



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Ausarbeitung

Gregory Föll

Entwicklung einer robusten Methode zur
Berechnung von Stereokorrespondenzen

Gregory Föll

Entwicklung einer robusten Methode zur
Berechnung von Stereokorrespondenzen

Ausarbeitung eingereicht im Rahmen der Veranstaltung Masterseminar
im Studiengang Informatik (Master)
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer : Prof. Dr. rer. nat. Kai von Luck
Zweitgutachter : Prof. Dr. rer. nat. Gunter Klemke

Abgegeben am 3. März 2011

Inhaltsverzeichnis

1	Einführung	4
2	Merkmalsextraktion	6
2.1	Aussortieren falscher Korrespondenzen	6
3	Projektive Transformation	8
4	Bildsegmentierung	10
4.1	Testen der Segmente auf lokale Ebenheit	10
4.2	Vorteile dieser Methode gegenüber üblichen Ansätzen	12
5	Dynamische Programmierung	13
6	Zusammenfassung, Ausblick und Risiken	14
	Literaturverzeichnis	15

1 Einführung

Seit vielen Jahren wird das Thema Stereo Vision in der Informatik intensiv erforscht. Es wird versucht aus zwei oder mehr Bildern, die ein Objekt gleichzeitig aus verschiedenen Blickwinkeln zeigen, die räumliche Ausdehnung des Objektes möglichst genau zu errechnen. Die große Herausforderung dabei ist die Herstellung der Korrespondenzen zwischen den Pixeln der einzelnen Bilder. Durch die sogenannte Rektifizierung, werden die Pixel der Bilder so angeordnet, dass die Korrespondierenden Punkte aller Bilder auf der gleichen horizontalen Linie (Epipolarlinie) liegen [Foell (2010)]. Obwohl die Rektifizierung den Suchraum nach Korrespondenzen drastisch reduziert, bleibt die korrekte Zuordnung der Pixel ein schwieriges Problem, da zum Beispiel Bildrauschen, Spiegelungen oder homogene Objektoberflächen die Aufgabe erheblich erschweren.

Spätestens seitdem Microsoft Kinect, eine Kamera die das Steuern der Videospielekonsole Xbox 360 ermöglicht, auf den Markt gebracht hat, ist die Time-of-Flight-Technologie für die breite Masse bezahlbar geworden. Bei diesem Verfahren werden Lichtimpulse von der Position der Kamera ausgesendet, welche vom betrachteten Objekt reflektiert und von der Kamera wieder aufgefangen werden. Durch das Messen der Zeit, die das Licht bis zum Objekt und wieder zurück benötigt, kann die Entfernung bis zu jedem einzelnen Bildpunkt sehr genau bestimmt werden. Es wird das komplette Bild gleichzeitig durch Lichtimpulse ausgeleuchtet. Aus den gewonnenen Einzeldistanzen kann eine akkurate Tiefenkarte erstellt werden. Obwohl die Time-of-Flight-Technologie verarbeitungstechnisch schnell ist und die Ergebnisse nicht durch Rauschen, Homogenität oder schlechte Bildausleuchtung beeinflusst werden, bleibt das Thema Stereo Vision weiterhin interessant, denn in einigen Bereichen ist Time-of-Flight nur begrenzt oder gar nicht einsetzbar.

In vorhergehenden Untersuchungen wurde festgestellt, dass die meisten gängigen Verfahren im Bereich Stereo Vision sehr unsaubere Tiefenkarten bei realem Bildmaterial erzeugen. In dieser Ausarbeitung wird eine Methode vorgestellt, die bekannte Verfahren kombiniert und versucht, bessere Ergebnisse für reale Bilder im Innen- und Außenbereich zu erzielen. Ziel ist es eine Methode zu entwickeln, die möglichst robust gegenüber Rauschen, Spiegelungen und homogenen Flächen ist. Die Geschwindigkeit des Verfahrens soll zunächst nicht von Bedeutung sein. Die vorgestellte Methode wird im Rahmen der Masterarbeit realisiert.

Zunächst wird in den Bildern nach robusten Merkmalen gesucht, durch die einige zuverlässige Korrespondenzen hergestellt werden können. Im zweiten Schritt wird das erste Bild an-

hand der gefundenen Korrespondenzen in kleine Bereiche aufgeteilt. Wird dann festgestellt, dass ein Teilbereich aus dem ersten Bild, im zweiten Bild dieselbe lokale Ebene beschreibt, dann können für jeden Pixel dieses Bereiches, mit Hilfe der projektiven Transformation, korrespondierende Pixel im zweiten Bild errechnet werden. Auf diese Weise werden bereits bei der Vorverarbeitung viele Korrespondenzen ohne aufwändige Berechnungen für jeden Pixel sehr zuverlässig hergestellt. Anschließend wird dynamische Programmierung (DP) eingesetzt um die verbleibenden Korrespondenzpaare zu finden. Der DP-Algorithmus kann sich an den Korrespondenzen, die im vorhergehenden Schritt gefunden wurden, orientieren. Er wird dadurch robuster und wird erheblich beschleunigt, da nicht mehr jedes einzelne Pixel des Bildes berechnet werden muss.

2 Merkmalsextraktion

Aus einem Bild könne verschiedene Merkmale extrahiert werden um z.B. anhand dieser Merkmale ein Objekt in einem anderen Bild wieder zu erkennen. Oft müssen für einen bestimmten Zweck die richtigen Merkmale ausgewählt werden. Merkmale können zum Beispiel Ecken, Kanten, Muster oder Konturen sein.

Für die gegebene Aufgabenstellung wird ein Extraktionsverfahren benötigt, welches die gefundenen Merkmale pixelgenau und sehr zuverlässig Objekten zuordnen kann. Zwei bekannte Verfahren die diese Voraussetzungen erfüllen und oft zum Einsatz kommen sind *Scale Invariant Feature Transform* (SIFT) [[Lowe \(2004\)](#)] und *Speeded Up Robust Features* (SURF) [[Bay u. a. \(2008\)](#)]. Die durch diese Verfahren extrahierten Merkmale verfügen über eine große Invarianz gegenüber Rauschen, Beleuchtung, Skalierung, Translation und Rotation.

Beide Verfahren zeigen eine gute Ausführungsgeschwindigkeit. Eigene und andere Untersuchungen (siehe beispielsweise [[Juan und Gwun](#)]) haben gezeigt, dass SURF deutlich schneller als SIFT arbeitet, SIFT aber im Gegensatz zu SURF, in den meisten Fällen, deutlich mehr Merkmale aus dem Bildmaterial extrahieren kann.

Das zu entwickelnde Verfahren soll das bestmögliche Ergebnis liefern, die Ausführungsgeschwindigkeit soll zunächst nicht von Bedeutung sein. Deshalb ist SIFT das Merkmalsextraktionsverfahren der Wahl, da hier mehr zuverlässige Korrespondenzpaare gefunden werden.

2.1 Aussortieren falscher Korrespondenzen

Die Pixelkorrespondenzen werden von SIFT entweder sehr genau oder ganz willkürlich zugeordnet. Um die wahllosen Zuordnungen herauszufiltern, kann der *Random Sample Consensus* (RANSAC) Algorithmus [[Fischler und Bolles \(1981\)](#)] benutzt werden. RANSAC ist ein robuster Algorithmus, welcher verschiedene Ausgleichsverfahren unterstützt, um ein Modell innerhalb einer Reihe von Messwerten mit Ausreißern zu schätzen.

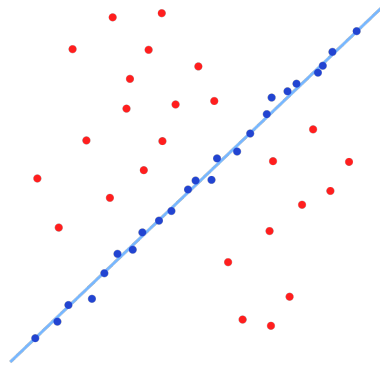
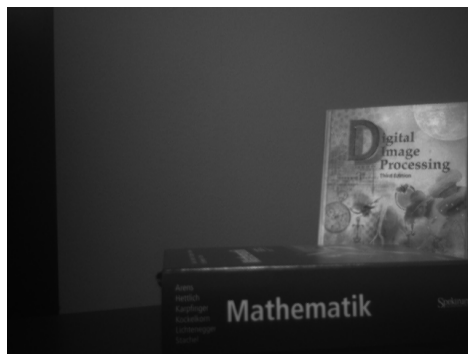


Abbildung 2.1: RANSAC zum Anpassen einer Geraden an Datenpunkte



(a)



(b)

Abbildung 2.2: Rektifizierte Aufnahmen einer Stereokamera (a) links (b) rechts

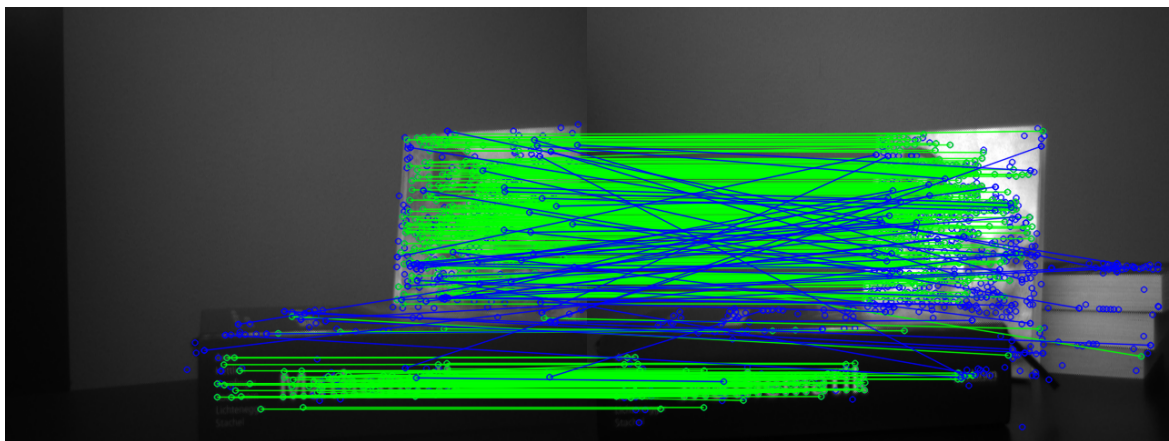
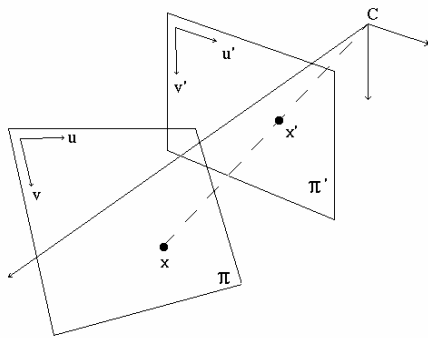


Abbildung 2.3: SIFT: richtige Zuordnungen (grün) durch RANSAC gefilterte Falschzuordnungen (blau)

3 Projektive Transformation



Wird ein Objekt von zwei Kameras aus unterschiedlichen Blickwinkeln aufgenommen, entstehen durch die Projektion eines 3D Objektes auf eine 2D Bildebene, in den Bildern unterschiedliche perspektivische Verzerrungen. Bei diesen Verzerrungen handelt es sich um projektive Transformationen, bei zweidimensionalen Flächen auch als Homographie bezeichnet. Die Aufnahmen lassen sich durch eine projektive Rücktransformation wieder entzerren. Zu einem Punkt p lässt sich sein korrespondierender, verzerrter Punkt p' berechnen, in dem p mit der 3x3-Matrix H multipliziert wird.

$$p' = Hp$$

Da die Punkte jeweils eine homogene Koordinate besitzen, gilt $p = (x, y, 1)^T$ und $p' = (x', y', k')^T$. Nach dem Einsetzen erhält man die folgende Darstellung:

$$\begin{pmatrix} x' \\ y' \\ k' \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Nach dem Ausmultiplizieren muss der 3D-Vektor p' durch Teilen durch die homogene Koordinate k' normalisiert werden. Es ergeben sich also folgende Gleichungen:

$$\begin{aligned} x' &= \frac{h_{11}x + h_{12}y + h_{13}}{k'} \Rightarrow h_{11}x + h_{12}y + h_{13} - x'k' = 0 \\ y' &= \frac{h_{21}x + h_{22}y + h_{23}}{k'} \Rightarrow h_{21}x + h_{22}y + h_{23} - y'k' = 0 \end{aligned}$$

Und mit $k' = h_{31}x + h_{32}y + h_{33}$ erhält man nach dem Einsetzen:

$$h_{11}x + h_{12}y + h_{13} - x'(h_{31}x + h_{32}y + h_{33}) = 0$$

$$h_{21}x + h_{22}y + h_{23} - y'(h_{31}x + h_{32}y + h_{33}) = 0$$

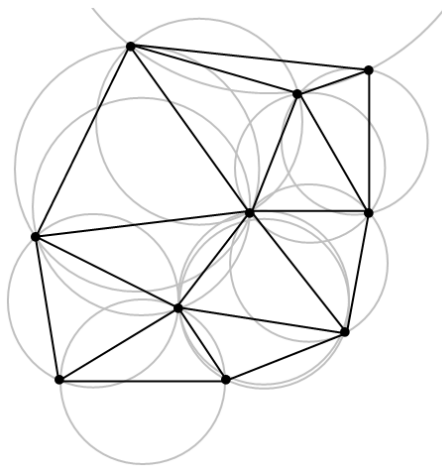
Durch Umformung ergibt sich ein Gleichungssystem der Art $A * h = 0$:

$$\begin{pmatrix} x & y & 1 & 0 & 0 & 0 & -x'x & -x'y & -x' \\ 0 & 0 & 0 & x & y & 1 & -y'x & -y'y & -y' \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ . \\ . \\ . \\ h_{33} \end{pmatrix} = 0$$

Bei n Punktpaaren hat A die Dimension $2n * 9$. Um das Gleichungssystem zu lösen, muss A den Rang 8 haben, d.h. es werden 8 Punktkorrespondenzen benötigt.

Mit Hilfe der Singulärwertzerlegung (SVD) lässt sich die Homographiematrix H auch mit nur 4 Punktkorrespondenzen berechnen (siehe [[Hartley und Zisserman \(2004\)](#)]).

4 Bildsegmentierung



Mit den im ersten Schritt gefundenen robusten Punkten, wird das erste (linke) Bild mit Hilfe der Delaunay-Triangulation in kleine Bereiche aufgeteilt. Die Delaunay-Triangulation ist ein Verfahren, das aus einer gegebenen Punktmenge ein Dreiecksnetz erstellt. Die Punkte werden so zu Dreiecken vernetzt, dass innerhalb des Kreises, auf dem die drei Dreieckspunkte liegen, keine anderen Punkte aus der vorgegebenen Punktmenge enthalten sind. Damit erfüllen alle Dreiecke des Dreiecksnetzes die sogenannte Umkreisbedingung. Die Dreiecke weisen so möglichst große Innenwinkel auf. Diese Eigenschaft minimiert Rundungsfehler und ist deswegen sehr erwünscht.

4.1 Testen der Segmente auf lokale Ebenheit

Da für die Delaunay-Triangulation SIFT-Korrespondenzen benutzt wurden, ist sichergestellt, dass die Bildbereiche, die durch die, bei der Triangulation gebildeten, Dreiecke beschrieben werden, in beiden Bildern vorhanden sind. Da aber nur das linke Bild segmentiert wird, ist nicht klar, ob alle Dreiecke aus dem ersten Bild auch die Umkreisbedingung im zweiten Bild erfüllen. Durch den anderen Blickwinkel, könnte es sein, dass ein Bereich, der im ersten Bild durch ein Dreieck beschrieben wird, im zweiten Bild durch ein anderes Objekt verdeckt wird und so nicht mehr dieselbe lokale Fläche wie im ersten Bild beschreibt. Sollte das verdeckende Objekt selbst SIFT-Merkmale aufweisen, so müssen diese Merkmale in dem Bereich liegen, der durch die projektive Transformation errechnet wird. Diese Bereiche/Dreiecke können nach solcher Prüfung gleich ausgeschlossen werden. Sollte das verdeckende Objekt

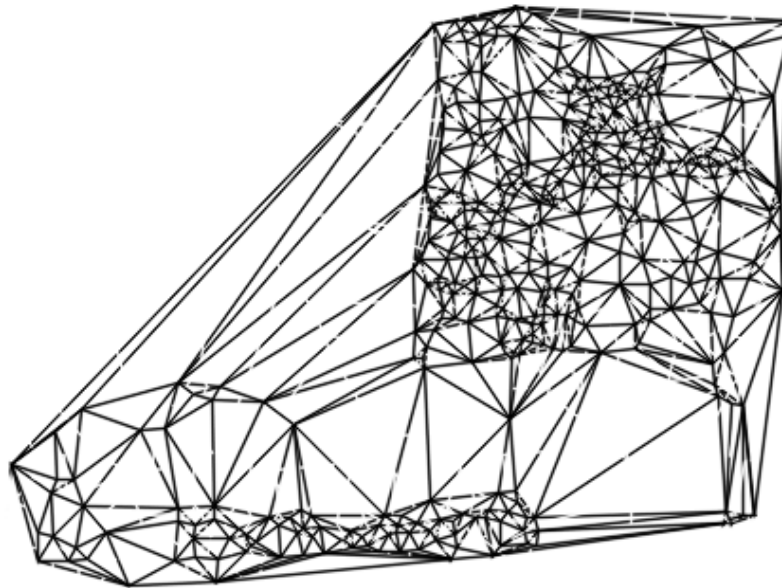


Abbildung 4.1: Delaunay-Triangulation für das Bild [2.2\(a\)](#)

keine SIFT-Merkmale aufweisen, dann kann geprüft werden, ob ein Pixel und sein korrespondierendes Pixel, welches durch die projektive Transformation errechnet wird, annähernd die gleiche Intensität besitzen. Ein großer Intensitätssprung weist auf ein anderes Objekt hin.

Nachdem alle falschen Zuordnungen für das zweite Bild ausgeschlossen sind, ist immer noch nicht bekannt, ob jedes Dreieck im ersten Bild auch wirklich lokal eine Ebene beschreibt. Es könnte sein, dass eines der Eckpunkte eines Dreiecks auf unterschiedlichen Objekten liegt, eine Projektion so eines Dreiecks in das zweite Bild wäre damit falsch. In der Literatur finden sich verschiedene Ansätze zur Lösung dieses Problems.

In [[Hartley und Zisserman \(2004\)](#)] wird eine sehr elegante Methode beschrieben mit der man testen könnte, ob ein Dreieck eine Ebene beschreibt.

$$H^T F + F^T H = 0$$

Die Gleichung besagt, dass die Matrix H für die projektive Transformation, kompatibel mit der Fundamentalmatrix F sein muss. Die Fundamentalmatrix beschreibt die Epipolargeometrie zwischen zwei Bildern für alle Punkte, während die Abbildungsmatrix H nur auf Punkte angewendet werden kann, die in einer Ebene liegen. Leider ist diese Methode sehr empfindlich gegenüber Rauschen und ist deswegen nicht für reales Bildmaterial geeignet.

In [[Li u. a. \(2006\)](#)] werden komplexe Bildanalysen durchgeführt, um die Geometrie des abgebildeten Objektes zu erkennen.

In [Chen u. a.] werden Tiefeninformationen mittels Triangulation für jeden Eckpunkt eines Dreiecks berechnet und durch einen Schwellwert wird geschätzt, ob alle Eckpunkte in derselben Ebene liegen könnten.

Sobald mit einem geeigneten Ansatz alle ebenen Dreiecke gefunden sind, wird für jedes Dreieck eine Abbildungsmatrix berechnet und die korrespondierenden Punkte im zweiten Bild können einfach durch die projektive Transformation ermittelt werden. Da man für die Berechnung der Abbildungsmatrix H mindestens vier Punktkorrespondenzen benötigt, kann H für ein Dreieck mit Hilfe eines nächstliegenden benachbarten Dreiecks berechnet werden. Dabei muss wieder geprüft werden, ob die benachbarten Dreiecke in derselben lokalen Ebene liegen.

4.2 Vorteile dieser Methode gegenüber üblichen Ansätzen

Ein großes Problem beim Stereo-Matching ist auch die Wahl der richtigen Disparitätsparameter. Üblicherweise wird davon ausgegangen, dass sich die Disparitäten [Foell (2010)] im Bildmaterial kontinuierlich ändern. Das Bild wird Zeile für Zeile nach möglichen korrespondierenden Punkten durchsucht. Wird dabei ein markanter Übergang, zum Beispiel eine Ecke, erkannt, muss der Disparitätsparameter neu angepasst werden. Diese Vorgehensweise erfordert eine sehr aufwendige und rechenintensive Bildanalyse um die Disparität dynamisch anzupassen. Da sogar im gleichen Bild viele sehr unterschiedliche Regionen vorkommen können, ist es in der Praxis nicht immer möglich die richtigen Parameter zu finden. In [Gong und Yang (2003)] wird ausführlicher auf dieses Problem eingegangen.

Mit der projektiven Transformation könne auf einen Schlag für einen ganzen Pixelbereich mit wenig Rechenaufwand die korrespondierenden Pixel berechnet werden. Ist sichergestellt, dass alle Pixel des transformierten Bereichs in derselben Ebene liegen, können die berechneten Korrespondenzen als sehr zuverlässig angesehen werden. Da die gefundenen Korrespondenzen über mehrere Bildzeilen verstreut liegen, kann sich die dynamische Programmierung an diesen bekannten, zuverlässigen Punkten orientieren um die Disparität entsprechend anzupassen. Damit erhält man eine automatische Anpassung der Disparitätsparameter ohne zusätzlichen Rechenaufwand.

5 Dynamische Programmierung

Algorithmen der dynamischen Programmierung (DP) werden oft für Optimierungsprobleme eingesetzt, die aus vielen gleichartigen Teilproblemen bestehen. Die Teilprobleme eines umfangreichen Problems werden unabhängig von einander gelöst. Die Lösungen der kleinsten Teilprobleme werden so lange zur Lösung eines nächstgrößeren Problems zusammengesetzt, bis das gesamte Problem gelöst ist. Ergebnisse der Teilprobleme werden in Tabellen gespeichert, auf die bei Bedarf zurückgegriffen wird. Dynamische Programmierung ist wegen der Einfachheit und Effektivität der Algorithmen sehr beliebt und wird auch bei vielen Stereo-Matching-Verfahren eingesetzt.

Nach Untersuchung einiger Stereo-Matching-Verfahren die mit dynamischer Programmierung arbeiten [[Gong und Yang \(2003\)](#); [Veksler](#); [Gonzalez u. a.](#); [Salmen u. a.](#)], wurde klar, dass kein vorhandener Algorithmus wirklich für die zu lösende Aufgabe passt und deshalb ein eigener Algorithmus entwickelt werden muss. Da durch die projektive Transformation bereits viele Punkte vorhanden sind, an denen sich der Algorithmus orientieren kann, sollte der zu entwickelnde Algorithmus bessere Ergebnisse liefern als herkömmliche DP-Verfahren.

6 Zusammenfassung, Ausblick und Risiken

In dieser Ausarbeitung wurde ein Verfahren zur Berechnung von Stereokorrespondenzen vorgestellt, welches im Rahmen der Masterarbeit implementiert werden soll. Das vorgestellte Verfahren kombiniert bekannte Methoden aus der Mathematik und Bildverarbeitung, wie Projektive Transformation, Delaunay-Triangulation, SIFT und Dynamische Programmierung, und verfolgt das Ziel, vor allem bei schlechtem Bildmaterial, robuster als herkömmliche Ansätze zu sein.

In vorhergehenden Arbeiten wurden die meisten Teilschritte des Verfahrens gründlich untersucht und getestet. Aus diesem Grund sind keine großen Risiken bei der Implementierung zu erwarten. Ein Risiko könnte der schlecht abschätzbare Zeitaufwand für die Entwicklung des DP-Algorithmus sein.

Literaturverzeichnis

- [Aziz 2009] AZIZ, Muhammad M.: *Colour Stereo Matching with Automated Camera Gain Adjustment in Different Lighting Conditions*, LITIS Laboratory, INSA de Rouen, France, Masters Thesis, 2009
- [Bay u. a. 2008] BAY, Herbert ; ESS, Andreas ; TUYTELAARS, Tinne ; GOOL, Luc V.: *Speeded-Up Robust Features (SURF)*. 2008. – Forschungsbericht
- [Bradski und Kaehler 2008] BRADSKI, Gary ; KAEHLER, Adrian: *Learning OpenCV*. 1. O'Reilly Media, Inc., 2008
- [Chen u. a.] CHEN, Chao-I ; SARGENT, Dusty ; TSAI, Chang-Ming ; WANG, Yuan-Fang ; KOPPEL, Dan: *Stabilizing Stereo Correspondence Computation Using Delaunay Triangulation and Planar Homography* / University of California Santa Barbara. – Forschungsbericht
- [Eugene 2011] EUGENE, Khvedchenya: *Comparison of feature descriptors*. 2011. – URL <http://computer-vision-talks.com/2011/01/comparison-of-feature-descriptors/>
- [Fischler und Bolles 1981] FISCHLER, Martin A. ; BOLLES, Robert C.: *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*. In: *Communications of the ACM* (1981), June, Nr. vol. 24 no. 6, S. 381–395
- [Foell 2010] FOELL, Gregory: *Stereo Vision: Vergleich verschiedener Algorithmen zur Lösung des Korrespondenzproblems* / Hamburg University of Applied Sciences. 2010. – Forschungsbericht
- [Geiger u. a.] GEIGER, Andreas ; ROSER, Martin ; URTASUN, Raquel: *Efficient Large-Scale Stereo Matching* / Dep. of Measurement and Control, Karlsruhe Institute of Technology. – Forschungsbericht
- [Gong und Yang 2003] GONG, Minglun ; YANG, Yee-Hong: *Fast Stereo Matching Using Reliability-Based Dynamic Programming and Consistency Constraints*. In: *Proceedings of the Ninth IEEE International Conference on Computer Vision* (2003), Nr. vol. 2

- [Gonzalez u. a.] GONZALEZ, Rafael C. ; CANCELAS, Jose A. ; ALVAREZ, Juan C. ; FERNANDEZ, Jose A. ; ALVAREZ, Ignacio: *Dynamic Programming Stereo Vision Algorithm for Robotic Applications* / Universidad de Oviedo. – Forschungsbericht
- [Hartley und Zisserman 2004] HARTLEY, R.I. ; ZISSERMAN, A.: *Multiple View Geometry in Computer Vision*. Second. Cambridge University Press, ISBN: 0521540518, 2004
- [Holzhäuser 2008] HOLZHÄUSER, Dennis: *Medizinische Bildverarbeitung: Eine Zusammenfassung der wichtigsten Vorlesungsthemen*. 2008
- [Juan und Gwun] JUAN, Luo ; GWUN, Oubong: *A Comparison of SIFT, PCA-SIFT and SURF* / Chonbuk National University. – Forschungsbericht
- [Li u. a. 2006] LI, Ping ; FARIN, Dirk ; GUNNEWIEK, Rene K. ; DE, Peter H. N.: *On Creating Depth Maps from Monoscopic Video using Structure from Motion*. 2006
- [Lowe 2004] LOWE, David G.: *Distinctive Image Features from Scale-Invariant Keypoints* / University of British Columbia. 2004. – Forschungsbericht
- [Mengelkoch 2008] MENGELKOCH, Marco: *Texturierte 3-D-Mesh Generierung aus Stereobildsequenzen*, Universität Koblenz Landau, Diplomarbeit, 2008
- [Miethe 2005] MIETHE, Martin: *Entwicklung eines Stereokamerasystems zur Objekterkennung und Entfernungsbestimmung in mobilen Robotersystemen*, Technische Universität Darmstadt, Diplomarbeit, 2005
- [Miethe 2009] MIETHE, Martin: *Graph-Cut basierte Stereorekonstruktion*, Technische Universität Dresden, Großer Beleg, 2009
- [Min u. a.] MIN, Dongbo ; YEA, Sehoon ; VETRO, Anthony: *Temporally Consistent Stereo Matching Using Coherence Function* / MITSUBISHI ELECTRIC RESEARCH LABORATORIES. – Forschungsbericht
- [Palme 2007] PALME, Elia: *Stereo Vision Library for Obstacle Avoidance Applications*, University of Fribourg, Masters Thesis, 2007
- [Salmen u. a.] SALMEN, Jan ; SCHLIPSING, Marc ; EDELBRUNNER, Johann ; HEGEMANN, Stefan ; LÜKE, Stefan: *Real-Time Stereo Vision: Making more out of Dynamic Programming* / Institut für Neuroinformatik, Ruhr-Universität Bochum. – Forschungsbericht
- [Short 2009] SHORT, Nathaniel J.: *3-D Point Cloud Generation from Rigid and Flexible Stereo Vision Systems*, Virginia Polytechnic Institute and State University, Masters Thesis, 2009
- [Sun 2002] SUN, Changming: *Fast Stereo Matching Using Rectangular Subregioning and 3D Maximum-Surface Techniques*. In: *International Journal of Computer Vision* (2002), May, Nr. vol. 47 no. 1/2/3, S. 99–117

[Tomasì und Birchfield] TOMASI, Carlo ; BIRCHFIELD, Stan: Depth Discontinuities by Pixel-to-Pixel Stereo / Stanford University. – Forschungsbericht

[Veksler] VEKSLER, Olga: Stereo Correspondence by Dynamic Programming on a Tree / University of Western Ontario. – Forschungsbericht