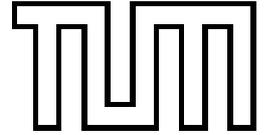


Institut für Informatik
der Technischen Universität München



Intraoperative fluoroskopiebasierte Patientenlageerkennung zur präzisen Unterstützung chirurgischer Eingriffe

Dissertation

Michael Roth

Institut für Informatik
der Technischen Universität München
Lehrstuhl Univ.-Prof. Dr. B. Radig

**Intraoperative fluoroskopiebasierte
Patientenlageerkennung zur präzisen
Unterstützung chirurgischer Eingriffe**

Michael Roth

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. Dr. Bernd Radig

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Achim Schweikard
2. Univ.-Prof. Dr. Gunther Reinhart

Die Dissertation wurde am 31.03.2000 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 16.08.2000 angenommen.

Kurzfassung

Beachtliche Fortschritte auf dem Gebiet der medizinischen 3D-Bildgebung (z.B. Computer-Tomographie) und insbesondere der computergestützten Visualisierung und Modellierung der gewonnenen Datensätze ermöglichen heutzutage in vielen Fällen die hochpräzise Planung eines operativen Eingriffs bereits im Vorfeld der eigentlichen Operation. Im Falle der Orthopädie ist eine solche Planung z.B. beim Einsatz von Hüft- und Kniegelenk-Implantaten oder auch bei einer gezielten Versteifung eines Teils der Wirbelsäule sinnvoll. Diese und ähnliche Operationen erfordern jeweils die exakte Durchführung einer Vielzahl komplizierter Arbeitsschritte zur gezielten Manipulation einzelner Knochen, z.B. dem Setzen von Bohrlöchern, der Entfernung von Knochenstücken durch entsprechende Sägeschnitte u.ä.

Ehe sich die Ergebnisse der präoperativen Planung während der Operation optimal verwenden lassen, ist es jedoch nötig, die Lage des Patienten bzw. der betrachteten Knochenstruktur präzise zu erfassen. Nur dann ist es überhaupt möglich, die bzgl. des präoperativen Datensatzes geplanten Arbeitsschritte auch relativ zum realen Knochen durchzuführen. Diese Lageerkennung wird im medizinischen Umfeld gemeinhin als Registrierung bezeichnet.

Das zentrale Anliegen dieser Arbeit bestand daher in der Entwicklung und Implementierung zweier neuartiger Verfahren zur intraoperativen Registrierung von Knochenstrukturen. Konzeptionelle Randbedingungen ergaben sich durch die Verwendung eines kalibrierten medizinischen Standard-Röntgenbildverstärkers ('C-Bogen', Fluoroskop) als bildgebende Komponente im Operationssaal (OP) sowie durch die Forderung, auf an den Patientenknochen angebrachte künstliche Hilfsmarken, die eine Registrierung erleichtern würden, zu verzichten. Verfahren unter Einsatz solcher Marken sind bekannt; sie erfordern jedoch einen – der eigentlichen Operation vorgeschalteten – zusätzlichen Eingriff.

Das erste der beiden Verfahren simuliert intraoperativ künstliche Röntgenbilder auf Basis eines Computer-Tomographie-Datensatzes des Patienten und vergleicht diese direkt mit realen Fluoroskopieaufnahmen, um so einen präzisen Rückschluß auf die Lage des Patienten zu ermöglichen.

Das zweite Registrierungsverfahren verzichtet hingegen auf den Einsatz einer präoperativen 3D-Bildgebung (Computer-Tomographie). Stattdessen dient ein formvariables 3D-Modell der patientenunabhängigen, anatomischen Modellierung zu erwartender Formausprägungen eines bestimmten Knochens. Ein solches Modell läßt sich intraoperativ an die – in den einzelnen Fluoroskopieaufnahmen segmentierten – Konturen bzw. Silhouetten des betrachteten Knochens anpassen.

Bei der Realisierung beider Verfahren standen neben einer hohen Präzision des Registrierungsergebnisses die Kernanforderungen Praktikabilität im OP-Umfeld, Robustheit und Geschwindigkeit im Vordergrund. Umfangreiche Experimente an Oberschenkelknochen-Phantomen und verschiedenen Wirbeln (Labor-Umgebung) sowie ein klinischer Test unter Verwendung eines pathologischen Kniepräparats (reale OP-Umgebung) runden die Arbeit ab und demonstrieren die Leistungsfähigkeit und Grenzen der entwickelten Verfahren und Ansätze.

Danksagung

Die vorliegende Dissertation entstand während meiner mehrjährigen Tätigkeit am Bayerischen Forschungszentrum für wissenschaftliche Systeme (FORWISS) in München.

Herrn Prof. Dr. Achim Schweikard danke ich für die wohlwollende Betreuung meiner Arbeit und insbesondere dafür, daß er zuletzt erstaunlich viel Geduld bewies und mir über Monate hinweg den dringend nötigen Freiraum verschaffte.

Herr Prof. Dr. Gunther Reinhart, Leiter des Instituts für Werkzeugmaschinen und Betriebswissenschaften (iwb) der TU München, verdient Dank für die bereitwillige Übernahme des Zweitgutachtens.

Meinem Freund und ehemaligen Kollegen, Herrn Christian Brack, möchte ich an dieser Stelle ebenfalls danken. Gemeinsam mit Herrn Hubert Götte, iwb, bewältigten wir in den letzten Jahren so manche Hürde des CARS-Forschungsprojektes (Computer-Aided Robotic Surgery). Christian stand aber auch während der Entstehung dieser Arbeit stets als diskussionsfreudiger Ansprechpartner zur Verfügung.

Herrn Manfred Dötter, ehemaliger Diplomand und zwischenzeitlich neuer Kollege, gebührt besonderer Dank für die herausragenden Leistungen im Rahmen seiner Diplomarbeit. Ohne ihn wären grundlegende Ideen aus Kapitel 6 nicht oder zumindest nicht in dieser Form umgesetzt worden.

Allen weiteren Studenten, die im Rahmen ihres Systementwicklungsprojektes bzw. ihrer Diplomarbeit direkt oder indirekt am Entstehen dieser Arbeit mitgewirkt haben, möchte ich an dieser Stelle ebenfalls meinen Dank aussprechen. Stellvertretend seien Daniel Frey und Florian Lindauer genannt.

Mein Dank gilt in jedem Fall Herrn Dr. Rainer Burgkart vom Münchener Klinikum Rechts der Isar, der durch sein Engagement und seine Begeisterung gegenüber neuen Ansätzen einen ersten vorklinischen Versuch der Verfahren ermöglichte.

Für die ohne Frage beschwerliche orthographische Durchsicht der Arbeit bedanke ich mich ganz herzlich bei meiner Mutter Ingrid. Überhaupt verdient auch meine restliche Familie Dank: Mein Vater Günter für die Vielzahl an Entlastungen sowie meine Schwester Simone und ihr Mann André für die Bereitstellung einer sehr praktischen, nahe gelegenen Übernachtungsmöglichkeit, wenn es mal wieder etwas länger dauerte im Büro.

Zu guter Letzt danke ich meiner Freundin Claudia für ihre Geduld und ihr Verständnis während der zuletzt doch sehr harten, entbehrungsreichen Zeit. Sie hat mir stets den dringend nötigen mentalen Beistand gewährt und am Ende insbesondere auch nicht davor zurückgeschreckt, meine Arbeit zeitaufwendig zu korrigieren. Danke!

Augsburg, 30. März 2000

Michael Roth

Inhaltsverzeichnis

1	Einführung	1
1.1	Thematisches Umfeld der Arbeit	1
1.2	Aufgabenstellung	2
1.2.1	Anforderungen und Zielsetzung	2
1.2.2	Lösungskonzept	4
1.3	Übersicht	5
2	Stand der Forschung	7
2.1	Orthopädische Anwendungsszenarien	7
2.2	Bekannte computerunterstützte Chirurgesysteme in der Orthopädie	9
2.2.1	Künstlicher Hüftgelenkrsatz	9
2.2.2	Künstlicher Kniegelenkrsatz	10
2.2.3	Wirbelsäulenversteifung	10
2.3	Diskussion bekannter Ansätze zur Registrierung	11
2.3.1	Starre Registrierung	11
2.3.2	Elastische Modellierung und Registrierung	13
2.3.3	Validierung – Verfahrensbewertung	15
2.3.4	Fazit	16
3	Systemüberblick	19
3.1	Das Röntgenbildverstärker-System (Fluoroskop)	19
3.2	Computer-Tomographie	22
3.3	Echtzeit-Tracking-System	23
4	Grundlagen der Registrierung	25
4.1	Problemstellung	25
4.2	Initialisierung	29
4.2.1	Vorgabe einzelner Aufnahmerichtungen	30
4.2.2	Verwendung anatomischer Details	33
4.2.3	Manuelle Modell-Plazierung	37
5	Intensitätsbasierte Registrierung	41
5.1	Zielsetzung und Ansatz	42
5.1.1	Motivation	42

5.1.2	Ansatz	44
5.2	CT-basierte Simulation von Fluoroskopie-Aufnahmen	45
5.2.1	Ray-Casting durch CT-Volumen	45
5.2.2	Helligkeits- und Kontrast-Adaption	48
5.2.3	Bildmaskierung – Berücksichtigung von Artefakten	49
5.2.3.1	Artefakt-Typen	49
5.2.3.2	Detektion von Artefakt-Bildregionen	50
5.2.3.3	Filterung/Interpolation detektierter Artefakte	52
5.3	Lokalisation	57
5.3.1	Initialisierung	57
5.3.2	Optimierung	57
5.4	Beschleunigte Simulation von Röntgenaufnahmen	60
5.4.1	Ansatz	61
5.4.2	Ansichtsmodelle	61
5.4.2.1	Erzeugung von Trainingsbildern	62
5.4.2.2	Eigenraum-Repräsentation	63
5.4.2.3	Interpolation im Eigenraum	69
5.4.2.4	Planare Approximation von 3D-Transformationen	70
5.4.2.5	Approximationsfehler-Analyse	72
5.4.3	Optimierung	74
5.5	Ergebnisse	75
5.5.1	Simulationsergebnisse	75
5.5.2	Anwendung auf reale Fluoroskopieaufnahmen	79
5.5.3	Zeitaufwand der Optimierung	85
6	Elastische Registrierung	89
6.1	Zielsetzung und Ansatz	90
6.1.1	Motivation	90
6.1.2	Ansatz	91
6.2	Modell-Repräsentation	94
6.3	Physikalisch motivierte Modellierung (<i>FEM</i>)	95
6.4	Statistische Modellierung durch Training	97
6.4.1	Ansatz	97
6.4.2	Berechnung der statistischen Kenngrößen	98
6.4.3	Erzeugung von Trainings-Datensätzen	100
6.4.4	Initiale Bestimmung weniger Basis-Korrespondenzen	101
6.4.4.1	Auswahl markanter und stabiler Merkmale	101
6.4.4.2	Automatische Korrespondenzfindung (Labelling)	105
6.4.5	Ausrichtung der Modelle	107
6.4.5.1	Rigide Ausrichtung	109
6.4.5.2	Elastisches Alignment (Abschätzung einer Deformation)	110
6.4.6	Vollständige Bestimmung von Korrespondenzen	113
6.5	Konturbasierte, elastische Lokalisation	115
6.5.1	Segmentation von Fluoroskopieaufnahmen	115

6.5.2	Optimierungskriterium	116
6.5.2.1	Effiziente Konturbestimmung	119
6.5.2.2	Korrespondenzsuche	120
6.5.3	Initialisierung	120
6.5.4	Optimierungsvarianten	121
6.5.4.1	Geschlossene Lösung der Deformation	121
6.5.4.2	Gradientenabstieg (Levenberg-Marquardt)	123
6.5.4.3	Randomisierte Suche	124
6.5.4.4	Rastersuche	124
6.5.5	Optimierungsstrategie	124
6.6	Ergebnisse	126
6.6.1	Simulationsergebnisse	126
6.6.2	Anwendung auf reale Fluoroskopieaufnahmen	131
7	Experimente in klinischer Umgebung	139
7.1	Vorklinischer Versuch (Klinikum Rechts der Isar)	139
7.2	Robotergestützter Kniegelenkersatz (Phantomstudie)	141
8	Zusammenfassung	145
A	Anatomische Grundlagen	147
B	Röntgenbildverstärker-Kalibrierung	151
C	Fehlerabschätzung bei planarer Approximation	157
D	Repräsentation von dreidimensionalen Oberflächenmodellen	161
E	Finite-Elemente-Methoden (FEM) – Grundlagen	163
	Literaturverzeichnis	167

Abbildungsverzeichnis

2.1	Orthopädische Anwendungsszenarien	8
3.1	OP-Szenario	20
3.2	Röntgenbildverstärker-System (Fluoroskop)	20
3.3	Idealisierte Kamera-Modellierung (Fluoroskop)	21
3.4	Computer-Tomographie	22
3.5	Optisches 6D-Trackingsystem	23
4.1	Festlegung von Koordinatensystemen	26
4.2	Vorgabe orthogonaler Aufnahmerichtungen	31
4.3	Unsicherheitsbereiche Relativpositionierung	33
4.4	Definition anatomischer Landmarken	34
4.5	Interaktive Zuordnung	35
4.6	Manuelle Modellplatzierung	38
5.1	Aussagefähigkeit von Konturen und Grauwertstrukturen	42
5.2	Intensitätsbasierte Registrierung (Prinzip)	45
5.3	Simulation von Röntgenbildern	47
5.4	Differenzbild aus realer und simulierter Röntgenaufnahme	49
5.5	Artefakte durch Metallmarken	51
5.6	Lokale Glättung von Artefakten	53
5.7	Artefakt-Behandlung: Lokale Glättung	54
5.8	Glättung/Interpolation großflächiger Bildteile	55
5.9	Lokale Glättung: Ausschnittsvergrößerung	56
5.10	Minimierungsalgorithmus	58
5.11	Unterschiedliche Projektionswirkung von 3D-Transformationen	62
5.12	Xray-View-Modell: Einschränkung der Ansicht	64
5.13	Ansichtstraining – Trainingsdaten	65
5.14	Bildapproximation durch Linear-Kombination	66
5.15	Durchschnittsbild und Eigenvektoren/bilder	66
5.16	Eigenwert-Verteilung mit akkumuliertem Variationsanteil	67
5.17	Parameter-Interpolation: B-Spline-Fläche im Eigenraum	69
5.18	Planare Approximation von R_z , T_x , T_y und T_z	70
5.19	Vergleich Ray-Casting und Approximation	74
5.20	Lokales Fehlergebirge	77

5.21	Konvergenzverhalten für simulierte Testdaten	78
5.22	Synthetisch verrauschte Testbilder	80
5.23	Konvergenzverhalten bei Rauschen	81
5.24	Restfehler-Streuung bei Konvergenz	82
5.25	Rausch-Sensitivität	82
5.26	Variation der Aufnahmerichtungen	83
5.27	T11-Brustwirbel (SIEMENS Siremobil 2)	85
5.28	Konvergenzverhalten für reale Testdaten	86
5.29	Zeitbedarf Optimierung	87
5.30	Variation der Eigenvektor-Anzahl	88
6.1	Abschätzung des Aufenthaltsvolumens	92
6.2	Anatomisch plausible Wirbelformen	93
6.3	FEM-Modellierung von Femur-Variationen	96
6.4	Dichtes <i>Resampling</i> der Oberflächenstruktur	102
6.5	Beispielobjekt (T6 Wirbel) – Verschiedene Ansichten	104
6.6	Dichtes Modell nach erfolgter Glättung	104
6.7	Ergebnis des Signifikanz-Tests	104
6.8	Landmarken-Detektion (Beispiele)	105
6.9	Korrespondenzsuche bzgl. Referenzobjekt-Vertices	108
6.10	Brustwirbel-Modelle	109
6.11	Landmarken- vs. vertexbasierte Ausrichtung	110
6.12	Alignment von drei Trainingsobjekten	111
6.13	Korrespondenzsuche nach lokaler Verformung	112
6.14	Elastisches Alignment des Dornfortsatzes	114
6.15	Modellbasierte Kontursegmentierung	117
6.16	Abstand zwischen Vertex und Projektionslinie	119
6.17	Formvariables Wirbelmodell (Verformungsmodus 1)	127
6.18	Formvariables Wirbelmodell (Verformungsmodus 2)	127
6.19	3D-Aufnahmeszene (Brustwirbelmodell)	128
6.20	Konvergenzverhalten (Simulation I)	130
6.21	Modellanpassung I (2D-Projektionen)	132
6.22	Konvergenzverhalten (Simulation II)	133
6.23	Modellanpassung II (2D-Projektionen)	134
6.24	Fehlerverlauf während Minimierung	135
6.25	Rigide Restfehler-Streuung	135
6.26	Restfehler-Streuung (Deformation)	136
6.27	Reale Wirbelaufnahmen (Modellanpassung)	137
6.28	Reale Femuraufnahmen (Modellanpassung)	138
7.1	Vorklinischer Versuch (Klinikum Rechts der Isar)	140
7.2	Versuchsaufbau (Phantomstudie)	142
7.3	Robotergestützte Femur-Präparation	143

A.1	Femur-Anatomie	148
A.2	Lendenwirbel-Anatomie	149
A.3	Teil der Brust- und Lendenwirbelsäule	150
B.1	Kamera-Koordinatensystem	152
B.2	C-Bogen-Kalibrierung	153
B.3	Fluoroskopieaufnahmen (Kalibrierung)	154
B.4	Kalibrierkörper 1	155
B.5	Kalibrierkörper 3	155
C.1	Planare Approximation – Definitionen	158
D.1	Komponenten eines 3D-Objektes	161
E.1	FEM: Knoten und Volumenelemente	163
E.2	Feder-Masse-Repräsentation	165

Tabellenverzeichnis

4.1	Restfehler (Relativpositionierung)	34
4.2	Restfehler (Interaktive Zuordnung)	37
4.3	Restfehler (Interaktive Modell-Plazierung)	38
5.1	Eigenwertanalyse Ansichtsmodell	67
6.1	Eigenwertanalyse Elastisches Wirbelmodell	129
D.1	Statistische Informationen zu dem Modell aus Abb. D.1	162

Kapitel 1

Einführung

1.1 Thematisches Umfeld der Arbeit

Seit einigen Jahren halten zunehmend computerunterstützte Systeme Einzug in den Operationsaal und ergänzen konventionelle Operationstechniken auf vielfältige Art und Weise. Das gemeinsame Hauptziel solcher Systeme – sei es auf dem Gebiet der Chirurgie (*Computer-assisted Surgery, CAS*) oder auch der Therapie (*Computer-assisted Therapy, CAT*) – besteht in der Regel darin, herkömmliche medizinische Eingriffe oder Behandlungsmethoden zu optimieren. Die Optimierung verteilt sich dabei auf die folgenden vier Schwerpunkte [TPD98]:

1. Optimale Planung eines Eingriffs oder einer Behandlung unter Einsatz multimodaler Bildgebung (z.B. Computer-Tomographie (kurz: CT), Magnetresonanz-Tomographie u.a.).
2. Hochgenaue Übertragung der Planungsergebnisse in den Operationsaal bzw. auf den Patienten sowie präzise Durchführung der geplanten Schritte.
3. Unterstützung minimal invasiver Eingriffe (*Schlüsselloch-Chirurgie*) durch eine verbesserte intraoperative Bildgebung (Lokalisation der anatomischen Zielstruktur) sowie eine deutlich verringerte Zerstörung umliegendes Gewebes beim Zugang zu der Zielstruktur (durch kleinste Schnitte bzw. körpereigene Öffnungen).
4. Erhöhung der Sicherheit bei komplizierten Eingriffen, z.B. durch die vielfältige automatische Überwachung des Patienten und seiner Lebensfunktionen.

Eine besondere Stellung nimmt Punkt 2 ein, da die präzise Planung operativer Schritte (Punkt 1) überhaupt nur dann Sinn macht, wenn sich diese intraoperativ auch in entsprechender Genauigkeit auf den Operationsaal übertragen lassen. Diese Übertragung erfordert eine intraoperative Lagerkennung des Patienten, da dessen anatomische Strukturen relativ zu einem Referenzsensorysystem innerhalb des Operationsaals erfaßt werden müssen. Dienen intraoperativ bildgebende Systeme – z.B. auf Ultraschall- oder Röntgenbasis – als Sensorysystem, so spricht man gemeinhin von bildgestützter Chirurgie (*Image-guided Surgery, IGS*). Dies gilt auch für den intraoperativen Einsatz tomographischer 3D-Bildgebung (Computer-Tomographie, offene Kernspin-Tomographie), wenngleich die zugehörigen Geräte aufgrund ihrer sperrigen Geometrie, des stationären Aufbaus sowie des sehr hohen Preises i.a. nur präoperativ genutzt werden. Steht jedoch ein Tomographiesystem intraoperativ zur Verfügung, so erübrigt sich dadurch

eine Patientenlokalisierung völlig. Sowohl dreidimensionale Planungsdaten als auch reale Patientendaten entstammen demselben Signalgeber [HB80, KHJH88].

Erste Wahl bei intraoperativer Bildgebung sind heutzutage sicherlich Fluoroskopieaufnahmen, die mit Hilfe eines mobilen Röntgenbildverstärker-Systems gewonnen werden. Solche Geräte sind flächendeckend vorhanden, sie befinden sich in tagtäglichem Gebrauch und ihr Einsatz ist klinische Routine. Allerdings bieten Röntgenaufnahmen lediglich zweidimensionale Ansichten der dreidimensionalen Realität. Daher ist letztendlich die i.a. sehr gute räumliche Vorstellungskraft des Chirurgen, kombiniert mit dessen Wissen über die Anatomie, essentiell für den Operationsverlauf. Präoperative Computer-Tomographie kann die fehlende 3D-Information bereitstellen, wobei die CT-Daten intraoperativ mit den Fluoroskopieaufnahmen registriert werden, was einer Lageerkennung des Patienten entspricht. Aufgrund der enormen Strahlenbelastung ist der Einsatz einer CT aber nicht unumstritten.

1.2 Aufgabenstellung

1.2.1 Anforderungen und Zielsetzung

Das primäre Ziel dieser Arbeit besteht in der Untersuchung und Entwicklung geeigneter Ansätze zur intraoperativen Patientenlageerkennung. Diese Lageerkennung berücksichtigt dabei nicht den gesamten Patienten, sondern konzentriert sich je nach Operationsszenario auf eine bestimmte anatomische Region des Körpers. Im vorliegenden Fall stehen konkret Knochenstrukturen der Wirbelsäule sowie des Hüft- und Kniegelenks im Mittelpunkt des Interesses (Anhang A).

Da die Röntgenaufnahmetechnik prinzipbedingt gut geeignet ist, um Knochenstrukturen abzubilden, soll die Lageerkennung unter Verwendung von intraoperativen Fluoroskopieaufnahmen erfolgen. Eine Kalibrierung des bildgebenden Sensorsystems wird dabei vorausgesetzt (Anhang B).

Das im nächsten Abschnitt vorgestellte Lösungskonzept der vorliegenden Arbeit wird wesentlich durch die nachfolgend aufgeführten Anforderungen und Randbedingungen beeinflusst (vgl. Abschnitt 2.3.3):

- **Verwendung natürlicher Merkmale:** Es sollen ausschließlich natürliche Merkmale einer Knochenstruktur für deren Lokalisation herangezogen werden, die sich durch eine geeignete Bildauswertung von Fluoroskopieaufnahmen detektieren lassen.

Im Gegensatz hierzu verwenden einige bekannte Verfahren – z.B. die kommerziellen Chirurgieroboter ROBODOC [Bö98] bzw. CASPAR [HRB⁺99] – mehrere sogenannte Referenzpins aus Metall, um die intraoperative Lokalisation bzw. Registrierung auf diese Weise extrem zu vereinfachen. Diese Pins werden während eines zusätzlichen Eingriffs vor der eigentlichen Operation an dem betrachteten Patientenknochen starr fixiert. Anschließend erfolgt eine Computer-Tomographie-Aufnahme des Patienten, wodurch sich die Pins sowohl in den 3D-Bilddaten als auch später intraoperativ wiederfinden lassen. Die Lokalisation reduziert sich damit auf die Berechnung einer einfachen 3D-3D-Transformation [HS93a].

Ob ein solcher zusätzlicher Eingriff, der den Patienten stark belastet, gerechtfertigt ist, erscheint zumindest fragwürdig, da die Einbringung der Metallpins ausschließlich der vereinfachten Registrierung des Patienten dient. Ein therapeutischer oder operativer Nutzen ist nicht erkennbar.

- **Präzision:** Die Lokalisation der Knochenstrukturen muß hochpräzise ausfallen, da die erzielbare Genauigkeit anschließender Operationsschritte – z.B. die geeignete Präparation eines Knochens als Vorbereitung einer Implantateinbringung – davon abhängt. Das für einen bestimmten operativen Eingriff tatsächlich benötigte Mindestmaß an Genauigkeit ist allerdings selbst von erfahrenen Chirurgen oftmals nur schwer zu quantifizieren [MV98]. Für orthopädische Anwendungen hat sich in den letzten Jahren jedoch die folgende Faustregel durchgesetzt, die eine obere Abschätzung für akzeptable translatorische bzw. rotatorische Restfehler ($\Delta t, \Delta r$) einer Lokalisation definiert: $\Delta t \leq 1mm$ bzw. $\Delta r \leq 1^\circ$.
- **Stabilität:** Ein Registrierungsverfahren muß hinsichtlich seines Konvergenzverhaltens stabil bzw. robust sein. Darunter ist die Garantie zu verstehen, daß sich im Falle kleiner Änderungen der Ausgangsdaten auch das Ergebnis der Registrierung nur geringfügig ändern wird. Stabilität oder Robustheit stellt in der Tat ein Schlüsselkriterium für die praktische Einsatzfähigkeit einer Registrierung dar, da sich Restfehler in den Meßdaten der eingesetzten Sensorik nie völlig ausschließen lassen. Für die vorliegende Arbeit ist z.B. stets von minimalen Restfehlern der Kalibrierung des Röntgenbildverstärkers (vgl. Anhang B) auszugehen, was letztlich dazu führt, daß die angenommene relative Ausrichtung mehrerer Aufnahmerichtungen nicht exakt der tatsächlichen Situation entspricht. Nichtzuletzt aufgrund des sehr geringen Signal/Rausch-Verhältnisses von Fluoroskopieaufnahmen sind diese selbst oftmals Ausgangspunkt für eine fehlerbehaftete Bildinterpretation (z.B. Ausreißer einer Segmentation). Ein robuster Algorithmus wird auch angesichts fehlerbehafter Eingabedaten weiterhin zuverlässig konvergieren und zumindest approximativ dasselbe Resultat liefern wie im Idealfall.
- **Geschwindigkeit:** Die Zeitkomplexität einer Registrierung sollte im angemessenen Verhältnis zur Gesamtdauer der betrachteten medizinischen Anwendung stehen. Im klinischen Umfeld ergeben sich hauptsächlich zwei Arten von Zeitrestriktionen. Werden im Rahmen eines klinischen Protokolls sukzessive mehrere Registrierungen benötigt, so sollten diese innerhalb eines vernünftigen Zeitrahmens durchgeführt werden können, ohne den klinischen Ablauf wesentlich zu verzögern. Die zweite Art von Zeitrestriktion ist noch kritischer zu sehen. Sie betrifft das Anwendungsszenario der vorliegenden Arbeit, d.h. die betrachtete Registrierung erfolgt intraoperativ mit dem Ziel einer Patientenlokalisierung. Da sich je nach Anwendung der Patient zum Zeitpunkt der Registrierung in Narkose befindet und auch das Operationsfeld bereits geöffnet wurde, ist die für die gesamte Operation zur Verfügung stehende Zeit meist knapp bemessen. Deshalb darf es in diesem Fall unter keinen Umständen zu unnötigen Verzögerungen durch die Registrierung kommen.
- **Anwendbarkeit im OP-Umfeld:** Die spezielle Umgebung eines Operationssaals stellt besondere Anforderungen an darin zum Einsatz kommende Systeme. Die prinzipielle Enge sowie die Abgrenzung eines sterilen Bereichs schränken die Anzahl an Personen, die sich gleichzeitig in dem Raum aufhalten können, drastisch ein. Um zusätzliches Bedienpersonal für die Kontrolle eines OP-Computers zu vermeiden, sollte daher der Ablauf einer Registrierung möglichst einfach gestaltet sein und sowohl den manuellen Bedienungsaufwand als auch das für die Durchführung benötigte Know-How gering halten.
- **Qualitätskontrolle:** Die Qualität der Transformation, die sich nach der Durchführung eines Registrierungsverfahrens ergibt, sollte für das OP-Team in gewissem Rahmen überprüfbar sein. Nur

auf diese Weise kann ein Fehlverhalten erkannt werden, ehe die berechneten Transformationen als Grundlagen weiterer Operationsschritte dienen. Je nach Anwendungsfall genügt ggf. bereits eine qualitative Beurteilung der erzielten Genauigkeit (z.B. durch visuelle Überlagerung von Röntgenbild und projizierten Modelldaten). Ansonsten sind quantitative Angaben über Restfehler sehr sinnvoll.

Im folgenden Abschnitt wird das in dieser Arbeit entwickelte Lösungskonzept für eine fluoroskopiebasierte Patientenlokalisierung vorgestellt.

1.2.2 Lösungskonzept

Die vorliegende Arbeit fundiert auf der Grundannahme, daß die hohen Präzisionsanforderungen nur durch die Integration geeigneter 3D-Information in den Registrierungsprozeß erfüllt werden können. Hier bietet sich in erster Linie die bereits erwähnte Computer-Tomographie an, da sich hiermit insbesondere Knochenstrukturen kontrastreich und in hoher Auflösung dreidimensional erfassen lassen. Allerdings wird der Patient durch eine Computer-Tomographie einer starken Strahlenbelastung ausgesetzt, die insbesondere bei Aufnahmen im Bereich der Wirbelsäule kritisch zu werten ist, da sich hier in unmittelbarer Umgebung eine Vielzahl sensibler Organe befindet. Ein weiteres Problem stellt sich im Zusammenhang mit Notfalloperationen: Hier ist es zeitlich oft nicht mehr möglich, vor einem Eingriff Computer-Tomographie-Aufnahmen des Patienten zu erstellen.

Aus diesen Überlegungen heraus stellt sich das Lösungskonzept dieser Dissertation zweigeteilt dar. Ein erstes Registrierungsverfahren nutzt die Vorteile der Computer-Tomographie. Seine Anwendung empfiehlt sich dann, wenn zum einen genügend Zeit für die Akquisition der Bilddaten zur Verfügung steht und zum anderen die Strahlenbelastung für den Patienten im konkreten Anwendungsfall vertretbar erscheint. Ein zweites Verfahren hingegen verzichtet auf präoperative Bildgebung. 3D-Informationen über den betrachteten Knochen werden in diesem Fall durch die patientenübergreifende Modellierung von Formvariationen gewonnen:

1. **Rigide Registrierung von Computer-Tomographie mit Fluoroskopieaufnahmen** (Kapitel 5): Ausgangsbasis für den ersten Registrierungsansatz ist zum einen ein präoperativ erstellter Computer-Tomographie-Datensatz des betrachteten Patientenknochens, zum anderen eine Reihe intraoperativ gewonnener Fluoroskopieaufnahmen des Knochens aus jeweils unterschiedlichen Richtungen. Die intraoperative Berechnung künstlicher Röntgenbilder auf Basis der Computer-Tomographiedaten ermöglicht den pixelweisen Vergleich dieser Bilder mit den realen Fluoroskopieaufnahmen. Eine geeignete Optimierung der Übereinstimmung zwischen korrespondierenden Aufnahmen liefert schließlich einen präzisen Rückschluß auf die Lage des betrachteten Knochens. Wesentliche Beiträge zu zentralen Problemstellungen in diesem Kontext beschäftigen sich mit einer adaptiven Kontrastanpassung und der Berücksichtigung von Bildartefakten (Abschnitt 5.2), der robusten Formulierung der Optimierungsproblematik (Abschnitt 5.3) sowie der beschleunigten Simulation von Röntgenaufnahmen (Abschnitt 5.4.2).
2. **Elastische Registrierung eines formvariablen 3D-Modells mit Fluoroskopieaufnahmen** (Kapitel 6): Im zweiten Ansatz zur fluoroskopiebasierten Registrierung dient ein formvariables 3D-Modell der patientenunabhängigen, anatomischen Modellierung zu erwartender Formausprägungen eines bestimmten Knochens. Position, Orientierung, Skalierung und Form eines solchen Modells lassen sich intraoperativ an die – in den einzelnen Fluoroskopieaufnahmen segmentierten

– Konturen bzw. Silhouetten des betrachteten Knochens anpassen. Das Kapitel beschäftigt sich zunächst intensiv mit der geeigneten Modellierung von Formvariation (Abschnitt 6.2). Vorgesprochen wird letztendlich die Kombination aus physikalisch motivierter Modellierung (*Finite Elemente*-Methode, Abschnitt 6.3) und statistischer Modellierung durch Training (Abschnitt 6.4). Anschließend wird ein Ansatz entwickelt, der die präzise Lage- und Formschatzung des betrachteten Patientenknochens unter Zuhilfenahme eines zuvor erstellten elastischen, formvariablen Modells ermöglicht.

1.3 Übersicht

Neben der zurückliegenden Einführung sowie einer abschließenden Zusammenfassung der Ergebnisse (Kapitel 8) beinhaltet die vorliegende Arbeit sechs weitere zentrale Kapitel.

Zunächst wird in Kapitel 2 ein breiter Überblick über den aktuellen Stand der Forschung auf dem Gebiet der Operationsunterstützung durch Computersysteme gegeben. Dabei stehen bisherige Ansätze zur präzisen intraoperativen Lageerkennung patientenspezifischer Knochenstrukturen im Vordergrund. Desweiteren werden die bekannten Nachteile eng verwandter Arbeiten diskutiert.

Kapitel 3 stellt anschließend drei medizinische Hardware-Komponenten vor, die im Kontext dieser Arbeit eine wesentliche Rolle spielen. Dabei handelt es sich in erster Linie um ein mobiles Röntgenkamerasystem (C-Bogen, Fluoroskop), das als zentraler intraoperativer Signalgeber dient. Ebenfalls intraoperativ kommt ein Infrarot-basiertes, echtzeitfähiges Trackingsystem zum Einsatz, um die Position und Aufnahmerichtung des Fluoroscops, die Navigation chirurgischer Instrumente sowie ggf. eine Bewegung des Patienten zu überwachen. Als präoperative 3D-Bildgebungsmodalität wird die Computertomographie genutzt, um daraus entweder künstliche Röntgenbilder zu generieren (Abschnitt 5.2) oder durch geeignete Segmentation dreidimensionale Modelle von Knochen unterschiedlicher Patienten als Ausgangsbasis für ein statistisches Formvariationstraining zu erhalten (Abschnitt 6.4). Ein Systemüberblick über die Integration der vorgeschlagenen Verfahren zur Patientenlokalisierung in ein prä- und intraoperatives Operationsszenario rundet das Kapitel ab.

Die mathematischen Grundlagen einer Lageerkennung von Knochen (Registrierung) sind Inhalt von Kapitel 4. Desweiteren werden dort mehrere Techniken, die sich für eine initiale Groblageschätzung eignen, präsentiert. Eine solche initiale Lagebestimmung ist nötig, da beide neu entwickelten Registrierungsverfahren ein eingeschränktes Konvergenzverhalten besitzen. Nur im Falle einer hinreichend guten Initialisierung gelingt eine präzise Lageschätzung. Die Auswahl der verschiedenen Initialisierungsverfahren orientiert sich dabei an den gegebenen Besonderheiten der Aufnahmesituation im Zusammenhang mit einem Fluoroskop.

Kapitel 5 und 6 stellen das Herzstück der vorliegenden Dissertation dar. Gemäß des im letzten Abschnitt vorgestellten Lösungsansatzes werden hier die beiden neuen Verfahren zur röntgenbasierten Lageerkennung knöcherner Strukturen jeweils motiviert, ausführlich vorgestellt und abschließend bewertet.

In Kapitel 7 werden schließlich erste Erfahrungen und Ergebnisse vorgestellt, die demonstrieren, wie sich die in dieser Arbeit entwickelten Verfahren zur Lokalisierung von Knochenstrukturen in klinischer Umgebung bewährt haben. Ein erstes Experiment sollte zeigen, wie präzise sich bestimmte anatomische Stellen eines Leichenkniees nach erfolgter Lagererkennung des Oberschenkelknochens durch die überwachte Navigation eines chirurgischen Zeigeinstrumentes wiederfinden lassen. Die betrachteten Stellen des Knochens waren dabei präoperativ durch das Setzen von Bohrlöchern speziell präpariert worden. Die

Zielsetzung eines weiteren, zweiten Experiments bestand hingegen darin, das distale Ende eines Oberschenkelknochenphantoms durch mehrere komplexe, roboterunterstützte Sägeschnitte geeignet zu bearbeiten, um anschließend ein Titan-Implantat möglichst paßgenau am Knochen zu befestigen.

Kapitel 2

Relevante Arbeiten – Stand der Forschung

2.1 Orthopädische Anwendungsszenarien

Nachfolgend werden einige orthopädische Eingriffe im Hüft-, Knie und Wirbelsäulenbereich, bei denen aufgrund der erforderlichen Genauigkeit eine präzise computergestützte Registrierung und Navigation sinnvoll ist, beispielhaft erläutert [Stu96].

Künstlicher Kniegelenksersatz (*total knee replacement, TKR*)

Degenerativen Veränderungen des Kniegelenks können vielfältige Ursachen zugrundeliegen. Am häufigsten jedoch liefert eine Osteoarthritis den Grund für einen operativen Eingriff mit dem Ziel eines vollständigen Kniegelenksersatzes. Bei Osteoarthritis handelt es sich um stark fortgeschrittenen Gelenkverschleiß, der z.B. ursächlich durch eine u.U. bereits sehr lange zurückliegende Gelenkverletzung ausgelöst worden sein kann. Insbesondere bei älteren Patienten tritt eine Osteoarthritis häufig auch ohne vorangegangene Verletzung auf. Die Gelenkknorpelschichten an Femur und Tibia sind dabei so stark abgenutzt, daß beide Knochen an bestimmten Stellen direkten Kontakt bekommen, was wiederum bei Bewegung und Belastung des Knies starke Schmerzen für den Patienten nach sich zieht.

Ein künstliches Kniegelenk besteht üblicherweise aus vier Einzelteilen. Eine zweiteilige tibiale Komponente ersetzt das proximale Ende der Tibia. Eine femurale Komponente übernimmt die Funktion der beiden Femurkondylen und stellt desweiteren eine künstliche *Gleitmulde* für die Kniescheibe bereit. Eine patellare Komponente ersetzt schließlich die nach innen gerichtete Knorpelfläche der Kniescheibe (Abb. 2.1 (links)). Die beiden Komponenten, die direkt mit Femur bzw. Tibia in Kontakt gebracht werden, bestehen i.a. aus einer Titan-Legierung, wobei die stabile Verbindung zum Knochen entweder durch eine spezielle Art von Klebstoff oder durch die Unterstützung natürlicher Verwachsungen (poröse Implantatoberfläche) hergestellt wird. Das zweite Teil der tibialen Komponente, das innerhalb der Gelenkanordnung in etwa die Stelle der Minisken einnimmt, besteht wie auch die patellare Kappe aus extrem widerstandsfähigem Plastik. Auf diese Weise wird verhindert, daß Metallteile mit anderen Metallteilen in Berührung kommen.

Die operative Durchführung des Eingriffs ist eine anspruchsvolle Aufgabe für den Chirurgen. Insbesondere die Präparation des distalen Femurendes sowie der proximalen Tibia nehmen den größten Teil der insgesamt oft zwei Stunden dauernden Operation ein. Dabei gilt es, eine Vielzahl an Randbedingungen zu beachten: optimale Paßgenauigkeit der Prothesenteile am Knochen, korrekte Gesamtstellung des

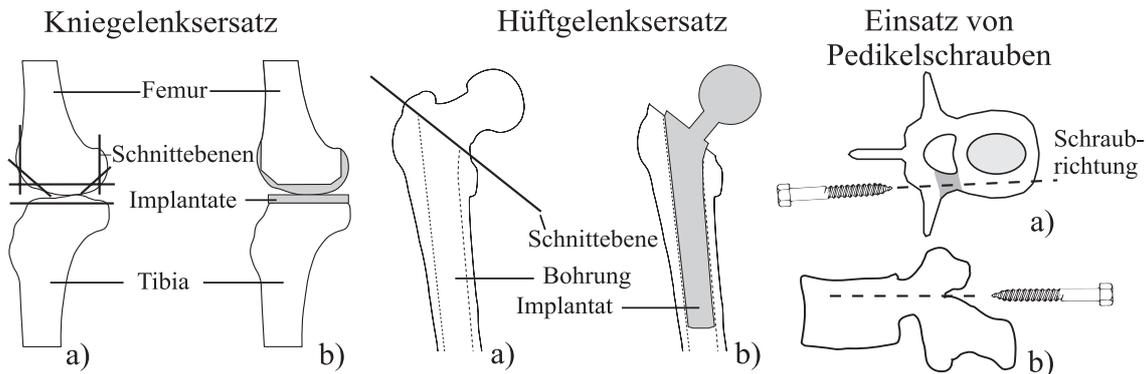


Abbildung 2.1: Orthopädische Anwendungsszenarien an Knie (links), Hüfte (Mitte) und Wirbelsäule (rechts). Gegenübergestellt wurden jeweils (a) Planungssituation und (b) Implantatanbringung (links/Mitte) bzw. zwei unterschiedliche Planungsansichten (rechts).

Knies u.v.m.

Künstlicher Hüftgelenksersatz (*total hip replacement, THR*)

Eine Indikation für einen kompletten Hüftgelenksersatz ist zumeist nach schweren Frakturen oder bei stark fortgeschrittener Arthritis angezeigt. Der operative Eingriff sieht zunächst das Aufbohren der Hüftpfanne und deren Ersatz durch einen künstlichen Einsatz vor. Zusätzlich wird der proximale Femurteil inklusive Femurkopf und -hals entfernt und ebenfalls durch eine anatomisch ähnlich geformte künstliche Komponente ersetzt (Abb. 2.1 (Mitte)). Die einzelnen Teile der Prothese bestehen aus einer Metall-Legierung, wobei in den meisten Fällen Titan verwendet wird. Beide Prothesenteile werden entweder wiederum über Klebstoff mit den übrigen Knochenstrukturen verbunden oder sie bieten durch eine aufgerauhte, poröse Oberfläche den umgebenden Knochen die Möglichkeit, langsam mit den Prothesen zu verwachsen. Nach heutigen Erfahrungen kann eine solche Hüftprothese durchaus bis zu 15 Jahre halten.

Wie auch bei einer TKR-Operation ist die Präparation des Femurknochens eine anspruchsvolle Tätigkeit. Entspricht die Aushöhlung des proximalen Femurendes nicht der Form des Titanimplantats, so entsteht keine optimale Haftung. Die Folge ist eine frühzeitige Lockerung des Implantats. D.h., es ist mit deutlich verkürzten Standzeiten zu rechnen.

Wirbelsäulen-Versteifung durch Pedikelschrauben

Häufigste Erkrankungen der Wirbelsäule sind degenerative Veränderungen sowie rheumatische und tuberkulöse Entzündungen. Versteifungen führen zur sogenannten *Bechterewschen Krankheit*. Daneben treten mitunter krankhafte Fehlkrümmungen der Wirbelsäule nach vorne (*Lordose*), hinten (*Kyphose*) oder zur Seite (*Skoliose*) auf. Schließlich verursacht auch das Zusammenfallen zerstörter Wirbelkörper – z.B. aufgrund von Tuberkulose – eine Fehlkrümmung. Eine mögliche Behandlungsmethode der genannten degenerativen Veränderungen ist die partielle Versteifung der Wirbelsäule. Hierzu werden an mehreren Wirbeln über und unter dem betroffenen Wirbelsäulenfragment sogenannte Pedikelschrauben aus Metall fixiert, die anschließend durch ein parallel zur Wirbelsäule verlaufendes Metallgestänge starr miteinander verbunden werden. Dadurch wird der betrachtete Teil der Wirbelsäule insgesamt versteift,

was dem Chirurgen die Möglichkeit gibt, zum einen krankhaft verkrümmte Teilstücke wieder auszurichten und zum anderen eine instabile Wirbelsäule neu zu stabilisieren.

Die große Schwierigkeit in Zusammenhang mit dieser Art von Eingriff liegt in der korrekten Platzierung der Pedikelschrauben. Diese Schrauben werden von hinten in die Pedikel des jeweils betrachteten Wirbels eingebracht. Die Pedikel sind dünne stielartige, den Markkanal seitlich begrenzende Knochenstrukturen, die den Wirbelkörper mit seinen hinteren Quer- und Dornfortsätzen verbinden (Anhang A). Erfolgt die Verschraubung nicht mit höchster Genauigkeit, kann es passieren, daß eine Pedikelschraube den Pedikel nach innen oder außen verläßt und dabei direkt benachbarte Nervenstränge oder gar das sensible Rückenmark selbst auf bedrohliche Art und Weise verletzt (Abb. 2.1 (rechts)). Dabei ist nicht nur der geringe Pedikeldurchmesser ein Problem. Auch das Blickfeld des Chirurgen ist während der Operation extrem eingeschränkt; der bearbeitete Wirbel ist nur von hinten zu sehen.

Ein geeigneter computerunterstützter Ansatz ermöglicht hier zum einen die präoperative Planung der optimalen einzelnen Schraub-Trajektorien. Zum anderen kann intraoperativ die aktuell durchgeführte Verschraubung unter ständiger Kontrolle und in Abgleich mit der Planung erfolgen.

2.2 Bekannte computerunterstützte Chirurgesysteme in der Orthopädie

Für jede der drei im letzten Abschnitt vorgestellten Operationstechniken (TKR, THR und Wirbelsäulenversteifung) gibt es bereits heute computerunterstützte Systeme, die sich in klinischer Anwendung oder zumindest in vorklinischer Erprobung befinden. Die wichtigsten dieser Systeme werden im folgenden kurz vorgestellt.

2.2.1 Künstlicher Hüftgelenkersatz

Das bekannteste System auf dem Gebiet des vollständigen Hüftgelenkersatzes ist sicherlich der Chirurgieroboter ROBODOC [Mit97], da dieser sich bereits vereinzelt in klinischer Anwendung wiederfindet [Bö98]. Das bereits seit 1987 in den USA entwickelte Konzept beinhaltet einen fünfschichtigen Roboter, der während der Operation das Ausfräsen des Femurschaftes übernimmt. Das Ausfräsen erfolgt mit Hilfe eines am Roboterarm angeflanschten, scharfen und rotierenden Messers [Tay96]. Ein zugehöriges Planungssystem [Jer99] ermöglicht es, präoperativ anhand von CT-Schichtaufnahmen des Patientenfemurs exakt festzulegen wie der Markkanal des Femurs später geweitet werden muß, um eine optimale Paßgenauigkeit zwischen Knochen und Implantat zu erreichen. Die Erfassung der intraoperativen Patientenlage erfolgt hierbei durch künstliche Metallpins, die bereits vor Aufnahme der CT in den Femur des Patienten implantiert werden müssen. Die intraoperative Abtastung dieser Marken durch den Roboter ermöglicht die Korrelation des Planungskoordinatensystems (CT) mit dem Roboterkoordinatensystem im Operationssaal. Seit 1997 befindet sich ein weiteres, sehr ähnlich ausgelegtes Robotersystem namens CASPAR im klinischen Einsatz [HRB⁺99]. Auch bei diesem System werden künstliche Metallpins implantiert, um die intraoperative Registrierung des Patienten bzgl. des Roboters zu vereinfachen.

Mit *HipNav* steht ein weiteres System für einen THR-Eingriff zur Verfügung [DiG97, Sim97]. Statt eines Roboters kommt hier ein optisches Trackingsystem zum Einsatz, um dem Chirurgen die genaue Navigation des verwendeten Instrumentariums zu erleichtern. Die Registrierung des Patienten bzgl. einer präoperativ erstellten CT erfolgt durch die intraoperative Oberflächenabtastung des teilweise freigelegten Beckenknochens mit Hilfe des optischen Trackingsystems.

In [Bra97] wird ein roboterbasiertes System vorgestellt, welches erstmals intraoperativ Fluoroskopieaufnahmen in den Ablauf einer computerunterstützten THR integriert. Die gleichzeitige Projektion eines virtuellen Implantats in mehrere zuvor aufgenommene Bilder erlaubt es dem Chirurgen, durch interaktive Veränderung der Position und Orientierung dieses Implantats die optimale Projektion zu bestimmen. Da der Roboter in eine zuvor durchgeführte Sensorkalibrierung des Fluoroscops integriert war, ist anschließend das Ausfräsen ohne eine weitere explizite Registrierung des starr fixierten Patientenbeines möglich.

Die Möglichkeit, patientenspezifische Individual-Prothesen anzufertigen [WH99], erweitert inzwischen die potentiellen Einsatzmöglichkeiten der vorgestellten Systeme.

2.2.2 Künstlicher Kniegelenksersatz

Aufgrund der deutlich höheren Komplexität eines TKR-Eingriffs im Vergleich zu THR sind erst in jüngster Zeit verstärkt Bestrebungen festzustellen, auch diese Operationstechnik per Computereinsatz geeignet zu unterstützen. Erste Ansätze hierzu werden in [Rot96, Gos97] und [Mar96] vorgestellt. Das Konzept von Roth et al. [Rot96] beinhaltet dabei die Präparation des distalen Femures durch eine spezielle Sägevorrichtung, die von einem Roboter in präoperativ geplanten Schnittebenen positioniert wird, ehe der Chirurg den eigentlichen Sägevorgang – beschränkt auf die vorgegebene Ebene – selbst durchführt. Die Registrierung des Patientenbeines erfolgt ohne die Verwendung künstlicher Metallpins nach einem in dieser Dissertation vorgestellten rigiden Verfahren unter Verwendung präoperativer Computer-Tomographie sowie mehrerer kalibrierter Fluoroskopieaufnahmen [Bra00a]. Das System befindet sich zur Zeit jedoch noch in der experimentellen Erprobung. Im Gegensatz hierzu wurden die Ansätze von Kienzle et al. [Kie96] bereits an diversen Kadaverknochen getestet. Allerdings benötigt dieses Konzept ebenso wie der in [Mar96] beschriebene aktive Robotereinsatz die bereits im Zusammenhang mit ROBODOC vorgestellten Metallpins für eine Referenzierung.

2.2.3 Wirbelsäulenversteifung

Nolte et al. stellten erstmals in [Nol94, Nol95] ihr computergestütztes Navigationssystem zur optimalen Ausrichtung von Pedikelschrauben vor. Dieses System befindet sich bereits seit einigen Jahren im klinischen Einsatz. Unter optischer Kontrolle durch ein auf Infrarotbasis arbeitendes Trackingsystems wird die freigelegte Rückseite eines Patientenwirbels durch eine Abtastung vieler Oberflächenpunkte im Bereich der Dorn-, Quer- und Gelenkfortsätze initial mit einer CT-Aufnahme des Patienten registriert. Da auch das zur Einbringung der Pedikelschrauben benötigte Instrumentarium optisch überwacht wird, erhält der Chirurg während der Operation ständig eine visuelle Rückkopplung über die aktuelle Ausrichtung der Schraubenachse relativ zu den CT-Schichtaufnahmen des Patienten. Mehrere virtuell berechnete Schnittbilder durch das CT-Volumen zeigen dabei zum einen die aktuelle Position der Pedikelschraubenspitze und zum anderen zwei bis drei *Was-wäre-wenn*-Aufnahmen, d.h. Schnittbilder, die veranschaulichen, wohin sich die Schraube bewegen wird, wenn der Chirurg sie in der aktuellen Ausrichtung um z.B. weitere 2, 4 oder 6 mm eindreht. Auf diese Weise kann rechtzeitig ein Ausbrechen der Schraube aus dem engen Pedikelkanal verhindert werden. In [Mer97] wird eine Studie veröffentlicht, die aufzeigt, wie sinnvoll eine solche Navigationshilfe bei der Pedikelverschraubung ist. Nur in fünf Prozent der untersuchten Fälle wurde die Außenhülle eines Pedikels verletzt. Bei konventioneller Vorgehensweise ohne visuelle Kontrolle sind mehr oder weniger starke Verletzungen eines Pedikels hingegen eher die

Regel.

2.3 Diskussion bekannter Ansätze zur Registrierung

Sehr gute, aktuelle Übersichtsartikel zum Thema Registrierung wurden sowohl von Maintz und Viergever [MV98] als auch von Lavallée [Lav96b] veröffentlicht. In [MV98] werden dabei zunächst sehr detailliert geeignete Kriterien zur Klassifikation bekannter Registrierungsansätze jeglicher Art vorgestellt. Anschließend werden die wichtigsten Veröffentlichungen aus den Jahren 1993 bis 1997 entsprechend der festgelegten Kriterien eingeteilt. In [Lav96b] hingegen konzentriert sich der Autor auf die erschöpfende Besprechung bekannter Verfahren auf dem Gebiet der computergestützten Orthopädie.

Unter Registrierung ist allgemein die Integration mehrerer Bilddatensätze (i.a. eines Patienten) zu verstehen, welche unterschiedlichen Bildmodalitäten entstammen können. Die Integration entspricht dabei der Suche nach den optimalen Parametern einer geeigneten Transformation zwischen den mit den betrachteten Modalitäten assoziierten Koordinatensystemen. Steht bei einer Registrierung – wie in der vorliegenden Dissertation – die intraoperative Erfassung des Patienten im Vordergrund, wobei Bilddaten einer der betrachteten Modalitäten lediglich als Zwischenergebnis Verwendung finden, so spricht man üblicherweise gleichwertig von einer direkten Patientenregistrierung, -lageerkennung oder auch -lokalisation.

2.3.1 Starre Registrierung

Eine starre bzw. rigide Registrierung erlaubt lediglich Drehungen und Verschiebungen zwischen den zu registrierenden Datensätzen.

In vielen Arbeiten wird die Verwendung von künstlich am oder gar im Patienten angebrachten Marken vorgeschlagen, um dessen Registrierung zu vereinfachen. Die Position der Marken wird jeweils entweder optisch, magnetisch oder mechanisch überwacht. Buchholz et al. [BSHM94] befestigen auf der Kopfhaut des Patienten mehrere Marken, ehe eine Computer-Tomographie erstellt wird. Anschließend werden diese Marken wieder entfernt und die einzelnen Befestigungsstellen mit einem Filzschreiber markiert. Intraoperativ wird der Patient mit einem optisch überwachten Ring aus Infrarot-Leuchtdioden versehen. Außerdem erfolgt mit Hilfe eines ebenfalls optisch überwachten Zeigeinstrumentes die Referenzierung der markierten Hautstellen durch Antasten. Im Schädelbereich halten sich dabei die resultierenden Genauigkeitsverluste in Grenzen, da sich die Positionen der markierten Hautstellen aufgrund ihrer unmittelbaren Nähe zum Schädelknochen kaum verändern. Für Anwendungen im Knie-, Hüft- oder Wirbelsäulenbereich ist das Verfahren jedoch zu fehlerträchtig. Deswegen ist hier die invasive Einbringung von Metallpins in den jeweiligen Knochen i.a. erste Wahl. Diese Vorgehensweise wurde bereits in Zusammenhang mit den vorgestellten Chirurgierobotern ROBODOC [Mit97] und CASPAR [HRB⁺99] erläutert. Bedenklich ist hierbei die Notwendigkeit eines zusätzlichen operativen Eingriffs, um die Pins vor einer CT-Aufnahme einzubringen.

Die beiden Studien von Jenkinson und Smith [JS99] bzw. Tanacs, Palagyi und Kuba [TPK99] belegen die hohe Robustheit und Genauigkeit, die mit der punktbasierten 3D-3D-Registrierung erreicht werden kann, solange man darauf achtet, daß die wenigen betrachteten Punkte einen genügend großen Raum aufspannen. Dies erklärt, warum die Verwendung von Pins u.ä. in der praktischen Anwendung nachwievor sehr beliebt ist.

Im Bereich der Forschung existiert allerdings durchaus bereits eine Reihe von Alternativen, die anstatt künstlicher Marken anatomische Merkmale der Patientendaten für die Registrierung heranziehen. Reduziert man die verwendeten Modalitäten auf die Dimensionalität ihrer Datensätze, so lassen sich die bekannten Verfahren entsprechend in Techniken zur 3D-3D- bzw. 2D-3D-Registrierung einteilen. 2D-2D-Verfahren sowie Ansätze zur Integration von Veränderungen des Datensatzes über die Zeit - in Kombination mit einer 3D-Bildgebung werden dadurch auch 4D-Datensätze möglich – spielen im Rahmen der vorliegenden Arbeit keine Rolle und werden daher im folgenden nicht weiter berücksichtigt. Neben der Dimensionalität macht die weitere Unterteilung von Ansätzen hinsichtlich der verwendeten Merkmale Sinn. Pelizzari et al. [Pel89], Ryan et al. [Rya95] sowie Simon, Hebert und Kanade [SHK95] registrieren 3D-Flächen mit anderen 3D-Flächen. Ebenfalls unter die Kategorie einer 3D-3D-Registrierung fallen Verfahren, die anstatt 3D-Flächen zwei 3D-Kurven betrachten. Beispiele hierzu finden sich in [JRH92] oder [Lav96a].

Die 2D-3D-Registrierung von Umrißkonturen mit 3D-Flächen ist für die vorliegende Arbeit von besonderem Interesse, wenngleich die nachfolgend besprochenen Arbeiten jeweils rigide Transformationen betrachten. Betting und Feldmar [BF95, FAB95] verwenden in ihrer Arbeit die folgende, auf Silhouetten basierende Technik für eine 2D-3D-Registrierung von Magnetresonanz- bzw. Computer-Tomographie mit Röntgenbildern: Einer Extraktion der äußeren Konturen der betrachteten Struktur in allen Röntgenbildern folgt eine rigide Transformation des zugrundeliegenden 3D-Datensatzes (MR/CT) bzw. des daraus segmentierten 3D-Modells. Die Projektion dieses transformierten Modells in die Bildebenen aller Röntgenaufnahmen liefert jeweils eine entsprechende Silhouette, deren Abstand zu den extrahierten Konturen schließlich bzgl. der rigiden Transformationsparameter minimiert werden kann.

In Lavallée et al. [Lav95, Lav96a] registrieren die Autoren ebenfalls einen 3D-Datensatz (CT) mit zwei Röntgenaufnahmen, wobei der Winkel zwischen den Aufnahmerichtungen als bekannt vorausgesetzt wird. Auch hier werden zunächst in beiden Aufnahmen die externen Konturen der betrachteten Knochenstruktur (menschlicher Wirbel) extrahiert, d.h. segmentiert. Anschließend werden die 2D-Bildebenen der Aufnahmen entsprechend ihrer relativen geometrischen Anordnung im dreidimensionalen Raum betrachtet. Die Menge aller Projektionsgeraden durch den Kameraursprung und je einen nun räumlich interpretierten Konturpunkt ergibt ein kegelförmiges Geradenbündel für jedes der Bilder. Beide Geradenbündel zusammen betrachtet definieren durch ihren räumlichen Schnitt ein Volumensegment. Das aus den CT-Daten gewonnene 3D-Modell wird nun initial innerhalb dieses Volumens plaziert. Anschließend wird der Abstand seiner Oberfläche zu den Geradenbüscheln bzgl. einer rigiden Transformation minimiert.

Sowohl Hamadeh et al. [HSLC95, HC97] als auch Gueziec et al. [Gue98] greifen diesen Ansatz auf und erweitern ihn geeignet. Hamadeh et al. integrieren hierzu die dynamische Detektion der Knochenkonturen in die Registrierung selbst, so daß ein kooperativer Ansatz entsteht [HSLC95]. Dieselben Autoren verwenden ihr Verfahren später für eine Studie der rigiden Kinematik der Lendenwirbelsäule [HC97]. Gueziec et al. [Gue98] hingegen beschreiben in ihrer Arbeit die effiziente Suche nach Korrespondenzen zwischen 3D-Modellpunkten und Projektionslinien sowie die Formulierung einer robusten Minimierung.

Neben der expliziten Ableitung von Merkmalen existiert auch eine Reihe von Techniken, die versuchen, den Inhalt der betrachteten Bilddaten *direkt* gegenseitig zu registrieren. Die Forschungsgruppe um Weese [Wee97b, Wee97a] beschreibt beispielsweise ein Verfahren, welches – ähnlich dem in dieser Dissertation entwickelten ersten Ansatz – einen vorsegmentierten CT-Datensatz mit einer einzelnen Fluoroskopieaufnahme registriert. Hierzu werden sukzessive simulierte Röntgenbilder berechnet und an-

schließlich in der Bildebene pixelweise mit den realen Aufnahmen verglichen. Der vorgeschlagene Algorithmus versucht dabei ein besonderes Fehlermaß (*pattern intensity*), welches die *Strukturiertheit* des betrachteten Differenzbildes bewertet, zu minimieren. Die Idee dahinter ist, daß sich im Falle einer optimalen Transformation innerhalb des Differenzbildes alle Strukturen der betrachteten Knochenstrukturen gegenseitig auslöschen sollten.

Inhärentes Problem dieses und ähnlicher Ansätze ist der hohe Zeitaufwand für die Generierung der künstlichen Röntgenbilder, da hierfür jedes Mal das zugrundeliegende CT-Volumen zumindest teilweise unter Verwendung eines *Ray-Casting*-Verfahrens durchlaufen werden muß. Mit dem Ziel, dieses Handycap zu beseitigen, wurde kürzlich in [Wee99] die Integration einer aus dem Bereich der 3D-Computer-Grafik (*Volume Rendering*) bekannten Vorverarbeitung des dreidimensionalen Datensatzes vorgeschlagen. Diese sogenannte *shear-warp*-Transformation bringt im vorliegenden Anwendungsfall allerdings nur dann signifikante Zeitvorteile, wenn lediglich ein einzelnes Röntgenbild betrachtet wird. Dies liegt daran, daß die genannte Transformation bei der Optimierung des Datensatzes hinsichtlich eines späteren Renderings die Blickrichtung bereits implizit berücksichtigen muß (vgl. Abschnitt 5.2.1).

Die Verwendung einer einzelnen Fluoroskopieaufnahme resultiert aufgrund der projektiven Bilddarstellung in einer herabgesetzten Genauigkeit in Blickrichtung. Da in der vorliegenden Arbeit deshalb im Gegensatz zu [Wee99] ein *Multi-view* bevorzugt wird – d.h. die gleichzeitige Berücksichtigung mehrerer Fluoroskopieaufnahmen aus unterschiedlichen Blickrichtungen –, ist die beschriebene Zeitoptimierung nicht übertragbar. In Abschnitt 5.4 wird ein neu entwickeltes Verfahren zur beschleunigten Simulation virtueller Röntgenbilder vorgestellt.

2.3.2 Elastische Modellierung und Registrierung

Verfahren zur elastischen Registrierung erweitern gegenüber rigiden Techniken die Art der Transformation zwischen den zu registrierenden Datensätzen. Nötig wird dies beispielsweise im Zusammenhang mit anatomischen Atlanten [GBHE91, STA98]. Hier werden nicht unterschiedliche Bilddaten eines Patienten zueinander registriert, sondern das Ziel besteht vielmehr in der Anpassung eines Referenzdatensatzes (Atlas) an die Bilddaten unterschiedlicher Patienten. D.h., die Transformation muß in der Lage sein, Formunterschiede geeignet zu modellieren. Was genau den Begriff der *Form* dabei auszeichnet, läßt sich gemäß eines Zitats von Bookstein anschaulich wie folgt beschreiben [Boo91]:

„*Shape is, what remains, after translation, rotation and scaling has been removed.*“

D.h. die Modelle zur Beschreibung von Formunterschieden sind i.a. translations-, rotations- und skalierungsinvariant. Bekannte Ansätze lassen sich in erster Linie durch die Art und Weise klassifizieren, wie der elastische Anteil der Transformation modelliert und parametrisiert wird. McInerney und Terzopoulos [MT96] geben in ihrer Arbeit einen detaillierten Überblick über die Anwendung formvariabler Modelle in der medizinischen Bildinterpretation, wobei der Schwerpunkt auf 2D- bzw. 3D-Bildsegmentation liegt.

Die ersten wesentlichen Forschungsarbeiten zur Formveränderung dreidimensionaler Flächen entstanden bereits im Laufe der Achtziger Jahre im Bereich der Computergrafik. Sowohl die Arbeiten von Barr [Bar84], Pentland und Williams [PW89] und Pentland [Pen92] beschreiben vordefinierte Verformungsmodi zur Anwendung globaler und lokaler Deformationen auf einfach geformte Körper. Die Arbeit von Székely et al. [Sze95] geht einen deutlichen Schritt weiter. Die vorgestellte Technik kombiniert physikalisch motivierte Eigenschaften bekannter formvariabler Segmentationsverfahren (vgl. *Snakes* [Kas87])

mit einer *Fourier*-Parametrisierung der Formbeschreibung. Anwendung findet das Verfahren sowohl in der 2D- als auch 3D-Segmentation menschlicher Gehirnstrukturen. Überhaupt sind physikalisch motivierte Ansätze zur dynamischen Modellierung von Formveränderungen weitverbreitet [Met98]. Terzopoulos et al. [TWK87] schlagen ein dreidimensionales Modell zur Rekonstruktion von Strukturen vor, welches sich durch eine ungewöhnlich hohe Anzahl an Freiheitsgraden beeinflussen läßt. Pentland und Sclaroff [PS91] verwenden für ihren Ansatz deutlich weniger Parameter, indem sie die veränderliche Form eines Objektes durch physikalische Deformationsmodi repräsentieren, die sich durch eine *Modale Analyse (Finite-Elemente-Methode)* aus der starren Objektbeschreibung selbst ableiten lassen. Ähnliche physikalische Modellierungsprinzipien finden sich auch in den Arbeiten von Nastar und Ayache [NA94], Sclaroff [Sci95] und Subsol et al. [STA98]. Letztgenannter präsentiert in seiner Arbeit, wie sich ein morphometrischer Atlas des menschlichen Schädels aufbauen läßt.

Einen anderen Weg verfolgt Bookstein [Boo91], indem er unabhängig von der betrachteten Struktur stets die physikalischen berechenbaren Deformationsmodi einer unendlich dünnen Platte (*thin-plate splines*) als Ausgangsbasis wählt, um die auftretenden Formvariationen zweidimensionaler Punktmengen zu modellieren. Auch die Arbeit von Cutting [Cut95] basiert auf der Verwendung von *thin-plate-Splines*. Im Gegensatz zu Bookstein wendet er das Verfahren jedoch auf 3D-Flächen an und untersucht damit degenerative Veränderungen der menschlichen Schädelform.

Neben physikalisch motivierten Ansätzen existiert eine weitere sehr wichtige Klasse von Verfahren zur Formmodellierung von Objekten. Diese Verfahren basieren auf der Approximation von Variationsmodi durch die statistische Varianzanalyse einer repräsentativen Stichprobe. Die Stichprobe enthält eine Menge von Beispiel- bzw. Trainingsformen, deren Abweichung von der Durchschnittsform aller Beispielformen unter Verwendung einer Hauptachsentransformation (PCA, *principal component analysis* [Jol86]) untersucht wird. Es resultiert schließlich eine handhabbare Menge an Deformationsmodi, die es erlauben, in Linearkombination mit der Durchschnittsform repräsentative, verformte Objektformen zu erzeugen. Bekanntester Vertreter dieses statistischen Ansatzes ist die Forschungsgruppe um Taylor und Cootes [CCTG92]. Deren Trainingsdaten bestehen aus beliebigen 2D- oder 3D-Punktmengen. Ihre Modelle werden daher konsequenterweise als *Point Distribution Models* bezeichnet. Wenngleich die Autorengruppe bereits die Erweiterung ihres Ansatzes auf 3D-Modelle beschrieben hat [HTT93], beschränken sich veröffentlichte Anwendungen der Modelle bisher vorrangig auf den Bereich der medizinischen 2D-Bildanalyse [Coo95b]. Fleute und Lavallée haben diesen Modellierungsansatz 1998 aufgegriffen und für die 3D-3D Registrierung eines formvariablen Femurmodells mit einer 3D-Punktwolke herangezogen [FL98]. Später wurde dieser Ansatz auf die 2D-3D-Registrierung des Femurmodells mit segmentierten Knochenkonturen erweitert [FL99]. Der Vorteil des statistischen Ansatzes liegt in erster Linie in der statistischen Aussagekraft des resultierenden Verformungsspielraumes. Die Modellierung erlaubt bei geeigneter Beschränkung der Deformationsmodi keine entarteten Verformungen. Außerdem läßt sich exakt festlegen, wieviel prozentualen Anteil der in den Trainingsdaten enthaltenen Formvariation das Modell letztlich beinhalten soll. Der Nachteil ist ein relativ hoher Modellierungsaufwand, da das Verfahren von bekannten 1:1-Korrespondenzen zwischen Punkten auf den einzelnen Trainingsobjekten ausgeht. Wie sich in der vorliegenden Arbeit zeigt, ist die Bereitstellung der Korrespondenzen insbesondere bei komplex geformten 3D-Objekten wie z.B. Wirbeln keineswegs trivial.

In den beiden Arbeiten von Martin, Pentland und Kikinis [MPK94] und von Cootes und Taylor [CT95] wird erstmals die geeignete Kombination von physikalisch motivierter und statistischer Formmodellierung durch experimentelles Training vorgeschlagen. Die erstgenannten Autoren untersuchen das dreidimensionale Formverhalten des ventrikularen Systems im menschlichen Gehirn. Ziel der Arbeit ist

die automatische Klassifikation krankhafter Veränderungen (z.B. Alzheimer) anhand der 3D-Form der Ventrikel. Deren Form resultiert jedoch nicht nur aus krankhaften Veränderungen, sondern wird auch wesentlich von der Form des gesamten Patientengehirns bzw. -schädels bestimmt. [MPK94] schlägt daher die lineare Kombination aus physikalischen und statistischen Modi vor. Erstere approximieren die elastischen Eigenschaften der Gehirnmasse, letztere hingegen pathologisch bedingte Formvariationen. Durch diese Trennung lassen sich später globale Effekte der Gesamtform des Gehirns herausrechnen, wodurch sich die Trefferquote der Klassifikation krankhafter Veränderungen deutlich verbessert. Die von [CT95] angestrebte Kombination dient einem anderen Ziel. Hier steht nachwievordie statistische Modellierung durch Training im Mittelpunkt. Da sich jedoch in praktischen Anwendungen oftmals keine genügend große Anzahl an Beispielformen gewinnen läßt, erzeugen die Autoren zusätzliche künstliche 2D-Trainingsformen, indem sie den bereits vorhandenen 2D-Beispielformen ohne jeglichen Bezug zur Realität ein physikalisches Feder-Masse-System unterstellen. Die *Modale Analyse* erlaubt dann die Generierung neuer Formbeispiele. Alle Trainingsformen zusammen werden anschließend einer PCA unterzogen. Die wesentliche Leistung des Verfahrens von [CT95] besteht darin, daß die Generierung der zusätzlichen, künstlichen Formen bei Bedarf automatisch erfolgt. Je weniger tatsächliche Formbeispiele vorhanden sind, desto mehr zusätzliche Formen werden erzeugt und umgekehrt.

2.3.3 Validierung – Verfahrensbewertung

Ehe ein Registrierungsverfahren bei Ärzten breite Akzeptanz erfährt und in realer klinischer Umgebung an lebenden Patienten eingesetzt wird, ist die umfangreiche Validierung des Verfahrens unverzichtbar. Eine leicht nachvollziehbare Anforderung; man denke etwa an die möglichen Folgen einer fehlerhaften oder ungenauen Registrierung auf die Gesundheit oder gar das Leben des Patienten. In diesem Abschnitt werden die wichtigsten Kriterien zur Beurteilung der Qualität eines Registrierungsverfahrens vorgestellt und diskutiert.

Das wohl bekannteste Bewertungskriterium ist die mit dem Verfahren erreichbare Genauigkeit hinsichtlich der berechneten Transformation. Eine Validierung sollte sich allerdings nicht ausschließlich auf die Untersuchung von Genauigkeit beschränken, sondern die Beantwortung einer ganzen Reihe weiterer wichtiger Fragestellungen beinhalten. Eine erschöpfende Verfahrensbewertung umfaßt die folgenden Bewertungskriterien [MV98]:

- Robustheit/Stabilität
- Zuverlässlichkeit
- Ressourcenbedarf
- Algorithmische Komplexität
- Präzision
- Genauigkeit
- Teststudien

Ein Verfahren gilt als robust oder stabil, falls geringfügige Veränderungen an den Eingabedaten auch nur geringfügige Änderungen nach sich ziehen. Zuverlässlichkeit hingegen bedeutet, daß sich ein Algorithmus wie erwartet verhalten soll – vorausgesetzt er wird mit sinnvollen Eingabedaten versorgt. Sowohl

der Bedarf an Ressourcen (Bildgebende Hardwarekomponenten etc.) als auch die Komplexität eines Verfahrens sollten in einem vernünftigen Verhältnis zu dem zu erwartenden Nutzen des Verfahrens für eine klinische Anwendung stehen. Die Verfahrenskomplexität beinhaltet dabei insbesondere auch den Zeitbedarf einer Registrierung. Wie lange eine Registrierung dauert, ist gerade dann von entscheidender Bedeutung, wenn diese intraoperativ, d.h. *online* durchgeführt werden muß.

Die Begriffe Präzision und Genauigkeit finden in der einschlägigen Literatur oftmals synonyme Verwendung, wenngleich zwischen beiden ein feiner aber entscheidender Unterschied besteht. Maintz und Viergever definieren Präzision als *den typischen systematischen Restfehler, den man erhält, wenn das Verfahren mit idealen Eingabedaten versorgt wird* [MV98]. Genauigkeit hingegen ist ein direktes Maß für den tatsächlich zu beobachtenden Restfehler während einer Registrierungssituation. D.h. Präzision repräsentiert eine Eigenschaft des betrachteten Systems oder Verfahrens, wohingegen sich Genauigkeit stets auf einen einzelnen konkret durchgeführten Registrierungsablauf bezieht.

Unter idealen Umständen sollte ein Verfahren jedes der genannten Kriterien erfüllen. Das zu erwarten ist jedoch nicht realistisch. Welchen Kriterien in der Praxis mehr Brisanz beigemessen wird, hängt letztlich von der betrachteten Anwendung ab und ist im Einzelfall jeweils neu zu entscheiden.

Üblicherweise findet die Validierung eines Registrierungsverfahrens hinsichtlich Präzision und Genauigkeit auf mehreren unterschiedlichen Ebenen statt:

1. Zunächst erlauben **Teststudien mit synthetischen Eingabedaten** die vollständige Kontrolle über sämtliche Einflußgrößen eines Systems. Bilddaten können simuliert und beliebig manipuliert werden. Außerdem sind im Falle synthetischer Testdaten die korrekten Transformationsparameter i.a. exakt bekannt. Das Restfehlerverhalten der Registrierung läßt sich daher ideal beurteilen.
2. **Phantomstudien** basieren auf speziell angefertigten Modellkörpern, die bei Anwendung bildgebender Verfahren zu weitgehend realistischem Bildmaterial führen. Die gesuchten Transformationsparameter sind dabei nicht mehr exakt bekannt. Sie lassen sich aber oftmals mit ausreichend hoher Genauigkeit indirekt ermitteln – z.B. durch das Einbringen von Marken in das Phantom.
3. Im Rahmen einer **vorklinischen Anwendung** wird das Verfahren entweder an Leichenpräparaten oder bereits an freiwilligen Testpersonen erprobt. Wie bei einer Phantomstudie ist auch hier das An- bzw. Einbringen von zusätzlichen Marken möglich, um die gesuchte Transformation abschätzen zu können.
4. Das letztlich angestrebte Ziel ist die **klinische Anwendung** des Verfahrens. In einer ersten Erprobungsphase kann dabei jedoch ein alternatives Registrierungsverfahren parallel verwendet werden, um auch hier eine Validierung zu ermöglichen.

2.3.4 Fazit

In den letzten Jahren ist ein deutlicher Trendwechsel auf dem Forschungsgebiet der bildbasierten Operationsunterstützung festzustellen: weg von Registrierungsverfahren, die das Einbringen künstlicher Marken erfordern und hin zu Verfahren, die ausschließlich auf patientenspezifischen Bildinformationen basieren. Interessanterweise zählen jedoch die meisten der bereits in klinischer Anwendung befindlichen Systeme nachwievor zur ersten Kategorie. Dies mag u.a. daran liegen, daß sehr viele der veröffentlichten Verfahren bisher nicht über die Phase einer Phantomstudie hinaus gekommen sind. Betrachtet man speziell die Literatur zur fluoroskopiebasierten Registrierung eines Patienten – das zentrale Themengebiet

dieser Dissertation –, so fällt auf, daß die Anzahl der Veröffentlichungen hierzu zwar stetig zunimmt, im Vergleich zu anderen Modalitäten aber weit zurückfällt. Angesichts der weiten Verbreitung dieses Standardgerätes verwundert dies zunächst, kann aber damit erklärt werden, daß erst in den letzten Jahren zunehmend Verfahren zur Sensorkalibrierung eines Fluoroskops bekannt wurden.

Vor diesem Hintergrund besteht das zentrale Anliegen dieser Dissertation darin, den Nachweis zu führen, daß eine hochpräzise, intraoperative Patientenregistrierung mit Hilfe eines kalibrierten Fluoroskops auch ohne Verwendung künstlicher Marken möglich ist – selbst dann, wenn zunächst keine patientenspezifischen 3D-Informationen vorliegen.

Kapitel 3

Systemüberblick

In diesem Kapitel werden drei medizinische Hardware-Komponenten vorgestellt, die im Kontext der vorliegenden Arbeit eine zentrale Rolle spielen (Abb. 3.1).

3.1 Das Röntgenbildverstärker-System (Fluoroskop)

Ein Fluoroskop ist ein mobiles Röntgengerät, bestehend aus einer Röntgenquelle sowie einer Detektoreinheit, dem Bildverstärker. Die relative Anordnung von Quelle und Detektor ist starr durch einen C-förmig gekrümmten Metallrahmen vorgegeben. Bezogen auf den frei verfahrbaren Sockel des gesamten Gerätes besitzt dieser Rahmen jedoch mehrere Freiheitsgrade. Er ist höhenverstellbar sowie bzgl. dreier Raumachsen drehbar (Abb. 3.2). Die Detektoreinheit hat den Zweck, einfallende Röntgenstrahlung sofort in ein verstärktes Videosignal umzuwandeln, welches anschließend auf einem Monitor betrachtet oder durch einen Rechner weiterverarbeitet werden kann. In dieser Bilddarstellung liegt der wesentliche Vorteil eines Fluoroscops im Vergleich zu herkömmlicher Röntgenaufnahmetechnik durch Belichtung von Photomaterialien: Das Bild ist jeweils sofort sichtbar. Eine manuelle Zwischenbearbeitung sowie Filmverbrauchsmaterial ist nicht nötig. Desweiteren ermöglicht eine kontinuierlich aktivierte Röntgenquelle sogar die Darstellung bewegter Bilder. Die direkte Bildumwandlung besitzt auf der anderen Seite jedoch auch Nachteile. Die Auflösung der Videobilder ist deutlich geringer. Entsprechendes gilt für den resultierenden Bildkontrast.

Angelehnt an das C-förmige Design, wird ein mobiles Röntgenbildverstärkersystem auch oft vereinfacht als *C-Bogen* bezeichnet¹. Durch die C-Form wird es möglich, das Gerät an einen Operationstisch heranzufahren, so daß sich die Röntgenquelle z.B. unterhalb und der Detektor entsprechend oberhalb des Tisches befindet. Liegt nun ein Patient auf dem Operationstisch, so kann dessen Anatomie entsprechend *durchleuchtet* werden. Die erwähnten Freiheitsgrade des Gerätes erlauben dabei eine sehr freizügige Wahl der Aufnahmeorientierungen. Ein Chirurg kann dadurch z.B. die relative Lage eines Instrumentes oder Implantats abschätzen.

Insbesondere bei orthopädischen Eingriffen ist ein mobiles Fluoroskop heutzutage intraoperativ als Bildgebungsmodalität zumeist erste Wahl. Allerdings dienen die damit aufgenommenen Röntgenbilder

¹Im weiteren Verlauf der Arbeit werden die Bezeichnungen Fluoroskop, Röntgenbildverstärker-System und C-Bogen synonym verwendet. Dasselbe gilt für die Begriffe Röntgenbild, -aufnahme und Fluoroskopie-Aufnahme.

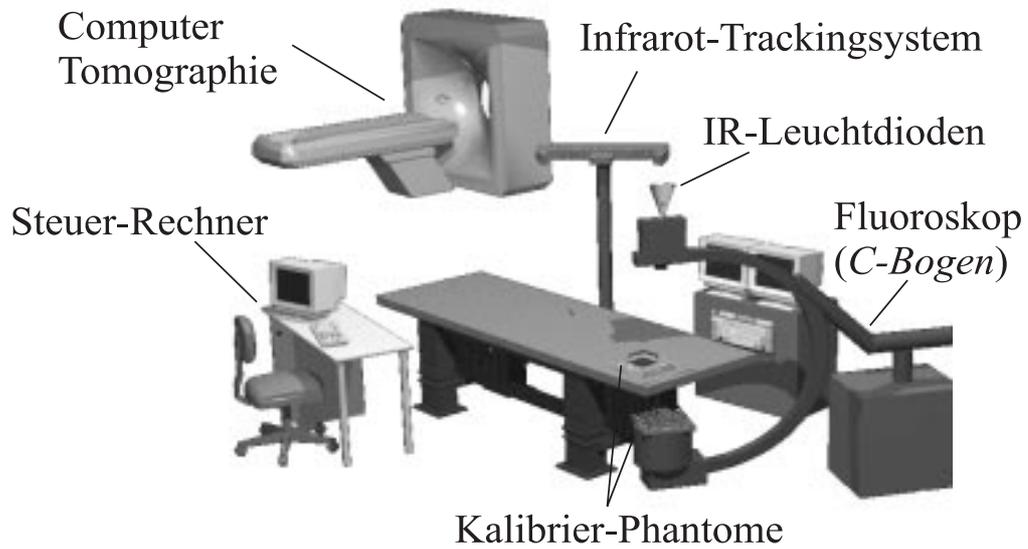


Abbildung 3.1: Operationsszenario: Computer-Tomographie (präoperativ), kalibriertes Fluoroskop und Infrarot-Trackingsystem (intraoperativ)

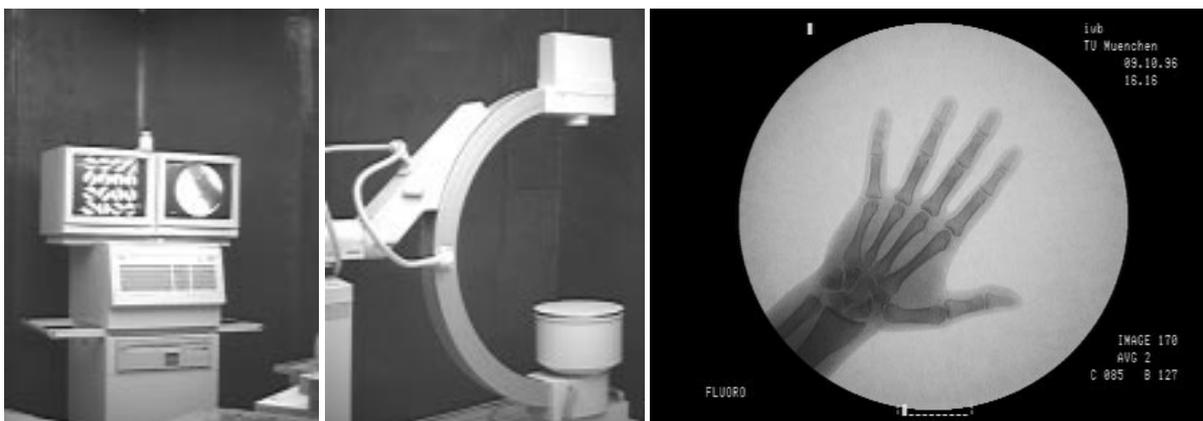


Abbildung 3.2: Flexibles intraoperatives Röntgen mit einem Röntgenbildverstärker-System (Fluoroskop)

in vielen Fällen lediglich einer qualitativen Beurteilung der Situation durch den Chirurgen. Eine quantitative Auswertung – z.B. die präzise Vermessung metrischer Abstände – findet nur selten statt. Ein Grund hierfür ist, daß Fluoroskopie-Aufnahmen signifikante geometrische Verzeichnungseffekte aufweisen, die eine direkte präzise quantitative Interpretation des Bildinhaltes unmöglich machen. Diese Verzeichnungseffekte haben vielfältige Ursachen, z.B. bewirkt die leicht gekrümmte Detektoroberfläche typischerweise eine kissenförmige Verzeichnung des Bildinhaltes. Desweiteren beeinflußt das Magnetfeld der Erde je nach Orientierung des C-Bogens die Flugbahn der Elektronen im Inneren des Bildverstärkers mehr oder weniger stark. Außerdem variiert die zentrierte Ausrichtung des Röntgenstrahlenkegels – ebenfalls in Abhängigkeit von der C-Bogen-Orientierung –, da sich aufgrund des enormen Gewichtes der Bildverstärker-Einheit der C-Bogen geringfügig verwinden kann.

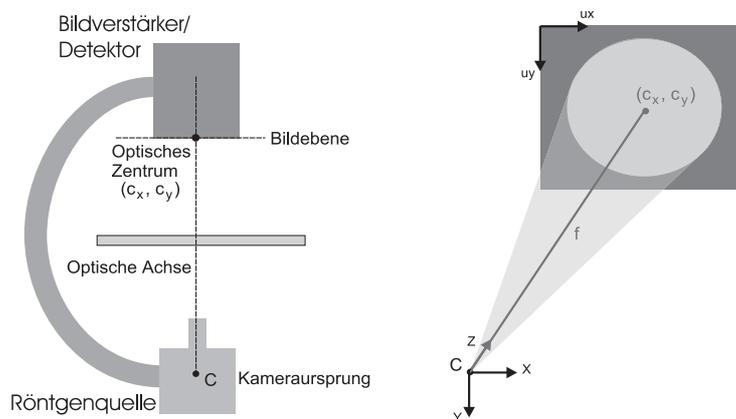


Abbildung 3.3: Idealisierte Kamera-Modellierung eines Fluoroscops

Diese Effekte lassen sich alle durch eine geeignete Sensor-Kalibrierung in den Griff bekommen. Die Behandlung dieser Verzeichnungs- und Kalibrierproblematik ist jedoch nicht Bestandteil der vorliegenden Arbeit. Stattdessen wird in den nachfolgenden Kapiteln jeweils eine Kalibrierung des verwendeten C-Bogen-Kamerasystems vorausgesetzt. Jede für die Patientenlokalisierung herangezogene Fluoroskopieaufnahme liegt bereits entzerrt vor und wird mit einem Parameterdatensatz assoziiert, der ihr ein eindeutiges perspektivisches Kameramodell zuordnet. Die Kamera-Modellierung unterteilt sich dabei in externe und interne Parameter. Die externen Parameter legen die Position und Orientierung der Röntgenquelle (Kamerarsprung) relativ zu einem fixen Bezugskoordinatensystem fest. Die internen Parameter hingegen modellieren die perspektivischen Eigenschaften des Kameramodells durch Festlegung von Brennweite und optischem Nullpunkt (Abb. 3.3).

Neue Verfahren zur Kalibrierung eines Röntgenbildverstärker-Systems sind zentraler Inhalt einer verwandten, zeitgleich entstandenen Dissertation [Bra00a]. Alle im Rahmen der vorliegenden Arbeit verarbeiteten Testdatensätze, die reale Fluoroskopie-Aufnahmen umfassen², entstanden unter Verwendung von Fluoroskopie-Geräten, die mit den dort beschriebenen Verfahren kalibriert wurden. Aus diesem Grund bietet Anhang B einen knappen Überblick über Ansatz und Vorgehensweise der Fluoroskopie-Kalibrierung. Dabei stehen zwei Aspekte im Vordergrund, die wesentlichen Einfluß auf die hier neu entwickelten Verfahren zur Patientenlokalisierung besitzen: die erreichbare Präzision der Kameramodellie-

²siehe z.B. Abschnitte 5.2.2, 5.5, 6.6 und 7

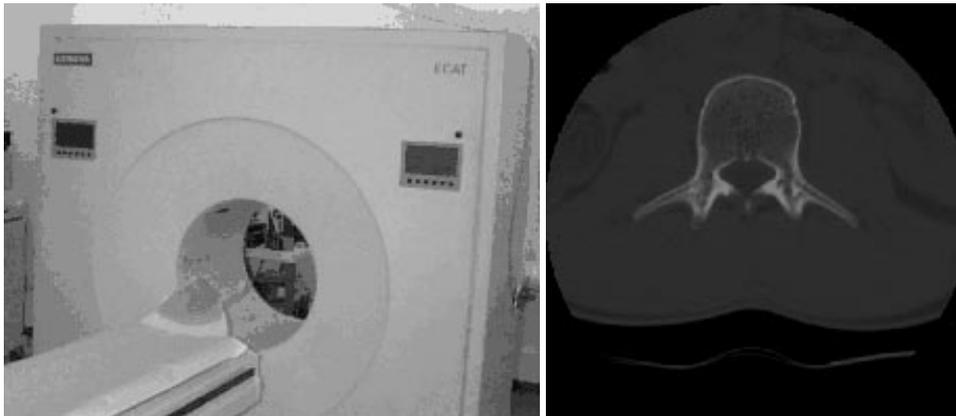


Abbildung 3.4: Dreidimensionale Bildgewinnung von körperinneren Strukturen durch Computer-Tomographie (CT)

zung sowie evtl. kalibrierungsbedingte Nebenwirkungen, z.B. verminderte Bildqualität, Einschränkung der möglichen Aufnahmerichtungen u.ä.

3.2 Computer-Tomographie

Computer-Tomographie, kurz CT, ist neben der vorgestellten Fluoroskopie ein weiteres Standard-Verfahren der Radiologischen Diagnostik. Es handelt sich hierbei um eine bildgebende Technik zur Darstellung dünner Querschnittsschichten eines dreidimensionalen Körpers. Eine einzelne solche Schicht wird durch ein zweidimensionales Pixel-Bild repräsentiert, wobei ein Pixel eine konstante physikalische Ausdehnung besitzt (z.B. $0.4\text{mm} \times 0.4\text{mm}$). Da jeder Schicht jedoch auch eine gewisse Dicke zugeordnet wird (z.B. 2mm), ergibt sich durch Aneinanderreihung vieler Einzelschichten schließlich ein Volumen. Dieses besteht aus einem dreidimensionalen Array rechteckiger Volumenelemente (im Beispiel der Größe $0.4\text{mm} \times 0.4\text{mm} \times 2\text{mm}$). Dieses Volumen wird im folgenden auch als CT-Datensatz bezeichnet.

In einem Computer-Tomographen wird eine Röntgenquelle auf einer Kreisbahn um den betrachteten Körper herum bewegt. Auf der jeweils gegenüber der aktuellen Position der Quelle liegenden Seite der Kreisbahn wird die abgeschwächte Strahlungsintensität durch entsprechende Detektoren gemessen und gespeichert. Aus allen Detektormessungen zusammen läßt sich anschließend ein gewebespezifischer Schwächungswert für jeden einzelnen Pixel des Schichtbildes rekonstruieren. Hierfür sind komplexe mathematische Verfahren wie z.B. direkte Rückprojektion, Algebraische Rekonstruktion oder diverse Fourier-Techniken nötig [Zon80].

Die berechneten Schwächungswerte werden standardmäßig derart auf einen ganzzahligen Bereich skaliert, so daß Luft einen Wert von -1000 , Wasser einen Wert von 0 und Knochen in etwa einen Wert von 1000 erhält. Diese Normierung bzgl. des linearen Abschwächungskoeffizienten von Wasser μ_{Wasser} wird als *Hounsfield-Skala* bezeichnet, einzelne Werte entsprechend als Hounsfield-Werte bzw. -Units (HU). Die Umrechnung eines Abschwächungskoeffizienten μ_x geschieht durch folgende Formel [Goe88,

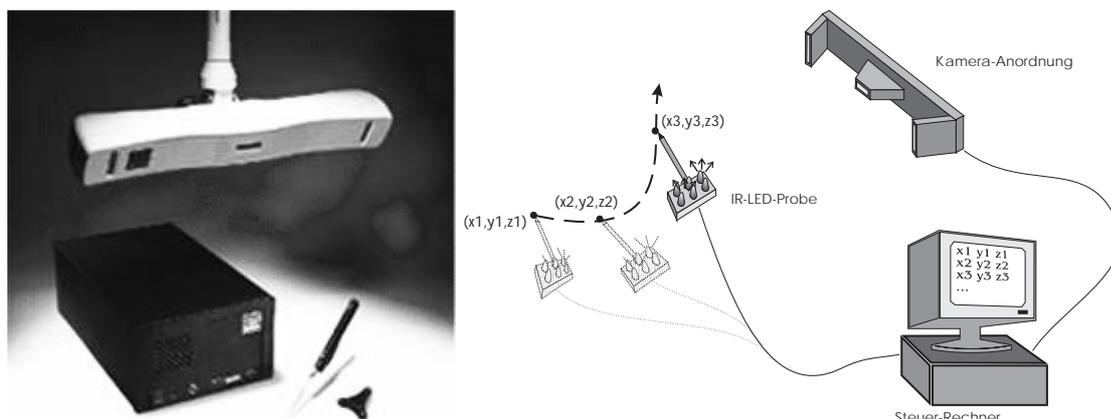


Abbildung 3.5: Ein aktives optisches 6D-Trackingsystem erlaubt die Echtzeit-Bestimmung räumlicher Positionen und Orientierungen von Objekten, die zuvor mit Infrarot-Leuchtdioden präpariert wurden. Die linke Aufnahme zeigt das kommerzielle FP5000-System der Firma *Image Guided Technologies (IGT)*.

Kre80]:

$$\mathbf{HU}_x = 1000 \frac{\mu_x - \mu_{Wasser}}{\mu_{Wasser}} \quad (3.1)$$

Computer-Tomographie spielt im Rahmen der vorliegenden Arbeit an zwei Stellen eine wesentliche Rolle. In Kapitel 5 wird ein neues Registrierungsverfahren vorgestellt, das aus einem vorsegmentierten patientenspezifischen CT-Datensatz künstliche Röntgenbilder generiert, um diese mit realen Fluoroskopie-Aufnahmen des Patienten zu vergleichen. Ohne den grundlegenden physikalischen Zusammenhang zwischen Fluoroskopie/Röntgen und Computer-Tomographie wäre dieses Verfahren überhaupt nicht möglich. Das in Kapitel 6 vorgeschlagene Verfahren zur formvariablen Registrierung hingegen benötigt explizit kein patientenspezifisches CT. Die dabei für das Formvariationstraining genutzten 3D-Modelle können jedoch z.B. durch Segmentation mehrerer bereits vorhandener CT-Datensätze unterschiedlicher Patienten gewonnen werden.

3.3 Echtzeit-Tracking-System

Wenngleich Tracking-Systeme heutzutage noch nicht zur klinischen Standardaustattung zählen, so sind sie zumindest dort, wo bereits operationsunterstützte Computersysteme innerhalb eines Operationssaals eingesetzt werden, nicht mehr wegzudenken. Derartige Systeme bieten die Möglichkeit, Objekte im Operationsfeld bzgl. eines Referenz-Koordinatensystems präzise zu lokalisieren und ggf. die räumliche Bewegung der Objekte in Echtzeit zu verfolgen. Die Echtzeitfähigkeit wird in der Regel dadurch erreicht, daß diese Objekte mit speziellen Marken versehen werden, deren Position über eine entsprechende Sensor-Anordnung eindeutig erfaßt werden kann.

In den letzten Jahren wurde im klinischen Umfeld eine Vielzahl an Tracking-Systemen entwickelt und getestet, die auf höchst unterschiedlicher Sensor-Technologie basieren: z.B. mechanisch (*ViewingWandTM* [RMo95]), magnetisch (*Flock-of-BirdsTM* [GFo95]), akustisch (*PegasusTM* [WB97]) und optisch (*FlashpointTM*, *OptotrakTM* [Cha98]). Am weitesten verbreitet und von mehreren

Anbietern kommerziell erhältlich sind jedoch in erster Linie optische Systeme. Diese erreichen je nach Qualität Lokalisationsgenauigkeiten von ca. 0.5mm [Cha98].

Abb. 3.5 zeigt ein solches optisches Trackingsystem (links, *FlashpointTM*) und verdeutlicht das Prinzip (rechts): Eine feste Kamera-Anordnung, bestehend aus drei Infrarot-CCD-Kameras, *beobachtet* Objekte in ihrem Sichtbereich, die jeweils mit mehreren Infrarotlicht emittierenden Leuchtdioden (IR-LEDs) versehen sind. Korrespondenzprobleme werden vermieden, indem ein Steuerrechner die einzelnen IR-LEDs in sehr schneller Folge nacheinander aktiviert. Auf diese Weise ist dem System zu jedem Zeitpunkt exakt bekannt, welche aktivierte IR-LED die drei Kameras gerade sehen. Wird eine der IR-LEDs von jeder einzelnen Kamera erfaßt, so kann durch einfache projektionsgeometrische Berechnungen die 3D-Position der LED relativ zu der Kamera-Anordnung berechnet werden. Wird ein Objekt mit mehreren IR-LEDs versehen, die gegenseitig starr angeordnet sind, so läßt sich zusätzlich zu der 3D-Position eine relative 3D-Orientierung berechnen. Daher spricht man i.a. auch von 6D-Tracking-Systemen.

Die Anwendungsmöglichkeiten eines solchen Systems sind vielfältig. Z.B. läßt sich die Position und Ausrichtung chirurgischer Instrumente in Echtzeit verfolgen. Anatomische Landmarken eines Patienten (z.B. Knochenvorsprünge) können abgetastet werden. Außerdem ist es möglich, für bildgebende Sensoren, z.B. ein Ultraschallgerät, deren relative Position und Orientierung zum Zeitpunkt einer Bildaufnahme zu ermitteln.

Im Anwendungskontext dieser Arbeit dient ein optisches Navigationssystem mehreren Zielen:

- Vereinbarung eines statischen Referenz-Koordinatensystems für Fluoroskopie-Aufnahmerichtung, Patientenlage und Instrumentpositionen,
- Erfassung der Fluoroskop-Orientierung (Sensor-Lage zum Zeitpunkt einer Röntgenaufnahme),
- Tracking nicht fixierter Knochen nach erfolgter Registrierung (Lageerkennung),
- Navigation von Instrumenten,
- Experimentelle Genauigkeitsuntersuchungen.

Kapitel 4

Grundlagen der Registrierung

4.1 Problemstellung

Die Kenntnis der Beziehung zwischen medizinischen Bilddaten einerseits und der räumlichen Position chirurgischer Instrumente andererseits ist die entscheidende Grundlage für den Einsatz eines jeden IGS-Systems. Die Vorgehensweise zur Bestimmung dieses mathematischen Zusammenhangs zwischen anatomischem Bildmaterial oder einem davon abgeleiteten Modell und z.B. einem Echtzeit-Tracking-System, das die Kontrolle von Instrumenten erlaubt, stellt eine Art Objekt-Lokalisation (vgl. z.B. [Wun97]) dar. Das Objekt ergibt sich dabei z.B. durch Segmentation aus präoperativen CT-Schichtaufnahmen. Die Lokalisation hingegen erfolgt schließlich unter Einbeziehung intraoperativen Bildmaterials. Im medizinischen Umfeld wird eine derartige Lokalisation stets als *Registrierung* bezeichnet.

In [Lav96b] wird ein allgemeingültiger dreistufiger Rahmen für beliebige Registrierungsverfahren vorgeschlagen:

1. Definition von Koordinatensystemen sowie von Relationen bzw. Transformationen dieser Systeme untereinander
2. Ableitung von Referenzmerkmalen sowie die Festlegung eines Ähnlichkeits- bzw. eines Fehlermaßes zwischen den extrahierten Merkmalen
3. Optimierung des Ähnlichkeits- (Maximierung) bzw. Fehlermaßes (Minimierung)

Dabei spiegelt Punkt 1 die vorgegebene Situation wieder (welche prä- und welche intraoperativen Modalitäten sind zu berücksichtigen?). Die Punkte 2 und 3 hingegen legen das algorithmische Verfahren zur Bestimmung der gesuchten Transformation fest.

Die vorliegende Arbeit folgt bei der Beschreibung der diversen Registrierungstechniken konzeptionell diesem Schema.

Definition von Koordinatensystemen und Transformationen

Ziel jeglicher Registrierung ist es, die Transformation \mathcal{T} zwischen KS_{PreOp} (z.B. 3D-Modell, CT-Daten o.ä.) und KS_{Sensor} zu schätzen und diese letztendlich in eine Transformation zwischen KS_{PreOp} und KS_{Ref} umzurechnen. Dabei wird \mathcal{T} gemeinhin als resultierende Lage des betrachteten Objektes, z.B. des Femur-Knochens, bezeichnet.

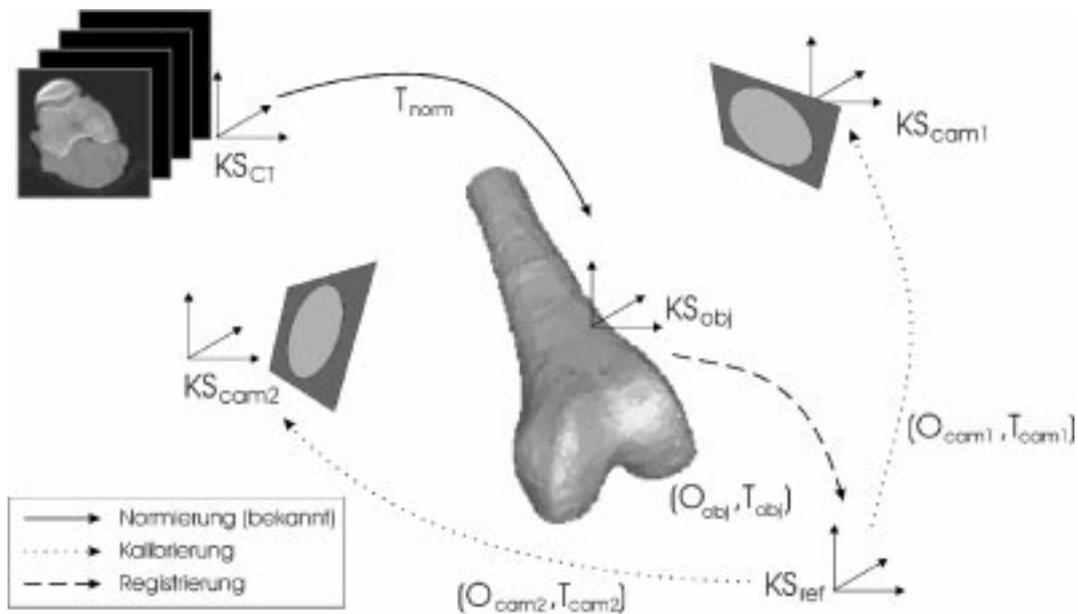


Abbildung 4.1: Definition von Koordinatensystemen und Transformationen

Im vorliegenden Kapitel repräsentiert diese Transformation \mathcal{T} stets eine Relation zwischen *denselben, unveränderten* Knochenstrukturen, die lediglich von unterschiedlichen Sensoren (CT einerseits, Fluoroskopie/Röntgen andererseits) und i.a. zu unterschiedlichen Zeitpunkten (präoperativ bzw. intraoperativ) erfaßt wurden. \mathcal{T} stellt in diesem Fall eine starre Transformation dar. Starre räumliche Transformationen lassen sich durch exakt sechs Parameter beschreiben, wobei die ersten drei Parameter t_x , t_y und t_z die Koordinaten eines Translationsvektors $\mathbf{t} = (t_x, t_y, t_z)^T$ definieren. Zusätzlich ermöglicht eine 3×3 -Rotationsmatrix \mathbf{O} beliebige räumliche Drehungen. Diese Matrix hängt dabei lediglich von drei weiteren Parametern ab.

Unter Verwendung dieser Notation läßt sich eine rigide Transformation (\mathbf{O}, \mathbf{t}) zwischen zwei Koordinatensystemen \mathbf{A} und \mathbf{B} folgendermaßen interpretieren: Ein Punkt \mathbf{P}_A , dessen Koordinaten bzgl. \mathbf{A} definiert sind, kann in einen Punkt \mathbf{P}_B , dessen Koordinaten bzgl. \mathbf{B} definiert sind, transformiert werden, indem folgende Rechenvorschrift angewandt wird:

$$P_B = \mathbf{O}P_A + \mathbf{t}$$

Faktisch verändert der Punkt dabei seine räumliche Position – z.B. bzgl. eines beliebigen dritten Koordinatensystems – nicht. Durch den Systemwechsel von \mathbf{A} nach \mathbf{B} ändert sich lediglich die *Blickrichtung*, aus der dieser Punkt gesehen wird.

Die drei Parameter des Translationsvektors \mathbf{t} sind bereits eindeutig interpretierbar. Aufgrund der redundanten Darstellung der drei rotatorischen Freiheitsgrade der rigiden Registrierung in Form einer 3×3 -Matrix gilt dies aber nicht für die räumliche Orientierung. Hier gibt es eine Vielzahl an verschiedenen Verfahren zur Ableitung von drei Parametern aus einer Rotationsmatrix bzw. deren Konstruktion, ausgehend von drei gegebenen Parametern.

Eine mögliche und häufig genutzte Darstellung ist die sogenannte Euler-Winkel-Darstellung

(α, β, γ) . Zwischen dieser Darstellung und der Rotations- bzw. Orientierungsmatrix besteht folgender Zusammenhang:

$$\mathbf{O} = \begin{pmatrix} \cos \gamma \cos \beta & -\sin \gamma \cos \beta & \sin \beta \\ \cos \gamma \sin \beta \sin \alpha + \sin \gamma \cos \alpha & -\sin \gamma \sin \beta \sin \alpha + \cos \gamma \cos \alpha & -\cos \beta \sin \alpha \\ -\cos \gamma \sin \beta \cos \alpha + \sin \gamma \sin \alpha & \sin \gamma \sin \beta \cos \alpha + \cos \gamma \sin \alpha & \cos \beta \cos \alpha \end{pmatrix}$$

Eulerwinkel lassen sich als Hintereinanderausführung dreier Drehungen um die x -, die y - und schließlich um die z -Achse interpretieren:

$$\mathbf{O} = \mathbf{O}_z(\gamma)\mathbf{O}_y(\beta)\mathbf{O}_x(\alpha)$$

Diese Darstellung ist minimal. Ihre Verwendung bringt aber zwei schwerwiegende Probleme mit sich: Zum einen sind die einzelnen Matrix-Koeffizienten hochgrad nicht-linear, zum anderen ergeben sich für $\beta = -\pi/2$ Singularitäten, d.h. die Darstellung ist an diesen Stellen nicht differenzierbar [Lav96b]. Eine Darstellung, die die beiden genannten Nachteile nicht besitzt, sind sogenannte Einheits-Quaternionen [FH86, Hor87]. Dabei handelt es sich um vier-dimensionale Vektoren $\mathbf{q} = (q_0 \ q_x \ q_y \ q_z)^T$, wobei die folgenden Nebenbedingungen erfüllt sein müssen:

$$q_0 \geq 0 \tag{4.1}$$

und

$$q_0^2 + q_x^2 + q_y^2 + q_z^2 = 1 \tag{4.2}$$

Unter diesen beiden Voraussetzungen existiert zwischen einer Rotationsmatrix \mathbf{O} und einem Einheits-Quaternion \mathbf{q} ebenfalls ein eindeutiger Zusammenhang:

$$\mathbf{O} = \begin{pmatrix} q_0^2 + q_x^2 - q_y^2 - q_z^2 & 2(q_x q_y - q_0 q_z) & 2(q_x q_z + q_0 q_y) \\ 2(q_x q_y + q_0 q_z) & q_0^2 - q_x^2 + q_y^2 - q_z^2 & 2(q_y q_z - q_0 q_x) \\ 2(q_x q_z - q_0 q_y) & 2(q_y q_z + q_0 q_x) & q_0^2 - q_x^2 - q_y^2 + q_z^2 \end{pmatrix}$$

Es fällt auf, daß sich die einzelnen Elemente der Matrix auf einfache Art und Weise ableiten lassen. Diese Darstellung verhindert zudem das Auftreten von Singularitäten. Leider ist sie aber nicht minimal (4 Parameter für drei Freiheitsgrade), wodurch sich in Zusammenhang mit den einzuhaltenden Nebenbedingungen (Gl. 4.1 u. 4.2) Probleme beim Einsatz üblicher Minimierungsverfahren ergeben. Derartige Verfahren variieren intern i.a. alle übergebenen, freien Parameter unabhängig voneinander, ohne dabei auf eventuelle Randbedingungen zu achten. Die Verwendung von Einheits-Quaternionen zur Rotationsrepräsentation macht es daher erforderlich, nach jedem Minimierungsschritt die resultierenden vier Parameter q_0, q_x, q_y und q_z entsprechend der Nebenbedingungen zu normieren.

Diese Probleme im Zusammenhang mit Eulerwinkeln bzw. Einheits-Quaternionen lassen sich umgehen, falls stattdessen – [Lav96b] folgend – eine Repräsentation verwendet wird, die erstens minimal ist und zweitens keine Singularitäten aufweist. Die Orientierung wird hierzu durch Definition einer einzelnen, beliebig orientierten räumlichen Rotationsachse beschrieben, die durch den Vektor $\mathbf{r} = (r_x, r_y, r_z)$ repräsentiert wird, wobei die Norm dieses Vektors den stets positiven Drehwinkel η um diese Achse festlegt [Lav96b].

Diese Darstellung läßt sich aus einer Rotationsmatrix \mathbf{O} direkt bzw. aus ihrer Quaternion-Repräsentation wie folgt ableiten¹:

¹Die Spur einer Matrix bezeichnet die Summe ihrer Diagonalelemente, d.h. $\text{Spur}(\mathbf{O}) = \sum o_{ii}$.

$$\cos \eta = \frac{1}{2}(\text{Spur}(\mathbf{O}) - 1) \quad (4.3)$$

$$\mathbf{r}' = \begin{pmatrix} r'_x \\ r'_y \\ r'_z \end{pmatrix} = \begin{pmatrix} q_x / \left(q_x^2 + q_y^2 + q_z^2 \right) \\ q_y / \left(q_x^2 + q_y^2 + q_z^2 \right) \\ q_z / \left(q_x^2 + q_y^2 + q_z^2 \right) \end{pmatrix} \quad (4.4)$$

$$\mathbf{r} = \eta \frac{\mathbf{r}'}{|\mathbf{r}'|} \quad (4.5)$$

In dieser Arbeit findet jede der drei Repräsentationen Anwendung. Die konkrete Auswahl erfolgt dabei je nach Zielsetzung des übergeordneten Verfahrens bzw. den Einschränkungen spezieller Optimierungsalgorithmen. In den entsprechenden Abschnitten wird jeweils auf die gewählte Orientierungs-Repräsentation explizit hingewiesen.

Ableitung von Referenzmerkmalen

Nachdem entsprechend der obigen Ausführungen Koordinatensysteme festgelegt wurden und insbesondere zwischen KS_{PreOp} und KS_{Sensor} eine Transformation definiert wurde, ist es nötig, in beiden Systemen korrespondierende Referenzmerkmale zu extrahieren.

Die Art der ausgewählten Referenzmerkmale und Ähnlichkeits- bzw. Fehlerfunktionen (siehe nächster Abschnitt) stellen i.a. das *Herzstück* eines Registrierungsverfahrens dar [Lav96b]. Beide eignen sich daher ideal als Schlüsselkriterien zur Klassifizierung von Registrierungsverfahren. Entsprechend lassen sich die im Rahmen dieser Arbeit entwickelten Registrierungstechniken einteilen. Sie verwenden die nachfolgend aufgeführten Referenzmerkmale:

- In Kapitel 5 wird auf die explizite Ableitung einer Menge von Referenzmerkmalen in den beiden Koordinatensystemen verzichtet. Stattdessen dienen die jeweiligen Bilddaten selbst implizit als Merkmale: präoperativ CT-Abschwächungswerte und intraoperativ die Intensitäten der Bildpunkte der Fluoroskopieaufnahmen.
- Kapitel 6 definiert präoperativ die Oberflächenpunkte eines formvariablen, d.h. deformierbaren Knochenmodells, das nicht notwendigerweise patientenspezifisch sein muß, als Referenzmerkmale (KS_{PreOp}). Intraoperativ dienen Konturpunkte von Knochensilhouetten, die in den Fluoroskopieaufnahmen segmentiert wurden, als Merkmale (KS_{Sensor}).

Auswahl und Optimierung des gewählten Maßes

Stehen die in beiden Koordinatensystemen definierten Referenzmerkmale einmal fest, so besteht – unter der Voraussetzung bekannter Korrespondenzen der Merkmale untereinander – der nächste Schritt darin, eine Funktion zu definieren, die ein Fehler- oder Ähnlichkeitsmaß zwischen den jeweiligen Merkmalen repräsentiert. Wie auch bei der Auswahl der verwendeten Referenzmerkmale gibt es für diese Funktion eine Reihe möglicher Definitionen (vgl. z.B. [Lav96b]).

Die Begriffe Referenzmerkmal und Optimierungsmaß erlauben die nachfolgende allgemeine Definition der *Registrierung*:

Definition 1 (Registrierung)

Zwei Koordinatensysteme KS_{PreOp} und KS_{Sensor} gelten als gegenseitig registriert, falls eine Ähnlichkeitsfunktion (bzw. Fehlerfunktion), die auf einer Menge korrespondierender Merkmale in beiden Systemen basiert, maximal (bzw. minimal) wird. Der Vorgang der Optimierung (Maximierung bzw. Minimierung) selbst wird als Registrierung bezeichnet.

In dieser Arbeit kommen ausschließlich Fehlerfunktionen – keine Ähnlichkeitsmaße – zum Einsatz. Die jeweilige Auswahl der unterschiedlichen Funktionen sowie deren Optimierung wird in den entsprechenden Abschnitten (in erster Linie 5.3.2 und 6.5.2) ausführlich vorgestellt und motiviert.

4.2 Initialisierung

Viele der in Kapitel 2 diskutierten Registrierungsmethoden basieren auf der Optimierung nicht trivialer, hochdimensionaler Fehler- oder Ähnlichkeitsfunktionen. Da derartige Funktionen neben dem in der Regel gesuchten globalen Optimum meist eine Vielzahl an lokalen Optima besitzen, ergibt sich bei einigen Optimierungsalgorithmen das Problem, daß die berechnete Lösung von Startbedingungen abhängig ist. In den weiteren Ausführungen dieses Abschnitts wird Optimum ohne Beschränkung der Allgemeinheit stets mit Minimum gleichgesetzt.

Solche unerwünschten lokalen Minima, die eine Registrierung scheitern oder zumindest unpräzise werden lassen, können auf der Basis experimenteller Erfahrung im klinischen Umfeld qualitativ in zwei Klassen unterteilt werden [Lav96b]:

Die erste Art tritt in der Regel in hoher Häufigkeit und jeweils sehr nahe am globalen Minimum auf. Solche lokalen Minima zeichnen sich zudem dadurch aus, daß sie – global gesehen – oft nur geringfügig höher liegen als das gesuchte globale Minimum. Die Erkennung und Behandlung solcher Minima gestaltet sich i.a. sehr schwierig. Da die Ursache für ihr Auftreten aber fast immer in einer schlecht gewählten Diskretisierung der zugrundeliegenden Daten – z.B. eines Oberflächenmodells – zu suchen ist, lassen sie sich mit entsprechender Sorgfalt oft vermeiden.

Die zweite Art umfaßt Minima, die in starker Abhängigkeit von der Form des zu registrierenden Objekts vereinzelt, isoliert und meist in großer Entfernung zum globalen Minimum auftreten. Solche Minima lassen sich leicht anhand eines signifikant erhöhten Restfehlers erkennen. Auch ihre Vermeidung ist im medizinischen Umfeld oftmals möglich, da in den meisten realistischen Anwendungsfällen *a-priori*-Wissen über die gesuchte Transformation vorhanden oder doch zumindest leicht bestimmbar ist. Solches Vorwissen kann in Form einer geeigneten Initialisierung der Start-Parameter Verwendung finden.

Registrierungsalgorithmen, bei deren zugrundeliegender Optimierungsfunktion mit isolierten lokalen Minima zu rechnen ist, sind daher meist nach einem zweistufigen Ansatz aufgebaut:

1. Bestimme eine approximative Transformation (Initialisierung)
2. Verfeinere bzw. verbessere die gefundene Lösung aus 1 (Optimierung)

Ob ein konkretes Verfahren in einem Minimum der zweiten Kategorie *stecken bleibt*, hängt direkt von dessen Konvergenzradius ab. In der Theorie gibt ein Konvergenzradius an, wie weit die initiale approximative Lösung (Schritt 1) von der globalen Lösung maximal entfernt sein darf, um den Algorithmus (Schritt 2) vor einem *Abgleiten* in ein lokales Minimum zu bewahren. In der Praxis lassen sich solche Parameter-Schranken analytisch aber nur schwer bestimmen, da die Lage und Ausprägung lokaler Minima neben der Art der Fehlerfunktion eben auch sehr stark von der Form des Objektes abhängt. Einen

Ausweg bietet die experimentelle Abschätzung des Konvergenzradius' anhand von Prototypen bestimmter Objektklassen, wie z.B. proximaler/distaler Femur, Lendenwirbel u.s.w. In Abhängigkeit von diesen Erkenntnissen kann dann über die Eignung oder Nicht-Eignung eines Initialisierungsverfahrens entschieden werden, sofern dessen Fehlertoleranzen bekannt sind.

Wenngleich die meisten Registrierungsverfahren konzeptionell nach dem beschriebenen zweistufigen Prinzip organisiert sind, so gibt es durchaus vielversprechende Ansätze, die eine der eigentlichen Registrierung vorgeschaltete Suche nach einer approximativen Lösung überflüssig machen. Ein Beispiel hierfür findet sich in [GBSN97]. Das durch die Autoren vorgeschlagene *restricted surface matching*-Verfahren kombiniert ein herkömmliches Oberflächen-Matching – abgetastete Oberflächenpunkte werden mit einer bekannten Oberfläche zur Deckung gebracht – mit einem Landmarken-basierten Ansatz. Als Landmarken dienen hierbei anatomisch eindeutige, markante Punkte am Knochen, deren intraoperative Erfassung durch Abtastung zumindest approximativ stets möglich ist. Die Korrespondenzen dieser abgetasteten 3D-Hilfspunkte mit entsprechenden 3D-Punkten am Oberflächenmodell sind dabei explizit bekannt, woraus sich bereits eine Transformation berechnen läßt [HS93a, BM92]. Wie die Bezeichnung *Hilfspunkte* bereits andeutet, haben diese Punkte aber keinen direkten Einfluß auf das Registrierungsergebnis. Sie (bzw. die damit definierte Transformation) dienen lediglich dazu, das verwendete Minimierungsverfahren auf die Suche nach Lösungen in der Nähe des globalen Minimums zu beschränken.

Die vorliegende Arbeit folgt jedoch dem weiter oben beschriebenen zweistufigen Ansatz, d.h. der Trennung zwischen Initialisierung und Optimierung. Daher werden nachfolgend mehrere praktikable Initialisierungsansätze für die im Rahmen dieser Arbeit im Vordergrund stehenden Algorithmen zur Registrierung der beiden Modalitäten Fluoroskopie (intraoperativ) und Computer-Tomographie (präoperativ) vorgestellt und diskutiert.

Im Gegensatz zu der eigentlichen Registrierung, d.h. den Optimierungsverfahren, muß dabei letztendlich keine der vorgeschlagenen Initialisierungen die anvisierte Präzision aufweisen. Es genügt bereits eine grobe Transformation zu schätzen, die gewährleistet, daß die Konvergenz der nachfolgenden Optimierung sichergestellt ist.

In den folgenden Abschnitten werden drei unterschiedliche Initialisierungstechniken vorgestellt, die im Rahmen dieser Arbeit erfolgreich eingesetzt wurden. Das erste Verfahren berücksichtigt die besondere Aufnahmesituation innerhalb eines OP-Saals, indem dem OP-Team eine der Aufnahmerichtungen relativ zu dem Patienten grob vorgegeben wird. Die zweite Technik basiert auf der präoperativen Festlegung anatomisch markanter 3D-Punkte, deren korrespondierende Röntgenbildpunkte intraoperativ bestimmt und zugeordnet werden. Schließlich wird ein weiteres Verfahren präsentiert, das eine interaktive Platzierung eines virtuellen Objektmodells vorsieht.

4.2.1 Vorgabe einzelner Aufnahmerichtungen

Die Grundidee dieses ersten Initialisierungsverfahrens ist einfach und wirkungsvoll zugleich. Sie besteht darin, bei der Akquisition der einzelnen Fluoroskopieaufnahmen aus unterschiedlichen Richtungen mindestens eine standardisierte Aufnahmesituation zu berücksichtigen. Ausgehend von dieser Aufnahme können alle für die anschließende Registrierung benötigten Transformationsparameter geschätzt werden.

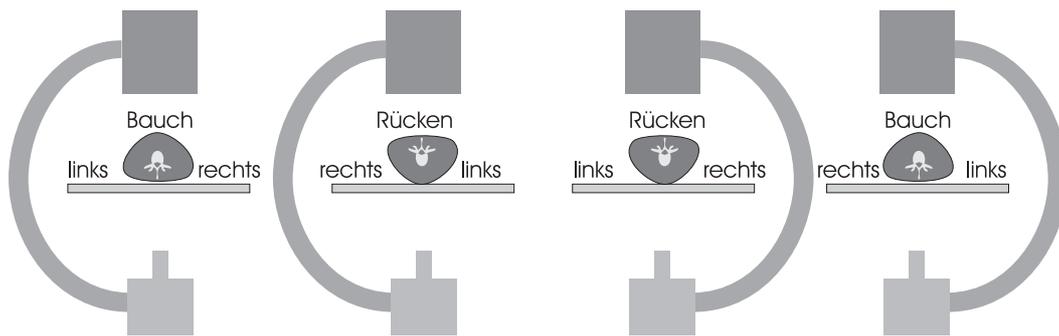


Abbildung 4.2: Vier verschiedene Möglichkeiten für die relative Positionierung von Patient und C-Bogen (Vorgabe orthogonaler Aufnahmeorientungen)

Verfahren

Die entscheidende Voraussetzung für die vorgeschlagene Technik ist, daß die Aufnahmeorientierung überhaupt beeinflußt werden kann. Für die meisten Situationen aus dem Gebiet der *Computer-Vision*, speziell dem Forschungsbereich der Objekt-Erkennung und -Lokalisation, ist dies gerade nicht der Fall. Hier gilt es u.a., Objekte zu erkennen, die beliebige Orientierungen und Positionen einnehmen können – und das oftmals autonom, d.h. sie unterliegen nicht der direkten Kontrolle des Kamerasystems bzw. des Anwenders.

Im krassen Gegensatz hierzu liegt im klinischen Umfeld während einer Operation eine besondere Situation vor. Zum einen liegt der Patient in jedem Fall in einer vordefinierten Art und Weise auf dem OP-Tisch. Dasselbe gilt für die präoperative Erstellung des patientenspezifischen CTs. Zum anderen besitzt das OP-Team natürlich die Möglichkeit, den verwendeten C-Bogen im Rahmen der Gegebenheiten vor Ort weitestgehend frei zu positionieren und orientieren. Von sich aus bevorzugt ein Arzt dabei je nach durchleuchtetem Körperbereich stets bestimmte orthogonale Aufnahmeorientungen: z.B. *lateral*, *medial*, *anterior-posteriori (AP)* u.a., um seine räumliche Interpretation der Bilder nicht unnötig zu erschweren. Im Hinblick auf eine konturbasierte Registrierung zählen genau diese Aufnahmeorientungen interessanterweise in den meisten Fällen zu den wenigen gut zu verarbeitenden Ansichten einer Zielstruktur. Insbesondere für Wirbel gilt, daß die meisten Ansichten aufgrund von Selbstüberlappung bzw. Überlagerung mit benachbarten Knochenstrukturen nur schwer interpretierbar sind.

Es bietet sich an, eine dieser intuitiven Aufnahmeorientungen als *Grundstellung* zu wählen, von der ausgehend die initialen Transformationsparameter geschätzt werden können. Diese *Grundstellung* des Röntgenbildverstärkersystems definiert sich wie folgt: Der Bildverstärker wird erstens parallel über dem OP-Tisch positioniert, so daß die virtuelle Achse zwischen der Röntgenquelle unter dem OP-Tisch und dem Bildverstärker/Detektor den Tisch in etwa senkrecht schneidet. Zweitens wird die Orientierung des Bildverstärkers bzgl. dieser virtuellen Achse derart gewählt, daß die resultierenden Bildachsen des Röntgenbildes möglichst parallel zu den Tischkanten verlaufen. Bzgl. des auf dem OP-Tisch befindlichen Patienten ist die Grundstellung jedoch noch nicht eindeutig, da dieser erstens entweder auf dem Bauch oder dem Rücken liegen kann (je nach Art der geplanten Operation) und sich zweitens der Röntgenbildverstärker wahlweise rechts oder links des OP-Tisches befindet. Daher ist es komfortabel und sinnvoll, anstatt einer einzelnen strikten Vorgabe, eine begrenzte Auswahl an möglichen initialen, zueinander orthogonalen Aufnahmeorientungen bzgl. des Patienten anzubieten. Es kommen insgesamt vier Fälle in

Frage (Abb. 4.2).

Die beschriebene Grundstellung legt zunächst lediglich die drei rotatorischen Parameter fest. Die restlichen drei translatorischen Freiheitsgrade werden durch die folgenden beiden Forderungen bestimmt:

- Der Bildwandler wird parallel zum OP-Tisch so positioniert, daß der Schwerpunkt des Zielobjektes in der Mitte des Bildes zu liegen kommt ($t_x \approx t_y \approx 0$).
- Die Höhe des Bildwandlers wird auf eine Standardhöhe bzgl. des OP-Tisches eingestellt ($t_z = t_{fix}$).

Eine Initialisierung mit der vorgestellten Technik ist einfach zu realisieren. Sie erfordert weniger Fachwissen als vergleichbare Ansätze und es ist insbesondere keine Interaktion am OP-Computer nötig, sieht man von der Auswahl der Grundstellung ab.

In den bisherigen Überlegungen wurden zwei Punkte außer Acht gelassen, deren präoperative Beachtung aber entscheidend für den praktikablen intraoperativen Einsatz der vorgestellten Initialisierungsstrategie ist.

- Die Lagerung des Patienten während der Computer-Tomographie muß standardisiert sein, d.h. er sollte z.B. in derselben Position und Orientierung wie später auf dem Operationstisch *gescannt* werden. Ansonsten schlägt eine Initialisierung der beschriebenen Art intraoperativ fehl oder wird zumindest sehr ungenau.
- Während der Kalibrierung des Röntgenbildverstärkers muß darauf geachtet werden, daß die Ausrichtung der Bildachsen orthogonal auf die physikalische Geometrie des C-Bogens abgestimmt wird. Die Achsen sollen genau dann parallel zu den beiden OP-Tischkanten verlaufen, wenn der C-Bogen so positioniert wurde, daß die durch die C-Form aufgespannte virtuelle Ebene den OP-Tisch senkrecht schneidet und zur kürzeren der beiden Tischkanten parallel verläuft. Durch das Gerät selbst und dessen Videoausgang ist dies nicht von vornherein garantiert. Während der Bild-Entzerrung im Rahmen der Kalibrierung kann dies aber in Form eines rotatorischen Offsets berücksichtigt werden.

Präzision

Wie in den Vorüberlegungen zu diesem Abschnitt bereits erläutert, ist es wichtig, für ein konkretes Initialisierungsverfahren zu wissen, wie nahe die damit geschätzte Transformation der wirklichen Lage der Zielstruktur bereits kommt bzw. wie weit sie bei korrekter Anwendung maximal davon abweicht. Für das vorliegende Verfahren ergaben sich experimentell folgende Erkenntnisse (Abb. 4.3):

- Orientierung:
Die beiden Winkel α und β lassen sich unproblematisch auf die gewünschten Werte von jeweils 0° einstellen. Hierfür ist an handelsüblichen C-Bogen-Systemen an der entsprechenden Achse eine Skala vorhanden. Die Achsen lassen sich also präzise in Nullstellung arretieren. Es ergeben sich minimale Restfehler von maximal 1° bis 2° . Für den dritten Winkel fällt die Unsicherheit größer aus: $d\gamma \leq 5^\circ$. Dies hängt damit zusammen, daß üblicherweise zwar auch hier eine Nullstellung existiert, diese aber aufgrund der Geometrie des Gerätes bzgl. des C-Bogen-Sockels definiert ist.

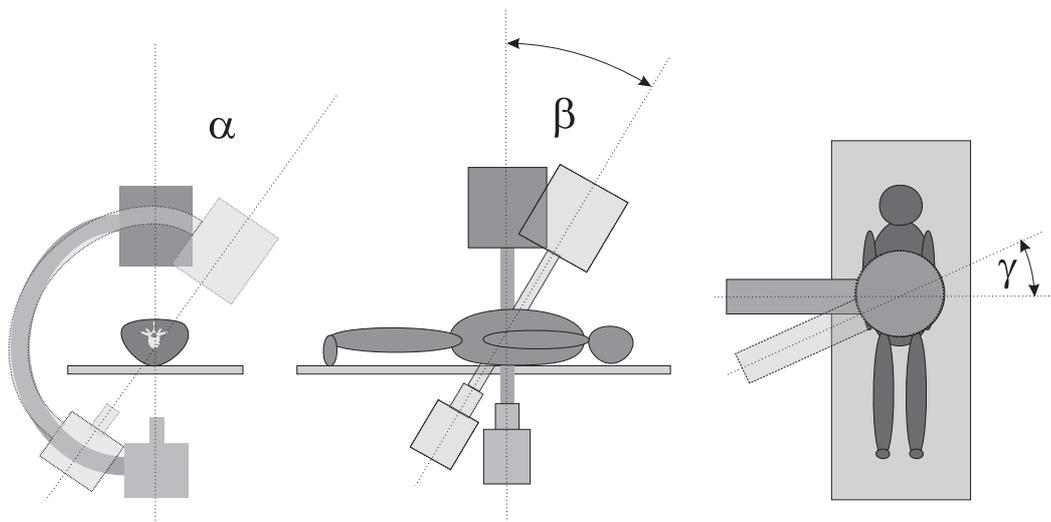


Abbildung 4.3: Definition von Unsicherheitsbereichen bei der Relativ-Positionierung des C-Bogens

Dieser Sockel ist im Raum frei verfahrbar, was dazu führt, daß eine perfekte Stellung oft nur schwer erreichbar ist.

- Translation:

Die Initialisierung der Translation durch entsprechende Positionierung des C-Bogens ist problematischer. Der Grund hierfür ist, daß die Translation bzgl. des Schwerpunktes der betrachteten Teilstruktur im CT definiert ist. Insbesondere im Falle von Femur oder Tibia wird oftmals kein komplettes CT erstellt. Zumindest wird in der Regel kein vollständiges Oberflächenmodell des Knochens rekonstruiert, da dieser mit Hilfe des C-Bogens sowieso nie komplett in einer Aufnahme erfasst werden kann. In diesem Fall bezieht sich die Translation auf den Schwerpunkt des partiellen Modells, wodurch eine präzise Einschätzung durch das OP-Personal deutlich erschwert wird. Ein Wirbel hingegen kann bzgl. seiner Verschiebung in t_x - und t_y -Richtung relativ genau geschätzt werden, sofern er nicht mit einem seiner Nachbarwirbel verwechselt wird²: $dt_x \approx dt_y \approx 10mm$. Die Richtung t_z hingegen, d.h. der Abstand des Wirbels zur Röntgenquelle, unterliegt einer größeren Unsicherheit: $dt_z \approx 50mm$.

Tabelle 4.1 faßt die experimentell bestimmten maximalen Restfehler bei Verwendung der vorgestellten Initialisierungsmethode nochmals in Abhängigkeit repräsentativer Knochenstrukturen zusammen.

4.2.2 Verwendung anatomischer Details

Die Ausnutzung diskreter, anatomisch markanter Details eines Knochens zum Zwecke der Vorregistrierung (Initialisierung) ist nicht neu (siehe z.B. [Str97]). Die übergreifende Gemeinsamkeit ansonsten unterschiedlicher Ansätze besteht darin, zunächst markante Punkte des Patientenknochens präope-

²Dieser Faupax einer Verwechslung ist nicht zu unterschätzen. Durch den beschränkten Blick des Röntgenbildes sind jeweils nur wenige Wirbel auf einmal zu sehen. Benachbarte Wirbel weisen naturgemäß eine sehr ähnliche Anatomie auf und sind daher selbst durch einen fachmännischen Blick in einer Fluoroskopieaufnahme oft nicht zweifelsfrei zu unterscheiden.

Anatomische Struktur	α	β	γ	t_x, t_y	t_z
Femur, proximal	2°	6°	13°	41 mm	74 mm
Femur, distal	4°	5°	10°	38 mm	59 mm
Tibia, proximal	3°	5°	12°	27 mm	62 mm
Wirbel (L2)	3°	3°	15°	12 mm	61 mm

Tabelle 4.1: Experimentell bestimmte, maximale Restfehler bei einer Initialisierung durch Vorgabe einer Aufnahme-richtung (vgl. Abb. 4.3 (Seite 33))

rativ zu definieren (z.B. innerhalb eines CT-Datensatzes) und dieselben Punkte später intraoperativ unter Verwendung einer anderen Signalgeber-Modalität (Fluoroskopie, IR-Trackingsystem) zu erfassen und einander eindeutig zuzuordnen.

Die in dieser Arbeit angewandten Verfahrensvarianten, deren Probleme sowie die damit zu erreichende Präzision einer Initialisierung werden im folgenden vorgestellt und diskutiert.

Verfahren

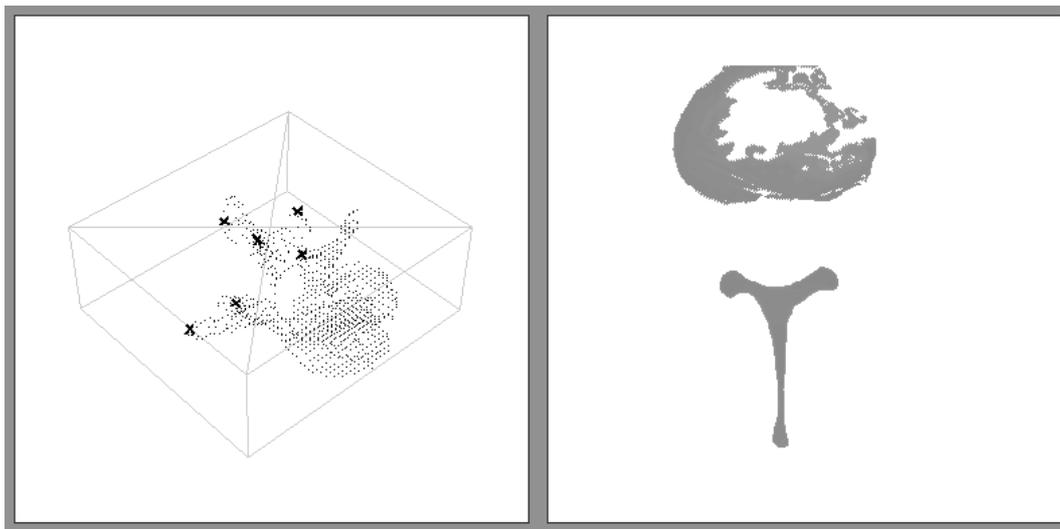


Abbildung 4.4: Definition anatomischer Landmarken

Nahezu jeder menschliche Knochen besitzt eine Reihe eindeutiger anatomischer Details, wobei an dieser Stelle lediglich punkt-basierte Oberflächenmerkmale betrachtet werden, d.h. Details auf der Außenseite eines Knochens, mit denen sich jeweils ein eindeutiger einzelner repräsentativer Punkt assoziieren lässt. Beispiele hierfür sind die Dorn- und Querfortsätze von Wirbeln, die Walzenhöcker eines distalen sowie die beiden Rollhügel eines proximalen Femurs (vgl. Anhang A). Im Sinne der Punkt-Assoziation ungeeignete anatomische Merkmale hingegen sind z.B. der Hals und die patellare Lauffläche eines Femurs.

Die manuelle, interaktive Definition der anatomisch markanten Punkte erfolgt in dieser Arbeit stets bzgl. eines zuvor aufgenommenen patientenspezifischen CT-Datensatzes (Abb. 4.4). Entscheidend für die

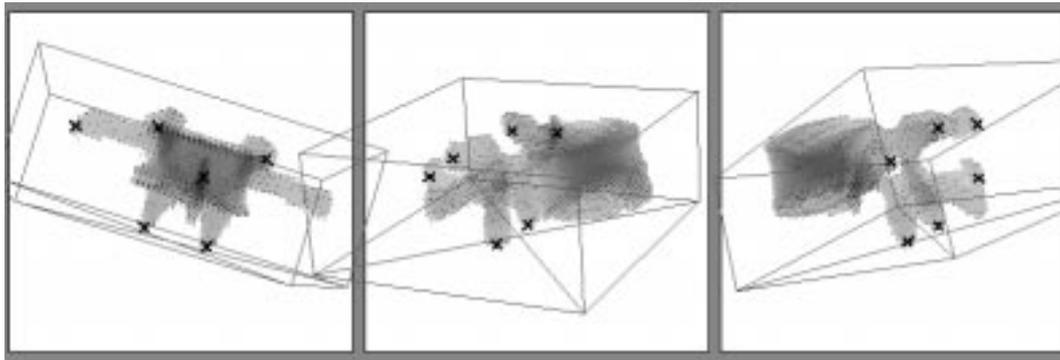


Abbildung 4.5: Interaktive Definition von Bildpunkten, die jeweils einer der zuvor definierten 3D-Landmarke entsprechen

Auswahl von Details ist dabei, daß sich entweder die Projektionen der markanten Punkte innerhalb der später aufzunehmenden Röntgenaufnahmen eindeutig wiederfinden lassen oder alternativ diese Punkte intraoperativ freigelegt und z.B. mit Hilfe eines optisch überwachten Zeigers abgetastet werden können.

Für die intraoperative Aquisition der präoperativ definierten Merkmale kommen im Rahmen des Szenarios dieser Arbeit drei unterschiedliche Varianten in Betracht:

- **2D-3D, Fluoroskopie (Einzelaufnahme):** Die mit den definierten Landmarken korrespondierenden Bildpunkte innerhalb eines aufgenommenen Fluoroskopie-Bildes werden interaktiv markiert und entsprechend zugeordnet. Auf der Basis von mindestens vier 2D-3D-Korrespondenzen (Bildpunkt (2D), CT-Merkmal (3D)) läßt sich anschließend eine initiale Lageschätzung berechnen [DeM92]. Dabei gilt es zu beachten, daß die ausgewählten 3D-Merkmalpunkte nicht koplanar sind, d.h. nicht in einer gemeinsamen Ebene liegen.
- **2D(3D)-3D, Fluoroskopie (mehrere Aufnahmen):** Die Verwendung mehrerer statt einer einzelnen Fluoroskopie-Aufnahme erlaubt eine stabile Initialisierung auf der Basis von 3D-3D-Korrespondenzen, vorausgesetzt, es finden sich mindestens drei anatomische Merkmale, die sich jeweils in mindestens zwei Röntgenbildern eindeutig (interaktiv) identifizieren lassen. Für je zwei Bildpunkte in unterschiedlichen Aufnahmen, die jedoch demselben anatomischen Merkmal entsprechen, läßt sich mittels Triangulierung³ ein zugehöriger 3D-Punkt bestimmen.

Liegen drei auf diese Weise berechnete 3D-Punkte vor, so kann nach eindeutiger Zuordnung zu den entsprechenden anatomischen 3D-Punkten innerhalb des CT-Datensatzes eine initiale Lageschätzung geschlossen bestimmt werden. Der in [HS93a] beschriebene lineare Lösungsansatz liefert stabile brauchbare Ergebnisse, solange die interaktiv markierten Bildpositionen nur schwach verrauscht sind, d.h. sehr präzise bestimmt wurden, und solange genügend Korrespondenzen verwendet werden. Ist dies nicht der Fall, so sind aufwendigere Verfahren, auf Quaternionen oder nichtlinearen Optimierungsansätzen basierend, zu bevorzugen [HS93b].

³Räumlicher Schnitt der mit den beiden Bildpunkten korrespondierenden Sichtgeraden im jeweiligen Kamerakoordinatensystem.

- **3D-3D, IR-Trackingsystem:**

Steht ein IR-Trackingsystem zur Verfügung, so können die anatomischen Merkmale intraoperativ mit Hilfe eines getrackten Zeigeeinstrumentes abgetastet werden. Voraussetzung hierfür ist, daß die anvisierten Stellen zuvor entsprechend *freigelegt* wurden. D.h. es kommen von vornherein nur Details in Frage, die durch das geöffnete Operationsfeld problemlos erreichbar sind. Im Falle einer Wirbelsäulen-Operation sind z.B. bei geöffnetem Rücken lediglich die Dorn- und Gelenksfortsätze einzelner Wirbel frei zugänglich. Allgemein gilt, daß bei vielen operativen Eingriffen nur sehr wenige Details eines Knochens zum Vorschein kommen und sich direkt abtasten lassen. Der Grund hierfür ist einleuchtend: Das Operationsfeld fällt auf der einen Seite so groß aus wie für die Durchführung des Eingriffs nötig, auf der anderen Seite aber gleichzeitig auch so klein wie möglich – nicht zuletzt, um nachfolgende Komplikationen zu vermeiden bzw. gering zu halten.

Für jedes erreichbare Oberflächendetail des betrachteten Knochens genügt ein einzelner Abtastpunkt, so daß die intraoperative Aquisition der 3D-Punkte sehr schnell vonstatten gehen kann. Die Bestimmung der initialen Lageschätzung erfolgt anschließend auf dieselbe Art und Weise wie im letzten Punkt beschrieben [HS93a].

Präzision

Mehrere experimentelle Untersuchungen zur Beurteilung der erzielbaren Präzision unter Verwendung der zuvor beschriebenen Verfahrensalternativen ergaben die nachfolgend zusammengefaßten Resultate.

Tabelle 4.2 faßt die erzielten Ergebnisse zusammen. Zunächst fällt auf, daß die beiden ersten Varianten (2D/3D und 2D(3D)/3D) für einen proximalen Tibia-Knochen nicht eingesetzt wurden. Der Grund hierfür ist, daß dieser Knochen zu wenige punkt-basierte anatomische Merkmale besitzt (vgl. Anhang A), die in einer Fluoroskopieaufnahme zweifelsfrei und präzise zu lokalisieren sind. Eine ähnliche Situation ist für ein distales Femurende gegeben. Hier bieten sich nur wenige Merkmale an (z.B. beide Walzenhöcker (Epicondylus lateralis/medialis), vgl. Anhang A), die zumindest aus speziellen Richtungen (lateral, AP) auch im Röntgenbild eindeutig markierbar sind. Aus diesem Grund ist die erste Verfahrensvariante prinzipiell anwendbar, die zweite hingegen nicht, da sich bei Betrachtung aus unterschiedlichen Richtungen i.a. nicht dieselben 3D-Merkmale in den einzelnen Fluoroskopieaufnahmen markieren lassen.

Alle drei Verfahrensvarianten liefern durchwegs für den L2-Wirbel bessere Ergebnisse als für einen der drei Stabknochen der unteren Extremitäten. Diese Erkenntnis ist intuitiv verständlich, da Wirbel durch ihren Aufbau bereits eine Vielzahl an punktförmigen anatomischen Landmarken besitzen⁴, die sich eindeutig erkennen und zuordnen lassen (vgl. Anhang A).

Es gilt zu beachten, daß die in der Tabelle zusammengefaßten Angaben maximaler Restfehler experimentell ermittelt wurden und keinen Anspruch auf Festlegung einer maximalen Fehlerobergrenze erheben können. Letzten Endes wird die in der Praxis tatsächlich zu erreichende Präzision im Einzelfall stets von der Bildqualität, den Aufnahme-Richtungen sowie in hohem Maße von den Anwendereingaben zur Markierung der anatomischen Merkmale in den einzelnen Fluoroskopiebildern abhängen. Die Tabelle dient lediglich der Bereitstellung repräsentativer Anhaltspunkte zur Orientierung.

⁴Insbesondere die jeweiligen Enden der diversen Dorn- und Gelenksfortsatz-Strukturen bieten sich hier an.

Anatomische Struktur	Δr	Δt
2D/3D, einzelne Fluoroskopie		
Femur, proximal	17°	14 mm
Femur, distal	12°	19 mm
Tibia, proximal	—	—
Wirbel (L2)	11°	9 mm
2D(3D)/3D, mehrere Fluoroskopien, Triangulierung		
Femur, proximal	10°	20 mm
Femur, distal	—	—
Tibia, proximal	—	—
Wirbel (L2)	2°	15 mm
3D/3D, IR-Trackingsystem, Abtastung		
Femur, proximal	11°	13 mm
Femur, distal	19°	18 mm
Tibia, proximal	12°	24 mm
Wirbel (L2)	7°	9 mm

Tabelle 4.2: Experimentell bestimmte, maximale Restfehler bei einer Initialisierung durch interaktive Zuordnung anatomischer Merkmale

4.2.3 Manuelle Modell-Plazierung

Abschließend soll ein drittes Initialisierungsverfahren vorgestellt werden, welches zwar einerseits die aufwendige Interaktion eines geübten Benutzers am OP-Computer erfordert, andererseits aber auch gute Ergebnisse zu liefern vermag.

Verfahren

Liegt von der zu registrierenden Knochenstruktur des Patienten ein dreidimensionales Modell vor, so läßt sich dieses für gegebene Lageparameter durch Projektion jeder Fluoroskopieaufnahme überlagern. Erfolgt die Berechnung der Projektion dabei in Echtzeit, kann ein Benutzer die Auswirkung einer Änderung der einzelnen Parameter auf die Qualität der Übereinstimmung in allen Aufnahmen sofort interaktiv kontrollieren. Die sukzessive Veränderung der Parameter führt letztlich zu einer initialen Transformation.

Prinzipiell eignet sich diese Vorgehensweise für beide im Rahmen dieser Arbeit entwickelten Registrierungstechniken. Zwar verzichtet die in Kapitel 5 vorgestellte intensitätsbasierte Registrierung explizit auf eine Modellbildung. Der dort stattdessen genutzte CT-Datensatz muß jedoch vor seiner Verwendung durch eine grobe Segmentierung geeignet maskiert (*geclippt*) werden, so daß er lediglich noch den betrachteten Einzelknochen, aber kein umgebendes Gewebe bzw. benachbarte Knochen mehr enthält. Aus einem derartig vorverarbeiteten CT-Datensatz läßt sich durch einfache Schwellwertbildung eine 3D-Punktmenge bestimmen, die die Oberfläche des enthaltenen Knochens beschreibt. Diese Punktmenge kann für die interaktive Initialisierung ersatzweise als Modell herangezogen werden (Abb. 4.6). Im Rahmen des Verfahrens zur elastischen Registrierung aus Kapitel 6 spielen hingegen 3D-Modelle selbst bereits eine zentrale Rolle. Ein solches Modell kann daher direkt für die Initialisierung herangezogen

werden.



Abbildung 4.6: Manuelle Modellplatzierung

Herausragender Nachteil einer manuellen Modell-Plazierung ist deren oftmals umständliche und zeitaufwendige Anwendung. Dies gilt insbesondere dann, falls durch das Benutzer-Interface keine sinnvolle Auswahl der die Transformation beeinflussenden Parameter möglich ist. Sind z.B. lediglich Rotationen und Translationen um die bzw. entlang der Achsen einer der virtuellen Kameras möglich, so reicht dies in vielen Fällen nicht für eine komfortable und schnelle interaktive Einstellung der Transformationsparameter aus.

Entsprechende Erweiterungen der Einflußmöglichkeiten sind jedoch in der Lage, den Bedienkomfort deutlich zu erhöhen. Beispiele hierfür finden sich u.a. in [Bra00a] (Drehung um beliebig orientierte, interaktiv definierte Achsen) und [Dö99] (3D Navigation).

Anatomische Struktur	Δr	Δt
Femur, proximal	10°	9 mm
Femur, distal	16°	11 mm
Tibia, proximal	12°	13 mm
Wirbel (L2)	7°	6 mm

Tabelle 4.3: Experimentell bestimmte, maximale Restfehler bei einer Initialisierung durch interaktive Platzierung eines 3D-Modells

Eine interaktive Platzierung des Modells ist auch in Kombination mit der eingangs vorgestellten automatischen Initialisierungsmethode (Vorgabe einer Aufnahme-richtung) denkbar: Sind aus irgendwelchen Gründen einzelne Parameter bei der Relativpositionierung des C-Bogens nicht ausreichend abschätzbar (z.B. die Entfernung Knochen-Kameraursprung/Röntgenquelle), so gewährleistet eine abschließende interaktive Nachkorrektur eine ausreichend gute Initialisierung. In diesem Fall reduziert sich der interaktive Aufwand im übrigen deutlich, da viele der Lageparameter durch die vorangegangene automatische Initialisierung bereits ausreichend gut approximiert wurden.

Präzision

Um die mit einer Initialisierung durch interaktive Modell-Plazierung erreichbare Präzision abschätzen zu können, wurden wie zuvor mehrere Experimente (je drei Aufnahmen aus unterschiedlichen Richtungen, Winkelabstand ca. 45°) mit verschiedenartigen Knochen durchgeführt. Tabelle 4.3 faßt die dabei beobachteten maximalen Restfehler – getrennt nach rotatorischen und translatorischen Abweichungen – zusammen. Es fällt insbesondere auf, daß sich Femur und Tibia im Vergleich zu dem L2-Lendenwirbel signifikant schlechter positionieren lassen. Lediglich der proximale Anteil des Femurknochens schnitt aufgrund seiner markanten Ausprägungen im Vergleich etwas besser ab. Die Ergebnisse insgesamt fallen auf der einen Seite deutlich besser aus als bei den beiden zuvor präsentierten Initialisierungsverfahren, insbesondere die translatorischen Restfehler betreffend. Dabei darf aber natürlich nicht in Vergessenheit geraten, daß sich auf der anderen Seite der interaktive Benutzeraufwand deutlich erhöht und daher je nach Anwendungsszenario u.U. nicht praktikabel erscheinen mag.

Kapitel 5

Intensitätsbasierte Registrierung von Fluoroskopieaufnahmen mit CT

Inhalt

5.1 Zielsetzung und Ansatz	42
5.1.1 Motivation	42
5.1.2 Ansatz	44
5.2 CT-basierte Simulation von Fluoroskopie-Aufnahmen	45
5.2.1 Ray-Casting durch CT-Volumen	45
5.2.2 Helligkeits- und Kontrast-Adaption	48
5.2.3 Bildmaskierung – Berücksichtigung von Artefakten	49
5.2.3.1 Artefakt-Typen	49
5.2.3.2 Detektion von Artefakt-Bildregionen	50
5.2.3.3 Filterung/Interpolation detektierter Artefakte	52
5.3 Lokalisation	57
5.3.1 Initialisierung	57
5.3.2 Optimierung	57
5.4 Beschleunigte Simulation von Röntgenaufnahmen	60
5.4.1 Ansatz	61
5.4.2 Ansichtsmodelle	61
5.4.2.1 Erzeugung von Trainingsbildern	62
5.4.2.2 Eigenraum-Repräsentation	63
5.4.2.3 Interpolation im Eigenraum	69
5.4.2.4 Planare Approximation von 3D-Transformationen	70
5.4.2.5 Approximationsfehler-Analyse	72
5.4.3 Optimierung	74
5.5 Ergebnisse	75
5.5.1 Simulationsergebnisse	75

5.5.2	Anwendung auf reale Fluoroskopieaufnahmen	79
5.5.3	Zeitaufwand der Optimierung	85

5.1 Zielsetzung und Ansatz

5.1.1 Motivation

Konturbasierte Ansätze (z.B. [Lav96a, FAB95]) stellen bereits seit einigen Jahren etablierte Techniken zur rigiden 2D/3D-Registrierung dar. Nichtsdestotrotz ergeben sich im Zusammenhang mit diesen Verfahren aber auch einige Probleme, die es in bestimmten Anwendungsfällen ratsam erscheinen lassen, nach einer Alternative Ausschau zu halten:

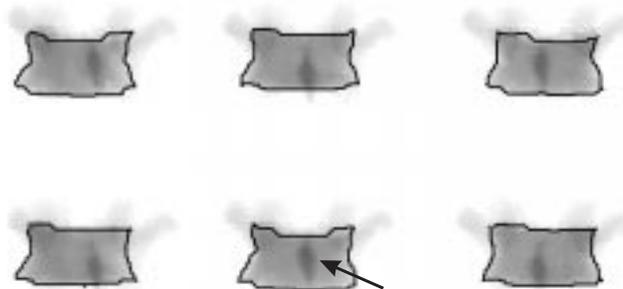


Abbildung 5.1: Umrißkonturen (schwarze Kurven) besitzen bei Röntgenbildern oftmals wesentlich weniger Aussagekraft als Grauwert-Strukturen. Erst die relative Lage des deutlich erkennbaren hinteren Dornfortsatzes (Pfeilmarkierung) läßt eine Rotation des T11-Brustwirbels erkennen (ca. 20° zwischen zwei benachbarten Ansichten).

- Solange keine vollautomatischen Verfahren zum Einsatz kommen können, ist die in jedem Röntgenbild vorzunehmende Segmentierung der umschließenden Kontur i.a. sehr aufwendig, was den Zeit- und Personalbedarf betrifft. Außerdem ist selbst bei automatischen Verfahren stets zumindest eine Kontrolle der Kontur durch einen erfahrenen Mediziner und daran anschließend in vielen Fällen eine manuelle Nachbesserung nötig. Dieselben Überlegungen gelten auch für die präoperative Erstellung des benötigten Oberflächenmodells, wobei insbesondere der hiermit nachwievor verbundene Zeitaufwand im klinischen Alltag oftmals nicht tragbar erscheint – mehrere Stunden pro Patienten-Datensatz sind keine Seltenheit. Angesichts aus der Literatur bekannter, zumindest teil-automatischer 3D-Segmentierverfahren mag dies zunächst verwundern. Doch eine solche Automatisierung, die z.B. bei durchschnittlichen CT-Datensätzen sehr gut funktioniert, ist im Falle eines realen Patientendatensatzes leider nur selten möglich. Im Hinblick auf die geplante Operation weichen solche CT-Aufnahmen oftmals deutlich von der Norm ab, da sie z.B. individuelle

Frakturen oder Knochenstrukturen mit stark fortgeschrittener patientenspezifischer Osteoporose enthalten.

- Es ist im praktischen Einsatz oftmals unmöglich, die vollständige Kontur, insbesondere markante Teile davon (z.B. Dornfortsätze bei Wirbeln), die für die Registrierung wertvolle Hinweise liefern, verlässlich zu erfassen. Bekannte Gründe hierfür sind z.B. Überlagerungen mit anderen Knochenstrukturen, der u.a. daraus resultierende geringe Kontrast sowie das niedrige Signal/Rausch-Verhältnis der intraoperativen Röntgenbilder.
- Der Ansatz konturbasierterer 2D-3D-Verfahren bedingt eine drastische Reduktion der vorhandenen Information, die damit der Registrierung für eine Optimierung der gesuchten Transformation nur teilweise zugänglich ist. Diese Reduktion findet präoperativ – ein einzelnes Oberflächenmodell repräsentiert das gesamte CT-Volumen – und intraoperativ – einzelne Konturen bzw. Silhouetten repräsentieren je ein Röntgenbild – statt. Problematisch hierbei ist, daß durch eine solche Reduktion mitunter wesentliche Informationen über innere Strukturen verworfen werden. Deren Fehlen kann letztendlich das Konvergenzverhalten sowie die Genauigkeit eines Registrierungsverfahrens negativ beeinflussen.
- Kleine Änderungen der Kameraausrichtung ändern oftmals die resultierenden Konturen (Silhouetten) in deutlich geringerem Umfang als die gesamte Grauwert-Region der betrachteten Knochenstruktur. Dies gilt insbesondere für Knochen, die Extrema mit signifikanter Ausdehnung in Blickrichtung der Kamera aufweisen (z.B. Wirbelkörper mit langen Dornfortsätzen). In diesem Fall stellen die sichtbaren Grauwert-Strukturen ein wesentlich sensitiveres Maß für eine veränderte Blickrichtung dar als Konturen (vgl. Abb. 5.1).
- Sowohl das aus dem CT-Volumen abgeleitete dreidimensionale Oberflächenmodell als auch die in den Röntgenbildern segmentierten Konturen stellen potentielle Quellen für Ungenauigkeiten dar, z.B. aufgrund von Fehlsegmentierungen. Desweiteren läßt sich die Frage, ob die segmentierten Konturen verfahrenstechnisch (Kantenfilter) überhaupt exakt mit den Begrenzungen des 3D-Modells korrespondieren, keineswegs trivial beantworten. Diese Korrespondenz ist aber essentielle Grundlage für die geforderte Genauigkeit der Registrierung.

Eine Technik zur 2D/3D-Registrierung von Röntgenbildern mit CT-Daten, die den Großteil der aufgeführten Probleme beseitigt, basiert auf der Berechnung von simulierten Röntgenbildern und deren Abgleich mit den realen Fluoroskopieaufnahmen. Die simulierten Bilder lassen sich hierfür direkt aus den CT-Daten rekonstruieren, indem der Röntgenprozeß vereinfacht nachgebildet wird. Solche simulierten Aufnahmen werden daher auch als digital rekonstruierte Röntgenbilder (*digitally reconstructed radiographs*, kurz DRRs) bezeichnet [BB96, Wee97b].

Wenngleich diese Art von Verfahren in der Lage ist, die genannten Probleme zu vermeiden, so ergibt sich doch andererseits im Vorfeld auch hier zunächst eine Vielzahl neuer kritischer Fragestellungen:

- Wie schnell können die DRRs während der Registrierung simuliert werden? Eine intraoperative Simulation ist bei Verwendung eines Röntgenbildverstärkers/Fluoroscops unabdingbar, da dieses Gerät durch seine Bewegungsfreiheit eine Vielzahl an Aufnahmerichtungen und -abständen erlaubt. Dies verhindert, daß bereits präoperativ eine Vielzahl an DRR-Aufnahmen für alle später intraoperativ zu erwartenden Blickrichtungen vorab rekonstruiert wird. Szenarien, wie z.B. in der

Radiochirurgie, wo anstatt eines beweglichen C-Bogen-Systems statische Röntgenquellen und -detektoren zur Lagekontrolle des Patienten Verwendung finden, erlauben dies [STAL94].

- Wie wirken sich intraoperativ u.U. im Bild sichtbare Elemente wie Fixationen, Kalibriermarken, Instrumente etc. auf die Qualität der Registrierung aus?
- Welchen Einfluß besitzt ein Kontrast- und Helligkeitsunterschied der realen Fluoroskopieaufnahmen im Vergleich zu den simulierten DRRs?
- Welche Auflösung der Röntgenbilder sowie der CT-Daten ermöglicht eine einerseits präzise, andererseits aber auch schnelle Registrierung?
- Wie gut muß die Registrierung initialisiert werden (Konvergenzradius des Verfahrens)?

Im folgenden wird nun ein neues Verfahren präsentiert, das in einigen grundlegenden Punkten dem erstmals in [Wee97b, Wee97a] vorgeschlagenen Ansatz folgt, diesen aber an anderer Stelle wesentlich erweitert und verbessert [Fre98, Rot99a, Rot99b]. Die einzelnen Abschnitte widmen sich dabei ausführlich den zuletzt aufgeworfenen Fragestellungen und präsentieren jeweils geeignete Lösungsansätze.

5.1.2 Ansatz

Im Gegensatz zu [Wee97b] erlaubt die in Anhang B vorgestellte Kalibrierung eines Fluoroskops die Etablierung eines *Multiview*-Ansatzes. D.h. die Registrierung setzt voraus, daß zuvor mehrere – i.a. zwei bis vier – reale Fluoroskopie-Aufnahmen aus möglichst unterschiedlichen Richtungen aufgenommen wurden. Die Aufnahmen werden entzerrt sowie deren virtuelle Lochkamera-Parameter bestimmt. Daraus ergibt sich schließlich eine *Szene* mit festen Transformationen zwischen den einzelnen virtuellen Kameragen. Durch den sukzessiven Vergleich der realen Fluoroskopien mit – während der Registrierung, d.h. intraoperativ berechneten – korrespondierenden simulierten Aufnahmen, die sich für jede einzelne Kamera für bestimmte Transformationen zwischen CT-Volumen und Referenz-Koordinatensystem ergeben, läßt sich die Lage der zu registrierenden Knochenstruktur durch entsprechende Variation der rigiden Transformationsparameter optimieren (Abb.5.2). Das Verfahren besitzt die folgenden beiden Voraussetzungen:

1. Eine kalibrierte Aufnahme-Szene (vgl. Abschnitt B), d.h. mehrere verzeichnungsfreie Fluoroskopieaufnahmen aus unterschiedlichen Richtungen, wobei alle internen und externen Kameramodell-Parameter bekannt sind.
2. Einen präoperativ erstellten patientenspezifischen CT-Datensatz. Dieser Datensatz muß bereits vorverarbeitet sein, d.h. durch eine Grobsegmentierung mit nachfolgendem *Clipping* enthält das Volumen lediglich noch die betrachtete Knochenstruktur, aber kein Gewebe und keine benachbarten Knochen mehr.

Ein Schlüsselkriterium, das letztlich über die Akzeptanz der Technik von Seiten der Chirurgen entscheidet, ist neben der Präzision der Registrierung nicht zuletzt deren Geschwindigkeit. Aus diesem Grund wird in Abschnitt 5.4 ein neues modellbasiertes Verfahren vorgestellt, das die ansonsten zeitintensive Rekonstruktion simulierter Röntgenaufnahmen deutlich beschleunigen kann. Die Idee basiert darauf, bereits präoperativ bestimmte Aufnahmerichtungen, die intraoperativ häufig genutzt werden, zu trainieren.

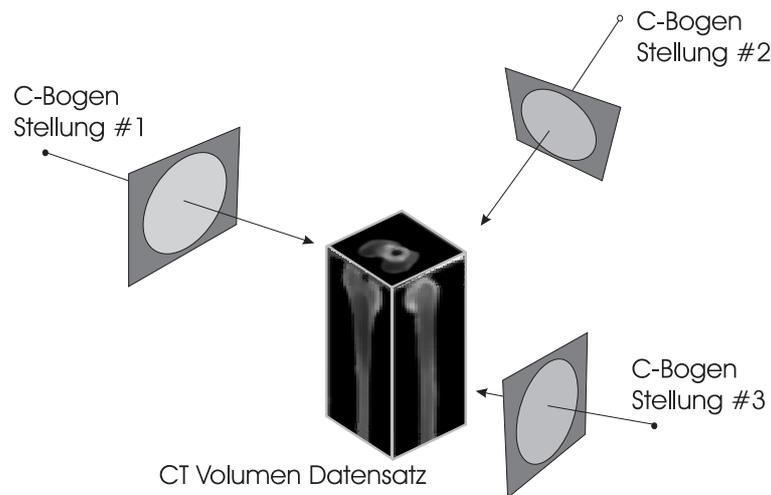


Abbildung 5.2: Intensitätsbasierte Registrierung aus mehreren Ansichten (Prinzip)

Es wird gezeigt, wie sich unter Einsatz dieser Ansichtsmodelle die Registrierung deutlich beschleunigen läßt, ohne dabei auf die erforderliche Genauigkeit verzichten zu müssen.

5.2 CT-basierte Simulation von Fluoroskopie-Aufnahmen

5.2.1 Ray-Casting durch CT-Volumen

In Abschnitt 3 wurden bereits ausführlich die physikalischen Grundlagen der Bildgewinnung bei den beiden Modalitäten Fluoroskopie und Computer-Tomographie vorgestellt. Der entscheidende Zusammenhang im Hinblick auf das Ziel einer CT-basierten Berechnung simulierter Fluoroskopie-Aufnahmen besteht darin, daß beiden bildgebenden Verfahren das Prinzip der Abschwächung von Röntgenstrahlung zugrundeliegt.

Unter Ausnutzung der physikalischen Zusammenhänge bei der Generierung der CT-Aufnahmen ist es möglich, den tatsächlichen Fluoroskopie-Bildgewinnungsprozeß durch einen *Ray-Casting*-Ansatz zu simulieren. *Ray-Casting*-Algorithmen stellen eine Unterklasse der bildgetriebenen *Volumen-Rendering*-Techniken¹ dar und dienen hauptsächlich Visualisierungszwecken [SML97]. *Bildgetrieben* bedeutet in diesem Zusammenhang, daß das zu berechnende Bild pixelweise generiert wird – im Gegensatz zu den hier nicht näher betrachteten modellgetriebenen Ansätzen. Die grundlegende Idee aller *Ray-Casting*-Verfahren besteht darin, den Wert eines Bildpunktes zu bestimmen, indem ein virtueller Strahl durch das betrachtete Volumen *geschickt* wird. Die Richtung des Strahls wird dabei durch vorgegebene Kameraparameter festgelegt. Entlang dieses Strahls wird anschließend eine Funktion in Abhängigkeit aller durch den Strahl passierten Volumenelemente (Voxel) ausgewertet. Nachdem der Strahl das Volumen wieder verläßt, liefert der resultierende Funktionswert den Grauwert des betrachteten Bildpixels.

Die vereinfachte Modellierung der Fluoroskopie-Bildentstehung alleine auf Basis des bekannten Abschwächungsgesetzes vernachlässigt mehrere physikalische Effekte, die bei dem Durchdringen von

¹Verfahren zur Berechnung von 2D-Bildern aus 3D-Datensätzen (Volumen)

Röntgenstrahlung durch Materie zu beobachten sind [Goe88]. Es zeigt sich, daß einige dieser Effekte im Rahmen der Simulation berücksichtigt werden müssen, andere hingegen ohne Konsequenzen vernachlässigt werden können [CAW96]:

- Quantenenergie-Niveau:

Zwar liegt sowohl der Fluoroskopie als auch der CT-Bildgewinnung Röntgenstrahlung zugrunde; die jeweils genutzten Energie-Niveaus sind jedoch unterschiedlich. Für Computer-Tomographien sind Scheitelspannungen von ca. $140kV$ üblich, wohingegen diese für Fluoroskopie-Aufnahmen im Bereich zwischen $60 \dots 110kV$ liegen und damit deutlich geringer ausfallen. Da es sich bei linearen CT-Schwächungskoeffizienten (μ) nicht um materialspezifische Konstanten handelt [Goe88], sondern diese nichtlinear von der verwendeten Strahlungsenergie abhängen, ist die direkte Verwendung der Schwächungskoeffizienten während des Ray-Castings problematisch. Wird dieser Zusammenhang bei der Generierung simulierter Bilder ignoriert, so fällt das Ergebnis unrealistisch kontrastarm aus [vZo97]. Insbesondere der für reale Röntgenbilder typische starke Kontrast zwischen Weichgewebe und Knochen geht verloren. Dieses Problem läßt sich durch Berücksichtigung energie- und gewebeabhängiger Korrekturfaktoren beheben. In [CAW96] z.B. werden solche Faktoren getrennt für vier Gewebetypen (Luft, Wasser, Muskelgewebe und Knochen) bestimmt und in Abhängigkeit von mehreren diskreten Fluoroskop-Scheitelspannungen (im Bereich von $32 \dots 161kV$) tabellarisiert. Werden Korrekturfaktoren für nicht in der Tabelle enthaltene Spannungen benötigt, so kommt eine einfache lineare Interpolation zum Einsatz.

Die Berechnung simulierter Fluoroskopieaufnahmen basiert in der vorliegenden Arbeit jeweils auf *geclippten*, d.h. vorsegmentierten CT-Datensätzen, die lediglich noch Knochen enthalten. Deshalb sind hier nur entsprechende Korrektur-Faktoren für Knochen-Strukturen nötig. Im Hinblick auf die praktische Anwendung energieabhängiger Korrekturfaktoren während der intraoperativen Registrierung stellt sich allerdings die Frage, wie das System die für eine Berechnung nötigen einzelnen Scheitelspannungen der realen Aufnahmen erhalten soll. Handelsübliche C-Bogen-Geräte bieten hierfür i.a. keinen direkten Signalausgang. Deshalb wird in dieser Arbeit ein konstanter Korrekturfaktor für eine übliche Scheitelspannung verwendet. Zusätzliche Anpassungen bzgl. Bildhelligkeit und Kontrast erfolgen dann entsprechend der im nächsten Abschnitt beschriebenen Vorgehensweise.

- Strahlen-Aufhärtung (*beam hardening*):

Für niedrige Quantenenergien ergeben sich tendenziell große und für hohe Energien entsprechend kleine lineare Schwächungskoeffizienten. Dieser Zusammenhang hat zur Folge, daß die niederenergetischen Anteile des Strahlenspektrums bei einer Gewebedurchdringung stärker und schneller absorbiert werden als hochenergetische Anteile. D.h. je weiter ein Strahl in das Innere eines Körpers vordringt, desto stärker nehmen die effektiven Schwächungskoeffizienten des Gewebes ab. Auch dieser als *Strahlen-Aufhärtung* bezeichnete Effekt ist hochgradig nicht-linear.

Auf eine Berücksichtigung des physikalischen Phänomens im Rahmen der Simulation kann allerdings verzichtet werden, da z.B. eine lineare Approximation des Effektes keine sichtbare Verbesserung der resultierenden Bilder bewirkte [CAW96]. Aufgrund des vorsegmentierten Volumens ergeben sich jeweils nur sehr kurze Strecken zwischen Eintritts- bis Austrittspunkt des Strahls, so daß der Effekt nicht zum Tragen kommt.

- Photonen-Rauschen:

Fluoroskopie-Aufnahmen sind meist stark verrauscht. Verantwortlich hierfür ist die relativ geringe Anzahl an abgespaltenen Photonen, was eine starke statistische Fluktuation der Quantenanzahl, die von einer Einheitsfläche des Bildverstärkers absorbiert wird, nach sich zieht. Dieser Effekt läßt sich leicht in die Simulation integrieren, indem eine berechnete Fluoroskopie-Aufnahme abschließend mittels einer gaußschen Rauschverteilung nachbearbeitet wird. Im Hinblick auf die angestrebte Optimierung erwies sich das künstliche Verrauschen der simulierten Aufnahmen jedoch als eher nachteilig. Daher wurde auch dieser physikalische Effekt vernachlässigt.

Abb. 5.3 stellt eine auf diese Weise simulierte Aufnahme (rechts) einer ähnlichen realen Fluoroskopie-Aufnahme (links) gegenüber.

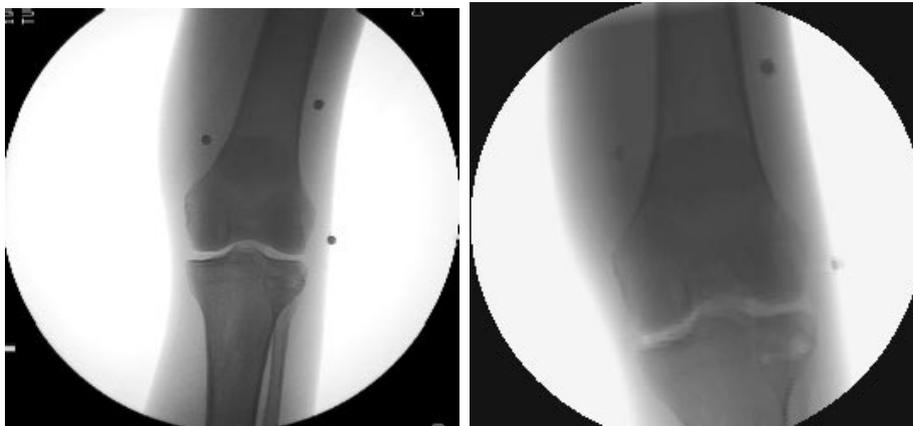


Abbildung 5.3: Reales (a) und simuliertes Röntgenbild (b) eines Patienten-Knies

Geradlinig realisierte *Ray-Casting*-Verfahren sind i.a. sehr langsam. Es gibt durchaus Anwendungen, bei denen dies keine Rolle spielt. Im vorliegenden Anwendungsfall jedoch stellt der signifikante Zeitbedarf für die Bildberechnung ein schwerwiegendes Problem dar, da während der angestrebten Registrierung eine Vielzahl an simulierten Aufnahmen berechnet werden muß. Die Zeit, die für die Generierung eines einzelnen Bildes benötigt wird, besitzt daher den größten Einfluß auf die Dauer der Registrierung. Letztendlich ist es unabdingbar, den *Ray-Casting*-Prozeß weitestgehend zu beschleunigen, um eine praktikable Registrierungsdauer zu ermöglichen.

Es gibt eine Reihe von Techniken, die allesamt auf die Performance-Verbesserung von *Ray-Casting* abzielen [SML97]. In [CAW96] und [GWS99] werden zwei unabhängige Verfahren zur beschleunigten Bildberechnung vorgestellt. Die erstgenannte Arbeit setzt eine spezielle 3D-Grafik-Hardware ein, um die Simulation der exponentiellen Abschwächung des Röntgenstrahls durch die elegante Verwendung von hardware-unterstütztem *3D-Texture Mapping* zu beschleunigen. Die zweite Arbeit schlägt die Verwendung einer *shear-warp*-Transformation vor. Dabei handelt es sich um eine in der einschlägigen Computer-Grafik-Literatur seit einigen Jahren hinlänglich bekannte hybride Methode zur Zeitoptimierung von *Ray-Casting*-Verfahren, die gleichzeitig bild- und modellgetrieben arbeitet. Die Grundidee besteht darin, in einem ersten Schritt das Voxel-Volumen dahingehend zu manipulieren, daß alle verfolgten Strahlen zueinander parallel und insbesondere senkrecht zu einer Begrenzungsfläche des Volumens verlaufen (*shearing*). Diese einmalige Transformation ermöglicht es, anschließend auf die Verfolgung der

einzelnen Strahlen zu verzichten. Die stark vereinfachte Geometrie erlaubt stattdessen die sehr schnelle Generierung eines Zwischenbildes in einer zu der Begrenzungsfläche parallelen Ebene. Hierzu genügt es, das modifizierte Voxel-Volumen sukzessive zu durchlaufen und den Zwischenbildbeitrag jedes einzelnen Voxels zu berechnen. Abschließend muß das resultierende Zwischenbild durch eine einfache planare Interpolation in das endgültige Ergebnisbild transformiert werden, um so die anfangs durchgeführte Volumentransformation auszugleichen (*warping*).

Die vorliegende Arbeit beschreitet einen anderen Weg, indem versucht wird, das *Ray-Casting* intraoperativ für häufig genutzte Ansichten weitestgehend zu vermeiden und stattdessen bereits präoperativ die Bildgenerierung für diese Ansichten zu trainieren und kompakt zu speichern. Abschnitt 5.4.2 setzt sich ausführlich mit dieser Thematik auseinander. Darüber hinaus wurde das *Ray-Casting* dadurch beschleunigt, daß jeweils zu Beginn der Strahlverfolgung eine erhöhte Schrittweite entlang des Strahls genutzt wird, solange bis der Strahl auf ein Voxel des Objektes im Zentrum des *geclippten* CT-Volumens trifft. Anschließend wird die Schrittweite verringert und solange beibehalten, bis der Strahl das Objektvolumen wieder verläßt. Dabei wird die Schrittweite auf den ursprünglichen hohen Wert zurückgesetzt.

In den nachfolgenden Abschnitten wird für die klare Trennung zwischen realen und simulierten Fluoroskopie-Aufnahmen eine unterschiedliche Namensgebung verwendet. Dabei steht $\mathcal{I}_{x_i}(u, v)$ für ein reales Bild, bezogen auf einen Kameradatensatz i , $\mathcal{I}_s(u, v)$ hingegen bezeichnet eine simulierte Aufnahme. Ggf. sorgen weitere Indizes für eine eindeutige Zuordnung, z.B. ein tiefgestellter Index i zur Auswahl der korrespondierenden Kamera sowie ein hochgestellter Index \mathbf{p} , um die der simulierten Aufnahme zugrundeliegende rigide Transformation (Objektlage) festzulegen: $\mathcal{I}_s^{\mathbf{p}}(u, v)$.

5.2.2 Helligkeits- und Kontrast-Adaption

Die im letzten Abschnitt beschriebene Technik liefert zunächst für eine vorgegebene Ansicht stets Aufnahmen gleicher Helligkeit sowie gleichen Kontrastumfangs. Dies stellt insofern ein Problem dar, als daß dadurch ein unmittelbarer intensitätsbasierter Vergleich mit realen Röntgenaufnahmen zunächst zum Scheitern verurteilt ist. Aufgrund der i.a. automatischen Einregelung eines handelsüblichen Röntgenbildverstärkers ist die Helligkeit sowie der Kontrast eines Bildes von Aufnahme zu Aufnahme sehr starken Schwankungen unterworfen und meistens nicht vorhersagbar. Es gibt eine Vielzahl an entscheidenden Einflußgrößen wie z.B. Fixationen, Instrumente, Kalibriermarken oder mit einem Kontrastmittel gefüllte Arterien, die sich im Strahlengang befinden.

Um diesen Unterschied zwischen realer und simulierter Aufnahme abzuschwächen und die beiden Bilder damit einem direkten, d.h. pixelbasierten Vergleich zugänglich zu machen, werden anhand jedes realen Röntgenbildes zwei skalare Korrekturwerte bestimmt. Ein Wert bg beschreibt dabei den durchschnittlichen *Hintergrund*-Grauwert der realen Fluoroskopie-Aufnahme außerhalb des betrachteten Objektes, d.h. z.B. in einem Bereich, in welchem sich nur Gewebe und Weichteile befinden, aber keine knöchernen Strukturen zu sehen sind. Dieser Wert wird innerhalb eines kleinen, aber repräsentativen – manuell oder durch geeignete Heuristiken festgelegten – Bildausschnitts R_{bg} (*region-of-interest, ROI* (Abb. 5.4 (a), linkes unteres Rechteck)) ermittelt:

$$bg = \frac{1}{|R_{bg}|} \sum_{(u,v) \in R_{bg}} \mathcal{I}_x(u, v)$$

Der zweite Wert, fg , erfaßt den mittleren Grauwert-Unterschied zwischen realer und simulierter Aufnahme. Er wird ebenfalls innerhalb eines repräsentativen Bildbereichs R_{fg} bestimmt, der in diesem Fall

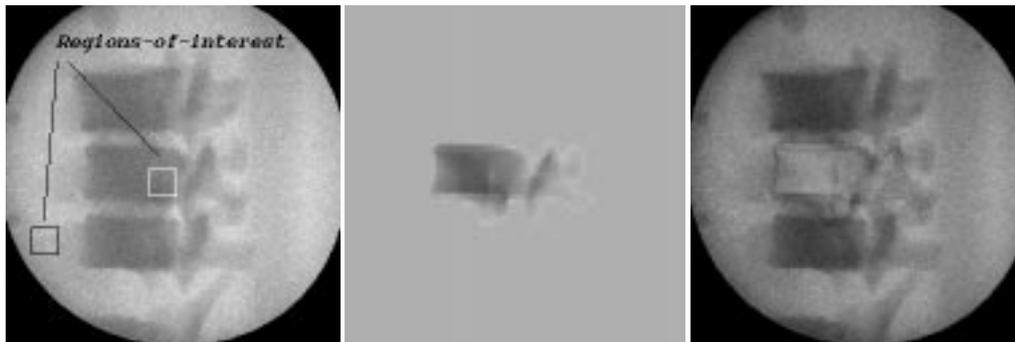


Abbildung 5.4: Reale Fluoroskopie-Aufnahme (a), Simulierte Aufnahme für eine ähnliche Kamera-Position (b), Differenzbild (c). In jedem realen Röntgenbild wird an zwei Stellen (einmal innerhalb, einmal außerhalb des betrachteten Objektes) ein durchschnittlicher Grauwert bestimmt, um die Helligkeit der simulierten Aufnahme daran anzupassen (a).

in *beiden* Bildern innerhalb des Objektes, z.B. des Wirbels liegt (Abb. 5.4 (a), rechtes oberes Rechteck). Abschließend wird der zuvor bereits ermittelte bg -Wert subtrahiert, um einen Differenzwert, der die betrachtete Knochenstruktur separat beschreibt, zu erhalten:

$$fg = \frac{1}{|R_{fg}|} \sum_{(u,v) \in R_{fg}} (\mathcal{I}_x(u, v) - \mathcal{I}_s(u, v)) - bg$$

Unter Verwendung dieser beiden Korrekturwerte fg und bg für Vorder- bzw. Hintergrundpixel läßt sich das aktuelle simulierte Röntgenbild $\mathcal{I}_s(u, v)$ entsprechend an die korrespondierende reale Aufnahme $\mathcal{I}_x(u, v)$ anpassen. Es ergibt sich ein modifiziertes Bild $\mathcal{I}_{\tilde{s}}(u, v)$:

$$\mathcal{I}_{\tilde{s}}(u, v) = \mathcal{I}_s(u, v) + \begin{cases} fg + bg & : \text{ falls } \mathcal{I}_s(u, v) \neq 0 \\ bg & : \text{ sonst} \end{cases} \quad (5.1)$$

5.2.3 Bildmaskierung – Berücksichtigung von Artefakten

Aufgrund der zu erwartenden Sensitivität eines pixelweisen Vergleichs simulierter Röntgenbilder mit realen Fluoroskopie-Aufnahmen ist die Berücksichtigung und Sonderbehandlung störender Artefakte unabdingbar. Unter Artefakten sind hierbei deutlich sichtbare Bildinhalte der realen Aufnahmen zu verstehen, die intraoperativ durch zusätzliche, im Strahlengang des C-Bogens befindliche, Objekte entstehen können. Da hierfür keine Entsprechung innerhalb des präoperativ aufgenommenen CT-Datensatzes des Patienten existiert, fehlen entsprechende Bildinhalte in den simulierten Aufnahmen. Insbesondere, falls es sich um Objekte aus Metall handelt, zieht dies enorme Probleme beim Bildvergleich nach sich.

5.2.3.1 Artefakt-Typen

Im Kontext dieser Arbeit ergaben sich die folgenden konkreten Beispiele für Objekte im Strahlengang, die jeweils störende Artefakte in den Fluoroskopie-Bildern verursachen:

- **Kalibrierphantome:**
Je nachdem, welcher der in Anhang B (C-Bogen Kalibrierung) vorgestellten Modi für die Eichung des verwendeten Fluoroscops zum Einsatz kam, finden sich intraoperativ mitunter die zahlreichen Metallmarken eines oder mehrerer Kalibrierphantome im Bild wieder (Abb. 5.5 (b-e)) [Bra00a].
- **Knochenfixationen:**
Für den Fall, daß die Knochenstruktur des Patienten zum Zeitpunkt der Registrierung durch eine entsprechende Vorrichtung starr fixiert wird [Gö00], ergeben sich hierdurch im Bild wenige großflächige Störregionen (Abb. 5.5 (d)).
- **Chirurgische Instrumente:**
Befinden sich zum Zeitpunkt der Aufnahmen chirurgische Instrumente im Strahlengang (Klammern, Zeigeinstrument etc.), so sind auch diese in den Fluoroskopien zu erkennen, sofern sie nicht aus weitestgehend röntgendurchlässigem Material bestehen (Abb. 5.5 (a,b)).

Der zuletzt genannte Objekttyp ist insofern unproblematisch, da sich hier ein Aufenthalt im Strahlengang i.a. vermeiden läßt, zumindest zum Aufnahmezeitpunkt der für die Registrierung verwendeten Röntgenbilder. Dasselbe gilt z.B. auch für eine Kontrastmittel-Einspritzung o.ä., was ebenfalls Artefakte zur Folge hätte. Kalibrierphantome und Fixationsvorrichtungen hingegen lassen sich oftmals nicht vermeiden, so daß die hierdurch hervorgerufenen Artefakte vor der Optimierung unbedingt berücksichtigt werden müssen.

In einem ersten Schritt sollen hierzu diejenigen Bildregionen ermittelt werden, die störende Artefakte der beschriebenen Art aufweisen. Die Pixel dieser Regionen definieren anschließend eine Art Maske für jedes aufgenommene Röntgenbild. Jeder maskierte Pixel erfährt anschließend eine Sonderbehandlung, sei es dadurch, daß die maskierten Bildbereiche während der Optimierung keine Berücksichtigung finden, oder dadurch, daß die Grauwerte dieser Bereiche aufgrund der umliegenden Bildinformationen *geglättet* bzw. zumindest teilweise rekonstruiert werden².

5.2.3.2 Detektion von Artefakt-Bildregionen

Je nach Art der störenden Artefakte bieten sich für die Ermittlung der betroffenen Bildteile zwei grundsätzlich verschiedene Vorgehensweisen an:

- **Virtuelle Rückprojektion bekannter Objekte:**
Handelt es sich bei dem Artefakte verursachenden Gegenstand um ein Kalibrierphantom, so ist zum Zeitpunkt der Registrierung dessen Position und Orientierung innerhalb des Kamera-Koordinatensystems aufgrund der bereits erfolgten Eichung (Anhang B) exakt bekannt. Dasselbe gilt für die geometrische Anordnung und Ausdehnung der Marken des Phantoms, wovon ein 3D-Modell existiert. Dieses Modell kann daher ohne Aufwand unter Verwendung des virtuellen Lochkamera-Modells in jede einzelne Fluoroskopie-Aufnahme projiziert werden. Die davon betroffenen Bildpixel liefern dann in ihrer Gesamtheit direkt die gesuchte Bildmaske aller resultierenden Artefakt-Regionen.

²Rekonstruktion bedeutet dabei, daß versucht wird, annähernd diejenigen Grauwerte zu bestimmen, die sich an den betroffenen Pixelpositionen ohne das Artefakte verursachende Objekt ergeben hätten. Selbstverständlich macht eine solche Rekonstruktion nur dann Sinn, falls die Artefakt-Regionen jeweils nur wenige Pixel groß ausfallen.

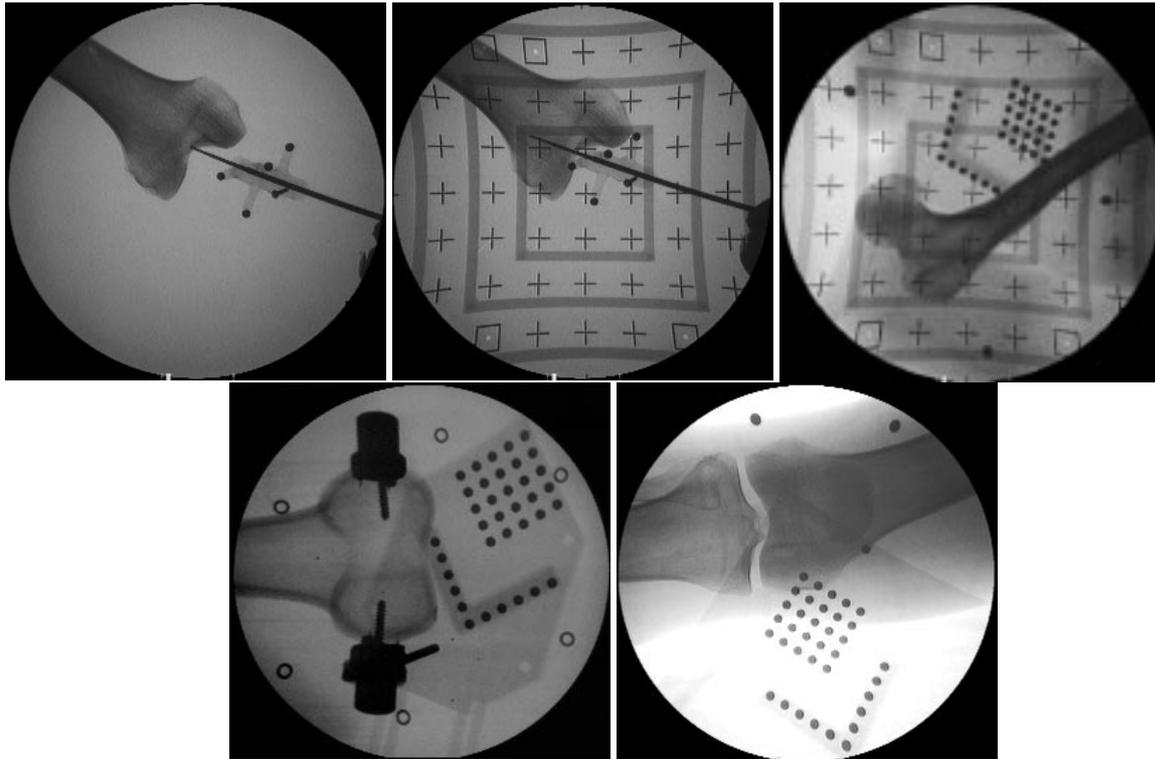


Abbildung 5.5: Fünf Beispiele zu berücksichtigender Artefakte aufgrund des intraoperativen Einsatzes von metallischen Marken, chirurgischen Instrumenten und/oder Kalibrierphantomen: (a) distaler Femur mit Instrument, (b) dito, mit Kalibrierphantom I, (c) proximaler Femur mit zwei Kalibrierphantomen (I und II), (d) Femurfixation, Randmarken und Kalibrierphantom II, (e) Kniephantom, Metallmarken und Kalibrierphantom II

In der Praxis hat es sich bewährt, die einzelnen Regionen zusammenhängender Maskenpixel vor der Weiterverwendung der Maske geringfügig zu vergrößern. Auf diese Weise wird eventuellen Ungenauigkeiten Rechnung getragen und sichergestellt, daß die Maske auch tatsächlich alle Artefakt-Pixel überdeckt.

- Bildanalyse zur Erkennung unbekannter Artefakt-Regionen:
Rühren die Artefakte von Objekten her, deren Geometrie, Position und Orientierung dem System i.a. nicht bekannt ist, so muß die Bestimmung der Artefaktregionen auf andere Weise erfolgen. Bezogen auf die Beispiele aus Abb. 5.5 gilt dies für das chirurgische Zeigelinstrument sowie die gezeigte Fixationsvorrichtung für Knochen.

Die hier betrachteten Artefakte lassen sich jeweils ausschließlich auf metallische Grundmaterialien der Objekte zurückführen. Da Metall Röntgenstrahlung stark abschwächt, fallen die betroffenen Bildteile im Vergleich zu ihrer Umgebung auch stets deutlich dunkler aus. Von einer Übersteuerung des C-Bogen-Gerätes abgesehen, ergeben sich jeweils tief-schwarze Bildanteile.

Unter Ausnutzung dieser Erkenntnis lassen sich Artefakt-Regionen in Bildern guter Aufnahmequalität (hoher Kontrast, geringes Rauschen) durch die Suche nach sehr dunklen Bildpixeln bestimmen. Da aber selbst Bilder neuer C-Bogen-Geräte prinzipbedingt ein niedriges Signal/Rausch-Verhältnis aufweisen, ist es nötig, während der Suche auf zusätzliches *a-priori*-Wissen über Größe und Form der zu suchenden Regionen zurückzugreifen. Ansonsten kommt es vor, daß auch andere dunkle – jedoch keine Artefakte darstellenden – Bildinhalte fälschlicherweise maskiert werden, was bei nachfolgender Glättung der maskierten Regionen fatale Auswirkungen haben kann.

Soweit anwendbar, ist die erstgenannte Vorgehensweise gegenüber der zweiten Methode zu bevorzugen. Die Vorteile liegen auf der Hand. Durch eine Projektion der Objekte werden alle Marken berücksichtigt, insbesondere auch solche, die z.B. aufgrund geringen Kontrasts, Überlagerungseffekten o.ä. zuvor im Bild nicht detektiert werden konnten. Außerdem wird auf diese Weise jegliche Fehldetektion umgangen, d.h. die resultierende Maske enthält tatsächlich nur Bildregionen, die entsprechende Objektteile zeigen.

Unabhängig von der gewählten Methode, ergibt sich jedoch in beiden Fällen schließlich ein binäres Maskenbild $\mathcal{I}_x^{mask}(u, v)$, welches dieselben Ausmaße besitzt wie das Ausgangsbild $\mathcal{I}_x(u, v)$. Maskierte Pixel erhalten den Wert 1, unmaskierte entsprechend 0.

5.2.3.3 Filterung/Interpolation detektierter Artefakte

Unter Verwendung dieses Maskenbildes erfolgt nun die lokale Glättung bzw. Interpolation, indem das zentrale Element einer quadratischen Filtermatrix ungerader Kantenlänge (Abb. 5.6) sukzessive über alle Pixel (u, v) des Ausgangsbildes bewegt wird. Ist der Pixel maskiert, so wird sein Grauwert nach folgender Rechenvorschrift auf Basis der durch die Filtermatrix abgedeckten lokalen Pixelumgebung neu berechnet:

$$\mathcal{I}_{\tilde{x}}(u, v) = \begin{cases} \mathcal{I}_x(u, v) & : \text{ falls } \mathcal{I}_x^{mask}(u, v) = 1 \\ \frac{\sum_{ij} [w_{ij} \mathcal{I}_x(u+i, v+j)]}{\sum_{ij} w_{ij}} & : \text{ sonst} \end{cases} \quad (5.2)$$

Dabei bezeichnet $\mathcal{I}_x(u + i, v + j)$ einen Bildpixel, der von einer – an der Position (u, v) zentrierten – Filtermatrix abgedeckt wird. Die ganzzahligen Indizes i und j beschreiben den horizontalen bzw.

0/1	0/2	0/3	0/2	0/1	1	2	0	0	0
0/2	0/3	0/4	0/3	0/2	2	0	0	0	0
0/3	0/4	0	0/4	0/3	3	4	0	0	3
0/2	0/3	0/4	0/3	0/2	2	3	0	0	2
0/1	0/2	0/3	0/2	0/1	1	2	3	0	1

Abbildung 5.6: Links: 5×5 -Filter-Maske zur gewichteten lokalen Glättung. Jedes Feld enthält Alternativwerte für die Gewichtung. Grauwerte benachbarter Pixel werden umso stärker gewichtet, je geringer ihr Abstand zum Zentrum der Filtermaske ist. Maskierte Pixel hingegen bleiben unberücksichtigt (Gewicht 0). Rechts: Beispiel-Situation, die dunkelgrauen Pixel seien maskiert. Der neue Grauwert des zentralen Pixels ergibt sich durch gewichtete Interpolation aus den restlichen nicht-maskierten Pixeln.

vertikalen Offset bzgl. (u, v) . Sie besitzen für eine Matrix der Dimension $m \times m$ den gemeinsamen Wertebereich $[-\frac{m-1}{2}, \dots, \frac{m-1}{2}]$, d.h. $(i, j) = (0, 0)$ entspricht dem zentralen Filterelement. Der Faktor w_{ij} legt fest, wie stark der Grauwert an Position $(u + i, v + j)$ zu der Grauwert-Interpolation für (u, v) beiträgt. Dabei gilt es folgendes zu beachten:

- Kommen Teile der Filtermatrix bei einer Platzierung derselben im Randbereich des Bildes außerhalb der Bildgrenzen zu liegen, so beträgt deren Gewicht 0.
- Pixel, die zu einem der Artefakte beitragen und deshalb maskiert sind, sollten keinen – die Rekonstruktion verfälschenden – Interpolations-Beitrag liefern. Die zugehörigen Gewichte für solche Pixel werden daher folgerichtig stets auf 0 gesetzt. Da die Interpolation nur erfolgt, falls die Filtermatrix auf einem maskierten Pixel plazierte wurde, hat dies zur Folge, daß das Gewicht für das Matrixzentrum selbst konstant gleich 0 ist: $w_{00} = 0$.
- Matrixelemente, die auf zulässigen, d.h. unmaskierten und im Bildbereich befindlichen Pixel-Koordinaten zu liegen kommen, werden in Abhängigkeit von ihrem Abstand zum Matrixzentrum gewichtet. Als Abstandsmaß dient anstelle des euklidischen Abstands die einfachere zu berechnende *Manhattan-Distanz* $d_{MAN}(i, j)$:

$$d_{MAN}(i, j) = |i| + |j|$$

Direkt benachbarte Pixel sollen ein höheres Gewicht und damit einen stärkeren Einfluß auf die Interpolation als weiter entfernte Pixel erhalten. Um dies zu erreichen, wird der ermittelte Distanzwert von der konstanten Spalten- bzw. Zeilenanzahl m der Filtermatrix subtrahiert: $w_{ij} = m - d_{MAN}(i, j)$. Damit erhalten die vom Zentrum am weitesten entfernten Filterelemente unabhängig von m jeweils das Gewicht 1. Im Gegensatz dazu werden dem Zentrum direkt benachbarte Elemente mit einem maximalen Gewicht von $(m - 1)$ assoziiert (Abb. 5.6).

Zusammenfassend ergibt sich die folgende Berechnung der Gewichtungsfaktoren ($-\frac{m-1}{2} \leq i, j \leq \frac{m-1}{2}$):

$$w_{ij} = \begin{cases} 0 & : \text{ falls } (u+i, v+j) \text{ ungültig} \\ 0 & : \text{ falls } \mathcal{I}_x^{mask}(u+i, v+j) = 1 \\ m - d_{MAN}(i, j) & : \text{ sonst} \end{cases} \quad (5.3)$$

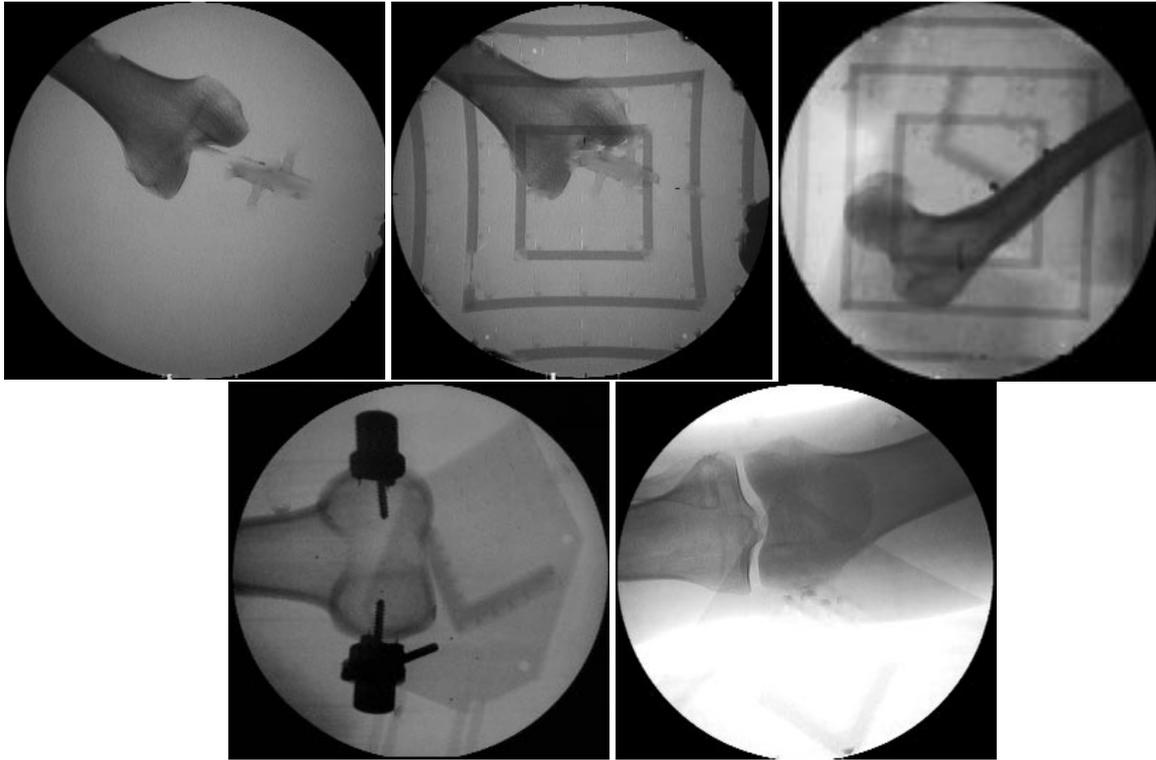


Abbildung 5.7: Aufnahmen aus Abb. 5.5, nach erfolgter lokaler Glättung von Artefakt-Regionen

Abb. 5.7 zeigt dieselben Fluoroskopie-Aufnahmen wie Abb. 5.5, nachdem diese durch das beschriebene Verfahren der lokalen Filterung bzw. Interpolation von Artefakten bearbeitet wurden. Zur Bestimmung zu maskierender Pixelbereiche wurde dabei für die kugel- und kreuzförmigen Metallmarken der Kalibrierphantome in Abb. 5.5 (b-e) jeweils die bevorzugte Technik der Rückprojektion angewandt. Die Detektion des ebenfalls mit Marken versehenen Zeigeeinstruments in Abb. 5.5 (Teilbilder a, b) sowie der vereinzelt an dem Kniephantom in den Teilbildern (e) und (f) angebrachten Metallkugeln erfolgte hingegen durch die zweite Technik der gezielten Bildanalyse.

Die beiden großflächigen Artefakte, die sich durch die Femur-Fixation in Teilbild (d) ergeben, wurden nicht behandelt. Mit entsprechend vergrößerter Filtermatrix (im Bsp. 31×31) ist eine Filterung der großflächigen Artefakte zwar prinzipiell möglich (Abb. 5.8, rechts), aber nicht sinnvoll. Der Grund hierfür ist die große Ausdehnung der betroffenen Bildteile, die eine verlässliche Rekonstruktion von Pixel-Grauwerten verhindert. Im Gegensatz zu der Situation im Falle kleiner Kalibriermarken könnten hier z.B. wesentliche Knochendetails verdeckt sein. Eine Pixel-Rekonstruktion auf der Basis umliegender Grauwerte käme einem *Stochern im Nebel* gleich. Dies ist nicht vertretbar. Für derartig flächendeckende

Artefakte ist es nötig, die berechnete Maskierung auch während der im nächsten Abschnitt beschriebenen Optimierung zu verwenden, indem maskierte Pixel nicht in die Bewertung eines Bildvergleichs einfließen.



Abbildung 5.8: Die Glättung/Interpolation großflächiger Bildteile (hier der Knochen-Fixationen, links) ist prinzipiell möglich, aber im Hinblick auf eine praktische Anwendung sehr fragwürdig. Evtl. verdeckte Knochenteile lassen sich nicht rekonstruieren, so daß das geglättete Bild (rechts) die Gegebenheiten der realen Situation u.U. stark verfälscht. Die negativen Auswirkungen auf nachfolgende Segmentierungs- und Registrierungsalgorithmen sind nicht kalkulierbar. Für Artefakte aufgrund von Kalibriermarken gilt dies nicht, da diese meist nur wenige Pixel große Regionen abdecken (Mitte). Hier fällt das Ergebnis der Interpolation plausibel aus.

Abb. 5.9 zeigt einige Ausschnittsvergrößerungen der behandelten Testbilder, wobei jeweils ein interessanter Teil des Originalbildes dem entsprechenden Ausschnitt nach der Artefakt-Filterung gegenübergestellt wird. Dabei ist deutlich zu erkennen, daß das Verfahren die Artefakt-Bildanteile nahezu rückstandslos zu entfernen vermag. Im Falle einer bildbasierten Maskenbildung bleiben allerdings auch nach der Filterung oft noch schwache Detailstrukturen sichtbar, sofern das Objekt, welches die Artefakte verursacht, nicht vollständig aus Metall besteht. Ein Beispiel hierfür ist das Instrument in Abb. 5.9 (oben). Die metallenen Anteile (Stabachse, Kugelmarken) werden sauber gefiltert, ein Teil des Instrumentes bleibt jedoch sichtbar. Durch eine modellbasierte Generierung des Maskenbildes ließe sich dies nur dann verhindern, falls Geometrie und Position des Instrumentes exakt bekannt wären.

Ob der Filter letztlich unerwünschte *Spuren* im resultierenden Bild hinterläßt, hängt von der Art des Bildinhaltes an den jeweiligen Anwendungsstellen ab. In unstrukturierter Umgebung gibt es keinerlei Probleme. Hier ist selbst in der Vergrößerung nicht mehr zu erkennen, an welchen Stellen der Filter zum Einsatz kam. Verdecken die einzelnen Artefakte jedoch Bildteile in Regionen mit ausgeprägter Struktur, z.B. Kantenverläufen, so bleibt die Filterung nicht ohne sichtbare Folgen. Es entstehen i.a. unerwünschte synthetische Strukturen, die auf der anderen Seite aber die Weiterverarbeitung der Fluoroskopieaufnahmen im Rahmen der Registrierung wesentlich weniger erschweren als die Metall-Artefakte ohne Filterung. In Abb. 5.7 (e) bzw. 5.9 (unten) wird dies besonders deutlich. Die dichte Reihung der Metallkugeln des Kalibrierphantoms überlagert sowohl die Grenze zwischen distalem Femurknochen und Oberschenkelgewebe als auch den Übergang von Gewebe zu hellem Hintergrund. An beiden Kantenverläufen kommt es durch die Filterung zu den erwähnten Fehlstrukturierungen.

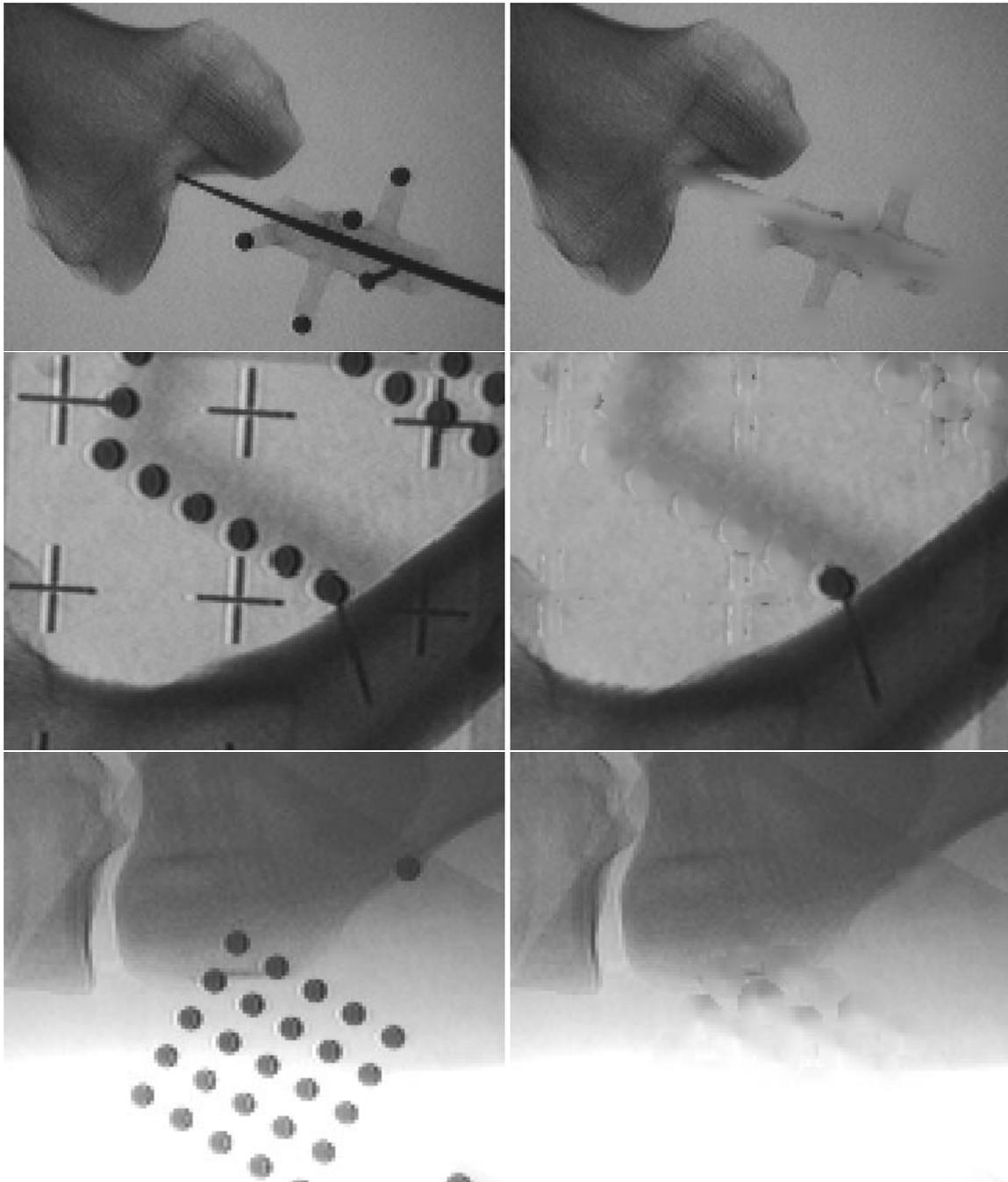


Abbildung 5.9: Gegenüberstellung von ausschnittsvergrößerten Bildteilen vor (links) und nach der Glättungsoperation (rechts). Die dunklen Artefakt-Regionen wurden erfolgreich gefiltert. Vereinzelt entstehen durch die Glättung lokal neue unerwünschte Grauwert-Strukturen. Diese können im Hinblick auf nachfolgende Schritte (Segmentierung, Registrierung) jedoch vernachlässigt werden, da sie stets deutlich kontrastärmer als die störenden Metall-Artefakte ausfallen.

5.3 Lokalisation

Ausgangsbasis für die Lokalisation eines Patientenknöchens stellt eine statisch gegebene *Szene* $[(\mathcal{I}_{x_1}, Cam_1), \dots, (\mathcal{I}_{x_n}, Cam_n)]$ dar – bestehend aus n verzerrungsfreien Fluoroskopieaufnahmen \mathcal{I}_{x_i} sowie den zugehörigen Aufnahme- bzw. Kameraparametern Cam_i , die sich aus der Kalibrierung des Röntgenbildverstärkers ergeben. Diese Aufnahmen werden intraoperativ nacheinander aus verschiedenen Richtungen gewonnen und zeigen jeweils die betrachtete Knochenstruktur. Während und zwischen den einzelnen Aufnahmen darf sich der Patient nicht bewegen. Daher ist es nötig, das entsprechende Körperteil ruhig zu stellen, ggf. auch geeignet starr zu fixieren.

5.3.1 Initialisierung

Die in diesem Abschnitt vorgestellte Registrierung basiert auf dem pixelweisen Vergleich von je zwei Aufnahmen. Wenngleich der eigentliche Algorithmus zur Optimierung der Objektlage und das damit verbundene planare Fehlermaß erst im nächsten Abschnitt vorgestellt werden, so ist es doch bereits an dieser Stelle offensichtlich, daß der Konvergenzradius der Registriertechnik deutlich begrenzt sein muß. Der pixelbezogene Vergleich zweier Aufnahmen, die aus signifikant unterschiedlichen Blickrichtungen auf das Objekt entstanden sind, läßt keinen sinnvollen Rückschluß zu, welche Parameter wie verändert werden sollten, um die aktuell angenommene Lage des Objektes besser an die Bilddaten anzugleichen. Nichtsdestotrotz sind alle drei in Abschnitt 4.2 vorgestellten Verfahren zur Initialisierung geeignet. Es ist lediglich darauf zu achten, daß entsprechend größere Sorgfalt gilt. D.h. je nach gewählter Technik sollte entweder die Relativ-Positionierung des Bildverstärkers, die manuelle Modell-Plazierung oder die Markierung anatomischer Punkte innerhalb der Röntgenbilder entsprechend genauer ausfallen.

5.3.2 Optimierung

Das Ziel der Optimierung besteht darin, diejenige rigide Transformation $\mathbf{p}(R, t)$ zu bestimmen, für die eine maximale Übereinstimmung zwischen allen gegebenen Fluoroskopie-Aufnahmen \mathcal{I}_{x_i} sowie den jeweils korrespondierenden, in Abhängigkeit von \mathbf{p} und Cam_i simulierten Aufnahmen $\mathcal{I}_{s_i}^{\mathbf{p}}$ erreicht wird. Dies kann in Form eines robusten *least-squares*-Ansatzes formuliert werden. Die Maximierung der Übereinstimmung ist dabei gleichbedeutend mit einer Minimierung des folgenden Fehlermaßes:

$$D(\mathbf{p}) = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} (\mathcal{I}_{x_i}(u, v) - \mathcal{I}_{s_i}^{\mathbf{p}}(u, v))^2 \right] \quad (5.4)$$

Die in den Abschnitten 5.2.2 (Kontrast-Adaption) bzw. 5.2.3 (Berücksichtigung von Artefakten) beschriebenen Problematiken erfordern es allerdings, das obige Maß nicht auf der Basis von \mathcal{I}_{x_i} und \mathcal{I}_{s_i} zu berechnen. Stattdessen müssen artefaktbereinigte Realaufnahmen $\mathcal{I}_{\hat{x}_i}$ einerseits mit kontrastadaptierten Simulationsaufnahmen $\mathcal{I}_{\hat{s}_i}$ andererseits verglichen werden. Das zu minimierende Fehlermaß lautet daher:

$$D'(\mathbf{p}) = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} (\mathcal{I}_{\hat{x}_i}(u, v) - \mathcal{I}_{\hat{s}_i}^{\mathbf{p}}(u, v))^2 \right] \quad (5.5)$$

Gl. 5.5 läßt sich aufgrund der einfließenden diskreten Größe $\mathcal{I}_{\hat{s}_i}^{\mathbf{p}}$, deren Wert für jede neue Auswertung der Fehlerfunktion algorithmisch berechnet werden muß, nicht geschlossen differenzieren und daher

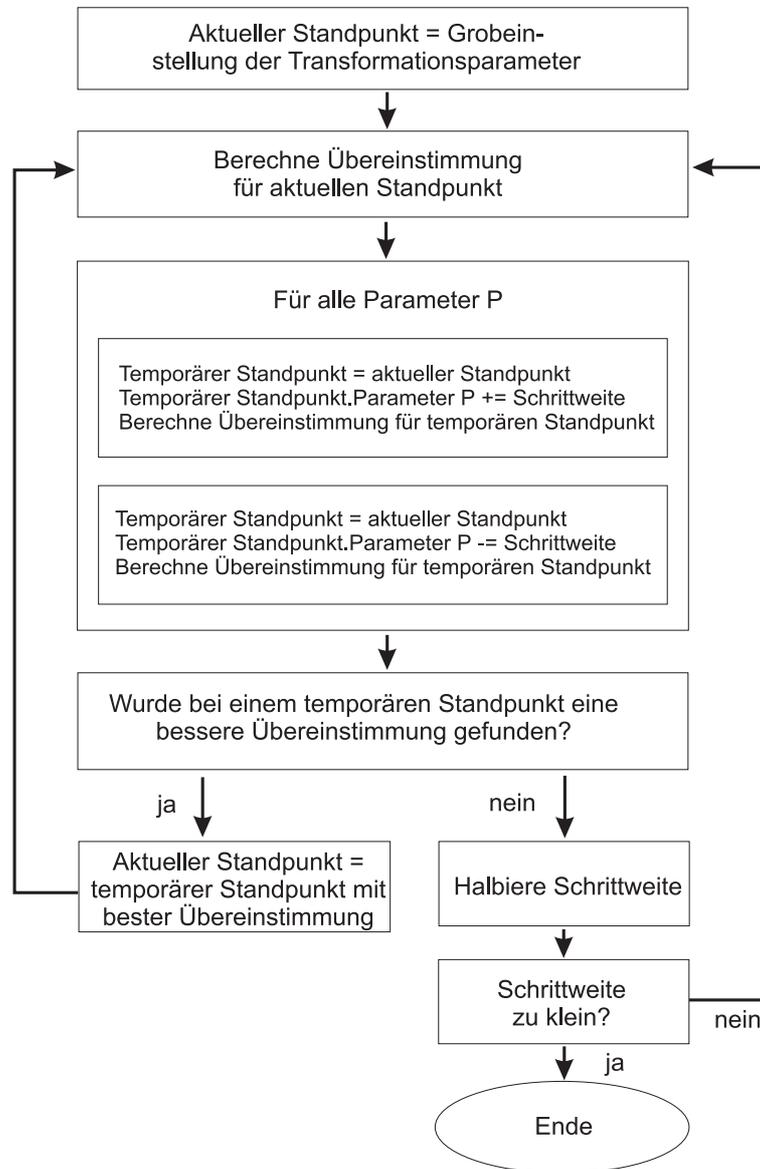


Abbildung 5.10: Alternativer Minimierungsalgorithmus für die intensitätsbasierte Registrierung

auch nicht geschlossen minimieren. Desweiteren kommen einige bekannte Verfahren zur mehrdimensionalen Minimierung wie z.B. *Genetische Algorithmen* [Sch93], *Simulated Annealing* [PTo94] oder auch *Diskrete Levenberg Marquardt-Minimierung* [Mor80] von vornherein für die Minimierung von Gl. 5.5 nicht in Frage. Der Grund hierfür ist, daß der Rechenaufwand für die Minimierung offensichtlich entscheidend durch die Anzahl nötiger Fehlerfunktionsaufrufe dominiert wird. Es gilt daher, die Anzahl an Funktionsaufrufen so niedrig wie möglich zu halten. Die zuletzt aufgezählten Verfahren sind jedoch implizit allesamt auf eine Vielzahl an Auswertungen der Fehlerfunktion angewiesen.

Eine mehrdimensionale Optimierungstechnik, die sowohl ohne Differenzierung auskommt als auch mit der Anzahl an Fehlerfunktionsaufrufen zuverlässig haushaltet, ist das *Powell-Verfahren* [PT⁺94a]. Diese Methode ist daher in der vorliegenden Arbeit erste Wahl für die Minimierung von Gl. 5.5. Das *Powell-Verfahren* ist der bekannteste Vertreter der Klasse sogenannter *direction-set-Verfahren*. Diese Verfahren besitzen die Gemeinsamkeit, daß sie sich für die Minimierung einer n -dimensionalen Funktion intern stets auf eine Sequenz n einfacher eindimensionaler Minimierungen abstützen. Die wesentliche Leistung besteht jeweils darin, eine Menge an geeigneten Richtungen zu finden, entlang derer anschließend die gegebene mehrdimensionale Funktion sukzessive eindimensional minimiert wird. Die Auswahl der Richtungen gilt dabei genau dann als geeignet, falls diese sich bei den aufeinanderfolgenden Minimierungen entlang der einzelnen Richtungen nicht gegenseitig *behindern*, d.h. die Minimierung in einer Richtung darf durch keine Minimierung der anderen Richtungen zunichte gemacht werden. Richtungen, die dies erfüllen, werden gemeinhin als *gegenseitig konjugiert* bezeichnet. Gelingt eine solche Auswahl, so genügt es, die gegebene Funktion entlang jeder der n Richtungen ein einziges Mal zu minimieren, um das Erreichen des gesuchten n -dimensionalen Minimums zu garantieren, sofern die betrachtete Funktion in der Nähe des Minimums eine annähernd quadratische Form besitzt. Ansonsten ist bei wiederholter Minimierung entlang der n Richtungen zumindest eine quadratische Konvergenz gewährleistet [PT⁺94a]. Die genaue von *Powell* vorgeschlagene Vorgehensweise zur Bestimmung einer Menge gegenseitig konjugierter Richtungen wird ausführlich in [PT⁺94a] beschrieben. Die eindimensionale Minimierung, auf die sich das Verfahren intern abstützt, übernimmt das Minimierungsverfahren von Brent, welches auf inverser parabolischer Interpolation basiert [Bre73, PT⁺94b].

Neben dem *Powell-Verfahren* wurde in der vorliegenden Arbeit eine weitere iterative Vorgehensweise zur Minimierung von Gl. 5.5 realisiert und getestet [Fre98]. Ausgehend von einer initialen Lageschätzung \mathbf{p}_0 werden hierbei zuvor festgelegte freie Parameter jeweils in positiver wie negativer Richtung um eine feste Schrittweite verändert. Für jeden der sich daraus ergebenden temporären Parametersätze wird das beschriebene Differenzbild-Fehlermaß $D'(\mathbf{p})$ berechnet. Ergibt sich dabei ein geringerer Fehler als bisher, so werden die zugehörigen Parameterwerte für die folgende Iteration übernommen. Ansonsten bleiben sie unverändert; stattdessen wird die Schrittweite halbiert. Die Iteration setzt sich solange fort, bis die Schrittweite eine vorgegebene Schranke unterschreitet.

Die zu beeinflussenden Parameter ergeben sich einerseits durch die Eulerwinkel-Repräsentation der gesuchten Transformation (drei rotatorische Parameter, vgl. Abschnitt 4.1), andererseits durch den translatorischen Anteil dieser Transformation (weitere drei Parameter). Desweiteren hat es sich als vorteilhaft für die Optimierung erwiesen, zumindest einen weiteren redundanten Parameter einzuführen: den Abstand zwischen einer der gegebenen Kameras und dem CT-Datensatz. Zwar läßt sich eine solche Abstandsänderung prinzipiell auch durch entsprechende simultane Veränderung der drei translatorischen Parameter erzielen. Dem Algorithmus ist dies jedoch ansonsten durch die Auswahl jeweils eines Parameters verwehrt. Trotz der scheinbar willkürlichen Vorgabe freier Parameter hat sich dieses Verfahren im vorliegenden Anwendungsfall als überraschend leistungsfähig erwiesen. Mehrere experimentelle Ver-

gleichstests ergaben, daß die beiden resultierenden Transformationen bei Anwendung des oben beschriebenen Verfahrens sowie Powells Algorithmus in den meisten Fällen sehr gut übereinstimmen. Vereinzelt konvergiert das Verfahren sogar schneller (siehe Abschnitt 5.5).

Bei iterativen Optimierungsverfahren, die wesentlich auf der Auswertung von Bildinformationen – z.B. Pixelintensitäten wie im vorliegenden Fall – beruhen, hat es sich bereits vielfach bewährt, die Bildauflösung durch Skalierung zunächst stark zu reduzieren und im weiteren Verlauf der Optimierung sukzessive bis auf ursprüngliche Originalgröße zu steigern [Lav96b]. Solche Verfahren werden gemeinhin als *Multi-Level*-Verfahren bezeichnet. Die kleine Bildauflösung zu Beginn der Optimierung unterdrückt insbesondere im Bild enthaltene Rauscheffekte. Desweiteren enthält ein in seiner Auflösung stark reduziertes Bild auch keine feinen Detailstrukturen mehr, sondern vermittelt vielmehr einen globalen Eindruck der markantesten Bildinhalte. Da sowohl Rauschen als auch feine Bilddetails insbesondere zu Beginn der Optimierung dazu führen können, daß sich das Verfahren in einem lokalen Minimum der Fehlerfunktion verfängt, sorgt die adaptive Auflösungsreduzierung in der Regel für ein robustes Konvergenzverhalten des entsprechenden Verfahrens.

Im Rahmen der vorliegenden Arbeit ergibt sich durch den *Multi-Level*-Ansatz ein zusätzlicher positiver Nebeneffekt. Da die Größe der während der Optimierung simulierten Röntgenbilder fest an die Auflösung der realen Fluoroskopieaufnahmen gekoppelt ist, reduziert sich auch deren Größe entsprechend. Dies wiederum hat eine enorme Geschwindigkeitssteigerung zur Folge.

Die adaptive Auflösungsreduzierung läßt sich in beide vorgestellten Optimierungsverfahren gleichermaßen gut integrieren. Powells Verfahren wird sukzessive auf die auflösungsreduzierten Bilddaten angewandt, wobei mit steigender Auflösung der Toleranzschwellwert der integrierten eindimensionalen Minimierung (Brents Algorithmus) reduziert wird. Dieser Schwellwert legt fest, bis zu welcher Genauigkeit der Algorithmus das gesuchte Minimum isoliert. D.h. im Verlauf der Iterationen wird mit ansteigender Bildauflösung sukzessive auch mehr Wert auf die Güte des Minimums gelegt. Im Falle des zweiten Optimierungsverfahrens erfolgt die Integration der *Multi-Level*-Strategie, indem die Größe der simulierten und realen Fluoroskopieaufnahmen sukzessive der aktuellen Schrittweite angepaßt wird. Dabei gilt: Je größer die Schrittweite, um so kleiner fallen die aktuellen Bilder aus (Abb. 5.10).

In Abschnitt 5.5 am Ende dieses Kapitels werden mehrere experimentelle Ergebnisse vorgestellt und ausführlich diskutiert.

5.4 Beschleunigte Simulation von Röntgenaufnahmen

Ein offensichtlicher Nachteil der bisher vorgestellten Methode ist der deutlich erhöhte Zeitbedarf für die gesamte Registrierung³ im Vergleich zu alternativen Ansätzen. Wesentliche Ursache hierfür ist die aufwendige intraoperative Berechnung der Vielzahl an simulierten Röntgenaufnahmen durch Ray-Casting. Die beschriebene Vorverarbeitung des CT-Volumens (Segmentierung, Clipping) reduziert den Zeitbedarf für das Ray-Casting zwar bereits signifikant, jedoch benötigt die Auswertung der zu minimierenden Fehlerfunktion nachwievord deutlich mehr Zeit als bei den Vergleichsalgorithmen (siehe Abschnitt 5.5).

Insbesondere die gewählte *Multiview*-Strategie verschärft das Problem der langsamen Simulation, wohingegen bei einem Ansatz wie in [Wee97a, Wee97b] die benötigte Zeit für die Registrierung u.U. im Einzelfall tolerierbar sein kann, da hier jeweils nur eine einzelne Röntgenaufnahme registriert wird.

³Bsp.: Die Registrierung eines L3-Lendenwirbels benötigt unter Verwendung dreier Ansichten auf einem Pentium II ca. 45 Sekunden. Dabei wurde bereits qualitativ gut initialisiert.

Im folgenden wird deshalb eine Technik vorgestellt, die es erlaubt, Röntgenbilder für bestimmte, häufig gewählte Ansichten um ein Vielfaches schneller zu berechnen als durch Ray-Casting.

5.4.1 Ansatz

Der vorgestellte Ansatz basiert darauf, die CT-Daten bereits präoperativ geeignet vorzuverarbeiten, um intraoperativ zeitintensives *Ray-Casting* weitestgehend zu vermeiden. Die Repräsentation der vorverarbeiteten Daten soll dazu erstens möglichst kompakt sein und zweitens insbesondere die schnelle Berechnung eines Röntgenbildes für eine konkret vorgegebene Kamera-Position erlauben. Hierfür wird das nachfolgend beschriebene Konzept für Ansichtsmodelle vorgeschlagen.

Die grundlegende Idee eines Ansichtsmodells besteht darin, einen Teil der insgesamt sechs freien Parameter, die die Transformation zwischen Kamera- und Objekt-Koordinatensystem festlegen, *präoperativ* für bestimmte, *intraoperativ* häufig genutzte Aufnahmerichtungen zu trainieren. Das Training umfaßt dabei zunächst nur Rotationen um eine der beiden, zur Bildebene parallelen Achsen R_x bzw. R_y und führt zu einer kompakten Eigenraum-Repräsentation der gewählten Ansicht. Die restlichen vier nicht trainierten Parameter (R_z , T_x , T_y und T_z) werden intraoperativ planar approximiert. D.h. ausgehend von einem Basisbild, das sich in Abhängigkeit von R_x und R_y aus der Eigenraum-Repräsentation ergibt, werden die räumlichen Transformationen durch entsprechende Transformationen in der Kamera-Bildebene angenähert.

Die beiden folgenden Abschnitte beschreiben zunächst ausführlich die Konstruktion solcher Ansichtsmodelle sowie anschließend deren Einsatz und die damit verbundenen Konsequenzen während der Optimierung.

5.4.2 Ansichtsmodelle

Ein Ansichtsmodell A läßt sich formal durch einen 6-Tupel beschreiben:

$$A = (\mathcal{V}_{CT}, \mathcal{C}_{int}, \mathcal{C}_{ext}, \mathcal{L}_x, \mathcal{L}_y, \mathcal{R}_{eig})$$

Die einzelnen Parameter besitzen dabei folgende Bedeutung:

- \mathcal{V}_{CT} entspricht dem vorverarbeiteten (*Clipping*) CT-Volumendatensatz, der das Objekt von Interesse beinhaltet und aus welchem die Röntgenbilder für das Ansichtstraining generiert werden sollen.
- $\mathcal{C}_{int} = (f, \begin{pmatrix} c_x \\ c_y \end{pmatrix})$
Hiermit werden die internen Kamera-Parameter (Brennweite f , optischer Nullpunkt $\begin{pmatrix} c_x \\ c_y \end{pmatrix}$) für die Erzeugung der Trainingsmenge festgelegt. Diese Parameter bleiben konstant und gelten daher für jedes der generierten Trainingsbilder.
- $\mathcal{C}_{ext} = (\mathbf{O}_{cam}, \mathbf{T}_{cam})$
Die Matrix \mathbf{O}_{cam} legt die Default-Orientierung und der Vektor \mathbf{T}_{cam} den konstanten Abstand d_{cam} der Kamera zum Objekt-Schwerpunkt (im CT-Volumen \mathcal{V}_{CT}) fest. Die Kamera ist dabei stets exakt auf den Objekt-Schwerpunkt ausgerichtet, d.h. es gilt $\mathbf{T}_{cam} = [0 \ 0 \ d_{cam}]^T$ (Abb. 5.12).

- $\mathcal{L}_x = (\Delta\varphi^-, \Delta\varphi^+, \phi)$
 \mathcal{L}_x legt fest, in welchem Bereich die Rotation um die R_x -Achse während des Trainings variiert wird. Die Angaben $\Delta\varphi^-$ und $\Delta\varphi^+$ beziehen sich dabei jeweils auf die Default-Orientierung der Kamera. Es werden insgesamt ϕ äquidistant auf diesen Bereich verteilte Winkel für das Training herangezogen (Abb. 5.12).
- $\mathcal{L}_y = (\Delta\vartheta^-, \Delta\vartheta^+, \theta)$
 Analog zu \mathcal{L}_x legt \mathcal{L}_y die Diskretisierung der Rotation um R_y fest (Abb. 5.12).
- $\mathcal{R}_{eig} = (\bar{\mathbf{x}}, \{\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{k-1}\}, \mathcal{S})$
 \mathcal{R}_{eig} beinhaltet eine kompakte Eigenraum-Repräsentation der gewählten Ansicht, d.h. \mathcal{R}_{eig} ermöglicht es, für konkrete Parameter (r_x, r_y) innerhalb der während des Trainings berücksichtigten Grenzen ein zugehöriges Röntgenbild zu rekonstruieren.

In den folgenden vier Unterabschnitten *Erzeugung von Trainingsbildern*, *Eigenraum-Repräsentation*, *Interpolation im Eigenraum* sowie *Planare Approximation von 3D-Transformationen* wird detailliert auf die Konstruktion eines Ansichtsmodells und damit die Verwendung bzw. Generierung der genannten Parameter eingegangen.

5.4.2.1 Erzeugung von Trainingsbildern

Die Erzeugung von Trainingsröntgenbildern erfolgt unter Anwendung des bereits in Abschnitt 5.2.1 vorgestellten *Ray-Casting*-Verfahrens auf das vorverarbeitete (*geclippte*) CT-Volumen \mathcal{V}_{CT} .

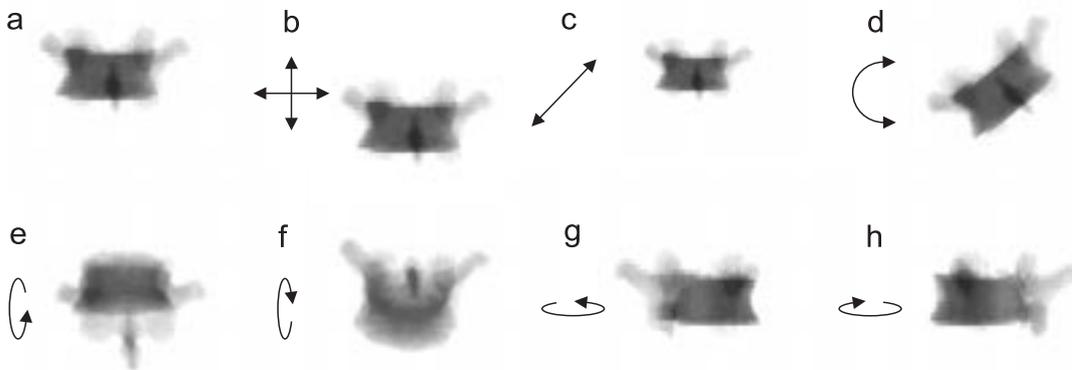


Abbildung 5.11: Unterschiedliche Auswirkungen von 3D-Transformationen auf das resultierende Projektionsbild: (a) Ausgangsbild, (b) Verschiebung parallel, (c) senkrecht zur Bildebene, (d) Rotation um optische Achse, (e,f) Rotation um X-Achse, (g,h) Rotation um Y-Achse. Die Ansichten (b),(c) und (d) lassen sich – ausgehend von (a) – planar approximieren.

Eine einfache Überlegung zeigt, daß es nicht nötig ist, während des Trainings alle sechs Freiheitsgrade der Transformation zwischen Kamera- und Objekt-Koordinatensystem zu berücksichtigen. Die

Überlegung basiert auf der realistischen Annahme, daß die maximale Ausdehnung des betrachteten Objektes in Richtung der Kamera (ΔZ_{obj}) deutlich geringer ausfällt als der Abstand zwischen Kamera und Objekt⁴: $\Delta Z_{obj} \gg d_{cam}$. Unter dieser Voraussetzung verändert eine kleine räumliche Verschiebung des Objektes (ohne Rotation) dessen Erscheinungsform in der Bildebene der Kamera nur unwesentlich. Dasselbe gilt für eine Rotation um die optische Achse R_z . Rotationen um die Achsen R_x bzw. R_y hingegen haben signifikante Auswirkungen auf das Bild (Abb. 5.11). Aus diesem Grund müssen diese beiden rotatorischen Parameter für die Modellbildung bzw. bei der Generierung der hierfür benötigten Trainingsbilder speziell berücksichtigt werden. Die virtuelle Kamera wird dabei für das gesamte Training mit konstanten internen Parametern ($f, [c_x c_y]^T$) assoziiert. Ihre Blickrichtung ist zentriert auf den Objekt-Schwerpunkt innerhalb des CT-Volumens ausgerichtet. Ferner bleibt der Abstand zwischen Kamera-Ursprung und Objekt-Schwerpunkt konstant d_{cam} . Die Kamera besitzt daher zunächst lediglich die Freiheit, sich auf einer Kugeloberfläche mit Radius d_{cam} um den Objekt-Schwerpunkt herum zu bewegen.

Die weitere Einschränkung dieser Freiheit legt den Blickbereich des zu bildenden Ansichtsmodells fest. Zunächst wird hierfür mittels der Matrix \mathbf{O}_{cam} die Default-Blickrichtung der Kamera festgelegt. Ausgehend von dieser Grundstellung werden anschließend zwei Rotationsparameter r_x und r_y variiert (Abb. 5.12):

$$\begin{aligned} \Delta\varphi^- &\leq r_x \leq \Delta\varphi^+ \\ \Delta\vartheta^- &\leq r_y \leq \Delta\vartheta^+ \end{aligned}$$

Innerhalb dieser Intervalle werden insgesamt ϕ (bzw. θ) gleichmäßig verteilte diskrete Werte φ_i (bzw. ϑ_j) ausgewählt. Für jede mögliche Parameter-Kombination (φ_i, ϑ_j) <sub>$i=1\dots\phi$
 $j=1\dots\theta$</sub> wird anschließend das zugehörige Röntgenbild generiert (Abb.5.13).

Um die weiteren Herleitungen zu vereinfachen, wird im folgenden für ein zweidimensionales Bild (Breite w , Höhe h) anstatt der zeilen- und spaltenorientierten Notation $\mathcal{I}_S(i, j)$ stets ein Vektor \mathbf{x} verwendet. Dieser Vektor besitzt die Dimension $w \times h$ und ergibt sich durch einfache, zeilenweise Linearisierung der Bildmatrix $\mathcal{I}_S(i, j)$, d.h. es gilt: $\mathbf{x}[i * w + j] = \mathcal{I}_S(i, j)$.

Bezeichnet man das Trainingsbild für die Parameter-Kombination (φ_i, ϑ_j) mit $\mathbf{x}_{i,j}$, so ergibt sich insgesamt die folgende Menge an Trainingsröntgenbildern:

$$\left\{ \begin{array}{l} \mathbf{x}_{1,1}, \mathbf{x}_{1,2}, \dots, \mathbf{x}_{1,\theta}, \\ \mathbf{x}_{2,1}, \mathbf{x}_{2,2}, \dots, \mathbf{x}_{2,\theta}, \\ \dots \\ \mathbf{x}_{\phi,1}, \mathbf{x}_{\phi,2}, \dots, \mathbf{x}_{\phi,\theta} \end{array} \right\} \quad (5.6)$$

5.4.2.2 Eigenraum-Repräsentation

Drei Anforderungen an die Repräsentation des zu bildenden Ansichtsmodells bestimmen im wesentlichen die weitere Vorgehensweise:

⁴Alle in der vorliegenden Arbeit verwendeten Knochen-Objekte (Femur, Tibia, Fibula sowie diverse Lenden- und Brust-Wirbelkörper) besitzen für eine sinnvolle Durchleuchtungsrichtung einen maximalen Durchmesser von wenigen Zentimetern. Demgegenüber liegt der übliche Abstand des Objektes von der Röntgenquelle (Kameraursprung) bei ungefähr 60 cm.

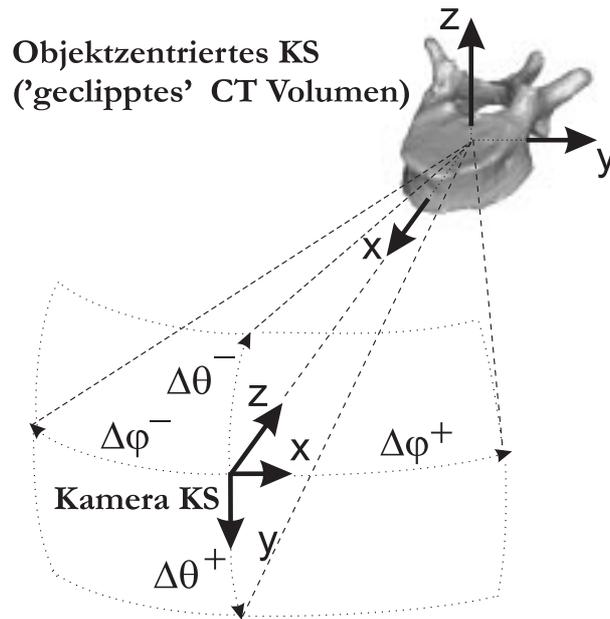


Abbildung 5.12: Xray-View-Modell: Einschränkung der Ansicht

1. Die Darstellung muß kompakt sein, d.h. es ist i.a. nicht sinnvoll, präoperativ eine sehr große Anzahl an Trainingsbildern zu generieren und diese intraoperativ während der Optimierung in Abhängigkeit von einer benötigten Parameter-Kombination (φ_i, ϑ_j) ständig nachzuladen.
2. Die Repräsentation sollte die intraoperative Simulation von Röntgenbildern für innerhalb der trainierten Ansicht liegende Konfigurationen stark beschleunigen. Dies ist die eigentliche Intention für die Bildung von Ansichtsmodellen, so daß sich ein erheblicher Geschwindigkeitsvorteil im Vergleich zum *Ray-Casting* ergeben sollte.
3. Die Generierung soll dabei auch für (φ, ϑ) -Konfigurationen, die nicht explizit in der Trainingsmenge enthalten waren, funktionieren. Hierfür ist eine geeignete Interpolationsmethode erforderlich.

Die im folgenden beschriebene Eigenraum-Repräsentation folgt im wesentlichen dem Ansatz aus [Mur95], da dieser jede einzelne der drei Anforderungen erfüllt.

Ausgehend von der Trainingsmenge (5.6) wird zunächst ein *Durchschnittsbild-Vektor* $\bar{\mathbf{x}}$ berechnet, der sich dadurch ergibt, daß auf jeder einzelnen Pixelposition der Durchschnitt über alle Trainingsbilder gebildet wird:

$$\bar{\mathbf{x}} = \frac{1}{\phi\theta} \sum_{\substack{\varphi=1\dots\phi \\ \vartheta=1\dots\theta}} \mathbf{x}_{\varphi,\vartheta} \quad (5.7)$$

Dieses Durchschnittsbild $\bar{\mathbf{x}}$ wird anschließend von jedem einzelnen Trainingsbild $\mathbf{x}_{\varphi,\vartheta}$ subtrahiert. Dadurch ergibt sich eine neue Menge von Bildern:

$$\{\mathbf{x}_{1,1} - \bar{\mathbf{x}}, \mathbf{x}_{1,2} - \bar{\mathbf{x}}, \dots, \mathbf{x}_{\phi,\theta} - \bar{\mathbf{x}}\} \quad (5.8)$$

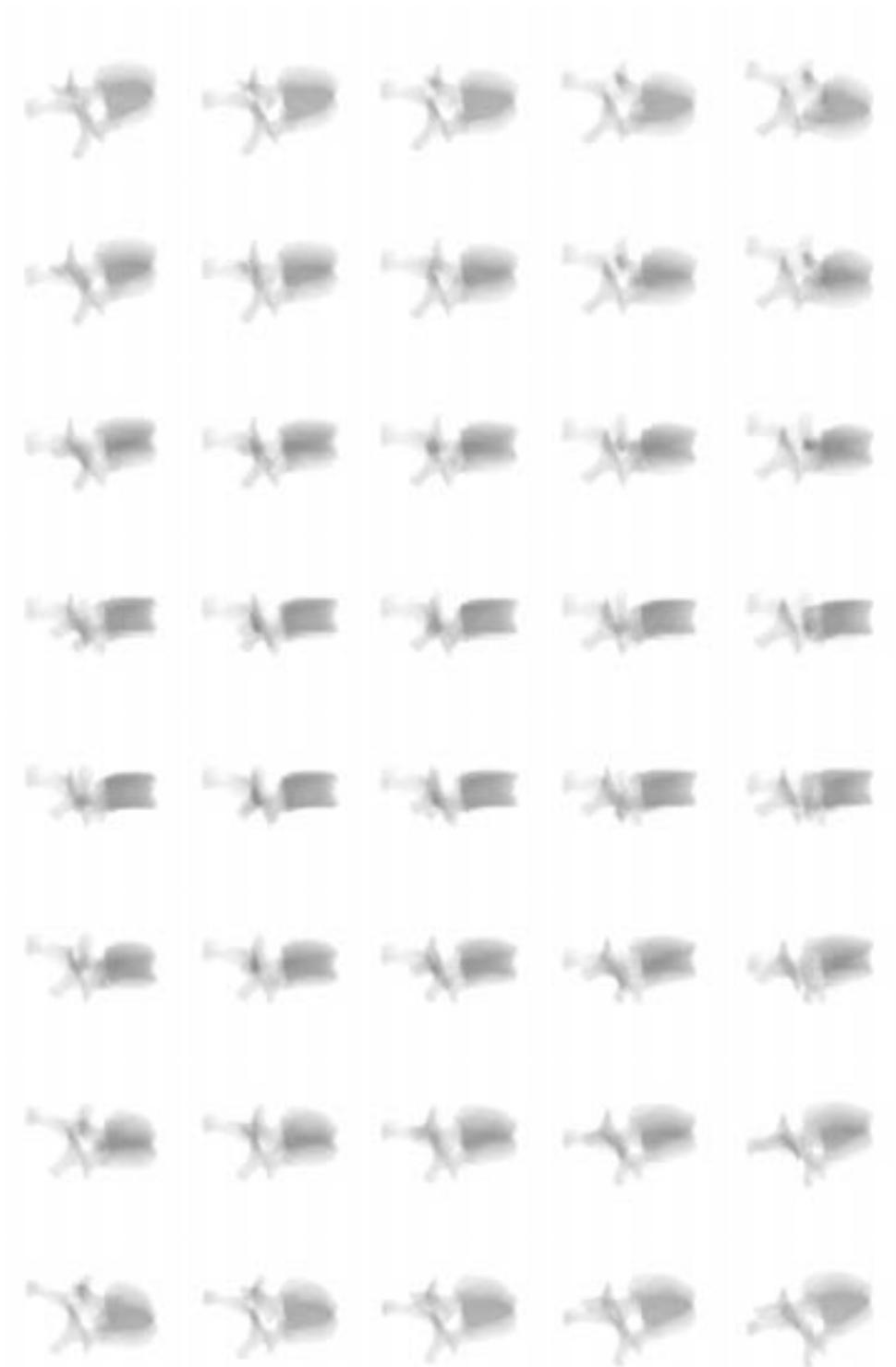


Abbildung 5.13: Trainingsmenge für Ansichtstraining (Winkelbereich: $\varphi = -\frac{\pi}{8} \dots \frac{\pi}{8}$, $\vartheta = -\frac{\pi}{4} \dots \frac{\pi}{4}$)

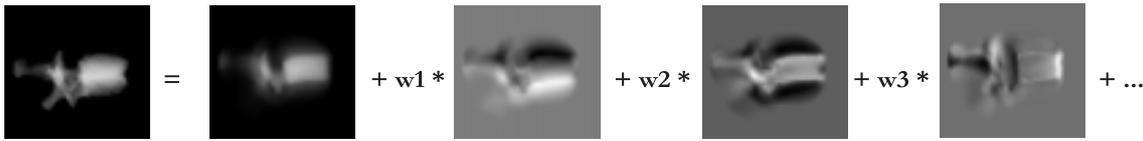


Abbildung 5.14: Bildapproximation durch Linear-Kombination

Jedes dieser Differenzbilder sagt aus, inwieweit das zugehörige Original-Trainingsbild vom Durchschnitt aller Trainingsbilder abweicht. Genau diese Eigenschaft kann nun für eine Hauptachsentransformation im Sinne einer starken Variation genutzt werden (*principal component analysis, PCA*, [Jol86]). Durch Zusammenfassung der einzelnen Differenzbild-Vektoren in der Matrix \mathbf{X} – jedes $\mathbf{x}_{\varphi, \vartheta}$ bildet eine Zeile – ergibt sich zunächst die Kovarianzmatrix \mathbf{C} :

$$\mathbf{C} = \mathbf{X}\mathbf{X}^T \tag{5.9}$$

Die Eigenwerte λ_i sowie die zugehörigen Eigenvektoren \mathbf{e}_i ergeben sich durch Lösung des bekannten linearen Eigenwert-Gleichungssystems:

$$\mathbf{C}\mathbf{e}_i = \lambda_i\mathbf{e}_i \quad \text{bzw.} \quad \mathbf{\Lambda} = \mathbf{E}^T\mathbf{C}\mathbf{E} \tag{5.10}$$

Dabei enthält die Diagonalmatrix $\mathbf{\Lambda} = \text{diag}(\lambda_0, \dots, \lambda_{N-1})$ auf Ihrer Diagonalen sämtliche Eigenwerte. Die Matrix \mathbf{E} ist entsprechend zeilenweise aus den Eigenvektoren \mathbf{e}_i aufgebaut: $\mathbf{E} = [\mathbf{e}_0 \mathbf{e}_1 \dots \mathbf{e}_{N-1}]^T$

Diese orthonormale Matrix definiert eine Koordinatentransformation, die lineare Korrelationen, die in den Trainingsdaten enthalten sind, beseitigt. Dabei werden die invarianten Unterräume von \mathbf{C} explizit dargestellt.

Durch Verwendung der linear unabhängigen Eigenvektoren als Basis läßt sich ein Bild-Vektor \mathbf{x} nun auch wie folgt darstellen:

$$\mathbf{x} = \bar{\mathbf{x}} + \sum_{i=0}^{N-1} (\omega_i \mathbf{e}_i) \tag{5.11}$$

Im Sinne einer *kompakten* Modellrepräsentation wäre dadurch aber noch nichts gewonnen, da nachwievor N Eigenvektoren benötigt würden, von denen jeder die Größe eines Trainingsbildes hat.

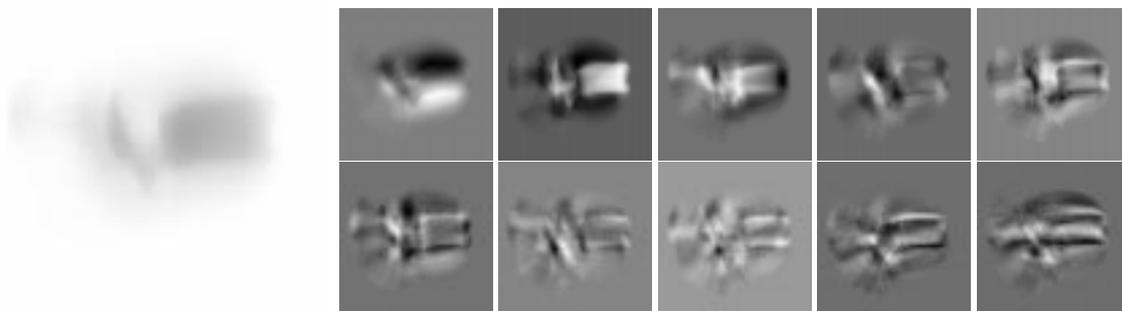


Abbildung 5.15: Durchschnittsbild und Eigenvektoren/bilder

k	Eigenwert λ_k	$\sum_1^k \lambda_i$	$\sum_1^k \lambda_i / \sum_1^N \lambda_i$	k	Eigenwert λ_k	$\sum_1^k \lambda_i$	$\sum_1^k \lambda_i / \sum_1^N \lambda_i$
1	1.44e+08	1.44e+08	33.74%	21	2.15e+06	4.07e+08	95.25%
2	6.75e+07	2.12e+08	49.55%	22	2.14e+06	4.09e+08	95.75%
3	6.36e+07	2.75e+08	64.43%	23	1.91e+06	4.11e+08	96.20%
4	2.30e+07	2.98e+08	69.80%	24	1.86e+06	4.13e+08	96.63%
5	1.88e+07	3.17e+08	74.20%	25	1.71e+06	4.15e+08	97.04%
6	1.61e+07	3.33e+08	77.97%	26	1.56e+06	4.16e+08	97.40%
7	1.36e+07	3.47e+08	81.15%	27	1.51e+06	4.18e+08	97.75%
8	8.12e+06	3.55e+08	83.05%	28	1.42e+06	4.19e+08	98.08%
9	5.99e+06	3.61e+08	84.45%	29	1.40e+06	4.21e+08	98.41%
10	5.80e+06	3.67e+08	85.81%	30	1.31e+06	4.22e+08	98.72%
11	5.59e+06	3.72e+08	87.12%	31	1.27e+06	4.23e+08	99.02%
12	5.39e+06	3.78e+08	88.38%	32	1.16e+06	4.24e+08	99.29%
13	4.83e+06	3.83e+08	89.51%	33	1.09e+06	4.25e+08	99.54%
14	4.47e+06	3.87e+08	90.56%	34	1.05e+06	4.26e+08	99.79%
15	3.51e+06	3.91e+08	91.38%	35	9.06e+05	4.27e+08	100.00%
16	3.38e+06	3.94e+08	92.17%	36	5.68e+02	4.27e+08	100.00%
17	3.31e+06	3.97e+08	92.94%	37	2.27e+02	4.27e+08	100.00%
18	2.67e+06	4.00e+08	93.57%	38	9.13e+01	4.27e+08	100.00%
19	2.60e+06	4.02e+08	94.17%	39	5.13e+01	4.27e+08	100.00%
20	2.45e+06	4.05e+08	94.75%	40	4.03e+01	4.27e+08	100.00%

Tabelle 5.1: Zugehörige Eigenwerte für die Eigenvektoren aus Abb. 5.15 (Ergebnis der PCA für die Trainingsdaten aus Abb. 5.13)

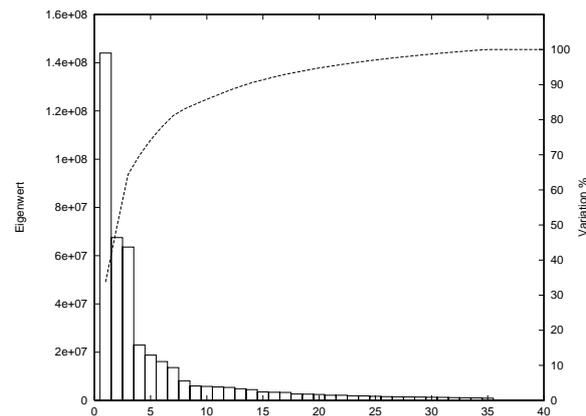


Abbildung 5.16: Eigenwert-Verteilung mit akkumuliertem Variationsanteil

Es zeigt sich in der Praxis jedoch, daß die einzelnen Eigenvektoren zu der Summe in Gl. 5.11 unterschiedlich stark beitragen. Insbesondere genügen i.a. bereits wenige Eigenvektoren, um nahezu den gesamten in den ursprünglichen Trainingsdaten enthaltenen Variationsspielraum zu repräsentieren (vgl. z.B. die Anwendungen in [MPK94, Coo95c, Nas94]). Wieviel der gesamten Variation auf einen konkreten einzelnen Eigenvektor \mathbf{e}_i entfällt, korreliert direkt mit der Größe des zugehörigen Eigenwertes λ_i . Unter Vorgabe der für das zu bildende Modell gewünschten Variation $\tau \in [0.0 \dots 1.0]$ (entsprechend 0...100% Variation) läßt nach folgender Formel die hierfür mindestens nötige Anzahl k der signifikantesten Eigenvektoren – die Vektoren mit den größten zugehörigen Eigenwerten – ermitteln:

$$\min_k \left(\frac{\sum_{i=0}^{k-1} \lambda_i}{\sum_{i=0}^{N-1} \lambda_i} \geq \tau \right) \quad (5.12)$$

Es gibt eine Vielzahl numerischer Methoden zur Berechnung von \mathbf{A} und \mathbf{E} , ausgehend von Gl. 5.10. Aus mehreren Gründen verbietet sich im vorliegenden Fall jedoch der Einsatz von Standardmethoden zur Berechnung *aller* Eigenwerte und -vektoren (z.B. Jacobi-Rotationen, [PTo94]). Erstens besitzt die Kovarianzmatrix \mathbf{C} eine nicht zu unterschätzende Größe⁵, was einen unnötigen Aufwand für die vollständige Berechnung von \mathbf{A} und \mathbf{E} zur Folge hätte, da lediglich $k \ll N$ Eigenvektoren benötigt werden⁶. Zweitens ist die Anzahl an Trainingsbildern in der Praxis stets deutlich geringer als die Dimension der Bildvektoren ($N \gg \phi\theta$). Dies hat zur Folge, daß \mathbf{C} singular ist und sich deshalb sowieso nur die ersten n ($k \leq n < \phi\theta$) Eigenwerte berechnen lassen. Methoden für die sukzessive Berechnung einzelner signifikanter Eigenwerte und -vektoren (z.B. *singular value decomposition (SVD)*) werden in [PTo94, Mur95] beschrieben.

Abbildung 5.15 zeigt das Durchschnittsbild sowie die 10 signifikanten Eigenvektoren (Signifikanz von links oben nach rechts unten abnehmend) für die Trainingsdaten aus Abbildung 5.13. Tabelle 5.1 und Abb. 5.16 enthalten eine Aufstellung der Eigenwerte sowie deren Anteil an der gesamten Variation derselben Daten.

Bei Vorgabe der gewünschten Variation τ kann das resultierende Ansichtsmodell umso kompakter gehalten werden, je weniger Variation in der ursprünglichen Trainingsmenge vorhanden ist. D.h. es ergibt sich ein geringeres k bzw. es sind vergleichsweise weniger Eigenwerte bzw. Eigenvektoren zu speichern.

Die reduzierte Basis \mathbf{E}_k erlaubt die approximative Darstellung eines Bild-Vektors \mathbf{x} (siehe auch Abb. 5.14):

$$\mathbf{x} \approx \bar{\mathbf{x}} + \sum_{i=0}^{k-1} (\omega_i \mathbf{e}_i) \quad (5.13)$$

Diese lineare Approximation eines Bild-Vektors erlaubt bei beliebiger Wahl der Gewichte ω_i zunächst noch die Erzeugung einer Vielzahl unerwünschter Bild-Vektoren. Das Ziel einer geeigneten Repräsentation ist es nun, die Wahl der ω_i derart einzuschränken, daß sich nur Bildvektoren, die den gewählten Trainingsbildern ähnlich sind, approximieren lassen.

⁵Bsp.: Bereits für eine Trainingsbildanzahl von lediglich 125 enthält die Kovarianzmatrix \mathbf{C} insgesamt $125^2 = 15.625$ Elemente.

⁶Die Zeit-Komplexität für die Berechnung von Eigenwerten und -vektoren mit Jacobi-Rotationen liegt in der Größenordnung $\mathbf{O}(n^3)$.

5.4.2.3 Interpolation im Eigenraum

Die Gewichte ω_i in Gl. 5.13 können als die Koordinaten eines Punktes \mathbf{m} innerhalb des Eigenraums $\mathbf{E}_k = [\mathbf{e}_0 \mathbf{e}_1 \dots \mathbf{e}_{k-1}]^T$ interpretiert werden. Dieser Punkt $\mathbf{m} = (\omega_0, \omega_1, \dots, \omega_{k-1})$ entspricht damit direkt dem zugehörigen approximierten Bild-Vektor \mathbf{x} und läßt sich durch Projektion von \mathbf{x} in den Eigenraum \mathbf{E}_k bestimmen. Diese Projektion wird für jeden Trainingsbild-Vektor $\mathbf{x}_{\varphi, \vartheta}$ durchgeführt, wodurch sich eine Punktmenge $\{\mathbf{m}_{\varphi, \vartheta}\}_{\substack{\varphi=1 \dots \phi \\ \vartheta=1 \dots \theta}}$ ergibt:

$$\mathbf{m}_{\varphi, \vartheta} = \mathbf{E}_k(\mathbf{x}_{\varphi, \vartheta} - \bar{\mathbf{x}}) \quad (5.14)$$

Es läßt sich zeigen, daß der Abstand zweier solcher Punkte im Eigenraum ein direktes Maß für die Korrelation der beiden zugehörigen Trainingsbilder ist [Mur95]. D.h. umso ähnlicher zwei Bilder sind, desto näher liegen auch ihre Projektionen im Eigenraum beieinander. Unter der Annahme, daß sich zwei Bilder innerhalb der Trainingsmenge, die benachbarte Ansichten des betrachteten Objektes zeigen, nur geringfügig voneinander unterscheiden, liegen damit insbesondere die Eigenraum-Projektionen von Trainingsbildern für ähnliche Parameter (φ, ϑ) sehr eng beisammen. Die Annahme ist im vorliegenden Anwendungsfall gerechtfertigt, da die Ansichten mit einer sehr kleinen Schrittweite (z.B. 5°) abgetastet werden. Außerdem sind kritische Problemfälle, die sich z.B. im Zusammenhang mit einer hochfrequenten Oberflächentextur des Objektes ergeben würden [Mur95], bei Röntgenbildern nicht zu erwarten.

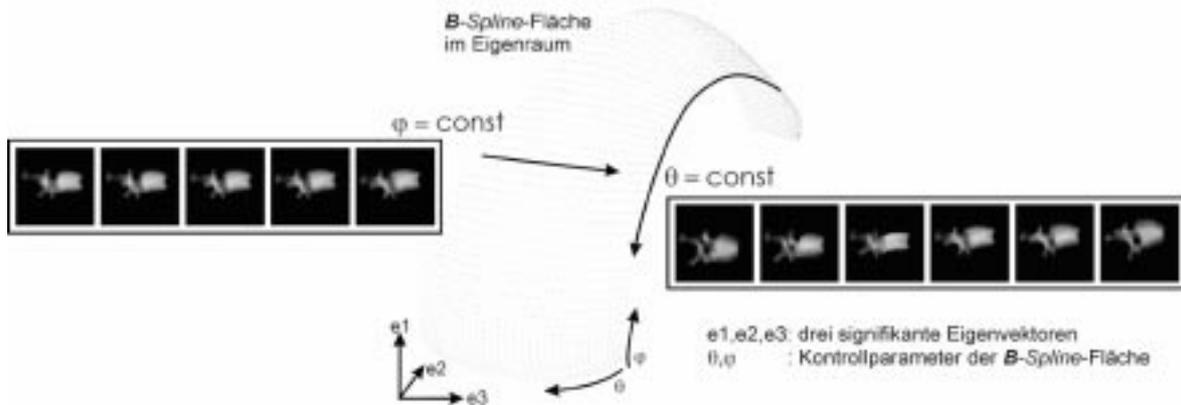


Abbildung 5.17: Parameter-Interpolation: B-Spline-Fläche im Eigenraum

Diese wichtige Korrelations-Eigenschaft der Eigenraum-Darstellung führt dazu, daß die Punktmenge $\mathbf{m}_{\varphi, \vartheta}$ einen *stetigen, zusammenhängenden* Teilraum innerhalb des Eigenraums beschreibt. Da zur Generierung der Trainingsbilder exakt zwei Parameter (φ, ϑ) variiert wurden, wird für die Modell-Repräsentation die Dimension dieses Teilraums ebenfalls auf zwei gesetzt. Der Teilraum entspricht damit einer Fläche und läßt sich aus der Punktmenge durch Interpolation direkt bestimmen (Abb. 5.17). In der vorliegenden Arbeit wurde hierfür eine bikubische Splinefunktion $S(\nu_{r_x}, \nu_{r_y})_{\nu_{r_x}, \nu_{r_y} \in [0 \dots 1]}$ verwendet [Bro89, PTo94].

Die Trainingsparameter r_x und r_y lassen sich durch lineare Abbildung der Intervalle $[\Delta\varphi^- \dots \Delta\varphi^+]$ bzw. $[\Delta\vartheta^- \dots \Delta\vartheta^+]$ auf das Intervall $[0 \dots 1]$ direkt in Spline-Parameter umrechnen:

$$\nu_{r_x}(r_x) = \frac{r_x - \Delta\varphi^-}{\Delta\varphi^+ - \Delta\varphi^-} \quad \text{bzw.} \quad \nu_{r_y}(r_y) = \frac{r_y - \Delta\vartheta^-}{\Delta\vartheta^+ - \Delta\vartheta^-} \quad (5.15)$$

Auf diese Weise kann aus einem Parameterpaar r_x, r_y unter Verwendung der Spline-Funktion $S()$ auf den korrespondierenden Punkt m_{r_x, r_y} innerhalb des Eigenraums \mathbf{E}_k geschlossen werden. Die Koordinaten dieses Punktes liefern schließlich die Gewichte ω_i , die, in Gl. 5.13 eingesetzt, die gewünschte Approximation eines Bild-Vektors ermöglichen.

5.4.2.4 Planare Approximation von 3D-Transformationen

Das in den vorangegangenen Abschnitten vorgestellte Ansichtstraining sowie die sich daraus ergebende Eigenraum-Repräsentation des Modells beinhaltet bisher lediglich zwei der insgesamt sechs Freiheitsgrade der relativen Objekt-Lage: Rotationen um eine der beiden Achsen R_x und R_y . Das Training des Ansichtmodells erfolgt dabei zeitlich getrennt von dessen eigentlichem Einsatz, der Registrierung, d.h. präoperativ. Intraoperativ liefert das Modell ein simuliertes Röntgenbild für eine vorgegebene Objekt-lage, wobei die beiden gelernten Rotationsparameter im Rahmen des präoperativ trainierten Spielraums variieren können. Die zuletzt beschriebene Interpolation im Eigenraum erlaubt stets die Rekonstruktion der entsprechenden Röntgenaufnahme. Die restlichen vier Freiheitsgrade, d.h. eine Rotation um R_z sowie eine beliebige räumliche Verschiebung des Objektes (T_x, T_y, T_z), bleiben dabei zunächst unberücksichtigt. Sie müssen exakt den zum Zeitpunkt des Trainings vereinbarten Konstanten entsprechen.

Das Modell muß nun dahingehend erweitert werden, daß es mit einer begrenzten Variation dieser vier 3D-Transformationen ebenfalls umgehen kann. Wie bereits auf S. 62 erwähnt wurde, lassen sich diese Transformationen, ausgehend von dem rekonstruierten Bild, unter gewissen Voraussetzungen durch geeignete Transformationen des rekonstruierten Bildes planar approximieren.

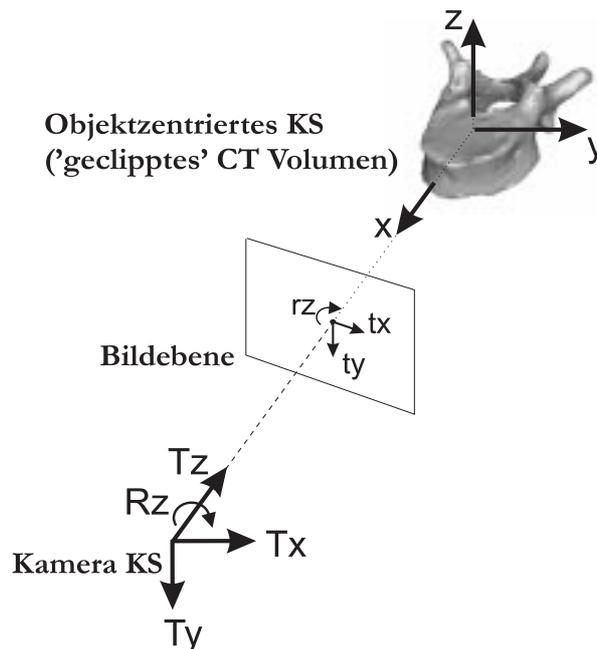


Abbildung 5.18: Planare Approximation von R_z, T_x, T_y und T_z

Das aus der Eigenraum-Darstellung rekonstruierte Bild korrespondiert direkt mit der Objekt-lage $(R_{x_0} + \Delta r_x, R_{y_0} + \Delta r_y, R_{z_0}, X_0, Y_0, Z_0)$, wobei die letzten vier Parameter noch nicht der gewünschten

Ziellage entsprechen. Im folgenden wird für jeden dieser Parameter eine entsprechende Approximationsvorschrift hergeleitet, die festlegt, wie sich die Bildpunkte des Ausgangsbildes in Abhängigkeit von einer Veränderung des betrachteten Parameters verschieben. Diese Approximationsformeln ermöglichen es schließlich, ein neues Bild zu generieren, das eine Näherung der Ansicht für die vorgegebene Ziellage repräsentiert.

Grundlage der Approximation ist das sogenannte *schwach-perspektivische* Kameramodell [XZo96]. Ausgehend von dem in Abschnitt 3.1 eingeführten perspektivischen Lochkameramodell werden dabei vor der Projektion alle individuellen Z -Koordinaten der Objektpunkte durch die Tiefe Z_c des Objektschwerpunktes ersetzt: $Z = Z_c = \text{const.}$ Das Kameramodell wird dadurch linear in X und Y :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + f \frac{1}{Z_c} \begin{pmatrix} X \\ Y \end{pmatrix} \quad (5.16)$$

Aus Sicht des Objektes führt diese Vereinfachung geometrisch zu einer zweistufigen Projektion. Zunächst werden alle Objektpunkte parallel in die zur Bildebene parallel verlaufende Ebene durch Z_c projiziert. Im Anschluß daran erfolgt eine perspektivische Projektion dieser Ebene in die Bildebene. Aufgrund der nun fehlenden Tiefe reduziert sich diese Projektion dabei auf eine gleichmäßige Skalierung des Bildinhaltes.

Die einzelnen Approximationsformeln ergeben sich wie folgt, wobei jeweils $(u, v)^T$ eine Pixelposition im Ausgangsbild, $(u', v')^T$ entsprechend die angenäherte Pixelposition im Zielbild, $(c_x, c_y)^T$ den optischen Nullpunkt sowie f die Brennweite darstellen:

1. Eine beliebige Translation parallel zur Bildebene $(\Delta X, \Delta Y)$ läßt sich durch eine mit $\frac{f}{Z_c}$ skalierte planare Verschiebung innerhalb der Bildebene annähern:

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} + \frac{f}{Z_c} \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \quad (5.17)$$

2. Eine Translation senkrecht zur Bildebene (ΔZ) wirkt sich in der Bildebene im wesentlichen als planare Skalierung mit optischem Nullpunkt als Zentrum aus:

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \frac{1}{1 + \frac{\Delta Z}{Z_c}} \left[\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} c_x \\ c_y \end{pmatrix} \right] \quad (5.18)$$

3. Rotationen um die optische Achse (Δr_z) lassen sich durch planare Rotationen in der Bildebene um den optischen Nullpunkt ersetzen⁷:

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \mathbf{R}_{2D}(\Delta r_z) \left[\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} c_x \\ c_y \end{pmatrix} \right] \quad (5.19)$$

Unter Beachtung der vorgeschriebenen Transformations-Reihenfolge (Rotation vor Translation entsprechend $P' = \mathbf{O}P + \mathbf{T}$) lassen sich die einzelnen Approximationsterme (Gl. 5.17, 5.18 und 5.19) kompakt zu einer einzelnen Formel zusammenfassen:

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \frac{1}{1 + \frac{\Delta Z}{Z_c}} \mathbf{R}_{2D}(\Delta r_z) \left[\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} c_x \\ c_y \end{pmatrix} \right] + \frac{f}{Z_c} \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \quad (5.20)$$

⁷ $\mathbf{R}_{2D}(\alpha)$ repräsentiert dabei die 2×2 -Rotationsmatrix für eine α -Grad-Drehung um den Nullpunkt.

Ausgehend von einem aus der Eigenraum-Darstellung rekonstruierten Bild $\mathcal{I}_{eig}(u, v)$ ergibt sich dadurch das endgültige Bild $\mathcal{I}_{res}(u', v')$ für die vorgegebenen Lageparameter $(R_{x_0} + \Delta r_x, R_{y_0} + \Delta r_y, R_{z_0} + \Delta r_z, X_0 + \Delta X, Y_0 + \Delta Y, Z_0 + \Delta Z)$.

Im Hinblick auf ihre algorithmische Umsetzung besitzt Gleichung 5.20 jedoch zwei entscheidende Nachteile. Indem für jeden Bildpunkt (u, v) des Ausgangsbildes der korrespondierende Pixel (u', v') im resultierenden Bild berechnet wird, bleiben einige Pixel in $\mathcal{I}(u', v')$ unbesetzt, da kein Pixel im Ausgangsbild direkt auf sie abgebildet wird. Andere Bildpunkte wiederum werden unnötig berechnet, da sie letztendlich an Positionen außerhalb der Bildgrenzen des Zielbildes transformiert werden. Diese beiden Probleme lassen sich umgehen, indem anstatt Gl. 5.20 die entsprechende Umkehrfunktion verwendet wird:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \left(1 + \frac{\Delta Z}{Z_c}\right) \mathbf{R}_{2D}(-\Delta r_z) \left[\begin{pmatrix} u' \\ v' \end{pmatrix} - \begin{pmatrix} c_x \\ c_y \end{pmatrix} - \frac{f}{Z_c} \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \right] \quad (5.21)$$

Diese Formel ermöglicht es, für jeden Zielbildpunkt (u', v') die Position (u, v) des korrespondierenden Ursprungspixels zu bestimmen und diesen Grauwert – gegebenenfalls durch Interpolation der Intensitäten direkt benachbarter Bildpunkte – anschließend zu übernehmen.

5.4.2.5 Approximationsfehler-Analyse

Die vorgestellten planaren Approximationsterme führen jeweils zu mehr oder weniger starken Unterschieden zwischen den approximierten Röntgenbildern im Vergleich zu den korrekten Ansichten, die sich für die betrachteten Lageparameter ergeben würden. Die Ursache hierfür ist insbesondere in der den einzelnen Näherungen zugrundeliegenden Reduktion der Tiefe des Objektes auf die Z-Koordinate Z_c des Schwerpunkts zu suchen. Es ergeben sich die nachfolgend aufgeführten Approximationsfehler⁸. Diese werden dabei jeweils in Form des euklidischen Abstands $\| \cdot \|_2$ zwischen korrektem Projektionspunkt ${}^I p_{kor}$ und approximiertem Projektionspunkt ${}^I p_{app}$ angegeben. Um die Aussagefähigkeit der resultierenden Abschätzungen nicht zu gefährden, erfolgt die Fehlerangabe analog zur Herleitung der Approximationsterme getrennt nach einer Verschiebung parallel zur Bildebene, einer Verschiebung senkrecht zur Bildebene sowie einer Rotation um die optische Achse. Z_c kennzeichnet die z-Koordinate des Objektschwerpunktes, ΔZ_{obj} die aktuelle maximale Ausdehnung des Objektes in z-Richtung:

1. Translation parallel zur Bildebene ($\Delta X, \Delta Y$):

$$\| {}^I p_{kor} - {}^I p_{app} \|_2 \leq \frac{f \sqrt{\Delta X^2 + \Delta Y^2}}{2 \frac{Z_c^2}{\Delta Z_{obj}} - Z_c} \quad (5.22)$$

Der durch die parallele Verschiebung bewirkte planare Approximationsfehler ist in erster Linie direkt proportional zu der Translationsstrecke (Term $\sqrt{\Delta X^2 + \Delta Y^2}$). Die restlichen Einflußgrößen wirken sich folgendermaßen aus:

- Besondere Beachtung verdient der Nenner des Terms; er wird dann gleich oder kleiner 0, falls Objektpunkte auf Höhe des Kamerazentrums ($Z = 0$) bzw. dahinter zu liegen kommen. Dieser Fall kann in der Praxis jedoch ausgeschlossen werden, da er physikalisch unmöglich ist – die Röntgenquelle befände sich dann im Patienteninneren. Vielmehr gilt i.a.: $Z_c \gg \Delta Z_{obj}$.

⁸Die ausführliche Herleitung dieser Abschätzungen findet sich im Anhang C.

- f : Der Fehlerwert ist direkt proportional zur Brennweite.
- Z_c : Je größer der Abstand des Objektes zur Röntgenquelle, desto geringer fällt der Fehler aus.
- ΔZ_{obj} : Eine geringere aktuelle Objektiefe (in z-Richtung) verkleinert den Fehler ebenfalls.

2. Translation in Richtung der optischen Achse (ΔZ):

$$\|I p_{kor} - I p_{app}\|_2 \leq \frac{f \Delta Z_{obj} \Delta Z \sqrt{X_p^2 + Y_p^2}}{2 Z_{min} (Z_{min} + \Delta Z) (Z_c + \Delta Z)} \quad (5.23)$$

Neben der Feststellung, daß sich eine größere Translation ΔZ in einem größeren Fehler niederschlägt, wirkt sich eine Veränderung der restlichen Einflußgrößen bei voneinander unabhängiger Betrachtung wie folgt aus:

- $\sqrt{X_p^2 + Y_p^2}$: Der Approximationsfehler für einen Objektpunkt wird umso größer, je weiter der betreffende Punkt von der optischen Achse entfernt ist. Punkte direkt auf der Achse werden fehlerfrei berechnet.
- f : Eine kleinere Brennweite hat eine direkt proportionale Verringerung des Fehlers zur Folge.
- Z_c, Z_{min} : Je größer der Abstand zwischen Objektzentrum und Kamerazentrum ausfällt, desto kleiner wird der Fehler. Dasselbe gilt für den der Kamera am nächsten gelegenen Punkt Z_{min} .
- ΔZ_{obj} : Der Zusammenhang zwischen aktueller Objektiefe und Approximationsfehler ist direkt proportional.

3. Rotation um die optische Achse (Δr_z):

$$\|I p_{kor} - I p_{app}\|_2 = 0 \quad (5.24)$$

D.h. die räumliche Rotation um die optische Achse kann fehlerfrei durch eine planare Rotation um das optische Zentrum approximiert werden. Aufgrund der Pixel-Diskretisierung des Digitalbildes kann es jedoch vereinzelt auch hier zu Abweichungen in der Größenordnung eines Pixels kommen.

Die Auswirkungen dieser Fehlerquellen können durch geschickte Wahl der realen Aufnahmerichtungen und -entfernungen eingeschränkt werden. Hierzu ist der C-Bogen jeweils derart zu plazieren, daß das Objekt im Zentrum des Bildes zu liegen kommt. Außerdem sollte der Detektor des Fluoroskops möglichst nahe an den Patienten gebracht werden, um dadurch den Abstand des Knochens zur Röntgenquelle zu maximieren.

Die in der Realität quantitativ tatsächlich zu erwartenden Fehler beim Einsatz der planaren Approximation sollen abschließend anhand eines konkreten Zahlenbeispiels verdeutlicht werden. Die folgenden Annahmen repräsentieren eine reale Aufnahmesituation für einen Brustwirbel (T6):

- Brennweite $f = 80.0$ cm.
- Z-Koordinate des Wirbel-Schwerpunktes $Z_c = 50.0$ cm.
- Maximale Ausdehnung des Wirbels in Z-Richtung $\Delta Z_{obj} = 6.0$ cm.

- Minimale z -Koordinate des Wirbels $Z_{min} = 46.0$ cm.
- Größter Abstand eines Punktes von der optischen Achse $\sqrt{X_p^2 + Y_p^2} = 4.0$ cm.

Unter diesen Annahmen ergeben sich für ein approximiertes Bild, dem eine Objektbewegung (relativ zur Objektlage des verwendeten Ansichtsmodells) von z.B. ($\Delta X = 2.0$ cm, $\Delta Y = 2.0$ cm, $\Delta Z = 5.0$ cm) zugrunde liegt, die folgenden Fehlerabschätzungen für eine Translation parallel (vgl. Gl. 5.22):

$$\|I_{p_{kor}} - I_{p_{app}}\|_2 \leq 0.29 \text{ cm}$$

bzw. senkrecht (vgl. Gl. 5.23) zur Bildebene:

$$\|I_{p_{kor}} - I_{p_{app}}\|_2 \leq 0.036 \text{ cm}$$

Bei einer üblichen Pixelauflösung von z.B. 0.4 mm pro Pixel ergeben sich damit im ersten Fall Maximalfehler von ca. sieben Pixeln, im zweiten Fall lediglich Fehler von ca. einem Pixel. Diese Werte sind für die vorliegende Anwendung sehr wohl akzeptabel, zumal es sich um Abschätzungen des maximalen Fehlers handelt. Der durchschnittliche Fehler fällt i.a. deutlich geringer aus, da zum einen die meisten Punkte eines üblichen Objektes in der Nähe des Schwerpunkts liegen und zum anderen nur wenige Punkte einen großen Abstand zur optischen Achse besitzen. Abb. 5.19 veranschaulicht den geringfügigen Unterschied zwischen einer korrekt simulierten und einer planar approximierten Aufnahme.



Abbildung 5.19: Vergleich zwischen einer durch Ray-Casting simulierten Röntgenaufnahme (links) und einer unter Verwendung eines Ansichtsmodells rekonstruierten (15 Eigenvektoren) und planar approximierten Aufnahme (Mitte). Die Aufnahmeparameter entsprechen dem im Text angegebenen Zahlenbeispiel für eine reale Aufnahmesituation (Aufnahmerichtung ca. 5° Grad von den nächstgelegenen Trainingsbildansichten entfernt). Das rechte Bild zeigt das invertierte Differenzbild aus den beiden berechneten Aufnahmen.

5.4.3 Optimierung

Die Verwendung von Ansichtsmodellen im Rahmen der Registrierung ist völlig transparent möglich. D.h. für die in Abschnitt 5.3.2 vorgestellten Algorithmen ist es letztlich unerheblich, ob der aktuelle Wert der Fehlerfunktion durch zeitaufwendiges Ray-Casting oder durch die schnelle Auswertung von einem oder mehreren Ansichtsmodellen erfolgt. Voraussetzung hierfür ist allerdings, daß sich die im letzten Abschnitt abgeschätzten Approximationsfehler im konkreten Fall nur geringfügig auswirken und so das

Konvergenzverhalten des Verfahrens nicht signifikant verschlechtern. Garantieren läßt sich diese Voraussetzung, indem man die Entscheidung darüber, ob ein vorhandenes Ansichtsmodell zur Bildberechnung herangezogen wird, jedesmal zusätzlich an eine Abschätzung des voraussichtlichen Approximationsfehlers koppelt.

Wird während der Optimierung die Röntgenbild-Simulation für eine Transformation \mathbf{p}_{cam} bzgl. einer der gegebenen Kameradefinitionen nötig, so erfolgt zunächst die Überprüfung der folgenden beiden Kriterien:

1. Existiert ein Ansichtsmodell, dessen beschränkter Ansichtsbereich die aktuelle Kamera-Blickrichtung beinhaltet?
2. Falls ja, unterschreiten die Abschätzungen aus Gl. 5.22 und Gl. 5.23 in Abhängigkeit von den aktuellen Werten für ΔT_x , ΔT_y und ΔT_z einen vorgegebenen Grenzwert für den maximal zulässigen Approximationsfehler?

Sind beide Kriterien erfüllt, so wird das angeforderte Röntgenbild durch Auswertung des Ansichtsmodells einschließlich planarer Approximation bestimmt. Ansonsten erfolgt die direkte Berechnung eines simulierten Bildes durch Ray-Casting. Abgesehen von der Erweiterung durch Integration von Ansichtsmodellen in den Ablauf der Optimierung verläuft diese analog zu einer der bereits in Abschnitt 5.3.2 vorgestellten Algorithmen-Varianten.

Aufgrund der durch die planaren Approximationen induzierten Fehler ist die resultierende Transformation nach Konvergenz des iterativen Verfahrens i.a. jedoch nicht optimal. Eine abschließende iterative Minimierung, bei der vollständig auf evtl. vorhandene Ansichtsmodelle verzichtet wird, beseitigt dieses Problem. Da dabei zum einen die Parameter-Schrittweiten der Minimierung sehr klein gewählt werden können und zum anderen keine adaptive Anpassung der Bildgröße – verbunden mit einer Vielzahl an Iterationen – mehr zu erfolgen braucht, konvergiert diese abschließende *Feinoptimierung* meist sehr schnell.

5.5 Ergebnisse

Um die Leistungsfähigkeit des im zurückliegenden Teil des Kapitels vorgestellten Registrierungsverfahrens isoliert, d.h. unabhängig von der Bildqualität des verwendeten Fluoroskops sowie der Präzision der Kalibrierung, beurteilen zu können, werden zunächst einige aussagekräftige Simulationsergebnisse vorgestellt. Es folgen mehrere Untersuchungen mit realen Fluoroskopieaufnahmen von Femur-Knochen und Wirbeln. Abschließend wird der Rechenaufwand des Verfahrens – insbesondere der vorgestellten Beschleunigung durch Ansichtsmodelle – analysiert.

5.5.1 Simulationsergebnisse

Um einen ersten qualitativen Eindruck über Beschaffenheit und Verlauf der sechsdimensionalen Fehlerfunktion 5.5 zu gewinnen, wurden ausgehend von exakt bekannten Lageparametern $(r_x, r_y, r_z, t_x, t_y, t_z)$ für einen Lendenwirbel-Datensatz je zwei Parameter ausgewählt und systematisch innerhalb vorgegebener Grenzen (Rotation: $\pm 60.0^\circ$, Translation: $\pm 20.0\text{mm}$) in beide Richtungen verändert. Der für jede sich ergebende Kombination resultierende Fehlerfunktionswert wurde schließlich gegen die beiden betrachteten Parameter in einem 3D-Diagramm angetragen. Abb. 5.20 faßt die ermittelten Diagramme für

alle 15 möglichen Parameterkombinationen zusammen. Dabei wurde im Sinne einer besseren Darstellung das Fehlermaß durch Negation als Ähnlichkeitsmaß angetragen. Der gewählte Datensatz beinhaltete insgesamt drei simulierte Aufnahmen eines isolierten L3-Lendenwirbels, wobei die Aufnahmerichtungen untereinander einen Zwischenwinkel von mindestens 45° und maximal 80° aufwiesen. In jedem der Diagramme hebt sich die optimale Parameterkombination deutlich von ihrer Umgebung ab. Desweiteren treten in der jeweiligen Umgebung der globalen Extrema überraschenderweise keine lokalen Extrema auf. Im Gegenteil, die Funktionsverläufe steigen mehr oder weniger steil in Richtung Optimum an. Wenngleich diese Erkenntnis zunächst noch wenig Aussagekraft bzgl. realer Fluoroskopieaufnahmen besitzt, so zeigt sich doch bereits deutlich die prinzipielle Eignung des gewählten pixelbasierten Fehlermaßes.

Für die Beurteilung des Konvergenzverhaltens der Optimierung in Abhängigkeit von der betrachteten Knochenstruktur wurden drei Experimente mit verschiedenen Phantomknochen (L3-Lendenwirbel, distaler und proximaler Femur) durchgeführt. Grundlage der Experimente war jeweils eine Szene, bestehend aus drei simulierten Röntgenaufnahmen, wobei die Zwischenwinkel der Aufnahmerichtungen wiederum im Bereich $45^\circ \dots 80^\circ$ lagen. Für jede dieser Szenen wurden insgesamt 100 Optimierungsläufe durchgeführt, wobei die initialen Lageparameter mit unterschiedlichen rotatorischen und translatorischen Abweichungen bezüglich der – durch die Simulation bekannten – korrekten Lageparameter besetzt waren. Die initiale translatorische Abweichung betrug maximal 15mm , die maximale initiale rotatorische Abweichung 60° . Die Beschränkung auf einen einzelnen Winkel zur Beschreibung der drei rotatorischen Freiheitsgrade deutet bereits auf die Verwendung der in Abschnitt 4.1 erwähnten Orientierungsrepräsentation (Einzelrotation um eine beliebig orientierte Achse) hin. Die Orientierung der Rotationsachse wurde jeweils zufällig ausgewählt. Entsprechendes gilt in diesem Abschnitt auch für die Beschreibung translatorischer Abweichungen und Fehler. Von Interesse war lediglich der Betrag der Verschiebung, jedoch nicht deren Richtung. Für die Initialisierung erfolgte die Auswahl ebenfalls zufällig. Abb. 5.21 präsentiert die Ergebnisse der Optimierungsläufe, d.h. die resultierenden rotatorischen bzw. translatorischen Restfehler in Abhängigkeit von initialen Abweichungen. Dabei fassen je zwei nebeneinander angeordnete Diagramme die Ergebnisse für eines der drei untersuchten Knochenphantome zusammen.

Für den L3-Lendenwirbel besitzt der Optimierungsalgorithmus ein ausgezeichnetes Konvergenzverhalten. Erst ab einer initialen rotatorischen Abweichung von ca. 50° beginnt das Verfahren schlagartig zu divergieren, was sich durch Restfehler von $2 \dots 10\text{mm}$ bzw. $10^\circ \dots 60^\circ$ bemerkbar macht. Bei kleineren rotatorischen Abweichungen ergeben sich ausnahmslos vernachlässigbare Restfehler von unter 0.2mm bzw. 0.5° . Das Konvergenzverhalten ist hier insbesondere unabhängig von der initialen translatorischen Abweichung (bis 15mm).

Für ein proximales Femurende ergibt sich ein etwas anderes Bild. Hier treten bereits bei etwa 30° initialer rotatorischer und ab etwa 5mm translatorischer Abweichung vereinzelt Divergenzfälle auf (Restfehler 7mm bzw. 3°). Ab etwa 50° ist der Optimierungsalgorithmus nicht mehr in der Lage, die gesuchten Parameter zu bestimmen, wobei sich dies interessanterweise stärker in translatorischen Restfehlern von $5 \dots 30\text{mm}$ niederschlägt. Orientierungsfehler von über 5° ergeben sich jedoch erst ab einer initialen translatorischen Abweichung von 13mm .

Fielen die bisherigen Ergebnisse insgesamt sehr ermutigend aus, so ändert sich dies bei Betrachtung der Restfehler-Diagramme für ein distales Femurende. Hier gelingt es der Optimierung lediglich bis zu einer initialen rotatorischen Abweichung von ca. 5 bis 7° , die korrekten Lageparameter präzise zu bestimmen. Fällt die Abweichung nur wenige Grad höher aus, so zieht dies unmittelbar translatorische Restfehler von $2 \dots 6\text{mm}$ und bis zu 20° nach sich. Dieses schlechte Abschneiden rührt vermutlich

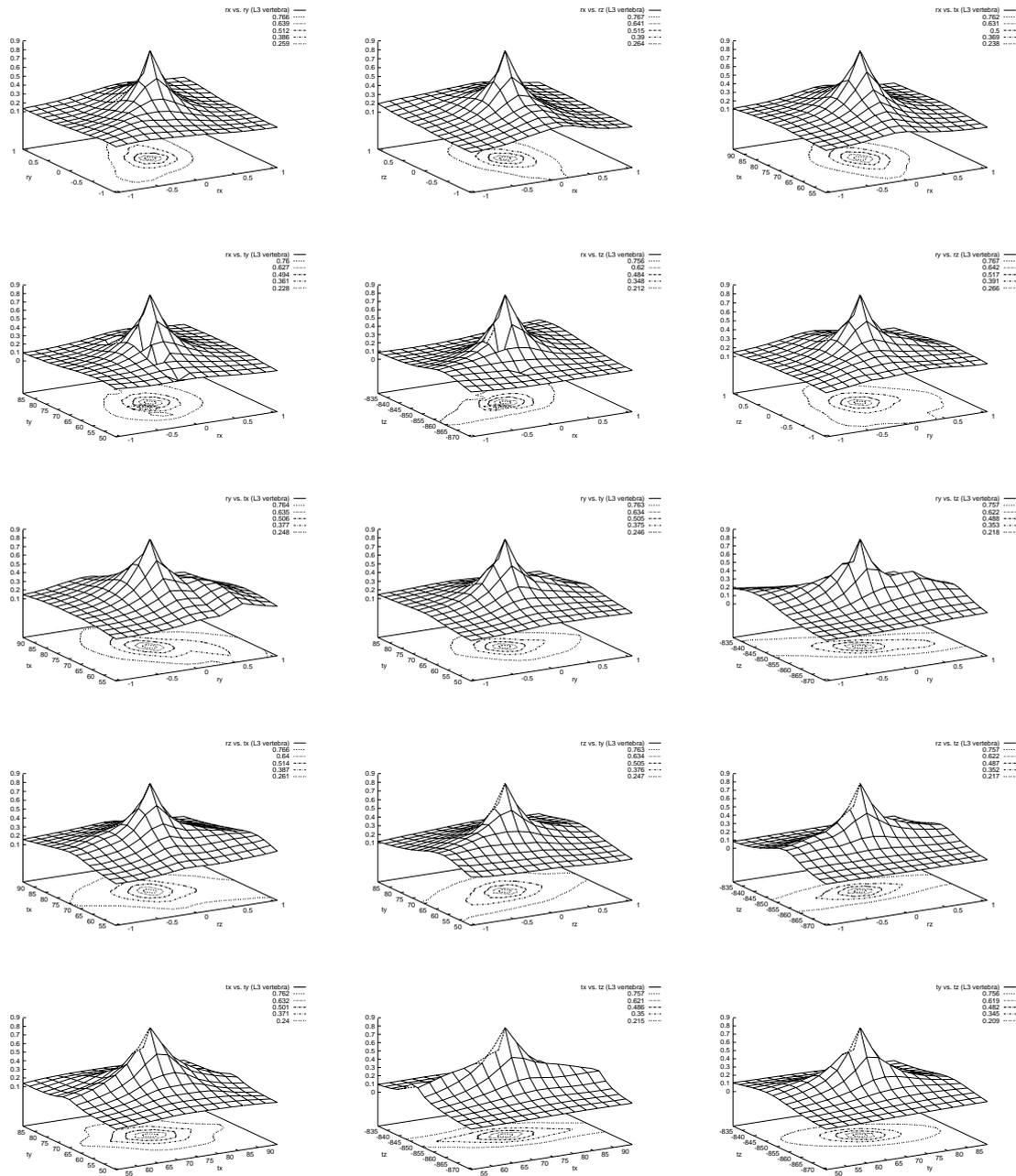


Abbildung 5.20: Grafische Veranschaulichung des lokalen *Fehlergebirges* für einen simulierten Lendenwirbel-Datensatz (L3), bestehend aus drei Aufnahmen mit einem Winkelabstand der Blickrichtungen zwischen 45° und 80° . Der besseren Darstellung wegen wurde das Fehlermaß aus Gl. 5.5 negiert und dadurch als Ähnlichkeitsmaß verwendet. Jeweils zwei der sechs Transformationsparameter (r_x, r_y, r_z) (Orientierung), (t_x, t_y, t_z) (Translation) sind gegeneinander angetragen.

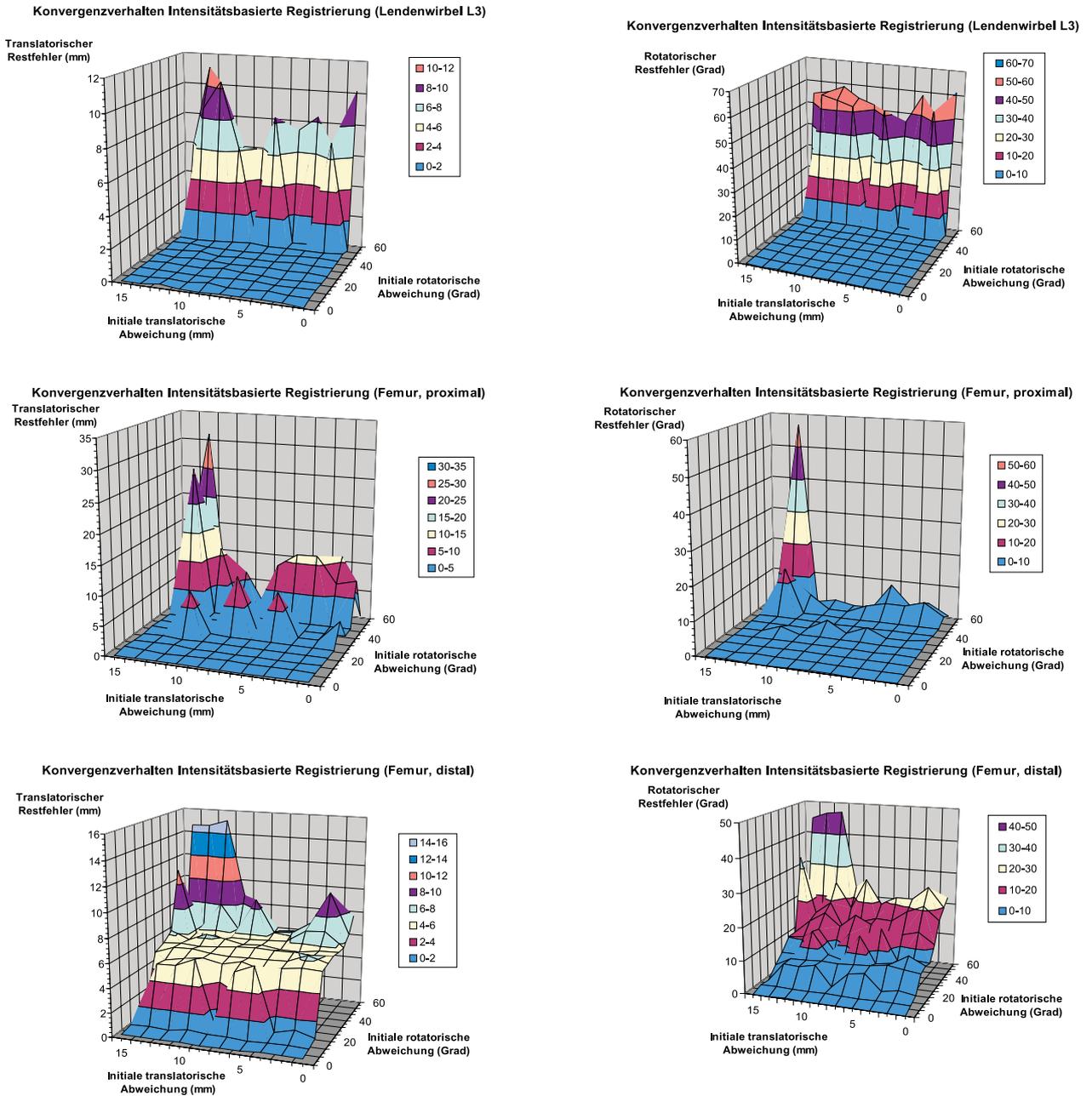


Abbildung 5.21: Konvergenzverhalten für simulierte Testdaten (jeweils drei Aufnahmen, Winkelabstand der Blickrichtungen zwischen 45° und 80°) für einen Lendenwirbel (L3) sowie einen distalen und einen proximalen Femur. Angetragen wurde jeweils der translatorische bzw. rotatorische Restfehler nach erfolgter Konvergenz der Optimierung, ausgehend von unterschiedlichen initialen translatorischen und rotatorischen Abweichungen.

daher, daß ein distales Femurende im Röntgenbild nur wenige markante innere Strukturen offenbart. Wie zu Beginn des Kapitels bereits angesprochen, sind es aber gerade diese inneren Strukturen (z.B. der Dornfortsatz eines Wirbels), die die intensitätsbasierte Optimierung präzise konvergieren lassen.

Reale Fluoroskopieaufnahmen sind i.a. stark verrauscht. Wenngleich sich starkes Rauschen durch Mittelung mehrerer unmittelbar hintereinander aufgenommener Bilder unterdrücken läßt, soll im folgenden die Rauschempfindlichkeit der Optimierung anhand des bereits beschriebenen L3-Datensatzes untersucht werden. Zunächst wurde für diesen Datensatz derselbe Konvergenzversuch wie zuvor dreimal wiederholt, wobei die Eingabebilder sukzessive stärker synthetisch verrauscht wurden (Gauß-verteilt Rauschen, Standardabweichung = 20,30,40). Abb. 5.22 (unten) veranschaulicht die Wirkung von Rauschen auf die Bildqualität. Abb. 5.23 beinhaltet die resultierenden Restfehlerdiagramme dieses Experiments. Es zeigt sich, daß das Rauschen den Konvergenzradius der Optimierung für einen Lendenwirbel auf etwa 20° Grad reduziert, allerdings nahezu unabhängig von der Stärke des Rauschens.

Wie der Vergleich der Restfehler-Streuung in Abb. 5.24 aufzeigt, führt Rauschen zum vermehrten Auftreten von Ausreißern, d.h. wenigen einzelnen isolierten Divergenzfällen. Der Mehrzahl aller Optimierungsläufe gelingt es jedoch, die korrekten Lageparameter in ausreichender Präzision zu bestimmen. Die initialen Abweichungen blieben bei diesem Versuch auf den eingangs ermittelten Konvergenzbereich für unverrauschte Bilder eines L3-Lendenwirbels beschränkt ($15\text{mm}/45^\circ$). Abb. 5.25 schließlich zeigt, wie sich das Rauschen auf die resultierenden Restfehler der Optimierung, ausgehend von einer einzelnen konkreten Initialisierung (Abweichung $2\text{mm}/10^\circ$), auswirkt. Ab einer Standardabweichung von 40 (50) steigt der rotatorische (translatorische) Restfehler stark an, wobei es der Optimierung immerhin gelingt, die initiale rotatorische Abweichung von 10° auf 2° zu reduzieren. Da die translatorische Abweichung im vorliegenden Beispiel mit 2mm bereits initial sehr gering ausfällt, verbessert die Optimierung bei stark verrauschtem Bildmaterial daran nichts mehr.

Aufgrund der Tatsache, daß die intraoperativen Aufnahmeorientierungen von vornherein nicht festgelegt sind, sondern innerhalb eines großen Bereichs durch entsprechende Manipulation der C-Bogen-Orientierung nahezu beliebig gewählt werden können, stellt sich folgende Frage: Wie stark wirken sich unterschiedliche Aufnahmeorientierungen auf die berechneten Lageparameter aus? Folgender Versuch soll hierauf eine Antwort liefern. Ausgehend von 15 unterschiedlichen Aufnahmen des L3-Wirbelphantoms wurden insgesamt fünfmal zufällig drei Aufnahmen ausgewählt, deren Aufnahmeorientierungen jedoch einen Zwischenwinkel von mindestens 45° aufweisen mußten. Für jede dieser Szenen wurden anschließend neun Optimierungsläufe mit unterschiedlichen initialen Abweichungen bis maximal $2.5\text{mm}/10^\circ$ Grad durchgeführt. Abb. 5.26 enthält die Restfehler-Streuung dieses Versuchs. Deutlich zu erkennen ist, daß die gemessenen translatorischen Restfehler gleichmäßig und in engen Grenzen von $\pm 0.1\text{mm}$ um 0.4mm streuen. Auch die rotatorischen Restfehler der unterschiedlichen Aufnahmeorientierungen unterscheiden sich nicht signifikant. Von einzelnen Ausreißern abgesehen, verteilen sie sich gleichmäßig von 0.0° bis 0.8° , wobei die Variation weniger von der gewählten Gruppierung als vielmehr von den im Einzelfall gewählten initialen Lageparametern abhängt.

5.5.2 Anwendung auf reale Fluoroskopieaufnahmen

Das im letzten Abschnitt untersuchte Konvergenzverhalten der vorgestellten Lageschätzung liefert bereits wichtige Anhaltspunkte für die Beurteilung der Leistungsfähigkeit und Grenzen des Ansatzes. Entscheidend im Hinblick auf einen praktischen, intraoperativen Einsatz der entwickelten Methoden ist letztlich jedoch deren Verhalten unter realen OP-Bedingungen. Dies gilt im vorliegenden Fall umso mehr, da

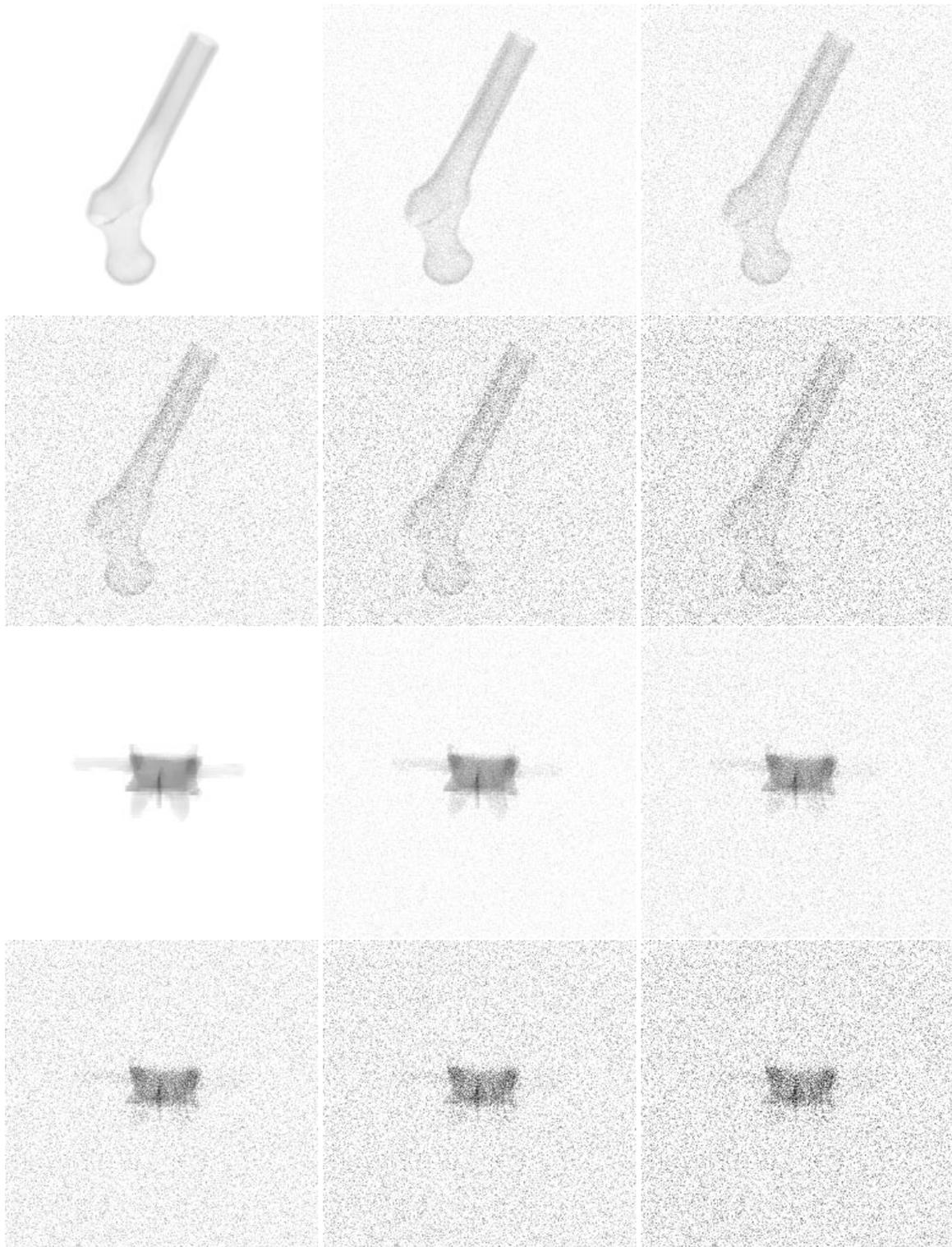


Abbildung 5.22: Synthetisch verrauschte Testbilder eines L3-Wirbels sowie eines proximalen Femurs (Gauß-Verteilung), Standardabweichungen (jeweils von links oben nach rechts unten): $\sigma = 0, 10, 20, 30, 40, 50$

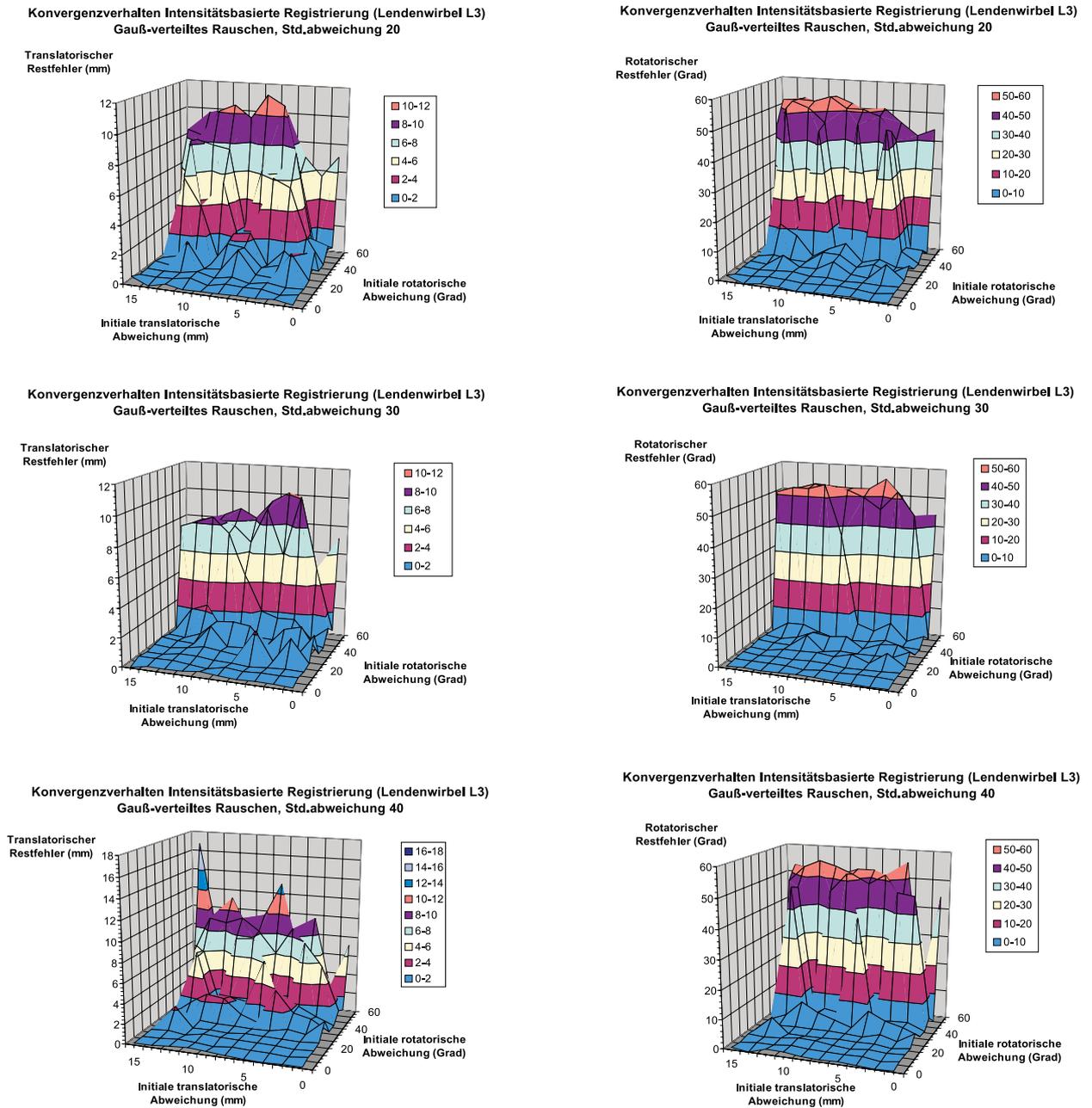


Abbildung 5.23: Konvergenzverhalten bei künstlich verrauschten Testdaten für einen Lendenwirbel (L3). Die ursprünglichen unverrauschten L3-Aufnahmen (vgl. Abb. 5.21) wurden mit einem Gauß-verteilt Rauschen überlagert, wobei die Standardabweichung der Verteilung sukzessive erhöht wurde. Die Darstellung der Ergebnisse entspricht Abb. 5.21.

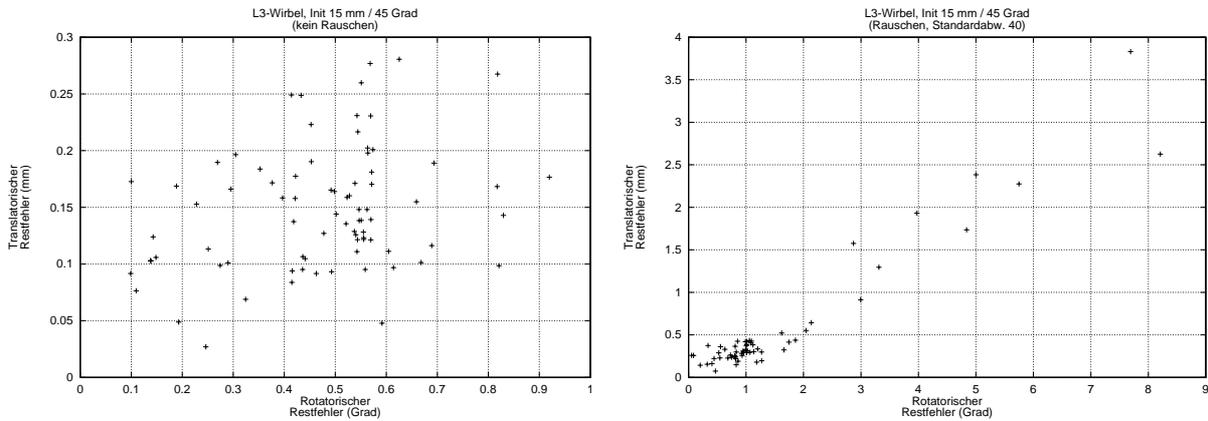


Abbildung 5.24: Restfehler-Streuung bei simulierten Testdaten eines L3-Wirbels. Initialisierungen erfolgten mit maximal 15mm translatorischer und maximal 45° Grad rotatorischer Abweichung. Links: unverrauschte Bilddaten, rechts: verrauschte Bilddaten (Standardabweichung der Gauß-Verteilung 40).

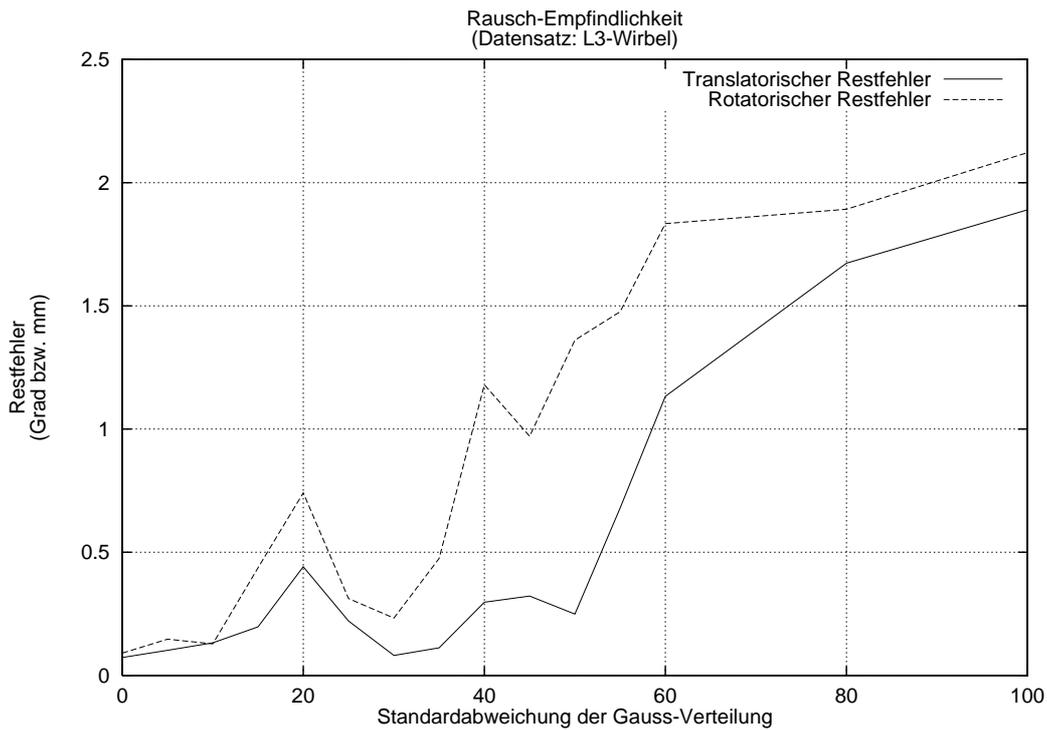


Abbildung 5.25: Sensitivität der Optimierung gegenüber Rauschen. Die der Abb. 5.23 zugrunde liegenden Datensätze (Lendenwirbel L3, sukzessive verrauscht) wurden nacheinander für die Optimierung herangezogen. Die initiale Abweichung betrug jeweils konstant 2mm und 10°.

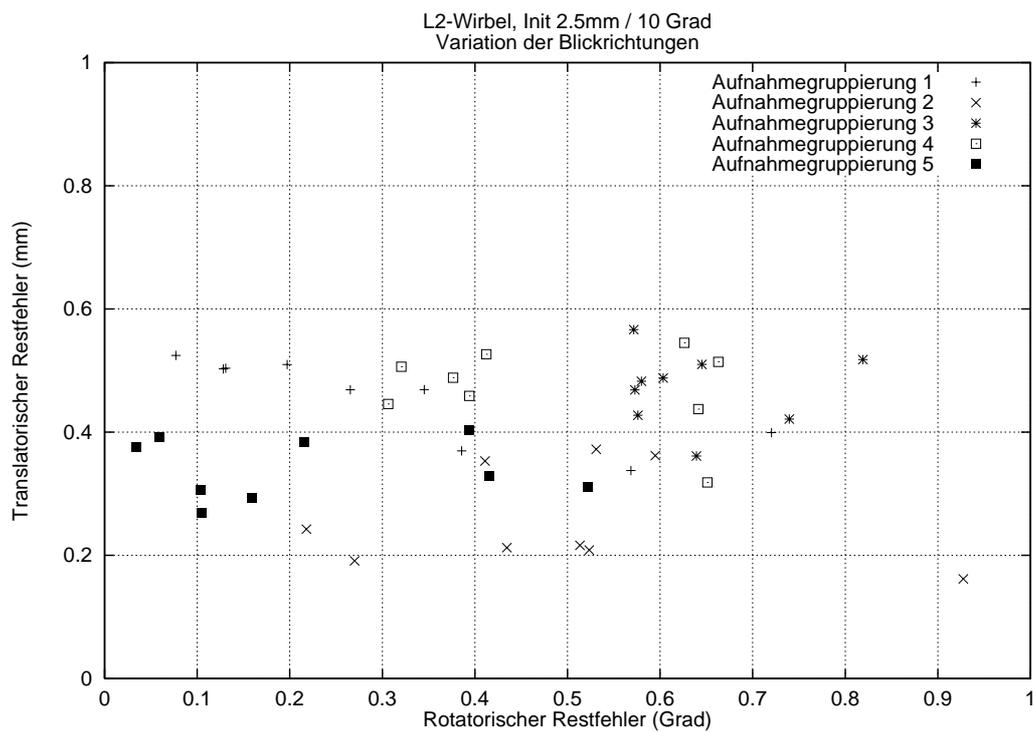


Abbildung 5.26: Restfehler-Streuung bei Verwendung unterschiedlicher Aufnahmerichtungen und mehrerer Initialisierungen bis maximal 2.5mm bzw. 10°

für die intraoperativ erstellten Fluoroskopieaufnahmen bewußt keine explizite Merkmalsextraktion (z.B. Konturen, Punkte etc.) erfolgt, sondern vielmehr die Pixeldaten selbst in die Optimierung einfließen. Daher kommt dem zu erwartenden signifikanten Qualitätsunterschied zwischen synthetischen, simulierten Röntgenaufnahmen und realen Fluoroskopieaufnahmen in Abhängigkeit des verwendeten C-Bogens eine besondere Bedeutung zu.

Um die Störeinflüsse einer realen Umgebung – wesentlich bestimmt durch die Bildqualität sowie die Kalibrierung des Fluoroscops – auf das Konvergenzverhalten der Lageparameter-Optimierung zu analysieren, wurden die nachfolgend dokumentierten Experimente durchgeführt. Grundlage der Versuche war erstens ein Brustwirbelphantom (T10-T12), wobei für die Registrierung der mittlere Wirbel (T11) ausgewählt wurde, zweitens ein Femurphantom, dessen distales Ende lokalisiert werden sollte, sowie drittens das proximale Ende eines realen Femur-Knochens. Als Aufnahmegerät dienten zwei unterschiedliche Fluoroskopie-Geräte: ein SIEMENS Siremobil 2 (T11-Brustwirbel, proximaler Femur) sowie ein ZIEHM Exposcop (distaler Femur). Beide Sensorsysteme wurden durch das in Anhang B beschriebene Verfahren präoperativ kalibriert [Bra00a].

Ein inhärentes Problem bei Experimenten mit realen Datensätzen stellt die Nicht-Kenntnis der optimalen Lageparameter dar, was die Definition von Restfehlern erschwert. Im Falle der simulierten Eingabedaten waren die Parameter hingegen exakt bekannt. Die drei realen Datensätze wurden aus diesem Grund zunächst unter Verwendung eines alternativen Registrierungsverfahrens lokalisiert. Hierzu dienten mehrere Metall- bzw. Teflonkugeln, die bereits vor Erstellung der Computer-Tomographie-Datensätze an den drei Phantomen angebracht wurden. Die 3D-Positionen der Kugeln in den CT-Daten wurden entsprechend dem Verfahren von [El196] ermittelt. Intraoperativ erfolgte eine manuelle Abtastung derselben Kugeln mit Hilfe eines Zeigeinstrumentes, dessen 3D-Position das Infrarot-basierte Trackingsystem verfolgte. Aus den beiden 3D-Punktmenge läßt sich schließlich die jeweils gesuchte rigide Transformation abschätzen [HS93a, BM92]. Diesen Lageparametern wurde unterstellt, daß sie optimal sind. Sie dienen anschließend als Zielparameter für die einzelnen Konvergenztests, deren experimentelle Durchführung ansonsten auf die bereits bei simulierten Datensätzen beschriebene Art und Weise erfolgte. Lediglich die maximale initiale rotatorische Abweichung wurde auf 20° verringert, da bei realen Daten eine Verbesserung der Konvergenz im Vergleich zu simulierten Daten nicht zu erwarten ist. Außerdem beträgt der maximale Zwischenwinkel zwischen zwei Aufnahmerichtungen eines Datensatzes nur noch 70° aufgrund einer zusätzlich eingeschränkten Bewegungsfreiheit der beiden verwendeten Röntgenbildverstärker-Systeme. Abb. 5.28 dokumentiert die erzielten Ergebnisse.

Für den T11-Brustwirbel ergeben sich translatorische Restfehler von unter 1mm , sofern die initiale rotatorische Abweichung nicht größer als 10° ausfällt. Der rotatorische Restfehler liegt in diesem Fall aber bereits bei etwa 3° . Bei einer maximalen rotatorischen Abweichung von 5° reduziert er sich auf unter 1° .

Die translatorischen Lageparameter des proximalen Femur-Datensatzes können durch die Optimierung robust bestimmt werden. Bei initialen Abweichungen von über 10mm und etwa 20° ergeben sich Restfehler von mehr als einem Millimeter. Jedoch auch bei diesem Datensatz wird der Konvergenzbereich wie bereits zuvor durch rotatorische Restfehler weiter eingeschränkt. Bei maximal $7.5\text{mm}/10^\circ$ initialer Abweichung liegen die resultierenden Orientierungsfehler unter einem Grad.

Die Ergebnisse, den realen distalen Femur-Datensatz betreffend, bestätigen das bereits in der Simulation festgestellte schlechtere Restfehlerverhalten. Insbesondere der translatorische Fehler steigt ab einer initialen rotatorischen Abweichung von etwa 10° deutlich stärker an als im Fall des T11-Wirbels bzw. des proximalen Femurendes. Ähnliches gilt für die rotatorischen Restfehler, die für über 10° initialer Ab-

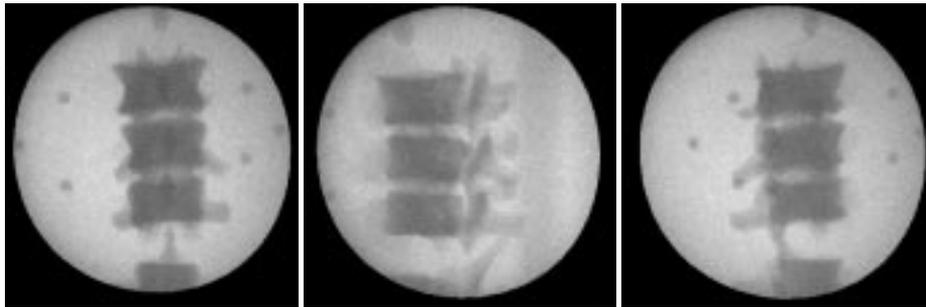


Abbildung 5.27: Drei reale Fluoroskopieaufnahmen (SIEMENS Siremobil 2) eines T11-Brustwirbelphantoms

weichung bei etwa 4° liegen. Bleiben die initialen rotatorischen Abweichungen jedoch auf etwa $5 - 10^\circ$ beschränkt, so lassen sich die Restfehler in Abhängigkeit von der initialen translatorischen Abweichung durchaus auf $0.5 - 1.0 \text{ mm}$ und $0.5 - 1.5^\circ$ reduzieren. Dieses Ergebnis ist überraschenderweise sogar besser als in der Simulation. Ein möglicher Grund hierfür mag ein zufällig *günstiger* gewähltes Arrangement der Aufnahmerichtungen sein.

Zusammenfassend läßt sich feststellen, daß der Konvergenzradius der intensitätsbasierten Lokalisation bei realen Fluoroskopieaufnahmen deutlich geringer ausfällt als in der Simulation. Angesichts der sehr schlechten Bildqualität des für die ersten beiden Versuche zur Verfügung stehenden *Siremobil 2*-Fluoroscops erscheint das eingeschränkte Konvergenzverhalten allerdings sehr wohl akzeptabel. Abb. 5.27 zeigt beispielhaft den Bilddatensatz für die Lokalisation des T11-Brustwirbels. Die präsentierten experimentellen Resultate sollten daher einerseits als *best-case*-Szenario (im Simulationsfall) bzw. *worst-case*-Szenario (Tests mit realen Aufnahmen) aufgefaßt werden. Insbesondere der Einsatz hochwertiger moderner C-Bogen-Systeme läßt eine deutliche Verbesserung des Konvergenzverhaltens erwarten.

5.5.3 Zeitaufwand der Optimierung

Bei den bisherigen Untersuchungen blieb die Verwendung von Ansichtsmodellen zur beschleunigten Berechnung synthetischer Röntgenbilder (Abschnitt 5.4.2) unberücksichtigt. Ein abschließender Versuch soll daher nun das Zeitverhalten der Optimierung bei Verwendung von Ray-Casting im Vergleich zum Einsatz der Ansichtsmodellierung untersuchen. Hierzu wurden zwei der simulierten Testdatensätze (Lendenwirbel L3, proximaler Femur, jeweils drei Aufnahmen) aus Abschnitt 5.5.1 herangezogen. Ausgehend von initialen Lageparametern innerhalb des jeweiligen Konvergenzradius⁹ erfolgte für jeden Datensatz je eine Optimierung ohne sowie eine mit Ansichtsmodellen. Die verwendeten Ansichtsmodelle wurden zuvor derart trainiert, daß sie den bekannten Ansichtsbereich der Testdaten vollständig abdecken. Auf diese Weise wurde verhindert, daß während der Optimierung vereinzelt Ray-Casting-Berechnungen nötig werden.

Abb. 5.29 faßt die Ergebnisse dieses Vergleichs zusammen⁹. Die Sprungstellen der einzelnen Kurven deuten auf die adaptive Anpassung der Bildgröße hin. Diese beträgt zu Beginn der Optimierungen jeweils 32×32 Pixel und steigert sich dann in drei Schritten bis auf 256×256 . Es zeigt sich deutlich, daß die Ansichtsmodelle bei den kleinen Bildauflösungen noch keinen Vorteil bringen. Hier hält sich der

⁹Die Zeitangaben beziehen sich auf CPU-Sekunden eines Intel Pentium III, 450 Mhz.

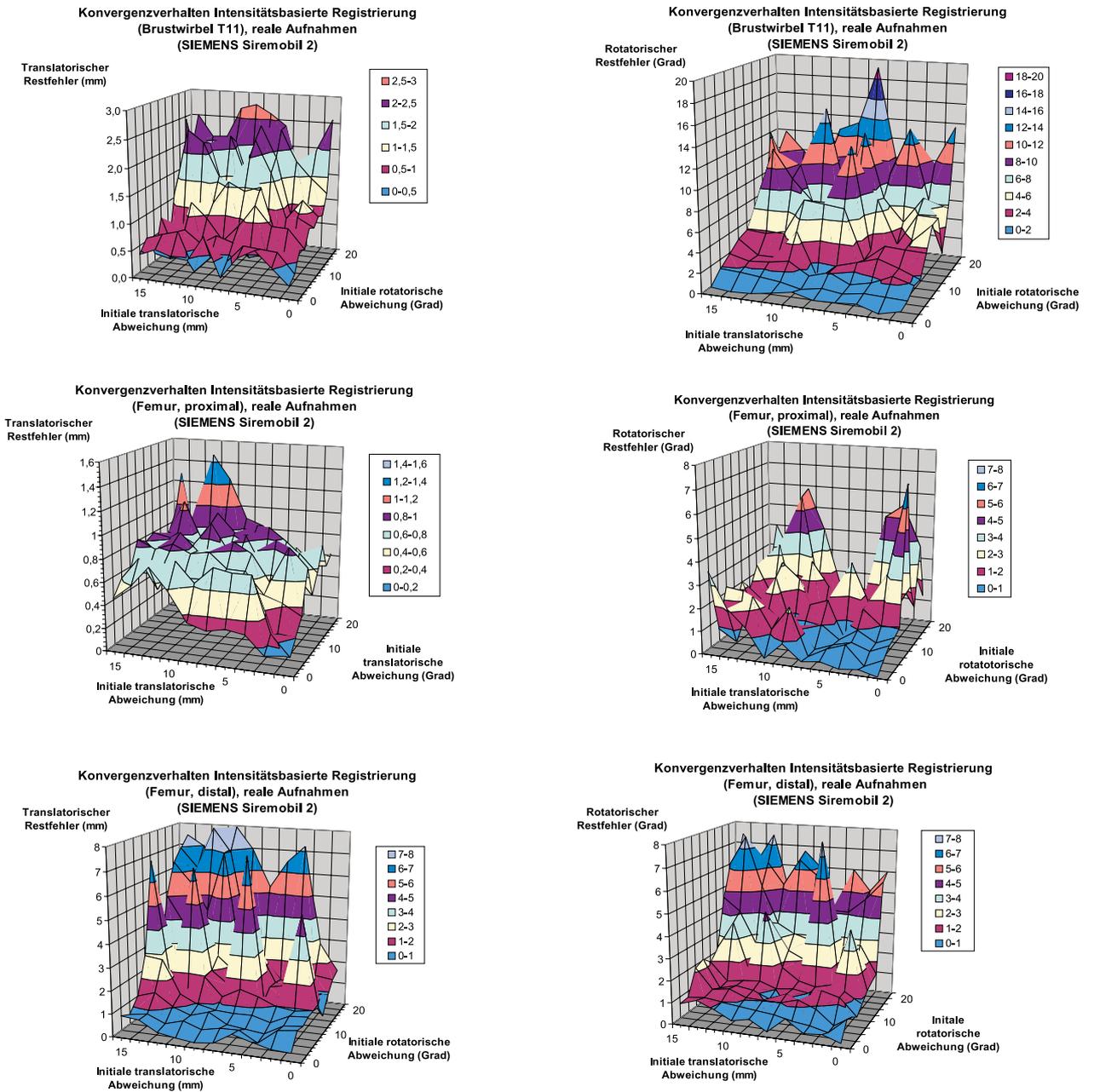


Abbildung 5.28: Konvergenzverhalten für reale Testdaten (jeweils drei Aufnahmen, Winkelabstand der Blickrichtungen zwischen 45° und 70°) für einen Brustwirbel (T11) sowie einen distalen und einen proximalen Femur. Die Fluoroskopie-Aufnahmen erfolgten einerseits mit einem *SIEMENS Siremobil 2* (T11, proximaler Femur), andererseits mit einem *ZIEHM Exposcop* (distaler Femur). Abgesehen von verkleinerten initialen Abweichungen (Rotation) entspricht die Darstellung der Ergebnisse der in Abb. 5.21.

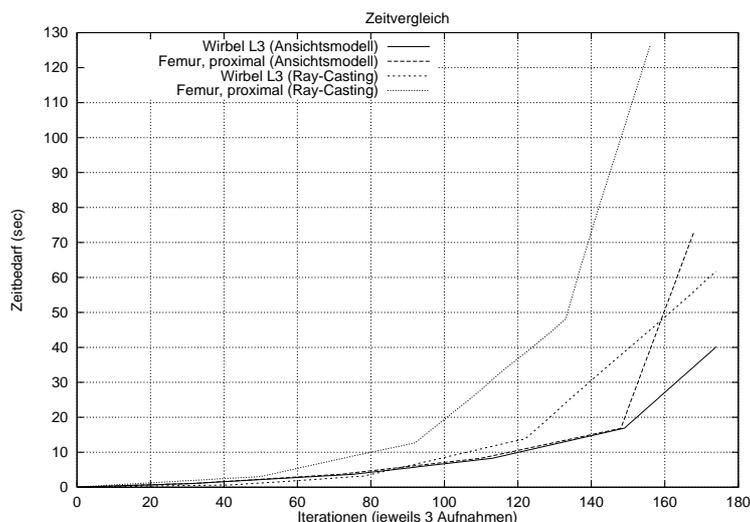


Abbildung 5.29: Vergleich des Zeitbedarfs für die Registrierung mit und ohne Verwendung von Ansichtsmodellen zur beschleunigten Berechnung von synthetischen Röntgenbildern

Aufwand für Ray-Casting und Eigenraum-Rekonstruktion mit anschließender planarer Approximation die Waage. Erst ab einer Bildgröße von 128×128 wird der Unterschied insbesondere für den Femur-Datensatz deutlich – bei der Lokalisierung des L3-Wirbels sogar erst bei voller Auflösung von 256×256 Pixeln.

Die Optimierung ohne Beschleunigung benötigt für den Femur-Datensatz insgesamt 126 Sekunden, für den L3-Datensatz mit gut 60 Sekunden knapp die Hälfte der Zeit. Die Verwendung der Ansichtsmodelle spart im ersten Fall 43 Prozent (54 Sekunden), im zweiten Fall 32 Prozent (22 Sekunden) der Rechenzeit ein. Dabei fällt auf, daß die Einsparung noch deutlich stärker ausfällt (86 bzw. 70 Prozent Einsparung), läßt man den letzten Iterationslauf der Optimierung mit Ansichtsmodellen zunächst unberücksichtigt. Wie in Abschnitt 5.4.3 beschrieben, muß dieser abschließende Lauf ohne Ansichtsmodelle nach Konvergenz der Optimierung erfolgen, um evtl. zuvor aufgetretene Approximationsfehler auszugleichen. Da das Ray-Casting dabei mit voller Bildauflösung erfolgt, benötigt dieser Schritt mit 54 (Femur) bzw. 44 Sekunden (L3-Wirbel) ein Vielfaches der gesamten vorangegangenen Optimierungszeit von etwa 18 Sekunden. Ein Großteil der Beschleunigung geht hierdurch also wieder verloren.

Sowohl für die Erstellung der Ansichtsmodelle, als auch für das Ray-Casting wurden CT-Datensätze mit einem Schichtabstand von 2mm herangezogen. Es gilt zu beachten, daß eine Verbesserung der Auflösung des verwendeten CT-Datensatzes durch die Verringerung des Schichtabstandes die Performance der Ray-Casting-Variante deutlich schmälern würde. Der Zeitbedarf für die intraoperative Auswertung eines entsprechenden Ansichtsmodells jedoch bliebe exakt derselbe. Hier würde sich lediglich die Dauer des präoperativen Trainings verlängern. In dieser Zeitkonstanz für die Rekonstruktion eines Röntgenbildes liegt eine der wesentlichen Stärken von Ansichtsmodellen. Die Verringerung des Schichtabstandes setzt natürlich voraus, daß ein entsprechend aufgenommenes CT des betrachteten Patienten überhaupt vorliegt. Alternativ lassen sich auch fehlende Zwischenschichten interpolieren. Die Interpolation bringt allerdings im Gegensatz zu einem echten CT-Datensatz keinen tatsächlichen Informationsgewinn. Daraus rekonstruierte Röntgenbilder enthalten jedoch weniger Artefakte, die im Fall signifikanter

Änderungen beim Übergang von einer CT-Schicht zur nächsten auftreten können, und wirken somit deutlich realistischer.

Für die Rekonstruktion und planare Approximation eines einzelnen Bildes der Auflösung 256×256 wurden im vorangegangenen Experiment jeweils ca. 0.08 Sekunden benötigt. Diese Zeit hängt u.a. von der Anzahl der in Gl. 5.11 eingerechneten Eigenvektoren ab. Im vorliegenden Fall waren es insgesamt 15. Laut Tabelle 5.1 werden hiermit gut 91 Prozent der in den ursprünglichen Trainingsdaten enthaltenen Variation modelliert. Abb. 5.30 veranschaulicht, wie sich die Anzahl der eingerechneten Eigenvektoren auf die Qualität des resultierenden Bildes auswirkt. Wieviele Eigenvektoren ein trainiertes Ansichtsmodell letztlich nutzt, ist im Einzelfall zu prüfen, kann aber z.B. an die zugehörige Variationsanalyse gekoppelt werden.

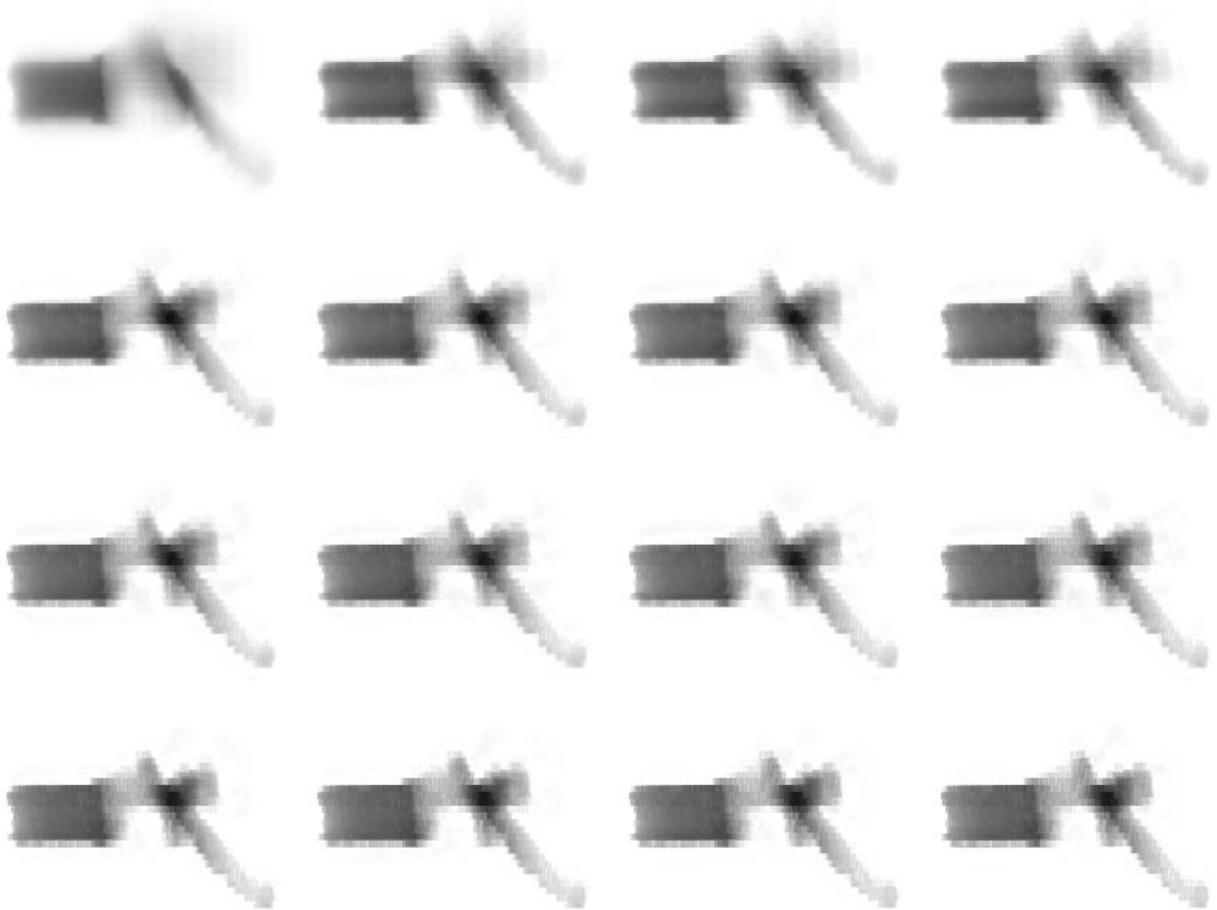


Abbildung 5.30: Rekonstruktion eines Röntgenbildes bei Verwendung von 1, 2, 3, ..., 11, 12, 14, 16, 18 bzw. 20 Eigenvektoren (von links oben nach rechts unten)

Kapitel 6

Elastische Registrierung eines formvariablen 3D-Modells mit Fluoroskopie-Aufnahmen

Inhalt

6.1 Zielsetzung und Ansatz	90
6.1.1 Motivation	90
6.1.2 Ansatz	91
6.2 Modell-Repräsentation	94
6.3 Physikalisch motivierte Modellierung (FEM)	95
6.4 Statistische Modellierung durch Training	97
6.4.1 Ansatz	97
6.4.2 Berechnung der statistischen Kenngrößen	98
6.4.3 Erzeugung von Trainings-Datensätzen	100
6.4.4 Initiale Bestimmung weniger Basis-Korrespondenzen	101
6.4.4.1 Auswahl markanter und stabiler Merkmale	101
6.4.4.2 Automatische Korrespondenzfindung (Labelling)	105
6.4.5 Ausrichtung der Modelle	107
6.4.5.1 Rigide Ausrichtung	109
6.4.5.2 Elastisches Alignment (Abschätzung einer Deformation)	110
6.4.6 Vollständige Bestimmung von Korrespondenzen	113
6.5 Konturbasierte, elastische Lokalisation	115
6.5.1 Segmentation von Fluoroskopieaufnahmen	115
6.5.2 Optimierungskriterium	116
6.5.2.1 Effiziente Konturbestimmung	119
6.5.2.2 Korrespondenzsuche	120
6.5.3 Initialisierung	120

6.5.4	Optimierungsvarianten	121
6.5.4.1	Geschlossene Lösung der Deformation	121
6.5.4.2	Gradientenabstieg (Levenberg-Marquardt)	123
6.5.4.3	Randomisierte Suche	124
6.5.4.4	Rastersuche	124
6.5.5	Optimierungsstrategie	124
6.6	Ergebnisse	126
6.6.1	Simulationsergebnisse	126
6.6.2	Anwendung auf reale Fluoroskopieaufnahmen	131

6.1 Zielsetzung und Ansatz

6.1.1 Motivation

Grundlegende Voraussetzung für das im letzten Kapitel vorgestellte intensitätsbasierte Registrierungsverfahren ist das Vorhandensein eines patientenspezifischen, dreidimensionalen Computer-Tomographie-Datensatzes. Die Volumendaten dienen dort direkt als Basis für die intraoperative Generierung simulierter Röntgenaufnahmen (Abschnitt 5). Auch bei einer Vielzahl anderer Verfahren spielt CT eine zentrale Rolle. Oftmals wird präoperativ aus dem CT-Datensatz des Patienten eine Zwischenrepräsentation in Form eines 3D-Oberflächenmodells erzeugt (vgl. Anhang D). Dieses Modell ermöglicht anschließend sowohl eine präzise Operationsplanung und -durchführung. Der Einsatz einer CT ist allerdings aus mehreren Gründen keineswegs unproblematisch:

- Der Patient wird einer zusätzlichen Strahlenbelastung ausgesetzt, die z.B. im Vergleich zur intraoperativen Verwendung eines Röntgenbildverstärkers signifikant höher ausfällt. Dies gilt im vorliegenden Fall umso mehr, da im Hinblick auf eine für die Registrierung ausreichend gute Auflösung des 3D-Modells bzw. der simulierten Röntgenaufnahmen der Abstand zwischen aufeinanderfolgenden Tomographieschichten sehr gering gewählt werden muß, i.a. im Bereich weniger Millimeter ($1 - 2\text{mm}$).
- Die Strahlenbelastung muß insbesondere dann kritisch gesehen werden, wenn die Aufnahmen nicht die Extremitäten (Arme, Beine), sondern z.B. Teile der Wirbelsäule betreffen, da sich hier in direkter Nähe eine Vielzahl sensibler Organe befindet.
- Die Aufnahme der tomographischen Bilddaten verlängert den stationären Aufenthalt des Patienten und damit die Gesamtdauer der Behandlung deutlich.
 - Nichtzuletzt aufgrund des zuvor genannten Punktes fallen zusätzliche Personal-, Raum- und Gerätekosten an, die die patientenbezogenen Gesamtkosten der Operation ansteigen lassen. Im Hinblick auf die aktuelle politische Diskussion über Gesundheitsreform und drastische Kosteneinsparungen der Krankenkassen, stellt sich in vielen Fällen die Frage nach der tatsächlichen Notwendigkeit einer Tomographie.
 - Bei Notfall-Patienten ist es zeitlich oft nicht mehr möglich, vor dem Eingriff eine Computer-Tomographie durchzuführen, obwohl die dadurch gewonnene 3D-Information mitunter von essentiellen Nutzen wäre.

Abgesehen von der Strahlenbelastung gelten die aufgeführten Kritikpunkte gleichermaßen auch für die Magnetresonanztomographie (MRT).

Es drängt sich letztendlich die Frage auf, ob mitunter der Einsatz einer 3D-Tomographie vermieden werden kann, ohne dabei auf die mit MRT oder CT verbundenen Vorteile für eine präzise räumliche Navigation zu Lasten des Patienten verzichten zu müssen.

Andererseits stellen z.B. die Fluoroskopie-Aufnahmen eines Röntgenbildverstärkers zunächst lediglich 2D-Information zur Verfügung. Daran ändern auch eine Kalibrierung des Gerätes (vgl. Anhang B), dessen Einbindung in ein Navigationskonzept und mehrere Aufnahmen aus unterschiedlichen Richtungen nichts (vgl. z.B. [Hof97]). Eine sehr gute räumliche Vorstellungskraft des Chirurgen, kombiniert mit dessen Wissen über die Anatomie, ist hier nachwievor unabdingbar.

Es ist einleuchtend, daß die vorgestellte intensitätsbasierte Registrierung (Kapitel 5) verfahrensbedingt nicht ohne CT-Daten funktionieren kann. Konturbasierte Ansätze wie z.B. in [Lav95] erfordern jedoch zunächst nur das Vorhandensein eines patientenspezifischen 3D-Oberflächenmodells. Sofern es gelingt, ein solches Modell unabhängig von CT- oder MRT-Aufnahmen zu erstellen, wird auch eine präzise räumliche Planung und Navigation ohne Tomographie möglich.

Aus dieser Motivation heraus wird im vorliegenden Kapitel eine neue Registrierungsmethode entwickelt, die auf der Generierung eines formvariablen anatomischen Modells basiert, welches sich ähnlich wie im Falle bekannter rigider Verfahren an in den Fluoroskopieaufnahmen segmentierte Konturen anpassen läßt, dabei aber ohne patientenspezifische Tomographiedaten auskommt.

6.1.2 Ansatz

Die Grundidee hinter dem im folgenden vorzustellenden Verfahren besteht darin, allein auf Basis mehrerer segmentierter Konturen eines Knochens in Röntgenaufnahmen unterschiedlicher Aufnahmerichtungen ein detailreiches, patientenspezifisches Modell dieses Knochens zu generieren.

Eine anatomisch vollständig korrekte 3D-Rekonstruktion aus wenigen Röntgenaufnahmen – bzw. den darin segmentierten Silhouetten – ist in der Regel aufgrund der spärlichen 2D-Information nur unter expliziter Verwendung von *a-priori*-Wissen möglich. Ohne solches Vorwissen ergibt sich ggf. lediglich eine grobe Volumenabschätzung, indem die – sich für jede Fluoroskopieaufnahme ergebenden – *Projektionslinienbüschel* räumlich geschnitten werden (Abb. 6.1) [Bra00a]). Das jeweils resultierende Volumen stellt eine grobe äußere Schranke für die tatsächliche Lage des betrachteten Knochens dar. Ein solches *Schnittvolumen* fällt dabei umso präziser aus, je größer die Anzahl der verwendeten Röntgenaufnahmen bzw. deren Abstand untereinander ist. Allerdings erreichen auf diese Weise erzeugte Oberflächen nie den Detaillierungsgrad vollwertiger Modelle. Insbesondere rundum konkave Vertiefungen der Oberfläche bleiben bei der Methodik zwangsläufig unentdeckt, da die zugehörigen Vertices aus keiner Blickrichtung einen Beitrag zur sichtbaren Außenkontur leisten.

Ein solches Schnittvolumenmodell ist daher aufgrund der beschriebenen Einschränkungen für eine präzise Navigation meist ungeeignet.

Wesentlich vielversprechender hingegen ist die Integration von vorhandenem Wissen über die Anatomie eines speziellen Knochentyps in Form eines elastischen Modells. Dieses Modell speichert dabei einerseits die durchschnittliche Oberflächenform des betrachteten Knochens, erlaubt aber andererseits auch die globale und lokale Verformung dieser Oberfläche, um eine bestmögliche Anpassung an die individuelle Situation für einen neuen Patienten zu erreichen.

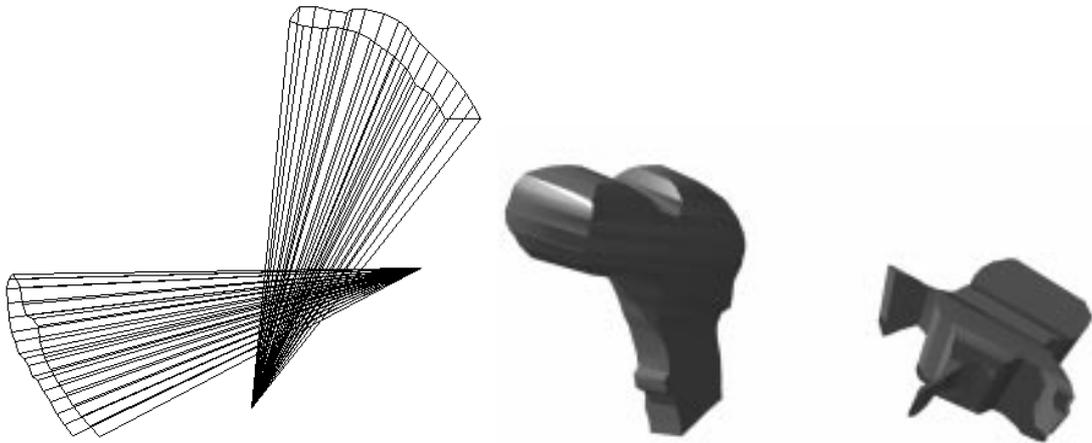


Abbildung 6.1: Der räumliche Schnitt mehrerer Projektionslinienbüschel (links) ergibt ein grobes Begrenzungsvolumen, das den zu registrierenden Knochen vollständig enthält, aber keinen weiteren Rückschluß auf dessen tatsächliche Oberfläche erlaubt. Selbst bei Verwendung vieler Röntgenaufnahmen aus sehr unterschiedlichen Richtungen (z.B. orthogonal) ist nicht sicher, ob der Knochen unentdeckte, konkave Oberflächenstücke enthält. Bsp. für Schnittvolumina: proximaler Femur (Mitte), L1-Lendenwirbel (rechts).

In der vorliegenden Arbeit werden zwei grundsätzlich konträre Vorgehensweisen zur Modellierung der gewünschten Verformung betrachtet:

- **Physikalisch motivierte Modellierung:** Dabei wird einem einzelnen Oberflächenobjekt eines repräsentativen starren Knochens ein gewisser Grad an *künstlicher* Elastizität unterstellt, indem die diskrete Menge der Objektvertices als physikalisches *Feder-Masse-System* aufgefaßt wird. Durch Anwendung der Modalen Analyse – einer Technik aus dem Bereich numerischer *Finite-Elemente-Methoden* (FEM) – ergeben sich dabei sogenannte physikalische Vibrationsmodi. Diese erlauben eine parametrisierte Verformung des einzelnen Ausgangsmodells. Diese Vorgehensweise wird in Abschnitt 6.3 näher beschrieben.
- **Statistische Modellierung:** Anstatt einem betrachteten Knochen willkürlich elastisches Material zu unterstellen, ist es auch möglich, keinerlei Grundannahmen zu treffen und stattdessen mehrere beispielhafte Oberflächenmodelle zu sammeln. Die statistische Analyse der Variation jeweils korrespondierender Oberflächenpunkte liefert mit Hilfe einer Hauptachsentransformation anschließend, wie im Fall der physikalischen Modellierung, ein parametrisiertes Verformungsmodell – ein sogenanntes Punktverteilungsmodell (*point distribution model*, PDM), dessen Freiheitsgrade als experimentelle Variations-Modi bezeichnet werden. Mit diesem Verfahren setzt sich Abschnitt 6.4 intensiv auseinander.

Der Vorteil des physikalischen Ansatzes ist, daß er relativ einfach zu realisieren ist, da lediglich ein einzelnes Modell bzw. Knochenobjekt benötigt wird, um eine Vielzahl möglicher Variationen zu erzeugen. Nachteilig wirkt sich jedoch aus, daß ein Knochen inhärent rigide ist, d.h. seine Verformbarkeit ohne

starke äußere Krafteinwirkung vernachlässigt werden kann. Die modellierende Verformung erfolgt hier also rein künstlich und damit in gewisser Weise willkürlich.

Dies ist bei dem statistischen Ansatz gerade nicht der Fall. Das Punktverteilungsmodell repräsentiert exakt die tatsächlich bzgl. unterschiedlicher Patienten zu beobachtenden Variationen eines Knochens, solange die Menge an beispielhaften Trainingsmodellen groß genug ausfällt und insbesondere für alle signifikanten Formunterschiede auch entsprechende Modelle enthält.

Im Gegensatz zu der FEM-Methodik ist dieser statistische Ansatz jedoch nicht ohne weiteres umzusetzen. Dies hat zwei Gründe. Zum einen ist die Beschaffung repräsentativen, klinischen Datenmaterials zur Generierung von Beispielmodellen mitunter keine leichte Aufgabe (Abschnitt 6.4.3). Zum anderen ist die Bestimmung korrespondierender Oberflächenpunkte auf den Trainingsmodellen ein kritischer und aufwendiger Vorverarbeitungsschritt. Die Beschreibung und Lösung der damit verbundenen Teilprobleme ist zentraler Inhalt der Abschnitte 6.4.4 bis 6.4.6.

Die vorliegende Arbeit empfiehlt die Kombination beider Methoden (FEM und PDM), indem grundsätzlich der statistische PDM-Ansatz Verwendung findet, jedoch erst *nachdem* die Menge an vorhandenen Trainingsobjekten durch Anwendung der FEM-Technik auf einzelne Objekte vergrößert wurde, falls die Menge nicht ausreichend mächtig ist (vgl. [CT95] (2D-Fall)).

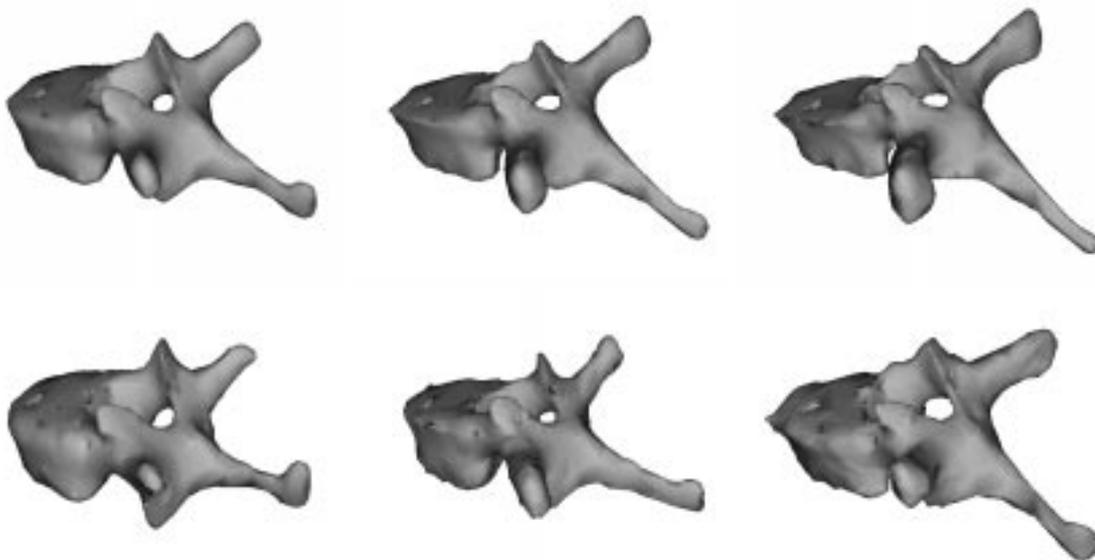


Abbildung 6.2: Beispiel für anatomisch plausible Formen menschlicher Wirbel, die durch Parametervariation eines elastischen Modells erzeugt wurden

Im Vergleich zu rigider Registrierung (Abschnitt 5) ist das hierbei entstehende verformbare Modell nicht mehr patientenspezifisch, sondern repräsentiert je nach Variation der verwendeten Trainingsbeispiel-Modelle eine mehr oder weniger breite Palette an möglichen Knochenformen, die anatomisch plausibel sind (Abb. 6.2).

Eine Spezialisierung auf ein konkretes rigides Modell ergibt sich erst wieder durch Anwendung des verformbaren Modells im Rahmen der Registrierung. Dabei werden neben den rigiden Parametern für

Translation und Orientierung zusätzlich die Skalierung sowie die vorhandenen Verformungsparameter des Modells bestimmt. Dies erfolgt konturbasiert, d.h. durch Anpassung des Modells an mehrere segmentierte Konturen in Röntgenaufnahmen unterschiedlicher Blickrichtung (Abschnitt 6.5). Da die Registrierung ambulant, also insbesondere ohne Öffnung des Patienten, erfolgen kann, ist deren Einsatz sowohl präoperativ als auch intraoperativ vorstellbar:

- **Präoperativ:** Das elastische Modell wird zunächst an die segmentierten Röntgenaufnahmen eines Patienten angepaßt. Die dabei resultierende Form des Modells wird anschließend als rigides patientenspezifisches Modell weiterverwendet – z.B. für präoperative Planungszwecke oder auch intraoperativ für eine *rigide* Registrierung.
- **Intraoperativ:** Die gleiche Vorgehensweise (Röntgenaufnahmen, Segmentierung und elastische Registrierung) intraoperativ angewandt, liefert sehr schnell eine räumliche Navigationshilfe, z.B. bei Notfalloperationen, bei denen keine zeitaufwendige Vorbereitung mehr möglich ist.

Grundsätzlich problematisch ist die Anwendung der elastischen Registrierung auf einen pathologischen, d.h. krankhaft veränderten Knochen. Da sich solche Veränderungen hinsichtlich ihrer Art und Ausbreitung individuell sehr stark unterscheiden, lassen sie sich statistisch nicht erfassen. In einem solchen Fall ist nach erfolgter Registrierung die lokale Nachbearbeitung des verformten Modells unabdingbar.

Die nachfolgenden vier Abschnitte beschäftigen sich zunächst mit der geeigneten Repräsentation eines formvariablen 3D-Modells (Abschnitt 6.2), anschließend mit der physikalisch (Abschnitt 6.3) und statistisch (Abschnitt 6.4) motivierten Modellgewinnung und schließlich mit der Anwendung formvariabler Modelle zur konturbasierten Registrierung (Abschnitt 6.5).

6.2 Modell-Repräsentation

Die mathematische Modellierung von Formvariation wird im Rahmen dieser Arbeit auf eine Topologieerhaltende Formvariation zunächst rigider, dreidimensionaler Oberflächenmodelle beschränkt. Derartige 3D-Modelle werden üblicherweise durch eine netzartige Struktur, die sich hierarchisch in 3D-Punkte (Vertices), Kanten und Flächen unterteilt, beschrieben, wobei die beiden letztgenannten Komponenten (Kanten, Flächen) für die Topologie des Modells verantwortlich sind (Anhang D). Eine parametrisierte, topologieerhaltende Deformation kann daher nur durch Manipulation der Vertex-Positionen selbst erfolgen:

$$f(\mathbf{x}, b_1, b_2, \dots, b_{n_b}) \rightarrow \mathbf{x}' \quad (6.1)$$

Dabei faßt der Vektor $\mathbf{x} = [x_1, y_1, z_1, \dots, x_n, y_n, z_n]^T$ die Koordinaten der insgesamt n Modellvertices $v_i = (x_i, y_i, z_i)$ zusammen. Entsprechend ergeben sich nach erfolgter Verformung durch die Funktion f neue Vertex-Koordinaten $\mathbf{x}' = [x'_1, y'_1, z'_1, \dots, x'_n, y'_n, z'_n]^T$, wobei die Deformationsparameter b_i die Art der Vertex-Beeinflussung festlegen.

Die Realisierung der Deformationsabbildung f wird in dieser Arbeit stets als Linearkombination von n_b Deformationsvektoren \mathbf{d}_i verstanden:

$$f(\mathbf{x}, b_1, b_2, \dots, b_{n_b}) = \mathbf{x} + b_1 \mathbf{d}_1 + \dots + b_{n_b} \mathbf{d}_{n_b} \quad (6.2)$$

Jeder Vektor \mathbf{d}_i besitzt die Dimension $3n$ und beschreibt die räumliche Bewegungsrichtung aller Vertices bei Veränderung des korrespondierenden Deformationsparameters b_i . Werden die Deformationsvektoren als Spalten einer Matrix \mathbf{P} interpretiert und die Deformationsparameter b_i in einem Vektor

$\mathbf{b} = [b_1, b_2, \dots, b_{n_b}]^T$ zusammengefaßt, so läßt sich Gl. 6.2 kompakt in Matrixschreibweise darstellen:

$$f(\mathbf{x}, b_1, b_2, \dots, b_{n_b}) = \mathbf{x} + \mathbf{P}\mathbf{b} \quad (6.3)$$

Neben der Einschränkung einer Topologie-erhaltenden Verformung ergibt sich durch die Wahl einer Linearkombination eine weitere Restriktion. Die Veränderung eines Parameters b_i resultiert aus Sicht eines einzelnen Vertex \mathbf{v}_j stets in seiner Verschiebung entlang einer Geraden $\mathbf{v}_j + \mu \mathbf{g}_{ji}$. Selbstverständlich können sich dabei unterschiedliche Vertices auch entlang unterschiedlicher Geraden bewegen. Die drei Komponenten des 3D-Richtungsvektors \mathbf{g}_{ji} ergeben sich direkt aus den Komponenten $3j, 3j + 1$ und $3j + 2$ des i -ten Deformationsvektors \mathbf{d}_i .

In der vorliegenden Arbeit steht die Modellierung patientenspezifischer Formunterschiede eines bestimmten Knochens im Vordergrund. Dabei können extreme topologische Unterschiede im Normalfall ausgeschlossen werden. Desweiteren erlaubt der Variationsspielraum der zu modellierenden Deformation deren Darstellung in Form einer Linearkombination. Beide Einschränkungen stellen folglich für die Zielsetzung des Kapitels kein Problem dar. Wie sich jedoch in den folgenden Abschnitten zeigen wird, vereinfacht diese Darstellung der Deformation auf der anderen Seite deren Handhabung enorm.

Hinsichtlich der Deformationsgleichungen 6.2 bzw. 6.3 besteht das gemeinsame Ziel der nachfolgenden beiden Abschnitte in der Bestimmung geeigneter Deformationsvektoren \mathbf{d}_i bzw. der Matrix \mathbf{P} . Die beiden vorgestellten Ansätze unterscheiden sich lediglich in ihrer spezifischen Vorgehensweise: physikalisch motiviert bzw. statistisches Training.

6.3 Physikalisch motivierte Modellierung (FEM)

Mit Hilfe der in Anhang E vorgestellten Finite-Elemente-Methode ist es möglich, ein einzelnes starres Knochenmodell so zu modellieren, als bestünde es aus einem beliebigen flexiblen Material. Durch die Modale Analyse läßt sich anschließend eine Menge an Schwingungsmodi ϕ_i berechnen, wobei diese Modi direkt als Deformationsvektoren entsprechend der im letzten Abschnitt definierten Repräsentation eines formvariablen Modells dienen können.

Die genannten Schwingungsmodi ergeben sich für den Fall einer freien und ungedämpften Schwingung eines elastischen Körpers. Mathematisch entsprechen sie den Eigenvektoren des nachfolgenden verallgemeinerten Eigenwertproblems (Gl. E.3):

$$\mathbf{K}\phi_i = \omega_i^2 \mathbf{M}\phi_i \quad (6.4)$$

Dabei repräsentiert \mathbf{M} die Massematrix und \mathbf{K} die Steifheitsmatrix des betrachteten Körpers.

Ausgehend von einem einzelnen starren Oberflächenmodell eines beliebigen Knochens läßt sich ein sehr einfaches Feder-Masse-System definieren. Zunächst wird mit jedem Modellvertex ein Knoten des Systems mit dem Gewicht einer Einheitsmasse assoziiert. Weiter wird angenommen, daß jeder Vertex mit jedem anderen Vertex des Modells über eine Feder konstanter Steifheit verbunden sei. Die Länge der einzelnen Federn soll dabei exakt dem Abstand zwischen den dadurch verbundenen Vertices entsprechen. Unter diesen Annahmen vereinfacht sich die obige Eigenwertgleichung, da die Massematrix \mathbf{M} zu einer Einheitsmatrix degeneriert. Die Konstruktion der $3N \times 3N$ -Steifheitsmatrix \mathbf{K} erfolgt durch geradlinige Erweiterung des in [Coo95a] vorgestellten 2D-Ansatzes auf 3D, wobei das Ausgangsmodell

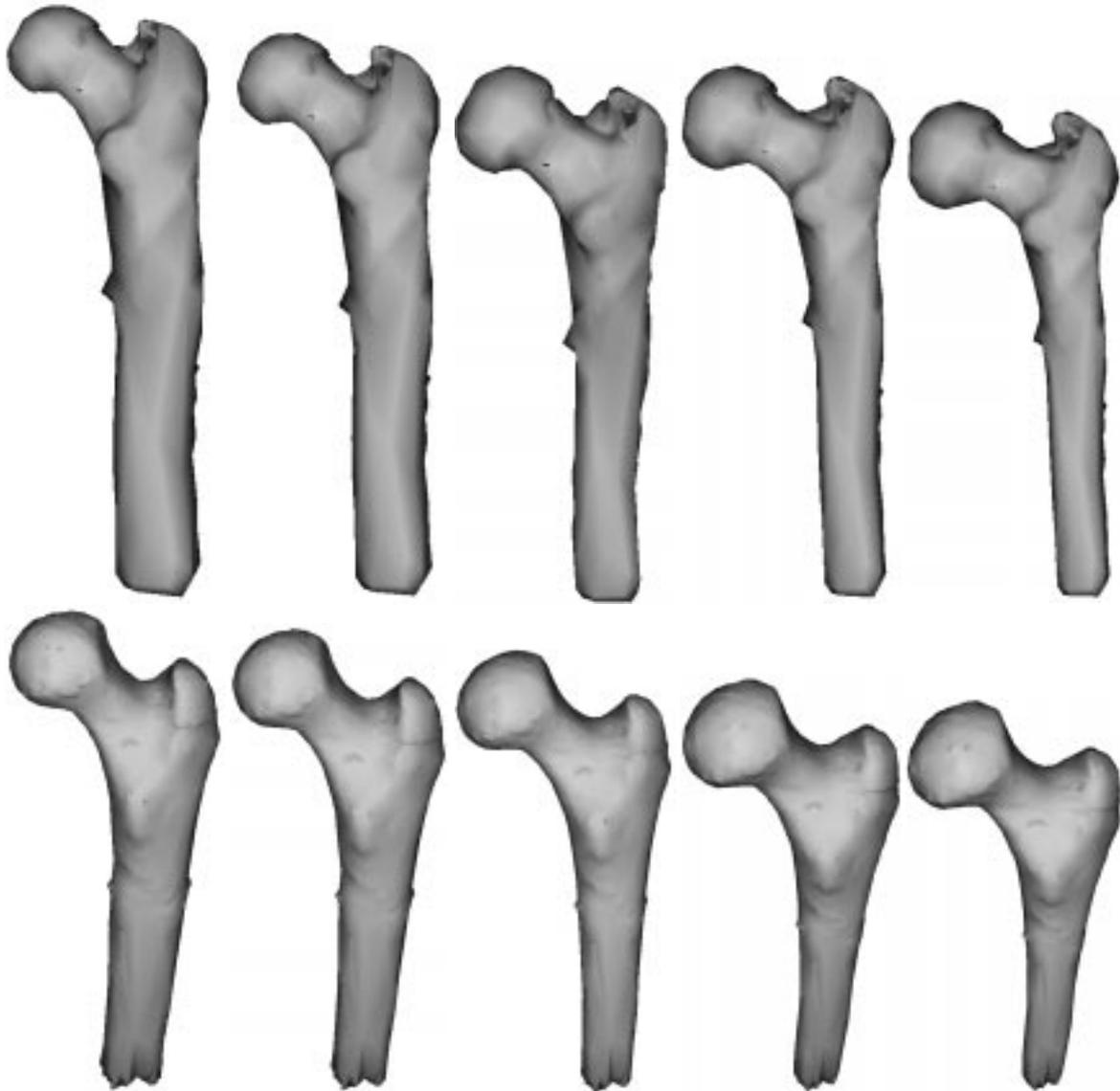


Abbildung 6.3: Jeweils stärkste FEM-Vibrationsmodi zweier als Feder-Masse-System modellierter proximaler Femur-Modelle

aus N Vertices $v_i(x_i, y_i, z_i)$ besteht:

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{xx} & \mathbf{K}_{yx} & \mathbf{K}_{zx} \\ \mathbf{K}_{xy} & \mathbf{K}_{yy} & \mathbf{K}_{zy} \\ \mathbf{K}_{xz} & \mathbf{K}_{yz} & \mathbf{K}_{zz} \end{pmatrix} \quad (6.5)$$

Die insgesamt neun $N \times N$ Untermatrizen $K_{\{x,y,z\}\{x,y,z\}}$ berechnen sich wie folgt:

$$\begin{aligned} \mathbf{K}_{yx} &= \mathbf{K}_{xy}^T & ; & & \mathbf{K}_{zx} &= \mathbf{K}_{xz}^T & ; & & \mathbf{K}_{zy} &= \mathbf{K}_{yz}^T \\ k_{xx_{ij}} &= \frac{d_{x_{ij}}^2}{d_{ij}^2} & ; & & k_{yy_{ij}} &= \frac{d_{y_{ij}}^2}{d_{ij}^2} & ; & & k_{zz_{ij}} &= \frac{d_{z_{ij}}^2}{d_{ij}^2} \\ k_{xy_{ij}} &= \frac{d_{x_{ij}}d_{y_{ij}}}{d_{ij}^2} & ; & & k_{xz_{ij}} &= \frac{d_{x_{ij}}d_{z_{ij}}}{d_{ij}^2} & ; & & k_{yz_{ij}} &= \frac{d_{y_{ij}}d_{z_{ij}}}{d_{ij}^2} \end{aligned}$$

wobei

$$\begin{aligned} d_{x_{ij}} &= x_i - x_j & ; & & d_{y_{ij}} &= y_i - y_j & ; & & d_{z_{ij}} &= z_i - z_j \\ d_{ij}^2 &= d_{x_{ij}}^2 + d_{y_{ij}}^2 + d_{z_{ij}}^2 \end{aligned}$$

Für die Diagonalkomponenten ($i = j$, $d_{ij}^2 = 0$) der Untermatrizen müssen die folgenden Sonderfälle berücksichtigt werden:

$$\begin{aligned} k_{xx_{ii}} &= k_{yy_{ii}} = k_{zz_{ii}} = 1 \\ k_{xy_{ii}} &= k_{xz_{ii}} = k_{yz_{ii}} = 0 \end{aligned}$$

Die gesuchten Vibrationsmodi ϕ_i ergeben sich anschließend direkt als Eigenvektoren der nach obigem Schema konstruierten Steifheitsmatrix \mathbf{K} .

Anhand zweier proximaler Femurmodelle veranschaulicht Abb. 6.3 den möglichen Variationsspielraum zweier repräsentativer Schwingungsmodi.

Der wesentliche Vorteil dieser physikalisch motivierten Modellierung von Formvariation besteht darin, daß bereits ein einziges Objektmodell genügt, um daraus durch unterschiedliche Gewichtungen der einzelnen Vibrationsmodi eine beliebige Anzahl an neuen verformten Modellen zu erzeugen. Allerdings sind die resultierenden Modelle oftmals anatomisch zu stark entartet oder unrealistisch proportioniert. Dies liegt daran, daß die berechneten Vibrationsmodi keinen Bezug zu den realen Gegebenheiten besitzen. Sie sind daher meist nicht repräsentativ für die bei Betrachtung einer Vielzahl an Patienten statistisch tatsächlich beobachtbaren Variationen.

Der folgende statistische Ansatz beseitigt diesen Nachteil. Allerdings erhöht sich damit zum einen der Aufwand für die Modellierung und zum anderen die hierfür benötigte Anzahl an Knochenmodellen signifikant.

6.4 Statistische Modellierung durch Training

6.4.1 Ansatz

Die FEM-Vibrationsmodi, die sich ausgehend von einem einzelnen Knochenmodell bestimmen lassen, sind leider in gewisser Weise willkürlich und repräsentieren oftmals nicht die gewünschten natürlichen Variationen, wie sie sich bei Betrachtung einer Vielzahl unterschiedlicher Patienten ergeben würden.

Dieses Problem kann umgangen werden, indem anstatt eines einzelnen Modells eine Vielzahl von repräsentativen Beispiel-Objekten gesammelt wird und man anschließend die Formabweichungen der einzelnen Modelle vom Durchschnitt aller Beispielmodelle mit Hilfe statistischer Methoden analysiert.

Der Vorteil dieser Vorgehensweise liegt darin begründet, daß hiermit eine kompakte Repräsentation der – im Hinblick auf unterschiedliche Patienten – tatsächlich beobachtbaren Variation modelliert wird. Dies gilt zumindest dann, falls sichergestellt ist, daß die Anzahl der Trainingsobjekte (Stichprobe) zum einen groß genug ist und zum anderen die in der statistischen Grundgesamtheit einer betrachteten Patientengruppe beobachtbaren Unterschiede ausreichend gut durch die Stichprobe approximiert werden.

6.4.2 Berechnung der statistischen Kenngrößen

Die statistische 3D-Formmodellierung dieses Kapitels basiert auf der geradlinigen und einfachen Dimensionserweiterung der *Point-Distribution-Model-Technik* [Coo95b] von 2D auf 3D. Ein einzelner Trainingsdatensatz wird dabei folgerichtig durch eine Menge eindeutig identifizierter 3D-Punkte repräsentiert. Diese Punkte entsprechen jeweils den Vertices des zugrundeliegenden triangulierten Trainings-Objektes, wobei durch die Objekt-Repräsentation zusätzlich dessen Topologie (Kanten- und Flächendefinitionen) festgelegt ist. Die Topologie der Trainingsmodelle spielt während der statistischen Analyse keine Rolle. Die Analyse konzentriert sich ausschließlich auf die räumliche Variation der Vertex-Positionen bzgl. aller Trainingsbeispiele.

Wenngleich sich der 2D-Ansatz zunächst sehr leicht auf den 3D-Fall übertragen läßt, so ergeben sich im Dreidimensionalen doch nichttriviale Problemstellungen sowohl bei der Bereitstellung der Trainingsdaten als auch bei der Gewährleistung bestimmter Voraussetzungen wie z.B. der vertextbezogenen 1:1-Korrespondenz von Trainingsbeispielen. Die Diskussion dieser Problemstellungen sowie der in vorliegender Arbeit gewählten Lösungsansätze ist eigentlicher Schwerpunktinhalt dieses Kapitels, wird aber im folgenden zunächst hinten angestellt.

Jedes der insgesamt m gegenseitig ausgerichteten Trainingsobjekte besitzt exakt n eindeutig benannte Punkte bzw. Vertices. Die einzelnen Punktkoordinaten werden zur Vereinfachung der mathematischen Darstellung für jedes Trainingsobjekt wie folgt in Form eines Vektors zusammengefaßt:

$$\mathbf{x}_i = (x_{i1}, y_{i1}, z_{i1}, \dots, x_{in}, y_{in}, z_{in})^T, i = 1, 2, \dots, m$$

Daraus läßt sich in einem ersten Schritt die Durchschnittsform $\bar{\mathbf{x}}$ (durchschnittliche Position aller Trainingspunkte) berechnen:

$$\bar{\mathbf{x}} = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i \quad (6.6)$$

Die relativen Positionen der Vertices jeder Trainingsform bezogen auf deren Positionen in der Durchschnittsform $\bar{\mathbf{x}}$ sind i.a. sehr stark miteinander korreliert. D.h. viele Vertices tendieren dazu, sich bei einer gleichmäßigen lokalen oder globalen Oberflächen-Deformation – wie z.B. der Verformung des Durchschnittsobjektes in Richtung eines der Trainingsobjekte – in gegenseitiger Abhängigkeit zu bewegen. Beispielsweise werden sich bei vielen Verformungen benachbarte Vertices in etwa in dieselbe Richtung bewegen. Andere Vertices wiederum bewegen sich exakt entgegengesetzt (Bsp.: Vertices auf gegenüberliegenden Seiten eines Wirbel-Dornfortsatzes bei dessen lokaler Vergrößerung).

Diese Korrelation zwischen einzelnen Vertices läßt sich nun wiederum quantitativ messen, indem die Kovarianzmatrix $\mathbf{C}_{\bar{\mathbf{x}}}$ bzgl. der Abweichungen aller Trainingsbeispiele von der Durchschnittsform $\bar{\mathbf{x}}$ wie folgt berechnet wird:

$$\mathbf{C}_{\bar{\mathbf{x}}} = \frac{1}{m} \sum_{i=1}^m [(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T] \quad (6.7)$$

bzw.

$$\mathbf{C}_{\bar{\mathbf{x}}} = \left(\frac{1}{m} \sum_{i=1}^m \mathbf{x}_i \mathbf{x}_i^T \right) - \bar{\mathbf{x}} \bar{\mathbf{x}}^T \quad (6.8)$$

Die beiden Berechnungsarten sind gleichwertig, wobei Gleichung 6.8 zu bevorzugen ist, falls die Menge der Trainingsobjekte später z.B. sukzessive durch neue Objekte erweitert wird. Durch die Trennung der \mathbf{x}_i - und $\bar{\mathbf{x}}$ -Terme läßt sich in solchen Fällen eine komplette Neuberechnung von $\mathbf{C}_{\bar{\mathbf{x}}}$ vermeiden. Stattdessen kann die bereits berechnete Kovarianzmatrix jeweils entsprechend leicht aktualisiert werden.

Die bereits in Abschnitt 5.4.2.2 angewandte PCA-Technik (*principal component analysis*, Hauptachsen- bzw. Karhunen-Loeve-Transformation, [Jol86]) erlaubt auch im vorliegenden Fall durch die Anpassung eines $3n$ -dimensionalen Ellipsoids an die Trainingsdaten (n Vertices je Trainingsmodell) linear korrelierte Vertex-Bewegungen zu erfassen. Jede Achse dieses Ellipsoids repräsentiert dabei einen linear unabhängigen *Variationsmodus*, der wiederum eine bevorzugte Art und Weise beschreibt, wie sich die Modellvertices im Falle einer Formveränderung des Modells gemeinsam bewegen.

Diese Variationsmodi ergeben sich durch Diagonalisierung der Kovarianzmatrix $\mathbf{C}_{\bar{\mathbf{x}}}$ als deren Eigenvektoren \mathbf{e}_i . Die n_b Eigenvektoren \mathbf{e}_i , ($i = 1, \dots, n_b$), die zu den n_b größten Eigenwerten λ_i gehören, bilden zusammen eine Menge orthogonaler Basisvektoren. Unter Verwendung dieser Basis läßt sich ein lineares Modell für ein verformbares Objekt definieren:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \quad (6.9)$$

Dabei enthält die $n \times n_b$ -Matrix \mathbf{P} spaltenweise die n_b Eigenvektoren \mathbf{e}_i : $\mathbf{P} = (\mathbf{e}_1 \mathbf{e}_2 \dots \mathbf{e}_{n_b})$. Der Vektor $\mathbf{b} = (b_1 b_2 \dots b_{n_b})^T$ vereinigt die freien Form-Parameter. Deren Variation ermöglicht durch Gl. 6.9 die Berechnung neuer, repräsentativer Modellformen \mathbf{x}' .

Die Beschränkung auf die n_b am stärksten ausgeprägten Eigenvektoren stellt sicher, daß aus den Trainingsdaten lediglich die wirklich signifikanten Variationen in das Modell übernommen werden – Rauscheffekte bleiben dabei außen vor. Störeinflüsse wie Rauschen führen i.a. sowieso dazu, daß sich Eigenvektoren höherer Ordnung nicht mehr numerisch stabil schätzen lassen. Eine Faustregel [MPK94] besagt: Nur die ersten 25% der Eigenvektoren können verlässlich berechnet werden. Beschränkt man daher n_b entsprechend, so ergibt sich eine sehr kompakte Darstellung der signifikanten Variation des betrachteten Objekts. Die Bestimmung des kleinstmöglichen n_b erfolgt analog zu Gleichung 5.12 aus Abschnitt 5.4.2:

$$\min_{n_b} \left(\frac{\sum_{i=0}^{n_b-1} \lambda_i}{\sum_{i=0}^{m-1} \lambda_i} \geq \tau \right) \quad (6.10)$$

Für den Schwellwert τ gilt wiederum: $\tau \in [0.0 \dots 1.0]$, entsprechend des zu modellierenden Variationsanteils von 0% bis 100%. Die Berechnung der 3D-Verformungsmodi selbst gestaltet sich damit nicht weiter kompliziert. Die eigentliche Schwierigkeit liegt vielmehr in den zu leistenden Vorarbeiten für die Bereitstellung der vertexweise korrespondierenden Trainingsdatensätze $\{\mathbf{x}_i\}$:

1. Akquisition von Trainingsmodellen

2. Gegenseitige Ausrichtung der Modelle (*Aligning*)
3. Etablieren von Korrespondenzen (*Labelling*)

Diese drei Themenbereiche nehmen jeweils eine Schlüsselstellung ein und entscheiden letztlich über die Qualität des resultierenden formvariablen Modells.

Der beschriebene Ansatz kompakt modellierter Formvariation findet bereits im letzten Kapitel im Zusammenhang mit dem Training von Ansichtsmodellen Anwendung (5.4.2). Dort stellen sich jedoch die drei Teilprobleme *Akquisition*, *Gegenseitige Ausrichtung* und *Korrespondenzsuche* als nicht existent oder zumindest als weitaus weniger schwierig dar. Dadurch, daß die Trainingsbilder des Knochens hier von vornherein mit konstanten internen Kameraparametern und insbesondere einem festem Abstand zwischen Kamera und Knochen generiert wurden, entsteht weder ein Alignment- noch ein Korrespondenzproblem. Die pixelweise Korrespondenz der Trainingsbilder ist bereits implizit durch die Pixel-Positionen vorgegeben. Die Akquisition von genügend Trainingsdaten ist ebenfalls problemlos möglich, da prinzipiell beliebig viele Röntgenbilder simuliert werden können.

In den folgenden vier Abschnitten werden Lösungsansätze für die genannten Problemstellungen im Zusammenhang mit 3D-Objekten vorgestellt.

6.4.3 Erzeugung von Trainings-Datensätzen

Die statistische Analyse basiert auf der Annahme, daß die Stichprobe der verwendeten Beispielobjekte repräsentativ ist, d.h. deren Verteilung der zugehörigen Grundgesamtheit entspricht. Nur dann ist gewährleistet, daß das resultierende formvariable Modell die bei verschiedenen Patienten tatsächlich beobachtbaren Variationen darzustellen vermag.

Leider ist es oftmals aber gerade nicht möglich, eine genügend große Anzahl an 3D-Trainingsdaten ein- und desselben Knochens verschiedener Patienten zu beschaffen¹. Aus diesem Grund wird in der vorliegenden Arbeit die Integration der vorgestellten FEM-Methodik (Abschnitt 6.3) in den PDM-Ansatz in folgendem Sinne vorgeschlagen:

1. Ausgangsbasis ist eine kleine Grundmenge repräsentativer 3D-Flächenmodelle M_i , die allesamt aus realen 3D-Patientendatensätzen generiert wurden: $M = \{M_1, M_2, \dots\}$.
2. Durch Anwendung der Modalen Analyse (FEM) läßt sich diese Menge erweitern, indem für jedes Beispielmodell M_i mehrere neue Modelle $M_i^{FEM} = \{M_i^{FEM_1}, M_i^{FEM_2}, \dots\}$ durch Variation der FEM-Vibrationsmodi erzeugt werden. Dabei ist unbedingt darauf zu achten, daß nur anatomisch plausible Modelle hinzugefügt werden. Diese *Qualitätskontrolle* kann entweder manuell oder auch automatisch durch sinnvolle Beschränkung der Formparameter erfolgen.
3. Durch Vereinigung der Grundmenge M mit den jeweils zusätzlich erzeugten FEM-Mengen M_i^{FEM} ergibt sich eine neue große Gesamtmenge an unterschiedlichen Modellen:

$$\{M_1, M_1^{FEM_1}, M_1^{FEM_2}, \dots, M_2, M_2^{FEM_1}, M_2^{FEM_2}, \dots, M_3, M_3^{FEM_1}, M_3^{FEM_2}, \dots\}$$

¹Zumindest galt diese Beschränkung auf sehr wenige Modelle während der Entstehung dieser Dissertation. Es ist jedoch durchaus denkbar, daß im klinischen Alltag gezielt eine umfangreiche Menge an 3D-Datensätzen einer Patientengruppe gesammelt wird. Ist dies der Fall, so kann auf den Einsatz der FEM-Verformung völlig verzichtet werden. Die Modellierung erfolgt dann ausschließlich auf Basis der vorhandenen 3D-Modelle.

Ehe jedoch auf diese erweiterte Menge an Trainingsmodellen die im letzten Abschnitt beschriebene statistische Variationsanalyse angewandt werden kann, ist es nötig, auf der Oberfläche jedes Modells eine feste Anzahl an Modell-Punkten (Vertices) zu bestimmen, zu denen auf jedem der restlichen Modelle entsprechende korrespondierende Punkte existieren.

6.4.4 Initiale Bestimmung weniger Basis-Korrespondenzen

Bei planaren formvariablen Modellen [Coo95c, Coo95b] u.U. noch vertretbar, ist eine manuelle, interaktive Definition von markanten Punkten an einer Vielzahl von 3D-Objekten sowie deren Korrespondenz ungleich umständlicher und insbesondere zeitraubender. Eine weitestgehende Automatisierung dieser Arbeitsschritte ist daher sehr wünschenswert. Im folgenden werden hierfür zwei Verfahren vorgestellt: zunächst eine automatische Detektion von anatomisch markanten Punkten (Landmarken) auf der Oberfläche eines Modellobjektes, anschließend eine automatische Korrespondenzbestimmung zwischen diesen Landmarken auf unterschiedlichen Modellen.

6.4.4.1 Auswahl markanter und stabiler Merkmale

Punkt-basierte Objektmerkmale, die sich für eine Charakterisierung der Formvariation des Objektes eignen, lassen sich u.a. in folgende Typ-Kategorien einteilen [Boo91, Coo95b]:

1. Punkte, die Teile eines Objektes markieren, wobei die Kennzeichnung modell- und anwendungsabhängig erfolgt. Ein Beispiel für einen solchen Punkt ist der Schwerpunkt des Hüftkopfes eines proximalen Femurknochens.
2. Punkte, die anwendungsunabhängig markante Objektdetails meist geometrischer Natur beschreiben, z.B. Oberflächenpunkte mit starker lokaler Krümmung oder Extrempunkte, die sich bei einer bestimmten definierten Objektausrichtung ergeben.
3. Sonstige Punkte, die sich durch bereits bekannte Punkte vom Typ 1 und 2 z.B. durch geeignete Interpolation bestimmen lassen.

In dieser Arbeit werden ausschließlich Modelle menschlicher Knochen betrachtet. Daher ergeben sich durch international standardisierte, anatomische Bezeichnungen markanter Knochendetails in den meisten Fällen bereits einige Merkmalspunkte vom Typ 1, z.B. die Walzenhöcker eines distalen Femurs oder auch die Gelenkflächen, Dorn- und Querfortsätze eines Wirbels (vgl. Anhang A). Interessanterweise lassen sich diesen anerkannten anatomischen Punkten oftmals auch geometrische Kriterien (Typ 2) zuordnen.

Im weiteren Verlauf dieses Kapitels werden solche hervorstechenden, anatomisch markanten Punktmerkmale vom Typ 1 oder 2 als anatomische Landmarken bezeichnet. Die folgende Definition präzisiert diese Bezeichnung:

Definition 2 (3D-Landmarke)

Eine Landmarke ist ein markantes, stabiles Oberflächenmerkmal eines Knochens, das sich stets eindeutig identifizieren läßt. Die eindeutige Kennzeichnung ist insbesondere auch an demselben Knochen einer Vielzahl unterschiedlicher Patienten möglich. Die Repräsentation einer Landmarke L_k^A eines Oberflächenmodells A des Knochens erfolgt durch eine Menge – nicht notwendigerweise benachbarter – Vertices $\{v_i\}$ bzw. durch deren räumlichen Schwerpunkt $\bar{L}_k^A = \frac{1}{n} \sum_1^n v_i$.

Das im weiteren Verlauf des Abschnitts präsentierte Verfahren stellt eine einfache automatische Alternative zur aufwendigen manuellen Bestimmung anatomischer Landmarken bzgl. eines gegebenen Objektmodells dar [Dö99]. Die Grundidee des Algorithmus basiert auf folgender Beobachtung: Die *Glättung* des betrachteten Objektes durch eine lokale arithmetische Mittelwertbildung der Objektvertices wird die Position solcher Vertices, die einem der gesuchten, markanten Oberflächendetails entsprechen, wesentlich stärker verändern als die aller restlichen Vertices. Dieser Effekt wirkt sich dabei umso stärker aus, je öfter die Glättung wiederholt wird.

Daher liefern alle Vertices des Objektes, deren Position sich durch iterative Glättung stärker verändert als die Positionen lokal benachbarter Vertices, einen ersten Hinweis auf das Vorhandensein einer Landmarke. Ausgehend von diesen Vertices lassen sich anschließend durch Gruppierung mit benachbarten Vertices Landmarken im Sinne der zuvor getroffenen Definition 2 ermitteln.

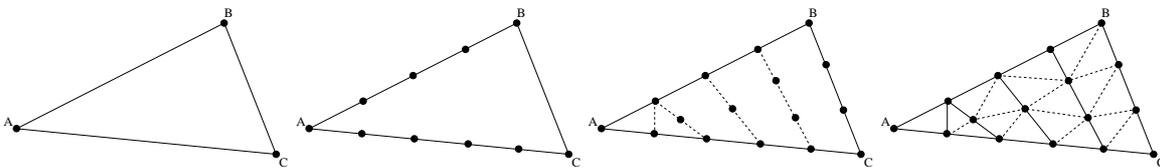


Abbildung 6.4: Re-Sampling der Objekt-Oberfläche: Dreiecke, deren Kantenlänge einen gegebenen Schwellwert überschreitet, werden gleichmäßig in kleinere Dreiecke unterteilt.

Das folgende Ablaufschema stellt die einzelnen Schritte des Algorithmus' zur Detektion von Landmarken ausführlich dar. Die Ergebnisse wichtiger Teilschritte werden dabei beispielhaft anhand eines menschlichen T6-Wirbels (Abb. 6.5) veranschaulicht.

1. Oberflächen-Resampling:

Die stabile Detektion von markanten Landmarken durch Glättung setzt voraus, daß die Vertices des Objektes auf dessen Oberfläche weitestgehend gleichverteilt sind. Ansonsten sind die durch die Glättung resultierenden Positionsveränderungen einzelner Vertices nicht direkt vergleichbar, da deren Nachbarvertices, die in die Mittelung mit einbezogen werden, unterschiedlich weit entfernt liegen können.

Gerade diese Voraussetzung der Gleichverteilung ist aber z.B. aufgrund zuvor angewandter Dezi- mierungsalgorithmen, die die Anzahl an Modell-Vertices, -Kanten und -Dreiecken adaptiv verringern, häufig nicht erfüllt [Gue96, Cia97, Coh96]. Deshalb muß die Oberfläche eines Modellobjektes M vor der eigentlichen Glättung zunächst gleichmäßig durch neue Vertices, Kanten und Dreiecke erweitert werden. Dieses dichte *Resampling* geschieht durch eine wiederholte Unterteilung bereits vorhandener Dreiecksflächen in kleinere Teildreiecke (Abb. 6.4) und resultiert schließlich in einem *dichten* Modellobjekt M' (Abb. 6.6).

2. Iterative Glättung:

Anschließend erfolgt die Glättung des dichten Objektes. Hierfür wird ein neues Objekt M'' gebildet, wobei zunächst sowohl die Anzahl der Vertices als auch die Topologie (Kanten, Flächendefinitionen) unverändert von M' übernommen wird. Die Koordinaten jedes Vertex v hingegen werden in M'' durch das arithmetische Mittel aus sich selbst sowie den entsprechenden Koordinaten der Nachbar-Vertices v_k – d.h. zwischen v und v_k verläuft eine Modellkante – ersetzt. Dieser Vorgang

wird iterativ wiederholt, wobei sich experimentell für die Anzahl der Durchläufe $\frac{3}{2}\sqrt{|M'|}$ ($|M|$: Vertexanzahl von M) als Anhaltswert bewährt hat [Dö99].

3. Abstandsberechnung:

Nach Abschluß der iterativen Glättung lassen sich die beiden Objekte M' (Ausgangsmodell) und M'' (geglättetes Modell) miteinander vergleichen, indem für jeden Vertex aus M' dessen Abstand zu seinem korrespondierenden Vertex aus M'' bestimmt wird. Große Abstände entsprechen dabei einer starken, kleine Abstände einer schwachen Glättung des Objektmodells an der Position des betrachteten Vertex. Die Abstandswerte werden daher im folgenden auch als Signifikanzwerte bezeichnet (Abb. 6.7).

4. Suche nach lokalen Maxima:

Die Grundidee des Algorithmus basiert, wie zuvor bereits angesprochen, auf der Erkenntnis, daß Vertices in anatomisch markanten Regionen der Modelloberfläche stärker auf die Glättung ansprechen als andere Vertices. Daher liefert eine Suche nach allen lokalen Maxima innerhalb der berechneten Signifikanzwerte eine Menge von Landmarken-Kandidaten in Form der zugehörigen Vertices.

5. Lokale Gruppierung:

Die Auszeichnung eines einzelnen Vertex als Landmarke führt i.a. zu instabilen Merkmalen. Kleinste Änderungen der Modelloberfläche können die Position der Landmarke bereits stark verändern. Es ist daher wünschenswert, statt einzelner Vertices jeweils zusammenhängende Oberflächen-Regionen von Vertices zu bestimmen, die gemeinsam hohe Signifikanzwerte aufweisen und daher zu einer stabilen Landmarke gruppiert werden können.

Diese Gruppierung erfolgt durch ein iteratives, oberflächenbasiertes Bereichswachstum, ausgehend von den Landmarken-Kandidaten aus dem vorangegangenen Schritt (*Seeded Region-Growing*). In jeder Iteration werden gleichzeitig für alle Landmarken-Kandidaten die jeweils benachbarten Vertices betrachtet und ggf. zu dieser Landmarke hinzugefügt, sofern die folgenden Kriterien bzgl. des betrachteten Vertex v erfüllt sind [Dö99]:

- v gehört noch keiner anderen Landmarke an,
- v besitzt einen hohen Signifikanzwert ($\geq s_{min}$),
- der Abstand zwischen v und dem Ausgangsvertex (Kandidat) überschreitet einen maximalen Abstand d_{max} nicht und
- die Anzahl aller Vertices der Landmarke beträgt maximal n_{max} .

Die drei Parameter s_{min} , d_{max} und n_{max} , die das Verhalten der Gruppierung steuern, sind dabei unkritisch, d.h. hier genügen Standardeinstellungen, die einmalig festgelegt werden.

Der vorgestellte Algorithmus ermöglicht die vollautomatische Detektion punktbezogener anatomischer Landmarken eines Oberflächenmodells. Allerdings liefert das Verfahren nur dann wirklich brauchbare Ergebnisse, wenn das betrachtete Objekt auch signifikante, lokal begrenzte Extrempunkte besitzt. Für das zuvor verwendete Beispielobjekt – den menschlichen Brustwirbel mit seinen Gelenkflächen (*Fovea costalis*), Dorn- (*Processus spinosus*) und Querfortsätzen (*Processus articularis/transversus*) – ist diese Voraussetzung erfüllt. Auch Hals- und Lendenwirbel besitzen eine Vielzahl an markanten Merkmalen,



Abbildung 6.5: Beispielobjekt (T6-Wirbel) für die Demonstration der automatischen Bestimmung anatomischer Landmarken (verschiedene Ansichten)



Abbildung 6.6: Beispielobjekt (Vertex-Darstellung) aus Abb. 6.5 nach erfolgtem *Resampling* (Verdichtung der Vertices) mit anschließender Glättung

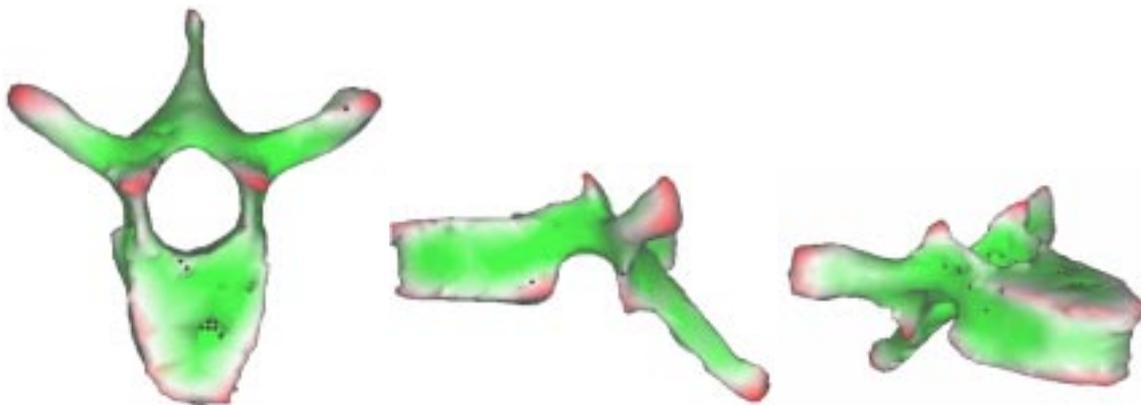


Abbildung 6.7: Ergebnis der Glättung des Objektes aus Abb. 6.5: Die Vertices in den rot eingefärbten Regionen wurden durch die Glättung stärker als der Durchschnitt aller Vertices verändert und werden daher als potentielle anatomische Landmarken identifiziert.

die sich automatisch detektieren lassen (Abb. 6.8 (links)). Bei einem proximalen Femur-Modell hingegen liefert das automatische Verfahren zunächst lediglich sehr wenige Landmarken. Es werden nur die beiden Rollhügel (*Trochanter minor/major*) erkannt (Abb. 6.8 (Mitte)), was i.a. nicht für das anschließende Ausrichten genügt. Bei einem distalen Femurende sieht es noch schlechter aus. Diese Knochenregion besitzt zwar sehr wohl anatomisch markante Teilgebiete (zwei Kniegelenkswalzen (*condylus lateralis/medialis*), zwei Walzenhöcker (*epicondylus lateralis/medialis*), den zwischen den Kondylen verlaufenden Gelenkgraben (*fossa intercondylaris*) sowie die patellare Lauffläche, vgl. Anhang A), aber diese sind streng genommen nicht punktbasiert und werden daher auch nicht stabil erkannt (Abb. 6.8 (rechts)).

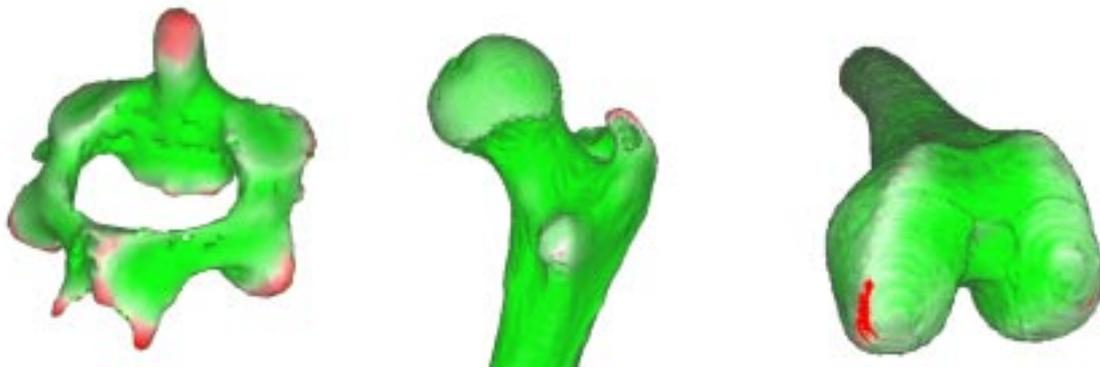


Abbildung 6.8: Die Leistungsfähigkeit des vorgestellten automatischen Verfahrens zur Detektion anatomischer Landmarken hängt sehr stark davon ab, inwieweit die Landmarken selbst lokal begrenzte Extrempunkte darstellen; je nachdem werden ausreichend viele (Halswirbel C2, links), nur wenige (proximaler Femur, Mitte) oder fast gar keine (distaler Femur, rechts) Marken detektiert.

Zusammenfassend kann gesagt werden, daß das vorgestellte Verfahren in der Lage ist, für viele Objektmodelle automatisch eine Landmarken-Menge handhabbarer Größe zu generieren. Je nach Ausprägung der Oberflächenstrukturen einzelner Objekte (vgl. Bsp. distaler Femur) ist es dabei jedoch mitunter nötig, das Ergebnis der automatischen Detektion anschließend manuell nachzubearbeiten. Drei mögliche Situationen und jeweils zugehörige Nachbearbeitungsschritte sind denkbar:

- An einigen Objekten wurden sporadisch Landmarken detektiert, die kein stabiles Merkmal repräsentieren und daher letztlich auch keine Entsprechung auf den Oberflächen der restlichen Objekte besitzen. Diese Landmarken werden entfernt.
- Ein wichtiges anatomisches Merkmal wurde an nahezu keinem der Objekte detektiert. In diesem Fall ist es nötig, an jedem der Objekte eine Landmarke für dieses Merkmal zu definieren. Ansonsten geht dieses Detail im Rahmen der weiteren Modellierung verloren.
- Ein wichtiges anatomisches Merkmal wurde an vereinzelten Trainingsobjekten nicht detektiert. Diese fehlenden Landmarken müssen an den betroffenen Objekten ebenfalls ergänzt werden.

6.4.4.2 Automatische Korrespondenzfindung (Labelling)

Nach der Bestimmung der Landmarken ist als nächster Schritt die Etablierung von Korrespondenzen zwischen Landmarken unterschiedlicher Trainingsobjekte nötig. Da die Landmarken auf der Oberfläche der

einzelnen Objekte i.a. sehr weit auseinanderliegen, genügt für die Korrespondenzbestimmung zwischen Landmarken oftmals eine einfache *nearest-neighbour*-Zuordnung. Voraussetzung hierfür ist allerdings, daß die einzelnen Objekte zuvor bereits mittels rigider Transformationen gut zur Deckung gebracht wurden.

Landmarken-basierte Koregistrierung

Ziel der Koregistrierung ist es, die zunächst *ungeordneten* Landmarken-Mengen von Trainingsobjekten durch Rotation und Translation der Objekte möglichst gut gegeneinander auszurichten, d.h. zur Deckung zu bringen.

Bzgl. zweier Objekte A und B mit zugehörigen Mengen an Landmarken – $L^A = \{L_1^A, L_2^A, \dots\}$ bzw. $L^B = \{L_1^B, L_2^B, \dots\}$ – läßt sich das gesetzte Ziel in Form des folgenden, zu minimierenden, Fehlermaßes E_{AB} zum Ausdruck bringen:

$$E_{AB}(\mathbf{O}, \mathbf{t}) = \sum_{L_i^B \in L^B} \min_{L_j^A \in L^A} \left\| (\mathbf{O} \bar{L}_i^B + \mathbf{t}) - \bar{L}_j^A \right\|_2 \quad (6.11)$$

Dabei entsprechen \bar{L}_i^A bzw. \bar{L}_i^B – der Definition 2 (Seite 101) folgend – den jeweiligen Schwerpunkten der zugehörigen Landmarken. Die beiden Parameter der Fehlerfunktion (\mathbf{O}, \mathbf{t}) mit je drei Freiheitsgraden stellen die gesuchte Transformation des Objektes B relativ zu A dar. Das Fehlermaß in Gl. 6.11 läßt sich z.B. durch Verwendung des Levenberg-Marquardt-Gradientenabstiegsverfahrens [Mor80] minimieren. Die damit i.a. verbundene Problematik lokaler Minima der Fehlerfunktion erweist sich in der vorliegenden Situation erfreulicherweise als unkritisch. Das liegt daran, daß die einzelnen Trainingsmodelle sämtlich aus dreidimensionalen Tomographiedaten gewonnen wurden. Dabei wird der Patient in der Regel einheitlich auf dem Tomographietisch positioniert, so daß auch einzelne Knochen jeweils in etwa gleich ausgerichtet sind. Zu beachten sind dabei lediglich vier unterschiedliche Konfigurationen, je nachdem ob der Patient auf dem Rücken oder dem Bauch liegend sowie kopf- oder fußwärts durch den Tomographen bewegt wurde (vgl. Abschnitt 4.2.1). Sollten daher einige der Trainingsobjekte zunächst in unterschiedlicher Konfiguration vorliegen, so lassen sie sich leicht durch einfache 180° -Drehungen gegeneinander ausrichten, ehe die Minimierung gestartet wird.

Gl. 6.11 muß tatsächlich nur für den Fall, daß initial keine Korrespondenzen bekannt sind, durch grobe Ausrichtung der Objekte mit anschließendem Gradientenabstiegsverfahren gelöst werden. Bereits bei der Kenntnis von Korrespondenzen für mindestens drei der auf allen Trainingsobjekten vorhandenen Landmarken läßt sich die Gleichung eindeutig und geschlossen lösen. In diesem Fall reduziert sich die Problemstellung auf die bekannte Fragestellung nach einer relativen 3D-3D-Orientierung [HS93a]. Dabei gilt es allerdings zu beachten, daß hiermit lediglich die Landmarken, für die die Korrespondenzen gegeben sind, gut miteinander ausgerichtet werden. Da es sich bei den Trainingsobjekten gezielt um Modelle unterschiedlicher Form handelt, resultiert die geschlossene Lösung von Gl. 6.11 i.a. in größeren Abweichungen für Landmarken ohne Korrespondenzen. Das Gradientenabstiegsverfahren hingegen verhindert dies.

Konsistentes Landmarken-Labeling

Nachdem alle Trainingsobjekte bzgl. ihrer Landmarken koregistriert wurden, lassen sich alle Korrespondenzen zwischen Landmarken auf unterschiedlichen Objekten bestimmen. Wie bereits angesprochen, läßt sich dieser Vorgang nun größtenteils durch eine bijektive *nearest-neighbour*-Zuordnung automatisieren. Für je zwei Landmarken L_i^A und L_j^B auf den Objekten A bzw. B sind hierfür die folgenden Kriterien zu überprüfen:

- L_i^A ist die zu L_j^B nächstgelegene Landmarke auf Objekt A .
- L_j^B ist die zu L_i^A nächstgelegene Landmarke auf Objekt B .
- Für L_i^A und L_j^B gibt es noch keine Korrespondenzen zu anderen Landmarken.

Sind alle drei Kriterien erfüllt, so wird die entsprechende Korrespondenz (L_i^A, L_j^B) gebildet. Eine wichtige Eigenschaft von Korrespondenzen ist deren Transitivität. Bestehen bei Betrachtung dreier Objekte A , B und C die beiden Korrespondenzen (L_i^A, L_j^B) und (L_i^A, L_k^C) , so folgt daraus direkt die Korrespondenz (L_j^B, L_k^C) . Unter Ausnützung dieser Eigenschaft genügt es daher, für die Landmarken eines ausgezeichneten Objektes Korrespondenzen zu allen restlichen Trainingsobjekten zu definieren. Daraus ergeben sich anschließend direkt alle restlichen noch fehlenden Korrespondenzen zwischen Landmarken auf allen Objekten (*transitive* Übertragung von Korrespondenzen).

6.4.5 Ausrichtung der Modelle

Strenggenommen sind die Trainingsobjekte bereits nach der Anwendung des im letzten Abschnitt beschriebenen Verfahrens gegenseitig ausgerichtet. Dabei wurde allerdings jedes Objekt auf seine wenigen anatomischen Landmarken reduziert. Zwar wäre prinzipiell bereits auf Basis dieser Landmarken eine statistische Formanalyse nach den Gleichungen 6.7 - 6.9 möglich, aber die Erfassung der Variation dieser wenigen Punkte (oftmals < 10) reicht noch in keinsten Weise aus, um die Variation der gesamten Objekt Oberfläche (oftmals mehrere 1000 Vertices) zu beschreiben.

Daher ist es nötig, eine Vielzahl an zusätzlichen vertexbasierten Korrespondenzen zwischen den Trainingsobjekten zu bestimmen. Dieser Vorgang muß automatisch erfolgen, da für die betroffenen Vertices zum einen deren große Anzahl und zum anderen die fehlenden anatomischen Anhaltspunkte einen manuellen, interaktiven Eingriff von vornherein ausschließen.

Die Bestimmung der neuen Korrespondenzen erfordert dabei zunächst eine erneute gegenseitige Ausrichtung aller Trainingsobjekte – dieses Mal unter Einbeziehung aller Vertices. Dabei wird sich insbesondere zeigen, daß erst durch die bereits etablierten Korrespondenzen zwischen markanten Landmarken, eine korrekte vertexbasierte Korrespondenzfindung möglich wird. Dies gilt umso mehr, je stärker sich zwei Trainingsobjekte in ihrer Form voneinander unterscheiden. Die vertexbasierte Ausrichtung beinhaltet die nachfolgend aufgeführten Teilschritte:

1. Auswahl eines Referenzobjektes:

Um sicherzustellen, daß sich trotz der i.a. unterschiedlichen Vertexanzahl der Trainingsobjekte jeweils die gleiche Anzahl an korrespondierenden Punkten ergibt, wird zunächst eines der Modelle als Referenzobjekt ausgewählt. Vertexanzahl und Topologie (Kanten- und Flächenverläufe) dieses Referenzobjektes M_{ref} sind damit für alle restlichen Trainingsobjekte im Hinblick auf die nachfolgenden Verarbeitungsschritte maßgebend. Das bedeutet insbesondere, daß im folgenden für jeden Vertex von M_{ref} auf allen anderen Objekten der jeweils korrespondierende Oberflächenpunkt

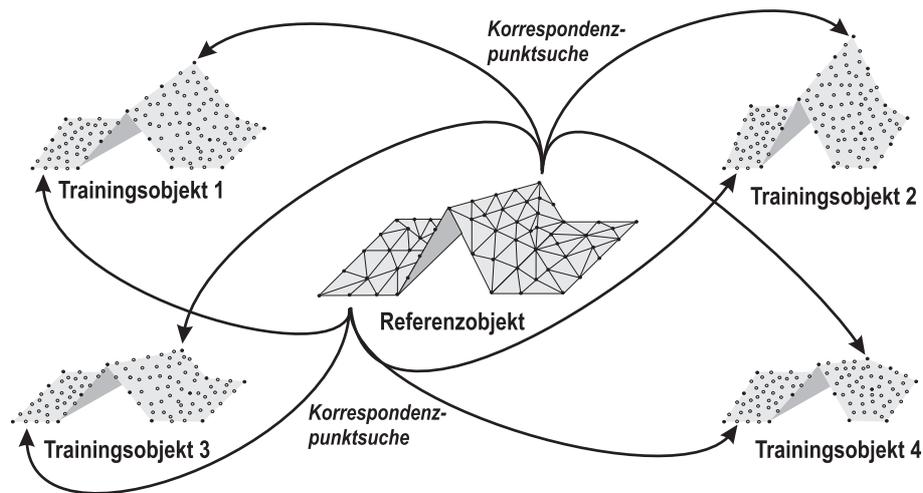


Abbildung 6.9: Das ausgezeichnete Referenzobjekt ist hinsichtlich Vertexanzahl und Topologie (Kanten und Flächen) maßgebend. D.h. für alle Referenzobjektvertices wird ein korrespondierender Oberflächenpunkt auf jedem der Trainingsobjekte gesucht.

gesucht wird. Dabei muß es sich nicht um bereits existierende Vertices handeln. Prinzipiell werden den Referenzvertices beliebige Punkte auf der Oberfläche des betrachteten Trainingsobjektes zugeordnet².

2. Rigide Ausrichtung:

Jedes der Trainingsobjekte wird derart bzgl. Orientierung, Translation und Skalierung auf das ausgewählte Referenzobjekt ausgerichtet, daß dabei ein vertexbasiertes Abstandsmaß minimal wird. Für eine Korrespondenzbildung durch einfache *nearest-neighbour*-Zuordnung ist diese rigide Ausrichtung aber in den wenigsten Fällen ausreichend. Insbesondere markante – in ihrer Form stark variierende – Merkmale (vgl. Landmarken, z.B. die Dorn- und Querfortsätze von Wirbeln) können oftmals durch rigide Ausrichtung nicht genügend zur Deckung gebracht werden.

3. Elastische Ausrichtung:

Eine zusätzliche elastische Ausrichtung ermöglicht, daß auch stark variierende Merkmale zur Deckung gebracht werden. Dabei ist es entscheidend, daß bereits einige Korrespondenzen zwischen Landmarken vorhanden sind (vgl. vorangegangenen Abschnitt). Diese Korrespondenzen können genutzt werden, um für jedes Paar korrespondierender Landmarken die Abschätzung einer lokalen Verformung zu berechnen. Diese bringt die Oberflächen von Referenz- und Trainingsobjekt dicht zueinander, so daß anschließend eine einfache *nearest-neighbour*-Zuordnung möglich wird. Die lokale Verformung dient dabei ausschließlich der Gewährleistung optimaler Korrespondenzen, hat ansonsten aber keinen Einfluß auf die anschließende statistische Variationsanalyse. Diese basiert einzig auf den unverformten Vertex-Koordinaten nach Abschluß der rigiden Ausrichtung.

²Ohne diese Eigenschaft käme nur das Trainingsobjekt mit den wenigsten Vertices als Referenzobjekt in Frage, da sich ansonsten für mindestens ein Trainingsobjekt nicht genügend Korrespondenzen bilden lassen.

Die beiden nachfolgenden Abschnitte setzen sich ausführlich mit den Teilschritten rigider und elastischer Ausrichtung auseinander.

6.4.5.1 Rigide Ausrichtung

Die starre Ausrichtung der Trainingsmodelle bzgl. eines ausgewählten Referenzobjektes dient zwei Zielen. Zum einen wird die Korrespondenzfindung zwischen Vertices auf beiden Objekten erleichtert, zum anderen stellt die Ausrichtung sicher, daß während der statistischen Analyse tatsächlich nur Formvariationen und keine globalen Translationen oder Rotationen modelliert werden. Neben globaler Verschiebung und Drehung ist auch die Modellierung gleichmäßiger globaler Größenunterschiede unerwünscht. Deshalb wird vor der eigentlichen Ausrichtung die Größe aller Trainingsobjekte normiert. Dies geschieht durch entsprechende Skalierungsfaktoren, so daß der durchschnittliche Abstand aller Vertices eines Objektes von dessen Schwerpunkt jeweils 1.0 beträgt.

Es bieten sich prinzipiell zwei mögliche Vorgehensweisen für das Ausrichten der Objekte an, basierend auf wenigen Landmarken mit bekannten Korrespondenzen oder unter Verwendung aller Vertices ohne explizite Korrespondenzbeziehungen:

- Wie bereits in Abschnitt 6.4.4.2 angedeutet wurde, ist die Ausrichtung geschlossen lösbar, sofern dabei ausschließlich die bereits bekannten Korrespondenzen zwischen Landmarken herangezogen werden. Dies führt allerdings im Hinblick auf die restlichen Oberflächenbereiche oftmals zu massiven Problemen, indem sich an diesen Stellen starke unerwünschte Abweichungen der Objekte untereinander ergeben. Ursache hierfür ist, daß das verwandte Fehlermaß auf alleiniger Berücksichtigung der Landmarken basiert. Dies gilt insbesondere dann, wenn die Formen der einzelnen Beispielobjekte starke *lokale* Variationen aufweisen, wie z.B. die beiden Trainingsobjekte aus Abb. 6.10: ein T6- (links) und ein T11-Brustwirbel (rechts). Bei Vergleich der beiden Wirbel fällt neben der schmäleren Körperform des T6-Wirbels besonders die stark unterschiedliche Neigung des Dornfortsatzes auf. Wird nun die rigide Ausrichtung auf z.B. fünf bis sechs Landmarken, die u.a. auch den Dornfortsatz mit beinhalten, beschränkt, so werden die nicht berücksichtigten Oberflächenstücke, in erster Linie die massiven Wirbelkörper, aufgrund der großen Neigung des Fortsatzes extrem schlecht zur Deckung gebracht (Abb. 6.11).

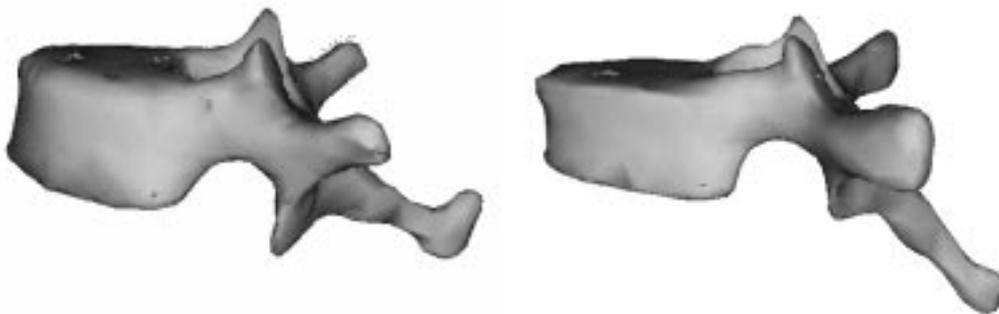


Abbildung 6.10: Brustwirbel-Modelle T11 (links), T6 (rechts)

Abgesehen von der damit verbundenen Korrespondenz-Problematik ist eine solche Ausrichtung

auch im Hinblick auf die statistische Auswertung der Verformung unerwünscht. Intuitiv würde man eher exakt das Gegenteil erwarten: daß die beiden massiven Wirbelkörper als stabiles Merkmal zur Deckung gebracht werden und dafür die Neigung und Form des Dornfortsatzes als Variation modelliert wird. Daher ist es in jedem Fall besser, den Großteil der Vertices zur Deckung zu bringen und dabei zunächst vereinzelte lokale Abweichungen in Kauf zu nehmen. Diese lokalen Abweichungen lassen sich zum Zwecke der Korrespondenzfindung durch die im nächsten Abschnitt beschriebene elastische Ausrichtung in den Griff bekommen. Im Sinne der Formmodellierung sind sie erwünscht und bleiben daher erhalten.

- Die Berücksichtigung aller Vertices für die Ausrichtung führt zu einer geänderten Problemformulierung, da in diesem Fall keine Korrespondenzen bekannt sind. Diese werden im Rahmen eines iterativen Verfahrens während jeder Iteration neu ermittelt (vgl. ICP-Algorithmus, *Iterative Closest Point*, [BM92]). Da hierbei wiederholt für jeden Vertex eines Objektmodells, dessen Abstand zu allen Dreiecksflächen des Referenzobjektes berechnet werden muß, ist aus Gründen der Effizienz eine zeitliche Optimierung der Abstandsberechnung unumgänglich. Hierzu wird die Vertexstruktur des Referenzobjektes vor den einzelnen rigiden Ausrichtungen einmalig gleichmäßig dicht mit neuen Vertices aufgefüllt. Anschließend wird für dieses Objekt, [Lav96a] folgend, eine sogenannte *Octree*-Datenstruktur gebildet. Diese ermöglicht die effiziente Berechnung des euklidischen Abstands für beliebige Punkte zu der Oberfläche des dichten Referenzobjektes. Die zeitaufwendige Berechnung der Datenstruktur muß dabei nur einmalig erfolgen. Abb. 6.12 zeigt das Ergebnis einer vertexbasierten Ausrichtung dreier unterschiedlich geformter Trainingsobjekte.

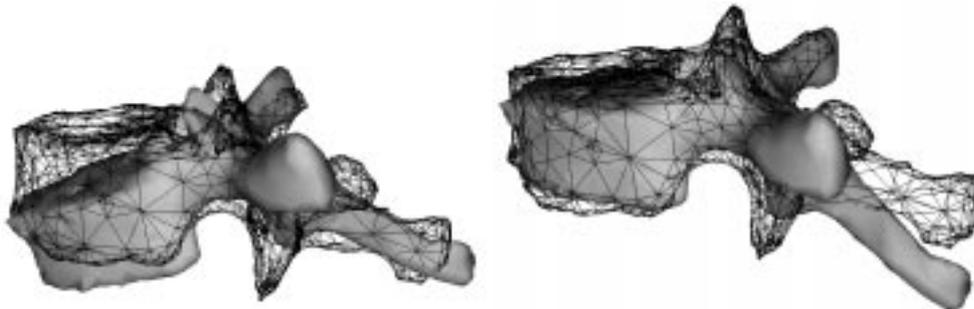


Abbildung 6.11: Eine Ausrichtung unter alleiniger Berücksichtigung weniger Landmarken führt oftmals zu unerwünschten Ergebnissen (links; als Landmarken dienen der Dornfortsatz sowie die beiden Quer- und artikulare Fortsätze). Eine brauchbare Ausrichtung hingegen ergibt sich durch die ungewichtete Berücksichtigung aller Objektvertices (rechts). Hier wird insbesondere der durch sehr viele Vertices repräsentierte Wirbelkörper gut zur Deckung gebracht – im Gegensatz zu den Dornfortsätzen.

6.4.5.2 Elastisches Alignment (Abschätzung einer Deformation)

Durch das im vorigen Abschnitt beschriebene Verfahren wird bereits ein rigides Alignment auf der Basis weniger gemeinsamer Landmarken bzw. alternativ unter Einbeziehung aller Vertices je zweier Objekte erreicht. Die nach dieser Ausrichtung zurückbleibenden lokalen Unterschiede der Trainingsobjekte (vgl. Abb. 6.12 u. 6.11) sind dabei Ursache des folgenden Dilemmas:

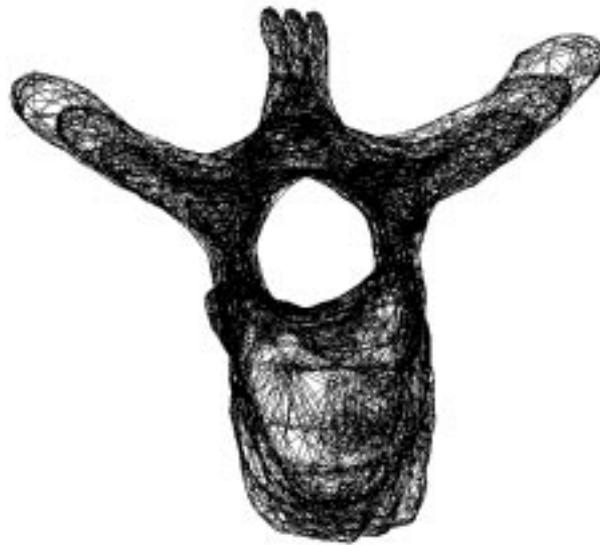


Abbildung 6.12: Drei überlagerte Trainingsobjekte nach erfolgter rigider Ausrichtung (Alignment) bzgl. aller Vertices. Dabei ist deutlich zu erkennen, daß die Position der Dorn- und Querfortsätze noch stark variiert, so daß eine *nearest-neighbour*-Zuordnung von Vertices an diesen Stellen ohne vorherige elastische Ausrichtung zum Scheitern verurteilt ist.

Auf der einen Seite ist eine starke lokale Variation innerhalb der Trainingsobjekte durchaus erwünscht, da das Ziel der statistischen Formanalyse gerade darin besteht, signifikante Formvariationen zu modellieren. Umso mehr repräsentative – d.h. in ihrer Form variierende – Objekte die Trainingsmenge enthält, desto besser fällt letztlich die Qualität der Modellierung aus. Auf der anderen Seite aber verhindern gerade diese starken Variationen – insbesondere bei signifikant konkaven Oberflächen – eine einfache Berechnung der für die statistische Analyse benötigten bijektiven '1:1'-Abbildung zwischen den Oberflächen zweier Trainingsobjekte.

Die sicherlich bekannteste anwendungsunabhängige Technik zur Bestimmung von Korrespondenzen ist eine einfache *nearest-neighbour*-Zuordnung. Dieses Verfahren hat bei komplex geformten Oberflächen aber nur dann Aussicht auf Erfolg, wenn alle einander entsprechenden Oberflächenstücke auf zwei Modellen bereits vor der Zuordnung sehr nahe beieinander liegen. Eine rigide Ausrichtung alleine reicht hierfür im vorliegenden Anwendungsfall nicht aus. Zwar werden große Teile der Oberflächen meist bereits sehr gut zur Deckung gebracht (im Beispiel insbesondere die massiven Wirbelkörper, Abb. 6.12), jedoch auf Kosten der erwähnten lokalen Detailunterschiede, die im Gegensatz hierzu eher hervorgehoben werden (z.B. die Dornfortsätze).

Um auch für diese Detailunterschiede eine einfache Korrespondenzbestimmung zu ermöglichen, werden ausgehend von der rigiden Ausrichtung sukzessive alle Oberflächendetails des Referenzobjekts, die noch nicht ausreichend mit dem betrachteten Trainingsobjekt zur Deckung kommen, durch die Berechnung einer temporären, lokal begrenzten Verformung des Referenzobjektes angenähert. Diese temporäre Deformation kann als eine Art elastische Ausrichtung aufgefaßt werden. Dabei wird für jede der bereits bekannten Landmarken-Korrespondenzen (vgl. Abschnitt 6.4.4.1) eine lokale Deformation des Referenzobjektes in der Umgebung dieser Landmarke abgeschätzt. Dabei gilt es zu beachten, daß

die temporäre Deformation des Referenzobjektes ausschließlich der Korrespondenzfindung dient. In die spätere Berechnung der statistischen Kenngrößen (Abschnitt 6.4.2) entsprechend Gl. 6.7 fließen hingegen die unveränderten Vertex-Koordinaten nach Abschluß der rigiden Ausrichtung ein.

Um die herausragende Leistungsfähigkeit dieses Ansatzes zu demonstrieren, wird im folgenden ein Trainingsdatensatz verwendet, dessen Objekte bewußt sehr starke lokale Unterschiede aufweisen. Der Datensatz beinhaltet u.a. einen Lenden- (*vertebra lumbaris*, L2) und einen Brustwirbel (*vertebra thoracica*, T11). Dabei gilt es zu beachten, daß es hinsichtlich eines Praxiseinsatzes der vorgestellten elastischen Modellierung im Gegensatz hierzu i.a. sinnvoll ist, das Variationstraining auf einen bestimmten vorgegebenen Knochentypus zu konzentrieren, wobei die Trainingsobjekte aus den Bilddatensätzen einer Vielzahl unterschiedlicher Patienten stammen. Das resultierende Modell erfaßt in diesem Fall die signifikantesten der zu erwartenden Formunterschiede bei Betrachtung mehrerer Patienten und nicht die anatomischen Formunterschiede ähnlicher Knochentypen (z.B. benachbarter Wirbel).

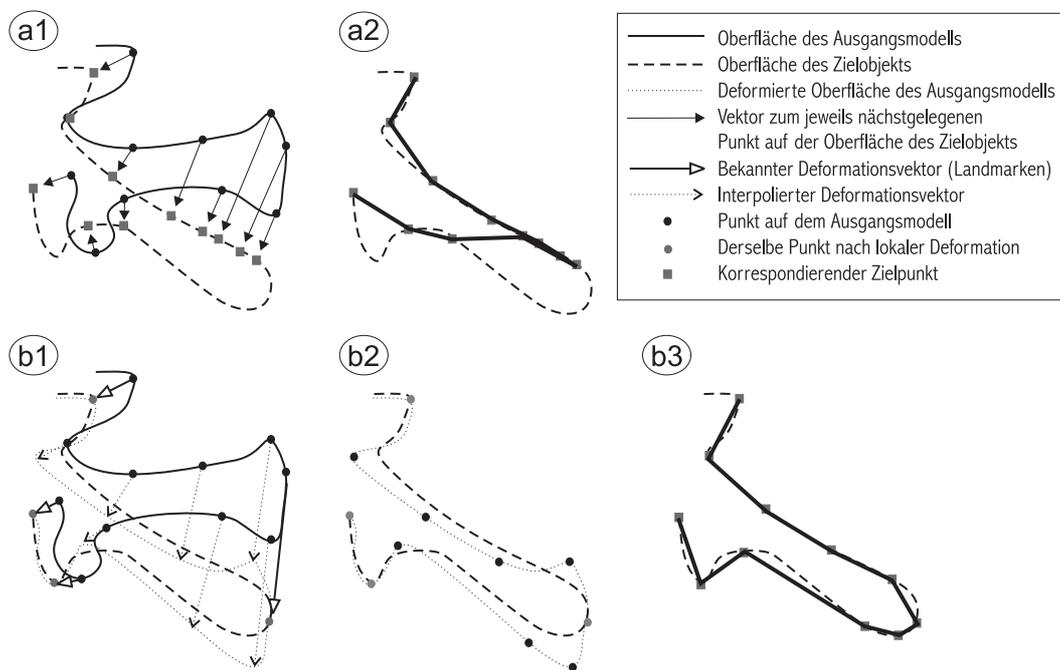


Abbildung 6.13: Naive Bestimmung korrespondierender Punkte auf der Oberfläche des Zielobjekts (a1), fehlerhaftes Resultat (a2); Lokale, interpolierte Verformung des Ausgangsmodells auf der Basis weniger bekannter Landmarken-Korrespondenzen (b1), Ergebnis der lokalen Verformung: die beiden Objekte stimmen nun deutlich besser überein (b2), korrektes Resultat nach Bestimmung korrespondierender Punkte (b3).

Die elastische Ausrichtung beinhaltet, ausgehend von der bereits erfolgten rigiden Ausrichtung, die folgenden beiden Teilschritte:

1. Berechnung der Deformationsvektoren für bekannte Landmarken-Korrespondenzen:
Für jedes Paar korrespondierender Landmarken ($L_i^{M_{ref}}$, $L_j^{B'}$) läßt sich auf einfache Weise ein Verschiebungsvektor berechnen, der den Schwerpunkt $\bar{L}_i^{M_{ref}}$ der Referenzobjekt-Landmarke in

den zugehörigen Schwerpunkt $\bar{L}_j^{B'}$ der Landmarke des gerade betrachteten Trainingsobjektes verschiebt:

$$\vec{d}_i = \bar{L}_j^{B'} - \bar{L}_i^{M_{ref}} \quad (6.12)$$

Die Anwendung dieses Vektors auf alle Landmarken-Vertices verschiebt die gesamte Landmarke in die Nähe ihrer zugehörigen Landmarke des Trainingsobjektes. Direkt benachbarte und auch alle restlichen Vertices, die selbst keiner Landmarke angehören, bleiben dabei aber zunächst unberücksichtigt. Für diese Vertices wird im nachfolgenden Schritt eine gleichmäßige Verformung abgeschätzt.

2. Abschätzung der restlichen Verschiebungsvektoren:

Durch den vorherigen Schritt sind die Teile des Deformationsvektors bekannt, welche die Landmarken betreffen. Für alle restlichen Vertices können die Deformationsvektoren nun aus den bekannten Vektoren interpoliert werden [Dö99].

Abb. 6.14 veranschaulicht die lokale Wirkung einer elastischen Ausrichtung anhand des Dornfortsatzes zweier Wirbelmodelle.

6.4.6 Vollständige Bestimmung von Korrespondenzen

Für die statistische Formanalyse ist es entscheidend, daß jeder Trainingsdatensatz gleichviele Vertices, die jeweils miteinander korrespondieren, enthält. Um dies zu gewährleisten, dient das ausgewählte Referenzobjekt als Bezugsobjekt, indem dessen Vertex-Anzahl und Topologie (Kanten- und Flächendefinitionen auf Basis der Vertices) die Anzahl modellierter Vertices einerseits sowie die Topologie des formvariablen Modells andererseits einmalig konstant festlegen. Für alle restlichen Trainingsobjekte sind daraufhin korrespondierende Oberflächenpunkte zu bestimmen, die i.a. nicht mit bereits vorhandenen Vertices zusammenfallen. Die tatsächliche Vertexanzahl der einzelnen Trainingsobjekte spielt dabei ebensowenig eine Rolle wie deren Topologie.

Die vorgestellten Ansätze zur rigiden und elastischen Ausrichtung ermöglichen es, das ausgewählte Referenzobjekt sukzessive mit jedem der restlichen Trainingsobjekte zur Deckung zu bringen, ungeachtet der ursprünglich vorhandenen lokalen Formunterschiede.

Für jeden Vertex des verformten Referenzmodells M_{ref}^l wird anschließend ein korrespondierender Oberflächenpunkt auf dem gerade betrachteten Trainingsobjekt M_i gesucht (vgl. Abb. 6.9), der den minimalen Abstand zu diesem Vertex aufweist. Die Auswahl des Punktes beschränkt sich dabei insbesondere nicht auf die Menge der Vertices von M_i . Prinzipiell kommt jeder Punkt auf der Oberfläche von M_i in Frage. Durch Berücksichtigung der Richtungen der jeweils zugehörigen Vertexnormalen kann verhindert werden, daß falsche Zuordnungen zwischen Punkten generiert werden, die auf stark unterschiedlich ausgerichteten Oberflächenstücken liegen. Ein sinnvolles Kriterium hierbei ist die Forderung nach einem Zwischenwinkel von weniger als 90° .

Sobald für jedes Trainingsobjekt alle gesuchten Vertex-Korrespondenzen bestimmt und alle zugehörigen Vertex-Koordinaten jeweils in Form eines Vektors \mathbf{x}_i zusammengefaßt wurden, läßt sich daraus eine Durchschnittsform berechnen. Dieses Modell ergibt sich durch arithmetische Mittelung aller Vertex-Koordinaten mit gleichem Index über alle m rigide ausgerichteten Trainingsformen inklusive

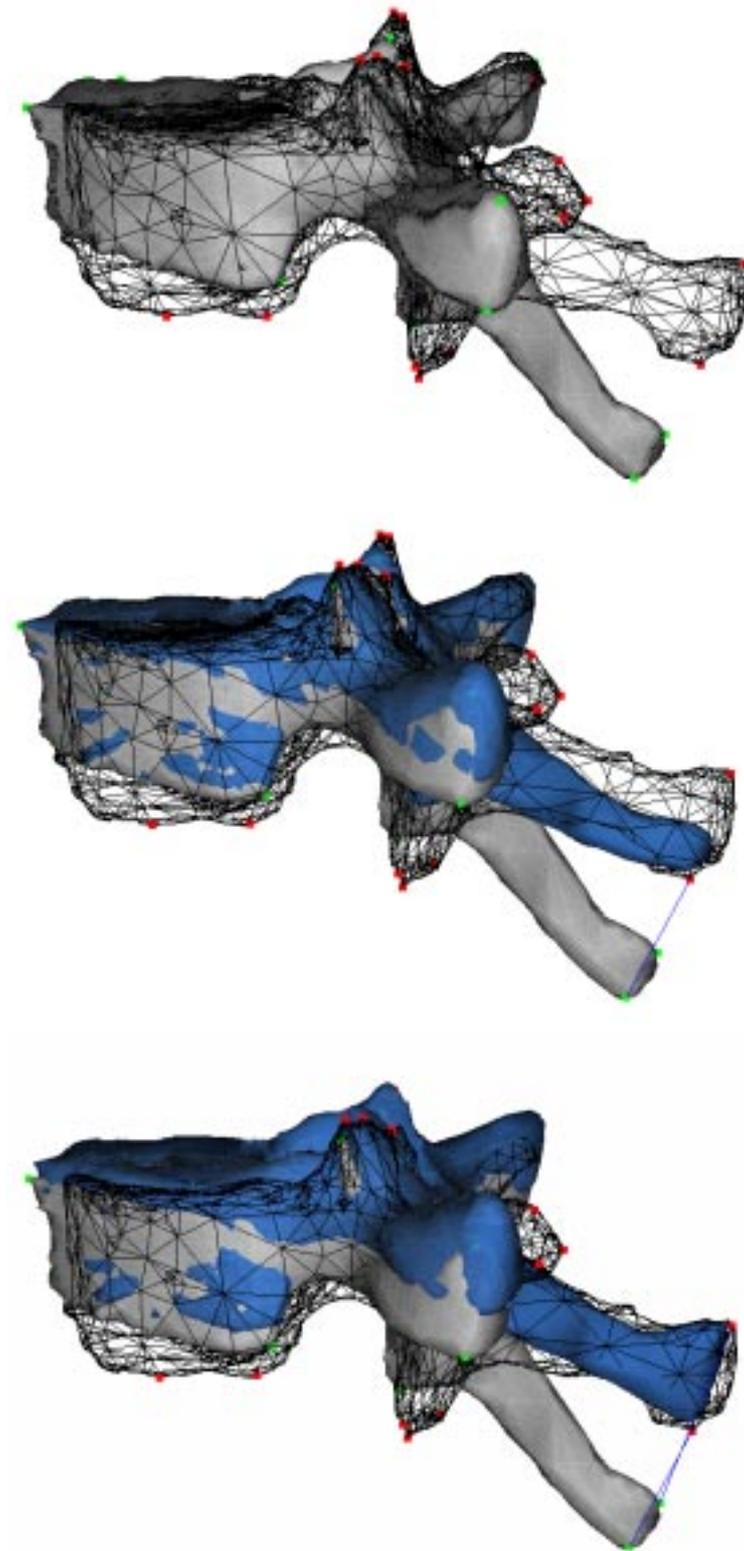


Abbildung 6.14: Elastische Ausrichtung des Dornfortsatzes: keine Korrespondenz (oben), eine Korrespondenz (Mitte) und zwei Korrespondenzen (unten). Das elastisch verformte Modell ist blau gezeichnet.

Referenzobjekt:

$$\bar{\mathbf{x}} = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i \quad (6.13)$$

Die Auswahl des Referenzobjektes aus dem Kreis aller vorhandenen Trainingsobjekte erscheint ein wenig willkürlich. Daher ist es sinnvoll, die gesamte beschriebene Prozedur zur Bestimmung der Trainingsformen \mathbf{x}_i , bestehend aus rigidem Ausrichten (Abschnitt 6.4.5.1), elastischem Ausrichten (Abschnitt 6.4.5.2) und vollständiger Korrespondenzbestimmung (Abschnitt 6.4.6) erneut durchzuführen. Als neues Referenzobjekt dient dann das soeben berechnete Durchschnittsmodell. Diese Neuausrichtung aller Objekte an einem Mittelwertobjekt kann solange wiederholt werden bis sich schließlich ein stabiles Durchschnittsmodell (Konvergenz) ergibt (vgl. [Coo95b], 2D-Situation).

6.5 Konturbasierte, elastische Lokalisation

Die Ausgangsbasis einer elastischen Registrierung ist zunächst dieselbe wie bereits in Abschnitt 5.3 (Intensitätsbasierte Registrierung): eine statisch gegebene Szene $[(\mathcal{I}_{x_1}, Cam_1), \dots, (\mathcal{I}_{x_n}, Cam_n)]$, bestehend aus n verzerrungsfreien Fluoroskopieaufnahmen \mathcal{I}_{x_i} sowie jeweils korrespondierenden Kameraparametern Cam_i , die sich aus der Kalibrierung und Lokalisation des Röntgenbildverstärkers ergeben. Diese Aufnahmen werden intraoperativ nacheinander aus verschiedenen Richtungen gewonnen – wobei sich der Patient zwischenzeitlich nicht bewegen darf – und zeigen jeweils die betrachtete Knochenstruktur.

Im Gegensatz zu der intensitätsbasierten Registrierung finden im vorliegenden Fall jedoch nicht die akquirierten Bilddaten selbst Verwendung während der Registrierung. Stattdessen erfolgt initial eine Kontursegmentierung in jeder der Fluoroskopieaufnahmen, um auf diese Weise die jeweils sichtbare Silhouette des betrachteten Knochens zu gewinnen (Abschnitt 6.5.1).

Die elastische Lokalisation zielt anschließend darauf ab, ein formvariables Knochenmodell – modelliert nach den zuletzt beschriebenen Verfahren – bzgl. seiner rigiden Lage (Position, Orientierung), seiner Größenskalierung sowie seiner variablen Form an die segmentierten Konturen anzupassen (Abschnitt 6.5.2).

6.5.1 Segmentation von Fluoroskopieaufnahmen

Die Detektion der für die elastische Registrierung benötigten Knochensilhouetten in jeder einzelnen Fluoroskopieaufnahme erfolgt in der vorliegenden Arbeit jeweils entsprechend eines der beiden folgenden alternativen 2D-Segmentationsverfahren. Enthalten die betrachteten Fluoroskopieaufnahmen Metallartefakte aufgrund von Kalibriermarken o.ä., so ist vor der Segmentierung eine Bildvorverarbeitung entsprechend des bereits in Abschnitt 5.2.3 besprochenen lokalen Glättungsverfahrens erforderlich.

- Das *Livewire/Intelligent Scissors*-Verfahren von Barrett und Mortensen [BM97] stellt ein flexibles interaktives Verfahren zur Kontursegmentierung dar. Es erlaubt insbesondere eine schnelle und präzise Segmentation der Bildkantenverläufe beliebiger Strukturen. Die Interaktion beinhaltet zunächst die Definition eines Startpunktes innerhalb des betrachteten Bildes. Anschließend wird bei jeder Mausbewegung ein optimaler Kurvenverlauf zwischen Startpunkt und aktueller Mausposition berechnet und angezeigt, wobei *optimal* bedeutet, daß die Kurve in ihrem Verlauf starke

Bildkanten bevorzugt. Führt der Benutzer die Maus daher in die nähere Umgebung einer Objektkante, so paßt sich der Kurvenverlauf unmittelbar an. Die Optimierung basiert darauf, das gesamte Bild als große Graphstruktur zu interpretieren, wobei jeder Bildpixel einen Knoten repräsentiert und über Kanten mit jedem seiner acht Nachbarpixel verbunden ist. Die Suche nach einem optimalen Kurvenverlauf entspricht dabei der Suche nach einem optimalen Pfad durch den Graph, wobei sich die mit den Graphkanten assoziierte Kostenfunktion auf die lokale Stärke des Gradienten abstützt. Ein hinlänglich bekanntes Verfahren zur Behandlung dieser Graphensuche ist der Dijkstra-Algorithmus [Dij59]. Die Folge von interaktiven Stützpunktfestlegungen und Mausbewegungen liefert schließlich einen geschlossenen Kurvenzug als Ergebnis der Segmentierung.

- Das zweite Verfahren basiert auf einem modellbasierten Ansatz [Tag97]. Das Modell bzw. *Template* besteht dabei aus insgesamt drei diskretisierten, sich nicht überschneidenden Kurven. Zwei der Kurven stellen eine innere bzw. äußere Begrenzungslinie dar, wohingegen die dritte Kurve zwischen den beiden anderen verläuft. Jede der drei Kurven enthält die gleiche Anzahl k an diskreten Kontrollpunkten. Senkrecht zu den drei bisherigen Kurven werden weitere n *orthogonale* Kurven definiert, die jeweils zwischen korrespondierenden Kontrollpunkten auf der inneren, mittleren und äußeren Begrenzungskurve verlaufen und ebenfalls eine diskrete Anzahl von Punkten enthalten. Ein solches *Template* ist in der Lage, alle Polygonzüge zu repräsentieren, die sich durch Auswahl je eines Punktes auf jeder der orthogonalen Kurven – d.h. zwischen den beiden Begrenzungskurven – ergeben können. Das Verfahren verhindert damit von vornherein Selbstüberschneidungen der resultierenden Kurven. Die Segmentierung erfolgt, indem ein solches *Template* zunächst geeignet orientiert und skaliert an einer bestimmten Stelle eines Bildes plaziert wird. Diese manuelle Initialisierung muß sicherstellen, daß die gesuchte Kontur des betrachteten Knochens vollständig innerhalb der beiden Begrenzungslinien zu liegen kommt. Anschließend wird eine auf *Dynamischer Programmierung* basierende Optimierung gestartet, die das Auffinden des besten Kurvenverlaufs innerhalb der Begrenzungen garantiert. Das Verhalten der Optimierung ist durch insgesamt drei Steuerparameter kontrollierbar. Diese gewichten wie stark das Verfahren auf den Grauwertgradienten des Bildes (Kanten), auf einen glatten Kurvenverlauf sowie auf eine Abweichung von der mittleren Kurve reagiert.

Abb. 6.15 zeigt die Ergebnisse der *Template*-Segmentierung eines Femurs (oben) sowie einer Tibia (unten). In beiden Bildern ist links jeweils das Template nach der Initialisierung zu sehen. Die rechten Aufnahmen hingegen zeigen das Ergebnis (weiße Punkte) im Vergleich zu einer interaktiv vorgenommenen *Intelligent-Scissors*-Segmentierung (durchgezogene schwarze Linien).

Unabhängig von den eingesetzten Verfahren resultiert jede Segmentierung einer Fluoroskopieaufnahme i letztlich in einer Liste S_i von s_i Konturpunkten:

$$S_i = [(x_1^s, y_1^s), \dots, (x_{s_i}^s, y_{s_i}^s)] \quad (6.14)$$

Diese Punkte repräsentieren die Silhouette des betrachteten Knochens in der jeweiligen Aufnahme und finden in der nachfolgend beschriebenen Optimierung Anwendung.

6.5.2 Optimierungskriterium

Die Anpassung eines gegebenen formvariablen Modells an segmentierte Knochenkonturen innerhalb mehrerer Fluoroskopieaufnahmen basiert in der vorliegenden Arbeit auf einer geeigneten Erweiterung

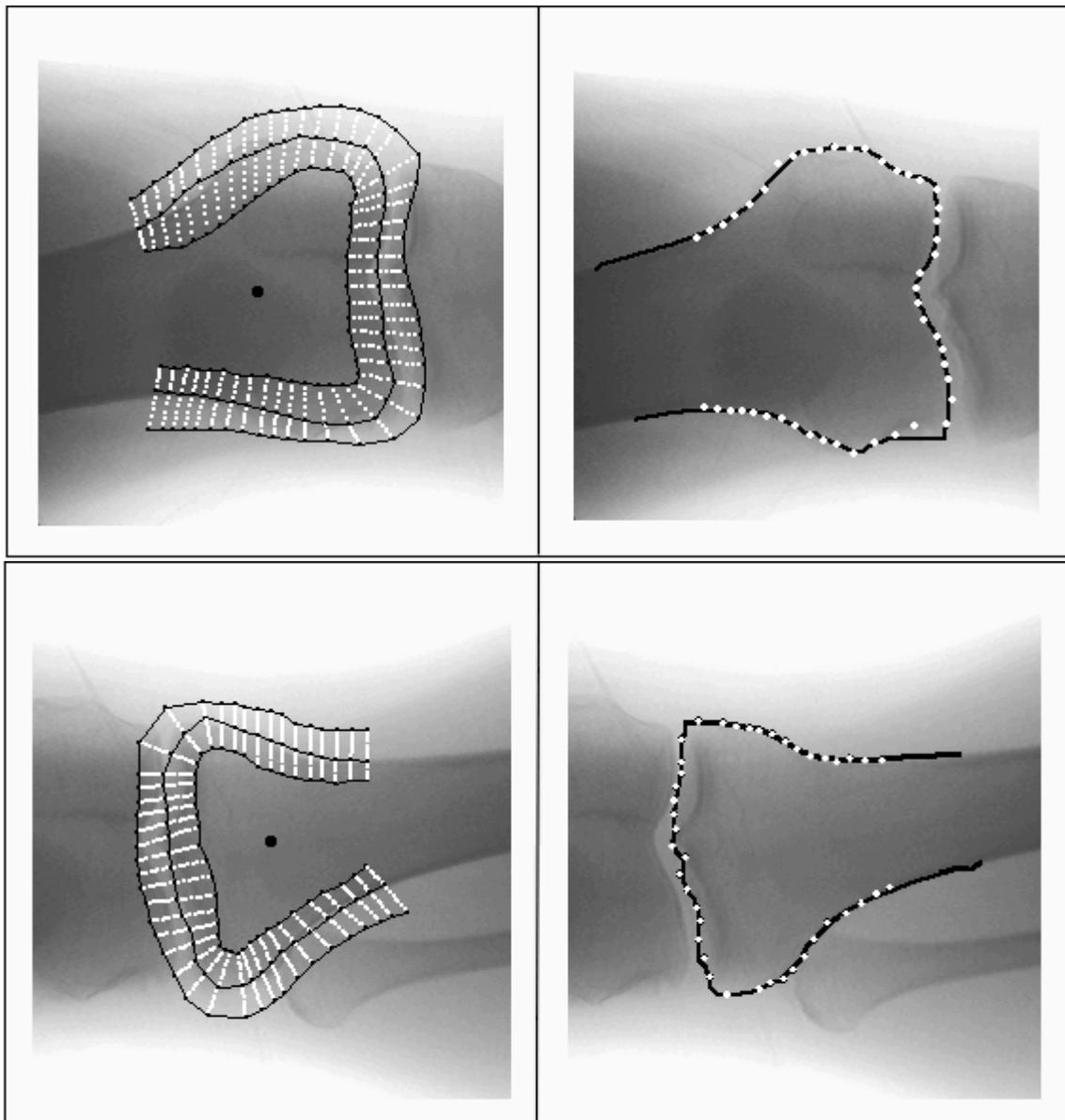


Abbildung 6.15: *Template*-basierte Segmentierung nach [Tag97]

des bekannten ICP-Verfahrens (*Iterative Closest Point* [BM92]). Das Ziel dieses Verfahren besteht in der iterativen, *rigiden* Registrierung zweier 3D-Punktmenge, wobei jede Iteration in zwei Phasen zerfällt: Zunächst werden Punkt-zu-Punkt-Korrespondenzen zwischen den beiden Punktmenge gesucht, anschließend wird die Summe der quadrierten Abstände zwischen korrespondierenden Punkten hinsichtlich der gesuchten *rigiden* Parameter minimiert. Damit bestehen zunächst zwei grundlegende Unterschiede im Vergleich zu der vorliegenden Situation:

1. Anstatt einer 3D-3D-Registrierung handelt es sich hier um eine 3D-2D-Problemstellung. Gesucht ist die gleichzeitige Anpassung einer 3D-Punktmenge – entsprechend der Vertices des formvariablen 3D-Modells – an mehrere 2D-Punktmenge, repräsentiert durch die segmentierten Konturen.
2. Die Verwendung eines formvariablen Modells erfordert eine *elastische Registrierung*, d.h. neben sechs *rigiden* Freiheitsgraden (Rotationsmatrix \mathbf{O} , Translationsvektor \mathbf{t}) sind weitere $n_b + 1$ modellbezogene Parameter zu optimieren: ein globaler Skalierungsfaktor s sowie n_b Deformationsparameter b_{n_1}, \dots, b_{n_b} .

Der in Punkt 1 beschriebene Unterschied läßt sich abschwächen, indem während einer Iteration statt der 2D-Konturpunkte selbst deren 3D-Projektionsgeraden herangezogen werden. Wie weiter unten gezeigt wird, erlaubt dies anschließend ebenfalls die Definition eines adäquaten 3D-3D-Abstandsmaßes zwischen den Projektionsgeraden und ausgewählten Modellvertices. Punkt 2 hat zur Folge, daß andere Optimierungsverfahren zur Minimierung dieser Abstände als in [BM92] eingesetzt werden müssen.

Für die nachfolgenden Überlegungen sei zunächst angenommen, daß in jeder Ansicht (Fluoroskopieaufnahme) bereits eine Menge an Modellvertices zusammen mit korrespondierenden Projektionsgeraden bekannt ist. Dann gilt es, einen optimalen Parametersatz $(\mathbf{O}, \mathbf{t}, s, b_1, \dots, b_{n_b})$ zu bestimmen, der das formvariable 3D-Modell hinsichtlich seiner Position \mathbf{t} , Orientierung \mathbf{O} , Skalierung s und Form (b_j) derart transformiert, daß dadurch jeder ausgewählte 3D-Vertex des Modells möglichst nahe an seine korrespondierende Projektionslinie herangebracht wird. Aus dieser Zielsetzung heraus läßt sich für jede Aufnahme i ein entsprechender Fehlerbeitrag E_i ableiten, der den durchschnittlichen Abstand eines Vertex von seiner zugeordneten Projektionsgeraden in Abhängigkeit von den gesuchten Parametern mißt:

$$E_i(\mathbf{O}, \mathbf{t}, s, b_1, b_2, \dots, b_{n_b}) = \frac{1}{m_i} \sum_{j=1}^{m_i} d(\mathbf{p}_j, (x_j^s, y_j^s)) \quad (6.15)$$

Dabei repräsentiert $d(\mathbf{p}_j, (x_j^s, y_j^s))$ den Abstand zwischen einem der insgesamt m_i ausgezeichneten Modellvertices \mathbf{p}_j und seiner korrespondierenden Projektionslinie, die durch den Konturpunkt (x_j^s, y_j^s) eindeutig bestimmt ist. Werden n Fluoroskopieaufnahmen aus unterschiedlichen Richtungen für die Registrierung herangezogen, so lassen sich die einzelnen Fehlerbeiträge E_i zu einem übergeordneten Gesamtfehlermaß zusammenfassen:

$$E(\mathbf{O}, \mathbf{t}, s, b_1, b_2, \dots, b_{n_b}) = \sum_{i=1}^n E_i(\mathbf{O}, \mathbf{t}, s, b_1, b_2, \dots, b_{n_b}) \quad (6.16)$$

Die Berechnung der Abstandsfunktion $d()$ zwischen einem Modellvertex \mathbf{p}_j und der zugehörigen Projektionsgeraden ist denkbar einfach (Abb. 6.16). Die Projektionsgerade verläuft zum einen durch das Zentrum C der zugehörigen virtuellen Kameradefinition und zum anderen durch den Konturpunkt (x_j^s, y_j^s) in der Bildebene. Unter Verwendung der internen Kameraparameter (optischer Nullpunkt $(c_x, c_y)^T$,

Brennweite f) läßt sich dieser Konturpunkt ebenfalls als 3D-Punkt in Kamerakoordinaten interpretieren: $(x_j^s - c_x, y_j^s - c_y, f)^T$. Hieraus ergibt sich direkt folgende Gleichung für die Projektionsgerade im Kamerakoordinatensystem:

$$\mathbf{q} = \lambda \begin{pmatrix} x_j^s - c_x \\ y_j^s - c_y \\ f \end{pmatrix} = \lambda \mathbf{g} \quad (6.17)$$

Die Länge des Differenzvektors zwischen dem 3D-Punkt \mathbf{p}_j und seiner senkrechten Projektion \mathbf{q} auf die Projektionsgerade g liefert schließlich den gesuchten Abstand:

$$d(\mathbf{p}_j, (x_j^s, y_j^s)) := d(\mathbf{p}_j, \mathbf{g}) = \left\| \mathbf{p}_j - \frac{\mathbf{p}_j \mathbf{g}}{\|\mathbf{p}_j\| \|\mathbf{g}\|} \mathbf{g} \right\|_2 \quad (6.18)$$

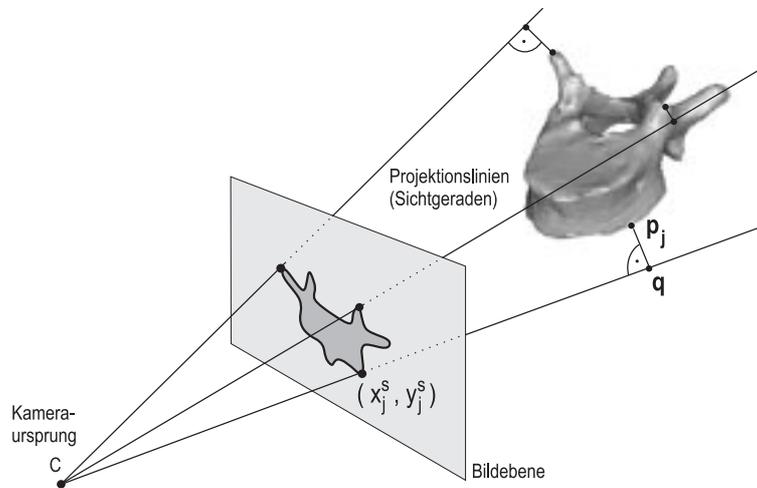


Abbildung 6.16: Abstandsdefinition zwischen Vertex und der Projektionslinie eines Konturpunktes

Auf welche Weise die betrachteten Vertices \mathbf{p}_j ausgewählt und insbesondere korrespondierenden Projektionslinien zugeordnet werden, blieb bisher unberücksichtigt. Im Hinblick auf eine effiziente Berechnung der Fehlerfunktion 6.16 kommt diesen Fragestellungen jedoch eine zentrale Bedeutung zu. Es wird vorgeschlagen, für eine Fluoroskopieaufnahme ausschließlich diejenigen Vertices des 3D-Modells zu berücksichtigen, die im Falle einer perspektivischen Projektion des Modells in die zugehörige Bildebene einen Beitrag zu der *äußeren* Kontur der resultierenden Silhouette liefern. Dabei ist zu beachten, daß sich i.a. für jede Aufnahme eine unterschiedliche Auswahl an Vertices ergibt. Anschließend werden 2D-Korrespondenzen zwischen den Projektionen dieser ausgewählten Vertices einerseits und segmentierten Konturpunkten andererseits in der Bildebene definiert. Die beiden nachfolgenden Abschnitte stellen für beide Schritte jeweils geeignete Verfahren vor.

6.5.2.1 Effiziente Konturbestimmung

Bei der perspektivischen Projektion eines 3D-Modells in die Bildebene einer virtuellen Lochkamera leisten nur sehr wenige Modellvertices und -kanten einen effektiven Beitrag zu inneren und äußeren Konturen in der Bildebene (*apparent contours, contour generators* [Gib98, Gue98]). Diese Modellanteile

lassen sich identifizieren, indem man die relative Lage und Orientierung des Modells bzgl. der Kamera berücksichtigt. Eine Dreiecksfläche des Modells (vgl. Anhang D) gilt aus Sicht der Kamera genau dann als *sichtbar*, falls die Flächennormale mit dem Richtungsvektor der Projektionslinie zwischen Kameraursprung und Dreieckszentrum einen stumpfen Winkel einschließt. Ansonsten ist die Fläche *nicht sichtbar*. Mit dieser Definition tragen genau diejenigen Modellkanten zu inneren und äußeren Konturen bei, die zwischen einer sichtbaren und einer nicht sichtbaren Dreiecksfläche verlaufen. Die Projektion aller Modellkanten mit dieser Eigenschaft liefert zunächst eine Menge von 2D-Linien, deren Endpunkte jeweils mit Vertices des Modells korrespondieren. Die geeignete Nachbearbeitung dieser Linienliste erlaubt es anschließend, alle Linien zu entfernen, die – für die vorliegende Arbeit nicht relevante – innere Konturen darstellen [Dö99]. Betrachtet man abschließend nur die Start- und Endpunkte der übrig gebliebenen Linien, so ergibt sich für jede Fluoroskopieaufnahme i eine zugehörige Konturpunktliste $V_i = \{(x_1^v, y_1^v), \dots, (x_{m_i}^v, y_{m_i}^v)\}$. Die Information darüber, welche Modellvertices für die Projektion der einzelnen Punkte verantwortlich sind, bleibt erhalten. Sie ist entscheidend für die spätere Berechnung des Fehlermaßes aus Gleichung 6.16.

6.5.2.2 Korrespondenzsuche

Nachdem für jede Fluoroskopieaufnahme eine Konturpunktliste V_i berechnet wurde, ist es nötig, Korrespondenzen zwischen den Punkten dieser Liste und Punkten der segmentierten Konturen S_i zu bilden. Hierzu wird ausgehend von jedem Punkt einer Liste der jeweils nächstgelegene Punkt der anderen Liste gesucht. Die Verwendung einer k-d-Baum-Datenstruktur (hier: $k = 2$) für die Speicherung der zu durchsuchenden Liste erlaubt einen effizienten Zugriff und beschleunigt die Suche [FBF77]. In welcher Richtung die Korrespondenzsuche erfolgt ($K_i \rightarrow S_i$ bzw. $S_i \rightarrow K_i$), hängt von folgenden Überlegungen ab:

- Stehen aus irgendeinem Grund durch die Segmentierung nur Teilfragmente der gesamten Knochensilhouette zur Verfügung, so ist eine ' $K_i \rightarrow S_i$ '-Zuordnung problematisch zu sehen, denn für einige der projizierten Modellvertices existieren in diesem Fall keine korrekten Korrespondenzen. D.h. diese Vertices werden zwangsläufig falschen Projektionslinien zugeordnet, was sich bei einer Häufung negativ auf die Optimierung auswirkt.
- Der umgekehrte Fall ergibt sich, wenn die Vertex-Verteilung des 3D-Modells nur sehr spärlich ausfällt. Eine ' $S_i \rightarrow K_i$ '-Korrespondenzsuche ordnet dann sehr viele Konturpunkte falsch zu.
- In der Regel ergeben sich aus der Kontursegmentierung deutlich mehr Konturpunkte als durch die Modellprojektion. Als Folge davon liefert eine ' $S_i \rightarrow K_i$ '-Zuordnung i.a. wesentlich mehr Korrespondenzen als umgekehrt. Trotz der im letzten Punkt angesprochenen Problematik wirkte sich dies in mehreren Versuchen sehr positiv auf die Optimierung aus. Für die Experimente am Ende dieses Kapitels kam daher ausschließlich eine ' $S_i \rightarrow K_i$ '-Zuordnung zum Einsatz.

Stehen die Korrespondenzen in der Bildebene fest, so ergeben sich daraus unmittelbar die für die Optimierung benötigten 3D-Korrespondenzen zwischen Modellvertices und Projektionslinien.

6.5.3 Initialisierung

Die drei bereits in Abschnitt 4.2 vorgestellten Verfahren zur Bestimmung einer hinreichend guten initialen Belegung der gesuchten Transformationsparameter wurden zunächst lediglich im Hinblick auf eine

CT-basierte, rigide Registrierung wie in Kapitel 5 entworfen. Sie eignen sich jedoch gleichermaßen für eine elastische Registrierung, solange die folgenden zusätzlichen Vereinbarungen getroffen werden:

1. Die Skalierung s des formvariablen Modells aus Gl. 6.16 wird während der Initialisierung auf einen konstanten Wert s_c gesetzt, so daß die Größe des Knochenmodells in etwa die zu erwartenden Größenverhältnisse des Patienten widerspiegelt.
2. Die Verformungsmöglichkeiten des Modells bleiben während der Initialisierung unberücksichtigt, d.h. für alle Verformungsparameter in Gl. 6.16 gilt: $b_i = 0$. In Zusammenhang mit Punkt 1 bedeutet dies, daß sich die Initialisierungsverfahren jeweils auf die rigide Lokalisation des auf Standardgröße skalierten Durchschnittsmodells konzentrieren.
3. Die ersten beiden Techniken aus Abschnitt 4.2 gehen von einer vorhandenen Computer-Tomographie-Aufnahme aus. Die daran gestellten Anforderungen lassen sich jedoch jeweils geradlinig auf das vorliegende formvariable Oberflächenmodell übertragen:
 - Die Vorgabe einer Aufnahmerichtung (Abschnitt 4.2.1) setzt implizit die Kenntnis der Patientenlage zum Zeitpunkt der präoperativen CT-Aufnahme voraus, um dadurch auf eine relative intraoperative Lage des Patienten auf dem OP-Tisch schließen zu können. Eine definierte Ausrichtung des formvariablen Modells bzgl. der Achsen des zugehörigen Objektkoordinatensystems erfüllt im vorliegenden Fall denselben Zweck.
 - Anstatt punktförmige anatomische Merkmale innerhalb des CT-Volumens festzulegen (Abschnitt 4.2.2), lassen sich hier direkt die während der Modellbildung gewonnenen Landmarken aus Abschnitt 6.4.4.1 heranziehen.

6.5.4 Optimierungsvarianten

Erste experimentelle Erkenntnisse hinsichtlich der Minimierung der Fehlerfunktion 6.16 zeigten, daß es sinnvoll ist, mehrere unterschiedliche Optimierungsverfahren in geeigneter Kombination zu verwenden. Nicht zuletzt aufgrund der erhöhten Dimension der Fehlerfunktion brachte keines der angewandten Verfahren alleine bei Betrachtung unterschiedlicher Testdatensätze stets optimale Resultate. Erst die sukzessive Anwendung mehrerer Minimierungsschritte, wobei in jedem Schritt ein anderes Verfahren einen zuvor festgelegten Teil der gesuchten Parameter optimiert, brachte sehr gute Resultate.

Im folgenden werden daher vier unabhängige Optimierungsverfahren vorgestellt. Zunächst zeigt sich, daß unter bestimmten Voraussetzungen für einen Teil der Parameter eine geschlossene Lösung existiert (Abschnitt 6.5.4.1). Bei den restlichen Verfahren handelt es sich im einzelnen um das Levenberg-Marquardt-Verfahren (Abschnitt 6.5.4.2), eine Randomisierte Suche (Abschnitt 6.5.4.3) sowie eine diskrete Rastersuche (Abschnitt 6.5.4.4). Daran anschließend wird eine Optimierungsstrategie vorgeschlagen, die im Rahmen dieser Arbeit stets erfolgreich angewandt werden konnte. Diese Strategie legt dabei fest, welches der Verfahren in welcher Reihenfolge zur Optimierung welcher Parameter herangezogen wird.

6.5.4.1 Geschlossene Lösung der Deformation

Unter der Annahme, daß die meisten aktuellen Korrespondenzen zwischen Modellvertices und Konturnpunkten bzw. Punkten auf den zugehörigen Projektionslinien bereits annähernd korrekt sind, läßt

sich die Fehlerfunktion 6.16 bzgl. aller Parameter außer der Orientierung \mathbf{O} geschlossen minimieren. D.h. die Deformationsparameter b_i , die Skalierung s sowie die Komponenten des Translationsvektors $\mathbf{t} = (t_x, t_y, t_z)^T$ können direkt durch Auflösen eines entsprechenden linearen Gleichungssystems berechnet werden. Dieses Gleichungssystem wird im folgenden hergeleitet.

Zunächst soll lediglich die Situation für einen einzelnen Vertex $\mathbf{v} = (v_x, v_y, v_z)$ untersucht werden. Dieser sei durch die vorangegangene Korrespondenzsuche bereits dem Konturpunkt $\mathbf{p} = (x_p, y_p)$ zugeordnet. Gesucht sind in diesem Fall diejenigen Parameter $(s, t_x, t_y, t_z, b_1, \dots, b_{n_b})$, die den betrachteten Vertex \mathbf{v} auf einen Zielpunkt \mathbf{q} abbilden. Diesen Zielpunkt \mathbf{q} erhält man durch die senkrechte Projektion von \mathbf{v} auf die zu \mathbf{p} gehörige Projektionslinie. In Form einer Gleichung zusammengefaßt ergibt sich:

$$\mathbf{O}s(\mathbf{v} + b_1\mathbf{d}_1 + \dots + b_{n_b}\mathbf{d}_{n_b}) + \mathbf{t} = \mathbf{q} \quad (6.19)$$

Die Orientierungsmatrix \mathbf{O} sowie die Deformationsvektoren \mathbf{d} werden dabei als konstant betrachtet. Durch Multiplikation mit der inversen Orientierungsmatrix \mathbf{O}^{-1} ergibt sich:

$$s(\mathbf{v} + b_1\mathbf{d}_1 + \dots + b_{n_b}\mathbf{d}_{n_b}) + \mathbf{O}^{-1}\mathbf{t} = \mathbf{O}^{-1}\mathbf{q} \quad (6.20)$$

Durch die beiden Substitutionen $\mu_i = sb_i$ und $\mathbf{q}' = (q'_x, q'_y, q'_z)^T = \mathbf{O}^{-1}\mathbf{q}$ vereinfacht sich die Gleichung entsprechend:

$$s\mathbf{v} + \mu_1\mathbf{d}_1 + \dots + \mu_{n_b}\mathbf{d}_{n_b} + \mathbf{O}^{-1}\mathbf{t} = \mathbf{q}' \quad (6.21)$$

Betrachtet man alle Einzelkomponenten der beiden dreidimensionalen Punkte \mathbf{v} und \mathbf{q}' sowie der n -dimensionalen Deformationsvektoren \mathbf{d}_i für sich, so resultiert daraus bereits ein lineares Gleichungssystem, bestehend aus drei Einzelgleichungen:

$$\begin{pmatrix} v_x & d_{x1} & \dots & d_{xn} & o_{11} & o_{21} & o_{31} \\ v_y & d_{y1} & \dots & d_{yn} & o_{12} & o_{22} & o_{32} \\ v_z & d_{z1} & \dots & d_{zn} & o_{13} & o_{23} & o_{33} \end{pmatrix} \begin{pmatrix} s \\ \mu_1 \\ \vdots \\ \mu_n \\ t_x \\ t_y \\ t_z \end{pmatrix} = \begin{pmatrix} q'_x \\ q'_y \\ q'_z \end{pmatrix} \quad (6.22)$$

mit

$$\mathbf{O} = \begin{pmatrix} o_{11} & o_{12} & o_{13} \\ o_{21} & o_{22} & o_{23} \\ o_{31} & o_{32} & o_{33} \end{pmatrix} \quad \text{und} \quad \mathbf{O}^{-1} = \mathbf{O}^T = \begin{pmatrix} o_{11} & o_{21} & o_{31} \\ o_{12} & o_{22} & o_{32} \\ o_{13} & o_{23} & o_{33} \end{pmatrix} \quad (6.23)$$

D.h. ein einzelner Vertex \mathbf{v} liefert zusammen mit einem korrespondierenden Zielpunkt \mathbf{q} entsprechend deren Dimension letztlich drei voneinander unabhängige Gleichungen. Da insgesamt $n + 4$ unbekannte Parameter gesucht sind, reicht dieses unterbestimmte Gleichungssystem jedoch für eine eindeutige Lösung nicht aus. Die Hinzunahme von insgesamt m weiteren Vertices mit jeweils zugehörigen Zielpunkten auf den korrespondierenden Projektionslinien liefert insgesamt $3m$ zusätzliche Gleichun-

gen, wodurch sich das folgende große Gleichungssystem ergibt³:

$$\begin{pmatrix} v_{x1} & d_{x11} & \dots & d_{x1n} & o_{11} & o_{21} & o_{31} \\ v_{y1} & d_{y11} & \dots & d_{y1n} & o_{12} & o_{22} & o_{32} \\ v_{z1} & d_{z11} & \dots & d_{z1n} & o_{13} & o_{23} & o_{33} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ v_{xm} & d_{xm1} & \dots & d_{xmn} & o_{11} & o_{21} & o_{31} \\ v_{ym} & d_{ym1} & \dots & d_{ymn} & o_{12} & o_{22} & o_{32} \\ v_{zm} & d_{zm1} & \dots & d_{zmn} & o_{13} & o_{23} & o_{33} \end{pmatrix} \begin{pmatrix} s \\ \mu_1 \\ \vdots \\ \mu_n \\ t_x \\ t_y \\ t_z \end{pmatrix} = \begin{pmatrix} q'_{1x} \\ q'_{1y} \\ q'_{1z} \\ \vdots \\ q'_{mx} \\ q'_{my} \\ q'_{mz} \end{pmatrix} \quad (6.24)$$

In kompakter Matrixdarstellung ergibt sich:

$$A \begin{pmatrix} s \\ \mu_1 \\ \vdots \\ \mu_n \\ t_x \\ t_y \\ t_z \end{pmatrix} = A\mathbf{x} = \mathbf{z} \quad (6.25)$$

Dieses System ist in der Regel überbestimmt, da meist nur wenige Formparameter b_i (bzw. μ_i) einer großen Zahl an vorhandenen Korrespondenzen zwischen Modellvertices und Projektionslinien gegenüber stehen ($m \gg n$). Aufgrund unvermeidlicher kleinster Rundungsfehler in den einfließenden Werten besitzt das lineare System in der Regel zunächst keine Lösung. Das numerische Verfahren der linearen Ausgleichsrechnung [Sto89] ermöglicht es dennoch, eine eindeutige approximative Lösung \mathbf{x}' des Systems zu bestimmen. Dabei wird garantiert, daß die Summe der quadratischen Abweichungen zwischen den Komponenten von \mathbf{z} und $A\mathbf{x}'$ für diese Lösung minimal ist. Die Ausgleichsrechnung entspricht dabei der Auflösung des folgenden Gleichungssystems:

$$A^T A\mathbf{x}' = A^T \mathbf{z} \quad (6.26)$$

Die einzelnen Komponenten des resultierenden Parametervektors \mathbf{x}' liefern anschließend direkt die optimale Skalierung s , den optimalen Translationsvektor \mathbf{t} sowie durch entsprechende Umrechnung ($b_i = \mu_i s^{-1}$) die optimalen Gewichtungskoeffizienten b_i der Verformungsmodi.

Die Qualität dieser berechneten Parameter hängt letztlich davon ab, wie gut die einfließenden Korrespondenzen zwischen Vertices und Projektionslinien bzw. Konturpunkten ausfallen. Da die Qualität dieser Korrespondenzen jedoch selbst von einer hinreichend guten initialen Belegung der gesuchten Parameter abhängt, erscheint eine Anwendung des vorgestellten geschlossenen Lösungsverfahrens nur zu einem Zeitpunkt sinnvoll, an dem bereits gute Parameterschätzungen bestimmt wurden.

6.5.4.2 Gradientenabstieg (Levenberg-Marquardt)

Das bekannte Levenberg-Marquardt-Verfahren eignet sich bekanntermaßen gut für die *least-squares*-Minimierung nicht linearer, mehrdimensionaler Funktionsverläufe. Dahinter verbirgt sich eine elegante

³Die ausgewählten Zielpunkte können dabei ohne Einschränkung von Projektionslinien verschiedener Ansichten bzw. Konturen stammen.

Technik, um während der iterativen Minimierung fließend zwischen zwei anderen bekannten Minimierungsverfahren zu wechseln: der inversen Hesse-Methode auf der einen Seite sowie der Methode des steilsten Abstiegs auf der anderen Seite [PTo94]. Die zuletzt genannte Methode findet Anwendung, solange sich die Minimierung noch weit entfernt von der gesuchten Lösung befindet. Bei Annäherung an die gesuchte Lösung erfolgt der Wechsel auf die zuerst genannte Methode. Im Rahmen dieser Arbeit kommt eine in [Mor80] beschriebene Implementierung des Levenberg-Marquardt-Verfahrens zum Einsatz. Der zugrundeliegende Gradientenabstieg macht das Verfahren prinzipbedingt anfällig für lokale Minima der Fehlerfunktion.

6.5.4.3 Randomisierte Suche

Einen prinzipiellen Ausweg aus lokalen Minima bieten diverse zufallsbasierte Optimierungsverfahren wie z.B. *Simulated Annealing* [PTo94] oder auch *Genetische Algorithmen* [Sch93]. In dieser Arbeit kommt hingegen das sehr einfache Konzept einer randomisierten Suche zum Einsatz. Ausgehend von einer initial gegebenen Parameterbelegung wird iterativ eine Vielzahl weiterer Parameterbelegungen getestet, die sich jeweils durch zufällige Veränderung einzelner Parameterwerte ergeben. Entstehen dabei bessere Parametersätze als bisher, so werden diese als Ausgangsbasis für die folgende Iteration herangezogen. Die Vorgabe einer maximalen Anzahl an Iterationen sowie eines zu unterschreitenden Fehlerschwellwertes beschränkt die Laufzeit der randomisierten Suche.

Dieses Verfahren funktioniert in der Regel dann recht gut, wenn es darum geht, ausgehend von einer gegebenen Parameterbelegung nach evtl. noch besseren Lösungen zu suchen, die dann als Ausgangsbasis für ein anderes Optimierungsverfahren dienen können.

6.5.4.4 Rastersuche

Die Idee einer Rastersuche besteht darin, ausgewählte Parameter innerhalb eines vorgegebenen Bereiches zu diskretisieren. Anschließend wird die Fehlerfunktion 6.16 im Rahmen einer erschöpfenden kombinatorischen Suche für alle möglichen Parameterkombinationen ausgewertet. Zwei grundsätzliche Problemstellungen sind bei dieser Art von Optimierung zu beachten. Die kombinatorische Komplexität erlaubt lediglich grobe Diskretisierungsschritte und die gleichzeitige Betrachtung einiger weniger Parameter, z.B. ausschließlich der drei rotatorischen Freiheitsgrade⁴. Alle übrigen Parameter sollten zu diesem Zeitpunkt bereits annähernd gut bestimmt worden sein. Außerdem muß sichergestellt sein, daß die ausgewählten Diskretisierungsbereiche die jeweils optimale Belegung eines Parameters überhaupt enthalten. Auf der anderen Seite garantiert dieses Verfahren das Auffinden des globalen Optimums bzgl. der vorgegebenen Diskretisierung.

6.5.5 Optimierungsstrategie

Es wurde bereits angesprochen, daß keines der im letzten Abschnitt vorgestellten Optimierungsverfahren, für sich alleine betrachtet, durchwegs akzeptable Registrierungsergebnisse liefert. Daher erscheint es sinnvoll, die einzelnen Verfahren im Rahmen einer Strategie geeignet zu kombinieren. Die Aufgaben der Optimierungsstrategie liegen dabei in

⁴Die Komplexität einer Rastersuche beträgt $O(s^n)$ (s Diskretisierungsschritte, n freie Parameter). Für den Fall, daß beispielsweise lediglich die Rotation (3 Freiheitsgrade) mit jeweils 10 Unterteilungen *gerastert* wird, ergeben sich bereits $10^3 = 1000$ Aufrufe der Fehlerfunktion.

1. einer Auswahl der Optimierungsverfahren,
2. der Festlegung einer Anwendungsreihenfolge, wobei das Ergebnis eines Optimierungsschrittes als initiale Parameterbelegung des nächsten Schrittes herangezogen wird und
3. der Entscheidung, zu welchem Zeitpunkt welche Parametergruppen getrennt voneinander optimiert werden.

Die nachfolgende Tabelle veranschaulicht das Prinzip einer Optimierungsstrategie. Die aufgeführte Vorgehensweise besitzt dabei nicht nur beispielhaften Charakter, sondern führt im praktischen Einsatz bereits zu sehr guten Registrierungsergebnissen. Die im nächsten Abschnitt präsentierten experimentellen Ergebnisse basieren weitestgehend auf der Verwendung exakt dieser Strategie.

Verfahren	Rotation (O)	Translation (t)	Skalierung (s)	Deformation (b_i)
1 Randomisiert	✓	✓	–	–
2 Levenberg-M.	✓	✓	–	–
3 Levenberg-M.	–	–	✓	✓
4 Levenberg-M./LGS	✓	✓ (LGS)	✓ (LGS)	✓ (LGS)

Diese Optimierungsstrategie basiert auf den folgenden Überlegungen. Zunächst soll eine randomisierte Suche (Schritt 1) die initiale Belegung der rigiden Parameter (Rotation, Translation) verbessern. Ausgehend von der besten gefundenen Lösung wird anschließend in einem zweiten Schritt das Levenberg-Marquardt-Verfahren angewandt, wobei sich die Optimierung wiederum auf Rotation und Translation beschränkt. Skalierung und Deformation bleiben in beiden Schritten zunächst deaktiviert, d.h. die zugehörigen Parameter werden bzgl. ihrer initialen Belegung *eingefroren*. Ansonsten kann es vorkommen, daß die Optimierung z.B. versucht, einen tatsächlichen Rotationsfehler fälschlicherweise durch eine skalierte Deformation auszugleichen. Nach Beendigung der ersten beiden Optimierungsschritte sollte die rigide Ausrichtung des formvariablen Modells bereits annähernd korrekt sein. Nun wird versucht, die Deformation und Skalierung des Modells durch Levenberg-Marquardt-Optimierung anzupassen (Schritt 3), wobei dieses Mal die Parameter für Rotation und Translation deaktiviert bleiben. Durch diesen Schritt sollten sich insgesamt sehr gute Korrespondenzen zwischen Modellvertices und Konturpunkten bzw. Projektionslinien ergeben, so daß in einem vierten und zugleich finalen Optimierungsschritt die geschlossene Berechnung von Skalierung, Translation und Deformation durch ein lineares Gleichungssystem eingesetzt werden kann. Die rotatorischen Parameter werden gleichzeitig durch Levenberg-Marquardt minimiert (Schritt 4).

Das vorgestellte vierstufige Optimierungsverfahren realisiert eine rein *statische* Strategie, d.h. bereits vor der Registrierung steht exakt fest, welche Verfahren in welcher Reihenfolge welche Parameter optimieren. Denkbar sind aber auch adaptive, dynamische Strategien. Ein *Strategie-Controller* könnte dabei selbstständig über Auswahl und Reihenfolge der angewandten Optimierungsverfahren entscheiden, wobei als Entscheidungskriterium u.a. der durch die Fehlerfunktion berechnete durchschnittliche Abstand zwischen allen Vertices und korrespondierenden Projektionslinien herangezogen wird. Beispielsweise erscheint es sinnvoll, durch eine randomisierte Suche eine bessere Parameterbelegung als Ausgangspunkt zu suchen, falls das Levenberg-Marquardt-Verfahren mit einem sehr hohen Restfehler konvergiert. Resultiert ein kleiner Restfehler, so könnte eine abschließende Rastersuche nach einer besseren Parameterbelegung in unmittelbarer Umgebung Abhilfe schaffen. Im Rahmen dieser Arbeit wurde dieser Ansatz

jedoch nicht weiterverfolgt.

Der konkrete Aufbau einer Strategie hängt entscheidend davon ab, welches Vorwissen über die Registrierungssituation bereits bekannt ist: Wie gut sind die initialen Belegungen der rigiden Parameter? Ist der zu registrierende Knochen im Vergleich zu der Durchschnittsform des verwendeten formvariablen Modells stark verformt? Insbesondere hat auch der jeweilige Anwendungszweck bzw. die damit u.U. verbundenen zeitlichen Restriktionen Einfluß auf die Auswahl der Algorithmen. Die präoperative Generierung eines rigiden patientenspezifischen Planungsmodells erlaubt z.B. die Mehrfachanwendung der randomisierten Suche. Hier kommt es i.a. nicht darauf an, ob die Modellbildung zwei, fünf oder 15 Minuten benötigt. Im intraoperativen Umfeld, z.B. im Falle einer *spontanen Notfallregistrierung* mit anschließender Navigation, ist jedoch eine Optimierungszeit von deutlich mehr als zwei bis drei Minuten meist nicht tragbar.

6.6 Ergebnisse

Um die Leistungsfähigkeit der vorgestellten elastischen Registrierung isoliert, d.h. unabhängig von der Qualität der segmentierten Konturen sowie der Präzision der Fluoroskop-Kalibrierung, beurteilen zu können, werden wie in Abschnitt 5.5 zunächst einige aussagekräftige Simulationsergebnisse vorgestellt. Es folgen mehrere Untersuchungen mit realen Fluoroskopieaufnahmen von Femur-Knochen und Wirbeln.

6.6.1 Simulationsergebnisse

Gemeinsame Grundlage der nachfolgend beschriebenen Experimente ist ein formvariables Wirbel-Modell, das aus insgesamt vier Trainingsmodellen generiert wurde. Um trotz der geringen Modelanzahl aussagekräftige Verformungen zu erhalten, wurden hierbei gezielt je zwei T11- und T6-Brustwirbelmodelle für das Training herangezogen, bei denen sich die Ausprägungen einzelner Formdetails (Wirbelkörperform, Dorn- und Querfortsätze) stark voneinander unterscheiden. Die Abbildungen 6.17 und 6.18 zeigen die beiden wichtigsten Verformungsmodi dieses Modells. Dabei fällt auf, daß der erste Modus bereits in der Lage ist, die wichtigsten Formunterschiede zwischen einem T11- und einem T6-Wirbel gleichzeitig zu modellieren. Demgegenüber beinhaltet der zweite Modus lediglich noch marginale Formänderungen, die in der Abbildung bereits kaum mehr wahrzunehmen sind. Belegt wird diese Feststellung durch das Ergebnis der Eigenwertanalyse in Tabelle 6.1. Es zeigt sich, daß der erste Verformungsmodus nahezu 70% der gesamten in den Trainingsdaten enthaltenen Formvariationen auf sich vereint. Auf den zweiten Modus entfallen lediglich noch etwa 15%. Um insgesamt knapp 99% der Deformation modellieren zu können, genügen bereits 5 Deformationsvektoren. Den folgenden Experimenten liegen daher stets $n_b = 5$ Deformationsparameter b_1, \dots, b_5 zugrunde. Angesichts der sehr wenigen Trainingsmodelle war dieses Ergebnis der Eigenwertverteilung zu erwarten. Dennoch wurde bewußt auf die physikalisch motivierte FEM-Modellierung zusätzlicher Trainingsbeispiele (Abschnitt 6.3) verzichtet, um das statistische Deformationsverhalten isoliert beurteilen zu können.

Ein erstes Experiment sollte qualitative und quantitative Aussagen über das Konvergenzverhalten der elastischen Optimierung liefern. Grundlage des Versuchs war eine Szene, bestehend aus zwei simulierten Fluoroskopieaufnahmen, die entsprechend projizierte Wirbelsilhouetten (Konturen) enthielten. Der Zwischenwinkel der beiden Aufnahmerichtungen betrug knapp 90° (Abb. 6.19).

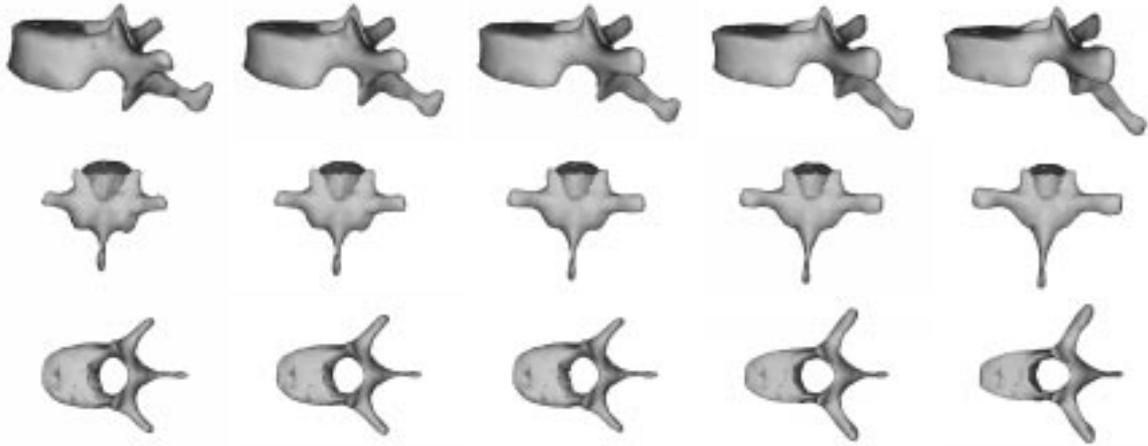


Abbildung 6.17: Formvariables Wirbelmodell (Verformungsmodus 1)

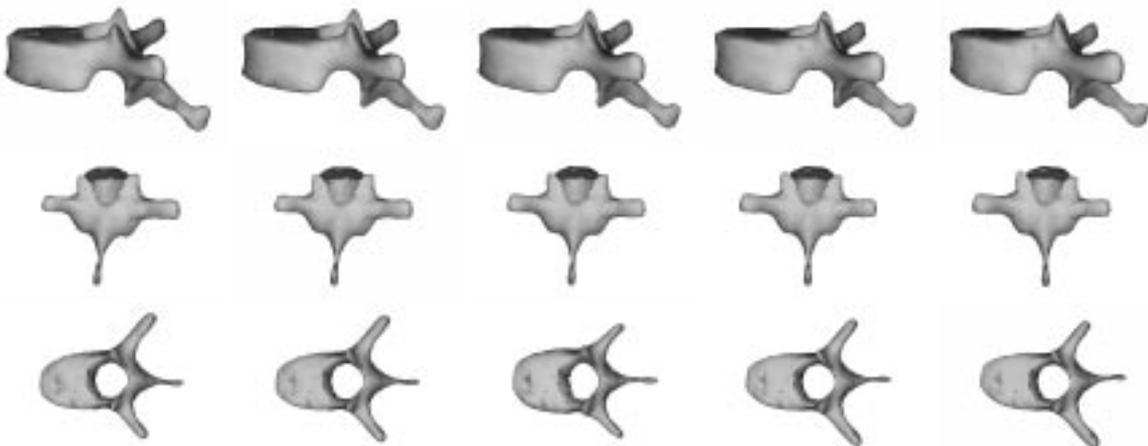


Abbildung 6.18: Formvariables Wirbelmodell (Verformungsmodus 2)

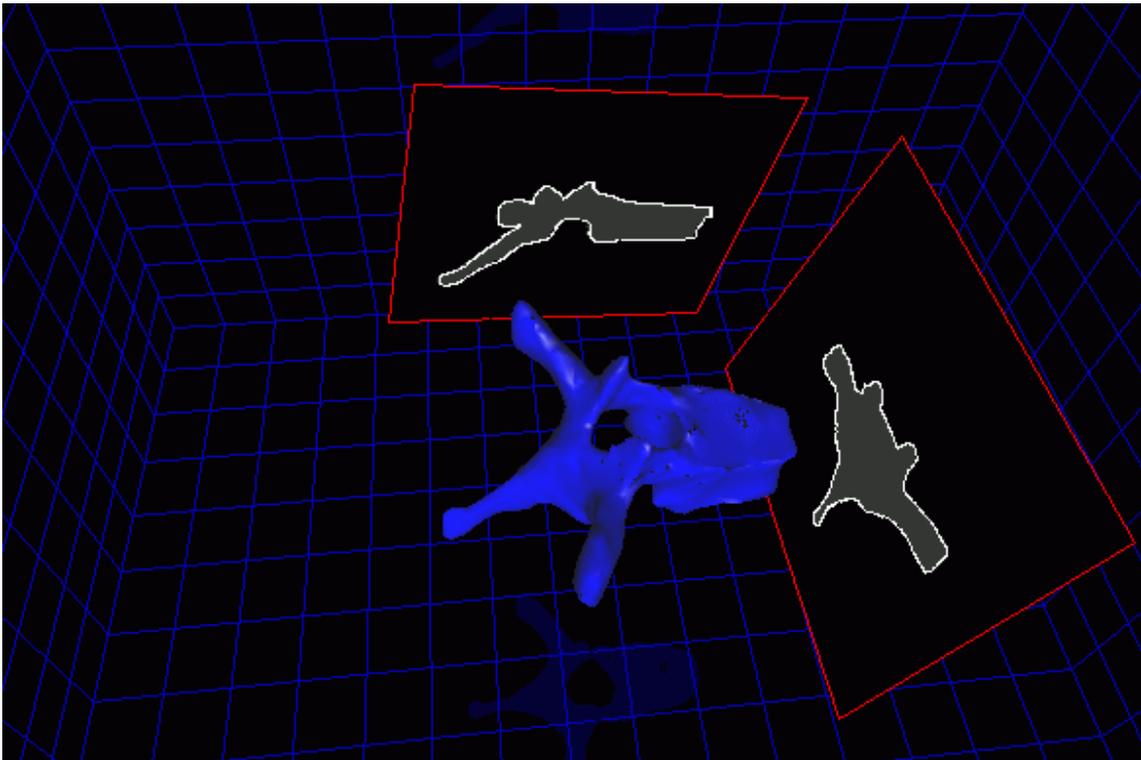


Abbildung 6.19: Virtuelle Darstellung der simulierten 3D-Aufnahmeszene zur Beurteilung des Konvergenzverhaltens der elastischen Optimierung

k	Eigenwert λ_k	$\sum_1^k \lambda_i$	Anteil an Gesamtverformung ($\sum_1^k \lambda_i / \sum_1^N \lambda_i$)
1	18.642690	18.642690	67.07%
2	4.434546	23.077235	83.02%
3	2.962815	26.040050	93.68%
4	1.163193	27.203243	97.86%
5	0.261785	27.465028	98.80%
6	0.154571	27.619600	99.36%
7	0.076208	27.695807	99.63%
8	0.034039	27.729846	99.76%
9	0.031813	27.761660	99.87%
10	0.029734	27.791394	99.98%
11	0.005995	27.797389	100.00%
12	0.000000	27.797389	100.00%

Tabelle 6.1: Eigenwertanalyse für das formvariable Modell aus Abb. 6.17 und 6.18

Insgesamt wurden 35 Optimierungsläufe durchgeführt, wobei die initialen Lageparameter mit unterschiedlichen rotatorischen und translatorischen Abweichungen bezüglich der korrekten Lageparameter besetzt waren. Durch die Simulation der Konturen auf Basis des unverformten, einheitlich skalierten Modells waren diese optimalen Parameter exakt bekannt. Die initiale translatorische Abweichung betrug maximal $50mm$, die größte initiale rotatorische Abweichung 45° . Wie bereits in Abschnitt 5.5.1 spielte dabei weder die Wahl der zugehörigen Rotationsachse noch die Richtung der Translation eine Rolle. Beide wurden jeweils zufällig ausgewählt. Die Registrierung erfolgte durch die in Abschnitt 6.5.5 vorgeschlagene vierstufige Optimierungsstrategie. Da die Konturen durch Projektion des unverformten Modells generiert wurden, sollte sich in diesem Versuch nach abgeschlossener Optimierung insbesondere keine Verformung gegenüber der Durchschnittsmodellform zeigen.

Abb. 6.20 präsentiert die Ergebnisse der Optimierungsläufe, d.h. die resultierenden Restfehler in Abhängigkeit von den initialen Abweichungen. Neben den bereits aus Abschnitt 5.5 vertrauten rotatorischen bzw. translatorischen Restfehlern lassen sich für die elastische Registrierung weitere intuitive Restfehlerwerte definieren, die insbesondere die Qualität des Deformationsverhaltens charakterisieren. Ein durch fehlerhafte Skalierung und Deformation hervorgerufener *Formfehler* soll unabhängig von Rotation und Translation betrachtet werden. Hierzu wird zunächst die Größe und Form des Durchschnittsmodells entsprechend der berechneten Parameter für Skalierung und Verformung verändert. Anschließend wird dieses verformte Modell Vertex für Vertex mit dem für die Kontursimulation verantwortlichen – ggf. ebenfalls verformten – Modell verglichen. Dies liefert sowohl einen durchschnittlichen als auch einen maximalen mm -Abstand, die zusammen ein intuitives Maß zur Beurteilung von Formfehlern darstellen.

Es zeigt sich, daß für maximale initiale Abweichungen von etwa $20mm$ bzw. 20° die Optimierung sehr gute Ergebnisse liefert: rotatorische Restfehler von maximal 1.6° stehen maximalen translatorischen Restfehlern von $1.1mm$ gegenüber. Der resultierende Formfehler lies sich dabei auf einen durchschnittlichen Vertex-Abstand von $0.6mm$ begrenzen, maximal ergaben sich $2.7mm$. Das minimierte Fehlermaß, d.h. der durchschnittliche Abstand von Modellvertices zu korrespondierenden Projektionslinien wies nach Beendigung der Registrierung jeweils Werte um $0.5mm$ auf.

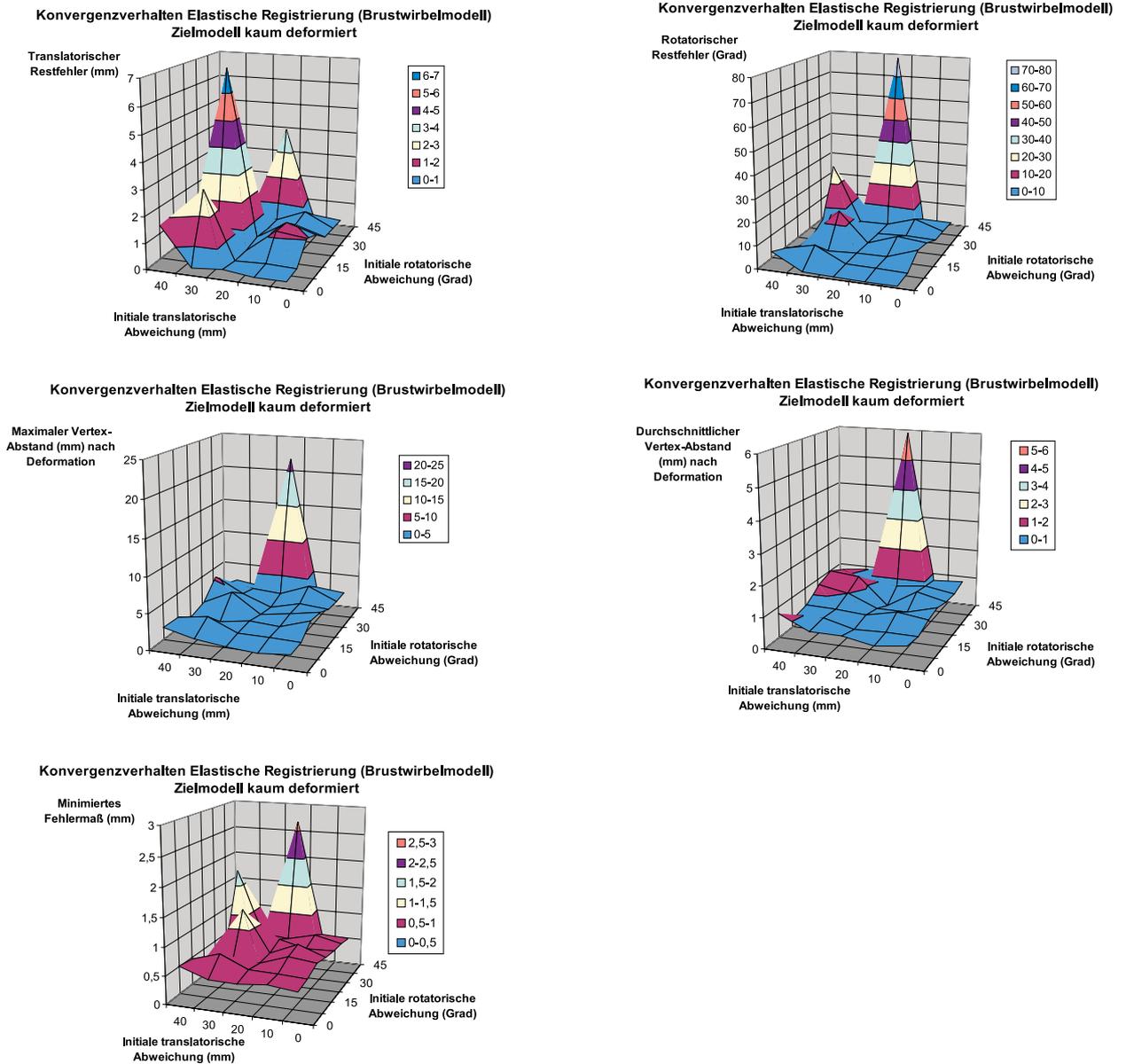


Abbildung 6.20: Konvergenzverhalten bei simulierten Testdaten (zwei Aufnahmen eines geringfügig deformierten Wirbels, Winkelabstand der Blickrichtungen 90° (lateral, AP)) für ein formvariables Brustwirbelmodell. Angetragen sind jeweils translatorische und rotatorische Restfehler, maximaler und durchschnittlicher Deformationsfehler (Vertex-Abstand) sowie das minimierte Fehlermaß (mittlerer 3D-Abstand Objekt/Projektionslinien) nach erfolgter Konvergenz der Optimierung, ausgehend von unterschiedlichen initialen translatorischen und rotatorischen Abweichungen.

Abb. 6.21 zeigt verschiedene Momentaufnahmen einer der obigen Optimierungsläufe in der Bildebene. Dargestellt wird jeweils die Zielkontur im Röntgenbild, projizierte Modellvertices sowie aktuell gebildete Korrespondenzbeziehungen. Dabei ist gut zu erkennen, wie sich die Modellprojektion sukzessive ohne sichtbare Verformung an die gegebenen Konturen anpaßt.

Ein weiteres Experiment sollte das Konvergenzverhalten im Falle signifikanter Deformationen untersuchen. Abgesehen davon, daß Konturen eines stark deformierten Formmodells vorgegeben wurden, entsprachen die Versuchsbedingungen exakt denen des letzten Experimentes, wobei geringere initiale Abweichungen untersucht wurden. Abb. 6.22 faßt die Ergebnisse zusammen. Sollen die rigiden Restfehler maximal 0.5mm bzw. 1.1° betragen, so resultiert ein eingeschränkter Konvergenzradius von etwa 10mm bzw. 15° initialer Abweichung. Die vorgegebene starke Deformation wurde durch die Optimierung einwandfrei erkannt. Der resultierende Formfehler betrug 0.9mm bzw. 1 bis 2.5mm (Durchschnitt/Maximum). Auch Abb. 6.23 veranschaulicht die Qualität der Verformung. Die Momentaufnahmen der Minimierung entstammen dabei dem dritten Optimierungsschritt der angewandten Strategie (Levenberg-Marquardt-Optimierung der Deformationsparameter inklusive Skalierung). Deutlich ist die Anpassung der Krümmung und Form des Dornfortsatzes zu erkennen.

Zur Beurteilung der vorgeschlagenen mehrstufigen Optimierungsstrategie empfiehlt sich ein Blick auf einen typischen Restfehlerverlauf des verwendeten Fehlermaßes (mittlerer Abstand Modell/Projektionslinien). Abb. 6.24 zeigt einen solchen Verlauf für eine der vorangegangenen Optimierungen. Abgesehen vom ersten Optimierungslauf (randomisierte Suche) wurde jeweils der bisher kleinste Fehler angetragen. Die randomisierte Suche reduziert den durchschnittlichen Abstand zwischen Modellvertices und zugehörigen Projektionslinien von initial 8.4mm auf 6.6mm . Anschließend gelingt es der, auf die rigiden Parameter beschränkten, Levenberg-Marquardt-Optimierung, den Fehler bis auf 4.8mm zu drücken. Die Optimierung der Deformationsparameter inklusive Skalierung bringt eine weitere Reduktion um 2.8mm , ehe die abschließende Optimierung aller Parameter einen Restfehler von unter 0.5mm ergibt.

Ein abschließender Versuch soll zeigen, welchen Einfluß unterschiedliche Zwischenwinkel bei der Auswahl von Aufnahmerichtungen auf die Qualität der Registrierung haben. Hierzu wurden Konturen für fünf Paare von Aufnahmerichtungen mit Zwischenwinkeln im Bereich von $30^\circ \dots 90^\circ$ generiert. Für jedes Paar wurden mehrere Optimierungen durchgeführt, wobei initiale Abweichungen bis maximal 5mm bzw. 15° fest ausgewählt wurden. Abb. 6.25 zeigt das Streudiagramm der rigiden Restfehler, Abb. 6.26 entsprechend die Deformationsfehlerstreuung. Ein Vergleich der beiden Diagramme zeigt deutlich, daß mit veringertem Zwischenwinkel die Größe der Restfehler tendenziell zunimmt. Zwar ergab die Aufnahmesituation für 30° überraschenderweise nur sehr geringe rigide Restfehler, ein Blick auf die hohen zugehörigen Deformationsfehler relativiert dieses vermeintlich gute Verhalten jedoch wieder.

6.6.2 Anwendung auf reale Fluoroskopieaufnahmen

Die elastische Lagerkennung basiert im Gegensatz zu der intensitätsbasierten Registrierung aus Kapitel 5 nicht auf den Fluoroskopieaufnahmen selbst. Stattdessen wird durch die vorgeschaltete Kontursegmentierung eine geeignete Zwischenrepräsentation gebildet. Als Konsequenz daraus lassen sich die Erkenntnisse des letzten Abschnitts nahezu direkt auf reale Fluoroskopieaufnahmen übertragen, solange ausschließlich fehlerfrei segmentierte Konturen vorliegen. Negative Auswirkungen der Röntgenbildverstärkerkalibrierung (Anhang B) auf die Registrierung konnten in ersten Versuchen nicht beobachtet werden. In der Tat stellt die Kontursegmentierung selbst oftmals ein Problem dar, wie folgendes Bei-

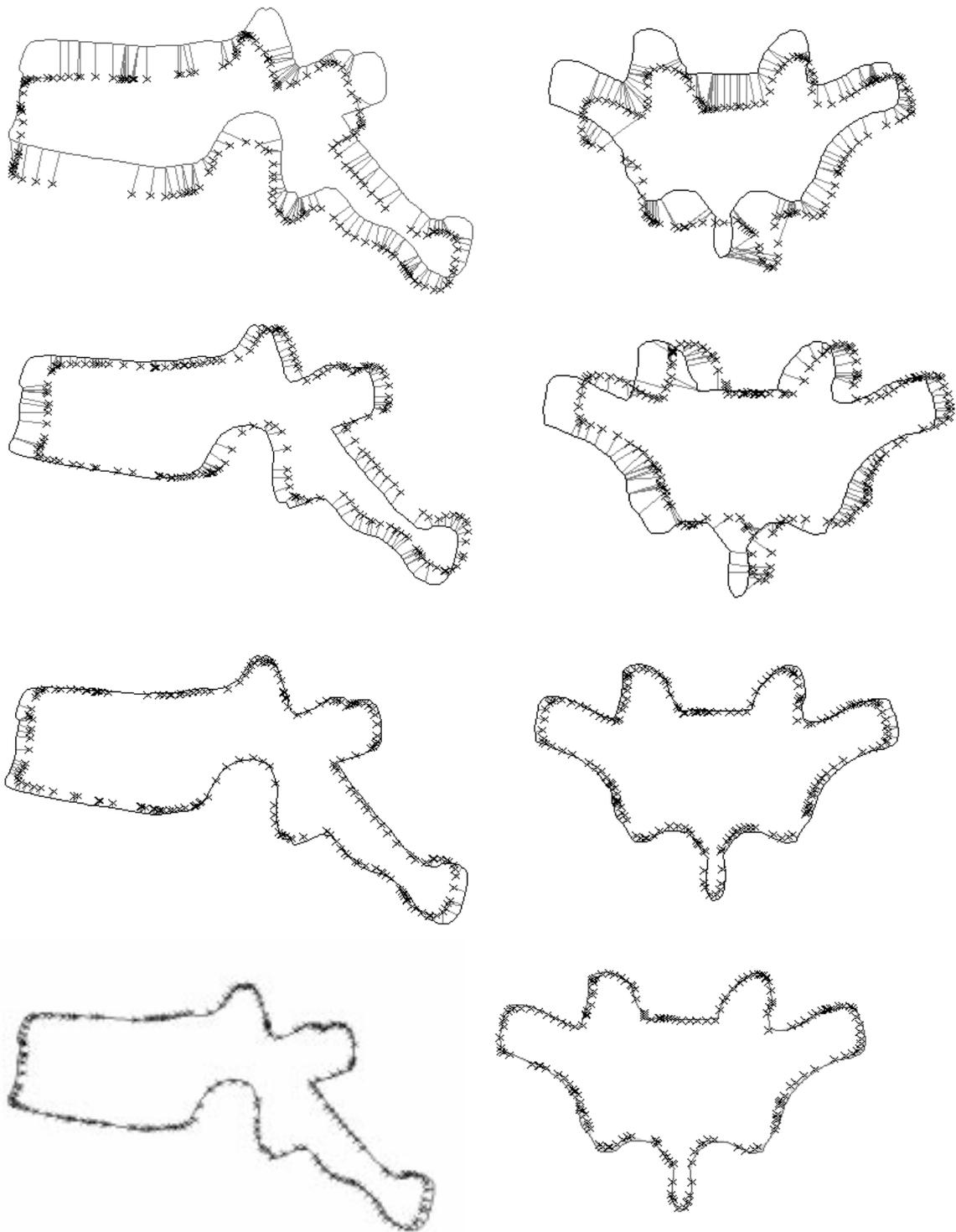


Abbildung 6.21: Anpassung des formvariablen 3D-Brustwirbelmodells an zwei vorgegebene Konturen (lateral, AP) eines nur *geringfügig verformten* Wirbels: Initialisierung (oben), Zwischenergebnisse (Mitte), Konvergenz (unten). Graue Linien veranschaulichen Korrespondenzen zwischen Konturpunkten und Projektionen der Modellvertices (Kreuzsymbole).

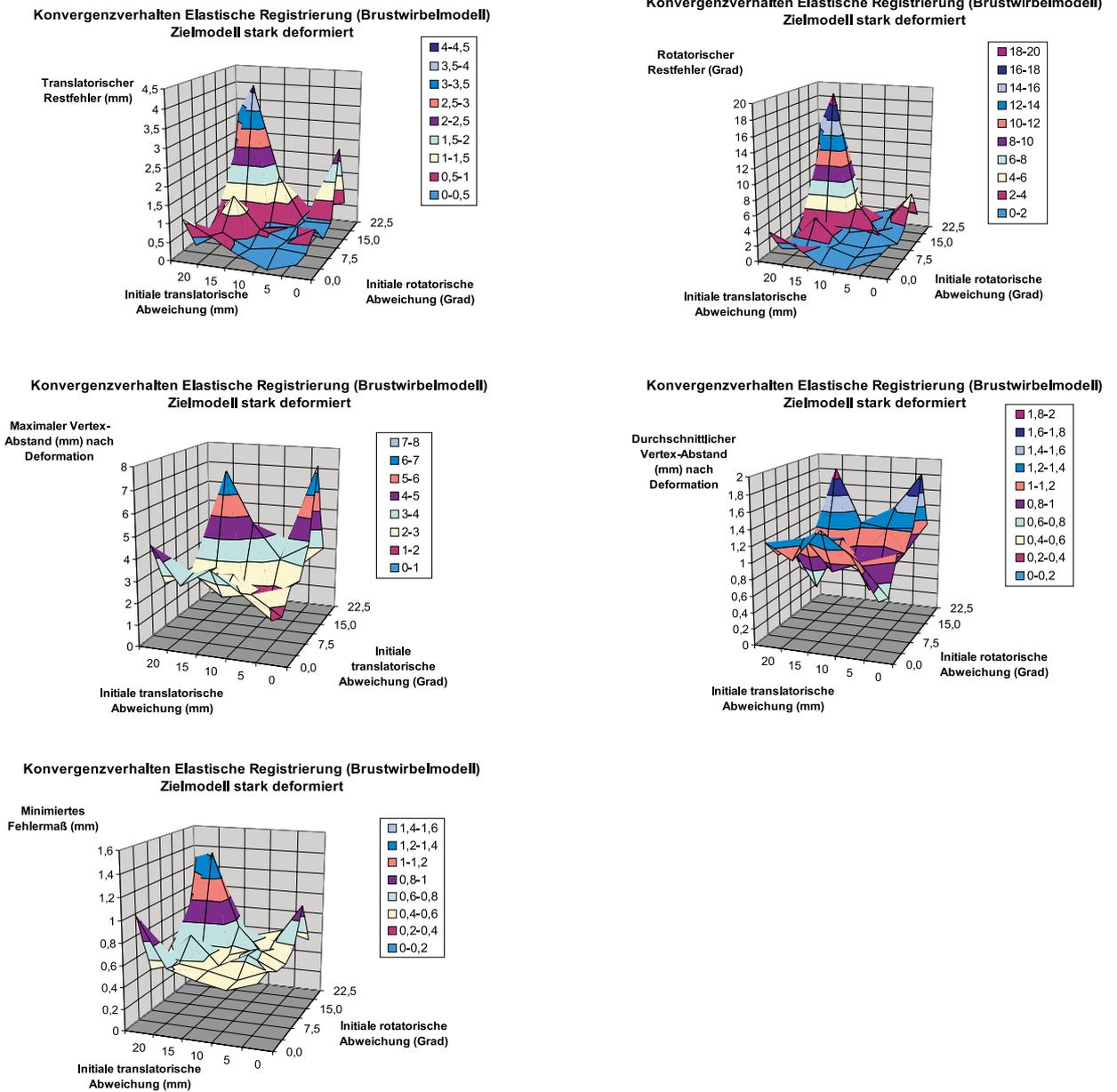


Abbildung 6.22: Konvergenzverhalten bei simulierten Testdaten (zwei Aufnahmen eines stark deformierten Wirbels, Winkelabstand der Blickrichtungen 90° (lateral, AP)) für ein formvariables Brustwirbelmodell. Angetragen sind jeweils translatorische und rotatorische Restfehler, maximaler und durchschnittlicher Deformationsfehler (Vertex-Abstand) sowie das minimierte Fehlermaß (mittlerer 3D-Abstand Objekt/Projektionslinien) nach erfolgter Konvergenz der Optimierung, ausgehend von unterschiedlichen initialen translatorischen und rotatorischen Abweichungen.

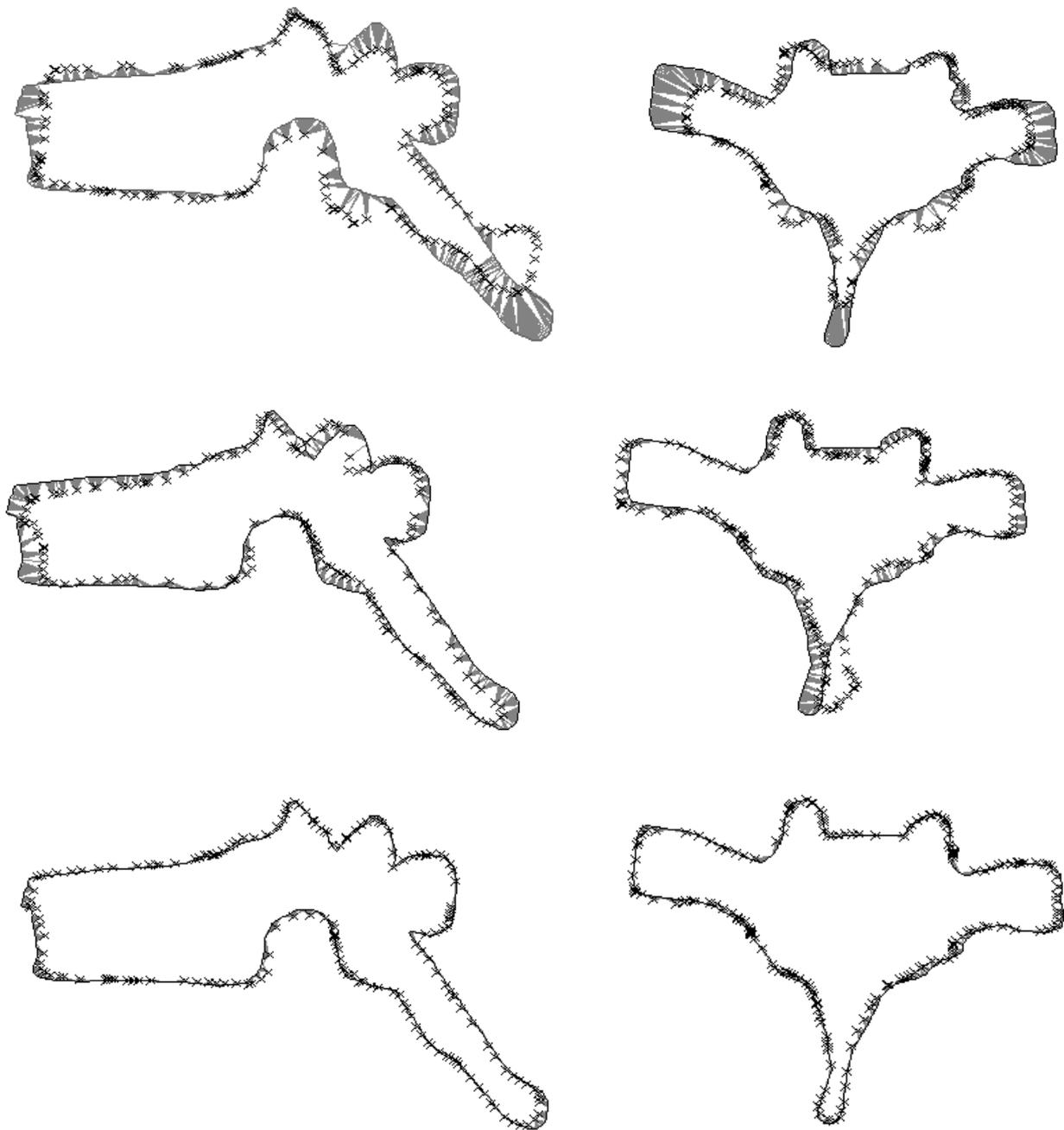


Abbildung 6.23: Anpassung der Deformationsparameter des formvariablen 3D-Brustwirbelmodells an zwei vorgegebene Konturen (lateral, AP) eines *stark verformten* Wirbels: Initialisierung (oben), Zwischenergebnis (Mitte), Konvergenz (unten). Graue Linien veranschaulichen Korrespondenzen zwischen Konturpunkten und Projektionen der Modellvertices (Kreuzsymbole).

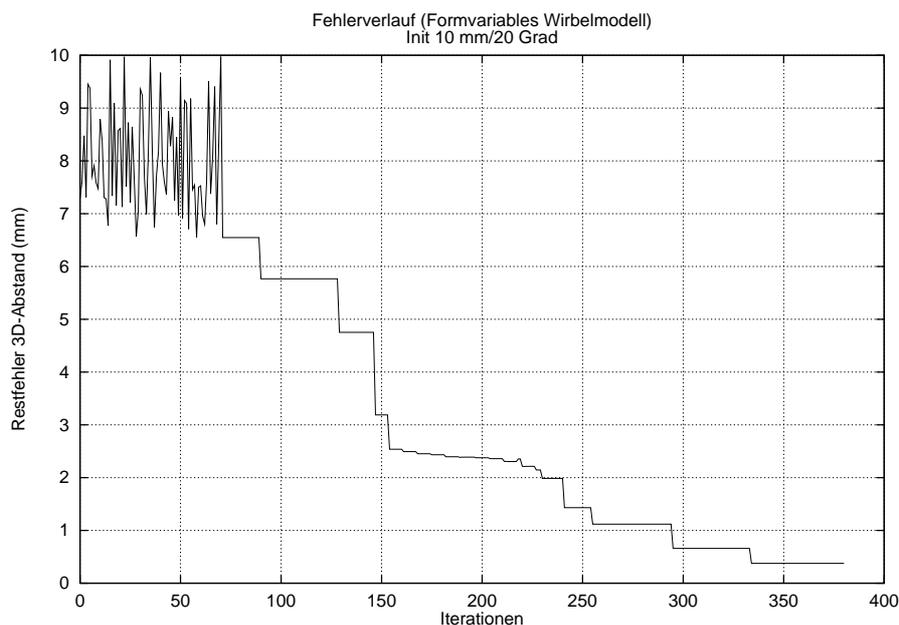


Abbildung 6.24: Beispielhafter Fehlerverlauf während der mehrstufigen Optimierung: (1) randomisiert (Iteration 0-67), (2) Levenberg-Marquardt (rigide, Iteration 68-140), (3) Levenberg-Marquardt (Deformation, Iteration 141-237), (4) Levenberg-Marquardt (rigide und Deformation, Iteration 238-380). Angetragen ist jeweils der minimalste (2-4) bzw. aktuelle Fehlerwert (1).

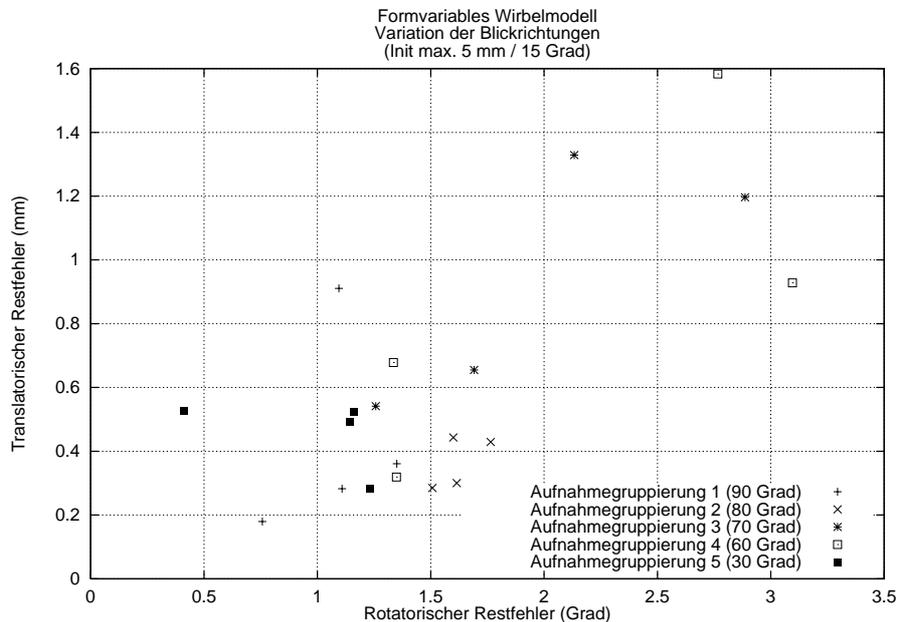


Abbildung 6.25: Restfehler-Streuung (Rotatorischer vs. translatorischer Restfehler) bei Verwendung unterschiedlicher Aufnahmerichtungen sowie mehrere Initialisierungen bis maximal 5 mm bzw. 15° .

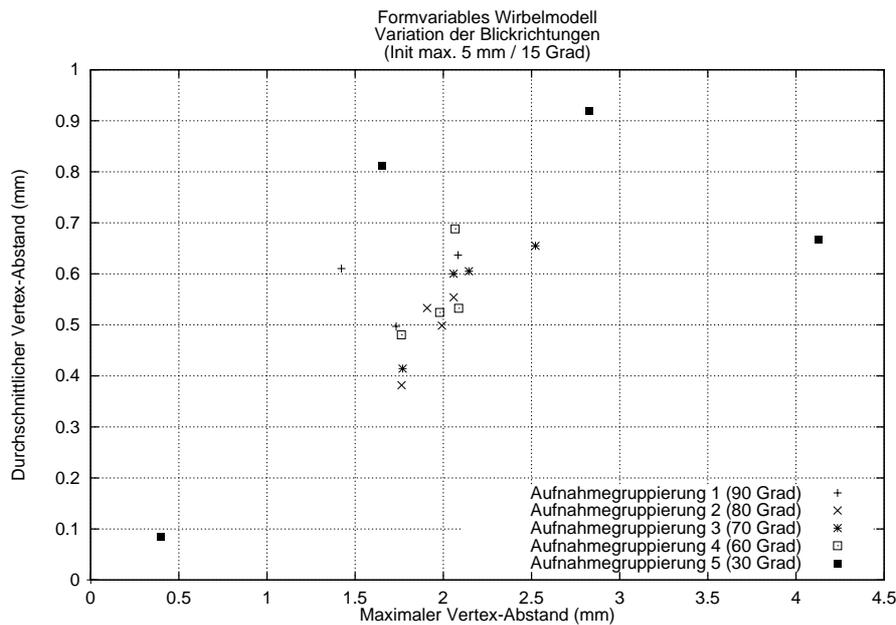


Abbildung 6.26: Restfehler-Streuung (Mittlerer vs. maximaler Deformationsfehler (Vertex-Abstand)) bei Verwendung unterschiedlicher Aufnahmerichtungen sowie mehrere Initialisierungen bis maximal 5mm bzw. 15°.

spiel demonstriert. Abb. 6.27 faßt das Ergebnis eines Experimentes mit realen Fluoroskopieaufnahmen eines kalibrierten C-Bogens des Typs *SIEMENS Siremobil 2* zusammen. Die Aufnahmen zeigen ein Wirbelsäulenphantom, wobei das bereits zuvor verwendete formvariable Wirbelmodell an den oberhalb der Bildmitte gelegenen T11-Wirbel angepaßt werden soll. Die Außenkonturen dieses Wirbels (rot) wurden in beiden Bildern bereits interaktiv segmentiert. Wie sich bei einem Vergleich mit den direkt benachbarten Wirbeln zeigt, ist die Segmentierung des rechten Bildes jedoch fehlerhaft. Insbesondere einer der kaum erkennbaren unteren Gelenkfortsätze (blauer Rahmen) wurde übersehen. Nichtsdestotrotz gelingt es der Optimierung, das formvariable Wirbelmodell gut an die korrekten Anteile der segmentierten Konturen anzupassen. Das Ergebnis der Rückprojektion des registrierten Modells liefert im übrigen Anhaltspunkte für weitere Segmentationsfehler bzgl. des zweiten unteren sowie beider oberen Gelenkfortsätze. Nicht zuletzt aufgrund der falschen Konturfragmente bleibt ein durchschnittlicher Restfehler pro Modellvertex und Projektionslinie von knapp 1.4mm bestehen [Dö99].

Abb. 6.28 zeigt Fluoroskopieaufnahmen eines proximalen Femurendes, die mit demselben C-Bogen-Gerät gewonnen wurden. In diesem Fall wurden die Konturen fehlerfrei segmentiert, so daß ein formvariables Femurmodell erfolgreich angepaßt werden konnte. Der durchschnittliche Restfehler pro zugeordnetem Vertex betrug nach der Optimierung 0.6mm.

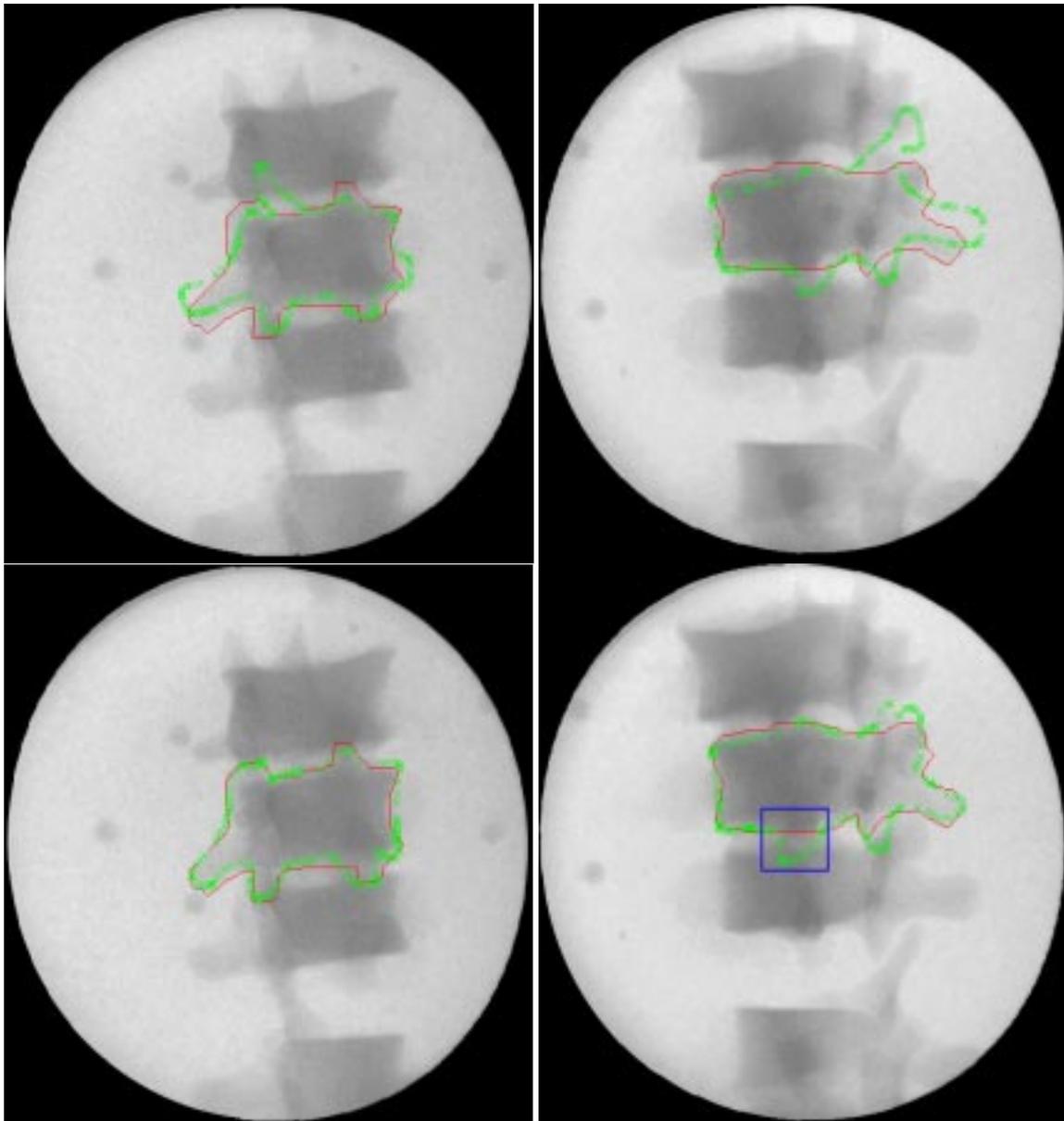


Abbildung 6.27: Elastische Registrierung mit realen Fluoroskopieaufnahmen eines Wirbelsäulenphantoms: Initialisierung (oben), Endergebnis (unten). Die blau markierte Stelle deutet auf einen Fehler aus der Kontursegmentierung hin.

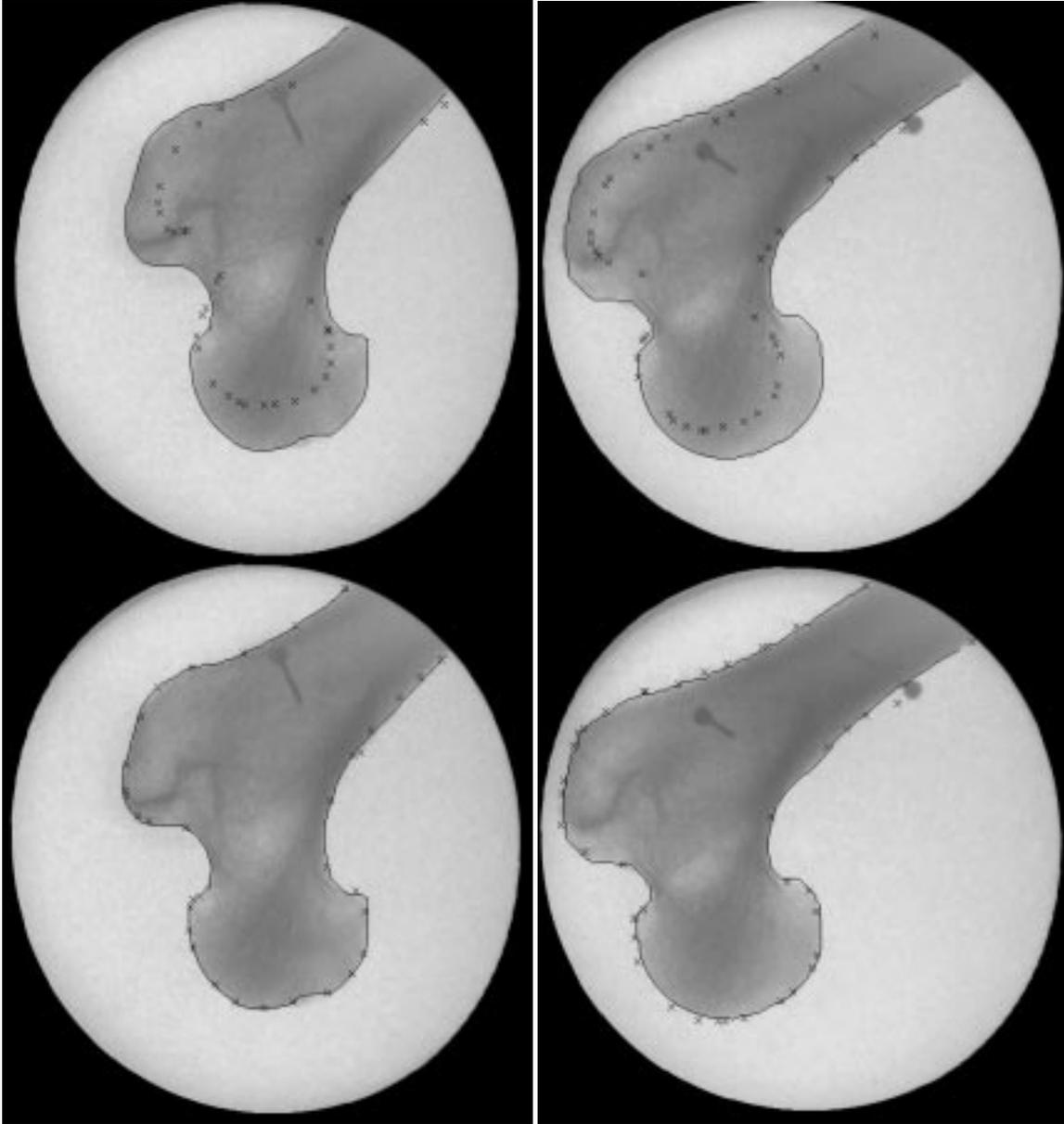


Abbildung 6.28: Elastische Registrierung mit realen Fluoroskopieaufnahmen eines proximalen Femurknochens: Initialisierung (oben), Endergebnis (unten)

Kapitel 7

Experimente in klinischer Umgebung

In den Abschnitten 5.5 und 6.6 wurden bereits experimentelle Ergebnisse präsentiert, die eine prinzipielle Eignung der beiden in dieser Arbeit neu entwickelten Registrierungstechniken demonstrieren. Nichtsdestotrotz ergeben sich in einer realen OP-Umgebung oftmals zusätzliche Fragestellungen und Randbedingungen. Beispielsweise stellt sich die Frage, welchen Einfluß die rotatorischen und translatorischen Restfehler der Registrierung auf die Präzision nachfolgender Operationsschritte besitzen. Um hiervon einen ersten Eindruck zu bekommen, kamen die vorgestellten Verfahren zusammen mit den Arbeiten aus [Bra00a] im Frühjahr 1998 am Klinikum Rechts der Isar, München, sowie im Herbst 1999 am Institut für Werkzeugmaschinen und Betriebswissenschaften (*iwb*) der Technischen Universität München in jeweils unterschiedlich realistischen Test-Szenarien zum Einsatz.

Im ersten Fall bestand das Ziel darin, präoperativ an einem anatomischen Kniepräparat angebrachte Bohrlöcher intraoperativ präzise zu lokalisieren. Im zweiten Fall wurde das distale Ende eines Femur-Phantoms durch mehrere robotergestützte Sägeschnitte präpariert, um anschließend ein entsprechendes Implantat anzubringen [Gö00]. Die Ergebnisse der beiden vorklinischen Versuche werden nachfolgend knapp zusammengefaßt.

7.1 Vorklinischer Versuch (Klinikum Rechts der Isar)

Im Frühjahr 1998 fand im Münchener Klinikum Rechts der Isar eine erste vorklinische Validierung der in [Bra00a] und der vorliegenden Dissertation erarbeiteten Konzepte und Verfahren statt. Unter engagierter Mitwirkung von Herrn Dr. Rainer Burgkart¹ gelang es, ein anatomisches Präparat (menschliches Leichenknie) fluoroskopiebasiert und ohne Verwendung künstlicher Marken zu registrieren. Im Vergleich zu den experimentellen Untersuchungen in Abschnitt 5.5.2 ergab sich bei diesem vorklinischen Versuch die zusätzliche Randbedingung einer realistischen Bildqualität durch das die Knochen umgebende Gewebe.

Abb. 7.1 zeigt die wesentlichen Stationen während des Versuchs. Präoperativ wurde ein Computer-Tomographie-Datensatz für das Präparat erstellt sowie eine Kalibrierung des verwendeten Fluoroskops vorgenommen (Anhang B). Intraoperativ erfolgte die Akquisition mehrerer Fluoroskopieaufnahmen. Nach Durchführung der intensitätsbasierten Registrierung aus Kapitel 6 ließen sich die präoperativ gesetzten Bohrlöcher mit Hilfe eines optisch überwachten Zeigerinstrumentes (siehe Abschnitt 3.3) bis auf

¹Klinik für Orthopädie und Sportorthopädie, Klinikum Rechts der Isar.

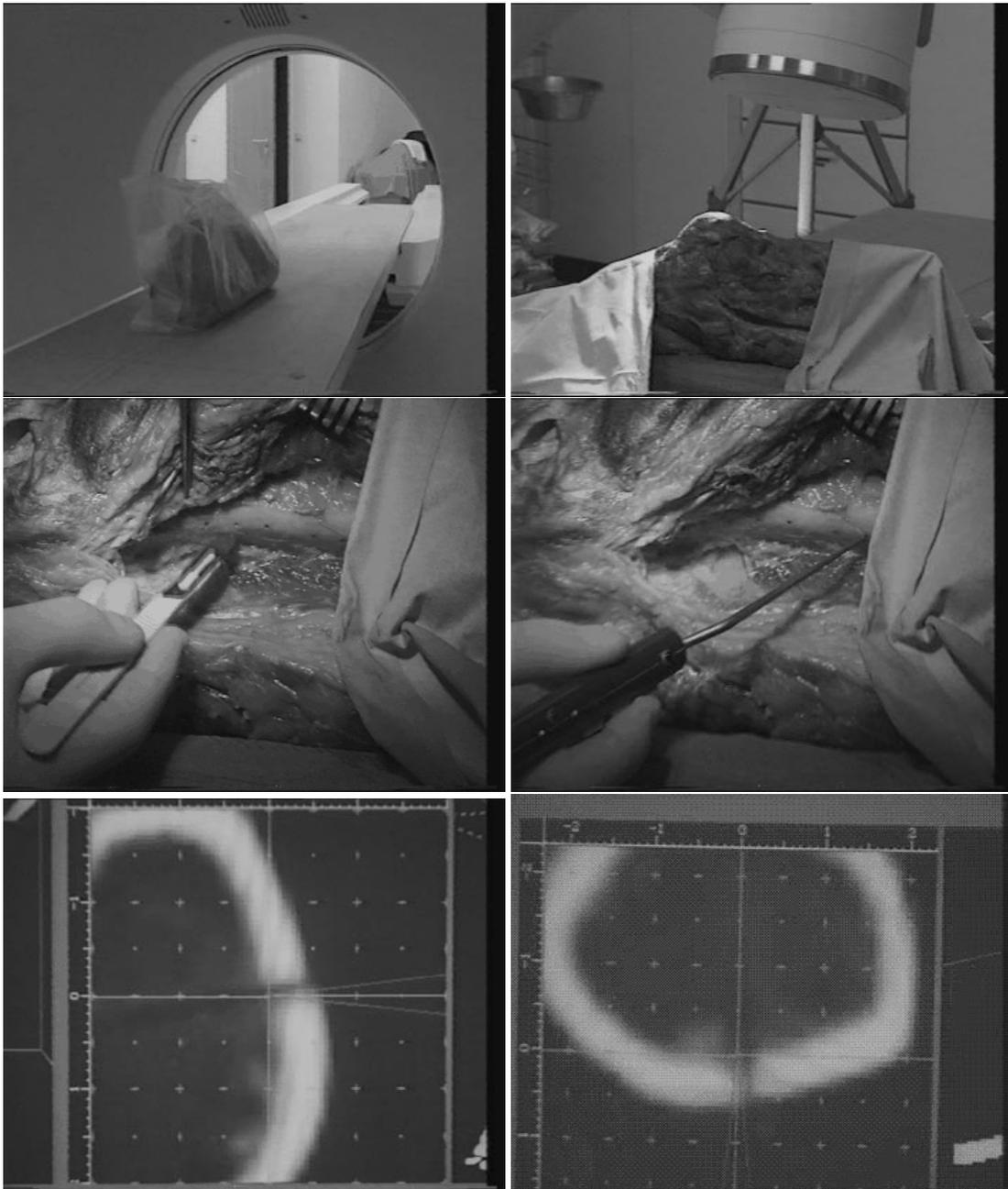


Abbildung 7.1: Momentaufnahmen einer Videodokumentation des am Klinikum Rechts der Isar durchgeführten vorklinischen Versuchs (von links oben nach rechts unten): (a) Computer-Tomographie-Aufnahme, (b) Akquisition mehrerer Fluoroskopieaufnahmen, (c) Freilegung des präparierten Femurschafts, (d) Navigation mit Hilfe eines optisch überwachten Zeigers, (e,f) Echtzeit-Visualisierung der aktuellen Zeigerposition innerhalb des CT-Datensatzes.

etwa einen Millimeter genau wiederfinden. Im Vergleich zu der Transformation, die mittels einer rigiden, konturbasierten Registrierung aus [Bra00a] bestimmt wurde, ergaben sich translatorische und rotatorische Abweichungen von 1.4mm und 0.6° . Da in diesem Versuchsszenario die optimale Transformation nicht bekannt war, dürfen diese Angaben nicht als Restfehler interpretiert werden. Die optische Kontrolle der navigierten Zeigerspitze innerhalb der visualisierten CT-Schichtaufnahmen vermittelt jedoch insgesamt ein überzeugendes Ergebnis. Das elastische Registrierungsverfahren aus Kapitel 6 wurde erst nach 1998 entwickelt und konnte deshalb zum damaligen Zeitpunkt leider noch nicht zum Einsatz kommen.

7.2 Robotergestützter Kniegelenksersatz (Phantomstudie)

Im Herbst 1999 wurde am Institut für Werkzeugmaschinen und Betriebswissenschaften (*iwb*) der Technischen Universität München in Garching eine Phantomstudie durchgeführt. Das Ziel bestand in der Validierung eines kompletten Prozeßdurchlaufs der im Rahmen des interdisziplinären Gemeinschaftsprojektes CARS² (*Computer-Aided Robotic Surgery*) entwickelten Verfahren für einen robotergestützten Kniegelenksersatz (TKR, vgl. Abschnitt 2.1). Das Konzept für die Patientenregistrierung sieht dabei präoperativ eine Computer-Tomographie des Patienten vor, auf deren Grundlage 3D-Modelle der involvierten Knochen generiert werden. Intraoperativ erfolgt eine konturbasierte Registrierung mit Hilfe eines kalibrierten Fluoroskops. Nichtsdestotrotz konnten beide Ansätze der vorliegenden Dissertation in diesem Umfeld ebenfalls erfolgreich getestet werden.

Grundlage der Phantomstudie war ein Femurphantom aus speziellem Kunststoff. Um quantitative Aussagen über die Genauigkeit der durchgeführten Arbeitsschritte treffen zu können, wurde das Knochenphantom distal, proximal und in der Mitte des Femurschaftes mit je einer transversal ausgerichteten Plexiglasscheibe präpariert. Auf diesen Scheiben befanden sich jeweils mehrere Referenzkugeln aus Teflon, die sich sowohl in den CT-Daten detektieren als auch mit Hilfe einer hochpräzisen Koordinatenmeßmaschine taktil vermessen ließen (vgl. hierzu Abschnitt 2.3.3). Das Femurphantom wurde zusätzlich für die Dauer der Registrierung auf einer Holzplatte starr verklebt.

Die Festlegung eines globalen Referenzkoordinatensystems erfolgte in diesem Experiment durch einen robotergeführten Kalibrierkörper, der für jede Fluoroskopieaufnahme durch manuelle Bewegung des Roboters neu im Strahlengang des C-Bogens – zwischen Femurphantom und Detektor des Bildverstärkers – positioniert wurde (Abb. 7.2). Die dabei ausgeführten Relativbewegungen des Roboters bezüglich einer definierten Nullstellung des Kalibrierkörpers wurden protokolliert, um diese im Rahmen der intraoperativen Kalibrierung und Lokalisation des Fluoroskops entsprechend zu berücksichtigen [Bra00a] (Anhang B). Nach Aufnahme von insgesamt fünf Fluoroskopieaufnahmen positionierte der Roboter den Kalibrierkörper wieder in Nullstellung, woraufhin dieser ebenfalls starr mit der Holzplatte verklebt und von der Roboterhand abgekoppelt wurde. Die beschriebene Vorgehensweise erlaubte es, eine optimale rigide Transformation unabhängig von den zu validierenden fluoroskopiebasierten Ansätzen zu bestimmen [HS93a]. Hierzu war es lediglich nötig, die Marken des festgeklebten Kalibrierkörpers zusammen mit den erwähnten Teflonkugeln taktil zu vermessen.

Ein erster Versuch im Rahmen des beschriebenen Szenarios sollte die Genauigkeit des in Kapitel 5 vorgestellten intensitätsbasierten Registrierungsverfahrens ermitteln. Zwei geänderte Randbedingungen im Vergleich zu den experimentellen Tests aus Abschnitt 5.5.2 ergaben sich: Zum einen enthiel-

²Projektpartner: Bayerisches Forschungszentrum für wissenschaftliche Systeme (FORWISS), Institut für Werkzeugmaschinen und Betriebswissenschaften (*iwb*) der TU München, Medizinische Hochschule Hannover (MHH)

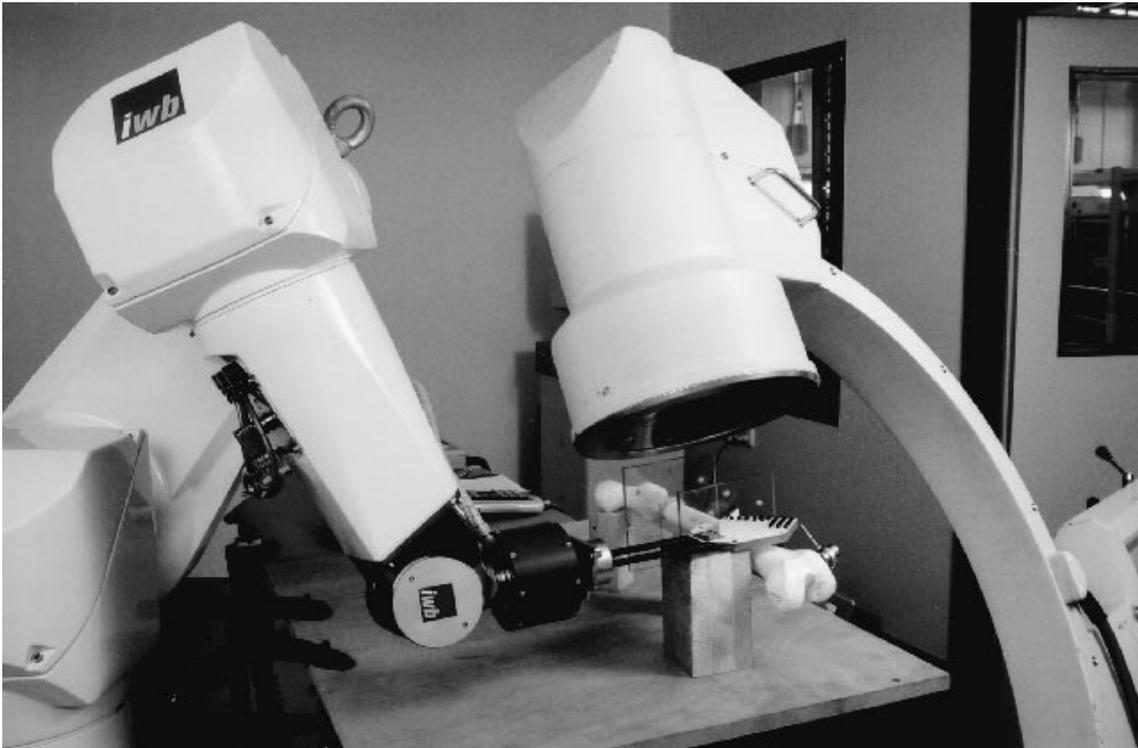


Abbildung 7.2: Versuchsaufbau der Phantomstudie [Gö00]

ten die Fluoroskopieaufnahmen großflächige Artefaktregionen aufgrund der verwendeten Femurfixation [Gö00], zum anderen beeinflusste die begrenzte Genauigkeit des Roboters die Ergebnisse der Kalibrierung. Im Vergleich zu der auf Basis der Metall- und Teflonkugeln vermessenen Transformation ergaben sich letztlich akzeptable Restfehler von 1.3mm bzw. 0.95° .

In einem zweiten Versuch fand die konturbasierte, elastische Registrierung aus Kapitel 6 Anwendung. Hierzu wurden zunächst interaktiv die Femurkonturen in den einzelnen Aufnahmen segmentiert (Abschnitt 6.5.1). Anschließend erfolgte eine einmalige rigide Anpassung des aus den vorhandenen CT-Daten segmentierten 3D-Modell an die Konturen, d.h. unter Deaktivierung der Skalierungs- und Verformungsparameter. Anschließend wurde zum Vergleich ein formvariables Femurmodell an die Konturen angepaßt. Im ersten Fall ergaben sich rigide Restfehler von 1.48mm bzw. 1.27° bei einem durchschnittlichen Abstand von 0.73mm zwischen Projektionslinien und Objektvertices. Im zweiten Fall hingegen ergaben sich translatorische und rotatorische Restfehler von 1.32mm bzw. 0.9° . Der durchschnittliche Abstand betrug 0.98mm . Die in Abschnitt 6.6.1 eingeführten Deformations- bzw. Formfehler ergaben 0.8mm durchschnittlichen bzw. 5.6mm maximalen Vertexabstand. Der relativ hohe maximale Formfehler läßt sich dadurch begründen, daß das formvariable Femurmodell aus nur vier Trainingsdatensätzen, die den Phantomfemur selbst nicht beinhalteten, aufgebaut worden war.

Neben der Validierung der Registrierung stand im Rahmen des Experimentes insbesondere die Genauigkeit der robotergestützten Femurbearbeitung für den Implantateinsatz im Mittelpunkt des Interesses [Gö00]. Abb. 7.3 zeigt links die im Vergleich zur Registrierung geänderte Kniestellung als Vorbereitung der durchzuführenden Sägeschnitte. Die Position des Femurs wird dabei durch ein optisches Trackings-

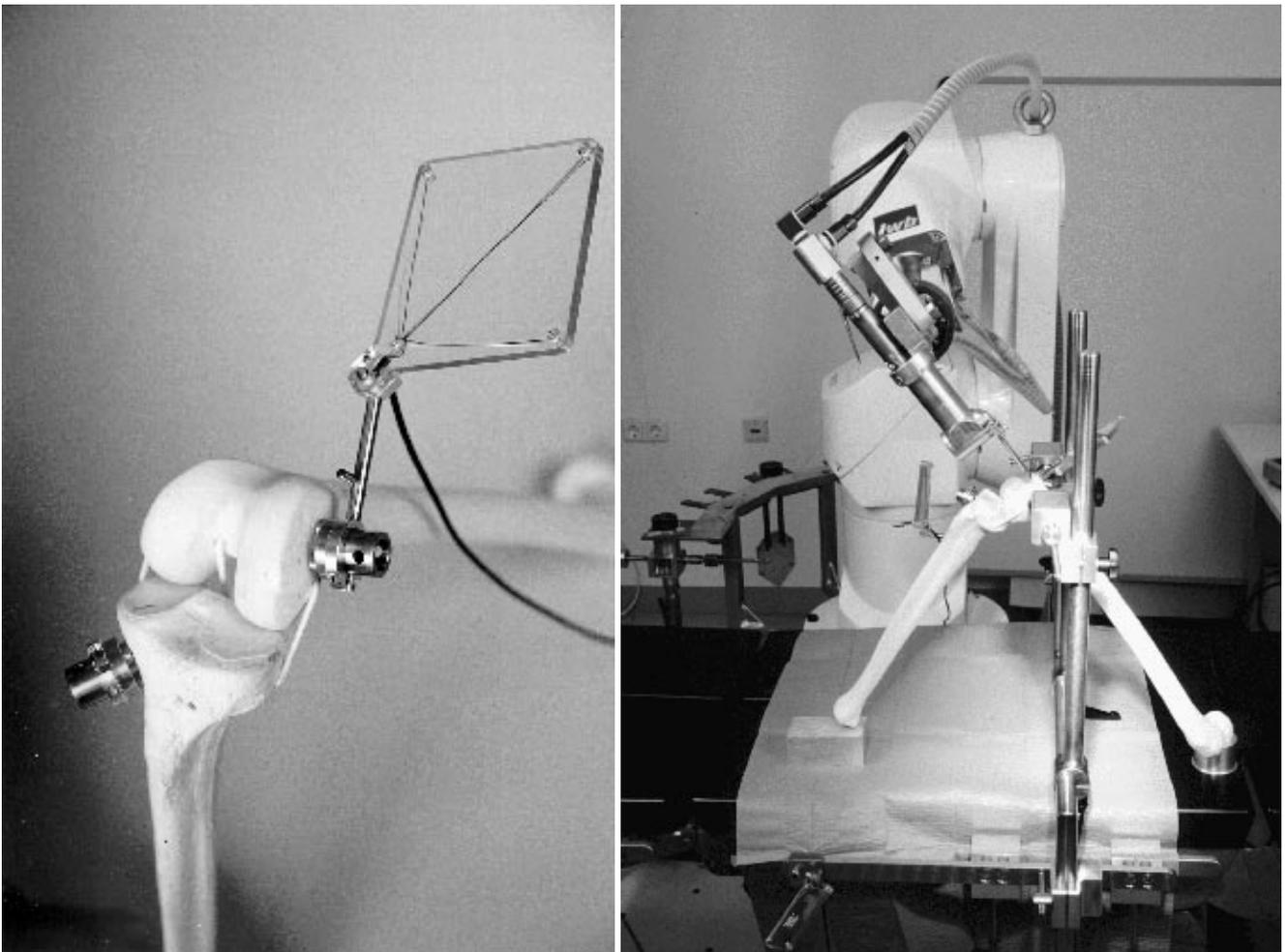


Abbildung 7.3: Links: Optisch überwachte Kniestellung, rechts: Robotergestützte Präparation des fixierten distalen Femurs [Gö00]

system überwacht. Das rechte Bild zeigt das in Beugstellung fixierte Knie sowie das durch den Roboter vorpositionierte Sägewerkzeug.

Die Anpassung des Implantats ergab nach Durchführung aller präoperativ auf Basis der CT-Daten geplanten Sägeschnitte rigide Abweichungen im Vergleich zu einer definierten Sollposition von 2.16mm bzw. 1.9° [Gö00]. Unter Berücksichtigung der gesamten Prozeßkette und insbesondere des verwendeten Roboters, der eine inakzeptable Absolutpositioniergenauigkeit aufwies, erscheint dieses Ergebnis durchaus akzeptabel.

Kapitel 8

Zusammenfassung

Mobile Röntgenbildverstärker stellen bereits seit langem den Standard intraoperativer Bildgewinnung dar. Jedoch erst seit einigen Jahren besteht verstärkt das Interesse, diese Fluoroskopie-Geräte auch in computerunterstützte Operationssysteme zu integrieren, nachdem die bekannten gerätespezifischen Probleme hinsichtlich einer Sensorkalibrierung weitestgehend gelöst zu sein scheinen.

Im Mittelpunkt der vorliegenden Arbeit stand die Entwicklung und Realisierung geeigneter Konzepte zur präzisen fluoroskopiebasierten Registrierung eines Patienten. Da sich in Röntgenbildern insbesondere Knochenstrukturen gut abbilden lassen, wurde die Arbeit speziell auf orthopädische Anwendungen im Bereich der Wirbelsäule sowie des Hüft- und Kniegelenks ausgerichtet.

Der erste Ansatz für eine fluoroskopiebasierte Registrierung setzt voraus, daß ein patientenspezifischer Computer-Tomographie-Datensatz der betrachteten Knochenstruktur vorliegt. Auf Basis dieser Volumendaten werden intraoperativ simulierte Röntgenbilder generiert und mit mehreren zuvor aufgenommenen Fluoroskopieaufnahmen aus unterschiedlicher Richtung verglichen. Eine geeignete Optimierung der Übereinstimmung zwischen korrespondierenden Aufnahmen liefert schließlich einen präzisen Rückschluß auf die Lage des betrachteten Knochens. Wesentliche Beiträge zu zentralen Problemstellungen in diesem Kontext beschäftigten sich mit einer adaptiven Kontrastanpassung, der Berücksichtigung von Bildartefakten, der robusten Formulierung der Optimierungsproblematik sowie der beschleunigten Simulation von Röntgenaufnahmen.

Der zweite Registrierungsansatz benötigt keine präoperativen Patientendaten, ohne dabei jedoch auf die Vorteile dreidimensionaler Patienteninformationen verzichten zu müssen. Er eignet sich daher besonders gut für Situationen, in denen keine Zeit für eine aufwendige Tomographie zur Verfügung steht. Ein Beispiel hierfür ist die spontane Notwendigkeit einer präzisen Patientenregistrierung bei komplexen Notfall-Operationen. Das vorgeschlagene Verfahren bedient sich eines formvariablen 3D-Modells, um die bei Betrachtung einer Vielzahl von Patienten zu erwartenden anatomischen Formausprägungen eines bestimmten Knochens *patientenunabhängig* zu modellieren. Eine für diesen Zweck geeignete Modellierung von Formvariationen bildete hierbei die zentrale Problemstellung. Vorgeschlagen wird die Kombination von statistischer Modellierung auf Basis einer Stichprobe repräsentativer Knochenformen und physikalisch motivierter Modellierung. Die Registrierung erfolgt durch intraoperative Anpassung von Position, Orientierung, Skalierung und Form eines solches Modells an die – in den einzelnen Fluoroskopieaufnahmen segmentierten – Konturen bzw. Silhouetten des betrachteten Knochens.

Die Leistungsfähigkeit der beiden neu entwickelten Registrierungstechniken konnte sowohl in der

Simulation, in realitätsnahen Phantomstudien sowie im Rahmen eines ersten vorklinischen Experimentes am Münchener Klinikum Rechts der Isar nachgewiesen werden.

Anhang A

Anatomische Grundlagen

In der vorliegenden Arbeit spielen verschiedene menschliche Knochenstrukturen eine zentrale Rolle. Dabei stehen einerseits Wirbel der Brust- und Lendenwirbelsäule und andererseits die am Aufbau des Kniegelenkes beteiligten Knochen (Oberschenkelknochen, Waden- und Schienbein) im Mittelpunkt des Interesses. Dieser Anhang vermittelt in knapper Form die wichtigsten anatomischen Bezeichnungen hinsichtlich der genannten Knochen. Für ein intensiveres Studium der menschlichen Anatomie sei auf [Leo91, Lip99] verwiesen.

Das Kniegelenk

Das Kniegelenk ist das größte, komplizierteste und vermutlich auch das verletzungsanfälligste Gelenk des menschlichen Körpers. Es ist ein komplexes Drehgelenk und erlaubt neben der Beugung (Flexion) und Streckung (Extension) auch eine eingeschränkte Gleit- und Drehbewegung. Der obere Teil des Gelenks wird durch das untere Ende des Oberschenkelknochens (distaler Femur), der untere Teil entsprechend durch das obere Ende des Schienbeines (proximale Tibia) gebildet. Das distale Femurende besteht aus zwei abgerundeten Flächen, den Kondylen. Dazwischen befindet sich eine grubenförmige Vertiefung, die intrakondyläre Fossa. In dieser Grube sind die beiden Kreuzbänder verankert. Deren Hauptaufgabe liegt in einer rotatorischen Stabilisierung des Kniegelenks. Am äußeren Rand der Kondylen liegen die sogenannten Epikondylen, die als Muskel- und Bandansatz dienen. Das gesamte Knie wird durch einen komplexen Bandapparat zusammengehalten und dient zugleich als Ansatz für einige der kräftigsten Muskeln des menschlichen Körpers.

Femur

Der Oberschenkelknochen (Femur) ist der längste Knochen des menschlichen Körpers (Abb. A.1). Das untere, distale Ende bildet zusammen mit dem oberen, proximalen Ende der Tibia das Kniegelenk. Das obere, proximale Femurende – der sogenannte Femurkopf – ist kugelförmig ausgeprägt. Dieser Femurkopf paßt sich direkt in eine Aushöhlung im Becken ein und bildet zusammen mit dieser das Hüftgelenk. Direkt unterhalb des Hüftkopfes verdünnt sich der Knochen rasch. Diese Stelle wird Femurhals genannt. In unmittelbarer Nähe des Femurhalses befinden sich desweiteren zwei markante Fortsätze: der große und kleine Trochanter (Rollbügel, *Trochanter major/minor*). Diese beiden Fortsätze dienen anatomisch unter anderem als Ansatzpunkte für Oberschenkel- und Gesäßmuskeln. Der Femurhals selbst sorgt einerseits für eine große Beweglichkeit des Hüftgelenks, ist andererseits aber zugleich ein Schwachpunkt

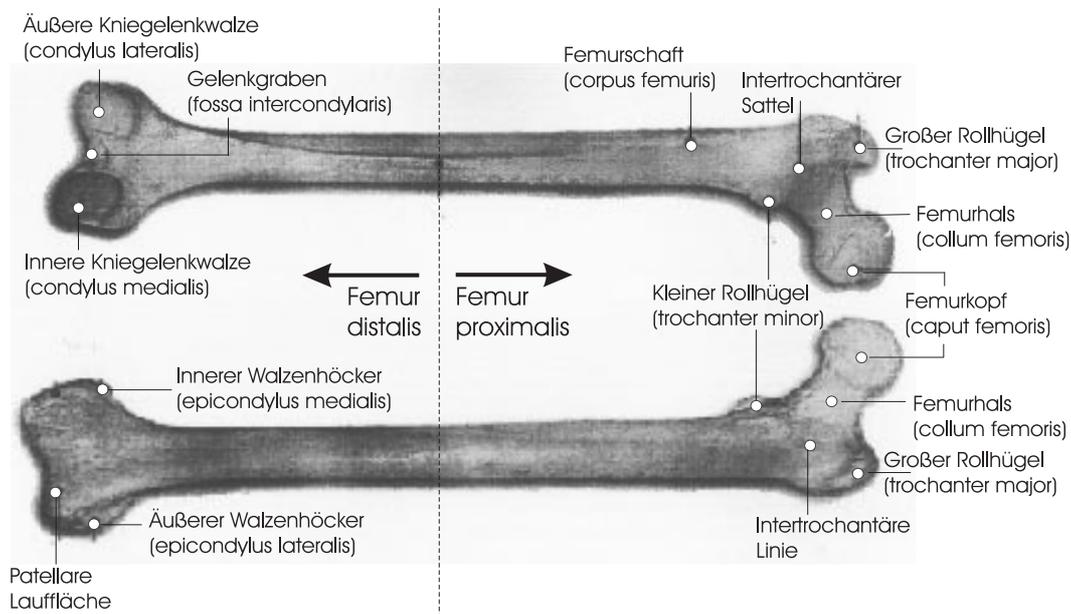


Abbildung A.1: Anatomie des Oberschenkels (Femur)

und damit sehr häufig von Frakturen betroffen (Oberschenkelhalsbruch).

Tibia

Die Tibia (Schienbein) ist der innen verlaufende, dickere Knochen der beiden langen Knochen des unteren Beines. Das obere Ende der Tibia unterteilt sich in zwei Kondylen (lateral und medial). Die Oberfläche dieser Kondylen ist konkav gewölbt und bildet dadurch das Gegenstück zu den beiden konvexen Femurkondylen. Getrennt werden die Kondylen durch die sogenannte intrakondyläre Eminenz, einen kammartigen Vorsprung. Die Tibia trägt die größte Last des unteren Beines und verläuft parallel zum Wadenbein (Fibula), dem zweiten, kleineren Knochen. Beide Knochen sind über Bänder miteinander verbunden. Der vordere Teil der Tibia läßt sich leicht durch die Haut ertasten. Das obere Ende bildet zusammen mit dem distalen Femur das Kniegelenk.

Fibula

Die Fibula (Wadenbein) ist der außen verlaufende, kleinere Knochen der beiden langen Knochen des unteren Beines. Er ist dabei wesentlich dünner als die parallel verlaufende Tibia (Schienbein) und mit dieser an beiden Enden über Bänder verbunden. Das obere Ende (Fibulakopf) erstreckt sich nicht bis in den Bereich des Kniegelenks, sondern endet unterhalb der lateralen Tibiakondyle. Das untere Ende hingegen liegt im Vergleich zur Tibia tiefer und ist in das Sprunggelenk integriert. Die Hauptaufgabe der Fibula liegt in der Bereitstellung von Muskelansatzpunkten. Im Gegensatz zur Tibia dient die Fibula weniger der Belastung und Stabilisierung des unteren Beines. Dies erklärt u.a., warum das Wadenbein häufig als Material für Knochentransplantationen herangezogen wird.

Die Wirbelsäule

Die Wirbelsäule (Columna vertebralis) stellt die Grundlage des menschlichen Stammes dar. Sie besteht aus 33 bis 34 einzelnen Knochenstücken, den Wirbeln (vertebrae) sowie jeweils dazwischenliegenden Bandscheiben (Disci intervertebrales). Die gesamte Wirbelsäule erstreckt sich von der Schädelbasis bis hinunter zum Becken. Die einzelnen Wirbel umschließen jeweils den spinalen Nervenstrang und schützen auf diese Weise dieses empfindliche System, dessen Nervenbahnen jede erdenkliche Körperregion mit dem Gehirn im Schädel verbinden. Je zwei benachbarte Wirbel sind über ein Gelenk miteinander verbunden, was einerseits die Stabilisierung des gesamten Oberkörpers sicherstellt, andererseits aber dessen freie Beweglichkeit erlaubt. Zwischen je zwei Wirbeln befindet sich eine der erwähnten Bandscheiben, bestehend aus fasrigem Knorpelmaterial mit einem geleeartigen Kern. Diese Scheiben dienen bei einer Bewegung und Belastung der Wirbelsäule als druckelastische Polster. Die einzelnen Wirbel sind über zwei dicke Bänder, die sich über die gesamte Länge der Wirbelsäule erstrecken, sowie eine Vielzahl kleinerer Bänder zwischen je zwei benachbarten Wirbeln miteinander verbunden. Desweiteren setzt direkt an den Wirbeln eine Vielzahl an Muskeln an, die zum einen die Bewegung der Wirbelsäule kontrollieren, zum anderen diese aber auch stabilisieren.

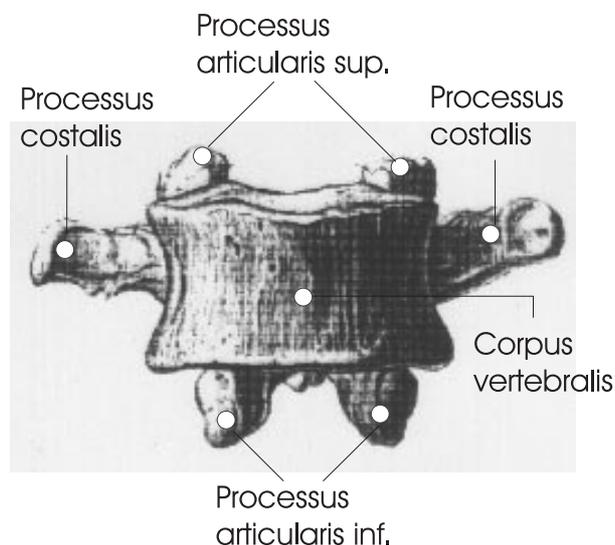


Abbildung A.2: Lendenwirbel (Ansicht von vorne)

Wirbel

Ein typischer Wirbel besitzt einen trommelförmigen Körper, der den größten Teil des gesamten Knochens ausmacht. Die Bandscheiben, die je zwei aufeinanderfolgende Wirbel voneinander trennen, sind direkt mit der rauhen und porösen Ober- bzw. Unterfläche der Wirbelkörper verwachsen. Diese Scheiben dämpfen und polstern äußere Krafteinwirkungen, die sich durch Bewegungsabläufe wie z.B. Heben, Laufen oder Springen ergeben. Ohne diese Kraftabschwächung würde eine Vielzahl einfacher Bewegungen und Belastungen bereits zu Knochenfrakturen bzw. lebensgefährlichen Verletzungen des spinalen Nervenstrangs führen. Die beiden Abbildungen A.2 und A.3 zeigen unterschiedliche Ansichten von Lenden- und Brustwirbeln. Aus dem hinteren Teil des Wirbelkörpers entwachsen in den meisten Fällen zwei kur-

ze stielartige Strukturen, die sogenannten Pedikel. Diese bilden die beiden seitlichen Begrenzungen des vollständig umschlossenen *Wirbellochs* (Foramen vertebrale). Im weiteren Verlauf der Pedikel ergibt sich auf jeder Seite eine plattenförmige Struktur (Lamina arcus vertebrae). Diese beiden Laminae vereinigen sich schließlich mittig zu dem nach hinten gerichteten Dornfortsatz (Processus spinosus). Damit bilden Pedikel, Laminae und Dornfortsatz gemeinsam das Wirbelloch, durch welches der spinale Nervenstrang verläuft. Bei einem durchschnittlichen Wirbel erstreckt sich zumeist zwischen den Pedikeln und den Laminae jeweils ein seitlich und nach hinten verlaufender Fortsatz (Processus transversus). Die Dornfortsätze der ersten neun Brustwirbel liegen – Dachziegeln gleich – geschichtet übereinander. Dadurch kommen die Enden der Dornfortsätze jeweils ein- bis eineinhalb Wirbelkörperhöhen tiefer zu liegen als der zugehörige eigene Wirbelkörper (Abb. A.3 (rechts oben)).

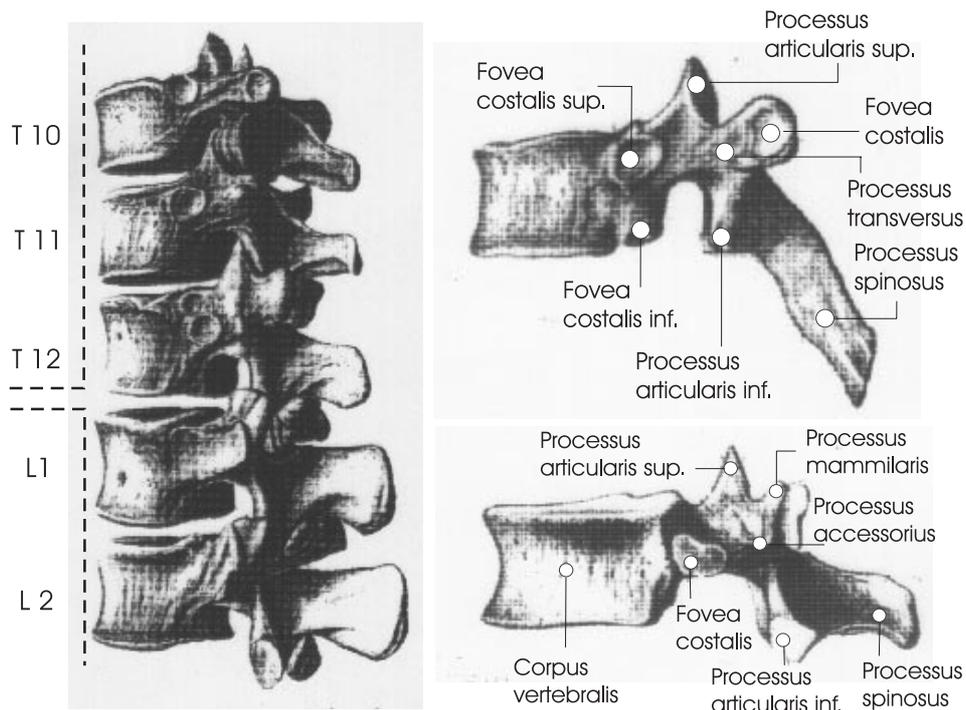


Abbildung A.3: Übergang der unteren Brustwirbelsäule T10-T12 in die Lendenwirbelsäule L1 und L2 (links), Brustwirbel T6 (rechts oben), Brustwirbel T12 (rechts unten)

Anhang B

Röntgenbildverstärker-Kalibrierung

Die zentrale Hardware-Komponente der vorliegenden Arbeit ist das in Abschnitt 3 eingeführte Röntgenbildverstärker-System (Fluoroskop). Die hiermit aufgenommenen Röntgenbilder dienen sowohl im Rahmen der intensitätsbasierten (Kapitel 5) als auch der elastischen Registrierung (Kapitel 6) als intraoperative Bildgebungsmodalität. Fluoroskopieaufnahmen sind jedoch aufgrund verschiedenster Störeinflüsse zunächst lediglich einer *qualitativen* Interpretation zugänglich. Eine für diese Arbeit wesentliche *quantitative* Verwendung, wie z.B. die Bestimmung räumlicher Information aus den Kamerabildern, wird erst nach einer geeigneten Sensorkalibrierung vor der eigentlichen Lageschätzung möglich. Die Sensorkalibrierung umfasst dabei einerseits die Bestimmung der Abbildungsparameter des Fluoroscops sowie andererseits zum jeweiligen Aufnahmezeitpunkt die Ermittlung der Lage und Orientierung des gesamten Aufnahmesystems bzgl. eines gemeinsamen Referenzkoordinatensystems.

Im Umfeld dieser Arbeit wird eine solche Sensorkalibrierung bereits als gegeben vorausgesetzt. Da die jeweilige Kalibrierungstechnik jedoch mitunter entscheidende Randbedingungen und Einschränkungen für später zum Einsatz kommende Registrierungsalgorithmen nach sich zieht, wird in diesem Anhang ein knapper Überblick über das angewandte Kalibrierungsverfahren gegeben, welches in einer eng verwandten und zeitgleich entstandenen Arbeit [Bra00a] entwickelt wurde. Konsequenterweise stehen dabei die Auswirkungen dieses Verfahrens für die in den Kapiteln 5 und 6 entwickelten Registrierungstechniken im Vordergrund. Für eine erschöpfende Beschreibung und Diskussion des Kalibrierungsvorgangs sei auf [Bra00a] verwiesen.

Problemstellung

Die in [Bra00a] vorgestellte Fluoroskop-Kalibrierung basiert auf einer idealen perspektivischen Lochkamera-Modellierung mit spezieller Berücksichtigung der nachfolgenden Eigenheiten eines C-Bogen-Systems:

- Verzeichnungseffekte:

Die mit Hilfe eines Fluoroscops erzeugten Videobilder weisen i.a. eine signifikante Verzeichnung auf. Diese Effekte besitzen zwei grundlegende Ursachen: erstens resultiert die Krümmung der in der Bildverstärkereinheit enthaltenen Photokathode in einer typischen Kissenverzerrung und zweitens entstehen ebenfalls innerhalb des Bildverstärkers *S-förmige* Verzeichnungen durch Wechselwirkungen mit externen Magnetfeldern, z.B. dem Erdmagnetfeld. Insbesondere die Auswirkungen

des letztgenannten Effekts sind nicht konstant, sondern abhängig von der jeweiligen Orientierung des C-Bogens.

- **Mechanische Instabilität:**

Bei einem Fluoroskop sind Röntgenquelle und Bildverstärker über ein halbkreisförmiges Trägergestell aus Metall fest miteinander verbunden. Das relativ hohe Gewicht des Bildverstärkers übt dabei jedoch eine enorme Kraft auf das Ende des Trägergestells aus und bewirkt so dessen mechanische Verwindung. Je nach Orientierung des C-Bogens ändern sich dabei Art und Richtung der Verwindung. Als Folge entsteht eine variable räumliche Beziehung zwischen Bildverstärker und Röntgenquelle. D.h. interne Kameraparameter, wie die Brennweite f sowie der optische Nullpunkt (c_x, c_y) sind nicht konstant, sondern variieren in Abhängigkeit von der C-Bogen-Orientierung.

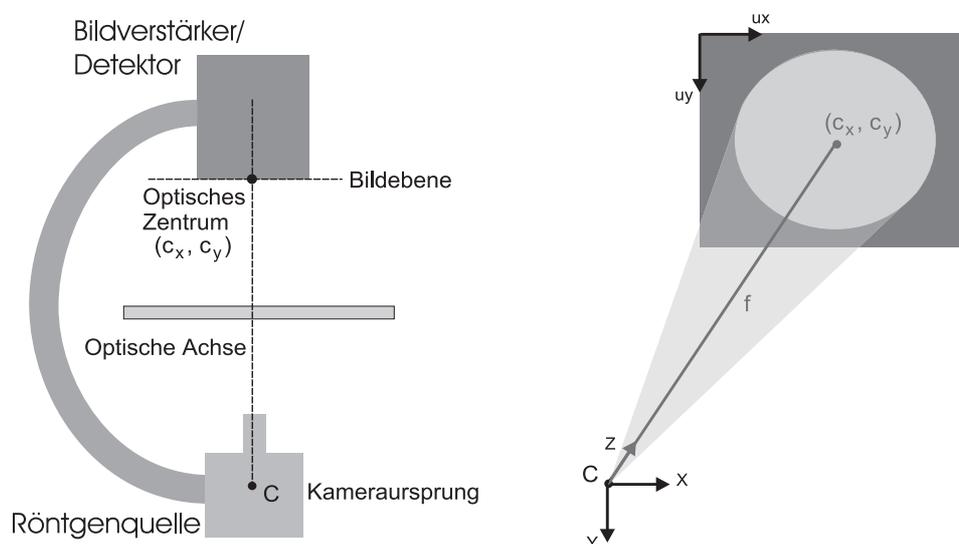


Abbildung B.1: Definition eines virtuellen Lochkameramodells

Die C-Bogen-Kalibrierung beinhaltet insgesamt drei Teilaufgaben:

- Korrektur der Verzeichnungseffekte (Entzerrung),
- Bestimmung der internen Modellparameter des Lochkameramodells (Brennweite f , optischer Nullpunkt (c_x, c_y)) unter Berücksichtigung der mechanischen Verwindung (Abb. B.1),
- Externe Kamera-Lokalisation, d.h. Bestimmung von Lage und Orientierung des mit der Röntgenquelle assoziierten Kameraursprungs.

Die einzelnen Teilaufgaben werden unter Einsatz spezieller Kalibrier- bzw. Referenzkörper gelöst. Diese Kalibrierkörper bestehen dabei jeweils aus einer exakt vermessenen Anordnung kugel- bzw. kreuzförmiger Metallmarken. Die Platzierung eines solchen Körpers im Strahlengang des C-Bogens sowie die nachfolgende Detektion der in der Fluoroskopieaufnahme sichtbaren Marken ermöglicht anschließend die Bestimmung der gesuchten Parameter.

In [Bra00a] werden im wesentlichen zwei unterschiedliche Vorgehensweisen zur Kalibrierung entwickelt, die sich in erster Linie darin unterscheiden, zu welchem Zeitpunkt – intra- oder präoperativ – Kalibrierkörper zum Einsatz kommen. Beide Ansätze werden zusammen mit den Auswirkungen auf die in der vorliegenden Arbeit vorgestellten Registrierungsverfahren im folgenden beschrieben.

Intraoperative Kalibrierung

Aufgrund der angesprochenen Orientierungsabhängigkeit von Verzerrungscharakteristik und Projektionsparametern bietet es sich an, intraoperativ das Fluoroskopiegerät für jede einzelne Aufnahme separat zu kalibrieren [Rot96, Bra96, Gos97].

Abb. B.2 veranschaulicht ein entsprechendes Szenario. Ein erster Kalibrierkörper (rechts oben) wird direkt vor dem Eingangsfenster des Bildverstärkerdetektors fixiert. Seine kreuz- und rechtecksförmigen Marken dienen ausschließlich der Bildentzerrung. Ein zweiter Kalibrierkörper (rechts unten) wird starr mit dem OP-Tisch verbunden und dient zugleich der Bestimmung von Projektionsparametern sowie als Referenzkoordinatensystem, bzgl. welchem sich die externe Kameraposition und -orientierung definiert. Beide Kalibrierkörper verbleiben bei einer Patientenaufnahme im Strahlengang, so daß sich z.B. für eine AP-Aufnahme des Knies Fluoroskopien wie in Abb. B.3 (links) gezeigt ergeben.

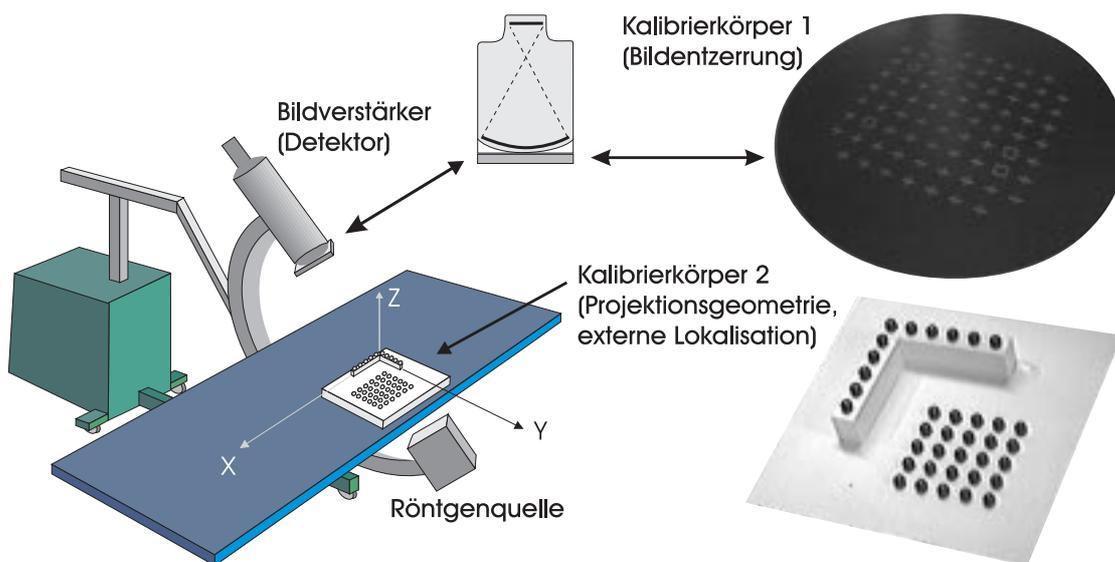


Abbildung B.2: Szenario der intraoperativen Kalibrierung

Präoperative Kalibrierung

Bei diesem Ansatz handelt es sich um ein zweistufiges Verfahren, das in der präoperativen Phase das interne Abbildungsverhalten des Fluoroskops unter Verwendung zweier Kalibrierkörper einmalig bestimmt

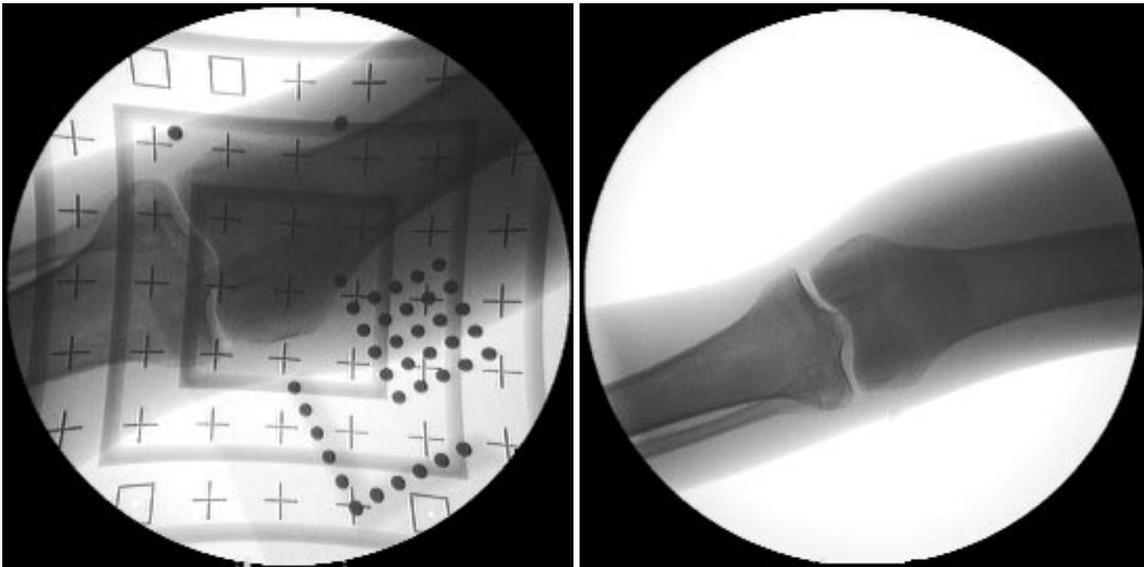


Abbildung B.3: Repräsentative Fluoroskopieaufnahmen (Kniegelenk, AP) im Falle einer intraoperativen (links) bzw. präoperativen Kalibrierung (rechts)

und in der intraoperativen Phase auf die präoperativ ermittelten Ergebnisse zurückgreift, ohne daß die dann erzeugten Aufnahmen durch Kalibriermarken beeinträchtigt werden [Bra98, Bra00b, Bra99].

Diese Kalibrierungsvariante setzt das Vorhandensein eines Infrarot-Trackingsystems voraus (vgl. Abschnitt 3.3). Letzteres dient in Verbindung mit einem am Gehäuse des Bildverstärkers angebrachten IR-LED-Phantom der Bestimmung der C-Bogen-Orientierung. Die beiden Kalibrierkörper unterscheiden sich gemäß ihrer Aufgabe. Die Modellierung der Bildverzerrung übernimmt wiederum die aus dem letzten Abschnitt bekannte Kohlefaserscheibe mit kreuzförmigen Marken (Abb. B.4). Die Bestimmung der Projektionsgeometrie hingegen erfolgt mit Hilfe eines neuen Kalibrierkörpers mit einer pyramidenförmigen Anordnung von Metallkugeln (Abb. B.5). Dieser Körper wird zu einem späteren Zeitpunkt ebenfalls direkt vor dem Detektor befestigt. Ein innerhalb des Arbeitsraumes starr angeordneter Kalibrierkörper, bzgl. dem sich die externe Kameraposition definiert, ist in diesem Szenario nicht länger nötig. Diese Aufgabe übernimmt das Infrarot-Trackingsystem.

Die präoperative Phase beinhaltet das zweimalige Abtasten des Arbeitsraumes des C-Bogens, einmal unter Zuhilfenahme von Kalibrierkörper 1 (Abb. B.4) und einmal mit Körper 3 (Abb. B.5), der ebenfalls direkt vor dem Eingangsfenster des Bildverstärker-Detektors befestigt wird. Abtastung bedeutet dabei ein gleichmäßiges Verschwenken des Bildverstärkers in Orientierungen, die später intraoperativ potentielle Aufnahmerichtungen darstellen. Der C-Bogen wird hierfür sukzessive in eine Vielzahl zulässiger Orientierungen innerhalb seines Arbeitsraumes gebracht, wobei für jede der Stellungen ein Röntgenbild aufgenommen wird. Die Aufnahmen der ersten Serie werden individuell entzerrt und liefern so eine orientierungsabhängige Modellierung der Bildverzeichnung. Die Aufnahmen der zweiten Serie hingegen erlauben die orientierungsabhängige Bestimmung der Projektionsparameter. Die jeweilige Orientierung des C-Bogens ergibt sich durch die mit Hilfe des Infrarot-Trackingsystems überwachte Position und Orientierung des IR-LED-Phantoms.

In der intraoperativen Anwendung dieser präoperativ gespeicherten Kalibrierungsdaten kann nicht

gewährleistet werden, daß die ausgewählte Orientierung exakt einer der Orientierungen aus der präoperativen Vorkalibrierung entspricht. Aus diesem Grund werden die Parameter für die Verzeichnung sowie die Projektionsgeometrie aus den einzelnen Kalibrierungsergebnissen der am nächsten gelegenen Orientierungen linear interpoliert. Die Berechnung der externen Kameraposition erfolgt indirekt über die Lage des am Detektor befindlichen IR-LED-Phantoms zum Zeitpunkt einer Aufnahme.

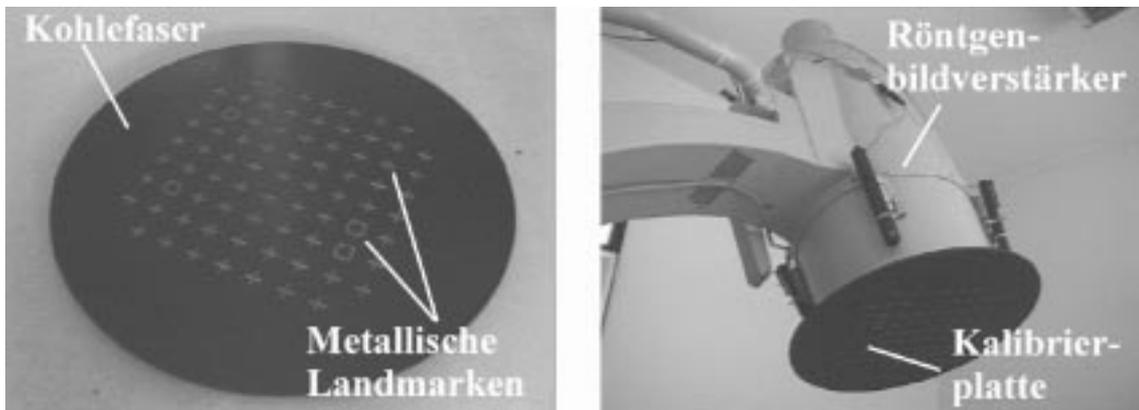


Abbildung B.4: Kalibrierkörper 1: Bestimmung der Verzeichnungscharakteristik

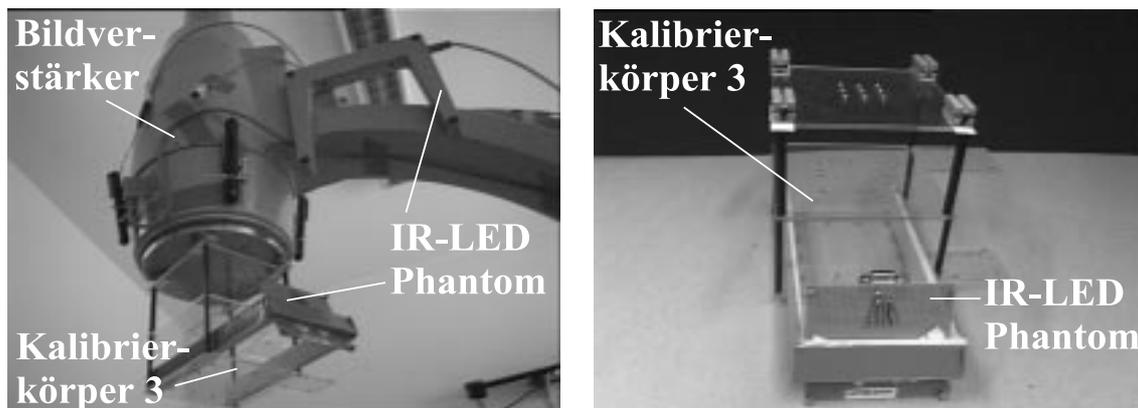


Abbildung B.5: Kalibrierkörper 3: Bestimmung der Projektionsgeometrie

Vor- und Nachteile

Der intraoperative Kalibrieransatz besitzt gegenüber der präoperativen Kalibrierung zwei wesentliche Vorteile. Zunächst ist kein präoperativer Installationsaufwand nötig, sondern die Anwendung der Methode kann intraoperativ ggf. sofort zum Einsatz kommen. Desweiteren wird, abgesehen von den beiden Kalibrierkörpern, keine zusätzliche Hardware benötigt. Dies gilt insbesondere für das Infrarot-Trackingsystem. Im Gegensatz zu einem C-Bogen gehört letzteres noch nicht zu der Standardausstattung eines OP-Saals. Demgegenüber stehen jedoch auch mehrere gewichtige Nachteile des Verfahrens.

Der statische Referenzkörper (Kalibrierkörper 2) schränkt die freie Wahl der Aufnahmerichtungen stark ein, da er sich in jeder Aufnahme gut sichtbar zusammen mit dem betrachteten Patientenknöchel im Strahlengang befinden muß. Zusätzlich verschlechtert sich die resultierende Bildqualität durch die beiden permanent sichtbaren Kalibrierkörper signifikant (Abb. B.3). Anatomische Detailstrukturen gehen mitunter vollständig oder zumindest teilweise verloren.

Die Vorteile einer präoperativen Kalibrierung ergeben sich direkt aus den genannten Nachteilen des intraoperativen Ansatzes und umgekehrt. D.h. den Vorteilen einer großen Freiheit bzgl. der zu wählenden Aufnahmerichtung sowie der unveränderten Bildqualität stehen die Nachteile des präoperativen Aufwands sowie die Notwendigkeit eines Trackingsystems gegenüber.

Aus der Sicht der in der vorliegenden Arbeit vorgestellten Algorithmen ist die präoperative Kalibrierung in jedem Fall zu bevorzugen. Da sowohl die intensitätsbasierte Registrierung in Kapitel 5 als auch die in Abschnitt 6.5.1 diskutierten Segmentierungsverfahren direkt auf den Bilddaten arbeiten, ist eine hohe Bildqualität wesentlich für die resultierende Präzision der Algorithmen. Kommt dennoch die intraoperative Kalibrierung zum Einsatz, so ist darauf zu achten, daß die Fluoroskopieaufnahmen vor einer Segmentierung oder Registrierung geeignet vorverarbeitet werden, z.B. durch die in Abschnitt 5.2.3 entwickelte Technik der Bildmaskierung.

Anhang C

Fehlerabschätzung bei planarer Approximation

In Abschnitt 5.4.2.4 wird die Approximation verschiedener 3D-Transformationen in der Bildebene des zugrundeliegenden Kameramodells vorgeschlagen. Die drei Ungleichungen 5.22, 5.23 und 5.24 liefern entsprechende *worst-case*-Abschätzungen für die durch die Approximation entstehenden Fehler. Im Hinblick auf eine anschauliche Interpretation der Abschätzungsterme werden dabei Translationen parallel und senkrecht sowie Rotationen um die optische Achse jeweils getrennt betrachtet. Diese Terme werden nachfolgend motiviert und hergeleitet.

Das betrachtete Fehlermaß E beurteilt jeweils den euklidischen Abstand $\| \cdot \|_2$ zwischen korrektem Projektionspunkt ${}^I p_{kor}$ und approximiertem Projektionspunkt ${}^I p_{app}$, nachdem ein gegebener 3D-Punkt räumlich verschoben und ggf. um die optische Achse gedreht wurde:

$$E = \| {}^I p_{kor} - {}^I p_{app} \|_2 \quad (\text{C.1})$$

Entsprechend des in dieser Arbeit verwendeten projektiven Lochkameramodells erfolgt die Projektion eines Objektpunktes $P(X_p, Y_p, Z_p)$ innerhalb des Kamerakoordinatensystems gemäß

$${}^I p(P) := \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \frac{f}{Z_p} \begin{pmatrix} X_p \\ Y_p \end{pmatrix} \quad (\text{C.2})$$

Dabei repräsentiert f die Brennweite und $(c_x, c_y)^T$ das optische Zentrum innerhalb der Bildebene.

Zunächst soll nun die Auswirkung der Approximation bei einer räumlichen Translation $\mathbf{t} = (\Delta X, \Delta Y, \Delta Z)^T$ untersucht werden. Die korrekte Projektion eines um \mathbf{t} verschobenen Punktes $P'(X_p + \Delta X, Y_p + \Delta Y, Z_p + \Delta Z)$ ergibt sich entsprechend zu

$${}^I p_{kor}(P') := \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \frac{f}{Z_p + \Delta Z} \begin{pmatrix} X_p + \Delta X \\ Y_p + \Delta Y \end{pmatrix} \quad (\text{C.3})$$

Wird die räumliche Verschiebung hingegen entsprechend Gl. 5.20 (Seite 71) in der Bildebene approximiert, so resultiert ein Bildpunkt ${}^I p_{app}(P')$:

$${}^I p_{app}(P') := \begin{pmatrix} u'_{app} \\ v'_{app} \end{pmatrix} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \frac{1}{1 + \frac{\Delta Z}{Z_c}} \mathbf{R}_{2D}(0) \left[\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} c_x \\ c_y \end{pmatrix} \right] + \frac{f}{Z_c} \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \quad (\text{C.4})$$

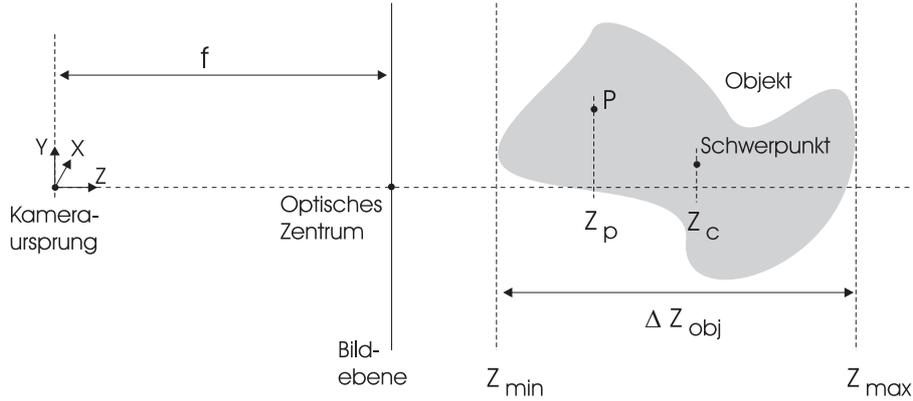


Abbildung C.1: Grundlegende Definitionen: Objektausdehnung in Z-Richtung (ΔZ_{obj}), kleinste und größte Z-Koordinate eines Objektpunktes (Z_{min} , Z_{max}), Schwerpunkt des Objektes (Z_c)

Dabei bleibt eine Rotation um die optische Achse zunächst unberücksichtigt, d.h. $\Delta r_z = 0$. Diese Transformation wird später getrennt betrachtet. Durch Einsetzen von Gl. C.2 ergibt sich schließlich:

$${}^I p_{app}(P') = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \frac{Z_c}{Z_c + \Delta Z} \frac{f}{Z_p} \begin{pmatrix} X_p \\ Y_p \end{pmatrix} + \frac{f}{Z_c} \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \quad (C.5)$$

Durch Subtraktion der beiden letzten Gleichungen läßt sich nun der resultierende Fehler E entsprechend Gl. C.1 berechnen bzw. abschätzen:

$$E = f \left\| \frac{Z_p(Z_c + \Delta Z) - Z_c(Z_p + \Delta Z)}{Z_p(Z_c + \Delta Z)(Z_p + \Delta Z)} \begin{pmatrix} X_p \\ Y_p \end{pmatrix} + \frac{Z_c - (Z_p + \Delta Z)}{Z_c(Z_p + \Delta Z)} \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \right\|_2 \quad (C.6)$$

Dieser Term wird im folgenden getrennt für die beiden Fälle einer parallelen ($\Delta Z = 0$) bzw. senkrechten ($\Delta X = \Delta Y = 0$) Bewegung bzgl. der Bildebene analysiert:

1. $\Delta Z = 0$:

Gleichung C.6 vereinfacht sich in diesem Fall wie folgt:

$$\| {}^I p_{kor} - {}^I p_{app} \|_2 = f \left\| \frac{Z_c - Z_p}{Z_c Z_p} \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \right\|_2 = f \left| \frac{Z_c - Z_p}{Z_c Z_p} \right| \sqrt{\Delta X^2 + \Delta Y^2} \quad (C.7)$$

Der Zähler des Bruches entspricht dabei dem Abstand der Z-Koordinaten des betrachteten Punktes vom Objektzentrum. Dieser Term läßt sich geeignet abschätzen. Entsprechend der Definitionen in Abb. C.1 ergibt sich für ein durchschnittliches Objekt maximal der Abstand $\frac{1}{2} \Delta Z_{obj}$, d.h. $|Z_c - Z_p| \leq \frac{1}{2} \Delta Z_{obj}$. Dies führt schließlich zu der Fehlerabschätzung aus Gl. 5.22:

$$\| {}^I p_{kor} - {}^I p_{app} \|_2 \leq \frac{\frac{1}{2} \Delta Z_{obj}}{Z_c - \frac{1}{2} \Delta Z_{obj}} f \sqrt{\Delta X^2 + \Delta Y^2} = \frac{f \sqrt{\Delta X^2 + \Delta Y^2}}{2 \frac{Z_c^2}{\Delta Z_{obj}} - Z_c} \quad (C.8)$$

2. $\Delta X = \Delta Y = 0$:

Auch in diesem Fall vereinfacht sich Gleichung C.6:

$$\| {}^I p_{kor} - {}^I p_{app} \|_2 = f \left\| \frac{Z_p(Z_c + \Delta Z) - Z_c(Z_p + \Delta Z)}{Z_p(Z_c + \Delta Z)(Z_p + \Delta Z)} \begin{pmatrix} X_p \\ Y_p \end{pmatrix} \right\|_2 \quad (C.9)$$

Dabei stellt der gesamte Bruch lediglich einen multiplikativen Faktor dar und die euklidische Norm von $(X_p, Y_p)^T$ ergibt sich zu $\sqrt{X_p^2 + Y_p^2}$. Wird nun Z_p wahlweise durch die kleinste mögliche z-Koordinate Z_{min} ($\approx Z_c - \frac{1}{2}\Delta Z_{obj}$) bzw. durch die größte z-Koordinate Z_{max} ($\approx Z_c + \frac{1}{2}\Delta Z_{obj}$) ersetzt, so läßt sich der Fehlerwert E wie folgt abschätzen:

$$\begin{aligned} \| {}^I p_{kor} - {}^I p_{app} \|_2 &\leq \frac{f}{2} \sqrt{X_p^2 + Y_p^2} \left| \frac{Z_{max}(Z_c + \Delta Z) - Z_c(Z_{min} + \Delta Z)}{Z_{min}(Z_c + \Delta Z)(Z_{min} + \Delta Z)} \right| \\ &= \dots = \frac{f}{2} \sqrt{X_p^2 + Y_p^2} \left| \frac{1}{Z_{min} + \Delta Z} \left(1 - Z_c \frac{Z_{min} + \Delta Z}{Z_{min}(Z_c + \Delta Z)} \right) \right| \end{aligned} \quad (C.10)$$

Dieser Term entspricht der Fehlerabschätzung aus Gl. 5.23.

Bisher blieben die Auswirkungen der planaren Approximation einer räumlichen Rotation des betrachteten Objektes um die optische Achse der Kamera unberücksichtigt. In Abschnitt 5.4.2.4 wurde bereits darauf hingewiesen, daß die vorgeschlagene Approximation hierbei fehlerfrei bleibt; abgesehen von Rundungsabweichungen aufgrund der unvermeidlichen Pixel-Diskretisierung der zugrunde liegenden Digitalbild-Repräsentation. Diese Behauptung läßt sich verifizieren, indem für die Projektion eines beliebigen Objektpunktes (X_p, Y_p, Z_p) nach einer Δr_z -Rotation um die optische Achse aus den Gleichungen C.2 bzw. 5.20 für den korrekten wie für den approximierten Projektionspunkt derselbe Term abgeleitet wird:

$${}^I p_{kor} = {}^I p_{app} = \begin{pmatrix} c_x \\ c_y \end{pmatrix} + \frac{f}{Z_p} \begin{pmatrix} X_p \cos \Delta r_z - Y_p \sin \Delta r_z \\ X_p \sin \Delta r_z + Y_p \cos \Delta r_z \end{pmatrix} \quad (C.11)$$

Auf eine explizite Darstellung der trivialen Herleitung wird an dieser Stelle verzichtet.

Anhang D

Repräsentation von dreidimensionalen Oberflächenmodellen

Viele Verfahren der vorliegenden Arbeit (z.B. in den Kapiteln 5 und 6) basieren auf der Verwendung dreidimensionaler Daten bzw. Modelle. Dabei beschränkt sich die Repräsentation jeweils auf diskrete, flächenbasierte 3D-Modelle. Dieser Anhang beschreibt im folgenden den Aufbau sowie die grundsätzlichen Eigenschaften derartiger Modelle.

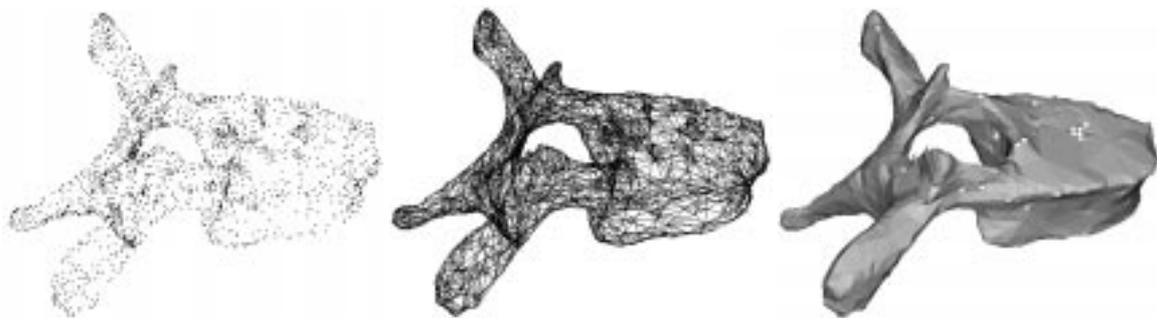


Abbildung D.1: Vertices, Kanten und Flächen: Einzeldarstellung der Datenstruktur-Komponenten eines 3D-Objekts

Ein diskretes 3D-Flächenmodell lässt sich durch die folgenden drei Strukturen hierarchisch beschreiben:

- **Vertex-Liste:**
Ein Vertex ist im Zusammenhang mit 3D-Objekten eine andere Bezeichnung für 3D-Knotenpunkt oder auch 3D-Stützpunkt. Die Liste an Vertices beinhaltet daher alle 3D-Punkte des Objektes. Diese Punkte stellen die geometrische Grundinformation des Objektes (Position einzelner Objektpunkte) dar. Sie liefern aber noch keine Information über die Topologie (d.h. den Flächenverlauf) des Objektes. Dies leisten erst die nachfolgend beschriebenen Listen von Kanten und Flächen.
- **Kanten-Liste:**
Der Verlauf einer Kante auf der Oberfläche des Objektes wird durch die Angabe zweier Vertices

(Start- und End-Vertex) eindeutig beschrieben. Die Kanten-Liste baut daher direkt auf der Vertex-Liste auf, indem jedes Element (v_i, v_j) durch die Indizes von Start- und End-Vertex innerhalb der Vertex-Liste angegeben wird. Dabei spielt die Richtung der Kante prinzipiell keine Rolle, d.h. (v_i, v_j) und (v_j, v_i) beschreiben dieselbe Kante.

Die Kanten-Liste ist im Hinblick auf die nachfolgend beschriebene Flächen-Liste redundant. D.h. Informationen über die Existenz sowie den Verlauf von Kanten können vollständig aus der zuletzt genannten Liste bestimmt werden. Eine redundante Verwaltung ist aus Performance-Gründen jedoch meist ratsam.

Um häufig benötigte Operatoren auf 3D-Objekten zeitlich zu optimieren, ist es darüber hinaus sinnvoll, zusätzliche redundante Verweisstrukturen zu speichern. Ein solcher Operator realisiert z.B. das Auffinden aller Kanten, die einen bestimmten Vertex v als Start- oder End-Punkt enthalten: $(v, *)$ oder $(*, v)$. Damit kann z.B. auf effiziente Weise bestimmt werden, welche Kanten von einem Vertex ausgehen und zu welchen Nachbar-Vertices diese führen.

- Flächen-Liste:

Auch diese Liste setzt – wie die Kanten-Liste – hierarchisch auf der Liste der Vertices auf. Ein Eintrag enthält jeweils mindestens drei Verweise (Indizes) auf die Vertex-Liste. Die Vertices definieren die Eckpunkte des beschriebenen Flächenstücks. Dabei ist die Reihenfolge der Verweise dieses Mal aber wesentlich, denn sie entscheidet über die Orientierung der Fläche.

Diese Orientierung spielt u.a. für eine Vielzahl an Visualisierungsverfahren eine bedeutende Rolle, z.B. innerhalb von *Hidden-Surface-Algorithm*en zur Entfernung nicht sichtbarer Flächenanteile oder auch bei der Berechnung einer optisch realistischen Schattierung [Fol90].

Die Beurteilung der Orientierung erfolgt über die Berechnung des Normalenvektors der Fläche. Dazu wird das Kreuzprodukt des Differenzvektors zwischen erstem und zweitem Vertex sowie des Differenzvektors zwischen erstem und drittem Vertex gebildet. Eine Veränderung der Reihenfolge der Vertices ändert dabei das Vorzeichen des Normalenvektors und damit die Orientierung.

Abb. D.1 demonstriert den Unterschied zwischen Vertices, Kanten und Flächen anhand des Modells eines menschlichen T6-Brustwirbels.

Die beschriebene Struktur erlaubt zunächst lediglich die Modellierung starrer Objekte (vgl. Kapitel 5). Das in Kapitel 6 vorgestellte Verfahren zur elastischen Registrierung erweitert daher diese Modelle so, daß sie geeignet verformt werden können. Die Verformung ist dabei derart definiert, daß durch ein entsprechendes Verschiebungsvektorfeld stets lediglich die Geometrie des Objektes, d.h. seine Vertex-Liste, verändert wird. Die Topologie hingegen (Kanten- und Flächen-Struktur) bleibt unverändert erhalten, was die möglichen Verformungsarten gewollt stark einschränkt.

Um einen Eindruck von der Größe der in dieser Arbeit verwendeten 3D-Modelle zu vermitteln, gibt Tabelle D.1 die wichtigsten Kennzahlen für das Wirbel-Modell aus Abb. D.1 an.

Vertices	Kanten	Flächen	Dateigröße
2582	6844	4489	149268 Byte

Tabelle D.1: Statistische Informationen zu dem Modell aus Abb. D.1

Anhang E

Finite-Elemente-Methoden (FEM) – Grundlagen

Die Methode der *Finiten Elemente* (FEM) ist ein universell einsetzbares, numerisches Verfahren zur Lösung komplizierter Differentialgleichungen aus Technik und Naturwissenschaft. Sie kann u.a. dazu genutzt werden, die kontinuierliche Bewegung und das materialabhängige Verhalten eines elastischen Körpers durch geeignete Diskretisierung desselben zu modellieren.

Für die Vielzahl bekannter Anwendungen auch im medizinischen Umfeld sei stellvertretend die präoperative Planung eines gesichtschirurgischen Eingriffs erwähnt [KGo96]. Hierbei wird das Zusammenspiel zwischen Haut, Muskeln und Knochen durch ein FEM-Modell simuliert, so daß der Patient bereits vor Beginn des Eingriffs einen realitätsnahen Eindruck erhält, wie sich seine Gesichtszüge durch die geplante Umstellung der Kiefer- und Wangenknochen verändern werden.

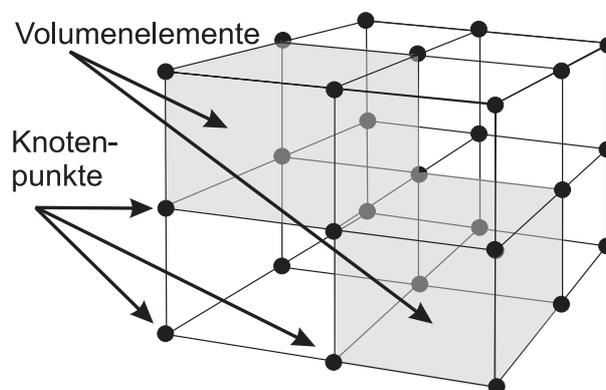


Abbildung E.1: Diskrete Repräsentation eines Objektes durch *Finite Elemente*

Das zentrale Prinzip der FEM-Methodik besteht darin, einen beliebig komplexen Körper durch das Zusammenspiel einer Vielzahl kleiner ein- oder mehrdimensionaler Grundbausteine (z.B. Federn, Würfel etc.), deren Geometrie und physikalisches Verhalten hinreichend exakt modellierbar ist, zu repräsentieren. Hierfür wird der jeweils betrachtete Körper zunächst z.B. in eine Vielzahl kleiner Basis-Volumenelemente unterteilt. Diese Elemente sind miteinander an diskreten Knotenpunkten verbunden, wodurch

sich ein räumliches Gitternetz ergibt (Abb. E.1). Nach Berücksichtigung und Festlegung physikalischer Materialeigenschaften wie Dichte oder Elastizität für jedes der Volumen-Elemente läßt sich eine Reihe von Gleichungen ableiten, die das Verhalten des Gesamtsystems als Reaktion auf eine äußere Krafteinwirkung beschreiben. Die einzelnen Gleichungen lassen sich zu einem Gleichungssystem zusammenfassen, welches die Verschiebung (\mathbf{U}), die Geschwindigkeit ($\dot{\mathbf{U}}$) und die Beschleunigung ($\ddot{\mathbf{U}}$) aller Knotenpunkte innerhalb des Systems modelliert (*Dynamische Gleichungsgleichung*):

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{D}\dot{\mathbf{U}} + \mathbf{K}\mathbf{U} = \mathbf{F} \quad (\text{E.1})$$

Dabei ist \mathbf{M} die *Masse*-, \mathbf{D} die *Dämpfungs*-, und \mathbf{K} die sogenannte *Steifheitsmatrix*. \mathbf{K} und \mathbf{M} repräsentieren die bereits erwähnten Materialeigenschaften des modellierten Körpers, entsprechend der verwendeten Anzahl und Position von Knotenpunkten diskretisiert. Der Vektor \mathbf{F} beschreibt die äußere Krafteinwirkung auf das System, ebenfalls zerlegt in Teilkräfte, die lokal auf einzelne Knoten des Systems wirken. Die Matrizen \mathbf{M} und \mathbf{K} ergeben sich durch geeignete Integration über das gesamte Modell-Volumen [MPK94].

Sind dynamische Prozesse bei der Betrachtung eines elastischen Körpers nicht relevant, so kann Gl. E.1 stark vereinfacht werden:

$$\mathbf{K}\mathbf{U} = \mathbf{F} \quad (\text{E.2})$$

Dieser Zusammenhang wird als *Statische Gleichungsgleichung* bezeichnet, da er den letztendlichen Ruhezustand beschreibt, den das gesamte System einnimmt, wenn äußere Kräfte entsprechend \mathbf{F} angelegt werden.

Gl. E.1 besitzt in der Praxis einige Nachteile. Zum einen läßt sie sich nicht geschlossen auflösen, zum anderen ist eine iterative Approximation der Lösung durch diskrete Zeitschritte für ein System mit sehr vielen Knoten extrem zeitaufwendig [PS91, NA94].

Modale Analyse

Die Technik der *Modalen Analyse* erlaubt die Vereinfachung von Gleichung E.1 durch die Betrachtung einer freien ($\mathbf{F} = \mathbf{0}$) und ungedämpften ($\mathbf{D} = \mathbf{0}$) Schwingung des modellierten Körpers. Diese Annahmen führen mit Gleichung E.1 auf das folgende verallgemeinerte Eigenwertproblem [GA94]:

$$\mathbf{K}\phi_i = \omega_i^2 \mathbf{M}\phi_i \quad (\text{E.3})$$

Dieses Gleichungssystem besitzt für einen dreidimensionalen Körper, der durch n Knotenpunkte diskretisiert wurde, insgesamt $3n$ Lösungen (ω_i^2, ϕ_i) .

Jeder Eigenwert ω_i^2 definiert dabei eine der natürlichen Schwingungsfrequenzen ω_i des Körpers. Die jeweils zugehörigen Eigenvektoren ϕ_i werden gemeinhin als *freie Schwingungs*- bzw. *Vibrationsmodi* des Körpers bezeichnet. Sie beschreiben die räumliche Verschiebungsrichtung jedes einzelnen Knotenpunktes des Gitternetzes in Abhängigkeit von der betrachteten Schwingungsfrequenz ω . Die natürlichen Schwingungsfrequenzen eines Körpers zusammen mit ihren Vibrationsmodi erlauben u.a. direkte Aussagen darüber, welche Frequenzen an welchen Stellen zu unerwünschten Resonanzen führen [GA94].

Im Rahmen dieser Dissertation werden die Vibrationsmodi jedoch für einen anderen Zweck verwendet. Die Eigenvektoren ϕ_i bilden zusammen eine orthogonale Basis. Dies läßt sich ausnutzen, um das

Gleichungssystem E.2 durch eine entsprechende Basistransformation in insgesamt $3n$ getrennte Gleichungssysteme zu zerlegen [MPK94]:

$$\mathbf{U} = \Phi \tilde{\mathbf{U}} \quad (\text{E.4})$$

Dabei repräsentiert Φ die Matrix mit den einzelnen Eigenvektoren ϕ als Spalten. Entsprechend ergeben sich durch einfache Transformation eine modale Massematrix $\tilde{\mathbf{M}}$, eine modale Steifheitsmatrix $\tilde{\mathbf{K}}$ sowie ein modaler Kraft-Vektor $\tilde{\mathbf{F}}$:

$$\tilde{\mathbf{M}} = \Phi^T \mathbf{M} \Phi \quad (\text{E.5})$$

$$\tilde{\mathbf{K}} = \Phi^T \mathbf{K} \Phi \quad (\text{E.6})$$

$$\tilde{\mathbf{F}} = \Phi^T \mathbf{F} \Phi \quad (\text{E.7})$$

Im statischen Fall führt dies zu

$$\tilde{\mathbf{K}} \tilde{\mathbf{U}} = \tilde{\mathbf{F}} \quad (\text{E.8})$$

Diese Diagonalisierung erlaubt es, beliebige Verschiebungen der Knotenpunkte des modellierten Systems als Linearkombination der Vibrations-Modi zu formulieren:

$$\mathbf{U} = \Phi \mathbf{a} = \sum_{i=1}^{3n} \phi_i \alpha_i \quad (\text{E.9})$$

Aus dem Bereich der Schwingungsanalyse ist hinlänglich bekannt, daß hochfrequente Anteile einer Vibration (entsprechend großen Eigenwerten ω_i) stets sehr stark gedämpft sind, nur kleine, lokale Verformungen bewirken und deshalb keinen wesentlichen Beitrag zur Gesamtverformung eines Körpers liefern. Interessant im Hinblick auf globale Verformungen sind daher lediglich die niederfrequenten Anteile der Schwingung. Deshalb kann die Linearkombination aus Gl. E.9 entsprechend gekürzt werden, indem nur ein kleiner Teil der Eigenvektoren und -werte mit einfließt.

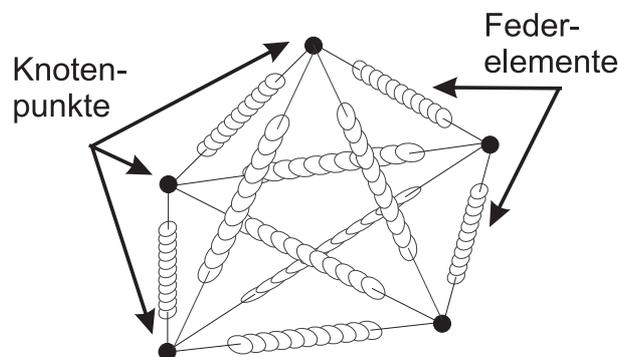


Abbildung E.2: Einfache Feder-Masse-Repräsentation eines Objektes

Literaturverzeichnis

- [Bar84] A.H. Barr. Global and Local Deformations of Solid Primitives. *Computer Graphics*, 18(3):21–30, 1984.
- [BB96] L.M.G Brown and T.E. Boulton. Registration of planar film radiographs with computed tomography. In *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MM-BIA96)*, pages 42–51, 1996.
- [BF95] F. Betting and J. Feldmar. 3d-2d projective registration of anatomical surfaces with their projections. In *Information Processing in Medical Imaging*, pages 275–286, 1995.
- [BM92] P.J. Besl and N.D. May. A method for registration of 3d-shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 14(2):239–256, 1992.
- [BM97] W. Barrett and E.N. Mortensen. Interactive live-wire boundary extraction. *Medical Image Analysis (Oxford University Press)*, 1(4):331–341, 1997.
- [Bö98] Börner M., Lahmer A., Bauer A., Stier U. Experiences with the ROBODOC system in more than 1000 cases. *Computer-Aided Radiology CAR98, Tokyo*, pages 689–693, 1998.
- [Boo91] F.L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, 1991.
- [Bra96] Brack C., Roth M., Schweikard A. et al. Towards accurate x-ray camera calibration in computer assisted robotic surgery. *Computer-Aided Radiology CAR96, Paris*, pages 721–728, 1996.
- [Bra97] Brandt. G. et al. A medical robot system for orthopedic interventions. In *Computer Assisted Radiology (CAR'97), Berlin*, 1997.
- [Bra98] Brack C., Roth M., Schweikard A., Burgkart R. et al. Accurate x-ray-based navigation in computer-assisted orthopedic surgery. *Computer-Aided Radiology CAR98, Tokyo*, pages 716–722, 1998.
- [Bra99] Brack C., Burgkart R., Czopf A., Götte H., Roth M., Radig B., Schweikard A. Radiological Navigation in Orthopaedic Surgery. In J. Jerosch, K. Nicol, and K. Pekenkamp, editors, *Rechnergestützte Verfahren in Orthopädie und Unfallchirurgie*, pages 452–460. Steinkopf Verlag, Darmstadt, 1999.

- [Bra00a] Brack C. *Röntgenbasierte Navigation von chirurgischen Werkzeugen in der Orthopädie*. PhD thesis, Institut für Informatik, Technische Universität München, 2000. to appear.
- [Bra00b] Brack C., Roth M., Schweikard A., et al. Accurate X-ray Camera Calibration and Navigation in Orthopedic Surgery. *IEEE Computer Aided Surgery*, 2000. submitted.
- [Bre73] R.P. Brent. *Algorithms for Minimization without Derivatives*, chapter 5. Prentice-Hall, Englewood-Cliffs, NJ, 1973.
- [Bro89] Bronstein, Semenajew. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, ISBN 3-87144-492-8, 1989.
- [BSHM94] R.D. Bucholz, K.R. Smith, J. Henderson, and L. MCDurmont. Intraoperative localization using a three dimensional optical digitizer. *Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, pages 283–290, 1994.
- [CAW96] John W.C. Chin-A-Woeng. Digital simulation of fluoroscopic images for clinical training. Master's thesis, Technical University Delft, August 1996. (Research done at Department of Image Sciences, Institute of Medical Image Sciences).
- [CCTG92] T.F. Cootes, D.H. Cooper, C.J. Taylor, and J. Graham. Trainable method of parametric shape description. *Image and Vision Computing*, 10(5):289–294, 1992.
- [Cha98] Chassat F., Lavallée S. Experimental Protocol of Accuracy Evaluation of 6-D Localizers for Computer-Integrated Surgery: Application to Four Localizers. *MICCAI 98, LNCS 1496, Springer-Verlag*, pages 277–284, 1998.
- [Cia97] Ciampalini A., Cignoni P., Montani C., Scopigno R. Multiresolution Decimation based on Global Error. *The Visual Computer, Springer Verlag*, 13(5), 1997.
- [Coh96] Cohen J., et al. Simplification Envelopes. *Proceedings of SIGGRAPH 96, (New Orleans, LA, August 4-9)*, pages 119–128, 1996. URL: <http://www.csunc.edu/~geom/envelope.html>.
- [Coo95a] Cootes T.F. and Taylor C.J. Combining point distribution models with shape models based on finite element analysis. *Image and Vision Computing*, 13(5), 1995.
- [Coo95b] Cootes T.F., Taylor C.J., Cooper D.H. and Graham J. Active Shape Models – Their Training and Application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
- [Coo95c] Cootes T.F., Taylor C.J., Cooper D.H. and Graham J. Training models of shape from sets of examples. *British Machine Vision Conference*, 1995.
- [CT95] T.F. Cootes and C.J. Taylor. Combining point distribution models with shape models on finite element analysis. *Image and Vision Computing*, 13(5):403–410, 1995.
- [Cut95] Cutting C. et al. A Three-Dimensional Smooth Surface Analysis of Untreated Crouzon's Disease in the Adult. *Journal of Craniofacial Surgery*, 1:1–10, 1995.
- [DeM92] DeMenthon D. F. Model-based object pose in 25 lines of code. *Proceedings, [DARPA] Image Understanding Workshop, San Diego, CA, January 26-29*, pages 753–761, 1992.

- [DiG97] DiGioia A.M., et al. HipNav: Pre-operative Planning and Intra-operative Navigational Guidance for Acetabular Implant Placement in Total Hip Replacement Surgery. Technical report, Center for Orthopaedic Research, Shadyside Hospital, Pittsburgh, 1997.
- [Dij59] Dijkstra. A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, 1:269–270, 1959.
- [Dö99] Dötter M. Projektive 2D/3D-Registrierung formvariabler, anatomischer Oberflächen-Modelle. Master's thesis, Technische Universität München, Informatik IX, 1999.
- [Ell96] Ellis R.E., et al. Use of a Biocompatible Fiducial Marker in Evaluating the Accuracy of CT Image Registration. *Investigative Radiology*, 1996.
- [FAB95] J. Feldmar, N. Ayache, and F. Betting. 3d-2d projective registration of free-form curves and surfaces. In *ICCV95*, pages 549–555, 1995.
- [FBF77] J.H. Friedman, J.L. Bentley, and R.A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Software*, 3(3):209–226, 1977.
- [FH86] O.D. Faugeras and M. Hebert. The representation, recognition and locating of 3d objects. *Int. Journal Robotic Res.*, 5(3):27–52, 1986.
- [FL98] M. Fleute and S. Lavalée. Building a complete surface model from sparse data using statistical shape models: Application to computer assisted knee surgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'99)*, pages 879–887, 1998.
- [FL99] M. Fleute and S. Lavalée. Nonrigid 3-d/2-d registration of images using statistical models. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'99)*, 1999.
- [Fol90] Foley J. D. *Computer Graphics - Principle and Practice*. Addison Wesley Publishing Company, 1990.
- [Fre98] Frey D. and Lindauer F. Realisierung eines Verfahrens zur grauwert-basierten 2D/3D-Registrierung von Computer-Tomographie-Daten mit Röntgenaufnahmen. Systementwicklungsprojekt (SEP), Informatik IX, TU München, Juli 1998.
- [GA94] J.I. Gobat and D.C. Atkinson. The FElt System – User's Guide and Reference Manual. Technical Report CS94-376, Dept. of Computer Science and Engineering, University of California, San Diego, 1994.
- [GBHE91] T. Greitz, C. Bohm, S. Holte, and L. Erikson. A computerized brain atlas: construction, anatomical content and some applications. *Computer Assisted Tomography*, 15:26–38, 1991.
- [GBSN97] J. Gong, R. Bächler, M. Sati, and L.-P. Nolte. Restricted Surface Matching, A New Approach to Registration in Computer Assisted Surgery. *CVRMed-MRCAS 97, LNCS 1205, Springer-Verlag*, pages 597–605, 1997.
- [GFo95] A.R. Gunkel, W. Freysinger, and other. Application of the ARTMA image-guided navigation system to endonasal sinus surgery. In *Computer Assisted Radiology (CAR'95), Berlin*, pages 1147–1151, 1995.

- [Gib98] P. Giblin. Apparent contours: an outline. *Phil. Trans. Royal Society, London*, A(356):1087–1102, 1998.
- [Gö00] Götte, H. *Computer- und Roboterunterstützung für die Knieendoprothetik*. PhD thesis, Institut für Werkzeugmaschinen und Betriebswissenschaften (iwb), Technische Universität München, 2000. geplant.
- [Goe88] Goerke H. *Medizin und Technik*. Callwey Verlag, München, 1988.
- [Gos97] Gossé F., Brack C., Roth M., Schweikard A., Götte H., Vahldiek M., Rühmann O. Roboterunterstützung in der Knieendoprothetik. *Der Orthopäde, Springer Verlag*, 26 (3), ISBN 0085-4530, 1997.
- [Gue96] Gueziec A. Surface Simplification Inside a Tolerance Volume. Technical Report RC 20440, IBM Research Report (Computer Graphics, Computational Geometry), 1996.
- [Gue98] Gueziec A., et al. Anatomy Based Registration of CT-scan and Intraoperative X-ray Images For Guiding a Surgical Robot. *IEEE Transactions on Medical Imaging, special issue on Image Guidance an Therapy*, 1998. to appear.
- [GWS99] R. Göcke, J. Weese, and H. Schuhmann. Fast Volume Rendering Methods for Voxel-base 2D/3D Registration - A Comparative Study. *Int. Workshop Biomedical Image Registration (WBIR99)*, pages 89–102, 1999.
- [HB80] W. Huck and U. Baer. A new targeting for stereotaxis procedures within the ct scanner. *Neuroradiology*, 19:13–17, 1980.
- [HC97] A. Hamadeh and P. Cinquin. Kinematic Study of Lumbar Spine Using Functional Radiographies and 3D/2D Registration. *CVRMed-MRCAS 97, LNCS 1205, Springer-Verlag*, pages 108–118, 1997.
- [Hof97] Hofstetter, R. et al. Fluoroscopy based surgical navigation-concept and clinical applications. In *Computer Assisted Radiology (CAR'97)*, Berlin, 1997.
- [Hor87] B. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.
- [HRB⁺99] P. Heeckt, M. Rühl, G. Buchhorn, H.G. Willert, and other. Computer assisted surgical planning and robotics mit dem caspar-system. In J. Jerosch, K. Nicol, and K. Pekenkamp, editors, *Rechnergestützte Verfahren in Orthopädie und Unfallchirurgie*, pages 414–433. Steinkopf Verlag, Darmstadt, 1999.
- [HS93a] R.M. Haralick and L.G. Shapiro. *Computer and Robot Vision*, volume 2, chapter 14. Analytic Photogrammetry. Addison Wesley Publishing Company, 1993.
- [HS93b] R.M. Haralick and L.G. Shapiro. *Computer and Robot Vision*, volume 2. Addison Wesley Publishing Company, 1993.

- [HSLC95] A. Hamadeh, P. Sautot, S. Lavalée, and P. Cinquin. Towards automatic registration between ct and x-ray images: cooperation between 3d/2d registration and 2d edge detection. In *2. Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95), Baltimore*, pages 39–46, 1995.
- [HTT93] A. Hill, A. Thornham, and C.J. Taylor. Model-Based Interpretation of 3D Medical Images. *British Machine Vision Conference*, 2:339–348, 1993.
- [Jer99] Jerosch, J. et al. Grundlagen zu Operationsrobotern in der Hüftendoprothetik und mögliche Ansätze zur Qualitätsverbesserung von Operationsplanung und Umsetzung. In J. Jerosch, K. Nicol, and K. Pekenkamp, editors, *Rechnergestützte Verfahren in Orthopädie und Unfallchirurgie*, pages 434–451. Steinkopf Verlag, Darmstadt, 1999.
- [Jol86] Jolliffe, L.T. *Principial Component Analysis*. Springer Verlag, New York, 1986.
- [JRH92] H. Jiang, R.A. Robb, and K.S. Holton. A New Approach to 3-D Registration of Multimodality Medical Images by Surface Matching. *SPIE Visualization in Biomedical Computing*, 1808:196–213, 1992.
- [JS99] M. Jenkinson and S. Smith. An investigation of the robustness of registration methods. *Int. Workshop Biomedical Image Registration (WBIR99)*, pages 200–210, 1999.
- [Kas87] Kass M., Witkin A., Terzopoulos D. Snakes: Active contour models. *International Conference on Computer Vision, London*, pages 259–268, 1987.
- [KGo96] R. M. Koch, M. H. Gross, and other. Simulating facial surgery using finite element models. In *Proceedings of the SIGGRAPH '96 conference*, page 8, 1996.
- [KHJH88] Y.S. Kwoh, J. Hou, E.A. Jonckheere, and S Hayati. A robot with improved absolute positioning accuracy for ct guided stereotactic brain surgery. *IEEE Transactions on Biomedical Engineering*, 35(2):153–160, 1988.
- [Kie96] Kienzle III C. T., et al. A Computer-Assited Total Knee Replacement Surgical System Using a Calibrated Robot. *Computer Integrated Surgery*, pages 409–416, 1996. edited by Taylor, et al., MIT, ISBN 0-262-20097-X.
- [Kre80] Krestel E. *Bildgebende Systeme für die medizinische Diagnostik*. Siemens, 1980.
- [Lav95] Lavalée S., et al. Recovering the position and orientation of free form objects from image contours using 3D distance maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 17(4):378–390, 1995.
- [Lav96a] Lavalée S. Anatomy-Based Registration of Three-Dimensional Medical Images, Range Images, X-Ray Projections, and Three-Dimensional Models Using Octree-Splines. *Computer Integrated Surgery*, pages 115–143, 1996. edited by Taylor, et al., MIT, ISBN 0-262-20097-X.
- [Lav96b] Lavalée S. Registration for Computer-Integrated Surgery: Methodology, State of the Art. *Computer Integrated Surgery*, pages 77–97, 1996. edited by Taylor, et al., MIT, ISBN 0-262-20097-X.

- [Leo91] H. Leonhardt. *dtv-Atlas der Anatomie*. DTV, Thieme Verlag, Stuttgart, 1991.
- [Lip99] H. Lippert. *Lehrbuch Anatomie*. Urban und Fischer, 1999.
- [Mar96] Marcacci M., Dario P., et al. Computer-Assisted Knee Arthroplasty. *Computer Integrated Surgery*, pages 417–423, 1996. edited by Taylor, et al., MIT, ISBN 0-262-20097-X.
- [Mer97] Merloz P., et al. Computer-Assisted Versus Manual Spine Surgery: Clinical Report. *CVRMed-MRCAS 97, LNCS 1205, Springer-Verlag*, pages 541–544, 1997.
- [Met98] D.N. Metaxas. *Physics-Based Deformable Models*. Kluwer Academic Publishers, 1998.
- [Mit97] Mittelstadt B.D., et al. The Evolution of a Surgical Robot from Prototype to Human Clinical Use. Technical report, Center for Orthopaedic Research, Shadyside Hospital, Pittsburgh, 1997.
- [Mor80] Moré J. J., Garbow B. S., Hillstrom K. E. User Guide for MINPACK-1. *Argonne National Laboratory Report ANL-80-74*, 1980.
- [MPK94] J. Martin, A. Pentland, and R. Kikinis. Shape analysis of brain structures using physical and experimental modes. Technical Report 276, M.I.T. Media Laboratory Perceptual Computing, 1994.
- [MT96] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis (Oxford University Press)*, 1(2):91–108, 1996.
- [Mur95] Murase H., Nayar S.K. Visual Learning and Recognition of 3D Objects from Appearance. *Int. Journal of Computer Vision*, 14:5–24, 1995.
- [MV98] J.B.A. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis (Oxford University Press)*, 2(1):1–37, 1998.
- [NA94] C. Nastar and N. Ayache. Classification of nonrigid motion in 3d images using physics-based vibration analysis. In *IEEE Workshop on Biomedical Image Analysis*, 1994.
- [Nas94] Nastar C., Moghaddam B. and Pentland A. Generalized Image Matching: Statistical Learning of Physically-Based Deformations. *4. European Conference on Computer Vision*, April 1994.
- [Nol94] Nolte L.P., et al. A Novel Approach to Computer Assisted Spine Surgery. *Proc. 1st Int Symp on Medical Robotics and Computer Assisted Surgery (MRCAS)*, pages 323–328, 1994.
- [Nol95] Nolte L. et al. Fixation of Spinal Implants. *Journal of Image Guided Surgery, Volume 1, Number 2*, 1995.
- [Pel89] Pelizzari, C.A. et al. Accurate Three-Dimensional Registration of CT, PET and/or MR Images of the Brain. *Computer Assisted Tomography*, 13(1):20–26, 1989.
- [Pen92] A. Pentland. Modal Description for Recognition and Tracking. *IAPR Workshop on Machine Vision Applications, Tokyo*, pages 435–444, 1992.

- [PS91] A. Pentland and S. Sclaroff. Closed-form solutions for physically based shape modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 13(7):715–729, 1991.
- [PT⁺94a] W.H. Press, S.A. Teukolsky, et al. *Numerical Recipes in C, 2.edition*, chapter 10.5 Direction Set (Powell's) Methods in Mulitdimensions. Cambridge University Press, 1994.
- [PT⁺94b] W.H. Press, S.A. Teukolsky, et al. *Numerical Recipes in C, 2.edition*, chapter 10.2 Parabolic Interpolation and Brent's Method in One Dimension. Cambridge University Press, 1994.
- [PTo94] W.H. Press, S.A. Teukolsky, and other. *Numerical Recipes in C*. Cambridge University Press, 1994.
- [PW89] A. Pentland and J. Williams. Good Vibrations: Modal Dynamics for Graphics and Animation. *Computer Graphics*, 23(3):215–222, 1989.
- [RMo95] R. Rohling, P. Munger, and other. Comparison of relative accuracy between a mechanical and an optical tracker for image-guided neurosurgery. *Image Guided Surgery*, 1(1):30–34, 1995.
- [Rot96] Roth M., Brack C., Schweikard A. et al. A New Less Invasive Approach To Knee Surgery Using A Vision-Guided Manipulator. *International Symposium on Robotics and Manufacturing (ISRAM'96), World Automation Congress (WAC'96), May, 27th-30th, Montpellier, France*, pages 731–738, 1996.
- [Rot99a] Roth M., Brack C., Burgkart R., Czopf A., Götte H., Schweikard A. et al. Multi-view contourless registration of bone structures using a single calibrated X-ray fluoroscope. *Computer-Aided Radiology CAR99, Paris*, 1999. to appear, accepted.
- [Rot99b] Roth M., Brack C., Schweikard A. Röntgenbasierte Registrierung von Knochenstrukturen in der Orthopädie. *1. Symposium: Neue Technologien für die Medizin, iwv, Garching 5.3.99, ISBN 3-89675-038-0*, pages 28–42, 1999.
- [Rya95] Ryan, M.J et al. Framless Stereotaxy with Real-Time Tracking of Patient Head Movement and Retrospective Patient-Image Registration. *Int. Symp. on Medical Robotics and Computer Assisted Surgery (MRCAS 95)*, pages 1–7, 1995.
- [Sch93] Schöneburg, E. and Heinzmann, F. and Feddersen, S. *Genetische Algorithmen und Evolutionsstrategien*. Addison-Wesley Publishing Company, 1993.
- [Sc195] Sclaroff, S.E. *Modal Matching: A Method for Describing, Comparing and Manipulating Digital Signals*. PhD thesis, Media Arts and Sciences, MIT, 1995.
- [SHK95] D.A. Simon, M. Hebert, and T. Kanade. Techniques for Fast and Accurate Intrasurgical Registration. *Image Guided Surgery*, 1(1):17–29, 1995.
- [Sim97] Simon D.A., et al. Development and Validation of a Navigational Guidance System for Acetabular Implant Placement. *CVRMed-MRCAS 97, LNCS 1205, Springer-Verlag*, pages 583–592, 1997.

- [SML97] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit – An Object-Oriented Approach to 3D Graphics*. 2.edition, Prentice Hall, ISBN 0-13-954694-4, 1997.
- [STA98] G. Subsol, J.-P. Thirion, and N. Ayache. A scheme for automatically building three-dimensional morphometric anatomical atlases: application to a skull atlas. *Medical Image Analysis (Oxford University Press)*, 2(1):37–60, 1998.
- [STAL94] A. Schweikard, R. Tombropoulos, J. Adler, and J.C. Latombe. Planning for image-guided radiosurgery. *IEEE Int. Conference Robotics and Automation*, 1994.
- [Sto89] J. Stoer. *Numerische Mathematik I*. Springer Verlag, 5.Auflage, 1989.
- [Str97] Strasters K.C., et al. Anatomical Landmark Image Registration: Validation and Comparison. *CVRMed-MRCAS 97, LNCS 1205, Springer-Verlag*, 1997.
- [Stu96] Stulberg S. D., Kienzle III T. C. Computer- and Robot-Assisted Orthopaedic Surgery. *Computer Integrated Surgery*, pages 373–378, 1996. edited by Taylor, et al., MIT, ISBN 0-262-20097-X.
- [Sze95] Szekely G. et al. Segmentation of 3D Objects from MRI Volume Data Using Constrained Elastic Deformations of Flexible Fourier Surface Models. *Lecture Notes in Computer Science 905 (CVRMED'95)*, pages 495–505, 1995.
- [Tag97] H.D. Tagare. Deformable 2-D Template Matching Using Orthogonal Curves. *IEEE Transactions on Medical Imaging*, 16(1):108–117, 1997.
- [Tay96] Taylor R.H., et al. An Image-Directed Robotic System for Precise Orthopaedic Surgery. *Computer Integrated Surgery*, pages 379–396, 1996. edited by Taylor, et al., MIT, ISBN 0-262-20097-X.
- [TPD98] J. Troccaz, M. Peshkin, and B. Davies. Guiding systems for computer-assisted surgery: introducing synergistic devices and discussing the different approaches. *Medical Image Analysis (Oxford University Press)*, 2(2):101–119, 1998.
- [TPK99] A. Tanacs, K. Palagy, and A. Kuba. Target registration error of point-based methods assuming rigid-body and linear motion. *Int. Workshop Biomedical Image Registration (WBIR99)*, pages 223–233, 1999.
- [TWK87] D. Terzopoulos, A. Witkin, and M. Kass. Symmetry-seeking models and 3d object reconstruction. *Int. Journal on Computer Vision*, 1:211–221, 1987.
- [vZo97] T. van Walsum, K.J. Zuiderveld, and other. CT-based Simulation of Fluoroscopy and DSA for Endovascular Surgery Training. *CVRMed-MRCAS 97, LNCS 1205, Springer-Verlag*, pages 273–282, 1997.
- [WB97] H.W. When and P.R. Beranger. Ultrasound-Based Robot Position Estimation. *IEEE Trans. on Robotics and Automation*, 13(5):682–692, 1997.

- [Wee97a] Weese, J., et al. 2D/3D Registration of pre-operative CT images and intra-operative x-ray projections for image guided surgery. *Computer-Aided Radiology CAR97, Berlin*, pages 833–838, 1997.
- [Wee97b] Weese, J., et al. An Approach to 2D/3D Registration of a Vertebra in 2D X-ray Fluoroscopies with 3D CT Images. *CVRMed-MRCAS 97, LNCS 1205, Springer-Verlag*, pages 119–128, 1997.
- [Wee99] Weese, J., et al. Fast voxel-based 2D/3D registration algorithm using a volume rendering method based on the shear-warp factorization. *SPIE Medical Imaging 1999, San Diego*, 1999.
- [WH99] A. Weipert and A. Hanusek. 3D-Planungs- und Herstellungsverfahren von Individualprothesen. In J. Jerosch, K. Nicol, and K. Pekenkamp, editors, *Rechnergestützte Verfahren in Orthopädie und Unfallchirurgie*, pages 472–481. Steinkopf Verlag, Darmstadt, 1999.
- [Wun97] P. Wunsch. *Modellbasierte 3-D Objektageschätzung für visuell geregelte Greifvorgänge in der Robotik*. PhD thesis, Shaker Verlag, Aachen, 1997.
- [XZo96] G. Xu, Z. Zhang, and other. *Epipolar Geometry in Stereo, Motion and Object Recognition – A Unified Approach*, chapter 2.2.1. Kluwer Academic Publishers, 1996.
- [Zon80] F.W. Zonneveld. *Computed Tomography*. Technical Report Technical Report 4522 984 51811, Philips Medical Systems, 1980.