

2D/3D-Multisensorkamera für Echtzeitanwendungen (MultiCam) **Abschlussbericht**



Dr. rer. nat.
Wolfgang Weihs

Prof. Dr.-Ing. habil.
Otmar Loffeld

Dr.-Ing.
Klaus Hartmann

Zentrum für Sensordysteme (ZESS)

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

Dynamisches 3D Sehen

Prof. Dr.-Ing.
Andreas Kolb

Zentrum für
Sensordysteme
(ZESS)

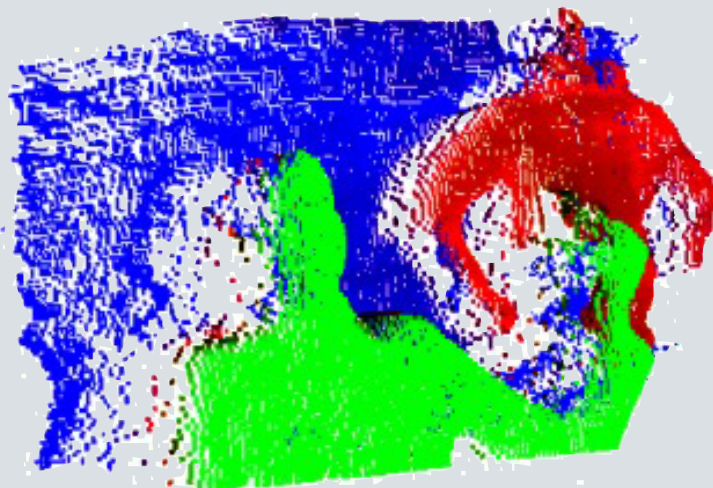
Universität Siegen
Paul-Bonatz-Str. 9-11
57068 Siegen



Prof. Dr.-Ing. habil.
Otmar Loffeld

Zentrum für
Sensordysteme
(ZESS)

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



2D/3D Multisensorkamera für Echtzeitanwendungen

Abschlussbericht zum DFG-Projekt
LO 455/10-1, -2
„MultiCam“

im Rahmen des DFG Forschungspakets PAK-73
"Dynamisches 3D Sehen mit PMD"

Inhaltsverzeichnis

1	Allgemeine Angaben	3
1.1	DFG-Geschäftszeichen	3
1.2	Antragsteller.....	3
1.3	Institut/Lehrstuhl.....	3
1.4	DFG-finanzierte wissenschaftliche Mitarbeiter	3
1.5	Thema	3
1.6	Berichts- und Förderzeitraum	3
1.7	Fachgebiet, Arbeitsrichtung.....	4
1.8	Verwertungsfelder.....	4
1.9	Am Projekt beteiligte Kooperationspartner (Name, Ort).....	4
2	Zusammenfassung	4
2.1	Wesentliche Ergebnisse und erzielte Fortschritte	4
2.2	Ausblick	5
3	Arbeits- und Ergebnisbericht.....	6
3.1	Ausgangslage	6
3.2	Beschreibung der durchgeführten Arbeiten	6
3.3	Darstellung der erzielten Ergebnisse.....	10
3.4	Ausblick auf zukünftige Arbeiten	14
3.5	Interdisziplinäre Weiterentwicklung.....	16
3.6	Verwertungspotenzial	16
3.7	Beteiligte Wissenschaftler.....	17
4	Liste der im Teilprojekt entstandenen Publikationen	18
4.1	Begutachtete Publikationen.....	18
4.1.1	Wissenschaftliche Zeitschriften	18
4.1.2	Fachkonferenzen	18
4.2	Andere Veröffentlichungen.....	19
4.3	Patente.....	19
5	Publikationen Dritter	19
6	Anhang: Das Kameramodell	19
6.1	Intrinsischen Parameter.....	20
6.2	Fusion der Daten	20

1 Allgemeine Angaben

Abschlussbericht im Rahmen des DFG Forschungspakets PAK-73 „Dynamisches 3D Sehen mit PMD“

1.1 DFG-Geschäftszeichen

LO 455/10-1, -2

DFG-Projekt "2D/3D Multisensorkamera für Echtzeitanwendungen (MultiCam)" im Rahmen des Antragspakets PAK-73 "Dynamisches 3D Sehen mit PMD"

1.2 Antragsteller

Prof. Dr.-Ing. habil. Otmar Loffeld
Nassauer Straße 21
57234 Wilnsdorf
Tel.: 02737 2299936

Dr.-Ing. Klaus Hartmann
Hoheroth 90
57234 Wilnsdorf
Tel.: 02739 2622

Dr. rer. nat. Wolfgang Weihs
Bismarckstr. 66A
56470 Bad Marienberg
Tel.: 02661/939544

1.3 Institut/Lehrstuhl

Zentrum für Sensorsysteme (ZESS)

Universität Siegen

Paul-Bonatz-Str. 9-11

57076 Siegen

Tel.: +49 271 740 3323

Fax.: +49 271 740 2336

E-Mail-Adresse: multicam@zess.uni-siegen.de

1.4 DFG-finanzierte wissenschaftliche Mitarbeiter

Herr Dipl.-Ing. Oliver Lottner und Dipl.-Ing. Jörg Grönweller wurden vom 1.3.2008-31.12.2010 ganz bzw. teilweise finanziert. Herr M.Sc. Seyed Eghbal Ghobadi und Herr Dipl.-Phys. Omar Löprrich in der Zeit vom 1.1.2009-30.09.2010 teilweise finanziert..

1.5 Thema

2D/3D Multisensorkamera für Echtzeitanwendungen (MultiCam)

1.6 Berichts- und Förderzeitraum

Förderzeitraum des Erstantrags: 01.03.2006 – 28.2.2008

Förderzeitraum des Folgeantrags: 01.03.2008 – 28.2.2010

Kostenneutrale Verlängerung der Laufzeit bis zum 28.2.2011

Berichtszeitraum: 01.03.2008 – 28.2.2011

1.7 Fachgebiet, Arbeitsrichtung

Informatik, System- und Elektrotechnik

1.8 Verwertungsfelder

2D/3D-Bildverarbeitung, multidimensionale, multimodale Sensorik, eingebettete Systeme

1.9 Am Projekt beteiligte Kooperationspartner (Name, Ort)

-keine-

2 Zusammenfassung

2.1 Wesentliche Ergebnisse und erzielte Fortschritte

In Multicam wurde unter den Aspekten Multisensorkamera, Multimodalität, Multi Exposures und der Multispektralanalyse eine Vielzahl von Detailproblemen analysiert und Lösungsansätze erarbeitet.

Modellierung des Messverfahrens

Parallel zu den Arbeiten in MultiCam/Dyn3D /1, 2, 4, 5/ haben auch andere Gruppen, außerhalb von Dyn3D, an Untersuchungen zu den Systemparametern und deren Interpretation, Einflüsse und Optimierung gearbeitet /12, 13, 14/. Diese Arbeiten basieren im Wesentlichen auf gegebenen Anordnungen von Standardmodulen (Kamera/Beleuchtung). Die Ergebnisse passen zu unseren Ergebnissen ohne aber die erweiterten Möglichkeiten der unterschiedlichen Modalitäten nutzen zu können. Aktuell von uns vorliegende Veröffentlichungen dazu sind noch nicht zitierfähig abgeschlossen.

Neben den Erkenntnissen über Detailabhängigkeiten von Parametern untereinander und deren Einflüsse auf das offensichtliche Ergebnis der Abstandsdaten und Bildinformationen konnten Verbesserungen für den Modulationskontrast und eine Reduzierung der Störeffekte durch das Hintergrundsignal erreicht werden. Durch einen rekursiven Ansatz der Phasenbilddauswertung konnte eine Echtzeitlösung entwickelt werden, die Abstandsbilder mit der vollen Einzelbildrate liefert.

Die Unterdrückung der Bewegungsartefakte ist ein wichtiges Kriterium für die Szenenanalyse TOF-Kameras. Die Bewegungsartefakte entstehen durch die Auswertung mehrerer, zeitlich aufeinanderfolgender Phasenbilder mit unterschiedlicher Phasenlage. Daraus ergibt sich der eindeutige Bezugspunkt für die Abstandsermittlung.

Eine raumzeitliche Analyse der Phasenbilder bietet eine deutliche Unterdrückung von Bewegungsartefakten, die durch eine flexible Synchronisation der Abstandsbilder und der hochauflösenden Farbbilder nochmal deutlich verbessert werden konnte.

Systemplattform

In Multicam entstand die Multimodalität auf der Kameraebene und wurde durch die sich verändernden Möglichkeiten der Chipene beeinflusst. Abgesehen von den Möglichkeiten der Vorverarbeitung bei den Sensor-Chips und der Kameraelektronik ergeben sich die Vorteile der Multisensorik/Multimodalitäten durch eine geeignete Algorithmik bei der Bilddatenverarbeitung. Aus diesem Grund wurden auch Untersuchungen z.B. zu Klassifikationsverfahren unter Berücksichtigung der Kamerasystemanalyse und -optimierung im Projekt durchgeführt.

Mit dem Verfahren der dynamischen, adaptiven Integrationszeiten (multi exposures) konnte auch bei dynamischen Szenen die Fehlerunterdrückung z.B. bei den Bewegungsartefakten verbessert werden bzw. führten die Arbeiten zu robusteren Systemverhalten. Eine allgemeinere, umfassendere Analyse konnte im Projekt dazu nicht durchgeführt werden.

Auf Basis der flexiblen und leistungsfähigen Kameraplattform konnten die vorgesehenen Lösungen für die Datenvorverarbeitung konsequent für eine Echtzeitleistung implementiert werden.

Nahfeld und Fernfeld Betrachtung

Im Nahfeld wie im Fernfeld hat sich der große Vorteil durch die entfernungsunabhängige Registrierung der Kamerapixel der beiden Sensorchips bestätigt. Durch die vielfältigen Analysen mit unterschiedlichen Aufbauten konnte eine auch physikalisch gut begründete Modellierung verschiedener Phänomene erzielt werden.

Im Fernfeld ergeben sich die Fragen insbesondere nach der Phaseneindeutigkeit (Phase Unwrapping). Die Realisierung eines Mehrfrequenzverfahrens als Standardlösung wurde analysiert. Auf Basis der Erfahrungen mit der Signaltheorie in der SAR-Signalverarbeitung konnten zu diesem Aspekt interessante Ergebnisse gefunden werden.

Multispektrale Szenenanalyse

Durch die Systemplattform stehen Bilddaten unterschiedlicher spektraler Bereiche zur Verfügung, die zeitgleich erfasst werden und räumlich durch eine spezifische Kalibrierung entfernungsunabhängig zueinander registriert sind. Neben den Farbbildern stehen die Bilddaten des NIR-Bereichs zur Verfügung. Dabei konnte durch die adaptive Beleuchtungssteuerung und das Multi Exposure Verfahren eine zusätzliche und insbesondere im Outdoorbereich sehr nützliche Information hinzugewonnen werden.

2.2 Ausblick

Die Systemplattform sowie Multimodale und Multispektrale Klassifikation und Tracking werden in verschiedenen neuen Projekten wie z.B. im Beitrag zum Graduiertenkolleg GRK 1564 "Imaging New Modalities" und in den Projekten LaotSe, P0mSe und Movedetect (s.a. Unterpunkt 3.4) bereits weiterentwickelt bzw. die Ergebnisse/Erkenntnisse aus MultiCam werden dort angewendet. Wir sehen eine Reihe weiterer relevanter Fragestellungen für die wir mit den erarbeiteten Ergebnissen Lösungskonzepte anbieten können.

Eine sehr interessante neue Entwicklung auf Basis der bisherigen Erkenntnisse bietet sich mit der Theorie des Compressive Sensing. Es erscheint uns sehr wahrscheinlich, dass sich z.B. die Empfindlichkeit und Auflösung der TOF-Technologie damit deutlich steigern lässt. Dazu sind wiederum spezifische Hardware- und algorithmische Lösungen zu entwickeln.

3 Arbeits- und Ergebnisbericht

3.1 Ausgangslage

Die Zielsetzung des Teilprojekts „MultiCam“ bestand in der Systementwicklung einer echtzeitfähigen monokularen 2D/3D-Kamerasystemplattform und der anschließenden Verifikation der wichtigsten Parameter dieser Plattform. Diese Plattform vereint eine chipbasierte Phasenmessung auf Basis der Continuous-Wave Time-of-Flight Technik (CW ToF) in Form eines PMD-Sensors mit einer dazu synchronisierten 2D-Bilddatenerfassung mittels gewöhnlicher CMOS-Farbbildsensoren.

Die Justierung eines solchen Systems ist fraglos komplizierter als die einer normalen Kamera. Es müssen jetzt nicht nur zwei Sensoren zueinander justiert werden, sondern es ist auch die Auftrennung des einfallenden Lichtes mit Hilfe eines Strahlteilers zu vollziehen: infrarotes Licht für die Phasenmessung und sichtbares Licht für den 2D-Farbsensor. Ausgehend von einer einfachen handelsüblichen Strahlteilerplatte, sollte im zweiten Teil des Projektes eine Prismenoptik unter Optimierung des Justageaufwandes und zur Minimierung kamerainternen Streulichtes untersucht werden.

Durch die FPGA-kontrollierte Ansteuerung beider Sensoren wird eine hochpräzise Synchronisierung auf Basis der Integrationszeit ermöglicht, wodurch neue Möglichkeiten der Bilddatenauswertung untersucht werden sollen. Neben der zeitlichen Synchronisierung erlaubt der monokulare Aufbau die koregistrierte Erfassung der 2D- und 3D-Bilddaten; die hieraus resultierenden Vorteile sind ebenfalls Bestandteil der Untersuchung. Darüber hinaus erlaubt die Eigenentwicklung der aktiven Beleuchtung und Beleuchtungsansteuerung für den 3D-Strang weitergehende Untersuchungen hinsichtlich Multi-Beleuchtungsanordnungen oder Beleuchtungen für große Entfernungen.

Als Konsequenz der Ansiedelung der Arbeiten zu Beginn des Projektes in die Nähe der Hardware und die damit erworbenen Kompetenzen in der Steuerungs- und Ausleseelektronik ergab sich automatisch die Funktion als Zulieferer für zweckorientiert modifizierte Aufbauten. So sollte z.B. eine spezielle Version der PMD Kamera für das Projekt „3DTHz“ bereitgestellt werden.

Auf Grund der internen Anordnung einer MultiCam ist gegenüber einem binokularen Aufbau vor allem eine vereinfachte Registrierung zu erwarten. Dementsprechend gut sollte das System auch z.B. zu Klassifizierungszwecken geeignet sein. Entsprechende Untersuchungen zur Verifizierung dieser Annahme wurden im zweiten Teil des Projektes angestrebt.

3.2 Beschreibung der durchgeführten Arbeiten

Parametrisierung der Basisplattform

Zur Verifikation der im Projekt „MultiCam“ entwickelten monokularen 2D/3D-Kamera wurde der Aufbau mit einer binokularen 2D/3D-Lösung verglichen. Ein wesentlicher Nachteil der binokulare Variante besteht darin, dass eine Registrierung nur dann möglich ist, wenn die 3D-Daten eine genügend hohe Messgenauigkeit aufweisen, was prinzipbedingt nur unter günstigen Beleuchtungsverhältnissen der Fall ist. In größeren Szenen und insbesondere im Outdoor-Bereich kann die Registrierung nur mit entsprechender Unsicherheit berechnet werden. Eine Verifikation des monokularen Kameramodells hingegeben ergab die Unabhängigkeit der Registrierung vom Abstand zu der Szene. Darüber hinaus ergeben sich keine durch verschiedene Aufnahmezeitpunkte bedingten Zuordnungsprobleme in dynamischen Szenen

PMD Modul für „3DTHz“ und Optimierung der Sensoransteuerung

Für das Teilprojekt „3DTHz“ wurde ein PMD Sensorfrontend mit modifizierten Ausleseelektronik und angepasster Auslesesoftware bereitgestellt. In dem experimentellen Aufbau wurde eine mit der Ausgangsfrequenz des Frequenzkomparators synchronisierte Ansteuerung des PMD Sensors benötigt. So bestand die Modifikation im Wesentlichen in der Implementierung dieser externen Modulationsmöglichkeit. Jedoch zeigten sich sehr schnell Defizite in der vom Sensorhersteller vorgeschlagenen und zunächst zu Beginn des Projektes übernommenen Ansteuerungselektronik. So wurden, ausgehend von den Bedürfnissen des „3DTHz“ Projektes, während des Berichtszeitraumes zahlreiche systematische Untersuchungen und Optimierungen hinsichtlich der elektrischen Beschaltung des PMD-Sensors und der Ausleseelektronik durchgeführt.

Weiterentwicklung der Optik

Nach der Fertigstellung eines F-Mount-Prototypen in der ersten Hälfte des Projekts stand die Anpassung an den C-Mount-Standard im Vordergrund der zweiten Projekthälfte, um die Kamera zu verkleinern und vor allem um von der breiten Auswahl an C-Mount-Objektiven mit hoher Lichtdurchlässigkeit und kleiner Brennweite, also großem Öffnungswinkel, profitieren zu können. Vor allem harmonisiert der Bildkreis dieser Objektive wesentlich besser mit den kleinen Chipabmessungen. Zu diesem Zweck wurden verschiedene optische Konfigurationen mit unterschiedlichen Prismenformen untersucht. Die Filtercharakteristika der verschiedenen Strahlteileranordnungen wurden verglichen, indem die Transmissionskurven über dem relevanten Wellenlängenbereich aufgenommen wurden.

Weiterentwicklung der Beleuchtung

Durch die C-Mount-Variante ergibt sich zuallererst kameraseitig die Möglichkeit eines größeren Öffnungswinkels, so dass sich die Frage stellt, wie das größere Blickfeld beleuchtet werden kann. Aus diesem Grund wurde im nächsten Schritt auch die Beleuchtung angepasst. Zu diesem Zweck wurden moderne, effiziente Hochleistungs-LEDs eingesetzt, die hinsichtlich der Modularität und dem thermischen Verhalten leicht an die hier gegebenen Verhältnisse adaptiert werden konnten. Sie emittieren eine optische Leistung von ca. 3,5W und unterscheiden sich somit um eine Größenordnung von den bislang verwendeten LEDs. Der große Abstrahlwinkel von ca. 120° bedingt jedoch den Einsatz von (z.T. kommerziell erhältlichen) Reflektoren oder Linsen, um (nur) das Sichtfeld der Kamera zu beleuchten. Je nach Anwendungszweck eignen sich hier unterschiedliche Konfigurationen (s.u.). Eine weitergehende Optimierung für noch größere Abstrahlwinkel wurde nicht betrieben, da sich Überschneidungen mit einem Teilprojekt von Lynkeus¹ abzeichnen. Stattdessen begannen wir mit Studien zur Verringerung des Beleuchtungsraumwinkels, um eine Erweiterung der Reichweite des Systems zu erzielen.

Mehrere Beleuchtungen

Zu Ausleuchtung größerer Raumbereiche und zum Erreichen besserer Signalverhältnisse bei größeren Abständen wurde darüber hinaus untersucht, inwieweit eine (verteilte) Anordnung eines oder mehrerer Beleuchtungsmoduls / -module zielführend ist. Der Standardfall wird dann als monostatische Anordnung mit einer Beleuchtung bezeichnet. Die bi-statische Anordnung, also die räumliche Separierung von der Kamera und der Beleuchtung, ist für den Sonderfall der gegenüberliegenden Konfiguration interessant, wenn die Reflektionscharakteristika eines zwischen Kamera und Beleuchtung liegenden Bodens ausgenutzt werden sollen. Beim Betrieb einer oder mehrerer Kameras mit mehreren Beleuchtungen ist zwischen kooperativem und unabhängigem Betrieb zu unterscheiden. Der unabhängige Betrieb erfordert, dass eine Kamera durch bestimmte Beleuchtungen zu einem möglichst geringen Grad in ihrer Distanzmessung beeinflusst wird, während eine andere Kamera eben diese Beleuchtungen benutzt. Im

¹ Mikointegrierte 3D Echtzeitkamarasysteme für intelligente Umgebungserfassung (Lynkeus), Rahmenprogramm „Mikrosysteme 2004-2009“ des Bundesministeriums für Bildung und Forschung (BMBF)

kooperativen Modus zielt die Zusammenschaltung mehrerer Beleuchtungen für eine oder mehrere Kameras darauf, die Szene oder einen Teil davon optimal synchron zu beleuchten. Die systematische Untersuchung für beide Verfahren zeigte zunächst, dass eine Separierung der Beleuchtungen über eine Parametrisierung der Modulationsfrequenzen einfach und schnell erreichbar ist; es gibt lediglich einen Zusammenhang mit der Integrationszeit, die in Abhängigkeit des Unterschieds der Modulationsfrequenzen gewählt werden muss. Im kooperativen Modus arbeiten die Beleuchtungen mit der gleichen Modulationsfrequenz um auch in größerer Entfernung bzw. in einem größeren Volumen ausreichend stabile Daten aufzunehmen. Die resultierende Mischphase hängt von den Amplituden und den Phasen der sich mischenden Wellen ab; es kommt zu einer Modulationsinterferenz. Dieser Effekt lässt sich auf der einen Seite ausnutzen, indem weitere Beleuchtungen so in das Phasenfeld der Kamera eingebracht werden, dass dieses nur in vernachlässigbarem Maße gestört wird, während die Modulationsamplitude und damit das Signal-zu-Rausch-Verhältnis deutlich verbessert wird. Auf der anderen Seite können weitere Beleuchtungen auch bewusst so eingebracht werden, dass eine Änderung der realen Distanz zu einer überproportionalen Änderung der gemessenen Phase führt, so dass zwar die Informationen über den realen Abstand verloren geht, die Änderungen dafür aber umso sicherer bzw. robuster erkannt werden kann.

Die Bewegung einer Kamera mit monostatischer Beleuchtung führt zum Auftreten von Bewegungsartefakten, da die einzelnen Phasenbilder hintereinander und somit bei unterschiedlichen Relativanordnungen von Kamera und Szene aufgenommen werden. Im Laufe des Projektes konnten verschiedene Lösungsansätze für das Problem gefunden werden, z.B. die rekursive Multi-Phasenauswertung, bei der ein Distanzbild bereits mit der Aufnahme eines Phasenbildes (statt nach vier Aufnahmen) berechnet wird, oder die parallele Phasendiskrimination, bei der zwei co-registrierte PMD-Sensoren zur gleichzeitigen Aufnahme verschiedener Phasen der gleichen Szene benutzt werden. Ist hingegen nur die Beleuchtung oder nur die Kamera ortsfest, können keine absoluten Distanzdaten ermittelt werden, allerdings bleiben die Distanzdaten in sich stimmig, d.h. die relativen Abstände von beobachteten Objekten ändern sich nicht. Jedoch unterliegen die relativen Daten dem bi-statischen Effekt, d.h. diese Daten werden von dem Winkel zwischen der Beleuchtung und der Kamera beeinflusst.

Beobachten mehrere MultiCams mit einer oder mehreren unabhängigen Beleuchtung(en) die gleiche Szene, können die Daten auch im Sinne einer Stereokonfiguration prozessiert werden. In dem Fall ist es möglich, die Distanzdaten der Szene im Sinne der lateralen Auflösung zu verfeinern, indem die Distanzdaten des 3D-Strangs als Initialisierungswerte für die Stereotiefenberechnung mit Hilfe der hoch aufgelösten 2D-Daten dienen. Die Stereoberechnung kann mit bekannten, etablierten Verfahren erfolgen und profitiert von der guten Initialisierung in der Weise, dass der Rechenaufwand deutlich reduziert wird. Der Vorteil dieser Kombination besteht darin, dass die beiden Verfahren sich gut ergänzen: das Stereoverfahren kann keine Distanzdaten für schwach texturierte Ebenen bilden, während das ToF-Verfahren gerade hier seine Stärke zeigt. An Kanten hingegen lassen sich gut Entfernungswerte über das Stereoverfahren ableiten, sofern sich diese Kanten als Distanz- und Intensitätskanten gleichzeitig zeigen, während beim ToF-Verfahren die relativ grobe laterale Auflösung zur Messung einer Mischphase (gemittelt aus Kanten- und Hintergrundabstand) führt.

Sensorik und Schnittstelle

Durch die Integration des neuen PMD-41kS-Sensors konnte die Leistung des 3D-Teils der MultiCam insbesondere hinsichtlich der Empfindlichkeit und der lateralen Auflösung verbessert werden. Zugleich wurde der bislang verwendete 2D-VGA-Sensor gegen einen höher auflösenden 1.3MP-Sensor (RGB) getauscht. Um der somit gestiegenen Datenmenge der beiden Sensoren Rechnung zu tragen, wurde die relative langsame USB2.0-Schnittstelle zum PC ausgewechselt gegen eine Gigabit Ethernet Verbindung. Entgegen der im Antrag angedachten Camera-Link Lösung, bietet Gigabit Ethernet neben der höheren Bandbreite insbesondere den Vorteil der einfacheren Systemintegration. Mit einer offenen UDP/IP-Implementierung konnte eine Nettodatenrate von etwa 110Mb/s erzielt werden. Dadurch konnten alle Modalitäten auch unter Echtzeitaspekten genutzt werden.

Scanning

Reicht für einen bestimmten Anwendungsfall die höhere laterale Auflösung des neuen PMD-Sensors nicht aus, ist es möglich, diese durch Micro-Scanning weiter zu verfeinern. Diese Technik bezieht sich auf die Aufnahme mehrerer Bilder, wobei der Sensor von Aufnahme zu Aufnahme durch eine μm -Positioniereinheit räumlich innerhalb der Grenzen, die durch die Abmessungen eines Pixels definiert sind, verschoben wird. Sind die Bewegungen des Sensors bekannt, entsteht somit ein Bild eines virtuellen Sensors mit höherer Auflösung. Durch Erhöhung der Verschiebungsinkremente ist es auch möglich, nicht nur die Auflösung, sondern auch das Sichtfeld des Sensors zu vergrößern, wenn die Optik für ein größeres Sensorformat ausgelegt wurde wie z.B. in der F-Mount-Variante der Hardware.

Multispektrale Szenenanalyse

Die Definition eines Merkmalsraums für die multispektrale Klassifikation umschließt die Wellenlängenbereiche VIS, Near-IR. Die verschiedenen Modalitäten sind natürliches Umgebungslicht einer Szene und/oder künstliches Licht in zumindest einzelnen Wellenlängenbereichen sowie die unterschiedlichen Sensorelemente als Detektoren dieser Strahlungsinformationen. Die Suche nach einer geeigneten Klassifikation für einen solchen Merkmalsraum, ist vielfältig mit der Systemanordnung verbunden um einen Nutzen für eine spezifische Anwendung zu generieren. In unseren Arbeiten hat sich auch an dieser Stelle gezeigt, dass der iterative Prozess einer Abstimmung von physikalischen Parametern und algorithmischen Lösungsansätzen zu interessanten und optimalen Ergebnissen führen kann. Beispielsweise sind die über die dichroitischen Schichten getrennten Spektralbereiche der beiden Sensormodalitäten hilfreich und hinderlich zu gleich, weshalb wir teilweise leichte Veränderungen bezüglich der Transmission vorgenommen haben. Die Kalibrierung wird durch einen gemeinsamen Signal/Spektralbereich einfacher. Für eine hohe Empfindlichkeit möchte man aber die Transmission im IR-Bereich optimieren. Für die Spektralanalyse will man die Bereiche ebenfalls optimal trennen. Wir haben für die Kalibrierung eine an die Filtergüte angepasste Beleuchtungsquelle verwendet um im Empfangskanal keine Kompromisse eingehen zu müssen. Beim anschließenden Wechsel der Beleuchtung werden die Daten der Kalibrierung nicht mehr verändert.

Klassifizierung

Ein wichtiges Anwendungsgebiet für die MultiCam sind Klassifikationsaufgaben z.B. in Mensch-Maschine-Interaktionssystemen. Die ToF-Technik hat in diesem Zusammenhang zunehmend an Bedeutung gewonnen. Hinsichtlich der MultiCam können die Verfahren von der gleichzeitigen Erfassung verschiedener Modalitäten profitieren, z.B. durch eine vereinfachte Objektsegmentierung. Im einfachsten Fall kann das Distanzbild selbst direkt eine Segmentierung durch Clusterbildung liefern. Dank der szenenunabhängigen Registrierung kann dieses Ergebnis direkt auf das 2D-Bild übertragen werden. Gegenüber einer reinen 2D-Lösung ist dieses Verfahren auf Grund der geringeren Auflösung weniger rechenaufwändig, und profitiert von robusteren Daten, da eine aktive modulierte Beleuchtung verwendet wird. Im nächsten Schritt wurde die 3D-basierte Segmentierung um eine Kantendetektion im 2D-Bild ergänzt. Die Kombination beider Modalitäten führt dazu, dass sich die Segmentierung einerseits nicht durch Texturierungen irreführen lässt, auf der anderen Seite können aber trotzdem Objekte, die sich in gleicher Distanz zur Kamera befinden, unterschieden werden. Schließlich wurde der Ansatz noch weiter verfeinert, indem der AdaBoost-Ansatz nach Viola und Jones verwendet wurde. Dabei handelt es sich um ein Objekterkennungssystem mit Haar-Merkmalen. Durch die Kombination der 2D- mit den 3D-Daten ist es möglich, das sonst stark störende Hintergrundrauschen weitgehend zu eliminieren, und die Anzahl an Suchfenstern signifikant zu verringern, da die Größe der Suchfenster auf Grund des bekannten Abstands bekannt ist. Dies führt zu einer deutlichen Verringerung des Rechenaufwands.

3.3 Darstellung der erzielten Ergebnisse

Systemplattform

PMD Modul für „3DTHz“ und Optimierung der Sensoransteuerung

Für das Teilprojekt „3DTHz“ wurde ein modifiziertes PMD Sensorfrontend bereitgestellt, das zur Synchronisation mit der übrigen Apparatur über einen externen Modulationseingang verfügt. Während der nachfolgenden Messungen zeigten sich jedoch problematische Unterschiede zwischen den einzelnen Phasenbildern, auch wenn der PMD-Sensor mit unmoduliertem Licht betrieben wird. Abbildung 1 zeigt die Problematik an einen PMD-Pixel. Da sie Abstandsmessungen vortäuscht, wurde der Begriff Pseudo-Autokorrelationsfunktion (Pseudo-AKF) für dieses Phänomen geprägt. Während der Projektlaufzeit wurden Untersuchungen mit hoher Statistik durchgeführt, um diesen systematischen Fehler zu lokalisieren. Dabei stellte sich zum einen die Symmetrie des Tastverhältnisses der Modulationssignale sowohl für den PMD-Sensor als auch für die Beleuchtung als kritischer Faktor heraus. Zum anderen müssen für die vier Phasenmessungen die Signale für Sensor und Beleuchtung jeweils um 90° verschoben werden. Auch hierbei darf sich das Tastverhältnis keinesfalls ändern.

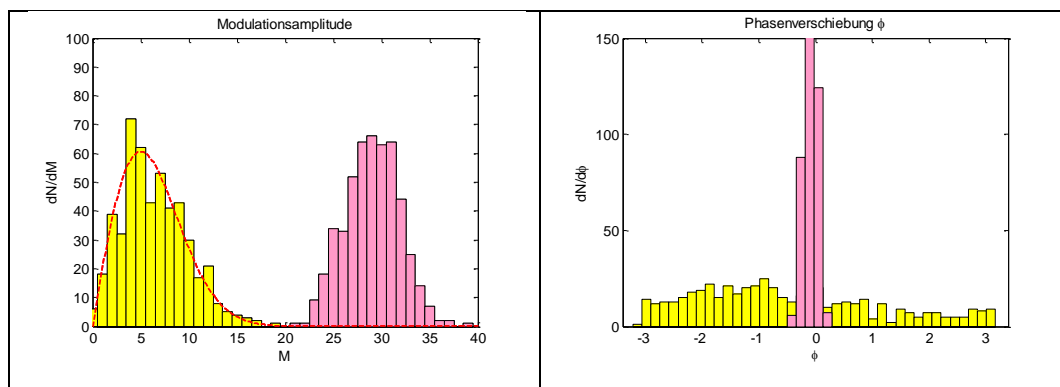


Abbildung 1: Modulationsamplitude und Phasenwert eines PMD Pixels. Histogrammteile in Magenta zeigen Daten vor der Modifikation. Es ist eine anomal hohe Modulationsamplitude und eine (pseudo) Phase erkennbar. Nach der Modifikation (Histogramme in Gelb) sehen wir die erwartete Rayleighverteilung² (rot gestrichelt) der Modulationsamplitude und annähernd gleichverteilte Phasenwerte.

Nach Durchführung der dazu nötigen Modifikationen stellte sich als Erfolg zunächst eine wesentlich reduzierte Umgebungslicht-Empfindlichkeit ein. In Folge können aber auch unjustierte Einstellungen leichter detektiert und korrigiert werden. Bedeutsam ist hier z.B. das Beleuchtungssignal. Um ein Tastverhältnis von 1:1 für das Lichtsignal zu bekommen, muss in der Regel das elektrische Ansteuersignal ein dazu modifiziertes Tastverhältnis vorweisen, welches schlechterdings auch von der verwendeten Frequenz abhängt. Muss während der Abstandsmessungen die Modulationsfrequenz verändert werden, um z.B. Distanzmehrdeutigkeiten aufzulösen, kann mit den im Projekt erarbeiteten Methode ein inkorrektes Tastverhältnis detektiert und (elektronisch und rechnergesteuert) nachjustiert werden.

Weiterentwicklung der Optik

² Als Sonderfall einer Rice-Verteilung mit Mittelwert Null

Hinsichtlich des optischen Aufbaus wurde nach längeren Tests und Erprobungen eine C-Mount-Variante auf Basis eines Bauernfeind-Prismas realisiert. Ein direkter Vorteil dieser Anordnung besteht darin, dass das 2D-Bild nicht gespiegelt aufgenommen wird. Die Transmissionskurve sieht der vorigen F-Mount-Lösung sehr ähnlich und erfüllt die Anforderungen an hoher Transmission im NIR-Bereich, hoher Reflektion im VIS-Bereich und möglichst steilem Anstieg dazwischen hinreichend gut. Ein nach unseren Vorgaben speziell gefertigter Prismensatz aus BK7 Glas und einer dichroitische Beschichtung erfüllte die Anforderungen schließlich sehr gut. Abbildung 22 zeigt den realisierten Aufbau mit der dazugehörigen Filterkurve. Bei Bedarf ist ein noch steilerer Anstieg realisierbar, der aber für die gegenwärtige Version (Wellenlänge der Beleuchtung: 850nm) nicht nötig ist.

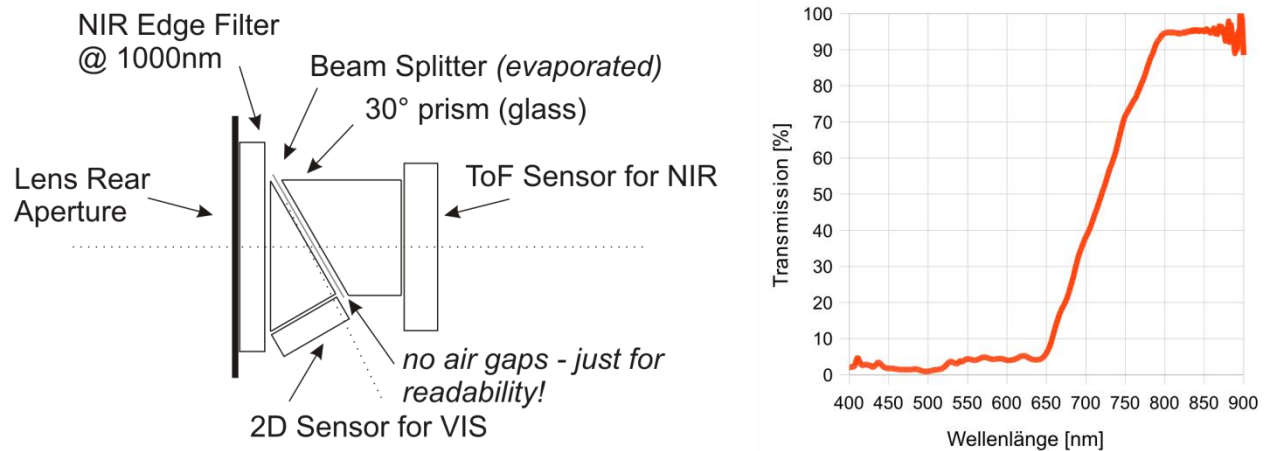


Abbildung 2: Realisierter C-Mount-Bauernfeind-Aufbau mit Transmissionskurve

Weiterentwicklung der Beleuchtung

Die Beleuchtung wurde entscheidend durch den Einsatz der Hochleistungs-LEDs verbessert. Zur Anpassung an das Sichtfeld der Kamera wurden TIR (total internal reflection) Optiken verwendet. Für die Beobachtung von Objekten in großer Entfernung ist der Einsatz von entsprechend großen parabelförmigen Reflektoren geeignet, die einen fast parallelen Lichtaustritt ermöglichen, wie die Simulation in 3 links oben zeigt. Die stark divergente Abstrahlcharakteristik der LED-Module kann somit in gebündeltes paralleles Licht gewandelt werden, ohne dass Inhomogenitäten in der Strahlverteilung erzeugt werden. Mit Hilfe von vier solcher Reflektoren (Abbildung 3 links unten) war es möglich, Personen und Objekte in einer Entfernung von 70m und mehr zu beobachten, wie im Beispielbild (Abbildung 33 rechts) zu sehen ist; auf dem Untergrund sind deutlich die Überschreitungen der Eindeutigkeitsbereiche zu erkennen. Die Beleuchtung ist so fokussiert, dass nicht der ganze Bildbereich des 41k-Sensors ausgeleuchtet wird. Natürlich muss die Brennweite bzw. die Optik i.A. auch an die Anwendung angepasst werden. Um noch höhere Reichweiten zu erzielen, ist es möglich, die emittierte Strahlung noch weiter zu bündeln, auch wenn dadurch nicht mehr der gesamte Bildbereich mit moduliertem Licht beleuchtet wird. Zu diesem Zweck kann eine Laserzeile oder ein Laserpunkt im Bildbereich bewegt werden. Aus verschiedenen Bildbereichen kann dann sequentiell ein 3D-Bild zusammengestellt werden.

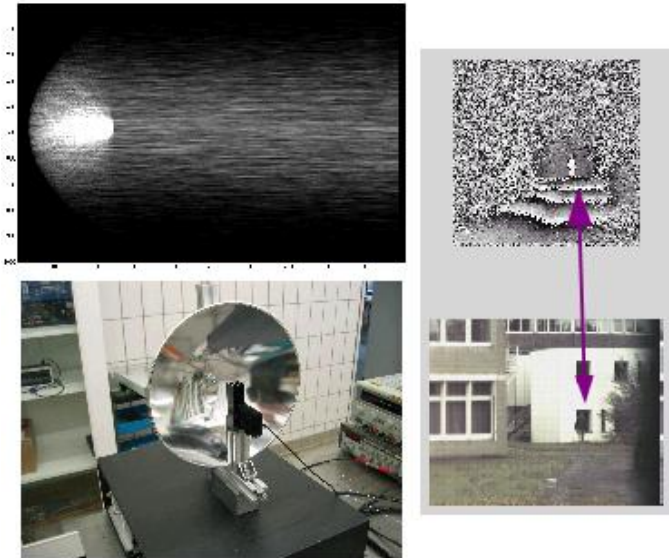


Abbildung 3: Beleuchtung für große Entfernungen

Mehrere Beleuchtungen

Die Zusammenschaltung mehrerer Beleuchtungen zur synchronen, optimierten Ausleuchtung einer Szene konnte mit insgesamt vier Beleuchtungsmodulen erfolgreich getestet werden. Abbildung 4 zeigt exemplarisch die Messergebnisse eines Aufbaus, in der eine Person mit Hilfe einer Kamera und einer bzw. mehreren Beleuchtungen beobachtet und verfolgt wird. Die Konfiguration mit nur einer Beleuchtung wird als „C1“, die mit mehreren Beleuchtungen als „C2“ bezeichnet. Als Maßstab für die Verbesserung dient die Modulationsamplitude, die als zuverlässiger Indikator für das Signal-zu-Rauschverhältnis zu sehen ist. Erfahrungsgemäß sind für die Klassifikation nutzbare Distanzdaten ab einer Modulationsamplitude von etwa 30..50 Zählereinheiten zu erwarten. Im linken Teil der Abbildung (Häufigkeit von Bereichen der Modulationsamplitude) ist deutlich zu sehen, dass mit mehreren Beleuchtungen die Beleuchtungsverhältnisse verbessert werden können, und nur noch wenige Aufnahmen ohne verwertbare Daten entstehen.

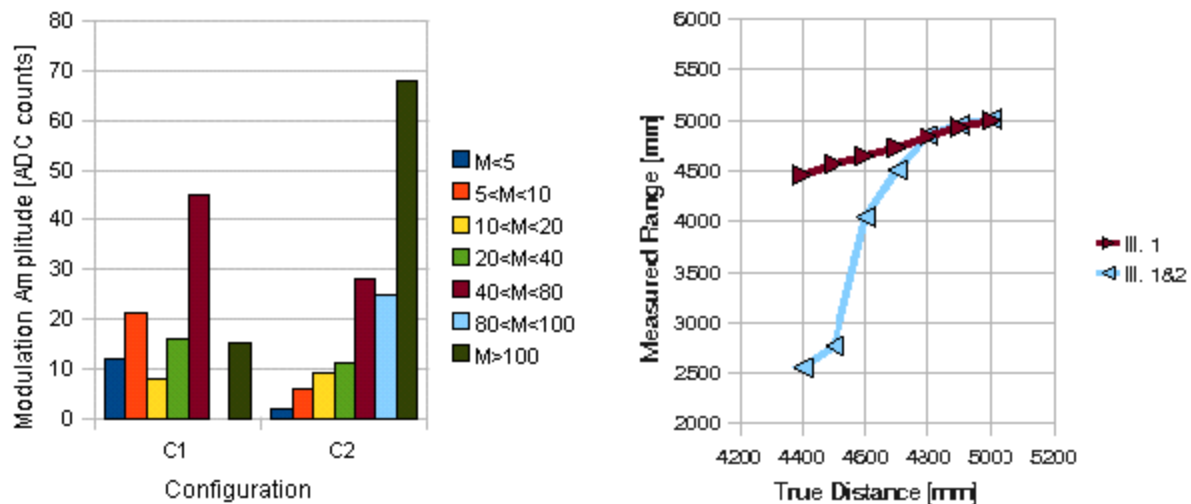


Abbildung 4: Messergebnisse von Multi-Beleuchtungs-Konfiguration

Der rechte Teil der Abbildung zeigt hingegen ein Beispiel, wo die Anordnung der Beleuchtungseinheiten bewusst so gewählt wurde, dass sich die gemessene Phase nicht mehr proportional zur realen Distanz

ändert. Dazu wurde eine zusätzliche Beleuchtung so angebracht, dass eine auf die Kamera zukommende Person innerhalb eines Teilbereichs seitlich beleuchtet wurde. Auf diese Weise entsteht eine überproportionale Phasenänderung, die zwar das reale Messergebnis verfälscht, eine Detektion der Änderung jedoch robuster macht.

Scanning

Das Konzept des Scannens, also das Bewegen des 3D-Sensors innerhalb der Sensorebene, wurde sowohl hinsichtlich des Micro-Scannings zur Verbesserung der Auflösung als auch hinsichtlich des Macro-Scannings, also das Vergrößern des Sichtfelds (dann in einem F-Mount-Aufbau mit einem Bildkreis, der dem Kleinbildformat entspricht), verifiziert. Zwar ist die Anwendbarkeit auf dynamische Szenen wegen der Bewegungsabläufe eingeschränkt, allerdings eignet sich das Konzept gut, um größere Szenen zu überwachen. Außerdem bleibt der Vorteil der tiefenunabhängigen Registrierung des monokularen Aufbaus auch hierbei erhalten. Als primäre Datenquelle ist dann der 2D-Sensor zu sehen, und sobald die Datenauswertung ergibt, dass zur Interpretation eines Teils des Bildinhalts eine weitere Modalität, die Abstandsdaten, herangezogen werden sollte, kann der 3D-Sensor entsprechend positioniert werden, so dass auf Basis der Datenfusion eine weitergehende Interpretation möglich ist. Bild 5 zeigt Beispiele für das Micro- und Macro-Scanning.

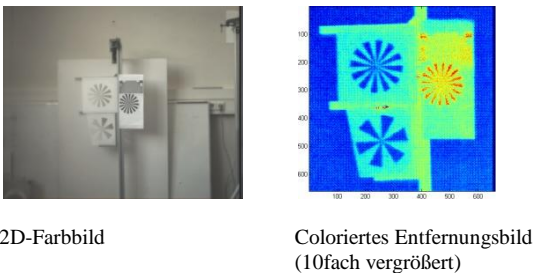


Abbildung 5a: Micro-Scanning Ergebnis für die Aufnahme zweier Böhlersterne und eines Siemenssterns

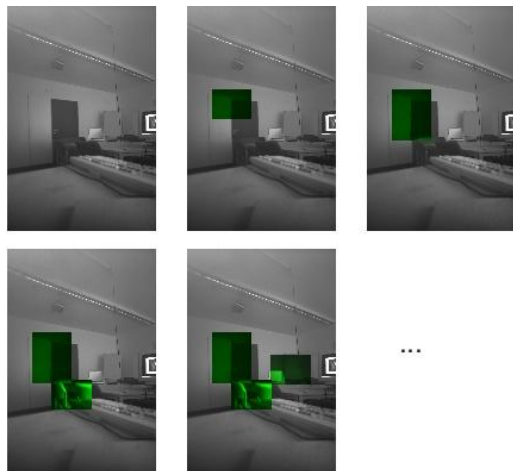


Abbildung 5b: Kombination eines 2D-Bildes (Seminarraum) mit 3D-Scanning Daten (grün)

Segmentierung und Klassifizierung

Zur Klassifizierung von Personen und Objekten wurde eine 2D/3D-Objektdetektionsverfahren nach Viola und Jones umgesetzt. Gegenüber konventionellen Algorithmen, die ohne die Fusion der 2D- und 3D-Bilddaten auskommen müssen, weist dieses Verfahren insb. zwei Vorteile auf. Zum einen kann störendes Hintergrundrauschen durch die Verwendung der 3D-Daten, die durch eine aktive modulierte Beleuchtung aufgenommen wurden, zu einem hohen Grad eliminiert werden. Dazu wird ein 3D-Volumen definiert, innerhalb dem sich das Objekt von Interesse befinden soll. Die so erhaltenen initiale Segmentierung kann auf Grund der einfachen Koregistrierung der Sensoren direkt auf das 2D-Bild übertragen werden. Zum anderen gestaltet sich die Anwendung der Haar-Merkmale einfacher, da die Größe der Suchfenster auf

Basis der 3D-Daten bekannt ist. Die Anzahl an zu überprüfenden Suchfenstern reduziert sich somit deutlich, und die Rechenkomplexität sinkt. Ein Beispiel ist in Abbildung 66 zu sehen.

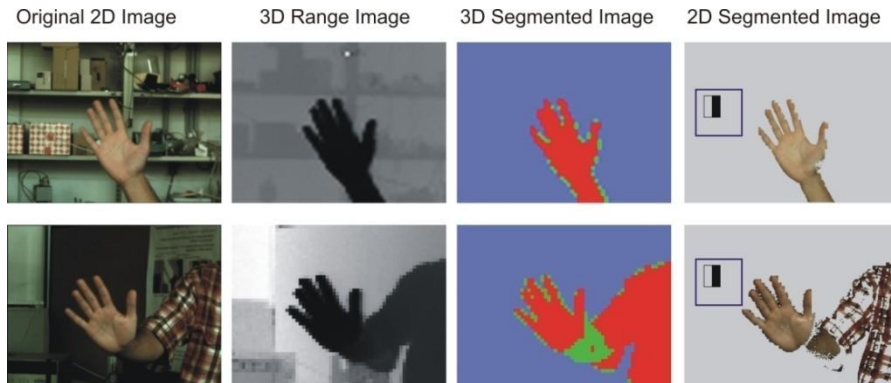


Abbildung 6: Beispiel zur Anwendung der 2D/3D-Segmentierung nach Viola und Jones

3.4 Ausblick auf zukünftige Arbeiten

In Multicam wurde unter den Aspekten Multisensorkamera, Multimodalität, Multi Exposures, Multidimensionalität und der Multispektralanalyse eine Vielzahl von Detailproblemen analysiert und Lösungsansätze erarbeitet. Die Ergebnisse mit der verteilten Beleuchtung und dem Mikroscanning zeigen z.B. interessante Möglichkeiten für die Echtzeitanalyse großer Szenen die Scananordnung durch eine MultiChip-Lösung zu ersetzen.

Die erreichten grundsätzlichen Ergebnisse und Erkenntnisse werden bereits bei den Arbeiten im Graduiertenkolleg verwendet und bieten darüber hinaus die Möglichkeiten auch unabhängig von der speziellen Sensor-/Kamerälösung des Multcam-Projekts in anderen, weiterführenden Projekten erfolgsversprechend eingesetzt zu werden. Im LaotSe-Projekt werden die Modalitäten z.B. auf den Therm-IR Bereich erweitert. Für den Outdoor-Bereich ist die Multispektral-Analyse eine wichtige Informationsquelle. Die Interaktion zwischen sich verändernden Systemparametern (Hardwarebasis) und einer möglichst flexiblen Auswertestrategie mit unterschiedlichen Algorithmenansätzen zeigt ein großes Potenzial. Die Raumzeitliche Analyse verteilter Daten in Verbindung mit Sensornetzwerken wurde in neuen Projekten bereits von uns aufgegriffen.

Für die zukünftige Arbeit ist ein wichtiger Aspekt auch darin zu sehen, inwieweit man mit Hilfe der verschiedenen Modalitäten nicht nur neue verbesserte Leistungsparameter, sondern auch eine höhere Robustheit, Zuverlässigkeit oder Nachhaltigkeit erzielen kann? Wie Nützlichkeit ist eine technische Entscheidungsfindung (z.B. Klassifikation) insbesondere bei der Interaktion mit Menschen. Wie kann z.B. die Komplexität/Dimensionalität des Merkmalsraums bei einer Anzahl von Modalitäten ggf. reduziert und eine genügende Transparenz hinsichtlich der Ordnung und der Relationen von Daten und physikalischen Phänomenen derart erreicht werden, dass auch eine Sicherheitsanalyse möglich ist. Dies ist z.B. ein ganz wichtiger Aspekt gerade bei Themen, die mit Sicherheit für Menschen in den verschiedensten Bereichen zu tun haben (z.B. das Sicherheitspersonal am Flughafen im LaotSe-Projekt). Durch die technologische Fortentwicklung verschieben sich die Modalitäten auch ständig zwischen den Ebenen Hardware, Software und Algorithmik. In Multicam entstand die Multimodalität auf der Kameraebene und wurde durch die sich verändernden Möglichkeiten der Chipebene beeinflusst. Abgesehen von den Möglichkeiten der Vorverarbeitung bei den Sensor-Chips und der Kameraelektronik ergeben sich die Vorteile der Multisensorik/Multimodalitäten durch eine geeignete Algorithmik bei der Bilddatenverarbeitung. Aus diesem Grund wurden auch Untersuchungen z.B. zu Klassifikationsverfahren unter Berücksichtigung der Kamerasystemanalyse und -optimierung durchgeführt.

Durch die Betrachtung von Ortssignalen in Verbindung mit Sensornetzen spannt sich ein deutlich wachsender Merkmalsraum auf Kosten einer entsprechend steigenden Komplexität auf. Die Verortung von Sensordaten und anderen damit zusammenhängenden relevanten Informationen werfen neue interessante Fragestellungen auf, die wir auf Basis der Arbeiten/Ergebnisse im Projekt Dyn3D z.B. in den unten aufgeführten Projekten bereits aufgegriffen haben.

Einer der wichtigen Fragestellungen bei der multimodalen und multisensoriellen Szenenanalyse ist die Frage nach der minimal notwendigen Auflösungen und Genauigkeiten der Sensoren bei einer zu erzielenden Detektivität/Sensitivität. Beispielanwendung ist zum der Straßenverkehr mit Fußgängern und unterschiedlichen Fahrzeugen.

Multimodale und Multispektrale Klassifikation und Tracking im Movedetect-Projekt.

Verbundprojekt „Sichere Detektion und Lokalisation in kritischen Arealen durch klassifizierende drahtlose Sensornetze“ (MOVEDETECT), 2010-2011, ZF-Projekt 668e

Hierbei handelt es sich um ein Verbundprojekt im Bereich der Zivilen Sicherheit (mit den Universitäten zu Lübeck und Passau sowie Industriepartnern, BMBF-Förderung über das BSI). Dabei können wir die Erfahrungen der Multimodalen Sensorsystemtechnik, der multispektralen Sensordatenanalyse und der auf diesen Daten basierenden Klassifikation in die räumlich-zeitliche Sensordatenanalyse für ein großes Sensornetzwerk einbringen. Auch in diesem Projekt zeigt sich, dass sich die erarbeiteten grundlegenden Technologien, Strategien und Ergebnisse z.B. auch auf verwandte Fragestellungen mit anderer Hardware-/Softwarebasis übertragen lassen.

Weiterentwicklung der Kameraelektronik zu einer Weitbereichskamera im Projekt LaotSe.
Hightech.NRW Projekt "Flughafen-Start- und Landebahnüberwachung durch multimodale, vernetzte Sensorik (LaotSe, (2009 - 2012), Vertragsnummer: w0805ht009b

Auf Basis der in MultiCam entwickelten Plattform ist in diesem Projekt die Entwicklung und Anwendung einer Weitbereichskamera eine wichtige Zielsetzung. Mit der Systemtechnik, die im Projekt in Verbindung mit der Industrie entwickelt wird, erhalten wir eine neue Plattform, mit der erstmals Bilder einer höheraufgelösten TOF-Kamera aus Entfernungen >100m aufgenommen werden können. Die Untersuchungen und Entwicklungen schließen die Plattform für netzwerkorientierte Multisensorik ein.

In diesem Projekt beschäftigen wir uns ergänzend mit den zusätzlichen Modalitäten therm. IR-Bilder und Radarbildern.

Compressive Sensing

In diesem Zusammenhang ist die Frage aufzugreifen, welche Auflösung wird für die Abstandsmessung in Lateraler Richtung benötigt. Da die Pixel bei höherer Auflösung kleiner werden, ist entweder die Empfindlichkeit entsprechend zu steigern oder die Entfernungreichweite wird eingeschränkt. Durch eine geschickte hardwarenahe Vorverarbeitung und Verfahren wie das compressiv sensing erscheint es uns in der Zukunft möglich, mit weniger aber genügend großen Pixeln Reichweite und Auflösung zu steigern. Erste, Ergebnisse (z.B. zu Super Resolution) zu den begonnen Arbeiten wurden bereits veröffentlicht /6/, /7/.

Datenbank mit TOF-Kamerabildern und registrierten 2D-Bildern

Als ein wesentliches Hindernis für den internationalen Austausch innerhalb der Community ist das generelle Fehlen gemeinsam nutzbarer Datenbestände zum Vergleich verschiedener 2D/3D Segmentierung- und Klassifizierungsverfahren aufgefallen. Eine freie Verfügbarkeit von Datenbeständen generischer Objekte (ähnlich COIL100³ oder HumanEva⁴) wäre hier von großem Nutzen. Es ist vorstellbar, eine entsprechende Webseite zur nicht kommerziellen Nutzung derartiger Daten zur Verfügung zu stellen. Konzepte und Überlegungen hinsichtlich geplanter Systematiken und eines einheitlichen Datenformats sind noch in der Design-Phase.

3.5 Interdisziplinäre Weiterentwicklung

Wir haben die Erfahrung gemacht, dass die Forschungsergebnisse und Systemtechnologie in einer Reihe von unterschiedlichen Anwendungsgebieten auf großes Interesse stoßen. Dies ist auch der Grund der vorgesehenen Verwertungsstrategie.

Konkret wurden bisher Projektanträge mit Architekten und Bauingenieuren erarbeitet und ein Antrag in deren fachspezifischen Schwerpunkt gemeinsam gestellt.

In Verbindung mit Multisensorlösungen z.B. für Sicherheitslösungen besteht das Interesse an einer gemeinsamen Weiterentwicklung für neue Anwendungsbereiche. Die Projekte LaotSe (2010-2012) POmSe (2009-2010) und Movedetect (2010-2011) haben von den Vorarbeiten und den parallel ablaufenden grundlegenden Analysen in Dyn3D deutlich profitiert. In diesen Projekten sind auch unterschiedliche Fachdisziplinen beteiligt.

3.6 Verwertungspotenzial

Die Ergebnisse haben gezeigt, dass das Konzept und die mehrfach weiterentwickelte und optimierte Elektronik erhebliche Vorteile in der Anwendung bringen. Durch die verschiedenen Modalitäten kann die Kamerasensorik zum einen wie eine bekannte Standardkamera verwendet werden. Bei Bedarf können die Abstandsdaten bzw. die RRI-Daten zum Farbbild hinzu genommen werden. Ein flexibles Beleuchtungskonzept ermöglicht die Anpassung an verschiedene Applikationen. Die Hürde und das notwendige Verständnis für den Einsatz der neuen 3D-Kameras in verschiedenen Anwendungsgebieten kann deutlich verringert werden. Die erzielten Optimierungen und Anpassungen sind für den Echtzeitbetrieb geeignet und bilden eine sehr interessante Lösung für vielfältige Anwendungen.

Ein weiterer Aspekt ist die Möglichkeit, Multispektrale Bilder, mit unterschiedlichen Ortsauflösungen zu den Zeitpunkten aufzunehmen, in denen die Eigenbeleuchtung abgeschaltet ist. Dies ist ein Effekt, der auch für die Hintergrundschätzung verwendet wird.

Durch die rekursive Phasenbildanalyse in der Kameraelektronik kann in Bildechtzeit eine neue Abstandsberechnung durchgeführt werden. Kombiniert mit dem Verfahren zur eindeutigen Phasenzuordnung (Phase-Unwrapping) und der Unterdrückung der Bewegungsartefakte in Verbindung mit den hochaufgelösten 2D-Daten ergeben sich robuste, hochwertige Informationen des Objektbereichs (Szenenbereichs) in vielen Anwendungen. Insbesondere bei unstrukturierten und nicht definierten Szenen und größeren Entfernungen hat die Technologie große Vorteile.

³ Columbia Object Image Library (COIL-100), S.A. Nene, S.K. Nayar and H. Murase, Technical Report CUCS-006-96, February 1996.

⁴ L. Sigal, A. Balan and M. J. Black. HumanEva: Synchronized Video and Motion Capture Dataset and Baseline Algorithm for Evaluation of Articulated Human Motion, In International Journal of Computer Vision, Vol. 87 (1-2), 2010.

Es besteht das Interesse eines Firmenkonsortiums, die Technologie für ein Standardprodukt zu verwenden. Derzeit laufen verschiedene Gespräche sowie eine Marktanalyse und es wird die Verifikation verschiedenen Details durchgeführt.

3.7 Beteiligte Wissenschaftler

Die Komplexität und Vielfältigkeit der Arbeiten (Elektronik, Optik, Optomechanik/Optoelektronik, Signalauswertung Datenanalyse, -vorverarbeitung der Kamera und Echtzeit-Auswertestrategien für spezifische Test- bzw. Anwendungsfälle machte die Zusammenarbeit unterschiedlicher Wissenschaftler der Arbeitsgruppe notwendig um das jeweilige Know-how bereitstellen bzw. einsetzen zu können.

Es waren keine externen Wissenschaftler am Projekt beteiligt. Die enge Zusammenarbeit mit den **Teilprojekten 2D3DProc und 3DTHz** waren zum einen bei den Fragen zur Datenverarbeitung und Kalibrierung sowie der Erarbeitung spezieller Messtechniken sehr wichtig. Das enge Feedback bezüglich der unterschiedlichen Anwendungs- und Nutzungsfragen mit den anderen Teilprojekten hat wesentlich dazu beigetragen, den Fokus immer wieder auf relevante Detailprobleme zu legen.

Name	Funktion	Beschreibung der Projektbeteiligung
Loffeld, Otmar, Prof. Dr.-Ing.	Leiter	Projektleitung Modellentwicklung für die Sensorik/Sensordatenerfassung, und Sensordatenauswertung, Projektleitung
Hartmann, Klaus, Dr.-Ing.	Gruppenleitung	Systemkonzeption, Systemanalyse, Rekursive Phasenbildauswertung 2D/3D-Auswertestrategie, Beleuchtung, Projektleitung
Weih, Wolfgang, Dr. rer. nat.	Gruppenleitung	Systemkonzeption, Optik, Hardwareentwicklung, Optimierung FPGA-Auswertung, Projektleitung
Lottner, Oliver, Dipl. Ing.	Wiss. MA	Olliver Lottner hat viele Experimente hinsichtlich der Kameraanordnungen und der verschiedenen Auswertestrategien um ein optimiertes Systemkonzept zu erreichen und ein besseres Verständnis für die Modellbildung zu gewinnen..Neben der Modellbildung war die Datenvorverarbeitung und die Echtzeitaspekte Schwerpunkt seiner Arbeiten.
Syed Ghobadi,	Wiss. MA	Syed Ghobadi hat sich mit der Datenfusion, Klassifikation und dem Tracking auf Basis der 2D-Farbildaten und der RRI-Daten. Mit den Auswertansätzen konnten die Einflüsse der Systemtechnik und der verschiedenen Vorverarbeitungsansätzen im Detail analysiert werden. Die Arbeiten bildeten eine gute Verbindung zu den Arbeiten z.B. in 2D/3D- PROC.

Löprrich, Omar, Dipl.-Phys.	Wiss. MA	Omar Löprrich hat zeitweise am Projekt mitgearbeitet und Beiträge hinsichtlich Mehrkameralösung, Kameranetzwerk und Echtzeitanwendungen geliefert. Echtzeitanwendung
Han, Juanjuan	Wiss. MA	Juanjuan Han hat dazu beigetragen die Datenmöglichkeiten der Kamera hinsichtlich neuer Auswerteverfahren, wie das Compressiv-Sensing zu analysieren und anzupassen sowie zur Auswertung der Kameradaten.

4 Liste der im Teilprojekt entstandenen Publikationen

4.1 Begutachtete Publikationen

4.1.1 Wissenschaftliche Zeitschriften

4.1.2 Fachkonferenzen

- /1/ O. Lottner, W. Weihs, K. Hartmann: "Systematic Non-Linearity of Multiple Distributed Illumination Units for Time-of-Flight (PMD) Cameras", International Conference on Systems 2008, Greece. ISBN: 9789606766831
- /2/ O. Lottner, W. Weihs, K. Hartmann: "Time-of-Flight Cameras with Multiple Distributed Illumination Units", Proceedings of the 8th Conference on Signal Processing, Computational Geometry and Artificial Vision, ISBN: 978-960-6766-95-4 pp. 40-45, 2008
- /3/ S. E. Ghobadi, O. Loepprich, O. Lottner, K. Hartmann, O. Loffeld, W. Weihs: "Improved Object Segmentation Based on 2D/3D Images", 5th IASTED International Conference on Signal Processing, Pattern Recognition and Applications, pp. 42-47, 2008, ACTA Press, Innsbruck, Austria, ISBN: 978-0-88986-717-8
- /4/ Lottner, O.; Weihs, W.; Hartmann, K.: "Multi-Lighting Aspects for ToF Cameras", EOS Conference on Frontiers in Electronic Imaging, Munich, Germany, June 15.-16., 2009, pp. 72-73.
- /5/ B. Langmann, K. Hartmann, O. Loffeld: „Depth Assisted Background Subtraction for Color Capable ToF-Cameras”, in: Proceedings of the 2010 International Conference on Image and Video Processing and Computer Vision (IVPCV-10), 2010, pages 75-82
- /6/ Han, J.; Loffeld, O.; Hartmann, K.; Wang, R.: "Multi image fusion based on compressive sensing", IEEE/IET International Conference on Audio, Language and Image Processing, Shanghai, China, Nov. 23-25, 2010, pp. 7.
- /7/ Han, J.; Hartmann, K.; Loffeld, O.; Wang, R.: "Moving object extraction using 2D/3D images", International Conference on Machine Vision, HongKong, China, 2010, pp. 5.
- /8/ V. Peters, S. Meckel, O. Loffeld, O. Lottner: "Modeling and Simulation of Light Propagation for Multistatic 3D PMD-Camera and Illuminator Constellations", Eurosime 2010

4.2 Andere Veröffentlichungen

- /9/ O. E. Löprrich, S. E. Ghobadi, O. Lottner, K. Hartmann, O. Löffeld: "Aspects of a Distributed Surveillance System Based on 2D/3D Image Data", internal report, University of Siegen, 2007
- /10/S. E. Ghobadi: "Real Time Object Recognition and Tracking Using 2D/3D Images", Dissertation, University of Siegen, 2010
- /11/O. Lottner: "Evaluation of Different Sensor / Sensor- and Illumination Configurations also in Dynamic Environments", Dissertation, University of Siegen, expected to be published in 2011

4.3 Patente

-keine-

5 Publikationen Dritter

An dieser Stelle sind die Veröffentlichungen der Partner im Projekt nicht referenziert, da auf die Arbeiten der Projektpartner an anderer Stelle im Dokument bzw. in den Berichten der Partner hingewiesen wird.

- /12/Erz, M., Jähne, B. "Radiometric and Spectrometric Calibrations, and Distance Noise Measurement of ToF Cameras", 3rd Workshop on Dynamic 3-D Imaging, Lecture Notes in Computer Science, 5742, 28-41, Springer, Berlin 2009
- /13/A.D. Payne, A.A. Dorrington and Michael J. Cree, "Illumination waveform optimization for time-of-flight range imaging cameras," Videometrics, Range Imaging, and Applications XI, Fabio Remondino, Mark R. Shortis Editors, Proc. SPIE 8085, 80850D (2011).
- /14/Payne, A.D., Dorrington, A.A., Cree, M.J. & Carnegie, D.A. (2010). Improved measurement linearity and precision for AMCW time-of-flight range imaging cameras. Applied Optics, 49(23), 4392-4403
- /15/Godbaz, J.P., Cree, M.J., Dorrington, A.A. & Payne, A.D. (2009). A fast Maximum Likelihood method for improving AMCW lidar precision using waveform shape. In proceedings of IEEE 2009 Sensors Conference, 25-28 Oct 2009, Christchurch, New Zealand. (pp. 735-738). Washington, USA
- /16/Büttgen, B. and Seitz, P., "Robust Optical Time-of-Flight Range Imaging Based on Smart Pixel Structures," IEEE Transactions on Circuits and Systems I: Regular Papers, 55, 1512-1525 (2008).
- /17/Valls Miro, J.; Dissanayake, G., Robotic 3D visual mapping for augmented situational awareness in unstructured environments, International Workshop on Robotics for Risky Interventions and Surveillance of the Environment (RISE'08)
- /18/Huhle, B.; Fleck, S.; Schilling, A., Integrating 3D Time-of-Flight Camera Data and High Resolution Images for 3DTV Applications, 3DTV CON - The True Vision, 2007.
- /19/Reulke, R., Combination of Distance Data with High Resolution Images, ISPRS, Commission V Symposium, Image Engineering and Vision Metrology, 2006.
- /20/Santrac, N.; Friedland, G., Rojas, R.: High Resolution Segmentation with a Time-of-Flight 3D-Camera using the Example of a Lecture Scene, Department of Computer Science, Freie Universität Berlin, 2006.
- /21/G. Ye: "Image Registration and Super-Resolution Mosaicing", Doctoral Thesis, University of New South Wales, Australia, 2005

6 Anhang

Das Kameramodell

In den folgenden beiden Abschnitten wird das Kameramodell unserer MultiCam zusammengefasst. Dies umfasst auf der einen Seite die räumliche Fusion und insbesondere die intrinsischen Parameter, um die Alltagstauglichkeit eines einfachen Lochkameramodells im Vergleich mit einem kommerziellen System zu zeugen. Auf der anderen Seite ergeben sich durch den speziellen Aufbau bedingt verschiedene zeitliche Synchronisierungsmöglichkeiten.

6.1 Intrinsischen Parameter

Die Modellierung der Kamera erfolgt nach dem Lochkameramodell mit radialen und tangentialen Verzerrungen zweiter Ordnung. Die intrinsischen Parameter dieses Modells beziehen sich im Falle der MultiCam auf die Sensoren, das Sicherheitsglas auf dem PMD-Sensor, dem IR-Filter vor dem 2D-Sensor, dem Strahlteiler, dem IR-Filter hinter der Linse und dem Objektiv. Eine Änderung eines Parameters bedingt eine Neukalibrierung. Die Anwendbarkeit des Modells auf ToF-Kameras wurde im Rahmen des Projekts durch Vergleich mit einer Standard-Industriekamera (Basler A601f) mit gleichem 2D-Sensor gezeigt (nur der innere optische Aufbau differiert in diesem Fall). Das Ergebnis ist in nachstehender Tabelle zu sehen. Der Bildmittelpunkt ist im Falle der Basler-Kamera in X-Richtung deutlich näher am erwarteten Mittelpunkt (halbe Pixelanzahl), was sich durch die automatisierte Montage erklärt. Die Brennweite ist vergleichbar, und die Verzerrungen sind nur geringfügig bedingt durch den komplizierteren inneren Aufbau größer. Die MultiCam verhält sich also sehr ähnlich einer kommerziellen Kamera, für die das Lochkameramodell schon sehr häufig erfolgreich angewendet wurde.

Objektiv	Pentax H1212, f=12mm, F=1:1,2		
Kamera	Basler A601f	MultiCam 2D/3D C-Mount	
Sensor	Micron MTV903A with 640*480 pixels of 9,9*9,9 μ m ²	PMD 3kS with 64*48 pixels of 100*100 μ m ²	
Brennweite	$f_x = 1241.3$; $f_y = 1239.6$	$f_x = 1206.4$; $f_y = 1208.3$	$f_x = 1206.4$; $f_y = 1208.3$
Mittelpunkt	$c_x = 314.0$; $c_y = 219.3$	$c_x = 253.4$; $c_y = 311.4$	$c_x = 253.3$; $c_y = 327.4$
Radiale Verzerrung	$k_1 = -0,057$; $k_2 = 0,136$	$k_1 = -0,365$; $k_2 = -0,228$	$k_1 = -0,363$; $k_2 = -0,243$
Tangentiale Verzerrung	$p_1 = 0,002$; $p_2 = 0,001$	$p_1 = -0,002$; $p_2 = 0,000$	$p_1 = 0,001$; $p_2 = 0,000$

Tabelle 6.1: Vergleich Kalibrierung MultiCam – Industriekamera (Sensor um 90° gedreht)

6.2 Fusion der Daten

Zeitliche Fusion

In der MultiCam übernimmt eine dediziert Elektronik (Field Programmable Gate Array, FPGA) die präzise Synchronisierung der 2D- und der 3D-Bildaufnahme. Die Synchronisierung kann sich dabei entweder auf den Start- oder den Endzeitpunkt der Integration beziehen. Im Standardfall wird ein 2D-Bild mit dem ersten Phasenbild getriggert und parallel zu den vier Phasenbildern aufgenommen. Es ist jedoch bei Bedarf möglich, z.B. ein 2D-Bild mit jedem Phasenbild aufzunehmen. Die nachstehende Abbildung zeigt die verschiedenen Möglichkeiten.

Der binokulare Aufbau mit zwei unabhängigen Kameras ist weder so flexibel noch so präzise. Fast keine kommerziell verfügbare ToF-Kamera verfügt überhaupt über einen externen Trigger, so dass eine Synchronisierung nur vage möglich ist, und selbst wenn, sind die genauen Startzeitpunkte der Integration nicht bekannt, woraus sich Zuordnungsprobleme in dynamischen Umgebungen ergeben können.

Räumliche Fusion

Die räumliche Registrierung der Bildinhalte beider Sensoren schließlich muss im binokularen Aufbau dadurch erfolgen, dass die Daten des 3D-Sensors zunächst von Bild- nach Weltkoordinaten umgerechnet werden müssen. Auf Grund der Messunsicherheit und der geringen lateralen Auflösung des 3D-Sensors ergibt sich hier eine – je nach Beleuchtungsverhältnissen – relativ große Unsicherheitszone, innerhalb der der genaue Punkt nicht bestimmt werden kann. Dieser Bereich muss schließlich auf den 2D-Sensor zurückprojiziert werden. Somit bestehen folgende Nachteile: zum einen muss die Registrierung für jede Kombination von Distanz- und 2D-Bild neu berechnet werden, zum anderen ist eine solche Berechnung nur bei zuverlässigen Distanzdaten möglich, die aber in sehr vielen Fällen nicht vorliegen, und zwar insbesondere in Outdoor-Szenarien.

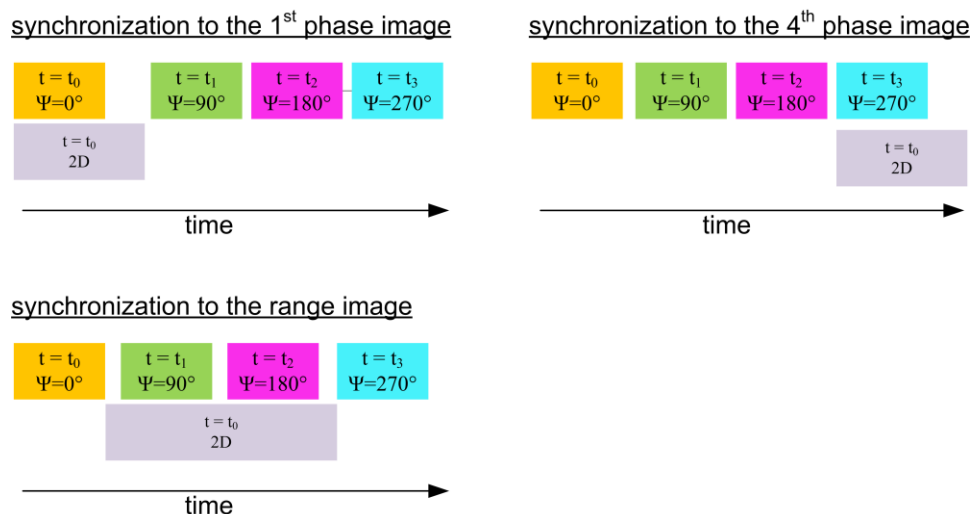


Abb. 6.1: Zeitliche Fusion der Sensoren der MultiCam

In der MultiCam teilen sich die beiden Sensoren ein und das gleiche Objektiv sowie die extrinsische Kalibrierung, so dass das gleiche Sichtfeld abgebildet wird. Auf diese Weise ergibt sich der Vorteil, dass die Registrierung nur aus einer leichten Verschiebung in x- und y-Richtung besteht, die unabhängig vom Bildinhalt immer gleich ist und somit leicht vorab berechnet werden kann und z.B. als Look-Up-Table immer wieder angewendet werden kann. Insbesondere sei darauf hingewiesen, dass die Registrierung nicht von der Distanz zur Szene abhängig ist, so dass genaue Tiefendaten keine Voraussetzung sind! Natürlich bedingt dies eine sorgfältige Justierung beider Sensoren insbesondere unter Vermeidung von Winkelfehlern. Diese ist mit den in der Konstruktion vorgesehenen Stellschrauben in der Praxis gut erreichbar. Eine automatische Justierung mittels eines mechanischen Anschlags oder einer µm-Positioniereinheit ist jedoch z.Zt. nicht vorgesehen, da es sich um einen Prototypen und nicht um ein Funktionsmuster handelt.