FGV - Fundação Getulio Vargas

Visão Computacional - Lista 5

Professor: Moacyr Alvim Horta Barbosa da Silva
Monitor: Tulio Koneçny

A lista deverá ser entregue no formato .pdf e o código em py. ou .ipynb
Entrega: 07/06/24

1 SIFT (Scale Invariant Feature Transform)

Ao final do exercício você deve costurar duas (ou mais) imagens obtidas por fotografias de um panorama com a posição da câmera no mundo um pouco deslocada (um pequeno giro da câmera, por exemplo). Você pode usar as imagens fornecidas como anexo a este exercício ou imagens que você mesmo capturar/escolher.

1. Use o SIFT para detectar os pontos de interesse e seus descritores nas duas imagens.

2. Exiba os pontos encontrados com círculos cujo tamanho indique a escala em uma das imagens. Sugestão: você pode usar a função cv2.drawKeypoints.

3. Escolha um ponto de interesse encontrado pelo SIFT em uma das imagens e exiba as informações deste ponto de interesse:

   (a) pt: Tupla contendo as coordenadas (x, y) do ponto de interesse.

   (b) size: Valor que representa a escala do ponto.

   (c) angle: Ângulo de orientação do ponto.

(d) response: Valor de resposta (relacionado à "qualidade" do ponto).

(e) octave: Número da oitava onde foi encontrado.

4. Determine a correspondência entre os pontos encontrados. Sugestão: ao invés por força bruta (todos os pontos contra todos os pontos) você pode usar cv2.FlannBasedMatcher.

   flann = cv2.FlannBasedMatcher(index_params, search_params) correspondencias = flann.knnMatch(descritores1, descritores2, k=2) Este trecho do código encontra os melhores pares de forma eficiente.

5. Estime a transformação projetiva H que leva pontos de destaque de uma imagem em seus correspondentes na outra imagem, usando o método RANSAC.

   Sugestão: você pode usar a função cv2.findHomography(pontos2, pontos1, cv2.RANSAC, 5.0), que já usa o método RANSAC para encontrar a transformação projetiva que melhor um conjunto de pontos alvo ao conjunto de pontos origem.

6. Transforme uma das imagens segundo a transformação projetiva H e costure as imagens para gerar uma imagem única.

   Sugestão: use imagem_unida = cv2.warpPerspective(imagem2, matriz_homografia, (imagem1.shape[1] + imagem2.shape[1], imagem1.shape[0])) imagem_unida[0:imagem1.shape[0], 0:imagem1.shape[1]] = imagem1 para transformar a imagem (warping) e costurar duas imagens.

2 Exercício teórico

Nesta parte vamos considerar a imagem em tons de cinza como uma função $I : \mathbb{R}^2 \to \mathbb{R}$ suave (isto é, infinitamente diferenciável). O espaço de escala é uma família de imagens $\hat{I}(t) = L(I, t)$, $t \leq 0$ onde a primeira imagem da família é a própria imagem original, ou seja, $\hat{I}(0) = L(I, 0) = I$.

A família de imagens é obtida por convolução com gaussianas de variâncias iguais à escala. Logo, $\hat{I}(x, y, t) = (g(t) \circledast I)(x, y)$. Esta convolução, como sabemos, é calculada pela integral:

$$\hat{I}(x, y, t) = \int\int_{R^2} g(x, y, t) \cdot I(s - x, r - y)\, ds\, dr$$

onde $g(x,y,t)$ é a função gaussiana com média $(0,0)$ com variância t:

$$g(x,y,t) = \frac{1}{2\pi t} e^{-\frac{x^2+y^2}{2t}}$$

Isto é equivalente a dizer que a família de imagens $\hat{I}(t)$ é a solução da equação do calor com valor inicial:

$$\hat{I}_t(x,y,t) = \hat{I}_{xx}(x,y,t) + \hat{I}_{yy}(x,y,t)$$

$$\hat{I}(x,y,0) = I(x,y)$$

Uma abordagem alternativa para os espaços de escala é axiomática. São descritas propriedades desejáveis que os espaços de escala devem possuir se conclui que a transformação $L(I,t)$ que satisfaz estas propriedades é a convolução com a gaussiana. Algumas das propriedades são:

1. Linearidade: $a \cdot \hat{I}_1(t) + b \cdot \hat{I}_2(t) = L(aI_1 + bI_2)$
2. Invariância por translação: $\hat{I}(x - \Delta x, y - \Delta y, t) = L(I(x - \Delta x, y - \Delta y), t)$
3. Propriedade de semigrupo: $L(I, t + s) = L(L(I,t), s)$
4. Não ampliação de máximos (e mínimos) espaciais:
   (a) $\partial_t \hat{I}(x_*, y_*, t) \leq 0$ se (x_*, y_*) for um ponto de máximo.
   (b) $\partial_t \hat{I}(x_*, y_*, t) \geq 0$ se (x_*, y_*) for um ponto de mínimo.
5. Invariante por rotação: o espaço de escala de uma imagem rotacionada é a rotação do espaço de escala da imagem original.

(Uma lista completa, com 9 axiomas, pode ser encontrada no artigo Lindeberg, T. Invariance of visual operations at the level of receptive fields, PLoS ONE, 2013 – ou no Wikipedia, claro!)

Neste exercício você deve mostrar que fazendo $L(I,t) = g(t) \circledast I$, onde $g(t)$ é a gaussiana com variância t, as 5 propriedades listadas acima são satisfeitas.

3 Extra

Utilizando as funções criadas nesta lista, faça uma função *find_wally* que receba duas imagens, uma referente a uma imagem do livro *Onde está Wally* (cenario.png), e a segunda imagem referente ao próprio Wally (wally.png). Sua função deve retornar a imagem do cenário com a imagem do Wally destacada (basta mudar as cores de wally.png e aplicar uma transformação H que leva a imagem em cenario_wally).

Perceba que a imagem dada de Wally NÃO é exatamente a imagem da própria figura (diferente da lista anterior). Em anexo possuem três versões do Wally, é a mesma imagem porém com dimensões diferente. Veja se o método funciona para as três. Comente os resultados.