ECE 619/645 – Computer Vision
Lab # 3 Geometric Stereo Reconstruction - Part 1: Epipolar Geometry and Stereo
(Issued 2/28 – Due 4/3)

In this part, the geometry of two views is investigated. These views may be acquired simultaneously as a stereo rig, two cameras at different positions, or by a camera moving relative to the scene. The two configurations are geometrically equivalent. Each view of them has its own projection matrix \( M_1 \) and \( M_2 \). A 3-D point \( P \) is projected into the two views as \( p_1 = M_1 X \), and \( p_2 = M_2 P \). Assuming that you have these two images and their projection matrices, how can you recover the depth of \( P \) from the given images? A first step to solve this problem is to find the projections of \( P; p_1 \) and \( p_2 \), in the two images. This matching process is time consuming, since for every point in one image you should search for its match in the entire second image. What is called the epipolar constraint is one of the constraints that can be applied to the images to limit the search to only lines not to the entire 2-D image. The epipolar geometry is then a powerful tool in describing the geometry of two views. Furthermore, we will see how the rectification of images can lead to easy manipulation of this epipolar constraint.

A. Epipolar Geometry

The epipolar geometry is the intrinsic projective geometry between two views. It is independent of the scene structure, and only depends on the camera’s internal parameters and relative pose [1]. The Fundamental matrix \( F \) encapsulates this intrinsic geometry. It is a 3\( \times \)3 matrix of rank 2. If a point \( P \) in 3-D space is projected in one view as \( p_1 \) and in the second view as \( p_2 \) the following constraint should be held:

\[
p_2^T F p_1 = 0
\]  

The epipolar geometry between two views is essentially the geometry of the intersection of the image planes with the pencil of planes having the baseline as axis (the baseline is the line joining the two cameras centers). This geometry is usually motivated by considering the search for corresponding points in stereo matching. As shown in Fig. 1, the 3-D point \( P \) is projected in two images as \( p_L \) and \( p_R \) (the subscripts stand for left and right views, respectively). The geometrical entities involved in the epipolar geometry shown in Fig. 1 are:

- \( e_L \) and \( e_R \) the left and right epipoles, where the epipole is the intersection of the line joining the two camera centers (the baseline) with the image plane. In other words, the epipole is the projection of one camera center into the other image. It is also the vanishing point of the baseline direction.

- An epipolar plane: is a plane containing the baseline. There is a one-parameter family (a pencil) of epipolar planes.

- An epipolar line: is the intersection of the epipolar plane with the image plane.

In addition, the epipole is the intersection of all epipolar lines in the image. A point in the left image has its matched point in the right image lies in an epipolar line, and vice versa. The epipolar line \( l_R \) corresponds to \( p_L \) is defined by two points on it, \( p_R \); the corresponding point of \( p_L \) and the right epipole \( e_R \). The point \( x_R \), defined up to scale, can be expressed in terms of the projection matrices \( M_L \) and \( M_R \) as follows:
\[ p_R = M_R M_L^+ p_L \]  

Then \( I_R \) is expressed as:

\[ l_R = [e_R]_r M_R M_L^+ p_L \]  

Where \( M_L^+ \) is the pseudo inverse of \( M_L \). Since, the point \( p_R \) lies on \( I_R \) then

\[ p_R^T l_R = p_R^T [e_R]_r M_R M_L^+ p_L = 0 \]  

comparing Equation 4 with Equation 1 then,

\[ F_{LR} = [e_R]_r M_R M_L^+ \]  

and

\[ F_{LR} p_L = [e_R]_r M_R M_L^+ p_L = l_R \]  

This means that the fundamental matrix \( F_{LR} \) transfers a point in the left image to a line in the right image on which the corresponding point of the left image lies. This is called the epipolar constraint.

**Properties of the Fundamental Matrix**

For two views the corresponding fundamental matrix has the following properties:
(i) Transpose: if $F_{LR}$ is the fundamental matrix of the pair of cameras $(M_L, M_R)$, then $F_{RL} = F_{LR}^T$ is the fundamental matrix of the pair in the opposite order: $(M_R, M_L)$.

(ii) Epipolar lines: an epipolar line in the right image is expressed as in Equation 6 whereas an epipolar line in the left image is expressed as:

$$l_L = F_{LR}^T p_R$$

(iii) The left and right epipoles are related to the fundamental matrix by the relation:

$$F_{LR} e_L = 0 \text{ and } F_{LR}^T e_R = 0$$

Hence, the epipoles are the null spaces of its corresponding matrices, so they can be computed using the singular value decomposition of $F$.

(iv) In general the fundamental matrix $F$ has 7 degrees of freedom. It is a 3×3 matrix with the ratios of its nine elements are only significant. Also, $F$ has zero determinant hence it removes another degree of freedom.

(v) If $H$ is a 4×4 matrix representing a projective transformation of 3-D space, then the fundamental matrices corresponding to the pairs of projection matrices $(M_1, M_2)$ and $(M_1H, M_2H)$ are the same. Although the pair $(M_1, M_2)$ can uniquely determine a fundamental matrix $F$, the converse is not true. However, the projection matrices can be recovered from $F$ up to 4×4 projectivity. Another explanation to this ambiguity; the two projection matrices have 22 degrees of freedom and $F$ has only 7, then we need another 15 degrees of freedom in addition to $F$’s to uniquely define the projection matrices. These 15 degrees of freedom are given by the 4×4 projectivity.

Computation of the Fundamental Matrix

There are two main methods for solving for the fundamental matrix $F$:

1- If the projection matrices are known, then the fundamental matrix can be computed from Equation 5. or as follows

$$s_L p_L = [A_L \ b_L] \begin{bmatrix} P \\ 1 \end{bmatrix}$$

$$s_R p_R = [A_R \ b_R] \begin{bmatrix} P \\ 1 \end{bmatrix}$$

then

$$s_L p_L = A_L A_R^{-1} (s_R p_R - b_R) + b_L$$

$$s_L p_L = s_R A_L A_R^{-1} p_R + (b_L - A_L A_R^{-1} b_R)$$
Cross multiply both sides by \( b_L - A_L A_R^{-1} b_R \) and then dot multiply both sides by \( p_L \)

\[
0 = p_L^T \left( b_L - A_L A_R^{-1} b_R \right) A_L A_R^{-1} p_R
\]

then

\[
F_{RL} = \left[ b_L - A_L A_R^{-1} b_R \right] A_L A_R^{-1}
\]

2- If 8 pairs of corresponding points are known, and using Equation 1 to form a system of linear equations of 8 unknowns, \( F \) can be recovered. A non-linear method can be used for the solution and only 7 pairs are needed. For more than 8 pairs, the SVD method can be used.

For \( N \) corresponding points, a system of linear equation can be expressed as:

\[
C f = 0
\]

or

\[
\begin{bmatrix}
u_{1L} u_{1R} & u_{1L} v_{1R} & u_{1L} & v_{1L} u_{1R} & v_{1L} v_{1R} & v_{1L} & u_{1R} & v_{1R} \\
u_{2L} u_{2R} & u_{2L} v_{2R} & u_{2L} & v_{2L} u_{2R} & v_{2L} v_{2R} & v_{2L} & u_{2R} & v_{2R} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
u_{1L} u_{1R} & u_{1L} v_{1R} & u_{1L} & v_{1L} u_{1R} & v_{1L} v_{1R} & v_{1L} & u_{1R} & v_{1R} \\
u_{1L} u_{1R} & u_{1L} v_{1R} & u_{1L} & v_{1L} u_{1R} & v_{1L} v_{1R} & v_{1L} & u_{1R} & v_{1R}
\end{bmatrix}
\begin{bmatrix}
f_{11} \\
f_{12} \\
f_{13} \\
f_{31} \\
f_{32} \\
f_{33}
\end{bmatrix} = 0
\]

Using SVD the matrix \( C \) can be decomposed as:

\[
C = U S V^T
\]

The solution is the eigenvector \( V \) corresponds to the smallest eigenvalue in the main diagonal of \( S \). The fundamental matrix is then decomposed, using SVD, as:

\[
F_{RL} = F_1 F_2 F_3^T
\]

To impose the constraint that \( |F_{RL}| = 0 \), the smallest eigenvalue in the main diagonal of \( F_2 \) is set to zero [2]. Then the fundamental matrix is recomputed using the imposed \( F_{2\text{imposed}} \) as:

\[
F_{RL} = F_1 F_{2\text{imposed}} F_3^T
\]

Finding the epipoles

From the fundamental matrix \( F_{RL} \) the right epipole is the eigenvector from \( F_3 \) that corresponds to the zero eigenvalue. The left epipole is the eigenvector from \( F_1 \) that corresponds to the same eigenvalue.

Given a pair of stereo images shown in Fig. 1 (image1.jpg and image2.jpg) below, and assuming that projection matrix is known for each view as follows:
1. From the given projection matrices, compute the fundamental matrix $F_{LR}$.

2. From the $F_{LR}$ compute the epipoles in the first and second images. Show the computed $F_{LR}$ and the epipoles in your report. Compare your $F_{LR}$ with

$$F_{LR} = \begin{bmatrix}
-3.4256742e-005 & -2.3840303e-004 & 9.1519386e-002 \\
1.8825529e-004 & -5.0879929e-005 & 3.1975842e-002 \\
-9.8852919e-002 & -1.8548191e-002 & 1
\end{bmatrix}$$

and the epipoles with

$$e_L = \begin{bmatrix}
-63.6295 \\
393.0282 \\
1.0000
\end{bmatrix} \quad \text{and} \quad e_R = \begin{bmatrix}
-182.7706 \\
4918417 \\
1.0000
\end{bmatrix}$$

3. Repeat the previous two steps however using the 8-point algorithm. Use the files image1.2D and image2.2D that contain a number of matched 2D-points in image1 and image2.

**B. Image Rectification**

Image rectification is an important component of stereo vision algorithms. We assume that a pair of 2-D images of a 3-D object or environment is taken from two distinct views and their epipolar geometry has been determined. Corresponding points between the two images must satisfy the epipolar constraint. For a given point in one image, we have to search for its correspondence in the other.
image along the epipolar line. In general, the epipolar lines are not parallel and don’t align with image axes. Such situation makes the searching process a little bit time consuming, even it is better than searching the entire 2-D images. The rectification process is supposed to parallelize and align the epipolar lines with one of the image axes. In such case, the search space will be reduced to only 1-D.

Image rectification can be view as the process of transforming the epipolar geometry of a pair of images into a standard form. This is accomplished by applying a homography to each image that maps the epipole to the point at infinity $i = [1,0,0]^T$ and the fundamental matrix for the rectified image pair is defined as:

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The rectified images should have the following properties:

- All epipolar lines are parallel to the x-axis,
- The corresponding points have the same y-coordinates

A homography $H$ is needed to parallelize epipolar lines and rotate them to be horizontal. This $H$ is decomposed into:

$$H = H_p H_a$$

Where $H_p$ is a projectivity and $H_a$ is affinity. The projectivity $H_p$ is assumed as:

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_x & w_y & 1 \end{bmatrix}$$

$w_x$ and $w_y$ are responsible for parallelizing the epipolar lines, i.e. transferring the epipoles into infinity. In addition, this should be done with minimum distortion. Also, the affinity $H_a$ can be decomposed into:

$$H_a = H_r H_s$$

Where $H_r$ is a similarity and $H_s$ is a shearing transform. The similarity $H_r$ is responsible for the rotation and the translation of the parallelized epipolar lines to align them with the x-axis. The shearing $H_s$ is another affinity however, it is only to preserve perpendicularity and aspect ratio that may be affected by the projectivity $H_p$ and is defined as:

$$H_s = \begin{bmatrix} a_1 & a_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Similar transformations can also be derived for the other image of the stereo pair. For more details about the rectification algorithm see [C. Loop and Z. Zhang: 99][3].
You are required to do:

In this part it is required to implement the rectification algorithm in [3]. Hint: see the file Rectification.pdf.


Acknowledgements: As with the Camera Calibration Lab, this Laboratory evolved over various iterations of offering of ECE 619(ECE645). Special thanks are due to Dr. Moumen Ahmed and Dr. Ahmed Eid who helped in putting this document together during 2002-2005. Their PhD work is readily available on the CVIP Lab website [www.cvip.uofl.edu](http://www.cvip.uofl.edu). The Laboratory was also implemented on the CardEye vision system by Asem Ali and was maintained throughout 2005-2008.