Non-Uniform Deformable Volumetric Objects for Medical Organ Segmentation and Registration

vom Fachbereich Informatik
der Technischen Universität Darmstadt
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
von

Dipl.-Inform. Marius Erdt

geboren in Hamburg, Deutschland

Referenten der Arbeit:	Prof. Dr.-Ing. Georgios Sakas Technische Universität Darmstadt Prof. Dr. techn. Dieter W. Fellner Technische Universität Darmstadt Prof. Dr. med. Thomas J. Vogl Universität Frankfurt
Tag der Einreichung:	17.04.2012
Tag der mündlichen Prüfung:	13.06.2012

Erscheinungsjahr 2012

Darmstädter Dissertation
D 17

Erklärung zur Dissertation

Hiermit versichere ich die vorliegende Dissertation selbständig nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den 17.04.2012 Marius Erdt

Acknowledgement

At this point, I would like to thank the whole team of the Department of Cognitive Computing & Medical Imaging at Fraunhofer IGD for their great support. Special thanks goes to Prof. Dr.-Ing. Georgios Sakas who supervised this thesis and accompanied my scientific work during the last years. I would also like to thank Prof. Dr. techn. Dieter W. Fellner and Prof. Dr. med. Thomas J. Vogl for acting as examiners of this thesis. Furthermore, I would like to thank PD Dr. Arjan Kuijper for proofreading this thesis. Many thanks also go to the consortium of the project Theseus Medico in which I had the chance to conduct much of the scientific work that is part of this thesis.

Abstract

In medical imaging, large amounts of data are created during each patient examination, especially using 3-dimensional image acquisition techniques such as Computed Tomography. This data becomes more and more difficult to handle by humans without the aid of automated or semi-automated image processing means and analysis. Particularly, the manual segmentation of target structures in 3D image data is one of the most time consuming tasks for the physician in the context of using computerized medical applications. In addition, 3D image data increases the difficulty of mentally comparing two different images of the same structure. Robust automated organ segmentation and registration methods are therefore needed in order to fully utilize the potentials of modern medical imaging.

This thesis addresses the described issues by introducing a new model based method for automated segmentation and registration of organs in 3D Computed Tomography images. In order to be able to robustly segment organs in low contrast images, a volumetric model based approach is proposed that incorporates texture information from the model's interior during adaptation. It is generalizable and extendable such that it can be combined with statistical shape modeling methods and standard boundary detection approaches. In order to increase the robustness of the segmentation in cases where the shape of the target organ significantly deviates from the model, local elasticity constraints are proposed. They limit the flexibility of the model in areas where shape deviation is unlikely. This allows for a better segmentation of untrained shapes and improves the segmentation of organs with complex shape variation like the liver.

The model based methods are evaluated on the liver in the portal venous and arterial contrast phase, the bladder, the pancreas, and the kidneys. An average surface distance error between 0.5 mm and 2.0 mm is obtained for the tested structures which is in most cases close to the inter-observer variability between different humans segmenting the same structure. In the case of the pancreas, for the first time, an automatic segmentation from single phase contrast enhanced CT becomes feasible.

In the context of organ registration, the developed methods are applied to deformable registration of multi-phase contrast enhanced liver CT data. The method is integrated into a clinical demonstrator and is currently in use for testing in two clinics. The presented method for automatic deformable multi-phase registration has been quantitatively and qualitatively evaluated in the clinic. In nearly all tested cases, the registration quality is sufficient for clinical needs.

The result of this thesis is a new approach for automatic organ segmentation and registration that can be applied to various clinical problems. In many cases, it can be used to significantly reduce or even remove the amount of manual contour drawing. In the context of registration, the approach can be used to improve clinical diagnosis by overlaying different images of the same anatomical structure with higher quality than existing methods. The combination of proposed segmentation and registration therefore saves valuable clinician time in dealing with today's 3D medical imaging data.

Zusammenfassung

Motivation

In der modernen medizinischen Bildgebung werden heutzutage enorme Mengen an Bilddaten erzeugt. Dies betrifft vor allem die dreidimensionalen Bildgebungstechniken wie Computer-Tomographie, Magnetresonanz-Tomographie oder 3D-Ultraschall. Eine CT-Aufnahme besteht heutzutage aus mehreren Hundert zweidimensionalen Schichtbildern. Die Handhabung der stetig steigenden Datenmengen in der klinischen Praxis wird für den Arzt mehr und mehr zum Problem. Beispielsweise müssen in der Radiotherapie einzelne Zielorgane manuell in jedem Schichtbild zeitaufwändig konturiert werden, in dem die Organe sichtbar sind. Aus den Konturen wird anschließend ein Bestrahlungsplan erstellt, in dem nur die Zielstruktur einer hohen Strahlungsdosis ausgesetzt ist. Ein weiteres Beispiel ist die bildgestützte Operationsplanung von Teilresektionen der Leber. Hier muss die Leber manuell in jeder Einzelschicht markiert werden, um Schnittlinien zu planen sowie das verbleibende Lebervolumen nach dem Eingriff zu bestimmen.

Die beschriebenen Anwendungsbeispiele aus der klinischen Praxis bedeuten für den Arzt einen hohen Zeitaufwand. Die Konturierung von mehreren Organen in der Radiotherapie kann mehr als eine Stunde an Zeit beanspruchen [DMJ09]. Zusätzlich weisen manuelle Konturierungen derselben Struktur — durchgeführt von unterschiedlichen Individuen oder von derselben Person zu verschiedenen Zeitpunkten — stets signifikante Abweichungen auf [CBA*03,SWM*06, RST*11]. Es besteht daher ein hoher wissenschaftlicher und praktischer Bedarf an robusten und zuverlässigen automatischen oder semiautomatischen *Segmentierungsverfahren*, welche den Arzt bei Konturierungsaufgaben in der computerassistierten Diagnose oder bei der Operationsplanung unterstützen.

Neben der Betrachtung einer einzelnen Bildaufnahme oder eines Einzelvolumens spielen heute Zeitserienbilder in einer immer höheren Anzahl klinischer Anwendungen eine große Rolle. In Zeitserienaufnahmen werden von demselben Patienten mehrere Aufnahmen zu unterschiedlichen Zeitpunkten erstellt. Um beispielsweise einen Lebertumor zu klassifizieren, werden häufig diverse CT-Aufnahmen zu unterschiedlichen Zeitpunkten der Kontrastmittelanreicherung zur Bewertung herangezogen. Diese CT-Aufnahmen zeigen dieselben anatomischen Strukturen sowie den Tumor. Durch die unterschiedliche Verteilung des Kontrastmittels im Körper beziehungsweise im Lebergewebe während der Aufnahme, werden allerdings bestimmte Strukturen unterschiedlich dargestellt. Nur durch Betrachtung aller Aufnahmen kann schließlich der Tumor korrekt klassifiziert werden.

Für den Arzt bedeutet der Umgang mit Zeitserienbildern ein Problem, da er die einzelnen Bilder üblicherweise nebeneinandergestellt betrachtet. Das heißt, der Arzt muss die Strukturen mental in Überlagerung bringen und so z.B. abschätzen, wie weit ein Gefäß, welches zu einem bestimmten Zeitpunkt zu sehen ist, von einem Tumor entfernt ist, der zu einem anderen Zeitpunkt kontrastiert ist (siehe Abbildung A). Dies ist allerdings meist sehr ungenau, da der Patient zwischen den Aufnahmen atmet oder sich bewegen kann. Beim Vergleich von prä-operativen

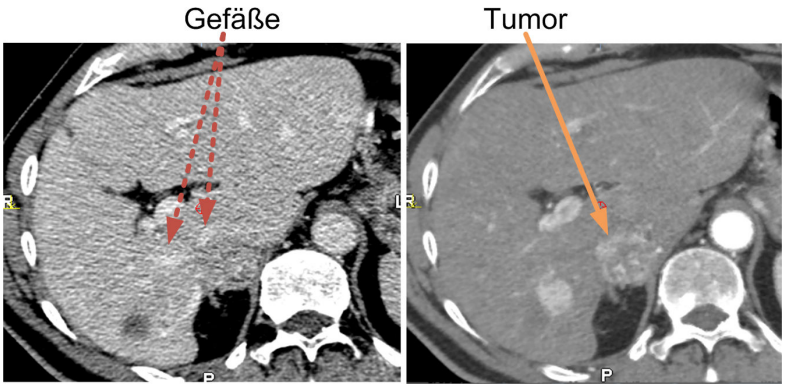


Abbildung A: CT-Bilder der Leber zu unterschiedlichen Zeitpunkten der Kontrastmittelaufnahme. In der linken Aufnahme sind zwei Gefäße zu erkennen. In der rechten Aufnahme ist ein Tumor sichtbar.

und post-operativen Bildern können zudem Veränderungen in der Anatomie hinzukommen — beispielsweise, wenn ein Tumor in der Behandlung entfernt wurde.

Um diesen Problemen zu begegnen, sind automatische *Registrierungsmethoden* notwendig, welche Bilder derselben anatomischen Struktur so deformieren, dass die Struktur in beiden Bildern in eine perfekte Überlagerung gelangt. Um nicht rigide Verformungen, wie sie etwa durch Atmung entstehen, zu korrigieren, müssen zudem lokal deformierende Registrierungsmethoden verwendet werden. In den überlagerten Bildern kann dann genau ausgemessen werden, wie weit die Zielstruktur von wichtigen Gefäßen entfernt ist.

Beide beschriebenen Probleme — die automatisierte Segmentierung und die Registrierung von anatomischen Strukturen in medizinischen Bilddaten — sind eng miteinander verknüpft, da Segmentierungsmethoden verwendet werden können, um den Registrierungsprozess zu unterstützen. Hierfür sind insbesondere modellbasierte Verfahren geeignet. Diese deformieren eine Referenzform einer anatomischen Struktur so, dass die Form mit der Struktur in Überlagerung kommt. Auf diese Weise wird eine Segmentierung der Struktur erzielt. Wird dieselbe Form an zwei Bilddatensätze angepasst, kann zwischen beiden Bildern eine Transformation der Koordinatensysteme berechnet werden. Dieser Vorgang stellt eine Registrierung dar.

Modellbasierte Ansätze haben sich in der Segmentierung von zahlreichen Organen, wie der Leber [HvGSea09, HM09, KLL07, WSH09] oder dem Herz [ZBG*07, EPS*08], bewährt. Modellbasierte Ansätze haben daher das Potenzial, die beschriebenen Herausforderungen sowohl im Bereich der Segmentierung als auch der Registrierung zu lösen.

Existierende modellbasierte Verfahren weisen jedoch Nachteile auf, die eine breite Anwendung im klinischen Alltag erschweren. So kann eine Transformation zwischen einzelnen Bildern nur für die bekannten Modellpunkte errechnet werden. Um ebenfalls eine hinreichende Registrierung in Bereichen außerhalb der Modellpunkte zu erreichen, ist die Entwicklung von weitergehenden Verfahren notwendig.

Ein weiterer Nachteil von modellbasierten Verfahren betrifft die Formbeschränkung des anzupassenden Modells. Das Ziel hierbei ist, eine genaue Anpassung an die Zielstruktur zu erreichen, ohne dass unplausible Formen entstehen. Im Fall komplexer Formvarianzen, wie sie etwa bei

der Leber zwischen Individuen auftreten [HM09, HvGSea09], werden häufig statistische Formmodelle verwendet [CHTH93, CTCG95]. Statistische Formmodelle modellieren die Formvarianz einer Struktur durch eine statistische Analyse von Beispielformen. Diese Beispielformen werden üblicherweise erzeugt, indem die Struktur in einigen repräsentativen Datensätzen manuell segmentiert wird. Das zu verformende Modell kann dann unter Berücksichtigung der modellierten Formvarianz so beschränkt werden, dass sich nur plausible Formen ergeben. Einige Organe, wie die Leber, weisen allerdings eine so hohe Formvarianz auf, dass üblicherweise nicht genügend Beispielformen erzeugt werden können, um eine genaue Anpassung des Modells an die Daten zu gewährleisten. Durch die verwendete Dimensionsreduktion der Daten verliert das Modell zudem an Genauigkeit, was eine exakte Anpassung, selbst unter Verwendung einer hohen Anzahl an Formen, verhindert. Abbildung B zeigt einige exemplarische Leberformen von unterschiedlichen Individuen[1]. Da statistische Formmodelle nicht alle Details modellieren können, wird die Formbeschränkung normalerweise zu einem gewissen Zeitpunkt gelockert oder aufgegeben [HM09]. Der Nachteil dieser Methode ist, dass die Form des Modells unplausibel werden kann, je länger die Formbeschränkung aufgehoben wird. Es ist daher notwendig, die bestehenden modellbasierten Verfahren weiterzuentwickeln, um eine robuste Segmentierung auch bei Organen mit hoher Formvarianz zu erreichen.

Ein zusätzlicher Nachteil modellbasierter Verfahren betrifft die geometrische Repräsentation der zu segmentierenden Strukturen. Die große Mehrzahl an Ansätzen verwendet oberflächenbasierte Modelle. Das heißt, es wird ein dreidimensionales Oberflächenmodell an die Zielstruktur in den Bilddaten angepasst. Es macht allerdings oft Sinn, volumetrische Modelle zu verwenden, welche über eine reine Anpassung der Modellgrenzen hinausgehen. Volumetrische Modelle berücksichtigen, neben den Bilddaten an den Modellgrenzen, zusätzlich die Bilddaten im Modellinneren und können so beispielsweise Einschlüsse organfremden Materials erkennen. In der Literatur werden verschiedene volumetrische Verfahren vorgeschlagen [TFCT98, PFJ*03, HM08, BPCO10, SLH11]. Diese Verfahren sind allerdings häufig komplex und daher rechenintensiv. Zudem können Standardverfahren zur Detektion von Organgrenzen oder etablierte Verfahren zur Formbeschränkung oft nicht mit diesen Ansätzen kombiniert werden. Es besteht daher ein Bedarf an der Entwicklung effizienter und generalisierbarer volumetrischer Verfahren.

Beiträge

In dieser Arbeit werden verschiedene Verfahren entwickelt, um die beschriebenen Probleme modellbasierter Ansätze zu adressieren und eine robuste Segmentierung und Registrierung von Organen in CT-Aufnahmen zu erreichen. Im Folgenden werden die wichtigsten Beiträge dieser Arbeit zusammengefasst.

Entwicklung einer neuen Klassifizierung für Segmentierungs- und Registrierungsverfahren

Im Rahmen dieser Arbeit wird ein neues Klassifikationsschema für Segmentierungs- und Registrierungsverfahren vorgestellt. In dem vorgestellten Schema werden Methoden in ein Kontinuum zwischen zwei Extremen eingeordnet: rein bildbasierten Verfahren und Verfahren, die detailliertes, globales Formwissen über die zu segmentierende Struktur voraussetzen. Durch Anwendung des Schemas auf bestehende Methoden wird gezeigt, dass Methoden mit globalem

[1] Die Datensätze stammen aus der 3D-IRCADb-01 Datenbank (http://www.ircad.fr).

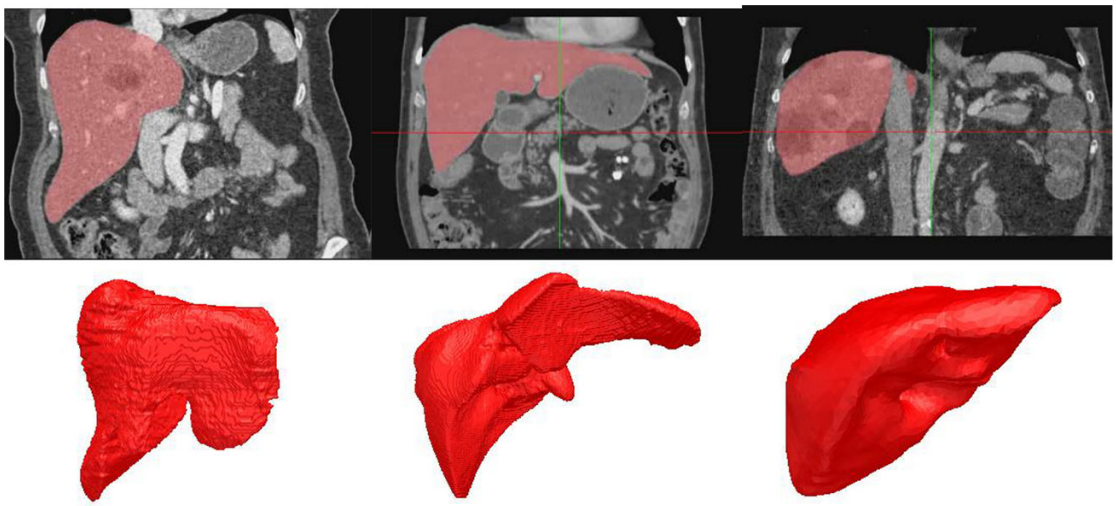


Abbildung B: Dargestellt sind Leberformen unterschiedlicher Individuen. Die Modellierung der Leber mit statistischen Formmodellen stößt durch die hohe anatomische Formvarianz an ihre Grenzen.

Vorwissen sowohl zur Segmentierung als auch zur Registrierung verwendet werden können. Im Gegensatz zu bestehenden Klassifikationsschemata kann so direkt die Menge an Methoden ermittelt werden, die sich zum Erreichen der in dieser Arbeit definierten Ziele anbieten.

Entwicklung einer Methode zur lokalen Formbeschränkung von deformierbaren Modellen

Ein wesentlicher Beitrag dieser Arbeit ist die Weiterentwicklung von deformierbaren Modellen zur Segmentierung medizinischer Bilddaten durch ein Verfahren zur lokalen Formbeschränkung. Hierbei wird die Formanpassung eines Modells lokal definiert, indem dem Modell an jedem Oberflächenpunkt unterschiedliche Steifheitswerte zugeordnet werden. Die Steifheitswerte werden durch Analyse von Beispielformen der jeweiligen Struktur bestimmt sowie durch Berücksichtigung der Krümmung der Form im Datensatz während der Anpassung. Abbildung C(a) zeigt ein solches formbeschränktes Modell der Leber.

Entwicklung von Verfahren zur automatischen modellbasierten Segmentierung

Das vorgestellte Modell wird in einer neuartigen Segmentierungsmethode zur Segmentierung von Organen in CT-Aufnahmen verwendet. Die Freiheitsgrade, welche das Modell während der Anpassung besitzt, werden hier Schritt für Schritt erhöht, um eine robuste Anpassung des Modells auch bei nicht optimaler initialer Positionierung zu erlauben. Im Unterschied zu bestehenden Anpassungsverfahren kann das Modell auf diese Weise mehrere Zentimeter von der optimalen Position entfernt initialisiert werden, ohne dass sich eine Verschlechterung der Segmentierungsgenauigkeit ergibt.

Zudem wird ein Verfahren zur Verbesserung bestehender Methoden zur Organgrenzen-Detektion entwickelt. Das Verfahren basiert auf der Entfernung von detektierten Punkten aus dem Anpassungsprozess, von denen es unwahrscheinlich ist, dass sie auf den Organgrenzen liegen.

Die Wahrscheinlichkeit eines Punktes auf einer Organgrenze zu liegen, wird dabei durch eine histogrammbasierte Methode bestimmt. Das entwickelte Verfahren lässt sich mit bestehenden Standardmethoden zur Organgrenzen-Detektion kombinieren.

Zur initialen Positionierung der Formmodelle in den Bilddaten wird ein zweidimensionaler histogrammbasierter Ansatz auf dreidimensionale medizinische Bilddaten erweitert und durch Intensitätsmerkmale ergänzt. Dadurch können grobe rechteckige Ausschnitte detektiert werden, in denen sich die Organe befinden. Im Fall von fehlerhaften Ausschnitten werden anschließend lernbasierte Verfahren benutzt, um jedem Organmodell genau eine initiale Position zuzuordnen.

Die entwickelten Verfahren zur Organsegmentierung werden anhand von CT-Aufnahmen der Leber, der Nieren und des Pankreas evaluiert. Die durchschnittlichen Oberflächenabweichungen zu manuell erstellten Segmentierungen betragen zwischen 0.5 mm und 1.7 mm. Die Abweichungen sind daher in den meisten Fällen mit den Varianzen zwischen zwei manuell erstellten Segmentierungen unterschiedlicher Individuen vergleichbar.

Entwicklung eines volumetrischen deformierbaren Modells

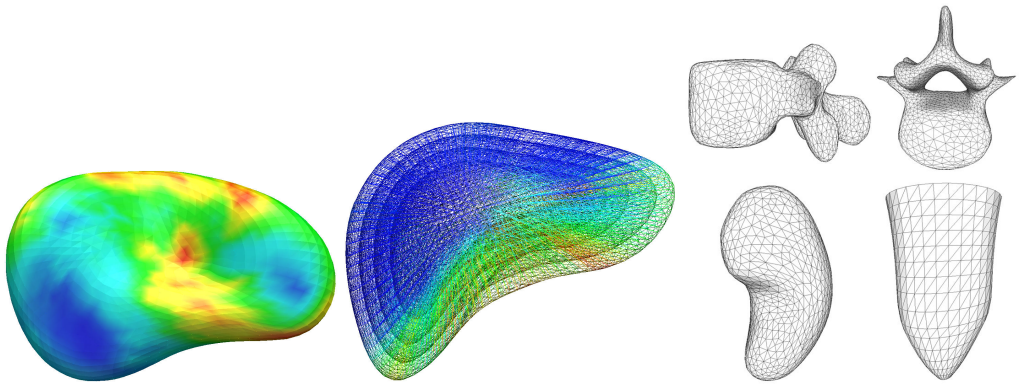
Im Rahmen dieser Arbeit wird ein Verfahren zur modellbasierten Segmentierung von Organen vorgestellt. Dabei wird ein neuartiges Volumenmodell verwendet. Das Modell besteht aus mehreren identischen, miteinander verbundenen Schichten. Durch Verwendung des Modells lassen sich Bildinformationen aus dem Inneren des Modells während der Anpassung berücksichtigen. Das vorgestellte Modell lässt sich zudem mit der im vorherigen Abschnitt beschriebenen lokalen Formbeschränkung kombinieren. Abbildung C(b) zeigt ein entsprechendes Modell der Leber. Durch die Verwendung des Modells lässt sich organfremdes Material im Inneren erkennen. Daraufhin können Form und Positionierung des Modells verändert werden. Auf diese Weise wird die Genauigkeit der Segmentierung vor allem in Datensätzen mit schlechtem Bildkontrast erhöht.

Im Vergleich zu bestehenden volumetrischen Methoden lässt sich das vorgestellte Modell mit derselben Aufwandsklasse wie ein punktbasiertes Oberflächenmodell optimieren. Durch Reduzierung der Schichtanzahl kann darüber hinaus eine weitere Beschleunigung erfolgen. Ein weiterer Vorteil der Methode ist, dass sich das Modell mit statistischen Formmodellen kombinieren lässt, was bei bestehenden Verfahren nicht möglich ist. Dadurch können auch Organe mit hoher Formvarianz segmentiert werden. Außerdem können Standardverfahren zur Detektion von Organgrenzen verwendet werden.

Das präsentierte Modell wird zur Segmentierung der Leber in der arteriellen Phase kontrastmittelverstärkten CTs sowie der Blase in nicht kontrastierten CT-Aufnahmen eingesetzt. Die durchschnittliche Oberflächendistanz zu manuell erstellten Segmentierungen liegt in den durchgeführten Tests bei 2 mm für die Leber und 1.3 mm für die Blase. Das vorgestellte volumetrische Modell hat sich zudem als robuster gegenüber Fehlpositionierungen erwiesen als Oberflächenmodelle.

Entwicklung eines Verfahrens zur Erzeugung von Formen mit Punktkorrespondenzen

Ein weiterer Beitrag dieser Arbeit ist die Entwicklung eines Verfahrens zur Erzeugung von Trainingsformen mit Punktkorrespondenzen für die Erstellung statistischer Formmodelle. Dabei werden Oberflächenmodelle eines Organs vom Benutzer an die Bilddaten angepasst. Das Ober-



(a) (b) (c)

Abbildung C: (a) Lokal formbeschränktes Lebermodell. Regionen hoher und niedriger Elastizität sind rot beziehungsweise blau eingefärbt [ESW10]. (b) Mehrschichtiges Formmodell der Leber. (c) Beispiele der erzeugten Modelle: Rückenwirbel (oben), linke Niere (links unten) und linker Herzventrikel (rechts unten) [EKW09].

flächenmodell wird dabei ständig optimiert, sodass die Modellpunkte jeweils korrespondierende Strukturen kennzeichnen. Die entstehenden Formen könnnen dann direkt zur Erzeugung statistischer Formmodelle verwendet werden.

Durch das vorgestellte Verfahren entfällt die zeitaufwändige Korrespondenzsuche zwischen Trainingsformen, die in bestehenden Verfahren notwendig ist. Diese Korrespondenzsuche kann je nach Verfahren mehrere Tage in Anspruch nehmen. Mit der beschriebenen Methode können darüber hinaus Modelle unterschiedlicher Topologie erstellt werden. Bestehende Verfahren sind hier meist auf Flächen des Geschlechts 0 beschränkt. Dies ist beispielsweise bei komplexen Strukturen wie den Rückenwirbeln nicht ausreichend. Durch das vorgestellte Verfahren zur Erstellung von Trainingsformen lässt sich die Erzeugung statistischer Formmodelle erheblich vereinfachen, da sich ein bestehendes Formmodell ohne aufwändige Neuoptimierung aller Korrespondenzen erweitern lässt.

Für die Evaluation des beschriebenen Verfahrens werden unterschiedliche Formmodelle verwendet. Es werden Modelle des linken Herzventrikels, der linken Niere sowie eines Rückenwirbels erzeugt (siehe Abbildung C(c)). Hierbei wird jeweils ein Modell mit der beschriebenen Methode erzeugt. Außerdem werden zum Vergleich für jede Struktur jeweils zwei Modelle mit Standardmethoden zur Korrespondenzoptimierung erzeugt. Die Qualität der Korrespondenzen wird sowohl qualitativ durch eine Visualisierung markanter anatomischer Merkmale evaluiert als auch durch eine quantitative Analyse unter Verwendung von Standardmaßen zur Beurteilung statistischer Formmodelle. Die Qualität der erzeugten Modelle liegt dabei meist über der Qualität der Modelle, die mit den Vergleichsverfahren erzeugt wurden.

Entwicklung einer Methode zur deformierbaren Registrierung von Leber-CT-Aufnahmen

Die in den vorigen Abschnitten beschriebenen modellbasierten Verfahren werden verwendet, um Leber-CT-Aufnahmen miteinander zu registrieren. Bei Untersuchungen von Lebertumoren werden üblicherweise mehrere Aufnahmen zu unterschiedlichen Zeitpunkten der Kontrastmittel-

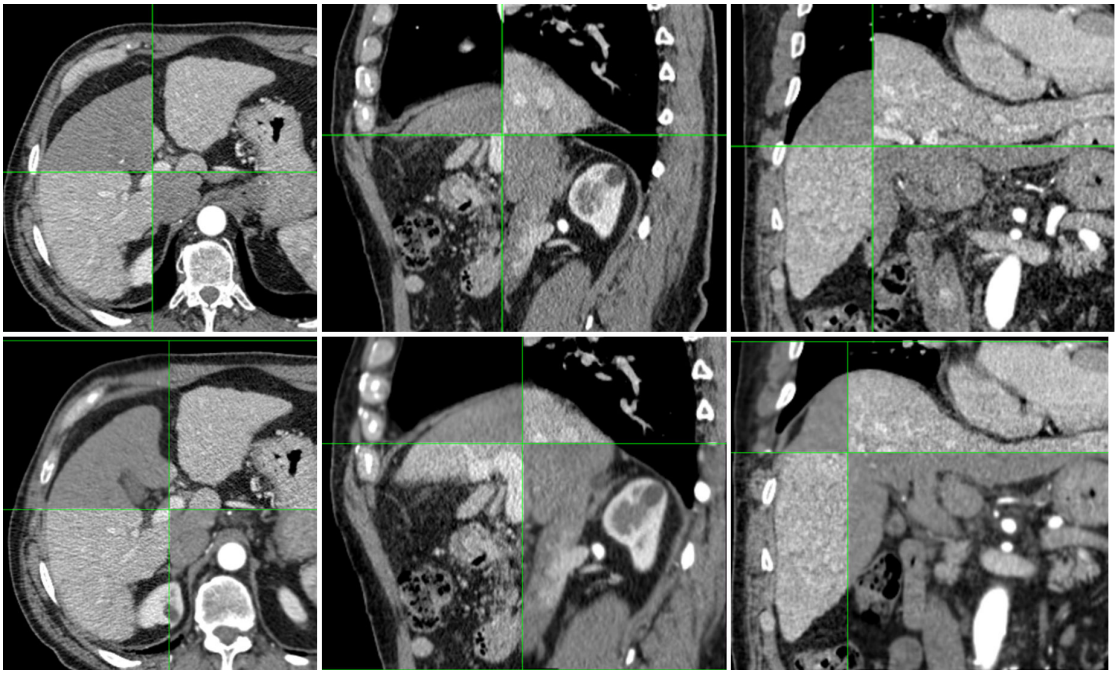


Abbildung D: Registrierung von kontrastmittelverstärktem CT der arteriellen Leberphase mit der portalvenösen Phase. Die obere Reihe zeigt den Zustand vor der Registrierung und die untere Reihe den Zustand nach der Registrierung. Beide Phasen sind hier jeweils wechselseitig dargestellt.

aufnahme erstellt. Die Überlagerung solcher Bilddaten kann helfen, einen klinischen Eingriff genauer als bisher möglich zu planen.

Das entwickelte Verfahren basiert auf der Extraktion der Leberformen aus der arteriellen Kontrastmittelphase sowie der portalvenösen Kontrastmittelphase. Hierbei wird für die Segmentierung der Leber in der portalvenösen Phase das vorgestellte lokal formbeschränkte Oberflächenmodell verwendet. Die Segmentierung der Leber in der arteriellen Phase wird durch Anpassung des vorgestellten volumetrischen Modells erreicht, da der Kontrast an den Organgrenzen oft nicht für eine hinreichend genaue Segmentierung mit Oberflächenmodellen ausreicht. Die Leberformen werden anschließend miteinander rigide registriert, um die CT-Bilder grob miteinander zu überlagern. Eine exakte Überlagerung findet anschließend durch eine neue Registrierungsmethode statt, welche sowohl Bildinformationen als auch Informationen über die Form der Leber in beiden Bildern einbezieht. Zunächst findet hier eine Anpassung nur an diejenigen Organgrenzen statt, die in beiden Phasen sichtbar sind. Die an den Rändern entstehenden Verformungskräfte werden dann in das Innere der Leber propagiert.

Abbildung D zeigt beispielhaft das Ergebnis einer Registrierung, wobei hier der Zustand vor und nach der Registrierung dargestellt ist. Zur Veranschaulichung der Registrierungsqualität sind beide Datensätze jeweils wechselseitig dargestellt. Das vorgestellte Verfahren ist von Medizinern evaluiert. In nahezu allen Fällen wird die Qualität der Registrierung als ausreichend für die klinische Verwendung bewertet.

Diskussion

In der vorliegenden Arbeit wird ein neues Verfahren zur Segmentierung und Registrierung von Organen in Aufnahmen der Computertomographie vorgestellt. Dabei wird zunächst ein neuartiges Klassifikationschema zur Einordnung von Segmentierungs- und Registrierungsmethoden entwickelt. Hierauf aufbauend, werden modellbasierte Verfahren ausgewählt und weiterentwickelt, um die Ziele dieser Arbeit zu erreichen. Hauptbeiträge sind die Entwicklung eines neuartigen volumetrischen Formmodells sowie die Entwicklung einer Methode zur lokalen Formbeschränkung von punktbasierten Formmodellen. Durch die vorgestellten Verfahren lassen sich zwei wichtige Probleme modellbasierter Segmentierung und Registrierung lösen: die robuste Segmentierung und Registrierung schwach kontrastierter Strukturen in CT-Aufnahmen sowie eine robuste Anpassung an komplexe Formen, die von der Menge gelernter Beispielformen abweichen.

Die entwickelten Methoden werden anhand wichtiger klinischer Fragestellungen evaluiert. Im Kontext der Segmentierung von Organen kann eine Genauigkeit erreicht werden, welche in den meisten Fällen ausreicht, um einen Großteil der manuellen Konturierung zu ersetzen. Eine Anwendung der Verfahren in der klinischen Praxis stellt für den Arzt eine Minderung des Zeitaufwandes für die Konturierung dar. Dies wiederum stellt in Aussicht, dass in Zukunft mehr Patienten von den Fortschritten in der medizinischen Bildgebung und in computergestützten Applikationen profitieren können.

Im Rahmen der Registrierung von Organen können Mehrphasen-CT-Aufnahmen der Leber robust miteinander registriert werden. Durch die präzise Überlagerung der Aufnahmen wird dem Arzt der kognitiv hochkomplexe Vergleich von wechselseitig sichtbaren Strukturen abgenommen. Dies stellt eine Verbesserung der Diagnose in der klinischen Praxis und somit eine patientenspezifischere Behandlung in Aussicht. Darüber hinaus können bestehende Planungssysteme durch den Einbezug von Mehrphasenplanungsdaten ergänzt und verbessert werden.

Contents

1. Introduction

1.1. Motivation

In medical imaging, huge amounts of data are created during each patient examination, especially using 3-dimensional or 2+1-dimensional image acquisition techniques such as Computed Tomography (CT), Magnetic Resonance Tomography (MR) or 3D ultrasound. For example, a single CT scan nowadays often comprises of several hundreds of 2-dimensional images. The increasing amount of data becomes more and more difficult to handle by humans without the aid of automated or semi-automated image processing means and analysis. As an example, in radiotherapy, the boundaries of the target organ must be manually outlined in each image where this organ appears. This is needed in order to create a safe treatment plan, where only the target organ is exposed to a high radiation dose. Another example is liver resection planning, where the liver boundaries must be outlined in order to calculate cutting lines and remaining organ volume after resection.

The mentioned work is very time-consuming for humans. Drawing contours of multiple organs in radiation therapy can take more than one hour [DMJ09]. Additionally, often a non trivial inter- and intra-observer variability rate exists. That means, that different individuals or even the same person may outline the boundaries of a structure differently, depending on the point in time or the level of experience [CBA*03, SWM*06, RST*11]. Therefore, there exists a strong scientific and practical need for robust and reliable automated and semi-automated *segmentation techniques* that support the human during diagnosis, treatment or operation planning.

Apart from examining a single data set, a second challenge in nowadays imaging is the handling of images taken from different points in time which becomes more and more important in all kinds of clinical applications. In this kind of imaging, several data sets of the same patient showing the same anatomical structure are created. For example, in order to stage liver tumors, often several images acquired at different stages of contrast agent saturation are used. These images show the same anatomical structure, i.e. the tumor, but usually provide complementary information due to the effect of the contrast agent on the physical imaging process.

The problem in handling these time series data is that the human has to mentally fuse the image information when he examines the different images. For example, he needs to estimate how far an important blood supply vessel shown in the portal venous contrast CT phase is away from a tumor which is visible in the arterial contrast CT phase scan. Changes in the patient's position, respiration state or anatomy between the scans (e.g. between pre- and post-treatment) can make this task very difficult or even impossible to perform correctly. Therefore, *registration methods* are needed that automatically deform corresponding scans such that the structure of interest is perfectly overlapping in all scans. In particular, deformable registration is needed to cope with the non-linear changes, e.g. caused by respiration.

The two described challenges — automatic segmentation and registration of anatomical structures in medical images — are closely related, because segmentation methods can be used to

guide registration. In particular, *model based segmentation* approaches deform a template of the structure to segment such that it fits to the image data. Once the template is adapted to the same structure in two data sets, a correspondence, i.e. a registration, between the structure in both data sets is already given by correspondence of the templates.

Model based approaches have also proven to be a robust technique for the segmentation of solid organs such as the liver [HvGSea09, HM09, KLL07, WSH09] or the heart [ZBG*07, EPS*08]. Therefore, modal based approaches have the potential to address the described challenges in both segmentation and registration.

However, open issues in model based approaches exist. Regarding registration, a correspondence between images is only given at points where template information is available, for example given by corresponding template surface points. In order to deform the images in regions without template information, further registration or extrapolation strategies are required.

Another issue of model based segmentation approaches concerns the shape restriction of the templates, i.e. the shape preserving strategy to avoid a non plausible deformation of the templates during adaptation. Usually statistical shape information as introduced by Cootes et al. [CHTH93, CTCG95] is used to restrict the template shape deformation in current approaches for segmentation of abdominal organs such as the liver [HM09, HvGSea09]. The idea of statistical shape models is that the typical shape of a structure can be statistically modeled using a set of so called training shapes which are usually example shape samples of the structure to segment. For example, a human may manually draw a contour of an organ in a couple of images to create those shape samples. Using the statistical shape information from the training shapes the template's shape can be restricted to be similar to the training shapes. In this manner, a non plausible deformation of the templates is avoided. However, the problem is that the shape restriction of statistical shape models is too restrictive for organs that show a high shape variability such as the liver. While it is expected that the generalizability of the templates increases with the amount of training shapes, it is not possible to model all patient specific details. For example, the amount of fat and surrounding tissue, the respiration state, treatment procedure and of course the disease itself have an impact on the shape of an organ (cf. Figure 1.1 for an exemplary overview of liver shapes from different individuals[1]). Statistical shape models are not able to model such detail. Therefore, the statistical shape restriction is usually relaxed at some point during deformation of the templates [HM09]. The drawback of this procedure and an unsolved problem so far is that the template shape may become non plausible, thus losing the information from the learned shape space. Therefore, techniques are needed that address this issue of model based segmentation in order to provide reliable tools for the segmentation challenges described above.

Another issue of model based segmentation concerns the templates of the structures to segment. The majority of current state of the art model based methods use templates based on surface representations, i.e. a 3-dimensional surface model is evolved towards the boundaries of the target organ or structure. However, it often makes sense to incorporate *volume information* during the adaptation process in order to make the method more robust against model initialization errors and to increase the robustness through detection of organ foreign tissue that may not be detected when considering image information near the model's surface only. Different kinds of volumetric segmentation approaches have been proposed in the literature [TFCT98, PFJ*03, HM08, BPCO10, SLH11]. However, those methods are often complex and computationally expensive. Moreover, existing standard adaptation and shape preserving strategies cannot be used within those models or must be adapted. Thus there is a need for

[1] Data sets are taken from the public 3D-IRCADb-01 database (http://www.ircad.fr).

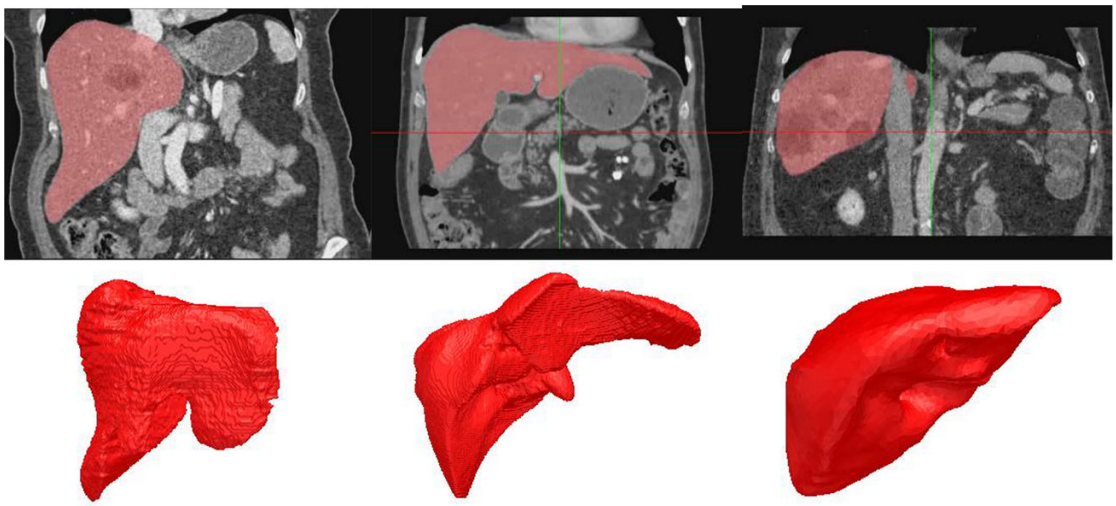


Figure 1.1.: Liver shapes from different individuals: Some parts of the organ may look completely different between individuals which makes it difficult to precisely model the liver using statistical shape methods.

computationally inexpensive 3-dimensional volumetric segmentation models that are easily combinable with said standard approaches.

1.2. Objectives of the thesis

The main objective of this thesis is the development of a 3-dimensional volumetric approach for segmentation and registration of organs in 3-dimensional medical images. The approach to be developed should address the open issues of current model based methods described in Section 1.1. In particular, the approach should be principally applicable for the segmentation of various organs. Since different human organs substantially differ in appearance, shape and tissue characteristics it is expected that adaptations for specific organs have to be made. Therefore, it is important that the approach allows that different existing standard adaptation and shape preserving strategies can be incorporated. For example, it should be possible to incorporate standard boundary detection methods based on trained gray level profiles [CT93] as well as different statistical shape model methods [DTT08].

In order to address the problem of model based segmentation regarding shape restriction and relaxation as described in Section 1.1, the approach should be flexible and robust in the sense that the approach adapts well to unseen data without showing much leakage into organ foreign areas.

Furthermore, the method to be developed should be an approach that incorporates volumetric information of the organ rather than only boundary information. That means, the method should incorporate information from the organ's interior during the adaptation process. This should lead to an increase of the robustness of the segmentation procedure regarding errors in model initialization. Additionally, it should help to detect organ foreign tissue inside the model template in order to improve the quality of the segmentation outcome. Current state of the art volumetric segmentation approaches are usually complex and therefore computationally expensive in comparison to surface based segmentation methods. It is expected that due to the increase of

information incorporated in a volumetric approach, the computation costs of the method to be developed will also be higher than a standard surface based approach. However, the class of complexity of the algorithm should be the same as in a standard surface based approach.

1.3. Main contributions

The main contributions of this thesis as presented in Chapter 3 and Chapter 4 are:

- the development of a new approach for building 3-dimensional shape templates with corresponding landmarks. The method combines manual segmentation and landmark correspondence establishment in a single approach. The built templates can be directly used for statistical shape model building without the need for applying time consuming optimization strategies [EKW09, EKW10]. The method has been applied to different organs such as the kidneys [ES10] and vertebra [WKEK10, WEKK11b, WEKK11a]
- the development of a new volumetric model based method for medical organ segmentation and registration — non-uniform deformable volumetric objects. The elasticity of the model is defined locally which makes it more flexible and robust in comparison to standard deformable models [ETS09, ES10, ESKW10]. The volumetric model extends existing parametric surface deformable models by using a layer based geometry [ESW10]. This makes the proposed model computationally fast to optimize and easily combinable with standard deformable model boundary search strategies.
- the development of a new generic segmentation method based on the developed model. The approach uses a rule based outlier removal which can be combined with any probabilistic boundary detection method. The outlier removal boosts the performance of simple boundary detection approaches to be on a par with complex learning based methods. A multi-tiered adaptation process with increasing degrees of freedom during adaptation is proposed that is shown to be more robust to model initialization errors than standard statistical shape model methods. The method is evaluated for various clinically important organs including liver [DOLCE10, ESKW10], kidneys [ES10], bladder and pancreas [EKD*11].
- the development of a new registration method for registration of liver CT scans from different phases of contrast agent saturation as well as for registration of pre- and post-treatment CT scans of the portal-venous phase based on non-uniform deformable volumetric objects [ESH*11, EOLD*12].

1.4. Structure of the thesis

This section lists the structures of this thesis sorted by chapters. The main topics are briefly described and put into the context of the thesis.

- Chapter 2 describes the relevant medical fundamentals of this work and presents a new classification scheme for segmentation and registration methods. Based on the scheme the current state of the art of segmentation and registration techniques of organs is classified and put into context. Suitable methods for organ segmentation and registration are selected and discussed in relation to the objectives listed in Chapter 1.2.
- Chapter 3 presents a new volumetric model based method for medical organ segmentation and registration — non-uniform deformable volumetric objects — in order to address the problem statements defined in Section 1.1 and Section 1.2.

- Chapter 4 presents an organ registration framework for registration of multi-phase contrast enhanced liver CT scans. Non-uniform deformable volumetric objects are used to guide a deformable Demons based registration in order to overlay CT scans of different phases of contrast agent saturation for clinical application.
- In Chapter 5, the methods proposed in this thesis are evaluated.
- Chapter 6 concludes the thesis by discussing and classifying the results. It also gives an overview of open problems as well as an outlook on how to address these problems in future work.

2. State of the art and related work

The main task of this thesis is the development of an automatic 3-dimensional volumetric approach for organ segmentation and registration in CT images. Various kinds of segmentation and registration approaches exist and are subject to ongoing research. In many cases, segmentation techniques are used for or in combination with registration methods and vice versa. This makes it difficult to distinguish between those methods and to classify them correctly in order to choose the right technique for the given problem task. Furthermore, the nomenclature in the literature is often not consistent or sometimes even misleading. For example, atlas based segmentation is used for segmenting objects but practically it is a registration method. The main goals of this chapter are twofold. Firstly, a proper classification scheme for existing segmentation and registration methods is found in order to be able to choose appropriate techniques for the problem task of this thesis. Secondly, based on the scheme, current state of the art methods in segmentation and registration will be described and classified. This chapter ends with a discussion about the strengths and drawbacks of these methods regarding their applicability for organ segmentation and registration.

2.1. Medical fundamentals

This section describes the medical fundamentals relevant for this work. First, the principle of *Computed Tomography* (CT) is outlined followed by a description of contrast agent enhancement techniques that are used in medical practice to pronounce certain structures like organs or vessels in order to allow for a better visual discrimination.

2.1.1. Computed Tomography

The principle of Computed Tomography is based on the estimation of the density of an object by sending X-rays from different directions through the object and measuring the attenuated radiation behind the object using a detector device. The measurements are then used to mathematically reconstruct density images that represent slice views of the object. Usually a stack of equally aligned slices is reconstructed in a top-down manner in order to build a 3-dimensional volume representation of the object.

Figure 2.1 shows the principle of an X-ray acquisition. A radiation source emits X-rays that are focused by a collimator. The X-rays penetrate the body and are partially absorbed by the tissue. At the opposite side, a detector measures the attenuated radiation. Source and detector are now translated in order to send several parallel rays which build a profile of attenuated intensity at the detector. Afterwards, a projection of the emitted radiation is computed as the logarithmic function of the relation between the non-attenuated intensity at the borders and the attenuated intensity after penetrating the body. This process is repeated at different angles of source and detector until the projection is computed at every position around the body. The smaller the angle intervals are chosen the better the reconstruction of the image slice gets.

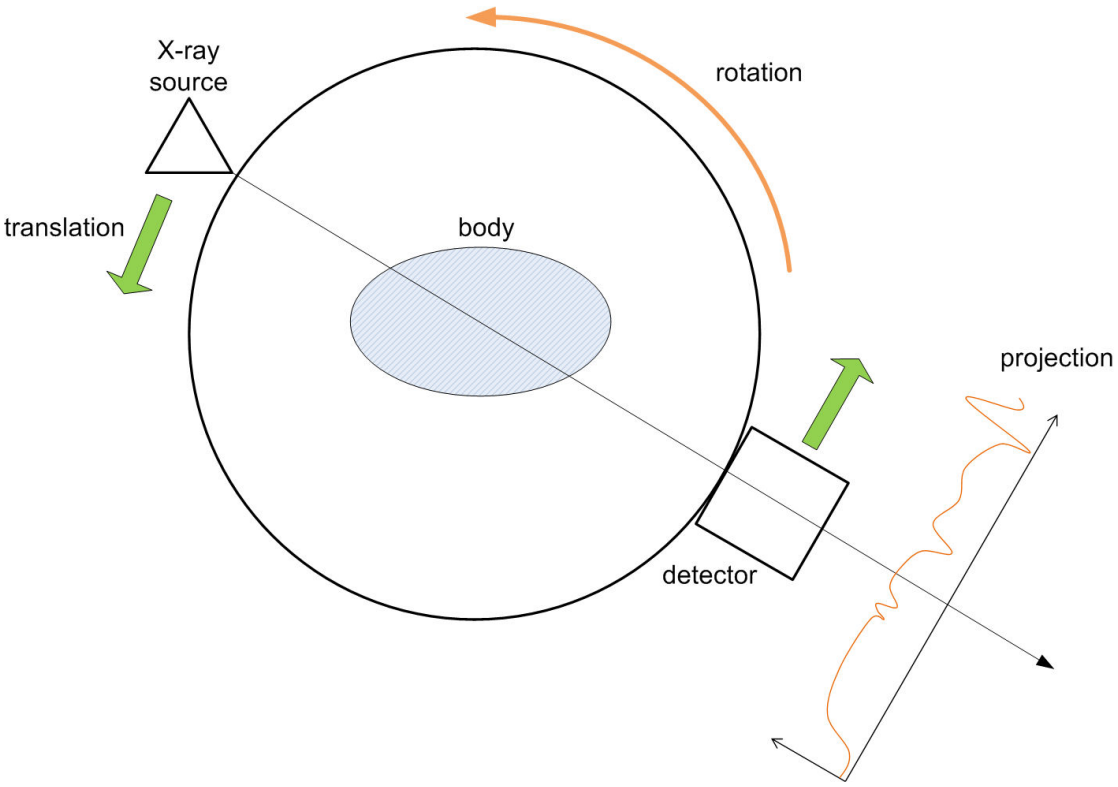


Figure 2.1.: Principle of CT image acquisition.

The projections from the different view points are now back-projected in order to compute the density function of the body area that has been penetrated by the X-rays. Usually filtered back-projection is used for that task. The resulting spatial distribution of the radiation attenuation is not displayed directly, since its amount depends on the used radiation energy. In order to compare images from different CT scanners, the attenuation is displayed relative to the attenuation of water in the so called Hounsfield Unit (HU) system.

Figure 2.2 shows the Hounsfield scale which maps the attenuation of human tissue to HU. By definition, water has a value of 0 HU. Air gets a value of -1000 HU. These two values define the fix points of the Hounsfield scale. All other attenuation values are linearly mapped to this scale. Using the Hounsfield scale, images that have been reconstructed using different radiation energies can be compared. This also eases automated image processing of such images since certain structures and organs are always located in the same intensity range. It must be noted though that the Hounsfield scale often only allows for a rough classification of organ tissue. This is mainly due to different tissue and anatomy characteristics between human individuals.

As can be seen in Figure 2.2, many organs share at least partially a common intensity interval. This problem can be alleviated by using contrast agents that are introduced into the metabolism in order to shift the intensity interval of a target structure upwards in the Hounsfield scale which allows for a better discrimination of the structure. Section 2.1.2 outlines the use of contrast agent in CT imaging for clinical diagnosis.

The Hounsfield scale is usually discretized to 12 Bit or 4096 gray values and ranges from -1024 to 3071 HU. Since most output devices and more particularly the human perception cannot

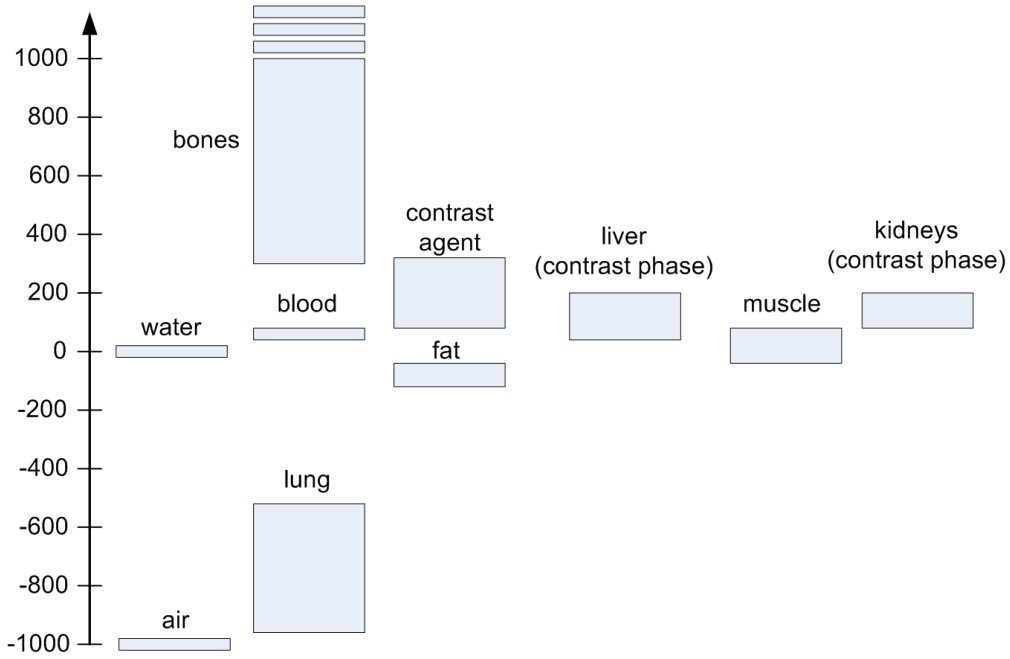
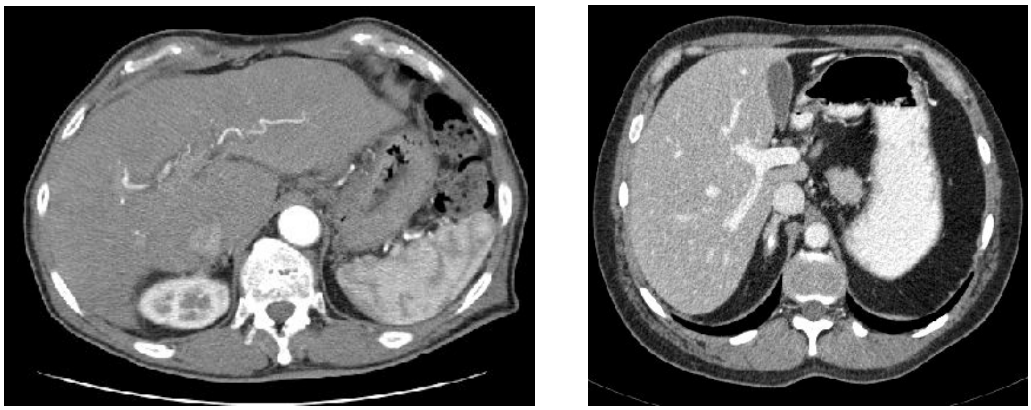


Figure 2.2.: The Hounsfield scale. The typical range of Hounsfield units is shown for selected anatomical structures.

differentiate between that many intensity values, so called *level-windowing* is used. Here, an intensity interval of a certain size (window) and a certain position (level) inside the Hounsfield scale is specified. The intensity interval is chosen such that the intensity range of the structure of interest, e.g. the liver, is fully contained in the interval. The HU values inside the chosen interval are now linearly mapped to the full displayable gray value range. HU below or above the interval are set to the lowest and highest displayable value, respectively. By this means, neighboring HU will be displayed with a higher contrast, which makes it easier for a human to visually distinguish between them.

2.1.2. Contrast agent application in Computed Tomography

As shown in Section 2.1.1, CT imaging is based on the principle of the interaction of X-rays with tissue. Depending on the body part to be examined, more or less radiation is necessary to penetrate the body and to generate an image of good visual quality. While body parts like the lung only require a moderate radiation dose to be accurately reconstructed, abdominal soft tissue CT usually requires a significantly higher radiation dose. The reason for this is that the density coefficients of different soft tissues like organs, fat or muscles are very similar. Therefore, more radiation is needed to generate images with sufficient contrast between those structures. However, in order to protect the patient's health, the radiation dose cannot be arbitrarily increased. Therefore, contrast agents are used in clinical practice which increase or decrease the amount of absorption of the X-rays in the medium they are introduced in. In CT, contrast agents are most frequently used to increase the intensity of vessels (angiography), tumors or organs like the bowel. Depending on the nature of the contrast agent (oral administration or injection into the blood stream), other structures than the target structure will also be contrast enhanced, e.g. the stomach for oral administration or vessels that are connected to the vessel system of the target



(a) (b)

Figure 2.3.: Two commonly used CT contrast enhancement phases: (a) arterial phase, (b) portal-venous phase.

structure in case of blood stream injection.

2.1.3. Liver angiography

One of the main target organs for the techniques developed in this thesis is the liver. In CT scans of the liver, contrast agent is frequently used for diagnosing oncologic liver diseases. Since the contrast agent is injected into the blood stream of the patient to visually enhance the different vessel systems of the liver, this technique is called liver angiography. Depending on the point in time the scan is performed after injection of the contrast agent, the appearance of liver vessels and liver parenchyma differs significantly which has to be considered when developing techniques for liver segmentation and registration.

In liver angiography usually iodine compounds are used as contrast agent since this element is characterized by a high radiation absorption. Such a contrast agent is injected into the patient's blood stream during examination — for example through an arm vein. Shortly afterwards, the contrast agent reaches the heart and is pumped into the aorta. From here, it reaches the liver artery which is directly connected to the aorta. This point in time is denoted as the *arterial phase*. Figure 2.3(a) shows a cross sectional view of a CT scan of the liver at this point in time. As can be seen, a small vessel, the liver artery, is contrast enhanced inside the liver while the surrounding tissue is not saturated. The arterial phase now fades. At the same time, contrast agent saturated blood from the abdominal area reaches the liver from its second supply vessel: the portal vein. This vein collects the contrasted blood that took a longer way from the aorta through the abdomen. Figure 2.3(b) shows a CT scan at this point in time which is called *portal venous phase*. In this phase, the large portal venous system of the liver is saturated by contrast agent. Additionally, the liver parenchyma also gets saturated with contrast agent, because a lot of blood reaches the liver through the portal vein. A later point in time when the portal vein saturation and the liver parenchyma saturation starts to fade is called *venous phase*. Here, the liver veins are relatively clearly saturated.

2.2. Nomenclature

This section constitutes the nomenclature of the terms *segmentation* and *registration* as it will be used within this thesis.

2.2.1. Segmentation

In this thesis, the term *segmentation* is interpreted as it is used in several literature survey publications [WZG07, HGM09, Erd09, ZOON11]. In general, the term segmentation denotes the process of assigning sets of voxels to one or more distinct groups that are defined by the needs of the respective image processing task. Regarding medical imaging, volumetric segmentation is based on the classification of voxels to regions, which usually correspond to objects or organs in the data set. For example, the visualization of the human body's skeleton can be realized by dividing a CT image into the two classes *bones* and *not bones* and then pass the result to a renderer. Finding criteria to decide which voxels in the volume are similar or share a common property is therefore the essential part of every segmentation technique. The result of segmentation is a classification that labels every voxel to be part of a certain region. This is referred to as binary segmentation since a voxel either shares a property with its neighbors or not. Because medical imaging techniques like CT or MRI produce discrete volume grids, certain voxels may represent two different materials, for example, on object boundaries. The so called partial volume effect leads to an uncertainty whether the voxel has to be assigned to the one object or the other. In contrast, fuzzy segmentation only computes a probability that a voxel belongs to a certain region. In the remainder of this thesis the focus will lie on binary segmentation since most medical imaging classification techniques target at a clear distinction of the detected structures. Furthermore, most computer aided medical applications like radiation therapy planning applications require a binary segmentation. For an overview of fuzzy segmentation techniques the reader is referred to the book by Terry Yoo [Yoo04].

2.2.2. Registration

The term *registration* is used within this thesis analogous to well known survey literature in the field [MV98, Zit03, Yoo04, CHH04].

In medical diagnosis and therapy, often several images of the same patient from different imaging modalities or from different points in time are used. Having information from CT, MRI, PET (positron emission tomography) or ultrasound combined enables the physician to make more precise diagnosis, since all those acquisition techniques are sensible to different components of human anatomy like bones, soft tissue or, as in the case of PET, show functional information of anatomical structures. However, there are some problems that make a direct comparison of the resulting images difficult. Usually the position and respiration state of patients change when moving from one imaging modality to another. In addition, sometimes a significant amount of time elapses between two recordings, so the patient may lose or gain weight. Furthermore, an operation itself will change appearance and anatomy of the operated structures.

The general goal of registration is to provide a mapping between two different domains or coordinate systems such that corresponding points in both domains are mapped. In the context of the described medical scenarios, registration provides a mapping between images that show comparable content such that for every voxel or point of interest in the first image, a corresponding voxel or corresponding point of interest in the second image can be identified. Figure 2.4 shows this

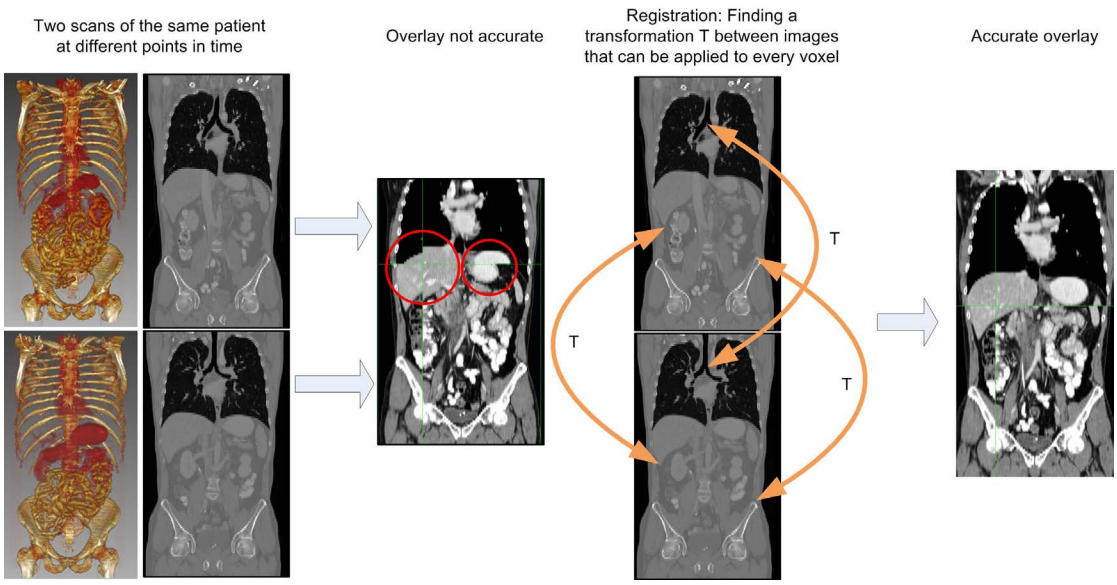


Figure 2.4.: The principle of registration in medical imaging. Two images, for example from the same patient at different points in time, cannot be accurately overlaid, because of differences in pose or respiration state. Registration finds a transformation T that defines a mapping for every voxel in one image to a corresponding voxel in the other images such that the same anatomical structures are matched.

basic principle. Finding an appropriate transformation between the images is the challenge of creating a registration algorithm. Usually one image is called the reference image and the other one the target image, whereby the target coordinates are mapped to the reference coordinates. There are several approaches that basically differ in the way what kind of transformations are used. For example, rigid registration only allows translation and rotation to map the images while an elastic registration can deal with local deformations. Appropriate transformation parameters are usually found by iteratively maximizing a similarity measure called *registration metric* between the images. A detailed overview of well known registration transformations, metrics and optimization schemes can be found in the surveys of Maintz and Viergever [MV98] and Zitova [Zit03]. A report on non-rigid registration techniques is given by Crum et al. [CHH04].

2.3. Taxonomy of organ segmentation and registration methods

Medical Image segmentation and registration have been important research topics over the last two to three decades and several state of the art surveys exist for segmentation [PP93, MT96, PXP00, Yoo04, HGM09, ZOON11] as well as for registration techniques [MV98, Zit03, CHH04]. However, in the scope of this thesis, three open question statements exist that have not been addressed so far.

The first question addresses possible similarities between segmentation and registration methods. Some registration approaches can be used to solve segmentation problems and vice versa. Since the goal of this thesis is the development of a technique that can be used to address segmentation *and* registration problems, it is necessary to determine the intersection set between segmentation and registration methods for a state of the art literature research. The first question

therefore is: what methods form the intersection set and which techniques are particularly suited for *automatic volumetric organ* segmentation and registration?

The second question is a direct consequence of the first question. If an intersection set of segmentation and registration methods exists, there is obviously an ambiguity in the current nomenclature of segmentation and registration approaches. This ambiguity makes it difficult both to properly classify an approach and to perform literature research. The second question therefore is: what is a proper taxonomy for segmentation and registration methods that does not have said ambiguity?

The third question regards the need for a classification scheme that can be used to determine the usability of a method for a given structure to segment or register. In the current literature, methods are often classified based on the technical class of technique used, but not based on the type of segmentation or registration problem that it can be applied for. In other words it is described what a method does and not for what a method can be used for. While many segmentation methods are very generic and can be applied to a variety of different segmentation problems other approaches are very specialized and can only be used in a small application domain. The third question is: what is an adequate classification scheme for determining the usability of a method for a given anatomical structure to segment or register.

The taxonomy presented in this section tries to address the open questions described above. In particular, a new classification scheme is proposed that comprises a new class of techniques that can be used for segmentation and registration. It is called *regmentation*. Based on the developed taxonomy, it is discussed what kind of methods are particularly suited to address the given task of this thesis.

2.3.1. Classification of segmentation techniques

There are three main characteristics which influence the segmentation of an object in an image: object boundaries, object homogeneity and object shape. Object boundaries and object homogeneity are image or signal based characteristics. Therefore they are affected by image specific disturbances like noise or reconstruction artifacts. Furthermore, they are modality dependent. For example, an object may have very dominant boundaries in a CT image and only poor boundary representation in an ultrasound image. An object's shape is image independent and in most imaging modalities — apart from small deviations like perspectival mapping distortions — also independent from the acquisition technique.

The concepts of object boundary, object homogeneity and object shape have a strong influence on the development of segmentation methods. Segmentation techniques try to detect boundaries and homogeneous regions in the images and incorporate shape information to restrict the shape of the resulting segmentation. Some methods like thresholding or region growing rely more on the image or signal information while other methods like model based approaches have a stronger focus on modeling the object's shape in the segmentation process. In the literature, segmentation approaches are therefore often classified according to the amount of boundary, homogeneity or shape knowledge they incorporate. Over the last three decades, several surveys about medical image segmentation have been published. Nikhil et al. [PP93] distinguish threshold methods from iterative pixel classification, surface based segmentation techniques, edge detection methods and fuzzy set theory methods. Hu et al. [HGM09] categorize segmentation techniques into four groups whereas each group is defined by the image features used by the segmentation technique: region-based, boundary-based, hybrid and atlas based. Zuva et al. [ZOON11] dis-

tinguish between threshold, edge and region based methods. Pham et al. [PXP00] use eight categories: thresholding approaches, region growing methods, classifiers, clustering methods, Markov random field models, artificial neural networks, deformable models and atlas guided methods. Thresholding, classifier, clustering and Markov random field methods are considered as pixel classification methods.

Although the nomenclature used in the described literature is not fully consistent and some single approaches have been assigned to different groups, two main classes of algorithms can be identified: *image based algorithms* and *shape based algorithms*. The number of subcategories used in the literature varies and hybrid categories are used in some articles to classify algorithms which show characteristics of multiple categories.

In this thesis, a more generic view on the classification of segmentation approaches is proposed. The proposed taxonomy is based on a continuum between two extremes: purely image based algorithms and strong shape based methods. All segmentation algorithms are classified inside this continuum according to the amount of shape information used by the method. An aspect that derives from this view is the shape generalizability and shape specializability of an algorithm. The more shape information an algorithm incorporates, the more specialized it gets. For example, a geometric active contour or *snake* with a low elasticity can only be used to segment objects which smooth boundaries. In contrast, a threshold can be used to segment arbitrary shapes with the same parameter setting. Figure 2.5 shows a schematic view of this taxonomy.

The proposed taxonomy contains four categories: voxel based methods, region based methods, methods with local shape priors and methods with global shape priors. The single categories will be explained in detail in the upcoming Sections 2.3.1.1 to 2.3.1.4. The categories have been chosen, because they form clearly identifiable groups in the continuum between image and shape based methods. Of course other ways to separate the continuum exist. There will also be segmentation algorithms which can be argued to fall into one or another category. However, in comparison to other classification schemes, all algorithms are embedded into the same continuum and can therefore be clearly distinguished from each other and set into context to other algorithms.

In many publications [PXP00, HGM09, ZOON11] machine learning techniques like clustering and classification methods represent either separate classification categories [PXP00] or form subcategories, for example sometimes they are considered as subcategories of region based methods [HGM09, ZOON11].

In this thesis, machine learning methods are not considered part of the segmentation classification scheme as described above, because they are not *per se* segmentation methods. Rather, they can be used to *support* segmentation mainly by finding appropriate *parameters* for a segmentation method. In Figure 2.5 it is indicated that the number of parameters of a segmentation method increases with the amount of shape domain knowledge used. For example, purely image based methods like thresholding or histogram based methods only need very few parameters — in a simple thresholding case only one parameter exists that represents the threshold. For such methods, it may be enough to examine some representative cases or to consult a domain expert. Model based approaches like mass spring models are characterized by many parameters, since they model complex shape knowledge. Machine learning methods can be used to automatically find appropriate parameters for such methods in order to increase their robustness on a diverse test data base. They can therefore help to automate complex segmentation algorithms. Furthermore, for structures with a strong shape variation like chromosomes [CW11], white mat-

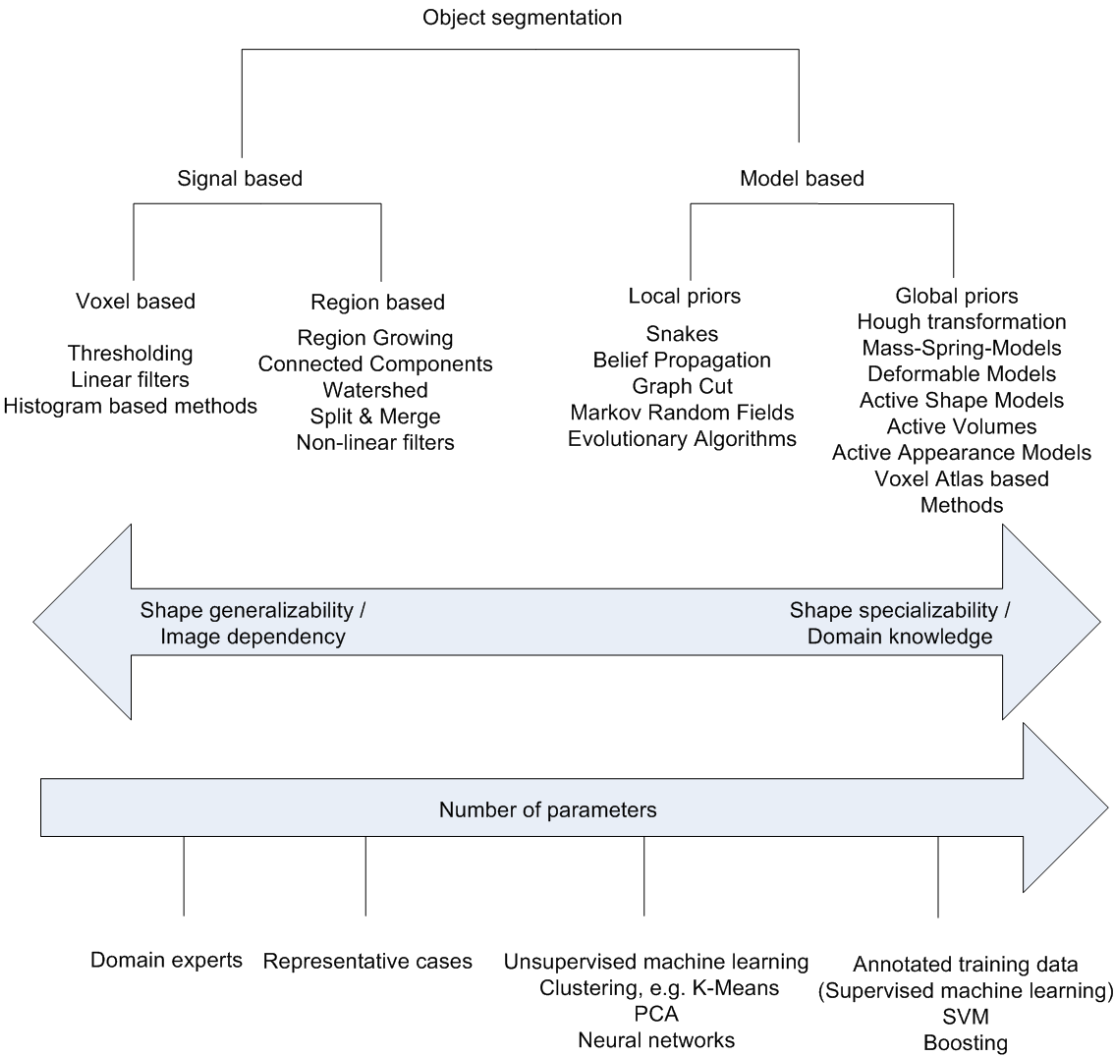


Figure 2.5.: Proposed classification scheme of methods for segmenting objects in medical images. The methods having the broadest application area concerning shape variability, i.e. showing the best shape generalizability are shown on the left. Those methods usually have only few parameters, because they do not incorporate much domain specific knowledge. The approaches shown on the right are highly specialized, i.e. they are often only suitable for one particular structure to segment and incorporate a high amount of domain specific parameters. Nowadays, typically supervised machine learning algorithms are used to cope with the high parameter count by incorporating a training base of known cases that serves as a model for the given problem statement [ESS12].

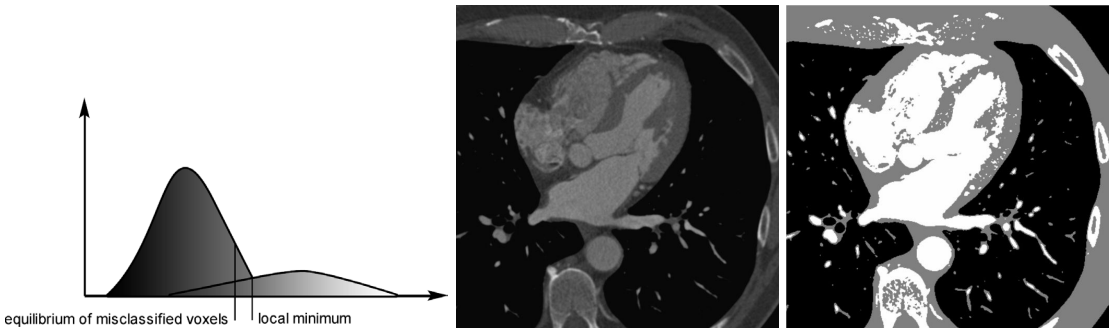


Figure 2.6.: (a) the histogram of two different objects which are characterized by a broad and a narrow distribution, respectively [Erd09]. (b) a CT image of the heart and (c) a segmentation using the Otsu method [Erd09].

ter [SCDC11] or prostatic glands [PJE*11] often multiple low parameter segmentation methods are combined to extract a set of simple image features. Machine learning algorithms are then used to cluster the image features such that a meaningful segmentation of the target structure can be achieved. Section 2.4 gives an overview of frequently used machine learning methods in medical imaging.

In the classification view presented in this thesis, machine learning algorithms are supportive tools to help to automate segmentation methods or to support segmentation methods by performing statistical analysis on a set of data that is originally generated by the segmentation algorithms. Machine learning methods can therefore be coupled with any existing segmentation algorithm. However, such a coupling does not add a new level of complexity to the data and it does not add any further image or shape domain knowledge to the existing segmentation algorithm. That means, it does not change the classification of an algorithm according to the taxonomy presented in this thesis. Therefore, machine learning methods are not part of the presented classification scheme.

In the following sections, the four classification groups as shown in Figure 2.5 are discussed in detail. A large number of medical image segmentation approaches have been proposed in the literature working either on two-dimensional or three-dimensional data. In each classification group, the most important methods for medical image segmentation are described. Often complex segmentation methods consist of a whole pipeline using algorithms from different classification groups. In such a case, a method is classified by the algorithm, that uses the most shape knowledge. For example, an approach that uses a thresholding followed by a model based segmentation is considered a model based method.

2.3.1.1. Voxel based methods

The first classification group for image segmentation according to the taxonomy presented in the previous section as well as in Figure 2.5 consists of *voxel based methods*. The voxel based methods group consists of methods that purely rely on image information and do not incorporate any prior shape knowledge about the structure to segment. That makes them suitable to segment structures that strongly vary in shape and at the same time show good images contrast. However, since they are purely based on image signal information voxel based methods are not very well

able to deal with image noise, reconstruction artifacts or low object contrast. The following gives an overview of well known voxel based methods that are used for medical image segmentation.

The simplest approach to address the segmentation problem is to classify a voxel solely based on its intensity. Such methods do not incorporate any local relationships between the voxels. From that perspective, a segmentation can be made by determining a value range that assumingly uniquely contains the gray values of the structure to be segmented. This approach is called *thresholding* and is often used by more sophisticated methods as a preprocessing step to build a coarse-grained segmentation. Some medical image acquisition techniques like CT have the advantage that it is roughly known a priori to which intensities different tissues will be mapped. Thresholds can therefore be directly determined. However, the value ranges of tissues slightly vary from scan to scan and between different patients, so a manual selection is generally not sufficient. In order to automatically find suitable thresholds, usually the histogram of the image is inspected and searched for two or more local maxima. The appropriate thresholds are then determined by the minimum between them. Unfortunately, most of the time the gray value ranges of different objects in images overlap. This has a direct impact on the appearance of the histogram, since the area of a local minimum is not necessarily the place with a minimum overlap of two corresponding distributions. For example, Figure 2.6(a) shows the histogram of an image that contains two different objects. One object has a very broad range of gray values (i.e., its histogram shows a wide distribution) while another is characterized by a high contrast to the background (and shows a narrow distribution). Here, the place in the cumulative histogram where both objects have an equal amount of misclassified voxels is not the local minimum between the two peaks.

Machine learning algorithms can be used to find an optimal threshold. The clustering algorithm *k-means* divides the histogram into k clusters such that a metric be tween the histogram's elements and the centroids of the clusters is minimized. For example, an image is divided into two clusters, one for an object and one for the background. First, the centroids of the clusters are, for example, randomly initialized. As a metric one may choose the minimum gray level distance. The intensity of each pixel is therefore compared to the mean intensity (centroid) of each cluster and is then assigned to the most similar one. Afterwards, the mean intensities are updated and the procedure starts anew until the algorithm converges. The result is a separation of the object from the background. One can think of generalizing this procedure to work with n-dimensional data vectors instead of just gray values. In addition, other metrics like the sum of the variance over all clusters or the total distance between every value and their centroids can be used. While having the advantage of being simple and fast, the major drawback of these algorithms is that they do not necessarily always converge to the same result — a consequence of the random placing of cluster centroids.

Another well known automatic thresholding technique is the Otsu method [Ots79] which automatically divides the image into two or several (multilevel Otsu method) classes. The idea is to find a threshold such that the gray value distribution in each cluster gets as narrow as possible. Since the distribution of one cluster gets wider when another gets tighter, the goal is to minimize the combined spread of all clusters. For example, a single object shall be segmented from the image background. For a given threshold T the mean μ_{back} and μ_{object} for the two resulting clusters is computed and weighted with the probabilities ω_{back} and ω_{object} of a voxel belonging to the respective cluster. The following term is called the between-class variance:

$$\sigma^2_{between}(T) = \omega_{back}(T)\omega_{object}(T)\left(\mu_{back}(T) - \mu_{object}(T)\right)^2. \tag{2.1}$$

The between-class variance is a measure of distance between the peaks of each distribution. In order to minimize the combined spread the between-class variance has to be maximized. However, practically every possible threshold would have to be tested by calculating the means and then choose the maximal result, so this algorithm means a heavy computational burden. Therefore the above formula is further simplified. It can be shown [sLsCcC01] that maximizing the between-class variance is equal to maximize a modified between-class variance denoted as:

$$\sigma^2_{betweenModified}(T) = \sum_{j=1}^{k} \omega_j \mu_j^2, \tag{2.2}$$

where k is the number of object classes. Liao et al. [sLsCcC01] showed that it is possible to precompute all sums of ω and μ in a look-up-table. The modified between-class variance can then be recursively computed by adding up the precomputed results of the table. Using the described improvements allows for a noticeable speedup making the algorithm interesting for medical image processing. In Figure 2.6(b) and 2.6(c) the segmentation result of a cardiac CT data set using the Otsu-method is shown.

In complex segmentation scenarios, thresholds and histogram based methods tend to produce either small islands that are not part of the object to be segmented (but share the same gray value) or result in segmentation holes (because the threshold was set too high).

Voxel based methods are very simple and are usually not suited for complex segmentation problems. They rely purely based on image content and do not include any shape specific knowledge about the structure to segment. However, they are useful for segmenting structures with a strong shape variation which at the same time show good image contrast. For example, Aarle et al. [vABS11] propose a threshold selection strategy to segment dense objects in tomographic images like phantom scans. Voxel based methods are used in cell image segmentation [YLKC10, YBCK10] where tens to hundreds of single cells with varying shape can be segmented simultaneously. Machine learning methods are often coupled with voxel based methods to increase their flexibility and robustness on diverse test bases. For example, Zhang et al. use voxel based methods together with several machine learning techniques to segment cervigram images [ZHM*10]. Yin et al. [YBCK10] apply classification methods together with voxel based methods for cell image segmentation.

Voxel based are often used as a pre-processing step in more complex segmentation pipelines. Sometimes, certain intensity areas in the image can be neglected, because the structure of interest does not contain any gray values from that intensity area. Voxel based methods are a simple and fast way to achieve that goal, helping to reduce the complexity of the segmentation problem for further processing steps. Furthermore, because voxel based methods usually have no prior shape knowledge and only few parameters to set, they can be easily automated.

2.3.1.2. Region based methods

The second classification group for image segmentation according to the taxonomy presented in Figure 2.5 consists of *region based methods*. Region based methods are mainly based on image signal information, but they incorporate local relationships between voxels, for example for building contiguous regions.

One of the most prominent region based methods in 2D and 3D is the *region growing algorithm*. Here, a segmentation grows from initially placed points called seeds by aggregating neighboring

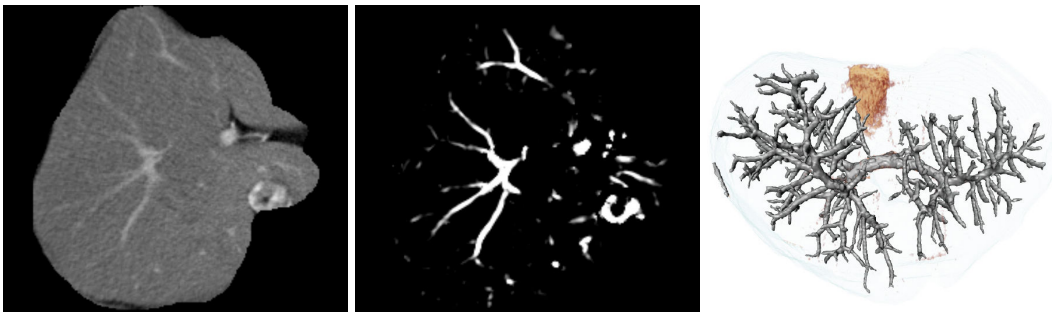
(a) (b)

Figure 2.7.: The region growing algorithm applied to segment the portal vein of the liver. (a) The original image is shown on the left. The right image shows the resulting segmentation that grew from a seed point at the position of the shown cross [ESS12]. (b) Three dimensional view of a leaking of region growing caused by a high threshold [ERS08].

pixels or regions according to some similarity criterion. Region growing is often used to segment homogeneous regions like vessels trees which vary in shape but share a similar intensity. Figure 2.7(a) (left) shows a cross sectional view of a CT image of the liver. The vessel tree is contrast enhanced and clearly visible. The region growing can be used on such an image to segment the vessels. First, a voxel inside a vessel is selected — either manually or automatically — as a seed point. In addition, a gray value range is defined that covers all intensities that occur within the vascular tree, i.e. a lower and an upper threshold, T_l and T_u, are defined as a measure of similarity. The region growing iterates over all voxels that have been segmented so far, comparing the current intensity with all direct neighbors and adding those voxels with an intensity between T_l and T_u. As a result, a connected segmentation of the vessel is obtained (cf. Figure 2.7(a) (right)). A problem arises if neighboring objects share a similar intensity to the structure to be segmented or if the intensity range is set incorrectly. The algorithm then leaks into those areas as Figure 2.7(b) shows. Nevertheless region growing is a widely used segmentation method due to its computational simplicity and the fact that the connectivity of all voxels grown from a seed point is ensured.

Another algorithm that iteratively groups pixels into regions is the so called *watershed transform*. Its idea is inspired by the observation of rain falling on a non flat area with peaks and valleys. Typically the water will pool at the local minima of the region and build small lakes. Applied to image processing, the gray values of an image can be thought of as height differences on a rectangular surface. A segmentation is now created by picking the minimum gray value v_{min} as the sources of basic catchment basins. In the next step all neighboring voxels with an intensity of $v_{min} + 1$ are added to the basins. If a voxel with the according intensity does not adjoin to a basin, a new one is created at that place. After some iterations the catchment basins will usually meet. The borders between them are called watershed lines and form the boundaries of the final segmentation. Because often the segmentation of whole objects is required, the algorithm is usually applied to gradient images, where the edges denote the local maxima. Smoothing the image is often used to remove local minima that may occur due to noise. More sophisticated approaches to further improve the watershed segmentation were introduced by Meyer and Beucher [BM93, Mey94] who developed methods to define a unique behavior in the presence of plateaus.

A fully automatic region based segmentation method is known as the *split and merge algorithm*. It starts considering the whole image as a single region. The region is then tested against



(a) (b) (c)

Figure 2.8.: Example of vesselness filtering on a liver CT image [ERS08]. (a) original image. (b) vesselness filter output. (c) rendered segmentation result of the portal venous vessel tree after applying a region growing algorithm.

a homogeneity criterion. If the test fails, the segmentation is split into four smaller regions (eight in three dimensions) of equal size. This procedure is recursively repeated until no further splitting is necessary. In a subsequent merging step, adjacent regions are tested for similarity and merged accordingly. The result is an irregular segmentation of single homogeneous regions. Since the algorithm is splitting the image into equally sized blocks, it is suitable for segmenting local and contiguous structures instead of fine objects like vessels which may be spread over the entire data set.

Another category of region based methods are filtering methods. In medical imaging, image filters are mainly used for point, edge or tube detection. In many literature overview publications [PP93, HGM09, ZOON11], such kind of methods form a separate category of segmentation algorithms. However, like the region based methods described above filters are image based methods that incorporate local relationships between voxels in order to classify a voxel in the image. In the context of the shape knowledge driven taxonomy presented in Figure 2.5, they are therefore classified as region based methods.

Edge detection is commonly used in medical image segmentation, because the boundaries of anatomical structures are often characterized by an intensity difference between the tissues. Such intensity discontinuities or edges can be found by computing the derivatives of the local image intensity function. Usually those derivatives are approximated by convolving the volume with filter masks . Several edge detection filters have been proposed, for example, the Canny edge detection filter [Can86]. However, filter results frequently needs to be post processed, because object contours are often not closed. While image filtering alone is in most cases not sufficient for image segmentation, the concept of edge detection plays a central role in algorithms that incorporate additional shape information. In particular, a large amount of local and global model based segmentation approaches as described in the upcoming Sections 2.3.1.3 and 2.3.1.4 incorporate edge detection methods for searching object boundaries.

Another well known type of filtering methods is the class of filters proposed by Sato et al. [SNA*97] and Frangi et al. [FNVV98]. Those filters are based on second order derivatives of the image and scale space theory [MVN06]. The approximated second order derivatives are combined to form the local Hessian matrix around a voxel. Using eigenanalysis, the Hessian matrix can be evaluated to detect tube like structures in an image. This property makes such filters ideally suited to detect all kinds of vessels in an image, for example pulmonary vessels [KKE*08],

coronary and retinal arteries [KCBP04] or liver vessels [ERS08]. They are therefore sometimes called *vesselness filters*. A subsequent segmentation can be achieved by thresholding the filter output or by applying region growing methods [ERS08]. Figure 2.8 shows an example of vesselness filtering on a liver CT image. The original image is shown in Figure 2.8(a) and the output of the filter is shown in Figure 2.8(b). A region growing is then used to segment the vessel tree in the image (cf. Figure 2.8(c)).

Region based methods have a lot of advantages. They are almost as flexible as voxel based methods and can be applied to a large variety of different segmentation problems. However, they incorporate neighborhood relations which is a very natural way of describing medical image content since neighboring voxels are in most cases related. This makes region based methods generally more robust than voxel based methods. They also do not incorporate complex shape knowledge which limits the amount of parameters to set for a region based method.

Apart from their application to vessel detection and segmentation, they are widely coupled with machine learning techniques to detect tumors or lesions of all kind. Usually filter based methods are used to describe the local texture around a voxel. Using a set of positive and negative examples, machine learning methods are then applied to automatically find the most characteristic features for that neighborhood. Such strategies have been applied to retinal lesion detection [QRA11], pulmonary nodule detection [MHR10], the detection of hepatocellular carcinoma [XS11] or the detection of white matter changes [SCDC11].

Generally, region based methods are very well suited for segmentation of objects that strongly vary in shape. Apart from pathological structures, for example, they have been applied to hippocampus segmentation [WSD*11] and neuron membrane segmentation [JPW*10]. However, since they only incorporate direct neighborhood relations, they are mainly based on image appearance. This limits their applicability to objects that are homogeneous in terms of intensity or texture pattern. They are therefore usually applied to small objects instead of complex structures like organs.

2.3.1.3. Shape methods with local priors

The third classification group for image segmentation according to the taxonomy presented in Figure 2.5 consists of segmentation methods that incorporate prior shape knowledge about the structure to segment. However, this prior shape knowledge is modeled *locally*, for example by enforcing that the surface of the object has a certain degree of smoothness. Furthermore, since local prior shape methods do not have a global idea of the shape to be segmented, they are not restricted to segment a specific shape but can be applied to various *types* of shapes.

A well known method for two dimensional image segmentation based on local shape priors is the *active contour* or *snake* approach [TF88]. A snake is a contour or curve parametrically defined in an image $B(x,y)$ on the image plane $(x,y) \in \Re^2$ as $\vec{s}(p) = (x(p), y(p))^T$, with $x(p)$ and $y(p)$ being the coordinate functions. $p \in [0,1]$ is the parametric domain. The shape of the snake is given by minimizing the energy functional

$$\Sigma(\vec{s}) = I(\vec{s}) + E(\vec{s}), \tag{2.3}$$

where $I(\vec{s})$ and $E(\vec{s})$ are representations of two energies: the internal snake energy I and the external snake energy E, respectively. The internal energy defines the rigidity and the tension of the contour, i.e. it defines how smooth and flexible the snake is. The internal energy therefore

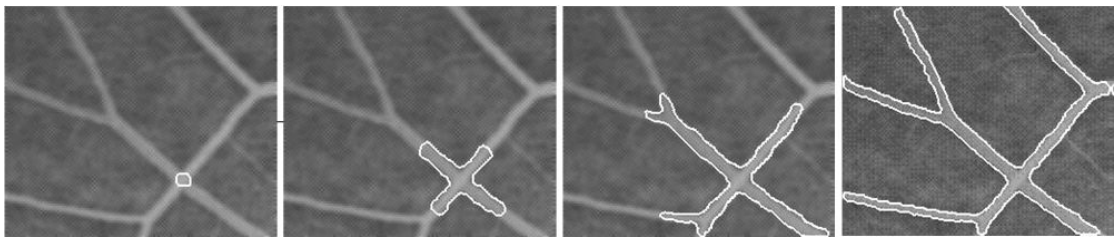


Figure 2.9.: Example of the snake algorithm applied to retinal vessel segmentation (image taken from [MT95]).

is a local shape prior, that determines how the object to be segmented should locally look like. The image driven part of the method is given by the external energy. It determines what kinds of image features attract the contour.

An example for an internal energy is

$$I(\vec{s}) = \int_0^1 w_1(p) \left| \frac{\partial \vec{s}}{\partial p} \right|^2 + w_2(p) \left| \frac{\partial^2 \vec{s}}{\partial p^2} \right|^2 dp, \tag{2.4}$$

where w_1 and w_2 define tension and rigidity of the snake, respectively. As an attracting image feature, often the image gradient is taken, i.e. the contour will evolve towards intensity differences in the image. The external energy then is

$$E(\vec{s}) = \int_0^1 -|\nabla (G_\sigma B(x,y))| \, dp. \tag{2.5}$$

Here, ∇ denotes the gradient of the image B smoothed by a Gaussian G. The standard deviation σ controls the extent of edges that attract the contour. The minimum of (2.3) is usually found using numerical algorithms as described in [MT96].

Snakes have been widely used for two dimensional image segmentation problems. The three dimensional generalization of snakes is called *deformable model* or *deformable surface* [MT96]. The advantage of the snake formalism is that the shape prior is very intuitive and easy to control. For example, if the structure to segment is a bone that does not contain any sharp edges, rigidity and tension of the snake can be set high. That way the snake is less dependent on the image signal and therefore also less affected by noise or discontinuous and weak edges. On the contrary, if the structure to segment does not comprise a smooth surface, the internal energy of the snake can be set more versatile in order to adapt to the structure. Figure 2.9 shows an example of a segmentation of retinal vessels using the very flexible snake variant called *topologically adaptable snake* [MT95] which is able to flow into complex shapes. The snake is initialized inside the vessel tree and iteratively adapts to the borders of the vessels until the whole tree is segmented.

The image driven part of the snake is generally very flexible, since many different external energies are possible. For example, snakes are used in ultrasound as well as in MR and CT which are imaging modalities that have very different image characteristics. In consequence of this flexibility, snakes are used to segment various kinds of anatomical structures including spicules in Mammography [MBG*10], tree structures [WNR11], the aorta in MRI [SsNC*10] and solid organs [MT96].

A disadvantage of snakes is their dependency on initialization. A snake will most likely get stuck in local minima if initialized far away from the structure to segment. This complicates automation of the method in comparison to most voxel and region based methods which are often

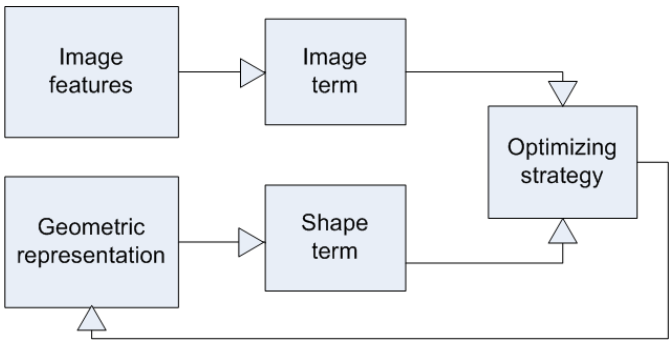


Figure 2.10.: Incorporation of local shape knowledge into the segmentation process. An image driven term and a shape preserving term are combined by an optimizer such that the shape representation is adapted to the image [ESS12].

easy to automate. Another drawback is that the flexibility of the snake is set globally for the whole contour, i.e. the stiffness of the snake is the same at all of its parts. This often means that the snake is not able to fully adapt to the structure in some parts while in other parts it may be already too flexible and leak into neighboring structures.

Apart from snakes, there are many other ways to incorporate local shape knowledge in the image segmentation process. The basic principle, however, remains the same. Figure 2.10 shows a schematic view of this process. Generally, an image driven term and a shape preserving term are defined. The image driven term is defined based on some image features. Image features can be, for example, intensity, edges, points of interest or regional homogeneity. The shape preserving term is defined based on the geometric shape representation used for adaptation. For example, the shape could be represented by an interconnected point cloud (cf. Section 2.5 for an overview of different shape representations). The shape preserving term could, for example, enforce that the distance between points should not change much.

Both image and shape preserving terms are combined and balanced using some kind of optimization strategy. There are many ways how such an optimization can be performed. Often the optimization problem is embedded into the well known frameworks of Graph Cuts [GPS89, FKE*10, EBK*10, SPV*10, APG11], Markov Random Fields [KSF*10, XMM10, BSX*10, FEK*11] or Graph based optimization [WLC*10, HUKE10, SLL*10, GGCH10, XNS*11]. After optimization, the geometric shape representation is updated. This process repeats until the structure of interest is segmented.

Local prior shape methods are widely used in image segmentation. Due to the incorporation of local shape knowledge and in comparison to the mainly image based voxel and region based methods, they can be used to segment objects with low contrast boundaries such as lymph nodes [BSX*10] in CT or the mitral annulus in 3D ultrasound [SPV*10]. As mentioned before in the example of snake segmentation, local shape prior methods are sensitive to initialization. However, since the shape priors are only defined locally — for example by enforcing a certain smoothness — they are usually flexible enough to adapt to a nearby structure if initialized closely. This makes these methods ideal for interactive segmentation [AHS10, EBK*10]. Other applications of local shape prior methods include bone segmentation [APG11], cell segmentation [DvCE*10] and detection of vessel-like structures [MST10, GGCH10].

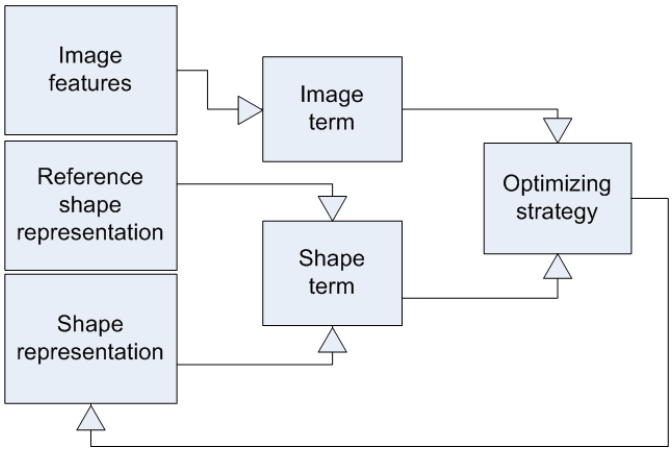


Figure 2.11.: Schematic view of the incorporation of global shape knowledge into the segmentation process. In comparison to local prior shape methods, the shape preserving term is additionally based on a reference shape representation that enforces the shape representation to be similar to a group of reference shapes [ESS12].

Local prior shape methods are also used for interactive and automatic segmentation of whole organs such as the kidneys [FKE*10], the prostate [FEK*11] or the bladder [SLL*10]. However, these organs have rather simple shapes and do not have strong shape variations between individuals. That is the reason why more complex organs like the liver or the heart are usually not segmented using local shape priors.

2.3.1.4. Shape methods with global priors

The fourth and last classification group for image segmentation according to the taxonomy presented in Figure 2.5 consists of segmentation methods that incorporate *global* shape knowledge about the structure to segment. These methods enforce the segmentation to be similar to one or a group of *reference shapes*. This way, complex objects can be robustly segmented even on low contrast images. Figure 2.11 outlines the principle of global prior shape methods which is similar to the scheme of local prior shape methods presented in Figure 2.10. An image driven term and a shape preserving term are combined by an optimizer such that the shape representation is adapted to the image. However, the shape preserving term is additionally based on a representation of reference shapes as mentioned above.

There are mainly two kinds of global prior shape methods: *geometric model based segmentation* approaches and *voxel atlas based segmentation* methods. They mainly differ in terms of the shape representation and adaptation process used.

Geometric model based segmentation In geometric model based segmentation, shape of single organs is represented by geometric objects like point clouds and polygonal surfaces [KLL07, HMW07, LZZ*08, ZBG*07, EPS*08, ZHB06], simplex meshes [CDNA07], B-spline representations [SBB09], level set representations [AFAX11, CBMS11, WSH09, KUA*09, CCBK07, FSK07], geometric grids [BP05] or finite element triangulations [SLH11]. Section 2.5.1 gives a detailed overview of shape representations used in geometric model based segmentation.

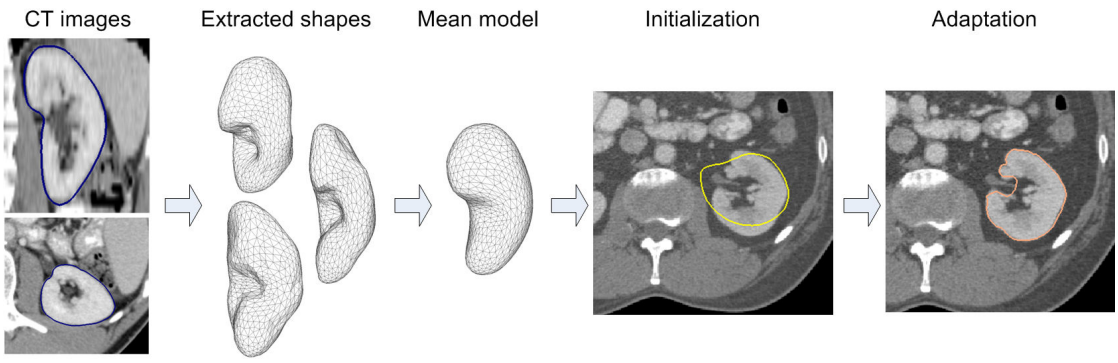


Figure 2.12.: Principle of geometric model based segmentation [ESS12]. An organ is manually delineated in a set of images. For each image a geometric representation of the organ is built. The built shapes are averaged to build a mean model. This model is then placed in an unseen data set and adapted to the image.

Figure 2.12 shows the principle of geometric model based segmentation. Based on the chosen shape representation, an instance of the object to be segmented — hereafter called model — is created that is used as an initial shape. For example, the initial shape can be the mean shape of some representative shapes of the structures to be segmented — hereafter called templates — or simply the shape of one particular case. The model is then placed in the image directly on the structure to be segmented or with significant overlap. This is necessary, because geometric model based segmentation is very sensitive to initialization. Unlike local prior shape methods, the model must also be roughly aligned with the structure to segment in terms of orientation. This is due to the fact that the adaptation of the model is in most cases performed based on a local boundary search which makes the handling of strong orientation mismatches difficult. Since geometric model based segmentation is usually used to segment complex shapes, often automatic model initialization methods are used which estimate the pose of the model in the image.

The initial model that has been placed in the data set adapts to the structure to be segmented according to the scheme presented in Figure 2.11. Like for local prior shape methods, image features like edges are searched in the image and integrated into an image term. Usually image features are used that describe the boundary of the structure to segment. A detailed discussion about boundary detection methods is given in Section 2.6.

As described above, the shape preserving term of geometric model based segmentation is based on the geometric shape representation and on the template shape representation. There are many ways to model a group of template shapes. The simplest way is to define a single object to be the template. Frequently, three dimensional deformable model based approaches [MT96] are initialized with a single template shape if the shape variance of the structure to segment is not high. For example, lymph nodes [DSP*06] are always spherical objects so in most cases it is sufficient to use a sphere as initialization. The external energy is then set such that the segmentation stays similar to the template shape.

However, for anatomical structures that strongly vary between individuals a single template shape is not sufficient. For such structures, a representative set of template shapes is necessary. Usually, several dozens of templates are used depending on the complexity of the structure to be segmented. However, using many templates also increases the complexity of the decision process of whether the current model is similar to the template shape set. There-

fore, usually dimension reduction techniques like principle component analysis are applied to the template shape set in order to extract a limited amount of significant modes of variation that sufficiently describe the template set. This technique is called statistical shape modeling and has been proposed by Cootes et al. [CHTH93, CTCG95]. It is the most frequently used method for geometric model based segmentation of complex anatomical structures with a high amount of shape variation. Statistical shape modeling has been applied to segment the liver in CT [KLL07, HMW07, LZZ*08, SBB09], the heart and heart chambers in CT [ZBG*07, EPS*08], the prostate in MR [TSM11] or bone structures in X-ray fluoroscopy [WKFH11]. Section 2.5.3 gives a detailed introduction into statistical shape modeling.

Because geometric model based segmentation depends on knowing the expected shape of the structure to segment a-priori, it is less suited for segmenting objects that can have arbitrary shapes like tumors. Here, local prior model based segmentation methods are more appropriate, since they only require the surface of the object to meet some local smoothness criteria. Instead, geometric model based segmentation is used to segment complex shapes that vary in shape, but are not completely arbitrary. Here, the incorporation of global shape knowledge prevents the segmentation to leak into neighboring structures and to generate non plausible shapes. They are therefore especially suited for organ segmentation, since organs typically have complex shapes that vary in certain limits between individuals.

However, open issues in model based approaches exist. An issue concerns the shape restriction of the model, i.e. the shape preserving method to avoid a non plausible deformation of the models during adaptation. Usually statistical shape information as described above is incorporated for organ segmentation. However, the problem is that the shape restriction of statistical shape models is too restrictive for organs that show a high shape variability such as the liver. While it is expected that the generalizability of the model increases with the amount of template shapes, it is not possible to model all patient specific details. For example, the amount of fat and surrounding tissue, the respiration state, treatment procedure and the disease itself influence the shape of an organ. Statistical shape models are not able to model such complex detail. Therefore, the statistical shape restriction is usually relaxed at some point during deformation of the model [HM09]. The drawback of this procedure and an unsolved problem so far is that the model shape may become non plausible, thus losing the information from the learned shape space.

Another issue of model based segmentation concerns the shape representation of the structures to segment. The majority of current state of the art model based methods use models based on surface representations, i.e. a 3-dimensional surface model is evolved towards the boundaries of the target organ or structure. However, it often makes sense to incorporate *volume information* during the adaptation process in order to make the method more robust against model initialization errors and to increase the robustness through detection of organ foreign tissue that may not be detected when considering image information near the model's surface only. Different kinds of volumetric segmentation approaches have been proposed in the literature [TFCT98, PFJ*03, HM08, BPCO10, SLH11]. However, those methods are often complex and computationally expensive. Moreover, existing standard adaptation and shape preserving strategies cannot be used within those models or must be adapted. In Section 2.5.2, a detailed outline of existing volumetric model based approaches is given.

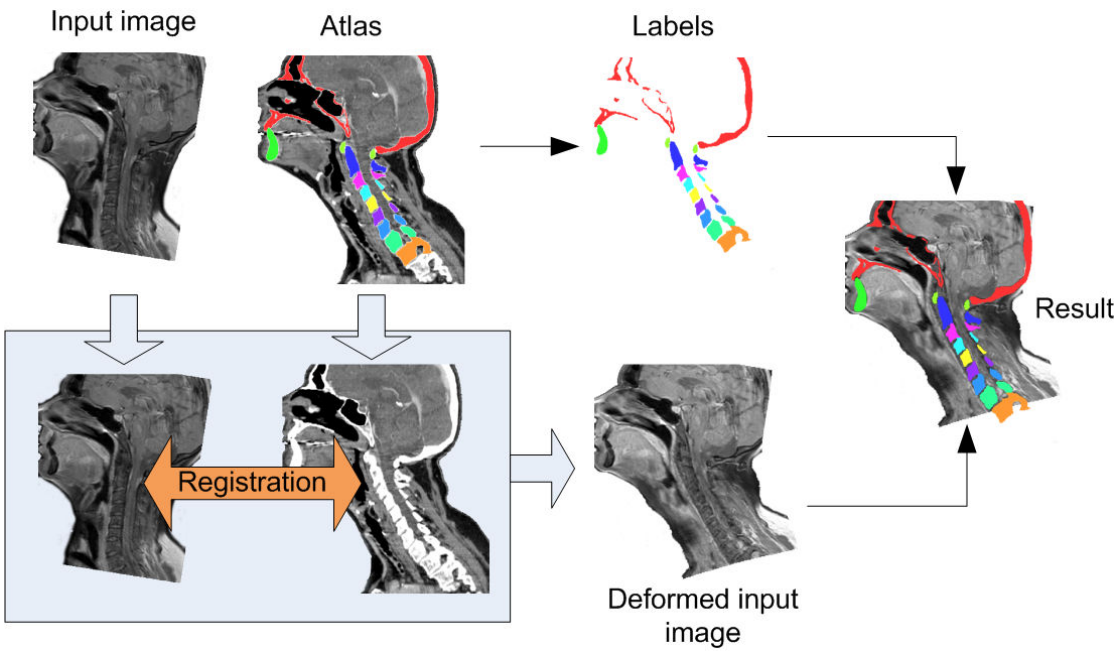


Figure 2.13.: Principle of voxel atlas based segmentation [ESS12]. A labeled image (the atlas) is registered with an input image. The labels from the atlas are then overlaid with the deformed input image so the segmented structures of the atlas are also available in the input image.

Voxel atlas based segmentation The second main type of global prior shape methods is voxel atlas based segmentation. Figure 2.13 outlines its basic principle[1]. In voxel atlas based segmentation, two images — a reference image called atlas and the input image to segment — are registered based on the voxel representation of both images. In the atlas image, the structures to segment are already contoured. After registration, both images are in alignment such that the segmented structures in the atlas image can be directly transferred to the input image. That way, all structures in the input image that are labeled in the atlas are segmented.

The quality of voxel atlas based segmentation mainly depends on two aspects: the atlas building strategy and the registration method that is applied to register an input image with the atlas [CM06]. There are many ways to build an atlas. In some applications, it is sufficient to use a single image as the atlas, for example in intra-patient atlas segmentation. Here, all images to be segmented stem from the same patient. The atlas is then created from the first image that is taken during treatment. All following images are directly registered with this first image. However, this means, that for each new patient, a new atlas has to be created. Therefore, usually the atlas is constructed more generically such that it fits not only to one patient but a group of individuals. A simple way to create a generic atlas is to average several representative non-pathologic images from different patients. However, it has been shown that patient variability is too high for most anatomical structures for an average atlas to work well [CGM08]. Possible solutions are population specific atlases where an average atlas is built for several population groups, for example based on gender or age [BM07]. Another approach tries to select the most similar image from a labeled data base of known cases as the atlas [CM07, WRLG*07]. Aljabar et al. [AHH*09] extend this method by selecting a set of the most similar images and registering each image individually

[1] Thanks to Sebastian Steger for providing the head and neck images.

with the input image. The segmentation results from every individual registration are afterwards combined to improve the overall segmentation for the input image.

After building an appropriate atlas, the registration with the input image is performed. Since this step is independent from the atlas building or labeling, generally any registration method can be used. In fact, this step solves a pure registration problem according to the definition in Section 2.2.2. Therefore, rigid, locally rigid or deformable registration methods can be applied depending on the structure to segment.

Voxel atlas based segmentation is the method with the highest amount of prior global shape knowledge according to the taxonomy presented in Figure 2.5. This is, because not only information about the shape of a single organ is considered, but implicitly there is information about *every* visible structure and its relation to all other visible structures available in the atlas. By registering the atlas with an input image, the relation between the structures — for example the heart is above the liver and is neighboring the lungs — is implicitly considered. That means, the segmentation of a single structure like the heart cannot leak too much into neighboring structures, because the neighboring structures themselves claim space in the input image during registration. Another interesting property of voxel atlas based segmentation is the ability to map arbitrary regions from the atlas to the input image. For example, in head and neck radio therapy, it is necessary to delineate the lymph node regions in order to limit radiation exposure to these areas. These lymph node regions do not correspond to any anatomical structures but are conglomerates of lymph nodes, fat, glandular tissue and vessels. However, the regions can be delineated in the label image of the atlas. Since the atlas is registered with the input image, arbitrary regions can then be mapped from the atlas label image to the input image. This strategy is used in many atlas based approaches for lymph node segmentation [CGM08, SIG*09].

Voxel atlas based segmentation has many advantages, because of the high amount of prior global shape knowledge they incorporate. However, this ability also makes them less flexible in practice. While a single structure can be modeled globally in an efficient way, for example using geometric model based segmentation as described above, a group of different structures and their relations are very difficult to model globally. Each image that is used for atlas building stems from only one specific patient. That means, it is just a snapshot of all possible variations in terms of organ shape, organ positions, organ orientations, respiration state or heart cycle. A complete global view of the whole body therefore would require an immense amount of images to build the atlas. However, even if enough images would be available, appropriate selection strategies are missing that can handle such amounts of data. Therefore, voxel atlas based segmentation is in practice limited to certain anatomical regions like the head and neck region [IDB*08, CGM08, SIG*09, RCM10] where the amount of variation is relatively low and a small amount of images is sufficient to build a complete atlas. Another way to utilize voxel atlas based segmentation is to use it as a coarse initialization method for other segmentation approaches. This way the missing accuracy of the atlas in areas of high variation is compensated, for example by a local prior shape method [GT11].

2.3.2. Regmentation: a new view of segmentation and registration methods

As mentioned in the beginning of this chapter, similarities between segmentation and registration methods exist. Some registration approaches can be used to solve segmentation problems and vice versa. In the literature, those hybrid approaches are usually either assigned to be registration or segmentation methods depending on the scope within the methods are used. Additionally,

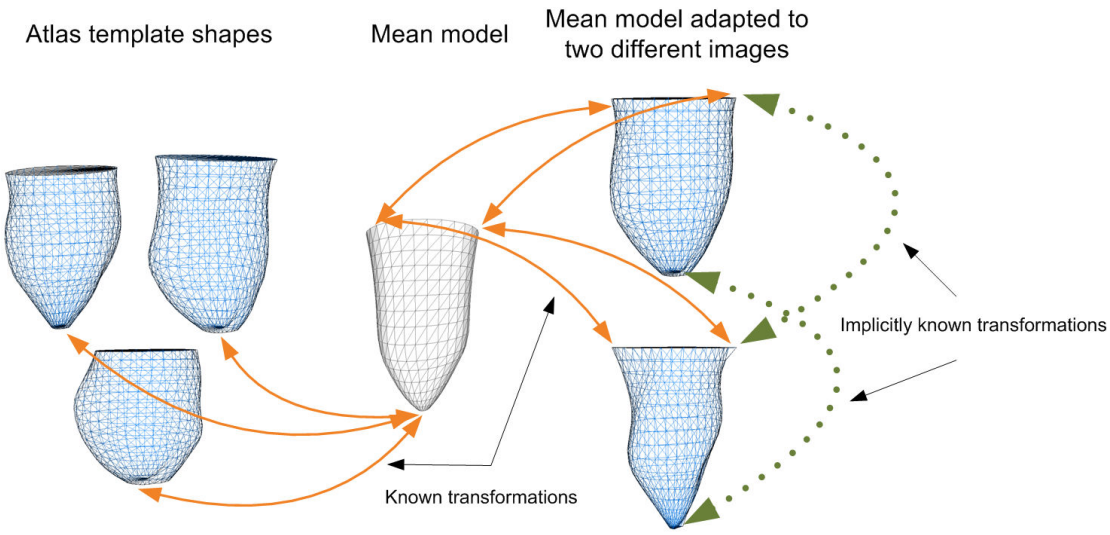


Figure 2.14.: Principle of solving a registration problem with geometric model based segmentation methods. The set of template shapes is regarded as an *atlas*. A mean model is constructed from the atlas and is adapted to two different data sets. Transformations from the atlas to the mean model are known as well as the transformations from the mean model to the adapted model (continuous arrows). Therefore, implicitly, the transformation between both adapted models is also known (dotted arrows), thus solving the registration problem between both images [ESS12].

sometimes the terms *segmentation for registration* and *registration for segmentation* are used to classify them in further categories. However, this point of view leads to an ambiguity, because some methods fall into multiple categories. For example, voxel atlas based segmentation would fall into the category *segmentation*, because it is used to solve a segmentation task. Furthermore, it would be classified as a *registration* method, because the technique used to solve the segmentation problem is a registration method. Lastly, it would also be classified as a *registration for segmentation* method for the same reason.

The reason why this ambiguity exists is that a method is classified based on two characteristics that are not related. The first characteristic is the problem statement that a methods tries to solve, for example a segmentation problem. The second characteristic is the technique that a method uses to solve the problem. Both characteristics are individual and unrelated but are often used at a single criterion for classification. In the scope of this thesis, this point of view on registration and segmentation is not appropriate, because the goal of this thesis is to develop a method for organ segmentation *and* registration. The classification scheme used to select appropriate methods should therefore be based on the problem statement to be solved.

The shape driven taxonomy for segmentation methods presented in the previous section allows for a new view on segmentation and registration methods. This new view resolves the ambiguity of current segmentation and registration classification schemes by proposing a new class of algorithms which will be called *regmentation* in the following.

The introduction of global shape knowledge in segmentation as described in Section 2.3.1.4 plays a key role in regmentation. Since information about the shape of one or multiple structures is available in a global context, global prior segmentation methods can also be used to solve the

problem statement of *registration* for these structures. Figure 2.14 illustrates this process. For example, model based segmentation uses geometric shape models that are adapted to objects which are visible in the image. The models are constructed based on template shapes that all have been aligned in a certain coordinate system. The set of template shapes defined in this coordinate system could therefore be regarded as an *atlas*. Furthermore, the geometry of the model is globally defined, that means, for example, the lower peak of the left cardiac ventricle stays the lower peak after adaptation and is not adapted to the top border of the ventricle. This means, if a model from the atlas is adapted to two different data sets, a correspondence between both data sets is given in the adapted state. Naturally, a correspondence between the model in the atlas and the adapted model is also given. Therefore, a mapping between both data sets can be established by using the correspondence chain from the adapted model in the first data set to the atlas and back to the adapted model in the second data set. Such a mapping is in fact a registration. Global prior segmentation methods are therefore both segmentation and registration methods. In other words, they are *regmentation* methods, because they cannot be differentiated from the problem statement solving point of view.

The presented view of regmentation leads to a new classification scheme for segmentation and registration methods that consists of three categories. The first category is represented by pure *segmentation* methods. According to the scheme presented in in Figure 2.11, these are voxel and region based segmentation methods as well as local shape prior methods. These approaches can solely be used to solve a segmentation problem, since they lack of global knowledge. No correspondence establishment and therefore no registration is possible. The second category consists of pure *registration* methods. These methods are classical registration approaches that are able to establish a mapping between two data sets, but lack an atlas in which structures are segmented. If such an atlas is added, they become *regmentation* methods which form the third category. This category consists of all global prior segmentation approaches.

The benefit of the described scheme is that the built categories are disjunct. All segmentation and registration algorithms can be classified as being exclusively part of one category. Furthermore, the proposed classification is a problem solving driven scheme, that means, it does not matter, which kind of technique is used to solve the problem. Appropriate techniques for solving a segmentation *and* registration problem can be directly taken from the regmentation category. The applicability of the proposed scheme is demonstrated in the following section by a comprehensive classification of articles from international medical imaging journals and conferences of the last two years (2010-2011).

2.3.3. Classification of existing methods

This section demonstrates the practical applicability of the taxonomy presented in Section 2.3.1 and Section 2.3.2 in comparison to the classical view of segmentation and registration methods. It also aims at demonstrating the current direction of research in the field of medical image segmentation and registration. 855 articles from 6 renowned international medical imaging journals and conferences of the last two years (2010-2011) have been investigated and classified [2]. The conferences and journals were mainly chosen by their impact factor and relevance to the medical

[2] The work of classifying state of the art methods from international journals and conferences of the last two years following the developed classification scheme presented in Section 2.3 has been shared with Sebastian Steger, Fraunhofer IGD.

imaging community.
The conferences investigated have been:

- Medical Image Computing and Computer-Assisted Intervention (MICCAI), in the year 2010 (96 articles). The MICCAI conference is one of the most renown conferences in the field of medical imaging.
- IEEE Conference on Computer Vision and Pattern Recognition (CVPR), in the years 2010 (13 articles) and 2011 (6 articles). CVPR is one of the top ranked conferences in the field of Computer Vision and Pattern Recognition. Usually a significant amount of articles in the field of medical imaging are published every year.
- IEEE International Symposium on Biomedical Imaging (ISBI), in the years 2010 (126 articles) and 2011 (220 articles). The ISBI conference is one of the biggest biomedical imaging conferences and contains comprehensive tracks on segmentation and registration regarding all imaging modalities.
- SPIE Medical Imaging, in the years 2010 (132 articles) and 2011 (116 articles). SPIE Medical Imaging is one of the biggest conferences which focus solely on advances in medical imaging.

The journals chosen for the study were:

- IEEE Transactions on Medical Imaging, volume 29, issues 1-12 in 2010 (56 articles) and volume 30, issues 1-8 in 2011 (33 articles). Transactions on Medical Imaging has an impact factor of 3.5 (2010) and is one of the most renown medical imaging journals.
- Medical Image Analysis, volume 14, issues 1-8 in 2010 and volume 15, issues 1-4 in 2011. Total articles: 57. Medical Image Analysis has an impact factor of 4.2 (2011) and is one of the most renown medical imaging journals.

Figure 2.15 (a) shows the classical four categories of segmentation and registration from the literature. The category *segmentation* denotes all segmentation approaches presented in Section 2.3.1 including voxel and region based methods as well as local and global prior segmentation approaches. The category *registration* consists of pure registration methods and is identical to the registration category of the scheme proposed in Section 2.3.2. The category *segmentation for registration* is formed by segmentation methods that are used to solve a registration task. The category *registration for segmentation* is formed by methods that address the inverted task. In order to ease the comparison of the classical with the proposed classification scheme, the described four classical categories have been made disjunct. Here, the last two categories had priority, that means, a method that would have been assigned to be a *segmentation for registration* method and a *segmentation* method is considered as a *segmentation for registration* method. The category *registration for segmentation* is treated analogous. As it can be seen in Figure 2.15 (a), in the classical scheme, the majority of investigated methods are segmentation methods followed by pure registration approaches. The hybrid categories consist of comparable few articles. This view suggests that the intersection set of segmentation and registration methods is relatively small. However, as it will be shown in the following, the opposite is true.

In the four described classical categories, the segmentation approaches are now further investigated by applying the proposed sub-classification of Figure 2.5. That means, in each classical category, the segmentation methods are classified as being either voxel based methods, region based methods, local prior or and global prior shape methods. Moreover, the sub-category *global prior shape methods* lists voxel atlas based segmentation approaches separately, since their incorporation of global shape knowledge significantly differs from the other global prior shape methods. The subdivision of the classical categories will answer the question of how much shape

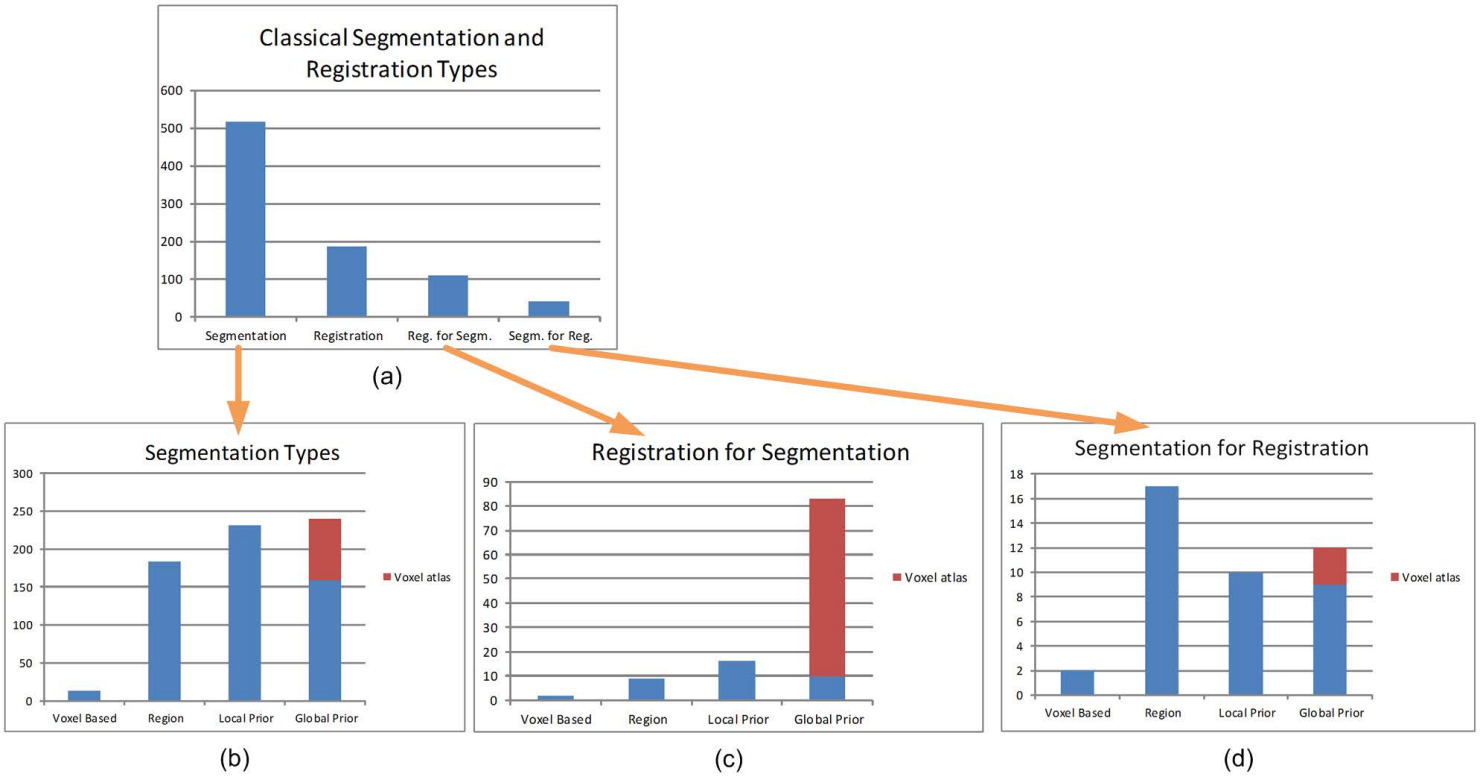


Figure 2.15.: (a) Amounts of articles from recent selected conferences and journals (cf. Section 2.3.3) classified as either segmentation, registration, segmentation for registration or registration for segmentation [ESS12]. The term segmentation for registration denotes work that target at solving a registration problem and make use of segmentation methods for guiding the registration, for example by segmenting corresponding areas in the images. The term registration for segmentation denotes the inverted problem: solving a segmentation problem guided by a registration as it is used for example in voxel based atlas registration. (b-d) Sub-classification of the categories of (a) which contain segmentation methods into the four segmentation types voxel based, region based, shape based with local priors and shape based with global priors as defined in Figure 2.5. Within the global prior methods, voxel based atlas approaches are denoted separately, since they operate at voxel level in comparison to most other global shape prior methods.

knowledge is used in current state of the art approaches. Finally, it will reveal that the classical classification scheme can be easily transformed into the proposed classification.

Figure 2.15 (b) shows the distribution of methods in the classical category *segmentation*. It can be seen that most of the current segmentation approaches make use of prior shape knowledge. Purely voxel based methods are only used in very few methods whereas the amount of region based approaches is comparable to local or global prior shape methods.

In Figure 2.15 (c), the segmentation methods used in the classical category *registration for segmentation* are sub-classified. Global prior shape methods form by far the largest class. All other classes are only represented by a small amount of methods. Among the global prior methods, voxel atlas based approaches are the most dominant technique.

Figure 2.15 (d) shows the distribution of segmentation approaches in the classical category *segmentation for registration*. Again, the majority of methods use prior shape knowledge while the largest single category is formed by region based methods. This is due to the fact that many registration methods rely on a detection of feature points in the image which in turn are often segmented using region based approaches. Voxel based methods only play a minor role in this category.

By building the proposed sub-categories of segmentation methods in each classical category, the classical classification view can be transferred into the scheme proposed in this thesis. Here, all global prior based methods from each classical category are moved to the new category *regmentation*, because they can be used to address segmentation *and* registration problems. The classical hybrid categories *segmentation for registration* and *registration for segmentation* are moved to the categories registration and segmentation, respectively, because they address either registration *or* segmentation problems. Figure 2.16 shows the resulting classification. In comparison to the classical scheme in Figure 2.15 (a), it is now evident that a large amount of recently published methods can be used to address both segmentation *and* registration problems. In fact, more articles about regmentation techniques have been published than classical registration approaches which shows that regmentation plays a key role in current medical imaging research.

2.3.4. Discussion

In this section, a new shape knowledge driven taxonomy for the classification of segmentation and registration methods has been presented. For the goal of this thesis — the development of an automatic volumetric organ segmentation and registration approach — several consequences can be drawn. Global prior segmentation methods are well suited for the segmentation of structures with complex shapes like organs. Moreover, through the incorporation of global shape knowledge, they are the best available technique for segmenting certain organs like the liver [CC07, HvGSea09] which are often not well separated from neighboring tissue. Those facts are also supported by the literature classification in Section 2.3.3 which showed that global prior shape methods are dominating the current research in the field of medical image segmentation.

The developed taxonomy also revealed that global prior shape methods are suited to solve both segmentation and registration problems. In other words, segmentation techniques become registration techniques if enough shape knowledge is incorporated. This makes global prior shape methods very well suited to address the goal of this thesis.

Among the global prior shape methods, two main categories have been identified: geometric model based segmentation approaches and voxel atlas based techniques. As it has been shown,

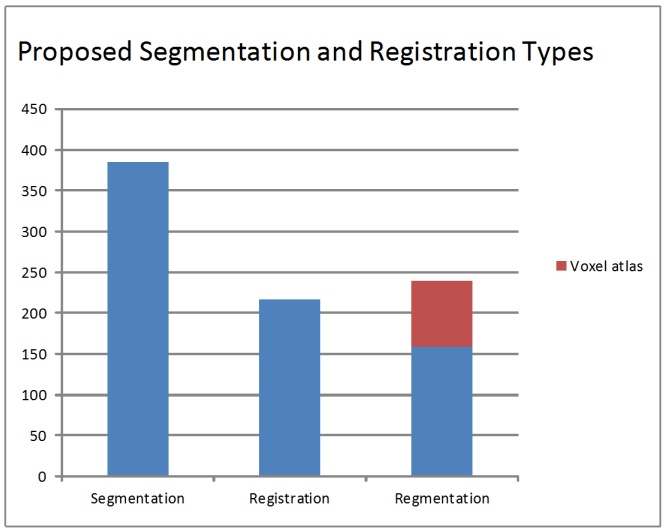


Figure 2.16.: Proposed classes of segmentation and registration approaches according to the taxonomy developed in this thesis. The *segmentation* category consists of methods that can be used to solve segmentation problems. The category *registration* classifies methods for registration tasks. Approaches that can be used to solve both problems fall into the category *regmentation* [ESS12].

voxel atlas based segmentation incorporate the highest amount of global shape knowledge, because they implicitly model relations and dependencies between different labeled structures to be segmented. However, it also has been shown that voxel atlas based methods are less flexible to segment complex soft tissue structures with a lot of shape variation (cf. Section 2.3.1.4). Therefore, they are in practice limited to the segmentation of certain anatomical regions like the head and neck region where the amount of variation is relatively low and a small amount of images is sufficient to build a complete atlas. Another often used application is to use a voxel atlas as a coarse initialization for other methods where high accuracy is not needed. Due to these limitations voxel atlas based segmentation is not considered for use in this thesis.

Geometric model based approaches are suited for addressing the goal of this thesis, because they can better cope with complex shape variation (cf. Section 2.3.1.4). However, as it has been shown, open issues in geometric model based approaches exist. One problem is that the shape variability of highly shape varying organs like the liver or the pancreas can not be sufficiently learned. Here, an extremely large number of representable reference shapes would be necessary for a complete modeling. However, even if enough reference shapes would be available, current statistical shape modeling methods are not able to retain all detail of the training set, because dimensionality reduction techniques are used to reduce the complexity of the shape modeling process. However, this prevents current geometric model based approaches to robustly segment shapes that differ a lot from the trained reference shapes.

Another issue of model based segmentation concerns the shape representation of the structures to segment. The majority of current state of the art model based methods use models based on surface representations, i.e. a 3-dimensional surface model is evolved towards the boundaries of the target organ or structure. However, it often makes sense to incorporate *volume information* during the adaptation process in order to make the method more robust against model initialization errors and to increase the robustness through detection of organ foreign tis-

sue that may not be detected when considering image information near the model's surface only. Different kinds of volumetric segmentation approaches have been proposed in the literature [TFCT98, PFJ*03, HM08, BPCO10, SLH11]. However, those methods are often complex and computationally expensive. Moreover, existing standard adaptation and shape preserving strategies cannot be used within those models or must be adapted. In Section 2.5.2, a detailed outline of existing volumetric model based approaches is given.

In order to give a deeper insight into the strengths and drawbacks of current state of the art geometric model based approaches, the following sections discuss the three basic elements of geometric model based segmentation. Dimensionality reduction and machine learning techniques used for modeling a set of training shapes or used for automatic model initialization methods are outlined in Section 2.4. Shape representation and preservation techniques are described in Section 2.5. Boundary detection methods are presented in Section 2.6.

2.4. Machine learning approaches

Machine learning methods have become very popular in the medical imaging community over the last decade. Pattern recognition techniques like clustering, support vector machines, principle component analysis or boosting help to cope with the analysis of nowadays massive amount of data produced with current medical imaging techniques. It is also a key enabler for coping with increasingly complex segmentation and registration techniques which are usually defined by a multitude of parameters. Determining those parameters for a given problem statement in a heuristic manner is time consuming and error prone. Machine learning techniques allow to estimate the optimal parameters by learning from previously seen data of the same problem statement. For example, in order to determine an appropriate threshold for segmenting the skull in a CT image, a human could investigate a couple of images trying different thresholds until the skull is sufficiently well separated from surrounding tissue. As the number of images increases, it will become more and more difficult to find the optimal threshold since a human usually has to iterate over all images again when trying a new value. As an alternative, the optimal threshold could be found using machine learning approaches.

A detailed introduction to the field of pattern recognition and machine learning methods can be found in the book by Christopher Bishop [Bis06]. This section describes the pattern recognition based approaches that have been used in this thesis. In particular, principle component analysis (PCA), boosting and binary decision trees are explained.

Machine learning approaches usually start with a set of input values, the so called *feature vector*. The goal is to assign the feature vector $\vec{v} \in D$, with D being the input domain, to one or multiple classes $c \in C$ based on the input values. This process is called *classification* if the classes are discrete, i.e. a discrete value is assigned to the feature vector for every class. If the classes are continuous, the process is called *regression*. In medical image segmentation, it is usually sufficient to work with two discrete classes $c \in \{0, 1\}$, for example organ and non-organ so the focus of this section lies on binary classification methods.

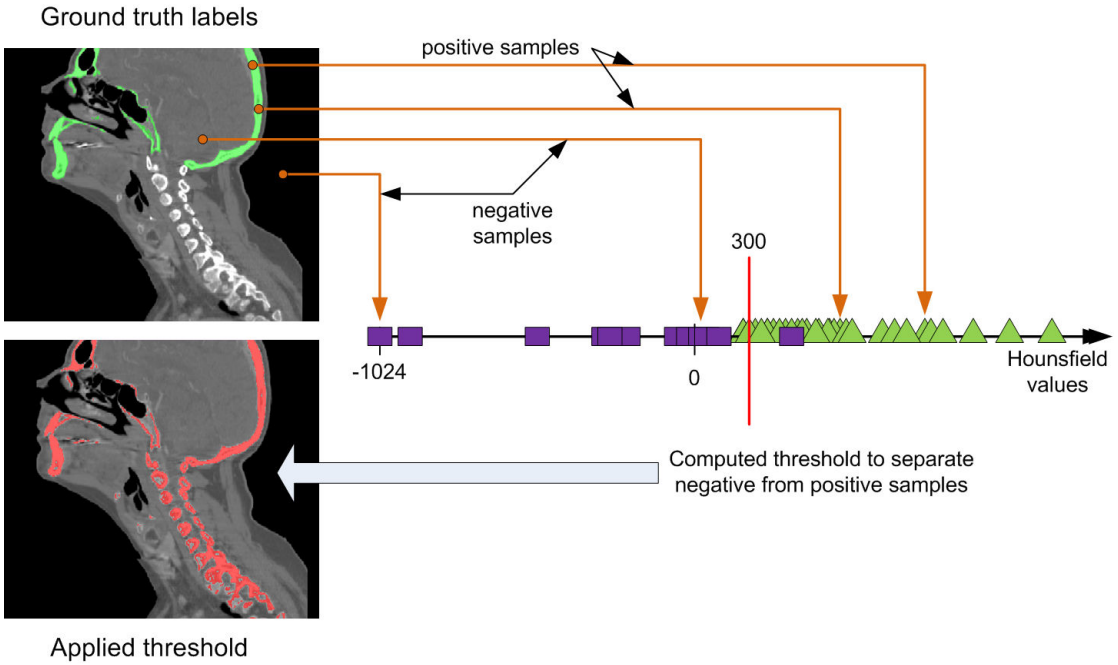


Figure 2.17.: Example of supervised learning for determining a threshold for skull segmentation in CT. From a labeled image, positive samples inside the labeling and negative samples outside the labeling are taken. Supervised learning determines a threshold that separates positive and negative samples. The found threshold is applied to the same image. The skull is in most areas well segmented in comparison to the ground truth labeling.

2.4.1. Supervised and unsupervised learning

There are two important categories of binary machine learning classification approaches frequently used in medical imaging: *supervised* and *unsupervised* methods.

Supervised learning In supervised learning, a set of n positive feature vectors $T^+ = \{(\vec{v}_1, 1), .., (\vec{v}_n, 1)\}$ and a set of m negative feature vectors $T^- = \{(\vec{v}_{n+1}, 0), .., (\vec{v}_{n+m}, 0)\}$ are built. The goal of the learning approach is now to find an appropriate function $g : D \rightarrow C$ from the input domain D to the output domain $C = \{0, 1\}$ to correctly classify all feature vectors as positive or negative. In the example of finding a threshold to segment the skull in a CT image, the input domain would be the 12 Bit Hounsfield range of gray values $D = -1024, .., 3071$. In this case, one feature vector $\vec{v}$ would be computed for every voxel and would contain only a single value — the gray value of the voxel. The assignment of the feature vectors to the positive and negative sets T^+, T^- could be done manually, by delineating the skull in a couple of CT images, i.e. by labeling every voxel as skull $c_1 = 1$ and as not skull $c_2 = 0$. A possible function $g : D \rightarrow 0, 1$ could be:

$$g(\vec{v}) = \begin{cases} +1 & if \quad v_1 \geq 300 \\ 0 & else \end{cases} \tag{2.6}$$

In this case the threshold for segmenting the skull would be 300 Hounsfield units which may be sufficient to correctly assign the majority of training feature vectors to their class. In Figure

2.17, an example of the described process is given. It is very unlikely though that a function g exists that performs the correct classification for all training vectors. Furthermore, there will be probably an amount of training vectors which contain the same Hounsfield value, but were assigned to different classes. This is due to the fact, that the classification problem in this example — segmenting the skull — is too complex to address using a single feature — a threshold — only. For example, in Figure 2.17, the spine is also segmented by applying the computed threshold. The aim of a supervised learning approach therefore is to find the function g that solves the classification problem using the available features in the best approximate way. What the best way is depends on the problem statement. For example, it could be desirable to maximize the distance between the centers of the positive and negative feature vector clusters. Other problems may require to have a low false positive rate, for example if the classification task concerns tumor detection.

Popular supervised learning methods are support vector machines, classification and regression trees or boosting methods.

Unsupervised learning In unsupervised learning, the input feature vectors used for training are unlabeled, i.e. there is only one training set $T = \{\vec{v}_1, .., \vec{v}_n\}$. That means, the function $g : D \to C$ from the input domain D to the output domain $C = \{0, 1\}$ cannot be estimated using the training feature vectors alone. Instead, there needs to be some a-priori assumption about the structure of the input data. For example, the assumption could be that there exist two classes of feature vectors (describing organ and not organ) and that the feature vectors of the same class are similar to each other. A metric defining the similarity could simply be the euclidean distance between the vectors in feature space. So the unsupervised learning algorithm would take the input vectors and group them in two classes such that the metric is optimized. In that way any feature vector can be classified to be in one of this two built classes.

Unsupervised learning methods include clustering methods like k-means clustering, dimensionality reduction techniques like principle component analysis and neural network models. However, some of those algorithms can also be adapted to work with labeled training data making them supervised techniques.

The advantage of unsupervised methods in comparison to supervised methods is that they do not need labeled training data. In terms of medical imaging this is an important aspect since here typically huge amounts of data is produced and the training data often has to be created manually which is very time consuming. Furthermore, since the labeling task itself is for many anatomical structures not trivial, often only medical experts are able to perform this task. However, in unsupervised learning, it is often not easy to define proper a-priori assumptions about the structure of the data. Sometimes the statistical nature of the data is not known so standard models like Gaussian data distribution are used which often are not the best solution.

Supervised learning methods have the advantage that the classification method can be trained to approximate the training data arbitrarily precise without knowing the statistical structure of the data[3]. The idea is that the more training data is available, the more accurately the method will also be able to classify unseen data of the same type. However, there are some practical limitations of this idea. First, it is often not possible to acquire enough training data for the problem statement, which might be complex and involving a lot of statistically independent parameters. Secondly,

[3] It is of course advantageous to know the statistical distribution of the data in order to choose an appropriate model for estimating the function $g : D \to C$.

supervised learning targets at precisely classifying the *training data*. A problem that arise in this context is *over-fitting*. Over-fitting means that the classifier is able to correctly classify the training data, but weakly performs on unseen data. This typically occurs if some of the chosen features are not statistically relevant to describe the problem. For example, using the Hounsfield value at a voxel as a feature for classification in skull and not skull in a CT image is a good feature, because the gray values of the skull are always inside a certain gray value range in CT images. This feature is therefore descriptive. However, using the same feature to classify a voxel in skull and not skull in a magnetic resonance image is a bad feature. In contrast to the normalized Hounsfield scale the absolute gray values in magnetic resonance imaging do not have a significance for describing a structure and may be completely different from one scan to the other. If the gray value is used as a feature, the classifier still might find some similarities for the skull in the training data. However, this similarity is most likely completely random if transferred to unseen images. In this case, the classifier is over-fitted and therefore worthless for a general application. A way to test whether a classifier is over-fitted is to compare the training error with the error in unseen data. If the error for the unseen data is significantly higher than for the training data over-fitting is likely.

Another source for over-fitting is the selection of non representative training data. For example, if many pathologic cases are used to train a classifier for detecting the hip bone, the classifier will likely perform badly on healthy bones. Furthermore, it is important that the labeling of the training data is consistent in all images. For example, if in the case of skull segmentation the nose is included in some training data and not in some others images, the performance of the classifier will suffer. A way to see if the data used for training is representative is cross fold validation where the training and the test data is mutually changed and the classification errors are compared. If the errors differ significantly, the training data is likely to be not representative for the given problem statement.

2.4.2. Binary decision trees

Binary decision trees are a well known concept in computer science and they are widely used in the context of supervised learning mainly because of their simplicity, their performance and their robustness for predicting useful outcomes based on sparse data. Popular types of decision trees are CHAID (chi-squared automatic interaction detection) [Kas80], classification and regression trees [BFOS84] or QUEST (quick, unbiased, efficient, statistical tree) [LS97].

The idea of binary decision trees in supervised learning is the following: As explained in Section 2.4.1, a set of n positive feature vectors $T^+ = \{(\vec{v}_1, 1), .., (\vec{v}_n, 1)\}$ and a set of m negative feature vectors $T^- = \{(\vec{v}_{n+1}, 0), .., (\vec{v}_{n+m}, 0)\}$ is built. All feature vectors have the same number k of entries (also called features). The first node in the decision tree — the root — is built by iterating over all training vectors $\vec{v} \in \{T^+, T^-\}$ and search for the feature v_i that is best suited for correctly grouping the training vectors into the classes $C = \{0, 1\}$. For example, let us assume the goal is to segment the skull in a CT image and $k = 2$ features are used for every feature vector $\vec{v}$. Those features could be the Hounsfield unit at a voxel and the image gradient at a voxel. That means for every voxel in the image a feature vector with those two entries is created. From the manual labeling it is known for every vector whether it belongs to $c = 0$ or $c = 1$. Over all training vectors, the feature is selected to perform the best possible split compared to the ground truth sets T^+ and T^-. Most likely, the Hounsfield units will be that best feature, because the skull has a very characteristic intensity in CT images. That means, in the root node, a threshold for the Hounsfield unit is selected, for example 300 that splits the training data into two parts. In an ideal situation,

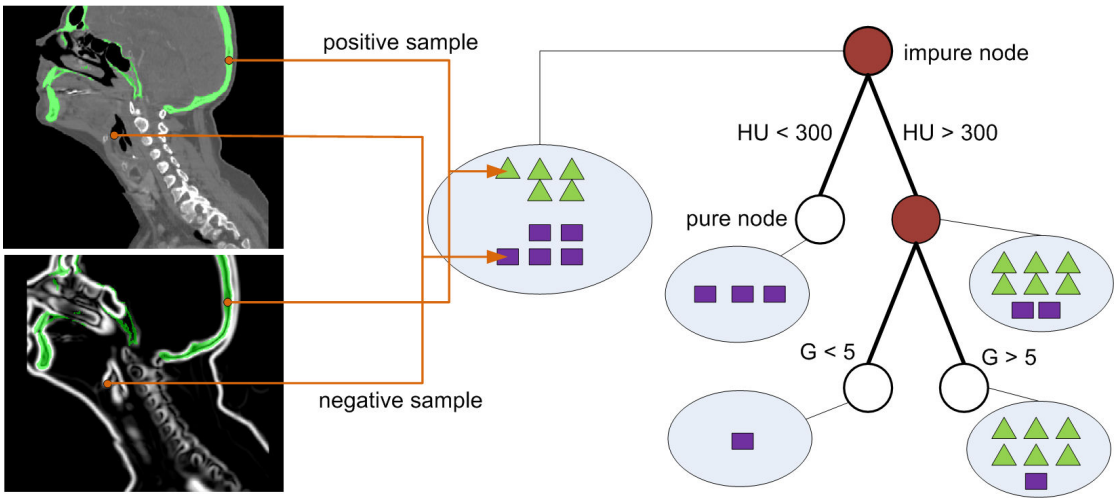


Figure 2.18.: Example of binary decision trees used for skull segmentation in CT. Positive and negative tuple samples are taken from labeled intensity and gradient images. A binary decision tree is built. The first split is done by using an intensity threshold of HU = 300 to separate positive from negative samples. One resulting node is considered impure and therefore split again. This time the split is done using the gradient feature G = 5 which results in two pure nodes. The segmentation can be obtained by traversing the tree for every voxel in the image.

this split is already sufficient to correctly classify all training vectors. In that case, the second features — the image gradient — would not appear in the tree and therefore would not play a role for the decision procedure when testing unseen data. This is one of the strengths of decision trees. They aim at selecting the most significant features for a proper prediction and neglect features that have only low predictive value. However, in the example of skull segmentation, it is unlikely that the image intensity alone is sufficient to correctly classify all training vectors, so some of the training vectors belonging to $c = 0$ would be classified as $c = 1$ and vice versa. The node is therefore called impure. A popular measure for this impurity is the Gini coefficient [BFOS84] which is 0 in case all vectors in the node fall into the same class. In case of an impure node, it is recursively split into two new nodes followed by again selecting the best feature to split the data. The procedure stops if a node after a split has a sufficient purity. In order to avoid over-fitting and to keep the tree reasonably small, additional criteria are used in praxis. Usually the maximum depth of a tree is specified a-priori. Also splitting is stopped if the number of training vectors in a node is smaller than a pre-defined threshold. If the number of training vectors in a node is low, the statistical significance of any selection criteria is low and further splitting would probably lead to over-fitting. Sometimes it happens that a split does not generate a better discrimination than a random choice. In that case further splitting is avoided as well, since the data obviously cannot be separated any more using the given features. After creation of the tree, any unseen feature vector can be classified by traversing the tree and following the splitting criteria in every node.

Binary decision trees have many advantages. They apply a simple set of splitting rules which are easy to understand by humans. This makes binary decision trees a *white box method*. Furthermore, they are relatively fast to train and very performant on classification of large datasets. By automatically selecting appropriate features for tree splitting, binary decision trees avoid over-

fitting. By selecting additional stopping criteria, they can also be flexibly adapted to different application scenarios, for example, if a certain runtime or accuracy is required.

2.4.3. Boosting

Boosting is a frequently used method to improve the quality of supervised machine learning techniques. Its idea is to combine several single classifiers to create a committee that is more powerful than each single classifier. In this context the single classifiers are usually called *weak classifiers*. A popular choice for weak classifiers are binary decision trees as described in the previous section.

There are several variants of boosting [Bis06]. One of the most popular boosting algorithms is adaptive boost (AdaBoost) proposed by Freund and Schapire [FS96]. Consider a set of n positive feature vectors $T^+ = \{(\vec{v}_1, 1), .., (\vec{v}_n, 1)\}$ and a set of m negative feature vectors $T^- = \{(\vec{v}_{n+1}, -1), .., (\vec{v}_{n+m}, -1)\}$ with $n+m = N$ and $T = \{T^+, T^-\}$. For each feature vector, a weight w_i exists that is set to $\frac{1}{N}$ at the beginning of the algorithm. Furthermore, a set of M weak classifiers $\{c_1(\vec{v}), .., c_M(\vec{v})\}$ exist which map the input vectors to $\{-1, 1\}$. The weak classifiers are now trained one by a time on the positive and negative feature vectors weighted by w. After training of the first classifier c_1, its classification error J_1 on all feature vectors is computed as

$$J_1 = \sum_{j=1}^{N} w_j I(c_1(\vec{v}_j)), \tag{2.7}$$

where $I(c_1(\vec{v}_j)$ is a function that yields 1 if the classification of $\vec{v}_j$ is correct and 0 otherwise. The error J_1 is now divided by the sum of all weights w which gives the weighted error

$$\varepsilon_1 = \frac{J_1}{\sum_{j=1}^{N} w_j}. \tag{2.8}$$

The weighted error is now used to weight c_1 by

$$\alpha_1 = ln\left\{\frac{1-\varepsilon_1}{\varepsilon_1}\right\}. \tag{2.9}$$

So far, the first classifier c_1 has been trained on the data set and — depending on its classification error — has been weighted by α_1. α_1 is now used to update the weights w^1 to become the weights w^2 as

$$w_i{}^2 = w_i{}^1 exp(\alpha_1 I(c_1(\vec{v}_i))). \tag{2.10}$$

That means, the weights for misclassified training vectors are increased so the next classifier c_2 will put emphasis on correctly classifying those vectors. The process of creating the next classifier, computing its weighted error and updating the weights continues until all weak classifiers have been built. Afterwards, $\{c_1(\vec{v}), .., c_M(\vec{v})\}$ are combined to form the strong classifier

$$C(\vec{v}) := sign\left(\sum_{i=1}^{M} \alpha_i c_i(\vec{v})\right). \tag{2.11}$$

C is a weighted vote of all weak classifiers. A positive sign means that a vector $\vec{v}$ is part of the positive feature class and a negative sign means the inverse.

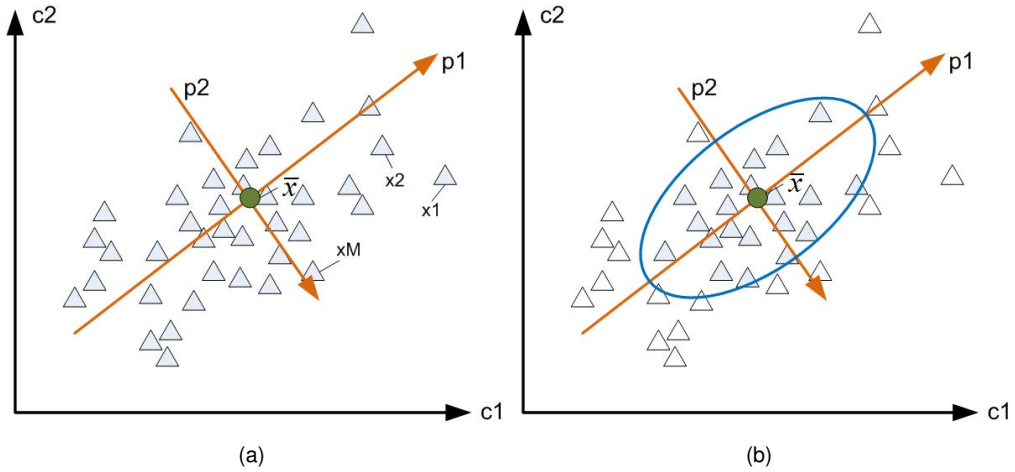


Figure 2.19.: (a) Principle component analysis on sample data $\{x_1, \ldots, x_M\}$. The coordinate system with the axes c_1, c_2 is transformed into the orthogonal coordinate system p_1, p_2. p_1 is aligned to the data such that it accounts for the highest variance. p_2 accounts for the second highest variance. (b) Samples in space p_1, p_2 can be restricted to have a maximum distance to the mean $\bar{x}$ on each principle component, that means they are projected into the shown ellipe.

2.4.4. Principle component analysis

Principle component analysis (PCA) is a mathematical technique widely used in medical imaging in the context of supervised learning. Particularly, in geometric model based segmentation, it is often used for dimensionality reduction of a set of reference shapes such that an arbitrary input shape can be efficiently compared to the reference shape set.

Principle component analysis uses an orthogonal coordinate transformation to transfer a coordinate system of possibly correlated coordinate axes to a coordinate system of uncorrelated coordinate axes given a number of data samples defined in the correlated coordinate system. The uncorrelated coordinate axes are called principle components. The principle components are orthogonal to each other and also ordered. They are defined such that the data samples have the highest variance along the first principle component. That means, along this axis, the largest distance between two data samples occurs (cf. Figure 2.19(a)). The second principle component maps the second highest variance and so on. Note that the data has to be normally distributed in order to guarantee that the principle components are uncorrelated.

Principle component analysis is often used for dimensionality reduction. Since the principle components are ordered by the amount of variance of the sample data they map, the idea is to discard some of the last principle components that only account for a small amount of variance in the data. The result is a coordinate system of lower dimension, that means, the same data can be represented by fewer variables which is beneficial for many applications. In geometric model based segmentation, PCA is used to reduce the dimensionality of a set of reference shapes in order to efficiently compare an arbitrary input shape with the reference shape set. This process is called statistical shape modeling and is explained in detail in Section 2.5.3.

Mathematically, the principle components can be found by computing the eigenvectors and eigenvalues of the covariance matrix of the data [DTT08]. Consider a set of data samples defined

in the same N-dimensional space. A single data sample is represented as the vector $x_i \in \mathbb{R}^N$. A total of M sample vectors $\{x_1, \ldots, x_M\}$ exist. The mean

$$\bar{x} = \frac{1}{M} \sum_{i=1}^{M} x_i \tag{2.12}$$

and the covariance matrix

$$C = \frac{1}{M-1} \sum_{i=1}^{M} (x_i - \bar{x})(x_i - \bar{x})^T \tag{2.13}$$

are computed. The eigenvectors $p_1, \ldots, p_N$ of C and their corresponding eigenvalues $\lambda_1 \geq \ldots \geq \lambda_N$ are obtained through diagonalization of C. Here, the eigenvectors are ordered by decreasing eigenvalues. The eigenvalue λ_i describes the variance along the principal component p_i.

In order to reduce the dimensionality of the principle component space, axes that account only for a small variance can be excluded. For example, the smallest dimension t can be computed such that $\sum_{i=1}^{t} \lambda_i$ captures a certain percentage of the variance of the sample data. Depending on the application, this value can be set for example to 90%.

Using the reduced principle component space, an arbitrary sample $\hat{x}$ in this space can be expressed as

$$\hat{x} = \bar{x} + Pb \tag{2.14}$$

where $P = (p_1 | \ldots | p_t)$ is a matrix that contains all remaining eigenvectors that have not been discarded by the dimensionality reduction. $\vec{b}$ is a t-dimensional parameter vector with parameters b_i. This vector of reduced dimensionality now represents $\hat{x}$.

The principle component space can not only be used for dimensionality reduction, but also for an easy way to compare input samples to the mean of the whole data set. Since the variance of the data along each principle component is given, a sample can be considered similar to the mean if its distance to the mean is, for example, three standard deviations on each principle component. In principle component space, this spans an ellipse inside which all samples are considered similar to the mean (cf. Figure 2.19(b)). Moreover, a sample that is outside of the ellipse can be projected to the borders of the ellipse. This can be done as follows. The parameters b_i in (2.14) can be restricted to be in a certain interval, whereas the interval defines the extents of the ellipse. An interval could be, for example, $[-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]$ which means, that the radii of the ellipse are three times the standard deviation of the data along each principle component.

2.5. Shape representation and modeling strategies

This section describes different state of the art strategies to represent and to model shape in model based segmentation. Generally, shape can be represented in different dimensions. For example, the contour of an organ in a two-dimensional image can be modeled by a one-dimensional spline. In a three-dimensional image, the surface of an organ can be modeled by a two-dimensional representation.

2.5.1. Shape representation

In Medical Imaging, mainly two kinds of shape representations are used: implicit shape representations and explicit or parametric shape representations. Implicit shape representations

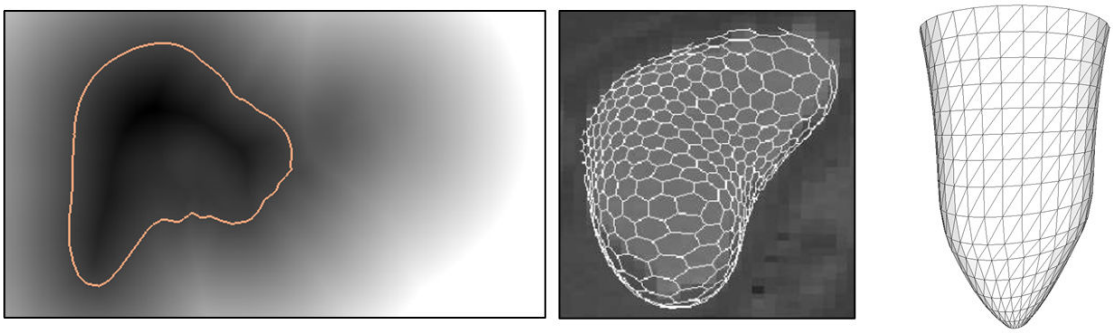


Figure 2.20.: Different representations of shape. From left to right: liver shape represented by zero entries in a distance map as it is used in level set segmentation. A bladder represented by a point cloud that is connected using simplex meshing (image taken from [CDA07]). A polygonal tessellation of a point cloud representing the cardiac left ventricle.

define the shape of the model through an embedding function while in explicit shape representation the shape is defined by a parametric representation such as a spline or a polygonal mesh. Both representations are widely used in medical image segmentation [HvGSea09] and both exhibit advantages and disadvantages. In the following, the two most popular implicit and explicit representations are outlined: level set methods and interconnected point clouds.

Level set methods The application of level sets in medical image segmentation has been proposed in the early 90's in the context of two-dimensional contour evolution by Malladi et al. [MSV94, MSV95] and Caselles et al. [Cas95]. Since then level sets have been extended to three-dimensional image segmentation and are nowadays applied to all kinds of medical segmentation problems. A detailed introduction into the field of different level set methods and their optimization can be found in the reviews by Suri et al. [SLS*02] and Cremers et al. [CRD07]. The idea of level sets is to represent the shape of an object implicitly by using a function that is of higher dimension than the representation of the object. For example, a three-dimensional surface can be represented by a four-dimensional function whereas one parameter of the four-dimensional function has a certain value, usually zero. The three dimensional surface is then called the zero level set of the higher dimensional function. In practice, a shape at a certain point in time during the adaptation is represented by a distance map and the zero level set or boundary of the shape is defined by the zero entries in this map (see Figure 2.20 left). A drawback of this representation is that signed distance maps do not form a linear space, which may result in non plausible shapes during adaptation.

The embedding of a shape in a higher dimensional space has the advantage that the shape may change its topology during adaptation. This dynamics make level set approaches very generic regarding the kind of structure to segment. In fact they have been applied to segment cellular images [WS11], genus-1 surfaces like the vertebra [AFAX11], thin surfaces like the bladder wall [CBMS11] as well as solid and quasi-solid organs like the liver [WSH09, KUA*09, CCBK07, FSK07] and the prostate [TYW*03]. Because of their flexibility, they are also very popular in vessel segmentation, for example in the context of carotid arteries [UAW*11].

Interconnected point clouds Many explicit shape representation are based on a set of points — also called control points or landmarks — which are somehow interconnected, for example by a spline or a polygonal tessellation. One of the first who proposed to use such interconnected point clouds for image segmentation were Terzopoulos et al. [TF88] with the introduction of snakes (cf. Section 2.3.1.3). They proposed a method for adapting two-dimensional contours which were defined by a set of points and adapted to the data set using image driven forces — the external forces — and shape preserving forces — the internal forces. The method has been later adapted to three-dimensional image segmentation [MT96] and is now used in several variants for segmentation of organs like the liver [KLL07, HMW07, LZZ*08, SBB09, HvGSea09] or the heart [ZBG*07, EPS*08, ZHB06]. Most of current approaches use a polygonal tessellation [KLL07, HMW07, LZZ*08, ZBG*07, EPS*08, ZHB06] or a simplex mesh [CDNA07] to interconnect the point clouds (cf. Figure 2.20 middle and right).

Comparison of implicit and explicit shape representations Both implicit and explicit shape representations have their advantages and drawbacks. Implicit representations can naturally deal with topological changes which is not straightforward to achieve with explicit representations. They are less dependent on initial positioning and do not have problems like self intersection. On the contrary, implicit representations are computationally complex and the mathematical framework of front evolution driven by partial differential equations makes it difficult to integrate local shape constraining functionalities. Explicit representations do not have such problems, since they are efficient to optimize and easily locally controllable through manipulation of control points. As noted by Suri et al. [SLS*02], implicit representations like level sets cannot deal with gaps in boundaries, as they are prominent in imaging modalities like ultrasound or MR. Furthermore, so called shocking problems may occur on places in the image like corners, protrusions and indentations.

Advantages of explicit representations are their low computational optimization cost, their ability to easily incorporate different shape preserving strategies through control point manipulation. The surface of an explicit shape is usually smooth which makes it possible to avoid leakage into gaps caused by the imaging modality. However, this makes them also less flexible. Furthermore, the amount of local adaptation is limited by the amount of control points. In order to alleviate this problem re-meshing strategies have to be applied, for example in order to increase the amount of control points in areas of high curvature. Another drawback of explicit shapes is self-intersection. Usually explicit shapes are adapted individually for every point and afterwards optimized to fulfill some smoothness criteria. This makes it very difficult to detect self-intersections of the surface during adaptation which is a problem for shapes that contain a lot of protrusions and indentations. Furthermore, explicit shapes are in praxis limited to segment genus-0 shapes. Although there exist approaches that can deal with higher order genus types [BP05] implicit shape representations are used in the vast majority of cases, because topological changes are easy to handle.

Many drawbacks of implicit representations are the strengths of explicit representations and vice versa. As noted before, implicit shape representations are commonly used for all kinds of segmentation problems including totally different structures like vessels, solid organs and other soft tissue. Explicit shape representations are mainly used for solid or quasi-solid organs like the heart, liver, bladder, prostate or lungs. For those organs, many of the drawbacks of explicit shapes become irrelevant, because most of the solid organs are genus-0 shapes and have a relatively smooth surface. Therefore, explicit representations are widely used in this area of medical image segmentation.

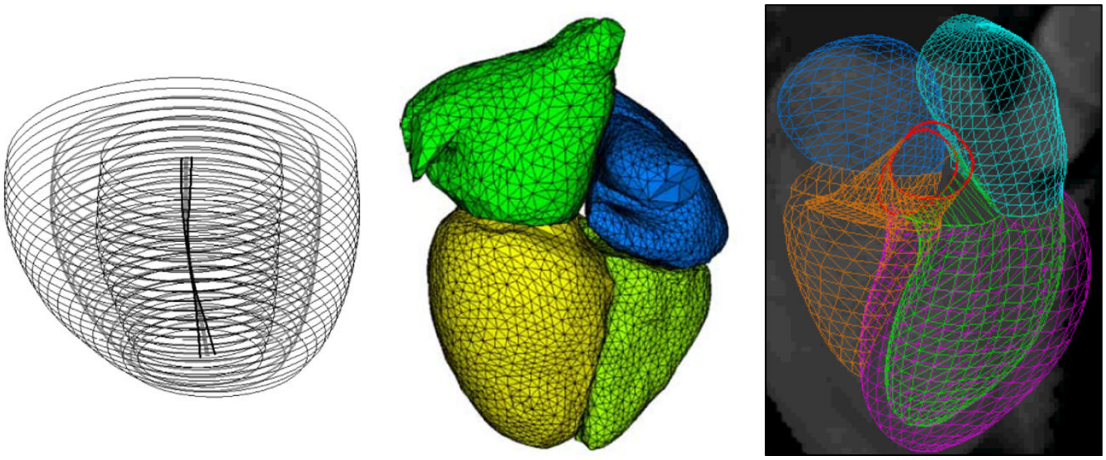


Figure 2.21.: Examples of multi-component surface models. From left to right: Two-component model of the cardiac left ventricle. Complete model of the heart (four chambers and main vessels). Four chamber heart model. Images are taken from [ZHB06, EPS*08, ZBG*07] (left to right).

2.5.2. Volumetric approaches

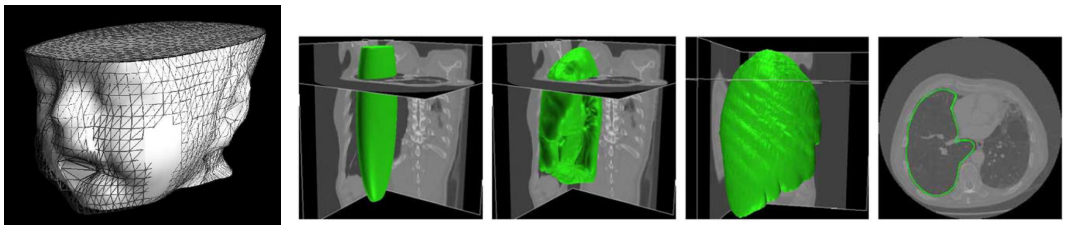
This section gives an overview about how implicit and explicit shape representations can be used to create volumetric models, i.e. models which go beyond a single surface representation.

2.5.2.1. Multi-component surface models

A natural way to extend common surface based model representation is to couple several single surfaces together in order to robustly segment complex structures. The most popular application of such multi-component surfaces is modeling the heart and its chambers. Zambal et al. [ZHB06] use a two-component model of the cardiac left ventricle that is based on two explicit parametric surface representations of the inner and outer wall of the ventricle. Figure 2.21 (left) shows this model. By coupling the outer an inner wall of the ventricle, additional global prior knowledge is incorporated which improves the segmentation quality on low contrast data. Zhen et al. [ZBG*07] created a four chamber model of the heart which showed a significant advantage over standard single surface chamber models in terms of segmentation accuracy (cf. Figure 2.21, right). Ecabert et al. [EPS*08] use single surfaces to built a complete model of the heart including the four chambers, myocardium and the main vessels (cf. Figure 2.21, middle). A general framework for coupling single surfaces in a common deformation has been proposed by Franz et al. [FWK*08].

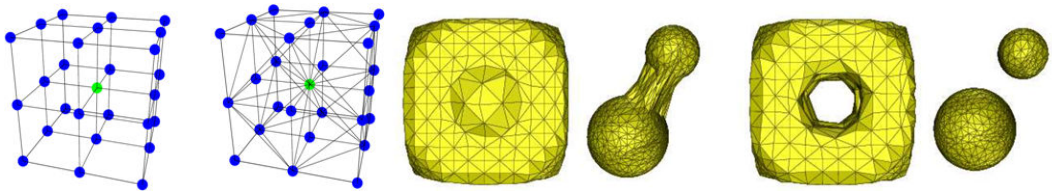
2.5.2.2. Active nets and topological active volumes

With the success of snakes and deformable surfaces, the idea emerged to extend the common parametric curve or surface representation to a net or grid like structure. A straight forward extension to deformable surfaces has been proposed by Morten Bro-Nielsen [Bn95] and is called *active cubes*. Like in deformable surfaces, an external and an internal energy are used to adapt the model to an object. However, the external energy is different for nodes that are inside the model than for nodes that are on the boundary of the model. For example, the internal nodes are



(a) (b)

Figure 2.22.: (a) 3D active cube mesh adapted to a head in a CT scan (image taken from [Bn95]. (b) Adaptation of a finite element deformation based active volume for segmentation of the lungs (images taken from [SLH11].



(a) (b)

Figure 2.23.: (a) Topological active volume represented by a cubic (left) and tetrahedral mesh (right). (b) Topological change of an active volume. Nodes not belonging to the object are removed in order to split the model or to adapt it to holes. Images are taken from [BPCO10].

attracted by high intensities in order to segment bone structures. Figure 2.22(a) shows the active cube grid adapted to a head in a CT scan. A two dimensional variant of this approach is the active net method [IBSP09]. It integrates region features for the internal nodes and edge features for the boundary nodes. Additionally, an active net is able to change its topology, that means, it can be split apart to segment multiple objects in one step and it can be used to segment complex shapes with holes. The extension to 3D is called topological active volume [BP05, BPCO10]. Figure 2.23(a) shows two different explicit shape representations: a cubic mesh (left) and a tetrahedral mesh (right). Figure 2.23(b) gives an example of the topology change abilities of topological active volumes. During adaptation, nodes are removed which do not correspond with the target object material.

Topological active volumes have been mainly applied to segment several types of bone structures [Bn95, IBSP09, BP05, BPCO10] or other structure that show high image contrast like the lung in CT [IBSP09].

2.5.2.3. Metamorphs

A class of volumetric deformable models using an implicit shape representation are *metamorphs* [HMC04, HM08]. Like in level sets, the model's shape is embedded into a higher dimensional space of distance transforms. The deformation of the model is parameterized using a free form deformation framework based on B-splines, i.e. the deformation is based on a regular grid of control points with B-spline interpolation. Metamorphs have been applied in 2D for lung seg-

mentation in MR, for cardiac left ventricle segmentation in MR and for lesion segmentation in ultrasound of the breast [HM08].

2.5.2.4. Active volume models

Recently, the so called *active volume models* [SLH11] have been proposed. Note that the naming is very similar to the *topological active volume* methods described above. However, both are independent approaches. Active volumes use an explicit shape representation in form of a polyhedron mesh (for example, tetrahedrons, octahedrons or icosahedrons can be used). Additionally, an implicit shape representation is used to incorporate boundary and regional information. Both representations are coupled in a finite element framework to adapt the model to the image. Figure 2.22(b) shows an example of an active volume segmentation. The model is initialized with a very generic shape and iteratively adapted to the lung. Active volumes cannot deal with topology changes so they are only used for solid objects. Active volumes have been extended to multiple surface active volumes [SH09] to be able to segment more complex shapes like the heart. Here, each chamber is modeled individually by a polyhedron mesh. The adaptation is then constrained such that the models do not overlap.

2.5.3. Statistical shape models

A special type of geometric model based approaches are the so called *statistical shape models* introduced by Cootes et al. [CHTH93, CTCG95]. Like in all model based approaches, their idea is to limit the shape of an object to be similar to a set of known reference shapes. They achieve this goal by drastically reducing the complexity of the reference set using statistical analysis. This way, a low dimensional shape space is created. An arbitrary shape of the same type as the reference shapes can be projected into this space. Inside the low dimensional shape space, the shape can be easily restricted to be similar to the reference shapes. During the adaptation of a model to an object in the image, the model is constantly projected into the learned shape space in order to avoid a non plausible deformation.

Statistical shape modeling starts with a set of reference shapes, often represented by point clouds. The reference shapes are usually created manually, by labeling a set of images of certain representable individuals. Hereby, it is important that all reference shapes need to have the same number of points. Moreover, the points of a single reference shape must correspond to points in the other reference, that means they must be placed at the same part of the object. For example, a point that is placed on the peak of the lower liver lobe must be placed on this peak in all other reference shapes as well. This is necessary in order to reduce the complexity of the reference set and in order to compare single shapes with each other.

The first statistical shape models were constructed from points which were manually placed on training images [CTCG95]. This manual annotation process is very time-consuming and regarded as intractable in 3D, due to the size and complexity of most shapes. Therefore, research focused on the development of algorithms that establish correspondence automatically or semi-automatically. Recent overviews of automatic correspondence algorithms can be found in [DTT08, HM09]. In typical semi-automatic methods, a sparse set of manual landmarks is defined which corresponds to predominant and unambiguously identifiable features. Additional landmarks are then automatically placed in between, either equally spaced according to the contour length in 2D [RGPA06] or by subdivision surfaces in 3D [SRN*03].

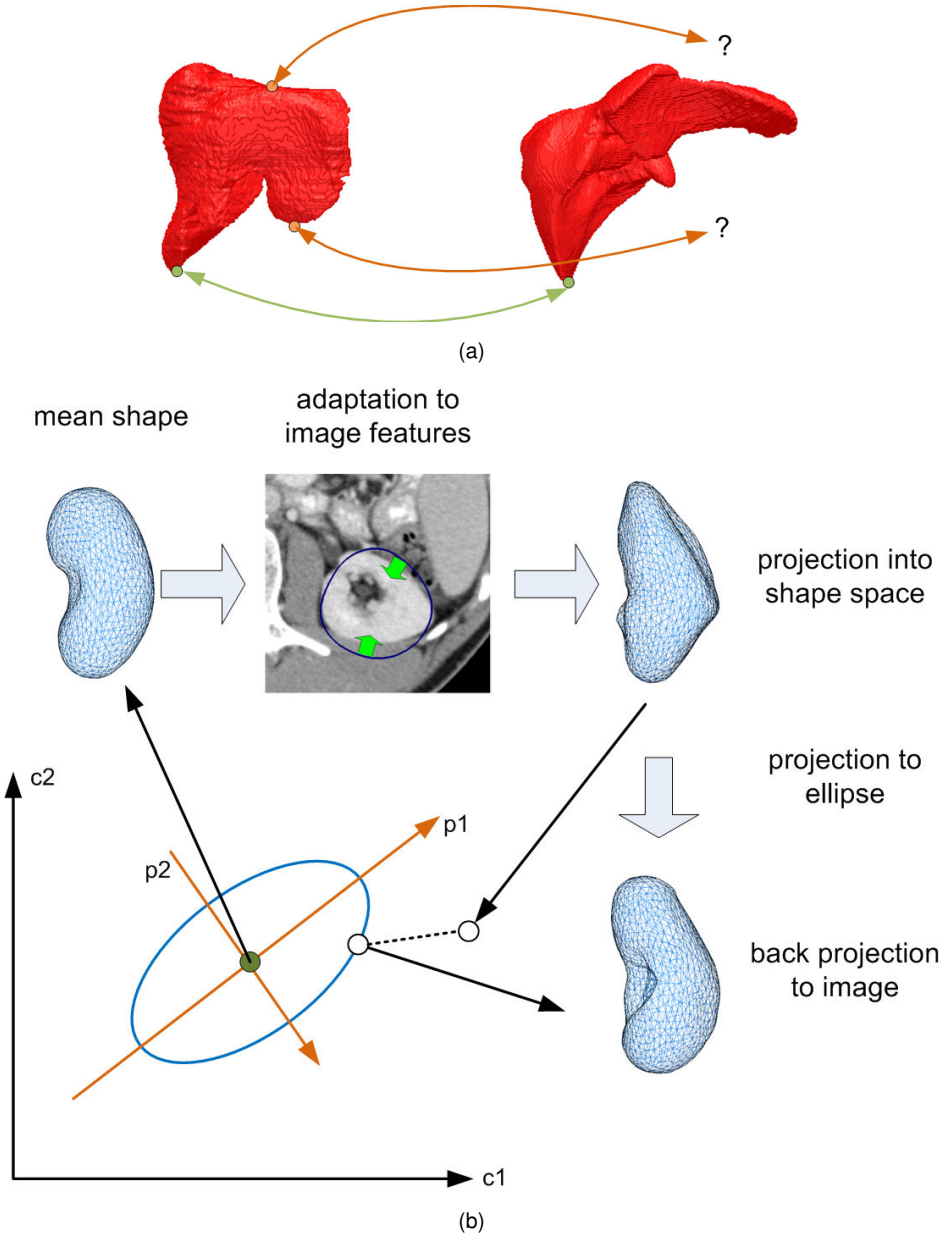


Figure 2.24.: (a) Problem of defining correspondences on complex shapes. (b) Principle of statistical shape model adaptation. The mean model is initialized in the image. The candidate shape is projected into the shape space and further projected to the ellipse of valid shapes. The result is back projected to the image which results in the restricted shape.

A basic problem that all correspondence algorithms share is that no definition exists what a good correspondence is. For simple shapes or for shapes that have clearly corresponding anatomical features, the correspondence might seem obvious. However, for complex shapes like the liver, this is not trivial. Figure 2.24(a) shows an example of the problem of defining correspondence on two different liver shapes. Often, measures like DetCov [KT97] or minimum description length MDL [DTC*02] are used for automatic optimization of correspondences. However, it remains unclear whether those methods really produce correct correspondences in the anatomical context.

After having defined a set of reference shapes with corresponding points, the complexity of the set is reduced. This is usually done using principle component analysis (cf. Section 2.4.4). Consider a set of reference shapes with N corresponding points. A single reference shape can be represented as a training vector $x_i \in \mathbb{R}^{3N}$, where x_i is the concatenation of the 3D coordinates of all points. This means, the space in which the reference shapes are defined is very high dimensional. Usually a structure like an organ is modeled by several thousands of 3D-points. In the common space, each point accounts for 3 additional dimensions. However, the single points are not independent from each other. Since organs are structures with relatively smooth surfaces, neighboring points are very likely to be close. In fact, it is impossible that they differ more than a certain distance. Therefore, PCA can be used to drastically reduce the dimensions of the shape space without losing much information about the reference set. Before applying PCA, the shapes are aligned into a common coordinate system, i.e. the representation becomes independent of translation, scaling and rotation. Afterwards, PCA is used according to Section 2.4.4 to extract the main principle components. For example, the first t components may be kept. All shapes $\hat{x}$ with the same number of landmarks as the reference shapes can now be modeled as $\hat{x} = \bar{x} + Pb$, where $P = (p_1|\ldots|p_t)$ is the matrix of retained eigenvectors, $\bar{x}$ is the mean of all reference shapes and $\vec{b}$ is the t-dimensional shape parameter vector that represents the shape in the low dimensional space (cf. Section 2.4.4).

The adaptation of the statistical shape model is done as shown in Figure 2.24(b). First, the model is initialized with the mean shape in the image. For every point, the best fitting image feature is searched for. Those candidate points form a new candidate shape x^c that is projected into the shape space according to (2.14). The shape parameters b_i are restricted to be in a certain interval in order to create an ellipse around the mean shape inside the shape space. This means, after projection, x^c will be inside the valid ellipse. Afterwards, x^c is projected back into the image. Due to its restriction in the shape space, the points of x^c will most likely not be identical with the found candidate points, but be somewhere between the initial model points and the candidate points. This process is now iteratively repeated until the shape of the model does not change significantly or a certain number of iterations has been reached.

2.6. Boundary detection

Among shape representation and shape preservation, boundary detection is a main factor which influences segmentation quality in model based segmentation. Generally, the more accurate the boundary detection model can predict the true boundaries of the structures to segment, the less prior shape knowledge is necessary to be included in the segmentation process. However, in most imaging modalities, typically a large number of anatomical structures exhibit only low image contrast. Therefore, numerous boundary detection methods have been proposed that aim at

dealing with the complex boundary appearance of such structures. Boundary detection is mainly based on two types of image features: *boundary image features* and *regional image features*, whereas boundary image features are defined in a local context of the model's boundary and regional image features are defined in a local or global context of the model's interior.

2.6.1. Boundary image features

Boundary image features are detected in a local neighborhood of the model. Using explicit shape representations, the neighborhood in which boundary features are determined is often tied to the control points of the shape. For example, a model that is represented by a tessellated point cloud will search for boundary image features in a neighborhood around every point. Usually the features are only searched along the surface normal of the point in order to avoid self-intersections of the surface during adaptation. The boundary detection than choses the best boundary image features for every point and the model will adapt towards these features.

The simplest way to detect the boundary of an object is to search for the strongest gradient in the local neighborhood. However, this leads to leakage of the segmentation in areas with low image contrast, for example, if a neighboring structure shares the same intensity values with the object to segment. Ecabert et al. [EPS*08] use simulated search on a set of training images to obtain an optimal gradient magnitude threshold separately for every model point.

Usually, additional image information around the boundary is integrated into the detection process. A very popular method is the sampling of so called *profiles* at a model point position. Such a profile is an array of image based values and certain length that is centered at the point and that is oriented along the point's normal direction. The image based values can be intensities, gradients or similar at corresponding positions in the image. During adaptation, a profile of a model point is shifted along the point's normal direction. The best fit of the profile with the boundary determines the new position of the point.

Kainmüller et al. [KLL07] created a heuristic to determine whether a sampled profile at a certain point lies on the boundary of the object or not. It is mainly based on defining valid intensity and gradient intervals in which the profile entries are required to fit. However, the created heuristic is especially designed for detecting the liver boundary only and has therefore little generalization capabilities.

The most popular way to determine whether a profile lies on the object boundary or not is based on profile training. Like in statistical shape modeling (cf. Section 2.5.3), a set of reference profiles is created, for example using manually delineated images. The dimensionality of the space in which the reference profiles are defined can also be reduced, for example using PCA (cf. Section 2.4.4). An appropriate metric for comparing profiles is defined in order to compare an input profile with the reference set. Cootes and Taylor [CT93] compared profiles of intensities with gradient profiles and their normalized counterparts. It turned out that normalized gradient profiles performed best. However, PCA based profile detection assumes a Gaussian distribution of the data. However, for complex anatomical structures this is not true. For example, parts of the upper liver area adjoin the lung as well as the heart. In CT, the mean intensity between both structures is not characteristic for either of the structures so the Gaussian assumption does not hold. Wimmer et al. [WSH09] select the k nearest neighbor profiles in the training set to create a probability of the profile being at the object boundary. In their work, they use intensity profiles. Zhen et al. [ZBG*07] use boosting to model non Gaussian boundary characteristics. Instead of one dimensional profiles, they sample a three dimensional grid around every point. Furthermore, they use steerable features as image features which are a set of scaled and oriented simplified

gradients. Heimann et al. [HMW07] also use a three dimensional sampling pattern around every point. They sample an cylinder oriented along the point's normal direction. The histograms of the parts of the cylinder that are inside and outside the reference shape boundary are concatenated into a single vector. Mean shift clustering is then used to determine profile clusters of similar appearance.

2.6.2. Region based image features

Region based image features are computed based on image characteristics of the model interior and/or exterior. The idea behind this is that a structure can be segmented more robustly if the intensity distribution inside the structure is considered in addition to the boundary features as described above. For example, some structures like the non contrast enhanced bladder in CT are very homogeneous. However, if the interior of the model is inhomogeneous, the model is probably misplaced. Region based image features can be defined as *global* or *local*. Global region based image features compute a single matching value for the whole model from its current position. Local region based image features compute several matching values for different areas or positions inside the model, i.e. they can be used to determine the exact position of a match or mismatch of image characteristics.

A popular method to incorporate global regional image features are active appearance models proposed by Cootes et al. [TFCT98, CT04]. The voxel intensity values of the interior model area in a set of reference images is sampled and normalized such that a single vector of a fixed length can be used to model the interior area in each reference image. The vector can then be treated analogous to a profile as described above, that means PCA can be used for dimensionality reduction and the same metrics can be used for comparing an input vector with the trained vectors. However, even if dimensionality reduction is used, active appearance model do not scale well on 3D data. In principle, each voxel of the model's interior region accounts for an own dimension. Therefore, active appearance models are usually either applied to images that are resampled to lower resolutions or to very small structure that do not contain much voxels. Furthermore, active appearance models use global regional image features. Therefore, they only have a global view on the accuracy of the current model fit. Wimmer et al. [WSH09] model the local texture around a voxel by using Haar features. The Haar features are trained on reference images using a boosted classifier cascade. This way a classification for every voxel of the model interior is given whether it is part of the object or not. Because of the huge amount of training data (one sample per voxel) training such a classifier cascade can take several days. Huang et al. [HM08] assign a probability to every interior voxel based on its similarity to all interior voxels. Here, a Gaussian distribution is considered for all interior voxels, so the probability is highest if a voxel's intensity is equal to the mean of all interior voxels.

Volumetric models that consist not only of surface points but also of internal points can be used to easily incorporate local regional information. In case of a detection of foreign tissue, it is therefore possible to determine the exact internal area of mismatch. Barreira et al. [BPCO10] sample a three dimensional cube around every internal node. The intensity of the local cube is integrated into the external energy in order to determine whether a point has to be removed from the model.

Table 2.1.: Overview of state of the art model based segmentation approaches.

Method	volumetric approach / optimization efficiency	boundary features	region features	shape preservation	generalizability / extensibility
1: Cootes et al. [CT93]	no	trained profiles	no	SSM	++ / ++
2: Kainmüller et al. [KLL07]	no	gradient/intensity heuristic	no	SSM + free form	– / ++
3: Heimann et al. [HMW07]	no	trained local histograms	no	SSM + free form	++ / ++
4: Wimmer et al. [WSH09]	no	trained profiles	local	SSM + free form	++ / +
5: Zhen et al. [ZBG*07]	no	trained steerable features	no	SSM	++ / ++
6: Ecabert et al. [EPS*08]	no	trained gradient threshold	no	affine transformation + SSM	+ / ++
7: Cootes et al. [CT04]	no	trained profiles	global	SSM	++ / ++
8:Huang et al. [HM08]	yes / medium	gradient/intensity heuristic	global	single reference	+ / –
9: Ibanez et al. [IBSP09]	yes / high	gradient/intensity heuristic	local	single reference	+ / –
10: Shen et al. [SLH11]	yes / medium	online intensity estimation	local	single reference	++ / –
11: Barreira et al. [BPCO10]	yes / medium	gradient/intensity heuristic	local	single reference	– / –

2.7. Discussion

In Table 2.1, state of the art model based segmentation approaches are classified based on model representation, boundary detection, incorporation of regional features, shape preservation strategy as well as generalizability and extensibility. Some of the chosen algorithms are representative for a class of similar algorithms that have been proposed in the literature. For example, the active appearance model approach by Cootes et al. [CT04] has inspired numerous methods that model regional appearance by statistical analysis. The category *volumetric approach* defines whether an algorithm uses a volumetric shape representation or not. In case such a volumetric representation is used, the optimization efficiency on three dimensional medical data is given. The generalizability category defines how well a method can be applied to the segmentation of other structures. The extensibility category determines the modularization capabilities of an approach, i.e. it determines whether single modules like boundary detection or shape preservation strategies can be exchanged without major changes to the general approach.

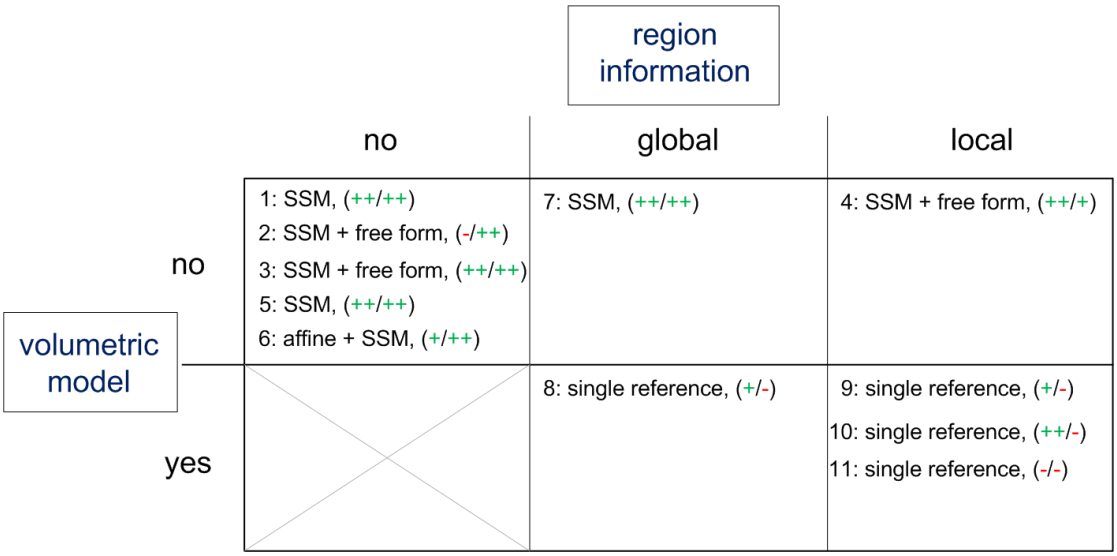


Figure 2.25.: Classification of the model based methods of Table 2.1 to surface based methods (top row) and volumetric methods (bottom row). The other axis shows the type of regional image appearance that is incorporated. The combination of a volumetric model without region features is excluded, since it stands in contrast with the idea of using volumetric models.

The optimal method for the goal of this thesis — a volumetric segmentation technique for organ segmentation and registration — has the following characteristics. It incorporates a *volumetric shape representation* in order to consider image information inside the structure to segment. In the best case, the *optimization complexity* of this shape representation is comparable to the optimization of surface based methods. Furthermore, the *regional information* to be incorporated is *local*. That means, it is possible to detect the position of object foreign material inside the object in opposition to global regional information that is only able to detect a mismatch of the whole interior intensity distribution. The shape preservation strategy *incorporates shape knowledge* learned from a set of shapes instead of simply using a single reference of the structure to segment. Moreover, the ideal method is *generalizable* to other structures and *extensible* such that the method can be easily extended to other application scenarios.

As can be seen in Figure 2.25, the majority of surface based segmentation methods is modular and well generalizable. However, since no volumetric shape representation is used, regional information is usually not incorporated. The volumetric approaches on the other hand all incorporate regional information, but have only weak generalization capabilities and are less modular. This is often due to the fact that volumetric methods are more complex and therefore use boundary detection and shape preserving strategies specifically adapted to the used model. Particularly, often well established boundary detection methods cannot be directly applied to such models which limits their application to well contrasted structures like bones or the lung in CT [HM08, BPCO10]. The biggest drawback of current volumetric model based segmentation however is that they cannot be easily coupled with statistical shape modeling methods. Statistical shape modeling has been a key enabler for segmentation of complex structure like the liver [HvGSea09]. This drawback relativizes many of the advantages of volumetric methods and prevents their broader application in current organ segmentation.

As it has been discussed in this chapter, several open issues in model based segmentation and registration exist. One problem is to find a proper way to leave the trained shape space during model adaptation in order to adapt the model to unseen cases without losing too much prior shape information during that process. This is especially challenging for structures of complex shape like the liver, where an adaptation in the learned shape space alone is not sufficient for an accurate segmentation. Another major challenge is to combine the flexibility, modularizibility and computational performance of current state of the art surface based methods with the capabilities of volumetric approaches to incorporate regional information in order to detect cavities and organ foreign tissue. The next chapter outlines the methods proposed in this thesis to address the described issues in order to create a robust volumetric segmentation technique for organ segmentation and registration.

Parts of this chapter are based on publications of the author [ERS08, Erd09, ESS12].

3. Non-uniform deformable volumetric objects

This chapter describes the methods developed in this thesis for automatic organ segmentation and registration. It is structured as follows: first, the proposed general segmentation and registration pipelines are outlined. An analysis about the variance of object boundaries, object homogeneity and object shape for the most important organs and structures as visible in CT is given in Section 3.2. Based on this analysis, the surface and volume based geometric models developed in this thesis are presented in Section 3.3. Afterwards, the model adaptation process is outlined in Section 3.4. The optimization of surface and volume based models is described in Section 3.4. The chapter concludes with the proposal of an automatic model initialization method in Section 3.6.

The main contributions of this chapter are

- the incorporation of local elasticity constraints in model based segmentation described in Section 3.3.2,
- a new method for combining ground truth label building and point correspondence establishment for statistical shape models outlined in Section 3.3.3,
- the proposal of a new volumetric multi-layer model for image segmentation and registration as described in Section 3.3.4,
- a rule based outlier removal for boundary detection of organs outlined in Section 3.4.2,
- the adaptation of an optimization scheme for surface and volumetric point based models presented in Section 3.5,
- a new automatic model initialization method based on 3D histograms of oriented gradients described in Section 3.6.

3.1. Overview

This section gives an overview about the general segmentation and registration pipelines developed in this work in order to address the challenges of model based organ segmentation and registration described in the previous chapter. Figure 3.1(a) shows the proposed model based segmentation pipeline. First, the organ to be segmented is coarsely detected in the image by detecting a bounding box that contains the organ. This process is described in detail in Section 3.6. The model is then initialized in the middle of the bounding box. It follows a step wise adaptation with increasing degrees of freedom of the model. The process of model adaptation and model optimization can be found in Section 3.4 and Section 3.5. The described segmentation pipeline can be executed using two different types of models: a surface model and a multi-layer model. The multi-layer model has been developed in order to incorporate information from the organ's interior during adaptation. The construction process for both models is outlined in Section 3.3.

The described model based segmentation pipeline can be used to register two images of the same patient. This process is shown in Figure 3.1(b). In each image the organ is extracted. The

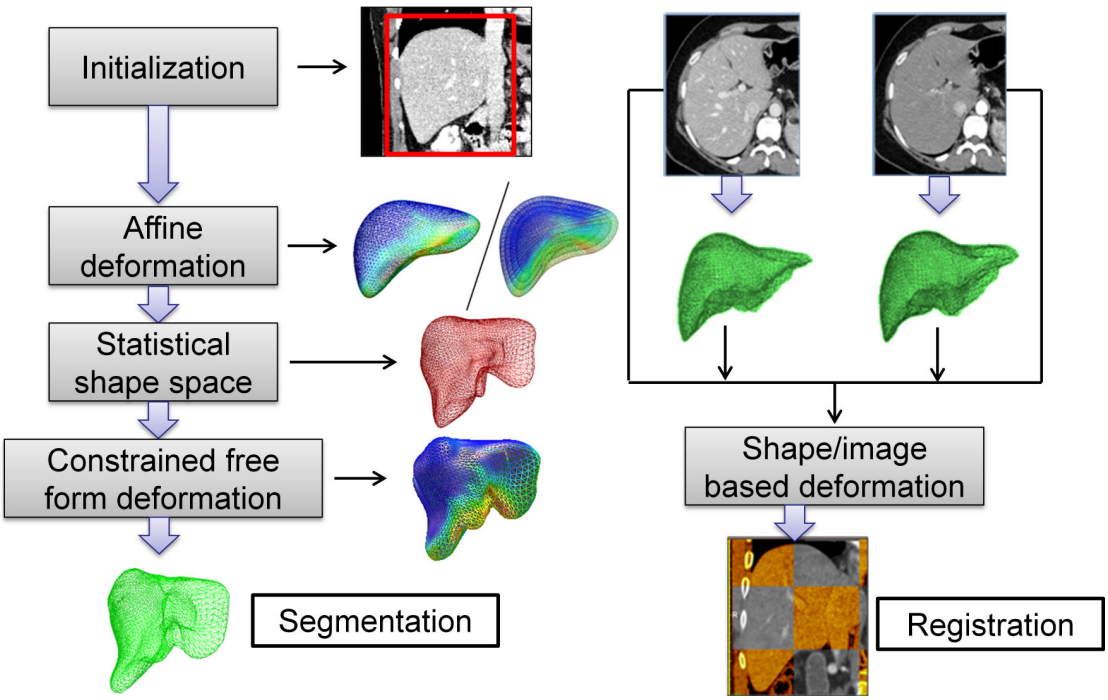


Figure 3.1.: (a) Overview of the proposed model based segmentation pipeline. The model is first initialized in the image. The adaptation of the model to the contours of the organ is done step wise with increasing degrees of freedom. (b) Overview of the registration process that uses the model based adaptation shown in (a). Shapes are extracted from two images. They are then used together with the image information to register both images.

shape information of both organs is then used to elastically register both images. This way, a precise registration can be performed without having to rely on matching image features inside of the organ. This is particularly important for organs with different appearances in both images. For example, in multi-phase CT of the liver, different vessel trees are saturated in the images leading to a different image contrast. A matching of image features in the interior of the live would therefore likely cause mismatches. Section 4.5 describes the registration process in detail.

3.2. Anatomical structures in CT

As it has been discussed in the previous chapter, there are mainly three main characteristics which influence the segmentation of a single object in an image: object boundaries, object homogeneity and object shape. Furthermore, in case of medical images, target structures are not always at the same position in the image but vary in terms of position according to the anatomical variance inside the human body between individuals. That means, the positional variance of the structure to segment also plays a role when it comes to the development of an *automatic* segmentation approach of that structure.

However, different organs also strongly differ in the described characteristics. For example, the skull in a CT image appears with a strong image contrast. Its shape variability is less complex than the shape variability of the liver and so on. Since the goal of this thesis is the development of an automatic method for segmentation and registration of organs, it is first necessary to give an overview of the shape variance, appearance variance and positional variance of different organs in the human body. Based on such an analysis, appropriate methods are developed to address the goal of this thesis for challenging organs that show strong variance in one or more of the described main characteristics.

This thesis focuses on Computed Tomography as the imaging modality, because it is one of the most widespread modalities for imaging of whole organs and the main imaging modality for important clinical applications such as computer aided diagnosis and treatment planning of cancer. The appearance based methods to be developed will therefore be based on that modality. However, since the concept of shape is generic, the shape based methods to be developed will be applicable to other imaging modalities such as MRI or 3D-ultrasound.

Figure 3.2 shows a classification of selected representative anatomical structures of the human body regarding their shape variability ans well as their relative positional variance in the body. Three types of anatomical structures are distinguished: 1.) bones, 2.) organs and 3.) other structures. Solid structures like the bones show the least shape variability since their shape does not vary in images of the same patient, but only between individuals due to their rigid characteristics. That means, they are not affected by respiration, weight loss or gain, cardiac movement or blood and body fluid movement. They also show the least positional variance, since the human skeleton defines a relative consistent reference frame. Soft tissue organs differ a lot in shape and position, but have generally less variability in comparison to structures like lymph nodes or vessels. Tumors represent the limits within this classification, since their shape and position inside the human body are close to being arbitrary.

As can be seen, inside the group of organs, the liver and the pancreas show the most shape variability while the kidneys, lung and bladder are usually of relatively stable shape. However, regarding the positional variance in the human body, the kidneys and the pancreas are the most challenging organs. Those organs are rather small and embedded into a soft tissue anatomical

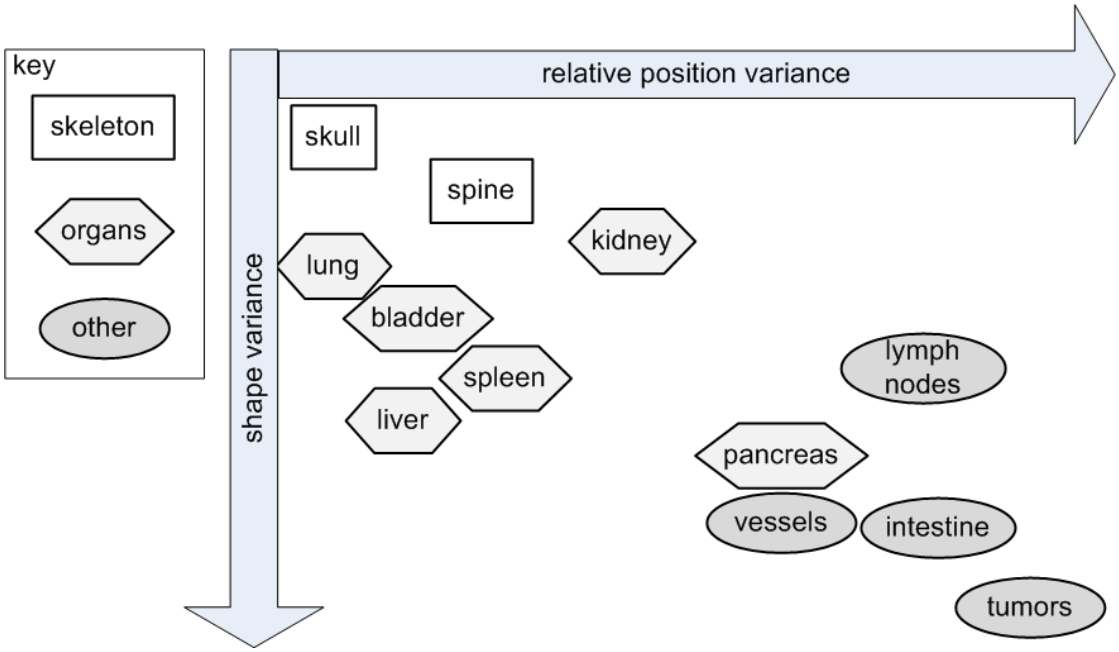


Figure 3.2.: Classification of selected representative anatomical structures regarding their shape variability as well as their relative positional variance in the human body. Three types of anatomical structures are distinguished: 1.) bones, 2.) organs and 3.) other structures. Solid structures like the bones show the least shape variability since their shape does not vary in images of the same patient, but only between individuals due to their rigid characteristics. They also show the least positional variance, since the human skeleton defines a relative consistent reference frame. Soft tissue organs differ a lot in shape and position, but have generally less variability in comparison to structures like lymph nodes or vessels. Tumors represent the limits within this classification, since their shape and position inside the human body are close to being arbitrary.

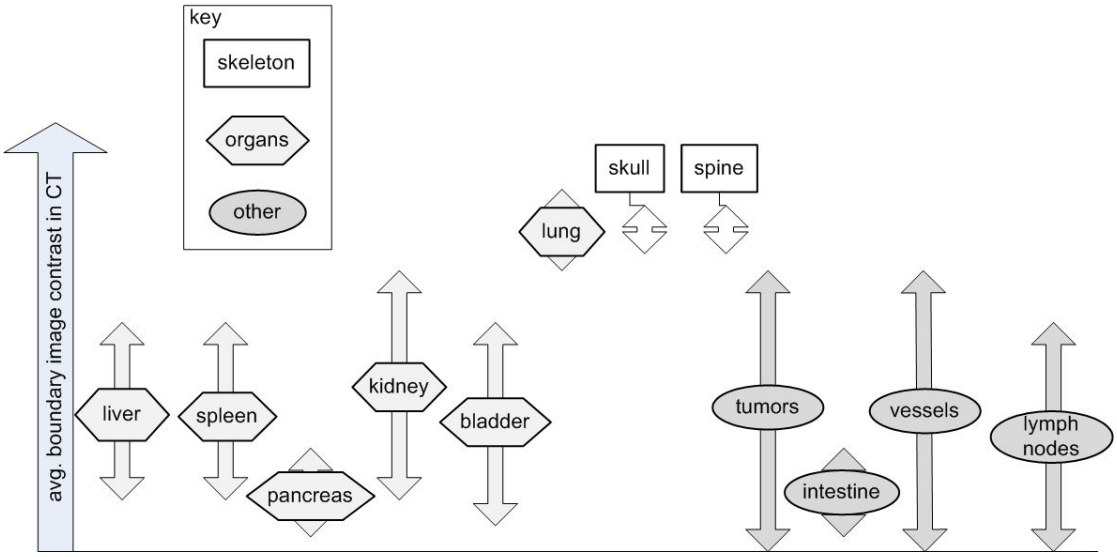


Figure 3.3.: Classification of anatomical structures regarding their image contrast at boundaries in CT. Some structures may have boundary parts that exhibit better contrast than other boundary parts. In this figure, the average boundary contrast for a structure is given. The arrows indicate the possible contrast span over different CT images. The contrast span is mostly influenced by contrast agent application and different embeddings of the structure. Structures like the lung or the bones usually have very good contrast at their boundaries. The main abdominal organs like liver, spleen and the kidney have medium to good contrast if contrast agent is used. The pancreas generally has very bad contrast at its boundaries even if contrast agent is used.

context which makes their position dependent on the shape of larger neighboring organs, respiration state, surrounding fat storages etc.

Figure 3.3 shows a classification of the same anatomical structures as shown before, but in the context of their average image contrast at organ boundaries in CT. The arrows indicate the possible contrast span over different CT images. The contrast span is mostly influenced by contrast agent application and different embeddings of the structure. For example, a lymph node may be embedded into fat or other soft tissue like muscles or glandular tissue which leads to different boundary contrasts. Structures like the lung or the bones usually have very good contrast at their boundaries. The main abdominal organs like liver, spleen and the kidneys have medium to good contrast if contrast agent is used. The pancreas generally has very bad contrast at its boundaries even if contrast agent is used.

Figure 3.4 shows the average amount of intensity heterogeneity of the discussed structures in CT, i.e. it shows whether the intensity inside of a structure varies a lot or not. Again, the arrows indicate the possible span over different CT images. For example, different types of tumors exist that may appear very homogeneous in CT or be totally heterogeneous. The use of contrast agent also has a huge impact on the homogeneity of a structure in CT. Generally, structures like the skull, vessels or pancreas are relatively homogeneous, i.e. their intensity does not vary much *inside* the structure. Organs like the kidney or the liver are often acquired using contrast agent which makes interior vessels visible thus leading to a certain degree of inhomogeneity. However,

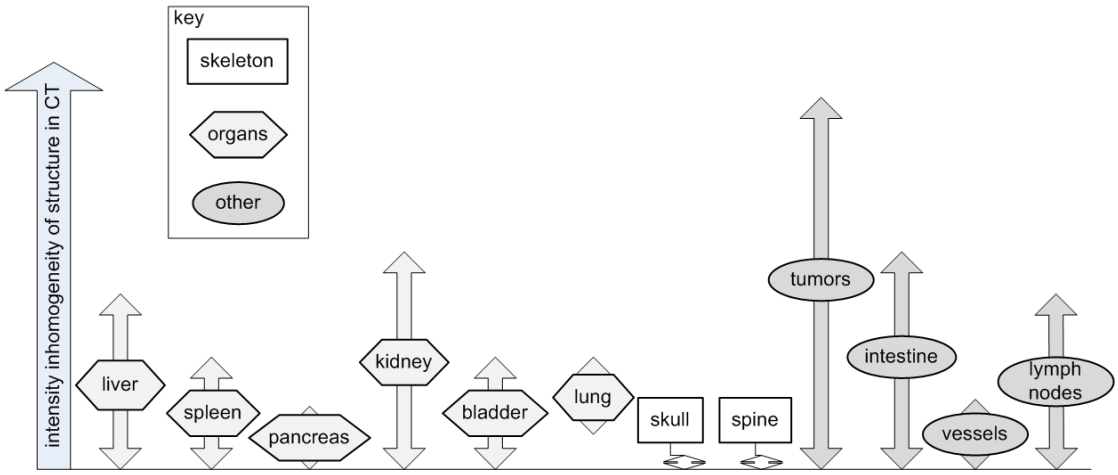
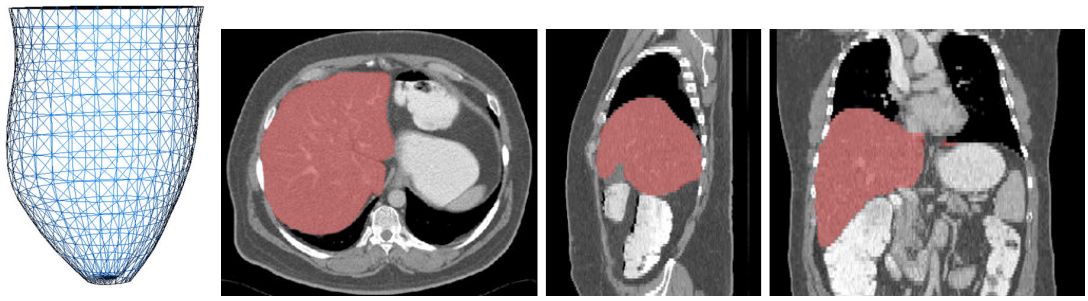


Figure 3.4.: Classification of anatomical structures regarding their intensity heterogeneity in CT images. The arrows indicate the possible span over different CT images. Structures like the skull, vessels or pancreas are relatively homogeneous. In contrast, organs like the kidney or the liver are often acquired using contrast agent which makes interior vessels visible thus leading to a certain degree of inhomogeneity. However, those organs can also be acquired without contrast agent. In this case they are very homogeneous.

those organs can also be acquired without contrast agent. In this case they are very homogeneous. Therefore, those organs are the most challenging structures regarding interior intensity variance.

The given analysis shows that the pancreas is in most categories the most challenging organ to segment because of its strong shape variance, poor image contrast and positional variance. The liver shows a little better boundary contrast but its interior is generally more heterogeneous depending on contrast agent saturation. The kidneys exhibit a strong positional variance and their image appearance is very dependent on contrast agent saturation. The lung is the easiest organ to segment, mainly because of its good image contrast and small variance in position. The methods to be developed in this thesis, therefore, will be mainly applied and evaluated on the challenging organs, in particular, the liver, the pancreas and the kidneys.

The analysis also shows that the image contrast of all considered organs except the lung can strongly vary between scans. This is mainly due to different states of contrast agent saturation. Without contrast agent, most organs will usually have very little or no boundary contrast in CT. For example, current research in liver segmentation [HvGSea09] focuses on the segmentation of the liver in the portal-venous phase where the parenchyma is saturated with contrast agent which leads to a better boundary contrast. However, the less saturated arterial phase also has a strong clinical value, for example in the case of cancer diagnosis. The methods to be developed will therefore also target at such difficult cases where so far no robust methods for segmentation exist.



(a) (b)

Figure 3.5.: (a) An organ model of the cardiac left ventricle represented by a three-dimensional point cloud and visualized by a polygonal tessellation. (b) Manually created binary segmentation of the liver shown as an overlay with the original CT image in three cross-sectional views.

3.3. Model construction

The first step in building a model based segmentation approach is the model construction, i.e. it must be decided how shape is represented and how reference shapes are incorporated in order to restrict the model to plausible shapes. In Section 2.5.1, commonly used shape representations in model based segmentation are described. In this thesis, the model is explicitly represented by a set of points defined in a three-dimensional space (cf. Figure 3.5(a)). In the literature, those model points are also often called landmarks. In the remainder of this thesis, both terms will be used synonymously. The 3D points are defined to lie on the surface of the model and are interconnected by a polygonal tessellation. As it has been discussed in Section 2.5.1 both explicit and implicit shape representations are used in the current state of the art of model based organ segmentation and have their strenghts and drawbacks. The reason for choosing an explicit landmark representation is the following: in contrast to implicit representations, interconnected point clouds are efficient to optimize. They are easily locally controllable through manipulation of single points, i.e. different shape preserving strategies can be easily applied. Furthermore, the strengths of implicit representations — topology changes and avoidance of self intersections of the model's surface — have only minor relevance to the segmentation of organs, since solid organs are genus-0 shapes and have a relatively smooth surface. Moreover, the best performing methods in model based segmentation of organs with complex shapes are often using interconnected point clouds as the model representation. For example, in a recent competition on liver segmentation [HvGSea09], the top performing method uses an interconnected point cloud [KLL07]. In current research on heart segmentation, the two approaches with the highest accuracy and the largest test base [ZBG*07, EPS*08] also use the same representation. Lastly, the method to be developed will also be used for registration of organs. Here, correspondences must be established after model adaptation. Those correspondences are naturally given in explicit representations but only difficult to obtain in implicit representations.

The second element in model construction is the choice of a suitable reference shape representation. As it has been discussed in Section 2.3.1.4, single representative shapes can be used as the reference. In case of complex shape variation, usually a group of reference shapes modeled by statistical analysis (cf. Section 2.5.3) is used. In Section 3.2, it has been shown that the shape variability strongly differs between different organs. Therefore, in this thesis, single

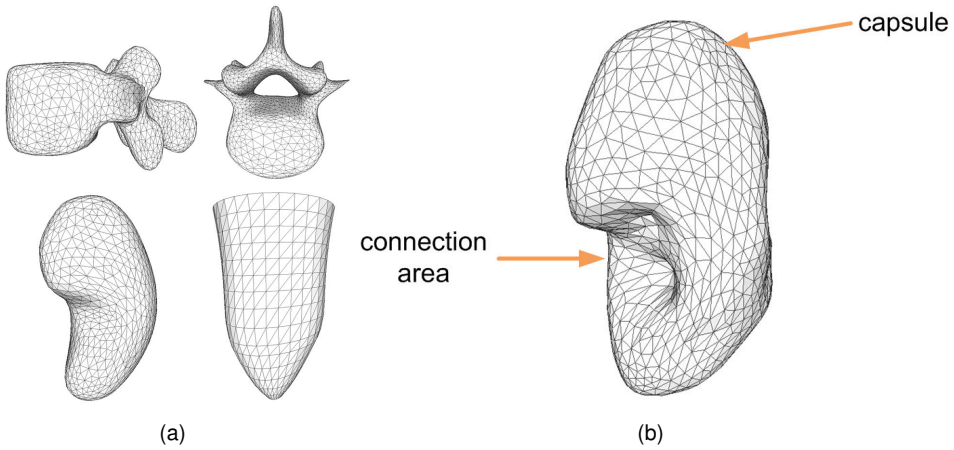


Figure 3.6.: (a) Exemplary generated mesh models from manually created reference segmentations: vertebra (top), kidney (lower left) and cardiac left ventricle (lower right) [EKW09]. (b) Reference shape of the left kidney. The surface of the renal capsule is significantly smoother than the region where vessels and ureter connect to the kidney.

reference shapes are used for organs with low shape variation like the bladder (cf. Section 3.2) and statistical shape modeling is used for organs with large shape deviation like the liver.

3.3.1. Surface model shape geometry

The geometry of the organ models is constructed based on clinically validated reference segmentations. Those reference segmentations are binary labels defined in 3D images of representative clinical cases (cf. Figure 3.5(b)). Only healthy organ shapes are considered as references, i.e. no exterior tumors or partial resections are visible. In case more than one reference segmentation for an organ is available, a model for each reference is created. Afterwards, the geometric mean of all models is taken as the reference shape for the organ.

The binary segmentation images are first resampled in order to remove typical staircase artifacts resulting from image reconstruction in the original CT or MRI. Secondly, morphological closing and opening is applied to close holes inside the organ (e.g. vessels that have not been classified as organ tissue by the expert). Afterwards, the *Marching Cubes Algorithm* [LC87] is used to generate a polygonal tessellation from the binary mask. Since the points of the resulting mesh are in most cases not regularly distributed, a Laplacian smoothing is iteratively applied until all polygons are of comparable size. Figure 3.6(a) shows the results of the model generation of vertebra, kidney and cardiac left ventricle.

3.3.2. Local shape constraints

As it has been discussed in Chapter 2 and Section 3.2, the target shapes may strongly differ from the shapes taken as the reference to build the model. Even if statistical shape modeling (cf. Section 2.5.3) is used to capture the anatomical variance of an organ over representative datasets, the trained shape space is usually left at some point in order to adapt to unseen cases. In order to avoid the formation of implausible shapes during model adaptation, in this thesis, a new

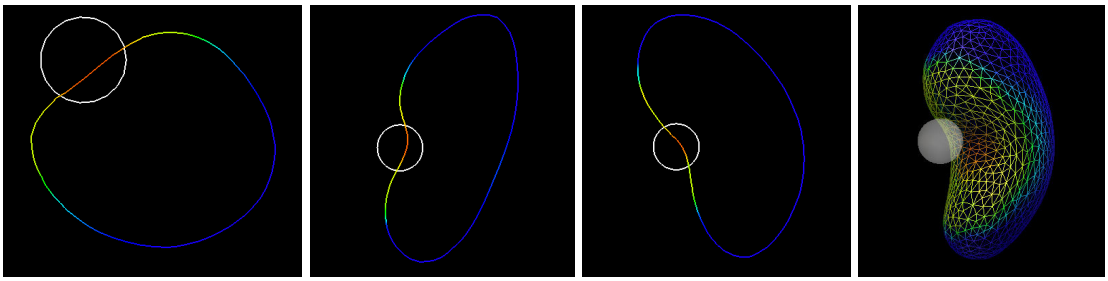


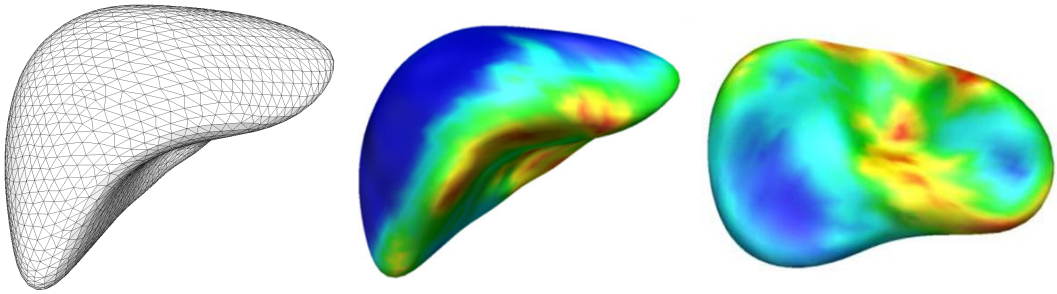
Figure 3.7.: Color coded editing of stiffness constraints [EKW10]. The local stiffness of the organ model is modified by a paint brush tool in the 2D and 3D views.

method for restricting the model to plausible shapes is proposed that can be used in conjunction with standard landmark based model based segmentation.

Some organs like the kidneys have clearly identifiable areas where the boundaries show high frequency oscillation while other regions are more or less smooth. For example, the kidney's capsule is a smooth area while the region where the ureter and vessels connect to the capsule is rather rough (cf. Figure 3.6(b)). Other organs like the liver have more complex shapes. However, in most cases, areas can be identified where the local curvature is relatively constant between individuals. Inspired by these anatomical properties, an extension to current model based segmentation approaches is proposed. Common models used for adaptation exhibit the same elasticity over the whole shape, i.e. the amount of shape restriction is not modeled locally but only globally. The shape preserving term of model based segmentation (cf. 2.3.1.4) therefore cannot distinguish between areas that should be smooth and regions that should be rough. In order to build a model with locally defined elasticity, for every landmark in the model, local weights $w_i \in [0, \infty)$ are defined that model the local stiffness of that point in relation to its neighbors. That means, in case of a high weight the relative position of the point to its neighbors should be preserved. In case of a weight of 0, the point is free to move anywhere without penalizing the shape preserving term.

For organs like the kidneys that consist of clearly identifiable regions of different curvature, the local weights can be set interactively. Therefore, a 2D/3D paint brush tool has been developed to add weights to every landmark of a model. Figure 3.7 shows the paint brush to set the stiffness constraints in the 2D views or in a 3D visualization for the left kidney. Here, weights from 0 to 100 can be set, where 0 (red) denotes a soft connection of a point to its neighbors and 100 (blue) denotes a stiff connection. The user defines a weight value and a paint brush size and edits the area of interest. In case of Figure 3.7, the area of the ureter and vessel connections of the kidney are defined to be more flexible than the region of the kidney capsule. In order to assure a smooth deformation behavior of the model between regions of different stiffness, new weights are alpha-blended with the current constraint values of the model (see color blending in Figure 3.7).

For organs with more complex shape variance like the liver, the local weights are automatically determined in a training step. As has been described in Section 2.4.4, statistical shape modeling methods cannot reconstruct local details of the training set due to the performed dimension reduction. Therefore, the reconstructed shapes will be smooth. This smoothness prevents a precise local adaptation of the model to the data. Therefore, the local weights have be set such that they allow the model to further adapt in areas where smoothness is not likely. Such areas can be identified by computing the mean curvature in every point of the model over the training set. A



(a) (b)

Figure 3.8.: (a) Geometric mean organ model of the liver generated from a set of reference shapes. (b) Color coded stiffness constraints of the same liver model [ESW10]. Dark blue areas are stiff regions where the model can only slightly deform whereas red areas are very flexible.

high mean curvature at a certain point indicates that the shape of the organ in this area is likely to be not smooth while a low mean curvature indicates that the area is generally smooth and does not need to be further adapted.

The process of automatic determination of the local weights using the mean curvature of the reference dataset is done as follows: for each manually labeled reference dataset, a model with corresponding landmarks is created. At this point, it is assumed that all created models comprise the same amount of points and that correspondence is already established (see Section 3.3.3 for the process of correspondence establishment on reference shapes). At each point t_i of all training shapes the Gaussian curvature $C_g(t_i) = \kappa_1\kappa_2$ is computed, where $\kappa_1\kappa_2$ is the minimum and maximum principle curvature. Afterwards the mean curvature over all training shapes is computed as weight w_i and stored for every point p_i in the geometric mean model $\bar{x}$ (cf. Figure 3.8(a)) of all training shapes to encode regions of different elasticity. Using this method, several anatomically plausible regions of similar curvature can be distinguished. Figure 3.8(b) shows the result for a model of the liver where the different regions are color coded. As can be seen, the upper areas of the liver are coded as dark blue which denotes low local curvature. In this region the lung adjoins the liver. Therefore, usually the liver surface is very smooth in this area. The ends of the liver lobes on the other hand show higher curvature. Also in the area where the portal vein enters the liver (yellow-red area in the middle of Figure 3.8(b), right) a high amount of curvature can be observed.

Section 3.5 shows, how the local weights can be integrated into a model based segmentation optimization framework. In Section 4.2, trained local curvature constraints are combined with constraints derived from the model adaptation process. An integration into the common framework of statistical shape modeling is shown in Section 4.2.3. The application of locally restricted deformable models to the segmentation of the kidneys, liver, pancreas and bladder will be demonstrated in Section 4.1, 4.2, 4.4, 4.3 and 4.5, respectively.

3.3.3. Shape model reference set creation using smart manual landmarking

As it has been discussed in Section 2.5.3, the so-called *correspondence problem* has to be solved in order to build a statistical shape model from a set of training shapes that represent the

organ to model. This means, that all point clouds representing the organ must contain the same number of points, and for two different point clouds each point of data set A is required to have a corresponding point in data set B, both representing the same anatomical feature.

The first statistical shape models were constructed from landmarks which were manually placed on training images [CTCG95]. This manual annotation process is very time-consuming and regarded as intractable in 3D, due to the size and complexity of the shapes. Therefore, many researchers focused on the development of algorithms that establish correspondence automatically or semi-automatically. Recent overviews of automatic correspondence algorithms can be found in [DTT08, HM09]. In typical semi-automatic methods, a sparse set of manual landmarks is defined which corresponds to predominant and unambiguously identifiable features. Additional landmarks are then automatically placed in between, either equally spaced according to the contour length in 2D [RGPA06] or by subdivision surfaces in 3D [SRN*03].

Beyond the tractability problem, the process of manual landmarking is often criticized for suffering from inter- and intra-observer variability. However, this does not imply that manually placed landmarks have lower quality than their automatically determined counterparts, as it remains unclear whether the objective functions applied in automatic methods really measure *true* correspondence. Evaluation studies give an inconsistent picture: In the study of Styner et al. [SRN*03], a semi-automatic method based on manually defined landmarks and subdivision surfaces produced worse landmarks than optimization algorithms based on DetCov [KT97] and MDL [DTC*02]. It is unclear whether the manually defined landmarks or rather the automatic subdivision scheme accounts for the poor performance of the semi-automatic method in this evaluation. On the other hand, in a study from Ericsson and Karlsson [EK07], models learned from manually defined landmarks performed better than those constructed from automatically established landmarks. Though the latter study was restricted to 2D shapes and therefore excluded all complications that occur in 3D, the results indicate that manual landmarks may be better than their reputation.

(Semi-)automatic algorithms require that the training shapes are provided as surfaces. In practice, these surfaces are reconstructed from segmentations, which requires that either an expert delineates contours on training images manually or that automatic segmentation algorithms are used. Manual delineation is time-consuming and tedious, though by far easier than consistently placing landmarks. If the segmentations are generated automatically, the resulting training shapes are restricted by the accuracy of the applied segmentation algorithm.

While segmentation of training shapes and establishing correspondence are treated independently in case of (semi-)automatic correspondence algorithms, manual placement of landmarks integrates both aspects. Hence it is promising to develop tools which support the user during placement of landmarks, thereby making the process tractable.

As a contribution of this thesis, a method for simultaneous segmentation and point correspondence establishment for statistical shape models is presented. In this approach, a reference mesh is manually deformed and at the same time optimized in real time to preserve point correspondence. The resulting meshes can be directly used for building a shape model. In addition, statistical shape models of arbitrary topology can be easily constructed using the method, whereas many automatic, parameterization-based methods are either restricted to shapes of specific topology, like genus-0 surfaces [HM09], or require an artificial decomposition of the training shapes into several patches [SKH*08]. In the latter approach, correspondence is established independently on the patches and the results are merged afterwards, which introduces discontinuities at the cuts.

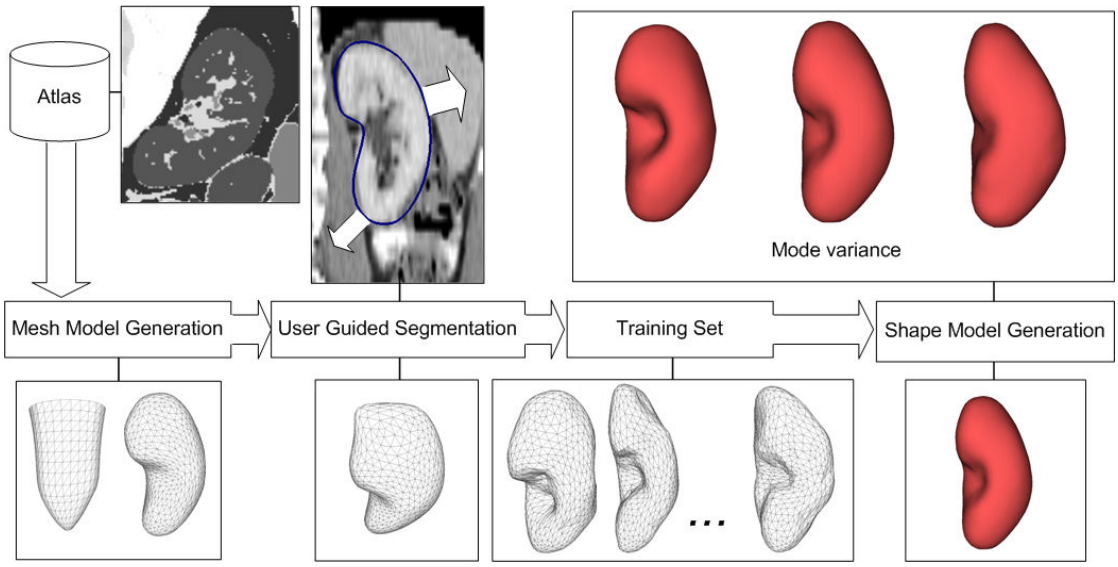


Figure 3.9.: Workflow of shape model generation [EKW09]. One triangulated reference mesh is created for every organ. The reference mesh is deformed by the user and simultaneously optimized to enforce correspondence of the deformed models with the reference shape. The resulting surfaces are then used as training data input for the shape model generation.

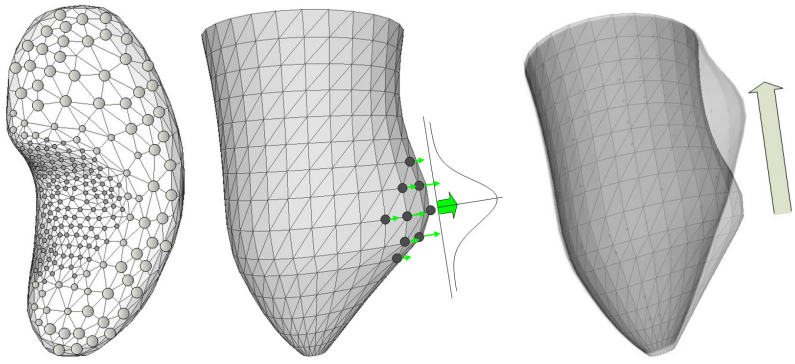
3.3.3.1. Workflow

An overview of the shape model generation process is given in Figure 3.9. The first step is the construction of a polygonal reference organ model that can be taken as a basis for all training data sets of the shape model. The next step is a user guided segmentation. Here, the mesh is three dimensionally deformed by the user to match the organ boundaries in the data set. In order to ensure that points in the reference mesh and the deformed mesh denote the same feature points and therefore correspond to the same region, the mesh is globally optimized in each deformation step. The results are deformed training meshes with regularly distributed points that can be directly taken for shape model generation in the last step.

3.3.3.2. Reference Mesh Construction

In order to create a segmentation by using manual mesh deformation, a reference model is needed for every organ that can be adapted to the data set by the user. This model should be represented by a regularly distributed point cloud with an adequate number of points in order to ensure that all local organ features in the current data set can be mapped to the shape model. The organ reference models are built as described in Section 3.3.1.

In order to preserve an anatomic correct shape during deformation and to prevent the user from mapping smooth regions to areas of high curvature, local weight constraints are added to the kidney model (compare Figure 3.10(a), larger points mean higher weight). This is done using the constraint editing method described in Section 3.3.2. Here, the kidney capsule is modeled with a 5 times higher rigidity than the area containing ureter and vessel connections. This leads to a more robust adaptation while being able to adapt to the regions of high frequency boundaries



(a) (b) (c)

Figure 3.10.: (a) 3D mesh deformation using a Gaussian weighting force [EKW09]. (b) Global shape preservation during lateral movement [EKW09]. (c) Local weights on the kidney shape: a soft area which maps the vessel and ureter connection region is embedded into a stiff capsule [EKW09].

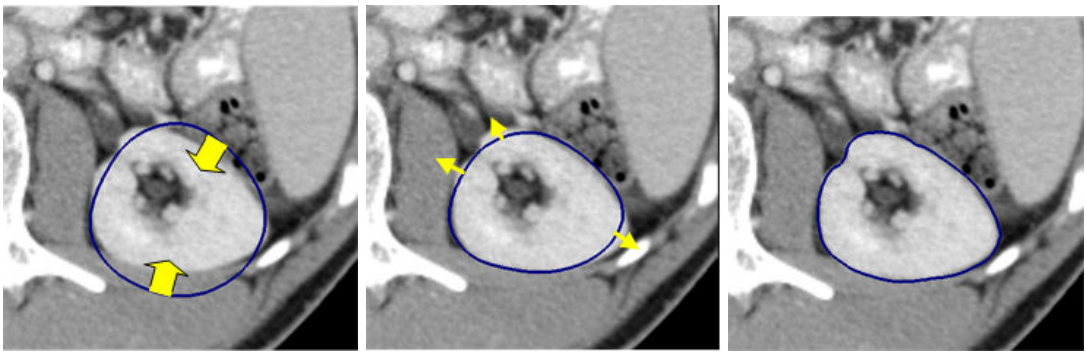
at the same time. In case of the cardiac left ventricle model, all constraints are set to 50, since the local curvature of the ventricle does not change much in the heart cycle.

3.3.3.3. User Guided Adaptation

The first step of the manual segmentation process is the selection of an organ and the placement of the according model in the data set by the user. After placement, the model can be scaled and rotated in order to ease the adaptation and to roughly align important feature points of the model to the underlying data (e.g. inferior and superior renal capsule, which denote the lower and upper boundaries of the kidney capsule, respectively).

The subsequent step is a fine grained segmentation by directly deforming the mesh. This is done by pulling the boundaries of the mesh towards the real image boundaries in the three 2D standard views of medical imaging (axial, sagittal and coronal image planes). The user driven force at a given point is propagated to adjacent points using a 3D Gaussian $G_\sigma(x,y,z)$ (compare Figure 3.10(b)). The user can switch between three different scales of the standard deviation σ of G_σ. Initially, it is suitable to select the highest value of σ, which results in a non-local or stiff deformation of the mesh around the user movement vector (Figure 3.11(a)). In order to adapt to areas of high curvature, σ may be lowered which results in a softer deformation until only points in a vicinity are affected (Figure 3.11(b)). This procedure is repeated in a couple of different slices using the standard views and changing the value of σ until the mesh is properly fitted to the data (Figure 3.11(c)).

In order to keep areas of corresponding points matched during the deformation process, lateral movement should not change the global shape of the organ. That means, if lateral user force is applied on the model's surface, local deformation should be propagated along the direction of the force (compare Figure 3.10(c)). This prevents self-intersection or folding of the surface. In order to achieve this, the user force propagation to the 3D Gaussian is computed for every interaction frame independently of the previous frame.



(a) (b) (c)

Figure 3.11.: Mesh adaptation in axial view [EKW09]. (a) Mesh after placement, scaling and rotation. The user forces the mesh into the real boundaries using stiff deformations (arrows). (b) Fine adaptation using soft mesh deformation. (c) End result of manual segmentation.

3.3.3.4. Mesh regularization

Relying only on the described mesh deformation would lead to highly irregularly distributed point clouds after adaptation. Moreover, the quality of the resulting meshes would directly depend on the number of refinement steps, since point correspondences would get worse with every step.

In order to use the user guided segmentation meshes as training input for the shape model generation, all point coordinates are optimized in real time such that the global shape of the reference mesh as well as the point distances are preserved.

Similar to [LB06], two energies E_{shape} and E_{force} are defined to regularize the mesh deformation. E_{shape} denotes a shape preservation energy defined as

$$E_{\text{shape}} = \sum_{i \in P} w_i \sum_{j \in N(i)} \left((p_i - p_j) - (r_i - r_j) \right)^2, \tag{3.1}$$

with p and r being the points of the deformed model and reference model respectively. The set of point indices P is the same for both r and p. $N(i)$ denotes the set of all direct neighbors of point p_i. w_i is a weight, that adds locally variant stiffness to the model as described in Section 3.3.3.2. The described term ensures that the point distances in the deformed mesh remain similar to the distances in the reference mesh.

The energy E_{force} contains the user movement force towards the boundaries of the organ and is defined as

$$E_{\text{force}} = \sum_{i \in P} (p_i - s_i)^2, \tag{3.2}$$

where s is the new point position resulting from the user force and weighted by G_σ *without* optimization.

The final point coordinates are now obtained minimizing

$$E = E_{\text{force}} + E_{\text{shape}}. \tag{3.3}$$

Equation (3.3) can be transformed into a linear system by setting the partial derivatives $\frac{\delta E}{\delta p_i}$ and $\frac{\delta E}{\delta p_j}$ to zero and bringing the resulting system to the form $A\vec{p} = \vec{b}$, with $\vec{p}$ containing the new point coordinates of the mesh. This overdetermined system can be solved in a least squares sense (cf. Section 3.5.1). Its solution is obtained by solving the normal equations $\left(A^{\mathrm{T}}A\right)\hat{\vec{p}} = A^{\mathrm{T}}\vec{b}$ with $\hat{\vec{p}}$ denoting the optimal solution vector for the new point coordinates. For performance reasons, x-,y- and z-coordinates are computed separately.

Since every point in the model mesh has a very limited number of direct neighbors (usually less than 7), most entries of A (and also $A^{\mathrm{T}}A$) will be zero. Therefore, sparse linear system solver [SG06] can compute the result very efficiently. Also note that matrix $A^{\mathrm{T}}A$ and A^{T} can be precomputed and loaded together with the model, because vector $\vec{b}$ contains all non-constant expressions. The whole optimization for a mesh of 2500 points is 70 ms on a 2.4 GHz Quad Core PC. The deformation can therefore be performed in real time which allows for a very smooth editing of the meshes. This is also very important for the use and acceptance of the application by the user since accuracy and speed should not be lower than a comparable conventional manual segmentation system.

3.3.3.5. Shape Model generation

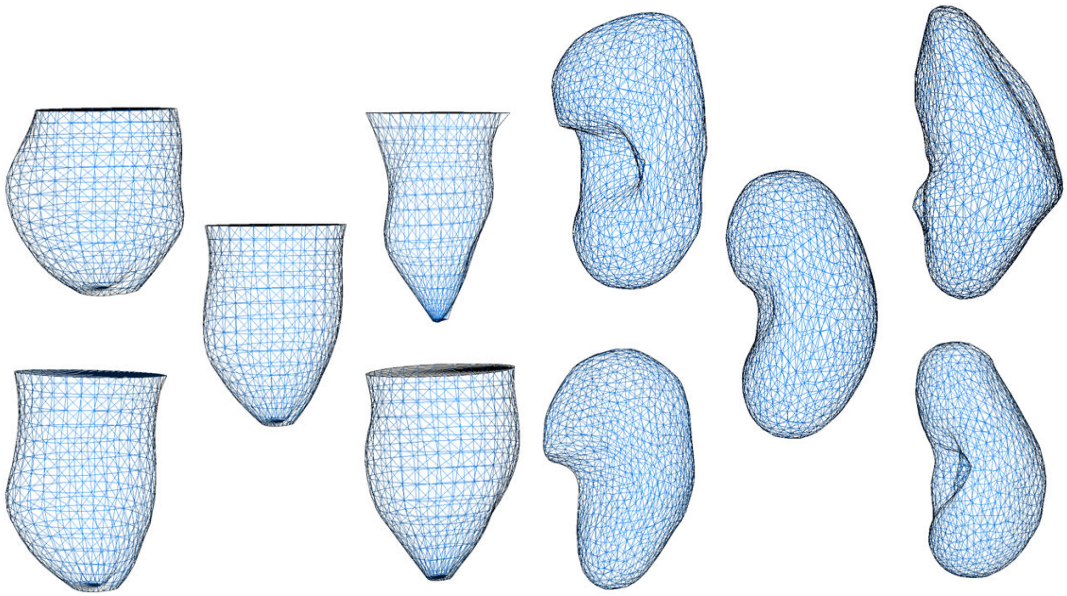
With the described procedure, a set of training meshes is created. As described in Section 2.5.3 a statistical shape model is built using the training meshes. Here, the shapes are aligned into a common coordinate system, i.e. the representation becomes independent of translation, scaling and rotation. Afterwards, Principal component analysis (PCA) is used in order to capture the statistics of the aligned training shapes (cf. Section 2.4.4 and Section 2.5.3).

In order to reduce the dimensionality of the model, axes with small variance are excluded from the model. The smallest dimension t is chosen such that $\sum_{i=1}^{t} \lambda_i$ captures 95% of the variance of the training data set, where eigenvalue λ_i describes the variance along the principal axis p_i. In model based organ segmentation, usually $\geq 95\%$ variance of the training set is kept [HM09], because the last principal modes are suspected to contain mainly noise. For a kidney model with 16 training meshes, PCA as described above leads to 11 principle axes. For the model of the cardiac left ventricle with 10 training meshes, 8 principle axes are retained.

The set of shapes modeled by the statistical shape model are all shapes $\hat{x}$ in the form $\hat{x} = \bar{x} + Pb$, where $P = (p_1|\ldots|p_t)$ is the matrix of retained eigenvectors and $\bar{x}$ is the mean shape. The shape parameters b_i are restricted to be in the interval $[-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]$ which is three times the standard deviation from the mean in each direction on each principle axis. As $\bar{x}$ and P are fixed, each shape $\hat{x}$ can be uniquely defined by the t-dimensional parameter vector $\vec{b}$.

Figure 3.12 shows the principal modes of variation for the built statistical shape models using smart landmarking. For each model, the mean (middle) is shown as well as the variation of the two largest Eigenmodes between $-3\sqrt{\lambda_i}$ (left) and $3\sqrt{\lambda_i}$ (right).

Assuming a Gaussian distribution of the anatomical shape variance of organs, the interval to restrict the shape parameters accounts for 99.7% of all possible shapes for the selected organ. Theoretically, this would be enough to precisely adapt to almost every plausible shape of the modeled organ. However, some information is lost by the dimensionality reduction of the linear PCA. Furthermore, usually not enough training shapes are available to accurately model the whole distribution of possible shapes. Lastly, it can be argued that the anatomical variance of organs is not Gaussian distributed. Kirschner et al. [KBW11] constructed a statistical shape model of a vertebra mixing lumbar with thoracic vertebrae in the training set. Here, a non Gaussian distribution



(a) (b)

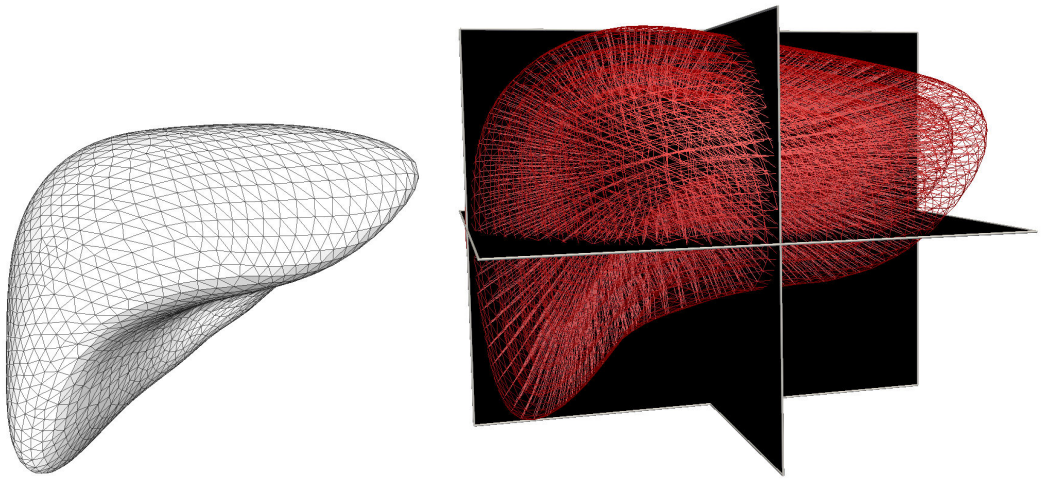
Figure 3.12.: Principal modes of variation for the SSMs of the LV (a) and kidney (b) using smart landmarking [EKW10]. The variation of the two largest eigenmodes between $-3\sqrt{\lambda_i}$ (left) and $3\sqrt{\lambda_i}$ (right) are shown together with the mean mesh (middle).

could be observed and therefore the use of non linear statistical shape modeling methods is proposed. It could be argued though, that modeling the lumbar and thoracic vertebrae independent from each other in a separate model could lead to a different result. In Section 5.3.4, the effect of applying non linear statistical shape models on landmark based models is discussed in detail. It will be shown that certain organs with high anatomical shape variation like the liver approximately follow a Gaussian distribution. In such cases, using linear statistical shape modeling is sufficient. Therefore, in this thesis linear PCA as described above is used for statistical modeling of organs.

The proposed smart manual landmarking method for building reference sets with corresponding landmarks in the context of statistical shape modeling as described in this section is evaluated in detail in Section 5.2.

3.3.4. Multi-layer geometry model

As it has been discussed in Chapter 2, surface model based approaches are the state of the art in segmentation of various organs [ZBG*07, ZHB06, CDNA07, EPS*08, WSH09, HvGSea09]. However, since only boundary information is used for deformation, a good initialization is needed and often problems with low contrast images occur. This is problematic since both points cannot be guaranteed in clinical scenarios. For example, in multi-phase CT of the liver usually several scans are taken at different time points of contrast agent saturation. At the early arterial phase, the liver parenchyma generally shows only very low contrast to other structures like muscles or stomach. Dealing with such kind of scans requires additional information beyond the standard surface-driven approaches used so far.

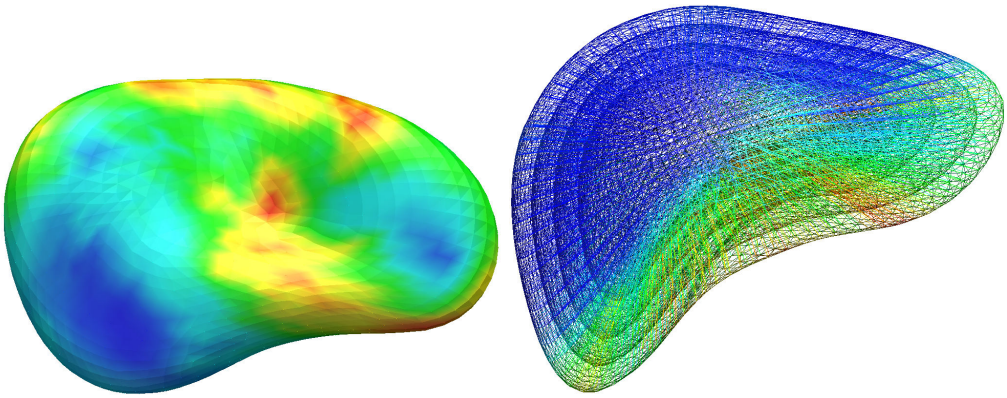


(a) (b)

Figure 3.13.: (a) Liver mean model created from a training base of reference shapes and (b) generated multi-layer model with depth links [ESW10].

As it has been shown in Chapter 2, several volumetric methods have been proposed that incorporate the organ's interior intensity information into the deformation process in order to make the segmentation process more robust. However, these approaches have only weak generalization capabilities and are often not compatible with standard boundary detection and shape preserving strategies. This often limits their applicability to well contrasted structures like bones or the lung in CT [HM08, BPCO10] which stands in contrast to the goal of volumetric models — making the segmentation of weakly contrasted structures more robust. Most methods are also computationally very demanding. Furthermore, current volumetric model based segmentation cannot be easily coupled with statistical shape modeling methods which has shown to be a requirement for robust segmentation of complex structures like the liver [HvGSea09]. Therefore, the majority of model based methods is still based on surface representations.

In this thesis, these points are addressed by the development of a multi-layer volumetric model that can detect inhomogeneities in the model's interior. In order to be able to combine the model with standard boundary and shape preserving strategies, it is based on an explicit landmark representation. Several ways exists to extend a landmark based surface model to incorporate image information from the model's interior. Barreira et al. [BPCO10] create a volumetric grid by adding additional points in the model's interior. Such a model has several advantages. It can be efficiently optimized and it allows for the direct local manipulation of the model in case a cavity or organ foreign tissue is detected around an interior point. However, since the number of interior points varies from shape to shape statistical shape modeling cannot be easily integrated. Additionally, exterior forces on the model's surface cannot be linearly propagated into the model's interior since often no straight connections exist between an interior point and a point on the surface where the force applies. The proposed model therefore consists of several layers with an identical amount of points in each layer. This enables the model to be coupled with statistical shape modeling, since every model for a certain organ consists of the same amount of points regardless of its shape. Corresponding points between layers are interconnected in order to propagate forces from the surface to the interior. The distance between layers and the total amount of layers can be limited in order to further reduce the computationally burden.



(a) (b)

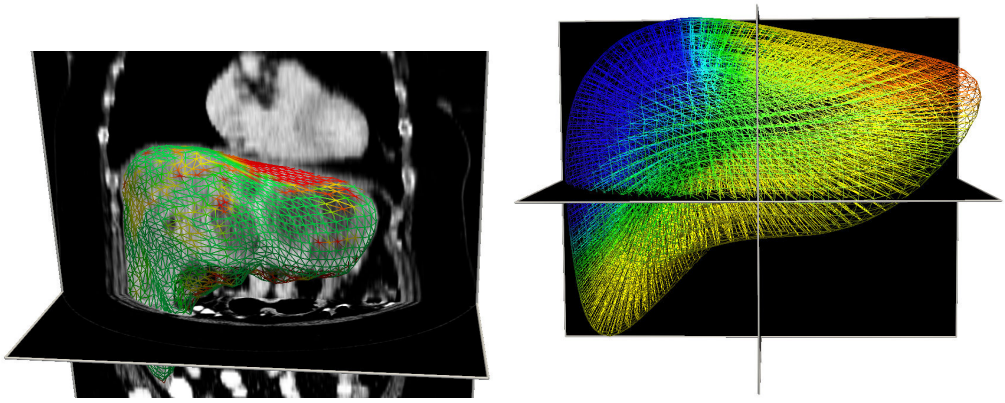
Figure 3.14.: (a) Mean local Gaussian curvature over all training shapes [ESW10]. (b) Local stiffness constraints mapped to all layers of the model (blue=stiff, red=flexible). It can be seen that shape variance can be differentiated well over the single areas of the liver [ESW10].

In order to build the multi-layer geometry model, a set of reference shapes is aligned using Procrustes alignment in order to transfer each shape into a common coordinate system. Afterwards, a geometric mean mesh is generated (see Figure 3.13(a)) that is used as the basis for the layer model generation.

Starting with the geometric mean, additional depth layers are added to the model by scaling the model with the factor f in inverse normal direction $-\vec{n}_i$ for every point p_i. The distance between the resulting layers should match the point normal's search length used for boundary detection (cf. Section 3.4) in order to guarantee a complete sampling of the model's interior. The new point coordinates for the kth layer L_k are given by $l_{ki} = (-\vec{n}_i f + p_i)$. In order to avoid self intersections of the resulting layers, L_k is manually optimized using the mesh deformation method of Section 3.3.3.

This procedure is repeated for the desired number of layers N_L. Afterwards corresponding points l_{ki} and $l_{k+1,i}$ in the layers are connected by depth links. The resulting layer model LM for a liver is shown in Figure 3.13(b).

In order to incorporate locally defined stiffness constraints into the multi-layer model, at each point of all training shapes the Gaussian curvature $C_g(t_i) = \kappa_1\kappa_2$ is computed as shown in Section 3.3.2 in context of the surface model. The mean Gaussian curvature over all training shapes is computed for every point (see Figure 3.14(a)). Since the multi-layer model consist of several layers, but the training data only provide information about the shape of the surface, the stiffness constraints from the outer surface layer are inherited in the internal layers. Furthermore, the multi-layer model distinguishes surface links between neighboring points from depth links between corresponding points of neighboring layers. In order to be able to treat both surface and depth links separately in the deformation and optimization stage, the stiffness constraints are now assigned to the links rather than to the model points. On that account, the stiffness values for two connected points in the layer model are interpolated. The corresponding layer link then stores this interpolated value as a link weight. The link weights control the local amount of stiffness of the model during deformation. Figure 3.14(b)) shows the resulting multi-layer model with local shape constraints.



(a) (b)

Figure 3.15.: (a) Color coded image gradient magnitude of a training image of the liver [ESW10]. Areas of fuzzy boundaries are shown in red. (b) Multi-layer model of the liver with color coded local intensity constraints denoting risk regions of potentially low contrast (e.g. the heart transition in the upper right lobe shown in yellow/red) [ESW10].

Setting the stiffness constraints for the internal layers as described above assumes that the elasticity of the organ's interior is the same as the elasticity of the organ's surface. While this assumption yielded satisfactory results for the cases tested in this work, it is a simplification. For example, the elasticity of interior vessels is likely to be different from the remaining organ tissue. Therefore, in the future, the stiffness constraints should be set based on elasticity experiments of the target organ to be segmented.

3.4. Model adaptation

In this section, the adaptation process of the surface model and the multi-layer model to the underlying image data is described. The key to robust model adaptation is finding appropriate image features that lie on the boundary of the structure to segment. In a landmark based model, usually for every point p_i one candidate point or attractor point a_i is defined based on the found image features. In the optimization stage, a new point position is computed that lies between p_i and a_i according to the defined shape preserving and the image driven forces (cf. Section 2.3.1.4).

3.4.1. Local appearance priors

In CT, many organs display areas that potentially contain a fuzzy boundary with the neighboring structures. In order to prevent the model from leaking, a local gradient threshold v_i is added to every model point p_i as well as an adaptation weight α_i which represents the quality of a found boundary point a_i in the deformation. Using non-pathologic cases of the training base, for every point t_i in each ground truth mesh the image gradient ∇g in the corresponding training image is computed (cf. Figure 3.15(a)). The threshold v_i for model point p_i is given by $\mu_i(\nabla g) - sd_i(\nabla g)$, with $\mu_i(\nabla g)$ and $sd_i(\nabla g)$ being mean and standard deviation of the gradient at t_i over the training set, respectively. Figure 3.15(b) shows the gradient thresholds color coded in a multi-layer model of the liver. $\alpha_i \in [0,1]$ is the current image gradient during deformation of p_i mapped linearly to

$[\mu_i(\nabla g) \pm sd_i(\nabla g)]$. During adaptation, α_i may be set to 0 by the boundary detection in order to deactivate the adaptation of point p_i if no valid boundary point is found or the quality of the boundary point is insufficient.

3.4.2. Rule based local boundary search

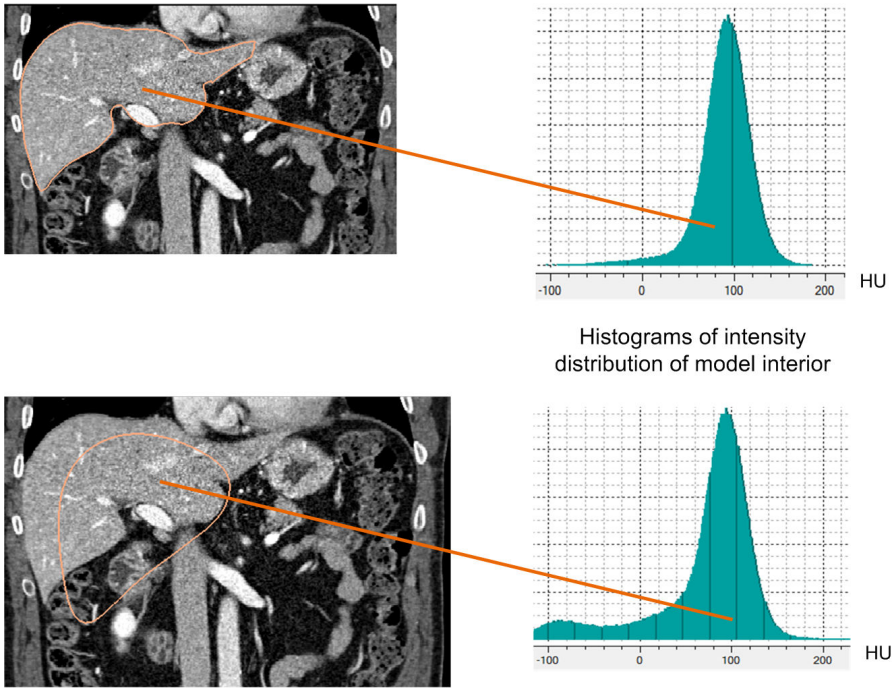


Figure 3.16.: Intensity distribution of the interior of a liver model that is perfectly adapted to the image (top) and intensity distribution of a liver model that is roughly aligned with the liver in the same CT image. The Gaussian distributed histogram peak denoting liver parenchyma is still prominent in the case of the roughly aligned model.

As it has been discussed in Section 2.6, several boundary detection methods exist. These methods are based on sampling local intensity patterns and the subsequent assignment of a fitness value to decide whether a point lies on the boundary or not [HM09, HvGSea09, WSH09, KUA*09, HMW07, ZBG*07]. In this thesis, this scheme is enhanced by a rule based system that deactivates boundary points in the adaptation in order to remove outliers. The remaining point positions are interpolated by the model optimization. The reason for developing such a scheme is the following: The rule based scheme can be combined with any standard boundary detection method that computes fitness values for boundary points. Therefore, different boundary detection methods for different application areas can be easily integrated. Furthermore, it boosts the performance of simple but generic and performant boundary detection methods.

The rule based scheme is based on a performant estimation of model points to be inside or outside the target organ. For that, image specific intensity constraints are calculated during adaptation, i.e. without prior training, by first thresholding the image to a rough intensity interval L_1 from that it is known that the target organ's intensity values will be inside. After examination of the intensity distribution of several organs like liver, kidneys, spleen, bladder and pancreas in ground truth segmentations, it is concluded that the intensity distribution of these organs in CT is

approximately Gaussian. Figure 3.16 shows an example of the intensity distribution of the liver. It can be seen that even if the model is only roughly aligned with the liver a prominent Gaussian can be observed in the corresponding histogram. From model initialization, it is assumed that there is a significant overlap of the model with the liver visible in the image so the histogram H of the model's interior clamped to L_1 contains a major peak representing liver tissue. Using H, a rough liver intensity range is estimated which is defined as $L_2 = [\mu_H \pm sd_H]$ with μ_H and sd_H being mean and standard deviation of H, respectively.

Figure 3.17 shows the distinguished cases of the rule based scheme. A model point p_i can either be in the state *inside* or *outside* (denoted as *I* and *O* in Figure 3.17) depending on whether $L_i = [\mu_{H_i} \pm sd_{H_i}]$ is $\subseteq L_2$, with H_i being the local histogram of an isotropically interpolated 9x9x9 neighborhood around p_i. Samples along the point's normal are then taken ($-/+$ denote the direction towards the model's interior and exterior in Figure 3.17). If a point is outside, only search direction '$-$' is applied. If the fitness value is too small, but $L_i \subseteq L_2$ at all sample positions, a default boundary point at the end of the normal (5 mm) is set (*D* in Figure 3.17) in order to avoid that the model's boundary stuck to local minima. If $L_i \subseteq L_2$ and the fitness value is sufficient, a boundary point (*B* in Figure 3.17) is assigned. The adaptation weight α_i of an *inside* point is set to 0 if $L_i \not\subseteq L_2$. This case denotes a contradiction during the search, i.e. the state of p_i is *inside* but something is found inside of the model that was not classified as organ tissue. Since it cannot be decided whether this is a vessel/tumor or some organ foreign structure, the point is deactivated in the adaptation.

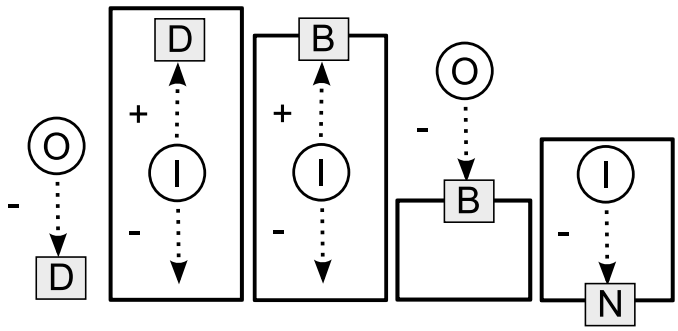


Figure 3.17.: Rule based boundary detection [ESKW10]. A model point either gets a default boundary point *D*, a regular boundary point *B* or a deactivated boundary point *N* assigned depending on the search direction $+/-$ and its state (*inside*/*outside* of estimated organ).

In this thesis, the local gradient is used as fitness criteria and the weights r from Section 3.4.1 are used as fitness thresholds. However, any other fitness criteria like the Mahalanobis distance between sampled profiles [DTT08] can be used without changing rules. This makes the system simpler and more generic in comparison to other rule based boundary approaches [KLL07]. Section 5.3.3 gives a comparison between standard profile based boundary search and gradient based boundary search in combination with the proposed rule based outlier removal. As it is shown, there is no benefit of using a profile based search over using the local gradient as fitness criteria if outlier removal is performed. This enables the use of the simple and performant gradient as fitness criteria.

3.4.2.1. Multi-layer model adaptation Logic

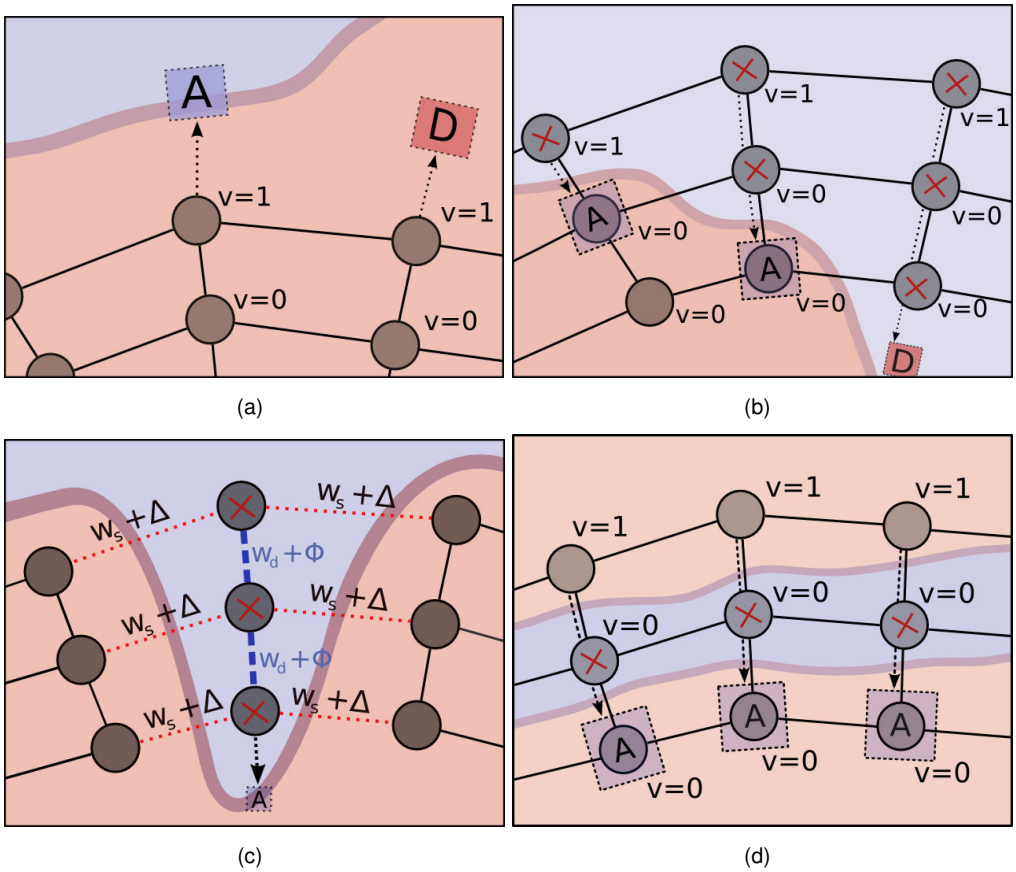


Figure 3.18.: Cases handled by the deformation logic according to point states (empty point = inside target tissue, cross-point = outside target tissue) [ESW10]. (a) Two cases where all points are inside of the target structure: (l.) an attractor A is found for a boundary point and (r.) no attractor is found and the default energy D is applied instead. (b) The target structure can not be found on one or more layers so attractors are assigned to points on lower layers which match the target structure. In the case where there is no match on any layer and within search range, a default attractor D is assigned (r.). (c) The flexibility of the model is dynamically adjusted to be able to adapt to a cavity. The flexibility w_s of surface links is increased while flexibility w_d for depth links is decreased. (d) The layer-model detects the boundary between two organs. Points on the lowest layer represent the new attractors.

Since the multi-layer model consists of several interior layers, a complexer adaptation logic can be applied in order to detect organ foreign tissue and in order to adapt the model to cavities. The adaptation logic evaluates samples taken from the model's interior and defines the forces for deforming the model's boundary: first, the outer layer searches for attractor points *A* on the target organ boundaries. The attractors may be found by standard gray value profile search [CTCG95] or gradient based boundary search [EPS*08, KLL07] as it has been discussed above. In case no attractor can be found along the surface's normal direction within 10 mm a default attractor *D* is assigned at the end of the normal.

The inner layer points now take gray value samples around an isotropic 9x9x9 neighborhood and indicate if they are inside or outside of the organ according to the same scheme as for the surface outlier removal. Figure 3.18 shows the cases handled according to the point states (inside/outside) defined before. The standard case is shown in Figure 3.18(a) where all interior points are inside of the organ. Surface points have either found an attractor *A* or got an default attractor *D* assigned. Since the model's interior is sufficiently homogeneous, all adaptation forces for internal points can be deactivated. Therefore, the force adaptation weight α_i is set to zero for the internal points and to one for surface points.

In Figure 3.18(b) the target structure cannot be found on one or more layers so attractors are assigned to points on lower layers which match the target structure. In the case where there is no match on any layer and within search range, a default attractor *D* at the end of the most inner point's normal is applied to the corresponding surface point.

In order to let the model adapt to small cavities, the flexibility of the mesh is dynamically adapted in case all points along a depth link are outside of the target structure (cf. Figure 3.18(c)). In this case the flexibility w_s of the surface links and w_d for the depth links is increased and decreased respectively by adding constants Δ and Φ. In the tests made in this thesis a change of 20% ($\Delta = -0.2, \Phi = 0.2$) was found to give the best results.

Using the described adaptation logic, the model is also able to recover from bad initialization by detecting the boundary in the model's interior as Figure 3.18(d) shows.

3.5. Model optimization

This section demonstrates how the surface model and the volumetric model with local shape priors as described in the previous sections can be optimized in a landmark model based segmentation regularization framework.

3.5.1. Surface model optimization

As described in Section 2.3.1.4, global prior shape methods incorporate a shape and an image term which are balanced by an optimization strategy in order to modify the shape representation, i.e. in order to deform the model. In terms of landmark model based segmentation, this relationship is commonly formulated mathematically as an energy minimization problem. The image term defines an external energy E_{ext} while the shape term represents an internal energy E_{int}. By minimizing the weighted sum of both energies

$$E = E_{\text{ext}} + E_{\text{int}}, \tag{3.4}$$

the new point coordinates of the model are determined. The internal energy is usually based on a point-neighbor regularization of the model's mesh. That means, a certain point of the model is only allowed to move according to some penalty function that depends on the first or nth direct neighbors of the point. For example, the penalty function can enforce that the angles between neighboring points are preserved [HMW07]. A popular type of point-neighbor regularization has been introduced by Lorenz et al. [LB06]. Here, the internal energy is defined as

$$E_{\text{int}} = \sum_{i \in P} \sum_{j \in N(i)} \left((p_i - n_j) - T\,(r_i - r_j) \right)^2, \tag{3.5}$$

with p and r being the points of the current model and reference model respectively. The set of point indices P is the same for both r and p. The points n are the direct neighbors of point

p_i. $N(i)$ is the set of point indices of n. T is a rigid transformation matrix and ensures that the deformed model can rotate and scale globally. This energy formulation has the advantage that it can be linearly optimized and does not contain any trigonometric expressions leading to a very efficient optimization even if the model consists of thousands or ten thousands of points. It also has been successfully applied to complex organ shapes like the heart [LB06, EPS*08]. Therefore, the model optimization proposed in this thesis is based on this formulation.

The external energy is denoted as

$$E_{\text{ext}} = \sum_{i \in P} \alpha_i \left(p_i - s_i\right)^2, \tag{3.6}$$

where s_i is the attractor point position for point p_i and α_i the confidence value for the same point according to the boundary search of Section 3.4.

The local shape constraints proposed in Section 3.3.2 are integrated into the described internal energy as

$$E_{\text{int}} = \sum_{i \in P} w_i \sum_{j \in N(i)} \left(\left(p_i - n_j\right) - T\left(r_i - r_j\right)\right)^2. \tag{3.7}$$

Here, w_i denotes the weight stored at point p_i in the training stage. Combining (3.6) and (3.7), an optimum for (3.4) can be found as follows: (3.4) can be written as

$$E = \sum_{i \in P} \left(\left(\left(\left(p_i - n_1\right) - d_{i,1}\right)^2 + \left(\left(p_i - n_2\right) - d_{i,2}\right)^2 + \ldots + \left(\left(p_i - n_{|N_i|}\right) - d_{i,|N_i|}\right)^2 \right) w_i + \alpha_i \left(p_i - s_i\right)^2 \right), \tag{3.8}$$

where $d_{i,j} = T\left(r_i - r_j\right)$. $|N_i|$ denotes the number of neighbors for point p_i. Since $d_{i,j}$ does not depend on p or n, it can be precomputed before every iteration. In order to determine the minimum of (3.7), the partial derivatives $\frac{\partial E}{\partial p_i}$ and $\frac{\partial E}{\partial n_j}$ of (3.4) are computed and set to zero. The partial derivative with respect to p_i is

$$\begin{aligned} \frac{\partial E}{\partial p_i} = \sum_{i \in P} & (2w_i(p_i - n_1 - d_{i,1}) + 2w_i(p_i - n_2 - d_{i,2}) + \ldots \\ & + 2w_i(p_i - n_{|N_i|} - d_{i,|N_i|}) + 2\alpha_i(p_i - s_i)) = 0. \end{aligned} \tag{3.9}$$

The partial derivative with respect to n_j is

$$\frac{\partial E}{\partial n_j} = \sum_{i \in P} \left(2w_i\left(p_i - n_j - d_{i,j}\right)\right) = 0. \tag{3.10}$$

(3.9) can be further written as

$$\sum_{i \in P} \left(\left(|N_i| w_i + \alpha_i\right) p_i - w_i n_1 - w_i n_2 - \ldots - w_i n_{|N_i|} \right) = \sum_{i \in P} \left(w_i \left(d_{i,1} + \ldots + d_{i,|N_i|} \right) + \alpha_i s_i \right) \tag{3.11}$$

and

$$\sum_{i \in P} \left(\left(|N_i| w_i + \alpha_i\right) p_i - w_i n_1 - w_i n_2 - \ldots - w_i n_{|N_i|} \right) = \sum_{i \in P} \left(w_i \sum_{j \in N(i)} d_{i,j} + \alpha_i s_i \right). \tag{3.12}$$

(3.10) can be written as

$$\sum_{i \in P} \left(w_i p_i - w_i n_j\right) = \sum_{i \in P} \left(w_i d_{i,j}\right). \tag{3.13}$$

The linear system resulting from (3.12) and (3.13) can be brought to the form $A\vec{p} = \vec{b}$, with $\vec{p}$ containing the new point coordinates of the mesh:

$$\begin{pmatrix} (|N\{i\}|w_1 + \alpha_1) & \dots & -w_1 & \dots & -w_1 & \dots \\ w_1 & \dots & -w_1 & \dots & \dots & \dots \\ w_1 & \dots & \dots & \dots & -w_1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} w_1 \sum_{j \in N(i)} d_{1,j} + \alpha_1 s_1 \\ w_1 d_{1,j} \\ \vdots \end{pmatrix}. \tag{3.14}$$

This overdetermined system can be solved in a least squares sense. Its solution is obtained by solving the normal equations $\left(A^{\mathrm{T}}A\right)\hat{\vec{p}} = A^{\mathrm{T}}\vec{b}$ with $\hat{\vec{p}}$ denoting the optimal solution vector for the new point coordinates.

A drawback of the point-neighbor regularization from Lorenz et al. [LB06] is that absolute lengths between points are used in the formulation. Using the rigid transformation T the model is only able to scale globally, but not locally. In case the model points have to move a lot from their position in the reference model such an absolute distance point-neighbor relation is too restrictive. As it will be shown in Section 4.2, some organs show strong local shape deviation from the learned training space. Therefore, in this thesis it is proposed to replace said point-neighbor relation with the Laplacian differential mesh operator

$$L(p_i) = p_i - \frac{1}{|N_i|} \sum_{j \in N(i)} n_j. \tag{3.15}$$

This operator was proposed by Lipman et al. [LSCO*04] for interactive mesh editing. L measures the deviation of a vertex from the centroid of its neighbors and thus preserves local detail properties of the surface. Furthermore, it is scale independent which allows for a better local adaptation. Integrating (3.15) into (3.7) yields

$$E = \sum_{i \in P} w_i \left(L(p_i) - L(r_i)\right)^2 - \alpha_i \left(p_i - s_i\right)^2. \tag{3.16}$$

$L(r_i)$ can be precomputed in every iteration. The partial derivatives of (3.16) with respect to p_i and n_j are

$$\frac{\partial E}{\partial p_i} = \sum_{i \in P} w_i \left(p_i - \frac{1}{|N(i)|}(n_1 + n_2 + \ldots + n_{|N(i)|}) - L(r_i) \right) + \alpha_i (p_i - s_i) = 0 \tag{3.17}$$

and

$$\frac{\partial E}{\partial n_j} = \sum_{i \in P} w_i \left(p_i - \frac{1}{|N(i)|}(n_1 + n_2 + \ldots + n_{|N(i)|}) - L(r_i) \right) \left(-\frac{1}{|N(i)|} \right) = 0, \tag{3.18}$$

respectively. The resulting linear system is

$$\begin{pmatrix} w_1+\alpha_1 & \dots & -\frac{1}{|N(i)|}w_1 & \dots & -\frac{1}{|N(i)|}w_1 & \dots \\ -\frac{1}{|N(i)|}w_1 & \dots & \frac{1}{|N(i)|^2}w_1 & \dots & \frac{1}{|N(i)|^2}w_1 & \dots \\ -\frac{1}{|N(i)|}w_1 & \dots & \frac{1}{|N(i)|^2}w_1 & \dots & \frac{1}{|N(i)|^2}w_1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \end{pmatrix}$$
$$= \begin{pmatrix} w_1 L(r_1)+\alpha_1 s_1 \\ -\frac{1}{|N(i)|}w_1 L(r_1) \\ -\frac{1}{|N(i)|}w_1 L(r_1) \\ \vdots \end{pmatrix}. \tag{3.19}$$

3.5.2. Multi-layer model optimization

Since the volumetric model proposed in Section 3.3.4 extends standard landmark based surface models by adding additional interconnected layers, it can be optimized using a point-neighbor relation analogous to the formulation in (3.7). However, since the multi-layer model needs the differentiation between surface and depth links, local weight constraints are defined as link weights rather than point weights. The proposed local Laplacian optimization of (3.16) for surface models cannot be applied to multi-layer models, because in the Laplacian formulation no explicit links between points exist. Therefore, a point-neighbor relation with link weights is proposed for optimizing all points of the multi-layer model.

Using $w_{i,j}$ as the link weight between point p_i and n_j (3.7) becomes to

$$E_{\text{int}} = \sum_{i \in P} \sum_{j \in N(i)} w_{i,j} \left((p_i - n_j) - T\left(r_i - r_j\right) \right)^2. \tag{3.20}$$

The partial derivative of (3.4) with respect to p_i is

$$\frac{\partial E}{\partial p_i} = \sum_{i \in P} w_{i,1}(p_i - n_1 - d_{i,1}) + w_{i,2}(p_i - n_2 - d_{i,2} + \ldots + w_{i,|N_i|}(p_i - n_{|N_i|} - d_{i,|N_i|} + \alpha_i(p_i - s_i) = 0 \tag{3.21}$$

and can be written as

$$\frac{\partial E}{\partial p_i} = \sum_{i \in P} (\alpha_i + w_{i,1} + w_{i,2} + \ldots + w_{i,|N_i|}) p_i - w_{i,1} n_1 - w_{i,2} n_2 - \ldots - w_{i,|N_i|} n_{|N_i|} = w_{i,1} d_{i,1} + \ldots + \alpha_i s_i. \tag{3.22}$$

The partial derivative with respect to n_j is

$$\frac{\partial E}{\partial n_j} = \sum_{i \in P} w_{i,j}(p_i - n_j - d_{i,j}) = 0. \tag{3.23}$$

The resulting linear system is

$$\begin{pmatrix} w_{1,1}+w_{1,2}+\ldots+w_{1,|N_1|}+\alpha_1 & -w_{1,1} & \ldots & \ldots & \ldots & -w_{1,|N_1|} \\ w_{1,1} & -w_{1,1} & \ldots & \ldots & \ldots & \ldots \\ \vdots & \ldots & \ldots & \ldots & \ldots & \vdots \\ w_{1,|N_1|} & \ldots & \ldots & \ldots & \ldots & -w_{1,|N_1|} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \end{pmatrix}$$
$$= \begin{pmatrix} w_{1,1}d_{1,1}+w_{1,2}d_{1,2}+\ldots+\alpha_i s_i \\ w_{1,1}d_{1,1} \\ w_{1,2}d_{1,2} \\ \vdots \end{pmatrix}. \tag{3.24}$$

3.6. Model initialization

An important step concerning the automization of a model based approach is model initialization. Though manual interaction is the most reliable method to place a model over the structure to segment and usually takes less than a minute to perform, automatic model initialization has many advantages. It enables the analysis of several scans in a row, for example in case the approach is applied to an existing clinical database. Furthermore, since no user interaction is required, segmentation or registration can be applied as part of data pre-processing. That means, such methods can be applied as soon as the image is acquired by the imaging modality which saves valuable clinician time. Lastly, automatic initialization enables accurate evaluation with reproducible results.

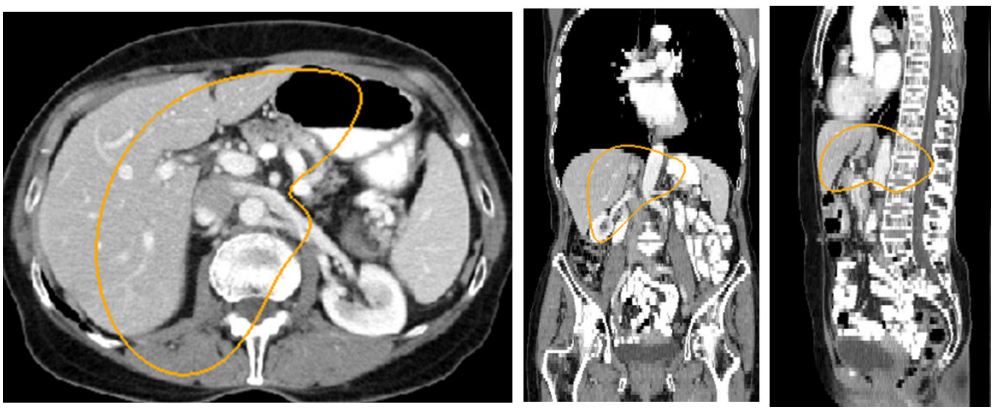


Figure 3.19.: Manual rigid placement of a liver model in a CT scan. The user moves, scales and rotates the three-dimensional model in the three standard cross-sectional views until it approximately fits to the data. Subsequently, the model may be further manually adapted using the mesh adaptation approach described in Section 3.3.3.

In this work, both manual and automatic initialization have been developed, since it cannot be guaranteed in a clinical scenario that automatic initialization always yields correct results. Figure 3.19 shows the manual initialization user interaction interface. The user first selects the model of the structure to segment. The three-dimensional model is then displayed in the center of the image in the three standard cross-sectional views. In each view, the model can be moved,

scaled and rotated until a sufficient overlay with the target structure is reached. In case the organ segmentation procedure fails or produces unsatisfactory results — e.g. in cases of strong pathologies — the mesh editing method described in Section 3.3.3 can be applied to manually adapt the model to the data. However, in most cases this is not necessary.

The goal of automatic initialization is to place the model in the image with at least a partial overlap with the target structure as well as with approximately correct orientation in the dataset. The needed amount of overlap and orientation quality depends on the robustness of the model based method used subsequently. Model based segmentation is based on detecting the boundaries of the structure to segment. As it will be shown in Chapter 5, standard statistical shape model adaptation is sensitive to model initialization errors. The multi-tiered adaptation method with increasing degrees of freedom during the adaptation process presented in Section 3.1 and Section 4.2.3, however, is more robust to such errors. Therefore, simpler, faster and more generic initialization methods in comparison to existing automatic model based approaches can be used.

In this work, a machine learning based initialization approach is proposed. Depending on the used features, learning based methods are very generic, computationally efficient and — if trained on a representative and sufficiently large training base — they are also very robust. In 3D medical imaging, usually the amount of training data is limited. The used features therefore must be discriminative and at the same time of relatively low dimensionality. The features used in this thesis are based on histograms of oriented gradients (HOG) which have been proposed by Dalal and Triggs [DT05] in 2D for human detection. The advanced variant of HOG features, called co-occurrence histograms of oriented gradients (Co-HOG) [RHZ*10] is not considered, because it would lead to a very high dimensional features vector if applied to 3D problems. The amount of required training data would therefore exceed the available resources in most clinical scenarios.

In the original approach, a 2D bounding box around the structure to detect is drawn manually in a set of training images. The bounding box is divided into a number of equally sized sub-boxes. In the case of human detection eight sub-boxes are used. For every pixel in each box, the local gradient direction is computed and discretized into one of eight directions ($d_1 = 0° - 45°$, $d_2 = 45° - 90°$, $d_3 = 90° - 135°$, $d_4 = 135° - 180°$, $d_5 = 180° - 225°$, $d_6 = 225° - 270°$, $d_7 = 270° - 315°$, $d_8 = 315° - 360°$). Afterwards, a normalized histogram of the discretized gradients is calculated for each sub-box. The feature vector representing the structure to segment consists of a concatenation of the histograms of all sub-boxes.

Being histogram based features, the advantage of HOG features in the context of 3D medical imaging is their robustness against scale changes and local deformations. Furthermore, in contrast to scale and rotation dependent features like Haar-features [SBZ*09], HOG features can be calculated in a single pass, since no rotated and scaled instances of the bounding boxes have to be applied to the image. This makes HOG features also computationally performant. It has to be noted though that HOG features are not rotationally invariant if the original bounding box is subdivided. However, in the context of CT or MRI imaging, the patient is usually recorded in a supine and head first manner which implies that all patients are roughly rotationally aligned. Additionally, in case of scans in prone position, the image can be mirrored before classification. If mirroring alone is not sufficient, a second classifier for prone position can be trained separately.

In this work, the original HOG approach is extended to 3D and applied to medical images. Since in CT a normalized intensity scale exist, the HOG feature vector is also enhanced by local intensity histograms. The resulting feature is therefore called 3D-HOGI (histograms of oriented gradients and intensities) in the remainder. The original bounding box is divided into 27 equally sized sub-cubes as shown in Figure 3.20. For each sub-box, a 26 bin histogram of discretized

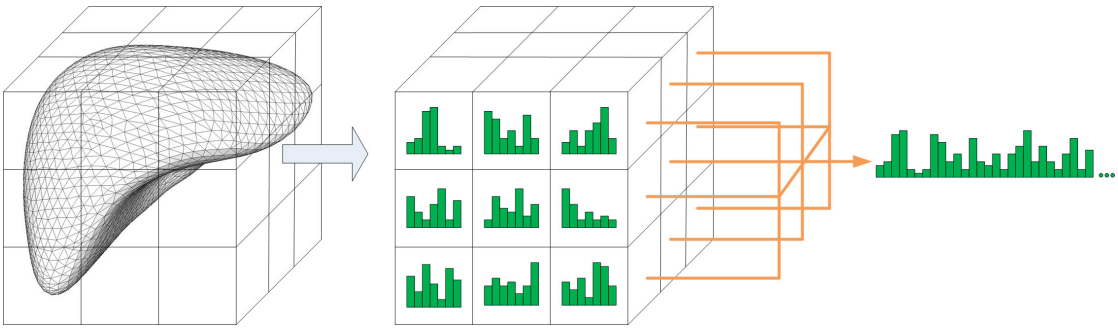


Figure 3.20.: Computation of 3D HOGI features. A bounding box around the structure to detect is subdivided into 27 sub-boxes. In each sub-box, a normalized histogram of gradient and intensity values is computed. The resulting feature vector consists of the concatenation of all histograms.

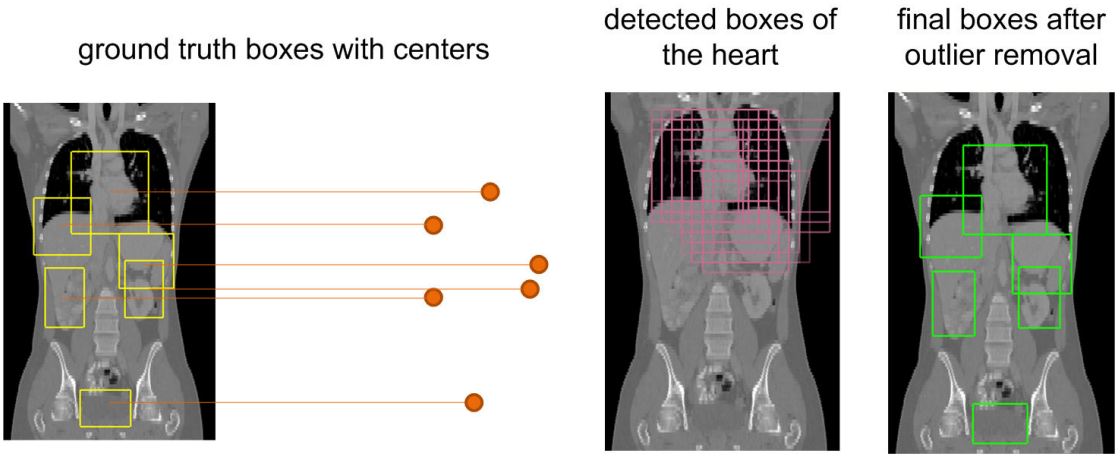


Figure 3.21.: Region detection of six exemplary structures in a CT image. Left: training image with ground truth bounding boxes and corresponding box centers. Middle: detected bounding boxes of the heart after running the classifier. Right: final detection result of all six structures after optimization.

gradients is computed. The directions are discretized analogous to the 2D variant in 45° steps. In order to suppress image noise and to encode homogeneous areas which often occur inside organs, an additional zero-direction bin is added to all local direction histograms. This bin stores all gradients which magnitudes are below a threshold. Furthermore, for each sub-box a 20 bin histogram of intensities is computed. In tests, this value yielded good results. An increase of the number of bins resulted in a decrease of detection performance which indicates an overfitting problem. The HU values are first clamped to the interval $[-500, 500]$ and then linearly mapped into the intensity histogram. It can be argued that a non-linear mapping with a closer sampling in the soft tissue range in which most organs are visible would lead to superior results. However, tests showed a slightly inferior performance using non-linear mapping. This could be an overfitting problem since the absolute HU values of CT images can vary on different scanners and on different patients. The described procedure leads to a vector of $26 + 1 + 20 = 47$ entries for each sub-box. Concatenation of all sub-box vectors yields the final feature vector for the structure to segment which consists of 1269 entries (cf. Figure 3.20).

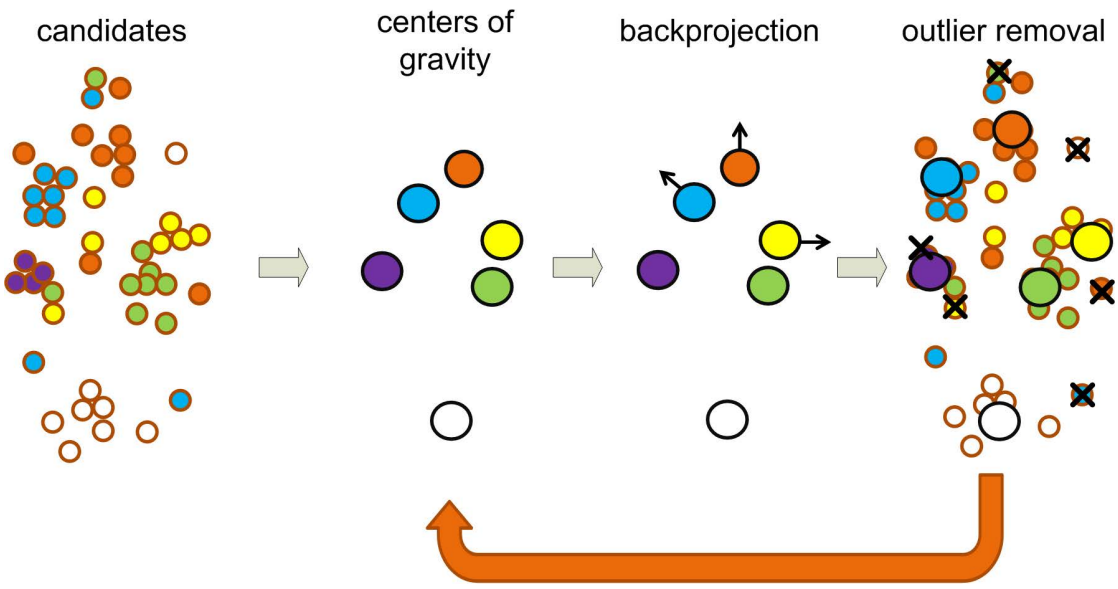


Figure 3.22.: Concept of the removal of outliers. From left to right: Centers of gravity of the candidates of all detected structures are computed. Centers of gravity are back-projected into the training space. Outlier removal is performed by discarding the box which is most distant to the back projected center of gravity (crossed circles). The process is repeated until only one box for each detected structure is left.

The 3D HOGI features are now used to train and detect structures in CT images. In a set of training images bounding boxes around the structures to segment are manually placed and adapted (cf. Figure 3.21 (left)). As an alternative, the ground truth bounding boxes can be automatically computed in case a binary ground truth segmentation of the structure of interest is already available. In this case, the bounding box is created axis aligned at the borders of the ground truth segmentation. The position, size and center of each box is stored. Negative feature samples for a certain structure are computed by shifting a bounding box of the same size as the ground truth box over the corresponding CT image. At each position, a 3D HOGI feature is computed and stored as a negative sample. During shifting, the box is not allowed to intersect with the ground truth box in order to avoid creating samples that partly show the structure of interest. Positive samples are created using the position of the ground truth box. In case a binary segmentation is available, additional positive samples are computed by randomly creating discrete directions in the area that is inside the ground truth box and outside the structure of interest. Using this procedure, artificial positive samples are created by variation of the neighborhood of the structure to segment inside the ground truth box. At the end, several hundreds positive and negative samples are created for each structure to segment in each training image.

For each structure, a classifier is trained in order to detect the bounding boxes in unseen images. Classification and regression trees [BFOS84] are used to create the classifier as described in Section 2.4.2. Classification and regression trees avoid over-fitting, are fast to train and performant on the classification of large datasets. Adaptive boost [FS96] is used as described in Section 2.4.3 in order to create a single strong classifier for every structure to detect. Here, 200 classification and regression trees with a maximum depth of 5 are combined to create the boosted classifier.

An exemplary detection result of a heart classifier can be seen in Figure 3.21 (middle). As can be seen, the heart has been detected. However, some false positives as well as many boxes with a partial overlap of the heart have also been detected. In order to remove such outliers, the relative positions between the structures to detect are taken into account. In the training step, the ground truth box centers of all structures have been computed for every training image. The concatenated center coordinates of all structures are stored in a single feature vector for every image. All feature vectors from training are aligned in a common coordinate system using Procrustes method. Principle component analysis is applied on the aligned vectors as described in Section 2.4.4. 95% variance of the training data is kept and the shape parameters are restricted to be in the interval $[-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]$ with λ_i being the eigenvalue of the ith principle axis. This allows for a back projection of arbitrary input vectors of the same size to the trained ellipsoid spanned by the given interval. Generally, the accuracy of the outlier removal process increases with the number of structures being added, since more information about the relative positions is available.

Figure 3.22 shows the outlier removal process. From the set of candidates of all detected structures (cf. Figure 3.22(left)) centers of gravity are computed (cf. Figure 3.22(middle-left)). The centers are projected into the trained space of box centers and further projected to the borders of the spanned ellipsoid in case they are outside of the ellipsoid. This leads to a movement of some of the centers of gravity (cf. Figure 3.22(middle-right)). Outliers are removed by discarding the box candidate which is most distant to the projected center of gravity for the corresponding structure (cf. Figure 3.22(right)). The process is repeated until only one box for each detected structure is left. Figure 3.21 (right) shows an exemplary detection result of six target structures (heart, kidneys, spleen, bladder and upper liver part).

3.7. Discussion

In this chapter, a new model based approach has been proposed that can be applied to segmentation and registration problems in medical imaging. As it has been discussed in Chapter 2, volumetric models are able to incorporate more image information during adaptation than surface based models. However, existing volumetric models are often computationally demanding, not easily extendable and not generalizable. Therefore, a multi-layered model has been proposed that can be combined with standard shape preserving methods and boundary detection strategies. Furthermore, it can be efficiently optimized at the same level of complexity as standard surface models.

Another common problem of model based methods is finding a proper way to leave the trained shape space during model adaptation such that the model may adapt to unseen cases without deforming into a non plausible shape. In order to address this problem, local shape constraints have been proposed that penalize the adaptation of the model locally, based on a prior statistical analysis of the training shapes. The local shape constraints can be either combined with standard surface based models or with the proposed multi-layered model. For both types of models, an optimization scheme has been described that incorporates the shape constraints.

Another issue of model based approaches is that the enhancement of shape models by the incorporation of new training shapes is usually very time-consuming. For each new training shape that is added, point correspondences have to be established again for the whole training set. A smart manual landmarking approach has been proposed to address this issue. It allows the

creation of reference sets where point correspondences have already been established. That way, new shapes can be easily incorporated to enhance an existing shape model.

Regarding model adaptation, a rule based outlier removal has been proposed that removes points with a low probability of being part of the organ's boundary. The outlier removal can be combined with any standard probabilistic boundary detection method.

Since the goal of this thesis is the development of an automatic model based method, an initialization scheme for the models has been proposed. It is based on a 3-dimensional extension of histograms of oriented gradients. The approach learns the relative positions between anatomical structures in order to determine the correct locations of organ centers and anatomical landmarks in the human body. The detected landmarks can then be used as starting positions for the models.

The developed methods are applied to common medical image segmentation and registration problems in the next chapter. Frameworks for automatic segmentation of the kidneys, the liver, the pancreas as well as low contrasted structures will be proposed. Furthermore, a deformable registration framework for multi-phase liver CT images based on the developed model based methods will be presented.

Parts of this chapter are based on publications of the author [EKW09, ETS09, EKW10, ES10, WKEK10, WEKK11b, WEKK11a, ETS09, ES10, ESKW10, ESW10].

4. Application to segmentation and registration

In this chapter, the application of the developed automatic model based methods to clinical use cases is demonstrated by solving segmentation and registration problems in CT. Several abdominal organs have been selected. In Section 4.1, a method for segmentation of the kidneys from contrast enhanced CT images is presented. Section 4.2 outlines a new approach for automatic segmentation of the liver in contrast enhanced CT data. Section 4.4 presents an automatic method for segmentation of the pancreas in contrast enhanced CT data.

The segmentation of low contrast structures is important for many clinical applications. Therefore, the applicability of the developed methods on such structures is demonstrated by segmenting the liver in arterial contrast phase images and the bladder in non contrasted CT scans (cf. Section 4.3).

In Section 4.5, the developed model based methods are used to register different contrast phase CT scans of the liver as well as pre-treatment and post-treatment CT liver scans.

The main contributions of this chapter are

- the proposal of a method for combining learned shape priors with observed shape deviation as demonstrated in Section 4.2.2,
- a multi-tiered adaptation method for model based organ segmentation outlined in Section 4.2.3,
- a method for automatic segmentation of the pancreas in contrast enhanced CT images presented in Section 4.4,
- an approach for segmentation of low contrast structures in CT data using a volumetric model as described in Section 4.3,
- an automatic approach for deformable registration of multi-phase liver CT images as well as pre-/post-treatment liver CT images outlined in Section 4.5.

4.1. Kidney segmentation in contrast enhanced CT scans

Kidney segmentation is an important medical image processing task needed for computer assisted navigation, computer aided diagnosis (CAD), or epidemiological research. However, time-consuming manual contouring is mostly used in practice due to the lack of automated and robust segmentation methods. In order to improve the clinical work flow, supportive tools for robust segmentation are therefore needed.

Model based techniques are frequently used to address this issue. Recently, Spiegel et al. [SHD*09] published a study of kidney segmentation using a *statistical shape model* that is initialized by a user defined seed point. A total of 41 CT scans of kidney pairs were available where 10 to 20 pairs were used to train the model and the remaining datasets were used for

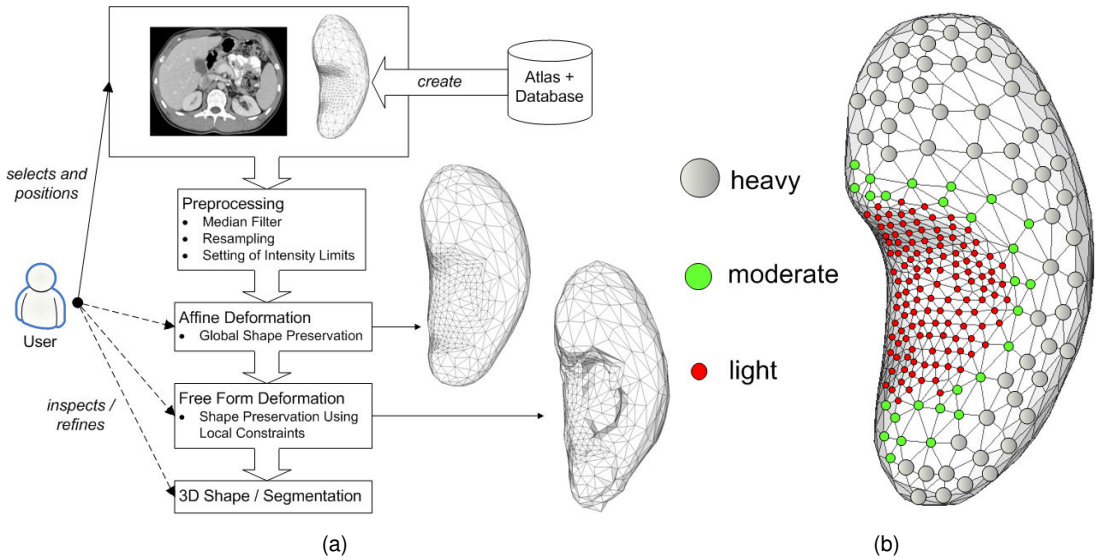


Figure 4.1.: (a) Overview of the kidney segmentation workflow [ES10]. Optional user interaction is possible at every main step of the workflow. (b) Kidney model with schematic internal point weights [ES10]. A dense and flexible area mapping ureter, veins and arteries is embedded into a stiff kidney capsule.

testing. An automatic hybrid region-based and model-based approach evaluated on a data base of 30 patients was published by Lin et al. [LLH06]. They report an average dice correlation coefficient of 88% compared to manual ground truth segmentation. Tsagaan et al. [TSK*01, TSKM02] proposed a deformable model of the kidney based on a non-uniform rational B-spline (NURBS) surface representation. They constrain their internal adaptation energy using statistical information about the local shape of the organ. In their work an average degree of correspondence to the gold standard of around 87% for a manual placement of the model is reported.

In this thesis, a segmentation framework based on locally constrained surface models as described in Chapter 3 is proposed. In order to prevent the model from deforming into implausible shapes, the local shape constraints described in Section 3.3.2 are used. They correspond to the local variability of the kidney, but do not constrain the global shape of the model. This knowledge is directly derived from the anatomy of the kidneys and the surrounding organs. The proposed framework also provides the possibility of real time manual refinement as described in Section 3.3.3 at every step of the segmentation process in order to allow an application in a clinical environment.

4.1.1. Workflow

Figure 4.1(a) gives an overview of the workflow of the kidney segmentation framework. A geometric mean model is generated and positioned in the data set. The positioning can be done by the user or using the automatic model initialization method presented in Section 3.6. Subsequently the model is automatically adapted using affine and constrained free form deformation steps. Since the shape variation of the kidney is not very high, a two step adaptation turned out to be sufficient in the tests. Therefore, no statistical shape information is incorporated.

4.1.2. Kidney model construction

Since in contrast enhanced CT the kidney appears as a highly inhomogeneous organ, a surface model representation is used. The kidney model consists of two parts: a geometric surface mean shape that is constructed from a set of ground truth manual delineations of the organ as described in Section 3.3.3.2 and a per-point definition of local constraints that are used to weight the internal energy in the adaptation process as described in Section 3.3.2.

In order to build a representative geometric mean model, 30 ground truth data sets of the left and right kidney are generated. As a result, 30 polygonal meshes for each organ with the same number of corresponding points are created. Using Procrustes method, the shapes are rigidly aligned in the same coordinate space. Afterwards, the point positions are averaged which results in the geometric mean (cf. Figure 4.1(b) for the left kidney). The final model consists of 1002 points.

Areas of low shape variance (encoded as *gray* weights in Figure 4.1(b)) are modeled with a 5 times higher rigidity $w = 100.0$ than regions containing high frequency boundaries $w = 20$ (shown as *red* in Figure 4.1(b)). The renal capsule (gray weights in Figure 4.1(b)) always has a smooth surface without any sharp frequencies. In addition, this area also adjoins structures of similar intensity (e.g. the transition to the spleen). This region needs a strong form preservation and a smaller influence of adaptation forces.

In addition to the shape preservation, the adaptation weight α penalizes the adaptation process per point (cf. Section 3.4.1). The value of α is determined using the set of ground truth segmentation meshes together with the original CT images. Images that show kidney pathologies have been sorted out. Since pathologies may occur at any point of the kidney, their intensity information is not useful for local image appearance analysis.

Local appearance priors v are computed according to Section 3.4.1. The coarse intensity interval for the rule based boundary search (cf. Section 3.4.2) is defined as $L_1 = [30\ \text{HU}, 290\ \text{HU}]$. In case of contrast agent saturated CT scans, the kidney's intensity values are usually inside this interval.

4.1.3. Kidney model adaptation

In order to remove noise and preserve the quality of edges, a median filter is applied to the input image. Afterwards, the image is rescaled to isotropic voxels. The new isotropic resolution is set to the minimum spacing of the x-,y- and z-direction in the original dataset.

The model adaptation is done in a two step process. First, only affine deformations are allowed. The point coordinates p are optimized by minimizing the following expression:

$$\sum_{i \in P} \left(T_A(p_i) - a_i\right)^2, \tag{4.1}$$

where a are the attractor points found in the boundary detection (cf. Section 3.4), T_A is an affine transformation and P is the set of point indices. In this way, the model coarsely adapts to the real boundaries without getting stuck in local minima. Figure 4.2 (a-f) shows an exemplary adaptation of the locally constraint kidney model using the proposed rule-based boundary search with the affine deformation described above. The initial placement of the model is shown in Figure 4.2 (a-c). The model after affine optimization is shown in Figure 4.2 (d-f).

After affine transformation, the model shape is coarsely aligned to the data set and is used as reference shape in the a constrained free form deformation step. The final segmentation is

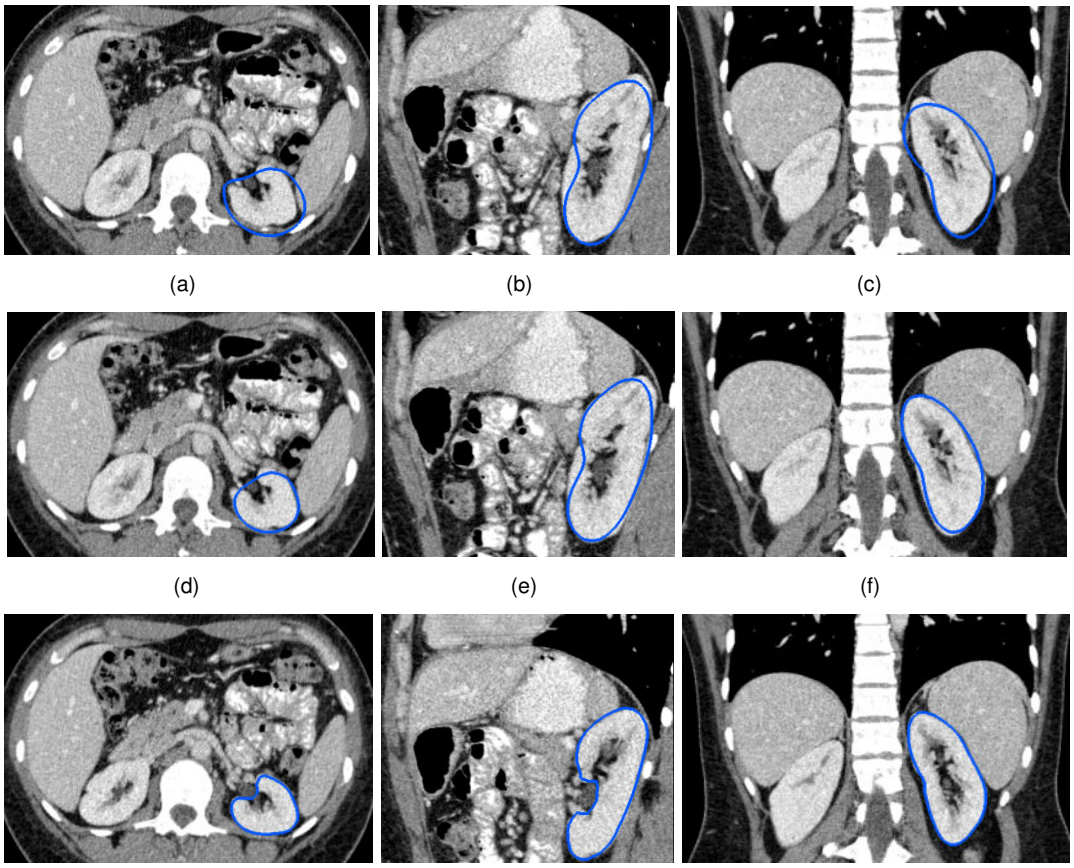


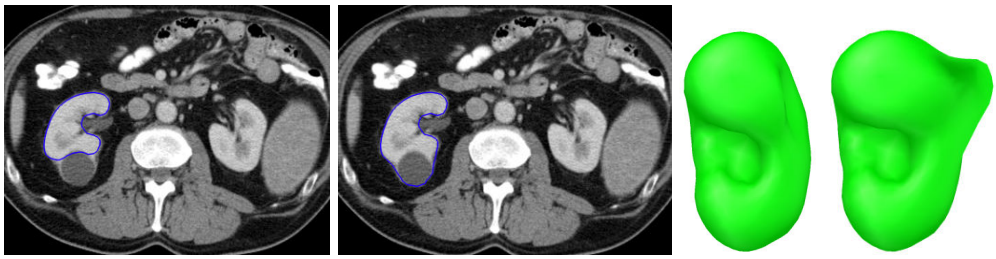
Figure 4.2.: Stepwise segmentation of the left kidney: (a-c) initial model positioning. (d-f) after affine transformation. (g-i) final result after constrained free form deformation [ES10].

gained by iteratively minimizing $E = E_{\text{ext}} + E_{\text{int}}$ as described in Section 3.5.1. Such a problem can be solved in a least squares manner. Using a sparse linear system solver [SG06] the result can be obtained in 40 ms on a 2.93 GHz Intel Quad Core for the given mesh of 1002 points. After some iterations of the free form deformation, the result only varies minimally and the adaptation process is stopped.

Figure 4.2 (g-i) shows an exemplary result of the constrained free form adaptation after the affine transformation step. At any step of this pipeline the user has the opportunity to refine the result if necessary using 3D mesh refinement tools (see Figure 4.3). This generally makes the segmentation more accurate in cases with prominent pathologies. The refinement tool is based on the interactive mesh editing method described in Section 3.3.3.

4.2. Automatic segmentation of the liver in contrast enhanced CT

Fully automatic liver segmentation in CT images plays an important role in medical imaging since it is a key enabler for many computer aided diagnosis applications. Furthermore, intervention planning and computer assisted navigation need a segmentation of the liver as a prerequisite for further procedures. In this Section, an automatic liver segmentation framework for contrast



(a) (b) (c) (d)

Figure 4.3.: (a) and (b) the user drags the mesh in the 2D standard views to add a lesion to the segmented kidney tissue [ES10]. (c) and (d) the resulting force is weighted by a Gaussian to deform the mesh three dimensionally [ES10].

enhanced CT images is outlined. In clinical diagnosis, usually the arterial and portal venous contrast phase are acquired. The portal venous contrast phase is easier to segment, since the liver parenchyma is typically contrast enhanced during that phase. In the arterial phase, usually only the liver artery is well contrast enhanced while the parenchyma shows only little saturation. This makes the segmentation of the arterial phase very difficult, since boundaries are often only barely visible. Therefore, most liver segmentation approaches focus on segmentation of the portal venous phase. In this thesis, a segmentation system for both arterial and portal venous phase is presented. The segmentation of the liver in the portal phase is addressed in this Section while the arterial phase is addressed in Section 4.3.

Despite the fact that portal venous liver segmentation is easier to perform than other contrast phases, robust automatic liver segmentation is still an unsolved problem. This is because shape and appearance of the liver may highly vary between scans due to the presence of pathologies, variation in patient pose and breathing cycle and due to a *per se* high anatomical variation. These issues are addressed by a multi-tiered model based framework that constantly increases the model's degrees of freedom. Figure 4.4 gives an overview of this workflow. As Ecabert et al. [EPS*08] do for the segmentation of the heart, the model evolution is started with an affine registration. Then classical statistical shape model adaptation [CTCG95, DTT08] is performed (cf. Section 2.5.3). The final segmentation is computed by a two-step adaptation that combines statistical shape modeling with learned and observed local curvature constraints.

4.2.1. Locally constrained shape model

A statistical shape model of the liver is built on 220 reference shapes of the liver with $N = 3612$ corresponding landmarks using the approach of Cootes et al. [CTCG95] (cf. Section 2.5.3). In order to reduce the dimensionality of the model, axes with small variance are excluded. The smallest dimension t has been chosen such that $\sum_{i=1}^{t} \lambda_i$ captures 98% of the variance of the training data set, where $\lambda_1 \geq \ldots \geq \lambda_{3N}$ are the eigenvalues of the according covariance matrix C. The set of shapes modeled by the statistical shape model are all shapes $\hat{x}$ in the form $\hat{x} = \bar{x} + Eb$, where $E = (e_1|\ldots|e_t)$ is the matrix of retained eigenvectors and $\bar{x}$ is the geometric mean, i.e. the initial shape to be adapted to the image. The shape parameters b_i are restricted to be in the interval $[-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i}]$ as discussed in Section 3.3.3.5.

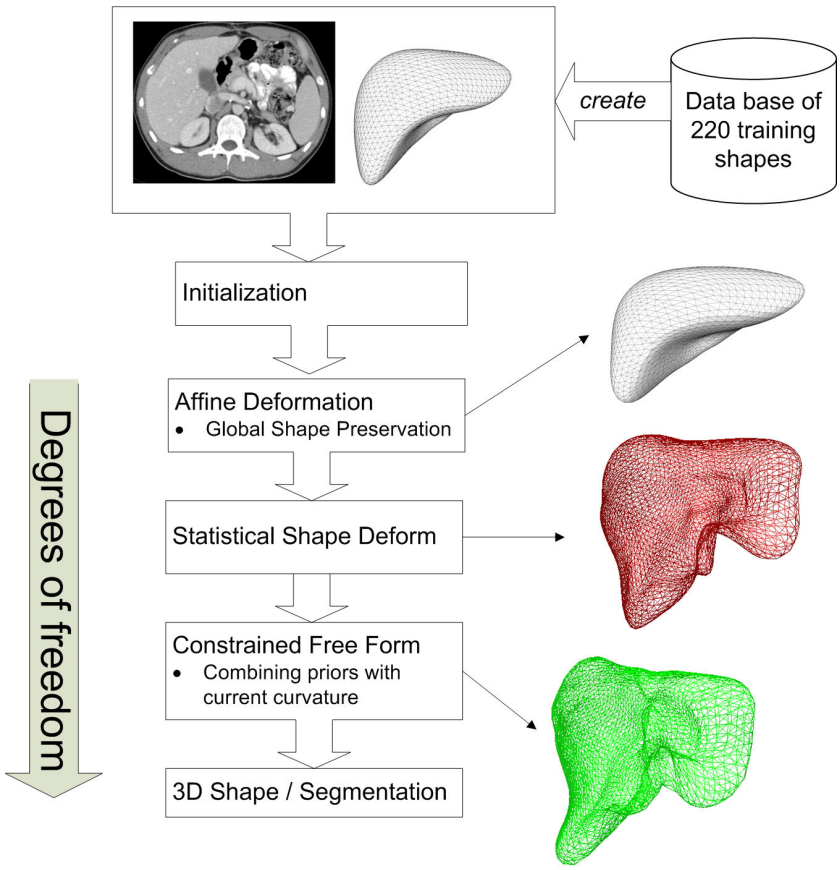


Figure 4.4.: Workflow of the proposed liver segmentation framework in contrast enhanced CT.

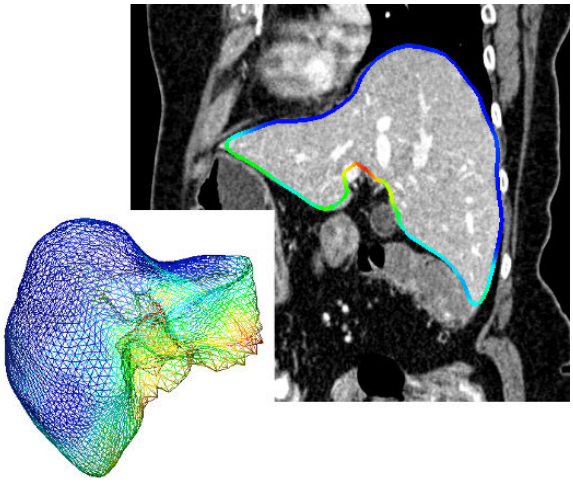
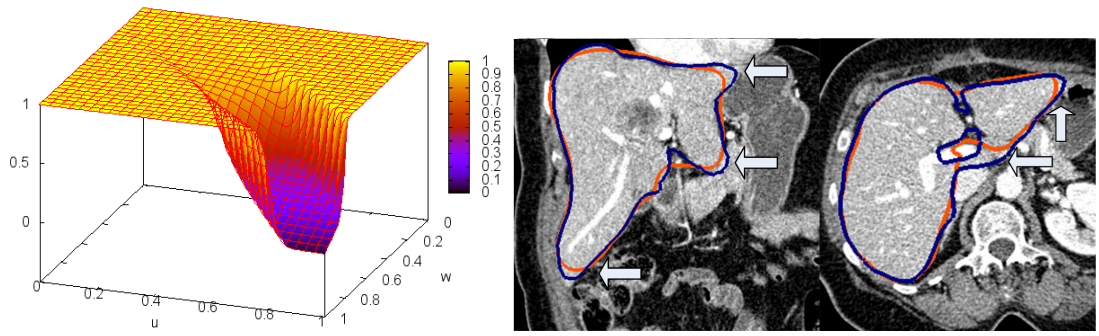


Figure 4.5.: Trained local shape constraints define regions of differing elasticity (model after adaptation) [ESKW10].



(a) (b)

Figure 4.6.: (a) Weighting function to combine trained local Gaussian curvature constraints *w* with observed local curvature *u* during adaptation [ESKW10]. (b) Using such local constraints, the multi-tiered model (dark outline) leaves the learned space (bright outline) without leaking into other neighboring structures [ESKW10].

4.2.2. Local shape and appearance priors

Local shape variation is incorporated into the deformation term according to Section 3.3.2. Since an explicit shape representation is used, correspondence is given through landmarks. At each point t_i of all 220 training shapes the Gaussian curvature $C_g(t_i) = \kappa_1\kappa_2$ is computed, where $\kappa_1\kappa_2$ is the minimum and maximum principle curvature. Afterwards the mean curvature over all training shapes is stored as weight w_i for every point p_i of $\bar{x}$ to encode regions of different elasticity (Figure 4.5).

Since a particular image may considerably deviate from trained samples, image specific shape constraints are directly derived from the current position of the shape model in the image. For every point p_i of the model during deformation the current local Gaussian curvature u_i in point p_i is combined with the trained curvature w_i and stored as

$$\delta_i = 1 - ((\pi/2 + \arctan(\beta(w_i + u_i - s_0))) \\ e^{-|(w_i-u_i)/\sigma|^\gamma}(w_i/\pi)). \tag{4.2}$$

The weight function of (4.2) is plotted in Figure 4.6(a). It ensures that the resulting weight δ_i becomes $1 - w_i$ if $u_i \approx w_i >> 0$ and 1 otherwise. That means the adaptation is restricted to those regions which show a high curvature w_i in the training *and* a similar high curvature u_i in the model shape after adapting to the statistical shape model space. High curvature is defined as $w_i, u_i > 0.4$ and the similarity of u_i and w_i as $u_i - w_i < 0.3$. This results in the parameter settings $s_0 = 1.2$ and $\sigma = 0.2$. In tests these values yielded the best results and were used for evaluation. γ and β control the smoothness of the function and are set to high values, e.g. 10 in order to clearly separate areas of high and low curvature.

In contrast to using only trained constraints the described combination of trained and observed Gaussian curvature allows the model to better adapt to shapes not fully covered by the training space (Figure 4.6(b)) while at the same time prevents the model from leaking into neighboring structures.

Both image specific shape and intensity constraints are updated in every iteration, i.e. with every deformation of the model.

4.2.3. Multi-tiered model adaptation

The model is initially placed in the dataset using the automatic initialization method presented in Section 3.6. In case no training data is available for the learning based approach, alternatively, the model can be placed axis-aligned below the right lower lung lobe. For automatic detection of the lobe simple voxel-counting operations similar to Kainmüller et al. [KLL07] are used. First, the background is extracted from the image by thresholding the image at -900 HU. The largest connected component is considered as the background. Another thresholding at -600 yields the lungs as the largest connected components. Assuming the patient is recorded in a head-first supine position, the component representing the right lung lobe can be directly selected.

A multi-tiered adaptation pipeline corrects translational errors and determines the correct orientation of the model as described in the following.

After initialization, the model is adapted to the image by a four-step evolution. The degrees of freedom are accordingly increased in every step in order to allow a robust adaptation to shapes which differ a lot from the initial model in terms of position, orientation and local shape.

Step one In the first step, only 12 degrees of freedom are allowed. That is, an affine transformation T_A to register the found boundary points a described in Section 3.4.2 with the current model points p is determined minimizing $\sum_{i\in P}(T_A(p_i)-a_i)^2$, with P being the set of point indices. This step is iteratively repeated until the sum of squared differences between corresponding points of two successive iterations does not differ significantly. The same stopping criteria will be used for the upcoming steps.

Step two The second step is a conventional statistical shape model guided segmentation using the restricted shape parameters b_i as defined before in order to approximate the real boundaries. The new model points $S(p)$ are determined by

$$S(p) = \bar{x} + Eb. \tag{4.3}$$

Step three The third step consists of step two coupled with the free-form deformation described in Section 3.5.1. However, the free-form deformation is additionally constrained by the extended shape priors δ_i defined in Section 4.2.2 which combines trained and observed local curvature constraints. The new coordinates for p are calculated by minimizing

$$\begin{aligned} E = \sum_{i\in P}\delta_i \sum_{j\in N(i)}\left((p_i-p_j)-(S(p_i)-S(p_j))\right)^2 \\ -\alpha_i\left(p_i-a_i\right)^2, \end{aligned} \tag{4.4}$$

with $N(i)$ denoting the set of all neighbors of point p_i. After convergence, the model already has a good alignment with the image. However, strong deviations from the trained space are still problematic. Some liver lobes are extremely long (cf. Figure 4.6(b) left). Since (4.4) uses a neighbor-distance-based regularization the model cannot evolve to such regions.

Step four For the fourth step, said regularization is replaced with the Laplacian differential mesh operator defined in 3.15 (cf. Section 3.5.1). The new model coordinates are calculated by minimizing

$$E = \sum_{i\in P}\delta_i\left(L(p_i)-L(S(p_i))\right)^2 - \alpha_i\left(p_i-a_i\right)^2 \tag{4.5}$$

After convergence of this last step, the segmentation is finished. Equations (4.4) and (4.5) form linear systems and can be solved very efficiently in a least squares sense (cf. Section 3.5.1).

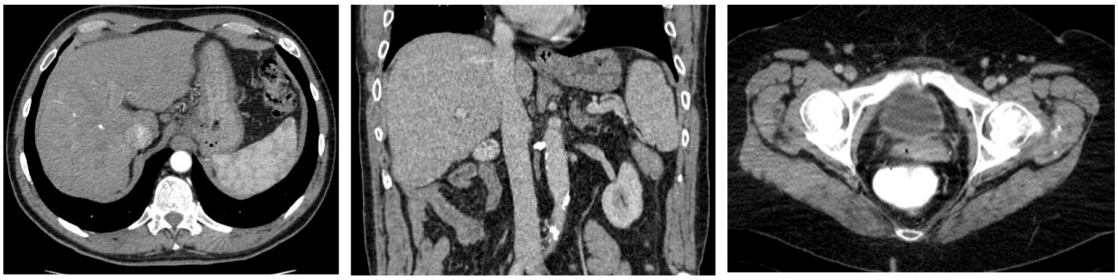


Figure 4.7.: Left and middle: arterial contrast phase CT scans of the liver. The liver parenchyma shows a weak contrast to neighboring structures such as muscles or the stomach. Right: CT scan of the bladder.

4.3. Segmentation of low contrast structures in CT

In CT scans, many organs show only low contrast at their boundaries to neighboring structures. Contrast agent application usually improves the contrast, but means a burden to the patient. Furhermore, for diagnosis of certain diseases no or little contrast agent is given.

This Section presents a framework for automatic segmentation of low contrast structures in CT based on two exemplary organs — the liver in the arterial contrast phase and the bladder in native non contrasted scans. Figure 4.7 (left and middle) shows a CT scan of the liver in the arterial phase. The liver interior is relatively homogeneous. Only some small vessels are contrast enhanced. Figure 4.7 (right) shows a CT scan of the bladder which also shows a relatively homogeneous intensity pattern in its interior. In order to segment these organs robustly in such low contrast images, the volumetric model presented in Section 3.3.4 is applied.

Figure 4.8 shows the built multi-layer models. For the bladder a 3-layer model has been created with a distance of 10 mm between layers which yielded the best results in the tests. Here, additional layers did not improve the segmentation results. The distance between layers has been chosen to match the search radius of the boundary detection in order to produce samples in the whole model's interior. For the liver model a 4-layer model has been created, because the liver is significantly larger than the bladder. 4 layers are not enough to completely sample the liver's interior. However, in tests more layers did not improve the results. The reason for this is probably that the area in the center of the liver does not add much information to the boundary adaptation of the model.

In case of the liver 220 training shapes were used. For the bladder 14 shapes were used. All training shapes have been created using the method described in Section 3.3.1. Using Procrustes alignment, each shape is transferred into a common coordinate system. Afterwards, a geometric mean mesh is generated that is used as the basis for the layer model generation described in Section 3.3.4. Local shape and appearance priors are set according to Section 3.4.1.

In clinical practice, usually both an arterial phase and a portal venous CT scan are acquired shortly after another. That means, both images are already coarsely aligned and the shape of the liver only differs by respiratory deformations. In such a case, the multi-layer liver model can be initialized with the shape of the portal venous segmentation result. This improves the segmentation quality of the liver in the arterial phase, because the portal phase segmentation is usually relatively accurate and can therefore be used as a good starting point. In case, only an arterial phase scan is available, the multi-layer model is initialized with the mean shape of

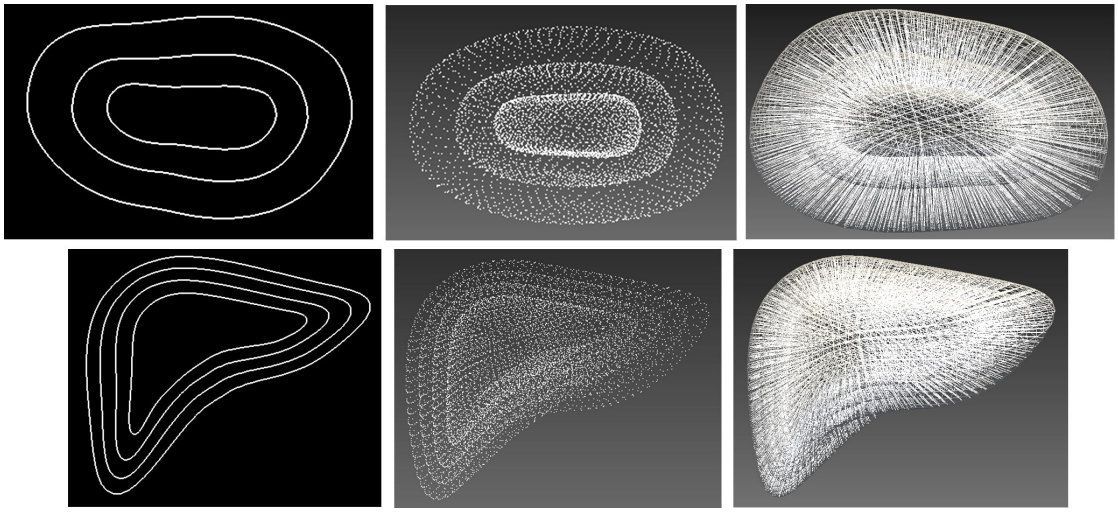


Figure 4.8.: Top row: 3-layer model of the bladder. Bottom row: 4-layer model of the liver. The first column shows the single layers of the models as a cross-sectional view. The second and third columns show the landmarks and links of both models.

the training set. Both liver and bladder models are positioned in the image using the method described in Section 3.6.

After placement in the datasets, the models are adapted to the image using the boundary detection procedure described in Section 3.4.2.1. During iterations, the models are optimized using the regularization term (3.20) defined in Section 3.5.2.

4.4. Segmentation of the pancreas in contrast enhanced CT scans

Automatic segmentation of the pancreas is a highly demanded tool for clinical practice, since it enables or can improve a variety of clinical applications. Pancreatic cancer like ductal adenocarcinoma has a high mortality rate (5-year survival below 5%) and is one of the most difficult cancers to treat [GCN08]. Patients are commonly examined using portal phase abdominal CT. Automatic delineation can support the clinician in tedious contouring work in order to cope with high-resolution data available nowadays. Furthermore, through exclusion of pancreas it can significantly help to develop automatic segmentation methods for other abdominal structures like intestine or enlarged abdominal lymph nodes for which no robust automatic segmentation solutions exist.

Segmentation of pancreas tissue in CT is difficult even for a human, since the pancreas head is always directly connected to the small bowel and can in most cases not be visually distinguished. This is also the area with the highest human observer variability. Furthermore, the small bowel may also contact the pancreas at any other place. Surrounding organs like liver, stomach and spleen are also problematic though contrast agent saturation can help to differentiate these organs from the pancreas. Figure 4.9 outlines the described problems on exemplary images.

Up to now, only few publications on automatic pancreas segmentation exist. Shimizu *et* al. [SOI*07] proposed a simultaneous segmentation framework for 12 organs including pancreas based on a combination of atlas-guided segmentation and level-sets. Evaluation on 10 non-contrast CT scans showed an average overlap of 32.5% for the pancreas. As the authors note,

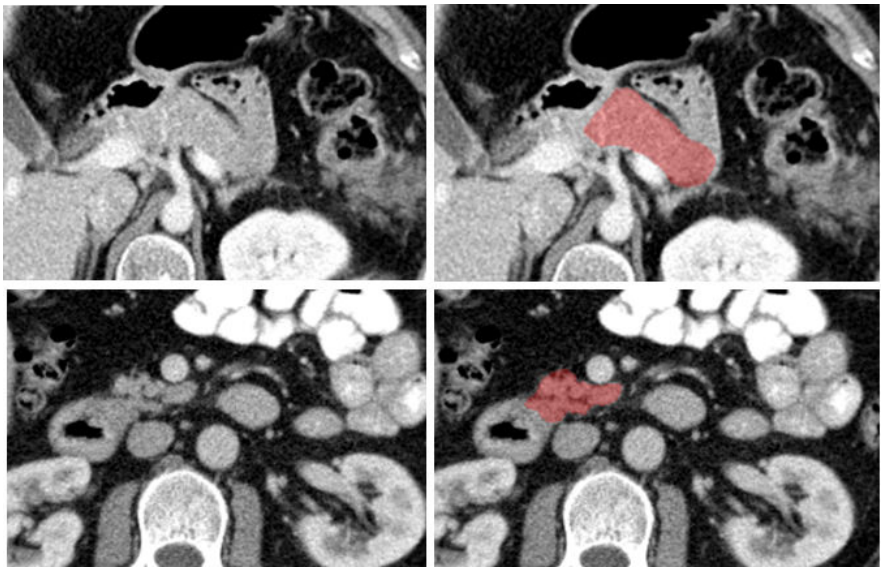


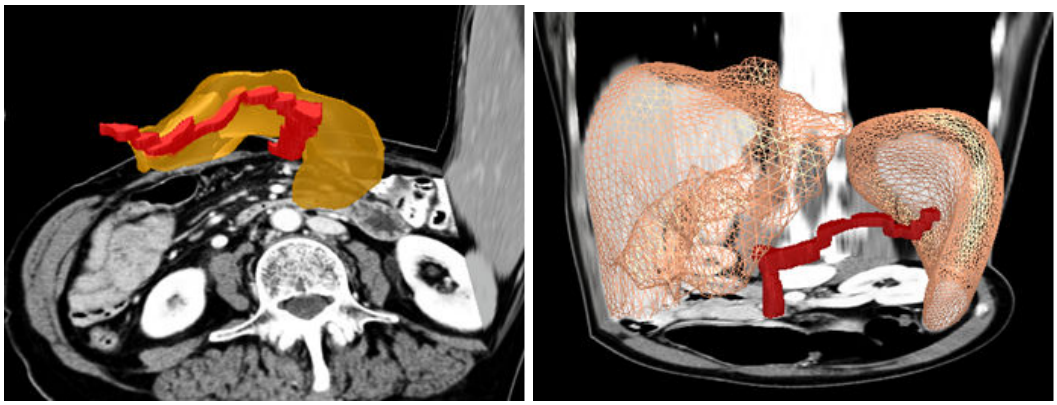
Figure 4.9.: Two exemplary axial views showing the difficulties in separating pancreas and contacting small bowel. Expert validated segmentation of the pancreas is marked in red [EKD*11].

the algorithm would need major revisions to make it applicable to contrast enhanced CT. The authors recently proposed a fully automatic pancreas segmentation system using three-phase contrast enhanced CT data [SKK*09]. They first register the three CT phases to a common space followed by a landmark based deformable registration with a certain patient chosen as reference. A rough segmentation is gained through a patient-specific probabilistic atlas guided segmentation. An intensity based classifier is used together with morphological operations for the final segmentation. This algorithm was also the winner of the medical image processing competition in Japan in the category pancreas [MSN*10]. Evaluation on 20 cases showed an average Jaccard index of 57.9%. As the authors conclude, their approach requires three-phase CT, because more signal information is available to guide segmentation. Kitasaka *et* al. [KSM*08] proposed a method to extract the pancreas region from four-phase CT based on estimation of organ distributions using expectation maximization and subsequent fine segmentation using a modified region growing algorithm. Using a three-point scale on 22 cases, segmentation quality was jugded based on visual inspection as FINE in 12 cases, MEDIUM in 6 cases and POOR in 4 cases, where FINE means little over- and under-extraction and POOR no overlap at all.

4.4.1. Support structure detection

Since pancreas tissue often cannot be visually distinguished from the small bowel, physicians incorporate prior knowledge about the

1. pathway of the pancreas relative to the splenic and superior mesenteric vein. The head of the pancreas always contact the superior mesenteric vein while body and tail follow the pathway of the splenic vein (cf. Figure 4.10(a)).
2. expected shape of the pancreas. The head of the pancreas is roughly spherical. Its body and tail are of tubular shape.



(a) (b)

Figure 4.10.: (a) Pancreas following the vessel path from superior mesenteric to splenic vein [EKD*11]. (b) Extracted liver, spleen and vessel path [EKD*11].

In this thesis, the same clinical knowledge is incorporated using a learning based approach. In order to segment the pathway of 1.), liver and spleen are segmented. Vessel enhancement filtering is then used to segment the vessel system between both organs. The system is transformed into a graph representation from which the desired pathway is extracted. Subsequently, a classifier is built that learns the spatial relation between pancreas tissue and the segmented pathway as well as incorporates local texture patterns around the pathway. In order to incorporate the shape knowledge of 2.), a constrained statistical shape model as described in Section 4.2 is used after classification to yield the final segmentation.

4.4.1.1. Automatic organ segmentation and vessel detection

In order to utilize the described anatomical information, liver and spleen are automatically detected and used as starting points for detection of splenic and mesenteric veins. Liver and spleen are automatically segmented by adapting the constrained surface models to the data using the approach described in Section 4.2. The axis-aligned bounding boxes $\vec{l}$ and $\vec{s}$ of the liver and spleen masks are computed and a pancreas region bounding box $\vec{p}$ is generated. $\vec{p}$ is parameterized by its minimum and maximum axes-coordinates:

$$\begin{aligned}\vec{p} = ((\vec{l}_{xmax} - \vec{l}_{xmin})/2, \min(\vec{l}_{ymin}, \vec{s}_{ymin}),\\ \min(\vec{l}_{zmin}, \vec{s}_{zmin}), (\vec{s}_{xmax} - \vec{s}_{xmin})/2,\\ \max(\vec{l}_{ymax}, \vec{s}_{ymax}), \max(\vec{l}_{zmax}, \vec{s}_{zmax})).\end{aligned} \tag{4.6}$$

All subsequent calculations are limited to the region described by $\vec{p}$.

The vessel system between liver and spleen is detected including portal, splenic, superior mesenteric vein and connected vessels in $\vec{p}$. Region-Growing approaches usually lead to leakage in the abdominal area. In order to enhance tube-like structures, the vessel enhancement filter described in [ERS08] is applied since it uses a parameter-free vesselness function and provides a mathematical basis to automatically compute optimal filter thresholds for the type of vessel to detect. The vesselness response for scale σ is computed as

$$S_{opt}(\sigma) = \sigma^{-\frac{3}{2}} \cdot \kappa \cdot (\frac{2}{3}\lambda_1 - \lambda_2 - \lambda_3), \tag{4.7}$$

where $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of the local Hessian sorted by size. Note that σ is equal to the vessel diameter to be detected. κ is an isotropy factor to avoid detection of plate like structures [ERS08]:

$$\kappa = 1 - \frac{||\lambda_2| - |\lambda_3||}{|\lambda_2| + |\lambda_3|} \qquad \in [0,1]. \tag{4.8}$$

Three different scales $\sigma \in \Sigma = \{2,3,4\}[mm]$ are used to ensure the whole vessel system of interest is detected. At each position the scale with the maximum output is taken as the final output. The optimal response [ERS08] computed for scale σ is:

$$r = d \cdot \gamma = d \cdot \kappa \cdot \sqrt{\frac{6}{5}} \pi^{\frac{3}{4}} \sigma^{\frac{3}{2}}. \tag{4.9}$$

The only unknown in (4.9) is d which is the radial vessel gradient. Since the filter is applied to the whole area in $\vec{p}$, d will vary depending on the surrounding tissue. A window of valid responses is defined as $\chi = [h_l, h_u]$. h_u can be easily computed as $h_u = (m_{max} - m_{mean}) \cdot \gamma$, with m_{max} and m_{mean} being the maximum and mean intensity inside the liver. This is the response for a small radial gradient from contrasted liver tissue to contrasted vessel. The lower bound $h_l = (m_{max} - f_{mean}) \cdot \gamma$ denotes the response for a big radial gradient from fat tissue f_{mean} to contrasted vessel. In the experiments f_{mean} was set to -120 HU. Using χ on the output of $\max(S_{opt}(\sigma_i))\forall i$ yields a binary mask of the vessel system while suppressing responses from other tube-like structures. In order to select only the vessel system connected to the portal vein, the largest connected component starting inside the liver is selected. The result is a binary mask M_v that contains the vessel system between liver and spleen.

4.4.1.2. Graph creation and landmark detection

The pancreas closely follows two vessel branches: the superior mesenteric vein and the splenic vein. These branches have to be extracted from the before segmented vessel system M_v. In particular, three branch points p_1, p_2, p_3 need to be identified. p_1 denotes the first major branch point of the superior mesenteric vein, p_2 is the branch point between portal and splenic vein and p_3 denotes the end of the splenic vein. p_3 can be directly computed without further processing as the intersection point of the binary mask M_v with the spleen segmentation mask. For detecting p_1 and p_2, a graph representation of M_v is built. Using the method described in [DOL10] M_v is decomposed to a set of sub-branches $B = (\vec{b}_0, \vec{b}_1, ..., \vec{b}_{n-1})$. Each branch contains two end points. From B, the branch $\vec{b}_l$ that origins from the liver mask is selected. The end point of $\vec{b}_l$ that is not inside the liver is selected as p_2. From p_2, there is one direct branch to p_3 and an additional branch denoting the superior mesenteric vein. Therefore, the end point of this second branch is p_1. M_v is accordingly cropped from p_1 to p_3 which yields the desired vessel pathway. Figure 4.10(b) shows the extracted vessel path together with the segmented organs.

4.4.1.3. Spatial anatomy descriptor

Global appearance based features successfully used for classification of other organs are generally inaccurate for pancreas detection, because its tissue often is indistinguishable from the small bowel. In order to overcome this problem, in this work, it is proposed to learn the spatial relationship between the before extracted vessel path and the pancreas position. The vessel path $\vec{v}(l) \in \Re^3$ from p_1 to p_3 is parameterized by its normalized length $l \in \{0,1,...,100\}$. Let $\vec{D}(\vec{t})$ be the vector field of the distance transform of M_v at position $\vec{t} \in \Re^3$ and $d(\vec{t})$ the signed distance

from $\vec{t}$ to the closest point in M_v. Let $\vec{v}(L(\vec{t}))$ be the intersection of $-\vec{D}(\vec{t})$ with the vessel path. A feature vector:

$$\vec{F}(\vec{t}) = (\vec{D}(\vec{t}), d(\vec{t}), L(\vec{t})) \tag{4.10}$$

encodes the local spatial relationship between vessel path and pancreas tissue needing only a small set of features. For efficiency reasons, $\vec{F}(\vec{t})$ is only computed for $d(\vec{t}) < N$ mm. In the experiments $N = 80$ always fully covered the pancreas and also included a large safety region around it. The advantage of this encoding in contrast to probabilistic atlas based registration [SKK*09] is two-fold. First, no time consuming and potentially error prone deformable registration is needed. Additionally, the feature space of $\vec{F}(\vec{t})$ exhibits a better de-correlation of the data compared to the Cartesian coordinate system of a probabilistic atlas. Since the distance of the pancreas to $\vec{v}(l)$ does not deviate much over l, the length-distance representation of (4.10) yields a much more compact data distribution. Therefore, simpler classification models can be used to achieve a good classification.

4.4.1.4. Appearance texture descriptors

In order to describe the pancreas tissue around the vessel path $\vec{v}(l)$, meaningful texture features are needed. In contrast to [SKK*09], local features are applied, i.e. the features are computed along perpendicular vectors around $\vec{v}(l)$ (cf. Figure 4.11(a)). The idea is that the pancreas always lies near $\vec{v}(l)$ so the variance of the local texture between $\vec{v}(l)$ and a position $\vec{t}$ should depend directly on the distance of $\vec{t}$ to $\vec{v}(l)$. Two descriptors are built that encode dominant frequency characteristics of the local texture around $\vec{v}(l)$. Let $\vec{P}(\vec{t})$ be a vector of length N with sampled intensities from $\vec{v}(L(\vec{t}))$ to $\vec{t}$. $\vec{P}(\vec{t})$ is padded with zeros if $d(\vec{t}) < N$. A frequency descriptor $\vec{\Phi}(\vec{t})$,

$$\vec{\Phi}_k(\vec{t}) = \sum_{n=1}^{N} \vec{P}_n(\vec{t}) \cos\left(\frac{\pi}{N} k(n+0.5)\right) \tag{4.11}$$

is built to map the intensity distribution along $\vec{D}(\vec{t})$ by computing the discrete cosine transform of the intensity profile. In the experiments, the largest $k = 0,..,10$ coefficients were kept to build $\vec{\Phi}$. Since the pancreas tissue is relatively homogeneous and lies nearby the vessel path $\vec{v}(l)$, the proportion of high frequencies should be low at small distances from $\vec{v}(l)$ and larger if the texture becomes inhomogeneous, e.g. between a transition from pancreas to small bowel.

The second texture descriptor $\vec{W}(\vec{t})$ encodes the amount of intensity variation peaks along $\vec{D}(\vec{t})$ by applying the Mexican hat wavelet on the volume V and integrating its negative and positive responses separately. Let the mapping $c : [0,1] \to \Re^3$ be defined as $\vec{c}(p) = \vec{t} - p\vec{D}(\vec{t})$.

$$\vec{W}_\eta^{\pm}(\vec{t}) = \int_{p=0}^{1} [V(\vec{c}(p)) * \nabla^2 G(\vec{c}(p))]^{\pm} dp, \tag{4.12}$$

where $G = \frac{1}{\sqrt{(2\pi)^3}\eta^3} e^{-\frac{x^2+y^2+z^2}{2\eta^2}}$ is a Gaussian and ∇^2 is the Laplace operator. The standard deviation of G is varied using two scales $\eta \in \{1,2\}[mm]$. The idea of this descriptor is that it will accumulate more responses the more different the tissue between $\vec{t}$ and $\vec{v}(l)$ is. Pancreas tissue usually only leads to two strong responses: one from vessel or fat to pancreas and the other from pancreas to other tissue.

The final pancreas tissue feature vector is built as

$$\vec{Z}(\vec{t}) = \{\vec{F}(\vec{t}), \vec{\Phi}(\vec{t}), \vec{W}_\eta^-(\vec{t}), \vec{W}_\eta^+(\vec{t}), \vec{H}(\vec{t})\}, \tag{4.13}$$

with $\vec{H}$ containing low level features at $\vec{t}$ like intensity, gradient and nonlinear combinations as described in [ZBG*07].

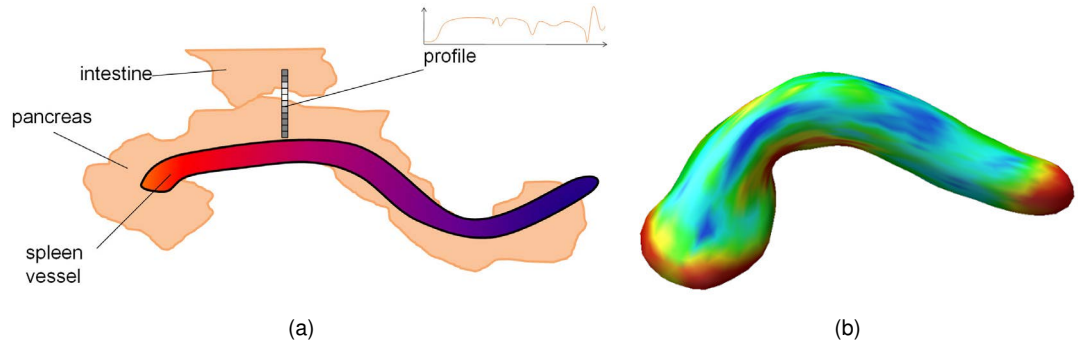


Figure 4.11.: (a) Sampling of intensity profiles along the splenic vessel in order to detect transitions between pancreas and other tissue. (b) Mean model of the pancreas with color coded local elasticity constraints.

4.4.1.5. Final detector

In a first detection stage, a boosted classifier is built using AdaBoost (cf. Section 2.4.3) with 200 classification and regression trees of maximum depth 5 as weak classifiers. The strong classifier is trained with $\vec{Z}(\vec{t})$ to learn the probability $p(l=1|\vec{Z}(\vec{t}))$. Figure 4.12 (upper right) shows a plot of this probability. Simple thresholding $p(l=1|\vec{Z}(\vec{t})) > \varepsilon$ leads to an incoherent labeling since no neighborhood relations between voxels are considered. In order to find a globally optimal classification belief propagation is used. A term Ξ is defined incorporating the probability of the classifier and global priors from the segmentation masks as well as a term Υ to incorporate voxel neighborhood relations. The labels $q \in \{0,1,2\}$ have to be found that satisfy

$$\underset{q}{\operatorname{argmin}} E(q) = \sum_{i}^{|I|} \Xi(\vec{I}_i, q_i) + \sum_{i}^{|I|} \sum_{j}^{N_i} \Upsilon(q_i, q_j), \tag{4.14}$$

with $|I|$ being the number of voxels in V and $\vec{I} \in \Re^{3|I|}$ being the set of voxel coordinates. N_i denotes the set of neighbors of the voxel at position $\vec{I}_i$. Ξ is defined as

$$\Xi(\vec{t}, q) = p(l=1|\vec{Z}(\vec{t})) + G(\vec{t}, q), \tag{4.15}$$

with

$$G(\vec{t}, q) \begin{cases} 1, & \text{if } q \in \{1,2\} \text{ and } \vec{t} \text{ in } M_v \\ 0, & \text{else.} \end{cases} \tag{4.16}$$

Υ is defined as $\Upsilon(q_1, q_2) = |q_1 - q_2|^2$ to penalize transitions between non-neighboring classes. A multi-scale approach [FH06] is used to solve (4.14) in linear time. The resulting label q_2 denotes the classification of the pancreas tissue (cf. Figure 4.12 for an example of the classification process on an unseen test case).

4.4.1.6. Shape space adaptation

In order to yield the final segmentation, knowledge about the expected pancreas shape is incorporated. A statistical shape model with $\Delta = 1692$ landmarks is built according to Section 2.5.3. The smallest dimension ρ is chosen such that $\sum_{i=1}^{\rho} \lambda_i$ captures 98% of the variance of the training data set, where $\lambda_1 \geq \ldots \geq \lambda_{3\Delta}$ are the eigenvalues of the corresponding covariance matrix.

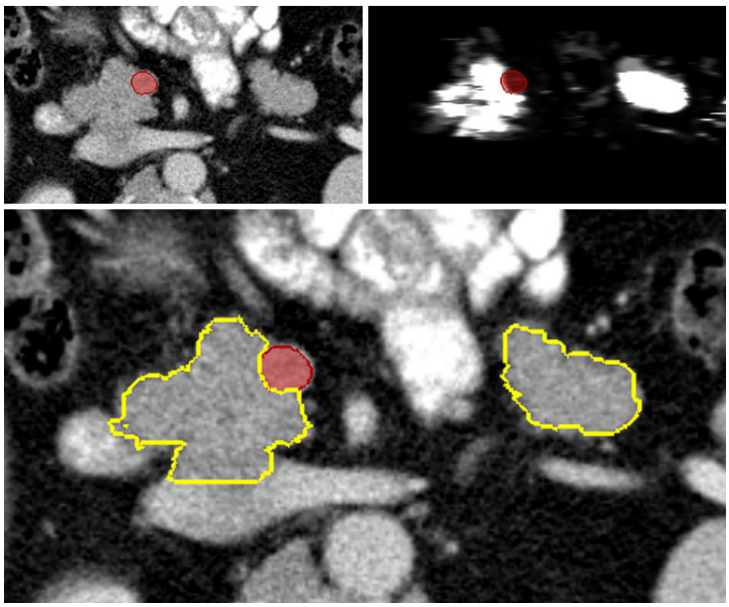


Figure 4.12.: Pancreas with segmented vessel (upper left). Classifier output (upper right) and final detector (bottom: bright outline) on an unseen test case [EKD*11].

Local shape constraints are added to the model according to Section 3.3.2. Figure 4.11(b) shows the mean of the statistical shape model with color coded shape constraints.

Label q_2 is transformed into a mesh using the Marching Cubes algorithm [LC87]. The mesh consists of a point set $\mathcal{M} = \{\mathbf{m}_1, \ldots, \mathbf{m}_K\}$. In order to register the statistical shape model with $\mathcal{M}$, an extended version of the Iterative Closest Point (ICP) algorithm is used. The classical ICP algorithm [BM92] rigidly registers a point set $\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_M\}$ with $\mathcal{M}$. In each iteration i of the classical approach, a correspondence function $C_i : \mathcal{P} \to \mathcal{M}$ between points from $\mathcal{P}$ and $\mathcal{M}$ is estimated, which is defined by the formula $C_i(\mathbf{p}) = \text{argmin}_{\mathbf{m}\in\mathcal{M}} \|(\mathbf{R}_i\mathbf{p} + \mathbf{t}_i) - \mathbf{m}\|$. Here, $\mathbf{R}_i$ and $\mathbf{t}_i$ denote the estimates of the rotation matrix and translation vector in iteration i. On the basis of C_i, the updated estimates $\mathbf{R}_{i+1}$ and $\mathbf{t}_{i+1}$ can be derived using the method of Horn [Hor87].

In the approach of Kirschner and Wesarg [KW11] used in this work, an extension of the ICP algorithm is able to register a complete statistical shape model with $\mathcal{M}$, instead of a static point set. It starts with initializing a vector $\mathbf{x} \in \mathbb{R}^{3p}$ with the mean shape $\bar{\mathbf{x}}$ of the statistical shape model. This vector $\mathbf{x}$ acts as the point set $\mathcal{P}$ in the original approach. In each iteration i, after the new pose parameters $\mathbf{R}_{i+1}$ and $\mathbf{t}_{i+1}$ have been estimated, $\mathbf{x}$ is deformed using the function C_i. The pose parameters are used to project the deformed vector $\mathbf{x}$ into the shape model space, where $\mathbf{x}$ is constrained to a plausible shape. This plausible shape is then reprojected to world space coordinates, and the next iteration begins. In contrast to the classical ICP algorithm, also a scale factor using Horns method [Hor87] is estimated. By using a k-d-tree to determine C_i in each iteration, the algorithm is very efficient.

After registration of the statistical shape model with $\mathcal{M}$, a constrained free-form deformation is applied as described in Section 4.2 to yield the final segmentation of the pancreas.

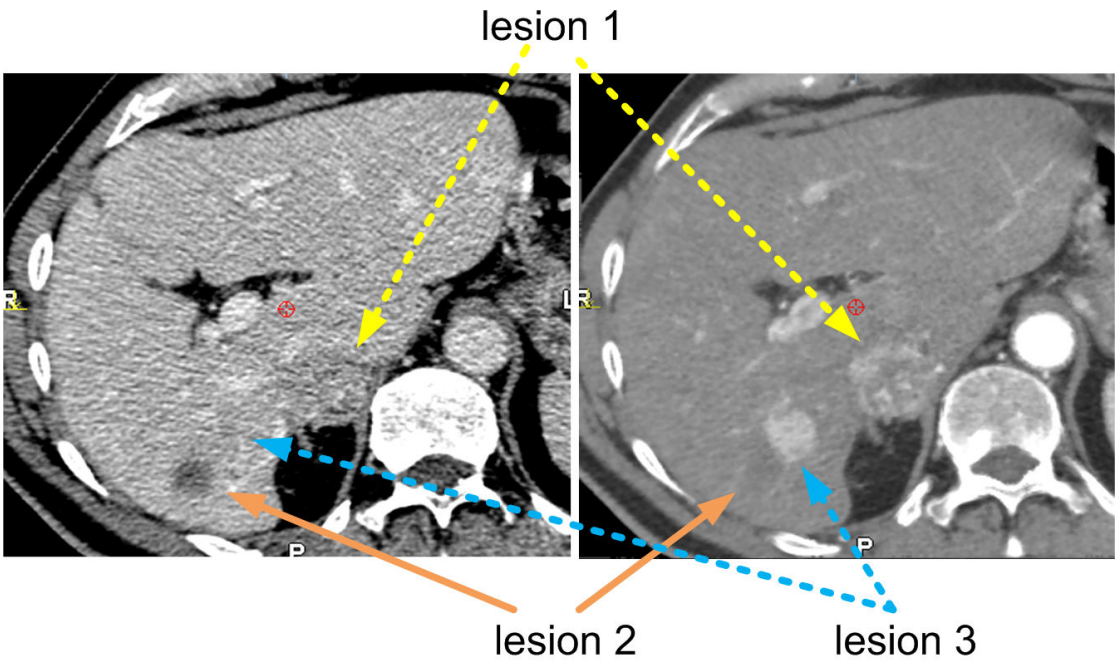


Figure 4.13.: Liver CT scan of the portal venous phase (left) and the arterial phase (right). 3 different lesions are shown in both images. Lesion 1 and lesion 3 are visible only in the arterial phase while lesion 2 is only visible in the portal venous phase.

4.5. Automatic deformable registration of the liver in CT data

Multi-phase liver registration has great potential improving computer aided diagnosis or operation planning since it allows for the fusion of complementary information from routinely gathered CT scans. Usually physicians have to compare the different phases slice-by-slice and map them mentally. This is especially problematic if precise knowledge about the location of structures is needed, e.g. the distance of a vessel to a tumor which are not visible in the same phase. Figure 4.13 shows an example of a liver CT scan in the portal venous phase (left) and the arterial phase (right). A total of three lesions is visible in the scans. Of these lesions, two are visible in the arterial phase and one is visible in the portal venous phase only. Furthermore, lesion 2 and lesion 3 are very close to each other. Lesion 1 is very close to a vessel which can only be seen in the portal venous phase. This example makes clear that a precise registration of both phases is necessary in order to plan an operative treatment of the lesions.

Most multi-phase registration methods rely on the liver's internal structure in the corresponding phases in order to register the volumes. Kwon et al. [KYLL08] use a B-spline elastic registration based on 3D features that are extracted and matched in different phases for 5 patients. However, since vessels like the liver arteries and the portal vein are very close to each other and are usually only visible in one phase it is likely that features will be selected that incorrectly match those different trees. Xiaoyang et al. [HWC*08] use a method based on a joint histogram to rigidly register the intra-hepatic vessels on 2 patients. However, since different respiration states will introduce local deformations rigid registration is not sufficient for a high accuracy matching. Heldmann and Zidowitz [HZ09] extract the vessel trees of one patient to define a penalty term in order to avoid overlapping of complementary vessels. However, segmenting a specific vessel

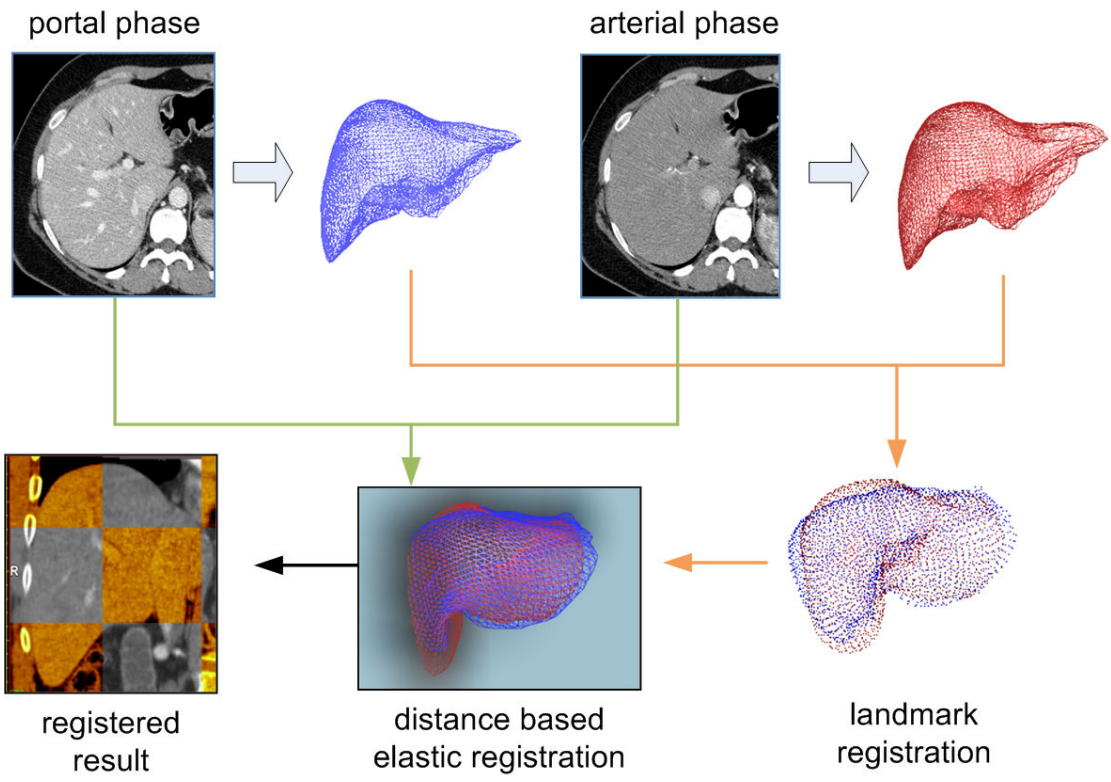


Figure 4.14.: Workflow of the automatic registration framework. The liver is extracted in the portal phase. The resulting mesh position is used as initialization for the extraction of the arterial phase. Afterwards shapes are registered based on their corresponding landmarks. The final registration is computed using a voxel based deformable registration that incorporates the distance of every voxel to the extracted shapes.

tree in one phase is very difficult since in most cases the contrast agent also saturates part of the other trees. Additionally, internal structures like tumors or ablations often prevent an accurate extraction of the vessel trees which may lead to false registration afterwards.

In this work a novel approach for automatic matching of multi-phase CT images is presented. First, the liver is automatically segmented in all phases. Subsequently, landmark based registration is performed to coarsely align the livers. The respiration forces are simulated by a Demon algorithm using a speed function that linearly penalizes the distance to the shape boundaries. Using the method, different phases can be registered fully automatically with high accuracy and without having to rely on error-prone matching of internal structures.

4.5.1. Workflow

Figure 4.14 outlines the registration workflow. First, the liver is extracted in the portal phase. The resulting mesh position is used as initialization for the extraction of the arterial phase. Afterwards shapes are registered based on their corresponding landmarks. The final registration is computed by a distance based deformable registration.

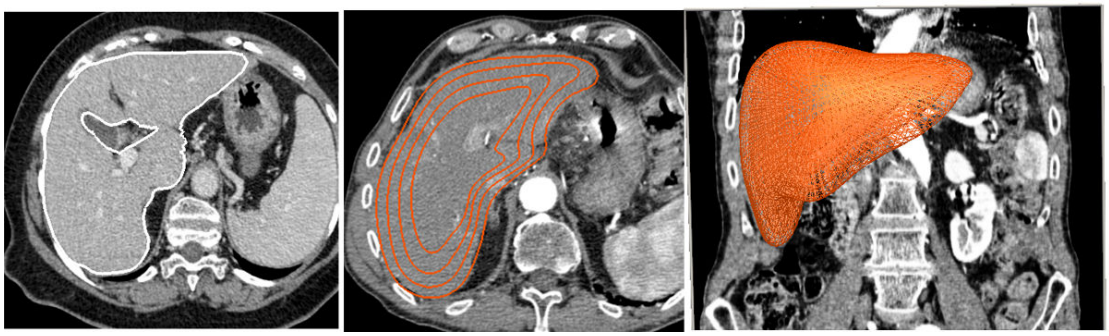


Figure 4.15.: Segmentation results of different liver phases [ESH*11]. The portal venous phase is segmented using a constrained statistical shape surface model (left). For the arterial phase, a multi-layered model is applied (middle, right).

4.5.2. Automatic liver segmentation

Liver segmentation of the portal venous phase is done using the approach described in Section 4.2. Figure 4.15 (left) shows an exemplary result of this segmentation step.

For segmenting the arterial phase, the method described in Section 4.3 has been applied (cf. Figure 4.15 (right)). Additionally, the model is initialized at the position of the portal phase segmentation result which helps to improve accuracy, since the liver boundaries are usually only shifted by some centimeters between phases. Figure 4.15 (middle) shows an exemplary result of the arterial phase segmentation.

4.5.3. Image Registration

Since the landmarks p^a and p^p of the extracted arterial and portal shape correspond to each other they can be efficiently registered minimizing $\sum_{i\in P}\left(T_A(p_i^a)-p_i^p\right)^2$, with P being the set of landmark indices and T_A being an affine transformation. The resulting transform is then applied to the arterial phase image. Both images are now coarsely aligned. In order to treat local deformations caused by respiration, elastic registration is performed based on the algorithm proposed by Thirion [Thi98]. However, its assumption that corresponding points to be matched share similar intensities does not hold for the liver's interior between phases. The original approach is modified as follows: a point's speed is penalized according to its position in fixed and moving image relative to the liver boundary. The displacement field at the N-th iteration is computed as

$$D^N(X) = D^{N-1}(X) - \frac{\Psi^{N-1}(X)(m(X+D^{N-1}(X)) - f(X))\nabla f(x)}{||\nabla f||^2 + (m(X+D^{N-1}(X)) - f(X))^2/K}, \tag{4.17}$$

with

$$\Psi^{N-1}(X) = \begin{cases} 1, & \text{if } \Phi_f^{N-1}(X) \text{ and } \Phi_m^{N-1}(X) > 0 \\ 0, & \text{if } \Phi_f^{N-1}(X) \text{ and } \Phi_m^{N-1}(X) < 0 \\ \frac{||\Phi_f^{N-1}(X)| - |\Phi_m^{N-1}(X)||}{|\Phi_f^{N-1}(X)| + |\Phi_m^{N-1}(X)|}, & \text{otherwise.} \end{cases} \tag{4.18}$$

$\Phi_f(X)$ and $\Phi_m(X)$ being the distance maps from fixed and moving image, f and m, respectively. K is the mean squared value of pixel spacings and ∇f is the fixed image's gradient. The dis-

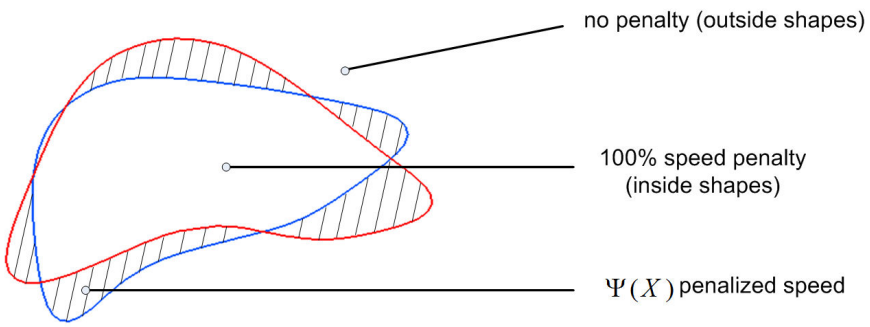


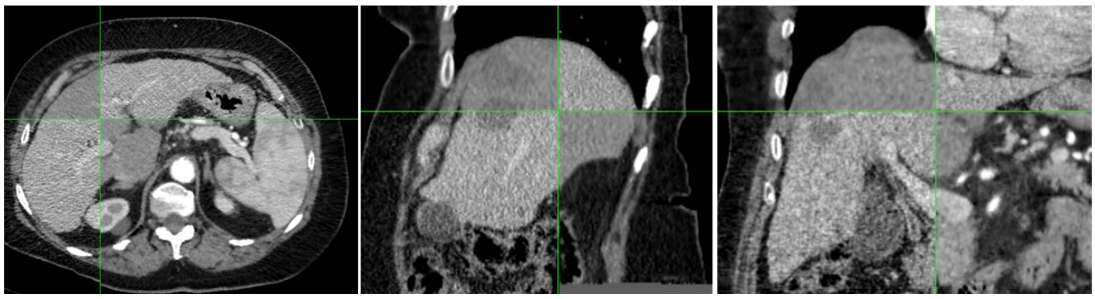
Figure 4.16.: Deformation penalty in different regions of overlapping liver shapes. Outside of both shapes the deformation speed function is not penalized. Inside both shapes no matching is allowed. In areas that are inside one shape only, the deformation speed function is penalized by $\Phi(X)$.

placement field is smoothed between iterations by a Gaussian using a standard deviation $\sigma = 3$ in order to propagate boundary forces to the liver's interior.

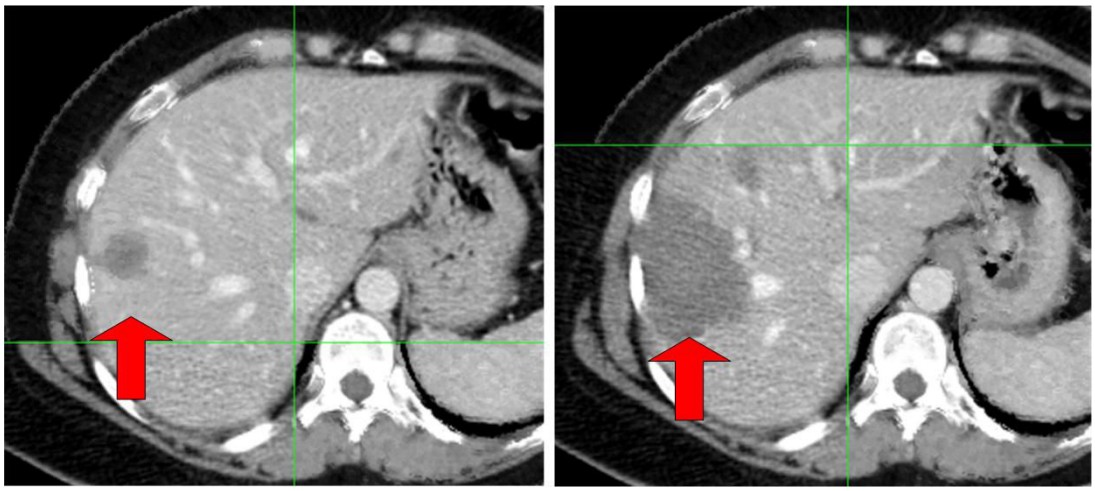
Figure 4.16 illustrates the deformation speed penalty (4.18) during one iteration. Both arterial and portal venous liver shapes are partly overlapping. Outside both shapes, the deformation speed is not penalized, i.e. a matching of neighboring structures is allowed. Inside both shapes, no deformation speed is allowed, in order to avoid the matching of unrelated structures. That means, the deformation field will not contain any displacements in this area. In the regions that are inside one shape only, the deformation speed of a voxel is penalized according to its distance to both shapes. This way, the deformation speed function is stronger penalized at positions that are far away from both boundaries. However, voxels that are close to both boundaries or directly on a boundary are less penalized, since boundaries are visible in both scans and should be matched. After computation of the deformation field for all areas, the deformation field is smoothed as described above. By repeating the above process of deformation force computations and smoothing, the deformation forces are propagated from the boundaries to the interior of the shapes. After a number of iterations, the deformation field only varies minimally and the registration is stopped. In the performed tests, usually an iteration count of 100 was sufficient.

The described elastic registration allows for a high accuracy matching of liver boundaries while internal structures are matched by the interpolation of the displacement field. This allows for a natural deformation of the organ's interior since respiration forces are also propagated approximately linearly to the liver's interior. An exemplary registration result of arterial phase and portal venous phase can be seen in Figure 4.17(a).

In addition to the multi-phase registration described in the previous Section, the registration can also be applied to register pre-treatment and post-treatment CT scans of the portal venous phase. Figure 4.17(b) shows an example of the registration of two portal venous CT images. In the pre-treatment image (left), a tumor is visible. This tumor has been ablated during treatment. In a follow-up CT scan, the result of the treatment is examined. Using the described registration framework, the physician can measure the distance of the original tumor area to the boundaries of the ablated area in order to assure that enough tissue has been burned to avoid a relapse.



(a)



(b)

Figure 4.17.: Split views of registration results: (a) Multi-phase registration of arterial and portal contrast enhancement phase. (b) Pre-treatment (left) and post-treatment (right) registration. Arrows indicate initial lesion in pre-treatment scan (left) and ablated area in post-treatment scan (right) [EOLD*12].

4.6. Discussion

In this chapter, the model based methods developed in this thesis have been applied to common medical image segmentation and registration problems. Segmentation frameworks for the kidney, the liver and the pancreas in contrast enhanced CT data have been proposed. A multi-tiered adaptation method has been presented with a stepwise increase of the model's degrees of freedom in order to allow for a robust segmentation in the case of model initialization errors. The local shape constraints proposed in the last chapter have been combined with the model's curvature during model adaptation. This way, deformation is limited to model areas where high curvature is excepted from the training stage and observed during current adaptation.

In the case of pancreas segmentation, new types of texture features have been proposed that classify pancreas tissue around the splenic vein. The features are integrated into a pancreas tissue classifier. Combined with the proposed model based approaches, an automatic segmentation of the pancreas from single phase contrast enhanced CT has been presented.

In order to allow the segmentation of low contrast structures, the proposed volumetric model has been used to segment the bladder in low contrast CT and the liver in the arterial contrast CT phase. The segmentation of the liver in the arterial contrast phase led to the proposal of a

fully automatic scheme for registration of multi-phase liver CT images. In this scheme, the liver is segmented from portal venous and arterial contrast CT phases. Afterwards, a voxel based deformable registration is proposed that incorporates the shape information of the liver during adaptation. This way, a matching of complementary structures as visible inside the liver during the different phases is avoided.

The next chapter presents an evaluation of the methods proposed in this thesis based on the segmentation and registration problems addressed in this chapter.

Parts of this chapter are based on publications of the author [ES10, WKEK10, WEKK11b, WEKK11a, ETS09, ES10, ESKW10, ESW10, ESH*11, EOLD*12, OLDE*12].

5. Evaluation

In this chapter, the methods developed in Chapter 3 and Chapter 4 are evaluated and put into context with existing methods. First, the quality measures used to assess the developed methods are described in Section 5.1. The smart manual landmarking method described in Section 3.3.3 is evaluated in Section 5.2. The model adaptation process is evaluated in Section 5.3. It includes evaluation of model initialization, evaluation of the robustness of the boundary detection as well as a comparison between linear and nonlinear shape models. The kidney segmentation framework presented in Section 4.1 is evaluated in Section 5.4. The multi-tiered liver segmentation method proposed in Section 4.2 is evaluated in Section 5.5. It follows the evaluation of the segmentation accuracy of bladder, arterial phase liver and pancreas in Section 5.6 and Section 5.7. Section 5.8 concludes the chapter with an evaluation of the registration of multi-phase liver CT images.

5.1. Quality measures

In order to evaluate the methods developed in this thesis, several well known quality measures are used. In the context of segmentation, the goal is to maximize the true positive area (TP) between the computed segmentation and a reference labeling as Figure 5.1 shows.

The used measures are:

- the volumetric overlap error (abbreviated as VOE in the remainder) in percent. The VOE is defined as $(1 - \frac{TP}{FN+TP+FP})100$. This value becomes 0 in case of a perfect overlap between computed segmentation and ground truth labeling and 100 in case no overlap exists.
- the volume difference (abbreviated as VD in the remainder) in percent. The VD is defined as $100((TP+FP)-(FN+TP))/(FN+TP)$. A value of 0 indicates that segmentation and reference have the exact same volume. Negative values denote under-segmentations and positive values denote over-segmentations. Since this measure only computes the difference between the total volume of a segmentation to the reference, a value of 0 does not necessarily characterize a good segmentation. However, in conjunction with other measures, it indicates whether an algorithm tends to oversegment or undersegment structures.
- the average symmetric surface distance (abbreviated as ASD in the remainder) in millimeters. The ASD is computed determining the border voxels of segmentation and reference. A border voxel is a voxel that has at least one neighbor (18 neighborhood) which is not labeled. For each voxel in one border set, the voxel with the minimum Euclidean distance in the other border set is determined. All minimum distances for both border sets are averaged to compute the ASD. The ASD is 0 for a perfect overlap between segmentation and reference.
- the root mean square symmetric surface distance (abbreviated as RMS in the remainder) in millimeters. The RMS is similar to the ASD, but computes the squared Euclidean distances between border voxel sets. The squared distances are averaged. The square root of the average is the RMS. The RMS is 0 for a perfect overlap of segmentation and reference.

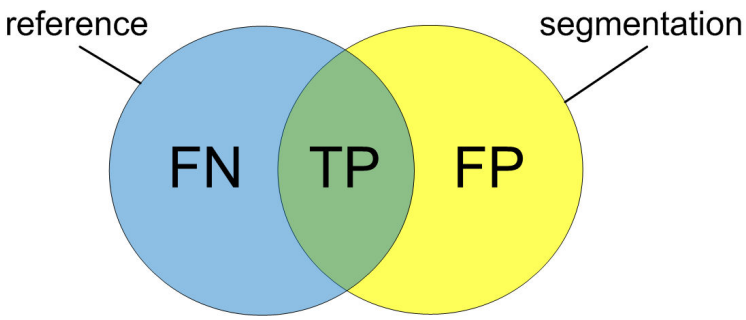


Figure 5.1.: Comparison of a computed segmentation (yellow circle) with a reference (blue circle). False negative (FN), true positive (TP) and false positive (FP) areas are shown.

- the maximum symmetric surface distance (abbreviated as MD in the remainder) in millimeters. The MD is defined as

$$MD(A,B) = max(h(A,B),h(B,A))$$

with $h(A,B) = max_{a\in A} min_{b\in B}||a-b||$ and with $A = TP+FP$ and $B = FN+TP$ are sets of voxels. $||a-b||$ denotes the Euclidean distance in millimeters and is computed in physical coordinates in order to account for non isotropic resolutions. The MD is 0 for a perfect overlap of segmentation and reference.

- the dice similarity coefficient (abbreviated as DC in the remainder). The DC is defined as

$$DC(A,B) = 2|A\cup B|/(|A|+|B|)$$

with $A = TP+FP$ and $B = FN+TP$ being sets of voxels. This measure becomes 0 in case no overlap between computed segmentation and reference exists and 1 in case of a perfect match.

- the positive predicive value (abbreviated as PPV in the remainder). The PPV is defined as $\text{PPV} = \text{TP}/(\text{TP}+\text{FP})$. The PPV becomes 0 in case no overlap between computed segmentation and reference exists and 1 in case of a perfect match.

5.2. Smart manual landmarking

In this section, the method presented in Section 3.3.3 for building a statistical shape model using smart manual landmarking is evaluated. First, the quality of the correspondence establishment is demonstrated on several anatomical structures using a visual correspondence scheme. It follows a quantitative evaluation using standard statistical shape model quality measures.

5.2.1. Correspondence visualization

The approach has been evaluated by building statistical shape models for three different anatomical structures of varying topology (left kidney, vertebra and cardiac left ventricle). For the vertebra and kidney, static volume data from CT was used (10 and 16 data sets, respectively) while the cardiac left ventricle is represented by 20 dynamic images provided as 3D cine MRI data reconstructed along the ventricle's main axis. An average user needs around 5 to 10 minutes for

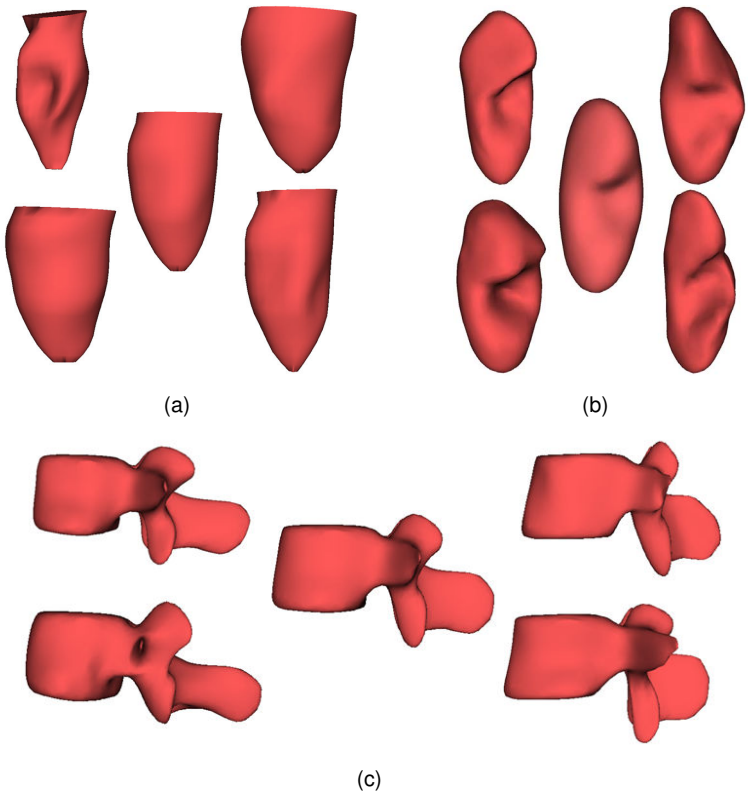


Figure 5.2.: Principal modes of variation for the statistical shape model of the cardiac left ventricle (a), left kidney (b) and vertebra (c). The variation of the two largest eigenmodes between $-3\sqrt{\lambda_i}$ (left) and $3\sqrt{\lambda_i}$ (right) are shown together with the mean mesh (middle) [EKW09].

segmenting a kidney and a vertebra sufficiently precise and around 7 minutes for a ventricle. These times are comparable to other manual segmentation systems. Therefore, the proposed segmentation step is suitable to be used in practice.

The deformed meshes were directly used for the shape model generation as described in Section 3.3.3.5. Figure 5.2 shows the principal modes of variation for all built statistical shape models. For each model, the mean (middle) is shown as well as the variation of the two largest eigenmodes between $-3\sqrt{\lambda_i}$ (left) and $3\sqrt{\lambda_i}$ (right). As can be seen in Figure 5.2(a), the first mode maps the deformation of the ventricle as it is visible in the cine MRI data. A comparison between Figure 5.2(c) and Figure 5.3 (showing the set of training shapes of the vertebra) shows that the variation of the training shapes is well mapped to the first two modes of the generated shape model.

In order to support visual evaluation of correspondence, a color coded visualization technique is used[1]. Coarse correspondence of regions is obtained by a color transfer function which maps landmark indices to hue in HSV space. Value is fixed to 1.0, and saturation to 0.4, which avoids loud colors. The technique requires that adjacent landmarks have similar indices. If this is not the case, a renumbering of the landmark indices is necessary. In the implementation, the index 0 is assigned to an arbitrarily selected landmark, and then a breadth first search is started from this

[1] The correspondence visualization has been developed by Matthias Kirschner (TU-Darmstadt).

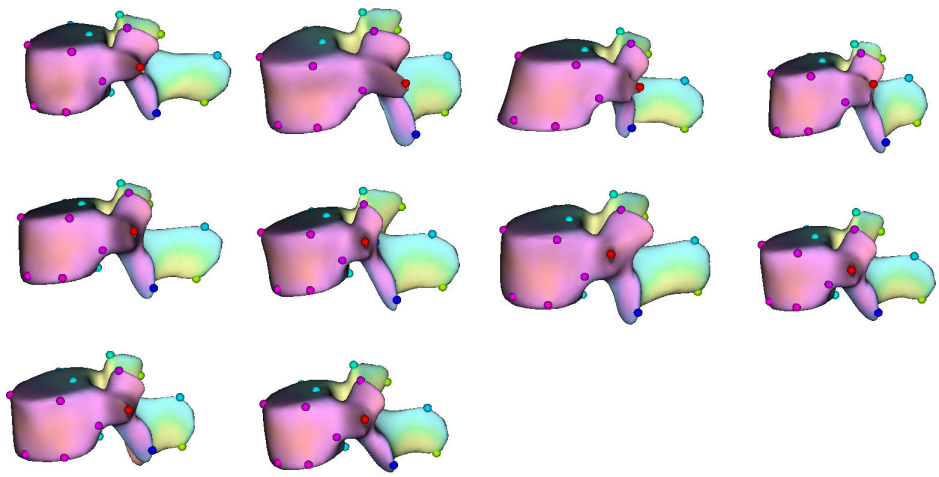


Figure 5.3.: Color coded visualization of point-to-point correspondence of the vertebra training set [EKW09]. Points evenly placed on features of locally high curvature are shown as glyphs.

landmark on the graph induced by the connectivity structure of the model. New landmark indices are determined by the order in which they are discovered during the breadth first search.

Visualization of correspondence on a fine-grained level of detail is achieved by means of glyphs. Since visualizing all landmarks would obstruct perception, a subset S is determined automatically. The approach aims at finding landmarks corresponding to predominant features, which at the same time cover approximately uniformly the whole surface of the training meshes. The landmarks are sorted by decreasing Gaussian curvature on a single, arbitrarily selected training mesh M. It follows an iteration over the sorted landmarks. A landmark is added to S if its Euclidean distance on M to the landmark that was previously added to S is larger than a specified threshold. All landmarks in S are visualized by spherical glyphs. In order to assign color to the glyphs, a similar transfer function as described above is used, where a saturation of 1.0 is set to obtain a better accentuation.

Figure 5.3 shows the correspondence visualization for the training sets of the vertebra. On the vertebra, prominent feature points are well visible and therefore mismatching points can be identified easily using the visualization approach. As can be seen, correspondence in terms of mapping areas from one shape to the other is established well. In addition, the feature points mapping local high curvature are also placed on corresponding positions over the set of training shapes. In case of the kidney, the correspondence cannot fully reach the quality of the vertebra models. This is due to the fact that the kidney does not have many uniquely identifiable points compared to the vertebra. The correspondence of the cardiac left ventricle models is comparable to the vertebra. Over all cine images, the position of the ventricle does not change. This eases the adaptation process and removes the error that may be introduced to a wrong initial orientation of the model.

5.2.2. Comparison with automatic correspondence optimization

The smart landmarking approach is compared against two fully automatic algorithms: The first algorithm, which is outlined in Section 5.2.2.1, is based on the *iterative closest point* (ICP) algo-

rithm [BM92]. ICP-based algorithms are frequently used in the literature and reported to produce sufficient results in many applications. In Section 5.2.2.2, a second algorithm is described which uses population-based optimization.

Both algorithms receive a set of training meshes as input. These input meshes were reconstructed from binary volumes – segmentations of the test data sets – using the marching cubes algorithm [LC87].

5.2.2.1. ICP-based correspondence

ICP-based correspondence algorithms establish correspondence by aligning the training meshes using the ICP point set registration algorithm [BM92]. A well-known shortcoming of ICP-based approaches is the use of the Euclidean distance as a correspondence measure, because pairs of closest points in registered shapes do not necessarily correspond. This shortcoming is alleviated to some extent by extracting the landmarks from consistent spherical parameterizations, which were computed using parameter space propagation [KW10]. While this approach also builds upon the ICP algorithm, it improves the quality of the correspondence through anisotropic scaling and fuzzy instead of fixed point correspondences. Furthermore, the parameter space representation of the meshes allows for easy interpolation of landmarks from the whole surface of the mesh, which means that the resulting landmark positions are not restricted to the mesh points of the input meshes.

The parameter space propagation method [KW10] works as follows: A reference mesh M_{ref} is selected from the training set, and an area-preserving spherical parameterization of M_{ref} is computed. All other meshes are anisotropically scaled such that the variances on their principal axes is identical to the variances on the principal axes of the reference meshes. The scaled meshes are aligned with the reference mesh using the ICP algorithm in order to derive a common coordinate system. Then a fuzzy correspondence between points of each mesh M and points of M_{ref} is established, where the degree of correspondence between two points is determined by weights which are dependent on the Euclidean distance of the points in the common coordinate system. Using this fuzzy correspondence, parameter space coordinates for M are computed by interpolating parameter space coordinates of corresponding points of M_{ref}. A subsequent correction step handles overlapping triangles on the spherical parameterization of M, such that the mapping from parameter space to the world space becomes bijective. Note that the anisotropic scaling is only used in order to derive the correspondence relation and is undone once the parameterization is propagated.

Corresponding landmarks are then extracted from the spherical parameterization using the simple icosahedron sampling technique: The faces of a unit icosahedron are subdivided, and all points of the subdivided icosahedron are then normalized to obtain sampling points on the unit sphere. Using the parameterization of a mesh M, sampling points can be mapped to world space in order to obtain the positions of the landmarks for the respective training example M.

5.2.2.2. Optimization-based correspondence

The optimization-based algorithm treats the problem of establishing correspondence as a problem of reparameterization, as proposed by Kotcheff and Taylor [KT98]. For a detailed description of the general approach, the reader is reffered to the book by Davies et al. [DTT08] and the survey of Heimann et al. [HM09]. The *DetCov* function [KT98] is used as target function of the

optimization algorithm, which explicitly favors compact models. This function performed well in the study of Styner et al. [SRN*03].

The procedure is started with computing consistent spherical parameterizations using parameter space propagation [KW10]. This means that the ICP-based correspondence solution from Section 5.2.2.1 is the starting solution of the optimization scheme. However, in contrast to the former approach, the positions of the sampling points which are used for remeshing are iteratively optimized. As proposed by Davies et al. [DTT08], only the sampling points of a single shape in one iteration are manipulated using clamped plate spline warps reparameterizations [DTT08]. For each data set, a fixed number of 3000 iterations is used after which the objective function did not change significantly.

5.2.2.3. Shape model evaluation

For a quantitative evaluation of smart manual landmarking, two statistical shape models of different anatomical structures have been built. The first model is a statistical shape model of the left kidney using 16 static CT datasets from a Siemens Somatom Sensation scanner as the basis. The in-plane spacing of the CT scans is 0.74 mm. The slices were reconstructed with a thickness of 5 mm. The second statistical shape model has been built from 10 dynamic images of the cardiac left ventricle's heart cycle. The scans are provided as 3D cine MRI data reconstructed along the ventricle's main axis and consist of 4 mm slices with an in-plane spacing of 1.36 mm.

The left kidney mesh consists of 1002 points while the cardiac ventricle mesh consists of 480 points. Generating the training set using ICP-based correspondence establishment takes less than a minute on a desktop computer with 2.4 GHz processor and 3 GB Ram. The population-based optimization requires 45 minutes for establishing correspondence on the ventricle data set and 95 minutes on the kidney data set using the same machine. Both automatic correspondence algorithms produce meshes of 1002 points (kidney) and 492 points (cardiac ventricle) in order to allow a fair comparison with the manual method.

In order to evaluate the quality of the constructed shape models, the common statistical shape model evaluation measures *specificity* S and *generalization* G are computed [DTT08]. They are defined as

$$S = \frac{1}{n_s}\sum_{A=1}^{n_s} \max_i(\Psi(A,i)) \quad \text{and} \quad G = \frac{1}{M}\sum_{i=1}^{M} \max_A(\Psi(A,i)). \tag{5.1}$$

$\hat{Y} = \{y_A : A = 1,...n_s\}$ is a set of shapes sampled from the model's probability density function and $\hat{X} = \{x_i : i = 1,...M\}$ is the set of training shapes. $\Psi(A,i)$ denotes a function to compare shape y_A with x_i. As Heimann et al. [HWM06] pointed out, it is more meaningful to calculate the similarity based on the resulting *binary segmentations* instead of using the *landmark positions*, because in image analysis one is most often interested in the volume encompassed by the model and not in the mesh itself. Therefore, the volumetric overlap $\Psi(A,i) = |Y_A \cap X_i|/|Y_A \cup X_i|$ is used as a similarity measure with X_A and Y_i being the sets of voxels enclosed by y_A and x_i. In this case $S = 1$ and $G = 1$ with $S \in [0,1]$ and $G \in [0,1]$ denote a perfect specificity and generalization respectively.

In order to have a comparable data basis, the training meshes generated with the smart landmarking method are converted to binary volumes. The binary masks are then taken as input to the automatic landmarking algorithms.

For the 6 largest modes of variation, a set of sample shapes $\hat{Y}$ is generated. Every set consist of 500 random samples. Using these sets, the generalization ability as well as the specificity

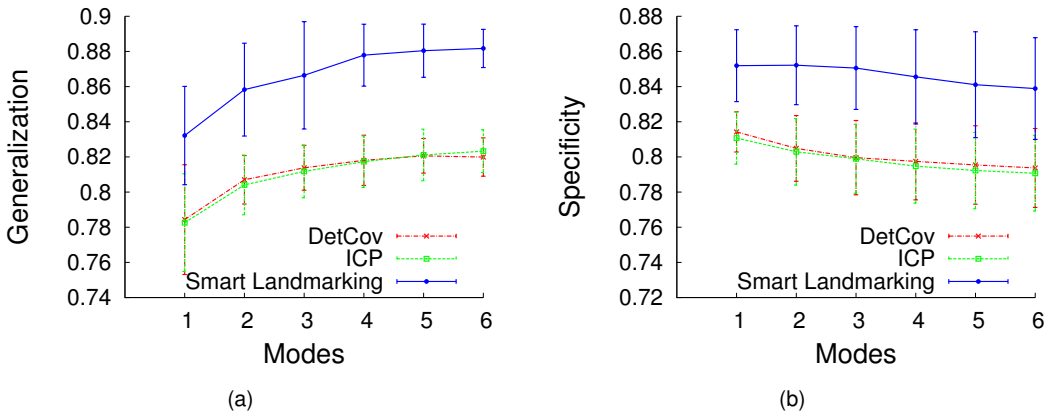


Figure 5.4.: (a) Generalization and (b) specificity measures for the built statistical shape model of the kidney [EKW10]. The metric used to compare the shapes is the volumetric overlap between the labeled voxel segmentations. A value of 1.0 denotes a perfect overlap.

of the generated statistical shape models is calculated. The analysis for each model takes 5.7 hours on a 2.4 GHz Quad Core Intel processor with Windows XP and 4GB of memory. Figure 5.4 shows the results of the evaluation for the built kidney model using smart landmarking in comparison to the ICP-based method and to the *DetCov* method while Figure 5.5 shows the same calculations for the cardiac left ventricle model. As can be seen, the smart landmarking approach performs better in terms of generalization (Figure 5.4(a)) and specificity (Figure 5.4(b)) in case of the statistical shape model of the kidney. The cardiac left ventricle model confirms the advantage of the smart landmarking for the generalization ability (Figure 5.5(a)) while it performs slightly worse in terms of specificity (Figure 5.5(b)). However, in the latter case the performance of all algorithms is very similar.

5.2.3. Discussion

The results show that smart landmarking is in most cases superior to the standard correspondence establishment methods using ICP-based correspondence and *DetCov* population-based optimization. An explanation for this outcome is that smart landmarking does not contain any preprocessing and directly work on the generated meshes. The other approaches rely on binary segmentations and therefore apply a smoothing step to remove staircase artifacts, which reduces the degree of match between the generated training sets and the binary segmentations.

It could be argued that other objective functions, for example based on *Minimum Description Length* [HWM06] would probably improve the results of the automatic correspondence algorithms. However, more sophisticated methods also need considerably more time — especially for large training sets. For example, the statistical shape models of the liver (cf. Section 4.2.1) used in this thesis were constructed from over 200 training meshes. Using smart landmarking, no additional work is required apart from the segmentation step. This is especially useful when an existing statistical shape model is improved by adding new training data. Here, new shapes can be directly included in the training base and the statistical shape model can be created. Since the process of creating the statistical shape model is very fast, this procedure can be repeated for every new training shape without a negative impact on the overall workflow. The developed tool

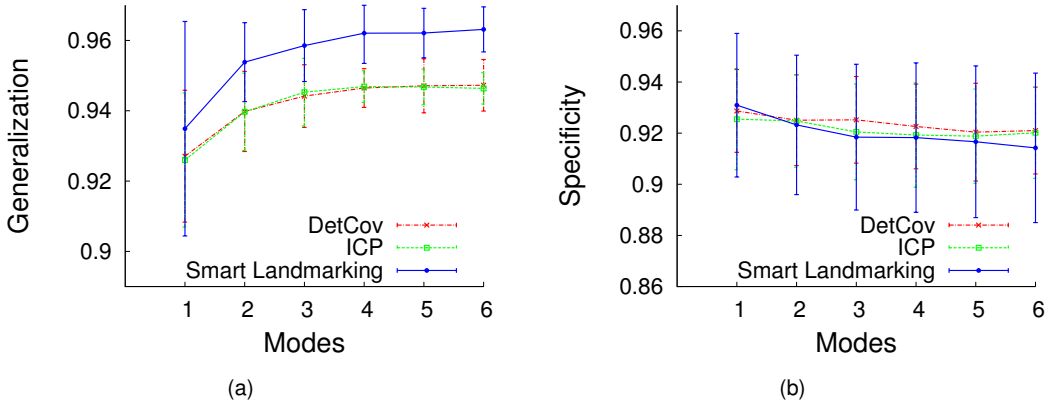


Figure 5.5.: (a) Generalization and (b) specificity measures for the built statistical shape model of the cardiac left ventricle [EKW10]. The metric used to compare the shapes is the volumetric overlap between the labeled voxel segmentations. A value of 1.0 denotes a perfect overlap.

can therefore be integrated into clinical practice when a physician needs to segment a structure manually. This way a very large amount of training meshes can be easily generated.

As the results point out, human guided landmarking is feasible in 3D and also suitable to produce statistical shape models of good quality. However, human interaction always introduce a potential error source. An unexperienced user may not be able to identify all corresponding features of a particular organ. Therefore, the quality of the training shapes may suffer. Furthermore, even the same user may generate different outcomes for the same organ.

5.2.4. Summary

A method for statistical shape model generation based on preserving point correspondences during user guided segmentation has been proposed. In order to perform a visual evaluation of the degree of correspondence between different data sets, a color coded visualization based on local curvature has been used. From tests on three different anatomical structures it is concluded that the resulting correspondences established during the manual mesh deformation can be directly used to build statistical shape models even for structures of varying topology.

A quantitative evaluation for the kidney and cardiac left ventricle confirmed the results of visual analysis. It has been shown that the proposed approach yields better results than standard ICP-based correspondence establishment both in terms of generalization capability and specificity and also performs well in comparison to a population-based optimization method. The process of shape model generation can therefore be reduced to a single step which eliminates potential errors introduced in the common landmarking and correspondence estimation steps and also means a considerably speed up of the whole procedure. Another main advantage of the method is its potential to handle arbitrary topologies. Most automatic approaches cannot deal with shapes of complex topologies like the surface of the vertebra. Since the presented method can use an arbitrary model as reference shape, topology changes are unproblematic.

Table 5.1.: Detection errors of selected anatomical structures averaged over 30 test images. Errors are measured between detected bounding box centers and ground truth bounding box centers.

Structure	average error [mm]	standard deviation [mm]
LEFT KIDNEY	25.9	12.0
RIGHT KIDNEY	34.6	11.3
TOP OF THE LIVER	55.7	16.2
BLADDER	32.5	14.9
HEART	42.5	11.1
LOWER LIVER LOBE	51.2	13.1
PORTAL VEIN / SPLENIC VEIN	34.6	12.8
SPLEEN	56.9	19.5

5.3. Model adaptation

In this section, the single model adaptation methods presented in the last chapters are evaluated. In Section 5.3.1, the automatic initialization of organ models is evaluated. Since it cannot be guaranteed that the model is always positioned at the exact location, the robustness of the model adaptation to model initialization errors is evaluated in Section 5.3.2 for the surface model as well as for the multi-layer model.

The ability of the presented outlier removal process to correct significant errors of the boundary detection is given in Section 5.3.3. The benefit of using nonlinear statistical shape models on the segmentation of the liver is demonstrated in Section 5.3.4.

5.3.1. Model initialization

The automatic model initialization described in Section 3.6 has been evaluated on 30 thoracic / abdominal CT scans with an axial spacing of 5 mm. Another 30 CT datasets have been used to train the organ classifiers and the relative organ center positions. Test and training set were disjunct. For evaluation, 8 structures have been trained (right kidney, left kidney, spleen, top of the liver, lower liver lobe, portal vein / splenic vein branch, heart and bladder. The detection time range from 1.7 seconds to 7.2 seconds depending on the structure. Larger objects like the heart have longer computation times, since more voxels are processed.

Figure 5.6 shows exemplary results of the detection of the heart, bladder and spleen (from top to bottom row). The crosses mark the detected organ centers. As can be seen, the selected structures are roughly enclosed by a bounding box. In Table 5.1, quantitative results are shown for the 30 test images. The average displacement error between detected bounding box center and ground truth bounding box center is between 25 mm for the left kidney and 56 mm for the spleen.

Given the high anatomical variance of the selected organs and structures, it can be expected that the error will further decrease if the classifiers are trained on more data. However, since the developed model based approaches only need a rough partial overlap to adapt to the target structure, the present accuracy is already sufficient.

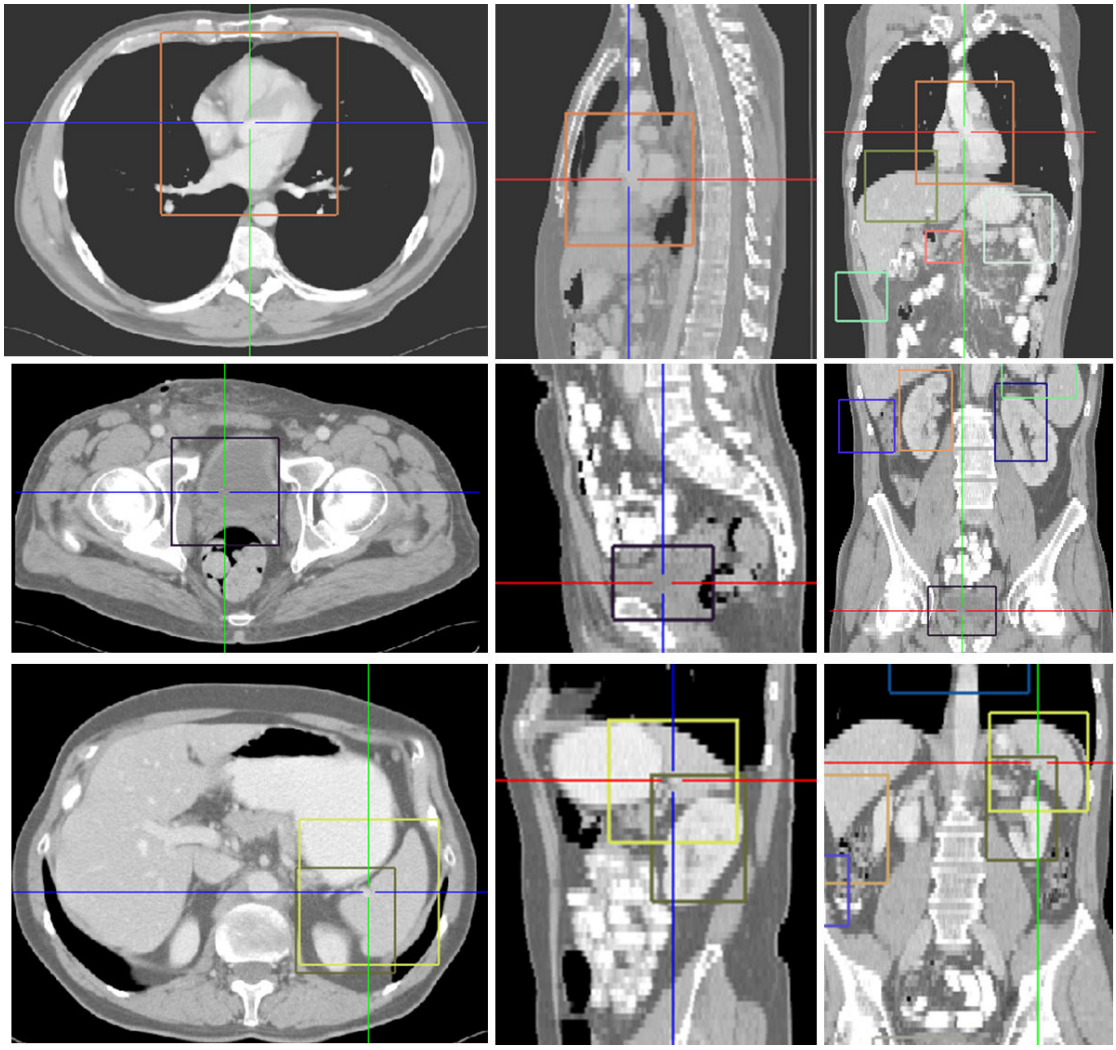


Figure 5.6.: Exemplary results of automatic organ initialization. The first row shows the detected bounding box of the heart. The second row shows the bounding box of the bladder and the last row shows the bounding box of the spleen (yellow box). The crosses denote the corresponding bounding box centers.

5.3.2. Robustness to model initialization errors

In this section, the proposed adaptation methods as described in Section 4.2.3 and Section 3.3.4 are evaluated regarding their robustness to initialization errors. Usually, automatic initialization approaches do not yield good results in every case. The model adaptation, therefore, must be robust enough to cope with a model misplacement of several centimeters.

5.3.2.1. Surface model

A benefit of the multi-tiered adaptation presented in Section 4.2.3 is its increased robustness to model initialization errors compared to a standard statistical shape model adaptation with full degrees of freedom. In order to evaluate this benefit, the segmentation quality using the proposed adaptation is compared to statistical shape model adaptation as follows: given a set of ground

Table 5.2.: Starting model deformation independently at the center of gravity of ground truth and at the automated calculated position results in the listed deviation of the final segmentation (cf. Section 5.1 for metrics). The first row shows the deviation for classical statistical shape model (SSM) adaptation, the second row shows the results for the multi-tiered approach. The results are averaged over 76 test scans [ESKW10].

Method	VOE Dev. [%]	VD Dev. [%]	ASD Dev. [mm]	RMS Dev. [mm]	MD Dev. [mm]
SSM	1.02	1.31	0.35	0.85	5.39
MULTI-TIERED	0.23	0.15	0.09	0.24	0.85

truth segmentations, the models are initialized at the center of gravity of the ground truth labelings. This positioning is assumed to be a ground truth initialization, since the model's surface is very close to the real boundaries of the structures to segment in the image [YS09]. The models are then adapted to the images using both adaptation approaches. Quality measures are computed after adaptation for each test image. Afterwards, the same test is performed again, but using the heuristic based automatic initialization method described in Section 4.2 for both approaches.

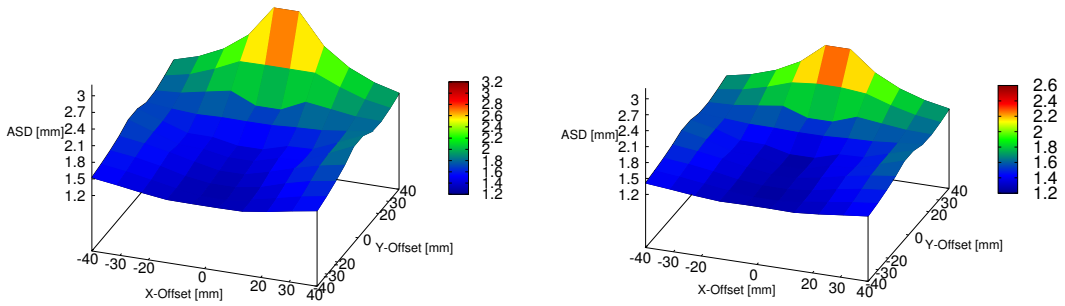
The positioning error of the heuristic initialization method compared to ground truth initialization has been computed based on 76 CT scans of the liver. The resulting average positioning error is 39.6 mm with a standard deviation of 14.6. That means, the heuristic initialization is around 4 cm away from the ground truth position which can be significant depending on the shape deviation of the model and the organ in the image. In order to evaluate how well both standard statistical shape model and the proposed multi-tiered adaptation can handle this error, the deviation of quality measures are computed between ground truth initialization and heuristic initialization. The results are listed in Table 5.2. The first row shows the deviation in segmentation quality for a classical statistical shape model using automatic and ground truth initialization. The average maximum distance error is 5 mm. The second row shows the same results for the multi-tiered approach. With an average maximum distance error below 1 mm and an average surface distance deviation below 0.1 mm, there is almost no segmentation quality difference between the ground truth initialization and the automatic initialization.

It can be concluded that the multi-tiered adaptation presented in Section 4.2.3 is more robust to model initialization errors than standard statistical shape model adaptation. Since current automatic initialization methods do not yield perfect initialization results, possible errors can be well handled by the proposed method.

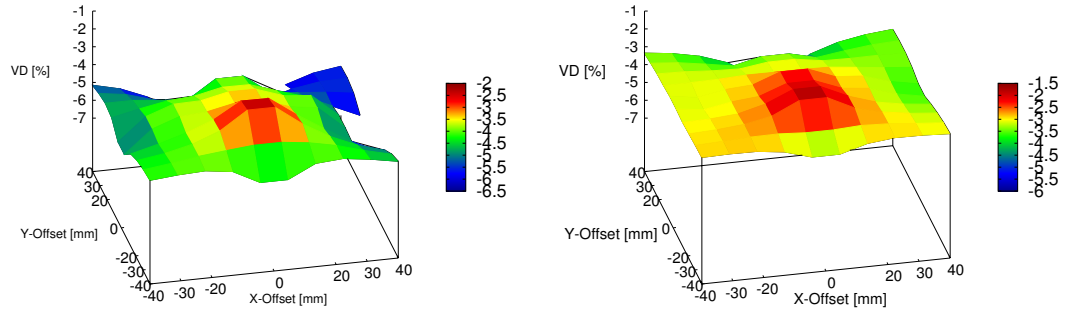
5.3.2.2. Multi-layer model

An advantage of the multi-layer model presented in Section 3.3.4 compared to standard surface based methods is their ability to detect organ foreign tissue in the model's interior during adaptation. That should increase their robustness to model initialization errors in difficult settings like low contrast datasets.

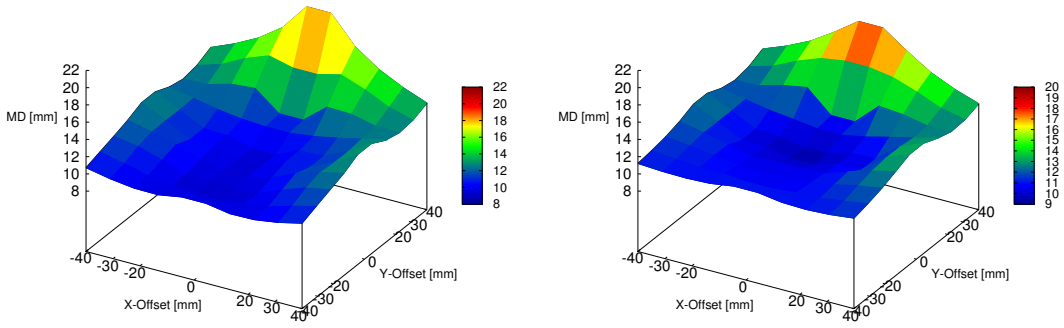
Therefore, for evaluation, the bladder as a structure with low contrast boundaries to neighboring tissue has been selected as test organ. 10 CT scans of the bladder with an axial spacing of 5 mm have been used. Segmentation accuracy is evaluated using VOE, PPV, ASD, RMS and MD (cf. Section 5.1 for quality measures). Smaller values mean better results except for PPV.



(a)



(b)



(c)

Figure 5.7.: Segmentation accuracy for different initial model positions (x-,y-offset of ± 40 mm). The ground truth initialization has an x-,y-offset of 0. The left column shows quality measures for a surface model while the right row shows the same measures for a multi-layer model (cf. Section 5.1 for a description of measures). (a) and (c): lower values are better. (b): higher values are better.

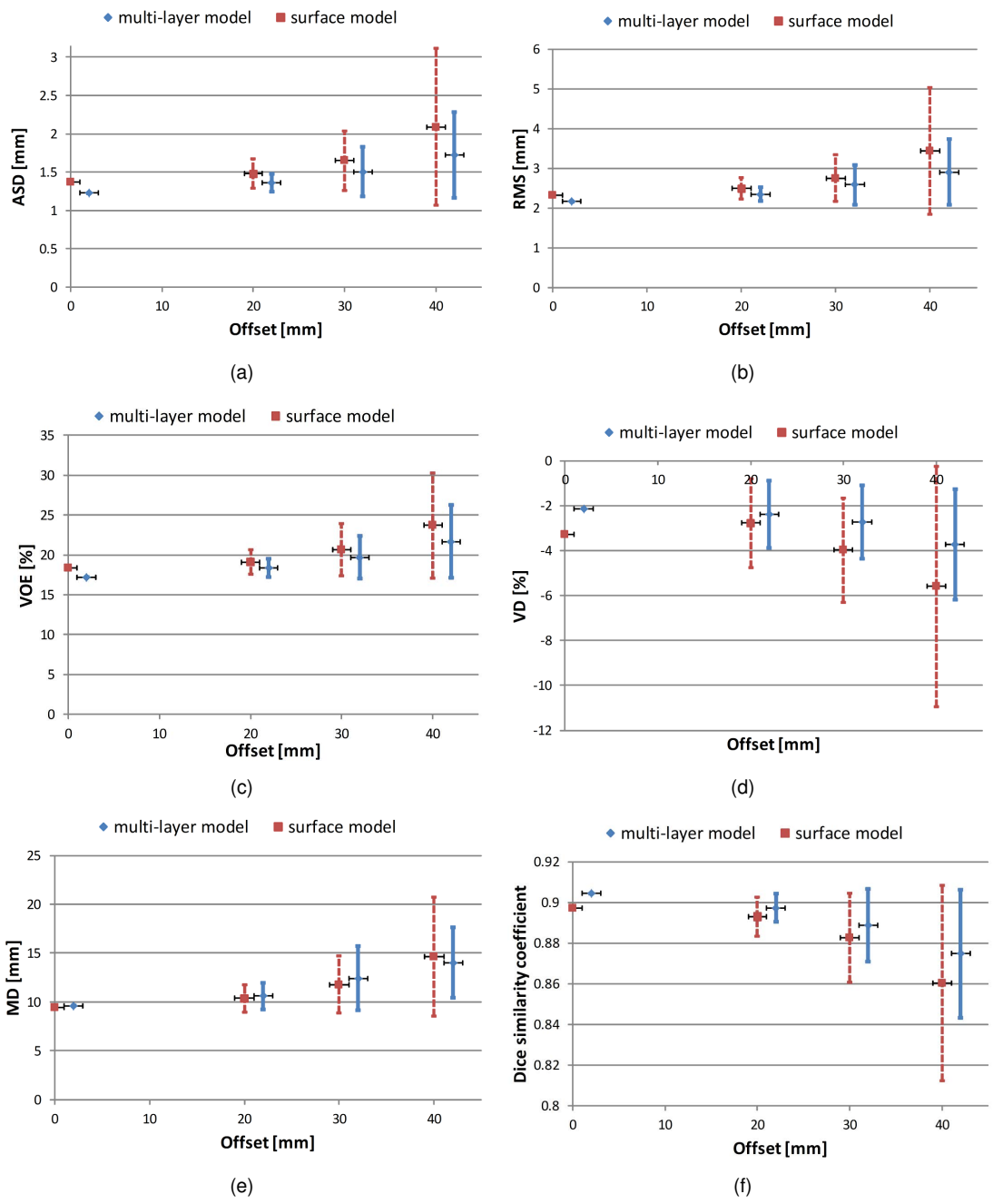


Figure 5.8.: Decrease of segmentation accuracy with increasing distance of the initial model placement to ground truth model initialization (x-,y-,z-offset of ± 40 mm). Results are averaged over all test scans and over x-,y-,z-direction. A multi-layer model and a standard surface model have been used. (a)-(e): a value of 0 denotes a perfect segmentation. (f): a value of 1 denotes a perfect segmentation.

For the tests, a 3-layer multi-layer model of the bladder has been built. Beginning with a ground truth initialization at the center of gravity of the reference labeling, the model is shifted in x-,y- and z-direction in the range of ± 40 mm. For each shift, segmentation quality measures are computed. The same procedure is done using a non volumetric surface model. Figure 5.7 shows plots of

quality measures ASD, VD and MD at the $z = 0$ plane for x/y shifts averaged over all test scans. The remaining measures showed similar results. As can be seen, the multi-layer model shows better segmentation accuracy in the whole shift range.

Figure 5.8 shows the decrease of segmentation accuracy with increasing distance of the model placement from the ground truth initialization. The results are averaged over all test scans and over the shift range of ± 40 in x-,y-,z-direction. The multi-layer model's accuracy shows a smaller drop of accuracy with increasing distance in comparison to the surface model. Additionally, the standard deviation of the multi-layer model is significantly lower which indicates a robust behavior over all test scans.

It can be concluded that the multi-layer model is more robust to model initialization on low contrast images like CT scans of the bladder than surface based models. Both overall segmentation accuracy and initialization robustness could be increased using the proposed multi-layer model.

5.3.3. Robustness of boundary detection

As it has been discussed in Section 3.4.2, the proposed rule based outlier removal method can be combined with any standard boundary detection approach that computes fitness values for boundary points. Furthermore, the multi-tiered adaptation presented in Section 4.2.3 prevents that the model deviates too much from the reference shape in the early stages of the adaptation process where the model's surface is typically far away from the real boundaries. The benefit of the proposed outlier removal and model adaptation is evaluated by using two different boundary detection methods. The first method is a simple, fast and generic gradient computation as a boundary fitness value. In contrast, a learning based boundary detection method is used that learns the distribution of sampled profiles along the surface's normal at each model point.

Usually, it is assumed that the learning based boundary detection outperforms the simple gradient boundary detection, because it incorporates more sophisticated knowledge about the appearance of the boundary. In the current literature, profile based boundary sampling is therefore used in the majority of applications. However, the proposed outlier detection and multi-tiered adaptation should decrease the difference between both approaches in terms of the resulting segmentation accuracy.

Therefore, 17 CT scans of the liver have been segmented using the method proposed in Section 4.2. The same scans were then segmented using the same method, but with a learning based boundary detection module. The learning based boundary detection is performed as follows: using a set of training data, at each point of the model an intensity profile is sampled along the surface's normal. These profiles are taken as positive samples for the corresponding points. Negative samples are acquired by sampling additional profiles with an offset at each point. For each point, a boosted classifier is built based on AdaBoost (cf. Section 2.4.3) and classification and regression trees (cf. Section 2.4.2). In the detection step on an unseen dataset instead of using a majority vote, the normalized positive and negative votes of the weak classifiers are taken as a probability for a point being on the boundary or not.

In comparison to standard profile based boundary detection [Coo01], the described approach does not assume a Gaussian distribution of profile samples, but is able to capture arbitrary distributions. This improves boundary detection in areas where the mean profile is not representative.

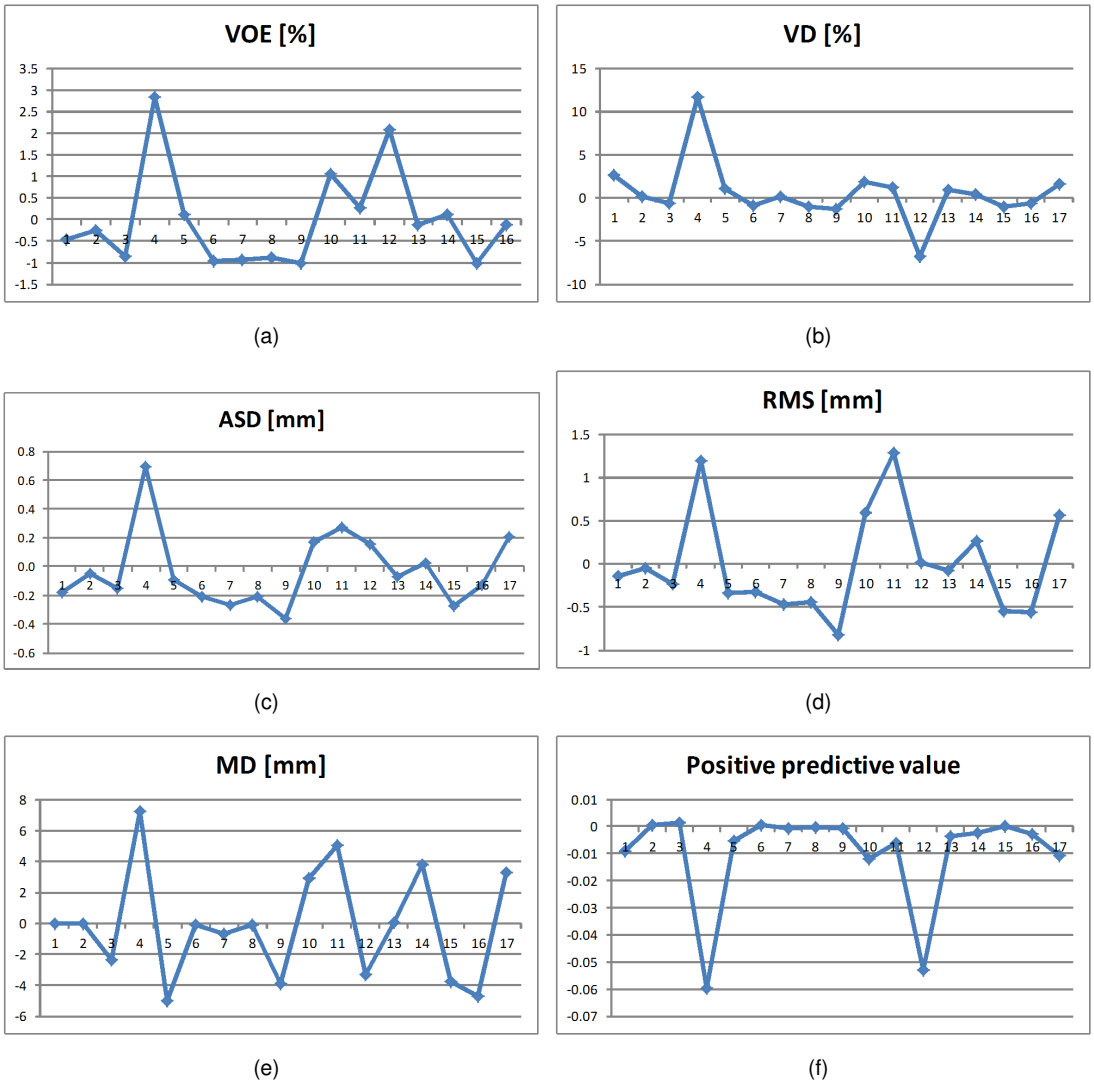


Figure 5.9.: Segmentation accuracy differences between gradient based boundary search and boosted profile based boundary search for 17 scans of the liver. Positive values denote better performance of the gradient boundary search and negative values denote the contrary. Both boundary search methods are combined with the outlier detection presented in Section 3.4.2 and the multi-tiered adaptation presented in Section 4.2.3. In all listed measures, only marginal differences can be observed.

Both boundary detection methods are applied to the test cases and VOE, VD, ASD, RMS, MD and PPV are computed on the segmentation results (cf. Section 5.1). Figure 5.9 outlines the differences between both methods in terms of segmentation accuracy. Positive values denote better performance of the gradient boundary search and negative values denote better performance of the learning based boundary detection. As can be seen, both approaches only marginally differ in terms of accuracy. For most datasets the average surface distance only varies less than 0.4 mm. Moreover, in some cases the gradient boundary search shows slightly better performance and in some cases the learning based boundary detection shows superior results.

Table 5.3.: Segmentation accuracy of different statistical shape model of the liver averaged over 46 datasets (cf. Section 5.1 for metrics).

Method	PPV	DC	VOE [%]	VD [%]	ASD [mm]	RMS [mm]	MD [mm]
PCA	0.88	0.86	19.66	-8.27	5.86	8.62	40.21
RELAXED PCA	0.88	0.87	18.97	8.83	5.45	8.17	41.01
KPCA	0.82	0.82	25.31	34.12	7.09	11.09	57.60

5.3.3.1. Discussion

The evaluation of this section shows, that there is almost no difference between a simple gradient boundary search and a state of the art learning based boundary detection in terms of segmentation accuracy if combined with the outlier detection presented in Section 3.4.2 as well as the multi-tiered adaptation presented in Section 4.2.3. It can therefore be concluded that the proposed methods allow the usage of very simple, fast and generic boundary detection methods without loosing segmentation accuracy. This result implies that the significant information about correct boundaries is encoded in the outlier detection and adaptation process and not in the local boundary detection method. Since local organ boundaries are often not unique in the image, local boundary detection methods may produce a lot of false positives. However, a large part of these false positives can be easily eliminated by incorporation of additional knowledge as presented in this work. In this regard, the result can be interpreted analogous to the result of the model initialization process presented in Section 5.3.1. Here, the main information about the position of the correct bounding boxes for each organ does not come from the simple local classifiers but from the global high level information about the relative positions of each organ to all other organs.

5.3.4. Linear and nonlinear shape models

As it has been discussed in Section 2.5.3, standard statistical shape models assume a Gaussian distribution of organ shapes. It can be argued though, that some organs with complex shape variation like the liver may not be Gaussian distributed. Recently, Kirschner et al. [KBW11] investigated the benefit of landmark based statistical shape model segmentation using standard linear PCA and nonlinear Kernel PCA (KPCA). They showed that nonlinear statistical shape models can improve the segmentation accuracy on structures where the mean is not representative for the distribution of shapes. In their tests, a statistical shape model mixing lumbar and thoracic vertebrae has been created, where the mean shape is neither a lumbar nor a thoracic vertebra.

In order to evaluate whether the model based segmentation methods proposed in this thesis benefit from a nonlinear statistical shape model, three different statistical shape models of the liver have been built based on 220 training meshes. The first model uses standard PCA as described in Section 2.5.3. The second model uses a recently proposed relaxed PCA model, which is able to leave the trained PCA subspace [KW11]. This approach has been evaluated by the authors on liver CT data and showed superior segmentation results compared to standard PCA. The third model uses a Kernel PCA shape energy as described in [KBW11]. It is therefore a nonlinear statistical shape model. The scale factor for the Gaussian kernel has been set to 0.3, since this value yielded the best results in the tests.

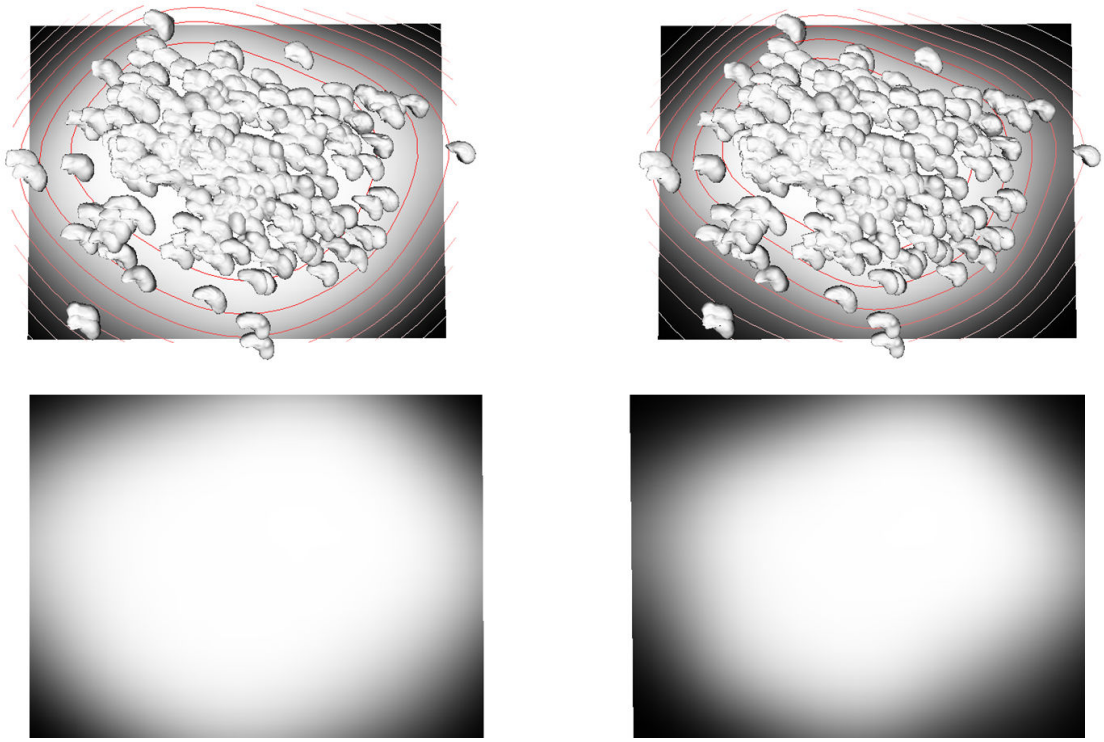


Figure 5.10.: Projection of the training shape set of the liver to the first two principle components. The first row shows the single projected training shapes. The second row shows the log-likelihood function of the shape energy where brighter shading means higher probability of a shape to be a plausible liver shape.

For evaluation, 46 CT scans of the liver with slicing between 1.0 mm and 5.0 mm have been used. The models were initialized at the center of gravity of the ground truth labellings in order avoid bias from bad model positioning. For boundary detection, the method described in Section 4.2 has been applied. Furthermore, no multi-tiered adaptation has been used, but solely statistical shape model evolution. This has been done in order to evaluate the benefit of the different statistical shape models only instead of the whole segmentation pipeline. Quality measures according to Section 5.1 have been computed for all three models. Table 5.3 shows the segmentation results of the tests. For most measures the relaxed PCA model shows slightly better performance than the standard PCA model. The KPCA model, however, performed worse than its nonlinear counterparts. This could indicate that the shape distribution of the liver is Gaussian distributed and that the KPCA models some non representative nonlinearities that appear in the training data.

Figure 5.10 shows the projection of the training shape set to the first two principle components. The shading encodes the probability of a shape to be a plausible liver shape according to the log-likelihood function of the shape energy. As can be seen, the training shapes cluster around the point with the highest probability. That means, the mean shape is very representative for the single training liver shapes. In this case, a linear model can be assumed to be sufficient to model the training set. The visualization of the distribution of the training shapes also supports the qualitative results presented in Table 5.3.

Table 5.4.: Quantitative results for segmentation of the left and right kidney averaged over 30 CT scans (cf. Section 5.1) [ES10].

LEFT KIDNEY	DC	VOE [%]	ASD [mm]	MD [mm]
Mean	0.94	0.12	0.40	5.72
Standard deviation	0.02	0.04	0.13	1.93
Worst	0.90	0.17	0.57	9.84
Best	0.97	0.05	0.18	3.16
RIGHT KIDNEY	DC	VOE [%]	ASD [mm]	MD [mm]
Mean	0.92	0.13	0.61	6.56
Standard deviation	0.05	0.08	0.59	4.77
Worst	0.77	0.37	2.46	20.85
Best	0.96	0.04	0.23	3.16
Mean total	0.93	0.12	0.51	6.14

5.3.4.1. Discussion

The tests showed that linear PCA is even for a complex shape like the liver sufficient to model the training distribution. It can be argued though, that nonlinear models may perform better when trained with a larger database. However, since the used training base already consists of 220 shapes, deviations from the Gaussian distribution are likely to be small.

Nonlinear statistical shape modeling did not show superior results to standard linear models. At the same time it led to computation times that are several magnitudes higher than for linear models, especially when many training shapes are used. Therefore, in this thesis, a linear statistical shape model has been applied.

5.4. Kidney segmentation

In this section, the kidney segmentation framework as proposed in Section 4.1 is evaluated. 30 patients have been randomly selected from a database (26 cases without visible lesions in the kidneys, 4 cases with lesions up to 3.5 cm in diameter) of CT images acquired by a Siemens Sensation scanner. The in-plane spacing of the CT data was 0.74 mm. The slices were reconstructed with an axial spacing of 5 mm. For the tests, the kidneys have been initialized using a user defined point in the middle of the kidneys.

The segmentation accuracy is evaluated by computing DC, VOE and MD (cf. Section 5.1). Table 5.4 lists the mean and standard deviation of the above measures for the left and right kidney averaged over the used data sets as well as the best and worst result for each category. Exemplary qualitative segmentation results are shown in Figure 5.11.

The total segmentation time on a system with a 2.93 GHz Intel Quad Core processor, 8 GB RAM and Windows 7 for the given datasets is below 6 seconds.

5.4.1. Discussion

The results show a slightly better performance for the left kidney. An explanation for this outcome is the neighborhood of the right kidney to the liver. The liver's shape has a strong influence on the

position and orientation of the right kidney. Furthermore, if the right kidney has direct contact with the liver, often no clear boundary is present which makes the adaptation process less robust.

Showing a mean dice similarity coefficient of 93% and a mean overlap of 87% on 30 patients, the segmentations generated by the proposed method are in most cases comparable to manual delineation. In comparison to the study of Lin et al. [LLH06] the dice similiarity coefficient on the tested datasets could be increased by around 4%. However, the approach of Lin et al. has been evaluated using an automatic initialization.

The proposed approach also yielded good results compared to Tsagaan et al. [TSKM02] who also evaluated their approach using manual placement of the model. They reported a mean overlap of 86%. It must be noted though that their experiments have been performed on low resolution data with an axial spacing of 10 mm.

A major advantage of the proposed framework is the speed and ease of use. Since a segmentation is done very quickly, it can be easily integrated into clinical practice. Even concerning complicated cases (e.g. with severe pathologies) the framework can reduce the clinician's time by providing an acceptable segmentation to start with. The user may then choose to edit the segmentation based on his needs.

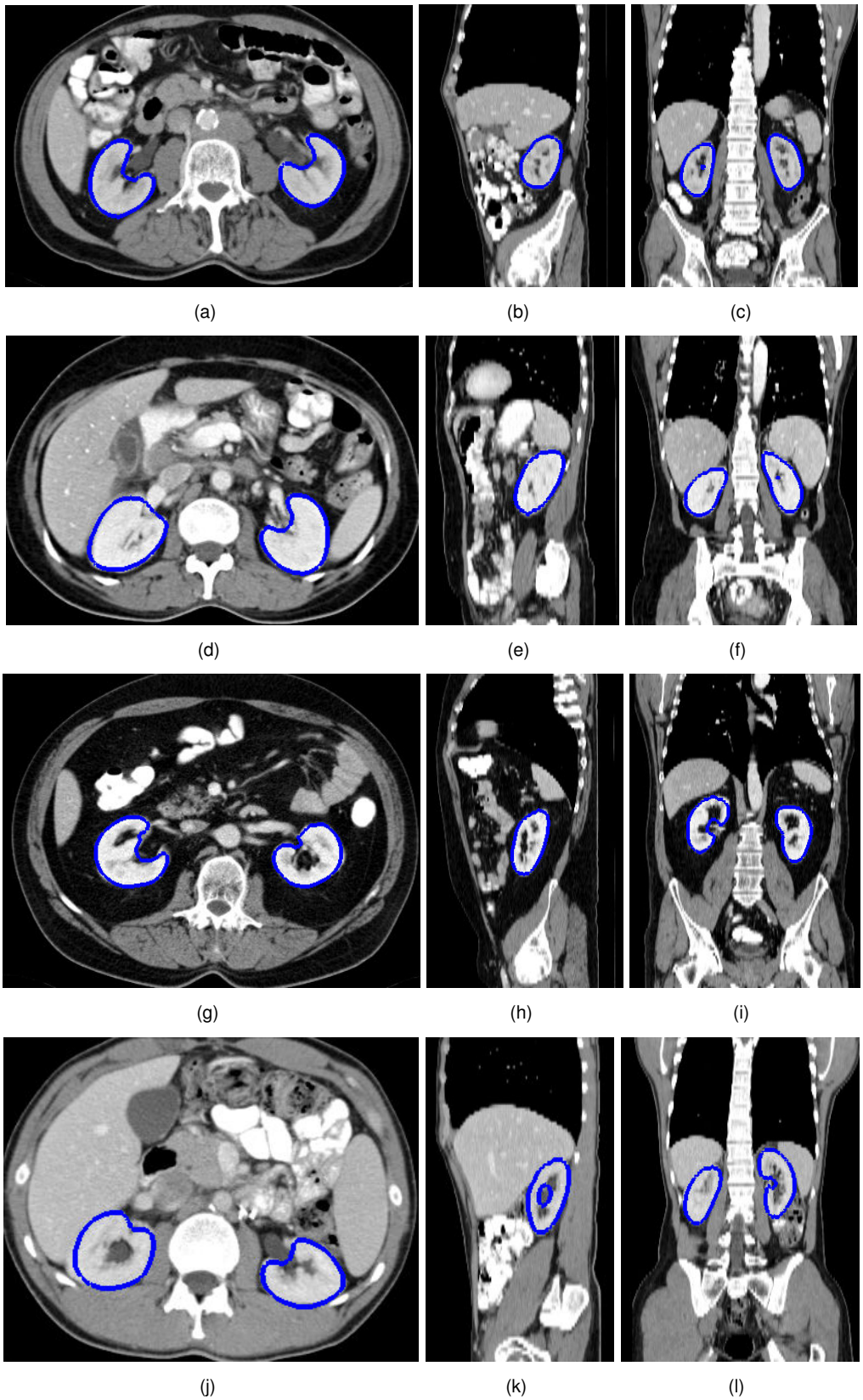


Figure 5.11.: Exemplary segmentation results of left and right kidney in axial (1st column), sagittal (2nd column) and coronal (3rd column) views [ES10].

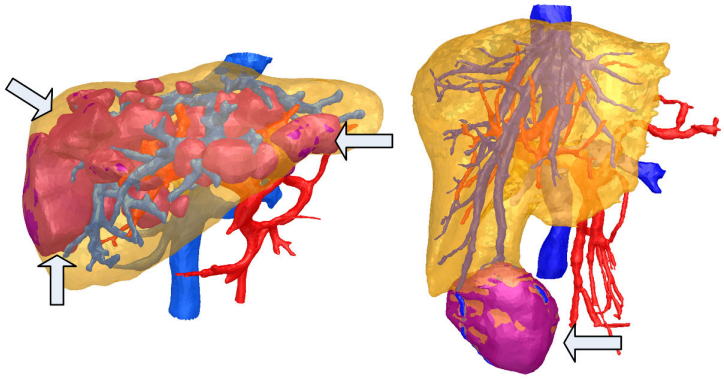


Figure 5.12.: Exemplary cases of the challenging IRCAD data base [ESKW10]. Left: scan with 20 tumors inside of the liver. Right: a huge tumor at the lower liver lobe leads to a very uncommon liver shape. In such a case a precise segmentation of the liver is impossible if only trained shape priors are used.

5.5. Liver segmentation

In the context of the liver segmentation competition held at a conference workshop of MICCAI 2007 [HvGSea09], a data base of CT images was made publicly available in order to let teams compare their approaches. The majority of the best automatic methods use statistical shape knowledge [HvGSea09]. Recent results show that both implicit [WSH09] and explicit shape representations [KLL07] are equally suited to segment the liver with high accuracy (1.37/1.02 mm mean error). Instead of using the 10 test cases for online competition, some teams are using the provided 20 training sets for evaluation [KUA*09, SBB09]. However, most methods are evaluated on a small amount of test cases (usually between 10 and 25 [WSH09, KUA*09, KLL07, HMW07, SBB09]). A test on a larger data base (174 scans) was performed in [LZZ*08] with a reported 1.59 mm mean error after a manual removal of outliers (1.76 mm without outlier exclusion). However, a non-public data base from a single vendor was used for evaluation so it is unclear how the method would perform on the highly heterogeneous scans from the public competition data bases. It is also unclear how many high and low resolution scans were used. This is especially crucial since only the average surface distance was used as a metric for evaluation.

The automatic liver segmentation approach presented in Section 4.2 has been evaluated using 4 non-overlapping sets consisting of 86 CT liver scans differing in resolution, contrast enhancement and present pathology. None of the test sets were used for training. Set MTEST is the test set of the MICCAI'07 liver challenge (10 scans), cf. [HvGSea09]. Set MTRAIN is the training set of the same challenge (20 scans). Instead of using this set as a training set it is used as an additional test set. This is done in order to compare the method to other approaches which also use this set for evaluation [KUA*09, SBB09].

In order to add very challenging datasets to the evaluation the public 3D-IRCAD data base (www.ircad.fr) built for comparison of segmentation algorithms has been used (test set IRCAD). 75% of the 20 provided scans (axial spacing between 1 and 4 mm) having hepatic tumors, which makes a robust segmentation extremely difficult. Figure 5.12 shows a 3D rendering of an examplary dataset of this testbase with large visible lesions.

Table 5.5.: Mean and standard deviation of volume and surface errors (cf. Section 5.1) for the 86 scans listed separately for each test set using the proposed fully automatic liver segmentation approach [ESKW10].

Num. scans / test set	VOE [%] $\pm SD$	VD [%] $\pm SD$	ASD [mm] $\pm SD$	RMS [mm] $\pm SD$	MD [mm] $\pm SD$
10/MTEST	8.62 ±1.49	1.32 ±2.04	1.54 ±0.44	3.13 ±0.99	25.90 ±7.28
20/MTRAIN	7.54 ±1.18	1.28 ±2.15	1.30 ±0.33	2.67 ±0.97	26.52 ±10.07
20/IRCAD	10.34 ±3.11	1.55 ±6.49	1.74 ±0.59	3.51 ±1.16	26.83 ±8.87
36/LOWRES	12.35 ±1.98	-4.53 ±4.05	1.85 ±0.34	3.27 ±0.68	23.09 ±5.62

Since most liver segmentation algorithms are only tested on high-resolution data a test set LOWRES has been added which consists of 36 thorax/abdomen scans from a Siemens Sensation scanner with the standard slice spacing for abdominal control scans of 5 mm.

A total of 86 scans were taken to evaluate the overall segmentation accuracy using the measures described in Section 5.1. For evaluating the initialization robustness, MTEST is excluded since the ground truth is only available to the challenge initiators.

The final segmentation results are listed in Table 5.5 and shown in Figure 5.13 for exemplary scans of the test sets. MTEST was submitted to the organizers of the challenge. It can be seen that the combination of trained shape constraints with observed shape deviation allows the model to robustly leave the trained space without leaking into neighboring structures. The fully automatic segmentation of one case takes in average 45 seconds on a 2.93 GHz Quad Core CPU.

5.5.1. Discussion

A fully automatic multi-tiered statistical shape model segmentation for the liver has been presented that combines learned local shape constraints with observed shape deviation during adaptation. The said combination allows the model to leave the trained space without leaking into neighboring structures. In comparison with using only trained local curvature constraints [KUA*09] the presented method yields significantly superior results (71.3 to 66.7 MICCAI challenge score). With a mean surface distance of 1.3 to 1.54 mm for the commonly used [HvGSea09] evaluation sets MTRAIN and MTEST the approach places among the top automatic methods [HvGSea09, WSH09, KUA*09, KLL07, HMW07, LZZ*08].

However, in contrast to other methods [HvGSea09, WSH09, KLL07, HMW07] (10 test scans), the approach has been evaluated using a large heterogeneous test set of 86 scans. Adequate comparability between methods is only possible if evaluation is performed on publicly available test cases. Therefore, the 3D-IRCAD data base (www.ircad.fr) has been included. This database has been used since it is a very challenging data base for segmentation. Some scans contain up to 46 tumors (Figure 5.12 left), some of them being extremely large (Figure 5.12 right). Nevertheless the approach is able to segment this set with high accuracy. Because most methods [HvGSea09, WSH09, KUA*09, KLL07, HMW07] are mainly tested on high-resolution data, the approach has been evaluated using 36 scans with the standard slice spacing for abdominal con-

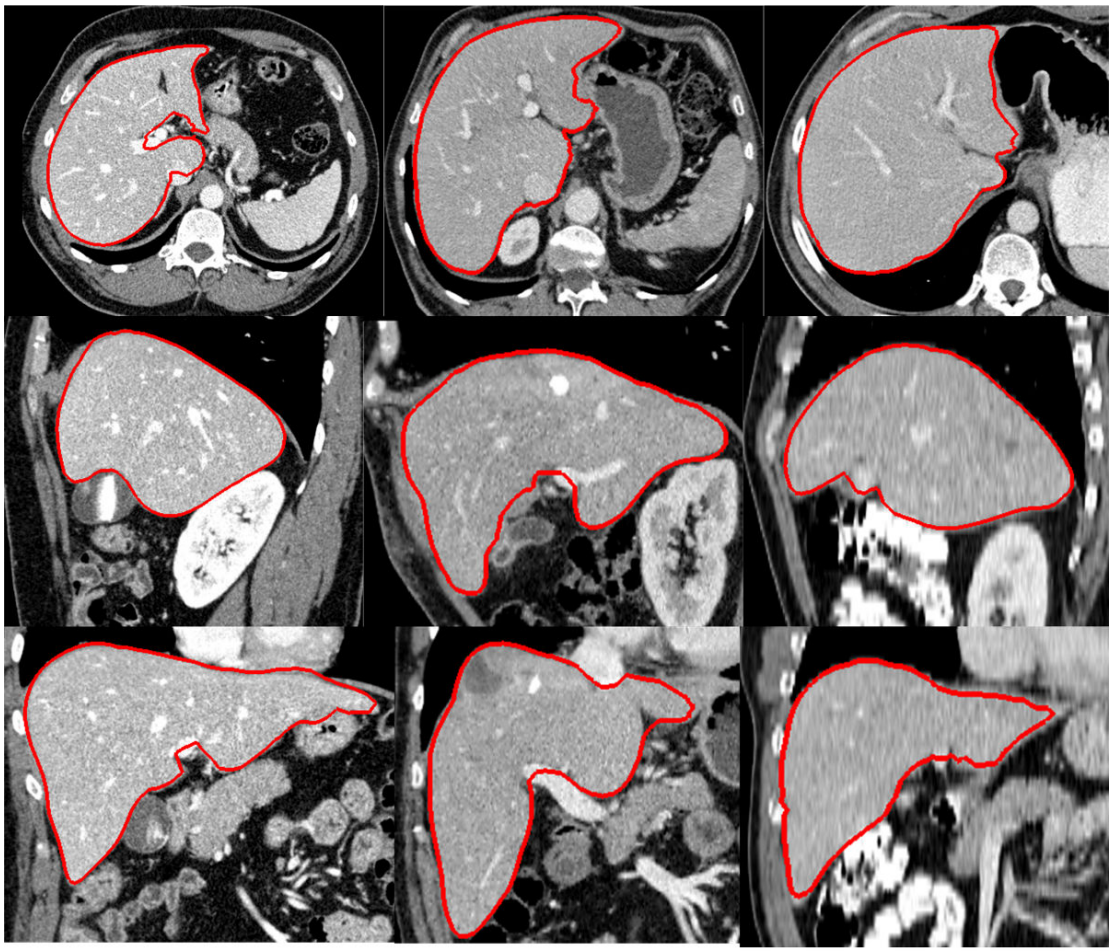
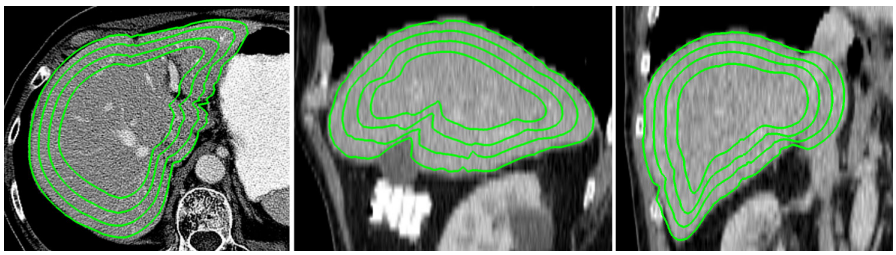


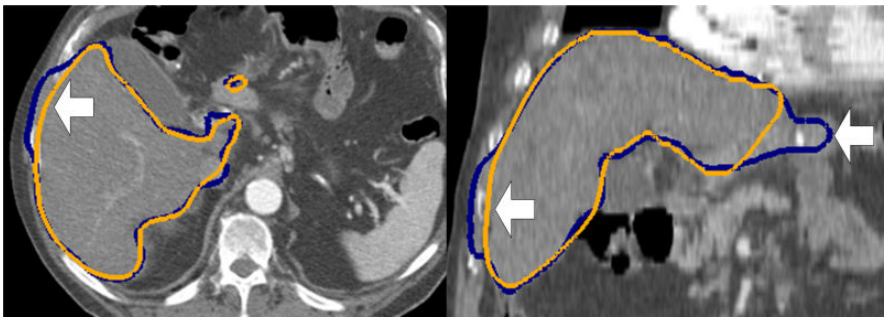
Figure 5.13.: From left to right: qualitative segmentation results for test sets MTRAIN, IRCAD and LOWRES. Each column shows transversal (top), sagittal (middle) and coronal (bottom) view [ESKW10].

trol scans of 5 mm. The approach is also able to segment this data with high accuracy which emphasizes the broad applicability of the method.

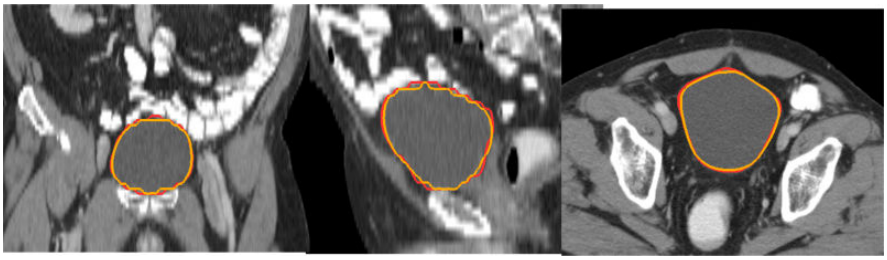
5.6. Segmentation of low contrast structures



(a)



(b)



(c)

Figure 5.14.: (a) Segmentation result for a 4-layer model [ESW10]. (b) Comparison of final segmentation using a single surface model (dark outline) and a 4-layer MLDM (bright outline) [ESW10]. (c) Segmentation result of a bladder using MLDM (bright outline) compared to ground truth (dark outline) [ESW10].

In order to evaluate the method for segmenting low contrast CT structures as proposed in Section 4.3, 20 CT scans of the liver with an axial spacing between 1 and 5 mm as well as 10 CT scans of the bladder with an axial spacing of 5 mm has been used. Segmentation accuracy has been evaluated using VOE, PPV, ASD, RMS and MD (cf. Section 5.1). Smaller values mean better results except for PPV. For the tests, a 3-layer bladder multi-layer model and a 4-layer multi-layer model of the liver have been built. Using more layers did not significantly improve segmentation accuracy but had a negative impact on the computation time. The optimization time of a layer

Table 5.6.: Mean and standard deviation of volume and surface errors (cf. Section 5.1) for 20 scans of the liver and 10 scans of the bladder using a 1-layer surface model and a multi-layer model [ESW10].

Method	VOE [%] $\pm SD$	PPV [%] $\pm SD$	ASD [mm] $\pm SD$	RMS [mm] $\pm SD$	MD [mm] $\pm SD$
1-LAYER (LIVER)	11.74 $\pm$3.09	94 $\pm$2	2.18 $\pm$0.80	3.87 $\pm$1.75	31.05 $\pm$12.81
4-LAYER (LIVER)	10.56 $\pm$3.39	95 $\pm$2	1.96 $\pm$0.84	3.81 $\pm$1.88	31.76 $\pm$12.56
1-LAYER (BLADDER)	19.21 $\pm$8.12	91 $\pm$9	1.61 $\pm$0.91	2.77 $\pm$1.39	11.52 $\pm$5.34
3-LAYER (BLADDER)	17.58 $\pm$6.46	95 $\pm$5	1.31 $\pm$0.53	2.35 $\pm$0.84	10.38 $\pm$3.74

model of 10000 points is 48 ms on a 2.93 GHz Intel Quad Core CPU. A single iteration takes around 7 seconds. The total adaptation takes an average of 121 seconds per case.

In order to evaluate the benefit of using multiple layers, all tests have also been run using a multi-layer model with a single layer. The models have been initialized at the same position using the center of gravity of the ground truth [YS09]. The results are shown in Table 5.6 for the liver and the bladder averaged over the test scans. It can be seen that the segmentation accuracy improves when multiple layers are used. Figure 5.14(a) shows the adapted multi-layer model after convergence. A qualitative comparison between a multi-layer model with 1 layer (dark outline) and 4 layers (bright outline) is shown in Figure 5.14(b). As can be seen, the 4-layer multi-layer model does not leak into neighboring structures (arrows). Figure 5.14(c) shows an exemplary segmentation result using a multi-layer model for a bladder scan (bright outline) compared to ground truth (dark outline).

5.6.1. Discussion

The multi-layer deformable model has been applied to the segmentation of low contrast structures in CT. Evaluation using 20 CT images of the liver and 10 bladder scans showed that the approach improves single surface segmentation especially in areas of low contrast boundaries and yields state of the art segmentation results [CDNA07, HMMW07, OSS*07]. Given the linear nature of the optimization framework and the possibility to limit the amount of layers in the model, the total computation time could be limited to an average of 121 seconds which is very fast compared to other volumetric methods (cf. [SH09] ca. 1000 seconds).

Furthermore, the incorporation of statistical shape knowledge is naturally given through corresponding landmarks between layers. Therefore, the model can be directly applied to enhance existing statistical shape model based approaches which only use a single surface [ZBG*07, KLL07, HMMW07, HM09]. This is especially important for structures with a strong shape variability like the liver.

Table 5.7.: Mean and standard deviation of volume and surface measures (cf. Section 5.1) for the proposed pancreas segmentation method using threefold cross-validation on 40 datasets. Results are shown using the tissue detector only and using the whole segmentation pipeline [EKD*11].

Process	VOE [%] $\pm SD$	VD [%] $\pm SD$	ASD [mm] $\pm SD$	RMS [mm] $\pm SD$	MD [mm] $\pm SD$	PPV [%] $\pm SD$
Detector only	47.8±10.31	10.53 ±3.73	2.58 ±0.77	4.69 ±1.15	23.62 ±5.45	74.9±10.7
Final result	38.8±9.08	5.62 ±3.47	1.70 ±0.71	3.10 ±1.13	16.13 ±5.18	81.07 ±11.2

Table 5.8.: Performance results of the proposed method compared to state of the art pancreas segmentation systems. Unpublished/unavailable information is marked as —. The best score in each category is marked bold [EKD*11].

Method	Tested scans	Run time	Needed CT scans per case	Accuracy
Shimizu *et* al. level-set [SOI*07]	10	—	**1** (non-contrast)	32.5 % overlap
Shimizu *et* al. atlas [SKK*09]	20	45 min.	3 (multi-phase)	57.9 % overlap
Kitasaka *et* al. [KSM*08]	22	—	4 (multi-phase + non-contrast)	visual inspection: 12 high overlap, 6 medium overlap, 4 no overlap
Proposed method	**40**	**20.4** min.	**1** (contrast)	**61.2** % overlap

5.7. Pancreas segmentation

In order to evaluate the pancreas segmentation framework proposed in Section 4.4, abdominal early parenchyma single phase CT data from 40 cases was acquired. Inter-slice spacing was 5 mm abdominal control standard protocol whereas spacing within an axial slice varied between 0.6 mm and 0.7 mm. Manual delineation from experienced radiologists was taken as the gold standard. Threefold cross validation was used for performance evaluation. For each fold, the statistical shape space as well as the classifiers were learned on the training data and evaluated on the test data.

Table 5.9.: Run time in seconds for each step of the proposed pancreas segmentation method [EKD*11].

organ segmentation	Vessel segmentation	Landmark detection	Feature computation + classification	Belief propagation	Model adaptation	**Total**
90	5	3	1069	35	25	**1224**

Figure 5.15 shows exemplary qualitative results separately for the detector and final segmentation. While using the detector only (middle row) already shows good accuracy compared to ground truth (bright outline), the incorporation of statistical shape space further improves the quality (bottom row). Generally, there is very limited leakage into neighboring structures (bottom row) given the low contrast of pancreas boundaries (upper row).

Table 5.7 shows quantitative results of the proposed method using the metrics described in Section 5.1. The results are shown using the detector only and using the full proposed processing pipeline including shape model adaptation. The average surface distance to the gold standard is 1.7 mm for the final segmentation. Given such a small distance error, the average overlap of 61.2% seems to be underperforming. However, considering that the pancreas is a very small organ mapped to the given resolution, few missclassified voxels can account for a big overlap difference.

A comparison to the state of the art is given in Table 5.8. The proposed single-phase method is about twice as accurate as the best method using single-phase CT [SOI*07] and even outperforms the best method using multiple phase scans [SKK*09]. It is also the method with the most datasets tested. Since no time consuming atlas-based registration is used, the proposed method is also more than twice as fast. Table 5.9 shows the runtime results in seconds for each processing step.

5.7.1. Discussion

An approach for fully automatic segmentation of the pancreas in single-phase contrast enhanced CT images has been presented. Pancreas segmentation is a very challenging task due to often non-visible borders to surrounding structures. In contrast to prior work, not only tissue appearance is considered, but also a high amount of anatomical knowledge — knowledge which is also used by the clinician to distinguish pancreas from tissue having the same texture. This is done by detecting clinically meaningful support structures and building a classifier that models local spatial relationships between the pancreas and the support structures.

Furthermore, performant texture descriptors based on wavelets and cosine transform have been proposed to model local appearance. The resulting classification is used to guide a statistical shape model for fine segmentation. Cross-validation on 40 datasets showed high quality results while needing only a single abdominal contrast CT scan per patient. This method even outperforms available automatic approaches that need multi-phase CT data and is in some cases similar to interactive segmentation. Since the approach is fully automatic and uses single-CT, it can be easily used to improve detection rates for other challenging abdominal structures, especially malign lymph nodes and intestine.

5.8. Automatic registration of liver CT scans

For evaluation of the automatic registration of liver CT scans described in Section 4.5, 22 CT volumes of 11 patients showing arterial and portal phase contrast agent saturation have been used. The inter-slice spacing varied between 1.0 and 1.25 mm. All volumes were acquired by a GE Light Speed Ultra CT scanner. Figure 5.16(a) (top row) shows the result of the portal phase segmentation. Figure 5.16(a) (bottom row) outlines the result of the deformed arterial phase. It can be seen that the transformed borders of the arterial phase now match the boundary shape of the portal phase. An exemplary registration result is shown using split views in Figure 5.16(b).

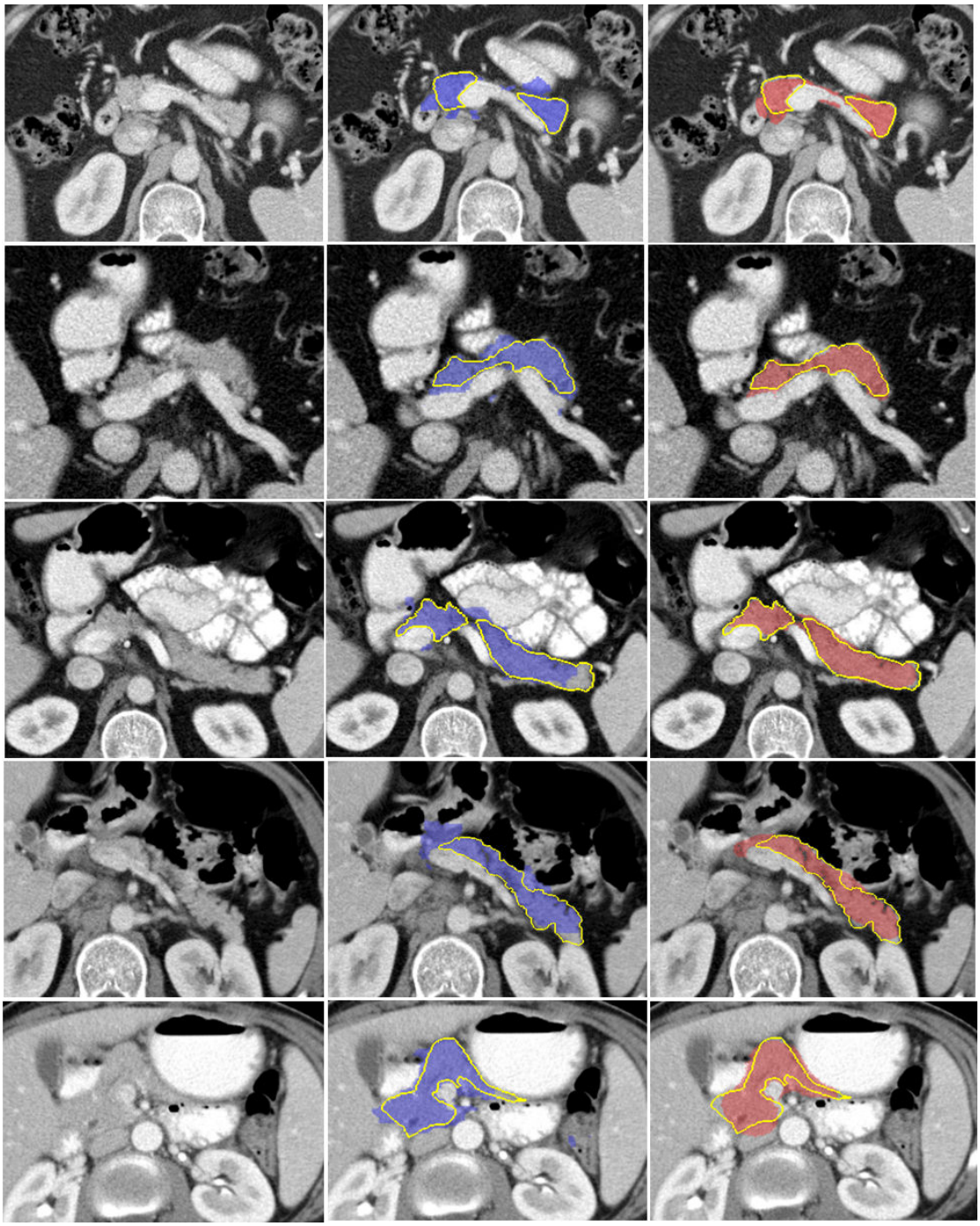
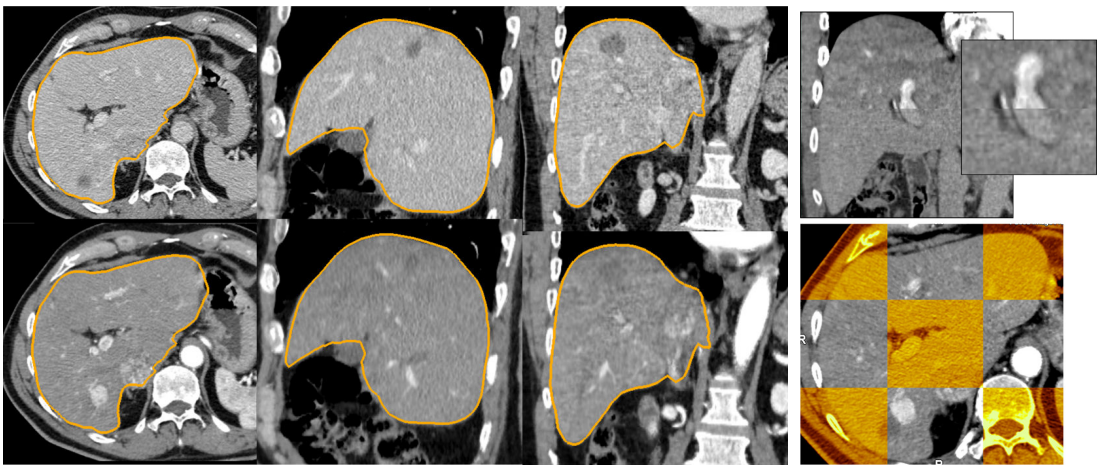


Figure 5.15.: Exemplary segmentation results of the pancreas on unseen data [EKD*11]. Original datasets (left column), detector output (blue, middle column) and final segmentation result (red, right column). The outline marks the ground truth.



(a) (b)

Figure 5.16.: (a) Segmented portal shape overlayed with the portal phase (top row) and with the deformed arterial phase (bottom row) [ESH*11]. (b) Split views of registration results. Interior vessels and outer liver shape are well matched [ESH*11].

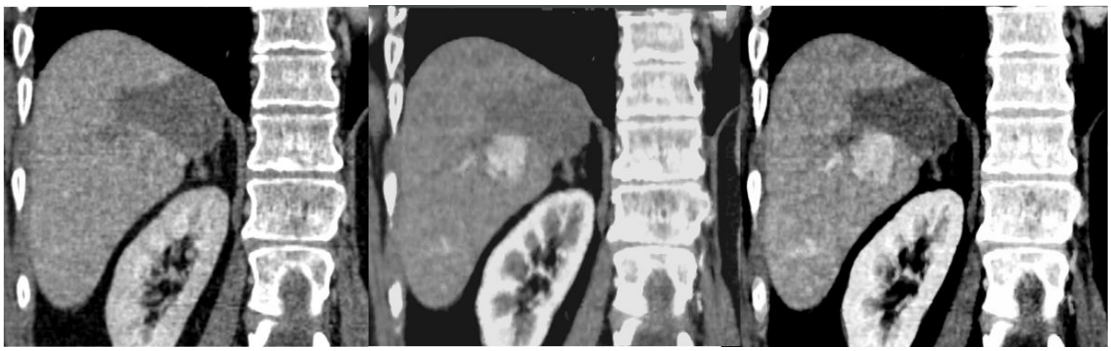


Figure 5.17.: Registration of liver phases allow the fusion of complementary structures in one image [ESH*11]. Left: portal phase. Middle: registered arterial phase. Right: fused result that shows burned liver tissue (dark) and liver lesion (bright) together.

Both vessels (top) and boundaries (bottom) are well aligned. Figure 5.17 shows a fusion of an arterial and portal phase CT scan. Both lesion (bright) and burned liver tissue (dark) are clearly visible in the fused image (Figure 5.17 right).

Since quantitative validation of deformable registration is generally problematic, two experiments have been performed by experienced radiologists to judge matching of outer shape and interior structures. The registration quality of the interior structures has been validated using a 5-point scale (1 = excellent alignment, 5 = insufficient alignment). Table 5.10 shows the corresponding scores. Most cases show good to high scores except of one case that included a lot of artifacts caused by breathing during scanning. However, the misalignment was higher towards the liver boundaries while the middle of the liver was well matched.

In another experiment, the registration quality of the outer shape has been evaluated. The liver was divided into three parts, covering the upper, middle and lower third of the liver in order to see how well the deformation is compensated in each area. Radiologists were then judging

Table 5.10.: Validation scores of deformable registration based on interior structure given by radiologists (1 = excellent alignment, 5 = insufficient alignment) [ESH*11].

Patient	1	2	3	4	5	6	7	8	9	10	11	Mean
Score	1	2	1	1	2	4	2	1	2	2	1	1.72

Table 5.11.: Average displacement error of the outer liver shapes in cm after deformable registration of lower (L), middle (M) and upper (U) liver parts judged by visual inspection of radiologists [ESH*11].

Patient	1	2	3	4	5	6	7	8	9	10	11	Mean
L	0.8	0.0	1.0	0.0	0.5	0.2	1.2	1.0	0.0	0.0	0.2	0.45
M	0.5	0.2	0.0	0.0	0.0	0.0	0.8	0.8	0.0	0.0	0.2	0.23
U	0.0	0.5	0.3	0.0	0.0	0.0	0.8	0.5	0.0	0.0	0.3	0.22

the average displacement error of the parts by visual inspection after deformable registration had been performed. Table 5.11 shows the results for the tested cases. As can be seen, the highest deviation between phases remains in the lower part while the middle and upper part are in most cases well registered. The reason that the registration accuracy of the lower liver part is inferior to the other parts is due to the lower segmentation accuracy at that position. The lower liver lobe usually shows a lot of inter-patient deviation which makes it difficult to segment that lobe in some cases. However, the middle and upper liver part are usually of higher medical importance since they contain the most volume and vessels.

The fully automatic segmentation and registration for one case takes approximately 7 minutes on a 2.93 GHz Intel Quad Core processor with 8 GB RAM and Windows 7 using a cropped area around both livers. A multi-scale image pyramid has been used for acceleration of the registration part. Using the full image content, the required time raises by a factor of 4-8 dependent on the original dimensions of the scan. However, for diagnosis and operation planning usually only the liver is of interest so working with the cropped area is sufficient.

5.8.1. Discussion

A novel multi-stage approach for automatic registration of the liver in contrast enhanced CT volumes has been presented. The method is based on an automatic pre-segmentation of the liver in the different phases utilizing the extracted shape information for registration. The method has been evaluated using 22 CT volumes from 11 patients. The matching quality of outer shape and internal structures was validated by experienced radiologists. The results suggest the applicability of the method in clinical practice.

In contrast to existing methods, the approach allows for a high accuracy natural deformation without having to rely on error-prone extraction or matching of the liver's internal structure. Furthermore, since shape information of the liver is given in all phases, registration can be significantly speed up. However, the registration accuracy of the method depends on the segmentation error of the liver extraction steps. While small segmentation errors can be compensated by the registration step, high deviations of the extracted shape to the real liver boundaries may introduce matching errors.

5.9. Discussion

In this chapter, the model based methods for segmentation and registration that have been presented in this thesis have been evaluated. Here, standard quality measures have been used to compare the segmentation accuracy of the proposed methods against ground truth labels created by humans. Clinical routine CT images of the liver in the portal venous and arterial contrast phase, the bladder, the pancreas, and the kidneys have been used. An average surface distance error between 0.5 mm and 2.0 mm has been obtained. This is close to the inter-observer variability between different humans segmenting the same structure [HvGSea09]. For most structures, the segmentation accuracy could be significantly increased by using the proposed methods in comparison to existing approaches. Furthermore, referring to the diversity of the test sets, it can be concluded that the presented methods deal well with anatomical shape variance between human individuals.

As it has been shown, the proposed volumetric model exhibited a higher robustness against model initialization errors than standard surface models and a higher segmentation accuracy in low contrast data sets. The use of the volumetric model also enabled the fully automatic registration of multiphase CT data. Here, a qualitative and quantitative clinical evaluation has been performed. In nearly all tested cases, the registration quality has been judged to be sufficient for clinical needs.

Based on the results of this chapter, it can be concluded that the goals of this thesis have been met. The proposed methods extend the state of the art of model based methods in significantly increasing the robustness of organ segmentation and registration in common medical imaging scenarios. Advances in model initialization robustness, boundary detection robustness and adaptation accuracy have the potential to save valuable clinician time in dealing with today's 3D medical imaging data.

Parts of this chapter are based on publications of the author [ES10, ETS09, ES10, ESKW10, ESW10, ESH*11].

6. Conclusions and future work

Today's high resolution 3-dimensional medical image acquisition techniques enable sophisticated clinical diagnosis and patient specific treatment that would have been infeasible two decades ago. With this great potential, however, also comes the burden of coping with the huge amount of data produced. Manual contouring of target structures and mental fusion of images from different time series are respectively two of the most time consuming and mentally most complex challenges in clinical practice.

In this thesis, significant advances to address these challenges have been made. Methods for the automatic organ segmentation and registration in Computed Tomography have been proposed to alleviate the time consuming manual contouring work of several organs as well as to allow a registration of CT liver images taken from different points in time.

A major contribution of this work is the development of a volumetric model based method for organ segmentation and registration. In contrast to other volumetric methods, the proposed model has the same class of complexity as standard surface based landmark models and can therefore be optimized very efficiently. It is generalizable and extensible such that it can be combined with statistical shape modeling methods and standard boundary detection approaches. The volumetric model shows superior segmentation performance on low contrasted structures like the liver in the arterial contrast phase in comparison to surface based methods.

As another contribution, a locally constrained statistical shape model has been proposed that addresses a major drawback of statistical shape modeling. The incorporation of local elasticity constraints combined with the curvature of the model during adaptation enables the model to robustly leave the trained space without leaking into neighboring structures. This allows for a better segmentation of unseen structures which is a benefit for organs with complex shape variation like the liver. The method has been integrated into an application framework and is currently being tested in clinical use on real patients. Tests show an average segmentation error that is comparable to the human inter-observer variability[1]. Similar results are obtained for segmenting the kidneys and the bladder.

Another methodological advancement is made in the context of automatic pancreas segmentation. The proposed method uses a new form of texture features that are integrated into a pancreas tissue classifier. Combined with the constrained shape model as described above, for the first time, an automatic segmentation of the pancreas from single phase contrast enhanced CT becomes feasible. The presented method may enable new applications in computer aided diagnosis and operation planning in the future. For example, it can be used for the detection of abdominal lymph nodes which are presently only classified by excluding surrounding abdominal structures.

[1] Inter-observer variability denotes the differences between the outcomes of different human individuals segmenting the same structure.

In the context of organ registration, a method for the automatic deformable multi-phase contrast enhanced liver CT registration has been proposed. The method combines the volumetric model based approach described above with an image based deformable registration scheme. The presented method for automatic deformable multi-phase registration has been quantitatively and qualitatively evaluated in the clinic. The method has been integrated into an application that is currently in use for testing in two clinics. In nearly all tested cases, the registration quality was sufficient for clinical needs.

In this work, several contributions to standard model based segmentation techniques in the context of shape model building, model adaptation and boundary detection have been made. These proposed methods can be either combined with standard surface based landmark models or with the volumetric model proposed in this thesis. In particular,

- a multi-tiered adaptation framework with increasing degrees of freedom has been proposed that is more robust to model initialization errors than standard statistical shape model adaptation frameworks.
- a method for combining ground truth label creation with shape model correspondence establishment has been presented. This method deals with arbitrary topologies and considerably speeds up the shape model generation, since there is no need for an independent correspondence establishment step.
- a rule based outlier removal for landmark boundary detection has been presented. The method can be combined with any probabilistic boundary detection method. In tests, it has been shown that the outlier removal boosts the performance of simple boundary detection approaches to be on a par with complex learning based methods. This allows for the application of more simple, fast and generic algorithms in boundary detection.

With the presentation of the described methods, the objectives of this thesis have been fulfilled. For most of the algorithms, the transfer into clinical demonstrators has already been made. The application in clinical practice will ensure sustainability and maturity of the presented methods. It will also display strengths and weaknesses in long term use. In the context of a thesis, only a limited amount of real images can be considered for evaluation. It is therefore important to see how the proposed methods perform in a clinical scenario with a large amount of patients and different diseases.

6.1. Improvement of the developed methods

In order to extend the methods presented in this work to other application scenarios as well as to broaden their applicability on large and diverse test bases, several improvements and extensions are possible.

The multi-layer model as described in this work can be combined with local elasticity constraints analogous to the surface model. The elasticity constraints for the internal links are inherited from the elasticity constraints of the first layer, i.e. they are copies of the surface link constraints. In the tests performed in this thesis for segmentation and registration of organs like the liver and the bladder, this procedure yielded satisfactory results. However, it is a simplification, because some organs are non-homogeneous in their interior. While the filled bladder without contrast agent saturation can be regarded as homogeneous, the interior of the liver consists of several tissues with potentially different elasticity. For example, vessel trees of varying

size represent a significant amount of the whole liver volume. Additionally, pathologies like cysts or tumors often exhibit a different density in comparison to the normal liver tissue. Furthermore, the density of the liver tissue differs between individuals and between scans depending on the amount of contrast agent saturation and the amount of fat stored in the liver.

In order to account for the different materials present inside of organs, experiments should be made in order to determine the correct elasticity constraints of the internal links of the multi-layer model. Such experiments may involve human or animal cadaver specimens of the organ. However, if such specimens are not provided, the elasticity for a specific patient could also be determined based on the adaptation of the multi-layer model in previous CT images of the patient. In case the multi-layer model is already adapted in two or more images of the same patient, the elasticity constraints for the internal links can be calculated based on the corresponding internal and external points of the adapted models. If another scan of the same patient is made, the same constraints can be used. Furthermore, this procedure can be repeated for a number of patients to estimate the best average parameters of the model.

An application scenario of the multi-layer model that has not been addressed in this work is the auto-detection of pathologies in the organ interior using the internal points. These points store information about the expected tissue characteristics at their position inside of the organ. In case a tumor is located somewhere in an organ, the majority of points will led to a proper adaptation of the model. However, some points will indicate a mismatch of tissue appearance. These areas could then be marked for further examination by the physician. Furthermore, tissue classification methods could be used to automatically classify the tissue as tumor as an alternative.

The multi-tiered adaptation pipeline step wise increases the degrees of freedom of the model during adaptation. The same principle could also be applied to the geometry of the model. Here, a level of detail hierarchy could be implemented where the number of points of the model increases with the degrees of freedom. This could lead to better adaptation in areas where only few points and sharp boundaries are present.

The initialization of the model in the image is done by detecting axis aligned bounding boxes of the structure to segment. In the context of CT or MRI imaging, the patient is usually recorded in a supine and head first manner which implies that all patients are roughly rotationally aligned. However, single organs like the kidney still may have different orientations between patients. This issue could be addressed by defining object aligned bounding boxes in the training stage. In the detection stage, the images can be rotated in order to account for different orientations. Another option is to train several classifiers for different positions. In case of scans in prone position, the image can also be mirrored before classification.

The bounding boxes used in the presented initialization method are rectangles, because the method should be applicable to a variety of different structures without changing parameters. However, organs are not rectangular. While it could be argued that the surrounding of an organ also exhibits unique characteristics that can be learned inside the box, for organs with a strong shape variance it makes sense to use other shapes for detection, for example the mean shape of the organ.

An important issue regards the absence of an organ. For example, the spleen or a single kidney may be operationally removed and therefore do not appear in the image. In this case, the presented method will select the box that has the highest probability of being the organ according to the classifier and at the same is near the estimated position relative to the other organs. This

box will not be the correct one since the organ is not present in the dataset. Such a case could be treated by evaluating the image characteristics of the final detected bounding box. In case the probability of this box is too low to be the organ, the correct bounding box could be marked red in order to be corrected or approved by the physician.

The extensibility of the presented methods is another important factor. In clinical application scenarios, new patients will be treated every day and several images will be generated that can be used to improve the existing training database. For example, after applying the model based segmentation methods, the physician could inspect the segmentation result and approve the shape to be added to the statistical shape model. Such a supervised strategy will improve the generalizability of the training shape set. The effort to include such new cases depends on the method. The statistical shape model can be easily extended. The training process of the initialization, however, may take some hours on very large training sets. Since this process is fully automatic, it can be performed e.g. once a day in the night.

Regarding the registration of organs, an improvement of the registration quality could be achieved for liver CT images that stem from the same contrast agent saturation phase. Here, the same vessel tree is contrasted which means additional information that can be used for matching the images. Based on the segmentation of the liver shapes, a segmentation of the vessel tree could be performed and branching points could be determined. These points could be used in the registration process as additional correspondences. This should improve the registration quality in the organ's interior, since vessel trees like the portal vein cover a large area inside the organ. However, since the contrast agent saturation will most likely not be identical in both images, methods are needed that can cope with different amounts of branching points. Additionally, other vessel trees may be slightly saturated in a single contrast phase and vessel segmentation may not be perfect for every case which also complicates correspondence establishment between images. Here, graph matching algorithms could be used to address these problems and to lead to a correspondence quality sufficient for registration needs.

An application scenario for the registration of multi-phase CT data that has not been addressed in this work is the detection of tissue abnormalities by incorporation of the contrast agent uptake value at a certain point in the organ. In the prostate, the contrast uptake ratio computed from different images of the same treatment is already a standard measure to detect certain cancer types. The registration of multi-phase liver CT data could be used to calculate an analogous measure for the detection of liver tumors.

The target imaging modality of this thesis is CT. However, many of the developed algorithms can be modified or directly used in other modalities like MRI or 3D ultrasound. Since the shape of an organ is image independent, all shape based methods do not need to be modified. Image based characteristics are used in the model initialization process as well as in the boundary search. In order to use the model initialization and boundary search for MRI images, a new training image database for MRI needs to be created. Moreover, the images need to be normalized, for example such that mean and variance of all images is the same. Additionally, bias field estimation methods could be used to correct irregular intensity distributions. In case of 3D ultrasound, intensity based training is likely to fail. Therefore, the initialization method could be modified to solely use the gradient part of the feature vector. That way only edges in the image would be considered for detection. Furthermore, in 3D ultrasound of large organs like the liver, it is likely that only parts of the organ are visible in the image. The position of model points that are

outside of the image could be left unchanged. Alternatively, an adaptation weight of zero could be assigned to the model points. In this case, the number of model points with a weight of non-zero needs to be larger than the number of columns in the system of linear equations that is built in the optimization.

6.2. Future work

The methods developed in this thesis build a basis towards a fully automatic segmentation and registration of the whole thoracic and abdominal area of the human body. Future work in this field points in several directions.

A general observation from reviewing the literature of the past two decades as well as from reviewing the results of this work is the improvement of segmentation and registration quality by incorporating prior knowledge. The first automatic approaches were mainly purely image based methods which did not take any prior shape or statistical knowledge into account. The accuracy of these approaches has usually been too low for clinical use. Nowadays, it is possible to model shape and appearance variability of complex organs like the liver accurately enough to allow an application in clinical practice. However, the quality of human cognition is still undisputed in many cases. Automatic methods often fail in the case of strong pathologies or cases that deviate a lot from trained instances. The reason for this is that most approaches target the segmentation or registration of a single structure. They therefore treat the problem in isolation and do not incorporate all information that is available in the image. In contrast, a physician always has knowledge about the anatomy of the whole body in mind which he can use to verify his decisions. For example, many low contrasted structures like lymph nodes, vessels or certain tumors are classified by exclusion and not because of prominent visual characteristics.

To this respect, a complete modeling of the whole body or at least a modeling of the area in which the target structure is embedded promises to further improve the quality of segmentation and registration methods. However, since the interrelation of soft tissue regarding shape, size and position in the human body is very complex, up till now, such solutions do not yet exist. In the latest literature, already some approaches have been proposed in the context of simultaneous multi-organ segmentation where the correlation of neighboring structures is incorporated into the segmentation process [CGHM11, SKWar]. These approaches are, however, either limited to rigid structures or only consider the interrelation of few selected organs. The extension of these methods to simultaneous segmentation of soft tissue organs like the liver, heart and spleen is promising to improve the segmentation quality in areas of low contrast, for example in the transition between liver and spleen.

A main problem for developing a method that accurately models the interrelation of many organs and other anatomical structures is the huge amount of training data that would be necessary to learn the interrelations. As it has been shown in this work, the anatomical variance of a single organ like the liver is hard to model accurately by statistical means alone — even if two hundred training images are incorporated. Furthermore, in the context of statistical shape modeling of a single structure, usually scale and rotational variance do not need to be modeled. This drastically reduces the degrees of freedom of the resulting model. In a complete multi-organ model, scale and rotational variance would have to be included in the model since every change in scale, rotation, position and shape influences the shape of the surrounding organs. Assuming only 20 structures are considered, a huge amount of training data would be needed to create an accurate model with such many degrees of freedom.

A solution could be to reduce the complexity of the modeling by building statistical meaningful groups of interrelations between structures. For example, the effect of deformation forces on an organ's shape caused by respiration could be modeled independently from the anatomical shape variance of organs between individuals. Furthermore, machine learning solutions already exist that are able to reveal statistical dependencies between the different degrees of freedom of the model. Such methods could be used to significantly reduce the complexity of the modeling problem. Furthermore, a stronger focus on building separate disease, age and sex specific models could reduce the amount of necessary training data and could lead to more robust and meaningful models. Still, a high amount of training data — probably in the scale of ten thousands or hundred thousands of images — would be needed for an accurate model. From the technical point of view, this problem could already be handled nowadays. Modern datacenters and computing clusters are able to provide the necessary storage space and computing power. Gathering such an amount of images and the cost involved are by far the bigger issues. However, given the possible benefits for society, these efforts seem to be worth it.

A. Publications and talks

The thesis is partially based on the following publications and talks of the author:

A.1. Publications

1. Erdt, Marius; Steger, Sebastian; Wesarg, Stefan: Deformable registration of MR images using a hierachical patch based approach with a normalized metric quality measure. In: IEEE Engineering in Medicine and Biology Society (EMBS): IEEE International Symposium on Biomedical Imaging: From Nano to Macro: ISBI 2012. New York: IEEE Press, 2012, to appear.
2. Erdt, Marius; Steger, Sebastian; Sakas, Georgios: Regmentation: A new view of image segmentation and registration. In: Journal of Radiation Oncology Informatics, Vol. 4, No. 1, 2012, pp. 1-23.
3. Erdt, Marius; Oyarzun Laura, Cristina; Drechsler, Klaus; De Beni, Stefano; Solbiati, Luigi: Improving diagnosis and intervention: A complete approach for registration of liver CT data. In: Yoshida H. (Ed.); Sakas G. (Ed.); Linguraru M. (Ed.); Abdominal Imaging: Computational and Clinical Applications, Berlin, Heidelberg, New York: Springer, 2012, pp. 108-115. (Lecture Notes in Computer Science (LNCS) 7029).
4. Oyarzun Laura, Cristina; Drechsler, Klaus; Erdt, Marius; Keil, Matthias; Noll, Matthias; De Beni, Stefano; Sakas, Georgios; Solbiati, Luigi: Intraoperative registration for liver tumor ablation. In: Yoshida H. (Ed.); Sakas G. (Ed.); Linguraru M. (Ed.); Abdominal Imaging: Computational and Clinical Applications, Berlin, Heidelberg, New York: Springer, 2012, pp. 133-140. (Lecture Notes in Computer Science (LNCS) 7029).
5. Erdt, Marius; Sakas, Georgios; Hammon, Matthias; De Beni, Stefano; Solbiati, Luigi; Cavallaro, Alexander: Automatic shape based deformable registration of multiphase contrast enhanced liver CT volumes. In: Dawant, Benoit M. (Ed.); Haynor, David R. (Ed.); The International Society for Optical Engineering (SPIE): Medical Imaging 2011: Image Processing. Part One: Progress in Biomedical Optics and Imaging. Proceedings Issue. Vol. 12, No. 31. Bellingham: SPIE Press, 2011, pp. 796212-1 – 796212-6. (Proceedings of SPIE 7962).
6. Wesarg, Stefan; Erdt, Marius; Kafchitsas, Konstantinos; Khan M. Fawad: Direct visualization of regions with lowered bone mineral density in dual-energy CT images of vertebrae. In: Summers, Ronald M. (Ed.); Ginneken, Bram van (Ed.); The International Society for Optical Engineering (SPIE): Medical Imaging 2011: Computer-Aided Diagnosis. Part Two: Progress in Biomedical Optics and Imaging. Proceedings Issue. Vol. 12, No. 32. Bellingham: SPIE Press, 2011, pp. 79633J-1 – 79633J-10. (Proceedings of SPIE 7963).
7. Erdt, Marius; Kirschner, Matthias; Drechsler, Klaus; Wesarg, Stefan; Hammon, Matthias; Cavallaro, Alexander: Automatic pancreas segmentation in contrast enhanced CT data using learned spatial anatomy and texture descriptors. In: IEEE Engineering in Medicine and Biology Society (EMBS): IEEE International Symposium on Biomedical Imaging: From Nano to Macro: ISBI 2011. New York: IEEE Press, 2011, pp. 2076-2082.

8. Steger, Sebastian; Ebert, Daniel; Erdt, Marius: Lymph node segmentation in CT slices using dynamic programming. In: IEEE Engineering in Medicine and Biology Society (EMBS): IEEE International Symposium on Biomedical Imaging: From Nano to Macro: ISBI 2011. New York: IEEE Press, 2011, pp. 1990-1993.
9. Wesarg, Stefan; Erdt, Marius; Kafchitsas, Konstantinos; Khan, M. Fawad: Quantifizierung und Visualisierung der Struktur des trabekulären Knochens in Wirbelkörpern. In: Handels, H. (Ed.) et al.; Gesellschaft für Informatik (GI): Bildverarbeitung für die Medizin 2011: Algorithmen, Systeme, Anwendungen. Berlin; Heidelberg; New York: Springer, 2011, pp. 309-313.
10. Wesarg, Stefan; Kafchitsas, Konstantinos; Erdt, Marius; Khan, M. Fawad: CAD of osteoporosis in vertebrae using dual-energy CT. In: Dillon, Tharam (Ed.) et al.; IEEE Computer Society Technical Committee on Computational Medicine: Twenty-Third IEEE Symposium on Computer-Based Medical Systems: CBMS 2010. Los Alamitos, Calif.: IEEE Computer Society, 2010, pp. 358-363.
11. Erdt, Marius; Sakas, Georgios: Computer aided segmentation of kidneys using locally shape constrained deformable models on CT Images. In: Karssemeijer, Nico (Ed.); Summers, Ronald M. (Ed.); The International Society for Optical Engineering (SPIE): Medical Imaging 2010: Computer-Aided Diagnosis: Progress in Biomedical Optics and Imaging. Proceedings Issue. Vol. 11, No. 34. Bellingham: SPIE Press, 2010, pp. 762419-1 – 762419-7. (Proceedings of SPIE 7624).
12. Erdt, Marius; Steger, Sebastian; Kirschner, Matthias; Wesarg, Stefan: Fast automatic liver segmentation combining learned shape priors with observed shape deviation. In: Dillon, Tharam (Ed.) et al.; IEEE Computer Society Technical Committee on Computational Medicine: Twenty-Third IEEE Symposium on Computer-Based Medical Systems: CBMS 2010. Los Alamitos, Calif.: IEEE Computer Society, 2010, pp. 249-254.
13. Steger, Sebastian; Erdt, Marius: Lymph node segmentation in CT images using a size invariant mass spring model. In: University of Ioannina: ITAB 2010: 10th International Conference on Information Technology and Applications in Biomedicine. New York: IEEE, Inc., 2010, 4 p.
14. Erdt, Marius; Schlegel, Patrice; Wesarg, Stefan: Multi-layer deformable models for medical image segmentation. In: University of Ioannina: ITAB 2010: 10th International Conference on Information Technology and Applications in Biomedicine. New York: IEEE, Inc., 2010, 4 p.
15. Drechsler, Klaus; Oyarzun Laura, Cristina; Chen, Yufei; Erdt, Marius: Semi-automatic anatomical tree matching for landmark-based elastic registration of liver volumes. In: Journal of Healthcare Engineering. 1 (2010), 1, pp. 101-123.
16. Erdt, Marius; Kirschner, Matthias; Wesarg, Stefan: Smart manual landmarking of organs. In: Dawant, Benoit M. (Ed.); Haynor, David R. (Ed.); The International Society for Optical Engineering (SPIE): Medical Imaging 2010: Image Processing. Part One: Progress in Biomedical Optics and Imaging. Proceedings Issue. Vol. 11, No. 33. Bellingham: SPIE Press, 2010, pp. 76234Y-1 – 76234Y-9. (Proceedings of SPIE 7623).
17. Graf, Norbert; Hoppe, Alexander; Georgiadi, Eleni; Belleman, Robert; Desmedt, Christine; Dionysiou, Dimitra; Erdt, Marius; Jacques, Julien; Kolokotroni, Eleni; Lunzer, Aran; Tsiknakis, Manolis; Stamatakos, Georgios: 'In silico' oncology for clinical decision making in the context of nephroblastoma. In: Klinische Pädiatrie. (2009), 221, pp. 141-149.

18. Erdt, Marius; Tulchiner, Roman; Sakas, Georgios: Erweiterung modellbasierter Segmentierung durch lokale Deformationskriterien. In: Meinzer, Hans-Peter (Ed.) et al.; Gesellschaft für Informatik (GI): Bildverarbeitung für die Medizin 2009. Proceedings: Algorithmen, Systeme, Anwendungen. Berlin; Heidelberg; New York: Springer, 2009, pp. 326-330. (Informatik aktuell).
19. Steger, Sebastian; Erdt, Marius; Chiari, Gianfranco; Sakas, Georgios: Feature extraction from medical images for an oral cancer reoccurrence prediction environment. In: Dössel, Olaf (Ed.); Schlegel, Wolfgang (Ed.): International Federation for Medical and Biological Engineering (IFMBE): World Congress on Medical Physics and Biomedical Engineering 2009. Proceedings DVD-ROM. Berlin, Heidelberg, New York: Springer, 2009, IFMBE 25/V; pp. 97-100. (IFMBE Proceedings).
20. Erdt, Marius; Kirschner, Matthias; Wesarg, Stefan: Simultaneous segmentation and correspondence establishment for statistical shape models. In: Magnenat-Thalmann, Nadia (Ed.): Modelling the Physiological Human: 3D Physiological Human Workshop. Proceedings. Berlin, Heidelberg, New York: Springer, 2009, pp. 25-35. (Lecture Notes in Computer Science (LNCS) 5903).
21. Erdt, Marius: Segmentation, registration, and fusion of medical images. In: Stergiopoulos, Stergios (Ed.); Advanced Signal Processing: Theory and Implementation for Sonar, Radar, and Non-Invasive Medical Diagnostic Systems. Boca Raton: Taylor & Francis, CRC Press, 2009, pp. 251-275. (The Electrical Engineering and Signal Processing Series).
22. Erdt, Marius; Raspe, Matthias; Sühling, Michael: Automatic hepatic vessel segmentation using graphics hardware. In: Dohi, Takeyoshi (Ed.); Sakuma, Ichiro (Ed.); Liao, Hongen (Ed.): Medical Imaging and Augmented Reality. Proceedings: 4th International Workshop. Berlin; Heidelberg; New York: Springer, 2008, pp. 403-412. (Lecture Notes in Computer Science (LNCS) 5128).
23. Kaftan, Jens N.; Kiraly, Atilla P.; Erdt, Marius; Sühling, Michael; Aach, Til: Fuzzy pulmonary vessel segmentation using optimized vessel enhancement filtering. In: Brown, Matthew (Ed.) et al.: The First International Workshop on Pulmonary Image Analysis. New York, 2008, pp. 233-242.
24. Erdt, Marius; Frank, Katrin; Henrich, Niklas; Jager, Fabian; Ohler, Patrick; Müller, Stefan: Towars: A mobile ar multi-player game. In: Müller, Stefan (Ed.); Zachmann, Gabriel (Ed.): Virtuelle und Erweiterte Realität. Shaker, Aachen, 2006, pp. 215-223.

A.2. Talks

1. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2012: Deformable registration of MR images using a hierachical patch based approach with a normalized metric quality measure. Barcelona, Spain, 2012.
2. MICCAI Workshop on Computational and Clinical Applications in Abdominal Imaging 2011: Improving Diagnosis and Intervention: A Complete Approach for Registration of Liver CT Data, Toronto, Canada, 2011.
3. Deutscher Röntgenkongress 2011: Automatische Segmentierung des Pankreas in kontrastmittelverstärkten CT-Daten. Hamburg, Germany, 2011.
4. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2011: Automatic Pancreas Segmentation in Contrast Enhanced CT Data Using Learned Spatial Anatomy and Texture Descriptors. Chicago, USA, 2011.
5. SPIE Medical Imaging: Automatic shape based deformable registration of multiphase contrast enhanced liver CT volumes. Orlando, USA, 2011.
6. European Congress of Radiology, ECR 2011: Automatic segmentation of the pancreas in contrast enhanced CT data. Vienna, Austria, 2011.
7. 23rd IEEE Symposium on Computer-Based Medical Systems, CBMS 2010: Fast Automatic Liver Segmentation Combining Learned Shape Priors with Observed Shape Deviation. Perth, Australia, 2010.
8. 10th IEEE International Conference on Information Technology and Applications in Biomedicine: Multi-Layer Deformable Models for Medical Image Segmentation. Corfu, Greece, 2010.
9. SPIE Medical Imaging: Computer Aided Segmentation of Kidneys Using Locally Shape Constrained Deformable Models on CT Images. San Diego, USA, 2010.
10. 3D Physiological Human Workshop: Simultaneous Segmentation and Correspondence Establishment for Statistical Shape Models. Zermatt, Swiss, 2009.
11. Medical Imaging and Augmented Reality: Automatic Hepatic Vessel Segmentation Using Graphics Hardware. Tokyo, Japan, 2008.

B. Supervising activities

The following list summarizes the student bachelor, Diplom and master theses supervised by the author in descending order.

B.1. Diplom and master theses

1. Automatische lernbasierte Landmarkendetektion in medizinischen Bilddaten. Brauer, Achim; Erdt, Marius (supervisor). Master thesis, University of Applied Sciences Darmstadt, Darmstadt, 2011.
2. Enhancement of deformable model based segmentation through a 3D layer model. Schlegel, Patrice; Erdt, Marius (supervisor). Master thesis, University of applied sciences Gießen-Friedberg, Gießen, 2009.
3. Atlasbasierte Segmentierung mittels elastischer Registrierung. Schmidt, Carola; Erdt, Marius (supervisor). Diplom thesis, University of Koblenz-Landau, Koblenz, 2009.
4. Volumetrische Organ-Segmentierung mithilfe von Semantically Shape Constrained Deformable Models. Tulchiner, Roman; Erdt, Marius (supervisor). Master thesis, TU-Darmstadt, Darmstadt, 2008.

B.2. Bachelor theses

1. Entwicklung von Boundary Detektoren für die Segmentierung von medizinischen Bilddaten. Gimbel, Stephan; Erdt, Marius (supervisor). Bachelor thesis, University of Applied Sciences Darmstadt, Darmstadt, 2009.
2. Entwicklung und Implementierung eines Verfahrens zur Lokalisierung und Registrierung von medizinischen Bilddaten. Yalnizer, Nilay; Erdt, Marius (supervisor). Bachelor thesis, University of Applied Sciences Darmstadt, Darmstadt, 2009.

C. Curriculum vitae

Personal Data

Name	Marius Erdt
Birth date & place	20.05.1980 in Hamburg
Family status	Single
Nationality	German

Education

2007	Graduation in Computational Visualistics (Dipl.-Inform.) at University of Koblenz-Landau in Koblenz, Germany
2007	Diplom thesis at Siemens Healthcare in Forchheim, Germany
2002 – 2007	Study at University Koblenz-Landau in Koblenz, Germany
2000 – 2002	Study at University Stuttgart in Stuttgart, Germany
1999 – 2000	Military service at German navy in Eckernförde and in Flensburg, Germany
1999	Abitur at Internatsgymnsasium Pädagogium Bad Sachsa in Bad Sachsa, Germany

Professional Experience

2007 – present	Researcher at the Cognitive Computing & Medical Imaging department at Fraunhofer IGD in Darmstadt, Germany.
2007	Diplom student at Siemens Healthcare in Forchheim, Germany

Bibliography

[AFAX11] ASLAN M. S., FARAG A. A., ARNOLD B., XIANG P.: Segmentation of vertebrae using level sets with expectation maximization algorithm. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 2010–2013. 24, 43

[AHH*09] ALJABAR P., HECKEMANN R., HAMMERS A., HAJNAL J., RUECKERT D.: Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage 46*, 3 (2009), 726–738. 27

[AHS10] ANDREWS S., HAMARNEH G., SAAD A.: Fast random walker with priors using precomputation for interactive medical image segmentation. In *Medical image computing and computer-assisted intervention* (Berlin, Heidelberg, 2010), Lecture Notes in Computer Science, Springer, pp. 9–16. 23

[APG11] ABABNEH S. Y., PRESCOTT J. W., GURCAN M. N.: Automatic graph-cut based segmentation of bones from knee magnetic resonance images for osteoarthritis research. *Medical Image Analysis 15*, 4 (2011), 438–448. 23

[BFOS84] BREIMAN L., FRIEDMAN J., OLSHEN R., STONE C.: *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984. 38, 39, 84

[Bis06] BISHOP C. M.: *Pattern Recognition and Machine Learning*. Springer Science+Business Media, 2006. 35, 40

[BM92] BESL P. J., MCKAY N. D.: A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 2 (1992), 239–256. 102, 113

[BM93] BEUCHER S., MEYER F.: *The Morphological Approach to Segmentation : The Watershed Transformation*. Marcel Dekker, New York, 1993, ch. 12, pp. 433–481. 19

[BM07] BLEZEK D. J., MILLER J. V.: Atlas stratification. *Medical Image Analysis 11*, 5 (2007), 443–457. 27

[Bn95] BRO-NIELSEN M.: Modelling elasticity in solids using active cubes - application to simulated operations. In *Computer Vision, Virtual Reality and Robotics in Medicine* (1995), pp. 535–541. 45, 46

[BP05] BARREIRA N., PENEDO M. G.: Topological active volumes. *EURASIP Journal on Applied Signal Processing 2005* (2005), 1939–1947. 24, 44, 46

[BPCO10] BARREIRA N., PENEDO M. G., COHEN L., ORTEGA M.: Topological active volumes: A topology-adaptive deformable model for volume segmentation. *Pattern Recogn. 43* (2010), 255–266. ix, 2, 26, 35, 46, 51, 52, 53, 71

[BSX*10] BARBU A., SUEHLING M., XU X., LIU D., ZHOU S., COMANICIU D.: Automatic detection and segmentation of axillary lymph nodes. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2010), vol. 6361 of *Lecture Notes in Computer Science*, Springer, pp. 28–36. 23

[Can86] CANNY J.: A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-8*, 6 (1986), 679 –698. 20

[Cas95] CASELLES V.: Geometric models for active contours. In *International Conference on Image Processing* (Washington, DC, 1995), vol. 3, IEEE Computer Society, pp. 9–12. 43

[CBA*03] COLLIER D. C., BURNETT S. S. C., AMIN M., BILTON S., BROOKS C., RYAN A., RONIGER D., TRAN D., STARKSCHALL G.: Assessment of consistency in contouring of normal-tissue anatomic structures. *Journal of applied clinical medical physics 4*, 1 (2003), 17–24. vii, 1

[CBMS11] CHI J. W., BRADY M., MOORE N., SCHNABEL J. A.: Segmentation of the bladder wall using coupled level set methods. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 1653–1656. 24, 43

[CC07] CAMPADELLI P., CASIRAGHI E.: Liver segmentation from CT scans: A survey. In *Applications of Fuzzy Sets Theory*. 2007, pp. 520–528. 33

[CCBK07] CHI Y., CASHMAN P. M. M., BELLO F., KITNEY R. I.: A discussion on the evaluation of a new automatic liver volume segmentation method for specified CT image datasets. In *MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge* (2007), pp. 167–175. 24, 43

[CDA07] COSTA M., DELINGETTE H., AYACHE N.: Automatic segmentation of the bladder using deformable models. In *Biomedical Imaging: From Nano to Macro, 2007 IEEE International Symposium on* (2007), pp. 252–260. 43

[CDNA07] COSTA J., DELINGETTE H., NOVELLAS S., AYACHE N.: Automatic segmentation of bladder and prostate using coupled 3D deformable models. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2007), Lecture Notes in Computer Science, Springer, pp. 252–260. 24, 44, 70, 133

[CGHM11] CHEN X., GRAHAM J., HUTCHINSON C., MUIR L.: Inferring 3D kinematics of carpal bones from single view fluoroscopic sequences. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2011), Fichtinger G., Martel A., Peters T., (Eds.), vol. 6892 of *Lecture Notes in Computer Science*, Springer, pp. 680–687. 145

[CGM08] COMMOWICK O., GRÉGOIRE V., MALANDAIN G.: Atlas-based delineation of lymph node levels in head and neck computed tomography images. *Radiotherapy Oncology 87*, 2 (May 2008), 281–289. 27, 28

[CHH04] CRUM W. R., HARTKENS T., HILL D. L. G.: Non-rigid image registration: theory and practice. *The British Journal of Radiology 77*, 2 (Dec. 2004), 140–153. 11, 12

[CHTH93] COOTES T. F., HILL A., TAYLOR C. J., HASLAM J.: The use of active shape models for locating structures in medical images. In *Information Processing in Medical Imaging* (1993), pp. 33–47. ix, 2, 26, 47

[CM06] COMMOWICK O., MALANDAIN G.: Evaluation of atlas construction strategies in the context of radiotherapy planning. In *SA2PM Workshop (From Statistical Atlases to Personalized Models)* (Oct. 2006), pp. 1–4. 27

[CM07] COMMOWICK O., MALANDAIN G.: Efficient selection of the most similar image in a database for critical structures segmentation. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2007), vol. 4792 of *Lecture Notes in Computer Science*, Springer, pp. 203–210. 27

[Coo01] COOTES T. F.: Statistical models of appearance for medical image analysis and computer vision. *SPIE Medical Imaging 4322* (2001), 236–248. 122

[CRD07] CREMERS D., ROUSSON M., DERICHE R.: A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *International Journal of Computer Vision 72*, 2 (Apr. 2007), 195–215. 43

[CT93] COOTES T. F., TAYLOR C. J.: Active shape model search using local grey-level models: A quantitative evaluation. In *4th British Machine Vision Conference* (1993), pp. 639–648. 3, 50, 52

[CT04] COOTES T. F., TAYLOR C. J.: *Statistical Models of Appearance for Computer Vision*. Tech. rep., University of Manchester, 2004. 51, 52

[CTCG95] COOTES T., TAYLOR C., COOPER D., GRAHAM J.: Active shape models - their training and application. *Computer Vision and Image Understanding 61*, 1 (1995), 38–59. ix, 2, 26, 47, 65, 76, 91

[CW11] CAO H., WANG Y.-P.: Segmentation of m-fish images for improved classification of chromosomes with an adaptive fuzzy c-means clustering algorithm. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 1442 –1445. 14

[DMJ09] DAS I. J., MOSKVIN V., JOHNSTONE P. A.: Analysis of treatment planning time among systems and planners for intensity-modulated radiation therapy. *Journal of the American College of Radiology 6*, 7 (2009), 514 – 517. vii, 1

[DOL10] DRECHSLER K., OYARZUN LAURA C.: Hierachical decomposition of vessel skeletons for graph creation and feature extraction. In *IEEE International Conference on Bioinformatics & Biomedicine (BIBM)* (2010), pp. 456–461. 99

[DOLCE10] DRECHSLER K., OYARZUN LAURA C., CHEN Y., ERDT M.: Semi-automatic anatomical tree matching for landmark-based elastic registration of liver volumes. *Journal of Healthcare Engineering 1*, 1 (2010), 101–123. 4

[DSP*06] DORNHEIM J., SEIM H., PREIM B., HERTEL I., STRAUSS G.: Segmentation of neck lymph nodes in CT datasets with stable 3D mass-spring models. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2006), Lecture Notes in Computer Science, Springer, pp. 904–911. 25

[DT05] DALAL N., TRIGGS B.: Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition* (june 2005), pp. 886–893. 82

[DTC*02] DAVIES R. H., TWINING C. J., COOTES T. F., WATERTON J. C., TAYLOR C. J.: A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging 21*, 5 (2002), 525–537. 49, 65

[DTT08] DAVIES R., TWINING C., TAYLOR C.: *Statistical Models of Shape - Optimization and Evaluation*. Springer Verlag, 2008. 3, 41, 47, 65, 75, 91, 113, 114

[DvCE*10] DZYUBACHYK O., VAN CAPPELLEN W., ESSERS J., NIESSEN W., MEIJERING E.: Advanced level-set-based cell tracking in time-lapse fluorescence microscopy. *Medical Imaging, IEEE Transactions on 29*, 3 (march 2010), 852 –867. 23

[EBK*10] EGGER J., BAUER M., KUHNT D., CARL B., KAPPUS C., FREISLEBEN B., NIMSKY C.: Nugget-cut: A segmentation scheme for spherically- and elliptically-shaped 3D objects. In *Pattern Recognition*, vol. 6376. 2010, pp. 373–382. 23

[EK07] ERICSSON A., KARLSSON J.: Measures for benchmarking of automatic correspondence algorithms. *Journal of Mathematical Imaging and Vision 28*, 3 (2007), 225–

241. 65

[EKD*11] ERDT M., KIRSCHNER M., DRECHSLER K., WESARG S., HAMMON M., CAVALLARO A.: Automatic pancreas segmentation in contrast enhanced CT data using learned spatial anatomy and texture descriptors. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 2076–2082. 4, 97, 98, 102, 134, 136

[EKW09] ERDT M., KIRSCHNER M., WESARG S.: Simultaneous segmentation and correspondence establishment for statistical shape models. In *Modelling the Physiological Human* (Berlin, Heidelberg, New York, 2009), vol. 5903 of *Lecture Notes in Computer Science*, Springer, pp. 25–35. xii, 4, 62, 66, 67, 68, 86, 111, 112

[EKW10] ERDT M., KIRSCHNER M., WESARG S.: Smart manual landmarking of organs. In *Medical Imaging 2010: Image Processing. Part One* (2010), Proceedings of SPIE; 7623, SPIE Press, Bellingham, pp. 76234Y–1–76234Y–9. 4, 63, 70, 86, 115, 116

[EOLD*12] ERDT M., OYARZUN LAURA C., DRECHSLER K., DE BENI S., SOLBIATI L.: Improving diagnosis and intervention: A complete approach for registration of liver CT data. In *Abdominal Imaging. Computational and Clinical Applications* (Berlin, Heidelberg, 2012), Yoshida H., Sakas G., Linguraru M., (Eds.), vol. 7029 of *Lecture Notes in Computer Science*, Springer, pp. 108–115. 4, 107, 108

[EPS*08] ECABERT O., PETERS J., SCHRAMM H., LORENZ C., VON BERG J., WALKER M. J., VEMBAR M., OLSZEWSKI M. E., SUBRAMANYAN K., LAVI G., WEESE J.: Automatic model-based segmentation of the heart in CT images. *IEEE Trans. Med. Imaging 27*, 9 (2008), 1189–1201. viii, 2, 24, 26, 44, 45, 50, 52, 61, 70, 76, 78, 91

[Erd09] ERDT M.: *Segmentation, Registration, and Fusion of Medical Images*, 2nd ed. The Electrical Engineering and Signal Processing Series. Taylor and Francis, CRC Press, Boca Raton, 2009, pp. 251–275. 11, 16, 54

[ERS08] ERDT M., RASPE M., SUEHLING M.: Automatic hepatic vessel segmentation using graphics hardware. In *Medical Imaging and Augmented Reality. Proceedings* (Berlin, Heidelberg, New York, 2008), vol. 5128 of *Lecture Notes in Computer Science*, Springer, pp. 403–412. 19, 20, 21, 54, 98, 99

[ES10] ERDT M., SAKAS G.: Computer aided segmentation of kidneys using locally shape constrained deformable models on CT images. In *Medical Imaging: Computer-Aided Diagnosis* (2010), Proceedings of SPIE; 7624, SPIE Press, Bellingham, pp. 762419–1–762419–7. 4, 86, 88, 90, 91, 108, 126, 128, 139

[ESH*11] ERDT M., SAKAS G., HAMMON M., DE BENI S., SOLBIATI L., CAVALLARO A.: Automatic shape based deformable registration of multiphase contrast enhanced liver CT volumes. In *Medical Imaging: Image Processing. Part One* (2011), Proceedings of SPIE; 7962, SPIE Press, Bellingham, pp. 796212–1–796212–6. 4, 105, 108, 137, 138, 139

[ESKW10] ERDT M., STEGER S., KIRSCHNER M., WESARG S.: Fast automatic liver segmentation combining learned shape priors with observed shape deviation. In *Twenty-Third IEEE Symposium on Computer-Based Medical Systems* (2010), IEEE Computer Society, Los Alamitos, Calif., pp. 249–254. 4, 75, 86, 92, 93, 108, 119, 129, 130, 131, 139

[ESS12] ERDT M., STEGER S., SAKAS G.: Regmentation: A new view of image segmentation and registration. *JROI-Journal of Radiation Oncology Informatics 4*, 1 (2012), 1–23. 15, 19, 23, 24, 25, 27, 29, 32, 34, 54

[ESW10] ERDT M., SCHLEGEL P., WESARG S.: Multi-layer deformable models for medical image segmentation. In *10th International Conference on Information Technology and Applications in Biomedicine* (2010), IEEE, Inc., New York, p. 4. xii, 4, 64, 71, 72, 73, 76, 86, 108, 132, 133, 139

[ETS09] ERDT M., TULCHINER R., SAKAS G.: Erweiterung modellbasierter Segmentierung durch lokale Deformationskriterien. In *Bildverarbeitung für die Medizin 2009. Proceedings* (Berlin, Heidelberg, New York, 2009), Informatik aktuell, Gesellschaft für Informatik (GI), Springer, pp. 326–330. 4, 86, 108, 139

[FEK*11] FIRJANI A., ELNAKIB A., KHALIFA F., GIMEL'FARB G. L., EL-GHAR M. A., SURI J., ELMAGHRABY A., EL-BAZ A.: A new 3D automatic segmentation framework for accurate segmentation of prostate from DCE-MRI. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 1476–1479. 23, 24

[FH06] FELZENSZWALB P. F., HUTTENLOCHER D. P.: Efficient belief propagation for early vision. *International Journal of Computer Vision 70*, 1 (2006), 41–54. 101

[FKE*10] FREIMAN M., KRONMAN A., ESSES S. J., JOSKOWICZ L., SOSNA J.: Non-parametric iterative model constraint graph min-cut for automatic kidney segmentation. In *Medical image computing and computer-assisted intervention* (Berlin, Heidelberg, 2010), Lecture Notes in Computer Science, Springer, pp. 73–80. 23, 24

[FNVV98] FRANGI A. F., NIESSEN W. J., VINCKEN K. L., VIERGEVER M. A.: Multiscale vessel enhancement filtering. In *Medical image computing and computer-assisted intervention* (Berlin, Heidelberg, 1998), Wells W. M., Colchester A. C. F., Delp S. L., (Eds.), vol. 1496 of *Lecture Notes in Computer Science*, Springer, pp. 130–137. 20

[FS96] FREUND Y., SCHAPIRE R. E.: Experiments with a new boosting algorithm. In *Thirteenth International Conference on Machine Learning* (1996), pp. 148–156. 40, 84

[FSK07] FURUKAWA D., SHIMIZU A., KOBATAKE H.: Automatic liver segmentation method based on maximum a posterior probability estimation and level set method. In *MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge* (2007), pp. 117–124. 24, 43

[FWK*08] FRANZ A., WOLZ R., KLINDER T., LORENZ C., BARSCHDORF H., BLAFFERT T., DRIES S. P. M., RENISCH S.: Simultaneous model-based segmentation of multiple objects. In *Bildverarbeitung für die Medizin* (2008), pp. 252–256. 45

[GCN08] GHANEH P., COSTELLO E., NEOPTOLEMOS J. P.: Biology and management of pancreatic cancer. *Postgraduate Medical Journal 84*, 995 (2008), 478–497. 96

[GGCH10] GRAHAM M., GIBBS J., CORNISH D., HIGGINS W.: Robust 3-d airway tree segmentation for image-guided peripheral bronchoscopy. *Medical Imaging, IEEE Transactions on 29*, 4 (april 2010), 982 –997. 23

[GPS89] GREIG D., PORTEOUS B., SEHEULT A.: Exact maximum a posteriori estimation for binary images. *Royal Journal on Statistical Society 51*, 2 (1989), 271–279. 23

[GT11] GAO Y., TANNENBAUM A.: Combining atlas and active contour for automatic 3D medical image segmentation. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 1401–1404. 28

[HGM09] HU A., GROSSBERG B., MAGERAS C.: Survey of recent volumetric medical image segmentation techniques. *Biomedical Engineering* (2009), 321–346. 11, 12, 13, 14, 20

[HM08] HUANG X., METAXAS D.: Metamorphs: Deformable shape and appearance models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 30*, 8 (2008), 1444–1459. ix, 2, 26, 35, 46, 47, 51, 52, 53, 71

[HM09] HEIMANN T., MEINZER H.-P.: Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis 13*, 4 (2009), 543–563. viii, ix, 2, 26, 47, 65, 69, 74, 113, 133

[HMC04] HUANG X., METAXAS D., CHEN T.: Metamorphs: Deformable shape and texture models. In *Computer Vision and Pattern Recognition* (2004), vol. 1, pp. 496–503. 46

[HMMW07] HEIMANN T., MÜNZING S., MEINZER H.-P., WOLF I.: A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. In *Information Processing in Medical Imaging* (2007), pp. 1–12. 133

[HMW07] HEIMANN T., MEINZER H.-P., WOLF I.: A statistical deformable model for the segmentation of liver CT volumes. In *MICCAI 2007 Workshop: 3D Segmentation in the Clinic: A Grand Challenge* (2007), pp. 161–166. 24, 26, 44, 51, 52, 74, 77, 129, 130

[Hor87] HORN B. K. P.: Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A: Optics, Image Science, and Vision 4*, 4 (1987), 629–624. 102

[HUKE10] HAMAMCI A., UNAL G., KUCUK N., ENGIN K.: Cellular automata segmentation of brain tumors on post contrast MR images. In *Medical image computing and computer-assisted intervention* (Berlin, Heidelberg, 2010), Lecture Notes in Computer Science, Springer, pp. 137–146. 23

[HvGSea09] HEIMANN T., VAN GINNEKEN B., STYNER M., ET AL.: Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Transactions on Medical Imaging 28*, 8 (2009), 1251–1265. viii, ix, 2, 33, 43, 44, 53, 60, 61, 70, 71, 74, 129, 130, 139

[HWC*08] HUANG X., WANG B., CHENG M., HUANG S., JU Y.: Image registration and data fusion for different phases of contrast enhanced liver CT data. In *Bioinformatics and Biomedical Engineering* (2008), pp. 2682–2685. 103

[HWM06] HEIMANN T., WOLF I., MEINZER H.-P.: Optimal landmark distributions for statistical shape model construction. In *Proc. SPIE Medical Imaging 2006: Image Processing* (2006), Reinhardt J., Pluim J., (Eds.), vol. 6144, pp. 518–528. 114, 115

[HZ09] HELDMANN S., ZIDOWITZ S.: Elastic registration of multiphase CT images of liver. In *Medical Imaging: Image Processing* (2009), Pluim J. P. W., Dawant B. M., (Eds.), vol. 7259, SPIE, p. 72591H. 103

[IBSP09] IBÁÑEZ O., BARREIRA N., SANTOS J., PENEDO M.: Genetic approaches for topological active nets optimization. *Pattern Recognition 42*, 5 (2009), 907–917. 46, 52

[IDB*08] ISAMBERT A., DHERMAIN F., BIDAULT F., COMMOWICK O., BONDIAU P.-Y., MALANDAIN G., LEFKOPOULOS D.: Evaluation of an atlas-based automatic segmentation software for the delineation of brain organs at risk in a radiation therapy clinical

context. *Radiotherapy Oncology 87*, 1 (Apr. 2008), 93–99. 28

[JPW*10] JURRUS E., PAIVA A. R. C., WATANABE S., ANDERSON J. R., JONES B. W., WHITAKER R. T., JORGENSEN E. M., MARC R. E., TASDIZEN T.: Detection of neuron membranes in electron microscopy images using a serial neural network architecture. *Medical Image Analysis* (June 2010), 770–783. 21

[Kas80] KASS G. V.: An exploratory technique for investigating large quantities of categorical data. *Applied Statistics 29*, 2 (1980), 119–127. 38

[KBW11] KIRSCHNER M., BECKER M., WESARG S.: 3D active shape model segmentation with nonlinear shape priors. In *Medical Image Computing and Computer-Assisted Intervention* (2011), vol. 6892 of *Lecture Notes in Computer Science*, Springer, pp. 492–499. 69, 124

[KCBP04] KOEHLER H., COUPRIE M., BOUATTOUR S., PAULUS D.: Extraction and analysis of coronary tree from single X-ray angiographies. In *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display. Proceedings of the SPIE* (May 2004), Galloway Jr. R. L., (Ed.), vol. 5367, pp. 810–819. 21

[KKE*08] KAFTAN J. N., KIRALY A. P., ERDT M., SUEHLING M., AACH T.: Fuzzy pulmonary vessel segmentation using optimized vessel enhancement filtering. In *The First International Workshop on Pulmonary Image Analysis* (2008), pp. 233–242. 20

[KLL07] KAINMUELLER D., LANGE T., LAMECKER H.: Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In *MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge* (2007), pp. 109–116. viii, 2, 24, 26, 44, 50, 52, 61, 75, 76, 94, 129, 130, 133

[KSF*10] KARIMAGHALOO Z., SHAH M., FRANCIS S., ARNOLD D., COLLINS D., ARBEL T.: Detection of gad-enhancing lesions in multiple sclerosis using conditional random fields. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2010), Jiang T., Navab N., Pluim J., Viergever M., (Eds.), vol. 6363 of *Lecture Notes in Computer Science*, Springer, pp. 41–48. 23

[KSM*08] KITASAKA T., SAKASHITA M., MORI K., SUENAGA Y., NAWANO S.: A method for extracting pancreas regions from four-phase contrasted 3D abdominal CT images. *International Journal of Computer Assisted Radiology and Surgery 3* (2008), 40. 97, 134

[KT97] KOTCHEFF A. C. W., TAYLOR C. J.: Automatic construction of eigenshape models by genetic algorithm. In *Information Processing in Medical Imaging* (1997), pp. 1–14. 49, 65

[KT98] KOTCHEFF A. C. W., TAYLOR C. J.: Automatic construction of eigenshape models by direct optimization. *Medical Image Analysis 2* (1998), 303–314. 113

[KUA*09] KOHLBERGER T., UZUNBAS M. G., ALVINO C. V., KADIR T., SLOSMAN D. O., FUNKA-LEA G.: Organ segmentation with level sets using local shape and appearance priors. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2009), Lecture Notes in Computer Science, Springer, pp. 34–42. 24, 43, 74, 129, 130

[KW10] KIRSCHNER M., WESARG S.: Construction of groupwise consistent shape parameterizations by propagation. In *Medical Imaging: Image Processing* (2010), Proceedings of SPIE; 7623, SPIE Press, Bellingham, pp. 762352–1–762352–12. 113, 114

[KW11] KIRSCHNER M., WESARG S.: Active shape models unleashed. In *Medical Imaging 2011: Image Processing. Part One* (2011), vol. 7962 of *Proceedings of SPIE*, The International Society for Optical Engineering (SPIE), SPIE Press, Bellingham, pp. 796211–1–796211–9. 102, 124

[KYLL08] KWON D., YUN I. D., LEE K. H., LEE S. U.: Efficient feature-based nonrigid registration of multiphase liver CT volumes. In *British Machine Vision Conference* (2008), Everingham M., Needham C. J., Fraile R., (Eds.), British Machine Vision Association. 103

[LB06] LORENZ C., BERG J.: A comprehensive shape model of the heart. *Medical image analysis 10*, 4 (2006), 657–670. 68, 77, 78, 79

[LC87] LORENSEN W. E., CLINE H. E.: Marching cubes: A high resolution 3D surface construction algorithm. In *SIGGRAPH* (1987), pp. 163–169. 62, 102, 113

[LLH06] LIN D., LEI C., HUNG S.: Computer-aided kidney segmentation on abdominal CT images. *IEEE transactions on information technology in biomedicine 10*, 1 (2006), 59–65. 88, 127

[LS97] LOH W.-Y., SHIH Y.-S.: Split selection methods for classification trees. *Statistica Sinica 7* (1997), 815–840. 38

[LSCO*04] LIPMAN Y., SORKINE O., COHEN-OR D., LEVIN D., RÖSSL C., SEIDEL H.-P.: Differential coordinates for interactive mesh editing. In *Shape Modeling International* (2004), IEEE Computer Society Press, pp. 181–190. 79

[LZZ*08] LING H., ZHOU S., ZHENG Y., GEORGESCU B., SUEHLING M., COMANICIU D.: Hierarchical, learning-based automatic liver segmentation. In *Computer Vision and Pattern Recognition* (2008), pp. 1–8. 24, 26, 44, 129, 130

[MBG*10] MURALIDHAR G., BOVIK A., GIESE J., SAMPAT M., WHITMAN G., HAYGOOD T., STEPHENS T., MARKEY M.: Snakules: A model-based active contour algorithm for the annotation of spicules on mammography. *Medical Imaging, IEEE Transactions on 29*, 10 (oct. 2010), 1768 –1780. 22

[Mey94] MEYER F.: Topographic distance and watershed lines. *Signal Processing 38*, 1 (July 1994), 113–125. 19

[MHR10] MESSAY T., HARDIE R. C., ROGERS S. K.: A new computationally efficient CAD system for pulmonary nodule detection in CT imagery. *Medical Image Analysis 14*, 3 (2010), 390 – 406. 21

[MSN*10] MAGJAREVIC R., SHIMIZU A., NAWANO S., SHINOZAKI K., TATENO Y.: Medical image processing competition in Japan. In *World Congress on Medical Physics and Biomedical Engineering* (Berlin, Heidelberg, 2010), vol. 25/4 of *IFMBE Proceedings*, Springer, pp. 1814–1817. 97

[MST10] MOHAN V., SUNDARAMOORTHI G., TANNENBAUM A.: Tubular surface segmentation for extracting anatomical structures from medical imagery. *Medical Imaging, IEEE Transactions on 29*, 12 (Dec. 2010), 1945 –1958. 23

[MSV94] MALLADI R., SETHIAN J. A., VEMURI B. C.: Evolutionary fronts for topology-independent shape modeling and recovery. In *European conference on Computer vision* (New York, 1994), Springer, pp. 3–13. 43

[MSV95] MALLADI R., SETHIAN J. A., VEMURI B. C.: Shape modeling with front propagation: a level set approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 17*, 2 (Feb. 1995), 158–175. 43

[MT95] MCINERNEY T., TERZOPOULOS D.: Topologically adaptable snakes. In *Medical Image Analysis* (1995), pp. 840–845. 22

[MT96] MCINERNEY T., TERZOPOULOS D.: Deformable models in medical image analysis: a survey. *Medical Image Analysis 1*, 2 (1996), 91–108. 12, 22, 25, 44

[MV98] MAINTZ J., VIERGEVER M.: A survey of medical image registration. *Medical Image Analysis 2*, 1 (1998), 1–36. 11, 12

[MVN06] MANNIESING R., VIERGEVER M. A., NIESSEN W. J.: Vessel enhancing diffusion: A scale space representation of vessel structures. *Medical Image Analysis 10* (2006), 815–825. 20

[OLDE*12] OYARZUN LAURA C., DRECHSLER K., ERDT M., KEIL M., NOLL M., DE BENI S., SAKAS G., SOLBIATI L.: Intraoperative registration for liver tumor ablation. In *Abdominal Imaging. Computational and Clinical Applications* (Berlin, Heidelberg, 2012), Yoshida H., Sakas G., Linguraru M., (Eds.), vol. 7029 of *Lecture Notes in Computer Science*, Springer, pp. 133–140. 108

[OSS*07] OKADA T., SHIMADA R., SATO Y., HORI M., YOKOTA K., NAKAMOTO M., CHEN Y. W., NAKAMURA H., TAMURA S.: Automated segmentation of the liver from 3D CT images using probabilistic atlas and multi-level statistical shape model. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2007), vol. 10 of *Lecture Notes in Computer Science*, Springer, pp. 86–93. 133

[Ots79] OTSU N.: A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics 9*, 1 (1979), 62–66. 17

[PFJ*03] PIZER S. M., FLETCHER P. T., JOSHI S., THALL A., CHEN J. Z., FRIDMAN Y., FRITSCH D. S., GASH A. G., GLOTZER J. M., JIROUTEK M. R., LU C., MULLER K. E., TRACTON G., YUSHKEVICH P., CHANEY E. L.: Deformable m-reps for 3D medical image segmentation. *International Journal of Computer Vision 55* (Nov. 2003), 85–106. ix, 2, 26, 35

[PJE*11] PENG Y., JIANG Y., EISENGART L., HEALY M., STRAUS F., YANG X.: Segmentation of prostatic glands in histology images. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 2091 –2094. 16

[PP93] PAL N. R., PAL S. K.: A review on image segmentation techniques. *Pattern Recognition 26*, 9 (1993), 1277–1294. 12, 13, 20

[PXP00] PHAM D. L., XU C., PRINCE J. L.: A survey of current methods in medical image segmentation. In *Annual Review of Biomedical Engineering*, vol. 2. 2000, pp. 315–338. 12, 14

[QRA11] QUELLEC G., RUSSELL S., ABRAMOFF M.: Optimal filter framework for automated, instantaneous detection of lesions in retinal images. *Medical Imaging, IEEE Transactions on 30*, 2 (Feb. 2011), 523 –533. 21

[RCM10] RAMUS L., COMMOWICK O., MALANDAIN G.: Construction of patient specific atlases from locally most similar anatomical pieces. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2010), vol. 6363 of *Lecture Notes in Computer Science*, Springer, pp. 155–162. 28

[RGPA06] RUEDA S., GIL J. A., PICHERY R., ALCAÑIZ M.: Automatic segmentation of jaw tissues in CT using active appearance models and semi-automatic landmarking. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2006), Lecture Notes in Computer Science, Springer, pp. 167–174. 47, 65

[RHZ*10] REN H., HENG C.-K., ZHENG W., LIANG L., CHEN X.: Fast object detection using boosted co-occurrence histograms of oriented gradients. In *IEEE International Conference on Image Processing* (2010), pp. 2705–2708. 82

[RST*11] ROBERGE D., SKAMENE T., TURCOTTE R., POWELL T., SARAN N., FREEMAN C.: Inter- and intra-observer variation in soft-tissue sarcoma target definition. *Cancer/Radiotherapie 15*, 5 (2011), 421–425. vii, 1

[SBB09] SONG Y., BULPITT A. J., BRODLIE K. W.: Liver segmentation using automatically defined patient specific b-spline surface models. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin Heidelberg, 2009), Lecture Notes in Computer Science, Springer, pp. 43–50. 24, 26, 44, 129

[SBZ*09] SEIFERT S., BARBU A., ZHOU S. K., LIU D., FEULNER J., HUBER M., SÜHLING M., CAVALLARO A., COMANICIU D.: Hierarchical parsing and semantic navigation of full body CT data. In *Proceedings of the SPIE* (2009), Pluim J. P. W., Dawant B. M., (Eds.), pp. 725902–725902–8. 82

[SCDC11] SAMAILLE T., COLLIOT O., DORMONT D., CHUPIN M.: Automatic segmentation of age-related white matter changes on flair images: Method and multicentre validation. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 2014 –2017. 16, 21

[SG06] SCHENK O., GÄRTNER K.: On fast factorization pivoting methods for sparse symmetric indefinite systems. *Electronic Transactions on Numerical Analysis 23* (2006), 158–179. 69, 90

[SH09] SHEN T., HUANG X.: 3D medical image segmentation by multiple-surface active volume models. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2009), Lecture Notes in Computer Science, Springer, pp. 1059–1066. 47, 133

[SHD*09] SPIEGEL M., HAHN D. A., DAUM V., WASZA J., HORNEGGER J.: Segmentation of kidneys using a new active shape model generation technique based on non-rigid image registration. *Computerized Medical Imaging and Graphics 33*, 1 (2009), 29–39. 87

[SIG*09] SIMS R., ISAMBERT A., GRÉGOIRE V., BIDAULT F., FRESCO L., SAGE J., MILLS J., BOURHIS J., LEFKOPOULOS D., COMMOWICK O., BENKEBIL M., MALANDAIN G.: A pre-clinical assessment of an atlas-based automatic segmentation tool for the head and neck. *Radiotherapy Oncology 93*, 3 (Dec. 2009), 474–478. 28

[SKH*08] SEIM H., KAINMUELLER D., HELLER M., LAMECKER H., ZACHOW S., HEGE H.-C.: Automatic segmentation of the pelvic bones from CT data based on a statistical shape model. In *Eurographics Workshop on Visual Computing for Biomedicine* (2008), pp. 93–100. 65

[SKK*09] SHIMIZU A., KIMOTO T., KOBATAKE H., NAWANO S., SHINOZAKI K.: Automated pancreas segmentation from three-dimensional constrast-enhanced computed tomography. *International Journal of Computer Assisted Radiology and Surgery 5* (2009), 85–98. 97, 100, 134, 135

[SKWar] STEGER S., KIRSCHNER M., WESARG S.: Articulated atlas for segmentation of the skeleton from head and neck CT datasets. In *IEEE International Symposium on Biomedical Imaging* (2012, to appear). 145

[SLH11] SHEN T., LI H., HUANG X.: Active volume models for medical image segmentation. *Medical Imaging, IEEE Transactions on 30*, 3 (Mar. 2011), 774 –791. ix, 2, 24, 26, 35, 46, 47, 52

[SLL*10] SONG Q., LIU Y., LIU Y., SAHA P., SONKA M., WU X.: Graph search with appearance and shape information for 3-d prostate and bladder segmentation. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2010), Jiang T., Navab N., Pluim J., Viergever M., (Eds.), vol. 6363 of *Lecture Notes in Computer Science*, Springer, pp. 172–180. 23, 24

[SLS*02] SURI J. S., LIU K., SINGH S., LAXMINARAYAN S. N., ZENG X., REDEN L.: Shape recovery algorithms using level sets in 2-d/3-d medical imagery: A state-of-the-art review. *Information Technology in Biomedicine, IEEE Transactions on 6*, 1 (2002), 8–28. 43, 44

[sLsCcC01] SUNG LIAO P., SHENG CHEN T., CHOO CHUNG P.: A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering 17* (2001), 713–727. 18

[SNA*97] SATO Y., NAKAJIMA S., ATSUMI H., KOLLER T., GERIG G., YOSHIDA S., KIKINIS R.: 3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. In *Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery* (Berlin, Heidelberg, 1997), vol. 1205 of *Lecture Notes in Computer Science*, Springer, pp. 213–222. 20

[SOI*07] SHIMIZU A., OHNO R., IKEGAMI T., KOBATAKE H., NAWANO S., SMUTEK D.: Segmentation of multiple organs in non-contrast 3D abdominal images. *International Journal of Computer Assisted Radiology and Surgery 2* (2007), 135–142. 96, 134, 135

[SPV*10] SCHNEIDER R., PERRIN D., VASILYEV N., MARX G., DEL NIDO P., HOWE R.: Mitral annulus segmentation from 3D ultrasound using graph cuts. *Medical Imaging, IEEE Transactions on 29*, 9 (Sept. 2010), 1676 –1687. 23

[SRN*03] STYNER M. A., RAJAMANI K. T., NOLTE L.-P., ZSEMLYE G., SZÉKELY G., TAYLOR C. J., DAVIES R. H.: Evaluation of 3D correspondence methods for model building. In *Information Processing in Medical Imaging* (2003), pp. 63–75. 47, 65, 114

[SsNC*10] SØRENSEN T., ØSTERGAARD NOE K., CHRISTOFFERSEN C., KRISTIANSEN M., MOURIDSEN K., ØSTERBY O., BRIX L.: Active contours in optical flow fields for image sequence segmentation. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on* (April 2010), pp. 916–919. 22

[SWM*06] SAMPAT M., WANG Z., MARKEY M., WHITMAN G., STEPHENS T., BOVIK A.: Measuring intra- and inter-observer agreement in identifying and localizing structures in medical images. In *International Conference on Image Processing* (Oct. 2006), pp. 81 –84. vii, 1

[TF88] TERZOPOULOS D., FLEISCHER K.: Deformable models. *The Visual Computer 4*, 6 (1988), 306–331. 21, 44

[TFCT98] T. F. COOTES G. J. E., TAYLOR C. J.: Active appearance models. In *5th European Conference on Computer Vision* (Berlin, Heidelberg, 1998), vol. 2 of *Lecture Notes in Computer Science*, Springer, pp. 484–498. ix, 2, 26, 35, 51

[Thi98] THIRION J. P.: Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis 2*, 3 (1998), 243–260. 105

[TSK*01] TSAGAAN B., SHIMIZU A., KOBATAKE H., KUNIHISA M., HANZAWA Y.: Segmentation of kidney by using a deformable model. In *International Conference on Image Processing* (2001), pp. 1059–1062. 88

[TSKM02] TSAGAAN B., SHIMIZU A., KOBATAKE H., MIYAKAWA K.: An automated segmentation method of kidney using statistical information. In *Medical Image Computing and Computer-Assisted Intervention* (London, UK, 2002), Lecture Notes in Computer Science, Springer, pp. 556–563. 88, 127

[TSM11] TOTH R., SPARKS R., MADABHUSHI A.: Medial axis based statistical shape model (MASSM): Applications to 3D prostate segmentation on MRI. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 1463–1466. 26

[TYW*03] TSAI A., YEZZI A., WELLS W., TEMPANY C., TUCKER D., FAN A., GRIMSON W. E., WILLSKY A.: A shape-based approach to the segmentation of medical imagery using level sets. *IEEE transactions on medical imaging 22*, 2 (Feb. 2003), 137–154. 43

[UAW*11] UKWATTA E., AWAD J., WARD A. D., BUCHANAN D., PARRAGA G., FENSTER A.: Coupled level set approach to segment carotid arteries from 3D ultrasound images. In *Biomedical Imaging: From Nano to Macro, 2012 IEEE International Symposium on* (2011), pp. 37–40. 43

[vABS11] VAN AARLE W., BATENBURG K., SIJBERS J.: Optimal threshold selection for segmentation of dense homogeneous objects in tomographic reconstructions. *Medical Imaging, IEEE Transactions on 30*, 4 (April 2011), 980–989. 18

[WEKK11a] WESARG S., ERDT M., KAFCHITSAS K., KHAN M. F.: Direct visualization of regions with lowered bone mineral density in dual-energy CT images of vertebrae. In *Medical Imaging: Computer-Aided Diagnosis* (2011), Proceedings of SPIE; 7963, The International Society for Optical Engineering (SPIE), SPIE Press, Bellingham, pp. 79633J–1–79633J–10. 4, 86, 108

[WEKK11b] WESARG S., ERDT M., KAFCHITSAS K., KHAN M. F.: Quantifizierung und Visualisierung der Struktur des trabekulären Knochens in Wirbelkörpern. In *Bildverarbeitung für die Medizin 2011* (Berlin, Heidelberg, 2011), Informatik aktuell, Gesellschaft für Informatik (GI), Springer, pp. 309–313. 4, 86, 108

[WKEK10] WESARG S., KAFCHITSAS K., ERDT M., KHAN M. F.: CAD of osteoporosis in vertebrae using dual-energy CT. In *Twenty-Third IEEE Symposium on Computer-Based Medical Systems* (2010), IEEE Computer Society, Los Alamitos, Calif., pp. 358–363. 4, 86, 108

[WKFH11] WANG L., KOHNEN M., FRIMAN O., HAHN H. K.: Fast automated segmentation of femoral heads in fluoroscopic X-ray images. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 984–988. 26

[WLC*10] WANG P., LIAO W.-S., CHEN T., ZHOU S., COMANICIU D.: Graph based interactive detection of curve structures in 2D fluoroscopy. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2010), Jiang T., Navab N., Pluim J., Viergever M., (Eds.), vol. 6363 of *Lecture Notes in Computer Science*, Springer, pp. 269–277. 23

[WNR11] WANG Y., NARAYANASWAMY A., ROYSAM B.: Novel 4-d open-curve active contour and curve completion approach for automated tree structure extraction. In *Computer Vision and Pattern Recognition* (June 2011), pp. 1105 –1112. 22

[WRLG*07] WU M., ROSANO C., LOPEZ-GARCIA P., CARTER C. S., AIZENSTEIN H. J.: Optimum template selection for atlas-based segmentation. *NeuroImage 34*, 4 (Feb. 2007), 1612–1618. 27

[WS11] WU X., SHAH S. K.: Cell segmentation in multispectral images using level sets with priors for accurate shape recovery. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 2117–2120. 43

[WSD*11] WANG H., SUH J. W., DAS S., ALTINAY M., PLUTA J., YUSHKEVICH P.: Hippocampus segmentation using a stable maximum likelihood classifier ensemble algorithm. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 2036 –2040. 21

[WSH09] WIMMER A., SOZA G., HORNEGGER J.: A generic probabilistic active shape model for organ segmentation. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2009), Yang G.-Z., Hawkes D., Rueckert D., Noble A., Taylor C., (Eds.), vol. 5762 of *Lecture Notes in Computer Science*, Springer, pp. 26–33. viii, 2, 24, 43, 50, 51, 52, 70, 74, 129, 130

[WZG07] WANG Y., ZHU Y., GUO Q.: Medical image segmentation based on deformable models and its applications. In *Deformable Models*, Topics in Biomedical Engineering. International Book Series. Springer, New York, 2007, pp. 209–260. 11

[XMM10] XU J., MONACO J., MADABHUSHI A.: Markov random field driven region-based active contour model (MaRACel): Application to medical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2010), Jiang T., Navab N., Pluim J., Viergever M., (Eds.), vol. 6363 of *Lecture Notes in Computer Science*, Springer, pp. 197–204. 23

[XNS*11] XU X., NIEMEIJER M., SONG Q., SONKA M., GARVIN M., REINHARDT J., ABRAMOFF M.: Vessel boundary delineation on fundus images using graph-based approach. *Medical Imaging, IEEE Transactions on 30*, 6 (June 2011), 1184–1191. 23

[XS11] XU J., SUZUKI K.: Computer-aided detection of hepatocellular carcinoma in hepatic CT: False positive reduction with feature selection. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (2011), pp. 1097–1100. 21

[YBCK10] YIN Z., BISE R., CHEN M., KANADE T.: Cell segmentation in microscopy imagery using a bag of local bayesian classifiers. In *IEEE International Symposium on Biomedical Imaging* (2010), pp. 125 –128. 18

[YLKC10] YIN Z., LI K., KANADE T., CHEN M.: Understanding the optics to aid microscopy image segmentation. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2010), Jiang T., Navab N., Pluim J., Viergever M., (Eds.), vol. 6361 of *Lecture Notes in Computer Science*, Springer, pp. 209–217. 18

[Yoo04] YOO T. (Ed.): *Insight Into Images 'Principles and Practice for Segmentation, Registration and Image Analysis'*. A K Peters Ltd., 2004. 11, 12

[YS09] YAO J., SUMMERS R. M.: Statistical location model for abdominal organ localization. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2009), Lecture Notes in Computer Science, Springer, pp. 9–17. 119, 133

[ZBG*07] ZHEN Y., BARBU A., GEORGESCU B., SCHEUERING M., COMANICIU D.: Fast automatic heart chamber segmentation from 3D CT data using marginal space learning

and steerable features. In *International Conference on Computer Vision* (2007), pp. 1–8. viii, 2, 24, 26, 44, 45, 50, 52, 61, 70, 74, 100, 133

[ZHB06] ZAMBAL S., HLADUVKA J., BÜHLER K.: Improving segmentation of the left ventricle using a two-component statistical model. In *Medical Image Computing and Computer-Assisted Intervention* (Berlin, Heidelberg, 2006), Lecture Notes in Computer Science, Springer, pp. 151–158. 24, 44, 45, 70

[ZHM*10] ZHANG S., HUANG J., METAXAS D., WANG W., HUANG X.: Discriminative sparse representations for cervigram image segmentation. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on* (April 2010), pp. 133–136. 18

[Zit03] ZITOVA B.: Image registration methods: a survey. *Image and Vision Computing 21*, 11 (Oct. 2003), 977–1000. 11, 12

[ZOON11] ZUVA T., OLUGBARA O. O., OJO S. O., NGWIRA S. M.: Image segmentation, available techniques, developments and open issues. *Canadian Journal on Image Processing and Computer Vision 2*, 3 (2011), 20–29. 11, 12, 13, 14, 20