



Universität Hamburg  
Fakultät für Mathematik,  
Informatik und Naturwissenschaften

## Bachelorarbeit

# Konzeption und Implementierung eines Verfahrens zur Messung von Verschiebungsvektoren in Multispektralbildern

**Oliver Bestmann**

---

7bestman@informatik.uni-hamburg.de

Studiengang Informatik

Matr.-Nr. 5945392

Fachsemester 6

Erstbetreuer: Prof. Dr. rer. nat. Leonie Dreschler-Fischer

Zweitbetreuer: Dr. Werner Hansmann

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Zielsetzung dieser Arbeit . . . . .	2
1.2	Gliederung und Vorgehen . . . . .	2
<b>2</b>	<b>Optischer Fluss</b>	<b>4</b>
2.1	Kreuzkorrelation . . . . .	5
2.2	Differentieller Ansatz . . . . .	6
2.3	Lucas-Kanade Verfahren . . . . .	8
2.3.1	Verbesserungen des Verfahrens . . . . .	10
2.3.2	Eigenschaften des Verfahrens . . . . .	11
2.4	Horn-Schunck Verfahren . . . . .	11
2.4.1	Eigenschaften des Verfahrens . . . . .	13
<b>3</b>	<b>Multispektralbilder</b>	<b>14</b>
3.1	Einführung . . . . .	14
3.2	Multispektrales Lucas-Kanade Verfahren . . . . .	16
3.3	Multispektrales Horn-Schunck Verfahren . . . . .	18
<b>4</b>	<b>Vergleich der Verfahren</b>	<b>22</b>
4.1	Testdaten . . . . .	22
4.2	Implementierung und Testumgebung . . . . .	23
4.3	Vergleichsmethodik . . . . .	24
4.4	Ergebnisse . . . . .	26
<b>5</b>	<b>Fazit und Ausblick</b>	<b>29</b>
	<b>Literaturverzeichnis</b>	<b>31</b>
	<b>Eidesstattliche Erklärung</b>	<b>33</b>

# 1 Einleitung

Die Schätzung von Bewegungs- und Verschiebungsvektoren aus Bildfolgen ist ein Problem, das sich in vielen Bereichen wiederfinden lässt. So spielt die Schätzung beispielsweise bei der automatischen Navigation von Robotern und Fahrzeugen eine wichtige Rolle. Auch in Alltagsanwendungen wie bei optischen Computermäusen oder bei der Videokompression werden diese Verschiebungsvektoren benötigt. Weiterhin bilden sie eine Ausgangsbasis für die Erkennung dreidimensionaler Strukturen in Szenen und zur Schätzung von Bewegungen im Raum, sowie für die Erkennung und Unterscheidung von verschiedenen bewegten Objekten. Diese Verschiebungsvektoren sind nach derzeitigen Erkenntnissen auch in der visuellen Informationsverarbeitung von Menschen und anderen Lebewesen von Bedeutung [5, 18, 19, 4].

Diese Vektoren, im Folgenden auch als optischer Fluss bezeichnet, sind eine Größe, die nicht direkt berechnet werden kann. Darum ist es nun notwendig, geeignete Verfahren zu finden, um diese Verschiebungsvektoren zu schätzen. Diese können sich nur auf die messbaren Helligkeitsmuster im Bild stützen. Zur Berechnung benötigte Größen wie die Gradientenbeträge lassen sich ebenfalls nur aus diesen diskreten Abtastwerten abschätzen. Weiterhin lässt sich aufgrund des Blendenproblems ohne weitere Informationen oder Annahmen über die Bewegung für jeden Bildpunkt grundsätzlich nur die Bewegung in Gradientenrichtung schätzen [17, 4].

*When only a featureless straight contour of a moving object is visible, one cannot tell its true velocity and the object seems to be moving perpendicularly to its orientation. [12]*

Da man jedoch an der Berechnung der Vektoren unabhängig von der Gradientenrichtung interessiert ist, bedient man sich geeigneter Modellvorstellungen über die zu schätzenden Vektorfelder. Beispielsweise wird häufig angenommen, dass das Verschiebungsfeld in einer kleinen Umgebung um jeden Bildpunkt konstant ist. Treten in dieser Umgebung dann unterschiedliche Gradientenrichtungen auf, so kann damit das Blendenproblem gelöst werden und somit die Verschiebung korrekt geschätzt werden. Auch können Annahmen bezüglich der Glattheit des resultierenden Feldes getroffen werden [7, 4, 17].

## 1.1 Zielsetzung dieser Arbeit

Ziel dieser Arbeit ist es, ein Verfahren zu entwickeln, welches zusätzliche Informationen aus Mehrkanalaufnahmen nutzt, um die Verschiebungsvektoren zwischen zwei Bildern zu schätzen. Dabei soll auf bereits bekannten Verfahren zur Messung von Verschiebungsvektoren aufgebaut und diese entsprechend erweitert werden. Durch die direkte Verwendung von Mehrkanalbildern ist dann eine bessere Schätzung der Verschiebung zu erhoffen. Die so hergeleiteten Verfahren sollen in der Sprache C++ mithilfe der Grafikbibliothek *VIGRA* implementiert werden.

*VIGRA stands for Vision with Generic Algorithms. It's a novel computer vision library that puts its main emphasize on customizable algorithms and data structures. [1]*

und sowohl an synthetischen Testbildern als auch an echten Mehrkanalaufnahmen getestet werden. Da für die synthetischen Testbilder die korrekten Verschiebungsvektoren für jeden Bildpunkt vorliegen, ist auch ein quantitativer Vergleich der entwickelten Verfahren untereinander und zu anderen bekannten, auf Grauwertbildern arbeitenden Verfahren möglich. Für die Mehrkanalaufnahmen stehen einige von einem Satelliten aufgezeichnete Fernerkundungsbilder der Ostsee zur Verfügung. Diese zeigen eine Algenblütenpopulation in der Ostsee, welche sich mit der Meeresströmung bewegt. Durch die Verschiebung der Algen kann direkt auf die Meeresströmungen geschlossen werden. Des weiteren stehen öffentlich zugängliche Mehrkanalbilder von Wettersatelliten zur Verfügung [6].

## 1.2 Gliederung und Vorgehen

Das erste Kapitel dieser Arbeit beschreibt die Wichtigkeit der Berechnung von Verschiebungsvektoren in Bildfolgen. Weiterhin werden Probleme aufgezeigt, die bei der Schätzung eben dieser auftreten, sowie kurze Hinweise auf Lösungsansätze beschrieben.

Im zweiten Kapitel werden die theoretischen Grundlagen erläutert. Es wird der Begriff des optischen Flusses eingeführt. Neben dem Schätzungsansatz über die Korrelation wird der differentielle Berechnungsansatz beschrieben. In diesem Zusammenhang wird die wichtige *motion-constraint equation* beschrieben. Darauf aufbauend werden zwei bedeutsame und bekannte Verfahren zur Schätzung des optischen Flusses auf Grauwertbildern hergeleitet und kommentiert.

Das dritte Kapitel diskutiert das Konzept von Mehrkanal, bzw. Multispektralbildern und dem zusätzlichen Nutzen, der sich durch die Verwendung mehrerer Kanäle bei der Schätzung von Verschiebungsvektoren erhoffen lässt. Hierbei werden die im zweiten Kapitel beschriebenen Grauwertverfahren als Grundlagen genutzt um Verfahren zu entwickeln, welche die Informationen aus Mehrkanalbildern zur Schätzung von Verschiebungsvektoren nutzen. Daraus motiviert sich dann die Implementierung der erarbeiteten Verfahren in C++.

In Kapitel vier wird kurz die Implementierung beschrieben. Um die entwickelten Algorithmen quantitativ vergleichen zu können, werden noch einige Bildsequenzen erzeugt und beschrieben. Weiter werden exemplarisch einige Vergleichsmaße erklärt. Zum Schluss werden die Testsequenzen dann mit den unterschiedlichen Verfahren ausgewertet und die Ergebnisse miteinander verglichen.

Abschließen wird diese Bachelorarbeit mit einem kurzen Fazit und einem Ausblick, wie die hier entwickelten Verfahren als Grundlage für weitere Forschung genutzt werden können.

## 2 Optischer Fluss

Es existieren verschiedene, zum Teil sich widersprechende Definitionen für den optischen Fluss. Es seien hier zwei aufgeführt:

*Optical flow is the distribution of apparent velocities of movement of brightness patterns in an image. [4]*

*The velocity field that represents the motion of object points across an image is called the optical flow field. [8]*

Als optischer Fluss wird ein 2-dimensionales Vektorfeld bezeichnet, welches ein Geschwindigkeitsfeld, und somit ein Verschiebungsfeld in einem gegebenen Zeitraum, darstellt. Uneinig sind die beiden Definitionen jedoch darin, welche Verschiebungen dargestellt werden. Im Folgenden soll der optische Fluss die durch Veränderungen von Grauwert- oder Farbmustern wahrgenommene Bewegung in einem Bild  $I$  bezeichnen. Jedem Bildpunkt  $I(x, y, t)$  wird dann durch den optischen Fluss ein Verschiebungsvektor  $\vec{u} = (u, v)^T$  zugeordnet. Diese Verschiebungsvektoren entsprechen, entgegen dem zweiten Zitat, nicht unbedingt der Projektion der echten 3-dimensionalen Verschiebungsvektoren auf die Bildebene. Als Beispiel wird häufig eine Kugel mit gleichmäßiger matter Oberfläche herangezogen, die von einer Punktlichtquelle beleuchtet wird (Abbildung 2.1). Dreht sich diese Kugel, so ändert sich die Helligkeitsverteilung nicht – es können also keine Verschiebungen von Helligkeiten wahrgenommen werden und somit sind alle Verschiebungsvektoren des optischen Flusses genau null. Bewegt sich jedoch die Lampe um die nun feste Kugel, so ist eine Änderung der Helligkeitsverteilung wahrnehmbar. Dadurch ergibt sich eine wahrgenommene Bewegung und es resultieren hieraus somit von null verschiedene Verschiebungsvektoren. [4, 2, 17, 3]

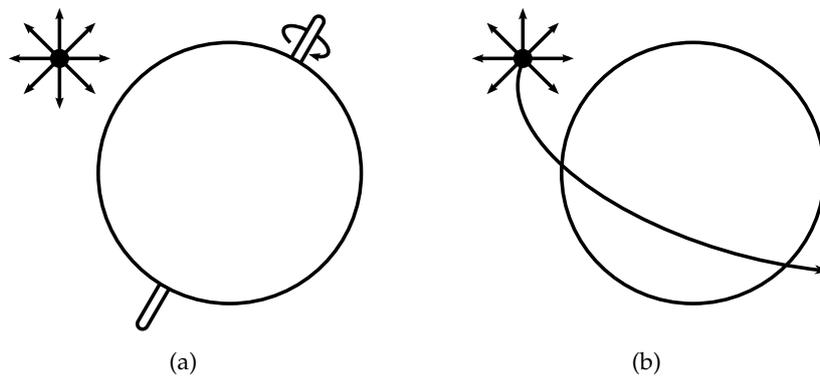


Abbildung 2.1: (a) rotierende Kugel ohne Änderung der Grauwerte, (b) Veränderung der Grauwerte durch bewegte Lichtquelle

Aus diesem Grund wird angenommen, dass die Beleuchtung auf den zu untersuchenden Bildern konstant ist und jegliche Änderung der Helligkeitsverteilung ausschließlich durch die gesuchte Bewegung entsteht. [17]

## 2.1 Kreuzkorrelation

Ein einfacher Ansatz zur Schätzung von Verschiebungsvektoren zwischen zwei Bildfunktionen geht über die Kreuzkorrelations-Funktion. Mit Hilfe der Kreuzkorrelation ist die Lokalisierung eines kleinen Bildausschnittes in einem größeren Bild möglich. Der Kreuzkorrelations-Term (2.1) beschreibt die Ähnlichkeit zwischen einer Schablone  $f$  an der Stelle  $u, v$  und einem Bild  $g$

$$c(u, v) = \sum_{x, y} f(x, y) \cdot g(x - u, y - v) \quad (2.1)$$

Eine Schätzung des optischen Flusses kann nun erfolgen, indem für jeden Bildpunkt im ersten von zwei aufeinanderfolgenden Bildern sein verschobenes Pendant gefunden wird. Dafür kann eine kleine Umgebung um den betrachteten Punkt als Schablone gewählt werden. Für diese wird dann in einer etwas größeren Umgebung, innerhalb einer gewählten Maximalverschiebung, für jede mögliche Verschiebung die Kreuzkorrelation im zweiten Bild ausgewertet. Der Punkt, an dem die Korrelation am größten ist, beschreibt die geschätzte Verschiebung.

Der Kreuzkorrelations-Koeffizient ist jedoch durch eine Reihe verschiedener Gründe fehleranfällig.

- Wenn sich die Helligkeit des Bildes stark mit der Position ändert, so kann die Korrelation zwischen der Schablone und der korrekten Zielregion geringer sein, als zwischen der Schablone und einer sehr hellen Zielregion.
- Der Wertebereich von  $c(u, v)$  ist abhängig von der Größe der betrachteten Region.
- Das Verfahren ist nicht invariant gegenüber Helligkeitsänderungen wie sie durch unterschiedliche Beleuchtung der Schablone und der Zielregion entstehen können.

Diesen Mängeln kann man durch die Verwendung des normierten Korrelations-Koeffizienten gerecht werden. Dabei wird der Mittelwert des Bildes bzw. der Schablone vom betrachteten Wert abgezogen und das Ergebnis der Kreuzkorrelation durch das Produkt der Standardabweichungen von  $f$  und  $g$  dividiert.

$$c_n(u, v) = \frac{\sum_{x,y} (f(x, y) - \bar{f}) \cdot (g(x - u, y - v) - \bar{g})}{\sqrt{\sum_{x,y} (f(x, y) - \bar{f})^2 \cdot \sum_{x,y} (g(x - u, y - v) - \bar{g})^2}} \quad (2.2)$$

Es ist jedoch nicht zu übersehen, dass der normierte Korrelations-Koeffizient immer noch einige Mängel hat. So ist das Verfahren nicht rotations- und skalierungsinvariant. Außerdem ist das Verfahren generell sehr rechenintensiv. Diese Probleme kann man jedoch vernachlässigen, wenn das Ausgangsmaterial in einer relativ hohen zeitlichen Auflösung vorliegt und somit zwischen zwei Bildern einer Sequenz nur minimale Veränderungen stattgefunden haben. Außerdem gibt es Verfahren, die die Kreuzkorrelation sehr effektiv über die Fourier-Transformation berechnen. Weiterhin ist jedoch ausschließlich die Schätzung von Bewegungsvektoren  $(u, v)^T$  mit ganzzahligen  $u$  und  $v$  möglich. [10, 6]

Aus diesen Gründen bedient man sich häufig eines differentiellen Ansatzes zur Schätzung von Verschiebungsvektoren.

## 2.2 Differentieller Ansatz

Für die Schätzung des optischen Flusses mit differentiellen Methoden wird angenommen, dass die Bildintensität der verschobenen Bildpunkte zwischen zwei aufeinander folgenden Bildern gleich bleibt [17]. Sei  $I(x, y, t)$  die Intensität eines Bildpunktes an  $x, y$  zum Zeitpunkt  $t$  der sich um  $\Delta x, \Delta y$  im Zeitraum  $\Delta t$  bewegt.

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (2.3)$$

Es wird angenommen, dass die Bewegung klein ist. Unter dieser Annahme kann die Bildfunktion  $I$  mittels einer Taylor-Reihe erster Ordnung angenähert werden.

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t \quad (2.4)$$

Aus (2.3) und (2.4) folgt dann direkt

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (2.5)$$

$$\frac{\partial I}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y} \frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} \frac{\Delta t}{\Delta t} = 0 \quad (2.6)$$

$$\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v = -\frac{\partial I}{\partial t} \quad (2.7)$$

wobei  $u$  und  $v$  die Komponenten des Verschiebungsvektors  $\vec{u}$  im optischen Flussfeld an  $x, y, t$  bezeichnen.

Es wird nun eine vereinfachende Schreibweise für partielle Ableitungen eingeführt. Die partielle Ableitung einer Funktion  $F$  nach der Variablen  $g$  wird im folgenden mit  $F_g$ , soweit dies sinnvoll ist, bezeichnet.

$$F_g := \frac{\partial F}{\partial g} \quad (2.8)$$

Mithilfe dieser Vereinfachung kann die oben hergeleitete Gleichung als

$$I_x u + I_y v = -I_t \quad (2.9)$$

geschrieben werden. Diese Gleichung (2.9) wird als *motion constraint equation* oder auch *gradient-constraint equation* bezeichnet [17, 6]. Die Schätzung des Verschiebungsvektors ist durch die *motion-constraint equation* jedoch allein nicht möglich, da sie zwei unbekannte Variablen besitzt und die Gleichung somit unterbestimmt ist. Sie bezeichnet eine Gerade im  $u, v$ -Raum, auf der die möglichen Lösungen für den Verschiebungsvektor  $(u, v)^T$  liegen müssen (Abbildung 2.2a) [17].

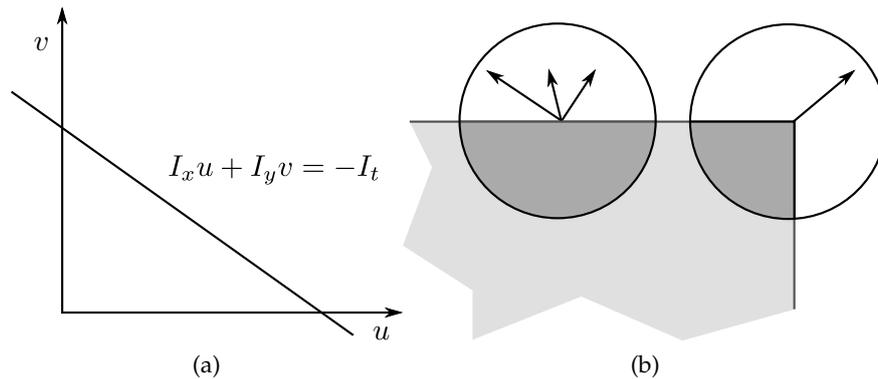


Abbildung 2.2: (a) Gerade im  $u, v$ -Raum, auf dem mögliche Verschiebungsvektoren liegen, (b) Unter der rechten Blende kann die Bewegung des grauen Rechtecks eindeutig ermittelt werden, unter der linken gibt es unendlich viele mögliche Bewegungen

Durch das Blendenproblem ist es nur möglich die Bewegung für einen Bildpunkt in Richtung seines Gradienten zu bestimmen. Für die Verschiebung  $u^\perp$  in Gradientenrichtung kann nun die folgende Formel hergeleitet werden [4]

$$u^\perp = \frac{-I_t}{\sqrt{I_x^2 + I_y^2}} \quad (2.10)$$

Gefragt bei der Schätzung von Verschiebungsvektoren ist jedoch nicht nur der Betrag in Richtung des Gradienten. Es ist außerdem von Interesse die Richtung der Bewegung sowie den Betrag in diese zu schätzen. Um das Blendenproblem zu überwinden, werden weitere Informationen zur Schätzung der Verschiebung benötigt. Dies kann jedoch nur durch das Treffen weiterer Annahmen an die Bewegungen oder über Modellwissen geschehen. Eine einfache Möglichkeit ist die Annahme einer lokalen Konstanz des optischen Flusses um jeden Bildpunkt. Betrachtet man so in einer kleinen Nachbarschaft um einen Punkt andere Punkte, die einen anderen Gradienten aufweisen, sich jedoch in die selbe Richtung bewegen, so entkommt man dem Blendenproblem [17, 7].

## 2.3 Lucas-Kanade Verfahren

Das Lucas-Kanade Verfahren ist eine relativ einfache Methode zur Schätzung des optischen Flusses in einer Sequenz aus zwei Bildern. Unter der Annahme, dass die Bewegung in einer kleinen Umgebung um einen Bildpunkt konstant ist, kann die Ver-

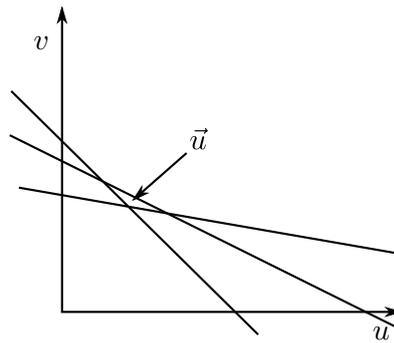


Abbildung 2.3: Viele sich im  $u, v$ -Raum kreuzende Linien und eine Schätzung für  $\vec{u}$ , die den Abstand zu allen Linien minimiert.

schiebung in diesem Bildpunkt geschätzt werden. Wichtig um dem Blendenproblem zu entkommen ist jedoch, dass innerhalb der betrachteten Umgebung mindestens zwei Punkte mit unterschiedlicher Gradientenrichtung vorkommen [?].

Wird nun für eine kleine  $m \times m$  Umgebung mit  $m > 0$  um den Bildpunkt  $x, y$  der optische Fluss als konstant angenommen, so muss die *motion-constraint equation* für jeden Punkt zugleich gelten. Die Bildpunkte werden mit  $n = 1, 2, \dots, m^2$  durchnummeriert und ein Gleichungssystem aus  $m^2$  Gleichungen wird aufgestellt. In diesem Gleichungssystem bezeichnet  $I_j$  die Intensität des Pixels mit der Nummer  $j$  und  $I_{jk}$  den Gradienten des  $j$ -ten Pixels in  $k$ -Richtung.

$$\begin{aligned}
 I_{1x}u + I_{1y}v &= -I_{1t} \\
 I_{2x}u + I_{2y}v &= -I_{2t} \\
 &\vdots \\
 I_{nx}u + I_{ny}v &= -I_{nt}
 \end{aligned} \tag{2.11}$$

Das entstandene Gleichungssystem ist nun überbestimmt und es gibt meistens keine eindeutige Lösung. Jede Gleichung des Gleichungssystems beschreibt eine Gerade im  $u, v$ -Raum, auf der die Lösung für den entsprechenden Punkt liegen muss (Abbildung 2.3). Der Schnittpunkt dieser Geraden beschreibt den zu schätzenden Verschiebungsvektor. Gibt es mehr als zwei Geraden, so gibt es häufig nicht nur einen Schnittpunkt. Es wird nun ein Vektor gesucht, der möglichst nahe an allen *constraint-lines* liegt und somit die *motion-constraint equations* für alle Bildpunkte am Besten erfüllt. Dafür werden die Quadrate der Abstände des gesuchten Vektors zu den Linien minimiert.

Die Summe der quadratischen Fehler kann nun, in Abhängigkeit von  $u$  und  $v$ , als

Fehlerfunktion  $E$  geschrieben werden.

$$E(u, v) = \sum_{j=1}^{m^2} \left( I_{j_x} u + I_{j_y} v + I_{j_t} \right)^2 \quad (2.12)$$

Um das Minimum unter Einfluss von  $u$  und  $v$  zu ermitteln, werden die partiellen Ableitungen der Fehlerfunktion nach  $u$  und  $v$  gebildet.

$$\begin{aligned} \frac{\partial E}{\partial u} &= 2 \sum_{j=1}^{m^2} \left( I_{j_x} u + I_{j_y} v + I_{j_t} \right) I_{j_x} \\ \frac{\partial E}{\partial v} &= 2 \sum_{j=1}^{m^2} \left( I_{j_x} u + I_{j_y} v + I_{j_t} \right) I_{j_y} \end{aligned} \quad (2.13)$$

Die partiellen Ableitungen werden nun in einem linearen Gleichungssystem mit Null gleichgesetzt. Durch Umformen können die Koeffizientenmatrix  $A$  und der Lösungsvektor  $b$  des linearen Gleichungssystem  $A\vec{u} = b$  bestimmt werden.

$$A = \begin{pmatrix} \sum_{j=1}^{m^2} I_{j_x}^2 & \sum_j I_{j_x} I_{j_y} \\ \sum_{j=1}^{m^2} I_{j_x} I_{j_y} & \sum_j I_{j_y}^2 \end{pmatrix} \quad b = \begin{pmatrix} - \sum_{j=1}^{m^2} I_{j_t} I_{j_x} \\ - \sum_{j=1}^{m^2} I_{j_t} I_{j_y} \end{pmatrix} \quad (2.14)$$

Durch Auflösen des Gleichungssystems nach dem Verschiebungsvektor  $\vec{u} = A^{-1} \cdot b$  kann dieser nun geschätzt werden.

### 2.3.1 Verbesserungen des Verfahrens

[7] Die Wahrscheinlichkeit, dass die Uniformität der Bewegung mit wachsender Entfernung zum betrachteten Punkt  $x, y$  abnimmt, ist in realen Anwendungen meist hoch. Eine Verbesserung des beschriebenen Verfahrens lässt sich durch die Gewichtung der einzelnen Quadrate während der Minimierung der Fehlerfunktion erreichen. Eine Gauß'sche Gewichtungsfunktion eignet sich gut, um die *constraint-lines* der Bildpunkte nahe  $x, y$  stärker zu gewichten, als die von weiter entfernten. Sei  $g_j$  das Gewicht für den  $j$ -ten Bildpunktes, so lässt sich die Fehlerfunktion wie folgt erweitern [7]

$$E(u, v) = \sum_{j=1}^{m^2} g_j \left( I_{j_x} u + I_{j_y} v + I_{j_t} \right)^2 \quad (2.15)$$

Nach der Minimierung wie oben beschrieben ergibt sich dann für die Koeffizientenmatrix  $A$  und für den Vektor  $b$

$$A = \begin{pmatrix} \sum_{j=1}^{m^2} g_j I_{j_x}^2 & \sum_j g_j I_{j_x} I_{j_y} \\ \sum_{j=1}^{m^2} g_j I_{j_x} I_{j_y} & \sum_j g_j I_{j_y}^2 \end{pmatrix} \quad b = \begin{pmatrix} - \sum_{j=1}^{m^2} g_j I_{j_t} I_{j_x} \\ - \sum_{j=1}^{m^2} g_j I_{j_t} I_{j_y} \end{pmatrix} \quad (2.16)$$

### 2.3.2 Eigenschaften des Verfahrens

Der *Lucas-Kanade*-Schätzer für das optische Flussfeld ist ein einfaches und leicht zu implementierendes Verfahren, welches auch heute noch eines der meist eingesetzten Verfahren ist. Auch die Annahme eines konstanten Flusses in einer kleinen Umgebung ist häufig gerechtfertigt. Da es sich jedoch um eine lokale Methode zur Schätzung der Verschiebung handelt, wird kein dichtes Vektorfeld geliefert. Die guten Flussinformationen, die an Objektträgern ermittelt werden können, verschwinden schnell mit wachsendem Abstand von diesen [2, 17].

## 2.4 Horn-Schunck Verfahren

Berthold Horn und Brian Schunck verwenden einen anderen Ansatz zur Lösung des Blendenproblems.

*In this case neighboring points on the objects have similar velocities and the velocity field of the brightness patterns in the image varies smoothly almost everywhere. [4]*

Wenn sich jeder Punkt unabhängig von den umgebenen Punkten bewegen würde, gäbe es keine richtige Möglichkeit, eine Bewegung zu schätzen. Deswegen nehmen Horn und Schunck an, dass sich das Vektorfeld über das Bild nur gleichmäßig ändert, woraus ein glattes Feld folgt. Um dieser Modellvorstellung gerecht zu werden, schlagen Horn und Schunck eine Glattheitsbedingung vor. Als Kriterium empfehlen sie, die Quadrate der Gradientenbeträge der Verschiebungsvektoren zu minimieren [4]:

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \quad \text{und} \quad \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \quad (2.17)$$

Neben diesem Glattheitsterm muss auch die *motion-constraint equation* minimiert werden. Daraus ergibt sich dann das global über alle Pixel  $\Omega$  zu minimierende Funktional

$$E(u, v) = \iint_{\Omega} \left( (I_x u + I_y v + I_t)^2 + \alpha^2 (|\nabla u|^2 + |\nabla v|^2) \right) dx dy \quad (2.18)$$

Der Faktor  $\alpha^2$  bestimmt, wie stark der Glättungsterm bei der Minimierung gewichtet wird. Größere Werte für  $\alpha^2$  führen zu einer stärkeren Glättung des Vektorfeldes. Horn und Schunck schlagen für  $\alpha^2$  einen Wert vor, der etwa dem Rauschen der Summe der Gradientenquadrate  $E_x^2 + E_y^2$  entspricht.

$E$  kann durch lösen der zugehörigen Euler-Lagrange Gleichungen minimiert werden (dazu in Kapitel drei mehr) [15]. Es ergeben sich zwei Gleichungen

$$\begin{aligned} I_x(I_x u + I_y v + I_t) - \alpha^2 \Delta v_x &= 0 \\ I_y(I_x u + I_y v + I_t) - \alpha^2 \Delta v_x &= 0 \end{aligned} \quad (2.19)$$

Der Laplace-Operator  $\Delta$  lässt sich nun numerisch approximieren und als  $\Delta s(x, y) = \bar{s}(x, y) - s(x, y)$  ausdrücken, wobei  $\bar{s}$  die gewichtete Summe der Intensitätswerte in einer Nachbarschaft um  $x, y$  bezeichnet [4, 17]. Die Gleichungen (2.19) lassen sich nun umformen zu

$$\begin{aligned} (I_x^2 + \alpha^2) u + I_x I_y v &= \alpha^2 \bar{u} - I_x I_t \\ I_x I_y u + (I_y^2 + \alpha^2) v &= \alpha^2 \bar{v} - I_y I_t \end{aligned} \quad (2.20)$$

Ein iteratives Lösungsverfahren bietet sich aufgrund der sehr hohen Zahl unbekannter Variablen und Gleichungen an. Mithilfe des Gauss-Seidel Verfahrens kann ein neuer Geschwindigkeitsvektor  $(u^{n+1}, v^{n+1})^T$  für jeden Bildpunkt aus den vorher ermittelten Schätzungen  $(u^n, v^n)^T$  berechnet werden [4, 17].

$$u^{n+1} = \bar{u} - \frac{I_x (I_x \bar{u} + I_y \bar{v} + I_t)}{\alpha^2 + I_x^2 + I_y^2} \quad (2.21)$$

$$v^{n+1} = \bar{v} - \frac{I_y (I_x \bar{u} + I_y \bar{v} + I_t)}{\alpha^2 + I_x^2 + I_y^2} \quad (2.22)$$

### 2.4.1 Eigenschaften des Verfahrens

Durch den iterativen Charakter ist das Horn-Schunck Verfahren sehr rechenaufwendig, da häufig dutzende oder gar hunderte Iterationen benötigt werden, um ein genügend dichtes Vektorfeld zu bekommen. Durch den Glattheitsterm werden dann jedoch die Verschiebungsvektoren von Stellen mit höheren Gradientenbeträgen in die Bereiche mit schwachen Gradienten propagiert. Dadurch ergibt sich auch ein dichtes Vektorfeld in Bereichen, in denen lokale Verfahren keine Verschiebungen registrieren (siehe Abbildung 2.4) [17, 2].

Barron und Fleet haben in ihrem Artikel verschiedene Verfahren zur Schätzung des optischen Flusses verglichen und dabei experimentell herausgefunden, dass die Anfälligkeit für Rauschen beim Horn-Schunck Verfahren größer ist, als für lokale Operatoren, wie dem Lucas-Kanade Verfahren [2].

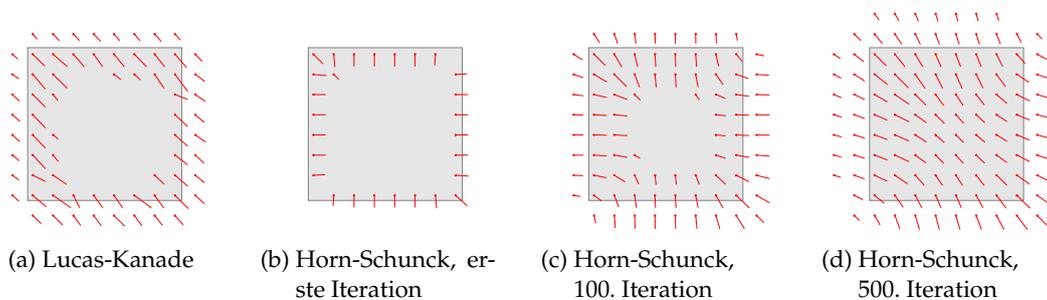


Abbildung 2.4: (a) lokaler Lucas-Kanade Operator, (b, c, d) Horn-Schunck mit unterschiedlicher Zahl Iterationen. Es ist gut zu erkennen, wie die Flussinformationen in die Flächen hineinpropagiert werden.

## 3 Multispektralbilder

### 3.1 Einführung

Ein Multispektralbild enthält im Gegensatz zu Grauwertbildern nicht nur einen Kanal. Es werden nun pro Bild mehrere Grauwertbilder gespeichert, welche die selbe Szene in der selben Auflösung zeigen, jedoch Strahlungsintensität in unterschiedlichen Wellenlängenbereichen bezeichnen. Abhängig davon, wie gut ein Objekt elektromagnetische Strahlung in einem bestimmten Wellenlängenbereich reflektiert, ist es auf den einzelnen Kanälen des Multispektralbildes auszumachen. Ein bekanntes Beispiel für Multispektralbilder sind Farbbilder, die aus den drei Kanälen *Rot*, *Grün* und *Blau* bestehen. Die Wellenlängenbereiche müssen jedoch nicht zwingend im sichtbaren Bereich des Lichtes liegen, sondern es können beispielsweise Radarwellen oder Wellen im Bereich des Infrarots für einen Kanal in einem Multispektralbild genutzt werden [11].

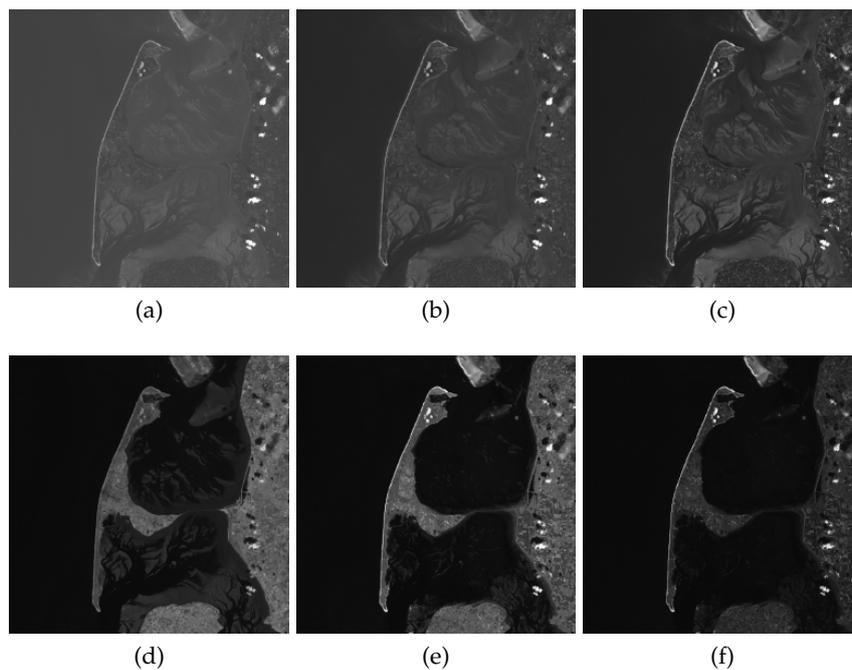


Abbildung 3.1: Einzelne Kanäle einer Multispektralaufnahme von Sylt. (a), (b) und (c) sind Aufnahmen aus dem sichtbaren Bereich, (d) und (e) nahes Infrarot und (f) mittleres Infrarot.

In Abbildung 3.1 sind einzelne Kanäle einer multispektralen Aufnahme von Sylt angeführt. (a), (b) und (c) zeigen die Intensitätsbilder im sichtbaren Bereich, wobei (a) blau, (b) grün und (c) rotes Licht zeigt. Dabei kann gut der Meeresboden im flachen Wasser zwischen der Insel und dem Festland erkannt werden. Er hebt sich jedoch trotz des Wassers kaum von den Landmassen ab. Die Kanäle (d) und (e) sind im Bereich des nahen Infrarots aufgenommen. Hier wird mit steigender Wellenlänge der Unterschied zwischen Wasser, welches die Strahlung absorbiert, und dem Land, das die Infrarotstrahlung von der Sonne besonders gut reflektiert, sehr deutlich. Der letzte Kanal (f) ist aus dem Bereich des mittleren Infrarots [20]. Wie gut zu erkennen ist, erhält man durch das gleichzeitige Betrachten mehrerer Kanäle häufig mehr Informationen über die aufgenommenen Bereiche. Unter der Annahme, dass eine aufgenommene Bewegung in mehreren oder allen Kanälen einer Multispektralaufnahme erkennbar ist, sollte es nun möglich sein, das Blendenproblem bei der Berechnung von Verschiebungsvektoren durch die zusätzlichen Informationen aus anderen Kanälen zu lösen. Dabei müssen auf den unterschiedlichen Kanälen natürlich unterschiedliche Bildstrukturen zu sehen sein, was jedoch später noch näher erläutert werden soll.

Es gibt nun unterschiedliche Ansätze Verschiebungsvektoren in Multispektralbildern zu berechnen. Zwei einfache Verfahren seien hier kurz angerissen [17]:

1. Es bietet sich an, die Kanäle der Multispektralaufnahme in eine Grauwertaufnahme zu kombinieren und auf diese die bereits bekannten Verfahren zur Schätzung der Flussvektoren anzuwenden. Zur Kombination eignet sich beispielsweise das arithmetische Mittel der Kanäle. Dies verbessert normalerweise das Signal-Rausch-Verhältnis des Ausgangsbildes, verwendet jedoch die eigentlichen zusätzlichen Informationen aus den Kanälen kaum.
2. Eine weitere Möglichkeit ist das Anwenden der bereits bekannten Verfahren auf jeden Kanal einzeln. Die daraus gewonnen Vektorfelder können dann geeignet miteinander verrechnet werden. So kann beispielsweise der Mittelwert berechnet werden oder es können nur die Verschiebungsvektoren übernommen werden, die in den meisten Kanälen mit einer gewissen Toleranz gefunden wurden.

Beide Verfahren lassen sich ohne großen Aufwand umsetzen, da bereits bekannte Algorithmen unverändert weiter genutzt werden können. Es werden jedoch die Mehrinformationen aus den Multispektralbildern kaum in den Schätzungsprozess einbezogen.

In diesem Kapitel sollen deshalb Verfahren entwickelt werden, welche die Informationen der Multispektralbilder besser nutzen. Als Grundlage zur Konzeption eines

lokalen Verfahrens wird die Lucas-Kanade Methode benutzt und entsprechend erweitert, um einen Schätzer zu finden, der den Fehler der *motion-constraint equation* im Kontext aller Kanäle gemeinsam minimiert. Als Zweites soll ein globales Verfahren hergeleitet werden. Dieses optimiert einerseits die Bewegungsgleichung und andererseits eine weitere Bedingung an das Vektorfeld und baut damit auf dem Horn-Schunck-Verfahren auf.

### 3.2 Multispektrales Lucas-Kanade Verfahren

Das bereits für Grauwertbilder vorgestellte Lucas-Kanade Verfahren lässt sich sehr einfach auf Multispektralbilder erweitern. Es müssen im Voraus jedoch drei Annahmen an die Quellbilder und das Flussfeld gestellt werden.

1. Der Fluss soll für jeden Kanal durch das selbe Verschiebungsfeld bezeichnet werden. Dies bedeutet also, dass auf jedem Kanal Bewegung in die selbe Richtung zu erkennen ist.
2. Die Gradientenrichtung sollte in jedem Punkt für eine gute Schätzung möglichst unterschiedlich in den Kanälen sein.
3. Der optische Fluss um einen beliebigen Punkt  $x, y$  kann als konstant angenommen werden, wie bereits für das Lucas-Kanade Verfahren für Grauwertbilder.

Ausgehend von der ersten Annahme muss die *motion-constraint equation* für den selben Verschiebungsvektor  $(u, v)^T$  auf allen Kanälen erfüllt werden. Da dies bei mehr als zwei Kanälen durch Rauschen und Quantisierung nicht zwingend möglich ist, wird wieder eine Ballungsanalyse durchgeführt, bei der ein Vektor  $(u, v)^T$  gesucht wird, der die Verschiebung angibt und zu dem kleinsten quadratischen Fehler beim Erfüllen der Bedingungsgleichungen führt. Dies erlaubt das Aufstellen einer Fehlerfunktion  $E(u, v)$ . Sei  $K = \{I_1, I_2, \dots\}$  die Menge der Bildfunktionen der zu betrachtenden Kanäle des Bildes

$$E(u, v) = \sum_{k \in K} (k_x u + k_y v + k_t)^2 \quad (3.1)$$

Die Minimierung dieser Fehlerfunktion allein reicht bei mehr als einem Kanal bereits zur Schätzung einer Verschiebung aus. Das Ergebnis würde jedoch durch Rauschen

im Ausgangsmaterial häufig stark von der erwarteten Verschiebung abweichen. Weiterhin muss dann für jeden Pixel die oben genannte zweite Bedingung, unterschiedliche Gradienten für jeden Kanal, in allen Punkten gelten. Aus diesen Gründen wird wie im Graustufen-Fall die lokale Konstanz der Verschiebung gefordert. Es wird angenommen, der Fluss sei in einer kleinen  $m \times m$ ,  $m > 0$  Umgebung um den Punkt  $x, y$  konstant. Dann kann ein Gleichungssystem mit einer Bedingungsgleichung für jeden Pixel  $j \in \{1, 2, \dots, m^2\}$  innerhalb der Umgebung für jeden Kanal  $k \in K$  aufgestellt werden. Dabei bezeichnet  $k_i$  den  $i$ -ten Pixel in der Umgebung um  $x, y$  in der Bildfunktion  $k$  und  $k_{ij}$  den Gradienten an dem Pixel in  $j$ -Richtung.

$$\begin{aligned} k_{1x}u + k_{1y}v + k_{1t} &= 0 \\ k_{2x}u + k_{2y}v + k_{2t} &= 0 \\ &\vdots \\ k_{m^2x}u + k_{m^2y}v + k_{m^2t} &= 0 \end{aligned} \quad (3.2)$$

Dieses Gleichungssystem ist überbestimmt und eine eindeutige Lösung nicht möglich, weswegen sich eine Schätzung für  $u, v$  mithilfe der Kleinste-Quadrate-Methode anbietet. Wie bereits für den Grauwert-Fall beschrieben wird eine Gewichtungsfunktion  $g_j$  verwendet, damit Pixel, die weiter von  $x, y$  entfernt sind, weniger stark in die Schätzung einbezogen werden. Die zu minimierende Funktion lautet nun

$$E(u, v) = \sum_{k \in K} \left[ \sum_j g_j (k_x u + k_y v + k_t)^2 \right] \quad (3.3)$$

Zur Berechnung des Minimums müssen die partiellen Ableitungen bestimmt und dann gemeinsam gleich Null gesetzt werden.

$$\begin{aligned} 0 &= 2 \sum_{k \in K} \left[ \sum_j g_j k_x (k_x u + k_y v + k_t) \right] \\ 0 &= 2 \sum_{k \in K} \left[ \sum_j g_j k_y (k_x u + k_y v + k_t) \right] \end{aligned} \quad (3.4)$$

Durch Umformen erhält man ein lineares Gleichungssystem der Form  $A\vec{u} = b$  mit

$$A = \sum_{k \in K} \begin{pmatrix} \sum_j^{m^2} g_j (k_x)^2 & \sum_j^{m^2} g_j k_x k_y \\ \sum_j^{m^2} g_j k_x k_y & \sum_j^{m^2} g_j (k_y)^2 \end{pmatrix} \quad (3.5)$$

$$b = \sum_{k \in K} \begin{pmatrix} \sum_j^{m^2} g_j k_t k_x \\ \sum_j^{m^2} g_j k_t k_y \end{pmatrix} \quad (3.6)$$

Die Lösung dieses Gleichungssystem kann dann mit  $\vec{u} = A^{-1}b$  bestimmt werden.

### 3.3 Multispektrales Horn-Schunck Verfahren

Das auf dem Lucas-Kanade aufbauende Verfahren aus dem vorigen Abschnitt ist ein lokales Verfahren und liefert deswegen nur an Stellen eine zuverlässige Schätzung von Verschiebungsvektoren, die bestimmten Forderungen genügen. In diesem Abschnitt soll ein globales Verfahren zur Schätzung des optischen Flusses anhand des Horn-Schunck Verfahren hergeleitet werden, welches die zusätzlichen Informationen aus Mehrkanalbildern verwendet. Dabei soll neben der Bedingung, die die *motion-constraint equation* für jeden Punkt in jedem Kanal stellt, noch eine weitere Bedingung an die Glattheit des resultierenden Vektorfeldes gestellt werden.

Das von Horn-Schunck vorgeschlagene Verfahren optimiert einerseits die *motion-constraint equation*  $I_x u + I_y v + I_t$ , die quadratisch für jeden Bildpunkt in der Punktmenge  $\Omega$  in das zu minimierende Fehlerfunktional  $E$  einfließt.  $\Omega$  bezeichnet dabei alle Bildpunkte des betrachteten Bildes. Es bietet sich nun an, diesen Term für jeden Kanal aus der Menge  $K$  quadratisch in  $E$  einfließen zu lassen. Der von der Bewegungsgleichung ausgehende zu minimierende Fehler  $\xi_m$  lautet dann

$$\xi_m = \sum_{k \in K} (k_x u + k_y v + k_t)^2 \quad (3.7)$$

Zum Anderen fließt ein mit einem Gewichtungsfaktor  $\alpha^2$  multiplizierter Glattheitsterm in das Fehlerfunktional ein. Dieser wird zusätzlich mit der Anzahl der betrachteten Kanäle gewichtet, damit seine Relevanz auch bei einer erhöhten Kanalzahl bestehen bleibt. Es wird die selbe Glattheitsbedingung an das resultierende Vektorfeld

gestellt, wie Horn und Schunck sie vorgeschlagen haben. Als  $\xi_c$  bezeichnet ergibt sich dafür das Fehlerfunktional

$$\xi_c = \alpha^2 |K| (u_x^2 + u_y^2 + v_x^2 + v_y^2) \quad (3.8)$$

Aus  $\xi_m$  und  $\xi_c$  ergibt sich dann das über alle Bildpunkte  $\Omega$  zu minimierende Fehlerfunktional  $E$ , also

$$E = \iint_{\Omega} \xi_m + \xi_c \, dx dy \quad (3.9)$$

$$= \iint_{\Omega} \sum_{k \in K} (k_x u + k_y v + k_t)^2 + \alpha^2 |K| (u_x^2 + u_y^2 + v_x^2 + v_y^2) \, dx dy \quad (3.10)$$

Über die Variationsrechnung und mit Hilfe des *Euler-Lagrange Formalismus* kann für  $E$  die Lagrangefunktion  $\mathcal{L}$  aufgestellt werden, aus welcher dann direkt ein Gleichungssystem abgeleitet werden kann, welches  $\mathcal{L}$  und somit auch  $E$  minimiert [4, 15].

$$\mathcal{L}(x, y, u, v, u_x, u_y, v_x, v_y) = \xi_m + \xi_c \quad (3.11)$$

Das Gleichungssystem, welches  $\mathcal{L}$  minimiert, lautet

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial u} - \frac{d}{du} \frac{\partial \mathcal{L}}{\partial u_x} - \frac{d}{du} \frac{\partial \mathcal{L}}{\partial u_y} &= 0 \\ \frac{\partial \mathcal{L}}{\partial v} - \frac{d}{dv} \frac{\partial \mathcal{L}}{\partial v_x} - \frac{d}{dv} \frac{\partial \mathcal{L}}{\partial v_y} &= 0 \end{aligned} \quad (3.12)$$

Einsetzen von  $\mathcal{L}$  in die Gleichungen und das anschließende partielle Ableiten ergibt das zu lösende Gleichungssystem

$$\begin{aligned} \sum_{k \in K} 2k_x (k_x u + k_y v + k_t) - 2\alpha^2 |K| u_{xx} - 2\alpha^2 |K| u_{yy} &= 0 \\ \sum_{k \in K} 2k_y (k_x u + k_y v + k_t) - 2\alpha^2 |K| v_{xx} - 2\alpha^2 |K| v_{yy} &= 0 \end{aligned} \quad (3.13)$$

Mit dem Laplace-Operator  $\Delta s = \nabla^2 s = s_{xx} + s_{yy}$  können die partiellen zweifachen

Ableitungen von  $u$  und  $v$  in  $x$ - und  $y$ -Richtung zusammengefasst werden.

$$\begin{aligned} \sum_{k \in K} k_x (k_x u + k_y v + k_t) - \alpha^2 |K| \nabla^2 u &= 0 \\ \sum_{k \in K} k_y (k_x u + k_y v + k_t) - \alpha^2 |K| \nabla^2 v &= 0 \end{aligned} \quad (3.14)$$

Der Laplace-Operator kann mittels einer Approximation direkt berechnet werden, so dass die partiellen Ableitungen wegfallen. Es gilt in einer guten Näherung  $\nabla^2 = \bar{s} - s$ , wobei  $\bar{s}$  den gewichteten Mittelwert von  $s$  in einer kleinen Nachbarschaft bezeichnet [4, 17]. Mit dieser Vereinfachung lässt sich die Gleichung schreiben als

$$\begin{aligned} \sum_{k \in K} k_x (k_x u + k_y v + k_t) - \alpha^2 |K| (\bar{u} - u) &= 0 \\ \sum_{k \in K} k_y (k_x u + k_y v + k_t) - \alpha^2 |K| (\bar{v} - v) &= 0 \end{aligned} \quad (3.15)$$

Die Klammer im rechten Summanden kann nun ausmultipliziert werden und  $\alpha^2 K \bar{u}$  bzw.  $\bar{v}$  in die Summe auf der linken Seite gezogen werden. Aufteilen der Summe und Ausklammern von  $u$  und  $v$  führen zu einem linearen Gleichungssystem mit den Unbekannten  $u$  und  $v$ . Es ist zu beachten, dass dieses Gleichungssystem für jeden Bildpunkt in  $\Omega$  aufgestellt werden muss.

$$\begin{aligned} \left[ \sum_{k \in K} (\alpha^2 + k_x^2) \right] u + \left[ \sum_{k \in K} k_x k_y \right] v &= \alpha^2 |K| \bar{u} - \sum k_t k_x \\ \left[ \sum_{k \in K} k_x k_y \right] u + \left[ \sum_{k \in K} (\alpha^2 + k_y^2) \right] v &= \alpha^2 |K| \bar{v} - \sum k_t k_y \end{aligned} \quad (3.16)$$

Die Koeffizientenmatrix  $A$  des Gleichungssystems und ihre Determinante können nun berechnet werden.

$$A = \begin{pmatrix} \sum_{k \in K} (\alpha^2 + k_x^2) & \sum_{k \in K} k_x k_y \\ \sum_{k \in K} k_x k_y & \sum_{k \in K} (\alpha^2 + k_y^2) \end{pmatrix} \quad (3.17)$$

$$\det(A) = \left[ \sum_{k \in K} (\alpha^2 + k_x^2) \right] \left[ \sum_{k \in K} (\alpha^2 + k_y^2) \right] - \left[ \sum_{k \in K} k_x k_y \right]^2 \quad (3.18)$$

Mit der Determinante von  $A$  lässt sich eine iterative Lösung für  $u^{(n+1)}$  und  $v^{(n+1)}$  angeben, die von dem gemittelten Wert  $\overline{u^{(n)}}$  und  $\overline{v^{(n)}}$  der letzten Iteration abhängt.

$$u^{(n+1)} = \frac{\left[ \alpha^2 K \overline{v^{(n)}} - \sum_{k \in K} k_t k_y \right] \left[ \sum_{k \in K} k_x k_y \right] + \left[ \sum_{k \in K} k_t k_x - \alpha^2 |K| \overline{u^{(n)}} \right] \left[ \sum_{k \in K} (\alpha^2 + k_y^2) \right]}{-\det(A)} \quad (3.19)$$

$$v^{(n+1)} = \frac{\left[ \alpha^2 K \overline{u^{(n)}} - \sum_{k \in K} k_t k_x \right] \left[ \sum_{k \in K} k_x k_y \right] + \left[ \sum_{k \in K} k_t k_y - \alpha^2 |K| \overline{v^{(n)}} \right] \left[ \sum_{k \in K} (\alpha^2 + k_x^2) \right]}{\det(A)} \quad (3.20)$$

## 4 Vergleich der Verfahren

In diesem Kapitel werden die in dieser Arbeit hergeleiteten Verfahren für Mehrkanalbilder mit den bekannten, grauwertbasierten Verfahren verglichen. Dazu werden zuerst die verwendeten synthetischen und realen Testbilder beschrieben. Weiterhin wird auf die Testumgebung und auf die Implementierung der Verfahren im Rahmen dieser Arbeit eingegangen, sowie auf die verwendeten Testmethoden und Vergleichsmaße. Abschließend werden die berechneten Ergebnisse vorgestellt und interpretiert.

### 4.1 Testdaten

Es existieren für Grauwertverfahren viele bekannte Testsequenzen wie die *Yosemite-Sequenz*, die *NASA-Sequenz* oder das *Hamburg Taxi*. Zu einigen dieser Bildfolgen existieren *Ground Truth*-Werte. Dabei handelt es sich um Vektorfelder mit den tatsächlichen Verschiebungsvektoren zwischen den Bildern. Diese können dann mit den geschätzten Daten unterschiedlicher Verfahren verglichen werden [2]. Zum Testen von Multispektralverfahren existieren jedoch nur einige Farbsequenzen. Auf die Verwendung dieser Farbsequenzen wird hier verzichtet, da es sich um synthetische Sequenzen handelt, die nicht ausreichend unterschiedliche Struktur und verschiedene Gradienten in den drei Kanälen aufweisen.

Um die entwickelten Verfahren dennoch testen zu können, wird ein eigenes Testbild entwickelt. Dabei wird darauf geachtet, dass die Gradientenrichtungen in den Kanälen

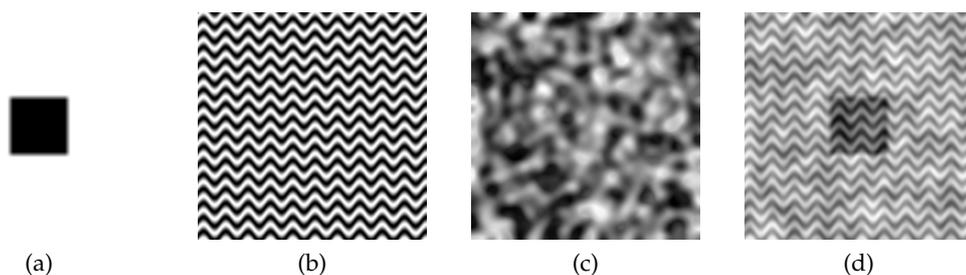


Abbildung 4.1: Generierte Testbilder. (a) unscharfes Rechteck, (b) Kombination aus sin- und cos-Funktion, (c) extrem hochskaliertes Rauschen, (d) Mittelwert der Bilder (a), (b) und (c)

möglichst verschieden sind. Das Testbild besteht dafür aus drei Kanälen, die in Abbildung 4.1 dargestellt sind. Es werden zwei Kopien angefertigt, wobei die zweite Kopie um einen Pixel nach unten und einen Pixel nach rechts verschoben ist. Auf beiden Bildern wird in jedem Kanal unabhängiges Gauß'sches Rauschen hinzugefügt, um eine realitätsnähere Testsituation zu erzeugen. Für die Analyse der grauwertbasierten Verfahren werden die drei Kanäle jeweils zu gleichen Teilen in ein gemitteltes Grauwertbild gespeichert (ebenfalls Abbildung 4.1).

Neben dieser generierten Sequenz werden die Verfahren an einem Ausschnitt eines Satellitenbildes vom 30.07.1999, aufgezeichnet mit einem *WiFS*-Sensor, getestet. Der Ausschnitt ist im zweiten Bild der Sequenz um  $1^\circ$  im Uhrzeigersinn um den Mittelpunkt des Bildes gedreht. Es handelt sich um eine Aufnahme mit zwei Kanälen. Für den Vergleich mit einem Grauwertverfahren wird wieder der Mittelwert der beiden Kanäle in ein Graustufen-Bild gespeichert (Abbildung ??).

Weiterhin wird das Verfahren noch qualitativ an einer Zweikanal-Sequenz des selben Satelliten getestet (Abbildung ??). Diese variiert jedoch von der ersten zur zweiten Aufnahme stark in der Helligkeit, wodurch der zeitliche Gradient nicht brauchbar ist. Aus diesem Grund wird der Hintergrund der einzelnen Bilder mittels eines Medianfilters mit geeignetem Filterradius herausgerechnet.

Eine weitere Quelle für multispektrale Satellitenbilder sind geostationäre Wettersatelliten, deren Bildmaterial im Internet häufig kostenlos verfügbar ist. Ein Ausschnitt vom 10. April 2010, der einen Windstrom über dem Nordpazifik zeigt, wird für einen weiteren qualitativen Vergleich gewählt. Die Daten stammen von dem japanischen Satelliten *MTSat-1R* [9] und bestehen aus drei unterschiedlichen Kanälen. Ein Kanal zeigt kurzwelliges Infrarot (um  $3.75\mu\text{m}$ ), ein weiterer zeigt Wasserdampf (aufgezeichnet um  $6.75\mu\text{m}$ ) und der letzte Kanal wurde im Bereich des fernen Infrarots um  $10.8\mu\text{m}$  aufgezeichnet [16].

## 4.2 Implementierung und Testumgebung

Die beiden hergeleiteten Algorithmen wurden im Rahmen dieser Bachelorarbeit in der Sprache C++ unter Verwendung der *VIGRA*-Bibliothek implementiert. Die entsprechenden Funktionen nehmen als Parameter jeweils die im Voraus berechneten Gradienten der Bildfunktion in den beiden Ortsrichtungen sowie in zeitlicher Richtung entgegen. Die Berechnung dieser kann nach verschiedenen Verfahren geschehen und wurde deshalb aus der eigentlichen Funktion ausgegliedert. Weiterhin werden zusätzliche Pa-

parameter wie die Gewichtungskonstante  $\alpha$  oder die Größe der Umgebung im Lucas-Kanade Algorithmus übergeben. Das resultierende Flussfeld wird dann pixelweise in eine Zweikanalgrafik geschrieben. Diese kann dann beispielsweise als SVG-Grafik ausgegeben werden oder anderweitig weiterverarbeitet werden.

Bei beiden Verfahren geschieht die Berechnung eines Punktes im Flussfeld unabhängig von der Berechnung der umgebenen Punkte. Im Lucas-Kanade Verfahren ist dies offensichtlich, da es sich um ein lokales Verfahren handelt. Im Horn-Schunck Algorithmus hängt die Berechnung zwar von den Werten in einer Nachbarschaft ab, jedoch beziehen sich diese Werte auf eine vorige Iteration. Daraus folgt, dass die beiden eingesetzten Verfahren auf moderner Hardware sehr gut parallelisierbar sind. In C und C++ kann dies durch die Verwendung von *OpenMP*, welches ausschließlich die `#pragma`-Präprozessordirektive nutzt, sehr einfach und portabel erreicht werden [14]. Die Performance der Algorithmen steigt durch die Verwendung von *OpenMP* in eigenen Tests etwa linear mit der Anzahl der CPU-Kerne. Denkbar ist auch die Implementierung der Verfahren auf einer modernen GPU, wie unter anderem von *Onera – The French Aerospace Lab* durchgeführt [13].

Die folgenden Tests wurden mit dem *GNU C++-Compiler* in der Version 4.4.4 und der Version 1.6.0 der *VIGRA*-Bibliothek auf einer Pentium 4 CPU durchgeführt.

### 4.3 Vergleichsmethodik

Da für die synthetischen Testsequenzen die genauen Verschiebungsvektoren bekannt sind, lassen sich quantitative Verfahren für die Güte der Schätzung definieren. Es gibt jedoch verschiedene Möglichkeiten, die Ground-Truth-Daten mit den berechneten Vektorfeldern zu vergleichen. Dabei sollen einige Maße exemplarisch herausgegriffen werden.

Im Folgenden wird unter anderem das von Barron, Fleet und Beauchemin vorgeschlagene Maß zur Messung des Fehlers der Geschwindigkeitsvektoren benutzt. Die Geschwindigkeit wird dabei als ein Raum-Zeit-Vektor der Form  $\vec{v} = (u, v, 1)^T$  in Einheiten von (Pixel, Pixel, Zeit) geschrieben. Um die Abweichungen im Vektorfeld zu messen empfiehlt es sich nun den Fehler als Winkelabweichung vom korrekten Geschwindigkeitsvektor zu messen. Der 2-dimensionale Verschiebungsvektor  $\vec{u} = (u, v)^T$  kann dann als 3-dimensionaler Geschwindigkeitsvektor  $\vec{v}$  geschrieben werden und der Winkelfehler

$\phi_E$  zur wahren Geschwindigkeit  $\vec{v}_c$  berechnet werden [2].

$$\vec{v} = \sqrt{\frac{1}{u^2 + v^2 + 1}} \cdot (u, v, 1)^T \quad (4.1)$$

$$\phi_E = \arccos(\vec{v} \cdot \vec{v}_c) \quad (4.2)$$

Es ist wichtig zu erwähnen, dass dieser Winkelfehler nicht mit dem Fehler der Richtungen der Verschiebungsvektoren im Ergebnis gleichzusetzen ist. Dieses Maß hängt sowohl von der Richtung der Vektoren als auch von deren Länge ab. Dadurch führen Richtungsabweichungen bei einer kleinen Geschwindigkeit zu einem nicht so großen Fehler, wie Abweichungen bei höheren Geschwindigkeiten. Insbesondere bewirken auch Vektoren, die zwar in die richtige Richtung zeigen, jedoch einen fehlerhaften Betrag haben, einen Fehler [2].

Auch kann eine Aussage zum Betrag der Verschiebungsvektoren gemacht werden. Zur Bestimmung der Güte kann hier ein einfaches Fehlermaß angewendet werden, wie der Mittelwert der quadratischen Fehler. Es gilt also für den Fehler  $|\vec{v}|_E$  des einzelnen Geschwindigkeitsvektors

$$|\vec{v}|_E = (|\vec{v}| - |\vec{v}_c|)^2 \quad (4.3)$$

Für den lokalen Lucas-Kanade Schätzer kann zusätzlich noch eine Angabe über die Dichte  $\rho$  des resultierenden Vektorfeldes gemacht werden. Mit der Dichte des Feldes wird in dieser Arbeit der Prozentsatz der Vektoren bezeichnet, der innerhalb einer Toleranz  $\tau$  von Null verschieden ist. Je größer die Dichte  $\rho$  ist, desto dichter ist das bezeichnete Vektorfeld. Aus allen Vektoren  $\vec{v}_n$  in dem Vektorfeld  $F$  folgt dann

$$\rho = \frac{1}{|F|} \sum_{v \in F} f(\vec{v}) \text{ mit } f(\vec{v}) = \begin{cases} 0 & \text{falls } |\vec{v}| < \tau \\ 1 & \text{sonst} \end{cases} \quad (4.4)$$

Es wird jeweils die Verschiebung zwischen zwei Mehrkanalbildern mit den entwickelten Verfahren geschätzt (Lucas-Kanade MS und Horn-Schunck MS). Weiterhin werden als Vergleich die Mittelwerte der Kanäle gebildet und auf den resultierenden Grauwertbildern mithilfe der bereits bekannten Verfahren ebenfalls die Verschiebung geschätzt (Lucas-Kanade GW und Horn-Schunck GW). Die resultierenden Vektorfelder werden dann qualitativ und quantitativ ausgewertet, soweit es möglich ist.

## 4.4 Ergebnisse

Für die synthetisch generierte Bildfolge liefern alle Verfahren relativ gute Ergebnisse (siehe Tabelle 4.1). Der Winkelfehler  $\phi_E$  ist bei den Mehrkanal-Verfahren minimal. Die Ursache dafür sind die zusätzlichen Informationen aus den multispektralen Bildern, die besser ausgewertet werden können. Erwähnenswert ist auch, dass das Lucas-Kanade Verfahren mit mehreren Kanälen in diesem Test einen geringeren Winkelfehler produzierte als das Horn-Schunck Verfahren auf Grauwertbildern. Dies ist offensichtlich auf die Kombination des Sinus/Cosinus-Kanals und des Rausch-Kanals zurückzuführen. Diese beiden Kanäle sorgen dafür, dass in fast jedem Punkt in den Kanälen ein starker Unterschied in den Gradientenrichtungen zu finden ist. Die relativ hohen Winkelfehler für das Horn-Schunck Verfahren sind auf die Randregionen zurückzuführen, in denen das sonst sehr glatte Vektorfeld einige Ausreißer aufweist.

Die zweite Testsequenz, der rotierte Ausschnitt aus einem Satellitenbild, bringt das Lucas-Kanade-Verfahren bereits an seine Grenzen. Das geschätzte Flussfeld ist in Abbildung ??, die Messwerte in Tabelle 4.2 zu sehen. Das Vektorfeld für das Graustufenbild ist sehr schwach besetzt. Nur für etwa 20 % der Bildpunkte konnten Verschiebungsvektoren geschätzt werden. Dies liegt an den großen, nahezu gradientenfreien Flächen. Die Mehrkanalversion des Algorithmus kann hier teilweise noch unterschiedliche Gradientenrichtungen in den verschiedenen Kanälen ausmachen, wodurch weitere Verschiebungsvektoren geschätzt werden. Das Vektorfeld ist nun etwa doppelt so dicht besetzt. Das vom Horn-Schunck Verfahren für diese Testsequenz geschätzte Vektorfeld ist für beide Felder mit einem Winkelfehler von etwa zehn Prozent sehr gut (Abbildung ??). Auf beiden Bildern haben die Verfahren die Verschiebungsvektoren von den gut ausmachbaren Algenfronten in die strukturloseren Bereiche propagiert. Durch die Verwendung von Multispektralbildern konnte der Winkelfehler der Geschwindigkeitsvek-

Methode	$\phi_E$ [°]	$ \vec{v} _E$ [px <sup>2</sup> ]	$\rho$ [%]
Lucas-Kanade, GW	14.79	0.47	99.9
Lucas-Kanade, MS	7.85	0.07	100.0
Horn-Schunck, GW	9.17	0.06	100.0
Horn-Schunck, MS	7.74	0.10	100.0

Tabelle 4.1: Quantitative Ergebnisse für die synthetische Bildsequenz

Method	$\phi_E$ [°]	$ \vec{v} _E$ [px <sup>2</sup> ]	$\rho$ [%]
Lucas-Kanade, GW	17.26	1.45	20.6
Lucas-Kanade, MS	14.87	1.10	41.1
Horn-Schunck, GW	11.13	1.82	100.0
Horn-Schunck, MS	9.26	1.47	100.0

Tabelle 4.2: Quantitative Ergebnisse für das rotierende Satellitenbild

toren im Verhältnis zum Grauwertverfahren um etwa 15 Prozent gedrückt werden. Aus den Flussfeldern ist ersichtlich, dass dieser Unterschied gering ist, da die Verfahren sehr ähnliche Ergebnisse erbringen.

Im dritten Test wurde die Strömung einer Algenpopulation in einem Satellitenbild geschätzt. Die Ergebnisse entsprechen denen des vorigen Tests: Auch hier hat der multispektrale Lucas-Kanade Schätzer ein besseres Ergebnis erbracht als das grauwertbasierte Verfahren. Da keine *Ground-Truth*-Werte vorliegen, bedeutet *besser* in diesem Kontext, dass das Vektorfeld dichter besetzt ist. Die Richtung des Vektorfeldes stimmt gut mit der wahrgenommenen Bewegung überein. Das Horn-Schunck Verfahren ist hier wieder die bessere Wahl. Das geschätzte Vektorfeld auf beiden Bildern ist dicht besetzt und beschreibt sehr glatt die vermeintliche Verschiebung der Algenblüte.

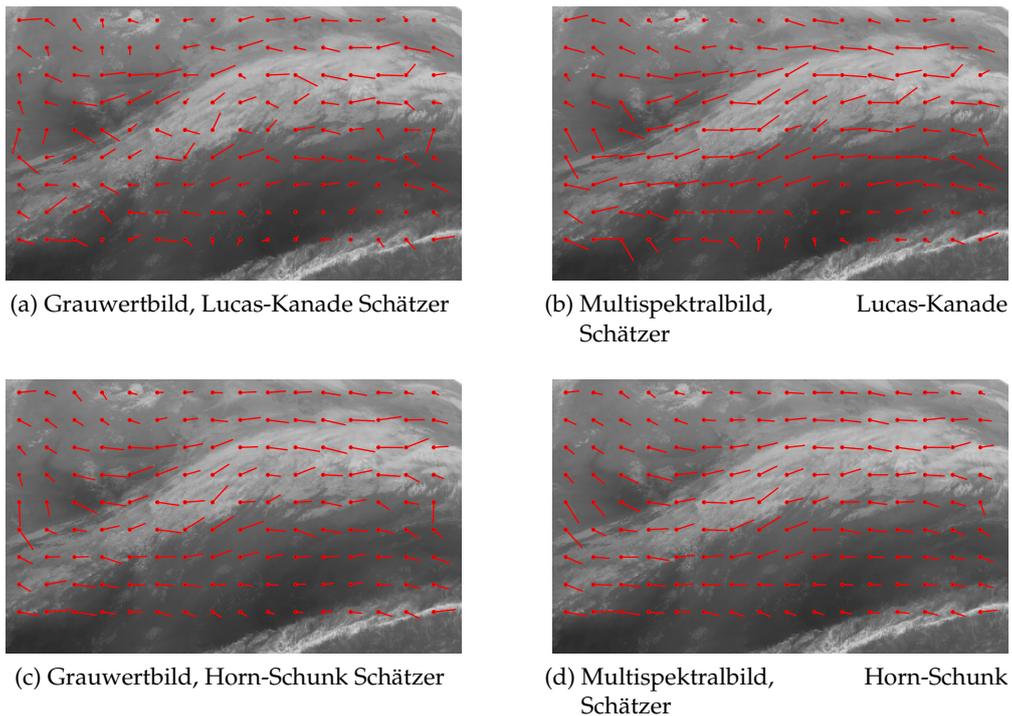


Abbildung 4.2: Ergebnisse für die Sequenz des geostationären Wettersatelliten

Das für das Satellitenbild des japanischen *MTSat*-Satelliten geschätzte Verschiebungsfeld ist in beiden Ausführungen des Lucas-Kanade Schätzers dicht besetzt. Auch ist jeweils die Bewegungsrichtung des hellen Wolkenstreifens, von links nach rechts, erkennbar. Das Graustufenverfahren weist jedoch starke Schwankungen in den bestimmten Richtungsvektoren auf, wohingegen der Schätzer auf dem Mehrkanalbild ein deutlich glatteres Feld berechnen konnte. Der Horn-Schunck Schätzer produziert ebenfalls in beiden Varianten ein gutes Feld. Auch hier ist das für die Mehrkanalaufnahme geschätzte Feld etwas glatter und zeichnet sich durch weniger Variationen in den Richtungen der Verschiebungsvektoren aus. Es fällt jedoch auf, dass die Bewegungen des schnellen Windstromes in die eigentlich recht ruhige, dunklere Region südlich des Windes propagiert wurden. Es befinden sich jedoch einige kleine Wolkenfragmente in dieser Region, die andeuten, dass dort kaum Strömung herrscht. Dies kann auch in dem vom Lucas-Kanade Schätzer geschätzten Vektorfeld nachvollzogen werden.

## 5 Fazit und Ausblick

Es wurden zwei Verfahren entwickelt, um bei der Schätzung des optischen Flusses die in Multispektralbildern vorhandenen Informationen besser nutzen zu können als die bekannten Grauwertverfahren es tun. Dabei wurde auf den bereits bestätigten Verfahren aufgebaut und die den Verfahren zugrundeliegenden Ideen auf die Verwendung mit Mehrkanalbildern erweitert.

Die Auswertung von Testdaten und der abschließende Vergleich ergaben, dass die hergeleiteten Verfahren, wenn die geforderten Bedingungen erfüllt werden, durchaus bessere Ergebnisse erbringen können.

Der aus dem Lucas-Kanade Verfahren abgeleitete Schätzer produzierte in den Vergleichen signifikant dichtere und genauere Vektorfelder als das normale Lucas-Kanade Verfahren und ist somit definitiv als eine Verbesserung anzusehen. Dies ist auch der Fall gewesen, wenn die Gradientenrichtungen in den verschiedenen Kanälen des Bildes in einem Punkt ähnlich sind. Es ist denkbar, dass das Verfahren auf Mehrkanalbildern auch gute Ergebnisse liefert, wenn die Größe der Umgebung, in der die Bewegung als konstant angenommen wird, verringert wird. Dies setzt natürlich voraus, dass die Gradientenrichtungen in den Kanälen nicht korrelieren. Das Verfahren ist im Gegensatz zum Horn-Schunck Verfahren performant, wodurch es auch für die Verwendung in Echtzeitanwendungen denkbar ist. So ist beispielsweise eine Verwendung in Computerprogrammen zur Gestenerkennung mittels einer Webcam, die Farbbilder aufzeichnet, möglich. Auch die in der Einleitung erwähnte automatische Navigation von Robotern kann durch die Verwendung von Mehrkanalbildern profitieren. So kann der Roboter neben den Informationen, die er im sichtbaren Bereich aufzeichnet, noch Daten aus dem Infrarot- oder UV-Bereich verwenden.

Bei dem vom Horn-Schunck Verfahren abgeleiteten Schätzer wurde in den betrachteten Testsequenzen jedoch nur eine geringe Verbesserung von etwa 10 % bis 20 % erzielt. In den resultierenden Vektorfeldern lässt sich über das reine Augenmaß hinaus somit kaum ein Unterschied ausmachen. Dies ist jedoch auch zum Teil auf die bei der Berechnung verwendeten Werte für die Gewichtungskonstante  $\alpha$ , sowie für den Radius der Nachbarschaft, in der  $\bar{u}$  und  $\bar{v}$  bestimmt werden, zurückzuführen. Diese Konstanten wurden in den Vergleichen für beide Bildarten gleich gewählt. Es ist jedoch denkbar, dass durch einen geringeren Wert für  $\alpha$  das Feld weniger geglättet wird

und dadurch der Vorteil bei der Verwendung von Multispektralbildern stärker hervortritt. Die Auswahl der synthetischen Testsequenzen spielt möglicherweise ebenfalls eine Rolle. Beide Sequenzen, für die eine quantitative Analyse möglich war, wurden nur durch einfache globale Transformationen wie Verschiebung und Rotation verändert. Für Vergleiche mit mehr Aussagekraft müssten Sequenzen mit verschiedenen lokalen Veränderungen erzeugt werden. Dazu könnte man nach einem bekannten Verschiebungsfeld ein Ausgabebild verzerren und diese bekannte Verschiebung danach erneut schätzen.

Die beiden in dieser Bachelorarbeit hergeleiteten Verfahren sollen als Grundlage für weitere Entwicklungen dienen. Eine Idee, die in einer anderen Arbeit weiter verfolgt werden kann, ist die unterschiedliche Gewichtung der betrachteten Kanäle bei der Schätzung der Verschiebungsvektoren. So könnten Kanäle, in denen die gesuchte Bewegung besonders deutlich wird, stärker in die Berechnung aufgenommen werden als Kanäle, in denen die Bewegung nicht zu registrieren ist. Dies könnte durch ein lokales Verfahren abhängig von der Struktur der betrachteten Region in den unterschiedlichen Kanälen geschehen. In der Fernerkundung ist auch eine Gewichtung der Kanäle anhand des Reflektionsspektrums des gesuchten Objektes oder Stoffes denkbar.

## Literaturverzeichnis

- [1] VIGRA Homepage. <http://hci.iwr.uni-heidelberg.de/vigra/>, abgerufen am 16.07.2010.
- [2] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12:43–77, 1994.
- [3] J. L. Barron and N. A. Thacker. Tutorial: Computing 2d and 3d optical flow. *Tina-Vision*, 2005.
- [4] Berthold K. P. Horn und Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [5] Miguel Tavares Coimbra. Approximating Optical Flow Within the MPEG-2 Compressed Domain. *IEEE Transactions on Circuits and Systems for Video Technology*, 15, 2005.
- [6] Gerald Fiedler. *Untersuchungen zur Bestimmung zweidimensionaler Strömungsfelder an der Meeresoberfläche mit Hilfe von multispektralen Satellitenbildern*. PhD thesis, Universität Hamburg, Dezember 2003.
- [7] Lorenz Gerstmayr. An Improvement of the Lucas-Kanade Optical-Flow Algorithm for every Circumstance, 2008. [http://www.ti.uni-bielefeld.de/downloads/lgerstmayr/lk\\_improvements.pdf](http://www.ti.uni-bielefeld.de/downloads/lgerstmayr/lk_improvements.pdf), abgerufen am 21.07.2010.
- [8] J. K. Kearney, W. B. Thompson, and D. L. Boley. Optical flow estimation: an error analysis of gradient-based methods with local optimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9(2):229–244, 1987.
- [9] Gunter Dirk Krebs. MTSat 1, 1R (Himawari 6). [http://www.skyrocket.de/space/doc\\_sdat/mtsatsat-1.htm](http://www.skyrocket.de/space/doc_sdat/mtsatsat-1.htm), abgerufen am 24.07.2010.
- [10] J. P. Lewis. Fast Normalized Cross-Correlation, 1995.
- [11] David Malin. Multispectral imaging. <http://encyclopedia.jrank.org/articles/pages/1163/Multispectral-Imaging.html>, abgerufen am 12.07.2010.

- 
- [12] Ikuya Murakami. The aperture problem in egocentric motion. *Trends in Neurosciences*, 27, 2004.
- [13] Onera - The French Aerospace Lab. GPU for Image: GPU-FOLKI. <http://www.onera.fr/dtim-en/gpu-for-image/folkigpu.php>, abgerufen am 20.07.2010.
- [14] OpenMP Architecture Review Board. The OpenMP API specification for parallel programming. <http://openmp.org/wp/>, abgerufen 22.07.2010.
- [15] Benjamin Seppke. Herleitung der Euler-Lagrange Gleichungen für Optical Flow Constraints, 2010.
- [16] Space Science and Engineering Center. Geostationary Image Browser. <http://www.ssec.wisc.edu/data/geo/index.php>.
- [17] Rainer Sprengel. *Untersuchung differentieller Ansätze zur Schätzung des optischen Flusses in Grauwert- und Farbbildfolgen*. PhD thesis, Universität Hamburg, Mai 1988.
- [18] Selim Temizer. Optical Flow Based Robot Navigation. [http://people.csail.mit.edu/lpk/mars/temizer\\_2001/Optical\\_Flow/](http://people.csail.mit.edu/lpk/mars/temizer_2001/Optical_Flow/), abgerufen am 15.07.2010.
- [19] Sarita Thakoor und John Morookian. Insect-inspired optical-flow navigation sensors. 2005. <http://www.techbriefs.com/content/view/216/32/>, abgerufen am 15.07.2010.
- [20] U.S. Geological Survey. LE71970222002196EDC00. Aufnahme des Landsat 7 ETM+ Satelliten.

## Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe angefertigt und mich anderer als der im beigefügten Verzeichnis angegebenen Hilfsmittel nicht bedient habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht.

Ich bin mit der Einstellung in den Bestand der Bibliothek des Fachbereichs einverstanden.

Oliver Bestmann

Hamburg, den 12. Februar 2012