5.1.1. Elimination of Structure Unknowns

Since we assume rigidity we know that the motion can be represented by a rotation and a translation

$$P' = RP + T \tag{5.1}$$

where P and P' are the position vectors of a world point before and after the motion, R is the rotation matrix and T is the translation vector. In the relative orientation problem all of them are unknowns. From the projective equations we have

$$p' = \frac{P'}{Z'}$$

and

$$p = \frac{P}{Z}$$

where p and p' are the projections of the world points on the image plane and Z and Z' are the z-coordinates of points P and P' respectively. If we eliminate the world points then Eq. (5.1) becomes

$$Z'p' = ZRp + T\,. \tag{5.2}$$

Now we have one equation where p and p' are knowns since we can measure them on the image and all the rest are unknowns. The unknowns are of two kinds: structure unknowns that are related to the structure of the scene, e.g. the position of the points viewed, and motion unknowns which describe the motion. Z and Z' belong to the first kind of unknowns and there are two of them for each point in the scene. The motion parameters belong to the second kind and are independent of the number of points in the scene.

To solve the problem we have to start eliminating unknowns. The strategy we follow is to eliminate all the per-point unknowns and get an equation independent of structure. Before we do this let us look at the balance of equations first. Eq. (5.2) is a vector equation that is equivalent to 3 scalar equations. If we eliminate the two structure unknowns Z and Z' we have only one equation left.

To eliminate the Z' we can use a property of the cross product that says that the cross product of a vector with itself is the zero vector:

$$a \times a = \vec{0}$$

and multiply both sides of Eq. (5.2) by p'

$$Z'p' \times p' = Z(Rp) \times p' + T \times p' = 0$$

and we still have a vector equation that is equavalent to 3 scalar ones (but they are not independent anymore) and one less unknown. We can eliminate the remaining unknown by using a property of the dot product that says that the dot product of two orthogonal vectors is equal to zero. We know that the cross product of T and p' is another vector that is orthogonal to both T and p'. So if we take the dot product of both sides with T we have

$$0 = Z\Big((Rp) \times p'\Big) \cdot T + \Big(T \times p'\Big) \cdot T = Z\Big((Rp) \times p'\Big) \cdot T.$$

Now assuming Z is not equal to zero (otherwise the object would be way too close to our camera) we get

$$\Big((Rp) \times p'\Big) \cdot T = 0. \tag{5.3}$$

We now have one scalar equation whose unknowns are the motion parameters. We can simplify things a bit more if we notice that Eq. (5.3) is a *triple scalar product*. Recall that one of the definitions of the cross product of two vectors

$$V_1 = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$$

and

$$V_2 = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$$

is equal to the following depterminant

$$\begin{aligned} V_1 \times V_2 = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = \\ (b_1c_2 - b_2c_1)\hat{x} + (c_1a_2 - c_2a_1)\hat{y} + (a_1b_2 - a_2b_1)\hat{z} = \\ \begin{bmatrix} b_1c_2 - b_2c_1 \\ c_1a_2 - c_2a_1 \\ a_1b_2 - a_2b_1 \end{bmatrix} \end{aligned} \tag{5.4}$$

where $\hat{x}$, $\hat{y}$, $\hat{z}$ are the unit vectors along the corresponding axes. If V_3 is

$$V_3 = \begin{bmatrix} a_3 \\ b_3 \\ c_3 \end{bmatrix} = a_3\hat{x} + b_3\hat{y} + c_3\hat{z}$$

then, the triple scalar product of V_1, V_2, V_3 can be written as

$$(V_1, V_2, V_3) = (V_1 \times V_2) \cdot V_3 = \begin{vmatrix} a_3 & b_3 & c_3 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}$$

which means that the well known properties of the determinants can be applied here. Every time we swap two rows the determinant changes sign and in our case if we swap two vectors then the left hand side of (5.3) should stay zero. If we apply this property a couple of times we get

$$p' \cdot (T \times (Rp)) = 0. \tag{5.5}$$

Time now for yet another representation of the cross product. Eq. (5.4) can be rewritten in matrix form as

$$V_1 \times V_2 = \begin{bmatrix} b_1 c_2 - b_2 c_1 \\ c_1 a_2 - c_2 a_1 \\ a_1 b_2 - a_2 b_1 \end{bmatrix} = \begin{bmatrix} 0 & -c_1 & b_1 \\ c_1 & 0 & -a_1 \\ -b_1 & a_1 & 0 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$$

We can also write the dot product in matrix form:

$$V_1 \cdot V_2 = V_1{}^T V_2 = V_2{}^T V_1$$

and we can write (5.5) as

$$p'^T \tilde{T} R p = 0$$

where

$$\tilde{T} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

and

$$T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

and if we do the substitution

$$E = \tilde{T} R \tag{5.6}$$

then

$$p'^T E p = 0 \tag{5.7}$$

which is the celebrated *epipolar constraint* and in various forms has been reinvented many times in the history of science and engineering.

The epipolar constraint Eq. (5.7) is a very convenient equation because it is linear in terms of the elements of E so we can compute it from a set of point correspondences.