CS376 Computer Vision
Lecture 13: Invariant Descriptors

Qixing Huang
March 6th 2019

Slide Credit: Kristen Grauman
Recap

• Harris corner detector

• Scale-invariant feature detector
Recall: Harris corner detector

\[ M = \sum w(x, y) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix} \]

1) Compute \( M \) matrix for each image window to get their \textit{cornerness} scores.
2) Find points whose surrounding window gave large corner response (\( f > \) threshold)
3) Take the points of local maxima, i.e., perform non-maximum suppression
Recall: Harris Detector: Steps
Scale-space blob detector: Example
Local features: main components

1) Detection: Identify the interest points

2) Description: Extract vector feature descriptor surrounding each interest point.

3) Matching: Determine correspondence between descriptors in two views

\[ x_1 = [x_1^{(1)}, K, x_d^{(1)}] \]

\[ x_2 = [x_1^{(2)}, K, x_d^{(2)}] \]

Slide credit: Kristen Grauman
Geometric transformations
e.g. scale, translation, rotation
Photometric transformations

Figure from T. Tuytelaars ECCV 2006 tutorial
Raw patches as local descriptors

The simplest way to describe the neighborhood around an interest point is to write down the list of intensities to form a feature vector.

But this is very sensitive to even small shifts, rotations.

Figure: Andrew Zisserman
Scale Invariant Feature Transform (SIFT) descriptor [Lowe 2004]

- Use histograms to bin pixels within sub-patches according to their orientation.

---

Final descriptor = concatenation of all histograms, normalize
Idea of SIFT

- Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters
Scale Invariant Feature Transform (SIFT) descriptor [Lowe 2004]

Interest points and their scales and orientations (random subset of 50)

SIFT descriptors

http://www.vlfeat.org/overview/sift.html
Making descriptor rotation invariant

- Rotate patch according to its dominant gradient orientation
- This puts the patches into a canonical orientation.
SIFT descriptor [Lowe 2004]

- Extraordinarily robust matching technique
  - Can handle changes in viewpoint
    - Up to about 60 degree out of plane rotation
  - Can handle significant changes in illumination
    - Sometimes even day vs. night (below)
  - Fast and efficient—can run in real time
  - Lots of code available, e.g. [http://www.vlfeat.org/overview/sift.html](http://www.vlfeat.org/overview/sift.html)
SIFT properties

• Invariant to
  – Scale
  – Rotation

• Partially invariant to
  – Illumination changes
  – Camera viewpoint
  – Occlusion, clutter
Example

NASA Mars Rover images
Example

NASA Mars Rover images with SIFT feature matches
Figure by Noah Snavely
Local features: main components

1) Detection: Identify the interest points

2) Description: Extract vector feature descriptor surrounding each interest point.

3) Matching: Determine correspondence between descriptors in two views

Slide credit: Kristen Grauman
Matching local features
Matching local features

To generate **candidate matches**, find patches that have the most similar appearance (e.g., lowest SSD)

Simplest approach: compare them all, take the closest (or closest k, or within a thresholded distance)

Slide credit: Kristen Grauman
Ambiguous matches

At what SSD value do we have a good match?

To add robustness to matching, consider **ratio**:

\[
\text{dist to best match} / \text{dist to second best match}
\]

If low, first match looks good.

If high, could be ambiguous match.

Image 1

Image 2

Slide credit: Kristen Grauman
Matching SIFT Descriptors

- Nearest neighbor (Euclidean distance)
- Threshold ratio of nearest to 2nd nearest descriptor

Lowe IJCV 2004
Scale Invariant Feature Transform (SIFT) descriptor [Lowe 2004]

Interest points and their scales and orientations (random subset of 50)

SIFT descriptors

http://www.vlfeat.org/overview/sift.html
SIFT (preliminary) matches

http://www.vlfeat.org/overview/sift.html
Value of local (invariant) features

- Complexity reduction via selection of distinctive points
- Describe images, objects, parts without requiring segmentation
  - Local character means robustness to clutter, occlusion
- Robustness: similar descriptors in spite of noise, blur, etc.
Applications of local invariant features

- Wide baseline stereo
- Motion tracking
- Panoramas
- Mobile robot navigation
- 3D reconstruction
- Recognition
- ...
Automatic mosaicing

Matthew Brown
http://matthewalunbrown.com/autostitch/autostitch.html
Wide baseline stereo

[Image from T. Tuytelaars ECCV 2006 tutorial]
Photo tourism [Snavely et al.]
Recognition of specific objects, scenes

Scale

Viewpoint

Lighting

Occlusion
Google Goggles
Summary

• Interest point detection
  – Harris corner detector
  – Laplacian of Gaussian, automatic scale selection

• Invariant descriptors
  – Rotation according to dominant gradient direction
  – Histograms for robustness to small shifts and translations (SIFT descriptor)
Coming up

• Additional questions we need to address to achieve these applications:
• Fitting a parametric transformation given putative matches
• Dealing with outlier correspondences
• Exploiting geometry to restrict locations of possible matches
• Triangulation, reconstruction
• Efficiency when indexing so many keypoints