Scale-Invariant Feature Transform

Hyun Soo Park
**Feature Matching**
Recall: HOG

\[ corr \begin{bmatrix} \text{left} \\ \text{right} \end{bmatrix} = 0.15 \]

What’s wrong?
CHALLENGES OF FEATURE MATCHING

- Scale
- Orientation
FEATURE NORMALIZATION
FEATURE NORMALIZATION
**Feature Normalization**

Scale normalization
**FEATURE NORMALIZATION**

Orientation normalization
**Scale-invariant Feature Transform (SIFT)**

```latex
\begin{align*}
\text{corr} &= 0.15 \\
\text{corr} &= 0.91
\end{align*}
```
SCALE-INVARINANT FEATURE TRANSFORM (SIFT)

1. Input image
**Scale-invariant Feature Transform (SIFT)**

2. Keypoint detection via scale space extrema
3. Orientation assignment
SCALE-INvariant FEATURE Transform (SIFT)

4. Descriptor extraction
**Keypoint Detection: Scale Space Extrema**

$L = \sigma^2 \nabla^2 G$

Laplacian of Gaussian (LoG)
KEYPOINT DETECTION: SCALE SPACE EXTREMA

$I(x, y)$
**Keypoint Detection: Scale Space Extrema**

\[ I(x, y) \ast L(\sigma) \]
**Keypoint Detection: Scale Space Extrema**

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

Scale space response
**RECALL: LAPLACIAN OF GAUSSIAN (LOG) ~ DOG**

\[ x(t) \ast (g(t;\sigma_1) - g(t;\sigma_2)) \approx \nabla \cdot \nabla g \]
\[ \text{Laplacian of Gaussian} \]

\[ X(f)(G(f;\sigma_1) - G(f;\sigma_2)) \]
RECALL: LAPLACIAN OF GAUSSIAN (LoG) ~ DoG
Laplacian ~ Difference of Gaussian
WHAT CHARACTERIZES KEYPOINT?

- Robust
- Repeatable
- Uniqueness

Local extrema (minima/maxima)
WHAT CHARACTERIZES KEYPOINT?

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

Scale space response

Local extrema (minima/maxima)

Characteristic scale
WHAT CHARACTERIZES KEYPOINT?

\[ D(x, y, \sigma + \Delta \sigma) \]

\( (x, y) \) is keypoint if

\[ D(x, y, \sigma) > D(x \pm \Delta x, y \pm \Delta y, \sigma \pm \Delta \sigma) \]

or

\[ D(x, y, \sigma) < D(x \pm \Delta x, y \pm \Delta y, \sigma \pm \Delta \sigma) \]
Subpixel Keypoint Localization

\[ D(x, y, \sigma) \]

Discrete local extremum
Subpixel Keypoint Localization

\[ D(x, y, \sigma) \]

Continuous local extremum
Discrete local extremum

\[ f(x) \]
**Subpixel Keypoint Localization**

Continuous local extremum
Discrete local extremum

\[
D(x, y, \sigma)
\]

\[
f(x + \Delta x) = f(x) + f'(\Delta x) + \frac{1}{2} f''(\Delta x^2)
\]

\[
\Delta x = -\frac{f'}{f''}
\]
**Subpixel Keypoint Localization**

Newton's method

\[ f(x + \Delta x) = f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)\Delta x^2 \]

\[ \Delta x = -\frac{f'}{f''} \]

\[ D(x + \Delta x) = D(x) + \frac{\partial D}{\partial x} \Delta x + \frac{1}{2} \Delta x^T \frac{\partial^2 D}{\partial x^2} \Delta x \]

Gradient

Hessian

\[ \Delta x = -\left( \frac{\partial^2 D}{\partial x^2} \right)^{-1} \frac{\partial D}{\partial x} \]
**Thresholding: Low Contrast**

\[ |D(x)| < 0.03 \]
**Thresholding: Edge**

Principal curvatures are eigenvalues of Hessian matrix:

\[
H = \begin{bmatrix}
  D_{xx} & D_{xy} \\
  D_{yx} & D_{yy}
\end{bmatrix}
\]

\[\lambda_1 \approx \lambda_2\]

\[\lambda_1 \gg \lambda_2\]
**Thresholding: Edge**

Principal curvatures are eigenvalues of Hessian matrix:

\[
H = \begin{bmatrix}
D_{xx} & D_{xy} \\
D_{yx} & D_{yy}
\end{bmatrix}
\]

\[\lambda_1 \approx \lambda_2\]

\[\lambda_1 \gg \lambda_2\]
Figure 5: This figure shows the stages of keypoint selection. (a) The 233x189 pixel original image. (b) The initial 832 keypoints locations at maxima and minima of the difference-of-Gaussian function. Keypoints are displayed as vectors indicating scale, orientation, and location. (c) After applying a threshold on minimum contrast, 729 keypoints remain. (d) The final 536 keypoints that remain following an additional threshold on ratio of principal curvatures.
**Orientation Assignment**

- Peak orientation ~ dominant orientation
- Histogram of orientation weighted by gradient magnitude and Gaussian
$\|I(x,y) * \nabla G(\sigma)\|

$\angle (I(x,y) * \nabla G(\sigma))$

Histogram of orientation weighted by gradient magnitude and Gaussian

**Orientation Normalization**
**Scale-invariant Feature Transform (SIFT)**

\[
\text{corr} = 0.15
\]

\[
\text{corr} = 0.91
\]
SCALE-IN Variant FEATURE Transform (SIFT)

2. Keypoint detection via scale space extrema
3. Orientation assignment
$\| I(x,y) \ast \nabla G(\sigma) \|$
**Scale-invariant Feature Transform (SIFT)**

4. Descriptor extraction