Der Harris-Eckendetektor und die Strukturmatrix

Wolfgang Konen
Institut für Informatik, FH Köln

08.11.2006

Abstract. Dieser kurze Technical Report enthält einige Anmerkungen zum Harris-Eckendetektor [Harris88] und zur Strukturmatrix. Ziel ist es, die ansonsten sehr gute didaktische Darstellung bei [Burger05] um Aspekte zur Bedeutung der Strukturmatrix und ihrer Eigenwerte zur ergänzen.

Inhalt

Einleitung

Sei I(u,v) das Intensitäts-"Gebirge" eines Bildes. Die Ableitungen $I_x(u,v)$ und $I_y(u,v)$ charakterisieren den lokalen Grauwertverlauf an der Stelle (u,v).

Nach [Burger05] ist die *lokale Strukturmatrix* **M** definiert als (sämtliche Ableitungen sind an der Stelle (u,v) genommen):

(1) $$\mathbf{M}_0 = \begin{pmatrix} I_x{}^2 & I_x I_y \\ I_x I_y & I_y{}^2 \end{pmatrix} = \begin{pmatrix} A_0 & C_0 \\ C_0 & B_0 \end{pmatrix}$$

[Man beachte: die Strukturmatrix hat eine gewisse Ähnlichkeit zur Hesse-Matrix

$$\mathbf{H} = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix},$$

aber die Hesse-Matrix ist etwas deutlich anderes, da sie Ableitungen 2. Ordnung enthält].

Weil die in Gl. (1) definierte Matrix $\mathbf{M_0}$ sehr empfindlich gegen Rauschen ist, nimmt man zweckmäßigerweise eine Glättung vor, indem man mit einer Fensterfunktion $w_{u'v'}$ faltet (die Indizes u',v' laufen dabei über einen symmetrischen Bereich um u und v herum):

(2) $$\mathbf{M} = \begin{pmatrix} A & C \\ C & B \end{pmatrix} \quad \text{mit} \quad \begin{cases} A = \sum_{u'v'} w_{u'v'} I_x(u',v')^2 \\ B = \sum_{u'v'} w_{u'v'} I_y(u',v')^2 \\ C = \sum_{u'v'} w_{u'v'} I_x(u',v') I_y(u',v') \end{cases}$$

Folgende Fragen stellen sich:

- Welche Bedeutung hat die Strukturmatrix? Wieso sagt sie etwas über das Ecken- oder Kantenverhalten am Punkt (u,v) aus?
- Welche Bedeutung haben die Eigenwerte der Strukturmatrix?

Welche Bedeutung hat die Strukturmatrix?

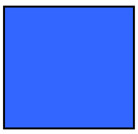
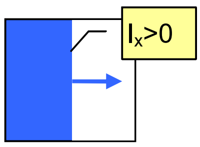
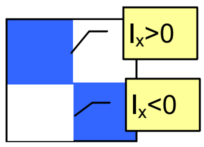
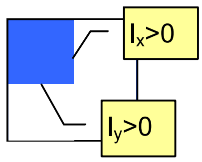
M charakterisiert die Struktur in der Umgebung des Punktes (u,v). Denn:

- Wenn (u,v) in einer strukturlosen Region des Bildes liegt (Intensitäts-"Gebirge" I ist flach, Plateau), dann sind sämtliche Ableitungen $I_x=I_y=0$, also ist **M** die Nullmatrix.
- Wenn durch (u,v) eine Kante mit Kantenvektor parallel zur x-Richtung verläuft, dann ist $I_x^2>0$, also auch A>0, aber $I_y=0$, mithin B=C=0. Wir haben also ein **M** in der Form

$$\mathbf{M} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$$

- Wenn durch (u,v) eine Ecke verläuft (z.B. die Ecke eines Schachbrettmusters), dann messen wir an einigen Stellen in der Umgebung von (u,v) ein $I_x^2>0$, an anderen Stellen ein $I_y^2>0$. Evtl. ist die Summe über alle Terme I_xI_y im Fensterbereich Null (Schachbrett), evtl. aber auch nicht (einseitige Ecke), wir haben also ein **M** in der Form

$$\mathbf{M} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \quad \text{oder} \quad \mathbf{M} = \begin{pmatrix} A & C \\ C & B \end{pmatrix}$$

keine Struktur	Kante (blauer Pfeil = Kantenvektor)	Ecke (Doppelecke "Schachbrett")	rechtwinklige Ecke
			
$\mathbf{M} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\mathbf{M} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$	$\mathbf{M} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$	$\mathbf{M} = \begin{pmatrix} A & C \\ C & B \end{pmatrix}$

Wenn man die Frage stellt: "Wenn ich für alle Pixel in einer Umgebung von (u,v) einen kleinen Schritt in Richtung (x_1,x_2) mache, welche (gemittelte) Intensitätsänderung habe ich dann, d.h. welchen Wert hat $E(x_1,x_2)$?" mit:

$$E\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \sum_{u,v} w_{uv}\left(I(u+x_1, v+x_2) - I(u,v)\right)^2 \tag{3}$$

dann ist die Antwort auf diese Frage, wie man mit einer Taylorentwicklung von $I(u+x_1,v+x_2)$ für kleine (x_1,x_2) zeigen kann, gegeben durch die sog. **quadratische Form** von **M** (s. auch [Harris88]):

$$E\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = (x_1, x_2)\,\mathbf{M}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{x}^T\mathbf{M}\mathbf{x} \tag{4}$$

Die Matrix **M** fasst die ganze Information über die Struktur von I in der Umgebung von (u,v) zusammen: Ich kann **M** mit kleinen Verschiebungen (x_1,x_2) verschiedenster Richtung "sampeln" und erhalte eine charakteristische Antwort zurück. Verschiebe ich z.B. genau senkrecht zu einem Kantenvektor, erhalte ich Null. An der gleichen Stelle erhalte ich aber bei Verschiebung parallel zum Kantenvektor eine maximale Antwort.

Dies ist die Begründung, wieso **M** die *lokale Strukturmatrix* heisst.

Welche Bedeutung haben die Eigenwerte der Strukturmatrix?

Ziel dieses Kapitels ist, darzulegen, dass mithilfe von Eigenwerten Kanten und Ecken in beliebigen Orientierungen durch eine einheitliche Terminologie charakterisiert werden können.

Eigenwerte rekapituliert

Bekanntlich sind Eigenwert λ und Eigenvektor **v** einer Matrix **M** ja diejenigen Elemente, für die die Eigenwertgleichung

(5) $$\mathbf{Mv} = \lambda\mathbf{v}$$

gilt. Eigenvektoren sind spezielle Richtungen im Raum, die die Abbildung **M** unverändert lässt.

Beispiel: Bei der Abbildung "Erddrehung" bleiben sämtliche Vektoren, die in Richtung der Rotationsachse Südpol-Nordpol zeigen in ihrer Richtung unverändert. Der zugehörige Eigenwert ist 1.

Es gilt der Satz, dass jede reelle, symmetrische N*N-Matrix genau N reelle Eigenwerte mit N zueinander senkrechten Eigenvektoren besitzt. Bei unserem **M** sind ferner alle Eigenwerte nichtnegativ.

Eigenwerte der Matrix M

Welche anschauliche Bedeutung haben Eigenwerte im Falle der Strukturmatrix **M**?

Beginnen wir zunächst mit einem einfachen Fall: C=0, d.h. **M** habe die Diagonalform

(6) $$\mathbf{M} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Dann sind die Eigenvektoren die Einheitsvektoren $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und die zugehörigen Eigenwerte sind λ_1 und λ_2. Wir interessieren uns nun für den Wert der quadratischen Form $\mathbf{x}^T\mathbf{Mx}$, wenn sie auf Einheitsvektoren **x** beliebiger Richtung angewendet wird. Mit ein wenig Algebra kann man zeigen

(7) $$\mathbf{x}^T\mathbf{Mx} = (x_1\mathbf{v}_1 + x_2\mathbf{v}_2)^T\mathbf{M}(x_1\mathbf{v}_1 + x_2\mathbf{v}_2) = \ldots = x_1^2\lambda_1 + (1 - x_1^2)\lambda_2$$

Dies ist die Gleichung für Punkte auf einer Ellipse, s. Abbildung 1. In Richtung des größeren Eigenwerts (dies sei o.B.d.A. λ_1) bewirkt **M** die größte Längenänderung, in Richtung von λ_2 die kleinste.

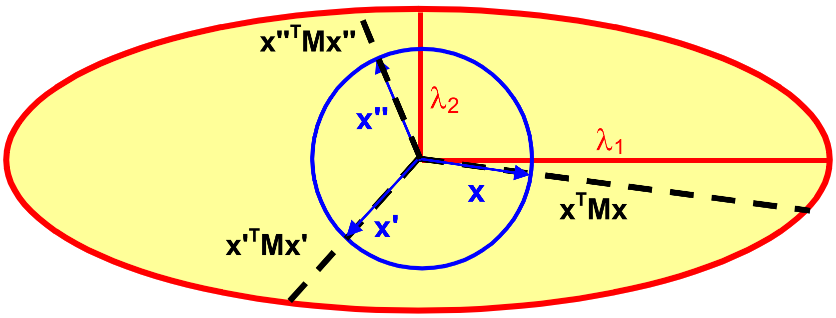


Abbildung 1: Die Abbildung $\mathbf{x}^T\mathbf{Mx}$ macht aus Vektoren x auf dem Einheitskreis einfache Zahlen, deren Länge durch die rote Ellipse charakterisiert ist. (Die Richtung von Mx ist i.a. ungleich der Richtung von x, nur bei den Hauptachsen gilt Mx || x)

Für die quadratischen Form aus Gl. (4) und die Einheits-Eigenvektoren $\mathbf{v_1}$, $\mathbf{v_2}$ gilt insbesondere:

$$E(\mathbf{v}_1) = \mathbf{v}_1^T \mathbf{M} \mathbf{v}_1 = \lambda_1 \tag{8}$$
$$E(\mathbf{v}_2) = \mathbf{v}_2^T \mathbf{M} \mathbf{v}_2 = \lambda_2$$

Eine Kante ist ein Sonderfall von Abbildung 1 mit λ_2=0: Vektoren in Richtung von $\mathbf{v}_2$ (also senkrecht zum Kantenvektor) werden durch **M** auf 0 abgebildet. Dies entspricht der Tatsache, dass senkrecht zum Kantenvektor keine lokale Struktur im Bild ist.

Was ist, wenn der Kantenvektor NICHT längs der x-Achse verläuft, sondern z.B. 35^0 geneigt dazu? – Eigentlich ist alles noch genauso, wenn wir in den Eigenvektoren $\mathbf{v}_1$ und $\mathbf{v}_2$ rechnen bzw. darstellen, s. Abbildung 2. Die Matrix **M** hat im v_1-v_2-Koordinatensystem weiterhin die Diagonalform nach Gl. (6).

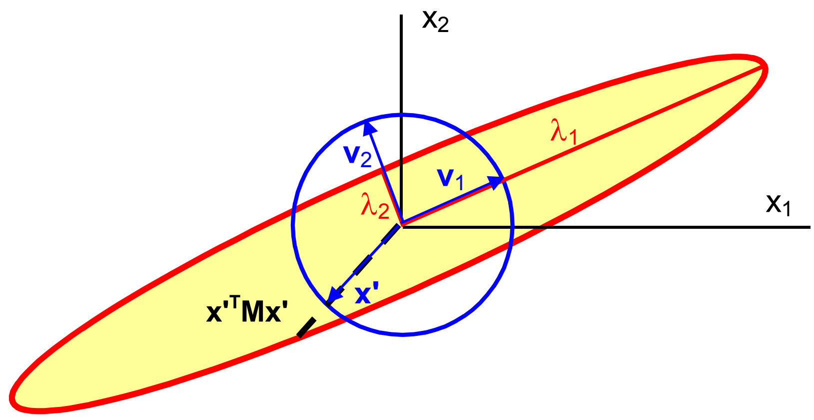


Abbildung 2: Sind die Eigenvektoren v_1 und v_2 gedreht, so gilt dasselbe wie in Abbildung 1 bezüglich dieser Vektoren. Im Koordinatensystem x_1, x_2 hat M aber dann eine andere Darstellung

Wenn wir aber die Matrix **M** im x_1-x_2-Koordinatensystem darstellen wollen, dann müssen wir sie drehen, und das geschieht bei Matrizen mit einer Drehmatrix **D**, die von links transponiert und von rechts "as-is" an die Diagonalform von **M** multipliziert wird (c=cos(ϕ), s= sin(ϕ)):

$$\mathbf{D} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix},$$

$$\mathbf{M} = \mathbf{D}^T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{D} = \begin{pmatrix} \lambda_1 c^2 + \lambda_2 s^2 & (\lambda_1 - \lambda_2) sc \\ (\lambda_1 - \lambda_2) sc & \lambda_1 s^2 + \lambda_2 c^2 \end{pmatrix} = \begin{pmatrix} A & C \\ C & B \end{pmatrix}$$

Man sieht, dass die Matrix **M** weiterhin symmetrisch ist. Aber die Matrix **M** ist im x_1-x_2-Koordinatensystem überhaupt nicht mehr diagonal, und auch bei einer Kante kann der Term B sehr wohl größer Null sein, obwohl "eigentlich" λ_2=0 ist.

Wir sehen also den großen **Vorteil der Beschreibung durch Eigenwerte**: Obwohl die Matrix **M** je nach Orientierung der Kante immer wieder anders aussieht, haben wir mit den Eigenwerten eine einheitliche Sprache, um alle Kanten beliebiger Richtung zu beschreiben:

"Kante" ist dann, wenn Eigenwert λ_1 > 0, aber λ_2=0 (bzw. nahe bei Null).

Dasselbe gilt für Ecken

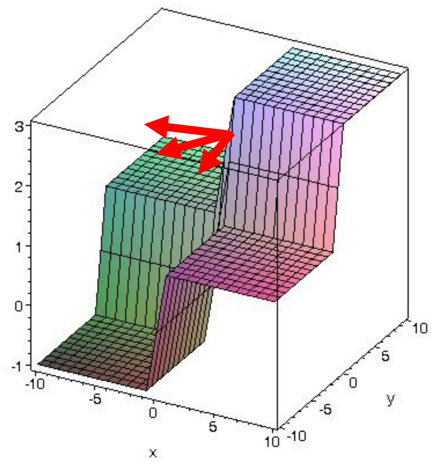
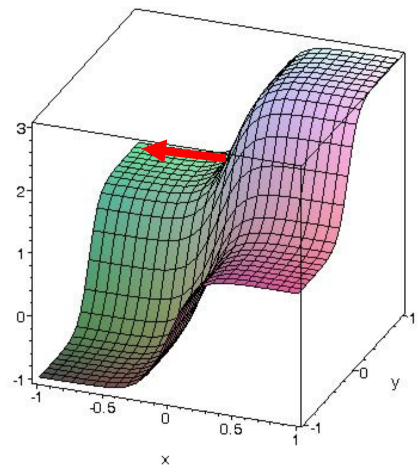
"Ecke" ist dann, wenn beide Eigenwerte in etwa gleich groß $\lambda_1 \approx \lambda_2$ (Ellipse wird zum Kreis)

Faltung ist wichtig für Eckendetektion (!)

Man beachte: Würde die Faltung (Übergang von $\mathbf{M}_0$ zu $\mathbf{M}$ in Gl. (2)) unterbleiben, dann gäbe es streng genommen gar keine Ecken: Denn mit den Werten A_0, B_0, $C_0^2 = A_0 \cdot B_0$ errechnet man mit der Eigenwertformel (8.7) in [Burger05]:

$$\lambda_{1,2} = \tfrac{1}{2}\left(A_0 + B_0 \pm \sqrt{A_0^{\,2} + 2A_0 B_0 + B_0^{\,2}} \right) \quad \Rightarrow \quad \lambda_1 = A_0 + B_0 , \quad \lambda_2 = 0$$

d.h. wir haben die Charakterisierung einer Kante mit Stärke A_0+B_0. Das macht auch Sinn, denn wenn wir die Stetigkeit des Intensitätsgebirge voraussetzen und an eine Ecke "ganz nah ran gehen", sieht diese Ecke lokal wie eine Kante in der Richtung aus, die die Mittelung beider dort "tatsächlich" zusammenlaufenden Kantenrichtungen ist.

	
Einer Ecke kann "von weitem" (x und y laufen von -10 bis 10) keine Kantenrichtung zugeordnet werden.	Wenn wir aber "ganz nah ran gehen" (x und y laufen von -1 bis 1), dann haben wir – beliebige Auflösung vorausgesetzt – auch an der Ecke eine Kante, d.h. eine Grauwertrampe vor uns.

Erst mit einer Faltung w_{uv}, die sich über mehrere Pixel erstreckt, erreichen wir, dass i.d.R. $C^2 \neq AB$ gilt, und nun können wir zwei von Null verschiedene Eigenwerte λ_1 und λ_2 haben. Die Faltung hat anschaulich die Wirkung, verschiedene Steigungsantworten in der Umgebung von (u,v) aufzusammeln. Auf diese Weise kann man **die Umgebung von (u,v)** genauer charakterisieren als wenn man nur den Funktionsverlauf an der Stelle (u,v) selbst betrachtet.

Literatur

[**Harris88**] C. Harris and M. Stephens (1988). "A combined corner and edge detector". *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151.

[**Burger05**] W. Burger, M. J. Burge, *Digitale Bildverarbeitung – Eine Einführung mit Java und ImageJ*. Kap. 8 *"Auffinden von Eckpunkten"*, S. 135–150, eXamen.press, 2005.