

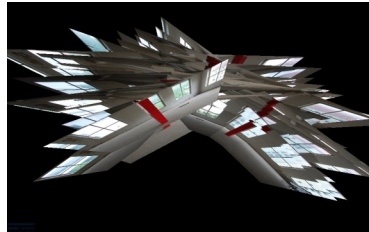
# 3D-Segmentierung in Bilderwelten

Frank Nagl, Paul Grimm

Fachhochschule Erfurt,  
Altonaer Str. 25, D-99085 Erfurt  
E-Mail: frank.nagl@fh-erfurt.de  
URL: <http://www.franknagl.de>



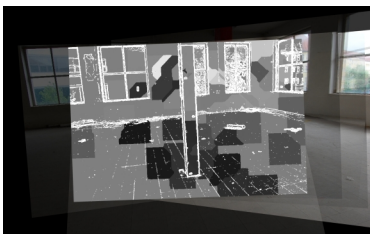
(a) Beliebige Fotos



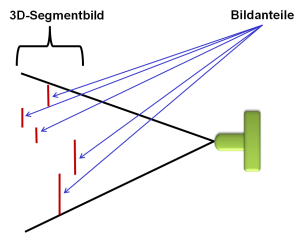
(b) Bilderwelt (Vogelperspektive)



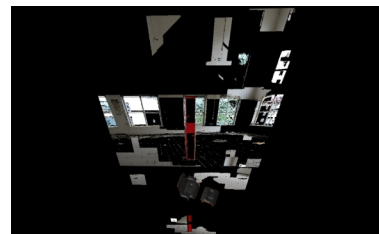
(c) Bilderwelt mit eingeblendeten 3D-Rekonstruktionspunkten



(d) 3D-Voronoi-Diagramm mit 3D-Informationen für jedes Segment



(e) Schematischer Aufbau des 3D-Segmentbildes



(f) 3D-Segmentbild (Vogelperspektive)

Abbildung 1: Erstellung eines 3D-Segmentbildes in einer Bilderwelt.

**Kurzfassung** Dieser Beitrag stellt eine neuartige 3D-Segmentierung von Fotos durch extrahierte 3D-Informationen eines Structure-From-Motion-Algorithmus vor. Mithilfe der 3D-Informationen wird eine 3D-Bilderwelt erzeugt für die Darstellung der Fotos sowie der 3D-Rekonstruktionspunkte im 3D-Raum (Abb. 1(b) und Abb. 1(c)). Als Grundlage dienen beliebige, unkalibrierte Fotos (Abb. 1(a)). Das Ziel dieser Arbeit ist die Zerlegung eines Fotos in logisch zusammenhängende Bereiche in Abhängigkeit der extrahierten 3D-Informationen der abgelichteten Szene (Abb. 1(d)). Die entstehenden 3D-Segmente werden an die entsprechende 3D-Position im Raum verschoben (Abb. 1(e)). Es entsteht ein *3D-Segmentbild* (Abb. 1(f)).

# 1 Einleitung

## 1.1 Motivation

Die Segmentierung ist ein bedeutendes Teilgebiet der digitalen Bildverarbeitung und des maschinellen Sehens. Hierbei werden inhaltlich zusammenhängende Regionen durch Zusammenfassung benachbarter Pixel extrahiert. Es existieren zwei Kategorien der Segmentierung. Konventionelle zweidimensionale Segmentierung verarbeitet Einzelbilder ohne Zusammenhänge zwischen mehreren Fotos. Im Gegensatz dazu existieren Ansätze für eine dreidimensionale Segmentierung, bei der Zusammenhänge zwischen mehreren Bildern und den darin abgebildeten Motiven für die Zusammenfassung inhaltlich zusammenhängender Regionen genutzt werden. Beispielsweise existieren Ansätze in [3] und [2] aus Bildern eines CT-Scans mittels 3D-Segmentierung das gescannte Objekt dreidimensional zu rekonstruieren. Jedoch setzen 3D-Segmentierungsverfahren in der Regel einige Anforderungen voraus. Die Aufnahmen der Bilder müssen mit exakt kalibrierten Kameras und speziellen Konfigurationsparametern (wie z.B. der Abstand zwischen zwei Schnittbildern bei einem CT-Scan) durchgeführt werden. Neben hohem Aufwand bringen spezielle Hardwareanforderungen hohe Kosten mit sich. Wünschenswert an dieser Stelle wäre eine 3D-Segmentierung aus beliebigen Fotos, aufgenommen mit beliebigen Consumer-Kameras ohne spezielle Konfigurationsansprüche.

Das hier vorgestellte 3D-Segmentierungsverfahren verarbeitet Farbbilder durch extrahierte 3D-Informationen der verwendeten Structure-From-Motion-Software Bundler ([11]). Als Grundlage dienen beliebig erstellte Fotos mit beliebigen Kameras ohne spezielle Anforderungen an Kalibrierung und Konfiguration der Kamera. Die erstellten Fotos müssen jedoch einige Überlappungen der abgelichteten Szene aufweisen, damit der Structure-From-Motion-Algorithmus korrekt arbeiten kann. Aus den Fotos werden 3D-Welten mit nahtlosen Übergängen von Foto zu Foto erstellt. Diese werden im weiteren Text als *Bilderwelten* bezeichnet. Zusätzlich werden markante Pixel, die in mehreren Fotos vorkommen (Eckpunkte, auffällige Kantenübergänge, etc.) 3D-rekonstruiert und zu einer 3D-Punktwolke zusammengeführt. Diese 3D-Rekonstruktionspunkte (im weiteren Text abgekürzt als: *RP*) dienen als Grundlage für die Erstellung des *3D-Segmentbildes*.

## 1.2 Anforderungen an die Farbtiefe der Bilder

Zwar können die Fotos ohne kalibrierte Kameras und spezielle Konfigurationsparameter aufgenommen werden, jedoch müssen die erstellten Bilder eine Farbtiefe von 24-(oder 32-)Bit aufweisen. Dies ist zum einen für den Structure-From-Motion-Algorithmus notwendig. Der Algorithmus verwendet die unterschiedlichen RGB-Kanäle für die Extraktion lokaler Bildmerkmale, die für das Erstellen der Bilderwelten notwendig sind. Ein Bild mit geringerer Farbtiefe kann nicht verarbeitet werden. Zum anderen wird die Farbtiefe von 24-(oder 32-)Bit von dem hier vorgestellten 3D-Segmentierungsverfahren vorausgesetzt. Die einzelnen Verarbeitungsschritte für die Erstellung des 3D-Segmentbildes (näher in Kapitel 3 erläutert) verwenden stets die RGB-Kanäle (24-Bit) für die Berechnungen/Verarbeitungen. Dies erhöht die Genauigkeit (aufgrund der größeren Bitzahl pro Pixel) im Vergleich zu einem Bild geringerer Farbtiefe. Ferner soll auf die Verwendung von Shader-Programmen bei der Implementierung des vorgestellten Verfahrens hingewiesen werden (Absatz 4.1). Ein Shader-Programm verarbeitet jedes Pixel eines Bildes als ein ARGB-Pixel. Somit ist eine geringere Farbtiefe ebenfalls nicht sinnvoll.

Aufgrund der besseren Lesbarkeit wird im weiteren Text auf die Angabe der vorausgesetzten Farbtiefe von 24-(oder 32-)Bit verzichtet und nur noch von Fotos oder Bildern gesprochen.

### 1.3 Gliederung

Die nächsten Abschnitte gliedern sich folgendermaßen: Zunächst werden existierende Arbeiten und Projekte aufgezeigt, die ein oder mehrere ähnliche Ziele dieser Arbeit verfolgen. Anschließend wird das Konzept zur Erstellung des 3D-Segmentbildes aus Bilderwelten erläutert. Die Umsetzung des erläuterten Verfahrens wird in Kapitel 4 beschrieben. Abschließend wird eine Zusammenfassung der Arbeit und ein Ausblick für weitere Entwicklungen des Verfahrens gegeben.

## 2 State-Of-the-Art

Dieser Beitrag beinhaltet Themen aus unterschiedlichen Forschungsgebieten der Grafischen Datenverarbeitung. Aus diesem Grund werden in diesem Kapitel bestehende Arbeiten aus den zu dieser Arbeit passenden Bereichen **2D-Segmentierung**, **3D-Segmentierung** und **Bilderwelten** vorgestellt.

Die **2D-Segmentierung** extrahiert inhaltlich zusammenhängende Regionen durch Zusammenfassung benachbarter Pixel. Die Arbeitsgruppe von Hengel [4] erarbeitete ein Verfahren, wie nahe liegende Kanten in einem Bild zu geschlossenen Segmenten zusammenwachsen und als Polygone genutzt werden können. Um ein Polygon-Modell zu erstellen, bedarf es der Interaktion eines Nutzers. Dieser hat die Aufgabe den Grundriss des Objektes mit einfachen 2D-Hilfsmitteln nachzuzeichnen. Daher eignet sich dieses Verfahren nicht in unserem System. Grundlegende Segmentierungs-Verfahren finden sich in allgemeiner Literatur zum Thema Bildverarbeitung. Als Beispiel wird an dieser Stelle auf die Ausführungen von Tönnies [14] verwiesen, der in seinem Buch einen detaillierten Überblick über grundlegende Segmentierungs-Verfahren gibt.

Verschiedene **3D-Segmentierungsverfahren** aus speziellen Fotos (wie MRT- oder CT-Bilder) wurden bereits in [6], [13], [3] und [2] vorgestellt. Da es sich bei diesen Arbeiten jedoch stets um kalibrierte Bildserien handelt, eignen sich diese Verfahren nicht für beliebige Kameras/Fotos und somit nicht für unser Verfahren.

Die Darstellung von unsortierten Fotos als **Bilderwelten** ermöglicht das Projekt Photo Tourism [12]. Damit können lose Fotosammlungen mit Überlappungen der Bildanteile zu Bilderwelten zusammengefügt werden. Aufbauend auf Photo Tourism wurde eine kommerzielle Variante namens Photosynth [7] von Microsoft veröffentlicht. Beide Projekte nutzen zum Extrahieren der Kamerapositionen und -ausrichtungen der geschossenen Fotos die Structure-From-Motion Software Bundler [11]. Auch das in dieser Arbeit vorgestellte Verfahren verwendet zum Extrahieren der Kameraparameter und der 3D-Rekonstruktionspunkte diese Software, wie noch in Kapitel näher 3 beschrieben wird. Abgrenzend zu Photo Tourism und Photosynth, können in unserem Verfahren Fotos aus den Bilderwelten 3D-segmentiert und dargestellt werden.

## 3 Konzept

### 3.1 Kurzfassung des Konzeptes

Ein Bilderwelten-Viewer und dazugehörige RPs sind gegeben (Abb. 1(b) und 1(c)). Aus der Bilderwelt wird das zu segmentierende Bild ausgewählt. Die RPs werden aus dem 3D-Raum in das Bild als 2D-Pixel zurück projiziert. Ihre Abstände zur Kamera im 3D-Raum dienen als Füllwerte für die Segmente (im weiteren Text als *Abstandswerte* bezeichnet). Ausgehend von jedem RP als Startpunkt wachsen Segmente solange bis das komplette Bild aus Segmenten besteht. Dieses Verfahren gleicht dem Erstellen eines Voronoi-Diagramms [1] im 3D-Raum und das resultierende Bild wird daher als 3D-Voronoi-Diagramm bezeichnet. Das 3D-Voronoi-Diagramm wird verwendet für das Verschieben der einzelnen Segmente des Originalbildes entsprechend der Abstandswerte in den einzelnen Voronoi-Zellen (Segmenten). Das Ergebnis ist ein 3D-Segmentbild.

### 3.2 Projektion der 3D-Rekonstruktionspunkte

Da die Structure-From-Motion-Software die Bilderwelten mit einer 3D-Punktwolke liefert, jedoch die RPs für die eigentliche Segmentierung des Bildes in 2D-Pixelkoordinaten und nicht in 3D-Weltkoordinaten benötigt werden, müssen die RPs in das Bild zurück projiziert werden.

Für jedes Bild ist die Brennweite ( $f$ ), die Rotationsmatrix ( $R$ ) und der Translationsvektor ( $t$ ) gegeben. Die Brennweite wird zuvor in Pixelkoordinaten umgerechnet (Formel 1).

$$f = w * \frac{fmm}{ccd} \quad (1)$$

Die Bildbreite ( $w$ ) in Pixeln, die Brennweite in Millimetern ( $fmm$ ), sowie die Breite des CCD-Sensors ( $ccd$ ) in Millimetern der verwendeten Kamera werden aus den Meta-Daten (Exif-Tags) des Bildes ausgelesen. Formel 2 bis 4 dienen zur Projektion eines RP auf ein Bild.

$$P = R * X + t \quad (2)$$

$$p = \frac{-P}{P.z} \quad (3)$$

$$p' = f * p \quad (4)$$

Formel 2 konvertiert die Weltkoordinaten des RP in Kamerakoordinaten. Formel 3 repräsentiert die perspektivische Division und Formel 4 überführt in Pixelkoordinaten.

### 3.3 Vorverarbeitungsschritte

Als Vorverarbeitungsschritt wird das Bild verschiedenen Glättungsfiltern unterzogen, um Bildrauschen entgegen zu wirken. Anschließend wird das Bild einem Verfahren zur Erkennung von parametrisierbaren geometrischen Figuren unterzogen. Beispielhaft sollen an dieser Stelle Kantendetektoren und die Hough-Transformation [5] genannt werden. Die detektierten Kanten bzw. Linien werden in das Bild eingezeichnet und dienen als Unterstützung bei der korrekten Einordnung ähnlicher Regionen im Segmentierungsprozess.

### 3.4 Erstellen eines 3D-Voronoi-Diagramms

Für die Erstellung des 3D-Segmentbildes werden die korrekten 3D-Positionen für jedes einzelne Segment benötigt. Diese 3D-Positionen kodiert als Farbwerte der Pixel in einem Bild wird *Depth-Map* (Tiefenkarte) genannt. Das Erstellen einer Depth-Map kann in dieser Arbeit als ein 3D-Voronoi-Diagramm [1] gesehen werden. Der Kern einer Voronoi-Zelle bildet dabei ein RP. Seine 3D-Position bzw. der Abstand zwischen Kameraposition und seiner 3D-Position wird als Farbfüllwert genutzt. Wie bereits in Absatz 3.1 erwähnt, wird dieser Wert als Abstandswert bezeichnet. Ausgehend von dem Kern werden benachbarte Pixel auf Zugehörigkeit zu einer Zelle untersucht (erste Abbruchbedingung). Ist dies nicht der Fall, werden sie Teil der eigenen Zelle und bekommen den entsprechenden Abstandswert als Farbfüllung. Dessen Nachbarn werden ebenfalls untersucht. Dies geschieht solange, bis das gesamte Bild vollständig mit Zellen ausgefüllt ist. Zusätzlich dienen die Kanten im Bild als zweite Abbruchbedingung der Voronoi-Zellbildung. Das so entstehende 3D-Voronoi-Diagramm (Abb. 2) wird für das Verschieben der einzelnen Segmente zu einem 3D-Segmentbild verwendet.

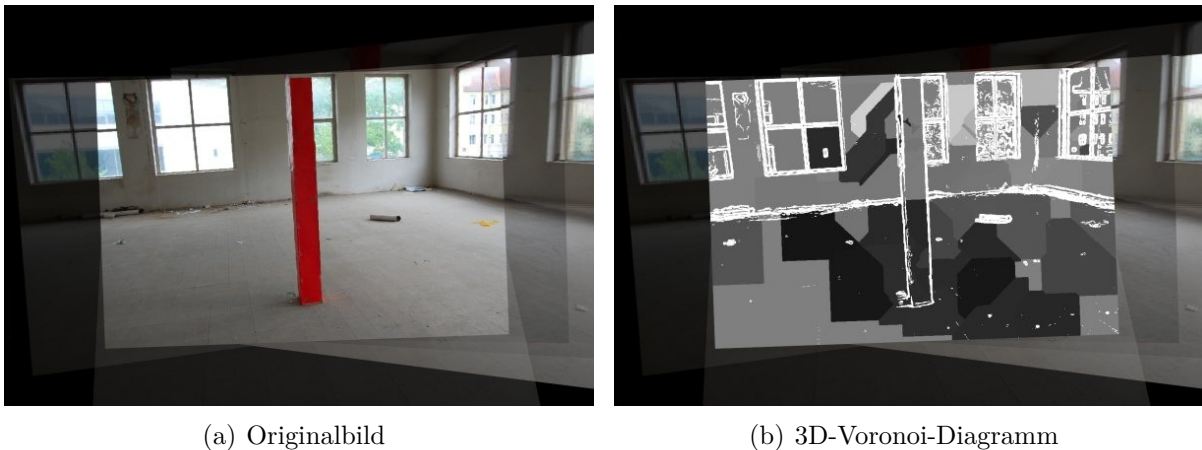
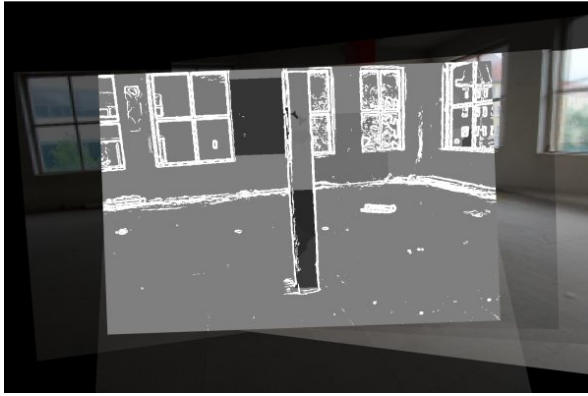
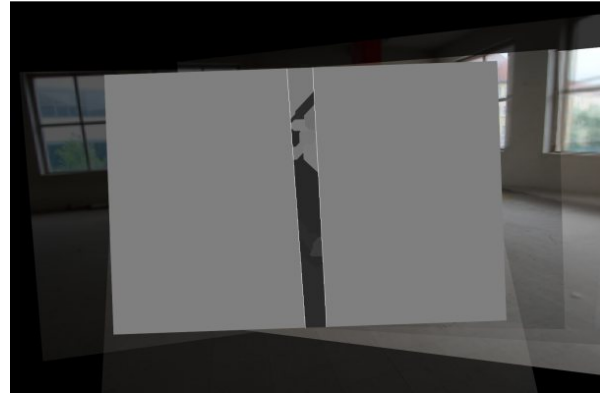


Abbildung 2: Gegenüberstellung Originalbild und das dazugehörige 3D-Voronoi-Diagramm.

Je nach Zielstellung ist es notwendig die Zellbildung des 3D-Voronoi-Diagramms zu modifizieren. Beispielsweise soll das Ziel sein, ausschließlich Vordergrundobjekte auf Fotos zu detektieren und zu segmentieren. Hierfür stellt ein Ansatz dar, eine dritte Abbruchbedingung für die Zellbildung des 3D-Voronoi-Diagramms einzuführen. Es wachsen nur Zellen, die sich zwischen parallelen Kanten befinden. Weiterführend kann die Größe der Körper bzw. Länge der Kanten der Körper mit einbezogen werden. Auf diese Weise lassen sich einfache geometrische Körper, die sich im Vordergrund eines Bildes befinden, segmentieren. Abb. 3 zeigt die Ergebnisse von erstellten 3D-Voronoi-Diagrammen mithilfe einer dritten Abbruchbedingung unter Verwendung zweier unterschiedlicher Verfahren zur Erkennung von geometrischen Figuren. Abb. 3(a) verwendet einen gradientenbasierten Kantendetektor ohne Berücksichtigung der Länge der Kanten. Dieses Verfahren ist effizient, kann jedoch zu fälschlich detektierten Segmenten führen, wie im Bild zu sehen ist. Abb. 3(b) verwendet eine Hough-Transformation und detektiert nur Kanten einer bestimmten Länge. Somit wird nur das Vordergrundobjekt eingezeichnet. Dieses Verfahren ist geeignet für Objekte, die sich in ihrer Größe von anderen Objekten im Bild unterscheiden.



(a) Verwendung des Sobel-Kantendetektors

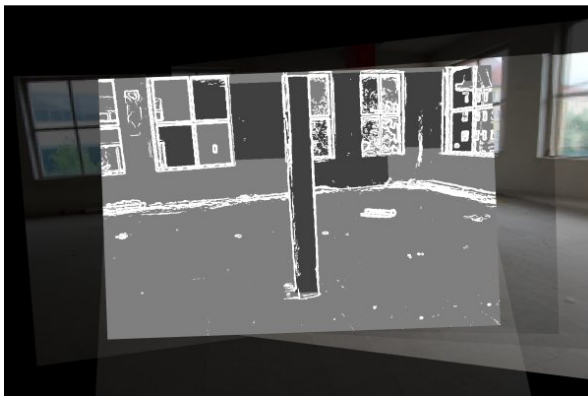


(b) Verwendung der Houghline-Transformation

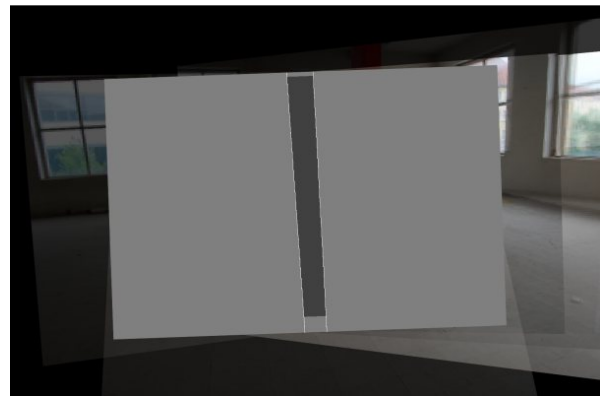
Abbildung 3: 3D-Voronoi-Diagramme mit unterschiedlichen Kantendetektoren.

Da der Structure-from-Motion-Algorithmus unter Umständen auch RPs detektiert, die sich räumlich hinter einem Vordergrundobjekt befinden (z.B. das Fenster hinter dem roten Balken im Beispielbild, dass vom Balken teilweise verdeckt wird), können aufgrund der Rückprojektion der RPs auf die Bildebene Voronoi-Zellen entstehen, die einen falschen Abstandswert kodieren. Dies ist in Abb. 3 im oberen Teil des Balkens zu sehen.

Um dieses Problem zu lösen, können unterschiedliche Segmente, die zu einem Objekt gehören, verschmelzen. Dazu werden die Abstandswerte der einzelnen Segmente auf Ausreißer untersucht und anschließend gemittelt. Abb. 4 zeigt die 3D-Voronoi-Diagramme aus Abb. 3, jedoch mit verschmolzenen Segmenten.



(a) Verwendung des Sobel-Kantendetektors



(b) Verwendung der Houghline-Transformation

Abbildung 4: 3D-Voronoi-Diagramme mit verschmolzenen Segmenten.

### 3.5 Erstellen des 3D-Segmentbildes

Mithilfe des 3D-Voronoi-Diagramms, das die korrekten 3D-Positionen bzw. Abstandswerte der einzelnen Segmente kodiert, werden die Segmente räumlich im 3D-Raum angeordnet. Dafür wird jedes Segment an die entsprechende 3D-Position verschoben. Damit alle Segmente zusammen dennoch den Eindruck eines gesamten Bildes vermitteln, müssen die verschobenen Segmente perspektivisch korrekt projiziert werden. Zur Kamera verschobene Segmente müssen verkleinert und von der Kamera weg verschobene Segmente müssen

vergrößert werden. Da die Fotos innerhalb der Bilderwelt immer orthogonal zur Kamera stehen, kann diese Skalierung durch eine lineare Verhältnisgleichung beschrieben werden (Formel 5).

$$\frac{z_{original}}{h_{original}} = \frac{z_{verschoben}}{h_{verschoben}} \quad (5)$$

Die Gesamtheit aller verschobenen, skalierten Segmente wird 3D-Segmentbild genannt. Abbildung 5 zeigt den schematischen Aufbau des 3D-Segmentbildes mit seinen verschobenen, skalierten Segmenten. Die Höhe aller Segmente ergibt die Höhe des Originalbildes.

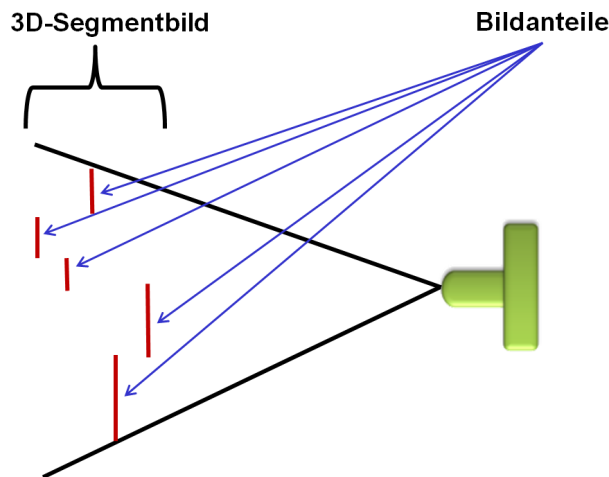


Abbildung 5: Schematischer Aufbau des 3D-Segmentbildes.

## 4 Implementierungsdetails und Ergebnisse

Zur Implementierung des Konzepts wurde ein 3D-Viewer für die Darstellung von Bilderwelten und virtuellen 3D-Objekten auf Basis von MS XNA 3.1 [8] als 3D-Grafik-Framework entwickelt. Weiterführend wurde das grafikartenbasierten Bildverarbeitungs-Framework SBIP [9] für Bildsegmentierungs-Algorithmen, sowie selbst entwickelte Klassen auf DotNet-Basis zur Erstellung des 3D-Voronoi-Diagramms und des 3D-Segmentbildes verwendet.

### 4.1 Grafikkartenbasierte Verarbeitung

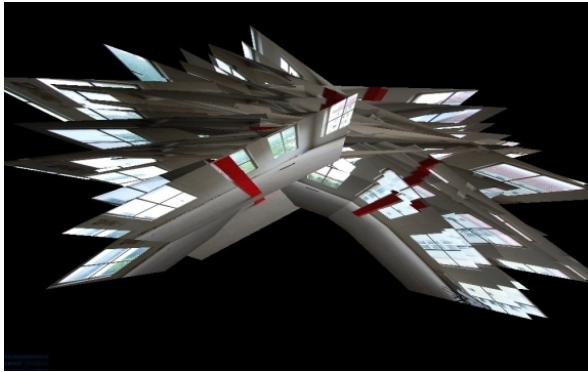
Die Implementierung der einzelnen Verarbeitungsschritte für die Erstellung des 3D-Segmentbildes arbeiten grafikartenbasiert. Über HLSL-Shader-Programme [10] werden die Berechnungen direkt auf der Grafikkarte durchgeführt. Aufgrund der Parallelisierung der Aufgaben auf der Grafikkarte können so die Berechnungen enorm performant abgearbeitet werden.

### 4.2 Umsetzung einer Bilderwelt

Der im Rahmen dieser Arbeit entwickelte 3D-Viewer liest die von der Structure-From-Motion-Software Bundler zuvor berechneten Kamerapositionen und -ausrichtungen der



Fotos, sowie die 3D-Positionen der RPs (3D-Punktwolke) ein und stellt die Bilderwelt dar (Abb. 6). Hierbei werden die Fotos mit einem berechneten Abstand orthogonal vor die Kameras positioniert. Der Abstand berechnet sich als Durchschnittswert der 3D-Positionen mittels RANSAC-Algorithmus [Bol81] aller zum Foto passenden RPs.



(a) Bilderwelt aus der Vogelperspektive



(b) Bilderwelt mit 3D-Rekonstruktionspunkten

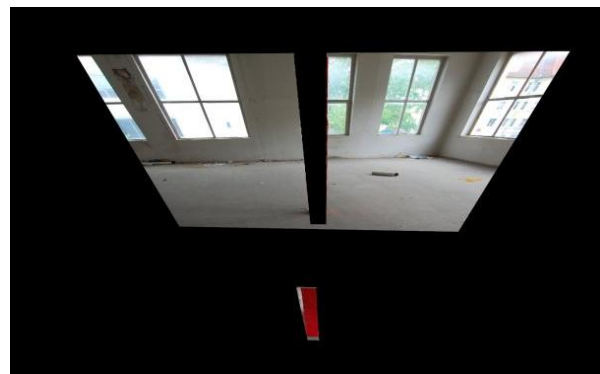
Abbildung 6: Aufbau und Darstellung einer Bilderwelt.

### 4.3 Ergebnisse

Bei der Realisierung für die Erstellung des 3D-Voronoi-Diagramms wurden zwei Verfahren zur Erkennung von geometrischen Figuren im Vorverarbeitungsschritt implementiert: Der Sobel-Kantendetektor und die Hough-Transformation für die Erkennung von Polygonlinien im Bild. Darauf aufbauend wurden zwei Verfahren (entsprechend dem Konzept) für die Erstellung eines 3D-Voronoi-Diagramms umgesetzt: Eine generelle Variante mit zwei Abbruchbedingungen und eine modifizierte Variante unter Verwendung einer dritten Abbruchbedingung für die ausschließliche Segmentierung von Vordergrundobjekten. Mithilfe des 3D-Voronoi-Diagramms wird das 3D-Segmentbild durch Verschieben und Skalieren der einzelnen Segmente des Originalbildes erstellt.



(a) Verwendung des 3D-Voronoi-Diagramms mit 2 Abbruchbedingungen



(b) Verwendung des 3D-Voronoi-Diagramms mit 3 Abbruchbedingungen, für die Segmentierung des Vordergrundobjektes

Abbildung 7: Darstellung eines 3D-Segmentbildes in einer Bilderwelt aus der Vogelperspektive.



Abb. 7 stellt beide Varianten eines 3D-Segmentbildes gegenüber.

#### 4.4 Anwendungsbeispiel Verdeckung

Ein Anwendungsszenario für die Verwendung eines 3D-Segmentbildes stellt die Behandlung potentieller Verdeckungsprobleme dar. Das Augmentieren einer Bilderwelt durch das Einblenden virtueller 3D-Objekte bringt eine Reihe von Anforderungen mit sich, um einen authentischen Eindruck der Mixed-Reality-Szene zu vermitteln. Beispielsweise können virtuelle 3D-Objekte in der Bilderwelt so positioniert werden, dass potentielle Verdeckungsprobleme durch reale Bildanteile entstehen (Abb. 8(a)).

Das resultierende 3D-Segmentbild (Abb. 7) kann zum Verdecken eingblendeter virtueller 3D-Objekte durch reale Bildanteile (einzelne Segmente) genutzt werden. Hierfür verdecken einzelne Segmente des 3D-Segmentbildes das virtuelle 3D-Objekt. Mit anderen Worten das virtuelle 3D-Objekt befindet sich innerhalb eines 3D-Segmentbildes, wobei einige Segmente vor und einige Segmente hinter dem 3D-Objekt positioniert sind. Dies steigert die Glaubwürdigkeit der Einblendung des virtuellen 3D-Objekts und somit den authentischen Eindruck der Mixed-Reality-Szene. Abb. 8 vergleicht das Einblenden eines virtuellen Stuhls mit und ohne Verwendung eines 3D-Segmentbildes.



(a) Originalbild ohne Verdeckung

(b) 3D-Segmentbild mit Verdeckung

Abbildung 8: Einblendung eines virtuellen 3D-Objektes in einer Bilderwelt.

## 5 Zusammenfassung und Ausblick

### 5.1 Zusammenfassung

In diesem Beitrag wurde ein neues Verfahren für die 3D-Segmentierung von Farbbildern beschrieben.

Im Gegensatz zu anderen 3D-Segmentierungsverfahren werden keinerlei Anforderungen an die Hardware gestellt. Als Grundlage dienen beliebig erstellte Fotos von beliebigen Kameras ohne spezielle Anforderungen an Kalibrierung und Konfiguration der Kamera. Das hier vorgestellte 3D-Segmentierungsverfahren verarbeitet Farbbilder durch extrahierte 3D-Informationen eines Structure-From-Motion-Algorithmus. Die so berechneten Kamerapositionen und -ausrichtungen der Bilder in einer 3D-Welt, sowie die 3D-Positionen der 3D-Rekonstruktionspunkten (RPs) werden in einer Bilderwelt dargestellt und für das

3D-Segmentierungsverfahren weiterverarbeitet. Aus der Bilderwelt wird das zu segmentierende Bild ausgewählt. Die RPs werden aus dem 3D-Raum in das Bild als 2D-Pixel zurück projiziert. Ihre Abstände zur Kamera im 3D-Raum dienen als Füllwerte für die Segmente (Abstandswerte). Ausgehend von jedem RP als Startpunkt wachsen Segmente solange bis das komplette Bild aus Segmenten besteht. Dieses Verfahren gleicht dem Erstellen eines Voronoi-Diagramms im 3D-Raum und das resultierende Bild wird daher als 3D-Voronoi-Diagramm bezeichnet. Das 3D-Voronoi-Diagramm wird verwendet für das Verschieben der einzelnen Segmente des Originalbildes entsprechend der Abstandswerte in den einzelnen Voronoi-Zellen (Segmenten). Das Ergebnis ist ein 3D-Segmentbild.

Für das Konzept des 3D-Segmentierungsverfahrens wurden Fotos mit einer Farbtiefe von 24-(oder 32-)Bit vorausgesetzt. Die Implementierung des Verfahrens arbeitet grafikkartenbasiert (Shader-Technologie), die ebenfalls Farbbilder bedingt.

Das hier vorgestellte Verfahren zur Erstellung eines 3D-Segmentbildes, dient zur Verbesserung der authentischen Produktpräsentation virtueller Möbel in Bilderwelten.

## **5.2 Ausblick**

Für eine zukünftige Erweiterung der hier vorgestellten 3D-Segmentierung sollen die Verfahren zur Erkennung von parametrisierbaren geometrischen Figuren als Vorverarbeitungsschritt optimiert werden. Eine vollständige Trennung der Vordergrundobjekte vom Hintergrund kann die 3D-Segmentierung erheblich verbessern.

Eine Alternative zu dem vorgestellten 3D-Voronoi-Diagramm könnte die Verwendung eines Stereo-View-Algorithmus sein. Da alle 3D-Kamerapositionen und -ausrichtungen der Fotos bekannt sind, könnte eine Depth-Map über Multi-View-Algorithmen erstellt werden.

## **6 Danksagung**

Diese Arbeit entstand im Rahmen des Projektes PoP-EYE, dass durch das Bundesministerium für Bildung und Forschung (Projekt-Nummer: 17N0909) gefördert wird. Darüber hinaus möchten wir uns bei der IP3D-Arbeitsgruppe der Fachhochschule Erfurt, bestehend aus Tobias Bindel, Bastian Birnbach, Konrad Kölzer und Stefan Rothe für ihre Unterstützung bei der Implementierung eines Bilderwelten-Viewers bedanken. Ferner gilt Ekkehard Beier (EasternGraphics GmbH) speziellen Dank für seine Hilfe auf wissenschaftlicher und technischer Basis.

## Literatur

- [1] AURENHAMMER, FRANZ: *Voronoi diagrams - A survey of a fundamental geometric data structure*. ACM Comput. Surv. Vol. 23. No. 3, Seiten 345–405, 1991.
- [2] DANJOU, MAKOTO, AYAKO TATEISHI, SAYAKA ITABASHI, YUKO MASUNAGA, KAORI KOISO, HIROKO UCHIYAMA, MASAHIRO ASANO, KAZUAKI HIRATA, TOMOHITO NAGAOKA und SAKUJI YOSHIMURA: *Facial reconstruction of Egyptian mummy 'Senu'*. SIGGRAPH '06: ACM SIGGRAPH 2006 Research posters, Seite 66, 2006.
- [3] FORTE, MAURIZIO: *3D facial reconstruction and visualization of ancient Egyptian mummies using spiral CT data*. SIGGRAPH '99: ACM SIGGRAPH 99 Conference abstracts and applications, Seite 223, 1999.
- [4] HENGEL, ANTON VAN DEN, ANTHONY DICK, THORSTEN THORMÄHLEN, BEN WARD und PHILIP H. S. TORR: *VideoTrace: rapid interactive scene modelling from video*. SIGGRAPH '07: ACM SIGGRAPH 2007 papers, Seite 86, 2007.
- [5] HOUGH, PAUL: *Method and means for recognizing complex patterns*. Patent, 1962.
- [6] KORFIATIS, PANAYIOTIS, SPYROS SKIADOPOULOS, PHILIPPOS SAKELLAROPOULOS, CHRISTINA KALOGEROPOULOU und LENA COSTARIDOU: *Automated 3D segmentation of lung fields in thin slice CT exploiting wavelet preprocessing*. CAIP'07: Proceedings of the 12th international conference on Computer analysis of images and patterns, Seiten 237–244, 2007.
- [7] MICROSOFT: *Photosynth - Your photos, automatically in 3D*. <http://photosynth.net>, 2010.
- [8] MICROSOFT: *XNA Game Studio 3.1*. <http://creators.xna.com/de-de>, 2010.
- [9] NAGL, FRANK, MARTIN FRIEDL, ANDREAS SCHÄFER und ANDRE TSCHENTSCHER: *Shader-Based-Image-Processor - Ein Framework für echtzeitfähige, grafikartenbasierte Bilderverarbeitung*. GI Seminars 9 Informatiktage 2010: Fachwissenschaftlicher Informatik-Kongress. Köllen Druck+Verlag GmbH, Bonn. 1. Edition, Seiten 223–226, 2010.
- [10] SHERIF, PATRIC: *HLSL-Shader Entwicklung unter Direct3D 9*. Vdm Verlag Dr. Müller, 2008. ISBN: 978-3639009514.
- [11] SNAVELY, NOAH: *Bundler 0.3: Structure from Motion for Unordered Image Collections*. <http://phototour.cs.washington.edu/bundler/>, 2009.
- [12] SNAVELY, NOAH, STEVEN M. SEITZ und RICHARD SZELISKI: *Photo tourism: exploring photo collections in 3D*. SIGGRAPH '06: Proceedings of the ACM SIGGRAPH 2006 Papers, Seiten 835–846, 2006.
- [13] TAWARA, TAKEHIRO und KENJI ONO: *Direct 3D manipulation for volume segmentation using mixed reality*. SIGGRAPH ASIA '09: ACM SIGGRAPH ASIA 2009 Posters, 2009.
- [14] TOENNIES, KLAUS D.: *Grundlagen der Bildverarbeitung*. Pearson Studium, 2005. ISBN: 978-3827371553.