Lezione sull'Harris corner detector

A. Masiero

January 31, 2007

1 Breve introduzione

Vista l'enorme quantità di dati contenuti in un'immagine, è fondamentale avere a disposizione un qualche algoritmo che a partire da questi dati riesca a riconoscere e selezionare solo quelli di interesse per lo scopo specifico che ci si è predisposti. In particolare in questa lezione supponiamo che lo scopo sia quello di rilevare gli angoli (o eventualmente gli "edge") dell'immagine. Questi diventeranno percio' le caratteristiche salienti, o *features*, che vogliamo poter estrarre dall'immagine stessa. A tal fine andremo ad esporre un metodo ormai diventato classico per l'estrazione di tali features.

2 Harris corner detector

Sia I un'immagine espressa in scala di grigi[1]. Siano inoltre le coordinate u e v le coordinate sul piano sul piano corrispondenti rispettivamente alla direzione orizzontale e verticale.

Si noti innanzi tutto che ad "edge" e "corner" corrispondono solitamente veloci variazioni nell'intensità dell'immagine (in degli intorni dei punti corrispondenti a tali features). Volendo rilevare delle "grosse" variazioni della I viene spontaneo andare ad esprimere la I stessa attraverso il suo sviluppo di Taylor (per semplicità troncato al primo ordine). Sia perciò x_0 un qualsiasi punto all'interno del dominio dell'immagine ed h un piccolo vettore spostamento, allora negli intorni di x_0 vale

$$I(x_0 + h) \approx I(x_0) + (\nabla I(x_0))^T h$$

dove con $\nabla I(x_0)$ abbiamo indicato il gradiente dell'immagine nel punto x_0, cioè:

$$\nabla I(x_0) = \begin{bmatrix} \frac{\partial I}{\partial u}(x_0) \\ \frac{\partial I}{\partial v}(x_0) \end{bmatrix}$$

[1] L'estensione al caso di un'immagine a colori è sostanzialmente immediata.

Per comodità di notazione useremo: $I_u(x_0) = \frac{\partial I}{\partial u}(x_0)$ e $I_v(x_0) = \frac{\partial I}{\partial v}(x_0)$. Si noti inoltre che essendo in realtà l'immagine descritta su di un dominio discreto avremo che: $I_u(x_0) = I_u(u, v) = I(u + 1, v) - I(u, v)$ e $I_v(u, v) = I(u, v + 1) - I(u, v)$.

La variazione dell'intensità dell'immagine in un intorno di x_0 può perciò essere scritta come:

$$I(x_0 + h) - I(x_0) \approx (\nabla I(x_0))^T h$$

Ora è anche chiaro che non siamo interessati al segno di tale variazione, ci interessa soltanto verificare se tale variazione è considerevole. Possiamo perciò considerarne il quadrato (questo ha il duplice effetto di eliminare il segno ed amplificare la differenza tra piccole e grosse variazioni):

$$(I(x_0 + h) - I(x_0))^2 \approx \left((\nabla I(x_0))^T h\right)^2 = h^T \nabla I(x_0)(\nabla I(x_0))^T h \ . \qquad (1)$$

Sfortunatamente però l'immagine a nostra disposizione sarà un'immagine tipicamente affetta da rumore: sarà perciò opportuno stimare $(I(x_0 + h) - I(x_0))^2$ "mediando" i valori che si trovano per tale indice in un intorno Ω_{x_0} di x_0 stesso:

$$(I(x_0 + h) - I(x_0))^2 \approx \sum_{x \in \Omega_{x_0}} w(x - x_0)(I(x + h) - I(x))^2 \qquad (2)$$

dove $w(x - x_0)$ è una qualche funzione peso. Tipicamente essa viene presa gaussiana. Si noti che la (1) continua a valere anche per un generico punto x:

$$(I(x + h) - I(x))^2 \approx \left((\nabla I(x))^T h\right)^2 = h^T \nabla I(x)(\nabla I(x))^T h$$

perciò andandola a sostituire in (2) otteniamo

$$(I(x_0 + h) - I(x_0))^2 \approx \sum_{x \in \Omega_{x_0}} w(x - x_0) h^T \nabla I(x)(\nabla I(x))^T h = E(x_0)$$

con

$$E(x_0) = \sum_{x \in \Omega_{x_0}} w(x - x_0) h^T \nabla I(x)(\nabla I(x))^T h \ .$$

È evidente che la h non dipende dalla x, per cui può essere portata fuori dalla sommatoria:

$$E(x_0) = h^T \left(\sum_{x \in \Omega_{x_0}} w(x - x_0) \nabla I(x)(\nabla I(x))^T \right) h = h^T C h$$

dove nell'ultima uguaglianza abbiamo posto:

$$C = \sum_{x \in \Omega_{x_0}} w(x-x_0) \nabla I(x) (\nabla I(x))^T = \sum_{x \in \Omega_{x_0}} w(x-x_0) \begin{bmatrix} I_u^2(x) & I_u(x)I_v(x) \\ I_u(x)I_v(x) & I_v^2(x) \end{bmatrix} .$$

Infine notando che la sommatoria può essere anche portata all'interno della parentesi si ottiene:

$$C = \begin{bmatrix} \sum_{x \in \Omega_{x_0}} w(x-x_0) I_u^2(x) & \sum_{x \in \Omega_{x_0}} w(x-x_0) I_u(x)I_v(x) \\ \sum_{x \in \Omega_{x_0}} w(x-x_0) I_u(x)I_v(x) & \sum_{x \in \Omega_{x_0}} w(x-x_0) I_v^2(x) \end{bmatrix} .$$

La matrice C così ottenuta è simmetrica e semi-definita positiva, può perciò essere decomposta tramite SVD in

$$C = USU^T = U \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} U^T .$$

Lo spostamento h può allora essere riscritto tramite le basi contenute nella U, cioè

$$U = \left[\begin{array}{c|c} u_1 & u_2 \end{array} \right] \Rightarrow h = Ua = u_1 a_1 + u_2 a_2$$

ed essendo la U una matrice unitaria si ha che $||a||^2 = ||h||^2$ (in pratica abbiamo effettuato solo una sorta di rotazione del sistema di coordinate). Ma allora:

$$E(x_0) = h^T C h = h^T USU^T h = a^T U^T USU^T Ua = a^T Sa = a_1^2 \sigma_1 + a_2^2 \sigma_2 .$$

Ricordando l'interpretazione geometrica dell'SVD questo ci dice che:

1. Quando $\sigma_1 \gg 0$ e $\sigma_2 \approx 0$ allora nei dintorni del punto x_0 ho una grossa variazione dell'intensità dell'immagine lungo la direzione $u_1 \Rightarrow$ in prossimità di x_0 passerà perciò un edge nella direzione *ortogonale* ad u_1.

2. Quando $\sigma_1 \gg 0$ e $\sigma_2 \gg 0$ allora nei dintorni del punto x_0 ho una grossa variazione dell'intensità dell'immagine sia lungo la direzione u_1 sia lungo la direzione $u_2 \Rightarrow$ avrò perciò un corner in prossimità di x_0.

3. Se $\sigma_1 \approx 0$ e $\sigma_2 \approx 0$ non sono in presenza nè di edge nè di corner.

Si noti che il calcolo dell'SVD di C andrebbe perciò fatto per ogni punto dell'immagine[2]: ciò può essere pesante dal punto di vista computazionale,

[2] Cioè C è da considerarsi funzione di $x_0 \Rightarrow C(x_0)$.

perciò al posto di calcolare direttamente l'SVD si può considerare la possibilità di utilizzare degli indici simili ma più veloci da calcolare: nella fattispecie si può considerare ad esempio[3]

$$R(x_0) = \det(C(x_0)) - k \ \mathrm{trace}^2(C(x_0))$$

dove k è una opportuna costante. Si può verificare che

$$\mathrm{trace}(C(x_0)) = \sigma_1 + \sigma_2$$

e

$$\det(C(x_0)) = \sigma_1 \sigma_2 \ .$$

Quindi $R(x_0) \gg 0$ se si è in presenza di un corner mentre $R(x_0) \ll 0$ se si è in presenza di un edge. Un altro possibile indice è ad esempio [4]

$$N(x_0) = \det(C(x_0)) / \ \mathrm{trace}^2(C(x_0)) \ .$$

La procedura così descritta va comunque raffinata in modo da selezionare solo i massimi locali: ad esempio nel caso di un possibile corner[5] in x_0 sarà sufficiente considerarlo effettivamente come corner solo se esso è il punto più probabile ad essere un corner in un tutto un suo intorno[6].

A Nozioni utili di algebra lineare: SVD

Sia $A \in \mathbb{R}^{m \times p}$ una matrice di rango n, dove $n \leq \{m, p\}$. Allora esistono due matrici U, V e una sequenza di numeri $\sigma_1, \ldots, \sigma_n$ tali che:

$$A = USV^T$$

con $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{p \times p}$,

$$UU^T = U^T U = I_m$$

$$VV^T = V^T V = I_p$$

[3]Proposto da Harris e Stephens.

[4]Proposto da Noble.

[5]Cioè in x_0 si ha che $\sigma_1 > \bar{\sigma}$ e $\sigma_2 > \bar{\sigma}$, dove $\bar{\sigma}$ è una opportuna soglia. NB Analogamente nel caso in cui si consideri $R(x_0)$ o un altro indice.

[6]Ad esempio se $(\sigma_1(x_0)\sigma_2(x_0))$ è un massimo locale di $(\sigma_1(x)\sigma_2(x))$. Si ragiona analogamente nel caso in cui si considerino altri indici.

$$S = \left[\begin{array}{cccc|ccc}
\sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & \sigma_2 & \dots & 0 & & & \\
0 & 0 & \dots & 0 & & & \\
0 & 0 & \dots & \sigma_n & 0 & \dots & 0 \\
\hline
0 & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0 & \dots & 0 & \vdots & \ddots & \vdots \\
0 & 0 & \dots & 0 & 0 & \dots & 0
\end{array}\right]$$

e

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0 \ .$$

Si ricordi che i vettori colonna "più significativi" che compongono U e V vengono presi tra gli autovettori di AA^T e A^TA.

È inoltre facile dimostrare che se $A = A^T \geq 0$ allora $A = USU^T$, si ha cioè che $V = U$.