Epipolar Geometry and Depth Map

Daniel Amirtharaj
damirtha@buffalo.edu

1 Objective

To detect keypoints in 2 given images, match them, find the Fundamental matrix between the 2 images, use this fundamental matrix to draw epipolar lines and also find the depth map of the image pair.

2 Analysis

2.1 Epipolar geometry

In stereo-vision problems, such as depth mapping or 3D reconstruction, features in different images obtained from multiple views must be matched to understand their relative positions in each image. This is essential for computing disparity and thus depth. Depth cannot be estimated from a single 2D image, but the knowledge that a point in an image lies at a certain point in another image obtained from a different view can be used to form a triangulation, this can later be used to precisely estimate the 3D coordinate of the real world point. This forms the basis of epipolar geometry.

Epipolar Lines A feature point on an image can be found in a line in another image formed from a different view. This is due to the fact that a point on an image can be from a real world point at any distance from the point, this ambiguous real world point may fall on any point on a certain line in the other image. This is the epipolar line.

Epipolar plane The plane formed by the 2 camera centers and the real world co-ordinate. The intersection of this plane with the image plane gives the epipolar lines.

Epipoles The intersection of all epilines that can be formed between 2 image planes. It can also be defined as the point of intersection of the line joiining the 2 camera centers and the epilines.

2.2 Fundamental Matrix

The fundamental matrix is a consequence of the properties found in epipolar geometry. It uses the rotation between the 2 image planes and the translation of the 2 camera centers (which together form the essential matrix), and the intrinsic camera parameters of the 2 cameras to define a relationship between corresponding points (which are formed from same real world point) in the 2 images.

$$x_1^T F x_2 = 0 \tag{1}$$

Using only 8 point correspondences (due to 8 unknowns in matrix) in the 2 images, the fundamental matrix can be computed. Knowing the fundamental matrix guarantees the knowledge of the epipolar line, since an image point multiplied by F gives the corresponding epipolar line in the other image.

$$\begin{bmatrix} x^{'} \\ y^{'} \\ 1 \end{bmatrix}^{T} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \tag{2}$$

3 Method

The following steps were followed to find the Fundamental matrix and the depth map from disparity between image1 and image2.

1. SIFT was used to compute keypoints and keypoint descriptors in image1 and image2. Keypoints in both images were captured.
2. Using the keypoint descriptors, good matches with distance lesser than 0.75 were filtered. Matching was done using brute force k-NN matcher with k=2 and L2 norm distances.
3. Good matches obtained were used to find the Fundamental matrix, using RANSAC with the find findFundamentalMat() function. 10 inlier pairs were randomly chosen out of this and plotted.
4. Using these inlier points, epilines on the other image for each image in the image pair were drawn and saved. This was done with the help of the computeCorrespondEpilines() function in OpenCV which accepts the feature points, the index of the image to which the feature points belog to, the fundamental matrix.
5. The depth map was calculated using the stereoBM() function. This is a block matching algorithm which matches corresponding windows/blocks in the image pairs and computes the disparity. The number of disparities and block size were both tuned to get an accepatable depth map.

4 Results

The Fundamental matrix for the given images was calculated using RANSAC with error threshold of 1 and rounded pixel co-ordinates,

$$F = \begin{bmatrix} 0 & -1.435e\text{-}4 & 3.876e\text{-}2 \\ 1.831e\text{-}4 & -1.222e\text{-}5 & -3.121e{+}13 \\ -4.224e\text{-}2 & 3.121e{+}13 & 1 \end{bmatrix}$$

Note : The fundamental matrix was first computed with error threshold of 3 and with floating point co-ordinates, as

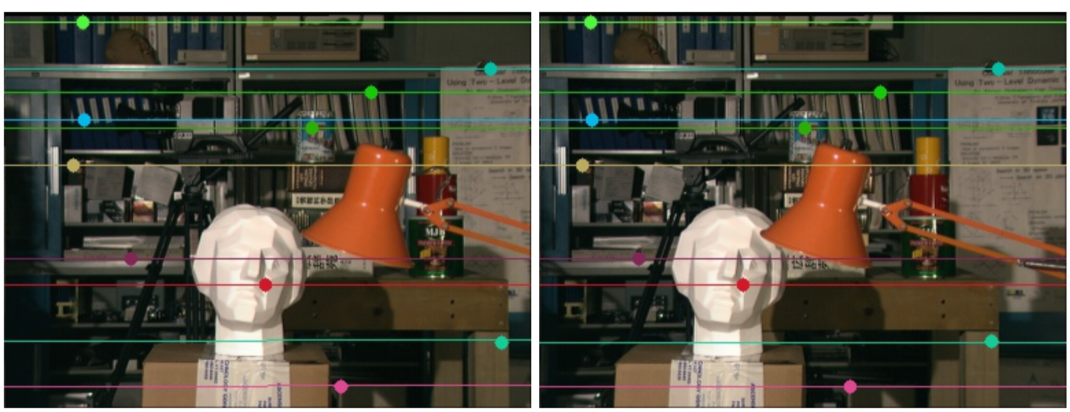
$$F = \begin{bmatrix} 4.341e\text{-}6 & -1.808e\text{-}4 & 3.88e\text{-}2 \\ 2.138e\text{-}4 & -1.825e\text{-}5 & 1.712e\text{-}1 \\ -4.251e\text{-}2 & -1.753e\text{-}1 & 1 \end{bmatrix}$$

This resulted in epilines that were not parallel (since given images are just x-shifted, epilines should have been parallel), since some keypoint matches returned by SIFT had minor changes in y co-ordinates. This is shown in figure 4. To overcome this, pixel values were rounded to nearest integer and RANSAC was run with error threshold of 1. Epilines obtained with this setting are shown in figure 3.

(a) SIFT keypoints detected in the left image. (b) SIFT keypoints detected in the right image.

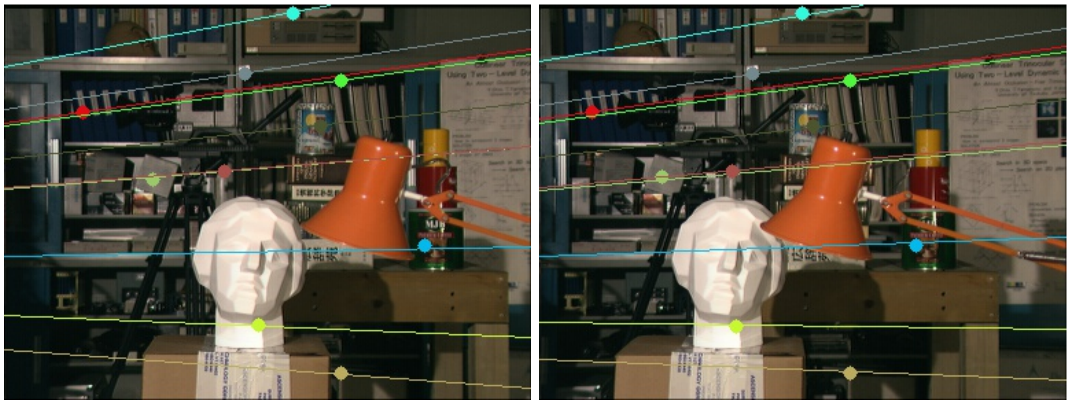
Figure 1: SIFT keypoints detected.

Figure 2: Matches detected with k-NN, with distance lesser than 0.75 between pairs.



(a) Epipolar lines for features detected in the right image, drawn on the left image. (b) Epipolar lines for features detected in the left image, drawn on the right image.

Figure 3: Epipolar lines, corresponding line pairs are drawn with the same color.



(a) Left Image. (b) Right Image.

Figure 4: Epilines detected with F calculated using RANSAC with error threshold 3 and floating pixel values. Not as accurate as epilines obtained in figure 3

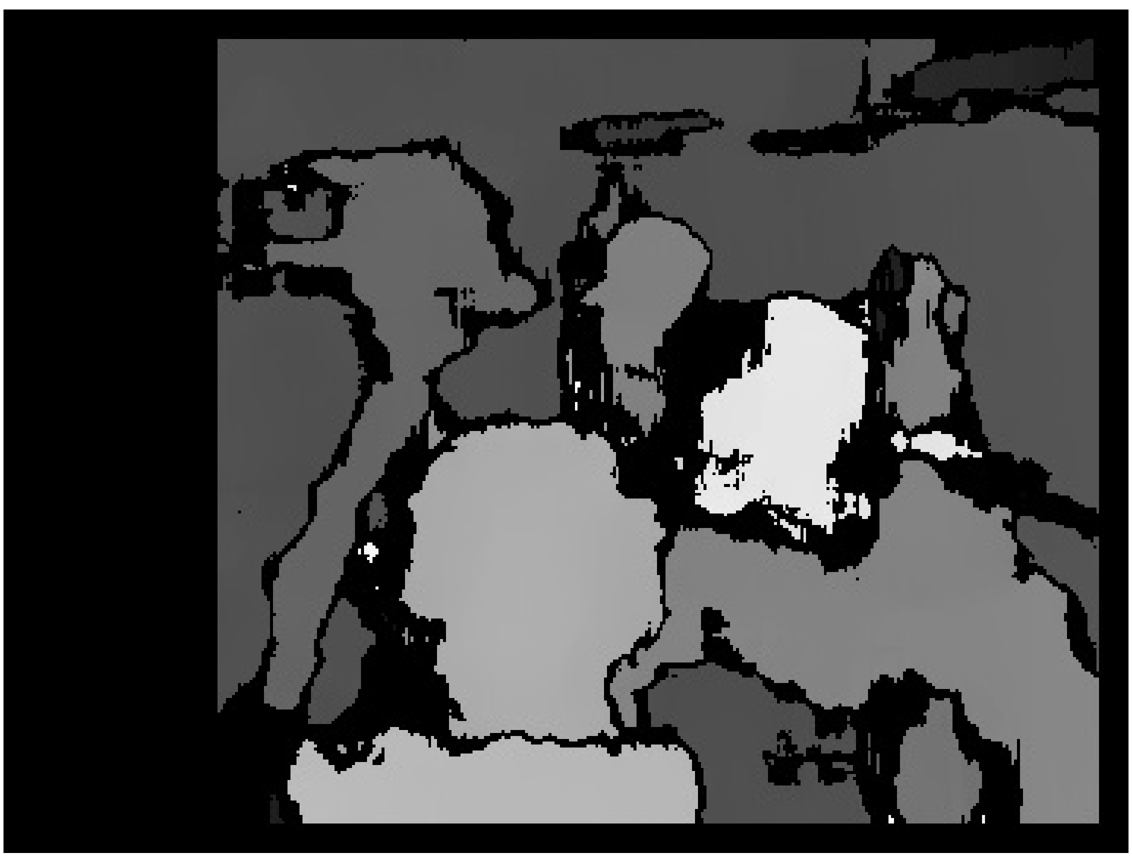


Figure 5: Depth map from disparity computed using block size 64 and number of disparities 21.