CS364/AM792 : Computer Vision 11 September 2017

Assignment 4: 3D Reconstruction

Please prepare a single, condensed and neatly edited document for submission. For each problem include a short description of what you did, results, and a brief discussion/interpretation. It is not necessary to include any code in your document (unless a snippet helps to explain your method). Hand in a printed copy **before 14:00** on the due date, and also e-mail me a zip-file containing your document and code.

You are free to use any programming language. I recommend Matlab or Python. If you are asked to implement a specific technique, the idea is that you do so from scratch; do not simply use a function from an image processing or computer vision library. Collaboration is restricted to the exchange of a few ideas. All code, results, and write-up that you submit must be your own work.

1. You are given an image pair (`et1.jpg` and `et2.jpg`), a number of SIFT feature correspondences between the pair (`etmatches.txt`, in the same format as those given in Assignment 2), and a 3×3 calibration matrix $\mathbf{K}$ applicable to both images (`etK.txt`). The aim is to estimate the spatial relationship between the two views and reconstruct the 3D coordinates of feature points.

(a) Display all the given feature correspondences over a greyscale version of image 1 (plot every match $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ as a clear line segment from $\mathbf{x}_i$ to $\mathbf{x}'_i$, and place a marker at $\mathbf{x}_i$ to distinguish it from $\mathbf{x}'_i$). Then remove "obvious" outliers by enforcing a maximum distance that a feature is allowed to move in image space, and display the remaining matches. Considering the rest of the reconstruction pipeline in parts (b)–(e) below, why is it better to use a lenient threshold here (one that prefers the retainment of incorrect matches over the removal of correct matches)?

(b) Determine a fundamental matrix from the matches remaining after part (a), by means of RANSAC. Display the final set of inlier matches, and remember to re-estimate the fundamental matrix using this entire set.

(c) Use your fundamental matrix and the given calibration matrix to determine an essential matrix $\mathbf{E}$. Calculate the SVD of $\mathbf{E}$, say $\mathbf{E} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, and check that the first two singular values are approximately equal and the third one is zero, as the theory predicts. We will be using $\mathbf{U}$ and $\mathbf{V}$ as rotation (orthogonal, determinant 1) matrices, but the SVD algorithm may return more general orthogonal matrices (determinant ± 1) which may lead to unwanted reflections. Implement the following fix:

```
[U,S,V] := svd(E)
if (det(U) > 0) and (det(V) < 0) then  E := -E, V := -V
elseif (det(U) < 0) and (det(V) > 0) then  E := -E, U := -U
```

Note that this code may multiply the homogeneous matrix $\mathbf{E}$ with a constant, which is perfectly acceptable. We need not worry about the case [$\det(\mathbf{U}) < 0$ and $\det(\mathbf{V}) < 0$] because we will be working with a product of $\mathbf{U}$ and $\mathbf{V}$. If both are reflections they cancel each other out.

(d) Determine two camera matrices from the essential matrix found in part (c) by the method discussed in section 6.5.4, and use triangulation (section 6.3.1) to pick the second camera matrix that produces 3D points in front of both cameras*.

*Further hints were given in class.

(e) Triangulate all the inlier matches, to create a feature-based 3D reconstruction of the objects in the images. You may triangulate by the basic SVD method (section 6.3.1). Put some effort into showcasing your reconstruction: scale the three axes equally, colour the 3D points according to their colours in the images, indicate the positions and orientations of the two cameras in the world coordinate system (like in Assignment 3), remove background points far away from the cameras, and show the 3D plot from three or four different viewpoints.

2. The reconstruction obtained in problem 1 is fairly sparse. Suppose we would like to upgrade it to a more dense representation of the observed scene. We will need to find more matching image points, necessitating **image rectification** (so that matching points have the same vertical coordinates in the two images).

(a) Use the camera matrices obtained in 1(d) to rectify `et1.jpg` and `et2.jpg`, by the procedure outlined in section 6.4. Once you have the necessary transformation matrices you may use your "apply homography" function from Assignment 2 to warp the images.

(b) Draw corresponding epipolar lines over the pair of rectified images, which should now be almost as straightforward as drawing horizontal lines. Just note that the top-left corner of a rectified image may no longer coincide with the origin of the image coordinate system (why not?), and you should take that into consideration when identifying corresponding horizontal lines.

Hint: the values of `miny` calculated in the "apply homography" function will be useful here.

Hand in: 26 September 2017