CS231A Course Notes 3: Epipolar Geometry

Kenji Hata and Silvio Savarese

Image Rectification

Recall that an interesting case for epipolar geometry occurs when two images are parallel to each other. Let us first compute the Essential matrix E in the case of parallel image planes. We can assume that the two cameras have the same K and that there is no relative rotation between the cameras ($R = I$). In this case, let us assume that there is only a translation along the x axis, giving $T = (T_x, 0, 0)$. This gives

$$E = [T_\times]R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T_x \\ 0 & T_x & 0 \end{bmatrix} \tag{17}$$

Once E is known, we can find the directions of the epipolar lines associated with points in the image planes. Let us compute the direction of the epipolar line ℓ associated with point p':

$$\ell = Ep' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T_x \\ 0 & T_x & 0 \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -T_x \\ T_x v' \end{bmatrix} \tag{18}$$

We can see that the direction of ℓ is horizontal, as is the direction of ℓ', which is computed in a similar manner.

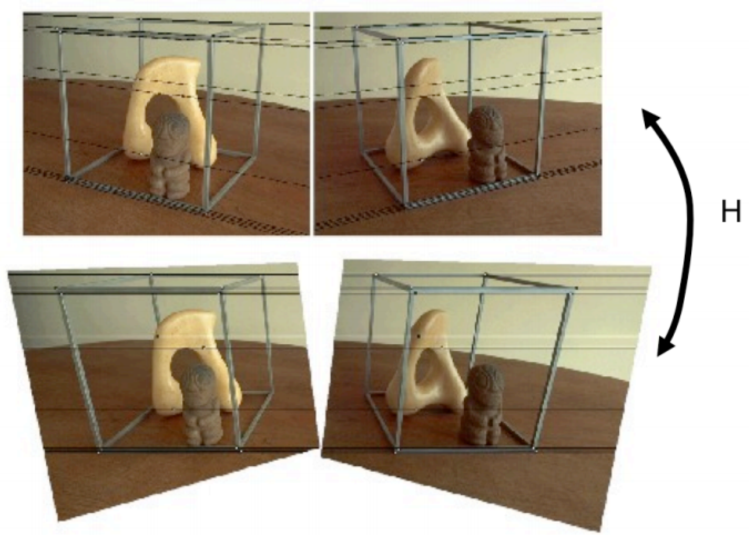

Figure 7: The process of image rectification involves computing two homographies that we can apply to a pair of images to make them parallel.

If we use the epipolar constraint $p^T E p' = 0$, then we arrive at the fact that $v = v'$, demonstrating that p and p' share the same v-coordinate. Consequently, there exists a very straightforward relationship between the corresponding points. Therefore, **rectification**, or the process of making any two given images parallel, becomes useful when discerning the relationships between corresponding points in images.

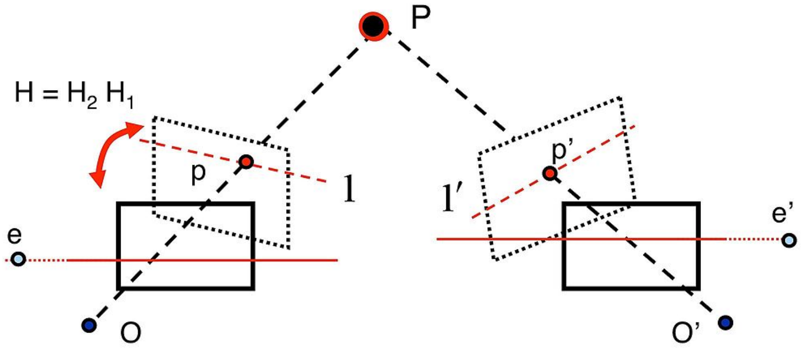

Figure 8: The rectification problem setup: we compute two homographies that we can apply to the image planes to make the resulting planes parallel.

Rectifying a pair of images does not require knowledge of the two camera matrices K, K' or the relative transformation R, T between them. Instead, we can use the Fundamental matrix estimated by the Normalized Eight Point Algorithm. Upon getting the Fundamental matrix, we can compute the epipolar lines ℓ_i and ℓ'_i for each correspondence p_i and p'_i.

From the set of epipolar lines, we can then estimate the epipoles e and e' of each image. This is because we know that the epipole lies in the intersection of all the epipolar lines. In the real world, due to noisy measurements, all the epipolar lines will not intersect in a single point. Therefore, computing the epipole can be found by minimizing the least squared error of fitting a point to all the epipolar lines. Recall that each epipolar line can be represented as a vector ℓ such that all points on the line (represented in homogeneous coordinates) are in the set $\{x|\ell^T x = 0\}$. If we define each epipolar line as $\ell_i = \begin{bmatrix}\ell_{i,1} & \ell_{i,2} & \ell_{i,3}\end{bmatrix}^T$, then we can we formulate a linear system of equations and solve using SVD to find the epipole e:

$$\begin{bmatrix} \ell_1^T \\ \vdots \\ \ell_n^T \end{bmatrix} e = 0 \tag{19}$$

After finding the epipoles e and e', we will most likely notice that they are not points at infinity along the horizontal axis. If they were, then, by definition, the images would already be parallel. Thus, we gain some insight into how to make the images parallel: can we find a homography to map an epipole e to infinity along the horizontal axis? Specifically, this means that we want to find a pair of homographies H_1, H_2 that we can apply to the images to map the epipoles to infinity. Let us start by finding a homography H_2 that maps the second epipole e' to a point on the horizontal axis at infinity $(f, 0, 0)$. Since there are many possible choices for this homography, we should try to choose something reasonable. One condition that leads to good results in practice is to insist that the homography acts like a transformation that applies a translation and rotation on points near the center of the image.

The first step in achieving such a transformation is to translate the second image such that the center is at $(0, 0, 1)$ in homogeneous coordinates. We can do so by applying the translation matrix

$$T = \begin{bmatrix} 1 & 0 & -\frac{\text{width}}{2} \\ 0 & 1 & -\frac{\text{height}}{2} \\ 0 & 0 & 1 \end{bmatrix} \tag{20}$$

After applying the translation, we apply a rotation to place the epipole on the horizontal axis at some point $(f, 0, 1)$. If the translated epipole Te' is located at homogeneous coordinates $(e'_1, e'_2, 1)$, then the rotation applied is

$$R = \begin{bmatrix} \alpha\frac{e'_1}{\sqrt{e'^2_1+e'^2_2}} & \alpha\frac{e'_2}{\sqrt{e'^2_1+e'^2_2}} & 0 \\ -\alpha\frac{e'_2}{\sqrt{e'^2_1+e'^2_2}} & \alpha\frac{e'_1}{\sqrt{e'^2_1+e'^2_2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{21}$$

where $\alpha = 1$ if $e'_1 \geq 0$ and $\alpha = -1$ otherwise. After applying this rotation, notice that given any point at $(f, 0, 1)$, bringing it to a point at infinity on the horizontal axis $(f, 0, 0)$ only requires applying a transformation

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{f} & 0 & 1 \end{bmatrix} \tag{22}$$

After applying this transformation, we finally have an epipole at infinity, so we can translate back to the regular image space. Thus, the homography H_2 that we apply on the second image to rectify it is

$$H_2 = T^{-1}GRT \tag{23}$$

Now that a valid H_2 is found, we need to find a matching homography H_1 for the first image. We do so by finding a transformation H_1 that minimizes the sum of square distances between the corresponding points of the images

$$\arg\min_{H_1} \sum_i \|H_1 p_i - H_2 p'_i\|^2 \tag{24}$$

Although the derivation[2] is outside the scope of this class, we can actually prove that the matching H_1 is of the form:

$$H_1 = H_A H_2 M \tag{25}$$

where $F = [e]_\times M$ and

$$H_A = \begin{bmatrix} a_1 & a_2 & a_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{26}$$

with (a_1, a_2, a_3) composing the elements of a certain vector $\mathbf{a}$ that will be computed later.

First, we need to know what M is. An interesting property of any 3×3 skew-symmetric matrix A is $A = A^3$ up to scale. Because any cross product matrix $[e]_\times$ is skew-symmetric and that we can only know the Fundamental matrix F up to scale, then

$$F = [e]_\times M = [e]_\times [e]_\times [e]_\times M = [e]_\times [e]_\times F \tag{27}$$

By grouping the right terms, we can find that

$$M = [e]_\times F \tag{28}$$

Notice that if the columns of M were added by any scalar multiple of e, then the $F = [e]_\times M$ still holds up to scale. Therefore, the more general case of defining M is

$$M = [e]_\times F + ev^T \tag{29}$$

for some vector v. In practice, defining M by setting $v^T = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ works very well.

To finally solve for H_1, we need to compute the $\mathbf{a}$ values of H_A. Recall that we want to find a H_1, H_2 to minimize the problem posed in Equation 24. Since we already know the value of H_2 and M, then we can substitute $\hat{p}_i = H_2 M p_i$ and $\hat{p}'_i = H_2 p'_i$ and the minimization problem becomes

$$\arg\min_{H_A} \sum_i \|H_A \hat{p}_i - \hat{p}'_i\|^2 \tag{30}$$

In particular, if we let $\hat{p}_i = (\hat{x}_i, \hat{y}_i, 1)$ and $\hat{p}'_i = (\hat{x}'_i, \hat{y}'_i, 1)$, then the minimization problem can be replaced by:

$$\arg\min_{\mathbf{a}} \sum_i (a_1 \hat{x}_i + a_2 \hat{y}_i + a_3 - \hat{x}'_i)^2 + (\hat{y}_i - \hat{y}'_i)^2 \tag{31}$$

[2] If you are interested in the details, please see Chapter 11 of Hartley & Zisserman's textbook *Multiple View Geometry*

Since $\hat{y}_i - \hat{y}'_i$ is a constant value, the minimization problem further reduces to

$$\arg\min_{\mathbf{a}} \sum_i (a_1 \hat{x}_i + a_2 \hat{y}_i + a_3 - \hat{x}'_i)^2 \tag{32}$$

Ultimately, this breaks down into solving a least-squares problem $W\mathbf{a} = b$ for $\mathbf{a}$ where

$$W = \begin{bmatrix} \hat{x}_1 & \hat{y}_1 & 1 \\ & \vdots & \\ \hat{x}_n & \hat{y}_n & 1 \end{bmatrix} \qquad b = \begin{bmatrix} \hat{x}'_1 \\ \vdots \\ \hat{x}'_n \end{bmatrix} \tag{33}$$

After computing $\mathbf{a}$, we can compute H_A and finally H_1. Thus, we generated the homographies H_1, H_2 to rectify any image pair given a few correspondences.