An Investigation of the Essential Matrix

Richard I. Hartley

G.E. CRD, Schenectady, NY, 12301.

Contents

1 Introduction **2**

2 Camera Models **2**

3 Calibrated Cameras **4**

3.1 The 8 point Algorithm 4

3.2 Numerical Considerations 6

3.3 Algorithm Outline 8

4 Epipolar Geometry **9**

4.1 Condition for Q to be an essential matrix. 11

4.2 Interpretation of the matrix Q 12

5 Computation of the Essential Matrix **13**

5.1 Solution given 7 image correspondences 14

5.2 Solution given 8 or more points 15

6 Realization of Essential Matrices. **16**

7 Special Realizations **19**

7.1 Determination of Camera Magnifications 20

7.2 Uniqueness of Solution 28

7.3 Failure to find a solution 30

7.4 Completing the algorithm 30

7.5 Algorithm Outline 31

8 Practical Results **32**

9 Compatible Matrices. **32**

9.1 Points lying in planes. 33

9.2 Computation of M. 34

9.3 Solution of the constrained minimization problem. 36

9.4 Joint Epipolar Projections. 37

9.5 Determination of the resampling transformation. 38

1 Introduction

Notation : The symbol $\mathbf{u}$ (in bold type) represents a column vector. We will use the letters u, v and w for homogeneous coordinates in image-space In particular, the symbol $\mathbf{u}$ represents the column vector $(u, v, w)^\top$. Object space points will also be represented by homogeneous coordinates x, y, z and t, or more often x, y, z and 1. The symbol $\mathbf{x}$ will represent a point in three-dimensional projective plane represented in homogeneous coordinates.

When a vector is represented by a single letter (for example $\mathbf{a}$), it is assumed to be a column vector. The correponding row vector is written $\mathbf{a}^\top$. On the other hand (a_1, a_2, a_3) represents a row vector. The corresponding column vector is denoted by $(a_1, a_2, a_3)^\top$

If two vectors or matrices, A and B are equal up to a non-zero scale factor, then we write $A \approx B$ to express this fact.

If A is a square matrix then its matrix of cofactors is denoted by A^*. The following identities are well known : $A^*A = AA^* = \det(A)I$ where I is the identity matrix. In particular, if A is an invertible matrix, then $A^* \approx (A^\top)^{-1}$.

2 Camera Models

The general model of a perspective camera that will be used here is that represented by an arbitrary 3×4 matrix, P, of rank 3, known as the *camera matrix*. The camera matrix transforms points in 3-dimensional projective space to points in 2-dimensional projective space according to the equation

$$\mathbf{u} = P\mathbf{x}$$

where $\mathbf{u} = (u, v, w)^\top$ and $\mathbf{x} = (x, y, z, 1)^\top$. The camera matrix P is defined up to a scale factor only, and hence has 11 independent entries. This was the representation of the imaging process considered by Strat [9]. As shown by Strat, this model allows for the modeling of several parameters, in particular:

1. The location and orientation of the camera.
2. The principal point offsets in the image space.
3. Unequal scale factors in two directions parallel to the axes in image space.

This accounts for 10 of the total 11 entries in the camera matrix. It may be seen that if unequal stretches in two directions **not** aligned with the image axes are allowed, then a further 11-th camera parameter may be defined. Thus, the imaging model considered here is quite general. In practical cases, the focal length (magnification) of the camera may not be known, and neither may be the principal point offsets. Strat [9] gives an example of an image where the camera parameters take on surprising values. Our purpose in treating general camera transforms is to avoid the necessity for arbitrary assumptions about the image.

In general it will be assumed that the camera matrix P can be subdivided into blocks as $P = (M \mid -MT)$ where M is a 3×3 non-singular matrix and T is a column vector. The assumption of non-singularity of M means precisely that the camera is not at infinity. That is, we are assuming perspective, rather than orthographic projection. This is not a serious restriction, since a projective transformation of 3-space can be applied to bring the camera to a non-infinite point without altering the validity of much of the discussion.

So, given that $P = (M \mid -MT)$ with M non-singular, the QR-factorization ([5]) method may be applied to matrix M to provide a factorization $M = KR$, where R is a rotation matrix and K is upper triangular. So we may write

$$P = (M \mid -MT) = K(R \mid -RT)\ . \tag{1}$$

The upper-triangular matrix K represents the "internal parameters" of the camera and may be thought of as describing a transformation in image space. The task of determining the internal parameters of the camera is known as "calibration". Matrix $(R \mid -RT)$ represents the "external parameters" of the camera and describes a transformation of coordinates in object space. The reason for writing the external parameter matrix in the form $(R \mid -RT)$ will be explained now.

The matrix of external parameters represents a simple pinhole camera model. If the camera is located at a point $T' = (t_x, t_y, t_z, 1)^\top$ with orientation represented by the rotation matrix R' relative to a fixed coordinate frame , then a simple pinhole camera model will take a point $\mathbf{x} = (x, y, z, 1)^\top$ to $\mathbf{u} = (u, v, w)^\top = R'\left((x, y, z)^\top - (t_x, t_y, t_z)^\top\right)$. Using homogeneous coordinates in both object and image space, this equation may be represented in matrix form as

$$\mathbf{u} = (R' \mid -R'T')\mathbf{x}$$

Comparing this equation with (1) it may be seen that vector $T = T'$ represents the location of the camera, and $R = R'$ is the rotation of the camera.

3 Calibrated Cameras

First, I will derive the 8-point algorithm of Longuet-Higgins in order to fix notation and to gain some insight into its properties. Alternative derivations were given in [11] and [12]. This section deals with calibrated cameras, that is, the matrix K of internal camera parameters is assumed to be the identity.

3.1 The 8 point Algorithm

We consider the case of two cameras, one of which is situated at the origin of object space coordinates, and one which is displaced from it. Thus the two camera matrices are assumed to be $P = (I \mid 0)$ and $P' = (R \mid -RT)$.

A transformation will now be defined between the 2-dimensional projective plane of image coordinates in image 1 and the pencil of epipolar lines in the second image. As is well known, given a point $\mathbf{u} = (u, v, w)^\top$ in image 1, the corresponding point $\mathbf{u'} = (u', v', w')^\top$ in image 2 must lie on a certain epipolar line, which is the image under P' of the set $\mathcal{L}$ of all points $(x, y, z, 1)^\top$ which map under P to $\mathbf{u}$. To determine this line one may identify two points in $\mathcal{L}$, namely the camera origin $(0, 0, 0, 1)^\top$ and the point at infinity, $(u, v, w, 0)^\top$. The images of these two points under P' are $-RT$ and $R\mathbf{u}$ respectively and the line that passes through these two points is given in homogeneous coordinates by the cross product,

$$(p, q, r)^\top = RT \times R\mathbf{u} = R\,(T \times \mathbf{u}) \quad . \tag{2}$$

Here $(p, q, r)^\top$ represents the line $pu' + qv' + rw' = 0$.

Now, a new piece of notation will be introduced. For any vector $\mathbf{t} = (t_x, t_y, t_y)^\top$ we define a skew-symmetric matrix, $S(\mathbf{t})$ according to

$$S(\mathbf{t}) = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} \quad . \tag{3}$$

An important fact about $S(\mathbf{t})$ is given by

Proposition 1. *For any vectors* $\mathbf{t}$ *and* $\mathbf{u}$, $\mathbf{t} \times \mathbf{u} = S(\mathbf{t})\mathbf{u}$.

Matrix $S(\mathbf{t})$ is a singular matrix of rank 2, unless $\mathbf{t} = 0$. Furthermore, the null-space of $S(\mathbf{t})$ is generated by the vector $\mathbf{t}$. This means that $\mathbf{t}^\top S(\mathbf{t}) = S(\mathbf{t})\,\mathbf{t} = 0$ and that any other vector annihilated by $S(\mathbf{t})$ is a scalar multiple of $\mathbf{t}$.

With this notation, equation (2) may be written as

$$(p, q, r)^\top = RS(T)\mathbf{u} \quad . \tag{4}$$

Since a point $\mathbf{u}'$ in the second image corresponding to $\mathbf{u}$ must lie on the epipolar line, we have the important relation

$$\mathbf{u}'^\top Q \mathbf{u} = 0 \tag{5}$$

where $Q = RS(T)$. This relationship is due to Longuet-Higgins ([11]).

As is well known, given 8 correspondences or more, the matrix Q may be computed by solving a (possibly overdetermined) set of linear equations. In order to compute the second camera transform, P', it is necessary to factor Q into the product RS of a rotation matrix and a skew-symmetric matrix. Longuet-Higgins ([11]) gives a rather involved, and apparently numerically somewhat unstable method of doing this. I will give an alternative method of factoring the Q matrix based on the Singular Value Decomposition ([**?**]). The following result may be verified.

Theorem 2. *A 3×3 real matrix Q can be factored as the product of a rotation matrix and a non-zero skew symmetric matrix if and only if Q has two equal non-zero singular values and one singular value equal to 0.*

A proof is contained in [10]. This theorem leads to an easy method for factoring any matrix into a product RS, when possible.

Theorem 3. *Suppose the matrix Q can be factored into a product RS where R is orthogonal and S is skew-symmetric. Let the Singular Value Decomposition of Q be $UDV^\top$ where $D = \mathrm{diag}(k, k, 0)$. Then up to a scale factor the factorization is one of the following:*

$$S \approx VZV^\top \ ; \ R \approx UEV^\top \ \text{ or } \ UE^\top V^\top \ ; \ Q \approx RS \ . \tag{6}$$

where

$$E = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \ Z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ . \tag{7}$$

Proof. That the given factorization is valid is true by inspection. It was shown in [11] that there are only two possible factorizations of the matrix Q, so they must be the ones given in (6). □

For the factorization given in (6) the translation vector T is equal to $V\,(0,0,1)^\top$ since this ensures that $ST = 0$ as required by (3). Furthermore $||T|| = 1$, which is a convenient normalization suggested in [11]. As remarked by Longuet-Higgins, the correct solution to the camera placement problem may be chosen based on the requirement that the visible points be in front of both cameras

([11]). There are four possible rotation/translation pairs that must be considered based on the two possible choices of R and two possible signs of T. Therefore, since $-RT = -UEV^\top V(0,0,1)^\top = -U(0,0,1)^\top$ the requisite camera matrix $P' = (R \mid -RT)$ is equal to $(UEV^\top \mid -U(0,0,1)^\top)$ or one of the obvious alternatives.

3.2 Numerical Considerations

In any practical application, the matrix Q found will not factor exactly in the required manner because of inaccuracies of measurement. In this case, the requirement will be to find the matrix closest to Q that does factor into a product RS. To quantify the notion of "closeness" of two matrices, we use the (Frobenius norm [**?**]), which is defined by

$$||X|| = \left(\sum X_{ij}^2\right)^{1/2}$$

for any matrix X. A useful property of the Frobenius norm is

Lemma 4. *If U is an orthogonal matrix, then $||X|| = ||UX||$ for any matrix X.*

The proof is straightforward, and is therefore omitted.

Given Q, we wish to find the matrix $Q' = RS$ such that $||Q - Q'||$ is minimized. The following theorem gives the solution to that problem, and shows that the factorization given 6 is numerically optimal.

Theorem 5. *Let Q be any 3×3 matrix and $Q = UDV^\top$ be its Singular Value Decomposition in which $D = \mathrm{diag}(r, s, t)$ and $r \geq s \geq t$. Define the matrix Q' by $Q' = UD'V^\top$ where $D' = \mathrm{diag}(k, k, 0)$ and $k = (r + s)/2$. Then Q' is the matrix closest to Q in Frobenius norm which satisfies the condition $Q' = RS$, where R is a rotation and S is skew-symmetric. Furthermore, the factorization is given up to sign and scale by (6).*

The proof of Theorem 5 will be given in a series of lemmas. The first lemma gives bounds for the singular values of a matrix in terms of the entries of the matrix.

Lemma 6. *Suppose X is a matrix with singular values $s_1, s_2, \ldots s_n$ where s_1 is the largest and s_n is the smallest. Let $\mathbf{c}$ be any column of X, then $s_n \leq ||\mathbf{c}|| \leq s_1$.*

Proof. Let $X = UDV^\top$ be the Singular Value Decomposition of X, and $D = \mathrm{diag}(s_1, s_2, \ldots, s_n)$. Let the i-th column of $DV^\top$ be $\mathbf{c}'_i$ and the i-th row of V be $\mathbf{v}_i{}^\top$. Then, $\mathbf{c}'_i = (s_1 v_{i1}, s_2 v_{i2}, \ldots, s_n v_{in})^\top$. Therefore, $s_n ||\mathbf{v}_i{}^\top|| \leq ||\mathbf{c}'_i|| \leq$

$s_1||\mathbf{v}_i{}^\top||$, and so $s_n \leq ||\mathbf{c}'_i|| \leq s_1$, since $||\mathbf{v}_i{}^\top|| = 1$. Now the i-th column of $X = UDV^\top$ is $\mathbf{c}_i = U\mathbf{c}'_i$. The required result now follows from the fact that $||U\mathbf{c}'_i|| = ||\mathbf{c}'_i||$. □

As a simple consequence of Lemma 6, we have

Lemma 7. *Suppose*

$$X = U\,\mathrm{diag}(r, s, t)\,V^\top = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}$$

with $r \geq s \geq t \geq 0$ *and* U, V *orthogonal. Then*

1. $t^2 \leq c^2 + f^2 + j^2$.
2. $rs \geq ae - bd$.

Proof. The first statement of Lemma 7 follows directly from Lemma 6. To prove the second statement, note that $ae - bd$ is one of the entries of the matrix X^*, whereas rs is the largest singular value of $X^* = U\,\mathrm{diag}(st, rt, rs)\,V^\top$. By Lemma 6, the largest singular value of any matrix is greater than any individual element in the matrix. □

We are now ready to prove the particular case of Theorem 5 for the case where Q is a diagonal matrix.

Lemma 8. *Let* $Q = \mathrm{diag}(r, s, t)$ *with* $r \geq s \geq t \geq 0$ *and* $Q' = \mathrm{diag}(\lambda, \lambda, 0)$ *where* $\lambda = (r + s)/2$. *If* Q'' *is any other matrix of the form* $Q'' = U^\top D'' V$ *where* $D'' = \mathrm{diag}(\lambda, \lambda, 0)$, *and* $\lambda > 0$, *then*

$$||Q - Q''|| \geq ||Q - Q'||.$$

Proof. Since $||Q - Q''|| = ||UQV^\top - UQ''V^\top|| = ||UQV^\top - D''||$ it is sufficient to prove that for all U, V and D'',

$$||UQV^\top - D|| \geq ||Q - Q'||\ .$$

Let $UQV^\top$ be the matrix denoted X in Lemma 7. By choosing a different Q if necessary to change the signs and order of the rows of $UQV^\top$, it may be assumed that $a \geq e \geq j \geq 0$, since this can be done without increasing the value of $||UQV^\top - D''||$. Furthermore, this norm takes on its minimum value when the non-zero singular values of D'' are equal to $(a+e)/2$, and so we may assume that this is the case. Let

$$\begin{aligned} E_1 &= ||Q - Q'||^2 \\ &= 2\left(\frac{r-s}{2}\right)^2 + t^2 \\ &= (r^2+s^2+t^2)/2 + (t^2/2 - rs) \end{aligned}$$

and

$$E_2 = 2\left(\frac{a-e}{2}\right)^2 + (b^2+c^2+d^2+f^2+g^2+h^2+j^2) \ .$$

We need to prove that $E_1 \leq E_2$. Now since U and V are orthogonal, $||UQV^\top|| = ||Q||, by Lemma 4$. In other words, $r^2+s^2+t^2 = a^2+b^2+\cdots+j^2$. Therefore,

$$\begin{aligned} E_1 &= \frac{a^2+b^2+\cdots+j^2}{2} + \frac{t^2}{2} - rs \\ &= 2\left(\frac{a-e}{2}\right)^2 + b^2+c^2+d^2+f^2+g^2+j^2 \\ &\quad +ae - (b^2+c^2+d^2+f^2+g^2+j^2)/2 + t^2/2 - rs \\ &= E_2 - (b^2+d^2)/2 - (c^2+f^2+j^2)/2 - g^2/2 + (ae-rs) + t^2/2 \end{aligned}$$

Now, using the inequalities of Lemma 7 we have

$$\begin{aligned} E_1 &\leq E_2 - (b^2+d^2)/2 - g^2/2 + bd \\ &= E_2 - (b-d)^2/2 - g^2/2 \\ &\leq E_2 \end{aligned}$$

as required. □

Now we can prove Theorem 5.

Proof. Suppose $Q'' = U''D''V''^\top$ and $D'' = \mathrm{diag}(\lambda, \lambda, 0)$. Then

$$\begin{aligned} ||Q - Q''|| &= ||UDV^\top - U''D''V''^\top|| \\ &= ||D - (U^\top U'')D''(V''^\top V)|| \\ &\geq ||D - D'|| \quad \text{by Lemma 8} \\ &= ||Q - Q'|| \quad \text{by Lemma 4} \end{aligned}$$

□

3.3 Algorithm Outline

The algorithm for computing relative camera locations for calibrated cameras is as follows.

1. Find Q by solving a set of equations of the form (5).

2. Find the Singular Value Decomposition $Q = UDV^\top$, where $D = \mathrm{diag}(a, b, c)$ and $a \geq b \geq c$.

3. The transformation matrices for the two cameras are $P = (I \mid 0)$ and P' equal to one of the four following matrices.

$$\begin{array}{lcr}
(UEV^\top & \mid & U(0,0,1)^\top) \\
(UEV^\top & \mid & -U(0,0,1)^\top) \\
(UE^\top V^\top & \mid & U(0,0,1)^\top) \\
(UE^\top V^\top & \mid & -U(0,0,1)^\top)
\end{array}$$

The choice between the four transformations for P' is determined by the requirement that the point locations (which may be computed once the cameras are known [11]) must lie in front of both cameras. Geometrically, the camera rotations represented by $UEV^\top$ and $UE^\top V^\top$ differ from each other by a rotation through 180 degrees about the line joining the two cameras. Given this fact, it may be verified geometrically that a single pixel-to-pixel correspondence is enough to eliminate all but one of the four alternative camera placements.

4 Epipolar Geometry

We consider two cameras C and C' in space. Data related to one of the cameras will be denoted using unprimed quantities, and the corresponding data for the second camera will be denoted using primed quantities. The images corresponding to the two cameras will be denoted by J and J' respectively. We suppose that the two cameras are imaging a common set of points. We investigate the correspondence between where a point is seen in one image and where it is seen in the other. Consider a point $\mathbf{x}$ in space which is imaged at position $\mathbf{u}$ in the first image, J. Knowing $\mathbf{u}$, it may be deduced that the point $\mathbf{x}$ must lie somewhere on a line in space, denoted $locus(\mathbf{x})$, extending radially from the camera centre of camera C. This line $locus(\mathbf{x})$ will be imaged by camera C' as a line, temporarily denoted $L(\mathbf{u})$, representing the locus of possible image points $\mathbf{u}'$ for a point that is imaged at $\mathbf{u}$ in the first image. Line $L(\mathbf{u})$ is known as the epipolar line corresponding to $\mathbf{u}$. Since the line $locus(\mathbf{x})$ passes through the camera centre of C, the corresponding epipolar line $L(\mathbf{u})$ must pass through the point in image J' at which the first camera is seen in the second image. In other words, as $\mathbf{u}$ varies, the lines $L(\mathbf{u})$ are all concurrent. The common point of intersection is known as the epipole, $\mathbf{p}'$. Reversing the roles of the two cameras we may define an epipole, $\mathbf{p}$ in the first image. Now, given a pair of coordinates, $\mathbf{u}$ and $\mathbf{u}'$ which are the images in J and J' of a common point $\mathbf{x}$ in space, then point $\mathbf{u}'$ must lie on the epipolar line $L(\mathbf{u})$, and $\mathbf{u}$ must lie on the epipolar line $L(\mathbf{u}')$.

We consider a general pair of camera matrices represented by $P = (M \mid -MT)$ and $P' = (M' \mid -M'T')$ and suppose that $\{\mathbf{u}\}$ and $\{\mathbf{u}'\}$ are matching sets of image points as seen by the two cameras. The essential matrix Q corresponding to the matching points was defined by Longuet-Higgins ([11]) to be a matrix defined by the relation

$$\mathbf{u}_i{}^\top Q \mathbf{u}_i = 0$$

for all i. We will determine what the form of the matrix Q is in terms of P and P'. First, however, we need some algebraic lemmas.

Proposition 9. *For any* 3×3 *matrix* M *and vector* $\mathbf{t}$

$$M^* S(\mathbf{t}) = S(M\mathbf{t})M$$

This formula is true without any restriction on M, as may be verified by direct computation, perhaps using a symbolic manipulation program such as Mathematica [13]. This proposition leads us to a property of the vector (cross) product.

Corollary 10. *If* M *is any* 3×3 *matrix, and* $\mathbf{u}$ *and* $\mathbf{v}$ *are column vectors, then*

$$(M\mathbf{u}) \times (M\mathbf{v}) = M^*(\mathbf{u} \times \mathbf{v}) \ . \tag{8}$$

Proof.

$$\begin{aligned} (M\mathbf{u}) \times (M\mathbf{v}) &= S(M\mathbf{u})M\mathbf{v} && \text{by Proposition 1} \\ &= M^* S(\mathbf{u})\mathbf{v} && \text{by Proposition 9} \\ &= M^*(\mathbf{u} \times \mathbf{v}) && \text{by Proposition 1} \end{aligned}$$

□

Now it is possible to give the form of the essential matrix for uncalibrated cameras.

Theorem 11. *The essential matrix corresponding to the pair of camera matrices* $(M \mid -MT)$ *and* $(M' \mid -M'T')$ *is given by*

$$Q \approx M'^* M^\top S(M(T' - T)) \ .$$

Proof. As with calibrated cameras (see Section 3.1), it is possible to determine the epipolar line corresponding to a point $\mathbf{u}$ in image 1. Two points that must lie on the ray from the first camera center extending through the point $\mathbf{u}$ are the camera center $\begin{pmatrix} T \\ 1 \end{pmatrix}$ of the first camera and the point at infinity $\begin{pmatrix} M^{-1}\mathbf{u} \\ 0 \end{pmatrix}$.

Transform P' takes these two points to the points $M'T - M'T'$ and $M'M^{-1}\,\mathbf{u}$. The line through these points is given by the cross product

$$M'(T - T') \times M'M^{-1}\mathbf{u} \ . \tag{9}$$

Using (8) it is easy to evaluate the cross product (9).

$$M'(T - T') \times M'M^{-1}\mathbf{u} \approx M'^{*}M^{\top}((M(T - T')) \times \mathbf{u}) \tag{10}$$

Now, writing $S = S(M(T-T'))$ as defined in (3), the epipolar line corresponding to the point $\mathbf{u}$ in image 1 is given by :

$$(p, q, r)^{\top} \approx M'^{*}M^{\top}S\mathbf{u} \ . \tag{11}$$

Furthermore, setting $Q = M'^{*}M^{\top}S$

$$\mathbf{u}'Q\mathbf{u}^{\top} = 0 \ . \tag{12}$$

□

We may use Lemma 9 to give a different form for the essential matrix.

Proposition 12. *The essential matrix corresponding to the pair of camera matrices $(M \mid -MT)$ and $(M' \mid -M'T')$ is given by*

$$Q \approx M'^{*}\,S(T' - T)\,.M^{-1}$$

Proof. If M is invertible, it follows from Lemma 9 that $M^{\top}\,S(MT) = S(T)\,M^{-1}$ and so the required formula follows immediately from Theorem 11. □

From this, we may deduce a useful fact.

Proposition 13. *If Q is the essential matrix corresponding to a pair of cameras (P, P'), then $Q^{\top}$ is the essential matrix corresponding to the pair (P', P).*

Proof. Taking the transpose of the matrix Q given in Proposition (12) simply reverses the roles of P and P'. □

4.1 Condition for Q to be an essential matrix.

We may call any matrix that arises from a pair of camera matrices according to the formula given in Proposition (12) an **essential matrix**. We can give the following characterization of essential matrices.

Proposition 14. *The following three conditions are equivalent for any 3×3 real matrix Q.*

1. *Q is an essential matrix.*

2. *Q can be factored as a product $Q = RS$, where S is a non-zero skew-symmetric matrix and R is non-singular.*

3. *Q has rank 2.*

Proof. Proposition 11 shows that an essential matrix can be factored as $Q = RS$. Conversely, if $Q = RS(T)$ where R is non-singular, then Q is the essential matrix corresponding to the pair $(I \mid 0)$ and $(R^* \mid -R^{*-1}T)$. Hence, conditions 1 and 2 are equivalent. Further, since a non-zero skew-symmetric matrix has rank 2, condition 2 implies condition 3. It remains to prove that condition 3 implies condition 2. Therefore, suppose that Q has rank 2, and let the singular value decomposition of Q be $Q = UDV^\top$ where U and V are orthogonal and D is a diagonal matrix $\text{diag}(r, s, 0)$, where r and s are non-zero. Defining matrices E and Z as in (7), it can be be seen that Q factors as

$$Q = \left(U\text{diag}(r, s, 1)EV^\top\right) . \left(VZV^\top\right) \tag{13}$$

The first bracketed term of this product is non-singular, and the second is skew-symmetric. This completes the proof. □

4.2 Interpretation of the matrix Q

We list some properties of the essential matrix Q.

Proposition 15. *Suppose that Q is the essential matrix corresponding to a pair of images (J, J').*

1. *If $\mathbf{u}$ is a point in image J, then the corresponding epipolar line, $L(\mathbf{u})$ in image J' is equal to $Q\mathbf{u}$.*

2. *If $\mathbf{u}'$ is a point in image J', then the corresponding epipolar line, $L(\mathbf{u}')$ in image J is equal to $Q^\top\mathbf{u}'$.*

3. *If $\mathbf{u}$ and $\mathbf{u}'$ are corresponding points in the two images, then $\mathbf{u}'^\top Q\mathbf{u} = 0$.*

4. *The epipole $\mathbf{p}$ is the unique point such that $Q\mathbf{p} = 0$*

5. *Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be points in J, neither of which is equal to the epipole $\mathbf{p}$. Points $\mathbf{u}_1$ and $\mathbf{u}_2$ are on the same epipolar line if and only if $Q\mathbf{u}_1 \approx Q\mathbf{u}_2$.*

Proof.

1. The derivation of Q given in Theorem 11 showed that Q has this property.

2. Reversing the roles of the two cameras is equivalent to replacing Q by $Q^\top$ as was shown in Proposition 13.

3. This property expresses the fact that the point $\mathbf{u}'$ must lie on the epipolar line $Q\mathbf{u}$.

4. The epipole $\mathbf{p}$ is characterized by the fact that for any point $\mathbf{u}'$, the epipolar line $L(\mathbf{u}') = Q^\top \mathbf{u}'$ passes through $\mathbf{p}$. This means that $\mathbf{p}^\top Q^\top \mathbf{u}' = 0$ for all $\mathbf{u}'$, and hence $\mathbf{p}^\top Q^\top = 0$. In using the phrase "unique point such that $Q\mathbf{p} = 0$", the solution $\mathbf{p} = (0, 0, 0)^\top$ is excluded as not being a valid set of homogeneous coordinates. Further, since we are dealing in homogeneous coordinates, non-zero scale factors are insignificant.

5. If $\mathbf{u}_1$ and $\mathbf{u}_2$ are on the same epipolar line, then there exist parameters α and β such that $\mathbf{u}_2 = \alpha\mathbf{u}_1 + \beta\mathbf{p}$, from which it follows that $Q\mathbf{u}_1 \approx Q\mathbf{u}_2$, since $Q\mathbf{p} = 0$. Conversely, suppose $Q\mathbf{u}_1 \approx Q\mathbf{u}_2$. Then, there exists α such that $Q(\mathbf{u}_1 + \alpha\mathbf{u}_2) = 0$. From this it follows by the characterization of $\mathbf{p}$ that $\mathbf{p} \approx \mathbf{u}_1 + \alpha\mathbf{u}_2$, and so $\mathbf{u}_1$ and $\mathbf{u}_2$ lie on the same line through $\mathbf{p}$.

□

In the same way that part 2 of Theorem 15 is derived from part 1, it is possible to derive from parts 4 and 5 of the theorem corresponding statements relating to points in the second image J' in which Q is replaced by $Q^\top$.

Part 5 of Proposition 15 shows that Q defines a one-to-one correspondence between the pencil of epipolar lines in one image and the pencil of epipolar lines in the other image. This correspondence may be understood geometrically as follows. If P is a plane in 3-space passing through the two camera centres, then the image of the plane in each camera is an epipolar line, and so P defines a pair of corresponding epipolar lines. As P varies, the complete pencil of epipolar lines is traced out in each image. It should be remarked that Q acts not by mapping the coordinates of a *line* in one image to the coordinates of a line in the other image, but rather by mapping the coordinates of an arbitrary *point* on an epipolar line in the first image to the coordinates of the corresponding epipolar line in the second image.

5 Computation of the Essential Matrix

We now consider the computation of the essential matrix Q from a set of image correspondences $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$. First, of all it will be seen that the essential matrix can be computed using only 7 points. Higgins [11] gave a method for computing Q from 8 points or more. The method of computing Q from only 7 points is non-linear in contrast to the method of [11] but it is sufficiently simple to be useful if the number of matched points is only 7. In using only 7 matched points to compute the essential matrix accuracy of the matched points is essential, otherwise the result can be badly wrong.

The essential matrix, Q, has 7 degrees of freedom. The matrix has 9 entries, but the scale is insignificant, so we are reduced to 8 degrees of freedom. Furthermore, the determinant of Q must equal zero, and this accounts for one more degree of freedom, reducing the number to 7. It should be no surprise therefore, that Q may be computed from only seven matched points. On the other hand, any number fewer than 7 of arbitrary matched points is insufficient to compute the essential matrix. This is in contrast with known methods for computing the essential matrix from only 5 matched points ([**?**]) in which the essential matrix computed is that of a calibrated camera, and is therefore more restricted, having only 5 degrees of freedom.

5.1 Solution given 7 image correspondences

Suppose that we have seven image-to-image correspondences. From this we can get seven linear equations in terms of the entries of the matrix Q, given by $\mathbf{u}_i'^{\top} Q \mathbf{u}_i = 0$. Since the matrix is defined only up to a scale factor one of the entries in the matrix could be chosen to have a value of 1, for instance $q_{11} = 1$. This choice could cause trouble if the actual value of q_{11} were equal to zero, and hence we prefer to proceed slightly differently.

We have a set of seven linear equations in nine unknowns. If these equations are not linearly independent, then the image-to-image correspondences have been badly chosen, and the system is underdetermined. There will be a family of solutions for the essential matrix. We will not investigate here the problem of specifying which systems of image correspondences are degenerate. This problem was considered for the case of callibrated cameras in [**?**] and [**?**]. Assuming that the equations are linearly independent, then the solution set is two-dimensional and it is possible to find the solution in terms of two parameters λ and μ. Each q_{ij} will be a homogeneous linear expression in λ and μ. Then substituting for each q_{ij}, equation $\det(Q) = 0$ gives a homogeneous cubic equation in λ and μ. This equation may be solved to find solutions (λ, μ) such that $\lambda^2 + \mu^2 = 1$. The solutions $(\lambda, \mu) = (1, 0)$ and $(\lambda, \mu) = (0, 1)$ may be checked individually. If neither solution exists, then we can solve for the ratio λ/μ by solving a cubic equation in one variable. The cubic may be solved either numerically or else by using an exact formula ([**?**]). Finally, the value or values for λ and μ may be used to compute the values of q_{ij}.

Since a cubic equation has at most three solutions of which at least one is real, we have the following existence theorem for essential matrices.

Theorem 16. *If* $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}_i'\}$ *is a set of seven image correspondences in general position, then there exist at least one and at most three essential matrices* Q *such* *that*

$$\mathbf{u}_i'^\top Q \mathbf{u}_i = 0$$

for all i.

The phrase "general position" is used as in [?] to exclude special degenerate cases in which the the linear equation set given by the point correspondences is rank deficient, or in which the resulting matrix Q has rank less than 2. Such degenerate cases are nowhere dense in the set of all possible image correspondences. Theorem 16 giving a count on the number of solutions is related to the main result of [?] where it is shown that for calibrated cameras, 5 image-to-image correspondences give rise to at most 10 solutions and the question is asked what happens when 6 or 7 image correspondences are known. This theorem solves that problem for 7 image correspondences under a less restrictive definition of the camera model. Theorem 16 relates to uncalibrated cameras, whereas [?] consider calibrated cameras.

5.2 Solution given 8 or more points

The method of solving for the essential matrix given 8 image correspondences or more is well known but will be repeated here for later reference. Each image correspondence gives rise to an equation $\mathbf{u}_i'^\top Q \mathbf{u}_i = 0$ in the entries of Q. The complete set of such correspondences may be written as a possibly overconstrained set of equations $A\mathbf{x} = 0$ where x is a vector containing the entries of Q. It must be assumed that A has rank at least 8, otherwise we are dealing with a degenerate set of image correspondences. Since we do not expect an exact solution, the problem becomes

Problem 17. Minimize $||A\mathbf{x}||$ subject to $||\mathbf{x}|| = 1$.

Solution This problem is easily solved by taking the Singular Value Decomposition $A = UDV^\top$ of A. Here, D is a diagonal matrix and V is a 9×9 orthogonal matrix. Writing $\mathbf{x}' = V^\top \mathbf{x}$, since V is orthogonal, $||\mathbf{x}|| = ||\mathbf{x}'||$ and the problem now becomes : minimize $||UD\mathbf{x}'||$ subject to $||\mathbf{x}'|| = 1$. Since U is orthogonal, this becomes : minimize $||D\mathbf{x}'||$ subject to $||\mathbf{x}'|| = 1$. Now D is diagonal so the solution is $\mathbf{x}' = \mathbf{e}_j$ where $\mathbf{e}_j$ is the vector containing all zeros, except for a 1 in the j-th position, and j is the index of column of D containing the smallest singular value. The solution to the original problem is now given by $\mathbf{x} = V\mathbf{x}' = V_j$ where V_j is the j-th column of V. □

The solution given by the method described here will not in general give rise to a matrix Q with determinant zero, hence not a valid essential matrix. If the matrix is used to analyze the epipolar geometry, then the fact that Q is not singular means that the lines $Q\mathbf{u}$ as $\mathbf{u}$ ranges over points of the first image will not be concurrent. Therefore, in many applications it will be appropriate to find the nearest valid essential matrix to the matrix Q found by the above algorithm. The following theorem shows how this is done.

Theorem 18. *If X is any non-singular 3×3 matrix with Singular Value Decomposition $X = U\,\mathrm{diag}(r,s,t)\,V^\top$ where $r \geq s \geq t \geq 0$, then the nearest (in Frobenius norm) rank 2 matrix to X is given by $X' = U\,\mathrm{diag}(r,s,0)\,V^\top$.*

Proof. As with the proof of Theorem 5, we begin by proving this theorem in the case where X is diagonal, say $X = \mathrm{diag}(r,s,t)$ Now if X'' is the matrix $U^\top DV$ where $D'' = \mathrm{diag}(r'', s'', t'')$, then writing

$$UQV^\top = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}$$

we see that

$$\begin{aligned} ||UQV^\top - D''||^2 &= (a-r'')^2 + b^2 + c^2 + d^2 + (e-s'')^2 + f^2 + g^2 + h^2 + j^2 \\ &\geq c^2 + f^2 + j^2 \\ &\geq t^2 \quad \text{by Lemma 7} \\ &= ||Q - Q'||^2 \end{aligned}$$

From this it follows that $||Q - Q''|| \geq ||Q - Q'||$ for the case where Q is diagonal. The general case may be proven using the same argument as in the proof of Theorem 5. □

6 Realization of Essential Matrices.

Given camera matrices P and P' and an essential matrix Q satisfying the relationship expressed in Theorem 11, we say that P and P' *give rise* to the matrix Q, or conversely that $\{P, P'\}$ is a *realization* of the essential matrix Q. The realization of an essential matrix by a pair of camera matrices is not unique, and the goal of this section is to see which camera matrix pairs may realize a given essential matrix.

As is indicated by Proposition (14), an essential matrix Q factors into a product $Q = RS$, where R is a non-singular matrix and S is skew-symmetric. The next lemma shows to what extent this factorization is unique.

Proposition 19. *Let the 3×3 matrix Q factor in two different ways as $Q \approx R_1 S_1 \approx R_2 S_2$ where each S_i is a non-zero skew-symmetric matrix and each R_i is non-singular. Then $S_2 \approx S_1 \approx S(\mathbf{t})$ for some vector $\mathbf{t}$ uniquely determined up to a non-zero scale factor. Further, $R_2 = R_1 + \mathbf{a}.\mathbf{t}^\top$ for some vector $\mathbf{a}$.*

Proof. Since R_1 and R_2 are non-singular, it follows that $Q\mathbf{t} = 0$ if and only if $S_i\mathbf{t} = 0$. From this it follows that the null-spaces of matrices S_1 and S_2 are equal, and so $S_1 \approx S_2$. Matrices R_1 and R_2 must both be solutions of the linear equation $Q \approx RS$. Consequently, they differ by the value $\mathbf{a}.\mathbf{t}^\top$ as required. □

We now prove a theorem which indicates when two pairs of camera matrices correpond to the same essential matrix.

Theorem 20. *Let $\{P_1, P_1'\}$ and $\{P_2', P_2'\}$ be two pairs of camera transforms. Then $\{P_1, P_1'\}$ and $\{P_2, P_2'\}$ correspond to the same essential matrix Q if and only if there exists a 4×4 non-singular matrix H such that $P_2H = P_1$ and $P_2'H = P_1'$.*

Proof. First we prove the **if** part of this theorem. To this purpose, let $\{\mathbf{x}_i^{(1)}\}$ be a set of at least 8 points in 3-dimensional space and let $\{\mathbf{u}_i\}$ and $\{\mathbf{u}_i'\}$ be the corresponding image-space points as imaged by the two cameras P_1 and P_1'. By the definition of the essential matrix, Q is defined by the condition

$$\mathbf{u}_i'^\top Q \mathbf{u}_i = 0$$

for all i. We may assume that the points $\{\mathbf{x}_i^{(1)}\}$ have been chosen in such a way that the matrix Q is uniquely defined up to scale by the above equation. The point configurations to avoid that defeat this definition of the Q matrix are discussed in [**?**], [**?**]. Suppose now that there exists a 4×4 matrix H taking P_1 to P_2 and P_1' to P_2' in the sense specified by the hypotheses of the theorem. For each i let $\mathbf{x}_i^{(2)} = H^{-1}\mathbf{x}_i^{(1)}$. Then we see that

$$P_2\mathbf{x}_i^{(2)} = P_1 H H^{-1}\mathbf{x}_i^{(1)} = P_1\mathbf{x}_i^{(1)} = \mathbf{u}_i$$

and the same holds for the primed system. In other words, the image points $\{\mathbf{u}_i\}$ and $\{\mathbf{u}_i'\}$ are a matching point set with respect to the cameras P_2 and P_2', corresponding to a set of object-space points $\{\mathbf{x}_i^{(2)}\}$. The essential matrix for this pair of cameras is defined by the same relationship

$$\mathbf{u}_i'^\top Q \mathbf{u}_i = 0$$

that defines the essential matrix of the pair P_1 and P_1'. Consequently, the two camera pairs have the same essential matrix.

Now, we turn to the **only if** part of the theorem and assume that two pairs of cameras give rise to the same essential matrix, Q. First, we consider the camera pair $\{(M_i \mid -M_iT_i), (M_i' \mid -M_i'T_i')\}$. It is easily seen that the 4×4 matrix

$$\begin{pmatrix} M_i^{-1} & T_i \\ 0 & 1 \end{pmatrix}$$

transforms this pair to the camera pair $\{(I \mid 0), (M_i' M_i^{-1} \mid -M_i'(T_i' - T_i))\}$. Furthermore by the **if** part of this theorem (or as verified directly using Lemma 11), this new camera pair gives rise to the same Q-matrix as the original pair.

Applying this transformation to each of the camera pairs

$$\{(M_1 \mid -M_1 T_1), (M_1' \mid -M_1' T_1')\} \quad \text{and} \quad \{(M_2 \mid -M_2 T_2), (M_2' \mid -M_2' T_2')\}$$

we see that there is 4×4 matrix transforming one pair to the other if and only if there is such a matrix transforming

$$\{(I \mid 0), (M_1' M_1^{-1} \mid -M_1'(T_1' - T_1))\} \quad \text{to} \quad \{(I \mid 0), (M_2' M_2^{-1} \mid -M_2'(T_2' - T_2))\}$$

Thus, we are reduced to proving the theorem for the case where the first cameras, P_1 and P_2 of each pair are both equal to $(I \mid 0)$. Consequently, simplifying the notation, let $\{(I \mid 0), (M_1 \mid -M_1 T_1)\}$ and $\{(I \mid 0), (M_2 \mid -M_2 T_2)\}$ be two pairs of cameras corresponding to the same essential matrix. According to Theorem 11, the Q-matrices corresponding to the two pairs are $M_1^* S(T_1)$ and $M_2^* S(T_2)$ respectively, and these must be equal (up to scale). According to Lemma 19, $T_1 \approx T_2$ and

$$M_2^* = M_1^* + \mathbf{a} T_1{}^\top$$

for some vector $\mathbf{a}$. Taking the transpose of this last relation yields

$$M_2{}^{-1} = M_1^{-1} + T_1 \mathbf{a}^\top \ . \tag{14}$$

At this point we need to interrupt the proof of the theorem to prove a lemma.

Lemma 21. *For any column vector* $\mathbf{t}$ *and row vector* $\mathbf{a}^\top$*, if* $I - \mathbf{t}\mathbf{a}^\top$ *is invertible then*

$$(I + \mathbf{t}\mathbf{a}^\top)^{-1} = I - k.\mathbf{t}\mathbf{a}^\top$$

where $k = 1/(1 + \mathbf{a}^\top \mathbf{t})$.

Proof. The proof is done by simply multiplying out the two matrices and observing that the product is the identity. One might ask what happens if $1 + \mathbf{a}^\top \mathbf{t} = 0$ in which case k is undefined. The answer is that in that case, $I + \mathbf{t}\mathbf{a}^\top$ is singular, contrary to hypothesis. Details are left to the reader. □

Now we may continue with the proof of Theorem 20. Referring back to (**??**), it follows that

$$\begin{aligned} M_2 &= (M_1^{-1} + T_1 \mathbf{a}^\top)^{-1} \\ &= (M_1^{-1}(I + M_1 T_1 \mathbf{a}^\top))^{-1} \\ &= (I - k.M_1 T_1 \mathbf{a}^\top) M_1 \\ &= M_1 - k.M_1 T_1 (\mathbf{a}^\top M_1) \end{aligned}$$

and

$$\begin{aligned} M_2 T_1 &= M_1 T_1 - k.M_1 T_1 (\mathbf{a}^\top M_1 T_1) \\ &= k'.M_1 T_1 \end{aligned} \tag{15}$$

where $k' = 1 - k.\mathbf{a}^\top M_1 T_1$. Since T_2 is a constant multiple of T_1 we have $M_2 T_2 = k'' M_1 T_1$. From these results, it follows that

$$(M_2 \mid -M_2 T_2) = (M_1 \mid -M_1 T_1) \begin{pmatrix} I & 0 \\ k.\mathbf{a}^\top M_1 & k'' \end{pmatrix}$$

This completes the proof of the theorem. □

7 Special Realizations

This section of the paper will be devoted to the question, given an essential matrix, Q, can one compute the camera transformations that give rise to it. In a general context, the answer is clearly no, since the matrix Q, does not contain enough information. First, we will investigate this from a point of view of degrees of freedom.

Each camera has 11 degrees of freedom, one for each of the entries of a 3×4 matrix, less one for indeterminate scale. Thus the two cameras have a total of 22 degrees of freedom, and clearly they can not all be determined from the 7 degrees of freedom of Q. In fact, Theorem 20 showed that Q determines the two cameras up to an arbitrary 3-dimensional projective transformation. There are 15 degrees of freedom belonging to an arbitrary 3-D projective transformation (one for each entry of a 4×4 matrix, less one for indeterminate scale). We verify that 22 degrees of freedom for the cameras, less 15 for the projective transform equals 7 degrees of freedom in the essential matrix.

Even without this analysis is it obvious that the two cameras can not be fully determined from Q, for Q may be derived from image correspondences, and there is nothing in a set of image correspondences that can be used to derive absolute position, or orientation of the two camera, or global scale. The best one might wish for is a determination of relative parameters of the two cameras, in other words determination of the cameras up to a common Euclidean (or similarity) transformation of 3-space. Such a transformation has 7 degrees of freedom (3 rotations, 3 translations and one scale) which may be used to fix the position and orientation of one of the cameras and the distance between them. The paper of Longuet-Higgins [**?**] shows that for calibrated cameras the determination of relative camera placements is possible. In our case, after factoring out similarity transforms, we are left with $22 - 7 = 15$ degrees of freedom, certainly more than can be determined from the essential matrix.

The 15 degrees of freedom of relative camera determination may be broken up into 5 "external" parameters giving the relative position and orientation of one camera with respect to the other and 10 "internal" parameters. The 5 external parameters are accounted for by the position and orientation of both cameras (a total of 12 parameters) modulo a similarity transformation (7 parameters). The 10 internal parameters are made up as follows : for each camera, two independent magnifications along image axis directions, two components of principal point offset and one skew. This makes 5 internal parameters for each camera, 10 in all.

Longuet-Higgins in his paper ([**?**]) introducing the essential matrix assumes that all the internal parameters are known (in fact assuming that internal camera transformation is the identity transformation). He then shows how to determine the 5 external camera parameters. However, the essential matrix has 7 degrees of freedom, and so there are two extra pieces of information in the essential matrix. It will be shown here how this extra information may be exploited.

7.1 Determination of Camera Magnifications

In many cases it may be assumed that many of the internal camera parameters are known. For instance usually one may assume that the skew parameters of the camera is zero and that the image magnifications in the two image coordinate directions are the same. Often also the position of the principal point is known, either by the presence of fiduccial marks or by an assumption that it lies at the centre of the image. For an image of unknown origin the assumption that the principal point is known is the one most likely to be faulty, and later we will examine the effect of this assumption.

Under the assumptions of the last paragraph, the only remaining internal parameter is the image magnification. The 5 relative external camera parameters plus the two image magnifications makes a total of 7 parameter, and we will show that these may all be derived from the essential matrix Q with its 7 degrees of freedom.

As remarked in [**?**] knowing the internal parameters of the two cameras is equivalent to being able to set them to zero (or other appropriate standard values). Let us make this more explicit.

Proposition 22. *Suppose that an essential matrix, Q corresponds to a pair of cameras with principal point offsets $(u_0, v_0, 1)^\top$ and $(u'_0, v'_0, 1)^\top$. Define a matrix*

$$G = \begin{pmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

and let G' be similarly defined. Then the matrix

$$Q' = G'^{\top} Q G$$

is the essential matrix corresponding to a pair of cameras with zero principal point offsets and all other parameters equal to those of the original cameras.

Proof. If P and P are the original camera matrices, then the matrices GP and $G'P'$ are the camera matrices of cameras with zero principal point offset and all other parameters the same. The result then follows from Proposition (**??**). □

In the same way if other internal camera parameters are known, it is possible to set them to zero values (or other appropriate standard values) by changing the Q matrix. Therefore, we assume during the rest of this section that the principal point offsets and skew of the two cameras are zero, and that the magnifications in the two image axial directions are equal. Let the magnifications of the two cameras be $1/k$ and $1/k'$. Define the matrix

$$K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{pmatrix}$$

and let K' be similarly defined. We assume that the first camera is placed at the origin of the coordinate system and is unrotated with respect to these coordinates. The two camera matrices, therefore are of the form

$$P = (K \mid 0) \quad \text{and} \quad P' = (K'R' \mid -K'R'T') \tag{16}$$

where R' is a rotation matrix. The next theorem gives a general condition for an essential matrix to have a realization of the type (16).

Theorem 23. *An essential matrix Q has a realization of the form 16 if and only if there is a factorization $Q = YS$ where Y is non-singular and S is skew-symmetric, such that*

$$Y^* = K'^2 Y^* K^{-2} \tag{17}$$

Proof. To prove the **only if** part of the theorem, suppose Q has a realization of the form (16). Then, according to Theorem **??**, Q can be written as $Q \approx K'^* R'^* K^{\top} S(KT')$. Since R is orthogonal and K and K' are diagonal, this becomes

$$Q \approx K'^{-1} R' K S(KT') \ .$$

Letting $Y = K'^{-1} R' K$, it follows that $Y^* \approx K' R' K^{-1}$ and so

$$R' \approx K' Y K^{-1} \approx K'^{-1} Y^* K$$

and so $K'^2 Y \approx Y^* K^2$. If in fact $K'^2 Y = xY^* K^2$ for some scalar multiple x, then $K'^2(xY) = (xY)^* K^2$ and so the equality (17) holds for the matrix xY.

Conversely, suppose that $Q = YS(\mathbf{t})$ where $Y^* = K'^2 Y K^{-2}$. Define $R' = K'YK^-1$ and $T' = K^-1\mathbf{t}$. It is easily verified that $R'^* \approx R'$. Therefore, there exists a constant x such that $(xR')^* = xR'$ and so xR' is a rotation matrix. Furthermore, the pair of matrices (16) is a realization of Q. □

Because matrices K and K' are of the form $\text{diag}(1, 1, k)$ and $\text{diag}(1, 1, k')$, multiplication by such matrices does not affect the top left-hand 2×2 block. Therefore, if Y is a matrix satisfying (16) then Y and Y^* are the same in their top left hand four elements. The converse of this is almost true, as the following lemma shows.

Lemma 24. *Let Y be any 3×3 matrix, then Y and Y^* agree in their top left hand 4 elements if and only if one of the following conditions hole.*

1. $\text{diag}(1, 1, 0)Y = Y^* \text{diag}(1, 1, 0)$
2. $Y\text{diag}(1, 1, 0) = Y^* \text{diag}(1, 1, 0)$
3. *there exist matrices* $K = \text{diag}(1, 1, k)$ *and* $K' = \text{diag}(1, 1, k')$ *such that* $Y^* = K'YK$.

Notice that unlike in (17), the matrices K and K' are not squared in option 3 of this lemma.

Proof. (Lemma 24.) Let

$$Y = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}$$

and

$$Y^* = \begin{pmatrix} a & b & c' \\ d & e & f' \\ g' & h' & j' \end{pmatrix}$$

First, we consider a few special cases. If $c = f = 0$, then $g' = 0$ and $h' = 0$. It follows that $\text{diag}(1, 1, 0)Y = Y^* \text{diag}(1, 1, 0)$ as required. On the other hand, if $g = h = 0$, then $c' = f' = 0$ and so $Y\text{diag}(1, 1, 0) = \text{diag}(1, 1, 0)Y^*$.

Excluding these special cases, we may assume that one of c and f is non-zero, and one of g and h is non-zero. Suppose then that $c \neq 0$ and $g \neq 0$. The other cases are handled by changing the roles of the first two rows or columns of the matrix in what follows.

Define $k = g'/g$ and $k' = c'/c$.

Next, we use the relation $Y^\top Y^* = \det(Y)I$ where I is the identity matrix. Computing the $(1,2)$ entry of $Y^\top Y^*$ yields $ab + de + gh' = 0$, and the $(2,1)$ entry gives $ab + de + g'h = 0$. From this it follows that $gh' = g'h$, and since $g \neq 0$ and $g' = kg$, it follows that $h' = kh$. In a similar fashion, we obtain $f' = k'f$. Similarly, computing the $(1,3)$ and $(3,1)$ entries of $Y^\top Y^*$ gives $ac' + df' + gj' = 0$ and $ac + df + g'j = 0$. Multiplying the second of these two equalities by k' and subtracting from the first gives $gj' - g'k'j = 0$, from which follows that $j' = kk'j$. From these equalities, it may be seen that

$$diag(1,1,g')\, Y\, \mathrm{diag}(1,1,c') = \mathrm{diag}(1,1,g)\, Y^*\, \mathrm{diag}(1,1,c)$$

Since g and c are non-zero, condition 3 of the lemma holds. □

We can get more information about the form of the two matrices K_1 and K_2 in Lemma 24.

Lemma 25. *If Y is a 3×3 matrix of rank at least 2, satisfying the condition $Y^* = K'YK$ for some matrices $K = \mathrm{diag}(1,1,k)$ and $K' = \mathrm{diag}(1,1,k')$, then if both k and k' are non-zero then they have the same sign.*

Proof. Assume that $k \neq 0$ and $k' \neq 0$. Let $k = \pm c^2$ and $k' = \pm c'^2$. Consider the matrix $Y' = \mathrm{diag}(c'^{-1}, c'^{-1}, 1) Y \mathrm{diag}(c^{-1}, c^{-1}, 1)$. Then

$$\begin{aligned} Y'^* &= \mathrm{diag}(c'^{-1}, c'^{-1}, c'^{-2}\, Y^*\, \mathrm{diag}(c^{-1}, c^{-1}, c^{-2}) \\ &= \mathrm{diag}(c'^{-1}, c'^{-1}, c'^{-2})\, K'YK\, \mathrm{diag}(c^{-1}, c^{-1}, c^{-2}) \\ &= \mathrm{diag}(c'^{-1}, c'^{-1}, \pm 1)\, Y\, \mathrm{diag}(c^{-1}, c^{-1}, 1) \\ &= \mathrm{diag}(1,1,\pm 1)\, Y'\, \mathrm{diag}(1,1,\pm 1) \end{aligned}$$

Thus, we have reduced to the case where both k and k' are ± 1. It remains to prove that it can not occur that $k = 1$ and $k' = -1$. Assume the opposite, namely that

$$Y'^* = \mathrm{diag}(1,1,-1)Y' \ .$$

Now, let $Y' = U\mathrm{diag}(r,s,t)V^\top$ be the Singular Value Decompositon of Y'. Then $Y'^* = \pm U\mathrm{diag}(st, rt, rs)V^\top$. On the other hand by assumption, $Y'^* = \mathrm{diag}(1,1,-1)U\mathrm{diag}(r,s,t)V^\top$. It follows that

$$\pm U\mathrm{diag}(st, rt, rs)V^\top = \mathrm{diag}(1,1,-1)U\mathrm{diag}(r,s,t)V^\top$$

and hence

$$\pm U\mathrm{diag}(st, rt, rs) = \mathrm{diag}(1,1,-1)U\mathrm{diag}(r,s,t) \ . \tag{18}$$

From this we get that

$$\mathrm{diag}(1,1,-1)U^\top\mathrm{diag}(1,1,-1)U\mathrm{diag}(st, rt, rs) = \pm\mathrm{diag}(r,s,-t) \ . \tag{19}$$

The matrix $\mathrm{diag}(1,1,-1)U^\top\mathrm{diag}(1,1,-1)U$ is easily verified to be a rotation matrix. Now it is exclude by assumption that Y' has rank 1. If Y' has rank 2, then we may assume that $t = 0$. In this case, the ranks of the two matrices on each side of 19 do not match. Therefore, Y' has rank 3, and none of r, s or t is zero. Therefore, from (19) it follows that the matrix

$$\pm\text{diag}(r, s, -t)\text{diag}((st)^{-1}, (rt)^{-1}, (rs)^{-1}) = \pm\text{diag}(r/(st), s/(rt), -t/(rs))$$

is a rotation matrix. Since r, s and t are all positive, the only possibility is that $r = s = t = 1$ and

$$\pm\text{diag}(r/(st), s/(rt), -t/(rs)) = \text{diag}(-1, -1, 1) \ .$$

However, this means that $Y' = UV$, and hence $Y'^{*} = \pm Y'$ which contradicts $Y'^{*} = \text{diag}(1, 1, -1)Y$. This completes the proof. ◻

According to Proposition (11), the essential matrix is of the form

$$Q = K'^{*} R K^{\top} S(KT)$$

Our task, given Q is to determine K, K', R and T.

Proposition (14) showed that an essential matrix factors into a product of a non-singular matrix and a skew-symmetric matrix and Proposition (**??**) showed that this factorization is not unique. Now, we need to look at this factorization again using the singular value decomposition.

Proposition 26. *Suppose $Q = UDW^{\top}$ is the Singular Value Decomposition of an essential matrix, where U and W are real unitary matrices and $D = \text{diag}(r, s, 0)$ is diagonal. Then Q factors into a product of a non-singular matrix and a skew-symmetric matrix as follows :*

$$Q = \left(U \begin{pmatrix} r & 0 & \alpha \\ 0 & s & \beta \\ 0 & 0 & \gamma \end{pmatrix} E^{\top} W^{\top} \right) . (WZW^{\top}) \tag{20}$$

where E and Z are the matrices given in (7).

Furthermore, up to arbitrary scale factors, any factorization must be of this form for suitable choice of α, β and γ.

Proof. By inspection, the factorization is correct and the two parts of the factorization are non-singular and skew-symmetric as required (as long as $\gamma \neq 0$). According to Proposition **??**, the skew-symmetric part of the factorization is uniquely determined up to scale factor. It remains only to show that the first half of 20 gives the general solution for the non-singular part of the factorization. Writing $R = UFE^{\top}W^{\top}$ where F is an arbitrary 3×3 matrix and multiplying out gives $Q = UF\text{diag}(1, 1, 0)V^{\top}$ from which it follows that

$$F\text{diag}(1, 1, 0) = U^{\top}QV = \text{diag}(r, s, 0)$$

and so F must be of the form

$$F = \begin{pmatrix} r & 0 & \alpha \\ 0 & s & \beta \\ 0 & 0 & \gamma \end{pmatrix}$$

as required. □

Since both E and W are orthogonal matrices, we write $V = WE$, and V is orthogonal also. Further, writing

$$X_{\alpha,\beta,\gamma} = \begin{pmatrix} r & 0 & \alpha \\ 0 & s & \beta \\ 0 & 0 & \gamma \end{pmatrix}$$

we obtain the factorization

$$Q = (UX_{\alpha,\beta,\gamma}V^\top).S$$

where S is skew-symmetric. Comparing this factorization with (**??**) gives the relationship

$$K'^{-1}RK \approx (UX_{\alpha,\beta,\gamma}V^\top)$$

for some suitable choice of α, β and γ.

In this equation, only the two orthogonal matrices U and V and the singular values r and s are known, and the goal is to determine K, K' and R. Taking the inverse transpose of both sides of this equation gives another equation

$$K'RK^{-1} \approx (UX^*_{\alpha,\beta,\gamma}V^\top)$$

since matrices R, U and V are equal (at least up to a factor of ± 1) to their inverse transposes, and K, K' are diagonal. The matrix $X^*_{\alpha,\beta,\gamma}$ is given by

$$X^* = X^*_{\alpha,\beta,\gamma} = \begin{pmatrix} s\gamma & 0 & 0 \\ 0 & r\gamma & 0 \\ -s\alpha & -r\beta & rs \end{pmatrix}$$

From **??** and **??** we derive

$$R \approx K'UX_{\alpha,\beta,\gamma}V^\top K^{-1} = K'^{-1}UX^*_{\alpha,\beta,\gamma}V^\top K$$

whence

$$K'^2UX_{\alpha,\beta,\gamma}V^\top = UX^*_{\alpha,\beta,\gamma}V^\top K^2$$

Let the entries of $UXV^\top$ be (f_{ij}) and those of $UX^*V^\top$ be (g_{ij}). Both f_{ij} and g_{ij} are linear expressions in the unknowns α, β and γ. Muliplying out equation **??** gives

$$\begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ k'^2 f_{31} & k'^2 f_{32} & k'^2 f_{33} \end{pmatrix} = x \begin{pmatrix} g_{11} & g_{12} & k^2 g_{13} \\ g_{21} & g_{22} & k^2 g_{23} \\ g_{31} & g_{32} & k^2 g_{33} \end{pmatrix}$$

where x is an unknown scale factor. The top left hand block of this equation set comprises a set of equations of the form

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = x \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

If the scale factor, x were known, then this system could be solved as a set of linear equations in the the variables α, β and γ. Unfortunately, x is not known, and it is necessary to find the value of x before solving for α, β and γ.

Since the entries of the matrices on both sides of **??** are linear expressions in α, β and γ, it is possible to rewrite (**??**) in the form

$$F(\alpha, \beta, \gamma, 1)^\top = x\, G(\alpha, \beta, \gamma, 1)^\top$$

where F and G are 4×4 matrices, each row of F or G corresponding to one of the four entries in the matrices of (**??**). Such a set of equations has a solution only if $\det(F - x\,G) = 0$.

This leads to a polynomial of degree 4 in $x : p(x) = \det(F - x\,G) = 0$. This equation may be solved to give a value of x. It turns out that the form of this equation is particularly simple, however and its solution may be found by taking a single square root. We now investigate the form of the polynomial $p(x)$.

Let $X_{\alpha,\beta,\gamma}$ be written in the form

$$\alpha.\Delta_{13} + \beta.\Delta_{23} + \gamma.\Delta_{33} + (r.\Delta_{11} + s.\Delta_{22}),$$

where Δ_{ij} is the matrix having a 1 in position i, j and zeros elsewhere. Then $UX_{\alpha,\beta,\gamma}V^\top$ is equal to

$$\alpha U\Delta_{13}V^\top + \beta U\Delta_{23}V^\top + \gamma U\Delta_{33}V^\top + rU\Delta_{11}V^\top + sU\Delta_{12}V^\top.$$

It may be verified that the p, q-th entry of the matrix $U\Delta_{ij}V^\top$ is equal to $U_{pi}V_{ql}$. Now suppose the rows of F are ordered corresponding to the entries f_{11}, f_{12}, f_{21} and f_{22} of $UXV^\top$. Then

$$\begin{bmatrix} f_{11} \\ f_{12} \\ f_{21} \\ f_{22} \end{bmatrix} = \begin{pmatrix} U_{11}V_{13} & U_{12}V_{13} & U_{13}V_{13} & r.U_{11}V_{11}+s.U_{12}V_{12} \\ U_{11}V_{23} & U_{12}V_{23} & U_{13}V_{23} & r.U_{11}V_{21}+s.U_{12}V_{22} \\ U_{21}V_{13} & U_{22}V_{13} & U_{23}V_{13} & r.U_{21}V_{11}+s.U_{22}V_{12} \\ U_{21}V_{23} & U_{22}V_{23} & U_{23}V_{23} & r.U_{21}V_{21}+s.U_{22}V_{22} \end{pmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ 1 \end{bmatrix} \tag{21}$$

and F is the matrix in this expression. Similarly, the matrix G is given by

$$G = \begin{pmatrix} -s.U_{13}V_{11} & -r.U_{13}V_{12} & r.U_{12}V_{12}+s.U_{11}V_{11} & rs.U_{13}V_{13} \\ -s.U_{13}V_{21} & -r.U_{13}V_{22} & r.U_{12}V_{22}+s.U_{11}V_{21} & rs.U_{13}V_{23} \\ -s.U_{23}V_{11} & -r.U_{23}V_{12} & r.U_{22}V_{12}+s.U_{21}V_{11} & rs.U_{23}V_{13} \\ -s.U_{23}V_{21} & -r.U_{23}V_{22} & r.U_{22}V_{22}+s.U_{21}V_{21} & rs.U_{23}V_{23} \end{pmatrix} \tag{22}$$

Now, it may be verified easily with a symbolic manipulation program such as Mathematica [13] that $\det(F) = \det(G) = 0$. This is done simply by writing out the matrices as given in (**??**) and (**??**) and computing their determinants. It does not rely on the fact that the matrices U and V are orthogonal. Since these are the constant and degree 4 terms of the polynomial $p(x) = \det(F - x\,G)$ it follows that $p(x)$ is of the form $p(x) = a_1 x + a_2 x^2 + a_3 x^3$. Symbolic manipulation may also be used to verify the following identity : $\det(F+G) + \det(F-G) = 0$, from which it follows that $p(1) + p(-1) = 0$ and hence $a_2 = 0$. Therefore, polynomial $p(x)$ has the form $p(x) = a_1 x + a_3 x^3$.

The polynomial determinant $p(x) = \det(F - x\,G)$ may be evaluated in various ways. It may be computed directly as a polynomial deteriminant in one variable x. This method involving polynomial arithmetic is the one that we have implemented. A somewhat simpler method is just to compute $a_1 + a_3 = p(1) = \det(F - G)$ and $2a_1 + 8a_3 = p(2) = \det(F - 2G)$, from which a_1 and a_3 are readily computed. This method has the advantage that it reqires only the computation of determinants involving real numbers.

Next, we solve the polynomial equation to find x. The solution $x = 0$ may be ignored, since according to (**??**) and (**??**) it would imply that $UX_{\alpha,\beta,\gamma}V^\top = 0$ and hence $X_{\alpha,\beta,\gamma} = 0$, which is not true. We meet a difficulty in finding the other roots of the equation. It is possible that the ratio a_1/a_3 is positive, in which case the two roots of $p(x)$ are imaginary. Since all matrices in equation **??** are intended to be real matrices, this means that there is no solution to equation **??** in this case, and this means that the matrix Q does not belong to a pair of cameras of the type supposed in **??**. More will be said about this point later. For the present we assume that $p(x)$ has two roots of equal magnitude and opposite sign. Let these solutions be x_0 and x_1 where $x_1 = -x_0$.

For each x_i we may now solve (**??**) to find the values of α, β and γ. The solutions will be different. Then from (**??**), the values of k^2 and k'^2 may be read off. In particular, according to (**??**),

$$k'^2 = xg_{31}/f_{31} = xg_{32}/f_{32}$$

$$k^2 = f_{13}/xg_{13} = f_{23}/xg_{23}$$

and as a further identity

$$k'^2 f_{33} = xk^2 g_{33}.$$

We get such a set of solutions for $x = x_0$ and another set for $x = x_1$. We seem to have several solution, namely at least two estimates of k and k' for each of the two values x_i. It will be shown next, however that the two values of k given by (**??**) must be equal, as must the two values of k'. Furthermore the identity (**??**) is always satisfied. Further it will be shown that the two solutions arising from the two different choices of x are in fact the same.

First, we turn to analyse the situation in which a matrix A has the property expressed in (**??**), namely that A and xA^* agree in their four top left entries. First of all, replacing A by $A' = xA$, it follows that A' and A'^* agree in their four top entries. Let us assume that this has been done, and designate A to be this new normalized matrix, dropping the notation A'.

Lemma 27. *If A is a matrix satisfying the condition that A and A^* agree in their top left hand four elements, then either*

i) $A^*_{13} = A^*_{23} = A_{31} = A_{32} = 0$

ii) $A^*_{13} = A^*_{23} = A^*_{33} = 0$ *and there exists a constant, k such that* $A^*_{13} = kA_{13}$ *and* $A^*_{23} = kA_{23}$

or

iii) there exist matrices K and K' of the form $K = \mathrm{diag}(1, 1, k)$ *and* $K' = \mathrm{diag}(1, 1, k')$*, such that* $KA = A^*K'$.

Note that parts (i) of the lemma may be considered a special case of part (iii) if which k and k' are both infinite. Similarly, part (ii) may be considered a special case in which k is finite and k' is infinite.

The effect of this lemma is to ensure that there is essentially only one solution for k^2 and k'^2 to be derived from (**??**).

Next, it will be shown that the solutions for k^2 and k'^2 derived from the different solutions x_0 and x_1 of the equation $p(x) = 0$ are the same.

7.2 Uniqueness of Solution

We reduce the problem to its essentials. Given an essential matrix Q, we define a valid factorization of Q to be a quadruple (G, T, K, L) where G, K, L are 3×3 matrices with $K = \mathrm{diag}(1, 1, k)$ and $L = \mathrm{diag}(1, 1, l)$ and T is a vector, and such a that the two following conditions hold

1. Q factors as $Q \approx GS(T)$
2. $K^2G = G^*L^2$.

This is precisely the problem that we have been solving and have shown that there are at most two solutions. The present goal is to prove that there is at most one valid factorization for Q. We start with a particular case.

Lemma 28. *If Q is a rank 2 matrix with two equal non-zero singular values, then there is only one valid factorization. This is the factorization in which K and L are both the identity matrix, and G is an orthogonal matrix.*

Proof. As we have seen, the general form for the matrix G in the factorization of Q as $GS(T)$ is

$$G \approx U \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & \gamma \end{pmatrix} V^\top .$$

Now, by inspection, we can find two valid factorizations, namely those for which $\alpha = 0, \beta = 0, \gamma = 1$ and $\alpha = 0, \beta = 0, \gamma = -1$. With these choices, we see that $G \approx G^*$ and so both L and K are the identity matrix. □

Now we may prove the general case

Theorem 29. *If Q is an essential matrix, then there is at most one valid factorization for Q.*

Proof. Suppose that we have one factorization given by (G_0, T_0, K_0, L_0). Define a new matrix Q' by $Q' = K_0 Q L_0$. First, we establish the relationship between the valid solutions for Q and those for Q'.

Suppose that (G', T', K', L') is a valid factorization for Q'. Then, $Q' = G'S(T')$ and $K'^2 G' = G'^* L'^2$. In this case,

$$\begin{aligned} Q &= K_0^* Q' L_0^* \\ &= K_0^* G' S(T') L_0^* \\ &= K_0^* G' L_0 S(L_0 T'). \end{aligned}$$

Now, writing $G = K_0^* G' L_0$, it follows that

$$\begin{aligned} K'^2 G &= K'^2 K_0^* G' L_0 \\ &= K_0^* K'^2 G' L_0 \\ &= K_0^* G'^* L'^2 L_0 \\ &= K_0^* G'^* L_0 L'^2 \end{aligned}$$

However, from $G = K_0^* G' L_0$ it follows that $G' = K_0 G L_0^*$ and hence, $G'^* = K_0^* G^* L_0$. Continuing the computation of $K'^2 G$:

$$K'^2 G = K_0^{*2} G^* L_0^2 L'^2$$

That is,

$$(K_0 K')^2 G = G^* (L_0 L')^2$$

and so the quadruple $(K_0^* G' L_0, L_0 T', K_0 K', L_0 L')$ is a valid factorization for Q.

Next, we show that the non-zero singular values of Q' are equal. In particular, it will be shown that $Q' = G'S(T')$ where $G' \approx G'^*$ from which it follows using

Proposition (??) that Q' has equal singular values. By definition,

$$\begin{aligned} Q' &= K_0 Q L_0 \\ &= K_0 G S(T) L_0 \\ &= K_0 G L_0^* S(L_0^* T) \end{aligned}$$

Now, writing $G' = K_0 G L_0^*$

$$\begin{aligned} G'^* &= K_0^* G^* L_0 \\ &= K_0^* K_0^2 G L_0^* \\ &= K_0 G L_0^* \\ &= G' \end{aligned}$$

According to Proposition (??), Q' has two equal non-zero singular values, and so by Lemma (??), the only valid factorizations of Q' are ones in which $L' = K' = I$, the identity matrix. Therefore, the only valid factorizations of Q is given by the quadruple G_0, T_0, K_0, L_0 as required. □

7.3 Failure to find a solution

In the previous section it was shown that the unique solution to our problem is given by any of the equations in the set (??). However, we may once again fail to find a solution if one of the ratios $k'^2 = xg_{31}/f_{31}$ or $k^1 = f_{13}/xg_{13}$ is negative, in which case we can not find a real square root. It is necessary to determine how this may come about.

It is possible that $p(x)$ has no real root, which indicates that no real solution is possible given the assumed camera model. This may mean that the position of the principal points have been wrongly guessed. For a different value of each principal point (that is, a translation of image space coordinates) a solution may be possible, but the solution will be dependent on the particular translations chosen.

7.4 Completing the algorithm

At this point, it is possible to continue and compute the values of the rotation matrix directly. However, it turns out to be more convenient, now that the values of the magnification are known, to revert to the case of a calibrated camera. More particularly, we observe that according to (??), Q may be written as $Q = K'^{-1} Q' K^{-1}$ where $Q' = RS$, and R is a rotation matrix. The original method of Section 3.3 may now be used to solve for the camera matrices derived from Q'. In this way, we find camera models $P = (I \mid 0)$ and $P' = (R \mid -RT)$ for the two cameras corresponding to Q'. Taking account of the magnification matrices K and K', the final estimates of the camera matrices are $(K \mid 0)$ and $(K'R \mid -K'RT)$.

In practice it has been observed that greater numerical accuracy is obtained by repeating the computation of k and k' after replacing Q by Q'. The values of k and k' computed from Q' are very close to 1 and may be used to revise the computed magnifications very slightly. However, such a revision is necessary only because of numerical round-off error in the algorithm and is not strictly necessary.

7.5 Algorithm Outline

Although the mathematical derivation of this algorithm is at times complex, the implementation is not particularly difficult. The steps of the algorithm are reiterated here.

1. Compute a matrix Q such that $(u'_i, v'_i, 1)^\top Q(u_i, v_i, 1) = 0$ for each of several matched pairs (at least 8 in number) by a linear least-squares method.

2. Compute the Singular Value Decomposition $Q \approx UDW^\top$ with $\det(U) = \det(V) = +1$ and set r and s to equal the two largest singular values. Set $V = WE$.

3. Form the matrices F and G given by (21) and (22) and compute the determinant $p(x) = \det(F - x\,G) = a_1 x + a_3 x^3$.

4. If $-a_1/a_3 < 0$ no solution is possible, so stop. Otherwise, let $x = \sqrt{-a_1/a_3}$, one of the roots of $p(x)$.

5. Solve the equation $(F - x\,G)(\alpha, \beta, \gamma, 1)^\top = 0$ to find α, β and γ and use these values to form the matrices $X_{\alpha,\beta,\gamma}$ and $X^*_{\alpha,\beta,\gamma}$ given by (**??**) and (**??**).

6. Form the products $UX_{\alpha,\beta,\gamma}V^\top$ and $UX^*_{\alpha,\beta,\gamma}V^\top$ and observe that the four top left elements of these matrices are the same.

7. Compute k and k' from the equations (**??**) where (f_{ij}) and (g_{ij}) are the entries of the matrices $UX_{\alpha,\beta,\gamma}V^\top$ and $UX^*_{\alpha,\beta,\gamma}V^\top$ respectively. If k and k' are imaginary, then no solution is possible, so stop.

8. Compute the matrix $Q' = K'QK$ where K and K' are the matrices $\mathrm{diag}(1, 1, k)$ and $\mathrm{diag}(1, 1, k')$ respectively.

9. Compute the Singular Value Decomposition of $Q' = U'D'V'^\top$.

10. Set $P = (K \mid 0)$ and set P' to be one of the matrices

$$\begin{array}{lcr} (K'U'EV'^{\top} & \mid & K'U'(0,0,1)^{\top}) \\ (K'U'E^{\top}V'^{\top} & \mid & K'U'(0,0,1)^{\top}) \\ (K'U'EV'^{\top} & \mid & -K'U'(0,0,1)^{\top}) \\ (K'U'E^{\top}V'^{\top} & \mid & -K'U'(0,0,1)^{\top}) \end{array}$$

according to the requirement that the matched points must lie in front of both cameras.

8 Practical Results

This algorithm has been encoded in C and tested on a variety of examples. In the first test, a set of 25 matched points was computed synthetically, corresponding to an oblique placement of two cameras with equal magnification values of 1003. The principal point offset was assumed known. The solution to the relative camera placement problem was computed. The two cameras were computed to have magnifications of 1003.52 and 1003.71, very close to the original. Camera placements and point positions were computed and were found to match the input pixel position data within limits of accuracy. Similarly, the positions in 3-space of the object points matched the known positions to within one part in 10^4.

The algorithm was also tested out on a set of matched points derived from a stereo- matching program, STEREOSYS ([**?**]). A set of 124 matched points were found by an unconstrained hierarchical search. The two images used were 1024×1024 aerial overhead images of the Malibu region with about 40% overlap. The algorithm described here was applied to the set of 124 matched points and relative camera placements and object-point positions were computed. The computed model was then evaluated against the original data. Consequently, the computed camera models were applied to the computed 3-D object points to give new pixel locations which were then compared with the original reference pixel data. The RMS pixel error was found to be 0.11 pixels. In other words, the derived model matches the actual data with a standard deviation of 0.11 pixels. This shows the accuracy not only of the derived camera model, but also the accuracy of the point-matching algorithms.

9 Compatible Matrices.

Definition : Let M be a non-singular 3×3 real matrix and Q be an essential matrix, we say that M is *compatible* with Q if $M^{\top}Q$ is skew-symmetric.

The motivation for this definition is as follows. We wish to define a projective transformation from the image C onto C' in such a way that any point u is mapped onto a point $\mathbf{u}'$ which lies on the epipolar line Qu corresponding to u. If the projective transform is denoted by the non-singular 3×3 matrix, M, then we require that $M\mathbf{u}$ lie on the line $Q\mathbf{u}$, and hence that for all $\mathbf{u}$, $\mathbf{u}^\top M^\top Q\mathbf{u} = 0$. If $M^\top Q$ is thought of as representing a quadratic form, then this last equation implies that the quadratic form is zero, and hence that $M^\top Q$ is skew-symmetric.

The following proposition lists some of the consequences of this definition.

Proposition 30. *Let Q be an essential matrix and p and p' the two epipoles. If M is a matrix compatible with Q, then*

1. $M^\top Q \approx S(p)$

2. $Mp = p'$.

3. M^{-1} *is compatible with* $Q^\top$.

4. *One realization of the essential matrix Q is given by the camera pair* $\{(I\ mid 0), (M \mid -Mp)\})$.

Proof.
1. If $S(T) = M^\top Q$, then $S(T)p = M^\top Qp = 0$. It follows that $p \approx S(T)$ and so $p \approx T$.

2. Since $M^\top Q \approx S(p)$, it follows that $p^\top M^\top Q \approx p^\top S(p) = 0$. Therefore, $Q^\top(Mp) = 0$ and it follows from (**??**) that $Mp \approx p'$.

3. If M is compatible with Q, then $M^\top Q = S(T)$, and hence, $Q \approx M^* S(T)$. Now, applying Proposition (9) gives $Q \approx S(MT)M$, and therefore $QM^{-1} = S(MT)$. It follows that QM^{-1} is skew-symmetric, and so is its transpose, $M^{-1T}Q^\top$. This states that M^{-1} is compatible with $Q^\top$ as required.

4. Follows from part 1 and Proposition (**??**). □

9.1 Points lying in planes.

Theorem 31. *Suppose that we have two cameras, P and P' and a set of points $\{\mathbf{x}_i\}$ lying in a plane not passing through either of the camera centres. Let $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ be the images of the points as imaged by the two cameras, then there is a 2-dimensional projective transform taking each $\mathbf{u}_i$ to $\mathbf{u}'_i$.*

Proof. Though this is a fairly well known fact, a quick proof will be given. By an appropriate choice of projective coordinates, it may be assumed that the points $\{\mathbf{x}_i\}$ lie in the plane at infinity, and hence that all points $\mathbf{x}_i$ are of the form $(x_i, y_i, z_i, 0)$. The camera matrices may be written in the form $(K \mid L)$ and $(K' \mid L')$, and since the camera centres do not lie on the plane at infinity, the matrices K and K' are non-singular. The point $\mathbf{u}_i = K(x_i, y_i, z_i)^\top$ and $\mathbf{u}'_i = K'(x_i, y_i, z_i)^\top$. It follows that $\mathbf{u}'_i = KK'^{-1}\mathbf{u}_i$ and this completes the proof. □

We may prove a converse to this theorem :

Theorem 32. *Suppose that two cameras P_0 and P'_0 are given, and let Q be the corresponding essential matrix. Let M be a non-singular 3×3 matrix compatible with Q. Then there exists a plane π_0 such that if $\mathbf{x}$ is any point lying in π and $\mathbf{u}$ and $\mathbf{u}'$ are the images of the point $\mathbf{x}$ in the two images, then $\mathbf{u}' = M\mathbf{u}$.*

Proof. Given compatible Q and M, an alternative realization of Q is given by the camera pair $\{P_1, P'_1\} = \{(I \mid 0), (M \mid -Mp)\}$, where p is the epipole. Let the plane π_1 be the plane at infinity, and suppose $\mathbf{x} = (x, y, z, 0)^\top$. Then $\mathbf{u} = P_1\mathbf{x} = (x, y, z)^\top$ and $\mathbf{u}' = P'_1\mathbf{x} = M(x, y, z)^\top$, and so $\mathbf{u}' = M\mathbf{u}$. Now, consider the camera pair $\{P_0, P'_0\}$. According to Theorem **??**, there exists a non-singular 4×4 matrix, H such that $P_0 = P_1H$ and $P'_0 = P'_1H$. Then define π_0 to be the plane $H^{-1}\pi_1$ and consider a point $H^{-1}\mathbf{x}$ in this plane Then, $P_0H^{-1}\mathbf{x} = P_1HH^{-1}\mathbf{x} = P_1\mathbf{x} = \mathbf{u}$, and similarly, $P'_0H^{-1}\mathbf{x} = \mathbf{u}' = M\mathbf{u}$. In other words, the plane π_0 satisfies the requirements of the theorem. □

According to Theorems **??** and **??**, the choice of a particular perspective transformation, M, compatible with Q is equivalent to choosing a plane in space. For points $\mathbf{x}$ on this plane, the mapping $u \leftrightarrow u'$ relating the coordinates in the two images is given by the projective transformation, M. For points above and below the plane, an image point $\mathbf{u}$ will correspond to a point $\mathbf{u}'$ that lies along the epipolar line through Mu.

9.2 Computation of M.

For point sets that lie close to a plane and are viewed from two cameras, the projective image-to-image transform denoted by a non-singular matrix M compatible with the essential matrix Q may be a close approximation to the actual image-to-image mapping given by point matching. For instance, in the standard stereo-matching problem where the task is to find a pixel u' in a second image to match the pixel u in the first image, a good approximation to u' may be given by Mu. The best match, u' may be found by an epipolar search centred at the point Mu.

Given a set of image-to-image correspondences $\mathbf{u}_i$ and $\mathbf{u}'_i$ and an essential matrix Q computed from these correspondences, our goal is to find a projective transformation given by a matrix M, compatible with Q, such that $\sum ||\mathbf{u}'_i - M\mathbf{u}||^2$ is minimized.

This may be considered as a constrained minimization problem. First, we consider the constraints. The condition that $M^\top Q$ should be skew-symmetric leads to a set of 6 equations in the 9 entries of M.

$$\begin{aligned}
m_{11}q_{11} + m_{21}q_{21} + m_{31}q_{31} &= 0 \\
m_{12}q_{12} + m_{22}q_{22} + m_{32}q_{32} &= 0 \\
m_{13}q_{13} + m_{23}q_{23} + m_{33}q_{33} &= 0 \\
m_{11}q_{12} + m_{21}q_{22} + m_{31}q_{32} &= -(m_{12}q_{11} + m_{22}q_{21} + m_{32}q_{31}) \\
m_{11}q_{13} + m_{21}q_{23} + m_{31}q_{33} &= -(m_{13}q_{11} + m_{23}q_{21} + m_{33}q_{31}) \\
m_{12}q_{13} + m_{22}q_{23} + m_{32}q_{33} &= -(m_{13}q_{12} + m_{23}q_{22} + m_{33}q_{32}) \qquad (23)
\end{aligned}$$

The entries q_{ij} of the matrix Q are known, so this gives a set of 6 known constraints on the entries of M. However, because of the fact that Q is singular of rank 2, there is one redundant restraint as will be shown next. The matrix of coefficients in the above array may be written in the form

$$\left(\begin{array}{ccc}
Q_1 & 0 & 0 \\
0 & Q_2 & 0 \\
0 & 0 & Q_3 \\
Q_2 & -Q_1 & 0 \\
-Q_3 & 0 & Q_1 \\
0 & Q_3 & -Q_2
\end{array} \right)$$

where Q_i represents the three entries $q_{1i}q_{2i}q_{3i}$ making up the i-th column of the matrix Q. First, we consider the case where $Q_1 = 0$. Since Q has rank 2, it can not be the case that Q_2 is a multiple of Q_3. In this case, the first row of the matrix is zero, but the remaining five rows are linearly independent. The same argument holds if $Q_2 = 0$ or $Q_3 = 0$.

Next, consider the case where $Q_2 = \alpha Q_1$. Since Q has rank 2, it can not be that $Q_3 = \beta Q_1$ or $Q_3 = \beta Q_2$. In this case, the fourth row of matrix (**??**) is dependent on the first two rows, but the other five rows are linearly independent.

Finally, consider the case where no column of Q is a simple multiple of another row, but $\alpha Q_1 + \beta Q_2 + \gamma Q_3 = 0$. Then, it may be verified that multiplying the rows of the matrix by factors $\alpha, -\beta^2/\alpha, -\gamma^2/\alpha, \beta, -\gamma, -\beta\gamma/\alpha$ and adding results in 0. In other words, the rows are linearly dependent. On the other hand, it can be shown by a straight-forward argument that any five of the rows are linearly independent.

Thus, in general, we have a set of 6 restraints on the entries of M, only 5 of which are linearly independent.

Next, consider the minimization of the goal function. We have a set of correspondences, $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$ and the task is to find the M that best approximates the correspondence. We can cast this problem in the form of a set of linear equations in the entries of M as follows. Let M_i be the i-th row of M. Then the basic equation $\mathbf{u}' = M\mathbf{u}$ can be written as

$$\begin{aligned} m_{11}u_i + m_{12}v_i + m_{13} &= w'_i u'_i \\ m_{21}u_i + m_{22}v_i + m_{23} &= w'_i v'_i \\ m_{31}u_i + m_{32}v_i + m_{33} &= w'_i \end{aligned}$$

We can substitute the third equation into the first two to get two equations linear in the m_{ij}

$$\begin{aligned} m_{11}u_i + m_{12}v_i + m_{13} &= m_{31}u_i u'_i + m_{32}v_i u'_i + m_{33}u'_i \\ m_{21}u_i + m_{22}v_i + m_{23} &= m_{31}u_i v'_i + m_{32}v_i v'_i + m_{33}v'_i \end{aligned}$$

The method used here is essentially that described by Sutherland [**?**]. It does not minimize exactly the squared error $\sum ||M\mathbf{u}_i - \mathbf{u}'_i||^2$, but rather a sum weighted by w'_i. However, we accept this limitation in order to use a fast linear technique. This has not caused any problems in practice.

Since the matrix, M is determined only up to a scale factor, we seek an appropriately normalized solution. In particular, we seek a solution to (**??**) for which $\sum m_{ij}^2 = 1$.

9.3 Solution of the constrained minimization problem.

Writing the constraint equations as $Bx = 0$ and the equations (**??**) as $Ax = 0$, our task is to find the solution x that fulfills the constraints exactly and most nearly satisfies the conditions $Ax = 0$. More specifically, our task is to minimize $||Ax||$ subject to $||x|| = 1$ and $Bx = 0$. One method of solving this is to proceed as follows.

Extend B to a square matrix B' by the addition of 3 rows of zeros. Let the Singular Value Decomposition of B' be $B' = UDV^\top$, where V is a 9×9 orthogonal matrix and D is a diagonal matrix with 5 non-zero singular values, which may be arranged to appear in the top left-hand corner. Writing $x' = V^\top x$, and $x = Vx'$, we see that $||x|| = ||x'||$. The problem now becomes, minimize $||AVx'||$ subject to $UDx' = 0$ and $||x'|| = 1$. The condition that $UDx' = 0$ means $Dx' = 0$, and hence the first five entries of x' are zero, since D is diagonal with its first five entries non-zero. Therefore, let A'' be the matrix formed from AV by dropping the first five columns. We now solve the problem : minimize $||A''x''||$ subject to $||x''|| = 1$. This is a straightforward unconstrained minimization problem and may be solved by the method used in Section (**??**). Once x'' is found, vector x' is obtained from it by appending 5 zeros. Finally, x is found according to the equation $x = Vx'$.

9.4 Joint Epipolar Projections.

Consider the situation in which we have two images J_0 and J'_0 of the same scene taken from different unknown viewpoints by cameras C and C'. Suppose that the essential matrix, Q_0 and a compatible projective transformation M_0 have been computed based on a set of image-to-image correspondences between the two images. In general, the epipolar lines in the two images will run in quite different directions, and epipolar lines through different points will not be parallel (since they meet at the epipole). The goal in this next section is to define two perspective transformations, F and F', to be applied to the two images so that if the image J_0 is transformed according to the perspective transformation F, and J'_0 is transformed according to the perpsective transformation F', then the resulting images J_1 and J'_1 correspond to each other in a particularly simple manner. In particular, the epipolar lines in the new images will be horizontal and parallel. Further, the resulting image-to-image transformation M_1 will be the identity mapping.

Let the epipole in the first image be p_0. The epipolar lines all pass through p_0 and hence are not parallel. Our goal is to transform the image so that the epipole is moved to the point $(1, 0, 0)^\top$ in homogeneous coordinates. This point is the point at infinity in the direction along the x axis. If this is the new epipole, then the epipolar lines will all be parallel with the x axis. Therefore, let F be a perspective transformation that sends the point p to $(1, 0, 0)^\top$. This is not by itself sufficient information to determine F uniquely, and F will be more exactly determined later.

Proposition 33. *Suppose there exist two images J_0 and J'_0 with corresponding essential matrix Q and compatible perspective transform M. Let p and p' be the two epipoles. Let F be a homogeneous transformation such that $Fp = (1, 0, 0)^\top$ and let $F' = FM^{-1}$. Then, $F'p' = (1, 0, 0)^\top$.*

Proof. $F'p' = F.M^{-1}p' = Fp = (0, 0, 1)^\top$. □

We now use the two projective transformations F and F' to resample the two images J_0 and J'_0 to give two new images J_1 and J'_1. By this is meant that J_1 is related to J_0 by the property that any point x in 3-space that is imaged at point u in image J_0 will be imaged at point Fu in image $J_1 = F(J_0)$.

Proposition 34. *The essential matrix for the pair of resampled images J_1 and J'_1 obtained by resampling according to the maps F and F' defined in the previous proposition is given by*

$$Q_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

The corresponding image-to-image translation M_1 is the identity matrix.

Proof. Let $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$ be a set of image-to-image correspondences between the images J_0 and J'_0 sufficient in number to determine the matrix Q_0 uniquely. Thus, $\mathbf{u}'_i Q_0 \mathbf{u}_i = 0$. However, in the images J_1 and J'_1, the points $\{F\mathbf{u}_i\}$ correspond with points $\{F'\mathbf{u}'_i\}$. It follows that

$$\mathbf{u}_i'^{\top} F'^{\top} Q_1 F \mathbf{u}_i = 0$$

and hence $Q_0 \approx F'^{\top} Q_1 F$ or

$$Q_1 = F'^{*} Q_0 F^{-1}.$$

However, since $F' = FM^{-1}$, we may write $F'^{*} = F^{*} M^{\top}$, and substituting in (**??**) gives

$$\begin{aligned} Q_1 &= F^{*} M^{\top} Q_0 F^{-1} \\ &= F^{*} S(p) F^{-1} \mathrm{b}y(\mathbf{??}) \\ &= F^{*} F^{\top} S(Fp) \mathrm{b}y(\mathbf{??}) \\ &= S((1,0,0)^{\top}) \end{aligned}$$

which is the required matrix.

To prove the second statement, let x be a point in space on the plane determined by the image-to-image projective transform M in the sense that if $\mathbf{u}$ and $\mathbf{u}'$ are the images of point x in the images J_0 and J'_0, then $\mathbf{u}' = M\mathbf{u}$ (see Proposition **??**). Then point $\mathbf{x}$ will be seen in the images J_1 and J'_1 at points $F\mathbf{u}$ and $F'\mathbf{u}'$. However, $F'\mathbf{u}' = F'M\mathbf{u} = F\mathbf{u}$. So, $F\mathbf{u}$ in image J_1 is mapped to $F'\mathbf{u}'$ in image J'_1 by the identity transformation. □

9.5 Determination of the resampling transformation.

The transformation F was described by the condition that it takes the epipole $\mathbf{p}$ to the point at infinity on the x axis. This leaves many degrees of freedom open for F, and if an inappropriate F is chosen, severe projective distortion of the image can take place. In order that the resampled image should look somewhat like one of the original images, we may put closer restrictions on the choice of F.

One condition that leads to quite good results is to insist that the transformation F should act as far as possible as a rigid transformation in the neighbourhood of a given selected point $\mathbf{u}_0$ of the first image. By this is meant that the neighbourhood of $\mathbf{u}_0$ may undergo rotation and translation only, and hence will look the same in the original and resampled image. An appropriate choice of point u_0 may be the centre of the image. For instance, this would be a good choice in an context of aerial photography if the first image is known not to be excessively oblique.

A projective transformation may be determined by specifying the destination of four points. Suppose that the epipole is already on the x axis at location $(1, 0, f)^\top$ and that be desire the projective transformation to approximate the identity map in the local neighbourhood of the origin $(0, 0, 1)^\top$. The desired map may be found by specifying the destinations of four points

$$
\begin{aligned}
(1, 0, f)^\top &\rightarrow (1, 0, 0)^\top \\
(0, 0, 1)^\top &\rightarrow (0, 0, 1)^\top \\
(\delta, \delta, 1)^\top &\rightarrow (\delta, \delta, 1)^\top \\
(\delta, -\delta, 1)^\top &\rightarrow (\delta, -\delta, 1)^\top
\end{aligned}
$$

and then letting $\delta \rightarrow 0$. The correct map is found to be expressed by the matrix

$$
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -f & 0 & 1 \end{pmatrix}
$$

It may be seen that if $af << 1$ then the point $(a, b, 1)^\top$ is mapped (almost) to itself by this transform.

Experimental results

The method of ...

References

[1] M.J. Hanna, *"Bootstrap stereo,"* Proc. Image Understanding Workshop, College Park, MD, April 1980, pp. 201–208.

[2] M.J. Hanna, *"A description of SRI's baseline stereo system,"* ARI International Artificial Intelligence Center Tech. Note 365, Oct. 1985. Workshop, College Park, MD, April 1980, pp. 201–208.

[3] H.C. Longuet-Higgins, *"A computer algorithm for reconstructing a scene from two projections,"* Nature, Vol. 293, 10 Sept. 1981.

[4] R. Hartley, *"Estimation of Relative Camera Positions for Uncalibrated Cameras,"*, Technical Report, GE Corporate R&D, 1 River Road, Schenectady, NY 12301, Oct., 1991.

[5] K.E. Atkinson, "An Introduction to Numerical Analysis," John Wiley and Sons, 2nd Edition, 1989.

[6] I.E. Sutherland, *"Three dimensional data input by tablet,"* Proceedings of IEEE, Vol. 62, No. 4, pp. 453–461, April 1974.

[7] L. H. Quam, *"Hierarchical Warp Stereo,"* Proc. DARPA Image Understanding Workshop, New Orleans, LA, pp. 149–155, 1984 (also in "Readings in Computer Vision, " M.A. Fischler and O. Firschein, Morgan Kaufmann Publishers, Inc., 1987).

[8] Strat, T. M., "Recovering the Camera Parameters from a Transformation Matrix," Readings in Computer Vision, pp 93-100, Morgan Kauffamann Publishers, Inc, Los Altos, Ca., 1987.

[9] T.M. Strat, *" Recovering the camera parameters from a transformation matrix,"* Proc. of DARPA Image Understanding Workshop, New Orleans, LA, pp. 264–271, 1984.

[10] Faugeras, O. and Maybank, S., "Motion from Point Matches : Multiplicity of Solutions," International Journal of Computer Vision, **4**, (1990), 225–246.

[11] Longuet-Higgins, H. C., "A computer algorithm for reconstructing a scene from two projections," Nature, Vol. 293, **10,** Sept. 1981.

[12] Porrill, J and Pollard, S., "Curve Fitting and Stereo Calibration," Proceedings of the British Machine Vision Conference, University of Oxford, Sept 1990, pp 37–42.

[13] Wolfram, S., Mathematica, "A System for Doing Mathematics by Computer," Addison-Wesley, Redwood City, California, 1988.