for every $\bar{\mathbf{p}}_r$. But since F is not identically zero, this is possible if and only if

$$F\bar{\mathbf{e}}_l = 0. \tag{7.21}$$

From (7.21) and the fact that F has rank 2, it follows that *the epipole, $\bar{\mathbf{e}}_l$, is the null space of F*; similarly, $\bar{\mathbf{e}}_r$ is the null space of $F^\top$.

We are now in a position to present an algorithm for finding the epipoles. Accurate epipole localization is helpful for refining the location of corresponding epipolar lines, checking the geometric consistency of the entire construction, simplifying the stereo geometry, and recovering 3-D structure in the case of uncalibrated stereo.

Again we present the algorithm in the case of the fundamental matrix. The adaptation to the case of the essential matrix is even simpler than before. The algorithm follows easily from (7.21): To determine the location of the epipoles, it is sufficient to find the null spaces of F and $F^\top$.

☞ These can be determined, for instance, from the singular value decomposition $F = UDV^\top$ and $F^\top = VDU^\top$ as column of V and U respectively corresponding to the null singular value in the diagonal matrix D.

Algorithm EPIPOLES_LOCATION

The input is the fundamental matrix F.

1. Find the SVD of F, that is, $F = UDV^\top$.
2. The epipole $\mathbf{e}_l$ is the column of V corresponding to the null singular value.
3. The epipole $\mathbf{e}_r$ is the column of U corresponding to the null singular value.

The output are the epipoles, $\mathbf{e}_l$ and $\mathbf{e}_r$.

☞ Notice that we can safely assume that there is exactly one singular value equal to 0 because algorithm EIGHT_POINT enforces the singularity constraint explicitly.

It has to be noticed that there are alternative methods to locate the epipoles, not based on the fundamental matrix and requiring as few as 6 point correspondences. More about them in the Further Readings.

7.3.7 Rectification

Before moving on to the problem of 3-D reconstruction, we want to address the issue of *rectification*. Given a pair of stereo images, rectification determines a transformation (or *warping*) of each image such that *pairs of conjugate epipolar lines become collinear and parallel to one of the image axes*, usually the horizontal one. Figure 7.7 shows an example. The importance of rectification is that the correspondence problem, which involves 2-D search in general, *is reduced to a 1-D search on a scanline identified trivially*. In other words, to find the point corresponding to (i_l, j_l) of the left image, we just look along the scanline $j = j_l$ in the right image.

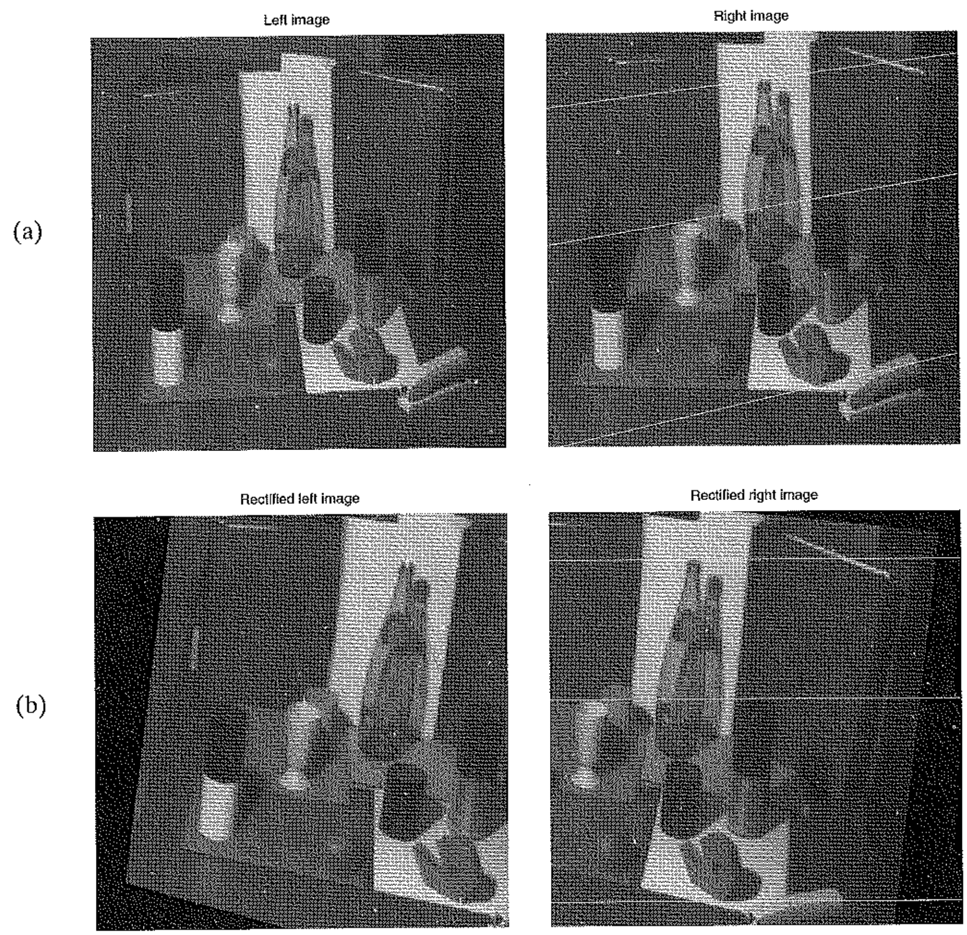

Figure 7.7 (a) A stereo pair. (b) The pair rectified. The left images plot the epipolar lines corresponding to the points marked in the right pictures. Stereo pair courtesy of INRIA (France).

Let us begin by stating the problem and our assumptions.

Assumptions and Problem Statement

Given a stereo pair of images, the intrinsic parameters of each camera, and the extrinsic parameters of the system, R and $\mathbf{T}$, compute the image transformation that makes conjugated epipolar lines collinear and parallel to the horizontal image axis.

The assumption of knowing the intrinsic and extrinsic parameters is not strictly necessary (see Further Readings) but leads to a very simple technique. How do we

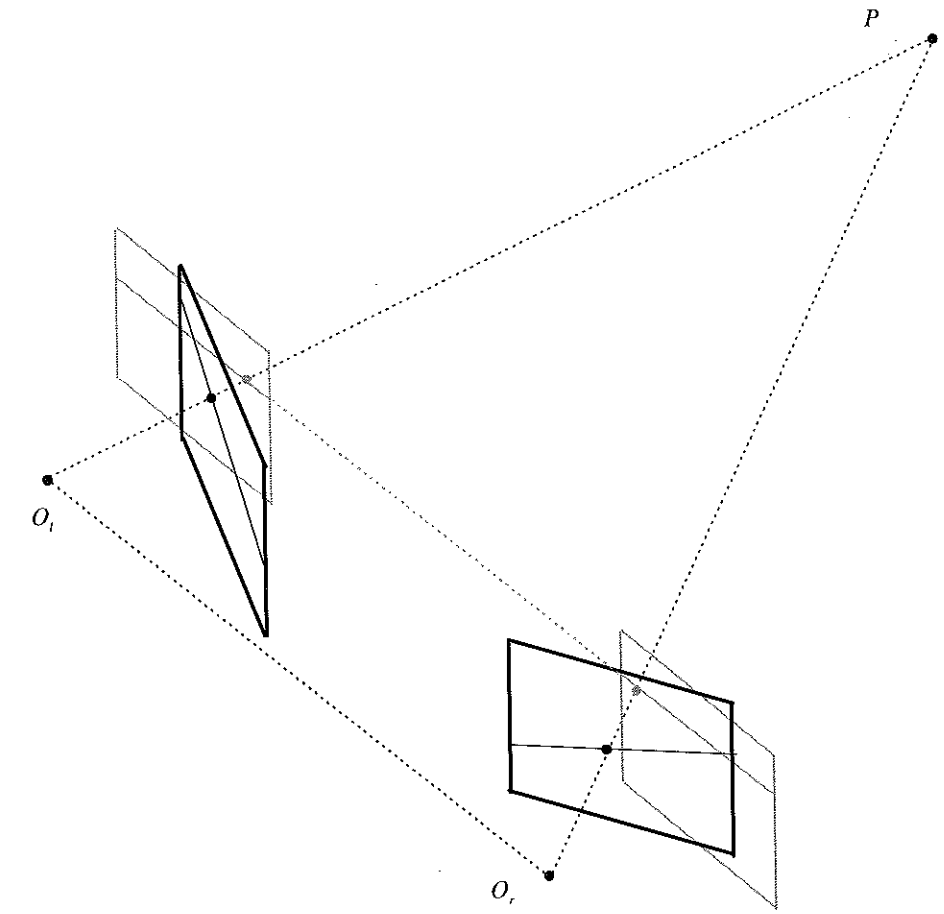

Figure 7.8 Rectification of a stereo pair. The epipolar lines associated to a 3-D point P in the original cameras (black lines) become collinear in the rectified cameras (light grey). Notice that the original cameras can be in any position, and the optical axes may not intersect.

go about computing the rectifying image transformation? The rectified images can be thought of as acquired by a new stereo rig, obtained by rotating the original cameras around their optical centers. This is illustrated in Figure 7.8, which shows also how the points of the rectified images are determined from the points of the original images and their corresponding projection rays.

We proceed to describe a rectification algorithm assuming, without losing generality, that in both cameras

1. the origin of the image reference frame is the principal point;
2. the focal length is equal to f.

The algorithm consists of four steps:

- Rotate the left camera so that the epipole goes to infinity along the horizontal axis.
- Apply the same rotation to the right camera to recover the original geometry.
- Rotate the right camera by R.
- Adjust the scale in both camera reference frames.

To carry out this method, we construct a triple of mutually orthogonal unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$. Since the problem is underconstrained, we are going to make an arbitrary choice. The first vector, $\mathbf{e}_1$, is given by the epipole; since the image center is in the origin, $\mathbf{e}_1$ coincides with the direction of translation, or

$$\mathbf{e}_1 = \frac{\mathbf{T}}{\|\mathbf{T}\|}.$$

The only constraint we have on the second vector, $\mathbf{e}_2$, is that it must be orthogonal to $\mathbf{e}_1$. To this purpose, we compute and normalize the cross product of $\mathbf{e}_1$ with the direction vector of the optical axis, to obtain

$$\mathbf{e}_2 = \frac{1}{\sqrt{T_x^2 + T_y^2}} \left[-T_y, T_x, 0\right]^\top.$$

The third unit vector is unambiguously determined as

$$\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2.$$

It is easy to check that the orthogonal matrix defined as

$$R_{rect} = \begin{pmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_2^\top \\ \mathbf{e}_3^\top \end{pmatrix} \tag{7.22}$$

rotates the left camera about the projection center in such a way that the epipolar lines become parallel to the horizontal axis. This implements the first step of the algorithm. Since the remaining steps are straightforward, we proceed to give the customary algorithm:

Algorithm RECTIFICATION

The input is formed by the intrinsic and extrinsic parameters of a stereo system and a set of points in each camera to be rectified (which could be the whole images). In addition, Assumptions 1 and 2 above hold.

1. Build the matrix R_{rect} as in (7.22);
2. Set $R_l = R_{rect}$ and $R_r = R R_{rect}$;

3. For each left-camera point, $\mathbf{p}_l = [x, y, f]^\top$ compute

$$R_l \mathbf{p}_l = [x', y', z']$$

and the coordinates of the corresponding rectified point, $\mathbf{p}'_l$, as

$$\mathbf{p}'_l = \frac{f}{z'}[x', y', z'].$$

4. Repeat the previous step for the right camera using R_r and $\mathbf{p}_r$.

The output is the pair of transformations to be applied to the two cameras in order to rectify the two input point sets, as well as the rectified sets of points.

Notice that the rectified coordinates are in general not integer. Therefore, if you want to obtain integer coordinates (for instance if you are rectifying the whole images), you should implement RECTIFICATION backwards, that is, starting from the *new* image plane and applying the *inverse* transformations, so that the pixel values in the *new* image plane can be computed as a bilinear interpolation of the pixel values in the *old* image plane.

☞ A rectified image is not in general contained in the same region of the image plane as the original image. You may have to alter the focal lengths of the rectified cameras to keep all the points within images of the same size as the original.

We are now fully equipped to deal with the reconstruction problem of stereo.

7.4 3-D Reconstruction

We have learned methods for solving the correspondence problem and determining the epipolar geometry from at least eight point correspondences. At this point, the 3-D reconstruction that can be obtained depends on the amount of *a priori* knowledge available on the parameters of the stereo system; we can identify three cases.[9] First, if both intrinsic and extrinsic parameters are known, you can solve the reconstruction problem unambiguously by triangulation, as detailed in section 7.1. Second, if only the intrinsic parameters are known, you can still solve the problem and, at the same time, estimate the extrinsic parameters of the system, but only *up to an unknown scaling factor*. Third, if the pixel correspondences are the only information available, and neither the intrinsic nor the extrinsic parameters are known, you can still obtain a reconstruction of the environment, but only *up to an unknown, global projective transformation*. Here is a visual summary.

---

[9] In reality there are several intermediate cases, but we concentrate on these three for simplicity.