

Kapitel 6

Ergebnisse der modellbasierten Bildverarbeitung

Der modellbasierte Bildverarbeitungsansatz setzt sich aus einer Vielzahl von funktionalen Modulen zusammen, die in ihrem Zusammenspiel über den Erfolg des Gesamtsystems entscheiden. Wie gut die modellbasierte Bildverarbeitung die definierten Zielsetzungen erfüllt, hängt einerseits davon ab, inwieweit die Module die vorgesehenen Teilfunktionen erfüllen, andererseits wie gut die einzelnen Module zur Erreichung des Gesamtziels zusammenwirken. Beide Aspekte werden maßgeblich dadurch beeinflusst, inwieweit das zu bearbeitende Bildmaterial die Annahmen des modellbasierten Bildverarbeitungsansatzes erfüllt.

Aus diesem Grund wird zunächst im Abschnitt 6.1 untersucht, welche Vorgaben durch die Testverfahren und das dabei verwendete Bildmaterial erfüllt werden müssen, um mit ihnen die Validität des Systems zu bestätigen. In Abschnitt 6.2 werden die Erfahrungen beschrieben, die mit Java als Programmiersprache gemacht wurden. In den Abschnitten 6.3.1 bis 6.3.4 wird anschließend untersucht, inwieweit die einzelnen funktionalen Module der modellbasierten Segmentierung die an sie gestellten Anforderungen erfüllen. Die Abschnitte 6.4 und 6.5 beschreiben die Testergebnisse für die modellbasierte Interpolation und Registrierung. Da die modellbasierte Bildverarbeitung mit einem starken Fokus auf die praktische Anwendbarkeit im klinischen Umfeld entwickelt wurde, wird im Abschnitt 6.6 die Verarbeitungspipeline der Strahlentherapie von Augentumoren unter Verwendung der modellbasierten Bildverarbeitung rekapituliert.

6.1 Testverfahren

Eine repräsentative Beurteilung von medizinischen Bildverarbeitungsansätzen ist oft schwer, da das zu verarbeitende Bildmaterial aufgrund der Vielseitigkeit der durch Krankheiten verursachten Veränderungen des Erscheinungsbilds der Anatomie und der nicht-deterministischen Einflüsse bei der Bildgebung sehr stark variiert. Während der Mensch durch intellektuelle Transferleistungen auf veränderte Aufgabenstellungen adaptiv reagieren kann, ist die computergestützte Verarbeitung in vielen Fällen auf die Anwendung von pauschalen Interpretations- und Verarbeitungsschemen angewiesen. Hohe Anforderungen an die Stabilität eines Bildverarbeitungsansatzes gegenüber der Variabilität des Bildmaterials verringern die Erfolgsaussichten, da das Variationsspektrum des Bildmaterials in der Regel nicht mehr zufriedenstellend berücksichtigt werden kann. Selbst wenn der Bildverarbeitungsansatz für eine gegebene Anzahl von beispielhaften Datensätzen gute Ergebnisse produziert, kann er bei dem nächsten hinzugekommenen Datensatz scheitern. Werden die vom Bildverarbeitungsalgorithmus vorausgesetzten Annahmen über das Bildmaterial nicht vollständig erfüllt, kann der Algorithmus oft nicht auf die veränderten Problemstellungen reagieren. Selbst ausführliche Tests können daher nur unvollständig aufzeigen, wo die Stärken und wo die Grenzen einer Bildverarbeitungssoftware liegen.

Bei der Beurteilung des Bildverarbeitungsansatzes darf die mögliche Leistungsfähigkeit der entwickelten Algorithmen nicht überschätzt werden. Selbst die Entwicklung eines Bildverarbeitungsverfahrens für

eine klar definierte Aufgabenstellung wie in dieser Arbeit ist nicht durch eine einzelne Forschungsleistung erbracht, sondern muß durch eine Vielzahl von aufeinander aufbauenden Fortschritten erreicht werden. Durch den drängenden Bedarf für eine solche Lösung besteht die Gefahr, daß hoffnungsvolle Ansätze desavouiert werden, weil sie nicht auf Anhieb die gewünschten Resultate liefern oder allen Problemstellungen gleichzeitig begegnen können.

Das in dieser Arbeit vorgestellte Verfahren zur modellbasierten Bildverarbeitung ist ein neuer Ansatz zur Interpretation und Verarbeitung von medizinischen Bilddaten, der über die in der aktuellen Literatur beschriebenen Ansätze hinausgeht und dadurch nicht auf bereits vorgeleistete Qualitätsanalysen anderer wissenschaftlicher Arbeiten zurückgreifen kann. Andererseits wirft allein die Auswertung von CT- und MRT-Daten der Orbita so viele Probleme auf, daß in dieser Arbeit nicht alle Aspekte in gebührender Ausführlichkeit berücksichtigt werden können.

Ein kritischer Punkt stellt in diesem Zusammenhang die folgende Frage dar: Wie kann das System mit durch Krankheiten verursachten Veränderungen der Bilddarstellung oder Artefakten umgehen? Eine pauschale Antwort kann auf diese Frage nicht gegeben werden, da sich die Fragestellung bei genauerer Betrachtung als so plakativ erweist, daß weder eine praktische noch eine theoretische Beantwortung möglich wäre. Es können sehr unterschiedliche Artefakte mit divergierenden Einflüssen auf die Bilddaten auftreten. Andererseits variieren Position und Erscheinungsformen intraokularer Tumoren von Patient zu Patient. Die Unterschiede sind dabei so stark, daß auch medizinisch geschulte Betrachter nicht ohne einen beträchtlichen Erfahrungsschatz eine endgültige Diagnose vornehmen können. Eine solche Leistung von einem computergestützten Bildverarbeitungsalgorithmus zu erwarten wäre zumindest bei dem aktuellen Stand der Forschung nicht angemessen.

Um die Frage nach dem Umgang mit solchen Variabilitäten im Bildmaterial beantworten zu können, müßten einerseits die Freiheitsgrade dieser Variabilitäten genau quantifiziert und kategorisiert werden. Andererseits müßten Reaktionsstrategien konzipiert und umgesetzt werden, die diese veränderten Anforderungen an die Bildverarbeitung detektieren und ihnen entsprechend begegnen können.

Bei der Auswertung von Datenmaterial der Orbita sind die für die Segmentierung relevanten Veränderungen der Darstellungseigenschaften in erster Linie im Bereich der Sklera zu erwarten. Während eine genaue Aussage über die Position des Tumors vor der Segmentierung nicht möglich ist, kann im Vorfeld festgestellt werden, daß die meisten intraokularen Tumoren auf der Sklera aufsetzen und fast immer in den Bulbus hinein wachsen. Nur in Ausnahmefällen wächst der Tumor auch außerhalb des Bulbus weiter. Um solchen Erkenntnissen Rechnung zu tragen, wäre eine Parametrisierung von Position und Erscheinungsform von Tumoren im Sklerabereich zu konzipieren, um sie gezielt detektieren und segmentieren zu können. Eine mögliche Vorgehensweise wäre in diesem Zusammenhang das Absuchen der ersten Feinsegmentierung nach unerwarteten Grauwertkonstellationen im Bereich der segmentierten Sklera. Wenngleich dies eine naheliegende Vorgehensweise wäre, können die Probleme bei einer solchen Analysestrategie im Vorfeld nur schwer abgeschätzt werden.

Weiterhin müßte definiert werden, wie das System damit umgehen soll, falls das Unerwartete doch eintrifft und der Tumor nach außen wachsen sollte. Zusätzlich können Probleme durch die Variabilität der Erscheinungsform des Tumors auftreten. Während die Berandungen der meisten intraokularen Tumoren klar und deutlich vom gesunden Teil des Glaskörpers abgesetzt sind, können bei einigen auch diffuse Ausläufer existieren, die selbst durch geübte Radiologen in ihren Ausmaßen nur schwer abgeschätzt werden können. Lediglich ein Augenarzt mit jahrelanger Erfahrung bei der Behandlung von Augentumoren kann aufgrund von umfangreichen Operationserfahrungen abschätzen, wie weit der Tumor tatsächlich gewachsen ist und wie sehr die Darstellung im Bildmaterial mit der Realität übereinstimmt. Kein aktueller Bildverarbeitungsansatz ist bisher in der Lage, Jahre der ophthalmologischen Erfahrung einzubeziehen. Ob dieses Ziel mittel- bis langfristig erreicht werden kann, ist heute noch nicht absehbar. Es muß dem-

nach festgestellt werden, daß weder durch andere Forschungsgruppen noch in dieser Arbeit das Problem der Variabilität des Bildmaterials konzeptionell abschließend gelöst werden konnte.

Diese Überlegungen zeigen, daß aufgrund der kaum zu begrenzenden Komplexität der medizinischen Bildverarbeitung im Rahmen der zugrunde liegenden Arbeit die Tests auf eng begrenzte und erreichbare Ziele fokussiert werden müssen. Als primäre Aufgabenstellung für die Evaluierung der modellbasierten Bildverarbeitung sei daher die Untersuchung der grundsätzlichen Funktionalitäten definiert, während eine umfassende Untersuchung der Reaktion auf die stark variierenden Bildeigenschaften für weiterführende wissenschaftliche Arbeiten überlassen werden muß. Nichtsdestotrotz werden bereits bei den hier durchgeführten Tests die wichtigsten Aspekte der gegebenen Problemstellungen angesprochen werden.

Zur Untersuchung der grundsätzlichen Eigenschaften der modellbasierten Bildverarbeitung wurden die folgenden Testverfahren umgesetzt:

1. Anhand von Tests mit Modelldaten sollen die grundlegenden quantitativen Eigenschaften der modellbasierten Bildverarbeitung untersucht werden. Die absoluten Ausmaße der Modelle und ihrer Teilobjekte sind bekannt und können mit den Bildverarbeitungsergebnissen verglichen werden (Abschnitt 6.1.1).
2. Durch das Anfertigen von Testdaten von Schweineaugen sollen erste realistische qualitative und quantitative Untersuchungen bezüglich der Bildverarbeitungseigenschaften gewonnen werden. Für die quantitativen Untersuchungen wurden auf den Schweineaugen Clips angehängt und anatomische Landmarken definiert, die an den Augen ausgemessen und anschließend im Bildmaterial rekonstruiert werden können.
3. Es wurden von zwei gesunden Probanden jeweils drei T1 MRT-Datensätze angefertigt. Zwei mit einer hohen Auflösung (Voxelgröße 0,9x0,9x3mm) und einem Field of View, der in erster Linie die Augen einschließt und einem Datensatz, der den gesamten Schädel mit einer geringen Auflösung (Voxelgröße 1,5x1,5x9mm) darstellt. Anhand dieser Datensätze sollen qualitative Untersuchungen bezüglich weitergehender Eigenschaften wie Artefakte und Stabilität gegenüber Rauschen vorgenommen werden. Darüber hinaus sollen anhand der Datensätze unterschiedlicher Auflösung und Orientierung grundsätzliche Eigenschaften der Registrierung untersucht werden.
4. In Zusammenarbeit mit Dr. Bechrakis von der Augenklinik im Universitätsklinikum Benjamin Franklin wurden 20 aktuelle Datensätze aus der Strahlentherapie von Augentumoren bearbeitet. Dabei wurden die Datensätze nach den Vorgaben von Dr. Bechrakis segmentiert. Die durch das Fachpersonal angefertigten Lösungen werden als Goldstandard für eine korrekte Segmentierung angenommen, da sie die verlässlichste Quelle für eine korrekte Segmentierung darstellen (Abschnitt 6.1.4). Sie werden einerseits als Referenzmodelle, andererseits als Bildmaterial zum Austesten der modellbasierten Bildverarbeitung unter realen Bedingungen genutzt.

Die modellbasierte Segmentierung des Auges wird quantitativ ausschließlich an Kernspindaten getestet, da hier die Grauwertaupflösung der relevanten Weichgewebestrukturen außerordentlich gut ist, während bei den CT-Daten eine Abgrenzung des Bulbus vom umgebenden Gewebe kaum möglich ist. Wenngleich qualitative Tests auch für die CT-Daten durchgeführt wurden, ist davon auszugehen, daß die grundsätzlichen Stärken und Schwächen der modellbasierten Bildverarbeitung von den MRT-Daten auf andere Bildmodi wie die Computertomographie übertragen werden können.

6.1.1 Modelldatensätzen

Zur Untersuchung der grundsätzlichen Eigenschaften der modellbasierten Bildverarbeitung anhand von quantitativen Beschreibungsparametern wurden Bilddaten von räumlich klar definierten künstlichen Modellen angefertigt. Die quantitative Auswertung der Bildverarbeitungsergebnisse wird dadurch möglich, daß die realen Ausmaße mit den berechneten Werten verglichen werden können.

Die Umsetzung eines künstlichen Modells wird durch ein grundlegendes Problem erschwert. Bei der Kernspintomographie spielt vor allem der Wasserstoff für die Bildgebung eine entscheidende Rolle. Materialien, die ähnliche Wasserstoffkonzentrationen aufweisen wie die anatomischen Strukturen des Auges sind flüssig oder zumindest plastisch leicht verformbar. Zum Aufbau eines Modells werden jedoch vor allem rigide Stoffe benötigt, die in der Zeit von der Erstellung des Modells bis zur Anfertigung der Aufnahmen ihre Form nicht verändern. Ausführliche Recherchen konnten keine Materialien ausfindig machen, die gleichzeitig die gewünschten Darstellungseigenschaften aufweisen und eine gute Modellierung erlauben. Es mußte daher ein Kompromiß zwischen der Ähnlichkeit der Darstellungseigenschaften und den Modellierungseigenschaften gefunden werden.



Abbildung 6.1: Materialien zur Herstellung von Modelldatensätzen

Für die Modelldatensätze wurden die beiden Hälften von Hartplastikkugeln (links unten) mit Agar gefüllt. Zur Simulation der Linse wurde zuvor eine Gotcha-Kugel in eine der Hälften gelegt (linke Abbildung, kleine, gelbe Kugel in der Mitte). Nach dem Erkalten des Agar wurden die beiden Hälften zusammengefügt (linke Abbildung, Mitte). Da während des Erkaltes das Agar leicht schrumpft, füllt das Agar das Volumen der beiden Kugelhälften nicht vollständig aus. Dadurch entsteht beim Zusammensetzen der beiden Hälften ein Lufteinschluß an der Schnittfläche beider Kugelsegmente (Abbildung, rechts). Für die Aufnahmen wurden die Modellkugeln in ein Gefäß mit Öl gehängt (rechte Abbildung), um bei den MRT-Aufnahmen einen Kontrast zwischen dem Kugelinneren und -äußeren herzustellen.

Als einen akzeptablen Kompromiß hat sich der folgende Modellaufbau herauskristallisiert (siehe Abbildung 6.1): Zur Herstellung einer sphärischen Augenform wurden runde, hohle Plastikkugeln benutzt. Sie weisen zwar keine hellen Grauwerte wie die Sklera auf, aber Form und Größe des Modells sind exakt definiert. Es wurden fünf, sieben und acht Zentimeter große Kugeln ausgewählt. In die hohlen Kugeln wurde Agar gefüllt, um den Glaskörper zu simulieren. Damit das Agar in den Kernspindaten sichtbar wird, wurde es mit Nickelchlorid (NiCl_2) dotiert. Zum Füllen der Kugeln wurden zwei unterschiedliche Techniken angewandt. Bei dem ersten Verfahren wurden die Plastikkugeln in zwei Hälften zerlegt, um das heiße Agar jeweils separat einzufüllen. Da das Agar während des Erkaltes leicht schrumpft, füllte das Agar das Volumen der beiden wieder zusammengefügt Kugelhälften nicht vollständig aus. Das separate Füllen der beiden Hälften hatte somit den Nachteil, daß beim Zusammensetzen Luft mit eingeschlossen wurde (siehe Abbildung 6.1, rechts). Andererseits konnte dadurch sichergestellt werden, daß keine Lufteinschlüsse innerhalb der Kugelhälften auftraten (siehe Abbildung 6.2). Darüber hinaus konnte eine definierte Kante zur Orientierung innerhalb des Modells geschaffen werden, mit der die unterschiedlichen Schichtorientierungen identifiziert werden können.

Bei der zweiten Methode zum Füllen der Plastikkugeln wurden in die geschlossenen Plastikkugeln ein Loch gebohrt, durch das dann das flüssige Agar eingefüllt wurde. Diese zweite Technik erwies sich im Nachhinein als vorteilhafter, da dadurch der Lufteinschluß zwischen den beiden Kugelhälften vermieden werden konnte (siehe Abbildung 6.3).

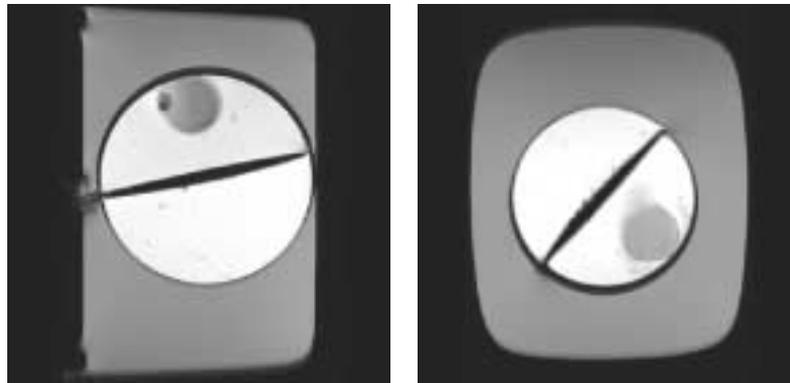


Abbildung 6.2: Zentraler Schnitt durch ein künstliches Modell

Links: sagittale Aufnahme des Modells. Rechts: axiale Aufnahme des Modells. Das Agar wird hell dargestellt. In der Mitte zwischen den beiden Hälften ist der Lufteinschluß zu erkennen. Am oberen Rand der Kugel ist die Gotcha-Kugel zu sehen. Aufgrund der temperatur-empfindlichen Gelatinehülle der Kugel kam es zu einem Lufteinschluß in der Kugel. Dieser zunächst unerwünschte Effekt kann den Umgang mit Artefakten in realem Bildmaterial testen. Sehr gut ist in den Abbildungen der Chemical Shift zu erkennen, der die Dicke der Plastikkugel ungleich groß erscheinen läßt. Dieser Effekt vergrößert in ähnlicher Form die Darstellung der Sklera bei klinischem Bildmaterial.

Zur Simulation der Linse wurden bei den ersten Modellen in das heiße Agar Gotcha-Kugeln gelegt. Sie weisen bei T1-Aufnahmen leicht dunklere Grauwerte auf als das Agar (siehe Abbildung 6.2). Wenngleich bei realem T1-MRT-Bildmaterial die Linse leicht heller als der Glaskörper dargestellt wird, stimmen die Grauwertdifferenzen der Modelldaten recht gut mit denen von realem Bildmaterial überein.

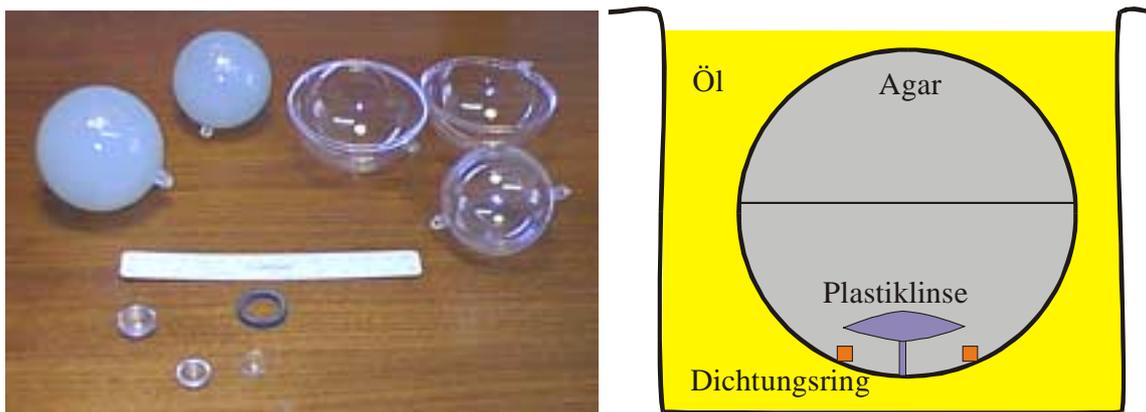


Abbildung 6.3: Materialien für die Herstellung eines künstlichen Modells mit einer Plastiklinse

Für die Konstruktion der weiteren Modelle wurden anstelle der Gotcha-Kugeln Linsen aus Plastik verwendet, um die Deformation durch das Aufweichen der Kugelhülle zu vermeiden. Außerdem wurden die Plastikkugeln nicht mehr separat gefüllt, sondern durch ein Loch, das in die Kugeloberfläche gebohrt wurde. Zusätzlich zu den MRT-Aufnahmen wurden von diesen Modellen auch CT-Aufnahmen angefertigt. Dabei unterstützt das Einbetten des Modells in einem Ölbad realitätsnahe Darstellungseigenschaften, da ähnlich wie bei realen Bildschichten des Menschen der Dichteunterschied zwischen umgebendem Öl und dem Agar im Bulbus vergleichsweise gering ist.

Einen Nachteil stellt die Konsistenz der Hüllen der Kugeln dar. Da sie aus Gelatine sind, werden sie während des Einfüllens des heißen Agars aufgeweicht. Dadurch kam es bei diesen Modellen zu einer Verformung der Kugeln, in einem Fall sogar zu einem Lufteschluß. Dieser Effekt konnte dazu genutzt werden, um die Stabilität des Verfahrens gegenüber Artefakten zu testen. Eine quantitative Auswertung wurde dadurch jedoch erschwert.

Bei den verbleibenden Modellen wurden anstelle der Gotcha-Kugeln Linsen aus Plastik verwendet. Dies hatte den Nachteil, daß die Darstellungseigenschaften der künstlichen Linsen deutlich von denen eines menschlichen Auges abwichen. Da jedoch eine plastische Verformung der Linse ausgeschlossen werden konnte, wurde eine quantitative Auswertung der Rekonstruktionsergebnisse für die Linse erst möglich. Die Linsen wurden auf einen ein Zentimeter langen Stift in der Kugel befestigt, damit zwischen ihnen und der Plastikhülle ein ausreichender Abstand vorhanden ist, der eine räumlich getrennte Abbildung gewährleistet (siehe Abbildung 6.3).

Zusätzlich zu der Plastiklinse wurden bei der zweiten Modellart zur Simulation der Iris Dichtungsringe konzentrisch zu der Linse verklebt. Von den Modellen mit Gotcha-Kugeln wurden lediglich sagittale und coronare MRT-Aufnahmen, von den Modellen mit Plastiklinsen zusätzlich noch CT-Aufnahmen angefertigt. Alle Modelle wurden während der Aufnahmen in einem Ölbad plaziert. Bei den MRT-Aufnahmen wurde durch dieses Ölbad ein Grauwertunterschied zwischen den Plastiklinsen und der Umgebung hergestellt. Dadurch konnte die Sklera einschließlich des üblichen Chemical Shift Artefakts nachgebildet werden (siehe Abbildung 6.2 und Abbildung 6.4). Bei den CT-Aufnahmen reduziert das Öl den Grauwertkontrast zwischen dem Glaskörper des Modells und dem umgebenden Volumen. Dadurch konnten die Darstellungseigenschaften von CT-Aufnahmen der menschlichen Orbita besser angenähert werden.

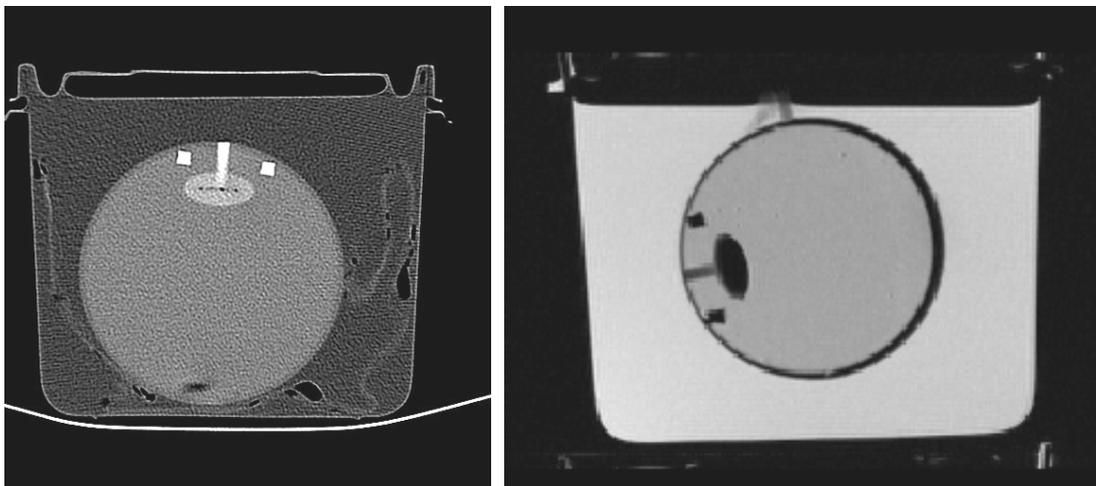


Abbildung 6.4: Beispielhafte CT- und MRT-Schichtbilder des künstlichen Modells mit Plastiklinse

Links ist eine beispielhafte CT-Aufnahme des künstlichen Modells im Ölbad zu sehen. Die weißen Vierecke repräsentieren den Dichtungsring, der die Iris simulieren soll. Rechts ist eine entsprechende Schicht des MRT-Datensatzes zu sehen. Linse, Dichtungsring und Plastikhülle werden dunkel abgebildet, da in Plastik kein beweglicher Wasserstoff vorhanden ist.

Die grundsätzlichen Zielsetzungen für die Auswertung der Modelldatensätze bestehen in der quantitativen und qualitativen Beurteilung der modellbasierten Segmentierung, Interpolation und Registrierung. Hierzu sollen die Ergebnisse der manuellen Segmentierung mit denen der semiautomatischen Verfahren verglichen werden. Besonderes Augenmerk soll dabei darauf gerichtet werden, inwieweit der dreidimensionale modellbasierte Segmentierungsansatz ein Vorteil gegenüber der weitgehend zweidimensionalen

manuellen Segmentierung von Bildschichten mit ausgeprägten Artefakten darstellt. Die quantitativen Ergebnisse mit den Modelldatensätzen sollen als ein Bewertungsrahmen genutzt werden, um die qualitativen Ergebnisse der verbleibenden Modelldatensätze besser einordnen zu können.

6.1.2 Datensätze von Schweineaugen

Um eine genaue Kenntnis der dargestellten Objekte mit einer realistischen Aufnahmesituation zu verbinden, wurden Datensätze von Schweineaugen angefertigt, die vor den Aufnahmen ausgemessen wurden. Die Verwendung von Schweineaugen als Modell bietet sich an, da Schweineaugen in ihren Ausmaßen und ihrer Form weitgehend mit dem menschlichen Auge übereinstimmen. Durch den Vergleich der Bildverarbeitungsergebnisse mit den zuvor ausgemessenen Ausmaßen können quantitative Analysen an realitätsnahen Bilddaten vorgenommen werden. Ähnlich wie für die Modelldaten wurden von den Schweineaugen sagittale, coronare und axiale Datensätze angefertigt. Aufnahmeparameter wie beispielsweise Aufnahme-sequenz oder Auflösung wurden korrespondierend zu den Aufnahmen von tatsächlichen Patienten eingestellt.

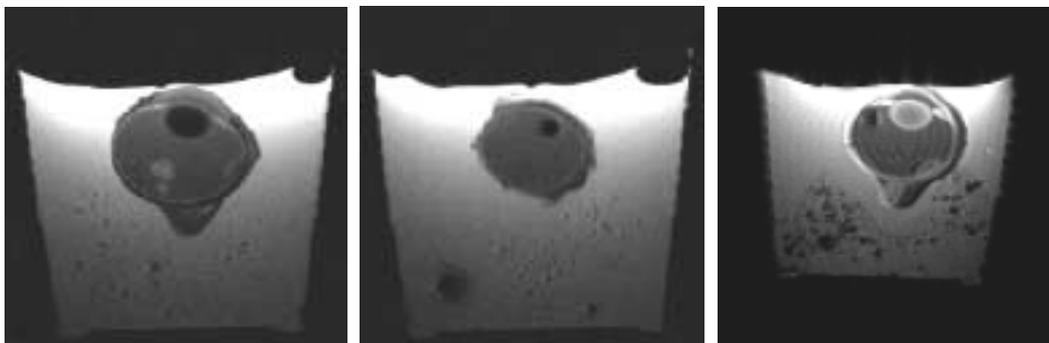


Abbildung 6.5: Testdaten von Schweineaugen

Die Abbildungen zeigen verschiedene T1- und T2-Schnittbilder der Schweineaugen, die in einem Ölbad auf einem Schwamm platziert wurden. Links ist ein zentraler Schnitt durch das Auge mit den Tantal-Clips zu sehen. Im unteren Bereich des Bulbus sind helle Punkte zu sehen. Sie werden durch die Creme verursacht, die zur Simulation eines Tumors in den Bulbus injiziert wurde. In der Mitte ist ein Schnitt durch eine Randschicht des Bulbus zu sehen. Der dunkle Punkt rechts oben im Bulbus der mittleren Abbildung ist der Keramik-Clip, der vor den Aufnahmen auf die Bulbusoberfläche aufgenäht wurde. Die Tatsache, daß der Clip innerhalb des Bulbus zu sehen ist, veranschaulicht die Artefakte, die durch die Feldinhomogenitäten um die Clips herum verursacht werden und das Erscheinungsbild der Clips deutlich vergrößern. Rechts ist eine T2 MRT-Aufnahme des Auges zu sehen, bei der die Linse nicht dunkel wie bei den T1-Aufnahmen, sondern hell dargestellt wird.

Um die Rahmenbedingungen für die Bildverarbeitung der Schweineaugen möglichst nah an die realen Bedingungen anzupassen, wurden die Schweineaugen am Morgen der Aufnahmen frisch vom Schlachter abgeholt. Vor Anfertigung der MRT-Bilder wurden auf ein Schweineauge Tantal-Clips, auf das andere Auge Keramik-Clips aufgenäht, um die Aufnahmesituation bei der Strahlentherapie von Augentumoren möglichst realitätsnah zu rekonstruieren. Anschließend wurden diverse Distanzen zwischen wichtigen anatomischen Strukturen, wie beispielsweise Kornea oder Sehnervansatz und den Clips ausgemessen, um die entsprechenden Distanzen mit dem Bildverarbeitungssystem zu verifizieren.

Bezüglich der Auswertung dieser Daten muß jedoch erwähnt werden, daß das Bildmaterial durch einige Faktoren verfälscht wird und eine quantitative Beurteilung der Ergebnisse nur unter Vorbehalt möglich ist. Obwohl die Augen kurz vor der Anfertigung der Aufnahmen entnommen wurden und daher sehr frisch waren, muß davon ausgegangen werden, daß sich die Form der Augen zwischen den verschiedenen Aufnahmen veränderte. Diese Formänderungen resultieren vor allem aus der physiologischen Verformbarkeit des Auges, aus der Verdunstung der Zellflüssigkeiten und aus einer nekrotischen Modifikation der

Zellstrukturen. Um diese Veränderungen zu minimieren, wurde kontinuierlich Formaldehyd in das Auge injiziert, wodurch das Auge stets mit einer gewissen Spannung gefüllt war. Dadurch konnte gegebenenfalls das Volumen des Auges stabilisiert werden. Bei der Bestimmung der Distanzen ist jedoch aufgrund der plastischen Verformbarkeit des Auges mit einer Variation der Distanzen in einem Bereich von fünf bis zehn Prozent zu rechnen.

Darüber hinaus existierten bei den Schweineaugen keine anatomischen Referenzpunkte, die für eine zuverlässige Distanzmessung genutzt werden könnten. Beispielsweise hat die Messung der Distanzen vom Ansatzpunkt des Sehnervs zu anderen Referenzpunkten gezeigt, daß der um den Sehnerv angesiedelte Fettansatz eine plastisch verformbare Barriere darstellte, die ein präzises Anlegen der Schiebelehre zwischen dem Sehnerv und den anderen Landmarken nahezu unmöglich machte. Andererseits wurde dieser Fettansatz in den Bilddaten nur ansatzweise und kaum erkennbar dargestellt. Im Bildmaterial durchgeführte Distanzmessungen können demnach diesen Fettansatz nicht mit einbeziehen, wodurch bereits eine nicht unbeträchtliche Unsicherheit bezüglich der korrekten Lokalisation der Meßpunkte verursacht wurde.

Der Sinn der Datensätze von den Schweineaugen besteht daher in erster Linie in einer groben Verifikation der Bildverarbeitungsalgorithmen. Dabei soll abgeschätzt werden, inwieweit die Ergebnisse der Modelldaten auf reales Bildmaterial übertragen werden können.

6.1.3 Probandendaten

Während die Modelldatensätze und die Schichtdaten von den Schweineaugen in erster Linie zur quantitativen Abschätzung der Eigenschaften der modellbasierten Bildverarbeitung angefertigt wurden, dienen Datensätze von freiwilligen Probanden vor allem zur Analyse von spezifischen Problematiken der Bildverarbeitungsalgorithmen. Hierzu gehören vor allem die folgenden Fragestellungen:

- Wie tolerant ist die modellbasierte Bildverarbeitung gegenüber Unterschieden in der Auflösung, dem Bildausschnitt und der Schichtorientierung?
- Wie wirken sich Unterschiede in Form oder Orientierung des Auges im Bildmaterial auf die Ergebnisqualität bei einer gegenseitigen Nutzung der Daten als Referenzmodell oder aktuellem Datensatz aus?

Um diese Aspekte der modellbasierten Bildverarbeitung auszutesten, wurden von zwei Freiwilligen T1-MRT-Datensätze angefertigt. Dazu gehörten Aufnahmen des ganzen Schädels mit axialer und sagittaler Ausrichtung und hochaufgelöste Aufnahmen des Auges in den Orientierungen axial, coronar und sagittal (vergleiche Abbildung 6.6).

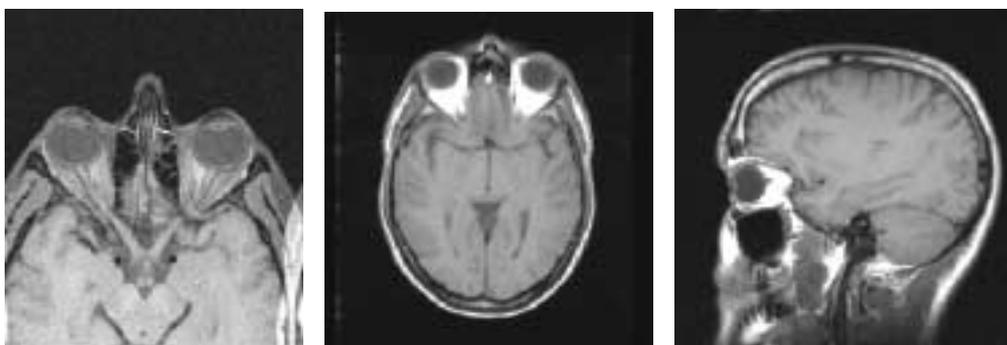


Abbildung 6.6: Aufnahmen von freiwilligen Probanden

Um die oben erwähnten Fragestellungen austesten zu können, wurden von freiwilligen Probanden Datensätze mit unterschiedlichen Aufnahmeparametern aufgenommen. Modifiziert wurden die Auflösung, der Bildausschnitt und die Schichtorientierung.

Um die Abhängigkeit der Darstellungseigenschaften der Orbita von der Form des Auges zu untersuchen, wurden ein normalsichtiger und ein weitsichtiger Proband ausgewählt. Diese unterschiedliche Sehfähigkeit kann in den Abbildung 6.6 und Abbildung 6.7 nachvollzogen werden. Indem die Datensätze der Probanden untereinander als Referenzmodell und aktuellem Datensatz angewandt werden, kann der Einfluß dieser unterschiedlichen Augenformen getestet werden.

Ein weiterer Einflußparameter, der durch die bisher beschriebenen Datensätze nicht abgedeckt wurde, ist die Ausrichtung des Auges innerhalb der Datenschichten. Zu diesem Zweck wurden die axialen T1-MRT-Aufnahmen der Probanden einmal mit gerade ausgerichteten Augen (Abbildung 6.6, links) und einmal mit stark nach oben ausgerichteten Augen aufgenommen (siehe Abbildung 6.7, rechts und zweite von rechts). Während bei dem ersten Datensatz in der zentralen Schnittebene durch das Auge Linse und Kornea gut abgezeichnet werden, sind diese Strukturen in der zentralen Schicht des Datensatzes mit nach oben orientierten Augen nicht erkennbar. Statt dessen werden Linse und Korneawölbung in einer deutlich höheren Schicht abgebildet, wobei die Darstellung durch die Partialvolumeneffekte beeinträchtigt wird. Von Interesse ist in diesem Zusammenhang, wie hoch die Toleranz der modellbasierten Bildverarbeitung gegenüber solchen Unterschieden ist.

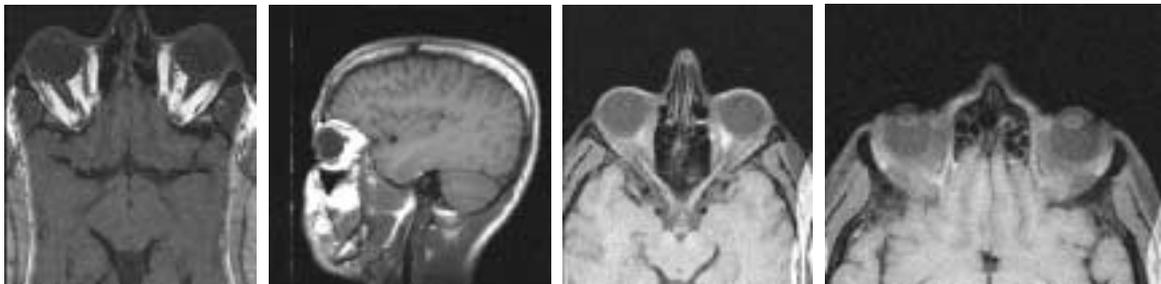


Abbildung 6.7: Datenmaterial von Probanden mit variierenden Aufnahmeparametern

Links: Auf der hochauflösenden axialen T1-MRT-Aufnahme des zweiten Probanden kann gut im Vergleich mit den Aufnahmen des normalsichtigen Probanden in Abbildung 6.6 die Weitsichtigkeit anhand der veränderten Form des Auges erkannt werden. Interessant ist bei der Durchführung der Tests, inwieweit solche Unterschiede die Bildverarbeitung beeinträchtigen. Außerdem soll die Toleranz gegenüber Unterschieden zwischen Referenzmodell und aktuellem Datensatz abgeschätzt werden, indem Datensätze mit unterschiedlichen Augenorientierungen aufgenommen werden. Neben den axialen Aufnahmen der Probanden mit geradeaus ausgerichteten Augen wurden zu diesem Zweck Aufnahmen angefertigt, bei denen die Probanden stark nach oben schauen (Abbildung rechts und zweite von rechts).

6.1.4 Medizinische Datensätze

Um außerdem die Eigenschaften der modellbasierten Bildverarbeitung an realem Bildmaterial zu testen, wurden mehr als dreißig Datensätze aus der Routinetherapie von Augentumoren einbezogen. Die Beurteilung der Bildverarbeitungsalgorithmen wird erreicht, indem die berechneten Ergebnisse mit Musterlösungen verglichen werden. Die Musterlösungen bestehen für die Tests der modellbasierten Segmentierung aus medizinisch korrekten Segmentierungen dieser Testdatensätze, die zuvor durch eine kompetente Arbeitskraft manuell angefertigt wurden. Ein Datensatz wird als medizinisch korrekt segmentiert angesehen, wenn die Segmentierung zuvor definierte Segmentierungsrichtlinien erfüllt.

Zur Definition dieser Segmentierungsrichtlinien wurden vier der Datensätze von einem erfahrenen Augenarzt der Augenklinik des Universitätsklinikums Benjamin Franklin manuell segmentiert. Diese Segmentierungen wurden durch einen Mitarbeiter beobachtet und die Entscheidungsgrundlagen für das Einzeichnen der Konturen vom Arzt erfragt. Durch die abstrakte Beschreibung einer in medizinischer Hinsicht korrekten Segmentierung und die sorgfältige Anfertigung der Mustersegmentierungen wurden die Segmentierungsrichtlinien in Form von Segmentierungsanweisungen und den beispielhaften Segmentie-

rungen festgehalten. Die verbleibenden Datensätze wurden von den Projektmitarbeitern entsprechend dieser Richtlinien und Vorlagen manuell segmentiert.

Die auf diese Weise angefertigten manuellen Segmentierungen werden für die anschließend durchgeführten Tests als Goldstandard angenommen, da sie als zuverlässigste Quelle für eine korrekte Segmentierung anzusehen sind. Es muß jedoch erwähnt werden, daß dieser Goldstandard nicht für quantitative Analysen genutzt werden kann. Wie nachfolgende Testergebnisse veranschaulicht werden, sind weder die abstrakten Beschreibungen noch die explizit durchgeführten Mustersegmentierungen konsistent oder reproduzierbar.

Bereits bei der Diskussion der manuellen Segmentierungsrichtlinien mit dem Arzt stellte sich heraus, daß die Segmentierung nicht nur ein Problem für den später angewandten Segmentierungsalgorithmus darstellt. Vielmehr bestätigten sich bereits bei der Definition einer korrekten Segmentierung und ihrer Umsetzung an den Datensätzen die systematischen Probleme der manuellen Segmentierung, die bereits durch eine Vielzahl von Veröffentlichungen erwähnt wurden. Bei der Bearbeitung des vorliegenden Bildmaterials fielen vor allem die folgenden Probleme der manuellen Segmentierung auf:

- Die manuelle Segmentierung von medizinischen Organen erfordert eine hohe Konzentration des segmentierenden Fachpersonals. Andererseits kann kaum ein Mensch seine Konzentration über einen längeren Zeitraum hinweg auf dem gleichen hohen Niveau halten. Erschwert wird das Aufrechterhalten der Konzentration außerdem durch den stark repetitiven Charakter einer manuellen Segmentierung. Da für die Strahlentherapie umfangreiche MRT- und CT-Datensätze segmentiert werden müssen, werfen die hohen Anforderungen an die Präzision der Segmentierung, der erforderliche Arbeitsaufwand und die Konzentrationsfähigkeit des Menschen einen Widerspruch auf. Sofern die segmentierenden Mitarbeiter nicht die Gelegenheit für eine Vielzahl von Pausen erhalten, ist davon auszugehen, daß manuelle Segmentierungen konzentrationsbedingt deutlichen intra- und interindividuellen Schwankungen unterliegen.
- Die Ergebnisse der manuellen Segmentierung werden maßgeblich durch die Kompetenz, den Erfahrungshorizont und durch individuelle Interpretationsschemen der segmentierenden Person bestimmt. Dabei unterliegt vor allem der individuelle Interpretationsspielraum beträchtlichen Schwankungen. Solche Schwankungen treten sowohl interindividuell auf, aber auch für ein und dieselbe Fachkraft über die Zeit hinweg. Als wesentlicher Grund für diesen subjektiven Interpretationsspielraum sind die komplexen und nicht quantifizierbaren Darstellungseigenschaften des Bildmaterials anzusehen. Selbst bei voller Konzentration ist eine eindeutige Interpretation oft nicht möglich. Dadurch wirken sich unter anderem auch nicht-quantifizierbare Faktoren wie Stimmung und momentanes Befinden der Arbeitskraft auf die Interpretation der Bildstrukturen und damit auf die Segmentierung der Organe aus. Hinzu kommt, daß der Mensch nicht wie eine Maschine auf reproduzierbare Handlungen ausgelegt ist. Eine nicht unbeträchtliche Streuung ist daher unvermeidlich.

Diese Schwankungen bei der Bildinterpretation wurden bereits bei der Segmentierung der Datensätze durch die ophthalmologische Fachkraft offenkundig. Selbst innerhalb eines Datensatzes wurden Organgrenzen mit starken Partialvolumeneffekten sehr unterschiedlich segmentiert (siehe Abbildung 6.8). Auf Nachfragen wurden einzelne Segmentierungen vom Arzt teilweise oder gänzlich modifiziert. Beobachtungen zeigten, daß es eine lange Liste von menschlichen Einflüssen gibt, die zu einer situativen, nicht reproduzierbaren Segmentierung führen. Eine vollständige Elimination dieser Faktoren erscheint dem Autor sowohl bei umfassenden Segmentierungsarbeiten, wie sie für die Strahlentherapie von Augentumoren erforderlich ist, als auch bei der Anfertigung der Musterlösungen nicht möglich zu sein.

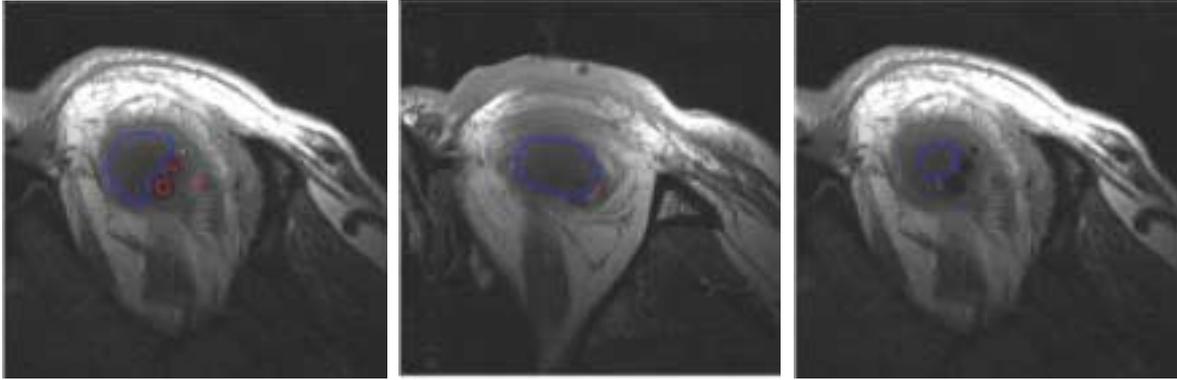


Abbildung 6.8: Subjektiver Spielraum bei der Segmentierung der medizinischen Bilddaten

Bei der Anfertigung der Mustersegmentierungen zeigte sich, daß ein sehr kritischer Punkt die Definition einer korrekten Segmentierung ist. Selbst innerhalb eines Datensatzes legte die ophthalmologische Fachkraft unterschiedliche Kriterien und Richtlinien für die Segmentierung der Bulbus-Kontur zugrunde. Während bei der linken Abbildung die Segmentierungskontur des Bulbus nur innerhalb des dunkelsten Bereichs liegt und die Marker nicht mit einschließt, wurde in der mittleren Abbildung ein weitergehender Bereich segmentiert. Der Arzt wurde bei der linken Abbildung maßgeblich durch die Darstellung der Clip-Artefakte beeinflusst. Wenngleich klar ist, daß nicht wirklich die Clips, sondern lediglich die Ausläufer der Artefakte dargestellt werden, beeinflussten sie die Segmentierung der betroffenen Schichten maßgeblich. Wiederum andere Ergebnisse erhielt man bei der Anfertigung der Musterlösungen, wenn bei der Segmentierung außerdem die dreidimensionale Form des Auges berücksichtigt wurde. Bei der ganz rechten Abbildung verglich der Benutzer die Segmentierung in der dargestellten Schicht mit den Segmentierungskonturen der angrenzenden Schichten. Nach einem rekursiven Korrigieren der Segmentierungskontur zeichnete der Benutzer letztendlich die Bulbus-Kontur deutlich kleiner als in den ersten Segmentierungsschritten ein.

- Segmentierungen, die ausschließlich in den zweidimensionalen Originalschichten angefertigt wurden, wiesen bei sämtlichen Testdatensätzen in der dreidimensionalen Rekonstruktion grundlegende Defizite auf. Die Betrachtung der groben Oberflächenformen nach einer dreidimensionalen Rekonstruktion der zweidimensionalen Segmentierungskonturen machten vor allem Segmentierungsfehler in Schichten mit starken Partialvolumeneffekten offensichtlich (siehe Abbildung 6.9, links). Die Einbeziehung der dreidimensionalen Oberfläche führte in nahezu allen Fällen zu grundlegenden Modifikationen der Segmentierung. Dementsprechend wurden bei der Segmentierung der Musterlösungen die zweidimensionalen Segmentierungsstrategien durch dreidimensionale Korrekturmechanismen ergänzt, um die medizinische Korrektheit der Segmentierungen zu erhöhen. Diese Fehler bei der Anfertigung der Mustersegmentierungen anhand der Originalschichten zeigten, daß eine rein zweidimensionale manuelle Segmentierung aufgrund der Diskrepanz zwischen den in den medizinischen Datensätzen dargestellten Strukturen und der tatsächlichen individuellen Anatomie zu inakzeptablen Fehlern führt.

All diese Nachteile der manuellen Segmentierung wurden von Anfang an auch durch die medizinische Fachkraft wahrgenommen. Dementsprechend äußerte die medizinische Fachkraft großes Interesse an einem konsistenten, automatisierten Segmentierungsverfahren. Eine rein manuelle Segmentierung wurde für die klinische Routine als inakzeptabel abgelehnt. Für die modellbasierte Bildverarbeitung stellen diese Feststellungen einerseits eine wichtige Rechtfertigung dar. Andererseits betonen sie die Wichtigkeit einer möglichst sorgfältigen Segmentierung der Datensätze zur Anfertigung der Referenzmodelle.

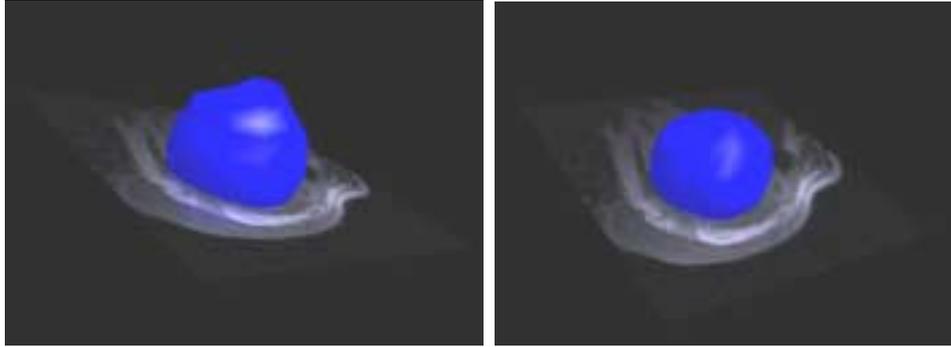


Abbildung 6.9: Berücksichtigung der dreidimensionalen Form bei der Segmentierung

Die Tests zeigten, daß eine zweidimensionale manuelle Segmentierungsstrategie zu grundsätzlich fehlerhaften dreidimensionalen Organoberflächen führt. Dies verdeutlicht, daß die zweidimensionalen Schnittbilder durch die systemimmanenten Darstellungsfehler so weit von der Realität abweichen, daß eine rein zweidimensionale Segmentierung nicht sinnvoll erscheint. Gerade in den Bereichen mit starken Partialvolumeneffekten muß festgestellt werden, daß selbst eine erfahrene Arbeitskraft keine korrekte Segmentierung auf Anhieb vornehmen kann (siehe linke Abbildung). Erst wenn bei der Segmentierung die dreidimensionale, globale Form des segmentierten Organs mit einbezogen wird, kann eine Segmentierung angefertigt werden, die auch dreidimensional plausibel erscheint.

Doch selbst eine noch so sorgfältige Segmentierung wird den subjektiven Interpretationsspielraum nicht vollständig eliminieren können. Eine quantitative Auswertung der Datensätze macht daher nur mit deutlichen Einschränkungen Sinn. Um nichtsdestotrotz eine verlässliche qualitative Auswertung zu ermöglichen, wurden für die medizinischen Datensätze die folgenden Segmentierungsrichtlinien festgelegt:

- In den Referenzmodellen werden die Organstrukturen Bulbus, Linse, Ziliarkörper, Iris und Sehnerv segmentiert.
- Der Bulbus wird an der Sklera-Außenkante und der Kornea-Außenkante segmentiert. In den Randbereichen mit starken Partialvolumeneffekten werden die Segmentierungskonturen so gestaltet, daß die dreidimensionale Oberflächenform möglichst plausibel die sphärische Form des Auges fortsetzt. Eine Segmentierung der Bulbus-Kontur auf der Basis des Bildmaterials erscheint in diesem Bereich aufgrund der augenscheinlich stark verfälschten Darstellung nicht sinnvoll zu sein. Vollständige Gewißheit über die Korrektheit dieser Segmentierungsstrategie gibt es nicht, da vor allem bei stark kurzsichtigen Patienten die Form des Bulbus beträchtlich von der sphärischen Form abweichen kann. Eine definitive Unterscheidung von Darstellungsproblemen und tatsächlichen Formanomalien ist dementsprechend nicht möglich.
- Die Segmentierung der Linse setzt an den erkennbaren Grauwertkanten auf. Besondere Aufmerksamkeit wird der Abgrenzung der Linse von der Iris beigemessen. Aufgrund der geringen Ausmaße beider anatomischer Strukturen kommt es oft zu einer Verschmelzung der Darstellung in den Randbereichen, die ohne entsprechendes anatomisches Vorwissen nicht erkannt werden kann. Bei der Segmentierung der Datensätze durch die medizinische Fachkraft konnte festgestellt werden, daß gerade bei den kleinen, nur schwer erkennbaren Organstrukturen das anatomische Vorwissen maßgeblich das Ergebnis der Segmentierung bestimmte. Während ein anatomisch unbedarfter Nutzer in zentralen Schichten Linse und Iris nicht voneinander unterschied und dadurch nicht-konvexe Konturen segmentierte, achtete der Ophthalmologe darauf, eine konvexe Linsenform einzuhalten (siehe Abbildung 6.10). Dieses Verhalten der Fachkraft resultiert aus dem Wissen, daß die Linse in jedem Fall konvex ist.

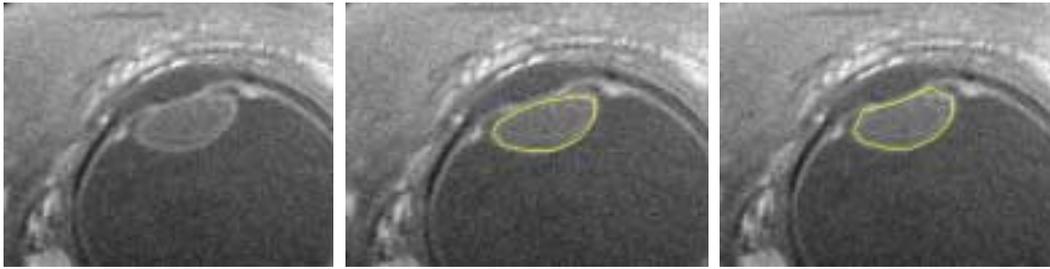


Abbildung 6.10: Beispiele für die manuelle Segmentierung der Linse

Die Abbildungen zeigen eine typische Situation, in der unterschiedliche Segmentierungen zwischen medizinischem und nicht-medizinischem Nutzer auftreten. Als korrekt wird die mittlere Segmentierung definiert, bei der eine konvexe Form der Linse eingehalten wird. Ein Mitarbeiter ohne die entsprechenden anatomischen Vorkenntnisse hatte die Linse wie in der rechten Abbildung segmentiert.

Darüber hinaus hat sich ähnlich wie bei der Segmentierung des Bulbus gezeigt, daß in den Randbereichen die Linse durch die Partialvolumeneffekte deutlich größer dargestellt wird, als es bei dreidimensionaler Begutachtung der Linsenoberfläche möglich ist. Für die Segmentierung der Referenzmodelle wird daher die runde Linsenform durch Berücksichtigung der dreidimensionalen Linsenoberfläche sichergestellt.

- Iris und Ziliarkörper werden ähnlich wie die Linse entsprechend der menschlichen Anatomie segmentiert. Dabei ist der Unterschied zwischen dem, was im Bildmaterial dargestellt wird und dem, wie segmentiert werden muß, deutlich größer als bei anderen Organen. Nach Vorgabe durch die ophthalmologische Fachkraft wird die Iris als eine kreisförmige Scheibe mit einem Loch in der Mitte, der Pupille, segmentiert. Der Ziliarkörper ist ebenfalls ein rotations-symmetrischer Körper, mit einem Loch in der Mitte. Die Zonularfasern werden nicht segmentiert, da sie im Bildmaterial nicht gesehen werden können. Eine Abgrenzung von Ziliarkörper und Sklera ist anhand des Bildmaterials nicht möglich.

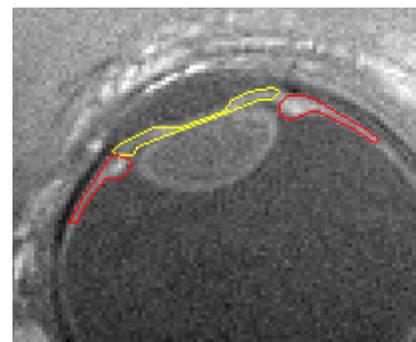
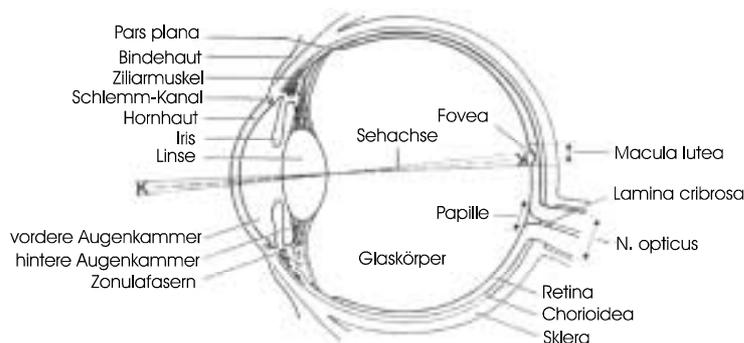


Abbildung 6.11: Segmentierung der Iris und des Ziliarkörpers

Die linke Abbildung zeigt die Anatomie von Iris und Ziliarkörper, wie sie in Lehrbüchern beschrieben wird. Die rechte Abbildung zeigt, wie diese Vorgaben bei der Segmentierung der Referenzmodelle einfließt.

Dr. Bechrakis, die ophthalmologische Fachkraft, hat aufgrund seiner Erfahrungen und Kenntnisse beim Ziliarkörper den Fortsatz und einen Teil der Sklera segmentiert. Wie weit der Sklera-Fortsatz geht, ist schwer zu definieren und muß einer individuellen Einschätzung überlassen werden. Die Referenzmodelle werden vergleichbar zur Abbildung 6.11 segmentiert. Ein sehr großes Problem stellt die Segmentierung beider Organe in den Randbereichen dar. Durch die Partialvolumeneffekte werden die Strukturen vollständig verwischt, wodurch

eine Rekonstruktion der Organe in diesem Bereich unmöglich wird. Dieser Effekt unterstreicht den Ansatz, zur Bearbeitung eines aktuellen Datensatzes nur solche Referenzmodelle zu verwenden, die in der grundsätzlichen Schichtorientierung übereinstimmen.

- Die Segmentierung des Sehnervs wird zumindest in den zentralen Schichten anhand der sichtbaren Grenzen vorgenommen. In den Randbereichen des Sehnervs verfälschen die Partialvolumeneffekte ähnlich wie bei allen anderen Organen die Darstellung der Organgrenzen. So scheinen die Konturen deutlich größer zu sein, als sie bei dreidimensionaler Betrachtung der Segmentierung sein können (vergleiche Abbildung 6.12). Bei der Segmentierung der Referenzmodelle werden daher lediglich die zentralen Schnittbilder entlang der dargestellten Kanten segmentiert. Bei den oberen und unteren Schichten wird unter Berücksichtigung der dreidimensionalen Form des Sehnervs eine möglichst runde Oberfläche angestrebt.

Durch die bei allen Organen auftretenden, grundlegenden Verfälschungen der dargestellten Konturen durch die Partialvolumeneffekte wurde die Notwendigkeit einer Modifikation des bisher beschriebenen modellbasierten Segmentierungsansatzes offensichtlich: Es muß zwischen Bildschichten unterschieden werden, in denen die Organkontur korrekt segmentiert werden kann und jenen Bildschichten, in denen eine korrekte Segmentierung aufgrund der Partialvolumeneffekte nicht möglich ist. Bei dem bearbeiteten Bildmaterial hat sich gezeigt, daß die oberste und unterste Bildschicht eines Organs in keinem Fall intuitiv segmentiert werden konnte. Ein zweidimensionaler Segmentierungsalgorithmus, wie es der in dieser Arbeit beschriebene modellbasierte Snake noch ist, kann in solchen Fällen aller Voraussicht nach keine korrekten Ergebnisse produzieren.

Eine mögliche Herangehensweise zur Lösung dieses Problems besteht darin, daß diese Schichten von der Segmentierung ausgespart werden. Da die Konturform bereits durch die Transformation der Referenzsegmentierung in den aktuellen Datensatz recht gut an die Organkonturen angepaßt wurde, ist nur ein geringer Fehler zu erwarten. Aller Voraussicht nach sind die Ergebnisse sogar besser als bei einer manuellen Segmentierung.

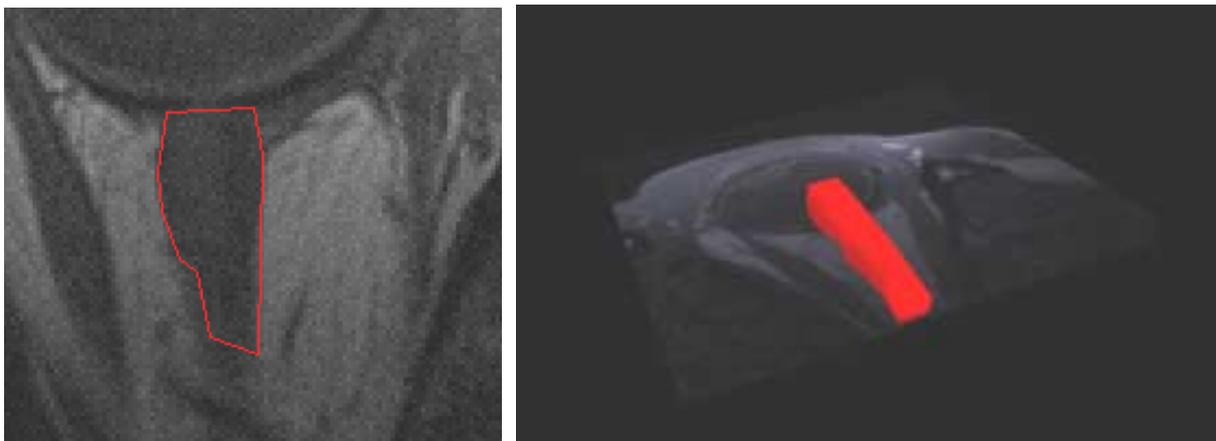


Abbildung 6.12: Verfälschung der Segmentierung des Sehnervs durch die Partialvolumeneffekte

Links: Werden die Organgrenzen des Sehnervs intuitiv im Bereich der stärksten Grauwertunterschiede eingegeben, stellt sich nach einer dreidimensionalen Rekonstruktion schnell heraus, daß die Kontur im Bildmaterial zu groß dargestellt wird. Die Folge ist, daß der dreidimensional rekonstruierte Sehnerv zu breit und sehr eckig erscheint (siehe rechte Abbildung). Es muß daher auch bei einem solch einfach zu segmentierenden Organ auf umfassende Erfahrungen zurückgegriffen und stets die dreidimensionale Rekonstruktion des segmentierten Objektes mit einbezogen werden.

Darüber hinaus besteht die Möglichkeit, die Segmentierung des Auges als eine sequentielle, aufeinander aufbauende Segmentierung umzusetzen. Es würde sich beispielsweise anbieten, die Anpassung und

Segmentierung zunächst nur auf den Bulbus zu beschränken. Anhand der Ergebnisse könnte dann eine recht präzise Lokalisation der angrenzenden Strukturen wie Linse, Iris, Ziliarkörper und Sehnerv vorgenommen werden. Da die Oberflächen- und Darstellungseigenschaften dieser Konturen deutlich von der des Bulbus abweichen, wären hierfür angepaßte Vorgehensweisen bei der modellbasierten Segmentierung notwendig. Gerade der Sehnerv müßte bei stark gekrümmter Form als Schlauch modelliert werden. Eine Anpassung des Sehnerv-Referenzmodells an die Daten, so wie sie in dieser Arbeit beschrieben wurde, wäre nicht in jedem Fall optimal. Eine solche Erweiterung liegt aufgrund des großen Aufwandes außerhalb dieser Arbeit und muß durch weitergehende Forschungsarbeiten abgedeckt werden. Für die modellbasierte Segmentierung der medizinischen Datensätze werden daher diese Unterstrukturen bei der globalen Anpassung des Referenzmodells an den aktuellen Datensatz berücksichtigt. Für die lokale Anpassung der Segmentierungskonturen wird jedoch allein die Segmentierungskontur des Bulbus berücksichtigt und auf die verbleibenden Strukturen übertragen.

Von der modellbasierten Segmentierung und Rekonstruktion der Gesichtsoberfläche wurde ebenfalls abgesehen, da die Ausschnitte des dargestellten Gesichts sehr unterschiedlich sind und damit die Verwendung eines Modells nur bedingt sinnvoll wäre. Darüber hinaus kann der Luftbereich außerhalb des Körpers problemlos mit einem Region Growing Algorithmus bestimmt werden und erfordert demnach keine aufwendige modellbasierte Bildverarbeitung.

Die segmentierten Datensätze wurden einerseits als Referenzmodelle genutzt, andererseits wurden aus ihnen die zu segmentierenden Datensätze rekrutiert. Diese Vorgehensweise resultiert einerseits aus der Tatsache, daß die medizinischen Arbeitskräfte nur sehr eingeschränkt für die Bearbeitung der Bilddaten in Anspruch genommen werden konnten. Andererseits ist gerade für die quantitative Auswertung der Segmentierungsergebnisse eine Mustersegmentierung notwendig. Um eine quantitative Auswertung der Datensätze trotz des möglichen Interpretationsspielraums möglich zu machen, wurden sämtliche Datensätze von einer Person segmentiert. Dabei wurden die Segmentierungen nach den bereits aufgeführten Richtlinien angefertigt. Würden die Segmentierungen aufgrund von interindividuellen Interpretationsunterschieden stark divergieren, wäre eine qualitative Auswertung der Segmentierungsergebnisse nicht möglich.

Eine quantitative Auswertung der modellbasierten Interpolation und Registrierung erschien nicht sinnvoll, da keine Vergleichswerte für eine Bestimmung der Abweichung von den erwarteten Werten vorlagen. Nichtsdestotrotz wurden die entsprechenden interpolierten Datensätze berechnet und eine Registrierung von ausgewählten Datensätzen vorgenommen, um anschließend die Meinung einer medizinischen Fachkraft zu erfragen.

6.2 Die Programmiersprache Java

Wenngleich sich die Verwendung von Java für die Umsetzung der Software insgesamt als sinnvoll und richtig herausgestellt hat, konnten einige Nachteile von Java während der Programmierarbeiten festgestellt werden. Negativ hat sich der in einigen Bereichen noch nicht ausgereifte Entwicklungsstand der Basisklassen von JDK 1.2 ausgewirkt. Die größten Probleme traten in diesem Zusammenhang bei der Speicherverwaltung auf. Nach dem Schließen von zuvor geladenen Datensätzen wurde der belegte Speicherplatz nicht mehr freigegeben. Bei der modellbasierten Segmentierung von großen Datensätzen führte dies bei Rechnern mit wenig Arbeitsspeicher während der schichtweisen Anwendung des modellbasierten Snake zu Systemabstürzen.

Eine entscheidende Rolle scheint in diesem Zusammenhang das Handling der Action-Events zu spielen. Wurden Bilddaten mit einem Action Event beispielsweise zur Aktualisierung der Visualisierung oder von gezeichneten Objekte verbunden, wird die Freigabe des Speicherplatzes für diesen Datensatz verhindert, selbst wenn der Datensatz geschlossen und alle Daten innerhalb der Hauptklassen auf Null gesetzt wurden. In anderen Programmiersprachen wie beispielsweise Delphi ist die Freigabe des Speichers auch

trotz solcher übriggebliebener Systemlinks möglich. Da die Benutzerinteraktionen innerhalb einer Bildverarbeitungssoftware schnell sehr komplex werden und viele Action Listener notwendig sind, können beim Freigeben von Klassen leicht einige Verbindungen zu Action Listnern übersehen werden. Ein sauberes und möglichst zentrales Management der Action Events erscheint daher für zukünftige Arbeiten sinnvoll zu sein, um dieses Problem besser angehen zu können.

Darüber hinaus scheinen aber auch einige Speicherprobleme entweder durch Fehler in JDK 1.2 oder durch die Kapselung von Java in Windows NT/98 verursacht worden zu sein. Da einige Speicherprobleme mit Windows 2000 nicht mehr auftraten, ist auch nicht auszuschließen, daß die Probleme von Microsoft durch eine entsprechende Implementation der Betriebssysteme provoziert wurden. Kurzfristig konnte diesem Problem durch eine Aufstockung des Arbeitsspeichers begegnet werden. Bei einem Arbeitsspeicher von 256 MByte traten keine Systemabstürze mehr aufgrund von Speichermangel auf.

Sehr positiv wirkte sich die Plattformunabhängigkeit von Java aus. Dadurch konnte die Software ohne zusätzlichen Arbeitsaufwand auf mehrere Betriebssysteme wie Windows und Linux portiert werden. Da die Basisklassen für andere Systeme wie Macintosh oder DEC Alpha erst mit einer gewissen Verspätung freigegeben werden, waren Tests auf diesen Plattformen nicht möglich. Es kann jedoch davon ausgegangen werden, daß nach der Veröffentlichung der entsprechenden Basisklassen der Übergang ebenso problemlos möglich ist wie der Wechsel zwischen Windows und Linux. Auch bezüglich der Performance von Java konnten aufgrund der immer schnelleren Computer keine Nachteile gegenüber anderen Programmiersprachen mehr festgestellt werden.

Die größten Vorteile der Programmiersprache Java sind jedoch in den ergänzenden Basisklassen zu sehen, die ein umfangreiches Angebot von Funktionalitäten zur Verfügung stellen, die nicht ohne weiteres durch eigene Programmierarbeiten aufgebracht werden könnten. Dabei ist die Nutzung dieser zusätzlichen Funktionen mit einem vergleichsweise geringen Entwicklungsaufwand möglich. Von besonderer Bedeutung für das Projekt waren in diesem Zusammenhang die 3D-API (3D Application Program Interface), die Klassen zum Lesen und Darstellen von HTML-Seiten und die Klassen zum Schreiben und Parsen von XML-Daten.

6.3 Modellbasierte Segmentierung

Zentrale Aufgabenstellung der Tests ist die umfassende Beurteilung der modellbasierten Segmentierung. Dabei sollen vor allem die Stärken und Schwächen der modellbasierten Segmentierung identifiziert und so weit wie möglich quantifiziert werden. Es können die folgenden Testkomplexe unterschieden werden:

1. Aufbau der Referenzdatenbank (Abschnitt 6.3.1),
2. Bestimmen des Bildkontextes (Abschnitt 6.3.2),
3. Übertragen des Vorwissens (Abschnitt 6.3.3) und
4. Wissensanwendung durch den modellbasierten Snake (Abschnitt 6.3.4).

6.3.1 Aufbau der Referenzdatenbank

Ein entscheidender Punkt bei der Realisierung eines wissensbasierten Bildverarbeitungsansatzes ist die Akkumulation des Wissens. Im Fall der modellbasierten Bildverarbeitung wird die Wissensakkumulation durch das Sammeln von Referenzmodellen erreicht. Wichtige Fragestellungen für die Realisierung einer solchen Referenzdatenbank sind einerseits der Arbeitsaufwand, der für den Aufbau einer adäquaten Wissensdatenbank erforderlich ist und andererseits wie der Zugriff auf das akkumulierte Wissen vorgenommen wird.

Der Arbeitsaufwand für die Wissensakkumulation wird in erster Linie durch die Konstruktion der Referenzmodelle und durch ihre Integration in einer Referenzdatenbank verursacht. Zur Einschätzung des Arbeitsaufwandes für den Aufbau einer Referenzdatenbank wurden die Phantomdatensätze, die Datensätze der Schweineaugen, 10 Datensätze von Probanden und 15 Datensätze aus der klinischen Routine segmentiert. Anschließend wurden sechs Landmarken für die Anpassung der Referenzmodelle hinzugefügt. Um die Referenzmodelle für die modellbasierte Bildverarbeitung nutzen zu können, wurden die Referenzdaten und -segmentierungen in definierten Verzeichnissen abgelegt, aus der sie für die modellbasierte Segmentierung gezielt geladen werden können.

6.3.1.1 Manuelles Segmentieren der Datensätze und Konstruktion der Referenzmodelle

Bei der manuellen Segmentierung der Datensätze zur Konstruktion der Referenzmodelle sind vor allem der Arbeitsaufwand zur Erstellung der Segmentierungen und die Korrektheit der Segmentierungen von Bedeutung. Entsprechend der Angaben im Abschnitt 6.1 wurden sämtliche Testdatensätze von verschiedenen Benutzern segmentiert. Vor der Segmentierung wurde den Testpersonen erklärt, wie die Software genutzt wird und wie die dargestellten Organe korrekt segmentiert werden. Die Testpersonen führten zunächst eine zweidimensionale Segmentierung aus, bei der die Eingaben in den ursprünglichen Schichten vorgenommen wurden. Im zweiten Schritt bezogen die Testpersonen die dreidimensionale Organoberfläche mit ein, indem die bisher vorgenommenen Segmentierungen dreidimensional rekonstruiert wurden. Anschließend wurden die Segmentierungen so korrigiert, daß eine sinnvolle dreidimensionale Oberfläche entstand.

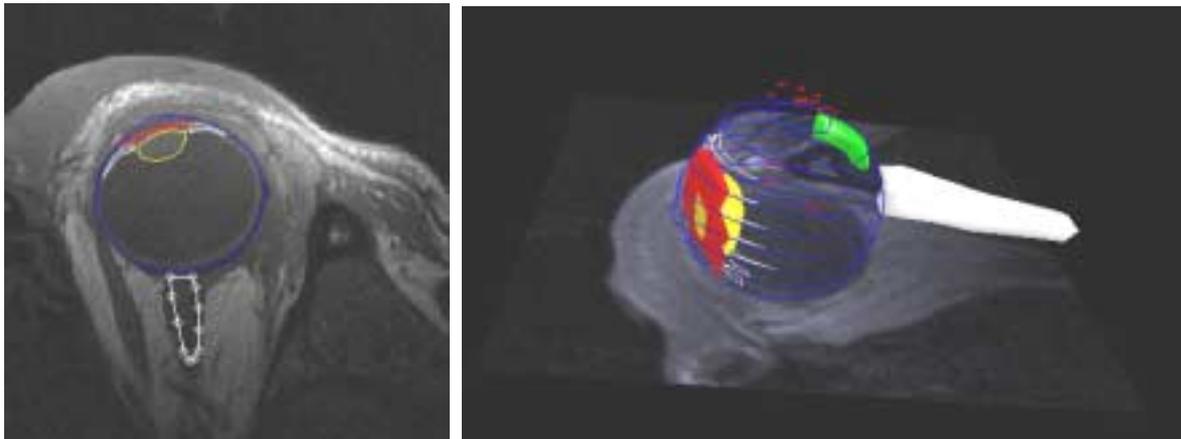


Abbildung 6.13: 2D-Segmentierung und 3D-Visualisierung der Organe

Links: Segmentieren der Organe in den zweidimensionalen Bildschichten. Die Kontur des Sehnervs wird gerade bearbeitet. Rechts: Durch die dreidimensionale Rekonstruktion der einzelnen Organe kann die Gesamtsegmentierung beurteilt und korrigiert werden.

Die Testdatensätze können unterteilt werden in Phantomdatensätze, Datensätze von Schweineaugen, Probandendaten und Datensätze aus der medizinischen Routine. Bei den Routinedatensätzen sind außerdem unterschiedliche Schichtauflösungen zu berücksichtigen, da sich die Auflösung direkt auf den Arbeitsaufwand auswirkt. Übliche MRT-Datensätze bestehen aus zirka 15 Schichten mit einer Bildmatrix von 256^2 Pixeln. CT-Datensätze weisen normalerweise 26 Schichten mit einer Bildmatrix von 512^2 Pixeln auf. Bei der Strahlentherapie besteht derzeit die Tendenz, die Auflösung der Daten über die der Routinedatensätze hinaus zu erhöhen. Diese Erhöhung der Auflösung wird einerseits durch eine Interpolation der Daten schon beim Bildgenerierungsprozeß erreicht, andererseits wird vor allem bei den MRT-Aufnahmen

eine Optimierung der Aufnahme-prozedur angestrebt. Der Arbeitsaufwand, der jeweils zur manuellen Segmentierung der Testdatensätze aufgebracht werden mußte, ist in Tabelle 6-1 zusammengefaßt.

Datensatz	Anmerkungen zum Datensatz	Segmentierte Objekte	Anzahl der Schichten	Gesamtdauer 2D-		Gesamtdauer 3D-		Arbeitsaufwand		Anmerkungen
				Mittelwert	Standardabweichung	Mittelwert	Standardabweichung	Min	Max	
Phantomdatensatz	19 Schichten 512 ² Bildmatrix	Korpus	16-17	11,50	2,50	1,50	1,00	17,00	36,00	
		Linsen-Imitat	5-6	4,50	2,00	1,50	0,75			
		Iris-Imitat	7-8	5,50	1,75	2,00	1,50			
Datensätze der Schweineaugen	14-17 Schichten 256 ² Bildmatrix	Bulbus	12-13	6,75	0,75	0,75	0,25	19,00	28,50	Die Iris wurde nach dem Ziliarkörper segmentiert
		Sehnerv	3	1,50	0,50	0,75	0,25			
		Linse	6-7	2,75	0,75	2,00	0,50			
		Ziliarkörper	8-9	5,25	0,75	0,25	0,25			
		Iris	8	3,50	0,50	0,25	0,25			
Probandendaten	11-14 Schichten 256 ² Bildmatrix	Bulbus	8-9	6,50	1,00	1,00	0,50	21,50	35,50	Die Iris wurde nach dem Ziliarkörper segmentiert
		Sehnerv	3	2,00	0,50	1,25	0,75			
		Linse	4	2,50	1,00	0,75	0,25			
		Ziliarkörper	6	6,75	1,25	3,75	0,75			
		Iris	5	3,25	0,75	0,75	0,25			
MRT-T1 axial	15 Schichten	Bulbus	11-12	10,00	3,80	5,42	5,46	23,97	69,64	Standarddatensatz
		Linse	4-6	3,58	1,06	3,50	2,08			
		Ziliarkörper	6-7	7,00	1,83	2,67	1,46			
		Iris	5-6	4,42	1,67	3,25	1,73			
		Sehnerv	3-4	2,00	0,76	1,93	1,27			
		Tumor	3-5	2,17	1,31	0,88	0,41			
MRT-T1 axial	80 Schichten	Bulbus	12	8,50	3,00	3,20	2,15	6,55	16,85	Höhere Auflösung bewirkte keine verbesserte Darstellung
CT axial	26 Schichten	Bulbus	10	16,00	10,00	11,65	3,50	14,15	41,15	Iris und Ziliarkörper nicht erkennbar
CT axial	50 Schichten	Bulbus	23	12,50	2,50	9,30	5,41	13,89	29,71	Künstlich erhöhte Auflösung
CT axial	131 Schichten	Bulbus	47	41,50	3,50	27,60	9,50	56,10	82,10	Schlecht erkennbare Kanten

Tabelle 6-1: Arbeitsaufwand für die Segmentierung der Testdatensätze in Minuten

Die Tabelle zeigt den durchschnittlichen Arbeitsaufwand in Minuten, der für die manuelle Segmentierung der unterschiedlichen Arten von Testdatensätzen aufgebracht werden mußte. Die Datensätze wurden von drei Testpersonen segmentiert. Bei der Angabe der Segmentierungsdauer wurden jeweils die mittlere Segmentierungsdauer und die Standardabweichung angegeben. Während die mittlere Segmentierungsdauer angibt, wie lange eine Segmentierung durchschnittlich dauerte, kann an der Standardabweichung abgelesen werden, wie stark sich interindividuelle Unterschiede und die Veränderungen im Bildmaterial auf die Segmentierungsdauer auswirkten. Dabei wird unterschieden zwischen dem Arbeitsaufwand für die zweidimensionale Segmentierung und die Modifikation der Segmentierung unter Berücksichtigung der dreidimensionalen Organoberflächenform. Bei „Anzahl der Schichten“ wird angegeben, in wie vielen Schichten das Organ segmentiert wurde. Bei dem CT-Datensatz mit 50 Schichten wurde die Auflösung künstlich erhöht, indem zwischen jeweils zwei Bildschichten eine Schicht interpoliert wurde.

Unter Berücksichtigung der Ergebnisse aus Tabelle 6-1 und aus Beobachtungen, die während der Segmentierungen gemacht wurden, können die folgenden Tendenzen bezüglich des Arbeitsaufwandes zur Segmentierung der Datensätze festgestellt werden:

- Der Zeitaufwand, der von den unterschiedlichen Testpersonen für die Segmentierung der Datensätze aufgebracht wurde, variierte zwischen den Testpersonen, zwischen den Datensätzen und über die Zeit hinweg.
- Unter realen Bedingungen ist für die Segmentierung eines üblichen MRT-Datensatzes mit 15 Bildschichten ein Arbeitsaufwand von mindestens einer halben Stunde einzurechnen. Dieser Arbeitsaufwand ist deutlich zu lang, um in die klinischen Routine integriert zu werden.
- Segmentierungsdauer und -qualität sind neben der Komplexität des Bildmaterials abhängig von der momentanen Konzentrationsfähigkeit, der körperlichen Verfassung und dem Erfahrungshorizont der Testperson. Während ein unerfahrener Nutzer seine Segmentierungen oft nachträglich korrigierte und anhand der dreidimensionalen Oberflächenform modifizierte, mußten bei einem erfahrenen Benutzer die Eingaben nur selten korrigiert werden. Dabei ist jedoch anzumerken,

daß der hohe Arbeitsaufwand für unerfahrene Nutzer zu einem nicht unbeträchtlichen Teil durch die täuschenden Darstellungseigenschaften des Bildmaterials verursacht wird. Gerade die Partialvolumeneffekte lassen das Auge deutlich größer erscheinen als es nach Beurteilung der dreidimensionalen Oberflächenform möglich oder sinnvoll ist (siehe Abbildung 6.14). Solche Probleme mußten dann nachträglich durch die 3D-Segmentierung korrigiert werden. Nach einer längeren Einarbeitungszeit konnten sich die Testpersonen jedoch auf diese Eigenschaften des Bildmaterials einstellen und gleich bei der zweidimensionalen Segmentierung berücksichtigen. Dadurch reduzierte sich vor allem der Aufwand der dreidimensionalen Segmentierung mit zunehmender Erfahrung.

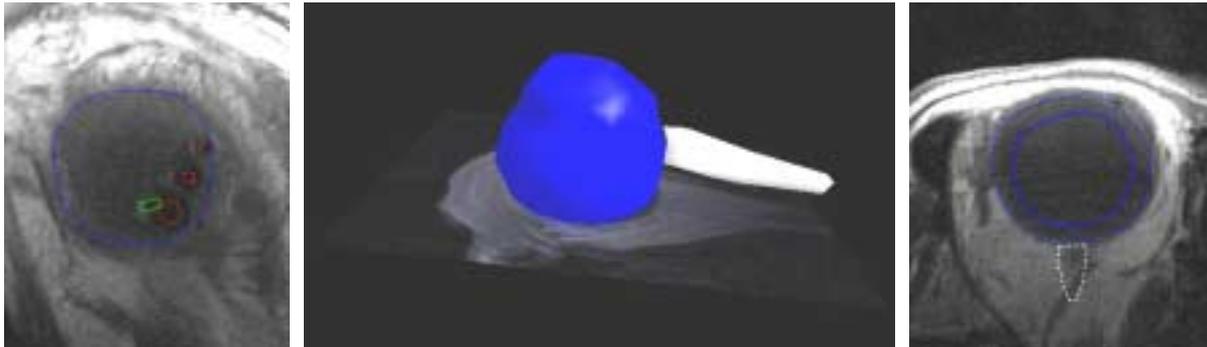


Abbildung 6.14: Segmentierungsprobleme durch Partialvolumeneffekte

Durch die Partialvolumeneffekte wird die Darstellung der Organe grundlegend verfälscht. Am deutlichsten läßt sich das Vergrößern von Organen am Bulbus verdeutlichen. Bei einer intuitiven Segmentierung des Bulbus zeichneten die Testpersonen die Organkontur am stärksten Grauwertübergang ein. Bereits eine genaue Betrachtung der zweidimensionalen Bildschicht zeigt, daß eine solche Segmentierung nicht sinnvoll sein kann, da beispielsweise auch die aufgenähten Clips im Bildmaterial zu sehen sind. Bei der dreidimensionalen Rekonstruktion in der Mitte ist deutlich die Beule am oberen Ende des Bulbus zu erkennen. Sie entsteht durch die vergrößerte Darstellung des Bulbus durch die Partialvolumeneffekte. Eine deutliche Vereinfachung der Segmentierung ermöglicht das Einblenden der Segmentierungen der Nachbarschichten (rechtes Bild). An den eingebblendeten Konturen (gestrichelte Konturen) kann durch Plausibilitätsüberlegungen erkannt werden, wie die Kontur in der aktuellen Schicht ungefähr aussehen muß.

- Eine höhere Auflösung bei der Darstellung der Organe erhöht den erforderlichen Zeitaufwand. Dieser Effekt ist nahezu unabhängig davon, ob die höhere Auflösung durch den Datensatz oder durch die Verwendung der Zoom-Funktion erreicht wurde. Als Hauptursache hierfür ist die Tatsache zu sehen, daß je größer die Darstellung des Organs ist, um so mehr Konturpunkte platzierten die Testanwender entlang der Organkante.
- Eine sehr wichtige Rolle bei der Segmentierung der Datensätze spielt die Bildqualität. Sowohl in den CT-, aber auch in den MRT-Daten sind vor allem die kleinen anatomischen Strukturen wie Linse, Iris und Ziliarkörper nur schlecht oder gar nicht zu erkennen. Es erfordert eine gute Kenntnis der Anatomie des Auges, um den Mangel an Dargestelltem anhand von plausibler Fortsetzung der Organform bei der Segmentierung zu kompensieren. Während die ophthalmologische Fachkraft umfassende Anatomiekenntnisse nutzen kann, muß sich ein Nutzer mit wenig Anatomiekenntnissen zunächst die zugrunde liegenden Organoberfläche mental verinnerlichen und in die zweidimensionalen Schichten übertragen. Sämtliche Testpersonen, einschließlich der medizinischen Fachkraft, hatten zusätzlich Probleme mit der Tatsache, daß die Darstellung von großen wie kleine Organstrukturen durch die Partialvolumeneffekte nicht nur verwischt, sondern auch verfälscht wurden. Am deutlichsten ist dies an den Randbereichen des Bulbus zu erkennen, der im Bildmaterial deutlich größer dargestellt wird, als es nach dreidimensionalen Oberflächenform sinnvoll wäre (siehe Abbildung 6.14). Dieses Problem führte vor allem bei der Segmentie-

rung der ersten Datensätze dazu, daß medizinische Fachkraft und Laie umfangreiche Korrekturen an der Segmentierung vornehmen mußten. Als sehr hilfreich erwies sich in diesem Zusammenhang die Erweiterung der Visualisierungssoftware durch eine Funktion, bei der die Segmentierungen der Nachbarschichten je nach Bedarf in gestrichelter Form in die aktuelle Schicht einblendbar werden können (vergleiche Abbildung 6.14, rechts). Nach der Segmentierung von mehreren Datensätzen stellten sich die medizinische Fachkraft und die anderen Testpersonen auf die täuschenden Eigenschaften des Bildmaterials ein. Dadurch nahm der Arbeitsaufwand für die notwendigen Korrekturen kontinuierlich ab.

- Die Segmentierung der CT-Daten erwies sich aufgrund des geringen Grauwertkontrastes als außerordentlich schwierig und konnte nur für die größeren Strukturen Bulbus, Sehnerv und Linse durchgeführt werden. Selbst für geübte Nutzer sind die Organgrenzen nur so schwer zu erkennen, daß in großen Bereichen die Organgrenze eher geschätzt und möglichst plausibel fortgesetzt werden muß, anstatt die Eingaben tatsächlich am Bildmaterial zu orientieren. Diese Unsicherheit wirkte sich direkt auf die Korrektheit der Ergebnisse und damit den Arbeitsaufwand für die 2D, in erster Linie aber für die 3D-Segmentierung aus (siehe Abbildung 6.16). Besonders hilfreich erwies sich auch in dieser Situation das Einblenden der Segmentierungen der Nachbarschichten. Es muß jedoch bei einer solchen Vorgehensweise in Frage gestellt werden, inwieweit solche Segmentierungsverfahren, die auf ästhetischen Aspekten und Plausibilitätsüberlegungen basieren, wirklich als korrekt angesehen werden können.
- Die Segmentierung der Datensätze mit hoher Ortsauflösung ist mit einem deutlich größeren Arbeitsaufwand verbunden. Die Segmentierung an sich wurde durch die höhere Auflösung nicht erleichtert, da die Organgrenzen nicht besser als bei den Routinedatensätzen dargestellt werden. Demgegenüber steht jedoch eine deutlich höhere Anzahl von Schichten, in denen die Organe segmentiert werden müssen (vergleiche Abbildung 6.15 und Abbildung 6.16). Dies stellt nicht nur erhöhte Anforderungen an die Konzentrationsfähigkeit des Nutzers, sondern auch an die Hardware des genutzten Computers. Einer der CT-Datensätze besteht aus 131 Bildschichten mit einer Bildmatrix von jeweils 512^2 Pixeln. Allein der Datensatz erfordert einen Speicheraufwand von 66 MByte. Sollen außer diesem CT-Datensatz noch ein oder mehrere entsprechende MRT-Datensätze geladen werden, ist schnell ein Arbeitsspeicher von über 256 MByte nötig, um mit den Datensätzen zu arbeiten. Eine solche Ausrüstung geht über die eines üblichen Arztes hinaus. Hinzu kommt der mit einer solchen Datenmenge verbundene Rechenaufwand. Um den 66 MByte Datensatz zu speichern wurden bei den Tests 5 Minuten benötigt. Eine solche Wartezeit ist für die klinische Routine nicht akzeptabel. Die Bearbeitung solcher Datensätze würde demnach speziell ausgerüstete Rechner erforderlich machen. Andererseits wird der erhöhte Aufwand für solche Datensätze nicht durch eine verbesserte Bildqualität gerechtfertigt. Die Datensätze erhöhter Auflösung weisen die gleichen Verwischungen und einen geringen Grauwertkontrast auf wie Routinedatensätze, die aus 15 Bildschichten bestehen (siehe Abbildung 6.15, links). Ein schwerwiegender Nachteil solcher hochauflösender Datensätze besteht zusätzlich darin, daß mit dem deutlich erhöhten Arbeitsaufwand für die Segmentierung auch eine überproportional gewachsene Frustration verbunden ist, da die Arbeitsfortschritte deutlich geringer sind und länger dauern. Darüber hinaus müssen die Konturen deutlich präziser eingegeben werden, um eine akzeptable Oberflächenqualität zu erreichen, da sich kleine Segmentierungsfehler deutlich in der dreidimensionalen Rekonstruktion widerspiegeln (siehe Abbildung 6.15, rechts). Die Konzentrationsanforderungen an die Nutzer sind dementsprechend deutlich höher als bei den Routinedatensätzen. Sollten sich solche Datensätze erhöhter Auflösung für die klinische Routine durchsetzen, müßte mit grundlegenden Motivationsproblemen bei den Nutzern gerechnet werden, da es sich letztendlich bei der Segmentierung um eine Tätigkeit mit geringer intellektueller Anforderung handelt.

Grundlegende Einbrüche bei der Segmentierungsqualität wären wohl kaum vermeidbar. Der Anreiz für die Nutzung der modellbasierten Segmentierung wäre andererseits um ein Vielfaches erhöht.

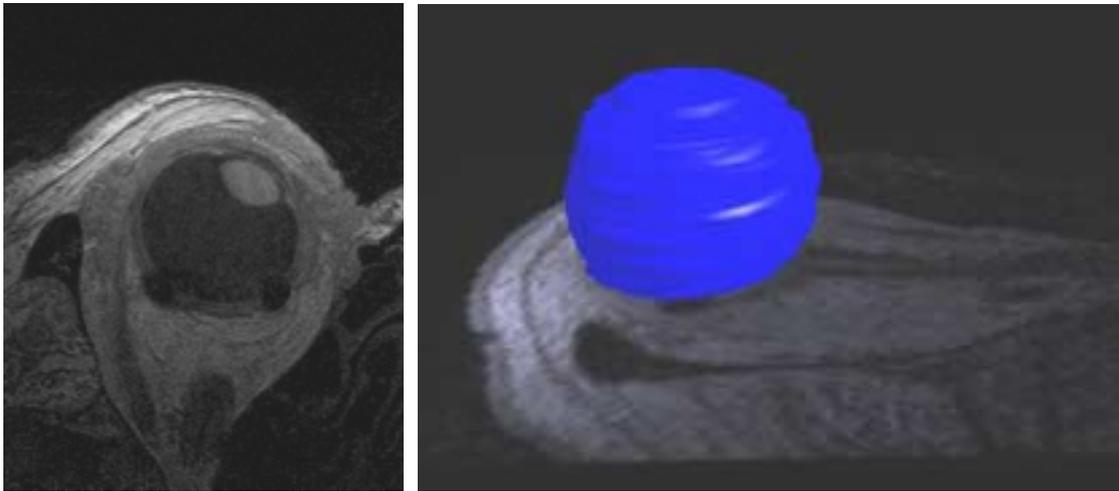


Abbildung 6.15: MRT-Aufnahmen mit erhöhter Auflösung

Links ist beispielhaft eine Schicht aus einem MRT-Datensatz mit erhöhter Auflösung dargestellt. Während zwar die Linse recht gut zu erkennen ist, bleibt die Darstellung der Sklera deutlich hinter der der Routinedatensätze zurück. Die Darstellungsqualität des Bildmaterials wurde durch die erhöhte Auflösung des Datensatzes nicht verbessert. Rechts ist die dreidimensionale Rekonstruktion einer zweidimensionalen Segmentierung des MRT-Datensatzes zu sehen. Der Segmentierungsaufwand ist deutlich angestiegen, da nun elf anstelle von 69 Schichten segmentiert werden mußten. Der Spielraum für Ungenauigkeiten ist dadurch deutlich größer, was sich in der kantigen Oberfläche des Bulbus widerspiegelt.

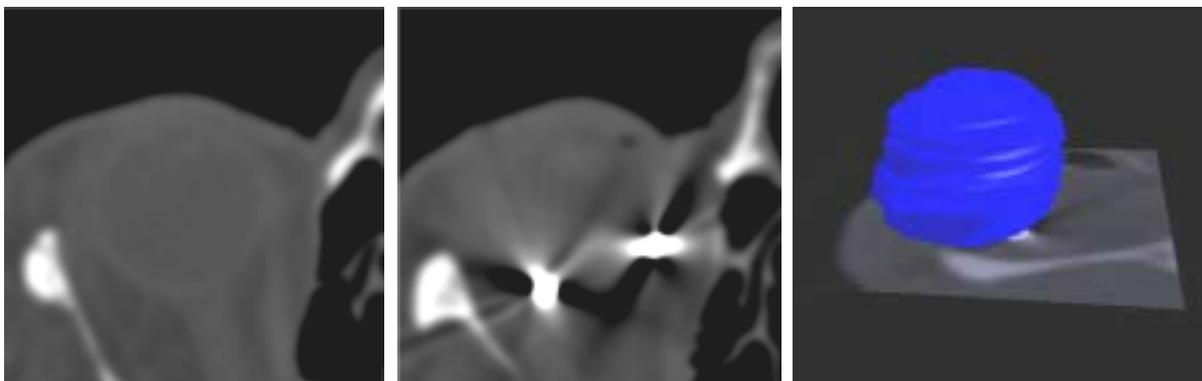


Abbildung 6.16: Probleme bei der Segmentierung von CT-Datensätzen erhöhter Auflösung

Auch die CT-Datensätze erhöhter Auflösung stellen keine Verbesserung gegenüber den herkömmlichen CT-Datensätzen dar. Der Grauwertkontrast ist in gleichem Maße gering und erschwert das Erkennen der Organkonturen. Metallartefakte erstrecken sich über deutlich mehr Schichten als bei den herkömmlichen Datensätzen. Dadurch wird die zweidimensionale Segmentierung weiter erschwert, da das Übernehmen von Konturformen aus den Nachbarschichten erschwert ist. Die Folge sind außerordentlich lange Bearbeitungszeiten, die eine vergleichsweise geringe Segmentierungsqualität mit sich bringen (siehe rechts eine dreidimensionale Rekonstruktion der zweidimensionalen Segmentierung des Bulbus).

Zusammenfassend kann festgehalten werden, daß der Arbeitsaufwand der manuellen Segmentierung für die klinische Routine deutlich zu groß ist und eine geringe Segmentierungsqualität aufgrund des Konzentrationsverlustes, der Frustration und der Motivationsprobleme zu erwarten ist. Die Verwendung von hochaufgelösten Datensätzen verschlechtert die Chancen auf eine sinnvolle Nutzung der Daten in der klinischen Routine und bietet keine erhöhte Korrektheit der Segmentierungsergebnisse.

Bei dem Aufbau einer Referenzdatenbank spielt neben dem Arbeitsaufwand auch die Segmentierungskorrektheit der manuellen Segmentierungen eine wichtige Rolle. Einerseits ist es das Ziel der modellbasierten Bildverarbeitung, nicht nur eine Reduzierung des Arbeitsaufwandes zu erzielen, sondern auch die Erreichung einer Segmentierungsqualität, die der manuellen Segmentierung so weit wie möglich ebenbürtig ist. Andererseits müssen bei der Einschätzung der Ergebnisqualität der modellbasierten Segmentierung die Ergebnisse im Zusammenhang mit den Möglichkeiten der manuellen Segmentierung gesehen werden. Da die Referenzmodelle durch manuelle Segmentierung angelegt werden, ist nicht zu erwarten, daß die Ergebnisse der modellbasierten Bildverarbeitung jene der manuellen Segmentierung in ihrer Korrektheit übertreffen. Allein in der Konsistenz der Ergebnisse ist ein Übertreffen der manuellen Segmentierung theoretisch möglich.

Um einen quantitativen Vergleich der manuellen und modellbasierten Segmentierung möglich zu machen, muß daher die Ergebnisqualität der manuellen Segmentierung quantifiziert werden. Hierzu wurden die Testdatensätze von unterschiedlichen Personen manuell segmentiert und anschließend die Ergebnisse miteinander verglichen. Die Übereinstimmungen zwischen den verschiedenen Segmentierungen wurden ermittelt, indem die von unterschiedlichen Mitarbeitern segmentierten Konturen miteinander geschnitten wurden. Anschließend wurden die Schnittflächen, die Flächen, in denen sich die Segmentierungen unterschieden und die Schnittvolumina berechnet. Diese Tests wurden sowohl für Segmentierungen, die von unterschiedlichen Testpersonen angefertigt wurden, durchgeführt, als auch für Datensätze, die von derselben Person zu unterschiedlichen Zeitpunkten segmentiert wurden.

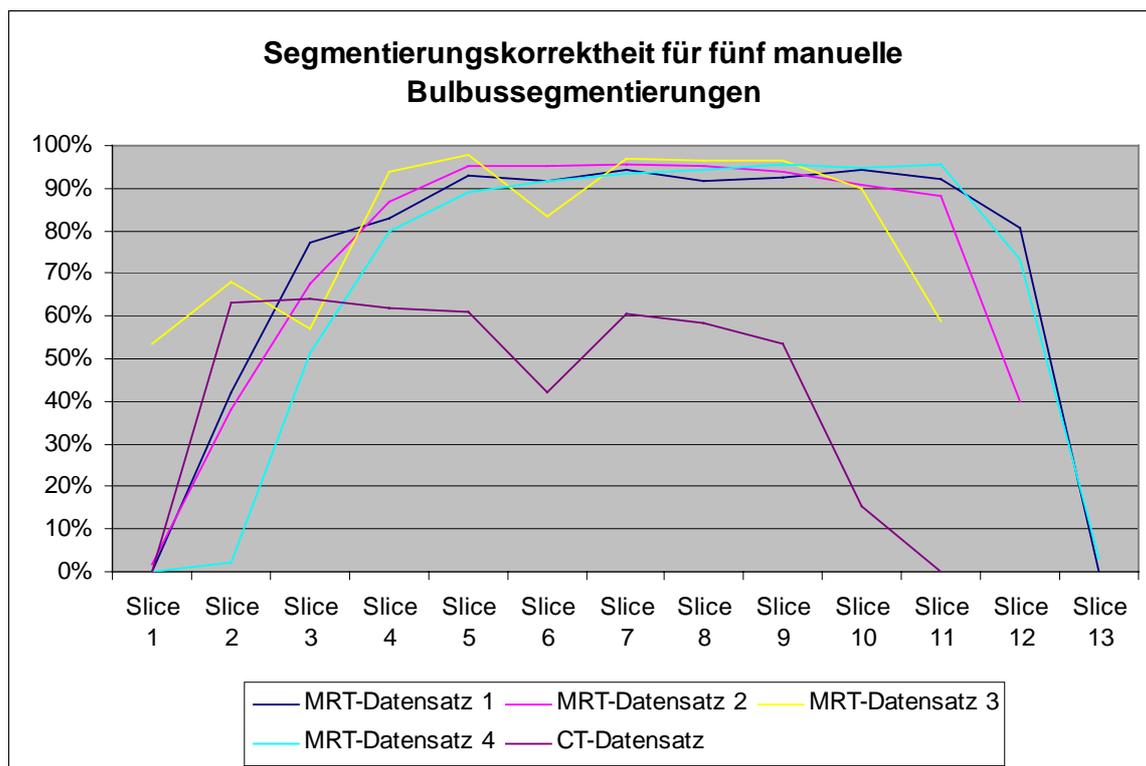


Abbildung 6.17: Interindividuelle Segmentierungsübereinstimmungen beim Bulbus

Zur Abschätzung der interindividuellen Segmentierungsübereinstimmungen wurden vier MRT-Datensätze und ein CT-Datensatz von zwei unterschiedlichen Mitarbeitern segmentiert. Anschließend wurden zwischen den beiden Segmentierungen die Schnittflächen und Schnittvolumina bestimmt. In der Abbildung wird die relative Übereinstimmung der Segmentierungen dargestellt. Sie wird berechnet, indem die Schnittfläche der Segmentierungen dividiert wird durch die mittlere Fläche der beiden Segmentierungskonturen. Während in den zentralen Bildschichten des Auges die Übereinstimmungen vergleichsweise groß sind, kommt es in den Randschichten zu beträchtlichen Abweichungen. Es ist vor allem auf die Partialvolumeneffekte zurückzuführen, daß in den Randschichten die Segmentierungen zum Teil gar nicht mehr übereinstimmen (relative Übereinstimmung 0%).

Aus medizinischer Sicht ist eine rein volumen- oder flächenorientierte Betrachtung der Segmentierungsunterschiede nur bedingt sinnvoll, da die Bedeutung von lokalen Abweichungen in den Segmentierungskonturen stark vom jeweiligen Organ und dem lokalen Bildkontext abhängig ist. Während eine Abweichung der Segmentierungskonturen von wenigen Pixeln im Falle des Bulbus durchaus tolerabel sein können, ist die Bedeutung solcher Abweichungen im Falle der Iris oder des Ziliarkörpers gegebenenfalls anders einzuschätzen. Für den grundsätzlichen Vergleich von unterschiedlichen Bildverarbeitungsansätzen ist der volumenorientierte Vergleich jedoch durchaus aussagekräftig, da medizinisch relevante Korrekturen später durch den Nachbearbeitungsaufwand berücksichtigt werden können.

Bei der Durchführungen von Tests zur Beurteilung der manuellen Segmentierung müssen die Testpersonen die Kompetenz, Erfahrung, Konzentration und Ausdauer mitbringen, um die Ergebnisse der Segmentierungen überhaupt miteinander vergleichen zu können. Zu diesem Zweck wurden die Testpersonen in die Durchführung einer medizinisch korrekten Segmentierung eingewiesen. Anhand von Beispielen wurde den Testpersonen erläutert, wie das Bildmaterial zu interpretieren und die Organkonturen einzutragen sind. Darüber hinaus konnten die Testpersonen während der Segmentierungsarbeiten stets Rückfragen stellen. Um eine Beeinträchtigung der Segmentierung durch mangelnde Konzentration zu vermeiden, wurden die Segmentierungen in Etappen durchgeführt. Da damit sämtliche Voraussetzung für die Anfertigung einer korrekten Segmentierung erfüllt waren, sollte davon ausgegangen werden, daß die Segmentierungsergebnisse der Testpersonen eng miteinander korreliert sein sollten.

Abbildung 6.17 veranschaulicht die Ergebnisse der Segmentierungen, indem beispielhaft die interindividuellen Unterschiede zwischen den Segmentierungen des Bulbus in vier MRT-Datensätzen und einem CT-Datensatz in den jeweiligen Bildschichten dargestellt werden. Abgebildet wird die relative Übereinstimmung der segmentierten Organflächen in Prozent. Aus den Ergebnissen können die folgenden Schlußfolgerungen gezogen werden:

- Die Übereinstimmungen zwischen den Segmentierungen in den zentralen Schichten des Bulbus sind recht groß, da in diesen Schichten die Ränder des Bulbus sehr gut lokalisiert werden können. Darüber hinaus ist die Gesamtfläche des Bulbus vergleichsweise groß, so daß kleine Segmentierungsunterschiede nicht so stark ins Gewicht fallen.
- In den Randschichten des Bulbus sind die Übereinstimmungen vergleichsweise gering und fällt in einigen Schichten sogar auf 0% ab, d.h. es gibt keine Schnittfläche zwischen den beiden Segmentierungen. Die Ursache hierfür ist einerseits in den starken Partialvolumeneffekten zu sehen, die eine Identifikation der Bulbus-Kontur deutlich erschweren, in einigen Fällen sogar unmöglich machen. Andererseits fallen aufgrund der geringen Gesamtflächen Unterschiede in den Segmentierungen stärker ins Gewicht.
- CT-Daten sind aufgrund des geringen Grauwertkontrastes generell schwerer zu segmentieren. Das Einblenden der Nachbarschichten vereinfacht zwar die Segmentierung von angrenzenden Schichten, in denen große Teile der Organkanten nicht erkannt werden können. Die Segmentierungskorrektheit wird dadurch jedoch nicht erhöht. Die Folge sind vergleichsweise große Abweichungen zwischen den Segmentierungen.

Die Segmentierung des Bulbus ist verglichen mit den anderen kleineren Strukturen des Auges als einfach anzusehen. Wie in Abbildung 6.18 zu sehen ist, wirken sich die erschwerenden Einflüsse bei der Segmentierung von kleineren Organen wie der Iris, dem Ziliarkörper oder der Linse noch deutlich stärker aus:

- Die Übereinstimmungen zwischen manuellen Segmentierungen von kleinen Organen wie Iris und Ziliarkörper sind außerordentlich gering und unterliegen sehr starken Schwankungen. Diese geringen Übereinstimmungen in den Segmentierungen resultieren aus den schlechten aber durchaus

normalen Darstellungseigenschaften des Bildmaterials und den ausgeprägten Artefakten. Beide Faktoren führen dazu, daß die Organe weniger anhand des Bildmaterials eingezeichnet werden können, sondern daß vor allem anatomisches Vorwissen die Segmentierung bestimmt.

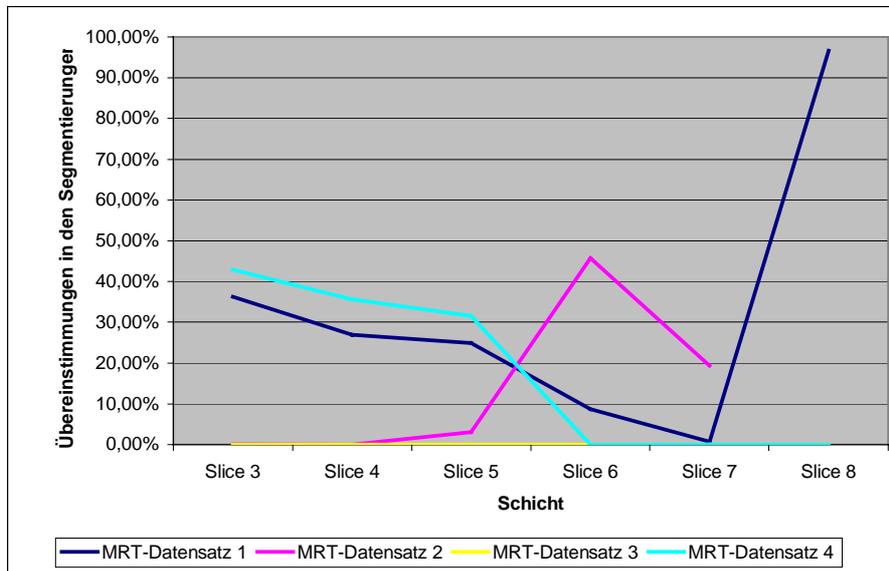


Abbildung 6.18: Interindividuelle Segmentierungsübereinstimmungen bei der Iris

Eine konsistente Segmentierung der Iris ist mit deutlich größeren Problemen behaftet. Wie in sämtlichen Kurven zu sehen ist, sind die mittleren Übereinstimmungen bei den Segmentierungen der Iris deutlich geringer als beim Bulbus in Abbildung 6.17. Darüber hinaus unterliegen die Übereinstimmungen auch deutlich stärkeren Schwankungen. In nicht wenigen Fällen gibt es keine Übereinstimmungen zwischen den Segmentierungen der zwei Testpersonen. Bei der Segmentierung des dritten MRT-Datensatzes gibt es in keiner Schicht eine Übereinstimmung zwischen den beiden Segmentierungen. Beim vierten Datensatz gibt es in den Schichten sechs, sieben und acht keine Übereinstimmungen.

- In CT-Daten sind diese kleinen Strukturen gar nicht zu erkennen. Ein manuelles Segmentieren dieser Strukturen ist daher nicht möglich.

Neben der Einschätzung von Detailproblemen bei der Segmentierung der Organe ist es zur Beurteilung der manuellen Segmentierung an sich wichtig, die zu erwartende Korrektheit der manuellen Segmentierung insgesamt abschätzen zu können. Zu diesem Zweck wurden die Übereinstimmungen zwischen den interindividuellen Segmentierungen gemittelt und in Tabelle 6-2 zusammengefaßt.

Organ	Durchschnittl. Volumen	Schnittvolumen	Übereinstimmung
Bulbus	6,66775	5,68625	86,07%
Iris	0,129375	0,00575	9,53%
Linse	0,12425	0,0715	75,56%
Sehnerv	0,2065	0,1075	51,70%
Ziliarkörper	0,077	0,0145	20,25%

Tabelle 6-2: Durchschnittliche Übereinstimmungen der Segmentierungen

In der Tabelle werden die Übereinstimmungen zwischen den Segmentierungen für die verschiedenen Organe in einem arithmetischen Mittelwert zusammengefaßt. Die maximale Übereinstimmung wird bei der Segmentierung des Bulbus mit 86% erreicht. Unter Berücksichtigung der Präzisionsanforderungen der Strahlentherapie ist ein relativer Unterschied von 14% zwischen den Segmentierungen als unzureichend anzusehen. Deutlich drastischer sind die Probleme bei der Segmentierung von kleinen Organen wie Iris und Ziliarkörper, bei denen lediglich eine maximale Übereinstimmung zwischen zehn und zwanzig Prozent erreicht werden kann. Da andererseits die Präzisionsanforderungen gerade auch für diese Organstrukturen sehr hoch sind, muß die manuelle Segmentierung für die gegebene Aufgabenstellung als unbrauchbar angesehen werden.

Die Abweichungen zwischen den Segmentierungen betragen beim Bulbus 14%, bei kleinen Organen wie Iris und Ziliarkörper 80 bis über 90%. Daß diese Unterschiede trotz der ausführlichen Einführung und der Möglichkeit, die eigene Segmentierung mit anderen Segmentierungen zu vergleichen, so schlecht ausfielen, stellt die manuelle Segmentierung an sich in Frage. Bevor jedoch die Anwendbarkeit der manuellen Segmentierung insgesamt beurteilt werden kann, muß untersucht werden, ob diese beträchtlichen Unterschiede aus den interindividuellen Interpretationsspielräumen resultieren, oder ob sie tatsächlich immanent zu diesem Segmentierungsansatz sind.

Abbildung 6.19 zeigt die relativen Übereinstimmungen der Segmentierungen des Bulbus, die von ein und derselben Person zu unterschiedlichen Zeitpunkten angefertigt wurden. Gerade in den zentralen Schichten des Bulbus sind die Übereinstimmungen recht groß. In den Randbereichen hingegen sind immer noch wesentliche Unterschiede festzustellen. Diese Unterschiede sind unmittelbar auf die starken Partialvolumeneffekte zurückzuführen, die eine eindeutige Segmentierung der Randbereiche unmöglich machen.

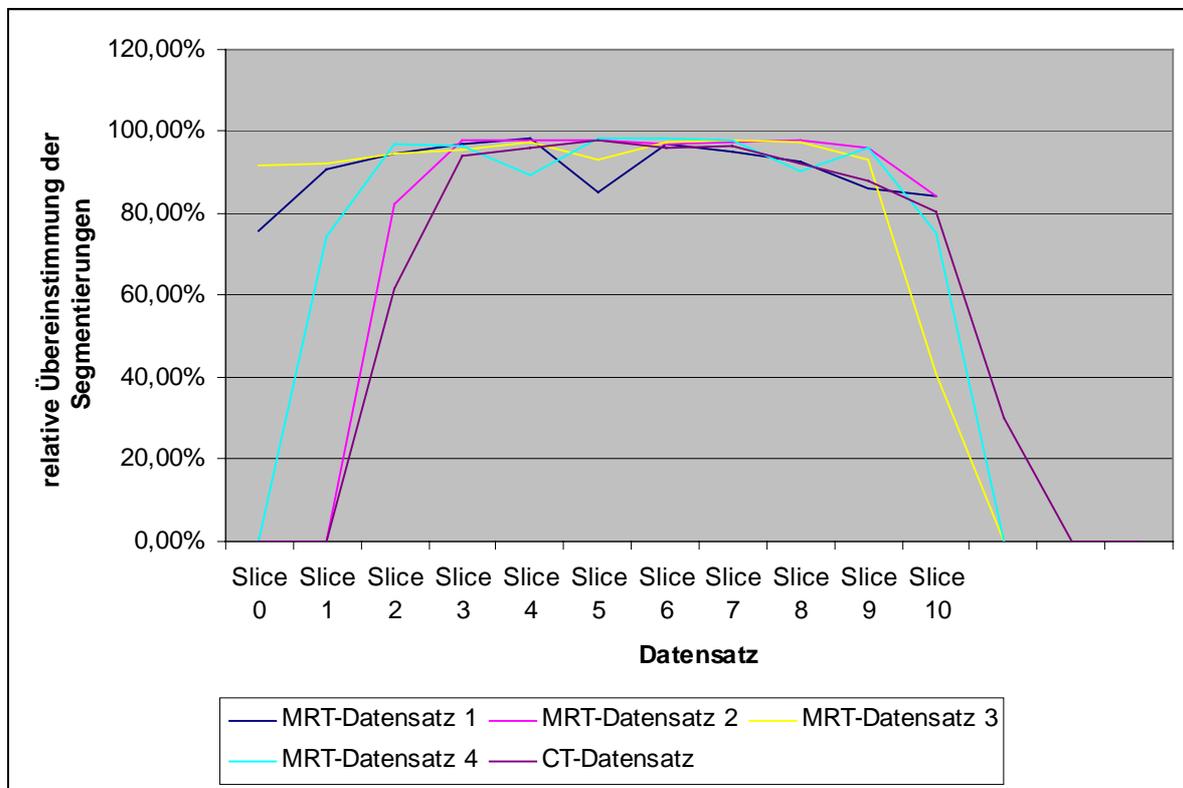


Abbildung 6.19: Intraindividuelle Übereinstimmungen bei der Segmentierung des Bulbus

Die Abbildung zeigt die relative Übereinstimmung der Segmentierungen des Bulbus, die von ein und derselben Person zu unterschiedlichen Zeitpunkten angefertigt wurden. Während die zentralen Schichten durchaus eine sehr hohe Übereinstimmung aufweisen, treten in den Randschichten durchaus beträchtliche Unterschiede auf.

Ähnliche Ergebnisse erzielte der Vergleich der Segmentierungen von kleinen Organen, die beispielhaft anhand der Iris in Abbildung 6.20 dargestellt werden. Die mittleren Übereinstimmungen der Segmentierungen, die von einer Testperson angefertigt wurden, sind sichtlich größer als bei Segmentierungen von unterschiedlichen Testpersonen. Ob die erzielten Übereinstimmungen den Anforderungen der Strahlentherapie von Augentumoren entsprechen, ist jedoch fraglich. Darüber hinaus unterliegen die Übereinstimmungen immer noch beträchtlichen Schwankungen.

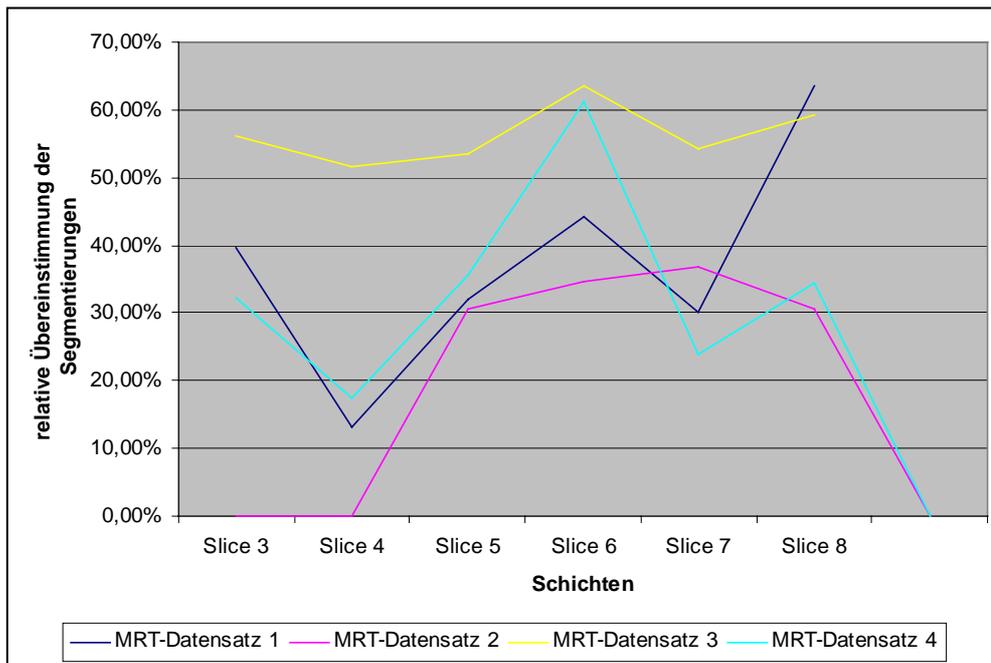


Abbildung 6.20: Intraindividuelle Übereinstimmungen bei der Segmentierung der Iris

Die Abbildung zeigt die relativen Übereinstimmungen zwischen den Segmentierungen der Iris, die von einer Testperson zu unterschiedlichen Zeitpunkten angefertigt wurden. Wenngleich die mittlere Übereinstimmung höher als beim Vergleich der interindividuellen Segmentierungen ausfallen, unterliegen die Übereinstimmungen auch für die Segmentierungen von einer Testperson deutlichen Schwankungen.

Tabelle 6-3 stellt die Durchschnittswerte für die interindividuellen und intraindividuellen Übereinstimmungen der Segmentierungen von Bulbus, Iris, Linse, Sehnerv und Ziliarkörper einander gegenüber. Ein Vergleich der Werte zeigt, daß die Übereinstimmungen außer bei der Linse bei allen Organen zunehmen, wenn die Segmentierung von einer Testperson angefertigt wurden. Diese Zunahme der Übereinstimmungen ist jedoch nicht so wesentlich, als daß man sagen könnte, daß damit das Problem der Reproduzierbarkeit der Segmentierungen überwunden wäre. Gerade für die kleinen Organstrukturen sind die Abweichungen immer noch so ausgeprägt, daß offensichtlich die Segmentierung stark vom Zufall beeinflusst wird.

Organ	Durchschnittl. Volumen	Schnittvolumen	Interindividuelle Übereinstimmung	Intraindividuelle Übereinstimmung
Bulbus	6,2253	5,6244	86,07%	90,91%
Iris	0,048625	0,018	9,53%	35,95%
Linse	0,11475	0,08325	75,56%	71,47%
Sehnerv	0,217875	0,13025	51,70%	60,15%
Ziliarkörper	0,0885	0,03175	20,25%	35,16%

Tabelle 6-3: Durchschnittliche Übereinstimmung der Segmentierungen von einer Testperson

Die intraindividuellen Übereinstimmungen übertreffen zwar im Durchschnitt die interindividuellen Übereinstimmungen. Jedoch treten immer noch so wesentliche Abweichungen auf, daß die Korrektheit der Segmentierungen nicht im Millimeterbereich liegen, wie es ursprünglich für die Bestrahlungstherapie von Augentumoren vorgesehen war.

Deutlich wird dieser Einfluß der Größe der Organkontur und der Partialvolumeneffekte auch in der genaueren Betrachtung der mittleren Übereinstimmungen zwischen den inter- und intraindividuellen Segmentierungen in Abbildung 6.21. Vor allem in den Randschichten des Bulbus, in denen die Kontur kleiner wird und zunehmend durch die Partialvolumeneffekte überlagert wird, stimmen inter- und intrain-

dividuelle Segmentierungen deutlich weniger miteinander überein. Im direkten Vergleich sind jedoch intraindividuelle Unterschiede deutlich geringer als interindividuelle, da personenbezogenen Interpretationsunterschiede wegfallen. Diese Interpretationsunterschiede kommen in den Randschichten anscheinend deutlich stärker zum Tragen als in den zentralen Schichten des Bulbus.

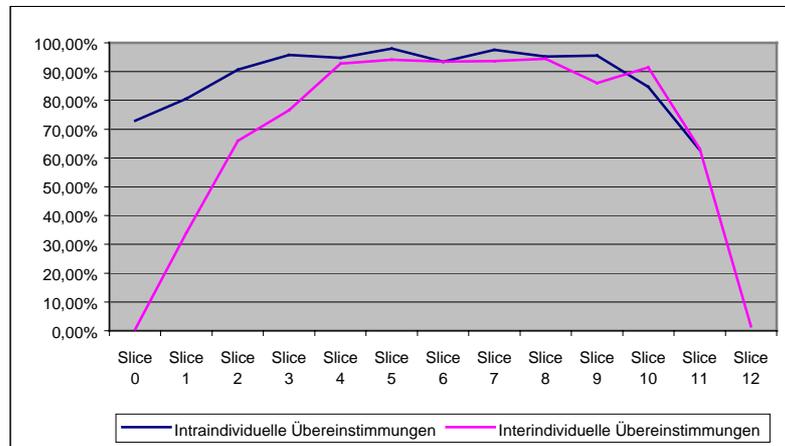


Abbildung 6.21: Inter- und intraindividuelle Übereinstimmungen in den Segmentierungen

Die Abbildung veranschaulicht graphisch die Übereinstimmungen zwischen interindividuellen und intraindividuellen Segmentierungen, die zu unterschiedlichen Zeitpunkten angefertigt wurden. Diese Mittelung der Übereinstimmungen ist nur eine grobe Beschreibung der tatsächlichen Übereinstimmungen, da die Segmentierungsschwierigkeiten von den Schichten nicht immer miteinander übereinstimmen. In einigen Datensätzen wird der Bulbus lediglich in neun Schichten abgebildet, während in anderen der Bulbus in bis zu zwölf Schichten segmentiert werden mußte. Eine grundsätzliche Einschätzung der Segmentierungen kann dadurch dennoch vorgenommen werden. Die Abbildung zeigt deutlich, daß die interindividuellen Segmentierung sich vor allem in den Randschichten deutlich unterscheiden. Doch auch wenn die Segmentierungen von einer Person angefertigt werden, ist diese Variabilität der Segmentierung nicht gänzlich auszuschließen. In den zentralen Schichten stimmen hingegen die Segmentierungen bei beiden Vorgehensweisen weitgehend überein.

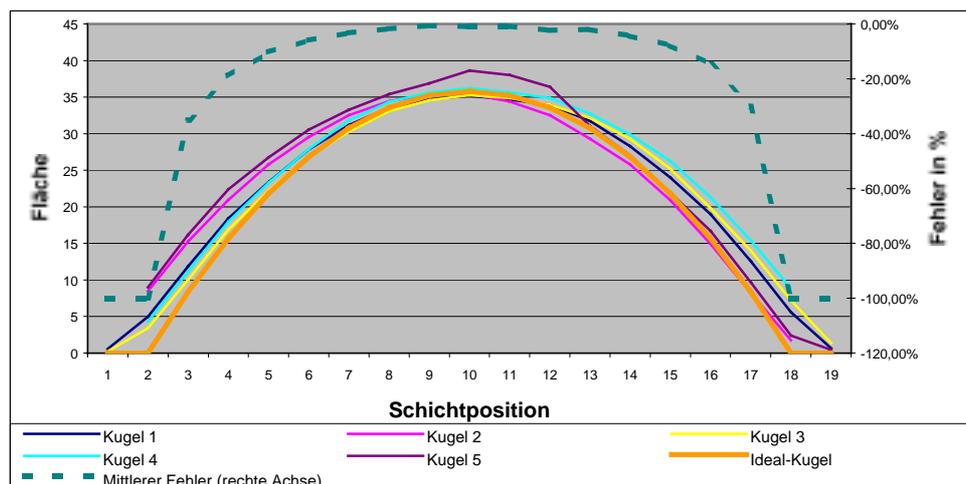


Abbildung 6.22: Grafische Darstellung der berechneten und idealen Flächen

Die Abbildung zeigt die berechneten und idealen Flächen für das künstliche Augenmodell mit einem Kugeldurchmesser von sieben Zentimetern. Die gestrichelte Linie stellt den mittleren Fehler bei den manuell segmentierten Flächen dar. Sämtliche manuell eingezeichneten Kurven nähern recht gut die Flächen an, die theoretisch aufgrund der bekannten Ausmaße des Modells erreicht werden müßten. Nichtsdestotrotz treten durchaus nicht unwesentliche Abweichungen auf. Während die zentralen Schichten des Bulbus erwartungsgemäß besser segmentiert wurden, sind die Abweichungen in den Randbereichen sehr deutlich.

Die vorangegangenen Betrachtungen und quantitativen Ergebnisse zur Beurteilung der manuellen Segmentierung betrachten grundsätzliche komparative Tendenzen bezüglich der Segmentierung von medizinischen Datensätzen. Zuverlässige quantitative Aussagen über die theoretische Genauigkeit der manuellen Segmentierung lassen sie jedoch nicht zu, da kein quantitativ eindeutig bestimmter Goldstandard vorliegt. Es ist nicht bekannt, wie die korrekte Segmentierung der Testdatensätze jeweils tatsächlich aussieht und es ist davon auszugehen, daß eine korrekte Segmentierung im wörtlichen Sinne gar nicht existiert.

Um dennoch abschätzen zu können, wie gut zumindest theoretisch manuell angefertigt Segmentierungen die realen Organformen rekonstruieren können, wurden die Phantomdatensätze manuell segmentiert und anschließend die rekonstruierten Oberflächen mit den bekannten Ausmaßen der Modelle verglichen. Abbildung 6.22 und Abbildung 6.23 zeigen beispielhaft die Ergebnisse für die Segmentierung und Rekonstruktion des künstlichen Bulbus, der aus sieben und acht Zentimeter großen Plastikugeln bestand.

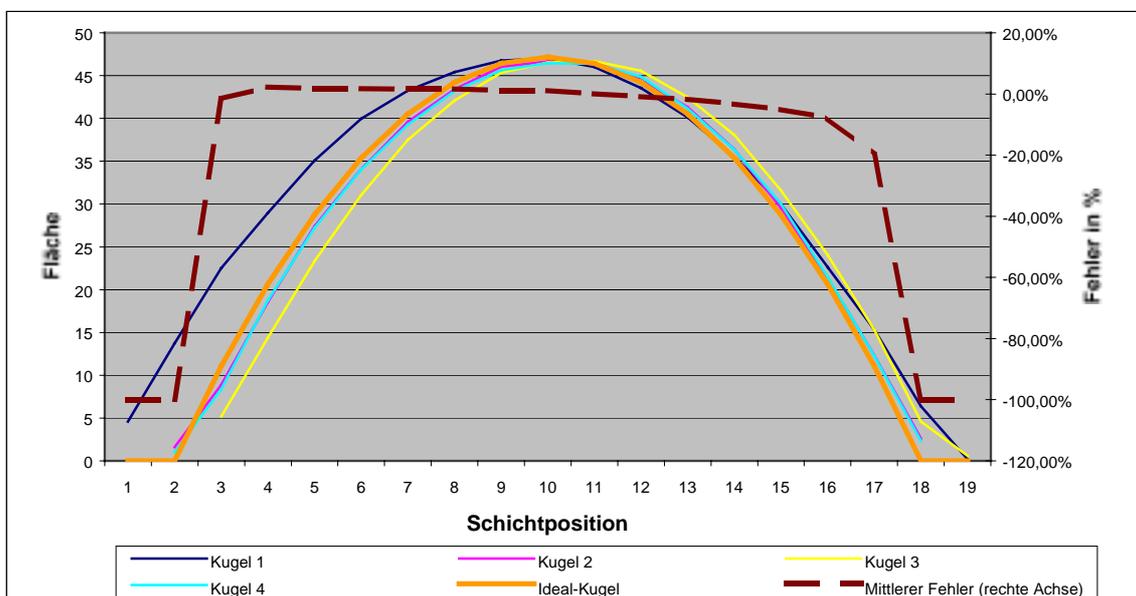


Abbildung 6.23: Grafische Darstellung der berechneten und idealen Flächen

Die Abbildung zeigt die berechneten und idealen Flächen für das künstliche Augenmodell mit einem Kugeldurchmesser von acht Zentimetern. Im Vergleich zu den Ergebnissen mit der Kugel mit einem Durchmesser von sieben Zentimetern ist die Korrektheit der segmentierten Bulbus-Konturen in vielen Schichten besser. Lediglich in den äußersten Randschichten geht der Fehler auf 100% hoch, da in Schichten mit starken Partialvolumeneffekten keine sinnvolle Segmentierung möglich ist.

Die Ergebnisse der acht und sieben Zentimeter großen Modelle zeigten, daß die zentralen Schnittebenen des Auges sowohl absolut als auch relativ auf die erwarteten Flächen bezogen besser segmentiert werden können als die Randschichten. Die Partialvolumeneffekte sind damit als eine der Hauptursachen für die Fehler bei der manuellen Segmentierung zu sehen. Diese Vermutung wird durch die in Tabelle 6-4 abgebildeten Ergebnisse bezüglich der Berechnung des Bulbusvolumens unterstützt. Der unterschiedliche mittlere Fehler bei dem acht und sieben Zentimeter großen Modell resultiert voraussichtlich aus der Tatsache, daß bei einem größeren Modell mehr Schichten des Auges durch die Partialvolumeneffekte betroffen sind.

Neben den quantitativen Ergebnissen konnten einige weitere Eigenschaften der manuellen Segmentierung festgestellt werden. So spielt die Reihenfolge, in der die einzelnen anatomischen Strukturen segmentiert werden, durchaus eine Rolle. Gerade bei der Segmentierung von Ziliarkörper und Iris beeinflusst

die Segmentierung des einen die Segmentierung des zweiten, da Form und Position beider Organe eng miteinander verbunden sind.

<i>Kugeldurchmesser</i>	<i>Durchschnittliche gemessenes Volumen</i>	<i>Ideal-Volumen</i>	<i>Mittlerer Fehler</i>	<i>Mittlere Standardabweichung</i>
7	165,651	172,0069102	3,84%	3,70%
8	242,3785	258,1546167	6,51%	2,46%

Tabelle 6-4: Zusammenfassung der Segmentierungskorrektheit bei den sieben und acht Zentimeter großen Phantommodellen

Die Tabelle zeigt die Ergebnisse des Vergleichs von manuellen Segmentierungsergebnissen und der aufgrund der bekannten Geometrie berechneten Volumina.

Darüber hinaus wurde durch alle Testpersonen bestätigt, daß die Durchführung der manuellen Segmentierungen eine nicht unbeträchtliche Belastung darstellt. Bei den Testpersonen zeigte sich nach ca. 20 bis 30 Minuten ein spürbares Nachlassen der Aufmerksamkeit und der Motivation. Gerade bei der medizinischen Fachkraft war zu spüren, daß nach zwei bis drei segmentierten Datensätzen die Bereitschaft für weitere Segmentierungsarbeiten deutlich abnahm. Verringerte Aufmerksamkeit und Motivation spiegelten sich vor allem darin wider, daß die Konturpunkte zunehmend weiter voneinander entfernt positioniert wurden und daß die Sorgfalt, mit der die Punkte plazierte wurde, spürbar nachließ. Darüber hinaus stellt die Segmentierung auch eine psychische Belastung dar, der sich die Testpersonen zunehmend ungern aussetzten.

Die Segmentierung der Testdatensätze hat bestätigt, daß eine manuelle Segmentierung der anfallenden Datensätze nicht praktikabel ist. Der erforderliche Arbeitsaufwand liegt bereits für die Routinedatensätze mit lediglich 15 Datenschichten im Bereich von mehr als einer halben Stunde und ist damit deutlich zu lang. Datensätze erhöhter Auflösung werden in der klinischen Routine nicht genutzt werden können. Darüber hinaus sind manuelle Segmentierungen so inkonsistent und fehlerbehaftet, daß die ursprünglich gestellten Anforderungen an das Bestrahlungsmodell nicht erfüllt werden können. Doch wenn bereits die manuelle Segmentierung nicht die Genauigkeitskriterien erfüllt, müssen gegebenenfalls die Anforderungen an das entwickelte Bildverarbeitungssystem modifiziert werden, um eine realistische Zielsetzung zugrunde zu legen. Dementsprechend hat nach den Segmentierungsarbeiten auch die medizinische Fachkraft festgestellt, daß bereits ein sehr wichtiges Ziel erreicht wäre, wenn die Segmentierungen konsistent wären, da dadurch eine Berechenbarkeit der Ergebnisse gegeben wäre. Gerade auch aufgrund des Mangels eines tatsächlichen, quantitativen Goldstandards, für den die manuelle Segmentierung offensichtlich nicht geeignet ist, ist die sogenannte Segmentierungskorrektheit als sekundär anzusehen.

Die Ergebnisse der manuellen Segmentierung haben einerseits die Bedeutung der modellbasierten Bildverarbeitung unterstrichen. Andererseits erscheint es sinnvoll zu sein, beim Aufbau der Referenzdatenbank möglichst wenige Datensätze manuell zu segmentieren, um die Konsistenz der Datenbank nicht zu gefährden. Statt dessen sollten die Referenzmodelle so weit wie möglich durch die modellbasierte Bildverarbeitung segmentiert werden, um dadurch eine möglichst einheitliche Segmentierung sämtlicher Referenzmodelle zu erreichen. Um interindividuelle Unterschiede auszuschließen, sollten außerdem die manuell angefertigten Datensätze ausschließlich von einer Person angefertigt werden.

6.3.1.2 Integration der Referenzdatenbank

Die Referenzdatenbank wird entsprechend dem Case-based Reasoning Ansatz aus einer Sammlung von manuell segmentierten Referenzmodellen zusammengesetzt. Bei der Umsetzung von fallbasierten Exper-

tensystemen mit großen Falldatenbanken ist eine hierarchische Strukturierung der Fälle üblich, um die Suche nach einem ähnlichen Referenzmodell zu beschleunigen. Wenngleich für die modellbasierte Bildverarbeitung zunächst eine vergleichsweise kleine Falldatenbank vorgesehen ist, bietet sich eine hierarchische Strukturierung der Referenzmodelle an. Auf der obersten Hierarchieebene werden die Referenzmodelle anhand der Bildmodi CT-, MRT-T1 und MRT-T2 unterschieden. Anschließend werden auf der darunter liegenden Ebene die Datensätze nach der Schichtorientierung (sagittal, coronar und axial) unterteilt. In jeder Kategorie wird ein Master-Referenzmodell dadurch markiert, daß das Datei-Suffix von „ref“ nach „mref“ verändert wird. Die Auswahl der Master-Referenzmodelle wurde durch eine geschulte Fachkraft vorgenommen. Die Auswahl des ähnlichsten Datensatzes innerhalb einer Kategorie wird anhand der Ähnlichkeitsparameter vorgenommen.

In dieser Arbeit wurde zunächst eine vergleichsweise einfache Form der Ähnlichkeitsberechnung vorgenommen. Eine detaillierte Untersuchung dieser Fragestellung findet sich in [Kun01]. Als primäre Ähnlichkeitsparameter wurden zunächst der mittlere Grauwert und die Standardabweichung genutzt. Es wird dabei jener Datensatz als Referenzmodell ausgewählt, dessen mittlere Abweichung in den Ähnlichkeitsparametern am geringsten ist. Die mittlere Ähnlichkeit wird berechnet, indem für jede Schicht vier möglichst weit auseinander liegende Punkte ausgewählt werden. Anschließend wird in einem Suchfenster der mittlere Grauwert und die Standardabweichung berechnet. Die Größe des Suchfensters wurde bei den Tests leicht variiert. Grundsätzliche Unterschiede konnten dabei jedoch nicht festgestellt werden. Abschließend werden die mittleren quadratischen Differenzen zwischen den Ähnlichkeitsparametern des aktuellen Datensatzes und jenen der Referenzmodelle der ausgewählten Kategorie berechnet. Es wird jener Datensatz als Referenzmodell ausgewählt, der die geringste mittlere quadratische Differenz in den Ähnlichkeitsparametern zum aktuellen Datensatz aufweist.

Um den ganzen Selektionsprozeß zu beschleunigen, werden die zur Verfügung stehenden Referenzmodelle und Master-Referenzmodelle und die dazu gehörigen Ähnlichkeitsparameter vor der Anwendung der modellbasierten Bildverarbeitung berechnet und in einer Datei abgespeichert. Beim Starten von JDisplay wird diese Datei eingeladen, so daß der Zugriff auf geeignete Referenzmodelle und Master-Referenzmodelle ohne Verzögerung vorgenommen werden kann. Sämtliche erforderlichen Informationen werden in der Datei im XML-Format abgelegt, da in diesem Format die Informationen gut strukturiert kodiert und gleichzeitig problemlos mit einem Browser betrachtet werden können. Da JDK 1.3 bereits mit einer XML-Schnittstelle ausgestattet ist und Parser und Schreibfunktionen für das Lesen und Schreiben von XML-Dokumenten anbietet, konnte der Zugriff auf die Referenzdatenbank komfortabel ohne größeren Implementationsaufwand umgesetzt werden.

Ähnlich wie bei dem Case-based Reasoning Ansatz können nur schwer Aussagen über die Vollständigkeit der Referenzdatenbank gemacht werden, da es in der Regel keinen eindeutig begrenzten Raum bezüglich der Variabilität der anfallenden Datensätze gibt. Diese Feststellung gilt für viele Anwendungsbereiche der künstlichen Intelligenz ebenso wie für die medizinische Bildverarbeitung. Es ist davon auszugehen, daß nicht ein einziger medizinischer Bilddatensatz eines Patienten mit dem eines anderen Patienten vollständig übereinstimmt, da sowohl die Anatomie als auch der Bildgebungsprozeß nicht-deterministischen Schwankungen mit einer nicht unbeträchtlichen Variabilität unterliegen.

Aus dieser Tatsache folgt eine wichtige Anforderung an das gesamte Bildverarbeitungssystem: Nicht die Anzahl der in der Referenzdatenbank gespeicherten Fälle stellt die Anwendbarkeit der modellbasierten Bildverarbeitung sicher, sondern die Toleranz des gesamten Bildverarbeitungsansatzes gegenüber Unterschieden zwischen aktuellem Datensatz und Referenzdatensatz. Diese Forderung erschwert jedoch die Frage nach der Vollständigkeit der Referenzdatenbank, da sie nun nur noch durch die Anwendung des Systems auf das anfallende Bildmaterial beantwortet werden kann. Die Nicht-Vollständigkeit wird immer dann bewiesen werden, wenn ein Datensatz nicht durch das bestehende System mit der geforderten Ge-

nauigkeit segmentiert werden kann. Ein wichtiger Parameter bei der Bestimmung der Vollständigkeit stellt demnach auch die Genauigkeitsanforderung dar, die durch das System erfüllt werden soll.

Diese Überlegungen veranschaulichen, daß keine systematischen Vorgaben gemacht werden können, um die Vollständigkeit zu gewährleisten. Daher wurde für die Austestung des Systems die Vorgehensweise angenommen, daß eine willkürliche Anzahl von Referenzmodellen der Datenbank hinzugefügt wurden. Aufgrund der geringen Anzahl von zur Verfügung stehenden Fällen wurden für die Phantomdaten, für die Probandendaten und die Aufnahmen der Schweineaugen keine Referenzdatenbanken angelegt. Für das medizinische Bildmaterial wurde eine Referenzdatenbank mit 25 Referenzmodellen angelegt.

6.3.2 Bestimmen des Bildkontextes

Das Bestimmen des allgemeinen Kontextes und die Auswahl des Master-Referenzmodells werden interaktiv durch den Benutzer vorgegeben oder automatisch anhand der Informationen im Bild-Header vorgenommen und müssen funktional nicht evaluiert werden. Von Relevanz ist bei der Bestimmung des Bildkontextes der Zeitaufwand, der für die Durchführung dieser Schritte einschließlich der Eingabe der Landmarken notwendig ist und die Korrektheit der landmarkenbasierten Registrierung.

Der Zeitaufwand für die interaktiven Schritte zur Bestimmung des Bildkontextes stellt eine wichtige Größe dar, da dies der gesamte interaktive Arbeitsaufwand für den Benutzer darstellt. Die Korrektheit der landmarkenbasierten Registrierung beeinflußt maßgeblich die Beschreibung des aktuellen Datensatzes, da anhand der transformierten Segmentierungskonturen die Kanten- und Regioneneigenschaften berechnet werden. Die Berechnung der Beschreibungsparameter wird in dieser Arbeit nur ansatzweise beschrieben, da die Auswahl des Referenzmodells durch Berechnung von Ähnlichkeitsparameter detailliert in [Kun01] behandelt wird.

6.3.2.1 Arbeitsaufwand

Der Arbeitsaufwand für die Bestimmung des Bildkontextes besteht einerseits aus dem Zeitaufwand für die interaktiven Eingaben, andererseits aus der Rechenzeit für die Anpassung des Master-Referenzmodells an den aktuellen Datensatz und die Berechnung der Bildbeschreibungparameter. Die Bedeutung dieser Zeitaufwände ist unterschiedlich einzustufen. Während die interaktiven Eingaben den tatsächlichen Arbeitsaufwand für den Benutzer darstellen, können die Berechnungen auch offline durchgeführt werden und sind damit für den Benutzer irrelevant. Zielsetzung des Bildverarbeitungsansatzes ist es, den Arbeitsaufwand für den Nutzer zu minimieren. Entscheidend sind daher vor allem die Verarbeitungsschritte zwischen der ersten und letzten interaktiven Eingabe des Benutzers. Für die Tests wurden daher die Zeiten zwischen dem Starten der modellbasierten Bildverarbeitung und der letzten eingegebenen Landmarke gemessen. Der Zeitaufwand für die Offline-Berechnung der landmarkenbasierten Registrierung und der Bildbeschreibungparameter wird zwar ebenfalls bestimmt, ist für die Gesamtbeurteilung der modellbasierten Bildverarbeitung irrelevant. Da ebenfalls alle nachfolgenden Verarbeitungsschritte automatisch durchgeführt werden, stellen sie keinen für den Nutzer relevanten Arbeitsaufwand dar.

Bei der Feststellung des Zeitaufwandes für die Benutzer wurden unterschiedliche Stufen der Vorkenntnisse berücksichtigt, um die Benutzerfreundlichkeit bei geringen oder keinen Vorkenntnissen mit dem System festzustellen. Es wurden die Zeiten von sechs Testpersonen gemessen. Vier Testpersonen hatten noch nie mit der modellbasierten Segmentierung gearbeitet. Zwei Testpersonen waren bereits routiniert mit dem Umgang der entwickelten Software. Abbildung 6.24 zeigt die Aufwände, die für die Eingabe der Landmarken bei fünf beispielhaften Datensätzen benötigt wurden. Die Tests wurden mit axialen MRT-T2 Datensätzen durchgeführt. Für die Bearbeitung von anderen Datensätzen bezüglich Schichtorientierung oder Bildmodus ist damit zu rechnen, daß der Benutzer sich zunächst auf das Bildmaterial ein-

stellen muß. Lernkurve und erforderlicher Zeitaufwand werden sich jedoch von den Testdatensätzen nicht grundsätzlich unterscheiden.

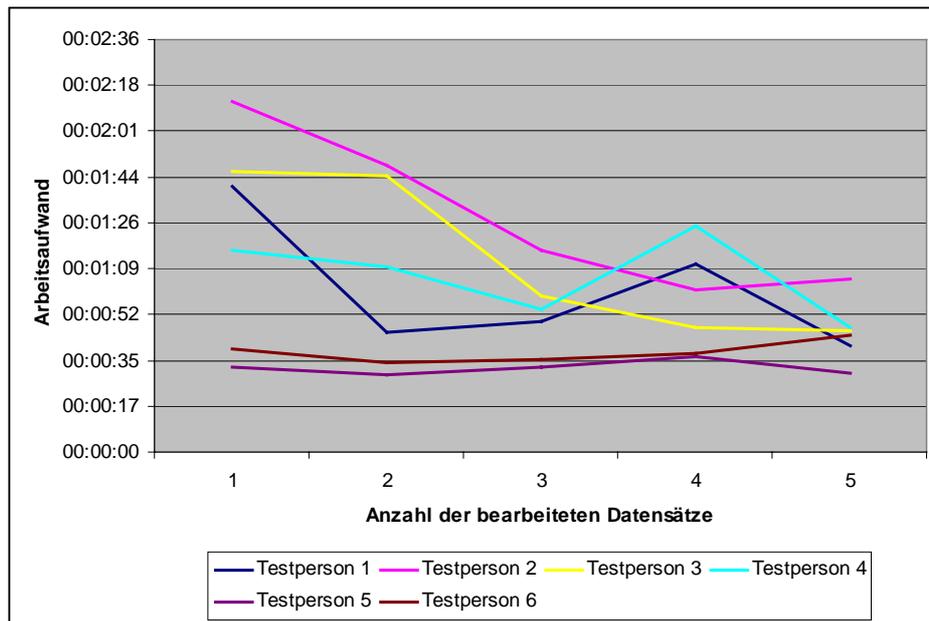


Abbildung 6.24: Zeitaufwände für die Eingabe der Landmarken

In der Abbildung werden die Zeiten dargestellt, die sechs unterschiedliche Testpersonen für die Eingabe der Landmarken benötigten. Während die Testpersonen 1 bis 4 die Software zum ersten Mal benutzten, waren die Testpersonen 5 und 6 bereits geübt in der Nutzung der Software. Die geübten Testpersonen benötigten zwischen 28 und 40 Sekunden für die Eingabe der Landmarken. Die ungeübten Testpersonen benötigten bei den ersten Anwendungen deutlich längere Zeiten für die Eingabe der sechs Landmarken. Die Lernphase bei sämtlichen Testpersonen war jedoch außerordentlich kurz. Schon nach sechs Datensätzen lagen die Zeiten für die Positionierung der Landmarken im Bereich der routinierten Testpersonen.

Wie in Abbildung 6.24 zu sehen ist, liegt der Arbeitsaufwand für einen geübten Nutzer bei zirka 30 Sekunden. Ungeübte Nutzer benötigen bei den ersten Malen zwischen einer und zwei Minuten. Doch bereits nach fünf segmentierten Datensätzen ist für nahezu alle Testpersonen der Arbeitsaufwand auf unter 40 Sekunden gefallen. Der steile Abfall in den Arbeitsaufwänden scheint die hohe Benutzerfreundlichkeit der Software zu belegen. Es waren lediglich Einführungen von wenigen Minuten notwendig, um den Testpersonen die Verwendung der Software zu erläutern. Schon nach dem zweiten Datensatz waren keine Erläuterungen mehr nötig. Geringe Schwankungen im Arbeitsaufwand resultierten nach dem dritten Datensatz lediglich aus der Komplexität der Bilddaten. Sämtliche Testpersonen empfanden die Nutzung der Software als komfortabel und angenehm. Es kam zu keinem Konzentrationsverlust. Statt dessen schien die Nutzung der Software sogar das Interesse zu wecken. Solche auf den ersten Blick trivialen Feststellungen sind gerade in Hinsicht auf eine Einführung der Software in die klinische Routine gegebenenfalls von großer Bedeutung.

Neben dem rein interaktiven Arbeitsaufwand für den Benutzer zur Eingabe der Landmarken müssen zur Bestimmung des allgemeinen Bildkontextes einige weitere Verarbeitungsschritte durchgeführt werden. Hierzu gehört das Laden des aktuellen Datensatzes, das Laden des Master-Referenzmodells (automatisch von der Software durchgeführt) und das Anpassen des Referenzmodells. Wenngleich der Benutzer bei diesen Schritten keine Eingaben machen muß, sind dies gegebenenfalls Wartezeiten, die mit eingerechnet werden müssen. Tabelle 6-5 zeigt beispielhaft für fünf Datensätze die Zeiten, die für die Durchführung jeweils benötigt wurden.

Datensatz	aktuellen Datensatz laden	Referenzmodell laden	Referenzmodell anpassen	Berechnung der Ähnlichkeitsparameter
MRT-Datensatz 1	00:00:05	00:00:13	00:00:16	00:00:06
MRT-Datensatz 2	00:00:04	00:00:10	00:00:18	00:00:08
MRT-Datensatz 3	00:00:04	00:00:11	00:00:17	00:00:07
MRT-Datensatz 4	00:00:03	00:00:21	00:00:16	00:00:05
MRT-Datensatz 5	00:00:05	00:00:12	00:00:16	04:00:07

Tabelle 6-5: Zeiten für die Bestimmung des Bildkontextes

Die Tabelle zeigt die gemessenen Zeiten für die Durchführung der verbleibenden Schritte, die zur Bestimmung des Bildkontextes durchgeführt werden müssen. Wenngleich diese Verarbeitungsschritte weitgehend automatisch durchgeführt werden, stellen sie gegebenenfalls Wartezeiten für den Nutzer dar, die mit eingerechnet werden müssen. Sämtliche Arbeitsschritte liegen im Sekundenbereich und sind somit für die Anwendung in der klinischen Routine geeignet.

Sämtliche in Tabelle 6-5 abgebildeten Zeiten für die Durchführung der verbleibenden Schritte zur Durchführung der Anpassung des Referenzmodells liegen im Sekundenbereich und sind damit für die Anwendung in der klinischen Routine absolut geeignet.

6.3.2.2 Landmarkenbasierte Registrierung

Die landmarkenbasierte Registrierung stellt eine zentrale Funktion für die modellbasierte Bildverarbeitung dar, da sie für die Anpassung des Master-Referenzmodells an den aktuellen Datensatz, für die Anpassung des Referenzmodells an den aktuellen Datensatz, für die modellbasierte Interpolation und für die modellbasierte Registrierung benötigt wird. Bei allen vier Anwendungen besteht die Zielsetzung der landmarkenbasierten Registrierung darin, den einen Datensatz oder die Segmentierung dieses Datensatzes möglichst sinnvoll an einen anderen Datensatz anzupassen. Relevante Fragestellungen für die Tests sind dementsprechend, wie präzise Datensatz und/oder Segmentierung an den anderen Datensatz durch Nutzung der landmarkenbasierten Registrierung angepaßt wird und wie gut dabei anatomische Vorgaben eingehalten werden. Der Rechenaufwand spielt aufgrund der Möglichkeit der Offline-Berechnung eine untergeordnete Rolle, wird aber für eine möglichst interaktive Nutzung der Registrierung auch untersucht.

Als Testgrundlage wurden die verschiedenen Testdatensätze herangezogen. Um die Ergebnisse der landmarkenbasierten Registrierung quantitativ auswerten zu können, wurden die manuellen Segmentierungen der Datensätze als Goldstandard zugrunde gelegt (vergleiche Abschnitt 6.3.1). Es wurden die vorhandenen Testdatensätze untereinander angewandt. Bei den Patientendaten wurde dabei nicht auf die Ähnlichkeit von Referenzmodell und aktuellen Datensatz geachtet, um einen Eindruck davon zu bekommen, wie hoch die Variabilität der Anpassungsergebnisse ist. Nach der landmarkenbasierten Registrierung wurde die Korrektheit der Anpassung abgeschätzt, indem die angepaßte Segmentierung mit dem Goldstandard, der manuellen Segmentierung des gleichen Datensatzes, verglichen wurde. Zur Beschreibung der Segmentierungsübereinstimmungen wurden die Schnittflächen in den Bildschichten, die Fehlerflächen, in denen die Segmentierungen nicht übereinstimmen und das Schnittvolumen berechnet.

Zur Disposition standen die vier in Abschnitt 4.1.6.4 beschriebenen Registrierungsverfahren: landmarkenbasierte Registrierung anhand von affinen Transformationen, landmarkenbasierte Registrierung anhand von affinen Transformationen und Scherung, landmarkenbasierte Registrierung anhand von affinen Transformationen und anschließendem lokalem Morphing und landmarkenbasierte Registrierung anhand einer proprietären Anpassungsstrategie.

Bereits bei ersten Tests zeigte sich, daß der Rechenaufwand für die affine Transformation mit anschließendem lokalem Morphing außerordentlich groß ist, während die Ergebnisse die bereits erwähnten spezifischen Probleme aufwiesen. Die Berechnung der Transformationsmatrix mittels Radial Basis Funktion benötigte in Einzelfällen mehr als 50 Minuten, während die Ergebnisse nicht besser waren als die der

affinen Transformationen, die innerhalb von wenigen Sekunden durchgeführt werden konnten. Darüber hinaus traten die bereits erwähnten Verzerrungen der Organoberflächen auf, die eine sinnvolle Nutzung der internen Energie deutlich erschweren würde. Dieses Mißverhältnis von hohem Rechenaufwand und geringer Anpassungsqualität führte dazu, daß die landmarkenbasierte Registrierung mittels affiner Transformation und anschließendem lokalem Morphing verworfen wurde. Bei den Tests wurden dementsprechend lediglich die verbleibenden drei Registrierungsansätze getestet:

- Landmarkenbasierte Registrierung anhand von affinen Transformationen (affine Anpassung),
- Landmarkenbasierte Registrierung anhand von affinen Transformationen und Scherung (nicht-affine Anpassung) und
- Landmarkenbasierte Registrierung anhand einer proprietären Anpassungsstrategie (proprietäre Anpassung).

Tabelle 6-6 faßt die Ergebnisse zusammen, die bei der Anwendung der Algorithmen auf den Bulbus in den Phantomdatensätzen erzielt wurden. Dabei wurden die Übereinstimmungen der Registrierungsverfahren in Verhältnis gesetzt zu den Ergebnissen der interindividuellen und intraindividuellen manuellen Segmentierungen aus Abschnitt 6.3.1.1. In der Tabelle ist zu sehen, daß ähnlich wie bei den manuellen Segmentierungsverfahren auch bei den Registrierungsverfahren die größten Unterschiede zwischen der Mustersegmentierung und den jeweiligen Ergebnissegmentierungen in den Randschichten des Bulbus auftreten. Im direkten Vergleich zeigt sich jedoch, daß bereits die landmarkenbasierten Registrierungsverfahren Segmentierungsergebnisse erzielen, die vergleichbar sind mit den Ergebnissen der interindividuellen Segmentierungen. Da die Form des Augenmodells sehr regelmäßig ist, stellt dieses Ergebnis noch keine Überraschung dar. Die guten Werte belegen jedoch, daß das Konzept einer Modellanpassung zur Unterstützung einer nachfolgenden Segmentierung sinnvoll funktioniert und bereits gute Ergebnisse erzielt.

	Mittelwert interindividuelle manuelle Segmentierung	Standard- abweichung interindividuelle manuelle Segmentierung	Mittelwert intraindividuelle manuelle Segmentierung	Standard-abweichung intraindividuelle manuelle Segmentierung	Mittelwert proprietäre Anpassung	Standard- abweichung proprietäre Anpassung	Mittelwert affine Anpassung	Standard- abweichung affine Anpassung	Mittelwert nicht- affine Anpassung	Standard- abweichung nicht- affine Anpassung
Slice 0	28,43%	35,32%	73,45%	12,54%	47,35%	15,34%	59,74%	14,38%	22,00%	3,10%
Slice 1	44,32%	12,37%	80,49%	4,87%	44,43%	13,78%	59,73%	13,80%	38,10%	8,47%
Slice 2	66,96%	8,75%	84,99%	3,21%	66,08%	13,08%	79,56%	6,46%	65,89%	7,89%
Slice 3	77,42%	7,58%	96,01%	1,19%	85,00%	5,29%	86,58%	5,05%	80,68%	3,76%
Slice 4	93,21%	8,51%	95,32%	0,94%	92,91%	2,04%	91,36%	2,99%	88,99%	1,42%
Slice 5	95,38%	4,34%	98,01%	0,68%	95,86%	1,29%	93,33%	2,27%	92,26%	1,10%
Slice 6	94,32%	3,54%	97,28%	3,51%	97,00%	0,92%	93,76%	1,79%	94,25%	0,80%
Slice 7	95,21%	2,94%	95,37%	1,45%	96,94%	0,82%	93,87%	1,63%	94,91%	0,63%
Slice 8	93,32%	1,54%	96,35%	2,04%	97,03%	1,15%	93,12%	2,01%	95,19%	0,90%
Slice 9	87,01%	2,76%	96,81%	0,69%	95,31%	2,53%	93,16%	2,05%	95,01%	1,10%
Slice 10	97,94%	5,12%	92,36%	4,07%	96,41%	1,70%	92,48%	2,57%	93,22%	2,46%
Slice 11	96,21%	0,56%	96,96%	0,31%	95,13%	1,92%	91,41%	3,35%	93,24%	1,66%
Slice 12	94,10%	5,54%	94,27%	4,51%	95,54%	2,20%	90,31%	4,28%	91,05%	2,21%
Slice 13	94,98%	4,34%	98,21%	1,56%	94,58%	2,57%	88,13%	5,91%	86,82%	3,39%
Slice 14	90,13%	9,60%	94,57%	2,65%	93,11%	2,90%	83,84%	9,15%	80,57%	8,11%
Slice 15	93,26%	5,34%	95,58%	1,18%	91,16%	3,05%	78,67%	11,66%	78,66%	3,45%
Slice 16	91,34%	6,54%	88,15%	4,23%	82,39%	5,74%	75,21%	5,24%	61,27%	6,88%
Slice 17	67,31%	7,32%	76,74%	9,78%	68,26%	8,24%	37,56%	15,94%	34,56%	8,02%
Slice 18	63,47%	18,34%	66,74%	10,65%	51,45%	8,65%	31,34%	0,00%	3,48%	0,00%

Tabelle 6-6: Ergebnisse der Tests mit Phantomdatensätzen

Die Tabelle zeigt die Ergebnisse der Tests mit den Phantomdatensätzen. Es wurden zur Berechnung dieser Tabelle 73 Tests durchgeführt. Abgebildet werden die Übereinstimmungen der manuell angefertigten Musterlösungen mit der interindividuellen Segmentierung, den intraindividuellen Segmentierungen, den angepaßten Segmentierungen mittels proprietärer Anpassung, affiner Transformation und der nicht affinen Transformation. Die Übereinstimmungen wurden berechnet, indem die Schnittfläche zu der mittleren Gesamtfläche in der jeweiligen Schicht in Verhältnis gesetzt wurde. Der Wert 100% entspricht einem vollständigen Übereinstimmen von Mustersegmentierung und der Ergebnissegmentierung. Bei 0% gibt es keine Schnittfläche zwischen Mustersegmentierung und der Ergebnissegmentierung. Die Standardabweichung zu den jeweiligen mittleren Übereinstimmungen beschreiben, wie stark die Übereinstimmungen zwischen den verschiedenen Tests variierten.

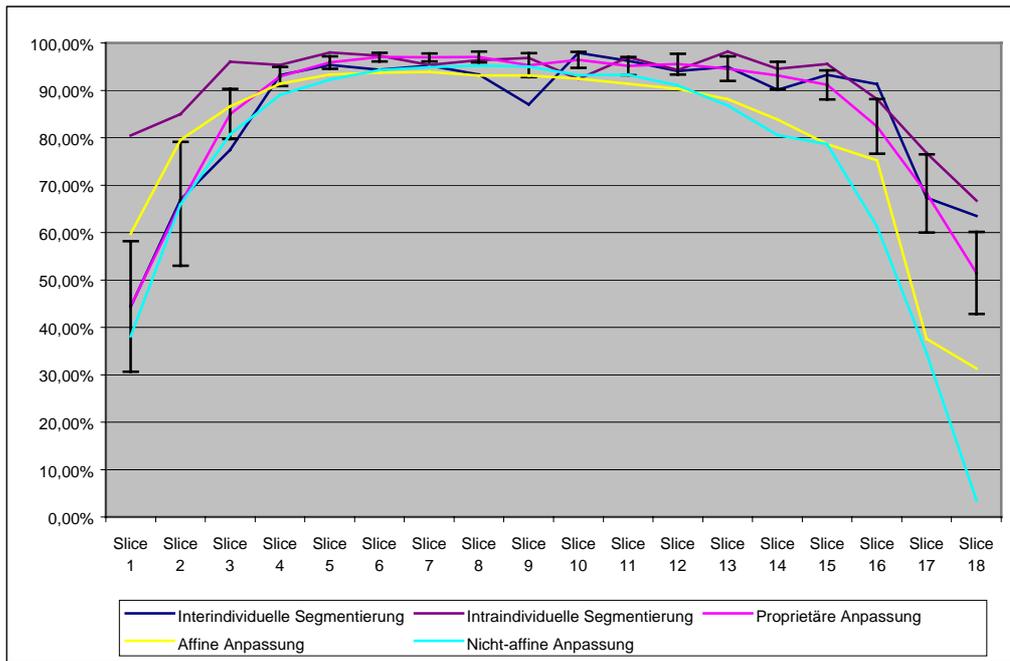


Abbildung 6.25: Grafische Darstellung der Ergebnisse aus Tabelle 6-6

Die Abbildung veranschaulicht die Ergebnisse aus Tabelle 6-6 grafisch, wodurch ein direkter Vergleich der einzelnen Ergebnisse über die Schichten hinweg erleichtert wird. Wie bei der manuellen Segmentierung fallen auch bei den Registrierungsverfahren die Übereinstimmungen in den Randschichten des Bulbus deutlich ab. Doch vor allem die proprietäre Anpassung liefert in nahezu allen Schichten gute Ergebnisse, die oft über denen der interindividuellen Segmentierung liegen und auch an die intraindividuellen Segmentierungen herankommen. Die Toleranzbalken bei der Kurve für die proprietäre Anpassung zeigen anhand der Standardabweichung der Ergebnisse, in welchem Bereich die Ergebnisse für die proprietäre Anpassungsstrategie variierten. Naturgemäß ist die Abweichung in den Randbereichen größer als in den zentralen Schichten. Die proprietäre Anpassung erreicht jedoch mit ihrem Toleranzbereich nahezu in allen Schichten die als Optimum anzusehende Korrektheit der intraindividuellen Segmentierung.

Abbildung 6.25 veranschaulicht die Ergebnisse für die Phantomdatensätze aus Tabelle 6-6 grafisch über die Schichten hinweg. Es ist sehr gut zu erkennen, daß die intraindividuellen manuellen Segmentierungen vor allem auch in den Randschichten die größten Übereinstimmungen aufweisen. Andererseits liegen die Ergebnisse der Registrierungsverfahren sehr eng bei den Ergebnissen der interindividuellen manuellen Segmentierungen. Dieses Ergebnis ist durchaus bemerkenswert, da eine Modellanpassung offensichtlich bereits ähnlich gute Segmentierungen produziert wie manuellen Segmentierungen von unterschiedlichen Personen ohne dabei den oben aufgeführten Arbeitsaufwand zu erfordern. Die besten Ergebnisse erzielte unter den Registrierungsverfahren die proprietäre Anpassung, die in fast allen Schichten besser ist als die interindividuelle Segmentierung und in vielen Bereichen auch an die intraindividuelle Segmentierung herankommt.

Schnittvolumen	Intraindividuelle Segmentierung	Interindividuelle Segmentierung	Proprietäre Anpassung	Affine Anpassung	Nicht-Affine Anpassung
Mittlere Übereinstimmung	97,91%	93,07%	93,98%	89,57%	88,93%
Standardabweichung	2,08%	4,62%	2,39%	5,55%	2,76%

Tabelle 6-7: Übereinstimmungen der Schnittvolumen der berechneten Segmentierungen mit den Mustersegmentierungen

Die besten Ergebnisse werden bei der intraindividuellen Segmentierung und mit der proprietären Registrierung erreicht. Das überraschende bei diesem Ergebnis ist, daß die proprietäre Anpassung der Referenzmodelle an den aktuellen Datensatz bereits bessere Ergebnisse lieferte als die interindividuelle Segmentierung.

Deutlich werden die guten Ergebnisse der Registrierungsverfahren auch bei der Betrachtung der mittleren Übereinstimmung in den Schnittvolumen in Tabelle 6-7. Vor allem die proprietäre Registrierung liefert bereits konsistentere Ergebnisse als die interindividuelle Segmentierung. Das bedeutet, daß allein durch die Registrierung einer Musterlösung mit einem aktuellen Datensatz bessere Ergebnisse erzielt werden können als durch die Anfertigung der Segmentierung durch mehrere unterschiedliche Personen.

	Mittelwert interindividuelle manuelle Segmentierung	Standard- abweichung interindividuelle manuelle Segmentierung	Mittelwert intraindividuelle manuelle Segmentierung	Standard- abweichung intraindividuelle manuelle Segmentierung	Mittelwert proprietäre Anpassung	Standard- abweichung proprietäre Anpassung	Mittelwert affine Anpassung	Standard- abweichung affine Anpassung	Mittelwert nicht- affine Anpassung	Standard- abweichung nicht- affine Anpassung
Slice 0										
Slice 1	0,52%	0,37%	73,00%	0,00%	26,73%	17,42%	17,40%	7,31%	1,41%	0,00%
Slice 2	34,02%	9,60%	80,58%	4,01%	46,16%	17,97%	39,06%	14,36%	36,12%	0,07%
Slice 3	65,97%	4,59%	90,69%	2,65%	56,02%	14,50%	48,35%	13,75%	50,02%	8,50%
Slice 4	76,64%	5,81%	95,76%	0,69%	83,93%	3,82%	81,70%	4,07%	81,48%	2,96%
Slice 5	92,80%	1,14%	94,80%	1,65%	95,18%	1,11%	73,31%	14,76%	85,14%	1,91%
Slice 6	94,08%	1,35%	98,03%	0,21%	95,28%	0,95%	76,82%	11,85%	85,39%	1,14%
Slice 7	93,42%	0,88%	93,42%	2,52%	96,07%	0,78%	87,20%	3,49%	79,48%	6,02%
Slice 8	93,69%	0,69%	97,56%	0,17%	94,73%	0,98%	86,73%	3,14%	80,41%	5,86%
Slice 9	94,51%	0,79%	95,22%	1,53%	91,92%	2,45%	77,85%	8,33%	82,06%	2,86%
Slice 10	86,05%	6,37%	95,58%	0,89%	88,40%	3,22%	72,09%	9,54%	78,96%	2,74%
Slice 11	91,44%	1,39%	84,75%	3,14%	85,03%	2,96%	60,48%	13,66%	60,08%	12,34%
Slice 12	63,14%	7,76%	62,74%	10,83%	75,87%	4,22%	69,85%	6,43%	48,72%	10,48%
Slice 13	1,60%	0,80%			51,50%	14,00%	47,23%	15,22%	30,53%	10,74%
Slice 14					20,75%	0,00%	51,35%	0,00%	1,73%	0,75%

Tabelle 6-8: Ergebnisse der Registrierungsverfahren bei Patientendaten

Die Tabelle zeigt die Ergebnisse der interindividuellen und intraindividuellen Segmentierung und stellt sie den Ergebnissen der Registrierungsverfahren gegenüber. Es wurden 67 Tests zur Erstellung dieser Tabelle durchgeführt. Die Übereinstimmungen in den Segmentierungen mit den Musterlösungen sind bei allen Ansätzen unter denen der Modelldatensätze. Nichtsdestotrotz setzen sich die Ergebnisse der Modelldatensätze auch in den Patientendatensätzen ähnlich fort.

Etwas andere Ergebnisse sind bei der Auswertung von Patientendaten zu erwarten, da hier die zu segmentierenden Strukturen nicht so regelmäßig sind wie bei den Modelldatensätzen. Tabelle 6-8 faßt die Ergebnisse von 67 Tests mit realen Patientendaten aus der klinischen Routine zusammen. Trotz der höheren Variabilität der Patientenanatomien setzen sich jedoch die grundsätzlichen Tendenzen der Test mit Phantomdatensätzen fort.

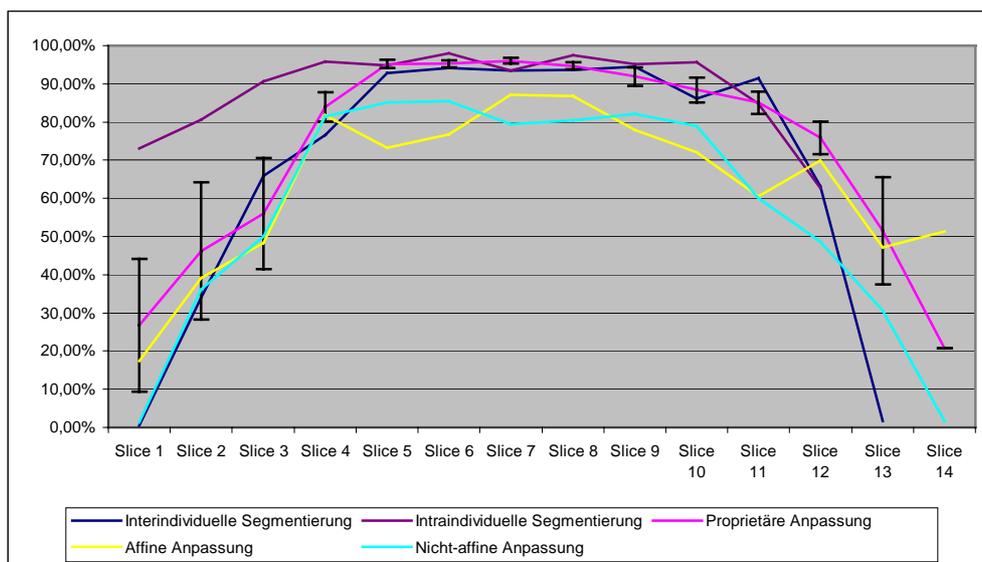


Abbildung 6.26: Ergebnisse der Registrierungsverfahren bei Patientendaten

Die Abbildung veranschaulicht die Ergebnisse mit Patientendaten aus Tabelle 6-8. Ähnlich wie bei den Phantomdatensätzen weist auch bei den Patientendatensätzen die intraindividuelle Segmentierung die konsistentesten Ergebnisse auf. Doch die proprietäre Anpassung der Referenzmodelle an den aktuellen Datensatz erzielt bereits bessere Ergebnisse als die interindividuelle Segmentierung.

Abbildung 6.26 veranschaulicht die Ergebnisse mit den Patientendaten aus Tabelle 6-8. Wie aus den Kurven gut zu erkennen ist, sind die Ergebnisse der proprietären Registrierung bereits besser als die interindividuellen Segmentierungen. Dieses Ergebnis ist durchaus erstaunlich, da mit dem geringen Arbeits- und Rechenaufwand bereits bessere Ergebnisse erzielt wurden als mit der sehr aufwendigen manuellen Segmentierung, die von unterschiedlichen Personen angefertigt wurden. Nichtsdestotrotz ist der Unterschied zu den intraindividuellen Segmentierungen in den Randschichten immer noch beträchtlich.

Der Vergleich der Schnittvolumina bestätigt die Ergebnisse, die durch den Vergleich der Schnittflächen gewonnen wurden. Die konsistentesten Ergebnisse produziert die intraindividuelle Segmentierung. Doch die proprietäre Registrierung von Referenzmodell und aktuellem Datensatz erreicht bereits bessere Ergebnisse als die interindividuelle manuelle Segmentierung.

Unter Berücksichtigung von sowohl den Phantom- als auch Patientendatensätzen kann festgestellt werden, daß insgesamt der Ansatz der modellbasierten Registrierung von Musterlösungen mit einem aktuell zu bearbeitenden Datensatz erfolgreich ist und die erwarteten Ergebnisse sogar übertrifft. Diese große Übereinstimmung der angepaßten Segmentierung mit der Musterlösung resultiert in erster Linie aus der Regelmäßigkeit des segmentierten Bulbus. Doch diese Regelmäßigkeit der Oberflächenform wird durch anderen Organe wie Leber, Lunge und Niere in ähnlichem Maß erfüllt wie durch das Auge.

	<i>Interindividuelle Übereinstimmung</i>	<i>Intraindividuelle Übereinstimmung</i>	<i>Proprietäre Anpassung</i>	<i>Affine Anpassung</i>	<i>Nicht-affine Anpassung</i>
Mittelwert	86,07%	90,91%	87,53%	77,25%	76,03%
Standardabweichung	1,27%	3,75%	4,39%	8,68%	5,17%

Tabelle 6-9: Globale Übereinstimmung der Registrierungsergebnisse im Schnittvolumen

Die Tabelle zeigt die Übereinstimmungen der manuellen Segmentierungen und der Registrierungsverfahren im Schnittvolumen. Dabei werden die Ergebnisse der einzelnen Schichten auch im Gesamtvolumen bestätigt. Die proprietäre Registrierung nähert die korrekte Segmentierung recht gut an und wird nur durch die intraindividuelle Segmentierung übertroffen.

Neben der allgemeinen Beurteilung des Ansatzes, eine aktuelle Segmentierung anzunähern, indem eine Musterlösung an einen aktuellen Datensatz angepaßt wird, sollte bei den Tests außerdem jenes Registrierungsverfahren ermittelt werden, das die besten Ergebnisse für die Segmentierung des Auges liefert. Dabei bestätigten die durchgeführten Tests die bereits in Abschnitt 4.1.6.4 beschriebenen Eigenschaften der jeweiligen Registrierungsansätze.

Die affinen Transformationen führen lediglich globale, geometrische Transformationen durch. Bei großer geometrischer Übereinstimmung beider Datensätze können mit diesem Ansatz gute Ergebnisse erzielt werden. So konnten vor allem bei den Phantomdatensätzen große Übereinstimmungen zwischen der transformierten Segmentierung und der Mustersegmentierung festgestellt werden. Die Standardabweichung ist vergleichsweise klein, was auf eine konsistente Ergebnisqualität hinweist. Bei den Patientendaten fällt die mittlere Übereinstimmung der Segmentierungen jedoch auf 77%, während die Standardabweichung auf fast 9% ansteigt.

Am deutlichsten erkennbar wird die schlechte Anpassung der Referenzmodelle vor allem durch die Abstände zwischen den Landmarken der angepaßten Segmentierung und der Landmarken der Mustersegmentierung. Die Ursache für die schlechten Ergebnisse ist bei dem Algorithmus zu suchen. Aufgrund der nur eingeschränkten Möglichkeiten der affinen Transformation ist ein Anpassen des Modells an alle Landmarken unmöglich (siehe Abbildung 6.27).

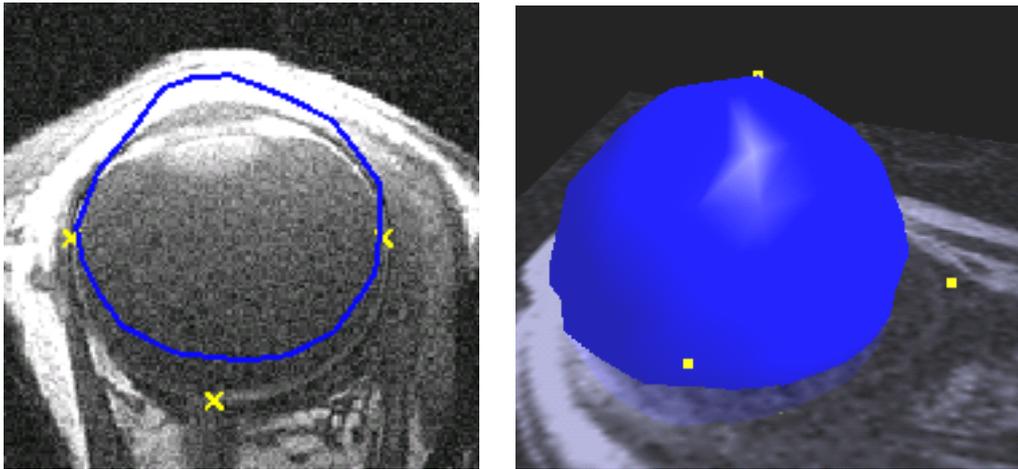


Abbildung 6.27: Beispiele für die Anpassungsfehler bei der affinen Transformation

Die Abbildungen zeigen beispielhaft das Problem, daß mit affinen Transformationen das Augenmodell nicht gleichzeitig an alle Landmarken angepaßt werden kann. Links ist die zweidimensionale Schicht zu sehen, in der die gelben Landmarken positioniert worden waren. Das in den Datensatz transformierte blaue Modell läuft nur durch die linke und rechte Landmarke. Die untere Landmarke ist noch weit von der Segmentierungskontur entfernt. Rechts ist die dreidimensionale Rekonstruktion der angepassten Segmentierung zu sehen. In der rechten Mitte des Bildes ist die gelbe Landmarke zu sehen, die weit von der Organoberfläche entfernt ist.

Selbst wenn die grundsätzliche Anatomie des Auges weitgehend einheitlich ist, treten bei jedem Patienten Abweichungen von der Optimalform des Gullstrandschen Auges auf. Diese Abweichungen sind in erster Linie lokal ausgeprägt, selbst wenn die grundsätzliche Form des Auges erhalten bleibt. Dabei scheinen vor allem leichte Asymmetrien, die sich auf einzelne Hemisphären, die Länge des Auges oder kleinere Oberflächensegmente erstrecken, üblich zu sein.

Die globale Anpassungsstrategie der affinen Transformationen kann diese Asymmetrien nicht kompensieren. Dadurch muß bei der Berechnung der Anpassungsfunktion ein Kompromiß gefunden werden, bei dem einige Landmarken besser in Übereinstimmung gebracht werden als andere. Besonders schlechte Kompromisse werden erzielt, wenn die obere oder untere Landmarke schlecht positioniert wurden, da sie den Anpassungskompromiß sehr stark beeinflussen. Da andererseits die Positionierung der oberen und unteren Landmarken durch die ausgeprägten Partialvolumeneffekte beeinträchtigt wird, sind schlechte Registrierungsergebnisse mit affinen Transformationen eher die Regel als die Ausnahme.

In Abschnitt 6.3.1 wurde erwähnt, daß ein wichtiger Anteil der Stabilität der modellbasierten Bildverarbeitung nicht aus der Vollständigkeit der Referenzdatenbank, sondern vor allem aus der stabilen und korrekten Anpassung der Segmentierung und damit des Referenzmodells an den aktuellen Datensatz resultiert. Die Ergebnisse der landmarkenbasierten Registrierung stellen die wichtigste Grundlage für eine solche korrekte Anpassung dar. Da gerade für das reale Bildmaterial die Ergebnisse der Anpassung durch affine Transformationen vergleichsweise großen Abweichungen auftreten können, gilt dieser Anpassungsansatz nicht als erste Wahl. Dabei ist der Hauptgrund für die schlechten Ergebnisse vor allem darin zu sehen, daß die Anpassung nur global durchgeführt wird, lokale oder asymmetrische Unterschiede aber keinesfalls kompensiert werden können.

Die affine Transformation mit Scherung (nicht-affine Anpassung) versucht die Anzahl der Landmarken, die bei der Anpassung des Modells einbezogen werden, zu erhöhen, indem die untere und die obere Hemisphäre des Bulbus an obere und untere Landmarken angepaßt werden. Bei den Tests hat sich jedoch gezeigt, daß diese Scherungen zu anatomisch nicht sinnvollen Verzerrungen des Auges führten. Diese Verzerrungen resultierten vor allem aus der Tatsache, daß die obere und untere Landmarke durch die starken Partialvolumeneffekte nur nach Gefühl plaziert werden können. Dabei treten oft translatorische

Verschiebungen innerhalb der Schicht auf, die zu einem Verzerren der globalen Bulbus-Oberflächen führt (vergleiche Abbildung 6.28).

Bei den Phantomdaten wirkte sich dieser Nachteil des Anpassungsverfahrens nicht so stark aus, da die Plazierung der Landmarken recht leicht war und die Oberfläche eine Kugeloberfläche ist. Sobald jedoch lokale Asymmetrien wie beim realem Bildmaterial auftreten und die Positionierung der oberen und unteren Landmarken deutlich erschwert ist, fällt die Übereinstimmung der angepaßten Segmentierung mit den Mustersegmentierungen auf 76% ab. Dabei resultieren die Fehlpositionierungen der Landmarken weniger aus der Inkompetenz des Nutzers als vielmehr aus der Tatsache, daß keine anatomischen Strukturen in diesem Bereich erkannt werden können. Dieses Problem ist daher systematisch und resultiert sowohl aus dem Bildmaterial als auch der Konzeption des Algorithmus.

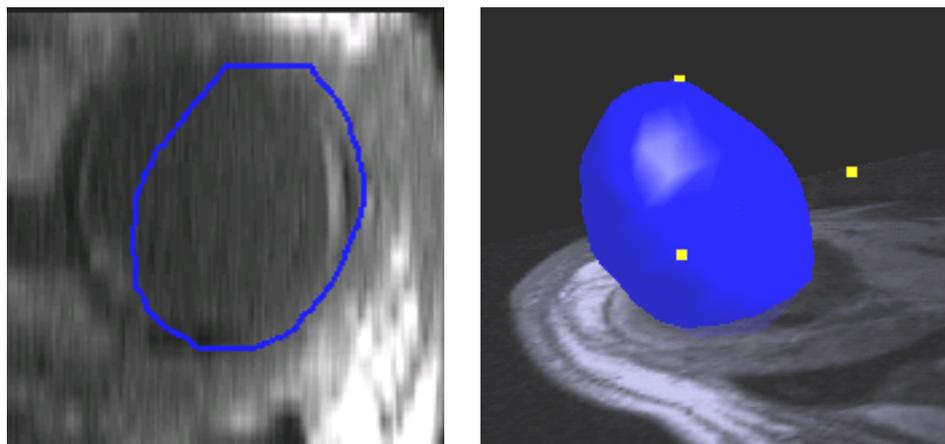


Abbildung 6.28: Typisches Problem bei der nicht-affinen Transformation

Die Abbildungen veranschaulichen eine typische Problemstellung, die bei der nicht-affinen Anpassung eines Referenzmodells an einen aktuellen Datensatz auftreten kann. Durch die Fehlpositionierung der oberen Landmarke wurde das gesamte Auge zu einem verzerrten Ei deformiert. Innerhalb der zentralen Schicht bleibt darüber hinaus immer noch eine Landmarke unberücksichtigt.

Die Ergebnisse der vorangegangenen Verfahren machen deutlich, daß die Schwierigkeit der landmarkenbasierten Registrierung darin besteht, daß einerseits eine individuelle lokale Anpassung des Referenzmodells notwendig ist, daß aber andererseits bei der Transformation die anatomische Korrektheit der globalen Organoberfläche erhalten bleiben muß. Für die landmarkenbasierte Registrierung von Datensätzen des Auges erscheinen daher Funktionen notwendig zu sein, die einen Kompromiß zwischen globalen und lokalen Transformationen gewähren. Hinzu kommt, daß die landmarkenbasierte Registrierung maßgeblich durch die Eigenschaften des Bildmaterials und der präsenten Artefakte beeinflusst wird.

Eine stabile landmarkenbasierte Registrierung sollte in seiner Konzeption sowohl den Eigenschaften des Bildmaterials als auch den anatomischen Eigenschaften entsprechen. Dieser Gedanke lag der Konzeption der proprietären landmarkenbasierten Registrierung zugrunde. Indem beispielsweise bei den Landmarken fünf und sechs lediglich die z-Komponente berücksichtigt wird, werden translatorische Fehler aufgrund von Partialvolumeneffekten vermieden. Durch die Anpassung der einzelnen Hemisphären werden Fehlpositionierungen der ersten und zweiten Landmarke korrigiert. Beide Konzepte führten dazu, daß die Qualität der Anpassung verbessert wurde, während die Anforderungen an eine korrekte Positionierung der Landmarken gesenkt werden konnten. Dadurch verringerte sich auch der Zeitaufwand, der für die Eingabe der Landmarken aufgebracht werden muß.

Nicht zuletzt durch diese konzeptionellen Veränderungen erzielt die proprietäre Anpassung für sämtliche Arten von Datensätze die besten Anpassungsergebnisse. Dabei zeigte sich, daß nicht nur die absolu-

ten Ergebnisse sich verbesserten, sondern daß auch die Standardabweichung der Registrierungsergebnisse sich verringerte. Die proprietäre landmarkenbasierte Registrierung erweist sich demnach als der stabilste Ansatz mit den besten Ergebnissen.

Für die modellbasierte Segmentierung muß ergänzend erwähnt werden, daß sich die Ergebnisqualität der Registrierung direkt auf den Rechenaufwand der nachfolgenden Schritte auswirkt. Je besser die zu erwartenden Registrierungsergebnisse sind, um so kleiner können beispielsweise die Fenster zur Berechnung der Mutual Information bei dem modellbasierten Snake gewählt werden. Dies wiederum reduziert den Rechenaufwand der modellbasierten Segmentierung drastisch. Ist hingegen die zu erwartenden Stabilität der Registrierung von Referenzmodell und aktuellem Datensatz sehr gering, muß mit deutlichen Abweichungen der groben Segmentierung von der tatsächlichen Segmentierung gerechnet werden. Der Suchraum für die externe Energie müßte demnach vergrößert werden, wodurch sich wiederum der Rechenaufwand deutlich erhöhen würde.

Die sehr kritische Beurteilung der ersten beiden Registrierungsverfahren soll nicht verdecken, daß auch diese Ansätze ihre Vorteile aufweisen und überzeugende Ergebnisse produzierten. Für die Segmentierung der ophthalmologischen Datensätze hat sich jedoch die proprietäre Anpassung nicht zuletzt aufgrund der hohen Stabilität der guten Ergebnisse als bestes Verfahren herauskristallisiert. Für zukünftige Anwendung scheint sich daraus abzuleiten, daß auch in anderen Anwendungsbereichen und bei anderen Organen eine angepaßte Registrierung einem allgemeinen Ansatz vorzuziehen ist. Die Eigenheiten von Bildmaterial und Organen wirken sich grundlegend auf die interaktiven Eingaben und die Besonderheiten der Anpassung aus. Allgemeine Registrierungsansätze können solchen Eigenheiten nicht Rechnung tragen. Auf der anderen Seite können proprietäre Anpassungsverfahren gezielt auf spezifische Problemstellungen eingehen und Fehler kompensieren. Wie das Beispiel des Auges zeigte, kann bei geeigneter Konzeption des Verfahrens Stabilität und Qualität der Ergebnisse verbessert, der Rechenaufwand gleichzeitig aber reduziert werden.

6.3.2.3 Berechnung der Referenzschichten

Das Ergebnis der landmarkenbasierten Registrierung ist eine Transformationsfunktion, die die Registrierung des Master-Referenzmodells mit dem aktuellen Datensatz definiert. Um die Segmentierung und die Bilddaten anhand dieser Transformationsfunktion auf den aktuellen Datensatz übertragen zu können, müssen noch die korrespondierenden Datenschichten und Segmentierungskonturen berechnet werden.

Die korrespondierenden Segmentierungskonturen können ohne jeden Präzisionsverlust aus der angepaßten Organoberfläche berechnet werden (vergleiche Abschnitt 4.1.7.2). Eine Analyse dieses Algorithmus ist daher nicht notwendig. Die zum aktuellen Datensatz korrespondierenden Bildschichten des Referenzdatensatzes werden durch Verwendung von Interpolationsverfahren berechnet. Dabei müssen ausgehend von den transformierten Referenzdaten die Grauwerte für jene Raumkoordinaten berechnet werden, die der Bildmatrix des aktuellen Datensatzes entsprechen. Hierzu können die in Abschnitt 3.10 beschriebenen Verfahren zur Interpolation von medizinischen Datensätzen angewandt werden.

Welches Interpolationsverfahren für die Berechnung der korrespondierenden Schichten angezeigt ist, hängt maßgeblich von den Rahmenbedingungen der individuellen Anpassung ab. Je stärker die Orientierung und Auflösung des aktuellen Datensatzes vom Master-Referenzmodell abweicht, um so stärker wirkt sich eine geringe oder andersartige Auflösung des Master-Referenzmodells aus. Je geringer die Auflösung des Master-Referenzmodells ist, desto größer sind die Anforderungen an die Qualität der Interpolation, falls die Voxelpositionen der korrespondierenden Schichten des aktuellen Datensatzes nicht mit jenen des originären Master-Referenzmodells übereinstimmen.

Da die Referenzmodelle immer wieder an unterschiedliche Datensätze angepaßt werden müssen, bietet es sich an, sämtliche Referenzmodelle vor der Nutzung zu einem isotropen Datensatz zu interpolieren, der in allen drei Raumachsen die gleiche Auflösung aufweist. Dadurch muß das Referenzmodell nur einmal vor der Einbindung in die Referenzdatenbank interpoliert werden.

Wenn das Referenzmodell vor Anwendung der modellbasierten Segmentierung bereits interpoliert wurde, erwies sich die Verwendung des nächsten Pixelwertes zur Berechnung der korrespondierenden Voxelwerte als zufriedenstellend. Eine Alternative zu dieser Methode stellt die gewichtete Mittelung der Grauwerte in der 8-Nachbarschaft dar. Bei diesem Ansatz werden die Grauwerte der nächsten acht Pixel bei der Berechnung des aktuellen Grauwertes einbezogen, wobei die einzelnen Grauwerte mit dem Kehrwert des Abstandes zur gesuchten Koordinate gewichtet werden. Die Ergebnisse der gewichteten 8-Nachbarschaft zeigten gegenüber der Nächste-Nachbar-Zuweisung jedoch keine deutlichen Verbesserungen der Bildqualität, so daß eine Anwendung dieser deutlich rechenaufwendigeren Interpolationsmethode nicht sinnvoll erscheint.

Wurde das Referenzmodell vor der Nutzung nicht interpoliert, sollten die Grauwerte durch eine oberflächenbasierte Interpolation berechnet werden, da sowohl die Nächste-Nachbar-Zuweisung als auch die Berechnung innerhalb der 8-Nachbarschaft die starken Verwischungen der linearen Interpolation aufweisen. Der Vorteil einer direkten Berechnung der korrespondierenden Schichten mittels oberflächenbasierter Interpolation der Originaldaten besteht darin, daß die Interpolation direkt auf den Originalwerten und nicht auf bereits interpolierten Werten aufsetzt. Quantifizierungseffekte, die bei der Nächste- und 8-Nachbar-Methode üblicherweise auftreten, werden dadurch vermieden. Der Nachteil dieser Vorgehensweise besteht in dem hohen Rechenaufwand, der voraussichtlich bei jeder Nutzung des nicht interpolierten Referenzmodells aufgebracht werden muß. Darüber hinaus besteht bei den Interpolationsverfahren immer die Schwierigkeit, die interpolierten Positionen für konkrete räumliche Koordinaten zu berechnen. Aus diesen Gründen wurden sämtliche Referenzmodelle vor ihrer Nutzung mittels der modellbasierten Interpolation zu einem isotropen Datensatz interpoliert. Zur Berechnung der korrespondierenden Datenschichten wurde aufgrund der zufriedenstellenden Ergebnisse die Nächste-Nachbar-Zuweisung angewandt.

Da die Interpolation des Master-Referenzmodells weitergehende Algorithmen zur Berechnung der korrespondierenden Schichten überflüssig macht, wird die Qualität der Berechnung der korrespondierenden Schichten maßgeblich durch das auf das Referenzmodell angewandte Interpolationsverfahren bestimmt. Da für die Interpolation der Referenzmodelle ausschließlich die modellbasierte Interpolation angewandt wurde, sind die entscheidenden Ergebnisse zur Berechnung der Referenzschichten bei der Beschreibung der modellbasierten Interpolationsergebnisse im Abschnitt 6.4 zu finden.

6.3.2.4 Berechnung der Ähnlichkeitsparameter

Nach der landmarkenbasierten Registrierung des Master-Referenzmodells mit dem aktuellen Datensatz können die Ähnlichkeitsparameter entlang der Kontur berechnet werden. Aus Gründen des Rechenaufwands werden für die Tests die Ähnlichkeitsparameter lediglich um die Landmarken herum berechnet. Als Ähnlichkeitsparameter wurden der mittlere Grauwert und die Standardabweichung berechnet. Für die in dieser Arbeit umgesetzten vergleichsweise kleine Referenzdatenbank ist diese Vorgehensweise als ausreichend anzusehen. Eine ausführlichere Untersuchung dieses Verfahrensbestandteils findet sich in [Kun01].

6.3.2.5 Zusammenfassung

Die durchgeführten Tests haben gezeigt, daß alle Vorgaben an die Bestimmung des Bildkontextes am besten durch die proprietäre landmarkenbasierte Registrierung erfüllt werden. Der Arbeitsaufwand für den Benutzer ist sehr gering, da lediglich die Landmarken plaziert werden müssen. Bei den Tests wurde darauf

geachtet, daß keine besondere Sorgfalt angewandt wird, um die optimalen Positionen der Landmarken zu finden. Statt dessen wurden sie eher mit einer gewissen Nachlässigkeit eingegeben, bei der Fehlpositionierungen im Bereich von bis zu 20 Pixeln bewußt in Kauf genommen wurden. Die bei den Tests berechneten Ergebnisse zeigten, daß solche Fehlpositionierungen vor allem bei der proprietären Anpassung nur einen geringen Einfluß auf die Ergebniskorrektheit hatten.

Der Rechenaufwand für die Anpassung des Referenzmodells lag in allen Fällen im Sekundenbereich. Durch das Einblenden der erläuternden HTML-Seiten ist die Benutzerfreundlichkeit so groß, daß selbst ungeübte Benutzer keine Einarbeitungszeit für die Nutzung der Software benötigten. Darüber hinaus ist auch das Plazieren der Landmarken nach einer kurzen Einleitung so leicht, daß keine besondere Übung dazu gehört, diese Eingaben in einem kurzen Zeitraum vorzunehmen.

Vor allem das grundlegende Konzept der groben Segmentierung durch Plazieren von Landmarken und Anpassen des Referenzmodells ist erfolgreich bestätigt worden. Durch die Eintragung von Landmarken konnte der Arbeitsaufwand zur Bestimmung aller notwendigen Informationen für eine grobe Segmentierung auf ein akzeptables Minimum reduziert werden, während die Korrektheit der Ergebnisse durch die Interaktion eines kompetenten Benutzers sichergestellt wird.

6.3.3 Übertragen des Vorwissens

Das Übertragen des Vorwissens setzt sich aus der Auswahl und der Anpassung des Referenzmodells zusammen. Die Zielsetzung dieser Schritte besteht darin, eine möglichst gut zum aktuellen Datensatz korrespondierende Vorlage zur Durchführung der modellbasierten Segmentierung herzustellen. Hierzu werden die folgenden Verarbeitungsschritte durchgeführt:

1. Im ersten Schritt wird das Referenzmodell anhand des Bildmodus und der im vorangegangenen Schritt berechneten Ähnlichkeitsparameter ausgewählt.
2. Das ausgewählte Referenzmodell wird an den aktuellen Datensatz angepaßt, indem die landmarkenbasierte Registrierung auf das Referenzmodell angewandt wird. Dabei werden die für die Bestimmung des Bildkontextes eingegebenen Landmarken erneut verwendet.

Neben der Beurteilung dieser beiden Algorithmen sind die Stabilität der Übertragung des Wissens und die Anwendbarkeit der Verfahren auf unterschiedliche Organe von Interesse.

6.3.3.1 Auswahl des Referenzmodells

Die Auswahl des Referenzmodells wird im ersten Schritt anhand des Bildmodus, der Schichtorientierung und der Augenseite (links oder rechts), im zweiten Schritt anhand der Darstellungseigenschaften vorgenommen. Da bei der Auswahl der möglichen Referenzmodelle im ersten Schritt Bildmodus und Schichtorientierung als rigide Kriterien angewandt werden und ein vollständiges Übereinstimmen von den Referenzmodellen mit dem aktuellen Datensatz in diesen Punkten vorausgesetzt wird, ist ein Testen dieses ersten Selektionsschrittes nicht notwendig.

Bei dem zweiten Selektionsschritt, bei dem die Ähnlichkeitsparameter verglichen werden, um das ähnlichste Referenzmodell auszuwählen, ist hingegen mit einem Spielraum für die Auswahl des geeigneten Referenzmodells zu rechnen. Dieser Spielraum resultiert aus der Unsicherheit, inwieweit die Ähnlichkeitsparameter den Ähnlichkeitsanforderungen des modellbasierten Snake entsprechen. Da es weder objektive Ähnlichkeitsparameter gibt, noch eine eindeutige Zuordnung durch verschiedene Betrachter zu erwarten sind, ist eine separate Untersuchung der Ergebnisqualität dieses Moduls nur bedingt sinnvoll. Kriterien, die für einen Menschen die Ähnlichkeit von Datensätzen gut beschreiben mögen, können für den modellbasierten Snake gegebenenfalls belanglos sein, da beispielsweise spezifische Grauwertprofile, -bereiche oder das überlagerte Rauschen von größerer Wichtigkeit sein könnten.

Um dennoch einen Eindruck von der Funktionstüchtigkeit der Auswahl des Referenzmodells anhand der Ähnlichkeitskriterien zu bekommen, wurden von drei Benutzern zu den Patientendatensätzen jeweils der Datensatz ausgewählt, der nach subjektivem Eindruck am besten mit ihm übereinstimmt. Anschließend wurden für die Datensätze und die zur Option stehenden Referenzmodelle die Ähnlichkeitsparameter mittlerer Grauwert und Standardabweichung berechnet. Im dritten Schritt wurde untersucht, inwieweit die Ähnlichkeitsparameter die Auswahl der Betrachter widerspiegeln.

Um festzustellen, welche Ähnlichkeitsparameter die Auswahl der Betrachter am besten widerspiegeln, wurden die unterschiedlichen Ähnlichkeitsparameter in verschiedenen Kombinationen getestet und anschließend mit der Auswahl verglichen. Da auch nicht davon auszugehen ist, daß die Ähnlichkeitsparameter konsistent die Auswahl nachzubilden vermögen, bietet es sich an, die einzelnen Parameter für unterschiedliche und ihre Variabilität innerhalb der Bildmodi zu untersuchen, da auf diese Weise zumindest festgestellt werden kann, wie der durch die Parameter aufgespannte Merkmalsraum konstituiert ist.

Die Tests zeigten, daß Ähnlichkeitskriterien wie mittlerer Grauwert und Standardabweichung nur bedingt geeignet sind, um die Ähnlichkeit von Datensätzen zu bewerten. Es konnten anhand dieser beiden einfachen Parameter in 64% der Fälle die vorausgesagten Datensätze selektiert werden. Ob diese Selektionsmethode eine zufriedenstellende Verfahrensweise darstellt, kann allein anhand der vorliegenden Ergebnisse nicht entschieden werden. Die nachfolgenden Tests mit dem modellbasierten Snake zeigten, daß die selektierten oder auch die durch Benutzer ausgewählten Datensätze nicht immer die besten Endergebnisse nach Anwendung des modellbasierten Snake erzielten.

Die Tests verdeutlichten, daß die Aufgabenstellung der Auswahl eines optimalen Referenzmodells offensichtlich durch eine Vielzahl weiterer, noch nicht berücksichtigter Faktoren beeinflusst wird, die im Rahmen dieser Arbeit nicht behandelt werden können. Aus diesem Grund muß dieses Thema durch zukünftige wissenschaftliche Arbeiten abgedeckt werden.

6.3.3.2 Modellanpassung

Die Modellanpassung kann in zwei Schritte unterteilt werden. Im ersten wird das Referenzmodell anhand der Landmarken durch die landmarkenbasierte Registrierung an den aktuellen Datensatz angepaßt. Im zweiten müssen beim Referenzmodell noch die zum aktuellen Datensatz korrespondierenden Schichten berechnet werden. Beide Verarbeitungsschritte stimmen vollständig mit der landmarkenbasierten Registrierung von Master-Referenzmodell und aktuellem Datensatz aus Abschnitt 6.3.2.2 überein und müssen daher an dieser Stelle nicht erneut beschrieben werden.

6.3.4 Wissensanwendung: modellbasierter Snake

Die Wissensanwendung mittels modellbasiertem Snake stellt den entscheidenden Schritt zur Berechnung einer Segmentierung dar, die im Unterschied zur Registrierung auf den Bildwerten des aktuellen Datensatzes aufsetzt. Die Präzisionsanforderungen an die Ergebnisqualität des modellbasierten Snake sind gerade in Hinsicht auf die bereits guten Ergebnisse der landmarkenbasierten Registrierung von Referenzmodell und aktuellem Datensatz hoch.

Die Schwierigkeit für den Algorithmus besteht auf der einen Seite darin, mit nicht-konkreten Angaben über eine korrekte Segmentierung eine korrespondierende Segmentierung im aktuellen Datensatz zu bestimmen. Da trotz der Auswahl eines ähnlichen Referenzmodells die Unterschiede zum aktuellen Datensatz in den meisten Fällen immer noch beträchtlich sein werden, muß der Algorithmus in der Lage sein, in einem unscharfen Lösungsbereich die korrekte oder zumindest eine anatomisch korrekte Lösung zu identifizieren. Auf der anderen Seite stellen gegebenenfalls bereits Abweichungen im Millimeterbereich von der Optimallösung einen Fehler dar, der durch Nachbearbeitung kompensiert werden muß. Ob dieser

Fehler immer durch eine Nachbearbeitung kompensiert werden kann, muß aufgrund der ebenfalls fehlerbehafteten Ergebnisse der manuellen Segmentierung jedoch in Frage gestellt werden.

Die Vorgaben an die Segmentierungsergebnisse stellen in ihrem Zusammenspiel mit der Komplexität des Bildmaterials außerordentlich hohe Anforderungen an die Stabilität und Präzision des Systems dar. Aus diesem Grund ist einer der wichtigsten Aspekte bei der Bewertung der modellbasierten Bildverarbeitung die Stabilität und Fehlertoleranz unter möglichst praxisnahen Versuchsbedingungen. Um diese praxisnahen Versuchsbedingung nachzuempfinden wurden die Phantomdatensätze und zwanzig Datensätze aus der routinemäßig durchgeführten Therapie von Augentumoren am Universitätsklinikum Benjamin Franklin getestet. Es wurden in erster Linie MRT-T1 Datensätze für die Tests segmentiert. Die Bildverarbeitungsalgorithmen wurden beispielhaft für die regelmäßigen anatomischen Strukturen auf den Bulbus angewandt. Eine Segmentierung des Tumors war nicht vorgesehen, da die Erscheinungsform okulärer Tumoren zu stark variiert.

6.3.4.1 Ergebnisse

Tabelle 6-10 zeigt die Ergebnisse, die mit dem modellbasierten Snake für die Phantomdatensätze erzielt wurden. Dabei erreicht der modellbasierten Snake sowohl mit als auch ohne Wichtungsmatrix leicht bessere Ergebnisse als die proprietäre Anpassung allein. Im Vergleich zu den manuellen Segmentierungsverfahren liegen die Übereinstimmungen der berechneten Segmentierungen mit den Mustersegmentierungen zwar unter denen der intraindividuellen Segmentierung, aber doch signifikant über denen der interindividuellen Segmentierungen.

	Mittelwert interindividuelle manuelle Segmentierung	Standard- abweichung manuelle Segmentierung	Mittelwert intraindividuelle manuelle Segmentierung	Standard- abweichung intraindividuelle manuelle Segmentierung	Mittelwert proprietäre Anpassung	Standard- abweichung proprietäre Anpassung	Mittelwert Snake mit Wichtungs- matrix	Standard- abweichung Snake mit Wichtungs- matrix	Mittelwert Snake ohne Wichtungs- matrix	Standard- abweichung Snake ohne Wichtungs- matrix
Slice 0	28,43%	35,32%	73,45%	12,54%	47,35%	15,34%	38,66%	16,25%	46,05%	13,30%
Slice 1	44,32%	12,37%	80,49%	4,87%	44,43%	13,78%	41,98%	14,56%	44,76%	12,63%
Slice 2	66,96%	8,75%	84,99%	3,21%	66,08%	13,08%	66,60%	12,87%	67,61%	12,63%
Slice 3	77,42%	7,58%	96,01%	1,19%	85,00%	5,29%	86,86%	4,82%	86,23%	5,39%
Slice 4	93,21%	8,51%	95,32%	0,94%	92,91%	2,04%	93,68%	1,87%	93,60%	1,97%
Slice 5	95,38%	4,34%	98,01%	0,68%	95,86%	1,29%	95,85%	1,27%	95,32%	1,05%
Slice 6	94,32%	3,54%	97,28%	3,51%	97,00%	0,92%	96,92%	0,68%	97,30%	0,63%
Slice 7	95,21%	2,94%	95,37%	1,45%	96,94%	0,82%	97,42%	0,60%	98,01%	0,60%
Slice 8	93,32%	1,54%	96,35%	2,04%	97,03%	1,15%	97,68%	0,65%	97,77%	0,69%
Slice 9	87,01%	2,76%	96,81%	0,69%	95,31%	2,53%	94,20%	5,55%	97,77%	0,64%
Slice 10	97,94%	5,12%	92,36%	4,07%	96,41%	1,70%	96,22%	1,51%	97,08%	0,88%
Slice 11	96,21%	0,56%	96,96%	0,31%	95,13%	1,92%	97,24%	0,80%	97,44%	0,68%
Slice 12	94,10%	5,54%	94,27%	4,51%	95,54%	2,20%	96,28%	0,98%	97,75%	0,69%
Slice 13	94,98%	4,34%	98,21%	1,56%	94,58%	2,57%	95,19%	0,65%	89,50%	9,78%
Slice 14	90,13%	9,60%	94,57%	2,65%	93,11%	2,90%	93,66%	1,18%	93,44%	1,48%
Slice 15	93,26%	5,34%	95,58%	1,18%	91,16%	3,05%	89,56%	1,88%	92,34%	1,30%
Slice 16	91,34%	6,54%	88,15%	4,23%	82,39%	5,74%	88,82%	4,23%	83,22%	4,72%
Slice 17	67,31%	7,32%	76,74%	9,78%	68,28%	8,24%	87,36%	9,51%	84,62%	9,09%
Slice 18	63,47%	18,34%	66,74%	10,65%	51,45%	8,65%	67,23%	9,89%	76,31%	5,79%

Tabelle 6-10: Ergebnisse des modellbasierten Snake für die Phantomdatensätze

Die Tabelle zeigt die Ergebnisse, die mit dem modellbasierten Snake für die Phantomdatensätze erzielt wurden. Dabei wurde der Snake sowohl mit als auch ohne Wichtungsmatrix getestet. Die Ergebnisse des Snake werden denen der interindividuellen Segmentierung, der intraindividuellen Segmentierung und der proprietären Anpassung gegenüber gestellt. Sämtliche Ergebnisse werden durch den Mittelwert und die Standardabweichung der Übereinstimmungen der berechneten Segmentierungen mit den Mustersegmentierungen beschrieben. An den Werten ist zu erkennen, daß der modellbasierte Snake leicht bessere Ergebnisse erzielte als die proprietäre Anpassung. Darüber hinaus fiel die Standardabweichung leicht ab, was eine Stabilisierung der Ergebnisqualität andeutet. Die Ergebnisse des Snake ohne Wichtungsmatrix scheinen darüber hinaus besser zu sein als jene mit Wichtungsmatrix.

Abbildung 6.29 veranschaulicht die Ergebnisse aus Tabelle 6-10, indem die Segmentierungskorrektheit über die Bildschichten hinweg eingetragen wurden. Deutlich zu sehen ist, daß die Ergebnisse des Snake sehr eng korreliert sind mit denen der interindividuellen Segmentierung und der proprietären Anpassung. Doch vor allem in den zentralen Schichten erreicht der Snake die besten Ergebnisse, die in mehreren Schichten sogar die Ergebnisse der intraindividuellen Segmentierung übersteigen.

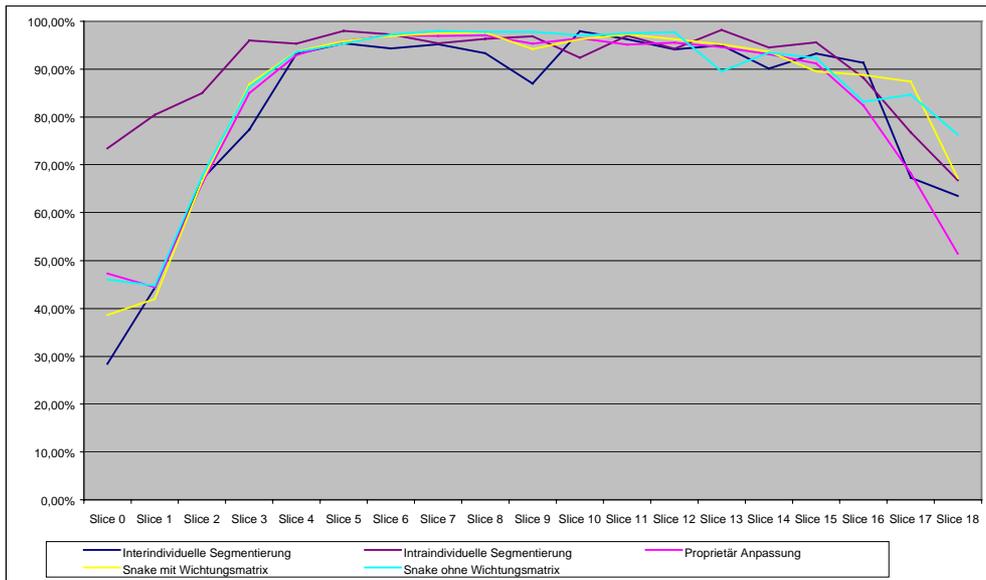


Abbildung 6.29: Ergebnisse des modellbasierten Snake für die Phantomdatensätze

Die Abbildung veranschaulicht die Ergebnisse des modellbasierten Snake für die Phantomdatensätze aus Tabelle 6-10. An den Kurven ist gut zu erkennen, daß die Ergebnisse des Snake sowohl mit als auch ohne Snake sehr eng mit den Ergebnissen der proprietären Registrierung und der interindividuellen Segmentierung korreliert sind. Nichtsdestotrotz sind leichte Verbesserungen erreicht worden, die in den zentralen Schichten sogar über den Ergebnissen der intraindividuellen Segmentierung liegen.

Die Auswertung der Schnittvolumina in Tabelle 6-11 bestätigt die Ergebnisse, die für die verschiedenen Schichten separat ermittelt wurden. Es wird zwar offensichtlich, daß die intraindividuelle Segmentierung immerhin noch drei Prozent besser ist als die der verbleibenden Ansätze. Doch die erreichten 94,5% Übereinstimmung der berechneten Segmentierung mit der Musterlösung ist als ein außerordentlich gutes Ergebnis anzusehen, da immer noch einbezogen werden muß, daß beim modellbasierten Snake Segmentierungen von einem Datensatz auf einen anderen Datensatz übertragen werden. Da bereits zwischen den Referenzmodellen nicht unbeträchtliche Unterschiede in den Segmentierungen existieren, können die Differenzen bei den berechneten Ergebnissen größtenteils aus diesen Unterschieden resultieren. Auch bei den Phantomdatensätzen gleicht vor allem in den Randschichten kein Datensatz einem anderen. Ein vollständiges Übereinstimmen der Referenzsegmentierungen war demnach auch bei den Phantomdatensätzen nicht möglich (vergleiche Abschnitt 6.3.1.1).

Schnittvolumen	Intraindividuelle Segmentierung	Interindividuelle Segmentierung	Proprietäre Anpassung	Snake mit Wichtungsmatrix	Snake ohne Wichtungsmatrix
Mittlere Übereinstimmung	97,91%	93,07%	93,98%	93,86%	94,50%
Standardabweichung	2,08%	4,62%	2,39%	1,61%	2,07%

Tabelle 6-11: Ergebnisse des modellbasierten Snake bezogen auf das Schnittvolumen

Die Tabelle zeigt die Übereinstimmungen der segmentierten Oberflächen bezogen auf die Schnittvolumina der berechneten Segmentierungen mit den Musterlösungen. Während die Überlegenheit der intraindividuellen manuellen Segmentierung immer noch mehr als drei Prozent beträgt, ist die erreichte Übereinstimmung des Snake ohne Wichtungsmatrix von über 94% als ein Erfolg zu werten.

Die Ergebnisse der Phantomdatensätze wurden durch die Tests mit den Patientendaten bestätigt. Wiederum waren die Ergebnisse des Snake mit und ohne Wichtungsmatrix eng mit den Ergebnissen der interindividuellen Segmentierungen und der proprietären Anpassung korreliert. Wieder lagen die Ergebnisse des Snake leicht über denen der beiden anderen Ansätze. Im Unterschied zu den Phantomdatensätzen scheint bei den Patientendatensätzen der Snake mit Wichtungsmatrix leicht bessere Ergebnisse zu

erzielen als der Snake ohne Wichtungsmatrix. Der Grund hierfür kann darin liegen, daß die Komplexität der Bulbus-Berandung in Patientendatensätzen deutlich größer ist und daher die Bedeutung der Grauwerte eine größere Rolle spielt.

	Mittelwert interindividuelle manuelle Segmentierung	Standard- abweichung manuelle Segmentierung	Mittelwert intraindividuelle manuelle Segmentierung	Standard- abweichung intraindividuelle manuelle Segmentierung	Mittelwert proprietäre Anpassung	Standard- abweichung proprietäre Anpassung	Mittelwert Snake mit Wichtungs- matrix	Standard- abweichung Snake mit Wichtungs- matrix	Mittelwert Snake ohne Wichtungs- matrix	Standard- abweichung Snake ohne Wichtungs- matrix
Slice 1	0,52%	0,37%			26,73%	17,42%	22,20%	15,74%	27,24%	17,05%
Slice 2	34,02%	9,60%	73,00%	0,00%	46,16%	17,97%	52,01%	17,24%	66,54%	9,50%
Slice 3	65,97%	4,59%	80,58%	4,01%	56,02%	14,50%	66,37%	13,35%	65,44%	9,41%
Slice 4	76,64%	5,81%	90,69%	2,65%	83,93%	3,82%	91,04%	2,44%	85,31%	4,66%
Slice 5	92,80%	1,14%	95,76%	0,69%	95,18%	1,11%	94,97%	0,79%	92,23%	1,56%
Slice 6	94,08%	1,35%	94,80%	1,65%	95,28%	0,95%	94,42%	1,97%	95,53%	2,87%
Slice 7	93,42%	0,88%	98,03%	0,21%	96,07%	0,78%	94,95%	0,59%	95,37%	1,84%
Slice 8	93,69%	0,69%	93,42%	2,52%	94,73%	0,98%	94,47%	1,11%	92,65%	0,84%
Slice 9	94,51%	0,79%	97,56%	0,17%	91,92%	2,45%	93,55%	1,43%	90,88%	2,86%
Slice 10	86,05%	6,37%	95,22%	1,53%	88,40%	3,22%	89,01%	2,83%	85,28%	7,31%
Slice 11	91,44%	1,39%	95,58%	0,89%	85,03%	2,96%	84,85%	3,02%	88,04%	1,53%
Slice 12	63,14%	7,76%	84,75%	3,14%	75,87%	4,22%	68,27%	14,74%	70,55%	14,72%
Slice 13	1,60%	0,80%	62,74%	10,83%	51,50%	14,00%	79,01%	3,81%	74,91%	7,33%
Slice 14					20,75%	0,00%	42,25%	0,00%	65,31%	0,00%

Tabelle 6-12: Ergebnisse des modellbasierten Snake für die Patientendaten

Die Tabelle zeigt die Ergebnisse des modellbasierten Snake mit und ohne Wichtungsmatrix für die Patientendatensätze und stellt sie den Ergebnissen der manuellen Segmentierungen und der proprietären Anpassung gegenüber. Wie bei den Phantomdatensätzen liegen die Ergebnisse aller Ansätze sehr eng beieinander.

Abbildung 6.30 stellt wiederum die Ergebnisse grafisch dar. Im Vergleich zu den Phantomdatensätzen können zwei Tendenzen festgestellt werden. Während der Snake ohne Wichtungsmatrix in den Randschichten anscheinend bessere Ergebnisse erzielt als der Snake mit Wichtungsmatrix, ist es in den zentralen Schichten des Auges eher umgekehrt.

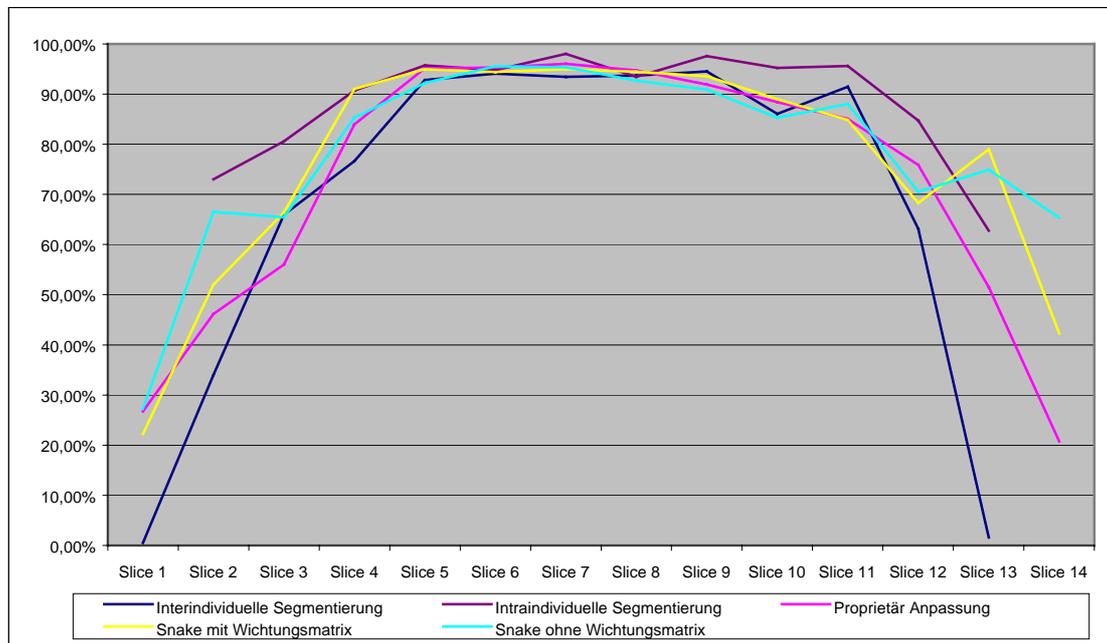


Abbildung 6.30: Darstellung der Ergebnisse des modellbasierten Snake für Patientendaten

Die Abbildung stellt die Ergebnisse aus Tabelle 6-12 graphisch dar. Wie bei den Phantomdatensätzen liegen die Ergebnisse des Snake mit und ohne Wichtungsmatrix über jenen der interindividuellen Segmentierung und der proprietären Anpassung. Darüber hinaus scheinen die Ergebnisse des modellbasierten Snake ohne Wichtungsmatrix in den Randschichten des Auges besser zu sein, als die des Snake mit Wichtungsmatrix. In den zentralen Schichten hingegen scheint der Snake mit Wichtungsmatrix leicht bessere Ergebnisse zu erzielen. Der Unterschied zwischen dem Snake und der intraindividuellen Segmentierung ist bei den Patientendatensätzen deutlich geringer als bei den Phantomdatensätzen.

Der Grund für diese Tendenz könnte darin liegen, daß in den Randschichten die Partialvolumeneffekte die Grauwerteigenschaften des Bulbus nicht-deterministisch verändern und so eine Ausnutzung von Grauwertähnlichkeiten nicht sinnvoll ist. In diesen Schichten sind textuelle Bildelemente gegebenenfalls von größerer Bedeutung. Andererseits sind die Grauwerteigenschaften in den zentralen Schichten vergleichsweise einheitlich, so daß eine Nutzung der Grauwertkonstellationen zur Verbesserungen der Segmentierung genutzt werden kann.

Um den Einfluß unterschiedlicher Segmentierungen in den Referenzmodellen abschätzen zu können, wurde außerdem der Snake getestet, indem zur Segmentierung der Datensätze die Mustersegmentierungen als Referenzmodell angewandt wurden. Diese Tests wurden sowohl für Patientendatensätze als auch für die Phantomdatensätze durchgeführt. Tabelle 6-13 beschreibt die Ergebnisse, die mit diesen Tests erzielt wurden. Eine Überraschung stellt dabei das Ergebnis dar, daß der modellbasierte Snake teilweise sogar bessere Ergebnisse erzielte als die intraindividuelle Segmentierung. Auch wenn die Musterlösungen auf den eigenen Datensatz angewandt wurden, war dieses Ergebnis nicht selbstverständlich. Durch das Platzieren der Landmarken im aktuellen Datensatz stellt die Anpassung der Konturen durchaus ein schwieriges Problem dar. So boten die Ergebnisse der proprietären Anpassung durchaus einen nicht unbeträchtlichen Spielraum zur Verbesserung der Segmentierung, den der Snake sowohl bei den Phantomdatensätzen als auch bei den Patientendaten nutzte.

	Mittelwert interindividuelle manuelle Segmentierung	Standard- abweichung manuelle Segmentierung	Mittelwert intraindividuelle manuelle Segmentierung	Standard- abweichung intraindividuelle manuelle Segmentierung	Mittelwert proprietäre Anpassung	Standard- abweichung proprietäre Anpassung	Mittelwert Snake bei Patientendaten	Standard- abweichung Snake bei Patientendaten	Mittelwert Snake bei Phantom- datensätzen	Standard- abweichung Snake bei Phantom- datensätzen
Slice 1	0,52%	0,37%			26,73%	17,42%	71,81%	0,00%	83,64%	3,10%
Slice 2	34,02%	9,60%	73,00%	0,00%	46,16%	17,97%	73,98%	1,08%	85,81%	3,80%
Slice 3	65,97%	4,59%	80,58%	4,01%	56,02%	14,50%	82,16%	4,22%	90,43%	2,99%
Slice 4	76,64%	5,81%	90,69%	2,65%	83,93%	3,82%	86,41%	1,90%	88,99%	2,66%
Slice 5	92,80%	1,14%	95,76%	0,69%	95,18%	1,11%	84,94%	4,94%	93,63%	1,85%
Slice 6	94,08%	1,35%	94,80%	1,65%	95,28%	0,95%	96,56%	0,94%	93,53%	1,12%
Slice 7	93,42%	0,88%	98,03%	0,21%	96,07%	0,78%	96,11%	1,01%	94,57%	2,19%
Slice 8	93,69%	0,69%	93,42%	2,52%	94,73%	0,98%	94,89%	0,92%	94,91%	1,73%
Slice 9	94,51%	0,79%	97,56%	0,17%	91,92%	2,45%	96,15%	0,42%	97,16%	1,14%
Slice 10	86,05%	6,37%	95,22%	1,53%	88,40%	3,22%	94,47%	1,46%	96,39%	1,41%
Slice 11	91,44%	1,39%	95,58%	0,89%	85,03%	2,96%	83,68%	4,18%	97,02%	1,19%
Slice 12	63,14%	7,76%	84,75%	3,14%	75,87%	4,22%	78,93%	6,14%	95,71%	1,56%
Slice 13	1,60%	0,80%	62,74%	10,83%	51,50%	14,00%	89,67%	3,17%	96,90%	1,17%
Slice 14	0,00%	0,00%			20,75%	0,00%	57,51%	0,00%	94,96%	0,88%

Tabelle 6-13: Ergebnisse des modellbasierten Snake bei Tests mit identischem Referenz- und aktuellem Datensatz

In der Tabelle werden die Ergebnisse zusammengefaßt, die dadurch erzielt wurden, daß auf die Testdatensätze die Musterlösungen als Referenzdatensätze angewandt wurden. Dadurch konnte bei den Tests der Einfluß von unterschiedlichen Segmentierungsstrategien zwischen den verschiedenen Datensätzen eliminiert werden. Es wurden wiederum die Ergebnisse des Snake mit Wichtigkeitsmatrix denen der manuellen Segmentierung und der proprietären Anpassung gegenüber gestellt.

Noch deutlicher werden die guten Ergebnisse des Snake in der grafischen Darstellung der Segmentierungsübereinstimmungen in Abbildung 6.31. Die Segmentierungsergebnisse für die Phantomdaten liegen deutlich über denen der intraindividuellen Segmentierung. Bei den Patientendaten liegen in nahezu 50% der Schichten die Übereinstimmungen über denen der intraindividuellen Segmentierung. Diese Ergebnisse deuten an, daß mit dem modellbasierten Snake unter günstigen Bedingungen zumindest theoretisch tatsächlich bessere Ergebnisse erzielt werden können, als mit einer manuellen Segmentierung. Erstaunlich ist, daß vor allem in den Randschichten bessere Ergebnisse erzielt wurden. Der Grund hierfür könnte darin liegen, daß Registrierungsfunktionen charakteristische Texturen qualitativ präziser auszuwerten vermögen als das eher qualitativ beurteilende menschliche Auge.

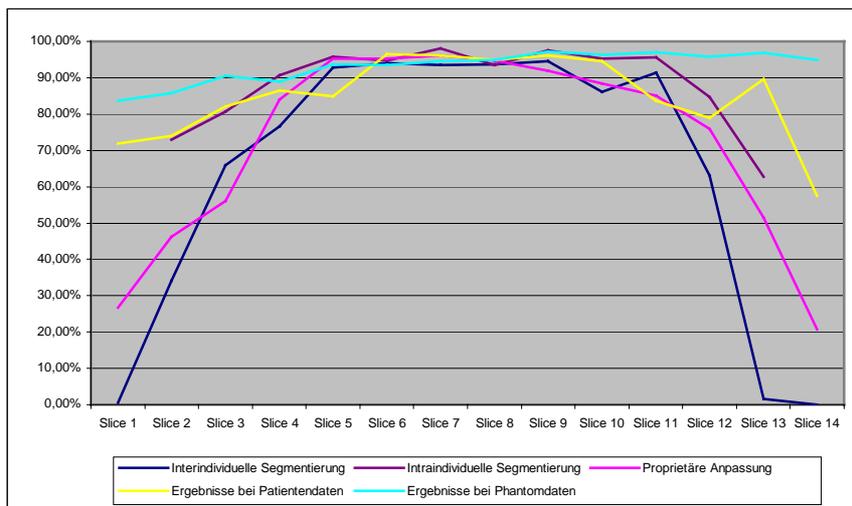


Abbildung 6.31: Ergebnisse mit identischem Referenz- und aktuellem Datensatz

Die Abbildung stellt die Ergebnisse aus Tabelle 6-13 grafisch dar. Es ist deutlich zu sehen, daß vor allem die Ergebnisse für die Phantomdaten deutlich über jenen der intraindividuellen liegen. Aber auch bei den Patientendaten übertrifft der Snake in knapp 50% der Schichten die Ergebnisse der intraindividuellen Segmentierung. Diese Ergebnisse scheinen anzudeuten, daß unter günstigen Bedingungen mit dem modellbasierten Snake sogar konsistentere Ergebnisse erzielt werden können als mit der manuellen Segmentierung.

Neben der Gesamtperformance des modellbasierten Snake ist auch die Funktionsweise der einzelnen Module von Bedeutung. Daher werden in den nachfolgenden Abschnitten Beobachtungen, die im Laufe der Tests gemacht wurden, aufgeführt.

6.3.4.2 Interne Kraft

Die Aufgabenstellung der internen Kraft besteht darin, die durch das angepaßte Referenzmodell vorgegebene anatomisch korrekte Form so weit wie möglich einzuhalten. Bei der Einschätzung der internen Kraft ist zu untersuchen, inwieweit die Umsetzung der internen Kraft die an sie gestellte Aufgabenstellung erfüllt. Dabei spielen drei Faktoren eine wichtige Rolle:

1. Wie wirken sich die Parameter Winkelkraft und Distanzkraft auf die Wirkung der internen Kraft aus?
2. Wie gut stellt die interne Kraft die ursprüngliche Form der Referenzsegmentierung wieder her?
3. Wie viele Iterationen benötigt die interne Kraft zur Wiederherstellung der ursprünglichen Form?
4. Wie wirken sich Variationen in der Parameterichtung auf die Performance der internen Kraft aus?

Die ebenfalls naheliegende Frage nach der optimalen Parametereinstellung für die interne Kraft kann erst im Zusammenspiel mit der externen Kraft beantwortet werden. Denn erst wenn die Gesamtkontur durch die externe Kraft an unterschiedliche Skalierungen, Positionen oder Formen angepaßt wird, kann beurteilt werden, welche Freiheitsgrade und Einschränkungen für die Anpassung an das Bildmaterial sinnvoll sind.

Um die internen Kräfte zu testen, wurden verschiedene Ausgangskonturen zunächst stark verzerrt. Anschließend wurde der Snake iterativ angewandt, um die Kontur durch die Kräfte wieder herzustellen.

Bei der Beurteilung des Ergebnisses wurde einerseits die Bewegung der Kontur, die Anzahl der Iterationen bis ein stabiler Zustand erreicht wurde und die Form der Kontur im stabilen Zustand beurteilt. Um die Wirkungsweise der einzelnen Komponenten beurteilen zu können, wurden diese Tests mit unterschiedlichen Gewichtungen von Winkel- und Distanzkraft durchgeführt. Zur Beurteilung der individuellen Eigenschaften von Winkel- und Distanzkraft werden die Einzelkräfte der internen Kraft zunächst einzeln angewandt, indem die andere Kraft vollständig ausgeschaltet wird.

Bei Tests, bei denen lediglich die Winkelkraft zur Anwendung kam, wurde die globale Form der Kontur nach zirka zwei bis vierhundert Iterationen vollständig wieder hergestellt. Zur Durchführung dieser Iterationen wurden zwischen drei und neun Sekunden benötigt. Je nach Ausmaß und Art der Veränderungen der Kontur vor der Anwendung des Snake kam es zu geringfügigen Abweichungen der Kontur in der Skalierung und der Orientierung der Ergebniskontur. Das bedeutet, daß die Kontur nach Anwendung der internen Energie kleiner oder größer als vorher war, oder daß die Kontur leicht rotiert wurde. Beide Effekte sind durchaus erwartet und bestätigten die von den Bilddaten unabhängige Erfüllung der Aufgabenstellung an die Winkelkraft.

Erwartungsgemäß bewirkte die ausschließliche Anwendung der Distanzkraft ein völlig anderes Verhalten. Wird die Winkelkraft ausgeschaltet, werden innerhalb weniger Iterationen die Distanzen zwischen den Konturen wieder hergestellt. Die Form der Kontur wich bei den Tests jedoch deutlich von der ursprünglichen ab und war maßgeblich von der Veränderung der Kontur vor der Anwendung des Snake geprägt.

Werden Distanz- und Winkelkraft gleichzeitig angewandt, kam es bei sämtlichen Tests nach drei bis fünfhundert Iterationen zu einer vollständigen Wiederherstellung der ursprünglichen Konturform. Die Skalierung der Kontur wurde durch den Einsatz der Distanzkraft, die Form durch die Winkelkraft wieder hergestellt. Leichte Rotationen der Ergebniskonturen waren möglich. Die Berechnungsdauer betrug zwischen fünf und zehn Sekunden.

6.3.4.3 Externe Kraft

Bei der Bewertung der externen Kraft müssen mehrere Einzelfaktoren berücksichtigt werden. Die wichtigste Frage ist in diesem Zusammenhang, inwieweit die Mutual Information in der Lage ist, die zum Referenzdatensatz korrespondierenden Grauwertkonstellationen im aktuellen Datensatz zu lokalisieren. Neben der Anwendung der Mutual Information in ihrer ursprünglichen Definition wurden bei den Tests auch die Eigenschaften der um die Wichtungsmatrix erweiterten Mutual Information untersucht. Ziel dieser Erweiterung war die Stabilisierung der Ergebnisse der Mutual Information bei der Detektion der korrespondierenden Punkte im aktuellen Datensatz. Durch den direkten Vergleich der Segmentierungsergebnisse des modellbasierten Snake sowohl ohne als auch mit Einbeziehung der Wichtungsmatrix sollte untersucht werden, ob dieses Ziel erreicht wurde.

Neben der Funktionsweise der Mutual Information spielt auch die Auswahl der Konturpunkte, für die die externe Energie berechnet wird, eine Rolle. Durch die Vorgabe, für welche Punkte die externe Energie berechnet und angewandt wird, wird das Zusammenspiel oder die Konkurrenz der internen und externen Kräfte bestimmt. Sowohl die Funktionsweise der Mutual Information als auch die getesteten Strategien bei der Auswahl der externen Energiepunkte werden in den folgenden beiden Abschnitten diskutiert.

6.3.4.3.1 Mutual Information

Die in Abschnitt 6.3.4.1 beschriebenen Ergebnisse des modellbasierten Snake mit und ohne Wichtungsmatrix belegten bereits die erfolgreiche Anwendung der Mutual Information als Registrierungsfunktion.

Die Tests lassen jedoch keine Rückschlüsse zu, inwieweit die guten Ergebnisse auf die Funktionsweise der Mutual Information oder auf die Gesamtkonzeption des modellbasierten Snake zurückzuführen sind. Um eine Bewertung der Mutual Information als geeignete Registrierungsfunktion vornehmen zu können, müßte untersucht werden, wie gut die korrespondierenden Positionen im aktuellen Datensatz unabhängig durch das Einwirken der internen Energie identifiziert wurden. Von Bedeutung wäre in diesem Zusammenhang einerseits, wie ausgeprägt die Maxima an den korrespondierenden Stellen sind im Verhältnis zum umgebenden Bereich und ob gegebenenfalls auch ungewollte Maxima ausgebildet werden. Eine detaillierte Untersuchung dieser spezifischen Eigenschaften der Mutual Information würde unter anderem komparative Analysen mit alternativen Registrierungsfunktionen wie beispielsweise der Kreuzkorrelation und dem Cepstrum erforderlich machen. Die in der aktuellen Literatur beschriebenen Ergebnisse bestätigen zwar die positiven Eigenschaften der Mutual Information. Ob diese Eigenschaften auch den Anforderungen des modellbasierten Snake entsprechen, läßt sich nur schwer abschätzen. Solche Untersuchungen liegen aufgrund des beträchtlichen Forschungsaufwandes außerhalb des Rahmens dieser Arbeit und müssen durch weiterführende Projekte untersucht werden.

Eine der wichtigsten Eigenschaften der Mutual Information besteht in der Berechnungsformel, bei der nicht wie bei anderen Registrierungsfunktionen die Grauwerte in ein direktes Verhältnis zueinander gesetzt werden, sondern lediglich die Häufigkeiten von den auftretenden Grauwertkombinationen bei den übereinander liegenden Bilddaten. Dadurch werden die Registrierungsergebnisse nicht durch die Grauwerte, sondern durch ihre Auftrittswahrscheinlichkeiten und ihre Anordnung zueinander beeinflusst. Diese Eigenschaft stellt die Stärke der Mutual Information dar, wenn Bilddaten mit divergierenden Grauwerteigenschaften miteinander registriert werden sollen. Da die Grauwerteigenschaften des ophthalmologischen Bildmaterials innerhalb eines gegebenen Toleranzbereichs durchaus einheitlich sind, erschien diese Grauwertindifferenz der Mutual Information nicht in jedem Fall positiv zu sein. Aus diesem Grund wurde die Verwendung der Wichtungsmatrix erwogen, mit der die Grauwertkonstellationen kontextabhängig einbezogen werden können.

Bei den Tests zeigte sich, daß weder die reine Mutual Information, noch die um die Wichtungsmatrix erweiterte Definition der Mutual Information entscheidende Vorteile oder Nachteile aufweisen. Es gab leichte Unterschiede zwischen den Ergebnissen der Phantom- und Patientendatensätze und bei den Tests, in denen die Musterlösungen auf die eigenen Datensätze angewandt wurden. Letztendlich muß jedoch festgestellt werden, daß die geringen Unterschiede in der Ergebnisqualität keine endgültigen Präferenzen erlauben.

Gerade bei einer geringen Anzahl von Referenzmodellen in der Referenzdatenbank erscheint die Verwendung einer Wichtungsmatrix eher kontraproduktiv zu sein. Durch die Einbeziehung von Grauwertkonstellationen beeinflussen unterschiedliche Grauwerteigenschaften der Bilddaten die Registrierungsergebnisse. Wenn wenige Referenzmodelle in der Referenzdatenbank vorliegen, kann es durchaus vorkommen, daß Datensätze aufeinander angewandt werden müssen, die sich stark in den Grauwerteigenschaften unterscheiden. Bei einer Anwendung der Wichtungsfunktion beim modellbasierten Snake können diese Grauwertunterschiede die Ergebnisse der externen Energie verfälschen. Dieser Effekt scheint sich bereits in den aktuellen Tests in den Randschichten abzuzeichnen. Durch die Partialvolumeneffekte wurden die Grauwerte in den Randschichten des Auges nicht-deterministisch verzerrt. Es existierten somit nicht die erwarteten regulären Grauwertkonstellationen. Die Ergebnisse des modellbasierten Snake mit Wichtungsfunktion waren dementsprechend für die Patientendaten in diesem Bereich deutlich schlechter als ohne Wichtungsfunktion.

Dementsprechend scheint die Nutzung von Wichtungsfunktionen bei der Berechnung der externen Energie nur dann sinnvoll zu sein, wenn einheitliche Grauwerteigenschaften des Bildmaterials sichergestellt werden können. Es wäre irrelevant, ob diese Einheitlichkeit langfristig durch eine Normierung der

Datensätze oder durch eine große Referenzdatenbank erreicht wird. Im Rahmen der durchgeführten Tests konnte diese Einheitlichkeit nicht sichergestellt werden, so daß die Anwendung der Wichtungsfunktion nicht empfohlen wird.

6.3.4.3.2 Automatische Auswahl der Punkte für die externe Energie

Die automatische Auswahl jener Punkte, für die die externe Energie berechnet wird, spielt für die Performance des modellbasierten Snake eine durchaus wichtige Rolle. Der anfangs vorgesehene Ansatz, pro Schicht lediglich vier bis sechs Punkte für die Anwendung der externen Energie heranzuziehen, erwies sich nicht als erfolgreich, da in den Zwischenräumen zwischen den selektierten Punkten oft beträchtliche Segmentierungsfehler auftraten. Die gegensätzliche Strategie, die externe Energie für sämtliche Konturpunkte zu berechnen, erwies sich ebenfalls als ungünstig, weil in diesem Fall externe und interne Energie in direkter Konkurrenz zueinander standen. Als Folge wurde die Kontur vorwiegend durch jene Kraft beeinflusst, die aufgrund der aktuellen Rahmenbedingungen stärker war. Die Kooperation der beiden Kräfte wurde in eine Konkurrenz gewandelt. Dadurch nahm die Bedeutung der Wichtung beider Energien stark zu. Die Stabilität des modellbasierten Snake schien aber eher nachzulassen, da eine individuelle Anpassung der Wichtung deutlich stärker die Ergebnisqualität beeinflusste als zuvor.

Als einen sinnvollen und stabilen Kompromiß stellte sich der Ansatz heraus, jeden zweiten Konturpunkt für die Anwendung der externen Energie anzuwenden. Damit bekam einerseits die interne Energie einen von der externen Energie unbeeinflussten Spielraum, um die globale Organkontur zu erhalten. Andererseits wurden mangelhafte Segmentierungen aufgrund fehlender Berücksichtigung des Bildmaterials vermieden.

Trotz dieser Weiterentwicklung der automatischen Auswahl der Punkte für die externe Energie ist das Selektionskonzept immer noch als suboptimal zu betrachten. Bei der Auswahl jedes zweiten Konturpunktes werden die Distanzen zwischen den externen Energiepunkten in der angepaßten Segmentierung einerseits durch die manuelle Segmentierung im Referenzmodell, andererseits durch die Anpassung des Referenzmodells an den aktuellen Datensatz vorgegeben. Beide Einflüsse können dazu führen, daß die Abstände zwischen den Konturpunkten sehr unregelmäßig sind. Bei der pauschalen Auswahlstrategie jedes zweiten Konturpunktes für die Berechnung der externen Energie können die unterschiedlichen Punktabstände bei der angepaßten Segmentierung dazu führen, daß die externen Energiepunkte stellenweise sehr nah beieinander, in anderen Bereichen wiederum sehr weit auseinander liegen. Eine solch ungleichmäßige Verteilung der Punkte kann sich im ungünstigen Fall negativ auf die Ergebnisse des modellbasierten Snake auswirken.

Der zunächst umgesetzte Ansatz, die Abstände zwischen den Punkten zu vereinheitlichen, indem die Punktabstände durch eine angepaßte Interpolation vereinheitlicht wird, hat sich nicht als vorteilhaft erwiesen, da dadurch die manuelle Nachbearbeitung der endgültigen Segmentierungsergebnisse deutlich erschwert wurde. Die Anzahl der Punkte hat sich in der Regel so stark erhöht, daß sich die Zeiten für die Nachbearbeitung deutlich erhöhten. Darüber hinaus plazierten die Testpersonen bei der manuellen Segmentierung der Referenzmodelle die Konturpunkte bei anatomisch markanten Landmarken. Diese anatomischen Landmarken zeichnen sich oft durch charakteristische Textureigenschaften aus. Die Nutzung solcher charakteristischen Texturen verbessert wiederum die Möglichkeiten der externen Energie zur korrekten Lokalisation korrespondierender Punkte im aktuellen Datensatz. Bei einer automatischen Neuberechnung der Konturpunkte wäre die Nutzung solcher markanten Bildelemente dem Zufall überlassen. Es wäre gegebenenfalls mit einer Verschlechterung der Ergebnisse zu rechnen. Hinzu kommt, daß Testpersonen in Bereichen stark verwinklelter Konturen mehr Punkte setzen und in Bereichen mit sehr kontinuierlichen Organkanten die Anzahl der Konturpunkte reduzierten. Die Beibehaltung solcher Punkteverteilungen scheint auch bei der Anwendung des modellbasierten Snake sinnvoll zu sein.

Es wäre daher empfehlenswert, bei weiteren Forschungsarbeiten im Bereich der modellbasierten Segmentierung diesen Aspekt der automatischen Auswahl der externen Energiepunkte weitergehend zu untersuchen und alternative Herangehensweisen zu entwickeln. Für die Umsetzung des aktuellen Systems erwies sich das direkte Übernehmen der Konturpunkte aus dem Referenzmodell und die Auswahl jedes zweiten Punktes für die externe Energie durchaus als stabil und wurde dementsprechend beibehalten.

6.3.4.4 Wichtung von interner und externer Kraft

Ein schwerwiegender Nachteil des herkömmlichen Snake-Ansatzes besteht in der hohen Empfindlichkeit der Segmentierungsergebnisse gegenüber den Einstellungen der internen und externen Energie und der Wichtung beider Energien zueinander. Die Variabilität der Ergebnisse des herkömmlichen Snake ist so groß, daß die Parameter für jeden Datensatz, oft sogar für jede Bildschicht individuell angepaßt werden müssen, um Segmentierungen mit akzeptabler Korrektheit zu erzielen. Dieses Anpassen der Parameter erweist sich dabei oft als so kompliziert, daß der Anpassungsaufwand den Arbeitsaufwand übersteigt, der für eine manuelle Segmentierung erforderlich wäre.

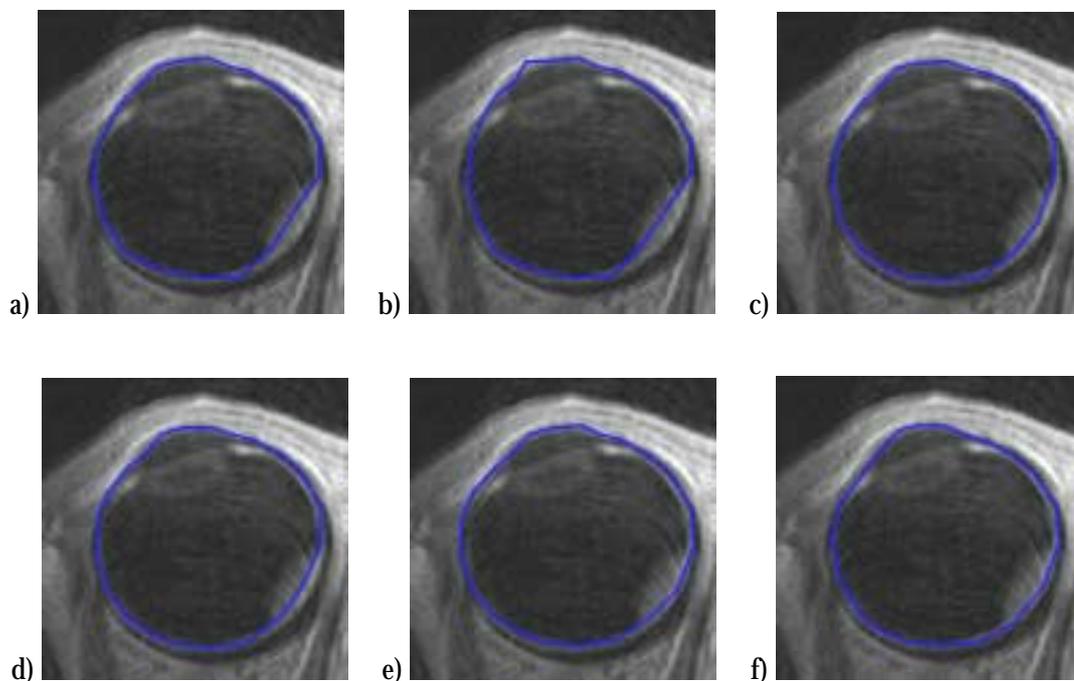


Abbildung 6.32: Ergebnisse des modellbasierten Snake bei unterschiedlichen Parametereinstellungen

Die Abbildungen zeigen die Ergebnisse, die bei verschiedenen Einstellungen für die Winkel-, Distanz- und externe Kraft berechnet wurden. Links oben (a) ist das Ergebnis zu sehen, das bei ausschließlicher Anwendung der externen Kraft berechnet wurde. Rechts daneben (b) ist die Ergebniskontur bei Anwendung der externen Kraft und der Distanzkraft zu sehen. Die Segmentierungskontur rechts oben (c) wurde durch Anwendung der externen Kraft und der Winkelkraft berechnet. Bei der Abbildung links unten (d) wurden Winkel-, Distanz- und externe Kraft gleichgewichtet angewandt. In der Mitte unten (e) wurde die Distanzkraft mit dem Faktor 0,4, die verbleibenden Kräfte mit eins gewichtet. Bei der rechten unteren Abbildung (f) wurde die Distanzkraft mit 0,5, die Winkelkraft mit 0,9 und die externe Kraft mit eins gewichtet.

Eine wichtige Zielsetzung des modellbasierten Snake war es daher, einen Algorithmus zu entwickeln, der so stabil gegenüber der Variabilität des Bildmaterials und den Parametereinstellungen ist, daß keine Anpassung für die verschiedenen Datensätze notwendig ist. Daß dieses Ziel erreicht wurde, haben die bereits oben beschriebenen Ergebnisse belegt. Obwohl die Parameter für keinen der Datensätze individuell angepaßt wurden, konnten Ergebnisse mit überzeugender Korrektheit erzielt werden. Nichtsdestotrotz

besteht auch bei dem modellbasierten Snake die Möglichkeit, die Performance anhand von Einstellungen an der internen und externen Energie zu modifizieren.

Um die optimalen Einstellungen des modellbasierten Snake zu ermitteln, wurden an mehreren Datensätzen verschiedene Einstellungen getestet. Abbildung 6.32 zeigt beispielhaft die Ergebnisse für einen T1-MRT-Datensatz, die mit verschiedenen Parametereinstellungen erzielt wurden. Insgesamt konnte festgestellt werden, daß die externe Kraft die Konturpunkte nahezu immer zu den korrekten Positionen zog. Eine hohe Wichtung der externen Energie erschien daher bei allen Tests sinnvoll zu sein. Bei einem ausschließlichen Anwenden der externen Energie wird jedoch die Form des Bulbus nicht immer eingehalten. Doch auch bei Anwendung von externer Kraft und Distanzkraft kam es zu deutlichen Abweichungen in der Bulbus-Form, da einerseits die Gesamtform des Bulbus durch die externe Energie verändert wurde. Andererseits versuchte die Distanzkraft die Abstände zwischen den Konturpunkten wieder herzustellen. Dadurch wurden Punkte, für die keine externe Energie berechnet wurde, an nicht sinnvolle Positionen verschoben (siehe Abbildung 6.32, obere Mitte).

Diese Verzerrungen der Bulbus-Form treten nicht auf, wenn statt dessen auch die Winkelkraft angewandt wird. Die besten Ergebnisse wurden bei den durchgeführten Tests erzielt, wenn die Distanzkraft mit 0,5, die Winkelkraft mit 0,9 und die externe Kraft mit eins gewichtet wurde.

Diese starke Wichtung der externen Kraft war einerseits sinnvoll, da die externe Kraft in den meisten Fällen die korrespondierenden Punkte tatsächlich identifizieren konnte. War eine eindeutige Lokalisation des korrespondierenden Punktes aufgrund der Bildeigenschaften nicht möglich, führte die starke Wichtung der externen Energie dazu, daß nicht unbeträchtliche Teile der Segmentierungskontur gegebenenfalls an eine falsche Position gezogen wurden. Die Frage, ob die externe Kraft zur Vermeidung von solchen Problemen weniger stark gewichtet werden sollte, kann nicht abschließend beantwortet werden. Denn eine stärkere Wichtung der internen Kräfte könnte dazu führen, daß die Segmentierungskontur stärker nach den Vorgaben des Referenzmodells ausgestaltet wird. Solche Veränderungen werden voraussichtlich nur im Ausnahmefall dem Bildmaterial sinnvoll Rechnung tragen.

Eine wirklich optimale Einstellung der Wichtungsparameter ist nicht möglich. Zu stark ist der Einfluß des Bildmaterials, der Patientenanatomie und der eingegebenen Landmarken auf die resultierenden Segmentierungsqualität. Eine optimale Wichtung der Energie könnte daher gegebenenfalls auch beim modellbasierten Snake individuell vom jeweiligen Datensatz abhängen. Im Unterschied zu dem herkömmlichen Snake sind jedoch die Schwankungen in der Segmentierungsqualität so gering, daß eine solche Anpassung an das zu bearbeitende Bildmaterial nicht notwendig ist.

Als optimale Einstellung wurde eine Wichtung der Distanzkraft mit 0,5, der Winkelkraft mit 0,9 und der externen Kraft mit eins übernommen. Die guten Testergebnisse des modellbasierten Snake haben gezeigt, daß diese Einstellungen für die unterschiedlichen Testdatensätze uneingeschränkt sinnvoll waren und eine anderweitige Anpassung zu keinem Zeitpunkt angezeigt schien. Darüber hinaus belegen die guten Ergebnisse, daß das Angehen von systematischen Problemen bei der Konzeption des Snake-Algorithmus deutlich wichtiger ist als die individuelle Optimierung von Parametern, die letztendlich von sekundärer Bedeutung sind.

6.3.4.5 Zeitaufwände

Neben der Segmentierungsqualität spielt auch der Zeitaufwand, der für die Erzielung dieser Ergebnisse notwendig ist, eine wichtige Rolle, da er die Einbindung der Bildverarbeitung in die klinische Routine maßgeblich beeinflusst. Tabelle 6-14 zeigt beispielhaft die Zeitaufwände für die modellbasierte Segmentierung des Bulbus in fünf T1-MRT-Datensätzen. Der für den Benutzer relevante Interaktionsaufwand setzt sich aus dem Laden des Datensatzes und des Referenzdatensatzes und dem Eingeben der Landmarken

zusammen. Bei allen Tests lag der hierfür erforderliche Zeitaufwand deutlich unter einer Minute. Das sind ungefähr 10 Prozent des Aufwandes, der für die manuelle Segmentierung der gleichen Daten aufgebracht werden müßte. Die verbleibenden Rechenschritte wie das Anpassen des Referenzmodells und die Anwendung des modellbasierten Snake werden automatisch durchgeführt und stellen keinen Arbeitsaufwand für die Arbeitskraft dar.

Datensatz	aktuellen Datensatz laden	Referenzmodell laden	Landmarken setzen	Referenzmodell anpassen	Modellbasierter Snakes	Gesamtaufwand
MRT-Datensatz 1	00:00:05	00:00:13	00:00:39	00:00:16	00:08:02	0:09:15
MRT-Datensatz 2	00:00:04	00:00:10	00:00:34	00:00:18	00:06:52	0:07:58
MRT-Datensatz 3	00:00:04	00:00:11	00:00:35	00:00:17	00:07:21	0:08:28
MRT-Datensatz 4	00:00:03	00:00:21	00:00:37	00:00:16	00:05:59	0:07:16
MRT-Datensatz 5	00:00:05	00:00:12	00:00:44	00:00:16	00:07:45	0:09:02

Tabelle 6-14: Zeitaufwände für die Durchführung der modellbasierten Segmentierungsschritte

Die Tabelle listet beispielhaft für fünf verschiedene T1-MRT-Datensätze die Zeitaufwände auf, die für die Durchführung der verschiedenen Verarbeitungsschritte erforderlich waren.

Neben dem interaktiven Arbeitsaufwand ist auch der Gesamtaufwand einschließlich der automatischen Berechnungen überraschend kurz, der nur ungefähr 80% des Zeitaufwandes der manuellen Segmentierung ausmacht (vergleiche Tabelle 6-14). Selbst wenn der Nutzer die Rechenzeiten nicht für eine andere Tätigkeit nutzen sollte, würde die Anwendung des modellbasierten Snake einen Zeitgewinn bedeuten.

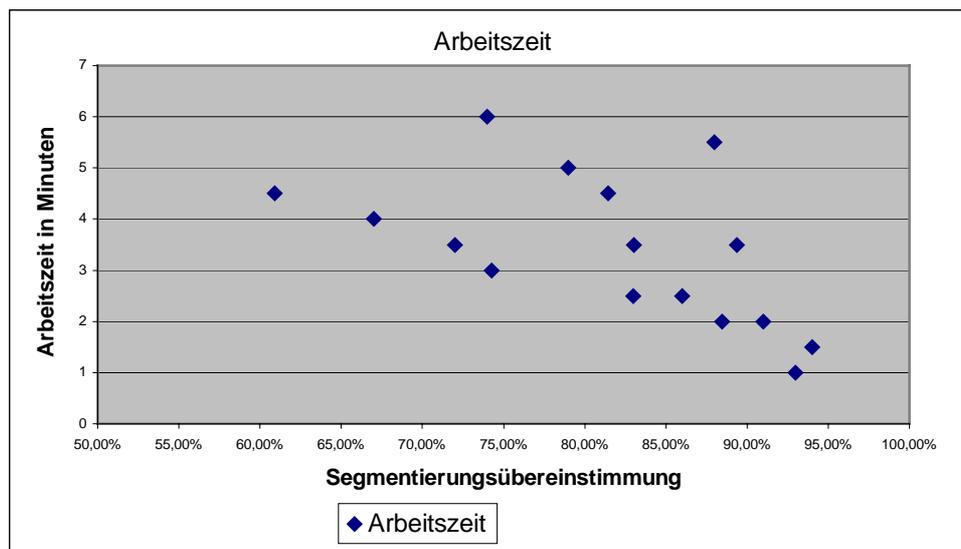


Abbildung 6.33: Zeitaufwände für Nachbearbeitung und Korrektur der Segmentierungsergebnisse

Die Abbildung trägt den Zeitaufwand, der für die Nachbearbeitung und Korrektur der modellbasierten Segmentierungsergebnisse aufgebracht werden mußte über den Übereinstimmungen der Segmentierung mit der Musterlösung auf.

Der effektive Arbeitsaufwand setzt sich jedoch nicht nur aus dem Arbeitsaufwand zusammen, der in Verbindung mit der modellbasierten Segmentierung aufgebracht werden muß. Zusätzlich muß der Arbeitsaufwand eingerechnet werden, der für die Nachbearbeitung und insbesondere für die Korrektur möglicher Fehler aufgebracht werden muß. Um auch diesen Arbeitsaufwand einzubeziehen, wurden die Segmentierungen der Patientendatensätze, die durch Anwendung des modellbasierten Snake gewonnen wurden, nachbearbeitet und Fehler in der Segmentierung korrigiert. Die Zeiten, die für diese Korrekturen

aufgebracht werden mußten, wurden gemessen und in Abbildung 6.33 im Verhältnis zu der Korrektheit der jeweiligen modellbasierten Segmentierungsergebnisse aufgetragen.

Wie in der Abbildung zu sehen ist, ist die Arbeitszeit, die zur Korrektur der Segmentierungsergebnisse aufgebracht werden muß, leicht mit der Übereinstimmung der Segmentierung mit der Musterlösung korreliert. Es besteht die Tendenz, daß je geringer die Übereinstimmungen sind, desto größer ist der aufzubringende Arbeitsaufwand. Diese Tendenz ist naheliegend, da mehr Fehler bei der Segmentierung auch einen erhöhten Arbeitsaufwand zur Korrektur dieser Fehler mit sich bringen. Dabei scheint es eine nahezu lineare Verbindung zwischen Übereinstimmungen und Arbeitsaufwand zu geben. Die Korrelation der beiden Größen ist aufgrund von mehreren Faktoren jedoch nicht sehr stark ausgeprägt.

Es sind nicht so sehr die Übereinstimmungen der Flächen, die über den Arbeitsaufwand entscheiden, sondern die Anzahl der Konturpunkte, die verschoben werden müssen, um die korrekte Segmentierung herzustellen. Leichte Fehler in den zentralen Schichten wirken sich aufgrund der höheren Anzahl von Punkten stärker auf den Arbeitsaufwand aus als Fehler in den Randschichten. Bei der Korrektur der Fehler ist es unerheblich, wie stark die Punkte von der korrekten Position entfernt sind. Entscheidend ist in erster Linie die Anzahl der Punkte, die verschoben werden müssen. Dadurch wird die direkte Korrelation von Übereinstimmung und Arbeitsaufwand aufgehoben. Selbst geringe Abweichungen können zu einem erheblichen Arbeitsaufwand führen, wenn die Punkte um noch so geringe Distanzen verschoben werden müssen. Andererseits können auch sehr erhebliche Unterschiede gegebenenfalls durch einen geringen Arbeitsaufwand korrigiert werden, wenn sich die Fehler auf wenigen Bereiche oder Schichten beschränken. Bei der Nachbearbeitung der Konturen hat sich außerdem eine Korrekturrichtlinie herauskristallisiert. Wenn mehr als ein Drittel der Konturpunkte korrigiert werden müssen, ist das Neuzeichnen der Kontur schneller als das Verschieben der einzelnen Konturpunkte.

In den Randschichten kommt es aufgrund der starken Partialvolumeneffekten oft zu sehr geringen Schnittflächen. Diese geringe Überschneidung der Segmentierungen hat jedoch nicht immer zu einer aufwendigen Korrektur oder zu einem Neuzeichnen der Kontur geführt. In vielen Fällen wurde die Form der Kontur durch die interne Energie erhalten, so daß mit geringem Aufwand die gesamte Kontur an die korrekte Stelle verschoben werden konnte. Ein Verschieben der einzelnen Konturpunkte war in einigen Fällen dann nicht mehr nötig. Wurde die Form nicht erhalten, erwies sich ein Neuzeichnen der Kontur oft als schneller als die Korrektur der einzelnen Konturpunkte.

Zusammenfassend kann festgestellt werden, daß die modellbasierte Segmentierung trotz der erforderlichen Nachbearbeitung einen beträchtlichen Zeitgewinn gegenüber der manuellen Segmentierung darstellt. Andererseits kann der Arbeitsaufwand für die Korrektur von Segmentierungsfehler im ungünstigen Fall schnell den zuvor erreichten Zeitgewinn beträchtlich schrumpfen lassen. Nichtsdestotrotz wurde bei allen Tests für die modellbasierte Segmentierung einschließlich der Nachbearbeitung weniger Zeit benötigt als für eine rein manuelle Segmentierung. Diese auf Anhieb sehr guten Ergebnisse der modellbasierten Segmentierung lassen die Hoffnung aufkommen, daß die Segmentierungskorrektheit durch weiterführende Forschungsarbeiten noch weiter gesteigert und damit der effektive Zeitgewinn für die klinische Routine noch mehr reduziert werden kann.

6.4 Modellbasierte Interpolation

Die modellbasierte Interpolation erlaubt die oberflächenbasierte Interpolation eines Datensatzes ohne zuvor eine manuelle Segmentierung durchgeführt zu haben. Bei der modellbasierten Interpolation stehen die beiden möglichen Vorgehensweisen zur Auswahl, entweder nur die Referenzsegmentierung an den aktuellen Datensatz anzupassen oder die gesamte modellbasierte Segmentierung einschließlich des modellbasierten Snake anzuwenden. Vorteilhaft für die Anwendung in der klinischen Routine ist vor allem die modellbasierte Interpolation basierend allein auf der Anpassung des Referenzmodells, da mit dieser

Methode ohne jede Wartezeit eine qualitativ hochwertige Interpolation durchgeführt werden kann. Aus diesem Grund wurde zur Beurteilung der modellbasierten Interpolation ausschließlich dieser Ansatz zur modellbasierten Interpolation getestet.

Zur Beurteilung der Interpolationsergebnisse wurde der Algorithmus auf die Patienten- und die Modelldaten angewandt und anschließend die Ergebnisse beurteilt. Die Teststrategien können in qualitative und quantitative Evaluationsverfahren unterschieden werden. Während sich bei der qualitativen Beurteilung der Interpolationsergebnisse eine gewisse Subjektivität nicht ausschließen läßt, ist bei der quantitativen Auswertung zunächst die Definition eines Korrektheitsmaßstabes erforderlich. Gerade da die optimale Lösung der Interpolation in der Regel nicht bekannt ist, können nur Evaluationsverfahren gewählt werden, die die Ergebnisse mit plausiblen Ergebnissen vergleichen. Für die qualitative und quantitative Bewertung der modellbasierten Interpolation wurden die folgenden Testreihen durchgeführt:

- Es wurden Patientendaten durch Anpassen von geeigneten Referenzmodellen segmentiert und anschließend unter Verwendung dieser Segmentierungen oberflächenbasiert interpoliert. Die Interpolationsergebnisse wurden qualitativ auf die Erhaltung von Kanten und das Fortführen von anatomisch sinnvollen Formen untersucht.
- Die Modelldaten wurden interpoliert und anschließend die Volumina für die interpolierten Organoberflächen berechnet. Zur quantitativen Beurteilung wurden die berechneten Volumina des ursprünglich segmentierten Datensatzes und des interpolierten Datensatzes mit dem bekannten, theoretischen Volumen der einzelnen Objekte verglichen.

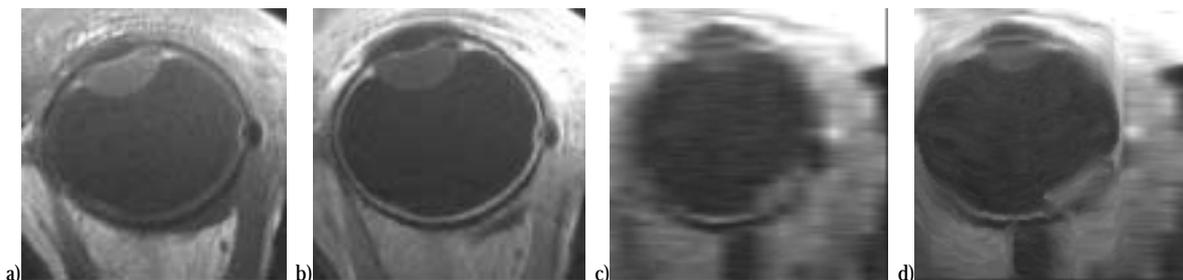


Abbildung 6.34: Ergebnisse der linearen und modellbasierten Interpolation bei Patientendaten

Die Abbildungen zeigen die Ergebnisse der linearen und modellbasierten Interpolation anhand eines T1-MRT-Datensatzes. Links (a) ist eine Zwischenschicht zu sehen, die mittels linearer Interpolation berechnet wurde. Daneben ist die gleiche Schicht abgebildet, die durch die modellbasierte Interpolation gewonnen wurde. Vor allem an den Kanten ist zu sehen, daß die oberflächenbasierte Interpolation deutlich besser die Kanten erhält als die lineare Interpolation. Noch deutlicher ist dieser Effekt in dem orthogonalen Schnitt durch die Datensätze zu sehen. Auf der zweiten Abbildung von rechts (c) ist der orthogonale Schnitt durch den linear interpolierten Datensatz zu sehen. Sämtliche Organgrenzen sind stark verwischt. Bei der modellbasierten Interpolation (d) hingegen konnten die Organkanten vergleichsweise gut erhalten werden.

Abbildung 6.34 zeigt am Beispiel eines T1-MRT-Datensatzes den Unterschied zwischen der linearen und der oberflächenbasierten Interpolation, für die zuvor die Referenzsegmentierung an den Datensatz angepaßt wurde. Wie deutlich zu sehen ist, erhält die oberflächenbasierte Interpolation die Kanten deutlich besser als die lineare Interpolation. Vor allem die Darstellung der Sklera wird bei der linearen Interpolation deutlich stärker verwischt als es bei der oberflächenbasierten Interpolation der Fall ist. Noch deutlicher wird der Unterschied zwischen linearer und oberflächenbasierter Interpolation bei der Betrachtung eines orthogonalen Schnitts durch die interpolierten Datensätze. Während bei dem linear interpolierten Datensatz die Konturen des Bulbus deutlich verwischt werden, werden bei dem oberflächenbasiert interpolierten Datensatz die Kanten zwischen Bulbus und umgebenden Fettgewebe deutlich besser herausgestellt.

Noch deutlicher werden die Vorteile der oberflächenbasierten Interpolation bei der Auswertung der interpolierten Phantomdatensätze in Abbildung 6.35. Vor allem in Schichten, in denen deutliche Unterschiede zwischen den angrenzenden Bildern bestehen, sind die für die lineare Interpolation üblichen Überblendeffekte zu sehen. Vor allem bei Kanten mit deutlichen Grauwertänderungen und starken Verschiebungen zwischen den Schichten werden die Kanten beider Schichten nebeneinander dargestellt. Solche verfälschenden Effekte können durch die Anwendung einer oberflächenbasierten Interpolation vermieden werden. In den orthogonalen Schichten durch die interpolierten Datensätze in Abbildung 6.35 ist sehr gut zu erkennen, daß diese Überblendeffekte bei der linearen Interpolation vor allem in den Randschichten ausgebildet werden. Bei der oberflächenbasierten Interpolation wird dieser Fehler gänzlich vermieden, wie an der sauberen und runden Kugelkante erkannt werden kann.

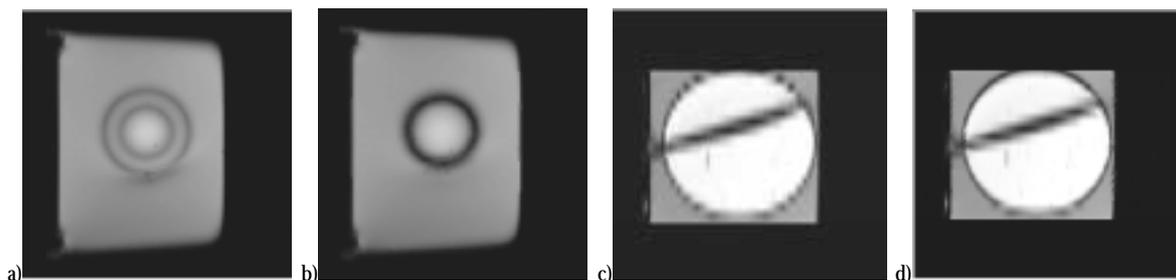


Abbildung 6.35: Ergebnisse der linearen und modellbasierten Interpolation von Phantomdaten

Die Abbildungen zeigen die Ergebnisse der linearen und modellbasierten Interpolation von einem beispielhaften Phantomdatensatz. Links ist die lineare Interpolation einer Randschicht zu sehen. Wie deutlich zu erkennen ist werden durch das einfache Überblenden zwischen den Schichten die dunklen Randkonturen beider Schichten nebeneinander dargestellt. Bei der modellbasierten Interpolation in der Abbildung rechts daneben (b) wird dieser Effekt vermieden und statt dessen der Rand sauber zwischen den Schichten weitergeführt. Der Spiegeleffekt der linearen Interpolation ist auch im orthogonalen Schnitt durch den Datensatz in Abbildung (c) gut zu erkennen. Bei dem Datensatz, der modellbasiert interpoliert wurde, wird die Kugeloberfläche sauber auf die Zwischenschichten übertragen. Lediglich der Lufteinschluß innerhalb der Kugel weist die gleichen Darstellungsprobleme auf wie bei der linearen Interpolation. Dieser Lufteinschluß wurde vor der modellbasierten Interpolation nicht segmentiert und wurde daher bei der Berechnung der modellbasierten Interpolationspfade nicht berücksichtigt.

Bei der Auswertung der berechneten Volumina zeigt sich, daß auch die Objektrekonstruktion durch die oberflächenbasierte Interpolation verbessert wird. Während ohne eine Interpolation das Bulbus-Volumen durchschnittlich mit einem Fehler von 6,5% behaftet ist, verringert sich dieser Fehler bei einer oberflächenbasierten Interpolation auf 4,27%.

<i>Volumenberechnung</i>	<i>Durchschnittliche gemessenes Volumen</i>	<i>Ideal-Volumen</i>	<i>Mittlerer Fehler</i>	<i>Mittlere Standardabweichung</i>
ohne Interpolation	242,3785	258,1546167	6,51%	2,46%
mit Interpolation	247,5437	258,1546167	4,27%	1,85%

Tabelle 6-15: Berechnete Volumina für die Phantomdatensätze mit und ohne Interpolation

Die Tabelle zeigt die Ergebnisse der Volumenberechnung mit und ohne vorherige Interpolation der Phantomdatensätze. Durch die Interpolation werden Ungenauigkeiten vor allem im Randbereich der Konturen vermieden. Dadurch wird nicht nur die Bildqualität der Zwischenschichten verbessert, sondern auch die Genauigkeit von Volumenberechnungen erhöht.

Bei der modellbasierten Interpolation konnte jedoch ein Problempunkt identifiziert werden, der im Rahmen dieser Arbeit nicht behandelt werden kann. Dabei handelt es sich um eine ähnliche Problemstellung, die bereits bei der Diskussion des modellbasierten Snake angesprochen wurde: der Abstand der Konturpunkte zueinander. Abbildung 6.36 veranschaulicht die Problemstellung, die daraus resultiert, daß

die Konturpunkte der Segmentierung oft recht weit voneinander entfernt sind. Bei der Korrelation der Segmentierungskonturen von zwei angrenzenden Schichten zur Berechnung der Interpolationspfade müssen die zweidimensionalen Konturen miteinander korreliert werden. Bei dieser Korrelation werden die Punkte der einen Kontur Punkten der anderen Kontur zugeordnet. Diese dadurch aufgespannten Korrespondenzvektoren geben vor, wie die Interpolationspfade zwischen den Schichten ausgerichtet werden. Durch den begrenzten Abstand der Konturpunkte zueinander richten sich die Pfeile stellenweise tangential zu den Segmentierungskonturen aus. Um eine kontinuierliche Transformation der einen Kontur in die anderen umzusetzen wären jedoch Korrespondenzvektoren nötig, wie sie in der mittleren Grafik von Abbildung 6.36 dargestellt werden. Um solche Vektoren zu berechnen müßten aber die Korrespondenzpunkte analytisch berechnet werden. Bei einer solchen Methode zur Bestimmung der Korrespondenzvektoren treten jedoch umfangreiche Probleme bezüglich der Eindeutigkeit der Korrespondenzen auf, wie sie in der rechten Grafik veranschaulicht werden.

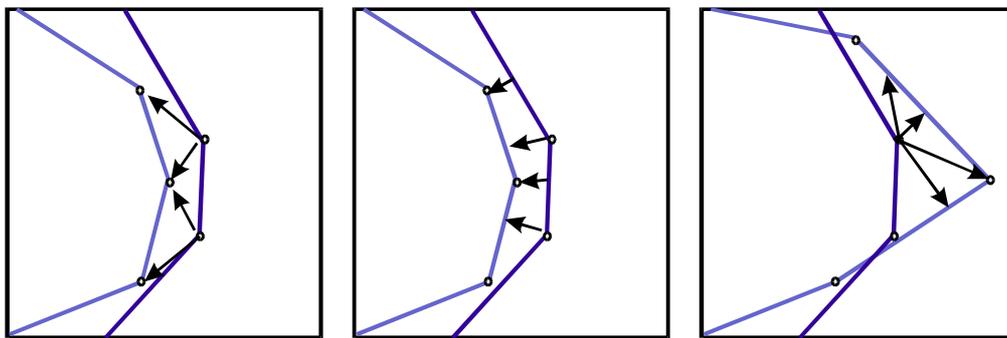


Abbildung 6.36: Problemstellung bei der Berechnung von Konturkorrespondenzen

Die Abbildung veranschaulicht anhand von einigen charakteristischen Konstellationen von angrenzenden Segmentierungskonturen die Problemstellungen bei der Berechnung von Korrespondenzvektoren zwischen den Konturen. Es zeigt zwei Kontursegmente, die miteinander korreliert werden sollen. Bei einer direkten Korrelation der Konturpunkte würden die Korrespondenzvektoren der linken Abbildung berechnet werden. Das Problem bei diesen Korrespondenzvektoren besteht darin, daß der Interpolationsfunktion tangentiale Interpolationspfade zugrunde gelegt werden. Diese Anpassung der Konturen ist jedoch nicht gewünscht. Statt dessen ist eher ein radiales Anpassen der Konturen aneinander gewünscht, wie es in der mittleren Abbildung dargestellt wird. Das Problem dabei ist die Berechnung der korrespondierenden Punkte zwischen den eigentlichen Konturpunkten. In der rechten Abbildung wird anhand einer anderen Konstellation der Segmentierungskonturen veranschaulicht, daß darüber hinaus Situationen entstehen können, in denen nicht ohne weiteres festgestellt werden kann, welches nun der geeignete Korrespondenzvektor ist. Gegebenenfalls wäre es sogar sinnvoll, mehrere Korrespondenzvektoren für einen Konturpunkt zu konstruieren. Die Frage ist nur, wann ein solche Konstruktion von mehreren Referenzvektoren angezeigt ist und wann nicht.

Eine Vermeidung dieser Probleme durch die Erhöhung der Punktzahl ist wiederum mit den gleichen Problemen behaftet wie bei dem modellbasierten Snake. Zu viele Punkte erhöhen den Arbeitsaufwand, falls die Kontur zu einem späteren Zeitpunkt manuell nachbearbeitet werden soll. Eine nachträgliche Elimination von Punkten ist problematisch, da der Benutzer die Konturpunkte oft bei anatomischen Landmarken positioniert und bei der Elimination die Gefahr besteht, daß diese wichtigen Konturpunkte ebenfalls eliminiert werden könnten. Die wahrscheinlich aussichtsreichste Methode wäre das temporäre Einfügen von Punkten an den numerisch berechneten optimalen Punkten für die Interpolation. Anschließend können diese temporären Punkte wieder entfernt werden. Doch wäre bei diesem Ansatz noch nicht das Problem der Mehrdeutigkeit der optimalen Korrespondenzen gelöst. Diese Überlegungen veranschaulichen, daß der Aspekt der Berechnung von Korrespondenzvektoren wissenschaftlich durchaus interessant ist und bei weiterführenden Arbeiten genauer behandelt werden sollte.

Es muß jedoch einschränkend erwähnt werden, daß bei den immer besser werdenden Auflösungen aktueller MRT-Datensätze die Interpolation der Daten an Bedeutung verliert. Nichtsdestotrotz kann

durch die Verwendung der Interpolation die Auflösung der Datensätze so weit erhöht werden, daß deren Verwendung als Referenzmodelle deutlich beschleunigt werden kann, indem keine Interpolation für die Anpassung der Referenzdaten mehr nötig wird. Darüber hinaus wird bei anderen medizinischen Aufgabenstellungen in der medizinischen Bildverarbeitung gegebenenfalls die Möglichkeit zur Anfertigung von hochaufgelösten Datensätzen gegebenenfalls nicht möglich oder nicht rentabel sein. In solchen Fällen ermöglicht die modellbasierte Interpolation von Datensätzen mit regulären Organen eine deutliche Verringerung des Arbeitsaufwandes, der bei herkömmlichen oberflächenbasierten Interpolationsverfahren vor allem aus der umfangreichen manuellen Segmentierung der dargestellten Organstrukturen resultiert.

6.5 Modellbasierte Registrierung

Bei der modellbasierten Registrierung werden die gleichen Algorithmen verwandt wie bei der landmarkenbasierten Registrierung von aktuellem Datensatz und Referenzmodell. Die in Abschnitt 6.3.2 beschriebenen Testergebnisse belegen daher bereits die guten Anpassungseigenschaften der modellbasierten Registrierung. Im Unterschied zu der bisher beschriebenen Aufgabenstellung, Referenzmodelle mit dem aktuellen Datensatz zu registrieren, besteht bei der modellbasierten Registrierung außerdem die Aufgabenstellung, MRT-Daten mit CT-Daten zu registrieren. Gerade in Hinsicht auf die Durchführung der Bestrahlungssimulation wäre eine Abschätzung der Genauigkeit einer solchen Kombination von CT- und MRT-Daten von großem Interesse.

Ähnlich wie bei der modellbasierten Interpolation stellt sich jedoch bei der modellbasierten Registrierung das Problem, daß die Präzision der Registrierung nur qualitativ abgeschätzt werden kann. Werden Schichtdaten unterschiedlicher Bildmodi miteinander registriert, besteht das Problem, daß nicht bekannt ist, welches Voxel des einen Datensatzes zu welchem Voxel des anderen Datensatzes korrespondiert. Vor allem bei den Patientendaten kann nicht von einer eindeutigen Korrelation der Grauwerte in den CT- und den MRT-Daten ausgegangen werden. Durch die unterschiedlichen Darstellungsprinzipien kann nicht einmal sichergestellt werden, daß das Dargestellte des einen Datensatz im anderen überhaupt abgebildet wird. Bei den CT-Daten kommt erschwerend hinzu, daß eine präzise Segmentierung der Datensätze aufgrund des geringen Grauwertkontrastes nicht möglich ist. Ein Vergleich der manuellen und der angepaßten Segmentierungen wie bei der landmarkenbasierten Registrierung in Abschnitt 6.3.2.2 wäre daher nicht sinnvoll. Eine quantitative Auswertung der Registrierungsergebnisse ist damit nur unter Vorbehalt möglich. Der Wert einer solchen Auswertung ist fraglich, da sie die Realität voraussichtlich deutlich verfälscht wiedergeben. Die Qualität der Multimodalen Registrierungsergebnisse von Patientendaten kann demnach nur grob abgeschätzt werden.

Etwas anders verhält es sich bei der Auswertung der Phantomdatensätze. Die bei der Konstruktion der Phantomdatensätze verwendeten Materialien wurden bewußt so ausgewählt, daß die einzelnen Teilkomponenten der Modelle sowohl in den MRT-Daten als auch in den CT-Daten gut erkannt werden können. Dadurch wird ein quantitativer Vergleich der manuellen Segmentierung in den CT-Daten und der aus den MRT-Daten registrierten Referenzsegmentierung möglich. Zur Beurteilung der modellbasierten Registrierung wurden daher die folgenden Tests durchgeführt:

- Es werden CT- und MRT-Datensätze der Phantommodelle miteinander registriert. Zur quantitativen Bewertung der Registrierungsergebnisse wurden die angepaßten Referenzsegmentierungen mit den manuellen Segmentierungen verglichen, indem die Schnittflächen und die Schnittvolumina berechnet wurden.
- Es wurden MRT-T1 Datensätze mit den CT-Datensätzen des gleichen Patienten registriert und die dabei berechneten Ergebnisse qualitativ beurteilt.

Tabelle 6-16 zeigt die Ergebnisse der modellbasierten Registrierung für einen beispielhaften Phantomdatensatz. In der Tabelle werden die durch die Registrierung in den CT-Datensatz übernommene Segmentierungen des MRT-Datensatz verglichen mit der manuellen Segmentierung des CT-Datensatzes. Ähnlich wie bei den Ergebnissen der anderen Module der modellbasierten Bildverarbeitung ist die Übereinstimmung in den zentralen Schichten am größten. Die größten Divergenzen treten in den Randschichten auf. Die Ursache hierfür ist in der kritischen Positionierung der Landmarken fünf und sechs zu sehen.

Schicht	Fläche Datensatz 1	Fläche Datensatz 2	Schnittfläche	Prozentuale Übereinstimmung
Slice 3	0,036	0,224	0,035	27,21%
Slice 4	3,669	5,96	3,669	76,21%
Slice 5	7,377	10,076	7,377	84,53%
Slice 6	12,652	14,095	12,567	93,97%
Slice 7	15,965	16,939	15,569	94,63%
Slice 8	19,209	20,165	18,719	95,08%
Slice 9	22,194	23,908	22,063	95,71%
Slice 10	24,878	25,151	24,008	95,98%
Slice 11	27,072	28,187	26,475	95,82%
Slice 12	29,505	30,581	28,966	96,42%
Slice 13	30,91	33,077	30,635	95,75%
Slice 14	32,429	34,521	32,323	96,56%
Slice 15	33,972	35,766	33,887	97,19%
Slice 16	35,003	36,74	34,919	97,34%
Slice 17	36,086	37,965	35,944	97,08%
Slice 18	37,393	38,14	36,933	97,79%
Slice 19	37,908	38,696	37,619	98,22%
Slice 20	38,158	38,686	37,632	97,95%
Slice 21	38,164	38,449	37,618	98,20%
Slice 22	38,31	38,401	37,483	97,72%
Slice 23	38,062	37,302	36,946	98,05%
Slice 24	37,303	36,82	36,462	98,38%
Slice 25	36,245	35,387	35,156	98,16%
Slice 26	35,091	34,028	33,741	97,63%
Slice 27	33,758	32,203	32,03	97,12%
Slice 28	31,948	30,289	29,809	95,79%
Slice 29	29,013	28,202	27,76	97,04%
Slice 30	26,518	25,66	25,334	97,11%
Slice 31	22,706	22,742	21,748	95,70%
Slice 32	19,158	20,618	19,014	95,61%
Slice 33	15,105	17,067	15,104	93,90%
Slice 34	10,92	13,576	10,92	89,15%
Slice 35	6,941	9,862	6,941	82,62%
Slice 36	2,756	5,876	2,756	63,85%
Slice 37	0,468	0,507	0,439	90,05%

Tabelle 6-16: Schichtweiser Vergleich der Registrierungsergebnisse von MRT- und CT-Daten eines Phantomdatensatzes mit der manuellen Segmentierung

Die Tabelle vergleicht die manuellen Segmentierungsergebnisse eines CT-Phantomdatensatzes mit denen, die bei der Registrierung eines korrespondierenden segmentierten MRT-Datensatzes mit dem CT-Datensatz berechnet wurden. In über 85% der Schichten liegt die Übereinstimmung zwischen der durch die Registrierung berechneten Segmentierungskonturen mit denen der manuellen Segmentierung bei über 95%.

Die Auswertung der Schnittvolumina bei 7 Registrierungstests zeigte, daß die durchschnittlichen Übereinstimmungen zwischen den registrierten Segmentierungen und den manuellen Segmentierungen bei über 96% liegen. Diese großen Übereinstimmungen traten bei sämtlichen Tests gleichermaßen auf, so daß die Standardabweichung bei diesen Tests lediglich 1,5% betrug.

Die quantitativen Tests belegten demnach die Korrektheit der Registrierung von MRT- und CT-Daten. Da diese multimodale Registrierung grundsätzlich mit der unimodalen Registrierung übereinstimmt, war dieses Ergebnis aufgrund der bereits zuvor beschriebenen guten Ergebnisse der landmarkenbasierten Registrierung keine Überraschung.

Die multimodalen Registrierungsergebnisse von MRT- und CT-Datensätzen lassen aus den bereits erwähnten Gründen nur qualitative Aussagen zu. Abbildung 6.37 zeigt die primären Ergebnisse der Registrierung von beispielhaften CT- und MRT-Daten eines Patienten. In der linken Abbildung ist die Segmentierungskontur zu sehen, die anhand der modellbasierten Registrierungen von dem segmentierten MRT-Datensatz in den CT-Datensatz übertragen wurde. Die gelben Kreuze zeigen die Landmarken, die für die Registrierung eingegeben wurden. In der rechten Grafik ist der angepaßte MRT-Datensatz zu se-

hen. Der hochaufgelöste Datensatz wurde bei der Registrierung verkleinert und leicht rotiert, um die beiden Datensätze anatomisch korrekt miteinander korrelieren zu können.

Abbildung 6.38 blendet die in Abbildung 6.37 dargestellten Registrierungsergebnisse gleichzeitig übereinander. In der Abbildung sind deutlich die weitgehenden Übereinstimmungen der beiden Datensätze durch die Registrierung zu sehen. Lediglich in den Randbereichen der eingeblendeten MRT-Daten sind bei der Nase und der Wange des Patienten Kanten zwischen den MRT- und CT-Daten zu erkennen. Diese Unterschiede zwischen den Datensätzen resultieren aus zwei Problemen, die die Registrierung beeinflussen.

Kurz bevor die MRT-Aufnahmen angefertigt werden, werden dem Patienten üblicherweise Tantal-Clips auf die Augenoberfläche aufgenäht, um bei der Bestrahlungstherapie eine Korrelation des Bestrahlungsplans mit der aktuellen Position und Orientierung des Auges am Patienten vornehmen zu können. Röntgenaufnahmen, die während der Bestrahlungsdurchführung diese Clips von der Seite darstellen, ermöglichen diese Korrelation. Um diese Clips aufzubringen, muß das Gewebe im Bereich der Orbita betäubt werden. Diese Betäubung geht immer mit einer nicht unbeträchtlichen Schwellung des Gewebes und vor allem der Augenlider einher. Diese Schwellung wird auf den MRT-Daten deutlich erkennbar an den dicken Lidern abgebildet. Bis die CT-Aufnahmen angefertigt werden, ist diese Schwellung in der Regel vollständig zurückgegangen. Demnach bestehen zwischen den CT- und MRT-Aufnahmen anatomische Unterschiede, die sich folgerichtig auch bei der Registrierung wie in Abbildung 6.38 zeigen.



Abbildung 6.37: Ergebnisse der Registrierung von MRT- und CT-Daten

Die Abbildungen zeigen einerseits die mittels Registrierung in den CT-Datensatz transformierte Segmentierung des MRT-Datensatzes (links), andererseits den an den CT-Datensatz angepaßten MRT-Datensatz (rechts). Der große schwarze Bereich um die MRT-Schicht entstand durch das Anpassen der MRT-Daten, die deutlich kleiner und leicht anders orientiert waren als die CT-Daten.

Gewisse anatomische Unterschiede lassen sich zwischen den CT- und MRT-Daten nicht vollständig ausschließen. In der Regel sind sie jedoch nicht sonderlich stark ausgeprägt und werden auch bei der Registrierung nicht auffallen.

Problematischer ist hingegen die Tatsache, daß die Orientierung des Patientenauges zwischen den Datensätzen deutlich variieren kann. Während beispielsweise der Patient bei den MRT-Daten geradeaus schaut, kann es sein, daß er bei den CT-Daten den Blick leicht nach links wendet. Solche Veränderungen der Augenstellung wirken sich direkt auf die Registrierung aus. Die Registrierung der beiden Datenmodi

wird anhand der anatomischen Strukturen des Auges vorgenommen. Wird das Auge gedreht, wird sich auch der registrierte MRT-Datensatz innerhalb der CT-Daten drehen. Dadurch werden zwar die anatomischen Strukturen des Auges wie Linse, Ziliarkörper und Bulbus in Übereinstimmung gebracht. Die umgebenden Strukturen wie Knochen, Fettgewebe und Gesichtsrand werden jedoch gegeneinander verdreht. Sowohl die Betäubung als auch eine leichte Rotation des Auges führten zu den Unterschieden, die in Abbildung 6.38 dargestellt werden.

Eine mögliche Strategie, mit dieser Problematik umzugehen, wäre die separate Registrierung von Bulbus und umgebenden Gewebestrukturen. Die Umsetzung eines solchen Registrierungsverfahrens ist jedoch alles andere als trivial. Denn durch die Bewegung des Auges wird nicht nur der Bulbus bewegt, sondern auch Teile des umgebenden Gewebes wie beispielsweise der Sehnerv und die Muskeln. Mit diesen Muskeln wird sich auch das umgebende Fettgewebe verändern. Demnach hat man es mit einer plastischen Verformung der betroffenen anatomischen Strukturen zu tun. Ob und wie eine solche Registrierung umgesetzt werden kann, müßte in einem eigenen Forschungsvorhaben behandelt werden. Es ist jedoch fraglich, inwieweit die Entwicklung einer solch aufwendigen Registrierung sinnvoll wäre. In Hinsicht auf den Kosten-Nutzen-Effekt ist es günstiger, Wege zu finden, wie man den Patienten dazu bringen kann, während der Aufnahmen in eine einheitlich definierte Richtung zu schauen.



Abbildung 6.38: Registrierte MRT- und CT-Daten

Die Abbildung zeigt das Ergebnis der Registrierung von CT- und MRT-Daten. Dabei werden die Daten gleichzeitig ineinander projiziert dargestellt. Es ist deutlich die gute Übereinstimmung beider Daten zu erkennen. Die leichten Differenzen zwischen den beiden Datensätzen, die vor allem an der Hautkante der Wange und der Nase zu erkennen sind, resultieren aus der leicht unterschiedlichen Orientierung des Auges und den in den MRT-Daten geschwollenen Augenlidern. Die Schwellung der Augenlider bei den MRT-Daten resultiert aus der Anästhesie der Orbita, die angewandt wird, um die Oberflächenclips anzubringen.

6.6 Ablauf eines beispielhaften Falls

Ausgangspunkt für die Entwicklung des in dieser Arbeit beschriebenen Bildverarbeitungsansatzes war die Verbesserung der Strahlentherapie von Augentumoren. Hierzu mußten wissenschaftliche Forschungsarbeiten in den Bereichen der medizinischen Bildverarbeitung, der Bestrahlungssimulation und der Bestrahlungsdurchführung unter Verwendung von Protonenstrahlen zur Behandlung der intraokularen Tumoren durchgeführt werden. Während der Bildverarbeitungsanteil durch die in dieser Arbeit beschriebenen Algo-

rithmen abgedeckt wird, wurde die Bestrahlungssimulation durch das Deutsche Krebsforschungszentrum in Heidelberg und die Bestrahlungsdurchführung durch das Hahn-Meitner-Institut in Berlin bearbeitet.

Um die Zweckmäßigkeit des in dieser Arbeit beschriebenen Bildverarbeitungsansatzes besser abschätzen zu können, soll in den nachfolgenden Abschnitten beschrieben werden, wie sich das Szenario einer verbesserten Strahlentherapie von Augentumoren gestalten könnte. Bedauerlicherweise muß in diesem Zusammenhang der Konjunktiv benutzt werden, da einerseits die Forschungsarbeiten der Kooperationspartner noch nicht den Abschluß gefunden haben, und andererseits aufgrund von diversen koordinativen Problemen die Integration der einzelnen Forschungssegmente noch nicht bis zu einem praktisch anwendbaren Forschungsstand fortgeschritten ist. Nichtsdestotrotz wird in den nachfolgenden Abschnitten beschrieben, welche Form die Bestrahlungstherapie in der klinischen Routine im voraussichtlichen Endstadium der Arbeiten annehmen könnte.

6.6.1 Aufnahme in der Klinik

Die erste Diagnose von intraokularen Tumoren wird in der Regel durch den behandelnden Augenarzt vorgenommen. Aufgrund des diffusen Wachstums in den Augapfel hinein, wird der Augentumor in der Regel durch eine weiße Trübung des Augapfels erkannt. Sofern sich der Tumor in der Nähe der Macula befindet wird er oft auch durch die betroffene Person bemerkt. Da es sich bei den von Retinoblastomen betroffenen Personen oft um Kinder zwischen drei und vierzehn Jahren handelt, werden die ersten Symptome der Sichteinschränkung anfänglich nicht mit der angemessenen Aufmerksamkeit aufgenommen. Dies führt dazu, daß der Tumor in vielen Fällen bereits sehr fortgeschritten ist. Nach einer ersten Diagnose des Tumors anhand einer einfachen Betrachtung des Augenhintergrunds wird der Patient in der Regel sofort an ein Klinikum verwiesen, das mit der Behandlung derartiger Tumoren vertraut ist. Ein Beispiel hierfür ist das Universitätsklinikum Benjamin Franklin.

6.6.2 Diagnostik

Um zunächst den allgemeinen Krankheitszustand abzuschätzen, werden in einem ersten diagnostischen Schritt Fundusaufnahmen, Ultraschallaufnahmen, CT-Schnittbilder und sowohl T1- als auch T2-gewichtete MRT-Datensätze aufgenommen. Um später bei der Bestrahlungsplanung die Position und Orientierung des Auges mit den Bilddaten und damit mit dem Bestrahlungsplan korrelieren zu können, werden dem Patienten vor der Anfertigung der Bilddaten Tantal-Clips auf die Bulbusoberfläche angehängt. Diese Clips werden üblicherweise um den Tumor herum positioniert und erlauben damit bereits eine grobe Lokalisation des Tumors während der Bestrahlungsdurchführung.

Die Fundusaufnahmen werden ausgewertet, um den Abstand des Tumors zur Makula genau abzumessen. Die Ultraschallaufnahmen erlauben eine schnelle Abschätzung der Prominenz des Tumors und die Ausprägung der Tumorbasis. Die MRT-T1 und T2 Bilddatensätze werden für eine umfassende Abschätzung des Tumolvolumens ausgewertet. Es wird dabei abgeschätzt, wie diffus der Tumor in den Glaskörper hinein wächst und welche Behandlungsoptionen zur Verfügung stehen. Durch die kombinierte Betrachtung von T1- und T2-gewichteten MRT-Aufnahmen kann außerdem eine Ablation der Sklera vom Tumolvolumen unterschieden werden. Während der Tumor und die flüssigkeitsgefüllte Ablation in den T1-Aufnahmen gleichermaßen hell dargestellt werden, bleibt die Ablation in den T2-Aufnahmen dunkel, während Tumor und Glaskörper hell sind.

Falls der Tumor recht weit vorne angesiedelt ist, ist auch eine operative Resektion des betreffenden Sklerasegments möglich, ohne dabei eine grundsätzliche Einschränkung der Sehfähigkeit zu riskieren. Befindet sich jedoch der Tumor weiter hinten (posterior) im Auge, ist eine Resektion nicht mehr möglich, da der Zugang zu dem betreffenden Sklerasegment nicht mehr möglich ist.

Wurde anhand der MRT-Daten festgestellt, daß sich der Tumor zu weit hinten befindet für ein operatives Entfernen, wird in der Regel die Entscheidung getroffen, den Tumor durch eine Bestrahlungstherapie zu behandeln. Hierzu müssen die aufgenommenen Bilddaten für die Rekonstruktion eines Bestrahlungsmodells genutzt werden, daß einer Bestrahlungstherapie mittels Protonenstrahlen zugrunde gelegt werden kann.

6.6.3 Rekonstruktion des Bestrahlungsmodells

Wichtige Bestandteile des Bestrahlungsmodells sind die genaue Beschreibung der Patientenanatomie und eine Beschreibung der Absorptionseigenschaften innerhalb der Patientenanatomie, um die Absorption der Protonenstrahlen anhand des Modells simulieren zu können. Zur Rekonstruktion der Patientenanatomie wird die modellbasierte Segmentierung auf die MRT-Daten angewandt. Aufgrund der guten Bildgebungseigenschaften werden hierzu in erster Linie die T1-MRT-Daten ausgewertet. Falls zusätzlich auch T2-Aufnahmen hinzugezogen werden sollten, werden auch diese Daten mit der modellbasierten Segmentierung ausgewertet. Nach der Nachbearbeitung beider Ergebnisse werden sie mittels modellbasierter Registrierung miteinander kombiniert. Dabei werden die T2-Aufnahmen an die T1-Aufnahmen angepaßt, da die T1-Aufnahmen eine deutlich präzisere Rekonstruktion der Patientenanatomie erlauben. In welcher Form die Segmentierungsergebnisse beider Datensätze miteinander kombiniert werden könnten, ist bei den aktuellen Forschungsarbeiten noch nicht geklärt worden.

Nach dieser Segmentierung der Patientenanatomie muß die Segmentierung noch mit den CT-Daten registriert werden. Hierzu werden wie für die modellbasierte Segmentierung im CT-Datensatz die sechs Landmarken an den korrespondierenden Stellen plaziert. Ob nun die übliche landmarkenbasierte Registrierung angewandt wird, bei der noch eine Skalierung des Referenzmodells in den drei Raumachsen möglich ist, oder ob eine rigide (affine) Transformation angewandt wird, bei der ausschließlich Translationen und Rotationen angewandt werden können, ist mit den Projektpartner noch zu besprechen. Beide Optionen sind mit der modellbasierten Registrierung möglich. Das Ergebnis dieser Registrierung ist bereits das fertige Bestrahlungsmodell.

Als verbleibender Schritt bleibt nun noch die Konvertierung der segmentierten und registrierten Datensätze in das vdx-Format, das durch die Bestrahlungssoftware VOXELPLAN eingelesen werden kann.

6.6.4 Bestrahlungsplanung

Entsprechend der bereits beschriebenen Schritte der Bestrahlungsplanung wird das zuvor berechnete Bestrahlungsmodell eingelesen. Anschließend werden verschiedene Bestrahlungssequenzen getestet und dahingehend optimiert, daß die absorbierte Strahlung innerhalb des Tumorgewebes maximiert wird, während sie außerhalb des Tumors minimiert wird.

Derzeit wird nicht zuletzt aufgrund der Tatsache, daß eine Bestrahlungsplanungssoftware zunächst entsprechend der Medizinischen Geräteverordnung eine umfangreiche und kostenintensive Zulassungsprozedur durchlaufen muß, noch mit der Bestrahlungsplanungssoftware EYEPLAN geplant. Die Bestrahlungspläne werden dabei durch die Mitarbeiter des Hahn-Meitner-Instituts erarbeitet und anschließend mit den Fachkräften der Ophthalmologischen Abteilung des Universitätsklinikums besprochen und gegebenenfalls modifiziert.

6.6.5 Bestrahlungsdurchführung

Zur Bestrahlungsdurchführung werden die Patienten zum Hahn-Meitner-Institut in Potsdam gebracht. Dort werden die Patienten in einem Stuhl fixiert, der den Kopf in einer definierten Position hält. Um zusätzlich das Auge in eine definierte Richtung auszurichten, wird ein kleines rotes Licht an eine vorgegebe-

ne Position bewegt. Der Patient wird gebeten, genau in dieses Licht zu schauen, wodurch die Orientierung des Auges stabilisiert wird. Gleichzeitig wird eine Videokamera frontal auf das Auge gerichtet und so eingestellt, daß ein Fadenkreuz in der Kamera genau die Pupille des Patienten zentriert aufnimmt. Zur Verifikation der korrekten Position des Auges werden von dem Patienten Röntgenaufnahmen angefertigt, die das Auge seitlich aufnehmen. In diesen Aufnahmen werden die Clips auf der Bulbusoberfläche lokalisiert. Die Positionen der Clips werden verglichen mit jenen, die entsprechend der Vorgaben durch den Bestrahlungsplan im Vorfeld berechnet wurden. Kommt es zu Abweichungen zwischen den Clip-Positionen in Röntgenaufnahmen und Bestrahlungsplan werden das rote Licht und die Kamera neu positioniert.

Die Bestrahlung wird durch zwei Einheiten freigeschaltet. Eine Person kontrolliert anhand der Kamera, ob der Patient die korrekte Position und Orientierung des Auges einhält. Ist dies gegeben, wird die Bestrahlung freigegeben. Gleichzeitig werden an einem Computer die genauen Bestrahlungsdaten eingegeben und anschließend vom Computer ausgeführt, solange die erste Person die Bestrahlung freigegeben hat. Neben der Beschleunigungsspannung werden zur Anpassung der Bestrahlungsparameter an die individuellen Anforderungen der Behandlung schnell rotierende Plexiglasplatten vor die Strahlenkanone montiert, die eine Tiefenmodulation der Protonenstrahlen erlauben.

6.6.6 Nachbehandlung

Um nach der Bestrahlungsdurchführung den Behandlungserfolg regelmäßig beobachten zu können, kommt der Patient regelmäßig in das Klinikum, wo zwei- und eindimensionale Ultraschallaufnahmen des betroffenen Auges angefertigt werden. Sie erlauben eine schnelle und zuverlässige Beobachtung der Entwicklung des Tumolvolumens. Dabei werden sowohl die Prominenz als auch die Größe der Tumorbasis beobachtet, um abzuschätzen, ob sich das Tumolvolumen verringert oder trotz der Behandlung weiter zunimmt. Statistiken über die Protonentherapie belegen jedoch die außerordentlich hohe Erfolgsquote dieses Therapieansatzes, so daß nur in Ausnahmefällen eine wiederholte Behandlung des Auges erforderlich ist.