

[CTT451] – [Nhập môn Thị giác Máy tính]

Tháng 4/2013

SIFT – HARRIS - BLOB



Bộ môn TGMT và KH Rô-bốt
Khoa Công nghệ thông tin
ĐH Khoa học tự nhiên TP HCM



MỤC LỤC

MỤC LỤC	1
1 Scale invariant feature transform (SIFT).....	3
1.1 Xây dựng không gian tỉ lệ	3
1.2 Dò tìm cực trị cục bộ	4
1.3 Loại bỏ keypoint có độ tương phản (contrast) thấp.....	4
1.4 Loại bỏ keypoint nằm trên biên cạnh	5
1.5 Gán hướng cho keypoint	5
1.6 Miêu tả đặc trưng.....	6
2 Thuật toán phát hiện góc Harris	7
3 Thuật toán phát hiện Blob	8
4 Bài tập.....	9
Tài liệu tham khảo	10

1 Scale invariant feature transform (SIFT)

1.1 Xây dựng không gian tỉ lệ

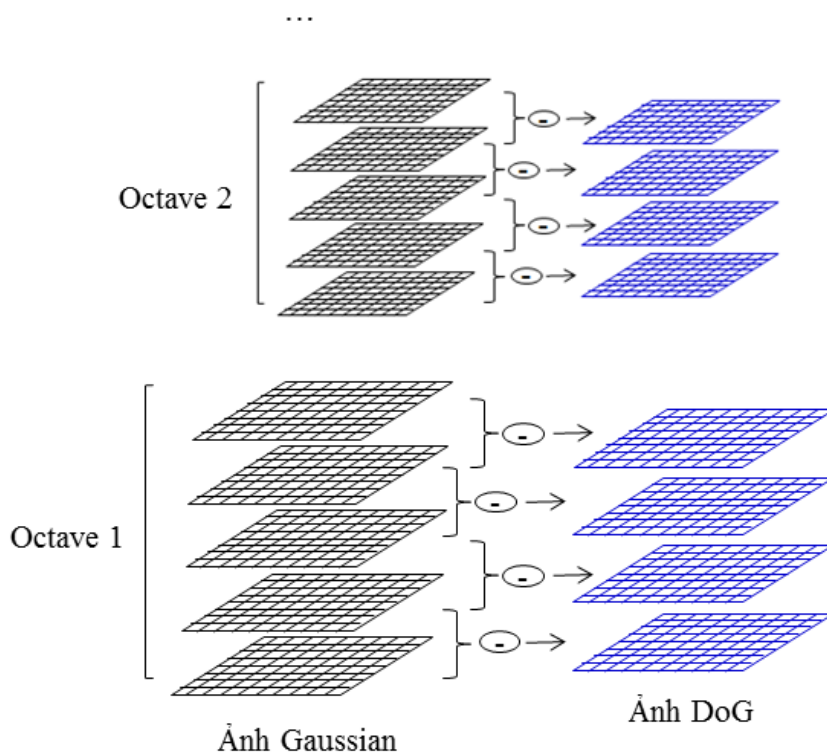
- Không gian tỉ lệ (scale space) gồm các ảnh $L(x, y, \sigma)$, tích chập (convolution) của lọc Gaussian $G(x, y, \sigma)$ với ảnh đầu vào $I(x, y)$.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

- Không gian tỉ lệ được chia thành từng octave, mỗi octave gồm 1 số nguyên s ảnh. Các ảnh tích chập với lọc Gaussian khác nhau ở tham số k , $k = 2^{1/s}$. Phải tạo ra $s + 3$ ảnh cho mỗi octave, ảnh được làm mờ thứ i trong octave sẽ là $L(x, y, k^i \sigma)$.
- Difference-of-Gaussian được sử dụng để phát hiện điểm trọng yếu (keypoint) trong không gian tỉ lệ, bằng cách trừ 2 ảnh L kế nhau trong octave như trong **Error! Reference source not found.**

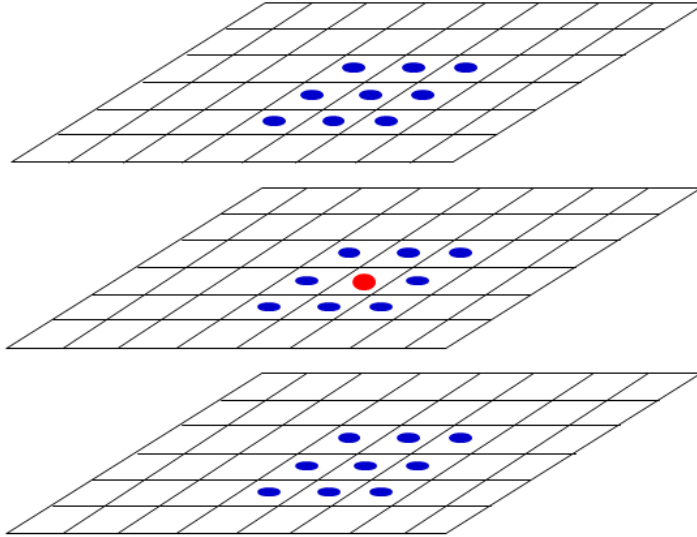
$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (1.1)$$

- Sau khi xử lý xong trên 1 octave, ảnh đầu tiên trong octave kế tiếp được giảm kích thước đi một nửa.



1.2 Dò tìm cực trị cục bộ

- Tìm cực đại và cực tiểu của các $D(x, y, \sigma)$, mỗi điểm được so sánh với 8 lân cận trong ảnh hiện tại, và 9 lân cận trong scale trên và scale dưới. Một điểm được chọn chỉ khi hoặc nó lớn hơn tất cả hoặc nhỏ hơn tất cả các lân cận.



1.3 Loại bỏ keypoint có độ tương phản (contrast) thấp

Áp dụng khai triển Taylor cấp 2 cho hàm $D(x, y, \sigma)$.

$$D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (1.2)$$

Trong đó: $x = (x, y, t)^T$

Đạo hàm cấp 1 của D tại x:

$$\frac{\partial D}{\partial x} = \left(\frac{\partial D}{\partial x} \quad \frac{\partial D}{\partial y} \quad \frac{\partial D}{\partial \sigma} \right)$$

Đạo hàm cấp 2 của D tại x:

$$\frac{\partial^2 D}{\partial x^2} = \begin{pmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial y \partial x} & \frac{\partial^2 D}{\partial \sigma \partial x} \\ \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial y^2} & \frac{\partial^2 D}{\partial \sigma \partial y} \\ \frac{\partial^2 D}{\partial x \partial \sigma} & \frac{\partial^2 D}{\partial y \partial \sigma} & \frac{\partial^2 D}{\partial \sigma^2} \end{pmatrix}$$

Lấy đạo hàm của D(x) và cho bằng không:

$$\hat{x} = - \frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (1.3)$$

Nếu $\hat{x} > 0.5$ tại bất kỳ hướng nào thì điểm cực trị nằm gần một keypoint khác hơn là keypoint đang xét, loại keypoint đang xét.

Giá trị của hàm $D(\hat{x})$ được sử dụng để loại những cực trị có độ tương phản thấp, thay \hat{x} vào D ta được:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x}$$

Những cực trị có giá trị $|D(\hat{x})| < 0.03$ sẽ được loại bỏ (giả sử giá trị độ xám của điểm ảnh nằm trong đoạn $[0,1]$).

1.4 Loại bỏ keypoint nằm trên biên cạnh

Tiến hành loại bỏ các điểm nằm trên biên cạnh, sẽ không ổn định nếu có một lượng nhiều nhỏ. Phương pháp Harris được dùng để xác định xem 1 keypoint nằm ở góc, biên cạnh hay trên vùng phẳng.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (1.4)$$

Đạo hàm được ước lượng thông qua hiệu số các điểm lân cận. Để tránh tính các giá trị riêng, mà chỉ cần quan tâm tới tỉ số của chúng. Gọi λ_1, λ_2 là hai giá trị riêng của H .

$$Tr(H) = D_{xx} + D_{yy} = \lambda_1 + \lambda_2$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \lambda_1\lambda_2$$

Gọi r là tỉ số giữa 2 giá trị riêng: $\lambda_1 = r\lambda_2$

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\lambda_1 + \lambda_2)^2}{\lambda_1\lambda_2} = \frac{(r\lambda_2 + \lambda_2)^2}{r\lambda_2^2} = \frac{(r+1)^2}{r}$$

Chỉ giữ lại những keypoint mà có:

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}$$

Ngưỡng đề nghị $r = 10$.

1.5 Gán hướng cho keypoint

Gán một hướng thích hợp cho keypoint dựa trên đặc tính cục bộ của ảnh, vì vậy bất biến với phép quay ảnh.

Với mỗi ảnh được làm mờ $L(x,y)$ ở scale gần nhất mà keypoint được phát hiện, độ lớn gradient $m(x,y)$ và hướng $\theta(x,y)$ được tính:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}$$

$$\theta(x,y) =$$

$$\tan^{-1}((L(x, y + 1) - L(x, y - 1)) / (L(x + 1, y) - L(x - 1, y)))$$

Lược đồ hướng gồm 36 bins để biểu diễn hết 360° của các hướng gradient. Các điểm xung quanh keypoint được thêm vào lược đồ bằng độ lớn gradient và cửa sổ tròn trọng số Gaussian.

Tìm đỉnh cao nhất trong lược đồ, các đỉnh còn lại mà đạt 80% so với đỉnh cao nhất, thì sẽ tạo keypoint ứng với hướng này.

1.6 Miêu tả đặc trưng

Tính độ lớn và hướng gradient của mỗi điểm xung quanh keypoint. Hàm Gaussian được dùng để gán trọng số độ lớn cho mỗi điểm. Kích thước vùng xung quanh keypoint là 16×16 , và được chia thành 4×4 vùng con.

Vector miêu tả đặc trưng chứa giá trị của tất cả lược đồ hướng. Vùng con 4×4 của lược đồ biểu diễn 8 hướng cho mỗi bin. Số thành phần của vector đặc trưng cho mỗi keypoint là $4 \times 4 \times 8 = 128$.

2 Thuật toán phát hiện góc Harris

Tính đạo hàm theo x và y của ảnh

$$I_x = G_\sigma^x * I; I_y = G_\sigma^y * I$$

Tính tích tại mỗi pixel của đạo hàm I_x và I_y

$$I_{x2} = I_x * I_x; I_{y2} = I_y * I_y; I_{xy} = I_x * I_y$$

Áp bộ lọc Gaussian lên I_{x2} , I_{y2} và I_{xy}

$$I_{xx} = G_{\sigma'} * I_{x2}; I_{yy} = G_{\sigma'} * I_{y2}; I_{xy} = G_{\sigma'} * I_{xy}$$

Tại mỗi pixel:

$$H(x, y) = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix}$$

Gọi:

$$A = I_{xx}$$

$$B = I_{xy}$$

$$C = I_{yy}$$

$$M = \det(H) - k * \text{Trace}(H)^2 > \text{ngưỡng}$$

Tính nonmax suppression

Tại mỗi pixel, giá trị $M(x,y)$ không lớn hơn các pixel xung quanh (3 x 3) thì loại pixel này.

3 Thuật toán phát hiện Blob

Xây dựng không gian tỉ lệ (scale space):

Từ ảnh ban đầu tạo ra các ảnh mờ với mức độ khác nhau. Sau đó thay đổi kích thước ảnh ban đầu xuống một nửa, và tiếp tục sinh ra các ảnh mờ. Quá trình cứ lặp lại như thế.

Để tạo ra các ảnh mờ có thể áp dụng bộ lọc Laplacian. Có thể xấp xỉ Laplacian bởi Difference of Gaussians.

Phát hiện các cực đại (cực tiểu) cục bộ: blob là các vị trí mà có cực đại cục bộ không chỉ trong cùng ảnh (8 pixel lân cận) mà còn với các pixel trong hai ảnh tỉ lệ lân cận (2 x 9 pixel lân cận), như vậy tổng cộng 8 + 18 pixel lân cận được xem xét.

4 Bài tập

- Viết chương trình tìm đặc trưng SIFT.
- Viết chương trình phát hiện góc Harris
- Viết chương trình phát hiện Blob.

Tài liệu tham khảo

- [1] David G. Lowe, "**Distinctive image features from scale-invariant keypoints**," *International Journal of Computer Vision*, 60, 2 (2004), pp. 91-110.
- [2] C. Harris and M. Stephens (1988). "A combined corner and edge detector".
Proceedings of the 4th Alvey Vision Conference. pp. 147–151.