

Untersuchung differentieller Ansätze
zur Schätzung des optischen Flusses
in Grauwert- und Farbbildfolgen

Rainer Sprengel

FBI-HH-B-137/88

17. Mai 1988

Fachbereich Informatik
Universität Hamburg
Bodenstedtstraße 16
D-2000 Hamburg 50

Zusammenfassung

Der optische Fluß stellt ein wichtiges Hilfsmittel zur Rekonstruktion der dreidimensionalen Geometrie einer Szene dar. Diese Arbeit enthält eine ausführliche Analyse der "motion constraint equation", die den meisten Verfahren zur Berechnung des optischen Flusses zugrundeliegt. Außerdem werden die Zusammenhänge von verschiedenen Verfahren zur Lösung des "Blendenproblems" dargestellt. Schließlich werden Möglichkeiten der Erweiterung auf Farbbildfolgen hergeleitet und die Ergebnisse mit entsprechenden Grauwertverfahren an Beispielszenen verglichen. Dabei wird auch die Aussagekraft von verschiedenen Größen untersucht, die die Zuverlässigkeit eines Verschiebungsvektors messen sollen.

Abstract

The optical flow field is an important aid in reconstructing the three-dimensional geometry of a scene. This paper contains a detailed analysis of the "motion constraint equation", which is the starting point of most techniques calculating the optical flow field. Moreover, the correlation between several methods to solve the "aperture problem" are explained. Finally, extensions of grey-value-algorithms for the application to color images are derived and the results are compared. Additionally, we examine the expressive power of some quantities designed to measure the reliability of each optical flow vector.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Definition des optischen Flusses	2
1.2	Berechnung von Verschiebungsvektoren	3
1.2.1	Merkmalszuordnungsverfahren	4
1.2.2	Gradientenbasierte Verfahren	5
1.3	Übersicht	6
2	Verwendung von Verschiebungsvektoren	7
2.1	Verwendung von einzelnen Punktkorrespondenzen	8
2.2	Verwendung von globalen Eigenschaften	11
2.3	Verwendung von Objektmodellen	11
2.4	Zusammenfassung	13
3	Der differentielle Ansatz	14
3.1	Die “motion constraint equation”	14
3.2	Diskussion der Voraussetzungen	16
3.3	Der diskrete Fall	19
3.3.1	Schätzung des Gradienten	20
3.3.2	Die Zeitabtastung	22
3.3.3	Zusammenfassung der Fehler	23
3.4	Herleitung als Optimierungsproblem	25
3.5	Betrachtungen im Frequenzraum	27
4	Schätzung der zweiten Komponente	31
4.1	Zusätzliches Wissen	31
4.2	Erweiterung der “motion constraint equation”	32
4.3	Annahme eines lokal konstanten Verschiebungsvektors	34
4.3.1	Erste Fehleranalyse	34
4.3.2	Least Square Fit	35
4.3.3	Approximative Lösung	37
4.3.4	Die Güte der Schätzung	39
4.3.5	Weitere Ansätze	40
4.4	Von der Konstanz zur Glattheit	41

4.4.1	Problemformulierung	42
4.4.2	Einfache Glattheitsmaße	44
4.4.3	Gerichtete Glattheitsmaße	48
4.5	Zusammenfassung	52
5	Verschiebungsvektoren aus Farbbildern	54
5.1	Definitionen	55
5.2	Von Grauwertbildern zu Farbbildern	55
5.2.1	Optimierung der "motion constraints"	57
5.2.2	Optimierung eines Farbabstandes	57
5.2.3	Interpretation im Farbraum	58
5.3	Lineare Farbtransformationen	60
5.4	Gütekriterien	62
5.5	Aufwandsabschätzung	63
6	Vergleich der Verfahren	65
6.1	Verwendete Bildfolgen	65
6.2	Reine Farbverfahren	66
6.3	Farbinformationen in Realweltbildern	69
6.4	Vergleich verschiedener Verfahren	71
6.5	Die Aussagekraft der Gütekriterien	76
6.6	Verschiebungsvektoren auf Realweltbildfolgen	80
7	Schlußbemerkungen	85
A	Zeichenerklärung	88
B	Quantitative Analyse der Beleuchtung	90
C	Die "motion constraint equation" bei Drehungen	92
D	Fehler der lokalen Konstanzannahme	94
E	Bilder	96
	Literaturverzeichnis	101

Kapitel 1

Einleitung

Ein Ziel bei der computergestützten Auswertung von Bildfolgen ist es, eine räumliche Beschreibung der abgebildeten Objekte zu erhalten.¹ Dazu gehört die Rekonstruktion der dreidimensionalen Form der Objekte, sowie die Bestimmung der Orte und der relativen Geschwindigkeiten zur Kamera. Ein Problem ist, daß bei der Abbildung der Objekte auf die Bildebene die Tiefeninformation nicht mehr explizit vorhanden ist.

Man kann aber aus *mehreren* zweidimensionalen Projektionen der Szene Schlußfolgerungen auf die dreidimensionale Struktur der Objekte ziehen. Bei dieser Art der geometrischen Tiefenbestimmung kann man die erforderliche Anzahl der Aufnahmen auf zwei verschiedene Arten gewinnen:

Stereoskopie: Im Falle zweier räumlich getrennter Aufnahmen zum gleichen Zeitpunkt, lassen sich aus der Kenntnis der Bildkoordinaten, die ein Objektpunkt in den zwei Ansichten der Szene hat, durch Triangulation dessen 3D-Koordinaten berechnen. Voraussetzung dafür ist allerdings die Kenntnis der Kameraparameter, dies sind u.a. die relative Lage der Kameras zueinander sowie deren Brennweiten.

Bewegungsstereo: Bewegt sich ein Objekt relativ zur Kamera, so kann man die zur Triangulation erforderlichen verschiedenen Ansichten des Objektes auch durch die Aufnahme von Bildern zu verschiedenen Zeitpunkten gewinnen. Die Kenntnis der Kameraparameter beim Stereosehen entspricht dann der Kenntnis der relativen Bewegung zwischen beobachtetem Objekt und Kamera, die allerdings in den meisten Fällen unbekannt ist. Es gibt nun die Möglichkeit, durch die Ermittlung von mehr als nur einem Punktepaar eines sich *starr* bewegenden Objektes die Bewegungsparameter mitzubestimmen.

Will man eine dieser Techniken der geometrischen Szenenrekonstruktion anwenden, so stellt sich zunächst das Problem, die Abbildungen eines Objektpunktes

¹Einen guten Einblick in das Gebiet der Bildfolgenauswertung, besonders der Bildverarbeitungsgruppe in Hamburg, geben die Übersichtsartikel *Nagel 83b*, *Nagel 85*, *Nagel 86*. In *Barron 84* finden sich Zusammenfassungen vieler wichtiger Artikel zu diesem Thema.

in den verschiedenen Bildern der Bildfolge einander zuzuordnen. Dieses Problem wird in der Literatur auch *Korrespondenzproblem* genannt. Bilder eines Objektpunktes in verschiedenen Abbildung heißen auch "(physikalisch) korrespondierende Punkte".

Hat man zu jedem Punkt eines Bildes den korrespondierenden Punkt im nächsten Bild gefunden, so erhält man durch Verbindung dieser Punkte ein Feld von Verschiebungsvektoren. Diese Arbeit beschäftigt sich mit der Ermittlung eines Vektorfeldes, welches "optischer Fluß" genannt wird, und unter bestimmten Annahmen zu dem oben genannten Verschiebungsfeld äquivalent ist. Zunächst soll der Begriff des optischen Flusses definiert, und damit der Unterschied zwischen diesen beiden Feldern näher erläutert werden. Dann werden wir uns kurz den verschiedenen Paradigmen zuwenden, die zur Lösung des Korrespondenzproblems vorgeschlagen worden sind.

1.1 Definition des optischen Flusses

Ein Blick in die Literatur kann zu einiger Verwirrung führen, denn es werden verschiedene, teilweise nicht übereinstimmende Definitionen des optischen Flusses angegeben. Hier seien einige Beispiele aufgeführt:

"Optical flow is the distribution of apparent velocities of movement of brightness patterns in an image." [Horn + Schunck 81]

"Optical flow is the velocity field in the image plane that arises due to the projection of moving points in the scene onto the image plane." [Schunck 84a]

"Optical flow is the field of apparent velocities of images of points on the surface of objects in space, due to relative motion of viewing system and objects." [Mitiche 84]

"Mit optischem Fluß werden die wahrgenommenen Bewegungen innerhalb der Bildebene oder bei visuellen Systemen auf der Retina bezeichnet." [Enkelmann 85]

"The velocity field that represents the motion of object points across an image is called the optical flow field." [Kearney u.a. 87]

Der optische Fluß ist also ein 2-dimensionales Vektorfeld von Geschwindigkeiten auf der Bildebene. Es scheint aber Uneinigkeit darüber zu bestehen, wie diese Geschwindigkeiten zu definieren sind. Man kann zumindest zwei Arten der Definition unterscheiden. Zum einen diejenige, die nur die Grauwertmuster auf der Bildebene zur Definition heranzieht und zum anderen diejenige, die die Bewegung

der Objekte im Raum zur Grundlage der Definition macht. Wir werden deshalb zwei Arten von Verschiebungsfeldern definieren.

Der Name "*optischer Fluß*" soll in dieser Arbeit die Bewegung von Grauwert- oder Farbmustern im Bild beschreiben. Eine solche wahrgenommene Bewegung kann mehrere Ursachen haben. In den meisten Anwendungen wird die Annahme gemacht, daß der optische Fluß mit der Projektion der 3-dimensionalen Geschwindigkeitsvektoren übereinstimmt, die die Bewegung der abgebildeten Objekte relativ zum Kamerakoordinatensystem beschreiben. Dieses Vektorfeld wird in der Literatur auch mit "image flow" [Waxman u.a. 87] oder "motion field" [Verri + Poggio 87] bezeichnet. Im weiteren werden wir dieses Feld *Bewegungsfluß* nennen.

Die Beziehung zwischen dem optischen Fluß und dem Bewegungsfluß ist nicht so offensichtlich. Stellen wir uns als Beispiel [Horn + Schunck 81] eine Kugel mit gleichmäßiger matter Oberfläche vor, die von einer Lichtquelle aus einer bestimmten Richtung beleuchtet wird. Wenn diese Kugel um eine Achse durch den Mittelpunkt rotiert, so ändert sich die Grauwertverteilung im aufgenommenen Bild (idealerweise) nicht, d.h. die Vektoren des optischen Flusses sind überall Null. Wenn aber die Kugel still steht und sich dafür nur die Lichtquelle bewegt, so registrieren wir auf der Bildebene einen optischen Fluß, der von Null verschieden ist, obwohl sich die Kugel selbst nicht bewegt hat. Der Zusammenhang zwischen projiziertem Bewegungsfeld und optischem Fluß hängt also von vielen Faktoren wie Beleuchtung und Oberflächeneigenschaften ab. Weitere Betrachtungen hierzu finden sich in Kapitel 3. Eine quantitative Analyse des Zusammenhangs ist im Anhang B zu finden.

Mit Verfahren, die nur die Ähnlichkeit von Grauwertmustern zur Korrespondenzbestimmung heranziehen, läßt sich aus den Bildern einer Bildserie nur der optische Fluß bestimmen, nicht jedoch das Bewegungsfeld, welches für die Rekonstruktionsalgorithmen benötigt wird. Eine Annäherung an das Bewegungsfeld läßt sich bestimmen, wenn man weiteres Wissen (spezieller oder allgemeiner Art) in das Verfahren einbezieht. Somit ist nicht immer eindeutig zu klären, ob ein bestimmtes Verfahren den optischen Fluß oder den Bewegungsfluß berechnet.

In der Praxis berechnet man nicht die *Geschwindigkeitsvektoren* des optischen Flusses selber, sondern die *Verschiebungsvektoren* zwischen den Grauwert- oder Farbmustern in aufeinanderfolgenden Bildern der Bildfolge. Man erhält ein diskretes Gitter von Vektoren, für welches wir ebenfalls den Namen optischer Fluß verwenden wollen.

1.2 Berechnung von Verschiebungsvektoren

Die meisten Verfahren, die versuchen das Korrespondenzproblem zu lösen, lassen sich in eine der beiden folgenden Klassen einordnen.

1.2.1 Merkmalszuordnungsverfahren

Eine Klasse von Methoden unterteilt das Problem der Verschiebungsvektorberechnung wiederum in zwei Schritte. In den Verfahren dieser Klasse wird zunächst eine Anzahl von möglichst markanten Merkmalen aus jedem Bild der Bildfolge extrahiert.

In einem zweiten Schritt müssen dann diejenigen Merkmale, die Abbildungen desselben Objektteils der Szene sind, einander zugeordnet werden. Für die Lösung dieses Problems ist es wichtig, daß die Merkmale möglichst "einzigartig" sind, d.h., daß sich die Merkmale durch möglichst viele Eigenschaften voneinander unterscheiden. Wegen der eindeutigen Lokalisation sind markante Punkte als Merkmale am beliebtesten [Moravec 80, Barnard + Thompson 80, Dreschler + Nagel 82]. Eine Diskussion von Punktfindern für die Korrespondenzanalyse unter besonderer Berücksichtigung von Farbe findet sich in Bartsch 87. Der Merkmalsfinder sollte unempfindlich gegen Rauschen und kleinere Verzerrungen der Grauwertstruktur sein und außerdem möglichst alle visuell markanten Punkte mit einer genauen Positionierung liefern.

Um die Zuordnungen zu bestimmen, werden meist mehrere Heuristiken angewendet. Häufig verwendet werden z.B. [Ballard + Brown 82]:

Maximale Verschiebung: Da sich Objekte nur mit einer endlichen Geschwindigkeit durch den Raum bewegen, legt man bei den meisten praktischen Anwendungen entsprechend der zu erwartenden Höchstgeschwindigkeit eine maximale Verschiebung fest.

Ähnliche Verschiebungen in benachbarten Bereichen: Glatte Oberflächen bewegter starrer Körper führen im allgemeinen zu einem glatten Verschiebungsvektorfeld. Diese Annahme ist allerdings an einigen Stellen, wie Objektkanten, nicht gerechtfertigt.

Glatte Bewegung und kleine Beschleunigungen: Da Körper mit endlicher Masse bei Anwendung endlicher Kräfte ihre Geschwindigkeiten nicht beliebig schnell ändern können, ergeben sich glatte Trajektorien im Raum. Auch die Projektionen dieser Trajektorien sollten außer in Spezialfällen auf glatte Kurven in der Bildebene abgebildet werden.

Ähnlichkeit der Merkmale: Weil sich Oberflächeneigenschaften von Objekten nicht abrupt ändern, ist es sinnvoll, eine gewisse Ähnlichkeit zwischen den Merkmalen in aufeinanderfolgenden Bildern zu verlangen.

Konsistente Zuordnungen: Aufgrund des Prinzips des generellen Standpunktes ist es sinnvoll anzunehmen, daß ein Merkmal in der Bildfolge erhalten bleibt. Daher sollten eindeutige Zuordnungen vorteilhaft bewertet werden.

Strukturelle Konsistenz: Die Beziehungen, die verschiedene Merkmale untereinander haben, sollten mit den Korrespondenzen konsistent sein. So sollten

z.B. die Endpunkte zweier korrespondierender Kanten ebenfalls korrespondieren.

Ob der Ansatz der Merkmalszuordnungen zur Lösung des Korrespondenzproblems seine Entsprechung im menschlichen Sehsystem hat, ist strittig. *Ullman 79* weist auf Experimente am menschlichen Sehsystem hin, die seine "minimal mapping" Theorie begründen sollen. *Jenkins + Kolers 86* schreiben hingegen:

"The correspondence of low-level tokens, determined by finding a matching between primitives of some type and dimensionality, is the major task of many computer algorithms. [...] The lack of success of these approaches may lie not in their choice of matching primitives, or in the rules used to combine the primitives, but rather it may lie in the approach itself: the assumption of matching as the fundamental operation in vision."

Im Gegensatz zu lokal arbeitenden Verfahren bleibt der Merkmalzuordnungsansatz auch für größere Verschiebungen anwendbar, wie sie z.B. beim Stereosehen vorkommen können. Der größte Nachteil dieser Verfahren ist, daß nur sehr dünn besetzte Verschiebungsvektorfelder berechnet werden können, da markante Merkmale recht selten in den Bildern auftreten. Der Aufwand der Korrespondenzermittlung steigt mit der Anzahl der Merkmale meist erheblich an.

1.2.2 Gradientenbasierte Verfahren

Die zweite Klasse bilden die gradientenbasierten Verfahren, mit denen wir uns in dieser Arbeit beschäftigen wollen. Sie beruhen auf der sogenannten "motion constraint equation", welche einen Zusammenhang zwischen den Verschiebungsvektoren und den Grauwertgradienten herstellt. Der Grauwertgradient gibt die Richtung der stärksten Grauwertänderung an. Mit seiner Hilfe läßt sich die Tangentialebene berechnen, von der angenommen wird, daß sie die Grauwertfunktion in einer lokalen Umgebung hinreichend genau approximiert. Mit der "motion constraint equation" läßt sich nun die Verschiebung dieser Ebene berechnen. Kapitel 3 enthält eine ausführliche Diskussion der Gleichung. Da sich die Gleichung nur auf lokale Größen bezieht, tritt das sogenannte *Blendenproblem* ("aperture problem" [*Marr + Ullman 81*]) in Erscheinung. Dieses besagt, daß lokal nur die Komponente des Verschiebungsvektors in Richtung des Gradienten bestimmt werden kann, da eine Bewegung der Grauwerte senkrecht zu dieser Richtung innerhalb einer kleinen Umgebung keine Änderung in der Bildstruktur hervorruft. Zusätzliche, meist heuristische Annahmen, müssen gemacht werden, um das Verschiebungsvektorfeld vollständig zu bestimmen. Diese laufen meist darauf hinaus, durch Glattheitsannahmen die Lokalität des Ansatzes aufzuheben. Einige dieser Verfahren werden im vierten Kapitel ausführlich hergeleitet und erläutert.

1.3 Übersicht

Bevor wir uns der Berechnung von Verschiebungsvektorfeldern mit dem gradientenbasierten Verfahren zuwenden, soll im folgenden Kapitel zunächst genauer beschrieben werden, wozu man Verschiebungsfelder einsetzen kann. Es gibt eine große Anzahl verschiedener Rekonstruktionsalgorithmen, die verschiedene Aspekte des Feldes ausnutzen oder Einschränkungen bzgl. der Bewegung der abgebildeten Objekte machen.

Sodann werden wir uns der Grundgleichung der gradientenbasierten Verfahren zur Verschiebungsfeldbestimmung zuwenden. Nach einer Herleitung der Gleichung wird ihr Gültigkeitsbereich geklärt. Danach werden die Fehler, die durch die Diskretisierung der Gleichung auftreten, hergeleitet. Schließlich wird der Zusammenhang zu verwandten Verfahren der Verschiebungsschätzung hergestellt.

Das Kapitel 4 enthält Vorschläge zur Lösung des Blendenproblems, welches bei der Verwendung der "motion constraint equation" auftritt. Die Verfahren werden in verschiedene Klassen eingeordnet, je nachdem, welche heuristischen Annahmen in die Berechnung des Feldes eingehen.

Dann wenden wir uns der Berechnung von Verschiebungsvektoren auf Farbbildern zu. Die Diskussion der Verfahren in den vorangehenden Kapiteln liefert dazu mehrere Ansatzpunkte. Der Haupt Gesichtspunkt dieses Kapitels richtet sich auf die Frage, welche Vorteile Farbbilder gegenüber Grauwertbildern für die Lösung des Blendenproblems haben. Eines der vorgeschlagenen Verfahren wird im anschließenden Kapitel mit entsprechenden Grauwertverfahren verglichen.

Diese Arbeit soll eine tiefere Einsicht in die Probleme vermitteln, die bei gradientenbasierten Verfahren zur Bestimmung des optischen Flusses in Grauwert- sowie in Farbbildern auftreten. Aus diesem Grund wird auch relativ viel Zeit auf die Analyse der vorgestellten Verfahren verwendet. Unser Ziel ist nicht die Implementation eines perfekten Verfahrens zur Verschiebungsvektorbestimmung. Es ist sinnvoll, nicht einfach nur alle Verschiebungsvektoren zu berechnen, sondern möglichst eine Auswahl von Vektoren zu treffen, die den Bewegungsfluß gut approximieren und sich deshalb für die Weiterverarbeitung besonders gut eignen. Im Hinblick auf dieses Ziel werden in Kapitel 5 einige Vorschläge gemacht und dann in Kapitel 6 anhand eines Beispiels erprobt.

Es sei hier noch darauf hingewiesen, daß die wichtigsten in den Formeln verwendeten Zeichen im Anhang A zusammengefaßt sind.

Kapitel 2

Verwendung von Verschiebungsvektoren

Unter der Annahme, daß das berechnete Verschiebungsvektorfeld mit den Projektionen der 3D-Bewegungsvektoren übereinstimmt, lassen sich daraus viele Eigenschaften der Szene rekonstruieren. So kann man die 3D-Bewegung der Objekte ermitteln, die Eigenbewegung der Kamera schätzen oder Oberflächenneigungen und relative Tiefeninformationen errechnen [Barron 84]. Verschiebungsvektoren lassen sich aber auch zur sparsamen redundanzfreien Kodierung von Bildfolgen verwenden [Robbins + Netravali 83]. Diese Anwendung wird hier nicht weiter betrachtet.

Die Algorithmen zur Szenenrekonstruktion benötigen im allgemeinen sehr genaue Verschiebungsvektoren, um korrekte Ergebnisse zu liefern. Da die Übereinstimmung des optischen Flusses mit dem Bewegungsfluß nicht immer gewährleistet ist, kann die geforderte Genauigkeit oft nicht erreicht werden. Der optische Fluß läßt sich aber auch dann noch sinnvoll einsetzen. So reicht die Qualität der Verschiebungsvektoren im allgemeinen dazu aus, Bereiche zu unterscheiden und zu segmentieren, die sich mit verschiedenen Geschwindigkeiten bewegen.

Zunächst werden wir Algorithmen beschreiben, die nur ein dünn besetztes Verschiebungsfeld benötigen und mit einer minimalen Anzahl von Punkten in möglichst wenig Bildern auskommen.

Es ist auch möglich, globale Eigenschaften des gesamten Bewegungsfeldes auszunutzen. Dies ist vor allem dann angebracht, wenn man die Eigenbewegung einer Kamera bei stationärem Hintergrund berechnen will, denn dann läßt sich das gesamte Bewegungsfeld durch eine Rotation und eine Translations beschreiben. Die Ausnutzung aller Verschiebungsvektoren verringert die Fehleranfälligkeit.

Am Schluß dieses Kapitels werden wir noch Verfahren betrachten, die lokale Parameter des Flusses zur Rekonstruktion nutzen. Diese Verfahren beruhen meist auf einem Modell der Form der betrachteten Objekte.

2.1 Verwendung von einzelnen Punktkorrespondenzen

Ullman stellte in *Ullman 79* fest, daß der Mensch in der Lage ist, allein aufgrund der Bewegung eines (unbekannten) Objektes, dessen Bewegung und Form zu erkennen. Dies gilt auch für den Fall, in dem die Einzelbilder keinerlei Informationen über die Objektform enthalten. Ein Beispiel hierfür ist die Zylinder-Demonstration: Eine Menge von nicht verbundenen Punkten wird auf einem Bildschirm so bewegt, als ob sie orthographisch projizierte Abbildungen von Objektpunkten seien, die auf der Oberfläche zweier ineinandergeschachtelter Zylinder liegen, die sich um ihre Achsen in entgegengesetzter Richtung drehen. Ullman weist darauf hin, daß die Punktkorrespondenzen allein nicht ausreichend sind, um dieses Problem der Rekonstruktion zu lösen. Es muß noch eine weitere Menge von Bedingungen geben, die die Menge der möglichen 3D-Interpretationen einschränkt. Ullman schlägt vor, die Starrheit der abgebildeten Objekte anzunehmen:

“Any set of elements undergoing a two-dimensional transformation which has a unique interpretation as a rigid body moving in space should be interpreted as such a body in motion.”

Diese Starrheitshypothese wird praktisch von allen weiteren Arbeiten zur 3D-Rekonstruktion aus Verschiebungsvektoren verwendet. Aus der Annahme folgt, daß sich die Bewegung eines Objektes durch 6 Parameter beschreiben läßt, nämlich einem 3D-Translationsvektor und einer Rotation, die sich entweder durch eine Rotationsmatrix beschreiben läßt, oder durch Angabe der Rotationsachse und des Rotationswinkels. Meistens wird die Rotation auf das Projektionszentrum bezogen. Wenn im folgenden von einer eindeutigen Lösung gesprochen wird, dann ist damit nicht gemeint, daß alle 6 Bewegungsparameter eindeutig bestimmt werden können. Vielmehr bleibt ein Skalierungsfaktor unbestimmt, der entweder die Tiefe eines der beteiligten Punkte oder die Komponente der Translationsgeschwindigkeit in die Tiefe angibt. Somit kann immer nur das Verhältnis von Translation zur Tiefe berechnet werden. Aus der Analyse des Bewegungsflusses kann Ullman nun sein “structure from motion theorem” herleiten:

“Given three distinct orthographic projections of four non-coplanar points in a rigid configuration, the structure and motion compatible with the three views are uniquely determined up to a reflection about the image plane.”

Die Angabe der minimal notwendigen Anzahl von Punkten und Abbildungen ist wichtig, um das Berechnungsschema möglichst lokal anwenden zu können, denn das Verfahren kann nicht davon ausgehen, daß sich im Bild nur ein starres Objekt bewegt. Die Annahme, daß alle Punkte zum selben Objekt gehören, ist außerdem umso eher erfüllt, desto näher sie beisammen liegen. Man kann auch

für jede lokale Gruppe von Punkten eine andere orthographische Projektion verwenden, deren Achse vom Kameraursprung zur Mitte der betrachteten Region verläuft. Dieses Schema, welches Ullman die polar-parallel-Methode nennt, kann so, in begrenztem Rahmen, auch perspektivische Effekte berücksichtigen. Die volle Berücksichtigung der perspektivischen Effekte durch eine Zentralprojektion ist zwar ebenfalls möglich, jedoch nach Ullman nicht ratsam, weil sie gerade bei lokaler Anwendung fehleranfällig ist. Außerdem haben Experimente ergeben, daß der Mensch wahrscheinlich keine direkten Interpretationsmethoden zur Erkennung von sich ändernden perspektivischen Projektionen verwendet:

“... the human visual system is not on par with the perspective interpretation scheme in its ability to infer structure from motion in depth, a fact that strongly suggest that it does not possess the capacity for directly interpreting structure from perspective projection.”

Aus einer einfachen Abzählung der Unbekannten gegenüber der Anzahl der Gleichungen ergibt sich eine untere Grenze für die minimale Anzahl der Punkte, die man bei einer bestimmten Anzahl von Ansichten benötigt, um alle Bewegungsparameter sowie die 3D-Koordinaten der abgebildeten Punkte zu erhalten. Wir haben pro Punkt und Ansicht 2 Gleichungen für die beiden Bildkoordinaten. Pro Punkt suchen wir die 3 Raumkoordinaten und pro Abbildung die relative Translation und Rotation zur Ausgangslage. Ein Skalierungsfaktor bleibt unbestimmt. Seien also m Punkte in n Ansichten gegeben, so gilt die Ungleichung [Meiri 80]:

$$2nm \geq 3m + 6(n - 1) - 1 \Leftrightarrow m \geq 3 + \frac{2}{2n - 3}$$

[Mitiche 84] behauptet hingegen, schon mit 4 Punkten aus 2 Ansichten alle Bewegungsparameter und Tiefeninformationen gewinnen zu können. Bei genauerem Hinsehen erkennt man jedoch, daß er in seinem Ansatz die Starrheitshypothese auf zwei verschiedene Weisen benutzt, so daß seine nichtlinearen Gleichungen voneinander abhängig sind.

Nagel + Neumann 81 leiten eine Verallgemeinerung der Polargleichung von Ullman 79 her. Es werden 3 nichtlineare Gleichungen für die Bestimmung der Rotationsmatrix angegeben, die sich aus 5 Punkten in 2 perspektivischen Ansichten ergeben. Geometrisch interpretiert, sagen die Gleichungen aus, daß sich die Ebenen, die durch jeweils einen verschobenen Punkt, den nur rotierten Punkt und das Kamerazentrum definiert sind, alle in einer durch den Translationsvektor gegebenen Geraden schneiden. Kennt man einmal die Rotationsmatrix, so kann man durch Einsetzen auch den Translationsvektor und die 3D-Punktkoordinaten bestimmen. Da es sich aber um nichtlineare Gleichungen handelt, ist eine eindeutige Lösung nicht in jedem Fall zu erwarten.

In der Näherung von kleinen Rotationswinkeln¹ zeigen Fang + Huang 84, daß 9 Punkte in 2 perspektivischen Ansichten zu einer eindeutigen Lösung führen,

¹Sei θ der Rotationswinkel, dann werden in der Arbeit die Näherungen $\cos \theta = 1$ und $\sin \theta = \theta$ gemacht. Praktisch zeigte sich, daß der Winkel 0.1 rad nicht übersteigen darf.

falls die Punkte nicht auf einer Fläche zweiter Ordnung durch den Nullpunkt liegen. Experimentelle Ergebnisse zeigen, daß meistens bereits 6 zufällig ausgewählte Punkte ausreichen, die Parameter zu berechnen. Typisch für die Ausnutzung perspektivischer Effekte ist, daß die numerischen Schwierigkeiten mit der Entfernung der Objekte zunehmen. Die Experimente mit verrauschten synthetischen Eingabedaten zeigen auch, daß die Genauigkeit des Verfahrens stark von den Eingangsfehlern abhängt. Außerdem erfordert die Verwendung iterativer numerischer Lösungsverfahren für die nichtlinearen Gleichungen einen guten Anfangswert.

Eine weitere Möglichkeit, das Problem durch Einschränkung der Bewegung zu vereinfachen, zeigen *Webb + Aggarwal 82*. Sie stellen die sog. "fixed axis assumption" auf, die besagt, daß die Rotationsachse während der Beobachtungszeit von mindestens 4 Bildern konstant bleibt. Dadurch beschreibt jede Bewegung eines Punktes des starren Objektes relativ zu jedem anderen gegebenen Objektpunkt einen Kreis, senkrecht zur festen Drehachse. Durch eine Parallelprojektion wird dieser Kreis auf eine Ellipse in der Bildebene abgebildet. Mindestens 4 Abbildungen von 2 Punkten reichen aus, um diese Ellipse zu bestimmen, aus der sich dann auch die relative Lage des Kreises und somit die Rotationsachse ermitteln läßt. Alle weiteren Parameter der Bewegung können ebenfalls ermittelt werden. Noch weiter wird die Bewegung beim Rekonstruktionsverfahren in *Bonde + Nagel 79* oder *Dreschler + Nagel 82* eingeschränkt. Die Aufnahme einer Straßenszene erlaubte es anzunehmen, daß sich ein Objekt nur innerhalb einer Ebene bewegt und die Rotationsachse somit nur senkrecht auf dieser Ebene stehen kann.

Tsai + Huang 84 zeigten, daß 7 Punkte in 2 perspektivischen Ansichten, unter bestimmten relativ schwachen einschränkenden Bedingungen bezüglich der Lage der Punkte im Raum, eine eindeutige Lösung garantieren. Sie berechnen aus 8 Punktepaaren eine Menge von 8 "essentiellen" Parametern. Diese hängen allein von der Translation und Rotation ab und nicht von den 3D-Koordinaten der Punkte. Somit erreicht man, daß die Bewegungsschätzung von der 3D-Rekonstruktion separiert wird. Hat man aber die Bewegungsparameter berechnet, so ist das Problem der 3D-Rekonstruktion durch Triangulation, ähnlich wie beim Stereosehen, möglich. Die 8 Parameter können mittels eines linearen Gleichungssystems aus den Eingabedaten gewonnen werden. Jedes Punktepaar liefert dabei eine Gleichung, die geometrisch interpretiert besagt, daß der verschobene Punkt, der nur rotierte Punkt und der Translationsvektor koplanar sind. Aus diesen essentiellen Parametern können dann in geschlossener Form (ohne Suche) der Translationsvektor und die Rotationsmatrix bestimmt werden. Für eine anschauliche Darstellung dieser Dekomposition siehe auch *Nagel 86*. In Experimenten mit fehlerbehafteten Daten zeigt sich, daß die Rotationsparameter zuverlässig bestimmt werden können, wenn man mittels quadratischer Minimierung mehr als nur 8 Punktepaare zur Berechnung der essentiellen Parameter heranzieht. Leider bleibt auch in diesen Fällen die Schätzung des Translationsvektors sehr ungenau.

2.2 Verwendung von globalen Eigenschaften

Will man die Eigenbewegung einer Kamera berechnen, wobei angenommen werden darf, daß die Umgebung unbewegt bleibt, so lassen sich auch globale Eigenschaften des Bewegungsflusses ausnutzen. Ist z.B. der Rotationsanteil des Verschiebungsvektorfeldes Null, so haben alle Verschiebungsvektoren im Raum dieselbe Richtung. Die Projektionen der Vektoren laufen dann entweder alle auf einen Punkt (Expansionspunkt oder "focus of expansion") zu oder scheinen von einem Punkt (Kontraktionspunkt) auszugehen. Nur die Länge des Verschiebungsvektors hängt von der Tiefe des abgebildeten Objektteils ab. *Prazdny 81* schlägt vor, die Rotation aus dem Bewegungsfeld herauszurechnen. Die Tatsache, daß das entstehende Vektorfeld einen Expansionspunkt enthalten muß, läßt sich für die Berechnung der Rotationsparameter ausnutzen. Die Translationsrichtung ist dann durch den Expansionspunkt ebenfalls eindeutig bestimmt. Ein solches Verfahren haben *Kories u.a. 86* auf natürliche Bilder angewendet.

Auch in *Bruss + Horn 83* wird ein globales Minimierungsverfahren zur Bestimmung der Eigenbewegung vorgeschlagen. Im Falle reiner Translation wird zunächst die Tiefe eliminiert, indem diejenige Tiefe in Abhängigkeit der Translation berechnet wird, die den gemessenen Verschiebungen am nächsten kommt. In einem zweiten Schritt wird dann der Translationsvektor durch ein Minimierungsverfahren gewonnen. Es wird eine Norm angegeben, die ein höheres Gewicht auf die Ähnlichkeit zu den großen Verschiebungen legt, so daß das Verfahren auf die Minimierung einer quadratischen Form und somit auf ein wohlbekanntes Eigenwertproblem zurückgeführt werden kann. Im Falle reiner Rotation hängen die Vektoren nicht von der Tiefe der abgebildeten Oberflächen ab, und so kann man den zu minimierenden Ausdruck direkt nach den Rotationsparametern ableiten und erhält, bei der üblichen Norm, ein lineares Gleichungssystem, welches sich nach den Rotationsparametern auflösen läßt. Bei beliebiger Bewegung muß eine numerische Methode zur Lösung des nichtlinearen Gleichungssystems angewendet werden.

2.3 Verwendung von Objektmodellen

Neueren Datums sind Ansätze, die den Bewegungsfluß analysieren, der von einer Ebene unter orthographischer [*Kanatani 86*] oder perspektivischer Projektion [*Subbarao + Waxman 86, Kanatani 87*] erzeugt wird. Es stellt sich heraus, daß es im allgemeinen zwei Lösungen zu diesem Problem gibt, die dual zueinander sind. Die beiden Lösungen fallen zusammen, wenn der translatorische Anteil der Verschiebungen entweder in Richtung der Ebenennormalen zeigt oder parallel zur Bildebene verläuft. Interessant ist auch, daß der Translationsanteil der Bewegung bei perspektivischer Projektion eindeutig bestimmt werden kann, wenn man den Schnittpunkt der Ebene mit der optischen Achse als Bezugspunkt der Rota-

tion nimmt [Kanatani 87]. Im Grenzfall der pseudo-orthographischen Projektion² gibt es ebenfalls nur noch eine mögliche Lösung [Kanatani 87]. Dies bedeutet, daß die zweite Lösung durch extrem perspektivische Effekte zustande kommt und daher eventuell fallengelassen werden kann. Der große Vorteil dieser Analysen ist, daß sich die Lösungen in geschlossener Form angeben lassen.³ Benutzt werden nur lokale Eigenschaften des Bewegungsflusses ("Deformationsparameter" oder "Flußparameter" genannt), nämlich die räumlichen Ableitungen bis zur ersten Ordnung (bei orthographischer Projektion) bzw. bis zur zweiten Ordnung (bei perspektivischer Projektion). Zur Bestimmung dieser Parameter werden mindestens 4 Punkte in 2 Ansichten benötigt. Hat man mehr Punkte zur Verfügung, so kann man auch entscheiden, ob ein gegebenes Verschiebungsfeld dem Bewegungsfluß einer Ebene entsprechen kann. Die richtige der dualen Lösungen kann ermittelt werden, wenn bei konstanter bzw. nur leicht veränderter Bewegung zwei Verschiebungsvektorfelder zu zwei aufeinanderfolgenden Zeitpunkten vorliegen oder wenn zwei starr miteinander verbundene Ebenen beobachtet werden. Entscheidbar ist auch die Frage, ob zwei Ebenen starr miteinander verbunden sind, d.h. ob die Bewegung der Ebenen mit dem gleichen Satz von Parametern beschrieben werden kann.

Einen Schritt weiter gehen Waxman u.a. 87, die das Modell einer Ebene auf gekrümmte Oberflächen ausgedehnt haben. Dazu werden die 12 Ableitungen des Verschiebungsfeldes bis zur zweiten Ordnung (Deformationsparameter genannt) in Beziehung gesetzt zu den Größen, die die Bewegung und die Form der Oberfläche bestimmen. Dies sind die Translationsgeschwindigkeit, die Rotationsparameter, die Neigung der Tangentialebene sowie der Krümmungstensor der Oberfläche (insgesamt 11 Größen). Durch diese Parameter wird die Bewegung und Form einer Quadrik beschrieben. Eine solche Quadrik erzeugt leider einen Bewegungsfluß, der auch Variationen dritten Grades enthält. Glücklicherweise tauchen die Krümmungsparameter schon in den Ableitungen zweiter Ordnung auf.

Um die 11 Größen aus den 12 nichtlinearen Gleichungen zu erhalten, rotiert man das Verschiebungsfeld (bzw. transformiert die Deformationsparameter direkt), so daß die Translationsgeschwindigkeit in X-Richtung verschwindet. Das Problem der Berechnung der Tangentialebene und der Bewegungsparameter ist dann separiert von der Berechnung der Krümmung, d.h. man kann zur Rekonstruktion der Ebene (Bewegung und Neigung) die bekannten Gleichungen auf die ersten 8 Deformationsparametern anwenden. Die Krümmungen und der Transformationswinkel ergeben sich dann aus den restlichen 4 Deformationsparametern. Insgesamt ist die Lösung analytisch und erfordert die Lösung von 2 kubischen Gleichungen (eine für die Berechnung der richtigen Transformation der Deforma-

²Sei f die Brennweite der Kamera, dann wird bei der pseudo-orthographischen Projektion der Term $1/f^2$, nicht jedoch der Term $1/f$, vernachlässigt. Diese Vereinfachung läßt sich in den üblicherweise gewählten Koordinatensystemen, die das optische Zentrum als Nullpunkt wählen, nicht gut durchführen. Besser ist es, den Nullpunkt des Koordinatensystems in die Bildebene zu legen.

³Es muß die mittlere Nullstelle einer kubischen Gleichung gefunden werden.

tionsparameter und eine für die Lösung der Tangentialebene). Auch hier können Mehrdeutigkeiten entstehen. Dies ist der Fall, wenn der Translationsvektor, die Ebenennormale und die Sichtlinie in einer Ebene liegen, wobei keine zwei dieser drei Richtungen parallel zueinander sind ("structure-motion coincidence"). Weitere davon unabhängige Möglichkeiten einer mehrdeutigen Lösung ergeben sich für zylindrische und sattelförmige Flächenstücke (2 bzw. 3 Lösungen). Eine Oberfläche, bei der beide Hauptkrümmungen das gleiche Vorzeichen haben, erlaubt eine eindeutige Lösung.

Die oben vorgestellten Verfahren zur Objekt- und Bewegungsrekonstruktion sind anfällig gegenüber fehlerhaften Eingangsdaten. Die geforderte Präzision des Verschiebungsfeldes kann meist nicht geliefert werden. Man kann jedoch auch im perspektivischen Fall allein aus den Ableitungen des Feldes bis zur ersten Ordnung nützliche Einschränkungen der Bewegungsparameter berechnen [Subbarao 87].

2.4 Zusammenfassung

Zum Abschluß sollen noch einmal die zugrundeliegenden Annahmen und Vereinfachungen der verschiedenen aufgezählten Verfahren aufgezählt werden. Je nach Anwendung kann das eine oder andere Verfahren ausgewählt werden.

Projektionsart: orthographisch, pseudo-orthographisch, perspektivisch, sphärisch, polar-parallel.

Verwendung des optischen Flusses: einige ausgewählte Stellen, lokale Eigenschaften wie Deformationsparameter, globale Eigenschaften des Feldes wie der Expansionspunkt mittels Minimierungsverfahren.

Einschränkung der Bewegung: Nur ein bewegtes Objekt, bewegte Kamera aber unbewegter Hintergrund, reine Translation oder reine Rotation, feste Rotationsachse während der ganzen Beobachtungsphase, kleiner Rotationswinkel, Bewegung nur innerhalb einer Ebene.

Objekteigenschaften: Starrer Körper, ebene Oberflächen, Quadriken.

Lösungsmethode: Iterative, numerische oder analytische Lösung, lineares oder nichtlineares System von Gleichungen.

Koordinatensystem bzw. Bezugspunkt: Ursprung in der Bildebene oder im Projektionszentrum. Kamerazentrierter oder objektzentrierter Standpunkt.

Kapitel 3

Der differentielle Ansatz

Wir beginnen mit einer ausführlichen Analyse der Grundgleichung, die den gradientenbasierten Verfahren zugrunde liegt. Dabei wird geklärt werden, welche Voraussetzungen für die Anwendung dieser Gleichung gemacht werden müssen. In der Literatur wird nur vereinzelt auf diese Grundannahmen eingegangen. Auch auf die Analyse der Rauscheinflüsse wird meistens verzichtet.

Zunächst leiten wir die Gleichung her und geben eine geometrische Interpretation an. Sodann werden die Bedingungen untersucht, unter denen die Gleichung den Bewegungsfluß richtig beschreibt (siehe dazu auch Anhang B). Zusätzliche Ungenauigkeiten kommen durch die Diskretisierung und Bildrauschen zustande. Im dritten Abschnitt werden die Fehler einer diskreten Version der Gleichung analysiert. Schließlich wird die Gleichung noch in Zusammenhang gebracht mit anderen Verfahren der Bewegungsanalyse. Zum einen ist dies die Minimierung der quadrierten Grauwertdifferenz und zum anderen die Anwendung von raumzeitlichen Filtern.

3.1 Die "motion constraint equation"

Der differentielle Ansatz erlaubt es, eine Komponente des Verschiebungsvektors direkt auszurechnen. Dazu werden die örtlich-zeitlichen Ableitungen der Bildfunktion $g(x, y, t)$ in Beziehung gesetzt zu den Verschiebungen (u, v) in der Bildebene. Damit der Bewegungsfluß durch diese Gleichung richtig beschrieben wird, muß die Annahme gemacht werden, daß die beobachteten Grauwerte aller Objektpunkte sich zeitlich nicht ändern. Das bedeutet, daß alle Änderungen durch reine Bewegung der Bildfunktion auf der Bildebene zustande kommen. Betrachten wir z.B. einen Punkt an der Stelle (x, y) zur Zeit t , der sich in dem Zeitintervall δt an die Stelle $(x + \delta x, y + \delta y)$ bewegt. Da die Intensität gleich bleibt, kann man schreiben:

$$g(x, y, t) = g(x + \delta x, y + \delta y, t + \delta t)$$

Durch Taylorentwicklung um den Punkt (x, y, t) erhalten wir:

$$g(x + \delta x, y + \delta y, t + \delta t) = g(x, y, t) + \frac{\partial g}{\partial x} \delta x + \frac{\partial g}{\partial y} \delta y + \frac{\partial g}{\partial t} \delta t + \dots$$

Teilt man durch den Zeitabschnitt δt und macht den Grenzübergang $\delta t \rightarrow 0$, so fallen alle Terme höherer Ordnung weg und man erhält die Gleichung:

$$g_x \frac{\delta x}{\delta t} + g_y \frac{\delta y}{\delta t} + g_t = g_x u + g_y v + g_t = \nabla g \cdot \mathbf{u} + g_t = 0 \quad (3.1)$$

wobei gilt (siehe auch Anhang A):

$$g_x = \frac{\partial g}{\partial x}, \quad g_y = \frac{\partial g}{\partial y} \quad \text{und} \quad g_t = \frac{\partial g}{\partial t}.$$

Eine ausführliche Ableitung findet man z.B. in *Horn + Schunck 81*. Noch kürzer kann diese Gleichung mit dem totalen Differential geschrieben werden: $dg/dt = 0$. Man kann diese Gleichung als Definition des optischen Flusses ansehen.

Die Gleichung wird in der Literatur mit verschiedenen Namen bezeichnet. So wird sie in *Schunck 84a* oder in *Barron 84* "motion constraint equation" genannt, in *Schunck 86a* heißt sie "optical flow constraint equation", und *Kearney u.a. 87* bezeichnen sie schließlich als "gradient constraint equation".

Bei dieser Herleitung wurde die Differenzierbarkeit der Bildfunktion g (für die Taylorentwicklung) vorausgesetzt. Man kann jedoch auch eine andere Herleitung für den *kontinuierlichen* Fall angeben, die diese Voraussetzung nicht macht und das Ergebnis auch an Unstetigkeiten des Grauwertbildes, bzw. Verschiebungsvektorfeldes bestätigt [*Schunck 86a*]. Bei dieser Herleitungen muß allerdings vorausgesetzt werden, daß man sich der Unstetigkeitsstelle beliebig dicht nähern kann. Da dies im *diskreten* Fall nicht gewährleistet ist, können an diesen Stellen erhebliche Abweichungen von der "motion constraint equation" auftreten.

Die "motion constraint equation" bestimmt eine Gerade im (u, v) -Raum (die sog. "constraint line"), auf der sich die möglichen Verschiebungen befinden müssen (siehe Abbildung 3.1). Dies drückt die Tatsache aus, daß wir mit einem lokalen Ansatz, wie ihn die "motion constraint equation" darstellt, nur die Komponente der Verschiebung in Richtung des Gradienten schätzen können. Diese Komponente der Verschiebung wird in der Literatur "normal flow" genannt. Der Grund für diese Namensgebung liegt darin, daß ein großer Gradient sich im Bild durch eine Kante bemerkbar macht. Die Gradientenrichtung liegt senkrecht zur "Kantenrichtung", und somit liegt die bestimmbare Komponente der Verschiebung normal zur "Kantenrichtung". Wir werden diese Komponente der Verschiebung mit u^\perp bezeichnen. Sie ergibt sich aus der "motion constraint equation" zu:

$$u^\perp = \frac{g_x u + g_y v}{|\nabla g|} = \frac{-g_t}{|\nabla g|} = \rho \cos(\alpha - \beta) \quad (3.2)$$

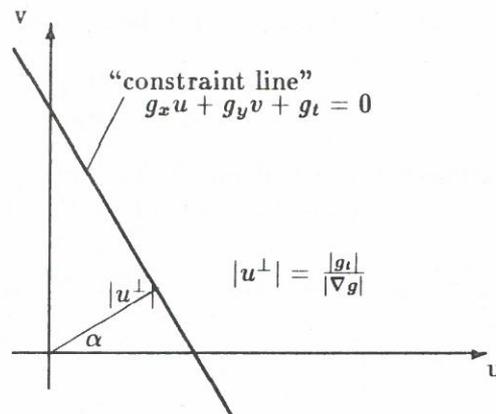


Abbildung 3.1: Die "constraint line". α ist die Gradientenrichtung, u^\perp die Komponente von u in Richtung des Gradienten.

Der letzte Term entspricht der "motion constraint equation" in Polarform, wie sie Schunck 86a vorgestellt hat. Dabei ist ρ der Geschwindigkeitsbetrag, β die Richtung der Geschwindigkeit und α bezeichnet den Winkel der "constraint line" also die Gradientenrichtung (siehe Abbildung 3.1). Diese Form der Gleichung setzt nur 4 Größen in eine Beziehung. Wir werden darauf in Kapitel 6 zu Darstellungszwecken zurückkommen.

Kritik an der "motion constraint equation" 3.1 kann man an zwei Punkten anbringen.

- Es gibt stark einschränkende Voraussetzungen.
- Es treten praktische Schwierigkeiten bei der Diskretisierung auf.

Außerdem bleibt das Problem bestehen, die zweite Komponente des Verschiebungsvektorfeldes zu bestimmen.

3.2 Diskussion der Voraussetzungen

Streng genommen ist die Annahme gleicher Intensitäten der Abbildungen derselben Objektpunkte zu unterschiedlichen Zeitpunkten kaum erfüllbar. Folgende Effekte haben einen Einfluß auf die Intensität, die ein Objektpunkt auf der Bildebene erzeugt (vergl. mit Grimson 83):

- Die *Flächenhelligkeit* des Objektes. (Die aus dem senkrecht zum Flächenelement dF stehenden Raumwinkelement $d\Omega$ eingestrahlte Leistung pro Fläche und Raumwinkel.) Diese ist abhängig von der *Entfernung* r der Lichtquelle, sowie deren *Intensität* I und außerdem von der Orientierung

der Fläche relativ zur Lichtquelle. Diese wird durch den Einfallswinkel φ beschrieben.

- Der Prozentsatz an reflektiertem Licht. Dieser hängt vom Material ab und wird durch die Albedo ρ festgelegt.
- Die Menge des zum Beobachter reflektierten Lichtes, abhängig von der Verteilung des reflektierten Lichtes über die verschiedenen Richtungen (Reflektanzkarte) und den Winkeln zwischen Beobachter und Flächennormalen (Ausfallswinkel ψ), sowie zwischen Lichtquelle und Beobachter (Phasenwinkel ϑ).

Hierbei wurde noch eine Einschränkung auf isotrope Materialien gemacht, bei denen eine Drehung der Oberfläche um die Flächennormale keine Auswirkung auf die Reflektanzeigenschaften hat. Man beachte, daß die Entfernung des Beobachters vom Objekt keinen Einfluß auf die Helligkeit des Bildpunktes hat, weil die Abnahme der Lichtmenge sich mit der Abnahme der Fläche die Waage hält. (Bei einer punktförmigen Lichtquelle spielt diese Entfernung wegen der fehlenden Ausdehnung sehr wohl eine Rolle, siehe z.B. in der Astronomie).

Wenn man eine, relativ zur Ausdehnung der Szene sowie zu den Verschiebungen, weit entfernte, zeitlich konstant strahlende Lichtquelle annimmt, so kann man den Term I/r^2 (das Verhältnis der Intensität zum Quadrat der Entfernung zur Lichtquelle) als konstant über die raum-zeitliche Ausdehnung der Szene ansehen. Nimmt man außerdem einen weit entfernten Beobachter an, so ist der Phasenwinkel ϑ über die Ausdehnung der Szene ebenfalls konstant. Die Einflüsse der beiden anderen Winkel (Einfallswinkel φ und Ausfallswinkel ψ) lassen sich wegen der Abhängigkeit zur Oberflächenorientierung nicht vernachlässigen. Soll die "motion constraint equation" erfüllt werden, so müssen die Bewegungen der Objekte oder deren Reflektanzeigenschaften sowie die Beleuchtung der Szene stark eingeschränkt werden. Um den Einfluß vom Ausfallswinkel ψ vernachlässigen zu können, brauchen wir z.B. eine Oberfläche, die in alle Richtungen die gleiche Leistung abstrahlt. Ein Körper mit dieser Eigenschaft wird Lambertscher Strahler genannt. Der Einfluß vom Einfallswinkel φ kann vernachlässigt werden, wenn das Licht aus allen Richtungen gleich einstrahlt. Am besten eignen sich also matte Oberflächen bei diffuser Beleuchtung zur Berechnung des optischen Flusses.

Im Anhang B ist eine Herleitung der Abhängigkeit der "motion constraint equation" von Bewegungsparametern und Beleuchtungsgrößen im allgemeinen Fall angegeben. Die Gleichung B.1 zeigt den Einfluß der verschiedenen Parametertypen auf die "motion constraint equation" auf. Weil eine volle Übereinstimmung des mit der "motion constraint equation" berechneten optischen Flusses mit den auf die Bildebene projizierten 3D-Geschwindigkeiten praktisch nur in Spezialfällen zu erreichen ist, haben verschiedene Autoren Zweifel angemeldet, ob sich der so berechnete optische Fluß für die 3D-Rekonstruktion der Objekte verwenden läßt. *Verri + Poggio 87* schlagen vor, den optischen Fluß nur zur qualitativen Bildauswertung, wie z.B. Bewegungsdetektion und Segmentierung, zu nutzen.

In mehreren Arbeiten ([Schunck 84b, Schunck 86a] oder [Barron 84, S. 2-18, 2-28]) ist behauptet worden, daß die "motion constraint equation" bei Rotation und perspektivischer Projektion nicht angewendet werden kann, weil Verzerrungen (divergente Terme) der Objektoberflächen auf der Bildebene entstehen. So schreibt Schunck in einem Abschnitt "Validity of the Image Flow Equation" des Artikels [Schunck 86a, S. 33f]:

"Only two conditions are required to ensure the validity of the image flow constraint equation¹: (1) the image must be smooth except at a finite number of discontinuities and (2) the perceived change in image irradiance at each point in the image plane must be entirely due to motion of the image pattern as opposed to changes in the pattern due to reflectance effects. [...]

It is necessary for the albedo to vary for motions such as rotational motion to be perceived. [...] If the variation in reflectance due to motion is ignored (as is often assumed), then $E(x, y, t)$ ² depends only on the albedo $\rho(\xi, \eta)$. In that case, the variation in $E(x, y, t)$ due to motion cannot be modeled by the image flow constraint equation (8) for the flow of image irradiance, because the albedo pattern will not simply be shifted within the image plane but rather the albedo pattern will be distorted by the change in surface orientation relative to the viewer."

Dies stimmt, jedenfalls für den kontinuierlichen Fall, nicht. Im Anhang C ist eine Ableitung für den Fall einer beliebig rotierenden Fläche unter orthographischer Projektion angegeben.

Es sei noch der Zusammenhang der "motion constraint equation" mit der Physik der Geschwindigkeitsfelder bewegter Flüssigkeiten angeführt. Die "motion constraint equation" weist eine gewisse Ähnlichkeit mit der Randbedingung für das Geschwindigkeitsfeld einer inkompressiblen Flüssigkeit auf. Die Massendichte ρ muß nämlich der einschränkenden Bedingung [Potter 73]

$$\frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho = 0$$

genügen. Diese Gleichung ist vollkommen analog zur "motion constraint equation" (man muß nur die Massendichte ρ durch die Grauwertfunktion g sowie die Geschwindigkeit \mathbf{v} durch \mathbf{u} ersetzen). In der Physik läßt sich aus dieser Gleichung auf die Divergenzfreiheit des Geschwindigkeitsfeldes schließen, denn es gilt zusätzlich die Kontinuitätsgleichung:

$$\frac{\partial \rho}{\partial t} + \nabla(\mathbf{v}\rho) = \frac{\partial \rho}{\partial t} + \rho \nabla \mathbf{v} + \mathbf{v} \nabla \rho = 0$$

¹Hier die "motion constraint equation" von Gleichung 3.1.

² $E(x, y, t)$ ist bei Schunck die Bildfunktion, die wir $g(x, y, t)$ genannt haben.

Diese Gleichung wird aus dem Erhaltungssatz für die Masse gewonnen. Die Änderung der Gesamtmasse innerhalb eines Gebietes S muß dem Fluß von Masse durch dessen Rand ∂S entsprechen:

$$\int_{\partial S} (\rho \mathbf{v}) \cdot \mathbf{n} \, dl + \int_S \frac{\partial \rho}{\partial t} \, dS = 0$$

Dabei ist \mathbf{n} der Einheitsvektor senkrecht zum Rand.

Diese Ableitung der Divergenzfreiheit des Geschwindigkeitsfeldes einer inkompressiblen Flüssigkeit kann aber nicht ohne weiteres auf den optischen Fluß übertragen werden. Man kann nur dann $\nabla \cdot \mathbf{u} = 0$ folgern, wenn sich die Intensität wie eine Dichte verhalten würde. Dies ist aber nicht der Fall, obwohl die Einheit der Intensität (Leistung pro Fläche) dies vielleicht vermuten läßt. Bei einer Verkürzung einer Oberfläche durch Rotation müßte sich die Helligkeit dann nämlich erhöhen, damit die integrierte Gesamthelligkeit der Oberfläche erhalten bleibt. *Schunck 86b* weist darauf hin, daß man die Kontinuitätsgleichung jedoch für die Berechnung des optischen Flusses verwenden sollte, wenn man auf Größen wie Textdichte oder Kantendichte arbeitet. Allerdings ist noch nicht geklärt, wie diese neue Gleichung in der Bildanalyse eingesetzt werden kann und ob sie Vorteile gegenüber der herkömmlichen Methode bietet.

3.3 Der diskrete Fall

Die bisherigen Erörterungen beziehen sich alle auf den Fall einer kontinuierlichen, rauschfreien Bildfolge. In der Praxis haben wir es aber mit einer Anzahl von diskreten Bildmatrizen zu tun. Es muß also eine diskrete Form der "motion constraint equation" gefunden werden. Eine Möglichkeit besteht darin, die Ableitungen durch Anwendung diskreter Operatoren zu schätzen und, statt der Geschwindigkeiten, die Verschiebungen in der Zeit Δt zu benutzen. Dem liegt die Annahme zugrunde, daß die Bildmatrix einer Abtastung der eigentlichen Bildfunktion entspricht, so daß Werte an den Zwischenstellen interpoliert werden können.

Nun können die Schätzungen der Ableitungen Fehler enthalten, die durch das Rauschen und die Nichtlinearitäten in der Bildfunktion entstehen. Durch die Verwendung endlicher statt infinitesimaler zeitlicher Abtastweiten ist die Vernachlässigung höherer als linearer Terme bei der Taylorentwicklung der Bildfunktion, wie sie bei der Herleitung der "motion constraint equation" angenommen wurde, im Falle von Krümmungen der Bildfunktion in Richtung der Verschiebung nicht mehr gerechtfertigt. Die Fehler, die durch diese Effekte in die diskrete Gleichung kommen, sollen nun untersucht werden. Man darf hoffen, aus solchen Analysen mehr über die Zuverlässigkeit von Verfahren zu erfahren, die die "motion constraint equation" benutzen, um Verschiebungsvektorfelder zu berechnen. Es lassen sich sogar Aussagen über die Güte der Schätzung aus den Daten selbst gewinnen.

Zunächst wird ein Operator zur Schätzung der räumlichen Ableitungen hergeleitet sowie seine Kovarianzmatrix berechnet. Dieser Operator liefert bei einem

mittelwertfreien Rauschen und einer Bildfunktion mit Variationen bis höchstens zweiten Grades im Mittel den richtigen Wert für die Ableitungen. Der Fehler, der auftritt, wenn Variationen dritten Grades nicht vernachlässigt werden können, wird ebenfalls angegeben. Danach wird eine diskrete Form der "motion constraint equation" angenommen und der systematische Fehler, der durch die Variationen zweiten Grades in die Komponente des Verschiebungsvektors in Gradientenrichtung kommt, hergeleitet. Schließlich wird noch der statistische Fehler der "motion constraint equation" angegeben.

Im folgenden ist g die gemessene Bildfunktion, die aus der unverrauschten Bildfunktion h und einem additiven Rauschanteil e besteht. Die durch die Operatoren approximierte Bildfunktion wird mit f bezeichnet.

3.3.1 Schätzung des Gradienten

Wir machen innerhalb einer quadratischen $(2k+1) \times (2k+1)$ -Umgebung um den Punkt $(0,0)$ die Annahme einer lokal bivariat quadratischen Bildfunktion $h(x,y)$,

$$h(x,y) = a_0 + a_1x + a_2y + \frac{1}{2}a_{11}x^2 + a_{12}xy + \frac{1}{2}a_{22}y^2 \quad (3.3)$$

die mit normalverteiltem, additivem, mittelwertfreiem, weißem Rauschanteil der Varianz σ^2 verfälscht wird:

$$g(x,y) = h(x,y) + e(x,y) \quad e \text{ normalverteilt} \quad (3.4)$$

Gemessen wird dann ein $n = (2k+1) \cdot (2k+1)$ dimensionaler Grauwertvektor $\mathbf{g} = (g_i) = (g(\mathbf{x}_i))$. Passen wir nun eine Approximationsfunktion

$$f_i = f(x_i, y_i) = f_0 + f_x x_i + f_y y_i + \frac{1}{2} f_{xx} x_i^2 + f_{xy} x_i y_i + \frac{1}{2} f_{yy} y_i^2 \quad (3.5)$$

so an den Grauwertvektor \mathbf{g} an, daß der mit der invertierten Kovarianzmatrix³ $\mathbf{W} = (W_{ij}) = (1/\sigma^2 \delta_{ij})$ gewichtete quadratische Gesamtfehler

$$E = (\mathbf{f} - \mathbf{g})^T \mathbf{W} (\mathbf{f} - \mathbf{g})$$

ein Minimum erreicht, so erhalten wir die Beaudet-Operatoren [Beaudet 78, Nagel 83] als Schätzer für den Parametervektor $\mathbf{p} = (f_0, f_x, f_y, f_{xx}, f_{xy}, f_{yy})^T$. Die Parameter sind nun ebenfalls normalverteilt mit dem Mittelwert

$$\bar{\mathbf{p}} = (a_0, a_1, a_2, a_{11}, a_{12}, a_{22})^T$$

³Die Kovarianzmatrix mißt die Abhängigkeiten der Rauschanteile des Grauwertvektors \mathbf{g} . Hier wird angenommen, daß die Rauschanteile an den verschiedenen Stellen des Vektors unabhängig voneinander sind (weißes Rauschen).

Geht man von der Herleitung aus, die *Nagel 83* angegeben hat, so kommt man schließlich zu folgendem Ausdruck für die Kovarianzmatrix des Parametervektors:

$$\mathbf{W}_p^{-1} = \left[\left(\frac{df}{dp} \right)^T \mathbf{W} \left(\frac{df}{dp} \right) \right]^{-1}$$

$$= \frac{\sigma^2}{n} \begin{pmatrix} \frac{14k^2+14k-3}{m} & 0 & 0 & \frac{-30}{m} & 0 & \frac{-30}{m} \\ 0 & \frac{3}{k(k+1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{k(k+1)} & 0 & 0 & 0 \\ \frac{-30}{m} & 0 & 0 & \frac{180}{k(k+1)m} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{9}{k^2(k+1)^2} & 0 \\ \frac{-30}{m} & 0 & 0 & 0 & 0 & \frac{180}{k(k+1)m} \end{pmatrix}$$

wobei gilt: $m = (2k - 1)(2k + 3) = n - 4$

In dieser Darstellung ist der Einfluß der Operatorgröße k explizit angegeben. Insbesondere gilt für die ersten Ableitungen:

$$\sigma_{f_x}^2 = \sigma_{f_y}^2 = \frac{3\sigma^2}{k(k+1)(2k+1)^2}$$

Man sieht, daß die Varianz von f_0 relativ hoch ist und sich auch bei größeren Operatorweiten leider nicht so verbessern läßt, wie es bei den anderen Parametern der Fall ist.

In der obigen Herleitung der Ableitungsoperatoren wurde angenommen, daß die unverrauschte Bildfunktion nur Terme bis zur zweiten Ordnung nach x und y enthält. Unter dieser Annahme wird das Ergebnis der Schätzung mit größer werdender Operatorweite immer besser. Je größer die Operatorweite jedoch ist, desto unwahrscheinlicher wird diese Annahme. Nehmen wir den Beaudet-Ableitungsoperator für die x -Richtung:

$$\widehat{g}_x = \frac{1}{nx^2} \sum_{i,j=-k}^k i \Delta x g(i\Delta x, j\Delta y) \tag{3.6}$$

Setzen wir in diese Gleichung eine unverrauschte Bildfunktion h mit Variationen bis zur dritten Ordnung in x und y ein, so erhalten wir den Fehlerterm:

$$\begin{aligned} \varepsilon_x &= \widehat{h}_x - h_x \\ &= \frac{1}{6} \left((\Delta x)^2 \frac{3k^2 + 3k - 1}{5} h_{xxx} + (\Delta y)^2 k(k+1)(2k+1) h_{xyy} \right) \end{aligned} \tag{3.7}$$

Diesmal erhöht sich der Fehler bei größer werdender Operatorweite. Es sollte somit eine optimale Operatorgröße geben, die einerseits den systematischen Fehler nicht zu stark werden läßt und andererseits nicht zu sehr verrauschte Ergebnisse liefert.

Bei großen Operatorweiten empfiehlt es sich, die weiter entfernten Pixel weniger stark zu wichten. Üblich ist es, dabei eine Gaußfunktion als Wichtung zu benutzen.

3.3.2 Die Zeitabtastung

Für die zeitlichen Ableitungen ist die Verwendung von Beudet-Operatoren unüblich, da meist nur jeweils zwei Bilder einer Folge zur Zeit analysiert werden. Die zeitliche Ableitung g_t wird daher häufig einfach durch die Differenz der Grauwerte zweier aufeinanderfolgender Bilder g_1 und g_2 an der gleichen Bildkoordinate approximiert. Dies birgt jedoch das Problem, daß räumliche und zeitliche Ableitungen nicht konsistent zueinander sind, da sie sich nicht auf den gleichen Bildpunkt in Raum und Zeit beziehen. Sind die beiden Bilder g_1 und g_2 in einem zeitlichen Abstand δt aufgenommen worden, so ergibt sich der Verschiebungsvektor zu $u\delta t$. Im folgenden setzen wir $\delta t = 1$, so daß wir für die Verschiebungen einfach u schreiben können. Die diskrete Form der "motion constraint equation" lautet nun:

$$g_2(x_0, y_0) - g_1(x_0, y_0) + u\widehat{g}_{1_x}(x_0, y_0) + v\widehat{g}_{1_y}(x_0, y_0) = 0 \quad (3.8)$$

Anstatt die Werte der Bildfunktion an der Stelle (x_0, y_0) zu nehmen, kann man auch den durch den Beudet-Operator \widehat{g}_0 geschätzten konstanten Anteil der Bildfunktion in den beiden Bildern verwenden, um die zeitliche Ableitung zu approximieren. Die Größen \widehat{g}_{1_x} und \widehat{g}_{1_y} seien die durch Operatoren geschätzten Ableitungen des Bildes zum Zeitpunkt t_1 . Die Fehler dieser Größen werden im folgenden nicht mehr betrachtet.

Wir nehmen nun als nächstes an, daß die unverrauschten Bilder h_1 und h_2 tatsächlich durch eine Verschiebung auseinander hervorgegangen sind:

$$h_2(x, y) = h_1(x - u, y - v)$$

Wir verwenden eine Bildfunktion h mit Variationen höchstens zweiten Grades, wie sie in Gleichung 3.3 aufgeschrieben ist. Somit gilt dann:

$$h_2(0, 0) = a_0 - a_1u - a_2v + \frac{1}{2}a_{11}u^2 + a_{12}uv + \frac{1}{2}a_{22}v^2$$

Ohne Beschränkung der Allgemeinheit drehen wir das Koordinatensystem so, daß der Verschiebungsvektor in x -Richtung zeigt, d.h. die Komponente v des Vektors verschwindet. Dies eingesetzt in die diskrete "motion constraint equation" ergibt (das Argument $(0, 0)$ lassen wir weg):

$$h_2 - h_1 + ua_1 + va_2 = \frac{1}{2}a_{11}u^2 \quad (3.9)$$

Dies ist der systematische Fehler, der entsteht, wenn man die diskrete "motion constraint equation" auf ein bivariat quadratisches Bildpaar anwendet, wobei das

zweite Bild aus dem ersten durch eine Verschiebung um den Vektor u hervorgegangen ist.

Schätzt man nun mit einer, analog zur Gleichung 3.8, diskretisierten Version der Gleichung 3.2 die Komponente u^\perp des Verschiebungsvektors u in Richtung des Gradienten, so erhält man:

$$\hat{u}^\perp = -\frac{h_2 - h_1}{|\nabla h_1|} = \frac{ua_1 + \frac{1}{2}u^2a_{11}}{|\nabla h_1|} = u^\perp + \frac{u^2a_{11}}{2|\nabla h_1|} \quad (3.10)$$

Der systematische absolute Fehler in der Schätzung der Komponente u^\perp des Verschiebungsvektors mittels der diskreten Form der Gleichung 3.2, der durch den Abbruch der Taylorreihe nach dem ersten Glied entsteht, ist proportional zu

- dem Verhältnis der zweiten Ableitung in Richtung der Verschiebung zum Betrag des Gradienten und
- dem Quadrat der Verschiebung u .

Die Inversen dieser Größen können als Gütekriterium für die geschätzte Komponente des Verschiebungsvektors in Richtung des Gradienten verwendet werden. Senkrecht zu dieser Richtung ist die Güte dann Null, weil keine Aussage über die Verschiebung in dieser Richtung gemacht werden kann.

Zu beachten bei dieser Herleitung ist, daß der Fehler, der durch eine Variation dritten Grades des Bildes entsteht, nicht berücksichtigt worden ist. Interessant ist auch, daß eine Variation zweiten Grades senkrecht zur Verschiebungsrichtung den Fehler nicht beeinflusst.

3.3.3 Zusammenfassung der Fehler

Wir wollen nun den Gesamtfehler der diskreten "motion constraint equation" untersuchen. Dazu nehmen wir eine, innerhalb einer Umgebung, die die Operatorweite plus der Verschiebung umfaßt, bivariat quadratische Bildfunktion h gemäß Gleichung 3.3 an. Die beiden zu den Zeiten t_1 und t_2 aufgenommenen Bilder g_1 und g_2 seien aus h durch eine Verschiebung um den Vektor u und einem additiven, normalverteilten Rauschanteil e_1 bzw. e_2 gewonnen. Die Varianz des Rauschens sei in beiden Fällen gleich σ^2 . Die Schätzung der Polynomfaktoren a_0, \dots, a_{22} geschehe mittels der Beaudet-Operatoren. Wir fragen nun nach der Verteilung des Fehlers ε

$$\varepsilon = \widehat{g_2}_0 - \widehat{g_1}_0 + \widehat{g_1}_x u + \widehat{g_1}_y v$$

Da dies eine Linearkombination von normalverteilten Größen ist, ist die Gesamtgröße ebenfalls wieder normalverteilt. Wir haben oben bereits den Mittelwert ausgerechnet (siehe 3.9):

$$\bar{\varepsilon} = \frac{1}{2}a_{11}u^2 + a_{12}uv + \frac{1}{2}a_{22}v^2 \quad (3.11)$$

Wir kennen auch schon die Varianzen der geschätzten Ableitungen. Diese setzen sich zur Varianz von ε folgendermaßen zusammen:

$$\sigma_\varepsilon^2 = \frac{\sigma^2}{n} \left(2 \frac{14k^2 + 14k - 3}{(2k-1)(2k+3)} + (u^2 + v^2) \frac{3}{k(k+1)} \right) \quad (3.12)$$

Die wichtigsten Ergebnisse lassen sich in folgenden Punkten zusammenfassen:

- Das Quadrat der Verschiebung geht sowohl in den systematischen als auch in den statistischen Fehler ein.
- Auch bei guter Schätzung der Gradienten und kleinen Verschiebungen kann der statistische Fehler nicht beliebig klein gemacht werden. Durch die relativ hohen Varianzen der Mittelwerte der Grauwerte bleibt eine untere Grenze bestehen.
- Die Varianz des statistischen Fehlers hängt nur von der Operatorgröße und der Rauschvarianz σ^2 ab. Von den Bildinhalten ist sie unabhängig.

Das Problem der großen Varianzen der Mittelwerte und der Konsistenz der Ableitungen in Raum und Zeit kann behoben werden, indem ein erweiterter Beaudet-Operator für alle drei Ableitungsrichtungen eingesetzt wird. Die neue diskrete "motion constraint equation" mit einem Operator der Größe $(2k+1) \times (2k+1) \times (2l+1)$ (l gibt die Größe des Operators in der Zeitrichtung an) lautet dann:

$$\widehat{g}_t(x_0, y_0) + u\widehat{g}_x(x_0, y_0) + v\widehat{g}_y(x_0, y_0) = 0$$

Der statistische Fehler dieser Gleichung hat dann die folgende Varianz σ_ε^2 :

$$\sigma_\varepsilon^2 = \frac{\sigma^2}{(2l+1)n} \left(\frac{3}{l(l+1)} + (u^2 + v^2) \frac{3}{k(k+1)} \right)$$

Der Aufwand bei der Faltung mit einer dreidimensionalen Operatormaske kann durch Separierung reduziert werden. Ein separierbarer Operator $O_{i,j,k}$ kann in drei Teile aufgespalten werden, die jeweils nur von einem der Indizes abhängen: $O_{r,c,t} = o_r p_c q_t$. Die Summation kann dann aufgespalten werden in drei Summen:

$$\sum_{r,c,t} O_{r,c,t} g_{r,c,t} = \sum_{r=-k}^k o_r \left(\sum_{c=-k}^k p_c \left(\sum_{t=-l}^l q_t g_{r,c,t} \right) \right)$$

Zur Faltung eines Bildes mit einem Operator benötigt man normalerweise $(2k+1)^2(2l+1)$ Multiplikationen pro Pixel, während mit Separierung nur $(4k+2l+3)$ Multiplikationen anfallen.⁴

⁴Formeln ohne Berücksichtigung von Effekten am Bildrand

Um den Fehler, der durch das Quadrat des Verschiebungsvektors entsteht, möglichst klein zu halten, kann man eine allgemeinere iterative Variante der "motion constraint equation" benutzen. Nehmen wir an, wir hätten schon eine Schätzung \mathbf{u}_0 des wahren Verschiebungsvektors \mathbf{u} . Die Aufgabe ist es jetzt einen Korrekturvektor $d\mathbf{u}$ zu finden, so daß gilt: $\mathbf{u} = \mathbf{u}_0 + d\mathbf{u}$. Mit den obigen Annahmen ergibt der geschätzte Vektor \mathbf{u}_0 eingesetzt in die "motion constraint equation" den Fehler:

$$h_2 - h_1 + u_0 a_1 + v_0 a_2 = -du a_1 - dv a_2 + \frac{1}{2} a_{11} u^2 + a_{12} uv + \frac{1}{2} a_{22} v^2$$

Aus dieser Gleichung können wir durch Umstellen der Terme und Ersetzung von \mathbf{u} durch \mathbf{u}_0 sofort die gesuchte verallgemeinerte "motion constraint equation" ablesen:

$$(g_2 - g_1 + (\nabla g_1)^T \mathbf{u}_0 - \frac{1}{2} \mathbf{u}_0^T \mathbf{H}_1 \mathbf{u}_0 + (\nabla g_1 - \mathbf{H}_1 \mathbf{u}_0)^T d\mathbf{u} = 0 \quad (3.13)$$

Hierbei ist $\mathbf{H}_1 = \nabla(\nabla g_1)^T$ die Hessematrix der Bildfunktion g_1 (siehe Anhang A). Die Schätzung \mathbf{u}_0 kann man aus den Verschiebungsvektoren des vorherigen Bildes der Bildfolge gewinnen oder aus vorangegangenen Berechnungen im gleichen Bild. Für $\mathbf{u}_0 = \mathbf{0}$ erhalten wir die alte Gleichung zurück. Eine Verbesserung des Ergebnisses ist nur zu erwarten, falls $|d\mathbf{u}| < |\mathbf{u}_0|$ ist. Kearney u.a. 87 verwenden eine Variante dieser Gleichung ohne Berücksichtigung der zweiten Ableitungen. Unter den hier gemachten Voraussetzungen ist dann jedoch keine Verbesserung des Ergebnisses zu erwarten. Es besteht allerdings die Möglichkeit, dieses Verfahren mit einer groben Auflösung zu beginnen, und dann iterativ auf immer feiner werdende Auflösungsebenen herabzusteigen. Diese Technik der hierarchischen Bildanalyse wird auch bei anderen Algorithmen häufig eingesetzt und ist unter dem Stichwort "Bildpyramide" bekannt.

3.4 Herleitung als Optimierungsproblem

Im nächsten Kapitel werden wir viele der Gleichungen aus Optimierungsproblemen gewinnen. Auch die "motion constraint equation" kann aus einem Optimierungsansatz gewonnen werden.

Wir gehen von zwei aufeinanderfolgenden Bildern g_1 und g_2 aus. Ein guter Verschiebungsvektor \mathbf{u} an einem Punkt \mathbf{x}_0 sollte die quadrierte Grauwertdifferenz minimieren.

$$M(\mathbf{u}) = (g_2(\mathbf{x}_0) - g_1(\mathbf{x}_0 - \mathbf{u}))^2 \longrightarrow \text{Minimum} \quad (3.14)$$

Um die Abhängigkeit von \mathbf{u} explizit zu machen, entwickeln wir g_1 um \mathbf{x}_0 in eine Taylorreihe erster Ordnung. Im folgenden lassen wir das Argument \mathbf{x}_0 der Kürze halber weg.

$$M(\mathbf{u}) \approx (g_2 - g_1 + (\nabla g_1)^T \mathbf{u})^2 \quad (3.15)$$

Setzen wir die Ableitung dieser Approximation zu Null, so erhalten wir die notwendige Bedingung

$$\nabla g_1(g_2 - g_1 + (\nabla g_1)^T \mathbf{u}) = \mathbf{0} \quad (3.16)$$

Ist der Gradient $\nabla g_1(\mathbf{x}_0)$ von Null verschieden und setzen wir $g_2 - g_1 = g_t$, so haben wir die "motion constraint equation" 3.1 vor uns.

Man kann ein Optimum von M auch durch ein Gradientenabstiegsverfahren berechnen. Daraus ergibt sich die "pel recursive gradient technique" von *Robbins + Neuvial* 83. Dieses Verfahren ist eng verwandt mit der iterativen Verbesserung der Verschiebungsvektoren, die im vorigen Abschnitt erwähnt wurde. Die allgemeine Iterationsformel für den Gradientenabstieg lautet [*Duda + Hart* 73, S. 140]:

$$\begin{aligned} \mathbf{u}_{k+1} &= \mathbf{u}_k - \rho_k \nabla_{\mathbf{u}} M(\mathbf{u}_k) \\ &= \mathbf{u}_k + 2\rho_k (g_2(\mathbf{x}_0) - g_1(\mathbf{x}_0 - \mathbf{u}_k)) \nabla_{\mathbf{u}} g_1(\mathbf{x}_0 - \mathbf{u}_k) \end{aligned}$$

Dabei ist \mathbf{u}_k der geschätzte Verschiebungsvektor nach dem k -ten Schritt. Die Schrittweite pro Iteration läßt sich durch den Parameter ρ_k einstellen. Ein großes ρ_k bewirkt eine schnelle Annäherung an das Minimum, kann allerdings auch dazu führen, daß man über das Ziel hinausschießt.

Bei Annahme einer linearen Bildfunktion können wir einen Zusammenhang zur "motion constraint equation" herstellen:

$$\begin{aligned} \mathbf{u}_{k+1} &= \mathbf{u}_k - 2\rho_k (g_2 - g_1 + (\nabla g_1)^T \mathbf{u}_k) \nabla g_1 \\ &= \mathbf{u}_k - 2\rho_k (g_t \nabla g_1 + (\nabla g_1)(\nabla g_1)^T \mathbf{u}_k) \end{aligned}$$

Und mit konstantem $\rho_k = \rho$ sowie dem Startwert $\mathbf{u}_0 = \mathbf{0}$ erhalten wir:

$$\begin{aligned} \mathbf{u}_1 &= -2\rho g_t \nabla g_1 \\ \mathbf{u}_2 &= -2\rho g_t \nabla g_1 - 2\rho (g_t \nabla g_1 - 2\rho g_t (\nabla g_1)(\nabla g_1)^T \nabla g_1) \\ &= -4\rho g_t \nabla g_1 (1 - \rho (\nabla g_1)^T \nabla g_1) \\ &= -4\rho g_t \nabla g_1 (1 - \rho |\nabla g_1|^2) \end{aligned}$$

Bei $\rho = 1/(2|\nabla g_1|^2)$ verändert sich der Wert von \mathbf{u}_k ab $k = 1$ nicht mehr. Mit diesem Wert erhält man schließlich die Komponente des Verschiebungsvektors in Richtung des Gradienten zu (Vergleiche mit Gleichung 3.10):

$$\mathbf{u}_1 = -\frac{g_t}{|\nabla g_1|^2} \nabla g_1 = \mathbf{u}^\perp$$

Die lokale Autokorrelationsfunktion gibt einen Eindruck von der Form von M . Sei $R_{g_i^2}$ die Autokorrelation von g_i und $R_{g_i g_j}$ die Kreuzkorrelation von g_i und g_j :

$$\begin{aligned} R_{g_i^2}(\mathbf{z}) &= E[g_i(\mathbf{x})g_i(\mathbf{x} - \mathbf{z})] \\ R_{g_i g_j}(\mathbf{z}) &= E[g_i(\mathbf{x})g_j(\mathbf{x} - \mathbf{z})] \end{aligned}$$

Hierbei bezeichnet $E[\cdot]$ den Erwartungswert der entsprechenden Größe. Es folgt somit:

$$E[M(\mathbf{u})] = R_{g_1^2}(\mathbf{0}) + R_{g_2^2}(\mathbf{0}) - 2R_{g_1g_2}(\mathbf{u})$$

Wollen wir also M minimieren, so müssen wir die Kreuzkorrelation $R_{g_1g_2}$ maximieren. Rechnen wir den Gradienten des lokalen Mittelwertes von M aus

$$\begin{aligned} \nabla_{\mathbf{u}} E[M(\mathbf{u})] &= -2\nabla_{\mathbf{u}} R_{g_1g_2} \\ &= -2\nabla_{\mathbf{u}} E[g_1(\mathbf{x} - \mathbf{u})g_2(\mathbf{x})] \\ &= -2E[\nabla_{\mathbf{u}} g_1(\mathbf{x} - \mathbf{u})g_2(\mathbf{x})] \\ &= 2R_{\nabla_{\mathbf{u}}g_1g_2}(\mathbf{u}) \end{aligned}$$

und setzen diesen in das Gradientenabstiegsverfahren ein, so ergibt sich ein Verfahren von *Bergmann 83*:

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \rho_k \nabla M(\mathbf{u}_k) \approx \mathbf{u}_k - 2\rho_k R_{\nabla_{\mathbf{u}}g_1g_2}(\mathbf{u}_k)$$

Mit der Anfangsschätzung von $\mathbf{u}_0 = \mathbf{0}$ erhalten wir somit:

$$\mathbf{u}_1 = -2\rho_k E[\nabla_{\mathbf{u}} g_1g_2]$$

Das Konvergenzverhalten dieser Methode läßt sich durch Anpassung der Schrittweite ρ_k an die Krümmung der Kreuzkorrelationsfunktion noch verbessern.

3.5 Betrachtungen im Frequenzraum

Um die "motion constraint equation" zu Arbeiten in Beziehung zu setzen, die versuchen, Geschwindigkeiten mittels räumlich-zeitlicher Filterung zu schätzen, ist es aufschlußreich, sie in den Frequenzraum zu transformieren.

Sei $G(\mathbf{k}, \omega)$ die Fouriertransformierte der Bildfolge $g(\mathbf{x}, t)$. Dabei ist \mathbf{k} die zweidimensionale Ortsfrequenz und ω die Zeitfrequenz. Es gilt für die partiellen Ableitungen:

$$\begin{aligned} \mathcal{F}\{g_x(\mathbf{x}, t)\} &= ik_1 G(\mathbf{k}, \omega) \\ \mathcal{F}\{g_y(\mathbf{x}, t)\} &= ik_2 G(\mathbf{k}, \omega) \\ \mathcal{F}\{g_t(\mathbf{x}, t)\} &= i\omega G(\mathbf{k}, \omega) \end{aligned}$$

Damit erhalten wir die "motion constraint equation" im Frequenzraum:

$$\mathbf{k}\mathbf{u} + \omega = 0 \quad (3.17)$$

Jede Geschwindigkeit \mathbf{u} definiert im Frequenzraum eine Ebene durch den Nullpunkt mit dem Normalenvektor $(u, v, 1)^T$. Der Tangens des Winkels zwischen

Normalenvektor und ω -Achse entspricht dem Geschwindigkeitsbetrag $\sqrt{u^2 + v^2}$. Der Winkel zwischen der Projektion dieses Vektors auf die \mathbf{k} -Ebene und der k_x -Achse entspricht der Geschwindigkeitsrichtung β .

Die gleiche Ebene erhält man auch, wenn man eine Bildfolge $g(\mathbf{x}, t)$ mit der Eigenschaft:

$$g(\mathbf{x}, t) = g_0(\mathbf{x} - \mathbf{u}t) \quad (3.18)$$

in den Frequenzraum transformiert. Dazu sei $G_0(\mathbf{k})$ die Fouriertransformierte von $g_0(\mathbf{x})$; dann folgt:

$$\begin{aligned} \mathcal{F}\{g(\mathbf{x}, t)\} &= \int g_0(\mathbf{x} - \mathbf{u}t) e^{-2\pi i(\mathbf{k}\mathbf{x} + \omega t)} d\mathbf{x} dt \\ &= \int e^{-2\pi i(\mathbf{k}\mathbf{u} + \omega)t} dt \int g_0(\mathbf{z}) e^{-2\pi i\mathbf{k}\mathbf{z}} d\mathbf{z} \\ &= \delta(\mathbf{k}\mathbf{u} + \omega) G_0(\mathbf{k}) \end{aligned}$$

Verfahren, die mit der "motion constraint equation" arbeiten, schätzen durch Kombination dreier Richtungsableitungen entlang der beiden Raumachsen und der Zeitachse den Gradienten (g_x, g_y, g_t) . Für die Bestimmung der Verschiebung wird allerdings nur die Richtung des Gradienten benötigt. Der Kontrast selber ist unerheblich (außer als Gütekriterium). Die Gradientenrichtung entspricht einer Geraden im Frequenzraum, d.h. einem konstanten Verhältnis der drei Frequenzen. Ein einziger solcher Gradient erlaubt also immer noch unendlich viele Verschiebungsvektoren, denn durch eine Gerade können unendlich viele Ebenen gelegt werden.

Nimmt man an, daß zwei Punkte mit verschiedenen Gradientenrichtungen den gleichen Verschiebungsvektor haben, so spannen die beiden entsprechenden Geraden im Frequenzraum die gesuchte Ebene auf. Genauere Analysen dieses Vorgehens findet man im nächsten Kapitel.

Wie bereits erwähnt, ist die Neigung der Ebene ein Maß für den Geschwindigkeitsbetrag. Die Neigung einer Geraden ist ein Maß für die Komponente der Geschwindigkeit in der Richtung, die durch die Projektion der Geraden auf die \mathbf{k} -Ebene bestimmt ist. An Stellen mit hoher Ortsfrequenz, also scharfen Kanten, läßt sich zwar die Neigung der Ebene, und damit die Geschwindigkeit, am genauesten bestimmen; da aber der Frequenzraum im diskreten Fall nur bis zu einer endlichen Frequenz reicht, kann man eine zu starke Neigung der Ebene, also eine zu hohe Geschwindigkeit, nicht mehr registrieren.

Um eine bessere Schätzung der dreidimensionalen Gradientenrichtung zu erreichen, kann man auch in mehr als nur drei Richtungen ableiten. Außerdem kann man darauf verzichten, den Kontrast mitzubestimmen, d.h. man braucht die absolute Gradientenstärke nicht zu kennen. Dazu braucht man allerdings dreidimensionale Filter. Ein solcher Filter sollte folgende Eigenschaften haben:

- gute Lokalisation im Orts-Zeit-Raum.
- Trennschärfe bezüglich der Gradientenrichtung.

Der Filter sollte also im Frequenzraum einer Geraden durch den Nullpunkt entsprechen. *Fleet + Jepson 85* versuchen, eine Filter-Kaskade auf diese Gesichtspunkte hin zu optimieren. Das Ergebnis sind Filter, die ihre größte Amplitude innerhalb eines Kegels im Frequenzraum haben.

Diese Filterungen können als Vorverarbeitung der Bildfolge angesehen werden. Auf dieser niedrigen Stufe der Verarbeitung werden noch keine Entscheidungen getroffen und alle Informationen, die die ursprüngliche Bildfolge enthält, bleiben im Prinzip erhalten. Das Ziel der Vorverarbeitung ist es, den weiteren Algorithmen Daten zu liefern, die diese dann direkter für ihre Zwecke verwenden können als die Grauwerte selber.

Heeger stellt einen Algorithmus vor, der mit den Ergebnissen ähnlicher Filterungen eine Ebene im Frequenzraum "optimal" anpaßt [*Heeger 87, Heeger 88*]. Er filtert die Bildfolge mit einer Anordnung von Gaborfiltern der Form:

$$f(\mathbf{x}, t) = e^{-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} + \frac{t^2}{2\sigma_t^2}\right)} \sin(2\pi k_{x0}x + 2\pi k_{y0}y + 2\pi\omega_{t0}t) \quad (3.19)$$

Ein Vorteil dieser Filter ist, daß sie im Frequenzraum eine einfache Form haben. Bei der Transformation in den Frequenzraum ergibt sich nämlich ein Paar von 3D-Gaußfunktionen mit den Mittelpunkten $(k_{x0}, k_{y0}, \omega_{t0})^T$ und $(-k_{x0}, -k_{y0}, -\omega_{t0})^T$. Der Filter ergibt ein maximales Ergebnis für eine bewegte Sinuskurve, deren Frequenz dem Mittelpunkt $(k_{x0}, k_{y0}, \omega_{t0})$ entspricht. Ein weiterer von Heeger genannter Vorteil dieser Filter ist, daß sie separierbar sind.

Zwölf dieser Filter sind im Frequenzraum so angeordnet, daß ihre Mittelpunkte auf einem Zylinder entlang der ω -Achse liegen. Um die Operatorsignale mit den Verschiebungsvektoren in Beziehung zu setzen, wird das Parseval Theorem benutzt, welches hier besagt, daß das Integral der quadrierten Operatorsignale über dem Orts-Zeit-Raum (die Gaborenergie E) gleich ist dem Integral der quadrierten Fourierkomponenten über dem Frequenzraum. Sei also F die Fouriertransformierte der Gaborfunktion, so gilt:

$$E = \int |f(\mathbf{x}, t) * g(\mathbf{x}, t)|^2 dxdt = \int |F(\mathbf{k}, \omega) \cdot G(\mathbf{k}, \omega)|^2 d\mathbf{k}d\omega \quad (3.20)$$

Nimmt man als Modell eine Bildfolge einer bewegten zufällig verteilten Textur an, die der Eigenschaft 3.18 genügt, so erhält man im Frequenzraum eine Ebene von konstanter, nur vom Kontrast abhängiger Dichte ρ . Somit läßt sich die Gaborenergie durch Integration im Frequenzraum bestimmen. Man erhält als Ergebnis für jeden der 12 Filter:⁵

$$E(u, v, \rho) = \frac{\rho^2}{8\pi\sqrt{\det \mathbf{A}}} e^{-4\pi\sigma_t^2\sigma_x^2\sigma_y^2 \frac{(k_{x0}u + k_{y0}v + \omega_{t0})^2}{\det \mathbf{A}}}$$

$$\det \mathbf{A} = (\sigma_x\sigma_y)^2 + (\sigma_t\sigma_y u)^2 + (\sigma_t\sigma_x v)^2$$

⁵Nach Berichtigung eines Fehlers in der Arbeit von Heeger

Diese Gleichung ist nur noch abhängig vom Verschiebungsvektor u und von der Dichte ρ . Die Dichte kann man aus der Gleichung eliminieren, indem man durch die mittlere Gaborenergie aller Filter teilt. Wir erhalten ein Gleichungssystem aus 12 nichtlinearen Gleichungen. Durch Minimierung des quadratischen Fehlers mit einem numerischen Gradientenabstiegsverfahren wird schließlich der Verschiebungsvektor bestimmt.

Dieses Verfahren hat den Vorteil, daß die gesamte Information der Filterung zur Schätzung der Verschiebung ausgenutzt wird, und nicht nur derjenige Filter mit maximalem Ergebniswert. Der Nachteil durch die Annahme einer zufallsverteilten Textur mit konstantem Kontrast in allen räumlichen Richtungen kann dadurch abgemildert werden, daß man die Filter in Gruppen einteilt, die in verschiedenen Richtungen wirken, und dann aus jeder Gruppe einzeln die Dichte eliminiert.

Kapitel 4

Methoden zur Schätzung der zweiten Komponente

Um die vollständigen Verschiebungsvektoren berechnen zu können, also das Blendenproblem zu lösen, braucht man zusätzlich zur "motion constraint equation" noch eine weitere Nebenbedingung. Dazu gibt es in der Literatur die verschiedensten Vorschläge. Man kann diese in mehrere Gruppen einteilen:

1. Zusätzliches geometrisches Wissen (z.B. epipolare Geometrie bei Stereoaufnahmen).
2. Erweiterung der "motion constraint equation" für nicht-lineare Bildfunktionen.
3. Annahme eines konstanten Verschiebungsvektors innerhalb einer lokalen Umgebung; Ballungsanalyse im (u, v) -Raum.
4. Annahme eines möglichst glatten Verschiebungsvektorfeldes.

Im folgenden werden für jede Klasse einige der verschiedenen Ansätze näher untersucht. Dabei sollen die zugrundeliegenden Annahmen und die charakteristischen Schwierigkeiten der einzelnen Methoden herausgearbeitet werden. Dazu werden die Zusammenhänge der Verfahren erläutert und zum Teil andere Herleitungen angegeben als sie in der Originalliteratur zu finden sind. Um einen Vergleich der verschiedenen Gleichungen zu ermöglichen, wird eine einheitliche Notation verwendet, die im Anhang A zusammengefaßt ist.

4.1 Zusätzliches Wissen

Hier soll nur kurz gezeigt werden, wie es möglich ist, zusätzliches Vorwissen über die betrachtete Szene in den Algorithmus zur Schätzung des Verschiebungsvektorfeldes einzubringen. Im allgemeinen muß die Zusatzinformation in Form einer mathematischen Nebenbedingung ausgedrückt werden.

Wenn wir z.B. eine rein translatorische Bewegung vor uns haben und den Expansionspunkt bereits kennen, brauchen wir nur noch den Betrag der Geschwindigkeit zu schätzen, denn alle Verschiebungsvektoren müssen sich im Expansionspunkt schneiden.

Ein weiteres Beispiel ergibt sich aus der Annahme, daß die Oberflächen der betrachteten Objekte eine bestimmte Form haben, z.B. ebene Flächen sind. In diesem Fall kann der gesamte Bewegungsfluß, der durch eine perspektivische Projektion entsteht, durch 8 Parameter beschrieben werden:

$$\begin{aligned} u(x, y) &= u_0 + Ax + By + (Ex + Fy)x \\ v(x, y) &= v_0 + Cx + Dy + (Ex + Fy)y \end{aligned} \quad (4.1)$$

Gemessen wird mit der "motion constraint equation" der Normalanteil dieser Geschwindigkeiten. Aus mindestens 8 dieser Gleichungen kann man die 8 Parameter des Bewegungsflusses bestimmen, wenn die Gradientenrichtungen genügend variieren. Waxman + Wohn 84 nennen diese Methode der Bestimmung von Flußparametern die "velocity functional method".

Der größte Nachteil aller Verfahren, die zusätzliches Wissen über die abgebildeten Objekte oder deren Bewegung einbringen, ist die stark eingeschränkte Anwendbarkeit. Falls die Anwendung jedoch möglich ist, z.B. in einer industriellen Umgebung, kann man gute Ergebnisse damit erzielen.

4.2 Erweiterung der "motion constraint equation"

In diesem Abschnitt sollen Erweiterungen der "motion constraint equation" diskutiert werden, die Variationen zweiten Grades in der Bildfunktion berücksichtigen.

Am besten lassen sich diese Gleichungen aus dem Optimierungsansatz 3.14 des letzten Kapitels gewinnen.

$$M = (g_2(\mathbf{x}_0) - g_1(\mathbf{x}_0 - \mathbf{u}))^2 \rightarrow \text{Minimum}$$

Statt die Taylorreihe nach dem ersten Glied abzuberechnen, wie in Gleichung 3.15, können wir auch noch die quadratischen Terme hinzunehmen.

$$M \approx \left(g_2 - g_1 + (\nabla g_1)^T \mathbf{u} - \frac{1}{2} \mathbf{u}^T \mathbf{H}_1 \mathbf{u} \right)^2$$

Setzen wir die Ableitung nach \mathbf{u} auf Null, so erhalten wir die Gleichung:

$$(\nabla g_1 - \mathbf{H}_1 \mathbf{u}) \left[g_2 - g_1 + \left((\nabla g_1)^T - \frac{1}{2} \mathbf{u}^T \mathbf{H}_1 \right) \mathbf{u} \right] = 0 \quad (4.2)$$

Man vergleiche dieses Ergebnis mit Gleichung 3.16, welche durch die lineare Approximation der Bildfunktion entstanden ist. Der Term in der eckigen Klammer ist eine Approximation des Grauwertes des ersten Bildes an der Stelle $\mathbf{x} - \mathbf{u}$, während

es sich beim ersten Term um eine Approximation des Gradienten von Bild 1 an der gleichen Stelle handelt. Dadurch wird der in Gleichung 3.9 berechnete Fehler vermieden. Man vergleiche die Gleichung auch mit der verallgemeinerten, iterativen "motion constraint equation" von Gleichung 3.13. Leider ist diese Gleichung nun nicht mehr linear, so daß sich eine Lösung nicht so einfach gewinnen läßt.

Nagel 83 geht noch einen Schritt weiter und mittelt den Term M über alle Pixel innerhalb der Umgebung, in der die ersten und zweiten Ableitungen geschätzt werden. Hier wird also auch die Annahme eines lokal konstanten Verschiebungsvektors gemacht. Es gibt daher eine Überschneidung mit Abschnitt 4.3. Vieles, was dort noch weiter analysiert werden wird, gilt auch für das nun folgende Ergebnis. Die Summation soll im folgenden immer über alle Pixel \mathbf{x}_i einer lokalen Umgebung um das Pixel $(0,0)$ gelten. Das Argument $(0,0)$ lassen wir im folgenden weg.

$$M \approx \sum \left(g2(\mathbf{x}_i) - g1 - (\nabla g1)^T(\mathbf{x}_i - \mathbf{u}) - \frac{1}{2}(\mathbf{x}_i - \mathbf{u})^T \mathbf{H}_1(\mathbf{x}_i - \mathbf{u}) \right)^2$$

Als notwendige Bedingung für das Optimum erhalten wir nun:

$$\begin{aligned} \sum (\nabla g1 + \mathbf{H}_1(\mathbf{x}_i - \mathbf{u})) \\ \cdot \left(g2(\mathbf{x}_i) - g1 - (\nabla g1)^T(\mathbf{x}_i - \mathbf{u}) - \frac{1}{2}(\mathbf{x}_i - \mathbf{u})^T \mathbf{H}_1(\mathbf{x}_i - \mathbf{u}) \right) = 0 \end{aligned}$$

Setzen wir ein quadratisches Fenster mit n Pixeln um den Punkt $\mathbf{x}_0 = \mathbf{0}$ voraus, so verschwinden alle Summen, die nur ungerade Potenzen von \mathbf{x}_i oder \mathbf{y}_i enthalten. Weiterhin folgt:

$$\begin{aligned} \frac{1}{n} \sum \left(g1 + (\nabla g1)^T \mathbf{x}_i + \frac{1}{2} \mathbf{x}_i^T \mathbf{H}_1 \mathbf{x}_i \right) &= g1 + \frac{1}{2n} \sum \mathbf{x}_i^T \mathbf{H}_1 \mathbf{x}_i \\ &= g1 + \frac{1}{2} \overline{\mathbf{x}^2} (g1_{xx} + g1_{yy}) \\ &= \overline{g1} \end{aligned}$$

$$\sum (\mathbf{u}^T \mathbf{H}_1 \mathbf{x}_i) \mathbf{x}_i = \overline{\mathbf{x}^2} \mathbf{H}_1 \mathbf{u}$$

$$\sum ((\nabla g1)^T \mathbf{x}_i) \mathbf{x}_i = \overline{\mathbf{x}^2} \nabla g1$$

Mit diesen Ergebnissen erhalten wir schließlich die Gleichung

$$\begin{aligned} (\nabla g1 - \mathbf{H}_1 g1)^T \mathbf{u} \left(\overline{g2} - \overline{g1} + (\nabla g1)^T \mathbf{u} - \frac{1}{2} \mathbf{u}^T \mathbf{H}_1 \mathbf{u} \right) \\ + \mathbf{H}_1 \left(\overline{\mathbf{x}^2} (-\nabla g1 + \mathbf{H}_1 \mathbf{u}) + \frac{1}{n} \sum g2(\mathbf{x}_i) \mathbf{x}_i \right) = 0 \end{aligned}$$

Bis zu diesem Punkt ist das Ergebnis unabhängig von den Operatoren, mit denen wir die Ableitungen der Bildfunktion schätzen. Wir können aber noch die spezielle

Form des Beaudet-Operators für die Schätzung der ersten Ableitung ausnutzen: Ein Beaudet-Operator mit ungerader Seitenlänge [Beaudet 78, Nagel 83] approximiert die erste Ableitung durch die Summe

$$\bar{x}^2 \widehat{\nabla} g = \frac{1}{n} \sum g(\mathbf{x}_i) \mathbf{x}_i$$

Mit dieser Eigenschaft der Beaudet-Operatoren können wir dann schreiben:

$$\begin{aligned} (\nabla g_1 - \mathbf{H}_1 \mathbf{u}) \left(\bar{g}_2 - \bar{g}_1 + \left((\nabla g_1)^T - \frac{1}{2} \mathbf{u}^T \mathbf{H}_1 \right) \mathbf{u} \right) \\ = -\bar{x}^2 \mathbf{H}_1 \left(\widehat{\nabla} g_2 - \nabla g_1 + \mathbf{H}_1 \mathbf{u} \right) \end{aligned} \quad (4.3)$$

Zusätzlich zu den Korrekturen der Gradienten an der Stelle $\mathbf{x}_0 - \mathbf{u}$ im ersten Bild, wie sie auch in der Gleichung 4.2 auftreten, bekommen wir noch eine zweite Korrektur, die eine Approximation der ersten Ableitung der "motion constraint equation" $\nabla g_t + \mathbf{H} \mathbf{u} = \mathbf{0}$ enthält, und damit die Übereinstimmung der Gradienten der beiden Bilder mitberücksichtigt.

Ein wesentlicher Nachteil dieser Gleichung ist die Nichtlinearität. Trotzdem läßt sich an speziellen wichtigen Punkten im Bild eine Lösung direkt angeben. Diese Punkte lassen sich durch folgende Bedingung ausdrücken:

$$\begin{aligned} g_{1_x} = \text{Extremum} \neq 0 \quad \text{und} \quad g_{1_y} = 0 \\ g_{1_{xx}} = g_{1_{xy}} = 0 \quad \text{sowie} \quad g_{1_{yy}} = \text{Extremum} \neq 0 \end{aligned} \quad (4.4)$$

Dies ist das Modell einer Grauwertecke, wie es in Nagel 83 angegeben ist. Die Bedingungen sind z.B. auf einem geeigneten Zylindermantel erfüllt. Man beachte auch, daß sich das Koordinatensystem immer so legen läßt, daß g_{xy} verschwindet. Diese Bedingung stellt also keine prinzipielle Einschränkung dar. Wie man leicht nachrechnet, läßt sich für dieses Eckenmodell folgende Lösung der Gleichung 4.3 angeben:

$$\begin{aligned} v &= -\frac{g_{1_{yy}}^2}{g_{1_{yy}}} \\ u &= -\frac{1}{g_{1_x}} \left(\bar{g}_2 - \bar{g}_1 - \frac{1}{2} g_{1_{yy}} v^2 \right) \end{aligned} \quad (4.5)$$

4.3 Annahme eines lokal konstanten Verschiebungsvektors

4.3.1 Erste Fehleranalyse

Nimmt man an, daß alle Verschiebungsvektoren innerhalb einer Umgebung gleich sind, so kann man die "motion constraint equation" an zwei Stellen simultan lösen und dadurch den vollständigen Verschiebungsvektor berechnen:

$$\mathbf{u}^T \nabla g = -g_t \quad (4.6)$$

Der Vektor \mathbf{g} enthält als Komponenten die Grauwerte an den beiden Stellen, an denen der Verschiebungsvektor den gleichen Wert haben soll. Dieses Gleichungssystem setzt voraus, daß die einzelnen Nebenbedingungen linear unabhängig sind, d.h. keine parallelen Geraden im (u, v) -Raum ergeben. Dies ist äquivalent zu einer Variation der Gradientenrichtung. Hier wird ein gewisses Dilemma deutlich, denn wie wir oben bei der Analyse der "motion constraint equation" sahen, kann eine Variation des Gradienten einen Fehler in den einzelnen Gleichungen hervorrufen. Um dieses Dilemma quantitativ zu fassen, müssen wir die Konditionierung des Gleichungssystems quantitativ messen. Für ihre Untersuchung wählen *Kearney u.a. 87* das folgende Maß:

$$\text{cond}(\mathbf{M}) = \|\mathbf{M}\| \|\mathbf{M}^{-1}\|$$

Wobei \mathbf{M} die Matrix des Gleichungssystems und $\|\cdot\|$ die Frobenius-Norm mit

$$\|(m_{ij})\| = \sqrt{\sum_{i,j} m_{ij}^2}$$

ist. Je größer $\text{cond}(\mathbf{M})$ ist, desto mehr werden kleine Fehler in den einzelnen Gleichungen das Ergebnis beeinflussen.

Mit diesen Festlegungen erhält man als Ergebnis für eine Umgebung von zwei Punkten:

$$\text{cond}(\nabla \mathbf{g}) = \frac{\|\nabla g_1\|^2 + \|\nabla g_2\|^2}{\|\nabla g_1\| \|\nabla g_2\| \sin \phi} = \frac{1}{\sin \phi} \left(\frac{\|\nabla g_1\|}{\|\nabla g_2\|} + \frac{\|\nabla g_2\|}{\|\nabla g_1\|} \right) \quad (4.7)$$

Dabei stehen die Indizes für die beiden Orte, an denen man die gleiche Verschiebung ansetzt. ϕ ist der Winkel zwischen den einzelnen Gradienten. Die Konditionierung des Gleichungssystem wird optimiert, wenn die Gradienten senkrecht aufeinander stehen und den gleichen Betrag haben. Wie man sieht, hat der Winkel ϕ einen weit größeren Einfluß auf die Konditionierung als der Betrag der Gradienten.

4.3.2 Least Square Fit

Will man mehr als zwei Orte zur Schätzung des Verschiebungsvektors berücksichtigen, so hat man ein überbestimmtes Gleichungssystem in der Form 4.6 zu lösen.

Anschaulich sucht man nach einem Punkt im (u, v) -Raum, an dem sich alle beteiligten "constraint lines" möglichst gut schneiden (siehe Abbildung 4.1). Dazu kann man z.B. eine Ballungsanalyse vornehmen. Verfahren dazu sind eine modifizierte Hough-Transformation [*Bandyopadyhay 84*] oder auch der Algorithmus von *Schunck 84a*. Vorteile dieser Verfahren sind, daß eventuelle Ausreißer unberücksichtigt bleiben. Dies ist besonders an Verdeckungskanten wichtig.

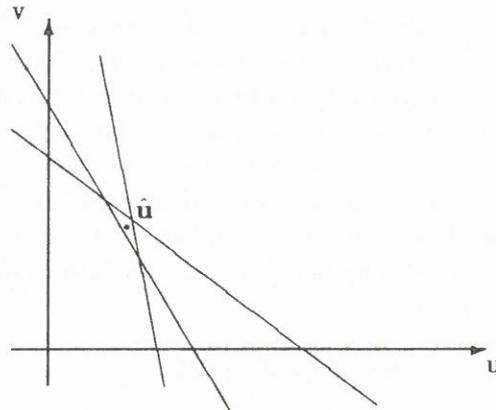


Abbildung 4.1: Optimaler "Schnittpunkt" der "constraint lines".

Wir nehmen im folgenden aber an, daß alle Punkte einer Umgebung den gleichen Verschiebungsvektor haben. Um den Verschiebungsvektor zu finden, können wir versuchen, die Summe der Fehlerquadrate zu minimieren:

$$M = (\mathbf{u}^T \nabla \mathbf{g} + \mathbf{g}_t)^2 \rightarrow \text{Minimum}$$

Der Vektor \mathbf{g} enthält $n > 2$ Komponenten.

Nach Nullsetzen der Ableitung von M nach \mathbf{u} erhält man

$$\nabla \mathbf{g} (\nabla \mathbf{g})^T \hat{\mathbf{u}} = -\nabla \mathbf{g} \mathbf{g}_t^T \quad (4.8)$$

bzw.

$$\hat{\mathbf{u}} = -(\nabla \mathbf{g} (\nabla \mathbf{g})^T)^{-1} \nabla \mathbf{g} \mathbf{g}_t^T = -(\nabla \mathbf{g})^\dagger \mathbf{g}_t^T \quad (4.9)$$

Man nennt $(\nabla \mathbf{g})^\dagger = (\nabla \mathbf{g} (\nabla \mathbf{g})^T)^{-1} \nabla \mathbf{g}$ auch Pseudoinverse von $\nabla \mathbf{g}$.

Der Vektor $\hat{\mathbf{u}}^T \widehat{\nabla \mathbf{g}} + \hat{\mathbf{g}}_t$ wird im allgemeinen einen vom Nullvektor verschiedenen Wert aufweisen (siehe dazu die Analysen im vorherigen Kapitel, insbesondere die Gleichungen 3.11 und 3.12). Wir werden diesen Vektor "Residuumsvektor" nennen und ihn mit dem Buchstaben \mathbf{r} bezeichnen. Sei \mathbf{r} normalverteilt mit dem Mittelwert $\mathbf{0}$ und der Kovarianzmatrix $\mathbf{S} = \sigma^2 \mathbf{I}$, wobei \mathbf{I} die Einheitsmatrix sei. Das bedeutet, daß wir annehmen, es gebe keine systematischen, sondern nur statistische Fehler. Der nach Gleichung 4.9 geschätzte Verschiebungsvektor $\hat{\mathbf{u}}$ ist dann ebenfalls eine normalverteilte Größe mit dem Mittelwert von Gleichung 4.9 und der Kovarianzmatrix

$$\mathbf{S}_{\hat{\mathbf{u}}} = \sigma^2 [\nabla \mathbf{g} (\nabla \mathbf{g})^T]^{-1} \quad (4.10)$$

Die Genauigkeit der Komponente des Verschiebungsvektors in Richtung des mittleren quadratischen Gradienten ist relativ gut im Verhältnis zur Richtung mit minimalem mittleren quadratischen Gradienten. Das Verhältnis in der Güte ist

umso unterschiedlicher, je größer der Unterschied in den beiden Eigenwerten von S_u ist. Am besten ist daher eine gleichmäßige Verteilung der Gradienten in alle Richtungen.

4.3.3 Approximative Lösung

Es soll nun der Zusammenhang mit einem Verfahren hergestellt werden, welches Nagel 83 verwendet hat, um die an Grauwertecken geschätzten Verschiebungsvektoren iterativ zu verbessern und in die Nachbarbereiche hinein zu propagieren. Dazu müssen wir in Gleichung 4.9 die Näherung einer quadratischen Bildfunktion einsetzen. Dies bedeutet, daß wir die Abhängigkeit des Gradienten ∇g und der zeitlichen Ableitung g_t von den Bildkoordinaten durch eine Taylorreihenentwicklung 1.Ordnung explizit machen können.

$$\begin{aligned}\nabla g(\mathbf{x}_i) &\approx \nabla g(\mathbf{0}) + \mathbf{H}(\mathbf{0}) \mathbf{x}_i \\ g_t(\mathbf{x}_i) &\approx g_t(\mathbf{0}) + (\nabla g_t)^T(\mathbf{0}) \mathbf{x}_i\end{aligned}$$

Im folgenden lassen wir der Kürze halber das Argument $\mathbf{0}$ weg. Wir nehmen weiterhin noch an, daß die Punkte \mathbf{x}_i symmetrisch um den Nullpunkt liegen, so daß

$$\sum_{k=1}^n \mathbf{x}_k = \mathbf{0} \quad \text{und} \quad \sum_{k=1}^n x_k^2 = \sum_{k=1}^n y_k^2 = n\bar{x}^2$$

gilt. Somit erhalten wir für $\nabla g(\nabla g)^T$ den Ausdruck:

$$\begin{aligned}\nabla g(\nabla g)^T &\approx \sum_{i=1}^n (\nabla g + \mathbf{H} \mathbf{x}_i)(\nabla g + \mathbf{H} \mathbf{x}_i)^T \\ &= \sum_{i=1}^n \left[\nabla g(\nabla g)^T + \nabla g(\mathbf{H} \mathbf{x}_i)^T + (\mathbf{H} \mathbf{x}_i)(\nabla g)^T \right. \\ &\quad \left. + (\mathbf{H} \mathbf{x}_i)(\mathbf{H} \mathbf{x}_i)^T \right] \\ &= n \nabla g(\nabla g)^T + \mathbf{H} \left(\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T \right) \mathbf{H} \\ &= n \nabla g(\nabla g)^T + n \bar{x}^2 \mathbf{H} \mathbf{H} = n\mathbf{C}\end{aligned}$$

mit

$$\begin{aligned}\mathbf{C} &= \nabla g(\nabla g)^T + \bar{x}^2 \mathbf{H} \mathbf{H} \\ &= \begin{pmatrix} g_x^2 + \bar{x}^2(g_{xx}^2 + g_{xy}^2) & g_x g_y + \bar{x}^2 g_{xy}(g_{xx} + g_{yy}) \\ g_x g_y + \bar{x}^2 g_{xy}(g_{xx} + g_{yy}) & g_y^2 + \bar{x}^2(g_{yy}^2 + g_{xy}^2) \end{pmatrix}\end{aligned}\tag{4.11}$$

Durch eine ähnliche Rechnung erhalten wir auch

$$\nabla g \mathbf{g}_t^T = n g_t \nabla g + n \bar{x}^2 \mathbf{H} \nabla g_t$$

Setzen wir dies in die Gleichung 4.8 ein und teilen durch n , so erhalten wir die gewünschte Approximation von Gleichung 4.9:

$$\mathbf{C}\mathbf{u} = \left(\nabla g (\nabla g)^T + \bar{x}^2 \mathbf{H} \mathbf{H} \right) \mathbf{u} = - \left(g_t \nabla g + \bar{x}^2 \mathbf{H} \nabla g_t \right) \quad (4.12)$$

Schreiben wir diese Gleichung noch einmal etwas um:

$$\nabla g \left((\nabla g)^T \mathbf{u} + g_t \right) + \bar{x}^2 \mathbf{H} (\mathbf{H} \mathbf{u} + \nabla g_t) = \mathbf{0} \quad (4.13)$$

Dies ist eine gewichtete Summe von zwei Termen. Der erste Term enthält die "motion constraint equation" gewichtet mit dem Gradienten. Somit ist sein Beitrag zur Gesamtsumme in der Richtung des Gradienten am größten. Der zweite Term enthält die erste Ableitung der "motion constraint equation". Während die "motion constraint equation" die Übereinstimmung der Grauwerte fordert, fordert die erste Ableitung die Übereinstimmung der Gradienten. Dieser Term ist gewichtet mit der Krümmungsmatrix. Sein Beitrag ist somit in Richtung der Hauptkrümmung mit dem größten Eigenwert am größten. Insgesamt wird also die Gleichheit der lokalen Grauwertstruktur gefordert. In *Tretiak + Pastor 84* wird vorgeschlagen, nur den zweiten Teil von 4.13, d.h. die Ableitung der "motion constraint equation", für die Schätzung des Verschiebungsvektors zu verwenden. Dort wird allerdings eine andere Herleitung der Gleichung angegeben, die einfach in der Differentiation der "motion constraint equation" besteht. Auch *Haralick + Lee 83* kommen durch die Forderung nach der Gleichheit der lokalen Grauwertstruktur auf das gleiche Ergebnis.

In der Form 4.13 erkennt man auch die Ähnlichkeit zur Gleichung 4.3 am besten. Jene enthält zusätzlich noch eine Korrektur der "motion constraint equation", die durch die Berücksichtigung von quadratischen Termen in der Bildfunktion zustande kommt. Dort ist das Modell einer bivariat quadratischen Bildfunktion schon bei der Herleitung der "constraint equation" eingeflossen, während wir nun die "motion constraint equation" unangetastet gelassen haben, und erst bei der Mittelung über ein Fenster eine Approximation mit einem Polynom zweiten Grades durchgeführt haben.

Durch den Vergleich des Verfahrens von *Nagel 83* mit dem Verfahren der Pseudoinversen sehen wir jetzt deutlich, daß es sich bei \mathbf{C} um eine Approximation der Matrix $\nabla g (\nabla g)^T$ handelt. Eine Untersuchung der Eigenschaften von \mathbf{C} , wie sie z.B. in *Nagel 85*, *Nagel 86* zu finden ist, zeigt, daß diese Matrix entscheidende Informationen über den lokalen Grauwertverlauf enthält. Die Eigenvektoren von \mathbf{C} geben die Richtung extremer Variationen im Grauwert bzw. Gradienten an. Wenn man einen Vergleich mit der Fehlerfortpflanzungsgleichung 4.10 anstellt, so sieht man, daß die Inverse von \mathbf{C} angibt, in welchen Richtungen die Unsicherheit (und damit der Fehler) am größten ist. Deshalb werden wir auf diese Matrix später, bei der Diskussion von gerichteten Glattheitsforderungen, noch einmal zurückkommen.

4.3.4 Die Güte der Schätzung

Um den Verschiebungsvektor \mathbf{u} aus Gleichung 4.12 bestimmen zu können, muß die Matrix \mathbf{C} invertiert werden. Zum Test für die Brauchbarkeit der Grauwertvariation zur Schätzung des Verschiebungsvektors reicht es aber nicht aus, einen konstanten Schwellwert an die Determinante der Matrix \mathbf{C} zu legen [Nagel 85, Nagel + Enkelmann 86]. Um das einzusehen, betrachten wir das Beispiel einer entlang der y -Achse leicht gekrümmten Kante mit (vergl. mit Eckenmodell 4.4)

$$g_x = \text{Maximum}, \quad g_y = 0, \quad g_{xx} = 0, \quad g_{yy} \approx 0 \quad \text{und} \quad g_{xy} = 0$$

Die Determinante von \mathbf{C} nimmt den Wert $\det \mathbf{C} = g_x^2 \cdot \bar{x}^2 g_{yy}^2$ an. Ist der Gradient g_x sehr groß, so kann er die Kleinheit von g_{yy} aufwiegen. Eine zuverlässige Schätzung beider Komponenten des Verschiebungsvektors ist jedoch nur möglich, wenn beide Eigenwerte möglichst groß sind. Entscheidend für die Beurteilung der Güte der Schätzung ist deshalb das Verhältnis der Eigenwerte von \mathbf{C} (siehe auch die Diskussion der Kovarianzmatrix $\mathbf{S}_{\mathbf{u}}$ von Gleichung 4.10). Seien κ und λ die Eigenwerte von \mathbf{C} mit $\kappa \gg \lambda$, so kann man die Abschätzung

$$q = \frac{\det \mathbf{C}}{(\text{Spur } \mathbf{C})^2} = \frac{\kappa \lambda}{(\kappa + \lambda)^2} \approx \frac{\lambda}{\kappa} \quad (4.14)$$

machen. Da sowohl die Determinante als auch die Spur invariant gegenüber Drehungen der Koordinatenachsen sind, brauchen wir die Diagonalisierung gar nicht explizit durchzuführen. Das Maximum des Verhältnisses q wird für $\lambda = \kappa$ mit $1/4$ erreicht. Wenn q wesentlich kleiner als $1/4$ wird, ist also Vorsicht geboten. Dann läßt sich nur eine Komponente von \mathbf{u} verwerten.

Wir können auch den genauen Fehler für eine quadratische unverrauschte Bildfunktion ausrechnen, die in einer lokalen Umgebung verschoben wurde. Approximieren wir die zeitliche Ableitung durch die Differenz der Grauwerte im Nullpunkt und nehmen weiterhin an, daß die räumlichen Ableitungen genau geschätzt werden können, so erhalten wir nach Einsetzen in die approximative Gleichung 4.12 und Drehung des Koordinatensystems in Richtung der Hauptkrümmungsachsen, so daß $g_{xy} = 0$ ist, den geschätzten Verschiebungsvektor (die genaue Herleitung ist in Anhang D zu finden):

$$\hat{\mathbf{u}} = \mathbf{u} - \frac{kr}{\det} \begin{pmatrix} g_x g_{yy}^2 \\ g_y g_{xx}^2 \end{pmatrix} |\mathbf{u}|^2 \quad (4.15)$$

wobei $kr = 1/2 g_{xx} e_u^2 + 1/2 g_{yy} e_v^2$ die Krümmung der Bildfunktion in Richtung der Verschiebung, und $\det = g_x^2 g_{yy}^2 + g_y^2 g_{xx}^2 + \bar{x}^2 g_{xx}^2 g_{yy}^2$ proportional zur Determinante von \mathbf{C} ist.

Zur weiteren Analyse können wir den Fehlervektor $\mathbf{u} - \hat{\mathbf{u}}$ in zwei Komponenten aufspalten. ε^\perp sei der Fehler in Richtung des Gradienten, also senkrecht zur

“Kantenrichtung”, und ε^{\parallel} der Fehler senkrecht zum Gradienten. Wir erhalten dann die beiden Terme:

$$\begin{aligned}\varepsilon^{\parallel} &= \frac{kr}{|\nabla g|} |\mathbf{u}|^2 \frac{g_x g_y (g_{xx}^2 - g_{yy}^2)}{g_x^2 g_{yy}^2 + g_y^2 g_{xx}^2 + x^2 g_{xx}^2 g_{yy}^2} \\ \varepsilon^{\perp} &= \frac{kr}{|\nabla g|} |\mathbf{u}|^2 \frac{g_x^2 g_{yy}^2 + g_y^2 g_{xx}^2}{g_x^2 g_{yy}^2 + g_y^2 g_{xx}^2 + x^2 g_{xx}^2 g_{yy}^2}\end{aligned}\quad (4.16)$$

Der erste Term in den beiden Gleichungen entspricht dem Fehler der “motion constraint equation”, wie wir ihn schon in Gleichung 3.10 ausgerechnet haben. Die zusätzlichen Faktoren kommen durch die Mittelung über das betrachtete Fenster zustande.

Man sieht nun, daß für den Fehler ε^{\parallel} in Richtung der Kante die Differenz der quadrierten Hauptkrümmungen $g_{xx}^2 - g_{yy}^2$ eine wesentliche Rolle spielt. Außerdem ist es gut, wenn die Achse mit der größten Krümmung in Kantenrichtung zeigt, denn falls eine der Hauptkrümmungen verschwindet und der Gradient in diese Richtung zeigt, so verschwindet der Fehler, während er andererseits ins unendliche wächst, wenn senkrecht zum Gradienten die Krümmung verschwindet.

Der Fehler in Richtung des Gradienten ist immer kleiner als der Fehler einer einzelnen “motion constraint equation”. Bei fehlender Krümmung verschwindet der Fehler sogar ganz, denn dann ist die “motion constraint equation” uneingeschränkt gültig.

4.3.5 Weitere Ansätze

Statt der Konstanzannahme kann man den lokalen Zusammenhang zwischen den Verschiebungsvektoren auch statistisch durch eine Markov-Kette beschreiben.

$$\mathbf{u}_i = \Phi \mathbf{u}_{i-1} + \mathbf{w}_i$$

Der Verschiebungsvektor \mathbf{u}_i am Ort \mathbf{x}_i stellt hier eine Zufallsvariable dar, die von dem Verschiebungsvektor \mathbf{u}_{i-1} abhängt. Der Zusammenhang ist allerdings nicht deterministisch, sondern durch eine Sequenz von normalverteilten, mittelwertfreien, unabhängigen Zufallsvariablen $\{\mathbf{w}_i : i = 1, 2, \dots\}$ additiv verrauscht. Eine solche Sequenz von Zufallsvariablen nennt man einen diskreten, weißen, Gaußschen Prozeß.

Wir können den Zusammenhang einer Messung z_i mit dem Verschiebungsvektor \mathbf{u}_i ebenfalls durch eine solche Gleichung beschreiben:

$$z_i = \mathbf{H}_i \mathbf{u}_i + v_i$$

Auch hier ist v_i ein diskreter, weißer, Gaußscher Prozeß. Eine Messung ist dabei ein Wert, der von den Grauwerten im Bild abhängt.

Um eine optimale Schätzung für ein \mathbf{u}_i zu erhalten, wobei die Messungen $\{z_1, \dots, z_i\}$ gemacht wurden, kann man den Modalwert der bedingten Wahrscheinlichkeitsdichte $p(\mathbf{u}_i | z_1, \dots, z_j)$ verwenden. In diesem linearen Fall stimmt der Modalwert mit dem Mittelwert überein. Das Ergebnis ist der Kalman-Bucy-Filter (siehe *Jazwinski 70* für die mathematischen Grundlagen der "Filtering Theory" und *Stuller + Krishnamurthy 83* für die Anwendung auf die Schätzung von Verschiebungsvektoren).

Als Verallgemeinerung der Konstanzannahme kann man Verfahren ansehen, die nicht die lokale Konstanz der Verschiebungsvektoren selbst fordern, sondern die Konstanz von Parametern, die das Verschiebungsvektorfeld lokal beschreiben. Da der optische Fluß einer bewegten ebenen Oberfläche bei orthographischer Projektion durch eine affine Abbildung dargestellt werden kann, ist es vernünftig, lokal eine lineare Variation des Verschiebungsvektorfeldes anzunehmen.

$$g_2(\mathbf{x}) = g_1(\mathbf{A}\mathbf{x} + \mathbf{b})$$

Schalkoff + Labuz 84 schätzen die 6 Parameter unter der Annahme einer linearen Bildfunktion über eine quadratische Fehlerminimierung (entspricht der Bildung der Pseudoinversen). Dies entspricht der im Abschnitt 4.1 besprochenen "functional velocity method" von *Waxman + Wohn 84* für den Fall einer orthographischen Projektion. Um die Fehler der Linearitätsannahme der Bildfunktion zu verringern, wird das Verfahren iterativ wiederholt. Der Aufwand bei der Berechnung der Pseudoinversen (es muß eine 6×6 -Matrix invertiert werden) steigt allerdings erheblich an. Beliebter sind deshalb Verfahren, die die Annahme konstanter Parameter fallen lassen, und stattdessen die schwächere Forderung der Glattheit des Verschiebungsvektorfeldes stellen.

4.4 Von der Konstanz zur Glattheit

Im vorherigen Abschnitt sahen wir, daß eine zuverlässige Schätzung eines Verschiebungsvektors mit der Voraussetzung der lokalen Konstanz des Verschiebungsvektors nur an Stellen des Bildes möglich ist, an denen die lokale Grauwertvarianz bestimmten Forderungen genügt. Um auch an den Stellen, an denen diese Forderungen nicht erfüllt sind, einen Verschiebungsvektor schätzen zu können, muß ein globales Verfahren verwendet werden. Man könnte daran denken, die an Ecken geschätzten Verschiebungen in homogene Bereiche auszudehnen. Dabei sollte das entstehende Verschiebungsvektorfeld einen möglichst glatten Verlauf haben [*Horn + Schunck 81*]. Diese Glattheit kann durch ein Funktional \mathcal{F} gemessen und dann optimiert werden. *Nagel 87* zeigte, daß an Stellen mit ausreichender Grauwertvarianz die Optimierung eines Glattheitsterms mit der Konstanzannahme äquivalent ist.

Das Glattheitsfunktional \mathcal{F} wird in der mathematischen Theorie über die Regularisierung von "ill-posed problems" auch als "stabilizing functional" bezeichnet

[Terzopoulos 86b]. Es handelt sich also um einen auch auf anderen Gebieten verwendbaren Ansatz, um a priori Information über mögliche Lösungen in ein Verfahren einzubauen (Einschränkung des Lösungsraums auf Funktionen mit glattem Verlauf). Ein weiteres Gebiet, auf dem dieser Ansatz erprobt wurde, ist die Rekonstruktion von Oberflächen aufgrund von Schattierungsinformationen ("shape from shading" siehe z.B. Ikeuchi + Horn 81).

4.4.1 Problemformulierung

Das Problem, das wir zu lösen haben, besteht in der Optimierung eines Funktionals, welches zwei Terme enthält. Zum einen die Bedingung, die die "motion constraint equation" an das Verschiebungsvektorfeld stellt, und zum zweiten einen Term, der die Glattheit des Verschiebungsvektorfeldes mißt:

$$M = \int_G [\mathbf{u}(\mathbf{x})^T \nabla g(\mathbf{x}) + g_t(\mathbf{x})]^2 + \alpha^2 \mathcal{F} \, d\mathbf{x} \rightarrow \text{Minimum} \quad (4.17)$$

Das optimale Verschiebungsfeld wird einerseits bestrebt sein, lokal möglichst gut die "motion constraint equation" zu erfüllen, und andererseits global einen möglichst 'glatten' Verlauf zu haben, d.h. das Glattheitsmaß zu minimieren. Der Wert der Konstanten α^2 bestimmt die relative Wichtung dieser beiden zu optimierenden Terme. Der zweite Term ist für alle Bildfolgen gleich und stellt deshalb eine Formulierung von a priori Wissen über die Form der im Anwendungsbereich vorhandenen Verschiebungsvektorfelder dar. Es handelt sich quasi um ein Modell eines Verschiebungsvektorfeldes, in der Art, wie z.B. das Verformungspotential einer elastischen Membran ein Modell für einen physikalischen Körper darstellt. Die Form einiger einfacher Glattheitsfunktionale hat in der Tat Ähnlichkeit mit diesem Potential einer elastischen Membran (siehe Abschnitt 4.4.2). Der erste Term wertet die eigentliche Bildinformation aus, und sieht in jeder Bildfolge anders aus. Er ist vergleichbar mit dem Potential einer externen Kraft, die auf die Membran wirkt.

Man kann statt der "motion constraint equation" auch die Grauwertdifferenz $g_2(\mathbf{x}) - g_1(\mathbf{x} - \mathbf{u})$ in diese Minimierungsgleichung einsetzen [Nagel + Enkelmann 86]. Dieser Ansatz reduziert sich auf den oben angegebenen, wenn man die Abhängigkeit von \mathbf{u} durch eine Taylorentwicklung bis zur ersten Ordnung von g_1 explizit macht (siehe Abschnitt 3.4).

Bei der Aufgabe, eine Funktion zu finden, die ein Integral optimiert, handelt es sich um ein Problem der Variationsrechnung (siehe z.B. [Courant + Hilbert 68a]). Falls wir annehmen, daß der Glattheitsterm \mathcal{F} nicht von der Funktion \mathbf{u} , sondern nur von dessen Ableitungen bis zu einem endlichen Grade n abhängt, so ergeben sich die Euler-Lagrange-Gleichungen für dieses Problem zu:

$$\begin{aligned} g_x (\nabla g)^T \mathbf{u} + g_x g_t &= \alpha^2 [\mathcal{F}]_u \\ g_y (\nabla g)^T \mathbf{u} + g_y g_t &= \alpha^2 [\mathcal{F}]_v \end{aligned} \quad (4.18)$$

mit

$$[\mathcal{F}]_u = \frac{d}{dx} \mathcal{F}_{u_x} + \frac{d}{dy} \mathcal{F}_{u_y} - \frac{d^2}{dx^2} \mathcal{F}_{u_{xx}} - 2 \frac{d^2}{dx dy} \mathcal{F}_{u_{xy}} + \dots + (-1)^n \frac{d^n}{dy^n} \mathcal{F}_{u_{y \dots y}}$$

und

$$[\mathcal{F}]_v = \frac{d}{dx} \mathcal{F}_{v_x} + \frac{d}{dy} \mathcal{F}_{v_y} - \frac{d^2}{dx^2} \mathcal{F}_{v_{xx}} - 2 \frac{d^2}{dx dy} \mathcal{F}_{v_{xy}} + \dots + (-1)^n \frac{d^n}{dy^n} \mathcal{F}_{v_{y \dots y}}$$

Lassen wir die Glattheitsforderung vollkommen fallen und setzen $\mathcal{F} = 0$, so erhalten wir als Ergebnis die kontinuierliche Form der Gleichung 3.16:

$$\nabla g (\nabla g)^T \mathbf{u} + \nabla g g_t = 0$$

Dies ist nichts weiter als die mit dem Gradienten multiplizierte "motion constraint equation". Da $\det \nabla g (\nabla g)^T$ verschwindet, ist diese Gleichung natürlich nicht nach \mathbf{u} auflösbar. Vergleicht man dieses Ergebnis mit Gleichung 4.18, so sieht man, daß die Glattheitsforderung eine Korrektur der einfachen "motion constraint equation" darstellt. Diese hängt nur von der unmittelbaren Umgebung des Verschiebungsvektorfeldes ab, da in die Terme $[\mathcal{F}]_u$ und $[\mathcal{F}]_v$ nur lokale Eigenschaften eingehen.

Man kann versuchen, dieses Differentialgleichungssystem 4.18 iterativ zu lösen. Dazu geht man von einer approximativen Lösung $\mathbf{u}_0(\mathbf{x})$ aus und versucht, mit dem Ansatz $\mathbf{u} = \mathbf{u}_0 + d\mathbf{u}$ einen Korrekturvektor $d\mathbf{u}(\mathbf{x})$ zu erhalten. Setzen wir $\mathbf{u} = \mathbf{u}_0 + d\mathbf{u}$ in die linke Seite der Differentialgleichung ein, so erhalten wir:

$$\nabla g (\nabla g)^T d\mathbf{u} = - \left(\nabla g (\nabla g)^T \mathbf{u}_0 + \nabla g g_t \right) + \alpha^2 [\mathcal{F}]_{\mathbf{u}_0} \quad (4.19)$$

Diese Gleichung läßt sich leider nicht nach $d\mathbf{u}$ auflösen, da die Matrix $\nabla g (\nabla g)^T$ nicht invertierbar ist. Nagel + Enkelmann 86 haben versucht, eine Invertierung durch eine Mittelung zu erreichen, wobei der Term $\nabla g (\nabla g)^T$ durch die Matrix \mathbf{C} aus Gleichung 4.11 ersetzt wird. Man erwartet jedoch, nach den Ergebnissen des letzten Abschnitts, daß man mit dieser Methode nur in der Nähe von Grauwert-ecken zu einem zuverlässigen Ergebnis kommen kann.

Wir müssen also den Einfluß von $d\mathbf{u}$ auch auf der rechten Seite der Differentialgleichung 4.18 berücksichtigen. Dazu ist es notwendig, sowohl das Funktional \mathcal{F} als auch die Form der diskreten Operatoren zu kennen, mit denen man die Ableitungen von $\mathbf{u}(\mathbf{x})$ berechnen will. Bei den in den nächsten Abschnitten zu besprechenden Verfahren wird aber im allgemeinen der Einfluß von $d\mathbf{u}$ nur an der Stelle des gerade betrachteten Ortes \mathbf{x} im Term $[\mathcal{F}]_u$ berücksichtigt.

Um einen Eindruck von der Größenordnung des Problems zu bekommen, stellen wir die Analogie zum Problem der Lösung eines linearen Gleichungssystems her. Im diskreten Fall kann man nämlich die Differentialgleichungen für ein Bild mit N Pixeln in einem linearen Gleichungssystem mit $2N$ Gleichungen der Form

$$\mathbf{A} \mathbf{z} = -\mathbf{b} + \alpha^2 \mathbf{F} \mathbf{z}$$

zusammenfassen. Dabei gilt:

$$\mathbf{z}_k = \begin{cases} u(\mathbf{x}_i) & \text{für } k = 2i - 1 \\ v(\mathbf{x}_i) & \text{für } k = 2i \end{cases}$$

$$\mathbf{A}_{kl} = \begin{cases} g_x^2(\mathbf{x}_i) & \text{für } k = l = 2i - 1 \\ g_x(\mathbf{x}_i)g_y(\mathbf{x}_i) & \text{für } k = 2i, l = 2i - 1 \\ g_x(\mathbf{x}_i)g_y(\mathbf{x}_i) & \text{für } k = 2i - 1, l = 2i \\ g_y^2(\mathbf{x}_i) & \text{für } k = l = 2i \\ 0 & \text{sonst} \end{cases}$$

$$\mathbf{b}_k = \begin{cases} g_x(\mathbf{x}_i)g_t(\mathbf{x}_i) & \text{für } k = 2i - 1 \\ g_y(\mathbf{x}_i)g_t(\mathbf{x}_i) & \text{für } k = 2i \end{cases}$$

Die Matrix \mathbf{F} ist vom Funktional und den verwendeten Operatoren abhängig und wegen der Lokalität im allgemeinen dünn besetzt. Die angegebene Methode zur Lösung der Differentialgleichung entspricht dann dem Jacobi-Verfahren zur Lösung dieses linearen Gleichungssystems [Marsal 76]. Man kann aber auch das Gauß-Seidel-Verfahren anwenden, wobei die Änderung \mathbf{du} des Verschiebungsfeldes an einer Stelle \mathbf{x} noch im gleichen Iterationsschritt zur Berechnung der Änderungen an anderen Stellen benutzt wird.

4.4.2 Einfache Glattheitsmaße

Die Frage ist nun, was für ein Glattheitsmaß man wählen sollte. Um das zu entscheiden, kann man zunächst einmal untersuchen, aus was für Komponenten sich ein Verschiebungsvektorfeld zusammensetzt. Nehmen wir ein lineares Verschiebungsvektorfeld an:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Dieses können wir in verschiedene, gegenüber Drehungen der (x, y) -Ebene möglichst invariante, Komponenten zerlegen:

$$\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \underbrace{\frac{v_x - u_y}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\text{Rotation}} + \underbrace{\frac{u_x + v_y}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\text{Divergenz}} + \underbrace{\begin{pmatrix} \frac{u_x - v_y}{2} & \frac{u_y + v_x}{2} \\ \frac{u_y + v_x}{2} & -\frac{u_x - v_y}{2} \end{pmatrix}}_{\text{Scherung}}$$

Der Rotations- und Divergenzanteil sowie die Determinante $(u_x - v_y)^2 + (u_y + v_x)^2$ des Scherungsanteils des Verschiebungsvektorfeldes sind invariant gegenüber Drehungen des Bildkoordinatensystems. Ein Glattheitsmaß \mathcal{F} , das nur erste Ableitungen des Verschiebungsvektorfeldes berücksichtigt, sollte diese verschiedenen Anteile in einer charakteristischen Art und Weise wichten, denn das Glattheitsmaß sollte nicht von der Lage des Bildkoordinatensystems abhängen. Je größer der Wichtungsterm eines invarianten Terms, desto weniger Anteile dieses Termes hat

später das Verschiebungsvektorfeld, denn es ist bei der Optimierung günstig, diesen Term möglichst klein zu halten. Da die oben genannten Terme invariant gegenüber Drehungen sind, kann man eventuell theoretische Eigenschaften des optischen Flusses berücksichtigen. Eine Analyse des optischen Flusses, den eine sich drehende Ebene bei orthographischer Projektion auf der Bildebene erzeugt, zeigt z.B., daß der Scherungsanteil immer mindestens genauso groß ist, wie der Divergenzanteil [Kanatani 86]:

$$(u_x + v_y)^2 \leq (u_x - v_y)^2 + (u_y + v_x)^2$$

Eine naheliegende Wahl für ein Glattheitsmaß \mathcal{F} haben [Horn + Schunck 81] getroffen:

$$\mathcal{F} = u_x^2 + u_y^2 + v_x^2 + v_y^2 \quad (4.20)$$

Dieser Glattheitsterm wichtet alle Anteile des optischen Flusses gleich stark, denn es gilt:

$$u_x^2 + u_y^2 + v_x^2 + v_y^2 = \frac{1}{2} \left\{ (u_x + v_y)^2 + \left((u_x - v_y)^2 + (u_y + v_x)^2 \right) + (u_y - v_x)^2 \right\}$$

Setzen wir diesen konkreten Glattheitsterm für \mathcal{F} in die Euler-Lagrange-Gleichungen 4.18 ein, so erhalten wir die partiellen Differentialgleichungen:

$$\begin{aligned} g_x(\nabla g)^T \mathbf{u} + g_x g_t &= \alpha^2(u_{xx} + v_{yy}) = \alpha^2 \Delta u \\ g_y(\nabla g)^T \mathbf{u} + g_y g_t &= \alpha^2(v_{xx} + u_{yy}) = \alpha^2 \Delta v \end{aligned}$$

Für den iterativen Lösungsansatz müssen wir jetzt den Operator finden, mit dem Δu und Δv berechnet werden, und dessen Einfluß auf \mathbf{du} explizit machen. Dazu setzen wir $\Delta \mathbf{u} = (\Delta u, \Delta v)^T \approx \kappa(\bar{\mathbf{u}} - \mathbf{u}) \approx \kappa(\bar{\mathbf{u}}_0 - \mathbf{u})$ wobei $\bar{\mathbf{u}}_0$ eine Mittelung des Startvektorfeldes $\mathbf{u}_0(\mathbf{x})$ in einer lokalen Umgebung sein soll. Der Proportionalitätsfaktor κ hängt vom Mittelungsoperator ab. Es gilt $\Delta \mathbf{u} \approx \kappa(\bar{\mathbf{u}}_0 - \mathbf{u}_0) - \kappa \mathbf{du}$ und wir erhalten

$$\begin{aligned} \mathbf{du} &= - \left(\nabla g(\nabla g)^T + \alpha^2 \kappa \mathbf{I} \right)^{-1} \left[\nabla g(\nabla g)^T \mathbf{u}_0 + \nabla g g_t - \alpha^2 \kappa (\bar{\mathbf{u}}_0 - \mathbf{u}_0) \right] \\ \mathbf{u}_{i+1} &= - \left(\nabla g(\nabla g)^T + \alpha^2 \kappa \mathbf{I} \right)^{-1} \left[\nabla g g_t - \alpha^2 \kappa \bar{\mathbf{u}}_i \right] \end{aligned}$$

In der letzten Gleichung wurde berücksichtigt, daß man die Verbesserung wiederum iterativ auf das erhaltene Verschiebungsfeld anwenden kann. Der Term $\alpha^2 \kappa \mathbf{I}$ stellt die Invertierbarkeit der Matrix sicher.

Der eben verwendete Glattheitsterm entspricht nach Anandan + Weiss 85 dem Modell einer Membran, d.h. der integrierte Glattheitsterm sollte der potentiellen Energie einer nach der Funktion $\mathbf{u}(\mathbf{x})$ verzerrten ebenen Membran entsprechen. Die Membran geht in den Zustand minimaler potentieller Energie über, welches dann der Minimierung des Integrals des Glattheitsterms entspricht. Dies stimmt

allerdings nur insoweit, als die potentielle Energie einer Membran nur von den ersten Ableitungen des Verschiebungsfeldes abhängt. Nur bei einer Auslenkung einer Membran senkrecht zur Oberfläche erhält man einen Term, der dem Betragsquadrat des Gradienten entspricht [Courant + Hilbert 68a, S. 214]. Mit den Lamé'schen Elastizitätskonstanten λ und μ erhält man für die potentielle Energie einer elastischen homogenen Membran ohne Berücksichtigung der Torsion nach [Feynman 64, Kap. 39]¹ den Ausdruck:

$$E_{pot} = \frac{1}{2} \int_{Vol.} (\lambda + \mu)(u_x + v_y)^2 + \mu \left((u_x - v_y)^2 + (u_y + v_x)^2 \right) dVol.$$

Der erste Term beschreibt den divergenten Anteil und der zweite Term den Scherungsanteil des Verschiebungsvektorfeldes. In der Gleichung für die potentielle Energie taucht hingegen kein Rotationsanteil auf. Die Energie einer Membran verändert sich nicht, wenn diese nur gedreht wird. Wie wir oben sahen, gilt dies für die Glattheitsforderung von Horn + Schunck 81 nicht.

Dengler verwendet zunächst diesen Glattheitsterm, der dem Modell einer elastischen Membran entspricht [Dengler 85a, Dengler 85b, Dengler u.a. 86]. Da er aber den Divergenzanteil durch Nullsetzen des Terms $\lambda + \mu$ wegläßt, um nur noch den Scherungsanteil des Verschiebungsfeldes zu minimieren, erhält er schließlich die gleiche Differentialgleichung wie Horn + Schunck 81. Denn das allgemeinste Glattheitsmaß erster Ordnung mit konstanter Wichtung der invarianten Anteile

$$\mathcal{F} = \frac{1}{2} \left[a(u_x + v_y)^2 + b(u_y - v_x)^2 + c((u_x - v_y)^2 + (u_y + v_x)^2) \right]$$

ergibt folgende Differentialgleichungen:

$$\begin{aligned} g_x (\nabla g)^T \mathbf{u} + g_x g_t &= \alpha^2 [(a+c)u_{xx} + (b+c)u_{yy} + (a-b)v_{xy}] \\ g_y (\nabla g)^T \mathbf{u} + g_y g_t &= \alpha^2 [(a+c)v_{yy} + (b+c)v_{xx} + (a-b)u_{xy}] \end{aligned}$$

Während Horn + Schunck 81 alle Konstanten gleich setzen ($a = b = c$), setzt Dengler die Wichtungsfaktoren für den Divergenz- und den Rotationsanteil ($a = b = 0$) auf Null und erhält so das gleiche Ergebnis.

Die Energie einer ebenen dünnen Platte, die in Normalenrichtung verbeult wird, ist proportional zu einer quadratischen Form der Hauptkrümmungen, der bei der Biegung entstandenen Fläche [Courant + Hilbert 68a, S. 217]:

$$E_{pot} \sim \int (u_{xx} + u_{yy})^2 - 2(1-\mu)(u_{xx}u_{yy} - u_{xy}^2) dx dy \quad (4.21)$$

Anandan + Weiss 85 schlagen als alternatives Glattheitsmaß eine Erweiterung dieses Energieterms für zwei Dimensionen vor:

$$\mathcal{F} = u_{xx}^2 + 2u_{xy}^2 + u_{yy}^2 + v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2$$

¹Wobei Formel (39.21) $C_{xyxy} = 2\mu$ in $C_{xyxy} = \mu$ korrigiert wurde. Vergleiche auch mit [Courant + Hilbert 68b, S. 530].

Im Gegensatz zum Membranmodell ist hier gefordert, daß die erste Ableitung des Verschiebungsvektorfeldes kontinuierlich sein muß. Während das Membranmodell also nur eine gewisse Stetigkeit des Vektorfeldes fordert, ist hier eine Glattheit im eigentlichen Sinne gefordert. Aus diesem Grunde findet dieser Term auch Verwendung bei der Approximation von Oberflächen [Grimson 83] oder der Interpolation mittels generalisierter Splines [Terzopoulos 86b]. Allerdings läßt sich die Wirkung dieses Glattheitsterms im konkreten Fall schlecht abschätzen. Insbesondere sind alle linear von den Bildkoordinaten abhängigen Verschiebungsvektorfelder zugelassen. Es ist daher sehr fraglich, ob die Verwendung dieses Glattheitsterms sinnvoll ist. Auch die Experimente von Anandan + Weiss 85 zeigen, daß dieser Term dem Glattheitsmaß von Horn und Schunck unterlegen ist.

Scott 86 stellt die Frage, unter welchen Umständen die Glattheitsbedingung eine gute Approximation an die "Wahrheit" ist. Zumindest in einfachen Fällen sollte das Ergebnis mit analytischen Werten übereinstimmen. Variationen des Verschiebungsfeldes in der Bildebene sind aber, jedenfalls im perspektivischen Fall, nicht von erster Ordnung, es sei denn das Objekt bewegt sich parallel zur Bildebene. Er definiert aus diesem Grunde den "motion vector"

$$\mathbf{p} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) = \frac{1}{Z} \left(\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt} \right)$$

Die Beziehung zu den Verschiebungsvektoren (u, v) wird durch die Gleichungen

$$\begin{aligned} u &= \frac{dx}{dt} - x \frac{dz}{dt} \\ v &= \frac{dy}{dt} - y \frac{dz}{dt} \end{aligned}$$

hergestellt. Während der Verschiebungsvektor \mathbf{u} im Falle einer sich bewegenden Ebene nicht mehr linear von den Bildkoordinaten abhängt (siehe Gleichung 4.1), ist dies beim "motion vector" \mathbf{p} durchaus der Fall. Daher schlägt Scott vor, nicht den Verschiebungsvektor selber mit dem Verfahren von Horn + Schunck 81 zu schätzen, sondern den "motion vector" \mathbf{p} . Dies ergibt dann den folgenden "constraint":

$$-g_t = g_x \frac{dx}{dt} + g_y \frac{dy}{dt} - (g_x x + g_y y) \frac{dz}{dt}$$

Bisher gab es eine "constraint line" im (u, v) -Raum, nun gibt es eine "constraint plane" im \mathbf{p} -Raum. Die Ebene steht senkrecht auf dem Vektor $(g_x, g_y, -g_x x - g_y y)$, d.h. wir schätzen in jedem Punkt der Bildebene die Komponente von \mathbf{p} , in dieser Richtung. Mit diesem Ansatz kann z.B. ein divergentes Feld, welches durch Translation in Z-Richtung entsteht zu einem konstanten \mathbf{p} -Feld reduziert werden. Die Einführung einer dritten Dimension kann also zu einem glatteren Vektorfeld führen.

4.4.3 Gerichtete Glattheitsmaße

Ein Problem mit den einfachen Glattheitsmaßen ist, daß sie die geschätzten Verschiebungen unabhängig vom Bildinhalt weitergeben. Dadurch kann es leicht geschehen, daß Verschiebungen über Diskontinuitäten hinweg z.B. in einen gleichmäßigen unbewegten Hintergrund hinein propagiert werden. Ein Glattheitsmaß, welches Glattheit nur in Gebieten fordert, die zum selben (starrten) Objekt gehören, wäre günstiger, da dort die Glattheitsannahme gerechtfertigt erscheint.

Cornelius + Kanade 83 verwenden zwar den Glattheitsterm 4.20 von *Horn + Schunck 81*, gehen aber in der Nähe von Kanten, die zuvor mit dem Mexican-Hat-Operator gefunden wurden, anders vor. Dort wird die Hypothese gleicher Verschiebung auf beiden Seiten nur getroffen, wenn der Fehlerterm aus der "motion constraint equation" und der Glattheit, sowie einer zusätzlichen Wichtung der Grauwertänderung kleiner ist, als bei getrennter Glattheit beiderseits der Kante. Dies hat den Nachteil, daß die Vorsegmentierung des Bildes einen erheblichen Einfluß auf die Berechnung des optischen Flusses hat. Man möchte jedoch die zusätzliche Information des optischen Flusses selbst gern für eine Segmentierung nutzen.

Noch extremeren Einfluß hat die Vorsegmentierung bei einem Verfahren von *Hildreth [Hildreth 82, Hildreth 83, Hildreth 84]*. Hier werden Verschiebungen nur entlang der Nulldurchgangslinien des Mexican-Hat-Operators geschätzt. Eine Schätzung der Normalkomponente \hat{u}^\perp der Verschiebungen entlang der Kante wird vorausgesetzt. Dazu kann z.B. Gleichung 3.2 verwendet werden. Im Verfahren wird dann die Tangentialkomponente durch die Minimierung eines über die gesamte Kante aufintegrierten Maßes bestimmt. Das Maß besteht dabei aus zwei Termen, von denen der eine die Abweichung zur geschätzten Normalkomponente (zweiter Term in 4.22), und der andere die Glattheit der Verschiebungsvektoren entlang der Kante (erster Term in 4.22) mißt.

$$\int \left| \frac{\partial \mathbf{u}}{\partial s} \right|^2 + \beta \left(\mathbf{u}^T \frac{\nabla g}{|\nabla g|} - \hat{u}^\perp \right) ds \rightarrow \text{Minimum} \quad (4.22)$$

s ist die Bogenlänge entlang der Kurve und β ein Wichtungsfaktor, der das relative Gewicht der geschätzten Normalkomponente zum Glattheitsterm angibt.

In *Hildreth 82* wird auch der Nachweis geführt, daß dieses Integral immer eine eindeutige Lösung besitzt, wenn es sich um eine differenzierbare, gekrümmte Kurve handelt. Bei der Verwendung anderer Glattheitsmaße, wie z.B. der quadratischen Differenz der Vektorlängen (ohne Berücksichtigung der Richtung der Verschiebungsvektoren), ist eine eindeutige Lösung nicht immer gesichert.

Interessant sind auch die Übereinstimmungen mit Eindrücken, die der Mensch bei der Drehung einer logarithmischen Spirale oder einem Barber-Pole (kreisförmigen Helix) hat. Der Mensch erliegt einer optischen Täuschung und sieht eine falsche Bewegung. Das Verfahren von *Hildreth* kommt zu den gleichen Ergebnissen. Der Nachteil dieses Verfahrens ist, daß, will man ein dichtes Vektorfeld

haben, die Verschiebungen noch in das Innere der Bereiche hinein propagiert werden müssen.

Gerichtete Glattheit mittels Wichtungsmatrix

Im Gegensatz zu diesen beiden Verfahren kann man auch das Funktional \mathcal{F} mit einer räumlichen Kontrolle über die Glattheitseigenschaften ausstatten [Terzopoulos 86b]. Nagel + Enkelmann 86 schlagen einen gerichteten Glattheitsterm vor, der keine Vorsegmentierung des Bildes voraussetzt. Diese Richtungsabhängigkeit wird in Form einer Wichtungsmatrix \mathbf{W} in den Glattheitsterm 4.20 eingebaut. Wir werden zunächst die Differentialgleichungen untersuchen, die sich aus diesem Ansatz ergeben, bevor wir im nächsten Schritt eine spezielle Form der Wichtungsmatrix angeben.

$$\begin{aligned}\mathcal{F} &= \text{Spur} \left\{ (\nabla \mathbf{u}^T)^T \mathbf{W} (\nabla \mathbf{u}^T) \right\} \\ &= w_{11}(u_x^2 + v_x^2) + (w_{12} + w_{21})(u_x u_y + v_x v_y) + w_{22}(u_y^2 + v_y^2)\end{aligned}\quad (4.23)$$

Im Fall $\mathbf{W} = \mathbf{I}$ erhalten wir die Glattheitsforderung von Horn + Schunck 81. Die Terme $[\mathcal{F}]_u$ und $[\mathcal{F}]_v$ aus der Gleichung 4.18 ergeben sich zu

$$[\mathcal{F}]_u = \nabla^T \left((\mathbf{W} + \mathbf{W}^T) \nabla u \right)$$

und

$$[\mathcal{F}]_v = \nabla^T \left((\mathbf{W} + \mathbf{W}^T) \nabla v \right)$$

Nehmen wir nun an, daß die Wichtungsmatrix \mathbf{W} symmetrisch ist und sich ihre Komponenten räumlich nur langsam ändern, so können die räumlichen Ableitungen von \mathbf{W} vernachlässigt werden, und wir erhalten die vereinfachten Terme:

$$\begin{aligned}\frac{1}{2}[\mathcal{F}]_u &= \text{Spur} \left\{ \mathbf{W} \nabla (\nabla u)^T \right\} = w_{11}u_{xx} + (w_{12} + w_{21})u_{xy} + w_{22}u_{yy} \\ \frac{1}{2}[\mathcal{F}]_v &= \text{Spur} \left\{ \mathbf{W} \nabla (\nabla v)^T \right\} = w_{11}v_{xx} + (w_{12} + w_{21})v_{xy} + w_{22}v_{yy}\end{aligned}\quad (4.24)$$

Nun müssen wir noch die Wirkung von $d\mathbf{u}$ auf das Zentralpixel berechnen, um die Differentialgleichung nach der Gauß-Seidel-Methode lösen zu können. Dazu nehmen wir an, daß die Ableitungen von \mathbf{u} mit den Beaudet-Operatoren berechnet werden. Die Beaudet-Operatoren für die zweiten Ableitungen haben die Form:

$$\begin{aligned}\widehat{g_{xx}} &= \sum_{i,j=-k}^k \frac{30(3i^2 - k(k+1))}{(2k+1)^2 k(k+1)(2k-1)(2k+3)} g(i\Delta x, j\Delta y) \\ \widehat{g_{xy}} &= \sum_{i,j=-k}^k \frac{9ij}{k^2(k+1)^2(2k+1)^2} g(i\Delta x, j\Delta y)\end{aligned}\quad (4.25)$$

Die Wichtung des Zentralpixels ist somit:

$$\text{Bei } \widehat{g_{xx}} : -\frac{30}{(2k+1)^2(2k-1)(2k+3)} \quad \text{und bei } \widehat{g_{xy}} : 0$$

Werden also u_{xx} , u_{xy} , u_{yy} und v_{xx} , v_{xy} , v_{yy} mit den Beudet-Operatoren berechnet, so gilt mit $\mathbf{u} = \mathbf{u}_0 + d\mathbf{u}$

$$\begin{aligned} & \text{Spur } \{ \mathbf{W} \nabla (\nabla u)^T \} \\ & \approx \text{Spur } \{ \mathbf{W} \nabla (\nabla u_0)^T \} - \frac{30}{(2k+1)^2(2k-1)(2k+3)} du \text{ Spur } \{ \mathbf{W} \} \\ & \text{Spur } \{ \mathbf{W} \nabla (\nabla v)^T \} \\ & \approx \text{Spur } \{ \mathbf{W} \nabla (\nabla v_0)^T \} - \frac{30}{(2k+1)^2(2k-1)(2k+3)} dv \text{ Spur } \{ \mathbf{W} \} \end{aligned}$$

Wir wollen jetzt eine ausführliche Herleitung einer Wichtungsmatrix angeben, wie sie Nagel + Enkelmann 86 verwendet haben. Dabei handelt es sich um eine neuartige Konstruktion des Glattheitsterms, aus der deutlich werden wird, in welche Richtungen die Glattheit propagiert wird.

Nehmen wir an, daß wir Glattheit in eine bestimmte Richtung \mathbf{e} fordern wollen. Wir müssen dazu die Variation von \mathbf{u} in dieser Richtung minimieren. Diese beträgt:

$$\Delta \mathbf{e} \mathbf{u} = (\nabla \mathbf{u}^T)^T \mathbf{e}$$

Das Quadrat dieses Vektors ist ein Maß für die Glattheit des Verschiebungsfeldes in Richtung von \mathbf{e} :

$$|\Delta \mathbf{e} \mathbf{u}|^2 = \text{Spur } \{ (\nabla \mathbf{u}^T)^T \mathbf{e} \mathbf{e}^T (\nabla \mathbf{u}^T) \}$$

Wenn wir diesen Term minimieren, so bedeutet das, daß wir gerade die Glattheit in die Richtung \mathbf{e} fordern. Durch einen Vergleich mit dem Glattheitsterm \mathcal{F} von Gleichung 4.23 erkennen wir, daß wir $\mathbf{W} = \mathbf{e} \mathbf{e}^T$ wählen müssen. Wir können \mathbf{W} auch noch mit einer, von der Grauwertstruktur abhängigen, Konstante multiplizieren. Außerdem kann man mehrere Matrizen additiv zu einer Gesamtmatrix kombinieren.

Glattheit senkrecht zum Gradienten

Zunächst ist es sinnvoll, Glattheit in Richtung einer "Kante", also senkrecht zum Gradienten zu fordern, denn bei einer solchen Diskontinuität der Grauwertfunktion könnte es sich um eine Objektkante handeln. Glattheit darf dann nicht über die Kante hinaus gefordert werden. \mathbf{e} ist somit der Einheitsvektor senkrecht zum Gradienten:

$$\mathbf{e} = \frac{1}{|\nabla g|} \begin{pmatrix} -g_y \\ g_x \end{pmatrix}$$

Wir können die Wichtungsmatrix, die sich aus der Glattheitsforderung in dieser Richtung ergibt, noch mit dem Betragsquadrat des Gradienten multiplizieren, denn je größer der Gradient, desto größer ist die Diskontinuität, und desto größer ist auch die Wahrscheinlichkeit, daß es sich um eine Objektgrenze handelt und die Glattheit nicht über die Kante hinweg propagiert werden sollte. Somit ergibt sich:

$$\mathbf{W}_{grad} = \begin{pmatrix} g_y^2 & -g_x g_y \\ -g_x g_y & g_x^2 \end{pmatrix}$$

Glattheit senkrecht zu den Hauptkrümmungen

Falls die Grauwertfunktion in eine Richtung stark gekrümmt ist, so sollte Glattheit nur senkrecht zu dieser Richtung gefordert werden. An jedem Punkt der Grauwertfunktion gibt es zwei zueinander senkrecht stehende Hauptkrümmungsrichtungen. In der einen Richtung ist die Krümmung maximal, in der anderen Richtung minimal. Man kann also fordern, daß die Glattheit in Richtung einer Hauptkrümmung mit der Krümmung senkrecht dazu gewichtet wird. Seien λ und κ die Hauptkrümmungen in den Richtungen \mathbf{l} und \mathbf{k} , so haben wir die Wichtungsmatrix:

$$\begin{aligned} W_{curv} &= W_{\mathbf{l}} + W_{\mathbf{k}} \\ &= \kappa^2 \mathbf{l} \mathbf{l}^T + \lambda^2 \mathbf{k} \mathbf{k}^T \end{aligned}$$

Nach kurzer Rechnung ergibt sich daraus die Wichtungsmatrix zu:

$$W_{curv} = \begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix} \begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix}$$

Zusammenfassung der Terme

Wir können die beiden oben hergeleiteten Matrizen nun zu einem gewichteten Mittel zusammenfassen:

$$\mathbf{F} = \begin{pmatrix} g_y \\ -g_x \end{pmatrix} \begin{pmatrix} g_y \\ -g_x \end{pmatrix}^T + b^2 \begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix} \begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix}$$

Die Wichtungsmatrix \mathbf{W} ergibt sich dann aus einer Normierung der Matrix \mathbf{F} . Dazu gibt es mehrere Möglichkeiten. Normieren wir \mathbf{F} mit der Determinante, so erhalten wir:

$$\mathbf{W} = \frac{\mathbf{F}}{\det \mathbf{F}} = \mathbf{C}^{-1}$$

wobei wir für die letzte Gleichheit $b^2 = \overline{x^2}$ setzen müssen. \mathbf{C} ist dabei die in Gleichung 4.11 definierte Matrix, die uns schon bei der lokalen Konstanzannahme begegnet war. Die Berechnung des optischen Flusses mit diesem Ansatz ergibt

jedoch nicht das gewünschte Resultat. Eine Analyse zeigt, daß der Glattheitsterm nicht genügend Einfluß auf das Endergebnis hat.² Im Grenzübergang von einer gekrümmten zu einer geraden Kante tendiert der Einfluß des Glattheitsterms gegen Null [Nagel + Enkelmann 86, Seite 572f].

Aus diesem Grunde sind noch zwei weitere Möglichkeiten der Normierung untersucht worden. Eine Möglichkeit besteht darin, die Matrix \mathbf{F} mit ihrer Spur zu normieren, und eine dritte Variante läßt diese Matrix unnormiert [Nagel + Enkelmann 86].

In dem Artikel Nagel 87 wird allerdings auf den erweiterten Term \mathbf{W}_{curv} der Glattheitsforderung verzichtet. Man kann nämlich zeigen, daß die Glättungsoperation selber die Variationen zweiten Grades berücksichtigt, so daß z.B. der Glattheitsterm von Horn + Schunck 81 an Grauwertecken einen konstanten optischen Fluß liefert. Das Weglassen dieses Terms führt dann zu sehr viel einfacheren Algorithmen (die zweite Ableitung der Bildfunktion muß nicht mehr ermittelt werden). Mit dem einfacheren Glattheitsmaß ist es dann auch möglich, die in Gleichung 4.24 gemachte Annahme einer lokal konstanten Wichtungsmatrix fallenzulassen.

4.5 Zusammenfassung

Wir wollen die verschiedenen eingeschlagenen Wege auf dem Weg zur Lösung des Blendenproblems noch einmal kurz bewerten. Die Benutzung von zusätzlichem geometrischem Wissen ist nur in sehr wenigen Fällen möglich. Diese Methode liefert zwar gute Ergebnisse, jedoch ist die Anwendung auf Spezialfälle beschränkt.

Die Ausnutzung von nichtlinearen Variationen der Bildfunktion liefert für sich allein genommen naturgemäß nicht in jedem Bildbereich eine eindeutige Schätzung der Verschiebung. Es mußte die zusätzliche Annahme eines lokal konstanten Verschiebungsvektors gemacht werden, um durch die Verbindung der "motion constraint equations" in verschiedenen Punkten zu einer eindeutigen Lösung zu gelangen.

Um noch dichtere Verschiebungsvektorfelder zu erhalten, kann man die lokale Konstanz der Verschiebungsvektoren in eine globale Glattheit abschwächen. Die Nachteile solcher Verfahren stecken in der zusätzlichen Hypothese, die naturgemäß z.B. an Verdeckungskanten nicht immer erfüllt ist. Das Aufkommen der einfachen Glattheitsmaße, insbesondere der Artikel Horn + Schunck 81, haben die Forschung auf diesem Gebiet sehr inspiriert. Einen vorläufigen Endpunkt bilden die schon recht aufwendigen Algorithmen, die eine gerichtete Glattheitsforderung beinhalten.

²Der Wichtungsfaktor α^2 müßte Werte in der Größenordnung von 1000 zugewiesen bekommen.

Ein Ansatzpunkt, der bisher noch sehr wenig Beachtung gefunden hat, ist die Verwendung von multispektralen Bildern. Deshalb werden wir uns in den nächsten Kapiteln mit den Möglichkeiten beschäftigen, die Farbbilder für die Lösung des Blendenproblems bieten. Wir werden weiterhin von der "motion constraint equation" ausgehen.

Kapitel 5

Verschiebungsvektoren aus Farbbildern

Die Verwendung von Farbbildern zur Bestimmung des optischen Flusses ist bisher in der Literatur kaum an praktischen Beispielen erprobt worden. Dies mag daran liegen, daß andere Prioritäten gesetzt werden, oder, daß nicht überall die Möglichkeit besteht, Farbbildfolgen aufzunehmen und zu digitisieren. Es gibt aber vereinzelt theoretische Erörterungen, wie z.B. das 2-Farben-Theorem von *Blicher 83*, welches besagt, daß zur eindeutigen Lösung des Korrespondenzproblems ohne Zusatzwissen mindestens 2 Farben benötigt werden. Der Satz wird mit differentialgeometrischen Methoden bewiesen, sagt aber nichts über die praktische Verwendbarkeit von Farbbildern aus.

In den nächsten zwei Kapiteln sollen folgende Fragen geklärt werden:

- Inwieweit ist es möglich, existierende Verfahren zur Ermittlung von Verschiebungsvektorfeldern für die Verwendung von Farbbildern abzuwandeln? Erlaubt die Verwendung von Farbinformation grundsätzlich neue Verfahren, die bei Grauwertbildern nicht angewendet werden können?
- Welche Vorteile bieten Farbbilder gegenüber Grauwertbildern? Lohnt sich der zusätzliche Aufwand?

Um brauchbare Verfahren zur Verschiebungsvektorbestimmung aus Farbbildern zu erhalten, werden wir im ersten Abschnitt einige Möglichkeiten betrachten, wie man die Ansätze des vorherigen Kapitels auf Farbbilder übertragen könnte. Obwohl die Verfahren für RGB-Bilder beschrieben werden, sind sie auf keine spezielle Farbdarstellung beschränkt. Nur die anschauliche Interpretation ist auf einen dreidimensionalen Farbraum beschränkt. Insbesondere kann man die Algorithmen auch für Bilder im nicht sichtbaren Teil des elektromagnetischen Spektrums verwenden.

Bei der Optimierung spezieller Abstandsmaße ergeben sich zum Teil Verfahren, die nicht mehr auf der üblichen "motion constraint equation" beruhen und auf die

Intensitätsinformation der Farbbilder fast vollkommen verzichten.

Schließlich schlagen wir vor, skalare Größen zu berechnen, die etwas über die Güte jedes einzelnen Verschiebungsvektors aussagen.

5.1 Definitionen

Unter einem Farbbild im allgemeinen Sinne verstehen wir eine Reihe von Grauwertbildern der gleichen Szene, die zum gleichen Zeitpunkt in unterschiedlichen Spektralbereichen aufgenommen worden sind. Man kann diese Art von Farbbildern auch Multispektralbilder nennen. Bei Farbbildern im eigentlichen Sinne handelt es sich um drei Bilder, die dem roten, grünen und blauen Spektralbereich entsprechen. Diese Bilder nennen wir auch RGB-Bilder. Der Grund für die Verwendung von gerade 3 Anteilen des sichtbaren Spektralbereichs, ist im menschlichen Sehsystem zu finden. Experimente zeigen nämlich, daß man durch additive Kombination verschiedener Intensitäten der 3 Farbkanäle fast jeden gewünschten Farbeindruck erzeugen kann. Das bedeutet, daß sich die Evolution auf 3 Komponenten für das Farbsehen beschränkt hat. Dies kommt daher, daß die einzelnen Teile des Spektrums für natürliche Materialien nicht unabhängig voneinander sind. Man kann also den Schluß ziehen, daß im Bereich des sichtbaren Lichtes durch zusätzliche Komponenten kaum weitere Informationen aus Realweltbildern zu gewinnen ist (siehe auch den Abschnitt 6.3). Für die rechnergestützte Auswertung von Bildern kann man aber auch beliebig viele andere Spektralbereiche auswählen, auch im nicht sichtbaren Teil des elektromagnetischen Spektrums. Im folgenden werden wir unsere Untersuchungen auf die Verwendung von RGB-Bildern beschränken, auch wenn die vorgeschlagenen Algorithmen sich leicht auf andere Fälle von multispektralen Bildern anpassen lassen.

Wir beschreiben ein Farbbild durch eine Vektorfunktion. An jedem Punkt des Bildes erhalten wir einen Farbvektor mit den Intensitäten der Farben Rot, Grün und Blau. Jedes Pixel besteht nun aus einem Zeilenvektor $\mathbf{c} = (c_R, c_G, c_B)$.

5.2 Von Grauwertbildern zu Farbbildern

Es gibt mehrere Möglichkeiten, Algorithmen für Grauwertbilder auf die Verarbeitung von Farbbildern anzupassen.

1. Man erzeugt aus dem Farbbild ein Grauwertbild und wendet dann ein Verfahren für Grauwertbilder auf dieses Bild an. Man kann z.B. das arithmetische Mittel aller drei Farbkanäle bilden, welches der Intensität des Farbvektors entspricht. Im allgemeinen führt dies zu rauschärmeren Bildern; die eigentliche Farbinformation wird jedoch nicht in den Verarbeitungsprozeß einbezogen.
2. Man wendet den Algorithmus auf alle drei Farbkanäle getrennt an und berechnet dann aus den drei einzelnen Ergebnissen ein Gesamtergebnis. Dies

könnte z.B. das arithmetische Mittel der Verschiebungsvektoren sein oder auch ein Verfahren, welches nur diejenigen Vektoren ins Gesamtergebnis übernimmt, welche mindestens in zwei Kanälen innerhalb einer gewissen Toleranz bestätigt worden sind.

3. Man berechnet in jedem Kanal getrennt mittels der "motion constraint equation" eine Bedingung für den Verschiebungsvektor. Sodann bestimme man denjenigen Verschiebungsvektor, der alle Bedingungsgleichungen möglichst gut erfüllt. Diesen kann man z.B. durch die Minimierung des quadratischen Fehlers berechnen. Dieses Verfahren hat im Gegensatz zu den bereits aufgezählten Möglichkeiten den Vorteil, daß innerhalb eines Farbkanals zusätzlich zur "motion constraint equation" keine weiteren Annahmen gemacht werden müssen, da bereits drei Gleichungen pro Pixel existieren. Somit sollte das Blendenproblem hier nicht auftauchen. Es entsteht auch kein Widerspruch zwischen der Forderung der Linearität einerseits und der Änderung der Gradientenrichtung andererseits, wie es bei Verfahren besteht, die auf der Konstanzannahme beruhen.
4. Man überlegt sich ein geeignetes Abstandsmaß für Farbvektoren und minimiere eine Funktion des Abstandes zwischen einem Pixel im ersten Bild und dem verschobenen Pixel im zweiten Bild. Als Ergebnis erhält man ähnlich wie bei der Minimierung der quadratischen Grauwertdifferenz eine Bedingungsgleichung für den Verschiebungsvektor.

Die aufgezählten Verfahren nutzen die Farbinformation in der Reihenfolge der Aufzählung immer direkter aus. Während bei den ersten beiden Möglichkeiten der Algorithmus für Grauwertbilder in einer Phase der Verarbeitung vollkommen unverändert eingesetzt wird, ergeben sich aus den weiteren Möglichkeiten neue Verfahren, die auch spezielle Eigenschaften von Farbinformationen nutzen können.

Die zweite Möglichkeit wird hier nicht weiter erörtert, da anzunehmen ist, daß die Ergebnisse von Verfahren, die auf dem dritten oder vierten Vorschlag beruhen, die Farbinformation besser nutzen können, ohne deswegen aufwendiger zu sein. Auch in das dritte Verfahren lassen sich weitere Annahmen, wie lokale Konstanz des Vektorfeldes einbauen. Während aber bei der zweiten Möglichkeit zunächst aus einer Reihe von Bedingungsgleichungen, den "motion constraint equations", ein Verschiebungsvektor bestimmt wird und dann im nachfolgenden Schritt ein zusammenfassendes Ergebnis berechnet werden muß, gehen bei der dritten Möglichkeit alle Bedingungsgleichungen auf einmal in den Verarbeitungsprozeß ein. Stellt man die "motion constraints" durch Geraden im u - v -Raum dar, so wird bei Vorschlag 2 aus einigen Geraden jeweils ein optimaler Schnittpunkt berechnet und dann die Schnittpunkte gemittelt. Bei der dritten Möglichkeit werden aber alle Geraden zur Ermittlung des optimalen Schnittpunktes herangezogen.

Es sollen jetzt entsprechend den beiden letzten Möglichkeiten Verfahren für Farbbilder entwickelt werden.

5.2.1 Optimierung der “motion constraints”

Verwendet man die “motion constraint equations” der einzelnen Farbkanäle, um den Verschiebungsvektor zu bestimmen, so stimmt dieses Verfahren formal vollkommen mit dem Verfahren der Pseudoinversen aus Abschnitt 4.3.2 überein. Wir müssen nur den Grauwertvektor \mathbf{g} durch den Farbvektor \mathbf{c} ersetzen. Das Ergebnis sei hier nochmals kurz angegeben:

$$\mathbf{u} = -(\nabla\mathbf{c}(\nabla\mathbf{c})^T)^{-1}\nabla\mathbf{c}\mathbf{c}_i^T = -(\nabla\mathbf{c})^\dagger\mathbf{c}_i^T \quad (5.1)$$

Dabei bezeichnet $(\nabla\mathbf{c})^\dagger = (\nabla\mathbf{c}(\nabla\mathbf{c})^T)^{-1}\nabla\mathbf{c}$ wiederum die Pseudoinverse von $\nabla\mathbf{c}$. Wir werden die zu invertierende Matrix $(\nabla\mathbf{c}(\nabla\mathbf{c})^T)$ mit dem Buchstaben \mathbf{E} abkürzen. Diesen Ansatz werden wir im folgenden auch “Pseudoinversionsansatz” nennen. Da die drei Farbauszüge denselben Ort zur selben Zeit zeigen, ist die Annahme eines gleichen Verschiebungsvektors für alle Auszüge an jedem Pixel gerechtfertigt. Trotz der formalen Ähnlichkeit mit dem Verfahren aus Abschnitt 4.3.2 hat dieses Verfahren den großen Vorteil, daß keine zusätzlichen Annahmen über die lokale Konstanz des Verschiebungsvektors gemacht werden müssen, wenn die Gradienten der einzelnen Farbkanäle unterschiedlich genug sind. Sollte die Pseudoinverse nicht existieren, d.h., die Ableitungen der einzelnen Farbkanäle sind voneinander abhängig, so kann man zusätzlich immer noch die Annahme eines lokal konstanten Verschiebungsvektors machen und diese Annahme problemlos in den Algorithmus einbeziehen.

Der zusätzliche Zeitaufwand für dieses Verfahren wird in Abschnitt 5.5 anhand der benötigten Multiplikationen abgeschätzt.

5.2.2 Optimierung eines Farbabstandes

Minimieren wir analog zum Vorgehen bei Grauwertbildern nach Gleichung 3.15 den euklidischen Farbabstand zwischen den Farbvektoren in aufeinanderfolgenden Bildern, wobei die lokale Umgebung des einen Bildes durch eine lineare Approximation angenähert wird

$$M \approx (\mathbf{c}_2 - \mathbf{c}_1 + (\nabla\mathbf{c}_1)^T\mathbf{u})^2 \rightarrow \text{Minimum} \quad (5.2)$$

so erhalten wir statt der Gleichung 3.16 die neue Vektorgleichung

$$\nabla\mathbf{c}_1(\nabla\mathbf{c}_1)^T\mathbf{u} = \mathbf{E}\mathbf{u} = -\nabla\mathbf{c}_1(\mathbf{c}_2 - \mathbf{c}_1)^T \quad (5.3)$$

Dies ist genau die Gleichung 5.1, die wir oben bei der Minimierung der quadratischen Fehler der einzelnen “motion constraint equations” bekommen haben.

Man kann sich jedoch auch andere Möglichkeiten der Verwendung von Farbinformation denken. Der quadratische Fehler, der oben minimiert wurde, kann auch anders geschrieben werden:

$$|\mathbf{c}_2(\mathbf{x}_0) - \mathbf{c}_1(\mathbf{x}_0 - \mathbf{u})|^2 = |\mathbf{c}_2(\mathbf{x}_0)|^2 + |\mathbf{c}_1(\mathbf{x}_0 - \mathbf{u})|^2 - 2\mathbf{c}_2(\mathbf{x}_0)\mathbf{c}_1^T(\mathbf{x}_0 - \mathbf{u})$$

Nimmt man an, daß die Bildenergien $|\mathbf{c}_1|^2$ und $|\mathbf{c}_2|^2$ überall gleich sind, so wird der euklidische Abstand minimiert, wenn der Term $\mathbf{c}_2(\mathbf{x}_0) \mathbf{c}_1^T(\mathbf{x}_0 - \mathbf{u})$ maximiert wird. Da diese Annahme nicht immer gerechtfertigt ist, kann man den Term auch mit den jeweiligen Bildenergien normieren.

$$M = \frac{\mathbf{c}_2(\mathbf{x}_0) \mathbf{c}_1^T(\mathbf{x}_0 - \mathbf{u})}{|\mathbf{c}_2(\mathbf{x}_0)| |\mathbf{c}_1(\mathbf{x}_0 - \mathbf{u})|} \rightarrow \text{Maximum} \quad (5.4)$$

Anschaulich läuft dies auf die Minimierung des Winkels zwischen den Farbvektoren $\mathbf{c}_2(\mathbf{x}_0)$ und $\mathbf{c}_1(\mathbf{x}_0 - \mathbf{u})$ hinaus. Wir werden dieses Verfahren "Korrelationsansatz" nennen. Der Vorteil dieses Ansatzes kann darin liegen, daß die Länge des Farbvektors für das Optimum nicht relevant ist und daher ein Helligkeitsunterschied zwischen aufeinanderfolgenden Bildern kaum eine Rolle spielt.

Entwickeln wir wiederum \mathbf{c}_1 in eine Taylorreihe erster Ordnung:

$$\mathbf{c}_1(\mathbf{x}_0 - \mathbf{u}) \approx \mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u}$$

Die erste Ableitung von M nach \mathbf{u} ergibt dann den Ausdruck:

$$\frac{\partial M}{\partial \mathbf{u}} \approx (\nabla \mathbf{c}_1) \left[\frac{(\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u})^T (\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u}) - |\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u}|^2 \mathbf{I}}{|\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u}|^3} \right] \frac{\mathbf{c}_2^T}{|\mathbf{c}_2|}$$

Durch Nullsetzen erhalten wir:

$$(\nabla \mathbf{c}_1) \left[\frac{(\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u})^T (\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u})}{|\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u}|^2} \right] \mathbf{c}_2^T = (\nabla \mathbf{c}_1) \mathbf{c}_2^T \quad (5.5)$$

Den Operator in den eckigen Klammern kann man als Projektionsoperator ansehen. Er beläßt von dem Vektor, auf den er angewendet wird, nur den Anteil, der in die Richtung von $(\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u})$ zeigt. Die obige Bedingung wird also dann erfüllt, wenn durch die Projektion von \mathbf{c}_2 auf diese Richtung nur der Anteil senkrecht zu den Vektoren \mathbf{c}_{1_x} und \mathbf{c}_{1_y} verändert wird. Wenn es einen Wert für \mathbf{u} gibt, so daß \mathbf{c}_2 und $\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u}$ in der Richtung übereinstimmen, so ist dies der gesuchte Verschiebungsvektor.

5.2.3 Interpretation im Farbraum

Bei RGB-Bildern gibt es noch eine andere, anschaulichere Interpretation der beiden besprochenen Optimierungsprobleme. Die Gleichung:

$$\mathbf{c}_1(\mathbf{x}_0 - \mathbf{u}) = \mathbf{c}_1 - u \mathbf{c}_{1_x} - v \mathbf{c}_{1_y}$$

kann man auch als mit den Parametern u und v parametrisierte Fläche im RGB-Farbraum auffassen (siehe Abbildung 5.1). Die Ebene geht durch den Punkt \mathbf{c}_1 und wird durch die Vektoren \mathbf{c}_{1_x} und \mathbf{c}_{1_y} aufgespannt.

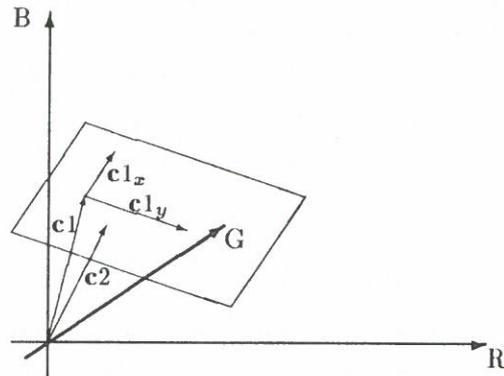


Abbildung 5.1: Optimierung eines Farbabstandes. Der Abstand zwischen c_2 und der Ebene, die durch die Vektoren c_{1x} und c_{1y} aufgespannt wird, muß minimiert werden.

Die Optimierung von

$$M = (\mathbf{c}_2 - \mathbf{c}_1 + (\nabla \mathbf{c}_1)^T \mathbf{u})^2 \rightarrow \text{Minimum}$$

läuft dann auf die Suche nach dem kürzesten euklidischen Abstand zwischen der Ebene und dem Punkt c_2 heraus. Die Tatsache, daß die kürzeste Verbindungsgerade zwischen der Ebene und dem Punkt c_2 senkrecht auf der Ebene steht, liefert die beiden Bedingungen

$$\begin{aligned} (\mathbf{c}_1 - u\mathbf{c}_{1x} - v\mathbf{c}_{1y} - \mathbf{c}_2)^T \mathbf{c}_{1x} &= 0 \\ (\mathbf{c}_1 - u\mathbf{c}_{1x} - v\mathbf{c}_{1y} - \mathbf{c}_2)^T \mathbf{c}_{1y} &= 0 \end{aligned}$$

Dies entspricht genau der Gleichung 5.3. Wird der Winkel α zwischen den die Ebene aufspannenden Vektoren c_{1x} und c_{1y} zu klein, so spannen diese beiden Vektoren keine Ebene, sondern nur noch eine Gerade im RGB-Raum auf. Die Tatsache, daß die kürzeste Verbindungsgerade zwischen der Geraden und dem Punkt c_2 senkrecht auf der Geraden liegt, liefert uns nur noch eine Nebenbedingung. Dies zeigt sich dann bei der Invertierung der Matrix $\mathbf{E} = \nabla \mathbf{c}_1 (\nabla \mathbf{c}_1)^T$, denn die Determinante dieser Matrix ist:

$$\det \mathbf{E} = (\mathbf{c}_{1x} \times \mathbf{c}_{1y})^2 = |\mathbf{c}_{1x}|^2 |\mathbf{c}_{1y}|^2 \sin^2 \alpha$$

Sie verschwindet also für kleine Winkel α .

Auch die zweite Optimierungsaufgabe (5.4) kann man sich so veranschaulichen. Wie bereits erwähnt, wird das Maximum genau dann erreicht, wenn der Vektor $\mathbf{c}_1 - (\nabla \mathbf{c}_1)^T \mathbf{u}$ parallel ist zum Vektor c_2 . Das bedeutet, daß wir nur den

Durchstoßpunkt der durch den Vektor \mathbf{c}_2 definierten und durch den Nullpunkt gehenden Geraden mit der Ebene bestimmen müssen, um den gesuchten Verschiebungsvektor zu erhalten. Wir suchen also eine Lösung für das Gleichungssystem

$$\mathbf{c}_1 - u \mathbf{c}_{1_x} - v \mathbf{c}_{1_y} = \lambda \mathbf{c}_2$$

mit der Lösung:

$$\begin{pmatrix} \lambda \\ u \\ v \end{pmatrix} = \begin{pmatrix} \mathbf{c}_2 \\ \mathbf{c}_{1_x} \\ \mathbf{c}_{1_y} \end{pmatrix}^{-1} \mathbf{c}_1^T$$

Dabei gibt λ den Faktor an, mit dem man \mathbf{c}_2 multiplizieren muß, um einen Punkt zu erhalten, der in der Ebene liegt. Ideal ist es also, wenn λ den Wert 1 hat, denn dann liegt der Punkt \mathbf{c}_2 auf der Ebene und es gibt somit einen Punkt in der lokalen Umgebung um \mathbf{x}_0 im ersten Bild, der die gleiche Farbe hat, wie der Punkt \mathbf{x}_0 im zweiten Bild. Die zu invertierende Matrix wird singulär, wenn $\mathbf{c}_{1_x} \parallel \mathbf{c}_{1_y}$ ist, d.h., die Gradientenrichtungen in allen Kanälen gleich sind. Die Ebene ist zu einer Geraden entartet. Probleme bereitet ebenfalls der Fall, in dem \mathbf{c}_2 parallel zur Ebene liegt, denn dann gibt es ebenfalls keinen Schnittpunkt zwischen der Ebene und der Geraden.

5.3 Lineare Farbtransformationen

Wie wir schon aus Gleichung 4.10 in Abschnitt 4.3.2 gesehen haben, ist die Matrix $\mathbf{E} = [\nabla \mathbf{c} (\nabla \mathbf{c})^T]$ entscheidend für die Güte der Schätzung der Verschiebungsvektoren. Wie aus Abschnitt 4.3.4 ersichtlich wurde, reicht es für die zuverlässige Schätzung eines Verschiebungsvektors nicht aus, daß die Determinante von \mathbf{E} eine gewisse Größe hat. Es ist vorteilhaft, wenn die Eigenwerte dieser Matrix möglichst gleich groß sind. Um dies zu erreichen, müssen die Gradienten in den einzelnen Farbkanälen in verschiedene Richtungen zeigen.

Um zu sehen, ob sich das Aussehen von \mathbf{E} durch eine lineare Farbtransformation verbessern läßt, wollen wir die Abhängigkeit von \mathbf{E} vom verwendeten Farbkoordinatensystem klären. Sei \mathbf{T} eine Transformationsmatrix, die einen Farbvektor in ein anderes Koordinatensystem überführt:

$$\mathbf{c}' = \mathbf{c}\mathbf{T}$$

Dann ergibt sich die neue Matrix \mathbf{E}' zu:

$$\begin{aligned} \mathbf{E}' &= [\nabla \mathbf{c}' (\nabla \mathbf{c}')^T] \\ &= [\nabla \mathbf{c} \mathbf{T} (\nabla \mathbf{c} \mathbf{T})^T] \\ &= (\nabla \mathbf{c}) \mathbf{T} \mathbf{T}^T (\nabla \mathbf{c})^T \end{aligned}$$

Für orthonormale Transformationen, also Drehungen im Farbraum, gilt $\mathbf{E} = \mathbf{E}'$. Es ist auch anschaulich einleuchtend, daß eine Drehung des Koordinatensystems im Farbraum keinen Einfluß auf die Genauigkeit des Verfahrens haben kann.

Bei nicht-orthogonalen Transformationen ist die Analyse etwas schwieriger. Wir führen die Winkel α, β, γ zwischen den neuen Basisvektoren und dem Vektor \mathbf{c}_x ein. Die Vektoren α', β' und γ' sind die entsprechenden Winkel zum Vektor \mathbf{c}_y . Falls wir annehmen, daß das neue Basissystem aus Einheitsvektoren besteht, so stellt sich die neue Matrix \mathbf{E}' wie folgt dar:

$$\mathbf{E}' = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

mit den Abkürzungen:

$$\begin{aligned} a &= |\mathbf{c}_x|^2 (\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma) \\ b &= |\mathbf{c}_x| |\mathbf{c}_y| (\cos \alpha \cos \alpha' + \cos \beta \cos \beta' + \cos \gamma \cos \gamma') \\ c &= |\mathbf{c}_y|^2 (\cos^2 \alpha' + \cos^2 \beta' + \cos^2 \gamma') \end{aligned}$$

Ohne Einschränkung der Allgemeinheit können wir annehmen, daß die beiden Vektoren \mathbf{c}_x und \mathbf{c}_y senkrecht aufeinander stehen, denn dies können wir durch drehen des Bildkoordinatensystems immer erreichen. Um die Spur und die Determinante der neuen Matrix \mathbf{E}' anschaulich interpretieren und damit auch abschätzen zu können, stelle man sich nun folgende Operation im RGB-Raum vor. Die 3 Basisvektoren der Transformation werden auf die Ebene projiziert, die durch die Vektoren \mathbf{c}_x und \mathbf{c}_y aufgespannt wird. Danach verzerren wir diese Ebene, indem wir sie entlang der Richtung von \mathbf{c}_x um den Faktor $|\mathbf{c}_x|$ und entlang der Richtung von \mathbf{c}_y um den Faktor $|\mathbf{c}_y|$ strecken. Der Spur von \mathbf{E}' entspricht nun die Summe der Längen dieser drei projizierten und verzerrten Basisvektoren. Die Determinante erhält man, indem man die Summe der Flächen der drei von jeweils zwei dieser Vektoren aufgespannten Parallelelogramme berechnet. Die Spur kann also höchstens den dreifachen Wert des längeren der beiden Vektoren \mathbf{c}_x bzw. \mathbf{c}_y annehmen, während die Determinante auf höchstens den zweifachen Wert der Determinante von \mathbf{E} beschränkt ist.

Es hat nun jedoch keinen Sinn, in jedem Einzelfall nach derjenigen Transformation zu suchen, die zu möglichst gleich großen Eigenwerten der Matrix \mathbf{E}' führt, denn die Transformation sollte einen anschaulichen Sinn haben, der vor der Anwendung festgelegt wird. Die obige Abschätzung sollte nur ein Gefühl dafür vermitteln, was bei einer solchen Transformation mit der den Fehler stark beeinflussenden Matrix \mathbf{E} geschieht. So kann es sein, daß man aus bestimmten Gründen eine bestimmte Komponente eines Farbvektors nicht berücksichtigen will. Es kann z.B. vorkommen, daß sich ein Objekt in einen Schatten hineinbewegt, wobei nur die Intensität der Pixel abnimmt, deren Farbe sich jedoch nicht ändert. In diesem Fall ist die Information, die die Intensität des Farbvektors liefert, unerwünscht, da sich daraus falsche Verschiebungen ergeben würden.

Um nun gezielt die Intensitätsinformation aus dem Algorithmus zu eliminieren, drehen wir das Koordinatensystem so, daß eine Achse in Richtung der Intensität

zeigt. Dies geschieht z.B. mit der Matrix:

$$\mathbf{T} = \frac{1}{\sqrt{6}} \begin{pmatrix} -2 & 0 & \sqrt{2} \\ 1 & \sqrt{3} & \sqrt{2} \\ 1 & -\sqrt{3} & \sqrt{2} \end{pmatrix}$$

Wir können nun die dritte Komponente der Transformation dadurch vernachlässigen, daß wir sie einfach auf Null setzen. Wir erhalten auf diese Weise die Transformation der Matrix \mathbf{E} zu

$$\mathbf{E}' = \frac{1}{3}(\nabla \mathbf{c}) \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} (\nabla \mathbf{c})^T$$

5.4 Gütekriterien

Wir wollen nun versuchen, Größen zu finden, die sich aus einer Bildfolge berechnen lassen und etwas über die Güte der Verschiebungsschätzungen aussagen. Mit einem solchen Gütemaß könnte man unzuverlässige Verschiebungsvektoren erkennen und für die weitere Analyse der Bildfolge unberücksichtigt lassen.

Zunächst bietet sich an, die Abweichung (das Residuum) von der "motion constraint equation" zu berechnen, denn bei einem großen Fehler in der "motion constraint equation" liegt entweder ein extremes Rauschen vor oder das Modell einer linearen Bildfunktion innerhalb des Verschiebungsgebietes ist falsch. In beiden Fällen ist der berechnete Verschiebungsvektor unzuverlässig. Sei $\hat{\mathbf{u}}$ der geschätzte Verschiebungsvektor, dann ist das Residuum r gegeben durch:

$$r = (\nabla g)^T \hat{\mathbf{u}} + g_t$$

Das Residuum in dieser Form sagt allerdings nur etwas über den Fehler der Komponente in Richtung des Gradienten aus, denn eine Veränderung der Schätzung senkrecht zu dieser Richtung verändert den Wert von r nicht. Bei Farbbildern erhalten wir auf diese Weise einen Residuumsvektor:

$$\mathbf{r} = (\nabla \mathbf{c})^T \hat{\mathbf{u}} + \mathbf{c}_t \approx (\nabla \mathbf{c}1)^T \hat{\mathbf{u}} + \mathbf{c}2 - \mathbf{c}1$$

Dies ist genau der Abstandsvektor zwischen dem Punkt $\mathbf{c}1 - (\nabla \mathbf{c}1)^T \hat{\mathbf{u}}$ auf der oben beschriebenen Ebene im Farbraum und dem Punkt $\mathbf{c}2$. Im Falle der Pseudoinversion ist \mathbf{r} auch der kürzeste Abstandsvektor zwischen der Ebene und dem Punkt $\mathbf{c}2$. Es erscheint daher sinnvoll, die euklidische Norm als skalares Maß der Zuverlässigkeit zu verwenden.

Von einem anderen Standpunkt aus betrachtet, kommen wir zu einem weiteren möglichen Gütemaß. Die Optimierung mittels Pseudoinverse kann im u - v -Raum als Suche nach dem besten Schnittpunkt aller "constraint lines" betrachtet werden.

Der Abstand einer "constraint line" vom geschätzten Verschiebungsvektor beträgt nun:

$$d = \frac{|(\nabla g)^T \hat{\mathbf{u}} + g_t|}{|\nabla g|} = \frac{|r|}{|\nabla g|}$$

Es besteht also ein enger Zusammenhang zwischen Residuum und dem Abstand zur "constraint line". Nun kann man bei Farbbildern oder der Benutzung einer lokalen Umgebung die Summe der Abstände zu allen beteiligten "constraint lines" als Gütemaß verwenden.

Die bisher betrachteten Gütemaße haben den Nachteil, daß sie erst nach der Berechnung des Verschiebungsvektors bestimmt werden können. Den Fehleranalysen in Abschnitt 3.3.3 konnten wir entnehmen, daß das Verhältnis von Gradientenstärke zur Krümmung in Verschiebungsrichtung einen Einfluß auf den systematischen Fehler der "motion constraint equation" hat. Eine Abschätzung dieser Größen könnte also auch als Kriterium für die Güte der Schätzung herangezogen werden.

Als weiteres Gütekriterium, welches bereits aus einem Bild berechnet werden kann, können wir die Determinante der Matrix \mathbf{E} verwenden. Wie wir in Abschnitt 4.3.4 sahen, reicht eine große Determinante alleine für die Bestimmung eines in allen Richtungen gleichermaßen zuverlässigen Verschiebungsvektors nicht aus; ausschlaggebend ist vielmehr, daß die Eigenwerte möglichst gleich groß sind. Dies kann durch das Verhältnis q von Determinante zum Quadrat der Spur von \mathbf{E} abgeschätzt werden, ohne daß eine Diagonalisierung explizit durchgeführt werden müßte.

Schließlich ist ein geschätzter Verschiebungsvektor, der aus dem Bereich herauszeigt, in dem die Ableitungen der Bildfunktion mittels Operatormaske geschätzt werden, mit Vorsicht zu genießen, denn die Gradientenstärke ist dann über den Rand des Gradientenoperators hinaus extrapoliert worden. Da die Grau- oder Farbwerte außerhalb des Operatorfensters nicht zur Schätzung des Gradienten herangezogen werden und nicht sichergestellt ist, daß die Gradienten sich lokal nur schwach ändern, kann eine solche Extrapolation natürlich zu großen Fehlern führen. Die Gleichungen 3.11 und 4.16 zeigen ebenfalls, daß lange Verschiebungsvektoren fehleranfällig sind, denn der systematische Fehlerterm ist proportional zum Quadrat des Verschiebungsvektors. Verschiebungsvektoren, deren Längen die Operatorweite erheblich überschreiten, sollten für die Weiterverarbeitung nicht verwendet werden.

5.5 Aufwandsabschätzung

Hier soll noch eine kurze Abschätzung des Aufwandes gemacht werden, der für die Minimierung des euklidischen Farbabstandes getrieben werden muß. Dazu nehmen wir an, wir hätten ein Bild mit n^2 Pixeln und verwenden zum Ableiten einen separierbaren $(2k+1) \times (2k+1)$ Operator. Für die Faltung eines Grauwertbildes mit einem dieser Operatoren brauchen wir $2n^2(2k+1)$ Multiplikationen

(siehe Abschnitt 3.3.3). Wir brauchen die Ableitungen in x und y -Richtung, sowie eine Mittelung über das erste und zweite Bild, um die zeitliche Ableitung bilden zu können. Das ergibt das folgende Ergebnis für die Bildung der erforderlichen Gradienten:

$$\begin{array}{ll} \text{Grauwertbild} & : \quad 8n^2(2k+1) \text{ Multiplikationen} \\ \text{Farbbild} & : \quad 24n^2(2k+1) \text{ Multiplikationen} \end{array}$$

Haben wir nun einen Grauwert- oder Farbvektor \mathbf{g} mit m Komponenten, so benötigen wir zur Bildung von $\nabla \mathbf{g}(\nabla \mathbf{g})^T$ und $\nabla \mathbf{g} \mathbf{g}_t^T$ pro Pixel $5m$ Multiplikationen. Wenn man sich hingegen die Werte $g_x(\mathbf{x}_i)^2$, $g_x(\mathbf{x}_i)g_y(\mathbf{x}_i)$, usw. merkt, dann braucht jeder dieser Werte nur einmal pro Pixel berechnet zu werden. Für die Invertierung der Matrix $\nabla \mathbf{g}(\nabla \mathbf{g})^T$ benötigt man nochmals 2 Multiplikationen sowie 3 Divisionen.

Insgesamt erhalten wir somit die folgende Abschätzung der Anzahl der Multiplikationen für die Berechnung des Verschiebungsvektorfeldes:

$$\begin{array}{ll} \text{Grauwertbild} & : \quad n^2[8(2k+1) + 7] \text{ Multiplikationen} \\ \text{Farbbild} & : \quad n^2[24(2k+1) + 17] \text{ Multiplikationen} \end{array}$$

Bei der Verwendung von 5×5 -Operatoren benötigen wir also beim Farbverfahren die 2.91-fache Anzahl von Multiplikationen gegenüber einem Grauwertverfahren.

Kapitel 6

Vergleich der Verfahren

Wir wollen nun die Ergebnisse unserer Verfahren zur Schätzung von Verschiebungsvektorfeldern miteinander vergleichen. Dabei soll der Versuch gemacht werden, die Ergebnisse zu erklären und zu bewerten.

6.1 Verwendete Bildfolgen

Die Brauchbarkeit der Verfahren kann auf mehrere Weisen getestet werden. Voraussetzung ist, daß das zu betrachtende Verschiebungsvektorfeld bekannt ist und mit den geschätzten Ergebnissen verglichen werden kann. So könnte man etwa daran denken, für diesen Zweck künstliche Bildserien per Rechner zu erzeugen, die Abbildungen bewegter dreidimensionaler Objekte darstellen. Dies hat den Vorteil, daß man im Anschluß an die Berechnung des Verschiebungsfeldes auch die Brauchbarkeit für einen Rekonstruktionsalgorithmus testen kann. Ein Nachteil ist allerdings der relativ große Aufwand, der zur Erzeugung einigermaßen realer Bilder getrieben werden muß. Da wir untersuchen wollen, ob sich Farbbilder für die Berechnung des optischen Flusses und zur Lösung des Blendenproblems eignen, ist es besonders wichtig, Farbbilder zu haben, die den realen Verhältnissen möglichst nahe kommen. Insbesondere sollten die lokalen Farbstrukturen und die Zusammenhänge zwischen den Gradienten in den verschiedenen Farbkanälen den natürlichen Verhältnissen entsprechen, denn diese Größen sind für die Lösung des Blendenproblems ausschlaggebend. Künstlich erzeugte Bilder kommen für diese Zwecke nicht in Betracht.

Wir werden deshalb eine viel einfachere, aber für unsere Untersuchungen ausreichende Methode verwenden, indem wir ein Farbbild einfach um eine bekannte Anzahl von Pixeln verschieben. Dies entspricht genau dem Bewegungsmodell aus dem wir die "motion constraint equation" mittels Optimierung hergeleitet haben. Ein weiterer Vorteil dieser Vorgehensweise ist, daß die Möglichkeit besteht, statistische Aussagen über die Güte der Messungen zu machen und in Form von Häufigkeitsverteilungen darzustellen, weil an allen Stellen der gleiche Verschiebungsvektor zu erwarten ist. So können wir auch den Einfluß der in Abschnitt 5.4

vorgeschlagenen Gütekriterien auf die Verteilung der geschätzten Verschiebungsvektoren untersuchen und darstellen (siehe Abschnitt 6.5).

Es muß allerdings auch bedacht werden, daß bestimmte wichtige Probleme wie z.B. Verdeckungen so nicht getestet werden können. An Stellen, wo die Bewegung der Farbwerte von dem zugrundegelegten Bewegungsmodell stark abweicht, wie es bei Verdeckungen der Fall ist, erwarten wir auch keine Verbesserungen gegenüber Verfahren, die mit Grauwertbildern arbeiten.

Wir werden zwei auf diese Art hergestellte Bildfolgen betrachten. Die Bilder 1 und 2 zeigen die Ausgangsbilder der beiden Bildfolgen "Blocksbild" und "Apfelbild". Das zweite Bild wurde jeweils um ein Pixel nach rechts und ein Pixel nach unten verschoben.

Das Apfelbild 2 ist auch das erste Bild einer realen Bildfolge. Bild 3 zeigt das zweite Bild dieser Reihe. Die Objekte wurden relativ zur Kamera in etwa horizontal verschoben. Die Ergebnisse, die mit dieser Bildfolge erzielt worden sind, werden zum Abschluß ebenfalls vorgestellt.

Alle räumlichen Ableitungen werden mit einem 5×5 -Beudet-Operator geschätzt (siehe Seite 21, bzw. 49). Die zeitlichen Ableitungen werden durch die Differenz des arithmetischen Mittels über eine 5×5 -Umgebung in zwei aufeinanderfolgenden Bildern ermittelt.

6.2 Reine Farbverfahren

Zunächst sollen hier die beiden vorgeschlagenen Verfahren für Farbbilder miteinander verglichen werden. Unter optimalen Bedingungen, bei denen einerseits in jedem Farbkanal eine lineare Bildfunktion vorliegt und andererseits in allen Farbkanälen unterschiedliche Gradientenrichtungen vorherrschen, ergeben beide Verfahren die richtigen Ergebnisse. Falls jedoch eine dieser Bedingungen nicht erfüllt wird, ergeben sich unterschiedliche Verschiebungsvektorschätzungen. Das Verfahren, welches die Farbkorrelation maximiert, stellt sich dabei als wesentlich schlechter heraus. Dies kommt hauptsächlich daher, daß nur zwei Parameter der Farbvektoren, nämlich die zwei Winkel, die die Orientierung des Farbvektors angeben, für die Optimierung genutzt werden. Wesentlich stabiler reagiert das Verfahren, welches den euklidischen Farbabstand mittels Pseudoinversion minimiert. Durch die Überbestimmtheit des Gleichungssystems wirken sich statistische Fehler nicht so stark auf das Endergebnis aus. Allerdings hat dieses Verfahren Schwierigkeiten, wenn die Gradientenrichtungen in allen Farbkanälen gleich sind. In diesem Falle ist die Konditionierung des Gleichungssystems schlecht, weil alle Gleichungen linear abhängig sind. Dies führt dazu, daß das Blendenproblem nicht gelöst werden und nur die Komponente in Richtung des Gradienten zuverlässig bestimmt werden kann.

In Diagramm 1 sind die Gradientenrichtungen in den verschiedenen Farbkanälen des Blocksbildes 1 gegeneinander aufgetragen. Man sieht deutlich, daß der Hauptanteil der Punkte auf der Diagonalen liegt. Es besteht also eine große Korre-

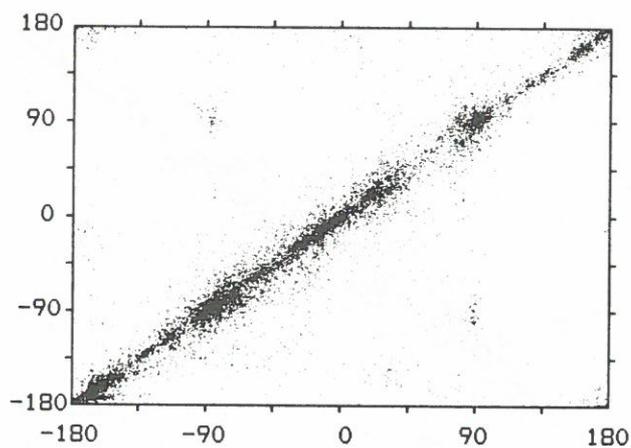
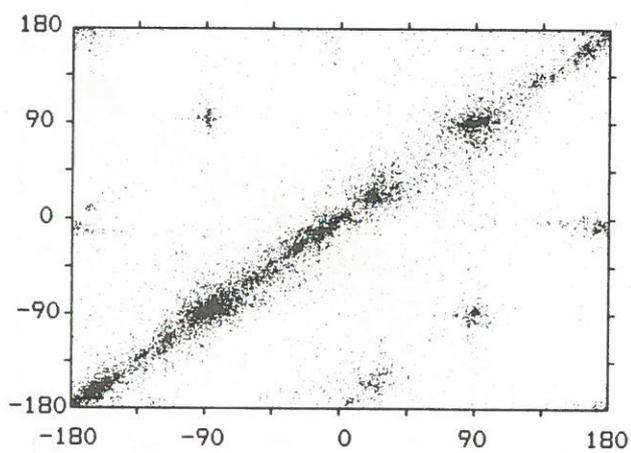
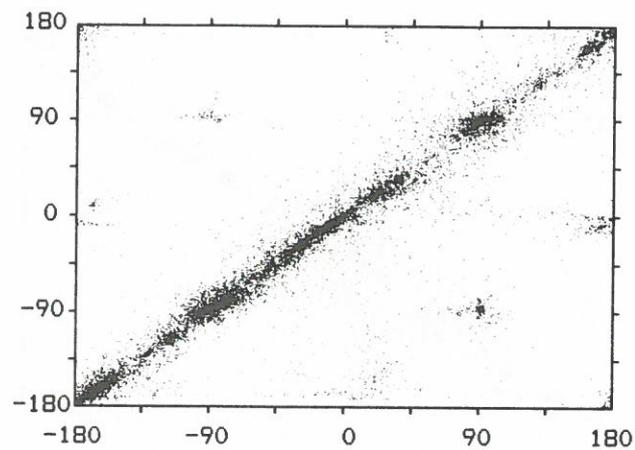


Diagramm 1: Gradientenrichtungen des Blockbildes 1 in den verschiedenen Farbkanäle gegeneinander aufgetragen. Von oben nach unten: Rotkanal gegen Grükanal, Rotkanal gegen Blaukanal und Grükanal gegen Blaukanal.

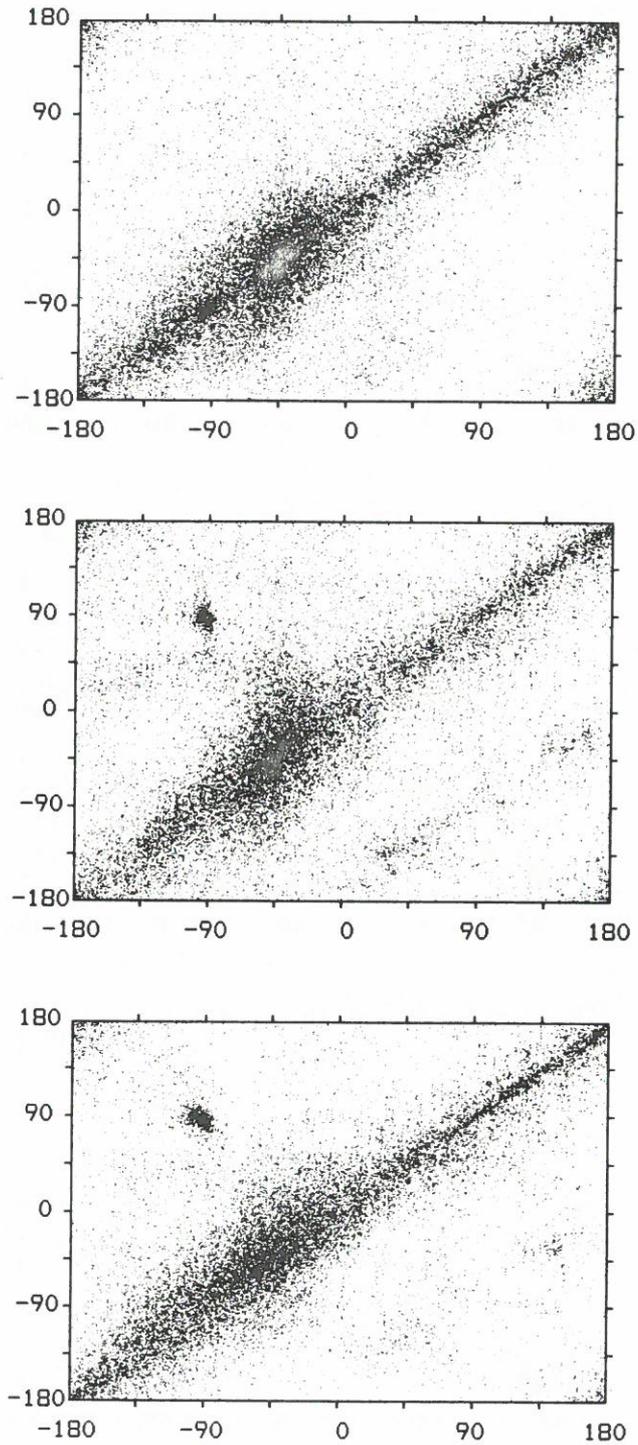


Diagramm 2: Gradientenrichtungen des Apfelbildes 2 in den verschiedenen Farbkanäle gegeneinander aufgetragen. Von oben nach unten: Rotkanal gegen Grünkanal, Rotkanal gegen Blaukanal und Grünkanal gegen Blaukanal.

lation zwischen den Gradientenrichtungen in den verschiedenen Farbkanälen. Aus diesem Grund handelt es sich beim Blocksbild um ein für Farbverfahren weniger geeignetes Bild. Wie das Diagramm 2 zeigt, sind die Korrelationen beim Apfelbild 2 nicht so stark. Die Bilder 4 bis 6 zeigen die verschiedenen Farbauszüge. Man kann deutliche Unterschiede in der Texturierung finden. Dies ist sehr vorteilhaft für die Algorithmen, die allein auf die Farbinformation angewiesen sind. Mit diesem Bild wurden deshalb auch die besten Resultate erzielt.

Die Korrelationen zwischen den Farbkanälen zeigen sich auch in den Eigenwerten der Matrix \mathbf{E}^{-1} . Abbildung 6.2 zeigt die Matrizen \mathbf{E}^{-1} innerhalb eines Ausschnitts des Blocksbildes 6.1 als Ellipsen. Die Längen der Hauptachsen entsprechen den Eigenwerten von \mathbf{E}^{-1} und deren Lagen den jeweiligen Eigenvektorrichtungen. Die größere Hauptachse der Ellipsen ist jeweils auf eine Einheitsgröße normiert, so daß die absoluten Längen der Achsen nichts über die Güte aussagen. Je kleiner der Abstand vom Mittelpunkt zum Ellipsenrand in einer Richtung ist, desto größer ist die Zuverlässigkeit des geschätzten Verschiebungsvektors in dieser Richtung. Es gibt also Fälle, in denen der Unterschied in den Eigenwerten extrem ist und darum nur eine Komponente des Verschiebungsvektors zuverlässig bestimmt werden kann.

Bevor wir uns dem Vergleich des Pseudoinversionsansatzes mit anderen Verfahren zuwenden, wollen wir noch weiter nach den Gründen für die schlechten Ergebnisse des Korrelationsansatzes suchen.

6.3 Farbinformationen in Realweltbildern

Wir wollen nun genauer untersuchen, woran es liegen könnte, daß die durch Maximierung der Farbkorrelation erzielten Ergebnisse im Allgemeinen nicht brauchbar sind. Dazu soll untersucht werden, wo die Hauptinformation in den Farbbildern steckt. Um dies zu erkennen, muß man diejenige Farbtransformation suchen, die in unkorrelierten Farbvektoren resultiert. Dies läuft auf eine Karhunen-Loève-Transformation der Farbvektoren hinaus [Rosenfeld + Kak 82, Seite 123ff]. Die Korrelationsmatrix Σ ist folgendermaßen zu berechnen:

$$\Sigma = \frac{1}{n-1} \sum_{\mathbf{x}_i} (\mathbf{c}(\mathbf{x}_i) - \bar{\mathbf{c}})^T (\mathbf{c}(\mathbf{x}_i) - \bar{\mathbf{c}})$$

Die Summe erstreckt sich dabei über die Pixel des ganzen Bildes. $\bar{\mathbf{c}}$ ist der über das Bild gemittelte Farbvektor. Um die Transformationsmatrix zu bestimmen, müssen wir das Eigenwertproblem

$$\Sigma \mathbf{t}^i = \sigma_i^2 \mathbf{t}^i$$

lösen, wobei \mathbf{t}^i der i -te Transformationsvektor ist. Die Eigenwerte σ_i^2 sind dann die Varianzen der Farbkanäle \mathbf{t}^i . Je größer der Eigenwert, desto größer ist die Varianz und damit auch die Information in diesem Kanal. Eigene Untersuchungen

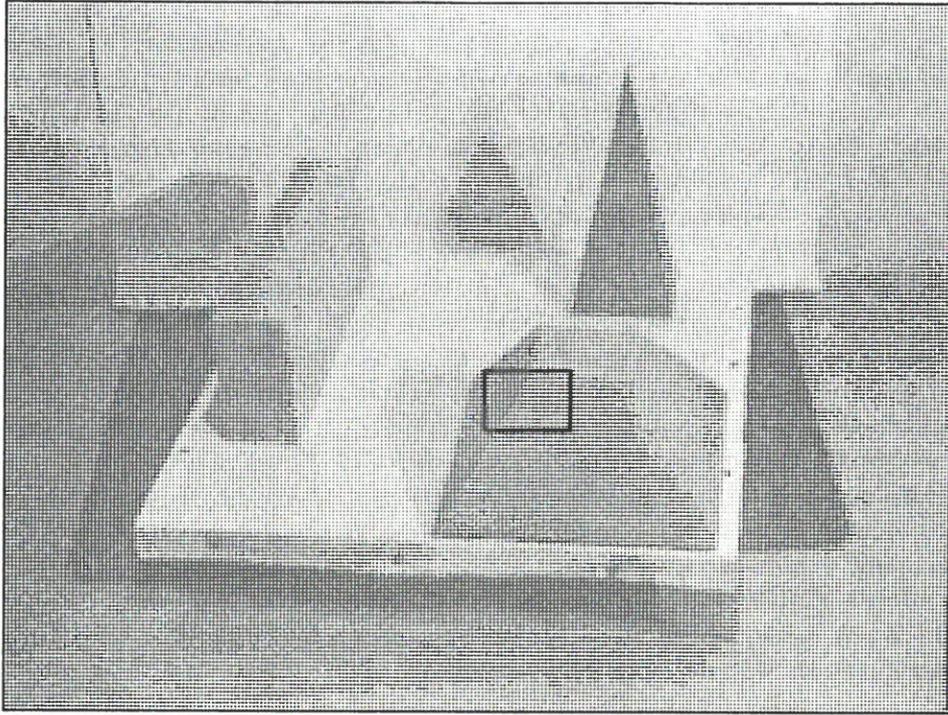


Abbildung 6.1: Blocksbild mit eingezeichnetem Ausschnitt.

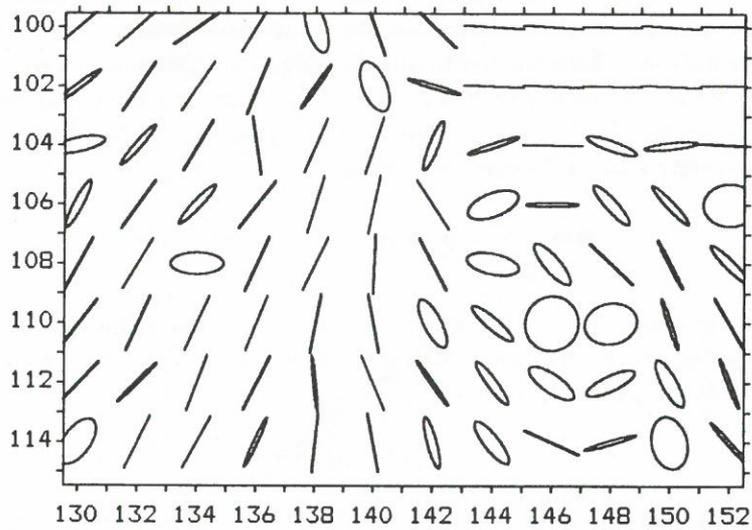


Abbildung 6.2: Darstellung der Matrizen E^{-1} in einem Ausschnitt des Blocksbildes (Erklärung siehe Text Seite 69).

an verschiedenen Farbbildern zeigen nun das auf den ersten Blick erstaunliche Ergebnis, daß sich die berechneten Transformationen recht ähnlich sind. Die durchschnittliche Transformationsmatrix $\bar{\mathbf{T}}$ hat etwa folgendes Aussehen:

$$\bar{\mathbf{T}} = \begin{pmatrix} 0.57 & 0.61 & -0.53 \\ 0.57 & 0.15 & 0.80 \\ 0.58 & -0.76 & -0.26 \end{pmatrix}$$

Die erste und wichtigste Komponente in der Matrix entspricht der Intensität. Der Eigenwert liegt um ein Vielfaches über dem der anderen Komponenten. Auch der Eigenwert des dritten Kanals liegt nochmals unter dem Eigenwert des zweiten Kanals.

Ähnliche Ergebnisse erzielen auch [Otha u.a. 80], die die Karhunen-Loève-Transformation für die Segmentation eines Bildes in verschiedene Bereiche nutzen. Sie stellen fest, daß man praktisch kaum schlechtere Ergebnisse erhält, wenn man die Information des letzten Kanals einfach wegläßt. Daß dies auch gilt, wenn wir auf die Spektralverteilung selber zurückgreifen und nicht von einem RGB-Bild ausgehen, zeigt Gershon 87. Er erwähnt eine Arbeit von Cohen, der diejenigen 3 Funktionen berechnet hat, mit denen sich die Spektralverteilung von 150 "Munsell Chips" optimal anpassen läßt. Diese Funktionen bilden nun die Basisvektoren für einen dreidimensionalen Vektorraum. Der Anteil der ersten Funktion am Spektrum enthält bereits 92.72% der Information und die beiden ersten Funktionen zusammen sogar 97.25%. Die erste Funktion entspricht wiederum der Intensität. Die zweite Kurve entspricht in etwa der Wichtung "2 · Grün - Rot - Blau" und der dritten Kurve kann man schließlich die Kombination "Rot - Blau" zuordnen. Dies entspricht ziemlich genau den Ergebnissen der Karhunen-Loève-Transformation.¹ Gershon hat auch die Daten von 370 natürlichen Materialien wie Erde, Vegetation oder Wasser nach diesen Funktionen entwickelt. Er erhält die drei Standardabweichungen 0.183, 0.028 und 0.027. Auch diese Ergebnisse zeigen also, daß die meiste Information in der Intensität enthalten ist.

Dies erklärt, warum mit dem Korrelationsansatz nicht so gute Ergebnisse zu erzielen sind, wie mit dem Pseudoinversionsansatz. Gerade die Hauptinformationsquelle, nämlich die Intensität, wird bei der Farbkorrelation kaum genutzt. Im Gegensatz dazu werden beim Pseudoinversionsansatz die Informationen aller Farbkanäle gleich genutzt.

6.4 Vergleich verschiedener Verfahren

Man kann den Ansatz der Minimierung des euklidischen Farbabstandes relativ einfach mit der Annahme eines lokal konstanten Verschiebungsvektors kombinieren. Dazu faßt man einfach alle Pixel und Farbwerte zu einem einzigen Vektor

¹Die Reihenfolge der beiden letzten Basisvektoren ist allerdings unterschiedlich.

zusammen und rechnet ansonsten genauso wie bisher. Bei einer lokalen Umgebung von 3×3 -Pixeln hat man dann bereits 27 lineare Gleichungen, die mittels Pseudoinversion optimal gelöst werden können. Alle folgenden Ergebnisse sind mit einer solchen 3×3 -Umgebung berechnet worden. Im Gegensatz zu dem bisher betrachteten Pseudoinversionsansatz auf jedem einzelnen Farbpixel, welches wir mit dem Namen "reines Farbverfahren" versehen wollen, werden auf dieses Verfahren mit dem Stichwort "kombiniertes Verfahren" verweisen. Wir werden den Namen \mathbf{E} für die zu invertierende Matrix beibehalten. Die Erweiterung läßt sich auch mit einem Verfahren vergleichen, welches allein die Intensität der Farbbilder verwendet. Bei dieser Methode berechnen wir die Verschiebungsvektoren mittels Annahme eines in einem 3×3 -Fenster konstanten Feldes; sie entspricht damit einer Realisierung des ersten Vorschlages aus Abschnitt 5.2. Wir werden diese Methode auch "Intensitätsverfahren" nennen.

Die mit dem kombinierten Verfahren erzeugten Verschiebungsvektoren sind sowohl besser gegenüber dem reinen Farbverfahren als auch gegenüber der reinen lokalen Konstanzannahme im Intensitätsbild. Dies zeigt Diagramm 3, in dem die für Bild 1 geschätzten Verschiebungsvektoren im u - v -Raum eingezeichnet worden sind. Berechnet man die Varianz σ^2 der Verschiebungsvektoren, so ergeben sich in u -Richtung die Werte $\sigma^2 = 15.1$ für das reine Farbverfahren, $\sigma^2 = 5.8$ für das Intensitätsverfahren und $\sigma^2 = 1.3$ für das kombinierte Verfahren. Dabei wurden alle Verschiebungsvektoren berücksichtigt, auch die offensichtlichen Ausreißer.

Man erkennt deutlich, daß die Fehler in den Verschiebungsvektoren mit den vorherrschenden Gradientenrichtungen zusammenhängen, denn die Verteilung der Verschiebungen ist nicht rotationssymmetrisch. An einer Kante, die entlang der x -Achse ausgerichtet ist, liefert der Algorithmus das richtige Ergebnis für v , aber das Resultat für u variiert stark. Somit entstehen im u - v -Raum Linienstrukturen entlang der u -Achse.

Die Gradientenrichtung in einem Farbbild soll in Verallgemeinerung zum Grauwertbild die Richtung sein, in der die Differenzvektoren den größten Betrag haben. Diese Richtung stimmt auch mit der Richtung des Eigenvektors mit dem größten Eigenwert von \mathbf{E} überein. Die Erweiterung des Verfahrens verkleinert nun vor allem den Fehler senkrecht zu dieser Gradientenrichtung. Dies zeigen die auf dem Apfelbild 2 berechneten Diagramme 4 und 5 sehr deutlich. Auf der x -Achse ist jeweils die Gradientenrichtung angegeben. Diese Richtung entspricht auch der stärksten Änderung der Farbvektoren bzgl. des euklidischen Abstandes. Auf der y -Achse ist nun genau die Verschiebungskomponente in dieser Richtung (Diagramm 4), bzw. senkrecht dazu (Diagramm 5) aufgetragen. Die Richtung ist dabei so gewählt, daß die berechneten Komponenten der Verschiebung positiv sind, d.h. bei negativer Verschiebung wird die Richtung um 180 Grad gedreht.

Die oberen Diagramme zeigen jeweils die Resultate für die reinen Farbverfahren. Darunter sind die Ergebnisse aufgetragen, die auf den Intensitätsbildern gerechnet wurden, und in den unteren Diagrammen sind die mit dem kombinierten Verfahren berechneten Verschiebungen eingetragen.

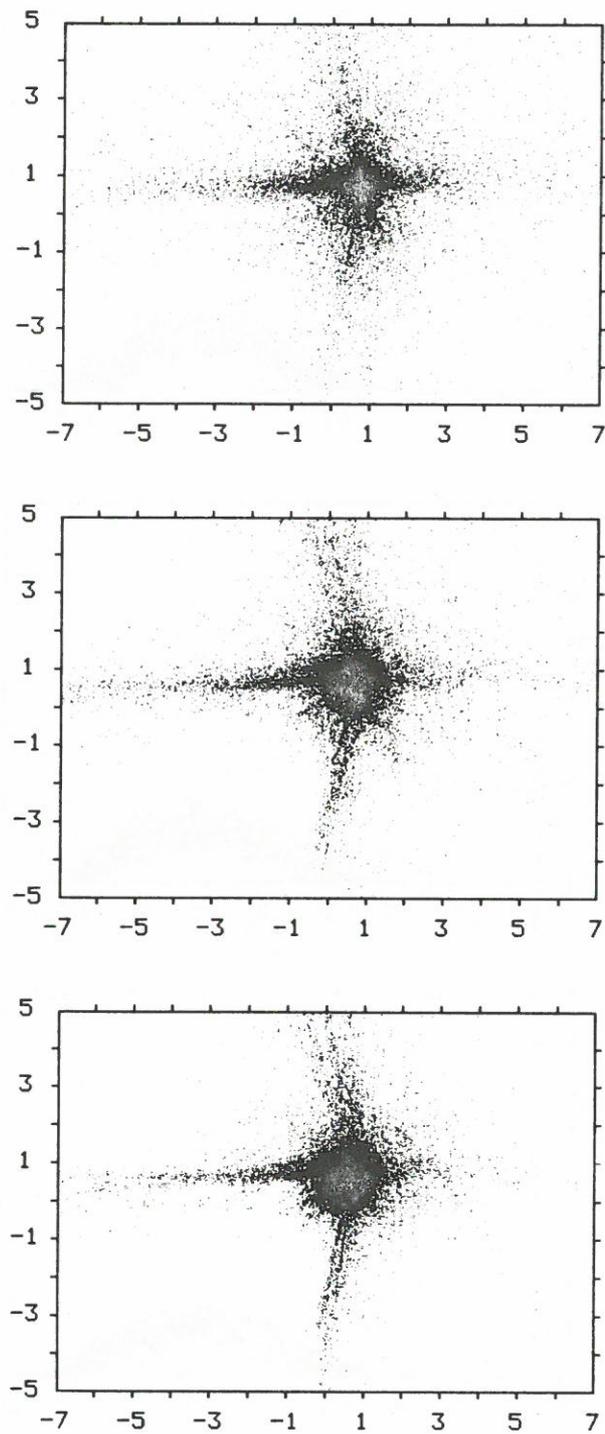


Diagramm 3: Für das Blocksbild berechnete Verschiebungsvektoren im u - v -Raum. Von oben nach unten: reiner Farbalgorithmus, Intensitätsverfahren und kombiniertes Verfahren. Der korrekte Verschiebungsvektor liegt bei $(u, v) = (1, 1)$.

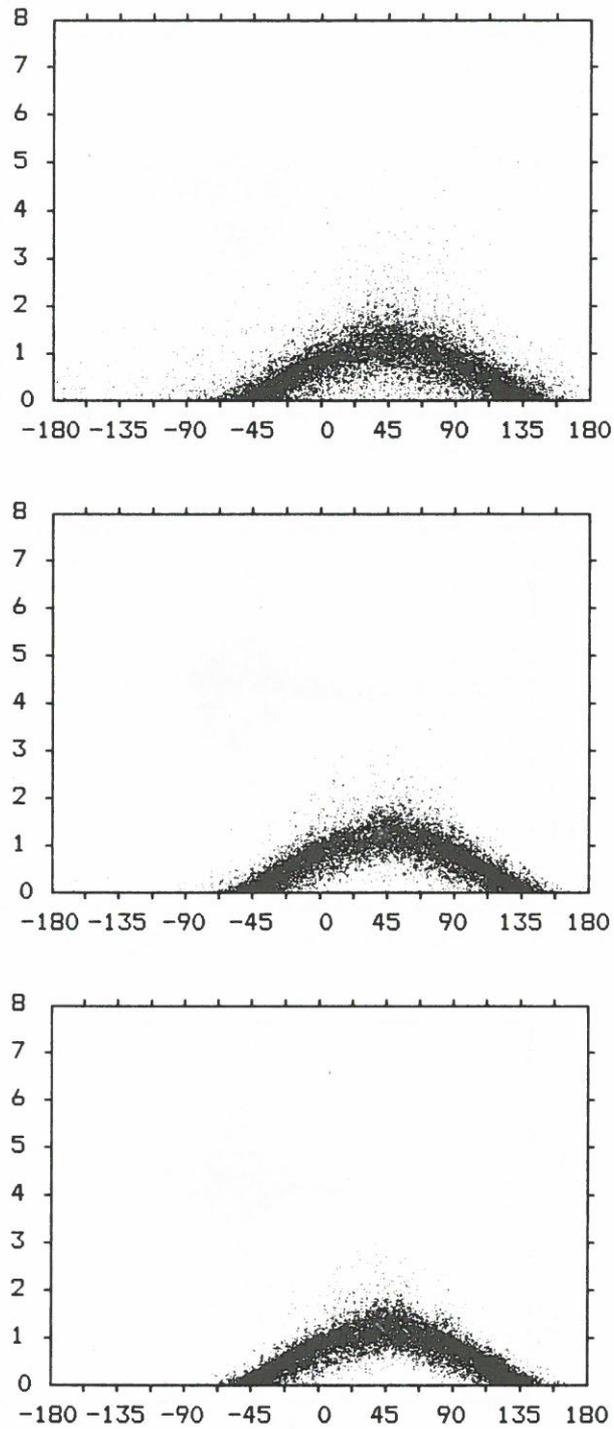


Diagramm 4: Die Normalkomponente des Verschiebungsvektors gegen die Gradientenrichtung aufgetragen. Von oben nach unten: reiner Farbalgorithmus, Intensitätsverfahren und kombiniertes Verfahren.

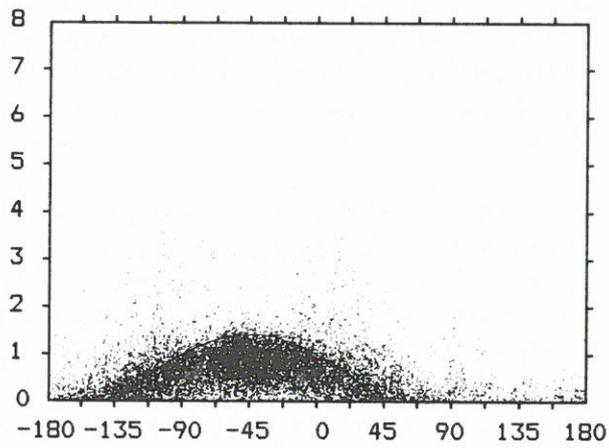
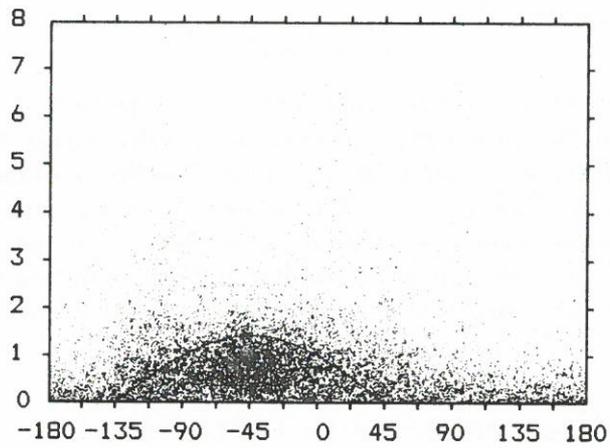
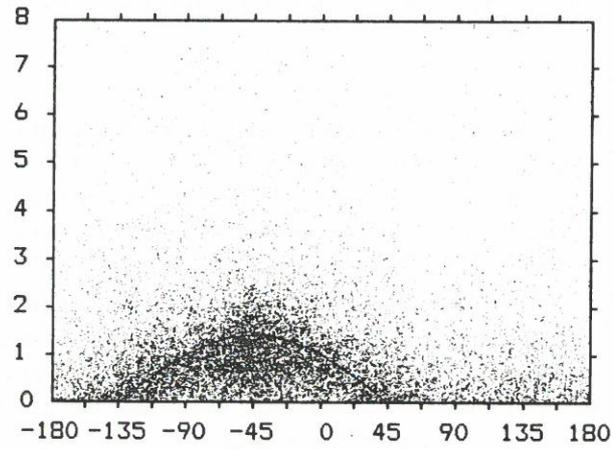


Diagramm 5: Die Parallelkomponente des Verschiebungsvektors gegen die Gradientenrichtung aufgetragen. Von oben nach unten: reiner Farbalgorithmus, Intensitätsverfahren und kombiniertes Verfahren.

Das Ergebnis, welches im Idealfall zu erwarten wäre, ergibt sich aus der in Gleichung 3.2 vorgestellten "motion constraint equation" in Polarform:

$$u^\perp = \rho \cos(\alpha - \beta)$$

Dieser Gleichung können wir sofort entnehmen, daß der Zusammenhang zwischen Gradientenrichtung α und Normalkomponente u^\perp im Idealfall durch eine Cosinusfunktion mit dem Maximum bei $\alpha = \beta$ beschrieben wird. Dabei gibt β die Richtung der Verschiebung und ρ den Geschwindigkeitsbetrag an. In unserem Fall hat der Winkel β den Wert 45 Grad und der Betrag der Verschiebung ist $\sqrt{2}$. Die erwartete Cosinuskurve ist in Diagramm 4 mit eingezeichnet, allerdings zum größten Teil verdeckt. Man sieht, daß die Fehler bei allen Verfahren nicht allzu groß und etwa gleich sind. Die Tatsache, daß die Verschiebungen im Mittel als zu klein eingestuft werden, läßt darauf schließen, daß Fehler bei den Modellannahmen zu leicht kleineren Verschiebungsvektoren führen.

Für die Komponente in Kantenrichtung, d.h. senkrecht zum Gradienten, ergibt sich der Zusammenhang:

$$u^\parallel = \rho \sin(\alpha - \beta)$$

Hier liegt an der Stelle $\alpha = \beta$ ein Nullpunkt. An Diagramm 5 ist zu erkennen, daß in dieser Richtung schon wesentlich größere Fehler auftreten als in Gradientenrichtung. Das reine Farbverfahren liefert dabei die schlechtesten Ergebnisse, da nur 3 Gleichungen zur Verfügung stehen. Die mit dem Intensitätsverfahren geschätzten Verschiebungen sind schon etwas besser, werden jedoch vom kombinierten Verfahren übertroffen. Trotzdem bleibt der Fehler in dieser Richtung wesentlich größer als senkrecht dazu. Auch hier fällt auf, daß die Verschiebungsvektoren eher zu klein geschätzt werden.

6.5 Die Aussagekraft der Gütekriterien

Um die Aussagekraft der in Abschnitt 5.4 vorgeschlagenen Gütemaße zu untersuchen, sind in den Diagrammen 6 bis 9 die Werte der Maße (x -Achse) gegen die geschätzten Verschiebungen in x -Richtung (y -Achse) des Blockbildes 1 aufgetragen. Es sei nochmals darauf hingewiesen, daß der korrekt Verschiebungsvektor im ganzen Bild die Größe (1,1) hat. Alle Diagramme wurden mit dem kombinierten Verfahren berechnet.

In Diagramm 6 sehen wir zunächst den Zusammenhang zwischen dem Quadrat des Gradienten und der Verschiebung in x -Richtung. Man erkennt deutlich, daß die kleineren Gradienten im Blockbild überwiegen. Leider nehmen die Abweichungen vom korrekten Verschiebungsvektor nicht mit stärker werdendem Gradienten ab. Der Gradient allein läßt sich daher kaum als Güteparameter einsetzen. Nach Gleichung 3.10 sollte aber das Verhältnis von Gradient zu Krümmung in Verschiebungsrichtung den Fehler der Komponente des Verschiebungsvektors in Gradientenrichtung bestimmen. Auch in den Gleichungen 4.16 kommt dieser

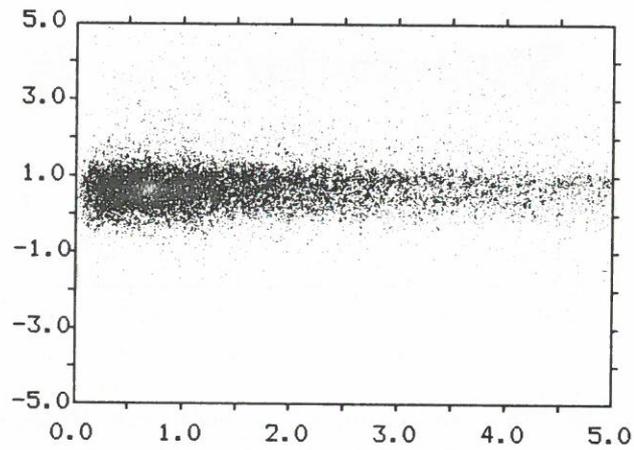


Diagramm 6: Das Quadrat des Gradientenbetrages gegen die gemessene Verschiebung in x -Richtung.

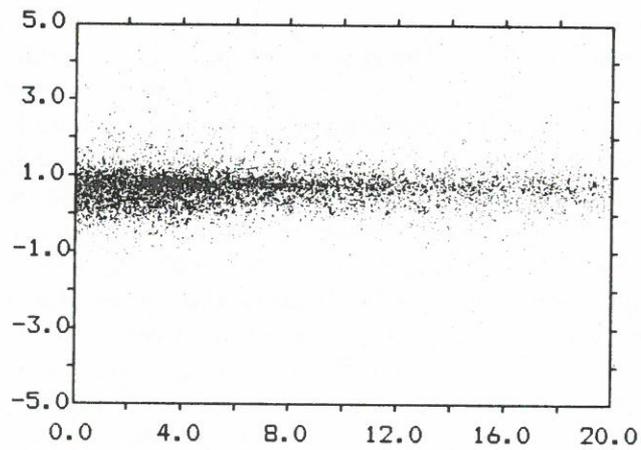


Diagramm 7: Das Verhältnis des Gradientenbetrages zur Krümmung gegen die gemessene Verschiebung in x -Richtung.

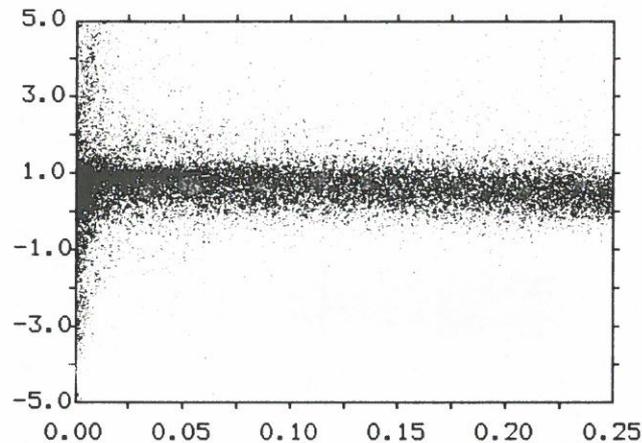


Diagramm 8: Das Verhältnis q von Determinante zum Quadrat der Spur der Matrix E gegen die gemessene Verschiebung in x -Richtung.

Term vor. Wie man allerdings im Diagramm 7 sieht, scheint eine Normierung des Gradienten mit der durch die Beudet-Operatoren geschätzten Krümmung in Verschiebungsrichtung kein besseres Maß für die Güte der Bilder zu sein.

In Diagramm 8 ist das Verhältnis q von Determinante und quadrierter Spur der Matrix E gegen die Verschiebung aufgetragen. Dieses Verhältnis wurde schon in Abschnitt 4.3.4 als Gütemaß vorgeschlagen, weil es das Verhältnis der Eigenwerte der Matrix E bzw. C approximiert (siehe Gleichung 4.14). Wie nun aus dem Diagramm 8 deutlich zu ersehen ist, treten bei sehr kleinen Werten von q relativ starke Variationen in der Verschiebung auf. Das Verhältnis eignet sich also als Gütemaß für die geschätzten Vektoren. Aber auch, wenn das Maß q unter eine gewisse Schwelle fallen sollte, so ist der Verschiebungsvektor nicht unbedingt vollkommen unbrauchbar, denn es ist nicht ausgeschlossen, daß die Komponente in Richtung des größten Eigenwerts von E einen korrekten Wert hat, der für die Weiterverarbeitung genutzt werden könnte. So kann man in den Bildern auch beobachten, daß die Verschiebungsvektoren an Kanten, an denen naturgemäß ein kleines Verhältnis q vorhanden ist, oft falsch liegen, die Komponenten senkrecht zu Kanten aber trotzdem brauchbar sind.

Auch das in Abschnitt 5.4 besprochene Residuum r sagt etwas über die Zuverlässigkeit der Verschiebungsvektoren aus. In Diagramm 9 ist die Summe der Quadrate (oben) und der Absolutbeträge (unten) der Residuen jeder beteiligten "motion constraint equation" aufgetragen. Die ungenauen Verschiebungsvektoren werden mit größer werdendem Residuum häufiger, wobei sich die Summe der Absolutbeträge besser für eine Schwellwertbildung zu eignen scheint. Wie wir ebenfalls

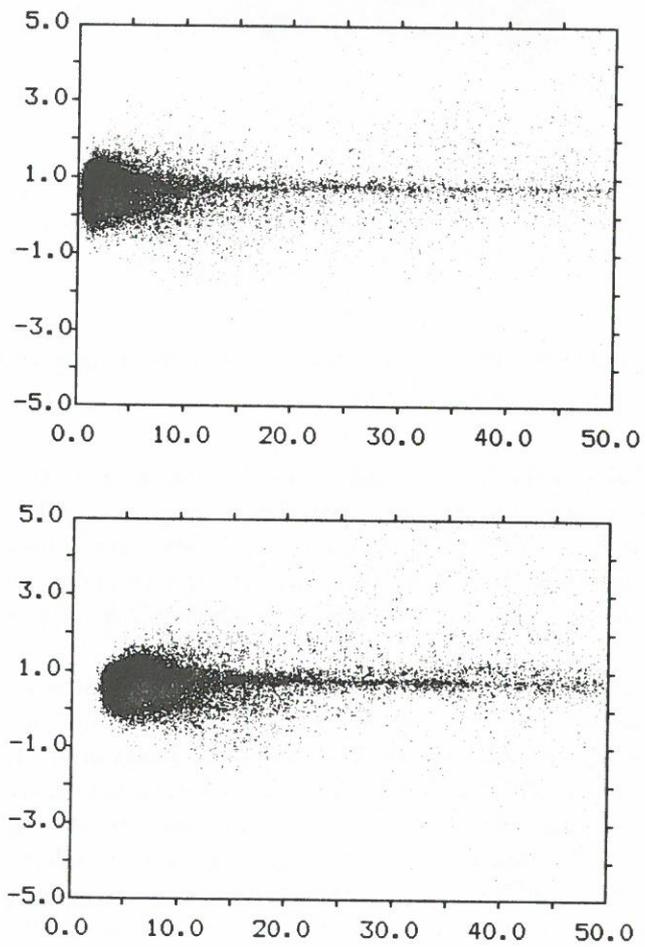


Diagramm 9: Die Summe der quadrierten Residuen (oben) bzw. die Summe der Absolutbeträge der Residuen (unten) gegen die gemessene Verschiebung in x -Richtung.

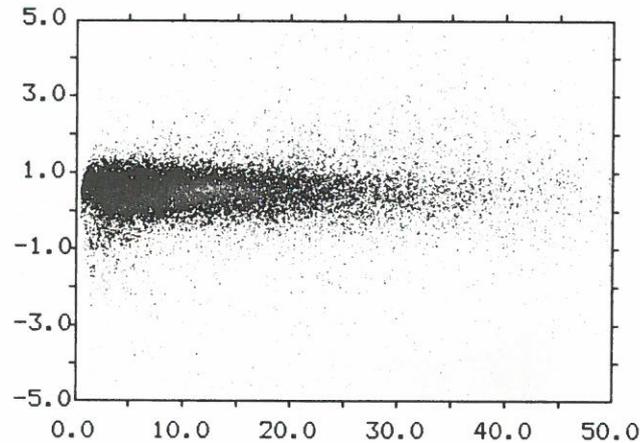


Diagramm 10: Die Summe der mit dem Gradientenbetrag normierten Residuen gegen die Verschiebung in x -Richtung.

in Abschnitt 5.4 gesehen haben, erhält man die Summe der Abstände d zu allen beteiligten "constraint lines", wenn man den Betrag eines jeden einzelnen Residuums r durch den Gradientenbetrag teilt und über diese Quotienten summiert. Die Aussagekraft dieses Kriteriums läßt sich Anhand von Diagramm 10 erkennen. Auch hier wächst die relative Häufigkeit der falschen Vektoren mit zunehmender Größe von d .

Berücksichtigt man nur diejenigen Vektoren des mit dem kombinierten Verfahren berechneten Verschiebungsvektorfeldes, deren Gütemaße eine Mindestbedingung erfüllen, so erhält man die Verteilung von Diagramm 11. Hier wurde ein Mindestwert von 0.01 für das Verhältnis q von Determinante und quadrierter Spur der Matrix E verlangt. Außerdem wurde die Summe der Absolutbeträge der Residuen r auf den Maximalwert von 40.0 festgelegt (man vergleiche diese Werte in den Diagrammen 8 und 9). Die Linienstrukturen, die in Diagramm 3 zu erkennen waren, sind nun größtenteils verschwunden. Die Varianz der Verschiebung in x -Richtung sinkt damit von $\sigma^2 = 1.3$ auf nur noch $\sigma^2 = 0.36$. Es bleiben immerhin noch 66% der ursprünglichen Vektoren übrig.

6.6 Verschiebungsvektoren auf Realweltbildfolgen

Die drei zuletzt vorgestellten Verfahren sind auch auf eine Realweltbildfolge angewendet worden. Das erste Bild dieser Bildfolge ist das bekannte Apfelbild 2. Bild 3 zeigt die Szene zum nächsten Zeitpunkt. Die Äpfel wurden mitsamt Unterlage um etwas mehr als ein Pixel in annähernd horizontaler Richtung verschoben.

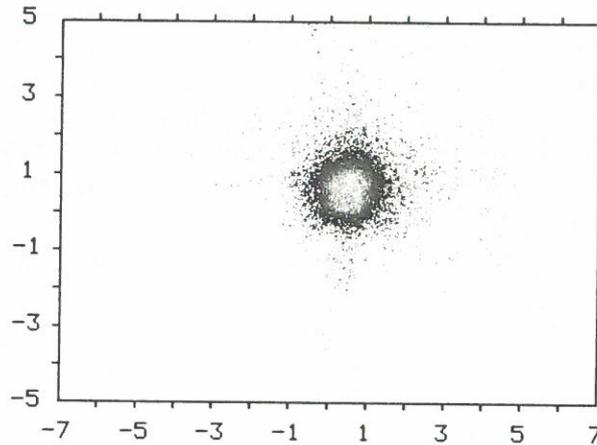


Diagramm 11: Die Verschiebungsvektoren, deren Güteparameter eine Mindestanforderung (siehe Text Seite 80) erfüllen, im u - v -Raum.

Die mit den verschiedenen Verfahren berechneten Verschiebungsfelder sind in den Abbildungen 6.4 bis 6.6 gezeichnet. Dabei wurden in jedem Bild Schwellen an die beiden folgenden Parameter gelegt:

- Das Verhältnis q von Determinante zum Quadrat der Spur der Matrix \mathbf{E} .
- Die Summe der Absolutbeträge der Residuen r der einzelnen "motion constraint equations" (im folgenden mit R bezeichnet).

Der Rechenausschnitt ist in Abbildung 6.3 eingezeichnet. Gezeigt wird immer nur jeder zweite Verschiebungsvektor, damit nicht so viele Überlappungen auftreten.

Abbildung 6.4 zeigt das Ergebnis des reinen Farbverfahrens. Die Bedingungen für die Gütemaße sind: $q \geq 0.03$ und $R \leq 40$. Die nächste Abbildung 6.5 zeigt das Ergebnis für das Intensitätsverfahren. Hier wurden die Schwellen so gesetzt, daß gilt: $q \geq 0.02$; $R \leq 40$. Schließlich zeigt Abbildung 6.6 das Resultat für das kombinierte Verfahren. Es sind nur dort Vektoren eingezeichnet, für die gilt: $q \geq 0.02$; $R \leq 100$.

Man sieht, daß vor allem an den Objektgrenzen die Verschiebungsvektoren fehlen. Hier ist das Verhältnis q sehr klein, da an Objektkanten die Gradienten in allen drei Farbkanälen und auch an benachbarten Pixeln alle in die gleiche Richtung zeigen. Untersucht man die dort ausgeblendeten Verschiebungsvektoren, so sieht man tatsächlich, daß sie fast alle entlang der Kante zeigen und viel zu lang sind. Schwierigkeiten macht dem reinen Farbverfahren auch das Glanzlicht auf dem linken Apfel, dort wird die Matrix \mathbf{E} sogar singulär. Außerdem fällt auf, daß die dunklen Bereiche rechts oben bzw. unten sowie in der Mitte zu schlech-

ten Ergebnissen führen. Dies könnte daran liegen, daß das Rauschen in dunklen Bildbereichen stärker ist als in den hellen Gebieten. Das Intensitätsverfahren liefert im allgemeinen leicht bessere Ergebnisse als das reine Farbverfahren. Eine Verbesserung gegenüber beiden vorherigen Verfahren kann durch das kombinierte Verfahren erreicht werden. Wie man der Abbildung 6.6 entnimmt, kommen hier kaum noch extreme Ausreißer vor. Das Vektorfeld hat einen "glatteren" Verlauf. Allerdings bleiben die Probleme beim Glanzlicht und an den Objektkanten bestehen. Dies rechtfertigt die Einführung der Gütemaße, die verhindern sollen, daß der Algorithmus fehlerhafte Verschiebungsvektoren an die weiterverarbeitenden Verfahren liefert.



Abbildung 6.3: Apfelbild, Rechnausschnitt des Intensitätsbildes.

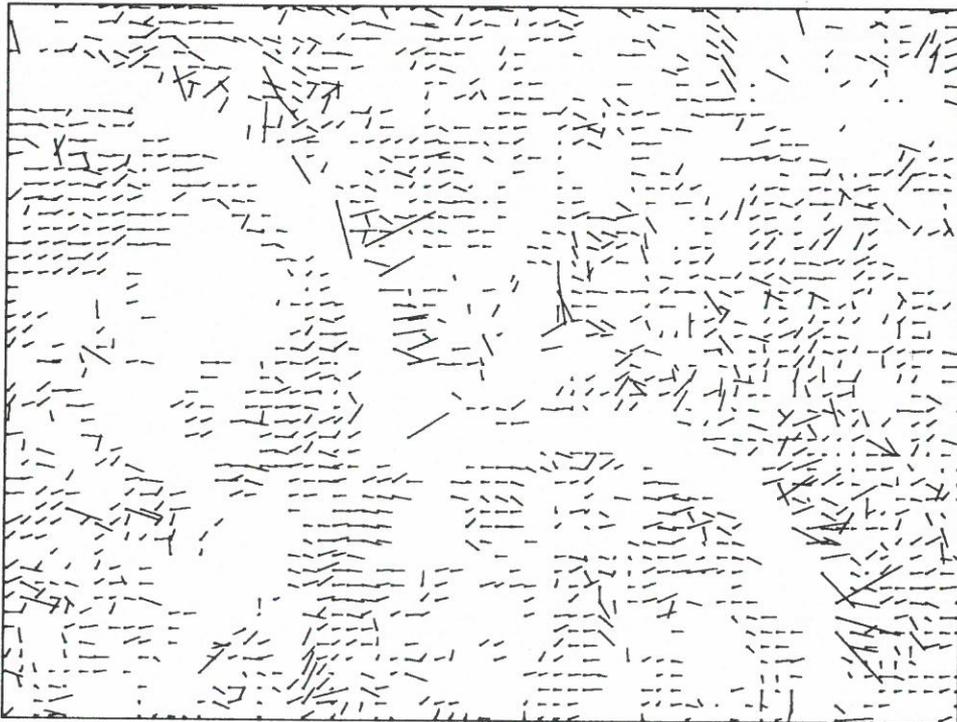


Abbildung 6.4: Verschiebungsvektoren von Bild 2 nach Bild 3, berechnet mit reinem Farbansatz



Abbildung 6.5: Verschiebungsvektoren von Bild 2 nach Bild 3, berechnet auf dem Intensitätsbild innerhalb einer 3×3 -Umgebung.



Abbildung 6.6: Verschiebungsvektoren von Bild 2 nach Bild 3, berechnet mit kombiniertem Verfahren innerhalb einer 3×3 Umgebung.

Kapitel 7

Schlußbemerkungen

Wir haben in dieser Arbeit eine Übersicht über verschiedene gradientenbasierte Verfahren zur Bestimmung des optischen Flusses gegeben. Eine nähere Untersuchung der diesen Verfahren zugrundeliegenden "motion constraint equation" zeigte, daß stark einschränkende Voraussetzungen gemacht werden müssen, damit der berechnete optische Fluß mit dem für die 3D-Rekonstruktion benötigten Bewegungsfluß übereinstimmt. Weiterhin konnten wir zeigen, daß im diskreten Fall erhebliche Fehler auftreten, wenn die Modellannahme einer verschobenen Grauwerttrampe nicht erfüllt wird.

Ein weiteres Problem der gradientenbasierten Verfahren ist das Blendenproblem. Verschiedene Ansätze zur Lösung dieses Problems wurden betrachtet. Nur die Einführung eines global zu optimierenden Glattheitsmaßes kann sicherstellen, daß an jedem Punkt des Bildes ein Verschiebungsvektor bestimmt wird. Leider können die entstehenden Differentialgleichungen nur durch eine relativ komplexe iterative Methode gelöst werden. Die Forschungen der letzten Jahre richtete sich deshalb auch auf eine Verminderung der Iterationsschritte durch spezielle Algorithmen, die z.B. auf einer Bildpyramide arbeiten.

In dieser Arbeit wurde nun der Vorschlag gemacht, die Information von Farbbildern zur Lösung des Blendenproblems zu nutzen. Die Hoffnungen richten sich dabei vor allem auf zwei Punkte.

- Bei Grauwertverfahren, die das Blendenproblem durch die Annahme eines lokal konstanten bzw. glatten optischen Flusses zu lösen versuchen, treten zwei Forderungen in einen gewissen Widerspruch. Zum einen sollte die Bildfunktion linear sein, damit die "motion constraint equation" gültig ist. Zum anderen muß die Bildfunktion gekrümmt sein, damit in benachbarten Pixeln ein anderer Gradient vorhanden ist. Bei der Verwendung von Farbbildern existiert dieser Widerspruch nicht mehr, wenn die Bildfunktion in jedem Kanal linear ist, aber in den verschiedenen Kanälen unterschiedliche Gradienten vorliegen.
- In Situationen, wo Beleuchtungseffekte eine Bewegung der Grauwerte in der

Bildebene erzeugen, stimmt der mit einem Grauwertverfahren berechnete optische Fluß nicht mit dem Bewegungsfluß überein. So kann es z.B. sein, daß sich durch eine Bewegung der Lichtquelle ein Schatten über ein Objekt bewegt. Dabei verändert sich die Intensität des Lichtes, und ein Grauwertverfahren registriert eine Bewegung des Objektes. Man könnte nun daran denken, die Farbinformation für die Erkennung dieser Situation heranzuziehen. Im einfachsten Fall kann man annehmen, daß sich die Farbe des Objektes nicht ändert. Verwendet man also nur die Farbinformation, ohne die Intensität zu berücksichtigen, so würde man keine Bewegung des Objektes detektieren.

In den praktischen Untersuchungen zeigte sich nun, daß sich diese beiden Hoffnungen nicht ganz erfüllen lassen.

- Die Verwendung der Farbinformation führt nicht in jedem Fall zu einem zuverlässigen Verschiebungsvektor, da die Gradientenrichtungen in den Farbkäneln bei den meisten Bildern korreliert sind. Es konnte aber gezeigt werden, daß die Verwendung der Farbe die Ergebnisse eines Grauwertverfahrens verbessern kann. So kommt man schon mit der Konstanzannahme in einer lokalen Umgebung von 3×3 Pixeln zu guten Ergebnissen.
- Der Verzicht auf die Intensität führt zu nicht befriedigenden Ergebnissen. Dies liegt daran, daß die Intensität im allgemeinen die wichtigste Information eines Farbbildes beinhaltet.

In der Arbeit wurde ein einfaches lokales Verfahren für die Verwendung von Farbbildern vorgestellt. Verbesserungen dieses Verfahrens sind vorstellbar, wie z.B. die Verwendung eines dreidimensionalen Beaudet-Operators für die Schätzung der räumlichen und zeitlichen Ableitungen. Mit einem solchen Operator können dann mehr als nur zwei Bilder einer Bildfolge verwendet werden. Die Ergebnisse dieses Verfahrens lassen sich auch als Startvektorfeld für ein globales Optimierungsverfahren verwenden, um an jedem Punkt einen Verschiebungsvektor zu erhalten. Man könnte allerdings auch die globalen Verfahren direkt für die Verwendung von Farbbildern erweitern. Insbesondere stellt sich die Frage, ob man die Farbinformation mit der gerichteten Glattheitsforderung kombinieren könnte.

Da die Genauigkeit der geschätzten Verschiebungsvektoren an den verschiedenen Stellen im Bild recht unterschiedlich ist, ist es sinnvoll, den weiterverarbeitenden Verfahren eine Bewertung jedes Verschiebungsvektors zur Verfügung zu stellen. Aus den Analysen konnten einige mögliche Kandidaten für ein Gütemaß hergeleitet werden. Die Tests mit einigen Bildern haben ergeben, daß sich vor allem zwei Parameter für die Messung der Zuverlässigkeit eignen:

- Das Verhältnis q von Determinante zum Quadrat der Spur der Matrix E .
- Die Summe der Absolutbeträge der Residuen r der einzelnen "motion constraint equations".

Ein offenes Problem ist noch die Weiterverarbeitung des entstehenden Verschiebungsvektorfeldes. Es läßt sich z.B. vorstellen, das Verfahren in ein umfangreicheres System, wie es in *Dreschler-Fischer 86* vorgeschlagen worden ist, als "Wissensquelle" zu integrieren. In einem solchen System könnten dann die Gütemaße differenzierter eingesetzt werden, als es bei den einfachen Schwellwertbildungen in dieser Arbeit gemacht wurde.

Anhang A

Zeichenerklärung

Hier werden alle Zeichen näher erklärt, die im Text häufiger auftauchen. Ich habe versucht, eine gewisse Konsistenz in der Schreibweise zu erreichen. Vektoren sind im allgemeinen durch kleine fette Buchstaben bezeichnet, Matrizen werden mit großen fetten Buchstaben bezeichnet, soweit sie nicht durch Operatoren entstehen. Der Nabla-Operator ∇ , der Bildkoordinatenvektor \mathbf{x} und der Verschiebungsvektor \mathbf{u} sind Spaltenvektoren. Dagegen handelt es sich beim Grauwertvektor \mathbf{g} und beim Farbvektor \mathbf{c} um Zeilenvektoren, da sich dann viele Gleichungen genauso schreiben lassen, wie im eindimensionalen Fall. Die normale Verknüpfung zwischen den Vektoren und Matrizen ist die Matrixmultiplikation. Falls die Skalarmultiplikation zwischen zwei Vektoren gemeint ist, steht ein Punkt \cdot zwischen den Vektoren. Ein hochgestelltes T bezeichnet die Transposition.

\mathbf{A}^T	Transponierte von \mathbf{A}	
\mathbf{A}^{-1}	Inverse von \mathbf{A}	
\mathbf{A}^\dagger	Pseudoinverse von \mathbf{A}	
∇	Ableitungsoperator	: $(\partial/\partial x, \partial/\partial y)^T$
\mathbf{I}	Einheitsmatrix	: (δ_{ij})
g	Bildfunktion $g(x, y)$: $\mathbf{R}^2 \rightarrow \mathbf{R}$
	Bildfolge $g(x, y, t)$: $\mathbf{R}^3 \rightarrow \mathbf{R}$
\mathbf{c}	Farbbildfunktion	: $\mathbf{R}^2 \rightarrow \mathbf{R}^3$
\mathbf{g}	Grauwertvektor	: (g_1, \dots, g_n) $(g(\mathbf{x}_1), \dots, g(\mathbf{x}_n))$
$(x, y)^T$	Bildkoordinatenvektor	
\mathbf{x}	Bildkoordinatenvektor	: $(x, y)^T$
g_x	partielle Ableitung	: $\partial g/\partial x$
\mathbf{g}_x	partielle Ableitung	: $(g_x(\mathbf{x}_1), \dots, g_x(\mathbf{x}_n))$
∇g	Grauwertgradient	: $(g_x, g_y)^T$
$\nabla \mathbf{g}$	Gradienten- oder Jacobi-Matrix	: $\begin{pmatrix} \mathbf{g}_x \\ \mathbf{g}_y \end{pmatrix}$
\mathbf{H}	Hessesche- oder Krümmungs-Matrix	: $\begin{pmatrix} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{pmatrix}$
\mathbf{C}	Approximation von $\nabla \mathbf{g}(\nabla \mathbf{g})^T$: $\nabla g(\nabla g)^T + x^2 \mathbf{H}\mathbf{H}$
\mathbf{E}	Kurzschreibweise von $\nabla \mathbf{c}(\nabla \mathbf{c})^T$	
$(u, v)^T$	Verschiebungsvektor	: $\left(\frac{dx}{dt}, \frac{dy}{dt}\right)^T$
\mathbf{u}	Verschiebungsvektor	: $(u, v)^T$
u^\perp	Komponente senkr. zur Kantenrichtung	
u^\parallel	Komponente in Kantenrichtung	
$\hat{\mathbf{u}}$	geschätzter Verschiebungsvektor	
$d\mathbf{u}$	Korrekturvektor	: $(du, dv)^T$
Δu	Laplace-Operator auf u	: $(u_{xx} + u_{yy})$
$\Delta \mathbf{u}$: $(\Delta u, \Delta v)^T$
$\nabla \mathbf{u}^T$	Gradientenmatrix	: $\begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix}$
r	Residuum	: $\nabla g \hat{\mathbf{u}} + g_t$
q	Gütemaß	: $\det E / (\text{Spur } \mathbf{E})^2$
k	Operatorweite	: Format: $(2k + 1)^2$
\mathbf{k}	Ortsfrequenz oder Wellenvektor	
ω	Zeitfrequenz	
(X, Y, Z)	Raumkoordinatenvektor	

Anhang B

Quantitative Analyse des Verhältnisses von Beleuchtung und "motion constraint equation"

Bei der Herleitung der "motion constraint equation" wurde die Annahme gemacht, daß sich der Grauwert eines auf die Bildebene projizierten Objektpunktes bei dessen Bewegung im Raum nicht ändert. Sei $\frac{dg}{dt}$ die mit dem Bewegungsfluß mitbewegte Ableitung der Bildfunktion, so kann man diese Annahme in der Form $\frac{dg}{dt} = 0$ schreiben. Die Aufgabe bei der quantitativen Analyse der Beleuchtungseffekte auf die "motion constraint equation" besteht also in der Berechnung dieser totalen Ableitung. Wir setzen zunächst noch kein spezielles Reflektanzmodell voraus.

Wir können zunächst einmal annehmen, daß die Bildhelligkeit g proportional zur ankommenden Leuchtkraft L ist. Dies gilt nur für kleine Winkel um die optische Achse der Kamera; am Rand eines Bildes können Vignettierungseffekte auftreten. Die Leuchtkraft L ist nun von folgenden Größen abhängig:

- der Oberflächennormalen $\mathbf{n}(\xi, \eta)$.
- von Größen α_i , die allein von den Flächenparametern $(\xi, \eta)^T$ und eventuell auch explizit von der Zeit abhängen. Ein Beispiel wäre die Albedo ρ .
- von Größen β_i , die allein von den Raumkoordinaten $(X, Y, Z)^T$ und eventuell auch explizit von der Zeit abhängen, wie z.B. die Beleuchtung in einem Raumpunkt.

Es gilt somit:

$$g(x, y) = L(\alpha_i(\xi, \eta), \beta_i(X, Y, Z), \mathbf{n})$$

$\frac{dg}{dt}$ ist die zeitliche Änderung der Bildintensität bei konstanten Oberflächenparametern $(\xi, \eta)^T$, also die Ableitung der beobachteten Helligkeit eines mitbewegten

Objektpunktes. Die Größen α_i , die von $(\xi, \eta)^T$ abhängen, liefern deshalb nur dann einen Beitrag zu $\frac{dg}{dt}$, wenn sie auch explizit von der Zeit abhängen. Bei den Größen β_i müssen wir hingegen eine totale Differentiation vornehmen. Dies liefert dann:

$$\frac{dg}{dt} = \sum_i \frac{\partial L}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial t} + \sum_i \frac{\partial L}{\partial \beta_i} \frac{d\beta_i}{dt} + \frac{\partial L}{\partial \mathbf{n}} \frac{d\mathbf{n}}{dt}$$

Die Oberfläche sei durch eine Funktion s beschrieben, die die Flächenparameter (ξ, η) auf die Raumkoordinaten abbildet. Wir machen nun noch die Annahme, daß sich die Bewegung der gesamten Objektoberfläche $s(\xi, \eta)$ durch nur eine Translationsgeschwindigkeit \mathbf{v}_t und eine Winkelgeschwindigkeit $\boldsymbol{\omega}$ beschreiben läßt. Das heißt, wir machen die Annahme, daß es sich um die Bewegung eines starren Körpers handelt. Die Gesamtgeschwindigkeit \mathbf{v} eines Objektpunktes in dem Raumpunkt \mathbf{X} ergibt sich zu $\mathbf{v} = \mathbf{v}_t + \boldsymbol{\omega} \times \mathbf{X}$. Dann gilt für die genannten Ableitungen:

$$\begin{aligned} \frac{d\beta_i}{dt} &= \frac{\partial \beta_i}{\partial X} \frac{dX}{dt} + \frac{\partial \beta_i}{\partial Y} \frac{dY}{dt} + \frac{\partial \beta_i}{\partial Z} \frac{dZ}{dt} + \frac{\partial \beta_i}{\partial t} \\ &= \nabla \beta_i \mathbf{v} + \frac{\partial \beta_i}{\partial t} \\ \frac{d\mathbf{n}}{dt} &= \frac{d}{dt} \left[\frac{\partial \mathbf{s}}{\partial \xi} \times \frac{\partial \mathbf{s}}{\partial \eta} \right] \\ &= \frac{\partial}{\partial \xi} (\mathbf{v}_t + \boldsymbol{\omega} \times \mathbf{s}) \times \frac{\partial}{\partial \eta} (\mathbf{v}_t + \boldsymbol{\omega} \times \mathbf{s}) \\ &= \boldsymbol{\omega} \times \mathbf{n} \end{aligned}$$

Setzen wir diese Ergebnisse in die obige Gleichung für die totale Ableitung ein, so erhalten wir:

$$\frac{dg}{dt} = \sum_i \frac{\partial L}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial t} + \sum_i \frac{\partial L}{\partial \beta_i} \left(\nabla \beta_i \mathbf{v} + \frac{\partial \beta_i}{\partial t} \right) + \frac{\partial L}{\partial \mathbf{n}} (\boldsymbol{\omega} \times \mathbf{n}) \quad (\text{B.1})$$

Nimmt man an, daß sich die Eigenschaften der Objektoberfläche selbst (wie die Albedo) mit der Zeit nicht ändern, so fällt der erste Term weg. Erst wenn man weiterhin annimmt, daß sich die Eigenschaften des Raumes mit der Zeit nicht verändern (keine bewegte Lichtquelle, keine indirekte Beleuchtung über bewegte Objekte, etc.) und auch noch räumlich konstant sind (weit entfernte Lichtquelle, keine Schatten, etc.), fällt der zweite Term weg. Um auch den dritten Term vernachlässigen zu können, müssen wir annehmen, daß das Objekt nicht rotiert oder daß die Reflektanzeigenschaften der Objektoberfläche nicht von der Kamerarichtung abhängen (matte Oberfläche) und, daß die Beleuchtung isotrop ist (diffuses Licht).

Konkrete Beispiele für verschiedene Oberflächenmodelle (lambertsche und spiegelnde Oberflächen) und eine weitere Herleitung der Gleichung B.1 findet man in Verri + Poggio 87.

Anhang C

Die "motion constraint equation" bei Drehungen

Eine Fläche, die durch eine Parametrisierung

$$\begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} = \begin{pmatrix} \xi \\ \eta \\ f(\xi, \eta) \end{pmatrix}$$

gegeben ist, werde um den Punkt $(0, 0, r)^T$ mit der Winkelgeschwindigkeit $\omega = (\omega_1, \omega_2, \omega_3)^T$ gedreht. Dabei legen die Parameter (ξ, η) einen Punkt auf der Oberfläche fest. Es soll untersucht werden, ob hier die "motion constraint equation" gültig ist. Um den Zeitpunkt 0 ergibt sich somit folgende Zeitabhängigkeit:

$$\begin{aligned} \begin{pmatrix} X(t) \\ Y(t) \\ Z(t) \end{pmatrix} &= \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + t \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \times \begin{pmatrix} \xi \\ \eta \\ f(\xi, \eta) - r \end{pmatrix} \\ &= \begin{pmatrix} 1 & -\omega_3 t & \omega_2 t \\ \omega_3 t & 1 & -\omega_1 t \\ \omega_2 t & \omega_1 t & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ f(\xi, \eta) \end{pmatrix} + r t \begin{pmatrix} -\omega_2 \\ \omega_1 \\ 0 \end{pmatrix} \end{aligned}$$

Jeder Punkt auf der Fläche ergibt auf der Bildebene einen Grauwert g , der sich mit der Zeit nicht ändern soll. Enthält die Oberfläche also ein Muster $P(\xi, \eta)$, so gilt: $g(x(t), y(t)) = P(\xi, \eta)$. Die Bildkoordinaten, auf die der Objektpunkt abgebildet wird, ändern sich natürlich mit der Zeit, da die Fläche rotiert.

Die Bildkoordinaten eines durch den Parametervektor $(\xi, \eta)^T$ bestimmten Flächenpunktes bei orthographischer Projektion $((x, y) = (XY))$, um den Zeitpunkt $t = 0$, ist gegeben durch

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -\omega_3 t \\ \omega_3 t & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \begin{pmatrix} t\omega_2(f-r) \\ -t\omega_1(f-r) \end{pmatrix}$$

Die Umkehrung ergibt sich zu

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \frac{1}{1 + t^2 \omega_3^2} \begin{pmatrix} 1 & \omega_3 t \\ -\omega_3 t & 1 \end{pmatrix} \begin{pmatrix} x - t\omega_2(f - r) \\ y + t\omega_1(f - r) \end{pmatrix}$$

Die räumlichen Ableitungen der Koordinaten ergeben sich zu

$$\begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} = \frac{1}{1 + t^2 \omega_3^2} \begin{pmatrix} 1 & \omega_3 t \\ -\omega_3 t & 1 \end{pmatrix} \begin{pmatrix} 1 - t\omega_2 f_x & -t\omega_2 f_y \\ t\omega_1 f_x & 1 + t\omega_1 f_y \end{pmatrix}$$

somit bekommen wir zum Zeitpunkt $t = 0$

$$\begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} \Big|_{t=0} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

und für die zeitliche Ableitung

$$\begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} \Big|_{t=0} = \begin{pmatrix} -\omega_2(f - r) \\ \omega_1(f - r) \end{pmatrix} + \omega_3 \begin{pmatrix} y \\ -x \end{pmatrix}$$

Die Verschiebungsvektoren berechnen sich zu

$$\begin{pmatrix} u \\ v \end{pmatrix} \Big|_{t=0} = \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} \Big|_{t=0} = \omega_3 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \omega_2(f - r) \\ -\omega_1(f - r) \end{pmatrix}$$

Die Ableitungen der Bildfunktion g sind:

$$\begin{aligned} \begin{pmatrix} g_x & g_y \end{pmatrix} &= \begin{pmatrix} P_\xi & P_\eta \end{pmatrix} \begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} \\ g_t &= \begin{pmatrix} P_\xi & P_\eta \end{pmatrix} \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} \end{aligned}$$

Mit diesen Termen können wir nun die "motion constraint equation" überprüfen:

$$\begin{aligned} \mathbf{u} \nabla g + g_t &= \begin{pmatrix} P_\xi & P_\eta \end{pmatrix} \left[\begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} \right] \\ &= \begin{pmatrix} P_\xi & P_\eta \end{pmatrix} \left[\omega_3 \begin{pmatrix} -y \\ x \end{pmatrix} + \begin{pmatrix} \omega_2 \\ -\omega_1 \end{pmatrix} (f - r) \right. \\ &\quad \left. + \begin{pmatrix} -\omega_2 \\ \omega_1 \end{pmatrix} (f - r) + \omega_3 \begin{pmatrix} y \\ -x \end{pmatrix} \right] \\ &= 0 \end{aligned}$$

Die "motion constraint equation" ist also erfüllt, obwohl das Verschiebungsvektorfeld divergente Teile enthält:

$$\nabla \cdot \mathbf{u} = u_x + v_y = \omega_2 f_x - \omega_1 f_y$$

Anhang D

Fehler der lokalen Konstanzannahme

Wir wollen den in Abschnitt 4.3.4 diskutierten Fehlerterm des geschätzten Verschiebungsvektors unter der Annahme eines lokal konstanten optischen Flusses herleiten. Die lokale Konstanzannahme führte auf die Gleichungen 4.9 bzw. 4.12:

$$\begin{aligned}\hat{\mathbf{u}} &= -(\nabla \mathbf{g} (\nabla \mathbf{g})^T)^{-1} \nabla \mathbf{g} \mathbf{g}_t^T \\ &\approx -\mathbf{C}^{-1} \left(g_t \nabla g + \bar{x}^2 \mathbf{H} \nabla g_t \right)\end{aligned}$$

Wir nehmen nun an, daß sich das Bild lokal durch ein Polynom zweiten Grades beschreiben läßt und die Bildfolge tatsächlich durch reine Verschiebung zustande gekommen ist:

$$g(\mathbf{x}, t) = g_0 + (\nabla g)^T (\mathbf{x} - \mathbf{u}t) + \frac{1}{2} (\mathbf{x} - \mathbf{u}t)^T \mathbf{H} (\mathbf{x} - \mathbf{u}t)$$

Die zeitlich Ableitung werde durch die Differenz der Grauwerte zweier aufeinanderfolgender Bilder gebildet. Wir verwenden im folgenden ein lokales Koordinatensystem. Die Zeitdifferenz zwischen zwei aufeinanderfolgenden Bildern sei $\Delta t = 1$.

$$g_t \approx -(\nabla g)^T \mathbf{u} + \frac{1}{2} (\mathbf{u}^T \mathbf{H} \mathbf{u})$$

Somit erhalten wir für den geschätzten Verschiebungsvektor:

$$\begin{aligned}\hat{\mathbf{u}} &= -\mathbf{C}^{-1} \left[\left(-(\nabla g)^T \mathbf{u} + \frac{1}{2} (\mathbf{u}^T \mathbf{H} \mathbf{u}) \right) \nabla g + \bar{x}^2 \mathbf{H} \nabla g_t \right] \\ &= \mathbf{C}^{-1} \left(\nabla g \left((\nabla g)^T \mathbf{u} - \frac{1}{2} (\mathbf{u}^T \mathbf{H} \mathbf{u}) \right) + \bar{x}^2 \mathbf{H}^T \mathbf{H} \mathbf{u} \right)\end{aligned}$$

Ohne Beschränkung der Allgemeinheit legen wir das Koordinatensystem so, daß die gemischte Ableitung g_{xy} verschwindet. Wir führen noch folgende Abkürzungen

ein:

$$\begin{aligned} \det &= g_x^2 g_{yy}^2 + g_y^2 g_{xx}^2 + \bar{x}^2 g_{xx}^2 g_{yy}^2 \\ kr &= \frac{1}{2} |\mathbf{u}|^2 (g_{xx} u^2 + g_{yy} v^2) \end{aligned}$$

Dabei ist \det proportional zur Determinante von \mathbf{C} und kr mißt die Krümmung der Bildfunktion in Richtung der Verschiebung. Mit diesen Größen folgt dann schließlich:

$$\begin{aligned} \hat{\mathbf{u}} &= \frac{1}{\bar{x}^2 \det} \begin{pmatrix} g_y^2 + \bar{x}^2 g_{yy}^2 & -g_x g_y \\ -g_x g_y & g_x^2 + \bar{x}^2 g_{xx}^2 \end{pmatrix} \\ &\quad \cdot \left[g_x u + g_y v - kr |\mathbf{u}|^2 \begin{pmatrix} g_x \\ g_y \end{pmatrix} + \bar{x}^2 \begin{pmatrix} g_{xx}^2 u \\ g_{yy}^2 v \end{pmatrix} \right] \\ &= \mathbf{u} - \frac{kr}{\det} \begin{pmatrix} g_x g_{yy}^2 \\ g_y g_{xx}^2 \end{pmatrix} |\mathbf{u}|^2 \end{aligned}$$

Anhang E

Bilder

Die verwendeten Bilder wurden mit einer RGB-Farbkamera der Firma BOSCH aufgenommen, die mit drei Plumbiconröhren ausgestattet ist. Zur Digitisierung der Apfelbildfolge wurde das Bildspeichersystem FB 768 P der Firma dvs-GmbH verwendet, welches bis zu 8 RGB-Farbbilder im Format von 512×512 Pixeln aufnehmen kann. Jedes Bild besteht aus zwei Halbbildern: zwei aufeinanderfolgende Zeilen haben einen zeitlichen Abstand von 20ms. Da dies bei bewegten Objekten zu störenden Effekten führt, haben wir ein solches Bild in die beiden Halbbilder aufgeteilt. Hiervon wurde jeweils nur das erste Halbbild weiter verwendet.

Anschließend wurden je zwei benachbarte Pixel einer Zeile zusammengefasst und ein Bildausschnitt ausgewählt, um auf das Seitenverhältnis von 4 : 3 zu kommen. Wir erhalten schließlich ein sogenanntes *Geo-Bild* im Format von 256×192 Pixeln.

Aus technischen Gründen konnten leider keine Farbbilder in die Arbeit aufgenommen werden. Stattdessen ist in den Bildern 1 bis Bild 3 die Intensität, d.h. das arithmetische Mittel der drei Farbkanäle, abgebildet. Für das Apfelbild 2 sind auch die Farbauszüge einzeln abgebildet (Bilder 4 bis 6).

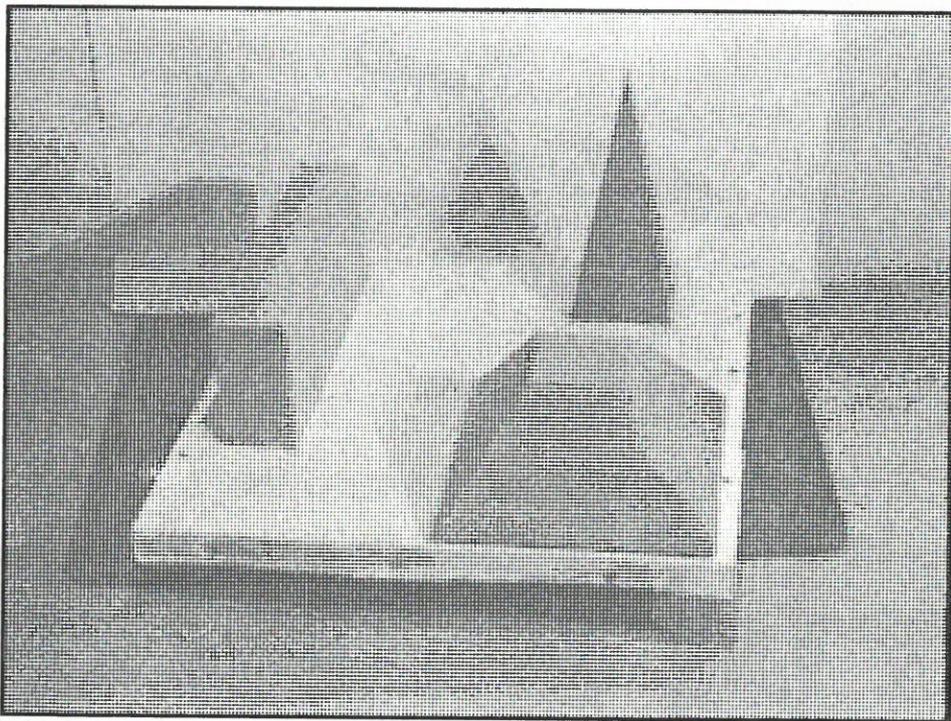


Bild 1: Blocksbild, Intensität



Bild 2: Apfelszene 1. Bild, Intensität

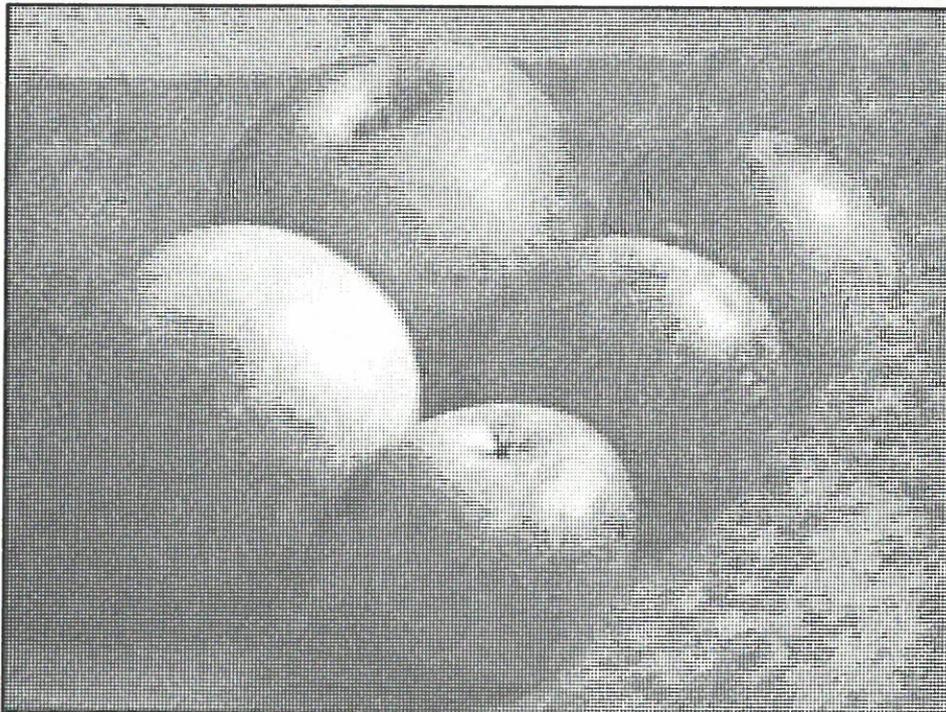


Bild 3: Apfelszene 2. Bild, Intensität

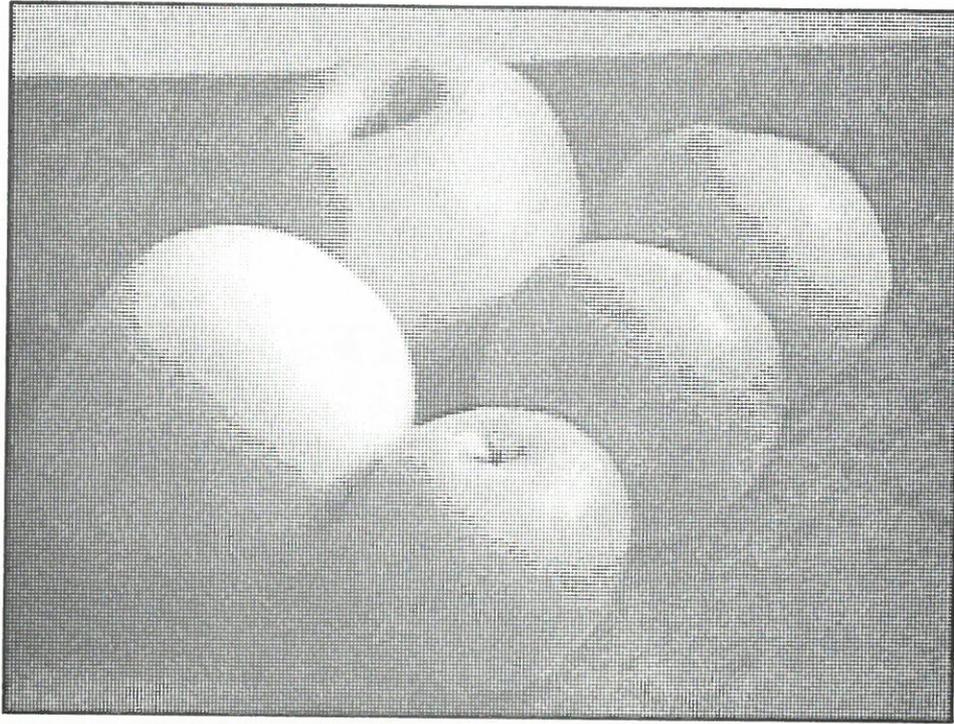


Bild 4: Apfelbild, Original, Rotauszug



Bild 5: Apfelbild, Original, Grünauszug



Bild 6: Apfelbild, Original, Blauauszug

Literaturverzeichnis

Anandan + Weiss 85

Introducing a Smoothness Constraint in a Matching Approach for the Computation of Optical Flow Fields.

P. Anandan, Richard Weiss.

Proc. Workshop on Computer Vision: Representation and Control, Bellaire/MI 1985, 186-194.

Ballard + Brown 82

Computer Vision.

Dana H. Ballard, Christopher M. Brown.

Prentice-Hall, Englewood Cliffs, New Jersey 1982.

Bandyopadhyay 84

A Multiple Channel Model for Perception of Optical Flow.

Amit Bandyopadhyay.

Proc. Workshop on Computer Vision: Representation and Control, Annapolis/MD 1984, 78-82.

Barnard + Thompson 80

Disparity Analysis of Images.

Stephen T. Barnard, William B. Thompson.

IEEE Trans. Patt. Anal. Mach. Intell. **PAMI-2** (1980), 333-340.

Barron 84

A Survey of Approaches for Determining Optic Flow, Environmental Layout and Egomotion.

John Barron.

Bericht *RBCV-TR-84-5*, Department of Computer Science, University of Toronto 1984.

Bartsch 87

Untersuchungen zur Detektion und Lokalisation markanter Punkte in Farbbildern.

Thomas Bartsch.

Bericht *FBI-HH-B-131/87*, Fachbereich Informatik, Universität Hamburg, 1987.

Beaudet 78

Rotationally Invariant Image Operators.

Paul R. Beaudet.

Proc. Int. Conf. Patt. Recogn. ICPR-1978, Kyoto, Japan, 579-583.

Bergmann 83

Analysis of Different Displacement Estimation Algorithms for Digital Television Signals.

Hans C. Bergmann.

in: T.S. Huang (Hrsg.): *Image Sequence Processing and Dynamic Scene Analysis*; NATO ASI Series Vol. F2; Springer-Verlag, Berlin - Heidelberg - New York - Tokyo 1983, 214-234.

Blicher 83

The Stereo Matching Problem From the Topological Viewpoint.

A. Peter Blicher.

Proc. Int. Joint Conf. on Art. Intell. IJCAI-1983, Karlsruhe, Germany, 1046-1049
siehe auch: *Stereo Matching From the Topological Viewpoint* in Proc. Image Understanding Workshop, Arlington, Virginia 1983.

Bonde + Nagel 79

Deriving a 3-D Description of a Moving Rigid Object from Monocular TV-Frame Sequences.

Thomas Bonde und Hans-Illmut Nagel.

Workshop on Computer Analysis of Time-Varying Imagery, Philadelphia/PA 1979, 44-45.

Bruss + Horn 83

Passive Navigation.

Anna R. Bruss, Berthold Klaus Paul Horn.

Computer Vision, Graphics, and Image Processing 21 (1983), 3-20; siehe auch *Horn 86*.

Cornelius + Kanade 83

Adapting Optical Flow to Measure Object Motion in Reflectance and X-ray Image Sequences.

Nancy Cornelius und Takeo Kanade.

Siggraph/Sigart Interdisciplinary Workshop on Motion: Representation and Perception, Toronto/Canada 1983, 579-583.

Courant + Hilbert 68a

Methoden der mathematischen Physik I, 3. Auflage.

Richard Courant, David Hilbert.

Heidelberger Taschenbücher Band 30; Springer-Verlag, Berlin - Heidelberg - New York - Tokyo 1968.

Courant + Hilbert 68b

Methoden der mathematischen Physik II, 2. Auflage.

Richard Courant, David Hilbert.

Heidelberger Taschenbücher Band 31; Springer-Verlag, Berlin - Heidelberg - New York - Tokyo 1968.

Dengler 85a

Methoden und Algorithmen zur Analyse bewegter Realweltszenen im Hinblick auf ein Blindenhilfesystem.

Joachim Dengler.

Dissertation, Naturwissenschaftlich-Mathematische Gesamtfakultät der Ruprecht-Karls-Universität Heidelberg, 1985.

Dengler 85b

Analyse bewegter Realwelt-Szenen.

Joachim Dengler.

Technical Report 5/1985, Abt. Medizinische und Biologische Informatik, Deutsches Krebsforschungszentrum Heidelberg, 1985.

Dengler u.a. 86

Lokale Bewegungsanalyse mit der dynamischen Pyramide.

Joachim Dengler, H.P. Meinzer und M. Schmidt.

8. DAGM-Symposium "Mustererkennung", Paderborn, 1986; G. Hartmann (Hrsg.); Informatik-Fachberichte Bd. 125; Springer-Verlag, Berlin - Heidelberg - New York - Tokyo 1986, 276-281.

Dreschler-Fischer 86

A Blackboard System for Dynamic Stereo Matching.

Leonie S. Dreschler-Fischer.

Proc. 1. Int. Conf. on Autonomous Mobile Systems, Amsterdam, Dezember 1986.

Dreschler + Nagel 82

Volumetric Model and 3D-Trajectory of a Moving Car Derived from Monocular TV-Frame Sequences of a Street Scene.

Leonie S. Dreschler, Hans-Hellmut Nagel.

Computer Graphics and Image Processing **20** (1982), 199-228

siehe auch: Proc. Int. Joint Conf. on Art. Intell. IJCAI-1981, Vancouver/BC., Canada, 692-697.

Duda + Hart 73

Pattern Classification and Scene Analysis.

Richard O. Duda, Peter E. Hart.

John Wiley & Sons, New York 1973.

Enkelmann 85

Mehrgitterverfahren zur Ermittlung von Verschiebungsvektorfeldern in Bildfolgen.

Wilfried Enkelmann.

Dissertation, Fachbereich Informatik, Universität Hamburg, FBI-HH-B-115/85.

Fang + Huang 84

Solving Three-Dimensional Small-Rotation Motion Equations: Uniqueness Algorithms and Numerical Results.

J.Q. Fang und Thomas S. Huang.

Computer Vision, Graphics, and Image Processing **26** (1984), 183-206.

Feynman 64

The Feynman Lectures on Physics, Volume 2.

Richard P. Feynman.

Addison-Wesley, Reading, Mass. 1964.

Fleet + Jepson 85

Velocity Extraction without Form Interpretation.

David J. Fleet, Allan D. Jepson.

Proc. Workshop on Computer Vision: Representation and Control, Bellaire/MI 1985, 179-185.

Gershon 87

The Use of Color in Computational Vision.

Ron Gershon.

Bericht RBCV-TR-87-15, Department of Computer Science, University of Toronto 1987.

Grimson 83

Surface Consistency Constraints in Vision.

W.E.L. Grimson.

Computer Vision, Graphics, and Image Processing 24 (1983), 28-51.

Haralick + Lee 83

The Facet Approach To Optic Flow.

Robert M. Haralick und Jong Soo Lee.

Proc. Image Understanding Workshop, Arlington/VA 1983, 84-93.

Heeger 87

Optical Flow from Spatiotemporal Filters.

David J. Heeger.

Proc. Int. Conf. on Comp. Vision ICCV-1987, London, 181-190.

Heeger 88

Optical Flow Using Spatiotemporal Filters.

David J. Heeger.

International Journal of Computer Vision 1 (1988), 279-302.

Hildreth 82

The Integration of Motion Information Along Contours.

Ellen C. Hildreth.

IEEE Workshop on Computer Vision: Representation and Control, Rindge, New Hampshire 1982, 83-91.

Hildreth 83

The Computation of the Velocity Field.

Ellen C. Hildreth.

AI Memo No. 734, MIT AI Lab, Sep. 1983.

Hildreth 84

The Measurement of Visual Motion.

Ellen C. Hildreth.

The MIT Press, Cambridge, Mass. 1984.

Horn 86

Robot Vision.

Berthold Klaus Paul Horn.

The MIT Press, Cambridge, Mass. 1986.

Horn + Schunck 81

Determining Optical Flow.

Berthold Klaus Paul Horn, Brian G. Schunck.

Artificial Intelligence 17 (1981), 185; siehe auch *Horn 86*.

Ikeuchi + Horn 81

Numerical Shape from Shading and Occluding Boundaries.

Katsushi Ikeuchi und Berthold Klaus Paul Horn.

Artificial Intelligence 17 (1981), 141-184; siehe auch *Horn 86*.

Jazwinski 70

Stochastic Processes and Filtering Theory.

Andrew H. Jazwinski.

Series: Mathematics in Science and Engineering, Volume 64, Academic Press, New York 1970.

Jenkins + Kolers 86

Some Problems with Correspondence.

Michael Jenkins und Paul A. Kolers.

Bericht RBCV-TR-86-10, Department of Computer Science / Department of Psychology, University of Toronto 1987.

Kanatani 86

Structure and Motion from Optical Flow under Orthographic Projection.

Ken-Ichi Kanatani.

Computer Vision, Graphics, and Image Processing 35 (1986), 181-199.

Kanatani 87

Structure and Motion from Optical Flow under Perspective Projection.

Ken-Ichi Kanatani.

Computer Vision, Graphics, and Image Processing 38 (1987), 122-146.

Kearney u.a. 87

Optical Flow Estimation: An Error Analysis of Gradient-Based Methods with Local Optimization.

Joseph K. Kearney, William B. Thompson, Daniel L. Boley.

IEEE Trans. Patt. Anal. Mach. Intell. PAMI-9 (1987), 229-244.

Kories u.a. 86

Experimente zur Bestimmung kinematischer Parameter aus dem optischen Fluß basierend auf einem robusten Verschiebungsschätzverfahren.

Ralf Kories, G. Zimmermann und H. Zimmer.

8. DAGM-Symposium "Mustererkennung", Paderborn, 1986; G. Hartmann (Hrsg.); Informatik-Fachberichte Bd. 125; Springer-Verlag, Berlin - Heidelberg - New York - Tokyo 1986, 190-195.

Marr + Ullman 81

Directional Selectivity and its Use in Early Visual Processing.

David Marr, Shimon Ullmann.

Proc. Roy. Soc. London, Ser. B 211 (1981), 151-180.

Marsal 76

Die numerische Lösung partieller Differentialgleichungen in Wissenschaft und Technik.

Dietrich Marsal.

Bibliographisches Institut, Mannheim - Wien - Zürich 1976.

Meiri 80

On Monocular Perception of 3-D Moving Objects.

A. Zri Meiri.

IEEE Trans. Patt. Anal. Mach. Intell. **PAMI-2** (1980), 582-583.

Mitiche 84

Computation of Optical Flow and Rigid Motion.

Amar Mitiche.

Proc. Workshop on Computer Vision: Representation and Control, Annapolis/MD 1984, 63-71.

Moravec 80

Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover.

Hans P. Moravec.

Report *STAN-CS-80-813*, Computer Science Dep., Stanford University, Stanford/CA, 1980

siehe auch: Report *CMU-RI-TR-3*, Robotics Institute, Carnegie-Mellon University, Pittsburgh/PA, 1980.

Nagel 83

Displacement Vectors Derived from Second-Order Intensity Variations in Image Sequences.

Hans-Hellmut Nagel.

Computer Vision, Graphics, and Image Processing **21** (1983), 85-117.

Nagel 83b

Overview on Image Sequence Analysis.

Hans-Hellmut Nagel.

in: T.S. Huang (Hrsg.): *Image Sequence Processing and Dynamic Scene Analysis*; NATO ASI Series Vol. F2; Springer-Verlag, Berlin - Heidelberg - New York - Tokyo 1983.

Nagel 85

Analyse und Interpretation von Bildfolgen.

Hans-Hellmut Nagel.

Informatik Spektrum **8** (1985), 178-200 (Teil 1), 312-327 (Teil 2).

Nagel 86

Image Sequences - Ten (octal) Years - From Phenomenology towards a Theoretical Foundation.

Hans-Hellmut Nagel.

Proc. Int. Conf. Patt. Recogn. ICPR-1986, Paris, France, 1174-1185.

Nagel 87

On the Estimation of Optical Flow: Relations between Different Approaches and

Some New Results.

Hans-Hellmut Nagel.

Artificial Intelligence **33** (1987), 299-324.

Nagel + Enkelmann 86

An Investigation of Smoothness Constraints for the Estimation of Displacement Vector Fields.

Hans-Hellmut Nagel, Wilfried Enkelmann.

IEEE Trans. Patt. Anal. Mach. Intell. PAMI-8 (1986), 565-593.

Nagel + Neumann 81

On 3-D Reconstruction from Two Perspective Views.

Hans-Hellmut Nagel und Bernd Neumann.

Proc. Int. Joint Conf. on Art. Intell. IJCAI-1981, Vancouver/BC., Canada, 661-663.

Otha u.a. 80

Color Information for Region Segmentation.

Yu-Ichi Otha, Takeo Kanade und Toshiyuki Sakai.

Computer Graphics and Image Processing **13** (1980), 222-241.

Potter 73

Computational Physics.

D. Potter.

John Wiley & Sons, New York 1973.

Prazdny 81

Determining the Instantaneous Direction of Motion from Optical Flow Generated by a Curvilinearly Moving Observer.

K. Prazdny.

Computer Graphics and Image Processing **17** (1981), 238-248.

Robbins + Netravali 83

Recursive Motion Compensation: A Review.

J.D. Robbins und A.N. Netravali.

in: T.S. Huang (Hrsg.): *Image Sequence Processing and Dynamic Scene Analysis*; NATO ASI Series Vol. F2; Springer-Verlag, Berlin - Heidelberg - New York - Tokyo 1983, 75-103.

Rosenfeld + Kak 82

Digital Picture Processing. Second Edition.

Azriel Rosenfeld, Avinash C. Kak.

Academic Press, New York 1982, 2 Bände.

Schalkoff + Labuz 84

An Integrated Spatio-Temporal Model and Recursive Algorithm for Image Motion Estimation.

Robert J. Schalkoff, Jefferey Labuz.

Proc. Int. Conf. Patt. Recogn. ICPR-1984, Montréal, Canada, 530.

Schunck 84a

Motion Segmentation and Estimation by Constraint Line Clustering.

Brian G. Schunck.

Proc. Workshop on Computer Vision: Representation and Control, Annapolis/MD 1984, 58-62.

Schunck 84b

The Motion Constraint Equation for Optical Flow.

Brian G. Schunck.

Proc. Int. Conf. Patt. Recogn. ICPR-1984, Montréal, Canada, 20-22.

Schunck 86a

The Image Flow Constraint Equation.

Brian G. Schunck.

Computer Vision, Graphics, and Image Processing **35** (1986), 20-46.

Schunck 86b

Image Flow Continuity Equations for Motion and Density.

Brian G. Schunck.

Proc. Workshop on Motion: Representation and Analysis, Charleston, South Carolina 1986, 89-94.

Scott 86

Smoothing the Optical Field under Perspective Projection.

G.L. Scott.

Comp. Vision and Pattern Recognition (1986), 504-509.

Stuller + Krishnamurthy 83

Kalman Filter Formulation of Low-Level Television Image Motion Estimation.

John Stuller und Goplan Krishnamurthy.

Computer Vision, Graphics, and Image Processing **21** (1983), 169-204.

Subbarao 87

Bounds on Translational and Angular Velocity Components from First Order Derivatives of Image Flow.

Muralidhara Subbarao.

Proc. Nat. Conf. on Art. Intell. AAAI-1987, Seattle, Washington, 744-748.

Subbarao + Waxman 86

Closed Form Solutions to Image Flow for Planar Surfaces in Motion.

Muralidhara Subbarao und Allen M. Waxman.

Computer Vision, Graphics, and Image Processing **36** (1986), 208-228.

Terzopoulos 86b

Regularization of Inverse Visual Problems Involving Discontinuities.

Demetri Terzopoulos.

IEEE Trans. Patt. Anal. Mach. Intell. **PAMI-8** (1986), 413-424.

Tretiak + Pastor 84

Velocity Estimation from Image Sequences with Second Order Differential Operators.

Oleh Tretiak, Louis Pastor.

Proc. Int. Conf. Patt. Recogn. ICPR-1984, Montréal, Canada, 16-19.

Tsai + Huang 84

Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces.

Roger Y. Tsai und Thomas S. Huang.

IEEE Trans. Patt. Anal. Mach. Intell. **PAMI-6** (1984), 13-26.

Ullman 79

The Interpretation of Visual Motion.

Shimon Ullman.

The MIT Press, Cambridge, Mass. 1979.

Verri + Poggio 87

Against Quantitative Optical Flow.

Alessandro Verri und Tomaso Poggio.

Proc. Int. Conf. on Comp. Vision ICCV-1987, London, 171-180.

Waxman + Wohn 84

Contour Evolution, Neighbourhood Deformation and Global Image Flow: Planar Surfaces in Motion.

Allen M. Waxman, K. Wohn.

Bericht *CAR-TR-58*, Center for Automation Research, Univ. of Maryland, April 1984.

Waxman u.a. 87

Closed-Form Solutions to Image Flow Equations.

Allen M. Waxman, Behrooz Kamgar-Parsi und Muralidhara Subbarao.

Proc. Int. Conf. on Comp. Vision ICCV-1987, London, 12-24.

Webb + Aggarwal 82

Structure from Motion of Rigid and Jointed Objects.

J.A. Webb und J.K. Aggarwal.

Artificial Intelligence **19** (1982), 107-130.

Danksagungen

Ich danke meiner Betreuerin Prof. Dr. Leonie Dreschler-Fischer für die Unterstützung bei der Durchführung dieser Arbeit. Sie hat es verstanden, mich durch ihre Zuversicht auch in schwierigen Phasen immer wieder neu zu motivieren.

Besonderer Dank gebührt auch Carsten Schröder und Christian Drewniok. Sie haben mich nach einer Durchsicht eines ersten Entwurfs der Arbeit auf Unklarheiten aufmerksam gemacht. Carsten Schröder möchte ich auch für die Hilfe bei TeXnischen Problemen danken. Christian Drewniok hat durch die Hilfe bei der Erstellung von Programmen zum Drucken von Bildern und Graphiken auf dem Laserdrucker viele Abbildungen in dieser Arbeit erst möglich gemacht.

Ingeborg Heer-Mück und Hartmuth Krüger waren stets hilfsbereit, wenn es Probleme mit der Verkabelung der Labor-Hardware bei der Digitisierung von Bildern gab. Auf Helmut Faasch und Volker Haarslev war bei Problemen mit der Software stets Verlaß.

Nicht zuletzt möchte ich an dieser Stelle auch meinen Eltern danken, die mir mein Studium und damit auch die Anfertigung dieser Arbeit ermöglicht haben.

