

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

J. Díaz Díaz¹, H. Eilers¹, A. Niemann², M. Leinung³, O. Majdani², T. Ortmaier¹

¹ Leibniz Universität Hannover, Institut für Mechatronische Systeme, Hannover, Germany

² Medizinische Hochschule Hannover, Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde, Hannover, Germany

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

Kontakt: 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.

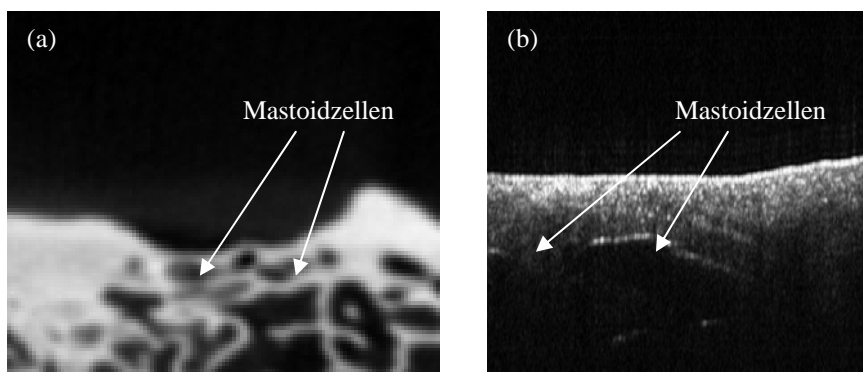


Abb. 1: Ausschnitte eines (a) fpVCT- (17,96 mm x 14,99 mm) und (b) OCT-Datensatzes (4 mm x 3 mm, letztere Angabe im Medium Luft) eines humanen Felsenbeinpräparats

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_φ 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.

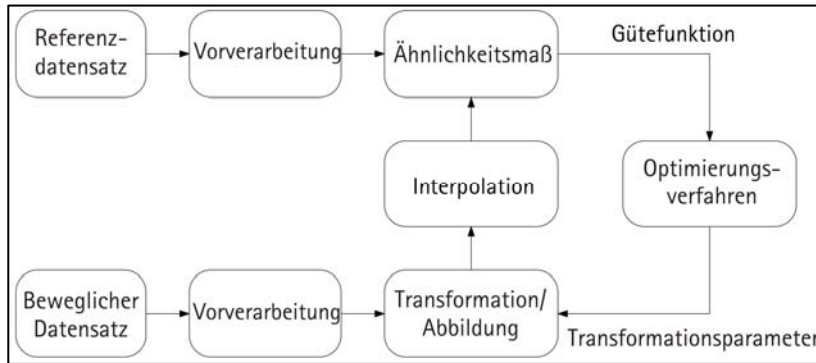


Abb. 2: Schematische Darstellung eines Registrierungsprozesses

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 ± 2 mm in jeder translatorischen und $\pm 2^\circ$ in jeder rotatorischen Komponente bestimmt.

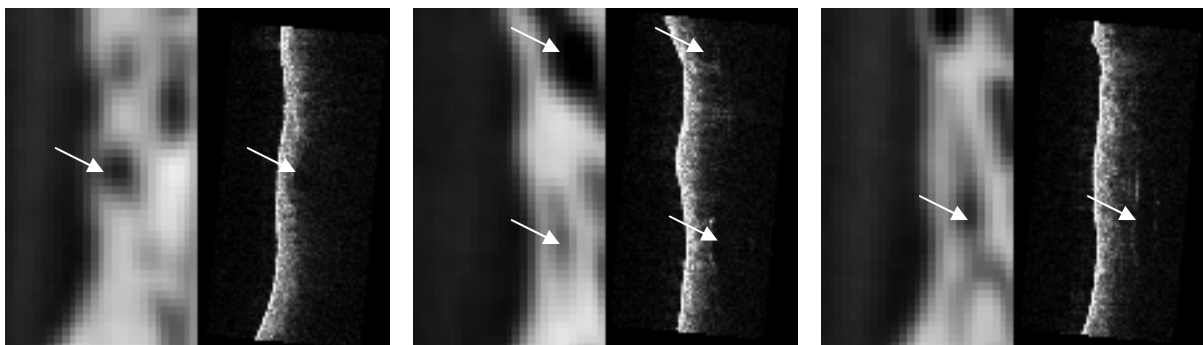


Abb. 3: Gegenüberstellung von 3 ausgewählten Schichten eines fpVCT-Datensatzes (jeweils links) mit den korrespondierenden transformierten Original OCT-Datensätzen (jeweils rechts)

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

$$e_x = \left\| x_{t,1} - x_{t,2} \right\|_2 \quad \text{für} \quad x_i = (t_{x,i}, t_{y,i}, t_{z,i}, \varphi_{z,i}, \varphi_{y,i}, \varphi_{z''i}) =: (x_{t,i}, x_{\varphi,i})$$

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_φ durch den Winkel in der Achsenwinkelstellung 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 Particle Swarm, Proceedings of the 6th 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