

Punktverfolgung in Farbbildsequenzen
Benno Heigl & Dietrich Paulus

Inhaltsverzeichnis

1	Einleitung	1
2	Verfolgung grauer Punkte	2
3	Verfolgung farbiger Punkte	3
4	Verfolgung in bekannter Richtung	4
5	Anwendung	5
6	Zusammenfassung und Ausblick	6
A	Farbbilder	8

Datei /home/paulus/text/papers/COLOR-97/PaHe/RCS/PaHe-1-97.tex,v
Version 1.24 vom 1997/09/12 12:18:48
L^AT_EX 7. Oktober 1997

Punktverfolgung in Farbbildsequenzen

Benno Heigl & Dietrich Paulus

Lehrstuhl für Mustererkennung, Universität Erlangen–Nürnberg,
heigl@informatik.uni-erlangen.de,
paulus@informatik.uni-erlangen.de

Übersicht

Zur Verfolgung von Punkten in Sequenzen von Grauwertbildern geben Tomasi und Kanade ein Verfahren an, das den Grad der Interessantheit eines Punkts davon abhängig macht, wie gut sich der Punkt in hochfrequenten Bildfolgen verfolgen läßt; die Forderung dabei ist, daß sich ein Bildpunkt zwischen zwei aufeinanderfolgenden Bildern nur wenig verschiebt.

In unserem Beitrag wird dieses Prinzip ausgedehnt auf Ströme von Farbbildern. Zunächst wird das Verfahren auf vektorwertige Meßdaten erweitert und in einer experimentellen Auswertung auf Farbbilder im RGB-Farbraum angewandt. Die Ergebnisse werden mit denen des ursprünglichen Verfahrens verglichen.

Ziel ist es, möglichst viele Punkte über längere Abschnitte verfolgen zu können, um aus den Parametern der Trajektorien Tiefe zu ermitteln. Hierzu werden Ergebnisse vorgestellt.

1 Einleitung

Um das Problem der Punktverfolgung in einer Bildfolge zu lösen, müssen die beiden folgenden Fragen beantwortet werden [8]:

- Welche Merkmale werden in den Einzelbildern gesucht?
- Wie werden diese Merkmale in den einzelnen Bildern der Sequenz weiterverfolgt?

Für die Auswahl von Merkmalen sind in der Literatur zahlreiche Verfahren vorgeschlagen worden. Die bekanntesten sind beispielsweise der “interest operator” in [2] und der Eckendetektor aus [5]. Eine Zusammenfassung bietet dazu auch [7].

Ebenso existieren zur Punktverfolgung verschiedene Verfahren, die meist die Korrelation zwischen lokalen Fenstern um die zu verfolgenden Punkte verwenden. Eine Übersicht hierzu findet man wiederum in [7].

¹Diese Arbeit wurde unterstützt durch die DFG im Rahmen des Sonderforschungsbereichs 182

In [8] werden die beiden getrennten Fragestellungen zu *einer* verschmolzen. Diese Ergebnisse gründen sich auf einer Serie von Veröffentlichungen dieser Autoren aus dem Bereich des Stereosehens, z.B. auf [1].

Die Anwendungsgebiete für die Punktverfolgung sind sehr verschiedenartig. In [9] werden beispielsweise die Trajektorien dazu verwendet, um die Kamera- und Objektbewegung zu ermitteln. Im folgenden wird die Punktverfolgung zur Tiefenermittlung in einer statischen Szene herangezogen, wobei die Kamerabewegung und die internen Kameraparameter bekannt sind.

2 Verfolgung grauer Punkte

Die zentrale Idee in [8] ist es, diejenigen Punkte zur Verfolgung auszuwählen, die sich gut verfolgen lassen. Damit werden die beiden Fragen, die in Abschnitt 1 unabhängig voneinander gestellt wurden, zu einem Problem zusammengefaßt.

Für die Formalisierung des Problems wird von einer zeitveränderlichen Bildfunktion $f(\mathbf{x}, t)$ des Ortes $\mathbf{x} = (x, y)^T$ und der Zeit t ausgegangen. Beobachtet man einen Objektpunkt zum Zeitpunkt t an der Position \mathbf{x} und zum Zeitpunkt $t + \tau$ an der Position $\mathbf{x} + \mathbf{d}$, so setzt man voraus, daß das Residuum $\epsilon = (f(\mathbf{x} - \mathbf{d}, t) - f(\mathbf{x}, t + \tau))^2$ minimal, im rauschfreien Fall sogar null ist. Eine ortsabhängige Gewichtungsfunktion $\omega(\mathbf{x})$ gestattet es, die Werte eines lokalen Fensters \mathcal{W} zu gewichten. Im einfachsten Fall definiert man $\omega(\mathbf{x}) = 1$. Damit wird ϵ definiert als

$$\epsilon = \int_{\mathcal{W}} (f(\mathbf{x} - \mathbf{d}, t) - f(\mathbf{x}, t + \tau))^2 \omega(\mathbf{x}) d\mathcal{W}. \quad (1)$$

Bezeichnet man mit $\mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x})$ den Gradienten der Grauwertfunktion an der Stelle \mathbf{x} , läßt sich für kleine Verschiebungen folgende Taylorentwicklung durchführen:

$$f(\mathbf{x} - \mathbf{d}, t) \approx f(\mathbf{x}, t) - \mathbf{g}(\mathbf{x}) \cdot \mathbf{d} \quad (2)$$

Unter Verwendung dieser Näherung kann durch Ableiten und Nullsetzen der Wert ϵ aus Gleichung (1) bezüglich der Verschiebung \mathbf{d} minimiert werden, wodurch sich folgendes Gleichungssystem ergibt:

$$\mathbf{G}\mathbf{d} = \mathbf{e} = \int_{\mathcal{W}} (f(\mathbf{x}, t) - f(\mathbf{x}, t + \tau)) \mathbf{g}(\mathbf{x}) \omega(\mathbf{x}) d\mathcal{W}$$

mit symmetrischer 2×2 -Matrix \mathbf{G}

$$\mathbf{G} = \int_{\mathcal{W}} \mathbf{g}(\mathbf{x}) \mathbf{g}^T(\mathbf{x}) \omega(\mathbf{x}) d\mathcal{W} = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix} \quad (3)$$

Dieses lineare Gleichungssystem läßt sich eindeutig lösen. Der hierdurch ermittelte Verschiebungsvektor \mathbf{d} ist eine Näherung der tatsächlichen Verschiebung. Durch Neuabtastung des Musters $\mathbf{f}(\mathbf{x} + \mathbf{d}, t)$ im beobachteten Fenster \mathcal{W} mittels bipolarer Interpolation kann

durch iteriertes Anwenden des obigen Verfahrens die Gesamtverschiebung in Subpixelgenauigkeit berechnet werden.

Die Eigenwerte λ_1 und λ_2 der Matrix \mathbf{G} berechnen sich folgendermaßen:

$$\lambda_{1,2} = \frac{1}{2} \left(g_{11} + g_{22} \pm \sqrt{(g_{11} - g_{22})^2 + 4g_{12}^2} \right) \quad (4)$$

Als Maß für die Verfolgbarkeit von Punkten dient in [8] der Wert $\min(\lambda_1, \lambda_2)$, welcher beurteilt, ob die Texturierung des Objektes an der betrachteten Stelle genügend stark ist, um über dem Signalrauschen zu liegen und ob sie andererseits dazu geeignet ist, das betrachtete Fenster in jede beliebige Richtung zu verfolgen. Der hierdurch definierte Operator sei als *Verfolgbarkeitsoperator* bezeichnet.

3 Verfolgung farbiger Punkte

Eine Erweiterung der Herangehensweise, die in Abschnitt 2 dargestellt wurde, bietet sich an für multispektrale Muster \mathbf{f} , im speziellen für Farbbilder, welche durch vektorwertige Bildfunktionen in einem vorgegebenen Farbraum repräsentiert sind. Beispielsweise für den *RGB*-Farbraum sei das Muster eine Funktion $\mathbf{f}(\mathbf{x}) = (f_r(\mathbf{x}), f_g(\mathbf{x}), f_b(\mathbf{x}))^T$. Die Ableitung der Gleichungen für die Verfolgung und für die Beurteilung der Verfolgbarkeit von Punkten geschieht analog dem Grauwertfall.

Das nach Verschiebung des Fensters \mathcal{W} auftretende Residuum wird entsprechend Gleichung (1) für Farbvektoren definiert:

$$\epsilon = \int_{\mathcal{W}} \|\mathbf{f}(\mathbf{x} - \mathbf{d}, t) - \mathbf{f}(\mathbf{x}, t + \tau)\|^2 \omega(\mathbf{x}) d\mathcal{W} \quad (5)$$

Die Taylorapproximation gestaltet sich ebenfalls ähnlich der Gleichung (2), nur daß an Stelle des Gradienten $\mathbf{g}(\mathbf{x})$ die Funktionalmatrix $\mathbf{J}(\mathbf{x})$ tritt:

$$\mathbf{f}(\mathbf{x} - \mathbf{d}, t) \approx \mathbf{f}(\mathbf{x}, t) - \mathbf{J}(\mathbf{x})\mathbf{d} \quad (6)$$

Setzt man Gleichung (6) in (5) ein und definiert $\mathbf{h}(\mathbf{x}) := \mathbf{f}(\mathbf{x}, t) - \mathbf{f}(\mathbf{x}, t + \tau)$, so erhält man:

$$\epsilon = \int_{\mathcal{W}} \|\mathbf{h}(\mathbf{x}) - \mathbf{J}(\mathbf{x})\mathbf{d}\|^2 \omega(\mathbf{x}) d\mathcal{W} \quad (7)$$

Die optimale Verschiebung \mathbf{d} , bei welcher der Fehler ϵ minimiert wird, läßt sich durch Nullsetzen des Gradienten für ϵ bezüglich der Verschiebung \mathbf{d} ermitteln:

$$\nabla \epsilon(\mathbf{d}) = \int_{\mathcal{W}} \left(2\mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x})\mathbf{d} - 2\mathbf{J}^T(\mathbf{x})\mathbf{h}(\mathbf{x}) \right) \omega(\mathbf{x}) d\mathcal{W} = \mathbf{0} \quad (8)$$

$$\underbrace{\left(\int_{\mathcal{W}} \mathbf{J}^T(\mathbf{x})\mathbf{J}(\mathbf{x})\omega(\mathbf{x}) d\mathcal{W} \right)}_{\mathbf{G}} \mathbf{d} = \underbrace{\int_{\mathcal{W}} \mathbf{J}^T(\mathbf{x})\mathbf{h}(\mathbf{x})\omega(\mathbf{x}) d\mathcal{W}}_{\mathbf{e}} \quad (9)$$



Bild 1: Vergleich der Ergebnisse des *Verfolgbarkeitsoperators* auf einem Grauwertbild (links) und einem Farbbild (rechts)

Dieses Gleichungssystem läßt sich entsprechend dem Grauwertfall eindeutig lösen und führt durch iterierte Anwendung zur subpixelgenauen Bestimmung von \mathbf{d} . Auch das Maß für die Verfolgbarkeit ist entsprechend dem Grauwertfall durch den kleinsten Eigenwert der Matrix \mathbf{G} bestimmt. Bild 1 zeigt Ergebnisse des Verfolgbarkeitsoperators jeweils für die Grauwert- und Farbdefinition. In dieser Szene lassen sich nur sehr geringe Unterschiede zwischen den beiden Ergebnisbildern erkennen, da hier die meisten Farbkanten auch im entsprechenden Grauwertbild deutlich hervortreten.

4 Verfolgung in bekannter Richtung

Will man Punkte einer starren Szene bei bewegter Kamera verfolgen, so lassen sich aufgrund der Epipolargeometrie [9] einschränkende Annahmen bezüglich des nachfolgenden Aufenthaltsortes machen. Ist beispielsweise die Lageveränderung rein translatorisch, so liegt für einen Bildpunkt der korrespondierende auf der Geraden, welche durch den Epipol und diesen Bildpunkt definiert ist. In diesem Fall ist eine Nutzung dieses Wissens über den zukünftigen Aufenthaltsort naheliegend, womit sich eine zusätzliche Erweiterung des Ansatzes aus [8] ergibt:

Die bekannte Verschiebungsrichtung eines Punktes \mathbf{x} sei mit \mathbf{r} bezeichnet, und es gelte o. B. d. A.: $\|\mathbf{r}\| = 1$. Ziel ist nun, die optimale Verschiebung $\mathbf{d} = u \mathbf{r}$ durch Bestimmung des Skalars u zu berechnen. Die Vorgehensweise ist ähnlich zu der bei beliebiger Verschiebungsrichtung, nur daß an Stelle des Gradienten die Richtungsableitung der Funktion \mathbf{f} in Richtung \mathbf{r} verwendet wird.

Unter Verwendung der Taylorapproximation

$$\mathbf{f}(\mathbf{x} - u \mathbf{r}) \approx \mathbf{f}(\mathbf{x}) - u \mathbf{J}(\mathbf{x}) \mathbf{r} \quad (10)$$

und der Fehlerabschätzung

$$\epsilon = \int_{\mathcal{W}} \|\mathbf{h}(\mathbf{x}) - u \mathbf{J}(\mathbf{x}) \mathbf{r}\|^2 \omega(\mathbf{x}) d\mathcal{W} \quad (11)$$

erhält man nach Differenzierung nach u und Nullsetzen

$$u \int_{\mathcal{W}} \|\mathbf{J}(\mathbf{x}) \mathbf{r}\|^2 \omega(\mathbf{x}) d\mathcal{W} = \int_{\mathcal{W}} \mathbf{h}^T(\mathbf{x}) \mathbf{J}^T(\mathbf{x}) \mathbf{r} \omega(\mathbf{x}) d\mathcal{W}. \quad (12)$$

Dasselbe Verfahren ist auch für Grauwertbilder anwendbar und führt zu folgender Gleichung:

$$u \int_{\mathcal{W}} (\mathbf{g}^T(\mathbf{x}) \mathbf{r})^2 \omega(\mathbf{x}) d\mathcal{W} = \int_{\mathcal{W}} h(\mathbf{x}) \mathbf{g}^T(\mathbf{x}) \mathbf{r} \omega(\mathbf{x}) d\mathcal{W} \quad (13)$$

In beiden Fällen kann als Maß für die Verfolgbarkeit eines Punktes der Wert des linken Integrals verwendet werden, das anschaulich der Inhomogenität in Richtung \mathbf{r} entspricht.

5 Anwendung

Zur Tiefenermittlung wurde eine Kamera auf einem Linearschlitten bewegt, wobei die horizontale Achse der Bildebene parallel zur Bewegungsrichtung war. Damit ergibt sich ein Aufbau, der vergleichbar zu einer Stereoanordnung mit parallelen optischen Achsen ist [4].

Das Korrespondenzproblem ist im Gegensatz zum Stereosehen jedoch nicht explizit zu lösen; stattdessen ergibt sich die Punktkorrespondenz aus der Verfolgung während der Bewegung. Wie auch im Fall der Stereoanordnung befinden sich korrespondierende Punkte vertikal immer auf gleicher Höhe, wodurch der Suchraum erheblich eingeschränkt werden kann. Deshalb wurde jeweils zusätzlich der Fall bekannter Richtung (Abschnitt 4) untersucht.

Für jeden verfolgten Punkt bleibt demnach die y -Koordinate über die Zeit hinweg konstant, die x -Koordinate ändert sich nach folgender Vorschrift:

$$x = F \frac{-vt + x_w}{z_w}, \quad (14)$$

wobei v die Geschwindigkeit des Linearschlittens bezeichnet, F die Brennweite der Kamera und x_w bzw. z_w die Koordinaten des beobachteten Punktes bezüglich des Kamerakoordinatensystems zum Zeitpunkt $t = 0$. Der Wert x ist somit linear von t abhängig. Die Lageparameter lassen sich für gegebene Trajektorien durch lineare Regression berechnen [6], wodurch zusätzlich ein Maß für die Unsicherheit der Verfolgung abfällt. Die Größe des Festers \mathcal{W} war 4×4 ; die Fensterfunktion $\omega(\mathbf{x})$ wurde als konstante Funktion gewählt.

Zum Test der Verfahren wurde eine Bildfolge verwendet, die einen Rechnerraum zeigt. Sie besteht aus 32 Einzelbildern, wobei die Pixelverschiebung für korrespondierende Punkte maximal 55 Pixel zwischen dem ersten und letzten Bild der Sequenz beträgt.

Bild 2 auf S. 8 zeigt das Ergebnis der Verfolgung von 200 Punkten mittels der in Abschnitt 2 und Abschnitt 3 gezeigten Verfahren; in beiden Fällen sind die Tiefenwerte im rechten Bild farbig codiert dargestellt. Für diese Szene sind nur geringe Unterschiede zwischen den Ergebnissen der beiden Verfahren festzustellen. Die Verarbeitung der Bildfolge benötigt auf einer SGI O2 73 Sekunden, für die entsprechende Grauwertfolge 39 Sekunden.

Bild 3 auf S. 8 zeigt das Ergebnis der Verfolgung bei horizontaler Bewegung unter Verwendung des Farbverfahrens aus Abschnitt 4. Man erkennt, daß wesentlich mehr Punkte (2283) zur Verfolgung ausgewählt werden können als bei beliebiger Bewegungsrichtung, wodurch das Tiefenbild an senkrechten Bildkanten sehr dicht ist. Die Berechnung für obige Bildfolge benötigt unter denselben Voraussetzungen 167 Sekunden.

Die Programme wurden in das Bildverarbeitungs-System ANIMALS [3] integriert.

6 Zusammenfassung und Ausblick

In diesem Artikel wurde gezeigt, daß sich das Verfahren [8] zur Punktverfolgung in Grauwertbildern auch auf Farbbilder erweitern läßt. Die Anwendung dieser Verfolgung bei der Tiefenermittlung zeigt qualitativ, daß die in diesem Artikel entwickelte Erweiterung zumindest für die hier beobachtete Szene ähnliche Ergebnisse wie das bisherige Verfahren erzeugt. Die neu eingeführte, zusätzliche Verwendung von Wissen über die Bewegungsrichtung ermöglicht bei der Tiefenermittlung gegenüber dem allgemeinen Fall dichtere Tiefenbilder.

Mittels einer quantitativen Auswertung wird zu beurteilen sein, ob die Verwendung von Farbbildern für die hier gezeigten Verfahren tatsächlich Vorteile gegenüber der von Grauwertbildern aufweist. Zusätzlich sind andere Szenarien und Farbräume zu untersuchen. Außerdem ist die Verwendung einer Auflösungshierarchie sinnvoll, um auch bei größeren Pixelverschiebungen eine Verfolgung zu ermöglichen.

Die aus der Tiefenermittlung erhaltenen Punkte sollen künftig zur Klassifikation von Objekten eingesetzt werden.

Literatur

- [1] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.
- [2] H. P. Moravec. Robot spatial perception by stereoscopic vision and 3d evidence grids. Technical Report CMU-RI-TR-96-34, The Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, September 1996.
- [3] D. Paulus and J. Hornegger. *Pattern Recognition of Images and Speech in C++*. Advanced Studies in Computer Science. Vieweg, Braunschweig, 1997.

- [4] Stefan Posch. *Automatische Tiefenbestimmung aus Grauwert-Stereobildern*. Dissertation, Deutscher Universitäts Verlag, Wiesbaden, 1990.
- [5] A. Rosenfeld and M. Thurston. Edge and curve detection for visual scene analysis. *IEEE Transactions on Computers*, 20:562–569, 1971.
- [6] G. Schmidt. Tiefe aus linearer Kamerabewegung . Technical report, Studienarbeit, Lehrstuhl für Mustererkennung (Informatik 5), Universität Erlangen–Nürnberg, Erlangen, 1995.
- [7] J. Shi and C. Tomasi. Good features to track. In *Proceedings of Computer Vision and Pattern Recognition*, pages 593–600, Seattle, Washington, Juni 1994. IEEE Computer Society Press.
- [8] C. Tomasi and T. Kanade. Detection and tracking of point features. Technical Report CMU-CS-91-132, Carnegie Mellon University, 1991.
- [9] G. Xu and Z. Zhang. *Epipolar Geometry in Stereo, Motion and Object Recognition — A Unified Approach*, volume 6 of *Computational Imaging and Vision*. Kluwer Academic Press, Dordrecht, 1996.

A Farbbilder

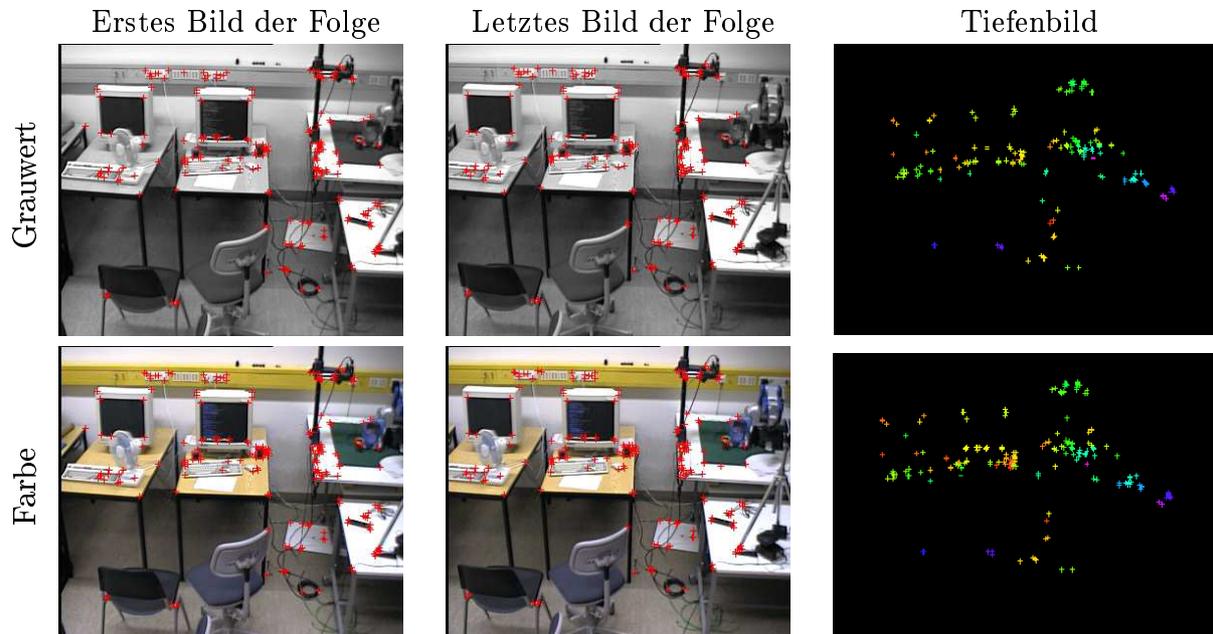


Bild 2: Vergleich der Punktverfolgung bei unbekannter Bewegungsrichtung; die Tiefenwerte sind im Tiefenbild durch Farben codiert: Punkte mit großer z -Koordinate sind durch rote Farbe, Punkte mit kleiner z -Koordinate durch blaue Farbe gekennzeichnet.



Bild 3: Punktverfolgung in Farbbildern bei bekannter Bewegungsrichtung; die Farbcodierung der Tiefenwerte entspricht der aus Bild 2.