Robot Vision: Features

Prof. Friedrich Fraundorfer

SS 2024
Outline

- The importance of feature matching
- Image similarity and viewpoint changes
- Challenges
- Properties of detectors and descriptors
- Detectors
  - What locations would be good
  - Point detectors (concept of Harris and FAST)
  - Blob detectors (DOG)
Image features

- The term “feature” or “image feature” is used with some variety of meaning.

- Set of properties, description of an image region (in this case including a specific location) or the whole image

- Strictly speaking the term “feature” only means a description, but any description needs a location. So the wider definition also means a location and region

- “Feature points” are the detected point locations in images that are used for image matching or geometric algorithms.

- Image features are a combination of the results of a detector method and a descriptor method.
The importance of feature matches

- Geometric algorithms need point correspondences i.e. image feature matches
- The quality of feature matches determines the outcome of geometric algorithms.
  - Location accuracy of feature matches
  - Correctness of feature matches (mis-matches)
- Image classification, image indexing, image search, image interpretation also need feature points and feature matches.
Image similarity and viewpoint changes
Image similarity and viewpoint changes
Two challenges

- How to select proper points (detectors)
- How to compute the similarity of image patches (descriptor)
Properties of detectors

- Accurate localization
- Useful locations
- High repeatability detection
Properties of descriptors

- Discriminative
- Descriptive
- Compact descriptions
- Invariance to image changes (brightness, rotation)
Detectors: Which locations would be good

- homogeneous area
- edge
- corner
Detectors

- Point detectors
  - Harris corners
  - FAST corners
- Blob detectors
  - DOG points
Harris corners

- Looks for locations in an image where the SSD changes strongly

\[ f(x, y) = \sum_{(x_k, y_k) \in W} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2 \]

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

\[ M = \sum_{(x, y) \in W} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \sum_{(x, y) \in W} I_x^2 \\ \sum_{(x, y) \in W} I_x I_y \\ \sum_{(x, y) \in W} I_x I_y \\ \sum_{(x, y) \in W} I_y^2 \end{bmatrix} \]
FAST corners

- Count the number $N$ of contiguous pixels around a center pixel $p$ that are brighter than the center pixel. If $N \geq$ some threshold this point is a feature location.
Harris corners vs. Fast corners
Harris corners vs. Fast corners

- Slower to compute
- Better control of number of detections with threshold

- Fast to compute
- Many detections
- Many corners next to each other
Difference of Gaussian (DOG) points

- Is a Blob detector, detections are not necessarily on image corners
- Is a scale invariant detector, high repeatability even for images of different scales (image resolution)
- Processes images at different resolutions (scales) and then selects a feature location in $x,y$ and a specific scale $s$ which has a high value for the sum of the squares of the second derivatives in all directions (Laplacian)
DOG filter mask

- Filter mask is composed of the subtraction of two Gaussian filter masks
- Is an approximation of the Laplace operator (Laplacian of Gaussian, LOG) which is a blob detector

\[
\text{DOG}(x, y) = \frac{1}{k} e^{-\frac{x^2 + y^2}{(k\sigma)^2}} - e^{-\frac{x^2 + y^2}{\sigma^2}}
\]
Computation of DOG’s measure

<table>
<thead>
<tr>
<th>first octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_0 \cdot k^5$</td>
</tr>
<tr>
<td>$\sigma_0 \cdot k^4$</td>
</tr>
<tr>
<td>$\sigma_0 \cdot k^3$</td>
</tr>
<tr>
<td>$\sigma_0 \cdot k \cdot k$</td>
</tr>
<tr>
<td>$\sigma_0 \cdot k$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>second octave</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2 \cdot \sigma_0 \cdot k^5$</td>
</tr>
<tr>
<td>$2 \cdot \sigma_0 \cdot k^4$</td>
</tr>
<tr>
<td>$2 \cdot \sigma_0 \cdot k^3$</td>
</tr>
<tr>
<td>$2 \cdot \sigma_0 \cdot k \cdot k$</td>
</tr>
<tr>
<td>$2 \cdot \sigma_0 \cdot k$</td>
</tr>
</tbody>
</table>

downsample (half resolution)

Gauss filtered images

Difference of Gauss filtered images
Selection of extrema

- Extrema are selected in 3D (x,y,scale)
- Center pixel needs to be larger or smaller than it’s 26 neighbors

https://aishack.in/tutorials/sift-scale-invariant-feature-transform-introduction/
DOG feature points