

3D-Poseschätzung und 3D-Mapping mittels PMD-Kamera (3DPoseMap)



Prof. Dr.-Ing.
Reinhard Koch

Multimediale Systeme
zur Informationsverarbeitung (MIP),
Institut für Informatik

Christian-Albrechts-Universität zu Kiel
Olshausenstr. 40
24098 Kiel



Prof. Dr.-Ing.
Hubert Roth

Zentrum für Sensorsysteme (ZESS)
und
Institut für Regelungs-
und Steuerungstechnik
(RST)

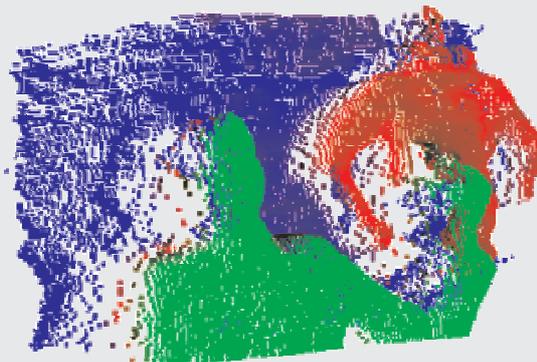
Universität Siegen
Hölderlinstr. 3
57068 Siegen

Antragspaket *Dynamisches 3D Sehen*

Prof. Dr.-Ing.
Andreas Kolb

Zentrum für
Sensorsysteme
(ZESS)

Universität Siegen
Hölderlinstr. 3
57068 Siegen



Prof. Dr.-Ing. habil.
Otmar Loffeld

Zentrum für
Sensorsysteme
(ZESS)

Universität Siegen
Paul-Bonatz-Str. 9-11
57068 Siegen

Inhaltsverzeichnis

1	Allgemeine Angaben	1
1.1	Antragsteller	1
1.2	Thema	1
1.3	Kennwort	1
1.4	Fachgebiet und Arbeitsrichtung	2
1.5	Voraussichtliche Gesamtdauer	2
1.6	Antragszeitraum	2
1.7	Gewünschter Beginn der Förderung	2
1.8	Zusammenfassung	2
2	Stand der Forschung, eigene Vorarbeiten	3
2.1	Stand der Forschung	3
2.1.1	Poseschätzung, Umgebungsmodellierung und SLAM	4
2.1.2	Poseschätzung und Umgebungsmodellierung mittels SfM	4
2.1.3	Binokular-Stereo	5
2.2	Eigene Vorarbeiten	7
2.2.1	Antragsteller 1: Prof. Dr.-Ing. Reinhard Koch, Universität Kiel	7
2.2.2	Antragsteller 2: Prof. Dr.-Ing. Hubert Roth, Universität Siegen	9
3	Ziele und Arbeitsprogramm	13
3.1	Ziele	13
3.2	Arbeitsprogramm	14
3.2.1	PI1: Transfer von SfM auf die AMR-Testplattform MOSRO.	14
3.2.2	PR1: Erweiterung der AMR-Testplattform MOSRO zur Einbindung von SfM.	15
3.2.3	PK1: SfM unter Einbeziehung von 3D-Daten aus einer PMD-Kamera.	16
3.2.4	PK2: Vergleich von 3D-Daten der PMD-Kamera mit 3D-Daten eines Stereokopfes und Einbeziehung in SfM.	17
3.2.5	PI2: Nutzung des 3D-Umgebungsmodells zur Kartengenerierung.	17
3.2.6	PR2: Erstellung einer 3D-Umgebungskarte mithilfe des SfM-Verfahrens und einer 2D/3D-Kamera.	18
3.2.7	PK3: Detektion und Verfolgung starrer bewegter Objekte.	18
3.2.8	PR3: Registrierung der aktuellen Position und Orientierung eines Fahrzeugs anhand der generierten 3D-Umgebungskarte.	19
3.2.9	PI3: Nutzung der erkannten Bewegungen für Kartierung und Navigation.	19
3.2.10	PK4: Detektion und Verfolgung sich deformierender bewegter Objekte.	19
3.2.11	PR4: Navigation eines teilautonomen Fahrzeugs anhand der generierten 3D-Umgebungskarte.	20
3.2.12	PR5: Einbeziehung und Klassifizierung von bewegten Objekten bei der 3D-Umgebungsmodellierung und der Hindernisvermeidung	21
3.2.13	PK5: Nutzung der Objektverfolgung zum Tracken von menschlichen Körpern mit Anwendungen in der Medizin.	21
3.2.14	Zeitliche Einordnung der Aufgaben	21

3.3	Untersuchungen am Menschen	22
3.4	Tierversuche	22
3.5	Gentechnologische Experimente	22
4	Beantragte Mittel	23
4.1	Personalkosten	23
4.2	Wissenschaftliche Geräte	23
4.2.1	Begründung des Bedarfs	23
4.3	Verbrauchsmaterial	24
4.4	Reisen	24
4.5	Publikationskosten	25
4.6	Sonstige Kosten	25
5	Voraussetzung für die Durchführung des Vorhabens	25
5.1	Zusammensetzung der Arbeitsgruppen	27
5.2	Zusammenarbeit mit anderen Wissenschaftlern	27
5.3	Arbeiten im Ausland und Kooperation mit ausländischen Partnern	28
5.4	Apparative Ausstattung	28
5.5	Laufende Mittel für Sachausgaben	29
5.6	Sonstige Voraussetzungen	29
6	Erklärungen	29
6.1	Max-Planck-Institut	29
7	Unterschriften	30
8	Verzeichnis der Anlagen	30

1 Allgemeine Angaben

Antrag auf Gewährung einer Sachbeihilfe
Neuantrag

1.1 Antragsteller

Antragsteller 1:

Reinhard Koch, Dr.-Ing., Universitätsprofessor (C3)
Arbeitsgruppe Multimediale Systeme zur Informationsverarbeitung (MIP)
(www.mip.informatik.uni-kiel.de),
Institut für Informatik, Christian-Albrechts-Universität zu Kiel
geb. 22.9.1958, Deutscher

Dienstadresse:

Olshausenstr. 40

D-24098 Kiel

Tel.: 0431/880-4484

rk@mip.informatik.uni-kiel.de

Privatadresse:

Domänenweg 3

D-24119 Kronshagen

Tel.: 0431/3053736

Antragsteller 2:

Hubert Roth, Dr.-Ing., Universitätsprofessor (C4)
Institut für Regelungs- und Steuerungstechnik (RST) (www.uni-siegen.de/rst),
Universität Siegen
geb. 21.03.1954, Deutscher

Dienstadresse:

Hölderlinstr. 3

D-57068 Siegen

Tel. : 0271/740-4439

hubert.roth@uni-siegen.de

Privatadresse:

Meretsreute 5

D-88273 Fronhofen

Tel.: 07505/9191

1.2 Thema

3D-Poseschätzung und 3D-Mapping mittels PMD-Kamera

1.3 Kennwort

3D-PoseMap

1.4 Fachgebiet und Arbeitsrichtung

Informatik, 3D-Bildverarbeitung

1.5 Voraussichtliche Gesamtdauer

4 Jahre

1.6 Antragszeitraum

24 Monate

1.7 Gewünschter Beginn der Förderung

01.01.2006

1.8 Zusammenfassung

In Echtzeit durchgeführte Positions- und Orientierungsschätzung (=Poseschätzung) einer 2D-Kamera ist Basis für zahlreiche Anwendungen im Bereich der Computer Vision. Für dieses Problem existieren bereits diverse Lösungen, die aber in Genauigkeit und Stabilität der Schätzung für viele Anwendungen nicht ausreichend sind. Zur Stabilisierung des Poseschätzproblems soll hier zusätzlich zu der 2D-Kamera eine neu entwickelte PMD-Kamera eingesetzt werden (s. Abbildung 1). Sie liefert gleichzeitig 2D-Grauwertbilddaten und 3D-Tiefendaten der Umgebung. Die auf diese Weise erzeugten Posedaten sollen genutzt werden, um die eigenständige Registrierung und Navigation eines Fahrzeuges in einer unbekanntem Umgebung zu realisieren. Die Posedaten werden dazu mit den Daten weiterer Sensoren (z.B. Odometer, Inclinometer) fusioniert und es wird eine 3D-Umgebungskarte erstellt, die eine konkrete Pfadplanung ermöglicht.

Simultan soll eine Detektion und Verfolgung sowohl starrer als auch sich deformierender bewegter Objekte realisiert werden. Diese Informationen sollen zur Navigation und Kollisionsvermeidung und im weiteren Verlauf des Projektes zum Motion-Capture, dem Tracken der Gestik von Menschen, genutzt werden.

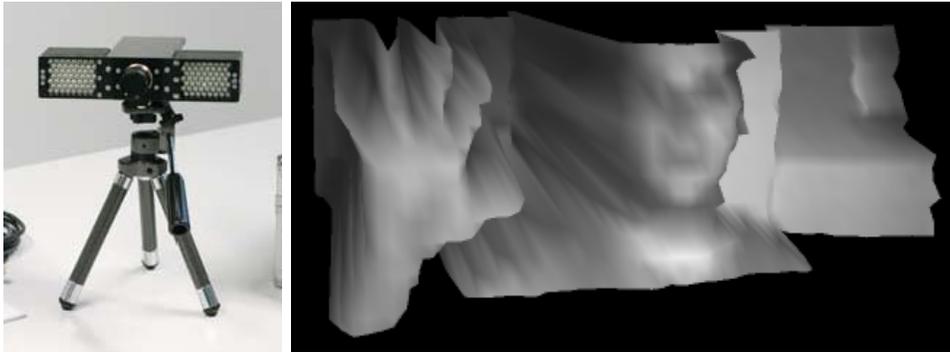


Abbildung 1: Die PMD-Kamera mit zwei symmetrisch angeordneten LED-Matrizen zur modulierten Beleuchtung (links) und deren 3D-Tiefenbild mit 64x16 Pixel (rechts).

2 Stand der Forschung, eigene Vorarbeiten

2.1 Stand der Forschung

Die Anwendungen in Echtzeit durchgeführter Positions- und Orientierungsschätzung (=Poseschätzung) und Umgebungsmodellierung sind vielfältig. Im Automotiv-Bereich stellt sie die Grundlage für autonom arbeitende Sicherheitssysteme eines Fahrzeugs dar. In der Augmented Reality (erweiterten Realität) wird sie genutzt, um Bilddaten mit rechnergenerierten 3D-Modellen passgenau zu überlagern. In der industriellen Handhabungsrobotik und mobilen Robotik ist sie Voraussetzung, um Handlungen und Dienstleistungen autonom ausführen zu können.

Umgebungsmodellierung setzt die Kenntnis der eigenen Pose voraus, auch wenn diese sich stetig durch die Eigenbewegung ändert. Bei der Poseschätzung z.B. eines frei bewegten Fahrzeuges, auch Lokalisierung bzw. Selbstlokalisierung genannt, gilt es, kontinuierlich dessen eigene Pose innerhalb der zu modellierenden Umgebung zu bestimmen.

Die Schätzung der 3D-Pose einer frei bewegten oder auf einem Fahrzeug fixierten Videokamera in 6 statischen $(x, y, z, \phi_x, \phi_y, \phi_z)$ oder dynamischen $(v_x, v_y, v_z, \omega_x, \omega_y, \omega_z)$ Freiheitsgraden (6 DoF) bei gleichzeitiger Umgebungsmodellierung wurde in den letzten Jahren eingehend untersucht und erforscht. Die aktuellen Systeme erzielen eine gute Schätzung der Kamerabewegung über einen längeren Zeitraum, oft sogar in Echtzeit mit bis zu 30 Bildern pro Sekunde [2][3][11].

In der Robotik hat sich hierzu der Begriff SLAM (Simultaneous Localization and Map Building) etabliert, dort wird Poseschätzung mit verschiedenen Sensoren, wie z.B. Odometrie, Ultraschall, 2D-Laserscanner, usw. durchgeführt [21][9][20]. Dieser Ansatz wurde kürzlich auch durch visuelles Tracking erweitert [2]. In der Computer Vision firmiert die Positions-, Orientierungsschätzung und Umgebungsmodellierung aus 2D-Bilddaten seit einigen Jahren unter der Bezeichnung Structure from Motion (SfM)[10]. Auch dieser Ansatz wurde schon früh für die Fahrzeugnavigation und Pfadplanung eingesetzt [1]. Beide Ansätze haben große Gemeinsamkeiten.

2.1.1 Poseschätzung, Umgebungsmodellierung und SLAM

Der Prozess der fortlaufenden Pose- bzw. Positions- und Orientierungsschätzung bei gleichzeitiger Generierung einer Umgebungskarte beschreibt den sog. SLAM Vorgang. Im SLAM-Verfahren wurden bisher eine Vielzahl von Sensoren in Kombination eingesetzt, so z.B. Odometer, Inklinometer, Gyroskop, Inertial-Sensoren, Kompass, 2D-, 3D-Laserscanner, Ultraschall, Radio- und Mikrowellen und nicht zuletzt die 2D-Kamera.

Viele Lösungsansätze arbeiten mit einem 2D-Laserscanner [6][7][5][4] in Verbindung mit Odometrie innerhalb eines 2D-Areals [13][19]. Dieser Laserscanner kann jedoch die zu modellierende Umgebung lediglich in einer einzelnen Ebene erfassen, d.h. es liegen damit nur 2D-Distanzdaten einer bestimmten Ebene der Umgebung vor. Aus der Folge der jeweiligen 2D-Einzelaufnahmen können kontinuierlich die Laserscannerposition und -orientierung errechnet werden. Gleichzeitig wird die jeweilige 2D-Einzelaufnahme datenreduziert und kontinuierlich in eine 2D-Karte eingepasst.

Beim 2D-Laserscanner wird ein einzelner Laserstrahl über einen mechanisch bewegten Spiegel periodisch abgelenkt. Durch zusätzliche mechanische Schwenk- bzw. Nickbewegungen des gesamten Laserscannersystems können mehrere Ebenen der Umgebung erfasst werden, so dass diese nun in 3D-Distanzdaten vorliegt [18]. Aufgrund der mechanischen Ablenkbewegungen benötigt jedoch die Erfassung eines einzelnen 3D-Bildes bei z.B. einem Scanwinkelbereich von $150^\circ \times 90^\circ$ und je nach Auflösung zwischen 4 und 12 Sekunden.

Dennoch konnte mit dieser relativ langsamen Technik erste Ansätze eines sog. 6D-SLAM unter Einbeziehung der o.g. 6 Freiheitsgraden realisiert werden. Auf der Basis des sog. HAYAI(Highspeed And Yet Accurate Indoor/Outdoor-tracking)-Algorithmus [14] wurden die jeweiligen Aufnahmen des 3D-Laserscanners zu einer 3D-Umgebungskarte zusammengesetzt. 6D-SLAM bezeichnet also die Berechnung der 6D-Poseänderung bei gleichzeitiger Generierung eines 3D-Umgebungsmodells.

2.1.2 Poseschätzung und Umgebungsmodellierung mittels SfM

Der aus der Computer-Vision stammende Structure-from-Motion-Ansatz (SfM) beschränkt sich bei der Schätzung der 6 DoF-Pose und der Umgebungsmodellierung auf die Auswertung des Bildstroms einer bewegten 2D-Kamera innerhalb einer starren Szene.

Die Poseschätzung im SfM-Ansatz lässt sich in zwei Schritte unterteilen: das Verfolgen von 2D-Merkmalen im untersuchten Bildstrom sowie die 3D-Geometrieschätzung aus den Informationen über korrespondierende Merkmale in aufeinander folgenden Bildern. Im Folgenden soll in einem kurzen Überblick die Funktionsweise von SfM dargestellt werden, reale Implementierungen gehen oft darüber hinaus, um die benötigte Robustheit und Genauigkeit erzielen zu können.

- **2D-Merkmalverfolgung:** In den Bildern einer Kamera werden gut identifizierbare 3D-Merkmale gesucht und ihre Positionsänderung wird von Bild zu Bild verfolgt, z.B.

mit dem KLT-Operator [16]. Hierbei ist es wichtig, die Merkmale möglichst genau zu lokalisieren, damit die zugrunde liegende Geometrie korrekt geschätzt werden kann.

- **3D-Initialisierung:** In den ersten beiden Bildern werden die 2D-Merkmalsskorrespondenzen dazu genutzt, die relative Pose der aufnehmenden Kamera aus der 2-Bild-Epipolargeometrie zu schätzen. Relativ heißt hier, dass die Orientierungsänderung und die Bewegungsrichtung der Kamera korrekt ermittelt werden und die Entfernung der Kamerazentren die Einheitsentfernung im Koordinatensystem definiert. Alle weiteren Kameraposen können nur relativ zu dem so definierten Koordinatensystem geschätzt werden. Die geschätzten Kameraposen werden genutzt, um 3D-Punkte aus den 2D-Korrespondenzen zu triangulieren.
- **3D-Merkmalschätzung und Positionsbestimmung:** Für alle weiteren Bilder können nun Korrespondenzen aus 2D-Punkten in den Bildern und bekannten 3D-Punkten genutzt werden, um die aktuelle Kamerapose zu bestimmen. Dabei werden die 3D-Punktpositionen immer weiter verbessert. Für verlorene Tracks werden kontinuierlich neue etabliert und neue 3D-Punkte erzeugt. Robustheit wird z.B. durch eine robuste statistische Evaluierung der geschätzten Pose aus den Positionen der verfolgten Merkmale mittels RANdom SAMpling Consensus (RANSAC)[10] erreicht.
- **Einschränkungen und Probleme:** Bei Beschränkung auf rein visuelle Sensorik ist die Referenzierung der Pose in einem absoluten metrischen Koordinatensystem bisher ungelöst, alle Kamerapositionen können ausschließlich mit einer unbekanntem relativen Skalierung bestimmt werden (Maßstabsproblem).

Problematisch ist auch eine auftretende Drift des geschätzten Kamerapfades gegenüber der tatsächlichen Kamerabewegung. Ein wesentlicher Grund für diese Drift ist eine ungenaue Schätzung der Objektkoordinaten in Kamerablickrichtung. Aufgrund der zunächst geringen Kamerabewegung liefert die Punkttriangulierung in dieser Richtung oft große Fehler.

Ein weiterer Grund für diese Drift ist die Annahme, dass alle verfolgten 3D-Raumpunkte „planare Features“ darstellen, d.h. auf einer Ebene liegen und aus allen auftretenden Blickwinkeln eine lediglich affin verzerrte Geometrie aufweisen. Dreidimensionale Ecken können eine ungenaue Positionsschätzung erzeugen, da aus unterschiedlichen Ansichten oft ein leicht verschobener 3D-Raumpunkt verfolgt wird.

Zudem gehen alle diese Systeme von statischen Szenen aus und erzeugen Schätzfehler, sobald sich Teile der Szene bewegen. Die Messungen bewegter Objekte können zwar durch oben erwähnte robuste statistische Verfahren aufwändig als Fehlmessungen klassifiziert werden, Bewegung in größeren Bildregionen beeinträchtigen die Messgenauigkeit aber stark.

2.1.3 Binokular-Stereo

Einige der in Abschnitt 2.1.2 angesprochenen Einschränkungen und Probleme lassen sich lösen, indem 3D-Daten einer Stereotiefenschätzung hinzugezogen werden, um die Poseschätzung zu verbessern. Die stereoskopische Tiefenschätzung versucht, aus den Bildern

zweier zueinander kalibrierter 2D-Kameras eine dichte Tiefenkarte, d.h. für jeden Punkt der Bilder die Tiefe, zu schätzen. Dazu wird i.A. versucht, Korrespondenzen in den Bildern zu finden und diese zur Schätzung zu nutzen. Durch die Kalibrierung wird das Skalierungsproblem gelöst, und die zeitlich synchrone Bildaufnahme erlaubt die Auswertung von nicht starren Szenen.

Da die Algorithmen auf 2D-Bildern arbeiten, haben die Tiefenkarten eine hohe Winkelauflösung und liefern besonders an Helligkeitssprüngen Daten, die im Nahbereich mit nur geringen Fehlern behaftet sind [12].

Problematisch ist aber, dass die stereoskopische Tiefenschätzung nur auf inhomogenen visuellen Strukturen gute Ergebnisse liefert. Homogene oder periodische Strukturen liefern keine oder falsche Korrespondenzen und führen damit zu Fehlern in den Tiefenkarten [8]. Zudem reicht die Geschwindigkeit existierender Algorithmen nur bedingt an die für die Navigation benötigte Echtzeitfähigkeit heran [15] und die von den momentan verfügbaren Echtzeit-Stereosystemen [17] erzeugten Tiefenkarten sind oft nicht dicht genug für eine exakte Umgebungsmodellierung.

Literatur zum Stand der Forschung

- [1] P.A. Beardsley, A. Zisserman, and D.W. Murray. Sequential updating of projective and affine structure from motion. In *International Journal of Computer Vision*, vol. 23, pages 235–259, 1997.
- [2] A. J. Davison. Real-time simultaneous localisation and mapping with a single camera. In *Proceedings International Conference Computer Vision, Nice*, 2003.
- [3] A. J. Davison, Y. González Cid, and N. Kita. Real-time 3D SLAM with wide-angle vision. In *Proc. IFAC Symp. on Intelligent Autonomous Vehicles, Lisbon*, July 2004.
- [4] Laserscanner der Fa. IBEO Automobile Sensor GmbH. Hamburg. www.ibeo-as.de.
- [5] Laserscanner der Fa. Schmersal Holding KG. Wuppertal. www.schmersal.de.
- [6] Laserscanner der Fa. Sick AG. Waldkirch. www.sick.de.
- [7] Laserscanner der Fa. Zoller + Fröhlich GmbH. Wangen. www.zf-laser.com.
- [8] L. Falkenhagen. Depth estimation from stereoscopic image pairs assuming piecewise continuous surfaces. In *Proc. of European Workshop on combined Real and Synthetic Image Processing for Broadcast and Video Production*, Hamburg, Germany, November 1994.
- [9] B. Graf, R. Schraft, and J. Neugebauer. A mobile robot platform for assistance and entertainment. In *Proceedings of ISR*, pages 252–253, 2000.
- [10] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [11] R. Koch, K. Koeser, B. Streckel, and J.-F. Evers-Senne. Markerless image-based 3d tracking for real-time augmented reality applications. In *Proc. WIAMIS*, Montreux, Switzerland, Apr. 2005.
- [12] R. Koch, M. Pollefeys, and L. Van Gool. Multi viewpoint stereo from uncalibrated video sequences. In *Proc. ECCV'98*, volume 1 of LNCS, pages 55 – 71, Freiburg, 1998. Springer-Verlag.
- [13] J. Kollmann and T. Röfer. Kartenaufbau mit einem 180°-laser-entfernungssensor. In *Autonome Mobile Systeme*, pages 121–128. Springer, 2000.
- [14] A. Nüchter, K. Lingemann, J. Hertzberg, H. Surmann, K. Pervözl, M. Hennig, K. R. Tiruchinapalli, R. Worst, and T. Christaller. Mapping of rescue environments with kurt3d. In *Proceedings of the International Workshop on Safety, Security and Rescue Robotics (SSRR '05)*, Kobe, Japan, 2005.
- [15] D. Scharstein, R. Szeliski, and R. Zabih. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In *Proceedings of IEEE Workshop on Stereo and Multi-Baseline Vision*, Kauai, HI, December 2001.

- [16] J. Shi and C. Tomasi. Good features to track. In *Conference on Computer Vision and Pattern Recognition*, pages 593–600, Seattle, June 1994. IEEE.
- [17] Fridtjof Stein. Efficient computation of optical flow using the census transform. In *DAGM-Symposium*, pages 79–86, 2004.
- [18] H. Surmann, K. Lingemann, A. Nüchter, and J. Hertzberg. Aufbau eines 3d-laserscanners für autonome mobile roboter. *GMD Report 126*, 2001.
- [19] H. Surmann, A. Nüchter, and J. Hertzberg. An autonomous mobile robot with a 3d laser range finder for 3d exploration and digitalization of indoor environments. *Robotics and Autonomous Systems*, 45:181–198, 2003.
- [20] T. Rupp T and P. Levi. Globale lokalisation mobiler roboter mit natürlichen landmarken in dynamischen umgebungen. In *Autonome Mobile Systeme*, pages 128–137. Springer, 1999.
- [21] S. Thrun. Robotic mapping: A survey. *Carnegie Mellon University, Technical Report CMU-CS-02-111*, Feb. 2002.

2.2 Eigene Vorarbeiten

2.2.1 Antragsteller 1: Prof. Dr.-Ing. Reinhard Koch, Universität Kiel

Der Antragsteller ist seit 16 Jahren auf dem Gebiet der computerbasierten Bildverarbeitung, 3D-Rekonstruktion und Visualisierung von 3D-Modellen tätig. Ein wesentlicher Forschungsschwerpunkt ist dabei die Echtzeit-Poseschätzung einer frei bewegten Videokamera durch Bestimmung der Kamerabewegung in allen 6 Freiheitsgraden [15][19][18]. Dabei wird gleichzeitig eine grobe Kartierung der Umgebung durch Speicherung und Wiedererkennung markanter Umgebungsmerkmale vorgenommen. Auf diesem Themengebiet wurden umfangreiche Vorarbeiten geleistet, auch im Rahmen von Drittmittelprojekten, die erfolgreich bearbeitet wurden und werden:

- Im EU-Projekt ORIGAMI (IST-2000-28436, 2000-2004) wurden Multikamera-Systeme zum 3D-Tracking und zur 3D-Hintergrundmodellierung mittels SfM und Mehrbild-Stereo entwickelt [1].
- Im EU-Projekt VANGUARD (AC074, 1995-1999) wurden sehr erfolgreich die SfM-Verfahren zur 3D-Umgebungsmodellierung aus frei bewegten, unkalibrierten Kamerabildern entwickelt [19].
- Das EU-Projekt MATRIS (IST-FP6, 2004-2007) entwickelt eine bildbasierte Echtzeit-Kameraverfolgung zur 3D-Überlagerung virtueller Objekte in TV-Studioanwendungen (Augmented Reality) [6].
- Im BMBF-Projekt ARTESAS (2004-2006) wird als Weiterentwicklung des Leitprojektes ARVIKA ein Prototyp zur markerlosen bildbasierten Echtzeitverfolgung für Augmented Reality im industriellen Umfeld entwickelt [2].

In den genannten Projekten wurde der Antragsteller als Spezialist für das Tracking von Kamerabewegungen und der Erstellung von 3D-Umgebungsbeschreibungen herangezogen. Speziell untersuchte Themen sind die Einbeziehung von Inertial-Sensoren, Kameraarrays [12]

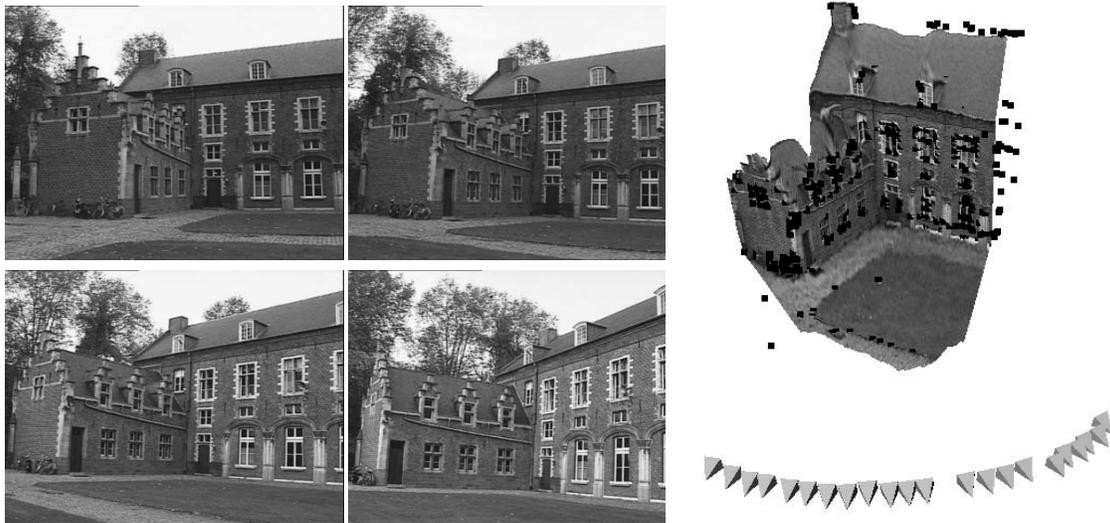


Abbildung 2: Links: 4 von 22 zur Rekonstruktion verwendeten Bildern. Rechts: Das rekonstruierte 3D-Modell mit 3D-Punkten (schwarz) und geschätztem Kamerapfad (Pyramiden).

oder Weitwinkel- bzw. sphärische Kameraobjektive [26][25] zur Stabilisierung des Kameratrackings. Im Rahmen der Projekte wurde von der Arbeitsgruppe eine Anwendung zur Kameraposeschätzung mittels 2D-Kamera und Inertialsensor in Echtzeit (30 Bilder pro Sekunde auf einem 3GHz PC) entwickelt [14]. Sie erzielt bereits eine gute Robustheit und kann Kamerabewegungen über einige Minuten verfolgen und schätzen, allerdings nur relativ zu einer unbekanntem Skalierung (vgl. Abschnitt 2.1).

Weiterhin existieren Vorarbeiten auf dem Gebiet der Stereo-Tiefenschätzung. So werden mittels Stereobildverarbeitung dichte Tiefenkarten generiert. Hierbei wurde besonderes die Echtzeitverarbeitung erforscht, 3D-Tiefenkarten können mittlerweile mit 5 Frames pro Sekunde in einem guten Detaillierungsgrad erzeugt werden [32][31]. Zur Steigerung der Geschwindigkeit und der Qualität der Tiefenkarten werden die 3D-Punkte aus dem Kameratracking als Zusatzinformation genutzt [13][24]. Abbildung 2 zeigt ein aus einer 22 Einzelbildern rekonstruiertes 3D-Modell, dargestellt sind auch die 3D-Punkte und der geschätzte Kamerapfad.

Ein weiterer Forschungsschwerpunkt ist die Detektion unabhängig bewegter Objekte aus dem Bildstrom einer bewegten Kamera, mit dem Ziel der Kollisionsvermeidung. Hierzu wurde im BMBF-Projekt INVENT [3] (als Subcontractor für DaimlerChrysler, 2002-2004) ein System zur automatischen bildbasierten Detektion bewegter Objekte im Kreuzungsbereich entwickelt [29][30][28].

Ebenso wird die Objekt- und Poseerkennung sich bewegender Objekte untersucht. Gegenstand ist sowohl die Bewegung kinematischer Objekte, die in sich starr sind und sich in der Szene bewegen, als auch die Bewegung von Objekten, die sich an ein Modell aus kinematischen Ketten anpassen lassen und sich im Rahmen des Modells deformieren und bewegen. Derzeit wird ein unbewegtes Stereokamerasetup verwendet, um die Tiefe zu schätzen. An die gewonnene Tiefeninformation wird das bekannte Modell eines Menschen angepasst [11][10].

Poseschätzung von Menschen wird oft auch als „Motion Capture“ bezeichnet, Abbildung 3 zeigt die Ergebnisse der Anpassung.

Es ist ein Simulator vorhanden, der es ermöglicht, 2D-3D-Kameradaten zu generieren. Damit können zu Beginn des Projektes einige 2D-3D-Testsequenzen erstellt werden, die ein kalibriertes synchrones Video-PMD-Kamera paar simulieren und für die exakte Posedaten bekannt sind. Diese Sequenzen sollen auch zur Untersuchung der erreichten Qualität der Poseschätzung genutzt werden.

Zudem ist ein Labor vorhanden, in dem großflächige 3D-Projektionen mit 4 Projektoren möglich sind. Hier kann eine dreidimensionale Visualisierung von fusionierten 2D/3D-Daten erfolgen, oder es kann ein beliebiges 3D-Modell interaktiv betrachtet werden [8][9]. Gleichzeitig kann im Labor die Pose der betrachteten Person erfasst werden.

2.2.2 Antragsteller 2: Prof. Dr.-Ing. Hubert Roth, Universität Siegen

Auf dem Gebiet der computerbasierten Bildverarbeitung wird am Institut RST in verschiedenen Bereichen der Robotik gearbeitet.

Medizin-Robotik: Im Rahmen des DFG-Projekts MODICAS [4] dient ein Handhabungsroboter zur Unterstützung bei der Hüftprothesen-Implantation. Mittels Stereo-Kamerasystem und geeigneter Bildverarbeitung gilt es, die genaue Position und Lage eines Patienten in Echtzeit zu bestimmen. Diese Daten werden zur schnellen Nachführungsregelung des Handhabungsroboters entsprechend der Lage des Patienten genutzt.

Autonome Mobile Robotik (AMR): Hierzu wurden verschiedene Roboter aufgebaut bzw. eingekauft, die sowohl im strukturierten 2D- und unstrukturierten 3D-Indoor-Areal,

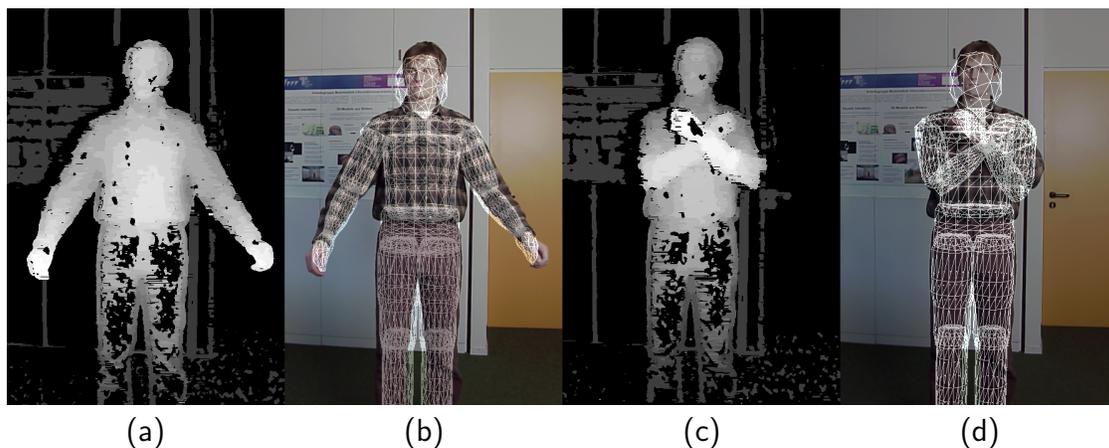


Abbildung 3: Motion Capture in der Arbeitsgruppe Koch. (a)+(c): Stereoskopische Tiefendaten, (b)+(d) Eingabebilder und überlagertes Menschmodell als Drahtgitter.

sowie im ungeordnetem 3D-Outdoor-Areal navigieren können. Mithilfe der Telemetrie sollen AMR ferngesteuert, teilautonom bis vollautonom agieren können.

Im Rahmen des jüngsten EU-Projekt PeLoTe [7][27] wurde ein Rettungssystem entwickelt, dass ein gemeinsames Team von Einsatzkräften und Rettungsrobotern unter der Leitung des Einsatzleiters bei der Suche nach Verletzten unterstützt. Kernthemen des Projekts waren Lokalisierungsmethoden für Mensch und Roboter, teilautonome mobile Roboter, kooperatives Planen, innovative Benutzerschnittstellen und Datenaustausch und -fusion. Im diesem Projekt wurde größten Wert auf die Zusammenarbeit mit potentiellen Nutzern gelegt, wie beispielsweise mit Feuerwehren und anderen Rettungsorganisationen.

AMR-Plattform MOSRO: Am Institut RST wird u. a. der mobile Roboter MOSRO eingesetzt, der von der Fa. Robowatch [16] als Bewachungsroboter konzipiert wurde (s. Abbildung 4(a)). MOSRO wurde als dreirädriges Fahrzeug aufgebaut, mit zwei Antriebsrädern und einem losen Nachlauftrad, und kann damit einfach und beliebig manövriert werden und sich z.B. auf der Stelle drehen. Ausgestattet mit zwei Rad-Encodern, mehreren akustischen und optischen Distanzsensoren, einem Microcontroller C167, einem Mini-Computer PC104, Wireless-LAN, u.v.m. verfügt der MOSRO über ausreichend geeignete Ausstattung, um in diesem Vorhaben als AMR-Plattform für ggf. mehrere PMD- und 2D-Kameras zu dienen. Darüber hinaus steht mit dem MOSRO und seiner Hardware-Architektur auch eine der weltweit größten Bibliotheken für Roboter-Software [20] zur Verfügung, deren Verwendung für dieses Vorhaben geprüft wird. Die Datenkommunikation ausgehend vom Microcontroller C167 vollzieht sich via RS232 zum Mini-Computer PC104 [22] innerhalb des mobilen Roboters. Über den PC104 und Wireless-LAN erhält der mobile Roboter eine LAN-Verbindung zum Internet. Der Benutzer bekommt mit einer Internet-Benutzerschnittstelle die Kontrolle über den mobilen Roboter.

Im frühen Stadium dieses Vorhabens soll zunächst die AMR-Plattform MOSRO zum Einsatz kommen. Später jedoch soll das sich noch im Aufbau befindliches mobiles Fahrzeug MERLIN eingesetzt werden, welches für den Einsatz innerhalb einer 3D-Umgebung besser geeignet ist (s. Abbildung 4(b)).

Standortbestimmung, Hinderniserkennung und -vermeidung mithilfe konventioneller Sensoren: In den am Institut RST aufgebauten und getesteten AMR kommen zahlreiche konventionelle Sensoren zur Standortbestimmung, Hinderniserkennung und -vermeidung zum Einsatz. Es werden Rad-Encoder als Odometer, Gyroskop, Inklinometer und magnetischer Kompass mithilfe von Extended Kalman Filtern (EKF) für die Standortbestimmung verwendet. Des Weiteren werden akustische und optische Distanzsensoren sowie mechanische haptische Berührungssensoren für die Hinderniserkennung verwendet. Für die Auswertung aller Sensordaten und Ansteuerung der Aktorik (zwei Antriebsmotoren) dient der Microcontroller C167. Zur Hindernisvermeidung wurde die Fahrregelung auf der Basis eines Fuzzy-Logik-Reglers realisiert [5][17].

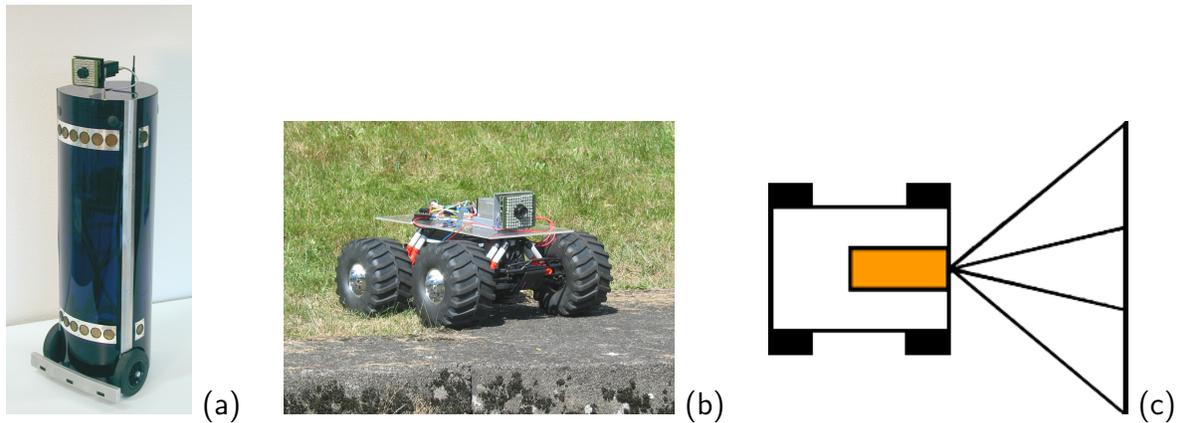


Abbildung 4: (a) AMR-Testplattform MOSRO mit 16×16 PMD-Kamera. (b) AMR-Testplattform MERLIN (Mobile Experimental Robot for Locomotion and Intelligent Navigation) mit 16×16 PMD-Kamera. (c) Mit MERLIN werden Hindernisse im linken, mittleren und rechten Sektor detektiert.

Visuelles 2D-Odometer auf der Basis einer 2D-Kamera: Zur Realisierung dieses Odometers wurde eine am Test-Roboter montierte 2D-Kamera senkrecht auf dessen Fahrweg gerichtet. Innerhalb eines ebenen 2D-Areals konnte mit entsprechender 2D-Bildverarbeitung der Bewegungsvektor des mobilen Roboters kontinuierlich in Echtzeit (20 fps) errechnet werden.

Hinderniserkennung und -vermeidung mithilfe der 16×16 PMD-Kamera: Eine PMD-Kamera wurde in Hauptfahrtrichtung eines Test-Roboters integriert. Mit einer 16×16 Pixel Auflösung liefert diese Kamera ein entsprechendes 16×16 Pixel Abstandprofilbild. Wegen der relativ geringen Auflösung bot sich an, die PMD-Kamera lediglich für die Hinderniserkennung zu benutzen. Daher wurde der Frontbereich des mobilen Roboters bzw. der Sichtbereich in drei Sektoren unterteilt, in 16×6 , 16×4 und 16×6 Pixel für den linken, mittleren und rechten Sektor (s. Abbildung 4(c)). In jedem Sektor wurde stets der jeweils kleinste Distanzwert ermittelt. Die so ermittelten drei Distanzwerte über den gesamten Sichtbereich wurden einer Fuzzy-Logic-Regelung zugeführt, zwecks Realisierung einer Hindernisvermeidung bzw. -umfahrung [23].

Standortbestimmung mithilfe der 16×16 PMD-Kamera anhand definierter Objekte (künstliche Landmarken): Innerhalb einer Roboterumgebung wurden künstliche Landmarken aufgestellt, d.h. bezüglich Geometrie, Position und Ausrichtung definierte Objekte. Mithilfe der auf dem Test-Roboter fixierten PMD-Kamera konnte die jeweilige Landmarke als solche erkannt, die Distanz zwischen Kamera und Objekt und deren perspektivische Ansicht ermittelt werden. Aus diesen Daten ließ sich nun der Standort des mobilen Roboters bestimmen [21].

Literatur zu den eigenen Vorarbeiten

- [1] G. Bazzoni, E. Bianchi, O. Grau, A. Knox, R. Koch, F. Lavagetto, A. Parkinson, F. Pedersini, A. Sarti, G. Thomas, and S. Tubaro. The ORIGAMI Project – advanced tools and techniques for high-end mixing and interaction between real and virtual conten. In *IEEE Proceedings of 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT'02)*, 2002.
- [2] BMBF-Project. Artesas. www.artesas.de.
- [3] BMBF-Project. Invent. www.invent-online.de.
- [4] DFG-Projekt. Modicas (modular interactive computer assisted surgery). www.modicas.de.
- [5] D. Driankov, H. Hellendoorn, and M. Reinfrank. *An Introduction to Fuzzy Control*. Springer-Verlag, 1993.
- [6] EU-Project. Matris (ist-002013). www.ist-matris.org.
- [7] EU-Projekt. Pelote (building presence through localization for hybrid telematic systems). <http://www7.informatik.uni-wuerzburg.de/pelote>.
- [8] Jan-Friso Evers-Senne, Jan-Michael Frahm, Jan Woetzel, Felix Woelk, and Reinhard Koch. Distributed realtime interaction and visualisation system. In *Proc. of 7th International Workshop on Vision, Modeling and Visualization (VMV)*, Erlangen, Germany, Nov. 2002.
- [9] Jan-Michael Frahm, Jan-Friso Evers-Senne, and Reinhard Koch. Distributed interaction processing and visualization of 3d scenes in realtime. In *Proc. of 3rd International Symposium on Image and Signal Processing and Analysis*, Rome, Italy, 2003.
- [10] B. Grest, J.M. Frahm, and R. Koch. A color similarity measure for robust shadow removal in real-time. In *Proc. VMV (to appear)*, Erlangen, Germany, Nov. 2005.
- [11] B. Grest, J. Woetzel, and R. Koch. Nonlinear body pose estimation from depth images. In *Proc. DAGM (to appear)*, Vienna, Austria, Aug. 2005.
- [12] R. Koch, J.M. Frahm, J.-F. Evers-Senne, and J. Woetzel. Pose estimation for multi-camera systems. In *Proc. DAGM*, Tübingen, Germany, Aug. 2004.
- [13] R. Koch, J.M. Frahm, J.F. Evers-Senne, and J. Woetzel. Plenoptic modeling of 3d scenes with a sensor-augmented multi-camera rig. In *Tyrrhenian International Workshop on Digital Communication (IWDC): proceedings*, Sept. 2002.
- [14] R. Koch, K. Koeser, B. Streckel, and J.-F. Evers-Senne. Markerless image-based 3d tracking for real-time augmented reality applications. In *Proc. WIAMIS*, Montreux, Switzerland, Apr. 2005.
- [15] Reinhard Koch and Jan-Michael Frahm. Tutorial: Visual-geometric 3d-scene reconstruction from uncalibrated image sequences. In *DAGM*, Munich, Germany, 2001.
- [16] Roboter MOSRO. Fa. rowatch technologies gmbh. www.rowatch.de.
- [17] Mini-Computer PC104. www.pc104.de.
- [18] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch. Visual modeling with a hand-held camera. *International Journal of Computer Vision*, 59(3):207–232, 2004.
- [19] Marc Pollefeys, Reinhard Koch, and Luc J. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. *International Journal of Computer Vision*, 32(1):7–25, 1999.
- [20] ActivMedia Robotics. Software, documentation and technical support. <http://robots.activmedia.com>.

- [21] H. Roth, R. Schwarte, N. Ruangpayoongsak, J. Kuhle, M. Albrecht, M. Grothof, and H. Heß. 3d vision based on pmd-technology for mobile robots. In *Proceedings SPIE Aerosense conference, Volume „Unmanned Ground Vehicle Technology“*, Orlando, Florida, Jul. 2003.
- [22] O. Rösch and C. Spilca. Autonomous guided vehicle with multiple sensors for virtual laboratory experiments. In *Proc. Smart Systems and Devices*, Hammamet, Tunisia, Mar. 2001.
- [23] Niramom Ruangpayoongsak, Hubert Roth, and Rudolf Schwarte. Position calibration of a mobile robot based on 3d vision. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic, Jul. 2005.
- [24] Jan-Friso Evers Senne, Jan Woetzel, and Reinhard Koch. Modelling and rendering of complex scenes with a multi-camera rig. In *1st European Conference on Visual Media Production (CVMP 2004)*, London, United Kingdom, March 2004.
- [25] B. Streckel, J.-F. Evers-Senne, and R. Koch. Lens model selection for a markerless ar tracking system. In *Proc. ISMAR (to appear)*, Vienna, Austria, Oct. 2005.
- [26] B. Streckel and R. Koch. Lens model selection for visual tracking. In *Proc. DAGM (to appear)*, Vienna, Austria, Aug. 2005.
- [27] Prof. Dr.-Ing. Hubert Roth und Prof. Dr. Klaus Schilling. Fa. ars (angewandte rechner- und softwaretechnologie), steinbeis-transferzentrum. <http://www.stw.de/k060/60030/162.htm>.
- [28] F. Woelk, S. Gehrig, and R. Koch. A monocular collision warning system. In *Proc. CRV*, Victoria, Canada, May 2005.
- [29] F. Woelk and R. Koch. Fast monocular bayesian detection of independently moving objects by a moving observer. In *Proc. DAGM*, Tübingen, Germany, Aug. 2004.
- [30] F. Woelk and R. Koch. Robust monocular detection of independent motion by a moving observer. In *Proc. IWCM*, Günzburg, Germany, Oct. 2004.
- [31] Jan Woetzel and Reinhard Koch. Multi-camera real-time depth estimation with discontinuity handling on pc graphics hardware. In *17th International Conference on Pattern Recognition (ICPR 2004)*, Cambridge, United Kingdom, August 2004.
- [32] Jan Woetzel and Reinhard Koch. Real-time multi-stereo depth estimation on gpu with approximative discontinuity handling. In *1st European Conference on Visual Media Production (CVMP 2004)*, London, United Kingdom, March 2004.

3 Ziele und Arbeitsprogramm

Im Zusammenwirken der verschiedenen Teilprojekte setzt der Teilantrag *3D-PoseMap* die Schätzung der Kamerapose in 6 Freiheitsgraden bei gleichzeitiger Generierung einer Umgebungskarte zur Fahrzeugnavigation um.

3.1 Ziele

Im Rahmen der beantragten Forschungen soll untersucht werden, wie sich durch Einsatz einer 3D-Tiefenkarte die Genauigkeit einer 6DoF-Poseschätzung verbessern lässt und wie sich diese genauere Pose dazu nutzen lässt, exakte 3D-Umgebungsbeschreibungen zu generieren, die als Basis für die Navigation eines Fahrzeugs dienen. Die 3D-Tiefenkarte stammt von einer PMD-Kamera [2] oder den Tiefeninformationen eines Stereokamerakopfes, hier

soll ein Vergleich der Datenqualität stattfinden. In der erstellten 3D-Umgebungskarte soll sich ein Fahrzeug autonom registrieren und navigieren, die Pfadplanung soll interaktiv mit dem Benutzer erfolgen.

Desweiteren sollen simultan mit der Kameraposeschätzung bewegte Objekte detektiert und verfolgt werden. Dies sind zunächst starre, später auch sich deformierende Objekte. Die generierten Objekttrajektorien sollen bei der Umgebungskartierung und zur Navigation genutzt werden, um Kollisionen zu vermeiden. In einem letzten Schritt sollen die Bewegungen von menschlichen Körpern verfolgt werden, mit Anwendungen beim Motion-Capture und in der Medizin.

Erreicht werden sollen diese Ziele durch die enge Zusammenarbeit der Lehrstühle Koch und Roth. Die Erfahrungen der Gruppe Koch in der Poseschätzung ergänzen sich direkt mit denen der Gruppe Roth bei der Datenfusion und der Fahrzeugnavigation auf Basis von Umgebungskarten. Geplant ist, dass die Gruppe Koch Kameraposen und 3D-Umgebungsinformationen und später Trajektorien von Objektbewegungen als Rohdaten generiert. Diese werden von der Gruppe Roth verarbeitet, d.h. sie werden mit den Daten anderer Sensoren geeignet fusioniert und es wird ihnen eine Semantik aufgeprägt, um Umgebungsinformationen in Form einer 3D-Karte zu erhalten, die sich zur autonomen Fahrzeugnavigation eignet.

Im folgenden Abschnitt 3.2 werden die einzelnen Ziele in Form von konkreten Arbeitspaketen detaillierter ausgeführt.

3.2 Arbeitsprogramm

Abbildung 5 stellt den zeitlichen Ablauf und die Interdependenzen der sich aus den Zielen ergebenden Arbeitspakete dar. Die Arbeitspakete PK1-5 und PR1-5 sind jeweils einer der beteiligten Arbeitsgruppen direkt zugewiesen, die Schnittstellenmodule PI1-3 werden von beiden Gruppen gemeinsam erbracht. Viele Arbeitspakete sind noch in einzelne Aktivitäten unterteilt, die gesondert aufgeschlüsselt werden.

3.2.1 PI1: Transfer von SfM auf die AMR-Testplattform MOSRO.

Um die Erfahrung der Arbeitsgruppen Koch auf dem Gebiet der Poseschätzung für die Fahrzeugnavigation zu nutzen, ist es zunächst nötig, die AMR-Plattform MOSRO der Arbeitsgruppe Roth mit der SfM-Software der Arbeitsgruppe Koch auszurüsten, so dass eine stabile Poseschätzung in Echtzeit auf einem mobilen Fahrzeug möglich ist. Um Copyright-Verletzungen beim Gebrauch der SfM-Software zu vermeiden soll diese von Mitarbeitern von Prof. Koch via Netzwerk(WAN, LAN), auf dem mobilen Fahrzeug installiert, konfiguriert, getestet und gewartet werden können.

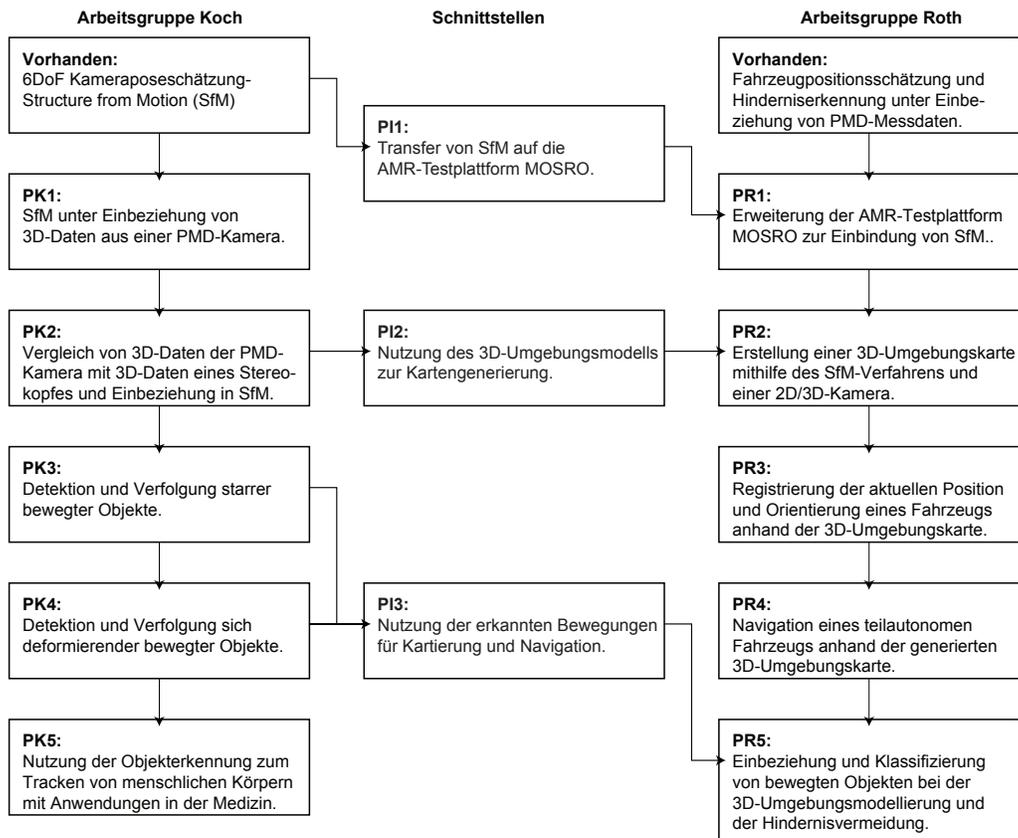


Abbildung 5: Zusammenhang zwischen den einzelnen Arbeitspaketen

3.2.2 PR1: Erweiterung der AMR-Testplattform MOSRO zur Einbindung von SfM.

Das SfM-Verfahren arbeitet in seiner jetzigen Form mit einer einzelnen 2D-Firewire-Kamera und soll in Arbeitspaket PK1 um eine PMD-Kamera und in PK2 um einen Stereokamerakopf, bestehend aus zwei 2D-Firewire-Kameras, erweitert werden. Für die Implementierung des SfM-Verfahrens ist eine ausreichende Anzahl an Firewire-Schnittstellen innerhalb der AMR-Plattform MOSRO vorzusehen. Des Weiteren soll der Zugriff auf die AMR-Plattform MOSRO via Internet, LAN und Wireless-LAN gegeben sein. Für Wireless-LAN wird eine PCMCIA- oder ebenfalls eine USB-Schnittstelle benötigt.

- **PR1.1: Erweiterung der AMR-Plattform MOSRO:** Eine ausreichende Anzahl geeigneter Schnittstellen kann mit einem sog. Mini-PC (z.B. Barebone- oder Embedded-PC mit Pentium 4) kostengünstig bereitgestellt werden, der passend zum mobilen Fahrzeug kleine geometrische Abmessungen besitzt. Ausgestattet mit dem Betriebssystem SUSE-Linux soll dieser Rechner in die AMR-Plattform eingebaut werden.
- **PR1.2: 2D/3D-Videobildserver als Alternative:** Statt der Vorort-Installation der SfM-Software auf dem mobilen Fahrzeug soll auch folgende Alternative getestet wer-

den: Auf dem mobilen Fahrzeug wird ein 2D/3D-Videobildserver installiert, der die Videobilder als Data-Stream anbietet. Damit ist die Möglichkeit gegeben, diesen Videodatenstrom via Wireless-LAN auch auf externe Rechner innerhalb des LAN oder Internet zu leiten. Die Videobilddaten können jetzt ausgehend von der AMR-Plattform auf Rechner der Arbeitsgruppe von Prof. Koch, übertragen und dort mittels SfM-Verfahren ausgewertet werden. Ergebnisse, wie Positions- und Orientierungsdaten, können auf umgekehrtem Weg wieder zurück zum mobilen Fahrzeug zwecks Navigation transferiert werden.

- **PR1.3: Datenfusion:** Die durch das SfM-Verfahren kontinuierlich ermittelten Positions- und Orientierungsdaten sollen nun mit den Daten der bisher verwendeten konventionellen Sensoren mittels der EKF fusioniert werden. Die so errechneten Schätzwerte werden schließlich der bereits bestehenden Fahrregelung zugeführt.

3.2.3 PK1: SfM unter Einbeziehung von 3D-Daten aus einer PMD-Kamera.

Ziel dieses Arbeitspaketes ist es, eine Poseschätzung durch Einbeziehung von 3D-Daten zu erzielen. Ausgegangen wird dabei von dem in der Arbeitsgruppe Koch existierenden Verfahren (s. Abschnitt 2.2.1) zur Schätzung der Kamerapose aus dem Bildstrom einer 2D-Kamera, die 3D-Daten kommen zunächst von der PMD-Kamera. Durch die Einbeziehung der 3D-Daten können folgende Probleme des Standard-SfM-Verfahrens untersucht und verbessert werden:

- **PK1.1: Metrische Skalierung** Durch Hinzunahme von 3D-Daten mit bekannter Skalierung kann die Kameraposeschätzung in einem absoluten metrischen Koordinatensystem durchgeführt werden. Ziel dieser Aktivität ist eine Untersuchung, inwieweit die Fusion von 2D-Video- und 3D-PMD-Daten eine maßstabsgetreue Poseschätzung ermöglicht.
- **PK1.2: Tiefeninformation zur Stabilisierung der 3D-Punkte** Die Genauigkeit der Poseschätzung hängt von den Fehlern der triangulierten 3D-Punkte ab, diese sind in Kamerablickrichtung im Allgemeinen recht groß. Durch Einbeziehung der 3D-PMD-Daten sollen die Punkte in dieser Richtung stabilisiert werden.
- **PK1.3: Oberflächenstruktur der verfolgten Merkmale** Bei der Verfolgung von Merkmalen in der Bildsequenz werden von Bild zu Bild ähnliche Regionen gesucht. Dabei wird davon ausgegangen, dass diese lokal planar und ausschließlich affin verzerrt sind. Ist dies nicht der Fall, kann es zu einer Drift der erkannten Merkmalspunkte kommen. Durch eine lokale Oberflächenmodellierung nahe den Merkmalen kann die Verzerrung geeigneter modelliert werden, dadurch wird eine genauere Lokalisierung der Merkmale möglich.
- **PK1.4: Initialisierung ohne Bewegung der Kamera** Da SfM auf einer Triangulierung vom 3D-Punkten aus unterschiedlichen Kamerabildern basiert, ist immer eine ausreichend große Bewegung zur Initialisierung der Poseschätzung notwendig. Durch Einbeziehung von 3D-Daten kann die Initialisierung ohne jede Bewegung stattfinden.

- **PK1.5: Auswertung und Einordnung der erzielten Verbesserungen** An dieser Stelle ist eine detaillierte Auswertung der erzielten Verbesserungen hinsichtlich der Qualität der geschätzten Pose sowie eine Kovarianzanalyse geplant.

3.2.4 PK2: Vergleich von 3D-Daten der PMD-Kamera mit 3D-Daten eines Stereokopfes und Einbeziehung in SfM.

Im Arbeitspaket PK1 wurde die PMD-Kamera genutzt, um die benötigten 3D-Daten zu generieren. Ebenso kann auch die Stereobildverarbeitung 2.1.3 gute Tiefeninformationen generieren und steht damit in Konkurrenz zur Verarbeitung von PMD-Daten. Daraus ergeben sich folgende Aktivitäten:

- **PK2.1: Vergleich der 3D-Daten aus Stereo und PMD** Hier soll ein Vergleich von 3D-Daten aus Stereobildverarbeitung und PMD-Sensor stattfinden. Die Daten sollen qualitativ verglichen werden und es soll erarbeitet werden, für welche Szenarien sich welches Verfahren eignet. Es wird erwartet, dass die Verfahren sich ergänzen. Während die Stereoalgorithmen hoch aufgelöste Tiefeninformationen in strukturierten Umgebungen liefern, erzeugt die PMD-Kamera besonders zuverlässige aber niedriger aufgelöste Daten in einer homogenen Szene. Unterschiede sind auch bei der Geschwindigkeit der Datengenerierung zu erwarten. Während es möglich ist, auf PMD-Daten in Echtzeit zu rechnen, ist die Stereobildverarbeitung nur eingeschränkt in Echtzeit möglich.
- **PK2.2: Nutzung der Tiefendaten aus der Stereoschätzung für SfM** Um einen Vergleich der 3D-Daten aus PMD und Stereo in ihrer Anwendbarkeit für SfM zu untersuchen, ist eine Einbindung der Stereobildverarbeitung in die Poseschätzung aus PK1 nötig. Es soll damit untersucht werden, in welchen Situationen welche der beiden Tiefenkarten die besseren Ergebnisse für die Poseschätzung liefert.
- **PK2.3: Nutzung der 3D-Informationen aus der Poseschätzung zur Verbesserung der Stereo-Tiefendaten** Während der Poseschätzung werden bekannte 3D-Punkte verfolgt und deren Positionen bei jeder erneuten Sichtung korrigiert. Dadurch kann die 3D-Position von Merkmalen, die über eine längere Zeit stabil erkannt wurden, sehr genau geschätzt werden. Diese 3D-Punkte liefern wertvolle Informationen zur Verbesserung der Stereo-Tiefenkarte.

3.2.5 PI2: Nutzung des 3D-Umgebungsmodells zur Kartengenerierung.

Bei der Schätzung der Kamerapose wird implizit eine 3D-Umgebungsbeschreibung aus einzelnen Raumpunkten ohne jede semantische Information generiert. Diese Umgebungsbeschreibung soll als Grundlage für die Generierung einer Umgebungskarte sowie das Registrieren der Fahrzeugposition in einer bekannten Umgebung herangezogen werden.

3.2.6 PR2: Erstellung einer 3D-Umgebungskarte mithilfe des SfM-Verfahrens und einer 2D/3D-Kamera.

Ein unbekanntes Areal soll als 3D-Umgebungsmodell erfasst werden. Dies ist durch die Nutzung des SfM-Verfahrens unter Verwendung einer PMD-Kamera und einer 2D-Kamera zu realisieren, die fest kombiniert werden. Die Einbeziehung eines Stereokamerakopfes zur Detaillierung der 3D-Daten kann bei Bedarf in schwierigen Situationen erfolgen. Dies hat eine langsamere Datengenerierung zur Folge.

- **PR2.1: 6D-Poseänderung** Die 3D-Umgebungsmodellierung erfolgt durch Bewegung der kombinierten 2D/3D-Kamera und der Stereokamera. Die erste Aufnahme und die momentane Pose der Kameras bilden den Startpunkt. Ab diesem Punkt werden die Kameras in das unbekannte Areal bewegt. Parallel dazu werden kontinuierlich und in definierten Zeitabständen die jeweiligen 2D/3D-Aufnahmen zusammen mit den dazugehörigen Posen in das 3D-Umgebungsmodell überführt. In komplexen Situationen, wenn eine erhöhte Genauigkeit erforderlich ist, kann diese durch Einbeziehung der Stereokamera bei Reduktion der Bewegungsgeschwindigkeit erreicht werden.
- **PR2.2: 6D-SLAM mit einem teilautonomen AMR** Die Bewegung der 2D/3D-Kamera ist Voraussetzung für 6D-SLAM. Die 6D-Kamera- bzw. 6D-Fahrzeugbewegung wird hierbei durch die Selbstmanövrierfähigkeiten des mobilen Fahrzeugs bestimmt. Beispielsweise müssen befahrbare Wege selbständig gefunden, Abgründe und Steigungen selbständig gemieden werden usw. Für diese Fähigkeiten sind Funktionalitäten, wie Hinderniserkennung und -vermeidung, notwendig. In diesem Vorhaben soll die 6D-Kamerabewegung und das 6D-SLAM mithilfe der teilautonomen AMR-Plattform realisiert werden, d.h. das mobile Fahrzeug soll auf einem vom Operator vorgegebenen Pfad entlang fahren. Beim Auftreten von Hindernissen muss das Fahrzeug stets diesen autonom ausweichen und anschließend auf den definierten Pfad zurückkehren und fortfahren. Parallel dazu wird das 6D-SLAM durchgeführt.

3.2.7 PK3: Detektion und Verfolgung starrer bewegter Objekte.

Der nächste Schritt ist die Erkennung von in sich starren, bewegten Objekten und eine Schätzung ihrer Bewegungsrichtung und Geschwindigkeit.

- **PK3.1: Detektion eines bewegten Objektes** In diesem Arbeitsschritt soll für Szenenteile untersucht werden, ob sie ein bewegtes Objekt darstellen. Hierzu bieten sich die 3D-PMD-Daten zur Klassifikation an, da sie für jeden Zeitpunkt eine 3D-Umgebungsbeschreibung liefern. Durch Kompensation der Eigenbewegung kann die Objektbewegung segmentiert werden [1]. Damit ist auch eine Klassifikation der verfolgten Merkmale in bewegte und unbewegte Merkmale möglich, und die Poseschätzung kann weiter stabilisiert werden.
- **PK3.2: Verfolgung eines bewegten Objektes** Nachdem ein bewegtes Objekt segmentiert und die Objektbewegung geschätzt wurde, kann es von Bild zu Bild verfolgt

werden und es kann eine Bewegungsschätzung über einen längeren Zeitraum erfolgen. Daraus soll eine 3D-Bewegungstrajektorie des Objektes erstellt werden.

- **PK3.3: Fehleranalyse** Durchzuführen ist an dieser Stelle wiederum eine detaillierte Kovarianzanalyse, die Information über Fehler in der Objektlokalisierung und -bewegung liefert.

3.2.8 PR3: Registrierung der aktuellen Position und Orientierung eines Fahrzeugs anhand der generierten 3D-Umgebungskarte.

Als Ergebnis des 6D-SLAM aus PR2 erhält man eine 3D-Umgebungskarte des ehemals unbekanntes Areals. Diese Karte soll nun dafür benutzt werden, um die eigene Position und Orientierung einer Person oder eines mobilen Fahrzeugs innerhalb dieses Areals zu bestimmen.

- **PR3.1: 3D-Teilkarte:** Dafür soll wiederum die 2D/3D-Kamera zur Anwendung kommen. Ein zu entwickelnder Algorithmus soll innerhalb der Momentaufnahme dieser Kamera Merkmale und Objekte als natürliche 3D-Land- bzw. Ortsmarken extrahieren. Nach einer Datenreduktion und einer Datentransformation, in ein der 3D-Umgebungskarte entsprechendes Format, wird daraus zunächst eine 3D-Teilkarte dieser Momentaufnahme erzeugt.
- **PR3.2: 3D-Matching-Verfahren:** Anhand der natürlichen Landmarken soll ein geeignetes 3D-Matching-Verfahren einwickelt werden, das die 3D-Teilkarte innerhalb der gesamten 3D-Umgebungskarte findet. Bei erfolgreicher Suche kann nun die Position und Orientierung der 2D/3D-Kamera bzw. der Person oder des mobilen Fahrzeugs errechnet werden. Dieser Prozess soll kontinuierlich und in regelmäßigen Zeitabständen erfolgen.

3.2.9 PI3: Nutzung der erkannten Bewegungen für Kartierung und Navigation.

Die in PK3 segmentierten Bildbereiche, in denen sich bewegte Objekte befinden, und die Trajektorien dieser Objekte sollen für die Kartengenerierung verfügbar gemacht werden.

3.2.10 PK4: Detektion und Verfolgung sich deformierender bewegter Objekte.

Aufbauend auf PK3 sollen in diesem Arbeitspaket auch Bewegungen von Objekten detektiert werden, die sich frei oder nach bestimmten Regeln (kinematischen Ketten) deformieren können. Dabei sollen zunächst Bewegungstrajektorien dieser Objekte bestimmt werden. In einem weiteren Schritt soll auch die Deformation erfasst und auf ein kinematisches Kettenmodell abgebildet werden. Dies alles soll bei gleichzeitiger Kamerabewegung und Kameraposeschätzung stattfinden.

3.2.11 PR4: Navigation eines teilautonomen Fahrzeugs anhand der generierten 3D-Umgebungskarte.

Die o.g. AMR-Plattform ist bereits in eine Netzstruktur eingebunden. Diese verfügt über eine Wireless-LAN- und Internetanbindung. Auf dem Host-Rechner des mobilen Fahrzeugs, über dessen WWW-Server, kann bereits eine internetbasierte Benutzerschnittstelle (s. Abbildung 6) angeboten werden, die ein Operator in seinen Internet-Browser als Java-Applet laden und darin einen virtuellen Pfad zeichnen kann.

Der virtuelle Pfad besteht aus Start-, Endpunkt und vielen Zwischenpunkten, die als gesamtes Datenpaket an das mobile Fahrzeug übergeben werden. Jeder dieser Punkte wird jeweils als Sollposition interpretiert, die der mobile Roboter nacheinander einnehmen soll. Durch den permanenten Vergleich zwischen Ist- und Sollposition sorgt ein bereits entwickelter Regelalgorithmus dafür, dass die AMR-Plattform alle diese Sollpositionen nacheinander abfährt. Das mobile Fahrzeug wird dabei in seiner Geschwindigkeit und Richtung bzw. Orientierung korrigiert, und zwar durch Drehzahlregelung der beiden Antriebsmotoren (Pfadkontrolle, Roboterkontrolle).

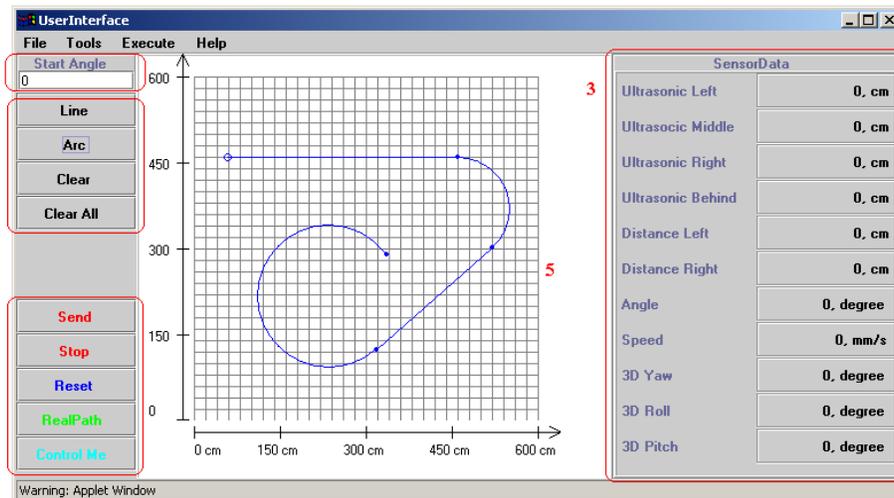


Abbildung 6: Internetbasierte grafische Benutzerschnittstelle.

- **PR4.1: Hinderniserkennung und -vermeidung mit der 2D/3D-Kamera:** Beim Auftreten eines mit der kombinierten 2D/3D-Kamera erfassten festen Hindernisses soll die AMR-Plattform den Pfad verlassen, die Behinderung umrunden und schließlich wieder auf dem gegebenen Pfad fortgefahren (Hinderniserkennung).
- **PR4.2: Pfadplanung:** In diesem Vorhaben soll die o.g. Benutzerschnittstelle dahingehend erweitert werden, dass der Operator die Möglichkeit bekommen, innerhalb dieser Benutzerschnittstelle die zuvor generierte 3D-Umgebungskarte zu laden. Anhand dieser 3D-Umgebungskarte kann der Operator einen beliebigen virtuellen Pfad zeichnen, den das mobile Fahrzeug innerhalb der realen 3D-Umgebung abfahren soll.

3.2.12 PR5: Einbeziehung und Klassifizierung von bewegten Objekten bei der 3D-Umgebungsmodellierung und der Hindernisvermeidung

Bisher wurde die zu modellierende Umgebung als starre Szene angenommen. Dies ist nur für den Indoor-Bereich leicht realisierbar, wenn während der Modellierungsphase keine sich bewegende Objekte, wie z.B. Personen oder Haustiere, innerhalb der Szene auftauchen.

Im Outdoor-Bereich sind diese idealen Bedingungen einer starren Umgebung kaum einzuhalten. Hier muss unterschieden werden zwischen:

- stehenden, jedoch sich verformenden Objekten,
- sich bewegenden, jedoch festen Objekten,
- sich bewegenden und sich verformenden Objekten.

Während stehende und sich verformende Objekte mit in das 3D-Umgebungsmodell Eingang finden sollen, müssen alle bewegte Objekte unberücksichtigt bleiben.

Bei der Pfadplanung können somit stehende und sich verformende Objekte mit berücksichtigt werden. Bei der Navigation eines mobilen Fahrzeugs, also das Fahren entlang des zuvor geplanten Pfads, muss mittels der Hinderniserkennung diese sich bewegenden Objekte auffindig gemacht werden. Die Hindernisvermeidung soll nun dahingehend reagieren, dass beim Entdecken eines solchen Objekts das mobile Fahrzeug seine Fahrt zunächst unterbricht, solange wartet, bis sich das Objekt außerhalb dessen Fahrbahn befindet und dann wieder die Fahrt fortsetzt.

3.2.13 PK5: Nutzung der Objektverfolgung zum Tracken von menschlichen Körpern mit Anwendungen in der Medizin.

In diesem Arbeitspaket soll die in PK4 entwickelte Verfolgung von sich deformierenden bewegten Objekten auf einem realen Anwendungsszenario getestet werden. Dazu soll die Registrierung und Verfolgung von menschlichen Körpern und deren Bewegungen untersucht werden (s. Abbildung 3 Abschnitt 2.2.1). Die Anwendungen einer solchen Verfolgung von menschlichen Bewegungen sind vielfältig. Sie reichen von Motion Capture in der Filmindustrie zur zahlreichen Anwendungen in der Medizin. So wird z.B. mittels Ganganalyse die Diagnose von Krankheiten wie Parkinson oder Schädel-Hirn-Traumata unterstützt, oder der Erfolg von Rehabilitationsmaßnahmen ausgewertet. Eine exakte Erkennung von Menschen wird auch in der Chirurgie zur Registrierung von Körpern benötigt, um Operationsroboter anzusteuern (s. Abschnitt 2.2.2).

3.2.14 Zeitliche Einordnung der Aufgaben

Die Abhängigkeiten der oben genannten Arbeitspakete sind in Abbildung 5 dargestellt. Tabelle 1 zeigt die zeitliche Einordnung der Einzelaufgaben in die verschiedenen Quartale im

Arbeitspakete	2006				2007				2008				2009			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
P-I1 Transfer SfM nach MOSRO	x	x														
P-K1 SfM mit PMD-Kamera	x	x	x													
P-R1 Erweiterung AMR für SfM	x	x	x													
P-K2 Vergleich PMD mit Stereo				x	x	x										
P-I2 Transfer von Umgebungsdaten					x	x										
P-R2 Erstellung der Umgebungskarte				x	x	x										
P-K3 SfM mit bewegten Objekten							x	x	x	x						
P-R3 Fahrzeugregistrierung in Karte							x	x	x							
P-K4 SfM mit deformierbaren bewegten Objekten									x	x	x	x				
P-R4 Fahrzeugnavigaion in Karte									x	x	x					
P-I3 Datentransfer bewegte Objekte									x	x						
P-K5 Motion Capture													x	x	x	x
P-R5 Navigation mit bewegten Objekten													x	x	x	x

Tabelle 1: Zeitliche Einordnung der Arbeitspakete nach Quartalen.

Laufe der Projektbearbeitung.

Literatur zu Zielen und Arbeitsprogramm

- [1] Stefan Heinrich. Real time fusion of motion and stereo using flow/depth constraint for fast obstacle detection. In *Proceedings of the 24th DAGM Symposium on Pattern Recognition*, pages 75–82, London, UK, 2002. Springer-Verlag.
- [2] H. Kraft, J. Frey, T. Moeller, M. Albrecht, M. Grothof, B. Schink, H. Hess, and B. Buxbaum. 3d-camera of high 3d-frame rate, depth-resolution and background light elimination based on improved pmd(photonic mixer device)-technologies. In *OPTO*, Nürnberg, Germany, May 2004.

3.3 Untersuchungen am Menschen

finden nicht statt.

3.4 Tierversuche

finden nicht statt.

3.5 Gentechnologische Experimente

finden nicht statt.

4 Beantragte Mittel

4.1 Personalkosten

1. Zwei wissenschaftliche Mitarbeiterstellen BAT IIa, zunächst für die Dauer von 2 Jahren.
2. Zwei studentische Hilfskräfte mit jeweils 65h/Monat, zunächst für die Dauer von 2 Jahren.

Jeweils ein Doktorand und eine Hilfskraft werden in der Arbeitsgruppe „Multimediale Informationsverarbeitung“ (Prof. Koch) und am Institut für Regelungs- und Steuerungstechnik (Prof. Roth) über die gesamte Projektlaufzeit angestellt. Die Doktoranden werden mit hoher Wahrscheinlichkeit aus den Reihen der Absolventen gewonnen, so dass fundierte praktische Kenntnisse im Arbeitsgebiet bereits zum Zeitpunkt der Anstellung vorhanden sind.

Geschätzter Aufwand im 3. und 4. Förderungsjahr:

1. Zwei wissenschaftliche Mitarbeiterstellen BAT IIa, für die Dauer von 2 Jahren.
2. Zwei studentische Hilfskräfte mit jeweils 65h/Monat, für die Dauer von 2 Jahren.

4.2 Wissenschaftliche Geräte

Im Rahmen des beantragten Projektes kann die vorhandene Infrastruktur der Arbeitsgruppen Koch und Roth genutzt werden, diese ist in Abschnitt 5.4 detailliert aufgeführt. Über diese Grundausstattung beider Gruppen hinaus sind Mittel für folgende Geräte nötig:

1.	1 PMD-Kamera mit der Auflösung 64x48 und Fremdlichtunterdrückung	6.949 EUR
2.	1 PMD-Kamera mit der Auflösung 160x120 ohne Fremdlichtunterdrückung	5.789 EUR
3.	2 Stereo-Kameraköpfe z.B. Bumblebee der Fa. Point Grey	4.000 EUR
3.	1 Firewire-Kamera z.B. Dragonfly der Fa. Point Grey	1.000 EUR
4.	Mini-PC (z.B. Barebone- oder Embedded-PC) und Sonstiges	1.000 EUR
	Summe (inkl. MwSt)	18.738 EUR

4.2.1 Begründung des Bedarfs

1. PMD-Kameras mit ausreichender Auflösung sind in den Arbeitsgruppen Koch und Roth nicht vorhanden und zur Durchführung des Projektes im beantragten Rahmen unerlässlich.

2. Für den Vergleich der PMD-Daten mit Tiefendaten aus der Stereobildverarbeitung ist ein kalibrierter Stereokamerakopf mit fester Kamerakonfiguration an jedem der beiden Lehrstühle nötig. Ein Kamerakopf soll im Verlauf des Projektes auf der mobilen Plattform der Gruppe Roth eingesetzt werden, ein weiterer ist als Referenzgerät für die Arbeitsgruppe Koch nötig. Eine reibungslose Integration von frei beweglichen, zu kalibrierenden Stereokameras auf einer mobilen Plattform ist sehr aufwändig und nicht Untersuchungsgegenstand dieses Projektes.
3. Um das SfM-Verfahren auf der AMR-Plattform der Gruppe Roth zu installieren, ist eine hochwertige CCD-Videokamera nötig. Eine solche ist dort nicht vorhanden.
4. Der Mini-PC wird zur Bereitstellung der Schnittstellen zum Anschluss der Kameras auf der AMR-Plattform MOSRO der Gruppe Roth benötigt.

Geschätzter Aufwand im 3. und 4. Förderungsjahr:

fallen voraussichtlich nicht an.

4.3 Verbrauchsmaterial

Es werden in geringem Umfang Druckkosten pro Jahr anfallen.

1.	Druckaufträge für Poster u.ä.	200 EUR
	Summe pro Jahr	200 EUR

Geschätzter Aufwand im 3. und 4. Förderungsjahr:

2.	Druckaufträge für Poster u.ä.	200 EUR
	Summe pro Jahr	200 EUR

4.4 Reisen

Als jährlichen Reisekosten werden in den ersten beiden Jahren veranschlagt:

1.	Präsentation von Ergebnissen auf nationalen und internationalen Tagungen	
1.1	ICCV oder ECCV (Internationale Computer Vision Tagungen) ¹	4.000 EUR
1.2	DAGM (Deutsche Arbeitsgemeinschaft für Mustererkennung)	2.000 EUR
1.3	IFAC (International Federation of Automatic Control)	2.000 EUR
2.	Arbeitstreffen der Projektpartner	2.000 EUR
	Summe pro Jahr	10.000 EUR

Geschätzter Aufwand im 3. und 4. Förderungsjahr:

1.	Präsentation von Ergebnissen auf nationalen und internationalen Tagungen	
1.1	ICCV oder ECCV	4.000 EUR
1.2	DAGM	2.000 EUR
1.3	IFAC (International Federation of Automatic Control)	2.000 EUR
2.	Arbeitstreffen der Projektpartner	2.000 EUR
	Summe pro Jahr	10.000 EUR

4.5 Publikationskosten

Für Publikationen werden für die ersten beiden Jahre veranschlagt:

1.	Kosten für Farbseiten in Journalen oder Tagungsbänden	750 EUR
	Summe pro Jahr	750 EUR

Geschätzter Aufwand im 3. und 4. Förderungsjahr:

1.	Kosten für Farbseiten in Journalen oder Tagungsbänden	750 EUR
	Summe pro Jahr	750 EUR

4.6 Sonstige Kosten

fallen voraussichtlich nicht an.

5 Voraussetzung für die Durchführung des Vorhabens

Abbildung 7 stellt die wesentlichen Verknüpfungen zwischen dem vorliegenden Projektantrag *3D-PoseMap* und den anderen Teilprojekten des Antragspaktes „Dynamisches 3D-Sehen“ dar. Die Vernetzungen wurden so gestaltet, dass durch direkte Kooperation Synergien zwischen den Teilprojekten genutzt werden können, aber das vorliegende Teilprojekt dennoch unabhängig ist vom Fortschritt in den anderen Teilprojekten. So wird zunächst auf selbst entwickelten 2D/3D-Testsequenzen sowie einer kombinierten 2D/3D-Kamera, bestehend aus einer 2D-Kamera und einer PMD-Kamera, gearbeitet. Damit können 2D/3D-Daten generiert werden, allerdings ohne eine explizite Fehlermodellierung in den Testsequenzen oder eine aufwändige Vorverarbeitung der 2D/3D-Daten.

Durch engen Austausch mit dem Teilprojekten *PMD-Sim*, *MultiCam* und *2D3DProc* wird sichergestellt, dass Ergebnisse aus diesen Teilprojekten genutzt werden können, sobald sie vorliegen. Das heißt, die kombinierte 2D/3D-Kamera wird durch Eingabedaten und Kamerabilder aus diesen Teilprojekten ersetzt bzw. ergänzt. Die Datengenietierung in *MultiCam* und

¹Mittlere Kosten für eine der Konferenzen für zwei Teilnehmer

2D3DProc verbessert direkt die Ergebnisse des vorliegenden Paketes, da die Poseschätzungsalgorithmen von der exakten 2D/3D-Kalibrierung profitieren und eine Datenvorverarbeitung die Qualität der Eingabedaten deutlich steigern kann. Die Daten von *PMD-Sim* versprechen eine erheblich realistischere Modellierung einer PMD-Kamera, als sie eine einfache 2D/3D-Testsequenz erbringen kann. Auch die in *MultiCam* zu entwickelnde omnidirektionale PMD-Kamera soll zur Poseschätzung genutzt werden. Durch eine Erweiterung des Blickwinkels wird die Genauigkeit der errechneten Pose deutlich gesteigert, Erfahrungen mit sphärischen Kameraobjektiven sind in der Arbeitsgruppe Koch bereits vorhanden (s. Abschnitt 2.2.1).

Eine erweiterte Zusammenarbeit mit dem Teilprojekt *MultiCam* ist im Bereich Stereobildverarbeitung und PMD-Kamera vorgesehen. In *MultiCam* wird eine binokulare Kamera zur Stereoschätzung mit einer PMD-Kamera kombiniert, und es ist eine Datenfusion zwischen 3D-Daten aus den Stereobildern mit denen der PMD-Kamera geplant. Hier kann die Arbeitsgruppe Koch Unterstützung leisten, da viel Erfahrung auf dem Gebiet der Stereoschätzung vorhanden ist. Die Ergebnisse des in *3D-PoseMap* durchgeführten Vergleiches sollen zur Optimierung der in *MultiCam* geplanten Datenfusion herangezogen werden. Geplant sind unter anderem Austausch von Know-How, Vergleiche der implementierten Verfahren und Unterstützung bei einer Echtzeitimplementierung. Die fusionierten 3D-Daten sind zudem exzellent als Eingabe für *3D-PoseMap* geeignet.

Die in *3D-PoseMap* errechnete Poseschätzung soll als Eingabe in das Teilprojekt *PMD Lumi* dienen. Zur Erkennung der Lage eines Objektes im Raum wird die Kamerapose und

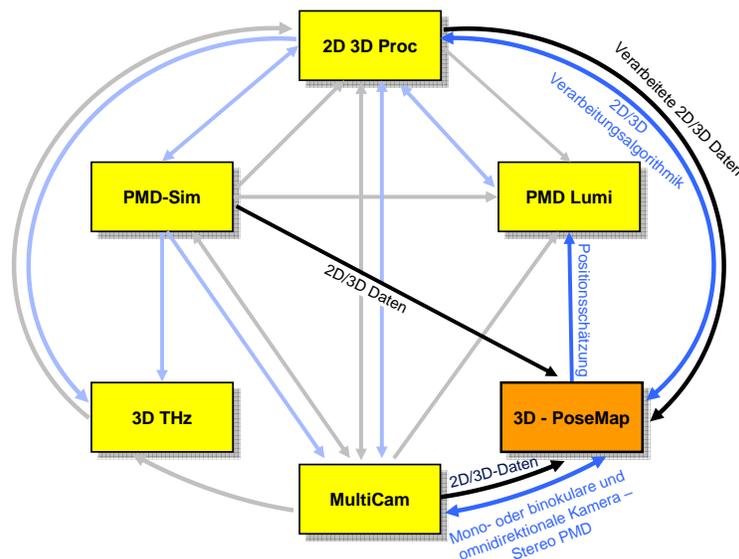


Abbildung 7: Zusammenhänge zwischen dem Teilprojekt *3D-PoseMap* und anderen Teilprojekten des Antragspaketes „Dynamisches 3D-Sehen“. Schwarze Pfeile weisen auf den Austausch von Daten hin, blaue Pfeile stellen den Austausch von Know-How oder anderer Informationen dar.

Dr.-Ing. Bernd Buxbaum: Fa. PMD Technologies GmbH, Siegen

Prof. Dr. rer. nat. Klaus Schilling: Lehrstuhl für Informatik VII, Robotik und Telematik, Universität Würzburg

Priv. Doz. Dr. med. Henning Stolze: Klinik für Neurologie, Universitätsklinikum Schleswig-Holstein, Campus Kiel, Leitung des Ganglors. Im Ganglor der Klinik für Neurologie wird zur Ganganalyse von Probanden markerbasiertes Tracking eingesetzt. Hier soll im Arbeitspaket PK5 getestet werden, inwieweit das Tracking unter Laborbedingungen auf das entwickelte markerlose Verfahren umgestellt werden kann.

5.3 Arbeiten im Ausland und Kooperation mit ausländischen Partnern

Kooperationen besteht mit folgenden ausländischen wissenschaftlichen Einrichtungen:

University of North Carolina, Chapel Hill: Prof. Dr. Marc Pollefeys und Dr. Jan-Michael Frahm.

Microsoft Research, Redmond: Dr. Richard Szeliski („Interactive Visual Media Group“).

Siemens Corporate Research, Princeton: Felix Woelk.

BBC Research, London: Dr. Graham Thomas, Dr. Oliver Grau.

King's College London: Prof. Dr. Althöfer.

National Institute of Standards and Technology (NIST): Prof. Dr. Maris Juberts.

Carnegie Mellon University: Prof. Dr. Melvin Siegel.

5.4 Apparative Ausstattung

Die Arbeitsgruppen Koch und Roth besitzen die räumliche und apparative Ausstattung zur Durchführung des beantragten Projektes. In der Arbeitsgruppe Koch sind insbesondere folgende Geräte vorhanden:

- Arbeitsplatzrechner für Mitarbeiter mit leistungsfähiger Graphik-Hardware.
- Serverrechner mit Dual-Prozessoren.
- Ein Labor mit einer großflächigen 3D-Projektionsleinwand mit 4 Projektoren.
- Mehrere 2D-Videokameras mit perspektivischen und sphärischen Kameraobjektiven.
- Pan-Tilt-Units für die automatische Kameranachführung.
- Inertial-Sensoren zur Bestimmung von 3D-Rotationen.

-bewegung benötigt. Ziel ist es, die in *PMD Lumi* bereits vorhandene markerbasierte Poseschätzung durch das im vorliegenden Antrag zu entwickelnde markerlose Verfahren zu ersetzen.

5.1 Zusammensetzung der Arbeitsgruppen

Die Arbeitsgruppe von Prof. Koch umfasst acht wissenschaftliche Mitarbeiter, diejenige von Prof. Roth besteht aus zwölf wissenschaftlichen Mitarbeitern. Zusätzlich zu den aus den beantragten Mitteln zu finanzierenden Mitarbeitern werden folgende Personen zur Bearbeitung des Projektes beitragen:

1. Reinhard Koch (MIP), Prof. Dr.-Ing., Universitätsprofessor
2. Hubert Roth (ZESS, RST), Prof. Dr.-Ing., Universitätsprofessor, Leiter des Instituts RST und ZESS Projektbereichsleiter des Projektbereichs 4 (PB 4)
3. Bogumil Bartczak (MIP), Dipl.-Ing., Arbeitsgebiet: Stereo-Rekonstruktion
4. Jan-Friso Evers-Senne (MIP), Dipl.-Ing., Arbeitsgebiet: Schnittstellen und Hardwareanbindung
5. Daniel Grest, Dipl.-Ing. (MIP), Arbeitsgebiet: Motion Capture
6. Solomon Kolawole Ige (ZESS), MSc., Arbeitsgebiet: Navigation and Lokalisation
7. Asif Iqbal (ZESS), MSc., Arbeitsgebiet: Telerobotic systems
8. Kevin Koeser, Dipl.-Ing. (MIP), Arbeitsgebiet: Structure from Motion
9. Jörg Kuhle (RST), Dipl.-Ing., Arbeitsgebiet: Autonome mobile Robotik
10. Alexander Prusak (RST), Dipl.-Ing., Arbeitsgebiet: Bildverarbeitung in der Robotik, Servicerobotik
11. Pinyo Puangmali (ZESS), MSc., Arbeitsgebiet: Automatic Control and Estimation Theory
12. Niramom Ruangpayoongsak (RST), MSc., Arbeitsgebiet: Bildverarbeitung in der Robotik, Autonome mobile Robotik
13. Birger Streckel, Dipl.-Inf. (MIP), Arbeitsgebiet: Structure from Motion mit sphärischen Kameras

5.2 Zusammenarbeit mit anderen Wissenschaftlern

Es wird mit den anderen Arbeitsgruppen im Rahmen dieses Paketantrages in enger Abstimmung kooperiert (siehe Präambel „Bistatic Exploration“).

Darüber hinaus bestehen weitere Kooperationen mit Wissenschaftlern dieser Arbeitsgruppen zu anderen Themenbereichen:

- PCs für die Studentenausbildung.

Auch das Institut RST kann auf bereits verfügbare Ausstattung zurückgreifen:

- Telematik- und Roboterlabor mit Wireless-LAN.
- Roboter MOSRO und weitere.
- Microcontroller C166, C167 und entsprechendes Programmierwerkzeug „ μ Vision“.
- 16 \times 16 PMD-Kamera, diverse 2D-Kamerasysteme und internetfähige Video-Server.
- Software „MATLAB“ und „Simulink“ für Regelungstechnik und für Bildverarbeitung.
- CAD-Software „EAGLE“, Schaltungs- und Platinenentwicklung.
- Software „World Tool Kit“, Erstellung dynamischer virtueller Welten innerhalb der Telematik.
- Software „ADAMS“, Modellbildung und Simulation mechanischer Systeme.

5.5 Laufende Mittel für Sachausgaben

Weitere Mittel für Sachausgaben, die über die speziellen Verbrauchsmaterialien unter 4.3 hinausgehen, werden vom jeweiligen Lehrstuhl getragen.

5.6 Sonstige Voraussetzungen

keine

6 Erklärungen

1. Es bestehen keine thematischen Zusammenhänge zwischen dem Vorhaben und Arbeiten eines am Ort befindlichen Sonderforschungsbereiches.
2. Ein Antrag auf Finanzierung dieses Vorhabens wurde bei keiner anderen Stelle eingereicht. Wenn wir einen solchen Antrag stellen, werden wir die Deutsche Forschungsgemeinschaft unverzüglich benachrichtigen.
3. Der Vertrauensdozent der Hochschule wird unverzüglich von dieser Antragstellung unterrichtet.

6.1 Max-Planck-Institut

Nicht anwendbar.

7 Unterschriften

7 Unterschriften



(Prof. Dr. Reinhard Koch)



(Prof. Dr. Hubert Roth)

8 Verzeichnis der Anlagen

1. Tabellarischer Lebenslauf Prof. Dr. Reinhard Koch
2. Tabellarischer Lebenslauf Prof. Dr. Hubert Roth