

## **Session 6**

# **Bilddatenanalyse**



# Segmentierung von Faserbahnsystemen mit einem strahlenbasierten Ansatz und Glättung der FA- und Winkel-Karten

M. H. A. Bauer<sup>1,2</sup>, S. Barbieri<sup>3</sup>, J. Egger<sup>1,2</sup>, D. Kuhnt<sup>1</sup>, J. Klein<sup>3</sup>, H. K. Hahn<sup>3</sup>, B. Freisleben<sup>2</sup>, Ch. Nimsky<sup>1</sup>

<sup>1</sup> Universitätsklinikum Giessen und Marburg, Klinik für Neurochirurgie, Marburg, Germany

<sup>2</sup> Philipps-Universität Marburg, Institut für Verteilte Systeme, Marburg, Germany

<sup>3</sup> Fraunhofer MEVIS, Bremen, Germany

Kontakt: baueremi@med.uni-marburg.de, sebastiano.barbieri@mevis.fraunhofer.de

## Abstract:

Für die Planung und Durchführung neurochirurgischer Eingriffe ist neben der Kenntnis der zu resezierenden Struktur auch Wissen über den Verlauf eloquenter Strukturen im menschlichen Gehirn unerlässlich, um die Gefahr postoperativer neurologischer Defizite minimal zu halten und so die Patientensicherheit zu erhöhen. Zur Lokalisation dieser Strukturen kann die Diffusions-Tensor-Bildgebung (DTI) eingesetzt werden. Gängige Verfahren wie Fiber Tracking erzeugen jedoch keine geschlossenen die Struktur begrenzenden 3D-Objekte, wie sie für die Integration in die Neuronavigation notwendig sind. Hier wird eine Erweiterung unseres strahlenbasierten Segmentierungsverfahrens vorgestellt und anhand von Software-Phantomen mit dem Dice-Koeffizienten evaluiert. Zur Segmentierung wird dabei an den einzelnen Evaluationspunkten eine lokale 3D-Umgebung betrachtet und anhand mehrerer Parameter und Schwellwerte eine Aussage über die Zugehörigkeit zur Struktur getroffen.

Schlüsselworte: Diffusions-Tensor-Bildgebung, DTI, Fiber Tracking, Segmentierung

## 1 Problem

Wissen über den Verlauf eloquenter Strukturen im menschlichen Gehirn ist für die Planung und Durchführung neurochirurgischer Eingriffe nahezu unerlässlich, um die Patienten möglichst schonend und ohne Risiko postoperativer neurologischer Defizite operativ zu behandeln. Neben beispielsweise zu entfernenden Strukturen sind wichtige Faserbahnsysteme, wie die Pyramidenbahn (Motorik) oder die Sprachbahnen als große Bahnsysteme besonders zu schützen.

Mit Hilfe der Diffusions-Tensor-Bildgebung (DTI) ist es möglich, den Verlauf und die Lokalisation solcher Bahnsysteme zu schätzen. Die Schätzung erfolgt dabei über die Diffusionseigenschaften von Wassermolekülen im menschlichen Gehirn. In einem hindernisfreien Raum verhält sich die Diffusion isotrop, entlang von Hindernissen jedoch stark anisotrop. Aus diesem Verhalten lassen sich Rückschlüsse auf die Faserverläufe ziehen. Um die Diffusionsparameter zu ermitteln, werden spezielle diffusionsgewichtete MRT-Sequenzen genutzt. Dabei werden neben einer nicht diffusionsgewichteten Aufnahme ( $b_0$ -Bild) mindestens 6 weitere Aufnahmen mit unterschiedlichen Gradienten gebraucht. Darüber lässt sich mit Hilfe der Stejskal-Tanner-Gleichung für jedes Voxel ein Tensor  $D$  zweiter Ordnung bestimmen, aus dem sich die drei Hauptdiffusionsrichtungen in diesem Voxel durch Diagonalisieren ableiten lassen [1, 2, 3].

$$D = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

Es gibt zahlreiche skalare Maße, die verwendet werden können, um die Gerichtetheit der Diffusion in einem Voxel zu charakterisieren. Das wohl bekannteste skalare Maß ist die fraktionelle Anisotropie (FA) [4], die den Anteil der Magnitude des Tensor angibt, der der gerichteten Diffusion zugeschrieben wird und sich zwischen 0 (vollständig isotrop) und 1 (anisotrop) bewegt. Der FA-Wert wie folgt definiert [1, 5]

$$FA = \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2}{2 \cdot (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}$$

In der Klinik eingesetzte Verfahren zur Rekonstruktion von Faserbündeln beruhen oft auf dem sogenannten *Streamline Fiber Tracking* [6, 7]. Diese Algorithmen erzeugen mehrere Linien, die entlang der Nervenbahnen verlaufen; zur Integration in den OP muss dazu ein korrespondierendes 3D-Objekt erzeugt werden. Dies geschieht in der Regel durch schrittweise Erzeugung von Hüllen in 2D-Schichten entlang des Verlaufs [8, 9]. Fehler im Tracking selbst haben dabei

maßgeblichen Einfluss auf das resultierende 3D-Objekt. Ein anderer Ansatz ist das *Directional Volume Growing* [2], das ausgehend von einer Saatregion direkt ein zur Struktur korrespondierendes Volumen erzeugt.

Der Beitrag ist wie folgt aufgebaut: Abschnitt 2 stellt den neuen Ansatz dar, in Abschnitt 3 werden die experimentellen Ergebnisse präsentiert. Abschnitt 4 fasst den Beitrag zusammen und gibt einen Ausblick auf potentielle Erweiterungen.

## 2 Methoden

In diesem Abschnitt wird eine Erweiterung des von uns entwickelten strahlenbasierten Segmentierungsverfahrens [10, 11] für eloquente Strukturen im menschlichen Gehirn vorgestellt.

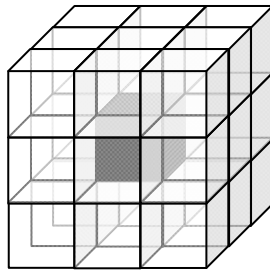
Basierend auf einem manuell gestarteten deterministischen Fiber Tracking wird zunächst durch die Platzierung von Include-Regionen das entstandene Fiber Tracking Ergebnis eingeschränkt (Ausreißer entfernen etc.) und anschließend durch die Berechnung der Mittellinie (Centerline) des Faserbündels nach [12] eine Grundlage zum Aussenden der Evaluationsstrahlen geschaffen. Entlang der Centerline werden an äquidistanten Stützstellen Ebenen senkrecht zum Centerlineverlauf berechnet. In jeder dieser Ebene werden gleich verteilt Strahlen von dem Centerlinestützpunkt ausgesendet und abgetastet, so dass eine Reihe von Evaluationspunkten für die Segmentierung entsteht.

Zu jedem Evaluationspunkt werden einige Parameter bestimmt, die später für die Segmentierungsentscheidung eine Rolle spielen. Diese umfassen den FA-Wert und zwei Winkel zwischen jeweils der Hauptdiffusionsrichtung des Centerlinepunktes und des Evaluationspunktes sowie des Evaluationspunktes und des vorangegangenen Evaluationspunktes.

Für jeden Strahl wird dann anhand von Schwellwertkriterien bzgl. der genannten Parameter ein Grenzpunkt für jeden Strahl bestimmt. Um fälschliche Aussagen durch Rauschen in den Daten einzudämmen, wurde bislang eine Folge von Evaluationspunkten entlang eines Strahls als Untersuchungsmenge zur Randpunktbestimmung für diesen Strahl genutzt [10]. Im Gegensatz dazu soll nun eine lokale 3D Umgebung um den aktuellen Untersuchungspunkt zur Entscheidungsfindung (Zuordnung des Punktes zu Faserbündel oder Umgebung) genutzt werden, wie sie in Abbildung 1 (links) schematisch dargestellt ist. Damit wird von der „linearen“ Evaluationsmenge Abstand genommen und auf eine 3D Umgebung umgestiegen um einen kompakten, aber in allen Richtungen expandierten Evaluationsbereich zu erhalten. Dadurch ist es möglich, einzelne verrauschte und sonst fehlerhaft klassifizierte Bildpunkte durch die Mittelungen mit den Masken zu schwächen, so dass diese das Ergebnis nicht maßgeblich beeinflussen. Gelangt man an eine Randzone, bei der dann mehrere Bildpunkte nicht mehr die Eigenschaften einer typischen Faserbahn aufweisen, schlägt dies trotz Mittelung zu Buche und kann detektiert werden. Als lokale dreidimensionale Umgebungen werden dabei nach dem Prinzip von Bild-Filteroperatoren gewichtete Würfel mit Kantenlänge  $k$  eingesetzt. Neben einem Würfel mit homogener Gewichtungsfunktion kommen auch noch Gauß-gewichtete Evaluationsschemata zum Einsatz (für die Gewichtung im zweidimensionalen Fall siehe Abbildung 1, rechts). Der entsprechende Evaluationswert  $e$  für die einzelnen Parameter wird dann gemäß der folgenden Formel für solche diskreten 3D-Masken (Zentrierung um Nullpunkt), der Größe  $(m+1) \times (n+1) \times (l+1)$  berechnet:

$$e = \sum_{i=-m/2}^{m/2} \sum_{j=-n/2}^{n/2} \sum_{k=-l/2}^{l/2} I(x+i, y+j, z+k) * \text{mask}(i, j, k)$$

Die resultierende Punktwolke, die die Grenze des Faserbündels skizziert, kann nun, wie auch in [10] beschrieben, noch einige optionale Korrekturverfahren durchlaufen (Korrekturen in einer Ebene, sowie Korrekturen bzgl. konsekutiver Ebenen). Da Schnitte der einzelnen 2D-Konturen (aus Grenzpunkten einer Ebene) eher unwahrscheinlich sind, kann hier ein einfaches Triangulationsverfahren eingesetzt werden. Für komplexere Strukturen als beispielsweise die mehr oder weniger kaum gekrümmte Pyramidenbahn können auch andere Triangulationsverfahren zum Einsatz kommen, wie zum Beispiel von Hornung und Kobbelt [13] beschrieben.



(a)	1	1	1
	27	27	27
	1	1	1
	27	27	27
(b)	2	4	2
	72	72	72
	4	16	4
	72	72	72
	2	4	2
	72	72	72

**Abb. 1:** Links: Prinzip der lokalen 3D Evaluationsumgebung der Größe 3x3x3 um das zentrale Voxel (schwarz). Rechts: Schema einer (a) homogenen und (b) Gauß-gewichteten Evaluationsmaske (exemplarisch für die mittlere vertikale Ebene einer 3x3x3 Evaluationsumgebung).

Der zeitliche Aufwand der Berechnung hängt dabei von verschiedenen Parametern ab – die Anzahl der Evaluationsebenen ( $p$ ), die Anzahl der Strahlen ( $r$ ) pro Ebene und die Anzahl der Evaluationspunkte ( $n$ ) pro Strahl.

Zur Evaluation des Verfahrens werden zwei Softwarephantome (Torus-Phantom und anatomisches Phantom) mit modellierten Faserverläufen eingesetzt [14, 15]. Das Wissen über Lage und Ausdehnung der Strukturen wird als Qualitätsgrundlage zur Bewertung des Verfahrens genutzt. In einem späteren Schritt soll dann die Evaluation mit realen Probanden- bzw. Patientendatensätzen folgen.

### 3 Ergebnisse

Alle Methoden wurden in C++ innerhalb der medizinischen Prototyping Plattform MeVisLab ([www.mevislab.de](http://www.mevislab.de)) implementiert. Die Rekonstruktion der Faserbahnen dauerte nach manueller Initialisierung (Saaregion und Parameter definieren) maximal 1 Minute auf einem Intel Core 2 Quad CPU, 3 GHz, 6 GB RAM, Windows XP Professional 2003, SP 2.

Zur Bewertung der Segmentierungsergebnisse wurde der Dice-Koeffizient (DSC) [16, 17] bestimmt, der die Qualität des Segmentierungsergebnisses aus dem Grad der Überlappung zweier segmentierter Objekte A (Referenz) und B (Segmentierung) mit folgender Formel ableitet:

$$DSC = \frac{2 \cdot |A \cap B|}{|A| + |B|} \quad DSC \in [0,1]$$

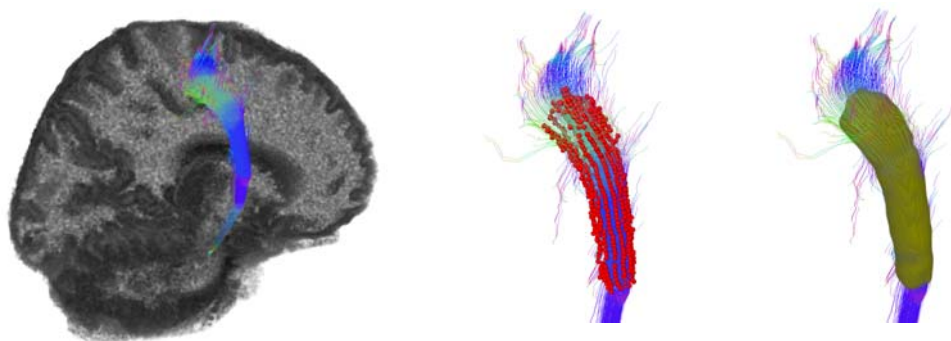
Für die beiden Software-Phantome (anatomisches Phantom siehe Abbildung 2, links) wurden jeweils 3D-Masken mit homogener Gewichtung und Gauß-Gewichtung unterschiedlicher Größe (1x1x1 (ohne Evaluationsumgebung), 3x3x3, 5x5x5) als lokale Evaluationsareale angewendet. Bei beiden Phantomen wurden 50 Evaluationsebenen mit jeweils 15 Strahlen genutzt. Die Strahlen wurden im Abstand von 0,5mm an 30 Punkten abgetastet, so dass insgesamt jeweils 22500 Evaluationspunkte zur Beschreibung der Faserbahnen zur Verfügung standen, was zu den in Tabelle 1 und Tabelle 2 angegebenen DSCs führt. Ein Rekonstruktionsergebnis für das anatomische Phantom ist in Abbildung 2 dargestellt.

Phantom 1	homogen gewichtete Evaluationsmaske			Gauß-gewichtete Evaluationsmaske		
	1x1x1	3x3x3	5x5x5	1x1x1	3x3x3	5x5x5
mittlerer DSC (%)	77,45	91,02	89,87	77,45	90,21	90,39
Standardabweichung	26,26	1,03	0,80	26,26	0,24	0,93

**Tabelle 1:** Ergebnisse des vorgestellten Segmentierungsverfahrens für das Torus-Phantom unter Verwendung der unterschiedlichen Evaluationsmasken.

Phantom 2	homogen gewichtete Evaluationsmaske			Gauß-gewichtete Evaluationsmaske		
	1x1x1	3x3x3	5x5x5	1x1x1	3x3x3	5x5x5
mittlerer DSC (%)	86,14	84,50	83,05	86,14	85,91	83,01
Standardabweichung	0,43	0,77	2,93	0,43	0,92	3,29

**Tabelle 2:** Ergebnisse des vorgestellten Segmentierungsverfahrens für das anatomische Softwarephantom mit modelliertem rechten Tractus corticospinalis und Verwendung der unterschiedlichen Evaluationsmasken.



**Abb. 2:** Links: Anatomisches Softwarephantom (Fiber Tracking der rechten Pyramidenbahn). Mitte: Faserbündel umhüllende Punktwolke zwischen Start- und Endbereich. Rechts: rekonstruiertes Faserbündelsegment zwischen Start- und Endbereich.

## 4 Diskussion

In dieser Arbeit wurde der Einfluss der Glättung von FA- und Winkel-Karten auf die Segmentierung von DTI Daten mit dem vorgeschlagenen strahlenbasiertem Ansatz analysiert. Bei der Segmentierung des Torus-Bündels wurden durch Anwendung der unterschiedlichen Evaluationsmasken (3x3x3 und 5x5x5) deutlich bessere Segmentierungsergebnisse erzielt. Ein signifikanter Unterschied zwischen den mit der homogen gewichteten Evaluationsmaske und der Gauß-gewichteten Evaluationsmaske produzierten Ergebnissen war jedoch nicht zu erkennen. Bei dem anatomischen Phantom ergab die Glättung keine höheren Dice-Koeffizienten. Dies könnte an nicht ausreichend optimierten Winkel- und FA-Schwellwerten liegen, was in nachfolgender Arbeit systematisch untersucht werden wird. Ferner sollten die optimale Größe der Evaluationsmasken unter verschiedenen Signal-zu-Rauschen (SNR) Bedingungen bestimmt und weitere Evaluationsmasken, wie zum Beispiel Medianfilter, getestet werden. Letzteres könnte den Einfluss von "noise-only"-Voxeln stark einschränken. In einem folgenden Schritt soll auch eine umfassende Evaluation auf weiteren Phantom-Datensätzen sowie auf realen MR-Datensätzen erfolgen. Zusammenfassend ist die hiermit vorgestellte Arbeit ein weiterer Beitrag zur genaueren Lokalisation eloquenter neuronaler Strukturen.

## 5 Referenzen

- [1] D. L. Bihan, et al., "Diffusion Tensor Imaging: Concepts and Applications," *Journal of Magnetic Resonance Imaging*, vol. 13, pp. 534–546, 2001.
- [2] D. Merhof, et al., "Directional Volume Growing for the Extraction of White Matter Tracts from Diffusion Tensor Data," in *SPIE - Medical Imaging 2005: Visualization, Image-Guided Procedures, and Display.*, vol. 5744, pp. 165–172, 2005.
- [3] Y. Ou and C. Wyatt, "Visualization of Diffusion Tensor Imaging data and Image Correction of Distortion Induced by Patient Motion and Magnetic Field Eddy Current," in *Virginia Tech - Wake Forest University School of Biomedical Engineering and Sciences 4th Student Research Symposium*, 2005.
- [4] P. J. Basser and C. Pierpaoli, "Microstructural and Physiological Features of Tissues Elucidated by Quantitative-Diffusion-Tensor MRI," *Journal of Magnetic Resonance, Series B*, vol. 111, pp. 209–219, 1996.
- [5] T. H. J. M. Peeters, et al., *Visualization and Processing of Tensor Fields*, ch. Analysis of Distance/Similarity Measures for Diffusion Tensor Imaging, pp. 113–136. Springer Berlin Heidelberg, 2009.
- [6] S. Mori and P. C. M. van Zijl, "Fiber tracking: principles and strategies - a technical review," *NMR Biomed.*, vol. 15, pp. 468–480, 2002.
- [7] M. Rueber, *Deterministic and probabilistic fibre tracking with diffusion tensor imaging under aversive diffusion conditions (in German)*. PhD thesis, Julius-Maximilians-University Wuerzburg, 2009.
- [8] Ch. Nimsky, et al., "Visualization strategies for major white matter tracts identified by diffusion tensor imaging for intraoperative use," in *Computer Assisted Radiology and Surgery*, vol. 1281, pp. 793 – 797, 2005.
- [9] Z. Ding, et al., "Case study: reconstruction, visualization and quantification of neuronal fiber pathways," in *VIS '01: Proceedings of the conference on Visualization '01*, pp. 453–456, IEEE Computer Society, 2001.
- [10] M. H. A. Bauer, et al., "Boundary Estimation of Fiber Bundles derived from Diffusion Tensor Images," *International Journal of Computer Assisted Radiology and Surgery (IJCARS)*, 2010.
- [11] M. H. A. Bauer, et al., "A Ray-based Approach for Boundary Estimation of Fiber Bundles Derived from Diffusion Tensor Imaging," in *Proceedings of Computer Assisted Radiology and Surgery (CARS 2010)*, 2010.
- [12] J. Klein, et al., "Automatic Quantification of DTI Parameters along Fiber Bundles," in *Proceeding of Image Processing for Medicine (BVM 2007)*, pp. 272–2, 2007.
- [13] A. Hornung and L. Kobbelt, "Robust reconstruction of watertight 3D models from non-uniformly sampled point clouds without normal information," in *SGP '06: Proceedings of the fourth Eurographics symposium on Geometry processing*, pp. 41–50, Eurographics Association, 2006.
- [14] S. Barbieri, et al., "Towards Image-Dependent Safety Hulls for Fiber Tracking," in *Proceedings of the Joint Annual Meeting ISMRM-ESMRMB 2010*, p. 1672, 2010.
- [15] S. Barbieri, et al. "Assessing Fiber Tracking Accuracy via Diffusion Tensor Software Models," in *Medical Imaging 2010: Image Processing* (B. M. Dawant and D. R. Haynor, eds.), vol. 7623, (San Diego, California, USA), p. 762326, SPIE, 2010.

- [16] M. P. Sampat, et al., "Measuring intra- and inter-observer Agreement in Identifying and Localizing Structures in Medical Images," *IEEE Inter. Conf. Image Processing*, 2006.
- [17] K. H. Zou, et al., "Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index," *Academic Radiology*, vol. 11, no. 2, pp. 178–189, 2004.





# **Funktionelles Modell der rotationsabhängigen Strukturverschieblichkeiten im humanen Hals: Datenakquisition, Methoden und erste Ergebnisse**

*L. Colter, L. A. Kahrs, J. Hirschfeld, J. Schipper*

*Universitätsklinikum Düsseldorf, Hals-Nasen-Ohrenklinik, Düsseldorf, Deutschland*

Kontakt: [lena.colter@uni-duesseldorf.de](mailto:lena.colter@uni-duesseldorf.de)

## **Abstract:**

*Die Bewegung anatomischer Strukturen im Kopf-Halsbereich rechnerisch vorhersagen zu können und für die medizinische Navigation nutzbar zu machen ist das langfristige Ziel dieser Modellierung. Bei chirurgischen Eingriffen am Hals wird der Kopf des Patienten so gelagert, dass sich das Halsgewebe im Vergleich zu der normalen Kopfstellung verformt und deshalb prä-operative Bilddaten für die medizinische Navigation unbrauchbar werden. Die Berücksichtigung der Verschieblichkeiten des Halsgewebes bei der Rotation des Kopfes ist ein erster Schritt um dieses Problem zu lösen. Auf Basis von MRT-Bilddaten, die von Probanden akquiriert werden, wird die Strukturverschiebung bei Kopfbewegungen gemessen und analysiert. Es zeigen sich deutlich erkennbar Unterschiede zwischen den unterschiedlichen Strukturen.*

*Schlüsselworte: minimal-traumatische Halschirurgie, Biomechanik, Tissue-Shift*

## **1 Problem**

Chirurgische Eingriffe im Halsbereich erfordern ein präzises Vorgehen um wichtige anatomische Strukturen nicht zu verletzen. Zukünftige Operationsverfahren werden eine dreidimensionale Computer-Visualisierung des OP-Situs benötigen, um eine gewebeschonende und zielgerichtete Vorgehensweise des Chirurgen im Sinne der medizinischen Navigation zu ermöglichen. Hierbei entsteht ein neuer Bedarf für die funktionelle Modellierung von Verschieblichkeiten, weil die Kopf-Hals-Position des Patienten während der präoperativen bildgebenden Verfahren nicht mit der Lagerung im OP übereinstimmt auch bilden einzelne (zusätzliche) Bildgebungsdatensätze in einer bestimmten verdrehten Kopf-Hals-Stellung niemals die Vielfalt der im OP eingenommen Kopf-Hals-Positionen ab. Eine funktionelle, winkelabhängige Modellierung wird demnach für die Anforderung der Navigierbarkeit benötigt.

Bei den Eingriffen am Hals - wie beispielweise der Neck-Dissection - findet die Lagerung zumeist mit überstrecktem Hals und gleichzeitig rotiertem Kopf statt, um dem Chirurgen eine höhere Gewebespannung, eine bessere Differenzierbarkeit und Erreichbarkeit der freizupräparierenden Strukturen zu präsentieren. Die Kombination mit intraoperativem, navigiertem Ultraschall zur dreidimensionalen Überlagerung des funktionellen Modells mit aktuellen Bildern würde die Methode der zielgerichteten Chirurgie im Halsbereich vervollständigen und die automatische Erkennung in den Sonographiebilddaten erleichtern. Eine atraumatische Passage von gefährdeten Strukturen wird somit möglich werden.

Auf dem Gebiet der Anästhesiologie wurde bereits vor Jahren die Auswirkung von Kopfrotation auf die Lage der V.jugularis interna zur A.carotis communis geprüft um die optimale Position zur Venenpunktion zu bestimmen [1,2]. Andere Forschungsgruppen untersuchten das Lageverhalten des Intubationstubus zur Luftröhre bei verschiedenen Kopf-positionen [3-5]. Relevant sind zudem Beobachtungen der Kopf-Hals-Anatomie (Kiefergelenk, A.vertrebralis) während Auffahrunfällen, bei denen der Kopf extreme Extensions- und Flexionsstellungen einnimmt [6,7]. Mittlerweile werden 3D-Visualisierungen genutzt um die normale Anatomie darzustellen, z.B. als Lernmodell für Medizinstudenten [8] und zur präoperativen Planung [9,10], oder um Pathologien, z.B. vergrößerte Lymphknoten [11] bestmöglich bewerten zu können. Eine Winkel-abhängige Modellierung von Verschieblichkeiten der Hals-Anatomie ist in der Literatur bisher nicht vollständig beschrieben worden.

Die vorliegende Arbeit setzt nun genau an dieser Stelle an: Die Anatomie des Halses wird visualisiert, allerdings nicht nur in der Neutral-Null-Stellung, sondern bei verschiedenen Kopfstellungen. So wird der Einfluss der Kopfposition auf die Anordnung der Strukturen im Hals modelliert und darstellbar.

## 2 Methoden

Um Gewebe- und Strukturverschieblichkeiten im menschlichen Hals analysieren zu können, wurden bei Probanden MRT-Aufnahmen durchgeführt (Siemens Trio, 3 Tesla, in T2-Gewichtung, 3mm Schichtabstand, Spulen: Neck 1 und 2, Spine 1 und 2, oberste Schicht: nasal, unterste Schicht: auf Höhe des Sternum). Zusätzlich zur Normalstellung erfolgte die Aufnahme bei mehreren Kopfrotationen (20°, 40°, 60°, maximal einnehmbarer Rotationswinkel und 40° zur entgegengesetzten Seite). Der jeweilige Proband nahm nach Anweisung die Winkelstellungen ein und hielt diese während der jeweiligen MRT-Aufnahmen (siehe Abb. 1). Die Vorgabe der Kopfposition erfolgte mittels eines Winkelmessers.

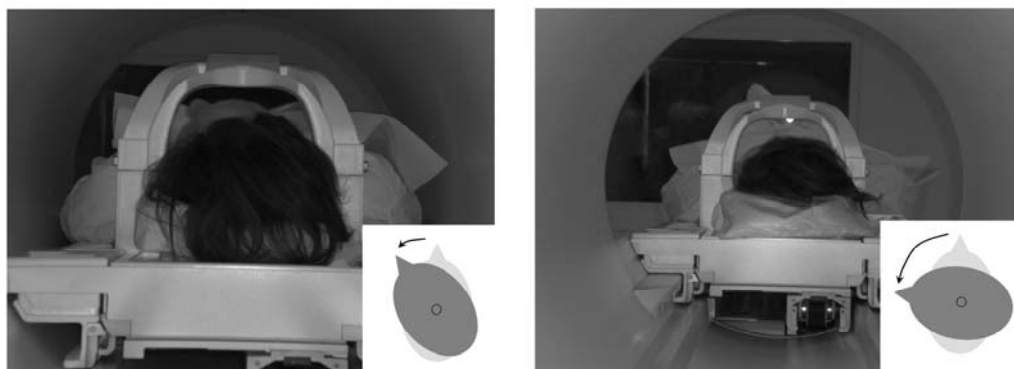


Abb. 1: Zwei beispielhafte Kopfpositionen im MRT, die beide nicht der Neutral-Null-Stellung entsprechen.

Nach der Datenakquisition wurden in den Schichtbildern beidseits relevante Strukturen (A.carotis communis, externa und interna; V.jugularis interna; Trachea; Larynx; Pharynx; Gl.thyroidea; M.sternocleidomastoideus; Sternum; Spinalkanal mit Rückenmark; Proc. spinosus; Hautoberfläche) in dem DICOM-Viewer OsiriX identifiziert und mit Hilfe eines Grafiktablets manuell segmentiert. Die anschließenden Untersuchungen und Auswertungen erfolgten sowohl anhand von 2D-Schichten als auch anhand von 3D-Visualisierungen.

Der Kern dieser Untersuchungen war die Bestimmung der Verschieblichkeit der Strukturen, die im Zweidimensionalen durch die Überlagerung von korrespondierenden Schichtbildern (z.B. auf Höhe eines bestimmten Wirbelkörpers) aus der Normalstellung und einer rotierten Kopfstellungen bestimmt werden kann. Dazu wurden in einem Grafikprogramm (GIMP) jeweils zwei der aus dem DICOM-Viewer exportierten Schichtbilder semitransparent übereinandergelegt und für die gemeinsame Rotationsachse das Zentrum des Spinalkanals festgelegt. Der Drehwinkel zwischen den jeweils gleichen Strukturen aus dem Bild in der Neutral-Null-Stellung und der rotierten Kopfposition konnte daraufhin gemessen werden (im Folgenden als Winkelverschieblichkeit bezeichnet). Abbildung 2 illustriert diese Messungen. Bisher konnten so 60 Messwerte generiert werden.

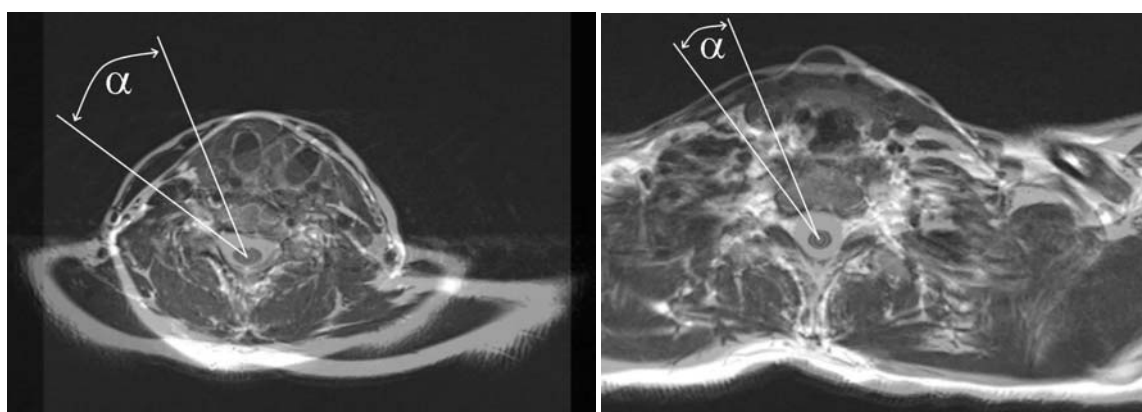


Abb. 2: Überlagerung von jeweils zwei MRT-Aufnahmen in axialer Schichtung. Im linken Bild ist die Höhe des sechsten cervikalen Wirbelkörpers (C6) und rechts die Höhe des ersten thorakalen Wirbelkörpers (Th1) dargestellt. Die untersuchte Struktur ist die rechte V.jugularis interna. Sie ist in den überlagerten semitransparenten Aufnahmen (Neutral-Null-Stellung und maximaler Rotation nach links) jeweils segmentiert und ihr Mittelpunkt markiert. Das Rotationszentrum ist durch einen weiteren Punkt visualisiert und wurde durch den Spinalkanal im Wirbelkörper (ebenfalls segmentiert) festgelegt. Die Verschiebung der Strukturen wird durch die Rotation  $\alpha$  gemessen, die in diesem Fall links  $\alpha = 31,6^\circ$  und rechts  $\alpha = 14,4^\circ$  beträgt.

Die Betrachtung der MRT-Daten im Dreidimensionalen ermöglichte einen realitätsnahen Überblick über die Halsstrukturen, die mit dem Programm 3D Slicer erzeugt wurde. Eine Überlagerung der segmentierten Strukturen bei verschiedenen Rotationswinkeln illustrierte die Veränderung (vgl. Abbildung 3).

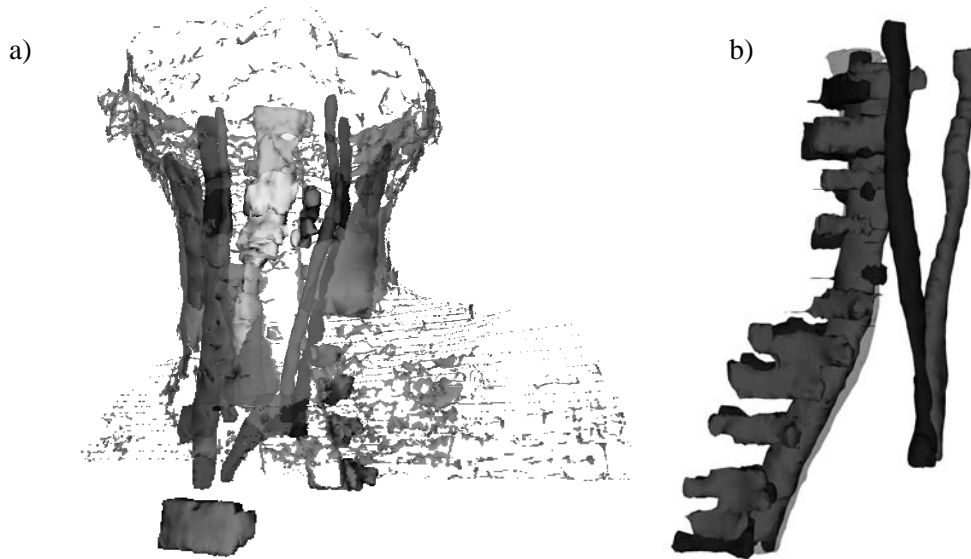


Abb. 3: a) Übersicht der segmentierten anatomischen Strukturen des Halses in einer um 51° nach links rotierten Kopf-Position. b) Beispielhafte Darstellung der rechten V.jugularis interna bei Neutral-Null (hellere Struktur) und links rotierten Kopf (77°, dunklere Struktur) mit gleichzeitiger Visualisierung der Wirbelsäule.

### 3 Ergebnisse

Um die selbstentwickelte Methodik zu evaluieren und für nachfolgende Probanden erweitern oder verbessern zu können, wurde bei einem ersten Probanden die komplette, oben beschriebene Vorgehensweise getestet. Die Winkelvorgaben müssen durch Winkelmessungen anhand der Nasenspitze in den MRT-Aufnahmen korrigiert werden. Tabelle 1 fasst diese Kalibrierung zusammen.

Tab. 1: Vorgabe, in den MRT-Schichtbildern gemessene Winkelstellung des Kopfes und deren Differenz.

Vorgabewinkel	0°	20°	40°	60°	max	40° n. re
Messung MRT	0,3° n. li	14,6°	38,3°	51,7°	76,9°	32°
Differenz	0,3°	5,4°	1,7°	8,3°	-	8°

Insgesamt wurden von 6 Datensätzen mit unterschiedlichen Rotationswinkeln jeweils in 65 Schichten die 19 oben genannten Strukturen segmentiert. Bei einem weiteren Probanden wurde die Normal- und Maximal-Rotations-Stellung untersucht. Zwei exemplarische Messwerte sind aus Abbildung 2 ablesbar. Es zeigte sich, dass Ähnlichkeiten bezüglich der Verschieblichkeit bestimmter Strukturen erkennbar sind. Abbildung 3 zeigt einen Ausschnitt der Messwerte und Abhängigkeiten, die sich auf die V.jugularis interna links und rechts beschränken. Generell ist erkennbar, dass bei der Links-Drehung des Kopfes die V.jugularis interna rechts eine größere Verschieblichkeit aufweist als die V.jugularis interna links.

Das linke Diagramm in Abbildung 4 zeigt den Verlauf der Struktur-Verschieblichkeit in Relation zum Drehwinkel des Kopfes. Zwischen 0° und 40° Kopffrotation unterscheiden sich die grafischen Verläufe von V.jugularis interna links und rechts stark; ab 40° zeigen diese wieder vergleichbare Abhängigkeit. Das rechte Diagramm in Abbildung 4 zeigt, dass die Lageveränderungen im caudalem Halsbereich nicht so ausgeprägt sind wie die weiter cranial. Die Verläufe für die V.jugularis interna links und rechts sind etwas verschoben aber ähnlich. Dargestellt ist die Verschieblichkeit der Strukturen auf unterschiedlichen Höhen der Wirbelkörper.

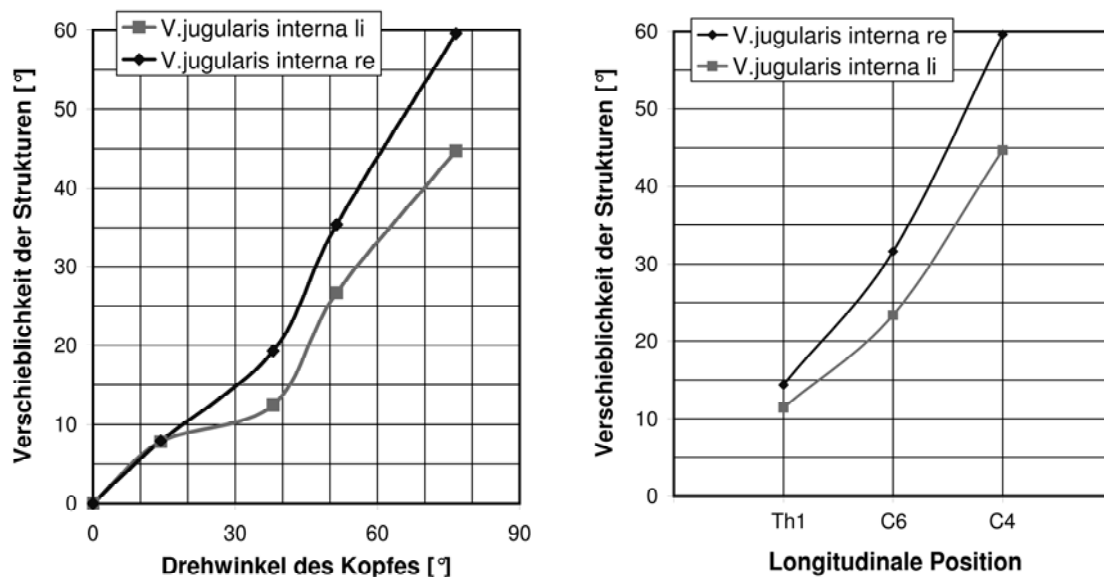


Abb. 4: Grafische Auftragung der Ergebnisse der Strukturbeweglichkeiten für die V.jugularis interna rechts und links. In der linken Grafik ist die Abhängigkeit der Kopffrotation in Relation zur Winkelverschiebung der Struktur auf Höhe des Wirbelkörpers C4 dargestellt. Die rechte Grafik zeigt die Abhängigkeit der Position in der longitudinalen Achse zur Winkelverschiebung der Strukturen.

## 4 Diskussion

Im Zentrum dieser laufenden Arbeit steht die Entwicklung eines durch Kopfbewegungen parametrisierbarem 3D-Modell des Halses. Um Winkelabhängigkeiten noch präziser und später generisch beschreiben zu können, werden eine größere Anzahl an Probanden benötigt. Durch das manuelle Segmentieren ist der Zeitaufwand im Moment allerdings sehr groß (pro Aufnahme, bestehend aus 65 Schichten, ca. 6 Std.). Methodisch sind außerdem die doch sehr großen Schichtabstände des MRTs sowie ein weiterer Fakt unbefriedigend: Beim Überlagern kann nicht gleichzeitig anhand der Wirbelkörper sowie den fazialen Oberflächen die korrespondierenden Schichten ausgewählt werden. Hierbei kommt es ggfs. zu Ungenauigkeiten. Hierfür könnte die Quantifizierung der bisher nur qualitativen Betrachtungen in 3D ein sehr interessanter Schritt sein, auch weil damit ein Vergleich innerhalb der 2D- und 3D-Daten möglich wird. Bislang wurden 3D-Messungen nicht intensiv untersucht, weil eine Definition von Messpunkten zur Auswertung der Gewebeverschiebung beispielsweise der V.jugularis interna (vgl. Abb. 3 rechts) sich als schwierig erweist. Dies ist begründet durch das Fehlen von wiederauffindbaren anatomischen Landmarken, wie z.B. Bifurkationen, signifikanten Dilatationen oder Stenosen.

Nach dem Segmentierungs- und Überlagerungsschritt wurden bisher 60 Datenpunkte durch die Messung der jeweiligen Strukturverschieblichkeit erzeugt, allerdings ist dies nur ein Bruchteil der insgesamt an einem Datensatz auswertbaren Strukturen. Beispielsweise lässt sich bereits aus den weiteren ausgewerteten Datenpunkten erkennen, dass generell die dickwandigen Arterien weniger beweglich sind als die Venen.

Die Analyse der Parameter der funktionellen Zusammenhänge ist zum momentanen Zeitpunkt noch nicht vollständig verstanden. Patientenindividuelle Werte wie z.B. Halslänge, BMI spielen mit Sicherheit eine Rolle, können aber noch nicht wissenschaftlich belegt werden. Eine Erweiterung der Untersuchung auf Patienten mit Pathologien und das dort sicherlich modifizierte Verhalten der Beweglichkeiten ist ebenfalls bislang nicht untersucht worden. Die Modifizierung des Modells hinsichtlich weiterer Beweglichkeiten (z.B. Reklination) ist der zunächst anstehende Erweiterungsschritt.

Schlussendlich erwarten wir, dass durch diese Untersuchungen zukünftig eine Möglichkeit der Navigation im Hals mit der Online-Berechnung der Weichgewebeverschiebung entsteht sowie für Forschung und Lehre ein flexibler Atlas für den Kopf-Hals-Bereich entsteht.

## 5 Danksagung

Die Autoren danken der Deutschen Forschungsgemeinschaft für die Förderung sowie dem Institut für Radiologie des Universitätsklinikums Düsseldorf für die Unterstützung und das Bereitstellen von MRT-Messzeit. PD Dr. Hans-Jörg Wittsack sowie Frau Erika Rädisch gebührt diesbezüglich ein besonderer Dank. Die Fragestellung dieses Beitrags wur-

de im Rahmen des DFG-geförderten Projektes „SACAS (Sonographic Aided Computer Assisted Surgery) im Halsbereich“ (Geschäftszeichen SCHI310/10-1) erarbeitet

## 6 Referenzen

- [1] Sulek CA, Gravenstein N, Blackshear RH, Weiss L. Head rotation during internal jugular vein cannulation and the risk of carotid artery puncture. *Anesth Analg* 1996;82(1):125-8.
- [2] Wang R, Snoey ER, Clements RC, Hern HG, Price D. Effect of head rotation on vascular anatomy of the neck: an ultrasound study. *J Emerg Med* 2006;31(3):283-6.
- [3] Conrardy PA, Goodman LR, Lainge F, Singer MM. Alteration of endotracheal tube position. Flexion and extension of the neck. *Crit Care Med* 1976;4(1):7-12.
- [4] Hartrey R, Kestin IG. Movement of oral and nasal tracheal tubes as a result of changes in head and neck position. *Anaesthesia* 1995;50(8):682-7.
- [5] Kim JT, Kim HJ, Ahn W, Kim HK, Bahk JH, Lee SC, Kim CS, Kim SD. Head rotation, flexion, and extension alter endotracheal tube position in adults and children. *Can J Anesth* 2009;56(10):751-6.
- [6] Schneider K, Zernicke RF, Clark G. Modeling of jaw-head-neck dynamics during whiplash. *J Dent Res* 1989;68(9): 1360-5.
- [7] Ivancic PC, Ito S, Tominaga Y, Carlson EJ, Rubin W, Panjabi MM. Effect of rotated head posture on dynamic vertebral artery elongation during simulated rear impact. *Clin Biomech* 2006;21(3):213-20.
- [8] Nguyen N, Wilson TD. A head in virtual reality: development of a dynamic head and neck model. *Anat Sci Educ* 2009;2(6):294-301.
- [9] Storck C, Gugatschka M, Friedrich G, Sorantin E, Ebner F, Fischer C, Wolfensberger M, Juergens P. Developing a 3D model of the laryngeal cartilages using HRCT data and MIMICS's segmentation software. *Logoped Phoniatr Vocol* 2010;35(1):19-23.
- [10] Mühler K, Preim B. Reusable visualizations and animations for surgery planning. *Computer Graphics Forum* 2010;29(3):1-10.
- [11] Strauss G, Hertel I, Dornheim J, Cordes J, Burgert O, Schulz T, Meixensberger J, Winkler D, Preim U, Dietz A, Preim B. Three-dimensional visualization of CT of the neck for surgical planning. *Laryngorhinootologie* 2006;85(10):746-54.



# Interaktive Techniken zur Korrektur medizinischer Segmentierungen auf Basis Stabiler Feder-Masse-Modelle

S. Rahner<sup>1</sup>, J. Dornheim<sup>1</sup>, L. Dornheim<sup>2</sup>, B. Preim<sup>1</sup>

<sup>1</sup> Universität Magdeburg, Institut für Simulation und Graphik, Magdeburg, Germany

<sup>2</sup> Dornheim Medical Images GmbH, Magdeburg, Germany

Kontakt: sebastian.rahner@student.uni-magdeburg.de

## Abstract:

*Die Auswertung digitaler Bilddaten für die Diagnose und Therapie von Krankheitsbildern nimmt heute einen hohen Stellenwert ein. Eine Computerunterstützung für die Segmentierung anatomischer Strukturen ist von Bedeutung, nach heutigem Stand der Technik aber nicht vollkommen exakt, was Fehlsegmentierungen zur Folge hat. Für die Planung von Halstumor-Operationen etwa werden Segmentierungen genutzt, um Maßzahlen wie Tumorgröße oder -ausdehnung, sowie Informationen über mögliche Infiltration und Sicherheitsabstände zu berechnen. Die Korrektheit der Segmentierungen muss in solchen Fällen gewährleistet sein und erfordert gerade bei automatischen Verfahren häufig manuelle Korrekturen.*

*Wir stellen vier Interaktionstechniken zur nachträglichen Korrektur modellbasierter Segmentierungen auf Basis Stabiler Feder-Masse-Modelle vor. Erstmals wird dabei das durch die Formmodelle integrierte Wissen über die zu segmentierende Struktur zur Korrekturdurchführung hinzugezogen. In einer ersten Nutzerstudie konnte gezeigt werden, dass die entwickelten Interaktionstechniken auch untrainierten Anwendern eine effektive und effiziente Durchführung von Nachbesserungen ermöglichen.*

*Schlüsselworte: Stabile Feder-Masse-Modelle, interaktive Korrekturwerkzeuge, Modellbasierte Segmentierung, CT*

## 1 Problem

Die computergestützte Bildgebung nimmt in der aktuellen Medizintechnik einen zunehmend höheren Stellenwert ein. Um die aufgenommenen Bilddaten effektiv und effizient analysieren zu können, ist häufig eine Segmentierung anatomischer Strukturen notwendig. Zur Unterstützung der chirurgischen Planung bei Hals-Tumoroperationen sind Segmentierungen des Tumors, der Lymphknoten sowie der umgebenden Muskeln, Gefäße und weiterer Risikostrukturen von Bedeutung [6], um u.a. Größe, Ausdehnung und Infiltrationsgebiete des Tumors sowie Sicherheitsabstände zu Nachbarstrukturen zu berechnen.

Gerade im Falle automatischer Segmentierungsverfahren sind aufgrund individueller bzw. pathologischer Abweichungen einzelner Strukturen vom Modell Fehler in den Segmentierungsergebnissen möglich, die manuell korrigiert werden müssen. Dazu müssen dem Anwender effiziente Interaktionstechniken bereitgestellt werden, die es ihm durch einfache Erlernbarkeit und intuitive Steuerung erlauben, seine Korrekturwünsche effektiv auf die Segmentierung zu übertragen.

Es existiert eine Vielzahl unterschiedlicher (halb)automatischer Verfahren, die mit individuellen Strategien Segmentierungen berechnen. Ein in [1] vorgestelltes Verfahren nutzt Stabile Feder-Masse-Modelle (abgekürzt SMSM von engl.: Stable Mass-Spring-Model) und deren integriertes Formwissen, um relevante Objekte zu identifizieren. Die Modelle konnten auf eine Reihe anatomischer Strukturen am Hals (wie z.B. Schildknorpel [4] und Lymphknoten [5]) erfolgreich angewendet werden.

Die Interaktion bei Segmentierungsverfahren wird in der Literatur selten thematisiert. Olabarriaga [7] untersucht Interaktionstechniken bei der Durchführung von Segmentierungen und empfiehlt, Benutzerinteraktionen bereits während eines Segmentierungsverfahrens durchzuführen. Im Falle dreidimensionaler modellbasierter Segmentierungstechniken gestattet jedoch die Performanz der Verfahren oftmals keine flüssige interaktive Steuerung des Vorgangs. Stattdessen müssen Verbesserungen der Resultate nachträglich vorgenommen werden. Kang et al. [2] beschreiben drei Verfahren, die auf Basis morphologischer Operationen die Struktur eines Segmentierungsergebnisses nachbessern. Proksch et al. [3] stellen Werkzeuge zur Korrektur medizinischer Segmentierungen vor, die durch Deformierung eines Polygonmeshes die Geometrie einer als Mesh vorliegenden Segmentierung direkt bearbeiten.

Alle genannten Techniken nutzen das in Form des Segmentierungsmodells in die Segmentierung eingebrachte Formwissen zur Nachkorrektur der Ergebnisse nicht aus. Daher erfordern sie oftmals komplexe Nutzereingaben und einen hohen Bearbeitungsaufwand für die Korrekturdurchführung oder liefern durch die Erzeugung von Diskontinuitäten in der Oberfläche des Segmentierungsergebnisses anatomisch nicht-plausible Ergebnisse. Die Entwicklung intuitiv bedienbarer Techniken, welche von in den Segmentierungsprozess integriertem Modellwissen Gebrauch machen, war daher das Ziel dieser Arbeit.

## 2 Methoden

Die hier konzipierten vier Korrekturwerkzeuge für modellbasierte Segmentierungsverfahren wurden auf der Grundlage einer Segmentierung mit *Stabilen Feder-Masse-Modellen (SMSM)* entworfen und umgesetzt. Dabei handelt es sich um physikalisch beschreibbare Formmodelle, bei denen eine Anzahl von Massepunkten durch elastische Federn miteinander verbunden ist. Verschiedene physikalische Parameter beschreiben die Eigenschaften des Modells und seiner Elemente, die in permanenter Wechselwirkung zueinander stehen. Die Dynamik des Modells entsteht durch die bekannte Wechselwirkung von *internen* und *externen Kräften*. Die internen Feder- und Torsionskräfte sollen die Formbewahrung im Rahmen der durch die Modellparameter kodierten zulässigen Variation gewährleisten, während die externen Kräfte, welche aus Bildinformationen generiert werden, das Modell zur gesuchten Struktur streben lassen. Bei modellbasierten Segmentierungen treten typischerweise zwei Ursachen für Segmentierungsfehler auf:

1. Ein Modell detektiert ein falsches lokales Extremum und segmentiert unerwünschte Objektgrenzen, oder
2. das Modell verfügt über unzureichendes Modellwissen, welches die gesuchte Struktur nicht optimal beschreibt.

Die entwickelten Korrekturmethode unterstützen eine Nachbesserung beider Fehlergruppen. Für die Fehler der Klasse 1 wird eine direkte Manipulation der Modellgeometrie unter Berücksichtigung des Formwissens ermöglicht, um das Modell vom falschen Extremum zur korrekten Strukturkante zu führen (Korrekturverfahren 1 und 2). Um Fehler der Klasse 2 korrigieren zu können, muss es dem Anwender möglich sein, zusätzliches Modellwissen in den Korrekturprozess zu integrieren – hierzu dienen die Korrekturverfahren 3 bzw. 4. Alle Verfahren verwenden für eine Korrektur das zur Segmentierung genutzte SMSM und verwenden so direkt das in dem Modell gespeicherte Formwissen über die jeweils gesuchte Struktur für eine Korrekturdurchführung aus. Hierin unterscheiden sich die Methoden von anderen in der Literatur vorgestellten Korrekturverfahren, wie etwa denen von Proksch et al. [3], die eine direkte Nachbearbeitung von als Polygonmodellen vorliegenden Segmentierungsergebnissen erlauben.

### 2.1 Einzelmassenverschiebung

Die erste Korrekturmethode ist die **Einzelmassenverschiebung**. Hierbei definiert der Nutzer Punkte auf nicht segmentierten Teilen der Strukturkante, indem er diese in einer Schichtansicht des Datensatzes anklickt (Abb. 1). Auf diese Punkte wird dann der jeweils nächstgelegene Massepunkt des zur Segmentierung genutzten SMSM verschoben und dort temporär fixiert. Anschließend werden Simulationsberechnungen der Modellbewegung unter Berücksichtigung der fixierten Massepunkte durchgeführt, bevor sie für eine zweite Simulation wieder freigelassen werden. Diese korrigierenden Simulationsläufe verbessern die Anpassung des Modells an die Bilddaten in Regionen, in denen der Nutzer Punkte gesetzt hat.



Abb. 1: Beispielhafte Korrektur eines fehlsegmentierten Lymphknotens durch Einzelmassenverschiebung  
Links: zu korrigierende Segmentierung, Mitte: Anklicken der Datensatzpunkte, Rechts: korrigierte Kontur



## 2.2 Massenbereichsverschiebung

Das Konzept der Einzelmassenverschiebung wurde bei der **Massenbereichsverschiebung** weiterentwickelt. Hier wird nicht ein einzelner Massepunkt auf den nächstgelegenen, durch den Nutzer angeklickten, Punkt umgesetzt. Stattdessen bestimmt die Methode auf Basis des Dijkstra-Algorithmus ein Massepunktnetz vordefinierter Größe, welches in Richtung des jeweiligen nutzerdefinierten Punktes verschoben wird. Der Vorteil gegenüber der Einzelmassenverschiebung besteht darin, dass die Gefahr, dass ein herausgezogener Massepunkt durch die inneren Modellkräfte einfach wieder in die Oberfläche des SMSM zurückgezogen wird, reduziert wird.

### 2.1.3 Lasso-Sensoren

Neben solchen Techniken, die für eine Korrektur eines Segmentierungsergebnisses direkt eine Änderung der Geometrie des jeweils genutzten SMSM herbeiführen, wurde ein Werkzeug entwickelt, welches dies indirekt durch die Einführung neuer Modellsensoren bewerkstelligt, die während des Segmentierungsprozesses auf die Massepunkte einwirkende Kräfte erzeugen. Diese Sensorart wird als **Lasso-Sensor** bezeichnet, da die Kräfte derart wirken, dass sie den Massepunkt, dem sie zugeordnet sind, zu sich heranziehen. Für die Verwendung der Sensoren klickt der Anwender in einer beliebigen Schichtansicht des Datensatzes auf Punkte in der Nähe des korrekten Verlaufs einer fehlsegmentierten Kante. An diesen werden dann Lasso-Sensoren erzeugt, die in einer korrigierenden Simulation der Modelldynamik die ihnen zugeordneten Massepunkte zu sich heranziehen und so eine verbesserte Anpassung des Modells an die Daten bewerkstelligen.

### 2.1.4 Torsionskraftverstärkung

Das vierte Korrekturwerkzeug macht sich die Wirkungsweise der Lasso-Sensoren zu Nutze. Bei dem **Torsionskraftverstärkungs-Werkzeug**, für das der Nutzer in der schon beschriebenen Art und Weise Punkte in den Bilddaten markiert (Abb. 2), wird an jedem dieser Punkte ein Lasso-Sensor erzeugt, welcher dem nächstgelegenen Modellpunkt zugewiesen wird. Zusätzlich wird für Nachbarmassepunkte, die den nächstgelegenen Massepunkt umgeben, sowie den Punkt selbst, die Torsionskonstante erhöht, bevor eine erneute Modellanpassung unter Berücksichtigung der neuen Parameter zur Korrektur berechnet wird.

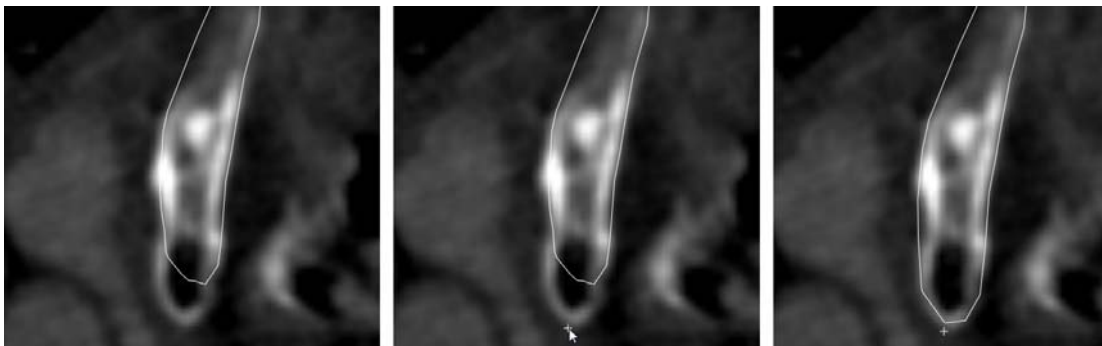


Abb. 2: Beispielhafte Korrektur durch Anwendung der Torsionskraftverstärkung:

Links: zu korrigierende Segmentierung, Mitte: Anklicken der Datensatzpunkte, Rechts: korrigierte Kontur

## 3 Ergebnisse

Die vorgestellten interaktiven Korrekturwerkzeuge wurden unter Nutzung der Programmiersprache C++ sowie verschiedener Softwarebibliotheken prototypisch implementiert.

In einer ersten Nutzerstudie wurde gezielt untersucht, wie leicht die Anwendung der Werkzeuge zu erlernen und zu beherrschen war, und wie schnell eine Korrektur mit ihnen vorgenommen werden konnte. Die objektive Qualität der Segmentierungsergebnisse stand dabei nicht im Vordergrund. Es standen vier Probanden zur Verfügung, von denen zwei mit den Korrekturtechniken vertraut waren, die anderen zwei Probanden waren fachfremd und hatten weder Erfahrungen mit den Korrekturwerkzeugen, noch mit medizinischen Bilddaten und Segmentierungen. Die Testnutzer waren nicht fachlich medizinisch ausgebildet.

Jeder Proband hatte die Aufgabe, vorberechnete SMSM-Segmentierungen von einem nekrotischen Lymphknoten und die rechte Platte dreier Schilddrüsennormale in vier verschiedenen CT-Datensätzen des Halsbereiches zunächst manuell

und anschließend unter Verwendung der vorgestellten Methoden zu korrigieren. Bei der Korrekturdurchführung wurden jeweils die benötigten Zeiten gemessen, sowie verschiedene subjektive Einschätzungen bzgl. der Anwenderfreundlichkeit und Effektivität der Korrekturwerkzeuge eingeholt.

Da modellbasierte Lymphknotensegmentierungen in der Regel bereits sehr genaue Ergebnisse liefern und deshalb meist wenig Korrekturbedarf besteht, wurde für diese erste Nutzerstudie lediglich ein nekrotischer Lymphknoten als Anwendungsfall ausgewählt. Bei diesem führt die Anwendung der SMSM aufgrund seiner besonderen Formeigenschaften zu Fehlsegmentierungen. Komplexere Strukturen, wie etwa der Schildknorpel, stellen eine größere Herausforderung für die SMSM dar. Hier kommt es etwa an den Hörnern häufig zu Untersegmentierungen, die korrigiert werden müssen. Aber es treten auch andere typische Arten von Fehlsegmentierungen auf, deren Ursache in den zwei oben dargestellten Problemklassen liegt und an denen die entwickelten Techniken getestet werden sollten. Es wurden deshalb drei unterschiedliche Schildknorpelsegmentierungen für die Evaluierung verwendet, die jeweils typische Segmentierungsfehler aufwiesen.

Es zeigte sich, dass alle vier entworfenen Techniken zur interaktiven Korrektur modellbasierter medizinischer Segmentierungen geeignet waren und einen raschen Einstieg in deren Anwendung ermöglichten. Außerdem konnten durch einen Einsatz der Werkzeuge von allen Probanden schneller die gewünschten Korrekturergebnisse erzielt werden als es die Nutzung des manuellen Verfahrens erlaubt (Tab. 1). In allen Testfällen gaben die Probanden an, dass die subjektive Übereinstimmung der korrigierten Segmentierung mit den jeweiligen Bilddaten zufriedenstellend war.

Struktur	Zeit (trainiert, in min)	Zeit (untrainiert, in min)	Zeit - manuell (trainiert, in min)	Zeit - manuell (untrainiert, in min)
Schildknorpel 1	1:30 - 2:23	1:04 - 1:35	22:53	15:25
Schildknorpel 2	1:27 - 2:47	1:26 - 2:31	15:42	8:04
Schildknorpel 3	1:22 - 2:23	2:42 - 3:16	8:54	4:30
Lymphknoten	0:28 - 0:47	0:38 - 0:54	3:50	4:38

Tabelle 1: Minimale und maximale durchschnittliche Zeiten der trainierten und untrainierten Probanden bei jedem Testfall für eine Korrektur unter Verwendung der Korrekturtechniken sowie durchschnittliche Zeiten der beiden Nutzergruppen für die manuelle Korrektur für jeden Testfall

## 4 Diskussion

Die vier interaktiven Techniken verfolgen eine neue Strategie zur Korrektur medizinischer Segmentierungen. Statt Werkzeuge zur Verfügung zu stellen, welche eine direkte Manipulation der Geometrie eines Segmentierungsergebnisses erlauben, integrieren die vorgestellten Techniken das über ein Formmodell in den Segmentierungsprozess eingebrachte Modellwissen in die Korrekturalgorithmen. So wird im Vergleich zur manuellen Korrektur weniger Zeit und Interaktion benötigt. Auch ungeübte Anwender lernten die Anwendung der Korrekturtechniken schnell und waren effektiv in der Lage, die gewünschten Ergebnisse zu erzielen. Segmentierungen müssen nicht mehr zeitaufwändig manuell nachgebessert werden, indem fehlsegmentierte Kantenzüge in jeder Schicht per Hand nachgezeichnet werden, sondern können durch einfaches Anklicken einiger Punkte in einer Schichtansicht des Datensatzes bearbeitet werden.

Im Ausblick ist besonders die Durchführung einer umfangreicheren Evaluierung mit medizinischem Fachpersonal interessant. Dabei ist zu untersuchen, inwiefern auch diese Zielgruppe ihre Korrekturwünsche effektiv auf die Segmentierungen übertragen kann und ob sich ebenfalls ein Zeitgewinn gegenüber der manuellen Korrektur einstellt. Zudem müsste die objektive Qualität der korrigierten Segmentierungsergebnisse anhand quantitativer Qualitätsmetriken genauer untersucht werden. Es ist außerdem von Interesse, welche der Korrekturwerkzeuge präferiert werden oder ob eine kombinierte Anwendung die besten Ergebnisse erzielt. Interessant ist außerdem eine vergleichende Evaluierung gegenüber direkten 3D-Objektmanipulationstechniken (wie etwa [3]), die keinerlei Formwissen berücksichtigen.

Die vorgestellten Korrekturwerkzeuge basieren auf der Segmentierung mit stabilen Feder-Masse-Modellen, jedoch sollte eine Übertragung auf andere kraftbasierte Formmodelle, wie etwa Active Contour Models, wegen der ähnlichen Funktionsweise der Modelldynamik, die ebenfalls durch Kraftwechselwirkungen generiert wird, leicht möglich sein.

## 5 Referenzen

- [1] L. Dornheim, K. Tönnies, J. Dornheim, Stable dynamic 3D shape models, IEEE International Conference on Image Processing, 2005
- [2] Y. Kang, K. Engelke, W. Kalender, Interactive 3D editing tools for image segmentation, Medical Image Analysis, 2004
- [3] D. Proksch, J. Dornheim, B. Preim, Interaktionstechniken zur Korrektur medizinischer 3D-Segmentierungen, Bildverarbeitung für die Medizin, 2010
- [4] J. Dornheim, L. Dornheim, B. Preim, I. Hertel, G. Strauss, Generation and Initialization of a Stable 3D Mass-Spring Model for the Segmentation of the Thyroid Cartilage, 28th Annual Symposium of the German Association for Pattern Recognition, 2006
- [5] J. Dornheim, H. Seim, B. Preim, I. Hertel, G. Strauss, Segmentation of Neck Lymph Nodes in CT Datasets with Stable 3D Mass-Spring Models, MICCAI 2006 Special Issue of *Academic Radiology*, 2007
- [6] J. Cordes, J. Dornheim, B. Preim and I. Hertel and G. Strauß, Preoperative Segmentation of Neck CT Datasets for the Planning of Neck Dissections, Proc. of SPIE Medical Imaging: Image Processing, 2006
- [7] S.D. Olabarriaga, A.W.M. Smeulders, Interaction in the segmentation of medical images: A survey, Medical Image Analysis, 2001



# Multimodale Bildregistrierung von fpVCT- und OCT-Daten zur Realisierung hochgenauer medizinischer Navigation

J. Díaz Díaz<sup>1</sup>, H. Eilers<sup>1</sup>, A. Niemann<sup>2</sup>, M. Leinung<sup>3</sup>, O. Majdani<sup>2</sup>, T. Ortmaier<sup>1</sup>

<sup>1</sup> Leibniz Universität Hannover, Institut für Mechatronische Systeme, Hannover, Germany

<sup>2</sup> Medizinische Hochschule Hannover, Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde, Hannover, Germany

<sup>3</sup> Klinikum der Johann Wolfgang Goethe-Universität, Klinik für Hals-Nasen-Ohrenheilkunde, Frankfurt am Main, Germany

Kontakt: [jesus.diazdiaz@imes.uni-hannover.de](mailto:jesus.diazdiaz@imes.uni-hannover.de)

## Abstract:

Medizinische Navigationssysteme stellen ein routinemäßig eingesetztes, intraoperatives Assistenzsystem dar. Moderne und insbesondere minimal invasive Operationstechniken stellen jedoch Anforderungen an die Genauigkeit der Lokalisation von Patient und Instrument, die von stereo-optischen oder elektro-magnetischen Navigationssystemen nicht mehr gewährleistet werden können. Durch die Integration eines hochauflösenden OCT-Sensors in den Lokalisationsprozess soll die Navigationsgenauigkeit für chirurgische Eingriffe an der seitlichen Schädelbasis erhöht werden. Der Kern dieses Navigationsprinzips ist die Registrierung von intraoperativ akquirierten OCT-Daten mit präoperativen CT-Daten. Dieser Beitrag stellt die Machbarkeit der multimodalen Registrierung vor. Es werden rigide Transformationen mit der „Nearest Neighbour“ Interpolation untersucht. Als Ähnlichkeitsmaß wird die in der multimodalen Bildregistrierung übliche „Mutual Information“ verwendet. Als stochastischer Optimierungsalgorithmus wird eine Partikel Schwarm Optimierung eingesetzt. Die qualitativen Ergebnisse zeigen eine hohe Güte der Registrierung.

Schlüsselworte: Cochlea Implantat, Navigation, Optische Kohärenztomographie, Multimodale Bildregistrierung

## 1 Problemstellung

Bei der Cochlea Implantat (CI) Operation handelt es sich um ein operatives Verfahren zur Behandlung von sensorischer Taubheit und an Taubheit grenzender Schwerhörigkeit. Im Rahmen der etablierten Operationstechnik zur Implantation eines CI wird in einem ersten Schritt das Felsenbein großvolumig abgetragen (Mastoidektomie), um im zweiten Schritt die sog. Cochleostomie, eine kleine Bohrung in die Wandung der Hörschnecke, durchzuführen. Die Cochleostomie sollte dabei möglichst mit einer Positionsabweichung von unter 0,5 mm ausgeführt werden [6]. Anschließend wird ein fadenförmiges Multielektrodenarray in die Gehörschnecke inseriert. Prinzipiell ist die Cochleostomie minimal invasiv, d.h. als direkter Stichkanal von dem äußeren Schädelknochen bis zur Cochlea, unter Einsatz geeigneter mechatronischer Assistenzsysteme möglich. Dieser Ansatz beinhaltet die Verwendung eines Navigationssystems um einen chirurgischen Bohrer mit Hilfe eines Roboters gemäß einer Operationsplanung zu führen. Bisherige Ergebnisse zeigen jedoch, dass die geforderte Genauigkeit mit einem herkömmlichen Navigationssystem nicht gewährleistet werden kann [2].

Um die Navigationsgenauigkeit zu erhöhen, erfolgt daher die Integration von Optischer Kohärenztomographie (OCT) als zusätzliche Bildgebungseinheit in das mechatronische Assistenzsystem. OCT ermöglicht die nicht-invasive Akquise von drei-dimensionalen und hochauflösenden Bilddatensätzen. Ziel dieses Ansatzes ist, einen OCT-Sensor mechanisch in der Nähe des Roboter-Endeffektors anzubringen, um intra-operative Datensätze des Situs zu erzeugen. Basierend auf einer Registrierung von OCT- und präoperativen CT-Daten können anschließend Korrekturen des Roboterarms errechnet und somit das ebenfalls vom Roboter geführte chirurgische Instrument hochgenau positioniert werden. Die prinzipielle Machbarkeit der Bildregistrierung stellt den Kern dieses Beitrags dar.

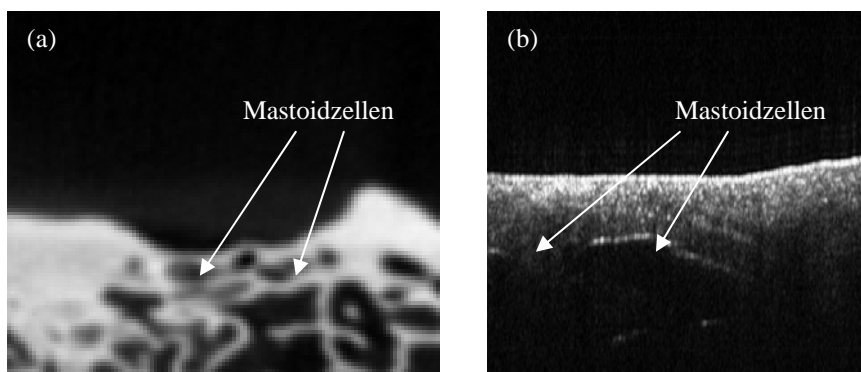
## 2 Methoden

OCT ist eine Bildgebung, die sich durch die Fähigkeit hochauflösende, tomographische Schnittbilder in Echtzeit zu erzeugen, auszeichnet. Sie ist vor allem dann interessant, wenn nicht-invasiv und zerstörungsfrei Tiefeninformationen einer Probe gewonnen werden sollen. Die prinzipielle Funktionsweise eines OCT basiert auf einem Michelson-Interferometer, bei der zurückreflektiertes Licht aus einer Probe mit einer Referenz interferiert wird. Reflektion findet

in einer Probe an Inhomogenitäten und damit insbesondere an Materialübergängen statt. Die Qualität und Eindringtiefe hängt damit stark von dem verwendeten Material ab; bei biologischen Proben beträgt die maximale Eindringtiefe ca. 2 mm [1]. Aufgrund der hohen Auflösung und dem beschränkten Arbeitsbereich eignet sich OCT insbesondere für die Bildgebung oberflächennaher Mikrostrukturen. Da die aktuelle Forschung die Systeme zunehmend auflösungsstärker und hochfrequenter in der Abtastung machen, ist zu erwarten, dass ein größeres Volumen einer biologischen Probe in-vivo und in Echtzeit mit OCT zu visualisieren ist.

## 2.1 Daten

Die diesem Beitrag zugrunde liegenden OCT-Daten wurden von einem Time Domain (TD) OCT [1] aufgenommen. Der abgebildete Arbeitsbereich wurde auf einen Bereich von 4 mm x 3 mm x 6 mm beschränkt, wobei 3 mm die maximale Scantiefe (in dem Medium Luft) beschreibt. Die räumliche Auflösung beträgt  $8,3 \mu\text{m} \times 6,3 \mu\text{m} \times 25,0 \mu\text{m}$ . Die CT-Bildgebung wurde mit einem „flat panel volume computerised tomography“ (fpVCT, General Electric, Schenectady, NY, USA) realisiert, da hiermit eine höhere örtliche Auflösung gegenüber konventionellen CTs realisierbar ist. Nach einer isotropen Rekonstruktion erhält man so eine räumliche Auflösung von  $0,148 \text{ mm} \times 0,148 \text{ mm} \times 0,148 \text{ mm}$ . Der Arbeitsbereich beinhaltet den Arbeitsbereich des OCT.



**Abb. 1: Ausschnitte eines (a) fpVCT- (17,96 mm x 14,99 mm) und (b) OCT-Datensatzes (4 mm x 3 mm, letztere Angabe im Medium Luft) eines humanen Felsenbeinpräparats**

Der abgebildete Bereich ist ein Teil des Felsenbeins eines Leichenpräparats. Um die im Felsenbein vorhandenen und in Abb. 1 dargestellten Strukturen im Knochen für die OCT-Bildgebung sichtbar zu machen, wurde die Oberfläche des Felsenbeins geringfügig angefräst. Im Vergleich zu einer kompletten Mastoidektomie handelt es sich hierbei nur um einen Bruchteil an Knochengewebe, das entfernt wurde. Dies ist aufgrund der hohen Dichte des Schädelknochens bzw. der geringen Eindringtiefe des OCT Laserstrahls in Knochen notwendig. Zu sehen ist in Abb. 1 das sog. Mastoid des Felsenbeins mit den charakteristischen Kavitäten im Knochen, den sog. Mastoidzellen.

Die Registrierungsgüte hängt im Besonderen von der Güte und der Vergleichbarkeit der Daten ab. Je größer der Informationsgehalt in den einzelnen Bildern, desto wahrscheinlicher lassen sich die Bilder korrekt registrieren. Notwendig ist aber auch, dass die gleichen Informationen in den Bildgebungen dargestellt werden. Dies ist nicht unmittelbar gegeben, da in einem CT-Bild die Grauwerte den *Absorptionsgrad* des Gewebes repräsentieren während sie in einem OCT-Bild für das *Reflektionsverhalten* des Gewebes stehen. Bei den Übergängen von Luft zu Knochen bzw. von Knochen zu Gewebe treten jedoch sowohl Materialinhomogenitäten als auch hohe Absorptionsgradienten auf. Folglich korrelieren die Positionen der Kanten in den beiden Bildgebungen miteinander.

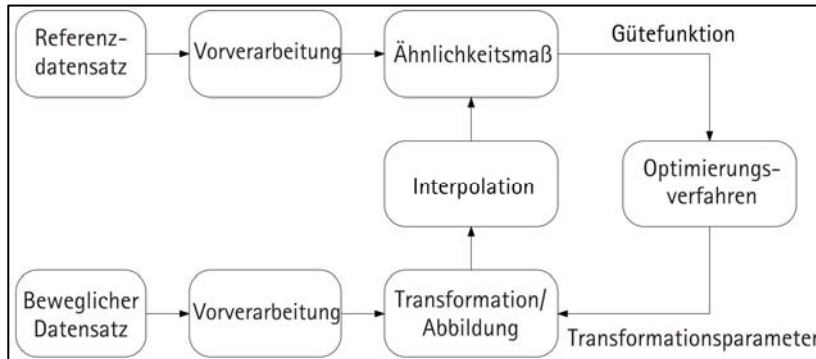
## 2.2 Registrierung

Registrierung beschäftigt sich mit der Suche nach der Transformation  $T$ , die den räumlichen Zusammenhang zwischen zwei Datensätzen herstellt. Die Einschränkung der Transformation  $T$  auf eine bestimmte Klasse von Abbildung ist Teil der guten Konditionierung der Problemstellung. In diesem Beitrag beschränken wir uns auf rigide Transformationen  $T$ , die durch einen sechs-dimensionalen Vektor  $x$  parametrisiert sind, für den gilt:

$$x = (t_x, t_y, t_z, \varphi_x, \varphi_y, \varphi_z) =: (x_t, x_\varphi).$$

Die Parameter  $x_t$  stellen dabei die Komponenten einer Translation dar, während  $x_\varphi$  die rotatorischen Komponenten in Euler-Winkel repräsentieren. Es wird nach einer der üblichen Konventionen um die Achsen  $z$ ,  $y'$ , und  $z''$  rotiert.

Die Einschränkung auf rigide Transformationen ist bei den gegebenen Datensätzen zulässig, da elastische Verformungen im Kadaverpräparat nicht zu erwarten sind und statische Effekte in einer Vorverarbeitung der Daten kompensiert werden können. Wendet man eine rigide Transformation auf einen Datensatz an, so besteht die Notwendigkeit der Interpolation, d.h. der Ausrichtung der transformierten Voxel an den Voxeln des Referenzdatensatzes. Die für diesen Beitrag implementierte Methode ist die „Nearest Neighbour“-Methode. Die damit erzielbaren Registrierungsgenauigkeiten sind im Vergleich zu anderen Interpolationsmethoden gering, jedoch für unsere Zwecke ausreichend.



**Abb. 2: Schematische Darstellung eines Registrierungsprozesses**

In Abb. 2 ist eine schematische Darstellung des gesamten Registrierungsprozesses zweier Datensätze  $D_i$  – Referenzdatensatz und transformierter beweglicher Datensatz – dargestellt. Ziel einer Registrierung ist die Wahl der Transformation, die das Ähnlichkeitsmaß maximiert. Da es sich in diesem Beitrag um Datensätze zweier unterschiedlicher Bildgebungen handelt – man spricht deshalb von *multimodaler* Bildregistrierung –, wurde die „Mutual Information“ [5] im Algorithmus implementiert, die sich wie folgt berechnet:

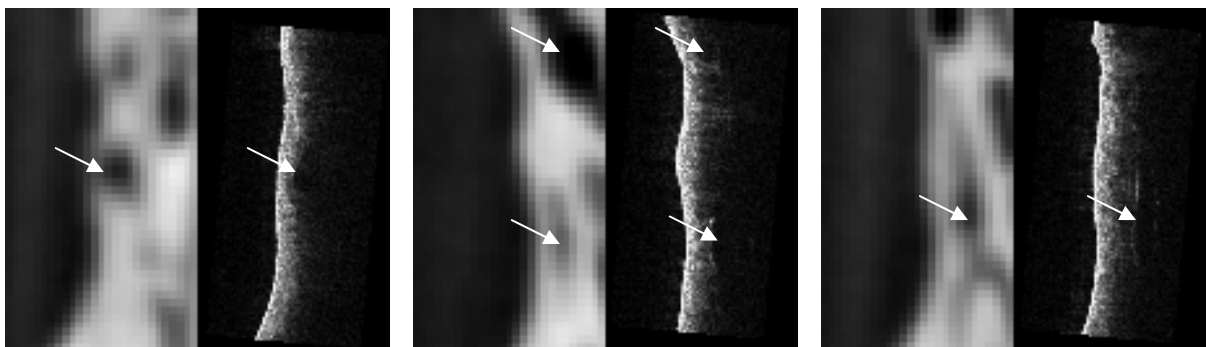
$$MI(D_1, D_2) = H(D_1) + H(D_2) - H(D_1, D_2).$$

„Mutual Information“ ist ein in der multimodalen Bildregistrierung übliches Ähnlichkeitsmaß, welches im Wesentlichen auf der Minimierung der (Shannon) Entropie  $H$  des gemeinsamen Histogramms  $H(D_1, D_2)$  basiert. Diese wird genau dann minimiert, wenn mittels Wahrscheinlichkeiten Korrespondenzen zwischen den Grauwerten der beteiligten Daten aufgestellt werden können. „Mutual Information“ ist ein intensitätsbasiertes Ähnlichkeitsmaß, dessen Vorteil die direkte Evaluation der Daten beinhaltet, wodurch eine Merkmalsdetektion umgangen wird. Die Anforderungen an die dem Ähnlichkeitsmaß übergebenen Daten sind damit sehr gering.

Ebenso wichtig wie das Ähnlichkeitsmaß ist die richtige Wahl des Optimierungsalgorithmus. In der multimodalen Bildregistrierung ist a priori die zu optimierende Funktion, die sog. Gütefunktion, hochgradig nicht konvex. Dadurch eignen sich deterministische Optimierungsalgorithmen nur bedingt, um das globale Maximum zu bestimmen. Es wurde daher auf eine stochastische und dadurch rechenintensive Partikel-Schwarm Optimierung (PSO) zurückgegriffen, die erstmalig von Kennedy und Eberhardt [4] benutzt wurde.

### 3 Ergebnisse

Der vorgestellte Algorithmus wurde implementiert und auf die vorliegenden Rohdatensätze angewendet. Die Grobregistrierung erfolgte anhand von künstlichen Landmarken an der Oberfläche des Felsenbeins und lieferte die Startparameter des Registrierungsprozesses. Die Genauigkeitsschätzung der Grobregistrierung lieferte die Schranken für die Optimierung und wurde mit  $\pm 2$  mm in jeder translatorischen und  $\pm 2^\circ$  in jeder rotatorischen Komponente bestimmt.



**Abb. 3: Gegenüberstellung von 3 ausgewählten Schichten eines fpVCT-Datensatzes (jeweils links) mit den korrespondierenden transformierten Original OCT-Datensätzen (jeweils rechts)**

Da mit Originaldaten gearbeitet wurde, fehlt der sog. Goldstandard. Damit ist eine Evaluation der optimierten Transformation und die Bestimmung des Transformationsfehlers nicht ohne weiteres möglich. Die Überprüfung der Transformationsgüte erfolgt somit zunächst rein qualitativ anhand der Gegenüberstellung der Datensätze. Wie in Abb. 3 anhand der Oberfläche und Mastoidzellen zu sehen ist, stimmt die Lage der Oberfläche in beiden Datensätzen gut überein. Sowohl die Lage als auch Größe und Form lassen den Schluss zu, dass eine zufrieden stellende Transformation erzielt wird. Die geringe Qualität des OCT-Datensatzes, insbesondere die begrenzte Eindringtiefe, verhindert jedoch eine bessere Gegenüberstellung.

Um die Ergebnisse weiter zu evaluieren soll ein Vergleich zu (leicht geänderten) in [3] erzielten Ergebnissen erfolgen. In der dortigen Veröffentlichung basierte die Registrierung auf originalen VCT-Daten und unter anderem mittels eines Kaskadenfilters, bestehend aus Median- und Mittelwertfilter, vorverarbeiteten und anschließend segmentierten OCT-Daten. Der Vergleich soll zum einen anhand der Differenz der translatorischen Komponenten

$$e_x = \|x_{t,1} - x_{t,2}\|_2 \quad \text{für} \quad x_i = (t_{x,i}, t_{y,i}, t_{z,i}, \varphi_{x,i}, \varphi_{y,i}, \varphi_{z,i}) =: (x_{t,i}, x_{\varphi,i})$$

erfolgen, wobei  $i=1$  für die Benutzung der segmentierten OCT-Daten, während  $i=2$  für die Benutzung der unsegmentierten Daten steht. Zum anderen soll die rotatorische Differenz  $e_\varphi$  durch den Winkel in der Achsen-Winkeldarstellung der Rotationsmatrix  $R_1 R_2^{-1}$  bestimmt werden, wobei  $R_i$  die Rotationsmatrix von  $T_i$  ist. Somit gilt

$$e_\varphi = \arccos\left(\frac{\text{Spur}(R_1 R_2^{-1}) - 1}{2}\right).$$

Diese Ausdrücke vermitteln einen guten Eindruck über den Unterschied der Ergebnisse der beiden Registrierungsverfahren. Es gilt  $e_x = 0,182$  mm bzw.  $e_\varphi = 2,298^\circ$ . Somit ist der Unterschied (zumindest in der translatorischen Komponente) ca. das 1,25 fache der fpVCT-Auflösung und damit relativ hoch. Er lässt sich jedoch durch eine kleine Änderung der OCT-Daten bzw. des Ähnlichkeitsmaßes stark mindern, nämlich unter der Annahme, dass ab einer festen Tiefe zur detektierten Oberfläche die Voxel aufgrund der geringen Eindringtiefe des OCT in der Registrierung nicht berücksichtigt werden. Der (translatorische) Unterschied lässt sich dann auf das 0,56 fache der fpVCT-Auflösung reduzieren und liegt dann bei  $e_x = 0,083$  mm bzw.  $e_\varphi = 0,633^\circ$ . Die in [3] noch als notwendig aufgefasste rechenzeitaufwendige Segmentierung der Daten ist somit a priori für die Registrierung der Daten überflüssig.

## 4 Diskussion

Die Idee einer hohen Navigationsgenauigkeit basierend auf der Integration eines OCT in ein bestehendes Navigationssystem zur Verfügung zu stellen, macht die Registrierung von hochauflösenden OCT-Daten mit präoperativen CT-Daten notwendig. Der Beitrag setzt sich mit dieser grundsätzlichen multimodalen Registrierungsproblematik auseinander und zeigt die Machbarkeit der Registrierung von originalen Daten. Auch wenn ein – nachzuholender – Vergleich zum Goldstandard fehlt, wird qualitativ eine Genauigkeitsaussage getroffen und quantitativ die Ergebnisse mit denen aus [3] verglichen. Es wird qualitativ eine zufrieden stellende Registrierung erzielt, die es in folgenden Beiträgen quantitativ zu bewerten gilt. Des Weiteren wird in folgenden Veröffentlichungen eine Aussage hinsichtlich der Robustheit dieses Verfahrens getroffen werden müssen. Die Autoren danken der Deutschen Forschungsgemeinschaft für die Förderung im Rahmen des Projektes „Einsatz der OCT-Bildgebung zur medizinischen Nahfeldnavigation“ und allen weiteren Institutionen und Personen, die das Projekt unterstützen.

## 5 Referenzen

- [1] Drexler, W., Fujimoto, J.G., Optical Coherence Tomography, Springer, 2008
- [2] Eilers, H. et al., Navigated, robot assisted drilling of a minimally invasive cochlear access, Proceedings of the 2009 IEEE International Conference on Mechatronics, IEEE, 2009
- [3] Eilers, H. et al., Multimodal image registration of VCT and OCT images: a step towards high accuracy in surgical navigation, International Journal of Computer Assisted Radiology and Surgery, 2009
- [4] Kennedy, J., Eberhardt, R., A New Optimizer Using Particles Swarm, Proceedings of the 6<sup>th</sup> International Symposium on Micro Machine and Human Science, 1995
- [5] Pluim, J.P.W., Mutual Information based Registration of Medical Images, 2000
- [6] Schipper, J. et al., Navigation-controlled cochleostomy. Is an improvement in the quality of results for cochlear implant surgery possible?, HNO, 2004



# Grenzflächendetektion im Felsenbein

T. Klenzner<sup>1</sup>, L.A. Kahrs<sup>1</sup>, R. Huber<sup>2</sup>, J. Raczkowsky<sup>3</sup>, H. Wörn<sup>3</sup>, J. Schipper<sup>1</sup>

<sup>1</sup> Universitätsklinikum Düsseldorf, Univ.-HNO-Klinik, Deutschland

<sup>2</sup> Ludwig-Maximilians-Universität München, Lehrstuhl für BioMolekulare Optik, Deutschland

<sup>3</sup> Karlsruher Institut für Technologie (KIT), Institut für Prozessrechentechnik, Automation und Robotik, Deutschland

Kontakt: Thomas.Klenzner@med.uni-duesseldorf.de

## Abstract:

Die seitliche Schädelbasis enthält sehr relevante, vital notwendige Strukturen, u.a. die innere Halsschlagader, der Gesichtsnerv, das Innenohr mit Gleichgewichts- und Hörorgan sowie Anteile des Geschmacksnervs. Zur Behandlungen von Pathologien in diesem Bereich sind präzise Interventionen für ein minimales Trauma beim Patienten gefordert. Bei solchen mikrochirurgischen Eingriffen ist ein definierter und vollständiger Knochenabtrag im Submillimeterbereich mit Schonung der darunter liegenden Nerven, Gefäße und Organe zwingend erforderlich. Als Anforderungen an ein neuartiges System ergeben sich die Detektion von Gewebeübergängen sowie der Stopp des Knochenabtrags vor dem Erreichen bzw. der Schädigung angrenzender Weichgewebestrukturen. Lösungsansätze bestehen in der Regelung des Knochenabtrags durch automatische Grenzflächendetektion mittels Videoüberwachung und/oder optischer Kohärenztomographie sowie Bildverarbeitung. Das System soll beispielhaft anhand der Eröffnung des Innenohrs (Cochleostomie) evaluiert werden.

**Schlüsselworte:** Schädelbasischirurgie, Cochleostomie, optische intraoperative Messmethoden, Bildverarbeitung, Laserknochenablation

## 1 Problem

Betrachtet man die chirurgische Anatomie im Bereich der Otobasis, so sind die meisten wichtigen Organe, Nerven oder Gefäße von Knochen umgeben. Beispielhaft sind hier das Innenohr, der Nervus Facialis für die Gesichtsmuskulatur, die Chorda Tympani als sensorische Afferenz für den Geschmack oder auch die großen venösen und arteriellen Blutleiter zu nennen. In den letzten Jahren besteht zunehmend die Absicht, minimalinvasive Operationsverfahren und eine navigierte Instrumentenführung zu einer Erhöhung chirurgischer Präzision im Bereich der Schädelbasis zu erreichen; diesbezüglich werden auch mechatronische Assistenzverfahren - teils experimentell – für den HNO Bereich entwickelt [1]. Allerdings ist es den bestehenden Systemen nicht möglich, eine Ortsauflösung im Submillimeterbereich zu erreichen, da das zugrunde liegende 3D-Modell des Patienten der klinisch eingesetzten bildgebenden tomographischen Verfahren (CT, MRT, Volumetomographie) mit der dazugehörigen Segmentierung nicht die entsprechende Auflösung hat [2,3]. Da aus ethischen Prinzipien vom Chirurgen eine möglichst minimale Traumatisierung des Gewebes gefordert werden muss, kommt einer Grenzflächendetektion als möglichem Sensorikansatz in der mechatronisch assistierten Chirurgie eine enorme Bedeutung zu. Als Grenzfläche wird die Gewebestruktur bezeichnet, die unmittelbar nach einer kompletten Entfernung des über ihr liegenden knöchernen Gewebes dargestellt ist. Der Erhalt dieser Gewebeformation ist für den Funktionserhalt des entsprechenden Gewebes von immenser Wichtigkeit, da diese, wenn nicht schon unmittelbar der Funktion dienend, so doch ernährend oder schützend wirkt. Hier sind als mögliche Schäden beispielsweise der Verlust des Hör- oder Gleichgewichtsorgans, eine Gesichtsnervenlähmung, Hirnblutung, Hirnwasseraustritt oder Geschmacksirritationen zu nennen.

Um einen definierten Knochenabtrag im Bereich der Schädelbasis zu untersuchen, wurde in unseren Experimentalansatz ein gepulster CO<sub>2</sub> Laser zur Eröffnung des Innenohres (Cochlea) verwendet. Dies ist ein Operationsschritt im Rahmen der Cochlea Implantation. Der Laserstrahl wird hierbei über einen Laserstrahlscanner gesteuert [4,5]. Hiermit war die Planung individuell angepasster Abtragungsmuster im Mikrometerbereich im Felsenbein möglich. Der CO<sub>2</sub>-

Laser bietet ein gutes Strahlprofil ( $TEM_{00}$ ) und eine gute Fokussierbarkeit ( $\varnothing < 250 \mu m$ ). Über die fein skalierbare Pulsenergie ist eine sehr gute Kontrollierbarkeit der Dicke der einzelnen Abtragschichten möglich. Zur Regelung der Laserablationsmuster und Abtragtiefe sind Verfahren zu einer sicheren Detektion von biologischen Grenzflächen erforderlich. Mögliche Ansätze werden im Folgenden beschrieben. Ein weiterführendes Ziel ist, die systematische Erfassung der biologischen Grenzflächen durch die zu entwickelnde Sensorik auf das Laserablationsverfahren zu übertragen und den Abtrag damit zu regeln. Beispielsweise würden entsprechend als Grenzfläche erkannte Gebiete des OP-Situs im nächsten Ablationszyklus ausgespart werden können. Der Chirurg würde hierfür einen entsprechend Vorschlag über die Software des Lasersystems erhalten mit der Möglichkeit der Modifikation des vorgeschlagenen Ablationsmusters.

## 2 Methoden

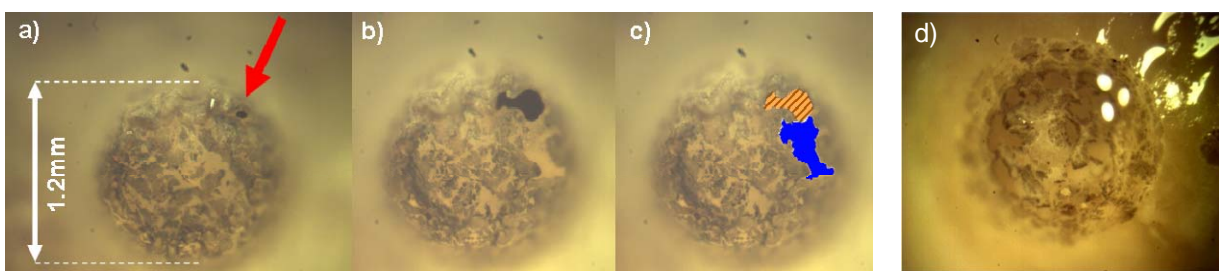
Prinzipiell können optische, akustische, haptische oder radiologische Methoden zur Grenzflächendetektion beim Übergang von knöcherner zu nicht-knöchernen Strukturen Verwendung finden. Als initiales Anwendungsbeispiel dient die Präparation der membranösen Auskleidung des Innenohrs während der Cochleostomie. Unsere beiden Ansätze sind optisch, um eine Integration in den Strahlengang der Laserablation zu ermöglichen.

Die erste Methode beruht auf einer mikroskopischen Videoüberwachung in Kombination mit Bildverarbeitungsmethoden. Freipräparierte Strukturen sollen anhand von einer automatischen Bildmustererkennung identifiziert werden. Für die Segmentierung dieser Regionen hat sich der Color Structure Code (CSC) als geeignet herausgestellt. Die segmentierten Bereiche stellen die Grenzflächen dar und dienen als Grundlage für die weitere Ablationspulsregelung. Der Vorteil dieses Ansatzes liegt in der vergleichbar einfachen Einbindung in den Operationsablauf. Komponenten dieses Systems sind in heutigen Operationssälen vorhanden (Mikroskop, Kamera, Laser) [6].

Die optische Kohärenztomographie (OCT) als weiterer methodischer Ansatz ermöglicht zudem eine Tiefendarstellung des Gewebes auf der Basis reflektierter Laserstrahlen. Die prinzipielle Abbildbarkeit von Knochen oder Innenohrstrukturen wurde bereits in anderen Forschungsarbeiten gezeigt [7,8]. Unser Ansatz zielt dabei auf die Kombination der Laserablation mit dem zusätzlichen dreidimensionalen Abrastern des Ablationsgebiets mit einem OCT-Messstrahl ab. Die Messergebnisse müssen dann dreidimensional rekonstruiert und mittels Bildverarbeitung die Grenzflächen erkannt werden.

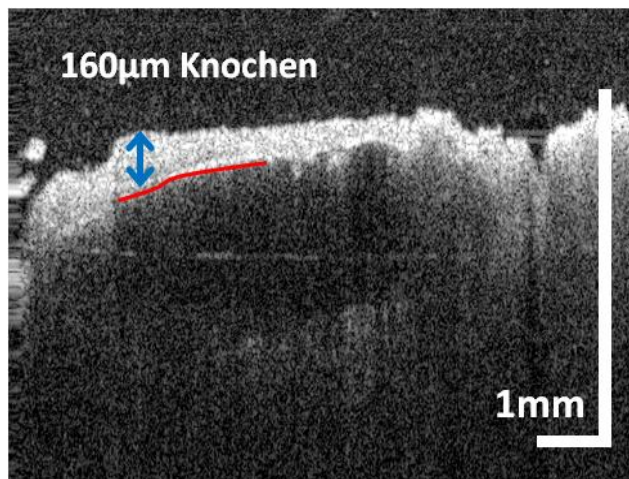
## 3 Ergebnisse

Im Gegensatz zu einer konventionellen Eröffnung des Innenohres mittels einer rotierenden Fräse, wo bei Mikroskopieaufnahmen das Endost als Grenzschicht deutlich identifizierbar ist, ist dieses bei Lasercochleostomien bisher nicht durchgehend gelungen (vgl. Abb. 1a). Bei den meisten Experimenten konnte die Ablation erst nach dem ersten Laserpuls gestoppt werden, welcher durch die Membran oder Perilymphe absorbiert wurde. Sind größere Bereiche der Membran (Endost) und des eröffneten Bereichs freigelegt, funktionierte die Segmentierung stabil (vgl. Abb. 1b/c). In späteren Experimenten gelang die Adaptation der Laserpulsmuster und der Reduzierung der Pulsenergie, um eine bis auf mit geringen Knochenfragmenten bedeckte Membran darzustellen (vgl. Abb. 1d).



**Abb. 1:** Videoüberwachungsaufnahmen während einer Lasercochleostomie. a) Ein erster (mit der Methode der Videoüberwachung oft nicht vermeidbarer) Laserpuls, der durch die Grenzschichten hindurch in die Cochlea appliziert wird. b) Darstellung eines Teils der membranösen Auskleidung unterhalb der Eröffnung. c) Automatisch segmentierte Regionen von Membran (blau) und Eröffnung (gestreift). d) Membran ca. zu einem Drittel freigelegt.

Die optische Kohärenztomographie ermöglicht hingegen ein Premonitoring, bei dem die Restknochendicke bekannt ist, die über der kritischen Struktur liegt. Unsere bisherigen Experimente zeigen, dass mindestens 150µm vor einer Perforation die Lage der Grenzfläche bekannt ist (vgl. Abb. 2). Die Proben wurden mit einem selbstgebauten OCT-System untersucht [9]. Außerdem wird es möglich sein, die Positionen der nächsten Ablationslaserpulse, die nach der OCT-Messung optimal verteilt werden, nicht nur wie bei der Videoüberwachung auf Basis von zweidimensionalen Informationen sondern dreidimensional im verbleibenden Restknochenvolumen zu setzen.



**Abb. 2:** Darstellung von Strukturen an einer humanen Cochlea mittels optischer Kohärenztomographie. In dem exemplarischen 2D-Schichtbild kann der Knochen mit einer Dicke von mindestens 160µm dargestellt werden. Entscheidend ist die Identifizierbarkeit einer klar erkennbaren Strukturunterkante, die für eine automatische Segmentierung und somit für die Grenzflächendetektion während der Laserknochenablation verwendet werden kann.

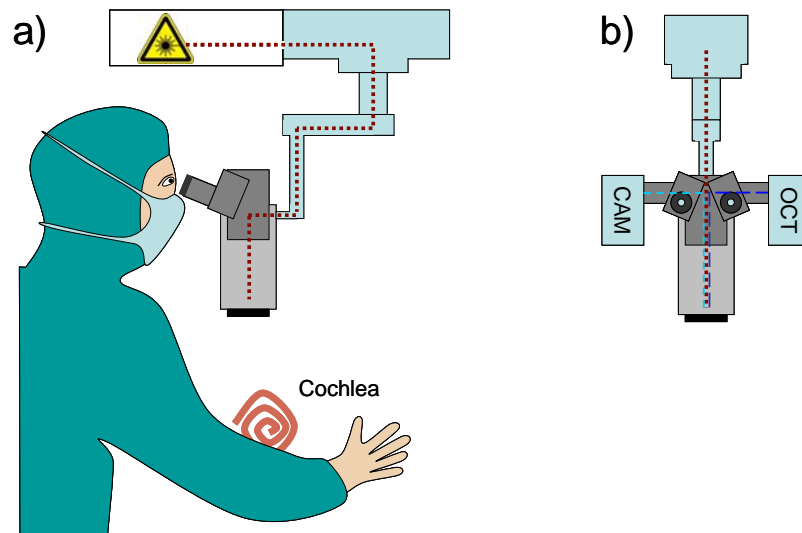
## 4 Diskussion

In der klinischen Routine bei Manipulationen an der Schädelbasis oder Innenohr finden bisher keine navigiert kontrollierten, reproduzierbaren Prozeduren für einen Knochenabtrag im Submillimeterbereich Verwendung. Als Konsequenz ist ein entsprechender Funktionsverlust beispielsweise des Hörorgans oder Nervenschäden möglich.

Für eine Detektion der Grenzfläche zwischen Knochen und Weichgewebe (Knochenhaut, Nerven, Blutgefäße, etc.) sind verschiedene Messmethoden möglich. Zwei optische Verfahren wurden anhand von Messergebnissen bei Experimenten am Innenohr in diesem Beitrag beschrieben. Die optische Kohärenztomographie zeigt sich in unseren initialen Untersuchungen der normalen Videoüberwachung bezüglich der Vorhersehbarkeit des Erreichens einer Grenzfläche bei der Laserablation überlegen. Untersuchungen bzw. Optimierung der Systeme bezüglich der maximal möglichen Eindringtiefe in Knochen sowie der automatischen Erkennung der Grenzfläche mittels Bildverarbeitung bei der optischen Kohärenztomographie stehen noch aus. Bezüglich notwendiger Redundanz können eventuell später beide Verfahren (Videoüberwachung und OCT) parallel eingesetzt werden. Dem Operateur stünde durch die hier beschriebene Kombination der Grenzflächenerkennung mit der Laserknochenablation ein neues Werkzeug zur Verfügung, welches nach erfolgtem Ausrichten auf die Zielstruktur eine Ein-Knopf-Lösung für das automatische Freipräparieren von Weichgewebestrukturen unter einer Knochenschicht ermöglichen würde. Die Abbildung 3 zeigt einen möglichen schematischen Aufbau für einen experimentellen Ansatz im Operationssaal mit einem angestrebten coaxialen Strahlengang unter Verwendung eines Operationsmikroskops. Zur Evaluation des Systems müssen aus ärztlicher Sicht noch Untersuchungen durchgeführt werden, die u.a. Sicherheitsaspekte wie Ausschluss von thermischen und akustischen Schäden durch die verwendeten Laser beinhalten oder die Integration in einen OP Ablauf ermöglichen.

## 5 Danksagung

Die Autoren danken der Deutschen Forschungsgemeinschaft, die diese Arbeit im Rahmen des SPP 1124 mit dem Projekt „Bildgestützte Navigation eines Laserstrahls zur mikrochirurgischen Knochenablation an der lateralen Schädelbasis“ unterstützt hat sowie im laufenden Normalverfahren „Optische Kohärenztomographie für eine geregelte Laserablation an Grenzschichten der Schädelbasis (OCT-LABS)“ unterstützt (Geschäftszeichen: KL 2113/1-1, KL 2113/3-1).



**Abb. 3:** Schematische Darstellung einer möglichen Integration der Grenzflächendetektion mittels Videoüberwachung (CAM) bzw. optischer Kohärenztomographie (OCT) in ein Operationsmikroskop für die Laserknochenablation. Abbildung a) zeigt die Seitenansicht und b) die Sicht von vorne auf das Mikroskop.

## 6 Referenzen

- [1] Bumm K, Federspil PA, Klenzner T, Majdani O, Raczkowsky J, Strauss G, Schipper J; ASKRA (Arbeitsgemeinschaft für Schädelbasis- und kraniofaziale Chirurgie der Dtsch. Ges. f. HNO, Kopf- u. Halschirurgie): Update in der navigiert kontrollierten und mechatronisch assistierten Kopf-Hals-Chirurgie in Deutschland. HNO. 56(9):908-15, 2008
- [2] Caversaccio M, Freysinger W: Computer assistance for intraoperative navigation in ENT surgery, Minim Invasive Ther Allied Technol. 12(1):36-51 . 2003
- [3] Schipper J, et al.: Navigation as a quality management tool in cochlear implant surgery. J Laryngol Otol.. 118: 764-70, 2004
- [4] Klenzner T, Aschendorff A, Arapakis I, Klenzner T, Teszler CB, Ridder GJ, Laszig R: High precision cochleostomy by the use of a pulsed CO<sub>2</sub>-Laser – an experimental approach, Cochlear Implants International 10 S1:58-62, 2009
- [5] Kahrs LA, Werner M, Knapp FB, Lu SF, Raczkowsky J, Schipper J, Ivanenko M, Wörn H, Hering P, Klenzner T: Video Camera Based Navigation of a Laser Beam for Micro Surgery Bone Ablation at the Skull Base - Setup and Initial Experiments. Advances in Medical Engineering, 219-23, 2007
- [6] Kahrs LA: Bildverarbeitungsunterstützte Laserknochenablation am humanen Felsenbein, KIT Scientific Publishing, 2009
- [7] Fujimoto JG: Laser Medicine and Biomedical Imaging, in Progress Report of the Research Laboratory of Electronics (RLE) at the Massachusetts Institute of Technology (MIT), No. 148 (12), 2005-2006
- [8] Pau HW, Lankenau E, Just T, Behrend D, Hüttmann G: Optical coherence tomography as an orientation guide in cochlear implant surgery?, Acta Otolaryngol. 127(9):907-13, 2007
- [9] Huber R, Wojtkowski M, Fujimoto JG: Fourier Domain Mode Locking (FDML): A new laser operating regime and applications for optical coherence tomography, Optics Express 14(8):3225-37, 2006

# Ultrasound Segmentation in Navigated Liver Surgery

Sylvain Anderegg, Matthias Peterhans, Stefan Weber

*University of Bern, Institute for Surgical Technology & Biomechanics (ISTB), Bern, Switzerland*

Contact: [matthias.peterhans@istb.unibe.ch](mailto:matthias.peterhans@istb.unibe.ch)

## **Abstract:**

*In computer assisted liver surgery, the use of pre-operative 3D computer tomography (CT) images provides basic orientation and valuable information about the patient liver. During the surgical intervention, the accuracy of this information is reduced by motion and deformation of the organ. The fusion of intra-operative ultrasound (US) imaging with the pre-operative data is the next step in order to improve this situation. As pre-requisite, the identification of corresponding structures in US and CT (such as blood vessels) is required. Within this paper, the integration of an ultrasound vessel segmentation algorithm in a navigation system for liver surgery is presented. Initial results obtained on patients undergoing liver resection are evaluated.*

*Keywords: Liver surgery, Ultrasound segmentation, Computer assisted surgery*

## **1 Problem**

Intra-operative ultrasound imaging is used during surgical interventions on the open liver, providing real-time information about the organ inner structure. Moreover, new tumors may be detected and the surgical intervention shall eventually be adapted on site. Within our navigation system for liver surgery [1], [2], surgeons are provided with the pre-operative 3D images, navigated US imaging and a set of optical trackers for the navigation of standard surgical tools. During the surgical intervention, the accuracy of the pre-operative images is reduced due to motions and deformations of the liver. Thus, updating the pre-operative images with the 3D information obtained from US imaging is an important step towards an increase of the navigation system accuracy.

Registration of CT images with US imaging is usually performed by detecting corresponding features in both data types. A non-rigid transformation can then be computed and applied onto the CT images like in [3]. The most obvious common features between CT- and US- images are liver vessels. They are easily segmented pre-operatively in the CT images, whereas the segmentation of vessels in US images is more challenging as these images have a low signal-noise ratio. Furthermore, it is strongly desirable that the segmentation happens in real-time in order to provide surgeons with direct feedback on the quality of the acquired US information. Several groups have reported on US segmentation algorithms in liver surgery [4] but most of them focused on tumor detection. Recently, [5] showed an algorithm for vessel segmentation and tracking, which requires an initial manual segmentation to start. A dynamic texture approach was presented in [6]. Good segmentation results are reported but again, user intervention is required for tissue classification. In [7] a vessel segmentation algorithm was developed and tested on several porcine livers under isolated perfusion. This segmentation algorithm gave promising results both in terms of efficiency and computational time and does not require user interaction. Within this paper we present the integration of this segmentation algorithm in our navigation system and the initial results of its use in real conditions.

## **2 Methods**

### **A. Navigation System and Ultrasound Device**

The surgical navigation system is built as a transportable setup containing an NDI Vicra Camera (Northern Digital Inc, Canada), an integrated Terason T3000 ultrasound system with an 8IOA intra-operative probe (Teratech Corporation, USA) and a Shuttle barebone PC (Shuttle Inc, Taiwan) with a touch-screen monitor (ELO Touchsystems, USA). Instrument tracking is enabled by a navigation toolset composed of adapters to the existing surgical tools (CUSA, ablation devices, US probe) and a pointer calibration unit.

## B. Workflow

Before each navigated liver surgery, pre-operative tri-phase CT data was processed by MeVis Distant Services. The MeVis analysis provides segmentation of all the important structures (vessels, tumors, surface) as well as several resection proposals. The resulting 3D models are evaluated by the surgeons and appropriate visualization models are selected and loaded into the navigation system.

The navigated surgery starts by an initial rigid registration of the liver. Four landmarks are manually defined on the virtual liver model. Then, the tip of the CUSA is calibrated with the pointer calibration unit. Using the CUSA tip, the surgeon points the four manually defined landmarks on the real liver, which are automatically recorded by the navigation system. Through the alignment of these two point sets, initial rigid registration is obtained.

In a second stage, US images are acquired and segmented in real-time. The US probe is pre-calibrated using an US calibration unit pre-operatively [8]. Using the probe, vessels are traced by the surgeon. The obtained segmentation and images are recorded with the corresponding position tracking information for further data analysis.

## C. Vessel Segmentation

The vessel segmentation algorithm is based on the approach proposed by Dagon in [7] may be shortly described as:

1. Vessel image mask generation
2. Vessel seed point detection
3. Vessel approximation using an elliptical model
4. Outliers removal

The vessel mask generation starts with a scale-space blob detection [9], which is performed by generating scale-space images  $I_l$  with eight different scale levels  $S_l$  using a Gaussian Kernel of:

$$\sigma = 3 \cdot S_l \quad (1)$$

The pixel intensities of the scale-space images are then summed in a scale-space map  $I_{Map}$  with the following equation:

$$I_{Map}(x, y) = \sum_{l=1}^n [I_n(x, y) - I_l(x, y)] \quad (2)$$

Thus, dark features in the US image that present a large scale-space lifetime are enhanced. A binary threshold  $t_{sc}$  is then applied on the scale-space map to obtain the features binary mask  $I_{Mask}$ .

The features mask is then cross-correlated with a binary circular mask of radius  $r_c$  and normalized. A threshold  $t_{corr}$  is applied on the obtained image and a connected component labeling algorithm is run in order to identify the highest intensity pixel in each region, which is then labeled as seed point.

To approximate the blood vessel, an elliptical model is applied. To detect the vessel border, rays are regularly casted around the seed point on the initial scale-space mask  $I_{Mask}$ . The edge points are detected at the first value change in the ray intensity profile. The edge points are then used to initialize a direct least square ellipse fitting algorithm [10], [11].

Finally, outliers are filtered by defining minimal and maximal ellipse axes ratio and axes average length.

The C++ implementation of the segmentation algorithm was based on the code provided by Dagon [7] which was improved by adding outlier removal in the vessel seed point detection step and by integrating a region of interest (ROI) provided by the direct interface to an US imaging device (described in Section 2.D)

## D. Software integration

The segmentation algorithm was integrated into our navigation software [1],[2]. In this way, the segmentation is applied in real-time onto the images generated by the US probe. The obtained segmentation is visualized in both the 2D US image and in the 3D world viewer. A touch-screen user interface was developed for easy use in the operation room environment. The user can switch the segmentation algorithm on and off, change segmentation parameters and record both

US images and the obtained segmentations. The segmentation steps can be visualized to better understand the parameters respective influences (Fig. 1). The whole sequence can be loaded and segmented again later with different parameters.

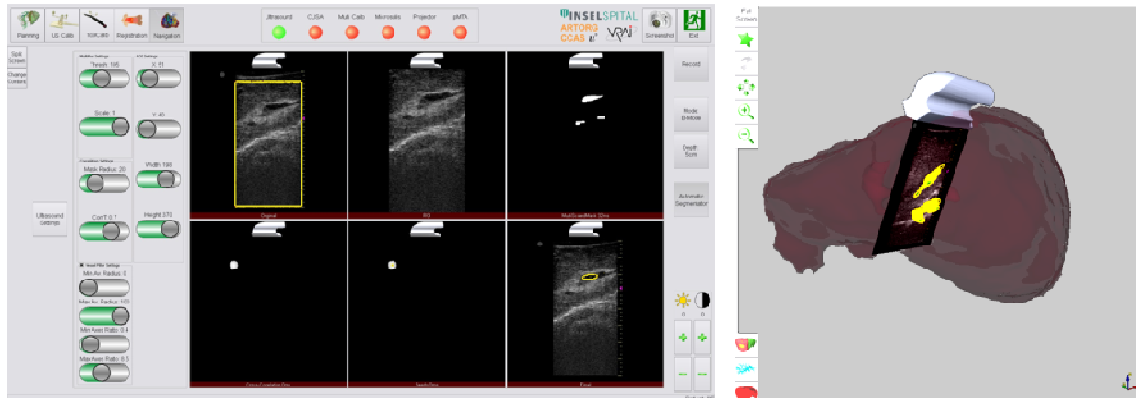


Fig. 1: Navigation software: On the left side, the user interface in detailed mode displays each step of the segmentation algorithm (original image, ROI, scale-space, correlation, seed points, ellipse fitting). On the right side, the 3D view shows the pre-operative image, tracked US probe and the segmentation (yellow).

## E. Validation

A validation application was developed for the result analysis. The recorded US images are loaded again with the obtained automatic vessel segmentation centers. The user then manually defines whether the detected vessel centers are correct, wrong (false positive) or missing (Fig. 2, right). The application then outputs the total number of each vessel center classification.

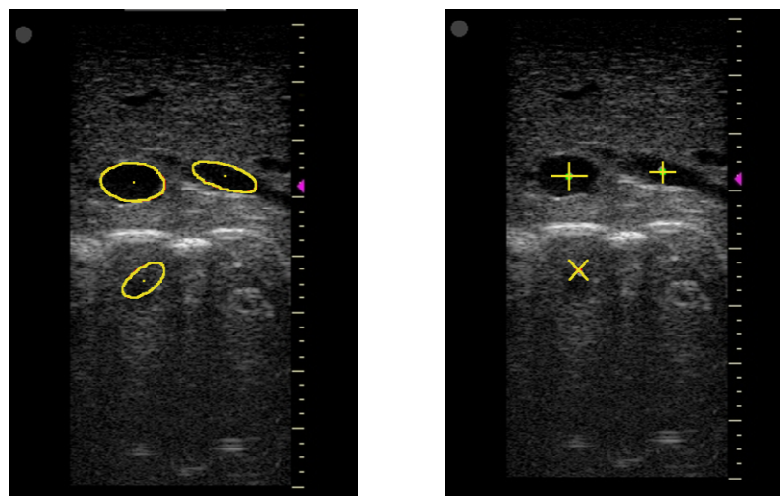


Fig. 2: Typical vessel detection produced by the vessel segmentation algorithm (left). Two segmented vessels are correct (+) and one is wrong (x) (right).

## 3 Results

About 2'200 images (300x430 pixels) were recorded and segmented during one surgical intervention at the Inselspital in Bern, Switzerland. The US probe parameters for brightness, contrast and time gain control were set to the default values. The average time to process one image was 70ms, which resulted in a frame rate of 7 frames per second. Six subsets including a total of 538 images were selected for post-operative analysis with optimized segmentation parameters. Table 1 and 2 present the results obtained with the parameters:  $t_{sc} = 65$ ,  $r_c = 10$  pixels,  $t_{corr} = 0.01$ , Region of Interest: 193x371 pixels.



	Processed images	Expected vessel centers	Correct centers	Missing centers
Set 0	109	551	250	301
Set 1	42	107	55	52
Set 3	147	463	223	240
Set 6	73	265	73	192
Set 10	13	26	22	4
Set 12	154	408	278	130

Table 1: Amount of correct- and missing- centers compared to the number of expected vessels.

	Processed images	Vessels detected	Correct classification	Wrong classification
Set 0	109	374	250	124
Set 1	42	158	55	103
Set 3	147	421	223	198
Set 6	73	228	73	155
Set 10	13	25	22	3
Set 12	154	295	278	17

Table 2: Amount of correct- and wrong- centers compared to the number of detected vessels.

The average value of correctly classified vessels is 60.03%. The average number of missing vessels per frame is 1.7 and the average value of wrongly classified vessels per frame is 1.1.

## 4 Discussion

The number of missing vessels per frame is relatively high and is mainly caused by small vessels. When the threshold value  $t_{sc}$  is set to a low value, small vessels are better detected but the amount of wrongly classified vessels from image noise increases as well. Wrongly classified vessels come from homogeneous noise regions in the images and from artefacts lying under the liver borders as illustrated in Fig. 2. When comparing our results with the ones presented by Dagon [7], we see improved outcomes in the amount of correctly segmented vessels (40% by Dagon, 60% in this work). However, the results are not directly comparable as different image data and slightly different validation methodologies were used. The images used herein were acquired in real surgery on human patients where Dagon used an ex-vivo porcine liver specimen under isolated perfusion. The validation approaches differ in that the distance between the manual and automatic segmented vessel centers was not considered for the vessel classification within this work. The higher computation time of 70ms vs 25ms in this work can be explained by the larger size of the US images used with the Terason imaging device and the added load due to the visualization of the segmentation steps as shown in Fig. 1.

As we aim to use the segmentation results for intra-operative registration, a further reduction of wrongly classified vessels is desirable. This should be partly solved by detecting the lower organ borders. Another way for reducing wrongly classified vessels is to use the 3D position information of the segmented vessels in the prior images to predict their positions in the next frame. In general, the large number of segmentation parameters made it difficult to obtain a good segmentation while in the operating room. Reducing their number or an automatic adjustment of some parameters might solve this issue.

Another effect observed when comparing the segmentation results from different acquisitions is the high sensitivity to the US image quality. We believe that by adjusting the US imaging parameters, in particular brightness, contrast and time gain control, the number of missing vessels per frame should decrease. The development of the vessel-based registration methods will show whether a reliable detection of large vessels or detection of smaller vessels with higher error rate are preferable.

Finally, the computation time of 70ms per frame is not yet acceptable for a real-time application. Considering that there should be more processing for the registration afterwards. This should be improved by code optimization and/or using GPU computing or multi-threading.



## 5 References

- [1] M. Peterhans, B. Dagon, L. Nolte, C. Baur, A. Vom Berg, D. Inderbitzin, and S. Weber, "Soft tissue navigation: Transferring operative planning data into the operation room", ECR 2009 Book of Abstracts, 2009.
- [2] A. Vom Berg, D. Candinas, D. Inderbitzin, M. Peterhans, S. Weber, and L. Nolte, "Computer assisted surgery and navigation in complex hepatic surgery and tumorablation: first clinical results of 10 patients", Proceedings of CARS 2010, 2010.
- [3] T. Lange, N. Papenberg, S. Heldmann, J. Modersitzki, B. Fischer, H. Lamecker, and P.M. Schlag, "3D ultrasound-CT registration of the liver using combined landmark-intensity information," IJMRCAS, 2009.
- [4] J. A. Noble, and D. Boukerroui, "Ultrasound Image Segmentation: a Survey", TMI, 2006.
- [5] J. Guerrero, S. E. Salcudean, J. A. McEwen, B. A. Masri, and S. Nicolaou, "Real-Time Vessel Segmentation and Tracking for Ultrasound Imaging Applications", TMI, 2007.
- [6] S. Milko, E. Samset, and T. Kadir, "Segmentation of the liver in ultrasound: a dynamic texture approach", IJMRCAS, 2008.
- [7] B. Dagon, C. Baur, and V. Bettschart, "Real-Time Update of 3D Deformable Models for Computed Aided Liver Surgery", Proceedings of IAPR, 2008.
- [8] M. Peterhans, S. Anderegg, P. Gaillard, T. Oliveira-Santos, and S. Weber, "A Fully Automatic Calibration Framework for Navigated Ultrasound Imaging", Accepted for Proceedings of EMBC 2010, 2010.
- [9] T. Lindeberg, "Detecting Salient Blob-Like Image Structures and Their Scales with a Scale-Space Primal Sketch: A Method for Focus-of-Attention", International Journal of Computer Vision, 1993.
- [10] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct Least Square Fitting of Ellipses", 13th International Conference on Pattern Recognition, 1996.
- [11] R. Halir and J. Flusser, "Numerically Stable Direct Least Squares Fitting of Ellipses", 1998.



# Elastische Registrierung der Leber mit laparoskopischem Ultraschall in drei Schritten

V. Martens<sup>1</sup>, A. Besirevic<sup>2</sup>, O. Shahin<sup>1,3</sup>, M. Kleemann<sup>2</sup>, A. Schlaefer<sup>1</sup>

<sup>1</sup> Universität zu Lübeck, Institut für Robotik u. Kognitive Systeme, Lübeck, Germany

<sup>2</sup> Universitätsklinikum Schleswig-Holstein, Klinik für Chirurgie, Lübeck, Germany

<sup>3</sup> Graduate School for Computing in Medicine and Life Sciences, Lübeck, Germany

Kontakt: {martens, shahin, schlaefer}@rob.uni-luebeck.de

## Abstract:

*Für die Navigation in laparoskopischer Leberchirurgie ist die Registrierung präoperativer Daten das zentrale Problem. Es wird das Konzept und die Umsetzung vorgestellt. Dabei kommt ausschließlich eine elektromagnetisch navigierte laparoskopische Ultraschallsonde zum Einsatz. Die Registrierung besteht aus drei Schritten: Einer Vorregistrierung basierend auf grob geschätzten Punkten an der Leberoberfläche, einer rigiden Oberflächenregistrierung und einer elastischen Registrierung. Letztere verwendet aus US-Bildern extrahierte Gefäße sowie die bereits erfasste Oberfläche. Es wird an Simulationen, am Phantom und einem echten Patientendatensatz gezeigt, dass die elastische Registrierung eine messbare und sichtbare Verbesserung bringt. Die Einbeziehung der Oberflächendaten sorgt dabei für vernünftige Ergebnisse, auch wenn nur wenige brauchbare Daten aus den Ultraschallbildern gewonnen werden können.*

*Schlüsselworte: Ultraschall, Laparoskopie, elastische Registrierung, TPS-ICP*

## 1 Problem

Die Resektion ist nach wie vor der erfolgversprechendste Heilungsversuch bei primären oder sekundären hepatischen Malignomen. Laparoskopie verspricht dabei, abgesehen vom kosmetischen Vorteil, weniger Trauma und bessere immunologische Parameter für den Patienten nach dem Eingriff. Allerdings geht dieses Verfahren einher mit dem Verlust der Tiefenwahrnehmung und der Taktilität, zumindest bei der in Europa verbreiteten Variante ohne Handport. Computer-assistierte Navigation zur Übertragung von Resektionsplanungen und Ansteuerung von Läsionen zur Ablation ist also in der Laparoskopie besonders hilfreich. Hierfür wurde von uns ein System entwickelt, welches ein Behandlungsinstrument und die bewegliche Spitze der laparoskopischen Ultraschallsonde elektromagnetisch erfasst [1].

In unserem Szenario wird eine CT- oder MRT-Aufnahme zur Diagnose und Operationsplanung erstellt. Erst danach entscheidet sich, welche Eingriffsart (interventionell, offen, laparoskopisch) am sinnvollsten ist, falls überhaupt. Diese präoperativen Daten unterscheiden sich aufgrund etlicher Faktoren wie Lagerung und Gasdruck von der intraoperativen Situation. Neben einer rigiden Registrierung zur Erfassung der Lage der Leber im Koordinatensystem des Navigationssystems ist also auch eine elastische Registrierung erforderlich.

Lange et al. [2] erfasst bei offenem Eingriff mit einer 3D-US-Sonde ein Volumen und identifiziert darin Landmarken aus dem portalvenösen Gefäßbaum. Danach findet eine Landmarken- und Intensitätsbasierte elastische Registrierung statt. Derzeit sind keine laparoskopischen 3D-Sonden auf dem Markt und bei Freihand-3D-US, also einer navigierten 2D-Sonde wie in unserem Fall, ist es nicht leicht, ein geschlossenes Volumen zu erzeugen. Daher wurde der Ansatz auf einzelne Schichtbilder im Raum erweitert [3]. Informationen über die Oberfläche werden jedoch nicht erfasst.

Die Gruppe um Chapman und Galloway aus Nashville präsentiert, ebenfalls für offene Eingriffe, verschiedene Ansätze für eine elastische Registrierung basierend auf Oberflächendaten, u.a. ein Atlas-basiertes Verfahren über mit FEM vorberechneten Deformationen [4]. Hierfür wird optisches Tracking und ein Laser-Range-Scanner verwendet. Arbeiten an einer minimal-invasiven Umsetzung laufen indes bereits [5].

Im Folgenden präsentieren wir ein dreistufiges Registrierungsverfahren, welches ausschließlich die navigierte laparoskopische Ultraschallsonde erfordert. Die Interaktion mit dem System ist auf wenige einfache Aktionen beschränkt (Lebermodell drehen, Knöpfe drücken, Schieberegler einstellen). Da die Leber nicht direkt getrackt wird, findet keine automatische Anpassung der Registrierung bei Bewegungen oder im Laufe der Resektion selbst statt. Allerdings konnte bei mehreren laparoskopischen Eingriffen an der Leber keine nennenswerte Atmungsbewegung festgestellt werden.

## 2 Methoden

Der ohnehin zur Diagnose aufgenommene CT- oder MRT-Patientendatensatz wird zuerst segmentiert (MeVis, Bremen). Diese Daten werden zu Beginn der Anwendung in das System eingelesen und 3D-Modelle der Tumore, Gefäße, Lebersegmente und Resektionsvorschläge erstellt. Nach einigen Schritten zur Kalibrierung der Instrumente beginnt die Registrierung der Leber zu den präoperativen Falldaten. Diese Registrierung gliedert sich in die folgenden 3 Schritte.

Der erste Schritt dient der groben Einmessung der Leber. Dazu muss das Modell auf dem Bildschirm so ausgerichtet werden, dass die sichtbare Seite auch in Wirklichkeit mit Instrumenten erreichbar ist (Interaktion I). Nach dem Start erscheinen automatisch 4 Landmarken an den Seiten der Leber, die nacheinander vom Chirurgen mit der navigierten Ultraschallsonde angesteuert und über eine Schaltfläche bestätigt werden müssen (Interaktion II). Dieser Vorgang ist zwar relativ grob, aber kann recht schnell durchgeführt werden und ist zudem sehr robust.

Als nächstes wird die Oberfläche erfasst, um sie mit dem ICP-Algorithmus (Iterative Closest Point) rigide auf das präoperative 3D-Modell zu registrieren mit der Grobregistrierung aus dem ersten Schritt als Startwert. Dies geschieht ebenfalls mit der navigierten Ultraschallsonde (Interaktion III), da deren flexible Spitze auch an die Seiten und hinter die Leber gelangt. Allerdings wird dabei die Oberfläche eingedrückt, weswegen die erfasste Punktposition leicht modifiziert wird. Wichtig ist, dass die aufgenommenen Punkte möglichst weit über die Leber verteilt sind, damit der Algorithmus nicht in ein lokales Optimum läuft. Daher werden die Anzahl und räumliche Ausdehnung der Oberflächenpunkte kontrolliert. Zudem wird anhand des Ultraschallbildes überprüft, ob überhaupt Kontakt zum Gewebe besteht. Bei Luft bilden sich charakteristische weiße Artefakte am oberen Bildrand, und es werden keine Oberflächenpunkte aufgenommen. Daher kann der Chirurg während der Aufnahme problemlos die Sonde umsetzen.

Als dritter und letzter Schritt folgt eine elastische Registrierung. Hierfür muss eine Sequenz von Ultraschallbildern aufgenommen werden (Interaktion IV). Dabei kann es sich sowohl um Farbdopplerbilder wie auch Grauwertbilder in allen verfügbaren Eindringtiefen handeln. Auch hier ist eine großräumige Aufnahme von Vorteil. Die elastische Registrierung besteht aus folgenden Schritten:

- **Segmentierung:** Alles Farbige im Farbdopplerbild. Bei Grauwertbildern adaptives Thresholding auf herunterskaliertem und geglättetem Bild, wobei über einen Schieberegler (Interaktion V) der Threshold feinjustiert werden kann. Dank einer Echtzeit-Segmentierungsvorschau kann die Segmentierung direkt auf der Leber so eingestellt werden, dass nur die sichtbaren Gefäße möglichst gut erkannt werden.
- **Gefäßvalidierung:** Alle segmentierten Bereiche werden überprüft, ob es sich um Gefäße handelt und nicht z.B. um Luft oder Schatten. Neben der bereits erwähnten Prüfung auf Luftkontakt werden zu kleine und zu große Segmente entfernt. Außerdem werden alle Bereiche entfernt, die den Bildrand berühren und ihren Schwerpunkt sehr nah am Bildrand haben, da vermutlich ein wesentlicher Teil ihrer Gestalt fehlt. Schließlich werden nur die Segmente als Gefäße akzeptiert, die größtenteils von helleren Regionen umgeben sind.
- **Merkmalsbestimmung:** Für eine schnelle Verarbeitung werden die Gefäße bzw. die Gefäßschnitte auf ihre wesentlichen Eigenschaften reduziert: Mittelpunkt und Radius. Bei ovalen Segmenten kann der Schwerpunkt genommen werden. Bei Gefäßaufzweigungen könnte dieser jedoch sogar außerhalb des Segments liegen. Daher wird eine Abstandskarte der Pixel eines Segments zusammen mit dem Skelett, wie in [6] beschrieben, erstellt. Liegt der Schwerpunkt nahe am Abstandsmaximum des Segments, wird er ausgewählt. Ansonsten werden die lokalen Maxima (als Teile des bereits extrahierten Skeletts) in der Abstandskarte verwendet. Liegen mehrere dicht beisammen, wird wiederum deren Schwerpunkt genommen. Dadurch können auch Gefäßaufzweigungen häufig gut erfasst werden (s. Abbildung. 1). Der zugehörige Gefäßradius ergibt sich direkt aus der Abstandskarte.

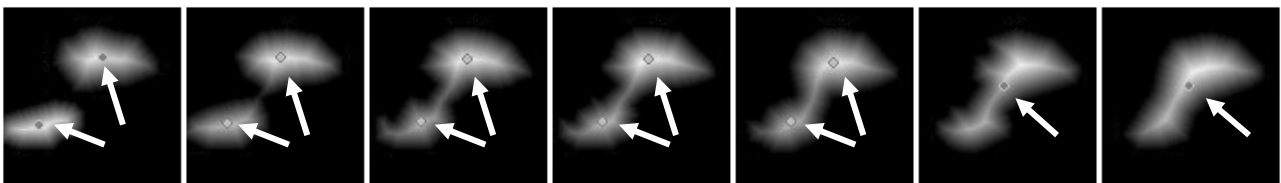


Abb. 1: Beispiel für Bestimmung der Gefäßmittelpunkte. Segmentierte Bereiche zur Anschaulichkeit mit Intensität proportional zum Abstand zum Segmentrand (Abstandskarte). Dunkle Punkte sind Schwerpunkte der Segmente, helle Punkte mit dunklem Rand die (aggregierten) Maxima. Im 2.-5. Bild von links wäre der Schwerpunkt keine gute Wahl.

- **Registrierung:** In einem iterativen Verfahren (TPS-ICP) wird eine elastische Verformung mittels TPS (Thin Plate Splines) durchgeführt. Dafür wird jedem im Ultraschall gefundenen Gefäßmittelpunkt ICP-ähnlich der nächstliegende Punkt im präoperativen Gefäßmodell (hepatische und portalvenöse Gefäße, sofern segmentiert auch Arte-

rien) zugewiesen, wobei auch der Gefäßradius in die Abstandsberechnung mit eingeht. Das gleiche geschieht parallel mit den im zweiten Schritt gefundenen Oberflächenpunkten und dem präoperativen Oberflächenmodell. In einem dem diskreten Abkühlen (discrete annealing) ähnlichen Verfahren wird der Regularisierungsparameter  $\alpha$  kontinuierlich verringert (elastischer). Wenn in einer inneren Schleife die Veränderung der Punktpositionen unter eine bestimmte Grenze fällt, wird diese verlassen und in einer äußeren die Temperatur und mit ihr die Elastizität verändert, bis eine minimale Temperatur erreicht ist. Dabei werden die wenigen erfassten Punkte verformt und erst zum Schluss mit einer inversen Transformation die präoperativen Daten auf die intraoperativen abgebildet.

Sowohl beim zweiten wie auch dritten Registrierungsschritt sollte die Beatmung für die Dauer der jeweiligen Aufnahme (Oberfläche bzw. US-Sequenz) angehalten werden, wenn die Bewegung der Leber zu stark ist (mehr als 5mm).

### 3 Ergebnisse

Es wurden drei Arten von Evaluationen durchgeführt: synthetisch (a), am Phantom (b) und an Patientendaten (c).

a) Der Schritt der elastischen Registrierung ohne Bildverarbeitung wurde an verschiedenen präoperativen Leberdatensätzen evaluiert. Dafür wurden zufällig Gefäß- und Oberflächenpunkte in einem Radius von 90mm um einen zufälligen Punkt nahe des Leberschwerpunkts ausgewählt. Zuvor wurden die Patientendaten zufällig rigide und elastisch (mit TPS) verformt. Tabelle 1 zeigt die mittlere quadratische und maximale Abweichung der registrierten Gefäßmittelpunkte und Oberflächenpunkte von der deformierten Grundwahrheit. Vor TPS-ICP (Gefäßmittel- und Oberflächenpunkte) wurde ICP auf die Gefäßmittelpunkte angewendet, um eine bessere Ausgangslage für TPS-ICP zu erhalten.

Verfahren	Mittl. Quadr. (Gefäße)			Max.			Mittl. Quadr. (Oberflächenpunkte)			Max.		
Initial	10,3	9,5	9,0	18,2	16,2	17,2	11,0	11,5	8,3	20,8	21,2	15,5
ICP (Gefäße)	6,1	6,7	6,7	12,4	13,8	12,7	7,3	9,1	6,8	15,2	18,6	14,6
TPS-ICP	3,8	3,8	4,6	11,1	14,8	14,9	6,7	8,1	5,5	16,6	25,8	15,8

Tabelle 1: Fehler in mm bei drei Testläufen (jeweils linke, mittlere und rechte Spalte) mit synthetischen Daten eines menschlichen Patientendatensatzes

b) An einem Leberphantom (Kagaku, Kyoto, Japan) wurde die ganze Registrierung durchgeführt, und zwar einmal ohne Verformung und mit zwei unterschiedlichen Verformungen. Für jeden der drei Testfälle wurden eine Grob- und Oberflächenregistrierung durchgeführt und 4-6 US-Sequenzen zur elastischen Registrierung aufgenommen. Die Grundwahrheit bildeten an der Oberfläche mit einem Zeiger und innen mit der Ultraschallsonde definierte Landmarken. Die Landmarken wurden zuvor in einem störfreien Umfeld und nach präziser Oberflächenregistrierung im unverformten Zustand definiert. Die tatsächlichen Aufnahmen fanden unter OP-ähnlichen Bedingungen statt (Tabelle 2).

Verfahren	Mittl. Quadr. (Gefäße)			Max.			Mittl. Quadr. (Oberflächenpunkte)			Max.		
Initial	7,4	11,6	10,0	11,2	19,2	20,1	1,9	2,3	2,0	6,6	7,2	10,8
ICP (Gefäße)	7,7	7,3	6,9	11,8	11,9	11,6	2,9	4,7	2,0	5,9	16,9	10,8
TPS-ICP	7,7	6,7	7,0	13,6	12,0	12,4	1,4	1,1	2,7	6,0	5,2	7,5

Tabelle 2: Fehler in mm am Phantom ohne Deformation (jeweils linke Spalte), gestaucht (jeweils mittlere Spalte) und gestreckt (jeweils rechte Spalte)

In der Tabelle findet sich jeweils nur die beste Registrierung pro Deformation. Bei der ersten Deformation (mittlere Spalte) waren zwei Registrierungen besser, eine brachte keine Veränderung, eine wurde schlechter. Bei der zweiten Deformation waren drei besser und zwei schlechter. Ohne Deformation blieben drei ungefähr gleich, drei wurden etwas schlechter (eine Verbesserung war hier allerdings auch kaum zu erwarten). Leider war die Segmentierung der präoperativen Daten nicht sehr gut (z.B. fehlt ein Gefäß), so dass Aufnahmen in diesen Bereichen zu Fehler führen konnten.

c) Schließlich standen navigierte Ultraschallsequenzen samt initialer Registrierung eines echten Patienten zur Verfügung, allerdings ohne Oberflächenpunkte. An diesen Daten wurde die elastische Registrierung erprobt (Tabelle 2). Als Grundwahrheit wurden postoperativ 10 Landmarken in den US-Sequenzen gesetzt.

Verfahren	Mittl. Quadr. Fehler	Max. Fehler
Intraop. Registrierung	38,7	44,7
ICP der Gefäßzentren mit Radius	6,5	13,5
TPS-ICP (nach ICP)	5,0	13,5

Tabelle 2: Ergebnisse der Registrierung am echten Patienten (alle Angaben in mm)

Der Unterschied zwischen dem Fehler nach ICP und nach TPS-ICP mag nicht groß sein, der visuelle Unterschied ist jedoch nicht unerheblich. Abbildung 2 zeigt einige ausgewählte (und in der Sequenz gleichmäßig verteilte) Ultraschallbilder nach ICP und nach darauf folgendem TPS-ICP. Für TPS-ICP zeigen die überblendeten Gefäße eine bessere Übereinstimmung mit den US-Bildern. Aufgrund der starken Gewebekompression wurden die ermittelten Gefäßdurchmesser mit dem Faktor 1.3 multipliziert. Dies ist beim nachfolgend beschriebenen Leberphantom nicht erforderlich. Die US-Segmentierung dauerte 64ms pro Bild (PC mit Intel Core 2 Duo, 2.4 GHz).

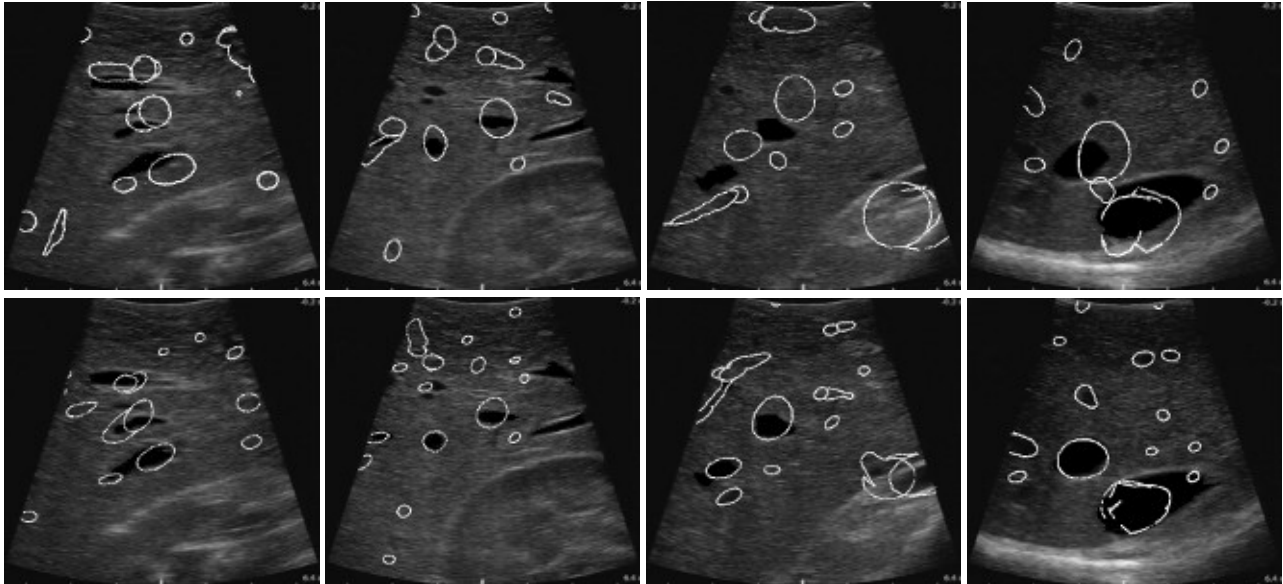


Abb. 2: Von links nach rechts vier Bilder aus der US-Sequenz vom echten Patienten. Die Eckpunkte der präoperativen Oberflächenmodelle sind als helle Punkte eingezeichnet. Oben: ICP über Gefäße. Unten: zusätzlich TPS-ICP.

## 4 Diskussion

Die Ergebnisse zeigen, dass die elastische Registrierung eine Verbesserung bringen kann. Stehen US-Daten für den gesamten Gefäßbaum zur Verfügung (Simulation), ist die Verbesserung erheblich und das Verfahren robust. Stehen nur wenige Daten zur Verfügung (wenige Gefäße im Sequenzbereich, schlechte Bildqualität), kann das TPS-ICP Verfahren auch zu schlechteren Lösungen führen. Hier helfen aber weiträumig verteilte Oberflächenpunkte. Der gezeigte Ansatz ist aufgrund der geringen Interaktion und der Verwendung weniger Merkmale bei den Ultraschallsequenzen (Gefäßmittelpunkt und Radius) schnell, es gibt fast keine Wartezeiten. Allerdings ist das Verfahren stark vom Startwert abhängig, und bei starker Deformation wäre eine elastische Vorregistrierung, z.B. über Landmarken, wohl hilfreich.

Das Projekt wird finanziert vom Bundesministerium für Bildung und Forschung (Förderkennzeichen 01IBE03A).

## 5 Referenzen

- [1] V. Martens, S. Schlichting, A. Besirevic and M. Kleemann, "LapAssistent - a laparoscopic liver surgery assistance system", Proceedings of the EMBEC, 121-125, Antwerpen, Belgien, 2008
- [2] T. Lange, N. Papenberg, S. Heldmann, J. Modersitzki, B. Fischer, H. Lamecker, P. Schlag, „3D ultrasound-CT registration of the liver using combined landmark-intensity information, Int. J. CARS, 4(1): 79-88, 2009
- [3] S. Heldmann und N. Papenberg, „A Variational Approach for Volume-to-Slice Registration”, Proceedings of the SSVM, 624–635, Voss, Norwegen, 2009
- [4] L. W. Clements, P. Dumpuri, W. C. Chapman, R. L. Galloway und M. I. Miga, "Atlas-based method for model updating in image-guided liver surgery", In K. R. Cleary und M. I. Miga, Hg., Medical Imaging 2007: Visualization and Image-Guided Procedures, Bd. 6509, SPIE 2007
- [5] T. P. Rauth, P. Q. Bao, R. L. Galloway, J. Bieszczad, E. M. Friets, D. A. Knaus, D. B. Kynor und A. J. Herline, "Laparoscopic surface scanning and subsurface targeting: implications for image-guided laparoscopic liver surgery", Surgery, 142(2): 207–214, 2007
- [6] S. Chang, "Extracting Skeletons from Distance Maps", IJCSNS International Journal of Computer Science and Network Security, 2007

# Lokalisierung der Leber mittels einer Diskriminativen Generalisierten Hough Transformation

Heike Ruppertshofen<sup>1,2</sup>, Cristian Lorenz<sup>3</sup>, Sarah Schmidt<sup>4,2</sup>, Peter Beyerlein<sup>4</sup>, Zein Salah<sup>2</sup>, Georg Rose<sup>2</sup>, Hauke Schramm<sup>1</sup>

<sup>1</sup> Fachhochschule Kiel, Institut für Angewandte Informatik, Kiel, Germany

<sup>2</sup> Otto-von-Guericke Universität, Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik, Magdeburg, Germany

<sup>3</sup> Philips Research Europe – Hamburg, Department Digital Imaging, Hamburg, Germany

<sup>4</sup> Technische Hochschule Wildau, Fachbereich Ingenieurwesen, Wildau, Germany

Kontakt: heike.ruppertshofen@fh-kiel.de

## Abstract:

*In diesem Beitrag soll ein für 2D Bilder bereits erfolgreich eingesetztes Verfahren für die automatische Objektlokalisierung auf 3D Problematiken erweitert und die Einsetzbarkeit des Verfahrens für 3D Daten gezeigt werden. Das Verfahren der Objektlokalisierung basiert dabei auf der Generalisierten Hough Transformation (GHT), bei der das dort verwendete Formmodell mit individuellen Punktgewichten ausgestattet wird, die im Abstimmverfahren der GHT zum Einsatz kommen und so die Genauigkeit des Verfahrens verbessern. Diese Punktgewichte werden mittels eines diskriminativen Lernverfahrens bezüglich eines minimalen Lokalisierungsfehlers in der GHT optimiert. Die trainierten Punktgewichte erhöhen dabei nicht nur die Genauigkeit der GHT, sondern bewerten auch die Wichtigkeit eines Punktes, so dass unwichtige Punkte aus dem Modell entfernt werden können. Kleinere Modelle führen zu kürzeren Laufzeiten, was den Einsatz des Verfahrens auch für 3D Aufgabenstellungen praktikabel macht. Die Stärke des Verfahrens soll hier am Beispiel der Lokalisierung der Leber, die meist eine recht starke Variabilität aufweist, gezeigt werden.*

*Schlüsselworte: Objektlokalisierung, Generalisierte Hough Transformation, Diskriminatives Training, Maschinelles Lernen*

## 1 Problem

Die Aufgabe der Objektlokalisierung kommt bei vielen Anwendungen in der computergestützten Medizin - im Rahmen der Diagnostik, der Radiologie und der Chirurgie - zum Tragen. Viele der Verfahren, insbesondere in der Operations- und Bestrahlungsplanung, stützen sich dabei auf eine vorherige Segmentierung von Organen oder Knochen, um voll automatisiert Volumen, Längen oder Winkel zu berechnen und die Planung einer Operation oder Bestrahlung am Computer zu ermöglichen.

Die meisten dieser automatischen Segmentierungsverfahren benötigen zunächst die ungefähre Position des Zielorgans, um ein Segmentierungsmodell initial zu positionieren. In vielen Fällen wird diese Position manuell bestimmt oder es kommen speziell auf das gestellte Problem zugeschnittene Verfahren zum Einsatz [4]. Die Anpassung dieser Verfahren auf neue Problemstellungen ist zeitaufwändig und häufig nicht machbar.

In diesem Beitrag wollen wir daher ein generelles und automatisches Verfahren zur Objektlokalisierung vorstellen, das problemlos auf weitere Objekte angewendet werden kann. Das Verfahren besteht aus zwei Modulen: Zum einen der Generalisierten Hough Transformation (GHT) [2] zur Objektlokalisierung, die hier insofern erweitert wurde, dass ein gewichtetes Modell verwendet wird, und zum anderen ein diskriminatives Lernverfahren (DMC) [3], mit dessen Hilfe die individuellen Gewichte für die einzelnen Modellpunkte ermittelt werden. Durch das Training wird erreicht, dass Bereiche des Modells, die besonders charakteristisch für das Zielobjekt sind, stärker gewichtet werden und so auch einen stärkeren Einfluss in der GHT haben.

Auf 2D Bildern kommt das Verfahren bereits erfolgreich zum Einsatz [1]; In diesem Beitrag soll nun die Einsetzbarkeit des Verfahrens auf 3D Bildern gezeigt werden. Als Anwendungsbeispiel wird die Lokalisierung der Leber in 3D CT-Aufnahmen verwendet. Um die Leistung der neuen Methode im direkten Vergleich zur Standard-GHT darstellen zu können, wird die Lokalisierung mit einem mittleren Formmodell, welches die durchschnittliche Form der Leber repräsentiert, durchgeführt. Die Erstellung eines in der GHT verwendbaren Modells für das Zielobjekt direkt aus den Daten, wie es in [1] für den 2D-Fall vorgestellt wurde, ist aber ebenfalls möglich.

## 2 Methoden und Material

Wie bereits in der Einleitung erwähnt wurde, besteht das Verfahren aus zwei Modulen, auf die im Folgenden näher eingegangen werden soll.

Für die Objektlokalisierung wird die GHT [2] eingesetzt. Das Verfahren eignet sich gut für medizinische Bildverarbeitung, da es robust gegenüber Bildrauschen und Überdeckungen oder fehlenden Objektteilen ist. Nachteil des Verfahrens ist die hohe Laufzeit, die vermutlich dafür verantwortlich war, dass die GHT auf 3D Problematiken kaum Anwendung fand, und die hier durch kleinere Modelle und stark eingeschränkte Transformationsparameter reduziert werden soll.

Die GHT führt die Lokalisierung mittels eines Abstimmverfahrens durch, bei dem das Bild mit Hilfe des Modells in den sogenannten Hough-Raum transformiert wird. Zu diesem Zweck wird ein Kantenbild des Originalbildes erstellt und für jeden Kantenpunkt  $e_j$  die Modellpunkte  $m_k$  mit ähnlicher Gradientenrichtung ermittelt. Über den Zusammenhang  $c_i = e_j - m_k$  wird die korrespondierende Hough-Zelle  $c_i$  ermittelt und ihr Wert erhöht. Nach Beendigung des Abstimmverfahrens wird die Zelle mit dem höchsten Wert gesucht und als wahrscheinlichste Objektposition ausgegeben. Das Verfahren kann weiterhin verwendet werden, um skalierte und rotierte Vorkommen des Zielobjekts zu lokalisieren. Dafür muss das Modell transformiert und das Verfahren wiederholt werden, was zu einem höher dimensional Hough-Raum führt. Aus Laufzeitgründen werden diese Möglichkeiten hier jedoch nicht in Betracht gezogen. Stattdessen soll das Modell lernen, welche Bereiche für die Lokalisierung am verlässlichsten und wichtigsten sind, so dass die Bestimmung der Position des Zielobjektes im Bild auch ohne Schätzung weiterer Parameter erfolgreich ist.

Die Erweiterung der GHT von 2D auf 3D ist unkompliziert und benötigt lediglich eine Erweiterung der Datenstrukturen und eine speichersparende Verwaltung des Hough-Raumes und der für das folgende Gewichtstraining benötigten Informationen in Heap- und Hashstrukturen.

In der Standard-GHT haben alle Modellpunkte einen gleich starken Einfluss auf die Lokalisierung. Bei jedem Objekt gibt es jedoch Gebiete, die charakteristisch sind und somit in allen Bildern ähnlich aussehen, und wiederum andere die recht stark variieren. Dieses soll hier im Modell berücksichtigt werden, indem für jeden Modellpunkt ein individuelles Gewicht bestimmt wird, welches im Abstimmprozess der GHT verwendet werden soll. Dabei sollen Punkte, die das Zielobjekt robust beschreiben, höhere Gewichte bekommen und Punkte, die nur bei sehr wenigen Bildern passen, ein niedriges Gewicht erhalten, so dass sie später über ihr Gewicht identifiziert und aus dem Modell entfernt werden können. Des Weiteren können auch negative Gewichte vergeben werden, um das Modell von verwechselbaren Strukturen abzustößen und so eine falsch-positive Lokalisierung zu vermeiden.

Für die Ermittlung dieser Modellpunktgewichte wird ein diskriminatives Lernverfahren [3] herangezogen, das erstmalig in der Spracherkennung eingesetzt wurde. Bei diesem Lernverfahren wird die Dimensionalität der Bilder nicht berücksichtigt, so dass es direkt für die drei dimensional Bilder anwendbar ist. In dem Verfahren wird jeder Modellpunkt mit seinem Beitrag zum Hough-Raum als individuelle Wissensquelle aufgefasst und diese Wissensquellen über einen log-linearen Ansatz kombiniert. Die über den log-linearen Ansatz eingeführten Gewichte für die einzelnen Wissensquellen sollen hier als Modellpunktgewichte verwendet werden. Um diese bestimmen zu können, wird eine Fehlerfunktion definiert, die den Fehler in Abhängigkeit der Gewichte über alle Trainingsbilder akkumuliert und so eine nahezu optimale Gewichtung der Wissensquellen ermittelt, bei der der Fehler über alle Bilder am kleinsten ist. Für eine detaillierte Herleitung und Erklärung der Bestimmung der Modellpunktgewichte verweisen wir den Leser auf [1].

Nach Ermittlung der Gewichte kann die Modellgröße reduziert werden, indem Modellpunkte mit einem kleinen absoluten Gewicht aus dem Modell entfernt werden. Gleichzeitig wird das Modell diskriminativer für das Zielobjekt, da robuste Modellpunkte einen stärkeren Einfluss bekommen und verwechselbare Modellbereiche entfernt werden.

Die Methode wird auf 38 CT-Aufnahmen des Rumpfes getestet mit der Aufgabe, die Leber zu lokalisieren. Die verwendeten Daten stammen aus einem Leber-Segmentierungs-Challenge [5] sowie einer öffentlich verfügbaren Datenbank [6], bei der auf etwa 75% der Bilder Lebertumore sichtbar sind. Die Bilder haben unterschiedliche Qualität und Auflösung, die zwischen  $0.56 \times 0.56 \times 1$  mm und  $0.87 \times 0.87 \times 4$  mm schwankt. Ein koronaler Schnitt durch einen der Datensätze ist in Abb. 1, links zu sehen. Aus Laufzeitgründen wurden die Bilder für das Verfahren zwei Mal heruntergesampelt; wird eine höhere Genauigkeit benötigt, kann dieser Schritt jedoch übersprungen oder ein mehrstufiger Ansatz verfolgt werden.

Von den gegebenen Bildern werden 8 im Training der Modellpunktgewichte eingesetzt und 30 zu Testzwecken zurück behalten, um zu evaluieren, wie gut die Lokalisierungseigenschaften des Modells auf unbekannten Daten sind. Als Modell wird in diesem Beitrag ein mittleres Punktmodell der Leber mit 5120 Modellpunkten verwendet (s. Abb. 2, links). Als Referenzpunkt, bzw. Zielpunkt für die Leber wird ihr Schwerpunkt verwendet, der aus vorliegenden Segmentierungen der Daten berechnet wurde.



### 3 Ergebnisse

In Abb. 1 wird ein Beispielergebnis der Lokalisierung vorgestellt. Das linke Bild zeigt dabei einen koronalen Schnitt durch den Körper an der Stelle des gefundenen Punktes. Das rechte Bild stellt den dazugehörigen Schnitt durch

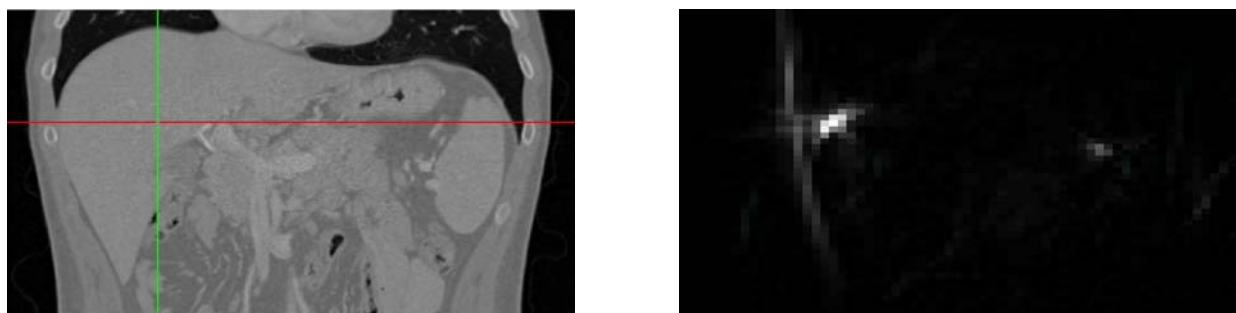


Abb. 1: Das linke Bild zeigt einen koronalen Schnitt durch eines der CT Bilder mit sichtbarer Leber. Der von der GHT gefundene Punkt wird durch die beiden Linien markiert. Das rechte Bild zeigt den dazugehörigen Hough-Raum der gleichen Schnittebene. Die Zellen mit den höchsten Werten sind deutlich zu erkennen.

den Hough-Raum dar mit einem klar sichtbaren Maximum. In diesem Beispiel wurde der Referenzpunkt in der Leber mit einer Genauigkeit von 0.7 cm lokalisiert. Abbildung 3 zeigt ein weiteres Beispiel, bei dem die Lokalisierung mit dem untrainierten und trainierten Modell verglichen wird.

Die weiteren Ergebnisse der Leberlokalisierung mit dem gewichteten und ungewichteten Modell sind in Tabelle 1 zu finden. Mit dem ungewichteten Modell wird die Leber in 18 von 38 Bildern gefunden mit einem mittlerem Fehler von 4.7 cm, bzw. 4.3 cm auf den Test-, bzw. Trainingsbildern. Durch die hohe Variabilität der Leber ist die Durchführung der Standard-GHT mit einem mittleren Modell in vielen Fällen nicht erfolgreich. Wird ein gewichtetes Modell verwendet erhöht sich die Genauigkeit der GHT auf 2.8 cm auf den unbekannten Testdaten und die Leber kann in allen Bildern lokalisiert werden. Durch eine Reduktion des Modells auf 500 Punkte werden unwichtige Punkte entfernt, wobei die Lokalisierungsgenauigkeit erhalten bleibt. Diese Ergebnisse sind sehr zufriedenstellend, insbesondere wenn man die grobe Auflösung des Hough-Raumes von durchschnittlich 0.5 cm und die vage Annotation der Zielpunkte, die sich nur am Schwerpunkt der Leber und nicht an klar definierten anatomischen Strukturen orientieren, in Betracht zieht.

Die gewichteten Modelle sind in Abb. 2 zu sehen. Dabei zeigt die linke Abbildung das komplette Modell, während die rechte Abbildung nur die 500 Punkte mit dem höchsten absoluten Gewicht darstellt. Aus dem rechten Bild wird ersichtlich, dass die obere Kante der Leber am robustesten für die Erkennung ist, da hier am wenigstens Variabilität auftritt. Hinzukommt, dass die Kante zwischen Leber und Lunge in allen Bildern gut zu erkennen ist. Die vordere Seite der Leber wurde mit negativen Gewichten versehen. Hierbei handelt es sich also um Punkte, die an einer anderen Stelle des Bildes besser passen und zu falschen Lokalisierungsergebnissen führen würden. Die laterale Seite der Leber erhielt nur niedrige Gewichte, da hier die Gefahr der Verwechslung mit der Körperaußenkontour besteht, wie in Abb. 3 für den Fall des untrainierten Modells gut zu erkennen ist.

Durch die Reduktion der Modellgröße verringert sich des Weiteren auch die Laufzeit des Verfahrens um 50%, wobei ca. 20s für Vorverarbeitungs- und Nachverarbeitungsschritte, wie z.B. das Ein- und Auslesen der Bilder, das Heruntersampeln und die Berechnung der Kanten, fallen. An dieser Stelle sollte dazu gesagt werden, dass das Verfahren bisher nicht laufzeitoptimiert ist. Insbesondere durch die gute Parallelisierbarkeit der verwendeten Methoden besteht hier noch ein hohes Potential für eine deutliche Reduktion der Laufzeit.

### 4 Diskussion

Der vorgestellte generelle Ansatz zur Objektlokalisierung zeigt auch auf 3D Daten zufriedenstellende Ergebnisse. Durch die individuelle Gewichtung der Modellpunkte konnte eine deutliche Erhöhung der Lokalisierungsgenauigkeit und eine

	Untrainiertes Modell	Trainiertes Modell	
Modellgröße	5120	5120	500
Fehler Trainingsbilder [cm]	$4.3 \pm 3.3$ (10.1)	$2.0 \pm 1.2$ (4.8)	$2.1 \pm 1.2$ (4.8)
Fehler Testbilder [cm]	$4.7 \pm 2.8$ (10.9)	$2.8 \pm 1.2$ (6.1)	$2.7 \pm 1.2$ (6.1)
Lokalisierungsrate [%]	47.4	100	100
Laufzeit [s]	49	49	24

Tabelle 1: Vergleich des ungewichteten und gewichteten Modells. Angegeben sind die Anzahl der Modellpunkte, der Lokalisierungsfehler der erfolgreichen Lokalisierungen in cm (Format: Mittelwert  $\pm$  Standardabweichung (Maximum)), die Anzahl der erfolgreichen Lokalisierungen in % und die durchschnittliche Laufzeit pro Bild in Sekunden.

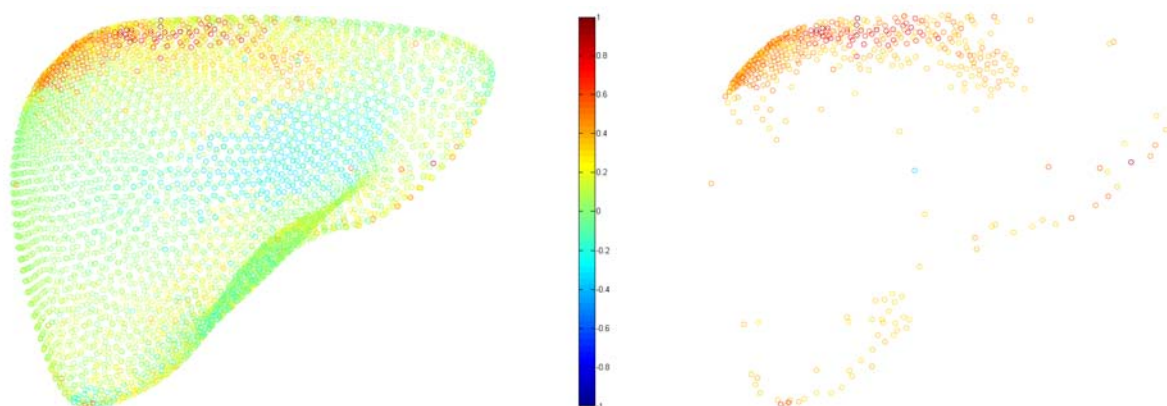


Abb. 2: Die Abbildung zeigt das mittlere Modell der Leber (links), wobei die Punkte entsprechend ihres Gewichts farbkodiert wurden (rot = 1, blau = -1). Das rechte Bild zeigt die 500 Punkte des Modells mit dem höchsten absoluten Gewicht.

Reduktion der falsch-positiven Lokalisierungen erreicht werden. Die Schätzung von Rotations- und Skalierungsparametern für die Lokalisierung des Zielobjektes wird dabei durch die Gewichtung des Modells redundant.

Ein großer Vorteil des Verfahrens besteht darin, dass es automatisch abläuft und ohne Anpassungen auf neue Aufgabenstellungen angewendet werden kann. Zwar wurde das Verfahren hier zu Vergleichszwecken unter Verwendung eines zuvor generierten, mittleren Modells vorgestellt, praktisch kann das Verfahren aber auch mit einer zufälligen Punktwolke arbeiten, da es selbständig die für das Zielobjekt relevanten Punkte identifiziert. Somit benötigt das Verfahren als Eingabe lediglich die Annotation des Zielpunktes in den Trainingsbildern, um ein diskriminatives Modell für die Lokalisierung zu erstellen.

Ein weiterer Vorteil der Methode liegt in der möglichen Reduktion der Modellgröße. Unwichtige Punkte, die keinen starken Beitrag zur Lokalisierung des Zielobjektes leisten, können anhand ihres niedrigen Gewichts identifiziert und aus dem Modell entfernt werden. Dieses führt zu einem klareren Ergebnis im Hough-Raum, da weniger störende Stimmen abgegeben werden, und einer deutlichen Reduktion der Laufzeit.

Das Verfahren wurde hier auf heruntergesampelten Bildern und einem relativ grob quantisierten Hough-Raum vorgestellt. Dieser hatte eine nahezu isotrope Auflösung von 0.5 cm im Mittel. Vor diesem Hintergrund ist die erreichte Lokalisierungsgenauigkeit von 2 cm im Mittel, was einem Fehler von 4 Voxeln entspricht, sehr zufriedenstellend. Sollte eine höhere Genauigkeit benötigt werden, kann jedoch auch mit voll aufgelösten Bildern gearbeitet oder ein mehrstufiger Ansatz verfolgt werden.

In diesem Beitrag sollte zunächst die Eignung des Verfahrens für 3D Bilder gezeigt werden. Um die Modelle, die in der GHT verwendet werden, weiter zu verbessern, soll als nächster Schritt die iterative Modellbildung wie in [1] zum Einsatz kommen, um die Stärken des Verfahrens voll nutzen zu können. Hierbei wird das Modell direkt aus den Daten generiert, um die vorliegende Variabilität in Rotation und Skalierung im Modell einfangen zu können. Gleichzeitig wird bei diesem Ansatz auch Information über verwechselbare Strukturen in das Modell gelernt, um die Gefahr falsch-positiver Lokalisierungen weiter zu minimieren.

## 5 Referenzen

- [1] H. Ruppertshofen, C. Lorenz, S. Strunk, P. Beyerlein, Z. Salah, G. Rose, H. Schramm, Fully Automatic Model Creation for Object Localization utilizing the Generalized Hough Transform, Proceedings of Bildverarbeitung für die Medizin, Springer, 2010
- [2] D. H. Ballard, Generalizing the Hough Transform to Detect Arbitrary Shapes, Pattern Recognition 13(2), 1981

- [3] P. Beyerlein, Discriminative Model Combination, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 1998
- [4] T. Heimann, B. van Ginneken, M. Styner et al., Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets, IEEE Transactions on Medical Imaging 28(8), 2009
- [5] B. van Ginneken, T. Heimann, M. Styner, 3D Segmentation in the Clinic: A Grand Challenge, Proceedings of the 3D Segmentation in the Clinic: A Grand Challenge Workshop of the 9th International Conference on Medical Image Computing and Computer-Assisted Intervention, 2007
- [6] <http://www.ircad.fr/software/3Dircadb/3Dircadb1/index.php?lng=en>

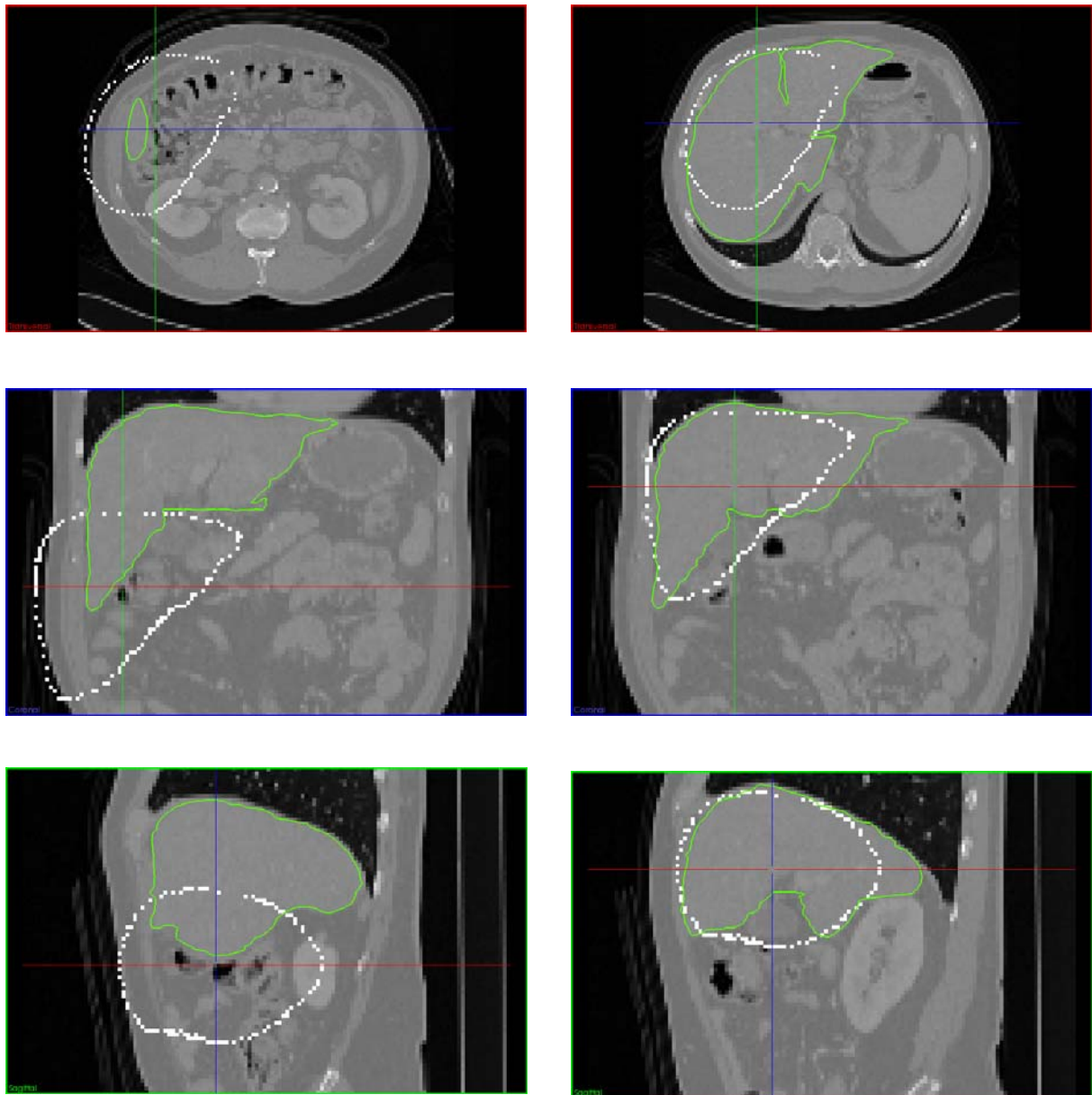


Abb. 3: Vergleich des Lokalisierungsergebnisses mit dem untrainierten Modell (links) und dem gewichteten (rechts) in axialer, koronaler und sagittaler Ansicht. Der von der GHT als bestes Ergebnis bestimmte Punkt wird durch das Fadenkreuz markiert. Zusätzlich sind das verwendete mittlere Modell (weiß) und die Segmentierung der Leber als Grundwahrheit (grün) eingezeichnet.