

# Kontinuierliche Echtzeit 3D-Modell Rekonstruktion im Kontext von kollaborativen Mixed-Reality-Anwendungen

Iwer Petersen

HAW Hamburg, Technik und Informatik,  
Berliner Tor 7, Hamburg, Germany  
iwer.petersen@haw-hamburg.de  
<http://www.haw-hamburg.de>

**Zusammenfassung.** In dieser Arbeit werden einige exemplarische Arbeiten die sich mit der 3D Echtzeit-Rekonstruktion und/oder der Rekonstruktion von 3D Modellen von Menschen beschäftigen. Unterschiedliche Herangehensweisen werden untersucht und im Hinblick auf die eigenen Arbeit bewertet.

**Schlüsselwörter:** Mixed Reality, 3D Scanning, 3D Rekonstruktion

## 1 Einleitung

Üblicherweise werden Teilnehmer in verteilten Virtual- und Mixed-Reality Umgebungen durch animierte, virtuelle Avatare repräsentiert, welche durch Skelettanimation bewegt werden. Zwar kann mittels Motion Capture Technologie die grobe Bewegung eines realen Menschen erfassen, und damit einen virtuellen Avatar steuern, jedoch fallen sie trotzdem leicht dem Phänomen des "Uncanny Valley" zum Opfer. Ein in Echtzeit eingescanntes 3D Modell der Teilnehmer könnte dieses Problem umgehen. Um das zu erreichen muss die benötigte Zeit zur Erstellung eines 3D Modells aus Scannerdaten soweit reduziert werden, dass sie wiederholt zu interaktiven Frameraten ausgeführt werden kann.

### 1.1 Ziele

Ziel dieser Ausarbeitung ist, verschiedene 3D Scan-Systeme, die auf eine Echtzeitrekonstruktion von 3D Modellen von Menschen hinarbeiten, auf ihre Methodiken zu untersuchen. Dabei soll ein Verständnis für geeignete Sensor-komponenten sowie Algorithmen zur Punkt- und Polygonmeshgenerierung erlangt werden. Desweiteren soll ein Einblick Parallelisierungsmöglichkeiten einzelner Arbeitsschritte innerhalb der 3 Teilaufgaben Sensordatenverarbeitung, Punktgenerierung und Meshgenerierung erhalten werden.



Abb. 1. Verarbeitungsschritte beim 3D Scan. Quelle: eigenes Bild

### 1.2 Aufbau

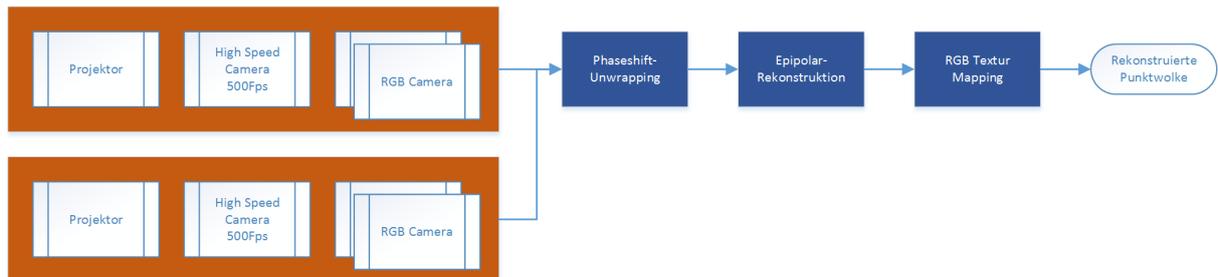
Im weiteren wird zunächst ein Structured-Light Scansystem betrachtet welches detaillierte 3D Informationen in Echtzeit liefert (Kapitel 2). Danach werden zwei Personenscanner betrachtet von denen ersterer eine höhere Qualität erreicht (Kapitel 3) und zweiter 3D Modelle in Echtzeit liefern kann (Kapitel 4).

## 2 Real-Time active Multiview 3D Reconstruction

In diesem Kapitel wird ein Echtzeit 3D-Scansystem untersucht, welches vor allem durch Performance hervorsteht. Ein vergleichsweise hohes Datenaufkommen wird hier dadurch bewältigt, dass die berechnungsintensiven Aufgaben auf Graphikprozessoren implementiert wurden.

## 2.1 Übersicht und Ziele

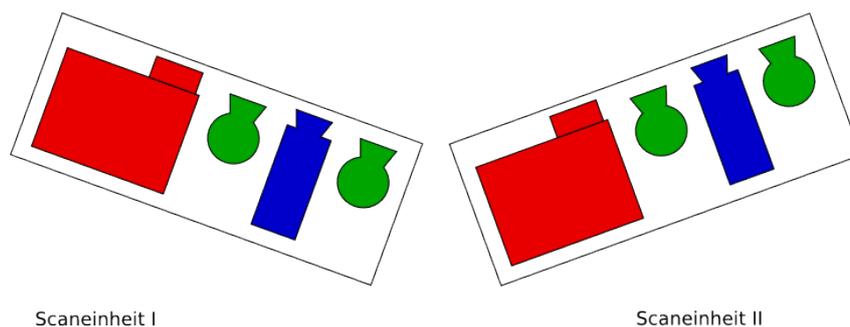
Das Ziel der Arbeit von Kai Ide und Thomas Sikora ([IS12]) war, ein System zu entwickeln, welches zeitlich variierende Geometrie in Echtzeit mit möglichst hoher Vollständigkeit und Genauigkeit scannen kann. Bei dem entwickelten System handelt sich um ein aktives Scansystem nach dem Prinzip der Structured Light Scanning Technik.



**Abb. 2.** Verarbeitungsschritte des 3D-Scansystems von Ide und Sikora. Quelle: eigenes Bild

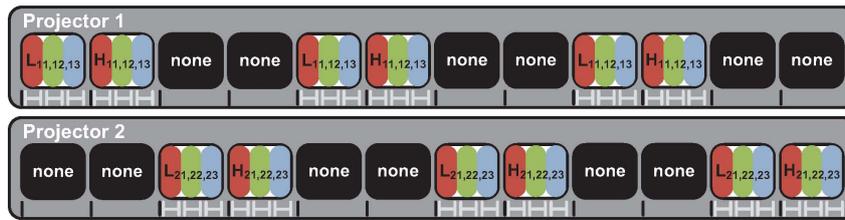
## 2.2 Scanner Komponente

Das Besondere an dem entwickelten System sind die Scaneinheiten, die je aus einem Videoprojektor, einer 500Hz Kamera sowie zwei 1.3 Megapixel RGB-Kameras bestehen (siehe Abbildung 3). Hierbei werden die Daten der Highspeed Kameras für die Punktgenerierung verwendet während die RGB-Kameras für die abschließende Farbgebung vorgesehen sind. Bei zwei verwendeten Scaneinheiten können stereoskopische Bilder von 4 unterschiedlichen Basislinien simultan aufgenommen werden, da beide Highspeedkameras je die Projektionen von den beiden abwechselnd projizierenden Projektoren aufnehmen.



**Abb. 3.** Schematischer Aufbau der Scaneinheiten mit Projektoren (rot), Highspeedkameras (blau) und RGB-Kameras (grün). Quelle: eigenes Bild

Um eine synchrone Aufnahme von Bildern zu gewährleisten, wird der Videostream, welcher an die Projektoren geschickt wird, mit einem Mikrokontroller auf regelmäßig eingestreute schwarze Frames überwacht. Diese dienen als Taktgeber für einen Kamera-Auslöse-Trigger. So wird eine präzise zeitliche Steuerung (siehe Abbildung 4) aller Komponenten gewährleistet. Die Highspeedkameras sind dabei so gesteuert, dass sie die von den Projektoren nacheinander ausgehenden Farbkomponenten (rot, grün, blau) einzeln erfassen können.



**Abb. 4.** Zeitliche Steuerung von zwei Scaneinheiten. Synchronisierungsframes sind schwarz. Schwarze Skalenstriche kennzeichnen Videoframes, graue Intervalle kennzeichnen Highspeedkamera-Frames. Quelle: [IS12]

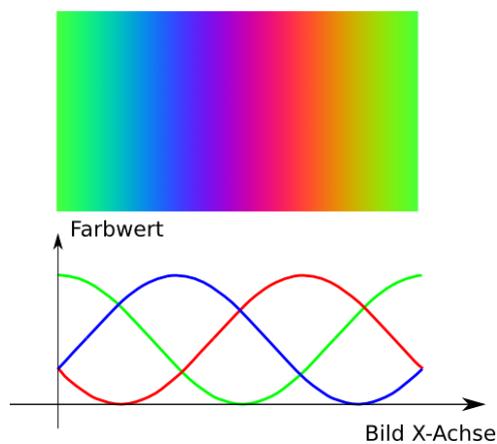
### 2.3 Kalibrierung

Zum Erlangen der Kalibrierungsparameter der vielen verschiedenen Kameras haben Ide und Sikora eine eigene Lösung entwickelt (siehe [ISS09]), welche Pixel-zu-Pixel Korrespondenzen findet indem in den Kamerabildern Features eines von einem Projektor projizierten Musters findet. Ihre Methode basiert auf der Arbeit von Svoboda et. al. ([SMP05]) welche auch in Abschnitt 4.3 als Basis dient.

### 2.4 Structured Light Pattern

Über die Projektoren werden farbige, sinusoidale Patterns auf die zu scannende Szene projiziert. Die Patterns bestehen aus horizontalen Farbverläufen in welchen die Rot-, Grün- und Blaukomponenten gegeneinander phasenverschoben, sinusförmig auf die X-Achse des projizierten Bildes abgebildet sind (siehe Abbildung 5). Die Lichtkodierung die hier verwendet wurde, wird als Phase-Shift-Pattern bezeichnet ([SPB04]).

Hierbei werden von den Projektoren abwechselnd ein niederfrequentes und ein hochfrequentes Pattern projiziert. Das hochfrequente Muster erlaubt eine höhere lokale Scanauflösung, wobei das niederfrequente Muster für die globale Positionszuordnung benötigt wird. Dafür bildet das niederfrequente Muster eine komplette Sinusperiode von  $2\pi$  auf die Breite des Projektorbildes ab. Das hochfrequente Muster wird hingegen mit einem ganzzahligen Vielfachen dieser Frequenz abgebildet.



**Abb. 5.** Ein niederfrequentes Phase Shift Pattern. Farben sind um  $\frac{2}{3}\pi$  gegeneinander phasenverschoben. Quelle: eigenes Bild

## 2.5 Phaseshift Unwrapping

Die Highspeedkameras liefern so für jedes der beiden projizierten Patterns ein Triplet von Bildern der Rot-, Grün- und Blaukanäle. Geleitet durch das jeweilige niederfrequente Pattern kann das sich wiederholende, hochfrequente Pattern in X-Richtung global aufgelöst werden. Dieses Verfahren ist als Phase-Shift-Unwrapping bekannt (siehe [GP]). Aus den beiden Bild-Triplets sowie den a-priori bekannten Projektionsmustern entsteht so ein stereoskopisches Bildpaar welches zur Epipolarrekonstruktion verwendet werden kann.

## 2.6 Punktgenerierung

Über die genaue Kenntnis der Zuordnung von Pixelkoordinaten der Projektoren zu Pixelkoordinaten der Kameras, welche über das Kalibrierungsverfahren erhalten werden (Abschnitt ??), können nun mittels epipolarer Triangulation 3D Punkte aus den Kamerabildern rekonstruiert werden.

## 2.7 Kolorierung

Abschließend werden mit Hilfe der Bilder der RGB-Kameras die 3D-Punkte mit Farbinformationen versehen, indem die Bilder der RGB Kameras mit Hilfe der Kalibrierungsparameter auf die Punkte projiziert werden.

## 2.8 Zusammenfassung

Das entwickelte System produziert 5.3 Millionen kolorierte 3D Punkte bei einer Framerate von 10Hz und erlaubt so die dreidimensionale Rekonstruktion feinsten Details. Durch die Verwendung von mehreren Scaneinheiten kann die Vollständigkeit der Abdeckung deutlich erhöht werden (siehe Abbildung 6). Um ein möglichst hohe Performance zu erreichen wurden sämtliche Berechnung als per-Pixel Operationen auf den Texturen der Kameras modelliert, welche auf einem Graphikprozessor durchgeführt werden.



**Abb. 6.** Rekonstruierte Punktwolken unter Benutzung von 2 (links) und 4 Scaneinheiten (rechts). Quelle: [IS12] (Ausschnitt)

Obwohl das entwickelte System nicht das Ziel hatte ein Polygonmesh zu erzeugen, zeigt es doch, wie wesentliche Verarbeitungsschritte auf dem Weg dahin effizient durch geeignete Modellierung beschleunigt werden können. Für eine Echtzeit-Rekonstruktion von Menschen wäre ein solches System allerdings trotzdem als eher ungeeignet zu bezeichnen, da der dauerhafte rapide Wechsel des sichtbaren Lichts für den Menschen sehr unangenehm werden dürfte.

## 3 Scanning 3d full human bodies using kinects

Die in diesem Kapitel untersuchte Arbeit zeigt exemplarisch die Verwendung mehrerer, zeitlich aufeinander folgenden Frames multipler Tiefenbildkameras zur Rekonstruktion eines vollständigen 3D Modells von Menschen. Zum einem ist hier die Vereinigung der Daten mehrerer Kameras interessant, zum anderen die Integration von zeitlich aufeinander folgender Frames. Beide Aufgaben bringen hier ihre eigenen Probleme mit sich.

### 3.1 Übersicht und Ziele

Das Ziel dieser Arbeit ([TZL<sup>+</sup>12]) war die Entwicklung eines Scanners mit Hilfe von Microsoft Kinect Tiefenbildkameras, welcher in der Lage sein sollte, ein vollständiges 3D Modell eines Menschen erstellen zu können. Diese Arbeit legt den Fokus auf die Polygonmeshgenerierung welche Bewegungen der zu scannenden Person während des Scanvorgangs algorithmisch ausgleichen muss, um ein geschlossenes Modell zu erhalten.

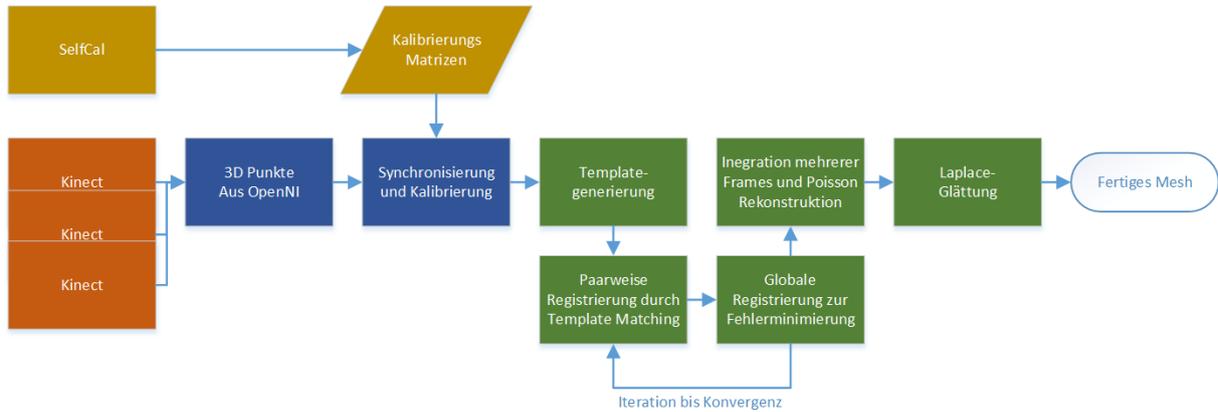


Abb. 7. Verarbeitungsschritte des 3D-Scansystems von Tong et. al. Quelle: eigenes Bild

### 3.2 Aufbau

Für den Scanvorgang steht die zu scannende Person auf einer sich langsam drehenden Plattform. An zwei Seiten der Plattform befinden sich Kinect-Kameras auf verschiedenen Höhen um Oberkörper, Unterkörper und Beine zu erfassen. Dadurch wird das Problem von Interferenzen zwischen den Structured-Light-Patterns verschiedener Kinects weitgehend umgangen. Die extrinsischen Parameter der Kameras, welche für die Überführung der Daten in ein gemeinsames Koordinatensystem benötigt werden, werden automatisch mit SelfCal ([Svo]) ermittelt. Innerhalb von 30 Sekunden dreht sich die Plattform um 360° während die Kameras simultan Bilder aufnehmen.

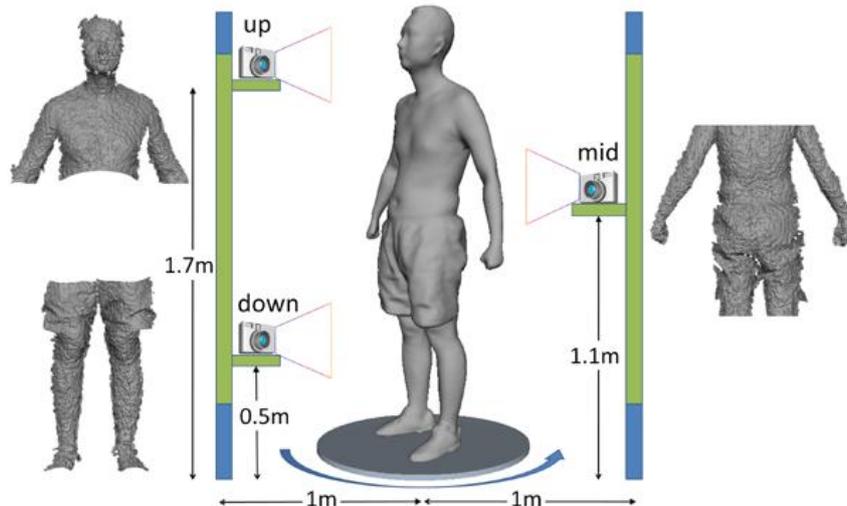


Abb. 8. Schematischer Aufbau der Scanvorrichtung. Quelle: <http://www.math.zju.edu.cn/ligangliu/CAGD/Projects/Kinects-CapturingHumans/>

### 3.3 Punktgenerierung

Über die OpenNI Middleware werden Tiefenbilder der Kameras geliefert. Ein auf die Tiefenbilder angewendeter Thresholding Filter hinterlässt, bedingt durch den Aufbau, lediglich Tiefenwerte von Teilen des Bildes, welche einen Menschen abbilden. Mit gegebenen intrinsischen Kameraparametern werden nun 3D Punkte aus diesen Tiefenwerten generiert. Mit Hilfe der extrinsischen Parametern werden die einzelnen Punktwolken in ein gemeinsames Koordinatensystem transformiert.

### 3.4 Sensordatenfusion

Durch die Aufnahmezeit von 30 Sekunden ist es sehr wahrscheinlich, dass leichte Bewegungen der zu scannenden Person zu Fehlern und Versätzen in sukzessiven Frames führen. Um diese Ungenauigkeiten auszugleichen wird mit dem ersten Frame ein sehr grobes Modell der Person aus 50-60 Vertizen erstellt, welches als Gerüst für nachfolgende Frames dient. Diese werden mit Hilfe von Iterative Closest Point (ICP) an dem Modell ausgerichtet. Dieses Verfahren ist als Punktwolken Registrierung bekannt ([BM92]). Das besondere an dem hier verwendeten Algorithmus ist, dass es sich um eine unstarre Registrierung handelt. Diese erlaubt eine Deformation des Meshes wenn dadurch eine bessere Ausrichtung erreicht werden kann.

### 3.5 Mesh Generierung

Nun werden die sukzessiven Frames der Aufnahme in ein einziges 3D Modell integriert. Die so entstandene Punktwolke wird mittels Poisson-Rekonstruktion zu einem 3D Polygongitter trianguliert. Schließlich wird das Modell einer Laplace-Glättung unterzogen, um eine glatte Oberfläche zu erhalten.

### 3.6 Zusammenfassung

Der hier entwickelte Scanner fertigt ein vollständiges, texturiertes 3D Modell eines Menschen in ungefähr 6 Minuten an. Das Verfahren bildet Oberflächenstrukturen mittlerer Größe überzeugend ab (siehe Abbildung 9, wobei feine Details vor allem durch den Glättungsfilter verloren gehen. Dies ist ein Kompromiss zwischen Oberflächendetails und Glätte, welcher abzuwägen ist.



**Abb. 9.** Rekonstruktionsergebnis als Komposition der Einzelmeshes (links) und als texturiertes 3D Modell in Vorder- und Rückansicht (rechts) Quelle: <http://www.math.zju.edu.cn/ligangliu/CAGD/Projects/Kinects-CapturingHumans/>

Während dieses System gut geeignet ist, personalisierte Avatare zu erstellen, welche in einer virtuellen Umgebung mittels Skelett kontrolliert werden würden, ist es offensichtlich nicht in der Lage 3D Modelle in Echtzeit zu liefern.

## 4 Real-time, full 3-D reconstruction of moving foreground objects from multiple consumer depth cameras

Auch in diesem Kapitel wird ein 3D Personenscanner betrachtet, der sich aber im Unterschied zu der im vorigen Kapitel betrachteten Arbeit auf die Integration einzelner Frames mehrerer Tiefenbildkameras beschränkt, um ein dynamisches Echtzeit 3D Modell des Menschen zu erhalten. Diese Arbeit ist besonders deshalb interessant, da sie interaktive Frameraten erreicht, indem sie die Zeit für die Erstellung eines 3D Modells so weit reduziert, dass sie für jeden Frame ausgeführt werden kann.

## 4.1 Übersicht und Ziele

In der Arbeit von Dimitrios S. Alexiadis, Dimitrios Zarpalas und Petros Daras ([AZD13]) sollte ein Personenscanner für Virtual-Reality Applikationen entwickelt werden, der ein Echtzeit 3D-Modell in einer Person liefern sollte. Hierbei sollten explizit kommerzielle Tiefenbildkameras zum Einsatz kommen. Im Vergleich zu der im Kapitel 3 beschriebenen Arbeit werden hier schon die separaten Punktwolken zu Polygonmeshes trianguliert. Dies erfordert einen größeren Aufwand bei der Zusammenführung der Meshes der einzelnen Kameras.

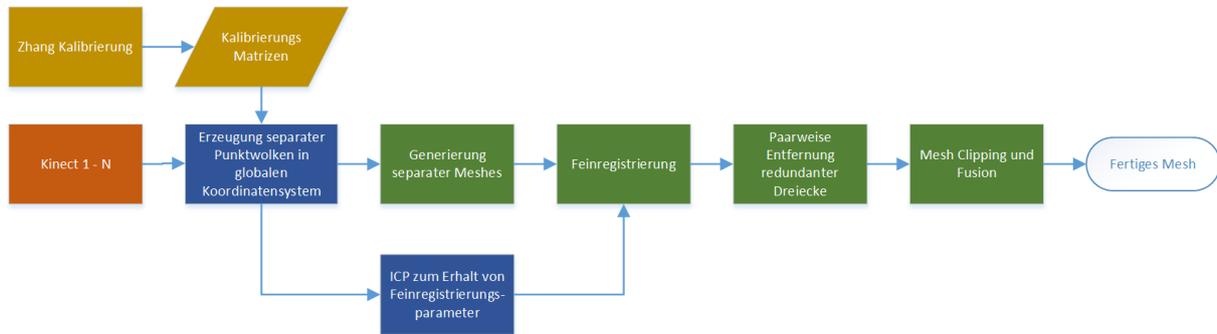


Abb. 10. Verarbeitungsschritte des 3D-Scansystems von Alexiadis et. al. Quelle: eigenes Bild

## 4.2 Aufbau

Das Scansystem besteht auf 4 Microsoft Kinect Kameras, welche 360° um ein Scanvolumen herum aufgebaut sind. Gegeben durch die Linsenparameter der Kameras ergibt sich ein Scanvolumen von  $1.4m * 1.4m * 1.2m$ , in dem ein Mensch ab Kniehöhe aus allen Richtungen erfasst wird.

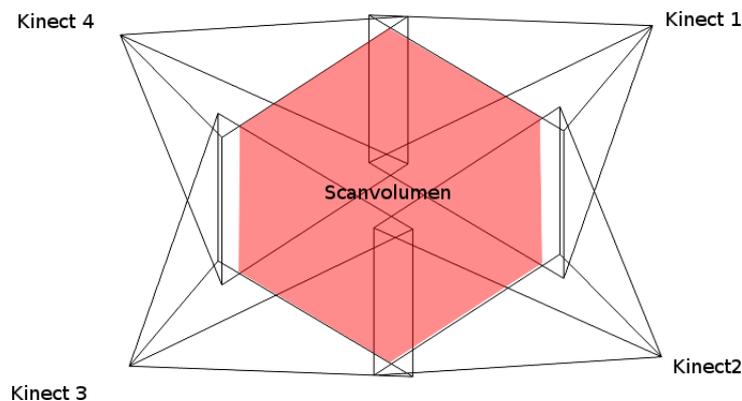


Abb. 11. Schematischer Aufbau. Quelle: eigenes Bild

## 4.3 Kalibrierung

Die intrinsischen Kalibrierungsparameter der Kinectkameras werden a-priori mit der Zhang-Methode (siehe [Zha00]) erhalten. Für die extrinsische Kalibrierung wird ein Kalibrierungstab mit definierter Länge mit LEDs an den Enden verwendet. Dieser wird im Scanbereich umhergeschwenkt, während die LEDs in den Kamerabildern detektiert werden, um Pixel-zu-Pixel Korrespondenzen zu erhalten, aus denen dann die extrinsischen Kameraparameter errechnet werden. Dieses Vorgehen kombiniert Ansätze die in [SMP05] und [KLB08] beschrieben werden.

## 4.4 Punktwolken Generierung

Wie schon in Kapitel 3.3 beschrieben, werden auch hier mit Hilfe der intrinsischen Parameter 3D-Punkte aus den Tiefenbildern generiert und mit Hilfe der extrinsischen Kalibrierungsparameter in ein gemeinsames Koordinatensystem transformiert. Die so zueinander ausgerichteten Punktwolken werden nun nebenläufig mittels ICP registriert um Parameter für eine spätere Feinausrichtung (siehe Kapitel 4.6) zu erhalten. Während diese Berechnung einige Frames lang dauert, werden die Punktwolken trianguliert.

## 4.5 Mesh Triangulation

Aus den separaten Punktwolken der einzelnen Kinectkameras werden nun Polygonmeshes trianguliert. Hierbei werden mittels Schwellenwerten Punkte mit Tiefenwerten außerhalb eines definierten Scanbereiches ( $[z_{min}, z_{max}]$ ) verworfen. In X-Y Ebene benachbarte Punkte werden so lange als verbunden betrachtet, wie ihre Tiefenwerte dicht genug beieinander liegen (Step-discontinuity constraint Triangulation [HSIW96]). So werden zum einen die Hintergründe aus den Punktwolken entfernt und zum anderen Teile des Meshes, welche aufgrund stark voneinander abweichender Tiefenwerte als nicht verbunden betrachtet werden, voneinander separiert.

## 4.6 Sensordatenfusion

Mit Hilfe der durch die nebenläufige ICP Registrierung erhaltenen Transformationsparameter werden die separaten Meshes in Position und Orientierung im gemeinsamen Koordinatensystem ausgerichtet, und in ein einziges Modell überführt.

## 4.7 Bereinigung

Das so erhaltene Modell ist im Prinzip überdefiniert, da einige Teile der zu scannenden Person sehr wahrscheinlich von zwei oder mehreren Kameras erfasst werden. In einem weiteren Schritt werden deshalb Dreiecke entfernt, welche sich mit anderen Dreiecken überlappen, wodurch allerdings Löcher im Polygonmesh entstehen. Diese Löcher werden anschließend durch eine vereinfachte Re-Triangulation der Rand-Polygone geschlossen. Anstatt dies im dreidimensionalen Raum zu berechnen werden die entsprechenden Polygone auf eine geeignete Ebene projiziert, so dass eine deutlich schnellere 2D-Triangulation vorgenommen werden kann.

## 4.8 Zusammenfassung

Das hier vorgestellte System erreicht eine Rekonstruktionsrate von knapp 10Hz unter Verwendung mehrerer CPU Kerne und einer GPU. Auch hier werden wieder einige Beispiele Möglichkeiten zur Beschleunigung von Algorithmen mittels Grafikkarten veranschaulicht. Im Vergleich zur Scanqualität der vorigen vorgestellten Arbeiten fällt die Scanqualität signifikant geringer aus.

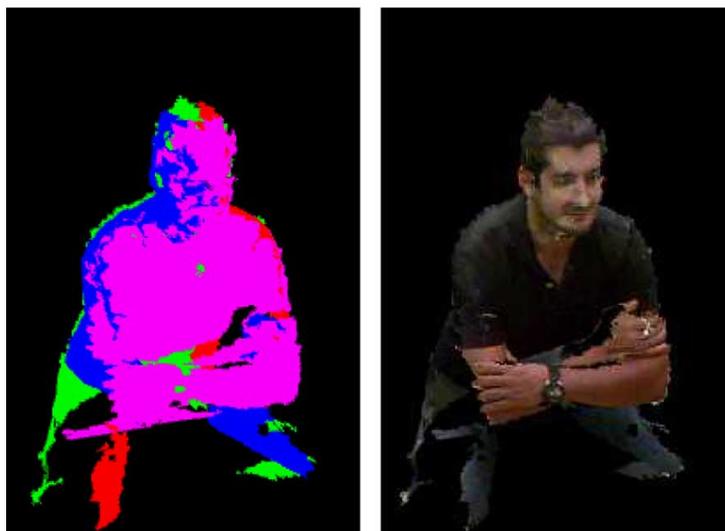


Abb. 12. Rekonstruktionsergebnis als Komposition der Einzelmeshes (links) und als texturiertes 3D Modell. Quelle: [AZD13]

## 5 Bezug zur eigenen Arbeit

In diesem Kapitel sollen die vorangegangenen Untersuchungen in einen Bezug zur eigenen Arbeit gesetzt werden.

## 5.1 Ide und Sikora

Obwohl der generelle Aufbau aufgrund der Verwendung von sichtbarem Licht recht ungeeignet ist um Personen dauerhaft zu scannen, bietet die Arbeit einige Anregungen. Vor allem die komplette Implementierung der verwendeten Algorithmen als Per-Pixel Operation auf den von den Kameras gelieferten Bildern scheint ein recht guter Ansatz zu sein um dem zu erwarteten Rechenaufwand Herr zu werden. Auch bei der eigenen Arbeit werden voraussichtlich ausschließlich Kamerasensoren eingesetzt welche in erster Linie Bilder liefern. Zusammengenommen mit der auf Per-Pixel Operationen ausgelegten Hardware in modernen GPUs scheint eine Parallelisierung auf diesem Wege der logische Schritt zu sein.

## 5.2 Tong et. al.

Die fehlende Echtzeitfähigkeit der von Tong et. al. vorgestellten Methode disqualifiziert sie zunächst für den in der eigenen Arbeit geplanten Zweck. Es lässt sich aber zumindest schließen, dass die Verwendung mehrerer, zeitlich aufeinander folgender Frames zwar zu einer besseren Scanqualität führt, aber die Berechnungszeit rasant ansteigen lässt. Die hier vorgeschlagene Pipeline kann somit noch bei einer Kompromissfindung zwischen Qualität und Geschwindigkeit hilfreich sein.

## 5.3 Alexiadis et. al.

Die von Alexiadis et. al. vorgestellte Lösung eignet sich prinzipiell als Scansystem für Personen. Hier wurde der Kompromiss zwischen Qualität und Geschwindigkeit stark in Richtung Geschwindigkeit verschoben, so dass die Qualität signifikant schlechter ist als in den anderen beiden Arbeiten.

# 6 Fazit

Die beiden vorgestellten Methoden zum Scannen von Personen unterschieden sich trotz teilweise gleicher verwendeter Komponenten und Algorithmen signifikant in der Vorgehensweise. Während bei Tong et. al. die Triangulation erst nach der Zusammenführung der Sensordaten erfolgt, triangulieren Alexiadis et. al. direkt nach Generierung der Punktwolken aus den Tiefenbildern. Dies resultiert in ganz unterschiedlichen Probleme welche dann natürlich separat gelöst werden. Auf diese Problemlösung fällt dann ein guter Anteil des Gesamtrechenaufwandes.

## 6.1 Ausblick

Für die eigene Arbeit ist es nötig weitere mögliche Vorgehensweisen zu beleuchten um daraus resultierende Probleme besser bewerten zu können. Hierzu sollen in der Projekt 2 Veranstaltung verschiedene Pipelines evaluiert werden, um einen besseren Kompromiss zwischen Geschwindigkeit und Qualität zu finden.

## Literatur

- [AZD13] ALEXIADIS, Dimitrios S. ; ZARPALAS, Dimitrios ; DARAS, Petros: Real-time, full 3-D reconstruction of moving foreground objects from multiple consumer depth cameras. In: *IEEE Transactions on Multimedia* 15 (2013), Nr. 2, S. 339–358
- [BM92] BESL, P.J. ; MCKAY, N.D.: A method for registration of 3-D shapes. In: *IEEE Transactions on pattern analysis and machine intelligence* 14 (1992), Nr. 2, S. 239–256
- [GP] GHIGLIA, Dennis C. ; PRITT, Mark D.: *Two-dimensional phase unwrapping: theory, algorithms, and software*
- [HSIW96] HILTON, Adrian ; STODDART, Andrew J. ; ILLINGWORTH, John ; WINDEATT, Terry: Reliable surface reconstruction from multiple range images. In: *Computer Vision—ECCV’96*. Springer, 1996, S. 117–126
- [IS12] IDE, Kai ; SIKORA, Thomas: Real-time active multiview 3d reconstruction. In: *Computer Vision in Remote Sensing (CVRS), 2012 International Conference on IEEE, 2012*, S. 203–208
- [ISS09] IDE, Kai ; SIERING, Steffen ; SIKORA, Thomas: Automating multi-camera self-calibration. In: *Applications of Computer Vision (WACV), 2009 Workshop on IEEE, 2009*, S. 1–6
- [KLB08] KURILLO, Gregorij ; LI, Zeyu ; BAJCSY, Ruzena: Wide-area external multi-camera calibration using vision graphs and virtual calibration object. In: *Distributed Smart Cameras, 2008. ICDS 2008. Second ACM/IEEE International Conference on IEEE, 2008*, S. 1–9
- [SMP05] SVOBODA, Tomáš ; MARTINEC, Daniel ; PAJDLA, Tomáš: A convenient multicamera self-calibration for virtual environments. In: *PRESENCE: teleoperators and virtual environments* 14 (2005), Nr. 4, S. 407–422
- [SPB04] SALVI, Joaquim ; PAGÈS, Jordi ; BATLLE, Joan: Pattern codification strategies in structured light systems. In: *Pattern Recognition* 37 (2004), Nr. 4, 827 - 849. <http://dx.doi.org/http://dx.doi.org/10.1016/j.patcog.2003.10.002>. – DOI <http://dx.doi.org/10.1016/j.patcog.2003.10.002>. – ISSN 0031–3203. – Agent Based Computer Vision
- [Svo] SVOBODA, TOMAS AND MARTINEC, DANIEL AND BOUGUET, JEAN-YVES AND WERNER, TOMAS AND CHUM, ONDREJ: *Multi-Camera Self-Calibration*. <http://cmp.felk.cvut.cz/~svoboda/SelfCal/>. – zuletzt besucht: 09.06.2014
- [TZL<sup>+</sup>12] TONG, J. ; ZHOU, J. ; LIU, L. ; PAN, Z. ; YAN, H.: Scanning 3d full human bodies using kinects. In: *IEEE Transactions on Visualization and Computer Graphics* 18 (2012), Nr. 4, S. 643–650
- [Zha00] ZHANG, Z.: A flexible new technique for camera calibration. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (2000), Nr. 11, S. 1330–1334

## Abbildungsverzeichnis

1	Verarbeitungsschritte beim 3D Scan. Quelle: eigenes Bild	1
2	Verarbeitungsschritte des 3D-Scansystems von Ide und Sikora. Quelle: eigenes Bild	2
3	Schematischer Aufbau der Scaneinheiten mit Projektoren (rot), Highspeedkamas (blau) und RGB-Kamas (grün). Quelle: eigenes Bild	2
4	Zeitliche Steuerung von zwei Scaneinheiten. Synchronisierungsframes sind schwarz. Schwarze Skalenstriche kennzeichnen Videoframes, graue Intervalle kennzeichnen Highspeedkamera-Frames. Quelle: [IS12]	3
5	Ein niederfrequentes Phase Shift Pattern. Farben sind um $\frac{2}{3}\pi$ gegeneinander phasenverschoben. Quelle: eigenes Bild	3
6	Rekonstruierte Punktwolken unter Benutzung von 2 (links) und 4 Scaneinheiten (rechts). Quelle: [IS12] (Ausschnitt)	4
7	Verarbeitungsschritte des 3D-Scansystems von Tong et. al. Quelle: eigenes Bild	5
8	Schematischer Aufbau der Scanvorrichtung. Quelle: <a href="http://www.math.zju.edu.cn/ligangliu/CAGD/Projects/Kinects-CapturingHumans/">http://www.math.zju.edu.cn/ligangliu/CAGD/Projects/Kinects-CapturingHumans/</a>	5
9	Rekonstruktionsergebnis als Komposition der Einzelmeshes (links) und als texturiertes 3D Modell in Vorder- und Rückansicht (rechts) Quelle: <a href="http://www.math.zju.edu.cn/ligangliu/CAGD/Projects/Kinects-CapturingHumans/">http://www.math.zju.edu.cn/ligangliu/CAGD/Projects/Kinects-CapturingHumans/</a>	6
10	Verarbeitungsschritte des 3D-Scansystems von Alexiadis et. al. Quelle: eigenes Bild	7
11	Schematischer Aufbau. Quelle: eigenes Bild	7
12	Rekonstruktionsergebnis als Komposition der Einzelmeshes (links) und als texturiertes 3D Modell. Quelle: [AZD13]	8