

Sensorfusion von Augmented Reality Komponenten für die medizinische Navigation

Tim SUTHAU

Zusammenfassung

Augmented Reality beschreibt eine Technik, welche es erlaubt virtuelle Daten mit der realen Welt zu verbinden. Dies kann beispielsweise mit Hilfe eines Head Mounted Display erfolgen. In der vorliegenden Arbeit steht die mathematische Modellierung der einzelnen Komponenten eines Augmented Reality Systems und deren geometrischer Bezug im Mittelpunkt. Dazu gehören insbesondere das Markertracking, die Kamerakalibrierung, die Head-Mounted Display (HMD) Kalibrierung, sowie das Eyetracking, um das Verrutschen des HMD zu korrigieren. Ergebnisse der Umsetzung werden präsentiert.

1 Einleitung

Seit ca. 3 Jahren wird am Fachgebiet Computer Vision & Remote Sensing der TU Berlin die Einsetzbarkeit von Augmented Reality Techniken in der Leberchirurgie untersucht. Bei einem Augmented Reality System werden reale Bilder durch zusätzliche Informationen ergänzt. Augmented Reality stellt somit die Verbindung zwischen realer und virtueller Welt dar. Im klinischen Alltag der Chirurgie stellt die Bereitstellung präoperativer Daten eine unverzichtbare Hilfe für den Operateur dar. Diese Daten werden aus präoperativen Aufnahmen (Computertomographie, Magnetresonanztomographie, Ultraschall u.a.) gewonnen. Segmentierungsergebnisse und zwei- bzw. dreidimensionale Darstellungen, welche aus diesen Aufnahmen gewonnen werden, dienen dem Chirurgen zur notwendigen Orientierung während der Operation. Diese Daten können in einer Augmented Reality Umgebung dem Arzt als virtuelle Daten zugänglich gemacht werden. Dem Operateur soll ein Hilfsmittel zur Seite gestellt werden, welches zweckmäßig zusätzliche Informationen darstellt und eine komfortable Navigation ermöglicht. Die exakte Überlagerung der virtuellen Daten mit dem realen Objekt stellt dabei ein wesentliches Grundproblem dar. Diese Aufgabe soll im Mittelpunkt dieser Arbeit stehen. Dies erfordert das Verständnis des praktischen klinischen Gesamt szenarios (SUTHAU2002), die Auswahl geeigneter Technologie (Hardware, wie Displays, Trackingsysteme und Sensoren) (SUTHAU2003), die mathematische Modellierung der Zusammenhänge und eine entsprechende Umsetzung. Im Folgenden werden die einzelnen Komponenten solch eines Augmented Reality Systems mathematisch modelliert und entsprechend erläutert.

Grundlage dieser Arbeit sind, neben den grundlegenden Techniken und Forschungsergebnissen der Photogrammetrie und Computer Vision, umfangreiche Forschungen und Umsetzungen der Abteilung Medizinische und Biologische Informatik des Deutschen Krebsforschungszentrums in Heidelberg. Die Abteilung Medizinische und Biologische Informatik und die Chirurgische und Radiologische Klinik der Universität Heidelberg haben einen Prototyp eines bildgestützten Navigations-Systems (image-guided surgery system, IGSS)

für die Anwendung in der onkologischen Leberchirurgie entwickelt. Dieses IGSS wird den Chirurgen in die Lage versetzen, sein Instrument in räumlicher Beziehung zu wichtigen Strukturen in der Leber zu sehen. Hierfür ist die Auswertung präoperativer Daten (Segmentierung der Lebergefäßbäume) erforderlich.

Für eine erfolgreiche Resektion eines Lebertumors ist die Kenntnis der genauen Lage des tumorösen Gewebes und des einzuhaltenden, umgebenden Sicherheitsabstandes erforderlich (HASSENFLUG 2001; HASSENFLUG 2001B; VETTER 2001). Die Chirurgen interessiert die möglichst exakte Übertragung der präoperativ am Computermodell geplanten Resektionsflächen auf den Operationssitus. Die bisherige Präsentation erfolgt mit Hilfe eines autostereoskopischen Bildschirms. Aufgabe des Fachgebiets Computer Vision & Remote Sensing ist die Bereitstellung eines Augmented Reality Szenarios für diese Umgebung im Sinne eines Nahbereichs-See-Through-Systems für die Viszeralchirurgie. Wie oben beschrieben sollen die virtuellen Daten der präoperativen Planung dem Chirurg bei Bedarf direkt in das Sichtfeld projiziert werden. Der Unterschied zwischen der heutigen und zukünftigen Arbeitsweise ist in Abbildung 1 illustriert.

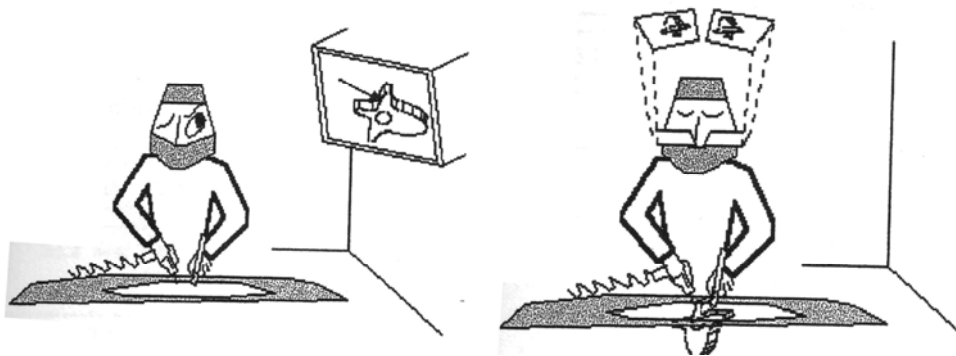


Abb.1: Heutige und zukünftige Arbeitsweise der computergestützten OP
nach (PEUCHOT1993)

2 Komponenten des Augmented Reality Systems

Das Herzstück eines Augmented Reality Systems ist das zu verwendende Display. Zur Visualisierung findet ein Virtual Retinal Display „Nomad“ der Firma Microvision Anwendung. Dieses muss dann in den Gesamtzusammenhang eingebunden werden. Für die mathematische Beschreibung von Augmented Reality Systemen steht die Berechnung der räumlichen Zusammenhänge einzelner Komponenten im Mittelpunkt. Zur Identifizierung der Lage des Objektes werden bekannte Objektpunkte in Form von Markern (Passpunkte) verwendet. Eine Kamera übernimmt dabei die ständige Positionsbestimmung (Tracking) des Markers. Soll diese Positionsbestimmung im Raum erfolgen ist eine Kalibrierung der Kamera notwendig. Mit diesen Komponenten kann der Bezug zwischen dem Objekt (mit festem Bezug zu den Markern) und den virtuellen Daten hergestellt werden. Sollen die virtuellen Daten positionsgenau auf einem Head-Mounted Display visualisiert werden, ist die Kalibrierung des HMD notwendig.

Voraussetzung für die Kalibrierung ist eine gleichzeitige Erkennung der Marker durch die Trackingkamera. Bei der Kalibrierung des HMD wird der geometrische Zusammenhang zwischen dem Objekt (Marker) dem HMD-Bildschirmkoordinaten und dem Auge hergestellt. Die Kalibrierung des HMD setzt eine feste Verbindung zwischen HMD und Kopf (Auge) voraus. Kann dies nicht gewährleistet werden muss nach einem Verrutschen des HMD die Kalibrierung automatisch verbessert oder neu bestimmt werden. Abbildung 2 zeigt das Kopfband mit einigen Komponenten.

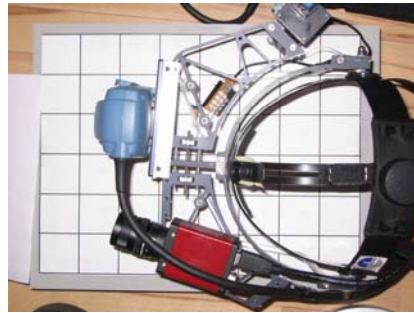


Abb.2: Kopfband

Zusammengefasst sind folgende Aufgaben mathematisch zu beschreiben:

- Positionsbestimmung (Tracking) von Markern (Passpunkte)
- Kamerakalibrierung (Orientierung / Verzeichnung)
- HMD Kalibrierung (Orientierung / Verzeichnung)
- Verrutschen des HMD (Korrektur)

Die grundlegenden Anforderungen sind Genauigkeit, Echtzeitfähigkeit und Nutzerakzeptanz. Die Photogrammetrie befasst sich klassisch mit der Geometrie der Aufnahme, der Orientierung und der Rekonstruktion von Objekten aus Bildern. Sie ist somit eine Technik zur Erfassung räumlicher Objekte aus Bildinformationen. Die grundlegenden Anforderungen bei einem Augmented Reality System, wie Echtzeitfähigkeit und Genauigkeit erfordern die Bereitstellung direkter Lösungsverfahren. Die Entwicklung der geometrischen Rekonstruktion im Bereich Computer Vision stützt sich ebenso auf die konsequente Nutzung der algebraischen projektiven Geometrie. Basis sind homogene Koordinaten, die die Repräsentation geometrischer Elemente, ihre Verknüpfung und ihre Transformation repräsentieren. Grundlage zur Lösung der allgemeinen Aufgabenstellung bilden Kenntnisse in der Bildverarbeitung, der linearen Algebra (lineare Gleichungssysteme, Eigenwert- und Singulärwertzerlegung), der Statistik und Ausgleichsrechnung sowie der Physik der geometrischen Optik.

3 Mathematische Modellierung der Komponenten

3.1 Markererkennung – Positionsbestimmung und Zielverfolgung

Der Zusammenhang zwischen den einzelnen Koordinatensystemen bzw. der Bezug zwischen Objektraum und Kopfraum wird durch die Verwendung von (Passpunkt-) Markern realisiert. Diese werden mit einer kalibrierten Kamera beobachtet. Hier sollen quadratische Marker, wie sie von (KATO 1999) und dem ARToolKit verwendet werden, ihre Anwendung finden. (Abb. 3) Diese Marker haben die Eigenschaften, dass es möglich ist, mit Hilfe nur eines Markers die räumliche Beziehung zwischen Objektraum (Markerraum) und Kopfraum (Kameraraum) herzustellen.

Nach erfolgter Identifizierung eines Markers kann dies über die Kenntnis der Abstände der Markereckpunkte und über Tests der Parallelität und Orthogonalität der Markerkanten erfolgen. Dabei liefern die Abstände der Markereckpunkte Informationen zur erfolgten Maßstabsveränderung und die Markerkanten Kenntnisse zur erfolgten Rotation. Das Feld im Inneren der Marker dient dabei der Identifizierung der Marker bzw. der internen Nummerierung bei Nutzung von mehreren Markern. Diese Art der Marker führt damit zu einer gewissen Geschwindigkeits- und Genauigkeitssteigerung.



Abb. 3: (Passpunkt-) Marker

Die Markererkennung setzt sich aus der Lösung folgender 3 Aufgaben zusammen:

- 1) Lokalisierung der Marker im Bild (Schwellwertbildung, Segmentierung)
- 2) Bestimmung der 4 Markereckpunkte und -kanten,
- 3) Berechnung der Transformationsmatrix

Markertracking ist die Suche nach der Transformation (\mathbf{T}_{TK_M} mit Rotation \mathbf{R} , Translation \mathbf{t}) zwischen dem Trackingkamerakoordinatensystem (\mathbf{X}_{TK}) und Objektraumkoordinatensystem der Marker (\mathbf{X}_M):

$$\mathbf{X}_{TK} \cong \mathbf{T}_{TK_M} \cdot \mathbf{X}_M \quad \text{bzw.} \quad \begin{bmatrix} X_{TK} \\ Y_{TK} \\ Z_{TK} \\ 1 \end{bmatrix} \cong \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} \quad (1)$$

3.2 Kamerakalibrierung

Zur Kamerakalibrierung wird ein regelmäßiges Gitter (Abb.4) in definierten Abständen zur Kamera angemessen. Die Koordinaten der angemessenen Kreuze sind somit im lokalen 3D Koordinatensystem bekannt. Mehrere Paare von Kamerabildkoordinaten x'_{TK}, y'_{TK} und den dazugehörigen lokalen 3D Koordinaten der Gitterkoordinaten X_G, Y_G, Z_G können zur Bestimmung der Projektionsmatrix \mathbf{P}_{TK} genutzt werden.

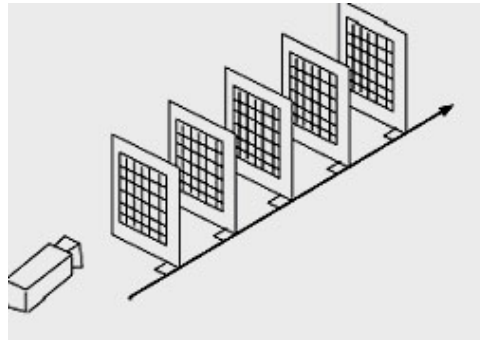


Abb. 4: Gitter zur Kamerakalibrierung

Die innere Orientierung ist definiert:

$$\begin{bmatrix} x'_{TK} \\ y'_{TK} \\ w'_{TK} \end{bmatrix} \cong \mathbf{K}_{TK} \begin{bmatrix} x_{TK} \\ y_{TK} \\ w_{TK} \end{bmatrix} \quad \text{mit} \quad \mathbf{K}_{TK} = \begin{bmatrix} m_x c & 0 & x'_h \\ 0 & m_y c & y'_h \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Es ergeben sich die Beziehungen zwischen dem Kamerabildkoordinatensystem x'_{TK}, y'_{TK} , dem Kamerakoordinatensystem X_{TK}, Y_{TK}, Z_{TK} und dem Gitterkoordinatensystem X_G, Y_G, Z_G :

$$\begin{bmatrix} x'_{TK} \\ y'_{TK} \\ 1 \end{bmatrix} \cong \mathbf{K}_{TK} \cdot \mathbf{\Pi}_0 \cdot \mathbf{H}_{TK,G} \begin{bmatrix} X_G \\ Y_G \\ Z_G \\ 1 \end{bmatrix} \quad \text{mit} \quad \mathbf{H}_{TK,G} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{X}_0 \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3)$$

Zusammengefasst zur Projektionsmatrix:

$$\begin{bmatrix} x'_{TK} \\ y'_{TK} \\ 1 \end{bmatrix} \cong \mathbf{P}_{TK} \begin{bmatrix} X_G \\ Y_G \\ Z_G \\ 1 \end{bmatrix} \quad \text{oder} \quad \begin{bmatrix} x'_{TK} \\ y'_{TK} \\ 1 \end{bmatrix} \cong \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & 1 \end{bmatrix} \begin{bmatrix} X_G \\ Y_G \\ Z_G \\ 1 \end{bmatrix} \quad (4)$$

Die Matrix \mathbf{P}_{TK} wird dabei durch mehrere Paare von Kamerabildkoordinaten x'_{TK}, y'_{TK} und lokalen 3D Koordinaten X_G, Y_G, Z_G der Gitterkreuze bestimmt. Die Bestimmung der 11 Parameter von \mathbf{P}_{TK} erfolgt durch die Lösung einer überbestimmten Direkten Linearen Transformation, mittels vermittelnder Ausgleichsrechnung. Nach deren Bestimmung kann \mathbf{P}_{TK} durch RQ-Zerlegung in $\mathbf{K}_{TK}, \mathbf{R}$ und \mathbf{X}_0 (damit $\mathbf{H}_{TK,G}$) zerlegt werden.

3.3 HMD Kalibrierung

Für die HMD Kalibrierung kann das gleiche mathematische Modell (projektive Abbildung) wie bei der Kamerakalibrierung genutzt werden. Als bekannt genutzt werden viele Paare der HMD-Bildkoordinaten x'_{HMD}, y'_{HMD} und deren zugehörige Kameraposition X_{TK}, Y_{TK}, Z_{TK} aus der Kamerakalibrierung. Die zugehörige Kameraposition wird durch das gleichzeitige Beobachten eines Kalibrierungsmarkers gewährleistet, d.h. der Beobachter betrachtet den Marker durch den Bildschirm eines HMD und bringt ein eingeblendetes Fadenkreuz im Display mit dem Marker in Übereinstimmung. Im gleichen Augenblick wird der Marker durch eine kalibrierte Kamera getrackt und somit geometrisch zugeordnet. Der direkte Zusammenhang zwischen HMD-Bildkoordinaten x'_{HMD}, y'_{HMD} und deren zugehörige Kameraposition X_{TK}, Y_{TK}, Z_{TK} ist gegeben.

Folgender mathematische Zusammenhang ist gegeben:

Beziehung Marker – Trackingkamera (Beobachtung von mindestens 3 Passpunkten):

$$\begin{bmatrix} x'_{TK} \\ y'_{TK} \\ 1 \end{bmatrix} \cong \mathbf{P}_{TK} \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} \quad (5)$$

Beziehung Marker – HMD Bildschirmkoordinaten (Fadenkreuzeinstellung):

$$\begin{bmatrix} x'_{HMD} \\ y'_{HMD} \\ 1 \end{bmatrix} \cong \mathbf{P}_{HMD} \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} \quad (6)$$

Mit einer Fadenkreuzeinstellung ergibt sich jetzt:

- \mathbf{x}'_{HMD} ... Fadenkreuzkoordinaten auf dem HMD-Display
- \mathbf{x}'_{TK} ... Koordinaten des angezielten (Pass-) Markerpunktes im Kamerabild
- \mathbf{X}_M ... Kameraobjektkoordinaten des angezielten Passpunktes im Kamerabild

Transformation des Punktes im Fadenkreuz:

Bei jeder Einstellung des Fadenkreuzes werden gleichzeitig die Objektkoordinaten der Marker bestimmt, da die kalibrierte Trackingkamera den Marker erfasst. Es sind \mathbf{P}_{TK} und \mathbf{x}'_{TK} zur Bestimmung von \mathbf{X}_M aus Gleichung (5) bekannt. Jetzt kann (äquivalent zur Kamerakalibrierung) \mathbf{P}_{HMD} durch mehrere Paare von HMD- Displaykoordinaten (Fadenkreuz) x'_{HMD}, y'_{HMD} und lokalen 3D Koordinaten X_M, Y_M, Z_M der Marker bestimmt. Dazu reichen wiederum 6 Punkte aus. Die Bestimmung der 11 Parameter von \mathbf{P}_{HMD} erfolgt durch die Lösung einer überbestimmten Direkten Linearen Transformation, mittels vermittelnder Ausgleichsrechnung. Nach deren Bestimmung kann \mathbf{P}_{HMD} durch RQ-Zerlegung in $\mathbf{K}_{HMD}, \mathbf{R}$ und \mathbf{X}_O zerlegt werden. Da \mathbf{P}_{HMD} jetzt bekannt ist, können die Objektdetails nun in das Display eingeblendet werden.

Wo muss das Fadenkreuz im Bildschirm erscheinen?

Der Anschauung nach, kann die Displayebene um die die mit einem einzigen Sehstrahl entstehende Achse rotieren, wenn das Kreuz immer im selben Displaypunkt erscheint. D.h. die \mathbf{x}'_{HMD} müssen sich verändern und am besten das ganze Display abdecken.

4 Ergebnis der Umsetzung und Ausblick

Die Umsetzung der Kalibrierung des oben beschriebenen Augmented Reality Systems und seiner Komponenten erfolgte auf der Grundlage der Open Source Software (ARToolKit). Die einzelnen Module wurden für verschiedene Genauigkeitstest (SCHÄFER 2003, PIETSCHER 2004, WACKROW 2004) entsprechend den Untersuchungszielen teilweise stark verändert. In der Abbildung 5 ist die Sicht durch das verwendete Virtual Retinal Display (VRD) nach erfolgter Kamera- und HMD Kalibrierung zu sehen. Die dort zu sehende Projektion wurde zur Genauigkeitsanalyse verwendet. Dabei sollten die roten Quadrate bei genauer Überlagerung den Ecken



Abb. 5: Sicht durch das VRD

der Marker entsprechen. Diese Überlagerungsgenauigkeit wurde für Abstände (Marker – Kamera) von 30cm bis 4m getestet. Dabei konnten gute Ergebnisse im interessierenden Entfernungsbereich von 50-80 cm (Abstand des Objektes bei einer Operation) erzielt werden. Hier erreichen die Abweichungen ihr Minimum und liegen nahe bei 10 Pixeln, was einer Genauigkeit im Objektraum unter 3mm entspricht. Fehler treten hauptsächlich durch fehlerbehaftetes Markertracking (Ausreißer) bei der HMD-Kalibrierung auf. Dieser Fehler ist dabei größer als der Fehler bei fehlender Verzeichnungskorrektur (WACKROW 2004) des VRD. Die Abbildungsgenauigkeit virtueller Daten auf dem VRD-Bildschirm ist auch von der Genauigkeit bzw. Ungenauigkeit des Lasers des VRD abhängig. Dabei kommt es zu Ungenauigkeiten bei der Darstellung von. Die Größe der Unregelmäßigkeiten liegt im Bereich von ca. 1-2 Pixel auf dem VRD oder entsprechend bis zu 5 Pixel im digitalen Bild (Aufgrund der höheren Auslösung der Aufnahmekamera). Eine Korrektur dieses gerätespezifischen Fehlers ist nur schwer möglich, da nicht erkennbar ist, wann der Laser in eine bestimmte Richtung verschoben ist.

Bisher gilt die Voraussetzung, dass das HMD fest mit dem Kopf verbunden ist. Verrutscht das HMD erfolgt eine Bewegung des HMD gegenüber dem Kopf also eine Bewegung, die sowohl Rotationen als auch Translationen enthalten kann. Schon kleine Bewegungen führen zu einer signifikanten Verschlechterung der Überlagerungsgenauigkeit. Dies lässt sich schnell mit dem Strahlensatz berechnen. So führt ein Verrutschen des HMD parallel zur Augenbildebene um 1 mm bei einem Abstand des HMD vom Auge von 2 (3) cm und einem Abstand des virtuellen Objektes vom HMD von 80 cm zu einer um 4 (3) cm falschen Einblendung des virtuellen Objektes. Durch die Einbindung eines Eye-Tracking Verfahrens soll zukünftig das Verrutschen erkannt und die HMD Kalibrierung entsprechend korrigiert werden

5 Literatur

- ARToolKit: Webseite: <http://www.hitl.washington.edu/artoolkit/> , neueste Windows Version unter <http://sourceforge.net/projects/artoolkit/> , Stand: Januar 2005
- Hassenpflug (2001): P. Hassenpflug, M.Vetter, M.Thorn, C. Cárdenas, G. Glombitza, W.Lamadé, G.M. Richter and H.-P. Meinzer, *Navigation in liver surgery – requirement analysis and possible solutions*, Proceedings of the Fifth Korea-Germany Joint Workshop on Advanced Medical Image Processing, M.H. Kim and H.-P. Meinzer, Ewha Womans University, Seoul, Korea, May 15th/16th, 2001
- Hassenpflug (2001b): Hassenpflug P, Vetter M, Cárdenas S, Thorn M, Meinzer HP: *Navigation in liver surgery - results of a requirement analysis*. CARS 2001 Proceedings: CARS2001 - Computer Assisted Radiology and Surgery, Elsevier Science, Amsterdam, vol. ICS 1230 (2001) 1162.
- Kato (1999): Kato, H.; Billinghamurst, M.; *Augmented Reality, 1999. Marker tracking and HMD calibration for a video-based augmented reality conferencing system (IWAR '99)* Proceedings. 2nd IEEE and ACM International Workshop on , 1999, Page(s): 85 -94
- Pietscher (2004): S. Pietscher, *Tracking für Augmented Reality*. Diplomarbeit, TU Berlin, Fachgebiet Photogrammetrie und Kartographie, Februar 2004, unveröffentlicht
- Peuchot (1993): B. Peuchot, *Camera virtual equivalent model: 0.01 pixel detectors*, Special issue on 3D Advanced Image Processing in Medicine in Computerized Medical Imaging and Graphics, 1993, 17 (4/5), pp.289-294
- Schäfer (2003): H. Schäfer, *Kalibrierungen für Augmented Reality*. Diplomarbeit, TU Berlin, Fachgebiet Photogrammetrie und Kartographie, Oktober 2003, unveröffentlicht
- Suthau (2002): T.Suthau, M. Vetter, P. Hassenpflug, H.-P. Meinzer, O. Hellwich *A concept work for Augmented Reality visualisation based on a medical application in liver surgery*. ISPRS, the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences;Volume XXXIV, Part5, Commission V, Proceedings of the ISPRS Commission V Symposium, close-range imaging, long-range vision, Corfu 2002, Greece, pp. 274-280
- Suthau (2003): T. Suthau *Augmented Reality Techniken für den Einsatz in der Leberchirurgie*. E. Seyfert (eds.), Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, Band 12, 23. Wissenschaftlich-Technische Jahrestagung der DGPF , 09.-11. September 2003, Bochum, S.301
- Wackrow (2004)R. Wackrow, *Berechnung und Korrektur der Verzeichnung eines Head Mounted Display*. Diplomarbeit, TU Berlin, Fachgebiet Computer Vision & Remote Sensing, 08/ 2004, unveröffentlicht
- Vetter (2001): Vetter M, Hassenpflug P, Cárdenas C, Thorn M, Glombitza G, Meinzer HP: *Navigation in der Leberchirurgie - Ergebnisse einer Anforderungsanalyse*. In: Lehmann T, Handels H (Eds). *Informatik Aktuell – Bildverarbeitung in der Medizin 2001 – Algorithmen, Systeme, Anwendungen*. Berlin, Heidelberg, New York: Springer, 2001.