

Large-Scale 3D Reconstruction: A Triangulation-Based Approach

von
Tobias Gleixner

Seminar WS 2002 / 03
3D Rekonstruktion

Lehrstuhl für Praktische Informatik IV
Prof. Dr. W. Effelsberg
Fakultät für Mathematik und Informatik
Universität Mannheim

Januar 2003

Betreuer: Dipl.-Inf. Dirk Farin

Inhaltsverzeichnis

| | |
|---|-----------|
| 1. Einleitung | 3 |
| 1.1. Themengebiet | 3 |
| 1.2. Motivation | 3 |
| 1.3. Lösungsansatz | 4 |
| 1.4. Vorgehen (Überblick) | 4 |
| 2. Grundlagen | 6 |
| 2.1. Annahmen | 6 |
| 2.2. Datenstrukturen | 6 |
| 2.3. Tracking-Algorithmen | 8 |
| 2.4. Reihenfolge der Hypothesenbildung | 9 |
| 3. Triangulation | 10 |
| 3.1. Triangulation in 2D | 11 |
| 3.2. Punkt-Triangulation..... | 11 |
| 3.3. Linienabschnitts-Triangulation | 13 |
| 4. Auswertung von Hypothesen | 15 |
| 4.1. Datenkonsistenz-Wahrscheinlichkeit | 16 |
| 4.2. Non-Accidentalness-Wahrscheinlichkeit..... | 16 |
| 4.3. A-posteriori-Wahrscheinlichkeit | 17 |
| 4.4. Bestätigung einer Hypothese..... | 18 |
| 4.5. Ablehnung einer Hypothese..... | 18 |
| 4.6. Update einer Hypothese..... | 18 |
| 5. Oberflächen | 19 |
| 5.1. Oberflächenerkennung..... | 20 |
| 5.2. Oberflächenberechnung | 20 |
| 5.3. Sichtbarkeitstest..... | 21 |
| 6. Ergebnisse | 22 |
| 7. Zusammenfassung | 23 |

1. Einleitung

1.1. Themengebiet

Thema dieser Ausarbeitung ist die Doktorarbeit von George Tao-Shun Chou mit dem Titel „Large-Scale 3D Reconstruction: A Triangulation-Based Approach“.

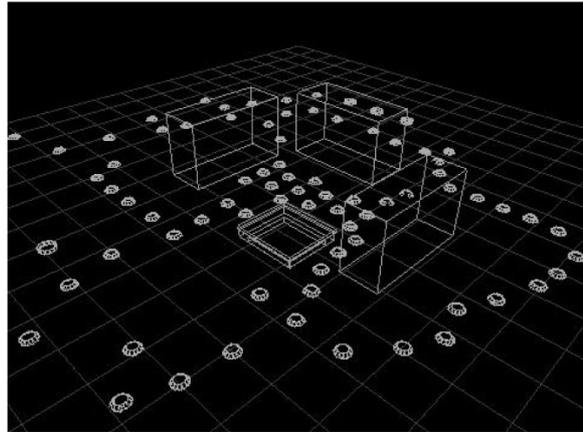


Abb. 1-1: Fertiges 3D-Modell

Large-Scale 3D Rekonstruktion beschäftigt sich mit der automatisierten Gewinnung einer 3D-Beschreibung von urbanen Umgebungen. Dabei soll aus einer großen Anzahl von digital vorliegenden Bildern ein 3D-Modell der darin vorkommenden Gebäude und Objekte vollautomatisch und effizient generiert werden. Das 3D-Modell sollte möglichst exakt die Umrisse der Gebäude enthalten, während die Details und Strukturen der Gebäudeoberflächen unberücksichtigt bleiben.

1.2. Motivation

Mit Hilfe der bekannten Kantendetektions-Algorithmen ist es möglich, in jedem Bild die so genannten Features (d.h. Eckpunkte und Kanten der Gebäude im 2D-Bild) zu finden. Durch Triangulation ließe sich daher relativ einfach aus bekannten Kameraparametern und den Features aus zwei oder mehr Bildern die 3D-Position dessen Elements bestimmen, das auf die Features abgebildet wurde. Leider müsste hierzu jedoch bekannt sein, welche Features das gesuchte Element beschreiben.

In den meisten Ansätzen für 3D-Rekonstruktion müssen daher zunächst andere Algorithmen verwendet werden, die übereinstimmende Features bestimmen (**Featurekorrespondenz**), bevor Triangulation eingesetzt werden kann. Diese Algorithmen schränken jedoch die Wahl

der Bilder stark ein, da sie voraussetzen, dass die Bilder geordnet eingegeben werden und die Kamerapositionen dicht beieinander liegen.

Gerade in Large-Scale Systemen ist diese Einschränkung jedoch unerwünscht, da zum einen keine Zwischenbilder nachträglich hinzugefügt werden können, und zum anderen die sowieso schon große Zahl an Bildern noch einmal vervielfacht werden müsste, nur um die Bestimmung der Featurekorrespondenz zu gewährleisten.

1.3. Lösungsansatz

Anstatt Triangulation nur zu verwenden, nachdem übereinstimmende Features bestimmt wurden, setzt der vorliegende Ansatz Triangulation nun gerade dazu ein, um solche Features zu finden. Damit übernimmt die Triangulation also sowohl die Aufgabe, Featurekorrespondenz herzustellen, als auch daraus eine Schätzung für das 3D-Modell abzuleiten.

1.4. Vorgehen (Überblick)

Ausgehend von einem leeren Initialzustand wird dem System nacheinander Bild für Bild in beliebiger Reihenfolge hinzugefügt. Mit jedem Bild werden vollständige Informationen über Kameraparameter (Position, Richtung, etc.) dem System zugeführt.

Jedes Bild löst dabei folgende Schritte aus:

1. **Featureextraktion:** Aus jedem Bild werden zwei Arten von Features extrahiert:
 - 2D Punkte
 - 2D Linienabschnitte
2. **Extrusionsberechnung:** Jedes Feature wird unter Zuhilfenahme der Kameraparameter eines Bildes ausgehend vom Kamerazentrum in den 3D-Raum projiziert. Diese Projektion wird dann Extrusion genannt.



Abb. 1-2: Beispielbild

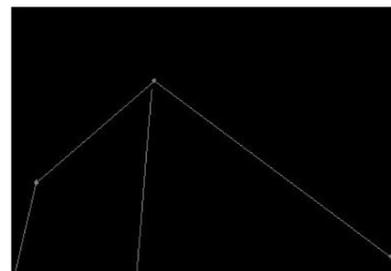


Abb. 1-3: Featureextraktion

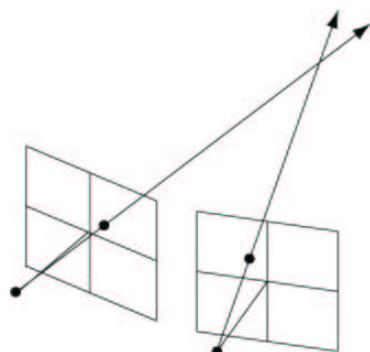


Abb. 1-4: Extrusionsberechnung

3. **Triangulation:** Jede Extrusion wird mit allen gleichartigen Elementen, Hypothesen und Features (genauer gesagt Feature-Extrusionen) aus früheren Schritten zu neuen Hypothesen kombiniert. Eine Schätzung für die 3D-Struktur der neuen Hypothese wird gewonnen, indem aus allen zur Hypothese kombinierten Extrusionen stark vereinfacht der „Schnitt“ berechnet wird (so genannte Triangulation).

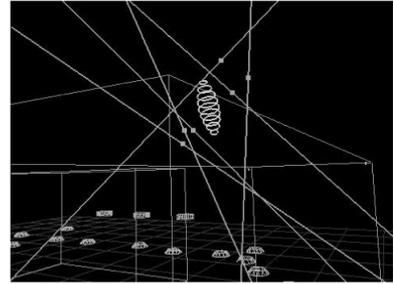


Abb. 1-5: Punkt-Triangulation

4. **Berechnung der A-posteriori-Wahrscheinlichkeit:** Da auch Hypothesen aus Extrusionen erzeugt werden können, die zu verschiedenen 3D-Elementen gehören, müssen diese herausgefiltert werden. Daher wird eine Wahrscheinlichkeit für die Gültigkeit der Hypothese berechnet, in die der Triangulationsfehler und die Anzahl der verwendeten Features einfließen.
5. **Beurteilung der Hypothese:** Liegt die A-posteriori-Wahrscheinlichkeit über einem bestimmten Schwellwert p_c , so wird aus der Hypothese ein 3D-Element, liegt sie unter einem bestimmten Schwellwert $p_r < p_c$, so wird die Hypothese verworfen; liegt sie zwischen p_r und p_c , so verbleibt die Hypothese im System und wird in den nächsten Iterationen weiterverwendet.

- (6.) **Oberflächenberechnung:** Wurde ein Linienabschnitt als 3D-Element bestätigt, wird untersucht, ob mehrere anliegende Linienabschnitte eine Oberfläche bilden. Falls ja, so werden die Parameter der Oberfläche aus den Daten der umschließenden Linienabschnitte geschätzt.

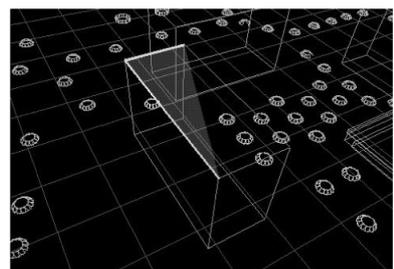


Abb. 1-6: Oberflächenberechnung

- (7.) **Sichtbarkeitstest:** Jede Hypothese, deren geschätzte Position hinter der neuen Oberfläche liegt, wird verworfen.

Im nächsten Kapitel werden zunächst einmal die Grundlagen und Begriffe des Systems näher erläutert. Während für Schritt 1 die bekannten Kantendetektions-Algorithmen angewandt

werden und Schritt 2 standardmäßig aus allen Eingabedaten abgeleitet werden kann, werden in den dann folgenden Kapiteln die Schritte 3 bis 7 detaillierter vorgestellt.

2. Grundlagen

2.1. Annahmen

Der vorliegende Ansatz geht davon aus, dass zu jedem Bild Informationen über Kamera-Position, Kamera-Richtung, Kamera-Brennweite und weitere Daten der Kamera zur Verfügung stehen. Dieser Annahme wird im praktischen Versuch dadurch entsprochen, dass die Kamera auf einer Spezialvorrichtung montiert ist, mit deren Hilfe unter anderem GPS-Daten (GPS: Global Positioning System) aufgezeichnet werden können.

Auch über die urbane Umgebung werden einige Annahmen getroffen: da die meisten Gebäude mit Hilfe von Punkten, Linienabschnitten und ggf. (ebenen) Oberflächen zu modellieren sind, werden lediglich diese Elemente im Model verwendet. Dadurch ist jedoch die Modellierung eines Gebäudes, das runde Elemente enthält (Turm oder Gebäude mit Kuppel), in diesem System nicht möglich.

Zusätzlich wird für Oberflächen angenommen, dass sie undurchsichtig sind. Diese Annahme dient v.a. dazu, hinter Oberflächen liegende Hypothesen schnell ausschließen zu können.

2.2. Datenstrukturen

Der untersuchte Ansatz definiert drei grundlegende Typen von 2D/3D-Modellierungsobjekten:

- **Punkt:** wird im Input-Bild erkannt als Endpunkt oder Kreuzung von Linienabschnitten.
- **Linienabschnitt:** wird im Input-Bild erkannt durch bekannte Kantendetektions-Algorithmen. Deren Output wird von Noise befreit, indem gerade Linien angelegt werden und zu kurze Linienabschnitte herausgefiltert werden.
- **Oberfläche:** wird nicht im Input-Bild erkannt, sondern erst im 3D-Modell, falls bestätigte Punkte und Linienabschnitte dort ein Polygon bilden.



Abb. 2-1: Punkte und Linienabschnitte im Inputbild

Des Weiteren werden die folgenden vier Datenstrukturen verwendet:

- **Features:** sie bilden, zusammen mit den Kameraparametern, den eigentlichen Input des Systems. Da Oberflächen nicht im Bild erkannt werden, gibt es zwei Typen von Features: Punkte und Linienabschnitte. Ein Punkt-Feature wird repräsentiert durch seine Bildposition und seine ausgehenden Linienabschnitte. Ein Linienabschnitts-Feature wird repräsentiert durch seine zwei Endpunkte. Beide Featuretypen enthalten Beziehungen zu den konkurrierenden Hypothesen, an denen sie beteiligt sind.
- **Extrusionen:** sie werden berechnet, indem Features durch das Kamerazentrum in den 3D-Raum projiziert werden. Entsprechend gibt es daher auch keine Oberflächen-Extrusionen. Punkt-Extrusionen werden als von einem Kamerazentrum ausgehender 3D-Strahl repräsentiert, während Linienabschnitts-Extrusionen durch einen 3D-Keil repräsentiert werden, der durch zwei von einem Kamerazentrum ausgehenden 3D-Strahlen begrenzt wird.
- **Hypothesen:** sind hypothetische 3D-Strukturen, die durch Triangulation von zwei oder mehr Extrusionen desselben Typs errechnet werden, mit Ausnahme von Oberflächen-Hypothesen. Diese werden aus bestätigten Punkten und Linienabschnitten abgeleitet. Jede Hypothese wird durch eine Schätzung ihrer 3D-Struktur und ihren Schätzfehler repräsentiert und enthält Beziehungen zu allen unterstützenden Features.
- **Elemente:** sind bestätigte Hypothesen. Jedes Element wird genauso beschrieben wie eine entsprechende Hypothese, im Unterschied zur Hypothese bleibt ein Element jedoch dauerhaft im Modell erhalten.

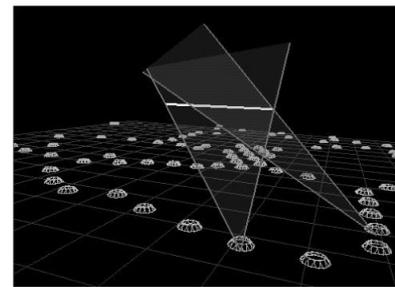


Abb. 2-2: Linienabschnitts-Extrusionen und daraus berechnete Hypothese

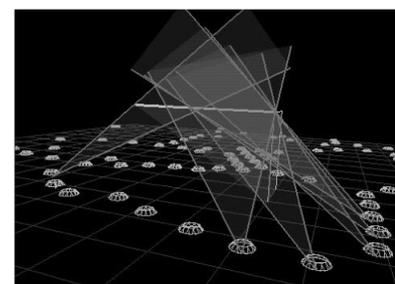


Abb. 2-3: aus der Hypothese wird ein Element

Im Idealfall sollte nun jedes erkannte Feature nur zu genau einer 3D-Struktur gehören, während die 3D-Struktur durch mehrere Features bestimmt wird (**unterstützende Features**).

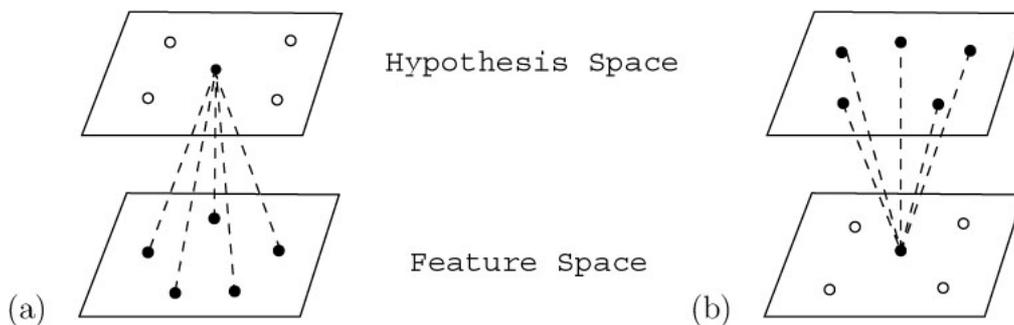


Abb. 2-4: (a) unterstützende Features einer Hypothese (1:n)

Abb. 2-4: (b) (um ein Feature) konkurrierende Hypothesen (n:1)

Da die Übereinstimmung von Features jedoch nicht manuell vorgegeben wird, sondern automatisch gefunden werden soll, müssen Features verschiedener Bilder erschöpfend miteinander zu Hypothesen kombiniert werden. Dadurch werden auch unzählige Hypothesen aus Features gebildet, die nicht genau einer tatsächlichen 3D-Struktur entstammen (diese gilt es dann jedoch zu bestimmen und auszusortieren).

Auf diese Weise taucht ein Feature in vielen Hypothesen auf (**konkurrierende Hypothesen**).

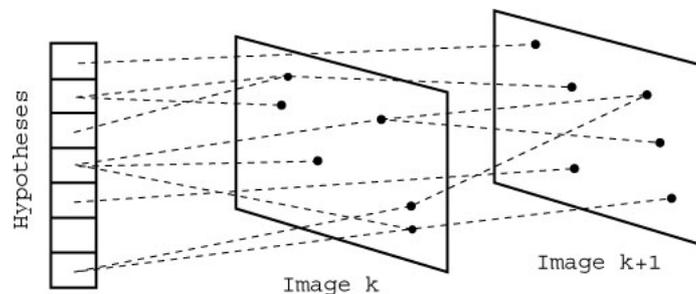


Abb. 2-5: Beziehungen zwischen Hypothesen und Features (n:n)

Sobald jedoch aus einer Hypothese ein Element bestätigt wird, werden aus den unterstützenden Features **bestätigte Features**, d.h. alle Links zu anderen Hypothesen eines Features werden abgebrochen.

2.3. Tracking-Algorithmen

Zur Erzeugung von ganz neuen Hypothesen werden Features des aktuell bearbeiteten Bildes erschöpfend mit Features aus bereits bearbeiteten Bildern kombiniert, so dass jede neue Hypothese aus einem Tupel von 2 Features besteht.

Für bereits aus früheren Schritten vorhandene Hypothesen (sogen. Langzeit-Hypothesen bzw. **Long-Term Hypotheses**) gibt es jedoch mehrere Möglichkeiten des Trackings (d.h. zur Kombination mit neuen Features):

1. Nur das am besten passende (d.h. die größte A-posteriori-Wahrscheinlichkeit erzeugende) Feature wird tatsächlich verwendet, um eine Hypothese zu aktualisieren (**Linear Exact Match**).
2. Ein gewichtetes Mittel von mehreren passenden (d.h. eine A-posteriori-Wahrscheinlichkeit oberhalb eines bestimmten Grenzwertes erzeugenden) Features wird verwendet, um eine Hypothese zu aktualisieren (**Linear Mixed Match**).
3. Bei mehreren passenden Features wird die Hypothese vervielfältigt und mit jedem passenden Feature aktualisiert, so dass jede mögliche Kombination weiterverfolgt wird (**Branching Exact Match**).

Es wurde jedoch nur die 1. Klasse tatsächlich implementiert.

2.4. Reihenfolge der Hypothesenbildung

Da auch bereits bestätigte Elemente noch mit Features neuer Bilder getestet und aktualisiert werden, gibt es mehrere denkbare Reihenfolgen der Hypothesenbildung. Als effizienteste Reihenfolge beim Abarbeiten eines Bildes hat sich jedoch die folgende herausgestellt:

- Zunächst wird das aktuell abzuarbeitende Feature mit jedem bestätigten Element kombiniert, getestet und ggf. das Element aktualisiert.
- Anschließend wird dasselbe Feature mit jeder Langzeit-Hypothese kombiniert zu einer kurzfristigen Hypothese, diese wird getestet und je nach Ausgang ein bestätigtes Element erzeugt, die Langzeit-Hypothese aktualisiert oder aber unverändert gelassen.
- Im letzten Schritt werden durch die oben beschriebene erschöpfende Kombination aller bereits vorhandenen Features mit dem aktuellen Feature neue Hypothesen generiert, getestet und entweder für die nächsten Runden beibehalten (neue Langzeit-Hypothesen) oder sofort wieder verworfen.

Diese Schritte werden für jedes Feature des aktuellen Bildes in der angegebenen Reihenfolge wiederholt.

Den genauen Ablauf für die Abarbeitung eines Bildes verdeutlicht folgende Abbildung:

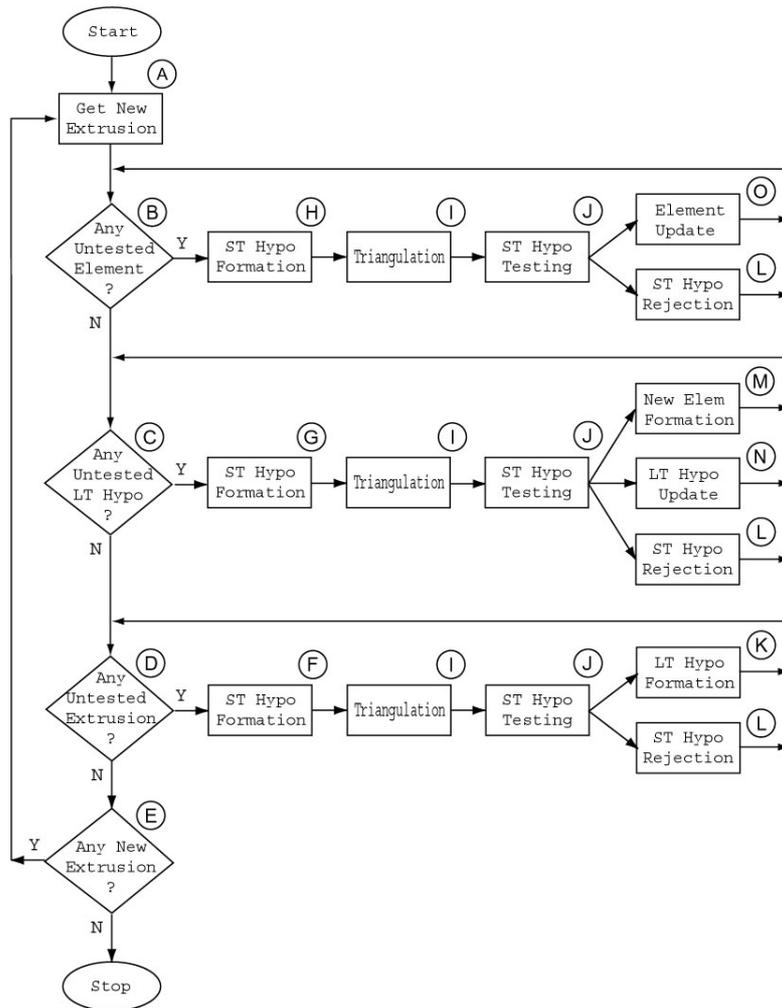


Abb. 2-6: Reihenfolge der Hypothesenbildung

3. Triangulation

Die Triangulation übernimmt zwei entscheidende Aufgaben des zu beschreibenden Ansatzes:

1. sie liefert eine Schätzung der 3D-Struktur von Elementen
2. sie liefert ein Maß, wie gut die Features übereinstimmen (und damit die Grundlage zur Berechnung einer A-posteriori-Wahrscheinlichkeit).

Zwei unterschiedliche Ereignisse können einen Triangulations-Vorgang auslösen: die Bildung einer neuen kurzfristigen Hypothese (durch Kombination eines neuen Features mit einem

Element, einer Langzeit-Hypothese oder einem anderen Feature aus bereits bekannten Bildern), oder das Update einer Langzeit-Hypothese (hervorgerufen durch Änderungen an deren unterstützenden Features).

Gefolgt wird ein Triangulations-Vorgang stets vom Test der jeweiligen kurzfristigen Hypothese (Berechnung der A-posteriori-Wahrscheinlichkeit, siehe Kapitel 4).

3.1. Triangulation in 2D

Triangulation in 2D ist ein Thema, das in Feldern wie der Astronomie oft zum Einsatz kommt und daher recht gut verstanden ist. Dabei wird ein Objekt X von zwei Punkten (Kamerazentren) A und B aus betrachtet. Aus den Betrachtungswinkeln α und β und dem Abstand d zwischen A und B lässt sich der Abstand z von X zur Grundlinie durch A und B berechnen:

$$z = \frac{d \sin \alpha \sin \beta}{\sin(\alpha + \beta)}$$

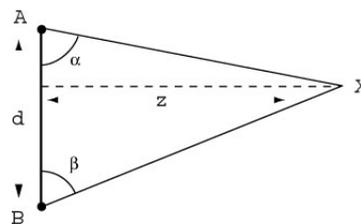


Abb. 3-1: Triangulation in 2D

Über diesen Umweg lassen sich dann auch leicht die Abstände zwischen A und X bzw. zwischen B und X bestimmen.

Anders betrachtet könnte man A, B und X jedoch auch als Punkte in einem Koordinatensystem auffassen. Kennt man die Koordinaten von A und B, so lässt sich X aus dem Schnitt der Geraden (AX) und (BX) berechnen. Letzteres Vorgehen liefert nun die Grundlage für Punkt- bzw. Linienabschnitts-Triangulation.

3.2. Punkt-Triangulation

Der Triangulations-Vorgang basiert auf der Erkenntnis, dass ein 3D-Punkt x in einem Set von Bildern I_i auf die Bild-Punkte m_i projiziert wird. Definiert man eine Punkt-Extrusion als den vom Kamera-Zentrum c_i ausgehenden Strahl durch m_i , so müsste im Umkehrschluss daher der Schnitt aller relevanten Extrusionen den Punkt x ergeben.

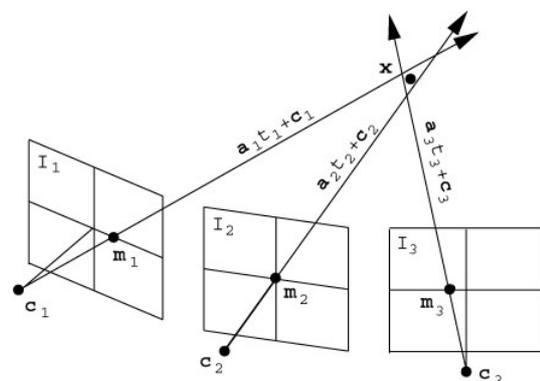


Abb. 3-2: Punkt-Triangulation

Dass dies in der Realität nicht exakt der Fall ist, liegt zum einen an der großen Zahl von Fehlerquellen (Messfehler bei Kameraparametern, Fehler beim Kantendetektions-Vorgang bzw. der anschl. Bearbeitung des Outputs, etc.), die die Berechnung der Punkt-Extrusionen beeinflussen können. Zum anderen liegt dies daran, dass eventuell Features miteinander kombiniert wurden, die in der Realität gar nicht einem Punkt x entsprechen (diese Fälle sollen dann später anhand der A-posteriori-Wahrscheinlichkeit unterschieden werden), so dass ein exakter Schnitt nicht zu erwarten ist.

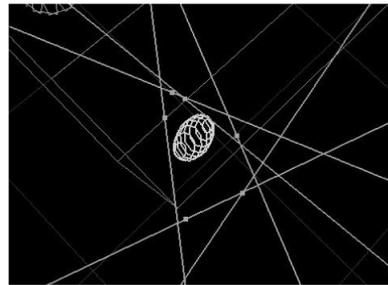


Abb. 3-3: ungenauer Schnitt bei der Punkt-Triangulation

Zunächst bedeutet dies daher, dass ein Modell entwickelt werden muss, das gewisse Fehler beim Schnitt der Extrusionen toleriert. Dies wird im vorliegenden Ansatz erreicht, indem eine zufällige Störgröße eingeführt wird. Statt zu erwarten, dass sich alle Extrusionen in einem Punkt schneiden, unterstellt man nun, dass jeder Abstand $d_i(x)$ zwischen Extrusion E_i und x (gemessen zwischen dem am nächsten liegenden Punkt y_i auf E_i und x) eine mit Kovarianz Λ_i und Mittelwert 0 normalverteilte Zufallsgröße ist.

Damit ergibt sich die bedingte Wahrscheinlichkeit, dass die Extrusionen einem gegebenen Punkt x entstammen, als Produkt der Abstands-Einzelwahrscheinlichkeiten:

$$p(Y | x) = \prod_i N(y_i - x, \Lambda_i)$$

Der Punkt x wird dann mit Hilfe der Maximum Likelihood-Methode geschätzt; nach einigen Umformungen ergibt sich, dass der wahrscheinlichste Punkt \hat{x}_{ML} bestimmt werden kann, indem die Summe der gewichteten quadrierten Abstände zwischen den Extrusionen und dem Punkt minimiert wird:

$$\hat{x}_{ML} = \arg \min_x D_v(x), \text{ wobei } D_v(x) \equiv \sum_i d_i^T(x) \Lambda_i^{-1} d_i(x)$$

Über die Bildung des Gradienten von $D_v(x)$ nach x lässt sich das gesuchte \hat{x}_{ML} daraus dann berechnen. Der Schätzfehler des Triangulations-Vorgangs ergibt sich aus $D_v(\hat{x}_{ML})$.

3.3. Linienabschnitts-Triangulation

Die Triangulation von Linienabschnitten basiert auf ähnlichen Beobachtungen wie die Punkt-Triangulation, verlangt verglichen mit dieser jedoch etwas mehr Aufwand. Da nicht jeder tatsächliche 3D-Linienabschnitt auch vollständig in den 2D-Bildern auftaucht, sondern eventuell nur ein Abschnitt davon, reicht es also nicht, nur den „Schnitt“ der Extrusionen zu betrachten. Daher ergibt sich folgendes dreistufiges Vorgehen:

1. Die Richtung der Lösungsgerade wird geschätzt (v).
2. Ein Punkt auf der Lösungsgerade wird bestimmt (p).
3. Der Abschnitt auf der Lösungsgerade wird bestimmt (p_1, p_2).

Da auch hier wieder Messfehler und Ungenauigkeiten ins Spiel kommen, bietet es sich an, zumindest Schritt 1 ebenfalls in einem wahrscheinlichkeitstheoretischen Rahmen zu entwickeln.

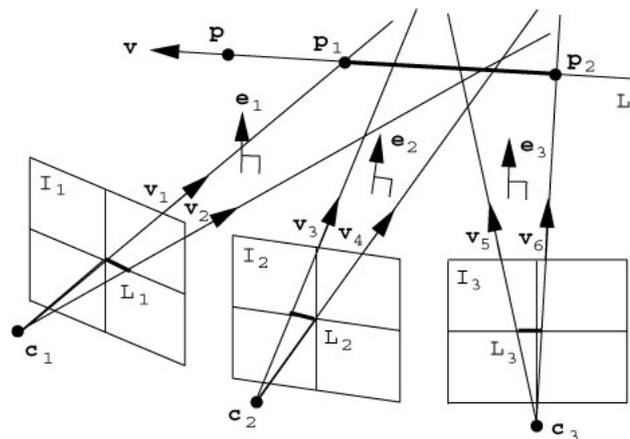


Abb. 3-4: Linienabschnitts-Triangulation

Bei einem exakten Schnitt aller Extrusions-Ebenen (diese sind eindeutig bestimmt durch die Linienabschnitts-Extrusionen) in einer Gerade läge diese Gerade in jeder Ebene, und somit wäre der Richtungsvektor v dieser Gerade orthogonal zum Normalenvektor e_i jeder Ebene. Dies ergibt das folgende Gleichungssystem:

$$e_i^T v = 0 \quad \forall i \quad \text{bzw. in Matrixschreibweise} \quad E \cdot v = 0$$

Aus oben erwähnten Gründen führt man nun erneut eine mit Kovarianz Λ_n und Mittelwert 0 normalverteilte Zufallsgröße n ein, die die von der Orthogonalität abweichenden Winkel

zwischen den Normalenvektoren und dem gesuchten Richtungsvektor modelliert:

$$E \cdot v = n$$

Damit ergibt sich die bedingte Wahrscheinlichkeit, dass die Extrusionen einem Linienabschnitt mit gegebenem Richtungsvektor v entstammen:

$$p(E | v) = N(E \cdot v, \Lambda_n)$$

Der Richtungsvektor v wird dann ebenfalls mit Hilfe der Maximum Likelihood-Methode geschätzt; nach einigen Umformungen ergibt sich, dass der wahrscheinlichste Vektor \hat{v}_{ML} durch folgende Minimierung bestimmt werden kann:

$$\hat{v}_{ML} = \arg \min_v \left\{ \frac{1}{2} (v^T \cdot E^T) \Lambda_n^{-1} (E \cdot v) \right\}$$

Über die Bildung des Gradienten des Terms nach v lässt sich das gesuchte \hat{v}_{ML} daraus dann berechnen.

Im nächsten Schritt gilt es einen Punkt auf der Lösungsgeraden zu bestimmen. Dazu wird eine Referenzebene definiert, die eine beliebige Kameraposition c_1 enthält und mit dem errechneten Richtungsvektor \hat{v}_{ML} als Normalenvektor eindeutig bestimmt ist.

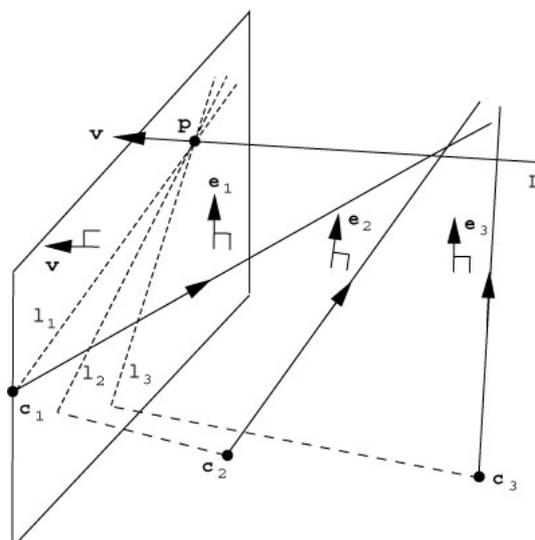


Abb. 3-5: Bestimmung des Punktes p

Mit dieser Referenzebene werden nun alle anderen Extrusions-Ebenen (e_i, c_i) geschnitten, wodurch sich Geraden l_i ergeben, die alle in der Referenzebene liegen. Der Schnitt dieser Geraden ergibt den Punkt p , der die gesuchte Schätzung für einen Punkt auf der Lösungsgeraden darstellt. Diese beiden ersten Schritte ergeben damit die Lösungsgerade $L(v, p)$.

Im letzten Schritt müssen nun noch die Endpunkte p_1 und p_2 bestimmt werden, die den Linienabschnitt eindeutig bestimmen. Dazu wird zunächst die Vereinigung aller Linienabschnitts-Extrusionen bestimmt und sichergestellt, dass diese Vereinigung stetig ist. Anschließend werden die beiden Strahlen, die die Vereinigung der Extrusionen begrenzen, mit der Lösungsgerade geschnitten. Dies ergibt schließlich die gesuchten Endpunkte p_1 und p_2 .

Der Schätzfehler errechnet sich bei der Linienabschnitts-Triangulation als Summe der quadrierten Abstände zwischen den die Extrusionen begrenzenden Strahlen und der Lösungsgeraden L .

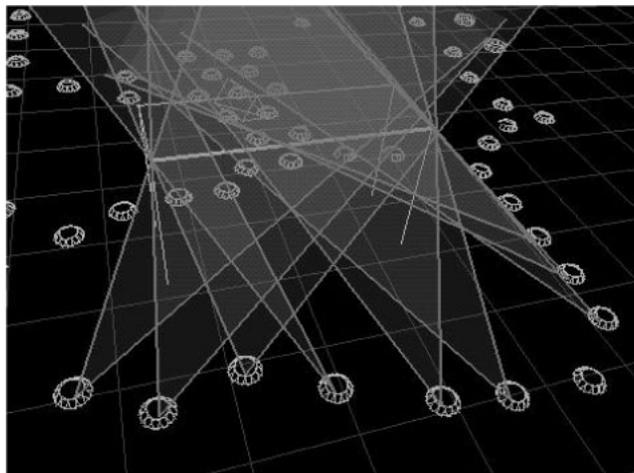


Abb. 3-6: Durch Triangulation wird eine Schätzung des ursprünglichen Linienabschnitts gewonnen

4. Auswertung von Hypothesen

Wie bereits im vorigen Kapitel angedeutet ist die A-posteriori-Wahrscheinlichkeit entscheidend, um gültige Hypothesen von ungültigen zu unterscheiden.

Da in einem Large-Scale Rekonstruktions-Projekt die Anzahl der generierten Hypothesen schnell enorm groß wird, ist es einerseits wichtig, unrichtige Hypothesen möglichst schnell zu streichen, andererseits sollte nicht voreilig eine Hypothese verworfen werden, die trotz sich anfänglich stark auswirkender Messfehler noch ein 3D-Element darstellen könnte.

Diesem Konflikt versucht der untersuchte Ansatz dadurch Rechnung zu tragen, indem zwei Wahrscheinlichkeitsmaße für die Gültigkeit einer Hypothese erarbeitet und zu einem Maß - der **A-posteriori-Wahrscheinlichkeit** - zusammengefasst werden.

Konkret sind diese beiden Maße die **Datenkonsistenz-Wahrscheinlichkeit** und die **Non-**

Accidentalness-Wahrscheinlichkeit. Diese bilden dann die Grundlage, um eine Hypothese zu beurteilen und weiter zu verarbeiten.

4.1. Datenkonsistenz-Wahrscheinlichkeit

Die Datenkonsistenz-Wahrscheinlichkeit wird abgeleitet aus der Forderung, dass Hypothesen für tatsächliche Elemente einen kleinen Triangulationsfehler aufweisen sollten. Daher wird für die Triangulationsfehler dieser Hypothesen eine Gauß-Funktion zugrunde gelegt. Für Hypothesen, die keinem tatsächlichen Element entsprechen, wird hingegen angenommen, dass der Triangulationsfehler gleichverteilt über alle möglichen Triangulationsfehler ist.

Gegeben einen beim Triangulations-Vorgang aufgetretenen Triangulationsfehler D lässt sich die Datenkonsistenz-Wahrscheinlichkeit einer Hypothese aufgrund dieser Annahmen und mit Hilfe des Bayesschen Theorems berechnen.

Zu bemerken ist hierbei, dass die Berechnung den durchschnittlichen Triangulationsfehler bzw. den absoluten Triangulationsfehler geteilt durch die Anzahl der unterstützenden Features zugrunde legt.

Das Schaubild der Datenkonsistenz-Wahrscheinlichkeit erhält die übliche Form einer traditionellen Threshold-Funktion.

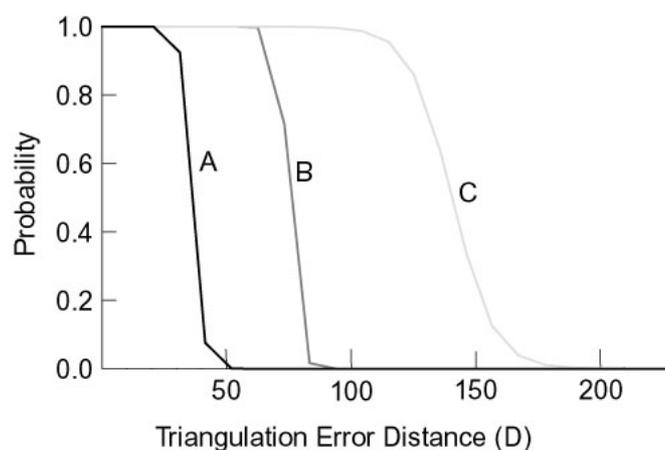


Abb. 4-1: Schaubild der Datenkonsistenz-Wahrscheinlichkeit

4.2. Non-Accidentalness-Wahrscheinlichkeit

Anschaulich scheint klar, dass mit steigender Zahl an unterstützenden Features die Wahrscheinlichkeit wächst, dass die zugehörige Hypothese nicht durch Zufall erzeugt wird. Diese Beobachtung wird formal in der Non-Accidentalness-Wahrscheinlichkeit erfasst.

Hierzu wird jede Feature-Extrusion als Ergebnis eines einstufigen Bernoulli-Experiments aufgefasst, wobei die Extrusion an einer Hypothese (zufällig oder weil sie zu einem tatsächlichen Element gehört) beteiligt oder unbeteiligt sein kann.

Definiert man F als die Anzahl der beteiligten Feature-Extrusionen, so lässt sich - gegeben, dass die Extrusionen einem tatsächlichen Element entsprechen - eine Wahrscheinlichkeit für F unterstützende Extrusionen bestimmen. Ebenso lässt sich eine Wahrscheinlichkeit für F unterstützende Extrusionen bestimmen gegeben, dass die Extrusionen keinem tatsächlichen Element entsprechen, also zufällig eine Hypothese bilden.

Unter Anwendung des Bayesschen Theorems ergibt sich hieraus die bedingte Wahrscheinlichkeit dafür, dass für eine gegebene Anzahl F von Extrusionen diese zu einem Element gehören - die Non-Accidentalness-Wahrscheinlichkeit.

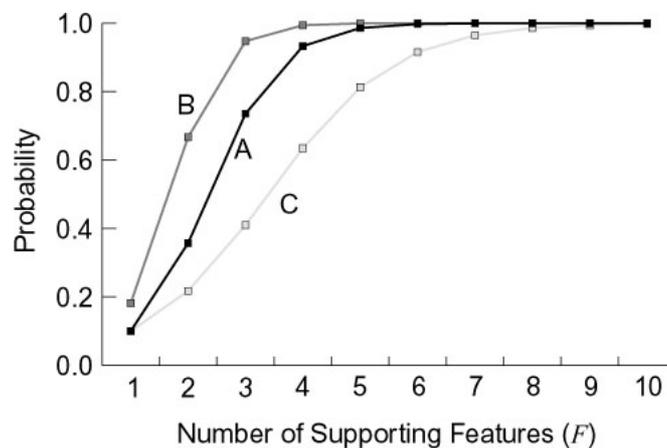


Abb. 4-2: Schaubild der Non-Accidentalness-Wahrscheinlichkeit

4.3. A-posteriori-Wahrscheinlichkeit

Erneut mit Hilfe des Bayesschen Theorems und unter Annahme der Unabhängigkeit zwischen Triangulationsfehler D und Anzahl der Feature-Extrusionen F lassen sich die beiden oben berechneten Wahrscheinlichkeiten zur A-posteriori-Wahrscheinlichkeit der Hypothese zusammenfassen.

Diese berechnet sich dann aus der folgenden Gleichung:

$$(A - \text{posteriori- Wahrsch.}) = \frac{(\text{Datenkonsistenz - Wahrsch.}) * (\text{Non - Accidentahess - Wahrsch.})}{(A - \text{priori- Wahrsch. der Hypothese})}$$

Die A-priori-Wahrscheinlichkeit der Hypothese ist dabei ein Systemparameter, der für Punkte bzw. Linienabschnitte getrennt festgelegt werden kann.

4.4. Bestätigung einer Hypothese

Jedes Mal, wenn eine Hypothese gebildet oder modifiziert wird, wird getestet, ob die A-posteriori-Wahrscheinlichkeit größer ist als ein bestimmter Schwellwert p_c . Ist dies der Fall, so wird die Hypothese als verlässlich angesehen und als ein Element bestätigt.

Dadurch werden mehrere Vorgänge ausgelöst:

1. Alle unterstützenden Features der bestätigten Hypothese werden mit dem neuen Element verbunden, Beziehungen zu alle konkurrierenden Hypothesen getrennt.
2. Konkurrierende Hypothesen gelten als modifiziert und werden daher einer Triangulation mit anschließender Hypothesen-Beurteilung unterzogen.
3. Falls das neue Element ein Linienabschnitt ist, könnte eine Oberfläche entstanden sein, und es wird die Oberflächenerkennung gestartet (s. Kap. 5.).

4.5. Ablehnung einer Hypothese

Im Verlauf ihrer Auswertung wird für jede nach 4.4. nicht bestätigte Hypothese getestet, ob ihre A-posteriori-Wahrscheinlichkeit kleiner ist als der Schwellwert p_r . Ist dies der Fall, so wird die Hypothese als so unwahrscheinlich eingestuft, dass sie ganz aus dem System entfernt werden kann. Dabei werden alle Beziehungen von unterstützenden Features zu dieser Hypothese abgebrochen, die Beziehungen dieser Features zu anderen Hypothesen bleiben jedoch unverändert. Anders als bei der Bestätigung einer Hypothese sind bei der Ablehnung also keine anderen Hypothesen betroffen.

4.6. Update einer Hypothese

Liegt die A-posteriori-Wahrscheinlichkeit einer Hypothese zwischen den oben festgelegten Schwellwerten p_c und p_r , kann zum Zeitpunkt ihrer Auswertung noch keine endgültige Entscheidung über sie gefällt werden. Dies hat zur Folge, dass die Hypothese mitsamt ihren Modifikationen (hinsichtlich unterstützenden Features, Triangulationsfehler, A-posteriori-Wahrscheinlichkeit) im System gespeichert bleibt.

In der tatsächlichen (Linear Exact Match-) Implementation wird hierbei die Auswertung aller Features eines Bildes in Kombination mit einer festen Hypothese abgewartet und nur ein Update mit dem Feature vorgenommen, das die beste A-posteriori-Wahrscheinlichkeit ergibt.

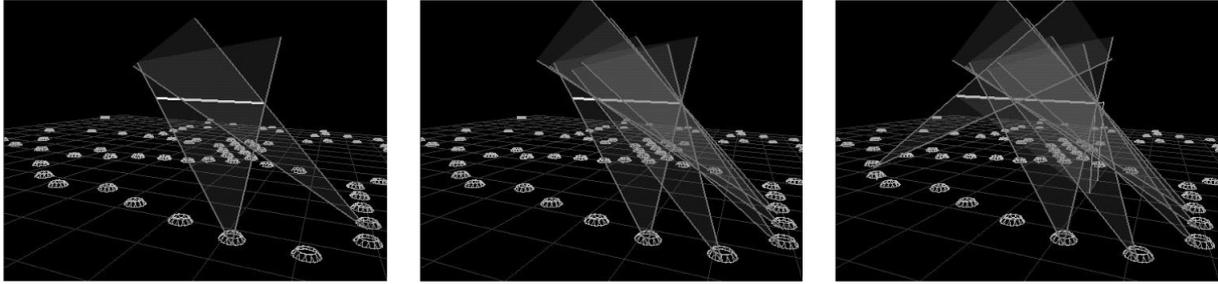


Abb. 4-3: Linienabschnitts-Hypothese von der Bildung über das Update bis zur Bestätigung

5. Oberflächen

Es werden zwei vereinfachende Annahmen über die Beschaffenheit von Oberflächen getroffen:

1. Ebenheit - erlaubt deren Modellierung durch ebene Polygone
2. Undurchsichtigkeit - erlaubt die Annahme, dass alles, was sich hinter einer Oberfläche befindet, nicht sichtbar ist

Erstere Annahme ermöglicht die Erkennung und Berechnung von Oberflächen, die zweite ermöglicht ihren Einsatz zur Erkennung und Eliminierung von ungültigen Hypothesen.

Diese zweite Art der Erkennung von ungültigen Hypothesen bietet im Vergleich zu Triangulation und Einschätzung durch A-posteriori-Wahrscheinlichkeit ein höheres Level zur Auswertung von Hypothesen. Auf diesem Level wird statt nur einer Hypothese potentiell eine ganze Menge von ungültigen Hypothesen auf einmal erkannt, die sonst möglicherweise noch lange im System verblieben wären.

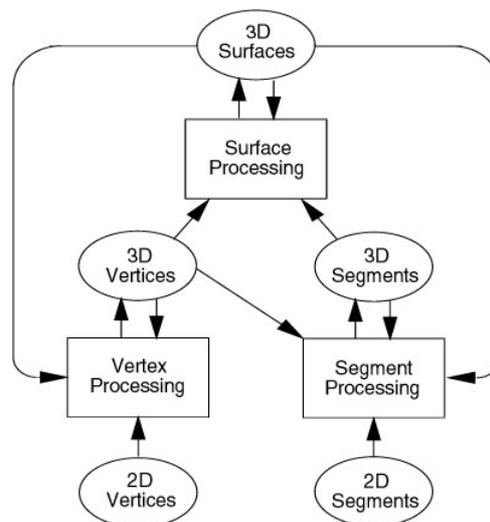


Abb. 5-1: Informationsfluss innerhalb des Systems

Durch die Ausführung der Oberflächenberechnung direkt im Anschluss an die Bestätigung eines Linienabschnitts wird eine parallele Abwicklung beider Level innerhalb des Gesamtfortschritts garantiert. Diese steigert die Effizienz des vorgestellten Systems im Vergleich zu traditionellen Systemen ohne höhere Level enorm und ermöglicht erst die Berücksichtigung von hundert oder mehr Bildern.

5.1. Oberflächenerkennung

Nach der Bestätigung eines neuen Linienabschnitts wird die Oberflächenerkennung gestartet. Diese prüft, ob an den Enden des aktuellen Linienabschnitts weitere Linienabschnitte anliegen. Ist dies der Fall, so wird eine neue Oberflächen-Hypothese gebildet. Diese wird dann mit jeder weiteren Oberflächen-Hypothese im System kombiniert und anhand der Normalenvektoren getestet, ob beide hypothetische (Teil-)Oberflächen zu einer vereinigt werden können. Verläuft der Test erfolgreich, wird die Vereinigung zu einer Hypothese durchgeführt.

Nach jeder Vereinigung wird geprüft, ob die neue hypothetische Oberfläche abgeschlossen ist. Geht auch dieser Test positiv aus, so wird die Oberflächen-Hypothese bestätigt als ein neues Oberflächen-Element.

5.2. Oberflächenberechnung

Da die Linienabschnitte, die ein Oberflächen-Element ergeben, nicht zwangsläufig exakt in einer Ebene liegen, muss eine Ebene geschätzt werden, die die Lage der Oberfläche möglichst gut beschreibt. Als Vorgehen bietet sich an, zunächst einen Punkt p in der gesuchten Ebene zu schätzen, und davon ausgehend dann den Normalenvektor n der Ebene. Hierfür werden die Endpunkte x_i der Linienabschnitte als Ausgangsdaten herangezogen.

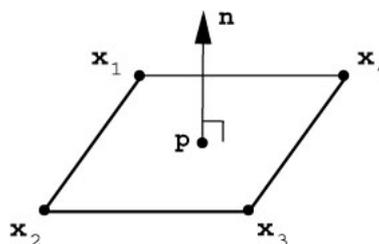


Abb. 5-2: Schätzung der Oberflächenebene durch die Endpunkte x_i der Linienabschnitte

Während die Herleitung der Schätzung für p im vorliegenden Ansatz wieder in einem wahrscheinlichkeitstheoretischen Rahmen und mit Hilfe der Maximum Likelihood-Methode geschieht, ergibt sich (unter der Annahme, dass den zufälligen Störgrößen identische Kovarian-

zen zugrunde liegen,) die tatsächliche Berechnung des Schätzwertes schlicht aus dem Mittel der Endpunkte x_i .

Benutzt man die so berechnete Schätzung für p und bestimmt denjenigen Normalenvektor n , der die Wahrscheinlichkeit des oben verwendeten allgemeinen Ansatzes maximiert, so erhält man die gesuchte Schätzung für n .

Diese Schätzungen für p und n werden dann im weiteren Verlauf zur Repräsentation des Oberflächen-Elements herangezogen.

5.3. Sichtbarkeitstest

Unter der Annahme der Undurchsichtigkeit von Oberflächen muss die geschätzte Position einer Hypothese von jedem Kamerazentrum c_i der unterstützenden Feature-Extrusionen aus sichtbar sein. Umgekehrt heißt dies, dass diejenigen Hypothesen abzulehnen sind, bei denen eine oder mehrere unterstützende Extrusionen eine Oberfläche schneiden, die geschätzte Position der Hypothese vom Kamerazentrum aus gesehen jedoch hinter dieser Oberfläche liegt. Daher werden nach der Berechnung einer neuen Oberfläche alle im System befindlichen Hypothesen auf diese Eigenschaft getestet.

Dabei wird jede unterstützende Extrusion einer Hypothese mit der betrachteten Oberfläche geschnitten. Ergibt sich kein Schnitt, oder liegt der Schnitt weiter entfernt vom Kamerazentrum c als die hypothetische Struktur, so gilt der Sichtbarkeitstest als bestanden. Andernfalls liegt eine Verletzung der Sichtbarkeits-Bedingung vor, und die Hypothese kann abgelehnt werden.

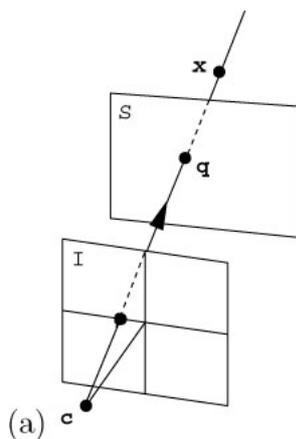


Abb. 2-4: (a) Sichtbarkeitstest für einen (hypothetischen) Punkt x

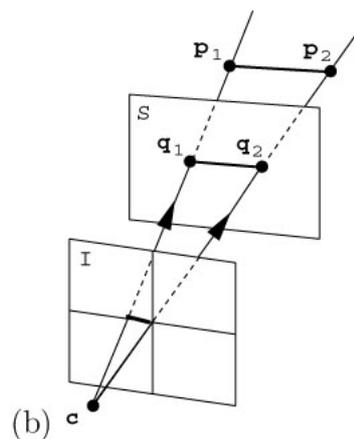


Abb. 5-3: (b) Sichtbarkeitstest für einen (hypothetischen) Linienabschnitt (p_1, p_2)

6. Ergebnisse

Die Implementation des betrachteten Ansatzes wurde mit zwei großen Datensätzen von Bildern getestet: zum einen wurden synthetische Bilder verwendet, zum anderen ein Datensatz mit echten Fotografien.

Die synthetischen Bilder sind nicht betroffen von unterschiedlichem Lichteinfall, Messfehlern bei der Kameradatenbestimmung oder von Störungen in den Bildern selbst. Fehler entstehen hier lediglich durch die Kantendetektion,



Abb. 6-1: echte Fotografie

als bei den synthetischen Bildern.

Bei der Durchführung der Tests wurde in beiden Fällen ein relativ genaues 3D-Modell der tatsächlichen 3D-Objekte generiert. Bei beiden Modellen wurden jedoch sehr vereinzelt Liniensegmente nicht gefunden, die aufgrund ihrer Lage nur in wenigen Bildern zu sehen sind. Generell wurde bei den Tests festgestellt, dass die Ergebnisse sehr stark von der Qualität der Kantendetektion abhängen.

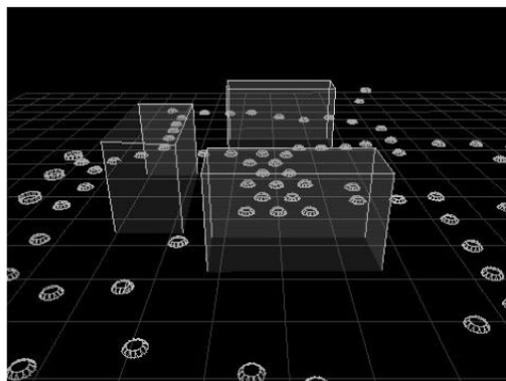


Abb. 6-3: aus den echten Fotografien gewonnenes Modell



Abb. 6-2: synthetisches Bild

durch die Abbildung in eine begrenzte Auflösung sowie durch

Ungenauigkeiten in der Computer-Arithmetik. Im Gegensatz dazu sind die echten Fotografien allen diesen Störfaktoren ausgesetzt. Aus diesem Grund wurde die Anzahl der Bilder bei diesem Datensatz auch bewusst größer gewählt

Bei einer genaueren Analyse der Teststatistiken ergab sich, dass die Genauigkeit des Modells wie erwartet bei den synthetischen Bildern höher ist. Erfreulicherweise lies sich jedoch auch feststellen, dass die Nutzung von Oberflächen für Sichtbarkeitstests eine deutlich reduzierte Zahl an am Testende verbleibenden unentschiedenen Hypothesen ergeben und somit die Qualität der Featurekorrespondenz -Bestimmung deutlich erhöht hat.

7. Zusammenfassung

In dieser Ausarbeitung wurde ein Large-Scale 3D Rekonstruktions-System vorgestellt, das durch schrittweise Abarbeitung von Bildern Featurekorrespondenz in 3D herstellt und dabei die Erzeugung von Oberflächen nutzt.

Nach einer kurzen Einführung und Motivation des Themas wurde der grobe Ablauf in einem solchen System skizziert. In Kapitel 2 wurden dann die grundlegenden Begriffe und verwendeten Datenstrukturen definiert. Im Anschluss daran wurde in Kapitel 3 die Bildung von Hypothesen und die Triangulation der beteiligten Extrusionen beschrieben, die das Kernstück des gesamten Systems darstellt. Ausgehend vom dadurch entstehenden Triangulations-Residuum und der Anzahl der beteiligten Extrusionen wurde in Kapitel 4 die Ermittlung der A-posteriori-Wahrscheinlichkeit vorgestellt, die die Grundlage zur Evaluation der Hypothesen liefert. In Kapitel 5 wurde die Erkennung und Berechnung von Oberflächen sowie deren Einsatz zur schnellen Elimination ungültiger Hypothesen beschrieben.

Im vorangehenden Kapitel wurden schließlich die Ergebnisse und Beobachtungen zusammengefasst, die bei der praktischen Umsetzung des Ansatzes erzielt werden konnten. Diese Erkenntnisse scheinen den Schluss zu erlauben, dass das vorgestellte Verfahren durchaus eines Tages praktische Anwendung finden könnte.