Local descriptors: Difference of Gaussians (DoG)

Course: Computer Vision

E. Francisco Roman-Rangel.
francisco.roman@itam.mx

Digital Systems Department.
Instituto Tecnológico Autónomo de México, ITAM.

September 2nd, 2020.

Last time

- Local descriptors.
- Points of interest.
- Harris corner detector.

Outline

Blob-like structure

Blob: An approximately uniform region.

Regions with smooth borders,
i.e., no sharp corners $\Longrightarrow$ circular-like regions.

Q: At which location $[x, y]$ is the PoI defined?

Blob-like structure

Blob: An approximately uniform region.

Regions with smooth borders,
i.e., no sharp corners $\implies$ circular-like regions.

Q: At which location $[x, y]$ is the PoI defined?
A: The center of the blob.

Blob detection

1. Build a scale-space image representation.
2. Detect points-of-interest (PoI) in space and scale.
3. Assign characteristic scale to each PoI (its region of interest).
4. Assign an orientation to each PoI.

Final representation: $p_i \leftarrow [x_i, y_i, \sigma_i, \theta_i]$,

where each PoI p_i is represented by:

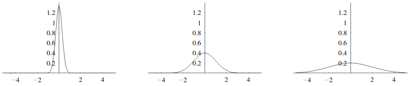
- Its spatial position $(x_i, y_i,)$.
- A characteristic scale σ_i.
- A local orientation θ_i.

Outline

Gaussian kernel

For one variable x:

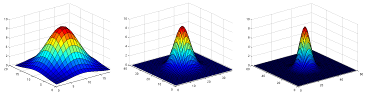
$$g(x;\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}.$$



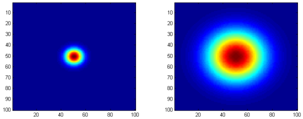
1D Gaussians with $\sigma = [0.3, 1, 2]$.

For two variables $[x, y]$:

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



2D Gaussians of varying σ.



Top view of 2D Gaussians.

Scale-space

Used to find the characteristic scale for a point of interest.

Characteristic scale (relates two concepts)

- Region of interest: spatial span used for local description.
- The need to get descriptors that are robust to scale variations.

The scale-space is constructed by convolving the input image I with a series of Gaussian kernels $g[\sigma]$ of increasing width (σ).

The output is a tensor of shape $[H \times W \times K]$,

where, H: height; W: width; K: number of kernels.

Scale-space

Methods

Most popular methods:

- Laplacian-of-Gaussian (LoG).
- Difference-of-Gaussian (DoG).
- Determinant of the Hessian.
- Hessian-Laplacian.

All based on the scale-space approach.

Outline

Intuition

Consider the center of a blob: (x_b, y_b).

Compute the average of the intensity values among all pixels within this blob (μ_b).

Intuition

Consider the center of a blob: (x_b, y_b).

Compute the average of the intensity values among all pixels within this blob (μ_b).

Now compute the average of the intensity values among all pixels within a smaller region inside the blob (μ_b^-).
Q: What is the difference between μ_b and μ_b^-?

Intuition

Consider the center of a blob: (x_b, y_b).

Compute the average of the intensity values among all pixels within this blob (μ_b).

Now compute the average of the intensity values among all pixels within a smaller region inside the blob (μ_b^-).
Q: What is the difference between μ_b and μ_b^-?
A: $= 0$.

Intuition

Consider the center of a blob: (x_b, y_b).

Compute the average of the intensity values among all pixels within this blob (μ_b).

Now compute the average of the intensity values among all pixels within a smaller region inside the blob (μ_b^-).
Q: What is the difference between μ_b and μ_b^-?
A: $= 0$.

Now compute the average of the intensity values among all pixels within a region slightly larger than the blob (μ_b^+).
Q: What is the difference between μ_b and μ_b^+?

Intuition

Consider the center of a blob: (x_b, y_b).

Compute the average of the intensity values among all pixels within this blob (μ_b).

Now compute the average of the intensity values among all pixels within a smaller region inside the blob (μ_b^-).
Q: What is the difference between μ_b and μ_b^-?
A: $= 0$.

Now compute the average of the intensity values among all pixels within a region slightly larger than the blob (μ_b^+).
Q: What is the difference between μ_b and μ_b^+?
A: $\neq 0$.

Definition

Q: What does the convolution $G_\sigma(I) = g[x, y; \sigma] * I[x, y]$ produce?

Definition

Q: What does the convolution $G_\sigma(I) = g[x, y; \sigma] * I[x, y]$ produce?
A: A sort of weighted average.

Q: Do you remember the definition of the Laplacian operator?
A: $\nabla^2 I = I_{xx} + I_{yy}$,
where, I_{aa} denotes the second partial derivative in the a direction.

Q: What does the Laplacian-of-Gaussian $(\nabla^2 G_\sigma(I))$ represent?

Definition

Q: What does the convolution $G_\sigma(I) = g[x, y; \sigma] * I[x, y]$ produce?
A: A sort of weighted average.

Q: Do you remember the definition of the Laplacian operator?
A: $\nabla^2 I = I_{xx} + I_{yy}$,
where, I_{aa} denotes the second partial derivative in the a direction.

Q: What does the Laplacian-of-Gaussian $(\nabla^2 G_\sigma(I))$ represent?
A: The local difference of average intensities.

Laplacian-of-Gaussian (LoG)

Q: What is the implication of $\nabla^2 G_\sigma(I) \approx 0$?
A: A uniform region.

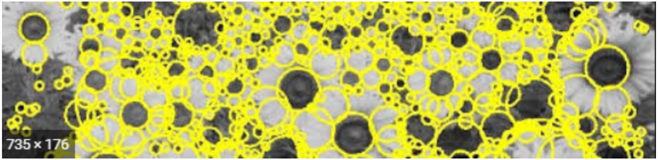
Q: How does the $LoG(I)$ look like for a non-uniform regions?
A: Far from 0.

Q: What does it mean if the $LoG(I)$ is local maximum/minimum?
A: The center of a blob with radius $r = \sqrt{2}\sigma$.

$LoG(I) >> 0$, for dark blobs, and
$LoG(I) << 0$, for bright blobs.

Scale invariance

Same numerical description at different scales.



- Strong dependence between size of blob and kernel.
- Decorrelate them to detect blobs at different scales.
- Apply Gaussians $G_\sigma(I)$ of increasing amplitude σ.

Scale-normalized LoG

$$\nabla_n^2 G_{\sigma_{(k)}}(I) = \sigma_{(k)}(\nabla^2 G_{\sigma_k}(I)).$$

Characteristic scale

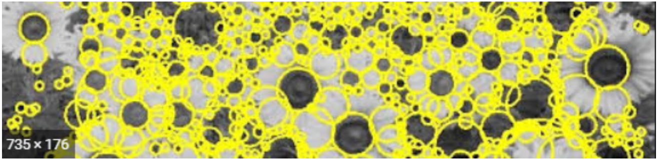
Span of the blob: region of interest around a PoI.

Q: Is it possible to have more than 1 characteristic scale per point?

Characteristic scale

Span of the blob: region of interest around a PoI.

Q: Is it possible to have more than 1 characteristic scale per point?
A: Yes. Concentric circular patterns.



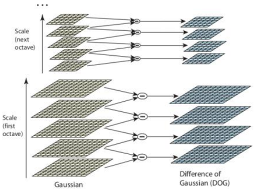
Outline

Difference-of-Gaussians (DoG)

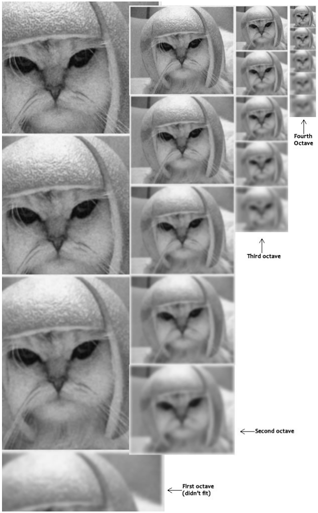
Efficient approximation to LoG.

$$DoG_k(I) = G_{\sigma_{(k)}}(I) - G_{\sigma_{(k-1)}}(I).$$

Octaves



- 4 octaves,
- 5 levels per octave,
- $\sigma_{(k)} = k\sigma_{(k-1)}$,
- $\sigma_0 = 1.6$,
- $k = \sqrt{2}$.



Local orientation

Canonical orientation θ per point is required to be robust against rotation transforms.

Q: Remember what is the "orientation" of a pixel?

Final point representation:

$$p = (x, y, \sigma, \theta)$$

Remember: one point might be associated with several characteristic scales → set of PoI might have repeated points with different scales.

Harris-Laplace

Combines Harris corner detector with the Gaussian scale-space.

$$M = \sigma_D^2 g(\sigma_I) * \begin{bmatrix} L_x^2(I, \sigma_D) & L_x L_y(I, \sigma_D) \\ L_x L_y(I, \sigma_D) & L_y^2(I, \sigma_D) \end{bmatrix}.$$

Note:

- Derivation scale σ_D.
- Integration scale σ_I.

Parameters to be adjusted, smoothing and normalization spans are not necessarily the same.

Q&A

Thank you!

francisco.roman@itam.mx