[7] On the Reconstruction of a Scene from Two Unregistered Images

Harit P Trivedi

GEC Research Limited
Hirst Research Centre, Wembley, HA9 7PP, UK

ABSTRACT

It is sometimes desirable to compute depth from unregistered pairs of images. I show that it is possible to calculate the two 'epicentres' and the relation governing pairs of epipolar lines, given 8 corresponding points in the two images in any coordinate system. This reduces the matching problem to one dimensional searches along pairs of epipolar lines and can be readily automated using any stereo algorithm. Depth, however, does not seem to be derivable without extra information. I show how to compute depth in two such instances, each involving two 'pieces' of information.

1. INTRODUCTION

One often encounters unregistered pairs of stereo images (e.g. in microscopy) from which three dimensional information is nevertheless desired. This provided the motivation for the work reported here. Longuet-Higgins (1981) has shown that the camera geometry is fixed (assuming perspective projection) by the coordinates of 8 corresponding points in a certain coordinate frame. The latter entails knowledge of the 'natural origins' (defined as the point where the respective optic axis meets the image plane) and the orientations of both the image coordinate systems - in other words, the registration information. He also gave an algorithm to compute depth given this information. When images are unregistered, however, neither the natural origins nor the relative image orientation may be known. To what extent can one then succeed in recovering structure (depth)?

I show that it is possible in the absence of any registration information whatever (i.e., given just the 8 corresponding points in *arbitrary* image coordinate systems) to work out the location of the 'epicentres' - where the interoccular axis intersects the image planes and through which all epipolar lines pass - and the relation governing pairs of epipolar lines (defined in section 4), one in each image. This reduces the rest of the matching problem to one dimensional searches along pairs of epipolar lines - which can be automated using any stereo algorithm. Although it seems that structure cannot be inferred from the image data alone in the absence of any registration information whatever, full registration information is also not necessary. For example, given either (a) the direction of displacement (two direction cosines) of one camera with respect to the optic axis of the other, or, (b) the orientation of the optic axis (two angles) of one camera with respect to that of the other, I show how structure can be recovered.

* Now with BP Research, Sunbury, Middlesex

2. BACKGROUND

I keep to the notation used by Longuet-Higgins (1981). Let a point in the scene have 3D coordinates (X_1, X_2, X_3) and (X'_1, X'_2, X'_3) with respect to the left and the right optic centres. Then its left and right image coordinates (measured from the natural origins) are $(x_1, x_2) = (X_1/X_3, X_2/X_3)$, and $(x'_1, x'_2) = (X'_1/X'_3, X'_2/X'_3)$, in the units of their respective focal lengths. Thus image coordinates $x_3 = 1 = x'_3$, so that $x_i = X_i/X_3$ and $x'_i = X'_i/X'_3$ (i,j = 1,2,3).

Let the right camera position and orientation be obtained by displacing the left camera by a vector **t** and then rotating it so that its new orientation can be obtained from the old by applying the the rotation matrix **R**. Then the two sets of 3D coordinates are related by $X'_j = R_{jk}(X_k - t_k)$, implicit summation convention implied hereinafter. Now from the cartesian components of **t**, construct an antisymmetric matrix

$$\mathbf{S} = \begin{bmatrix} 0 & t_3 & -t_2 \\ -t_3 & 0 & t_1 \\ t_2 & -t_1 & 0 \end{bmatrix} .$$

Longuet-Higgins shows that the matrix $\mathbf{Q} = \mathbf{RS}$ satisfies the relations

$$X'_i Q_{ij} X_j = 0, \qquad (i,j=1,2,3) \qquad (1)$$

and hence

$$x'_i Q_{ij} x_j = 0 \qquad (i,j=1,2,3) \qquad (2)$$

for any point. Notice that (1) and (2) continue to hold under image magnification and length-scale changes to the displacement **t**. For convenience, one chooses $|\mathbf{t}| = 1$. Given eight independent pairs of corresponding points - barring special cases (see Longuet-Higgins (1984)) -, it is straightforward to compute the 8 independent ratios of the elements of **Q** as solutions to an 8 by 8 linear simultaneous system of equations. In the same paper, Longuet-Higgins also shows how to extract **R** and **t** (from **Q**), and hence structure.

3. TRANSFORMATION UNDER ROTATION AND TRANSLATION

Now consider a rotation of the right image (described by the rotation matrix $\mathbf{R}_{z'}(g)$) about its optic axis - the z' axis - by some angle 'g' as introducing registration error in the orientation. By writing (2) as a matrix equation

$$\mathbf{x}'^T \mathbf{Q} \mathbf{x} = 0; \qquad (3)$$

i.e.,

$$(\mathbf{R}_{z'}(g)\mathbf{x}')^T (\mathbf{R}_{z'}(g)\mathbf{Q})\mathbf{x} = 0, \qquad (4)$$

we immediately see that the image pair still satisfies an equation of the form (2) but with

$$\mathbf{Q} \rightarrow \mathbf{Q}' = \mathbf{R}_{z'}(g)\mathbf{Q}. \qquad (5)$$

All that needs to be done to get things right is to absorb the extra rotation in **R**, i.e.,

$$\mathbf{R} \rightarrow [\mathbf{R}_{z'}(g)\mathbf{R}]. \qquad (6)$$

Next we consider the effect of displacing the image origins by (u_1,u_2) and (u'_1,u'_2) in the left and the right images respectively. Then $x_i \rightarrow \xi_i = x_i - u_i$, and $x'_i \rightarrow \xi'_i = x'_i - u'_i$, $(i=1,2,3;\ u_3 = u'_3 = 0)$. Starting with (2) yields, by algebraic manipulation, the relation

$$\xi'_i Q''_{ij} \xi_j = 0, \ or, \ (\xi')^T \mathbf{Q}'' \xi = 0; \qquad (7)$$

where

$$u_3 = 0 = u'_3, \qquad (7a)$$

$$Q''_{ij} = Q_{ij}, \quad (i,j=1,2) \qquad (7b)$$

$$Q''_{13} = Q_{13} + r, \qquad (7c)$$

$$Q''_{23} = Q_{23} + s, \qquad (7d)$$

$$Q''_{31} = Q_{31} + r', \qquad (7e)$$

$$Q''_{32} = Q_{32} + s', \qquad (7f)$$

$$Q''_{33} = Q_{33} + t_0 + t'_0 + v; \qquad (7g)$$

i.e.,

$$\mathbf{Q} \rightarrow \mathbf{Q}'' = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13}+r \\ Q_{21} & Q_{22} & Q_{23}+s \\ Q_{31}+r' & Q_{32}+s' & Q_{33}+t_0+t'_0+v \end{bmatrix}. \qquad (8)$$

Here

$$\begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \qquad (9a)$$

$$\begin{bmatrix} r' & s' \end{bmatrix} = \begin{bmatrix} u'_1 & u'_2 \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \qquad (9b)$$

$$t_0 = Q_{31}\, u_1 + Q_{32}\, u_2, \qquad (9c)$$

$$t'_0 = u'_1 Q_{13} + u'_2 Q_{23}, \qquad (9d)$$

and

$$v = r' u_1 + s' u_2 = u'_1 r + u'_2 s. \qquad (9e)$$

Combined rotations and translations of the image coordinate systems can be readily described by replacing **Q** in (7)-(9) with **Q′** of (5). The image coordinates, therefore, always obey a relation of the form (2), or equivalently, (3), whatever the coordinate system. Using this observation, I show how to work out the locations of the epicentres and the relation governing pairs of epipolar lines.

4. EPICENTRES AND EPIPOLAR LINES

Where the interoccular axis intersects the image planes are the two epicentres. Now imagine a family of planes passing through the interoccular axis. Each such plane intersects each image plane in a straight line (which naturally passes through the respective epicentre), giving rise to pairs of epipolar lines. Let the left and the right epicentres be located at (π_1, π_2) and (π'_1, π'_2). The equation of a straight line of slope m passing through (π_1, π_2) is $(\xi_2 - \pi_2) = m(\xi_1 - \pi_1)$. Similarly, denoting by m' the slope of the corresponding epipolar line, the equation of the latter is $(\xi'_2 - \pi'_2) = m'(\xi'_1 - \pi'_1)$. [The geometric motivation presented here is not essential. One can simply postulate the existence of epicentres and epipolar lines and the arguments go through.] Now any point on a certain epipolar line in one image can match any point on the corresponding epipolar line in the other image. Given that all matched points obey (7), one obtains by inserting for ξ'_2 and ξ_2 from the linear equations above into the matrix representation of (7), that

$$\begin{bmatrix} \xi'_1, & m'(\xi'_1 - \pi'_1) + \pi'_2, & 1 \end{bmatrix} \mathbf{Q}'' \begin{bmatrix} \xi_1 \\ m(\xi_1 - \pi_1) + \pi_2 \\ 1 \end{bmatrix} = 0 \qquad (10)$$

for all values of ξ_1 and ξ'_1. The left hand side is a second order inhomogeneous polynomial in ξ_1 and ξ'_1 and can vanish identically if and only if the coefficient of each term vanishes. This yields four equations. The first of them, arising from the vanishing coefficient of the $(\xi_1\xi'_1)$ term, immediately gives the relation

$$m = -(Q''_{11} + m'Q''_{21})/(Q''_{12} + m'Q''_{22}) \tag{11}$$

governing the slopes of a pair of epipolar lines. Note that it is independent of the normalisation of $\mathbf{Q}''$. The solution to the rest of the matching problem can be mechanised by the use of any stereo algorithm.

The condition that the coefficient of the term in ξ'_1 must vanish yields, after substituting (11) for m, a polynomial in m' which must vanish. Equating the coefficient of each power of m' to zero gives two linear inhomogeneous equations in the two unknowns π_1 and π_2:

$$\begin{bmatrix} Q''_{11} & Q''_{12} \\ Q''_{21} & Q''_{22} \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} = - \begin{bmatrix} Q''_{13} \\ Q''_{23} \end{bmatrix}. \tag{12}$$

Similarly, the condition that the coefficient of the term in ξ_1 vanish yields

$$\begin{bmatrix} \pi'_1 & \pi'_2 \end{bmatrix} \begin{bmatrix} Q''_{11} & Q''_{12} \\ Q''_{21} & Q''_{22} \end{bmatrix} = - \begin{bmatrix} Q''_{31} & Q''_{32} \end{bmatrix}. \tag{13}$$

That the constant term also vanishes can be verified by inserting the coordinates of the two epicentres in (7) and using (12) and (13). In the process, one obtains two interesting equations - one for each epicentre:

$$[\pi'_1, \pi'_2, 1]\mathbf{Q}'' = 0, \tag{14}$$

and

$$\mathbf{Q}''[\pi_1, \pi_2, 1]^T = 0; \tag{15}$$

implying that

$$\det|\mathbf{Q}''| = 0. \tag{16}$$

This serves as a check on the accuracy of the data and the calculations.

Alternatively, observing that the last row and column of $\mathbf{Q}''$ in (8) are linear combinations of the rows and columns of $\mathbf{Q}$, it is readily seen that $\det|\mathbf{Q}''| = 0$ if and only if $\det|\mathbf{Q}| = 0$. That $\det|\mathbf{Q}| = 0$ follows from the fact that $\det|\mathbf{Q}| = \det|\mathbf{R}|.\det|\mathbf{S}|$, and it can be verified that $\det|\mathbf{S}| = 0$.

Starting with (3) and using the equivalents of (14) and (15) in the 'natural' coordinate system, i.e.,

$$[p'_1, p'_2, 1]\mathbf{Q} = 0, \tag{14a}$$

and

$$\mathbf{Q}[p_1, p_2, 1]^T = 0, \tag{15a}$$

where (p_1, p_2) and (p'_1, p'_2) are the epicentres in the natural coordinate system, an alternative form of $\mathbf{Q}''$ can also be given:

$$\begin{aligned} Q''_{ij} &= Q_{ij}, \quad (i,j = 1,2) \\ Q''_{i3} &= Q_{i1}(u_1 - p_1) + Q_{i2}(u_2 - p_2), \quad (i=1,2) \\ Q''_{3i} &= (u'_1 - p'_1)Q_{1i} + (u'_2 - p'_2)Q_{2i}, \quad (i=1,2) \\ Q''_{33} &= (u'_1 - p'_1)Q''_{13} + (u'_2 - p'_2)Q''_{23}, \\ &= Q''_{31}(u_1 - p_1) + Q''_{32}(u_2 - p_2). \end{aligned} \tag{17}$$

5. SCENE RECONSTRUCTION

Longuet-Higgins gives a method of recovering structure from $\mathbf{Q}$. He also points out three equations relating the diagonal and the off-diagonal elements of the matrix $\mathbf{Q}^T\mathbf{Q}$ (his eqn. (17)), the rotation matrix dropping out in the process. Three equations are not sufficient to determine the four unknowns u_1, u_2, u'_1 and u'_2 needed to recover $\mathbf{Q}$ from (8) or (17). Thus given $\mathbf{Q}''$ alone, it does not seem possible to recover $\mathbf{Q}$ (whence structure).

It is possible to recover structure, however, given either (a) the direction of displacement of one camera with respect to the optic axis of the other, or, (b) the orientation of the optic axis of one camera with respect to that of the other. Note that $Q_{ij} = Q''_{ij}$, $(i,j=1,2)$. From the image data, therefore, one can obtain three ratios between these four elements. Now, from $\mathbf{Q}=\mathbf{RS}$,

$$\begin{aligned} Q_{11} &= t_2R_{13} - t_3R_{12}, \quad & Q_{12} &= t_3R_{11} - t_1R_{13}, \\ Q_{21} &= t_2R_{23} - t_3R_{22}, \quad & Q_{22} &= t_3R_{21} - t_1R_{23}. \end{aligned} \tag{18}$$

Given $\mathbf{R}$, and using $\mathbf{R}_i \times \mathbf{R}_j = \mathbf{R}_k$, $(i,j,k$ cyclic permutations of 1,2,3), where $\mathbf{R}_m$ refers to the mth row of $\mathbf{R}$ regarded as a vector, (18) yields

$$\begin{aligned} t_1 &= (R_{11}Q_{22} - R_{21}Q_{12})/R_{32}, \\ t_2 &= (R_{12}Q_{21} - R_{22}Q_{11})/R_{31}, \\ t_3 &= (R_{13}Q_{21} - R_{23}Q_{11})/R_{31}, \\ &= (R_{13}Q_{22} - R_{23}Q_{12})/R_{32}. \end{aligned} \tag{19}$$

The two expressions for t_3 in (19) provide an accuracy check. More importantly, it can happen that the right image (say) was rotated about its original position. This corresponds to an unknown rotation about the z' axis - represented by $\mathbf{R}_{z'}(g)$, g being the angle. The two expressions for t_3 then force a constraint on $tan(g)$. To see this, write the final rotation matrix as

$$\mathbf{R} \rightarrow [\mathbf{R}_{z'}(g)\mathbf{R}],$$

where R is known (for example, (Arfken 1970)). That is,

$$\mathbf{R} \rightarrow \begin{bmatrix} cos(g) & sin(g) & 0 \\ -sin(g) & cos(g) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}. \quad (20)$$

Equating the two expressions for t_3 in (19) and substituting for the new **R** from (20), one obtains

$$tan(g) = -a/b, \quad (21)$$

$$a = (R_{13}Q_{21} - R_{23}Q_{11})/R_{31} - (R_{13}Q_{22} - R_{23}Q_{12})/R_{32},$$

$$b = (R_{23}Q_{21} + R_{13}Q_{11})/R_{31} - (R_{23}Q_{22} + R_{13}Q_{12})/R_{32}.$$

There are two possible solutions for g, given $tan(g)$. If the two images are coarsely aligned (by eye, say) then the small angle solution is the desired solution.

Next consider known displacement (t_1, t_2, t_3). Denoting the ratio Q_{11}/Q_{12} by a_x (computed from data measurement), and setting $R_{11}/R_{12} = a_1$ and $R_{13}/R_{11} = a_2$, it can be readily shown that

$$a_2 = t_3(a_1 + a_x)/(t_2 + a_x t_1) = f_1(a_1) \quad (22)$$

is a linear function of a_1. Similarly, denoting the ratio Q_{22}/Q_{21} by a_y (measured), and setting $R_{21}/R_{22} = b_1$ and $R_{23}/R_{22} = b_2$, it can be verified that

$$b_2 = t_3(b_1 + a_y)/(t_1 + a_y t_2) = f_2(b_1) \quad (23)$$

is a linear function of b_1. Then

$$R_{11}^2 + R_{12}^2 + R_{13}^2 = 1, \quad \text{and} \quad R_{21}^2 + R_{22}^2 + R_{23}^2 = 1$$

imply

$$R_{11}^2 = (1 + a_1^2 + f_1^2(a_1))^{-1} \quad (24)$$

and

$$R_{22}^2 = (1 + b_1^2 + f_2^2(b_1))^{-1}. \quad (25)$$

The rotation matrix **R** is characterised by the four unknowns R_{11}, R_{22}, a_1 and b_1, and has the form

$$\mathbf{R} = \begin{bmatrix} R_{11} & a_1 R_{11} & f_1(a_1)R_{11} \\ b_1 R_{22} & R_{22} & f_2(b_1)R_{22} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}, \quad (26)$$

where

$$R_{31} = R_{11}R_{22}[a_1 f_2(b_1) - f_1(a_1)], \quad (26a)$$

$$R_{32} = R_{11}R_{22}[b_1 f_1(a_1) - f_2(b_1)], \quad (26b)$$

and

$$R_{33} = R_{11}R_{22}(1 - a_1 b_1). \quad (26c)$$

Every relationship following from the equation $\mathbf{R}_i \times \mathbf{R}_j = \mathbf{R}_k$, ($i,j,k$ cyclic permutations of 1,2,3) gives

$$a_1 + b_1 + f_1(a_1)f_2(b_1) = 0. \quad (27)$$

Since $f_1(a_1)$ and $f_2(b_1)$ are linear functions of a_1 and b_1 respectively, (27) takes the form

$$c_1 a_1 b_1 + c_2 a_1 + c_3 b_1 = c_4,$$

or

$$b_1 = (c_4 - c_2 a_1)/(c_3 + c_1 a_1); \quad (28)$$

where

$$c_1 = t_3^2, \quad (28a)$$

$$c_2 = (t_2 + a_x t_1)(t_1 + a_y t_2) + a_y t_3^2, \quad (28b)$$

$$c_3 = (t_2 + a_x t_1)(t_1 + a_y t_2) + a_x t_3^2, \quad (28c)$$

and

$$c_4 = -a_x a_y t_3^2. \quad (28d)$$

There is now the last piece of unused information, the ratio $Q_{22}/Q_{11} = a_{yx}$ (measured). Writing this out explicitly, squaring it [to get rid of the square- roots from (24) and (25)], and using (22)-(28), one obtains a fourth degree polynomial equation in a_1 :

$$\begin{aligned} &a_1^4[a_{yx}^2 h_1 h_2 - h_3 h_4] + \\ &a_1^3[a_{yx}^2(e_1 h_2 + h_1 e_2) - (e_3 h_4 + h_3 e_4)] + \\ &a_1^2[a_{yx}^2(d_1 h_2 + e_1 e_2 + h_1 d_2) - (d_3 h_4 + e_3 e_4 + h_3 d_4)] + \\ &a_1[a_{yx}^2(d_1 e_2 + e_1 d_2) - (d_3 e_4 + e_3 d_4)] + \\ &[a_{yx}^2(d_1 d_2) - d_3 d_4] = 0 ; \end{aligned} \quad (29)$$

where

$$n_1 = a_x n_2, \quad n_2 = t_3/(t_2 + a_x t_1),$$

$$n_3 = a_y n_4, \quad n_4 = t_3/(t_1 + a_y t_2); \quad (29a)$$

$$d_1 = (t_3 - t_1 n_1)^2,$$

$$d_2 = (1 + n_3^2)c_3^2 + 2n_3 n_4 c_3 c_4 + (1 + n_4^2)c_4^2,$$

$$d_3 = (1 + n_1^2),$$

$$d_4 = [(t_3 - t_1 n_4)c_4 - n_3 c_3 t_1]^2; \quad (29b)$$

$$e_1 = -2t_1 n_2(t_3 - t_1 n_1),$$

$$e_2 = 2[(1+n_3^2)c_1 c_3 + n_3 n_4(c_1 c_4 - c_2 c_3) - c_2 c_4(1+n_4^2)],$$

$$e_3 = 2n_1 n_2,$$

$$e_4 = -2[(t_3 - t_1 n_4)c_4 - n_3 c_3 t_1][c_1 n_3 t_1 + c_2(t_3 - t_1 n_4)]; \quad (29c)$$

and

$$h_1 = (t_1 n_2)^2,$$

$$h_2 = (1 + n_3^2)c_1^2 - 2c_1c_2n_3n_4 + (1 + n_4^2)c_2^2,$$

$$h_3 = (1 + n_2^2),$$

$$h_4 = [c_1n_3t_1 + c_2(t_3 - t_1n_4)]^2. \quad (29d)$$

Efficient subroutines exist (e.g. NAG) for obtaining the four roots of the polynomial. Having obtained a_1, **R** can be calculated using (22)- (26) and (28). Since **R** is real, only real roots are of interest. Of the real roots, only that which yields positive depth (both X_3 and $X'_3 > 0$) for all points is acceptable. Empirically, the polynomial always appears to have two real roots. Each root has a single combination of the signs of R_{11} and R_{22} which yields positive depths for all data points. The nonveridical solution, however, produces a large origin shift (typically five times the image width) in one image, and small depths (typically a few tenths of the interoccular distance). If the positions of the natural origins are known even roughly (e.g., they may be known to lie somewhere within the pictures), the veridical solution can be chosen quite unambiguously.

Given **t** it is thus possible to compute **R**, and vice versa. Hence Q can also be computed. From (8) or (17), after rescaling Q'', the unknown coordinates (u_1, u_2) and (u'_1, u'_2) of the natural origins can also be obtained. The image coordinates can then be appropriately transformed into their natural systems, whence depth can be calculated by the method prescribed by Longuet-Higgins:

$$X_3 = [(\mathbf{R}_1 - x'_1\mathbf{R}_3)\cdot\mathbf{t}]/[(\mathbf{R}_1 - x'_1\mathbf{R}_3)\cdot\mathbf{x}], \quad (30)$$

$$X_1 = x_1X_3, \quad X_2 = x_2X_3, \quad (31)$$

and

$$X'_j = R_{jk}(X_k - t_k). \quad (j,k = 1,2,3) \quad (32)$$

Note that $\mathbf{x}, \mathbf{x}', \mathbf{X}, \mathbf{X}'$ are now in the natural image coordinate system.

6. SUMMARY

Given 8 corresponding points in two images without any registration information whatsoever, it is possible to calculate the two epicentres and the relation governing the pairs of epipolar lines. The rest of the matching problem reduces to one dimensional searches along the epipolar lines and can be automated using any stereo matching algorithm.

Although it would appear that structure cannot be inferred from the image data alone in the absence of any registration information whatever, full registration information is also not necessary. For example, given either (a) the direction of displacement (two direction cosines) of one camera with respect to the optic axis of the other, or, (b) the orientation of the optic axis (two angles) of one camera with respect to that of the other, methods were described to obtain structure.

7. ACKNOWLEDGEMENTS

It is a pleasure to thank Bernard Buxton for his critical reading of the manuscript.

8. REFERENCES

[1] Longuet-Higgins H C, "A computer algorithm for reconstructing a scene from two projections", **Nature**, 293 (1981) 133-5.

[2] Longuet-Higgins H C, "The reconstruction of a scene from two projections - configurations that defeat the 8-point algorithm", in the *Proceedings of the First International Conference on the Applications of Artificial Intelligence* , Denver, Colorado, USA (1984) 395-7.

[3] Arfken G, *Mathematical methods for physicists.* New York, USA: Academic Press. (1970) pp 178-181 (second edition).

[4] The NAG (Numerical Algorithms Group) Fortran Library, subroutine C02AEF.