CSE 252B: Computer Vision II

Lecturer: Serge Belongie
Scribe: Hamed Masnadi Shirazi, Solmaz Alipour

LECTURE 5
Relationships between the Homography and the Essential Matrix

5.1. Introduction

In practice, especially when the scene is piecewise planar, we often need to compute the essential matrix E given a homography H computed from some four points known to be coplanar. The terminology is that H is *induced* by a plane P. In some other cases, the essential matrix E may have been already estimated using points in general position, and we then want to compute the homography for a particular set of coplanar points. There are special relationships between H and E that let us go between them easily.

5.2. Special Properties between H and E

Recalling the definition of essential matrix, $E = \widehat{T}R$, and

$$H = R + \frac{1}{d}\boldsymbol{T}\boldsymbol{N}^\top$$

[1]Department of Computer Science and Engineering, University of California, San Diego.

April 12, 2004

Without loss of generality we can write:

$$H = R + \boldsymbol{T}\boldsymbol{u}^\top$$

where $R \in \mathbb{R}^{3\times 3}$ is a general nonsingular 3×3 matrix – not necessarily a rotation matrix – so the following results apply to both the calibrated and uncalibrated cases. Also assume $\boldsymbol{T}, \boldsymbol{u} \in \mathbb{R}^3$ and $\|\boldsymbol{T}\| = 1$. For the assumed H and E, we have:

(a) $E = \widehat{T}H$
(b) $H^\top E + E^\top H = 0$
(c) $H = \widehat{T}^\top E + \boldsymbol{T}\boldsymbol{v}^\top \quad$ for some $\boldsymbol{v} \in \mathbb{R}^3$.

The proofs for each item are as follows.

(a)

$$\begin{aligned}
\widehat{T}H &= \widehat{T}(R + \boldsymbol{T}\boldsymbol{u}^\top) \\
&= \widehat{T}R + \widehat{T}\boldsymbol{T}\boldsymbol{u}^\top \quad (\text{Note: } \widehat{T}\boldsymbol{T} = \boldsymbol{0}) \\
&= \widehat{T}R \\
&= E
\end{aligned}$$

(b)

$$\begin{aligned}
H^\top E &= (R + \boldsymbol{T}\boldsymbol{u}^\top)^\top \widehat{T}R \\
&= R^\top \widehat{T}R + \boldsymbol{u}\boldsymbol{T}^\top \widehat{T}R \quad (\text{again, } \widehat{T}\boldsymbol{T} = \boldsymbol{0}) \\
&= R^\top \widehat{T}R
\end{aligned}$$

also

$$\begin{aligned}
E^\top H &= (H^\top E)^\top \\
&= (R^\top \widehat{T}R)^\top \\
&= R^\top \widehat{T}^\top R \quad (\widehat{T} \text{ skew-symmetric}) \\
&= -R^\top \widehat{T}R
\end{aligned}$$

The above result shows that $H^\top E$ is a skew symmetric matrix.

Alternatively, this can be shown as follows. Start with the epipolar constraint $\boldsymbol{x}_2^\top E\boldsymbol{x}_1 = 0$. Substituting $\boldsymbol{x}_2 = H\boldsymbol{x}_1$, we get $\boldsymbol{x}_1^\top H^\top E\boldsymbol{x}_1 = 0$ for any $\boldsymbol{x}_1$, which means $H^\top E$ is skew symmetric. (See MaSKS Exercise 2.5.)

(c) Notice that

$$\begin{aligned} \widehat{T}H &= \widehat{T}R \\ &= \widehat{T}\widehat{T}^\top\widehat{T}R \quad (\text{when } \|\boldsymbol{T}\| = 1 \text{ then } \widehat{T}\widehat{T}^\top\widehat{T} = \widehat{T}) \\ &= \widehat{T}\widehat{T}^\top E \quad (\widehat{T}R = E\ ) \end{aligned}$$

(Recall that $\widehat{T}$ is rank deficient, so $\widehat{T}\widehat{T}^\top \neq I$.) Therefore $\widehat{T}(H - \widehat{T}^\top E) = 0$. That is, all the columns of $(H - \widehat{T}^\top E)$ are parallel to $\boldsymbol{T}$, and hence we have $(H - \widehat{T}^\top E) = \boldsymbol{T}\boldsymbol{v}^\top$ for some $\boldsymbol{v} \in \mathbb{R}^3$.

$$\boldsymbol{T}\boldsymbol{v}^\top = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$$

$$\Rightarrow H = \widehat{T}^\top E + \boldsymbol{T}\boldsymbol{v}^\top \quad \text{for some } \boldsymbol{v} \in \mathbb{R}^3$$

Since $R \in \mathbb{R}^{3\times 3}$ is not necessarily in $SO(3)$ these results hold for E and F, where $F = K^{-\top}EK^{-1}$ is the (uncalibrated) fundamental matrix.

5.3. From Homography to the Essential Matrix

Given the homography H our task is to find E given two points p^1 and p^2 not on the plane P from which H was induced. We denote the corresponding images of p^j by $\boldsymbol{x}_i^j$ in view $i = 1, 2$. This is illustrated by figure 1.

From item (a) we have

$$E = \widehat{T}H$$

where $\boldsymbol{T} \sim \widehat{\boldsymbol{l}_2^1}\boldsymbol{l}_2^2$ and $\|\boldsymbol{T}\| = 1$.

We know that $\boldsymbol{l}_2^i \sim \widehat{\boldsymbol{x}_2^i}H\boldsymbol{x}_1^i$, $i = 1, 2$. This is expressing the epipolar lines $\boldsymbol{l}_2^1$ and $\boldsymbol{l}_2^2$ by crossing two points on each line, $H\boldsymbol{x}_1^1$ with $\boldsymbol{x}_2^1$ and $H\boldsymbol{x}_1^2$ with $\boldsymbol{x}_2^2$, respectively. We also know that $\widehat{\boldsymbol{l}_2^1}\boldsymbol{l}_2^2$ is the intersection of the two epipolar lines and remember that all epipolar lines intersect at the epipole. All that remains is to realize that the epipole is $\boldsymbol{T}$ up to a scale factor:

$$\boldsymbol{e}_2 = \widehat{\boldsymbol{l}_2^1}\boldsymbol{l}_2^2 \sim \boldsymbol{T}$$

The above method can be called the $4+2$ point algorithm since it uses 4 points on a plane and 2 points off the plane. Let's look at an example. In figure 2 we see that applying the homography matrix to points on the left image and averaging the result with the right image gives us the third image and it is seen that the coplanar points on the paper are correctly transferred but the points not on the paper surface, like the mug, form a "ghosted" image. Points that are farther away from the plane have more disparity along the epipolar lines. This is known as plane-induced parallax.

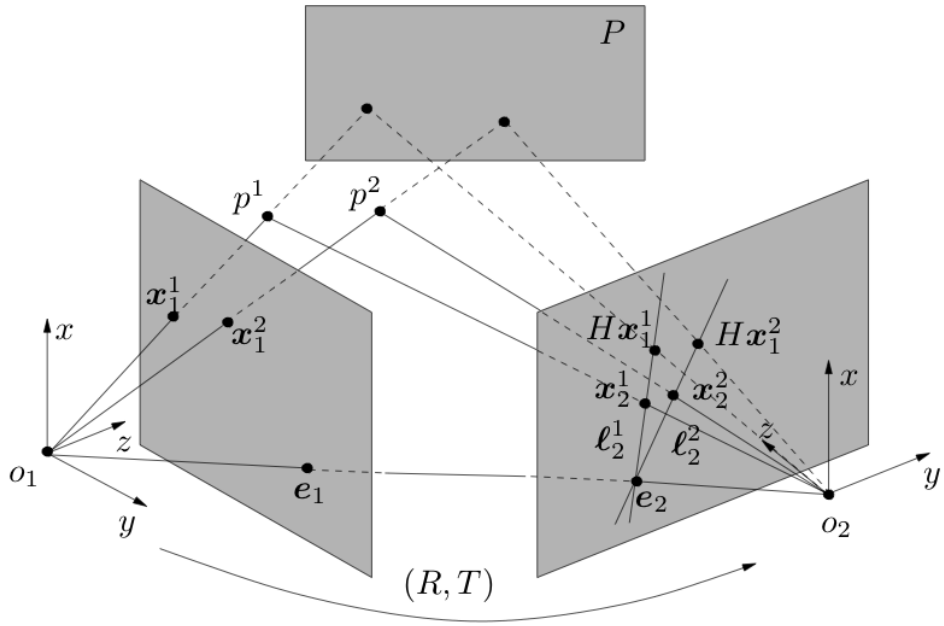

Figure 1. Figure from *Ch. 5 of MaSKS*. The original caption reads: "A Homography H transfers two points $\boldsymbol{x}_1^1$ and $\boldsymbol{x}_1^2$ in the first image to two points $H\boldsymbol{x}_1^1$ and $H\boldsymbol{x}_1^2$ on the same epipolar lines as the respective true images $\boldsymbol{x}_2^1$ and $\boldsymbol{x}_2^2$ if the corresponding 3-D points p^1 and p^2 are not on the plane P from which H is induced. "

5.4. From the Essential Matrix to Homography

Now consider the opposite situation, where the essential matrix E and three points $(\boldsymbol{x}_1^i, \boldsymbol{x}_2^i), i = 1, 2, 3$ are given and we want to compute the homography. The homography induced by the plane specified by the three points is

$$H = \widehat{T}^\top E + \boldsymbol{T}\boldsymbol{v}^\top$$

where $\boldsymbol{T}$ can be found from E as previously shown and $\boldsymbol{v} \in \mathbb{R}^3$ solves the following system of equations

$$\widehat{\boldsymbol{x}_2^i}(\widehat{T}^\top E + \boldsymbol{T}\boldsymbol{v}^\top)\boldsymbol{x}_1^i = \boldsymbol{0} \quad i = 1, 2, 3.$$

Notice that the above equation is similar to $\widehat{\boldsymbol{x}_2^i} H \boldsymbol{x}_1^i = \boldsymbol{0}$, except that H is constrained to have the given form. The method for solving for $\boldsymbol{v}$ is left as an exercise.

While one can find an H for any four points on the image (using the four point algorithm), one can only guarantee H will map a point to an epipolar line if it has the form $\widehat{T}^\top E + \boldsymbol{T}\boldsymbol{v}^\top$. Such an H is "consistent" or "compatible" with E. We will use this method for epipolar (stereo) rectification.

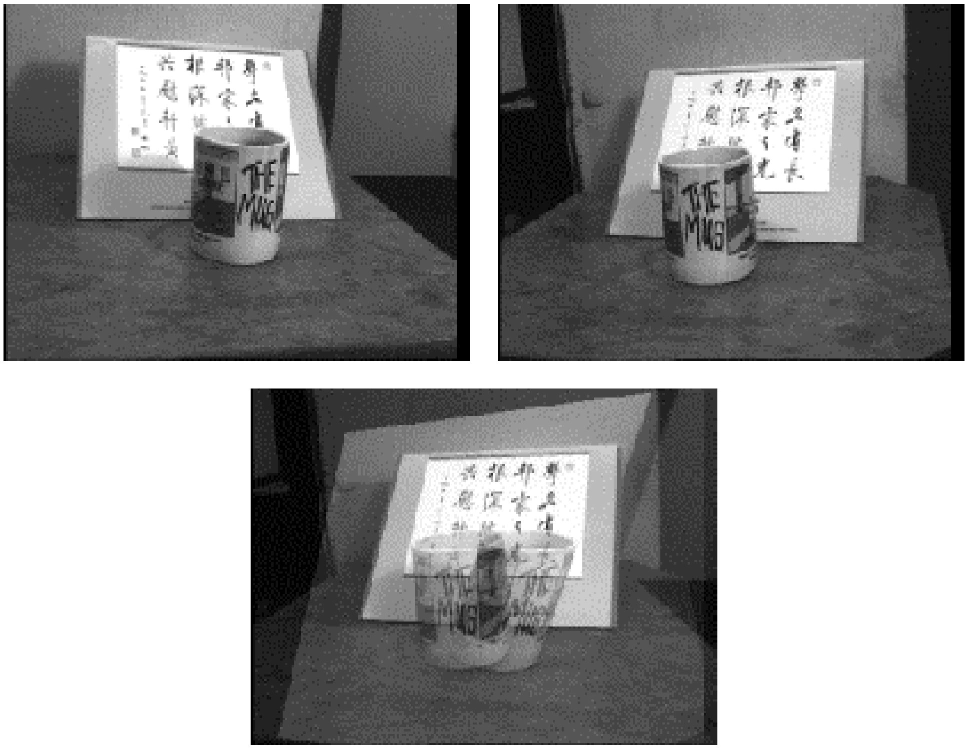

Figure 2. Example of $4+2$ point algorithm. (Figure from Hartley & Zisserman.)

5.5. Epipolar Rectification

In a stereo rectification problem we have two views of a scene. In order to further simplify the search for corresponding points, it is desirable to apply projective transformations to both images so that all epipolar lines correspond to the horizontal scan lines. This entails finding two linear transformations, say H_1 and H_2, that map the epipoles to infinity.

MaSKS Algorithm 11.9 shows us how to do epipolar rectification.

(a) Compute E (or F) and $\boldsymbol{e}_2$. Finding epipoles from E can be done by finding the right and left singular vectors of E.

(b) Map $\boldsymbol{e}_2$ to infinity to make epipolar lines parallel (Hartley, 1997) using H_2. There is a family of H_2's that will do this, parametrized by $\mathbf{v} \in \mathbb{R}^{\mathbf{3}}$.

$$H = \widehat{T}^\top E + \boldsymbol{T}\boldsymbol{v}^\top$$

Which one to choose? Find the H_2 such that

$$H_2 \boldsymbol{e}_2 \sim \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\top$$

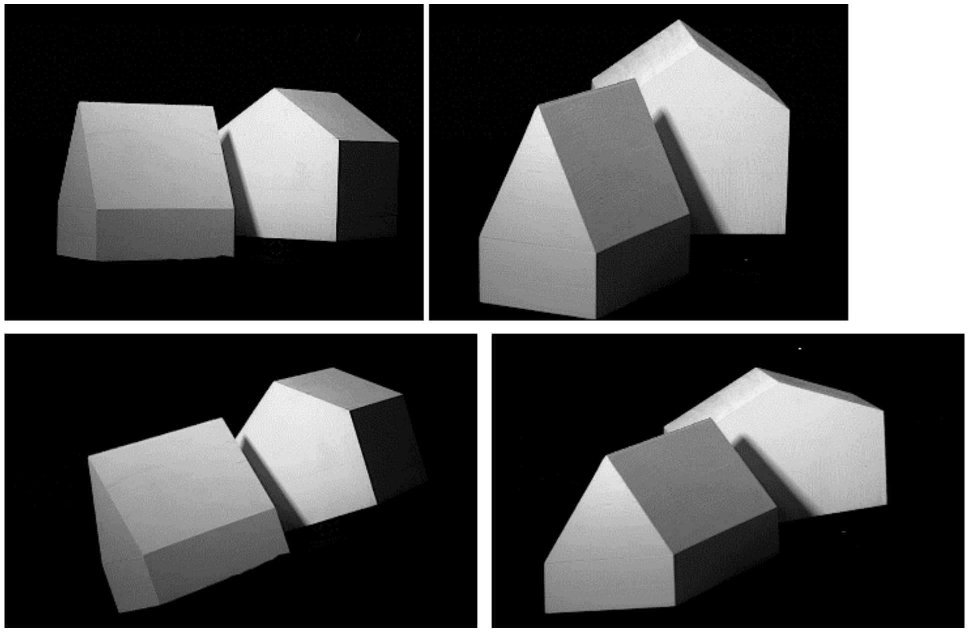

Figure 3. Example of Epipolar Rectification

and where H_2 is as close as possible to a rigid body transformation. So first define $G_T \in \mathbb{R}^{3\times3}$ as

$$G_T = \begin{bmatrix} 1 & 0 & -o_x \\ 0 & 1 & -o_y \\ 0 & 0 & 1 \end{bmatrix}$$

which translates the image center $(o_x, o_y)^\top$ to the origin. Now choose $G_R \in SO(3)$ to do a rotation around the Z-axis so as to put the translated epipole onto the x-axis

$$G_R G_T \boldsymbol{e}_2 = \begin{bmatrix} x_e & 0 & 1 \end{bmatrix}^\top$$

Finally pick $G \in \mathbb{R}^{3\times3}$ defined as

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/x_e & 0 & 1 \end{bmatrix}$$

This sends the epipole to infinity. Choosing $H_2 = G G_R G_T \in \mathbb{R}^{3\times3}$ completes rectification for the second view.

(c) Find an H compatible with E (or F). Here is where we use the method of Section 5.4 for finding H from E. Use the the least squares version of $H = \widehat{T}^\top E + \boldsymbol{T}\boldsymbol{v}^\top$ for multiple points and choose an H that minimizes the distortion induced by the rectification transformation.

(d) Compute the "matching" homography: $H_1 = H_2 H$.
(e) Apply H_1 and H_2 to the left and right images, respectively.

Figure 3 shows the original and rectified versions of two images. It can be seen that corresponding points in the two images are on horizontal scan lines.

The above algorithm will not work if the camera is moving toward the image, in which case the epipole is inside the image. Pollefeys has developed methods to deal with that case.