Keypoint Features: Scale Invariance and SIFT

CS 6384 Computer Vision
Professor Yu Xiang
The University of Texas at Dallas

Some slides of this lecture are courtesy Kris Kitani
Geometry-aware Feature Matching for Structure from Motion Applications. Shah et al, WACV’15

Applications: stereo matching, image stitching, 3D reconstruction, camera pose estimation, object recognition
Feature Detectors

• How to find image locations that can be reliably matched with images?
Harris Corner Detector

\[ f(\Delta x, \Delta y) \approx \sum_{x,y} w(x, y)(I_x(x, y)\Delta x + I_y(x, y)\Delta y)^2 \]

\[ f(\Delta x, \Delta y) \approx (\Delta x \quad \Delta y)M \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \]

\[ M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \sum_{x,y} w(x, y)I_x^2 & \sum_{x,y} w(x, y)I_x I_y \\ \sum_{x,y} w(x, y)I_x I_y & \sum_{x,y} w(x, y)I_y^2 \end{bmatrix} \]
Invariance

• Can the same feature point be detected after some transformation?
  • Translation invariance
    Are Harris corners translation invariance?
  • 2D rotation invariance
    Are Harris corners rotation invariance?
  • Scale invariance
    Are Harris corners scale invariance?

No
Scale Invariance

• Solution 1: detection features in all scales, matching features in corresponding scale (for small scale change)

Multi-scale oriented patches (MOPS) extracted at five pyramid levels (Brown, Szeliski, and Winder 2005)
Scale Invariance

• Solution 2: detect features that are stable in both location and scale

Intuition: Find local maxima in both position and scale

What filter can we use for scale selection?

Consider Harris corner detector
Recall Derivative Filter

Central difference

\[ f'(x) = \lim_{h \to 0} \frac{f(x + 0.5h) - f(x - 0.5h)}{h} \]

-1 0 1

X derivative

Find edge
Image Gradient

Gradient in x only: \( \nabla f = \left[ \frac{\partial f}{\partial x}, 0 \right] \)

Gradient in y only: \( \nabla f = \left[ 0, \frac{\partial f}{\partial y} \right] \)

Gradient in both x and y: \( \nabla f = \left[ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right] \)

Gradient direction: \( \theta = \tan^{-1} \left( \frac{\partial f}{\partial y} / \frac{\partial f}{\partial x} \right) \)

Gradient magnitude: \( ||\nabla f|| = \sqrt{\left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2} \)
Signal Noises

• Derivative filters are sensitive to noises

How to deal with noises?
Gaussian Filter

- Smoothing

1D

\[ g(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \]

2D

\[ g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]

Convolution \( h \ast f \)

Derivative \( \frac{\partial}{\partial x}(h \ast f) \)

Peak = edge location
Derivative of Gaussian Filter

- Derivative Theorem of Convolution
  \[
  \frac{\partial}{\partial x}(h \ast f) = \left(\frac{\partial}{\partial x} h\right) \ast f
  \]

Smoothing and derivative

[Graphs showing the effects of smoothing and derivative operations on a signal]
Derivative of Gaussian Filter

- Derivative Theorem of Convolution
  \[
  \frac{\partial}{\partial x} (h \ast f) = \left( \frac{\partial}{\partial x} h \right) \ast f
  \]

  \[
  g_x(x,y) = \frac{\partial g(x,y)}{\partial x} = \frac{-x}{2\pi \sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}}
  \]

  \[
  g_y(x,y) = \frac{\partial g(x,y)}{\partial y} = \frac{-y}{2\pi \sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}}
  \]

  \[
  g(x, y) = \frac{1}{2\pi \sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}
  \]
Laplace Filter

first-order finite difference

\[ f'(x) = \lim_{h \to 0} \frac{f(x + 0.5h) - f(x - 0.5h)}{h} \]

Derivative filter

-1 0 1

second-order finite difference

\[ f''(x) \approx \frac{\delta_h^2 [f](x)}{h^2} = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = \frac{f(x + h) - 2f(x) + f(x - h)}{h^2} \]

Laplace filter

1 -2 1
Laplace Filter

- 2D

\[ \nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \]

1D Laplace filter

| 1 | -2 | 1 |

2D Laplace filter

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Laplacian of Gaussian Filter

\[ \nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \]

\[ \nabla^2 I \circ g = \nabla^2 g \circ I \]

\[ \nabla^2 g = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} g(x, y) \]

Smoothing and second derivative

Mexican Hat Function

Laplacian of Gaussian
Laplacian of Gaussian Filter

\[ f \]

\[ \frac{\partial}{\partial x} h \]

Derivative of Gaussian

\[ \left( \frac{\partial}{\partial x} h \right) \ast f \]

Zero crossings

\[ \frac{\partial^2}{\partial x^2} h \]

Laplacian of Gaussian

\[ \left( \frac{\partial^2}{\partial x^2} h \right) \ast f \]
Laplacian of Gaussian for Scale Selection

Highest response when the signal has the same characteristic scale as the filter
Laplacian of Gaussian for Scale Selection

characteristic scale

Search over different scales $\sigma$

Laplacian of Gaussian

$\nabla^2 h_\sigma(u,v)$
Laplacian of Gaussian for Scale Selection

Multi-scale 2D Blob detection
Laplacian of Gaussian for Scale Selection

cross-scale maximum

local maximum

local maximum

local maximum
Scale Invariance Feature Transform (SIFT)

• Keypoint detection

• Compute descriptors

• Matching descriptors

David Lowe, Distinctive Image Features from Scale-Invariant Keypoints. IJCV, 2004
SIFT: Scale-space Extrema Detection

• Difference of Gaussian (DoG)

\[ G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \]

\[ L(x, y, \sigma) = G(x, y, \sigma) \ast I(x, y) \]

\[ D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \ast I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma). \]

Approximate of Laplacian of Gaussian (efficient to compute)
SIFT: Scale-space Extrema Detection

- Gaussian pyramid

- Gaussian filters

\[ L(x, y, \sigma) = G(x, y, \sigma) \ast I(x, y) \]

\[ G_{\sigma_1} \ast G_{\sigma_2} = G_{\sigma} \quad \sigma^2 = \sigma_1^2 + \sigma_2^2 \]

- Sub-sampling by a factor of 2
  - Multiple the Gaussian kernel deviation by 2
SIFT: Scale-space Extrema Detection

Maxima and minima of DoG images

\[ L(x, y, \sigma) = G(x, y, \sigma) \ast I(x, y) \]
\[ G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \]
\[ D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \ast I(x, y) \]
\[ = L(x, y, k\sigma) - L(x, y, \sigma). \]
SIFT Descriptor

• Image gradient magnitude and orientation

\[ L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \]

\[
m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2}
\]

\[
\theta(x, y) = \tan^{-1}((L(x, y + 1) - L(x, y - 1))/(L(x + 1, y) - L(x - 1, y)))
\]
SIFT Descriptor

- Divide the 16x16 window into a 4x4 grid of cells (2x2 case shown below)
- Compute an orientation histogram for each cell
- 16 cells * 8 orientations = 128 dimensional descriptor

Using the scale of the keypoint to select the level of Gaussian blur for the image
SIFT: Rotation Invariance

• Rotate all orientations by the dominant orientation
SIFT: Rotation Invariance

• Rotate all orientations by the dominant orientation
SIFT Properties

• Can handle change in viewpoint (up to about 60 degree out of plane rotation)

• Can handle significant change in illumination

• Relatively fast < 1s for moderate image sizes

• Lots of code available
  • E.g., https://www.vlfeat.org/overview/sift.html
SIFT Matching Example

https://www.vlfeat.org/overview/sift.html
SIFT Matching Example
Further Reading

• Section 7.1, Computer Vision, Richard Szeliski

• David Lowe, Distinctive Image Features from Scale-Invariant Keypoints. IJCV, 2004

• ORB: An efficient alternative to SIFT or SURF. Rublee et al., ICCV, 2011