

## Computer Vision 4. Interest Points

#### Lars Schmidt-Thieme

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

《口》《聞》《臣》《臣》 王言 '오오오

## 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

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少へつ

## 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



# Smoothing / Blurring / Averaging

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

$$egin{aligned} & I_{ ext{smooth}}(x,y;w) := \sum_{\Delta x,\Delta y} w(-\Delta x,-\Delta y) I(x+\Delta x,y+\Delta y) \ & = \sum_{x',y'} w(x-x',y-y') I(x',y') \end{aligned}$$

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

## Gaussian Kernels

► Precomputed weights: (clipped) Gaussian density values

$$egin{aligned} & ilde{w}(\Delta x,\Delta y) := egin{cases} \mathcal{N}(\sqrt{\Delta x^2 + \Delta y^2}; 0, \sigma^2), & ext{if } |\Delta x| \leq \mathcal{K}, |\Delta y| \leq \mathcal{K} \ 0, & ext{else} \end{aligned} \ & w(\Delta x,\Delta y) := rac{ ilde{w}(\Delta x,\Delta y)}{\sum_{\Delta x',\Delta y'} ilde{w}(\Delta x,\Delta y)} \end{aligned}$$

- ► 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:  $\mathcal{N}(x; \mu, \sigma^2) := \frac{1}{2\sigma^2} e^{-\frac{1}{2\sigma^2}}$ . Lars Schmidt-Thieme, Information of Hildesheim, Germany Lab (ISMLL), University of Hildesheim, Germany







#### original:







シック 正所 エボットボット きょう



#### original:







シック 비론 《로》《로》《唱》《曰》



#### original:







・ 日 ト 《 四 ト 《 三 ト 《 三 ト 《 四 ト 《 四 ト



#### original:







シック 正則 エル・エリ・エリ・

## 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

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

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

・ロト < 団ト < ヨト < ヨト < 国ト < ロト</li>

## Image Derivatives / Example



#### original (grayscale):



#### derivative in x-direction:



#### シック 三回 ヘルマ・ルマ・トロマ

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

5 / 47

## Image Derivatives / Example



#### original (grayscale):



#### derivative in y-direction:



#### 

## Convolutions



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:





## 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  $\begin{pmatrix} -1 & 2 & -1 \end{pmatrix}$ :  $\rightarrow$  just convolve with  $\begin{pmatrix} -1 & 2 & -1 \end{pmatrix} * w$ 

$$\left(\begin{array}{cccc} -1 & 2 & -1 \end{array}\right) * w = \left(\begin{array}{cccc} -0.007 & 0.002 & 0.017 & 0.002 & -0.007 \\ -0.033 & 0.008 & 0.077 & 0.008 & -0.033 \\ -0.054 & 0.077 & 0.128 & 0.077 & -0.054 \\ -0.033 & 0.008 & 0.077 & 0.008 & -0.033 \\ -0.007 & 0.002 & 0.017 & 0.002 & -0.007 \end{array}\right)$$

▲□▶▲母▶▲글▶▲글▶ 三百 めんの



## 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

・日・《四・《川・《田・《日・

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

《日》《四》《王》《王》 레티 이익()



## 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.,

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

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少々で

## Gradient Magnitude / Example



### original (grayscale):



## gradient magnitude:



#### 

## Gradient Magnitude / Example



### original (grayscale):



#### overlay with 500 interest points:



・ロト・4日ト・4日ト・4日ト 日日・990



# 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

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

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:

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

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany



◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少々で

## Shiversiter Fildesheiff









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

[Sze11, p. 187] • □ ▶ • 륜 ▶ • 토 ▶ 토 = ∽ ९...

# Universiter.

<u>-</u>.

## 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

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

and define interestingness measure:

$$\begin{split} i_{\mathsf{Shi-Tomasi}}(x,y) &:= \sigma_2 \\ i_{\mathsf{Harris}}(x,y) &:= \sigma_1 \sigma_2 - \alpha (\sigma_1 + \sigma_2)^2 = \det H - \alpha \operatorname{trace}(H)^2, \qquad \alpha := 0.06 \\ i_{\mathsf{Triggs}}(x,y) &:= \sigma_2 - \alpha \sigma_1, \qquad \alpha := 0.05 \\ i_{\mathsf{Brown}}(x,y) &:= \sigma_1 \sigma_2 / (\sigma_1 + \sigma_2) = \det H / \operatorname{trace}(H) \end{split}$$

- the larger  $\sigma_{1:2}$ , the steeper the autocorrelation surface *E*.
- ► Harris and Brown avoid computing σ<sub>1</sub>, σ<sub>2</sub> explicitly (which requires computing a square root).



## Harris Corner Detector / Algorithm



1: procedure INTERESTPOINTS-HARRIS( $I \in \mathbb{R}^{N \times M}$ ;  $w \in \mathbb{R}^{-K:K \times -L:L}$ ,  $\alpha \in \mathbb{R}$ ) 2:  $I_X := D_X * I$  $I_{\mathbf{V}} := D_{\mathbf{V}} * I$ 3:  $I_X^2 := I_X \cdot I_X$ 4:  $I_Y^2 := I_Y \cdot I_Y$ 5: 6:  $I_X I_Y := I_X \cdot I_Y$  $\triangleright \text{ compute } H(x, y) = \begin{pmatrix} A(x, y) & C(x, y) \\ C(x, y) & B(x, y) \end{pmatrix}$  $A := w * I_{x}^{2}$ 7:  $B := w * I_{v}^{2}$ 8:  $C := w * I_X I_Y$ 9:  $i := A \cdot B - C \cdot C - \alpha (A + B) \cdot (A + B)$ 10: 11: 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.

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

Computer Vision 2. Edges, Corners, and Interest Points

## Harris Corner Detector / Example





(a)

(b)

(c)

#### a) original, b) Harris corners, c) DoG interest points

[Sze11, p. 213] 《 □ ▷ 《 큔 ▷ 《 壴 ▷ 《 壴 ▷ 훅 트 》 역 ()



## 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:
  - $\blacktriangleright$  each octave is represented by a downsampling by a factor 2
  - scales within an octave are σ<sub>s</sub> := 2<sup>s/S₀</sup>σ
     (with S₀ the number of scale levels within an ocatve)

《曰》《聞》《臣》《臣》 토言 '오오오



## Interest Points at Different Scales (SIFT Detector)



19 / 47

QC

# Shiversiter

## 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:

$$i'(p) := egin{cases} i(p), & ext{if } i(p) > i(p') \ orall p' \in N(p) \ 0, & ext{else} \end{cases}, \quad p \in [N] imes [M]$$

with neighborhood

$$\begin{split} &N_{K}(p) := \{ p' \in [N] \times [M] \mid |p_{x} - p'_{x}| \leq K, |p_{y} - p'_{y}| \leq K, p' \neq p \} & \text{recta} \\ &N_{K}(p) := \{ p' \in [N] \times [M] \mid ||p - p'|| \leq K, p' \neq p \} & \text{circu} \end{split}$$

<ロ> < 団> < 団> < 三> < 三> 三三 のへで

## Non-Maximum Suppression / Example





(c) ANMS 250, r = 24 (d) ANMS 500, r = 16Note: ANMS = adaptive non-maximum suppression; see the book for details = 1, p. 214]



## Non-Maximum Suppression / At Different Scale

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

$$egin{aligned} &\mathcal{N}_{\mathcal{K}}(p,s):=&\{(p',s')\in [\mathcal{N}] imes [\mathcal{M}] imes [\mathcal{S}]\mid |p_x-p_x'|\leq \mathcal{K}, |p_y-p_y'|\leq \mathcal{K}, \ &|s-s'|\leq 1, p'
eq p\} \end{aligned}$$



[Sze11, p. 216] ∢□▷∢률▷∢≧▷∢≧▷ ≜⊨ ୬९୯.

## SIFT Interest Points



SIFT refines interest points by further steps:

- non-maximum suppression at different scale
- ► localization of interest points at sub-pixel granularity

うせん 正則 スポッスポッス セッ

## 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

・ロト・(聞・・ヨト・(問・・ロト)

## Sniversiter Fildesheim

## Image Patch Descriptors

- ► Which properties from a patch to extract?
  - ► grayscale intensities, color intensities, gradient directions
- Which patches to extract?
  - ► orientation of the patch w.r.t. the image frame
  - ▶ offset of the patch w.r.t. the interest point (cells)

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少々で

## Histograms

- ▶ the most simple patch:
  - a square centered on the interest point



\* □ > \* @ > \* E > \* E > E = \* 9 < 0</p>

## Histograms

- ▶ the most simple patch:
  - a square centered on the interest point
- ► properties:
  - most simple: grayscale intensities of the pixels

- ▶ how to represent?
  - as a matrix or a vector



#### もしゃ 正明 ふばやえばや ふむやくしゃ
- ▶ the most simple patch:
  - a square centered on the interest point
- ► properties:
  - most simple: grayscale intensities of the pixels

- ▶ how to represent?
  - as a matrix or a vector
    - ► is affected by rotations



#### ・日本・四本・山本・山本・山本・日本・日本

- ▶ the most simple patch:
  - a square centered on the interest point
- ► properties:
  - most simple: grayscale intensities of the pixels

- ▶ how to represent?
  - as a matrix or a vector
    - ► is affected by rotations
  - ▶ by some scalar properties (mean, standard deviation)

もうすい 正則 ふかす ふやす ふきゃくりゃ



- ▶ the most simple patch:
  - a square centered on the interest point
- ► properties:
  - most simple: grayscale intensities of the pixels

- ▶ how to represent?
  - as a matrix or a vector
    - is affected by rotations
  - ▶ by some scalar properties (mean, standard deviation)
    - represents only little information



シック 비토 《王》《王》《臣》《曰》

- ▶ the most simple patch:
  - a square centered on the interest point
- ► properties:
  - most simple: grayscale intensities of the pixels

- ▶ how to represent?
  - as a matrix or a vector
    - is affected by rotations
  - ▶ by some scalar properties (mean, standard deviation)
    - represents only little information
  - by its histogram



Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少へつ

- ▶ the most simple patch:
  - a square centered on the interest point
- ► properties:
  - most simple: grayscale intensities of the pixels
    - is affected by global intensity fluctuations
  - gradient directions
- ▶ how to represent?
  - as a matrix or a vector
    - is affected by rotations
  - ▶ by some scalar properties (mean, standard deviation)
    - represents only little information
  - ► by its **histogram**



Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少へつ

#### Universiter Fildesheim

#### Histograms / Intensities

► represent interest point (x, y) by its B-dimensional intensity histogram features φ(x, y):

$$\begin{split} \phi(x,y)_b &:= |\{(x',y') \in \mathcal{N}(x,y) \mid I(x',y') \in bin_b\}|, \quad b = 0, \dots, B-1\\ bin_b &:= [\frac{b}{B}I_{\max}, \frac{b+1}{B}I_{\max}[\\ \mathcal{N}(x,y) &:= \{(x',y') \in [N] \times [M] \mid |x'-x| < K, |y'-y| < K\} \end{split}$$

for intensities I(x, y) in range  $[0, I_{max}]$ .

・日本・4日本・4日本・4日本・996で

### Histograms / Smoothed Counting

- To avoid non-continuous changes if a value crosses bin boundaries, values can be counted
  - in both closest bins,
  - ► antiproportional to their distance from the bin center

$$\mathsf{binc}_b := \frac{b+0.5}{B} I_{\mathsf{max}}$$

$$\mathsf{bin}_b := \sum_{(x',y') \in \mathcal{N}(x,y)} \mathsf{max}(0, 1 - \frac{|I(x',y') - \mathsf{binc}_b|}{I_{\mathsf{max}}/B})$$

sometimes called trilinear counting.

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少へつ



#### Histograms / Gradient Directions



► represent interest point (x, y) by its B-dimensional gradient directions histogram features φ(x, y):

$$\begin{split} \phi(x,y)_b &:= |\{(x',y') \in \mathcal{N}(x,y) \mid d(x',y') \in bin_b\}|, \quad b = 0, \dots, B-1 \\ d(x,y) &:= \tan^{-1}((D_Y * I)(x,y)/(D_X * I)(x,y)) \\ bin_b &:= [\frac{b}{B}2\pi, \frac{b+1}{B}2\pi[ \end{split}$$

・日本・西本・山田・山田・山市・今日・

#### Histograms / Gradient Directions



► represent interest point (x, y) by its B-dimensional gradient directions histogram features φ(x, y):

$$\begin{split} \phi(x,y)_b &:= |\{(x',y') \in \mathcal{N}(x,y) \mid d(x',y') \in bin_b\}|, \quad b = 0, \dots, B-1 \\ d(x,y) &:= \tan^{-1}((D_Y * I)(x,y)/(D_X * I)(x,y)) \\ bin_b &:= [\frac{b}{B} 2\pi, \frac{b+1}{B} 2\pi] \end{split}$$

► variant: weight gradients by their magnitude:

$$\phi(x,y)_b := \sum_{(x',y') \in \mathcal{N}(x,y), d(x',y') \in \mathsf{bin}_b} (D_X * I)(x',y')^2 + (D_Y * I)(x',y')^2$$

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少へつ

Computer Vision 3. Image Patch Descriptors

#### Histograms / Gradients / Example





[Sze11, p. 217] < □ ▶ < 륜 ▶ < 토 ▶ 로 = ∽ 의 ೕ

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

29 / 47

#### Block Descriptors



 Describe an interest point not just by features of the surrounding patch,

but by the features of several neighboring patches (blocks, cells):

$$\phi(x,y) := \bigoplus_{(x',y') \in \mathcal{C}(x,y)} \phi'(x',y')$$
$$\mathcal{C}(x,y) := \{x + c\Delta X, y + d\Delta Y \mid c, d \in \{-C, \dots, C\}\}$$

- Often a simple partition of a large  $(2C+1)(2K+1) \times (2C+1)(2K+1)$  patch is used  $(\Delta X = \Delta Y = 2K+1)$ .
- Features have dimensions  $(2C + 1)^2 B$ .

Note:  $(x_1, \ldots, x_N) \oplus (y_1, \ldots, y_M) := (x_1, \ldots, x_N, y_1, \ldots, y_M)$  concatenation.

#### **Block Descriptors**





[Low04, p. 15] 《 ロ ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ 《 문 ▷ ③

うへで 31 / 47

# Align Patches by the Gradient Direction of the Interest

- Extract features from the image rotated by
  - the negative gradient direction at the interest point
  - around the interest point

(afterwards the gradient at the interest point (x, y) points towards positive x-direction):

### SIFT descriptors



- ► patches:
  - $\blacktriangleright$  extract from the scaled image the interest point has been detected on
  - align patch by the gradient direction of the interest point
  - $16 \times 16$ , partitioned into 16 blocks a  $4 \times 4$
- block features:
  - gradient directions
  - weighted by a Gaussian of the distance to the interest point
- block feature aggregation:
  - smoothly counted histograms
  - ► 8 bins
- $\blacktriangleright \ \leadsto$  feature vector  $\phi \in \mathbb{R}^{128}$
- normalization in 3 steps:

$$\phi'_i := \phi_i / ||\phi||_2, \qquad \phi''_i := \min(0.2, \phi'_i), \qquad \phi'''_i := \phi''_i / ||\phi''||_2$$

◇ ▷ ♡ □ □ ■ ▲ □ ▷ ▲ □ ▷ ▲ □ ▷ ▲ □ ▷

Image Descriptors

To describe a whole image (not just a patch), two main approaches are used:

- 1. Concatenate patch descriptors of equally spaced "interest points"
  - 1.1 e.g., used in Histograms of Oriented Gradients (HoG)

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少々で

#### Universiter Hildesheim

Image Descriptors

To describe a whole image (not just a patch), two main approaches are used:

- 1. Concatenate patch descriptors of equally spaced "interest points"
  - 1.1 e.g., used in Histograms of Oriented Gradients (HoG)

#### 2. Bag of words descriptors:

- 2.1 compute interest points and their descriptors for a set of images
- 2.2 discretize the descriptors
  - e.g., clustering in K clusters using k-means
- 2.3 represent each image by the  ${\cal K}$  cluster frequencies of their interest point descriptors

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少々で

Computer Vision 3. Image Patch Descriptors

#### Histograms of Oriented Gradients (HoG)





**Figure 13.17** HOG descriptor. a) Original image. b) Gradient orientation, quantized into nine bins from 0 to  $180^{\circ}$ . c) Gradient magnitude. d) Cell descriptors are 9D orientation histograms that are computed within  $6 \times 6$  pixel regions. e) Block descriptors are computed by concatenating  $3 \times 3$  blocks of cell descriptors. The block descriptors are normalized. The final HOG descriptor consists of the concatenated block descriptors.

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

35 / 47

[Pri12, p. 343]

#### 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

・日・《四・《王・《田・《日・

#### Settings, Assumptions, Distances

Two settings:

- match interest points in different scenes
  - goal: detect similar objects (object identification)
  - coordinates of the points do not matter

$$d(\left(\begin{array}{c}x_1\\y_1\end{array}\right),\left(\begin{array}{c}x_2\\y_2\end{array}\right)):=d'(\phi(x_1,y_1),\phi(x_2,y_2))=||\phi(x_1,y_1)-\phi(x_2,y_2)||_2$$

$$d\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) := \alpha d'\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + \beta d'(\phi(x_1, y_1), \phi(x_2, y_2))$$
$$= \alpha || \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} - \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} ||_2 + \beta ||\phi(x_1, y_1) - \phi(x_2, y_2)||_2$$



### Settings, Assumptions, Distances

Two settings:

- match interest points in different scenes
  - goal: detect similar objects (object identification)
  - coordinates of the points do not matter

$$d(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) := d'(\phi(x_1, y_1), \phi(x_2, y_2)) = ||\phi(x_1, y_1) - \phi(x_2, y_2)||_2$$

- match interest points in two views of the same scene
  - goal: detect corresponding points in different views of the same scene (required for SLAM)
  - coordinates of corresponding points also should be close, e.g.,

$$d\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) := \alpha d'\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + \beta d'(\phi(x_1, y_1), \phi(x_2, y_2))$$
$$= \alpha ||\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} - \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} ||_2 + \beta ||\phi(x_1, y_1) - \phi(x_2, y_2)||_2$$

#### Simple methods



To match two sets P and Q of interest points:

match interest points by distance threshold

$$p \sim q : \Leftrightarrow d(p,q) < d_{\mathsf{max}}, \quad p \in P, q \in Q$$

► distance threshold d<sub>max</sub> can be estimated from known matches and non-matches

#### Simple methods



To match two sets P and Q of interest points:

match interest points by distance threshold

$$p \sim q : \Leftrightarrow d(p,q) < d_{\mathsf{max}}, \quad p \in P, q \in Q$$

- ► distance threshold d<sub>max</sub> can be estimated from known matches and non-matches
- match interest points by nearest neighbor

$$p \sim q : \Leftrightarrow q = \operatorname*{arg\,min}_{q \in Q} d(p,q)$$

シック 비로 《로》《로》《唱》《曰》

#### Nearest Neighbor Distance Ratio



► match interest points by nearest neighbor distance ratio (NNDR)

$$p \sim q :\Leftrightarrow i) \ q = \operatorname*{arg\,min}_{q \in Q} d(p,q) \text{ and}$$
  
 $ii) \ \operatorname{NNDR}(p,q) := \frac{d(p,q)}{d(p,q')} < \operatorname{NNDR}_{\min}, \quad q' := \operatorname*{arg\,min}_{q' \in Q \setminus \{q\}} d(p,q')$ 



Computer Vision 4. Interest Point Matching



## Comparison of Different Descriptors & Matchings a) fixed threshold:



Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

39 / 47

Computer Vision 4. Interest Point Matching



# Comparison of Different Descriptors & Matchings b) nearest neighbor:





# Comparison of Different Descriptors & Matchings c) nearest neighbor distance ratio:



#### Mutual Nearest Neighbors



► match interest points if they **mutually** are nearest neighbors

$$p \sim q : \Leftrightarrow i) \; q = rgmin_{q \in Q} d(p,q) \; ext{and}$$
 $ii) \; p = rgmin_{p \in P} d(p,q)$ 

► also for more than two views P<sub>1</sub>, P<sub>2</sub>,..., P<sub>V</sub> (called closed chains)

$$(p_1, p_2, \dots, p_V)$$
 corresponding tuple  
 $:\Leftrightarrow i) p_{v+1} = \underset{q \in P_{v+1}}{\operatorname{arg min}} d(p_v, q), \quad v = 1, \dots, V-1 \text{ an}$   
 $ii) p_1 = \underset{q \in P_V}{\operatorname{arg min}} d(p_1, q)$ 

シック・目前 (ボッ・(ボッ・(型)) (ロ)

#### 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

・ロト・中国・ 山田・ 山田・ 人日・

#### Image Stitching



 join several images depicting overlapping parts of the same real scene to one large image

ペロト 《四ト 《三ト 《三ト 》目上 のへで

#### Image Stitching



- ▶ join several images depicting overlapping parts of the same real scene to one large image
- ► algorithm:
  - 1. detect interest points in all images and extract their descriptors
  - 2. match interest points between every two images
  - 3. form a tree linking the best matching image pairs
  - 4. estimate a similarity transform between each two such images
  - 5. transform all images to joint coordinates
  - 6. average overlapping image regions

- 今夕今 正則 《田》《田》《田》《日》

#### Image Stitching



- ▶ join several images depicting overlapping parts of the same real scene to one large image
- ► algorithm:
  - 1. detect interest points in all images and extract their descriptors
  - 2. match interest points between every two images
  - 3. form a tree linking the best matching image pairs
  - 4. estimate a similarity transform between each two such images
  - 5. transform all images to joint coordinates
  - 6. average overlapping image regions
- also called panography

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

#### Image Stitching / Example





#### [Sze11, p. 312] < □ ▷ < @ ▷ < ≣ ▷ < ≣ ▷ < ≣ ▷ < ≡ ▷ < < ♡ < ♡

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

42 / 47

Computer Vision 5. A Simple Application: Image Stitching

#### Image Stitching / Different Transforms





(a) translation [2 dof]



(b) affine [6 dof]





(c) perspective [8 dof] (d) 3D rotation [3+ d

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

43 / 47

### Image Stitching / Example





#### Summary



- Small intensity fluctuations can be damped by smoothing, intensity changes can be captured by image derivatives, both being convolutions.
- ► Interest points are found as maxima of an interestingness measure,
  - gradient magnitude, Laplacian of Gaussian (LoG), Different of two Gaussians (DoG)
  - Harris corners:
    - large eigenvalues of the Hessian
    - can be approximated efficiently: det  $H \alpha (traceH)^2$
  - ► SIFT:
    - detected interest points at different scale
    - several further tweaks
- non-maximum suppressions:

ignore large values in the vicinity of a maximum

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少々で

### Summary (2/3)



- Interest points are characterized by local image information (descriptors)
- Descriptors often describe several patches (blocks/cells)
- Patches are described by histograms
- Histograms usually do not count pixel intensities, but gradient directions
- Descriptors sometimes
  - ► align patches with the orientation of the gradient at the interest point
  - weight gradient directions by their
    - gradient magnitude and/or
    - distance of the location to the interest point
- Common descriptors:
## Summary (3/3)

Jniversiter Tildesheif

- Whole images can be described two ways:
  - by the descriptors on a fixed grid of "interest points"
  - by the cluster frequencies of descriptors of variably located interest points

Both is useful, e.g. for image classification.

- Interest points are matched by their descriptors
  - for geometric tasks: also by their positons
- ► To match interest points, nearest neighbors are used
  - with a maximal distance threshold to avoid wrong matches e.g. of points occluded in one view
  - Nearest Neighbor Distance Ratio
  - mutual nearest neighbors, closed chains in multiple views.
- Corresponding points can be used for
  - image stitching
  - ► SLAM, camera auto-calibration, ...

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

◆□▶ ◆□▶ ★∃▶ ★∃▶ ★目★ 少へつ

## Further Readings



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

- (日) - (四) - (三) - (三) - (四) - (2

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany

## References





## David G. Lowe.

Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2):91–110, 2004.



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.

うせん 正則 ふばやふぼやふ 見て ひょう

Lars Schmidt-Thieme, Information Systems and Machine Learning Lab (ISMLL), University of Hildesheim, Germany