Image Registration
2D, 3D, Rigid and Deformable Scenes

Adrien Bartoli
CNRS – LASMEA
Clermont-Ferrand, France

Lecture 1.05 – 2D Configurations I:
Feature-Based Homography Estimation
[Torr and Zisserman, WVA’99]

with some images and panoramas courtesy of
S. Hauberg, J. Sloth and D. R. Jørgensen
Roadmap

- The homographic model
  - Algebraic formulation
  - In which cases does it apply?
- Feature-based estimation
  - Which features can be used?
  - Keypoint detection
  - Keypoint matching (using appearance)
  - Robust homography estimation
  - Guided keypoint matching
The Homographic Model: Algebraic Formulation

Formulation for points

\[ q' \sim Hq \quad \lambda \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} H_{1,1} & H_{1,2} & H_{1,3} \\ H_{2,1} & H_{2,2} & H_{2,3} \\ H_{3,1} & H_{3,2} & H_{3,3} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \]
The Homographic Model: Algebraic Formulation

- Based on homogeneous coordinates
- 8 degrees of freedom (3$\times$3-1)
- Defined by 4 pairs of points
- Linear for the homogeneous coordinates
- Nonlinear for the affine coordinates

\[
\begin{pmatrix}
  x' \\
  y'
\end{pmatrix} = \frac{1}{H_{3,1}x + H_{3,2}y + H_{3,3}} \begin{pmatrix}
  H_{1,1}x + H_{1,2}y + H_{1,3} \\
  H_{2,1}x + H_{2,2}y + H_{2,3}
\end{pmatrix}
\]
In Which Cases Does it Apply?

- **Perspective camera**
  - Simplifies to affine transformations for affine cameras
  - Does not hold in the presence of nonlinear optical effects such as radial distortion

- **Two cases**
  - Camera motion: pure rotation ; Scene: arbitrary
  - Camera motion: arbitrary ; Scene: planar
The Rotating Camera Case
The Planar Scene Case

Planar scene

Camera 1

Camera 2
Which Features can be Used?

- Not all pixels can be used in feature-based approaches.
- Choose those points which are as salient as possible, to ease the matching task.
- These points are edge or corner points, i.e. points where the intensity varies significantly.

Original image  Canny edges  Förstner-Harris interest points
What are Keypoints?

- **Intuitively**
  - Junction of contours
  - Point with significant variation in their neighborhood

- **Stable features over changes of viewpoint, in general**

- Therefore, we should be able to find such points by looking through ‘small’ windows

- Also termed: corner points, interest points and salient points
Basis Idea for Keypoint Detection

- Move a small window around each image pixel
- No change $\rightarrow$ flat, homogeneous region
- Change along one direction $\rightarrow$ edge
- Significant change in all directions $\rightarrow$ corner

Some other methods:
- Parametric image model fitting
- Local energy / phase congruency
- Morphology
Example

Original image $\mathcal{I}$

Cornerness

Keypoints (1120)
The Windowing Function and the Scale Selection Issue

\[ \sigma = 0.5 \]
1130 points

\[ \sigma = 2.5 \]
438 points

\[ \sigma = 5 \]
146 points

\[ \sigma = 7.5 \]
69 points

\[ \sigma = 10 \]
32 points
Repeatability

Does the detector find the same keypoints despite the change of viewpoint?
Scale Detection using SIFT

SIFT is an integrated keypoint and scale detector and descriptor.
For more, see [Lowe, IJCV’04]
Keypoint Matching

- The matching problem is a search problem
  - Given a point in the first image, find the match in the second image
  - Geometric constraints to limit the search are needed
- We must define
  - A search strategy / matching algorithm
  - A point similarity measure
Matching Algorithm

Let $q_{j_1}$ and $q'_{j_2}$ be the interest points in images $\mathcal{I}$ and $\mathcal{I}'$ respectively, with $j_1 = 1, \ldots, m_1$ and $j_2 = 1, \ldots, m_2$

We define $\mathcal{R}(q_{j_1}, q'_{j_2}, \mathcal{I}, \mathcal{I}')$ as a similarity function between the two points

Winner-Takes-All (WTA) Algorithm
For $j_1 = 1, \ldots, m_1$:

- For $j_2 = 1, \ldots, m_2$: compute $c_{j_1,j_2} = \mathcal{R}(q_{j_1}, q'_{j_2}, \mathcal{I}, \mathcal{I}')$

Repeat:

- Find $(j_1, j_2) = \arg \min_{j_1,j_2} c_{j_1,j_2}$ with $c_{j_1,j_2}$ unmarked
- Mark $c_{j_1,j_2}$
- Add $(j_1, j_2)$ in the correspondence list if it does not violate the unicity constraint
Matching Criteria

SSD (Sum of Square Differences)
\[ \mathcal{R}(q, q', I, I') = \sum_{\theta \in \mathbb{N}^2} w(\theta)(I(q + \theta) - I'(q' + \theta))^2 \]

Comparison of SIFT descriptors
A SIFT descriptor \( s(q, I) \) is a normalized histogram of gradient orientations
\[ \mathcal{R}(q, q', I, I') = \| s(q, I) - s(q', I') \|^2 \]
Example
Sum of Squared Differences (SSD)

- SSD value: 72.55
- Reduced SSD value: 15.13
Motion Vectors From SSD+WTA
Result

For an $11 \times 11$ window and SSD
Point matches by comparing SIFT descriptors

Transferred image frames with the LLS estimated homography
Robust Methods

- Deal with erroneous data
- Classify each data as inlier or outlier
- Some robust methods:
  - M-estimators
    - handles few percents of outliers
    - require a threshold
  - LMedS (Least Median of Squares)
    - handles 50% outliers (requires to know the proportion of erroneous data)
    - does not require a threshold
  - RANSAC (RANdom SAmple Consensus)
    - handles more than 50% outliers
    - requires a threshold
Line Fitting Example

- Least Squares fail to find the ‘right’ line
The ‘right’ line is close to all valid points

Two problems:
  - Line fitting
  - Data classification
RANSAC: Basic Idea

- Select two points randomly
- Measure the support for the line they define
- The support is the number of points lying within a distance threshold
- Repeat this a number of times and select the line with most support
- The points within the distance threshold are the inliers
RANSAC: Basic Idea

Line \( \langle A, B \rangle \) has a support of 7
Line \( \langle C, D \rangle \) has a support of 2
RANSAC: Algorithm

★ Repeat $N$ times

1. Randomly select $s$ data points and instantiate the model from this subset

2. Determine the set of inliers as the data points lying within a distance threshold $t$ of the model

★ Select the largest set of inliers and re-estimate the model
Choosing the Two Parameters

- **The distance threshold** $t$
  - Empirically, *e.g.* 1 pixel
  - Assuming that the measurement error has some distribution, *e.g.* Gaussian centred with known standard deviation

- **The number of trials** $N$
  - Infeasible to try every possible sample
  - $N$ must be large enough to ensure a high probability of success
The Number of Trials

★ $\epsilon$ is the proportion of erroneous data (wrong correspondences)

★ $(1 - \epsilon)$ is the probability of picking a non erroneous data

★ $(1 - \epsilon)^s$ is the probability of picking $s$ non erroneous data

★ $(1 - (1 - \epsilon)^s)^N$ is the probability of failure for all trials

★ Let $p$ be the desired probability of success, we have:

$$1 - p = (1 - (1 - \epsilon)^s)^N,$$

giving:

$$N = K(\epsilon, p) = \frac{\log(1 - p)}{\log(1 - (1 - \epsilon)^s)}$$
Adaptive Number of Samples

Idea: estimate the proportion of erroneous data from the samples

★ Set $N \leftarrow \infty$ and sample_count $\leftarrow 0$

★ While $N >$ sample_count repeat

1. Randomly select $s$ data points and instantiate the model from this subset

2. Determine the set of inliers as the data points lying within a distance threshold $t$ of the model

3. Compute the proportion $\epsilon$ of outliers

4. Set $N \leftarrow \min(N, K(\epsilon, p))$

5. Increment sample_count by 1

★ Select the largest set of inliers and re-estimate the model
Using RANSAC for Estimating Two-View Relationships

- RANSAC can be used as is
- The samples must:
  - Be non degenerate (e.g. 3 colinear points for an homography)
  - Include well spread points over the images
- The distance for each correspondence is the reprojection error and might require nonlinear least squares (c.f. lecture 1.8 on triangulation)
- RANSAC is often followed by a guided matching step (i.e. the estimated relationship is used to constrain the matching)
RANSAC for an homography: inliers

RANSAC for an homography: outliers
Transferred image frames with the RANSAC+NLS estimated homography

The panorama
Some more panoramas...
Closure

- Interest points are detected by looking where the intensity varies significantly.
- Interest points are matched by comparing descriptors, such as the whole surrounding patches (SSD).

References

- Interest point detection
  - [ Förstner, 86], [Harris and Stephen, 88]

- Interest point matching
Closure

- RANSAC has proven very successful
- Idea: draw minimal samples of data at random and maximize the support
- References:
  - [Fischler and Bolles, 81]
  - 25 Years of RANSAC, workshop held in conjunction with CVPR’06, includes online tutorial notes