Advanced Image Analysis – 02506
Local Image Features – Spring 2020

Professor Anders Bjorholm Dahl
DTU Compute

Advanced Image Analysis

February 2020

SIFT

Matching

Transformation

Visualization

Exercise

Reading material

- David Lowe: Distinctive Image Features from Scale-Invariant Keypoints, (IJCV 2004).

Local Image Features

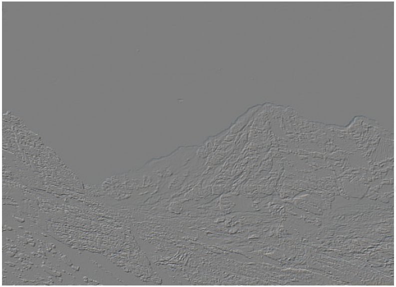
Leaning objectives

- Image features (detector and descriptor)
- Matching of image features
- Computing rotation, translation and scale of 2D point sets
- Visualization of feature matching

Similarity

Pixel-wise comparison

- Shift of single pixel vs. two views



Similarity

Basic idea

- Locally appearance between views is the same
- Variation can be handled via normalization

Local image features

SIFT – key elements

- Features localized at interest points
- Adapted to scale – based on scale-space theory
- Invariance to appearance changes in the image

SIFT – scale invariant feature transform (Lowe, 2004)

- Scale-space blob detection – difference of Gaussians
- Interest point localization
- Orientation assignment
- Interest point descriptor
- Note – SIFT is one example of interest point feature

SIFT

Scale invariance

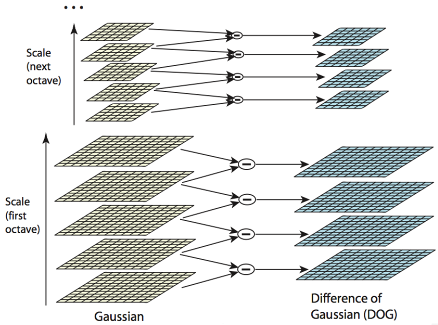
- Using a difference of Gaussians for blob detection

$$
\begin{aligned}
D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) && (1) \\
&= L(x, y, k\sigma) - L(x, y, \sigma) && (2)
\end{aligned}
$$

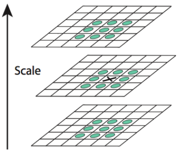
- The scale normalized Laplacian is given by (note that the DoG is scale normalized)

$$
\sqrt{t}\nabla^2 G = \sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \qquad (3)
$$

SIFT – Estimation of DoG



(a) Difference of Gaussians



(b) Extrema localization

SIFT – Interest point localization

- Setting the derivative of D to zero

$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2}\frac{\partial D}{\partial \mathbf{x}} \tag{4}$$

- We get

$$D(\hat{\mathbf{x}}) = D + \frac{1}{2}\frac{\partial D^T}{\partial \mathbf{x}}\hat{\mathbf{x}} \tag{5}$$

- $|D(\hat{\mathbf{x}})| > 0.03$ otherwise the point is discharged

SIFT – Interest point along edges discharted

- The eigenvalues of the Hessian are proportional to the principal curvatures

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial x \partial y} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} \tag{6}$$

Interest point along edges discharted

- α and β are the eigenvalues of the Hessian

$$\mathrm{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta \tag{7}$$

$$\mathrm{Det}(\mathbf{H}) = D_{xx} D_{yy} - D_{xy}^2 = \alpha\beta \tag{8}$$

- Points are kept if

$$\frac{\mathrm{Tr}(\mathbf{H})^2}{\mathrm{Det}(\mathbf{H})} < \frac{(r+1)^2}{r} \tag{9}$$

- where $r = 10$ (found to be a good heuristic)

SIFT – Orientation assigment

- Build a histogram of gradient magnitudes and orientations

$$m(x, y) = \sqrt{L_x^2 + L_y^2} \tag{10}$$

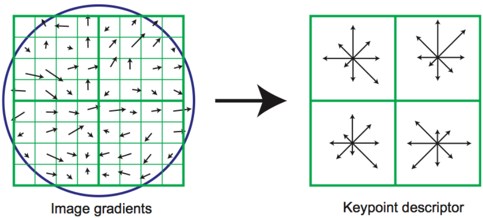
- Where $L_x = L(x+1, y) - L(x-1, y)$ and $L_y = L(x, y+1) - L(x, y-1)$

$$\theta(x, y) = \tan^{-1} \frac{L_y}{L_x} \tag{11}$$

- 36 orientations are used – local patch sampled according to dominant orientation

SIFT – Descriptor

- Build a histogram of local gradient orientations
- Normalized using L_2 norm: $\mathbf{d}_n = \frac{1}{\sqrt{\sum_{i=1}^{128} \mathbf{d}(i)^2}} \mathbf{d}$



SIFT – Matching of detscriptors

- Use Euclidian distance between normalized vectors

$$\delta(\mathbf{d}_i, \mathbf{d}_j) = \sqrt{\sum_{n=1}^{128} (\mathbf{d}_i(n) - \mathbf{d}_j(n))^2}$$

- Note – for comparison the square root is not needed
- Use vector angles

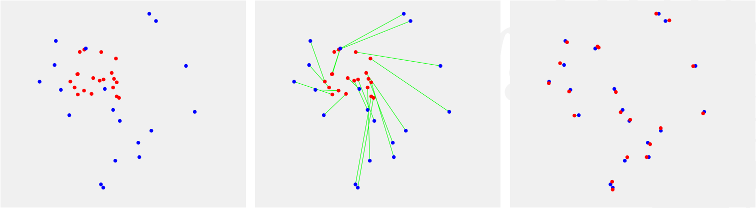
$$\theta = \arccos\left(\frac{\mathbf{d}_i^T \mathbf{d}_j}{||\mathbf{d}_i|| ||\mathbf{d}_j||}\right)$$

SIFT – Summary

- Image in pixel space $r \times c \rightarrow$ image in 128-dimensional SIFT space
- Allows matching of images invariant to: Scale, illumination, viewpoint
- Partly visible objects can be matched

Least squares transformation

- Rotation $\mathbf{R}$
- Translation $\mathbf{t}$
- Scale s



Least squares transformation

- Scale

$$s = \frac{\sum_{i=1}^{n} ||\mathbf{q} - \mu_q||}{\sum_{i=1}^{n} ||\hat{\mathbf{p}} - \mu_p||}$$

Least squares transformation

- Rotation and translation

$$\arg\min_{\mathbf{R},\mathbf{t}} \sum_{i=1}^{n} ||\mathbf{q}_i - \mathbf{R}\mathbf{p}_i - \mathbf{t}||$$

- Covariance of centered point sets

$$\mathbf{C} = \sum_{i=1}^{n} (\mathbf{p}_i - \mu_p)(\mathbf{q}_i - \mu_q)^T$$

- SVD

$$\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{C}$$

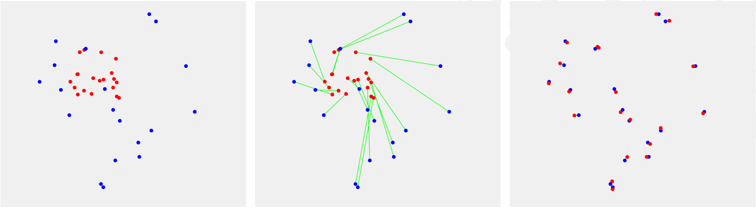
Least squares transformation

- Compute rotation

$$\mathbf{R} = \mathbf{U}\mathbf{V}^T$$

- Compute translation

$$\mathbf{t} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{q}_i - \mathbf{R}\mathbf{p}_i) = \mu_q - \mathbf{R}\mu_p$$



Exercise

1. Function for rotation, translation and scale
2. Compute and match SIFT
3. Transform the matched features

Local Image Features

Questions for the exam

- Explain the principle of feature-based image registration and how SIFT can be used for feature matching.