Lecture 2
Local Interest Point Detectors

Lin ZHANG, PhD
School of Software Engineering
Tongji University
Spring 2019
Content

• Local Invariant Features
  • Motivation
  • Requirements
  • Invariance
• Harris Corner Detector
• Scale Invariant Point Detection
  • Automatic scale selection
  • Laplacian-of-Gaussian detector
  • Difference-of-Gaussian detector
Motivation

- Global representations have major limitations
- Instead, describe and match only local regions
- Increased robustness to
  - Occlusions
  - Articulation
  - Intra-category variations
Motivation

Application: Image Matching
Motivation

Application: Image Matching

Harder Case

by Diva Sian

by scgbt
Motivation

Application: Image Matching

NASA Mars Rover Images
Motivation

Application: Image Matching  (Look for tiny colored squares)

NASA Mars Rover images with SIFT matches
Motivation

• Panorama stitching
  • We have two images – how do we combine them?
Motivation

• Panorama stitching
  • We have two images – how do we combine them?

• Procedure:
  – Detect feature points in both images
Motivation

• Panorama stitching
  • We have two images – how do we combine them?

• Procedure:
  – Detect feature points in both images
  – Find corresponding pairs
Motivation

• Panorama stitching
  • We have two images – how do we combine them?

• Procedure:
  – Detect feature points in both images
  – Find corresponding pairs
  – Use these pairs to align the images
General Approach for Image Matching

1. Find a set of distinctive keypoints
2. Define a region around each keypoint
3. Extract and normalize the region content
4. Compute a local descriptor from the normalized region
5. Match local descriptors

Source: B. Leibe
Characteristics of Good Features

- **Repeatability**
  - The same feature can be found in several images despite geometric and photometric transformations

- **Saliency**
  - Each feature has a distinctive description

- **Compactness and efficiency**
  - Many fewer features than image pixels

- **Locality**
  - A feature occupies a relatively small area of the image; robust to clutter and occlusion
Invariance: Geometric Transformations
Level of Geometric Invariance
Invariance: Photometric Transformations

- Often modeled as a linear transformation:
  - Scaling + Offset
Applications

Feature points are used for:

- Motion tracking
- Image alignment
- 3D reconstruction
- Object recognition
- Indexing and database retrieval
- Robot navigation
Content

• Local Invariant Features
  • Motivation
  • Requirements
  • Invariance

• Harris Corner Detector

• Scale Invariant Point Detection
  • Automatic scale selection
  • Laplacian-of-Gaussian detector
  • Difference-of-Gaussian detector
Finding Corners

My office,
5:30PM, Sep. 18, 2011
Finding Corners

• Key property: in the region around a corner, image gradient has two or more dominant directions

• Corners are repeatable and distinctive

Corner Detection: Basic Idea

- We should easily recognize the point by looking through a small window.
- Shifting a window in any direction should give a large change in intensity.

“flat” region: no change in all directions
“edge”: no change along the edge direction
“corner”: significant change in all directions
Harris Detector: Basic Idea

Demo of a point + with well distinguished neighborhood. Moving the window in any direction will result in a large intensity change.
Harris Detector: Basic Idea

Demo of a point + with well distinguished neighborhood.

Moving the window in any direction will result in a large intensity change.
Harris Detector: Basic Idea

Demo of a point + with well distinguished neighborhood.
Moving the window in any direction will result in a large intensity change.

Difference = 5
Harris Detector: Basic Idea

Demo of a point + with well distinguished neighborhood. Moving the window in any direction will result in a large intensity change.
Harris Detector: Basic Idea

Demo of a point $+$ with well distinguished neighborhood.
Moving the window in any direction will result in a large intensity change.
Harris Detector: Basic Idea

Demo of a point + with well distinguished neighborhood.
Moving the window in any direction will result in a large intensity change.
Harris Detector: Basic Idea

Demo of a point + with well distinguished neighborhood. Moving the window in any direction will result in a large intensity change.

Difference = 3
Harris Detector: Basic Idea

Demo of a point + with well distinguished neighborhood.
Moving the window in any direction will result in a large intensity change.

Difference = 2
Harris Corner Detection: Mathematics

Change in appearance of a local patch (defined by a window \( w \)) centered at \( p \) for the shift \((\Delta x, \Delta y)\):

\[
S_w(\Delta x, \Delta y) = \sum_{(x_i, y_i) \in w} (f(x_i, y_i) - f(x_i + \Delta x, y_i + \Delta y))^2
\]
Harris Corner Detection: Mathematics

\[ S_w(\Delta x, \Delta y) = \sum_{(x_i, y_i) \in w} (f(x_i, y_i) - f(x_i + \Delta x, y_i + \Delta y))^2 \]  

\[ \approx \sum_{(x_i, y_i) \in w} \left( f(x_i, y_i) - f(x_i, y_i) - \left[ \frac{\partial f(x_i, y_i)}{\partial x}, \frac{\partial f(x_i, y_i)}{\partial y} \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \]  

\[ = \sum_{(x_i, y_i) \in w} \left( \begin{bmatrix} \frac{\partial f(x_i, y_i)}{\partial x} \\ \frac{\partial f(x_i, y_i)}{\partial y} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \]  

\[ = [\Delta x, \Delta y] \sum_{(x_i, y_i) \in w} \begin{bmatrix} \frac{\partial f(x_i, y_i)}{\partial x} \\ \frac{\partial f(x_i, y_i)}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial f(x_i, y_i)}{\partial x} & \frac{\partial f(x_i, y_i)}{\partial y} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \]  

\[ = [\Delta x \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \]
Harris Corner Detection

\[ M = \left[ \sum_{(x_i, y_i) \in w} \left( \frac{\partial f(x_i, y_i)}{\partial x} \right)^2 \right] \left[ \sum_{(x_i, y_i) \in w} \left( \frac{\partial f(x_i, y_i)}{\partial x} \cdot \frac{\partial f(x_i, y_i)}{\partial y} \right) \right] \]

\[ \left[ \sum_{(x_i, y_i) \in w} \left( \frac{\partial f(x_i, y_i)}{\partial x} \cdot \frac{\partial f(x_i, y_i)}{\partial y} \right) \right] \left[ \sum_{(x_i, y_i) \in w} \left( \frac{\partial f(x_i, y_i)}{\partial y} \right)^2 \right] \]

- It is real symmetric
Harris Corner Detection

\[ S(\Delta x, \Delta y) \approx \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \]

\[
M = \begin{bmatrix}
\sum_{(x_i, y_i) \in w} (I_x)^2 & \sum_{(x_i, y_i) \in w} (I_x I_y) \\
\sum_{(x_i, y_i) \in w} (I_x I_y) & \sum_{(x_i, y_i) \in w} (I_y)^2
\end{bmatrix}
\]

\[ S(\Delta x, \Delta y) = 1 \] actually is the ellipse equation.

The shape of the ellipse is determined by \( M \).
The “cornerness” of the window $w$ is reflected in $M$.

Suppose there are two local windows $w_1$ and $w_2$; consider the cases when the moving of the two windows leads to the intensity change equals to $1$. The moving vector $[\Delta x, \Delta y]$ of each window satisfies the ellipse equation. Thus,

For $w_1$,
$$\begin{bmatrix} \Delta x, \Delta y \end{bmatrix} M_1 \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = 1$$

For $w_2$,
$$\begin{bmatrix} \Delta x, \Delta y \end{bmatrix} M_2 \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = 1$$
Harris Corner Detection

Diagonalization of $M$: 

$$M = R^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R$$

The axis lengths of the ellipse are determined by the eigenvalues and the orientation is determined by $R$.
Interpreting the eigenvalues

Classification of image points using eigenvalues of $M$:

- **Corner**: $\lambda_1$ and $\lambda_2$ are large, $\lambda_1 \sim \lambda_2$; $S$ increases in all directions
- **Edge**: $\lambda_2 \gg \lambda_1$ and $\lambda_1 \gg \lambda_2$
- **Flat** region: $\lambda_1$ and $\lambda_2$ are small; $S$ is almost constant in all directions
Corner response function

Measure of corner response:

\[
R = \det \mathbf{M} - k (\text{trace} \mathbf{M})^2
\]

\[
\det \mathbf{M} = \lambda_1 \lambda_2 \\
\text{trace} \mathbf{M} = \lambda_1 + \lambda_2
\]

\( (k \text{ – empirical constant, } k = 0.04\text{-}0.06) \)
Harris corner detector--illustration

Ellipse with equation: $\begin{bmatrix} \Delta x, \Delta y \end{bmatrix} M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = 1$

- **Flat region**: both eigenvalues small
- **Edge**: one small, one large
Harris corner detector--illustration

Ellipse with equation: \[ [\Delta x, \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = 1 \]

corner
both eigenvalues large
Harris corner detector - Algorithm

1. Compute second moment matrix (autocorrelation matrix)

\[ M(\sigma_i, \sigma_D) = g(\sigma_i) \ast \begin{bmatrix} I_x^2(\sigma_D) & I_xI_y(\sigma_D) \\ I_xI_y(\sigma_D) & I_y^2(\sigma_D) \end{bmatrix} \]

2. Square of derivatives

3. Gaussian filter \( g(\sigma_i) \)

4. Cornerness function - two strong eigenvalues

\[ R = \det[M(\sigma_i, \sigma_D)] - \alpha[\text{trace}(M(\sigma_i, \sigma_D))]^2 \]

\[ = g(I_x^2)g(I_y^2) - [g(I_xI_y)]^2 - \alpha[g(I_x^2) + g(I_y^2)]^2 \]

5. Perform non-maximum suppression
Harris Detector: Steps
Harris Detector: Steps

Compute corner response $R$
Harris Detector: Steps

Find points with large corner response: $R > \text{threshold}$
Harris Detector: Steps

Take only the points of local maxima of $R$
Models of Image Change

Photometric
  - Affine intensity change \((l \rightarrow a \cdot l + b)\)

Geometric
  - Rotation
  - Scale
  - Affine
Harris Detector: Some Properties

Rotation invariance

Ellipse rotates but its shape (i.e. eigenvalues) remains the same

*Corner response* $R$ is invariant to image rotation
Harris Detector: Some Properties

Not invariant to *image scale*!

The underlying reason is that Harris corner detection scheme does not provide an automatic and appropriate window size selection method!
Local Descriptors for Harris Corners

- Descriptor for a Harris corner point
  - Take a region with a fixed size around it
  - Stack the region into a vector
  - This vector serves as the descriptor
  - When matching two descriptors in two different images, usually the correlation coefficient is used

\[
\begin{bmatrix}
v_1 \\
v_2 \\
\vdots \\
v_n
\end{bmatrix}
\]
Local Descriptors for Harris Corners

- Descriptor for a Harris corner point
  - Take a region with a fixed size around it
  - Stack the region into a vector
  - This vector serves as the descriptor
  - When matching two descriptors in two different images, usually the correlation coefficient is used

Correlation coefficient can be used to measure the similarity of two descriptors

\[ \rho = \frac{E[v_1 - E(v_1)][v_2 - E(v_2)]}{\sqrt{D(v_1)} \sqrt{D(v_2)}} \]
Local Descriptors for Harris Corners

- Descriptor for a Harris corner point
  - Take a region with a fixed size around it
  - Stack the region into a vector
  - This vector serves as the descriptor
  - When matching two descriptors in two different images, usually the correlation coefficient is used
- Deficiencies of such simple descriptors
  - Not rotation invariant
  - Not scale invariant

Lin ZHANG, SSE, 2019
Local Descriptors for Harris Corners

• Descriptor for a Harris corner point
  • Take a region with a fixed size around it
  • Stack the region into a vector
  • This vector serves as the descriptor
  • When matching two descriptors in two different images, usually the correlation coefficient is used

• We want:
  • Rotation and scale invariant feature points
  • Rotation and scale invariant feature descriptors
Content

- Local Invariant Features
  - Motivation
  - Requirements
  - Invariance
- Harris Corner Detector
- Scale Invariant Point Detection
  - Automatic scale selection
  - Laplacian-of-Gaussian detector
  - Difference-of-Gaussian detector
From Points to Regions

• The Harris corner detector defines interest points
  • Precise localization
  • High repeatability

• In order to match those points, we need to compute a descriptor over a region
  • How can we define such a region in a scale invariant manner?
  • That is how can we detect sale invariant regions?
Naïve Approach: Exhaustive Search

• Multi-scale procedure
  – Compare descriptors while varying the patch size
Scale Invariant Region Selection

Naïve Approach: Exhaustive Search

- Multi-scale procedure
  - Compare descriptors while varying the patch size
Scale Invariant Region Selection

Naïve Approach: Exhaustive Search

- Multi-scale procedure
  - Compare descriptors while varying the patch size

\[ d(f_A, f_B) \]

Slide credit: Krystian Mikolajczyk

Lin ZHANG, SSE, 2019
Scale Invariant Region Selection

Naïve Approach: Exhaustive Search

- Comparing descriptors while varying the patch size
  - Computationally inefficient
  - Inefficient but possible for matching
  - Prohibitive for retrieval in large databases
  - Prohibitive for recognition
What do we want to do next?

• Naïve approach for scale invariant local description is not efficient (Detect Harris corners first, and then exhaustively searching for regions with appropriate sizes)

• Now we want to:
  • Find scale invariant points in the image (location)
  • At the same time, we want to know their characteristic scales (used to determine the neighborhood for local description)
Achieving scale covariance

- Goal: independently detect corresponding regions in scaled versions of the same image
- Need *scale selection* mechanism for finding characteristic region size that is *covariant* with the image transformation
Automatic Scale Selection

- Solution:
  - Design a function on the region, which is “scale invariant” *(the same for corresponding regions, even if they are at different scales)*
  - For a point in one image, we can consider it as a function of region size (patch width)
Automatic Scale Selection

- Common approach
  - Take a local extremum of this function
  - Observation: region size for which the extremum is achieved should be covariant to image scale; this scale covariant region size is found in each image independently

![Graphs showing scale selection](image)
Automatic Scale Selection

- Function responses for increasing scale (scale signature)
Automatic Scale Selection

- Function responses for increasing scale (scale signature)
Automatic Scale Selection

- Function responses for increasing scale (scale signature)
Automatic Scale Selection

- Function responses for increasing scale (scale signature)
Automatic Scale Selection

- Function responses for increasing scale (scale signature)
Automatic Scale Selection

- Function responses for increasing scale (scale signature)

\[ f(I_{h\cdots i_{m}}(x, \sigma)) \]

\[ f(I_{h\cdots i_{m}}(x', \sigma')) \]
Automatic Scale Selection

- Normalize: Rescale to fixed size
Automatic Scale Selection

• A good function for scale selection
  • It should have one stable sharp peak response to region size
What is a useful signature function for scale?

- Laplacian-of-Gaussian = “blob” detector
We define the *characteristic scale* as the scale that produces peak of Laplacian response.

**Spatial selection**: the magnitude of the Laplacian response will achieve an extremum at the center of the blob, provided the scale of the Laplacian is “matched” to the scale of the blob.
Scale-Invariant Point Detection

- Interest points:
  - Local extremum in scale space of Laplacian of Gaussian

\[ L_{xx}(\sigma) + L_{yy}(\sigma) \]
Scale-Invariant Point Detection

- **Interest points:**
  - Local extremum in scale space of Laplacian of Gaussian

\[ L_{xx}(\sigma) + L_{yy}(\sigma) \]

Lin ZHANG, SSE, 2019
Scale-Invariant Point Detection

- Interest points:
  - Local extremum in scale space of Laplacian of Gaussian

\[ L_{xx}(\sigma) + L_{yy}(\sigma) \]
Scale-Invariant Point Detection

- Interest points:
  - Local extremum in scale space of Laplacian of Gaussian

\[ L_{xx} (\sigma) + L_{yy} (\sigma) \]

\[ \Rightarrow \text{List of } (x, y, \sigma) \]

(Positions of extrema in the scale-spatial space)
We have got want we want!

Note: local extrema is obtained by comparing the examined location with all the other 26 points around it in the scale-space.

If the local extrema of LoG is achieved at $p$, two things of $p$ can be determined: its spatial location and characteristic scale!
Scale normalization

• The response of a derivative of Gaussian filter to a perfect step edge decreases as $\sigma$ increases
• To keep response the same (scale-invariant), must multiply Gaussian derivative by $\sigma$
• Laplacian is the second Gaussian derivative, so it must be multiplied by $\sigma^2$
Blob detection in 2D

Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D

\[ \nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} , \quad g \text{ is the Gaussian function} \]
Blob detection in 2D

Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D

Scale-normalized:

\[
\nabla^2_{\text{norm}} g = \sigma^2 \left( \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} \right)
\]
Scale-Invariant Point Detection: Example
Scale-Invariant Point Detection: Example

\[
sigma = 11.9912
\]
Scale-Invariant Point Detection: Example
Efficient implementation

Approximating the Laplacian with a difference of Gaussians:

\[ L = \sigma^2 \left( G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma) \right) \]

(Laplacian)

\[ DoG = G(x, y, k\sigma) - G(x, y, \sigma) \]

(Difference of Gaussians)

where Gaussian is

\[ G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \]
DoG

• Difference of Gaussians as approximation of the LoG
  – This is used e.g. in Lowe’s SIFT pipeline for feature detection.

• Advantages
  – No need to compute 2\textsuperscript{nd} derivatives
  – Gaussians are computed anyway, e.g. in a Gaussian pyramid.
Scale-Invariant Point Detection

- **Given:** Two images of the same scene with a large *scale difference* between them.
- **Goal:** Find the same interest points *independently* in each image.
- **Solution:** Search for *maxima* of suitable functions in *scale* and in *space* (over the image).

- Two strategies
  - Laplacian-of-Gaussian (LoG)
  - Difference-of-Gaussian (DoG) as a fast approximation

  - *These can be used either on their own, or in combinations with single-scale keypoint detectors (Harris, Hessian).*
Examples
Examples

Interest points found by DoG extrema

What does the arrows mean?

*Next lecture!!*
Thanks for your attention