbf{q}_2^{c_2} + \mathbf{t}_2^{c_1} = \mathbf{q}_2^{c_2} + \mathbf{o}_{c_2}^{c_1} \quad (3.4)$$

Abbildung 3.3 visualisiert den Sachverhalt: \mathbf{q}_2^{c1} , der Koordinatenvektor von Q_2 im Koordinatensystem c_1 , wird berechnet durch Addition des Verschiebungsvektors \mathbf{t}_2 mit dem Koordinatenvektor von Q_2 im Koordinatensystem c_2 (\mathbf{q}_2^{c2}).

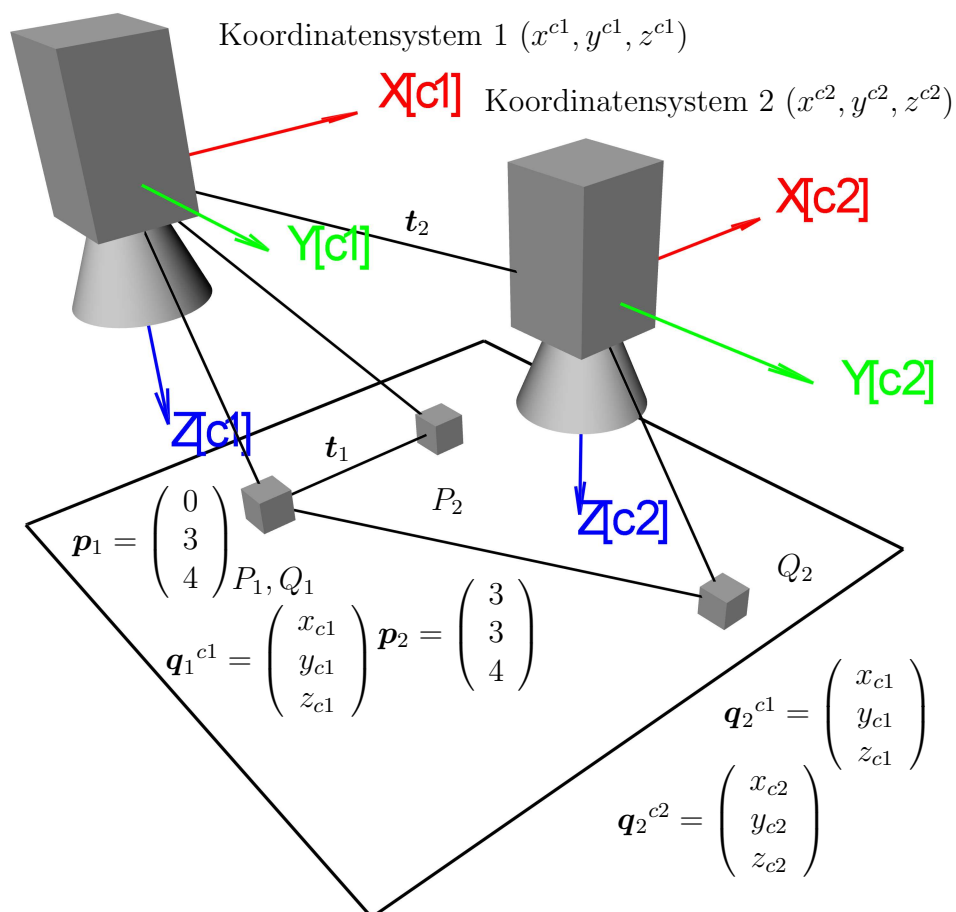
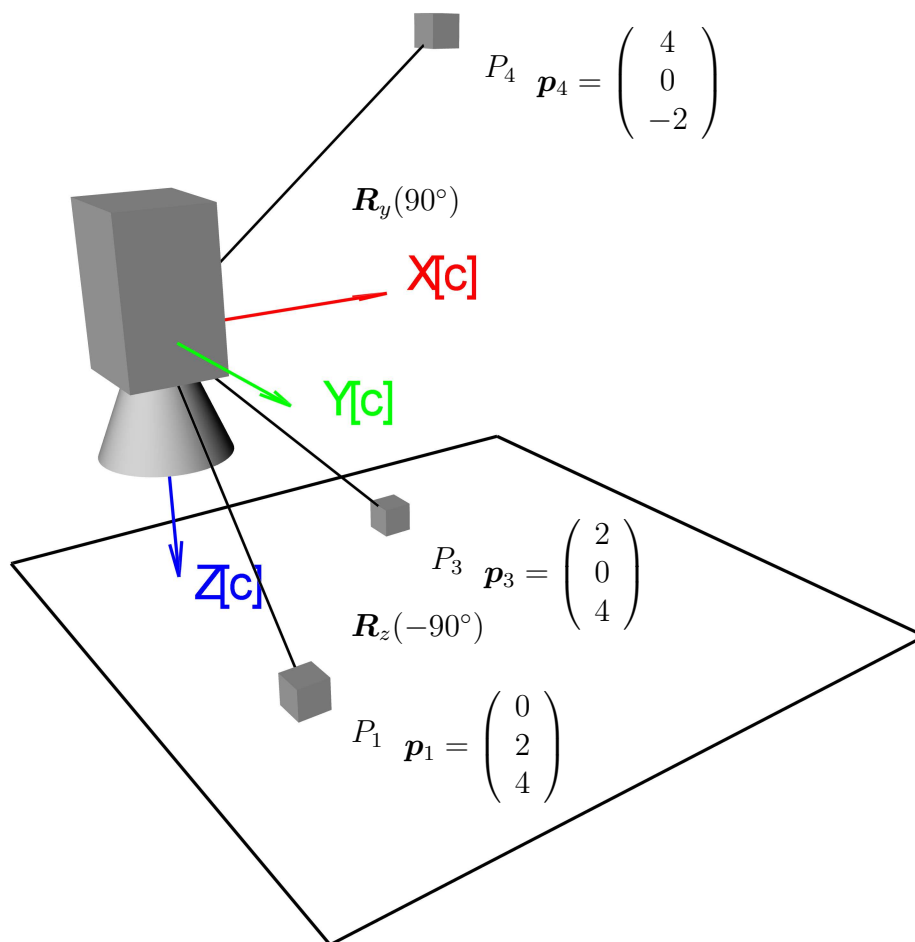


Abbildung 3.3: Verschiebung von Punkten und Koordinatensystemen

3.3.3 Rotationen (Drehungen)

Rotation von Punkten Punkte können in kartesischen Koordinatensystemen um bestimmte Achsen des Koordinatensystems oder um eine beliebige Achse rotiert werden. Die Drehung eines Punktes um eine Achse des Koordinatensystems wird durch eine Multiplikation seines Koordinatenvektors mit einer 3×3 Rotationsmatrix \mathbf{R} ausgedrückt. Die Drehung eines Punktes P_1 um die z -Achse (und P_3 als Ergebnis) sieht beispielsweise folgendermassen aus:

$$\mathbf{p}_3 = \mathbf{R}_z(\gamma) \cdot \mathbf{p}_1 = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} x_{p_1} \\ y_{p_1} \\ z_{p_1} \end{pmatrix} = \begin{pmatrix} \cos \gamma \cdot x_{p_1} - \sin \gamma \cdot y_{p_1} \\ \sin \gamma \cdot x_{p_1} + \cos \gamma \cdot y_{p_1} \\ z_{p_1} \end{pmatrix} \quad (3.5)$$

Abbildung 3.4: Drehung um z - und y -Achse

Die entsprechenden Rotationsmatrizen für die x - und y -Achse lauten:

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad \mathbf{R}_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \quad (3.6)$$

Ketten von Rotationen Wenn nun der gedrehte Punkt weiter um die y -Achse gedreht wird, so kann dies mittels einer Kette von Rotationsmatrizen sehr elegant beschrieben werden:

$$\mathbf{p}_4 = \mathbf{R}_y(\beta) \cdot \mathbf{p}_3 = \mathbf{R}_y(\beta) \cdot \mathbf{R}_z(\gamma) \cdot \mathbf{p}_1 \quad (3.7)$$

Zu beachten ist, dass im Unterschied zur Multiplikation von Skalaren, die Multiplikation von Matrizen nicht kommutativ ist, d.h. die Vertauschung der Reihenfolge der Rotationsmatrizen ein anderes Ergebnis liefert.

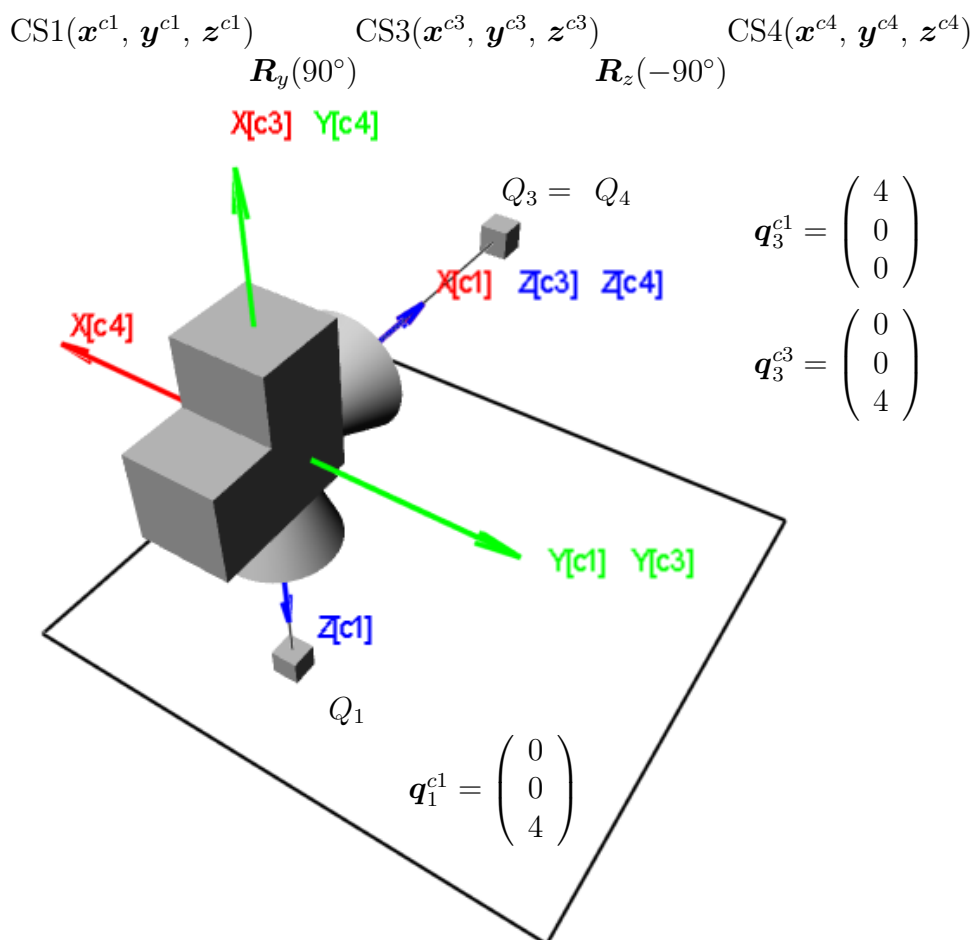


Abbildung 3.5: Drehung vom Koordinatensystem (und Punkt) um z - und y -Achse

Rotation von Koordinatensystemen Koordinatensysteme haben im Gegensatz zu Punkten auch eine Orientierung relativ zu einem anderen Koordinatensystem. Diese Orientierung ändert sich wenn das Koordinatensystem gedreht wird. Ähnlich wie bei der Position eines Koordinatensystems, gegeben durch einen Verschiebungsvektor, ist die Drehung in einer Rotationsmatrix enthalten. Die Spalten der Rotationsmatrix entsprechen den Achsenvektoren des gedrehten Koordinatensystems in Koordinaten des ursprünglichen:

$$\mathbf{R} = \begin{bmatrix} \mathbf{x}_{c_3}^{c_1} & \mathbf{y}_{c_3}^{c_1} & \mathbf{z}_{c_3}^{c_1} \end{bmatrix} \quad (3.8)$$

Beispielsweise können die Axisvektoren des Koordinatensystems c_3 aus Abbildung 3.5 aus der entsprechenden Rotationsmatrix $\mathbf{R}_y(90^\circ)$ berechnet werden:

$$\mathbf{R}_y(90^\circ) = \begin{bmatrix} \cos(90^\circ) & 0 & \sin(90^\circ) \\ 0 & 1 & 0 \\ -\sin(90^\circ) & 0 & \cos(90^\circ) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad (3.9)$$

Daraus folgt nun:

$$\mathbf{x}_{c_3}^{c_1} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \quad \mathbf{y}_{c_3}^{c_1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{z}_{c_3}^{c_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad (3.10)$$

Koordinatentransformationen Wie bei der Translation, um Punktkoordinaten aus dem rotierten Koordinatensystem c_3 in das ursprüngliche Koordinatensystem c_1 zu transformieren, wendet man die selbe Transformation auf Punktkoordinaten an, wie sie auch beim Koordinatensystem c_3 zum Einsatz kam. Man multipliziert die Punktkoordinaten mit der Rotationsmatrix, wie sie auch für die Transformation von c_1 nach c_3 verwendet wurde:

$$\mathbf{q}_3^{c_1} = {}^{c_1}\mathbf{R}_{c_3} \cdot \mathbf{q}_3^{c_3} \quad (3.11)$$

Die zur Abbildung 3.5 gehörende Kette von Rotationen entspricht folgenden Gleichungen:

$$\mathbf{q}_4^{c_1} = {}^{c_1}\mathbf{R}_{c_3} \cdot {}^{c_3}\mathbf{R}_{c_4} \cdot \mathbf{q}_4^{c_4} = \mathbf{R}_y(\beta) \cdot \mathbf{R}_z(\gamma) \cdot \mathbf{q}_4^{c_4} = {}^{c_1}\mathbf{R}_{c_4} \cdot \mathbf{q}_4^{c_4} \quad (3.12)$$

Reihenfolge der Achsen in Ketten von Rotationen Wenn man die Ketten von Rotationen aus Abbildung 3.4 und 3.5, und die dazugehörigen Gleichungen 3.7 und 3.12 betrachtet, dann fällt auf, dass zwei verschiedene Abfolgen von Rotationen mit der selben Kette von Rotationsmatrizen beschrieben werden. So wurde in Abbildung 3.4 der Punkt erst um die z -Achse, und dann um die y -Achse, während in Abbildung 3.5 das Koordinatensystem erst um die y -Achse und dann um die z -Achse rotiert wurde. Doch beide Transformationen werden durch $\mathbf{R}_y(\beta) \cdot \mathbf{R}_z(\gamma)$ beschrieben. Der Grund hierfür ist, dass Ketten von Rotationsmatrizen in verschiedene Richtungen gelesen werden können: in Abbildung 3.4 von rechts nach links, und in Abbildung 3.5 von links nach rechts. Es muss jedoch einen Unterschied geben, zumal die Multiplikation von Matrizen nicht kommutativ ist. Den Unterschied machen die Reihenfolgen der Achsen aus, um welche rotiert wird.

Bei der zweiten Rotation des Koordinatensystems aus Abbildung 3.5 gibt es zwei Möglichkeiten für die Reihenfolge der Achsen, um welche rotiert wird: um die des “alten” Koordinatensystems c_1 , und um die des bereits rotierten “neuen” Systems c_3 . Im Beispiel wurde die zweite Drehung um die “neue” z -Achse durchgeführt. Demgegenüber gibt es bei der Drehung von Punkten aus Abbildung 3.4 nur eine Möglichkeit, und zwar um die Achsen des “alten” Koordinatensystems. Zusammenfassend kann man nun sagen:

- Wenn wir eine Kette von **links nach rechts** lesen, so wird um die “**neuen**” Achsen rotiert.
- Wenn wir eine Kette von **rechts nach links** lesen, so wird um die “**alten**” Achsen rotiert.

Wie bereits erwähnt werden Punktrotationen immer von rechts nach links gelesen. Im Falle von Koordinatensystemen hat man die Wahl, doch in den meisten Fällen ist es intuitiver von links nach rechts zu lesen.

3.3.4 Rigide Transformationen

Wenn man Translationen und Rotationen kombiniert, bekommt man sogenannte rigide² Transformationen. In Abbildung 3.6 werden beispielsweise die Translation aus Abbildung 3.3 und die Rotationen aus Abbildung 3.4 kombiniert. Eine solche Transformation wird folgendermassen beschrieben:

$$\mathbf{p}_5 = \mathbf{R} \cdot \mathbf{p}_1 + \mathbf{t} \quad (3.13)$$

Für mehrere Transformationen können solche Gleichungen schnell unübersichtlich werden, wie folgendes Beispiel zweier Transformationen zeigt:

$$\mathbf{p}_6 = \mathbf{R}_a \cdot (\mathbf{R}_b \cdot \mathbf{p}_1 + \mathbf{t}_b) + \mathbf{t}_a = \mathbf{R}_a \cdot \mathbf{R}_b \cdot \mathbf{p}_1 + \mathbf{R}_a \cdot \mathbf{t}_b + \mathbf{t}_a \quad (3.14)$$

3.3.5 Homogene Transformationsmatrizen

Eine elegante Alternative sind sogenannte *homogene Transformationsmatrizen* und die entsprechenden *homogenen Vektoren*. Eine homogene Transformationsmatrix \mathbf{H} enthält sowohl die Rotationsmatrix als auch den Translationvektor. So kann beispielsweise die rigide Transformation aus Gleichung 3.13 folgendermassen geschrieben werden:

²soll auch noch die Skalierung erlaubt sein, so heißt die Transformation affin

$$\begin{pmatrix} \mathbf{p}_5 \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} \mathbf{p}_1 \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R} \cdot \mathbf{p}_1 + \mathbf{t} \\ 1 \end{pmatrix} = \mathbf{H} \cdot \mathbf{p}_1 \quad (3.15)$$

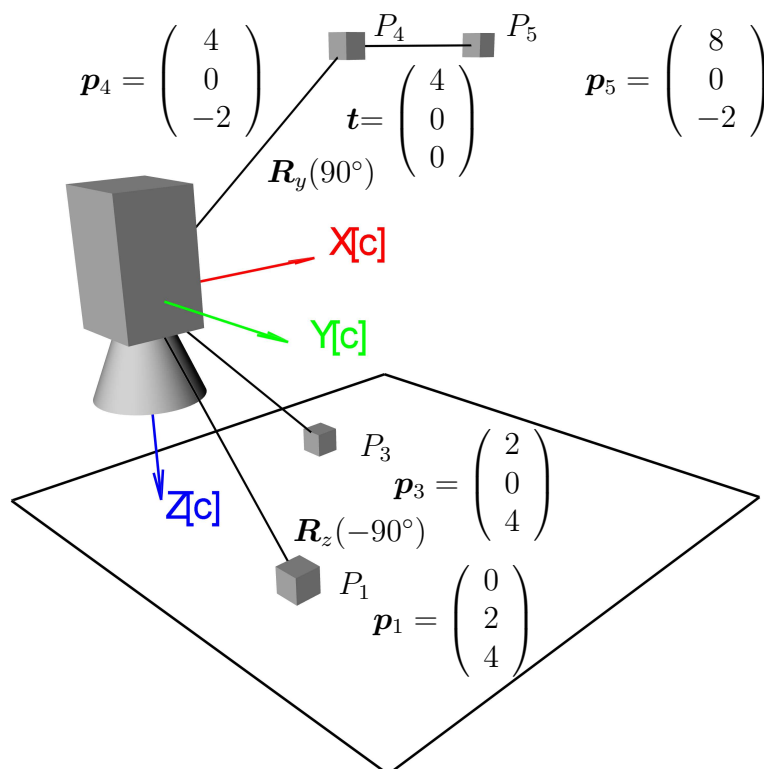


Abbildung 3.6: Die Translation aus Abbildung 3.3 und die zwei Rotationen aus Abbildung 3.4 werden kombiniert und bilden eine rigide Transformation

Der Nutzen dieser Notation wird ersichtlich, wenn man mit Sequenzen von rigiden Transformationen zu tun hat. Diese können, analog zu Ketten von Rotationen, auch entsprechend beschrieben werden:

$$\mathbf{H}_1 \cdot \mathbf{H}_2 = \begin{bmatrix} \mathbf{R}_a & \mathbf{t}_a \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R}_b & \mathbf{t}_b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_a \cdot \mathbf{R}_b & \mathbf{R}_a \cdot \mathbf{t}_b + \mathbf{t}_a \\ 0 & 1 \end{bmatrix} \quad (3.16)$$

Wie bereits bei Ketten von Rotationen beschrieben, können auch Ketten von rigiden Transformationen in beiden Richtungen gelesen werden. Wenn von links nach rechts gelesen, erfolgen die Transformationen um die “neuen” Achsen, und wenn von rechts nach links gelesen, um die “alten” Achsen.

Eigentlich ist eine rigide Transformation bereits eine Kette, weil sie ja aus einer Translation und einer Rotation besteht:

$$\mathbf{H} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \\ 0 & 1 & 1 & \mathbf{t} \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} & & & 0 \\ \mathbf{R} & & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{H}(\mathbf{t}) \cdot \mathbf{H}(\mathbf{R}) \quad (3.17)$$

Wenn eine Rotation aus mehreren Rotationen um verschiedene Achsen wie in Abbildung 3.6 besteht, so können die einzelnen Rotationen auch als homogene Transformationsmatrizen geschrieben werden:

$$\mathbf{H} = \begin{bmatrix} \mathbf{R}_y(\beta) \cdot \mathbf{R}_z(\gamma) & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \\ 0 & 1 & 1 & \mathbf{t} \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} & & & 0 \\ \mathbf{R}_y(\beta) & & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} & & & 0 \\ \mathbf{R}_z(\gamma) & & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.18)$$

Von rechts nach links gelesen kann man die Transformation des Punktes P aus Abbildung 3.6 nachvollziehen: erst wird um die z -Achse, dann um die (“alte”) y -Achse rotiert, und zum Schluss translatiert.

Rigide Transformation von Koordinatensystemen Rigide Transformationen von Koordinatensystemen funktionieren auf die gleiche Weise wie bereits für separate Rotationen und Translationen beschrieben. Demnach beschreibt die homogene Transformationsmatrix ${}^{c_1}\mathbf{H}_{c_5}$ die Transformation des Koordinatensystems c_1 in das Koordinatensystem c_5 . Gleichzeitig beschreibt sie Position und Orientierung des Koordinatensystems c_5 relativ zum Koordinatensystem c_1 : ihre Spaltenvektoren enthalten die Koordinaten der Achsenvektoren und des Ursprungs.

$${}^{c_1}\mathbf{H}_{c_5} = \begin{bmatrix} \mathbf{x}_{c_5}^{c_1} & \mathbf{y}_{c_5}^{c_1} & \mathbf{z}_{c_5}^{c_1} & \mathbf{o}_{c_5}^{c_1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.19)$$

Wie bei Rotationen bereits erwähnt, werden Ketten von rigiden Transformationen von Koordinatensystemen von links nach rechts gelesen. Demnach wird die Transformationskette aus Gleichung 3.18 wie folgt gelesen: erst wird das Koordinatensystem verschoben, dann wird um die “neue” y -Achse rotiert, und anschließend um die “neueste” z -Achse rotiert.

Koordinatentransformationen Wie bereits für getrennte Translationen und Rotationen beschrieben, werden Koordinaten von Punkten aus einem transformierten Koordinatensystem c_5 in das ursprüngliche Koordinatensystem c_1 dadurch transformiert, dass man die selbe Transformation anwendet, welche auch für c_5 benutzt

wurde. Man multipliziert einfach den Koordinatenvektor mit der homogenen Transformationsmatrix:

$$\begin{pmatrix} \mathbf{p}_5^{c_1} \\ 1 \end{pmatrix} = {}^{c_1} \mathbf{H}_{c_5} \cdot \begin{pmatrix} \mathbf{p}_5^{c_5} \\ 1 \end{pmatrix} \quad (3.20)$$

Typischerweise werden die homogenen Vektoren weggelassen, falls keine Verwechslungsgefahr besteht, und wir schreiben einfach:

$$\mathbf{p}_5^{c_1} = {}^{c_1} \mathbf{H}_{c_5} \cdot \mathbf{p}_5^{c_5} \quad (3.21)$$

3.4 Kameramodell und Parameter

3.4.1 Binokularer Aufbau

Mit Hilfe eines entsprechend angeordneten Aufbaus von zwei Kameras (siehe Abbildung 3.7) ist es möglich, 3D-Koordinaten einer beobachteten Szene zu gewinnen. Dabei liefert eine Kette bestehend aus Linse, Bildsensor (CCD-Chip) und optionalem Framegrabber Bilder von optisch überlappenden Bereichen, woraus mit Hilfe geometrischer Werkzeuge kartesische Koordinaten berechnet werden können. Um genaue 3D-Koordinaten aus einem Stereobildpaar zu erhalten, ist eine Kalibrierung des Stereokamerasystems notwendig. Unter Kalibrierung versteht man hier die Erstellung eines möglichst genauen Modells, die kontinuierliche 3D-Weltkoordinaten auf diskrete 2D-Pixelkoordinaten abbildet (und umgekehrt). Für die 3D-Rekonstruktion ist es notwendig, dass man bei einer kalibrierten Kamera jedem Bildpunkt einen Strahl zuordnen kann, auf dem der abgebildete Weltpunkt liegen muss. Erst dann können die Strahlen zweier korrespondierender Bildpunkte miteinander geschnitten und die Position des Weltpunktes genau bestimmt werden. Wir betrachten zunächst die Eigenschaften einer einzelnen Kamera, die Besonderheiten eines stereoskopischen Aufbaus werden in einem späteren Kapitel behandelt. Wenn wir von einer Kamera sprechen, so ist immer die Kombination von Bildsensor (CCD-Chip) und Linse gemeint.

3.4.2 Bildgebende Sensoren und Linsen

Digitalkameras funktionieren im Wesentlichen wie klassische Fotoapparate mit dem Unterschied, dass die Lichtstrahlen anstatt auf einen lichtempfindlichen Film, auf einen Sensor treffen. Dreidimensionale Szenen werden dabei nach den gleichen geometrischen Prinzipien auf eine Ebene abgebildet. Die Sensorfläche besteht entweder aus einem CCD-Chip, oder wird in CMOS-Technologie gefertigt, wobei die Funktionsweise in beiden Fällen auf den inneren photoelektrischen Effekt basiert. Dabei

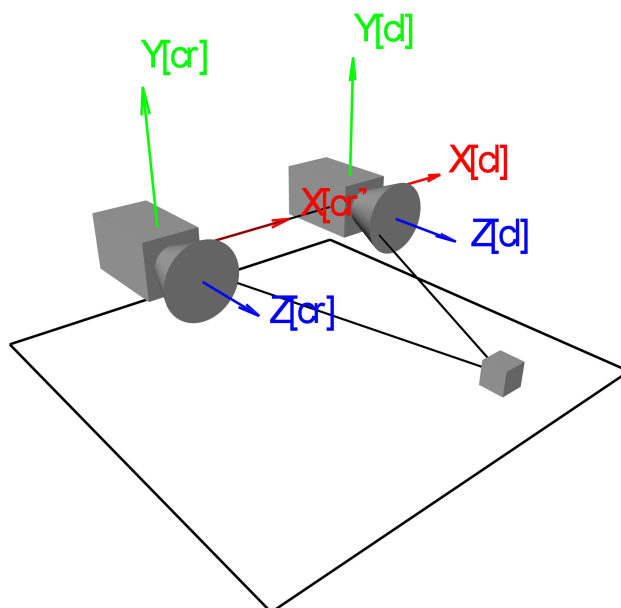


Abbildung 3.7: Stereokamerasystem zur genauen 3D-Rekonstruktion

werden beim Auftreffen von Photonen Elektronen freigesetzt, welche anschließend abgetastet, verstärkt und digitalisiert werden können. Die Verschlusszeit wird elektronisch geregelt. Der Sensor an sich ist nicht auf die verschiedenen Farben empfindlich, sondern nur auf Lichtintensität. Farbfiltermasken direkt vor der Sensorfläche sorgen für die entsprechende Aufbereitung eines Farbbildes. Aufwendigere Konstruktionen verwenden spektrale Strahlenteiler-Prismen und getrennte Sensoren für die Grundfarben.

Der Ort der Entstehung des eigentlichen Bildes bestimmt, ob es sich um ein analoges oder digitales Kamerasystem handelt. Wird eine Wandlerkarte (Framegrabber) zur Aufbereitung des Bildes für die weitere Verarbeitung im Rechner benötigt, so spricht man von einem analogen System. Die Nachteile dieser Methode sind zum einen bedingt durch die langen analogen Leitungswege, und zum anderen durch die zwei verschiedenen Stellen der Bildentstehung. Bei rein digitalen Kamerasystemen mit digitalen Schnittstellen (*FireWire*, *USB*) findet der komplette Bildaufbau in der Kamera statt und wird über eine genormte Schnittstelle zum Rechner übertragen. Obwohl zunächst für den low-budget Sektor bestimmt, bieten heute immer mehr Hersteller Kameras mit diesen Schnittstellen für den professionellen Einsatz an.

Es gibt grundsätzlich zwei relevante Arten von Linsen, welche bei der rechnergestützten Bildverarbeitung zum Einsatz kommen. Die erste Art bewirkt eine perspektivische Projektion von Weltkoordinaten ins Kamerabild, wie dies auch beim menschlichen Auge der Fall ist. Dadurch werden Objekte umso kleiner im Bild, desto weiter entfernt von der Kamera sie sich befinden. Eine solche Kombination von Bildsensor und Linse bezeichnet man als *Lochkameramodell* (engl. *pinhole camera model*). Die Bezeichnung ist darauf zurückzuführen, dass diese Art von perspektivi-

scher Projektion auch dadurch erreicht werden kann, indem ein kleines Loch in ein planares Objekt gebohrt wird, welches nun parallel vor ein anderes planares Objekt (die Bildebene) plaziert wird. Das Lochkameramodell war auch ursprünglich das erste Modell eines Fotoapparates (vgl. *camera obscura*).

Die zweite Art von Linsen, welche für maschinelle Bildverarbeitung noch relevant ist, sind telezentrische Linsen. Der Hauptunterschied zum vorherigen Linsenmodell ist, dass sie eine parallele Projektion von Weltkoordinaten auf die Bildebene bewirken, zumindest für einen bestimmten Entfernungsbereich zwischen Objekt und Kamera. Das hat zur Folge, dass im Bild alle Objekte gleich groß sind, unabhängig von ihrer Entfernung zur Linse. Diese Kombination von Linse und Kamera nennt man das *telezentrische Kameramodell*.

3.4.3 Das Lochkameramodell

Eines der am häufigsten verwendeten Kameramodelle in der Bildverarbeitung ist das Lochkameramodell. Unter Berücksichtigung der Verzerrung realer Systeme hat es Lenz [83] bereits 1987 beschrieben. Weng et al. [84] beschreiben später ein komplexeres Modell, welches sich insbesondere den Verzerrungen widmet. Wir betrachten zunächst das einfache verzerrungsfreie Modell, und erweitern es anschließend nur um notwendige Verzerrungsmodelle. Abbildung 3.8 zeigt schematisch die perspektivische Projektion einer Lochkamera. Der Punkt P wird durch das optische Zentrum der Linse auf den Punkt P' in der Bildebene projiziert, welches sich im Abstand f (Brennweite) hinter dem optischen Zentrum befindet. Das Bild des Punktes ist demnach der Schnittpunkt der Verbindungsgeraden zum Zentrum mit der Bildebene. Obwohl sich die Bildebene in Wirklichkeit hinter dem optischen Zentrum der Linse befindet, vereinfacht es manche Berechnungen, wenn man festlegt, dass sie im Abstand f vor dem optischen Zentrum (wie in Abbildung 3.9) liegt. Dadurch erreicht man, dass das Bildkoordinatensystem wie das Pixelkoordinatensystem ausgerichtet ist (Zeilenkoordinaten nehmen nach unten zu, und Spaltenkoordinaten nach rechts). Das Kamerakoordinatensystem ist so definiert, dass die x und y Achsen parallel zu den Spalten und Zeilen des Bildes ausgerichtet sind, und die z Achse senkrecht zur Bildebene. Mit dieser Konvention sind wir nun in der Lage, Projektionen von Objekten mit ihren 3D-Koordinaten in die 2D Bildebene zu beschreiben, die hierzu notwendigen Kameraparameter vorausgesetzt.

Transformationskette der Lochkamera

Wie bereits eingangs erwähnt, muss jedem Bildpunkt, dessen 3D-Koordinate ermittelt werden soll, ein Strahl zugeordnet werden, welcher durch das optische Zentrum der Kamer verläuft. Wir werden diesen Strahl zunächst nicht wie beim eigentlichen Verarbeitungsprozess vom Bildpunkt aus verfolgen, sondern gehen den umgekehrten Weg vom Objektpunkt einer 3D-Szene aus. Die Projektion eines

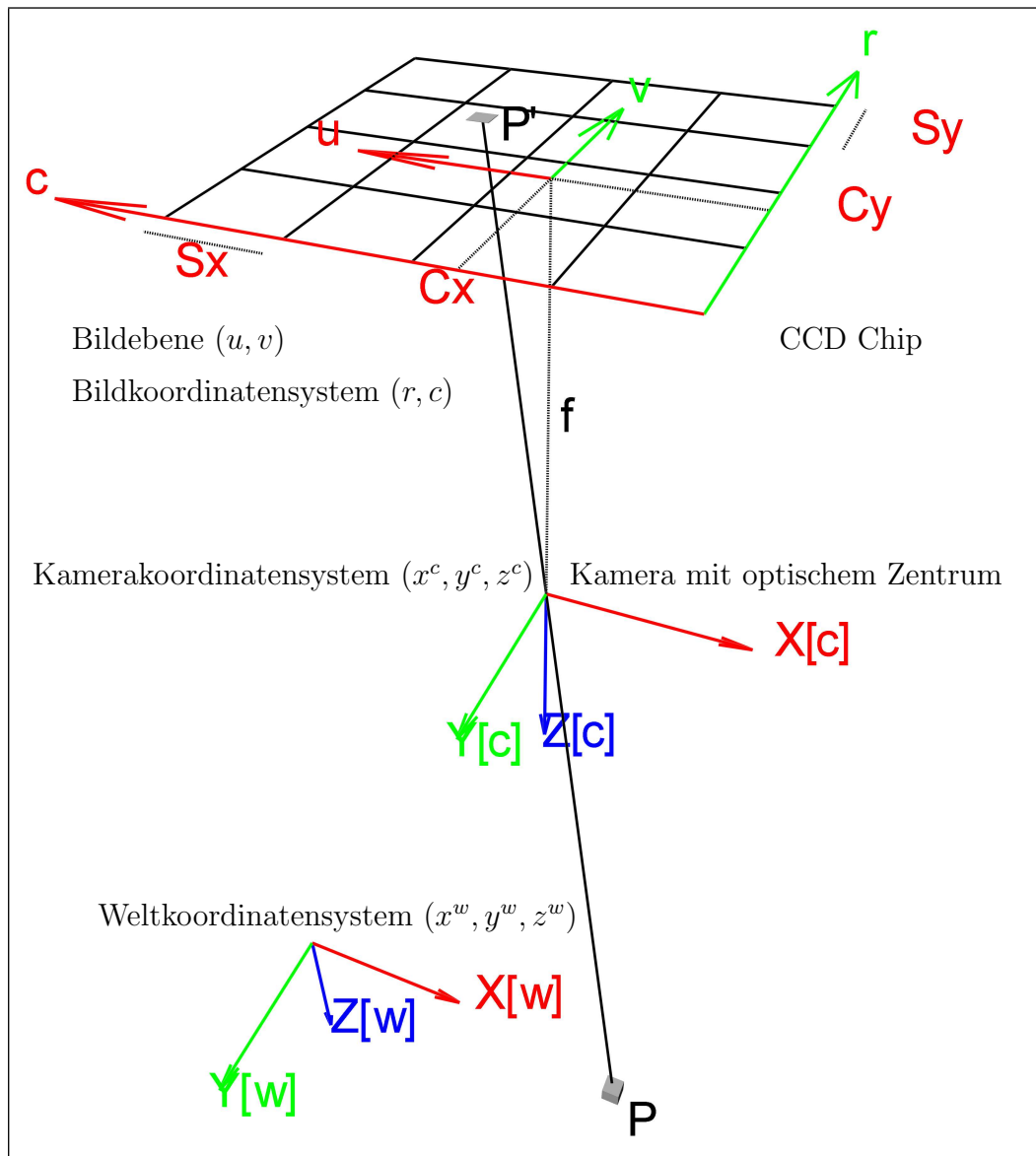


Abbildung 3.8: Perspektivische Projektion einer Lochkamera

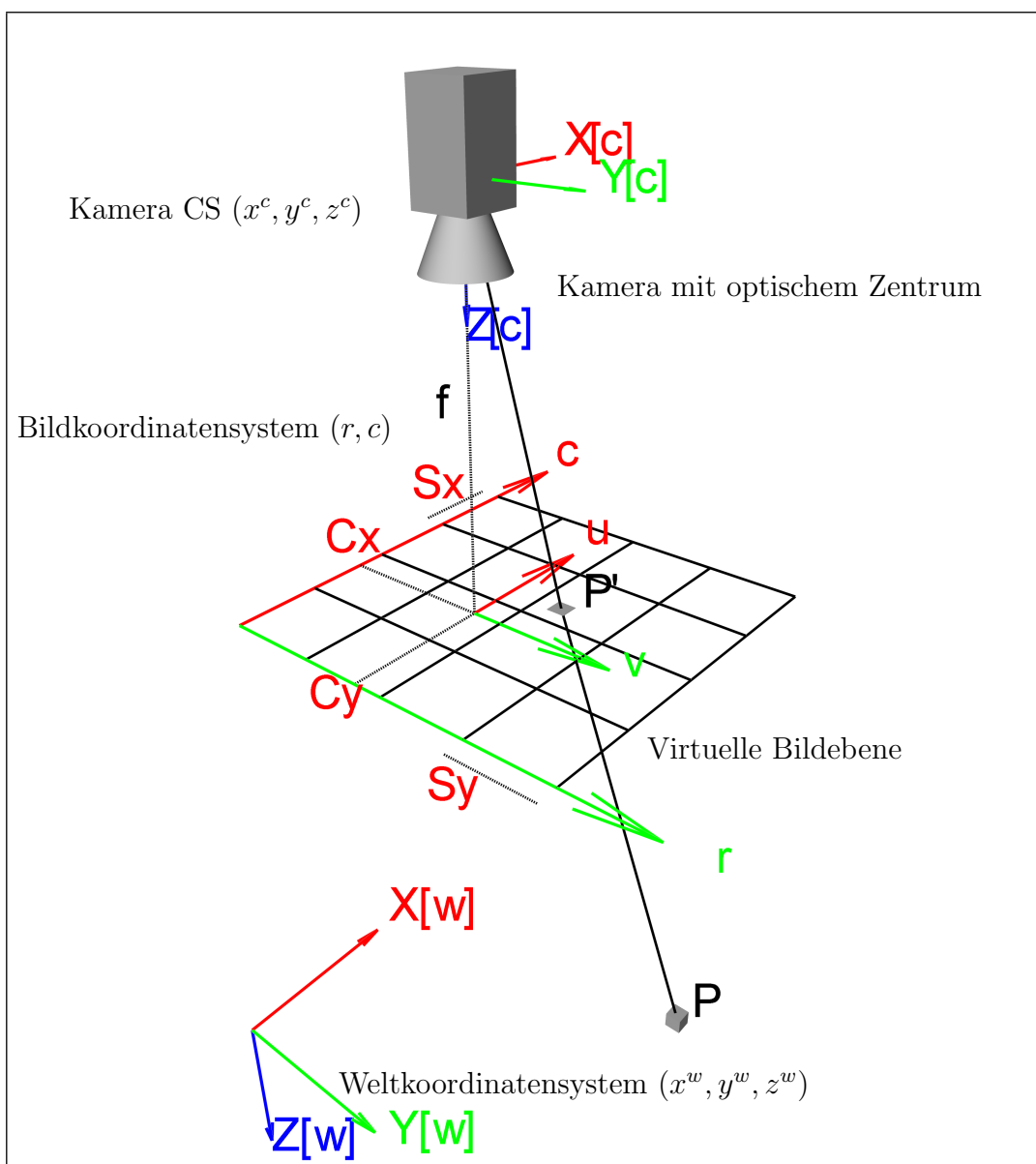


Abbildung 3.9: Virtuelle Bildebene

Punktes P mit seinen Weltkoordinaten in das gerätespezifische Pixel P' kann durch eine Kette von geometrischen Transformationen beschrieben werden. Dabei wird zunächst $p^w = (x^w, y^w, z^w)^T$ aus dem Weltkoordinatensystem (WCS) in den Punkt $p^c = (x^c, y^c, z^c)^T$ des Kamerakoordinatensystems (CCS) transformiert. Die verzerrungsfreie perspektivische Projektion der Lochkamera bildet p^c auf $p^i(u, v)$ des Bildkoordinatensystems (ICS) ab. Unter Berücksichtigung etwaiger Verzerrungen muss der Punkt p^i auf $p^d(u^d, v^d)^T$ korrigiert werden. Anschließend können die diskreten Koordinaten des Pixels $p^p(r, c)$ im Pixelkoordinatensystem (PCS) ermittelt werden. Diese Transformationskette kann (mit bestimmten Einschränkungen) in beide Richtungen durchlaufen werden:

$$p^w \mapsto p^c \mapsto p^i \mapsto p^d \mapsto p^p \quad (3.22)$$

Jede Stufe benötigt bestimmte Parameter, die für jeden Aufbau unterschiedlich ausfallen, und während der Kalibrierung ermittelt werden. Die Parameter lassen sich in externe (extrinsische) und interne (intrinsische) aufteilen. Erstere geben die Lage der Kamera bezüglich eines globalen Koordinatensystems (Weltkoordinatensystem) an. Die internen Parameter beschreiben die Abbildung der Weltpunkte vom lokalen Koordinatensystem der Kamera in das Bild. Für die Abbildungskette aus Gleichung 3.22 gibt es zunächst kein einheitliches mathematisches Modell, weil hier affine Transformationen in homogenen Koordinaten ($p^w \mapsto p^c$), perspektivische Projektionen mittels inhomogenen Koordinaten ($p^c \mapsto p^i$), nichtlinearen Näherungen ($p^i \mapsto p^d$) und Diskretisierungsschritte ($p^d \mapsto p^p$) miteinander kombiniert werden müssen. Im Folgenden werden die einzelnen Transformationsschritte und die notwendigen Parameter näher beschrieben.

1. *Weltkoordinaten* \mapsto *Kamerakoordinaten* Als Erstes erfolgt die Transformation des Punktes P aus dem Weltkoordinatensystem (WCS) in das Kamerakoordinatensystem (CCS). Diese ist eine rigide Transformation, und kann deshalb mit Hilfe einer Transformationsmatrix beschrieben werden. Der Zusammenhang zwischen dem Punkt P mit seinen Weltkoordinaten $p^w = (x^w, y^w, z^w)^T$ und seiner Projektion in die Kamerakoordinaten $p^c = (x^c, y^c, z^c)^T$ ist gegeben durch

$$\begin{pmatrix} x^c \\ y^c \\ z^c \end{pmatrix} = \mathbf{R} \cdot \begin{pmatrix} x^w \\ y^w \\ z^w \end{pmatrix} + \mathbf{t}, \quad (3.23)$$

wobei $\mathbf{t} = (t_x, t_y, t_z)^T$ ein Verschiebungsvektor ist, und \mathbf{R} eine orthonormale 3×3 Rotationsmatrix der Form

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad (3.24)$$

wobei die Elemente $r_{11} \cdots r_{33}$ die Eulerwinkel *roll*, *pitch*, *yaw* enthalten. Jede Drehung im 3D-Raum wird dabei durch eine Drehung um die Z -Achse (γ), gefolgt von einer Drehung um die neue Y -Achse (β) und einer abschließenden Drehung um die neue X -Achse (α) beschrieben. Die sechs Parameter dieser Transformation (drei Translationen t_x, t_y, t_z , und drei Rotationen α, β, γ) nennt man *externe Kameraparameter*, und werden bei der Kalibrierung ermittelt.

2. *Kamerakoordinaten* \mapsto *ideale Bildkoordinaten* Der nächste Schritt ist nun die Projektion des 3D Punktes gegeben im CCS in das Koordinatensystem der Bildebene ICS . Für das ideale, verzerrungsfreie Lochkameramodell ist dies eine *perspektivische Projektion*, und wird mit Hilfe des Strahlensatzes beschrieben durch

$$\frac{f}{z^c} = \frac{u}{x^c} = \frac{v}{y^c} \quad (3.25)$$

und wir erhalten für die Koordinaten in der Bildebene (u, v) :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{z^c} \begin{pmatrix} x^c \\ y^c \end{pmatrix} \quad (3.26)$$

Abbildung 3.10 zeigt diesen Sachverhalt schematisch. Die Brennweite f ist ein interner Kameraparameter und wird ebenfalls bei der Kalibrierung ermittelt. Für das telezentrische Kameramodell, welche eine parallele Projektion beinhaltet, erhalten wir:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x^c \\ y^c \end{pmatrix} \quad (3.27)$$

Wie man sieht, enthält die Gleichung die Brennweite nicht mehr, da sie für telezentrische Linsen nicht definiert ist. Desweiteren hat die Entfernung z zwischen Objekt und Kamera keinen Einfluss auf die Bildkoordinaten.

3. *Ideale Bildkoordinaten* \mapsto *verzerrte Bildkoordinaten* Das Lochkameramodell ist eine idealisierte Annäherung für Kameraprojektionen, und erlaubt eine einfache mathematische Modellierung der Beziehung zwischen Objekten einer Szene und ihren Bildkoordinaten. Reale optische Systeme, und Endoskope insbesondere, leiden jedoch zwangsläufig unter einer Vielzahl von Abbildungsfehlern (auch Verzerrungen oder Verzeichnungen genannt), und genügen nicht dem Modell. Die häufigsten Fehler werden verursacht durch die Bauart der Linsen, die ungenaue Positionierung des Bildsensors (CCD-Chip) und sogar durch Temperaturschwankungen und Vibrationen.

Auch die Abbildung einer 3D-Szene auf ein 2D-Bild (*perspektivische Projektion*) kann als eine Verzerrung betrachtet werden, zumal sie Entfernungen und Winkel

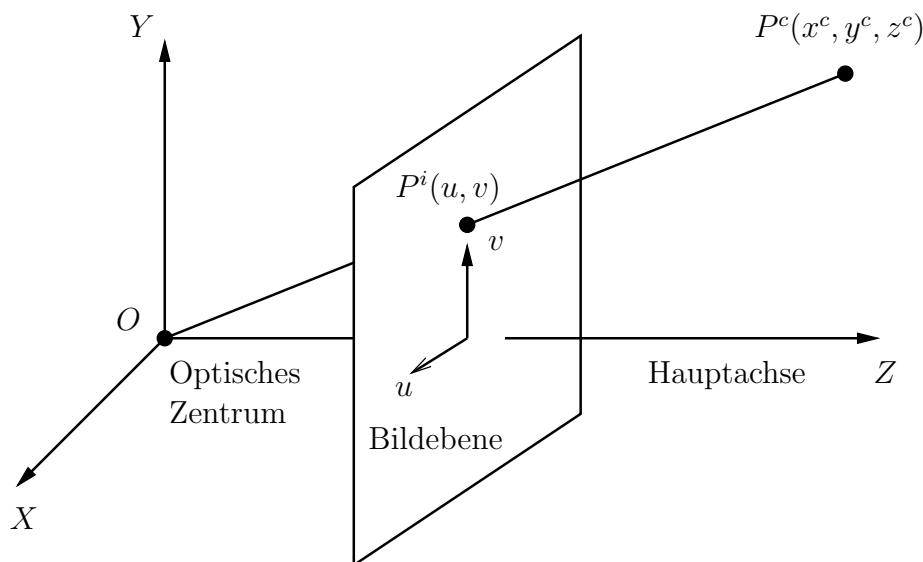


Abbildung 3.10: Perspektivische Projektion

nicht erhaltend abbildet. Diese Art von Verzerrung ist linear, weil sie mittels linearer (Matrix-) Algebra ausgedrückt werden kann. Doch es gibt auch nichtlineare geometrische Verzerrungen, die ebenfalls modelliert werden können. Sie sind jedoch ein rein zweidimensionales Problem, sodass keinerlei 3D-Information für ihre Behebung notwendig ist (Abbildung 3.11(a)). Man unterscheidet folgende Verzerrungen:

1. Radiale Verzerrung: verursacht durch fehlerhafte radiale Krümmung der Linse, wodurch ein Versatz der idealen Position vom Bildhauptpunkt nach innen oder außen auftritt.
2. Tangentiale Verzerrung: verursacht durch eine optische “Unwucht” (imperfekte Zentrierung), die bei fehlerhafter Fertigung auftritt, und kommt wegen der konzentrischen Schleiftechnik heutiger Linsen nur noch selten vor.
3. Dezentrierende Verzerrung: verursacht durch Fehlausrichtung der optischen Zentren der Linsen und des CCD-Sensors, sodass sie nicht kollinear sind.
4. Verzerrung des dünnen Prismas: verursacht durch die Abweichung von der idealen senkrechten Ausrichtung des CCD-Sensors zur optischen Achse.

Die Hauptverzerrung ist die Radialverzerrung, die je kleiner die Brennweite eines Objektivs (d.h. je größer der Öffnungswinkel), desto gravierender ausfällt. Dabei werden Punkte, die im Lochkameramodell auf eine Gerade auf der Bildebene abgebildet werden, auf eine gekrümmte Linie abgebildet. Das entsteht dadurch, dass die

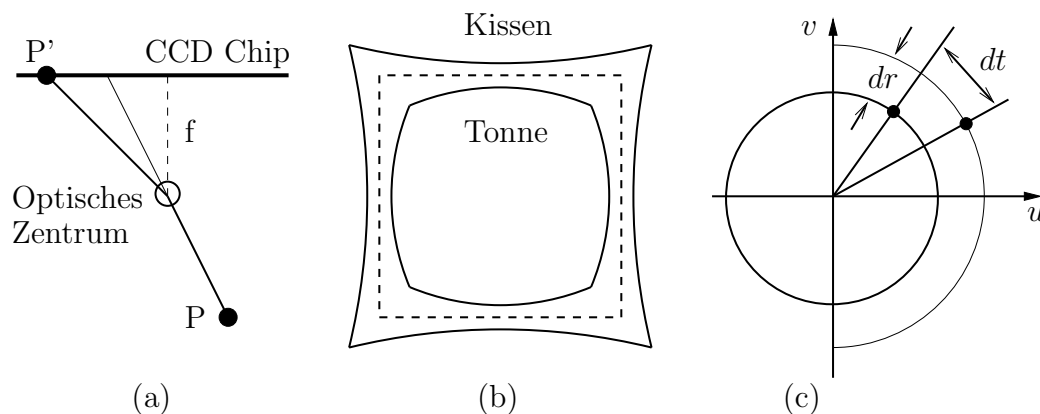


Abbildung 3.11: Schematische Darstellung von Bildverzerrungen

Linse Punkte in der Bildebene entweder zum optischen Zentrum hin, oder von ihm weg verschiebt. Die Verschiebung ist abhängig von der Entfernung zum Zentrum und ist meistens kreissymmetrisch. Werden die Punkte in Richtung Zentrum verschoben, so nennt man sie *tonnenförmig*, andernfalls *kissenförmig*. Die Namensgebung beruht auf die Art wie ein Rechteck abhängig von der Verzerrung abgebildet wird (Abbildung 3.11(b)). Wenn die Verzerrung auch noch eine tangentielle Komponente hat (Abbildung 3.11(c)), wird die Behebung des Abbildungsfehlers nochmal schwieriger.

Um solche Abbildungsfehler zu kompensieren wird das Lochkameramodell dahingehend erweitert, dass es auch die Verzerrungen berücksichtigt. Die Radialverzerrung kann als eine stetige Funktion beschrieben werden, und skaliert mit dem Abstand des Bildpunktes zum Zentrum der Verzerrung. Sie kann durch Polynome der Form $p = k_0x^0 + k_1x^1 + \dots + k_nx^n$ beliebig genau angenähert werden. Wenn (u, v) die vom Lochkameramodell gelieferten verzeichnungsfreien Bildkoordinaten sind, so modelliert man die korrespondierenden verzerrten Koordinaten (u^d, v^d) folgendermaßen:

$$\begin{pmatrix} u^d \\ v^d \end{pmatrix} = L(r) \begin{pmatrix} u \\ v \end{pmatrix}, \quad r^2 = u^2 + v^2 \quad (3.28)$$

$L(r)$ ist dabei nur für positive Werte definiert, und es gilt $L(0) = 1$. Eine Approximation für eine beliebige Funktion $L(r)$ kann durch eine Taylorreihe angegeben werden:

$$L(r) = 1 + \kappa_1 r + \kappa_2 r^2 + \kappa_3 r^3 + \kappa_4 r^4 + \dots \quad (3.29)$$

Zusammen mit dem Zentrum der Radialverzerrung gehören die Koeffizienten $\kappa_1, \dots, \kappa_n$ mit zu der intrinsischen Kamerakalibrierung. Weil die meisten Linsen rotationssymmetrisch geschliffen sind, müssen nur Polynome mit ausschließlich ge-

raden Potenzen berücksichtigt werden, wobei die Praxis zeigt, dass Polynome vierten Grades meistens ausreichen:

$$u^d = u (1 + \kappa_1 r^2 + \kappa_2 r^4) \quad (3.30)$$

$$v^d = v (1 + \kappa_1 r^2 + \kappa_2 r^4) \quad (3.31)$$

Sowohl Brown [85] als auch Tsai [86] weisen darauf hin, dass komplexere Modelle kaum Verbesserungen bringen und bei ungeeigneter Implementierung sogar numerische Instabilitäten auftreten können. Bricht man die Reihenentwicklung aus den Gleichungen 3.30 und 3.31 nach dem ersten Glied ab, so ergeben sich für die Umkehrungen geschlossene Lösungen:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{1 + \kappa(u^{d^2} + v^{d^2})} \begin{pmatrix} u^d \\ v^d \end{pmatrix} \quad (3.32)$$

$$\begin{pmatrix} u^d \\ v^d \end{pmatrix} = \frac{2}{1 + \sqrt{1 - 4\kappa(u^2 + v^2)}} \begin{pmatrix} u \\ v \end{pmatrix} \quad (3.33)$$

Der Parameter κ modelliert den Ausmaß der Radialverzerrung. Ist $\kappa < 0$, so bezeichnet man die Verzerrung *tonnenförmig*, und für $\kappa > 0$ heißt sie *kissenförmig*. Abhängig von der jeweiligen Anwendung gibt es zwei grundsätzliche Möglichkeiten, wie die Radialkorrektur in der Praxis angewendet werden kann:

- Das Kamerabild wird zunächst radial-korrigiert und alle nachfolgenden Berechnungen finden auf dem korrigierten Bild statt. Dadurch erscheinen Geraden auch wirklich als geraden im Bild, wodurch beispielsweise das modellbasierte Segmentieren von Objekten erleichtert wird.
- Etwaige Berechnungen, wie beispielsweise das Segmentieren sehr feiner Strukturen, die bei der Entzerrung verwischt werden könnten, finden vor der Korrektur statt. Anschließend können die segmentierten Bildkoordinaten zur weiteren Verarbeitung radial-korrigiert werden.

4. *Verzerrte Bildkoordinaten* \mapsto *Pixelkoordinaten* Nun können die Punkte $(u^d, v^d)^T$ aus der Bildebene in das gerätespezifische Pixelkoordinatensystem transformiert werden.

Abbildung 3.12 zeigt schematisch eine Sensorzelle des CCD-Chips. Dabei sind S_x und S_y die horizontalen und vertikalen Abmessungen eines Pixels, γ ist der *Schrägeparameter* (engl. *skew factor*), der unter normalen Umständen 0 ist, kann aber auch Werte $\gamma \neq 0$ annehmen, wenn die x -Achse auf dem CCD-Sensor nicht genau rechtwinklig zur y -Achse ist.

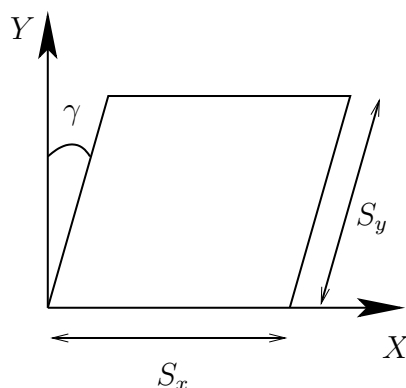


Abbildung 3.12: Geometrisches Modell eines Pixels

Geht man jedoch von rechteckigen Pixel aus, was meistens der Fall ist, so erhalten wir:

$$\begin{pmatrix} r \\ c \end{pmatrix} = \begin{pmatrix} v^d/S_y + C_y \\ u^d/S_x + C_x \end{pmatrix} \quad (3.34)$$

Dabei kann man S_x und S_y auch als Skalierungsfaktoren betrachten. Beim Lochkameramodell repräsentieren sie die Höhe und die Breite der Sensorzelle auf dem CCD-Chip. Der Punkt $(C_x, C_y)^T$ ist der Bildmittelpunkt. Er ist eine senkrechte Projektion des optischen Zentrums auf die Bildebene, also der Punkt im Bild, von welchem ein Strahl durch das optische Zentrum senkrecht auf die Bildebene ist. Er markiert ausserdem den Ursprung der Radialverzerrung. Die sechs Parameter $(f, \kappa, S_x, S_y, C_x, C_y)$ nennt man *intrinsische Kameraparameter*, und beschreiben die Projektion von 3D nach 2D durch die Kamera.

Um Informationen, wie Entfernungen zwischen Objekten, ihre Form und Farbe, aus dem Bild einer Szene gewinnen zu können, ist eine Kalibrierung der Kamera notwendig. Kamerakalibrierung ist ein Prozess, bei der Kenntnis über die Relation zwischen dem Bild einer Szene und die Szene selbst erlangt wird, einschließlich der Position der Kamera in der Szene und ihrer internen Charakteristika.

Außer den geometrischen Verzerrungen gibt es noch radiometrische Verzerrungen, welche diverse Verfälschungen im sichtbaren Bereich des Lichts charakterisieren. Eine vollständige Kamerakalibrierung beinhaltet sowohl die geometrische, als auch die radiometrische Kalibrierung. Die radiometrische Kalibrierung [87] wird durchgeführt um zu erfahren, wie die Kamera die Helligkeit, Farbe und Leuchtdichte einer Szene beeinflusst. Die geometrische Kalibrierung ermittelt Position und Orientierung der Kamera (extrinsische Kameraparameter) und interne Parameter (intrinsische Parameter) wie Brennweite, Bildhauptpunkt, Scherung und Eigenschaften des CCD-Sensors.

3.5 Kamerakalibrierung

3.5.1 Mathematische Grundlagen

Wie wir gesehen haben, sind für die Abbildung der Punkte einer 3D-Szene auf Pixelkoordinaten durch eine Kamera eine Reihe von Transformationen notwendig. Wünschenswert ist hierbei die einzelnen Stufen der Kette von Transformationen in eine einheitliche mathematische Form zu bringen. Wie schon bereits eingeführt sei der Punkt P im Weltkoordinatensystem WCS gegeben durch $p^w = (x^w, y^w, z^w)^T$ und seine Entsprechung im Kamerakoordinatensystem CCS $p^c = (x^c, y^c, z^c)^T$. Die durch die Gleichung 3.23 angegebene Transformation kann auch mittels einer homogenen Transformationsmatrix angegeben werden, welche aus Rotationsmatrix und Translationsvektor gebildet werden kann:

$$\begin{pmatrix} x^c \\ y^c \\ z^c \\ 1 \end{pmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} x^w \\ y^w \\ z^w \\ 1 \end{pmatrix} \quad (3.35)$$

oder durch die kompaktere Schreibweise:

$$\begin{pmatrix} x^c \\ y^c \\ z^c \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0^T & 1 \end{bmatrix} \cdot \begin{pmatrix} x^w \\ y^w \\ z^w \\ 1 \end{pmatrix} \quad (3.36)$$

beziehungsweise

$$\mathbf{p}^c = {}^c\mathbf{H}_w \cdot \mathbf{p}^w \quad (3.37)$$

Gleichung 3.26 lässt sich ebenfalls in eine Form homogener Koordinaten bringen, denn in homogenen Koordinaten repräsentieren alle Vektoren $(wu, wv, w)^T$ mit $w \neq 0$ den selben Punkt $(u, v)^T$. Die inhomogenen Koordinaten ergeben sich durch Normierung des Vektors auf $w = 1$. Ein weiterer Vorteil homogener Koordinaten ist, dass sich damit auch Punkte im Unendlichen darstellen lassen. Der Vektor $(u, v, 0)^T$ steht für denjenigen Punkt, der in Richtung $(u, v)^T$ im Unendlichen liegt. In homogenen Koordinaten gilt für die Überführung in die Bildebene bis auf einen Skalierungsfaktor:

$$\begin{pmatrix} U \\ V \\ S \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{pmatrix} x^c \\ y^c \\ z^c \\ 1 \end{pmatrix}, \quad u = U/S, v = V/S, S \neq 0 \quad (3.38)$$

Bei einem gegebenen Vektor $\mathbf{x} = (x_1, x_2, \dots)^T$ bezeichnen wir mit $\tilde{\mathbf{x}}$ den um das Element 1 erweiterte Vektor. Sei \mathbf{P} die 3×4 Matrix

$$\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (3.39)$$

welche wir *perspektivische Projektionsmatrix* nennen werden. Die Transformation des 3D-Punktes $p^c = (x^c, y^c, z^c)^T$ in den Bildpunkt $p^i = (u, v)^T$ kann nun in der Matrixschreibweise geschrieben werden

$$s\tilde{p}^i = \mathbf{P}\tilde{p}^c, \quad (3.40)$$

wobei $s = S$ ein von 0 verschiedener Skalar ist.

Berücksichtigt man jedoch den Umstand, dass Pixelkoordinaten von einer Ecke des Bildes gemessen werden, fällt der Hauptpunkt nicht mehr mit dem Ursprung des Bildkoordinatensystems zusammen, sondern befindet sich an der Stelle $(C_x, C_y)^T$. Außerdem muss noch berücksichtigt werden, dass die Pixelgrößen in horizontaler und vertikaler Richtung unterschiedlich sein können, und dass die Achsen nicht notwendigerweise orthogonal sind (siehe Abbildung 3.12). Das Bildkoordinatensystem (O, u, v) , mit dem Hauptpunkt als Ursprung, hat in beide Richtungen der u - und v -Achse die gleiche Einheit. Das Pixelkoordinatensystem $(0, r, c)$, üblicherweise mit dem Ursprung in der linken oberen Ecke, hat bedingt durch nicht rechteckige Pixel, verschiedene Einheiten für beide Achsen. Für einen gegebenen Punkt $p^i = (u, v)^T$ kann die Transformation in das Pixel $p^p = (r, c)^T$ des gerätespezifischen Pixelkoordinatensystems angegeben werden als

$$\tilde{p}^p = \mathbf{K}\tilde{p}^i, \quad (3.41)$$

wobei

$$\mathbf{K} = \begin{bmatrix} S_x & S_x \cot \gamma & C_x \\ 0 & S_y / \sin \gamma & C_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3.42)$$

Laut Gleichungen 3.40 und 3.41 gilt

$$s\tilde{p}^p = \mathbf{K}\mathbf{P}\tilde{p}^c, \quad (3.43)$$

und es ergibt sich die neue Kameramatrix

$$\mathbf{M} = \begin{bmatrix} fS_x & fS_x \cot \gamma & C_x & 0 \\ 0 & fS_y / \sin \gamma & C_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3.44)$$

Mit Hilfe der Matrix cH_w aus Gleichung 3.37 und der Kameramatrix aus 3.44 kann nun die Transformation zwischen Welt- und Pixelkoordinatensystem in homogenen Koordinaten geschrieben werden als

$$\begin{pmatrix} x^p \\ y^p \\ z^p \end{pmatrix} = \mathbf{M} {}^cH_w \begin{pmatrix} x^w \\ y^w \\ z^w \\ 1 \end{pmatrix} \quad (3.45)$$

Die Kalibrierung besteht nun aus der Berechnung dieser Matrizen, beziehungsweise, wenn man sie zusammenfasst, aus Berechnung und entsprechender Dekomposition. Wegen des hohen Bedarfs sowohl in der klassischen Bildverarbeitung als auch in der Robotik wurde die Thematik der exakten und effizienten Kamerakalibrierung umfangreich behandelt. Alle bekannten Verfahren zur Kamerakalibrierung stützen sich auf die Korrespondenzen zwischen bekannten Raumpunkten einer Szene und ihren beobachteten Koordinaten im aufgenommenen Bild. Dabei wird die Kalibrierung als ein Minimierungsproblem formuliert, welches die Parameter der Kamera optimiert. Die nachfolgenden Abschnitte stellen einige der bekanntesten Verfahren vor. Die Besonderheiten bei der Kalibrierung eines Stereokamerasystems werden in einem späteren Kapitel behandelt.

3.5.2 Kalibriermethoden

Direkte lineare Methode

Eine der ersten Verfahren zur Kamerakalibrierung wurde 1971 von Abdel-Aziz und Kara [88] vorgestellt. Ihrer DLT-Methode (direct linear transformation) liegt das einfache, verzerrungsfreie Lochkameramodell zugrunde. Die Projektionsmatrix einer Kamera wird anhand der Korrespondenzen von Bild- und Szenenpunkten nicht-iterativ ermittelt. Dabei werden lineare Gleichungssysteme aufgestellt, sodass die Lösung mit Hilfe von Singulärwertzerlegung in geschlossener Form angegeben werden kann. Die Vorteile durch Einfachheit und Effizienz werden durch das Fehlen eines Verzerrungsmodells und durch vergleichsweise ungenauen Kalibrierergebnissen zunichtegemacht. Die Methode erfordert einen präzise gefertigten dreidimensionalen Kalibrierkörper, wobei die Kalibrierpunkte idealerweise ein Volumen aufspannen sollten, das gleichzeitig auch dem Arbeitsraum entsprechen sollte.

Das Verfahren von Tsai

Um die Problematik volumetrischer Kalibrierkörper zu umgehen, entwickelte Tsai [89, 86] ein Verfahren, welches mit einer, oder zwei orthogonal angeordneten planaren Ebenen als Kalibrierkörper auskommt. Die Ebenen enthalten leicht zu extrahierende

Merkmale in Form eines Schachbrettmusters, und können mit heutiger Laserdrucktechnik in ausreichender Genauigkeit gefertigt werden. Für die Linsenverzerrung wird ein rein radiales Modell verwendet, bei dem der Hauptpunkt unverändert bleibt und alle anderen Punkte entlang der Strahlen wandern, die sie mit dem Hauptpunkt verbinden. Es wird nur der quadratische Term der Verzerrung geschätzt in der Annahme, dass er den größten Teil der Verzerrung erklärt. Die meisten Parameter des Modells können in geschlossener Form angegeben werden, andere werden iterativ ermittelt. Das Verfahren macht sich das sogenannte “Radial Alignment Constraint” (RAC) zunutze. Es beschreibt formal, dass die drei Vektoren p^i , p^d und p^c linear abhängig sind, und zwar von den Kameraparametern.

Die Methode von Zhang

Ebenfalls Kalibriermuster auf einer ebenen Fläche verwendet das Verfahren von Zhang [90]. Die Technik erweist sich in der Praxis als stabiler und flexibler im Vergleich mit den zuvor beschriebenen. Hier werden auch leicht zu extrahierende (quadratische) Muster auf ein planares Objekt gedruckt und Aufnahmen aus mehreren beliebigen Orientierungen gemacht. Für die Kameraparameter gibt es zunächst eine Lösung in geschlossener Form, gefolgt von einer nichtlinearen Optimierung. Obwohl ursprünglich nur die Radialverzerrung berücksichtigt wurde, bezieht die *OpenCV*³ Implementierung auch nicht rein radiale Komponenten mit ein, welche bei nicht-paralleler Ausrichtung zwischen Linse und CCD-Chip auftreten. Eine Besonderheit des Verfahrens ist, dass die internen Parameter der Kamera in Pixelkoordinaten bestimmt werden. Dadurch muss man die physikalischen Parameter des Bildsensors nicht mehr ermitteln.

Selbstkalibrierung nach Faugeras

Einen völlig neuen Ansatz verfolgt Faugeras in seiner wegweisenden Arbeit [91] bei der Kalibrierung von Stereokamerasystemen. Seine Methode verzichtet gänzlich auf Kalibrierkörper jeglicher Art, die Kameraparameter werden anhand von Punktkorrespondenzen in beliebig aufgenommenen Bildern ermittelt. Der Hauptnachteil dieser Methode ist ihre Instabilität, bedingt durch die Tatsache, dass alle Parameter zunächst unbekannt sind. Ist zumindest einer der Parameter bekannt, so können die anderen leichter geschätzt werden. Diese Probleme treten nicht auf, wenn das Kamerasystem in Bewegung ist, und diese Bewegung bekannt ist.

3.5.3 Kalibrierergebnisse

Wir stellen die Ergebnisse der Kalibrierung von zwei medizinischen Stereoendoskopen vor, die während der Entstehung dieser Arbeit bei der Durchführung von

³Open Computer Vision Library

Experimenten zur Verfügung standen. Zum einen handelt es sich um ein Endoskop der Firma Richard Wolf GmbH [92], und zum anderen um das Stereoendoskop des *da Vinci*-Systems, gefertigt von der Firma Schölly Fiberoptic GmbH [93]. Es wurden mehrere Kalibrierungen mit Hilfe zweier unterschiedlicher Toolkits (*HALCON* und *MATLAB*) durchgeführt, die Ergebnisse können den Tabellen aus den Abbildungen 3.13 und 3.14 entnommen werden. Dabei liegt beiden Paketen die Kalibriermethode von Zhang zugrunde.

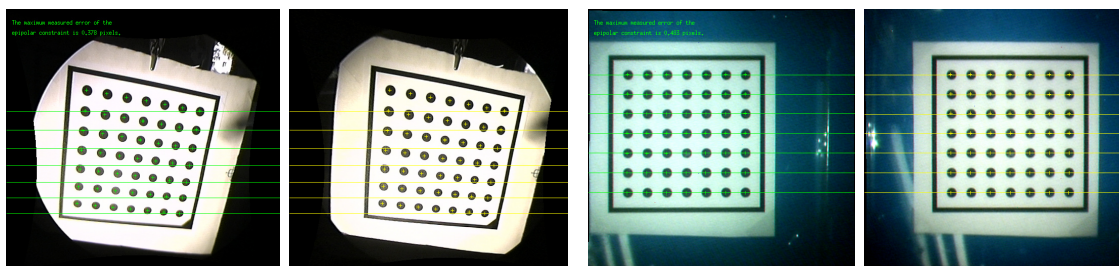
MATLAB			HALCON		
Parameter	<i>links</i>	<i>rechts</i>	Parameter	<i>links</i>	<i>rechts</i>
f_x	447.10785	444.04779	f	0.00337668	0.00334208
f_y	445.59725	441.79863	S_x	7.43272e-06	7.41861e-06
C_x	261.59271	303.96516	S_y	7.4e-06	7.4e-06
C_y	281.13238	278.25671	C_x	264.451	309.995
κ_1	-0.37794	-0.38591	C_y	285.153	276.26
κ_1	0.15735	0.17126	κ	-39342.2	-40324.8

Abbildung 3.13: Ergebnisse der Kalibrierung für das R. Wolf-Endoskop

MATLAB			HALCON		
Parameter	<i>links</i>	<i>rechts</i>	Parameter	<i>links</i>	<i>rechts</i>
f_x	1170.58439	1161.18908	f	0.00872584	0.0087306
f_y	1183.95785	1167.02620	S_x	7.4153e-06	7.34975e-06
C_x	216.79494	282.34498	S_y	7.4e-06	7.4e-06
C_y	292.23064	271.84053	C_x	138.045	272.292
κ_1	-0.11506	-0.21060	C_y	249.245	201.239
κ_1	0.43055	3.09866	κ	-1920.95	-1169.79

Abbildung 3.14: Ergebnisse der Kalibrierung für das Schölly-Endoskop

Abbildung 3.15 visualisiert jeweils den verbleibenden vertikalen Ausrichtungsfehler, auch Epipolarfehler genannt (R. Wolf 0.378 Pixel, Schölly(IS) 0.483 Pixel). Im Abschnitt 6.2.2 des Kapitels “3D-Rekonstruktion des Situs” werden die Parameter im Kontext der Tiefengenaugigkeit nochmal gegenübergestellt.



(a) 3D-Endoskop von der Richard Wolf GmbH

(b) 3D-Endoskop des *da Vinci* (Schöllly)

Abbildung 3.15: Epipolarfehler nach Kalibrierung (Rektifizierung)

3.6 Hand–Auge–Kalibrierung

Bei einem komplexen Aufbau bestehend aus Roboterarmen und Kameras ist außer der Kalibrierung der Robotergelenke und der Kameras auch noch das Problem der unbekannt relativen Lage aller Komponenten zueinander zu lösen. Dies ist deshalb wichtig, damit die externen Kameraparameter nicht immer wieder neu kalibriert werden müssen, wenn sich Roboter und Kamera bewegen. Die Ermittlung dieser relativen Lage bezeichnet man in der Literatur als Hand–Auge–Kalibrierung. Dadurch dass das Endoskop an einem der Roboterarme starr befestigt ist, und seine Position relativ zum Flansch nun bekannt ist, sind beliebige Kamerafahrten durchführbar. Die Möglichkeit, zuvor erstellte 3D–Modelle zu manipulieren, oder sie zu erweitern bleibt somit erhalten.

Eine weit verbreitete Methode zur Hand–Auge–Kalibrierung verwendet analog zur Kamerakalibrierung planare Kalibriermuster, welche vom Roboter in verschiedene Lagen positioniert und von der Kamera aufgenommen werden. Diese Technik ist beim ARAMIS²–Aufbau aus verschiedenen Gründen nicht praktikabel. Zum einen wäre die Fertigung eines entsprechenden Kalibrierkörpers schwierig, der von den chirurgischen Instrumenten gegriffen und positioniert werden kann. Zum anderen ist diese Methode nur für Chirurgierobotersysteme geeignet, die eine Vor– und Rückwärtsrechnung mit Zugriffsmöglichkeit auf Gelenkwinkel und kartesische Koordinaten mit hoher Genauigkeit ermöglichen. Dies ist beim ARAMIS²–System zwar möglich, doch kommerziell erhältlichen Systemen fehlt diese Eigenschaft.

Auch die kinematische Integration der chirurgischen Instrumente in das Robotersystem ist wegen der Nachgiebigkeit und dem nicht vollständig reproduzierbaren Verhalten der Instrumentenservos mit Ungenauigkeiten behaftet. Die Hand–Auge–Kalibrierung der Roboterarme und des Stereoendoskops und die anschließende Vorwärtsrechnung anhand der technischen Daten der chirurgischen Instrumente ist aus diesem Grund nicht praktikabel. Im Folgenden wird eine einfache Methode vorgestellt, die es mit Hilfe eines 3D–Kalibrierkörpers ermöglicht, sowohl die chirurgischen Instrumente, als auch die von Stereoendoskop gelieferten Rekonstruktionsdaten in einem gemeinsamen Koordinatensystem darzustellen.

3.6.1 Von Kamera– nach Weltkoordinaten

Bei dem Kalibrierkörper handelt es sich um pyramidenartig angeordnete Balken, welche aus einem Aluminiumblock mittels CNC–Technik mit hoher Genauigkeit ausgefräst wurden. Abbildung 3.16(a) zeigt eine Originalaufnahme durch das Endoskop. Für die Kalibrierung des Gesamtsystems muss der Kalibrierkörper im Arbeitsraum der Instrumente und im Sichtfeld des Endoskops positioniert werden. Anschließend fährt der Bediener mit jedem der Instrumentenarme bestimmte Punkte des Kalibrierkörpers an. Aus den bekannten 3D–Weltkoordinaten dieser Punkte, und aus den zu berechnenden 3D–Kamerakoordinaten werden die für die Umrechnung notwendigen Transformationen automatisch ermittelt. Der Grundgedanke des Verfahrens ist die Bereitstellung eines gemeinsamen metrischen Weltkoordinatensystems im Arbeitsraum, und entsprechende Transformationen für die beteiligten Komponenten. Ein weiterer Vorteil der vorgestellten Methode ist die anschließende Verfügbarkeit von genauen metrischen Daten, welche sowohl zur Kollisionserkennung, als auch zur Vermessung des Operationssitus herangezogen werden können.

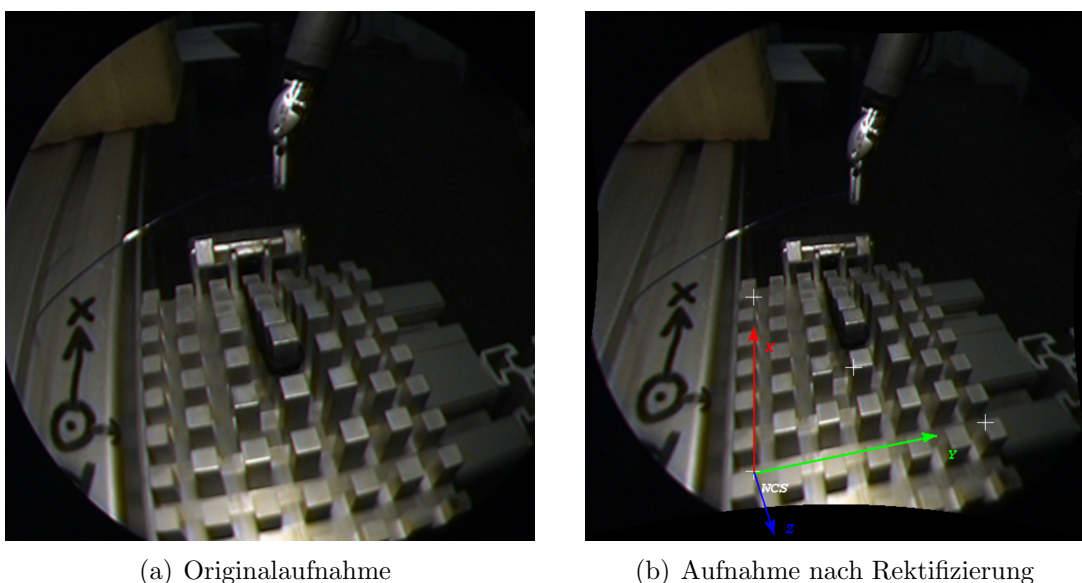


Abbildung 3.16: 3D–Kalibrierkörper

Das gemeinsame kartesische Weltkoordinatensystem wird wie in Abbildung 3.16(b) festgelegt, sodass ein rechtshändiges System entsteht. Die Stereorekonstruktion liefert ihre Ergebnisse im Koordinatensystem der linken rektifizierten Kamera. Zur Berechnung einer homogenen Transformationsmatrix, welche die Überführung aus dem linken rektifizierten Kamerakoordinatensystem in das festgelegte Weltkoordinatensystem ermöglicht, sind drei Punkte des Kalibrierkörpers notwendig. Wir werden die Spalten der Transformationsmatrix anhand normierter Vektoren angeben, welche mit Hilfe einfacher Vektoroperationen berechnet werden.

Dafür bezeichnen wir mit $(O_x, O_y, O_z)^T$ den Ursprung, mit $(X_x, X_y, X_z)^T$ einen Punkt auf der X -Achse und mit $(XY_x, XY_y, XY_z)^T$ einen Punkt in der XY -Ebene. Zunächst müssen die Koordinaten dieser Punkte im Koordinatensystem der linken Kamera ermittelt werden. Bei der Rekonstruktion werden korrespondierende Punkte aus rechtem und linkem Bild manuell ausgewählt. Seien \vec{X} , \vec{Y} , \vec{Z} drei Vektoren, und \vec{X}^n , \vec{Y}^n , \vec{Z}^n ihre Normalen. Wir berechnen zunächst \vec{X} und \vec{X}^n wie folgt:

$$\vec{X} = \begin{pmatrix} X_x - O_x \\ X_y - O_y \\ X_z - O_z \end{pmatrix}, \quad \vec{X}^n = \frac{\vec{X}}{\|\vec{X}\|} \quad (3.46)$$

und sei \vec{V} gegeben durch

$$\vec{V} = \begin{pmatrix} XY_x - O_x \\ XY_y - O_y \\ XY_z - O_z \end{pmatrix} \quad (3.47)$$

Um \vec{Z} zu erhalten, bilden wir das Kreuzprodukt von \vec{X}^n und \vec{V} :

$$\vec{Z} = \vec{X}^n \times \vec{V}, \quad \vec{Z}^n = \frac{\vec{Z}}{\|\vec{Z}\|} \quad (3.48)$$

\vec{Y}^n lässt sich nun berechnen durch

$$\vec{Y}^n = \vec{Z}^n \times \vec{X}^n \quad (3.49)$$

Wir bilden nun aus den entsprechenden Elementen der Vektoren \vec{X}^n , \vec{Y}^n , \vec{Z}^n die Matrix M :

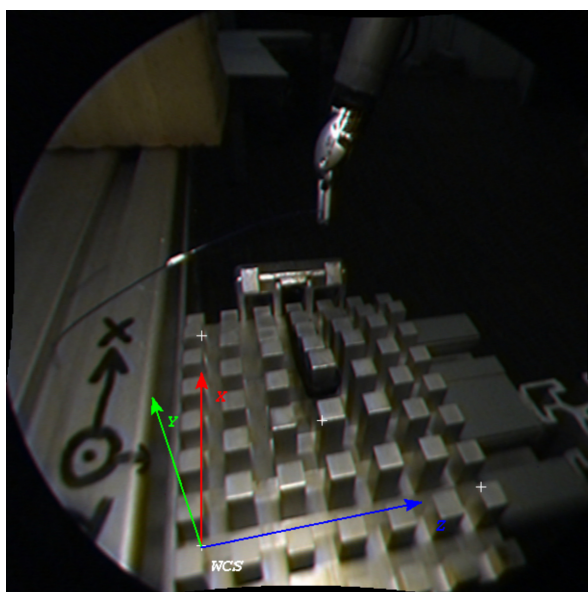
$$M = \begin{pmatrix} \vec{X}_x^n & \vec{Y}_x^n & \vec{Z}_x^n & O_x \\ \vec{X}_y^n & \vec{Y}_y^n & \vec{Z}_y^n & O_y \\ \vec{X}_z^n & \vec{Y}_z^n & \vec{Z}_z^n & O_z \end{pmatrix} \quad (3.50)$$

Nun können Transformationen aus dem Kamerakoordinatensystem in das festgelegte Weltkoordinatensystem mittels ${}^{lrc}\mathbf{H}_w$ erfolgen:

$${}^{lrc}\mathbf{H}_w = M^{-1} \quad (3.51)$$

Um die tendenziell gleiche Ausrichtung der Achsen, wie die vom Robotersystem zu bekommen, rotieren wir noch das erhaltene Koordinatensystem um -90° um die X -Achse. Die Abweichungen und die ebenfalls unbekannt verschobene Ursprünge werden in einem nächsten Schritt berechnet. Abbildung 3.17 zeigt die endgültige Lage des gemeinsamen Weltkoordinatensystems. Die Tabelle enthält die Koordinaten

der drei Eckpunkte sowie einen zusätzlichen Punkt zur Kontrolle. Die Abweichungen (Fehler) zwischen realen Werten auf dem Kalibrierkörper und transformierten Koordinaten sind der letzten Spalte zu entnehmen.



Kamera	Welt [mm]		
	<i>WCS</i>	<i>REAL</i>	<i>ERR</i>
-2.97	17.81	18.0	0.19
26.82	12.30	12.0	0.3
104.20	29.29	30.0	0.71
-28.19	0.0	0.0	0.0
50.51	0.0	0.0	0.0
92.83	0.0	0.0	0.0
-42.93	65.69	66.0	0.31
8.74	0.0	0.0	0.0
141.35	0.0	0.0	0.0
35.73	0.32	0.0	0.32
43.68	0.0	0.0	0.0
106.79	65.78	66.0	0.22
			0.17

Abbildung 3.17: Kalibrierkörper und Weltkoordinatensystem

Roboterkoordinatensystem und Weltkoordinatensystem weichen nun durch eine geringe Transformation voneinander ab. Eine einfache und elegante Methode die einzelnen Rotationen der Achsen und die Translation zwischen den Ursprüngen zu ermitteln bietet der Richtungskosinus. Wir stellen zunächst dieses mathematische Werkzeug vor, um anschließend die Transformation aus Weltkoordinatensystem ins Roboterkoordinatensystem anzugeben.

3.6.2 Richtungskosinus

Wie bereits erwähnt, müssen Koordinatensysteme in komplexen Modellen nicht notwendigerweise rechtwinklig zueinander angeordnet sein. Oft ist es der Fall, dass (beispielsweise wegen Konstruktionsungenauigkeiten) Koordinatensysteme durch eine geringe aber unbekannt Transformation (Translation und Rotation) voneinander abweichen, auch wenn sie theoretisch rechtwinklig zueinander sein müssten.

Abbildung 3.18(a) zeigt zwei kartesische Koordinatensysteme mit unterschiedlichem Ursprung und unterschiedlicher Orientierung. Die Aufgabe besteht nun darin, eine Transformation zu finden, mit deren Hilfe die zwei Koordinatensysteme in Deckung gebracht werden können. Dafür stellen wir den Richtungskosinus als weiteres mathematisches Werkzeug vor. Demnach beschreiben wir eine bestimmte Richtung \tilde{r}

im Raum als Kosinuse der Winkel, die dieser Richtungsvektor mit den Achsen X ($l_{\tilde{r}x} = \cos \theta_{\tilde{r}x}$), Y ($l_{\tilde{r}y} = \cos \theta_{\tilde{r}y}$) und Z ($l_{\tilde{r}z} = \cos \theta_{\tilde{r}z}$) des Koordinatensystems bilden. Abbildung 3.18(b) zeigt \tilde{r} und die Winkel $\theta_{\tilde{r}x}$, $\theta_{\tilde{r}y}$ und $\theta_{\tilde{r}z}$.

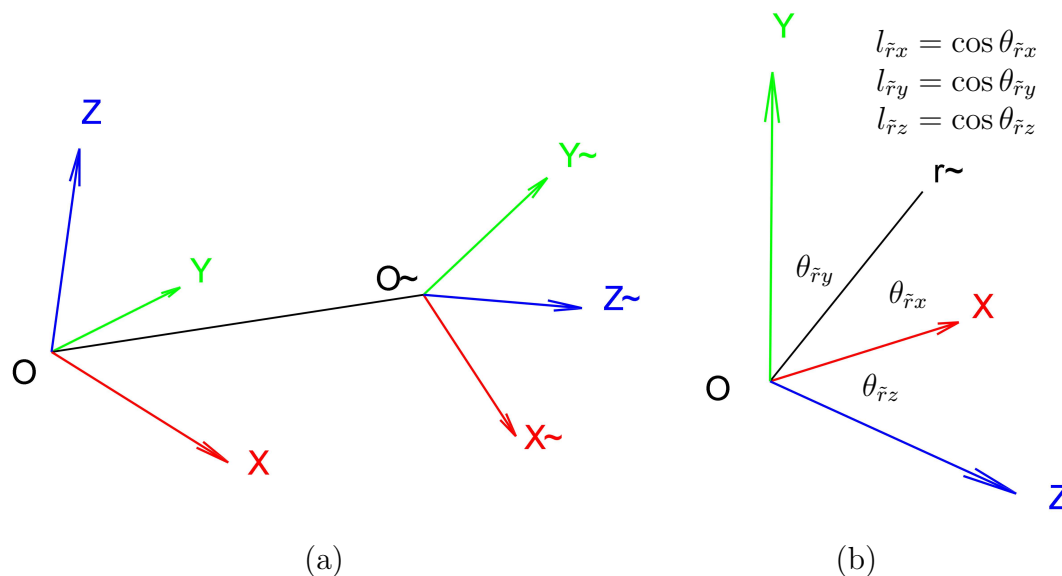


Abbildung 3.18: Kartesische Koordinatensysteme und Richtungskosinus

Wenn wir nun eine Koordinatentransformation mit Hilfe dieser Größen beschreiben, so geben wir jede Achse des neuen Koordinatensystems (z.B. \tilde{Z}) mit Hilfe des Richtungskosinus bezogen auf die ursprüngliche Achse an. Die Transformation zwischen Koordinatensystemen kann in diesem Fall als Matrixform geschrieben werden als:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} l_{\tilde{x}x} & l_{\tilde{x}y} & l_{\tilde{x}z} \\ l_{\tilde{y}x} & l_{\tilde{y}y} & l_{\tilde{y}z} \\ l_{\tilde{z}x} & l_{\tilde{z}y} & l_{\tilde{z}z} \end{pmatrix} \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \quad (3.52)$$

Dabei ist der Ursprung des (X, Y, Z) -Koordinatensystems $O = (0, 0, 0)$, und die des $(\tilde{X}, \tilde{Y}, \tilde{Z})$ -Systems $\tilde{O} = (x_0, y_0, z_0)$. Die Elemente der Rotationsmatrix sind jeweils ein Richtungskosinus für die entsprechenden Achsen. Man beachte, dass nur sechs Elemente erforderlich sind um die Matrix anzugeben, die verbleibenden drei können mit folgendem Gleichungssystem berechnet werden:

$$\begin{aligned} l_{\tilde{x}x}^2 + l_{\tilde{y}x}^2 + l_{\tilde{z}x}^2 &= 1 \\ l_{\tilde{x}y}^2 + l_{\tilde{y}y}^2 + l_{\tilde{z}y}^2 &= 1 \\ l_{\tilde{x}z}^2 + l_{\tilde{y}z}^2 + l_{\tilde{z}z}^2 &= 1 \end{aligned} \quad (3.53)$$

3.6.3 Von Welt– nach Roboterkoordinaten

Für die Anwendung der Methode des Richtungskosinus sind die Eckpunkte der beiden beteiligten Koordinatensysteme notwendig. Die Eckpunkte des Koordinatensystems des Kalibrierkörpers sind aus den CAD–Daten bekannt. Diese Punkte werden nun mit den Instrumenten angesteuert, und ihre Koordinaten aufgezeichnet. Der Ursprung und die Punkte auf der X – und Z –Achse (Abb. 3.17) können direkt angesteuert werden, der Punkt auf der Y –Achse muss anhand der anderen berechnet werden. Dieser befindet sich auf einer Senkrechten durch den Ursprung, für ihre Berechnung ermitteln wir die implizite Ebenengleichung der XZ –Ebene und das Lot darauf. Die explizite Darstellung einer Ebene im 3D–Raum kann mittels den drei nicht–kollinearen Punkten (O_x^r, O_y^r, O_z^r) (Ursprung), (X_x^r, X_y^r, X_z^r) (Punkt auf X –Achse) und (Z_x^r, Z_y^r, Z_z^r) (Punkt auf Z –Achse) angegeben werden. Für die Berechnung des Lots ist jedoch die implizite Ebenengleichung notwendig:

$$ax + by + cz + d = 0. \quad (3.54)$$

Die Koeffizienten a, b, c, d können nun folgendermaßen berechnet werden:

$$\begin{aligned} a &= (X_y^r - O_y^r) \cdot (Z_z^r - O_z^r) - (X_z^r - O_z^r) \cdot (Z_y^r - O_y^r) \\ b &= (X_z^r - O_z^r) \cdot (Z_x^r - O_x^r) - (X_x^r - O_x^r) \cdot (Z_z^r - O_z^r) \\ c &= (X_x^r - O_x^r) \cdot (Z_y^r - O_y^r) - (X_y^r - O_y^r) \cdot (Z_x^r - O_x^r) \\ d &= -X_x^r \cdot a - X_y^r \cdot b - X_z^r \cdot c \end{aligned}$$

Mit den normierten Koeffizienten

$$a^n = \frac{a}{n}, \quad b^n = \frac{b}{n}, \quad c^n = \frac{c}{n}, \quad n = \sqrt{a^2 + b^2 + c^2} \quad (3.55)$$

und den Abmessungen der Achsen des Kalibrierkörpers d_x, d_y, d_z erhalten wir die Koordinaten des Punktes auf der Y –Achse:

$$\begin{aligned} Y_x^r &= O_x^r + d_x \cdot a \\ Y_y^r &= O_y^r + d_y \cdot b \\ Y_z^r &= O_z^r + d_z \cdot c \end{aligned}$$

Nun können die einzelnen Winkel zwischen den Achsen der Koordinatensysteme berechnet werden. Der Winkel θ zwischen zwei Vektoren $\vec{P}_1 = (x_1, y_1, z_1)^T$ und $\vec{P}_2 = (x_2, y_2, z_2)^T$ mit bekannten kartesischen Koordinaten lässt sich aus dem Zusammenhang

$$\vec{P}_1 \cdot \vec{P}_2 = \text{norm}(\vec{P}_1) \text{norm}(\vec{P}_2) \cos \angle(\vec{P}_1, \vec{P}_2) \quad (3.56)$$

folgendermaßen berechnen:

$$\theta = \arccos \frac{d}{v_1 \cdot v_2}, \quad (3.57)$$

wobei d das Skalarprodukt und v_1, v_2 jeweils die normalisierten Vektoren sind. Wir geben eine analytisch geschlossene Form für die Werte $l_{\tilde{x}x}$, $l_{\tilde{x}y}$ und $l_{\tilde{x}z}$ an, die Berechnung der anderen Werte der Matrix aus Gleichung 3.52 erfolgt analog:

$$\begin{aligned} l_{\tilde{x}x} &= \frac{(X_x^r - O_x^r)}{\sqrt{(O_x^r - X_x^r)^2 + (O_y^r - X_y^r)^2 + (O_z^r - X_z^r)^2}} \\ l_{\tilde{x}y} &= \frac{(X_y^r - O_y^r)}{\sqrt{(O_x^r - X_x^r)^2 + (O_y^r - X_y^r)^2 + (O_z^r - X_z^r)^2}} \\ l_{\tilde{x}z} &= \frac{(X_z^r - O_z^r)}{\sqrt{(O_x^r - X_x^r)^2 + (O_y^r - X_y^r)^2 + (O_z^r - X_z^r)^2}} \end{aligned}$$

Bei gegebenen Koordinaten (x^w, y^w, z^w) in gemeinsamen Weltkoordinatensystem lassen sich die entsprechenden Roboterkoordinaten (x^r, y^r, z^r) wie folgt berechnen:

$$\begin{aligned} x^r &= (O_x^r + x^w) \cos l_{\tilde{x}x} + (O_y^r + y^w) \cos l_{\tilde{y}x} + (O_z^r + z^w) \cos l_{\tilde{z}x} \\ y^r &= (O_x^r + x^w) \cos l_{\tilde{x}y} + (O_y^r + y^w) \cos l_{\tilde{y}y} + (O_z^r + z^w) \cos l_{\tilde{z}y} \\ z^r &= (O_x^r + x^w) \cos l_{\tilde{x}z} + (O_y^r + y^w) \cos l_{\tilde{y}z} + (O_z^r + z^w) \cos l_{\tilde{z}z} \end{aligned}$$

Kapitel 4

Bildvorverarbeitung

Zusammenfassung Bei der laparoskopischen Chirurgie werden Eingriffe innerhalb der Bauchhöhle mit Hilfe eines optischen Instruments, dem Endoskop, vorgenommen. Man unterscheidet zwischen monokularen und binokularen, starren und flexiblen Endoskopen, bzw. zwischen Ausführungen mit einfachem Okular zur Direktbetrachtung, und Videoendoskopen mit digitaler Bildgebungs- und Aufzeichnungstechnik. In dieser Arbeit werden schwerpunktmäßig starre, binokulare Videoendoskope mit entsprechender stereoskopischer Betrachtungsmöglichkeit behandelt. Für die robotergestützte minimalinvasive Chirurgie ergeben sich durch diese Kombination von Komponenten einige Vorteile. Zum einen ermöglicht die starre Bauweise die kinematische Integration des Endoskops in das Robotersystem. Zum anderen lassen sich durch eine binokulare Ausführung zusätzliche Techniken aus der maschinellen Bildverarbeitung anwenden.

Bei Endoskopen wird die Bildinformation des zu untersuchenden Objektes bzw. Raumes durch ein Linsensystem im Inneren des dünnen Endoskopschaftes an das Bildgebungssystem geleitet. Licht wird über Glasfaserbündel ebenfalls durch den Schaft übertragen. Durch diese spezielle, platzbedarfoptimierte Konstruktion ergeben sich sowohl aus bildverarbeiterischer, als auch aus Sicht des Chirurgen mehrere Probleme. Die Linsen haben meistens eine große Verzeichnung, wodurch Form und Dimension anatomischer Strukturen erschwert geschätzt werden können. Sie neigen außerdem stark zu chromatischen Fehlern, welche oft eine korrekte Farbauswertung verhindern. Wegen der konstruktionsbedingten hohen Packungsdichte der CCD-Sensoren von Endoskopkameras ist häufig Bildrauschen zu beobachten. Dies verursacht eine verminderte räumliche Auflösung, welche Ungenauigkeit und Ermüdung zur Folge hat. Durch die helle, stark fokussierte Lichtquelle entstehen außerdem auf den meist nassen Oberflächen der inneren Organe störende Spiegelreflexionen. Demzufolge trägt jede Komponente der Kette bestehend aus Lichtquelle, Linsen, CCD-Chip und Framegrabber unterschiedlich dazu bei, dass das aufgenommene Bild nicht der Realität entspricht, oder zumindest nicht dem, was der Mensch als solche wahrnimmt.

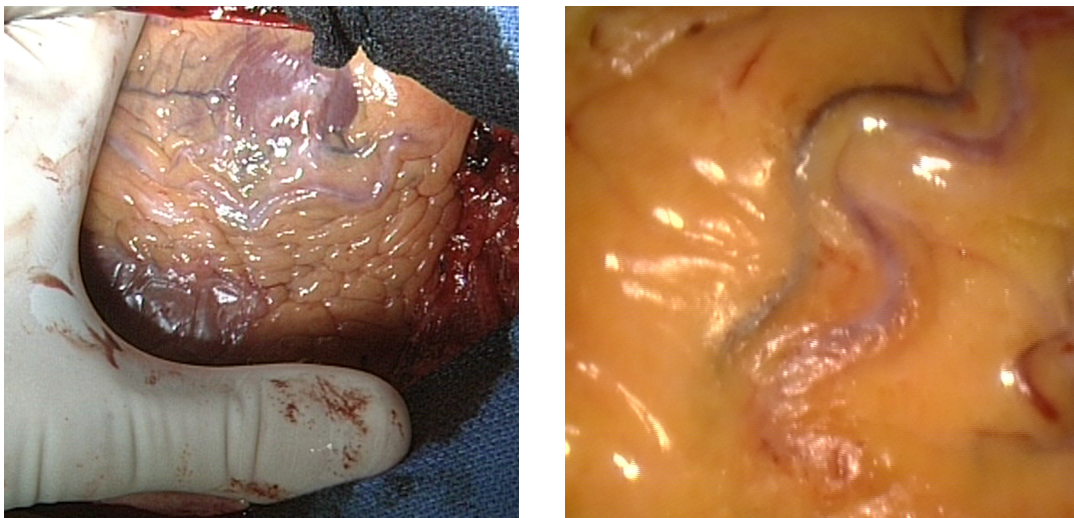
In dieser Arbeit werden für einige der oben aufgeführten Probleme Lösungsansätze vorgestellt. Ein neues, auf Kettencodes basierendes Füll-Verfahren wurde entwickelt, welches Spiegelreflexionen entfernt und zugrundeliegende Strukturen größtenteils Rekonstruiert. Das hierdurch ermöglichte Abblenden sehr heller Bereiche des Operationssitus entlastet den Chirurgen, ohne dass er dabei auf Details verzichten muss. Es besteht die Möglichkeit sowohl einer automatischen Anpassung an die Lichtverhältnisse, als auch einer feinstufigen manuellen Regelung. Der Hauptvorteil des hier vorgestellten Algorithmus ist, dass er kontinuierliche Folgen von Stereobildpaaren in Realzeit bearbeiten kann. Das Verfahren kann dadurch transparent in beliebige stereoskopische Sichtsysteme integriert werden, sodass die Korrektur bei Bedarf dynamisch (de-)aktiviert werden kann.

Die meisten heutigen Endoskopiesysteme haben in der Bildgebung einen digitalen Zwischenschritt, welcher einen Eingriff von außen auf den Bilddatenfluss ermöglicht. Dies erscheint auch sinnvoll, zumal es zahlreiche Verfahren zur nachträglichen Verbesserung der Bildqualität gibt. Einige Hersteller bieten Bildprozessoren als optionale Ausstattung ihrer Systeme an, welche Aufgaben wie Kontrasterhöhung, oder Hervorhebung bestimmter Farben übernehmen. Alle unterliegen jedoch der Einschränkung, dass sie realzeitfähig auf der vorhandenen Hardware sein müssen, um eine latenzfreie Darstellung zu gewährleisten. Doch manchmal kann es sinnvoll sein, auch nicht-realzeitfähige Maßnahmen zur Bildaufbereitung zu ergreifen, wenn diese eine substantielle Verbesserung bringen. So kann beispielsweise die Tiefeninformation erheblich verstärkt werden, wenn das Stereobildpaar vom Rauschen komplett befreit wird, ohne dabei natürlich die Kanten (Strukturen) zu glätten. Ein auf vektorwertige Regularisierung und partielle Differentialgleichungen basierendes Verfahren zur Rauschentfernung wurde ausgewählt, welche obige Anforderungen bestens erfüllt. Dessen parallelisierte, verteilte Implementierung ist zwar nicht echtzeitfähig, doch es eignet sich gut für gelegentliche Inspektionen problematischer Befunde.

4.1 Spiegelreflexionen

4.1.1 Einführung

Spiegelreflexionen (engl. *specular reflection*, *specularity*) entstehen primär auf glatten, reflektierenden Oberflächen, wenn bestimmte Voraussetzungen betreffend den Lichteinfall, Beobachtungswinkel und Materialeigenschaften erfüllt sind. Ihr Vorhandensein, bzw. ihre Intensität hängt in erster Linie von den Ein- und Austrittswinkel der einfallenden Lichtstrahlen ab. Sie sind dann besonders intensiv, wenn parallel eintreffende Lichtstrahlen auch nahezu parallel eine Oberfläche verlassen. Spiegelreflexionen zerstören das darunterliegende Bildmaterial meistens vollständig, und erweisen sich für die meisten Algorithmen der Bildverarbeitung als hinderlich.



(a) Offener Eingriff

(b) Endoskopischer Eingriff

Abbildung 4.1: Spiegelreflexionen auf der Herzoberfläche während realer Eingriffe (Quelle: Deutsches Herzzentrum München)

In der medizinischen Endoskopie sorgen die nassen Oberflächen der inneren Organe fast immer für ausgeprägte Glanzlichter. Die stark fokussierende Lichtquelle strahlt das Operationsgebiet parallel an, die Verwendung von diffusem Licht mit ausreichender Helligkeit ist technisch nicht möglich. Abbildung 4.1 zeigt Momentaufnahmen aus zwei herzchirurgischen Eingriffen. Links ist ein offener Eingriff unter Verwendung einer typischen OP-Lampe, rechts die endoskopische Sicht eines minimalinvasiven Eingriffs zu sehen. In beiden Aufnahmen ist die Auslöschung von Strukturen auf der Herzoberfläche zu beobachten.

Erschwerend kommt hinzu, dass bedingt durch die Eigenbewegungen der Organe oder der Atembewegung des Patienten, die Oberflächen sich verformen, wodurch

sich auch die Spiegelreflexionen unvorhersagbar verlagern. Ihre Helligkeit und ihre schnellen Lagewechsel sind nicht nur aus bildverarbeiterischer Sicht störend, sie bewirken außerdem eine schnellere Augenermüdung des Chirurgen. Vollständigkeithalber seien hier auch Schleierreflexionen erwähnt, die in geringerem Maße ebenfalls auftreten können (Abbildung 4.1(b), untere Bildmitte). Sie sind eine besondere Art von Spiegelreflexionen, welche jedoch die Bildinformation nicht gänzlich zerstören, sondern eine solche Verminderung des Kontrastes bewirken, dass Objekteinzelheiten nur noch schwer erkennbar sind.

Die Notwendigkeit der Behandlung von Spiegelreflexionen ist zum einen durch ergonomische Überlegungen motiviert, die zur visuellen Entlastung des Chirurgen führen sollen. Zum anderen profitieren die Algorithmen zur Modellierung des Operationssitus (Fadenerkennung und 3D-Rekonstruktion) von der Entfernung der Glanzlichter. Es ergeben sich dabei zwei wichtige Anforderungen für ein Verfahren zur Beseitigung von Spiegelreflexionen: Realzeitfähigkeit und ausreichende Wiederherstellung von Oberflächenstrukturen.

4.1.2 Vorarbeiten auf dem Gebiet

Es gibt zahlreiche Vorarbeiten, die sich mit der Behandlung von Reflexionsartefakten in der Bildverarbeitung beschäftigen. Es folgt ein Überblick relevanter Arbeiten.

Jin et al. behandeln in [94] und [95] das Problem der 3D-Rekonstruktion von Oberflächen aus mehreren kalibrierten Ansichten in Anwesenheit von Spiegelreflexionen, ohne jedoch das klassische Bild-zu-Bild Korrespondenzproblem zu lösen. Unter der Annahme, dass die Oberflächen der Objekte nicht lambertinsche¹ Reflexionen aufweisen, und dass jeder Punkt in mindestens drei Ansichten eine Spiegelreflexion aufweist, ist die Schätzung eines geeigneten Modells der Szene möglich. Der Nachteil dieser Methode ist, dass sie mehr als zwei kalibrierte Ansichten zur Rekonstruktion eines Objekts benötigt, und sich zunächst nur für Flächen eignet.

Die Motivation von Klinker et al. war zunächst die Segmentierung und Klassifikation von Bildern, unter Farbschwankungen infolge von Glanzlichtern und Schatten. Sie zeigten in [96], [97] und [98], dass die Farbe jedes Pixels eines Objekts als Linearkombination aus Objektfarbe und Glanzlichtfarbe beschrieben werden kann. Dadurch lässt sich die Farbe eines Pixels in eine Matt- und Glanzlichtkomponente trennen. Die Arbeiten basieren auf das dichromatische Reflexionsmodell, eine Mischung aus Lambertinisches- und Spiegelmodell. Prinzipiell erlaubt die Methode eine Trennung der Spiegelreflexionen vom Ausgangsbild, sieht jedoch keine Rekonstruktionsmöglichkeit vor.

Bhat et al. behandeln in [99] das Problem akkurater Tiefenschätzung mittels Stereoskopie in Anwesenheit von Spiegelreflexionen. Dadurch dass sie betrachtungswin-

¹Lambertinsche Oberflächen sind aus jeder Richtung aus betrachtet gleich hell, und reflektieren alles eintreffende Licht

kelabhängig sind, sorgen Spiegelreflexionen für große Helligkeitsunterschiede an korrespondierenden Stellen von Stereobildpaaren. Die Folgen sind Falschzuordnungen und erhebliche Tiefenfehler. Der vorgestellte Ansatz versucht eine optimale und szenenunabhängige Stereoanordnung zu bestimmen, welche Falschzuordnungen minimiert und die Tiefenauflösung maximiert. Das Hauptziel ist nicht Spiegelreflexionen zu entfernen oder zu vermeiden, sondern ihren Einfluss auf das Korrespondenzproblem zu minimieren. Später erweitern sie ihre Arbeit auf trinokulare Anordnungen [100, 101], wobei sie sich stets auf Oberflächen beschränken.

Gröger et al. stellen in [102] gleich zwei Verfahren zur Rekonstruktion von Bildstrukturen in Anwesenheit von Spiegelreflexionen vor. Ein Verfahren basiert auf Linearinterpolation, das andere ist ein iteratives Füllverfahren mittels anisotroper Diffusion. Das Ziel ihrer Arbeiten war es, die Robustheit ihrer Methode [103, 104] zur Verfolgung natürlicher Landmarken auf der bewegten Herzoberfläche zu erhöhen.

Lin et al. machen sich in [105] und [106] den Umstand zunutze, dass in binokularen Kameraanordnungen manche Pixel nur in einer Ansicht von Spiegelreflexion betroffen sind. Sie ermitteln die diffuse Komponente von Glanzlichtern einer Ansicht aus der korrelierenden Ansicht unter Zuhilfenahme der Epipolarbedingung zur Berechnung der Tiefeninformation.

Tan et al. stellen in [107] eine *inpainting*-Technik zur Glanzlichtentfernung vor, welche jedoch auf Benutzerinteraktion angewiesen ist. Chan et al. entwickeln in [108] mathematische Modelle für das lokale Inpainting texturloser Bilder. Ballester et al. führen in [109] einen neuen Ansatz ein, wodurch die Restauration von Bereichen mit fehlenden Daten in Grauwertbildern ermöglicht wird.

Wenn es nicht um die explizite Entfernung vom Glanzlichtern geht, sondern die Verbesserung der Korrelation bei der Stereorekonstruktion im Vordergrund steht, so werden in der vorgestellten Arbeiten nur Flächen behandelt. Handelt es sich jedoch primär um die Behebung von Artefakten, dann beschränkt man sich auf Schleierreflexionen. Diese Techniken sind aus Performanzgründen und wegen dem geringen Nutzen in unserem Fall nicht angebracht. Kein radiometrischer, sondern ein pragmatischer algorithmischer Ansatz ist notwendig, um die Realzeitanforderung zu erfüllen. Die Linearinterpolationsmethode nach Gröger et al. schien zunächst geeignet zu sein, bei der Nachimplementierung erwies sie sich auch als sehr schnell, brachte jedoch keine ausreichend gute Strukturwiederherstellung. Wir werden eine neue Methode vorstellen, die bei sehr guten Laufzeiteigenschaften zerstörte Strukturen größtenteils wiederherstellen kann.

4.1.3 Erkennung von Spiegelreflexionen

Wie bereits erwähnt, beschränken wir uns bei der Erkennung und Rekonstruktion nur auf Spiegelreflexionen, die aufgrund ihrer Intensität das darunterliegende Bildmaterial vollständig zerstören. Wir gehen folglich von der Annahme aus, dass

innerhalb der Spiegelreflexion kein verwertbares Bildmaterial mehr vorhanden ist. Dadurch dass Spiegelreflexionen größtenteils homogene und sehr helle Regionen sind, eignen sich insbesondere Schwellwertverfahren zu ihrer Detektion. Weil das Ergebnis der Segmentierung die Rekonstruktionsgüte direkt beeinflusst, greifen wir etwas vor, und zeigen für die verschiedenen Segmentierungstechniken auch das jeweilige Ergebnis der hier vorgestellten Rekonstruktion. Die verwendete Rekonstruktionsmethode wird in einem späteren Abschnitt genauer beschrieben.

Schwellwertbildung

Wir definieren im allgemeinen Fall eine segmentierte Region \mathcal{R} , deren Pixel Helligkeitswerte oberhalb einer bestimmte Schwelle besitzen, formal als:

$$\mathcal{R} = \{ p \mid I(p) \geq s \} \quad (4.1)$$

Dabei stellt der Parameter s den unteren Schwellwert dar, während $I(p)$ die Bildfunktion des Punktes p ist. Bei Grauwertbildern bildet die Funktion I jede Bildkoordinate eines m Pixel hohen und n Pixel breiten Bildes auf einen Grauwert aus dem Grauwertbereich mit der oberen Schranke g ab:

$$I : \{0 \dots m - 1\} \times \{0 \dots n - 1\} \rightarrow \{0 \dots g - 1\} \quad (4.2)$$

Bei mehrkanaligen Bildern, beispielsweise gegeben im RGB - oder YUV -Farbraum, stehen bei der Schwellwertbildung mehrere Kanäle zur Auswahl. Im Falle des YUV -Farbraumes können die Luminanzwerte (Y) gleich zur Schwellwertbildung herangezogen werden. Ist ein Bild im RGB -Farbraum gegeben, so besteht grundsätzlich auch die Möglichkeit einer Konvertierung in reine Helligkeitswerte. Die Gewichtung der Komponenten r, g, b erfolgt dabei unter Berücksichtigung einiger Aspekte des Farbsehens des menschlichen Auges mittels folgender (*per definitionem* exakten) Umrechnungsformel:

$$Y = 0.299 r + 0.587 g + 0.114 b \quad (4.3)$$

Abbildung 4.2 zeigt den gleichen segmentierten Ausschnitt des Bildes 4.1(a), jeweils mit dem gleichen Schwellwert ($s = 240$) angewendet auf das konvertierte Grauwertbild und die entsprechenden Farbkanäle R, G, B . Wie es Gleichung 4.3 durch die anteilmäßig stärkere Berücksichtigung des grünen Kanals bereits vermuten lässt, fallen die Ergebnisse der Segmentierung für das Grauwertbild und den grünen Farbkanal sehr ähnlich aus. Die entsprechende Anpassung der Schwellwerte und die Benutzung der Farbkanäle R oder B liefern immer schlechtere Segmentierungsergebnisse. So kann in den meisten Fällen aus Effizienzgründen ohne Einbuße in der Segmentierungsqualität der grüne Farbkanal zur Schwellwertbildung herangezogen werden.

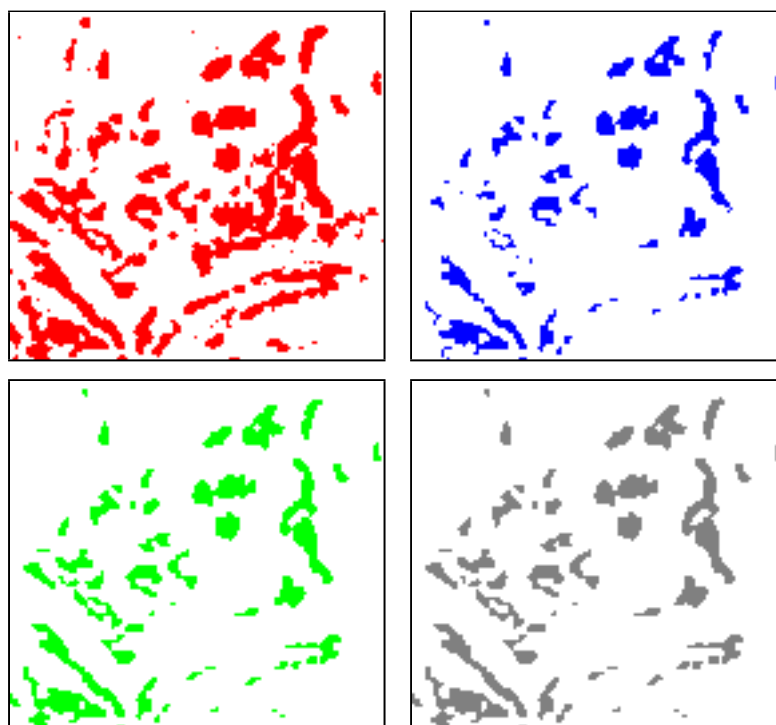


Abbildung 4.2: Ausschnitt aus Abbildung 4.1(a) nach Anlegen eines Schwellwerts an die drei Farbkanäle R, G, B und das Grauwertbild

Multischwellwertverfahren

Befinden sich auch chirurgische Instrumente im Kamerafeld, so werden diese wegen ihren Materialeigenschaften mittels obiger Methode ebenfalls als Spiegelreflexionen erkannt. Abhilfe schafft hier die Einführung eines zweiten (oberen) Schwellwertes. Etwaige verwendete Instrumente befinden sich immer näher an der Kamera und Lichtquelle, wodurch sie sich durch ein deutlich helleres Reflexionsbild auszeichnen. Wir schließen solche Falschzuordnungen dadurch aus, indem wir Gleichung 4.1 um einen zusätzlichen Schwellwert erweitern:

$$\mathcal{R} = \{ p \mid s_u \leq I(p) \leq s_o \} \quad (4.4)$$

Abbildung 4.3 zeigt das Ergebnis der Erkennung und Rekonstruktion einer Szene nach Anwendung des Multischwellwertverfahrens mit den Schwellwerten $s_u = 240$ und $s_o = 250$. Die große Fläche des chirurgischen Instruments wird von der Rekonstruktion ausgeschlossen, wodurch ein erheblicher Vorteil in der Ausführungsgeschwindigkeit entsteht. Lediglich einige kleine Bereiche am Rand werden noch falsch zugeordnet, die jedoch keinen nennenswerten Einfluss auf die Performanz haben.

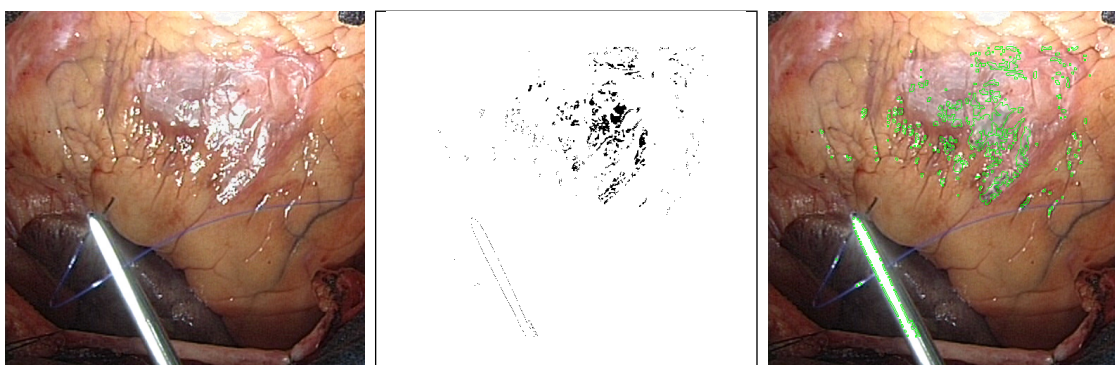


Abbildung 4.3: Segmentierung mittels eines Multischwellwertverfahrens bei chirurgischen Instrumenten im Sichtfeld

Dynamische Schwellwertermittlung

Die Festlegung statischer Schwellwerte führt nur unter konstanten Beleuchtungsverhältnissen zu akzeptablen Ergebnissen. Abbildung 4.4 zeigt Helligkeitsschwankungen, die innerhalb kurzer Zeit bedingt durch die Repositionierung der Lichtquelle aufgetreten sind. Um auch in solchen Fällen eine robuste Segmentierung zu gewährleisten, kann für jedes Bild ein eigener Schwellwert bestimmt werden.

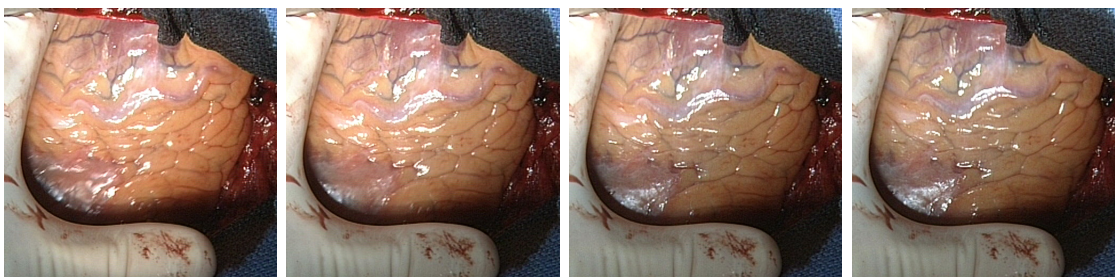


Abbildung 4.4: Helligkeitsschwankungen durch Neupositionierung der Lichtquelle

Dies wird bewerkstelligt, indem man den Schwellwert um einen konstanten Offset von dem gemittelten Helligkeitswert des jeweiligen Bildes festsetzt. Die untere Schranke aus Gleichung 4.4 berechnet sich demnach wie folgt:

$$s_u = \frac{\sum_{i=0}^{m \cdot n} I(i)}{m \cdot n} + o \quad (4.5)$$

Der Offset o lässt sich nicht empirisch berechnen, er kann jedoch für einen bestimmten Aufbau bestehend aus Kamera und Lichtquelle experimentell bestimmt werden. Diese Methode erweist sich bei moderaten Helligkeitsschwankungen als robust und effizient. Abbildung 4.5 zeigt das Ergebnis der Rekonstruktion, wobei jedes Bild mit einem eigenen Schwellwert segmentiert wurde. Bei nahezu unveränderter Szene und

einem Helligkeitsunterschied von 10% zwischen dem ersten und letzten Bild ergeben sich größtenteils segmentierte Bereiche gleicher Dimension.

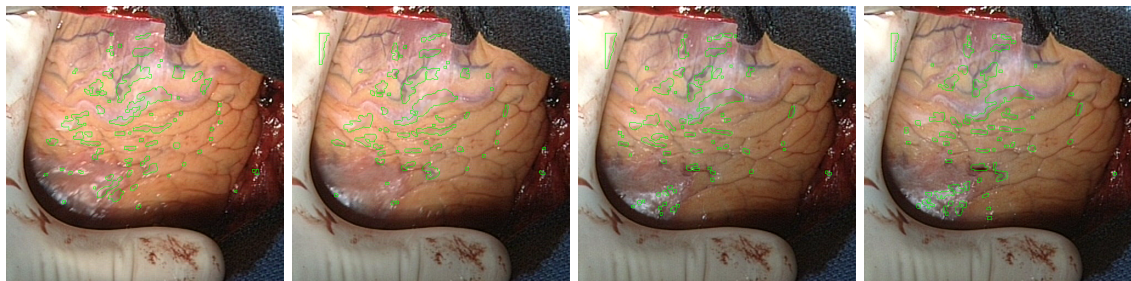


Abbildung 4.5: Rekonstruktion nach dynamischer per-Frame Schwelwertbildung

Adaptive Schwellwertermittlung

Der nächste logische Schritt ist die Bestimmung eines lokalen Schwellwerts für jedes einzelne Pixel. Sei $\{I(i, j); 0 \leq i \leq m - 1, 0 \leq j \leq n - 1\}$ das zu segmentierende Bild. Für jedes Pixel i, j werden Mittelwert m_{ij} und Varianz v_{ij} wie folgt berechnet:

$$m_{ij} = \frac{1}{p^2} \sum_{s=-p}^p \sum_{t=-p}^p I(i + s, j + t) \quad (4.6)$$

$$v_{ij} = \sum_{s=-p}^p \sum_{t=-p}^p (I(i + s, j + t) - m_{ij})^2 \quad (4.7)$$

Dabei erfolgen die Berechnungen in einer $p \times p$ großen Nachbarschaft des Pixels. Der lokale Schwellwert s_{ij} für den Bildpunkt $I(i, j)$ wird nun wie folgt bestimmt:

$$s_{ij} = \begin{cases} m_{ij} + v_{ij} & \text{falls } v_{ij} > v_{min} \\ s_{ij-1} & \text{falls } v_{ij} \leq v_{min} \end{cases} \quad (4.8)$$

Der zunächst offensichtliche Nachteil dieser Methode ist der gestiegene Berechnungsaufwand. Um auch ausgedehnte Bereiche korrekt segmentieren zu können, sind hohe Werte für die Nachbarschaft definierende Fenstergröße notwendig. Abbildung 4.6 zeigt die Rekonstruktion nach einer per-Pixel Schwelwertbildung. Eine Fenstergröße von 80 Pixel war notwendig, um alle Spiegelreflexionen korrekt zu entfernen. Trotz des betriebenen Aufwands ist diese Methode nicht optimal. Sie segmentiert zwar alle Spiegelreflexionen korrekt, doch es finden zahlreiche Falschzuordnungen statt, infolge großer lokaler Helligkeitsunterschiede, die jedoch nicht von Spiegelreflexionen verursacht wurden.

Zusammenfassend lässt sich sagen, dass der Aufwand einer adaptiven Schwellwertermittlung nicht gerechtfertigt ist. Die Ermittlung dynamischer Schwellwerte liefert bereits sehr gute Ergebnisse, oft ist sogar ein statischer Schwellwert ausreichend.

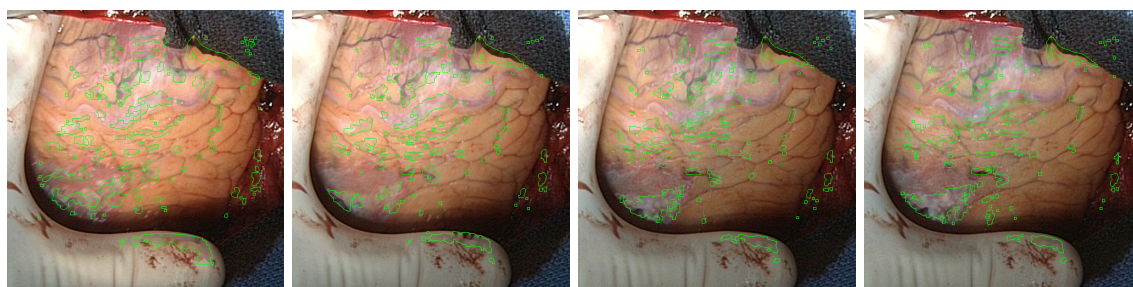


Abbildung 4.6: Rekonstruktion nach adaptiver per-Pixel Schwellwertbildung

Artefakte am Rand

Bei manchen Aufnahmen lassen sich an den Rändern von Spiegelreflexionen dunkle Artefakte beobachten. Diese entstehen offensichtlich nicht aufgrund chromatischer Fehler der Linsen oder Konstruktionsungenauigkeiten des CCD-Chips. Vielmehr handelt es sich dabei vermutlich um Kompressionsartefakte infolge von Chroma Subsampling. Einige neuere Endoskopiesysteme verwenden für die Übertragung und Speicherung von Bildsequenzen den Digital Video (DV) Standard, welcher verschiedene Stufen der Komprimierung der Farbkanäle vorsieht. Abbildung 4.7 zeigt exemplarisch Spiegelreflexionen in starker Vergrößerung, aufgenommen jeweils mit verschiedenen Bildaufzeichnungssystemen, zweimal im DV-Format und zweimal unkomprimiert mittels analoger Framegrabberkarte. Während die Ausschnitte 4.7(a) und 4.7(b) starke Störungen aufweisen, sind 4.7(c) und 4.7(d) unbetroffen.

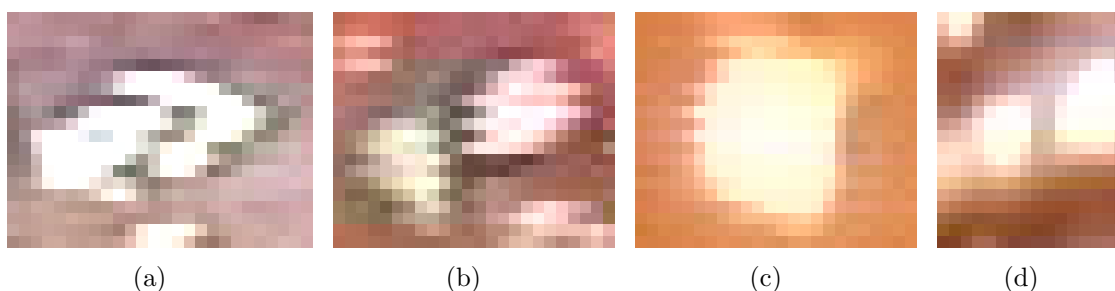


Abbildung 4.7: Spiegelreflexionen in Vergrößerung aufgenommen mittels DV (a,b) und analoger (c,d) Aufzeichnungstechnik

Betrachtet man den Grauwertverlauf entlang einer Diagonalen quer durch eine Spiegelreflexion, lässt sich der Effekt besonders gut beobachten. Abbildung 4.8 zeigt eine Zeile aus dem vergrößerten Ausschnitt 4.7(a), wobei man eine Helligkeitsschwankung von bis zu 50 Grauwertstufen erkennen kann. Diese Regionen müssen, obwohl strenggenommen nicht mehr zur Spiegelreflexion gehörend, ebenfalls als zu rekonstruierende Bereiche betrachtet werden. Die äußere Kontur einer Spiegelreflexion muss somit ausgedehnt werden, um diese Artefakte einzuschließen. Eine gezielte

Erkennung und Beseitigung erweist sich aus Performanzgründen als nicht gerechtfertigt. Nahezu gleiche Ergebnisse können mittels einer einfachen Dilatation erreicht werden. Durch die geeignete Wahl des strukturierenden Elements reicht meistens ein einmaliger Dilatationsvorgang. Einziger Nachteil der Dilatation ist, dass es auch Bereiche erweitert werden, die nicht betroffen sind (beispielsweise Abb. 4.7(a) links unten, 4.7(b) rechts oben).

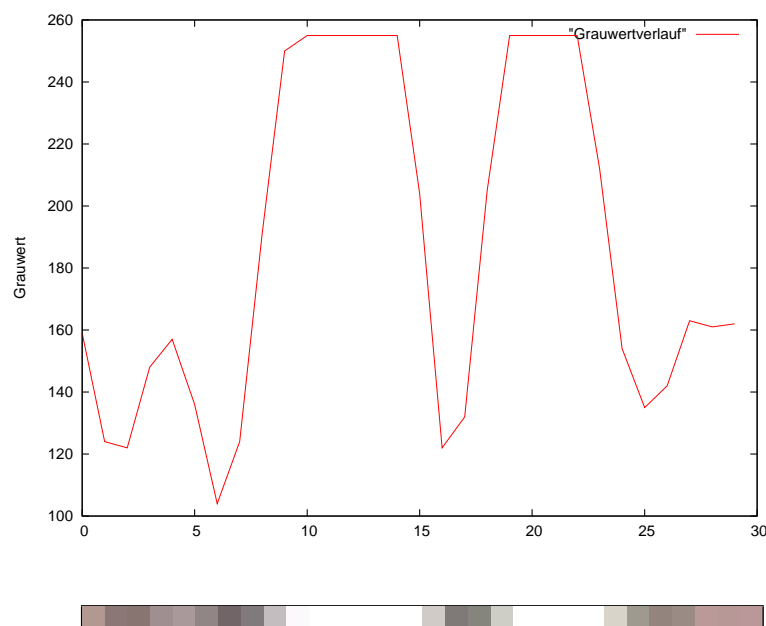


Abbildung 4.8: Grauwertverlauf am Rande von Spiegelreflexionen

Abbildung 4.9 zeigt die Ergebnisse der Schwellwertbildung und anschließender Dilatation. Als strukturierendes Element wurde ein rechteckiger Bereich von 3×3 Pixel gewählt, und es wurden zwei Dilatationsvorgänge durchgeführt.



Abbildung 4.9: Schwellwertbildung und Dilatation

4.1.4 Kettencodes zur Speicherung von Kontourdaten

Nach der Schwellwertbildung und nach Bedarf durchgeführter Dilatation stehen die segmentierten Spiegelreflexionen zunächst in Form eines Binärbildes zur Verfügung. Für die nachfolgenden Verarbeitungsschritte ist jedoch ein direkter Zugriff auf die äußere Kontour jeder Spiegelreflexion notwendig. Eine besonders effiziente Speicherung, sowohl den Platzbedarf als auch die Zugriffszeit betreffend, bieten Kettencodes. Sie wurden bereits 1974 von Freeman [110] vorgeschlagen, und werden deshalb auch Freeman-Codes genannt. Kettencodes ermöglichen eine Polygondarstellung in Pixelgenauigkeit von Objektkontouren in gerasterten digitalen Bildern. Die Idee des Verfahrens besteht darin, nur die Koordinaten eines Anfangspunktes der Kontour zu speichern, die restlichen werden durch relative Richtungsangaben kodiert. Es kann sowohl die Vierer-, als auch die Achter-Pixelnachbarschaft verwendet werden.

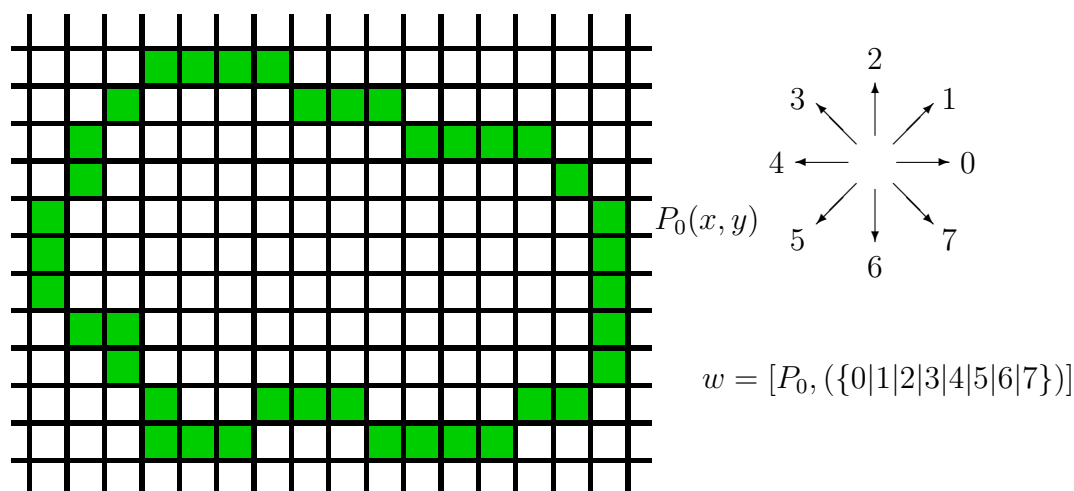


Abbildung 4.10: Kettencode unter Achter-Pixelnachbarschaft

Wenn p und q zwei bezüglich der Achter-Pixelnachbarschaft N_8 adjazente Bildpunkte sind, dann ergeben sich die Codes der relativen Richtungsänderungen $r(p, q)$ wie in Abbildung 4.10 skizziert. Die äquivalente Beschreibung eines Weges $w = (p_0, p_1, \dots, p_n)$ mit $(p_j, p_{j+1}) \in N_8$ entlang einer Kontour durch einen Freeman-Code lautet somit: $w = [p_0, C]$ mit $C = (c_1, c_2, \dots, c_n)$ und $c_j = r(p_{j-1}, p_j)$. Dabei ist $n = |w|$ die Länge, und p_0 der Startpunkt des Weges mit den Bildkoordinaten (x, y) . Das Rechnen mit Richtungen entspricht bei Kettencodes Additionen und Subtraktionen *modulo* 8 (oder *modulo* 4 bei N_4), wodurch eine sehr effiziente Implementierung der Kantenverfolgung ermöglicht wird.

Bei einer Achter-Pixelnachbarschaft und einem Ausgangspunkt $P_0(x, y)$ ergibt sich für die obenstehende Kontour folgende Beschreibung mittels Kettencodes: $w = [(x, y), (33444344434444556566706760010070001012222)]$. Diese Art der Kontourbeschreibung benachteiligt zwar den Direktzugriff auf die einzelnen Bildkoordinaten, ermöglicht aber dafür eine sehr effiziente Traversierung vom Startpunkt aus.

4.1.5 Rekonstruktion mittels Linearinterpolation

Entlang horizontaler Sehnen

Die Grundidee dieses Verfahrens besteht darin, zwischen zwei horizontal entgegengesetzten, und sich am Rande einer Spiegelreflexion befindlichen Bildpunkten die fehlenden Farbwerte linear zu interpolieren. Abbildung 4.11 zeigt schematisch die Vorgehensweise: es werden für jeden Bildpunkt zwischen $P_1[r_1, g_1, b_1]$ und $P_2[r_2, g_2, b_2]$ die jeweiligen Komponenten der Farbkanäle im *RGB*-Farbraum iterativ berechnet. Dabei bestimmt die horizontale Entfernung d im Bild zwischen P_1 und P_2 die Anzahl der unterschiedlichen Farbintervalle. Die Schrittweiten s_r, s_g und s_b werden für jeden Farbkanal getrennt ermittelt, beispielsweise für den roten Farbkanal als $s_r = (r_1 - r_2)/d$. Angefangen bei $[r_1, g_1, b_1]$ werden bei jeder Iteration die Schrittweiten solange aufaddiert, bis $[r_2, g_2, b_2]$ erreicht wird.

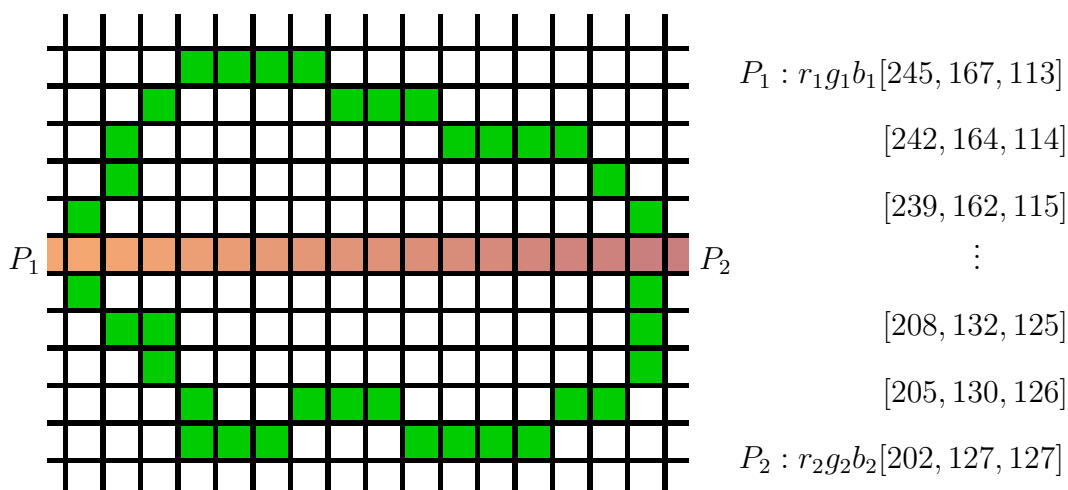


Abbildung 4.11: Linearinterpolation zwischen horizontal entgegengesetzten Pixel

Der Zugriff auf die Kontourdaten einer Spiegelreflexion erfolgt über ihre Kettencodarstellung. Wie bereits erwähnt, erlaubt diese Art der Darstellung keinen direkten Zugriff auf Bildkoordinaten einzelner Pixel. Diagonal entgegengesetzte Punkte müssten somit mittels Traversierung aller Kontourpunkte bis zum Erreichen eines Punktes mit gleicher Zeilenkoordinate ermittelt werden. Eine einfachere Möglichkeit entsteht, wenn auch das segmentierte Binärbild verwendet wird. Dazu iteriert man ausgehend von einem Kontourpunkt (P_1) horizontal solange in eine Richtung (beispielsweise in Richtung steigender Spaltenkoordinaten), bis die Spiegelreflexion verlassen wird. Der erreichte Punkt (P_2) kann nun zusammen mit P_1 für die Linearinterpolation verwendet werden.

Die Methode vermag keine Strukturen innerhalb der Spiegelreflexion zu rekonstruieren, und eignet sich wegen des Unschärfefeffektes nur zur Behandlung kleiner Regionen. Zudem entsteht durch die Linearinterpolation der oft unerwünschte Effekt, dass

zusätzliche Farben ins Bild eingefügt werden. Abbildung 4.12 zeigt das Ergebnis der Rekonstruktion für einen Ausschnitt der Szene aus Abbildung 4.1(a).



Abbildung 4.12: Rekonstruktion von Spiegelreflexionen mittels Linearinterpolation entlang horizontaler Sehnen

Der große Vorteil des Verfahrens ist seine Einfachheit und dadurch bedingt die hohe Ausführungsgeschwindigkeit. Auf der Referenzhardware² erreicht man eine Rekonstruktionsrate von etwa 55 Bildern (512×512 @ 24bit Farbtiefe) pro Sekunde. Abbildung 4.13 zeigt das Ergebnis der Rekonstruktion für das gesamte Operationsgebiet. Negativ fällt der horizontale Verwischungseffekt auf, der bei ausgedehnten Regionen besonders ausgeprägt, und auf die Nichtberücksichtigung lokaler Strukturinformation zurückzuführen ist.

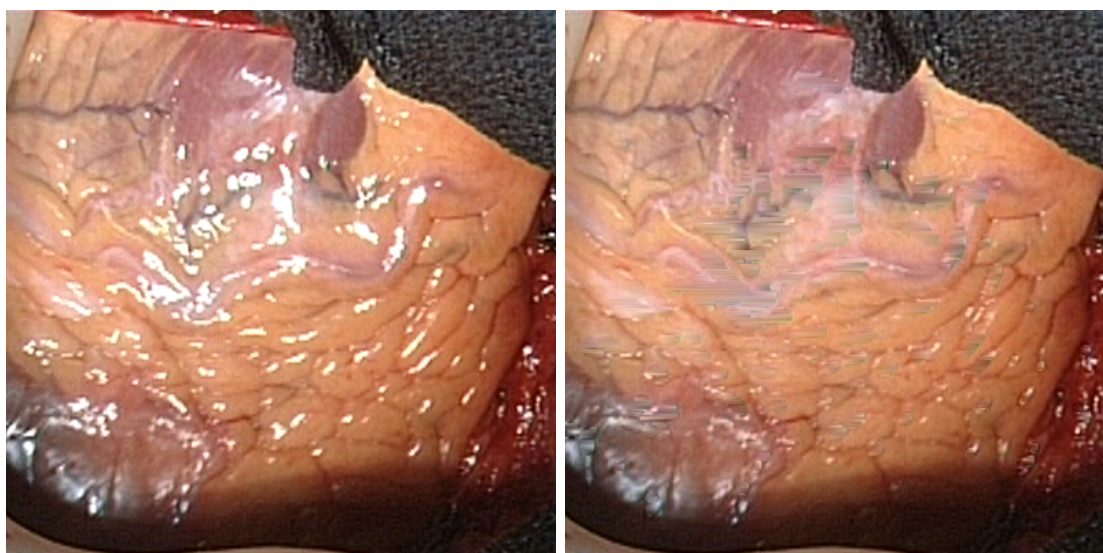


Abbildung 4.13: Entfernung von Spiegelreflexionen für das gesamte Operationsgebiet

²AMD Athlon 64 Prozessor, 2.2 GHz Taktfrequenz

Entlang gerichteter Sehnen

Das obige Verfahren kann dahingehend erweitert werden, dass die lokale Ausrichtung der Bildstruktur am Rand der Spiegelreflexion berücksichtigt wird. Gröger et al. schlagen in [102] die Verwendung des Strukturtenors zur Bestimmung von Textureigenschaften vor. Dabei wird für jeden Bildpunkt die Kovarianzmatrix direktionaler Gradienten in einer rechteckigen Nachbarschaft bestimmt. Direktionale Gradienten erhält man durch die Faltung des Bildes \mathcal{I} mit zwei orthogonalen Filtern G_x und G_y . Die Faltung $\mathcal{I} \otimes G_x$ ergibt den Gradienten in x -, $\mathcal{I} \otimes G_y$ den Gradienten in y -Richtung, wodurch sich für jedes Pixel ein Vektor mit Gradienten $\mathbf{g} = (g_x, g_y)^T$ ergibt. Für \mathbf{g} lässt sich anschließend sowohl Gradientenstärke $|\mathbf{g}|$ als auch die Richtung der schnellsten Grauwertzunahme θ bestimmen:

$$|\mathbf{g}| = \sqrt{g_x^2 + g_y^2}, \quad \theta = \tan^{-1} \left(\frac{g_y}{g_x} \right) \quad (4.9)$$

Eine optimale Komposition von Differenzialoperator und Tiefpassfilter bietet die Gauß-Funktion:

$$\begin{aligned} G(x, y : \sigma) &= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \\ G_x(x, y) = \frac{\partial}{\partial x} G(x, y : \sigma) &= \frac{-x}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \\ G_y(x, y) = \frac{\partial}{\partial y} G(x, y : \sigma) &= \frac{-y}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \end{aligned}$$

Mit Hilfe von σ lässt sich die Rauschempfindlichkeit steuern. Bei kleinem σ sind feine Strukturen gut erkennbar, Textur und Rauschen bleiben erhalten. Große Werte für σ hingegen bewirken eine Rauschunterdrückung, wobei nur grobe Strukturen hervorgehoben werden. Nun ergibt sich für jedes Pixel der Strukturtenor \mathbf{M} als folgende Matrix:

$$\mathbf{M} = \begin{pmatrix} \sum_W g_x^2 & \sum_W g_x g_y \\ \sum_W g_y g_x & \sum_W g_y^2 \end{pmatrix} \quad (4.10)$$

W ist die Nachbarschaft des Pixels, dessen Tensor berechnet wird, und entspricht einem quadratischen Fenster mit Seitenlänge N_W um das Pixel. \mathbf{M} gleicht einer Kovarianzmatrix, die mit dem Mittelwert $\mu = 0$ berechnet wird, und beschreibt die Verteilung der Gradienten um den Ursprung $\mathbf{g} = (0, 0)^T$. Die Eigenwerte λ_1 und λ_2 der Matrix \mathbf{M} geben Aufschluss über die Gradientenstärken in den Richtungen der Eigenvektoren \mathbf{v}_1 und \mathbf{v}_2 . Der zum größeren absoluten Eigenwert ($|\lambda_1|$) gehörende Eigenvektor \mathbf{v}_1 zeigt die Richtung der größten Grauwertänderung, auch

die Rekonstruktionssehne nun der Strukturorientierung entspricht. Die Anzahl der Punkte zwischen zwei zu interpolierenden Pixel und ihre Bildkoordinaten werden mit Hilfe des Bresenham-Algorithmus zur iterativen Rasterung einer Linie ermittelt. Abbildung 4.15 stellt schematisch die Vorgehensweise dar, während Abbildung 4.16 das Ergebnis der Rekonstruktion für einen kleinen Ausschnitt der Szene aus Abbildung 4.1(a) zeigt. Auf der Referenzhardware erreicht man eine Rekonstruktionsrate von etwa 25 Bildern (512×512 @ 24bit Farbtiefe) pro Sekunde.

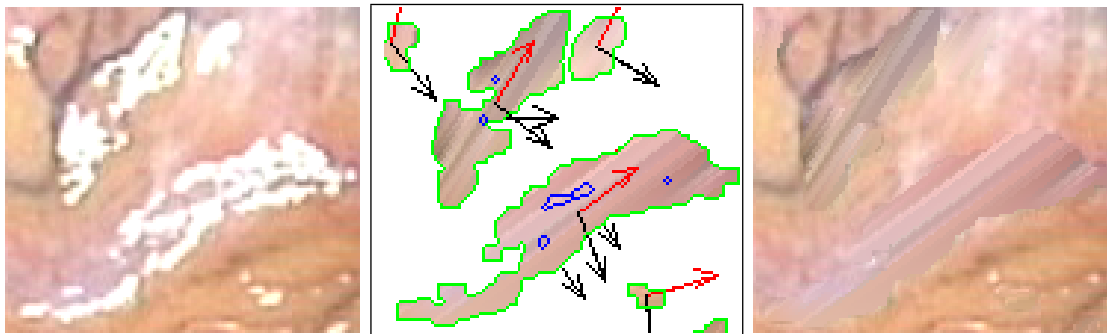


Abbildung 4.16: Rekonstruktion einer Spiegelreflexion mittels Linearinterpolation unter Berücksichtigung der Strukturorientierung

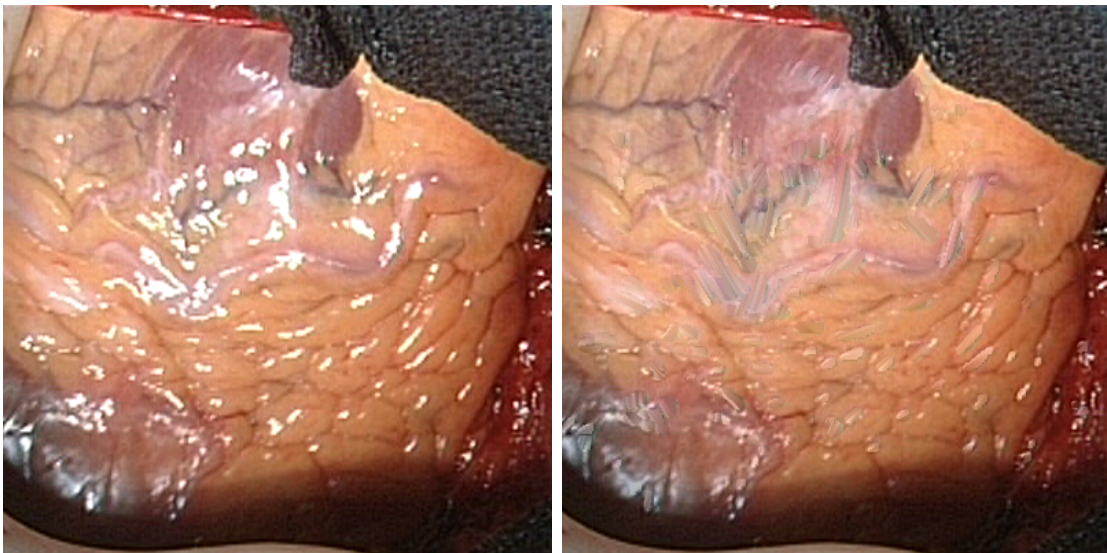


Abbildung 4.17: Entfernung von Spiegelreflexionen für das gesamte Operationsgebiet

Das Bild wirkt, wie in Abbildung 4.17 zu sehen, auch bei ausgedehnten rekonstruierten Bereichen natürlicher als bei der Verwendung horizontaler Sehnen. Alternativ ist es möglich, statt der Strukturorientierung die Richtung der größten Kontraständerung für die Rekonstruktionssehnen zu verwenden. Die Abbildungen 4.18 und 4.19 zeigen die Ergebnisse nach der Entfernung der Spiegelreflexionen der gleichen Szene

wie in 4.16 und 4.17. Hier wurde jedoch bei der Rekonstruktion die Richtung der jeweils größten Kontraständerung berücksichtigt. Entgegen den Erwartungen führt in manchen Fällen die Berücksichtigung des Kontrastes zu natürlicher aussehenden Ergebnissen, als bei der Verwendung der Strukturorientierung. Vergleicht man die zwei rekonstruierten Ausschnitte aus 4.18 und 4.16, ist zumindest für den menschlichen Betrachter das Ergebnis aus 4.18 subjektiv besser, bei einem nahezu identischen Laufzeitverhalten.

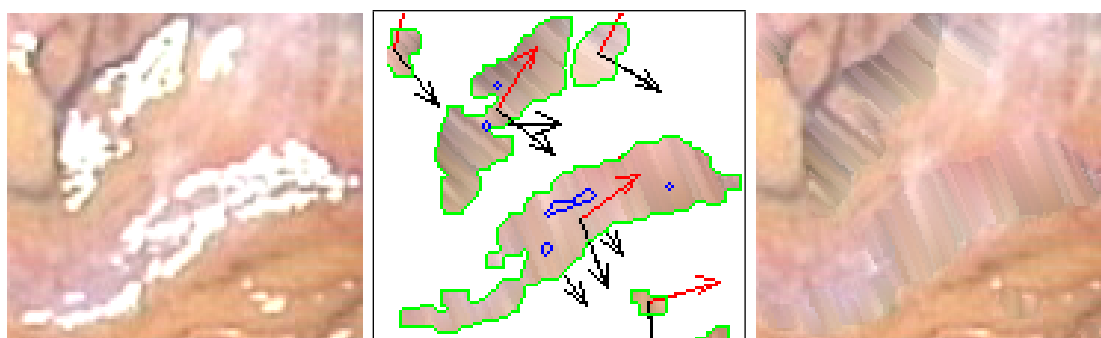


Abbildung 4.18: Rekonstruktion einer Spiegelreflexion mittels Linearinterpolation unter Berücksichtigung der Richtung der größten Kontraständerung

Der subjektive Eindruck kann jedoch irreführend sein, relevant sind nur messbare Kriterien, welche für die anschließende Bildverarbeitung von Nutzen sind. Im letzten Abschnitt dieses Kapitels wird die Rekonstruktionsgüte der einzelnen Verfahren empirisch ermittelt und verglichen.

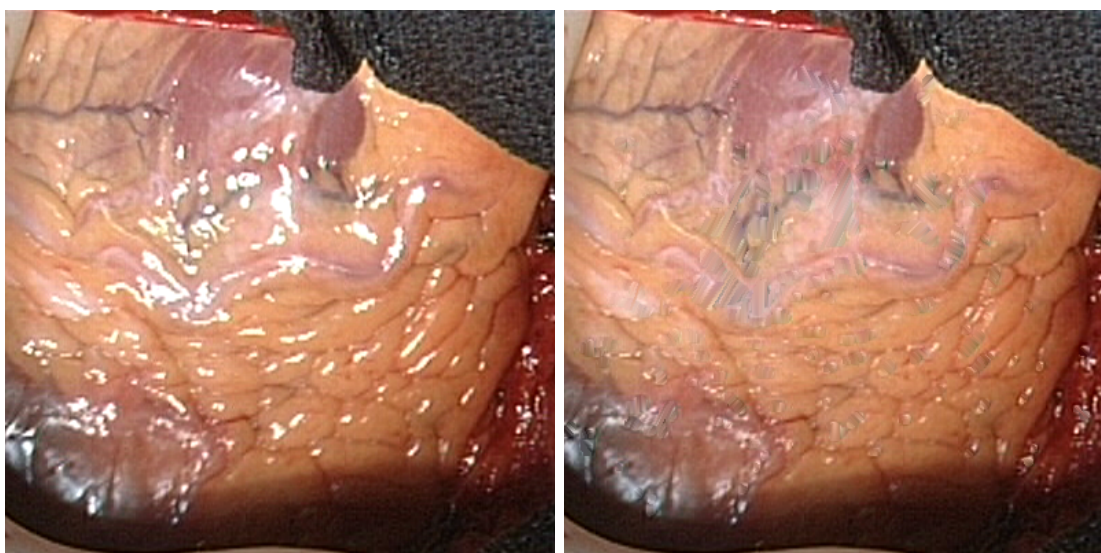


Abbildung 4.19: Entfernung von Spiegelreflexionen für das gesamte Operationsgebiet

Die Bewegungsunschärfe, bedingt durch die ruckartige Natur des Herzschlages, führt

insbesondere bei den Übergängen zwischen den Phasen des Herzzyklus zu verfälschten Werten bei der Berechnung der Eigenwerte und Eigenvektoren. In solchen Fällen sind die Unterschiede zwischen der Rekonstruktion anhand der Strukturorientierung und Richtung der größten Kontraständerung zufällig.

4.1.6 Rekonstruktion mittels Collapsing Chains

Die Grundidee dieses hier neu vorgestellten Verfahrens besteht darin, die äußere Kontour einer Spiegelreflexion iterativ um einen Pixel nach innen zu verlagern, indem Kontourpunkte anhand benachbarter Bildpunkte ersetzt werden. Die Rekonstruktion einer Spiegelreflexion gilt als beendet, wenn das Innere des Artefakts nur noch aus einem Pixel besteht, der ebenfalls substituiert wird. Bei der Ersetzung der Kontourpunkte werden diejenigen Pixel aus einer Achter-Nachbarschaft ausgewählt, die als nicht zur Spiegelreflexion gehörend segmentiert wurden. Abhängig von der Form der Kontour, schwankt die Anzahl solcher Pixel entsprechend.

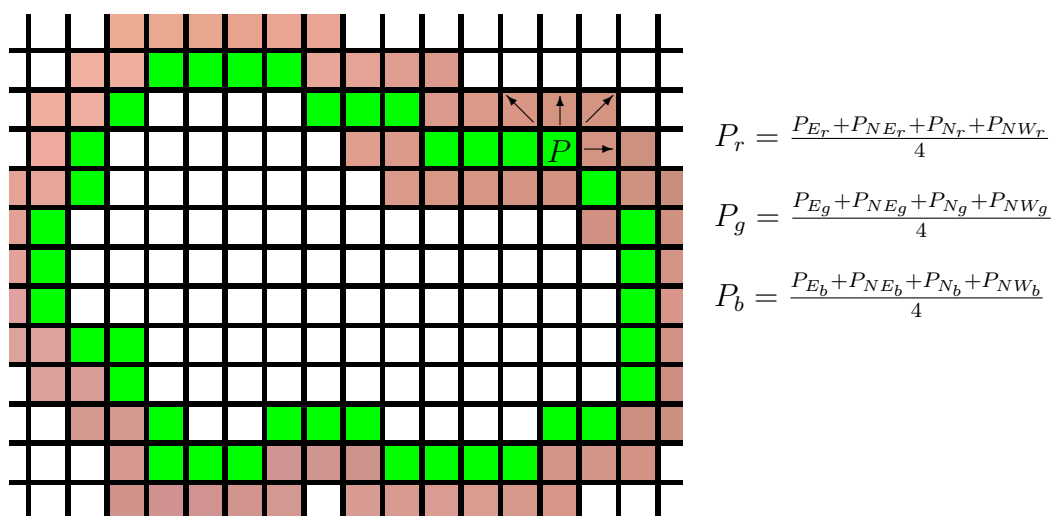


Abbildung 4.20: Funktionsprinzip Collapsing Chains

Abbildung 4.20 zeigt schematisch die vorgehensweise. So hat beispielsweise der Punkt P in einer N_8 -Nachbarschaft vier verwendbare Bildpunkte als Nachbarn, sodass sich sein RGB -Wert aus den jeweiligen gemittelten Komponenten der RGB -Werte der Nachbarpixel bilden. Nachdem alle Pixel des äußeren Verlaufs berechnet wurden, dienen sie nun als geeignete Nachbarn für die nächste Kontour. Sie hat sich um eine Pixelbreite in Richtung Objektinnere verlagert, der Umfang verringert sich bei jeder weiteren Iteration linear. Wie bereits vorgestellt, werden die Kontouren einzelner Spiegelreflexionen als Kettencodes gespeichert. Sie werden bei jeder Iteration für die stetig sich verkürzenden Kontouren neu berechnet. Die Namensgebung des Verfahrens geht auf diese schrumpfenden Umlaufketten zurück.

Ein großer Vorteil dieses Verfahrens ist die ausgeprägte Lokalität bei der Ersetzung von Kontourpunkten. Dadurch werden etwaige Strukturen aus der näheren Umgebung ins Innere fortgesetzt. Dieser Effekt lässt sich in Abbildung 4.21 anhand der rekonstruierten Gefäße besonders gut beobachten. Alle Linearinterpolationsverfahren versagen bei der Rekonstruktion solcher feinen Strukturen, lediglich der im nächsten Abschnitt vorgestellte Verfahren erzielt ähnlich gute Ergebnisse.

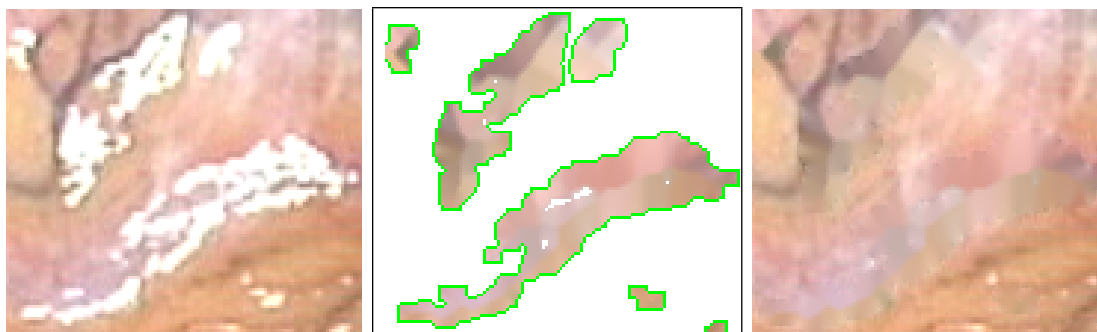


Abbildung 4.21: Rekonstruktion einer Spiegelreflexion mittels Collapsing Chains

Dank der effizienten Speicherung der Kontourverläufe als Kettencodes verfügt das Verfahren über ein sehr gutes Laufzeitverhalten. Auf der Referenzhardware erreicht sie eine Rekonstruktionsrate von 45 bis 50 Bildern (512×512 @ 24bit Farbtiefe) pro Sekunde, abhängig von der Anzahl der Spiegelreflexionen. Somit ist die Behandlung von Stereobildpaaren in Realzeit möglich.

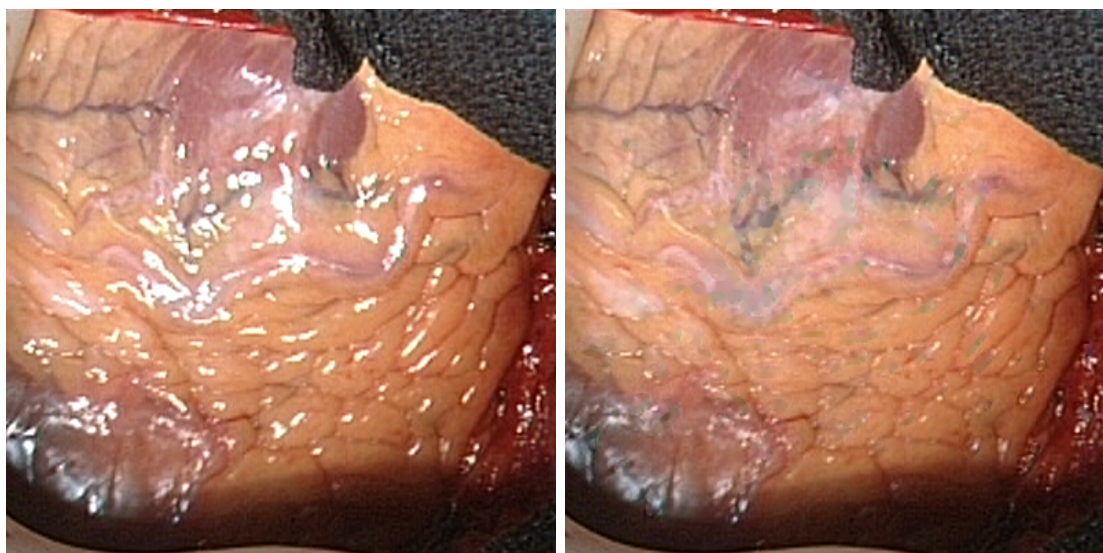


Abbildung 4.22: Entfernung von Spiegelreflexionen für das gesamte Operationsgebiet

Auch bei ausgedehnten Regionen werden Strukturen gut wiederhergestellt. Der Verlauf der Koronararterie quer durch die Mitte des Bildes aus Abbildung 4.22 wird

nach der Rekonstruktion gut sichtbar. Damit eignet sich die Methode zur transparenten Einbindung in stereoendoskopische Sichtsysteme.

4.1.7 Regularisierung vektorwertiger Daten mittels partiellen Differenzialgleichungen (RVWPDGL)

Das Ziel bei der Regularisierung von Daten ist es, ein verrauschtes Signal so zu glätten, dass wichtige Informationen, wie beispielsweise markante Unstetigkeiten, nicht verlorengehen. In den späten 80ern stieg das Interesse an anisotropen regularisierenden partiellen Differenzialgleichungen in der Bildverarbeitung. Ihre Fähigkeit zur Glättung, und dabei globale Merkmale wie Kontouren und Kanten zu erhalten eröffneten neue Wege zur Behandlung klassischer Aufgabenstellungen aus der Bildverarbeitung wie Restauration, Segmentierung und Registrierung.

Abbildung 4.23(a) zeigt ein stark verrauschtes Rechtecksignal, das in Abbildung 4.23(b) mit einem Gauß'schen Weichzeichenfilter behandelt und in Abbildung 4.23(c) regularisiert wurde. Als Folge der Gauß-Glättung wurden die Sprungstellen stark abgerundet und die Flankensteilheit reduziert, während die Regularisierung das Signal nahezu fehlerfrei wiederherstellt.

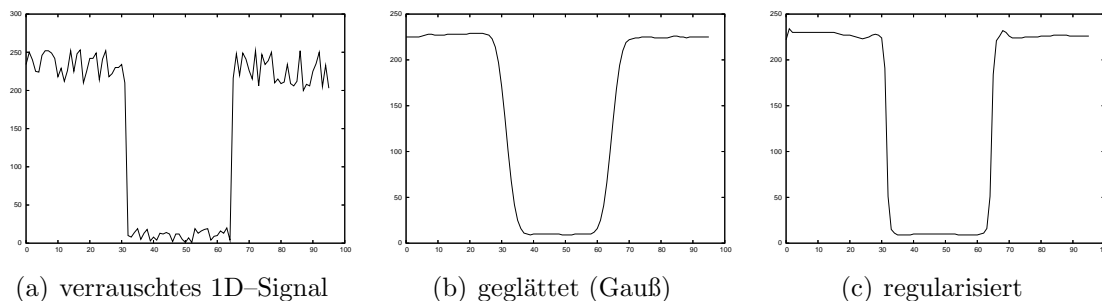


Abbildung 4.23: Verrauschtes 1D-Signal und unterschiedliche Glättungstechniken

Die meisten in der Literatur ursprünglich vorgestellten Methoden zur Regularisierung eignen sich zur Behandlung skalarer Daten, beispielsweise von Grauwertbildern der Form $I : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. In der Bildverarbeitung verwendet man partielle Differenzialgleichungen (PDGL) der Gestalt

$$\frac{\partial I}{\partial t} = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (4.11)$$

zur anisotropen Glättung, wobei I die zu verarbeitenden Daten (1D-Signal, oder 2D/3D-Bild) darstellt. Die Zeitvariable t entspricht der Anzahl von Iterationen des Algorithmus. Man startet mit einem Bild $I_{(t=0)}$, welches bis zur Konvergenz, oder bis Erreichen einer endlichen Anzahl von Iterationen $t = t_{end}$ entwickelt wird:

$$\begin{cases} I_{(t=0)} = I_0 \\ \frac{\partial I_{(x,y,t)}}{\partial t} = \beta_{(x,y,t)} \end{cases} \quad (4.12)$$

Es gab viele verschiedene Ansätze und Vorgehensweisen, wobei wir zwei prominente Arbeiten herausgreifen. Die ursprüngliche Motivation von Koenderink war der Zusammenhang zwischen Bildstrukturen und unterschiedliche Auflösungsstufen. Er schlägt in [111] die Einbettung des Originalbildes in eine Klasse abgeleiteter Bilder, welche durch Faltung mit einem Gauß-Kern entstehen:

$$I_{(t)} = I_{(t=0)} * G_\sigma, \text{ wobei } G_\sigma = \frac{1}{4\pi t} e^{-\frac{x^2+y^2}{4t}} \Leftrightarrow \frac{\partial I}{\partial t} = \Delta I = \operatorname{div}(\nabla I) \quad (4.13)$$

Er wies darauf hin, dass diese Familie abgeleiteter Bilder auch als Lösungen der Wärmeleitung oder Diffusion gesehen werden können. Perona und Malik stellen in [112] eine einfache Möglichkeit zur Erweiterung des *space-scale*-Paradigma vor, und verwenden nichtlineare PDGLen in der anisotropischen Diffusionsgleichung

$$\frac{\partial I}{\partial t} = \operatorname{div}(c \nabla I) = c \Delta I + \nabla c \cdot \nabla I, \quad c: \mathbb{R} \rightarrow \mathbb{R}, \quad (4.14)$$

wobei div den Divergenzoperator darstellt, mit ∇ und Δ werden jeweils Gradient- und Laplace-Operator bezeichnet. Für konstante c vereinfacht sich Gleichung 4.14 zu $I_{(t)} = c \Delta I$. Sie führen den nichtkonstanten Diffusionskoeffizienten c ein, der idealerweise 1 innerhalb von Regionen ist, und 0 an den Rändern. Weil das von vornherein nicht bekannt ist, schätzen sie den aktuellen Wert anhand des Helligkeitsgradienten.

Erweiterungen dieser Verfahren auf vektorwertige Bilder $I: \Omega \rightarrow \mathbb{R}^n$, beispielsweise zur Behandlung von Vektorfeldern, wurden später ebenfalls vorgeschlagen ([113], [114], [115]). In den so entwickelten PDGLen taucht nun die Kopplung zwischen den einzelnen Bildkanälen auf, unter Berücksichtigung des Strukturtenors, wie er bereits in Abschnitt 4.1.5 vorgestellt wurde.

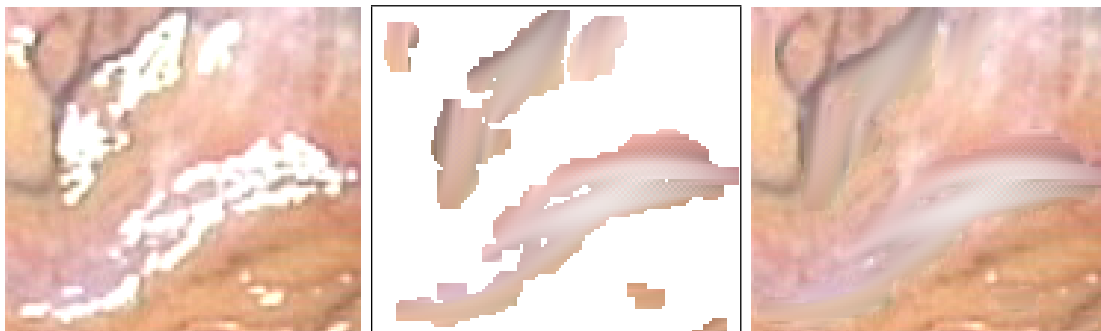


Abbildung 4.24: Rekonstruktion einer Spiegelreflexion mittels RVWPDGL

Die Regularisierung mittels PDGLen kann auch für das Inpainting in (Farb)Bildern verwendet werden. Dabei werden die Gleichungen nur auf vordefinierte Gebiete (definiert durch Masken) angewendet. Durch anisotropische Regularisierung werden die ausgewählten Bereiche strukturerhaltend rekonstruiert. Während Groeger et. al. in [102] eine Implementierung nach Weickert [113] vorstellen, verwenden wir hier die Implementierung nach Tschumperlé und Deriche [114, 115], welche auch die Behandlung von Farbbildern erlaubt. Abbildung 4.24 zeigt die Ergebnisse der Diffusion nach 500 Iterationen. Dabei wurden die als Spiegelreflexionen segmentierte Bereiche als Masken verwendet.

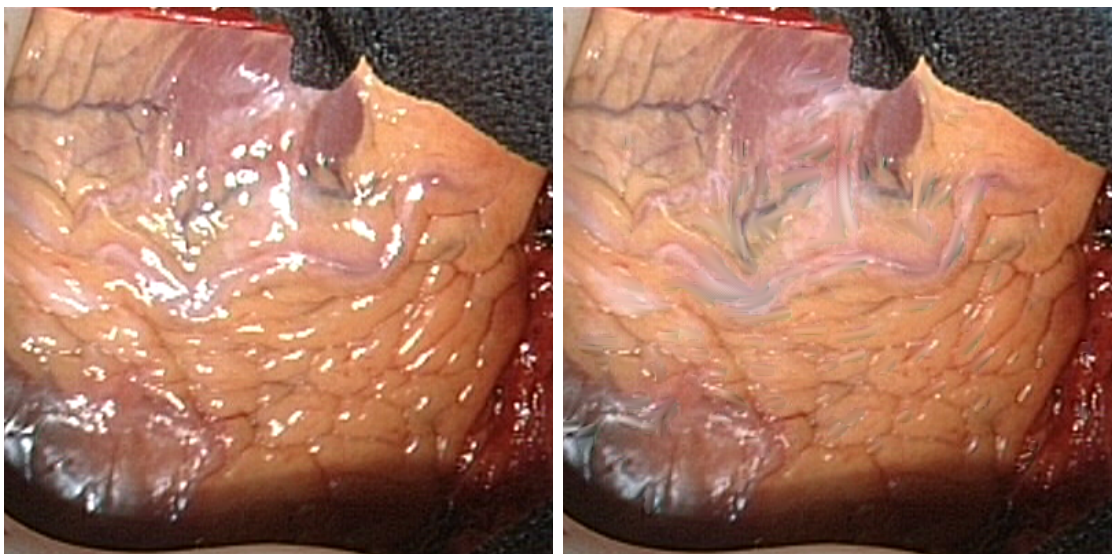


Abbildung 4.25: Entfernung von Spiegelreflexionen für das gesamte Operationsgebiet

Der größte Nachteil aller Diffusionsfilter ist der enorm hohe Berechnungsaufwand. So dauert beispielsweise die Berechnung von 500 Iterationen je nach Anzahl und Größe der Spiegelreflexionen eine halbe bis ganze Minute auf der Referenzhardware. Das subjektive visuelle Ergebnis ist dafür besser als bei allen Verfahren basierend auf Linearinterpolation, bleibt jedoch hinter dem des Collapsing Chains Verfahrens.

Perona und Malik schlagen in [112] auch eine mögliche analoge Hardwareimplementierung vor, welche die gute Parallelisierbarkeit des Verfahrens ausnutzt. Eine parallele Implementierung auf modernen CISC/RISC-Prozessoren ist ebenfalls möglich, auf einem Rechnercluster bestehend aus 4×4 Prozessoren gelang es somit ein Stereobildpaar pro Sekunde zu rekonstruieren. Dies scheint auch die sinnvolle Grenze für die Anzahl verteilter Rechenknoten zu sein, denn der Kommunikationsaufwand zwischen den Knoten steigt ebenfalls. Unter Umgehung der Netzwerkkommunikation, beispielsweise auf Mehrprozessor-/Mehrkern-Systemen, wären bei entsprechender Implementierung weitaus höhere Geschwindigkeitsvorteile erreichbar.

4.1.8 Evaluation der vorgestellten Verfahren

Wie bereits erwähnt, ist der subjektive Eindruck eines menschlichen Betrachters kein ausschlaggebendes Kriterium für die Beurteilung der Rekonstruktionsgüte. Idealerweise würde man in solchen Fällen mit Hilfe eines geeigneten Ähnlichkeitsmaßes zwischen rekonstruierten Bildern, und Bildern ohne jegliche Artefakte der selben Szene, auf die Qualität des jeweiligen Verfahrens schließen. Dafür müsste man die Lichtquelle so repositionieren, dass zuvor von Spiegelreflexionen betroffene Bereiche davon frei werden. Dadurch dass die Lichtquelle in das Endoskop integriert ist, wirken sich Änderungen am Lichteinfall direkt auf den beobachteten Ausschnitt aus, sodass keine Vergleichbarkeit bestimmter Bereiche mehr möglich ist. Aus diesem Grund wurde hier ein anderer Ansatz für die Beschaffung von *ground-truth*-Daten³ gewählt, welche sich gut für die Durchführung der Vergleiche eignet. Dabei wurde aus dem linken Bild eines Stereobildpaares eine typische Spiegelreflexion segmentiert und in eine Region desselben Bildes mit intakter Oberflächenstruktur kopiert (Abbildung 4.26). Die Zielregion wurde so gewählt, dass sie in ausreichendem Maße texturiert ist. Anschließend wurde dieser Bereich mit Hilfe der vorgestellten Verfahren rekonstruiert, und mit dem Original verglichen.

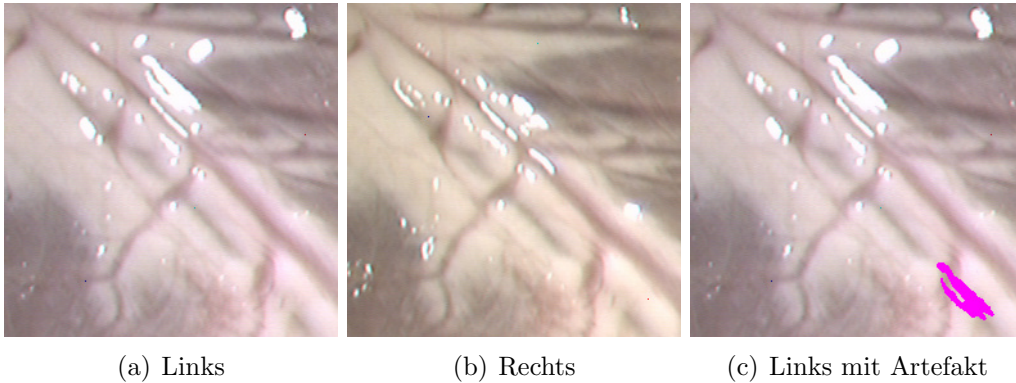


Abbildung 4.26: Endoskopisches Stereobildpaar für die Evaluation

Als Ähnlichkeitsmaß wurde eine Art der normierten Kreuzkorrelation (engl. *zero mean normalized cross correlation*, ZNCC) verwendet, welche zum einen Beleuchtungsschwankungen gut ausgleicht, und zum anderen auch bei der Disparitätsberechnung der Stereorekonstruktion eingesetzt werden kann:

$$ZNCC(r, c) = \frac{\sum_{r'=r-m}^{r+m} \sum_{c'=c-n}^{c+n} (g_1(r', c') - \bar{g}_1(r, c)) \cdot (g_2(r', c') - \bar{g}_2(r, c))}{\sqrt{\sum_{r'=r-m}^{r+m} \sum_{c'=c-n}^{c+n} (g_1(r', c') - \bar{g}_1(r, c))^2 \cdot (g_2(r', c') - \bar{g}_2(r, c))^2}} \quad (4.15)$$

³mangels eines deutschen Äquivalents verwenden wir hier den englischen Ausdruck für genaue, unverfälschte Daten, wie er insbesondere in der Stereorekonstruktion verwendet wird

Mit Hilfe des ZNCC ist es möglich, die Ähnlichkeit von rechteckigen Bereichen aus zwei Grauwertbildern g_1 und g_2 auf das reelle Intervall $[-1, 1]$ abzubilden, wobei der Wert 1 identischen Regionen entspricht. In einer Nachbarschaft der Breite $2m+1$ und Höhe $2n+1$ gibt $\bar{g}(r, c)$ den gemittelten Grauwert eines Pixels an der Position (r, c) an, wobei $g(r', c')$ den Grauwerten innerhalb der untersuchten Region entspricht. Zur besseren Visualisierung der Korrelationsergebnisse wurde das Ergebnisintervall $[-1, 1]$ mit folgender Berechnungsvorschrift auf die ganzzahligen Werte aus $[0 - 255]$ abgebildet:

$$g = 50 + (score \cdot 100)^{1.115} \quad (4.16)$$

Das Ergebnis ist als Grauwertbild dargestellt, wobei helle Bereiche guten Korrelationscores entsprechen, dunkle hingegen deuten auf erhebliche Abweichungen vom Original hin.

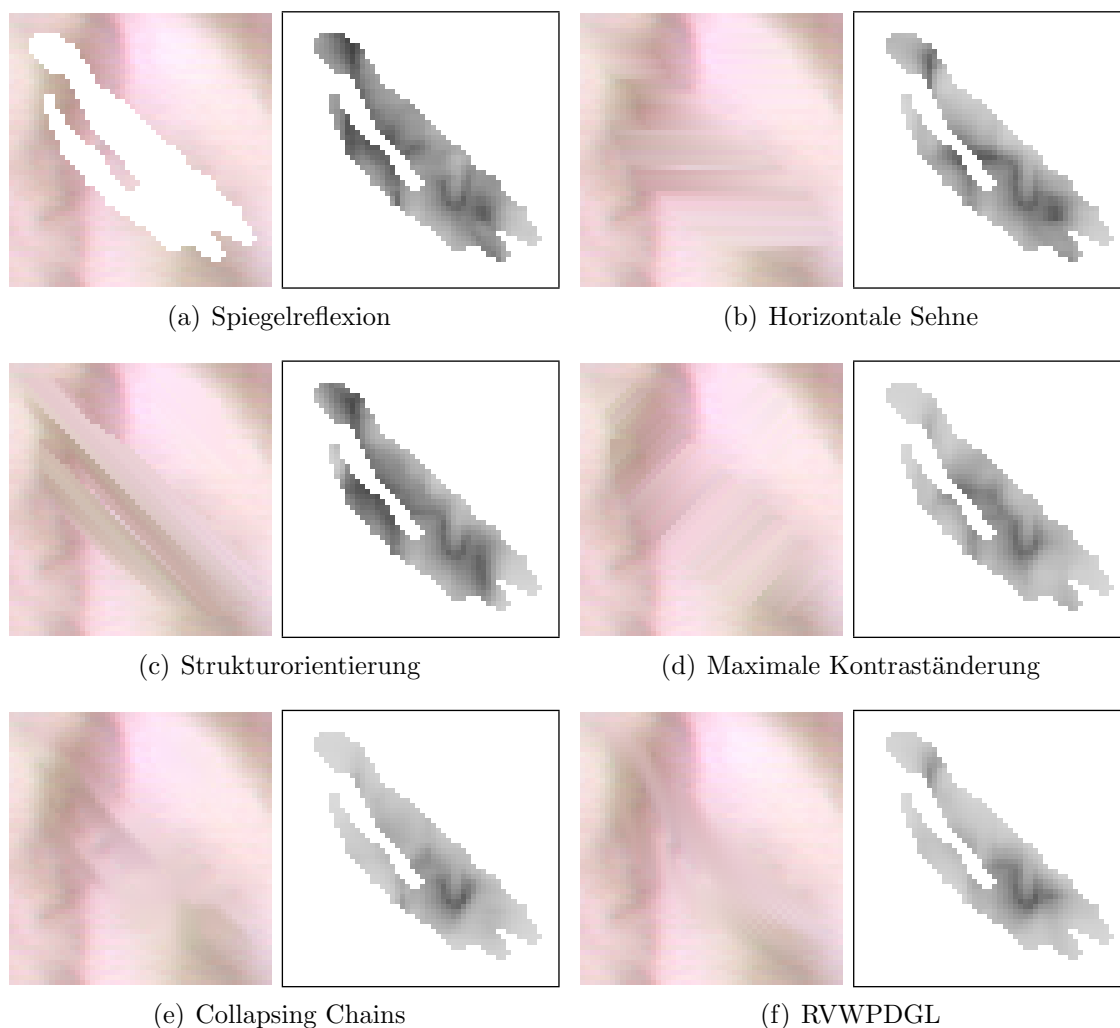


Abbildung 4.27: Ergebnisse der Rekonstruktion und entsprechend visuell aufbereitete ZNCC-scores der Vergleiche mit dem Original

Abbildung 4.27 zeigt für die vorgestellten Methoden jeweils die Ergebnisse der Rekonstruktion und die ZNCC-scores als Grauwertbilder. Um eine bessere regionale Differenzierbarkeit der Ergebnisse innerhalb der rekonstruierten Bereiche zu erhalten, wurde für jeden Bildpunkt in einer 11×11 Umgebung ein ZNCC-score ermittelt. Die gemittelten ZNCC-scores können dem Balkendiagramm aus Abbildung 4.28 entnommen werden. Das beste Korrelationsergebnis erzielt man mit dem Collapsing-Chains-Verfahren, dicht gefolgt von der RVWPDGL-Methode. Interessanterweise fällt die Linearinterpolationsmethode entlang der Strukturorientierung sogar hinter die Rekonstruktion entlang horizontaler Sehnen zurück.

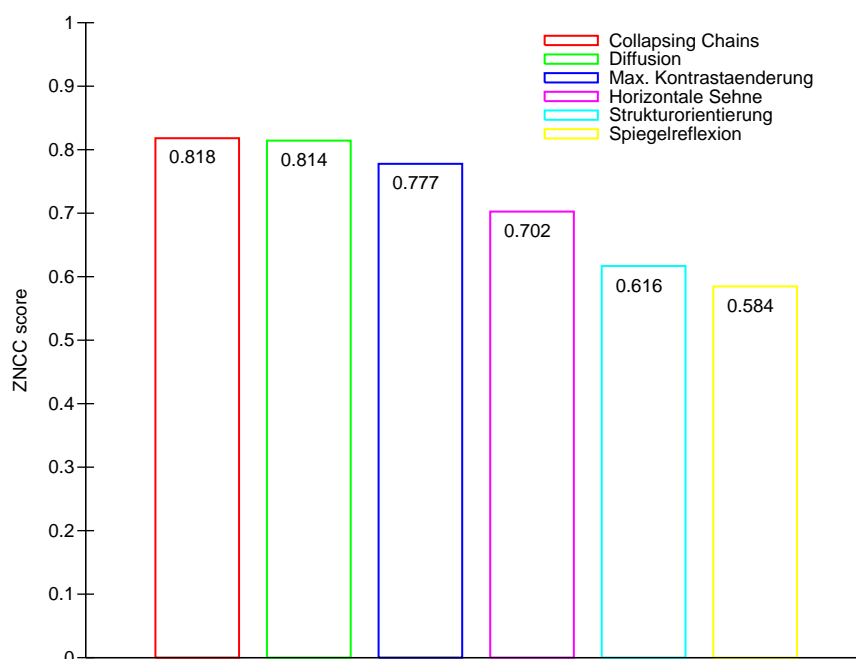


Abbildung 4.28: Mittelwerte der ZNCC-scores der Autokorrelation innerhalb des rekonstruierten Bereichs

Ein weiteres wichtiges Kriterium bei der Beurteilung der Spiegelreflexionsentfernung ist der Einfluss der jeweiligen Methode auf die Stereokorrelation. Sie bestimmt die Qualität der Erstellung der Disparitätskarte, anhand derer die 3D-Rekonstruktion des Operationsgebietes erfolgt. Bei diesem Vergleich dienten die Korrelationsscores zwischen dem intakten linken und rechten Bild innerhalb des künstlich eingefügten Artefakts als *ground-truth*-Daten. Bei einer Fenstergröße von ebenfalls 11 Pixel wurden die Korrelationsscores berechnet und mit Hilfe der Abbildungsfunktion aus Gleichung 4.16 als Grauwertbilder visualisiert. Auch hier liefern die RVWPDGL- und Collapsing Chains Methoden die besten Ergebnisse. Die Rekonstruktion mittels Linearinterpolation entlang einer Sehne in Richtung maximaler Kontraständerung übertrifft erneut die Methode basierend auf die Strukturorientierung.

Die Tabellen aus Abbildung 4.32 fassen die Ergebnisse der Vergleiche nochmal zu-

sammen, wobei Minimal- und Maximalwerte ebenfalls aufgelistet werden. Vor allem sind die Maximalwerte der Korrelationsscores der einzelnen Verfahren relevant, weil vorzugsweise ihnen zugeordnete Punkte aus der Disparitätskarte zur 3D-Rekonstruktion ausgewählt werden.

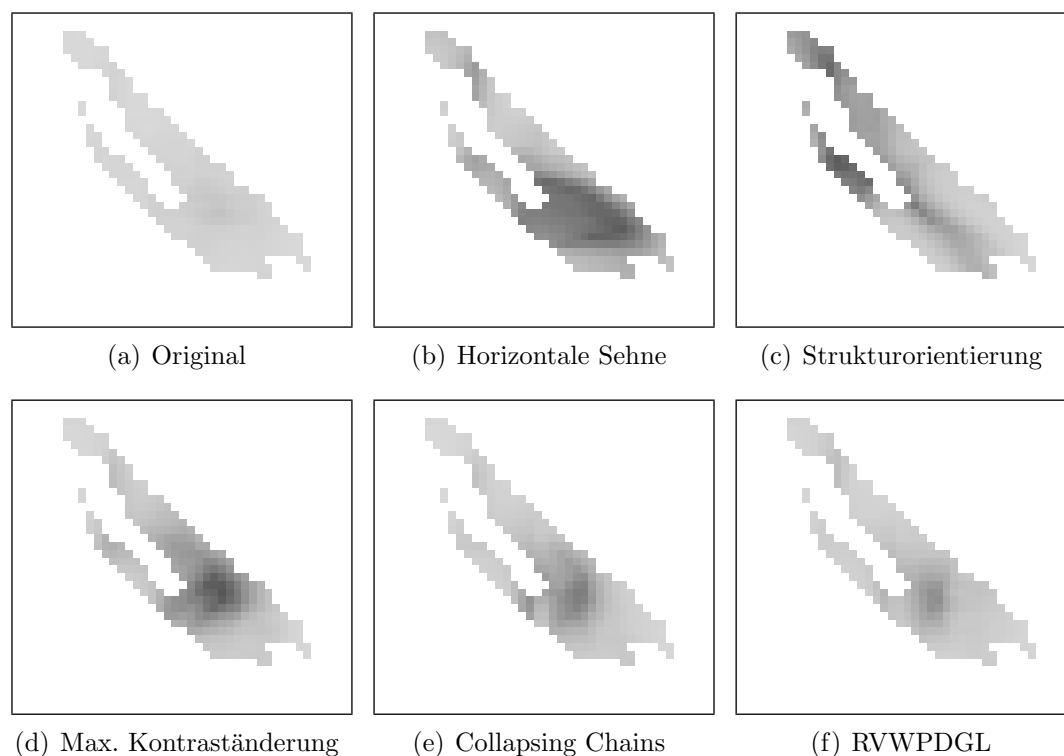


Abbildung 4.29: ZNCC-scores der Stereokorrelation

Zusammenfassend lässt sich sagen, dass die Methode der Collapsing Chains sowohl visuell betrachtet als auch empirisch gemessen bessere Ergebnisse liefert als alle Linearinterpolationsmethoden. Sie ist der RVWPDGL-Methode sowohl vom Erscheinungsbild her, als auch nach messbaren Kriterien ebenbürtig, jedoch bei einem erheblichen Geschwindigkeitsvorteil in der Ausführung. In der Tabelle aus Abbildung 4.30 werden die Laufzeitverhalten der verschiedenen Methoden gegenübergestellt.

Methode	Rekonstruktionsrate
<i>Horizontale Sehne</i>	55 fps
<i>Collapsing Chains</i>	45 fps
<i>Max. Kontraständerung</i>	25 fps
<i>Strukturorientierung</i>	25 fps
<i>RVWPDGL</i>	0.0027 fps

Abbildung 4.30: Ausführungsgeschwindigkeiten auf der Referenzhardware

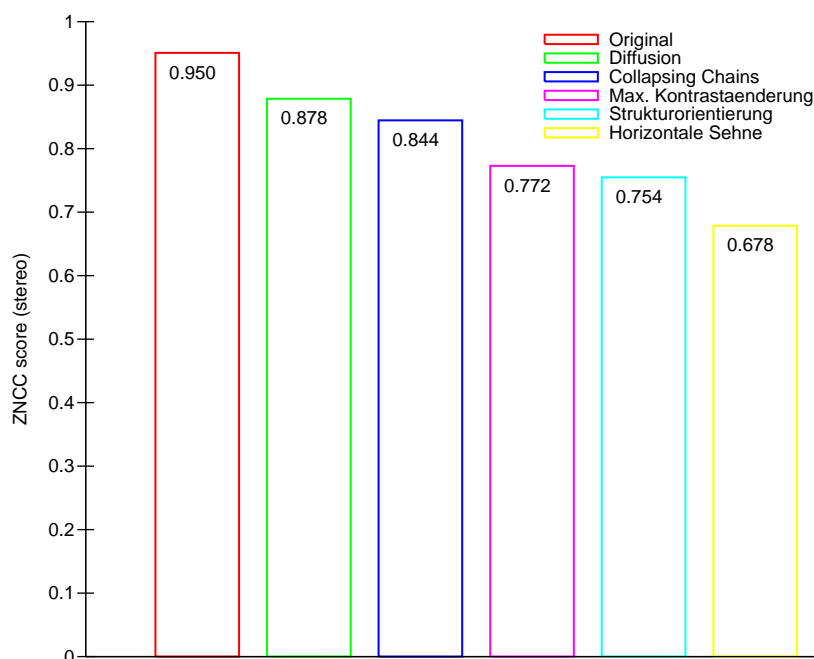


Abbildung 4.31: Mittelwerte der ZNCC-scores der Stereokorrelation innerhalb des rekonstruierten Bereichs

ZNCC-scores, Autokorrelation			
Methode	<i>Min.</i>	<i>Max.</i>	<i>Mtl.</i>
<i>Collapsing Chains</i>	0.315258	0.993718	0.818067
<i>RVWPDGL</i>	0.315258	0.990029	0.814221
<i>Max. Kontraständerung</i>	0.372020	0.979419	0.777773
<i>Horizontale Sehne</i>	0.276923	0.977959	0.702458
<i>Strukturorientierung</i>	0.184880	0.979073	0.616745
<i>Spiegelreflexion</i>	0.128426	0.948317	0.584669

ZNCC-scores, Stereokorrelation			
Methode	<i>Min.</i>	<i>Max.</i>	<i>Mtl.</i>
<i>Original</i>	0.871843	0.988610	0.950741
<i>RVWPDGL</i>	0.587839	0.969110	0.878477
<i>Collapsing Chains</i>	0.485630	0.983761	0.844557
<i>Max. Kontraständerung</i>	0.256095	0.977785	0.772815
<i>Strukturorientierung</i>	0.280209	0.977222	0.754906
<i>Horizontale Sehne</i>	0.295317	0.949230	0.678572

Abbildung 4.32: Ergebnisse der Auto- und Stereokorrelationen: Minimal-, Maximal- und Mittelwerte der ZNCC-scores

4.1.9 Visueller Vergleich anhand von Videosequenzen

In diesem Abschnitt hat der Leser die Möglichkeit, in der elektronischen Online-Version dieser Arbeit die vorgestellten Methoden anhand einer längeren Videosequenz zu vergleichen. Insbesondere wird in Abbildung 4.33 die Funktionsweise der Collapsing Chains und der RVWPDGL Methode anhand der Art der Fortsetzung von Strukturen ins Innere der Spiegelreflexion gegenübergestellt.

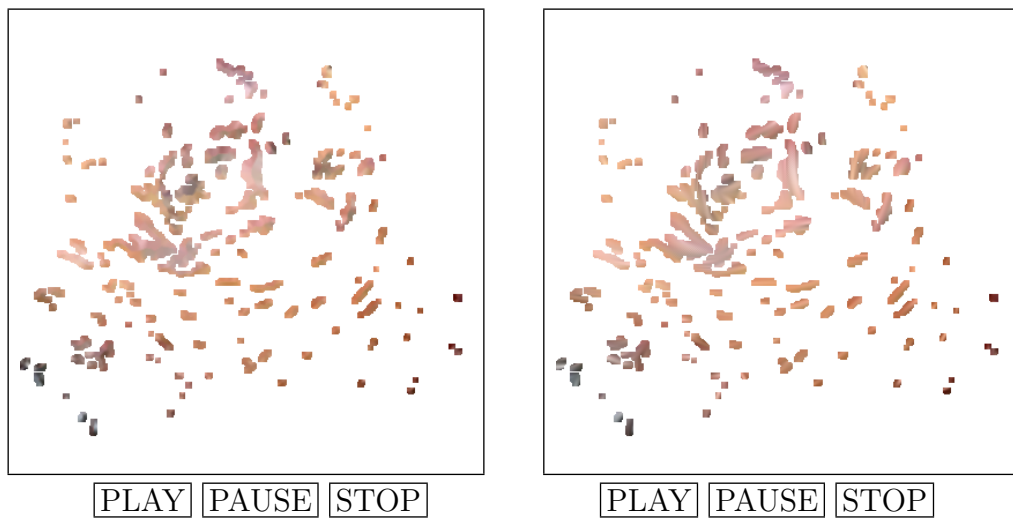


Abbildung 4.33: Collapsing Chains vs. RVWPDGL

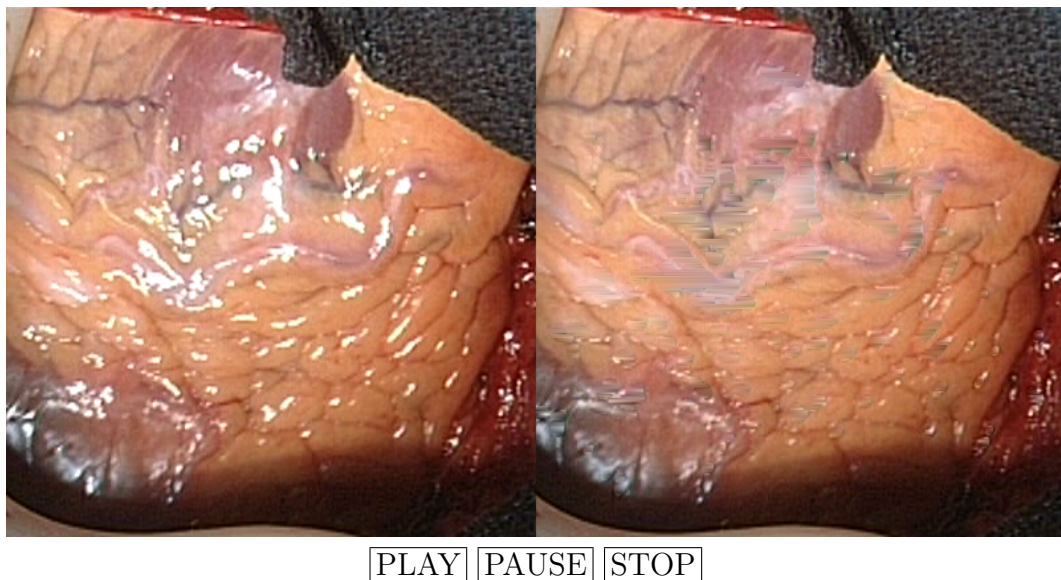
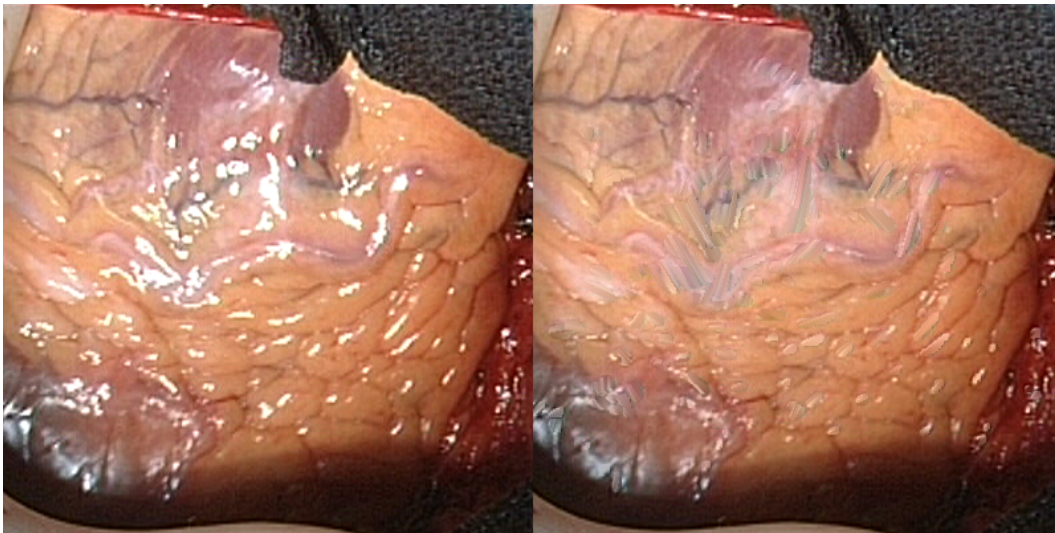
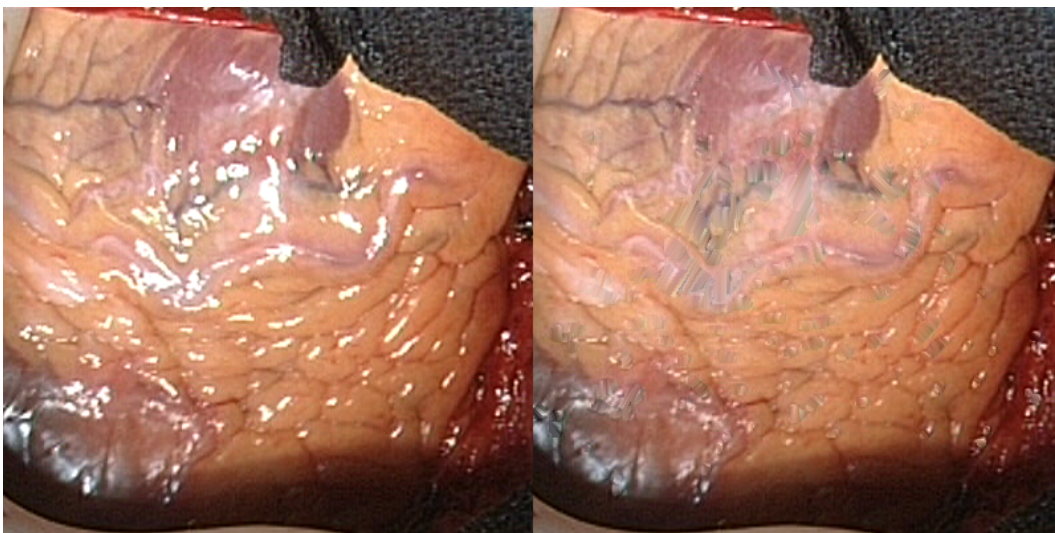


Abbildung 4.34: Entfernung von Spiegelreflexionen (Horizontale Sehne)



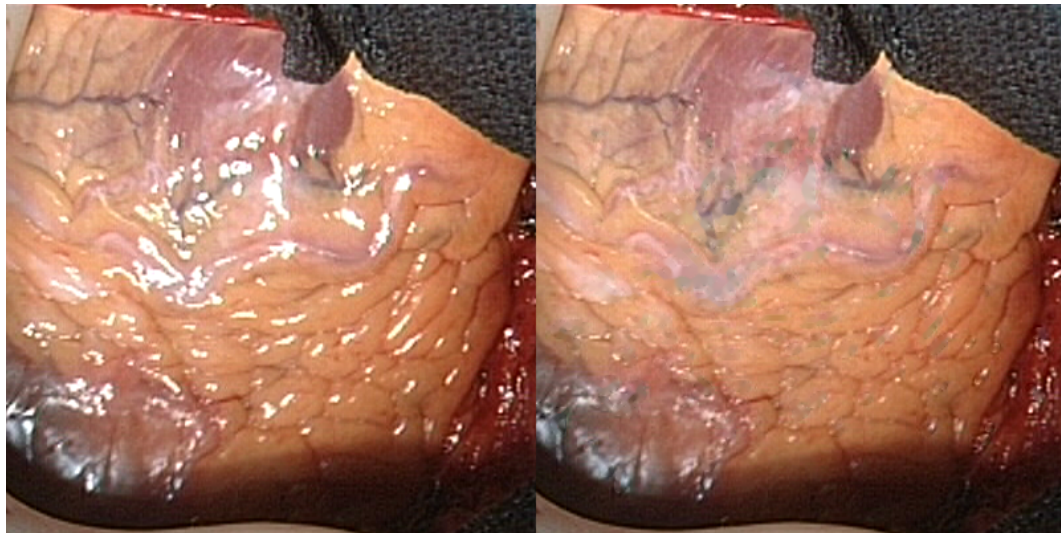
PLAY PAUSE STOP

Abbildung 4.35: Entfernung von Spiegelreflexionen (Strukturorientierung)



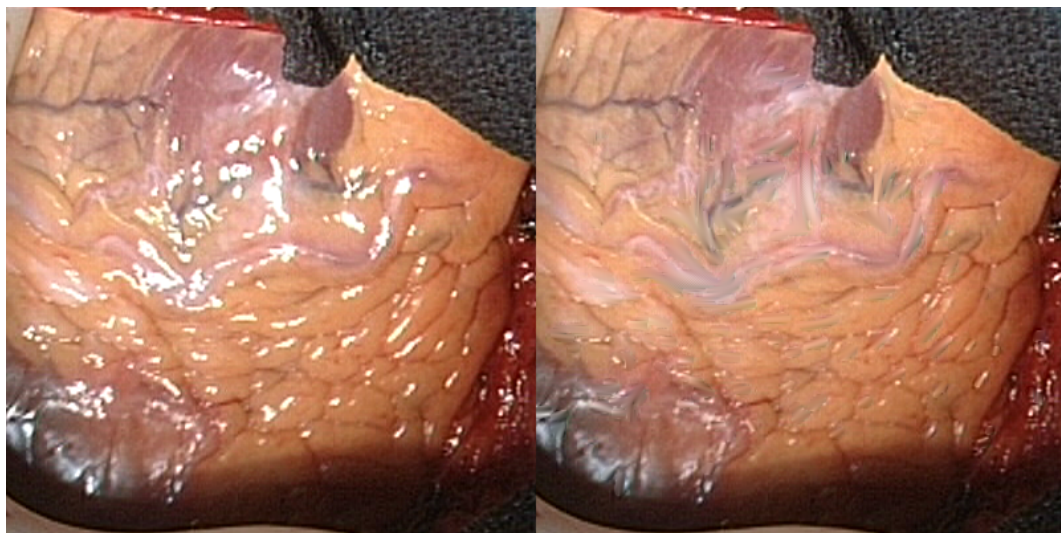
PLAY PAUSE STOP

Abbildung 4.36: Entfernung von Spiegelreflexionen (Max. Kontraständerung)



PLAY PAUSE STOP

Abbildung 4.37: Entfernung von Spiegelreflexionen (Collapsing Chains)



PLAY PAUSE STOP

Abbildung 4.38: Entfernung von Spiegelreflexionen (RVWPDGL)

4.2 Anwendung der Regularisierung von Bildern mittels PDGLen zur Rauschunterdrückung

Insbesondere bei analogen Endoskopiesystemen ist teilweise erhebliches Bildrauschen zu beobachten. Dies ist zum einen auf die kleinen CCD-Sensoren mit hoher Packungsdichte zurückzuführen. Zum anderen verursacht der nachgelagerte Digitalisierungsschritt einen gewissen Anteil an Quantisierungsrauschen. Aufgrund sehr feiner anatomischer Strukturen in medizinischen Aufnahmen sind klassische Verfahren zur Bildglättung wegen ihrer kantenzerstörenden Wirkung nicht geeignet. Im Abschnitt 4.1.7 wurde eine Methode zur Regularisierung von Bildern mittels partiellen Differentialgleichungen beschrieben, die zunächst für die Entfernung von Spiegelreflexionen eingesetzt wurde. Wir zeigen nun vollständigshalber auch die Möglichkeit der Anwendung zur Rauschunterdrückung, was eigentlich das spezifische Haupteinsatzgebiet solcher Verfahren ist.

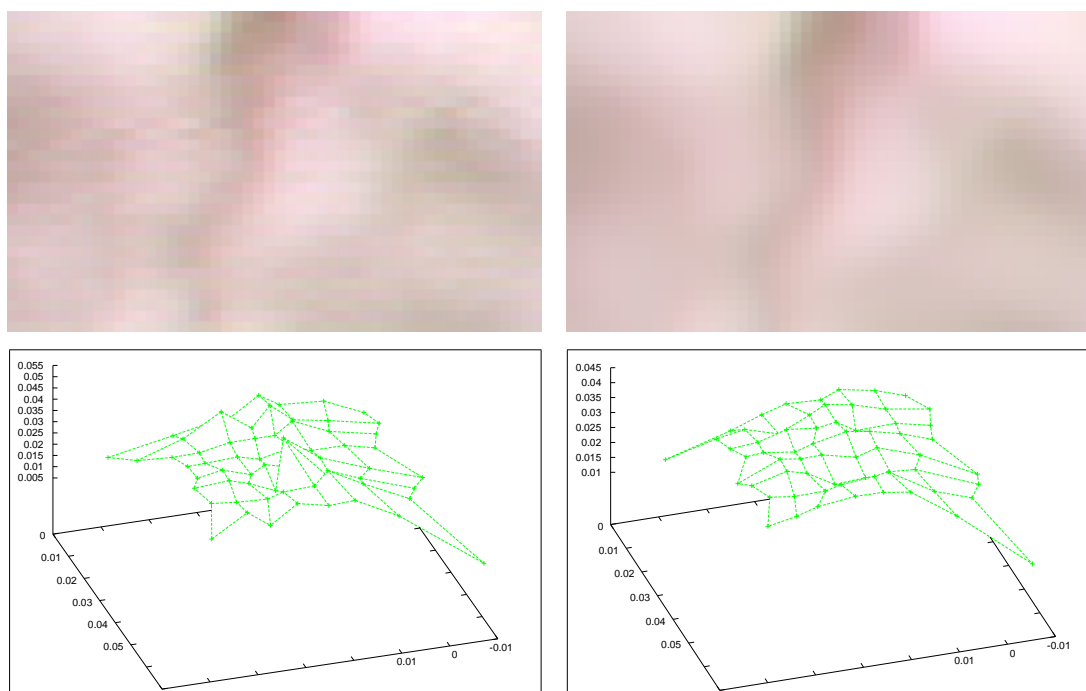


Abbildung 4.39: 3D-Rekonstruktion von Punktwolken vor und nach Regularisierung

Besonders hinderlich ist das Bildrauschen bei der 3D-Rekonstruktion von Oberflächen. Abbildung 4.39 zeigt die Auswirkung auf die Rekonstruktion eines Ausschnitts der Herzoberfläche. Es wurden jeweils für 64 Punkte aus einem rechteckigen Bereich (ca. 5×5 cm) 3D-Koordinaten ermittelt, und daraus eine polygonale Fläche erzeugt. Auf der linken Seite verursacht das verrauschte Bildmaterial ein zerklüftetes Höhenprofil der ansonsten größtenteils glatten Herzoberfläche. Nach der Regularisierung hingegen verschwinden starke Ausreißer (rechts), das Gewebe wird von der

4.2. ANWENDUNG DER REGULARISIERUNG VON BILDERN MITTELS PDGLEN ZUR RA

rekonstruierten Fläche besser angenähert. Wegen dem hohen Berechnungsaufwand der ursprünglichen Implementierung der Autoren [115] (über eine Minute für ein Stereobildpaar) ist die Anwendung in der regulären Steuerungsschleife eines Robotersystems nicht geeignet. Unter Zuhilfenahme der veröffentlichten Bibliothek wurde im Rahmen dieser Arbeit eine verteilte mehrprozessorfähige Variante entwickelt, die auf einem Rechnercluster bestehend aus 4×4 Prozessoren die Behandlung von einem Stereobildpaar pro Sekunde ermöglicht.

Dadurch ist das Verfahren zumindest in zwei Anwendungsfällen sinnvoll einsetzbar. Zum einen kann während eines Eingriffs zur Inspektion problematischer Befunde die Bildverbesserung kurzzeitig zugeschaltet werden, wodurch auch der Tiefeneindruck deutlich verbessert wird. Zum anderen kann zu Beginn eines Eingriffs anhand von aufbereiteten Aufnahmen ein genaues Übersichtsmodell vom Operationsgebiet mit qualitativ hochwertiger Texturierung erstellt werden. Zum Schluss zeigt Abbildung 4.40 anhand einer längeren Sequenz das Potential dieser Methode.



PLAY PAUSE STOP

Abbildung 4.40: Rauschunterdrückung mittels Regularisierung

Kapitel 5

Fadenerkennung

Zusammenfassung Obwohl seit fast 10 Jahren für den klinischen Einsatz verfügbar, erreichte die robotergestützte minimalinvasive Chirurgie bis zum Zeitpunkt der Erstellung dieser Arbeit keinen allgemeinen Durchbruch. Lediglich für gewisse Eingriffe im Bauch und kleinen Becken häufen sich die Erfolgsmeldungen. Die dringendste Notwendigkeit Eingriffe minimalinvasiv durchzuführen ergibt sich jedoch in der Thoraxchirurgie, insbesondere Herzchirurgie, weil konventionelle Operationstechniken durch das vollständige Öffnen des Brustkorbs hier das größte Trauma verursachen. Die Gründe für das Meiden der Technik in der Herzchirurgie sind allem Anschein nach die fehlende Taktile und haptische Rückkopplung, und der erhöhte Zeitaufwand der ohnehin langen Eingriffe. Der Zeitbedarf ergibt sich einerseits aus Planungs- und Vorbereitungsphase, währenddessen sich der Patient schon in der belastenden Narkose befinden muss. Zeitaufwendige und motorisch komplexe Aufgaben verlängern ebenfalls den Eingriff, wobei der Patient an eine Herz-Lungen-Maschine angeschlossen ist, die ihrerseits auch negative Auswirkungen haben kann. Durch eine Verkürzung der Operationsdauer könnte die Akzeptanz der robotergestützten Chirurgie erhöht werden. Um das zu erreichen wurde die zeitaufwendigste, aber gleichzeitig gut automatisierbare Aufgabe, das Knotenlegen, exemplarisch ausgesucht, zumal die häufigsten Eingriffe am Herzen, Rekonstruktion und Ersatz der Herzklappen und Bypass legen, sehr nahintensiv sind.

In der minimalinvasiven Chirurgie beschränkt man sich notgedrungen meistens auf nur eine bestimmte Art von Instrumentenknoten. Dabei muss das Nahtmaterial nur ein oder zweimal um den Greifer gewickelt, und einmal das Fadenende gegriffen und angezogen werden. Der Einstich der Nadel sollte aus offensichtlichen Gründen immer dem menschlichen Operateur vorbehalten bleiben, allerdings sind Umwickeln, Greifen und Ziehen gut automatisierbar. Doch dafür muss dem Robotersystem die genaue Lage des Fadens jederzeit bekannt sein. Die Wickelvorgänge können zwar im optimalen Fall "blind" vom System durchgeführt werden, jedoch für das Vollenenden des Knotens durch Anziehen muss das Fadenende gegriffen werden. Zusätzliche Informationen, wie die relative Lage zwischen Nahtmaterial, Gewebe und Greifer

sind ebenfalls notwendig, um das Gewebstrauma zu minimieren. Die Voraussetzung hierfür ist das Vorhandensein eines Stereoendoskops, mit Hilfe dessen sowohl die Erkennung als auch die dreidimensionale Rekonstruktion der relevanten Teile des Operationssitus durchführbar ist. Eine über die Bildverarbeitung geschlossene Steuerungsschleife Sensorik \leftrightarrow Aktorik gewährleistet, dass auf alle Veränderungen in der Szene, Fehlerfälle eingeschlossen, reagiert werden kann, wofür die robuste Segmentierung des Fadens unabdingbar ist.

Chirurgisches Nahtmaterial gibt es hauptsächlich in den Farben blau und grün, wobei die Fadenstärke durch verschiedene Farbtöne der Grundfarben kodiert wird. Eine Segmentierung ausschließlich über die Farbe erweist sich aus zwei Gründen als nicht praktikabel. Zum einen liefern Endoskope entfernungs- und belichtungsabhängig starke chromatische Abweichungen, sodass nur aufwendige, endoskopspezifische Farbkalibrierungen zufriedenstellende Ergebnisse liefern. Zum anderen sorgen anatomische Strukturen auf der Herzoberfläche oft für falsche Zuordnungen. Die nächste naheliegende Möglichkeit sind Kantenfilter, wobei es in der Fachliteratur eine ganze Reihe verschiedener Ansätze beschrieben sind, welche sich grob in drei Kategorien einordnen lassen. Der erste Ansatz detektiert Kanten anhand von lokalen Grauwertunterschieden im Bild. Der Nachteil solcher einfachen Verfahren ist, dass die zahlreichen Falschzuordnungen mittels rechenaufwendiger Verfahren entfernt werden müssen. Außerdem bieten sie keine Subpixelgenauigkeit. Der zweite Ansatz betrachtet Linien als Objekte mit parallelen Rändern. Zunächst wird hierbei die lokale Richtung einer Linie für jeden Pixel ermittelt. Dann wendet man zwei Kantenfilter in der Richtung senkrecht auf die Linie an, wobei sie auf die Erkennung der jeweils linken und rechten Seite optimiert sind. Der Nachteil dabei ist, dass rechenintensive direktionale Kantenfilter notwendig sind. Die dritte Kategorie betrachtet Kanten als Schluchten und Kämme in der Bildfunktion und extrahiert sie unter Zuhilfenahme verschiedener geometrischer Eigenschaften. Der Hauptvorteil solcher Verfahren ist, dass sowohl Kanten, als auch ihre Breiten subpixelgenau ermittelt werden können. Weil chirurgische Fäden auch im vergrößerten Endoskopbild oft nur einige Pixel breit sind, empfiehlt es sich ein Verfahren dieser Art anzuwenden.

Die hier vorgestellte Fadenerkennung erfolgt mit Hilfe eines mehrstufigen Ansatzes, an dessen Anfang ein subpixelgenauer Kantenfilter steht. Grobe Ausreißer können bereits in diesem frühen Stadium über die Kantenbreite eliminiert werden. Für die zweite Stufe wird die Berechnung einer Disparitätskarte notwendig, welche später auch für die Positionsbestimmung herangezogen wird. Die Tiefeninformation dient als zusätzliches Validierungskriterium. Die Forderung der Stetigkeit des Disparitätsverlaufs entlang des Fadens ist durch die materielle Beschaffenheit motiviert und vermeidet etwaige Ausreißer. Insgesamt ergibt sich eine robuste Segmentierung des Fadenverlaufs, der betriebene Aufwand wirkt sich jedoch negativ auf die Ausführungsgeschwindigkeit aus.

5.1 Motivation

Binokulare Sicht ist für herkömmliche minimalinvasive endoskopische Eingriffe nicht zwingend erforderlich [81, 82] und hat sich trotz der Vorteile nicht durchgesetzt. Der Hauptgrund dafür ist wahrscheinlich die problematische Integration der Geräte zur Bilddarstellung in die Operationsräume. Robotergestützte Systeme für minimalinvasive Chirurgie hatten gleich von den Anfängen diese Technik integriert, und berücksichtigten dabei auch ergonomische Aspekte. Durch die kinematische Integration eines Stereoendoskops in das Robotersystem ergeben sich weitere Vorteile und Möglichkeiten, welche über die bessere Darstellungsqualität und Ortsauflösung hinausgehen. Stattet man die üblicherweise ferngesteuerten Arme eines solchen Systems mit (teil-) autonomen Fähigkeiten aus, sind Positions- und Größenangaben über das Operationsgebiet absolut notwendige Voraussetzungen für einen sicheren Betrieb. Diese Daten können jedoch nur über binokulare Sicht in ausreichender Genauigkeit zur Verfügung gestellt werden.

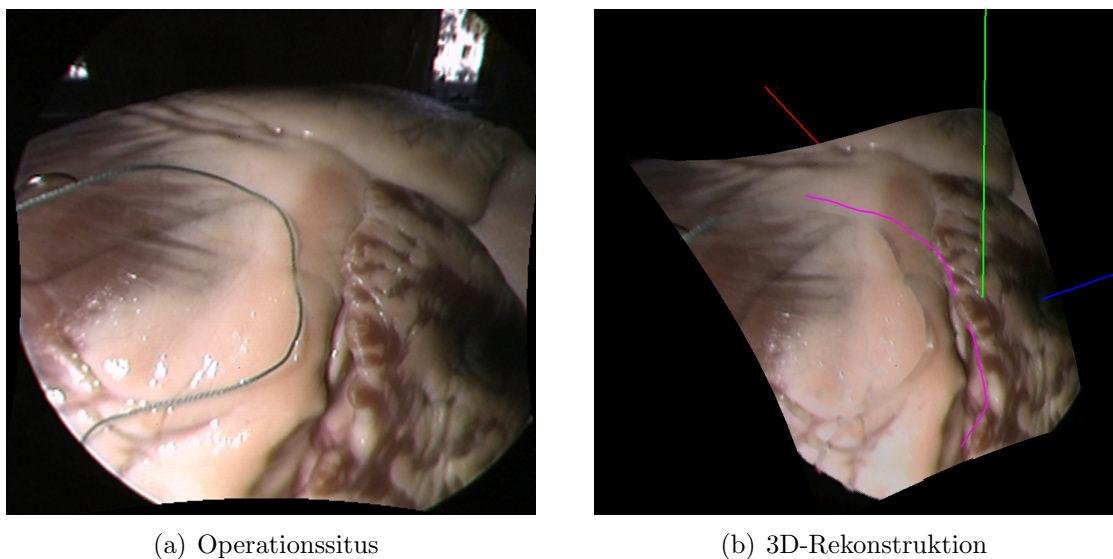


Abbildung 5.1: 3D-Rekonstruktion von Nahtmaterial und Herzgewebe anhand von binokularen Endoskoptaufnahmen

Abbildung 5.1(a) zeigt eine typische Szene aus der Herzchirurgie mit einem Polyester-Faden im Vordergrund und Herzgewebe als Hintergrund. Die Aufnahme wurde mit einem für minimalinvasive chirurgische Eingriffe zugelassenen 3D-Endoskop der Firma Richard Wolf [92] erstellt, die Szene zeigt jedoch einen offenen Situs eines Schweineherzens.

Durch die rasante Entwicklung der bildgebenden medizinischen Verfahren, und der digitalen Bildverarbeitung im allgemeinen, erscheint die Integration eines Subsystems für Bildverarbeitung in robotergestützte Chirurgesysteme als sinnvoll. Die

Aufgabe eines solchen Subsystems könnte unter anderem darin bestehen, anhand von Bildpaaren eines Stereoendoskops ein 3D-Modell des Operationsgebietes zu erstellen, und möglichst oft zu aktualisieren. Die Bereitstellung der Lage des Nahtmaterials und des Gewebshintergrunds im Koordinatensystem der Instrumente würde teilautonome Techniken ermöglichen und die Sicherheit eines solchen Systems erhöhen.

Wir greifen etwas vor, und zeigen in Abbildung 5.1(b) das Ergebnis einer 3D-Rekonstruktion: das Nahtmaterial wurde segmentiert und als Freiformkurve approximiert. Die Organoberfläche wurde ebenfalls mit Hilfe einer Freiformfläche angenähert und texturiert. Sowohl Spiegelreflexionen als auch Nahtmaterial wurden aus dem Texturbild mit Hilfe des bereits vorgestellten Collapsing Chains Verfahrens entfernt, der Faden unter Zuhilfenahme einer entsprechenden Maske, die bei der Segmentierung erstellt wurde. In den folgenden Abschnitten werden die einzelnen Verarbeitungsschritte, die zur Erstellung eines 3D-Modells des Operationssitus notwendig sind, ausführlich beschrieben. Die am Anfang einer solcher Kette befindliche Bildvorverarbeitung haben wir in diesem Kontext bereits behandelt, die nächste Stufe stellt die Segmentierung des Nahtmaterials dar. Doch zunächst stellen wir im folgenden Abschnitt einige der am häufigsten verwendeten Nahtmaterialien in der (minimalinvasiven) Herzchirurgie vor.

5.2 Chirurgisches Nahtmaterial

Dem Chirurgen stehen heutzutage für die Magendarmtrakt-, Gefäß-, Muskel-, Faszien-, Subkutan- und Hautnaht eine ganze Reihe verschiedener Fäden in adäquaten Stärken zur Verfügung. Man unterscheidet zwischen nichtresorbierbaren, resorbierbaren und schnell resorbierbaren Fäden, beschichtet oder unbeschichtet, geflochten multifil, pseudomonofil oder monofil [116]. Die zwei wichtigsten Merkmale, nach denen chirurgische Fäden kategorisiert werden, sind die Resorbierbarkeit und der Fadenaufbau. Nichtresorbierbare Nahtmaterialien umfassen Fäden, welche im Gegensatz zu resorbierbaren im Gewebe persistieren, und oft noch nach Jahren im Gewebe nachzuweisen sind. Nichtresorbierbare Materialien dienen in der Chirurgie zum Verschluss von Wunden, die entweder auf die langfristige Adaptation durch diese angewiesen sind (z.B. Gefäßanastomosen, Schrittmacherimplantate) oder deren Nahtentfernung nach entsprechender Heilungszeit möglich ist (z.B. Hautnähte).

Eigenschaften chirurgischer Nahtmaterialien

Der Fadenaufbau, darunter ist die Struktur der chirurgischen Fadens zu verstehen, hat ebenfalls einen bedeutenden Einfluss auf den Anwendungsbereich und die Einsetzbarkeit eines bestimmten Nahtmaterials. Drei Filamentarchitekturen werden heute in der Chirurgie eingesetzt: multifilamentär, monofilamentär und pseudomonofilamentär. Einen multifilamentären Faden erhält man durch Verflechten dünner

Garne, die Einzelfäden liegen dabei mehr oder weniger quer zur Fadenlängsachse. Der Hauptvorteil dieser Technik ist, dass die Zugfestigkeit die Summe der einzelnen Fäden erreicht, nicht aber die Härte, die einem soliden Faden von gleichem Querschnitt zukäme. Geflochtene nichtresorbierbare Polyesterfäden werden in der Herzchirurgie insbesondere bei Klappenoperationen eingesetzt.

Monofile Fäden bestehen aus einem Faden, besitzen eine glatte Oberfläche und im Vergleich zu multifilen Fäden des gleichen Rohstoffes eine höhere Drahtigkeit. Bei dickeren Fäden verschlechtert diese Drahtigkeit, die alle monofilen Fäden aufweisen, die Handhabung, insbesondere lassen sie sich schlechter kneten. Die Bereitstellung feiner monofiler Fäden war neben der Verfügbarkeit von optischen Hilfen sowie geeigneten chirurgischen Instrumenten Voraussetzung für die Verbreitung der Mikrogefäßchirurgie. Monofile Fäden aus Polypropylen werden in der Herzchirurgie insbesondere bei Gefäßanastomosen verwendet.

Zwischen den beiden Gruppen monofiler und polyfiler Fäden stehen pseudomonofile Fäden, die wir nur vollständigheitshalber erwähnen, da sie in der Herzchirurgie keine Anwendung finden. Hier besteht das Fadeninnere aus einem multifilen Geflecht, welches einen mantelartigen Überzug erhält. Diese glatte Oberfläche kann aus dem selben Rohstoff wie das Innere bestehen, oder es wird nach Fertigstellung als eine Art Beschichtung aufgesprüht.

Die meisten chirurgischen Fäden erhalten eine charakteristische Färbung, zum einen um bessere Sichtbarkeit im Gewebe zu erreichen, aber auch um die verschiedenen Fadenstärken zu kennzeichnen. Die Bezeichnungen für Fadenstärken (3-0, 4-0, ...) waren anfangs nach der United States Pharmacopea (Amerikanisches Arzneibuch, U.S.P.) teilweise willkürlich gewählt und weitergegeben worden. Viel einfacher ist die Stärkenbezeichnung nach der Europäischen Pharmacopoe, welche ab dem 1.1.1976 die Fadenstärken in Zehntelmillimeter angibt und *metric* genannt wird (z.B. *metric* 0.7 = $7 \cdot 1/10$ mm = 0.07 mm \emptyset). Um Verwechslungen auszuschließen, geben die Hersteller auf allen Packungen die Fadenstärken in *metric* und in der herkömmlichen U.S.P.-Bezeichnung an.

Verwendete Fäden

In der Tabelle aus Abbildung 5.2 werden die in dieser Arbeit verwendeten chirurgischen Fäden aufgelistet. Ausgesucht wurden die in der Herzchirurgie hauptsächlich eingesetzten monofilen und geflochtenen Fäden aus Polypropylen und Polyester in verschiedenen Stärken. Die Angaben über den Durchmesser weichen jedoch in der Realität oft von der Norm ab [117], sodass sie sich nur bedingt für die Parametrierung von Segmentierungsalgorithmen eignen. Abbildung 5.3 zeigt endoskopische Aufnahmen dieser Fäden mit einem realistischen Gewebshintergrund.

Augenscheinlich können alle Fäden, bis auf den 7-0er Faden (0.05 mm \emptyset) aus 5.3(g), gut vom Hintergrund separiert werden, trotz ähnlicher Strukturen auf der Herzoberfläche. Einige wichtige Merkmale helfen einem menschlichen Betrachter bei dieser

Bezeichnung	Hersteller	Material	U.S.P.	\emptyset
ETHIBOND EXCEL	J&J Intl.	Polyester (geflochten)	3-0	0.2 mm
ETHIBOND EXCEL	J&J Intl.	Polyester (geflochten)	5-0	0.1 mm
PROLENE	J&J Intl.	Polypropylen (monofil)	3-0	0.2 mm
PROLENE	J&J Intl.	Polypropylen (monofil)	4-0	0.15 mm
PROLENE	J&J Intl.	Polypropylen (monofil)	5-0	0.1 mm
–	Lukens	Polypropylen (monofil)	6-0	0.07 mm
PROLENE	J&J Intl.	Polypropylen (monofil)	7-0	0.05 mm

Abbildung 5.2: Verwendete Nahtmaterialien aus Abbildung 5.3

Aufgabe: der typische Verlauf, die scharfe Abgrenzung des Linienprofils (Umgebung und Tiefe) und die Farbe. Nun gilt es diese semantischen Vorgaben in maschinennahe Spezifikationen umzusetzen. Es hat sich herausgestellt, dass die Berücksichtigung nur eines der drei weiter oben erwähnten Merkmale allein nicht ausreicht, um den Fadenverlauf robust zu segmentieren. Dies motiviert einen hybriden mehrstufigen Ansatz, den wir im folgenden grob skizzieren, und dessen Stufen in den nächsten Abschnitten detailliert darstellen.

Am Anfang steht ein subpixelgenauer Kantenfilter, der in der Lage ist, auch die Linienbreiten präzise zu extrahieren. Dadurch können grobe Ausreißer bereits in einem frühen Stadium eliminiert werden. In einer nächsten Stufe wird der Farbraum der Szene betrachtet, wobei die typische Farbgebung chirurgischen Nahtmaterials als Ausschlusskriterium für falsche Zuordnungen dienen kann. Die dritte Stufe ist auf die Berechnung einer Disparitätskarte angewiesen, welche auch für die nachgelagerte 3D-Rekonstruktion herangezogen wird. Die Tiefeninformation entlang segmentierter Strukturen dient als zusätzliches Validierungskriterium. Die Forderung nach Stetigkeit des Disparitätsverlaufs entlang des Fadens ist durch dessen materielle Beschaffenheit motiviert, und soll etwaige verbliebene Ausreißer erkennen.

5.3 Kantenfilter

Die Segmentierung von Kanten ist eine der häufigsten Operationen in der digitalen Bildverarbeitung. In den meisten Fällen detektieren Kantenfilter Diskontinuitäten der räumlichen Grauwertfunktion einer Bildebene. Solche Diskontinuitäten können zum einen durch einfache Übergänge zwischen verschiedenen homogenen Bereichen entstehen, denen allerdings kein reales Objekt zugeordnet werden kann. Zum anderen werden auch die Verläufe von "eindimensionalen" Strukturen, die wir im folgenden als Linien bezeichnen werden, oft als Kanten bezeichnet. Es gibt keinen Universalfilter für Kanten, weil die Ziele auf der Merkmalsebene, die ein Kantenoperator erfüllen muss, nicht exakt und allgemein spezifizierbar sind. Infolgedessen wurden in der Literatur eine große Anzahl von Operatoren publiziert, jeweils auf einen bestimmten

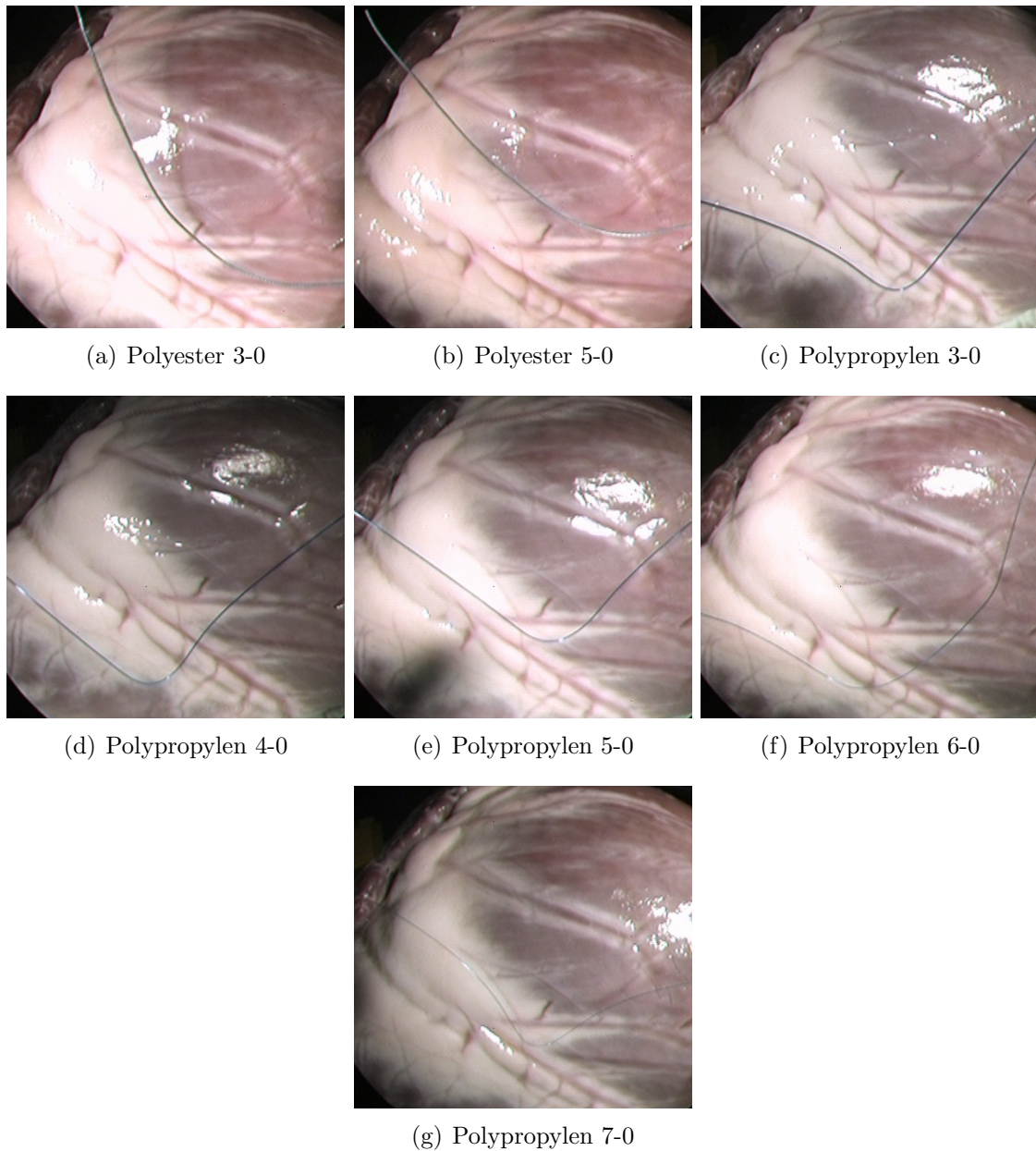


Abbildung 5.3: Nichtresorbierbare geflochtene (*ETHIBOND*[©] EXCEL, a-b) und monofile (*PROLENE*[©], c-e, g) Fäden in verschiedenen Stärken der Firmen Johnson & Johnson Intl. [118] und Lukens Medical Corporation

Anwendungsbereich zugeschnitten. Die symbolischen (semantischen) Zielvorgaben, die sich bei der Suche an menschlichen Spezifikationen orientieren, müssen von Fall zu Fall in numerische Gütevorgaben transformiert werden.

In dieser Arbeit wird ausschließlich die Segmentierung von dünnen Linien behandelt, d.h. Abbildungen von real existierenden Objekten, deren Breite jedoch nur wenige Bildpunkte beträgt. Konkret wollen wir uns der Erkennung und symbolischer Repräsentation von chirurgischem Nahtmaterial in endoskopischen Aufnahmen widmen. Diese konkrete Domäne erlaubt anhand der physikalischen Eigenschaften chirurgischer Fäden eine weitere wichtige Annahme, und zwar einen kurvilinearen Verlauf nahezu konstanter Breite. Es gibt in der Literatur bisher keine Arbeiten, die sich explizit mit der Segmentierung von chirurgischem Nahtmaterial beschäftigen. Die Erkennung kurvilinearere Strukturen spielt jedoch in anderen Bereichen, wie in der Analyse von Satellitenaufnahmen (Straßen und Wege wie in [119]) oder der Medizin (Blutgefäße in Röntgenaufnahmen wie in [120]), eine wichtige Rolle. Dementsprechend gibt es eine Reihe von Veröffentlichungen, wobei Steger [121] die vorgestellten Ansätze in drei Kategorien einteilt.

Zur ersten Kategorie gehören Verfahren, die im Grauwertbild mit differentiellen Operatoren arbeiten. Sie betrachten ausschließlich lokale Grauwertunterschiede, und sind entsprechend ungenau. Sie neigen außerdem zu Falschzuordnungen, deren Entfernung oft aufwendiger ist, als die Segmentierung selbst [122].

Der zweite Ansatz betrachtet Linien als Objekte mit parallelen Rändern [123]. Zunächst wird hierbei die lokale Richtung einer Linie für jeden Pixel ermittelt. Dann wendet man zwei Kantenfilter in der Richtung senkrecht auf die Linie an, wobei sie auf die Erkennung der jeweils linken und rechten Seite optimiert sind. Der Nachteil dabei ist, dass rechenintensive direktionale Kantenfilter notwendig sind.

Bei der dritten Kategorie von Filtern wird das Grauwertbild als eine zweidimensionale Funktion $g(x, y)$ aufgefasst, wobei der Wert an der Koordinate (x, y) der Höheninformation eines Geländes entspricht. Kanten werden bei dieser Analogie als Schluchten und Käme einer Gebirgslandschaft betrachtet, und unter Zuhilfenahme verschiedener differentialgeometrischer Eigenschaften extrahiert [124]. Die Ermittlung sowohl eines subpixelgenauen Verlaufs, als auch der Breite der Kanten ist möglich.

Im Folgenden werden wir diese dritte Kategorie genauer betrachten, und die drei am weitesten in der Literatur verbreiteten Unterkategorien vorstellen. Eine erste einfache Möglichkeit besteht darin, auf den Höhenlinien eines Bildes die Punkte mit maximaler Krümmung zu suchen und anschließend zu einer kontinuierlichen Kante zu verbinden [125]. Dafür extrahiert man die Höhenlinien explizit, und sucht nach Maxima, was jedoch bei Kanten mit einem flachen Linienprofil fehlschlägt. Auch die zweite Subkategorie, die nach einem lokalen Maximum einer der zwei Hauptkrümmungen sucht, versagt bei flachen Linienprofilen [126]. Bei der dritten Unterkategorie wird zunächst die Bildfunktion mittels eines Taylor-Polynoms zweiten

oder dritten Grades angenähert. Die Koeffizienten werden dabei unter Verwendung des Facet-Modells mit Hilfe der Methode der kleinsten Quadrate aus einer Pixelnachbarschaft entsprechender Größe berechnet [127]. Die Richtung einer Kante wird anschließend anhand der Hesse-Matrix des Taylor-Polynoms bestimmt. Die einzelnen Punkte einer Kante sind diejenigen Pixel, die ein lokales Maximum in der zweiten Richtungsableitung senkrecht zur Linienrichtung besitzen.

Der Hauptvorteil dieses Ansatzes ist, dass subpixelgenaue Verläufe ermittelt werden können, ohne dabei direktionale Filter anwenden zu müssen. Die bei der Berechnung der Taylor-Koeffizienten verwendeten Faltungsmasken sind jedoch für die Schätzung der ersten und zweiten partiellen Ableitungen eher mangelhaft. Dadurch können, insbesondere bei größeren Masken (> 5), für eine Linie mehrere Antworten entstehen, was die Breite segmentierbarer Strukturen auf die Maskengröße beschränkt. Abhilfe schafft in solchen Fällen die Verwendung von Gauß-Masken [128], sodass sie mittels der Wahl des Parameters σ auf bestimmte Kantenbreiten abgestimmt werden können. Dadurch dass die Umgebung der Kanten nicht explizit modelliert wird, schwindet jedoch die Positionsgenauigkeit der Segmentierung für steigende Werte von σ .

Nur wenige Verfahren ermitteln sowohl den genauen Verlauf, als auch die Breite einer Linie, wobei die meisten das durch Iterieren im Skalenraum bewerkstelligen [123]. Der Skalenraum, entstanden durch Hinzunahme der gefalteten Bilder zum Ursprungsbild, liefert ein kontinuierliches Spektrum in Abhängigkeit von der Varianz σ , die somit ein Skalenraumparameter ist. Ein zunehmender Skalenraumparameter lässt Strukturen immer weiter verschwinden, sodass gezielt nach Linien bestimmter Breite gesucht werden kann. Da die Iteration durch den Skalenraum sehr rechenaufwendig ist, wird diese Vorgehensweise selten angewendet.

Von allen vorgestellten Ansätzen der dritten Kategorie schien das Verfahren von Steger [121, 124] die besten Voraussetzungen zur robusten und präzisen Segmentierung von chirurgischem Nahtmaterial zu bieten. Dies ist in erster Linie darauf zurückzuführen, dass es zusätzlich zum differentialgeometrischen Ansatz ein explizites Modell der zu segmentierenden Linien verwendet. Dadurch dass Gauß-Masken für die Berechnung der Ableitungen benutzt werden, skaliert die Methode für beliebige Linienbreiten, ohne dabei Mehrfachantworten zu generieren. Die Modellierung der Wechselwirkung zwischen Linie und ihre Ränder erlaubt außerdem eine subpixelgenaue Extraktion der Lage und Breite segmentierter Strukturen.

Ursprünglich für die Erkennung von Straßen und Wegen in Luftbildern entworfen, eignet sich die Methode auch für die Segmentierung anatomischer Strukturen, insbesondere Gefäße, aus Röntgenaufnahmen. Das Gemeinsame in diesen Anwendungsbereichen ist das bekannte charakteristische Linienprofil der zu segmentierenden Strukturen. Über ähnliche Voraussetzungen verfügt man bei der Segmentierung von chirurgischem Nahtmaterial. Sowohl Linienprofil, als auch die Verlaufseigenschaften chirurgischer Fäden sind a priori bekannt. In gewissem Umfang stehen sogar Hinweise auf die zu erwartende Linienbreite zur Verfügung, die sich aus dem bekannten

Fadendurchmesser und der geschätzten Entfernung zur Linse aus den kalibrierten Parametern des Stereoendoskops berechnen lassen. Wir stellen im nächsten Abschnitt das Verfahren vor, und zeigen anschließend ihre Anwendbarkeit auf unsere Domäne.

5.3.1 Segmentierung von Linien nach Steger

Linienprofile

Die zunächst naheliegende Annahme betreffend das Profil einer Linie der Breite $2w$ und Höhe h ist, dass diese die Form eines Balkens wie in Abbildung 5.4 besitzt, und formal beschrieben werden kann als:

$$f_b(x) = \begin{cases} h, & \text{falls } |x| \leq w \\ 0, & \text{falls } |x| > w \end{cases} \quad (5.1)$$

In der Praxis ist dies jedoch nur selten der Fall, denn infolge von Abtastfehlern des Bildsensors weisen Linien in der Praxis vielmehr ein parabelförmiges Profil (Abbildung 5.4) auf. Unter dieser Annahme kann das Profil im Idealfall folgendermaßen beschrieben werden:

$$f_p(x) = \begin{cases} h(1 - (x/w)^2), & \text{falls } |x| \leq w \\ 0, & \text{falls } |x| > w \end{cases} \quad (5.2)$$

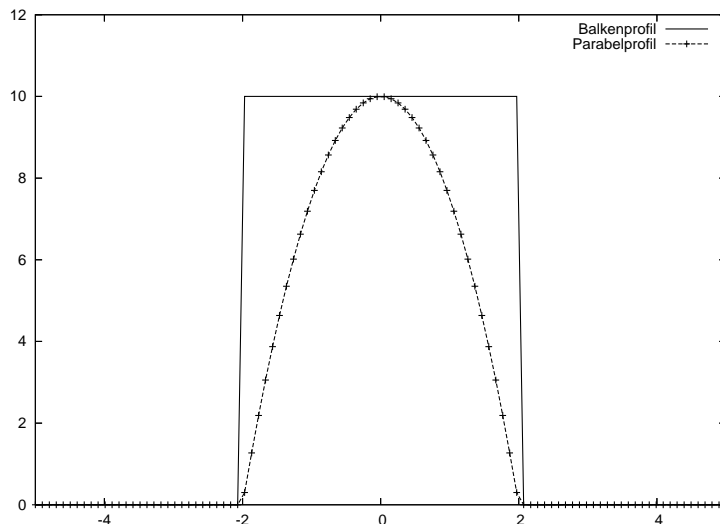
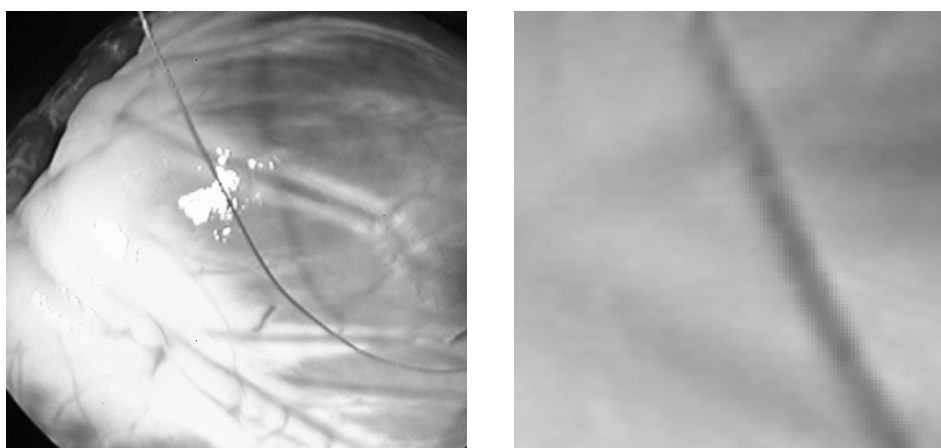


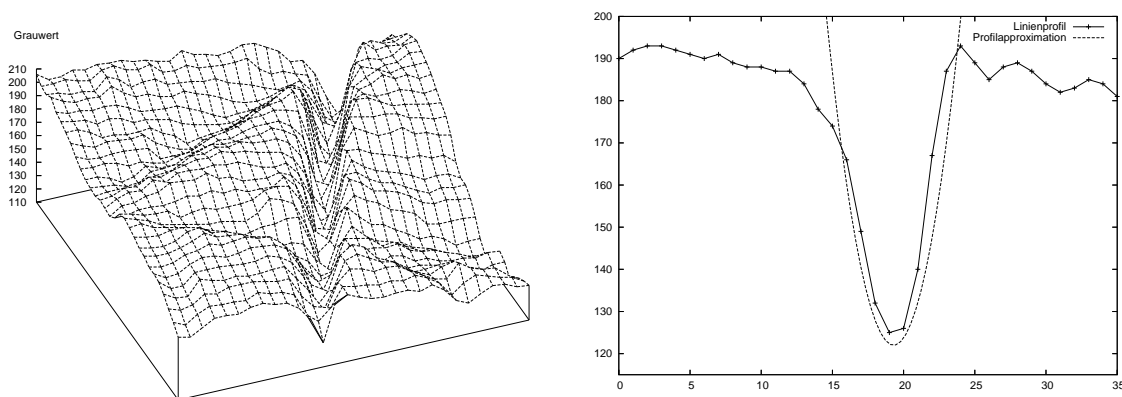
Abbildung 5.4: Idealisiertes balken- und parabelförmiges Profil von Linien

Dieser Umstand trifft in der Praxis insbesondere auf Profile chirurgischer Nahtmaterialien zu. Die zylindrische Form der Fäden bewirkt in der Draufsicht einen

weichen Übergang zwischen Linienrand und Hintergrund. Mit Hilfe dieses a priori Wissens kann ein spezialisierter Filter entworfen werden, der sowohl eine höhere Genauigkeit aufweist, als auch robuster ist im Vergleich mit allgemeinen Verfahren. Abbildung 5.5(a) zeigt die typische Aufnahme eines geflochtenen Polyesterfadens. Im vergrößerten Ausschnitt ist der Unterschied zu den Gefäßen der Herzoberfläche gut zu beobachten. Sowohl das entsprechende dreidimensionale Linienprofil, als auch die einzelne Pixelzeile (Abbildung 5.5(b)) bestätigen die vorherige Annahme eines parabelförmigen Profils.



(a) Faden im Grauwertbild und vergrößerter Ausschnitt



(b) 3D-Linienprofil und einzelne Pixelzeile

Abbildung 5.5: Profil des geflochtenen Polyesterfadens ETHIBOND EXCEL 3-0

Erkennung von Linien in 1D

Unter Idealbedingungen, d.h. ohne Sensorrauschen, können die Punkte einer Linie mit dem Profil aus Gleichung 5.2 einfach mit Hilfe der ersten Ableitung ermittelt werden. Gesucht werden diejenigen Punkte des Bildes $z(x)$, für die der Wert $z'(x)$ sehr klein wird. Eine Möglichkeit, die Suche nur auf besonders ausgeprägte Linien zu

beschränken, bietet die Berücksichtigung der zweiten Ableitung $z''(x)$ an den Stellen mit $z'(x) = 0$. Für helle Linien auf dunklem Hintergrund gilt $z''(x) \ll 0$, während für dunkle Linien auf hellem Hintergrund $z''(x) \gg 0$. Für das ideale Linienprofil aus Gleichung 5.2 gilt dabei:

$$f_p''(x) = -\frac{2h}{w^2}, \quad \forall |x| \leq w. \quad (5.3)$$

Bilder aus der Praxis weisen jedoch ein erhebliches Maß an Rauschen auf, wodurch obiges Modell nicht ausreichend genau ist. In solchen Fällen empfiehlt es sich, die ersten und zweiten Ableitungen mittels Faltung mit einem Gauß-Kern zu ermitteln. Die Kerne und die entsprechenden Rückantworten (Ableitungen) lauten jeweils:

$$g_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad r_p(x, \sigma, w, h) = g_\sigma(x) * f_p(x) \quad (5.4)$$

$$g_\sigma'(x) = \frac{-x}{\sqrt{2\pi}\sigma^3} e^{-\frac{x^2}{2\sigma^2}}, \quad r_p'(x, \sigma, w, h) = g_\sigma'(x) * f_p(x) \quad (5.5)$$

$$g_\sigma''(x) = \frac{x^2 - \sigma^2}{\sqrt{2\pi}\sigma^5} e^{-\frac{x^2}{2\sigma^2}}, \quad r_p''(x, \sigma, w, h) = g_\sigma''(x) * f_p(x) \quad (5.6)$$

Für eine helle Linie auf dunklem Hintergrund mit $w = 1$ und $h = 1$ ergibt sich jeweils für r_p' und r_p'' :

$$r_p'(x, \sigma, w, h) = 0, \quad r_p''(x, \sigma, w, h) = \min \quad \text{für } x = 0, \quad \forall \sigma \quad (5.7)$$

Demzufolge ist es möglich, die Position einer Linie für beliebige Werte von σ genau zu bestimmen. Bei steigenden Werten für σ werden Strukturen als Folge der Glättung abgeflacht, sodass Schwellwerte bei der Segmentierung entsprechend gesetzt werden müssen.

Für die Segmentierung von Linien mit einem Balkenprofil können keine einfachen Kriterien basierend auf Werte von $z'(x)$ und $z''(x)$ angegeben werden, zumal diese im Intervall $[-w, w]$ verschwinden. Faltet man jedoch das Profil aus Gleichung 5.1 mit entsprechenden Gauß-Kernen, so erhält man immer glatte Funktionsverläufe. Die Ecken des Balkenprofils werden abgerundet, wobei die erste Ableitung bei $x = 0$ nur für $\sigma > 0$ verschwindet. Die zweite Ableitung wird jedoch an dieser Stelle nicht ihren kleinsten Wert annehmen, sondern für kleine σ sehr nahe bei Null sein. Sie wird zwei verschiedene Minima im Intervall $[-w, w]$ haben. Erstrebenswert ist jedoch ein Minimum bei $x = 0$, was ausgeprägten Linien entspricht. Steger zeigt, dass dies für Werte von $\sigma \geq w/\sqrt{3}$ zutrifft. Desweiteren hat die zweite Ableitung ihre maximale negative Rückantwort im Skalenraum bei $\sigma = w/\sqrt{3}$. Unter Einhaltung dieser Einschränkung kann der obige Ansatz folglich auch für Linien mit einem Balkenprofil erfolgreich angewendet werden.

Bei den bisherigen Betrachtungen wurde davon ausgegangen, dass der Kontrast auf beiden Seiten einer Linie gleich ist. Sehr oft ist das jedoch in der Praxis nicht der Fall. Wir erweitern das Modell um auch Linien mit einem Profil wie in Abbildung 5.6 beschreiben zu können:

$$f_a(x) = \begin{cases} 0, & \text{falls } x < -w \\ 1, & \text{falls } |x| \leq w \\ a, & \text{falls } x > w \end{cases} \quad (5.8)$$

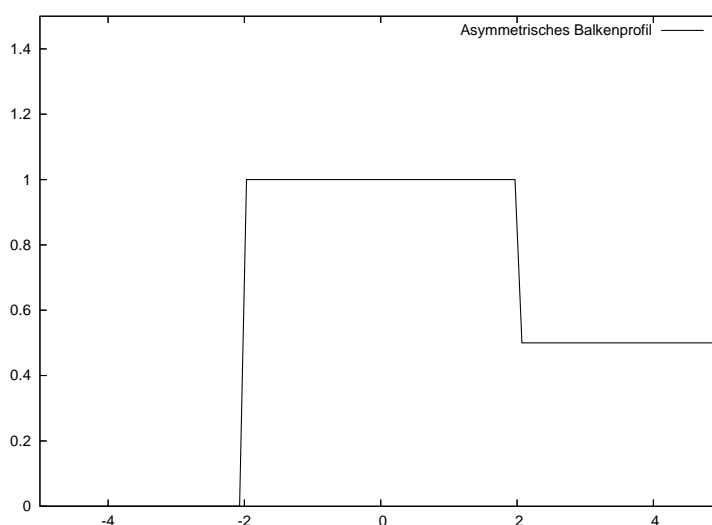


Abbildung 5.6: Asymmetrisches Balkenprofil

Anstelle der Höhe h berücksichtigt man nun den Kontrastunterschied a , und definiert Linien der Höhe h durch Skalierung asymmetrischer Profile mittels $hf_a(x)$. Die Nullstelle der Rückantwort $r'_p(x, \sigma, w, a)$, d.h. die genaue Lage der Linie, ist somit gegeben durch:

$$l = -\frac{\sigma^2}{2w} \ln(1 - a) \quad (5.9)$$

Diskretisierung

Für digitalisierte Bilder, wie sie üblicherweise für die Bildverarbeitung zur Verfügung stehen, sind reell-analytische Funktionen $z(x)$, wie wir sie bisher verwendet haben, nicht mehr geeignet. Das diskrete Bild z_n wird als eine stückweise konstante Funktion $z(x) = z_n$ betrachtet, für $x \in (n - \frac{1}{2}, n + \frac{1}{2}]$. Die genaue Position der Linie im diskreten Fall wird mit Hilfe des Taylorpolynoms zweiten Grades von z_n ermittelt. Wenn r , r' und r'' die im Punkt n ermittelten lokalen Ableitungen sind, so wird das Taylorpolynom $p(x)$ gegeben durch:

$$p(x) = r + r'x + \frac{1}{2}r''x^2 \quad (5.10)$$

Die genaue Lage eines Punktes auf der Linie, d.h. der Punkt wo $p'(x) = 0$ gilt, ist somit:

$$x = -\frac{r'}{r''} \quad (5.11)$$

Der Punkt n wird nur dann als ein Punkt der Linie betrachtet, wenn $x \in [-\frac{1}{2}, \frac{1}{2}]$ und die zweite Ableitung r'' einen bestimmten Schwellert überschreitet.

Übergang 1D→2D

Die Modellierung kurvilinearere Strukturen in 2D kann mit Hilfe von Kurven $s(t)$ erfolgen, die ein charakteristisches Linienprofil, beispielsweise f_a , senkrecht zu $s'(t)$ aufweisen. Wenn wir diese senkrechte Richtung mit $n(t)$ bezeichnen, müssen Linienpunkte folgende Bedingungen erfüllen: (i) die erste Richtungsableitung sollte verschwindend klein sein entlang von $n(t)$ und (ii) die zweite Richtungsableitung sollte einen hohen Absolutwert haben.

Die Berechnung der Linienrichtung für die einzelnen Punkte erfolgt mit Hilfe der partiellen Ableitungen r_x, r_y, r_{xx}, r_{xy} und r_{yy} . Diese werden wieder mittels Faltungsmasken ermittelt. Die Richtung, in der die zweite Richtungsableitung von $z(x, y)$ einen maximalen Absolutwert erreicht, wird als $n(t)$ benutzt. Diese kann aus den Eigenwerten und Eigenvektoren der Hessematrix berechnet werden:

$$H(x, y) = \begin{pmatrix} r_{xx} & r_{xy} \\ r_{xy} & r_{yy} \end{pmatrix} \quad (5.12)$$

Der zum absolut größten Eigenwert gehörende Eigenvektor (n_x, n_y) , mit $\|(n_x, n_y)\|_2 = 1$ gibt nun die Richtung einer Senkrechten zur gesuchten Linie an. Wie bereits im 1D-Fall, wird ein quadratisches Polynom verwendet, um zu ermitteln, ob in dem Punkt die erste Richtungsableitung entlang von (n_x, n_y) verschwindet. Dieser Punkt ist gegeben durch:

$$(p_x, p_y) = (tn_x, tn_y), \quad (5.13)$$

wobei

$$t = \frac{r_x n_x + r_y n_y}{r_{xx} n_x^2 + 2r_{xy} n_x n_y + r_{yy} n_y^2}. \quad (5.14)$$

Wieder muss $(p_x, p_y) \in [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$ gelten, damit ein Punkt als Linienpunkt betrachtet werden kann. Die zweite Richtungsableitung entlang von (n_x, n_y) , der größte Eigenwert, kann als Schwellwert zur Identifikation ausgeprägter Linien verwendet werden.

Linienbreiten

Linien mit einem Balkenprofil werden auf beiden Seiten von jeweils einer Kante vom Hintergrund getrennt, welche die Breite einer Linie markieren. Wie wir bereits gesehen haben, wird sich die genaue Position l der Linie (üblicherweise Mitte des Balkenprofils) an dem Nulldurchgang der Rückantwort $r'_p(x, \sigma, w, a)$ befinden. Die Linienbreite, gemessen von linken bis zum rechten Rand, kann man anhand der Nulldurchgänge von $r''_p(x, \sigma, w, a)$ bestimmen.

Die richtige Wahl der Parameter und Schwellwerte ist für eine optimale Funktionsweise des Filters sehr wichtig. A priori Wissen über Beschaffenheit der Kanten, Linienprofil und Liniendicke erlauben in unserem Fall eine ideale Parametrierung des Filters. Anhand der zu erwartenden Liniendicke w in Pixelkoordinaten und deren Kontrast h abhängig vom Hintergrund kann der Parameter σ aus dem Zusammenhang $\sigma \geq w/\sqrt{3}$ bestimmt werden. Im folgenden Abschnitt wird die vorgestellte Methode für die Segmentierung der chirurgischen Fäden aus Abbildung 5.3 verwendet, und ihre Eignung für verschiedene Fadenstärken gezeigt.

5.3.2 Anwendungsbeispiele für chirurgisches Nahtmaterial

Experimentalaufbau

In Abschnitt 5.2 haben wir bereits chirurgische Fäden vorgestellt, welche vorzugsweise in der minimalinvasiven Herzchirurgie Anwendung finden. In diesem Abschnitt werden wir die Aufnahmen aus Abbildung 5.3 segmentieren und die Ergebnisse diskutieren. Dafür wurden monofile und geflochtene Fäden verschiedener Stärken aus Polypropylen und Polyester mit einem Stereoendoskop der Firma R. Wolf (Abbildung 2.21) aus typischen Entfernungen aufgenommen. Auf einen realistischen geschlossenen minimalinvasiven Aufbau wurde verzichtet, das Ziel war in erster Linie die Ermittlung der Genauigkeit des Verfahrens.

Die Fadenerkennung dient primär der anschließenden 3D-Rekonstruktion des Fadenverlaufs. Wegen der schwierigen Erstellung von *ground truth* Daten für chirurgische Fäden, wurde eine andere Methode zur Schätzung der Rekonstruktionsgenauigkeit verwendet. Über die Ermittlung der Fadenstärke mittels Bildverarbeitung kann indirekt auf die Genauigkeit der 3D-Rekonstruktion geschlossen werden. Um metrische Größenangaben über die Fadenstärke machen zu können wird der im Kapitel 3 bereits vorgestellte Kalibrierkörper eingesetzt. Abbildung 5.7 zeigt den verwendeten

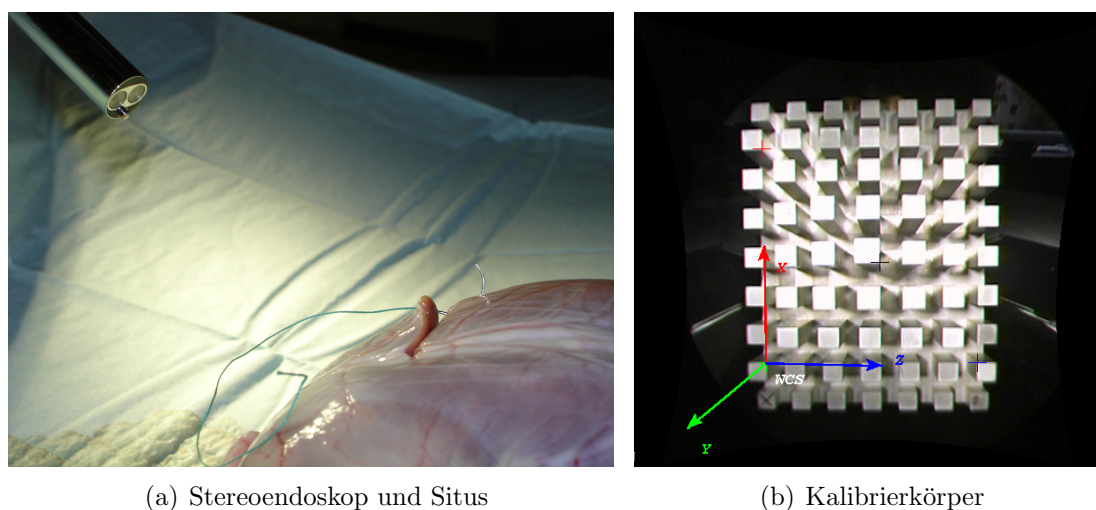


Abbildung 5.7: Aufbau zur Auswertung der Ergebnisse der Segmentierung

Aufbau. Mit Hilfe des Kalibrierkörpers wurde ein kartesisches Weltkoordinatensystem definiert, das sich ungefähr in gleicher Entfernung wie der Situs befindet. Es wurde großen Wert darauf gelegt, die Messungen für die einzelnen Fadenstärken unter möglichst gleichen Bedingungen durchzuführen. Eine Messreihe bestand dabei aus folgenden Teilschritten:

1. Rektifizierung des Stereobildpaares anhand der Kameraparameter
2. Segmentierung des Fadenverlaufs im linken rektifizierten Bild
3. Ermittlung der Fadenstärke in Bildkoordinaten
4. 3D-Rekonstruktion im Koordinatensystem der linken rektifizierten Kamera
5. Transformation ins Koordinatensystem des Kalibrierkörpers
6. Ermittlung der Fadenstärke in 3D-Koordinaten mittels zwei Methoden

Beleuchtungs- und Reflexionskanten

Eine Hauptschwierigkeit der Kantendetektion, insbesondere bei dreidimensionalen Objekten, ist die Unterscheidung von Reflektionskanten, die auf Eigenschaften des Objektes beruhen, und Beleuchtungskanten, die auf Eigenschaften der Beleuchtung beruhen (z.B. Schatten, Lichtkegel). Abbildung 5.8 zeigt das Ergebnis der Segmentierung zweier Abschnitte desselben Fadens, wobei die Visualisierung subpixelgenauer Positionen mittels zehnfacher Vergrößerung und folgender Umrechnungsformel erfolgte:

$$x_i = scale \cdot (\text{int})\text{round}(x_f) + (\text{int})scale \cdot (x_f - \text{round}(x_f)) \quad (5.15)$$

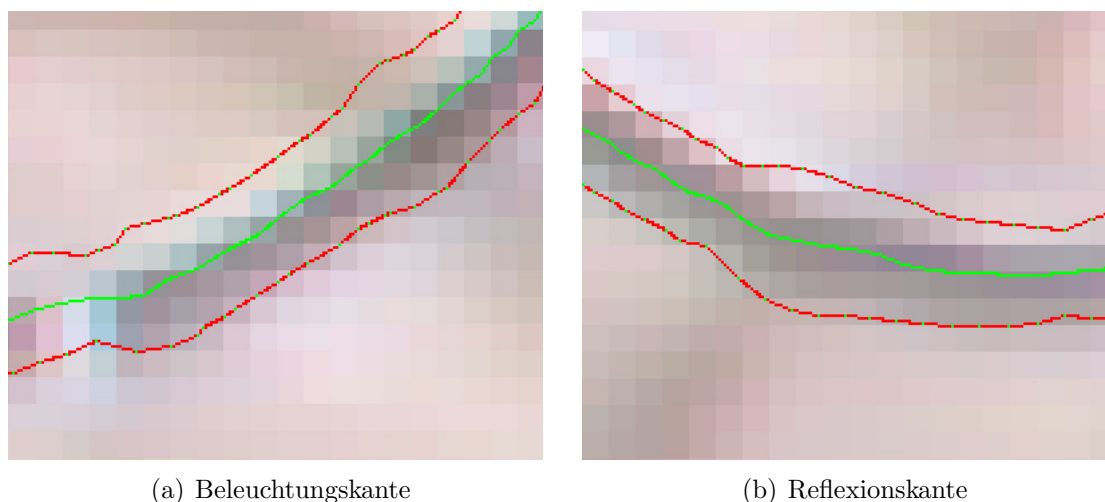


Abbildung 5.8: Beleuchtungs- und Reflexionskanten

Das Maß der Vergrößerung (*scale*) definiert, mit welcher Auflösung die subpixelgenauen Linienpunkte x_f auf ganzzahlige Bildpunkte x_i abgebildet werden. Mit der oben angegebenen Skalierung ist eine zehntelpixelgenaue Darstellung möglich. Im rechten Bild wurden sowohl Verlauf als auch die Ränder des Fadens korrekt segmentiert. Im linken Bild hingegen erzeugt die ungünstig positionierte Lichtquelle eine leichte Verschiebung der Kanten. Dadurch dass Ränder und Kern in gleichem Maße von der Verschiebung betroffen sind, wirken sich Beleuchtungskanten nicht negativ auf die Messung des Fadendurchmessers aus.

Ein zusätzlicher Hinweis auf die veränderten Lichtbedingungen in Abbildung 5.8(a) ist die Spiegelreflexion in der linken unteren Ecke. Spiegelreflexionen verfälschen im Unterschied zu Beleuchtungskanten die Messung der Linienbreite, indem sie zu einer verengten Randkontour führen. Im ungünstigsten Fall verursachen sie sogar eine Unterbrechung im Linienverlauf. Eine Entfernung der Spiegelreflexionen, wie es im Abschnitt 4.1 vorgestellt wurde, führt hingegen zu einer Verdickung an der entsprechenden Stelle. Wir verzichten deshalb bei den Messungen auf die Glanzlichtentfernung komplett, und berechnen die Linienbreiten einzelner Fadentypen als die Mittelung aller gemessenen Werte entlang eines Fadenverlaufs. Dadurch werden auch entfernungsabhängige Schwankungen ausgeglichen, die infolge der variierenden Tiefenauflösung des Endoskops entstehen.

Segmentierungsparameter und Ergebnisse

In der Tabelle aus Abbildung 5.9 wurden die verwendeten Parameter für die Segmentierung der Bilder aus Abbildung 5.3 zusammengefasst. Statt im Originalbild,

wurde die Segmentierung jeweils im linken rektifizierten Bild durchgeführt, um gültige Pixelkoordinaten für den späteren Zugriff auf die Disparitätskarte zu erhalten. Die Bildentzerrung hat in diesem Fall keine negative Auswirkung auf die Segmentierung. Für Linsen mit sehr hoher Radialverzerrung kann infolge von Interpolationsartefakten im entzerrten Bild eventuell eine Segmentierung im Originalbild notwendig sein. Segmentierte Punkte können anschließend anhand der Kameraparameter in die Pixelkoordinaten des rektifizierten Bildes transformiert werden.

Die Werte von σ müssen, wie erwartet, für abnehmende Fadendurchmesser entsprechend verringert werden. Untere und obere Schwellwerte sind abhängig vom Situs, insbesondere von der Texturierung und Beschaffenheit der Herzoberfläche.

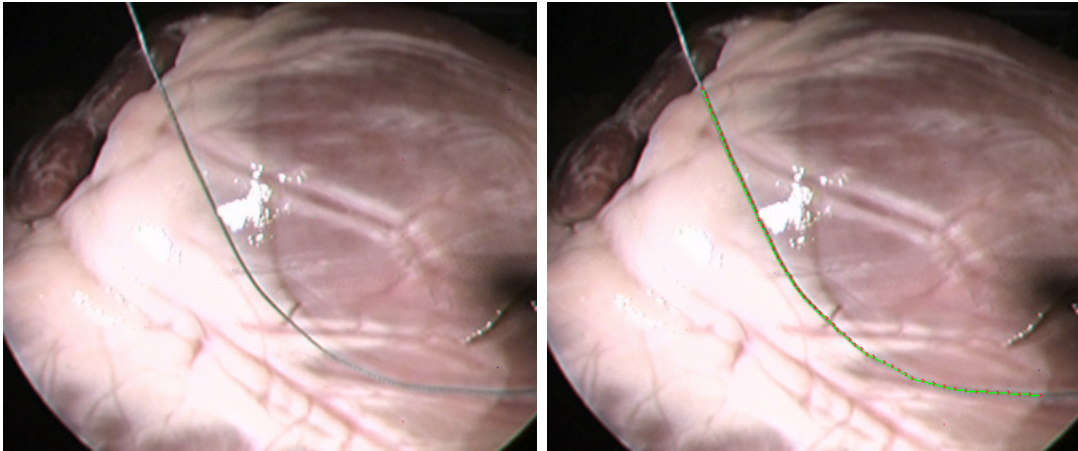
Material	U.S.P.	\emptyset	sigma	low	high
Polyester (geflochten)	3-0	0.2 mm	2.5	0.5	5.0
Polyester (geflochten)	5-0	0.1 mm	2.1	0.0	4.0
Polypropylen (monofil)	3-0	0.2 mm	2.5	1.5	5.0
Polypropylen (monofil)	4-0	0.15 mm	2.15	0.5	4.0
Polypropylen (monofil)	5-0	0.1 mm	1.9	2.0	4.0
Polypropylen (monofil)	6-0	0.07 mm	1.35	3.0	4.5
Polypropylen (monofil)	7-0	0.05 mm	0.775	3.0	4.5

Abbildung 5.9: Verwendete Parameter für die Segmentierung

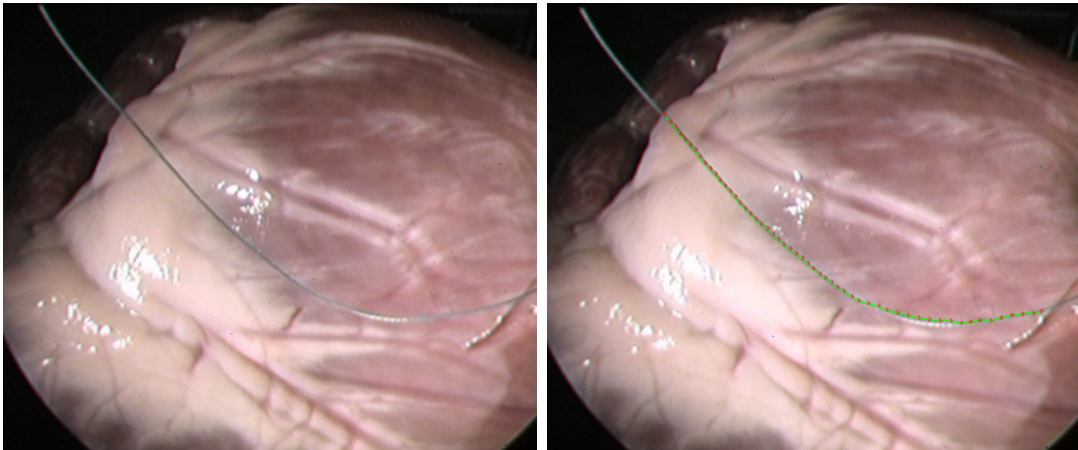
Die Abbildungen 5.10, 5.11 und 5.12 zeigen die Ergebnisse der Segmentierung. Die Fadenverläufe wurden größtenteils korrekt erkannt, lediglich die zwei dünnsten Polypropylenfäden (6-0, 7-0) weisen größere Lücken auf. Bei Fadendurchmessern von unter einem Zehntelmillimeter kommt das optische System bei diesen Entfernungen an die Grenzen ihrer Auflösungsfähigkeit. Nichtsdestotrotz werden auch hier genügend Punkte des Fadenverlaufs segmentiert, die für eine 3D-Rekonstruktion herangezogen werden können.

Messung des Fadendurchmessers

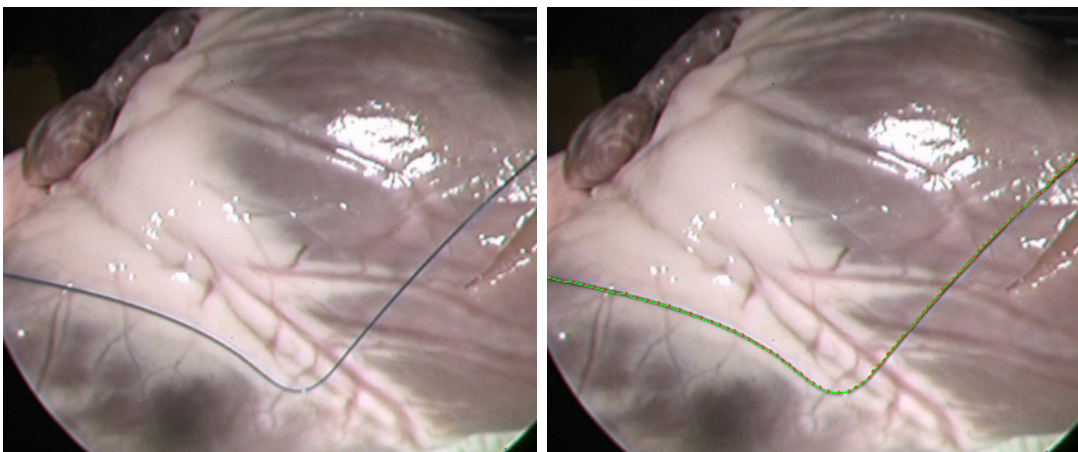
Für die 3D-Rekonstruktion der Fäden in einem Operationssitus ist die explizite Segmentierung ihrer Ränder, und somit ihres Durchmessers, nicht erforderlich. Für das Greifen mit Hilfe chirurgischer Instrumente bietet eine zusätzliche Dimensionsangabe im 3D-Raum keine Vorteile, zumal die Backen der Greifer nicht über die erforderliche Genauigkeit verfügen, um diese Information nutzen zu können. Nichtsdestotrotz verwenden wir die ermittelten Linienbreiten als ein Maß für die Segmentierungs- und Rekonstruktionsgenauigkeit im allgemeinen, denn die Fadenstärken sind bei chirurgischen Eingriffen a priori bekannt. Eine geringe Abweichung von den Herstellerangaben deutet sowohl auf eine genaue Segmentierung der



(a) Polyester 3-0

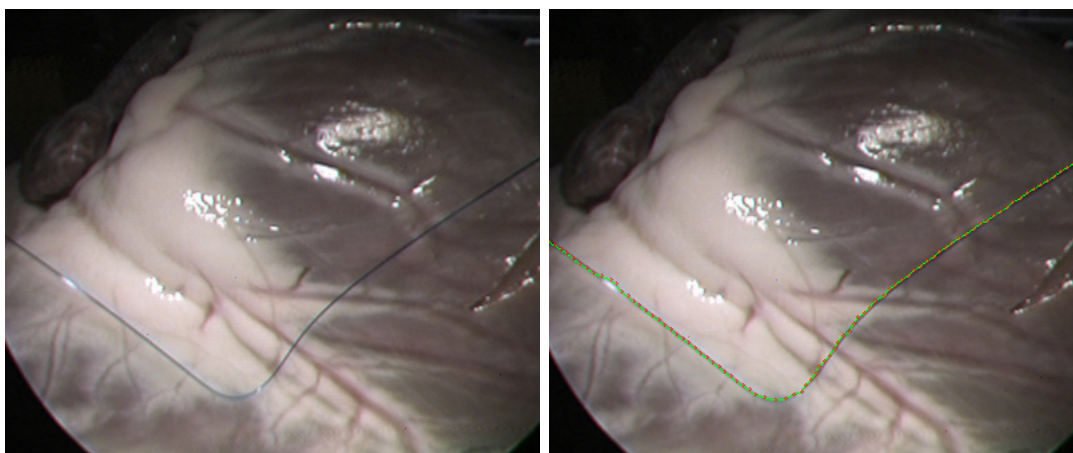


(b) Polyester 5-0

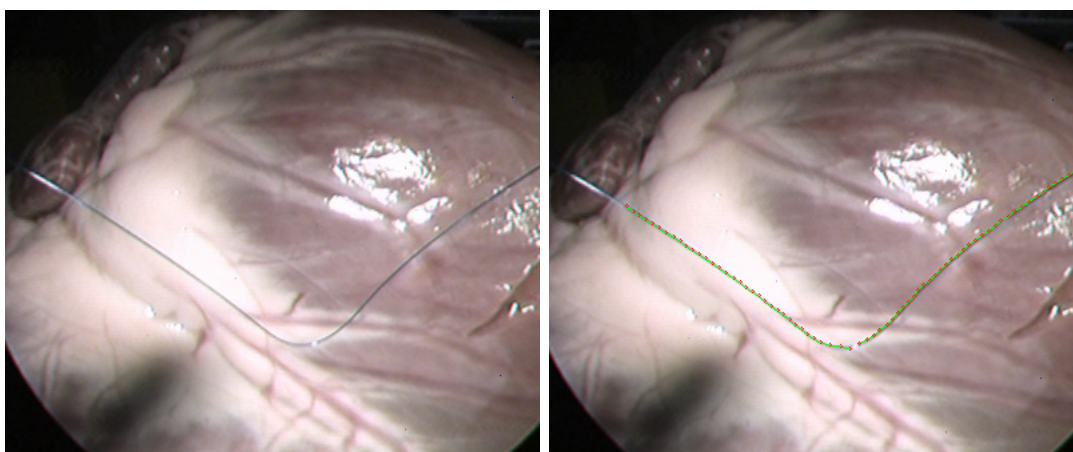


(c) Polypropylen 3-0

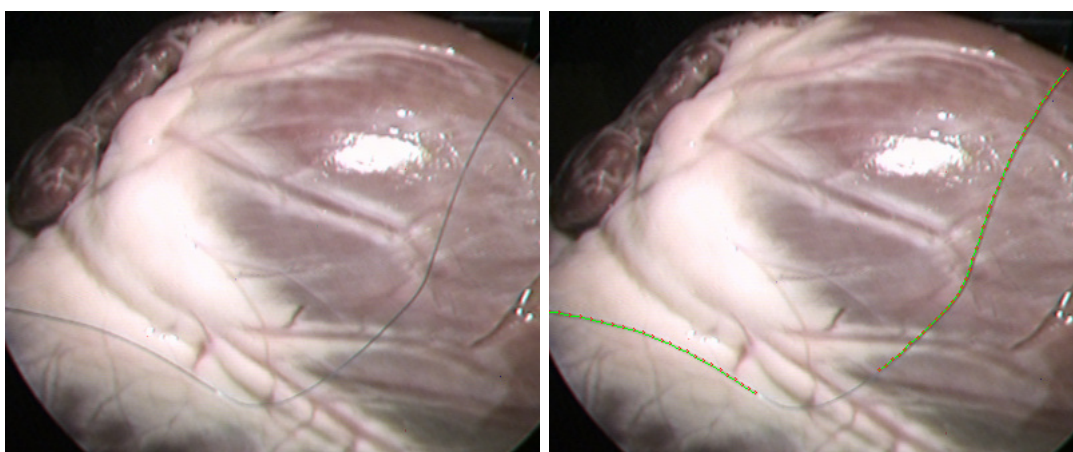
Abbildung 5.10: Jeweils das linke rektifizierte Bild und der segmentierte Fadenverlauf für die Polyesterfäden 3-0 und 5-0 und den Polypropylenfaden 3-0.



(a) Polypropylen 4-0

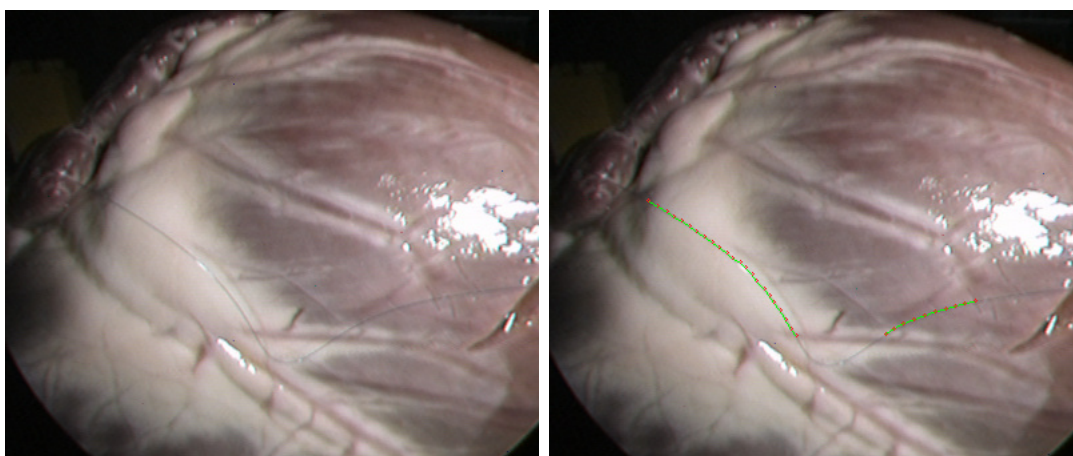


(b) Polypropylen 5-0



(c) Polypropylen 6-0

Abbildung 5.11: Jeweils das linke rektifizierte Bild und der segmentierte Fadenverlauf für die Polypropylenfäden 4-0, 5-0 und 6-0.



(a) Polypropylen 7-0

Abbildung 5.12: Jeweils das linke rektifizierte Bild und der segmentierte Fadenverlauf für den Polypropylenfaden 7-0.

Ränder in Pixelkoordinaten hin, als auch auf eine akkurate Rekonstruktion im 3D-Raum. Die Tabelle aus Abbildung 5.13 fasst die berechneten Werte für die Bilder aus den Abbildungen 5.10, 5.11 und 5.12 zusammen. Es wurden dabei für jeden Linienpunkt eines Verlaufs jeweils der linke und rechte Randpunkt berechnet, der euklidische Abstand zwischen ihnen entspricht somit der gesuchten Breite.

Material / U.S.P.	Mtl. \emptyset	Min. \emptyset	Max. \emptyset	Mtl. Kontrast
Polyester (geflochten) / 3-0	4.511	2.025	7.970	-74.897
Polyester (geflochten) / 5-0	4.849	0.057	10.084	-41.833
Polypropylen (monofil) / 3-0	4.238	2.789	10.069	-64.814
Polypropylen (monofil) / 4-0	3.821	2.545	5.835	-41.652
-//-	3.975	1.433	7.088	-35.343
Polypropylen (monofil) / 5-0	3.557	2.253	4.970	-44.453
-//-	3.834	0.645	7.318	-44.842
Polypropylen (monofil) / 6-0	3.516	1.733	6.625	-35.731
-//-	3.506	1.707	4.084	-34.460
Polypropylen (monofil) / 7-0	2.238	0.377	3.527	-22.862
-//-	2.908	1.801	3.869	-24.280

Abbildung 5.13: Fadendurchmesser in Bildkoordinaten (subpixelgenau) und gemittelter Kontrast entlang des Verlaufs

Wie bereits eingangs erwähnt, weist der ermittelte Durchmesser aufgrund von Sensorrauschen und unterschiedlichen Lichtverhältnissen relativ große Schwankungen auf. Die Tabelle listet deshalb Minimal-, Maximal- und Durchschnittswerte für die errechneten Durchmesser, als auch den durchschnittlichen Kontrast entlang des ge-

samten Linienverlaufs. Ein hoher Kontrastwert kann in manchen Fällen als zusätzliches Ausschlusskriterium für Falschzuordnungen (wie Gefäße oder Oberflächenebenen) dienen. Die durchschnittlichen Linienbreiten in Bildkoordinaten unterschreiten für keinen der Fäden das Minimum von zwei Pixeln, was üblicherweise in der Bildverarbeitung als notwendige Mindestgröße für eine zu segmentierende Struktur gesehen wird.

Die ermittelten Pixelkoordinaten können nun unter Zuhilfenahme der Kameraparameter in beliebige Koordinatensysteme umgerechnet werden. Es wurden zwei unterschiedliche Methoden zur Bestimmung metrischer Dimensionsangaben über den Fadendurchmesser verwendet. Eine erste naheliegende Methode ist die 3D-Rekonstruktion aller Randpunkte l und r und ihre Transformation in das Weltkoordinatensystem des Kalibrierkörpers aus Abbildung 5.7(b). Anschließend kann der euklidische Abstand zwischen den zwei Punkten ${}^{WCS}L$ und ${}^{WCS}R$ berechnet werden:

$${}^{WCS}d = \sqrt{({}^{WCS}x_L - {}^{WCS}x_R)^2 + ({}^{WCS}y_L - {}^{WCS}y_R)^2 + ({}^{WCS}z_L - {}^{WCS}z_R)^2} \quad (5.16)$$

Der Nachteil dieser Methode ist, dass für die Randpunkte unter Umständen keine genauen Disparitätswerte zur Verfügung stehen. Für einen subpixelgenauen Randpunkt muss nämlich aus der diskreten Disparitätskarte ein Wert an einer gerundeten ganzzahligen Bildkoordinate verwendet werden. Ein subpixelgenauer Disparitätswert kann auch nicht anhand von Nachbarpixel mittels Interpolation berechnet werden, da diese in unserem Fall immer zum Hintergrund gehören und das Ergebnis massiv verfälschen würden. Um trotzdem die höchstmögliche Genauigkeit zu erreichen, wurden für die Berechnung nur diejenigen Punkte der Ränder herangezogen, die bei der Stereokorrelation mittels der ZNCC-Methode (siehe Gl. 4.15) einen Score von mindesten 0.9 aufweisen.

Eine zweite Methode ergibt sich aus der Tatsache, dass infolge des zylindrischen Profils des Fadens der Durchmesser aus jeder Blickrichtung theoretisch gleich ist. Dadurch kann das Problem der 3D-Rekonstruktion auf ein einfacheres 2D-Projektionsproblem reduziert werden. Abbildung 5.14 zeigt schematisch das Lochkameramodell, das auch der Kalibrierung für die Ermittlung der Kameraparameter zugrunde liegt.

Dabei erlaubt der Zusammenhang

$$\text{Brennweite} = \frac{\text{Gebrauchsdistanz} \times \text{Sensorgröße}}{\text{Sichtfeld}} \quad (5.17)$$

die Berechnung von Größen von Strukturen, die sich in einer Ebene parallel zur Linse befinden. Der Vorteil dieser Methode ist, dass sie subpixelgenaue Bildkoordinaten verwendet. Die Abmessung einer Struktur auf der Sensorfläche des CCD-Chips ist

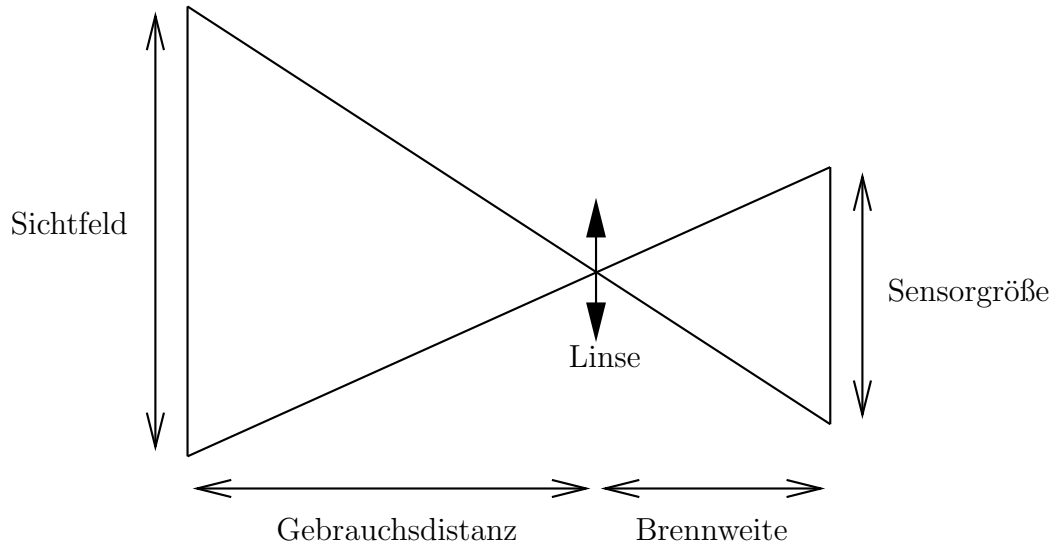


Abbildung 5.14: Schematische Darstellung des Lochkammermodells

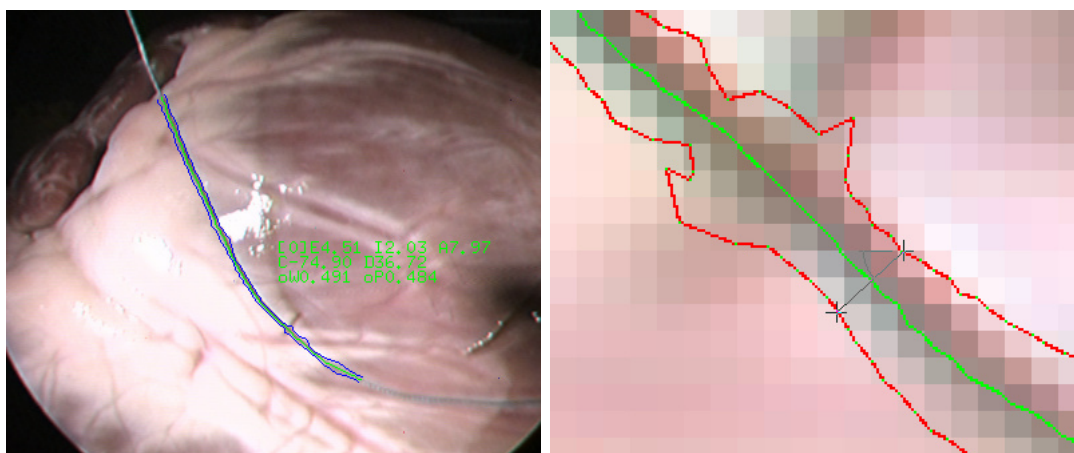
somit das Produkt zwischen der Größe einer Sensorzelle und die (nicht notwendigerweise ganzzahlige) Anzahl der belegten Pixel. Damit ergibt sich als metrische Größe für den Fadendurchmesser

$${}_{CCS}d = \frac{\sqrt{[S_x(Ix_L - Ix_R)]^2 + [S_y(Iy_L - Iy_R)]^2} \cdot dist}{f}, \quad (5.18)$$

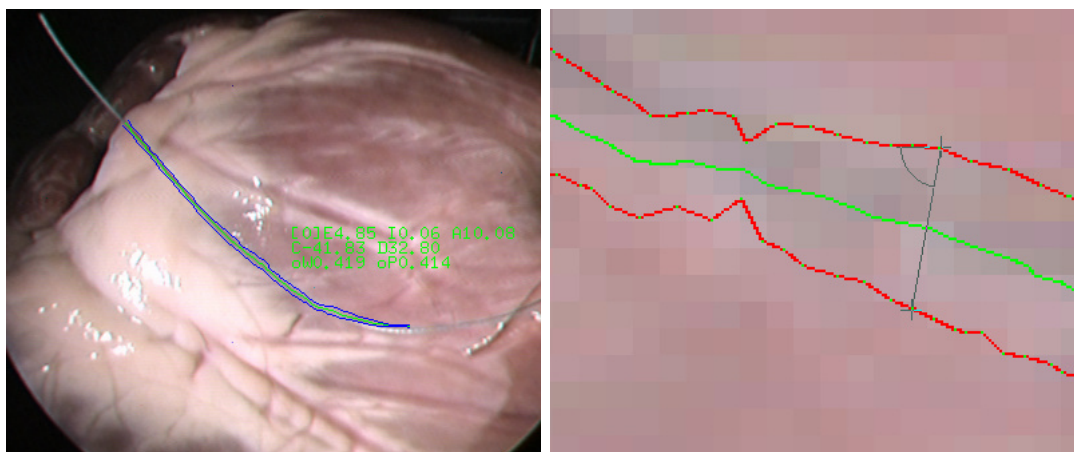
wobei $Ix_{L,R}$ und $Iy_{L,R}$ die subpixelgenauen Bildkoordinaten der Randpunkte, S_x und S_y jeweils die Breite und Höhe einer Sensorzelle sind. Desweiteren ist f die Brennweite und $dist$ die Entfernung zur Linse. Für ein Kamerasystem mit quadratischen Pixeln der Größe S , was bei uns der Fall ist, vereinfacht sich die Gleichung zu:

$${}_{CCS}d = \frac{S \cdot \sqrt{(Ix_L - Ix_R)^2 + (Iy_L - Iy_R)^2} \cdot dist}{f} \quad (5.19)$$

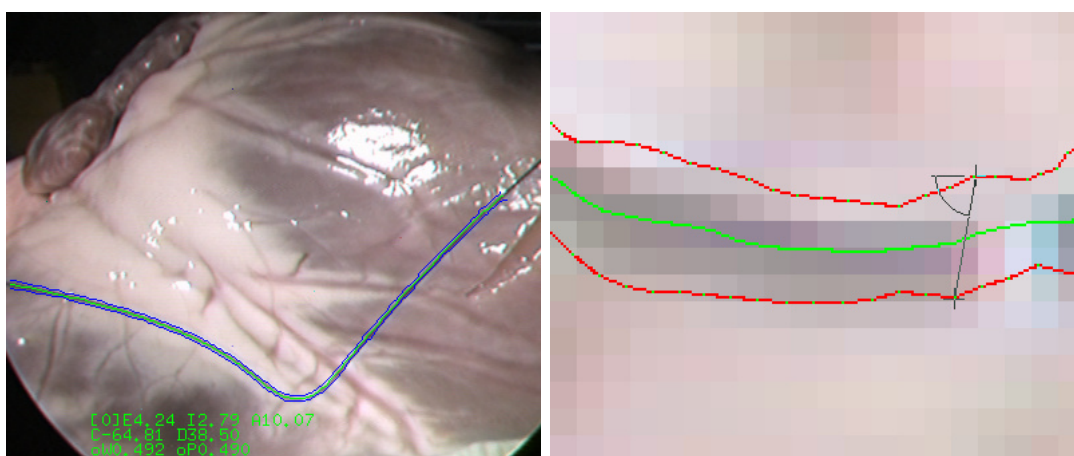
Für die Berechnung der Entfernung $dist$ werden nicht die Randpunkte, sondern der zugeordnete Linienpunkt verwendet, da für ihn aus bereits erwähnten Gründen eine genauere Tiefenbestimmung möglich ist. Das Koordinatensystem CCS ist das der linken rektifizierten Kamera, und WCS das des Kalibrierkörpers aus Abbildung 5.7(b). Die nachfolgenden Abbildungen (5.15, 5.16 und 5.17) zeigen jeweils die Ergebnisse der Segmentierung und Ermittlung der Linienbreiten für die verschiedenen Fadenstärken. Die Bilder enthalten überdies die berechneten Werte, wobei E , I und A jeweils den mittleren, minimalen und maximalen Durchmesser in Pixelkoordinaten bezeichnen.



(a) Polyester 3-0

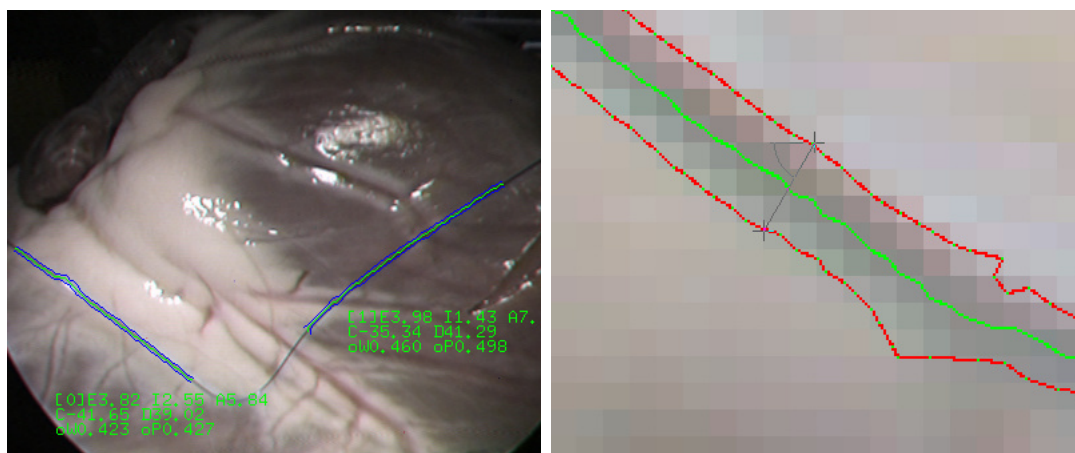


(b) Polyester 5-0

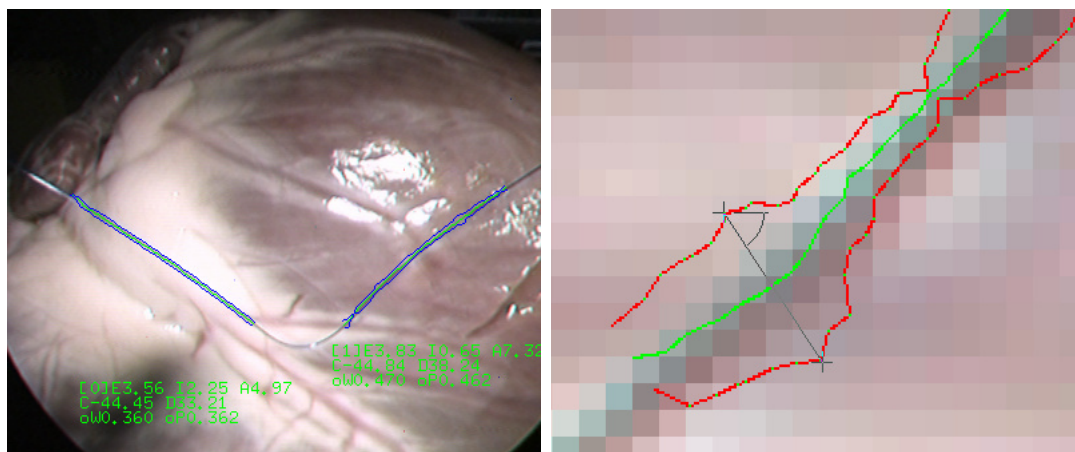


(c) Polypropylen 3-0

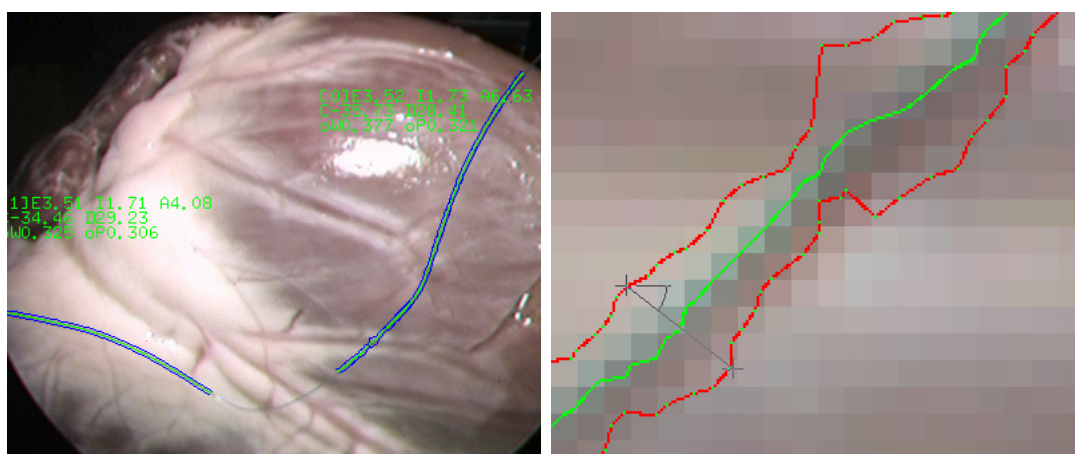
Abbildung 5.15: Segmentierte Fadenverläufe mit eingezeichneten Linienparametern



(a) Polypropylen 4-0

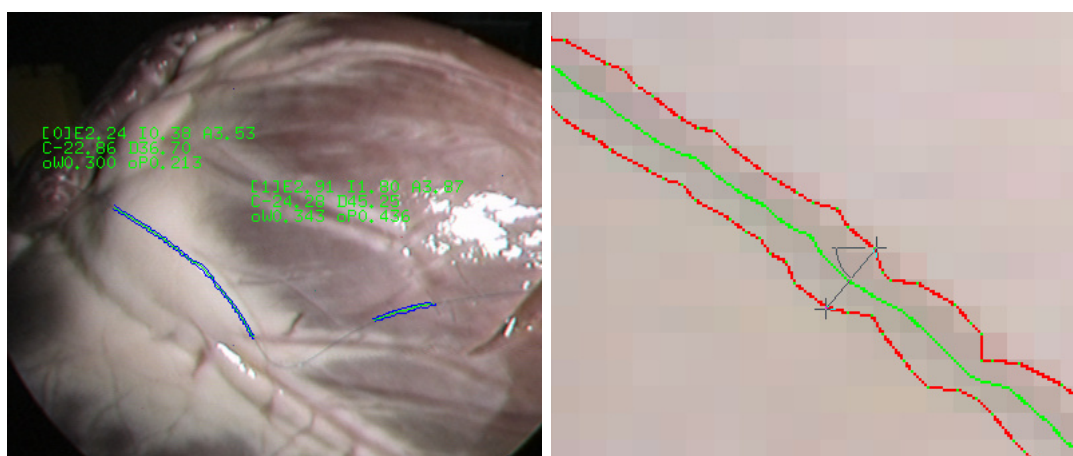


(b) Polypropylen 5-0



(c) Polypropylen 6-0

Abbildung 5.16: Segmentierte Fadenverläufe mit eingezeichneten Linienparametern



(a) Polypropylen 7-0

Abbildung 5.17: Segmentierte Fadenverläufe mit eingezeichneten Linienparametern

Der durchschnittliche Kontrast und die mittlere Entfernung zum Endoskop werden mit jeweils C und D beschriftet. Für den im Koordinatensystem des Kalibrierkörpers gemessene Durchmesser steht $\circ W$, für den im Koordinatensystem der linken rektifizierten Kamera $\circ P$. Weil diese Aufnahmen zur Visualisierung der segmentierten Liniensbreiten nur auf einen Pixel genau sind, wurden ihnen auch vergrößerte Ausschnitte beigefügt, die vorzugsweise problematische Bereiche und Ausreißer darstellen.

Die Hinzunahme der Bedingung eines hohen Scores bei der Stereokorrelation hat zur Folge, dass sich nun im Vergleich zu den segmentierten Verläufen aus den Abbildungen 5.10, 5.11 und 5.12 die Anzahl gültiger Linienpunkte verringern, sowohl am Rand als auch im Inneren des Fadenverlaufs. Desweiteren sind an den nahezu horizontal verlaufenden Abschnitten dünner Fäden Lücken zu beobachten. Dies ist darauf zurückzuführen, dass Punkte horizontal verlaufender Strukturen trotz Epipolareinschränkung im linken und rechten Bild nur schlecht einander zugeordnet werden können.

Material / U.S.P.	\emptyset (WCS)	\emptyset (CCS)	Herst. \emptyset	ERR (mm)
Polyester (geflochten) / 3-0	0.491	0.483	0.2 mm	0.291/0.283
Polyester (geflochten) / 5-0	0.419	0.413	0.1 mm	0.319/0.313
Polypropylen (monofil) / 3-0	0.491	0.490	0.2 mm	0.291/0.290
Polypropylen (monofil) / 4-0	0.441	0.461	0.15 mm	0.291/0.311
Polypropylen (monofil) / 5-0	0.414	0.412	0.1 mm	0.314/0.312
Polypropylen (monofil) / 6-0	0.351	0.313	0.07 mm	0.281/0.243
Polypropylen (monofil) / 7-0	0.320	0.324	0.05 mm	0.270/0.274

Abbildung 5.18: Fadendurchmesser in Welt- und Kamerakoordinaten

Die Tabelle aus Abbildung 5.18 fasst nochmal die Ergebnisse für die verschiedenen

Fadenstärken zusammen. Sie enthält neben den berechneten Linienbreiten in den zwei Koordinatensystemen auch die jeweiligen Abweichungen von den Herstellerangaben. Es ist für alle Fadenstärken ein nahezu konstanter Messfehler von etwa 0.3 Millimetern zu beobachten. Abbildung 5.19 verdeutlicht nochmal diesen Umstand. Links wurden die gemessenen Fadenstärken zusammen mit den Herstellerangaben aufgezeichnet, rechts hingegen jeweils die Abweichungen von der Spezifikation. Die Messungenauigkeit scheint nicht von der Fadenstärke abhängig zu sein, sie ist höchstwahrscheinlich auf den Tiefenfehler des Stereoendoskops zurückzuführen. Auch bei sehr sorgfältiger Kalibrierung können nämlich die Kameraparameter nur für einen beschränkten Entfernungsbereich mit ausreichender Präzision ermittelt werden. Außerhalb dieses Bereiches leidet die Rekonstruktionsgenauigkeit mit zunehmender Entfernung.

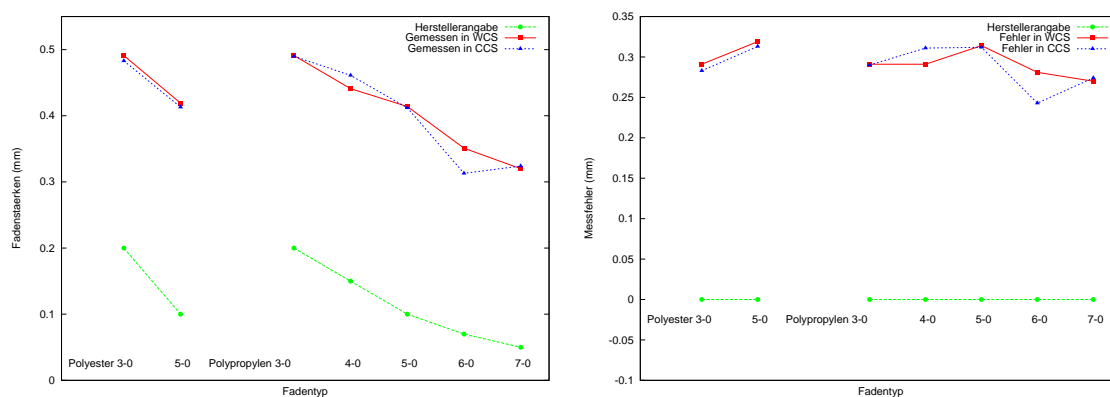


Abbildung 5.19: Gemessene Fadenstärken und Messfehler

Nichtsdestotrotz deutet der beobachtete Messfehler, der für alle Fadenstärken und beide Methoden zur Ermittlung der Linienbreite nahezu konstant ist, auf die Robustheit der Segmentierung hin. Dadurch dass die Messfehler die zu ermittelnden Größen bei weitem überschreiten, eignet sich jedoch das verwendete Endoskop nicht für Messungen im Submillimeterbereich, zumindest nicht aus Entfernungen, wo der Operationsbereich noch zu überblicken ist. Das nächste Kapitel widmet sich ausführlich der 3D-Rekonstruktion, und untersucht Möglichkeiten zur Kompensierung von Tiefenfehlern.

Disparitätskarten und 3D-Rekonstruktionen

In Abbildung 5.20 sind die Disparitätskarten dargestellt, die zur 3D-Rekonstruktion der Randpunkte dienen. Eingezeichnet wurden außer den segmentierten Linienverläufen (grün) auch diejenigen Punkte (rot), wo eine Messung der Linienbreite stattfand. Voraussetzung dafür waren ZNCC-Korrelationscores von über 0.9, sowohl für Rand- als auch Kernpunkte der jeweiligen Linie. Die nebenstehenden 3D-Verläufe stellen den rekonstruierten Faden dar, und zeigen dass trotz lückenhafter

Verfügbarkeit von 3D-Punkten eine glatte Kurve berechnet werden kann. Insbesondere bei sehr dünnem Nahtmaterial kommt es häufig vor, dass eine Unterbrechung im segmentierten Verlauf entsteht.

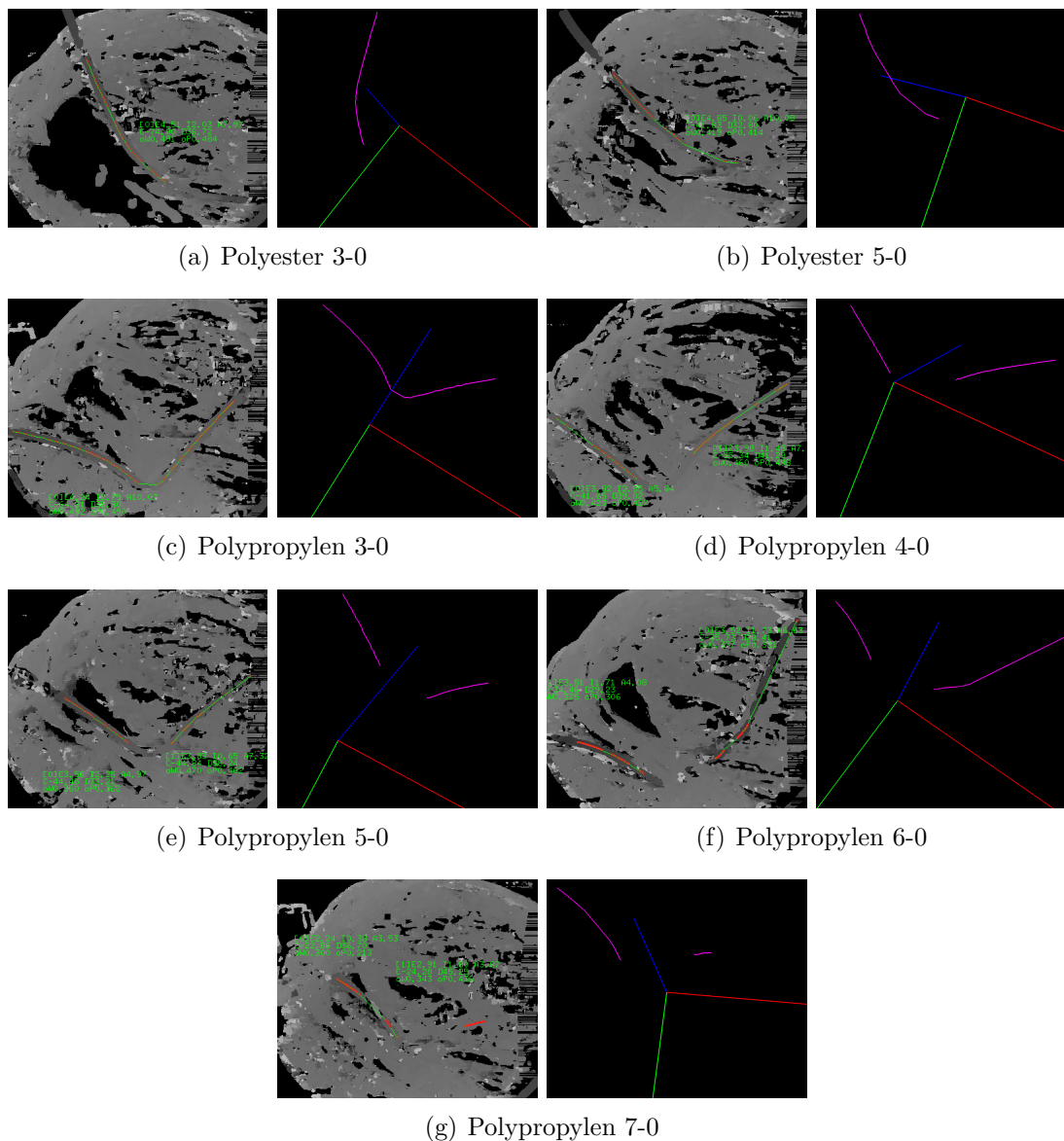


Abbildung 5.20: Disparitätskarten und rekonstruierte Fadenverläufe

Die physikalischen Eigenschaften chirurgischer Nahtmaterialien erlauben jedoch, mittels Freiformkurven auch solche Lücken zu schliessen. Dabei kann sowohl ein glatter Übergang als auch eine genügend hohe Anzahl von Linienpunkten gewährleistet werden. Das Kapitel 6 “3D-Rekonstruktion des Situs” widmet sich ausführlich der Modellbildung vom Nahtmaterial und Herzoberfläche.

Kapitel 6

3D-Rekonstruktion des Situs

Zusammenfassung Obwohl binokulare Sicht für minimalinvasive endoskopische Eingriffe nicht zwingend erforderlich ist, benötigt ein (teil-) autonomes robotergestütztes Chirurgesystem aus Sicherheitsgründen zusätzlich Positions- und Größenangaben über das Operationsgebiet, welche nur über binokulare Sicht in ausreichender Genauigkeit zur Verfügung stehen. Der Fadenerkennung nachgelagert, besteht die primäre Aufgabe dieser Stufe der Bildverarbeitung darin, anhand von Bildpaaren eines Stereoendoskops die Lage des Nahtmaterials als 3D-Modell im Koordinatensystem der chirurgischen Instrumente bereitzustellen. Die Erstellung eines hinreichend genauen Modells der Herzoberfläche ist ebenfalls notwendig. Die 3D-Rekonstruktion wird mit Hilfe der Lösung des Korrespondenzproblems in überlappenden Bildbereichen (Disparitätsberechnung) und anhand der bekannten Geometrie des Kamerasystems (Stereo-Triangulation) durchgeführt.

Es gibt mehrere Aspekte, die eine Modellierung und symbolische Repräsentation des Operationssitus motivieren. Zum einen benötigen zwei Stadien des teilautonomen Knotenlegens exakte Positionsangaben über das Nahtmaterial. So muss während des Umwickelns gewährleistet sein, dass die Schlaufen nicht vom Greifer rutschen. Um das Fadenende automatisch greifen zu können bedarf es ebenfalls einer genaueren Kenntnis der Lage. Mit Hilfe der Steuerung des Robotersystems können Kollisionen zwischen den Instrumenten vermieden werden. Um jedoch auch unbeabsichtigte Zusammenstöße zwischen Instrumenten und Herzoberfläche zu vermeiden, ist ihre Rekonstruktion und Modellierung erforderlich. Wie bereits eingangs erwähnt, führt die eingeschränkte Sicht der Endoskope oft zu Orientierungsproblemen während des Eingriffs. Dieser Umstand kann dadurch kompensiert werden, dass zu Beginn ein 3D-Übersichtsmodell des Operationsgebietes erstellt, und um den jeweils aktuell beobachteten Ausschnitt eingeblendet wird. Die kinematische Einbindung des Kamerasystems in das Robotersystem ermöglicht außerdem eine kontinuierliche Anpassung des Modells bei Bewegungen der Kamera. Desweiteren ist Modellierung die Voraussetzung für Augmentierungstechniken, wie beispielsweise die Überblendung realer Szenen mit klinisch relevanten Zusatzinformationen.

Klassische Verfahren der Stereorekonstruktion bilden Dreiecksnetze aus berechneten 3D-Punktwolken, wobei die Hauptschwierigkeit darin besteht, die richtigen Punkte zu einem Dreieck zusammenzufassen, bzw. die nachträgliche Bearbeitung des so entstandenen Gittermodells. Wenn es jedoch a priori Wissen über die zu rekonstruierende Szene verfügbar ist, besteht die Möglichkeit mit Hilfe von Heuristiken sowohl die Rekonstruktionsgenauigkeit, als auch die Ausführungszeit positiv zu beeinflussen. So hat sich beispielsweise herausgestellt, dass die physikalischen Eigenschaften von chirurgischem Nahtmaterial bestimmte Entsprechungen zu geometrischen Merkmalen von Freiformkurven haben. Insbesondere scheinen kubische Splines das notwendige Kontinuitätskriterium bei der Approximation des Fadenverlaufs zu erfüllen. Um möglichst hohe räumliche Genauigkeit zu erzielen, werden für die 3D-Rekonstruktion nur diejenigen Punkte aus der Disparitätskarte entlang des Fadenverlaufs herangezogen, für welche das Stereo-Korrespondenzproblem sehr gut gelöst werden konnte. Hierdurch entsteht zwar ein lückenhafter Punkteverlauf, doch die physikalischen Eigenschaften von chirurgischem Nahtmaterial erlauben es, durch die Interpolation einer Spline-Kurve auch dann ein genaues Modell zu erstellen, wenn Bereiche des Fadens verdeckt sind.

In gewissem Maße ist eine ähnliche Analogie zur Modellierung der Herzoberfläche verwendbar. Der betrachtete Ausschnitt ist überwiegend konvex, Selbstverdeckungen treten praktisch nicht auf. Die stellenweise stark texturierte Oberfläche des Herzens ermöglicht eine optimale Auswahl von Kandidaten als Stützpunkte für die Annäherung mittels einer Freiformfläche. Dadurch dass manche Bereiche jedoch fast keinerlei Textur aufweisen (beispielsweise wegen Fetteinlagerung), ist die Verwendung eines Dreiecksnetzes für die Modellierung zunächst nachteilig. Freiformflächen ermöglichen hier einen glatten Übergang zwischen gut rekonstruierbaren Bereichen. Bei Bedarf kann der Hintergrund auch in mehrere Quadranten aufgeteilt werden, wodurch lokale Unebenheiten besser angenähert werden können. Weil effiziente Algorithmen zur Kollisionserkennung ein Modell auf Dreiecksbasis benötigen, sind Freiformflächen in ihrer mathematischen Beschreibung hierfür zunächst nachteilig. Doch betrachtet man die Problematik global, so ergeben sich Vorteile für Freiformflächen und Tessellation (Zerlegung in primitive Flächen) gegenüber ursprünglichen Dreiecksnetzen. Ein weiterer Vorteil, der sich durch die Tessellation ergibt ist, dass die Darstellung angepasst an die Anforderungen in beliebiger Auflösung erfolgen kann. Die Texturierung solcher Flächen gestaltet sich im Vergleich mit polygonbasierten Flächen ebenfalls einfacher, und wird von moderner Grafikhardware unterstützt.

6.1 Motivation

Zur Demonstration möglicher teilautonomer Fähigkeiten eines robotergestützten minimalinvasiven Chirurgesystems, wie es im Abschnitt 2.6 anhand des ARAMIS²-Systems vorgestellt wurde, haben wir das chirurgische Knotenlegen exemplarisch ausgesucht. Die Notwendigkeit der Fadenerkennung- und Rekonstruktion ergibt sich aus offensichtlichen Sicherheitsanforderungen. Das System beobachtet in der Lernphase, wie der Chirurg speziell angepasste Instrumentenknoten knüpft, und ist anschließend in der Lage, die vom menschlichen Operateur initiierte Knoten kontextabhängig zu vervollständigen [73, 74, 75].

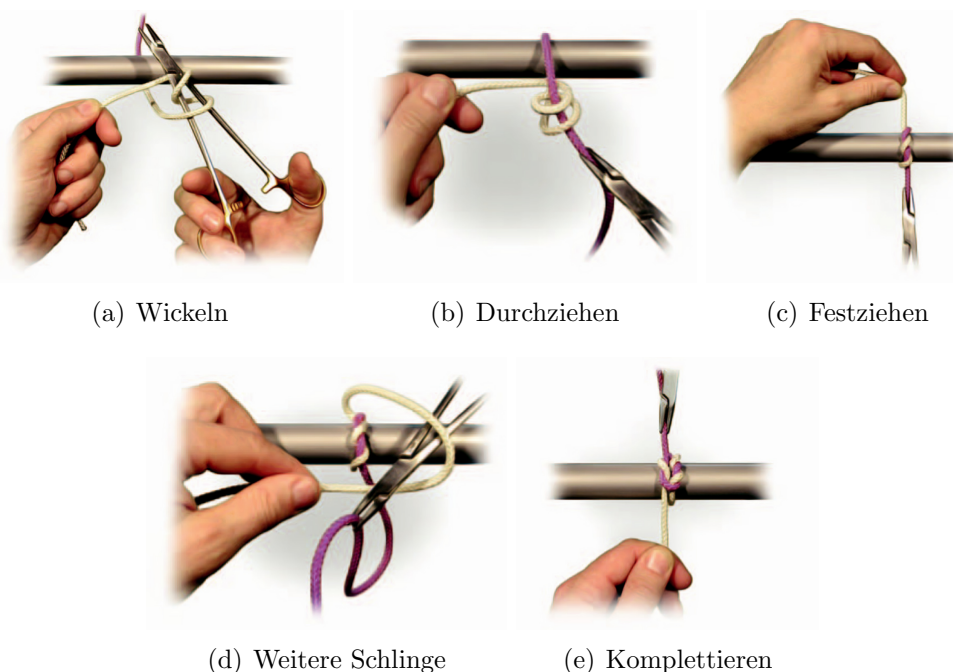


Abbildung 6.1: Stadien eines Instrumentenknotens (Quelle: Johnson & Johnson [118])

Instrumentenknoten werden unter Zuhilfenahme chirurgischer Zangen und Nadelhaltern hauptsächlich bei minimalinvasiven Eingriffen angewendet, wo die Platzverhältnisse das Knüpfen mit beiden Händen verhindern. Abbildung 6.1 zeigt die Stadien eines typischen Instrumentenknotens, wobei alle Teilschritte auch mit zwei Instrumenten durchführbar sind. Die ersten beiden Schlingen werden durch Legen des Fadens um den Nadelhalter herum gebildet (a). Das freie Fadenende wird nun gegriffen, und durch die beiden Schlingen gezogen (b). Die erste Hälfte des Knotens wird durch Ziehen der Fadenenden vollendet (c). Eine weitere Schlinge wird gelegt (d), und anschließend nochmals festgezogen (e).

Bei robotergestützten minimalinvasiven Eingriffen beschränkt man sich wegen der schwierigeren Handhabung meistens auf eine vereinfachte Ausführung, bei der auf

Teilschritt (d) verzichtet wird. Dabei wird mit einem Instrument die Nadel gehalten, und an der gewünschten Stelle ein Stich gesetzt. Mit dem anderen Instrument wird die freiwerdende Spitze der Nadel gegriffen, und der Faden entsprechend herausgezogen. Allerdings achtet man immer darauf, dass die Nadel ständig gegriffen bleibt. Ein Verlust der Nadel birgt große Gefahren und verursacht im günstigsten Fall nur eine Verzögerung. Nun kann das Nahtmaterial um die geöffneten Backen des freien Instruments gewickelt werden. Nach meistens zweimaligem Umwickeln greift dieses Instrument das Fadenende und zieht es durch die entstandene Schlaufe. Abbildung 6.2 zeigt die drei Stadien des Knotenlegens bei einem robotergestützten Eingriff mit dem *da Vinci*-System von Intuitive Surgical [15]. Wie bereits bei der Beschreibung des ARAMIS²-Systems erwähnt, kommen auch bei unserem Experimentalaufbau die gleichen minimalinvasiven Instrumente zum Einsatz.

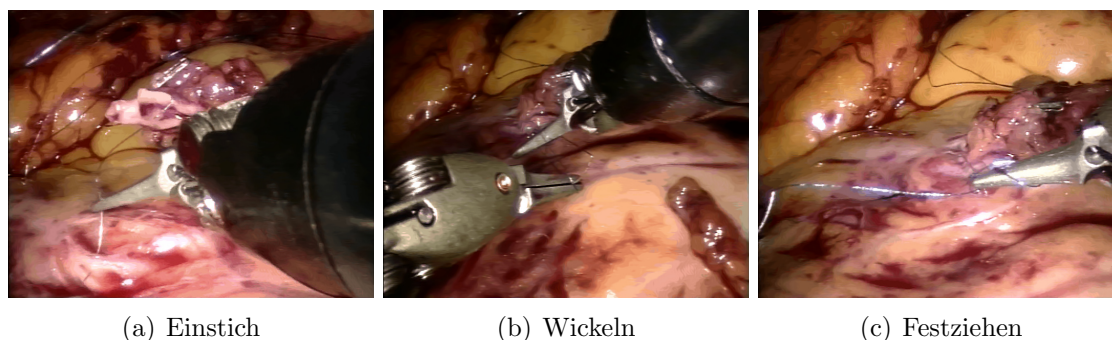


Abbildung 6.2: Instrumentenknoten im robotergestützten minimalinvasiven Fall (Quelle: Deutsches Herzzentrum München)

Der Einstich der Nadel wird aus offensichtlichen Gründen immer dem Chirurgen vorbehalten bleiben, allerdings sind die Teilaufgaben Umwickeln, Greifen und Anziehen gut automatisierbar. Doch dafür muss dem Robotersystem die genaue Lage des Fadens und der Herzoberfläche bekannt sein. Die Wickelvorgänge können zwar im optimalen Fall “blind” vom System durchgeführt werden, doch für die Vervollständigung des Knotens durch Anziehen muss das Fadenende autonom gegriffen werden. Anschließend kann der Knoten durch kraftgeregeltes Ziehen mit einer für das Nahtmaterial optimalen Kraft komplettiert werden. Zusätzliche Informationen, wie die relative Lage zwischen Nahtmaterial, Gewebe und Greifer sind ebenfalls notwendig, um das Gewebstrauma zu minimieren. Dafür muss ein hinreichend genaues Modell des Operationsgebiets angefertigt, und möglichst oft aktualisiert werden. Die Voraussetzung hierfür ist das Vorhandensein eines Stereoendoskops, mit Hilfe dessen sowohl die Erkennung als auch die dreidimensionale Rekonstruktion der relevanten Teile des Operationssitus durchführbar ist. Eine über die Bildverarbeitung geschlossene Steuerungsschleife Sensorik \leftrightarrow Aktorik gewährleistet, dass auf alle Veränderungen in der Szene, Fehlerfälle eingeschlossen, reagiert werden kann. Im nächsten Abschnitt werden wir einige Aspekte der binokularen Stereobildverarbeitung vorstellen, insbesondere im Hinblick auf die Besonderheiten der Stereoendoskopie.

6.2 Binokulare Stereobildverarbeitung

Ein Bild ist die Projektion einer dreidimensionalen Szene auf eine Fläche. Bei diesem Abbildungsprozess geht jedoch die Entfernungsinformation größtenteils verloren. Um diese Information wiederzugewinnen sind mindestens zwei Projektionen desselben Punktes notwendig. Der Mensch und die meisten höherentwickelten Tiere besitzen zwei Augen, um eine volumetrische Repräsentation ihrer Umgebung zu erlangen. Analog kann man in technischen Anwendungen mit Hilfe von zwei Kameras und unter Kenntnis der Geometrie des Aufbaus und einiger Parameter der beteiligten Sensoren dreidimensionale Koordinaten einer beobachteten Szene berechnen. Diesen Prozess nennt man in der Bildverarbeitung *3D- oder Stereo-Rekonstruktion*. Obwohl Stereobildverarbeitung auch mit einem beweglichen monokularen Aufbau möglich ist, werden wir fortan den Begriff "Stereo" für starre binokulare Anordnungen von Kameras verwenden. Der nächste Abschnitt stellt das notwendige mathematische Gerüst vor, um dreidimensionale Information aus Bildern zweier örtlich versetzten Kameras zu gewinnen.

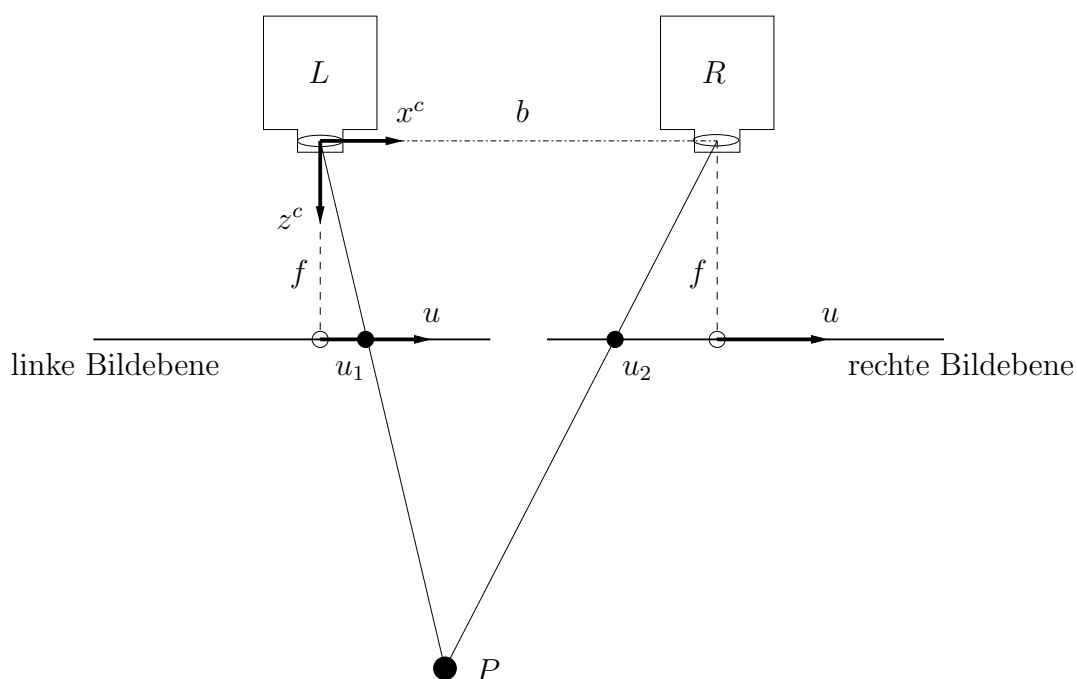


Abbildung 6.3: Schematische Darstellung eines binokularen Aufbaus von Kameras

6.2.1 Abbildungsgeometrie

Betrachten wir zunächst den stark vereinfachten binokularen Aufbau von zwei identischen 1D-Kameras aus Abbildung 6.3. Ihre optischen Achsen sind parallel, und die Verbindungsgerade zwischen ihren optischen Zentren, die Basis, fällt mit der

x -Achse der linken Kamera zusammen. Sind die Brennweiten f und die Basisbreite b bekannt, so kann der Punkt $P(x^c, z^c)$ in die zwei Bildebenen projiziert werden. Es ergeben sich folgende Bildkoordinaten:

$$u_1 = f \frac{x^c}{z^c}, \quad u_2 = f \frac{x^c - b}{z^c} \quad (6.1)$$

Die zwei Punkte u_1 und u_2 , die aus der Projektion desselben Punktes in die Bildebenen entstehen, nennt man homologe Punkte. Die Entfernung zwischen zwei solchen Punkten in der Bildebene nennt man die Disparität d :

$$d = (u_1 - u_2) = \frac{f \cdot b}{z^c} \quad (6.2)$$

Bei gegebenen Kameraparametern und Bildkoordinaten der homologen Punkte, kann die Entfernung z^c eines Punktes P zum Kamerasystem berechnet werden als:

$$z^c = \frac{f \cdot b}{d} \quad (6.3)$$

Bereits anhand dieses einfachen Modells lassen sich die zwei wichtigsten Aufgaben der Stereobildverarbeitung formulieren: (i) die möglichst genaue Ermittlung der internen (u.a. Brennweite f) und externen (u.a. Basisbreite b) Kameraparameter, sowie (ii) die Bestimmung homologer Punkte aus Bildpaaren. Doch bevor wir uns der 3D-Rekonstruktion einer Szene mit Hilfe eines realen stereoskopischen Kameraaufbaus widmen, wollen wir zunächst einen wichtigen Aspekt untersuchen, und zwar die theoretischen Grenzen der Rekonstruktionsgenauigkeit.

Im vorherigen Kapitel haben wir die Genauigkeit der Segmentierung anhand von Herstellerangaben zum Fadendurchmesser untersucht. Dabei wurden etwaige Fehler, die in erster Linie auf Ungenauigkeiten im ermittelten Lochkameramodell zurückzuführen sind, außer Acht gelassen. Die Kalibrierung eines Stereokamerasystems anhand der Abbildungsgeometrie des Lochkameramodells erfüllt nämlich wegen Fertigungsdefekten nicht immer die erwarteten Ansprüche an Genauigkeit.

6.2.2 Genauigkeitsermittlung mittels 3D-Kalibrierkörper

In diesem Abschnitt werden wir die Rekonstruktionsgenauigkeit des verwendeten 3D-Endoskops (R. Wolf GmbH) empirisch untersuchen. Insbesondere interessiert es uns, wie akkurat sich die Tiefeninformation in Abhängigkeit von Kameraparametern und Entfernung zwischen Objekt und Endoskop ermitteln lässt.

Aus Gleichung 6.3 lässt sich die Tiefenauflösung eines kalibrierten Aufbaus bestimmen, d.h die Genauigkeit, mit der die Entfernung von Objekten zum Kamerasystem

bestimmt werden können. Für das Modell aus Abbildung 6.3 berechnet sich der absolute Fehler anhand von Gleichung 6.3 wie folgt:

$$\Delta z = \frac{z^2}{f \cdot b} \cdot \Delta d. \quad (6.4)$$

Die Tabelle aus Abbildung 6.4 stellt die hierfür benötigten Parameter für die 3D-Endoskope der Firmen R. Wolf und Schölly (IS) gegenüber. Sie sind jeweils das Ergebnis einer Stereokalibrierung, und entsprechen im entzerrten und rektifizierten Fall dem Modell aus Abbildung 6.3. Desweiteren gehen wir davon aus, dass Disparitäten nur bis auf einen Pixel genau bestimmt werden können.

Parameter [m]	R. Wolf	Schölly (IS)
f	0.00334143	0.00873355
b	0.00457138	0.0057099
S_x, S_y	1.036e-05	7.4e-06
κ	0.0	0.0

Abbildung 6.4: Parameter der rektifizierten Endoskope R. Wolf und Schölly (IS)

Nun ist es möglich, für bestimmte Entfernungsbereiche die theoretische Tiefenauflösung anzugeben. Abbildung 6.5 zeigt für beide Endoskope den Verlauf in einen typischen Arbeitsbereich bis 12 cm.

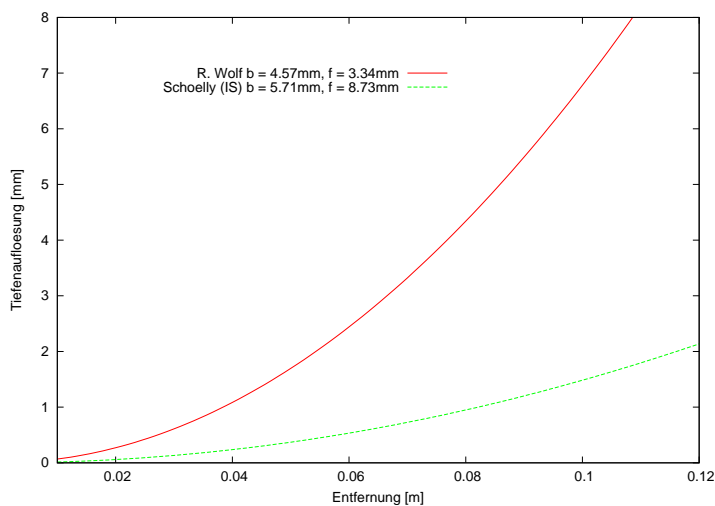


Abbildung 6.5: Entfernungsabhängige Tiefenauflösung

Je größer die Basisbreite b und die Brennweite f sind, desto genauer kann die Tiefeninformation extrahiert werden. Das Schölly-Endoskop des *da Vinci* bleibt theoretisch bis auf fast 10 cm Entfernung submillimetergenau. Beim Wolf-Endoskop

hingegen ist eine Verwendung der Tiefeninformation zu Zwecken der Bildverarbeitung scheinbar lediglich in dem Bereich 4-7 cm sinnvoll. Nichtsdestotrotz könnte ein menschlicher Operateur auch bei solchen Abweichungen noch sicher arbeiten.

Gleichung 6.4 gibt lediglich Aufschluss über die zu erwartende *Abweichung* von einem Idealwert, aber nicht darüber, in welcher Richtung die Verfälschung auftritt. Um auch dies zu ermitteln und zur Überprüfung der Ergebnisse aus Abbildung 6.5 wurde der bereits vorgestellte 3D-Kalibrierkörper (Abb. 6.6) verwendet. Dazu wurden die Eckpunkte der pyramidenförmigen Balken im linken und rechten entzerrten Bild manuell pixelgenau markiert und anschließend rekonstruiert.

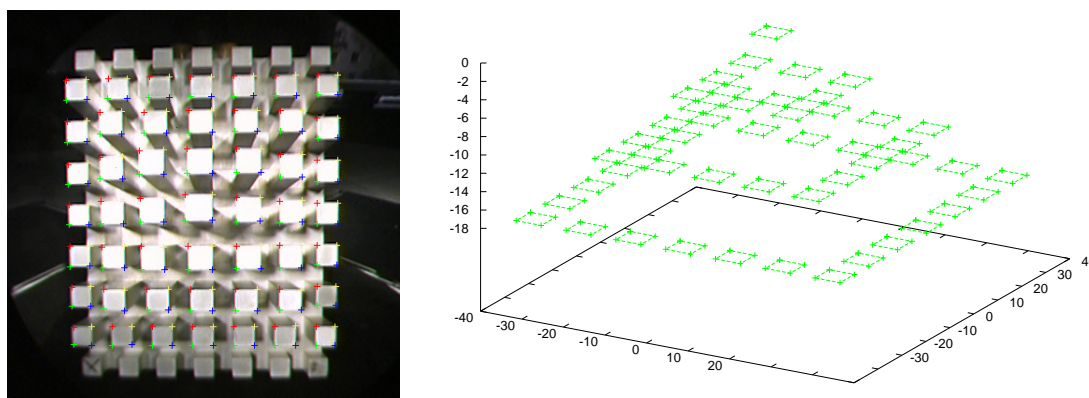


Abbildung 6.6: 3D-Kalibrierkörper

Die Spitze der Pyramide wurde in einer Entfernung von etwa 66 mm vom Endoskop mit ihrer Basis parallel zur Linse aufgestellt. Die Quaderförmigen Balken mit einer Seitenlänge von 6 mm sind jeweils 6 mm voneinander entfernt, die Höhe der Stufen beträgt ebenfalls 6 mm. Die unterste Ebene wurde bei den Berechnungen nicht berücksichtigt, da sie sich zu nah am Bildrand, und dadurch im stark verzerrten Bereich befindet. Das Volumen des vermessbaren Arbeitsraumes beträgt $78 \times 78 \times 18$ mm, wenn wir die Entfernung zum Endoskop nicht mitzählen.

Die Punkte an der Spitze der Pyramide (etwa in der Bildmitte) dienten dabei als Referenz, denn hier ist die Rekonstruktion erwartungsgemäß am genauesten. Anhand der Referenzpunkte, die als "fehlerfrei" rekonstruiert betrachtet wurden, und anhand der genauen Abmessungen des Kalibrierkörpers, können die Idealpositionen der restlichen Punkte ebenfalls im Koordinatensystem der Kamera angegeben werden. Diese Referenzpunkte können nun mit den rekonstruierten verglichen werden. Zusätzlich zur Tiefenabweichung kann natürlich auch der Rekonstruktionsfehler in den Ebenen parallel zur Linse ermittelt werden.

Abbildung 6.7 zeigt den rekonstruierten Kalibrierkörper im Koordinatensystem der linken rektifizierten Kamera. Die leichten Tiefenabweichungen zwischen den Eckpunkten eines Quaders sind darauf zurückzuführen, dass ihre Positionen in Bildkoordinaten lediglich mit Pixelgenauigkeit angegeben wurden. Dies wurde bewusst

so gehandhabt, damit die Ergebnisse mit dem Verlauf aus Abbildung 6.5 verglichen werden können. Interessanterweise sind die Rekonstruktionsergebnisse besser als von Gleichung 6.4 vorausgesagt.

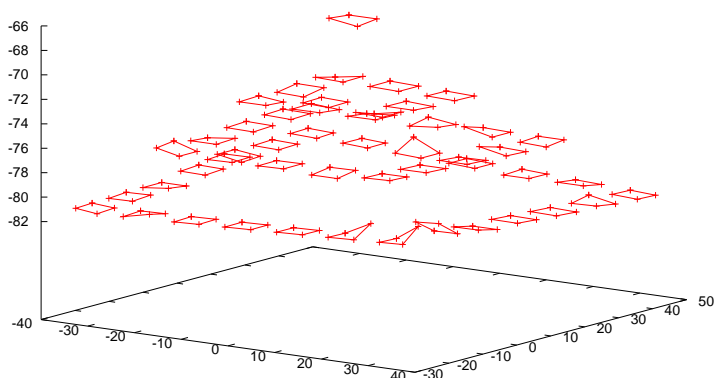


Abbildung 6.7: Stereorekonstruktion des Kalibrierkörpers

Betrachtet man die berechneten Positionen mit den entsprechenden Referenzpunkten im gleichen Koordinatensystem, so können die Abweichungen in allen drei kartesischen Richtungen ermittelt werden. Abbildung 6.8 zeigt sowohl Rekonstruktionsfehler in der Ebene als auch in der Tiefe.

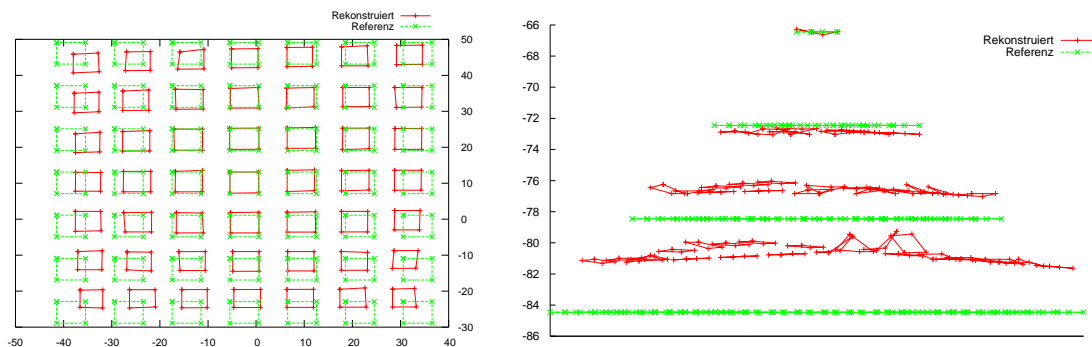


Abbildung 6.8: Abweichungen der Rekonstruktion von der Referenz

Innerhalb eines Kreises mit einem Durchmesser von etwa 20 mm in der Bildmitte bleibt die durchschnittliche Abweichung in der Ebene unter einem Millimeter. Danach kann man bis zu einem Durchmesser von 40 mm immerhin mit Genauigkeiten im Millimeterbereich rechnen, größtenteils unabhängig von der Tiefe. Diese Ergebnisse sind in etwa deckungsgleich mit den Beobachtungen aus dem Abschnitt 5.3.1 zur Fadenerkennung. Um einiges höher fällt hingegen der entfernungabhängige Tiefenfehler aus. Er beträgt in der ersten Ebene unter der Pyramidenspitze im

Schnitt 0.42 mm. Auf den nachfolgenden Ebenen wächst die Abweichung auf nahezu 2 bzw. 4 mm. Dabei werden Punkte immer als näher zum Endoskop rekonstruiert, als tatsächlich der Fall war. Wäre der aufgespannte Volumen des Kalibrierkörpers größer, so könnte man eine feiner aufgelöste Korrekturkarte erstellen, die eine Anpassung der rekonstruierten Positionen ermöglichen würde. Die Tabelle aus Abbildung 6.9 fasst die Ergebnisse nochmal zusammen.

Entfernung $Z/(X, Y)$	X (horizontal)			Y (vertikal)			Z (Tiefe)		
	<i>Min.</i>	<i>Max.</i>	<i>Mtl.</i>	<i>Min.</i>	<i>Max.</i>	<i>Mtl.</i>	<i>Min.</i>	<i>Max.</i>	<i>Mtl.</i>
7.2/12 mm	0.026	0.874	0.272	0.058	1.309	0.595	0.164	0.602	0.420
7.8/24 mm	0.000	2.647	1.029	0.128	2.821	1.131	1.423	2.407	1.853
8.4/36 mm	0.129	4.897	2.159	0.037	4.518	2.070	2.821	5.187	3.718

Abbildung 6.9: Entfernungsabhängige Rekonstruktionsfehler

Festzuhalten bleibt, dass bis zu einer Entfernung von etwa 8 cm die Tiefengenauigkeit des Wolf-Endoskops für das autonome Greifen des Fadens ausreicht. Geringe Abweichungen von der realen Position können noch aufgrund der Abmessungen der Greiferbacken toleriert werden. Der Fehler in der Ebene ist im typischen Arbeitsbereich vernachlässigbar, und kann nur unter extremer Schiefelage des Endoskops zu Problemen führen.

6.2.3 Rekonstruktion von 3D-Punktwolken

Epipolargeometrie

Wird eine Szene gleichzeitig von mindestens zwei Kameras aus unterschiedlichen Richtungen beobachtet, so können anhand der aufgenommenen Bilder für ausgewählte Punkte dreidimensionale Koordinaten berechnet werden. Wie bereits eingangs erwähnt, ist dafür die genaue Kenntnis der Geometrie des Kamerasystems, und das Auffinden korrespondierender Punkte in zeitgleichen Aufnahmen notwendig. Bei einer starren Anordnung von Kameras, wie dies bei medizinischen Endoskopen der Fall ist, kann die Kalibrierung vor einem Eingriff durchgeführt werden, und bleibt üblicherweise auch während der Operation erhalten. Die Hauptschwierigkeit bleibt nun das Auffinden homologer Punkte, auch als Stereo-Korrespondenzproblem genannt, in einer sich ständig verändernden Umgebung.

In einer statischen und regelmäßig strukturierten Szene, wie dies beim Kalibrierkörper der Fall war, konnten korrespondierende Punkte leicht einander zugeordnet werden. Wird jedoch ein realer Operationssitus beobachtet, gestaltet sich die Lösung des Korrespondenzproblems erheblich komplizierter. Eine Vereinfachung des Problems ermöglicht die Kenntnis der Geometrie der Stereoanordnung. In einem ersten Schritt werden die Bilder entzerrt, wodurch gekrümmte Abbildungen gerader Linien berichtigt werden. Der zweidimensionale Suchraum des Korrespondenz-

problems kann durch die epipolare Einschränkung auf eine Linie reduziert werden, diesen Prozess nennt man *Rektifizierung*. Sie bewirkt, dass durch Anpassung der Brennweiten und durch Rotationen um die Brennpunkte die Bildebenen der beiden Kameras koplanar werden. Abbildung 6.10 zeigt schematisch die Vorgehensweise.

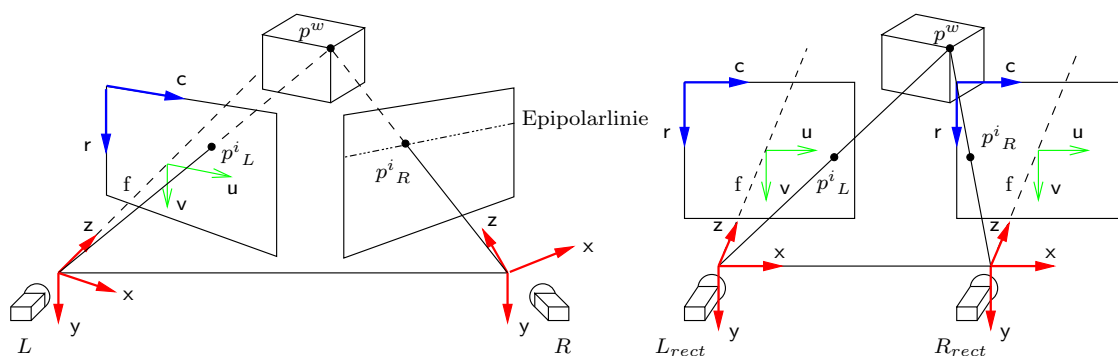


Abbildung 6.10: Epipolargeometrie und Rektifizierung

Korrespondenzproblem

Homologe Punkte haben nun im Idealfall gleiche Zeilenkoordinaten, leichte Abweichungen können jedoch aufgrund suboptimal ermittelten Kameraparameter auftreten. Für die Bestimmung der Disparität, was der Differenz der Spaltenkoordinaten korrespondierender Punkte entspricht, sind verschiedene Verfahren anwendbar. Die zwei prominentesten sind die merkmalsbasierten und die flächenbasierten Verfahren. *Merkmalsbasierte Verfahren* versuchen eine möglichst vollständige symbolische Repräsentation relevanter Bildinhalte zu erreichen, um anschließend nach korrespondierenden Merkmalen der Objekte zu suchen. Im einfachsten Fall liefern Kantendetektoren die notwendige Abstraktion. Für die generelle Lagebestimmung von chirurgischem Nahtmaterial würde sich dieser Ansatz prinzipiell eignen, zumal der Faden noch vor der Rekonstruktion segmentiert wurde. Für das Greifen werden jedoch genaue Positionen einzelner Punkte benötigt, die auch entsprechend attribuiert sind (Anfang, Ende, Mitte). Bei der Rekonstruktion der Organoberfläche hingegen würde ein merkmalsbasierter Ansatz in den meisten Fällen versagen.

Flächenbasierte Verfahren eignen sich in nahezu jeder Domäne zur Lösung des Korrespondenzproblems. Sie gehen lediglich von der Annahme aus, dass sich die Bilder der beiden Kameras grundsätzlich ähneln. Ist diese Voraussetzung erfüllt, d.h. die Blickwinkel beider Kameras weichen nicht entscheidend voneinander ab, so können durch geeignet gewählte photometrische Ähnlichkeitsmaße korrespondierende Regionen gefunden werden. Diese Vorgehensweise benötigt kein a priori Wissen über die beobachtete Szene, und liefert eine dichte Disparitätskarte mit einem Disparitätswert pro Bildpunkt. In dieser Arbeit wurde durchgehend eine bereits vorgestellte

Art der normalisierten Kreuzkorrelation (ZNCC, siehe Gleichung 4.15) als Ähnlichkeitsmaß verwendet, trotz des vergleichsweise hohen Berechnungsaufwands. Sie ist immun gegen leichte Beleuchtungsunterschiede und erwies sich in dieser Domäne als besonders robust. Ein weiterer Vorteil flächenbasierter Ähnlichkeitsmaße ist der Umstand, dass sie eine qualitative Einordnung ermittelter Disparitäten ermöglichen.

Stereorekonstruktion

Aus bereits bestimmten Korrespondenzen können mit einfachen geometrischen Werkzeugen dreidimensionale Koordinaten für ausgewählte Bildpunkte berechnet werden. Dabei ist die externe Kalibrierung des Stereosystems Voraussetzung für die Gewinnung metrischer Daten. Für rektifizierte Systeme lässt sich die Rekonstruktionsaufgabe mit Hilfe des Strahlensatzes lösen. Als Ergebnis erhält man 3D-Punkte in einem kartesischen Koordinatensystem, üblicherweise der linken rektifizierten Kamera. Diese Punkte haben zunächst keine semantische Bedeutung und müssen erst Objekten der Szene zugeordnet werden. Dies geschieht anhand einer vorangegangenen Segmentierung und Bezeichnung von Objekten. Die Robustheit der Fadenerkennung garantiert, dass jegliches Nahtmaterial der Szene erkannt wird. Dadurch dass in unserem Anwendungsfall nur zwei Klassen von relevanten Objekten gibt (Faden und Gewebe), lässt sich die Herzoberfläche durch einfaches Maskieren des Fadens in der Disparitätskarte rekonstruieren.

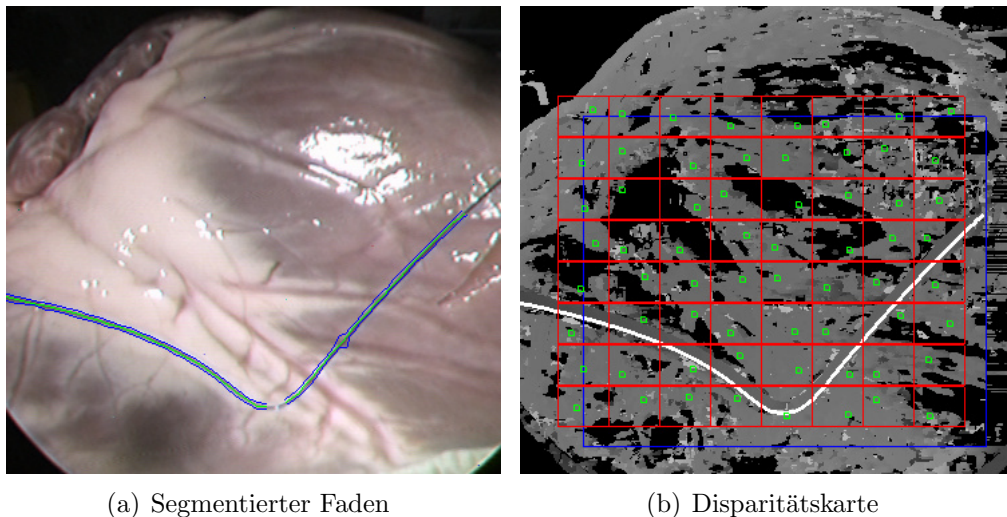


Abbildung 6.11: Segmentierter Fadenverlauf und dichte Disparitätskarte

Abbildung 6.11 zeigt die Vorstufen der Rekonstruktion. Zunächst wird das Nahtmaterial im linken rektifizierten Bild (6.11(a)) segmentiert. Dadurch dass die Segmentierung im bereits rektifizierten Bild erfolgt, können ermittelte Bildkoordinaten direkt für den Zugriff auf die Disparitätskarte herangezogen werden. Es werden nur Punkte

entlang des Fadenverlaufs rekonstruiert, die mindestens ein ZNCC-Korrelationscore $s > 0.9$ aufweisen, wobei $s \in [-1, 1]$. Diese Bedingung wird in der Hoffnung aufgestellt, dass dadurch nur Bereiche des Fadens rekonstruiert werden, deren segmentierte Linienmittelpunkte sich nicht nahe an Pixelgrenzen befinden. Für solche Punkte ist die Bestimmung der Disparität mit einem Fehler behaftet. Aufgrund dieser Tatsache, und einer Spiegelreflexion genau in der Mitte, entsteht im Beispiel aus Abbildung 6.11(a) eine größere Lücke im rekonstruierten 3D-Fadenverlauf. In einem späteren Abschnitt wird gezeigt, wie diese unter Gewährleistung eines glatten Übergangs geschlossen werden kann.

Obwohl eine dichte Disparitätskarte (Abb. 6.11(b)) zur Verfügung steht, wird auf die Rekonstruktion der Herzoberfläche mit Hilfe eines Dreiecksnetzes bewusst verzichtet. Statt der Berechnung einer vollständigen 3D-Punktvolke und ihrer Triangulierung wird der Hintergrund in Quadranten aufgeteilt, und aus jedem von ihnen der Bildpunkt mit dem höchsten ZNCC-Korrelationscore ausgewählt. Dabei werden zum Fadenverlauf gehörende Punkte und ihre unmittelbaren Nachbarn ausgeschlossen, indem ihre Korrelationscores auf Null gesetzt werden.

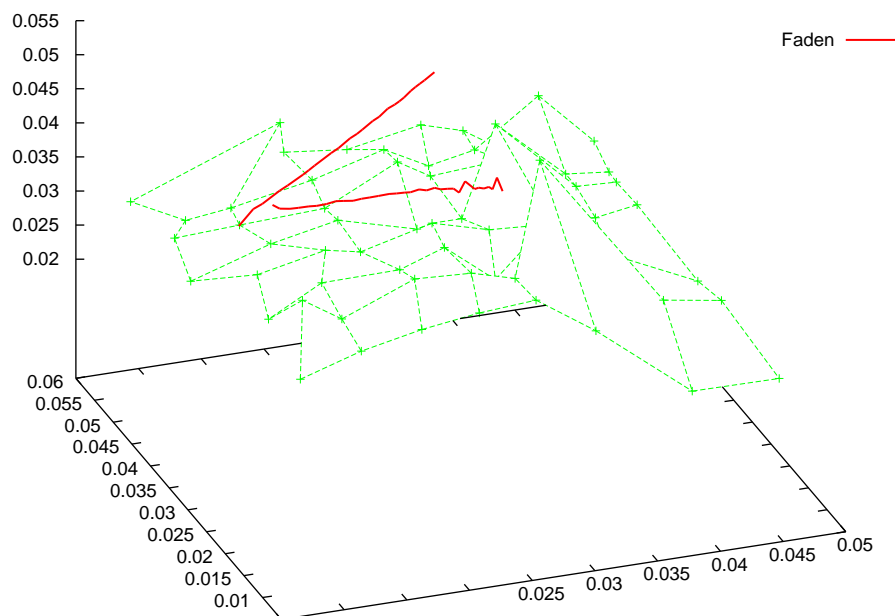


Abbildung 6.12: Rekonstruierte 3D-Punktvolken von Faden und Herzoberfläche

Die Nachteile einer polygonalen Fläche können anhand von Abbildung 6.12 angedeutet werden. Die leicht gewölbte, aber ansonsten größtenteils glatte Herzoberfläche, erscheint in der Polygonaldarstellung bereits bei 8×8 rekonstruierten Flächenpunkten stark zerklüftet. Außerdem stellt eine korrekte Texturierung der Fläche einen erheblichen Aufwand dar. Weil die (nichttriviale) Verfeinerung und Texturierung von Polygonnetzen nicht den Schwerpunkt dieser Arbeit darstellen soll, wird hier

ein anderer, auf Freiformflächen basierender Ansatz verwendet, der in einem späteren Abschnitt vorgestellt wird.

6.3 Symbolische Repräsentation des Fadens

In Abbildung 6.12 wurden zwecks der Darstellung die rekonstruierten Punkte des Fadenverlaufs einfach mit geraden Linien verbunden. Somit gibt es für Regionen zwischen den Punkten zunächst keine geometrische Repräsentation. Eine erste Näherung könnte man durch die stückweise lineare Interpolation von Zwischenpunkten mit Hilfe eines 3D-Bresenham-Algorithmus erreichen. Der offensichtliche Nachteil dieser Methode ist die starke Diskontinuität bei ausgedehnten Lücken im Fadenverlauf, und die große Anzahl zu speichernder Koordinaten. Eine symbolische Repräsentation mit Hilfe einer Polynominterpolation im Lagrangeschen Sinne scheidet ebenfalls aus, weil es mit wachsendem Polynomgrad zu starken Oszillationen zwischen den Stützpunkten kommt. Außerdem bewirken kleine (lokale) Änderungen in den Stützwerten große (globale) Änderungen im Kurvenverlauf. Die wichtigsten Anforderungen an eine symbolische Repräsentation lassen sich demnach wie folgt formulieren:

1. Durchlauf aller Stützstellen
2. Glatter Übergang an Unterbrechungsstellen
3. (Speicher-)Effiziente Handhabung des Fadenverlaufs

Die erste Bedingung ist durch die Tatsache motiviert, dass die Stützstellen (rekonstruierte Punkte des Fadenverlaufs) aufgrund der strengen Vorauswahl der Disparitätsberechnung sehr genau rekonstruiert werden können. Ein glatter Übergang an etwaigen Unterbrechungsstellen ist notwendig, damit Kurvenparameter wie Krümmung und Torsion an jeder Position unverfälscht berechnet werden können. Die effiziente Handhabung des rekonstruierten Fadenverlaufs wird durch den Verzicht auf die explizite Speicherung aller Kurvenpunkte erreicht. Dadurch ergeben sich mehrere Vorteile: zum einen ist dadurch die latenzfreie Netzwerkübertragung an die Robotersteuerung möglich, zum anderen können bei Bedarf für ausgesuchte Bereiche beliebig fein aufgelöste 3D-Koordinaten berechnet werden. Affine Transformationen der Kurve sind anhand einer geeigneten symbolischen Repräsentation ebenfalls effizienter durchführbar.

Eine geeignete Möglichkeit zur Interpolation unregelmäßiger Stützpunkte, die alle oben genannten Bedingungen erfüllen, bieten Splines. Insbesondere wird ihre Verwendung durch die Materialeigenschaften chirurgischer Fäden motiviert. Wir stellen in den nächsten Abschnitten das notwendige mathematische Gerüst vor, und evaluieren den gewählten Ansatz anhand realistischer Aufnahmen.

6.3.1 Splines

Motivation

Der englische Begriff “Spline” bedeutet soviel wie “biegsames Lineal”, und wurde ursprünglich im Schiffsbau verwendet, um die Rümpfe zu zeichnen. Eine lange dünne Latte (*Straklatte*) wurde an bestimmten Punkten durch Nägel fixiert, und ist dabei bestrebt, ihre durch die Biegungen hervorgerufene innere Spannung zu minimieren bzw. zu verteilen. Anhand dieser Analogie wurden in der Mathematik Splines als Funktionen eingeführt, die sich stückweise aus Polynomen zusammensetzen. Eine Spline n -ten Grades besteht dabei aus Polynomen mit maximalem Grad n . An den vorgegebenen Verbindungsstellen (Stützstellen) verlangt man, dass die Kurvenstücke ohne Sprung (stetig) und ohne Knick (differenzierbar) aneinanderversetzen. Außerdem wird gefordert, dass die Kurvenstücke in den Verbindungsstellen identische Krümmungen haben, d.h. die Krümmung soll dort stetig verlaufen, wodurch auch die zweite Ableitung existieren und stetig sein muss.

Die Biegung einer Straklatte ist der einer kubischen Spline sehr ähnlich, wenn die Randpunkte bestimmte Bedingungen erfüllen. Chirurgisches Nahtmaterial, insbesondere das aus monofilem Polypropylen gefertigt, hat ein ähnliches (dem Maßstab entsprechendes) Verhalten. Wir wollen die Analogie ausnutzen, dass bei der 3D-Rekonstruktion des Fadenverlaufs oft nur in unregelmäßigen Abständen geeignete Punkte zur Verfügung stehen. Dadurch dass die Positionen dieser wenigen Punkte (“Nägel”) sehr genau sind, erwarten wir dass die interpolierende Spline dem realen Fadenverlauf sehr ähnlich sein wird. Desweiteren sollen getrennt segmentierte Teilstücke desselben Fadens mittels glatter Übergänge verbunden werden.

Kubische Splineinterpolation

Wir definieren eine Raumkurve als eine eindimensionale, zusammenhängende Punktmenge im dreidimensionalen Raum, und fordern dass sie zumindest stückweise glatt ist. Es gibt grundsätzlich drei Möglichkeiten eine Kurve mathematisch zu beschreiben. Die Tabelle aus Abbildung 6.13 fasst diese für die Beschreibung eines Kreises in der Ebene zusammen.

implizit	explizit	parametrisiert
$h(x, y) = x^2 + y^2 - r^2 = 0$	Graph von $f(x) = \sqrt{r^2 - x^2}$	$F(t) = \begin{pmatrix} r \cdot \cos t \\ r \cdot \sin t \end{pmatrix}$
$h : \mathbb{R}^2 \rightarrow \mathbb{R}$	$f : \mathbb{R} \rightarrow \mathbb{R}$	$F : \mathbb{R} \rightarrow \mathbb{R}^2$

Abbildung 6.13: Mathematische Beschreibungsalternativen für einen 2D-Kreis

Durch diese Darstellungsformen anhand der a priori Kenntnis einer Kurve kann ihr Verlauf generiert werden. Oft stellt sich jedoch das Problem umgekehrt, und muss

aus einem vorgegebenen Verlauf eine mathematische Repräsentation generiert werden. Hierfür eignet sich insbesondere die parametrisierte Darstellungsform. Je nach Aufgabenstellung müssen Funktionen ermittelt werden, welche durch die vorgegebenen Punkte verlaufen (Interpolationsproblem), oder diese annähern (Approximationsproblem). Betrachten wir dazu eine beliebige Funktion

$$\mathbf{f} : D \rightarrow \mathbb{R}^m \quad (6.5)$$

mit dem Definitionsbereich $D = [t_{min}, t_{max}]$ und dem Wertebereich bestehend aus m -Tupeln mit $m \geq 1$. Insbesondere gegeben sei o.B.d.A für $m = 3$ eine Raumkurve, die in Form diskreter (Mess-)Punkte $\{t_i, \mathbf{x}_i\}_{i=0}^n$ angegeben ist. Dabei stellen die Werte $\mathbf{x}_i = \{x_i(t_i), y_i(t_i), z_i(t_i)\}$ kartesische Koordinaten dar, und unsere einzige Information über \mathbf{f} ist, dass $\mathbf{x}_i = \mathbf{f}(t_i)$ und $t_{min} \leq t_0 < t_1 < \dots < t_{n-1} < t_n \leq t_{max}$. Nun wollen wir die unbekannte Funktion \mathbf{f} mit Hilfe einer stückweise kubischen Polynomspline S der Form

$$S(t) = \begin{cases} S_0(t), & t \in [t_0, t_1] \\ S_1(t), & t \in [t_1, t_2] \\ \vdots \\ S_{n-2}(t), & t \in [t_{n-2}, t_{n-1}] \\ S_{n-1}(t), & t \in [t_{n-1}, t_n] \end{cases} \quad (6.6)$$

bestehend aus Polynomen dritten Grades der allgemeinen Form

$$S_i(t) = a_i(t - t_i)^3 + b_i(t - t_i)^2 + c_i(t - t_i) + d_i \quad (6.7)$$

für $0 \leq i \leq n - 1$ interpolieren. Bei n Polynomen mit jeweils vier Koeffizienten sind $4n$ Unbekannte zu bestimmen, wofür ein System aus $4n$ Gleichungen aufzustellen und zu lösen ist. Die Interpolationsbedingung besagt, dass die Polynome die Stützstellen durchlaufen müssen, wodurch folgende Bedingungen aufgestellt werden können:

$$S_i(t_i) = \mathbf{x}_i, \quad 0 \leq i \leq n - 1, \quad \text{und} \quad S_{n-1}(t_n) = \mathbf{x}_n \quad (6.8)$$

Dadurch ergeben sich zunächst $n + 1$ Gleichungen. Da wir für $S(t)$ außerdem eine C^2 -Kontinuität (keine Sprungstellen, gleicher Tangentenvektor und gleicher Schmiegkreis) an den Knotenpunkten verlangen, müssen $S(t)$, $S'(t)$ und $S''(t)$ im Intervall $[t_0, t_n]$ stetig sein. Jedes $S_i(t)$ ist C^2 -stetig im offenen Intervall (t_i, t_{i+1}) , sodass zunächst nur das Verhalten an den Stützstellen von Interesse ist. Die Polynome, ihre erste und zweite Ableitungen müssen somit an den Übergängen der Stützstellen gleich sein:

$$S_i(t_{i+1}) = S_{i+1}(t_i), \quad S'_i(t_{i+1}) = S'_{i+1}(t_i), \quad S''_i(t_{i+1}) = S''_{i+1}(t_i), \quad 0 \leq i \leq n - 2 \quad (6.9)$$

Dadurch ergeben sich weitere $3(n - 1)$ Gleichungen, wodurch zunächst insgesamt $4n - 2$ zur Verfügung stehen. Für die verbleibenden zwei Freiheitsgrade können Bedingungen an das Verhalten an den äußeren Randpunkten aufgestellt werden. Die

natürliche Randbedingung verlangt, dass die zweiten Ableitungen an den Randpunkten verschwinden:

$$S''_0(t_0) = 0, S''_{n-1}(t_n) = 0 \quad (6.10)$$

Das nun vollständige lineare Gleichungssystem ist tridiagonal und kann mit einer Laufzeitkomplexität von $O(n)$ gelöst werden, im Gegensatz zu allgemeinen Lösern mit einer Komplexität von $O(n^3)$. Eine weitere nützliche Eigenschaft der natürlichen Randbedingung ist, dass dadurch die Steigungen an den Randpunkten beliebige Werte annehmen können. Sie minimieren die Oszillation der Kurve mit der Folge, dass unter allen interpolierenden C^2 -Funktionen die natürliche kubische Spline die "glatteste" ist.

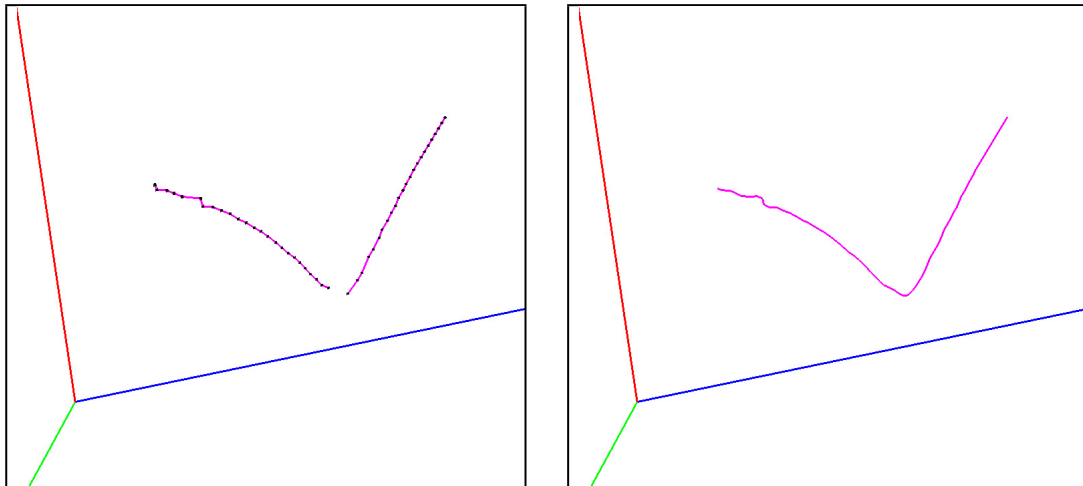


Abbildung 6.14: 3D-Stützpunkte und Splineinterpolation

Abbildung 6.14 zeigt den rekonstruierten Fadenverlauf aus Abbildung 6.11. Links sind die einzelnen Punkte des Fadenverlaufs segmentiert, welche die Vorauswahl anhand der ZNCC-Scores passiert haben. Sie wurden für die Darstellung lediglich durch Liniensegmente verbunden. In den meisten Fällen ergibt sich keine äquidistante Verteilung der Stützpunkte, was durch eine entsprechende Anpassung des Parameters t kompensiert werden kann. Der Definitionsbereich der kubischen Spline wird auf das reelle Intervall $[t_{min} = 0.0, t_{max} = 1.0]$ festgelegt. Weil wir ursprünglich davon ausgehen, dass für alle n Punkte des segmentierten Fadenverlaufs genaue 3D-Positionen berechnet werden können, unterteilen wir den Definitionsbereich in $n - 1$ Intervalle und setzen die Schrittweite des Parameters t auf $t_{step} = 1.0/(n - 1)$. Nun wird der Reihe nach für alle segmentierten Liniensegmente anhand des ZNCC-Scores entschieden, ob er als Stützpunkt herangezogen werden kann. Wird ein Punkt disqualifiziert, so findet eine Erhöhung von t_{step} trotzdem statt.

Auf der linken Seite von Abbildung 6.14 ist das Ergebnis der Splineinterpolation zu sehen. Der nächste Abschnitt skizziert die Vorgehensweise zur glatten Fortsetzung von Unterbrechungen im segmentierten Fadenverlauf.

Schließen von Lücken im Kurvenverlauf

Obwohl rekonstruierte Linienpunkte aus bereits erwähnten Gründen nicht immer äquidistant ausfallen können sie, eine lückenlose Segmentierung vorausgesetzt, desselben Fadenverlaufs zugeordnet und entsprechend interpoliert werden. Insbesondere bei sehr dünnem Nahtmaterial kann es jedoch vorkommen, dass bereits die Segmentierung Unterbrechungen aufweist. In solchen Fällen ist es wünschenswert, die Teilstücke trotzdem als eine zusammenhängende interpolierende Raumkurve zu modellieren. Dadurch dass bei einem minimalinvasiven Eingriff niemals zwei chirurgische Fäden gleichzeitig gehandhabt werden, kann man davon ausgehen, dass alle getrennt segmentierten Verläufe zum selben Faden gehören.

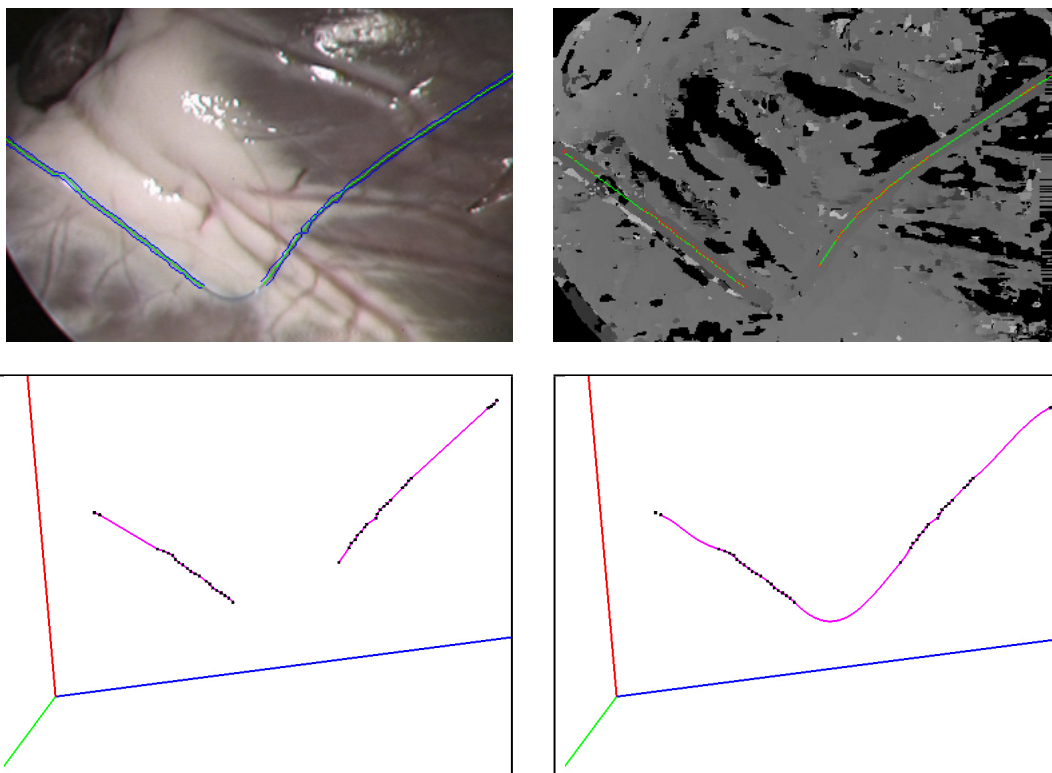


Abbildung 6.15: Fortsetzung des Fadenverlaufs bei einer Unterbrechung

Abbildung 6.15 zeigt Segmentierung, Rekonstruktion und Splineinterpolation eines 4-0er Polypropylenfadens. Während größere Lücken *innerhalb* eines Teilstücks von der interpolierenden Spline auf natürliche Weise überbrückt werden, benötigt die Unterbrechung *zwischen* den Teilstücken besondere Behandlung. Das Prinzip bleibt jedoch dasselbe: es muss eine geeignete Aufteilung des Definitionsbereichs $[t_{min}, t_{max}]$ mit der entsprechenden Zuordnung von Stützpunkten gefunden werden. Zunächst müssen die Randpunkte der Unterbrechung(en) einander zugeordnet werden. Dafür wird die Szene wie gewohnt segmentiert, und die gefundenen Linienverläufe splineinterpoliert. Werden mindestens zwei Fadenstücke gefunden, so erfolgt

eine Berechnung aller euklidischer Entfernungen zwischen Anfangs- und Endpunkten der Raumkurven untereinander, und anschließend aufsteigend sortiert. Bei zwei gefundenen Linienstücken werden die der kleinsten Entfernung entsprechende Randpunkte als zusammengehörend betrachtet. Bei drei Linienstücken sind es nun die ersten zwei (minimalen) Abstände zu betrachten, wobei sich das Prinzip für eine beliebige Anzahl Teilstücke fortsetzen lässt.

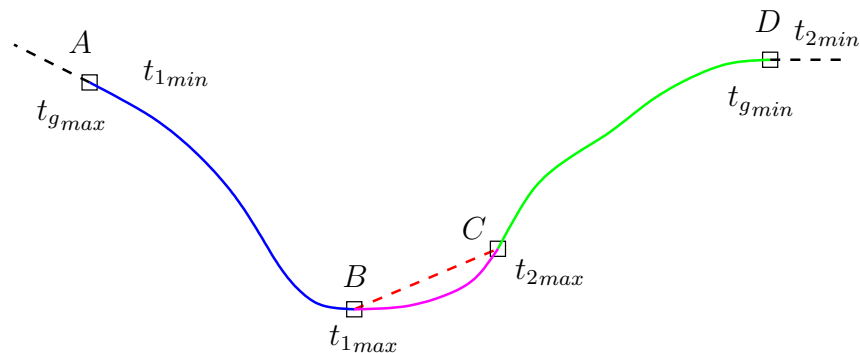


Abbildung 6.16: Lokale und globale Splineinterpolation

Abbildung 6.16 zeigt die Vorgehensweise im Falle zweier Raumkurven. Die Splineinterpolationen liefern zwei Kurven mit den vier Randpunkten A, B, C, D . Anhand von $\min(\min(|AC|, |AD|), \min(|BC|, |BD|))$ wird die Lücke zwischen B und C detektiert. Dabei kann durch ein Schwellwert die maximale Entfernung in kartesischen Weltkoordinaten zwischen zwei beliebigen Kurvenrandpunkten angegeben werden, damit noch ein Zusammenfügen stattfindet. Bei der Festlegung des Definitionsbereichs wird durchgehend t_{min} dem Linienendpunkt mit der niedrigsten Zeilenkoordinate zugeordnet und t_{max} der höchsten. Nach dieser Konvention kann im obigen Fall der Definitionsbereich der global interpolierenden Spline festgelegt werden: $t_{gmin} = t_{2min}$ bzw. $t_{gmax} = t_{1min}$. Für die Aufteilung in einzelne Intervalle ist nun die Gesamtlänge der Kurve zu bestimmen. Sie setzt sich zum einen aus den Einzellängen bereits segmentierter und interpolierter Stücke zusammen ($|\tilde{AB}|, |\tilde{CD}|$), und zum anderen aus der zunächst unbekanntem Länge des fehlenden Stücks. Diese Länge nähern wir als der euklidische Abstand $|BC|$ in Pixelkoordinaten an. Somit ergibt sich für die Schrittweite des Parameters $t_{step} = 1.0 / (|\tilde{AB}| + |\tilde{CD}| + |BC| - 1)$. Bei der Abbildung der Stützpunkte auf den neuen Definitionsbereich muss lediglich darauf geachtet werden, dass beim Übergang von C nach B der Parameter t um den Wert $|BC| \cdot t_{step}$ erhöht wird.

6.3.2 Evaluation der Genauigkeit der Splineinterpolation

Die Genauigkeit der 3D-Rekonstruktion haben wir in einem vorherigen Abschnitt anhand eines Kalibrierkörpers bereits behandelt. Wie akkurat die Splineinterpolation nicht unterbrochener Fadenverläufe ist, lässt sich bedingt durch das Fehlen

entsprechender *ground-truth*-Information nur schwer abschätzen. Die prinzipielle Eignung des Ansatzes für diese Domäne kann jedoch anhand künstlich verursachter Unterbrechungen in ansonsten intakten Verläufen verifiziert werden. Wir nehmen dabei an, dass die Größenordnung der räumlichen Abweichung zwischen realem Fadenverlauf und interpolierender Spline etwa der Abweichung zwischen überbrückter (künstlicher) Lücke und ihrer ursprünglichen Interpolation entspricht.

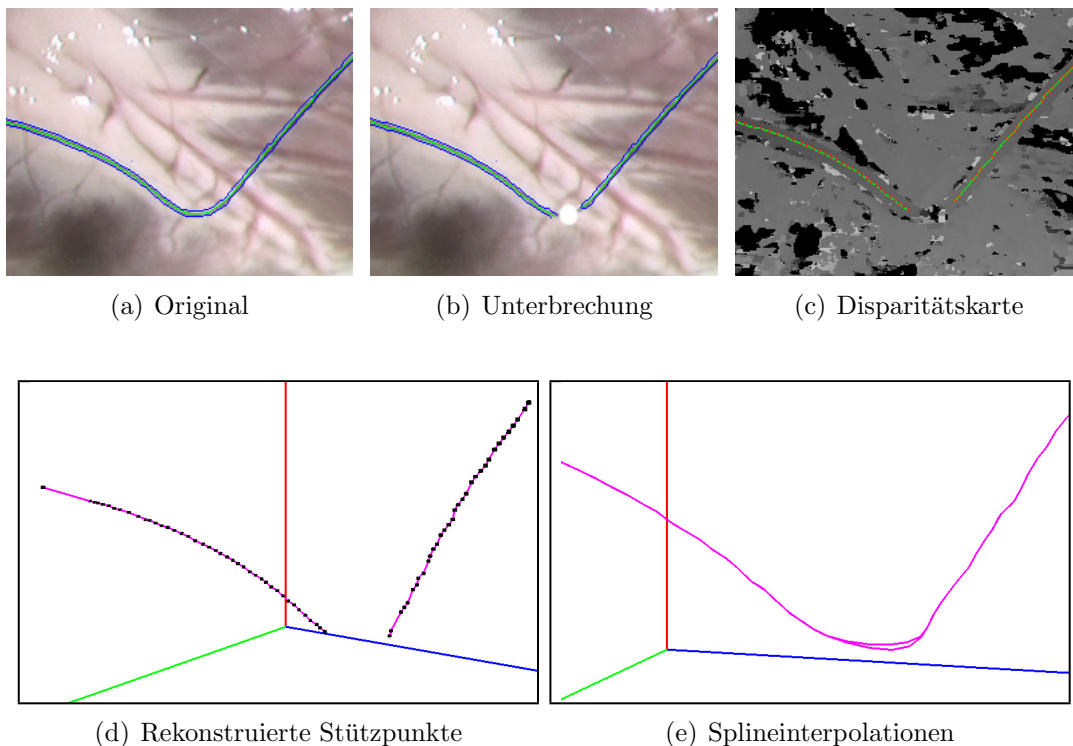


Abbildung 6.17: Schließen einer künstlichen Unterbrechung

Abbildung 6.17 zeigt die Segmentierung, Rekonstruktion und Splineinterpolation einer künstlichen Unterbrechung. Sie wurde in etwa an die Stelle maximaler Krümmung im Fadenverlauf platziert, weil sie erwartungsgemäß hier die größte Störung der Kurve erzeugt. Die Lücke wurde anschließend mit der soeben beschriebenen Methode geschlossen, und die resultierende Kurve mit dem Original verglichen. Jeweils 80 äquidistante Punkte wurden aus den Parameterdarstellungen beider Kurven generiert, und der euklidische Abstand in Weltkoordinaten zwischen paarweise entsprechenden Positionen berechnet. Bei 96% aller Punktpaare lag die Abweichung unter einem Millimeter, bei einem Mindestabstand von 0 mm. Die maximale und durchschnittliche Abweichung betrug 1.777 mm bzw. 0.158 mm. Wie in Abbildung 6.17(e) zu beobachten, beschränken sich die größeren relativen Abweichungen auf die unmittelbare Umgebung der Unterbrechung. Als wichtigstes Ergebnis halten wir fest, dass der Ansatz die Auswirkung lokaler Störungen auf die gesamte Kurve verhindert.

6.4 Modellierung der Herzoberfläche

Die Rekonstruktion und Modellierung der Operationsumgebung in unmittelbarer Nähe des Nahtmaterials (und damit des Arbeitsbereichs) wird durch Sicherheitsüberlegungen motiviert. Während der Ausführung (teil)autonomer Tätigkeiten muss gewährleistet sein, dass keine unbeabsichtigten Kollisionen zwischen chirurgischen Instrumenten und Organgewebe auftreten. Demzufolge müssen durch maschinelles Lernen generierte Trajektorien vor ihrer Ausführung auf Kollisionsfreiheit überprüft werden. Dabei müssen jedoch nicht die selben hohen Anforderungen an die Rekonstruktionsgenauigkeit erfüllt werden, wie dies beim Faden der Fall war.

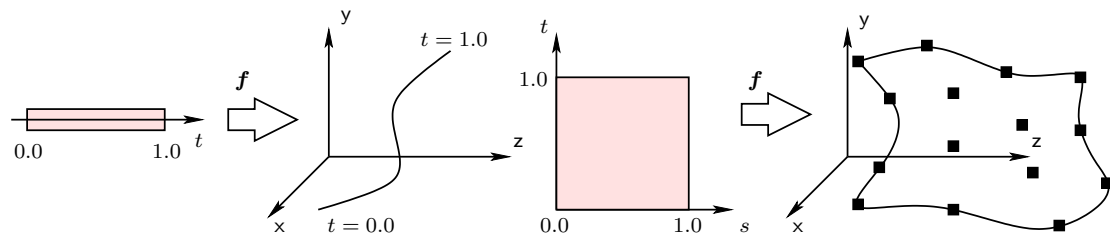


Abbildung 6.18: Abbildung parametrischer Kurven und Flächen in den 3D-Raum

Wie bereits anhand der Polygonaldarstellung aus Abbildung 6.12 angedeutet, ist diese Art der Rekonstruktion für unseren Anwendungsfall nicht geeignet. Bei der Rekonstruktion des Fadenverlaufs haben wir darauf bestanden, dass ihre symbolische Repräsentation segmentierte und rekonstruierte Linienpunkte genau passiert. Das vom Hintergrund scharf hervorgehobene Linienprofil ermöglichte eine sehr genaue Bestimmung der Disparität, und damit die Ermittlung präziser 3D-Koordinaten. Das Auffinden mehrerer, und möglichst gleichverteilter Punkte auf der Herzoberfläche als Kandidaten für eine Rekonstruktion gestaltet sich hingegen schwieriger. Abhängig von der Texturierung gibt es Bereiche mit zahlreichen guten Kandidaten, aber auch größtenteils homogene Flächen, welche die Korrespondenzfindung erschweren. Dementsprechend weitverstreut sind die Tiefenkomponenten der resultierenden 3D-Punktwolke.

6.4.1 Bézier-Flächen

Bei derart stark verrauschten Punktwolken erweist sich ein Interpolationsansatz zunächst als nachteilig, denn wir wollen auf eine aufwendige Nachbearbeitung einer triangulierten Fläche aus Effizienzgründen verzichten. Stattdessen soll eine geeignete mathematische Beschreibungsform gefunden werden, die auf natürliche Weise für eine Annäherung und Glättung der Fläche sorgt. Ähnlich dem Ansatz zur Modellierung des Fadenverlaufs bietet sich insbesondere für das Organgewebe eine parametrisierte Beschreibungsform an. Dabei erweitern wir den eindimensionalen

Parameterbereich für die Darstellung von Flächen um eine zusätzliche Dimension (siehe Abbildung 6.18). Eine einfache und elegante Art anhand von Kontrollpunkten (Stützstellen) eine glatte Fläche zu generieren bieten Bézier-Flächen [129]. Eine rechteckige Bézier-Fläche vom Grad (n, m) wird durch $(n+1)(m+1)$ Kontrollpunkte $\mathbf{b}_{i,j}$ definiert. Dabei wird das Einheitsquadrat auf eine Fläche in einem Raum abgebildet, welcher die gleiche Dimensionalität wie $\mathbf{b}_{i,j}$ hat. Die Beschreibung einer Bézier-Fläche $P(s, t)$ erfolgt mit Hilfe des Produkts skalarer Gewichtsfunktionen (*blending functions*):

$$P(s, t) = \sum_{i=0}^n \sum_{j=0}^m \mathbf{b}_{i,j} \cdot B_i^n(s) \cdot B_j^m(t) \quad (6.11)$$

Dabei sind $\mathbf{b}_{i,j} \in \mathbb{R}^3$ die Bézierpunkte der Fläche, sie spannen ein Kontrollnetz für die Fläche auf. Als Basisfunktionen dienen die sogenannten Bernsteinpolynome:

$$B_i^n(s) = \binom{n}{i} s^i (1-s)^{n-i}, \quad \binom{n}{i} = \frac{n!}{i!(n-i)!} \quad (6.12)$$

Abbildung 6.19 zeigt eine Bézierfläche vom Grad $(7, 7)$ wobei rekonstruierte Punkte von der Herzoberfläche als Kontrollpunkte dienten. Die Kontrollpunkte einer Bézierfläche liegen im allgemeinen nicht auf der Fläche, lediglich die vier Eckpunkte des Kontrollnetzes werden berührt. Für die so entstandene Fläche gilt jedoch, dass sie komplett innerhalb der konvexen Hülle ihrer Kontrollpunkte liegt [130].

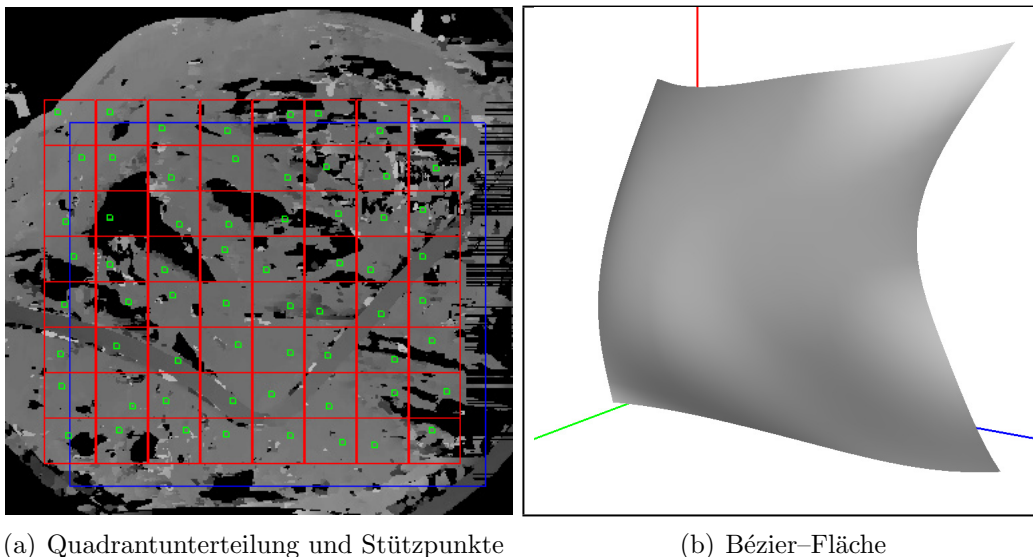


Abbildung 6.19: Modellierung der Herzoberfläche

Die Verteilung der Basisfunktionen im Parameterraum (s, t) , und damit der Einfluss der Kontrollpunkte auf die Fläche, wird durch sogenannte Knotenvektoren

geregelt. Diese sind bei der obigen Fläche identisch für s und t und lauten jeweils $[0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1]$. Eine wichtige geometrische Eigenschaft von Bézier-Flächen ist, dass die Transformation der Fläche auf die Transformation ihrer Kontrollpunkte reduziert werden kann. Für ihre Darstellung bieten 3D-Toolkits wie *OpenGL* und *Open Inventor* native Unterstützung, eine entsprechende Texturierung wird von moderner Grafikhardware ebenfalls effizient erledigt. Die zwei erwähnten Bibliotheken ermöglichen jedoch eine Darstellung von Freiformflächen nur bis zu einem maximalen Grad von sieben [131], wodurch zunächst nur eine Fläche mit maximal 8×8 Kontrollpunkten generiert werden kann. Sollte eine höhere räumliche Auflösung anhand weiterer Kontrollpunkte notwendig sein, so gibt es zwei Möglichkeiten. Zum einen können mehrere Flächen mit hohem Grad durch eine feinere Unterteilung in Quadranten berechnet werden. Zum anderen kann eine Fläche niedrigen Grades mit Hilfe verschiedener Approximationsmethoden anhand einer hohen Zahl von Kontrollpunkten generiert werden.

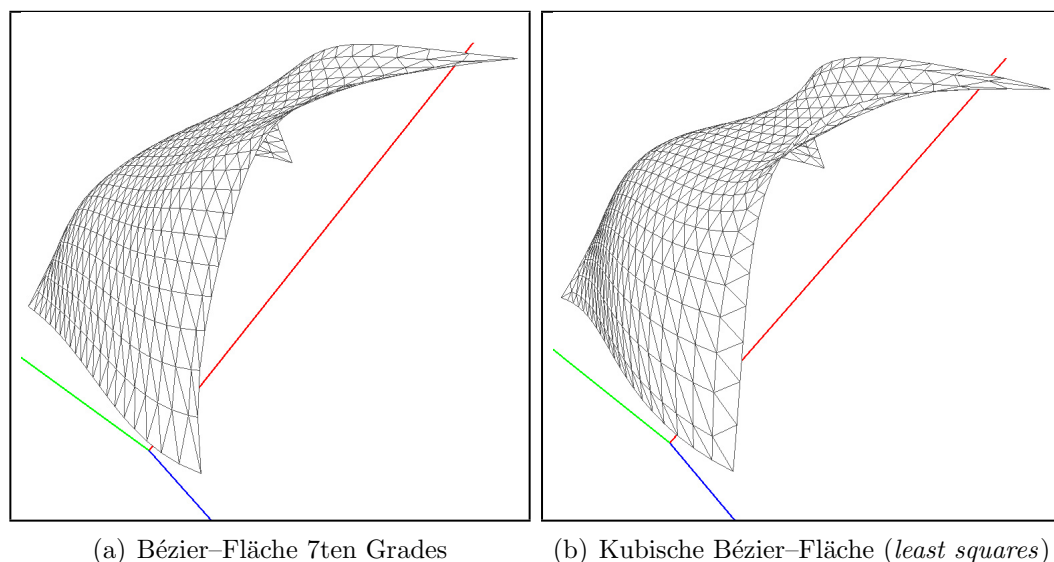


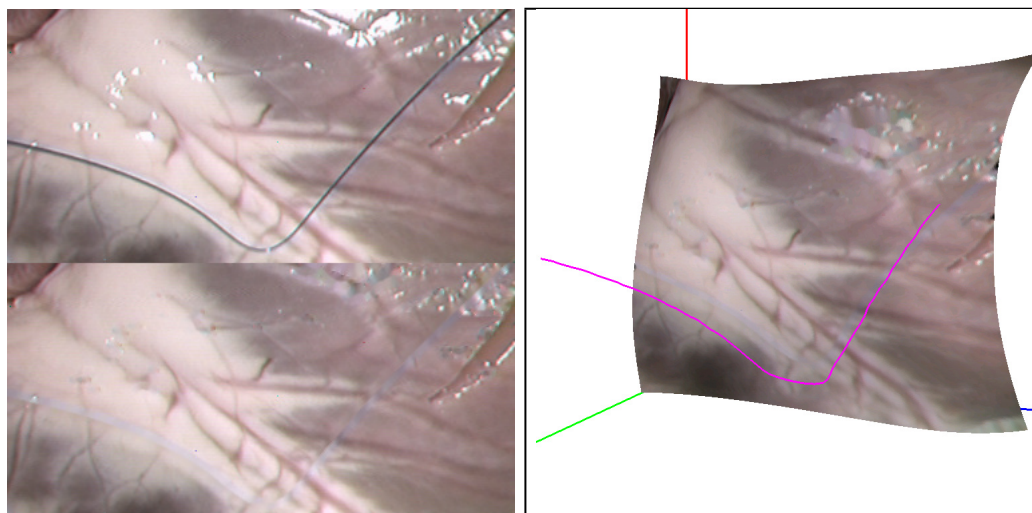
Abbildung 6.20: Approximation mittels kleinster Fehlerquadrate und native Bézier-Fläche vom Grad 7

Die zweite Methode empfiehlt es sich besonders dann, wenn es sich um eine Menge von Punkten handelt, die durch fehlerbehaftete Messungen ermittelt wurden. Eine genaue Interpolation würde dementsprechend auch die Fehler darstellen, eine ausgleichende Fläche hingegen entspräche eher dem gewünschten Ergebnis. Ein sehr verbreiteter Ansatz zum Ausgleich von Messwerten ist die Methode der kleinsten Fehlerquadrate (engl. *least square fitting*, *LSF*). Dabei wird der globale Fehler zwischen Messwerten und Ausgleichsfunktion minimiert. Abbildung 6.20(a) zeigt die triangulierte Form der Bézier-Fläche aus Abbildung 6.19(b). Sie wurde anhand von 64 Kontrollpunkten generiert, und nähert auf natürliche Weise die zu modellierende Herzoberfläche an. Die kubische Bézier-Fläche aus Abbildung 6.20(b) hingegen

ist das Ergebnis einer globalen Approximation mittels der kleinsten Fehlerquadrate, ebenfalls anhand von 64 Kontrollpunkten. Die Unterschiede in der feinaufgelösten triangulierten Darstellungen sind marginal, sodass in den meisten Fällen der gestiegene Berechnungsaufwand mittels *LSF* nicht gerechtfertigt ist.

6.4.2 Texturierung

Die ursprüngliche Zielsetzung bei der Modellierung der Herzoberfläche war die Erkennung von Kollisionen zwischen Instrumenten und Gewebe während autonomer Ausführungen. Hierfür ist die Geometrie in der triangulierten Form ausreichend, die mittels Tesselation aus der mathematischen Beschreibung gewonnen wird. Somit ist das Rendern der Fläche auf der Ebene der Systemsteuerung zunächst nicht notwendig. Eine visuelle Darstellung der vom Nahtmaterial und der Operationsumgebung erstellten Modelle kann jedoch in manchen Fällen angebracht sein. Zum einen können sie in die Simulationsumgebung der Robotersteuerung integriert werden, um beispielsweise autonome Tätigkeiten in einer sicheren Umgebung zu verifizieren. Zum anderen ermöglichen solche Modelle die Integration von Augmentierungstechniken. Eine korrekte Texturierung ist dabei notwendig, um anatomische Strukturen darzustellen, und dadurch die Orientierung des Operateurs zu erleichtern.



(a) Entfernung der Glanzlichter/Fäden

(b) Texturierte Bézier-Fläche

Abbildung 6.21: Photorealistisches Modell des Operationsgebietes

Die Texturierung von Flächen, und von Freiformflächen insbesondere, gestaltet sich mit Hilfe von 3D-Toolkits wie *OpenGL* und *Open Inventor* relativ einfach. Beachtet werden muss lediglich, dass die Größe des Bildausschnitts richtig ermittelt wird, damit keine Stauchung oder Streckung der Textur auftritt. Der rechteckige Bereich der Disparitätskarte (Abb. 6.19(a)), anhand dessen die Flächenrekonstruktion erfolgt,

definiert ebenfalls den Bildausschnitt für die Textur. Besonders störend ist dabei der Faden, dessen Anwesenheit auf der Textur einen falschen Tiefeneindruck vermittelt. Auch die Glanzlichter auf der Herzoberfläche sind hinderlich, zumal sie unabhängig vom Betrachtungswinkel und Beleuchtungsmodell der gerenderten Szene ein starres Erscheinungsbild aufweisen.

Mit Hilfe des im Abschnitt 4.1.6 bereits vorgestellten *Collapsing Chains* Verfahrens können sowohl die Spiegelreflexionen, als auch der Faden aus der Textur ohne nennenswerten Performanzverlust entfernt werden. Anhand des segmentierten Fadenverlaufs wurde eine Maske generiert, die zunächst in der Disparitätskarte (Abb. 6.11(b)) zur Identifikation des Hintergrunds diente. Sie kann auch bei der Entfernung des Fadens aus der Textur verwendet werden, indem der Faden als eine ausgedehnte Spiegelreflexion betrachtet wird. Abbildung 6.21(a) zeigt einen Ausschnitt vom Operationsgebiet jeweils vor und nach der Aufbereitung. Sowohl Nahtmaterial als auch Glanzlichter wurden entfernt, die ursprüngliche Oberflächenstruktur blieb dabei größtenteils erhalten. Die rekonstruierte Szene kann nun photorealistisch dargestellt, augmentiert und in verschiedene Koordinatensysteme transformiert werden.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Zusammenfassung

Die vorliegende Dissertation stellt die Ergebnisse der Arbeiten im Rahmen des Teilprojekts “Teilautonomie und multimodale Instruktion zum Transfer von Fertigkeiten” des DFG-geförderten [8] Sonderforschungsbereichs (SFB) 453 “Wirklichkeitsnahe Telepräsenz und Teleaktion” [9] vor. Das Projekt hatte zwei Hauptziele, die in Zusammenarbeit mit dem Projektpartner von der Klinik für Herz- und Gefäßchirurgie am Deutschen Herzzentrum München [10] verfolgt wurden. Zum einen sollte evaluiert werden, inwieweit die Integration von Kraftsensorik in robotergestützte minimalinvasive Chirurgieszenarien den Operateur entlastet und die Qualität der Eingriffe erhöht. Zum anderen war die Erprobung und Integration von teilautonomen Fähigkeiten mittels Techniken des maschinellen Lernens Gegenstand der Forschung. Es entstanden zwei realitätsnahe Experimentalaufbauten für robotergestützte minimalinvasive Eingriffe, wobei zunächst Kraftsensorik und danach auch teilautonome Abläufe basierend auf maschinelle Bildverarbeitung integriert wurden. Das aktuelle System befindet sich zum Zeitpunkt der Erstellung dieser Arbeit am Deutschen Herzzentrum München, und wird dort im Großtierversuch unter klinischen Bedingungen von Chirurgen evaluiert.

Als exemplarische automatisierbare Aufgabe wurde das Knotenlegen ausgewählt, da laparoskopische Instrumentalknoten wegen ihrer Komplexität einen erhöhten Zeitbedarf haben, und somit ein erhebliches Zeitersparnispotenzial aufweisen. Dabei kann der Wickelvorgang und das Komplettieren des Knotens der Maschine überlassen werden. Zur sicheren und robusten Ausführung ist jedoch eine symbolische Repräsentation des Operationssitus notwendig. Diese kann nur mit Hilfe eines Stereoadsensors mit ausreichender Genauigkeit erstellt werden. Der Schwerpunkt dieser Arbeit liegt somit auf Stereobildverarbeitung im Kontext eines mehrarmigen Robotersystems. Das Ziel war die Entwicklung eines Subsystems für Bildverarbeitung, mit dessen Hilfe die Steuerungsschleife Aktorik \leftrightarrow Sensorik geschlossen werden kann.

Bei der Verarbeitung medizinischer Endoskopieaufnahmen ergeben sich einige spezielle Schwierigkeiten. Durch die helle, stark fokussierte Lichtquelle entstehen auf den meist nassen Oberflächen der inneren Organe störende Spiegelreflexionen. Außerdem neigen Endoskopkameras wegen der konstruktionsbedingten hohen Packungsdichte der CCD-Sensoren häufig zu Bildrauschen. Ein neues Verfahren zur effizienten Entfernung von Glanzlichtern wurde hier vorgestellt und mit existierenden Verfahren verglichen. Die neue Methode ist aufwendigen Diffusionsverfahren qualitativ ebenbürtig, bei erheblich besserem Laufzeitverhalten. Die Problematik des Bildrauschens wurde mittels einer verteilten mehrprozessorfähigen Implementierung eines Diffusionsfilters gelöst.

Für das gewählte Szenario des teilautonomen Knotenlegens ist insbesondere die Kenntnis der Lage des Nahtmaterials notwendig. Um die genaue Geometrie anhand korrespondierender Bilder des Stereoendoskops zu rekonstruieren, muss der Verlauf zunächst in einem der Aufnahmen segmentiert werden. Die Robustheit der Fadenerkennung wurde in realistischen endoskopischen Aufnahmen anhand von chirurgischen Fäden aus verschiedenen Materialien und unterschiedlichen Stärken gezeigt. Desweiteren wurde die Genauigkeit des Verfahrens anhand von Herstellerangaben zum Nahtmaterial ermittelt.

Die anschließende 3D-Rekonstruktion des Fadenverlaufs wurde mit Hilfe der Lösung des Korrespondenzproblems in überlappenden Bildbereichen (Disparitätsberechnung) und anhand der bekannten Geometrie des Kamerasystems (Stereo-Triangulation) durchgeführt. Die symbolische Repräsentation der resultierenden Raumkurve erfolgt mit Hilfe einer Interpolationstechnik basierend auf natürliche kubische Splines. Sie ist durch physikalische Materialeigenschaften chirurgischer Fäden motiviert, und gewährleistet glatte Verläufe und Übergänge. Bei der Modellierung der Herzoberfläche kommen Bézier-Freifformflächen zum Einsatz, die aufgrund aufbereiteter Texturen photorealistisch dargestellt werden können. Die Tiefenaufösung des verwendeten medizinischen Endoskops, und damit die Genauigkeit der Stereorekonstruktion, wurde mit Hilfe eines 3D-Kalibrierkörpers ermittelt.

Für das autonome Greifen des Fadens muss die Steuerung über eine symbolische Repräsentation des Verlaufs im Koordinatensystem der chirurgischen Instrumente verfügen. Um das zu erreichen ist die Kalibrierung des Gesamtsystems bestehend aus Stereoendoskop und Roboterarme notwendig. Ein neuartiger 3D-Kalibrierkörper wurde vorgestellt, der sowohl die sogenannte "Hand-Auge"-Kalibrierung auf einfache Weise ermöglicht, als auch zur Überprüfung der Rekonstruktionsgenauigkeit verwendet werden kann. Unter diesen Voraussetzungen ist es gelungen, einfache autonome Greifvorgänge ausschließlich über die Bildverarbeitung durchzuführen.

7.2 Ausblick

Bei der Vorverarbeitung medizinischer Endoskopieaufnahmen wurde bereits eine verteilte mehrprozessorfähige Implementierung eines Diffusionsfilters verwendet. Mit der zunehmenden Verfügbarkeit von Multiprozessor/Multicore-Architekturen sollten insbesondere laufzeitintensive Algorithmen aus der Bildverarbeitung auf Parallelisierbarkeit untersucht werden. So könnte beispielsweise das hier neu vorgestellte Verfahren zur Entfernung von Glanzlichtern effizient parallelisiert werden. Durch die hohe Datenlokalität bei der Behandlung einzelner Spiegelreflexionen ist die Aufteilung auf einzelne Prozessorkerne zweckmäßig.

Das verwendete medizinische Endoskop wies entfernungs- und belichtungsabhängig starke chromatische Abweichungen auf, wodurch die Verwendung der Farbinformation für das Segmentieren des chirurgischen Nahtmaterials nur bedingt von Nutzen war. Mit der Verfügbarkeit von Endoskopen mit besseren Linsensystemen und höher auflösenden CCD-Sensoren sollte dem Farbraum einer Szene größere Bedeutung beigemessen werden.

Die Auflösung der lediglich pixelgenauen Disparitätskarte ist in gewissem Umfang für Ungenauigkeiten bei der Stereorekonstruktion verantwortlich. Zu untersuchen wäre, inwieweit eine mittels Interpolationstechniken erstellte subpixelgenaue Disparitätskarte die Genauigkeit erhöht. Besonderes Augenmerk ist dabei auf die Tiefendiskontinuität in unmittelbarer Nähe des Fadens zu richten. Eine merkmalsbasierte Zuordnung einzelner segmentierter Linienpunkte im linken und rechten Bild könnten ebenfalls subpixelgenaue Disparitätswerte liefern, jedoch bei einem höheren Berechnungsaufwand.

Die Modellierung der Herzoberfläche erfolgte mittels einer einzelnen Bézier-Freifformfläche. Um etwaige lokale Unebenheiten besser annähern zu können, ist eine feinere Aufteilung in mehrere Flächenabschnitte möglich. Dabei muss insbesondere das Problem der Diskontinuität an den Übergängen einzelner Abschnitte gelöst werden.

Der verwendete 3D-Kalibrierkörper kann aufgrund seiner Materialeigenschaften im fokussierten Licht des Endoskops nicht automatisch segmentiert werden. Die Markierung einzelner Punkte bleibt somit dem Bediener überlassen. Eine entsprechende Eloxierung an relevanten Stellen würde eine automatische Segmentierung ermöglichen. Andere Techniken zur "Hand-Auge"-Kalibrierung, insbesondere Methoden basierend auf Infrarotmarker, sollten ebenfalls evaluiert werden.

Literaturverzeichnis

- [1] Ralph J. Damiano, Walter J. Ehrman, Christopher T. Ducko, Harold A. Tabaie, Edward R. Stephenson, Charles P. Kingsley, and Charles E. Chambers. Initial United States clinical trial of robotically assisted endoscopic coronary artery bypass grafting. *The Journal of Thoracic and Cardiovascular Surgery*, (119):77–82, 2000.
- [2] L. Wiley Nifong, Victor F. Chu, B. Marcus Bailey, David M. Maziarz, Vincent L. Sorrell, Donald Holbert, and W. Randolph Chitwood. Robotic mitral valve repair: experience with the da Vinci system. *The Annals of Thoracic Surgery*, 75:438–443, 2003.
- [3] Friedrich W. Mohr, Volkmar Falk, Anno Diegeler, Thomas Walther, Jan F. Gummert, Jan Bucerius, Stefan Jacobs, and Rüdiger Autschbach. Computer-enhanced “robotic” cardiac surgery: experience in 148 patients. *The Journal of Thoracic and Cardiovascular Surgery*, 121:842–853, 2001.
- [4] Ashok K. Hemal and Mani Menon. Robot-assisted (da Vinci) urologic surgery: An emerging frontier. *Recent Advances in Endourology*, 6:27–38, 2005.
- [5] D. H. Boehm, H. Reichenspurner, C. Detter, M. Arnold, H. Gulbins, B. Meiser, and B. Reichart. Clinical use of a computer-enhanced surgical robotic system for endoscopic coronary artery bypass grafting on the beating heart. In *Thorac. Cardiovasc. Surg.*, number 48, pages 198–202, 2000.
- [6] Colin G. Knight, Alex Cao, Attila Lorincz, Kelly Gidell, Scott E. Langenburg, and Michael D. Klein. Application of a surgical robot to open microsurgery: The equipment. In *Pediatric Endosurgery & Innovative Techniques*, volume 7, pages 227–232, 2003.
- [7] J. Bodner, H. Wykypiel, G. Wetscher, and T. Schmid. First experiences with the da Vinci(TM) operating robot in thoracic surgery. *European Journal of Cardio-Thoracic Surgery*, 25:844–851, 2004.
- [8] Deutsche Forschungsgemeinschaft. <http://www.dfg.de>.
- [9] Sonderforschungsbereich 453 *Wirklichkeitsnahe Telepräsenz und Teleaktion*. <http://www.lsr.ei.tum.de/sfb453>.
- [10] Deutsches Herzzentrum München. <http://www.dhm.mhn.de>.

- [11] O. Fleig, F. Devernay, J. Scarabin, and P. Jannin. Surface reconstruction of the surgical field from stereoscopic microscope views in neurosurgery. In *International Journal of Computer Assisted Radiology and Surgery (CARS)*, 2001.
- [12] Hai Sun, Hany Farid, Kyle Rick, Alex Hartov, David W. Roberts, and Keith D. Paulsen. Estimating cortical surface motion using stereopsis for brain deformation models. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2003.
- [13] F. Mourgues, F. Devernay, and E. Coste-Maniere. 3D reconstruction of the operating field for image overlay in 3D-endoscopic surgery. In *Proceedings IEEE and ACM International Symposium on Augmented Reality*, pages 191–192, 2001.
- [14] F. Devernay, F. Mourgues, and E. Coste-Maniere. Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery. In *Proceedings of Medical Imaging and Augmented Reality*, 2001.
- [15] Integrated Surgical Systems Inc. <http://www.robodoc.com>.
- [16] U. Wiesel and M. Boerner. First experiences using a surgical robot for total knee replacement. In *Journal of Bone and Joint Surgery*, volume 86-B, 2001.
- [17] Universal Robotic Systems Ortho GmbH. <http://www.urs-ortho.de>.
- [18] M. Naito, S. Kurose, M. Oyama, and Y. Sugioka. Anterior cervical fusion with the Caspar instrumentation system. In *International Orthopaedics*, number 17, pages 73–76, 1993.
- [19] ProSurgics. <http://www.prosurgics.com>.
- [20] Patrick A. Finlay and Paul Morgan. PathFinder image guided robot for neurosurgery. In *Industrial Robot: An International Journal*, volume 30, pages 30–34, 2003.
- [21] T.R.K. Varma, P.R. Eldridge, A. Forster, S. Fox, N. Fletcher, M. Steiger, P. Littlechild, P. Byrne, A. Sinnott, K. Tyler, and S. Flintham. Use of the NeuroMate stereotactic robot in a frameless mode for movement disorder surgery. In *Image-Guided Surgery*, number 80, page 132135, 2003.
- [22] Q. H. Li, L. Zamorano, A. Pandya, R. Perez, J. Gong, and F. Diaz. The application accuracy of the NeuroMate robot – a quantitative comparison with frameless and frame-based surgical localization systems. In *Computer Aided Surgery*, number 7, pages 90–98, 2002.
- [23] S. Aiono, J.M. Gilbert, B. Soin, P.A. Finlay, and A. Gordan. Controlled trial of the introduction of a robotic camera assistant (EndoAssist) for laparoscopic cholecystectomy. In *Surgical Endoscopy*, volume 16, pages 1267–1270, September 2002.
- [24] L.R. Kavoussi, R.G. Moore, J.B. Adams, and A.W. Partin. Comparison of robotic versus human laparoscopic camera control. In *The Journal of urology*, volume 154, pages 2134–2136, 1995.

- [25] Computer Motion Inc. <http://www.computermotion.com>.
- [26] Y. Wang J.M. Sackier. Robotically assisted laparoscopic surgery. From concept to development. In *Surgical Endoscopy*, volume 8, pages 63–66, January 1994.
- [27] Jan Martin Proske, Ibrahim Dagher, and Dominique Franco. Comparative study of human and robotic camera control in laparoscopic biliary and colon surgery. In *Journal of Laparoendoscopic & Advanced Surgical Techniques*, volume 14, pages 345–348, 2004.
- [28] L. Mettler, M. Ibrahim, and W. Jonat. One year of experience working with the aid of a robotic assistant (the voice-controlled optic holder AESOP) in gynaecological endoscopic surgery. In *Human Reproduction*, volume 13, pages 2748–2750, 1998.
- [29] Andrew A. Wagner, Ioannis M. Varkarakis, Richard E. Link, Wendy Sullivan, and Li-Ming Su. Comparison of surgical performance during laparoscopic radical prostatectomy of two robotic camera holders, EndoAssist and AESOP: A pilot study. In *Adult Urology*, volume 68, pages 70–74, 2006.
- [30] P.B. Nebot, Y. Jain, K. Haylett, R. Stone, and R. McCloy. Comparison of task performance of the camera-holder robots EndoAssist and Aesop. In *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*, volume 13, pages 334–338, 2003.
- [31] Sirona. <http://www.cereconline.com>.
- [32] D. Schauer, A. Hein, and T. C. Lueth. Dynamic force control for a miniaturised medical robot system. In *Advanced Intelligent Mechatronics, 2003. AIM 2003. Proceedings. 2003 IEEE/ASME International Conference on*, volume 2, pages 1090–1095, 2003.
- [33] Dirk Engel, Jörg Raczkowski, and Heinz Wörn. A safe robot system for craniofacial surgery. In *IEEE Transactions on Robotics and Automation (ICRA)*, pages 2020–2024, 2001.
- [34] Dirk Engel, Jörg Raczkowski, and Heinz Wörn. RobaCKa: Ein Robotersystem für den Einsatz in der Chirurgie. In *Rechner- und sensorgestützte Chirurgie, Proceedings zum Workshop*, pages 279–286. GI, 2001.
- [35] Intuitive Surgical Inc. <http://www.intuitivesurgical.com>.
- [36] The KISMET (Kinematic Simulation Monitoring and Off-Line Programming Environment for Telerobotics) 3D-Simulation Software. <http://www-kismet.iai.fzk.de>.
- [37] U. Voges, E. Holler, B. Neisius, M. Schurr, and T. Vollmer. Evaluation of ARTEMIS, the Advanced Robotics and Telem manipulator System for Minimally Invasive Surgery. In *Proceedings IARP 2nd Workshop on Medical Robotics, Forschungszentrum Karlsruhe*, pages 137–148, 1997.
- [38] Hermann Rininsland. ARTEMIS. A telem manipulator for cardiac surgery. In *European Journal of Cardio-Thoracic Surgery*, pages 106–111, 1999.

- [39] Gerhard F. Buess, Marc O. Schurr, and Sabine C. Fischer. Robotics and allied technologies in endoscopic surgery. In *ARCHSURG*, volume 135, pages 229–235, February 2000.
- [40] SensAble Technologies Inc. <http://www.sensable.com>.
- [41] Murat Cenk Cavusoglu. *Telesurgery and Surgical Simulation: Design, Modeling, and Evaluation of Haptic Interfaces to Real and Virtual Surgical Environments*. PhD thesis, University of California, Berkeley, August 23 2000.
- [42] M. C. Cavusoglu, I. Villanueva, and F. Tendick. Workspace analysis of robotic manipulators for a teleoperated suturing task. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2001)*, October 29–November 3 2001.
- [43] M. C. Cavusoglu, F. Tendick, M. Cohn, and S. S. Sastry. A laparoscopic telesurgical workstation. In *IEEE Transactions on Robotics and Automation (ICRA)*, volume 15, pages 728–739, August 1999.
- [44] M. C. Cavusoglu, W. Williams, F. Tendick, and S.S. Sastry. Robotics for telesurgery: Second generation Berkeley/UCSF laparoscopic telesurgical workstation and looking towards the future applications. In *Proceedings of the 39th Allerton Conference on Communication, Control and Computing*, October 3–5 2001.
- [45] Dong-Soo Kwon, Ki Young Woo, and Hyung Suck Cho. Haptic control of the master hand controller for a microsurgical telerobot system. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1722–1727, May 1999.
- [46] Dong-Soo Kwon, Ki Young Woo, Wan Soo Kim Se-Kyong Song, and Hyung Suck Cho. Microsurgical telerobot system. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 945–950, October 1998.
- [47] Dong-Soo Kwon and Se-Kyong Song. A microsurgical telerobot system with a 6-dof haptic master device. In *Proceedings of the 2000 International Symposium on Mechatronics and Intelligent Mechanical System for 21. Century*, pages 65–71, 2000.
- [48] Mamoru Mitsuishi, Yasuhiro Iizuka, Hiroyoshi Watanabe, Hiroyuki Hashizume, and Kazuo Fujiwara. Remote operation of a micro-surgical system. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1013–1019, May 1998.
- [49] Mamoru Mitsuishi, Shin Tomisaki, Takumi Yoshidome, Hiroyuki Hashizume, and Kazuo Fujiwara. Tele-micro-surgery system with intelligent user interface. In *IEEE Transactions on Robotics and Automation (ICRA)*, pages 1607–1614, April 2000.
- [50] Jumpei Arata, Mamoru Mitsuishi, Shin'ichi Warisawa, Katsuya Tanaka, Takasi Yoshizawa, and Makoto Hashizume. Development of a dexterous minimally-invasive surgical system with augmented force feedback capability. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3207– 3212, August 2005.

- [51] Mamoru Mitsuishi, Jumpei Arata, Katsuya Tanaka, Manabu Miyamoto, Takumi Yoshidome, Satoru Iwata, Shin'ichi Warisawa, and Makoto Hashizume. Development of a remote minimally-invasive surgical system with operational environment transmission capability. In *IEEE Transactions on Robotics and Automation (ICRA)*, pages 2663–2670, September 2003.
- [52] P. Berkelman, P. Cinquin, J. Troccaz, JM. Ayoubi, C. Letoublon, and F. Bouchard. A compact compliant laparoscopic endoscope manipulator. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 1870–1875, May 2002.
- [53] P. Berkelman, P. Cinquin, J. Troccaz, JM. Ayoubi, and C. Letoublon. Development of a compact cable-driven laparoscopic endoscope manipulator. In *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 17–24, September 2002.
- [54] P. Berkelman, E. Boidard, P. Cinquin, and J. Troccaz. Control and user interface design for compact manipulators in minimally-invasive surgery. In *Proceedings of IEEE Conference on Control Applications*, pages 25–30, August 2005.
- [55] Ji Ma and Peter Berkelman. Control software design of a compact laparoscopic surgical robot system. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2345–2350, October 2006.
- [56] Koji Ikuta, Keiji Sasaki, Keiichi Yamamoto, and Takayuki Shimada. Remote microsurgery system for deep and narrow space-development of new surgical procedure and micro-robotic tool. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 163–172, September 2002.
- [57] Keiichi Yamamoto Koji Ikuta and Keiji Sasaki. Development of remote microsurgery robot and new surgical procedure for deep and narrow space. In *IEEE Transactions on Robotics and Automation (ICRA)*, pages 1103–1108, September 2003.
- [58] Koji Ikuta, Takahiko Hasegawa, and Shinichi Daifu. Hyper redundant miniature manipulator "hyper finger" for remote minimally invasive surgery in deep area. In *IEEE Transactions on Robotics and Automation (ICRA)*, September 2003.
- [59] Koji Ikuta, Shinichi Daifu, Takahiko Hasegawa, and Humihiro Higashikawa. Hyper-finger for remote minimally invasive surgery in deep area. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 2488, pages 173–181, 2002.
- [60] I. Nagy, H. Mayer, and A. Knoll. Endo[PA]R - an overview of an experimental system for robotic surgery. Technical report, Technische Universität München, Germany, December 2003.
- [61] I. Nagy, H. Mayer, and A. Knoll. Application of skill transfer in robotic surgery. Technical report, Technische Universität München, Germany, December 2003.

- [62] H. Mayer, I. Nagy, and A. Knoll. Skill transfer and learning by demonstration in a realistic scenario of laparoscopic surgery. In *Proceedings of the IEEE International Conference on Humanoids [CD-ROM]*, Munich, Germany, October 2003.
- [63] I. Nagy, H. Mayer, and A. Knoll. Application of force feedback in robot assisted minimally invasive surgery. In *Proceedings of the 4th International Conference Eurohaptics*, pages 240–245, Munich, Germany, June 2004.
- [64] H. Mayer, I. Nagy, and A. Knoll. Inverse kinematics of a manipulator for minimally invasive surgery. Technical report, Technische Universität München, Germany, February 2004.
- [65] H. Mayer, I. Nagy, and A. Knoll. Kinematics and modeling of a system for robotic surgery. In J. Lenarcic and C. Galletti, editors, *On Advances in Robot Kinematics (Conf. Proceedings)*, pages 181–190, Sestri Levante, Italy, June 2004. Kluwer Academic Publishers.
- [66] H. Mayer, I. Nagy, A. Knoll, E.U. Schirmbeck, and R. Bauernschmitt. Integration of force feedback in an open robot platform for robotic surgery. In *Proceedings of the 10th Mediterranean Conference on Medical and Biological Engineering [CD-ROM]*, Ischia, Italy, July 2004.
- [67] Boris Vakić. Design und Entwicklung einer intelligenten, hotplug-fähigen Messverstärkerschaltung zur Greifkraftbestimmung chirurgischer Instrumente. Master's thesis, Technische Universität München, Betreuer István Nagy, 2006.
- [68] H. Mayer, I. Nagy, A. Knoll, E.U. Schirmbeck, and R. Bauernschmitt. Robotic system to evaluate force feedback in minimally invasive computer aided surgery. In *Proceedings of the ASME International Design Engineering Technical Conferences [CD-ROM]*, Salt Lake City, Utah, September 2004.
- [69] E.U. Schirmbeck, I. Nagy, H. Mayer, A. Knoll, R. Lange, and R. Bauernschmitt. Evaluation of force feedback in minimally invasive robotic surgery. *Biomedical Engineering*, 49(2):108–109, 2004.
- [70] R. Bauernschmitt, E.U. Schirmbeck, C. Haßelbeck, F. Freyberger, H. Mayer, A. Knoll, S. Wildhirt, and R. Lange. Evaluation of haptic feedback in an experimental telemanipulator system. In *Proceedings of the IEEE 3rd European Medical and Biological Engineering Conference [CD-ROM]*, Prague, Czech Republic, November 2005.
- [71] I. Nagy, H. Mayer, A. Knoll, E.U. Schirmbeck, and R. Bauernschmitt. Endo[PA]R: An open evaluation system for minimally invasive robotic surgery. In *Proceedings of the IEEE Conference "Mechatronics & Robotics"*, pages 1464–1467, Aachen, Germany, September 2004.
- [72] H. Mayer, I. Nagy, A. Knoll, E.U. Schirmbeck, and R. Bauernschmitt. The Endo[PA]R system for minimally invasive robotic surgery. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3637–3642, Sendai, Japan, September 2004.

- [73] R. Bauernschmitt, A. Knoll, H. Mayer, I. Nagy, E.U. Schirmbeck, S. Wildhirt, and R. Lange. Towards robotic heart surgery: Introduction of autonomous procedures into an experimental surgical telemanipulator system. In *Proceedings of the 3rd Annual Meeting of the German Society for Computer and Robot Assisted Surgery, CURAC [CD-ROM]*, Munich, Germany, October 2004.
- [74] H. Mayer, F. Gomez, D. Wierstra, I. Nagy, A. Knoll, and J. Schmidhuber. A system for robotic heart surgery that learns to tie knots using recurrent neural networks. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 543–548, Beijing, China, October 2006.
- [75] H. Mayer, I. Nagy, A. Knoll, E.U. Braun, R. Lange, and R. Bauernschmitt. Adaptive control for human-robot skilltransfer: Trajectory planning based on fluid dynamics. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1800–1807, Rome, Italy, April 2007.
- [76] Michael Stegherr. Design und Implementierung einer dreidimensionalen haptischen Benutzerschnittstelle. Anwendung in robotergestützten endoskopischen Chirurgieszenarien. Master's thesis, Technische Universität München, Betreuer István Nagy, 2006.
- [77] Stefan Hueber. Erweiterung einer Simulationsumgebung zur Robotersteuerung. Master's thesis, Technische Universität München, 2005.
- [78] H. Mayer, I. Nagy, A. Knoll, E.U. Schirmbeck, and R. Bauernschmitt. Upgrading instruments for robotic surgery. In *Proceedings of the Australasian Conference on Robotics and Automation [CD-ROM]*, Canberra, Australia, December 2004.
- [79] Boris Vakšić. Realzeitsteuerung von Servomotoren für den Antrieb minimalinvasiver Instrumente, Technische Universität München, Systementwicklungsprojekt, Betreuer István Nagy, 2005.
- [80] H. Mayer, I. Nagy, A. Knoll, E.U. Braun, and R. Bauernschmitt. Human-computer interfaces of a system for robotic heart surgery. In *Proceedings of the Second IASTED International Conference on Human-Computer Interaction*, pages 31–36, Chamonix, France, March 2007.
- [81] Urs Müller-Richter, Andreas Limberger, Peter Weber, Klaus Wilhelm Ruprecht, Wolfgang Spitzer, and Martin Schilling. Possibilities and limitations of current stereo-endoscopy. *Surgical Endoscopy*, 18(6):942–947, 2004.
- [82] Urs Müller-Richter, Andreas Limberger, Peter Weber, and Martin Schilling. Vergleichende Untersuchung räumlicher Darstellungsverfahren in der Stereoendoskopie. *Mund-, Kiefer- und Gesichtschirurgie*, 7(3):157–163, 2003.
- [83] Reimar Lenz. Linsenfehlerkorrigierte Eichung von Halbleiterkameras mit Standardobjektiven für hochgenaue 3D-Messungen in Echtzeit. In *Mustererkennung 1987, 9. DAGM-Symposium, Braunschweig*, volume 149 of *Informatik-Fachberichte*, pages 212–216. Springer, 1987.

- [84] Juyang Weng, Paul Cohen, and Marc Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(10):965–980, 1992.
- [85] D. C. Brown. Close-range camera calibration. In *Photogrammetric Engineering*, volume 37, pages 855–866, 1971.
- [86] Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. In *Proceedings of IEEE Journal of Robotics and Automation*, pages 323–344, 1987.
- [87] Craig Kolb, Don Mitchell, and Pat Hanrahan. A realistic camera model for computer graphics. In *Proceedings of the International Conference on Computer Graphics and Interactive Techniques*, pages 317 – 324, 1995.
- [88] Y.I. Abdel-Aziz and H.M. Karara. Direct linear transformation into object space coordinates in close-range photogrammetry. *Proc. Symp. Close-Range Photogrammetry*, pages 1–18, 1971.
- [89] Roger 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.
- [90] Zhengyou Zhang. A flexible new technique for camera calibration. In *Technical Report, Microsoft Corporation, MSR-TR-98-71*, 1998.
- [91] Olivier D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig. In *Proceedings of the Second European Conference on Computer Vision (ECCV)*, pages 563–578, London, UK, 1992. Springer-Verlag.
- [92] Richard Wolf GmbH. <http://www.richard-wolf.com>.
- [93] Schöolly Fiberoptic GmbH. <http://www.schoelly.de>.
- [94] Hailin Jin, Anthony J. Yezzi, and Stefano Soatto. Variational multiframe stereo in the presence of specular reflections. *3dpvt*, 00:626, 2002.
- [95] H. Jin, S. Soatto, and A. Yezzi. Multi-view stereo beyond lambert, 2002.
- [96] G.J. Klinker, S.A. Shafer, and T. Kanade. Using a color reflection model to separate highlights from object color. In *DARPA87*, pages 614–619, 1987.
- [97] G.J. Klinker, S.A. Shafer, and T. Kanade. Image segmentation and reflection analysis through color. In *DARPA88*, pages 838–853, 1988.
- [98] G.J. Klinker, S.A. Shafer, and T. Kanade. A physical approach to color image understanding. *IJCV*, 4(1. January 1990):7–38, January 1990.
- [99] D.N. Bhat and S.K. Nayar. Binocular stereo in the presence of specular reflection. In *ARPA94*, pages II:1305–1315, 1994.

- [100] Dinkar N. Bhat and Shree K. Nayar. Stereo in the presence of specular reflection. In *ICCV*, pages 1086–1092, 1995.
- [101] D. Bhat and S. Nayar. Stereo and specular reflection, 1998.
- [102] Martin Gröger, Wolfgang Sepp, Tobias Ortmaier, and Gerd Hirzinger. Reconstruction of image structure in presence of specular reflections. In *Proceedings of the 23rd DAGM-Symposium on Pattern Recognition*, pages 53–60, London, UK, 2001. Springer-Verlag.
- [103] T. Ortmaier, M. Gröger, D. H. Boehm, F. Falk, and G. Hirzinger. Motion estimation in beating heart surgery. In *IEEE Transactions on biomedical Engineering*, volume 52, 2005.
- [104] T. Ortmaier, M. Gröger, and G. Hirzinger. Multisensorielle Schätzung der Herzbewegung in der minimal invasiven Chirurgie. In *CURAC*, 2002.
- [105] Stephen Lin, Yuanzhen Li, Sing Bing Kang, Xin Tong, and Heung-Yeung Shum. Diffuse-specular separation and depth recovery from image sequences. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part III*, pages 210–224, London, UK, 2002. Springer-Verlag.
- [106] Yuanzhen Li, Stephen Lin, Hanqing Lu, Sing Bing Kang, and Heung-Yeung Shum. Multibaseline stereo in the presence of specular reflections. In *ICPR '02: Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02) Volume 3*, page 30573, Washington, DC, USA, 2002. IEEE Computer Society.
- [107] Ping Tan, Stephen Lin, Long Quan, and Heung-Yeung Shum. Highlight removal by illumination-constrained inpainting. In *ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision*, page 164, Washington, DC, USA, 2003. IEEE Computer Society.
- [108] Tony F. Chan and Jianhong Shen. Mathematical models for local nontexture inpaintings. In *SIAM Journal on Applied Mathematics*, volume 62, pages 1019–1043, 2002.
- [109] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. In *IEEE Transactions on Image Processing*, volume 10, pages 1200–1211, 2001.
- [110] Herbert Freeman. Computer processing of line-drawing images. In *ACM Computing Surveys (CSUR)*, volume 6, pages 57 – 97, 1974.
- [111] Jan J. Koenderink. The structure of images. *Biological Cybernetics*, 50(5):363–370, 1984.
- [112] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
- [113] Joachim Weickert. *Anisotropic Diffusion in Image Filtering*. PhD thesis, 1996.

- [114] D. Tschumperlé and R. Deriche. Constrained and unconstrained PDE's for vector image restoration. *Proceedings of the 10th Scandinavian Conference on Image Analysis*, pages 153–160, 2001.
- [115] D. Tschumperlé and R. Deriche. Diffusion PDE's on vector-valued images. *IEEE Signal Processing Magazine*, 19(5):16–25, 2002.
- [116] Monika Franziska Maria Flury. *Die Entwicklung chirurgischen Nahtmaterials als Voraussetzung und Folge operativer Tätigkeiten und wissenschaftlicher Forschung*. PhD thesis, Julius–Maximilians–Universität Würzburg, Würzburg, 2002.
- [117] L. Nobile, L. Checchi, and G. Monaco. Experimental analysis of tensile properties of some suturing materials. *Journal of Materials Science: Materials in Medicine*, 8(1):53–56, 1997.
- [118] Johnson & Johnson Intl. <http://www.jnj.com>.
- [119] Albert Baumgartner, Carsten Steger, Christian Wiedemann, Helmut Mayer, and Wolfgang Eckstein. Update of Roads in GIS from Aerial Imagery: Verification and Multi-Resolution Extraction. In *International Archives of Photogrammetry and Remote Sensing*, volume XXXI, Part B3, pages 53–58, 1996.
- [120] G. Coppini, M. Demi, R. Poli, and G. Valli. An artificial vision system for X-ray images of human coronary trees. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(2), 1993.
- [121] Carsten Steger. An unbiased detector of curvilinear structures. Technical report, Technische Universität München, Germany, July 1996.
- [122] Donald Geman and Bruno Jedynek. An active testing model for tracking roads in satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):1–14, 1996.
- [123] Thomas Koller, Guido Gerig, Gabor Szekely, and Daniel Dettwiler. Multiscale detection of curvilinear structures in 2D and 3D image data. In *Fifth International Conference on Computer Vision*, pages 864–869, 1995.
- [124] Carsten Steger. An unbiased detector of curvilinear structures from 2D and 3D images. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 20, pages 113–125, 1998.
- [125] In So Kweon and Takeo Kanade. Extracting topographic terrain features from elevation maps. *CVGIP: Image Underst.*, 59(2):171–182, 1994.
- [126] D. Eberly, R. Gardner, B. Morse, S. Pizer, and C. Scharlach. Ridges for image analysis. *J. Math. Imaging Vis.*, 4(4):353–373, 1994.
- [127] Li Wang and Theo Pavlidis. Detection of curved and straight segments from gray scale topography. *CVGIP: Image Underst.*, 58(3):352–365, 1993.

- [128] Tony Lindeberg. Edge detection and ridge detection with automatic scale selection. In *Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition (CVPR '96)*, pages 465–470, Washington, DC, USA, 1996. IEEE Computer Society.
- [129] Hans-Joachim Bungratz, Michael Griebel, and Christoph Zenger. *Einführung in die Computergraphik*. Vieweg, 1996.
- [130] Les Piegel and Wayne Tiller. *The NURBS Book*. Springer, 1997.
- [131] Josie Wernecke. *The Inventor Mentor*. Addison Wesley, 1994.