

Reconhecimento de padrões

Bruno Fernandes Carvalho - 15/0007159

Dep. Ciência da Computação - Universidade de Brasília (UnB)

Princípios de Visão Computacional - Turma A

Data de realização: 11/06/2017

brunofcarvalho1996@gmail.com

Abstract

Pattern and objects recognition are one of the main applications in computer vision and it can be done from key points relations detected in a scene. In this project it is compared two methods (Harris and SURF) that detects these key points, as the computational cost and the robustness of each one to find objects in a video.

Abstract

O reconhecimento de padrões e objetos é uma das principais aplicações em visão computacional e ela pode ser feita a partir da comparação de pontos chaves detectados numa cena. Nesse projeto é feita uma comparação entre dois métodos (Harris e SURF) para detecção desses pontos, assim como o custo computacional e a robustez de cada um para achar objetos em um vídeo.

1. Objetivos

Essa atividade tem como base entender e implementar algoritmos que permitam realizar reconhecimento de objetos em vídeo a partir de pontos chaves detectados.

2. Introdução

O conceito de pontos chaves de uma imagem é útil para resolver diversos problemas em visão computacional, dentre eles o reconhecimento de objetos a partir da comparação desses pontos entre duas imagens. Isso quer dizer que determinado objeto deve ter uma coleção específica de pontos chave que o caracterizam, e a partir de um modelo previamente obtido, pode-se detectar o objeto em um vídeo, conforme será feito nesse projeto.

Esses pontos chaves na maioria das vezes são os cantos da imagem, e existem diversas técnicas para detectá-los. Será apresentado dois métodos: Harris e SURF.

O primeiro é menos sofisticado, já que não é invariante

à mudança de escala do objeto na cena. Seu funcionamento consiste em basicamente procurar por pixels que tenham a derivada alta em uma direção e em uma direção perpendicular a esta, caracterizando um canto na imagem. Um exemplo desse detector pode ser observado na imagem 1.

Figura 1. Pontos chaves usando detector Harris



Já o segundo apresenta um método consistente e que consegue identificar pontos chaves mesmo que haja variação na escala em relação ao modelo previamente obtido. O método SURF (Speeded-Up Robust Features) é derivado do SIFT (Scale-Invariant Feature Transform), representando uma versão mais rápida e robusta, sem perder a propriedade de ser invariante à escala. Além disso, após a detecção dos cantos, é possível usar um descritor que irá tirar informações do ponto chave e de seus vizinhos, caracterizando melhor a região para posterior comparação e reconhecimento de padrões. Um exemplo desse método é mostrado na figura 2

Figura 2. Pontos chaves usando detector SURF



3. Materiais e Metodologia

A ideia desse projeto é comparar os dois métodos explicados na seção anterior, com o objetivo de verificar custo computacional e robustez em reconhecimento de objetos. Para isso, durante o streaming de vídeo, o usuário deve clicar com o mouse para escolher o objeto que vai ser reconhecido. Isso deve ser feito a partir de dois cliques que englobem o objeto. Com o modelo capturado e seus pontos chaves detectados a partir dos dois métodos, o programa começará a procurar os pontos correspondentes na sequência de imagens e irá ligar aqueles que forem semelhantes. No final do processamento, é mostrado ao usuário a ligação desses pontos entre o modelo e o frame atual para o dois métodos, sendo possível comparar aquele que melhor detectou o objeto.

Além disso, para comparar o custo computacional, a quantidade de pontos chaves utilizadas e a quantidade de ligações que foram feitas entre o modelo e o frame atual, foi feita a variação destes pontos chaves de 1 até 1000 (com passos incrementais de 10 em 10) e foi salvo num arquivo txt os resultados de ambos detectores. Vale deixar claro que as vezes os detectores não foram capazes de achar 1000 pontos na imagem, então nesses casos o arquivo txt irá conter os resultados para o máximo de pontos achados. Ainda, para facilitar a comparação dos resultados, foram criados dois arquivos txt, um para cada método.

Para esse projeto foi utilizado a câmera do computador para capturar a imagem de modelo e também para realizar a dinâmica de reconhecimento de objetos mencionada acima. Foi utilizado o próprio rosto como objeto para comparação dos métodos.

Vale deixar claro que, para o desenvolvimento dos algoritmos, foi utilizado a linguagem de programação C++, assim como técnicas de orientação a objetos. É importante lembrar que o usuário deve definir o objeto modelo a par-

tir de dois cliques com o mouse e que ele pode alterar a qualquer momento o template capturando um novo objeto na tela de vídeo.

4. Resultados

Conforme dito anteriormente, foi usado como modelo o próprio rosto para comparar a correspondência de pontos chaves achados no vídeo. Abaixo é possível ver nas figuras 3 e 4 ligação desse pontos.

Figura 3. Correspondência entre os pontos chaves usando Harris

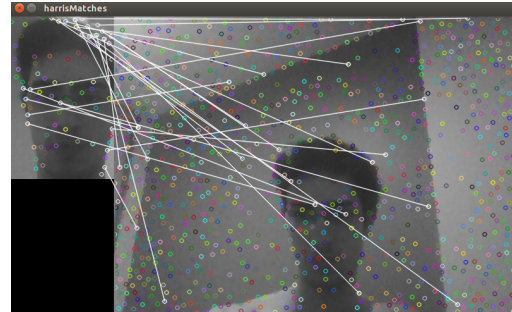
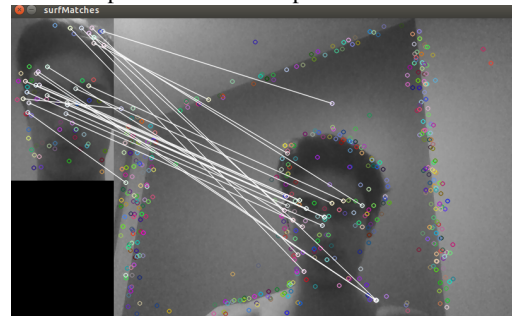


Figura 4. Correspondência entre os pontos chaves usando SURF



Observando as duas figuras, vê-se claramente que o método SURF foi superior ao método Harris, já que houve muito mais pontos correspondentes do objeto de modelo. Isso ocorre pois SURF é muito mais robusto e consegue caracterizar melhor os pontos chaves detectados e seus vizinhos, mesmo que estes tenham suas escalas alteradas. Ainda, Harris detectou mais pontos que SURF, mesmo que eles não fossem cantos da cena, gerando confusão na correspondência dos pontos que realmente pertenciam ao objeto.

Analisando os resultados apresentados nos dois arquivos txt gerados, observa-se que SURF, apesar de sua robustez e melhor precisão no reconhecimento de objetos, ainda apresentou menor tempo de processamento em relação à Harris. Foi comparado também o tempo de processamento com a quantidade de pontos chaves analisados, chegando à conclusão esperada: quanto mais pontos utilizados, maior o custo computacional.

Em relação à quantidade de pontos correspondidos entre o modelo e o frame atual do vídeo, observa-se que esse valor

é saturado rapidamente em ambos os casos com o aumento de pontos chaves utilizados, já que os detectores começam a confundir outros objetos da cena com o objeto desejado, não conseguindo melhorar esse valor. Ainda, como Harris foi capaz de detectar mais pontos na imagem, ele apresentou um número maior de correspondências.

5. Conclusão

Ficou claro como pode ser feito o reconhecimento de objetos a partir da detecção de pontos chaves, assim como a comparação dos dois métodos apresentados. Explicitamente o método SURF apresentou resultados satisfatórios, tanto computacionalmente quando na correspondência dos pontos entre o modelo e o vídeo. Já Harris não conseguir realizar a detecção do objeto na cena devido à grande quantidade de pontos indevidos detectados, gerando confusão na hora das correspondências.

Referências

- [1] G. Bradski and A. Kaehler, *Learning OpenCV*, 1st ed. O'Reilly.
- [2] Forsyth and Ponce, *Computer Vision - A Modern Approach*, 1st ed. Pearson.
- [3] OpenCV API Reference. (2017, 31 Março). [Online]. Available: <http://docs.opencv.org/2.4/modules/refman.html>
- [4] R. Laganière, *OpenCV 2 Computer Vision Application Programming Cookbook*, 1st ed. Packt Pub Limited, 2011.

[1] [2] [3] [4]