COMP90086 Assignment 3 Report

1 Key Points and Correspondences

To identify key points and correspondences between two images, SIFT descriptors with FLANN based matchers and ratio tests were used. The ratio value is set at 0.6 to achieve a good standard in correspondence selection.

2 Normalized 8-point Algorithm

A. Scaling and Shifting Pixel Coordinates

Scaling:

As the pixel coordinates in the provided images range from 1 to 1600, the design matrix may contain an even wider range of numbers from 1 (i.e., the p1 entry) to 10^6 (i.e., the p1q1 entry). Therefore, applying scaling reduces the effect of order-of-magnitude differences between columns, providing better conditions for the design matrix and leading to a better estimate of the fundamental matrix F.

Shifting:

By generating positive and negative pixel coordinates with (0,0) in the middle, shifting can also improve the conditioning of the problem and yields better F estimates.

Design choices:

The image coordinates are scaled relative to the height and width of image (in pixel values) and shifted to retain the x or y value within the range[-1:1], with a zero mean and unit variance. Given the input coordinate (x,y), the scaled and shifted coordinates are computed using the formula:

$$[\frac{\mathrm{x}-\mathrm{r}}{\mathrm{r}}, \frac{y-\mathrm{c}}{\mathrm{c}}] = [\frac{\mathrm{x}}{\mathrm{r}} - 1, \frac{y}{c} - 1]$$

where r= 1/2 number of pixel rows , c= 1/2 number of pixel columns in the image.

When implementing the normalization procedure, we used the array of data points multiplied by the transformation matrix $[[\mathrm{r}, 0], [0, \mathrm{c}]]$.

C&D. SVD & Null Space Detection on Design Matrix and Draft F Calculation

Using the NumPy random choice generator, 8 correspondences are randomly chosen to compute the 8x9 design matrix, with each row generated with the coordinates (p,q) from each correspondence. SVD is performed to decompose the design matrix into three matrices: U, S, VT. The null space is retrieved from the last row of VT (1x9). Since multiplication of design matrix and the flattened F equals 0 given by the formula $\mathrm{Df} = 0$, the flattened F is null space of the design matrix. As a result, the draft F is obtained by reshaping the f to a 3x3 array.

E. SVD on Draft F and Resemblance on F

The draft fundamental matrix is then applied with SVD to generate the middle vector S, whose values are the singular values of the original matrix. The smallest singular value in S is set to 0 to create rank deficiency in F with a determinant of 0. By multiplying the three matrices together, the fundamental matrix is reassembled.

F. Inlier Detection Using F

To detect if a matching point on the right side image (x', y', z') is an inlier, the point can be examined to see if it lies along the epipolar line (l) of the other image. The epipolar line is computed using the F and coordinates of the other image's correspondence $l = FX$, where $X = (x, y, z)$. The resulting epipolar line is a vector (a, b, c), representing the line equation $l = ax + by + c$. The matching point lies on the epipolar line when it satisfied the condition $l = ax' + by' + c = 0$.

Design choices:

As the matching point is allowed to be off by 1 or 2 pixels for error tolerance, a universal array is hardcoded to specify all possible shifts on the matching point, resulting in 13 scenarios covering no shift, shifting x or y coordinate only and shifting both coordinates. The scaling and shifting are then accounted for by applying the transformation matrix defined in part A to the shift array. When checking whether or not the matching point is an inlier, the 13 variations are generated to check if any of them meets the condition $ax' + by' + c = 0$ and thus is an inlier. If so, the original point is added as an inlier. An additional error tolerance is implemented to slightly relax the condition. If $ax' + by' + c \leq 0.01$, the point also registers as an inlier to compensate the rounding problems.

G. Implementing a RANSAC Loop

The steps in B-F are wrapped in a RANSC loop to obtain the inliers detected in the same sample when the inlier count exceeds the MIN_MATCH_NUM (the least number of matches required to find the object). This procedure ensures that inliers from the same sample are more likely to form epipolar lines that intersect at one epipole, yielding the best estimated F.

Design choices:

First, the MIN_MATCH_NUM is set to be 4, allowing some error tolerance while ensuring that at least half of the points are inliers in the sample. Second, the NUM_SAMPLES (number of samples required to have probability > 99% of finding 8 inliers) is set to 1177. It is based on the in-class assumption that half the data is outliers. According to Hartley and Zisserman (2009), the formula for computing the number of samples (N) is as follows:

$$N = log(1 - p)/ log(1 - (1 - \text{Ɛ})^s)$$

where p= 0.99,

Ɛ = probability of any selected data point being an outlier = 0.5,

s= number of points in each sample = 8.

Therefore, $N = log(0.01)/ log(1 - (1 - 0.5)\hat{}8) = 1176.6 = 1177\ samples$

H&I. Re-estimating and Denormalizing F

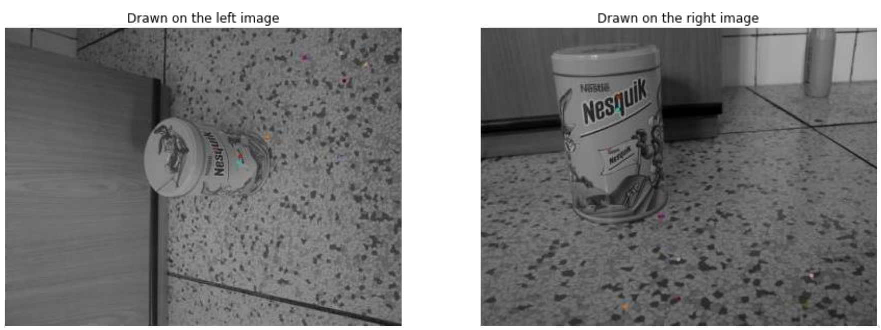
The F is recomputed using the collected inliers, following the steps in C&D. To properly display epipolar lines on the image using F, it is then denormalized to undo the effects of scaling and shifting, using the formula $F = T_R{}' * F * T_L$. T_R and T_L are transformation matrices of the right and left images defined in part A.

Observations

Images with many clear and distinct features with higher contrast and a clear background tend to have correctly identified correspondences, with epipolarlines intersecting towards one point, and thus perform better. For example, tower edges in the *Kyoto* images, blue sky background and house roof edges in the *valbonne* images and the wall background in the *zoom* image.

Zoom images

On the other hand, images with clustered background and many objects with varying colors tend to have illy identified correspondences which do not qualify as distinctive features and non-intersecting epipolar lines, hence performing poorly. For example, the dot patterns in the floor from *box* images can be mistakenly identified as matching points, as with many similar dots in the image, finding the correct correspondences is very unlikely.



Box images

Citation

Hartley, R.I., & Zisserman, A. (2009). Multiple View Geometry. *Encyclopedia of Biometrics*.