

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



Prof. Dr.-Ing.
Reinhard Koch

Multimediale
Informations-
verarbeitung
(MIP)

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

Dynamisches 3D Sehen

Prof. Dr.-Ing.
Andreas Kolb

Zentrum für
Sensorsysteme
(ZESS)

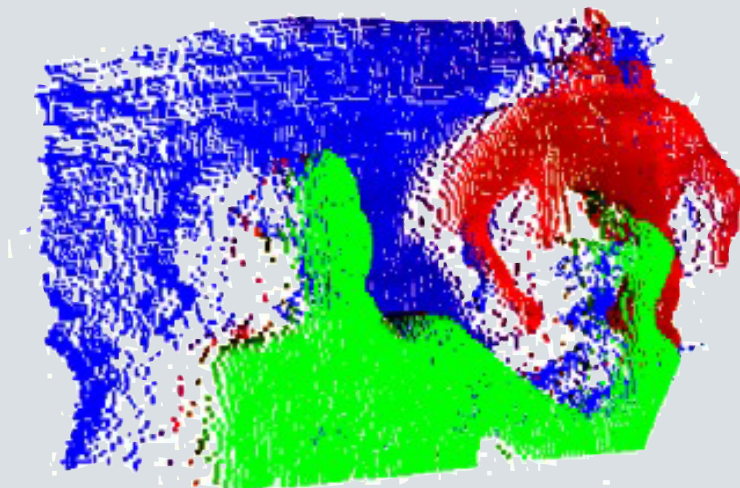
Universität Siegen
Paul-Bonatz-Str. 9-11
57068 Siegen



Prof. Dr.-Ing. habil.
Otmar Loffeld

Zentrum für
Sensorsysteme
(ZESS)

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



Ergebnisbericht

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

Prof. Dr.-Ing. Reinhard Koch

Institut für Informatik
Multimediale Informationsverarbeitung (MIV)
Christian-Albrechts-Universität zu Kiel
Olshausenstr. 40
24098 Kiel

Kiel
2010

Inhaltsverzeichnis

1	Allgemeine Angaben	1
1.1	DFG-Geschäftszeichen	1
1.2	Antragsteller	1
1.3	Institut/Lehrstuhl	1
1.4	DFG-finanzierte wissenschaftliche Mitarbeiter	1
1.5	Thema des Projekts	2
1.6	Berichtszeitraum, Förderungszeitraum insgesamt	2
1.7	Fachgebiet und Arbeitsrichtung	2
1.8	Verwertungsfelder	2
1.9	Am Projekt beteiligte Kooperationspartner	2
2	Zusammenfassung	2
2.1	Wesentliche Ergebnisse und erzielte Fortschritte	2
2.2	Ausblick	3
3	Arbeits- und Ergebnisbericht	4
3.1	Ausgangslage	4
3.2	Beschreibung der durchgeführten Arbeiten	5
3.2.1	Kalibrierung	5
3.2.2	Umgebungs- und Objektmodellierung	5
3.2.3	Objektdetektion und Verfolgung	6
3.2.4	Poseschätzung und Erweiterte Realität	6
3.3	Darstellung der erzielten Ergebnisse	7
3.4	Ausblick auf zukünftige Arbeiten	9
3.5	Interdisziplinäre Weiterentwicklung	10
3.6	Verwertungspotenzial	11
3.7	Beteiligte Wissenschaftler	12
4	Publikationen	13
4.1	Begutachtete Publikationen	13
4.2	Andere Publikationen	15

1 Allgemeine Angaben

Ergebnisbericht im Rahmen des DFG Forschungspakets PAK-73 "Dynamisches 3D Sehen mit PMD", Teilprojekt "3DPoseMap".

1.1 DFG-Geschäftszeichen

KO 2044/3-2

1.2 Antragsteller

Reinhard Koch, Dr.-Ing., Universitätsprofessor (C3)
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

1.3 Institut/Lehrstuhl

Arbeitsgruppe "Multimediale Systeme zur Informationsverarbeitung" (MIP)

Institut für Informatik

Christian-Albrechts-Universität Kiel

Hermann-Rodewald-Str. 3

D-24118 Kiel/Germany

Tel: +49-431 880-4474

Fax: +49-431 880-4845

rk@mip.informatik.uni-kiel.de

www.mip.informatik.uni-kiel.de

1.4 DFG-finanzierte wissenschaftliche Mitarbeiter

Dipl.-Ing. Ingo Schiller: Februar 2008 - August 2009

Dr.-Ing. Jan-Friso Evers-Senne: Februar 2010 - März 2010

Dipl.-Ing. Bogumil Bartczak: Oktober 2010 - Dezember 2010

1.5 Thema des Projekts

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

1.6 Berichtszeitraum, Förderungszeitraum insgesamt

Berichtszeitraum: 01.02.2008 - 31.12.2010

Förderungszeitraum insgesamt (einschließlich KO 2044/3-1): Januar 2006 - Dezember 2010

1.7 Fachgebiet und Arbeitsrichtung

Informatik, 3D-Bildverarbeitung

1.8 Verwertungsfelder

- Selbstlokalisierung und Navigation von Robotern
- Rekonstruktion von Innenräumen
- 3D Objektrekonstruktion und Vermessung
- Inhaltsproduktion für 3D Film und Fernsehen
- Bewegungserfassung des Menschen

1.9 Am Projekt beteiligte Kooperationspartner

- Prof. Dr. Andreas Kolb, Universität Siegen, Deutschland, Lehrstuhl für Computergrafik
- Die Firma PMDTechnologies GmbH aus Siegen, Deutschland war durch Leihstellungen von Kameras und preisreduzierte Kameras ein wichtiger Kooperationspartner.

2 Zusammenfassung

2.1 Wesentliche Ergebnisse und erzielte Fortschritte

Die Ziele des Projekts 3D PoseMap während der ersten Projektlaufzeit waren die Posebestimmung der Time-of-Flight- (ToF-) Kamera mittels Structure-from-Motion (SfM) und die Erkennung und Vermeidung von Hindernissen. Des Weiteren wurde ein Vergleich von Stereo und ToF-Kamera durchgeführt und Ansätze zur Fusion der beiden Messtechniken untersucht. Bei der Entwicklung der Ansätze für die gleichzeitige Navigation und Kartierung der

Umgebung wurde festgestellt, dass eine festmontierte ToF-Kamera mit festem Neigungswinkel ungeeignet für die Kartierung ist, da nur ein ungenügend kleiner Öffnungswinkel zur Verfügung steht. Daher sollte der Ansatz durch die Verwendung einer Schwenk-Neige Einheit zur Umgebungsmodellierung erweitert werden. Diese sollte ebenfalls zur Detektion und Verfolgung von dynamischen Objekten verwendet werden. Durch den Wegfall der Arbeiten der Arbeitsgruppe von Herrn Professor Roth wurde der Schwerpunkt des Projektes verändert. Der zweite Teil des Projektes auf welchen sich dieser Ergebnisbericht bezieht, beschäftigt sich daher weniger mit der Navigation, Kartierung und Hindernisvermeidung auf einer Mobilen Plattform, sondern mit der Poseschätzung im Allgemeinen, der Umgebungsmodellierung und der Objektrekonstruktion sowie mit der Erkennung und Verfolgung dynamischer Objekte. Die Hauptarbeiten in der zweiten Projektphase beinhalten die verbesserte Kalibrierung von Mehrkamarasystemen, bestehend aus ToF- und Standardkameras. Diese Arbeiten wurden in enger Kooperation mit dem Projekt 2D3DProc durchgeführt. Dabei wurden entscheidende Fortschritte erzielt um solche Kamerasysteme über die Projektgrenze hinaus einzusetzen. Die Software wurde öffentlich zugänglich gemacht und wird international eingesetzt.¹

Weiterer Schwerpunkt der Arbeit war die Umgebungsmodellierung unter Verwendung einer Schwenk-Neige Einheit. Damit konnte eine präzise Rekonstruktion realisiert werden, so dass vollständige, geschlossene Modelle von Räumen erstellt werden können, die außerdem für andere Zwecke wie beispielsweise Positionsschätzung Verwendung finden. Zusätzlich können so erstellte Modelle bei der Detektion und Verfolgung dynamischer Objekte eingebunden werden. Neben der Modellierung war die Poseschätzung der Kamera ein wichtiger Punkt der weiter untersucht werden sollte. Neben einem Verfahren dass in der Lage ist die Position der Kamera nur aus den Tiefenbildern der PMD-Kamera zu schätzen wurde ein Verfahren entwickelt um die Position einer ToF-Kamera relativ zu zwei oder mehr festinstallierten CCD-Kameras zu bestimmen.

Die Detektion und Segmentierung von Objekten wurde ebenfalls untersucht. Durch eine kombinierte Verwendung von Tiefen- und Farbinformation konnte die Segmentierung von dynamischen Objekten stark verbessert werden. Die Segmentierung war auch bei der Entwicklung eines Systems für virtuelle Studios und Video-Postproduktion von grosser Bedeutung. Dabei wurde ein System entwickelt das es erlaubt mit einem Kamerasystem aus Farb- und ToF-Kamera Inhalte zu produzieren, in denen reale und künstliche Inhalte so miteinander kombiniert werden dass gegenseitige Verdeckung und Schattenwurf möglich ist. Neben starren Körpern sind sich dynamisch deformierende Objekte von großem Interesse. Deshalb wurde in diesem Projekt außerdem die Eignung von ToF-Kameras für die Aufnahme von Objekten die sich in Form eines kinematischen Kettenmodells deformieren und für die Aufnahme sich frei deformierender Objekte studiert. Es wurden insgesamt 11 Publikationen veröffentlicht.

2.2 Ausblick

Die bisherigen Arbeiten haben die Grundlage zur Kalibrierung und Auswertung von ToF-Daten gelegt und bereits erste Ansätze zur 3D-Umgebungsmodellierung aus Tiefendaten

¹<http://www.mip.informatik.uni-kiel.de/tiki-index.php?page=Calibration>

dargestellt. Die modellierten Szenen waren weitgehend statisch oder beschränkt dynamisch und bewegt. Erste Ansätze zur dynamischen Modellierung sich verformender Oberflächen über Freiformflächendarstellung (NURBS) zeigen vielversprechende Ergebnisse. Zukünftige Arbeiten werden sich auf die Analyse frei deformierbarer Oberflächen konzentrieren und untersuchen, welche Freiheitsgrade bei den Objektdeformationen erlaubt sind und ob intrinsische Deformationsparameter, wie Oberflächensteifigkeit oder Elastizität, aus den Tiefendaten geschätzt werden können.

Ein weiterer wichtiger Forschungsgegenstand wird die Kombination von visueller Merkmalsfindung (z.B. über invariante Deskriptoren wie SIFT) mit tiefenbasierten Merkmalen, wie differentialgeometrischen Invarianten, sein. Hier bietet sich an zu untersuchen, inwiefern durch Verknüpfung von Farb- und Tiefendaten neue robuste Merkmale berechnet werden können. Diese neuen Merkmale könnten insbesondere in der Innenraumnavigation erfolgversprechend eingesetzt werden.

3 Arbeits- und Ergebnisbericht

3.1 Ausgangslage

In den ersten zwei Jahren der Projektlaufzeit konzentrierte sich die Arbeit hauptsächlich auf die Integration der ToF-Kamera in das Structure-from-Motion Verfahren. Dieses Verfahren nutzt im wesentlichen Bildinformation um die Position der Kamera zu bestimmen. Mit der zusätzlichen Verwendung der ToF-Kamera konnte ein systematisches Problem des SfM, nämlich das der Skalierung gelöst werden, da die Messungen der ToF-Kamera metrische Informationen darstellen. Der Vergleich von ToF-Kamera und Stereoverfahren [BBK07b] [BBK07a] erbrachte, dass die ToF-Kameras in mancher Hinsicht den Stereoverfahren überlegen sind. Vor allem in strukturlosen Gebieten bieten die ToF-Kameras verlässliche Messungen was mit Stereoverfahren nicht möglich ist. Allerdings sind ToF-Kameras durch die verwendete Technik reichweitenbegrenzt was in realen Szenen zu Problemen führen kann. Für den zweiten Förderungszeitraum wurde nur ein Teil des Antrags dieses Projektes bewilligt. Der Antrag der Arbeitsgruppe RST der Universität Siegen von Prof. Hubert Roth wurde nicht bewilligt, was die Ausrichtung des Projektes stark beeinflusste. Die ursprüngliche Zielsetzung beinhaltete vor allem die Navigation, die Pfadplanung und die Hindernisvermeidung auf einer mobilen Plattform. Dies sollte durch den zusätzlichen Einsatz einer Schwenk-Neige Einheit unterstützt werden. Die Zielsetzung des Projektes wurde nach der Bewilligung insofern geändert, dass nicht mehr die Navigation im Vordergrund stehen sollte, sondern die Umgebungs- und Objektmodellierung und die Detektion und Verfolgung dynamischer Objekte. Dieses Projekt verfolgte daher in der zweiten Förderphase die Fragestellung inwieweit diese neuartige Technik der ToF-Kameras auf PMD-Basis geeignet sind um in der Kamerapositionsschätzung, der Umgebungs- und Objektmodellierung, der Detektion dynamischer Objekte und der Bewegungserfassung eingesetzt zu werden. Nicht nur die Eignung dieser Kameras, sondern auch die Präzision die damit erreichbar ist sollten dabei bewertet werden. Durch die oben angesprochene Überlegenheit der ToF-Kameras über Stereoverfahren wurde erwartet, dass eine schnelle und präzise Umgebungsmodellierung realisierbar ist. Außerdem

wurde durch die Verfügbarkeit von Tiefeninformationen erwartet, dass eine zuverlässige Detektion von Bewegung und deren Verfolgung möglich sein sollte. Durch den Wegfall einiger zeitintensiver Integrationsarbeiten auf der Roboterplattform wurde außerdem die Untersuchung der Eignung der ToF-Kamera für Kollisionsdetektion und Augmented Reality Anwendungen in die Zielsetzung aufgenommen.

3.2 Beschreibung der durchgeführten Arbeiten

3.2.1 Kalibrierung

Eine wesentliche Voraussetzung für die Verwendung der ToF-Kamera für die präzise Modellierung der Umgebung oder von Objekten aber auch bei der Erfassung von Objektdeformationen ist die präzise Kalibrierung dieser Kameras. In enger Kooperation mit dem Projekt 2D3DProc der Universität Siegen wurde eine Kalibrierung entwickelt die diese Aufgabe erfüllt. Da die ToF-Kamera nicht nur allein sondern im Verbund mit anderen Kameras verwendet werden sollte war die gleichzeitige Schätzung aller relevanten Parameter Teil der Aufgabe. Da die ToF-Kameras neben Tiefenkarten auch Amplitudenbilder bereitstellen, auf denen Intensitätsunterschiede erkennbar sind, konnten Schachbrettmuster zur Kalibrierung herangezogen werden was ein Standardverfahren der Kamerakalibrierung ist. Zur Kalibrierung müssen interne (Brennweite und Bildverzerrung) und externe Kameraparameter (Rotation und Translation) bestimmt werden. ToF-Kameras machen, bedingt durch ihre Funktionsweise, Fehler bei der Messung der Entfernung. Diese Fehler teilen sich in systematische und umgebungsabhängige Fehler auf. Die Kalibrierung der ToF-Kameras schließt deshalb die Modellierung dieser Fehler und die Schätzung der jeweiligen Parameter der Korrekturfunktion mit ein. Im Zuge des Projektes wurde eine Kalibrierung entwickelt und implementiert, die alle relevanten Parameter bestimmt. Dazu werden Aufnahmen eines Schachbretts verwendet und ein Analyse-durch-Synthese-Verfahren zusammen mit einer nichtlinearen Optimierung eingesetzt. Die Parameter der Korrekturfunktion für die Tiefenabweichung werden in einem zweistufigen Verfahren bestimmt. Zuerst werden mithilfe des Schachbretts interne und externe Kameraparameter der ToF-Kamera bestimmt. Daraus ist die Entfernung und die Orientierung des Schachbretts von der Kamera bekannt. Basierend auf der Differenz zwischen der Messung der ToF-Kamera und der tatsächlichen Entfernung werden die Parameter der Korrekturfunktion optimiert. ([LSKK10, SBK08])

3.2.2 Umgebungs- und Objektmodellierung

Die Umgebungsmodellierung hat zum Ziel ein möglichst präzises Abbild der Umgebung zu erstellen. Aufgrund des kleinen Öffnungswinkels und der damit verbundenen Ungenauigkeiten bei der Positionsschätzung der ToF-Kamera sollte ein Verfahren implementiert werden welches eine aktiv gesteuerte Kamera zur Modellierung verwendet. Dazu wurde eine Kombination aus 2D und ToF-Kamera auf eine Schwenk-Neige Einheit montiert und mit dieser ein systematisches Abscannen der Umgebung realisiert. Die dabei gewonnenen Aufnahmen wurden in einem 2.5D Panorama fusioniert, aus welchem wiederum ein vollständiges 3D

Modell erstellt werden konnte. Mithilfe dieses Verfahrens war es möglich eine präzise Umgebungsmodellierung zu realisieren. ([SK09]) Die Objektmodellierung ist mit der Umgebungsmodellierung eng verwandt. Allerdings ist es ein Ziel der Objektmodellierung Aufnahmen von allen Seiten eines Objekts zu fusionieren und daraus ein Modell zu erstellen. Daher ist es nicht möglich das oben verwendete 2.5D Panorama zu verwenden. Stattdessen musste eine andere Form der Repräsentation gefunden werden. Aufgrund der hierarchischen Organisation und dem einfachen Zugriff auf und der einfachen Fusion von Daten wurde dazu ein Octree verwendet. Mithilfe dieser Struktur ist es möglich Daten von allen Ansichten in ein gemeinsames Objekt zu fusionieren, wobei eine Zellgröße festgelegt wird ab der Daten miteinander fusioniert werden. Bei diesem Ansatz ist es nicht mit vertretbarem Aufwand möglich die Kamera aktiv zu steuern, stattdessen muss die Position der Kamera mithilfe der entwickelten SfM-Ansätze bestimmt werden, so dass zu jedem Bild eine Kamerapose vorliegt mit der die Daten in das gemeinsame Modell eingebracht werden. Neben der Modellierung starrer Objekte ist die Modellierung dynamischer Objekte ein Ziel dieses Projektes. Dazu wurde die Octree Datenstruktur erweitert um dynamische Inhalte zu halten und darzustellen. Jedes Element in diesem Octree wurde um Zeitstempel und die Möglichkeit mehrere Repräsentation eines 3D Punktes zu speichern erweitert. Dadurch wird es möglich 3D Szenen als Space-Time-Video aus völlig neuen Perspektiven darzustellen. ([SK09])

3.2.3 Objektdetektion und Verfolgung

Dieser Teil der Arbeit beschäftigte sich mit der Detektion, der Segmentierung und der Verfolgung sich bewogender Objekte. Zur Detektion von dynamischen Vordergrundobjekten wurden verschiedene Verfahren evaluiert. Der einfachste Ansatz ist die Verwendung eines Hintergrundbildes, das sowohl Tiefen- als auch Farbinformation beinhaltet. Die aktuellen Informationen werden dann mittels eines Schwellwertverfahrens verglichen und damit dynamische Vordergrundobjekte detektiert. Ein erweiterter Ansatz ist die Verwendung sog. Adaptiver Hintergrundmodelle wie beispielsweise das Mixture-of-Gaussians. Dieses Verfahren wurde um einen Tiefenkanal erweitert was signifikante Verbesserungen der Detektion erbrachte. Die ToF-Kamera liefert außerdem ein Amplitudenbild. Durch geeignete Kombination von Tiefen-, Farb- und Amplitudenbild konnte die Segmentierung nochmals verbessert werden. Liegt ein Umgebungsmodell vor, können Abweichungen vom Modell detektiert werden, obwohl sich die Kamera in der Szene bewegt. Dazu wird das Modell mit der momentanen Kamerapose auf der Grafikkarte gerendert und die Detektion mittels eines Schwellwertverfahrens durchgeführt. Durch die Verfolgung eines Objektes über die Zeit können sie Bewegungstrajektorien aufgezeichnet werden. ([SBKK10])

3.2.4 Poseschätzung und Erweiterte Realität

Die Schätzung der Position und Orientierung der ToF-Kamera war zentraler Bestandteil der ersten Projektphase. Die Arbeit daran wurde im zweiten Teil des Projektes fortgesetzt. Dabei wurden 3 verschiedene Verfahren entwickelt. Das erste Verfahren nutzt das erstellte Umgebungsmodell und einen Analyse-durch-Synthese Ansatz um die Pose der Kamera zu

bestimmen ([BSBK08]). In einem zweiten Verfahren wurde die Posebestimmung ausschließlich auf den Tiefendaten der ToF-Kamera untersucht ([BSK08b]) und in einem dritten Verfahren wurde die inkrementelle Nachschätzung der Pose einer ToF-Kamera bezüglich mehrerer Standardkameras mittels Gradientenabstiegsverfahren entwickelt ([BSK08a]). Die oben beschriebenen Entwicklungen lassen sich zu einem System für die Generierung und Darstellung von erweiterter Realität kombinieren. Erweiterte Realität beschreibt die Ergänzung natürlicher Aufnahmen durch künstliche Inhalte. Diese müssen abhängig von der Kamera sichtenrichtig in das Kamerabild eingeblendet werden. Dabei muss neben der korrekten gegenseitigen Verdeckung eine natürliche Beleuchtung realisiert werden. Im Zuge dieses Projektes konnte das erste System für erweiterte Realität geschaffen werden, dass in Echtzeit reale und künstliche Inhalte miteinander kombiniert. ([BSBK08, SBKK10])

3.3 Darstellung der erzielten Ergebnisse

Kalibrierung

Die Kalibrierung war die Voraussetzung um die anderen Arbeiten in diesem Projekt durchführen zu können. Dabei ist vor allem die gemeinsame Arbeit an diesem Teil mit Marvin Lindner und Andreas Koch von der Arbeitsgruppe CG der Universität Siegen hervorzuheben. Die systematischen Tiefenfehler wurden durch eine BSpline-Funktion approximiert und minimiert. Damit kann über den ganzen Einsatzbereich der ToF-Kamera von einer korrekten Tiefe ausgegangen werden. Details zu dieser Arbeit finden sich in [SBK08, LSKK10].

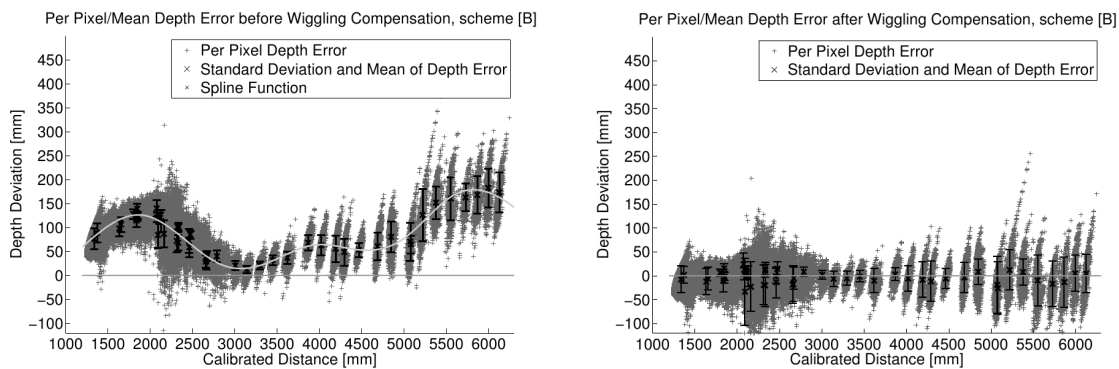


Abbildung 1: Tiefenfehler vor (links) und nach (rechts) der Korrektur durch die geschätzten Parameter. Einzelne Pixelfehler sind in grau, mittlerer Fehler und Standardabweichung in schwarz und die B-Spline-Fehlerfunktion ist in hellgrau dargestellt.

Umgebungs- und Objektmodellierung

Die Umgebungsmodellierung mittels einer Schwenk-Neige Einheit erlaubte die Erzeugung sehr genauer Umgebungsmodelle basierend auf 2.5D Panoramabildern. Diese wurden aus bis zu 100 Aufnahmen zusammengesetzt und Pixel die dabei mehrmals aufgenommen wurden, wurden durch Mittelung der Werte bestimmt. Abbildung 2 zeigt links das gemittelte Tiefenpanorama. Aus diesen Panoramen konnte dann mittels Triangulierung ein 3D Modell der Umgebung erstellt werden welches in Abbildung 2 rechts gezeigt ist.

Die Rekonstruktion der Umgebung konnte mit guter Genauigkeit realisiert werden, bei einer

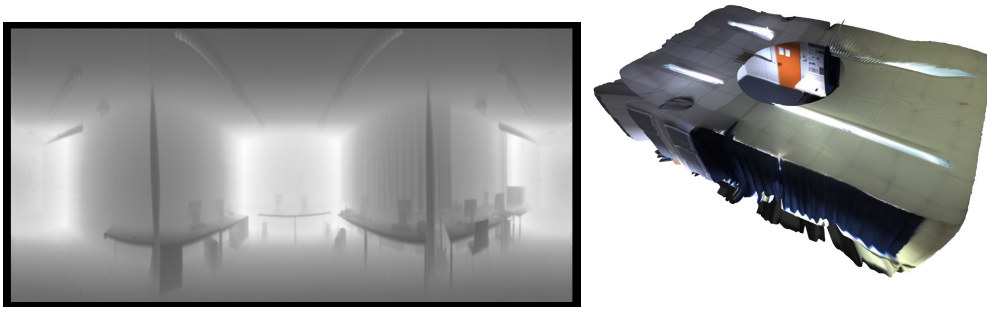


Abbildung 2: 2.5D Tiefenpanorama (links) und komplettes Modell des Dreiecksnetzes, generiert aus den Panoramabildern, in der Vogelperspektive (rechts).

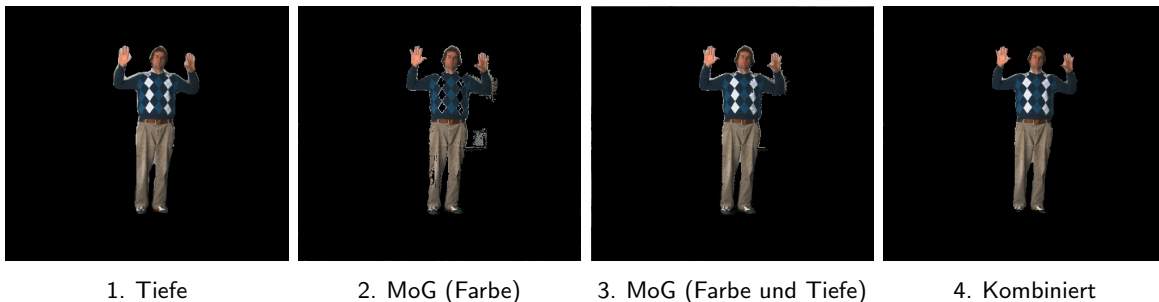


Abbildung 3: Segmentierte Bilder. Das vierte Verfahren verwendet MoG auf Farbinformation und Tiefenschwellwerte, kombiniert mit einer Gewichtungsfunktion die aus den Tiefendiskontinuitäten der Tiefenkarte errechnet wird.

Raumgröße von $8,5\text{m} \times 3\text{m} \times 5,5\text{m}$ betrug der mittlere Fehler lediglich 44mm-177mm, was 1,25%-2,03% entspricht. Die Modellierung einzelner starrer und dynamischer Objekte war ebenfalls Teil des Projektes. Durch die oben beschriebenen Verfahren gelang es, sowohl starre als auch sich dynamisch verformende Objekte zu rekonstruieren. Details wurden in [SK09] beschrieben.

Objektdetektion und Verfolgung

Zum Zweck der Objektverfolgung werden hier 4 verschiedene Verfahren verglichen die im Projekt entwickelt wurden. In Abbildung 3 werden diese Ansätze verglichen. Ansatz 1 ist dabei tiefenbasierte Schwellwertsegmentierung, Ansatz 2 farbbasiertes Mixture-of-Gaussians, Ansatz 3 farb- und tiefenbasiertes Mixture-of-Gaussians und Ansatz 4 eine gewichtete Kombination von Ansatz 1 und 3, bei der Informationen über Tiefendiskontinuitäten in die Gewichtung einfließen. Alternativ wurde hier das Amplitudenbild der ToF-Kamera verwendet, was die Ergebnisse nochmals leicht verbesserte.

Bild 4 in Abbildung 3 zeigt, dass durch eine geeignete Kombination von Tiefen- und Farbsegmentierung eine deutliche Verbesserung der Segmentierung dynamischer Objekte erreichbar ist. In [SBKK10] wurde die tiefenbasierte Segmentierung verwendet um eine Person in Bildsequenzen zu isolieren und zu verfolgen. Diese Information wurde dann verwendet um

zusätzliche Objekte in der Szene zu platzieren.

Poseschätzung und Erweiterte Realität

In [BSBK08] und [SBKK10] wurde die modellbasierte Kameraposeschätzung verwendet um die Position der sich bewegenden Kamera zu ermitteln. Außerdem wurde die oben genannte Tiefensegmentierung verwendet um zusätzliche Inhalte in die Szene einzubringen (siehe Abbildung 4). Die Verwendung der ToF-Kamera erlaubte diese Entwicklung, was gegenüber dem Stand der Technik vor dem Projekt eine entscheidende Weiterentwicklung darstellt. Davor konnten solche Effekte nur durch komplexe Studioaufbauten und manuelle Nachbearbeitung erzielt werden.



Abbildung 4: Eingabebilder und komponierte Ausgabebilder des Systems zur Erzeugung von Erweiterter Realität.

3.4 Ausblick auf zukünftige Arbeiten

Eine Erweiterung des Projektes, der angefangen aber noch nicht abgeschlossen werden konnte stellt das markerlose Human Motion Capture (Bewegungserfassung von Menschen) dar. Menschen können als artikulierte Objekte verstanden werden, die durch ein kinematisches Kettenmodell beschrieben werden können. In einer Diplomarbeit [Ang10] wurde die Basis für die Schätzung des Parameter eines Menschmodells gelegt. Darin enthalten ist auch die Kalibrierung von mehreren ToF-Kameras in einem Kameraverbund mittels eines mehrdimensionalen Kalibrierobjekts. Diese Schätzung von menschlichen Bewegungen bietet vielfältige Anwendungen in der Medizin und für Filmproduktionen.

Bei der Verwendung der ToF-Kamera konnte verschiedene Effekte beobachtet werden, die teilweise auf Materialeigenschaften zurückzuführen sind. Beispielsweise sorgen bestimmte organische Materialien dafür, dass anstatt einer Erhebung eine Vertiefung gemessen wird.

Eine systematische Untersuchung dieser Effekte und ein Verfahren zur Kompensation dieser Effekte würde den Einsatzbereich von ToF-Kameras entscheidend erweitern.

Die Navigation und die Registrierung einer Kamera in einer bekannten Umgebung sind viel untersuchte Verfahren in der Computer Vision. Im Laufe des Projektes konnte eine verbesserte Poseschätzung mit ToF-Kameras entwickelt werden. Die ToF-Kamera wurde als Tiefensensor verwendet um Tiefe für einzelne Punkte zu messen. Eine logische Weiterentwicklung ist die Verwendung von geometrischen Merkmalen in Form von 3D Features in Verbindung mit 2D Merkmalen um die Deskriptivität zu erhöhen und somit die Lokalisierung zu vereinfachen. In Verbindung mit schnell durchsuchbaren Baumstrukturen zur Organisation der Features soll hier eine zuverlässigere Navigation und Selbstlokalisierung entwickelt werden.

3.5 Interdisziplinäre Weiterentwicklung

Die im Projekt untersuchte Laufzeitkamera zeichnet sich durch die Fähigkeit aus verlässliche Entfernungsinformation mit hoher Frequenz aufnehmen zu können. Im Projekt konnte vielfältiges Wissen über mögliche Anwendungen und Beschränkungen erworben werden. Durch die geleisteten Arbeiten im Projekt kann diese Kameratechnik nun auch für andere Zwecke eingesetzt werden.

- **Tierzucht und Verwertung**

Die Aufzucht und Verwertung von Tierprodukten ist ein Fachgebiet in dem ein großes Automatisierungspotential vorhanden ist. Außerdem werden beständig Verfahren gesucht um bestimmte Untersuchungen, wie die Beurteilung der körperlichen Kondition etwa bei Kühen automatisch zu erfassen. Diese körperliche Verfassung lässt sich an bestimmten Merkmalen ablesen. Es konnte in einer Kooperation mit der Agrar- und Ernährungswissenschaftlichen Fakultät der Universität Kiel nachgewiesen werden, dass Laufzeitkameratechnik die erforderliche Genauigkeit liefert um solch eine Klassifizierung durchzuführen.

Ein weiterer Bereich in dem Laufzeitkameras aufgrund unserer Erfahrungen und Empfehlungen eingesetzt und untersucht werden ist die automatisierte Verarbeitung organischer Materialien wie beispielsweise Fleisch. Aufgrund der großen Variabilität der Konsistenz von Fleisch ist es bisher nur sehr eingeschränkt möglich diese Verarbeitung zu automatisieren. Mithilfe der Laufzeitkamera kann in Echtzeit die Deformation der Oberfläche erfasst werden und damit eine Prädiktion der Deformation erfolgen. Dies wird verwendet um die automatische Verarbeitung von Fleisch durch einen Roboter zu steuern. (Zusammenarbeit mit den Mitarbeitern im Projekt IRFO ²)

- **3D Fernsehen und Produktion von Medieninhalten**

Die neueste Generation von 3D Fernsehgeräten ist in der Lage 3D autostereoskopisch darzustellen, d.h. ein Betrachter kann ohne eine Brille zu tragen dargestellte Inhalte räumlich wahrnehmen. Dazu ist es nötig von jedem Punkt in der Szene die Entfernung

²<http://www.interreg-robot.eu>

zu kennen um die entsprechenden Ansichten der Szene erzeugen zu können. Aufgrund der positiven Erfahrungen mit der Laufzeitkameratechnik im Projekt 3DPoseMap wurde diese Technik auch bei der Aufnahme von 3D Inhalten für die Medienproduktion evaluiert. Es konnte unter Verwendung der erarbeiteten Verfahren gezeigt werden dass diese Technik die nötigen Fähigkeiten besitzt und konkurrierende Techniken übertrifft. Die verfolgten Ansätze werden im Journal "Transactions on Broadcasting" (IEEE) publiziert, eine entsprechende Einreichung wurde zur Veröffentlichung akzeptiert. (Zusammenarbeit mit den Mitarbeitern im Projekt EU-3D4YOU ³)

- **Archäologie, Digitalisierung von Fundstücken**

Eine weitere Fachdisziplin die von der Laufzeitkameratechnik und den Ergebnissen dieses Projektes profitieren könnte ist die der Archäologie, z.B. bei der Konservierung und Darstellung von Ausgrabungen und bei der Digitalisierung von Fundstücken. Momentan werden Ausgrabungsstätten durch Fotografien und Karten dokumentiert. Vermessungen werden durch Laserverfahren punktweise durchgeführt. Es existieren teilweise schon sehr erfolgreiche Ansätze eine automatisierte 3D Digitalisierung durch Methoden des Structure-from-Motion zu realisieren. Die Laufzeitkameratechnik bietet hier die Möglichkeit diese Verfahren zu ergänzen oder zu ersetzen. Gerade in der Archäologie ist eine metrische Aufnahme der Daten sehr wichtig was durch die Laufzeitkameratechnik gegeben ist.

Ein weiterer Schwerpunkt ist die Digitalisierung von Fundstücken um sie zu vermessen und darzustellen. Es konnte im Projekt gezeigt werden dass durch Schätzung der Kamerapose und Fusion der Messungen der Laufzeitkamera eine solche Rekonstruktion von Objekten möglich ist.

3.6 Verwertungspotenzial

Die Laufzeitkameratechnik bietet aufgrund ihrer Eigenschaften vielfältige Anwendungsmöglichkeiten. Durch die im Projekt erzielten Ergebnisse bezüglich Kalibrierung, Poseschätzung und Segmentierung von Vordergrundobjekten eröffnen sich vielversprechende Möglichkeiten der Verwertung der Resultate.

Die Verwertung der Ergebnisse aus dem Projekt 3DPoseMap erfolgte in anschließenden und teils auf parallel laufenden Projekten. Im EU-Projekt 3D4YOU, welches sich mit der Entwicklung einer Aufnahme- und Verbreitungskette für 3D Fernsehen beschäftigte, wurden nach den erfolgversprechenden Ergebnissen im Projekt 3DPoseMap Laufzeitkameras zur Aufnahme von 3D Informationen angewandt. Dabei wurde stark auf die Ergebnisse von 3DPoseMap zurückgegriffen, vor allem in Bezug auf Kalibrierung und Vorverarbeitung der Mehrkamerasysteme und Entfernungsdaten.

Die Entwicklungen im Bereich 3D Fernsehen erfordern eine neue Herangehensweise bei der Produktion der Inhalte. Wie bereits im Abschnitt 3.5 erläutert bieten die Laufzeitkameras die nötige Zusatzinformation um diesen neuen Herausforderungen zu begegnen. Mithilfe der Laufzeitkameras, einer geeigneten Kalibrierung und weiteren Verfahren kann jedem Pixel

³<http://www.3d4you.eu>

Tiefe zugeordnet werden und Verdeckungsinformation kann berechnet werden. Dies eröffnet neben der 3D Darstellung auch die Möglichkeit die Aufnahmen um künstliche Information zu ergänzen. Dabei kann es sich um Logos oder beliebige Objekte handeln. Eine entsprechende Ideenskizze wurde beim Ideenwettbewerb 2010 der WTSH Schleswig-Holstein GmbH mit dem 2. Platz ausgezeichnet. Die Gründung eines Unternehmens zur Vermarktung der Ergebnisse wird derzeit geprüft.

Auch andere Projekte konnten stark vom Projekt 3DPoseMap profitieren. Durch das Einfließen der Ergebnisse in das Projekt IRFO (INTERREG) konnten Fortschritte viel schneller erreicht werden. Dadurch ergaben sich Kooperationen mit anderen Fachrichtungen, wie beispielsweise der Agrar- und Ernährungswissenschaftlichen Fakultät der Universität Kiel.

Ein weiterer Interessanter Bereich in dem die erreichten Ergebnisse Verwertung finden könnten ist der Bereich der Überwachung von Innenräumen. Laufzeitkameras eignen sich hierfür, da sie nicht auf Farbinformationen sondern auf Tiefeninformationen basieren und auch im Dunkeln funktionieren.

Das Bereich der Bewegungserfassung (Motion Capture) ist ebenfalls ein Bereich in dem Verwertung stattfinden könnte. Durch die direkte Aufnahme von Tiefe können Bewegungen besser zugeordnet werden was beispielsweise in der Filmindustrie oder in der Bewegungsanalyse im medizinischen Sektor Verwendung finden könnte.

In Kooperation mit Prof. Andreas Kolb der Universität Siegen ist 2011 ein Workshop über Laufzeitkameras auf Schloss Dagstuhl⁴ (Leibniz-Zentrum für Informatik) geplant, bei dem das erworbene umfassende Wissen auf diesem Gebiet an interessierte Wissenschaftler weitergegeben werden soll.

3.7 Beteiligte Wissenschaftler

- **Prof. Dr.-Ing. Reinhard Koch** als Projektleiter.
- **Dipl.-Ing. Ingo Schiller** beschäftigte sich während der ganzen Laufzeit des Projektes mit dessen verschiedenen Themen. Er war in der Entwicklung der Kalibrierung, der Poseschätzung, der Umgebungsmodellierung, der Bewegungserfassung und -wiedergabe und der Anwendung in der Erweiterten Realität maßgeblich beteiligt.
- **Dr.-Ing. Christian Beder** war an der Entwicklung der Kalibrierung und der Poseschätzung beteiligt.
- **Dipl.-Ing. Bogumil Bartczak** leistet Beiträge zum System der Erweiterten Realität.
- **Dipl.-Inf. Falko Kellner** leistet Beiträge zum System der Erweiterten Realität und der Umgebungsmodellierung.

⁴<http://www.dagstuhl.de>

4 Publikationen

4.1 Begutachtete Publikationen

In dieser zweiten Projekthälfte konnten 11 Publikationen über die erzielten Fortschritte der Arbeiten erfolgreich publiziert werden. Die wichtigsten Publikationen werden hier mit einer Zusammenfassung vorgestellt. Im Literaturverzeichnis sind die weiteren Publikationen gelistet.

1. "High-Resolution Object Deformation Reconstruction with Active Range Camera" [JSBK10]

Zusammenfassung:

This contribution discusses the 3D reconstruction of deformable freeform surfaces with high spatial and temporal resolution. These are conflicting requirements, since high-resolution surface scanners typically cannot achieve high temporal resolution, while high-speed range cameras like the Time-of-Flight (ToF) cameras capture depth at 25 fps but have a limited spatial resolution. We propose to combine a high-resolution surface scan with a ToF-camera and a color camera to achieve both requirements. The 3D surface deformation is modeled by a NURBS surface that approximates the object surface and estimates the 3D object motion and local 3D deformation from the ToF and color camera data. A set of few NURBS control points can faithfully model the motion and deformation and will be estimated from the ToF and color data with high accuracy. The contribution will focus on the estimation of the 3D deformation NURBS from the ToF and color data.

2. "Calibration of a PMD camera using a planar calibration object together with a multi-camera setup" [SBK08].

Zusammenfassung:

We discuss the joint calibration of novel 3D range cameras based on the time-of-flight principle with the Photonic Mixing Device (PMD) and standard 2D CCD cameras. Due to the small field-of-view (fov) and low pixel resolution, PMD-cameras are difficult to calibrate with traditional calibration methods. In addition, the 3D range data contains systematic errors that need to be compensated. Therefore, a calibration method is developed that can estimate full intrinsic calibration of the PMD-camera including optical lens distortions and systematic range errors, and is able to calibrate the external orientation together with multiple 2D cameras that are rigidly coupled to the PMD-camera. The calibration approach is based on a planar checkerboard pattern as calibration reference, viewed from multiple angles and distances. By combining the PMD-camera with standard CCD-cameras the internal camera parameters can be estimated more precisely and the limitations of the small fov can be overcome. Furthermore we use the additional cameras to calibrate the systematic depth measurement error of the PMD-camera. We show that the correlation between rotation and translation estimation is significantly reduced with our method.

3. "Photoconsistent Relative Pose Estimation Between a PMD 2D/3D-Camera and Multiple Intensity Cameras" [BSK08]

Zusammenfassung:

Active range cameras based on the Photonic Mixer Device (PMD) allow to capture low-resolution depth images of dynamic scenes at high frame rates. To use such devices together with high resolution optical cameras (e.g. in media production) the relative pose of the cameras with respect to each other has to be determined. This task becomes even more challenging, if the camera is to be moved and the scene is highly dynamic. We will present an efficient algorithm for the estimation of the relative pose between a single 2D3D-camera with respect to several optical cameras. The camera geometry together with an intensity consistency criterion will be used to derive a suitable cost function, which will be optimized using gradient descend. It will be shown, how the gradient of the cost function can be efficiently computed from the gradient images of the high resolution optical cameras. We will show, that the proposed method allows to track and to refine the pose of a moving 2D3D-camera for fully dynamic scenes.

4. "Datastructures for Capturing Dynamic Scenes with a Time-of-Flight Camera" [SK09].

Zusammenfassung:

To capture 3D dynamic scenes, a suitable 3D data structure is needed that can represent the dynamic scene content. In this contribution we analyse and evaluate different data structures for capturing time-varying depth and color data of a dynamic scene obtained with Time-of-Flight and color cameras. The comparison of depth-panoramas, layered depth images and volumetric structures shows that a volumetric octree is best suited to fuse time-varying 3D scene data. We exploit the octree data fusion capabilities for different application scenarios like 3D environment building, volumetric object reconstruction and geometric interaction.

Literatur

- [BSK08] Christian Beder, Ingo Schiller, and Reinhard Koch. Photoconsistent Relative Pose Estimation Between a PMD 2D3D-Camera and Multiple Intensity Cameras. In *Proceedings of DAGM Symposium 2008, LNCS 5096*, pages 264–273, Munich, Germany, June 2008.
- [JSBK10] Andreas Jordt, Ingo Schiller, Johannes Brünger, and Reinhard Koch. High-Resolution Object Deformation Reconstruction with Active Range Camera. In Michael Goesele, Stefan Roth, Arjan Kuijper, Bernt Schiele, and Konrad Schindler, editors, *Pattern Recognition*, volume 6376 of *Lecture Notes in Computer Science*, pages 543–552. Springer Berlin / Heidelberg, 2010.
- [SBK08] Ingo Schiller, Christian Beder, and Reinhard Koch. Calibration of a PMD camera using a planar calibration object together with a multi-camera setup. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVII. Part B3a, pages 297–302, Beijing, China, 2008. XXI. ISPRS Congress.

[SK09] Ingo Schiller and Reinhard Koch. Datastructures for Capturing Dynamic Scenes with a Time-of-Flight Camera. In *Dynamic 3D Imaging DAGM 2009 Workshop, Dyn3D, LNCS 5742*, pages 42–57, Jena, Germany, September 2009.

Der State-of-the-Art report in [KBKL10] gibt einen Überblick über die erzielten Fortschritte im Bereich der Computer Grafik in Verbindung mit Laufzeitkameras. Poseschätzung aus Tiefenbildern wurde in [BSK08b] veröffentlicht. Das System für Erweiterte Realität wurde schrittweise in [BSBK08], [KSB⁺09] und [SBKK10] erweitert und publiziert. Die Kalibrierung wurde in [LSKK10] in Zusammenarbeit mit der Universität Siegen publiziert. In [BK09] wurde eine Methode vorgeschlagen um hochaufgelöste Tiefenkarten zu korrespondierende Farbbildern aus PMD-Bildern zu berechnen.

4.2 Andere Publikationen

1. Dissertation von Ingo Schiller 2011 (eingereicht)
"Dynamic 3D Scene Analysis and Modeling with a Time-of-Flight Camera"
2. Diplomarbeit von Johannes Brünger 2010 [Brü10]
"Approximation der Deformation dynamischer Oberflächen aus Tiefendaten" Diese Arbeit wurde mit dem Prof. Dr. Werner Petersen - Preis der Technik 2010, 2. Preis ausgezeichnet.
3. Diplomarbeit von Benjamin Angerer 2010 [Ang10]
"Modellbasierte volumetrische Rekonstruktion artikulierter Körper aus Tiefendaten mehrerer 3D-Kameras"
4. Diplomarbeit von Falko Kellner 2008 [Kel08]
"Umgebungsmodellierung und Detektion bewegter Objekte unter Verwendung einer aktiv gesteuerten PMD-Kamera" Die Arbeit wurde auf dem "Young Researchers Forum" auf der DAGM 2009 vorgestellt und mit dem 1. Preis ausgezeichnet.

Literatur

- [Ang10] Benjamin Angerer. Modellbasierte volumetrische Rekonstruktion artikulierter Körper aus Tiefendaten mehrerer 3D-Kameras, 2010. Diploma thesis, University of Kiel.
- [BBK07a] Christian Beder, Bogumil Bartczak, and Reinhard Koch. A Combined Approach for Estimating Patchlets from PMD Depth Images and Stereo Intensity Images. In *29th Annual Symposium of the German Association for Pattern Recognition, DAGM*, Heidelberg, September, 11th 2007.
- [BBK07b] Christian Beder, Bogumil Bartczak, and Reinhard Koch. A Comparison of PMD-Cameras and Stereo-Vision for the Task of Surface Reconstruction using Patchlets. In *The second international ISPRS workshop, BenCOS in Conjunction with IEEE CVPR*, Minneapolis, June 23 2007.

- [BK09] Bogumil Bartczak and Reinhard Koch. Dense depth maps from low resolution time-of-flight depth and high resolution color views. In George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Yoshinori Kuno, Junxian Wang, Renato Pajarola, Peter Lindstrom, André Hinkenjann, Miguel Encarnação, Cláudio Silva, and Daniel Coming, editors, *Advances in Visual Computing*, volume 5876 of *Lecture Notes in Computer Science*, pages 228–239. Springer Berlin / Heidelberg, 2009.
- [Brü10] Johannes Brünger. Approximation der Deformation dynamischer Oberflächen aus Tiefendaten, 2010. Diploma thesis, University of Kiel.
- [BSBK08] Bogumil Bartczak, Ingo Schiller, Christian Beder, and Reinhard Koch. Integration of a Time-of-Flight Camera into a Mixed Reality System for Handling Dynamic Scenes, Moving Viewpoints and Occlusions in Real-Time. In *Proceedings of the 3DPVT Workshop*, Atlanta, GA, USA, June 2008.
- [BSK08a] Christian Beder, Ingo Schiller, and Reinhard Koch. Photoconsistent Relative Pose Estimation Between a PMD 2D3D-Camera and Multiple Intensity Cameras. In *Proceedings of DAGM Symposium 2008, LNCS 5096*, pages 264–273, Munich, Germany, June 2008.
- [BSK08b] Christian Beder, Ingo Schiller, and Reinhard Koch. Real-Time Estimation of the Camera Path from a Sequence of Intrinsically Calibrated PMD Depth Images. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVII. Part B3a, pages 45–50, Beijing, China, 2008. XXI. ISPRS Congress.
- [KBKL10] A. Kolb, E. Barth, R. Koch, and R. Larsen. Time-of-Flight Sensors in Computer Graphics. *Computer Graphics Forum*, 29(1):141–159, 2010.
- [Kel08] Falko Kellner. Umgebungsmodellierung und Detektion bewegter Objekte unter Verwendung einer aktiv gesteuerten PMD-Kamera, 2008. Diploma thesis, University of Kiel.
- [KSB⁺09] Reinhard Koch, Ingo Schiller, Bogumil Bartczak, Falko Kellner, and Kevin Koeser. MixIn3D: 3D Mixed Reality with ToF-Camera. In *Dynamic 3D Imaging DAGM 2009 Workshop, Dyn3D, LNCS 5742*, pages 126–141, Jena, Germany, September 2009.
- [LSKK10] Marvin Lindner, Ingo Schiller, Andreas Kolb, and Reinhard Koch. Time-of-Flight sensor calibration for accurate range sensing. *Computer Vision and Image Understanding*, 114(12):1318 – 1328, 2010. Special issue on Time-of-Flight Camera Based Computer Vision.
- [SBK08] Ingo Schiller, Christian Beder, and Reinhard Koch. Calibration of a PMD camera using a planar calibration object together with a multi-camera setup. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVII. Part B3a, pages 297–302, Beijing, China, 2008. XXI. ISPRS Congress.

- [SBKK10] Ingo Schiller, Bogumil Bartczak, Falko Kellner, and Reinhard Koch. Increasing Realism and Supporting Content Planning for Dynamic Scenes in a Mixed Reality System incorporating a Time-of-Flight Camera. *Journal of Virtual Reality and Broadcasting* 7(2010), *CVMP 2008 Special Issue*, no. 4, August 2010. urn:nbn:de:0009-6-25786, issn 1860-2037.
- [SK09] Ingo Schiller and Reinhard Koch. Datastructures for Capturing Dynamic Scenes with a Time-of-Flight Camera. In *Dynamic 3D Imaging DAGM 2009 Workshop, Dyn3D, LNCS 5742*, pages 42–57, Jena, Germany, September 2009.