Extração Automática de Características

Prof. Jefferson T. Oliva
jeffersonoliva@utfpr.edu.br

Processamento de Imagens
Engenharia de Computação
Departamento Acadêmico de Informática (Dainf)
Universidade Tecnológica Federal do Paraná (UTFPR)
Campus Pato Branco

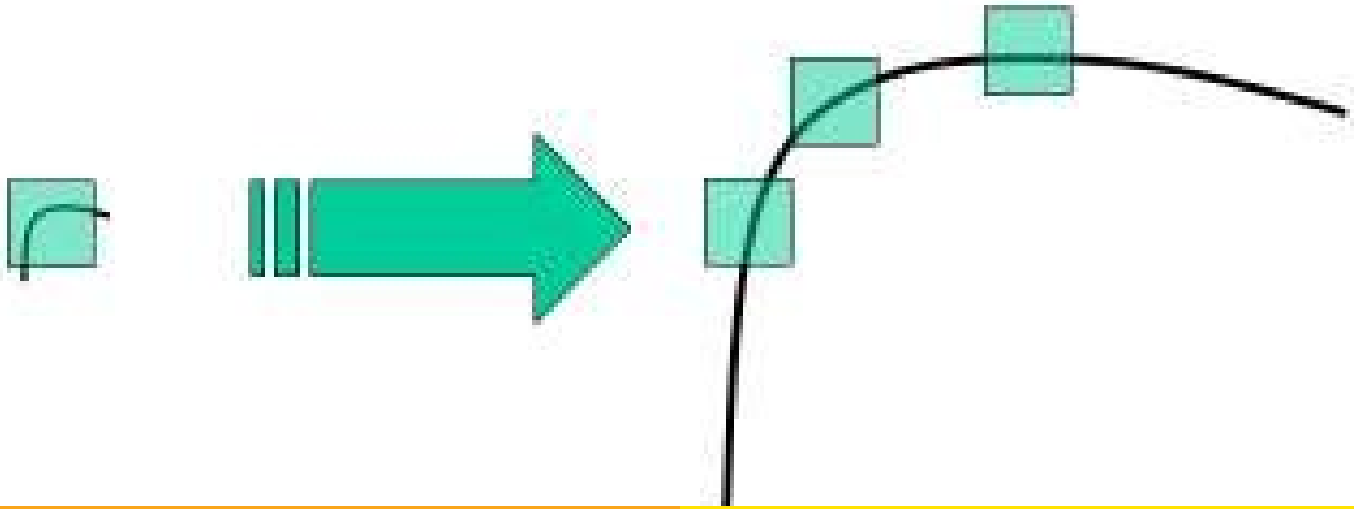
Sumário

- Descritores SIFT

Sumário

Descritores SIFT

- SIFT: *Scale Invariant Feature Transform*
- Extração automática de características
- São invariantes à rotação, o que significa que, mesmo que a imagem seja girada, podemos encontrar os mesmos cantos
  - Os cantos também permanecem cantos na imagem rotacionada
- Mas e a escala? Um canto pode não ser um canto se a imagem for dimensionada



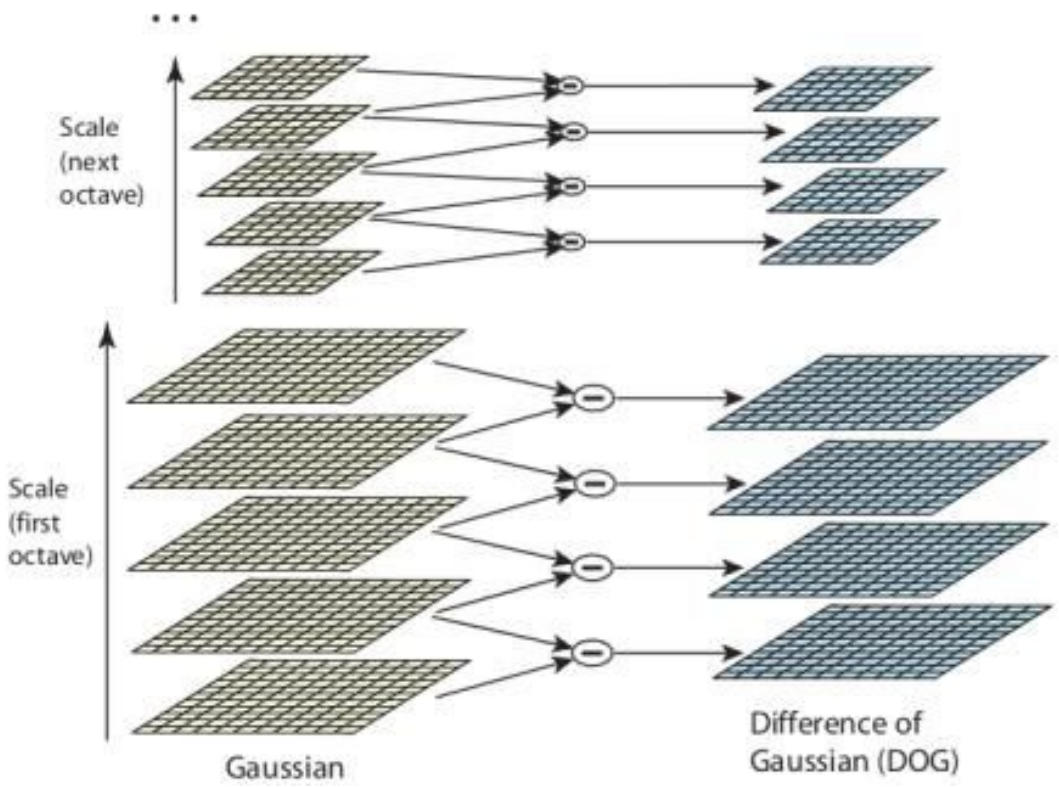
Descritores SIFT

- Os descritores SIFT são pontos-chave que são invariantes em escala
  - Extraem pontos-chave e calculam seus descritores
- É utilizada a filtragem por espaço de escala, baseado em "Laplace of Gaussian" (LoG) com vários valores de σ
- Assim, podemos encontrar os máximos locais na escala e no espaço, o que nos dá uma lista de valores (x, y, σ), o que significa que existe um ponto-chave potencial em (x, y) na escala σ
- Mas o LoG é caro computacionalmente e, dessa forma, o algoritmo SIFT usa diferença de Gaussianas (DoG – *difference of Gaussians*), que é uma aproximação de LoG

Descritores SIFT

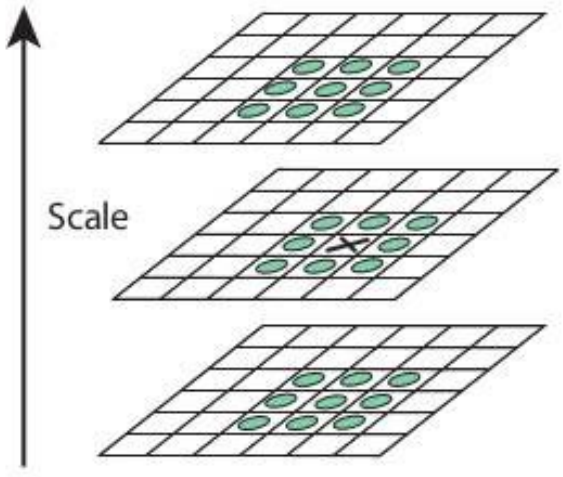
- O processo de diferença de Gaussianas é feito para diferentes oitavas da imagem na pirâmide Gaussiana

- A diferença de gaussiana é obtida como a diferença de desfoque gaussiano de uma imagem com dois σ diferentes, sejam σ e $k\sigma$



Descritores SIFT

- Uma vez que a diferença de Gaussianas é encontrada, as imagens são procuradas por extremos locais sobre escala e espaço
  - Por exemplo, um pixel em uma imagem é comparado com seus 8 vizinhos, bem como 9 pixeis na próxima escala e 9 pixeis nas escalas anteriores
  - Se for um extremo local, é um ponto-chave potencial

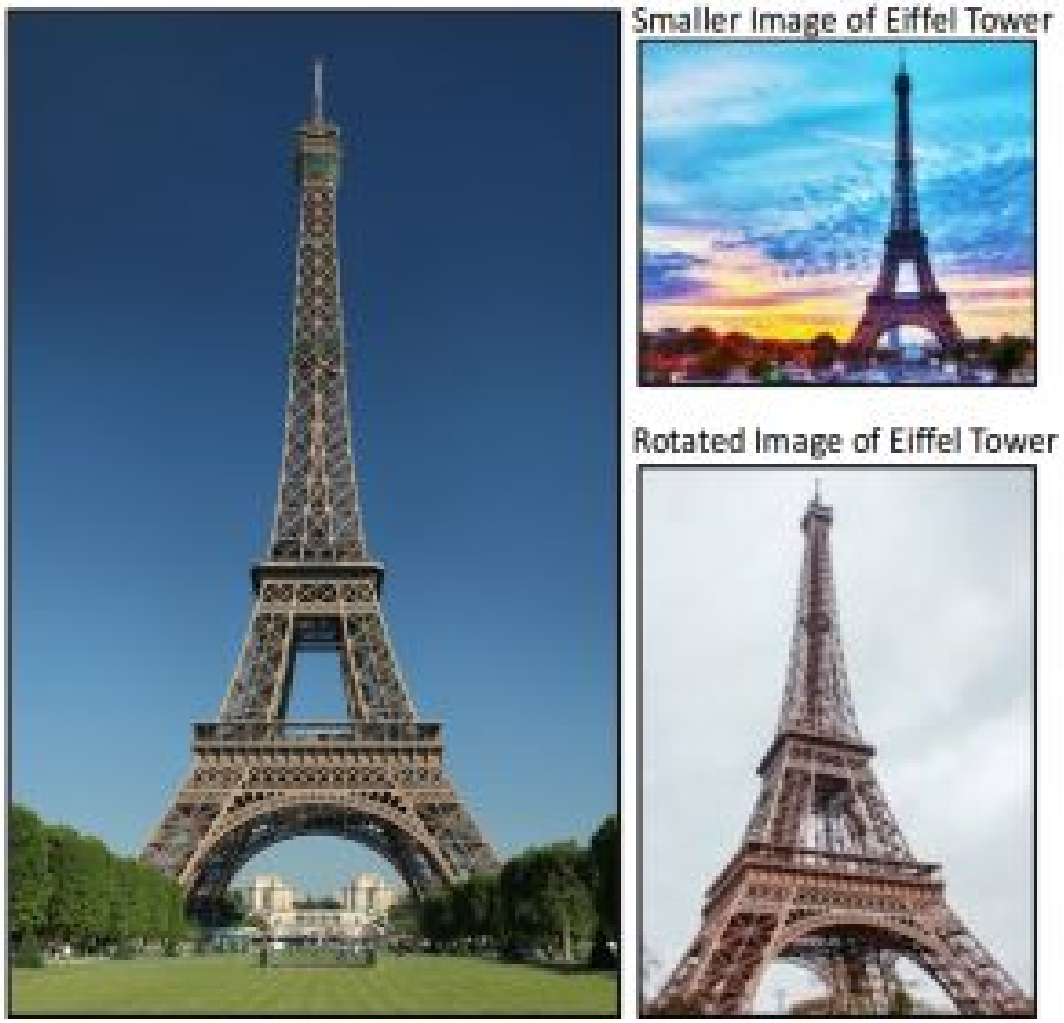


Descritores SIFT

- Em relação a diferentes parâmetros, são indicados
  - Número de oitavas = 4
  - Número de níveis de escala = 5
  - σ inicial = 1.6
  - k = $\sqrt{2}$
- Agora o descritor de ponto-chave é criado
  - Uma vizinhança de 16x16 ao redor do ponto-chave é tomada
  - Divisão em 16 sub-blocos de tamanho 4x4
  - Para cada sub-bloco, um histograma de orientação de 8 bins é criado
  - Assim, um total de 128 valores bin estão disponíveis
  - Representação como um vetor para formar o descritor de ponto-chave
  - Além disso, várias medidas são tomadas para obter robustez contra mudanças de iluminação, rotação etc

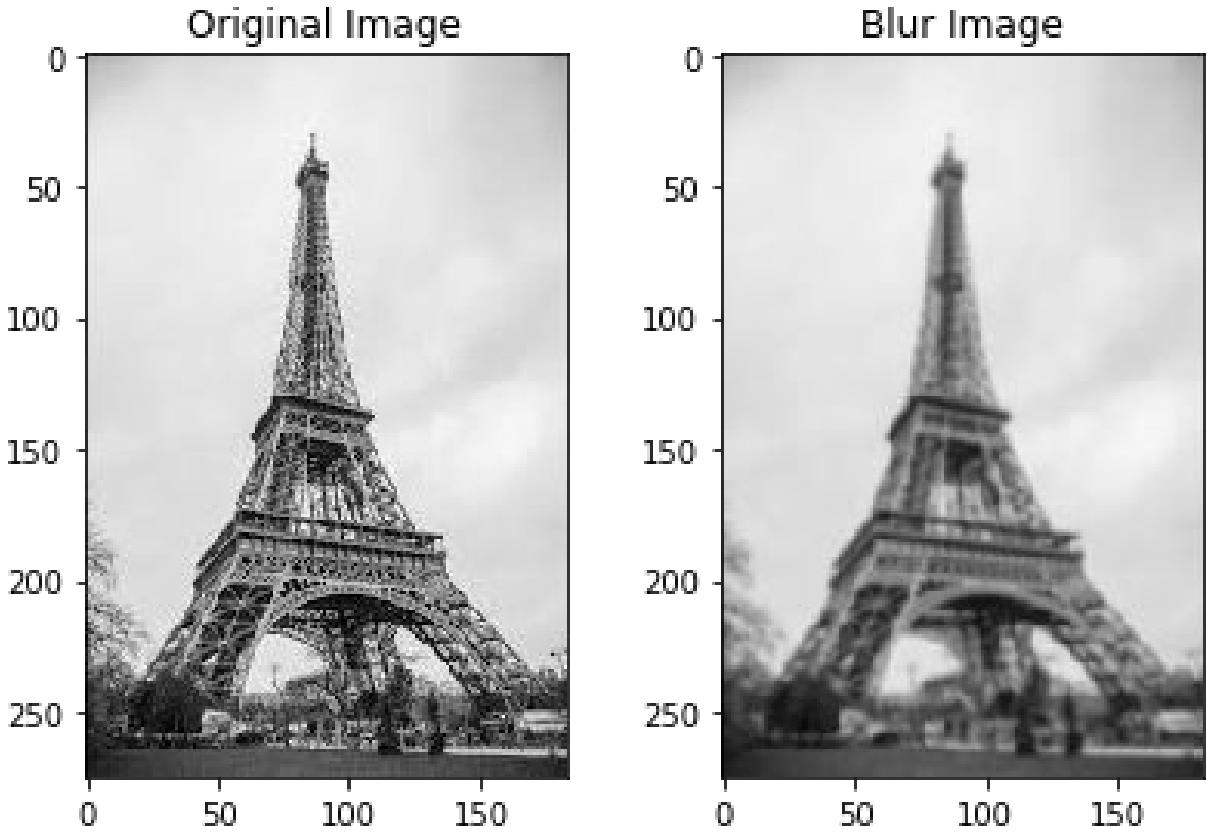
Descritores SIFT

- Exemplo abaixo: em comum temos a torre Eiffel



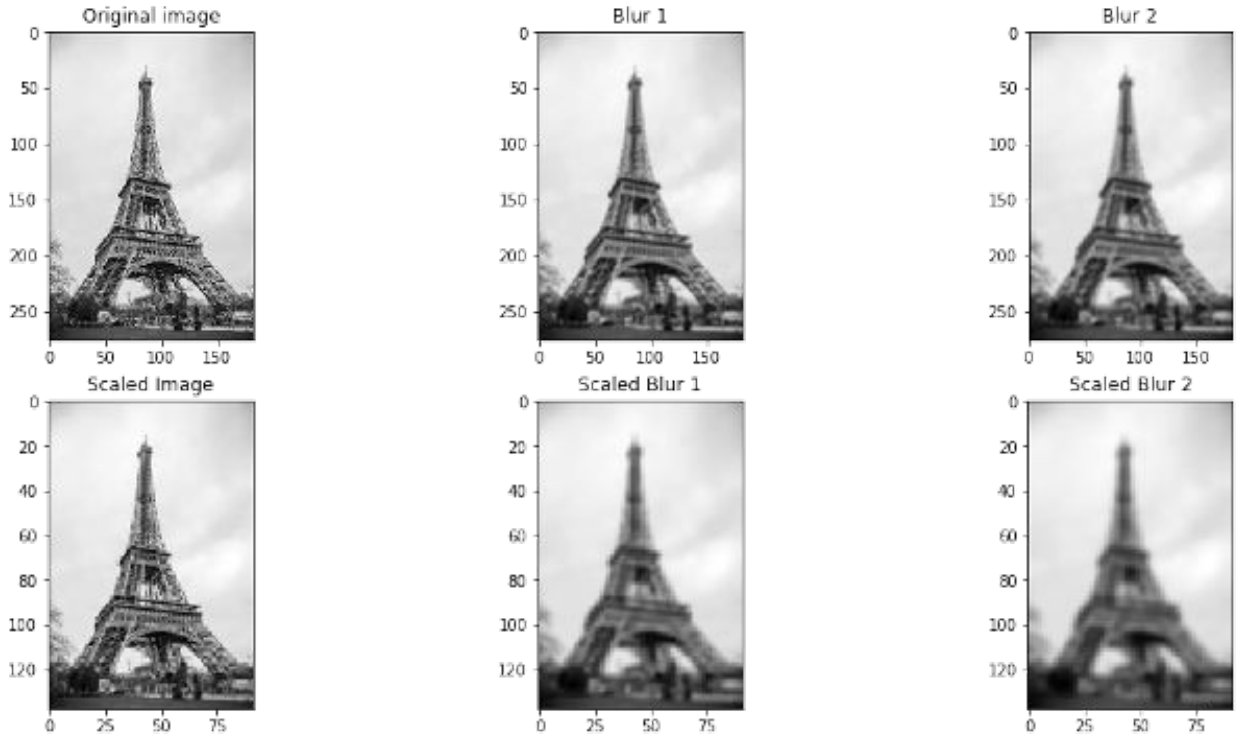
Descritores SIFT

- Construindo o filtro Gaussiano



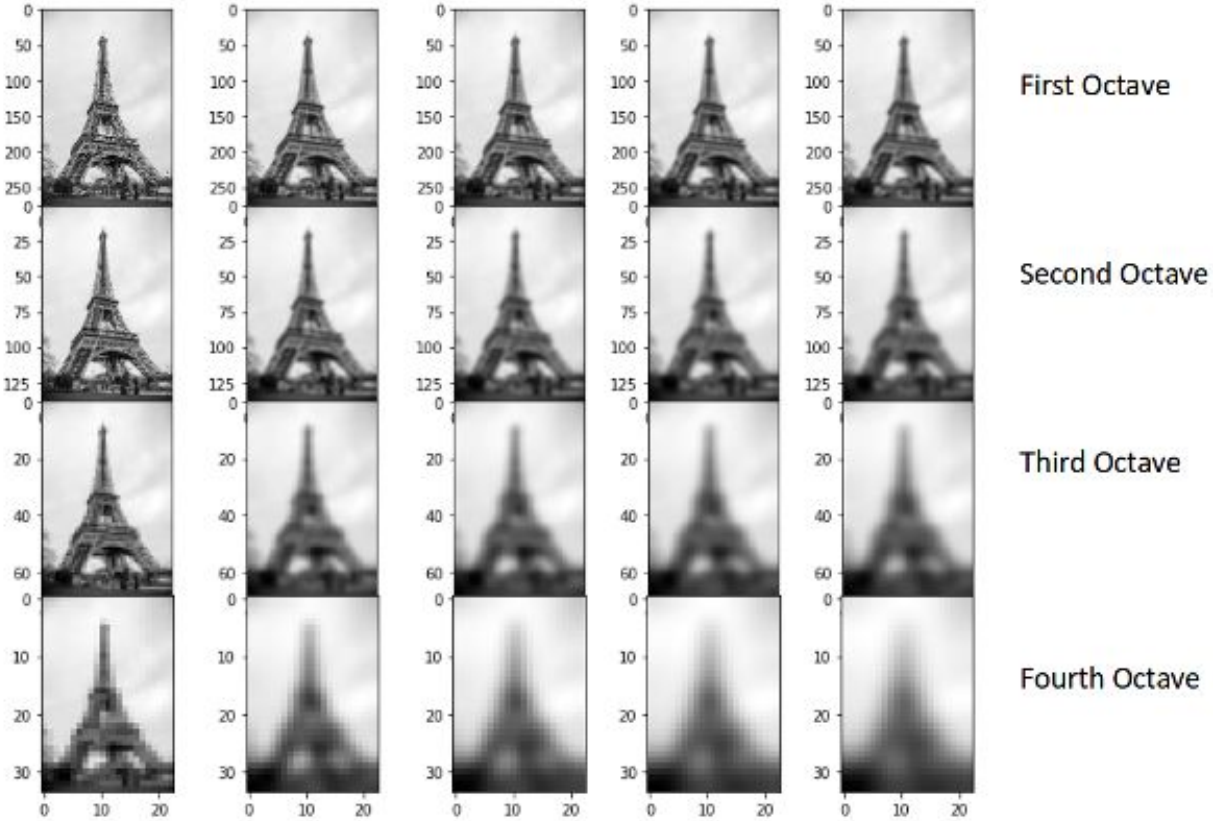
Descritores SIFT

- Construindo o filtro Gaussiano



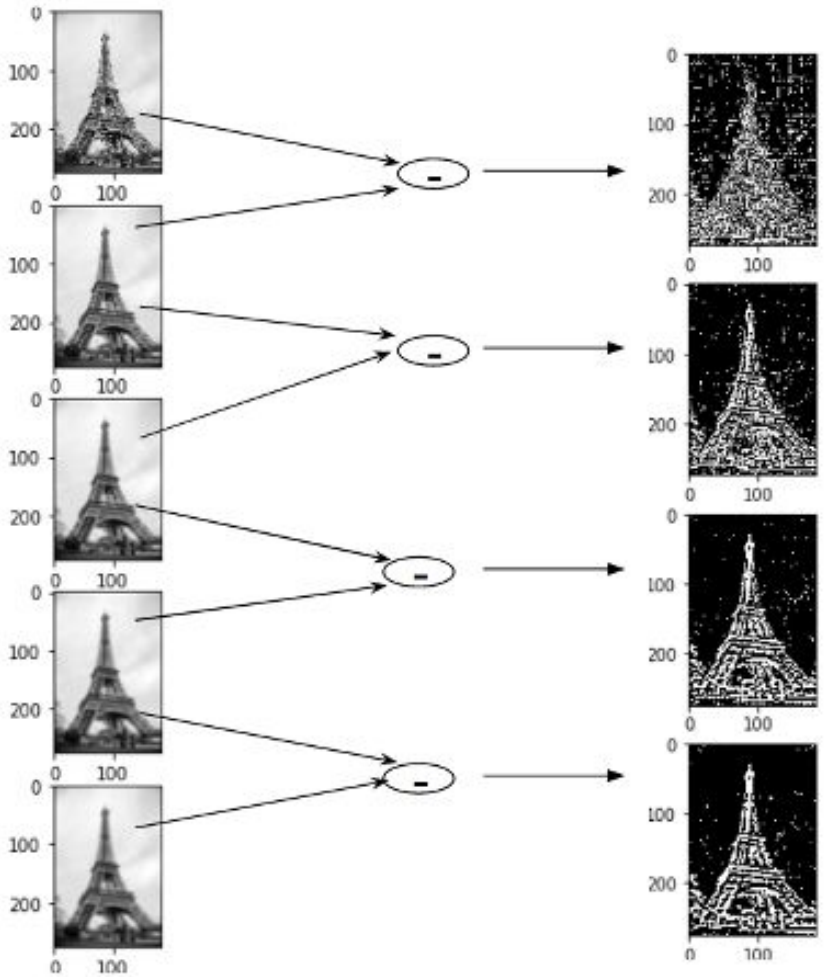
Descritores SIFT

- Construindo o filtro Gaussiano: 4 oitavas e 5 vezes o filtro
  - As imagens foram divididas pela metade a partir da original
  - Para cada imagem filtrada, novas versões foram criadas



Descritores SIFT

- Diferença Gaussiana: algoritmo de aprimoramento de características que envolve a subtração de uma versão desfocada de uma imagem original de outra versão menos desfocada do original

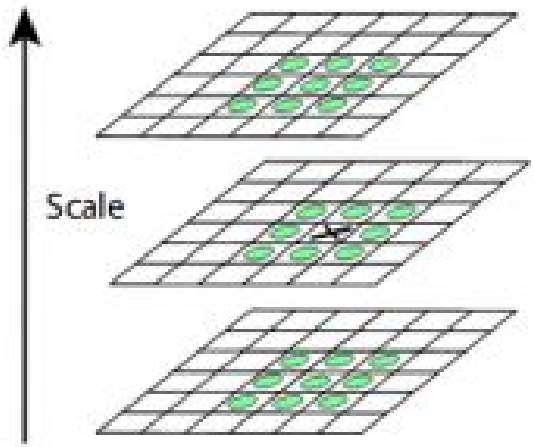


Descritores SIFT

- Localização dos pontos-chave
  - Depois que as imagens forem criadas, a próxima etapa é encontrar os pontos-chave importantes da imagem que podem ser usados para correspondência de características
  - A ideia é encontrar os máximos e mínimos locais para as imagens em duas etapas:
    - Encontre os máximos e mínimos locais: para isso, percorremos cada pixel da imagem e os comparamos com os pixeis vizinhos
    - Remover pontos-chave de baixo contraste (seleção de ponto-chave)
  - Isso significa que cada valor de pixel é comparado com 26 outros valores de pixel para descobrir se é o máximo/mínimo local chamado extremo

Descritores SIFT

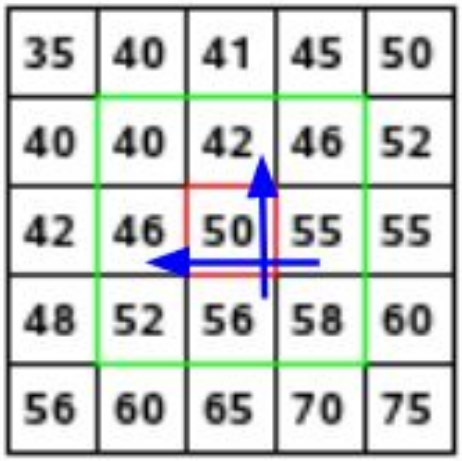
- O pixel marcado com x é comparado com os pixeis vizinhos (em verde) e é selecionado como ponto-chave ou ponto de interesse se for o maior ou o menor entre os vizinhos



Descritores SIFT

- Alguns pontos-chave podem não ser robustos ao ruído, sendo necessária uma verificação final para garantir que temos os pontos-chave mais precisos para representar as características da imagem
  - Para lidar com os pontos-chave de baixo contraste, uma expansão de Taylor de segunda ordem é calculada para cada ponto-chave
  - Se o valor resultante for menor que 0,03 (em magnitude), rejeitamos o ponto-chave
- Atribuição de orientação: nesta fase, temos um conjunto de pontos-chave estáveis para as imagens
  - Atribuiremos agora uma orientação a cada um desses pontos-chave para que sejam invariantes à rotação
  - Podemos novamente dividir esta etapa em duas etapas menores: cálculo do módulo e da orientação

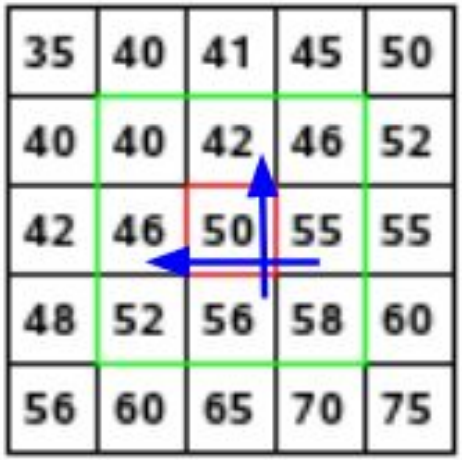
- Atribuição de orientação:
  - Digamos que queremos encontrar a magnitude e a orientação do valor do pixel em vermelho
    - Para isso, calcularemos os gradientes nas direções x e y tomando a diferença entre 55 e 46 e 56 e 42, resultando em $G_x = 9$ e $G_y = 9$, respectivamente



- Atribuição de orientação:
  - Digamos que queremos encontrar a magnitude e a orientação do valor do pixel em vermelho
    - Uma vez que temos os gradientes, podemos encontrar a magnitude (M) e a orientação (Φ) usando as seguintes equações:

$$M = \sqrt{G_x^2 + G_y^2} = 16,64$$

$$\Phi = \arctan \frac{G_y}{G_x} = 57,17$$

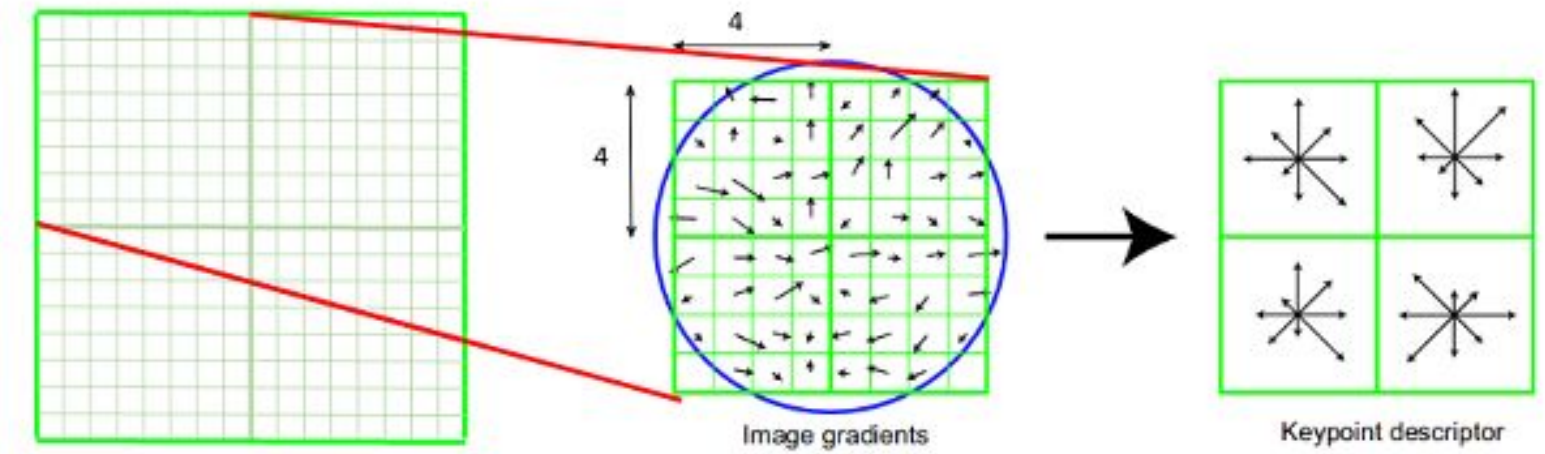


Descritores SIFT

- Descritores:
  - Etapa final do SIFT: temos pontos-chave estáveis que são invariantes à escala e à rotação
  - São usados os pixeis vizinhos, suas orientações e sua magnitude para gerar uma 'impressão digital' única para este ponto-chave chamado de 'descritor'
  - Além disso, como usamos os pixeis circundantes, os descritores serão parcialmente invariantes à iluminação ou brilho das imagens
  - Primeiro, é utiliza uma vizinhança de 16×16 ao redor do ponto-chave
    - Este bloco de 16×16 é dividido em sub-blocos de 4×4 e para cada um desses sub-blocos, geramos o histograma usando magnitude e orientação

Descritores SIFT

- Descritores



Descritores SIFT

- Exemplo OpenCV

```
import cv2
import matplotlib.pyplot as plt
```

```
img1 = cv2.imread('eiffel_2.jpeg')
gray1 = cv2.cvtColor(img1, cv2.COLOR_BGR2GRAY)
```

```
sift = cv2.xfeatures2d.SIFT_create()
keypoints_1, descriptors_1 =
sift.detectAndCompute(img1,None)
img_1 = cv2.drawKeypoints(gray1,keypoints_1,img1)
plt.imshow(img_1)
```

Descritores SIFT

- SIFT é um algoritmo lento
- A alternativa é usar o algoritmo SURF, contudo, ele possui menor acurácia
- Melhor resultado em imagens de escala de cinza
- SIFT é um algoritmo patenteado, o que significa que seu uso é restrito em alguns contextos
  - Em contraste, o SURF não é patenteado e pode ser usado livremente

Referências I


Gonzalez, R. C.; Woods R. E.
Processamento de imagens digitais.
Editora Blucher, 2000.


Piemontez R. A.
Detecção de pontos faciais (Facemark) com OpenCV.
Disponível em: `https://visaocomputacional.com.br/deteccao-de-pontos-faciais-facemark-com-opencv/`, 202024.


Prateek, J., Millan, E. D., e Vinivius, G.
OpenCV by Example.
Packt Publishing, 2016.


Villán, A. F.
Mastering OpenCV 4 with Python.
Packt Publishing, 2019.

Referências II


Wiggers, K. L.
Notas de aula – Processamento de Imagens: segmentação de imagens.
UTFPR. Pato Branco, PR, 2024.