Invariant Local Features

Tuesday, February 6

Invariant local features

Subset of local feature types designed to be invariant to
- Scale
- Translation
- Rotation
- Affine transformations
- Illumination

1) Detect distinctive interest points
2) Extract invariant descriptors

[Good] invariant local features

• Reliably detected
• Distinctive
• Robust to noise, blur, etc.
• Description normalized properly

Classes of transformations

• Euclidean/rigid: Translation + rotation
• Similarity: Translation + rotation + uniform scale
• Affine: Similarity + shear
• Projective: Affine + projective warps

Case study: panorama stitching

[These slides are from Darya Frolova and Denis Simakov]

How do we build panorama?

• We need to match (align) images

[These slides are from Darya Frolova and Denis Simakov]
Matching with Features
• Detect feature points in both images

Matching with Features
• Detect feature points in both images
• Find corresponding pairs

Matching with Features
• Detect feature points in both images
• Find corresponding pairs
• Use these pairs to align images

Matching with Features
• Problem 1:
  – Detect the same point independently in both images

We need a repeatable detector

Matching with Features
• Problem 2:
  – For each point correctly recognize the corresponding one

We need a reliable and distinctive descriptor

Interest operators: an introductory example
Harris corner detector

The Basic Idea

- We should easily recognize the point by looking through a small window
- Shifting a window in any direction should give a large change in intensity

Harris Detector: Basic Idea

“flat” region: no change in all directions
“edge”: no change along the edge direction
“corner”: significant change in all directions

Harris Detector: Basic Idea

Corner: significant change in all directions.

Harris Detector: Mathematics

Window-averaged change of intensity for the shift \([u, v]\):

\[
E(u, v) = \sum_{x,y} w(x, y) \left[ I(x+u, y+v) - I(x, y) \right]^2
\]

Window function \(w(x, y)\):

- 1 in window, 0 outside
- Gaussian

Harris Detector: Mathematics

A bilinear approximation for average intensity change for small shifts in direction \([u, v]\):

\[
E(u, v) \equiv \begin{bmatrix} u \\ v \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix}
\]

where \(M\) is a 2x2 matrix computed from image derivatives:

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

Intensity change in shifting window: eigenvalues tell us how intensity changes in different directions

\[
E(u, v) \equiv \begin{bmatrix} u \\ v \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix}
\]

\(\lambda_1, \lambda_2\) – eigenvalues of \(M\)
Classification of image points using eigenvalues of $M$:

- **Corner**: $\lambda_1$ and $\lambda_2$ are large, $\lambda_1 \approx \lambda_2$; $E$ increases in all directions.
- **Edge**: $\lambda_1 >> \lambda_2$ or $\lambda_2 >> \lambda_1$; $E$ is almost constant in all directions.
- **Flat** region: $\lambda_1$ and $\lambda_2$ are small; $E$ is almost constant in all directions.

Measure of corner response:

$$R = \text{det} M - k (\text{trace } M)^2$$

$\text{det } M = \lambda_1 \lambda_2$

$\text{trace } M = \lambda_1 + \lambda_2$

(k is empirical constant, $k = 0.04-0.06$)

• The Algorithm:
  - Find points with large corner response function $R$ ($R > \text{threshold}$)
  - Take the points of local maxima of $R$
**Harris Detector: Workflow**

Find points with large corner response: $R > \text{threshold}$

**Harris Detector: Workflow**

Take only the points of local maxima of $R$

---

**Harris Detector: Some Properties**

- Rotation invariance

  ![Diagram](https://example.com/diagram.png)

  Ellipse rotates but its shape (i.e. eigenvalues) remains the same

  *Corner response $R$ is invariant to image rotation*

**Harris Detector: Some Properties**

- Not invariant to *image scale*!

  ![Diagram](https://example.com/diagram.png)

  All points will be classified as *edges*

---

**Scale Invariant Detection**

![Images from T. Tuytelaars](https://example.com/images.png)
Scale Invariant Detection

- Consider regions (e.g. circles) of different sizes around a point
- Regions of corresponding content will look the same in both images

Scale Invariant Detection

- The problem: how do we choose corresponding circles independently in each image?

Scale Invariant Detection

- Solution:
  - Design a function on the region (circle), which is "scale invariant" (the same for corresponding regions, even if they are at different scales)
  Example: average intensity. For corresponding regions (even of different sizes) it will be the same.
  - For a point in one image, we can consider it as a function of region size (circle radius)

Scale Invariant Detection

- Common approach:
  Take a local maximum of this function
  Observation: region size, for which the maximum is achieved, should be invariant to image scale.

  Important: this scale invariant region size is found in each image independently!

Scale Invariant Detection

- A "good" function for scale detection: has one stable sharp peak

- For usual images: a good function would be a one which responds to contrast (sharp local intensity change)

Scale Invariant Detection

Functions for determining scale

Kernels:

- L = σ²(G(x, y, σ) + G(x, y, σ))
  (Laplacian)
- DoG = G(x, y, kσ) - G(x, y, σ)
  (Difference of Gaussians)

where Gaussian

\[ G(x, y, σ) = \frac{1}{2\piσ^2} e^{-\frac{x^2+y^2}{2σ^2}} \]

Note: both kernels are invariant to scale and rotation.
Scale Invariant Detectors

1. Harris-Laplacian
   - Find local maximum of:
     - Harris corner detector in space (image coordinates)
     - Laplacian in scale

2. SIFT (Lowe)
   - Find local maximum of:
     - Difference of Gaussians in space and scale

Scale Invariant Detection: Summary

- **Given:** two images of the same scene with a large scale difference between them
- **Goal:** find the same interest points independently in each image
- **Solution:** search for maxima of suitable functions in scale and in space (over the image)

Affine Invariant Detection

- **Intensity-based regions (IBR):**
  - Start from a local intensity extremum
  - Consider intensity profile along rays
  - Select maximum of f(t) along each ray
  - Connect local maxima
  - Fit an ellipse

Affine Invariant Detection

- **Maximally Stable Extremal Regions (MSER):**
  - Threshold image intensities:
    - |I| > ℓ₀
  - Extract connected components ("Extremal Regions")
  - Seek extremal regions that remain "Maximally Stable" under range of thresholds
Point Descriptors

• We know how to detect points
• Next question: How to describe them for matching?

Point descriptor should be:
1. Invariant
2. Distinctive

Rotation Invariant Descriptors

• Harris corner response measure:
  depends only on the eigenvalues of the matrix $M$

Rotation Invariant Descriptors

• Find local orientation
  Dominant direction of gradient

• Rotate description relative to dominant orientation

Rotation Invariant Descriptors

• Use the scale determined by detector to compute descriptor in a normalized frame

Affine Invariant Descriptors

• Compute rotation invariant descriptor in the affine normalized frame (deskew)

Applications

• Wide baseline stereo
• Motion tracking
• Panoramas
• Mobile robot navigation
• 3D reconstruction
• Recognition
  – Specific objects
  – Textures
  – Categories
• …
Wide baseline stereo

Panorama stitching

Recognition of specific objects, scenes

Recognition of categories

Comparative evaluations

Issues

- For specific-level recognition — scaling the search?
  - Complexity
  - Distinctiveness

- For category-level recognition — are features most appropriate?
  - Sparse
  - Strict appearance description
  - Texture vs. shape

- Expense of detecting interest points