

PROJETO DEMONSTRATIVO 6

JONATHAN ALIS S. LIMA*

*Disciplina: Visão computacional

Professor Flávio Vidal

Universidade de Brasília

Brasília, DF, Brasília 14 de Maio de 2015

Email: jonathanalis@gmail.com

Keywords— Recognition, Harris, SURF

Palavras-chave— Reconhecimento, Harris, SURF

1 Objetivos

O objetivo deste projeto é a exploração e utilização de algoritmos para análise e comparativo do processo de reconhecimento de padrões em imagens.

Para tal, é exigido que se utilize 2 algoritmos de reconhecimento de padrões em imagens. Um dos métodos é conhecido como Harris e o outro como SURF. É exigido que aplique os dois métodos na parte de uma imagem que contenha um objeto de interesse, e depois fazer o *match* de pontos do objeto com os pontos da captura de câmera. É esperado que quando mostrado o objeto na captura da câmera, os pontos de interesse encontrados sejam os mesmos que aqueles encontrados no objeto de interesse. Serão avaliados robustez a variação de escala, a rotação e o tempo de execução de cada método.

2 Introdução

Para a realização deste projeto, precisamos descrever o funcionamento dos métodos de detecção e reconhecimento segundo o método de Harris e o SURF. Esta seção apresenta a base teórica para o entendimento destes métodos.

2.1 Detecção e reconhecimento de padrões

As técnicas de detecção e reconhecimento se baseiam em encontrar características nas imagens consideradas boas para se que seguir. Em especial, uma das características, em que os métodos se baseiam, é em encontrar quinas (corners, encontro de duas bordas) (Laganière, 2011).

2.2 Harris

O algoritmo de Harris (Harris and Stephens, 1988) para detecção de quinas se baseia na busca de variações médias de intensidade em diferentes direções em volta de uma ponto de interesse. Se considerarmos o vetor de deslocamento (u, v) , a alteração média de intensidade é

$$R = \sum (I(x + u, y + v) - I(x, y))^2, \quad (1)$$

uma soma sobre a vizinhança de (x, y) . A definição de quina utilizada se baseia em que essa alteração média é maior que um certo limiar em mais de uma direção. A partir dessa definição, Harris desenvolveu seu método. Primeiramente, obtém a direção de maior variação de intensidade média, e em seguida, observa se a direção ortogonal também apresenta uma alta variação de intensidade média, caso positivo, o ponto de interesse é considerado uma quina.

Matematicamente, essa condição pode ser testada utilizando uma aproximação da expansão de Taylor a partir da Eq. 1:

$$\begin{aligned} R &\approx \sum \left(I(x, y) + \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v - I(x, y) \right) \\ &= \sum \left(\left(\frac{\partial I}{\partial x} u \right)^2 + \left(\frac{\partial I}{\partial y} v \right)^2 + \frac{\partial I}{\partial y} \frac{\partial I}{\partial x} vu \right), \end{aligned} \quad (2)$$

que escrita na forma matricial é a matriz de covariância que caracteriza a taxa de mudança de intensidade em todas as direções.

Essa definição envolve computar as primeiras derivadas $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$, que geralmente são calculadas utilizando filtros Sobel (*Sobel Derivatives*, 2015).

Pode ser mostrado que os dois auto-valores da matriz de covariância fornecem a maior mudança de intensidade média e a variação de intensidade média da direção ortogonal. A partir disso, se os auto-valores forem baixos, então a região é considerada homogênea. Se um dos valores é baixo e outro alto, é uma borda. E se os dois auto-valores forem altos, é uma quina. Então, o método de Harris considera borda cada ponto da imagem em que os auto-valores são maiores que um certo limiar.

2.3 SURF

O método SURF (*Speed-up Robust Features*) é um método de detecção de características de ima-

gens a se rastrear de forma que é invariante a escala e computacionalmente eficiente.

O método SURF computa uma matriz Hessiana para cada pixel. Essa matriz mede a curvatura local de uma função, e é definida por

$$H(x, y) = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial x \partial y} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} \quad (3)$$

O determinante desta matriz nos fornece a força da curvatura. Os pontos que consideramos quinas, neste caso, são os pontos de maiores curvaturas locais. Como a matriz é composta por derivadas de segunda ordem, a matriz pode ser calculada utilizando filtros Laplacianos Gaussianos com diferentes valores de desvio padrão (σ) da Gaussiana. Essa Hessiana então se torna uma função de 3 variáveis: $H(x, y, \sigma)$. A característica invariante a escala¹ é declarada quando o determinante da Hessiana alcança um máximo local, tanto espacial (nas variáveis (x, y)) como na escala (variável σ).

O cálculo dessas derivadas em diferentes escalas é computacionalmente custoso. Para tornar o processo computacionalmente mais eficiente, o método SURF utiliza filtros gaussianos aproximados que envolvem apenas poucas adições inteiras. A Figura 3(b) ilustra a estrutura de filtros que o SURF utiliza.

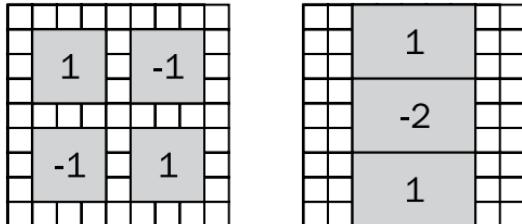


Figura 1. Representação dos filtros utilizados no cálculo da matriz Hessiana do método SURF. O da direita ilustra o termo $\frac{\partial^2 I}{\partial x \partial y}$, e o da direita ilustra o termo $\frac{\partial^2 I}{\partial y^2}$. O termo $\frac{\partial^2 I}{\partial x^2}$ é obtido pela rotação da máscara da direita em 90 graus.

Para diferentes valores de escala σ , o processo apenas aumenta o tamanho dos filtros, sendo que a menor escala é com filtros 9x9 e a maior de filtros 99x99. Note que o processo é realizado calculando somas de pixels de regiões da imagem. Portanto, cada filtragem em cada região pode ser obtida com apenas 3, se utilizarmos imagens integrais (Laganière, 2011).

¹Para simular a mudança de escala utilizando filtragem, é alterado o desvio padrão dos filtros gaussianos, e isso altera o tamanho dele, por isso o valor de escala é proporcional ao desvio padrão dos filtros utilizados.

3 Materiais e Metodologia empregada

Os materiais foram um notebook com uma câmera acoplada e um livro de origami para ser reconhecido.

O método Harris foi implementado utilizando a classe GoodFeaturesToTrackDetector do OpenCV, com a opção de utilizar o método de Harris para detecção. O método SURF foi implementado utilizando a classe SurfFeatureDetector do OpenCV.

O procedimento (desenvolvido no arquivo Projeto6.cpp) consiste em obter uma imagem de captura da câmera do notebook, em seguida o usuário deve apertar em dois pontos na tela para definir a região de interesse correspondente ao objeto que se quer reconhecer. Neste caso, o livro de origami foi colocado a mostra na frente da câmera, e o usuário selecionou os pontos correspondentes aos limites do livro.

Em seguida, o programa abre duas janelas, uma para cada método. Em cada janela, é mostrado o quadro atual do vídeo e ao lado o objeto de interesse, assim como os pontos de interesse detectados por cada método, tanto no objeto a ser reconhecer como no quadro do vídeo. Além disso, é traçada uma linha entre os pontos correspondentes no objeto a ser reconhecido e os quadros capturados, identificando a posição do objeto no vídeo.

A avaliação dos resultados é feita de forma subjetiva, comparando visualmente os pontos que fazem parte do template e do objeto no vídeo. Para uma avaliação mais justa, foram escolhidos o número máximo de 25 matches entre os pontos do objeto e os pontos correspondentes no vídeo. Como os métodos nem sempre retornam o mesmo número de pontos, foram escolhidos os 25 ponto de maior relevância em ambos os métodos, desta forma é mais fácil comparar visualmente os resultados.

Há também uma avaliação quanto ao tempo de execução de cada método. Para calculá-lo, é medido o tempo que se passou no mundo real durante o processo de detecção e de extração das características de cada método e esses tempos são comparados.

4 Resultados

Os primeiros resultados estão ilustrados na Figura 2, tanto para o método Harris quanto para o SURF.

Para ambos os métodos, os pontos reconhecidos foram em sua maioria correspondentes as letras que formam a palavra 'origami' do livro. E aparentemente todos os pontos foram corretamente identificados nesta primeira posição do objeto. A partir disto, fizemos testes em que inclinamos o livro e alteramos a escala, movendo o livro



Figura 2. Segmentação por limiarização. Esquerda: imagens originais. Meio: Segmentação utilizando o limiar na intensidade da imagem. Direita: Segmentação utilizando o limiar na cromaticidade da imagem.

para uma posição mais distante da câmera. E os dois tipos ao mesmo tempo.

Como esperado, o SURF mostrou melhor robustez a variação de escala, enquanto o Harris logo se perdia na mudança de escala. Entretanto, quando alterava-se muito a escala, o SURF também se perdia. Quanto a rotação, ambos apresentavam uma robustez razoável, com leve vantagem para o SURF. A Figura 3 ilustra um quadro do vídeo capturado que mostra a rotação e mudança de escala ao mesmo tempo. Note que o método SURF apresentou resultado bastante superior ao método Harris.

Os tempos de execução no notebook em que foi realizado o experimento para esta configuração, fora, em média 12,1 milissegundos para o método Harris e 65,0 milissegundos para o método SURF. Essa diferença provavelmente acontece porque o método SURF faz filtragens para vários tamanhos

de escala.

5 Discussão e Conclusões

Neste laboratório utilizamos 2 métodos de reconhecimento de padrões em imagens. O método Harris utiliza o conceito de encontro de bordas, dada por grandes variações de intensidade em direções ortogonais para definir pontos de interesse. Já o método SURF calcula a matriz Hessiana que mede a curvatura local do ponto de interesse, o método é executado várias vezes para cada escala diferente, e utiliza o conceito de imagens integrais para o rápido cálculo das filtragens.

Os critérios avaliados foram a robustez a variação de escala, robustez a rotação e tempo de execução. O método SURF mostrou melhores resultados em relação a robustez a variação de escala, isso devido ao método SURF fazer filtragens



Figura 3. Segmentação por limiarização. Esquerda: imagens originais. Meio: Segmentação utilizando o limiar na intensidade da imagem. Direita: Segmentação utilizando o limiar na cromaticidade da imagem.

em vários fatores de escala. Quanto a rotação, nenhum dos métodos se sobressaiu significantemente, segundo minha avaliação subjetiva visual. Quanto ao tempo de execução, o método Harris foi mais eficiente. Isso é devido ao método SURF efetuar o processamento em vários fatores de escala.

Referências

Harris, C. and Stephens, M. (1988). A combined corner and edge detector., *Alvey vision conference*, Vol. 15, Manchester, UK, p. 50.

Laganière, R. (2011). *OpenCV 2 Computer Vision Application Programming Cookbook: Over 50 recipes to master this library of programming functions for real-time computer vision*, Packt Publishing Ltd.

Shi, J. and Tomasi, C. (1994). Good features to track, *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, IEEE, pp. 593–600.

Sobel Derivatives (2015). http://docs.opencv.org/doc/tutorials/imgproc/imgtrans/sobel_derivatives/sobel_derivatives.html. Acessado em Maio, 2015.