Recovering Unknown Focal Lengths in Self-Calibration: An Essentially Linear Algorithm and Degenerate Configurations

G.N. Newsam D.Q. Huynh M.J. Brooks H.-P. Pan
{gnewsam,du,mjb,heping}@cssip.edu.au

Centre for Sensor Signal and Information Processing
Signal Processing Research Institute
Technology Park, The Levels, SA 5095, Australia

KEY WORDS: Stereo vision, degeneracy, epipolar equation, fundamental matrix, self-calibration, relative orientation.

ABSTRACT

If sufficiently many pairs of corresponding points in a stereo image pair are available to construct the associated fundamental matrix, then it has been shown that 5 relative orientation parameters and 2 focal lengths can be recovered from this fundamental matrix. This paper presents a new and essentially linear algorithm for recovering focal lengths. Moreover the derivation of the algorithm also provides a complete characterisation of all degenerate configurations in which focal lengths cannot be uniquely recovered. There are two classes of degenerate configurations: either one of the optical axes of the cameras lies in the plane spanned by the baseline and the other optical axis; or one optical axis lies in the plane spanned by the baseline and the vector that is orthogonal to both the baseline and the other axis. The result that the first class of configurations (i.e. ones in which the optical axes are coplanar) is degenerate is of some practical importance since it shows that self-calibration of unknown focal lengths is not possible in certain stereo heads, a configuration widely used for binocular vision systems in robotics.

1 INTRODUCTION

Relative orientation is the problem of recovering the parameters defining the rotation and translation direction relating two calibrated camera views from a set of corresponding image points. This has long been studied by photogrammetrists [4, 16], and more recently by the computer vision community [3, 7, 9, 17]. 5 parameters suffice to define the relative orientation of the two cameras, 3 describing the rotation and 2 the direction of translation. An essentially linear algorithm for their recovery was proposed by Stefanovic [15] and revived by Longuet-Higgins [9], based on the coplanarity constraint between corresponding points. This involves computing the 3×3 *essential matrix* [3] associated with the stereo pair, either by solving a linear system of equations derived from eight or more pairs of corresponding points [9], or by finding the vector associated with the smallest singular value of the associated system [3, 17]. Once the essential matrix is available, it can be decomposed into a product of a rotation matrix and a skew symmetric matrix derived from the baseline (see [3] for a comprehensive discussion of algorithms for this).

Given these solutions, the question of how to deal with uncalibrated cameras has come to the fore in the last few years. Here, the problem is to simultaneously calibrate the cameras and recover the viewing geometry from a given set of images, i.e. to recover both the *intrinsic* (interior orientation) and *extrinsic* (relative orientation) parameters. We shall refer to this process as *self-calibration*. Faugeras et al. [2] developed a model for the interior orientation of a general uncalibrated pinhole camera in terms of 5 intrinsic parameters. Moreover a generalisation of the essential matrix, the *fundamental matrix* [10, 11], can be defined and derived from the data that is a function of both the intrinsic and extrinsic parameters. Even if the same camera is used to take both images in a stereo pair, however, these cannot be recovered simultaneously along with the 5 extrinsic parameters from the fundamental matrix as this matrix is determined by only 7 independent parameters in total. If three or more images of the same scene taken by the same camera are available though, then the 10 orientation parameters can be recovered from the set of fundamental matrices derived from all possible pairs of images [2]. More recently Niini [12] has presented a promising new approach for recovering these parameters based upon setting up linear systems for certain intermediate quantities.

An interesting special case in which self-calibration is possible from a single pair of images attains when the cameras taking the images are calibrated up to unknown focal lengths. In this case there are only 2 unknown intrinsic parameters (the unknown focal length in each image) in addition to the 5 extrinsic parameters. Since the fundamental matrix has 7 independent parameters there would seem to be sufficient information to carry out self-calibration. Hartley [6] has shown that this is indeed possible through an approach based on the singular value decomposition (SVD) of the fundamental matrix. Pan et al. [13] presented an alternative closed-form expression for the focal lengths as roots of certain cubics, along with an iterative least-squares technique for refining parameter estimates [14]. Recently, however, Huynh et al. [8] have shown that both algorithms fail for certain degenerate camera configurations: in particular, when the optical axes of the cameras are coplanar the problem becomes degenerate and unique focal lengths for each camera cannot be determined. This has some practical importance since in the stereo heads used in robotic vision it is common for camera motion to be restricted to vergence only, so that the optical axes and the baseline between them are confined to the horizontal plane.

The purpose of this paper is to present a new and simpler algorithm for self-calibration of focal lengths through the solution of a linear system of equations. The approach is similar in spirit to that of Niini [12] and rests very much on a linear algebraic formulation of the problem, but it is based on a different representation than that used in [12]. It follows Hartley in using the SVD, but is considerably simpler in that it does not require solution of any further higher order systems. Moreover a by-product of the approach is a complete characterisation of degenerate configurations in which self-calibration is not possible: these are identified as those configurations which make the linear system singular. In particular we shall show that a necessary and sufficient condition for the system to be singular is for the configuration to belong to one of two classes: configurations in which the optical axes and baseline are coplanar; and configurations in which there is coplanarity of one optical axis, the baseline and the vector perpendicular to both the baseline and the other optical axis.

The paper is organised as follows. The next section defines the notation, sets out the geometry of the problem and reviews some basic concepts. The third section shows that self-calibration can be reduced to solving a linear system from whose solution the focal lengths can be easily calculated. The fourth section derives conditions under which the linear system is singular (so no unique solution exists), and shows that these correspond to the geometric configurations described in the previous paragraph. Finally the last section presents a solution for the special case when both focal lengths are known a priori to be equal: here, the unknown focal length can be read off from the roots of a quadratic.

2 PROBLEM FORMULATION

The following notation will be used in the paper. World points are written in upper case, image points in lower case, vectors in bold lower case, and matrices in bold upper case. $\mathbf{a}_j$ denotes the j-th column of the corresponding matrix $\mathbf{A}$, while A_{ij} denotes the ij-th element. In particular, $\mathbf{I}$ denotes the 3×3 identity matrix and $\mathbf{i}_j$ denotes the j-th column of $\mathbf{I}$, i.e. the unit vector with a one in the j-th position and zeroes elsewhere. (Note that by definition $\mathbf{a}_j = \mathbf{A}\mathbf{i}_j$.) The symbol ^{T} denotes matrix and vector transpose, while $^{-T}$ denotes the transpose of the inverse matrix. The notation $\mathbf{u} \sim \mathbf{v}$ ($\mathbf{U} \sim \mathbf{V}$) indicates that the vectors $\mathbf{u}$ and $\mathbf{v}$ (matrices $\mathbf{U}$ and $\mathbf{V}$) are the same up to an arbitrary scale factor. Finally entities related to the right image are marked with a $'$.

We now define the geometry of the model problem. Figure 1 shows a scene being stereoscopically imaged. The global reference system $O' - X'Y'Z'$ is taken to coincide with the coordinate system of the right image, with the origin at the optical centre. $\mathbf{R} = \mathbf{R}(\alpha, \beta, \gamma)$ denotes the matrix associated with a rotation of angles α, β, γ about the X'-, Y'-, and Z'-axes, respectively, that renders the left image parallel to the right image. $\mathbf{b} = (b_x, b_y, b_z)^T$ denotes the baseline vector connecting the optical centres of the cameras. The parameters α, β, γ, b_x, b_y, b_z are said to specify the relative orientation of a stereo pair of images. Since the scene can only reconstructed to within an overall scale factor, it is usual to remove this ambiguity by assuming that $\|\mathbf{b}\| = 1$, leaving 5 independent parameters to be determined in relative orientation.

We next assume that the location of the principal point (the intersection of the optical axis with the image plane) is known in each image, and that the image coordinate system is Euclidean (i.e. no skewness or differing scales on different axes). In this case a simple translation of all image points will ensure that the principal points coincide with the origins in the image plane coordinates. Thus the only unknown intrinsic parameter associated with the formation of each image is the focal length.

Now let M be a visible point in the scene, and $m = (x, y)^T$ and $m' = (x', y')^T$ be its projections onto the left and right image planes. Relative to the global reference system $O' - X'Y'Z'$ at C' and the coordinate system $O - XYZ$ at C, the points m' and m can be expressed in vector form as

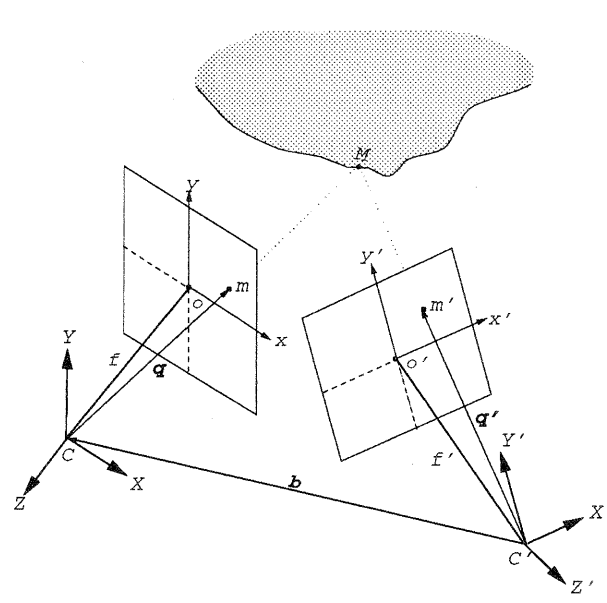

Figure 1: Stereo camera setup.

$\mathbf{q}' = (x', y', -f')^T$ and $\mathbf{q} = (x, y, -f)^T$. It is clear from Figure 1 that the three vectors $\mathbf{q}$, $\mathbf{q}'$, and $\mathbf{b}$ are coplanar. This relationship is encapsulated in the *epipolar equation* which relates corresponding image points by

$$\mathbf{q}' \cdot (\mathbf{b} \times \mathbf{R}\mathbf{q}) = 0. \tag{1}$$

Let $f > 0$ and $f' > 0$ be the two focal lengths, $\mathbf{A}$ and $\mathbf{A}'$ be the intrinsic parameter matrices of the two cameras, and $\mathbf{B}$ be the skew-symmetric matrix containing elements of $\mathbf{b}$. Then under the above assumptions these matrices have the form

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/f \end{pmatrix}, \quad \mathbf{A}' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/f' \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{pmatrix}, \tag{2}$$

and (1) may be expressed in matrix form as

$$\mathbf{q}'^T \mathbf{E} \mathbf{q} = 0, \tag{3}$$

where $\mathbf{E}$ is the *essential matrix* given by

$$\mathbf{E} = \mathbf{B}\mathbf{R}. \tag{4}$$

Now let $\mathbf{m}' = (x', y', 1)^T$ and $\mathbf{m} = (x, y, 1)^T$ be alternative representations of the image points in homogeneous coordinates. Observe that

$$\mathbf{q} = \mathbf{A}^{-1}\mathbf{m} \tag{5}$$

$$\mathbf{q}' = \mathbf{A}'^{-1}\mathbf{m}'. \tag{6}$$

We may now immediately infer that

$$\mathbf{m}'^T \mathbf{F} \mathbf{m} = 0, \tag{7}$$

where $\mathbf{F}$ is the *fundamental matrix* [2] embodying both extrinsic and intrinsic imaging parameters, and is given by

$$\mathbf{F} = \mathbf{A}'^{-T}\mathbf{B}\mathbf{R}\mathbf{A}^{-1}. \tag{8}$$

If sufficiently many corresponding points can be located in the two images, it is possible to obtain a numerical estimate, $\mathbf{F}_{est}$, of the matrix $\mathbf{F}$ [11]. Note that $\mathbf{F}_{est}$ is only defined up to an arbitrary scale factor (we shall show in the next section that a sensible choice for this is to scale any estimate so that $\|\mathbf{f}_3\| = 1$ as this is likely to optimise the conditioning of the linear system defining the solution). In what follows we shall not distinguish between $\mathbf{F}$ and $\mathbf{F}_{est}$ (i.e. we assume a sufficiently accurate estimate of $\mathbf{F}$ has been obtained), so that (8) is replaced by

$$\mathbf{F} \sim \mathbf{A}'^{-T}\mathbf{BRA}^{-1}. \tag{9}$$

Thus self-calibration reduces to the problem of estimating the various extrinsic and intrinsic parameters from $\mathbf{F}$.

3 DERIVING A LINEAR SYSTEM FOR FOCAL LENGTHS

In this section we shall construct a linear system whose coefficients are numerical values derived from the fundamental matrix, and from whose solution we can easily read off the values of the focal lengths. We begin the construction by recalling the following characterisations of an essential matrix given in [3].

Proposition 1 *The following three conditions are equivalent:*

(i) $\mathbf{E}$ *is an essential matrix;*

(ii) $\mathbf{E}$ *has one zero singular value and two non-zero and equal singular values;*

(iii) $\mathbf{EE}^T = \mathbf{I} - \mathbf{bb}^T$.

We next introduce the singular value decomposition (SVD) of $\mathbf{F}$ (see [5] for details). That is, $\mathbf{F}$ can be written as

$$\mathbf{F} \;=\; \mathbf{U\Sigma V}^T \;=\; \sum_{k=1}^{3} \sigma_k \mathbf{u}_k \mathbf{v}_k^T. \tag{10}$$

Here $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices and $\mathbf{\Sigma}$ is a diagonal matrix with entries $\Sigma_{kk} = \sigma_k$. Since $\mathbf{A}$ and $\mathbf{A}'$ are non-singular, it follows from condition *(ii)* of Proposition 1 that $\sigma_1, \sigma_2 \neq 0$ and that $\sigma_3 = 0$, so that

$$\mathbf{Fv}_3 \;=\; \mathbf{F}^T\mathbf{u}_3 \;=\; \mathbf{0}. \tag{11}$$

It also follows from the proposition that $\mathbf{b}$ is the unique vector for which $\mathbf{E}^T\mathbf{b} = \mathbf{0}$. Therefore from (8) and (11) we have that

$$\mathbf{F}^T\mathbf{u}_3 \;=\; \mathbf{A}^{-T}\mathbf{E}^T\mathbf{A}'^{-1}\mathbf{u}_3 \;=\; \mathbf{0}, \tag{12}$$

so it follows that

$$\mathbf{b} \;=\; \frac{\mathbf{A}'^{-1}\mathbf{u}_3}{\|\mathbf{A}'^{-1}\mathbf{u}_3\|}. \tag{13}$$

We may now combine (13) with condition *(iii)* of Proposition 1 to derive the following key equation:

$$\begin{aligned}\mathbf{A}'^T\mathbf{FAA}^T\mathbf{F}^T\mathbf{A}' &\sim \mathbf{EE}^T \\ &= \mathbf{I} - \frac{\mathbf{A}'^{-1}\mathbf{u}_3\mathbf{u}_3^T\mathbf{A}'^{-T}}{\mathbf{u}_3^T\mathbf{A}'^{-T}\mathbf{A}'^{-1}\mathbf{u}_3}.\end{aligned} \tag{14}$$

Multiplying on the left by $\mathbf{A}'^{-T}$ and on the right by $\mathbf{A}'^{-1}$ and then multiplying through by the denominator gives

$$\begin{aligned}\mathbf{FAA}^T\mathbf{F}^T \sim\; &(\mathbf{u}_3^T\mathbf{A}'^{-T}\mathbf{A}'^{-1}\mathbf{u}_3)\,\mathbf{A}'^{-T}\mathbf{A}'^{-1} \\ &-\mathbf{A}'^{-T}\mathbf{A}'^{-1}\mathbf{u}_3\mathbf{u}_3^T\mathbf{A}'^{-T}\mathbf{A}'^{-1}.\end{aligned} \tag{15}$$

We now reduce (15) to a linear system involving quantities that are directly related to the unknown focal lengths. To achieve this, we note from (2) that the matrices $\mathbf{AA}^T$ and $\mathbf{A}'^{-T}\mathbf{A}'^{-1}$ can be written as

$$\mathbf{AA}^T = \mathbf{I} + (f^{-2} - 1)\mathbf{i}_3\mathbf{i}_3^T \tag{16}$$

$$\mathbf{A}'^{-T}\mathbf{A}'^{-1} = \mathbf{I} + (f'^2 - 1)\mathbf{i}_3\mathbf{i}_3^T, \tag{17}$$

where, as noted above, $\mathbf{i}_3$ is the unit vector $(0,0,1)^T$. If we define the quantities μ and ν by

$$\mu = f^{-2} - 1 \tag{18}$$

$$\nu = f'^2 - 1, \tag{19}$$

then (15) can be expressed in terms of μ and ν as:

$$\begin{aligned}\mathbf{F}(\mathbf{I} + \mu\mathbf{i}_3\mathbf{i}_3^T)\mathbf{F}^T \sim\; &(1 + \nu(\mathbf{u}_3^T\mathbf{i}_3)^2)(\mathbf{I} + \nu\mathbf{i}_3\mathbf{i}_3^T) - \\ &(\mathbf{u}_3 + \nu(\mathbf{u}_3^T\mathbf{i}_3)\mathbf{i}_3)(\mathbf{u}_3 + \nu(\mathbf{u}_3^T\mathbf{i}_3)\mathbf{i}_3)^T.\end{aligned} \tag{20}$$

On the face of it this equation is quadratic in ν, but a little algebra shows that in fact the quadratic terms cancel. Indeed, if we introduce the arbitrary and as yet unknown scale factor λ, the system reduces to:

$$\begin{aligned}\mathbf{FF}^T + \mu\mathbf{f}_3\mathbf{f}_3^T = \;&\lambda(\mathbf{I} - \mathbf{u}_3\mathbf{u}_3^T) + \\ &\lambda\nu\left[(\mathbf{u}_3^T\mathbf{i}_3)^2\mathbf{I} - (\mathbf{u}_3^T\mathbf{i}_3)(\mathbf{u}_3\mathbf{i}_3^T + \mathbf{i}_3\mathbf{u}_3^T) + \mathbf{i}_3\mathbf{i}_3^T\right].\end{aligned} \tag{21}$$

Gathering up terms and defining the new variables

$$\omega_1 = -\mu, \quad \omega_2 = \lambda\nu, \quad \omega_3 = \lambda \tag{22}$$

now gives:

$$\begin{aligned}\mathbf{FF}^T = \;&\mathbf{f}_3\mathbf{f}_3^T\omega_1 + \\ &\left[(\mathbf{u}_3^T\mathbf{i}_3)^2\mathbf{I} - (\mathbf{u}_3^T\mathbf{i}_3)(\mathbf{u}_3\mathbf{i}_3^T + \mathbf{i}_3\mathbf{u}_3^T) + \mathbf{i}_3\mathbf{i}_3^T\right]\omega_2 + \\ &(\mathbf{I} - \mathbf{u}_3\mathbf{u}_3^T)\,\omega_3\end{aligned} \tag{23}$$

Formally this system can be written as

$$\mathbf{G}_1\omega_1 + \mathbf{G}_2\omega_2 + \mathbf{G}_3\omega_3 = \mathbf{G}_4, \tag{24}$$

where the $\mathbf{G}_k$ are symmetric 3×3 matrices constructed from the data. This gives six linear equations in three unknowns. The unknowns could be found by a least squares solution, but more insight comes from making a change of coordinates from the standard basis defined by the vectors $\mathbf{i}_j$ to the orthonormal basis defined by the $\mathbf{u}_j$. This transforms the matrices $\mathbf{G}_k$ to the matrices $\tilde{\mathbf{G}}_k$ with entries

$$[\tilde{\mathbf{G}}_k]_{ij} = \mathbf{u}_i^T\mathbf{G}_k\mathbf{u}_j. \tag{25}$$

It is easy to verify that $[\tilde{\mathbf{G}}_k]_{i3} = [\tilde{\mathbf{G}}_k]_{3j} = 0$ for all i, j and all k: this stems from the fact that $\mathbf{u}_3$ corresponds to the zero singular value. Thus in the new coordinate system five of the equations are identically zero, leaving the following three equations to be solved for ω_1, ω_2 and ω_3.

$$\begin{aligned}\sigma_1^2 &= (\mathbf{u}_1^T\mathbf{f}_3)^2\,\omega_1 + ((\mathbf{u}_1^T\mathbf{i}_3)^2 + (\mathbf{u}_3^T\mathbf{i}_3)^2)\,\omega_2 + \omega_3 \\ 0 &= (\mathbf{u}_2^T\mathbf{f}_3)(\mathbf{u}_1^T\mathbf{f}_3)\,\omega_1 + (\mathbf{u}_1^T\mathbf{i}_3)(\mathbf{u}_2^T\mathbf{i}_3)\,\omega_2 \\ \sigma_2^2 &= (\mathbf{u}_2^T\mathbf{f}_3)^2\omega_1 + ((\mathbf{u}_2^T\mathbf{i}_3)^2 + (\mathbf{u}_3^T\mathbf{i}_3)^2)\,\omega_2 + \omega_3\end{aligned} \tag{26}$$

These equations are formed by computing $\mathbf{u}_i^T\mathbf{G}_k\mathbf{u}_j$ for $i, j = 1, 2$ (and noting that symmetry implies $[\tilde{\mathbf{G}}_k]_{12} = [\tilde{\mathbf{G}}_k]_{21}$).

The coefficients in (26) can all be calculated directly from the data once the SVD of $\mathbf{F}$ has been computed. Moreover once (26) has been solved, the focal lengths f and f' can be read off directly from the values of ω_1, ω_2 and ω_3 through (18) and (19). Finally there is a unique correspondence between focal lengths and coefficients once the arbitrary scale factor λ has been eliminated.

To conclude this section, note that if $\mathbf{F}$ is rescaled to $\kappa\mathbf{F}$, then the coefficients in the vector on the left hand side and in the first column on the right hand side scale as κ^2, while the coefficients in the remaining two columns are invariant. Hence any numerical estimate of $\mathbf{F}$ should be rescaled to ensure that (26) is a well-conditioned system. A simple sensible choice for this is to attempt to ensure that elements in the first column have roughly the same magnitude as elements in the other columns (i.e. are of order unity) by scaling $\mathbf{F}$ so that $||\mathbf{f}_3|| = 1$ as proposed in the discussion leading up to (9).

4 CHARACTERISING DEGENERATE CONFIGURATIONS

We now explore degenerate configurations for which multiple factorisations of the form in (8) may exist with $\mathbf{A}$ and $\mathbf{A}'$ having the form in (2). Assuming the underlying camera model is correct, by Proposition 1 such a factorisation is possible if and only if $\mathbf{A}, \mathbf{A}'$ and $\mathbf{F}$ satisfy (14). As each step in the remainder of the derivation is reversible, it follows that, given the special form of $\mathbf{A}$ and $\mathbf{A}'$, a unique factorisation exists for (8) if and only if (26) has a unique solution.

For convenience, let us express (26) formally as $\mathbf{s} = \mathbf{C}\mathbf{w}$, where $\mathbf{s} = \left(\sigma_1^2, 0, \sigma_2^2\right)^T$ and $\mathbf{w} = (\omega_1, \omega_2, \omega_3)^T$. Then it is a fundamental result of linear algebra that (26) has a unique solution if and only if $\mathbf{C}$ is invertible. If $\mathbf{C}$ is not invertible, then (26) will have either multiple solutions or no solutions depending on whether or not $\mathbf{s}$ is in the range of $\mathbf{C}$. The case where (26) has no solutions is of little interest. Since we are assuming we have data from a real world system there must be at least one solution: assuming that the underlying camera model is essentially correct, then the fact that $\mathbf{s}$ is outside the range of $\mathbf{C}$ can be put down to measurement errors and appropriate allowances made (e.g. (26) could be solved in a total least squares sense as described in [5]).

The condition that $\mathbf{C}$ is not invertible is equivalent to $\det(\mathbf{C}) = 0$. From (26) and a little algebra we have that

$$\begin{aligned}\det(\mathbf{C}) = &\left[(\mathbf{u}_1^T\mathbf{f}_3)(\mathbf{u}_1^T\mathbf{i}_3) + (\mathbf{u}_2^T\mathbf{f}_3)(\mathbf{u}_2^T\mathbf{i}_3)\right] \\ &\left[(\mathbf{u}_1^T\mathbf{f}_3)(\mathbf{u}_2^T\mathbf{i}_3) - (\mathbf{u}_2^T\mathbf{f}_3)(\mathbf{u}_1^T\mathbf{i}_3)\right]. \end{aligned} \tag{27}$$

Thus $\det(\mathbf{C})$ vanishes if and only if either of the factors on the right hand side of (27) is identically zero. The next proposition now interprets these conditions in terms of the geometry of the imaging system.

Proposition 2 *Two focal lengths cannot be uniquely identified from the fundamental matrix if and only if the geometry of the imaging system is in either of the following classes of configurations:*

(i) The optical axes of the two cameras and the baseline between them are coplanar.

(ii) One optical axis, the baseline and the vector perpendicular to the baseline and the other optical axis are coplanar.

Proof. We have established that a unique solution is not possible if and only if one of the factors on the right hand side of (27) is zero. It remains to give the geometric interpretation of each of the conditions

$$\left[(\mathbf{u}_1^T\mathbf{f}_3)(\mathbf{u}_1^T\mathbf{i}_3) + (\mathbf{u}_2^T\mathbf{f}_3)(\mathbf{u}_2^T\mathbf{i}_3)\right] = 0 \tag{28}$$

$$\left[(\mathbf{u}_1^T\mathbf{f}_3)(\mathbf{u}_2^T\mathbf{i}_3) - (\mathbf{u}_2^T\mathbf{f}_3)(\mathbf{u}_1^T\mathbf{i}_3)\right] = 0. \tag{29}$$

We first show that (28) is equivalent to the condition that $F_{33} = 0$. To see this, note that since the $\mathbf{u}_k$ form an orthonormal basis

$$\mathbf{f}_3 = \sum_{k=1}^{3}(\mathbf{u}_k^T\mathbf{f}_3)\mathbf{u}_k. \tag{30}$$

Since $\mathbf{F}^T\mathbf{u}_3 = 0$, it follows that $\mathbf{f}_3^T\mathbf{u}_3 = 0$ and that

$$(\mathbf{u}_1^T\mathbf{f}_3)(\mathbf{u}_1^T\mathbf{i}_3) + (\mathbf{u}_2^T\mathbf{f}_3)(\mathbf{u}_2^T\mathbf{i}_3) = \mathbf{f}_3^T\mathbf{i}_3 = F_{33}. \tag{31}$$

Thus (28) is equivalent to $F_{33} = 0$. Since $\mathbf{A}^{-1}\mathbf{i}_3 \sim \mathbf{i}_3$ and $\mathbf{A}'^{-1}\mathbf{i}_3 \sim \mathbf{i}_3$, we have that

$$\begin{aligned} 0 = F_{33} = \mathbf{i}_3^T\mathbf{F}\mathbf{i}_3 &= \mathbf{i}_3^T\mathbf{A}'^{-T}\mathbf{B}\mathbf{R}\mathbf{A}^{-1}\mathbf{i}_3 \\ &\sim \mathbf{i}_3^T\mathbf{B}\mathbf{R}\mathbf{i}_3 \\ &= \mathbf{i}_3 \cdot (\mathbf{b} \times \mathbf{R}\mathbf{i}_3). \end{aligned} \tag{32}$$

But $\mathbf{i}_3$ is also the unit vector in the direction of the optical axis in each image expressed in the local coordinate system of each image. Thus (32) is equivalent to $\mathbf{z}' \cdot (\mathbf{b} \times \mathbf{z}) = 0$, where $\mathbf{z}$ and $\mathbf{z}'$ are unit vectors in the direction of the optical axes expressed in the global coordinate system. This in turn implies that the vectors $\mathbf{z}, \mathbf{z}'$ and $\mathbf{b}$ are coplanar.

We next examine (29). If we define the vector $\mathbf{g}$ by

$$\mathbf{g} = (\mathbf{u}_1^T\mathbf{f}_3)\mathbf{u}_2 - (\mathbf{u}_2^T\mathbf{f}_3)\mathbf{u}_1, \tag{33}$$

then it is straightforward to check that $\mathbf{g}$ is orthogonal to $\mathbf{f}_3$. Moreover since the $\mathbf{u}_k$ form an orthonormal basis, $\mathbf{g}$ is orthogonal to $\mathbf{u}_3$. Finally (29) effectively states that $\mathbf{g}$ is orthogonal to $\mathbf{i}_3$. Since $\mathbf{g}$ is non-zero as long as $\mathbf{f}_3$ is non-zero, this implies that the vectors $\mathbf{f}_3, \mathbf{i}_3$ and $\mathbf{u}_3$ must be linearly dependent (i.e. they are coplanar).

We now show that this coplanarity condition is equivalent to condition *(ii)* of the proposition. To see this, it suffices to show that $\mathbf{b}$ and $\mathbf{B}\mathbf{R}\mathbf{i}_3$ are in the subspace spanned by $\mathbf{f}_3, \mathbf{i}_3$ and $\mathbf{u}_3$. To establish this first for $\mathbf{b}$, note that (2) and (13) imply

$$\begin{aligned} \mathbf{b} &\sim \mathbf{A}'^{-1}\mathbf{u}_3 = (\mathbf{I} - (f'+1)\mathbf{i}_3\mathbf{i}_3^T)\mathbf{u}_3 \\ \Rightarrow \qquad \mathbf{b} &= \lambda_1\mathbf{u}_3 + \lambda_2\mathbf{i}_3, \end{aligned} \tag{34}$$

for some constants λ_1 and λ_2. To establish a similar result for $\mathbf{B}\mathbf{R}\mathbf{i}_3$, we first note that the particular form of $\mathbf{A}$ implies that $\mathbf{A}^{-1}\mathbf{i}_3 \sim \mathbf{i}_3$. Thus

$$\begin{aligned} \mathbf{B}\mathbf{R}\mathbf{i}_3 &\sim \mathbf{B}\mathbf{R}\mathbf{A}^{-1}\mathbf{i}_3 \\ &= \mathbf{A}'^T\mathbf{A}'^{-T}\mathbf{B}\mathbf{R}\mathbf{A}^{-1}\mathbf{i}_3 \\ &= \left(\mathbf{I} - (f'^{-1}+1)\mathbf{i}_3\mathbf{i}_3^T\right)\mathbf{A}'^{-T}\mathbf{B}\mathbf{R}\mathbf{A}^{-1}\mathbf{i}_3 \\ &= \mathbf{A}'^{-T}\mathbf{B}\mathbf{R}\mathbf{A}^{-1}\mathbf{i}_3 - (f'^{-1}+1)(\mathbf{i}_3^T\mathbf{A}'^{-T}\mathbf{B}\mathbf{R}\mathbf{A}^{-1}\mathbf{i}_3)\,\mathbf{i}_3 \\ &= \lambda_3\mathbf{f}_3 + \lambda_4\mathbf{i}_3, \end{aligned} \tag{35}$$

for some constants λ_3 and λ_4.

Since $\mathbf{b}$ and $\mathbf{BRi}_3$ can be expressed as linear combinations of $\mathbf{f}_3, \mathbf{i}_3$ and $\mathbf{u}_3$ and these three vectors are linearly dependent, it follows that $\mathbf{b}, \mathbf{BRi}_3$ and $\mathbf{i}_3$ are also linearly dependent. But by construction $\mathbf{BRi}_3 = \mathbf{b} \times \mathbf{z}$ where $\mathbf{z} = \mathbf{Ri}_3$ is the optical axis for the first image. Thus $\mathbf{BRi}_3$ is the vector perpendicular to both the baseline and the optical axis of the first image. Since $\mathbf{i}_3$ is the direction of the optical axis in the second image, we have that the optical axis in the second image, the baseline and the vector perpendicular to the optical axis in the first image and to the baseline are all coplanar. This completes the proof. □

In conclusion, we have shown that there are two classes of degenerate imaging configurations in which two focal lengths cannot be uniquely recovered from the fundamental matrix, but that in all other cases unique recovery is possible. As noted in the introduction, the fact that the first class (in which the optical axes and the baseline are coplanar) is degenerate is of practical significance, since many existing artificial vision systems are restricted to such configurations. The form of the second class has a pleasing symmetry with respect to the form of the first, but it is of little practical importance: imaging systems are usually constructed so that the optical axes tend to be roughly parallel with each other and roughly orthogonal to the baseline. Note, however, that it is still possible for there to be significant overlap of the camera fields of view for configurations in this class.

5 DERIVING A SOLUTION FOR A SINGLE FOCAL LENGTH

We now turn to the special case where the two focal lengths are known a priori to be equal: this will occur if a single camera with an unknown focal length is used to take both images. For general configurations the problem can obviously be solved by the algorithm presented above: we are interested in deriving a solution that will still recover the focal length in some of the degenerate configurations identified under Proposition 2.

To achieve this we return to (14). In the case where the focal lengths are equal a priori, we have $\mathbf{A} = \mathbf{A}' = \mathbf{A}^T$, so multiplying each side of the equation on the left by $\mathbf{A}^{-1}$ and on the right by $\mathbf{A}$ gives

$$\mathbf{FA}^2\mathbf{F}^T\mathbf{A}^2 = \lambda \left[\mathbf{I} - \frac{\mathbf{A}^{-2}\mathbf{u}_3\mathbf{u}_3^T}{\mathbf{u}_3^T\mathbf{A}^{-2}\mathbf{u}_3} \right], \qquad (36)$$

where λ is an arbitrary scale factor. We now show how to derive a quadratic from this system that will uniquely identify the focal length f in almost all configurations. To do this we note that forming the inner products of the matrices in (36) w.r.t the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ gives:

$$\mathbf{u}_1^T\mathbf{FA}^2\mathbf{F}^T\mathbf{A}^2\mathbf{u}_1 \quad = \quad \lambda \quad = \quad \mathbf{u}_2^T\mathbf{FA}^2\mathbf{F}^T\mathbf{A}^2\mathbf{u}_2. \qquad (37)$$

Recalling (16), setting $\mu = (f^{-2} - 1)$ and noting that $\mathbf{u}_k^T\mathbf{F} = \sigma_k\mathbf{v}_k^T$ shows that (37) can be written as

$$\begin{aligned}\sigma_1\mathbf{v}_1^T(\mathbf{I} + \mu\,\mathbf{i}_3\mathbf{i}_3^T)\,\mathbf{F}^T(\mathbf{I} + \mu\,\mathbf{i}_3\mathbf{i}_3^T)\mathbf{u}_1 = \\ \sigma_2\mathbf{v}_2^T(\mathbf{I} + \mu\,\mathbf{i}_3\mathbf{i}_3^T)\,\mathbf{F}^T(\mathbf{I} + \mu\,\mathbf{i}_3\mathbf{i}_3^T)\mathbf{u}_2.\end{aligned} \qquad (38)$$

Expanding out both sides of the equations and collecting terms gives the following quadratic in μ

$$\begin{aligned}0 = {} & \sigma_1^2 - \sigma_2^2 + \\ & \left[\left((\mathbf{u}_1^T\mathbf{i}_3)^2 + (\mathbf{v}_1^T\mathbf{i}_3)^2 \right) \sigma_1^2 - \left((\mathbf{u}_2^T\mathbf{i}_3)^2 + (\mathbf{v}_2^T\mathbf{i}_3)^2 \right) \sigma_2^2 \right] \mu + \\ & \left[(\mathbf{u}_1^T\mathbf{i}_3)(\mathbf{v}_1^T\mathbf{i}_3)\sigma_1 - (\mathbf{u}_2^T\mathbf{i}_3)(\mathbf{v}_2^T\mathbf{i}_3)\sigma_2 \right] F_{33}\,\mu^2.\end{aligned} \qquad (39)$$

This quadratic has at most two solutions unless all the coefficients vanish identically. If the highest order coefficient does not vanish we have observed that the roots μ_1, μ_2 of the quadratic always appear to be real and satisfy $\mu_1 < -1 < \mu_2$, but have not been able to prove this. If this does indeed hold true, (18) shows that this will give one real and positive focal length and one imaginary focal length: the latter can be discarded immediately leaving the former as the unique solution.

Thus more than two solutions will only occur if all coefficients vanish. A necessary condition for this is that $\sigma_1 = \sigma_2$, i.e. that the matrix be an essential matrix. This by itself is not sufficient, however a complete analysis of exactly what configurations will achieve this is not attempted here: we simply note that they are likely to be rather special configurations that are already known to be degenerate, such as when there is no rotation (see [1, 2] for a more detailed analysis of these situations).

REFERENCES

[1] M. J. Brooks, L. de Agapito, D. Q. Huynh, and L. Baumela, "Direct Methods for Self-Calibration of a Moving Stereo Head," *Proc. European Conference on Computer Vision*, Cambridge, UK. 14-18 Apr 1996. To appear.

[2] O. D. Faugeras, Q. T. Luong, and S. J. Maybank, "Camera Self-Calibration: Theory and Experiments," *Proc. European Conference on Computer Vision*, Italy. pp. 321–334, May 1992.

[3] O. Faugeras. *Three-Dimensional Computer Vision — A Geometric Viewpoint.* The MIT Press, Cambridge, 1993.

[4] H. G. Fourcade, "A New Method of Aerial Surveying," *Transactions of the Royal Society of South Africa*, XIV(1):93–112, 1926.

[5] G. H. Golub and C. F. V. Loan. *Matrix Computations.* Johns Hopkins series in the mathematical sciences. Johns Hopkins University Press, 2nd edition, Baltimore, 1989.

[6] R. I. Hartley, "Estimation of Relative Camera Positions for Uncalibrated Cameras," *Proc. European Conference on Computer Vision*, Italy. pp. 579–587, May 1992.

[7] B. K. P. Horn, "Relative Orientation," *International Journal of Computer Vision*, 4:59–78, 1990.

[8] D. Q. Huynh, M. J. Brooks, L. de Agapito, and H.-P. Pan, "Stereo Cameras with Coplanar Optical Axes: A Degenerate Configuration for Self-Calibration," Publication Report No. 96/20, Centre for Sensor Signal and Information Processing, Adelaide, Australia, Mar 1996.

[9] H. C. Longuet-Higgins, "A Computer Algorithm for Reconstructing a Scene from two Projections," *Nature*, 293(10):133–135, Sep 1981.

[10] Q. T. Luong and O. D. Faugeras, "A Stability Analysis of the Fundamental Matrix," *Proc. European Conference on Computer Vision*, vol. 1, pp. 577–588, May 1994.

[11] Q.-T. Luong, R. Deriche, O. Faugeras, and T. Papadopoulo, "On Determining the Fundamental Matrix: Analysis of Different Methods and Experimental Results," Report No. 1894, INRIA, Apr 1993.

[12] I. Niini, "Relative Orientation of Multiple Images Using Projective Singular Correlation," *ISPRS Commission III Symposium: Spatial Information from Digital Photogrammetry and Computer Vision*, Munich, Germany. vol. 30, Part 3/2, pp. 615–621, Sep 1994.

[13] H.-P. Pan, M. J. Brooks, and G. N. Newsam, "Image Resituation: Initial Theory," *SPIE Videometrics IV*, Philadelphia, Pennsylvania, USA. vol. 2598, pp. 162–173, 22-27 Oct 1995.

[14] H.-P. Pan, D. Q. Huynh, and G. K. Hamlyn, "Two-Image Resituation: Practical Algorithm," *SPIE Videometrics IV*, Philadelphia, Pennsylvania, USA. vol. 2598, pp. 174–190, 22-27 Oct 1995.

[15] P. Stefanovic, "Relative Orientation – A New Approach," *ITC Journal*, pp. 417–448, 1973.

[16] E. H. Thompson, "A Rational Algebraic Formulation of the Problem of Relative Orientation," *The Photogrammetric Record*, 3(14):152–159, 1959.

[17] J. Weng, T. S. Huang, and N. Ahuja, "Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(5):451–476, May 1989.