

## **Lehrstuhl für Informatik 5 (Mustererkennung)**

**Anschrift:** Martensstrasse 3, 91058 Erlangen

**Tel.:** +49.9131.85.27775      **Fax.:** +49.9131.303811

**E-Mail:** info@i5.informatik.uni-erlangen.de

### **Leitung:**

Prof. Dr.-Ing. Heinrich Niemann (bis 30.09.2005)

Prof. Dr.-Ing. Joachim Hornegger (ab 01.10.2005)

### **Professor:**

Prof. Dr.-Ing. Joachim Hornegger (bis 30.09.2005)

### **Emeritus:**

Prof. em. Dr.-Ing. Heinrich Niemann (ab 01.10.2005)

### **Sekretariat:**

Iris Koppe

Kristina Müller

Silvia Zettelmaier

### **Sekretariat SFB 603:**

Martina Montel-Kandy

### **Leitung Bildverarbeitung:**

Dipl.-Inf. Jochen Schmidt (bis 30.09.2005)

Dipl.-Inf. Benjamin Deutsch (ab 01.10.2005)

### **Bildverarbeitung:**

Ing. Radim Chrastek

Dipl.-Ing. Christian Derichs

Dipl.-Inf. Christoph Gräßl

Dipl.-Inf. Marcin Grzegorzek

Dipl.-Inf. Ingo Scholz

Dipl.-Inf. Florian Vogt

Dipl.-Math. (FH) Stefan Wenhardt

Dipl.-Inf. Timo Zinßer

**Leitung Medizinische Bildverarbeitung:**

Dipl.-Inf. Marcus Prümmer

**Medizinische Bildverarbeitung:**

Dipl. -Inf. Rüdiger Block  
Dipl. -Inf. Anja Borsdorf  
Dipl. -Inf. Volker Daum  
Dipl. -Inf. Dieter Hahn  
Dipl.-Inf. Volker Daum  
Jingfeng Han, M. Sc.  
Dipl.-Ing. Kurt Höller  
Dipl.-Ing. (FH) Stefan Hoppe  
Dipl.-Inf. Florian Jäger  
Dipl. Med-Inf. Jochen Penne  
Dipl.-Inf. Holger Scherl

**Leitung Sprachverarbeitung:**

Dr.-Ing. Elmar Nöth

**Sprachverarbeitung:**

Dr. phil. Anton Batliner  
Dipl.-Math. Silvia Dennerlein  
Dipl.-Inf. Christian Hacker  
Dipl.-Inf. Tino Haderlein  
Dipl. Inf. Florian Hönig  
Dipl.-Inf. Axel Horndasch  
Dipl.-Inf. Michael Levit  
Dipl.-Inf. Andreas Maier  
Dipl.-Inf. Stefan Steidl  
Dipl.-Inf. Viktor Zeissler

**Lehrbeauftragte:**

Dr.-Ing. Thomas Wittenberg

**Nichtwiss. Personal:**

Walter Fentze  
Iris Koppe  
Martina Montel-Kandy  
Kristina Müller  
Friedrich Popp

# 1 Einführung

Der Lehrstuhl für Mustererkennung (LME) ist Teil des Instituts für Informatik der Friedrich-Alexander-Universität Erlangen-Nürnberg. Von 1975, dem Gründungsjahr des Lehrstuhls, bis September 2005 war Prof. Dr.-Ing. H. Niemann Lehrstuhlinhaber des LME. Im Oktober 2005 hat Prof. Dr.-Ing. J. Hornegger die Leitung des Lehrstuhls übernommen.

Das Ziel der Mustererkennung ist die Erforschung der mathematischen und technischen Aspekte der Perzeption von Umwelteindrücken durch digitale Rechensysteme. Die Umwelt wird dabei durch Sensoren erfasst - die gemessenen Werte bezeichnet man als Muster. Die automatische Transformation der gewonnenen Muster in symbolische Beschreibungen bildet den Kern der Mustererkennung. Ein Beispiel hierfür sind automatische Sprachdialogsysteme, bei denen ein Benutzer an ein System per natürlicher gesprochener Sprache Fragen stellt: Mit einem Mikrophon (Sensor) werden die Schallwellen (Umwelteintrücke) aufgenommen. Die Auswertung des Sprachsignals mit Hilfe von Methoden der Mustererkennung liefert dem System die notwendigen Informationen, um die Frage des Benutzers beantworten zu können. Die Mustererkennung befasst sich dabei mit allen Aspekten eines solchen Systems von der Akquisition der Daten, bis hin zur Repräsentation der Erkennungsergebnisse.

Die Anwendungsgebiete der Mustererkennung sind sehr breit gefächert und reichen von Industrieller Bildverarbeitung über Handschriftenerkennung, Medizinischer Bildverarbeitung, sprachverstehenden Systemen bis hin zu Problemlösungen in der Regelungstechnik. Die Forschungsaktivitäten am Lehrstuhl werden dabei in die drei Bereiche

- Medizinische Bildverarbeitung
- Rechnersehen
- Sprachverarbeitung

gegliedert, wobei der Anwendungsschwerpunkt im Bereich der Medizin liegt.

## 1.1 Sprachverarbeitung

Neben der automatischen Merkmalberechnung und der darauf aufbauenden Spracherkennung beschäftigt sich der Lehrstuhl mit den folgenden Aufgabengebieten der Spracherkennung: Erkennung und Verarbeitung von unbekanntem Wörtern, Automatische Prosodieanalyse, sowie Erstellung von automatischen Dialogsystemen. Weiterer Schwerpunkt ist seit einigen Jahren die automatische Erkennung von emotionalen Benutzerzuständen mit Hilfe akustischer und linguistischer Merkmale. Neu hinzugekommen sind die Erkennung solcher Benutzerzustände anhand physiologischer Parameter sowie die multimodale Erkennung des Aufmerksamkeitsfokus von Benutzern bei der Mensch-Maschine-Interaktion.

## 1.2 Rechnersehen

Der Bereich Rechnersehen bearbeitet die Objektverfolgung, Objekterkennung und Objektrekonstruktion aus Kameradaten. Ein zentrales, darauf aufbauendes Thema ist die aktive Sensordatenauswahl. Dabei werden die informationstheoretisch optimalen Kameraaktionen für diese

Probleme a priori geschätzt. Ein weiterer Punkt ist die plenoptische Modellierung von Objekten und die Integration dieser Daten in reale Aufnahmen mit dem Ziel der Erweiterten Realität. In der Objekterkennung werden aktuell erscheinungsbasierte, statistische Klassifikatoren mit Farb- und Kontextmodellierung untersucht.

### **1.3 Medizinische Bildverarbeitung**

Die Forschungsarbeiten im Bereich der Medizinischen Bildverarbeitung beschäftigen sich mit Fragestellungen der Bildregistrierung, Rekonstruktion, Segmentierung und Bildanalyse. Im Rahmen des SFB 539 wird ein Verfahren zur Früherkennung von Glaukomerkrankungen weiterentwickelt. Hierbei wird die Segmentierung des optischen Sehnervenkopfes ebenso untersucht wie die segmentierungsfreie Klassifikation. Erweiterte Rekonstruktionsalgorithmen zur 3D/4D Herzkonstruktion unter Verwendung von C-Arm CT werden untersucht und entwickelt. Im Rahmen des Sonderforschungsbereiches 603 besteht ein Ziel darin, Augmented Reality während minimal-invasiv durchgeführter Eingriffe in der Bauchhöhle bereitzustellen. Weiterhin werden neuartige bildgebende Verfahren sowie exakte Rekonstruktionsalgorithmen in der Computertomographie (CT) entwickelt und deren Realisierung mittels unterschiedlichen Hardwarearchitekturen untersucht. Eine weitere Problemstellung ist die Detektion und Segmentierung von Lymphknoten in Ganzkörper Magnetresonanzaufnahmen und Kantenerhaltende Rauschreduktion in der CT auf Basis von Korrelationsanalysen.

### **1.4 Forschungsschwerpunkte**

- nicht-starre Registrierung multimodaler Bilddaten
- monomodale Bildfusion zur Verlaufskontrolle bei der Tumor Therapie
- Verfahren zur Schwächungskorrektur bei der SPECT-Rekonstruktion
- Rekonstruktion bewegter Objekte bei bekannter Projektionsgeometrie
- Berechnung und Visualisierung des Blutflusses in 3D-Angiogrammen
- Segmentierung von CT-Datensätzen
- schnelle Bildverarbeitung auf Standardgrafikkarten
- Diskrete Tomographie
- Sprachsteuerung interventioneller Werkzeuge
- 3D Objekterkennung
- Objektverfolgung
- Aktive Sensordatenverarbeitung
- 3D Rekonstruktion und Kamerakalibrierung
- Plenoptische Modellierung
- Erweiterte Realität
- Autonome, mobile Systeme
- Mimik- und Gestik

- Sprachverarbeitung
- Prosodie
- Dialog
- Benutzerzustandserkennung (von Ärger über Müdigkeit bis Zögern)

## 1.5 Forschungsrelevante apparative Ausstattung

- Mobiles System MOBSY mit aktivem Stereokamera-Kopf <http://www5.informatik.uni-erlangen.de/~mobsy/>
- Drehteller und Schwenkarm zur Bildaufnahme
- Smartkom-Zelle <http://www5.informatik.uni-erlangen.de/SmartKom/page.html>
- Head-Mounted Display mit integriertem Stereokamera-System
- Pan-Tilt Einheiten

Aufgrund der engen Kooperation der Arbeitsgruppe mit den Kliniken und der Industrie besteht Zugriff auf sämtliche Modalitäten, die in der modernen Medizin heute zum Einsatz kommen. Die verfügbare Entwicklungsumgebung erlaubt die schnelle Überführung der neu entwickelten Methoden in den klinischen Test.

## 1.6 Kooperationsbeziehungen

- Charité Universitätsmedizin Berlin: Klinik und Hochschulambulanz für Radiologie und Nuklearmedizin  
<http://www.medizin.fu-berlin.de/radio/>
- Deutsches Forschungszentrum für Künstliche Intelligenz  
<http://www.dfki.de/web/index.de.html>
- Harvard University, USA: Department of Radiology at Brigham and Women's Hospital  
<http://brighamrad.harvard.edu/index.html>
- ITC-irst, Trento, Italia: Sistemi sensoriali interattivi (Interactive Sensory System Division)  
<http://ssi.itc.it/>
- LIMSI-CNRS, Orsay, France: Groupe Traitement du Langage Parlé (Spoken Language Processing Group)  
<http://www.limsi.fr/Scientifique/tlp/>
- LMU München: Institut für Phonetik und Sprachliche Kommunikation  
<http://www.phonetik.uni-muenchen.de/>
- Queen's University Belfast, UK: School of Psychology  
<http://www.psych.qub.ac.uk/>

- Stanford University, USA: Radiological Sciences Laboratory  
<http://rsl.stanford.edu/>
- Szegedi Tudományegyetem, Magyarország (University of Szeged, Hungary): Képfeldolgozás és Számítógépes Grafika Tanszék (Department of Image Processing and Computer Graphics)  
<http://www.inf.u-szeged.hu/tanszekek/kepfeldolgozasesszg/starten.xml>
- TU München: Lehrstuhl für Mensch-Maschine-Kommunikation  
<http://www.mmk.ei.tum.de/>
- Universität Bielefeld: Angewandte Informatik <http://www.techfak.uni-bielefeld.de/ags/ai/> , Neuroinformatik  
<http://www.techfak.uni-bielefeld.de/ags/ni/>
- Universität Bonn: Institut für Numerische Simulation  
<http://www.ins.uni-bonn.de/>
- Universität des Saarlandes: Lehrstuhl für Sprachsignalverarbeitung  
<http://www.lsv.uni-saarland.de/index.htm>
- Universität Jena: Lehrstuhl Digitale Bildverarbeitung  
<http://www.inf-cv.uni-jena.de/>
- Universität Koblenz-Landau: Institut für Computervisualistik  
<http://www.uni-koblenz.de/FB4/Institutes/ICV>
- Universität Mannheim: Bildverarbeitung, Mustererkennung und Computergrafik  
<http://www.cvgpr.uni-mannheim.de/>
- Universität Marburg: Diskrete Mathematik und Optimierung  
[http://www.mathematik.uni-marburg.de/forschung/arbeitsgebiete\\_mathe/diskret.php](http://www.mathematik.uni-marburg.de/forschung/arbeitsgebiete_mathe/diskret.php)
- Universitätsklinikum Erlangen:  
Nuklearmedizinische Klinik  
<http://www.nuklearmedizin.klinikum.uni-erlangen.de/> ,  
Radiologisches Institut  
<http://www.idr.med.uni-erlangen.de/> ,  
Medizinische Klinik 1  
<http://www.medizin1.klinikum.uni-erlangen.de/> und  
Medizinische Klinik 2  
<http://www.medizin2.klinikum.uni-erlangen.de/>,  
Phoniatrie und Pädaudiologische Abteilung  
<http://www.phoniatrie.klinikum.uni-erlangen.de/>
- Universität Würzburg: Abteilung für Neuroradiologie,  
<http://www.neuroradiologie.uni-wuerzburg.de/>
- University of Utah, USA: Utah Center for Advanced Imaging Research  
<http://www.ucair.med.utah.edu/>

## 1.7 Wissenschaftliche Tagungen

- Vision, Modeling, and Visualization 2005  
<http://www.vmv2005.uni-erlangen.de/>

## 1.8 Veröffentlichungsreihen

Die Veröffentlichungen des Lehrstuhls befinden sich auf der lehrstuhleigenen Homepage unter <http://www5.informatik.uni-erlangen.de/literature/German/Literature.html>

# 2 Forschungsprojekte

## 2.1 Aktive 3-D Rekonstruktion

### **Projektleitung:**

Prof. em. Dr.-Ing. Heinrich Niemann

### **Beteiligte:**

Dipl.-Math. (FH) Stefan Wenhardt

**Stichwörter:** 3-D Rekonstruktion; Aktives Sehen; Rechnersehen

**Beginn:** 1.1.2004

Die 3-D Rekonstruktion befasst sich mit der Berechnung der 3-D Geometrie eines Objekts. Diese Geometrie kann in weiteren Anwendungen wie z. B. dem bildbasierten Rendering oder der Erzeugung von erweiterter Realität (Augmented Reality) genutzt werden. Dabei werden die Informationen für die aktive Rekonstruktion aus Kamerabildern gewonnen - das Wort "aktiv" bedeutet nicht, dass es sich hierbei um aktive Sensoren, wie z. B. Lasersensoren oder strukturiertes Licht, handelt. Vielmehr werden die Kameras aktiv gesteuert, d. h. für die 3-D Rekonstruktion werden Kameraparameter gesucht, die den zu erwartenden Fehler der Rekonstruktion minimieren.

In den bisherigen Arbeiten wurde der Einfluss der Kameraparameter (Translation, Rotation und Brennweite) in Stereoaufnahmen untersucht. Grundlage dieses Ansatzes ist das Fehlermodell: Die Koordinaten der 2-D Bildpunkte, die von einem bestimmten 3-D Objektpunkt stammen, sind verrauscht. Bei der Triangulation ergibt sich durch dieses Rauschen eine Abweichung zum tatsächlichen 3-D Punkt, in Abhängigkeit der bei der Stereoaufnahme verwendeten Kameraparameter. Diese Vorgehensweise wird jedoch bei drei oder mehr Bildern sehr komplex und rechenaufwändig, da die Kameraparameter aller Kameras gleichzeitig optimiert werden müssen. Somit lässt sich dieser Ansatz nicht für eine effiziente Optimierung der Kameraparameter bei der 3-D Rekonstruktion nutzen.

Statt die Kameraparameter aller Kameras gleichzeitig zu optimieren wurde deshalb ein iterativer Algorithmus entwickelt, der die beste Parameterkonfiguration für die jeweils nächste Kamera ermittelt. Der Ansatz verwendet hierzu ein erweitertes Kalman Filter. Das Kalman Fil-

ter liefert aufgrund der bereits aufgenommenen Bilder eine probabilistische Schätzung der Lage der 3-D Punkte in Form einer mehrdimensionalen Normalverteilung. Der Erwartungswert dieser Normalverteilung ist der beste lineare Schätzwert - im Sinne des mittleren quadratischen Fehlers - für die 3-D Koordinaten der Punkte. Die Kovarianzmatrix ist ein Maß für die Unsicherheit der Schätzung.

Mittels des Kalman Filters lässt sich zudem vorhersagen, wie sich die Integration eines neuen Bildes mit bestimmten Kameraparametern auf die Kovarianzmatrix auswirken wird, ohne tatsächlich das Bild aufnehmen zu müssen. Man kann also aufgrund der bisherigen Schätzung diejenige nächste Konfiguration von Kameraparametern bestimmen, die die stärkste Reduktion der Unsicherheit bewirkt. Hat man diese gefunden, wird tatsächlich ein entsprechendes Bild aufgenommen und die Schätzung mittels Kalman Filter aktualisiert.

Der beschriebene Ansatz wurde realisiert und mit realen Daten getestet. In den Experimenten sind jedoch zusätzlich zwei weitere Restriktionen zu beachten:

- Zur Positionierung der Kameras wurde ein Roboterarm verwendet. Das Bewegungsmodell dieses Arms muss berücksichtigt werden, um nur die Positionen zu analysieren, die tatsächlich erreichbar sind.
- Verdeckungen von Objektregionen durch das Objekt selbst müssen berücksichtigt werden. Die Selbstverdeckung kann sonst ein Bild für die 3D-Rekonstruktion unbrauchbar machen. Die Modellierung von Selbstverdeckungen ist noch nicht in den Ansatz integriert und bildet einen Schwerpunkt im zukünftigen Fortgang des Projekts.

Die Kameraparameter, die mit dem erweiterten Kalman Filter errechnet wurden, stimmen bei Stereoanordnungen und einfachen Objekten mit den vorherigen Untersuchungen zu Stereoaufnahmen überein und entsprechen somit den Erwartungen. Der Algorithmus stellt somit eine Erweiterung des Ansatzes für Stereoaufnahmen dar und ermöglicht darüber hinaus die effiziente Berechnung optimaler Kameraparameter für die nächste Ansicht.

## **Publikationen**

1. Wenhardt, Stefan; Denzler, Joachim; Niemann, Heinrich: On Minimizing Errors in 3-D-Reconstruction for Stereo Camera Systems. In: Geppener, V.V.; Gurevich, I.B.; Ivanova, S.E.; Nemirko, A.P.; Niemann, Heinrich; Puzankov, D.V.; Trusova, Yu.O.; Zhuravlev, Yu.I. (Hrsg.): 7th International Conference on Pattern Recognition and Image Analysis 2004: New Information Technologies (7th International Conference on Pattern Recognition and Image Analysis 2004: New Information Technologies St. Petersburg, Russia). St. Petersburg : SPbETU, 2004, S. 562-565.



## 2.2 Detektion von Lymphknoten in Ganzkörper Magnetresonanzaufnahmen

### **Projektleitung:**

Prof. Dr.-Ing. Joachim Hornegger

**Beteiligte:** Dipl.-Inf. Florian Jäger, Prof. Dr. med. Frank Wacker,  
Dr. med. Bernd Frericks

**Stichwörter:** MRT; nichtstarre Registrierung; Segmentierung;

**Beginn:** 1.7.2005

### **Mitwirkende Institutionen:**

Charité Universitätsmedizin Berlin, Campus Benjamin Franklin, Klinik für Nuklearmedizin und Radiologie

### **Kontakt:**

Dipl.-Inf. Florian Jäger

Tel.: +49.9131.85.27894, Fax: +49.9131.303811,

E-Mail: jaeger@informatik.uni-erlangen.de

Maligne Lymphome stellen die siebt häufigste Todesursache in der westlichen Welt dar. Die Therapie der Patienten sowie die Prognose hängen entscheidend vom Ausbreitungsmuster der Erkrankung ab, was die wiederholte bildgebende Diagnostik des gesamten Körpers erfordert. Zukünftig wird vermehrt die Ganzkörper-Magnetresonanztomographie an Bedeutung gewinnen, weil damit Aufnahmen ohne Repositionierung während der Akquisition möglich sind. Allerdings umfasst ein typischer Datensatz einer solchen Ganzkörper MRT im Durchschnitt ein Volumen von 512x410x1400 Voxel. Derartige Datensätze können in der klinischen Routine ohne rechnergestützte Hilfe nicht mehr vollständig einer zeitnahen und zuverlässigen Evaluierung unterzogen werden, insbesondere wenn diese mit vorangegangenen Untersuchungen verglichen werden müssen. Das Projekt befasst sich mit der Entwicklung effizienter Methodiken zur rechnergestützten Auswertung großer medizinischer Datensätzen sowie zeitlicher Sequenzen. Durch das Hervorheben medizinisch relevanter Bereiche in den Bilddaten wird der Mediziner bei der Diagnostik unterstützt und somit eine höhere Effektivität und Kosteneffizienz im klinischen Alltag erreicht. Der Schwerpunkt liegt dabei auf der Behandlung von Lymphompatienten, wobei eine Verallgemeinerung der entwickelten Verfahren möglich sein soll.

Die Bearbeitung dieses Projekts erfordert eine sehr enge interdisziplinäre Zusammenarbeit von Informatikern und Medizinern. Die beteiligten Gruppen sind einerseits der Lehrstuhl für Mustererkennung (Informatik 5), der Friedrich-Alexander Universität Erlangen-Nürnberg, sowie die Radiologie und Nuklearmedizin der Charité, Campus Benjamin-Franklin, Berlin. Der Aufgabenbereich des Lehrstuhls bezieht sich auf die Entwicklung neuer effizienter Methodiken zur Bearbeitung von großen medizinischen Datensätzen, wobei diese auf die Anwendbarkeit im klinischen Umfeld und die Validität von den beteiligten Medizinern untersucht werden.

Strukturell kann das Projekt in zwei nahezu disjunkte Ansätze untergliedert werden: Zunächst wird die Detektion von Lymphomen in MRT Aufnahmen einer Untersuchung betrachtet. In der zweiten Phase wird dann die Lokalisation von Knoten in zeitlichen Sequenzen von MRT Aufnahmen bearbeitet.

### **Detektion von Lymphknoten in einer Studie**

Die Detektion von Lymphknoten innerhalb einer MRT Studie basiert auf der Untersuchung mehrerer Wichtungen von MRT Datensätzen. Bei den in Frage kommenden Sequenzen handelt es sich primär um solche, die bei Routineuntersuchungen verwendet werden, z.B. T1-gewichtet, T2-gewichtet, FLAIR oder TIRM Sequenzen. Bei der Auswahl spielt die benötigte Akquisitionszeit eine wichtige Rolle. Erste Experimente zeigten, dass vor allem T1-gewichtete und TIRM Aufnahmen für die Segmentierungs- und Lokalisationsalgorithmen vielversprechend sind. Um beide Datensätze vergleichen zu können werden diese in einem initialen Vorverarbeitungsschritt registriert. Hierbei wird vorausgesetzt, dass die beiden Volumina bereits nahezu perfekt zueinander ausgerichtet sind, da sich der Akquisitionszeitpunkt nur marginal unterscheidet. Trotz allem wird, um kleinere Bewegungen des Patienten auszugleichen, eine nicht-starre Registrierung der Daten vorgenommen. Da hierbei zwar Datensätze der gleichen Modalität, aber unterschiedlicher Wichtungen betrachtet werden, wird auf multi-modale Ansätze zurückgegriffen. Allerdings muss dabei die Plausibilität der Ergebnisse (z.B. die Stärke der Deformation) im Auge behalten werden, um das Problem der Detektion nicht weiter zu erschweren. Zur Lokalisation der Lymphknoten werden ausschliesslich statistische Methoden verwendet. Dies hat zwei Vorteile: Erstens liefern diese im Allgemeinen Wahrscheinlichkeiten über das Vorhandensein von Lymphknoten, was sich direkt mit dem Projektziel deckt, Zweitens sind diese oftmals generischer einsetzbar und damit die entwickelten Methodiken weitgehend von der Anwendung unabhängig. Hierbei werden verschiedene Klassen von Ansätzen betrachtet. Diese basieren einerseits auf der Clusterbildung der Datensätze durch eine Klassifikation der Voxel eines Datensatzes (z.B. mittels Fuzzy C-Means oder Markov Zufallsfelder basierter Methoden) und andererseits der Vorverarbeitung mit statistischen Methoden durch beispielsweise probabilistische Differenzbildung und probabilistische Grauwertadaption.

### **Detektion von Lymphknoten in zeitlichen Sequenzen**

Ein weiterer Schwerpunkt des Projekts ist die Detektion von Lymphomen in zeitlichen Sequenzen von Ganzkörper MRT Aufnahmen. Hier erweist sich eine automatische Vorverarbeitung für den Mediziner als sehr wünschenswert, da er andernfalls Datensätze mehrerer Zeitpunkte sichten muss, was in der Regel sehr zeitintensiv ist. Da die einzelnen Volumina zu verschiedenen Zeitpunkten akquiriert wurden, werden diese zunächst starr transformiert, so dass sie weit möglichst deckungsgleich sind. Darauffolgend wird eine nicht-starre Registrierung durchgeführt. Als Ergebnis erhält man ein Vektorfeld, welches die Deformation zwischen den Datensätzen charakterisiert, so dass diese bezüglich eines Abstandsmaßes ideal zueinander passen. Damit beschreibt dieses Deformationsfeld auch die Volumenänderung von sich entwickelnden Strukturen, wie beispielsweise Läsionen. Wachsende Strukturen sind als mathematische Quelle und schrumpfende als Senke erkennbar. Zusammen mit den Informationen über die Position von Lymphknoten, welche durch die Lokalisation in Datensätzen eines Zeitpunktes bestimmt wur-

den, werden die Veränderungen innerhalb des Deformationsfeldes zur Detektion verwendet. Um Informationen aus Differenzbildern zugänglich zu machen müssen die Datensätze ebenso nicht-starr registriert werden. Allerdings wird dabei eine weit stärkere Regularisierung des Deformationsfeldes benötigt, als im Falle der Detektion innerhalb einer Studie.

### **Präsentation der Ergebnisse**

Das Ziel des Projektes ist nicht das Treffen einer endgültigen medizinischen Aussage, sondern der Verweis auf für die Diagnose interessante Bereiche innerhalb der Datensätze um die benötigte Zeit der Sichtung zu reduzieren. Hierfür werden die Ergebnisse der Lokalisation mit Hilfe einer Wahrscheinlichkeitskarte dem Anwender zugänglich gemacht. Dieser kann die Darstellung durch die Wahl eines Konfidenzintervalls seinen Ansprüchen anpassen.

## **2.3 Erscheinungsbasierte, statistische Objekterkennung**

### **Projektleitung:**

Prof. Dr.-Ing. Heinrich Niemann

### **Beteiligte:**

Dipl.-Inf. Marcin Grzegorzek

### **Stichwörter:**

Objekterkennung; Objektlokalisierung; Objektklassifikation; statistische Modellierung; Wavelets; Farbmodellierung, Kontextmodellierung

**Laufzeit:** 1.6.1999 - 30.6.2006

**Förderer:** Deutsche Forschungsgemeinschaft

### **Kontakt:**

Dipl.-Inf. Marcin Grzegorzek

Tel.: +49.9131.85.27275, Fax: +49.9131.303811,

E-Mail: grzegorz@informatik.uni-erlangen.de

Ziel dieses Projekts, das im Rahmen des Graduiertenkollegs "Dreidimensionale Bildanalyse und -synthese" durchgeführt wird, ist die Klassifikation und Lokalisation von dreidimensionalen Objekten in Bildern. Dazu wird ein erscheinungsbasierter Ansatz eingesetzt. Dementsprechend wird kein vorhergehender Segmentierungsprozess benötigt, der geometrische Merkmale wie Kanten oder Ecken detektiert, sondern die Merkmale werden direkt aus den Intensitäten der Bildpunkte bestimmt. Dieser Ansatz verwendet 2-D lokale Merkmale, die von Grauwertbildern mit Hilfe der Wavelet Multi-Skalen-Analyse berechnet werden. Statistische Modellierung der Merkmale mit Hilfe von Dichtefunktionen macht den Ansatz robust gegenüber Beleuchtungsschwankungen und Rauschen.

Bei praktisch relevanten Anwendungen sind die Objekte oft partiell verdeckt oder können sich vor beliebigem Hintergrund befinden. Deshalb wird auch der Hintergrund explizit modelliert

und in der Erkennungsphase für jeden Merkmalsvektor entschieden, ob er zum Objekt oder zum Hintergrund gehört.

Auch mehrere Objekte können im Bild auftreten, die lokalisiert und klassifiziert werden sollen. Im Allgemeinen ist die Anzahl der Objekte, die sich im Bild befinden, unbekannt. Auf Basis eines Abbruchkriteriums wird entschieden, wann die Suche nach Objekten beendet werden soll. Ein robustes Kriterium wurde bereits experimentell bestimmt.

Um Objektklassen zu trainieren, werden mehrere Bilder von Objekten in bekannten Lagen und Orientierungen gebraucht. Üblicherweise werden die Objekte auf einem Drehteller gestellt, um einen bekannten Winkel gedreht und von einer Kamera, die auf einem Roboterarm montiert ist, aufgenommen. Bei größeren oder sich bewegenden Objekten ist der Einsatz vom Drehteller nicht möglich. In diesem Falle wird der Aufnahmeprozess mit einer handgeführten Kamera durchgeführt. Die unbekannt Lagen und Orientierungen von den Objekten, die für das Training benötigt werden, werden mit Hilfe eines "Struktur aus Bewegung" Algorithmus rekonstruiert.

Zur Evaluierung des Objekterkennungssystems entstand eine neue umfangreiche Stichprobe 3D-REAL-ENV. Bei über 30000 Trainingsbildern und 8000 Testbildern mit einem realen heterogenen Hintergrund von 10 Objekten kann man verschiedene Erkennungsalgorithmen objektiv vergleichen. Da die Beleuchtung in den Testbildern anders als in der Trainingsphase ist, konnte auch die Beleuchtungsunabhängigkeit des Systems nachgewiesen werden.

In der letzten Zeit wurde die Objektfarbe bei der Berechnung der Merkmale berücksichtigt. 6-D lokale Merkmalsvektoren werden von Farbbildern berechnet, wobei die Wavelet Multi-Skalen-Analyse separat für den roten, grünen, und blauen Kanal durchgeführt wird. Bei Testbildern mit heterogenem Hintergrund stieg die Klassifikationsrate von 55.4% (Grauwertmodellierung) auf 87.3% (Farbmodellierung). Die Lokalisationsrate verbesserte sich von 69.0% (Grauwertmodellierung) auf 77.1% (Farbmodellierung)

## **Publikationen**

1. Grzegorzec, Marcin; Niemann, Heinrich: Statistical Object Recognition Including Color Modeling. In: Kamel, Mohamed; Campilho, Aurelio (Hrsg.): Proceedings of the 2nd International Conference on Image Analysis and Recognition (2nd International Conference on Image Analysis and Recognition Toronto 28-30.09.2005). Berlin : Springer, 2005, S. 481-489. - ISBN 10 3-540-29069-9.
2. Reinhold, Michael; Grzegorzec, Marcin; Denzler, Joachim; Niemann, Heinrich: Appearance-Based Recognition of 3-D Objects by Cluttered Background and Occlusions. In: Pattern Recognition 38 (2005), Nr. 5, S. 739-753.
3. Grzegorzec, Marcin; Scholz, Ingo; Reinhold, Michael; Niemann, Heinrich: Fast Training for Object Recognition with Structure-from-Motion. In: Pattern Recognition and Image Analysis: Advanced in Mathematical Theory and Applications 15 (2005), Nr. 1, S. 183-186.

## 2.4 Exakte Bildrekonstruktion aus Kegelstrahlprojektionen für Spezialtrajektorien

### Projektleitung:

Prof. Dr.-Ing. Joachim Hornegger, Dr. Günter Lauritsch (Siemens Med. Solutions),  
Prof. Dr. Frederic Noo (UCAIR)

**Beteiligte:** Dipl.-Ing. (FH) Stefan Hoppe, Dipl.-Inf. Frank Dennerlein

**Stichwörter:** exakte Bildrekonstruktion; Kegelstrahl-Computertomographie; Kreis und Kreisbogen; Kreis und Linie; geometrische Kalibrierung

**Beginn:** 1.3.2005

**Förderer:** Siemens Medical Solutions

**Mitwirkende Institutionen:** University of Utah, Utah Center for Advanced Imaging Research (UCAIR)

### Kontakt:

Dipl.-Ing. (FH) Stefan Hoppe  
Tel.: +49.9131.85.28977, Fax: +49.9131.85.28990,  
E-Mail: hoppe@informatik.uni-erlangen.de

Im Jahr 2005 wurde ein weiteres Projekt auf dem Gebiet der dreidimensionalen Rekonstruktion aus zweidimensionalen Bildfolgen auf den Weg gebracht. Das Projekt beschäftigt sich mit der Rekonstruktion eines Volumens aus einer Sequenz von Röntgenbildern. Ziel des Forschungsprojekts ist die Entwicklung, Verbesserung und Evaluierung von Verfahren zur exakten Bildrekonstruktion aus Kegelstrahlprojektionen für Spezialtrajektorien. Neben der Auswahl und Untersuchung neuartiger Abtastbahnen steht die Entwicklung von Strategien zur korrekten Handhabung realer Daten und zur Minimierung von Rekonstruktionsartefakten im Vordergrund. Hierfür wurden zwei Verfahren für die geometrische Kalibrierung einer C-Bogen Anlage für eine Abtastbahn bestehend aus zwei Teilkreisen entworfen, implementiert und an realen Daten getestet. In enger Zusammenarbeit mit der Forschungseinrichtung UCAIR der University of Utah wurde ein Algorithmus zur Rekonstruktion mit einer Abtastbahn bestehend aus einem Teilkreis und einer Linie implementiert. Die Implementierung erfolgte im Rahmen eines zweimonatigen Forschungsaufenthalts von Frank Dennerlein und Stefan Hoppe an der University of Utah und wurde dort von Professor Frederic Noo betreut. Ergebnisse mit simulierten Daten zeigen die Überlegenheit des Verfahrens gegenüber einem Feldkamp Algorithmus hinsichtlich entstehender Kegelstrahlartefakte. Im Anschluss wurde das Verfahren auf die Abtastbahn bestehend aus zwei Teilkreisen übertragen und zunächst an simulierten Daten getestet. Im Dezember 2005 konnte ein erster Prototyp fertiggestellt werden, der mit Hilfe der geometrischen Informationen aus dem Kalibrierungsschritt in der Lage ist reale Daten zu verarbeiten. Hierfür wurde eine Reihe realer Projektionsdaten verschiedener Phantomkörper mit einem Siemens AXIOM Artis dBA C-Bogen System akquiriert und Referenzrekonstruktionen mit

einem aktuellen Feldkamp-basierten Rekonstruktionsalgorithmus durchgeführt. Das Projekt wird von unserem industriellen Partner Siemens Medical Solutions finanziell unterstützt. Darüber hinaus stellt Siemens Medical Solutions den Zugang zu den neuesten Röntgenanlagen zur Verfügung und ermöglicht so die Entwicklung unter realen Bedingungen. Das Ziel ist der Einsatz des Verfahrens im klinischen Umfeld.

## **2.5 HUMAINE (Human-Machine-Interaction Network on Emotion)**

### **Projektleitung:**

Dr.-Ing. Elmar Nöth

### **Beteiligte:**

Dr. phil. Anton Batliner, Dipl.-Inf. Stefan Steidl, Dipl.-Inf. Christian Hacker,  
Dipl.-Inf. Florian Hönig

### **Stichwörter:**

Emotionserkennung; Mensch-Maschine-Interaktion

**Laufzeit:** 1.1.2004 - 31.12.2007

**Förderer:** EU

### **Mitwirkende Institutionen:**

<http://emotion-research.net/Members/>

### **Kontakt:**

Dr.-Ing. Elmar Nöth

Tel.: +49.9131.85.27888, Fax: +49.9131.303811,

E-Mail: [noeth@informatik.uni-erlangen.de](mailto:noeth@informatik.uni-erlangen.de)

HUMAINE (Human-Machine-Interaction Network on Emotion) ist ein NoE (Network of Excellence) im europäischen Rahmenprogramm FP6 und läuft vom 1.1.2004 bis zum 31.12.2007; Vertrags-Nr. 507422. HUMAINE will in der Europäischen Union die Grundlagen für die Entwicklung von Systemen schaffen, die menschliche emotionale oder emotions-ähnliche Zustände und Prozesse erkennen, modellieren und/oder beeinflussen können - sog. "emotionale Systeme". Solche Systeme sind grundlegend für die zukünftige Mensch-Maschine-Kommunikation; allerdings sind ihre Grundlagen noch nicht genügend erforscht. Einer der Gründe dafür ist sicherlich, dass das einschlägige Wissen auf unterschiedliche Disziplinen verteilt ist (Psychologie, Künstliche Intelligenz, Robotik, Spracherkennung, Bildverarbeitung, usw.). In HUMAINE kooperieren Experten aus all diesen Disziplinen. Sechs unterschiedliche thematische Gebiete werden bearbeitet: Emotionstheorie, Signale und Zeichen, emotionale Interaktion, Emotion in Kognition und Aktion, Emotion in Kommunikation und Überzeugung sowie Benutzerfreundlichkeit emotionaler Systeme. Gemeinsame Arbeitsprojekte werden auf Workshops sowie in Plenartreffen abgesprochen; darunter fällt auch die notwendige Infrastruk-

tur (Korpora, die kulturelle und geschlechtsspezifische Unterschiede widerspiegeln, ein ethisches Rahmenwerk sowie ein Web-Portal). Ein Ziel ist das Verfassen eines Handbuchs sowie Empfehlungen zur Methodik.

Die FAU Erlangen-Nürnberg ist Partner im Teilprojekt WP4 (Signale und Zeichen) sowie in WP5 (Korpora). Im ersten Jahr konzentrierten sich die Aktivitäten auf das initiale Plenartreffen, eine Sommerschule über Korpora sowie Workshops über Theorie, Signale und Zeichen sowie Korpora. Gegen Ende 2004 wurden sog. "Exemplars" genauer definiert und beschrieben. Darunter ist ein beispielhaftes Vorgehen zu verstehen, mit dem gezeigt wird, auf welche Weise ein Problem gelöst werden kann, und wie eine interdisziplinäre Herangehensweise auszusehen hat. Die spezielle Form eines solchen Exemplars ist nicht vorgegeben; es kann sich dabei um einen Demonstrator handeln, um ein Korpus, um ein experimentelles Design oder um spezifische Methoden. Der Beitrag unseres Instituts zu diesen "Exemplars" bestand 2005 in einer so genannten "forced co-operation" Initiative CEICES (Combining Efforts for Improving automatic Classification of Emotional user States): Wir stellen ein annotiertes emotionales Korpus zur Verfügung (Sprachdaten, phonetisches Lexikon, handkorrigierte Segmentierung, Emotionslabels) und definieren Training- und Test-Stichprobe. Alle teilnehmenden HUMAINE-Partner stellen allen anderen Partnern ihre eigenen extrahierten Merkmale zur Verfügung. Ziel ist es, mit dieser Kooperation eine Verbesserung der Erkennungsraten sowie einen Einblick in die Relevanz unterschiedlicher Merkmale und Merkmalstypen zu erreichen.

## Publikationen

1. Batliner, Anton; Hacker, Christian; Steidl, Stefan; Nöth, Elmar; D'Arcy, S.; Russell, M.; Wong, M.: "You stupid tin box" - children interacting with the AIBO robot: A cross-linguistic emotional speech corpus. In: ELRA (Hrsg.): Proceedings of the 4th International Conference of Language Resources and Evaluation LREC 2004 (LREC Lisbon). 2004, S. 171-174.
2. Steidl, Stefan; Levit, Michael; Batliner, Anton; Nöth, Elmar; Niemann, Heinrich: "Of All Things the Measure is Man" - Classification of Emotions and Inter-Labeler Consistency. In: IEEE (Veranst.): Proceedings of ICASSP 2005 - International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005 - International Conference on Acoustics, Speech, and Signal Processing Philadelphia, Pennsylvania, U.S.A. March 18-23, 2005). Bd. 1. 3833 S. Texas Ave., Ste. 221 Bryan, TX 77802-4015 : Conference Management Services, Inc., 2005, S. 317-320. - ISBN 0-7803-8874-7.
3. Batliner, Anton; Steidl, Stefan; Hacker, Christian; Nöth, Elmar; Niemann, Heinrich: Tales of Tuning - Prototyping for Automatic Classification of Emotional User States. In: ISCA (Hrsg.): Proceedings of the 9th European Conference on Speech Communication and Technology (9th European Conference on Speech Communication and Technology - Interspeech 2005 Lisbon 4-8.9.2005). Bonn : ISCA, 2005, S. 489-492. - ISBN 1018-4074.
4. Batliner, Anton; Steidl, Stefan; Hacker, Christian; Nöth, Elmar: Private Emotions vs. Social Interaction - towards New Dimensions in Research on Emotion. In: Carberry, Sandra; de Rosis, Fiorella (Hrsg.): Adapting the Interaction Style to Affective Factors (Workshop on Adapting the Interaction Style to Affective Factors, 10th International Conference on User Modelling Edinburgh 25.7.2005). 2005, (8 pages, no pagination).

## 2.6 Koronarangiographie unter Verwendung von C-Arm CT

**Projektleitung:**

Prof. Dr.-Ing. Joachim Hornegger

**Beteiligte: Dipl.-Inf. Marcus Prümmer**

**Stichwörter:** nicht starre Registrierung; Koronarangiographie

**Laufzeit:** 1.10.2003 - 30.9.2006

**Kontakt:**

Dipl.-Inf. Marcus Prümmer

Tel.: +49.9131.85.27826, Fax: +49.9131.303811,

E-Mail: pruemmer@informatik.uni-erlangen.de

Unter Koronarangiographie versteht man die Durchleuchtung von Herzkranzgefäßen oder Herzkammern mittels Röntgenstrahlung. Häufig wird ein Angiogramm des Herzens aufgenommen, um zu entscheiden, ob eine weitere Behandlung notwendig ist. Beispielsweise kann eine Gefäßverengung dazu führen, dass einzelne Herzkranzgefäße nur wenig oder nicht mehr mit Blut versorgt werden. Während der Behandlung wird ein Katheder (dünner flexibler Schlauch) in eine Arterie am Arm oder in der Leiste eingesetzt und bis zur erkrankten Gefäßposition durchgeschoben. Blutgefäße des Herzens werden dann durch Einspritzen von Kontrastmittel durch den Katheder besser sichtbar gemacht. Eine Folge von Niedrigkontrastaufnahmen (Fluoroscopes) wird zur Visualisierung ständig während des minimal-invasiven Eingriffes aufgenommen. Eine 3-D Visualisierung der Herzkammern oder der Kranzgefäße wäre dabei wünschenswert. Dazu bedarf es einer 3-D Rekonstruktion unter Verwendung von Röntgenaufnahmen des sich bewegenden Herzens.

Die Vision dieses Projektes ist es, eine rein datengetriebene 3-D Rekonstruktion aus Aufnahmen herkömmlicher C-Arm Systeme zu ermöglichen. Mit C-Arm CT Systemen können mehrere EKG getriggerte Angiogramme aufgenommen werden. Dies erlaubt eine Selektion der aufgenommenen Röntgenbilder (für jeden Projektionswinkel), die jeweils das schlagende Herz möglichst nahe zu einer festgelegten Herzphase darstellen. Herzschlagvariabilität kann jedoch zu einer weniger optimalen Triggerung multipler C-Arm CT Durchläufe führen, was zu einer Erniedrigung der zeitlichen Auflösung führen kann, was wiederum starke Bewegungsartefakte in der 3D Rekonstruktion hervorrufen kann.

Deshalb müssen zusätzlich Bewegungskorrekturen in den aufgenommenen Röntgenaufnahmen und/oder 3D Rekonstruktionen vorgenommen werden. Die folgenden im Jahr 2005 entwickelten Teilprojekte beschäftigen sich mit verschiedenen Ansätzen zur Bewegungskorrektur um die Bildqualität für Koronarangiographie für C-Arm CT Systeme zu verbessern. Zur Erhöhung der zeitlichen Auflösung einer 3D Koronarangiographie wurden verschiedene nicht-starre Registrierungsverfahren entwickelt und getestet.



## **Sinogramm Interpolation**

In diesem Ansatz wird ein Variationsansatz verwendet, um fehlende Sinogrammdaten zu interpolieren. Fehlende Sinogrammdaten können entstehen, wenn für einzelne Projektionswinkel keine Aufnahme existiert, die das Herz in der gewünschten Herzphase zeigt. Die Idee zur Interpolation kommt vom sogenannten Video-Impainting, wo mehrere Schlüsselbilder einer Bildsequenz gegeben sind und dazwischenliegende Bilder interpoliert werden. Mittels Diffusion wird zwischen beobachteten Röntgenbildern die fehlende Sinogramminformation interpoliert. Da hohe Frequenzen entlang einer Detektorzeile mit Objektstrukturen (z.B. Knochen, Blutgefäße) korrespondieren, wird nur entlang einer sichtbaren sinnogrammförmigen Kontourlinie geglättet (diffundiert) um somit Objektstrukturen zu erhalten. Da wichtige Objektinformation im gesamten 2-D bzw. 3-D Sinogramm enthalten ist, wird vorzugsweise der komplette Sinogramminhalt berücksichtigt.

## **2-D - 2-D nicht-starre Registrierung von Angiogrammen**

Mittels nichtstarrer 2-D - 2-D Registrierung wird versucht die projektive Herz- und Sinogrammbewegung zu berechnen um fehlende Sinogrammdaten für einzelne Projektionswinkel zwischen aufgenommenen Projektionsdaten desselben Winkels zu interpolieren. Dazu wird ein dichtes Deformationsfeld zwischen einzelnen Projektionsbildern berechnet.

## **3-D - 3-D nicht-starre Registrierung**

Da die Herzbewegung nicht mittels einer affinen Abbildung beschrieben werden kann, wird die Bewegung zwischen 3-D rekonstruierten Bildern mittels nicht-starrer 3-D - 3-D Registrierung berechnet. Für die Registrierung werden State-of-the-Art Algorithmen verwendet, wobei der Hauptfokus auf der Bildqualität des rekonstruierten Bildes unter Anwendung des berechneten Deformationsfeldes liegt.

## **Multi-modale 2-D - 3-D nicht-starre Registrierung**

Die Idee dieses Ansatzes beruht auf der nicht-starren Registrierung eines initial rekonstruierten 3D Bildes einer definierten Herzphase mit einer Sequenz aus aufgenommenen Röntgenbildern einer anderen Herzphase. Das berechnete Deformationsfeld des Herzens wird bei der Rekonstruktion berücksichtigt. Hierzu wurde speziell ein multi-modaler 2D-3D nicht-starrer Registrierungsalgorithmus entwickelt und bisher an dynamischen Herzphantomdaten evaluiert.

## **Publikationen**

1. Prümmer, Marcus; Han, Jingfeng; Hornegger, Joachim: 2D-3D Non-rigid Registration using Iterative Reconstruction. In: Greiner, Günther; Hornegger, Joachim; Niemann, Heinrich; Stamminger, Marc (Hrsg.): Vision Modeling and Visualization (Workshop Vision Modeling and Visualization in Erlangen 2005 Erlangen 16.-18. November 2005). Erlangen : Akademische Verlagsgesellschaft Aka GmbH, Berlin, 2005, S. 187-194. - ISBN 3-89838-068-8.

2. Prümmer, Marcus; Köstler, Harald; Hornegger, Joachim; Rude, Ulrich: A full multigrid technique to accelerate an ART scheme for tomographic image reconstruction. In: Hülsemann, Frank; Kowarschik, Markus; Rude, Ulrich (Hrsg.): *Frontiers in Simulation (Simulationstechnique 18th Symposium in Erlangen 2005 Erlangen 12.-15. September 2005)*. Erlangen : SCS Publishing House e.V., 2005, S. 632-637. - ISBN 3-936150-41-9.

## 2.7 Lautbildungsstörungen bei Lippen-Kiefer-Gaumenspalte

### **Projektleitung:**

PD Dr. med Maria Schuster, PD Dr. med. Dr. med. dent. Emeka Nkenke,  
Dr.-Ing. Elmar Nöth

**Beteiligte:** Dipl.-Inf. Andreas Maier, Dipl.-Inf. Tino Haderlein,  
Dipl. Med.-Inf. Jochen Penne

**Stichwörter:** Phoniatrie und Pädaudiologie; Mund-Kiefer-Gesichtschirurgie;  
medizinische Bildverarbeitung; automatische Sprachanalyse

**Laufzeit:** 1.1.2005 - 1.8.2006

**Förderer:** Johannes-und-Frieda-Marohn-Stiftung

### **Kontakt:**

Dipl.-Inf. Andreas Maier  
Tel.: +49.9131.85.27825,  
E-Mail: andreas.maier@informatik.uni-erlangen.de

Lippen-Kiefer-Gaumenspalten stellen eine der häufigsten Fehlbildungen im Gesichtsbereich dar. Auch nach erfolgreicher operativer Behandlung bleiben morphologische Veränderungen der oberen Atmungs- und Sprechorgane zurück und führen zu Lautbildungsstörungen. Die Diagnostik der Lautbildungsstörungen erfolgt derzeit lediglich subjektiv oder erfasst nur Teilbereiche. Eine neue und für den Patienten nicht belastende Möglichkeit der objektiven und umfassenden Diagnostik der verschiedenen Lautbildungsstörungen stellt der Einsatz eines automatischen Spracherkennungssystems dar, welches sich bereits bei der Diagnostik anderer Störungen der lautsprachlichen Kommunikation bewährt hat. Hiermit sollen Lautbildungsstörungen erkannt, unterschieden und quantifiziert werden. Morphologische Besonderheiten des Nasenresonanzraumes, die bei dieser Patientengruppe Einfluss auf die Nasalität und damit die Sprachverständlichkeit haben, sollen mittels einer dreidimensionalen Rekonstruktion anhand von Endoskopiebildern vollständig und quantitativ erfasst werden. Damit lassen sich kostenintensive und belastende radiologische Untersuchungen zur Verlaufskontrolle oder zur Planung von operativen Eingriffen vermeiden. Diese Analysen von Funktion und Morphologie sollen die Grundlage für einen Ansatz zu Klärung der ursächlichen Zusammenhänge der Lautbildungsstörungen liefern.

## 2.8 Registrierung von funktionellen (CT, MRT) und morphologischen (PET, SPECT) Datensätzen

### Projektleitung:

Prof. Dr.-Ing. Joachim Hornegger, PD Dr. med. Wolfgang Römer,  
Prof. Dr. med. Torsten Kuwert, Prof. Dr. med. Werner Bautz

**Beteiligte:** Jingfeng Han, M. Sc.

**Stichwörter:** Multimodale Registrierung

**Laufzeit:** 1.11.2004 - 31.12.2006

Die Registrierung von funktionellen (CT, MRT) und morphologischen Daten ist immer noch eine der großen Herausforderungen in der medizinischen Bildverarbeitung. Die Vereinigung der komplementären Daten aus zwei unterschiedlichen Modalitäten liefert wertvolle Informationen für Ärzte in der Tumordiagnose. Die grundlegende Aufgabe dieses Projektes ist die Entwicklung von neuartigen und robusten Registrierungsmethoden und ihre effiziente Implementierung in automatischen Registrierungssystemen. Zusätzlich müssen diese natürlich auch ausreichend in einem klinischen Umfeld evaluiert werden. Im Folgenden wird der Fortschritt dieses Projektes sowie seine Motivationen, die eingesetzten Methoden und die erzielten Resultate kurz zusammengefasst.

### Die starre Registrierung

beinhaltet ein Volumen bezüglich eines anderen so zu drehen, und zu verschieben, dass eine möglichst gute Registrierung erzielt wird. Die daraus resultierende globale Transformation verändert dabei nichts an der zugrundeliegenden Anatomie. Die Anwendung der starren Registrierungstechniken zur Registrierung multimodaler Bilddaten hat im klinischen Umfeld eine weite Verbreitung erfahren. Auch in der Registrierung von PET/CT oder SPECT/CT Daten die eine nicht-starre Registrierung erfordert, ist üblicherweise ein starrer Vorregistrierungsschritt notwendig.

In vielen Fällen sind die Bilddaten aus einer Modalität nur Subvolumen/-bilder der Daten der anderen Modalität, und entsprechen folglich nur einem Teil dieser Daten, zum Beispiel einem einzelnen Organ. Bereits existierende gradientenbasierte starre Registrierungsverfahren sind üblicherweise nur in der Lage vollständige Datensätze miteinander zu registrieren. Um sie erfolgreich für Subvolumenregistrierung anwenden zu können ist deshalb eine grobe manuelle Vorregistrierung notwendig. Um diese aufwendige Benutzerinteraktion zu eliminieren wurde ein stochastischer Algorithmus entwickelt der die optimale Transformation selbst bei einer schlechten Initialisierung finden kann. Mit den entsprechenden Ähnlichkeitsmaßen lässt sich dieser Ansatz sowohl auf mono- also auch auf multimodale Daten anwenden.

Weiterhin wurde eine starre Registrierungsmethode basierend auf auffallenden Bildmerkmalen entwickelt und implementiert. Eine solche auffallende Bildregion bietet zwei Arten von Informationen: Sie beschreibt die Einzigartigkeit des zugrundeliegenden Bildinhaltes durch ihre Helligkeitswerte und sie enthält geometrische Informationen über den Mittelpunkt und die Aus-

maße der Region. In diesem Projekt haben wir die Verwendung solcher auffallenden Bildregionen in ein hybrides Registrierungsframework für starre Körper integriert. Die Methode besteht aus drei Schritten: Eine automatische Extraktion von auffälligen Bildmerkmalen, eine robuste Schätzung von Korrespondenzen zwischen diesen Regionen und eine Subpixelverfeinerung die auch einen Schritt zum Entfernen von Ausreißern beinhaltet. Zusätzlich wurde auch die Robustheit des Algorithmus in Bezug auf Bildüberlappungen, fehlende Information und Bildartefakte gezeigt.

### **Die nicht-starre Registrierung**

erlaubt die Deformation einer anatomischen Struktur so, dass sie mit einer entsprechenden Struktur in einem weiteren Datensatz übereinstimmt. Dieser Prozess ist dafür gedacht einen Vergleich zwischen Anatomien zu ermöglichen, die zu unterschiedlichen Zeitpunkten, von unterschiedlichen Patienten oder mit unterschiedlichen Modalitäten aufgenommen wurden. Für dieses Teilprojekt sind die entsprechenden nicht-starren, multimodalen Registrierungsmethoden wichtig für die Fusion von funktionellen und morphologischen Bilddaten.

Ein wichtiges Kriterium für die Registrierung von funktionellen und morphologischen Modalitäten ist die korrekte Zuordnung von besonderen Merkmalen, die sich meist durch deutliche Kanten auszeichnen. Die manuelle Annotierung solcher zueinander korrespondierenden Kanten in multimodalen Bilddaten ist jedoch sehr aufwendig und ungenau. Deshalb wurde versucht die Kantendetektion und die nicht-starre Registrierung gleichzeitig durchzuführen. Dazu wurden beide Vorgänge in eine gemeinsame, mathematische Formulierung basierend auf der Ambrosio-Tortorelli Approximation des Mumford-Shah Modells zur Segmentierung eingebettet. Um zu garantieren, dass die errechnete Transformation bijektiv ist wird nicht eine, sondern parallel zwei entgegengesetzte Transformationen berechnet. Dabei sorgt ein Konsistenzfunktional das die Transformationen einschränkt dafür, dass beide die inverse Transformation der jeweils anderen sind. Die Optimierung wird mittels eines Gradientenfluss Algorithmus durchgeführt der für glatte Transformationen sorgt. Eine multilevel Implementierung sichert schließlich eine gute Effizienz des Algorithmus. Vorläufige medizinische Auswertungen auf T1 und T2 MRT Daten lieferten ermutigende Resultate.

Die Multimodale Bildfusion muss bestimmte Bedingungen für die Transformation die sie errechnet aufstellen, um zu vermeiden, dass degenerierte Abbildungen auftreten. Wir haben dieses Problem der Notwendigkeit von Nebenbedingungen durch die Einbringung von medizinischem Wissen behandelt: Anatomische Fixpunkte, wie die Basis der Blase oder die Konturen der Nieren und der Leber können von einem Arzt in beiden Datensätzen identifiziert und markiert werden. Unser Algorithmus zur Bildregistrierung verwendet dann ein intensitätsbasiertes Abstandsmaß und errechnet die entsprechende Abbildung mit einem Variationsansatz in den sich die erwähnten Merkmalskorrespondenzen als Randbedingungen einbringen lassen. Um diesen Ansatz im klinischen Umfeld zu evaluieren wurden CT und PET Daten von 42 Melanompatienten manuell mit einem kommerziellen Programm registriert. Die Resultate dieser manuellen Registrierung wurden dann mit den Ergebnissen unserer Algorithmen verglichen. Die neue Registrierungsmethode erzielte dabei gute Fusionsergebnisse Die erzielte Bildqualität

war bei einer visuellen Beurteilung einem existierenden, kommerziellen Fusionspaket überlegen.

Dank der Installation des SPECT/Spiral CT hybrid Scanners in der nuklearmedizinischen Klinik hat dieses Projekt Zugriff auf eine riesige Menge registrierter SEPCT/CT Volumendatensätze. Diese nahezu registrierten Volumendatensätze werden als Trainingsdaten zur Erstellung einer Abbildung von SPECT auf CT Volumen genutzt. Es wurde eine lernfähige Methode entwickelt die diese Trainingsdaten nutzt um die softwarebasierte Bildregistrierung zu verbessern. Hierzu werden Trainingsdatensätzen benutzt um eine a priori Wahrscheinlichkeit für die Korrespondenzen Intensitätswerten in der nicht-starren Registrierung zu bestimmen. In den von uns durchgeführten Experimenten war unsere lernfähige Methode in der Lage die lokale Fehlregistrierung einiger anatomischer Strukturen in der Registrierung von SPECT/CT Ganzkörperdatensätzen zu beheben.

### Publikationen

1. Han, Jingfeng; Hornegger, Joachim; Kuwert, Torsten; Bautz, Werner; Römer, Wolfgang: Feature Constrained Non-rigid Image Registration. In: Hülsemann, Frank; Kowarschik, Markus; Råde, Ulrich (Hrsg.): *Frontiers in Simulation (18th Symposium on Simulationstechnique, Erlangen 2005 Erlangen 12.-15. September 2005)*. Erlangen : SCS Publishing House e.V., 2005, S. 638-643. - ISBN 3-936150-41-9.
2. Römer, Wolfgang; Hornegger, Joachim; Han, Jingfeng; Bautz, Werner; Kuwert, Torsten: Non-rigid Fusion of Morphological and Functional Images Using Anatomical Fix Points and Contours - A New Approach to Overcome the Current Drawbacks of Retrospective Image Fusion. In: *The Radiological Society of North America (Veranst.): RSNA2004 (RSNA2004)*. Chicago, USA : RSNA, 2005, S. 356.
3. Han, Jingfeng; Qiao, Min; Hornegger, Joachim; Kuwert, Torsten; Bautz, Werner; Römer, Wolfgang: Partial Image Data Registration using Stochastic Optimization. In: Hülsemann, Frank; Kowarschik, Markus; Råde, Ulrich (Hrsg.): *Frontiers in Simulation (18th Symposium on Simulationstechnique, Erlangen 2005 Erlangen 12.-15. September 2005)*. Erlangen : SCS Publishing House e.V., 2005, S. 644-645. - ISBN 3-936150-41-9.

## 2.9 Schlaganfallrisikoindexberechnung

**Projektleitung:**

Prof. em. Dr.-Ing. Heinrich Niemann

**Beteiligte:** Ing. Radim Chrastek, Dipl.-Inf. Rüdiger Bock

**Beginn:** 1.1.1998

**Mitwirkende Institutionen:** Augenklinik mit Poliklinik

**Kontakt:**

Dipl.-Inf. Rüdiger Bock

E-Mail: ruediger.bock@informatik.uni-erlangen.de

Das Ziel dieses Projektes ist die Erkennung eines nahenden Schlaganfalles. Ein Forschungsteam der Universitäten Wisconsin-Madison und North Carolina hat belegt, dass eine Verengung der Arterien des Augenhintergrundnetzes mit einem bevorstehenden Hirninfarkt verbunden ist. Als Maßzahl für arterielle Verengung in der Retina und gleichzeitig als Gefährdungsindex wurde von unseren medizinischen Partnern (Augenklinik Erlangen) vorgeschlagen, die gewichteten durchschnittlichen Durchmesser der Arterien durch die gewichteten durchschnittlichen Durchmesser der Venen zu dividieren. Im Gegensatz zu dem US-Forschungsteam, das den Gefährdungsindex halbautomatisch ermittelt, soll der Gefährdungsindex im Rahmen dieses Projektes automatisch berechnet werden.

Der Algorithmus besteht aus 6 Modulen: Einlesen und Vorverarbeiten des Bildes, Papillen- und Makulasegmentierung (Papille bezeichnet die Stelle, an der der Sehnerv das Auge verlässt; die Makula bezeichnet die Stelle des schärfsten Sehens), Segmentierung des Gefäßbaumes, Klassifikation der Gefäße in Arterien und Venen, Vermessung der Gefäßdurchmesser und Berechnung des Gefährdungsindex. Nachdem das Netzhautfoto eingelesen wird, werden die Gefäße durch die sog. Isolinie (in Analogie zur Geographie) grob segmentiert. Sie werden dann für die Papillensegmentierung verwendet. Die Papillen- und Makulasegmentierung dient zur Bestimmung der Vermessungszonen und zur Ausblendung der Bereiche, in denen die Gefäße nicht segmentiert werden sollen. Die Papille wird aufgrund der Helligkeitsunterschiede und der Konvergenz des Gefäßbaumes ermittelt. Die Makula wird mit Hilfe des anatomischen Wissens ermittelt. Sie liegt 2 bis 3 Papillenträger temporal (d.h. Richtung Schläfen) von der Papillengrenze. Dann wird der gesamte Gefäßbaum segmentiert. Die Segmentierung basiert auf Isoliniendetektion, Parallelverlauf des Gefäßrandes, Kontrasten und anatomischen Eigenschaften (z.B.: Arterien kreuzen nie andere Arterien; Venen kreuzen nie andere Venen). Für die Klassifikation der Gefäße in Arterien und Venen wird der Kontrast im roten Kanal und die Information aus dem Gefäßbaum genutzt. Falls sich 2 Gefäße kreuzen, ist das dunklere Gefäß fast immer die Vene. Für die Berechnung des Gefährdungsindex werden die 6 größten Arterien und die 6 größten Venen herangezogen. Die Segmentierung des gesamten Gefäßbaumes wird demnächst abgeschlossen werden. Die Entwicklung der anderen Module ist abgeschlossen.

## 2.10 Segmentierung und Multi-Modale 3D Registrierung mit CT, SPECT und Ultraschall

### **Projektleitung:**

Prof. Dr.-Ing. Joachim Hornegger

### **Beteiligte:**

Dipl.-Inf. Dieter Hahn

**Stichwörter:** Segmentierung, Registrierung

**Laufzeit:** 1.1.2005 - 31.12.2005

**Mitwirkende Institutionen:** Nuklearmedizinische Klinik mit Poliklinik  
Professur für Pränataldiagnostik und Perinatalmedizin

### **Kontakt:**

Dipl.-Inf. Dieter Hahn

Tel.: +49.9131.85.27894, Fax: +49.9131.303811,

E-Mail: dieter.hahn@informatik.uni-erlangen.de

Als Aufgabenstellungen im Jahr 2005 haben sich unter anderem die Segmentierung und Registrierung multi-modaler Volumendaten ergeben. Im Rahmen des Segmentierungsprojekts lautet die Zielstellung, Schilddrüsendaten in 3D Ultraschallbildern (US) zu klassifizieren. Wegen der schlechten Bildqualität, bedingt durch verrauschte Bildsignale, ist dieser Ansatz ohne eine adäquate Vorverarbeitung nicht zu bewältigen. Daher konzentrierten sich die anfänglichen Arbeiten darauf, verschiedene rauschunterdrückende Filteroperationen zu finden, um einerseits die Bildqualität zu verbessern und andererseits die vorhandene Information nicht zu verlieren. Beim Ansatz der Segmentierung mit Level Sets kommt es dabei vor allem auf die Kantenerhaltung an. Eine quantitative Analyse geeigneter Filter ließ hierbei einen klaren Vorteil auf Seiten des anisotropen, kantenverstärkenden Diffusionsfilters erkennen. Auf Basis der vorverarbeiteten Daten wird schließlich versucht, eine geeignete Segmentierungsmethode für die möglichst vollautomatische Klassifikation der Schilddrüse zu implementieren. Die hiermit zusätzlich zu den Intensitäten gewonnene Information wird im zweiten Projekt eingesetzt: der multi-modalen Registrierung von CT (Computed Tomography), SPECT (Single Photon Emission Computed Tomography) und US. Die 3D Ultraschallbildgebung ist im Vergleich zu CT oder SPECT ein relativ kostengünstiges Verfahren, das zunehmend in den Kliniken eingesetzt wird. Derzeit liegt ein Hauptanwendungsgebiet in der Pränataldiagnostik, die Technik ist jedoch vielseitig für die Aufnahme morphologisch begrenzter Regionen geeignet. Im Projekt: „Dreidimensionalen Subtraktions-Ultraschallbildgebung und -verarbeitung zur verbesserten Diagnostik von Schilddrüsenerkrankungen und insbesondere von Schilddrüsenkrebs" ist ein entsprechender DFG Antrag formuliert, der einen Bearbeitungszeitraum von drei Jahren vorsieht. Bisher symptomatische Knoten der Schilddrüse kommen in Deutschland bei etwa 23 % der Bevölkerung vor. Nur ein kleiner Teil hiervon ist behandlungsbedürftig, insbesondere ist die Inzidenz von Schilddrüsen-

krebs mit etwa 3: 100.000 im Vergleich zu dieser Zahl außerordentlich gering. Viele dieser sogenannten Inzidentalome werden nach Ausschluss einer klinisch relevanten Autonomie oder einer im Hinblick auf den Nachweis von Malignität ergebnislosen Feinnadelbiopsie longitudinal über mehrere Jahre hinweg mit Ultraschall (US) kontrolliert. Dieser wird in der Regel in zweidimensionaler (2D) Technik durchgeführt und als 2D-Bild dokumentiert. Die Reproduzierbarkeit von Befunden des 2D-US ist bekanntlich eingeschränkt. Dieses gilt insbesondere für das Vermögen dieser Technik, Veränderungen von Volumen und Echotextur von Schilddrüsenknoten longitudinal über mehrere Jahre hinweg zu erfassen. Eine größere Genauigkeit ist von 3D-US zu erwarten, da mit dieser Technik 3D-Datensätze der thyreoidalen Echotextur generiert, analysiert und dokumentiert werden können. Ein Ziel besteht deshalb in der Registrierung zeitlich versetzter 3D-US Aufnahmen mit einer initialen SPECT-CT Aufnahme. Dadurch lassen sich quantitative Rückschlüsse über morphologische Veränderungen ziehen. Es wird in diesem Projekt vor allem Wert auf die Zusammenarbeit des Lehrstuhls für Mustererkennung (Prof. Dr.-Ing. J. Hornegger) mit klinischen Instituten gelegt. Dazu zählen unter anderem die Nuklearmedizinische Klinik mit Poliklinik (Prof. Dr. med. T. Kuwert) und die Abteilung Spezial-Ambulanz für Pränatale Diagnostik m. spez. Ultraschall (Prof. Dr. med. R. Schild) der Frauenklinik (Prof. Dr. med. W. Beckmann) an der Universität Erlangen-Nürnberg.

## **2.11 SFB 539, A4: Automatisches Glaukom-Screening**

### **Projektleitung:**

Prof. Dr.-Ing. Heinrich Niemann, Prof. Dr. Georg Michelson,  
PD Dr.rer.nat. Berthold Lausen, Prof. Dr.-Ing. Joachim Hornegger

### **Beteiligte:**

Ing. Radim Chrastek, Dr. med. Victor Derhartunian, Dipl.-Inf. Werner Adler,  
Dipl.-Inf. Rüdiger Bock

### **Stichwörter:**

Glaukom; optischer Sehnervenkopf; Segmentierung; morphologische Operationen;  
Hugh-Transformation; aktive Konturen

**Laufzeit:** 1.7.2003 - 30.6.2006

**Förderer:** Deutsche Forschungsgemeinschaft

Der Lehrstuhl für Mustererkennung befasst sich im Rahmen des SFB 539 (<http://www.sfb539.forschung.uni-erlangen.de>) (Teilprojekt A.4, Automatisches Glaukom -Screening (<http://www.imbe.med.uni-erlangen.de/~teilpa4/>) ) mit der Segmentierung des optischen Sehnervenkopfes für eine Glaukomfrüherkennung. Der Sehnervenkopf (der Austritt des Sehnervs aus dem Auge) wird mit dem Heidelberg Retina Tomograph (HRT) aufgenommen. Hierbei handelt es sich um ein Gerät, welches auf dem Prinzip eines konfokalen Mikroskops basiert. Anschließend werden verschiedene Oberflächen- und Volumenparameter (Merkmale) berechnet, die zur Glaukomklassifizierung dienen. Bis jetzt mussten hierfür die Grenzen des optischen



Sehnervenkopfes (sog. Elschnigscher Sklerering) von einem Augenarzt aufwändig manuell markiert werden. Es wurden zwei Verfahren für die Segmentierung des optischen Sehnervenkopfes entwickelt.

Das erste Verfahren verwendet die HRT-Reflektivitätsbilder, d.h. Bilder, die durch Mittelwertbildung aus den Schichtbildern berechnet werden. Dieses Verfahren basiert auf morphologischen Operationen, Hough -Transformation und aktiven Konturen. Da die HRT-Bilder Beleuchtungsunterschiede aufweisen, werden sie durch Korrekturkoeffizienten verbessert. Die Korrekturkoeffizienten werden aus der approximierten Hintergrundbeleuchtung geschätzt. In den nächsten Schritten werden die Lage und der Radius des optischen Sehnervenkopfes geschätzt. Der optische Sehnervenkopf gehört zusammen mit den Gefäßen zu den dunkelsten Stellen im Bild. Sie werden durch einfache Binarisierung extrahiert. Die Regionen des optischen Sehnervenkopfes werden dann anhand der geometrischen Eigenschaften von den Gefäßen getrennt. Außerdem werden relevante anatomische Strukturen des optischen Sehnervenkopfes segmentiert (sog. parapapilläre Atrophie und neuroretinaler Randsaum). Anhand der gewonnenen Informationen werden dann die genauen Grenzen des optischen Sehnervenkopfes mittels der aktiven Konturen ermittelt. Das Verfahren wurde durch Klassifikationsraten von verschiedenen Klassifikatoren anhand der Fallkontrollstudie mit von einem erfahrenen Augenarzt eingezeichneten Grenzen des optischen Sehnervenkopfes validiert. Folgende Ergebnisse hinsichtlich der Glaukomdiagnose (normal/glaukomatös) wurden erzielt: linear discriminant analysis mit **27.7%** der geschätzten Fehlerrate für automatische Segmentierung (*aut*) und **26.8%** der geschätzten Fehlerrate für manuelle Einzeichnung (*man*), classification trees mit **25.2%** (*aut*) und **22.0%** (*man*) und bootstrap aggregation mit **22.2%** (*aut*) und **13.4%** (*man*).

Die oben aufgeführten Fehlerraten zeigen, dass die Ergebnisse des Algorithmus mit denen eines Augenarztes vergleichbar sind, er aber trotzdem noch verbessert werden muss. Diese Verbesserungen wurden durch die Fusion von HRT - Bildern und Farbfotos erzielt. Die beiden Bildmodalitäten beinhalten komplementäre Informationen des optischen Sehnervenkopfes. Die Segmentierung in HRT - Bildern wird durch Information aus den Farbfotos gesteuert. Zum einen werden die segmentierten Gefäße in Farbfotos zur Eliminierung von Fehlern der aktiven Kontur verwendet. Zum anderen wird der Suchraum der aktiven Konturen durch die Kontur des optischen Sehnervenkopfes, die in den Farbfotos segmentiert wurde, beschränkt. Dieses multimodale Verfahren brachte eine qualitative Verbesserung der Erkennungsrate von ursprünglich ca. 74% (für das monomodale Verfahren) auf 89%. Die quantitative Evaluation wird noch durchgeführt werden.

## **2.12 SFB 603, B2: Optimierungsansatz für die Integration von Kamerabildern bei der Klassifikation**

### **Projektleitung:**

Prof. Dr.-Ing. Heinrich Niemann

### **Beteiligte:**

Dipl.-Ing. Christian Derichs, Dipl.-Inf. Benjamin Deutsch

### **Stichwörter:**

Optimierung; Sensordatenfusion; Objektverfolgung; Objektklassifikation

**Laufzeit:** 1.1.1998 - 31.12.2006

**Förderer:** Deutsche Forschungsgemeinschaft

### **Kontakt:**

Dipl.-Inf. Benjamin Deutsch

Tel.: +49.9131.85.27891, Fax: +49.9131.303811,

E-Mail: deutsch@informatik.uni-erlangen.de

Das langfristige Ziel des Teilprojektes B2 ([http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/B2/tpB2\\_home\\_g.html](http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/B2/tpB2_home_g.html)) des Sonderforschungsbereiches 603 (<http://sfb-603.uni-erlangen.de/>) ist die Entwicklung von optimierungsbasierten Verfahren für die Integration von Kamerabildern bei der Klassifikation, Lokalisation und Verfolgung von Objekten. Die Darstellung des Wissens über das System basiert dabei immer auf einem probabilistischen Ansatz, welcher durch die Fusion aller zur Verfügung stehenden Informationsquellen generiert und ständig erneuert wird. Entscheidend für die Güte von Schätzungen bezüglich Objektklasse, -lage und -position ist dabei das aktive Vorgehen bei der Sensordatenakquisition.

In der Vergangenheit wurde der Großteil der Arbeit in die theoretische Optimierung der Methoden zur aktiven Ansichtenauswahl sowie zur aktiven Objektverfolgung investiert. Obwohl identische informationstheoretische Ansätze verwendet werden, wurden beide Bereiche des Teilprojektes B2 hauptsächlich unabhängig voneinander optimiert. Mit der zunehmenden Integration von Objektverfolgung und -klassifikation in ein gemeinsames System ist nun auch der wechselseitige Informationsaustausch und damit die gegenseitige Unsicherheitsreduktion der beiden Teilbereiche ermöglicht worden.

Hinsichtlich der praktischen Anwendung in einem gemeinsamen Demosystem musste vor allem auf dem Gebiet der Ansichtenplanung die vereinfachende Einschränkung von eindimensionalen Sensorbewegungen um das Objekt aufgehoben werden. Im vergangenen Jahr wurde dazu die Möglichkeit geschaffen, Sensoren nicht mehr nur auf einer Kreisbahn, sondern auf einer Halbkugel auf das Objekt auszurichten. Auf diese Weise kann jede in der Objektverfolgung mögliche Ansicht so verarbeitet werden, dass sie von der Ansichtenplanung genutzt werden kann. Die Herausforderung bestand dabei darin, geplante aber unmögliche Kamerabewegungen unter die Halbkugelebene möglichst geschickt aufzufangen. Die Belegung solcher Aktionen mit harten

Straftermen während des Reinforcement Learnings erwies sich dabei als suboptimal, da so eine ungerechtfertigte Bevorzugung bestimmter Ansichtenpositionen erzielt wurde. Das Problem konnte allerdings durch die systematische, eindeutige Zuordnung von ausführbaren Aktionen zu jeder unmöglichen Aktion behoben werden. Durch eine darauf adaptierte Modifizierung der Wissensrepräsentation kann das System nun gänzlich ohne illegale Aktionen beschrieben werden.

Ein weiterer Arbeitspunkt ergab sich aus der Forderung, die Ansichtenplanung nicht nur einzusetzen, um auf einem vorgegebenen Objektmodell mit wenigen Ansichten eine hohe Klassifikationsrate zu erzielen, sondern gleichzeitig dieses Modell optimal zu verbessern. Bezüglich dieses Anspruchs, aktives Lernen zu ermöglichen, wurde zuletzt ein Verfahren implementiert, welches Merkmale mittels linearer Regression an Positionen schätzt, zu denen keine Sensoraufnahme existiert. Die lokale Beschränkung hinsichtlich der Stützstellen für diese Approximation geschieht durch deren gaussförmig abfallende Gewichtung mit zunehmendem Abstand vom Evaluationspunkt. Auf diese Weise kann angegeben werden, an welchen Positionen eine erhöhte Klassenunsicherheit bzw. ein maximales Potential zur Unsicherheitsreduktion vorliegt. Durch primäres Anfahren solcher Ansichten konnte gezeigt werden, dass sich die Güte des zugrundeliegenden Modells vergleichsweise schnell erhöhen lässt.

In der Objektverfolgung wurden die Arbeiten bei der Mehrschritt-Aktionsplanung weiter fortgeführt. Durch die in realen Umgebungen entstehenden, variablen Kosten einer Kameraaktion ist es nicht mehr ausreichend, nur den Informationsgewinn einer einzigen zukünftigen Aktion zu betrachten. Stattdessen muss eine Folge von Aktionen betrachtet werden. Ausgehend von der bereits erfolgten Erweiterung des bisherigen, auf dem *Kalman Filter* basierenden Ansatz zur Bewertung einer Folge von Aktionen, konnte die Recheneffizienz im vergangenen Jahr deutlich erhöht werden.

Bislang wurde bei einer Folge von  $t$  Aktionen ein Sichtbarkeitsbaum aufgespannt, in dem sämtliche  $2^t$  möglichen Folgen der Sichtbarkeit des Objektes in den Kameras evaluiert wurden. Durch die Fusion der beiden auftretenden Zwischenergebnisse (mit und ohne Objektverlust in der Kamera) zu einer einzigen Normalverteilung kollabiert dieser Sichtbarkeitsbaum zu einer Sichtbarkeitsliste mit  $t$  Knoten. Dadurch wird die Rechenzeit linear in der Anzahl der Kameras und die Berechnung kann in Echtzeit durchgeführt werden. Durch diese Linearisierung kann in Verbindung mit dem im vergangenen Jahr untersuchten *Sequentiellen Kalman Filter* die optimale Aktionsfolge für alle beteiligten Kameras gemeinsam bestimmt werden, was bislang nicht möglich war.

Weitere Arbeiten wurden im Bereich der optimalen Auswahl von Merkmalen in der Objektverfolgung durchgeführt. Dadurch wurde die bisher verwendete, auf Intensitätsgradienten basierende Objektverfolgung mit nur 10 Bildpunkten stabil möglich. Tiefergehende quantitative Evaluierungen, sowie eine Beschleunigung der Suche nach optimal auszuwählenden Merkmalen, sind jedoch noch in Arbeit.

Die projektübergreifende Anwendbarkeit der im Teilprojekt B2 entwickelten Methoden wurde in vergangenen Jahr durch die Kooperation mit A1 ([http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/A1/tpA1\\_home\\_g.html](http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/A1/tpA1_home_g.html)) deutlich. Hier konnte das Reinforcement Learning

durch eine Anpassung der Belohnungsfunktion auch für die Aufgabe der Objektrekonstruktion eingesetzt werden. Wiederum wurde die Anzahl der benötigten Ansichten, um eine zuverlässige Rekonstruktion zu generieren, deutlich reduziert. Sowohl ein zufälliges Vorgehen in der Ansichtenauswahl als auch die manuelle Kameraführung konnten diesbezüglich nicht an die erzielten Werte heranreichen.

### **Publikationen**

1. Derichs, Christian; Deinzer, Frank; Niemann, Heinrich: Cost Integration in Multi-step Viewpoint Selection for Object Recognition. In: Perner, Petra; Imiya, Atsushi (Hrsg.): Machine Learning and Data Mining in Pattern Recognition (4th International Conference, MLDM 2005 Leipzig 9-11.7.2005). Heidelberg : Springer, 2005, S. 415-425. - ISBN 3-540-26923-1.
2. Deutsch, Benjamin; Niemann, Heinrich; Denzler, Joachim: Multi-Step Active Object Tracking with Entropy Based Optimal Actions Using the Sequential Kalman Filter. In: IEEE (Veranst.): International Conference on Image Processing - ICIIP'05 (IEEE International Conference on Image Processing, 2005 Genua, Italien 11.09.-14.09.2005). Bd. 3. 2005, S. 105-108. - ISBN 0-7803-9135-7.
3. Deutsch, Benjamin; Deinzer, Frank; Zobel, Matthias; Denzler, Joachim: Active Sensing Strategies for Robotic Platforms, with an Application in Vision-Based Gripping. In: Araújo, H.; Vieira, A.; Braz, J.; Encarnaçã, B.; Carvalho, M. (Hrsg.): INSTICC (Veranst.): Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics (1st International Conference on Informatics in Control, Automation and Robotics Setúbal). Bd. 2. Setúbal, : INSTICC Press, 2004, S. 169-176.

## 2.13 SFB 603, B6: Rechnergestützte Endoskopie des Bauchraums

### **Projektleitung:**

Prof. Dr. med. Werner Hohenberger, Prof. Dr.-Ing. Joachim Hornegger,  
Prof. Dr. Günther Greiner, OA PD Dr. med. Thomas Horbach,  
Dr. med. Sophie Krüger

**Beteiligte:** Dipl. Med.-Inf. Jochen Penne, Dipl.-Inf. Marco Winter

**Stichwörter:** Endoskopie; Minimal-invasive Operation; 3-D-Visualisierung; Lichtfeld;  
Registrierung; Bildverbesserung

**Laufzeit:** 1.1.2000 - 31.12.2006

**Förderer:** Deutsche Forschungsgemeinschaft

### **Mitwirkende Institutionen:**

Chirurgische Universitätsklinik Erlangen  
Lehrstuhl für Graphische Datenverarbeitung Erlangen

### **Kontakt:**

Dipl. Med.-Inf. Jochen Penne  
Tel.: +49.9131.85.25246, Fax: +49.9131.303811,  
E-Mail: penne@informatik.uni-erlangen.de

Im Teilprojekt B6 "Computerunterstützte Endoskopie des Bauchraumes" des Sonderforschungsbereiches 603 besteht das Ziel darin, Augmented Reality während minimal-invasiv durchgeführter Eingriffe in der Bauchhöhle bereitzustellen. Nach der Integration eines optischen Trackingsystems zur Lagebestimmung des Endoskopes in das Gesamtsystem, wurden von Seiten der Mustererkennung zwei maßgebliche Ziele verfolgt: die Evaluation des Systems unter realistischen OP-Bedingungen und die Untersuchung verschiedener Ansätze zur Berechnung bzw. Akquirierung von Tiefeninformation.

Lichtfelder wurden sowohl im Labor als auch im OP rekonstruiert. Im Labor wurden Lichtfelder eines Silikonmodells der Leber und Gallenblase erstellt; Im OP wurden Lichtfelder während einer minimal-invasiv durchgeführten Gallenblasenentfernung rekonstruiert. Insgesamt wurden im Rahmen des Teilprojekts 50 Lichtfelder generiert. Grundlegende Idee zur Lichtfeldevaluation war der Vergleich von Bildpaaren. Erfahrenen Operateuren wurden Bildpaare gezeigt: Ein mit einem Endoskop aufgenommenes Bild und ein künstlich generiertes Bild, das aus dem zugehörigen Lichtfeld mit derselben Kameraposition berechnet wurde. Für die Berechnung des Lichtfeld-Bildes durften die drei am besten dafür geeigneten Bilder nicht verwendet werden, so dass das Bild aus Bildern benachbarter Kamerapositionen interpoliert werden musste. Die Qualität des Lichtfeld-Bildes hing somit stark von den berechneten Kameraparametern und der Tiefeninformation ab. Die Operateure konnten zur Evaluation Noten von 1 (kein Unterschied zwischen den Bildern sichtbar) bis 4 (große Unterschiede sichtbar; Lichtfeld-Bild nicht geeignet um damit arbeiten zu können) vergeben. Als Durchschnittsnote für die Lichtfeld-Bilder ergab

sich dabei der Wert 2,9. Das heißt, ein Unterschied ist bemerkbar, allerdings kann mit den Lichtfeld-Bildern gearbeitet werden. Der Durchschnittswert für die Labor-Lichtfelder betrug 2,5; der Durchschnittswert der OP-Lichtfelder war 3,3. Beim Vergleich zweier bewegter Ansichten (eine endoskopisch aufgenommene Bildfolge und eine generierte Videosequenz, die aus einem Lichtfeld berechnet wurde) verbesserte sich die Durchschnittsnote der Lichtfelder auf 1,9. Das zeigte die Ebenbürtigkeit der Lichtfeldvisualisierung gegenüber den realen Bildströmen aus dem Endoskop.

Ein weiterer wichtiger Punkt war, verschiedene Möglichkeiten zur Tiefendatenakquisition zu untersuchen. Die Benutzung eines Stereo-Endoskopes zur Berechnung von Tiefeninformationen aus korrespondierenden Bildpaaren lieferte nur schlechte Ergebnisse. Obwohl die Anzahl der gefundenen Punkt-Korrespondenzen in entsprechenden Bild-Paaren ausreichend groß war, konnte auf Grund der kleinen Stereo-Basis des Endoskopes keine hinreichend exakte Tiefeninformation berechnet werden. Es wurde gezeigt, dass bei einer Stereo-Basis von etwa 5 cm die Tiefenwerte entsprechend genau berechnet werden können; allerdings ist diese Basis ungeeignet für den Einsatz bei minimal-invasiven Eingriffen.

Aus den in einer anatomischen Datenbank gespeicherten segmentierten CT-Daten wurden 3-D-Modelle in Form von Oberflächennetzen erzeugt. Diese wurden anschließend in ihrer Komplexität reduziert und geglättet. Für die Visualisierung wurden Rendermodule realisiert, welche sowohl die Dreiecksnetze der 3-D-Modelle als auch deren zugrundeliegende CT-Datensätze mittels Volumenvisualisierung darstellen können. Darauf aufbauend wurde die angestrebte erweiterte 3-D-Lichtfeld-Visualisierung implementiert. Diese erlaubt die fusionierte Darstellung des aufgenommenen Lichtfelds des Operationsgebiets zusammen mit präoperativ gewonnenen CT-Daten bzw. 3-D-Modellen. Auf diese Weise wurde ein wichtiger Schritt in Richtung Augmented Reality vollzogen. Desweiteren wurden Untersuchungen bezüglich der sinnvollen Aufbereitung dieser fusionierten Darstellung durchgeführt. Ergebnis dieser Forschung war zum Einen eine komplexe Überblendtechnik, die es u.a. erlaubt, vom Lichtfeld verdeckte Bereiche der CT-Daten semi-transparent darzustellen. Andererseits wurde eine sog. Abstandsvisualisierung realisiert, welche die räumlichen Abstände zwischen Lichtfeld und CT-Daten mittels einer Farbkodierung visuell darstellt. Frei definierbare Farbtabelle erlauben dabei eine Anpassung der Visualisierung an konkrete Anwendungsszenarien.

## **Publikationen**

1. Vogt, Florian; Krüger, Sophie; Winter, Marco; Niemann, Heinrich; Hohenberger, Werner; Greiner, Günther; Schick, Christoph: Erweiterte Realität und 3-D Visualisierung für minimal-invasive Operationen durch Einsatz eines optischen Trackingsystems. In: Meinzer, H.-P.; Handels, H.; Horsch, A.; Tolxdorff, T. (Hrsg.): Proceedings Bildverarbeitung für die Medizin (Bildverarbeitung für die Medizin Heidelberg 13.-15.3.2005). Berlin : Springer, 2005, S. 217-221. - ISBN 3-540-25052-2.
2. Winter, Marco; Greiner, Günther; Vogt, Florian; Niemann, Heinrich; Krüger, Sophie: Visualizing distances between light field and geometry using projective texture mapping. In: Greiner, Günther; Hornegger, Joachim; Niemann, Heinrich; Stamminger, Marc (Hrsg.): Proceedings of the Workshop Vision, Modeling, and Visualization (Workshop Vision, Modeling, and Visualization Erlangen 16.-18.11.2005). St. Augustin : Infix, 2005, S. 257-264. - ISBN 3-540-25052-2.

## 2.14 SFB 603, C2: Analyse, Codierung und Verarbeitung von Lichtfeldern zur Gewinnung realistischer Modelldaten

### Projektleitung:

Prof. Dr.-Ing. Heinrich Niemann, Prof. Dr. Günther Greiner

### Beteiligte:

Dipl.-Inf. Ingo Scholz, Dipl.-Inf. Christian Vogelgsang

### Stichwörter:

SFB 603, bildbasierte Modellierung; und Verarbeitung von Lichtf...

**Laufzeit:** 1.1.1998 - 31.12.2006

**Förderer:** DFG

**Mitwirkende Institutionen:** Lehrstuhl für Graphische Datenverarbeitung Erlangen

### Kontakt:

Dipl.-Inf. Ingo Scholz

Tel.: +49.9131.85.27275, Fax: +49.9131.303811,

E-Mail: scholz@informatik.uni-erlangen.de

Das Teilprojekt C2 ([http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/C2/tpC2\\_home\\_g.html](http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/C2/tpC2_home_g.html)) des Sonderforschungsbereichs 603 ([http://sfb-603.uni-erlangen.de/HTML/sfb603\\_g.html](http://sfb-603.uni-erlangen.de/HTML/sfb603_g.html)) bearbeitet zusammen mit dem Lehrstuhl für Graphische Datenverarbeitung (<http://www9.informatik.uni-erlangen.de/>) (LGDV) das Thema der "Analyse, Modellierung und Verarbeitung von Lichtfeldern zur Gewinnung realistischer Modelldaten". Die in diesem Teilprojekt zentrale Datenstruktur des Lichtfelds erlaubt es, durch sog. bildbasierte Modellierung beliebige Ansichten einer realen Szene zu generieren, deren Aussehen durch eine Sammlung an Bilddaten bekannt ist. Die benötigten Informationen über Parameter und Positionen der verwendeten Kamera werden über Verfahren der "Struktur aus Bewegung" (Structure from Motion) direkt aus den Bilddaten ermittelt, bei denen es sich im Allgemeinen um einen Bildstrom einer handgeführten Kamera handelt. Weitere Daten können beispielsweise Bilder einer auf einem Roboterarm montierten Kamera sein. Das Teilprojekt C2 wird bereits seit 1998 von der Deutschen Forschungsgemeinschaft (<http://www.dfg.de/>) (DFG) gefördert.

Die bei Lichtfeldern übliche Betrachtung von ausschließlich statischen Szenen wurde in den letzten Jahren immer mehr auch auf bewegte Szenen erweitert. Die bereits in 2003 begonnenen Arbeiten zur Rekonstruktion von Lichtfeldern aus Szenen, die bewegte, aber in sich starre Objekte enthalten, wurden auch 2005 fortgesetzt. Nach der Trennung von Merkmalen in "Hintergrund" und "bewegtes Objekt" können nun die beiden unabhängig voneinander berechneten 3-D-Rekonstruktionen mit ausreichender Genauigkeit registriert werden. Daraus lässt sich nun die Eigenbewegung des Objekts berechnen. Ähnliche Objektpositionen werden mittels einer Vektorquantisierung jeweils einem Zeitschritt zugeordnet. Die Visualisierung der so entstehen-

den *dynamischen Lichtfelder* erfolgt durch sog. *Vertrauenskarten* (Confidence Maps), die das Ausblenden des bewegten Objekts zu nicht sichtbaren Zeitschritten erlauben.

Ein neu entwickeltes Verfahren zur Evaluation der Lichtfeldqualität über die Berechnung des Signal-Rausch-Verhältnisses (SNR) ermöglicht einen objektiven Vergleich verschiedener Verfahren zur Lichtfeldrekonstruktion. Es wurde bereits für Untersuchungen zur Berechnung intrinsischer Kameraparameter und für den Vergleich verschiedener Verfahren zur Berechnung von Tiefeninformation eingesetzt. Dabei wurden auch die erstmals zur Lichtfeldvisualisierung eingesetzten *lokalen Proxies*, also nur für jede Ansicht lokal gültige Dreiecksnetze, mit herkömmlichen Tiefenkarten verglichen. Des Weiteren wurde das SNR dazu genutzt, die Tiefenkarten eines vorhandenen Lichtfeldes über die Rückkopplung aus der Lichtfeldsynthese zu optimieren.

Die Anwendbarkeit der entwickelten Verfahren wurde außerdem in mehreren Kooperationen mit anderen Projekten gezeigt. In Zusammenarbeit mit dem Graduiertenkolleg "Dreidimensionale Bildanalyse und -synthese" (<http://www9.informatik.uni-erlangen.de/Research/gk244/index.html>) wurden 3-D-Rekonstruktionsalgorithmen zum schnellen Training für die Objekterkennung eingesetzt. Im Teilprojekt A7 ([http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/A7/tpA7\\_home\\_g.html](http://sfb-603.uni-erlangen.de/HTML/TEILPROJEKTE/A7/tpA7_home_g.html)) des Sonderforschungsbereichs 603 wurden diese Algorithmen für die 3-D-Rekonstruktion von mittels flexibler Endoskope aufgenommenen Bildfolgen eingesetzt.

## Publikationen

1. Scholz, Ingo; Vogelgsang, Christian; Denzler, Joachim; Niemann, Heinrich: Dynamic Light Field Reconstruction and Rendering for Multiple Moving Objects. In: Katsushi Ikeuchi (Hrsg.): Proceedings of the Ninth IAPR Conference on Machine Vision Applications (IAPR Conference on Machine Vision Applications Tsukuba Science City, Japan 16-18.5.2005). Tokyo, Japan : IAPR MVA Conference Committee, 2005, S. 184-188. - ISBN 4-901122-04-5.
2. Niemann, Heinrich; Scholz, Ingo: Evaluating the Quality of Light Fields Computed from Hand-held Camera Images. In: Pattern Recognition Letters 26 (2005), Nr. 3, S. 239-249.
3. Scholz, Ingo; Denzler, Joachim; Niemann, Heinrich: Calibration of Real Scenes for the Reconstruction of Dynamic Light Fields. In: IEICE Transactions on Information & Systems E87-D (2004), Nr. 1, S. 42-49.



## 2.15 SmartWeb

### **Projektleitung:**

Dr.-Ing. Elmar Nöth

### **Beteiligte:**

Dr. phil. Anton Batliner, Dipl.-Inf. Christian Hacker, Dipl.-Inf. Florian Hönig,  
Dipl.-Inf. Axel Horndasch

### **Stichwörter:**

Spracherkennung; OOV-Verarbeitung; Biosignale; Benutzerzustandsklassifikation;  
multimodale Informationsquellen

**Laufzeit:** 1.4.2004 - 31.3.2007

**Förderer:** Bundesministerium für Bildung und Forschung

### **Kontakt:**

Dr.-Ing. Elmar Nöth

Tel.: +49.9131.85.27888, Fax: +49.9131.303811,

E-Mail: noeth@informatik.uni-erlangen.de

## **Leitinnovation SmartWeb**

Das Ziel des BMBF-geförderten Projekts *SmartWeb* (<http://www.smartweb-projekt.de>) ist der intelligente, multimodale, mobile Zugriff auf das Semantische Web. An der Umsetzung dieser Vision arbeiten seit Mitte 2004 vierzehn Konsortialpartner unter der Führung des Deutschen Forschungsinstituts für künstliche Intelligenz (DFKI). Neben einer Reihe von Universitäten und Forschungseinrichtungen sind auch kleinere und mittlere Unternehmen sowie die Großindustrie beteiligt. Die Planungen sehen eine Laufzeit von ca. drei Jahren bis 2007 vor.

Ein großes Teilprojekt in SmartWeb bildet das *Semantische Web*. Dieses *Internet der nächsten Generation* basiert auf der inhaltlichen Beschreibung digitaler Dokumente mit standardisierten Vokabularien, die eine maschinell verarbeitbare Semantik haben. In SmartWeb wird vor allem angestrebt, die maschinenlesbaren Inhalte des World Wide Web auch maschinell verarbeitbar zu machen. Die Arbeitspakete, die am *Lehrstuhl für Mustererkennung (Informatik 5)* bearbeitet werden, sind Teil des zweiten Teilprojekts, das sich mit dem multimodalen, mobilen Zugang zu den Inhalten des Semantischen Webs beschäftigt. Das Architekturmodell von SmartWeb sieht dabei ein Client-Server-Szenario vor, bei dem die Benutzereingaben von einem Endgerät (PDA, Smartphone o.Ä.) zu einem Server geschickt werden, auf dem die Dialogverarbeitung stattfindet. Einen Teil des Dialog-Servers bildet ein multimodaler Erkenner (MME), der neben einem Spracherkenner weitere Software-Module zur Verarbeitung der verschiedenen Eingabemodalitäten (u.a. Sprach-, Videosignale) umfasst. Drei der Module werden am *Lehrstuhl für Mustererkennung (LME)* entwickelt: die Detektion und Verarbeitung von Out-Of-Vocabulary-Wörtern (OOV-Wörtern, also Wörter, die nicht im Erkennungsvokabular sind) bei der Spracherkennung

(in enger Zusammenarbeit mit der Firma Sympalog), die Klassifikation des Benutzerzustands auf der Basis von Sprach-, Video- und Biosignalen und die Klassifikation des Benutzerfokus mit Hilfe multimodaler Informationsquellen.

Nach der Integration einer ersten Version des OOV-Moduls in das SmartWeb-System im Jahre 2005 werden zur CeBIT 2006 auch die Einzelkomponenten zur Bestimmung des Benutzerfokus (Klassifikation von On-View/Off-View, On-Talk/Off-Talk) eingebunden. Die Ergebnisse der Arbeiten zur Benutzerzustandsklassifikation werden im Rahmen der Messe als separates Demonstrationssystem vorgestellt. Gemeinsam mit den Partnern aus der Industrie werden mögliche Szenarien, vor allem zum Einsatz der Verarbeitung von Biosignalen im Auto und beim Motorradfahren, evaluiert.

### **Erkennung und Verarbeitung von Out-Of-Vocabulary Wörtern**

Die in Absprache mit dem Projektpartner Sympalog erarbeitete Architektur des OOV-Moduls wurde im Rahmen der Integrationsarbeiten für den ersten SmartWeb-Demonstrator im Juni 2005 umgesetzt: Neben einem Worterkenner, dessen Sprachmodell auch Klassen für unbekannte Wörter enthält, wurde ein Monophonerkenner zur Generierung von Lauthypothesen für OOV-Bereiche herangezogen. Komplettiert wurde die erste Version des OOV-Moduls für das deutsche SmartWeb-System im September durch eine Komponente zur automatischen Umsetzung der Lauthypothesen in Buchstabenfolgen (Sound-to-letter-Problem).

Die Komponente zur Umsetzung der Lauthypothesen basiert auf einem neu entwickelten Ansatz zur Konvertierung von Symbolketten verschiedener Alphabete. Der Ansatz sieht eine vollautomatische Zuordnung einer Menge korrespondierender Zeichenketten (z.B. orthographisch und lautschriftlich vorliegende Lexikoneinträge) zu einem Trainingskorpus vor. Auf Basis der erlernten statistischen Parameter erzeugt ein Programm für eine (unbekannte) Symbolfolge eine Reihe von Hypothesen im gewünschten Zielalphabet. Das Gesamtsystem mit dem Namen MASSIVE (Mapping Arbitrary Symbol Sets Is Very Easy) kann neben der in SmartWeb benötigten Umwandlung von Lauthypothesen zum Beispiel auch zur Phonetisierung eingesetzt werden (Letter-to-sound-Problem). Bei Training und Tests mit dem deutschen Celex-Lexikon des Max-Planck-Instituts für Psycholinguistik in Nijmegen wurden sowohl beim Sound-to-letter- als auch beim Letter-to-sound-Problem über 94% der Wörter korrekt konvertiert.

#### **2.15.1 Benutzerzustandsklassifikation**

Für eine effiziente Interaktion zwischen SmartWeb und dem Benutzer kann es für das System von Vorteil sein, Informationen über den psychischen Zustand des Benutzers zu besitzen (etwa gestresst vs. entspannt oder ärgerlich vs. zufrieden). Besonders augenfällig wird dies im Auto- und Motorradszenario, wo das System idealerweise in der Lage sein sollte, in kritischen Situationen die Interaktion mit dem Fahrer durch eine Einschränkung des Informationsflusses anzupassen.

Sprache, Mimik und Gestik des Benutzers lassen häufig Rückschlüsse auf dessen Zustand zu, unterliegen jedoch meist einer mehr oder weniger starken Maskierung. Einen relativ direkten und unverfälschten Zugang zum Benutzerzustand können physiologische Parameter wie Haut-

leitwert, Puls usw. bieten. Für entsprechende Messungen existieren mobile Systeme, und es besteht die Hoffnung, dass sich wenigstens einige der Sensoren in Zukunft in die Kleidung, das Lenkrad oder ein mobiles Endgerät integrieren lassen. Mit dem Ziel, einen automatischen Klassifikator von Stress anhand physiologischer Signale zu erstellen, wurden im Jahr 2005 eine geeignete Stichprobe gesammelt und Untersuchungen zur Merkmalsauswahl und Klassifikation durchgeführt. Auf der CeBIT 2006 wird dieser Klassifikator in einem Live-System demonstriert.

### **Multimodale Erkennung des Benutzerfokus**

Im SmartWeb-Projekt kann der Benutzer gesprochene Anfragen an das System stellen. Mit einem mobilen Endgerät (z.B. dem T-Mobile MDA Pro) ist er dabei ortsungebunden: die Anfrage ("Wo ist das nächste italienische Restaurant?") wird an einen Server geschickt und dort analysiert. Dabei soll das System automatisch erkennen, ob es überhaupt angesprochen war. Vielleicht galten die gesprochenen Worte ja einer menschlichen Begleitung, vielleicht war es ein Selbstgespräch? Mit dem Mikrofon und der Kamera am MDA-Pro kann man alle notwendigen Informationen erhalten und der Benutzer muss keine lästige Push-To-Talk-Taste drücken. Wird ein Gesicht im Kamerabild detektiert, klassifiziert das System den Fokus der Aufmerksamkeit anhand der Blickrichtung: On-View, wenn der Benutzer direkt auf das Display schaut, oder Off-View, wenn der Benutzer nicht auf das Gerät, sondern beispielsweise zu einem Gesprächspartner blickt. Analog wird für ein Sprachsignal bewertet, ob der Benutzer direkt mit dem System spricht (On-Talk) oder mit sich selbst bzw. einem Dritten (Off-Talk).

Im Berichtszeitraum wurde ein Demonstrator implementiert, der auf der CeBIT 2006 vorgestellt wird. Für die Off-View-Detektion wurde der Viola-Jones-Algorithmus zur Gesichtsdetektion untersucht. Die Off-Talk-Erkennung basiert auf einer wortweisen Klassifikation mit 99 prosodischen Merkmalen. On-Talk und gelesener Off-Talk lassen sich besonders gut durch Dauermerkmale unterscheiden; On-Talk und der restliche Off-Talk, wie etwa Selbstgespräch oder Beiseitesprechen, durch Energiemerkmale. Die verschiedenen Off-Talk-Sprechregister unterscheiden sich auch durch Pausenmerkmale. Insgesamt ergibt sich für die On-Talk/Off-Talk-Klassifikation eine klassenweise gemittelte Erkennungsrate von 72%.

## **2.16 Techniken der Hardware-Beschleunigung für die 3D Bildrekonstruktion aus Projektionen in Kegelstrahlgeometrie**

### **Projektleitung:**

Prof. Dr.-Ing. Joachim Hornegger, Dr.-Ing. Markus Kowarschik

**Beteiligte:** Dipl.-Inf. Holger Scherl

**Stichwörter:** Computertomographie; 3D-Rekonstruktion; Hardwarenahe Programmierung; Parallele Algorithmen; Performanzanalyse

**Laufzeit:** 1.11.2004 - 31.10.2007

**Förderer:** Siemens Medical Solutions (Components Division)

### **Kontakt:**

Dipl.-Inf. Holger Scherl

Tel.: +49.9131.85.28977, Fax: +49.9131.85.28990,

E-Mail: [scherl@informatik.uni-erlangen.de](mailto:scherl@informatik.uni-erlangen.de)

Im Rahmen einer Kooperation des Lehrstuhls für Mustererkennung und Siemens Medical Solutions (Geschäftsgebiet Components) werden seit November 2004 neuartige bildgebende Verfahren in der Computertomographie (CT) und deren Realisierung mittels unterschiedlichen Hardwarearchitekturen untersucht. Im besonderen Interesse stehen Algorithmen für die 3D-Rekonstruktion in der Spiral-CT und in C-Bogen-CT-Systemen inklusive notwendiger Vor- und Nachverarbeitungsschritte.

Ein Schwerpunkt innerhalb des Projekts ist die Erforschung von Rekonstruktionsverfahren, die in zukünftigen Generationen von CT-Systemen eingesetzt werden können und eine deutliche Verbesserung der Bildqualität mit sich bringen. Es wurden neben iterativen Verfahren insbesondere exakte gefilterte Rückprojektionsalgorithmen untersucht.

Neben der Erforschung von Rekonstruktionsalgorithmen ist die Realisierung solcher Verfahren auf unterschiedlichen Hardwarearchitekturen ein weiterer Forschungsschwerpunkt des Projekts. Die Rekonstruktionsverfahren in der CT erfordern einerseits aufgrund der zu verarbeiteten Menge von Projektionsdaten und andererseits aufgrund der Größe und Anzahl der zu berechnenden Schnittbilder eine extrem hohe Rechenleistung des eingesetzten Rekonstruktionssystems. Die vorherrschenden Anforderungen an die Rechenzeit in Produktivsystemen erfordern nach dem heutigen Stand der Technik weiterhin den Einsatz skalierbarer Hardwarekomponenten. Im Fokus dieses Projekts steht die parallele Umsetzung von verschiedenen Rekonstruktionsalgorithmen auf Multiprozessorsystemen, Grafikkbeschleunigerkarten, Spezialprozessoren (wie beispielsweise die Cell Broadband Engine Architecture), sowie rekonfigurierbarer Beschleunigerhardware, die auf Field Programmable Gate Array (FPGA) Technologie basiert.

## 2.17 Untersuchung und Bewertung der Ersatzstimme Laryngektomierter (SVcheck)

### Projektleitung:

Prof. Dr. med. Frank Rosanowski

**Beteiligte:** Prof. Dr.med., Dr.rer.nat. Ulrich Eysholdt, Dr.-Ing. Elmar Nöth,  
Dipl.-Inf. Tino Haderlein

**Stichwörter:** Laryngektomie; Ersatzstimme; automatische Sprachverarbeitung

**Laufzeit:** 1.4.2005 - 31.3.2007

**Förderer:** Deutsche Krebshilfe

### Kontakt:

Dr.-Ing. Elmar Nöth

Tel.: +49.9131.85.27888, Fax: +49.9131.303811,

E-Mail: noeth@informatik.uni-erlangen.de

Die Anbahnung einer tracheoösophagealen Ersatzstimme (TE-Stimme) ist eine Möglichkeit, Patienten nach einer totalen Laryngektomie, d.h. Kehlkopfentfernung, die Fähigkeit zu sprechen zurück zu geben.

Ein Ventil zwischen Luft- und Speiseröhre erlaubt es, den Luftstrom aus der Lunge umzuleiten und Gewebeschwingungen in der Speiseröhre zur Ersatzstimmgebung zu nutzen. Die Betroffenen durchlaufen eine Therapie, in der wiederholt evaluiert werden muss, ob und wie sich ihre Ersatzstimme hinsichtlich Kriterien wie Lautstärke, Verständlichkeit oder Prosodiefähigkeit entwickelt hat. Da die Beurteilung subjektiv erfolgt und das Verfahren für Arzt und Patienten aufwändig ist, erscheint eine Automatisierung und Objektivierung in diesem Bereich sinnvoll.

In unserer Arbeit untersuchen wir, wie gut tracheoösophageale Sprache von einem automatischen Spracherkennungssystem erkannt wird und ob die Ermittlung der Qualität einer Ersatzstimme zumindest teilweise automatisiert erfolgen kann. Dazu müssen die Bewertungen der Maschine und einer Vergleichsgruppe von Experten korrelieren. Die Selbstbewertung der Patienten (SF-36, V-RQOL, VHI, Trierer Skalen) wird ebenfalls Eingang in die Berechnung eines kompakten Globalmaßes finden, welches automatisch erstellt wird und eine Aussage über die Qualität der Ersatzstimme trifft.

Die im Projekt erhobenen und analysierten Daten umfassen kontinuierliche Sprache, die sowohl mit einem Nahbesprechungsmikrofon als auch über das Telefon aufgenommen werden.

Bisher konnten wir zeigen, dass die von menschlichen Experten vergebene Verständlichkeitsnote für die TE-Aufnahmen sehr gut mit der von einem Spracherkennungssystem errechneten Wortakkuratheit korreliert ( $r=-0,84$ ). Ähnlich gute Korrelationen ergaben sich zwischen automatisch berechneten prosodischen Merkmalen und dem Kriterium "Übereinstimmung von Atem- und Sinneinheiten".

## Publikationen

1. Schuster, Maria; Nöth, Elmar; Haderlein, Tino; Steidl, Stefan; Batliner, Anton; Rosanowski, Frank: Can you Understand him? Let's Look at his Word Accuracy - Automatic Evaluation of Tracheoesophageal Speech. In: IEEE (Veranst.): Proceedings of ICASSP 2005 - International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005 - International Conference on Acoustics, Speech, and Signal Processing Philadelphia (USA) 18.-23.3.2005). 2005, S. 61-64. - ISBN 0-7803-8874-7.

## 2.18 Visual Active Memory Processes and Interactive Retrieval (VAMPIRE)

### Projektleitung:

Prof. Dr.-Ing. Heinrich Niemann, Prof. Dr.-Ing. Joachim Denzler

### Beteiligte:

Dipl.-Inf. Timo Zinßer, Dipl.-Inf. Christoph Gräßl, Dipl.-Inf. Jochen Schmidt

### Stichwörter:

Aktive Speicherprozesse; Kognitives Sehen; Rechnersehen; Erweiterte Realität; Objektverfugung; Objekterkennung; Bildbasierte Modellierung

**Laufzeit:** 1.5.2002 - 30.4.2005

**Förderer:** EU 5. Rahmenprogramm

### Mitwirkende Institutionen:

Bielefeld I Applied Computer Science, Bielefeld II Neuroinformatics, TU Graz, Surrey

### Kontakt:

Dipl.-Inf. Timo Zinßer

E-Mail: zinsser@informatik.uni-erlangen.de

Im Rahmen des EU-Projekts VAMPIRE (<http://www5.informatik.uni-erlangen.de/vampire>) (Visual Active Memory Processes and Interactive Retrieval) wird seit Juni 2002 Forschung im Bereich der automatischen Analyse von Videosequenzen durchgeführt. Im Vordergrund stehen dabei die Erkennung von Objekten und Bewegungsabläufen, sowie das Lernen neuer Objekt- und Bewegungsmodelle. Neben dem Lehrstuhl für Mustererkennung sind auch die Universität Bielefeld, die Technische Universität Graz und die University of Surrey am Projekt beteiligt.

Ein Schwerpunkt der Arbeit des Lehrstuhls für Mustererkennung innerhalb des Projekts VAMPIRE liegt in den Bereichen *3-D Rekonstruktion* und *bildbasierte Objektmodelle*. Insbesondere auf dem Gebiet der bildbasierten Objektmodelle besteht mit dem Teilprojekt C2 des SFB 603 (<http://sfb-603.uni-erlangen.de>) eine enge Zusammenarbeit. Um die Erzeugung der Objektmodelle noch robuster und effizienter zu gestalten, wurde ein *Struktur-aus-Bewegung* Verfahren implementiert, das selbst Bildsequenzen mit mehr als 200 Bildern innerhalb weniger Sekunden

verarbeitet. Die Eingabedaten für dieses Verfahren stammen von dem ebenfalls in VAMPIRE entwickelten Punktverfolger, der auf dem *Kanade-Lucas-Tomasi Tracker* basiert. Bisher durchgeführte Optimierungen erlauben bei der Verwendung einer digitalen VGA-Kamera das Verfolgen von 250 Merkmalen mit 30 Bildern pro Sekunde. Anwendungsspezifische Verbesserungen für eine von der TU Graz entwickelte Hochgeschwindigkeitskamera erhöhten die Bildrate bei der Verfolgung von zehn Merkmalen von 30 Bildern pro Sekunde auf 200 Bilder pro Sekunde.

Die Rekonstruktion von 3-D Punkten mittels Struktur-aus-Bewegung erlaubt es nicht, die Koordinaten der Punkte und der Kameras eindeutig zu bestimmen, da eine globale Skalierung der gesamten Szene die Eingabedaten nicht verändert. Wenn zwei Rekonstruktionen miteinander verglichen werden sollen, muss also neben der relativen Lage der Rekonstruktionen zueinander auch der relative Skalierungsfaktor ermittelt werden. Zu diesem Zweck wurde im Rahmen des Projekts VAMPIRE der bekannte Registrierungsalgorithmus ICP (*Iterative Closest Point*) um die Möglichkeit zur Schätzung dieses Skalierungsfaktors erweitert.

Neben der 3-D Rekonstruktion sind die *Objektverfolgung* und *Objekterkennung* ein weiterer Schwerpunkt im Projekt VAMPIRE, der vom Lehrstuhl für Mustererkennung untersucht wird. Bei der Objektverfolgung muss zwischen *datengetriebenen* und *modellbasierten* Verfahren unterschieden werden. Datengetriebene Verfahren besitzen die Eigenschaft, dass sie ohne a priori Wissen (beispielsweise Geometrieinformationen) über das zu verfolgende Objekt arbeiten können. Dies hat den Vorteil, dass sofort nach der Detektion einer Bewegung in einer Szene die Verfolgung eines unbekanntes Objektes starten kann. Nachteilig ist allerdings, dass bei starken externen Rotationen das Objekt verloren wird, weil kein Wissen vorhanden ist, wie das Objekt aus einer anderen Ansicht aussieht. Diesen Nachteil kompensieren Ansätze, die sehr robust gegenüber Erscheinungsänderungen sind, bzw. modellbasierte Verfahren, die während eines Trainings signifikante Informationen über das Objekt sammeln. Nachdem ein Objekt mit Hilfe eines Klassifikationsverfahrens erkannt wurde, kann von einem datengetriebenen auf ein modellbasiertes Verfahren umgeschaltet werden.

Sowohl bei der datengetriebenen als auch bei der modellbasierten Objektverfolgung wurden vorhandene Verfahren weiter verbessert und neue Verfahren entwickelt. Als Grundlage für neue Verfahren zur datengetriebenen Objektverfolgung diente ein Verfahren, welches auf Farbhistogrammmerkmalen und dem CONDENSATION Algorithmus basiert. Die Erweiterung dieses Verfahrens auf mehrere Kameras ermöglicht die robuste Schätzung der 3-D Position eines verfolgten Objekts mit einem datengetriebenen Verfahren. Durch die wahrscheinlichkeitsbasierte Schätzung der 3-D Position toleriert das Verfahren eine temporäre Verdeckung des Objekts in einer Kamera. Eine andere Erweiterung, bei der der oben beschriebene Punktverfolger integriert wurde, macht das Verfahren robuster bei der Verfolgung von Objekten vor einem ähnlich gefärbten Hintergrund. Auch zur modellbasierten Objektverfolgung mit 3-D Objektmodellen wurde ein neues Verfahren entwickelt. Die Kombination von SIFT-Merkmalen zur Initialisierung der Lageschätzung mit dem Punktverfolger zur effizienten Verfolgung der SIFT-Merkmale erlaubt eine vollständige 3-D Lageschätzung des verfolgten Objekts in Echtzeit.

### 3 Weitere Veröffentlichungen und Vorträge

1. Bajramovic, Ferid; Gräßl, Christoph; Denzler, Joachim: Efficient Combination of Histograms for Real-Time Tracking Using Mean-Shift and Trust-Region Optimization. In: Kropatsch, Walter; Sablatnig, Robert; Hanbury, Allan (Hrsg.): Pattern Recognition, 27th DAGM Symposium (27th Annual meeting of the German Association for Pattern Recognition Wien 31.8.-02.09.2005). Berlin : Springer, 2005, S. 254-261. - ISBN 3-540-28703-5.
2. Batliner, Anton: Automatic classification of emotional states: purpose, possibilities, prospects. Vortrag: Invited talks, Università Degli Studi di Bari, Bari, 07.06, 09.06.2005.
3. Batliner, Anton: Emotion, Interaktion, und Erkennungsraten. Vortrag: Austrian Research Institute for Artificial Intelligence (OFAI), Wien, 07.10.2005.
4. Batliner, Anton: Labelling Emotional User States and some Ideas w.r.t. Physiological Data. Vortrag: HUMAINE Physiology-Meeting, Universität Augsburg, Augsburg, 11.10.2005.
5. Batliner, Anton: Private emotions vs social interaction - Towards new dimensions in Research on emotions. Vortrag: ADAPTING THE INTERACTION STYLE TO AFFECTIVE FACTORS, Workshop in conjunction with User Modeling 05, Edinburgh, 25.07.2005.
6. Batliner, Anton; Möbius, Bernd: Prosodic Models, Automatic Speech Understanding, and Speech Synthesis: Towards the Common Ground? In: Barry, William J.; Dommelen, Wim A. van (Hrsg.): The Integration of Phonetic Knowledge in Speech Technology. Dordrecht : Springer, 2005, S. 21-44. - ISBN 1-4020-2635-8.
7. Batliner, Anton: Tales of Tuning - Prototyping for Automatic Classification of Emotional User States. Vortrag: Interspeech 2005, ISCA, Lissabon, 05.09.2005.
8. Batliner, Anton; Blomberg, Mats; D'Arcy, Shona; Elenius, Daniel; Giuliani, Diego; Gerosa, Matteo; Hacker, Christian; Russell, Martin; Steidl, Stefan; Wong, Michael: The PF-STAR Children's Speech Corpus. In: ISCA (Hrsg.): Proceedings of the 9th European Conference on Speech Communication and Technology (9th European Conference on Speech Communication and Technology - Interspeech 2005 Lisbon 4-8.9.2005). Bonn : ISCA, 2005, S. 2761-2764. - ISBN 1018-4074.
9. Batliner, Anton: The PF-STAR Children's Speech Corpus. Vortrag: Interspeech 2005, 07.09.2005.
10. Dennerlein, Frank; Katsevich, Alexander; Lauritsch, Günter; Hornegger, Joachim: Exact and efficient cone-beam reconstruction algorithm for a short-scan circle combined with various lines. In: The International Society for Optical Engineering (Veranst.): Proc. SPIE (Medical Imaging San Diego 17.2.2005). Bd. 5747. 2005, S. 388-399.
11. Derichs, Christian: Cost Integration in Multi-step Viewpoint Selection for Object Recognition. Vortrag: Machine Learning and Data Mining 2005, Renaissance Hotel, Leipzig, 11.07.2005.
12. Deutsch, Benjamin; Gräßl, Christoph; Bajramovic, Ferid; Denzler, Joachim: A Comparative Evaluation of Template and Histogram Based 2D Tracking Algorithms. In: Kropatsch, Walter; Sablatnig, Robert; Hanbury, Allan (Hrsg.): Pattern Recognition, 27th DAGM Symposium (27th Annual meeting of the German Association for Pattern Recognition Wien 31.8.-02.09.2005). Berlin : Springer, 2005, S. 269-276. - ISBN 3-540-28703-5.



13. Deutsch, Benjamin: A Comparative Evaluation of Template and Histogram Based 2D Tracking Algorithms. Vortrag: 27th Annual meeting of the German Association for Pattern Recognition, Wien, 01.09.2005.
14. Deutsch, Benjamin: Multi-Step Active Object Tracking with Entropy Based Optimal Actions Using the Sequential Kalman Filter. Vortrag: IEEE International Conference on Image Processing, Institute of Electrical and Electronics Engineers, Genua, Italien, 14.09.2005.
15. Deutsch, Benjamin: Teilprojekt B2 "Optimierungsansatz für die Integration von Kamerabildern bei der Klassifikation". Vortrag: SFB-Kolloquium, Friedrich-Alexander-Universität Erlangen-Nürnberg, Pommersfelden, 07.07.2005.
16. Frank, Carmen; Nöth, Elmar: Sprechen: Ein Hindernis in der modernen Mensch-Maschine-Kommunikation? In: DEGA (Veranst.): Fortschritte der Akustik - Proc. DAGA 05 (DAGA 05 München 14.-17.3.2005). München : DEGA, 2005, S. 729-730.
17. Greiner, Günther; Hornegger, Joachim; Niemann, Heinrich; Stamminger, Marc (Hrsg.): Vision, Modeling, and Visualization 2005, Proc. of the International Workshop, Erlangen, Germany. (Vision, Modeling, and Visualization 2005) Berlin : Akademische Verlagsgesellschaft, 2005. - ISBN 3-89838-068-X.
18. Grzegorzec, Marcin: Appearance-Based Statistical Object Recognition. Vortrag: Erasmus-Sokrates Austausch "Teaching Staff Mobility", Silesian University of Technology, Gliwice, 25.04.2005.
19. Grzegorzec, Marcin: Erscheinungsbasierte, statistische Objekterkennung mit Farb- und Kontextmodellierung. Vortrag: Einladung des Forschungsinstituts für Optronik und Mustererkennung, FGAN-Forschungsinstitut für Optronik und Mustererkennung, Ettlingen, 10.10.2005.
20. Grzegorzec, Marcin; Reinhold, Michael; Niemann, Heinrich: Feature Extraction with Wavelet Transformation for Statistical Object Recognition. In: Kurzynski, Marek; Puchala, Edward; Wozniak, Michal; Zolnierek, Andrzej (Hrsg.): Proceedings of the 4th International Conference on Computer Recognition Systems (4th International Conference on Computer Recognition Systems Rydzyna 22-25.05.2005). Berlin : Springer, 2005, S. 161-168. - ISBN 10 3-540-25054-9.
21. Grzegorzec, Marcin: Feature Extraction with Wavelet Transformation for Statistical Object Recognition. Vortrag: 4th International Conference on Computer Recognition Systems, Wroclaw University of Technology, Rydzyna, 23.05.2005.
22. Grzegorzec, Marcin: Statistical Object Recognition Including Color Modeling. Vortrag: 2nd International Conference on Image Analysis and Recognition, University of Waterloo, Toronto, 29.09.2005.
23. Gräßl, Christoph; Zinßer, Timo; Scholz, Ingo; Niemann, Heinrich: 3-D Object Tracking with the Adaptive Hyperplane Approach Using SIFT Models for Initialization. In: Ikeuchi, Katsushi (Hrsg.): Proceedings of the Ninth IAPR Conference on Machine Vision Applications (IAPR Conference on Machine Vision Applications Tsukuba Science City, Japan 16-18.5.2005). Tokyo, Japan : IAPR MVA Conference Committee, 2005, S. 5-8. - ISBN 4-901122-04-5.
24. Gräßl, Christoph: 3-D Object Tracking with the Adaptive Hyperplane Approach Using SIFT Models for Initialization. Vortrag: IAPR Conference on Machine Vision Applications, Tsukuba Science City, Japan, 16.05.2005.

25. Görz, Günther; Hornegger, Joachim: Sprach- und Bilderkennung. In: Bruns, Kai; Meyer-Wegener, Klaus (Hrsg.): Taschenbuch der Medieninformatik. München : Carl Hanser, 2005, S. 194-219.
26. Hacker, Christian; Batliner, Anton; Steidl, Stefan; Nöth, Elmar; Niemann, Heinrich; Cincarek, Tobias: Assessment of Non-Native Children's Pronunciation: Human Marking and Automatic Scoring. In: Kokkinakis, G.; Fakotakis, N.; Dermatas, E.; Potapova, R. (Hrsg.): SPEECOM 2005 Proceedings, 10th International Conference on SPEECH and COMPUTER (10th International Conference on Speech and Computer (SPECOM 2005) Patras, Greece 17.10.2005 - 19.10.2005). Bd. 1. Moscow, Patras : Moskow State Linguistics University, 2005, S. 123 - 126. - ISBN 5-7452-0110-x.
27. Hacker, Christian: Assessment of Non-Native Children's Pronunciation: Human Marking and Automatic Scoring. Vortrag: 10th International Conference on Speech and Computer, SPECOM 2005, University of Patras, Patras, Greece, 17.10.2005.
28. Hacker, Christian; Cincarek, Tobias; Gruhn, Rainer; Steidl, Stefan; Nöth, Elmar; Niemann, Heinrich: Pronunciation Feature Extraction. In: Kropatsch, Walter; Sablatnig, Robert; Hanbury, Allan (Hrsg.): Pattern Recognition, 27th DAGM Symposium (27th Annual meeting of the German Association for Pattern Recognition (DAGM 2005) Wien 31.08.2005 - 02.09.2005). Berlin : Springer, 2005, S. 141-148. - ISBN 3-540-28703-5.
29. Haderlein, Tino: Menschliche und automatische Verständlichkeitsbewertung bei tracheoösophagealen Ersatzstimmen. Vortrag: 31. Deutschen Jahrestagung für Akustik DAGA 2005, Deutsche Gesellschaft für Akustik e.V. (DEGA), München, 15.03.2005.
30. Haderlein, Tino; Steidl, Stefan; Nöth, Elmar; Schuster, Maria: Menschliche und automatische Verständlichkeitsbewertung bei tracheoösophagealen Ersatzstimmen. In: Fastl, Hugo; Fruhmann, Markus (Hrsg.): Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 31. Deutschen Jahrestagung für Akustik DAGA 2005, München (31. Deutsche Jahrestagung für Akustik, DAGA '05 München 14.-17.3.2005). Berlin : Deutsche Gesellschaft für Akustik e.V., 2005, S. 243-244. - ISBN 3-9808659-1-6.
31. Haderlein, Tino; Nöth, Elmar; Herbordt, Wolfgang; Kellermann, Walter; Niemann, Heinrich: Using Artificially Reverberated Training Data in Distant-Talking ASR. In: Václav Matousek; Pavel Mautner; Tomáš Pavelka (Hrsg.): Text, Speech and Dialogue; 8th International Conference TSD 2005 (8th International Conference on Text, Speech and Dialogue (TSD 2005) Karlovy Vary, Czech Republic 12.-16.9.2005). Berlin : Springer, 2005, S. 226-233. - ISBN 3-540-28789-2.
32. Haderlein, Tino: Using Artificially Reverberated Training Data in Distant-Talking ASR. Vortrag: 8th International Conference on Text, Speech and Dialogue (TSD 2005), University of West Bohemia in Pilsen, Czech Republic, Karlovy Vary, Czech Republic, 13.09.2005.
33. Hahn, Dieter; Hornegger, Joachim; Bautz, Werner; Kuwert, Torsten; Römer, Wolfgang: Unbiased Rigid Registration Using Transfer Functions. In: International Society for Optical Engineering (Veranst.): Proceedings of the SPIE Medical Imaging (SPIE International Symposium - Medical Imaging San Diego 12-17.02.2005). 2005, S. 151-162.
34. Hahn, Dieter: Unbiased Rigid Registration using Transfer Functions. Vortrag: SPIE International Symposium - Medical Imaging 2005, San Diego, 14.02.2005.

35. Hamberger, Maik; Moser, Michael; Nöth, Elmar; Levit, Michael; Deventer, Rainer: Der Einsatz unterschiedlicher Klassifikationssysteme zur Rutschungserkennung. In: Lehrstuhl für Angewandte Geologie (Veranst.): Berichte zur 15. Tagung für Ingenieurgeologie 2005 (Tagung für Ingenieurgeologie 2005 Erlangen). 2005, S. 179-185.
36. Horndasch, Axel; Gallwitz, Florian; Haas, Jürgen; Nöth, Elmar: Der mixed-initiative Ansatz als Basis für benutzerfreundliche Sprachdialogsysteme. In: Künstliche Intelligenz 19 (2005), Nr. 3, S. 38-41.
37. Horndasch, Axel: eMotion and Speech - Von der Theorie zur Praxis. Vortrag: Future Talk, CeBIT 2005, future parc, Hannover, 11.03.2005
38. Hornegger, Joachim: Bildregistrierung. Vortrag: Informatik Kolloquium der Technischen Universität Darmstadt, Technische Universität Darmstadt, Darmstadt, 02.06.2005.
39. Hornegger, Joachim: Bildtechnik. Vortrag: 27. Sitzung der Kompetenzinitiative Medizin und Pharma, Forchheim, 26.01.2005.
40. Hornegger, Joachim: Bildverarbeitung in der Medizin. Vortrag: Informationstechnologien in den bayerischen Universitätskliniken, Erlangen, 01.04.2005.
41. Hornegger, Joachim: Computed Tomography in the Presence of Unknown Motion. Vortrag: 1st Bavarian Russian Workshop, München, 14.10.2005.
42. Hornegger, Joachim: Registration and Segmentation. Vortrag: Jahrestagung der Gesellschaft für Biomedizinische Technik, Nürnberg, 15.09.2005.
43. Hornegger, Joachim: Registrierung funktioneller und morphologischer Volumendatensätze. Vortrag: Faszination Technik, Fraunhofer Institut für Integrierte Schaltungen II, Erlangen-Tennenlohe, 06.01.2005.
44. Hornegger, Joachim: Registrierung funktioneller und morphologischer Volumendatensätze. Vortrag: Jahrestagung der Gesellschaft für Informatik, Bonn, 22.09.2005.
45. Hornegger, Joachim: Registrierung funktioneller und morphologischer Volumendatensätze. Vortrag: Bayerischer Radiologenkongress, Amberg, 15.10.2005.
46. Hornegger, Joachim: Technical Devices: Can They Learn? Vortrag: Symposium Kognitive Technische Systeme, Nürnberg, 12.09.2005.
47. Hornegger, Joachim: Von der Rekonstruktion bis zur Registrierung. Vortrag: SoftMed-Innovationsforum Medizinische Bildverarbeitung, Erlangen, 08.12.2005.
48. Hönig, Florian; Stemmer, Georg; Hacker, Christian; Brugnara, Fabio: Revising Perceptual Linear Prediction (PLP). In: ISCA (Veranst.): Proceedings of the 9th European Conference on Speech Communication and Technology (9th European Conference on Speech Communication and Technology - Interspeech 2005 Lisbon 4-8.9.2005). Bonn : ISCA, 2005, S. 2997-3000. - ISBN 1018-4074.
49. Jäger, Florian; Hornegger, Joachim; Hahn, Eckhart Georg: Formbasierte Segmentierung des Bronchialbaumes. In: Meinzer, Hans-Peter; Handels, Heinz; Horsch, Alexander; Tolxdorff, Thomas (Hrsg.): Bildverarbeitung für die Medizin 2005 (Workshop Bildverarbeitung für die Medizin 2005 Heidelberg 13.-15. März 2005). Berlin : Springer, 2005, S. 108-112. - ISBN 3-540-25052-2.
50. Jäger, Florian: Formbasierte Segmentierung des Bronchialbaumes. Vortrag: Workshop Bildverarbeitung für die Medizin 2005, Heidelberg, 14.03.2005.

51. Lell, Michael; Hornegger, Joachim; Scherl, Holger; Prümmer, Marcus; Anders, Katharina; Baum, Ulrich; Bautz, Werner: CTA: Semiautomatische Quantifizierung von Karotisstenosen. In: Wolf, Karl-Jürgen (Hrsg.): RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren (86. Deutscher Röntgenkongress Berlin 4.-7.5 2005). Stuttgart : Georg Thieme Verlag KG, 2005, S. PO\_139.
52. Maier, Andreas; Hacker, Christian; Steidl, Stefan; Nöth, Elmar: Helfen "Fallen" bei veräuschten Daten? - Spracherkennung mit TRAPs. In: Fastl, Hugo; Fruhmann, Markus (Hrsg.): Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 31. Deutschen Jahrestagung für Akustik DAGA 2005, München (31. Deutsche Jahrestagung für Akustik, DAGA '05 München 14. bis 17. März 2005). Bd. 1, 1. Aufl. Berlin : Deutsche Gesellschaft für Akustik e.V., 2005, S. 315-316. - ISBN 3-9808659-1-6.
53. Maier, Andreas: Helfen "Fallen" bei veräuschten Daten? - Spracherkennung mit TRAPs. Vortrag: 31. Deutsche Jahrestagung für Akustik, DAGA '05, TU München, DEGA, München, 16.03.2005.
54. Maier, Andreas; Hacker, Christian; Steidl, Stefan; Nöth, Elmar; Niemann, Heinrich: Robust Parallel Speech Recognition in Multiple Energy Bands. In: Kropatsch, Walter G.; Sablatnig, Robert; Hanbury, Allan (Hrsg.): Pattern Recognition, 27th DAGM Symposium, Vienna, Austria, August/September 2005, Proceedings (27th DAGM Symposium, Vienna, Austria, August/September 2005 Wien 30. August - 2. September). Bd. 1, 1. Aufl. Berlin : Springer, 2005, S. 133-140. (Lecture Notes in Computer Science Bd. 1, Nr. 3663) - ISBN 3-540-28703-5.
55. Maier, Andreas: Robust Parallel Speech Recognition in Multiple Energy Bands. Vortrag: Pattern Recognition 27th DAGM Symposium Vienna, Austria, Tu Wien, Pattern Recognition and Image Processing (PRIP) Group, Wien, 31.08.2005.
56. Münzenmayer, Christian; Wilharm, Sylvia; Hornegger, Joachim; Wittenberg, Thomas: Illumination Invariant Color Texture Analysis based on Sum- and Difference-Histograms. In: Kropatsch, Walter G.; Sablatnig, Robert; Hanbury, Allan (Hrsg.): Pattern Recognition - Proceedings of the 27th DAGM Symposium (27th DAGM Symposium Wien 30.8.-2.9.2005). Berlin, Heidelberg : Springer, 2005, S. 17-24.
57. Niemann, Heinrich: Aufnahme und Nutzung von Lichtfeldern. Vortrag: Fakultät für Mathematik und Informatik, Universität Jena, Jena, 24.01.2005.
58. Niemann, Heinrich: Towards AR in Laparoscopic Surgery. Vortrag: Istituto di Scienza e Tecnologia dell'Informazione "A. Faedo" (ISTI, CNR), Pisa, 05.04.2005.
59. Nömayr, Anton; Römer, Wolfgang; Hothorn, Torsten; Pfahlberg, Annette; Hornegger, Joachim; Bautz, Werner; Kuwert, Torsten: Anatomical accuracy of lesion localization. In: Nuklearmedizin 44 (2005), Nr. 4, S. 149-155.
60. Nöth, Elmar: Analyse pathologischer Stimmen. Vortrag: Kolloquium des Lehrstuhls für Mustererkennung, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, 19.01.2005. Nöth, Elmar: Can you Understand him? Let's Look at his Word Accuracy - Automatic Evaluation of Tracheoesophageal Speech. Vortrag: ICASSP 2005 - International Conference on Acoustics, Speech, and Signal Processing, ICASSP, Philadelphia (USA), 21.03.2005.
61. Nöth, Elmar: Das sprachgesteuerte ISA-Haus. Vortrag: Fortbildungsveranstaltung, Volkshochschule Erlangen, Erlangen, 17.11.2005.

62. Nöth, Elmar: Multimodale Erkennung in SmartWeb: Sprecheraktivität und On-Fokus/Off-Fokus. Vortrag: Projektlenkungssitzung des BMBF-Verbundprojektes SmartWeb, European Media Lab, Heidelberg, 16.12.2005.
63. Nöth, Elmar: Noten für Benoter: Maße zur Inter-Labeler-Beurteilung. Vortrag: Kolloquium der Abteilung für Phoniatrie und Pädaudiologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, 08.06.2005.
64. Nöth, Elmar: Phonetisches Wissen in der Sprachtechnologie. Vortrag: Kolloquium, Ludwig-Maximilians-Universität München, München, 18.04.2005.
65. Nöth, Elmar: Pronunciation Feature Extraction. Vortrag: DAGM 2005, DAGM, Wien, 31.08.2005.
66. Nöth, Elmar: Spracherkennung mit offenem Wortschatz in Smartweb. Vortrag: Projektstandssitzung des BMBF-Verbundprojektes SmartWeb, Siemens AG, München, 14.06.2005.
67. Nöth, Elmar: Sprachgesteuerte 3D-Angiographie - Die Software folgt dem Arzt aufs Wort. Vortrag: Kolloquium, IMKAI, Wien, 29.08.2005.
68. Prümmer, Marcus: 2D-3D Non-rigid Registration using Iterative Reconstruction. Vortrag: Workshop Vision Modeling and Visualization in Erlangen 2005, Friedrich-Alexander University, Erlangen, 17.11.2005.
69. Prümmer, Marcus: A full multigrid technique to accelerate an ART scheme for tomographic image reconstruction. Vortrag: Simulationstechnique 18th Symposium in Erlangen 2005, Friedrich-Alexander University, Erlangen, 13.09.2005.
70. Prümmer, Marcus: Increasing Temporal Resolution for 3D Cardiac Angiography using Non-Rigid Registration. Vortrag: 4th world congress of Cellular and Molecular Biology., World Society of Cellular and Molecular Biology, Poitiers, 07.-12.2005.
71. Prümmer, Marcus; Nöth, Elmar; Hornegger, Joachim; Horndasch, Axel: Mensch-Maschine Interaktion für den interventionellen Einsatz. In: Meinzer, Hans-Peter; Handels, Heinz; Horsch, Alexander; Tolxdorff, Thomas (Hrsg.): Bildverarbeitung für die Medizin 2005 (Workshop Bildverarbeitung für die Medizin 2005 Berlin 13.-15. März 2005). Berlin : Springer, 2005, S. 485-489. - ISBN 3-540-25052-2.
72. Rohlfing, Torsten; Denzler, Joachim; Gräßl, Christoph; Russakoff, Daniel; Maurer, Calvin: Markerless Real-Time 3D Target Region Tracking by Motion Backprojection from 2D Projection Images. In: IEEE Transactions on Medical Imaging 24 (2005), Nr. 11, S. 1455-1468.
73. Römer, Wolfgang; Hornegger, Joachim; Han, Jingfeng; Bautz, Werner; Kuwert, Torsten: Tumor Therapy Monitoring Based on Monomodal Image Fusion. In: The Radiological Society of North America (RSNA) (Veranst.): Proceedings of the Radiological Society of North America (RSNA 2004) (RSNA2004 Chicago, USA 28.11. - 3.12.2004). 2005, S. 355.
74. Scherl, Holger; Kowarschik, Markus; Hornegger, Joachim: Bit-Accurate Simulation of Convolution-Based Filtering on Reconfigurable Hardware. In: Hülsemann, Frank; Kowarschik, Markus; Rüde, Ulrich (Hrsg.): Frontiers in Simulation (Simulationstechnique 18th Symposium in Erlangen 2005 Erlangen 12.-15. September 2005). Erlangen : SCS Publishing House e.V., 2005, S. 662-667. - ISBN 3-936150-41-9.

75. Schmidt, Jochen; Vogt, Florian; Niemann, Heinrich: Calibration-Free Hand-Eye Calibration: A Structure-from-Motion Approach. In: Kropatsch, W.G.; Sablatnig, R.; Hanbury, A. (Hrsg.): Pattern Recognition, 27th DAGM Symposium (Pattern Recognition, 27th DAGM Symposium Wien 30.8.-2.9.2005). Berlin, Heidelberg, New York : Springer, 2005, S. 67-74. (Lecture Notes in Computer Science Bd. 3663) - ISBN 3-540-28703-5.
76. Scholz, Ingo: Analyse, Modellierung und Verarbeitung von Lichtfeldern.Vortrag: Berichtskolloquium des SFB 603, Sonderforschungsbereich 603, Universität Erlangen-Nürnberg, Pommersfelden, 07.07.2005.
77. Scholz, Ingo: Dynamic Light Field Reconstruction and Rendering for Multiple Moving Objects.Vortrag: Machine Vision Applications 2005, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Science City, Japan, 16.05.2005.
78. Wenhardt, Stefan: Ansichtenauswahl für die 3D Rekonstruktion.Vortrag: Einladung zum Lehrstuhl für Digitale Bildverarbeitung (Prof. Denzler), Friedrich-Schiller-Universität Jena, Institut für Informatik, Jena, 24.11.2005.
79. Winter, Christian; Scholz, Ingo; Rupp, Stephan; Wittenberg, Thomas: Reconstruction of tubes from monocular fiberoptic images -- Application and first results. In: Greiner, Günther; Hornegger, Joachim; Niemann, Heinrich; Stamminger, Marc (Hrsg.): Vision, Modeling, and Visualization 2005 (10th International Fall Workshop Vision, Modeling, and Visualization Erlangen 16-18.11.2005). Berlin : Aka / IOS Press, Berlin, Amsterdam, 2005, S. 57-64. - ISBN 3-89838-068-8.
80. Zhuravlev, Y.; Gurevitch, I.; Puzankov, D.; Nemirko, A.; Niemann, Heinrich: Special Issue Proceedings of PRIA-7-2004, 7th International Conference on Pattern Recognition and Image Analysis: New Information Technologies, Part 1. In: Pattern Recognition and Image Analysis (PRIA) 15 (2005), Nr. 1, S. 1-271.
81. Zinßer, Timo; Gräßl, Christoph; Niemann, Heinrich: High-Speed Feature Point Tracking. In: Greiner, Günther; Hornegger, Joachim; Niemann, Heinrich; Stamminger, Marc (Hrsg.): Vision Modeling and Visualization (Workshop Vision Modeling and Visualization in Erlangen 2005 Erlangen 16.-18. November 2005). Erlangen : Akademische Verlagsgesellschaft Aka GmbH, Berlin, 2005, S. 49-56. - ISBN 3-89838-068-8.
82. Zinßer, Timo: High-Speed Feature Point Tracking.Vortrag: Konferenz - Workshop Vision Modeling and Visualization, DFG SFB 603 und andere, Erlangen, 16.11.2005.
83. Zinßer, Timo; Schmidt, Jochen; Niemann, Heinrich: Point Set Registration with Integrated Scale Estimation. In: Sadykhov, Rauf; Ablameiko, Serge; Doudkin, Alexander; Podenok, Leonid (Hrsg.): Proceedings of the Eighth International Conference on Pattern Recognition and Image Processing (Eighth International Conference on Pattern Recognition and Image Processing Minsk, Belarus 18.-20. Mai 2005). 2005, S. 116-119. - ISBN 985-6329-55-8.
84. Zinßer, Timo: Point Set Registration with Integrated Scale Estimation.Vortrag: Eighth International Conference on Pattern Recognition and Image Processing, Belarusian State University of Informatics and Radioelectronics, International Association for Pattern Recognition, Minsk, Belarus, 18.05.2005.

## 4 Studien- und Diplomarbeiten, Master Theses

### 4.1 Studienarbeiten

1. Stocker, Thomas: Aufbau einer Dual-Energy Referenzdatenbank.
2. Herrmann, Eva: Automatische Optimierung von Prüfparametern.
3. Kollorz, Eva: Bewegungsschätzung der Koronarien in kardiographischen Bildsequenzen.
4. Gerhard, Friedericke: Development and Validation of a Software Algorithm for Analyzing Laser Speckle Interferometric Images.
5. Heßler, Andre: Entwicklung einer Englisch-Lernsoftware mit integrierter Spracherkennung.
6. Sünkel, Thomas: Evaluation und Optimierung von Lichtfeldern mittels Ansichtenvergleichs aus der Bildsynthese.
7. Mayer, Markus: Klassifikation von Emotionen mit linguistischen Merkmalen.
8. Euring, Matthias: Konzeption und Bewertung einer Kamerakalibrierung.
9. Gagesch, Gerlinde: Merkmalsextraktor für die Klassifikation von Leukozyten.
10. Bock, Rüdiger: Modellierung und Klassifikation von Retinaaufnahmen.
11. Eskofier, Björn: Non-Rigid Image Registration of Myocardial SPECT/CT Datasets.
12. Deuerling-Zheng, Yu: Non-Rigid Registration for Digital Subtraction Angiography in Cardiology.
13. Sickel, Konrad: Rapid 3-D Cone-Beam Reconstruction with the Simultaneous Algebraic Reconstruction Technique (SART) Using 2-D Texture Mapping Hardware.
14. Martin, Sandra: Reconstruction of the Human Ear Canal from Optical Coherence Tomography (OCT) Scans.
15. Maier, Andreas: Robuste Erkennung von verrauschten oder verhallten Daten durch parallele Erkennung in unterschiedlichen Energiebändern.
16. Swadzba, Agnes: Segmentierung des optischen Sehnervenkopfes mittels des Aktiven Form-Modells.
17. Zorn, Dominik: Vergleichende Klassifikation des Benutzerzustands "Müdigkeit" in Sprache.
18. Siller, Christian: Vergleich von Ähnlichkeitsmassen für eine multimodale Bildregistrierung unter Berücksichtigung anatomischer Merkmale.

### 4.2 Diplomarbeiten

1. Dennerlein, Frank: 3-D Bildrekonstruktion aus Kegelstrahl-Projektionsdaten einer Abtastbahn bestehend aus Kreis und Linie.
2. Maier, Andreas: Adaption bestehender Erkenner durch Nachschätzen der akustischen Modelle auf einer kleinen Trainings-Stichprobe des Anwendungsgebiets.
3. Heil, Tobias: Ansichtenplanung bei der aktiven 3-D Rekonstruktion.

4. Hahn, Dieter: A practical saliency region based hierarchical method for aligning 3D medical images.
5. Weber, Marco: Augmented Reality für Stereo-Endoskope.
6. Rehak, Markus: Automatische Segmentierung von 3-D-Computertomographie-Daten zur Charakterisierung von pillierten Rübensamen.
7. Deuerlin-Zheng, Yu: Computer-aided Detection and Evaluation of Brian Lesions in MR Images.
8. Poljak, Maria: Differenzbildgebung und -verarbeitung für die Katheternavigation in der Neuroradiologie.
9. Bauer, Elena: Digital Subtraction Imaging for Renal Perfusion.
10. Payer, Gunnar: Helical Cone-Beam Reconstruction using High-Performance Processors.
11. Kloeters, Christoph: Klassifikation der Modulationsart von analogen Funksignalen.
12. Höning, Florian: Modifications of Perceptual Linear Prediction and the Mel-Frequency Cepstrum.
13. Borsdorf, Anja: Reduktion von Quantenrauschen in CT-Bildern durch Identifikation von Korrelationen.
14. Urzenska, Sylvia: Segmentierung von Metallimplantaten in 2D-Röntgenbildern.

### **4.3 Master Theses**

1. Elter, Matthias: 3D Shape and Appearance Recovery from Image Sequences.
2. Vangara, Jaya Bhargavi: Adaption of ART Algorithms and Comparison with Cone-Beam Reconstruction for 3D C-Arm Imaging.
3. El Allali, Soufyane: Appearance Based Classification of Glaucoma in Retina Images.
4. Schmidt, Achim Florian: Catheter Localization in Coronary Biplane X-ray Imaging.
5. Prause, Andreas: Digital Image Processing and Corneal Surface Reconstruction for Videokeratoscopes.
6. Xie, Weigou: Efficient Restoration of Specular Reflection Areas in Endoscopic Images.
7. Friedl, Sven: Head-Tracking and Pose Estimation in medical applications.
8. Zhang, Peng: High Performance X-Ray Image Processing using Standard Graphics Cards.
9. Qiao, Min: Mono-modal Image Registration for Partial Volume Data.
10. Raghavendra, Eeratta: Quantification of Stenosed Blood Vessels in CTA Data Sets.
11. Dudam, Mithun Kumar: Tomosynthesis in Digital Mammography.