

Alles im Fluss

Optischer Fluss für industrielle Anwendungen

Was ist optischer Fluss und wofür kann man ihn verwenden? Einfach zusammengefasst, schätzt ein Algorithmus zur Bestimmung des optischen Flusses an jedem Pixel eines Bildes seine Verschiebung bezüglich eines Referenzbildes. Klingt einfach? Auf den ersten Blick schon. Bevor wir nun tiefer in diese Definition einsteigen und all die Feinheiten betrachten, die dieses seit 30 Jahren studierte Problem schwer machen, möchte ich Ihnen einige Anwendungen vorstellen. Sie werden sich vielleicht wundern, dass ich Ihnen keine Beispiele aus der Machine Vision nenne. Das liegt daran, dass optischer Fluss in diesem Gebiet gerade erst bekannt wird. Deshalb konzentriere ich mich mit meinen Beispielen auf andere kommerzielle Bereiche und hoffe Sie damit zu inspirieren, dieses Verfahren kreativ auf Ihre eigenen Anwendungen zu übertragen.



Die bekannte „Trinity Szene“ im Film Matrix wurde zunächst mit mehreren Kameras aufgenommen um anschließend neue Ansichten durch Flussfelder zu interpolieren.

Was können wir also mit Flussfeldern anfangen? Bill Silver von Cognex glaubt, dass optischer Fluss eine wichtige Komponente für Subpixelgenauigkeit darstellt: niedrigauflösende Kameras könnten so verwendet werden, um gleichzeitig virtuell die Auflösung zu erhöhen und die Hardwarekosten zu senken. In anderen großen Firmen wie der Robert Bosch AG und der Daimler AG werden Fahrerassistenzsysteme entwickelt, die entweder auf Stereo-Verfahren, optischem Fluss oder beidem basieren. So können Abstände zu Hindernissen geschätzt, sich bewegende Objekte erkannt, Kollisionen vorhergesagt und viele andere Aussagen über die Verkehrssituation getroffen werden. Ein anderer Wirtschaftszweig ist die medizinische Bildverarbeitung: hier wird der optische Fluss „Registrierung“ genannt. Firmen wie Siemens und General Electric konzentrieren einen großen Teil ihrer Forschung darauf, z.B. Zeitreihen von Magnetresonanz-Bildern auf Tumorzunahme zu untersuchen oder Röntgenvideos zu tauschen.

Unterhaltung ist ein anderes Gebiet, das von Firmen wie Sony, Microsoft und Canon betrachtet wird. Eingabegeräte, die menschliche Bewegungen in Kommandos für Spielekonsolen umsetzen, sind momentan ein heißes Thema. Etwas unbekannter ist die Tatsache, dass Fernseher anhand des optischen Flusses normale Bildraten von 25 Hz auf 100 Hz hochinterpolieren. Firmen wie Industrial Light and Magic benutzen ähnliche Methoden für Spezialeffekte in Kinofilmen: So wurde die bekannte „Trinity Szene“ im Film Matrix (1999) zunächst mit mehreren Kameras aufgenommen, um anschließend neue Ansichten durch Flussfelder zu interpolieren. Während der Postproduktion von 3D-Filmen wird

anhand des Flusses die Entfernung jedes Pixels zur Kamera berechnet. Dadurch kann nachträglich virtuell der Abstand der Stereokameras variiert werden (Nuke, The Foundry Visionmongers Ltd.).

Da Flussfelder an Tiefenkanten häufig scharfe Kanten aufweisen, werden diese zur Segmentierung verwendet. Videokompressionsstandards wie MPEG speichern so pro Objekt nur dessen Bewegung. Findet in der aufgenommenen Szene keine Bewegung statt oder wird nur ein Objekt aufgenommen, um das sich die Kamera dreht, entsteht ein 3D-Scanner. Die Ergebnisse sind meist relativ ungenau; trotzdem kann der Nutzen groß sein in B2C-Szenarien, wo Hardware nur wenig kosten darf. Zu guter Letzt kennen Sie sicherlich Google Street View, wo Bewegungen und Stereodaten zusammengeführt werden, um ganze Städte in 3D zu rekonstruieren.

Sehr häufig wird optischer Fluss auch als Messtechnik verwendet. So analysiert man z.B. Blutpumpen, indem man kleine Partikel in eine Flüssigkeit gibt und deren Bewegung mit einer Hochgeschwindigkeitskamera aufnimmt. Die gemessenen Bewegungen geben Aufschluss über Strömungen, die die Qualität des Blutes beeinträchtigen. Auch Luftströmungen werden so gemessen, um die Aerodynamik von Autos oder Flugzeugen zu optimieren. Ein weiteres spannendes Anwendungsfeld ist die Analyse von Flussfeldern in Überwachungsszenarien. Verdächtiges Verhalten kann hier z.B. an der Gangart eines Menschen erkannt werden.

Herausforderungen der Flusserschätzung

Es existiert eine Vielzahl weiterer Anwendungen in Wissenschaft und Wirtschaft. Auf den ersten Blick ist das Ver-

fahren sehr mächtig. Der Teufel liegt aber bekanntlich im Detail. Wenn Sie basierend auf dem optischen Fluss eigene Anwendungen entwickeln wollen, sollten Sie sich der Beschränkungen dieses Verfahrens bewusst sein. Was sind also die wichtigsten Fragen, die Sie sich stellen müssen, um zu entscheiden, ob und wie Sie einen Flussschätzer implementieren sollen?

Die erste Frage ist jene nach dem Helligkeitsmodell. Eine typische Annahme ist, dass die Intensität eines Pixels sich entlang seiner Bewegungstrajektorie nicht ändert. Sollten sich Objekte einer Sequenz typischerweise einer Lichtquelle nähern oder von ihr entfernen, wäre es angebracht, eine quadratische Helligkeitsänderung anzunehmen. Ändert sich die Beleuchtung völlig unkontrolliert, wird man sich meist gegen Flussschätzung entscheiden.

Die zweite Frage stellt sich bezüglich des sog. Regularisierers oder Bewegungsmodells. Angenommen, die Beleuchtung ist konstant und wir betrachten einen einzelnen Pixel im ersten Bild. Wir werden viele verrauschte Pixel mit der gleichen Intensität im zweiten Bild finden. Auf den ersten Blick sind alle möglichen Verschiebungsvektoren dieses Pixels gleich wahrscheinlich. Ohne Vorwissen kann der optische Fluss deshalb nicht berechnet werden. Eine Lösung dieses Problems ergibt sich, wenn man Vorwissen annimmt, das als Randbedingung während der Flussschätzung genutzt wird. Eine einfache Variante kann z.B. die Annahme sein, dass der Fluss von einem Pixel zu seinen Nachbarn nur geringfügig variiert. Dadurch können Abweichungen mit einem Energieterm bestraft werden: das Flussfeld wird regularisiert. Je mehr Sie bereits über die möglichen Bewegungen in Ihrer Anwendung wissen, desto genauer und schneller können Sie Flussfelder berechnen.

Die dritte Frage bezieht sich auf die Optimierung: Sind die Entscheidungen für Helligkeits- und Bewegungsmodell gefällt, können Sie eine Energie ableiten, die eine Aussage über die Qualität eines Flussfeldes bei gegebenem Bildpaar erlaubt. Eine Optimierungsmethode soll Ihnen nun mitteilen, wie Sie den Fluss verändern müssen, damit er besser zum Modell passt. Diese Wahl eines Optimierers ist entscheidend für die Frage, ob Ihr Flussschätzer schnell genug für Ihre

Anwendung sein wird. Abhängig von der benötigten Genauigkeit, der Komplexität der formulierten Energie und natürlich der Hardware kann die Berechnung eines Flussfeldes der Größe 640 x 480 mal wenige Millisekunden und mal einige Minuten dauern.

Ein anderer Faktor ist der Bildinhalt. Optischer Fluss kann nur verlässlich geschätzt werden, wenn die sich bewegenden Strukturen entweder hinreichend groß oder klein sind. Als Daumenregel sollte bei vielen Verfahren der größte Flussvektor einen Pixel lang sein. Auf Kosten der Rechenzeit können auch Flüsse bis ca. fünf Pixel Länge geschätzt werden. Andererseits können auf Kosten der Genauigkeit mit anderen Verfahren



Bill Silver von Cognex glaubt, dass optischer Fluss eine wichtige Komponente für Subpixelgenauigkeit darstellt.

auch sehr viel größere Flüsse bestimmt werden. Sollten sowohl sehr kleine als auch sehr große Verschiebungen auftauchen, versagen viele Verfahren.

Sehr oft werden Bildverarbeitungs-algorithmen auf Daten angewendet, die nicht sorgfältig in Hinsicht auf ihre automatische Analyse aufgenommen wurden. So können viele Probleme des optischen Flusses durch einfache Modifikationen des Aufnahmeverfahrens vermieden werden. Zum Beispiel beeinflussen die Texturen im Bild stark das Ergebnis. Unter idealen Umständen sind alle Objekte im Bild lambert'sche Strahler mit fixem Abstand bezüglich der Lichtquelle. Schatten erzeugen Helligkeitsänderungen, die oft vom Helligkeitsmodell nicht richtig behandelt werden können. Deshalb sollten Sie falls möglich Farbspritzer auf ansonsten einfarbige Objekte anbringen, oder Verunreinigungen erst später entfernen. Haben Sie Einfluss auf die Lichtquelle, deren Form, Spektrum und Einfallswinkel, sollten Sie das unbedingt ausnutzen.

Verdeckungen sind ebenfalls schwer zu behandeln. Die größte Herausforderung liegt hier in der häufigen Annahme, dass Pixel in einer lokalen Nachbarschaft alle den gleichen Fluss aufweisen. Können Sie also die Bewegungen in Ihrer Anwendung auch nur ansatzweise kontrollieren, kann dies in dramatisch reduzierten Rechenzeiten oder erhöhter Zuverlässigkeit resultieren.

Ein letzter Punkt ist die Frage, wie Sie Ihre Ergebnisse validieren. In den vier wichtigsten wissenschaftlichen Magazinen (IJCV, PAMI, CVIU, IP) wurden seit 1981 mehr als 1.500 Artikel über den optischen Fluss veröffentlicht. Erstaunlicherweise untersuchen aber nur vier davon die Qualität einiger weniger Verfahren. Es bleibt offen, ob die vorgeschlagenen Benchmarks relevant für Ihre Anwendung sind. Hinzu kommt, dass es nur sehr wenig frei verfügbaren Code gibt. Da ein sehr guter Master-Student etwa ein halbes Jahr zur Programmierung relativ fortgeschrittener Flussschätzer benötigt, ist es sehr schwer zu entscheiden, welche Methode für Ihre Anwendung am geeignetsten ist.

Fazit

Sie müssen fünf Fragen beantworten, um ein Verfahren zur Schätzung des optischen Flusses zu entwickeln: Zunächst sollten Sie ein Modell der Helligkeitsänderung eines Pixels entlang seiner Trajektorie entwickeln. Zweitens benötigen Sie einen Regularisierer, der Vorwissen ausnutzt. Der dritte Schritt besteht darin, ein Optimierungsverfahren zu verwenden, das schnell genug brauchbare Ergebnisse liefert. Sie sollten sicherstellen, dass bereits bei der Aufnahme der Daten bestmögliche Bedingungen herrschen. Da verlässliche Benchmarks nicht existieren, wäre es ratsam, selbst Ihre spezifische Anwendung zu validieren.

Ich glaube, dass heute bereits viele Flussschätzungsprobleme schnell und zuverlässig gelöst werden können. Andererseits ist der optische Fluss keine eierlegende Wollmilchsau und wird wahrscheinlich in mindestens genauso vielen Szenarien versagen. Die Wahl der algorithmischen und technischen Methoden sollte deshalb in jedem Fall sorgfältig getroffen werden; Black-Box- oder General-Purpose-Schätzer führen nur sehr selten zum Ziel. (pe)

► **Autor**
Dr. Daniel Kondermann, Postdoc
unter Prof. Bernd Jähne
am Heidelberg Collaboratory
for Image Processing



► **Kontakt**
Heidelberg Collaboratory for Image Processing
Universität Heidelberg
Tel.: 06221/54-8875
Fax: 06221/54-8846
daniel.kondermann@iwr.uni-heidelberg.de
www.iwr.uni-heidelberg.de

