



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Studienarbeit

Nikolaus Rusitska

Multiple View Geometry

3D-Umgebungsmodell anhand
einer bewegten Kamera

Inhalt

1	Einleitung	3
2	Grundlagen.....	5
2.1	Kameramodellierung.....	5
2.2	Epipolargeometrie	5
2.3	Fundamentalmatrix.....	7
3	Multiple View Geometry.....	9
3.1	Auffinden von Merkmalen und Korrespondenzen	9
3.2	Verfahren zur Schätzung der Fundamentalmatrix.....	10
3.2.1	8-Punkt Algorithmus	10
3.2.2	RANSAC.....	10
3.2.3	PROSAC	10
3.2.4	LO-RANSAC.....	11
3.2.5	R-RANSAC.....	12
3.2.6	BEEM.....	12
4	Fazit	15
	Literaturverzeichnis	16

1 Einleitung

Die dreidimensionale Bildverarbeitung findet in vielen verschiedenen Bereichen Anwendung und ermöglicht verschiedene Verwendungen.

Werden verschiedene Bilder analysiert, die die gleiche Szene zeigen, wird auch von Multiple View Geometry gesprochen.

Eine große Herausforderung für die digitale Bildverarbeitung besteht in der Rekonstruktion von dreidimensionalen Objekten, von denen nur zweidimensionale Bilder vorliegen. Es kann mit kalibrierten Kameras gearbeitet werden, wodurch viele Parameter der Kamera, wie zum Beispiel die Brennweite und die Verzerrung der Linse, bekannt sind. Dieser Umstand vereinfacht die Rekonstruktion, aber schränkt das Quellmaterial stark ein und erfordert zusätzliche Metadaten.

Auch können Szenen mit stereoskopischen Kameras aufgenommen werden: Dabei sind zwei Kameras nebeneinander angeordnet, wodurch die menschliche Wahrnehmung mit zwei Augen imitiert wird. Aus den resultierenden Bildpaaren wird eine Tiefensicht abgeleitet, die dem dreidimensionalen Eindruck der menschlichen Wahrnehmung entspricht.

Eine nicht kalibrierte Rekonstruktion erweitert die Anwendungsmöglichkeiten, da bei ihr mit anderem Quellmaterial gearbeitet werden kann. Zum einen ist es leichter aufzuzeichnen, da die Kalibrierung nicht erst erstellt werden muss, zum anderen kann auch Material verwendet werden, das für einen anderen Zweck bestimmt war. So könnte beispielsweise aus mehreren Urlaubsbildern eine fotografierte Sehenswürdigkeit rekonstruiert werden.

Ein weiteres Themengebiet ist die Kameraverfolgung, bei der die Positionsänderung der Kamera während der Aufnahme von Filmen gesucht ist. Oftmals kommt aufwendige zusätzliche Technik zum Einsatz, um Transformation der Kamera im Raum zu verfolgen. Dabei wird auf Techniken wie GPS, Sonar oder Infrarotlokalisierung zurückgegriffen, um die Position der Kamera zu verfolgen. Soll eine Rotation der Kamera verfolgt werden, so verkomplizieren sich die angewandten Techniken noch weiter, bis das teilweise in die aufzunehmende Szene Markierungen eingebracht werden, die in einer späteren Nachbearbeitung wieder entfernt werden müssen.



**Ein computergeneriertes Objekt
eingefügt in eine reale Kameraaufnahme.**

(<http://www.ssangba.com/Video/Camera%20Tracking/Camera%20Tracking.html>)

Um die Bewegung einer Kamera zu verfolgen, wird eine kontinuierliche Sequenz von Bildern – ein Video – analysiert. Aus der Beziehung eines Bildes zu seinem Nachfolger wird die Bewegung der Kamera zwischen beiden Bildern abgeleitet.

Die Bewegungsverfolgung der Kamera kann bei der Produktion von Filmen eingesetzt werden, um Animationen in eine Szene einzubetten. Die Bewegungsverfolgung findet allerdings auch Anwendung im Bereich Computer Vision. Sie kann verwendet werden, um Bewegungsinformationen für autonome Systeme und Roboter zu liefern.

Die Medien bieten ein immer größer werdendes Maß an Interaktion. So haben sich Augmented Reality genannte Überschneidungspunkte zwischen Computergrafik und Realität gebildet. Ziel ist es, computergenerierte Inhalte in Bildaufnahmen der Realität einzubetten, um somit eine neue Art von Mensch-Maschine-Kommunikation anzubieten. Um dieses sinnvoll durchzuführen, ist eine sehr präzise Einbettung der computergenerierten Inhalte wichtig. So muss dabei die Beschaffenheit der Szene bekannt sein, damit beim Hinzufügen des künstlichen Objektes die Position, Drehung, Größe und Verdeckung durch andere Objekte berücksichtigt wird.

Die 3D-Rekonstruktion beschäftigt sich mit der Erstellung von dreidimensionalen Modellen einer Szene, die aus verschiedenen Blickwinkeln aufgenommen wurde. Aus Übereinstimmungen der verschiedenen Ansichten werden dreidimensionale Raumpunkte generiert, die die Grundlage eines 3D-Modells sind.

Ziel ist die Erstellung eines vollständigen dreidimensionalen Modells aus vielen Bildern eines Objektes, die eine Rundumsicht ergeben. Je mehr Bilder für das Modell verwendet werden, desto genauer und detailreicher kann es werden.

Diese Arbeit soll einen Einblick in die aktuellen Verfahren bieten, durch die Szenen analysiert werden können, von denen Bilder aus mehreren Blickwinkeln verfügbar sind. Den aktuellen Ansätzen ist gemeinsam, dass zunächst markante Punkte in den bereitgestellten Bildern gesucht werden. Diese Punkte werden dann verglichen, um Übereinstimmungen zwischen den Bildern zu erkennen.

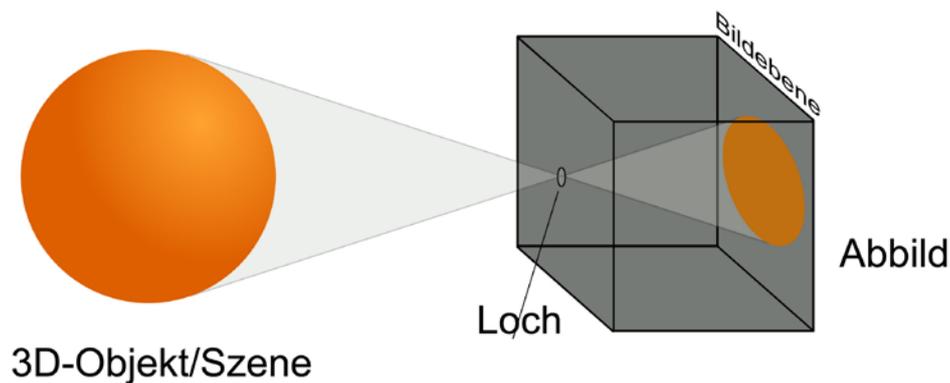
Anhand dieser Übereinstimmungen wird dann versucht eine gemeinsame Geometrie der Bilder herzuleiten. Die Stellung der verschiedenen Ansichten zueinander wird durch die sogenannte Fundamentalmatrix beschrieben.

Diese Matrix ist das zentrale Element der Multiple View Geometry: ist sie bekannt, so können verschiedene Berechnungen durchgeführt werden, mit denen das gewünschte Ziel (Kameraverfolgung, 3D-Rekonstruktion etc.) erreicht werden kann.

2 Grundlagen

2.1 Kameramodellierung

Eine Kamera kann geometrisch durch eine Lochkamera modelliert werden. Sie verfügt über ein Projektionszentrum C , über das die Punkte eines aufgenommenen Objektes auf einer Bildebene abgebildet werden. Das Projektionszentrum ist – wie bei dem realen Vorbild, der camera obscura – ein Loch in der Fokalebene der Kamera. Allerdings wird dieses Loch als unendlich klein, also als Punkt, modelliert. Orthogonal zur Fokalebene und ausgehend vom Projektionszentrum verläuft die optische Achse Z . Die Bildebene liegt parallel zur der Fokalebene hinter dem Projektionszentrum. Der Abstand von Bild- und Fokalebene wird als Kamerakonstante f beschrieben. Der Mittelpunkt der Bildebene wird Bildhauptpunkt genannt. Ein aufgenommener Raumpunkt, das Projektionszentrum und der entstehende Bildpunkt liegen auf einer Geraden. Ein aufgenommenes Objekt wird gespiegelt auf der Bildebene abgebildet.



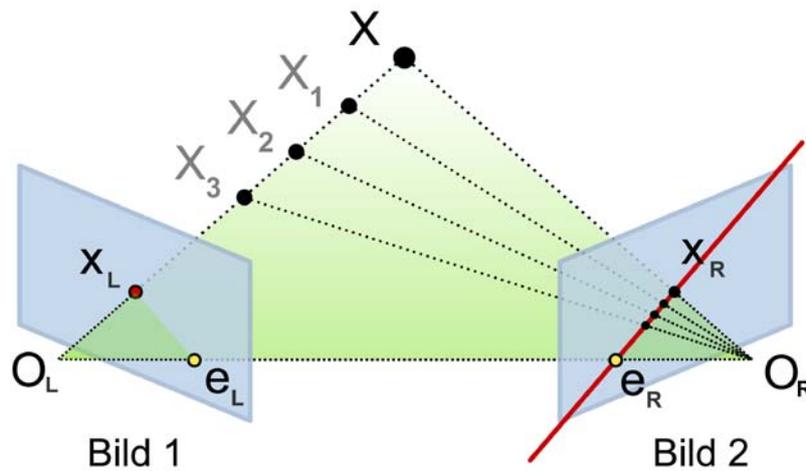
Prinzip einer Lochkamera

Das Modell kann vereinfacht werden, indem die Bildebene bei gleichbleibendem Abstand vor die Fokalebene gelegt wird. Dadurch werden die aufgenommenen Objekte nicht mehr gespiegelt auf der Bildebene abgebildet; die geometrischen Beziehungen werden dadurch jedoch nicht verändert.

Ein betrachteter dreidimensionaler Raumpunkt mit den Koordinaten (X,Y,Z) , bezogen auf das Projektionszentrum, wird auf einen zweidimensionalen Bildpunkt mit den Bildkoordinaten (x,y) , bezogen auf den Bildhauptpunkt, abgebildet.

2.2 Epipolargeometrie

Zu einem zweidimensionalen Punkt auf der Bildebene kann im dreidimensionalen Raum nur noch ein Strahl zugeordnet werden. Wird diese Szene aus einem anderen Blickwinkel mit einer weiteren Kamera betrachtet, so kann dieser Strahl als Linie auf der Bildebene der zweiten Kamera abgebildet werden. Diese Linie wird Epipolarlinie genannt. Der

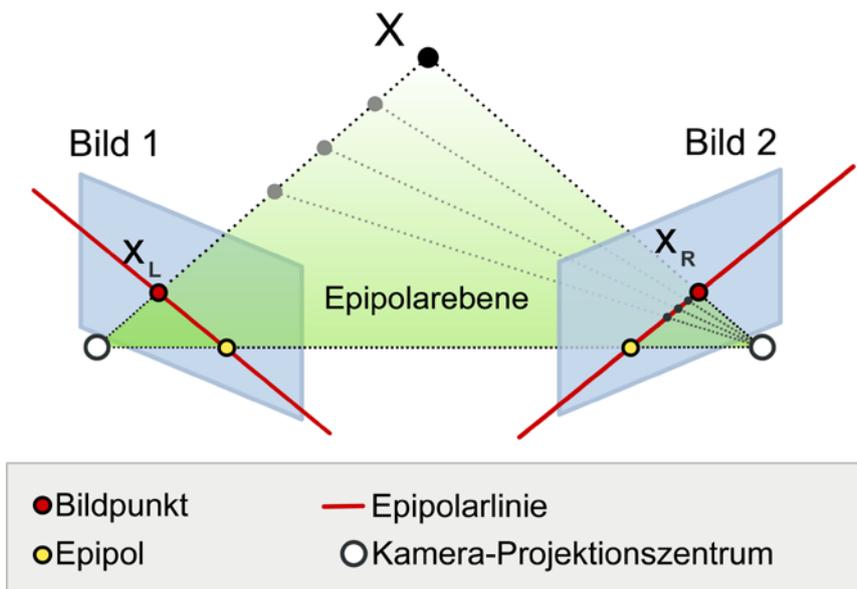


Epipolarlinie in der rechten Bildebene

Anfangspunkt dieser Linie wird Epipol genannt. Der Epipol entspricht der Abbildung des Projektionszentrums der ersten Kamera auf der Bildebene der zweiten Kamera.

Alle Strahlen, die zu betrachteten Raumpunkten zugeordnet werden, beginnen im Projektionszentrum der ersten Kamera. Daher schneiden sich auch alle Epipolarlinien der zweiten Kamera im Epipol.

Wird ein Raumpunkt auf den Bildebenen von zwei Kameras abgebildet, so bilden dessen dazugehörigen Strahlen jeweils eine Epipolarlinie auf der anderen Bildebene ab. Die Epipole der beiden Kameras liegen auf der sogenannten Grundlinie, die die beiden Kamera- bzw. Projektionszentren verbindet. Zusammen mit dem Raumpunkt spannt diese die Epipolarebene auf. Die Epipolarebene schneidet die Bildebenen wiederum in den Epipolarlinien.



Die Epipolarebene schneidet die Bildebenen in den Epipolarlinien

Jeder Raumpunkt, der auf beiden Bildebenen betrachtet wird erzeugt eine Epipolarebene. Alle Epipolarlinien schneiden sich dabei in der Grundlinie und bilden zusammen ein Ebenenbündel.

Ein Raumpunkt und sein Abbild liegen beide in der Epipolarebene, und die Epipolarebene schneidet die beiden Bildebenen in den Epipolarlinien. Daher muss die Projektion eines Bildpunktes der ersten Kamera auf der dazugehörigen Epipolarlinie in Bildebene der zweiten Kamera liegen.

2.3 Fundamentalmatrix

Mit der Fundamentalmatrix F lässt sich die Beziehung zwischen den beiden Bildpunkten eines Raumpunktes beschreiben. Zu einem Bildpunkt m der ersten Kamera lässt sich mit der Fundamentalmatrix die dazugehörige Epipolarlinie l' in der Bildebene der zweiten Kamera berechnen.

$$l' = F * m$$

Analog dazu gilt:

$$l = F^T * m'$$

Dabei gilt außerdem:

$$m'^T * F * m = 0$$

Unter Verwendung von homogenen Koordinaten lässt sich ein Gleichungssystem aufstellen um die Fundamentalmatrix zu bestimmen.

$$(x' \quad y' \quad 1) * \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

Bzw. ausmultipliziert zu:

$$x' x f_{11} + x' y f_{12} + x' f_{13} + y' x f_{21} + y' y f_{22} + y' f_{23} + x f_{31} + y f_{32} + f_{33} = 0$$

Dieses System wird durch Auffinden von Punktkorrespondenzen gelöst. Eine Punktkorrespondenz bezeichnet zwei Bildpunkte (je einen pro Kamera), die Abbildungen des selben Raumpunktes sind. Bei n gefundenen Punktkorrespondenzen ergibt sich das folgende homogene Gleichungssystem.

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} * f = A * f = 0$$

$$\text{Mit } f = \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix}$$

Im optimalen Fall würde das System mit acht Korrespondenzen gelöst werden. Allerdings werden die Korrespondenzen anhand von Pixelgrafiken gefunden, daher ergeben sich Ungenauigkeiten.

Aus diesem Grund kann die Fundamentalmatrix stets nur abgeschätzt werden. Zur Verringerung der Ungenauigkeiten kann versucht werden, das System mit mehr als acht Korrespondenzen zu lösen. Die dabei entstehende Matrix ist allerdings nicht mehr quadratisch, da sie nicht mehr genügend Spalten hat. Daher kann in diesem Fall keine inverse Matrix A^{-1} existieren, durch die der Lösungsvektor f bestimmt wird. Stattdessen muss versucht werden eine Pseudoinverse der Matrix zu finden, um dennoch zu einer Lösung zu gelangen.

3 Multiple View Geometry

3.1 Auffinden von Merkmalen und Korrespondenzen

Ein Verfahren, um prägnante Merkmale in Bildern aufzufinden, ist der SIFT-Algorithmus. Er ist in der Lage, Merkmale zu extrahieren, die dann in einem anderen Bild wiedergefunden werden können.

SIFT (Scale-invariant feature transform) ist ein Algorithmus der Merkmale in Bildern auffindet, extrahiert und beschreibt. Es ist invariant gegenüber Translation, Rotation, Skalierung, Beleuchtung und Bildrauschen.

Ein SIFT-Merkmal besteht aus Position, Orientierung und Skalierung sowie einem Deskriptor. Der Deskriptor ist ein Vektor mit 128 Elementen, durch den ein Merkmal beschrieben wird und durch den Merkmale miteinander verglichen werden können. Zwei Deskriptoren werden miteinander verglichen, indem der Euklidische Abstand der beiden Vektoren berechnet wird. Je geringer der Abstand der Vektoren ist, desto ähnlicher sind sich die Merkmale.

Korrespondenzen zwischen Merkmalen werden aufgestellt, indem die Merkmale beider Bilder miteinander verglichen werden. Dabei bilden jeweils die beiden Merkmale eine Korrespondenz, deren Ähnlichkeit am größten ist.



Markante Punkte in einem Bild

3.2 Verfahren zur Schätzung der Fundamentalmatrix

3.2.1 8-Punkt Algorithmus

Dieser Algorithmus verwendet acht oder mehr korrespondierende Punktpaare. Werden genau acht Korrespondenzen verwendet, werden ihre Koordinaten in das Gleichungssystem eingesetzt um dieses aufzulösen und die Fundamentalmatrix auszurechnen. (Vgl. Abschnitt 2.3 Fundamentalmatrix)

Werden mehr als acht Punktpaare benutzt, ergeben sich allerdings Probleme bei der Lösung des Gleichungssystems.

Um das Gleichungssystem dennoch zu lösen, kann eine Singulärwertzerlegung durchgeführt werden. Allerdings muss dabei auf Verfahren zur Singulärwertzerlegung aus der numerischen Mathematik zurückgegriffen werden. Bei der Berechnung der Singulärwertzerlegung könnten sonst starke Rundungsfehler auftreten.

Durch die Singulärwertzerlegung ergibt sich eine Pseudoinverse A^{-1} der Matrix A (A enthält die Punktkorrespondenzen). Dadurch kann das Gleichungssystem gelöst und die Faktoren der Fundamentalmatrix bestimmt werden.

3.2.2 RANSAC

Der 8-Punkt Algorithmus kann nur dann eine korrekte Fundamentalmatrix bestimmen, wenn alle Korrespondenzen, die verwendet werden, auch korrekt sind.

Der Random Sample Consensus ist im Stande auch bei fehlerbehafteten Daten eine gute Fundamentalmatrix zu bestimmen.

Es wird aus der verfügbaren Menge an Korrespondenzen zufällig eine kleine Menge ausgewählt und anhand dieser eine Fundamentalmatrix errechnet.

Fundamentalmatrix und Korrespondenzen können überprüft werden, indem mit einem Punkt einer Korrespondenz und der Fundamentalmatrix die dazugehörige Epipolarlinie errechnet wird. Der zweite Punkt der Korrespondenz muss dann auf der Epipolarlinie liegen. Ist dieses nicht der Fall, ist entweder die Fundamentalmatrix oder die Korrespondenz falsch.

Auf diese Weise wird die berechnete Fundamentalmatrix gegen alle gefundenen Korrespondenzen überprüft. Es wird dabei gezählt, wie viele Korrespondenzen die aktuelle Fundamentalmatrix unterstützen. Die Menge der unterstützenden Korrespondenzen wird Support-Set genannt. An der Größe des Support-Sets wird die Qualität der Matrix bestimmt.

Dieser Vorgang wird wiederholt, bis eine genügend qualitative Fundamentalmatrix gefunden werden konnte. Hierbei können allerdings viele Iterationen nötig sein, bis ein ausreichend gute Fundamentalmatrix gefunden wurde.

3.2.3 PROSAC

Der Progressive Sample Consensus ist eine Verbesserung des RANSAC-Verfahrens durch Einschränkung und Spezialisierung. Das PROSAC-Verfahren ist auf Umsetzungen

festgelegt, bei denen die Punktkorrespondenzen mithilfe des SIFT-Verfahrens gefunden wurden.

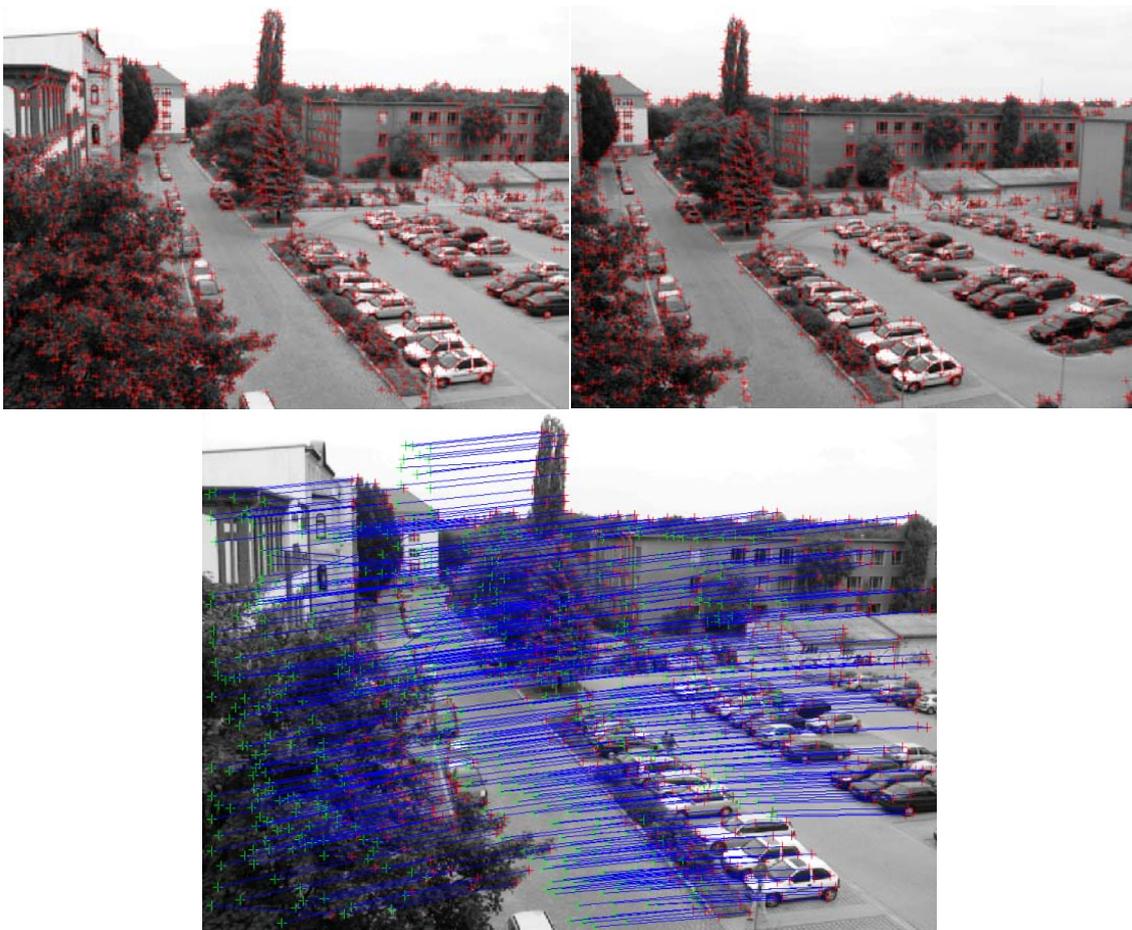
Während bei RANSAC die zu verwendenden Korrespondenzen rein zufällig gewählt werden, betrachtet PROSAC die Ähnlichkeit der Deskriptoren der SIFT-Merkmale, die die Korrespondenz bilden. Anstatt zufällig zu wählen, werden Korrespondenzen bevorzugt, bei denen sich beide Merkmale besonders ähnlich sind.

Das weitere Verfahren verläuft analog zur Vorgehensweise des RANSAC-Verfahrens.

3.2.4 LO-RANSAC

Der Locally Optimized RANSAC Ansatz enthält einen zusätzlichen Schritt zur Optimierung der Fundamentalmatrix. Wenn festgestellt wird, dass die erzeugte Matrix die momentan beste ist, wird eine neue Fundamentalmatrix aus dem gesamten Support-Set der Matrix berechnet. Dieses dient zur Verbesserung der gefundenen Matrix.

Es wird nicht von vornherein mit mehr als acht Korrespondenzen gearbeitet, da dieses die Zahl der Iterationen stark erhöhen würde, da damit auch die Wahrscheinlichkeit der Berücksichtigung von falschen Korrespondenzen steigen würde.



**Zwei Ansichten der gleichen Szene mit markanten Punkten (oben)
und das Support-Set der F-Matrix, die mit RANSAC berechnet wurde (unten)**

3.2.5 R-RANSAC

Randomized RANSAC enthält eine Abbruchbedingung um die Einschätzung der Qualität der berechneten Fundamentalmatrix zu beschleunigen.

Bevor die errechnete Fundamentalmatrix gegen sämtliche Korrespondenzen getestet wird, wird zuerst eine kleine Menge von Korrespondenzen zufällig ausgewählt. Nur wenn die Fundamentalmatrix erfolgreich gegen diese Korrespondenzen getestet wurde, wird sie danach auch gegen alle Korrespondenzen getestet um die Qualität einzuordnen, ansonsten wird sie verworfen.

Durch dieses Vorgehen werden schlechte Fundamentalmatrizen schnell aussortiert, wodurch Zeit bei der Überprüfung gespart wird, da aussortierte Matrizen nur ein sehr geringes Support-Set gehabt hätten.

3.2.6 BEEM

Der BEEM-Algorithmus (Balanced Exploration and Exploitation Model) verwendet verschiedene Zustände, aus denen heraus die Fundamentalmatrix abgeschätzt wird. Dabei werden der aktuelle Zustand, das aktuelle Modell der Fundamentalmatrix und der Vorgängerzustand berücksichtigt. BEEM ist beschränkt auf Korrespondenzen, die durch das SIFT-Verfahren gefunden wurden.

Zusätzlich wird eine neue Methode zur Berechnung der Fundamentalmatrix eingeführt, die allerdings auf Korrespondenzen beschränkt ist, die mithilfe von SIFT gefunden wurden. Dieses 2-SIFT genannte Verfahren ist im Stande, die Fundamentalmatrix aus zwei SIFT-Korrespondenzen zu berechnen.

Der BEEM-Algorithmus kennt fünf verschiedene Zustände:

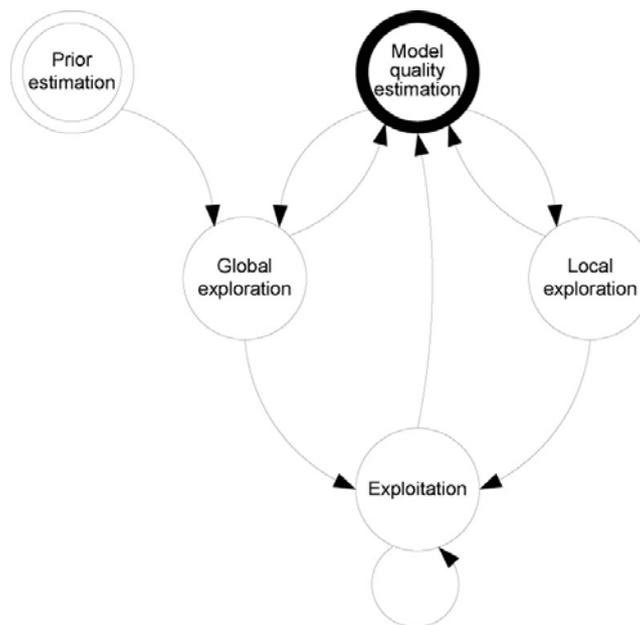
- Prior estimation
- Global exploration
- Model quality estimation
- Local exploration
- Exploitation

Prior estimation: Dieses ist der Startzustand, den das System auch ausschließlich am Anfang annimmt. Mithilfe der Ähnlichkeiten der SIFT-Korrespondenzen wird versucht, die Wahrscheinlichkeit zu ermitteln, mit der die Korrespondenz korrekt ist. Diese Wahrscheinlichkeit wird in den späteren Zuständen wieder aufgegriffen. Danach geht das System in den „Global exploration“-Zustand über.

Global exploration: In diesem Zustand wird eine Fundamentalmatrix abgeschätzt und danach ihre Qualität geprüft. Vergleichbar zum PROSAC-Verfahren werden für die Berechnung Korrespondenzen ausgewählt, die mit großer Wahrscheinlichkeit korrekt sind.

Es wird das 2-SIFT-Verfahren angewendet, wodurch nur zwei Korrespondenzen benötigt werden, um die Fundamentalmatrix zu errechnen.

Ist das Support-Set der berechneten Matrix größer als die der bisherigen, wird danach in



Zustandsdiagramm des BEEM-Algorithmus

den „Exploitation“-Zustand gewechselt, ansonsten in den „Model quality estimation“-Zustand.

Model quality estimation: Es wird die Qualität der bisher besten gefundenen Fundamentalmatrix abgeschätzt. Basis der Schätzung ist die Größe des Support-Sets, das zur Matrix gehört, sowie die Anzahl der Durchgänge, die der Algorithmus bisher durchlaufen hat.

Anhand dieser Schätzung wird das weitere Vorgehen bestimmt. Der Algorithmus kann in den Zustand „Global exploration“ oder „Local exploration“ wechseln oder terminieren.

Local exploration: Auch in diesem Zustand wird eine Fundamentalmatrix berechnet. Es werden Korrespondenzen aus dem Support-Set der bisher besten Matrix verwendet, sowie Korrespondenzen die nicht darin enthalten sind. Aus beidem wird gemeinsam eine Fundamentalmatrix errechnet und geprüft. Dabei werden je vier Koeffizienten des Gleichungssystems aus dem Support-Set entnommen und vier aus dem Rest. Bei der Verwendung der Restlichen Korrespondenzen können die vier Koeffizienten aus einer Korrespondenz entnommen werden (wie beim 2-SIFT-Verfahren beschrieben), im Support-Set darf aus einer Korrespondenz nur ein Koeffizient entnommen werden. Ist das Support-Set der neuen Matrix das größte, so wird in den „Exploitation“-Zustand gewechselt, anderenfalls in den „Model quality estimation“-Zustand.

Exploitation: In diesem Zustand wird wie im LO-RANSAC-Verfahren die Fundamentalmatrix verbessert. Das Vorgehen wurde dabei allerdings noch erweitert. Wurde eine neue Matrix berechnet und ihr Support-Set bestimmt, so startet der Zustand erneut. Dieses wird wiederholt, bis mehrere Durchgänge keine weitere Verbesserung gebracht haben.

Danach wird in den „Model quality estimation“-Zustand gewechselt.

2-SIFT: Das 2-SIFT-Verfahren verwendet zusätzliche Informationen der SIFT-Merkmale, die eine Korrespondenz bilden. Aus Position und Orientierung eines SIFT-Merkmals werden vier Punkte generiert. Zusammen mit den generierten Punkten aus dem korrespondierenden SIFT-Merkmal ergeben sich so vier Punktkorrespondenzen aus einer SIFT-Korrespondenz.

Somit können aus zwei SIFT-Korrespondenzen alle acht Koeffizienten abgeleitet werden, die zur Bestimmung der Fundamentalmatrix notwendig sind. Die daraus resultierende Matrix hat zwar keine gute Qualität, allerdings wird sie beim BEEM Algorithmus auch nur dazu verwendet, ein erstes Support-Set zu erhalten. Danach werden Matrizen im „Exploitation“-Zustand verbessert, oder es werden im „Local exploration“-Zustand verbesserte Matrizen abgeleitet.

4 Fazit

Die betrachteten Algorithmen und Erweiterungen bieten verschiedene Möglichkeiten der Spezialisierung auf schnelle oder akkurate Berechnung. Bei der Verwendung muss abgewägt werden, welche Eigenschaft wichtiger ist.

Es erscheint durchaus erfolgversprechend, bestehende Verfahren mit Erweiterungen aus anderen Verfahren zu verknüpfen, um einen verbesserten Kompromiss aus Geschwindigkeit und Genauigkeit zu erreichen.

Zur Verbesserung wäre es möglich, eine Erweiterung für den BEEM-Algorithmus zu suchen, um den Algorithmus zu beschleunigen.

Auch kann versucht werden, selbst einen effizienten und genauen Algorithmus zu entwerfen. Ansatzpunkte dafür könnten die Methoden PROSAC und 2-SIFT sein, da sie spezielle Eigenschaften der mit SIFT gefundenen Korrespondenzen berücksichtigen. Der LO-RANSAC Ansatz kann dabei hilfreich sein, die gefundenen Ergebnisse zu verbessern.

Desweiteren kann es großen Nutzen bringen, die Überprüfung der gefundenen Fundamentalmatrix zu beschleunigen, da diese einen großen Aufwand bedeutet.

Abschließend lässt sich sagen, dass es bereits gute Methoden zur Ermittlung der Geometrie mehrerer Bilder gibt, die allerdings auf Grund des großen Aufwands und der großen Menge an Informationen die Bilder beinhalten noch Raum für Verbesserungen bieten.

Literaturverzeichnis

- Chum, O., & Matas, J. (2005). Matching with PROSAC - Progressive Sample Consensus. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Los Alamitos: IEEE Computer Society.
- Goshen, L., & Shimshoni, I. (2006). Balanced Exploration and Exploitation Model Search for Efficient Epipolar Geometry Estimations (BEEM). In *9th European Conference on Computer Vision (ECCV 2006)* (pp. 151-164).
- Hartley, R. (1997). In Defense of the Eight-Point Algorithm. In *Pattern Analysis and Machine Intelligence 19* (pp. 580–593).
- Hartley, R., & Zisserman, A. (2003). *Multiple View Geometry in Computer Vision* (2. Ausg.). Cambridge: Cambridge University Press.
- Heyden, A., & Pollefeys, M. (2004). Multiple View Geometry. In G. Medioni, & S. B. Kang, *Emerging Topics in Computer Vision* (pp. 45-107). Prentice Hall.
- Košecká, J. (2004). Multiple-View Geometry for Image-Based Modeling. *Siggraph Course #23*. Los Angeles.
- Lowe, D. (2004). Distinctive Image Features from Scale-Invariant Keypoints. In *International Journal of Computer Vision 60* (pp. 91-110).
- Thormählen, T., & Seidel, H.-P. (2008). 3D-Modeling by Ortho-Image Generation from Image Sequences. In *ACM Transactions on Graphics, Vol. 27, No. 3, Article 86*.
- Wan, C., & Sato, J. (2008). Computing Multiple View Geometry in Space-Time from Mutual Projections of Multiple Cameras. *19th International Conference on Pattern Recognition (ICPR) 2008*, (pp. 1-4).