

3D-INFORMATION AUS MEHRFACHEN ANSICHTEN

Bernd Neumann

Fachbereich Informatik
Universitaet Hamburg
Schlueterstr.70, D-2000 Hamburg

Zusammenfassung

Dieser Bericht gibt einen Ueberblick ueber die Moeglichkeiten, raeumliche Information aus mehreren Ansichten einer Szene zu extrahieren, ohne dabei Vorwissen ueber den Szeneninhalt auszunutzen. Zwei Verfahrensgruppen werden erlaeutert, Binokularstereo und Bewegungsstereo. In der ersten Gruppe werden zwei Ansichten aus bekannten Blickrichtungen ausgewertet. Hier liegen die Schwierigkeiten vorwiegend in der Bestimmung korrespondierender Bildelemente. Bewegungsstereo basiert auf bewegungsbedingten Unterschieden in einer Bildfolge. Es wird ueber Ansaetze zur formalen Analyse der prinzipiellen Moeglichkeiten und ueber erste Experimente mit praktischen Verfahren berichtet.

1. Einleitung

Eines der grundlegenden Probleme der maschinellen Bildinterpretation besteht darin, die bei der Abbildung einer dreidimensionalen Bildvorlage verlorengegangene dritte Dimension wiederzugewinnen. Das menschliche Sehvermoegen beweist, dass dies unter verschiedenen Umstaenden moeglich sein kann: mit zwei Augen oder mit einem, bei bekannten oder unbekanntem Bildinhalten, aus fester oder bewegter Betrachterposition, in Szenen mit oder ohne Bewegung, und unter unterschiedlichsten Beleuchtungsverhaeltnissen. Mehrere Mechanismen sind heute bekannt [ROCK 75], die zum raeumlichen Sehvermoegen des Menschen beitragen. Man kann sie in zwei Gruppen einteilen: monokulare und binokulare, je nach dem, ob ein oder zwei Bilder gleichzeitig ausgewertet werden. Zur ersten Gruppe gehoeren Tiefenhinweise wie Texturgradienten, Schattierung und Schattenwurf, sowie perspektivische Effekte bei Gegenstaenden, deren 3D-Form bekannt ist. Einige dieser monokularen Tiefenhinweise werden in BAJCSY 80 diskutiert. KANADE 79 untersucht monokulares raeumliches Sehen anhand von Annahmen ueber Eigenschaften der realen Welt. Er zeigt an eindrucksvollen Beispielen, dass untypische Gegenstaende, die den Annahmen nicht entsprechen, raeumlich falsch interpretiert werden.

Zur zweiten Gruppe gehoeren das bekannte Binokularstereo sowie verschiedene Mechanismen, die auf Bewegung beruhen. Die zusammenfassende Bezeichnung "binokulare Tiefenhinweise" ist verbreitet, jedoch nicht genau zutreffend - beim bewegungsbedingten raeumlichen Sehen sind keineswegs "zwei Augen" erforderlich! Charakteristisch fuer diese Gruppe ist vielmehr die Tatsache, dass raeumliche Informationen aus mehreren Ansichten gewonnen werden. Bei Binokularstereo handelt es sich in der Regel um gleichzeitige Ansichten aus verschiedenen Positionen, bei Bewegungsstereo um zeitlich aufeinanderfolgende Ansichten.

Die vorliegende Arbeit befasst sich ausschliesslich mit Verfahren dieser zweiten Gruppe, und zwar nicht im Hinblick auf eine Erklarung oder Simulation menschlicher Fahigkeiten sondern aus der Sicht der maschinellen Bildinterpretation. Es wird versucht, einen Ueberblick ueber Theorie, Implementierung und Eigenschaften der wichtigsten bisher untersuchten Ansaetze zu vermitteln.

Kapitel 2 behandelt Binokularstereo. Das zugrundeliegende Konzept der Triangulierung ist zwar sehr einfach, jedoch muessen vorher Bildpunkte in Korrespondenz gebracht werden, und dieses Problem ist bis heute noch nicht befriedigend geloest.

Bewegungsstereo stellt in mancher Hinsicht eine Verallgemeinerung von Binokularstereo dar und schliesst letzteres auch ein. Denn zwei Ansichten aus verschiedenen Blickwinkeln koennen natuerlich auch zeitlich nacheinander durch Bewegung eines monokularen Beobachters erzeugt werden. Bewegungsstereo kann sich jedoch auch auf wesentlich komplexere Situationen beziehen, wo z.B. mehrere Objekte mit unbekannter Form auf unbekanntem Trajektorien in Bewegung sind und eine zeitliche Folge von vielen Ansichten zur Verfuegung steht.

In Kapitel 3 werden Verfahren vorgestellt, in denen die Positionsveraenderungen einiger markanter Bildelemente zur raeumlichen Analyse herangezogen werden. Eine typische Situation besteht z.B. darin, dass einzelne Punkte auf der Oberflaeche eines starren Objektes verfolgt werden und daraus die relativen 3D-Abstaende dieser Punkte sowie die Objekttrajektorie berechnet werden. Die mathematischen Zusammenhaenge sind wesentlich komplexer als die Triangulierungsrechnungen bei Binokularstereo. Unter welchen Umstaenden eine eindeutige Loesung existiert und wie man sie zuverlaessig findet, ist bis heute nicht vollstaendig erforscht. Eine neue geschlossene Loesung wird fuer eingeschraenkte Bewegungen in orthographischer Projektion vorgestellt. Sobald mehr als ein Objekt oder eine nicht-starre Objektkonfiguration auftritt, muessen zusaetzliche Probleme geloest werden, z.B. das Gruppieren der bewegten Bildelemente nach ihrer Zugehoerigkeit zu einzelnen Objekten oder Objektteilen. Hier ist erst ueber wenige Loesungsvorschlaege zu berichten.

Eine andere Klasse von Verfahren, die sich aber letzten Endes auch auf mehrere Ansichten stuetzt, ist durch das Stichwort "optischer Fluss" gekennzeichnet. Dieser Begriff wurde urspruenglich fuer die Bildveraenderungen auf der Retina eines bewegten menschlichen Beobachters gepraegt [GIBSON 50] und bezeichnet das Vektorfeld von Pixelgeschwindigkeiten, das durch Bewegung (von Beobachter oder Szenenteilen) induziert wird. Verfahren zur Bestimmung und Auswertung von optischem Fluss [KOENDERINK und VAN DOORN 77, CLOCKSIN 78+80, WILLIAMS 80, HORN und SCHUNCK 80, PRAZDNY 79+80] unterscheiden sich wesentlich von den vorher angesprochenen Verfahren; eine ausfuehrliche Diskussion muss hier unterbleiben, um den Rahmen dieser Uebersicht nicht zu sprengen.

Ebenfalls nicht betrachtet werden photometrische Verfahren zur 3D-Analyse. Hierzu zaehlen alle Ansaetze, die 3D-Information aus den Intensitaetsveraenderungen (und nicht den geometrischen Verschiebungen) von korrespondierenden Pixeln in mehrfachen Ansichten ableiten. An dieser Stelle sei nur auf die Arbeiten von Woodham [WOODHAM 78+80] und HORN et al. 78 hingewiesen, die 3D-Information aus den durch Beleuchtungsveraenderungen bedingten Unterschieden zwischen mehreren Ansichten

einer statischen Szene gewinnen. Eine Anwendung desselben Grundgedankens auf industrielle Objekterkennung findet sich in VAMOS und BATHOR 80. Auch bewegungsbedingte Beleuchtungsänderungen können zur 3D-Analyse herangezogen werden, siehe dazu WOODHAM 80 und NEUMANN 80a.

2. Binokularstereo

Gegenstand dieses Kapitels sind Ansätze zur räumlichen Analyse von Szenen durch Auswerten zweier Ansichten, die aus verschiedenen - aber bekannten - Blickwinkeln erzeugt werden. Im allgemeinen handelt es sich dabei nicht um radikal verschiedene Blickwinkel sondern um Anordnungen, die im Prinzip dem menschlichen zweiaugigen Sehen entsprechen.

Die Bedeutung von zwei unterschiedlichen Ansichten für das räumliche Sehen war schon Leonardo da Vinci bekannt - für ihn eine wichtige Einsicht bezüglich der Möglichkeit, "lebensechte" Gemälde zu schaffen. Die genauen geometrischen Beziehungen, die Binokularstereo zugrundeliegen, wurden jedoch erst im 19. Jahrhundert aufgedeckt, nachdem der Brite Wheatstone das Stereoskop erfunden hatte. Die Entwicklung von maschinellem Binokularstereo erfolgte zunächst vorwiegend für Zwecke der Photogrammetrie. Das Bildmaterial wurde dabei meist zunächst gefilmt und dann später mit Rechnerhilfe ausgewertet. Diesbezügliche Referenzen finden sich in HANNAH 74 und YAKIMOVSKY und CUNNINGHAM 78.

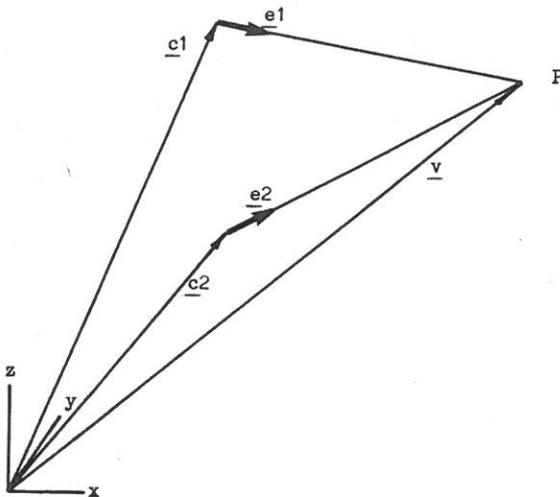
Als wichtige Teilprobleme stellten sich die folgenden Aufgaben heraus:

- (i) Bestimmen der Kameraparameter (Kalibrierung)
- (ii) Generieren von korrespondierenden Punktpaaren (Korrespondenzproblem)
- (iii) Berechnen der 3D-Koordinaten

Der letzte Schritt beinhaltet die eigentliche Entfernungsberechnung durch Triangulierung. Die mathematischen Zusammenhänge werden im folgenden kurz erläutert (siehe Fig. 1). Zwei Kameras befinden sich in den Positionen \underline{c}_1 und \underline{c}_2 (in einem globalen Koordinatensystem). Ein Punkt P mit unbekanntem Ortsvektor \underline{v} erscheint in den Kamerabildern auf Abbildungsstrahlen mit unterschiedlichen Einheitsvektoren \underline{e}_1 und \underline{e}_2 . Im Idealfall schneiden sich die Abbildungsstrahlen, und aus

$$\underline{v} = \underline{c}_1 + s_1 \underline{e}_1 = \underline{c}_2 + s_2 \underline{e}_2$$

lassen sich Faktoren s_1 und s_2 ermitteln, die die Position von P auf den Strahlen festlegen und \underline{v} bestimmen. Die obige Beziehung ist ein überbestimmtes Gleichungssystem, das bei fehlerhaften Ausgangsdaten im allgemeinen nicht befriedigt werden kann - die Abbildungsstrahlen schneiden sich dann nicht. Aus verschiedenen Gründen ist dieses eher die Regel als die Ausnahme. Man nimmt als Lösung meist ein \underline{v} , das mittig zwischen den Strahlen an der Stelle ihrer engsten Annäherung liegt. Die entsprechenden Ausdrücke für s_1 und s_2 finden sich in DUDA und HART 73.



Figur 1: Entfernungsberechnung durch Triangulierung

Eine der ersten Untersuchungen zum Einsatz von Binokularstereo in der Robotik wurde von HANNAH 74 durchgeführt. Sie verwendet von Hand fotografierte Diapositive, mit einem Abstand von 1 bis 50 Metern zwischen den Kamerastandpunkten. Davon werden Ausschnitte mit einer typischen Auflösung von 150×150 Pixeln digitisiert. Im Kalibrierungsschritt werden 2 Brennweiten sowie 5 weitere Parameter bestimmt, die die Position und Orientierung der zweiten Kamera relativ zur ersten festlegen. Hannah wendet dazu die Triangulierungsgleichungen auf manuell selektierte Punkte an, deren Korrespondenz genau bekannt ist. In einem iterativen Verfahren werden die Kameraparameter so lange veraendert, bis die Summe aller Abstände zwischen den zu den Punkten gehoerenden Abbildungsstrahlen zu einem Minimum wird. Dazu werden mindestens 14 Kalibrierungspunkte benoetigt. Ein aehnliches Verfahren wird auch in SOBEL 74 und GENNERY 77 beschrieben.

Der Hauptteil von Hannahs Untersuchungen bezieht sich auf Verfahren zur Korrespondenzermittlung. Dabei sollen diejenigen Bildteile in den zwei Ansichten identifiziert werden, die dasselbe Stueck Realwelt darstellen. Als "Bildteile" werden rechteckige Fenster in verschiedenen Groessen (z.B. 15×15) verwendet, der Vergleich erfolgt mit verschiedenen Korrelations- und Differenzmassen. Folgende Aspekte und Ergebnisse dieser Untersuchungen scheinen bedeutsam.

1. Die Suche nach einem korrespondierenden Bildteil braucht lediglich auf einem schmalen Streifen zu erfolgen, der durch die Projektion des Abbildungsstrahls aus dem ersten Bild in das zweite Bild bestimmt wird.
2. Die Intensitaeten korrespondierender Pixel koennen sich aufgrund der unterschiedlichen Blickwinkel sowohl durch einen Skalierungsfaktor als auch einen konstanten Zuschlag unterscheiden. Die normalisierte Korrelation ist gegen beides invariant und ist deshalb anderen Vergleichsmassen, die diese Eigenschaft nicht haben, vorzuziehen.

3. Das wiederholte Berechnen von Korrelationen ist zeitaufwendig. Zur Beschleunigung des Verfahrens eignet sich ein Aehnlichkeitstest, der nicht-kompatible Bereiche anhand von einfachen statistischen Eigenschaften erkennt und von der Korrelation ausschliesst.

Das Korrespondenzverfahren von Hannah wird in THOMPSON 75 weiter verbessert. Thompson beruecksichtigt bei der Korrelationsberechnung, dass eine 3D-Oberflaeche in den zwei Ansichten auf unterschiedliche Weise perspektivisch verzerrt wird, um so mehr, je staerker die Blickwinkel differieren. Er vergleicht den Korrelationswert zwischen zwei Fenstern mit der "Autokorrelation" eines der Fenster, die durch Verschieben um 1 Pixel in alle Richtungen berechnet wird. Dadurch wird beruecksichtigt, dass die Korrelation strukturierter Bildteile auch bei perfekter Korrespondenz haeufig geringer ist als die von gleichfoermigen Bildteilen.

In YAKIMOVSKY und CUNNINGHAM 78 wird ein Stereokamerasystem beschrieben, das fuer die automatische Steuerung eines Roboterfahrzeuges entwickelt wurde. Das System wertet die Bilder zweier hochlinearer CID-Kameras (188*244 Pixel) direkt aus und kann die 3D-Koordinaten eines 2 Meter entfernten Punktes auf 5 mm genau berechnen. Die Kalibrierung der Kameras erfolgt mithilfe eines Roboterarms, der Raumpunkte mit bekannten 3D-Koordinaten vorgeben kann. Durch Verknuepfen der Abbildungsgleichungen von vier Punkten ergibt sich eine Beziehung fuer drei der acht Kameraparameter, deren Werte dann durch iterative Optimierung auf der Basis von mehreren Punktquadrupeln bestimmt werden.

Korrespondierende Punkte werden aehnlich wie in THOMPSON 75 mithilfe von Korrelationstests ermittelt. Um Echtzeitanforderungen besser nachkommen zu koennen, wird das adaptive Suchverfahren weiter ausgebaut. Bis zu sieben Fenstergroessen mit zunehmender Diskriminierungsfahigkeit koennen selektiert werden, damit fuer einen Vergleich nicht mehr Rechenaufwand als erforderlich eingesetzt werden muss.

Der fuer die Korrelationsoperationen benoetigte Rechenaufwand stellt einen entscheidenden Engpass fuer praktische Anwendungen dar, sei es in der Photogrammetrie oder der Robotik. Einen moeglichen Ausweg bietet die Entwicklung spezieller Hardware [GEMMAR 79]. Eine andere Moeglichkeit ist die drastische Einschraenkung der Vergleichsoperationen auf wenige markante Punkte. Dabei wird natuerlich ein zusaetzlicher Verfahrensschritt erforderlich, in dem geeignete Punkte lokalisiert werden. Zahlreiche Punktefinder sind vorgeschlagen worden, auch im Hinblick auf andere Anwendungen. DRESCHLER 81 enthaelt eine ausgezeichnete Uebersicht ueber diese Verfahren und auch interessante Ergebnisse mit einem von ihr untersuchten Operator, der auf starke Kruemmungen im Intensitaetsverlauf eines Bildes anspricht.

Ein weiteres grundsatzliches Problem stellt die Fehlerempfindlichkeit der bisher betrachteten Verfahren dar. Ist die Stereobasis klein im Verhaeltnis zur Entfernung eines Punktes, so fuehren kleine Ungenauigkeiten (die allein durch die diskrete Rasterung bedingt sein koennen) bereits zu erheblichen Fehlern in den berechneten Raumkoordinaten. Eine grosse Stereobasis dagegen erschwert eine genaue Korrespondenzbestimmung wegen der starken Disparitaet der Ansichten.

Wie in vielen anderen Situationen kann mangelnde Qualitaet durch Quantitaet teilweise kompensiert werden, indem die Resultate mehrerer Messungen vereinigt werden. Dies wird in NEVATIA 76 fuer Bildfolgen untersucht, die mithilfe einer Drehscheibe vor einer festen Kamera erzeugt werden. Wesentlich ist dabei, dass Kameraposition und -orientierung fuer jede Ansicht bekannt sind. Nevatia nennt sein Verfahren "Bewegungsstereo" - dieser Begriff sollte jedoch fuer die bewegungsbasierenden Stereoverfahren vorbehalten bleiben, die sich konzeptionell von Binokularstereo unterscheiden (siehe Kapitel 3). Nevatia nutzt das Vorhandensein von vielen aehnlichen Bildpaaren vorwiegend fuer die Korrespondenzbestimmung aus. Nur solche Punkte werden akzeptiert, deren Korrespondenz waehrend der gesamten Bildfolge verifiziert werden kann.

Das Stereokamerasystem in MORAVEC 80 [siehe dazu auch GENNERY 77, MORAVEC 77+79] wertet zur Verbesserung der Genauigkeit neun Ansichten aus, die durch kontrollierte Verschiebung einer einzigen Kamera erzeugt werden. Daraus lassen sich 36 Bildpaare kombinieren, aus denen Raumkoordinaten fuer jeweils dieselben Raumpunkte berechnet werden. Aus den Einzelergebnissen werden dann sehr genaue Endergebnisse unter Beruecksichtigung der individuellen Zuverlaessigkeiten abgeleitet.

Marr und Mitarbeiter [MARR und POGGIO 79, GRIMSON und MARR 79] stellen in ihrem theoretischen Modell der menschlichen Stereowahrnehmung ein Verfahrenskonzept vor, das sich in mancher Hinsicht von den bisher beschriebenen Verfahren unterscheidet. Sie lokalisieren in den Bildern Kantenelemente mithilfe von richtungsabhaengigen Filtern verschiedener Durchlasscharakteristik. Niedrige Frequenzanteile koennen zur Grobbestimmung und hoehere zur Feinbestimmung der jeweiligen Kantenpositionen herangezogen werden. Zusammen mit der Richtungsinformation wird auf diese Weise ein effektiver und genauer Korrespondenzprozess ermoeeglicht. Inwieweit dieses Verfahren auch fuer technische Anwendungen geeignet ist, laesst sich anhand der experimentellen Untersuchungen, die erst im Anfangsstadium sind, nicht absehen.

3. Bewegungsstereo

Bewegung kann auf natuerliche Weise dazu fuehren, dass sich Objekte in mehreren verschiedenen Ansichten praesentieren. Der entscheidende Unterschied zu Binokularstereo besteht jedoch nicht in der groesseren Zahl von Ansichten sondern in den - im allgemeinen - unbekanntem Bewegungsparametern, die fuer die Unterschiede zwischen den Ansichten verantwortlich sind. Bewegungsstereo-Verfahren muessen deshalb in der Regel sowohl eine Bewegungsanalyse der Bildfolge als auch eine raeumliche Analyse einzelner Ansichten leisten. Dies ist nur dann moeglich, wenn ueber die Beziehungen von Punkten untereinander zusaetzliche Annahmen gemacht werden koennen. Die am haeufigsten zugrundegelegte Annahme ist die Starrheitsbedingung. Man geht also davon aus, dass sich mehrere Punkte als starrer Verband bewegen, z.B. wenn sie Fixpunkte an der Oberflaeche eines starren Koerpers sind.

Als eine fuer die hier dargestellten Verfahren prototypische Situation kann eine Szene angesehen werden, in der mehrere starre Objekte, die sich unabhaengig voneinander bewegen, von einer festen Kamera beobachtet werden. Aus den bisherigen Untersuchungen zeichnet sich ab, dass folgende Teilprobleme geloest werden muessen,

um 3D-Information ueber die bewegten Objekte und ihre Trajektorien gewinnen zu koennen:

- (i) Extraktion von markanten Bildelementen
- (ii) Korrespondenz von Bildelementen in aufeinanderfolgenden Ansichten
- (iii) Gruppieren der Bildelemente nach ihrer Zugehoerigkeit zu unabhangig bewegten Objekten
- (iv) 3D-Analyse der Bildelemente eines einzelnen Objektes

Die ersten beiden Teilprobleme haben viel mit den entsprechenden Schritten bei Binokularstereo gemeinsam, und in der Tat gibt es Verfahren [BARNARD und THOMPSON 80], die fuer beide Anwendungen gleichermaßen konzipiert wurden. Als "Bildelemente" werden meist Merkmale berechnet, die einem koerperfesten Punkt in der Szene entsprechen sollen. Alternativ koennen auch Kantenelemente verwendet werden, die geometrisch als Gerade fungieren. Dies fuehrt zu voellig anderen mathematischen Beziehungen, die bisher nur in NEUMANN 79 untersucht worden sind. Im folgenden wird stets nur von Punkten die Rede sein.

Eine Reihe von Arbeiten geht speziell auf das Korrespondenzproblem bei Szenen mit bewegten Objekten ein. ULLMAN 79 untersucht in zahlreichen Experimenten, wie der Mensch das Korrespondenzproblem loest. Seine Theorie der minimalen Zuordnung (minimal mapping) ist jedoch noch nicht als Rechnerverfahren erprobt worden. Kraasch und Mitarbeiter [KRAASCH et al. 79a+79b] beschreiben ein interessantes Vergleichsverfahren auf der Basis von relational repraesentierten Segmentationsergebnissen. Ein aehnliches Verfahren wird von JACOBUS et al. 80 vorgeschlagen. DRESCHLER 81 benutzt zur Korrespondenzermittlung eine verbesserte Version des Relaxationsverfahrens von BARNARD und THOMPSON 80. In Dreschlers Arbeit findet sich auch ein kritischer Vergleich der wichtigsten Methoden.

Das Gruppierungsproblem - Teilproblem (iii) - ist bisher erst in wenigen Arbeiten behandelt worden. DRESCHLER und NAGEL 81 gehen davon aus, dass Objektmasken durch ein Differenzenverfahren [JAIN und NAGEL 79] gewonnen werden und dadurch zusammengehoeerige Punkte identifiziert werden koennen. Interessant ist der Vorschlag von O'ROURKE 81, gemeinsam bewegte Punktgruppen durch eine Art Hough-Transformation zu ermitteln. Dazu muesste ein 6-dimensionales Zaehlerfeld eingerichtet werden, in dem die moeglichen Bewegungsparameter eines Punktes registriert werden. Zaehlermaxima zeigen moegliche Gruppierungen an. In NEUMANN 80b wird das Gruppierungsproblem gleichzeitig mit der 3D-Analyse geloest. Hierueber wird weiter unten berichtet.

Der Hauptteil dieses Kapitels befasst sich mit Teilproblem (iv). Hier geht es beispielsweise um eine Situation, wie sie Fig.2 zeigt. Ein Objekt, repraesentiert durch koerperfeste Punkte, ist in mehreren Ansichten zu sehen, die Korrespondenz der Punkte sei bekannt. Inwieweit kann man von den 2D-Bildkoordinaten der Punkte auf ihre 3D-Raumkoordinaten schliessen?

Figur 2: Drei Ansichten von Eckpunkten eines Quaders

ULLMAN 79 hat fuer diese Frage eine erste Antwort in Gestalt seines "Structure from Motion Theorem" vorgelegt. Das Theorem besagt, dass man raeumliche Struktur und Trajektorie von vier nicht-koplanaren Punkten eindeutig (bis auf eine Reflektion bezueglich der Bildebene) aus drei orthographischen Projektionen berechnen kann. Unter "raeumlicher Struktur" sind die relativen raeumlichen Abstaende der Punkte zu verstehen. Absolute Werte, insbesondere der Abstand zum Beobachter, lassen sich nicht angeben, weil sich eine Bewegung in die Tiefe (vom Beobachter weg) bei orthographischer Projektion nicht bemerkbar macht. Ullman gibt einen konstruktiven Beweis an, der gleichzeitig eine Berechnungsmoeglichkeit aufzeigt. Experimentelle Ergebnisse sind jedoch bis heute nicht bekannt geworden.

Ein aehnlicher Ansatz, der auch auf orthographischen Projektionen basiert, findet sich in ASADA et al. 80. Hier wird gezeigt, dass drei Ansichten von lediglich drei Punkten erforderlich sind, um ihre raeumliche Struktur und Bewegung ermitteln zu koennen. Natuerlich muessen auch hier absolute raeumliche Koordinaten unbekannt bleiben.

Der Berechnungsgang geht zunaechst von zwei Ansichten einer starren 3-Punkte-Konfiguration aus. Die beiden Ansichten lassen sich durch eine Translation in der Bildebene sowie eine Rotation um einen beliebig waelzbaren Drehpunkt ineinander ueberfuehren. Als Drehpunkt wird einer der drei Punkte gewaehlt, der Verschiebungsvektor dieses Punktes gibt damit trivialerweise die gesuchte Translation an. Eliminiert man den translatorischen Bewegungsanteil und legt den Ursprung eines Koordinatensystems auf den Drehpunkt, so entsteht die in Fig. 3 illustrierte Situation. Die xy-Ebene ist parallel zur Bildebene gewaehlt. Damit sind die x- und y-Koordinaten bekannt, waehrend die z-Koordinaten zu bestimmen sind. Aus der Starrheitsbedingung folgt, dass Laengen und Winkel erhalten bleiben. Somit gilt

$$x_{11} + y_{11} + z_{11} = x_{21} + y_{21} + z_{21}$$

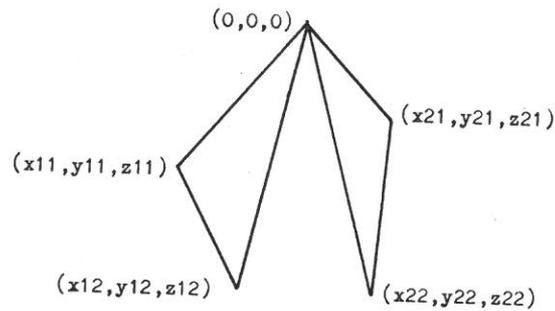
$$x_{12} + y_{12} + z_{12} = x_{22} + y_{22} + z_{22}$$

$$x_{11} x_{12} + y_{11} y_{12} + z_{11} z_{12} = x_{21} x_{22} + y_{21} y_{22} + z_{21} z_{22}$$

Daraus lässt sich eine quadratische, nichtlineare Gleichung für z_{11} und z_{12} ableiten.

$$A z_{11} - 2B z_{11} z_{12} + C z_{12} = AC - B$$

A, B und C sind Ausdrücke, die nur aus den beobachteten x- und y-Koordinaten zusammengesetzt sind. Eine analoge Gleichung kann für die Größen der zweiten und dritten Ansicht aufgestellt werden. Dadurch sind insgesamt zwei quadratische Ortslinien für z_{11} und z_{12} gegeben, deren Schnittpunkte die gesuchte Lösung und ihre Reflektion darstellen. Hieraus können dann leicht alle anderen unbekanntes z-Koordinaten sowie die Rotationsparameter berechnet werden.



Figur 3: Rotation von 2 Punkten um den Ursprung

Um den Einfluss fehlerhafter Ausgangsdaten zu minimieren, schlagen Asada und Mitarbeiter vor, durch Hinzunahme eines vierten Punktes ein überbestimmtes Gleichungssystem zu erzeugen. Aus dem Fehler, mit dem dieses Gleichungssystem bestenfalls befriedigt werden kann, lässt sich dann ein Zuverlässigkeitswert für die betrachtete Punkt Kombination ableiten. Durch die Auswahl der zuverlässigsten Punkt Kombinationen aus einer möglicherweise größeren Zahl von Punkten kann insgesamt das Ergebnis verbessert werden.

Bei der Wahl eines der Punkte als Koordinatenursprung wird dessen z-Koordinate willkürlich auf Null gesetzt. Dieser durch die orthographische Projektion bedingte Freiheitsgrad kann genutzt werden, um unter den möglichen Trajektorien eine besonders "natuerliche" zu selektieren. In einer Weiterentwicklung ihrer Arbeit [ASADA et al. 81] weisen die Autoren nach, dass bei Bewegungen mit konstant geneigter Rotationsachse z-Verschiebungen derart angenommen werden können, dass sich die Rotationsachse mit konstanter Geschwindigkeit bewegt. Hier wird also ein zusätzliches Kriterium eingeführt, um zwischen mehrdeutigen 3D-Interpretationen zu entscheiden.

NEUMANN 80 wählt für dieselbe Situation - bewegte Punkte in orthographischer Projektion - ein völlig anders Vorgehen. Er formuliert Kompatibilitätestests, die auf zwei Ansichten eines einzelnen Punktes angewandt werden können und zeigen, ob hypothetische Annahmen über dessen Rotation, Translation und räumliche Position mit den Beobachtungen verträglich sind. Dadurch wird es möglich,

Gruppierungsentscheidungen, also Zuordnungen von Punkten zu bestimmten Objekten, erst waehrend der 3D-Analyse auf der Basis von 3D-Interpretierbarkeit zu faellen, im Gegensatz zu 2D-Heuristiken, wie sie z.B. von THOMPSON 80 vorgeschlagen werden. Fuer das punktweise Vorgehen muss jedoch mit erhoehetem Rechenaufwand bezahlt werden, denn statt der formelmaessigen Berechnung der Bewegungsparameter findet hier eine Suche statt. Eine weitere positive Eigenschaft dieses Ansatzes liegt in der expliziten Beruecksichtigung einer Fehlertoleranz mithilfe von Intervallrechnung. Dadurch kann der Unsicherheitsbereich der Resultate genau angegeben werden.

Ueber eine Anwendung der bisher beschriebenen Verfahren auf Realweltszenen liegen keine Angaben vor, auch der Autor hat bei den genannten Untersuchungen nur simulierte Daten benutzt. Dadurch laesst es sich nur schwer abschaezen, welchen Einfluss die Annahme orthographischer Projektion, die diesen Verfahren zugrunde liegt, auf die Qualitaet der Ergebnisse hat. Es ist allgemein bekannt, dass der Abbildungsvorgang mit einer Kamera genauer durch perspektivische Projektion beschrieben wird (siehe z.B. DUDA und HART 73). Die Annahme orthographischer Projektion fuehrt jedoch haeufig zu einfacheren mathematischen Beziehungen (dies wird weiter unten noch demonstriert) und wird deshalb gelegentlich bevorzugt.

Vielleicht ist es bezeichnend, dass die erste Bewegungsstereo-Analyse einer Realweltszene mit einem Verfahren erfolgte, dem perspektivische Projektion zugrunde liegt [BONDE 79, BONDE und NAGEL 79, NAGEL 81]. Es handelt sich dabei um eine Strassenverkehrsszene, die mit einer Fernsehkamera von schraeg oben aus ca. 25 Metern Entfernung aufgenommen wurde. In der Szene vollfuehrt ein abbiegendes Auto eine Translation von ca. 10 Metern, gekoppelt mit einer Rotation um seine Hochachse von ca. 60 Grad. Die Bildfolge besteht aus ca. 30 Ansichten. Vor der eigentlichen 3D-Analyse wird zunaechst ein Kameramodell berechnet. Dies geschieht hier durch iterative Anpassung der Kameraparameter (Position, Orientierung, Brennweite, u.a.) mithilfe von Punkten, deren raeumliche Koordinaten bekannt sind. Das 3D-Analyseverfahren benutzt die inverse perspektivische Transformation (DUDA und HART 73, HARALICK 80) unter der einschraenkenden Annahme, dass eine ebene Bewegung vorliegt, d.h. dass eine Rotation nur um die Hochachse und eine Translation nur in einer Ebene senkrecht dazu erfolgt. Damit verbleiben noch 3 der 6 moeglichen Bewegungsfreiheitsgrade. Trotz dieser Vereinfachung lassen sich fuer die unbekanntenen Bewegungsparameter und Raumkoordinaten keine geschlossenen Loesungsformeln angeben. Aehnlich wie bei der Kalibrierung wird deshalb in einem iterativen Verfahren eine Fehlerfunktion minimiert, in der die Fehlerquadrate von gemessenen Bildkoordinaten gegen die aus den geschaezten Parametern errechneten Koordinaten aufsummiert werden.

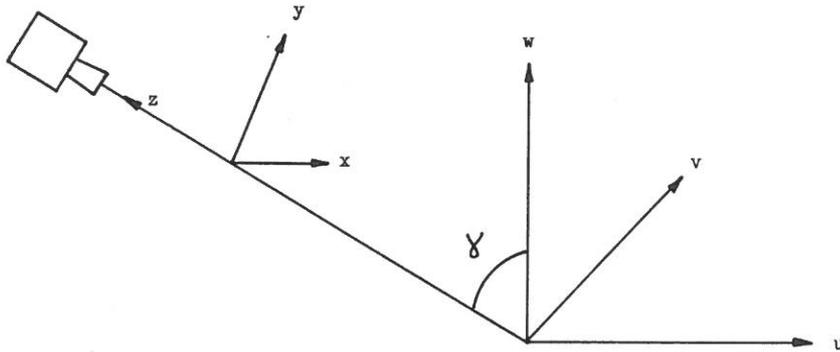
Es ist interessant, dass fuer dieselbe Situation eine geschlossene Loesung existiert, wenn man orthographische Projektion zugrunde legt. Dem Autor ist keine Veroeffentlichung hierueber bekannt, deshalb sei die Ableitung kurz skizziert. Fig. 4 zeigt die verwendeten Koordinatensysteme. Zwei starr verbundene Punkte p und q rotieren um die w -Achse des Weltkoordinatensystems bei gleichzeitiger beliebiger Translation. Ihre Positionen zu einem Zeitpunkt 2 haengen wie folgt mit den Positionen zu einem Zeitpunkt 1 zusammen.

$$p_2 = R p_1 + t$$

$$\underline{q}_2 = R \underline{q}_1 + \underline{t}$$

$$\text{mit } R = \begin{vmatrix} \cos r & \sin r & 0 \\ -\sin r & \cos r & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad \text{und} \quad \underline{t} = \begin{vmatrix} t_u \\ t_v \\ t_w \end{vmatrix}$$

Die Bildkoordinaten der Punkte werden im Kamerakoordinatensystem in der xy -Ebene gemessen. Die z -Achse faellt mit der optischen Achse der Kamera zusammen. Das Weltkoordinatensystem sei so gewaehlt, dass die z -Achse in der wv -Ebene liegt und durch den Ursprung geht. Somit verbleibt als einziger Parameter, der die Lage des Kamerakoordinatensystems gegenueber dem Weltkoordinatensystem spezifiziert, der Winkel γ zwischen z - und w -Achse. Der Abstand der Kamera entlang der z -Achse ist wegen orthographischer Projektion irrelevant.



Figur 4: Weltkoordinaten uvw und Kamerakoordinaten xyz

Ein Punkt (u,v,w) wird mit den einfachen Beziehungen

$$x = u$$

$$y = v \cos \gamma + w \sin \gamma$$

in die Bildebene transformiert. Fuer zwei Ansichten des Punktes p gilt also

$$px_1 = pu_1$$

$$py_1 = pv_1 \cos \gamma + pw_1 \sin \gamma$$

$$px_2 = pu_1 \cos r + pv_1 \sin r + t_u$$

$$py_2 = \cos \gamma (-pu_1 \sin r + pv_1 \cos r + t_v) + \sin \gamma (pw_1 + t_w)$$

Entsprechende Beziehungen gelten fuer q . In mehreren Eliminationsschritten koennen aus diesen Gleichungen alle Unbekannten bis auf r und γ beseitigt werden. Substituiert man schliesslich

$$\tan(r/2) = (1 - \cos r) / \sin r$$

so folgt die huebsche Beziehung

$$\tan(r/2) \cos \gamma = - \frac{(py_2 - py_1) - (qy_2 - qy_1)}{(px_2 + px_1) - (qx_2 + qx_1)}$$

Ist der Winkel γ durch Kalibrierung bekannt, ergibt sich der Drehwinkel r aus zwei Ansichten von zwei Punkten. Aehnliche Beziehungen lassen sich fuer die Differenzen der Weltkoordinaten $pu_1 - qu_1$, $pv_1 - qv_1$ und $pw_1 - qw_1$ ableiten. Hieraus folgt der raeumliche Abstand der beiden Punkte.

Mit der Einfachheit von obiger Beziehung verbindet sich gleichzeitig auch Analysierbarkeit. Z.B. kann der Einfluss ungenauer Ausgangsdaten auf den berechneten Drehwinkel r leicht untersucht werden. Die Ergebnisse sind natuerlich nur insofern realistisch, als orthographische Projektion den Grenzfall von perspektivischer Projektion fuer grosse Entfernungen darstellt. Dennoch koennen weitere Untersuchungen mit dem orthographischen Modell durchaus nuetzlich sein.

Fuer Bewegungsstereo bei perspektivischer Projektion koennen im allgemeinen keine expliziten Loesungsformeln angegeben werden, so dass iterative Verfahren herangezogen werden muessen. ROACH und AGGARWAL 80 wenden ein Minimierungsverfahren direkt auf die inversen perspektivischen Transformationsbeziehungen an [DUDA und HART 73]. Sie betrachten zunaechst 5 Punkte in 2 Ansichten. Dies fuehrt zu 20 Gleichungen in 27 Unbekannten. Durch willkuerliches Festlegen eines Skalierungsfaktors (dieser laesst sich durch Bewegungsstereo niemals ermitteln) und durch Wahl von geeigneten Koordinatenkonventionen ergibt sich daraus ein System von 18 Gleichungen in 18 Unbekannten. Experimente mit simulierten Daten zeigen, dass sich in diesem hochdimensionalen Parameterraum tatsaechlich die korrekte Loesung finden laesst, wenn man die Anfangsschaetzwerte sorgfaeltig bestimmt (Roach und Aggarwal geben hierzu eine Verfahren an) und ein geeignetes Gradientenabstiegsverfahren benutzt. Die Ergebnisse sind jedoch sehr fehlerempfindlich, wenn sie nur auf 5 Punkten basieren. Erst wenn 12 oder mehr Punkte in 2 Ansichten ausgewertet werden, ergibt sich genuegend Ueberbestimmtheit, um trotz ungenauer Ausgangsdaten zu sinnvollen Ergebnissen zu kommen.

In bestimmten Situationen, wenn die Punkte in reiner Drehung um das optische Zentrum begriffen sind, konnte von Roach und Aggarwal keine Loesung gefunden werden. Es ist bis heute unklar, welche Bedingungen insgesamt erfuellt sein muessen, damit eine Loesung garantiert ist. MEIRI 80 zeigt auf, wieviel Punkte P in wieviel Ansichten A betrachtet werden muessen, um mindestens ebensoviele Gleichungen wie Unbekannte zu haben. Die Ungleichung lautet

$$P \geq 3 + 2 / (2A - 3)$$

Man kann daraus entnehmen, dass fuer 2 Ansichten in der Tat mindestens 5 Punkte und bei mehr Ansichten mindestens 4 erforderlich sind.

NAGEL 81 leitet eine kompakte mathematische Beziehung fuer die 3D-Analyse von 2 Ansichten ab, fuer die in NAGEL und NEUMANN 81 eine einfache geometrische Interpretation angegeben wird. Gegeben sei ein Koordinatensystem mit Ursprung im optischen Zentrum der Kamera. Die Bewegung der Punktkonfiguration kann in eine Translation \underline{t} , gefolgt von einer Rotation R um den Ursprung zerlegt werden. Sei \underline{e}_m der Einheitsvektor fuer den Abbildungsstrahl, auf dem der m -te Punkt in der n -ten Ansicht liegt. Dreht man den Abbildungsstrahl \underline{e}_m um R zurueck in die Richtung $\underline{e}_{m2R'}$ (der Apostroph bedeutet Transposition), so spannen \underline{e}_m und $\underline{e}_{m2R'}$ eine Ebene auf, die \underline{t} enthalten muss. Das Vektorprodukt

$$(\underline{e}_m \times \underline{e}_{m2R'}) \quad m = 1, 2, \dots$$

definiert den Normalenvektor dieser Ebene. Die Ebenen fuer $m=1$ und $m=2$ schneiden sich in einer Geraden, deren Richtung durch

$$(\underline{e}_{11} \times \underline{e}_{12R'}) \times (\underline{e}_{21} \times \underline{e}_{22R'})$$

gegeben ist und mit der Richtung von \underline{t} uebereinstimmt. Alle weiteren Punkte muessen damit kompatibel sein, d.h. die Normalenvektoren ihrer Ebenen muessen senkrecht auf der Richtung von \underline{t} stehen, die durch die ersten zwei Punkte definiert wird. Dies kann mit dem Skalarprodukt ausgedrueckt werden.

$$[(\underline{e}_{11} \times \underline{e}_{12R'}) \times (\underline{e}_{21} \times \underline{e}_{22R'})] (\underline{e}_m \times \underline{e}_{m2R'}) = 0$$

Die obige Gleichung ist eine nichtlineare Beziehung zwischen den 3 unbekanntem Parametern, die die Rotationsmatrix R spezifizieren. Bei 5 Punkten ergeben sich 3 solcher Gleichungen ($m=3,4,5$), so dass sich Loesungen fuer R bestimmen lassen. Daraus koennen dann ueber einfache Beziehungen der Translationsvektor \underline{t} und die raeumlichen Punktkoordinaten ermittelt werden. Der 18-dimensionale Parameterraum, in dem ROACH und AGGARWAL 80 nach einer Loesung suchen, laesst sich also auf 3 Dimensionen reduzieren. NAGEL und NEUMANN 81 leiten aus Gleichung (4) einige Bedingungen fuer die Existenz von Loesungen ab, u.a. auch die Notwendigkeit einer translatorischen Komponente (s.o.).

Zum Abschluss dieses Kapitels soll auf Arbeiten hingewiesen werden, die sich auf nicht-starre Bewegung beziehen. Ein einfacher Fall von nicht-starrer Bewegung liegt vor, wenn mehrere starre, aber beweglich zusammenhaengende Objektteile individuelle Bewegungen ausfuehren. Die Situation unterscheidet sich prinzipiell nicht von einer, in der sich starre Objekte unabhaengig bewegen, soweit man nur an raeumlichen Koordinaten interessiert ist. Das Gruppierungs- und 3D-Analyseverfahren von NEUMANN 80 koennte also auch hier angewendet werden. ASADA et al. 80 versuchen zusaetzlich, eine hierarchische Bewegungsinterpretation zu geben. Sie nutzen dabei die Annahme aus, dass die Hauptkoerper eine Rotation mit konstant geneigter Drehachse ausfuehren und dadurch identifiziert werden koennen. Die Bewegungen der verbleibenden Objekte werden dann relativ zu den Hauptbewegungen untersucht und auf aehnliche Weise hierarchisch dekomponiert.

HOFFMAN und FLINCHBAUGH 80 untersuchen die Moeglichkeit, "biologische Bewegung" raeumlich zu interpretieren. Darunter sind Geh- und Laufbewegungen von Zwei- oder Vierbeinern zu verstehen, (siehe dazu auch RASHID 80). Hoffman und Flinchbaugh

gehen wie andere davon aus, dass eine Bildfolge mit bewegten Punkten vorliegt, deren Korrespondenz bekannt ist. Sie schlagen vor, alle Punkte, die sich nicht als starrer Verband interpretieren lassen, in starr verbundene Paare (Glieder) zu zerlegen. Glieder koennen untereinander beweglich zusammenhaengen, indem sie einen Punkt gemeinsam haben. Nimmt man an, dass ein Gliederpaar nur eine ebene Bewegung (in der Ebene der zwei Glieder) ausfuehren kann, so laesst sich eine 3D-Interpretation aus 2 orthographischen Ansichten ableiten. Fuer ein einzelnes Glied sind 3 Ansichten erforderlich, wenn man auch hier eine ebene Bewegung annimmt.

4. Ausblick

Es wurde ueber Verfahren berichtet, die raeumliche Information aus mehreren Ansichten einer Szene ableiten, ohne spezifisches Vorwissen ueber den Szeneninhalt auszunutzen. Eine Methode, Binokularstereo, wird schon seit laengerer Zeit erforscht und auch schon ebenso lange in der Photogrammetrie angewendet, konkrete Anwendungen in der Robotik stehen noch aus. Der hohe Rechenaufwand, der fuer eine zuverlaessige Korrespondenzermittlung erforderlich ist, stellt nach wie vor ein Problem dar. Spezielle Hardware scheint erforderlich, um Echtzeitanforderungen befriedigen zu koennen.

Bewegungsstereo steckt noch in den Kinderschuhen. Zusaetzlich zum Korrespondenzproblem sind mindestens zwei weitere nichttriviale Probleme zu loesen: das Gruppieren von Bildelementen zu Objektkandidaten und das Bestimmen von raeumlichen Trajektorien mithilfe eines komplexen nichtlinearen Gleichungssystems. Erste Loesungsansaeetze wurden vorgestellt.

Naheliegende Anwendungsgebiete der hier diskutierten Verfahren sind Photogrammetrie und Robotik. Hier sind vielerlei Situationen denkbar, in denen raeumliche Entfernungen festgestellt oder unbekannte Objekte vermessen werden muessen. Aber auch bei modellgesteuerter Bildinterpretation, wo Vorwissen ueber eine Szene und die darin zu erwartenden Objekte zur Verfuegung steht, kann eine raeumliche Analyse mit Binokularstereo oder Bewegungsstereo ihren Platz haben, da auf diese Weise 3D-Merkmale gewonnen werden koennen, die eine Objektidentifikation erleichtern [NEUMANN und RADIG 79]. Derartige Systeme werden jedoch nicht in naerer Zukunft zur Anwendung kommen.

Literatur

- Asada et al. 80
 Three Dimensional Motion Interpretation for the Sequence
 of Line Drawings
 M. Asada, M. Yachida, S. Tsuji
 ICPR-80 (1980) 1266-1273
- Asada et al. 81
 Understanding of Three-Dimensional Motion in Blocks World
 M. Asada, M. Yachida, S. Tsuji
 (eingereicht zur Veroeffentlichung in IEEE-PAMI)
- Bajcsy 80

- Three-Dimensional Scene Analysis
 R. Bajcsy
 ICPR-80 (1980) 1064-1074
- Barnard and Thompson 80
 Disparity Analysis of Images
 S.T. Barnard and W.B. Thompson
 IEEE-PAMI-2 (1980) 333-340
- Bonde 79
 Untersuchungen zur dreidimensionalen Modellierung
 bewegter Objekte durch Analyse von Formveraenderungen
 der Objektbilder in TV-Aufnahmefolgen
 T. Bonde
 Diplomarbeit (Januar 1979)
 Fachbereich Informatik, Universitaet Hamburg
- Bonde and Nagel 79
 Deriving a 3-D Description of a Moving Rigid
 Object from Monocular TV-Frame Sequences
 T. Bonde and H.-H. Nagel
 WCATVI-79, pp. 44-45
- Clocksinn 78
 Determining the Orientation of Surfaces from Optical
 Flow
 W.F. Clocksinn
 Proc. AISB/GI-78 on Artificial Intelligence
 Hamburg, July 18-20, 1978, pp. 93-102
- Clocksinn 80
 The Effect of Motion Contrast on Surface Slant
 and Edge Detection
 W.F. Clocksinn
 Proc. AISB-80 Conference on Artificial Intelligence
 St. Hardy (ed.), Amsterdam, July 1-4, 1980
- Dreschler 81
 Ermittlung markanter Punkte auf den Bildern bewegter Objekte
 und Berechnung einer 3D-Beschreibung auf dieser Grundlage
 L. Dreschler
 Dissertation (1981)
 Fachbereich Informatik, Universitaet Hamburg
- Dreschler und Nagel 81
 Volumetric Model and 3D-Trajectory of a Moving Car
 Derived from Monocular TV-Frame Sequence of a Street Scene
 L. Dreschler, H.-H. Nagel
 IJCAI-81 (1981)
- Duda and Hart 73
 Pattern Classification and Scene Analysis
 R.O. Duda and P.E. Hart
 John Wiley _Sons, New York, 1973
- Gemmar 79
 Ein digitales on-line Stereokorrelationssystem
 P. Gemmar
 Proc. DAGM Symposium 1979

- Informatik Fachberichte pp. 315-321
Springer Verlag, Berlin-Heidelberg-New York, 1979
- Gennery 77
A Stereo Vision System for an Autonomous Vehicle
D.B. Gennery
IJCAI-77, pp. 576-582
- Gibson 50
The Perception of the Visual World
J.J. Gibson
Houghton Mifflin Co., Boston/MA 1950
reprint by Greenwood Press, Westport/CT, 1974
- Grimson and Marr 79
A Computer Implementation of a Theory of Human
Stereo Vision
W.E.L. Grimson and D. Marr
Proc. Image Understanding Workshop, pp. 41-47
L.S. Baumann (ed.), Palo Alto/CA, April 24-25, 1979
Science Applications, Inc., Arlington/VA 22209
- Hannah 74
Computer Matching of Areas in Stereo Images
M.J. Hannah
Ph.D. Thesis, Memo AIM 239 (July 1974)
Stanford University, Stanford/CA
- Haralick 80
Using Perspective Transformations in Scene Analysis
R.M. Haralick
Computer Graphics and Image Processing 30, pp. 191-221 (1980)
- Hoffman und Flinchbaugh
The Interpretation of Biological Motion
D.D. Hoffman, B.E. Flinchbaugh
MIT-AI-MEMO No. 608 (1980)
Massachusetts Institute of Technology, Cambridge
- Horn and Schunck 80
Determining Optical Flow
B.K.P. Horn and B.G. Schunck
AI Memo 572 (April 1980)
Artificial Intelligence Laboratory
Massachusetts Institute of Technology, Cambridge/MA
- Horn et al. 78
Determining Shape and Reflectance Using Multiple
Images
B.K.P. Horn, R.J. Woodham, and W.M. Silver
AI-Memo 490 (August 1978), MIT, Cambridge/MA
- Jacobus et al. 80
Motion Detection and Analysis of Matching Graphs of
Intermediate-Levels Primitives
C.J. Jacobus, R.T. Chien, and J.M. Selander
IEEE-PAMI-2 (1980) 495-510
- Jain and Nagel 79
On the Analysis of Accumulative Difference Pictures

- from Image Sequences of Real World Scenes
R. Jain and H.-H. Nagel
IEEE-PAMI-1 (1979) 206-214
- Kanade 79
Recovery of the Three-Dimensional Shape of an Object
from a Single View
T. Kanade
CMU-CS-79-153 (1979)
Carnegie-Mellon University, Pittsburgh, Pa.
- Koenderink und van Doorn 77
How an ambulant observer can construct a model of the
environment from the geometrical structure of the visual inflow
J.J. Koenderink, A.J. van Doorn
in Hauske and Butenandt (Hrsg.), Kybernetik, pp. 224-247
Oldenburg Verlag, Muenchen (1977)
- Kraasch et al. 79a
Automatic Three-Dimensional Description of
Simple Moving Objects
R. Kraasch, B. Radig, and W. Zach
WCATVI-79, pp. 42-43
- Kraasch et al. 79b
Automatische Dreidimensionale Beschreibung
bewegter Gegenstaende
R. Kraasch, B. Radig, W. Zach
in: Angewandte Szenenanalyse, J.P. Foith (ed.)
Informatik Fachberichte 20, pp. 208-215
Springer Verlag, Berlin-Heidelberg-New York 1979
- Marr and Poggio 79
A Computational Theory of Human Stereo Vision
D. Marr and T. Poggio
Proc. Royal Society of London B 204 (1979) 301-308
- Moravec 77
Towards Automatic Visual Obstacle Avoidance
H.P. Moravec
IJCAI-77, p. 584
- Moravec 79
Visual Mapping by a Robot Rover
H.P. Moravec
IJCAI-79, pp. 598-600
- Moravec 80
Obstacle Avoidance and Navigation in the Real World
by a Seeing Robot Rover
H.P. Moravec
Ph.D. Thesis, Department of Computer Science
STAN-CS-80-813, Stanford University
available, too, as CMU-RI-TR-3 (September 1980)
Robotics Institute, Carnegie-Mellon University
Pittsburgh/PA
- Nagel 81
On the Derivation of 3D Rigid Point Configurations

- from Image Sequences
H.-H. Nagel
Proc. PRIP-81, Austin, Texas (1980)
- Nagel und Neumann 81
On 3D Reconstruction from two Perspective Views
H.-H. Nagel, B. Neumann
Proc. IJCAI-81, Vancouver, Canada (1981)
- Neumann 79
Raemliche Analyse von Bildsequenzen mithilfe
korrespondierender Kanten
B. Neumann
in: Angewandte Szenenanalyse, J.P. Foith (ed.)
Informatik Fachberichte 20, pp. 216-221
Springer Verlag, Berlin-Heidelberg-New York 1979
- Neumann und Radig 79
Strategien zur Analyse von Szenen mit Bewegung
B. Neumann
4. Arbeitstagung der GI-Fachgruppe fuer Kuenstliche Intelligenz
Bad Honnef (1979), Proc. als Bericht des Inst. f. Informatik,
Universitaet Bonn
- Neumann 80a
Exploiting Image Formation Knowledge for Motion Analysis
B. Neumann
IEEE-PAMI-2 (1980) 550-554
- Neumann 80b
Motion Analysis of Image Sequences for
Object Grouping and Reconstruction
B. Neumann
ICPR-80 (1980) 1262-1265
- Nevatia 76
Depth Measurement by Motion Stereo
R. Nevatia
Computer Graphics and Image Processing 5 (1976) 203-214
- O'Rourke 81
Dynamically Quantized Spaces Applied to Motion Analysis
J. O'Rourke
JHU-EE 81-1 (1981), The Johns Hopkins University,
Baltimore, Maryland
- Prazdny 79
Motion and Structure from Optical Flow
K. Prazdny
IJCAI-79, pp. 702-704
- Prazdny 80
Egomotion and Relative Depth Map from Optical Flow
K. Prazdny
Biological Cybernetics 56 (1980) 87-102
- Rashid 80
Towards a System for the Interpretation of Moving Light Displays
R.F. Rashid
IEEE-PAMI-2 (1980) 574-581

- Roach und Aggarwal 80
 Determining the Movements of Objects from a Sequence of Images
 J.W. Roach, J.K. Aggarwal
 IEEE-PAMI-2 (1980) 554-562
- Rock 75
 An Introduction to Perception
 I. Rock
 MacMillan Publishing Company, N.Y. (1975)
- Sobel 74
 On Calibrating Computer Controlled Cameras for
 Perceiving 3-D Scenes
 I. Sobel
 Artificial Intelligence 5 (1974) 185-198
- Thompson 75
 Depth Perception in Stereo Computer Vision
 C. Thompson
 Memo AIM-268, Stanford Artificial Intelligence Laboratory
 Stanford University (1975)
- Ullman 79a
 The Interpretation of Visual Motion
 S. Ullman
 The MIT Press, Cambridge/Mass., 1979
- Vamos and Bathor 80
 3-D Complex Object Recognition Using Programmed Illumination
 T. Vamos, M. Bathor
 ICPR-80 (1980) 1091-1093
- Williams 80
 Depth from Camera Motion in a Real World Scene
 T.D. Williams
 IEEE-PAMI-2 (1980) 511-516
- Woodham 78
 Photometric Stereo: A Reflectance Map Technique
 for Determining Surface Orientation from Image Intensity
 R.J. Woodham
 Proc. SPIE 155 (August 1978)
- Woodham 80
 Photometric Method for Determining Surface Orientation
 from Multiple Images
 R.J. Woodham
 Optical Engineering 19 (1980) 139-144
- Yakimovsky and Cunningham 78
 A System for Extracting Three-Dimensional Measurements
 from a Stereo-Pair of TV-Cameras
 Y. Yakimovsky and R.T. Cunningham
 Computer Graphics and Image Processing 7 (1978) 195-210