Computer Vision

4. Interest Points

Lars Schmidt-Thieme

Information Systems and Machine Learning Lab (ISMLL)
Institute for Computer Science
University of Hildesheim, Germany

Outline

1. Smoothing, Image Derivatives, Convolutions
2. Edges, Corners, and Interest Points
3. Image Patch Descriptors
4. Interest Point Matching
5. A Simple Application: Image Stitching

Outline

Smoothing / Blurring / Averaging

- **Smoothing**: Replace each pixel by the weighted average of its surrounding patch:

$$I_{\text{smooth}}(x, y; w) := \sum_{\Delta x, \Delta y} w(-\Delta x, -\Delta y) I(x + \Delta x, y + \Delta y)$$
$$= \sum_{x', y'} w(x - x', y - y') I(x', y')$$

  - **padding** with 0 at the image boundaries.
- example: **box kernel**

$$w_{-2:2,-2:2}(\Delta x, \Delta y) := \frac{1}{25} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

- **Gaussian smoothing**: smoothing with a Gaussian kernel.

Gaussian Kernels

- Precomputed weights: (clipped) Gaussian density values

$$\tilde{w}(\Delta x, \Delta y) := \begin{cases} \mathcal{N}(\sqrt{\Delta x^2 + \Delta y^2}; 0, \sigma^2), & \text{if } |\Delta x| \leq K, |\Delta y| \leq K \\ 0, & \text{else} \end{cases}$$

$$w(\Delta x, \Delta y) := \frac{\tilde{w}(\Delta x, \Delta y)}{\sum_{\Delta x', \Delta y'} \tilde{w}(\Delta x, \Delta y)}$$

- clipped: small support, **window size K**.
- example ($K = 2, \sigma^2 = 1$):

$$w_{-2:2,-2:2} := \begin{pmatrix} 0.003 & 0.013 & 0.022 & 0.013 & 0.003 \\ 0.013 & 0.060 & 0.098 & 0.060 & 0.013 \\ 0.022 & 0.098 & 0.162 & 0.098 & 0.022 \\ 0.013 & 0.060 & 0.098 & 0.060 & 0.013 \\ 0.003 & 0.013 & 0.022 & 0.013 & 0.003 \end{pmatrix}$$

Note: $\mathcal{N}(x; \mu, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Image Derivatives

- **Image Derivative**: How does the intensity values change in x or y direction?

$$I_X(x,y) := I(x,y) - I(x-1,y)$$
$$I_Y(x,y) := I(x,y) - I(x,y-1)$$

or symmetric

$$I_X(x,y) := 2I(x,y) - I(x-1,y) - I(x+1,y)$$
$$I_Y(x,y) := 2I(x,y) - I(x,y-1) - (x,y-2)$$

Convolutions

- Smoothing, Image Derivatives and further operations such as filtering can be represented by a
  - **convolution**: an image where each pixel (x, y) represents the weighted sum around (x, y) in image I weighted with w:

$$(w * I)(x, y) := \sum_{x', y'} w(x - x', y - y') I(x', y')$$

- Examples:

$$\begin{aligned}
I_{\text{smooth}} & & &= w * I \\
I_X(x, y) &:= I(x, y) - I(x - 1, y) & &= \begin{pmatrix} 1 & -1 \end{pmatrix} * I \\
I_Y(x, y) &:= I(x, y) - I(x, y - 1) & &= \begin{pmatrix} 1 \\ -1 \end{pmatrix} * I \\
\text{or } I_X(x, y) &:= 2I(x, y) - I(x - 1, y) - I(x + 1, y) & &= \begin{pmatrix} -1 & 2 & -1 \end{pmatrix} * \\
I_Y(x, y) &:= 2I(x, y) - I(x, y - 1) - (x, y - 2) & &= \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} * I
\end{aligned}$$

Convolutions / Associativity

- Convolutions are associative:

$$I * (J * K) = (I * J) * K$$

- Example:
  First smooth an image with Gaussian w from slide 2,
  then compute its x-derivative with $\begin{pmatrix} -1 & 2 & -1 \end{pmatrix}$:
  $\rightsquigarrow$ just convolve with $\begin{pmatrix} -1 & 2 & -1 \end{pmatrix} * w$

$$\begin{pmatrix} -1 & 2 & -1 \end{pmatrix} * w = \begin{pmatrix} -0.007 & 0.002 & 0.017 & 0.002 & -0.007 \\ -0.033 & 0.008 & 0.077 & 0.008 & -0.033 \\ -0.054 & 0.077 & 0.128 & 0.077 & -0.054 \\ -0.033 & 0.008 & 0.077 & 0.008 & -0.033 \\ -0.007 & 0.002 & 0.017 & 0.002 & -0.007 \end{pmatrix}$$

Outline

Edges, Corners, and Interest Points

- good candidates for points that are easy to recognize and match in two images are
  - points on edges
  - corners

  i.e., points with sudden intensity changes.

- two stage approach: given an image $I \in \mathbb{R}^{N \times M}$,
  1. compute an interestingness measure $i \in \mathbb{R}^{N \times M}$ for points,
  2. select a useful set of points $p_1, \ldots, p_K \in [N] \times [M]$
     - with high interestingness measure
     - not too close to each other.

- many names: corners, interest points, keypoints, salient points, ...

Note: $[N] := \{1, \ldots, N\}$.

Gradient Magnitude (Canny Edge Detector)

- Simply use the **magnitude of the gradient** as interestingness measure:

$$i(x,y) = \sqrt{(D_X * I)(x,y)^2 + (D_Y * I)(x,y)^2}$$

- D_X, D_Y: differentiation kernels, e.g.,

$$D_X := \begin{pmatrix} -1 & 2 & -1 \end{pmatrix}, \quad D_Y := \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

Laplacian of Gaussian and Difference of Gaussian

Further simple interestingness measures:

- **Laplacian of Gaussian** (**LoG**):

$$i(x,y) = (((D_X * D_X + D_Y * D_Y) * G) * I)(x,y)$$

  - uses second order information

- **Difference of two Gaussians** (**DoG**):

$$i(x,y) = ((G_{\sigma_1} - G_{\sigma_2}) * I)(x,y), \quad \sigma_1 \neq \sigma_2$$

  - uses variations at different scales
  - often interpreted as limit of Laplacian of Gaussians

$$((D_X * D_X + D_Y * D_Y) * G_\sigma) * I \approx \frac{\sigma}{\Delta\sigma}((G_{\sigma+\Delta\sigma} - G_{\sigma-\Delta\sigma}) * I)$$

Harris Corner Detector

- Represent a **corner** by its **patch** surrounding it, represent such a **patch** by a **weight function**

$$w : [N] \times [M] \to \mathbb{R},$$

i.e.,

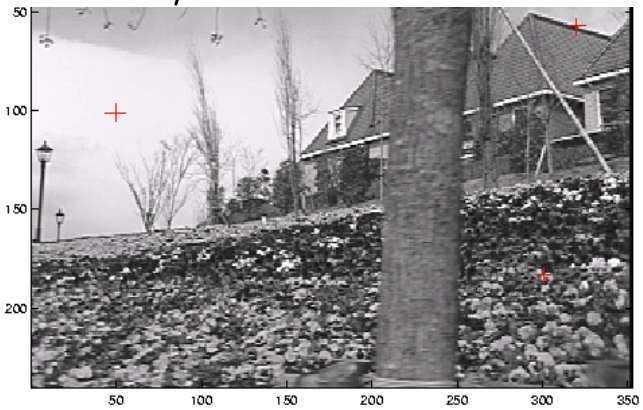
$$w(x, y) := \begin{cases} 1, & \text{if } |x - x_0| < 3 \text{ and } |y - y_0| < 3 \\ 0, & \text{else} \end{cases}$$

for a rectangular patch of size 5 centered around (x_0, y_0).

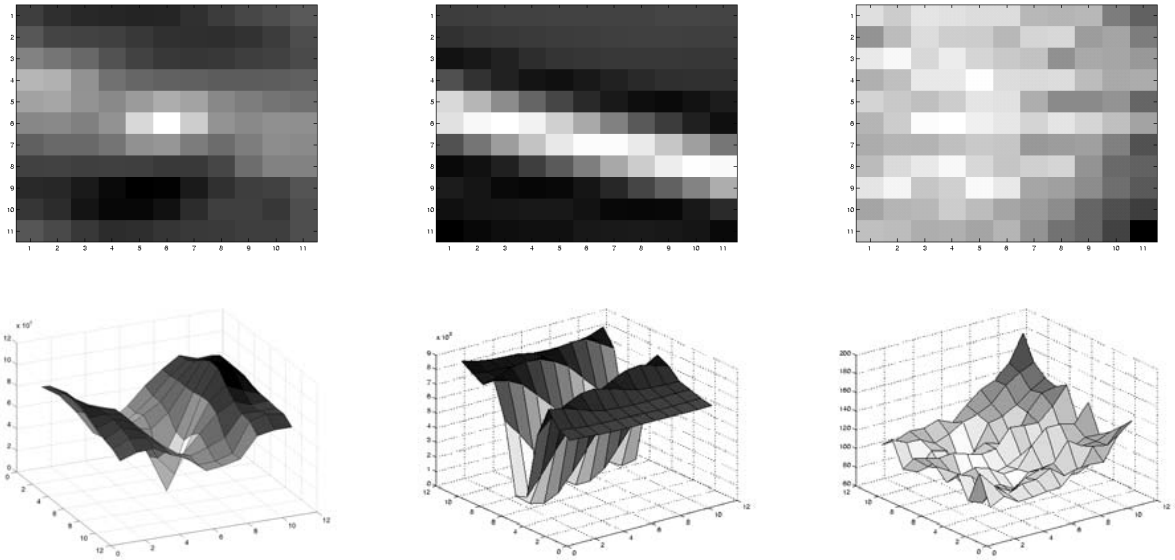
- A point is easy to identify, if its minimum in the **autocorrelation surface** is pronounced:

$$E(\Delta x, \Delta y; w) := \sum_{x,y} w(x, y)(I(x + \Delta x, y + \Delta y) - I(x, y))^2$$

Harris Corner Detector / Autocorrelation Surface



(a)



Note: left to right: flower bed, roof edge, cloud.

[Sze11, p. 187]

Harris Corner Detector

$$E(\Delta x, \Delta y; w) := \sum_{x,y} w(x,y)(I(x+\Delta x, y+\Delta y) - I(x,y))^2$$

with Hessian at minimum:

$$H(0,0;w) \approx 2 \sum_{x,y} w(x,y) \nabla I|_{(x,y)} \nabla I|_{(x,y)}^T, \quad \text{for } \frac{\partial^2 I}{\partial^2 (x,y)} := 0$$

$$= 2w * \begin{pmatrix} (I_X)^2 & I_X I_Y \\ I_X I_Y & (I_Y)^2 \end{pmatrix},$$

$$I_X(x,y) := I(x+1,y) - I(x,y) \approx \frac{\partial I}{\partial x}(x,y)$$

$$I_Y(x,y) := I(x,y+1) - I(x,y) \approx \frac{\partial I}{\partial y}(x,y)$$

Note: $I * J(x,y) := \sum_{x',y'} I(x-x', y-y') J(x',y')$ **convolution** of two images.

Harris Corner Detector

use SVD to assess steepness

$$H = U \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} U^T, \quad \sigma_1 \geq \sigma_2 \geq 0, UU^T = I$$

and define interestingness measure:

$$i_{\text{Shi-Tomasi}}(x, y) := \sigma_2$$

$$i_{\text{Harris}}(x, y) := \sigma_1\sigma_2 - \alpha(\sigma_1 + \sigma_2)^2 = \det H - \alpha\,\text{trace}(H)^2, \qquad \alpha := 0.06$$

$$i_{\text{Triggs}}(x, y) := \sigma_2 - \alpha\sigma_1, \qquad \alpha := 0.05$$

$$i_{\text{Brown}}(x, y) := \sigma_1\sigma_2/(\sigma_1 + \sigma_2) = \det H/\text{trace}(H)$$

- the larger $\sigma_{1:2}$, the steeper the autocorrelation surface E.
- Harris and Brown avoid computing σ_1, σ_2 explicitly (which requires computing a square root).

Harris Corner Detector / Algorithm

procedure INTERESTPOINTS-HARRIS($I \in \mathbb{R}^{N \times M}; w \in \mathbb{R}^{-K:K \times -L:L}, \alpha \in \mathbb{R}$)

$I_X := D_X * I$

$I_Y := D_Y * I$

$I_X^2 := I_X \cdot I_X$

$I_Y^2 := I_Y \cdot I_Y$

$I_X I_Y := I_X \cdot I_Y$

$A := w * I_X^2$ ▷ compute $H(x,y) = \begin{pmatrix} A(x,y) & C(x,y) \\ C(x,y) & B(x,y) \end{pmatrix}$

$B := w * I_Y^2$

$C := w * I_X I_Y$

$i := A \cdot B - C \cdot C - \alpha(A+B) \cdot (A+B)$

return i

- D_X, D_Y: differentiation kernels, e.g.,

$$D_X := \begin{pmatrix} -1 & 2 & -1 \end{pmatrix}, D_Y := \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

Note: $\cdot$ denotes the element/pixelwise product.

Interest Points at Different Scales (SIFT Detector)

- Interest points also can be identified **at different scales** in parallel:

$$i(p,s) := (G_{\sigma_{s+1}} * I - G_{\sigma_s} * I), \quad s \in [S]$$

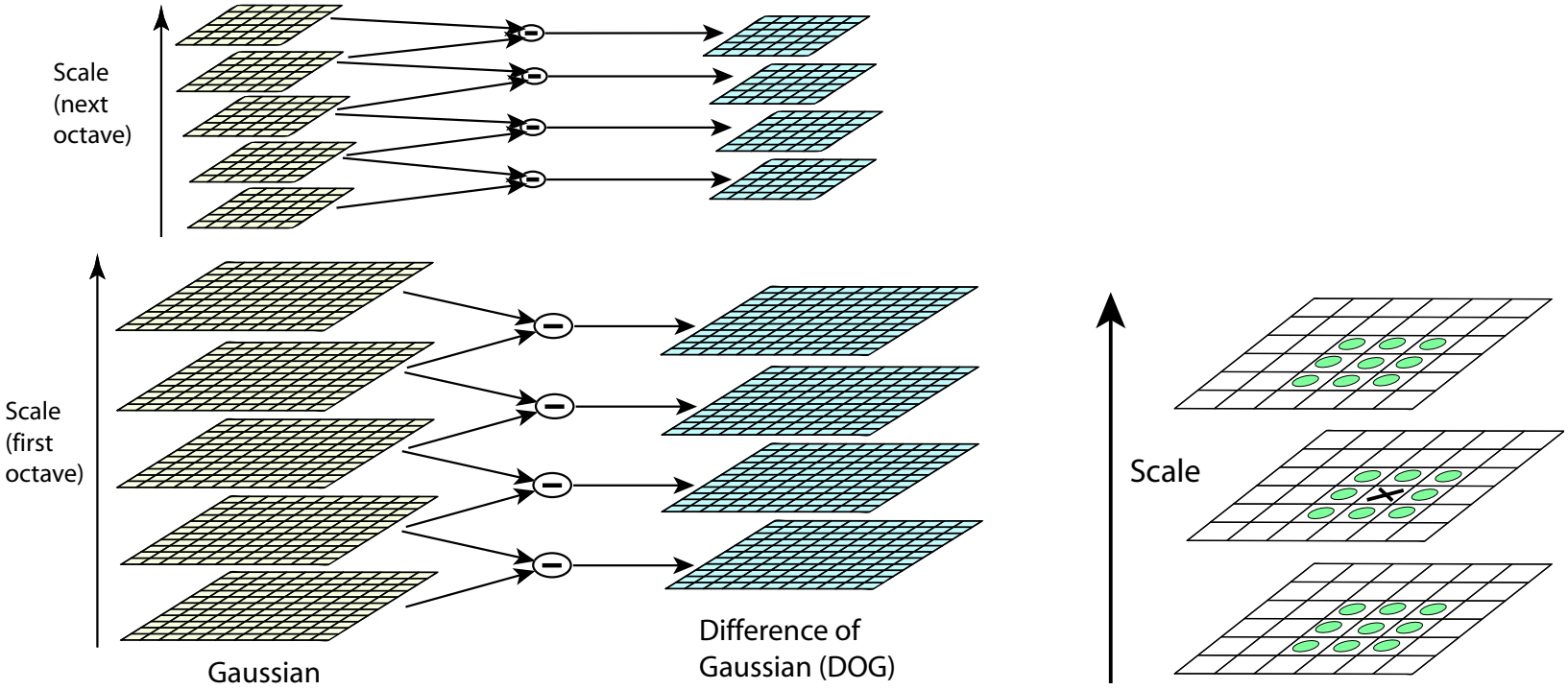
where

$$\sigma_1 > \sigma_2 > \cdots > \sigma_S$$

where $S \in \mathbb{N}$ is the **number of scale levels**

- Often scale levels are grouped by octaves:
  - each octave is represented by a downsampling by a factor 2
  - scales within an octave are $\sigma_s := 2^{s/S_o}\sigma$
    (with S_o the number of scale levels within an ocatve)

Interest Points at Different Scales (SIFT Detector)



(a) (b)

[Sze11, p. 216]

Non-Maximum Suppression

- Often neighbors of interest points have similar high interestingness, yielding redundant close-by interest points.

- Keep only interest points that are **local maxima** in their neighborhood:

$$i'(p) := \begin{cases} i(p), & \text{if } i(p) > i(p') \; \forall p' \in N(p) \\ 0, & \text{else} \end{cases}, \quad p \in [N] \times [M]$$

with **neighborhood**

$$N_K(p) := \{p' \in [N] \times [M] \mid |p_x - p'_x| \leq K, |p_y - p'_y| \leq K, p' \neq p\} \quad \text{recta}$$

$$N_K(p) := \{p' \in [N] \times [M] \mid ||p - p'|| \leq K, p' \neq p\} \quad \text{circu}$$

Non-Maximum Suppression / At Different Scale

- Non-Maximum Suppression also can be extended to work on interest points at different scale:

$$N_K(p,s) := \{(p',s') \in [N] \times [M] \times [S] \mid |p_x - p'_x| \leq K, |p_y - p'_y| \leq K, \\ |s - s'| \leq 1, p' \neq p\}$$

Non-Maximum Suppression / At Different Scale

- SIFT refines interest points by further steps:
  - localization at sub-pixel granularity

Outline

Histograms

SIFT descriptors

Histograms of Oriented Gradients (HoG)

Bag of Words Descriptors

?? Dimensionality Reduction

Outline

...

Outline

...

Summary

- ...

Further Readings

- Interest points and patch descriptors: [Pri12, ch. 13], [Sze11, ch. 4].
- Image stitching: [Sze11, ch. 9].

References

Simon JD Prince.
Computer vision: models, learning, and inference.
Cambridge University Press, 2012.

Richard Szeliski.
Computer vision: algorithms and applications.
Springer Science & Business Media, 2011.