Computer Vision AI – Assignment 2
Epipolar Geometry and Chaining

Thursday 16$^{\text{th}}$ April, 2015

The results, the analysis and the source code must be included in the final delivery. Students are supposed to work on this assignment for two weeks.

1 Fundamental Matrix

Without any extra assumption on the world scene geometry, you cannot affirm that there is a projective transformation (homography) between two views. In this assignment you will write a function that takes two images as input and computes fundamental matrix by Normalized Eight-point Algorithm with RANSAC. You will work with supplied teddy bear and house images. The overall scheme is very similar to what we have seen in the previous assignment for homography estimation:

1. Detect interest points in each image.
2. Characterize the local appearance of the regions around interest points.
3. Get a set of supposed matches between region descriptors in each image.
4. Estimate the fundamental matrix for the given two images.

The first three steps can also be performed using VLFeat functions. **Note:** Eliminating detected interest points on background would help.

In the next stage, we will introduce a method for estimating fundamental matrix [1]. For $n \geq 8$ known corresponding points' pairs in two stereo images, we can formulate a homogenous linear equation as follows:

$$[x_i' y_i' 1] \underbrace{\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}}_{F} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = 0 \ , \tag{1}$$

where x_i and y_i denote x and y coordinates of the i^{th} point p_i, respectively. Equation 1 can also be written as

$$\underbrace{\begin{bmatrix} x_1{x_1}' & x_1{y_1}' & x_1 & y_1{x_1}' & y_1{y_1}' & y_1 & {x_1}' & {y_1}' & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n{x_n}' & x_n{y_n}' & x_n & y_n{x_n}' & y_n{y_n}' & y_n & {x_n}' & {y_n}' & 1 \end{bmatrix}}_{A} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{bmatrix} = 0 \ , \qquad (2)$$

where F denotes the fundamental matrix.

1.1 Eight-point Algorithm

- Construct the $n \times 9$ matrix A
- Find the SVD of A: $A = UDV^T$
- The entries of F are the components of the column of V corresponding to the smallest singular value.

An important property of fundamental matrix is that it is singular, in fact of rank two. The estimated fundamental matrix F will not in general have rank two. The singularity of fundamental matrix can be enforced by correcting the entries of estimated F:

- Find the SVD of F: $F = U_f D_f {V_f}^T$
- Set the smallest singular value in the diagonal matrix D_f to zero in order to obtain the corrected matrix ${D'}_f$
- Recompute F: $F = U_f D'_f {V_f}^T$

1.2 Normalized Eight-point Algorithm

It turns out that a careful normalization of the input data (the point correspondences) leads to an enormous improvement in the conditioning of the problem, and hence in the stability of the result [1]. The added complexity necessary for this transformation is insignificant.

1.2.1 Normalization:

We want to apply a similarity transformation to the set of points $\{p_i\}$ so that their mean is 0 and the average distance to the mean is $\sqrt{2}$.

Let $m_x = \frac{1}{n}\sum_{i=1}^{n} x_i$, $m_y = \frac{1}{n}\sum_{i=1}^{n} y_i$, $d = \frac{1}{n}\sum_{i=1}^{n} \sqrt{(x_i - m_x)^2) + (y_i - m_y)^2)}$, and $T = \begin{bmatrix} \sqrt{2}/d & 0 & -m_x\sqrt{2}/d \\ 0 & \sqrt{2}/d & -m_y\sqrt{2}/d \\ 0 & 0 & 1 \end{bmatrix}$, where x_i and y_i denote x and y coordinates of a point p_i, respectively.

Then $\widehat{p_i} = Tp_i$. Check and show that the set of points $\{\widehat{p_i}\}$ satisfies our criteria. Similarly, define a transformation T' using the set $\left\{\widehat{p_i}'\right\}$, and let $\widehat{p_i}' = T'p_i'$.

1.2.2 Find a fundamental matrix:

- Construct a matrix A from the matches $\widehat{p_i} \leftrightarrow \widehat{p_i}'$ and get $\widehat{F}$ from the last column of V in the SVD of A.
- Find the SVD of $\widehat{F}$: $\widehat{F} = U_{\widehat{f}} D_{\widehat{f}} V_{\widehat{f}}^T$
- Set the smallest singular value in the diagonal matrix $D_{\widehat{f}}$ to zero in order to obtain the corrected matrix $D'_{\widehat{f}}$
- Recompute $\widehat{F}$: $\widehat{F} = U_{\widehat{f}} D'_{\widehat{f}} V_{\widehat{f}}^T$

1.2.3 Denormalization:

Finally, let $F = T'^T \widehat{F}' T$.

1.3 Normalized Eight-point Algorithm with RANSAC

Fundamental matrix estimation step given in Section 1.2.2 can also be performed via a RANSAC-based approach. First pick 8 point correspondences randomly from the set $\left\{\widehat{p_i} \leftrightarrow \widehat{p_i}'\right\}$, then, calculate a fundamental matrix $\widehat{F}'$, and count the number of inliers (the other correspondences that agree with this fundamental matrix). Repeat this process many times, pick the largest set of inliers obtained, and apply fundamental matrix estimation step given in Section 1.2.2 to the set of all inliers.

In order to determine whether a match $p_i \leftrightarrow p_i'$ agrees with a fundamental matrix F, we typically use the Sampson distance as follows:

$$d_i = \frac{(p_i'^T F p_i)^2}{(Fp_i)_1^2 + (Fp_i)_2^2 + (F^T p_i')_1^2 + (F^T p_i')_2^2}, \tag{3}$$

where $(Fp)_j^2$ is the square of the j^{th} entry of the vector Fp. If d_i is smaller than some threshold, the match is said to be an inlier.

To check the fundamental matrix estimation we can also plot their corresponding epipolar lines. The epipolar line can be thought of as the projection of the line on which the point in the other image could have originated from. Draw the epipolar lines based on your estimated fundamental matrix.

2 Chaining

The matching process described so far is performed across pairs of views. These matches can be represented in a single match graph structure. Intuitively, the set of views of the same surface point forms a connected component of the match graph, which can in turn be used to form a sparse point-view matrix whose columns represent surface points, and rows represent the images they appear in. Construct point-view matrix for chaining multiple views with the matches found in last step using all consecutive teddy-bear images (1-2, 2-3, 3-4, ..., 15-16, 16-1). Rows of the point-view matrix will be representing your images while columns will be points.

1. Start from any two consecutive image matches. Add a new column to point-view matrix for each newly introduced point.

2. If a point which is already introduced in the point-view matrix and another image contains that point, mark this matching on your point-view matrix using the previously defined point column. Do not introduce a new column.

References

[1] Hartley, R.: In defense of the eight-point algorithm. TPAMI (1997)