Effiziente Bewegungserfassung des Augenhintergrundes

Raimund Lakmann

Universitäts-Augenklinik, Abt. II
D–72076 Tübingen
Email: raimund.lakmann@uni-koblenz.de

Zusammenfassung. Es wird ein Verfahren zur automatischen Detektion von Augenbewegungen vorgestellt, mit dem Untersuchungen zur Erfassung der dynamischen Blicksteuerung effizient ausgewertet werden können. Die automatische Bewegungsanalyse ist prinzipiell anwendungsunabhängig konzipiert und modular in verschiedenen augenmedizinischen Experimenten einsetzbar. Das Verfahren wird auf Bildsequenzen des Augenhintergrundes angewendet, die mit einem Scanning Laser Ophthalmoskop (SLO) aufgezeichnet werden. Die automatische Auswertung der Bewegungen erfolgt offline durch Methoden zur automatischen Bildauswertung, die auf einem modifizierten Block-Matching-Algorithmus basieren. Eine nachgeschaltete interaktive Komponente zur effizienten manuellen Fehlerkontrolle trägt den hohen Ansprüchen eines klinischen Einsatzes des Verfahrens Rechnung.

Schlüsselwörter: Bewegungsanalyse, Scanning Laser Ophthalmoskop, SLO, Augenbewegung

1 Einleitung

Eine exakte Erfassung der Augenbewegungen ist unabdingbare Voraussetzung für verschiedenartige medizinische Auswertungen von Aufnahmen des Augenhintergrundes. Einerseits sind dabei die Bewegungen der Augen selbst der Gegenstand der Untersuchungen. So können die Augenbewegungen eines Patienten Störungen bei der stabilen Fixierung (Nystagmus) anzeigen und Aufschluss über die individuelle Blicksteuerung bei der visuellen Erfassung von Objekten und beim Lesen von Textzeilen geben. Andererseits besteht durch eine Bewegungserfassung auch die Möglichkeit zur Bewegungskompensation. Damit können die störenden Augenbewegungen von Patienten ausgeglichen werden, die schon bei kurzen Aufnahmezeiten unvermeidlich sind, z.B. bei perimetrischen und angiographischen Untersuchungen der Retina. Die Erfassung der Augenbewegungen anhand von Fundus-Aufnahmen gehört zu den Bildverarbeitungsdisziplinen Bildsequenzanalyse und Bildregistrierung. Bei den in der Literatur vorgestellten Verfahren zur Bewegungserfassung des Augenhintergrundes handelt es sich zum einen um adaptive Verfahren, die im Hinblick auf eine spezielle Anwendung optimiert sind, und zum anderen um schnelle Methoden für Echtzeitanwendungen [1, 2, 3, 4]. Das im folgenden beschriebene Verfahren ist prinzipiell für den

Einsatz in verschiedenen Anwendungen konzipiert. Zur Erfassung der Augenbewegung wird ein Scanning Laser Ophthalmoskop (SLO) vorausgesetzt, mit dem der Augenhintergrund als Videosequenz aufgezeichnet wird. Diese SLO-Videosequenzen werden durch ein Verfahren zur automatischen Bildauswertung nach Abschluss der Untersuchung offline ausgewertet und interpretiert. Damit unterliegt das System grundsätzlich keinen Echtzeitanforderungen, der Hauptaspekt liegt auf der Qualität der Bewegungsdetektion und auf einer einfachen Einsetzbarkeit des Verfahrens im klinischen Alltag.

2 Methode

Bei den mit dem SLO gewonnenen Aufnahmen handelt es sich um Bildsequenzen beider Augen, in denen jeweils die Fovea und die Gefäße der Netzhaut sichtbar sind. Mit dem SLO können durch optische Einblendungen verschiedene Muster präsentiert werden, die der Proband fixieren soll. Die Art dieser Muster reicht in den untersuchten Bildsequenzen von einfachen optischen Markierungen (Fixationskreuz) bis hin zu Textfragmenten. Die eingeblendeten Muster werden durch die zugrundeliegende Aufnahmetechnik gleichzeitig mit dem Augenhintergrund aufgezeichnet, so dass eine sichere Zuordnung der relativen Positionen der Markierungen zur Fovea, dem Zentrum des scharfen Sehens, möglich ist [5]. Die Analyse der Bewegungen kann prinzipiell durch einen menschlichen Betrachter erfolgen, der die Videosequenz bildweise anschaut und die relativen Bewegungen der Netzhaut manuell erfasst. Aufgrund der Länge der Untersuchungen ist die manuelle Auswertung der Bildsequenzen jedoch mit einem erheblichen Arbeitsaufwand und einem erhöhten Fehlerrisiko verbunden. Bei der Verfolgung der Bewegungen des Augenhintergrundes kommen mangels signifikanter Objektmerkmale bevorzugt Regionenmerkmale infrage [4]. Regionenmerkmale sind daher Hauptbestandteil der eingesetzten Methode zur automatischen Bildauswertung (s. Abs. 3). In einer initialen Phase wählt der Benutzer des Programms interaktiv ein typisches Bild der Sequenz als Referenzbild aus, in dem kleine Bildausschnitte markiert werden, die in der gesamten Bildsequenz verfolgt

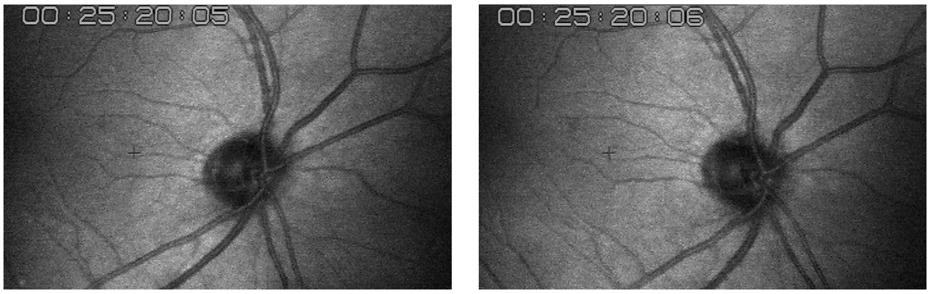

Abb. 1. Beispiele von zwei aufeinanderfolgenden Fundus-Bildern einer SLO-Videosequenz mit einer deutlichen Bildverzerrung.

werden sollen. Diese Referenzregionen werden so definiert, dass mindestens eine dieser Regionen in allen Bildern der Sequenz sichtbar ist, um eine kontinuierliche Bewegungsverfolgung zu ermöglichen. Der Algorithmus zur Bewegungsdetektion ermittelt in der gesamten Bildsequenz die Bildregionen, die den Referenzregionen entsprechen.

3 Bewegungsdetektion

Zur automatischen Bewegungsdetektion wird ein Korrelationsverfahren verwendet, das es erlaubt, eine Bildregion in der Bildsequenz zu verfolgen. Es handelt sich um einen modifizierten Block-Matching-Algorithmus, der die Robustheit und Einfachheit dieses statistischen Verfahrens mit Komponenten zur Erhöhung der Qualität und Schnelligkeit verbindet. Als Gütekriterium für die Ähnlichkeit von Bildregionen dient die normierte Kovarianz der Grauwerte der beteiligten Bildpunkte. Prinzipiell ermittelt der Matching-Algorithmus dieses Ähnlichkeitskriterium für alle infrage kommenden Bildpositionen der Referenzregionen. Die Bildregion mit der maximalen Kovarianz wird als korrespondierende Bildregion akzeptiert, wenn ein spezifischer Schwellwert erreicht ist. Ein Vorteil dieses Verfahrens liegt in seiner inhärenten Unempfindlichkeit gegen Helligkeitsunterschiede zwischen den zu vergleichenden Bildregionen, die auf Helligkeitsschwankungen in der SLO-Videosequenz zurückzuführen sind. Ein weiterer Vorteil basiert auf der Robustheit der Methode gegen Grauwertstörungen einzelner Bildpunkte und gegen die eingeblendeten Muster.

Der bedeutende Nachteil des Verfahrens ist vor allem sein hoher Rechenaufwand, der sich jedoch durch Modifikationen reduzieren lässt. Zur Beschleunigung des Programms trägt u.a. die Definition von Suchbereichen bei, die einerseits statisch vom Benutzer je nach Grad der zu erwartenden Bewegungen festgelegt werden können und andererseits dynamisch aus der aktuellen Bewegung der Bilder abgeleitet werden. Die Suche nach den korrespondierenden Bildregionen beginnt jeweils in Abhängigkeit der Bewegung und der vorangehenden Bildposition an einer geschätzten Position im Bild. Diese Suche wird solange in konzentrischen Kreisen innerhalb des Suchbereiches fortgesetzt, bis ein signifikantes Maximum des Ähnlichkeitskriteriums erreicht ist. Eine weitere Beschleunigung konnte durch eine hierarchische Korrelationsanalyse erreicht werden, bei der die Kovarianzberechnungen nicht für alle Bildpunkte der Bildregionen, sondern zunächst nur für gröbere Auflösungsstufen durchgeführt werden. Ein weiterer Nachteil des Verfahrens betrifft Bildverzerrungen, die in SLO-Videosequenzen auftreten. Dabei handelt es sich um horizontale Verzerrungen (s. Abb. 1), die aufgrund schneller Augenbewegungen bei konstanter Videoabtastrate der Bildzeilen verursacht werden. Um auch bei den auf diese Weise verzerrten Bildern ein Regionen-Matching zu ermöglichen, werden die Verzerrungen bei der Berechnung des Ähnlichkeitskriteriums durch ein vereinfachtes lineares Bewegungsmodell kompensierend berücksichtigt.

In kritischen Aufnahmesituationen, die beispielsweise durch das Schließen der Augen oder durch kurze Unterbrechungen der Untersuchung bedingt sind, können Fehler bei der automatischen Bewegungsanalyse auftreten. Durch die Wahl geeigneter Schwellwerte für das Ähnlichkeitskriterium fällt die Zahl der Fehler jedoch weitgehend gering aus. Eine statische Definition der Schwellwerte ist aufgrund starker Unterschiede in der Bildqualität der SLO-Videosequenzen nicht ratsam. Daher können die Schwellwerte je nach Aufnahmequalität adaptiv auf optimale Werte eingestellt werden. Außerdem gewährleistet eine zusätzliche manuelle Fehlerkontrolle die fehlerfreie Bewegungserfassung der Bildsequenzen. Nach Abschluss der automatischen Bewegungsdetektion kann der Benutzer die Szenen mit den niedrigsten Korrelationswerten visualisieren und das Ergebnis gegebenenfalls manuell korrigieren. Die Kontrolle der Bilder wird dabei in der Reihenfolge mit zunehmender Korrelationsgüte vorgenommen, so dass auf diese Weise nur ein Bruchteil der Bildsequenz gesichtet werden muss, um die Qualität der automatischen Bewegungsdetektion beurteilen zu können.

4 Ergebnisse

Die Ergebnisse des Verfahrens zur Bewegungsdetektion wurden durch Vergleiche mit den Ergebnissen einer rein manuellen Bewegungserfassung überprüft. Zur qualitativen Bewertung wurden mehrere Testsequenzen mit dem System ausgewertet. In den Sequenzen, in denen aufgrund der Bildqualität auch eine manuelle Auswertung möglich ist, werden bei der automatischen Bewegungsdetektion hohe Erfolgsquoten erzielt. Die Bewertung des Systems ist prinzipiell schwierig, weil die Qualität der automatischen Bewegungsdetektion signifikant von der Wahl der Schwellwerte und von der Definition des Erfolgsfalls abhängt. Ein objektiver Bewertungsmaßstab muss für das gesamte System gelten, d.h. für die automatische Bewegungsdetektion einschließlich der Komponente zur manuellen Qualitätskontrolle. Aufgrund dieses zweistufigen Vorgehens sind größere Abweichungen bei der Bewegungsdetektion sowie Fehler 1. Art (falsch negativ) und 2. Art (falsch positiv) ausgeschlossen. Die Beurteilung des Verfahrens konzentriert sich daher auf die Güte der Lokalisierung der verfolgten Bildregionen. Dazu wurden die Abstände zwischen den Bildpositionen für die Bilder der Testsequenzen gemessen, die einerseits durch das System zur Bewegungsdetektion und andererseits rein manuell erfasst wurden. Als Abstandsmaße wurden dabei die horizontalen und vertikalen Distanzen (in Pixel) sowie die zweidimensionale Euklidische Distanz verwendet. In Tabelle 1 sind die Abstände der Bildpositionen einer verfolgten Bildregion für eine Testsequenz exemplarisch

Abstandsmaß	$0 \le d < 2$	$2 \le d < 4$	$4 \le d < 6$	$6 \le d < 8$	$8 \le d < 10$	$d \ge 10$
horizontal	90.8	9.0	0.2	–	–	–
vertikal	65.3	34.2	0.5	–	–	–
euklidisch	58.8	40.5	0.7	–	–	–

Tabelle 1. Statistische Verteilung (in %) der Abweichungen d der ermittelten Bildpositionen (in Pixel) einer verfolgten Bildregion in einer Testsequenz mit 1050 Halbbildern.

dargestellt. Da eine rein manuelle Bewegungsdetektion nicht eindeutig erfolgen kann, wurden die Testsequenzen mehrfach manuell erfasst und für jedes Bild der Median der erfassten Bildpositionen als Referenz festgelegt. Auf der Basis dieses Referenzmaßes konnten nun die Abweichungen der Ergebnisse des Systems genau ermittelt werden. Dabei zeigte sich, dass die Abweichungen in mehr als 50 % der Bilder weniger als zwei Pixel betragen, bei mehr als 99 % der Bilder liegen die Abweichungen bei weniger als 4 Pixel. Abweichungen von 4 oder mehr Pixel traten bei den untersuchten Sequenzen selten auf.

5 Realisierung

Die Methode zur Bewegungsdetektion wurde als unabhängiges Modul für den Einsatz in verschiedenen Softwareumgebungen implementiert. Die Dateneingabe besteht aus digitalen Bildsequenzen, die in verschiedenen Formaten vorliegen können (z.B. BMP, PGM, AVI). Dazu werden die SLO-Videosequenzen zunächst digitalisiert, wobei entsprechend der Videonorm (interlaced mode) Sequenzen mit 50 Halbbildern pro Sekunde entstehen. Die automatische Bewegungsdetektion wird offline ohne ständige Kontrolle des Benutzers für mehrere Bildsequenzen durchgeführt. Die effiziente manuelle Fehlerkontrolle kann anschließend aufgrund der protokollierten Ergebnisse der Bewegungsdetektion zu einem späteren Zeitpunkt erfolgen. Die Ausgabe des Programms besteht aus Listen von Bewegungsvektoren, in denen die horizontalen und vertikalen Bewegungen von Referenzregionen protokolliert werden, aus denen die translatorischen Bewegungen und Rotationen der Bilder abgeleitet werden. Die Ergebnisse können graphisch mit Hilfe einer ergonomischen Benutzerschnittstelle dargestellt werden und stehen direkt für weitergehende automatische Diagnosen zur Verfügung.

Literatur

1. S.F. Barrett, M.R. Jerath, H.G. Rylander, A.J. Welch. Digital tracking and control of retinal images. *Optical Engineering*, 33(1):150-159, 1994.
2. B. Petrig, J. Bigun, M.G. Curchod. Motion estimation of ocular fundus images. In *Proc. International Conference on Image Processing (ICIP)*, S. 691-694, 1996.
3. A. Dölemeyer, H. Liebau, F. Toonen, S. Wolf, D. Meyer-Ebrecht. Real-Time-Tracking von Augenbewegungen: Automatische Mikroperimetrie mit dem Scanning-Laser-Ophthalmoskop. In P. Levi, R.-J. Ahlers, F. May, M. Schanz (Hrsg.), *Mustererkennung 1998*, S. 227-233, 20. DAGM-Symposium, 1998.
4. K. Voss, W. Ortmann, H. Süße. Bildmatching und Bewegungskompensation bei Fundus-Bildern. In P. Levi, R.-J. Ahlers, F. May, M. Schanz (Hrsg.), *Mustererkennung 1998*, S. 439-446, 20. DAGM-Symposium, 1998.
5. S. Trauzettel-Klosinski, J. Reinhard. The vertical field border in hemianopia and its significance for fixation and reading. *Investigative Ophthalmology and Visual Science (IOVS)*, 39(11):2177-2186, 1998.