

## Projeto 2

### Tema: Localização de objeto com variação de escala.

#### Alunos:

Leonardo Donderi Rodrigues, RA: 754756

Luís Felipe Corrêa Ortolan, RA: 759375

Marco Antonio Bernardi Grivol, RA: 758619

Reynold Navarro Mazo, RA: 756188

Neste trabalho foi implementado a localização de um objeto na pirâmide gaussiana de uma imagem, ou seja, calcular as diferenças quadráticas entre cada nível da pirâmide e a imagem, visando identificar o menor valor de diferença em todos os níveis. O intuito em utilizar a pirâmide é cobrir tamanhos (resoluções) diferentes de imagens para buscar o objeto, já que podemos estar lidando com um objeto menor do que o apresentado na imagem original.

A técnica da diferença quadrática consiste em, para cada pixel da imagem ( $r$ ) e do objeto ( $c$ ), encontrar a distância entre eles, que espera-se ser mínima na região em que o objeto está localizado na imagem.

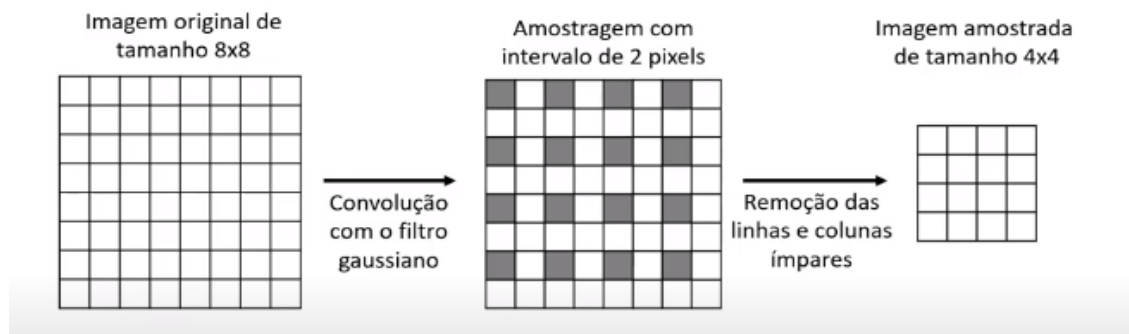
$$d(r, c) = \sum_{s=0}^{R-1} \sum_{t=0}^{C-1} [I_g(r + s - R/2, c + t - C/2) - I_o(s, t)]^2$$

Para a implementação, também pode ser utilizado a correlação-cruzada para encontrar a distância quadrática, por meio da correlação cruzada entre a imagem (cena com todos os valores ao quadrado) e um filtro  $w$  de mesmo tamanho que o objeto, com valor 1 em todas as posições, somando o resultado disso a soma de todos os valores do objeto ao quadrado e subtraindo o dobro da correlação cruzada entre a imagem e o objeto.

$$d(r, c) = I_g(r, c)^2 \circ w + S_{I_o} - 2I_g(r, c) \circ I_o$$

O uso da pirâmide gaussiana se dá quando o objeto é menor do que sua representação na cena, e há a necessidade de produzir imagens menores a fim de

“encaixar” o objeto na cena reduzida. Para isso, diminuimos o tamanho da imagem “pulando” pixels, ou seja, para metade do tamanho.



Na pirâmide gaussiana, reduz-se a imagem para metade da sua resolução, e aplica-se um filtro gaussiano para suavizar a imagem antes de a reamostrar (diminuindo efeitos como aliasing).

A estratégia para implementação da localização de objeto com variações de escala se deu da seguinte forma:

1. Construir a pirâmide de imagens
  - a. Para isso foi implementada a função `build_gaussian_pyramid`, que retorna uma pirâmide com `size` imagens, partindo da original (maior resolução)
  - b. A função `downsample` utilizada internamente na função anterior, é responsável por criar a imagem menor, amostrando a imagem anterior da pirâmide a cada 2 pixels e aplicar o filtro gaussiano para suavizar o problema de aliasing.
2. Para cada nível da pirâmide encontramos a diferença quadrática (imagem diferença) entre a imagem e o objeto.
  - a. Isso é feito pela função `quadratic_difference`, que recebe a imagem e o objeto e nos retorna a imagem diferença.
  - b. A partir da imagem diferença, procuramos o valor mínimo da mesma (onde há a chance de estar o objeto que estamos buscando), assim, armazena-se o menor valor e sua posição.
3. Por fim, desenha-se um retângulo do tamanho do objeto em volta da posição de mínimo, utilizando a função `draw_rectangle` que ajusta os parâmetros para a função `cv2.rectangle()` (da biblioteca *OpenCV*).
4. Após o desenho, comparamos o mínimo obtido com o mínimo global até agora e caso o mínimo obtido seja menor que o global, o atualizamos.

Por fim, após todas as imagens da pirâmide serem analisadas, exibimos o mínimo global e a qual imagem ele foi encontrado, e para confirmação visual, as imagens com o desenho do retângulo no objeto.

Como exemplo, mostramos abaixo a identificação de um gato em uma dada imagem, onde o melhor resultado foi obtido na Imagem 1:

