

Markert, Steffen; Berkes, Sebastian; Götze, Andreas; Plagwitz, Kai-Uwe; Husar, Peter; Henning, Günter:

## Echtzeitblickrichtungserfassung mit CMOS-Zweikamerasystem

---

<i>Zuerst erschienen in:</i>	Biomedizinische Technik = Biomedical Engineering. - Berlin [u.a.] : de Gruyter. - 47 (2002), S1b, S. 652-655. Jahrestagung der Deutschen Gesellschaft für Biomedizinische Technik (DGBM) im VDE ; 36 (Karlsruhe) : 2002.09.25-27
<i>Erstveröffentlichung:</i>	2002
<i>Datum Digitalisierung:</i>	2009-11-26
<i>ISSN (online):</i>	1862-278X
<i>ISSN(print)</i>	0013-5585
<i>DOI:</i>	<a href="https://doi.org/10.1515/bmte.2002.47.s1b.652">10.1515/bmte.2002.47.s1b.652</a>
<i>[Zuletzt gesehen:</i>	2019-12-06]

„Im Rahmen der hochschulweiten Open-Access-Strategie für die Zweitveröffentlichung identifiziert durch die Universitätsbibliothek Ilmenau.“

“Within the academic Open Access Strategy identified for deposition by Ilmenau University Library.”

„Dieser Beitrag ist mit Zustimmung des Rechteinhabers aufgrund einer (DFG-geförderten) Allianz- bzw. Nationallizenz frei zugänglich.“

„This publication is with permission of the rights owner freely accessible due to an Alliance licence and a national licence (funded by the DFG, German Research Foundation) respectively.“



# ECHTZEITBLICKRICHTUNGSERFASSUNG MIT CMOS-ZWEIKAMERASYSTEM

S. Markert, S. Berkes, A. Goetze, K.U. Plagwitz, P. Husar, G. Henning

Institut für Biomedizinische Technik und Informatik, Technische Universität Ilmenau, Deutschland

steffen.markert@tu-ilmenau.de

**Abstract**— Es wurde ein Eyetracking-Konzept entworfen, welches sich auf die VOG unter Verwendung eines Zweikamerasystems und spezieller Bildverarbeitungs- und 3D-Algorithmen stützt. Entwickelt wurde ein Versuchsaufbau, bestehend aus einem Zweikamerasystem, IR-Beleuchtung und einem dreh- und verschiebbar gelagerten Augenmodell. Der Algorithmus zur Berechnung der Lage des Messobjektes wurde für den Echtzeitbetrieb auf PC-Technik optimiert. Messergebnisse am Augenmodell zeigen, dass die angestrebten Genauigkeits- und Geschwindigkeitsanforderungen erfüllt werden. Das wichtigste Ergebnis der Messungen ist die Tatsache, dass die räumliche Lage eines pupillenähnlichen Messobjektes ohne vorherige Kalibration ermittelt werden kann. Weiterführende Untersuchungen müssen zeigen, ob dies auch für Messungen am menschlichen Auge gilt.

**Keywords**—Eyetracking, Blickrichtungserfassung, objektive Perimetrie, Kalibration

## Einleitung

Objektive Perimetrie beschäftigt sich mit der Analyse des Gesichtsfeldes. Dabei werden die nach optischer Stimulation im ERG bzw. EEG auftretenden Reizantworten ausgewertet. Die Zuordnung von Reizmarke und Stimulationsort auf der Retina ist von entscheidender Bedeutung, deshalb müssen Informationen über die Blickrichtung der untersuchten Person vorliegen. Mit Blickrichtungserfassungssystemen (Eyetracker) können die Blickbewegungen des menschlichen Auges untersucht werden.

Alle bereits existierenden Eyetracking-Systeme, die für die objektive Perimetrie in Frage kommen, weisen z.T. erhebliche Nachteile auf.

In der Elektrookulographie (EOG) werden durch Augenbewegungen verursachte Potentialdifferenzen mit Biosignalverstärkern aufgezeichnet. Messbereich und Geschwindigkeit sind recht groß, allerdings gibt es viele Störgrößen (EEG, EMG, Drift der Elektrodenpotentiale).

Search-Coil-Verfahren nutzen Induktionsspulen, die in Kontaktlinsen eingebettet sind. Abhängig von der räumlichen Lage des Auges in definierten Magnetfeldern können induzierte Spannungen gemessen werden. Hoher Genauigkeit und guter zeitlicher Auflösung stehen der große technische Aufwand und die relative hohe Patientenbelastung gegenüber.

Ein verbreitetes Verfahren ist die Video-Okulographie (VOG). Hier wird versucht, neben der Pupillenlage die Lage eines durch eine Lichtquelle erzeugten Cornealreflexpunktes aus Bildsequenzen zu ermitteln. Die Verknüpfung beider Informationen liefert Aussagen zur ungefähren Blickrichtung.

Durch den Einsatz von verbreiteter Kamera- und PC-Technik sind diese Systeme preiswert. Nachteilig neben der geringen zeitlichen Auflösung ist die erforderliche Kalibration vor jeder Messung. Bereits geringe Kopfbewegungen während einer Untersuchung führen zu großen Messfehlern. Für objektive perimetrische Untersuchungen sind die Anforderungen an Eyetracking-Systeme besonders hoch. Eine zeitaufwendige Kalibration ist insbesondere bei unkooperativen Patienten oder bei Fixationsproblemen (zentraler Gesichtsfeldausfall o.ä.) nicht praktikabel. Die Blickrichtung soll in hoher zeitlicher Auflösung (>100 Messungen pro Sekunde) und mit ausreichender Genauigkeit (Fehler < 3° bzw. < 1° bei Untersuchungen des zentralen Gesichtsfeldes [1]) ermittelt werden. Es sollte keine Patientenbelastung auftreten. Der technische Aufwand ist zu minimieren.

## Methoden

Das zur Lösung der o.g. Probleme entworfene Konzept stützt sich auf VOG, jedoch soll das Messobjekt gleichzeitig aus zwei unterschiedlichen Richtungen aufgenommen werden, um aus der zusätzlichen Bildinformation auf die absolute Lage des Messobjektes im Raum schließen zu können. Dieses Messprinzip ist bekannt (z.B. Luftbildauswertung), wurde jedoch für genaue Blickrichtungsmessungen am Auge unter Verwendung von schnellen Bildsequenzen noch nicht genutzt.

Für die theoretischen Abschätzungen galten einige Vereinfachungen: Es existieren zwei Projektionsebenen (Bildsensoren), die definiert im Raum positioniert sind. Sie bilden das Messobjekt aus unterschiedlichen Richtungen ab. Die Abbildung sei durch Parallelprojektion beschreibbar. Der Einfluss vorgelagerter Optik auf den Strahlengang sei bekannt. Werden im Bild beider Sensoren Punkte des Messobjektes gefunden, die gleichen Ursprungs sind, lässt sich deren Lage im Raum durch Strahlenrückverfolgung rekonstruieren. Mit der Lage der Punkte sind auch Position und Lage des Messobjektes bestimmt. Wird als Messobjekt die Pupille des menschlichen Auges angenommen, können unter Voraussetzung weiterer Vereinfachungen (Pupille kreisrund, verzerrungsfrei) Position und Ausrichtung ermittelt werden. Hieraus lassen sich Aussagen zur Blickrichtung ableiten.

Das System soll aus zwei Bildsensoren bestehen, deren Flächennormalen, ausgehend vom Schwerpunkt der Sensorflächen, sich im Koordinatenursprung des Gesamtsystems treffen. Jeder Bildsensor kann als Teil einer Ebene angenommen werden, welche durch die allgemeine Ebenengleichung

$$E: \mathbf{a}r + \mathbf{b}s + s_0 = 0 \quad (1)$$

beschrieben wird. Die beiden Vektoren  $\mathbf{a}$  und  $\mathbf{b}$  sollen ortho-

gonal zueinander stehen. Der Punkt  $s_0$  stellt den Mittelpunkt eines Sensors dar.

In Referenzposition soll gelten:

$$a_{ref} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, b_{ref} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, s_{0ref} = \begin{pmatrix} 0 \\ 0 \\ d_{cam} \end{pmatrix} \quad (2)$$

Somit sind in Referenzposition die Vektoren  $a$  und  $b$  mit den Einheitsvektoren in  $x$ - bzw.  $y$ -Richtung identisch. Der Punkt  $s_{0ref}$  liegt auf der  $z$ -Achse mit dem Abstand  $d_{cam}$  vom Koordinatenursprung entfernt.

Die Positionierung der Kameras kann durch zwei Drehungen um den Koordinatenursprung beschrieben werden. Für die Ebenenparameter ergibt sich dann:

$$a = M_{rot} a_{ref}, b = M_{rot} b_{ref}, s_0 = M_{rot} s_{0ref} \quad (3)$$

Die Sensorfläche ist definiert als:

$$|r| \leq \frac{numCols}{2} \frac{\beta}{pxS}, |s| \leq \frac{numRows}{2} \frac{\beta}{pxS} \quad (4)$$

Die Anzahl der Spalten wird durch  $numCols$ , die Anzahl der Zeilen durch  $numRows$  festgelegt. Die Pixelgröße definiert sich durch  $pxS$ , und wird in mm angegeben.

Ein beliebiger Punkt auf dem Messobjekt bildet bei angenommener Parallelprojektion mit dem Normalenvektor der Sensorfläche eine Gerade. Diese Gerade trifft die dem Sensor vorgelagerte Optik in genau einem Punkt. Das optische System bildet diesen Punkt auf der Sensorfläche ab. Durch Gleichsetzen der Ebenengleichung (Gl.1) mit der Geradengleichung erhält man den Schnittpunkt (Projektionspunkt).

$$E = p_{pupil} + nt \mid n = a \times b \quad (5)$$

Einsetzen der Ebenengleichung in E:

$$ar + bs - nt = p_{pupil} - s_0 \quad (6)$$

Lineares Gleichungssystem in Matrixschreibweise:

$$A = \begin{bmatrix} a & b & -n \end{bmatrix}, b = (p_{pupil} - s_0) \quad (7)$$

Lösung für das Gleichungssystem:

$$\begin{pmatrix} r \\ s \\ t \end{pmatrix} = A^{-1}b \quad (8)$$

Das Gleichungssystem ist immer dann lösbar, wenn gilt:

$$\det(A) \neq 0, \text{rang}(A) = 3 \quad (9)$$

Dies ist gegeben, da die Vektoren  $a$ ,  $b$  und  $n$  ein Rechtssystem bilden.

Für den auf dem Sensor abgebildeten Punkt gilt:

$$p_{sensor} = \begin{pmatrix} r \\ s \end{pmatrix} \beta / pxS \quad (10)$$

wobei  $\beta$  der Abbildungsmaßstab, und  $pxS$  die Größe eines Sensorpixels ist.

Auf der Grundlage dieser Betrachtungen wurde ein Algorithmus zur Lage- und Positionsbestimmung eines kreisrunden (in der Projektion meist elliptischen abgebildeten) Objektes entworfen:

1. Finden des Objektrandes in beiden Sensorbildern
2. Bestimmung des Mittelpunktes (Schwerpunktes) des Objektes
3. Suchen von 4 Punkten gleichen Ursprungs auf dem Objektrand in beiden Bildern
4. Rekonstruktion der 4 Randpunkte sowie des Mittelpunktes im 3D-Raum durch Strahlrückverfolgung
5. Ermittlung der Lage des Objektes im Raum

Um Punkte gleichen Ursprunges nach einer Drehung des Messobjektes in beiden Bildern wiederfinden zu können, müssen sie eindeutige Merkmale besitzen. Dies ist nicht unproblematisch für das hier betrachtete Messobjekt, lässt sich mit Wissen um die Kamerapositionen und unter gewissen Vorgaben jedoch ermitteln.

### Ergebnisse

Um die Realisierbarkeit der Konzeption zu belegen, erfolgten zunächst Messungen an einem stark vereinfachten Augenmodell, welches um die horizontale und vertikale Achse drehbar gelagert aufgehängt wurde. Alle Drehwinkel waren variierbar, die Drehung um die vertikale Achse war zehntelgradgenau einstellbar. Die Pupille wurde durch ein kreisrundes schwarzes Scheibchen vor weißem Hintergrund nachgebildet.

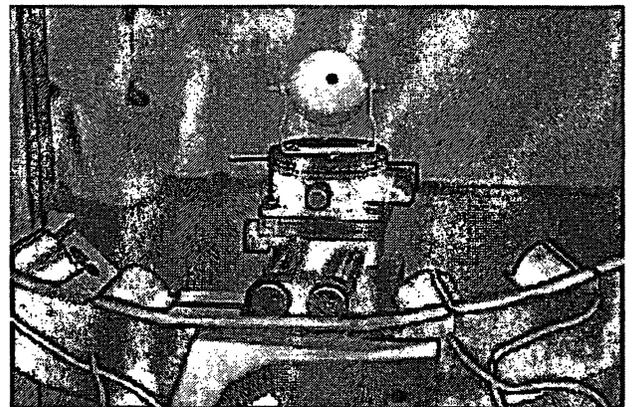


Abbildung 1: Augenmodell mit Nachbildung der Pupille

Zur Bilderfassung kamen zwei CMOS-Kameras (Graustufenbilder, 1280x1024 Pixel) mit IEEE 1394a-Schnittstelle zum Einsatz. Die wesentlichen Vorteile:

- kein Framegrabber, keine Bindung an analoge Video-Normen
- Auslesen von Teilbereichen des Sensors (Windowing), Bildrate steigt, Erfassung schneller Objektbewegungen möglich
- CMOS-Sensoren prinzipiell ohne Blooming (Verschmierungseffekte bei Überbelichtung von Pixeln)
- CMOS-Sensoren sind weit in den IR-Bereich hinein empfindlich

Als optisches System wurden zwei gleichartige Objektive mit Festbrennweite genutzt. Um Störungen durch andere Lichtquellen (z.B. Stimulationslicht) ausschliessen zu können, wurde das System auf den Betrieb im Infrarot-Bereich ausgelegt. Die vor den Objektiven befestigten IR-Filter lassen Licht ab einem Wellenlängenbereich von 780nm passieren. Dieses für den Menschen unsichtbare Licht wird beim praktischen Einsatz am menschlichen Auge die Ergebnisse perimetrischer Untersuchungen nicht beeinflussen. Zur Beleuchtung der Szenerie wurden Infrarot-LED-Strahler im Wellenlängenbereich um 880nm eingesetzt. Da aufwendige Vorverarbeitungsalgorithmen zunächst nicht implementiert werden sollten, wurde die Bildqualität durch die Erhöhung der Beleuchtungsintensität reguliert. Störende Bildelemente verschwinden auf diese Weise in der Sättigung des CMOS-Sensors, ohne dass Größe oder Form der interessierenden Abschnitte beeinflusst werden.

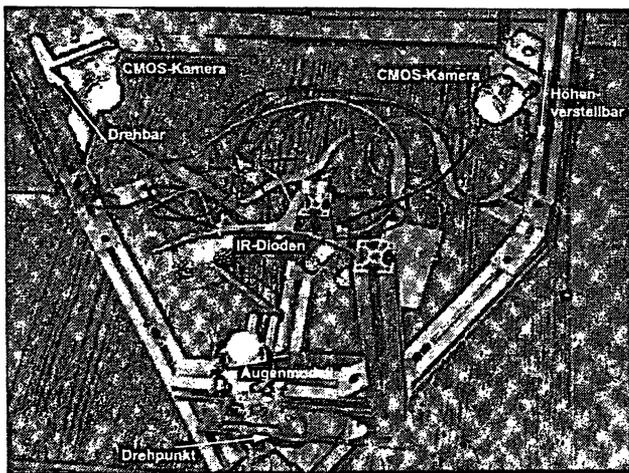


Abbildung 2: Aufbau der Versuchsanordnung

Der mechanische Aufbau wurde auf hohe Positioniergenauigkeiten (Kamera, Augenmodell, Beleuchtung) und Stabilität optimiert.

Die Verarbeitung der Datenströme beider Kameras erfolgte auf einem Standard-PC mit FireWire-PCI-Interface. Die Nutzung von geschwindigkeitsoptimierten Open-Source-Softwarebibliotheken (C++) von Intel® für Imageprocessing-Anwendungen erleichterte und beschleunigte die Umsetzung der vorgeschlagenen Algorithmen erheblich. Als Referenzsystem diente ein Intel® Pentium®3 (1GHz) mit 256 MByte RAM. Betriebssystem war Microsoft® Windows® 2000.

Das GUI der selbstentwickelten Software ermöglichte die Einstellung aller Kameraparameter, das Abspeichern von beliebigen Bildsequenzen sowie die Überwachung des Tracking-Algorithmus durch numerische und visuelle Darstellung aller Ergebnisse in Echtzeit (Bildwiederholrate der Kameras betrug bei der gewählten Bildgröße von 400x200 Pixeln 48fps).

Die durchgeführten Messreihen konzentrierten sich auf den Vergleich zwischen eingestelltem Drehwinkel  $\theta$  (um die vertikale Achse) des Objekts und dem berechneten Winkel. Für eine horizontale Drehung von  $\phi = 0^\circ$  zeigte sich, dass der maximale absolute Winkelfehler für  $\theta$  über den gesamten Messbereich unter  $0.6^\circ$  lag. Die mittlere

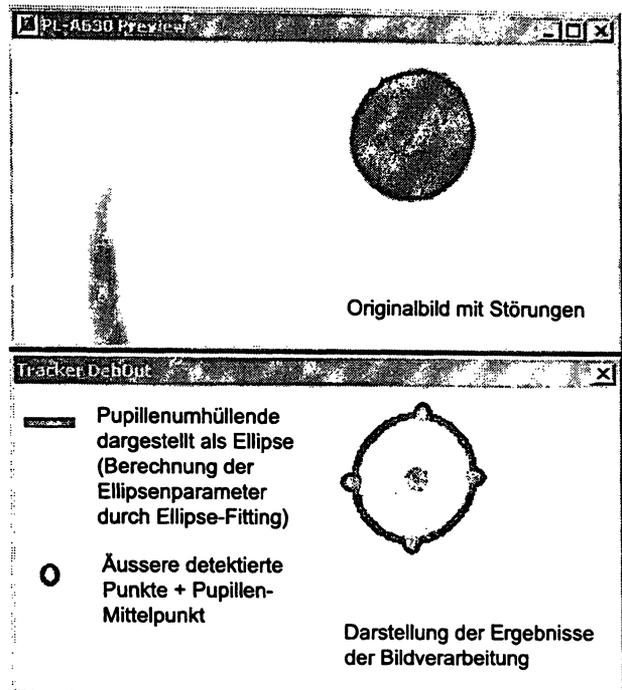


Abbildung 3: Darstellung des Trackings im GUI

Abweichung betrug  $0.3^\circ$ . Bei Veränderung von  $\theta$  mit einem festgelegten Winkel von  $\phi = 10^\circ$  erhöhte sich die mittlere Winkelabweichung über den gesamten Messbereich auf  $0.5^\circ$ , bei einem maximalen Winkelfehler von ca.  $1.2^\circ$ .

Für Verschiebungen des Augenmodells in x- bzw. y-Richtung bestätigten entsprechende Untersuchungen, dass der Einfluss auf die Drehwinkelbestimmung gering ist ( $<1^\circ$  für Verschiebungen um max. 10 mm in jede Richtung).

Für Einlesen und Verarbeitung beider Kamerabilder sowie die Berechnung der räumlichen Lage benötigte das Referenzsystem in allen Fällen weniger als 10ms (theoretisch 100fps).

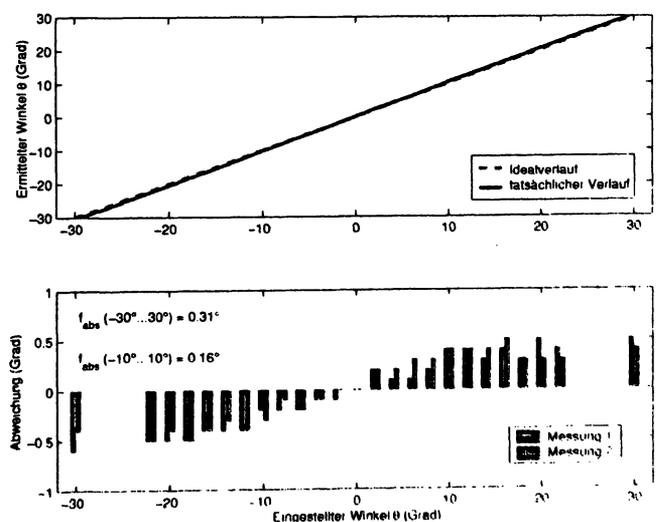


Abbildung 4: Messreihe  $\phi = 0^\circ, \theta = -30^\circ \dots 30^\circ$

## Diskussion

Die Messungen am Augenmodell zeigten, dass die Nutzung der Zweikameraanordnung in Verbindung mit den entworfenen Algorithmen zu dem angestrebten Ergebnis führt: die genaue räumliche Lage eines pupillenähnlichen Objektes kann aus der Information zweier aus unterschiedlichen Richtungen aufgenommenen Bildsequenzen des Objektes ermittelt werden. Die geforderte Genauigkeit der Messung wurde im gesamten Messbereich erreicht. Translative Bewegungen des Messobjekts hatten - sofern sie gewisse Grenzen nicht überschritten - nur geringen Einfluss auf die Ergebnisse. Die angestrebte Verarbeitungsgeschwindigkeit kann erreicht werden.

Nachfolgende Untersuchungen müssen zeigen, ob sich ähnliche Ergebnisse bei Messungen am menschlichen Auge ermitteln lassen. Ist dies der Fall, ließe sich kalibrationsfrei die Blickrichtung bestimmen. Außerdem würden geringe Kopfbewegungen keine Auswirkungen auf die Genauigkeit der Blickrichtungserfassung haben. Bilder des menschlichen Auges sind jedoch durch Lichtreflexe, Lidschlag, Wimpern etc. von geringerer Qualität als Bilder des Augenmodells, deshalb ist die Realisierung einer robusten Pupillendetektion im Vorfeld der erläuterten Verarbeitungsschritte notwendig. Ob hierbei die Übersteuerungstechnik angewendet werden kann, hängt einerseits von der Empfindlichkeit der Kamera, von der Wellenlänge IR-LED und nicht zuletzt von den Grenzwerten für die Bestrahlung des menschlichen Auges ab.

## Literaturverzeichnis

- [1] M.D. Menz, C.M. Poloschek, M.K. Menz, M. Wang, P. Penrose, E.E. Sutter, "The Effect Of Fixation Error On The Multifocal VEP", Proceedings of ARVO 2002 CD-ROM, 2002, Presentation Nr.: 4740
- [2] K.J. Ciuffreda, B. Tannen, "Eye Movement Basics for the Clinician", Mosby-Year Book, Inc. 1995
- [3] D. Strauß, "Grundlagenuntersuchung eines Verfahrens zur Blickrichtungsbestimmung", Diplomarbeit TU Ilmenau, 2001: Inventarisierungsnummer: 2001-12-14/0049/EI96/2221
- [4] A.S. Aguado, M.S.Nixon, "A new Hough Transform Mapping for Ellipse Detection", University of Southampton, June 1996
- [5] A.W. Fitzgibbon, M. Pilu, R.B. Fischer, "Direct least squares fitting of ellipses", Technical Report DAIRP-794, Department of Artificial Intelligence, The University of Edinburgh, January 1996
- [6] R. Halif, J. Flusser, "Numerically Stable Direct Least Squares Fitting of Ellipses", The Sixth International Conference in Central Europe on Computer Graphics and Visualization 1998
- [7] K. Kim, R.S. Ramakrishna, "Vision-Based Eye-Gaze Tracking for Human Computer Interface", IEEE International Conf. on Systems, Man, and Cybernetics, Tokyo, Japan, Oct. 1999

- [8] D. Young, H. Tunley, R. Samuels, "Specialized Hough Transform and Active Contour Methods for Real-Time Eye Tracking", Cognitive Science and Research Paper, no. 386 University of Sussex, England, U.K. 1995
- [9] K. Schreiber, "Erstellung und Optimierung von Algorithmen zur Messung von Augenbewegungen mittels Video-Okulographie-Methoden", Diplomarbeit University of Tübingen, 1999.
- [10] S. Shih, Y. Wu, J. Liu, "A Calibration Free Gaze Tracking Technique", Proceedings of the International Conference on Pattern Recognition, 2000