

Variationsansätze zur Bewegungsschätzung: Präzise Modellierung und effiziente Numerik

Andrés Bruhn

Lehrstuhl für Mathematische Bildanalyse
Fakultät für Mathematik und Informatik, Geb. E1.1
Universität des Saarlandes, 66041 Saarbrücken
bruhn@mia.uni-saarland.de

Abstract: Die Schätzung von Bewegungsinformation aus Bildfolgen ist eines der zentralen Probleme des Maschinensehens. In diesem Zusammenhang spielen Variationsansätze eine große Rolle, die eine vollständige und genaue Bestimmung des Verschiebungsfelds zwischen aufeinander folgenden Bildern einer Bildfolge ermöglichen. In der hier zusammengefassten Dissertation [Bru06] werden in zweierlei Hinsicht wichtige Beiträge auf dem Gebiet der Bewegungsanalyse geleistet: (i) Zum einen wird ein systematischer Rahmen zur Entwicklung hochgenauer Variationsansätze vorgestellt. Dieses erlaubt den Entwurf der zur Zeit präzisesten Schätzverfahren in der gesamten Literatur. (ii) Zum anderen wird ein allgemeiner Ansatz zur Konstruktion hochperformanter Mehrgitteralgorithmen präsentiert. Dieser ermöglicht die Berechnung der Ergebnisse in Echtzeit und macht damit erstmals diese äußerst präzisen Variationsansätze praktisch einsetzbar. Experimente zur Evaluation der Qualität und der Effizienz belegen die Überlegenheit der entwickelten Schätzverfahren.

1 Einführung

Zahlreiche Fragestellungen des Maschinensehens erfordern die Schätzung von Bewegungsinformation aus Bildfolgen. Typische Ziele dieser Schätzung sind die Berechnung von Objektbewegungen innerhalb einer Szene oder die Bestimmung der Eigenbewegung der Kamera. Jedoch ist es ohne Vorwissen über statische Objekte nicht möglich, die absolute Bewegung von Objekten zu ermitteln: Es kann nur ihre Bewegung *relativ* zur möglicherweise ebenfalls bewegten Kamera bestimmt werden. Darüber hinaus erlaubt die Verwendung einer einzelnen Kamera nur die Berechnung der *Projektion* der relativen 3D-Bewegung auf die 2D-Bildebene: Diese Projektion wird in der Literatur als *optischer Fluss* bezeichnet. Abbildung 1 illustriert die daraus resultierende Problemstellung: Es muss aus zwei aufeinander folgenden Bildern einer Bildfolge ein Verschiebungsvektorfeld bestimmt werden, das die Pixel des ersten Bildes in ihre neue Position im zweiten Bild überführt. Aufgrund der Tatsache, dass zur Lösung dieser Aufgabe die Zuordnung einander entsprechender Pixel erforderlich ist, wird die Bestimmung des optischen Flusses häufig auch als *Korrespondenzproblem* bezeichnet.

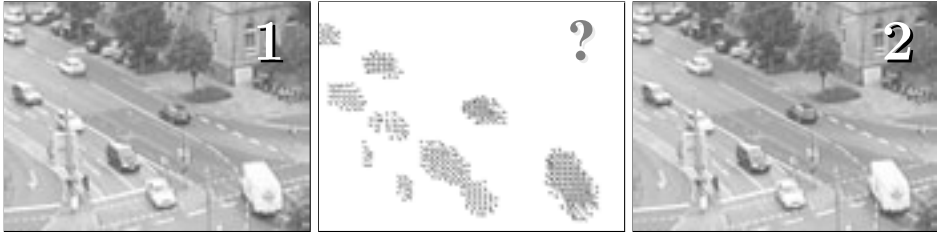


Abbildung 1: Das Optische-Fluss-Problem. *Links und rechts:* Zwei aufeinander folgende Bilder der *Karl-Wilhelm-Straßen-Sequenz* [IKS] (Ausschnitt). *Mitte:* Gesuchtes Verschiebungsvektorfeld (Optischer Fluss). Wie kann dieses Verschiebungsvektorfeld effizient und genau bestimmt werden?

Zur Lösung dieses Korrespondenzproblems wurden in der Literatur diverse Verfahren vorgeschlagen, die sich im wesentlichen in lokale und globale Methoden unterteilen lassen [BFB94, BWS05]. Während lokale Verfahren nur eine feste Pixelnachbarschaft verwenden und dadurch an Stellen ohne Texturinformation keine Schätzung ermöglichen, propagieren globale Verfahren durch geeignete Glattheitsannahmen bezüglich der Lösung implizit Informationen durch das gesamte Bild und führen in jedem Fall zu einer vollständigen (dichten) Schätzung der Ergebnisse. Besonders wichtige Vertreter der globalen Methoden – nicht zuletzt auf Grund ihrer hohen Genauigkeit bei Vergleichen in der Literatur – sind sogenannte Variationsansätze, die den optischen Fluss als Minimierer eines geeigneten Energiefunktionals berechnen [HS81, NE86, BA96, MP02, PBB⁺06]. Im Allgemeinen haben solche Energiefunktionale die Form

$$E(\mathbf{u}) = E_D(\mathbf{u}) + \alpha E_G(\mathbf{u})$$

wobei $\mathbf{u} = (u, v, 1)^\top$ das gesuchte Verschiebungsvektorfeld ist, $E_D(\mathbf{u})$ den Datenterm bezeichnet, $E_G(\mathbf{u})$ für den Glattheitsterm steht und $\alpha > 0$ einen Regularisierungsparameter darstellt, durch den sich der Glattheitsgrad der Lösung steuern lässt. Während im *Datenterm* Annahmen über die Konstanz von Objektmerkmalen in aufeinander folgenden Bildern formuliert werden, z.B. die Konstanz des Grauwerts oder der ersten räumlichen Ableitungen des Grauwerts, verkörpert der *Glattheitsterm* Annahmen über die (stückweise) Glattheit der Lösung, d.h. die Glattheit des gesuchten Verschiebungsvektorfelds. Durch Minimierung dieses Funktionals wird dann dasjenige Flussfeld als Lösung bestimmt, das optimal bezüglich allen berücksichtigten Modellannahmen ist.

Seit dem Entwurf des ersten Variationsansatzes zur Berechnung des optischen Flusses vor mehr als zwei Jahrzehnten (Horn und Schunck [HS81]) fand eine immense Entwicklung dieser Verfahrensklasse statt: Es wurden Glattheitsterme vorgestellt, die Bewegungskanten modellieren können [NE86, WS01], es wurden Datenterme entworfen, die die Berechnung robust gegenüber Rauschen und Ausreißern in den Bilddaten machen [BA96] und es wurden Minimierungsstrategien entwickelt, die selbst bei großen Verschiebungen genaue Schätzergebnisse ermöglichen [BA96, MP02]. Allerdings wurden nur selten Ansätze vorgeschlagen, die mehrerer dieser Konzepte innerhalb eines einzigen Energiefunktionals vereinigen. Beleuchtungsschwankungen während der Aufnahme fanden ebenfalls kaum

Beachtung. Darüber hinaus galten Variationsansätze als völlig ungeeignet für den Einsatz in Echtzeitapplikationen, da ihre Minimierung das Lösen großer nichtlinearer Gleichungssysteme erfordert.

Das Ziel der hier zusammengefassten Dissertation war es deshalb, sowohl die Modellierung als auch die algorithmische Umsetzung bisheriger Verfahren deutlich zu verbessern. Zu diesem Zweck wurden auf Modellierungsseite zahlreiche existierende Konzepte für Daten- und Glattheitsterme untersucht sowie neue Strategien vorgeschlagen, die gegen Beleuchtungsänderungen und Rauschen robust sind (z.B. in [BWS05, PBB⁺06]). Dabei wurde auch eine einheitliche Beschreibung der einzelnen Komponenten in Form von Bewegungstensenoren (Datenterm) und Diffusionstensenoren (Glattheitsterm) vorgestellt [BWKS06]. Auf algorithmischer Seite wurde der Einsatz sogenannter *bidirektionaler Mehrgitterverfahren* untersucht. Diese numerischen Methoden zählen zu den schnellsten Algorithmen zur Lösung linearer und nichtlinearer Gleichungssysteme überhaupt [Bra77]. Da die Komplexität beider Teilaspekte eine knappe Abhandlung im Rahmen dieser Zusammenfassung nicht zulässt, soll im folgenden die schematische Vorgehensweise anhand eines konkreten Beispiels verdeutlicht werden. Dabei handelt es sich um den Ansatz von *Bruhn und Weickert* [BW05], der zu den weltweit genauesten Verfahren zählt.

Der weitere Aufbau dieses Beitrags ist wie folgt gegliedert: In Abschnitt 2 wird auf die genaue Modellierung dieses Ansatzes eingegangen und die zu Grunde liegenden Konzepte besprochen. Abschnitt 3 widmet sich dann seiner algorithmische Umsetzung und geht dabei insbesondere auf die Optimierung und deren algorithmische Realisierung ein. Experimente zur Evaluation der Schätzgenauigkeit und Vergleiche mit der Literatur werden dann in Abschnitt 4 vorgestellt. Der Beitrag endet schließlich mit einer Zusammenfassung in Abschnitt 5.

2 Präzise Modellierung

Wir möchten zunächst die einzelnen Annahmen besprechen, die dem Verfahren von Bruhn und Weickert zu Grunde liegen. Diese Vorgehensweise soll einen weitgehenden Einblick in die Modellierung hochgenauer Variationsansätze ermöglichen. Dazu betrachten wir zwei aufeinander folgende Bilder $f(x, y, t)$ und $f(x, y, t + 1)$ einer Grauwertbildfolge f , wobei $(x, y)^\top$ den Ort innerhalb des Bildes beschreibt und t die Zeit angibt. Unter diesen Voraussetzungen lassen sich die Annahmen des Verfahrens von Bruhn und Weickert wie folgt beschreiben:

- *Konstanz des Grauwerts.* Es wird angenommen, dass zueinander gehörende Pixel in aufeinander folgenden Bildern einer Bildfolge den gleichen Grauwert besitzen, also dass $f(x, y, t + 1) = f(x, y, t)$ gilt [HS81]. Dies kann in Vektorschreibweise auch kompakter als $f(\mathbf{x} + \mathbf{u}) = f(\mathbf{x})$ formuliert werden.
- *Konstanz der räumlichen Grauwertableitungen.* Es wird angenommen, dass zueinander gehörende Pixel in aufeinander folgenden Bildern einer Bildfolge darüber hinaus auch die gleichen räumlichen Grauwertableitungen haben, also dass zusätzlich $\nabla_2 f(x + u, y + v, t + 1) = \nabla_2 f(x, y, t)$ gilt, wobei $\nabla_2 f = (\partial_x f, \partial_y f)^\top$. In Vek-

torform erhält man dann $\nabla_2 f(\mathbf{x} + \mathbf{u}) = \nabla_2 f(\mathbf{x})$. Gegenüber der Grauwertkonstanz bietet diese Gradientenkonstanz den Vorteil, dass sie deutlich robuster gegenüber Helligkeitsschwankungen ist [PBB⁺06]. Der Einsatz solcher Konstanzannahmen zum Zweck der Beleuchtungsinvarianz wurde zum ersten mal im Rahmen der hier vorgestellten Dissertation untersucht.

- *Große Verschiebungen.* Es wird außerdem angenommen, dass große Verschiebungen, d.h. Verschiebungen, die größer als ein Pixel sind, auftreten können. Unter dieser Annahme macht es Sinn, bewusst auf die übliche Linearisierung der Konstanzannahmen im Modell (siehe [HS81]) zu verzichten und eine wesentlich komplexere Optimierung in Kauf zu nehmen [PBB⁺06]. Die theoretische Herleitung einer solchen Optimierung war ebenfalls Bestandteil dieser Dissertation.
- *Statistische Robustheit im Datenterm.* Um starken Verletzungen der Konstanzannahmen durch Ausreißer weniger Bedeutung beizumessen, werden diese statistisch robust bestraft [BA96, PBB⁺06], d.h. die übliche nicht-robuste quadratische Bestrafung $\Psi_D(s^2) = s^2$ wird durch eine lineare Bestrafung mittels $\Psi_D(s^2) = \sqrt{s^2 + \epsilon^2}$ ersetzt, wobei $\epsilon = 0.001$ eine feste Regularisierungskonstante darstellt. Auch die Frage nach einer gemeinsamen oder getrennten Bestrafung im Fall mehrerer Konstanzannahmen konnte im Rahmen dieser Dissertation geklärt werden [BW05].
- *Diskontinuitätenerhaltung im Flussfeld.* In Bezug auf die Lösung wird angenommen, dass diese stückweise glatt ist. Während die eigentliche Glattheit durch Bestrafungen der ersten Flussableitungen $|\nabla_2 u|^2 + |\nabla_2 v|^2$ erreicht wird, sorgt die Verwendung einer robusten (linearen) Straffunktion $\Psi_D(s^2) = \sqrt{s^2 + \epsilon^2}$ mit $\epsilon = 0.001$ für die gewünschten Erhaltung von Bewegungskanten im zu berechnenden Flussfeld. Dieser Typ des Glattheitsterms wird als flussgetrieben isotrop bezeichnet [WS01].

Unter der Berücksichtigung all dieser Annahmen lässt sich schließlich das Energiefunktional für das Verfahren von Bruhn und Weickert formulieren. Es lautet

$$E(\mathbf{u}) = E_D(\mathbf{u}) + \alpha E_S(\mathbf{u}),$$

wobei der Datenterm durch

$$E_D(\mathbf{u}) = \int_{\Omega} \left(\underbrace{\psi_D(|f(\mathbf{x} + \mathbf{u}) - f(\mathbf{x})|^2)}_{\text{Grauwertkonstanz}} + \gamma \underbrace{\psi_D(|\nabla_2 f(\mathbf{x} + \mathbf{u}) - \nabla_2 f(\mathbf{x})|^2)}_{\text{Gradientenkonstanz}} \right) dx dy$$

und der Glattheitsterm durch

$$E_S(\mathbf{u}) = \int_{\Omega} \psi_S(\underbrace{|\nabla_2 u|^2 + |\nabla_2 v|^2}_{\text{Glattheitsannahme}}) dx dy.$$

gegeben ist. Hierbei ist $\alpha > 0$ der schon zuvor beschriebene Regularisierungsparameter, der die Glattheit des Flussfelds steuert, und $\gamma > 0$ bezeichnet ein Gewicht, das den Einfluss der Gradientenkonstanz bestimmt.

3 Effiziente Numerik

Um das eigentliche Verschiebungsvektorfeld zu bestimmen, muss das zuvor modellierte Energiefunktional minimiert werden: Es wird also diejenige Lösung gesucht, die bezüglich aller getroffenen Annahmen optimal ist. Bei dieser Minimierung spielen sogenannte *Coarse-to-Fine-Warping-Strategien* [BA96, MP02, PBB⁺06] eine wichtige Rolle, da sie die Schätzung großer Verschiebungen erlauben. Damit sind solche Strategien ein geeignetes Mittel, um den bewussten Verzicht auf eine Linearisierung der Konstanzannahmen im Datenterm algorithmisch umzusetzen.

Um große Verschiebungen schätzen zu können, machen Coarse-to-Fine-Warping-Strategien von zwei Konzepten Gebrauch, die beide eng miteinander verzahnt sind: Die *sukzessive Verfeinerung* der Problemstellung (Coarse-to-Fine) und die *fortlaufende Kompensation* des aktuellen Bildpaares um bereits berechnete Verschiebungen (Warping). Algorithmisch lassen sich solche Coarse-to-Fine-Warping-Strategien wie folgt formulieren:

- 1) Zunächst werden beide Bilder des aktuellen Bildpaares durch Glättung und Unterabtastung auf eine sehr grobe Auflösungsstufe gebracht.
- 2) Dann wird das Verschiebungsvektorfeld auf dieser groben Auflösung berechnet.
- 3) Dieses Vektorfeld wird auf der nächst feineren Auflösungsstufe benötigt: Es wird dort aus dem zweiten Bild der Bildfolge *herausgerechnet*, d.h. die Problemstellung auf der feineren Auflösungsstufe wird um das bereits berechnete Flussfeld korrigiert. Dieser Schritt wird auch als *Warping* bezeichnet.
- 4) Das so modifizierte Problem (Differenzproblem) wird nun auf der feineren Auflösungsstufe durch Berechnung eines weiteren Verschiebungsvektorfelds gelöst.
- 5) Die Schritte 3-4 werden wiederholt, bis die ursprüngliche Auflösung erreicht wird.
- 6) Das letztendliche Ergebnis wird durch Addition der Verschiebungsvektorfelder aller Auflösungsstufen gebildet.

Diese *inkrementelle Berechnung* des Verschiebungsvektorfelds bietet folgenden Vorteil: Während das Coarse-to-Fine-Konzept sicherstellt, dass die Verschiebungen auf der größten Auflösungsstufe sehr klein sind, sorgt die Warping-Strategie dafür, dass dies auch für die zu berechnenden Verschiebungskorrekturen (Verschiebungsvektorfelder der Differenzprobleme) so bleibt. Da kleine Verschiebungen viel genauer berechnet werden können als größere, nimmt durch den Einsatz einer solchen Coarse-to-Fine-Warping-Strategie die Schätzqualität im Allgemeinen deutlich zu. Jedoch muss anstelle eines einzelnen Korrespondenzproblems eine *Hierarchie* von ebensolchen Problemen gelöst werden. Im Rahmen dieser Dissertation wurde eine solche Vorgehensweise zum ersten Mal direkt aus einem zu Grunde liegenden Funktional hergeleitet und damit theoretisch fundiert [PBB⁺06].

Des Weiteren ist die Minimierung von Funktionalen aus mathematischer Sicht sehr eng mit der Minimierung von Funktionen verwandt: So wie eine Nullstelle der ersten Ableitung eine notwendige Bedingung für ein Minimum einer Funktion darstellt, ist die Erfüllung der sogenannten *Euler-Lagrange-Gleichungen* eine notwendige Bedingung für die minimierende Funktion eines Funktionals (die minimierende Funktion entspricht hier dem

gesuchten Verschiebungsvektorfeld). Durch eine geeignete Diskretisierung dieser Euler-Lagrange-Gleichungen mit Hilfe finiter Differenzen entsteht ein *großes dünnbesetztes nichtlineares Gleichungssystem*, das es zu lösen gilt. Unter Verwendung von Coarse-to-Fine-Warping-Strategien – so wie in unserem Fall – muss ein solches Gleichungssystem sogar für jede einzelne Auflösungsstufe, d.h. für *jedes Warpinglevel* errechnet werden.

In der Literatur werden zum Lösen dieser Gleichungssysteme im Allgemeinen einfache Algorithmen wie das explizite Verfahren (Gradientenabstieg) oder ein nichtlineares Gauß-Seidel-Verfahren eingesetzt. Diese sind relativ einfach zu implementieren, benötigen jedoch Tausende von Iterationen, um die gewünschte Genauigkeit zu erreichen. Deshalb galten Variationsansätze als völlig ungeeignet für den Einsatz in Echtzeitapplikationen. Im Rahmen dieser Dissertation wurden hingegen die Klasse der *bidirektionalen Mehrgitterverfahren* untersucht, die für bestimmte Probleme lineare Komplexität besitzen, und deshalb aus numerischer Sicht zu den schnellsten Verfahren zum Lösen linearer und nichtlinearer Gleichungssysteme überhaupt gehören [Bra77]. Im Gegensatz zu herkömmlichen nicht-hierarchischen iterativen Verfahren, wie z.B. den verschiedenen linearen und nichtlinearen Gauß-Seidel-Varianten, bieten solche Mehrgitterstrategien den Vorteil, Korrekturen der Lösung effizient auf größeren Auflösungsstufen berechnen zu können. Dies führt wiederum zu einer erheblich schnelleren Konvergenz. Im Fall des Verfahrens von Bruhn und Weickert wurde darüber hinaus eine Coarse-to-Fine Initialisierung der Lösung vorgeschlagen. Das so erhaltene *volle Mehrgitterverfahren* ist in Abbildung 2 schematisch skizziert, die den Wechsel zwischen den verschiedenen Auflösungsstufen illustriert.

Allerdings erfordert der Entwurf eines solchen Mehrgitteralgorithmus ein problemspezifisches und aufwändiges Abstimmen der einzelnen Verfahrenskomponenten aufeinander. Dabei spielen Fragen nach einer geeigneten *Grobitterrepräsentation* des Problems ebenso eine zentrale Rolle wie die Wahl des zu Grunde liegende nicht-hierarchische Verfahrens sowie der Transferoperatoren zwischen den Auflösungsstufen. Stellt sich eine der Komponenten als ungeeignet heraus, kann das Mehrgitterverfahren beliebig langsam werden oder sogar divergieren. Auf alle diese Fragen wurde in dieser Dissertation ausführlich eingegangen. Dabei wurden sowohl lineare als auch nichtlineare Verfahren abgedeckt [BWKS06].

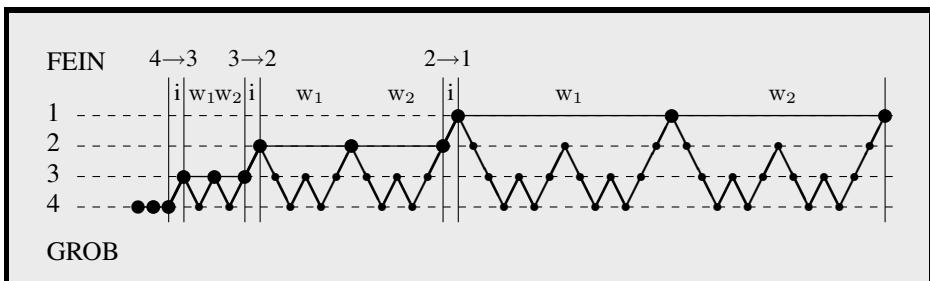


Abbildung 2: Beispiel für ein volles Mehrgitterverfahren mit zwei W-Korrekturzyklen pro Auflösungsstufe der hierarchischen Initialisierung. Die Coarse-to-Fine-Interpolationsschritte der Lösung von einer Stufe zur nächsten sind mit 'i' bezeichnet, die beiden W-Korrekturzyklen mit ' w_1 ' und ' w_2 '. Während Iterationen auf dem Originalproblem mit großen Markierungen symbolisiert werden, stehen kleine Markierungen für Iterationen auf Fehlerkorrekturproblemen.

Tabelle 1: Vergleich mit den genauesten Verfahren in der Literatur bezüglich der Qualität der Schätzergebnisse für die *Yosemite-Sequenz mit Wolken*. MWF = mittlerer Winkelfehler. Die mit * markierten Verfahren wurden ebenfalls im Rahmen dieser Dissertation (mit-)entwickelt.

Verfahren	MWF
Singh (1990)	10.44°
Nagel (1986)	10.22°
Horn/Schunck, mod. (1981)	9.78°
Uras <i>et al.</i> (1988)	8.94°
Liu <i>et al.</i> (1998)	6.85°
Alvarez <i>et al.</i> (2000)	5.53°
Farnebäck (2001)	4.84°
Mémin/Pérez (1998)	4.69°

Verfahren	MWF
Bruhn <i>et al.</i> * (2005)	4.17°
Wu <i>et al.</i> (1998)	3.54°
Teng <i>et al.</i> (2005)	2.70°
Unser Verfahren (2-D)	2.42°
Bab Hadiashar/Suter (1998)	2.05°
Papenberg <i>et al.</i> * (2006)	1.78°
Amiaz/Kiryati (2005)	1.73°
Unser Verfahren (3-D)	1.72°

4 Experimente

Nachdem in den beiden vorangegangenen Abschnitten die wesentlichen Merkmale der Modellierung und Optimierung besprochen worden sind, soll in diesem Abschnitt sowohl die Schätzgenauigkeit als auch die algorithmische Effizienz des vorgestellten Verfahrens verdeutlicht werden. Zu diesem Zweck wurden zwei verschiedene Experimente durchgeführt.

In einem ersten Experiment wollen wir zunächst die Genauigkeit der berechneten Flussfelder evaluieren. Dazu haben wir das Verschiebungsvektorfeld zwischen Bild 8 und Bild 9 der *Yosemite Sequenz mit Wolken* berechnet (siehe Abbildung 3). Diese synthetische Sequenz zeigt einen Flug durch den Yosemite-Nationalpark kombiniert das durch die Eigenbewegung der Kamera entstehende divergente Bewegungsfeld mit einer translatorischen Bewegung der Wolken. Zudem liegt im Himmel durch die partielle Verdeckung der Sonne eine geringe Helligkeitsänderung vor. Auf Grund dieser hohen Komplexität und der Tatsache, dass ein korrektes Verschiebungsvektorfeld verfügbar ist, gehört diese Sequenz zu den beliebtesten Testszenarien zur Evaluation der Schätzqualität.

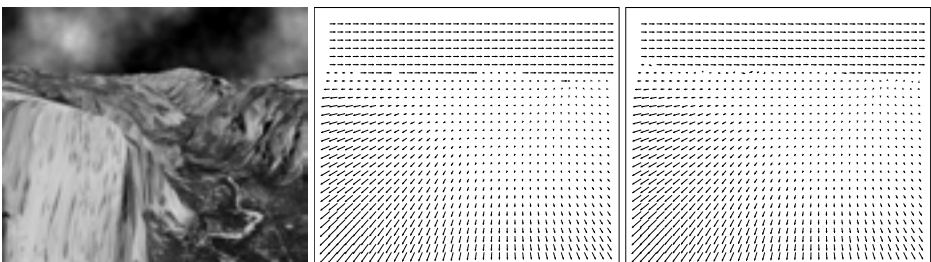


Abbildung 3: Vergleich des geschätzten und des korrekten Verschiebungsvektorfelds für die *Yosemite-Sequenz mit Wolken* von Quam. Links: Bild 8 der Sequenz. Mitte: Korrektes Flussfeld. Rechts: Berechnetes Flussfeld.

Tabelle 2: Effizienzvergleich zwischen verschiedenen numerischen Lösungsstrategien für das vorgestellte Verfahren. Die Laufzeiten beziehen sich auf die Berechnung *eines* Flussfelds der *Rheinhafen-Sequenz* der Größe 160×120 auf einer Pentium 4 CPU mit 3.06 GHz. BPS= Bilder pro Sekunde.

Löser für jedes Warpinglevel	Iterationen	Laufzeit [s]	BPS [s^{-1}]	Speedup
Explizites Verfahren	174217	1906.546	0.001	1
Gauß-Seidel	9798	74.816	0.013	25
Coarse-to-Fine-Gauß-Seidel	1646	16.190	0.062	118
Semiimplizites Verfahren	25×5	0.664	1.506	2868
Linearisiert + SOR Verfahren	25×5	0.617	1.621	3089
Volles Mehrgitterverfahren	1	0.349	2.861	5454

Wie auf dem Gebiet der Bewegungsschätzung üblich [BFB94], haben wir zur Beurteilung der Genauigkeit den mittleren räumlich-zeitlichen Winkelfehler (MWF) zwischen unserer Schätzung und dem korrekten Flussfeld ermittelt. Durch die zeitliche Komponente dieses Fehlermaßes wird dabei im Übrigen auch die Länge der geschätzten Vektoren berücksichtigt. Wie aus Tabelle 1 hervorgeht sind die erzielten Winkelfehler nicht nur extrem niedrig, sondern zum Zeitpunkt der Dissertation *die kleinsten in der gesamten Literatur überhaupt*. Dabei ist anzumerken, dass die 3-D Variante des vorgestellten Verfahrens alle Flussfelder der Sequenz gleichzeitig berechnet, pro Bild jedoch ähnlich effizient arbeitet wie die bereits im Detail erläuterte 2-D Methode. Das Verfahren von Amiaz und Kiryati, das ebenfalls sehr gute Ergebnisse liefert, basiert direkt auf der Methode von Papenberg *et al.* [PBB⁺06], die ebenfalls während dieser Dissertation (mit-)entwickelt wurde. Das dem besten Ergebnis von 1.72° entsprechende Verschiebungsvektorfeld ist in Abbildung 3 zu sehen, in der es mit dem korrekten Flussfeld verglichen wird. Hier wird der sehr kleine Winkelfehler visuell bestätigt: Trotz der Helligkeitsänderung im Himmel ist die Schätzung im gesamten Bild sehr genau. Auch der Horizont, der die beiden verschiedenen Bewegungsmuster – Divergenz und Translation – voneinander trennt, wurde akkurat geschätzt.

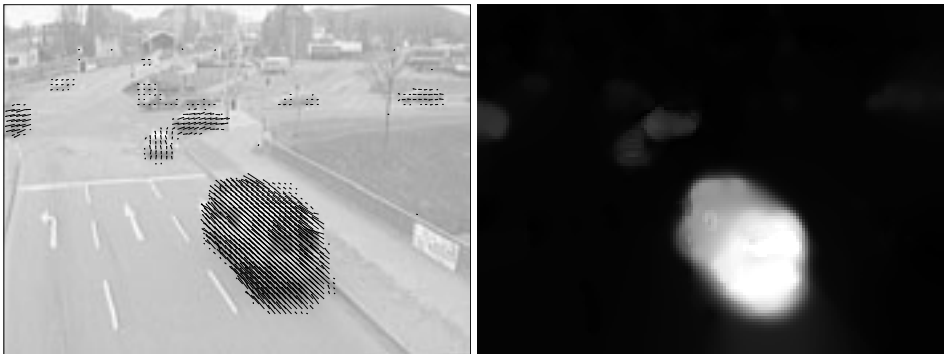


Abbildung 4: Ergebnis der Echtzeitberechnung für die *Rheinhafen-Sequenz* [IKS]. Links: Bild 1130 mit überlagertem Verschiebungsvektorfeld. Rechts: Betrag des Verschiebungsvektorfelds.

In einem zweiten Experiment wollen wir schließlich die Effizienz der Berechnung selbst überprüfen. Hierzu haben wir das implementierte volle Mehrgitterverfahren mit einigen anderen numerischen Methoden verglichen, darunter auch das häufig verwendete explizite Verfahren und eine nichtlineare Gauß-Seidel-Variante. Die gemessenen Laufzeiten, die in Tabelle 2 angegeben sind, beziehen sich auf ein Flussfeld der Größe 160×120 und eine relative Genauigkeit von einem Prozent (d.h. der Fehler beträgt weniger als ein Prozent der Lösung). Beim Vergleich der Zeiten wird deutlich, dass das entwickelte volle Mehrgitterverfahren nicht nur bis zu vier Größenordnungen schneller ist als herkömmliche in diesem Zusammenhang verwendete Algorithmen, sondern auch die Berechnung von drei dichten Flussfeldern mit je 20.000 Verschiebungsvektoren (40.000 Unbekannte) pro Sekunde erlaubt. Dies bedeutet, dass zum ersten Mal in der Literatur eine Berechnung von optischem Fluss in solch *hoher Qualität in Echtzeit* möglich ist. Im Rahmen dieser Dissertation wurden auch andere, etwas weniger genauere Variationsansätze implementiert. Diese erreichten sogar bis zu 60 Flussfeldern dieser Größe pro Sekunde [BWKS06]. Das entsprechende Flussfeld, das zur einer unterabgetasteten Version der Rheinhafen-Sequenz gehört, ist in Abbildung 4 zu finden. Wie zu sehen ist, wird die Bewegung aller Fahrzeuge korrekt wiedergegeben. Selbst die durch die geringere Entfernung zur Kamera größere Bewegung des Hecks des weißen Lieferwagens im Vordergrund wird genau erfasst. Dies zeigt, dass sich bei Verwendung geeigneter Algorithmen eine genaue Schätzung eine effiziente Berechnung nicht widersprechen müssen.

5 Zusammenfassung und Diskussion

Durch die einzelnen Beiträge der hier zusammengefassten Dissertation wurde das Gebiet der Schätzung von Bewegungsinformation aus Bildsequenzen in mehrfacher Hinsicht vorgebracht. Es wurden sowohl die weltweit genauesten Verfahren entwickelt als auch echtzeitfähige Algorithmen entworfen, die eine praxisrelevante Anwendung der Verfahren zulassen. Insbesondere im Bereich Fahrerassistenzsysteme und der mit der Bewegungsanalyse eng verwandten Stereorekonstruktion besteht bereits industrielles Interesse. Dies zeigt, dass eine konsequente Entwicklung von der Modellierung bis hin zur tatsächlichen algorithmischen Umsetzung unter Berücksichtigung der jeweils besten Konzepte zu Ergebnissen führen kann, die bisherigen Verfahren sowohl in der Qualität als auch in der Effizienz überlegen sind.

Literatur

- [BA96] M. J. Black und P. Anandan. The robust estimation of multiple motions: parametric and piecewise smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, Januar 1996.
- [BFB94] J. L. Barron, D. J. Fleet und S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, Februar 1994.
- [Bra77] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation*, 31(138):333–390, April 1977.

- [Bru06] A. Bruhn. *Variational Optic Flow Computation – Accurate Modelling and Efficient Numerics*. Dissertation, Fakultät für Mathematik und Informatik, Universität des Saarlandes, Deutschland, Juli 2006.
- [BW05] A. Bruhn und J. Weickert. Towards ultimate motion estimation: Combining highest accuracy with real-time performance. In *Proc. Tenth International Conference on Computer Vision*, Seiten 749–755, Beijing, China, Juni 2005. IEEE Computer Society Press.
- [BWS06] A. Bruhn, J. Weickert, T. Kohlberger und C. Schnörr. A multigrid platform for real-time motion computation with discontinuity-preserving variational methods. *International Journal of Computer Vision*, 70(3):257–277, Dezember 2006.
- [BWS05] A. Bruhn, J. Weickert und C. Schnörr. Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231, 2005.
- [HS81] B. Horn und B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [IKS] Universität Karlsruhe Institut für Algorithmen und Kognitive Systeme. Digitalisierte Video-Bildfolgen. http://i21www.ira.uka.de/image_sequences/.
- [MP02] E. Mémin und P. Pérez. Hierarchical estimation and segmentation of dense motion fields. *International Journal of Computer Vision*, 46(2):129–155, 2002.
- [NE86] H.-H. Nagel und W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:565–593, 1986.
- [PBB⁺06] N. Papenberg, A. Bruhn, T. Brox, S. Didas und J. Weickert. Highly accurate optic flow computation with theoretically justified warping. *International Journal of Computer Vision*, 67(2):141–158, April 2006.
- [WS01] J. Weickert und C. Schnörr. A theoretical framework for convex regularizers in PDE-based computation of image motion. *International Journal of Computer Vision*, 45(3):245–264, Dezember 2001.



Andrés Bruhn erhielt 2001 sein Diplom in Technischer Informatik an der Universität Mannheim und 2006 seinen Dokortitel in Informatik an der Universität des Saarlandes. Seit seiner Promotion arbeitet er als Postdoktorand in der Arbeitsgruppe Mathematische Bildanalyse an der Universität des Saarlandes. Zu seinen Forschungsinteressen zählen die Bewegungsanalyse in Bildfolgen, die 3D-Rekonstruktion aus Stereobildpaaren sowie schnelle numerische Algorithmen im Bereich Bildverarbeitung und Maschinensehen. Dr. Bruhn ist in den Programmkomitees der wichtigsten internationalen Konferenzen im Bereich Com-

puter Vision (ICCV, ECCV und CVPR) tätig und begutachtet regelmäßig Beiträge für die bedeutendsten Fachzeitschriften. Seine Arbeiten wurden bereits mehrfach mit Preisen ausgezeichnet, unter anderem mit dem Longuet-Higgins Best Paper Award (ECCV), dem höchsten europäischen Computer-Vision-Preis, und dem Olympus-Preis, der höchsten Auszeichnung der Deutschen Arbeitsgemeinschaft für Mustererkennung (DAGM).