16720-B HW5

Jacob Miller

November 2020

1.1 Let $\mathbf{x}$ be the homogeneous coordinate in image 1 and $\mathbf{x}$' be the homogeneous coordinate in image 2. If the image coordinates are normalized so that the coordinate origin coincides with the principal point, then $\mathbf{x} = \mathbf{x'} = [0, 0, 1]^T$.

$$\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$$

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

$$f_{33} = 0$$

1.2 Let the translation between the two cameras be $\mathbf{t} = [dx, 0, 0]^T$. Since there is no rotation between the cameras, $R = I_{3x3}$. The essential matrix can be decomposed into:

$$E = [t]_x R$$

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

$$E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -dx \\ 0 & dx & 0 \end{bmatrix}$$

The essential matrix can be used to calculate the epipolar lines for both cameras as follows:

$$
\begin{aligned}
l &= E^T x' \\
l &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & dx \\ 0 & -dx & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \\
l &= \begin{bmatrix} 0 \\ dx \\ -y'dx \end{bmatrix} \\
l' &= Ex \\
l' &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -dx \\ 0 & dx & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\
l' &= \begin{bmatrix} 0 \\ -dx \\ ydx \end{bmatrix}
\end{aligned}
$$

Since l and l' have zero slope in the x-direction, the epipolar lines are parallel in the x direction.

1.3 Given the rotation and translation $\mathbf{R}_i$ and $\mathbf{t}_i$ at times i and i+1, we can find the relative rotation and translation as follows:

$$
\begin{aligned}
\mathbf{R}_{i+1} &= \mathbf{R}_i \mathbf{R}_{rel} \\
\mathbf{R}_{rel} &= \mathbf{R}_i^{-1} \mathbf{R}_{i+1} \\
\mathbf{t}_{i+1} &= \mathbf{t}_i + \mathbf{t}_{rel} \\
\mathbf{t}_{rel} &= \mathbf{t}_{i+1} - \mathbf{t}_i
\end{aligned}
$$

The essential matrix can be expressed as $\mathbf{E} = [\mathbf{t}_{rel}]_x \mathbf{R}_{rel}$. Given the camera intrinsics $\mathbf{K}$, the fundamental matrix can be expressed as $\mathbf{F} = \mathbf{K}^{-1}\mathbf{E}\mathbf{K}^{-1}$.

1.4 Let $\mathbf{x}$ and $\mathbf{x}'$ be the homogeneous coordinates of points on the image plane. The epipolar contraint for reflection is then given by

$$
\begin{aligned}
\mathbf{x}'\mathbf{F}\mathbf{x} &= 0 \\
\mathbf{F} &= \mathbf{K}^{-T}\mathbf{E}\mathbf{K}^{-1}
\end{aligned}
$$

We can write the relationship between the 3D points $\mathbf{X}$ and $\mathbf{X}'$ as follows, where d is the distance from the camera to the reflective plane and $\mathbf{n}$ is the normal:

$$
\begin{aligned}
\mathbf{X}' &= \mathbf{S}\mathbf{X} + 2d\mathbf{n} \\
S &= \mathbf{I} - 2\mathbf{n}\mathbf{n}^T \\
\lambda' \mathbf{K}^{-1}\mathbf{x}' &= \lambda \mathbf{S}\mathbf{K}^{-1}\mathbf{x} + 2d\mathbf{n} \\
\mathbf{x}'^T \mathbf{K}^{-T}(2d[\mathbf{n}]_x \mathbf{S})\mathbf{K}_{-1}\mathbf{x} = 0 & \\
\mathbf{E} &= 2d[\mathbf{n}]_x \mathbf{S}
\end{aligned}
$$

Therefore, the essential and fundamental matrices are skew symmetric.

2.1 The dot product is taken between n and l because the image irradiance depends on the angle between the lighting source and the surface normal. The viewing direction does not matter because the surface normal does not depend on where the surface is viewed from. Only the lighting source direction needs to be known. The projected area plays a role in the albedo, which is a factor in the image irradiance equation.

2.2 The surface normal **n** is related to the partial derivatives of the 3D depth map f_x and f_y because $-\frac{n_1}{n_3}$ can be interpreted as the slope of the depth map in the x direction and $-\frac{n_2}{n_3}$ can be interpreted as the slope of the depth map in the y direction. This is due to the fact that the surface normal is a vector that describes the direction that is perpendicular to that surface.

3.2.1 See Figure 1 for the epipolar line visualization using the funamental matrix. The estimated fundamental matrix was:

$$F = \begin{bmatrix} -8.33149231e-09 & 1.29538462e-07 & -1.17187851e-03 \\ 6.51358337e-08 & 5.70670061e-09 & -4.13435038e-05 \\ 1.13078765e-03 & 1.91823637e-05 & 4.16862081e-03 \end{bmatrix}$$

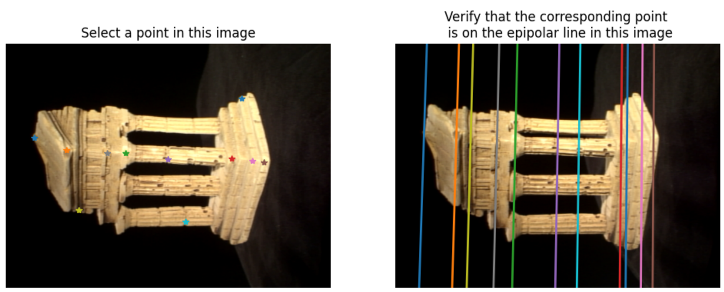

Figure 1: A visualization of epipolar lines using the fundamental matrix.

3.2.2 See Figure 2 for a visualization of epipolar lines using the fundamental matrix obtained from the 7 point algorithm. The estimated fundamental matrix was:

$$F = \begin{bmatrix} 2.28357993e-075.78139007e-063.76540886e-03 \\ -5.95192300e-06-9.30724632e-075.33324961e-04 \\ -3.79932405e-032.34231648e-05-6.14642185e-02 \end{bmatrix}$$

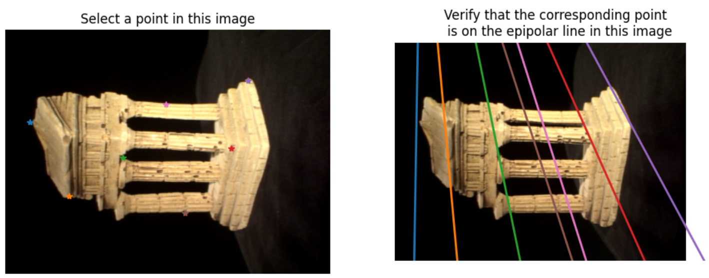

Figure 2: A visualization of epipolar lines using the fundamental matrix obtained from the 7 point algorithm.

3.3.1 The essential matrix estimated from the fundamental matrix was:

$$E = \begin{bmatrix} -1.92592123e-02 & 3.00526429e-01 & -1.73693252e+00 \\ 1.51113725e-01 & 1.32873151e-02 & -3.08885272e-02 \\ 1.73986815e+00 & 9.11774762e-02 & 3.90697726e-04 \end{bmatrix}$$

3.3.2 For solving the triangulation problem of the form $\mathbf{A}_i w_i = 0$, $\mathbf{A}_i$ is defined as follows:

$$A_i = \begin{bmatrix} y\mathbf{p}_3^T - \mathbf{p}_2^T \\ \mathbf{p}_1^T - x\mathbf{p}_3^T \end{bmatrix}$$

When implementing triangulation on the temple images, the reprojection error was 119.457.

3.4.1 See Figure 3 for epipolar matches on the temple images.

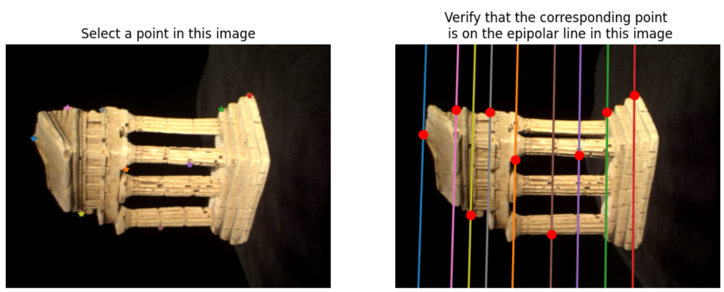

Figure 3: A visualization of epipolar matches using the fundamental matrix on the temple images.

3.4.2 See Figure 4 for a 3D visualization of the temple.

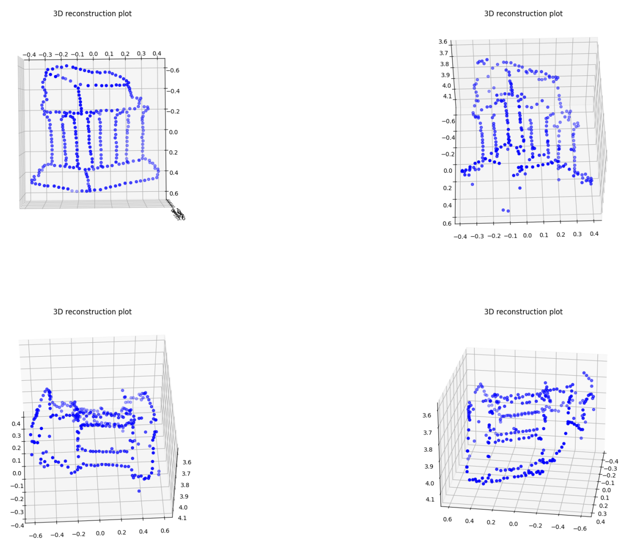

Figure 4: A 3D visualization of the temple using triangulation.

3.5.1 To check if a point is an inlier, I tested if $x'Fx < \epsilon$ for some tolerance ϵ. Figure 5 shows a visualization of the epipolar lines using the fundamental matrix obtained from RANSAC. The results look similar to that of the eight-point algorithm. The epipolar lines accurately run through the corresponding points in the second image.

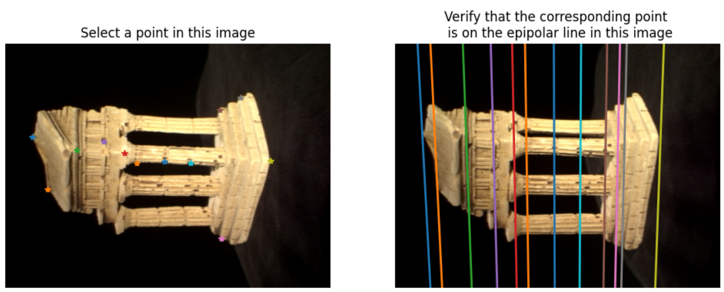

Figure 5: A visualization of epipolar lines using the fundamental matrix obtained from the RANSAC.

4.1 See Figure 6 for the n-dot-l rendering of a sphere with various lighting directions.

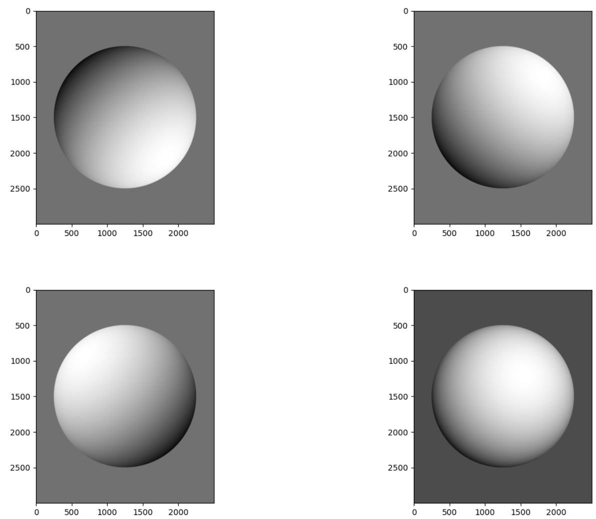

Figure 6: The n-dot-l sphere rendering for lighting directions 1 (top-left), 2 (top-right), 3 (bottom-left), and the combination of the 3 lighting sources.

4.2.3 and 4.2.4 See Figure 7 for the albedos map and a visualization of the image normals.

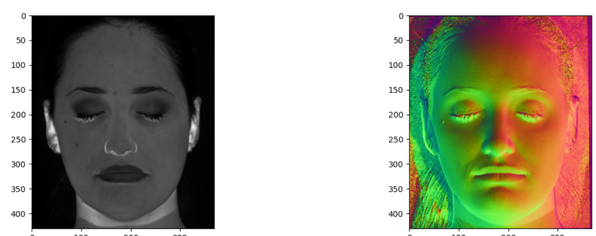

Figure 7: Visualizations of the albedo map and image normals generated from the given images and lighting directions.

4.3.1 See Figure 8 for a 3D visualization of the face using depths estimated from normals. The Frankot-Chellappa algorithm enforces integratability through the following:

1. Compute the 2D discrete Fourier Transforms Z_x and Z_y using the image gradients f_x and f_y.

2. Shift the zero-frequency component of the Fourier Transform to the center of the frequency spectrum.

3. Perform projection

$$z = \frac{-jw_x Z_x - jw_y Z_y}{w_x^2 + w_y^2}$$

   where w_x and w_y are frequency grids from $-\pi$ to π.

4. Shift z such that the mean depth is equal to zero.

5. Compute Inverse Fourier Transform to obtain the real parts of z.

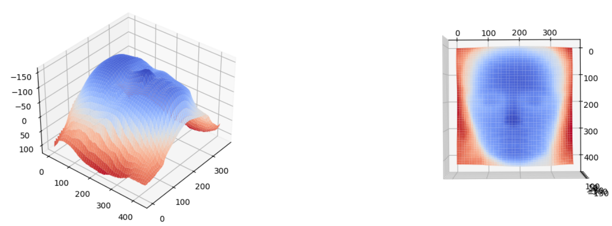

Figure 8: A 3D visualization of the face using depths estimated from normals.

Study Group I formed a study group with Ben Kolligs on 11/15/2020 with a peak attendance of 2.