Direct methods for self-calibration
of a moving stereo head

M.J. Brooks[1,3]   L. de Agapito[2]   D.Q. Huynh[1]   L. Baumela[3]

[1] Centre for Sensor Signal and Information Processing, Signal Processing Research
Institute, Technology Park, Adelaide, SA 5095, Australia
[2] Instituto de Automática Industrial, CSIC, La Poveda, Madrid 28500, Spain
[3] Departamento de Inteligencia Artificial, Facultad de Informática, Universidad
Politécnica de Madrid, 28660 Boadilla del Monte, Spain
Email correspondence: mjb@cs.adelaide.edu.au

Abstract. We consider the self-calibration problem in the special context of a stereo head, where the two cameras are arranged on a lateral rig with coplanar optical axes, each camera being free to vary its angle of vergence. Under various constraints, we derive explicit forms for the epipolar equation, and show that a static stereo head constitutes a degenerate camera configuration for carrying out self-calibration in the sense of Hartley [4]. The situation is retrieved by consideration of a special kind of motion of the stereo head in which the baseline remains confined to a plane. New closed-form solutions for self-calibration are thereby obtained, inspired by an earlier discrete motion analysis of Zhang et al. [11]. Key factors in our approach are the development of explicit, analytical forms of the fundamental matrix, and the use of the vergence angles in the parameterisation of the problem.

Keywords: Self-calibration, stereo head, degeneracy, epipolar equation, fundamental matrix, ego-motion.

1   Introduction

The simultaneous recovery of intrinsic and extrinsic parameter values from both static and dynamic imagery has in recent times attracted considerable attention (see [2], [3], [4], [7], [8], [9], [10], [11], [12], [13]). In this paper, we consider this self-calibration problem in the special context of a stereo head, perhaps the most commonly adopted binocular camera configuration in robotics. First, however, we recall the epipolar geometry which underpins the analysis.

We adopt a notation similar, but not identical, to that of Faugeras et al. [3]; see the Appendix for a summary of the differences. Let $\mathbf{m}$ and $\mathbf{m}'$ denote corresponding points, in homogeneous coordinates, in the left and right images, respectively. We may express the epipolar equation as

$$\mathbf{m}^T \mathbf{F}\, \mathbf{m}' = 0, \tag{1}$$

where $\mathbf{F}$ is the fundamental matrix [3, 6], defined as

$$\mathbf{F} = \mathbf{A}^T \, \mathbf{T} \, \mathbf{R} \mathbf{A}'. \qquad (2)$$

Here, $\mathbf{R}$ embodies the pure rotation that renders the left image parallel with the right image, $\mathbf{T}$ is a skew-symmetric matrix formed from the baseline vector connecting the left and right optical centres, and $\mathbf{A}$ and $\mathbf{A}'$ are the intrinsic parameter matrices of the left and right cameras. (Note again the use in equations (1) and (2) of non-standard, but convenient, definitions of the various matrices. See the Appendix.)

As is well known, no more than 7 imaging parameters may be recovered from the epipolar equation, due to the special properties of the fundamental matrix [3]. In particular, Hartley [4] and Pan et al. [8] have shown that, under favourable conditions, two focal lengths and 5 relative orientation parameters may be recovered by a process of self-calibration.

In the next section, we consider the special situation in which two cameras form a stereo head assembly. We show that, in the absence of motion, this commonly adopted configuration is degenerate in that self-calibration may no longer be carried out in the sense of Hartley [4]. Direct methods of self-calibration are then explored in the context of a moving stereo head. Note that an extended version of this work appears in [1].

2   Stereo head assembly

Consider the special case of a stereo head in which a pair of cameras is mounted on a lateral rig. The cameras are free to vary their angles of vergence. The y-axes of the two images are parallel, and are orthogonal to the baseline vector, as depicted in Figure 1. The optical axes and the baseline are therefore coplanar. The matrices $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{A}$ now take the forms

$$\mathbf{R} = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}, \qquad \mathbf{T} = \begin{pmatrix} 0 & -t_z & 0 \\ t_z & 0 & -t_x \\ 0 & t_x & 0 \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & -f \end{pmatrix}.$$

Here, β is the angular rotation about the y-axis that renders the left image parallel with the right image; $\mathbf{T}$ is formed out of the baseline vector $\mathbf{t} = (t_x, 0, t_z)^T$; and, for each camera, the focal length and the principal point are the only unknown intrinsic parameters, denoted by f and (u_0, v_0), with the image coordinate system axes assumed to be orthogonal and similarly scaled.

In view of (2), the fundamental matrix is now given by

$$\mathbf{F} = \begin{pmatrix} 0 & -t_z & t_z v_0' \\ \phi & 0 & -u_0'\phi - f'\sigma \\ -v_0\phi & u_0 t_z - f t_x & u_0' v_0 \phi + v_0'(f t_x - u_0 t_z) + f' v_0 \sigma \end{pmatrix}, \qquad (3)$$

where $\phi = t_z \cos\beta - t_x \sin\beta$, and $\sigma = -t_x \cos\beta - t_z \sin\beta$.
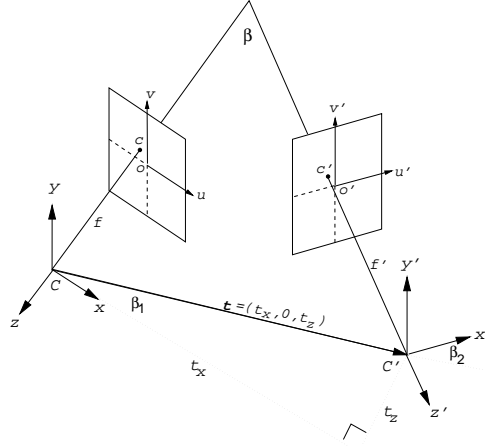
Fig. 1. Stereo head configuration.

As is well known, absolute dimensions of depth cannot be determined solely from knowledge of corresponding points and the associated fundamental matrix. Accordingly, without loss of generality, we set the baseline length to unity, and note that the direction of the baseline vector is now effectively described by 1 parameter. There are therefore 8 unknowns encoded within $\mathbf{F}$, these being $\beta, u_0, v_0, f, u_0', v_0', f'$ and either t_x or t_z.

2.1 Vergence-angle parameterisation

The form of $\mathbf{F}$ is simplified if an adjustment is made to the parameterisation by incorporating the left and right vergence angles β_1 and β_2, where

$$t_x = \cos \beta_1, \qquad t_z = \sin \beta_1, \qquad \beta = \beta_1 + \beta_2. \tag{4}$$

Here, β_1 and β_2 specify the extent to which the left and right optical axes point inwards from the direction 'straight-ahead'. (Note, therefore, that the left and right vergence angles are measured in an opposite sense. See Figure 2.) Relative orientation in this situation is now determined by the pair β_1, β_2, instead of β, t_x. Equation (3) is then expressed as

$$\mathbf{F} = \begin{pmatrix} 0 & -\sin\beta_1 & v_0' \sin\beta_1 \\ -\sin\beta_2 & 0 & (u_0' \sin\beta_2 + f' \cos\beta_2) \\ v_0 \sin\beta_2 & (u_0 \sin\beta_1 - f\cos\beta_1) & \begin{array}{c} -v_0(u_0' \sin\beta_2 + f' \cos\beta_2) \\ -v_0'(u_0 \sin\beta_1 - f\cos\beta_1) \end{array} \end{pmatrix}, \tag{5}$$

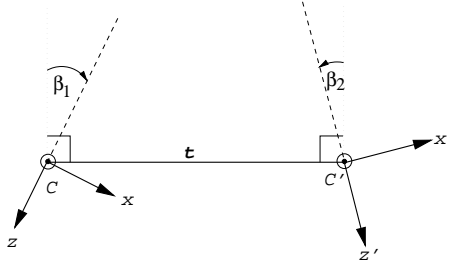in which the 8 unknowns are $\beta_1, \beta_2, u_0, v_0, f, u_0', v_0', f'$.

Fig. 2. Plan view of stereo head showing vergence angles.

In the event that sufficiently-many corresponding points can be located in the two images, it may be possible to obtain a numerical estimate, $\mathbf{F}_{est}$, of the matrix $\mathbf{F}$. Let

$$\mathbf{F}_{est} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_3 \\ \delta_4 & \delta_5 & \delta_6 \\ \delta_7 & \delta_8 & \delta_9 \end{pmatrix}. \tag{6}$$

Noting that $\mathbf{F}_{est}$ may only be determined up to a scale factor, we may form the equation

$$\mathbf{F} = \lambda \mathbf{F}_{est}. \tag{7}$$

Here, the unknown λ aligns the scales of the two matrices. We can now obtain 7 equations by linking respective elements of the matrices. However, these equations are not all independent, since $F_{33} = F_{23}F_{31}/F_{21} + F_{32}F_{13}/F_{12}$. Thus we may obtain up to 6 independent equations, one of which will be utilised in eliminating λ. We therefore observe that, of the 8 unknowns encoded within $\mathbf{F}$, at most 5 may be determined provided the remaining 3 are known. Further constraints are therefore needed if we are to solve for the various parameters.

3   Self-calibration of a static stereo head

We now consider how (7) may be solved under additional assumptions. In doing so, we shall gain insight into the feasibility of self-calibration of a static rig. Our aim is to develop closed-form expressions for the various parameters.

Case 1: u_0, v_0, u_0', v_0' known

If the locations of the principal points are known, we may without loss of generality offset the coordinates of image points. This simple form of image rectification then permits the setting of $u_0 = v_0 = u_0' = v_0' = 0$ in (7), giving

$$\mathbf{F} = \begin{pmatrix} 0 & -\sin\beta_1 & 0 \\ -\sin\beta_2 & 0 & f'\cos\beta_2 \\ 0 & -f\cos\beta_1 & 0 \end{pmatrix} = \lambda \begin{pmatrix} 0 & \delta_2 & 0 \\ \delta_4 & 0 & \delta_6 \\ 0 & \delta_8 & 0 \end{pmatrix}. \tag{8}$$

This case is relevant either to a pair of static stereo cameras having independent, unknown focal lengths, or to a single mobile camera in which the focal length may be varied. We now have 5 unknown parameters, including λ, but are able to generate only 4 independent equations. Thus, whereas in general we may obtain via self-calibration 2 focal lengths and 5 relative orientation parameters, we are unable to fix any of the unknown parameters in this special situation. The camera configuration is therefore degenerate for Hartley self-calibration.

Case 2: u_0, v_0, u_0', v_0' known; $\quad f = f'$

Here we assume that the left and right focal lengths are equal, and seek only the 3 unknown parameters f, β_1, β_2. Equation (7) now reduces to

$$\mathbf{F} = \begin{pmatrix} 0 & -\sin\beta_1 & 0 \\ -\sin\beta_2 & 0 & f\cos\beta_2 \\ 0 & -f\cos\beta_1 & 0 \end{pmatrix} = \lambda \begin{pmatrix} 0 & \delta_2 & 0 \\ \delta_4 & 0 & \delta_6 \\ 0 & \delta_8 & 0 \end{pmatrix}, \tag{9}$$

yielding 4 independent equations. All 3 imaging parameters can now be determined, viz:

$$f = \sqrt{(\delta_8^2 - \delta_6^2)/(\delta_4^2 - \delta_2^2)}, \quad \tan\beta_1 = f\,\delta_2/\delta_8, \quad \tan\beta_2 = -f\,\delta_4/\delta_6. \tag{10}$$

Note here that f is computed in the same units as the coordinates of the corresponding points used to estimate the fundamental matrix.

Case 3: u_0, v_0, u_0', v_0' known; $\quad \beta_1 = \beta_2 = \beta/2$

We now seek to determine β, f and f', given the very special situation in which the vergence angles are equal, with the principal axes of the cameras and the baseline forming an isosceles triangle. Our equation is now

$$\mathbf{F} = \begin{pmatrix} 0 & -\sin\beta/2 & 0 \\ -\sin\beta/2 & 0 & f'\cos\beta/2 \\ 0 & -f\cos\beta/2 & 0 \end{pmatrix} = \lambda \begin{pmatrix} 0 & \delta_2 & 0 \\ \delta_4 & 0 & \delta_6 \\ 0 & \delta_8 & 0 \end{pmatrix}. \tag{11}$$

Noting that $F_{12} = F_{21}$, we see that none of the unknown parameters may be determined without more information being provided. Remarkably, if the focal lengths are known to be equal, it remains impossible to recover any of the parameters. Note, however, that the ratio of the focal lengths may be determined.

Case 4: $(u_0, v_0, f) = (u_0', v_0', f')$

Here we assume left and right cameras have identical focal length and principal point locations. This also corresponds to a mobile camera moving horizontally. The 5 parameters $\beta_1, \beta_2, u_0, v_0, f$ are now free, (7) reducing to:

$$\begin{pmatrix} 0 & -\sin\beta_1 & v_0\sin\beta_1 \\ -\sin\beta_2 & 0 & u_0\sin\beta_2 + f\cos\beta_2 \\ v_0\sin\beta_2 & \begin{matrix} u_0\sin\beta_1 \\ -f\cos\beta_1 \end{matrix} & \begin{matrix} -u_0 v_0(\sin\beta_1 + \sin\beta_2) \\ +v_0 f(\cos\beta_1 - \cos\beta_2) \end{matrix} \end{pmatrix} = \lambda \begin{pmatrix} 0 & \delta_2 & \delta_3 \\ \delta_4 & 0 & \delta_6 \\ \delta_7 & \delta_8 & \delta_9 \end{pmatrix}. \tag{12}$$

Note that $F_{31} = F_{21}F_{13}/F_{12}$ and $F_{33} = (F_{23} + F_{32})F_{13}/F_{12}$, and so only 5 independent equations may be generated. Given the need to eliminate λ, at most 4 of the 5 parameters may be determined, provided the remaining parameter is known.

Further discussion of self-calibration and the effect of either the tilting or rotation about the optical axis of one camera is to be found in [1].

4 Self-calibration of a horizontally moving stereo head

Having seen that a static stereo head, with coplanar optical axes, is a degenerate configuration for self-calibration, we now assess the consequences of moving the head. Specifically, we permit:

- motion of the head such that the optical axes of the cameras are confined to a plane. This therefore captures the situation in which an upright robot head may translate or rotate in the horizontal plane.
- independent vergence angles of the head that may vary with the motion.
- each camera to have an unknown but fixed focal length.

The following analysis adopts a technique of Zhang et al. [12] in which various fundamental matrices are utilised.

4.1 Formulating the fundamental matrices

Let the rig move from an initial position to a final position. Let the left-right pair of images in the initial position be termed I_1 and I_2, and let the left-right images in the final position be termed I_3 and I_4 (see Figure 3). The left camera is thus responsible for the successive images I_1 and I_3. Assume that the determining of corresponding points has led to estimates for the fundamental matrices linking the following image pairs: (I_1, I_2), (I_3, I_4), (I_1, I_3), (I_2, I_4). Let the associated analytical fundamental matrices be termed $\mathbf{F}^{12}$, $\mathbf{F}^{34}$, $\mathbf{F}^{13}$, $\mathbf{F}^{24}$. We shall not here make use of $\mathbf{F}^{14}$ and $\mathbf{F}^{23}$. As before, we aim to solve for the parameters embedded within these matrices by exploiting the fact that the analytical and the estimated forms of the fundamental matrix are directly proportional. Note that, in this regard, the approach pursued in Zhang et al. [12] is quite different in that a least-squares approach is used to solve a more general problem in which motion is not confined to the plane (although, unlike here, the relative orientation of the head is assumed fixed).

Recalling (8), the initial position of the rig gives rise to the fundamental matrix, $\mathbf{F}^{12}$, given by

$$\mathbf{F}^{12} = \begin{pmatrix} 0 & -\sin\beta_1^{12} & 0 \\ -\sin\beta_2^{12} & 0 & f'\cos\beta_2^{12} \\ 0 & -f\cos\beta_1^{12} & 0 \end{pmatrix}. \tag{13}$$

Assuming that the focal lengths of the respective cameras remain fixed, and that the vergence angles are free to shift, we obtain the following fundamental matrix,
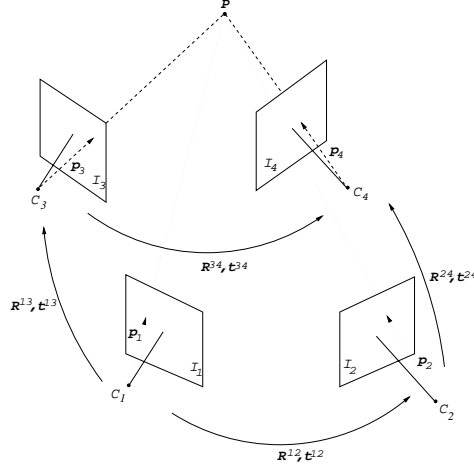
Fig. 3. Motion of the stereo head.

$\mathbf{F}^{34}$, capturing the epipolar relationship between the left and right images of the rig in its final position:

$$\mathbf{F}^{34} = \begin{pmatrix} 0 & -\sin\beta_1^{34} & 0 \\ -\sin\beta_2^{34} & 0 & f'\cos\beta_2^{34} \\ 0 & -f\cos\beta_1^{34} & 0 \end{pmatrix}. \tag{14}$$

Similarly, the fundamental matrix relating the image pair (I_1, I_3) is given by

$$\mathbf{F}^{13} = \begin{pmatrix} 0 & -\sin\beta_1^{13} & 0 \\ -\sin\beta_2^{13} & 0 & f\cos\beta_2^{13} \\ 0 & -f\cos\beta_1^{13} & 0 \end{pmatrix}, \tag{15}$$

with the fundamental matrix for the image pair (I_2, I_4) being

$$\mathbf{F}^{24} = \begin{pmatrix} 0 & -\sin\beta_1^{24} & 0 \\ -\sin\beta_2^{24} & 0 & f'\cos\beta_2^{24} \\ 0 & -f'\cos\beta_1^{24} & 0 \end{pmatrix}. \tag{16}$$

Here we note that the focal lengths of the respective cameras remain unchanged in the movement of the rig from its initial to final position. We observe that, under the above parameterisation, image I_1 undergoes a rotation of $(\beta_1^{13} + \beta_2^{13})$, relative to its own local coordinate system, in becoming oriented in parallel with image I_3.

4.2   Solving the fundamental matrix equations

It is now necessary to further enhance our notation so as to be able to deal simultaneously with various fundamental matrices. Let the numerical estimate, $\mathbf{F}^{ij}_{est}$, of fundamental matrix $\mathbf{F}^{ij}$ be represented as

$$\mathbf{F}^{ij}_{est} = \begin{pmatrix} \delta^{ij}_1 & \delta^{ij}_2 & \delta^{ij}_3 \\ \delta^{ij}_4 & \delta^{ij}_5 & \delta^{ij}_6 \\ \delta^{ij}_7 & \delta^{ij}_8 & \delta^{ij}_9 \end{pmatrix}, \tag{17}$$

and let $\mu^{ij}_k = (\delta^{ij}_k)^2$. Let a right bar and superscript indicate the fundamental matrix from which the elements derive. Thus, for example, note that

$$\mu_k + \mu_l\big|^{ij} = (\delta^{ij}_k)^2 + (\delta^{ij}_l)^2. \tag{18}$$

In view of the earlier analysis, resulting in (10), we may immediately infer that

$$f = \sqrt{\frac{\mu_8 - \mu_6}{\mu_4 - \mu_2}}\Bigg|^{13}, \qquad f' = \sqrt{\frac{\mu_8 - \mu_6}{\mu_4 - \mu_2}}\Bigg|^{24}$$

$$\tan\beta^{13}_1 = f\frac{\delta_2}{\delta_8}\Bigg|^{13}, \ \tan\beta^{13}_2 = -f\frac{\delta_4}{\delta_6}\Bigg|^{13}, \ \tan\beta^{24}_1 = f'\frac{\delta_2}{\delta_8}\Bigg|^{24}, \ \tan\beta^{24}_2 = -f'\frac{\delta_4}{\delta_6}\Bigg|^{24}$$

It then follows that

$$\tan\beta^{12}_1 = f\frac{\delta_2}{\delta_8}\Bigg|^{12}, \ \tan\beta^{12}_2 = -f'\frac{\delta_4}{\delta_6}\Bigg|^{12}, \ \tan\beta^{34}_1 = f\frac{\delta_2}{\delta_8}\Bigg|^{34}, \ \tan\beta^{34}_2 = -f'\frac{\delta_4}{\delta_6}\Bigg|^{34}$$

We therefore have closed-form solutions for the 2 focal lengths, and the rotations between images. Implicit in the above are the directions of the various translations between perspective centres. Note that we have so far not made any assumption about the rigidity or otherwise of the rig.

4.3   Solving the baseline constraint equation

We have yet to completely determine the relative orientation of all image pairs as we have still to compute the relative magnitudes of the baselines. (As noted earlier, it is not possible to compute absolute scale of the baselines only from corresponding points.) These relative magnitudes will complete the description of the motion of the head.

Let the magnitude of the head's baseline vector in the initial position be unity. The baseline vector, $\mathbf{t}^{12}$, may therefore be written as

$$\mathbf{t}^{12} = (\cos\beta^{12}_1, 0, \sin\beta^{12}_1)^T. \tag{19}$$

Letting L^{ij} denote the length of the baseline vector $\mathbf{t}^{ij}$, we may immediately write down the remaining baseline vectors as

$$\mathbf{t}^{13} = L^{13}(\cos\beta^{13}_1, 0, \sin\beta^{13}_1)^T \tag{20}$$

$$\mathbf{t}^{24} = L^{24}(\cos\beta^{24}_1, 0, \sin\beta^{24}_1)^T \tag{21}$$

$$\mathbf{t}^{34} = L^{34}(\cos\beta^{34}_1, 0, \sin\beta^{34}_1)^T. \tag{22}$$

Our task is now to determine the lengths L^{13}, L^{24}, L^{34}. Returning to Figure 3, we observe (after [12]) the baseline constraint equation

$$\mathbf{R}^{12}\,\mathbf{t}^{24} = \mathbf{t}^{13} - \mathbf{t}^{12} + \mathbf{R}^{13}\mathbf{t}^{34}. \tag{23}$$

Expanding this, we have

$$L^{24}\begin{pmatrix} \cos(\beta^{12} + \beta_1^{24}) \\ 0 \\ \sin(\beta^{12} + \beta_1^{24}) \end{pmatrix} = \begin{pmatrix} L^{13}\cos\beta_1^{13} - \cos\beta_1^{12} + L^{34}\cos\beta_1^{34}\cos\beta^{13} \\ -L^{34}\sin\beta_1^{34}\sin\beta^{13} \\ 0 \\ L^{13}\sin\beta_1^{13} - \sin\beta_1^{12} + L^{34}\cos\beta_1^{34}\sin\beta^{13} \\ +L^{34}\sin\beta_1^{34}\cos\beta^{13} \end{pmatrix}.$$

Here β^{13} rotates I_1 parallel to I_3 and is such that $\beta^{13} = \beta_1^{13} + \beta_2^{13}$. Recall that $\beta^{12} = \beta_1^{12} + \beta_2^{12}$. Clearly, the 3 unknown lengths may not be determined from the above equation. But, on the assumption that the baseline length of the rig remains constant, so that L^{34} also has unit length, we may readily infer that

$$L^{13} = \left(\sin(\omega - \beta^{13}) - \sin\phi\right)/\sin\tau, \tag{24}$$

where $\omega = \beta^{12} + \beta_1^{24} - \beta_1^{34}$, $\phi = \beta_2^{12} + \beta_1^{24}$ and $\tau = \beta_1^{13} - \beta^{12} - \beta_1^{24}$. The formula for L^{24} then follows directly from the baseline constraint equation. We have therefore described the motion of the rig. Further discussion is given in [1].

5   Adding camera tilt to the moving stereo head

Our analysis here is a generalisation of that considered in the previous section in that the head may now tilt up or down, by a rotation about the baseline. We note that the baseline remains confined to a plane, and that the optical axes of the two cameras are at all times coplanar, but are not confined to the same plane in consecutive head positions. Critically, in the analysis presented here, either the initial or final position of the head should have zero tilt.

We now consider how a rotation of θ about the baseline maps the left image to a new position. A rotation of θ about the baseline is equivalent, in the left image's coordinate system, to three composite rotations: a rotation of β_1^{34} about the y-axis, followed by a rotation of θ about the x-axis, and then a rotation of $-\beta_1^{34}$ about the y-axis. In addition to this tilting, the previous rotation in the plane may still take place. The fundamental matrix may therefore be expressed as $\mathbf{F}^{13} = \mathbf{A}^T\,\mathbf{T}\,\mathbf{R}\mathbf{A}$, where

$$\mathbf{T} = \begin{pmatrix} 0 & -\sin\beta_1^{13} & 0 \\ \sin\beta_1^{13} & 0 & -\cos\beta_1^{13} \\ 0 & \cos\beta_1^{13} & 0 \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -f \end{pmatrix}, \tag{25}$$

and $\mathbf{R} = R_y(\beta_1^{13} + \beta_2^{13})R_y(\beta_1^{34})R_x(\theta)R_y(-\beta_1^{34})$. Here, we adopt the convention that $R_m(\psi)$ signifies a rotation of ψ about the m-axis.

A rather complex fundamental matrix results (given in [1]) from which the following equations may now be derived:

$$\tan \beta_1^{13} = f \frac{\delta_2}{\delta_8} \bigg|^{13}, \qquad \tan \beta_1^{34} = f \frac{\delta_1}{\delta_3} \bigg|^{13}, \qquad \tan \theta = - \frac{\delta_7}{\delta_8 \sin \beta_1^{34}} \bigg|^{13}$$

$$f^2 = \frac{-\delta_3 \delta_6}{\delta_2 \delta_5 + \delta_1 \delta_4} \bigg|^{13}, \quad \sec^2 (\beta_1^{34} + \beta_2^{13}) = \sin^2 \theta \left(1 + \frac{\mu_4 + \mu_6/f^2}{\mu_5} \right) \bigg|^{13}$$

We therefore have closed-form solutions for the 5 unknowns f, θ, β_1^{34}, β_1^{13}, β_2^{13}.

Consideration of the fundamental matrix $\mathbf{F}^{24}$ yields symmetric formulae for the right camera vergence angles. The analysis is completed when we note that the previous formulae for baseline lengths are precisely applicable here, since the moving baseline has remained confined to a plane.

6   Experimental Results

We now describe synthetic tests carried out on the method of self-calibration. A cloud of 35 points was randomly generated within a cubic volume of side 2400mm lying approximately 600mm in front of the stereo head. These points were then projected onto each of the 4 image planes arising in the two positions of the stereo head. The location of each image point was then perturbed in a random direction by a distance governed by a Gaussian distribution with zero mean and standard deviation, σ, expressed in pixel units. Such a distribution results in an expected value for the perturbation distance of approximately $0.8\,\sigma$. As a matter of interest, in the many tests carried out here, the highest perturbation distance was found to be $3.7\,\sigma$.

Left and right focal lengths were set at 6mm and 8mm, with a fixed baseline length of 300mm. Vergence angles were 15 deg and 17 deg in the initial position, and 18 deg and 22 deg in the final position. The motion of the head was such that the upward tilt was 10 deg, rotation of the baseline in the plane was 12 deg, with the length of the translation vector mapping the left camera from initial to final position being approximately the same as the baseline length of the head. Image sizes were 1000×1000 pixels.

Experiments were conducted with σ varying from 0.0 to 1.2 in steps of 0.1. For each value of σ, self-calibration was run 20 times (each time operating on a different set of images) and the root-mean-square (rms) error of each parameter was computed. Figure 4 gives a brief summary of how self-calibration is affected by increasing noise, in the case considered. Errors (rms) in lengths and tilt rotation are given as percentages of the true values, while errors (rms) in the vergence angles are expressed in degrees. We can see from the figures that errors in the estimates of the various parameters vary approximately linearly with the extent of the introduced noise, over the range considered. The lengths of the translation vectors L^{13} and L^{24} are the parameters most affected by noise, with relative errors of up to 14% occurring with noise $\sigma = 1.2$. At this high noise level, the rms error of the right camera's focal length is 5.2%, the error in the
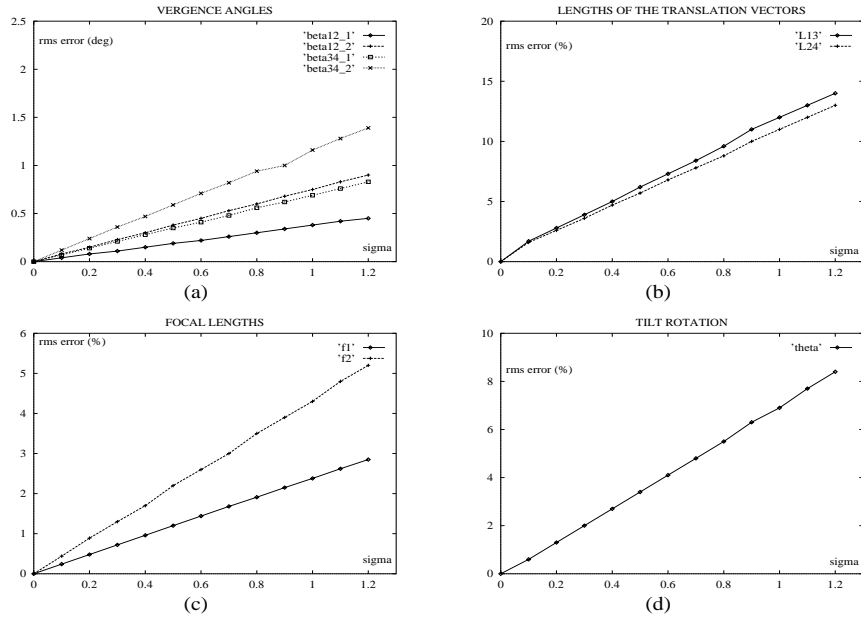
Fig. 4. Results of the experiments.

estimated tilt rotation is 8.5%, and the maximum rms error of the vergence angles is 1.4 deg. Comparable results were obtained for similar head movements.

Given that points can routinely be located by automated techniques with an accuracy of better than $\sigma = 0.5$, the above results suggest that this approach holds promise. Note that no special efforts have been made to optimise the estimates obtained via this process of self-calibration. Thus, for example, no effort has been expended to generate more accurate estimates of fundamental matrices by reducing the deleterious impact of poor localisation of points, nor has any optimisation of the estimates been attempted as a post-process.

Acknowledgements

References

1. Brooks, M. J., Agapito, L., Huynh, D. Q. and Baumela, L. Direct Methods for Self-Calibration of a Moving Stereo Head. Tech. Rep. 1/96, CSSIP, January, 1996.
2. Brooks, M. J., Baumela, L. and Chojnacki, W. An Analytical Approach to Determining the Egomotion of a Camera having Free Intrinsic Parameters. Tech. Rep. 96-04, Dept. Computer Science, University of Adelaide, January, 1996.
3. Faugeras, O. D., Luong, Q. T. and Maybank, S. J. Camera Self-Calibration: Theory and Experiments. In *Proc. European Conference on Computer Vision* (1992), pp. 321–334.
4. Hartley, R. I. Estimation of Relative Camera Positions for Uncalibrated Cameras. In *Proc. European Conference on Computer Vision* (1992), pp. 579–587.
5. Huynh, D. Q., Brooks, M. J., Agapito, L. de and Pan. H-P. Stereo Cameras With Coplanar Optical Axes: a Degenerate Configuration for Self-Calibration. Tech. Rep. 2/96, CSSIP, January, 1996.
6. Luong, Q.-T., Deriche, R., Faugeras, O. and Papadopoulo, T. On Determining the Fundamental Matrix: Analysis of Different Methods and Experimental Results. Tech. Rep. 1894, INRIA, 1993.
7. Maybank, S. J. and Faugeras, O. D. A Theory of Self-Calibration of a Moving Camera. *International Journal of Computer Vision*, 8, 2 (1992), 123–151.
8. Pan, H.-P., Brooks, M. J. and Newsam, G. N. Image Resituation: Initial Theory. In *SPIE Videometrics* (1995), vol. 2598, pp. 162–173.
9. Viéville, T., Luong, Q. and Faugeras, O. Motion of Points and Lines in the Uncalibrated Case. *International Journal of Computer Vision* 17, 1 (1994).
10. Viéville, T. and Faugeras, O. Motion Analysis with a Camera with Unknown, and Possibly Varying Intrinsic Parameters. In *ICCV'95* (Cambridge, MA, June 1995), IEEE, pp. 750–756.
11. Zhang, Z., Deriche, R., Faugeras, O. and Luong, Q. T. A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry. Tech. Rep. 2273, INRIA, 1994.
12. Zhang, Z., Luong, Q. T. and Faugeras, O. D. Motion of an Uncalibrated Stereo rig: Self-Calibration and Metric Reconstruction. Tech. Rep. 2079, INRIA, 1993.
13. Zisserman, A., Beardsley, P. A. and Reid, I. Metric Calibration of a Stereo Rig. Proc. IEEE Workshop on Representation of Visual Scenes, Boston, June, 1995.

A Notation semantics

Our notation differs from the standard notation of Faugeras et al. [3] (henceforth termed the Faugeras notation). Symbols $\mathbf{F}$, $\mathbf{T}$, $\mathbf{R}$ and $\mathbf{A}$ denote in this work the fundamental, translation, rotation and intrinsic-parameter matrices, respectively. Let the corresponding matrices in Faugeras notation be denoted F, T, R and A. Herein, the epipolar equation has the form $\mathbf{m}^T \mathbf{F} \mathbf{m}' = 0$, where $\mathbf{F} = \mathbf{A}^T \mathbf{T} \mathbf{R} \mathbf{A}'$. This contrasts with Faugeras notation, where $\mathbf{m}'^T F \mathbf{m} = 0$, and $F = A'^{-T} T R A^{-1}$. The full list of notational relationships is now given:

$$\mathbf{F} = \sqrt{\det(A)\,\det(A')}\,F^T, \quad \mathbf{A} = -\sqrt{\det(A)}\,A^{-1}, \quad \mathbf{R} = R^T, \quad \mathbf{T} = -R^T\,T\,R.$$

See [1] for further discussion.

This article was processed using the LaTeX macro package with ECCV'96 style