Stereo Geometry

COSC 470: Special Topic
Computer Vision | 3D Reconstruction
Steven Mills

Stereo and Mosaicing

- We saw last time that in specific cases images are related by Homographies
- This is not the case in general
  - Two (perspective) images of the same scene are in general related by a *Fundamental Matrix*
  - For calibrated cameras this can be reduced to the *Essential Matrix*
  - These can be computed from feature matches

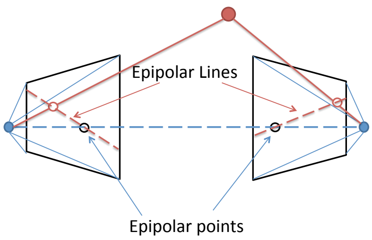
The Fundamental Matrix

- Matching points, *p* and *p'*, in two views are related by the Fundamental Matrix
  $$p'^T F p = 0$$
- *F* has 9 elements, so looks like it has 9 degrees of freedom
  - Homogenous equation, so just 8 degrees
  - *F* has *rank* 2, so just 7 degrees of freedom

Epipolar Geometry

- *F* defines the *epipolar geometry* for two views
  - The *epipolar points* are the projections of each camera centre into the other view
- Given the projection of a 3D point in one view
  - This corresponds to a ray from the camera centre
  - This projects to an *epipolar line* in the other view
  - This line passes through the image of the 3D point and the epipolar point in the second view

Epipolar Geometry



Epipolar Geometry and *F*

- The Fundamental matrix is closely related to the epipolar geometry, suppose $p'^T Fp = 0$
  - $l' = Fp$ is the epipolar line in the second view corresponding to point *p* in the first view
  - $l = F^T p'$ is the epipolar line in the first view corresponding to point *p'* in the second view
  - Here a line, $l = [a, b, c]^T$ is interpreted as the line with formula $ax + by + c = 0$

The Essential Matrix

- Recall our camera projection model
  $$P = K[R|t]$$
- *K* is the camera calibration matrix. If we know the calibration matrices we can form the *Essential matrix*
  $$E = K'^T F K$$
- This depends only on *R* and *t*

The Essential Matrix

- The Essential matrix has 9 elements but only 5 degrees of freedom
  - There are 3 degrees of freedom in *R*
  - There are 3 degrees of freedom in *t*
  - Less one for scale – we don't know the distance between the two cameras

Computing *F* and *E*

- There are several important algorithms based on different numbers of correspondences
  - The 8-point algorithm is a linear method for estimating *F*, but is over-constrained
  - The 7-point algorithm for *F* is minimal but can give 3 possible solutions from solving a cubic equation. These can be checked for consistency
  - The 5-point algorithm for *E* requires solving an order 10 polynomial and can give up to 10 solutions which need to be checked

The 8-point algorithm

- In practice the 8-point algorithm is often used
  - Similar in essence to estimating a homography, but we need 8 points to get 8 equations
  - Similar issues of normalisation and outliers
- Each point correspondence gives us one homogeneous equation in the 9 elements of *F*
- 8 of these let us solve for *F* (up to a scale) using the SVD as we did for finding *H*

Point Correspondence Equations

- Expanding out $p'^T F p = 0$ gives

$$[x' \quad y' \quad 1] \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$

$$\begin{array}{c} x'xf_{11} + y'xf_{21} + xf_{31} + \\ x'yf_{12} + y'yf_{22} + yf_{32} + = 0 \\ x'f_{13} + y'f_{23} + f_{33} \end{array}$$

Basic 8-point Algorithm

- Given 8 (or more) correspondences $p' \leftrightarrow p$:
  - Form a linear equation for each correspondence
  - Stack these up to from $Af = 0$
  - Find the Singular Value Decomposition $A = USV^T$
  - *S* should have one zero (or very small) value
  - The column of *V* corresponding to this value is *f*
  - Reshape *f* into the 3×3 matrix *F*

Robust 8-point Algorithm

- We need to normalise the data as before
  - Shift all the points in each view so that their mean value is zero
  - Scale them so that their average distance from the origin is $\sqrt{2}$
- There will be outliers in the data
  - Use more than 8 points and RANSAC to find an inlier set, then use inliers to find least squares fit

RANSAC and the 8-point Algorithm

- Recall that the number of iterations needed is

$$n = \frac{\log(p)}{\log(1 - \phi^s)}$$

- Here p is the probability of getting a bad result, ϕ the inlier rate, and s the sample size
- Sample size here is 8, so really want ϕ to be close to 1 – need reliable correspondences

Refining *F*

- The solution we get from the 8-point algorithm usually isn't quite right
  - F should have rank 2, which means that the each column is a linear combination of the other two
  - This is equivalent to having a zero singular value
  - So we can correct our estimate by taking the SVD, $F = USV^T$, and setting the smallest value in S to 0 to give S', our corrected result is then $F' = US'V^T$

Estimating *E*

- If we know the calibration matrices we can compute E from F
- This means we can use the 8-point algorithm and have a simple linear solution
- It is possible to compute E from just 5 points
  - This is much more complicated, and not as stable
  - However, it means that $s = 5$ in RANSAC, so can be more robust to larger outlier rates

Non-Linear Optimisation

- As with homographies, the solution we get from the linear equations doesn't quite minimise what we want
- We want to minimise the reprojection error for some unknown 3D point
- Next lecture we will see how to find 3D points from our 2D correspondences and E