

# Intra-operative Augmented Reality(AR) mit magnetischer Navigation und multitexturbasiertem Volume Rendering in der minimal-invasiven Chirurgie

M. Scheuring<sup>1,2</sup>, C. Rezk-Salama<sup>1</sup>, H. Barfuß<sup>2</sup>, K. Barth<sup>2</sup>, A. Schneider<sup>3</sup>,  
G. Greiner<sup>1</sup>, G. Wessels<sup>3,2</sup>, H. Feussner<sup>4</sup>

<sup>1</sup>Lehrstuhl für Graphische Datenverarbeitung,  
Universität Erlangen-Nürnberg

<sup>2</sup>Siemens Medical Solutions, Grundlagenentwicklung, Erlangen

<sup>3</sup>Institut für minimal invasive interdisziplinäre Therapie und Intervention,  
Klinikum rechts der Isar, TU München

<sup>4</sup>Chirurgische Klinik und Poliklinik,  
Klinikum rechts der Isar, TU München

Email: [scheuring@informatik.uni-erlangen.de](mailto:scheuring@informatik.uni-erlangen.de)

<http://www9.informatik.uni-erlangen.de/Persons/Scheuring>

**Zusammenfassung** Augmented Reality(AR) Systeme in der Medizin erlauben die Überlagerung von laparoskopischen Videobildern mit dreidimensionalen Volumendaten. Unter Einsatz eines leistungsfähigen Volume Rendering Systems mit Multitexturen ist es möglich, Echtzeit-Frameraten bei der Überlagerung zu erzielen. Die Registrierung zwischen Patient und CT-Datensatz erfolgt anhand künstlicher Landmarken. Die Navigation erfolgt mit einem magnetischen Navigationssystem. Das System wurde intra-operativ in minimal-invasiven Interventionen eingesetzt und evaluiert.

## 1 Einleitung und medizinischer Hintergrund

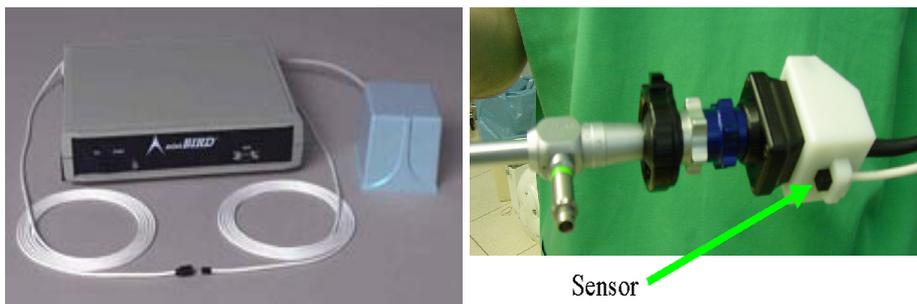
Minimal-invasive Eingriffe gewinnen in der Visceralchirurgie zunehmend an Bedeutung aufgrund der deutlichen Traumareduktion für den Patienten beim laparoskopischen Vorgehen. In der Diagnostik leistet auch die Laparoskopie ergänzend zu anderen modernen bildgebenden Verfahren wertvolle Dienste. Die Laparoskopie allein weist jedoch methodische Einschränkungen auf. So kann der Arzt zwar auf seine Erfahrung zurückgreifen, dennoch fehlt explizit die Volumeninformation. Die Beurteilung der dritten Dimension basiert somit im wesentlichen auf den prätherapeutisch gewonnenen Informationen, die mental auf die anatomischen Strukturen übertragen werden müssen. Dies betrifft unter anderem die Positionierung der Trokare für ein gezieltes Aufsuchen des Interventionsgebietes.

Einen deutlichen Informationszugewinn für den Chirurgen schaffen hier Augmented Reality (AR) Methoden, die es erlauben, Organe dreidimensional darzustellen und tiefliegende Gewebestrukturen durch Überlagerung des laparoskopischen Videobildes mit CT-Bildgebung zu visualisieren. In den letzten Jahren

wurden verschiedene AR-Methoden vorgestellt, die überwiegend bei neurochirurgischen und orthopädischen Eingriffen zum Einsatz kommen [3, 6, 8]. Diese Systeme bringen jedoch oft den Nachteil mit sich, daß teure Hardware eingesetzt werden muß, um eine interaktive Volumenvisualisierung bei gleichzeitig guter Bildqualität zu erzielen. Die verwendeten optischen Navigationssysteme (Polaris<sup>TM</sup>, Optotrak<sup>TM</sup>), die zur Bestimmung der Orientierung und Position des therapeutischen Instrumentes eingesetzt werden, bieten zwar hohe Genauigkeit, erfordern allerdings den Sichtkontakt zwischen Sender und Empfänger.

## 2 Navigation

Das hier vorgestellte AR-System für minimal-invasive Eingriffe im Abdomen stellt dem Chirurgen ein integriertes System zur Verfügung, das zusätzliche Informationen aus den Volumendatensätzen der präoperativen Bildgebung (CT) mit dem laparoskopischen Livebild im therapeutischen Ablauf zusammenführt, ohne dabei den Interventionsablauf zu beeinträchtigen. Um dies zu realisieren, wird ein *magnetisches* Positionserfassungssystem (Ascension miniBIRD<sup>TM</sup>, siehe Abb. 1 links) eingesetzt. Dieses System besitzt Sensoren mit sechs Freiheitsgraden und kann bei  $1.8\text{mm}$  Meßgenauigkeit mit unterschiedlichen Sensorgrößen von  $10\text{mm} \times 5\text{mm} \times 5\text{mm}$  und  $18\text{mm} \times 8\text{mm} \times 8\text{mm}$  bestückt werden. Soll die Position und die Orientierung des Laparoscops im Raum bestimmt werden, wird der *Sensor* an der Laparoskopiekamera befestigt (siehe Abb. 1 rechts). Um anschließend intra-operativ Positionswerte am Patienten zu erhalten, wird der *Sender* des magnetischen Navigationssystems aufgrund seiner geringen Ausmaße zwischen den Beinen des Patienten plaziert (siehe Abb. 5). Das klinische Personal wird somit bei der Intervention nicht behindert.



**Abbildung 1.** Navigationssystem Ascension Technology Corp. miniBIRD<sup>TM</sup> 800 (links), Befestigung des Sensors am Laparoskop (rechts).

### 3 Visualisierung

#### 3.1 Texturbasierte Interpolation

Bei der medizinischen Volumenvisualisierung ist nicht nur die Performance wichtig, sondern auch die Qualität der resultierenden Bilder, d.h. die Exaktheit bei der Darstellung der im Datensatz enthaltenen Information. Diese wird einerseits bestimmt durch das verwendete Interpolationsverfahren, andererseits durch die Abtastrate, d.h. die Anzahl der Abtastpunkte entlang eines Sehstrahls. Trilineare Interpolation wird hier als Standardanforderung angesehen. Zur Darstellung kleiner Strukturen – Nerven oder Blutgefäßen beispielsweise – muß darüber hinaus eine Möglichkeit gegeben sein, die Abtastrate entsprechend zu erhöhen.

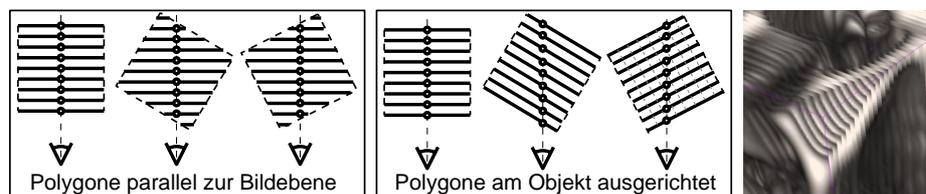
Texturbasierte Verfahren versuchen die Fähigkeit der Textureinheit zur hardwarebeschleunigten Interpolation auszunutzen. Da Graphikkarten jedoch keine volumetrischen Rendering-Primitive unterstützen, muß das Volumen durch eine bestimmte Anzahl polygonaler Schichten approximiert werden. Je nach vorhandener Hardware gibt es verschiedene Strategien.

#### 3.2 Verwendung von 3D-Texturen

Das Standardverfahren zum texturbasierten Volume-Rendering wurde 1994 von Cabral et al. [2] vorgestellt und erfordert eine Unterstützung von 3D-Texturen mit trilinearer Interpolation durch die Graphikhardware. Dieses in der Regel auf High-End Workstations implementierte Verfahren gilt als Referenz sowohl für Bildqualität als auch für Performanz. Das Volumen wird dann wie in Abb. 2 (links) dargestellt in Texturschichten zerlegt, die abhängig von der Blickrichtung immer parallel zur Bildebene verlaufen. Da sich in diesem Szenario die Schichtpolygone bei Änderung des Augpunktes ebenfalls verändern, müssen die Texturen für jedes Bild neu interpoliert werden. Da die Hardware trilineare Interpolation unterstützt, ist dies bei interaktiver Bildrate möglich.

#### 3.3 Verwendung von 2D-Texturen

Trilinear interpolierte 3D-Texturen werden von PC-Graphikkarten kaum unterstützt. Daher will man zumindest die bilineare Interpolation der 2D-Texturen



**Abbildung 2.** Ausrichtung der Polygone an der Bildebene (links) und am Volumen (Mitte). Bei zu geringer Abtastrate führen beide Verfahren zu visuellen Artefakten (rechts).

ausnutzen. Wie in Abb. 2 (Mitte) dargestellt, kann dies geschehen, indem man die Schichtpolygone anstatt an der Bildebene am Objekt selbst ausrichtet. Auf diese Weise wird trilineare durch bilineare Interpolation ersetzt. Ändert sich der Winkel zwischen Blickrichtung und Schichtnormale um mehr als  $45^\circ$ , so wird auf einen anderen Schichtstapel umgeschaltet. Demzufolge müssen die meisten der texturbasierten Algorithmen das Volumen dreifach im Speicher halten, jeweils ein Schichtenstapel für jede Raumachse. Dies ist auch das Grundprinzip beim *Shear-Warp-Algorithmus* [5], der als ein Standardverfahren für Software-Implementierungen gilt.

### 3.4 Verwendung von 2D-Multitexturen

Unter Multitexturen versteht man die Möglichkeit, ein einzelnes Polygon innerhalb der Rendering-Pipeline gleichzeitig mit mehreren Texturen zu versehen. In der Rasterisierung werden Farbe und Opazität eines Fragments dann als Ergebnis programmierbarer arithmetischer Operationen zwischen der Primärfarbe und mehreren Texturwerten bestimmt (*Texturkombination*). Wie in [9] beschrieben, kann diese flexible Framebuffer-Arithmetik effizient ausgenutzt werden, um echte trilineare Interpolation auch mit 2D-Texturen zu berechnen.

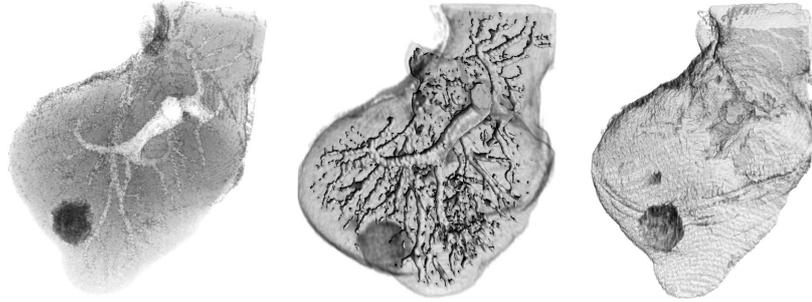
Die trilineare Interpolation einer Texturschicht  $S_{i+\alpha}$  kann beschrieben werden als Überblenden zwischen zwei aufeinanderfolgenden Schichten  $S_i$  und  $S_{i+1}$ :

$$S_{i+\alpha} = (1 - \alpha) \cdot S_i + \alpha \cdot S_{i+1}. \quad (1)$$

Die beiden Schichten  $S_i$  und  $S_{i+1}$  werden als Multitexturen aktiviert. Die bilineare Interpolation innerhalb der Texturschicht wird hierbei durch die Texturumgebung berechnet. Der fehlende dritte Interpolationsschritt in Richtung der Schichtnormalen erfolgt anschließend gemäß Gleichung 1 bei der Texturkombination. Berechnungen wie diese werden auf PC-Graphikkarten mittlerweile standardmäßig unterstützt, beispielsweise im neuen Standard *DirectX* 8.0 oder unter *OpenGL* durch NVidia's *Register Combiners* [1].

Moderne Multitextur-Hardware unterstützt weitaus komplexere Möglichkeiten der Texturkombination. Beispielsweise können die RGB-Kanäle einer Textur als Vektoren interpretiert und ein Skalarprodukt berechnet werden. Dies kann genutzt werden, um das Phong-Beleuchtungsmodell in die texturbasierte Volumendarstellung zu integrieren (siehe Abb. 3), wie es beispielsweise in [11] zur Darstellung *nicht-polygonaler* Isoflächen genutzt wurde. Die Basis dieses Verfahrens ist die Vorberechnung der Gradientenvektoren, deren  $x$ -,  $y$ - und  $z$ -Komponenten in einer RGB-Textur gespeichert werden. In der Rasterisierung wird dann das für die Beleuchtung benötigte Skalarprodukt zwischen Gradient und Lichtrichtung (bzw. *Halfway*-Vektor) bestimmt. Die Berechnung des Phong-Beleuchtungsmodells kann somit vollständig durch die Rasterisierungshardware erfolgen [9].

Durch Einsatz eines derartigen Volume Rendering Systems ist man nun in der Lage, sensorgesteuert durch das Navigationssystem entsprechend der Position und Orientierung des therapeutischen Instrumentes 3D-Ansichten aus dem prätherapeutischen Datensatz zu erzeugen.



**Abbildung 3.** CT-Leber mit Tumor: direktes Volume-Rendering ohne Beleuchtung als Standardverfahren (links), Isofläche mit diffuser und spekulärer Beleuchtung (Mitte), transparente Isofläche mit hervorgehobenem Tumor (rechts).

### 3.5 Multi-Planare-Rekonstruktion

Um den Informationszugewinn intra-operativ im OP zu erhöhen, wird zusätzlich zum direkten Volume Rendering die Visualisierung durch freie Multi-Planare-Rekonstruktion (MPR) unterstützt. Bei dieser Darstellung des prätherapeutischen Datensatzes werden ebenfalls sensorgesteuert die drei orthogonalen Schnittebenen parallel oder orthogonal zur Spitze des Laparoscops visualisiert.

## 4 Registrierung

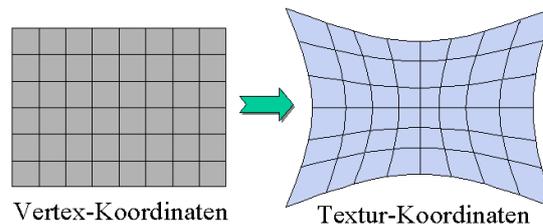
Die Registrierung zwischen Patient und präoperativ gemessenem CT-Datensatz erfolgt anhand künstlicher oder anatomischer Landmarken. Sie wird kurz vor der Operation einmalig durchgeführt. Dabei werden die Landmarken auf der Körperoberfläche mit dem Laparoskop angefahren, wobei parallel hierzu vom Navigationssystem die Positionswerte des Laparoscops ermittelt werden. Zusätzlich müssen die dazu korrespondierenden Volumenelemente im CT-Datensatz bestimmt werden. Anschließend wird mit Methoden der quadratischen Fehlerminimierung [7]) die Transformation berechnet. An dieser Stelle ist jedoch anzumerken, daß mit dieser Vorgehensweise ausschließlich eine rigide Transformation errechnet wird. Die Bewegung der inneren Organe kann in diesem Stadium nicht berücksichtigt werden.

## 5 Kalibrierung

Um ein realitätsnahes und genaues Video-Overlay zu erreichen, müssen zunächst zwei Kalibrierungsschritte durchgeführt werden. So gilt es einerseits die sogenannte *Tip-Transformation* vom Sensor des Navigationssystems zur Spitze des therapeutischen Instruments zu bestimmen, zum anderen die Kalibrierung der Laparoskopiekamera durchzuführen.

Die Bestimmung der Tip-Transformation kann experimentell durch Messung oder durch Algorithmen der *Hand-Eye-Calibration* bestimmt werden. Methoden hierzu werden in [4] beschrieben.

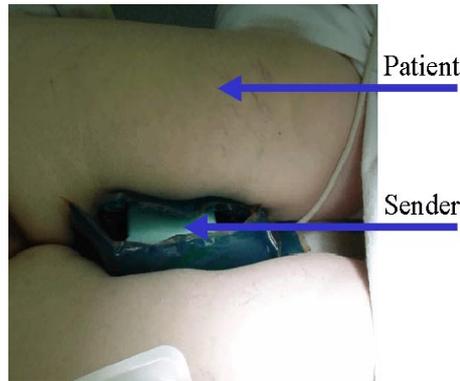
Die Kalibrierung der Laparoskopiekamera kann mit bekannten Algorithmen nach Tsai [10] durchgeführt werden, die die intrinsischen Parameter als Resultat liefern. Sollen nun mit Hilfe des Volume Rendering Systems die Linseneigenschaften der Kamera nachgebildet werden, müssen die intrinsischen Parameter eingesetzt werden, um das Viewing-Frustum in OpenGL nachzubilden. Eine sich anschließende exakte Überlagerung des Videobildes kann jedoch nur gelingen, wenn gleichzeitig die sehr starke radiale und tangentielle Linsenverzerrung mitberücksichtigt und somit aus dem Videobild herausgerechnet wird. Da dies jedoch bei Standard PAL bei einer Auflösung von  $768 \times 576$  Bildpunkten erheblichen Rechenaufwand bedeutet, wurde eine schnelle Hardware-Implementierung mit Hilfe von 2D-Texturen in OpenGL eingesetzt (siehe Abb. 4). Bei diesem Verfahren wird das Videobild als OpenGL-Textur definiert und in  $n \times m$  Patches aufgeteilt. Mit Hilfe der radialen und tangentialen Verzerrungskoeffizienten, die durch die Kamerakalibrierung bestimmt werden, können anschließend hochperformant durch gezieltes Verschieben der Texturkoordinaten die Videobilder entzerrt werden. Die Ermittlung der Bildpunkte der Videotextur innerhalb der Patches erfolgt dann durch bilineare Interpolation der Hardware. Verwendet man  $32 \times 32$  Patches, erhält man 10-15 Frames bei sehr guter Bildqualität.



**Abbildung 4.** Hardwarebasierte Linsenentzerrung mit Hilfe von 2D-Texturen. Die Videotextur wird in  $n \times m$  Patches aufgeteilt.

## 6 Implementierung und OP-Einsatz

Um medizinischem Fachpersonal den Zugriff und den Einsatz des beschriebenen Volume Rendering Systems zu ermöglichen, wurde eine Integration in die medizinische Softwareplattform *syngo* der Firma Siemens AG zu Evaluierungszwecken vorgenommen, die jedoch auf Forschungs- und Testzwecke am Institut beschränkt ist und nicht als Produkt vorliegt. Durch die Realisierung ähnlich einer *syngo*-Komponente ist es möglich, die Applikation in den Workflow des Mediziners einzubinden. So erlaubt es die Anwendung nun, direkt vom CT-



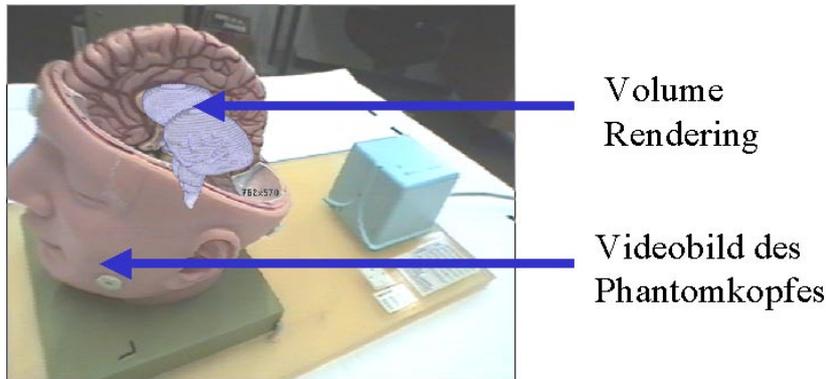
**Abbildung 5.** Lagerung des Magneten zwischen den Beinen des Patienten.

oder MR-Scanner aufgenommene Tomographiedaten per DICOM-Protokoll zu transferieren und in einem Patientbrowser auszuwählen. Anschließend können die Daten mit den in den vorigen Kapiteln beschriebenen Visualisierungstechniken dargestellt werden. Als Basisplattform dient ein Standard-PC (Pentium III 800 MHz) unter Windows NT 4.0<sup>TM</sup>. Zur Visualisierung wird eine Graphikkarte mit einem NVidia GeForce 2 GTS<sup>TM</sup>-Chip mit 64 MB DDR verwendet.

In Genauigkeitsstudien wurden zunächst Testobjekte (Phantomkopf, Plastiktorso, Plastikwürfel) vermessen, um den Einfluß metallischer Gegenstände und Instrumente im OP auf das magnetische Navigationssystem zu ermitteln. Abb. 6 zeigt die Überlagerung des Phantomkopfes mit Strukturen des Gehirns. Dabei ergab sich bei allen Testobjekten eine Genauigkeit der Überlagerung von besser als 5mm. Anschließend erfolgten vier Anwendungen an Patienten in Ergänzung des Routineverfahrens. Dabei wurde der Sender des Navigationssystems zwischen den Beinen des Patienten gelagert (siehe Abb. 5). Um während des Eingriffs sowohl das originale laparoskopische Videobild, als auch das überlagerte Bild optimal zur Verfügung zu haben, wurde das virtuelle Bild mittels eines Flachbildschirms dargestellt.

## 7 Fazit

Das vorhandene System bietet Chirurgen eine wertvolle Unterstützung bei Interventionen. Das magnetische Navigationssystem behindert den Eingriff nicht. Mit preisgünstiger Hardware wird intra-operativ zusätzliche 3D-Information mit MPR und direktem Volume Rendering in Echtzeit bereitgestellt. Bei verschiedenen Organen sind noch keine Genauigkeitsaussagen aus der Praxis möglich. Um in solchen Fällen die Genauigkeit zu verbessern, soll das System um eine hardwarebeschleunigte nichtlineare Registrierung erweitert werden. Dies wird Gegenstand zukünftiger Forschungsarbeit sein.



**Abbildung 6.** Überlagerung des Phantomkopfes mit volumen-gerendertem CT-Bild.

## Literatur

1. GeForce 256 and RIVA TNT Combiners. <http://www.nvidia.com/Developer>.
2. B. Cabral, N. Cam, and J. Foran. Accelerated Volume Rendering and Tomographic Reconstruction Using Texture Mapping Hardware. *ACM Symp. on Vol. Vis.*, 1994.
3. E. Grimsen, M. Leventon, G. Ettinger, A. Chabrerie, F. Ozlen, S. Nakajama, H. Atsumi, R. Kikinis, and P. Black. Clinical experience with a high precision image guided neurosurgery system. *MICCAI 98*, pages 63–73, 1998.
4. R. Horaud and F. Dornaika. Hand-eye calibration. *IJRR*, 14(3):195–210, June 1995.
5. P. Lacroute and M. Levoy. Fast Volume Rendering Using a Shear-Warp Factorization of the Viewing Transform. *Comp. Graphics*, 28(4), 1994.
6. M. Blackwell, C. Nikou, A. M. DiGioia, and T. Kanade. An image overlay system for medical data visualization. *Proc. of SPIE Medical Imaging Analysis 2000*, 4:62–72, 2000.
7. William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge (UK) and New York, 1st edition, 1986.
8. R. Shahidi, B. Wang, M. Epitoux, J. Adler, and G. Steinberg. Intraoperative videos and volumetric image fusion. *CARS 99*, pages 625–630, 1999.
9. C. Rezk-Salama, K. Engel, M. Bauer, G. Greiner, and T. Ertl. Interactive Volume Rendering on Standard PC Graphics Hardware Using Multi-Textures and Multi-Stage Rasterization. In *Eurographics/SIGGRAPH Worksh. on Graphics Hardware*. ACM, 2000.
10. R. Tsai. A versatile camera calibration technique for high-accuracy 3-D machine vision metrology using off-the-shelf TV cameras and lenses. In L. Wolff, S. Shafer, and G. Healey, editors, *Radiometry – (Physics-Based Vision)*. Jones and Bartlett, 1992.
11. R. Westermann and T. Ertl. Efficiently Using Graphics Hardware in Volume Rendering Applications. In *Proc. of SIGGRAPH*, Comp. Graph. Conf. Series, 1998.