

Die Hough-Transformation als Verfahren zum robusten Eye Tracking

Frank Klefenz, Valiantsin Hardzeyeu

Fraunhofer IDMT

Zusammenfassung

Die Hough-Transformation ist in der Bildverarbeitung als Instrument der Formenerkennung nicht mehr wegzudenken und hat sich dahingehend zu einem Standardverfahren entwickelt. Die Grundzüge gehen in die 50er-Jahre zurück, aber erst mit dem Aufkommen leistungsfähigerer Rechner und der algorithmischen Parallelisierung wurde die Hough-Transformation für Echtzeitanwendungen nutzbar. Die Hough-Transformation erschließt viele Anwendungsfelder wie Fahrbahnranderkennung, Verkehrs- und Schriftzeichenerkennung als auch Eye Tracking. Jedoch ist der Algorithmus an die jeweiligen Aufgabenstellungen im Detail anzupassen. Aus diesem Grund wird die Hough-Transformation hier speziell als Werkzeug zum robusten Eye Tracking vorgestellt.

1 Einleitung

Die Hough-Transformation wurde von P.V.C Hough zur Detektion von Teilchenspuren, die auf einer Kreisbahn liegen, beschrieben (Hough 1959). Die mathematischen Lösungsgleichungen für Geraden, Kreise und Ellipsen sind exakt spezifizierbar, die Lösung kann grafisch erfolgen. Das Hough-Verfahren ist ein sehr rechenaufwendiges Histogramm-Update-Verfahren. Von seinen Ursprüngen her wurde der Hough-Algorithmus in vielen Triggersystemen der Hochenergiephysik eingesetzt. Das entscheidende Kriterium, um Echtzeitanwendungsfelder zu erschließen, ist die Performancesteigerung des Algorithmus. Dies kann dadurch erzielt werden, dass der Algorithmus massiv parallelisiert wird. Epstein et al. gibt die Realisierung eines VLSI ASIC zur parallelen Hough-Transformation an (Epstein et al. 2002). Chern und Lu und Chen sind weitere Referenzen zur Parallelisierung des Hough-Algorithmus (Chern & Lu 2005; Chen et al. 2004). Um die Echtzeitkriterien zu erfüllen, werden meist eingeschränkte Hough-Transformationen gerechnet, bei der nicht alle Bildpunkte berücksichtigt werden. Oft werden Randbedingungen gezielt genutzt, dass zum Beispiel der Ursprung des zu findenden Kreises fest in der Bildmitte liegt. Bei realistischen Bildverarbeitungsszenarien kann von diesen vereinfachenden Annahmen nicht ausgegangen werden, da sich die Bildinhalte durch die Bewegung verändern und gesuchte Objekte wie Augen, Verkehrsschilder und

Fahrbahnränder sich mit gewissen Einschränkungen frei im Raum befinden können. Kreise sind eine hypothetische Approximation, da runde Verkehrsschilder und das kugelförmige Auge durch eine seitliche Kameraposition in der Projektion als Ellipsen abgebildet werden. Für viele Spezialfälle wird daher eine echtzeitfähige Hough-Transformation für Ellipsen benötigt.

2 Elliptisches Eye Tracking

In erster Näherung kann das Auge als kugelförmig angesehen werden, wobei in der Feinanalyse auch die sphärischen Aberrationen und Krümmungen zu berücksichtigen sind. Abbildung 1 ist eine exemplarische Aufnahme der Pupille, wie sie im Routinebetrieb aufgenommen wurde. Durch die nichtzentrale Kameraprojektion wird die Aufnahme der Pupille ellipsenförmig.

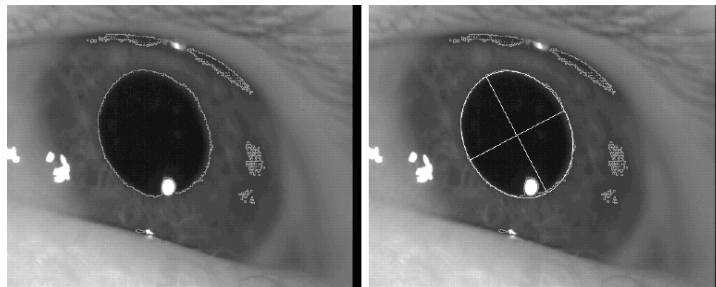


Abbildung 1: Bild links: Pupillenkontur, Bild rechts: eingezeichnete, berechnete Ellipse

2.1 Elliptischer Hough-Core

Johannes Katzmann und Christian Holland-Nell haben das in (Klefenz 1992) beschriebene Delayline Verfahren auf Ellipsen erweitert (Katzmann 2005; Holland-Nell 2005). Die Ellipse besteht aus fünf Freiheitsgraden (große, kleine Halbachse, (x,y) -Mittelpunktskoordinaten der Schnittpunkte der großen und kleinen Halbachse und Rotationswinkel θ). Bei der Kreislösung sind nur zwei Freiheitsgrade Krümmungsradius und Tangentialwinkel zu bestimmen, da der Kreis ortsfest im Zentrum (x_0, y_0) ruht. Bei der Delayline-Methode handelt es sich um eine parallele Berechnungsmethode für einfache Hough-Transformationen, wie die zur Detektierung von Linien oder Kreiskurven. Sie wird in (Klefenz 1992) für die Auswertung von Ereignissen in Teilchendetektoren beschrieben. Die Methode unterscheidet sich prinzipiell von anderen Hough-Transformations-Methoden, da zur Beschreibung keine Transformationsgleichung benötigt wird. Kern dieses Verfahrens ist eine digitale Struktur aus einfach verschalteten Delayelementen. Der Kern der Struktur besteht aus n Spalten (Delaylines) mit jeweils m Elementen, die die Bilddaten in einer bestimmten Geschwindigkeit von Anfang bis zum Ende transportieren. Eine Spalte kann pro Takt einen Pixel aufnehmen und transportieren, während viele nebeneinander liegende Spalten pro Takt eine Bildzeile aufnehmen können. Diese Bildzeile bewegt sich dann vertikal durch die Gesamtstruktur. Als Ausgangssig-

nal werden horizontal die Bildpunkte pro Bildzeile $N(i)$ gezählt und als Histogramm ausgewertet. Übersteigt ein Histogrammsummenwert einer Zeile des Hough-Cores eine vorher festgelegte Schwelle, wird ein Ereignis in dieser Zeile ausgelöst. Es steht für ein gefundenes Extremum.

Abbildung 2 zeigt die vier Hough-Cores für die vier verschiedenen Schieberichtungen (Abb. 2 – 4), in die jeweils eine Bildzeile (Bildspalte) pro Takt angelegt wird. Helle Punkte geben die aktivierten Delayelemente an. Das Computersimulationsprogramm ist über eine Vielzahl von Parametern einstellbar (Abb. 2 – 1, 2, 3, 5, 6, 7) und Parameter wie Bildgröße, maximale, minimale Halbachse können vorgegeben werden, um das System zu individualisieren. In Eigenversuchen konnte das Programm WinDelayLine durch Variation weniger Parameter zum Auffinden von Schriftzeichen, Verkehrsschildern, Fahrbahnrandern etc. eingestellt werden. In der rechten Spalte unten werden die gefundenen Ellipsenparameter ausgegeben.

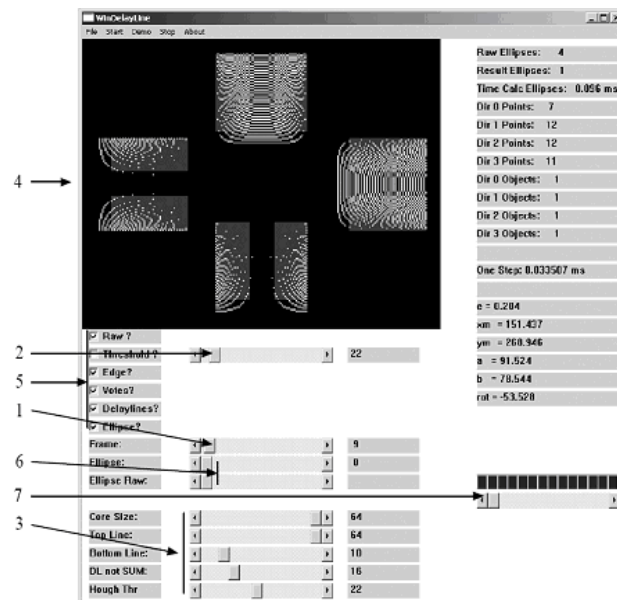


Abbildung 2: Grafische Oberfläche des Ellipsenfinder-Computerprogramms

2.2 Extrempunkte

Beim spalten- bzw. zeilenweisen Durchschieben der Daten durch den Hough-Core werden lokale Extrema bestimmt, die zu einem Kreis bestimmter Krümmung gehören. Dies ist eine numerische Lösung des ersten Hauptsatzes der Kurventheorie, die besagt, dass jede Funktion aus dem Extrempunkt und der zweiten Ableitung (Krümmung) entwickelbar ist. Durch die Entwicklung mit Kurvenscharen kann derselbe Hough-Core wie zur Kreiserkennung benutzt werden.

2.3 Echtzeitfähigkeit

Da der Hough-Core pro Takt mit einer Bildzeile versorgt wird, ist die Bearbeitung nach n Schritten + d beendet, wobei n die Zahl der Bildzeilen und d der Delay ist, bis der Hough-Core völlig leergetaktet ist. Tabelle 1 zeigt die Berechnungszeiten und Bildraten bei unterschiedlichen Bildgrößen.

Bilddimensionen in Pixel	128×128	256×256	512×512	400×500
Anzahl von Shifts im „Hough-Core“	8.192	49.152	229.376	168.000
Berechnungszeit für ein-zelles Bild	59 μ s	352 μ s	1.639 μ s	1.200 μ s
Zeitliche Auflösung	17.080 fps	2.840 fps	610 fps	830 fps

Tabelle 1: Berechnungszeiten und Bildraten bei unterschiedlichen Bildgrößen

3 Zusammenfassung und Ausblick

Der Algorithmus wurde als VHDL Netzliste synthetisiert und auf einem FPGA-Demoboard getestet. Im nächsten Schritt wird das System in einem FPGA-Framegrabbersystem integriert und danach seine Funktionalität praktisch getestet.

Literatur

- Chen, Ling, Chen, Hongjian, Pan, Yi, Chen, Yixin (2004): „A Fast Efficient Parallel Hough Transform Algorithm on LARPBS”, The Journal of Supercomputing 29, no. 2 (2004), pp 185-195
- Ming-Yang Chern, Yi-Hsiang Lu (2005): „Design and Integration of Parallel Hough-Transform Chips for High-speed Line Detection“, Proceedings of the 2005 11th International Conference on Parallel and Distributed Systems (ICPADS'05), IEEE Computer Society, 2005
- A. Epstein, G.U. Paul, B. Vettermann, C. Boulin, and F. Klefenz (2002): „A parallel systolic array ASIC for real-time execution of the Hough transform“, *IEEE Trans. Nucl. Scie.* vol.49, pp.339-346
- Holland-Nell, Christian (2005): „Implementierung eines Pupillen-Detektionsalgorithmus basierend auf der Hough-Transformation für Kreise“, Diplomarbeit, TU Ilmenau, 2005
- Paul V.C. Hough (1959): „Machine Analysis of Bubble Chamber Pictures“, Proc. of the Int. Conf. on High-energy accelerators and instrumentation, L. Kowarski, ed., Geneva, 14th – 19th Sep. 1959
- Katzmann, Johannes (2005): „Echtzeitfähige, auf der Hough-Transformation basierende Methoden der Bildverarbeitung zur Detektion von Ellipsen“, Diplomarbeit, TU Ilmenau, 2005
- Klefenz, F. (1992): „Ein systolischer Arrayprozessor zur Spurrekonstruktion als Trigger in Driftkammern“, Dissertation, Ruprecht-Karls-Universität, Heidelberg, 1992