INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Camera Self-Calibration Using the Singular Value Decomposition of the Fundamental Matrix: From Point Correspondences to 3D Measurements

Manolis I.A. LOURAKIS and Rachid DERICHE

N° 3748

Août 1999

THÈME 3

Rapport de recherche

ISSN 0249-6399

Camera Self-Calibration Using the Singular Value Decomposition of the Fundamental Matrix: From Point Correspondences to 3D Measurements

Manolis I.A. LOURAKIS and Rachid DERICHE

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Robotvis

Rapport de recherche n° 3748 — Août 1999 — 45 pages

Abstract: This paper deals with a fundamental problem in motion and stereo analysis, namely that of determining the camera intrinsic calibration parameters. A novel method is proposed that follows the autocalibration paradigm, according to which calibration is achieved not with the aid of a calibration pattern but by observing a number of image features in a set of successive images. The proposed method relies upon the Singular Value Decomposition of the fundamental matrix, which leads to a particularly simple form of the *Kruppa* equations. In contrast to the classical formulation that yields an over-determined system of constraints, the derivation proposed here provides a straightforward answer to the problem of determining which constraints to employ among the set of available ones. Moreover, the derivation is a purely algebraic one, without a need for resorting to the somewhat non-intuitive geometric concept of the *absolute conic*. Apart from the fundamental matrix itself, no other quantities that can be extracted from it (e.g. the epipoles) are needed for the derivation. Experimental results from extensive simulations and several image sequences demonstrate the effectiveness of the proposed method in accurately estimating the intrinsic calibration matrices. It is also shown that the computed intrinsic calibration matrices are sufficient for recovering 3D motion and performing metric measurements from uncalibrated images.

Key-words: *Self-Calibration, Kruppa Equations, 3D Measurements, Motion Analysis, Stereo, Structure from Motion.*

This work was funded in part under the VIRGO research network (EC Contract No ERBFMRX-CT96-0049) of the TMR Programme.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 — Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Auto-Calibration par Décomposition en Valeurs Singulières de la Matrice Fondamentale: De l'Appariement de Points aux Mesures 3D

Résumé : Ce rapport traite du problème fondamental de l'auto-calibration d'une caméra à partir d'un ensemble de points appariés entre différentes images. Une méthode basée sur les équations de Kruppa, bien connues dans le cadre de cette application, est développée. On fait usage de la décomposition en valeurs singulières de la matrice fondamentale pour dériver de manière purement algébrique des équations de Kruppa remarquablement simplifiées. Ceci permet en particulier de resoudre le problème du choix des deux équations de Kruppa à utiliser parmi l'ensemble plus grand des équations dérivées par la méthode classique. Dans cette méthode, les équations sont dérivées très simplement, on ne fait nullement usage de l'interprétation géometrique à base de la conique absolue, ni de celle liée au plan à l'infini, et on n'utilise pas explicitement les épipoles, dont l'estimation est connue pour être instable. Enfin et surtout, cette méthode est mise en oeuvre, comparée et testée avec succès pour retrouver les paramètres intrinsèques de différentes caméras à partir de données synthétiques bruitées et de plusieures images réelles. On montre aussi que la qualité des résultats obtenus permet de valider remarquablement l'approche jusqu'à l'obtention de mesures 3D fiables à partir d'images.

Mots-clés : *Auto-Calibration, Equations de Kruppa, Mesures 3D, Analyse du mouvement, Stéréo-vision, Structure à partir du mouvement.*

1 Introduction

It is well-known that a pinhole camera acts like a projective transformation device [15, 14]. This means that from a sequence of images, 3D structure can be recovered modulo a projectivity of 3D space [15, 22]. Numerous vision tasks, however, ranging from visual navigation and 3D reconstruction to novel view synthesis and augmented reality, require that this ambiguity is raised so that metric (i.e. Euclidean) quantities can be computed from images. In order to facilitate this, the problem of determining the intrinsic calibration parameters of the camera needs to be solved [16]. Early approaches for coping with this problem rely upon the presence of an artificial calibration object in the set of captured images [48, 49]. Knowledge of the 3D shape of the calibration object supplies the 3D coordinates of a set of reference points in a coordinate system attached to the calibration object. Thus, the transformation relating the 3D coordinates to their associated image projections can be recovered through an optimization process. Despite that intrinsic calibration can be accurately computed with such approaches, their major drawback is that they are suitable for off-line calibration only. In other words, they are inapplicable in cases where the intrinsic parameters undergo constant changes due to focusing, zooming or mechanical and thermal variations. This is particularly true in the context of active vision, where the camera optical parameters undergo continuous, purposive changes [1].

In a seminal paper, Maybank and Faugeras [35] have shown that the calibration problem can be solved without resorting to a calibration object. By tracking a set of points among images of a rigid scene, captured while the camera is pursuing unknown, unrestricted motion with constant intrinsic calibration parameters, the latter can be estimated by determining the image of the *absolute conic*. The absolute conic is a special conic lying at the plane at infinity, having the property that its image projection depends on the intrinsic parameters only. This fact is expressed mathematically by the so-called *Kruppa* equations [31]. In following years, several researchers have investigated the application of the Kruppa equations for solving the calibration problem. This paradigm, has become known under the names of *self* or *auto* calibration. For example, Zeller [51, 52] and Heyden and Åström [23] propose variants of the basic approach. Pollefeys and Van Gool [41, 39] describe a stratified approach to self-calibration, which starts from projective calibration, augments it with the homography of the plane at infinity to yield affine calibration and finally upgrades to Euclidean calibration. Luong and Faugeras [33] use the Kruppa equations to derive systems of polynomial equations, which are of degree four in five unknowns (i.e. the camera intrinsic parameters). These systems are solved with the use of numerical continuation methods. The main disadvantage of this approach is that it involves high computational costs. In addition, it is difficult to take into account the uncertainty associated with the estimates of the epipolar geometry and any a priori knowledge regarding the intrinsic parameters that might be available. All the methods mentioned above have demonstrated that intrinsic calibration can be recovered fairly accurately, provided that image features can be well localized and reliably tracked among images. In the opposite case, however, these methods can easily run into difficulties due to the nonlinearity of the equations and the large number of unknowns involved. To improve the stability of self-calibration, Sturm [44] proposes a scheme that captures the interdependence of the intrinsic parameters in an off-line pre-calibration process. By assuming a moving camera equipped with a zoom lens, he models all but one intrinsic parameters as a function of the remaining one. Thus, self-calibration is reduced to estimating only one parameter, something that can be done by finding the closest common root of several low degree polynomials. At this point, it should be mentioned that most self-calibration methods implicitly assume that the camera undergoes a general rigid motion, i.e. a combination of general translation and rotation. This is because an inherent problem in self-calibration is the fact that sequences of *critical* camera motions lead to ambiguities in the recovered calibration parameters. Classes of such motion sequences are studied by Sturm [43], Zisserman et al [56] and Demirdjian et al [11]. However, it is interesting to note that Hartley [21] has shown that in the case of purely rotating cameras, the self-calibration problem can be solved by a linear algorithm. A special case is also considered by Armstrong et al [3]. They show that using three images, the intrinsic calibration of a camera undergoing planar motion can be recovered up to a two fold ambiguity.

Due to the widespread use of binocular heads in robotics, the problem of autocalibrating a stereo rig from a sequence of images has also received considerable attention [55, 26, 6, 27, 25]. Owing to the redundancy of the information contained in a sequence of stereo images, such methods have the potential of being considerably more robust compared to self-calibration methods employing monocular image sequences.

Recently, attempts to solve the self-calibration problem in the case of varying camera intrinsic parameters have started to appear. Heyden and Åström [24], for example, have shown that self-calibration in the case of continuously focusing/zooming cameras is possible when the aspect ratio is known and no skew is present. Extending their work, Pollefeys et al [40] proved that the absence of skew in the image plane is sufficient to allow for self-calibration. They also proposed a framework based on the concept of the absolute quadric [46] for recovering the intrinsic parameters. Techniques for achieving self-calibration assuming restricted types of camera motion and non constant intrinsic parameters have also appeared [10, 9].

In this work, we propose a simplification of the Kruppa equations and show how it can be employed for self-calibration. The simplification is derived in a purely algebraic manner and is based solely on the fundamental matrix. Estimates of the epipoles, which are known to be difficult to compute accurately[1], are not needed. Therefore, compared to existing self-calibration methods, the proposed one has the potential of being more stable and robust with respect to measurement noise.

The rest of the paper is organized as follows. Section 2 reviews some background material and introduces the notation that is used in the remainder of the paper. Section 3 gives an overview of constraints on the intrinsic parameters that historically preceded the development of the Kruppa equations. The potential advantages and disadvantages of each approach are also discussed. Using a purely algebraic scheme, the classical Kruppa equations are derived in section 4 and a geometric in-

[1] This is particularly true in the case that the epipoles lie at infinity.

terpetation is supplied. The simplified Kruppa equations are derived in section 5. Section 6 describes in detail the proposed self-calibration method and discusses some implementation issues. Experimental results from a prototype implementation are presented in section 7. The paper concludes with a brief discussion in section 8.

2 Notation and Background

The projection model assumed for the camera is the projective one. The formalism capturing the details of the projection of 3D points on a planar retina is based on projective geometry. This section gives a brief review of some basic projective geometry concepts. For more detailed treatments regarding the application of projective geometry to computer vision, the reader is referred to the tutorial by Mohr and Triggs [36] or the relevant chapters of the books by Faugeras [14], Kanatani [30] or Mundy and Zisserman [37].

A 3D point $\mathbf{M} = [x, y, z]^t$ projects to a 2D image point $\mathbf{m} = [u, v]^t$ through a 3×4 projection matrix **P**, as follows:

$$s\hat{\mathbf{m}} = \mathbf{P}\hat{\mathbf{M}}, \tag{1}$$

where s is a nonzero scale factor and the notation $\hat{\mathbf{p}}$ is such that if $\mathbf{p} = [x, y, \ldots]^t$ then $\hat{\mathbf{p}} = [x, y, \ldots, 1]^t$.

In the case of a binocular stereo system, every physical point **M** in space yields a pair of 2D projections $\mathbf{m_1}$ and $\mathbf{m_2}$ on the two images. Those projections are defined by the following relations:

$$\begin{aligned} s_1\hat{\mathbf{m_1}} &= \mathbf{P_1}\hat{\mathbf{M}} \\ s_2\hat{\mathbf{m_2}} &= \mathbf{P_2}\hat{\mathbf{M}} \end{aligned} \tag{2}$$

Assuming that the two cameras are identical and that the world coordinate system is associated with the first camera, the two projection matrices are given by:

$$\mathbf{P_1} = [\mathbf{A}|\mathbf{0}] \tag{3}$$

$$\mathbf{P_2} = [\mathbf{A}\mathbf{R}|\mathbf{A}\mathbf{t}], \tag{4}$$

where **R** and **t** represent respectively the rotation matrix and the translation vector defining the rigid displacement between the two cameras. Note that the same relations still hold when considering a single moving camera instead of a binocular rig. Matrix **A** is the 3×3 intrinsic parameters matrix, having the following well-known form [16]:

$$\mathbf{A} = \begin{bmatrix} \alpha_u & -\alpha_u \cot\theta & u_0 \\ 0 & \alpha_v / \sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

The parameters α_u and α_v correspond to the focal distances in pixels along the axes of the image, θ is the angle between the two image axes and (u_0, v_0) are the coordinates of the image principal point. The ratio $\frac{\alpha_v}{\alpha_u}$ is known as the *aspect ratio*. In practice, θ is very close to $\frac{\pi}{2}$ for real cameras [14]. Additionally, modern cameras have almost rectangular pixels, therefore in this case the aspect ratio is very close to being 1. At this point, it should be mentioned that given the intrinsic calibration parameters matrix **A**, 3D angles and length ratios of 3D line segments can be directly computed from images. Further details on achieving this are given in appendix A.

In the sequel, **K** will denote the symmetric matrix $\mathbf{A}\mathbf{A}^t$. By eliminating the scalars s_1 and s_2 associated with the projection equations (2) as well as the point **M**, the following equation relating the pair of projections of the same 3D point is obtained:

$$\hat{\mathbf{m}_2}^t \mathbf{F} \hat{\mathbf{m}_1} = 0 \tag{5}$$

In this equation, matrix **F** is the fundamental matrix, given by

$$\mathbf{F} = \mathbf{A}^* [\mathbf{t}]_\times \mathbf{R} \mathbf{A}^{-1} \tag{6}$$

where $\mathbf{A}^* = (\mathbf{A}^{-1})^t$ is the adjoint matrix of $\mathbf{A}$ and $[\mathbf{x}]_\times$ denotes the antisymmetric matrix of vector **x** that is associated with the cross product. This matrix has the property $[\mathbf{x}]_\times \mathbf{y} = \mathbf{x} \times \mathbf{y}$ for each vector **y** and has the following analytic form:

$$[\mathbf{x}]_\times = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}$$

The fundamental matrix **F** describes the epipolar geometry between the pair of views considered. It is the equivalent to the essential matrix $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$ in the uncalibrated case, as dictated by (see also Eq.(6))

$$\mathbf{F} = \mathbf{A}^* \mathbf{E} \mathbf{A}^{-1} \tag{7}$$

Due to the above relation, **E** can be written as a function of **F** as follows:

$$\mathbf{E} = \mathbf{A}^t \mathbf{F} \mathbf{A} \tag{8}$$

As pointed out by Trivedi [47], the symmetric matrix $\mathbf{E}\mathbf{E}^t$ is independent of the rotation **R** since

$$\mathbf{E}\mathbf{E}^t = [\mathbf{t}]_\times \mathbf{R}\mathbf{R}^t ([\mathbf{t}]_\times)^t = [\mathbf{t}]_\times ([\mathbf{t}]_\times)^t \tag{9}$$

Substitution of Eq.(8) into the above equation yields

$$\mathbf{F}\mathbf{K}\mathbf{F}^t = \mathbf{A}^* [\mathbf{t}]_\times ([\mathbf{t}]_\times)^t \mathbf{A}^{-1} \tag{10}$$

This last equation will be employed in subsequent sections for algebraically deriving the Kruppa equations.

3 The Constraints of Huang-Faugeras and of Trivedi

Before presenting the simplified version of the Kruppa equations, this section briefly presents the constraints of Huang-Faugeras [28] as well as those of Trivedi [47]. These constraints, known since the late eighties, preceded the appearance of autocalibration methods based on the Kruppa equations. Apart from making this paper more self-contained, the recall of those previous works enable us to raise some questions of particular importance to the problem of autocalibration.

3.1 The constraints of Huang-Faugeras

Based on the constraints that an essential matrix must satisfy, equations that are equivalent to those of Kruppa but having a higher degree, can be derived. It is straightforward to note that since the essential matrix is the product of an antisymmetric matrix by a rotation matrix, its rank always equals two. Indeed, according to Huang and Faugeras [28], in order for a 3×3 matrix to be an essential one, a sufficient and necessary condition is that it is of rank two and its two nonzero singular values are equal. This last constraint is equal to

$$trace^2(\mathbf{E}\mathbf{E}^{\mathbf{t}}) - 2\,trace((\mathbf{E}\mathbf{E}^{\mathbf{t}})^2) = 0 \tag{11}$$

By employing Eq. (8), the following constraint which involves only the fundamental matrix **F** and the unknown intrinsic calibration matrix **A** is obtained:

$$trace^2(\mathbf{A}^{\mathbf{t}}\mathbf{F}\mathbf{A}\mathbf{A}\mathbf{F}^{\mathbf{t}}\mathbf{A}) - 2\,trace((\mathbf{A}^{\mathbf{t}}\mathbf{F}\mathbf{A}\mathbf{A}\mathbf{F}^{\mathbf{t}}\mathbf{A})^2) = 0 \tag{12}$$

Owing to the fact that the fundamental matrix is such that $det(\mathbf{F}) = 0$, the rank constraint for **E** is always verified. On the other hand, Eq. (12) is a polynomial constraint of order eight in the elements of matrix **A**, i.e. the intrinsic parameters. Thus, this constraint can be exploited for autocalibration: Assuming a mobile camera observing a certain number of characteristic points, each camera motion yields a fundamental matrix which gives rise to a polynomial constraint in the elements of $\mathbf{A}$. Therefore, it is theoretically possible to recover the intrinsic parameters of the camera with a sufficient number of camera movements.

3.2 The Constraints of Trivedi

Another way to establish a constraint between matrices **F** and **A** is to proceed in the manner proposed by Trivedi [47]. Specifically, he exploited the particular form of the symmetric matrix $\mathbf{E}\mathbf{E}^{\mathbf{t}}$, which depends only on the translation vector **t**, to derive the following set of independent polynomial equations:

$$\mathbf{E}\mathbf{E}^t = [\mathbf{t}]_\times \mathbf{R}\mathbf{R}^t([\mathbf{t}]_\times)^t = [\mathbf{t}]_\times([\mathbf{t}]_\times)^t \tag{13}$$

After some algebraic manipulation, the matrix $\mathbf{S} = \mathbf{E}\mathbf{E}^t$ is shown to have the following form:

$$\mathbf{S} = \begin{bmatrix} t_2^2 + t_3^2 & -t_1 t_2 & -t_1 t_3 \\ -t_2 t_1 & t_3^2 + t_1^2 & -t_2 t_3 \\ -t_3 t_1 & -t_3 t_2 & t_1^2 + t_2^2 \end{bmatrix}$$

Since the symmetric matrix $\mathbf{S}$ depends only on the three parameters of the translation vector **t**, its six elements must be constrained by the following equations:

$$4S_{ij}^2 - (trace(\mathbf{S}) - 2S_{ii}) * (trace(\mathbf{S}) - 2S_{jj}) = 0, \quad 1 \le i < j \le 3 \tag{14}$$

The above three constraints are polynomial and of degree four in the elements of **E** and thus of degree eight in the elements of **A**. Only two out of the three constraints are linearly independent. Those two constraints can be employed within an autocalibration framework, following a scenario similar to that delineated previously for the case of the Huang-Faugeras constraints. However, it is important to notice that in this case there are two equations for each camera movement, instead of a single one that was available in the previous case.

3.3 Discussion

Luong [34, 33] has demonstrated that the polynomial constraint expressed by Eq. (12) can be decomposed in two independent polynomial relations that are equivalent to the two equations of Trivedi given by Eq. (14). It can be argued that both the Huang-Faugeras and the Trivedi constraints share the disadvantage of being of degree eight in the elements of the intrinsic parameters matrix **A**. This is in contrast to the Kruppa equations, which as will become clear in section 4, are of degree two in the elements of matrix $\mathbf{K} = \mathbf{A}\mathbf{A}^t$, and therefore of degree four in the elements of the intrinsic parameters matrix **A**. Relying on numerical considerations associated with the practical difficulties of solving high order polynomial equations, researchers have invariably preferred the Kruppa equations for autocalibration.

There is, however, a very important reason for favoring the Huang-Faugeras and Trivedi constraints over the Kruppa equations. More specifically, the latter involve the fundamental matrix **F** which encodes the epipolar geometry between the two views considered as well as the epipole in the second image. It is well-known that the determination of the epipoles is very sensitive to measurement noise. Therefore, poorly estimated epipoles might significantly affect the accuracy of subsequent computations based on the Kruppa equations. On the other hand, the constraints of Eqs. (12) and (14) employ just the fundamental matrix **F** and not any other entities extracted from it, such as the epipoles. Thus, it may be advantageous in terms of numerical stability and robustness to noise to use them for autocalibration. It is the belief of the authors that these questions are worth posing and detailed experiments should be carried out to reach a definite conclusion.

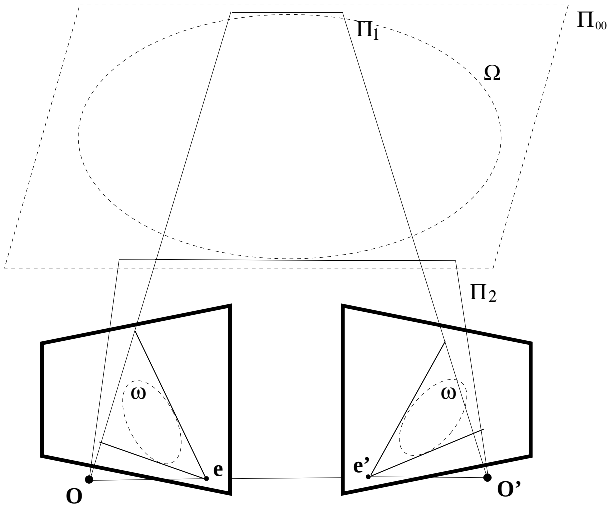


Figure 1: Geometric interpretation of the Kruppa equations; see text for details.

4 Deriving the Classical Kruppa Equations

In this section, the well-known Kruppa equations are presented. Although not needed in the remainder of the paper, we start by providing a geometric interpretation of the Kruppa equations for the sake of completeness. Then, the Kruppa equations are re-derived in a simple and purely algebraic manner, i.e. without making use of the absolute conic [17, 35] or the plane at infinity [51, 39]. Part of this derivation will be later employed for obtaining the simplified equations proposed in this paper.

4.1 Geometric Interpretation of the Kruppa Equations

Let Ω be the absolute conic, i.e. the virtual conic lying on the plane at infinity Π_∞ and having equation $x^2 + y^2 + z^2 = 0$ [16]. It is easy to see that each point $\mathbf{p}$ belonging to the projection ω of Ω in the second image satisfies $\mathbf{p}^t\mathbf{A}^{-t}\mathbf{A}^{-1}\mathbf{p} = 0$. Since Ω is invariant to rigid transformations, its image ω remains unchanged after a rigid motion of the camera, provided that the intrinsic calibration parameters do not change. As can be seen in Fig.1, there exist two planes Π_1 and Π_2 tangent to Ω and containing the baseline OO'. Thus, those two planes intersect the image planes at two pairs of corresponding epipolar lines which are tangent to ω. This fact is best expressed with the aid of duality: Conics are self-dual objects, i.e. the envelope of lines tangent to a conic is also a conic, known as the *dual* conic. In the case of the dual of the image of absolute conic ω^*, its equation is $\mathbf{l}^t\mathbf{A}\mathbf{A}^t\mathbf{l} = 0$, where $\mathbf{l}$ is a line tangent to ω. This implies that for a point $\mathbf{q}$ on any of the tangents to

ω in the second image, the following holds:

$$(\mathbf{e}' \times \mathbf{q})^t \mathbf{A}\mathbf{A}^t (\mathbf{e}' \times \mathbf{q}) \quad = \quad 0$$

The epipolar line $\mathbf{F}^t\mathbf{q}$ corresponding to $\mathbf{q}$ in the first image is also tangent to ω. Therefore, the invariance of ω under rigid transformations yields:

$$(\mathbf{F}^t\mathbf{q})^t \mathbf{A}\mathbf{A}^t (\mathbf{F}^t\mathbf{q}) \quad = \quad 0$$

The above two equations are thus equivalent to

$$\mathbf{F}\mathbf{A}\mathbf{A}^t\mathbf{F}^t \quad = \quad \beta\,([\mathbf{e}']_\times)^t \mathbf{A}\mathbf{A}^t ([\mathbf{e}']_\times), \tag{15}$$

where β is an arbitrary, nonzero scale factor. In other words, the Kruppa equations express the constraint that epipolar lines in the second image that correspond to epipolar lines of the first image that are tangent to ω, are also tangent to ω and vice versa.

4.2 An Algebraic Derivation

We start by using Eq.(6) to compute the epipole $\mathbf{e}^{'}$ in the second image. Given that $\mathbf{F}^t\mathbf{e}^{'} = 0$, $\mathbf{e}^{'}$ must satisfy

$$\mathbf{A}^{-t}\mathbf{R}^t([\mathbf{t}]_\times)^t\mathbf{A}^{-1}\mathbf{e}^{'} = 0 \tag{16}$$

Owing to the fact that $([\mathbf{t}]_\times)^t\mathbf{t} = 0$, the following solution for $\mathbf{e}^{'}$ is obtained:

$$\mathbf{e}^{'} = \lambda\mathbf{A}\mathbf{t}, \tag{17}$$

where λ is a nonzero scalar. This equation also supplies the following expression for **t**:

$$\mathbf{t} = \lambda'\mathbf{A}^{-1}\mathbf{e}^{'}, \tag{18}$$

where $\lambda' = 1/\lambda$. Equation (18) leads to the following relation for the matrix $[\mathbf{t}]_\times$[2]:

$$[\mathbf{t}]_\times = \lambda'\; det(\mathbf{A}^{-1})\mathbf{A}^t[\mathbf{e}^{'}]_\times\mathbf{A} \tag{19}$$

Substitution of this last relation into Eq.(10), yields directly the Kruppa equations:

$$\mathbf{F}\mathbf{K}\mathbf{F}^t = \gamma[\mathbf{e}^{'}]_\times\mathbf{K}([\mathbf{e}^{'}]_\times)^t, \tag{20}$$

[2] By making use of the equation $[\mathbf{M}\mathbf{u}]_\times = det(\mathbf{M})\mathbf{M}^*[\mathbf{u}]_\times\mathbf{M}^{-1}$, where $\mathbf{M}$ is a nonsingular matrix.

where γ is an unknown, nonzero scalar. Note that Eq.(20) is identical to Eq.(15) which was derived using a geometric argument. Since $\mathbf{FKF}^t$ is a symmetric matrix, Eq.(20) corresponds to the following equations obtained by eliminating γ:

$$\frac{\mathbf{FKF}^t_{11}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{11}} = \frac{\mathbf{FKF}^t_{12}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{12}} = \frac{\mathbf{FKF}^t_{22}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{22}} = \\ = \frac{\mathbf{FKF}^t_{13}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{13}} = \frac{\mathbf{FKF}^t_{23}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{23}} = \frac{\mathbf{FKF}^t_{33}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{33}} \tag{21}$$

These equations, however, are linearly dependent since

$$(\mathbf{FKF}^t - \gamma[\mathbf{e}^{'}]_\times \mathbf{K}([\mathbf{e}^{'}]_\times)^t)\mathbf{e}^{'} = \mathbf{0} \tag{22}$$

As shown in [51], there are only two independent equations among the set of the six equations given by Eq.(21). These equations are second order polynomials in the elements of $\mathbf{K}$, and therefore of order four in the elements of **A**. This is a significant order reduction compared to the Huang-Faugeras [28] or Trivedi [47] constraints which, as already mentioned, are of degree eight in the elements of **A**. When using the Kruppa equations for self-calibration, it is common to start by estimating $\mathbf{K}$ and then using Cholesky decomposition[3] [18] to obtain **A**.

At this point, it should be noted that the question of deciding which two equations out of the total six to use, remains open. Up to now, this problem has been resolved either by employing a specific parameterization of the epipolar geometry as in [17, 35, 33], or by randomly selecting one equation for estimating the scale factor and then substituting the result into two others that are arbitrarily chosen among the remaining five ones [52, 4]. The alternative of taking into account all equations simultaneously can also be considered, although numerical methods usually fail to produce a solution in the case of high order, over-determined polynomial systems such as this. In section 5, a simple answer to the above question is provided by an approach which directly leads to three linearly dependent equations, out of which two are linearly independent.

5 The Simplified Kruppa Equations

This section develops a simpler variant of the Kruppa equations. The principal motivation is twofold: First, to directly derive less equations than the six of the original formulation, so that the task of choosing the ones to employ for self-calibration becomes simpler. Second, to avoid employing the epipole $\mathbf{e}^{'}$, since its accurate estimation is difficult in the presence of noise and/or degenerate motions. Towards this end, the Singular Value Decomposition (SVD) [18] of the matrix **F** is employed:

$$\mathbf{F} = \mathbf{UDV}^t \tag{23}$$

[3] The Cholesky decomposition of a positive definite matrix **B**, is a matrix **C** s.t. $\mathbf{B} = \mathbf{C}^t\mathbf{C}$.

Recalling that **F** is of rank 2, the diagonal matrix **D** has the following form:

$$D = \begin{bmatrix} r & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where r and s are the eigenvalues of the matrix $\mathbf{F}\mathbf{F}^t$, whereas $\mathbf{U}$ and $\mathbf{V}$ are two orthogonal matrices. By making use of this relation, the epipole in the second image $\mathbf{e}^{'}$ can be deduced very simply. Specifically,

$$\mathbf{F}^t\mathbf{e}^{'} = \mathbf{V}\mathbf{D}^t\mathbf{U}^t\mathbf{e}^{'} = 0 \tag{24}$$

Since **D** is a diagonal matrix with a last element equal to zero, the following direct solution for $\mathbf{e}^{'}$ is obtained:

$$\mathbf{e}^{'} = \delta\mathbf{U}\mathbf{m}, \;\; \delta \neq 0 \tag{25}$$

with $\mathbf{m} = [0, 0, 1]^t$. Therefore, the matrix $[\mathbf{e}^{'}]_\times$ is equal to

$$[\mathbf{e}^{'}]_\times = \mu\mathbf{U}\mathbf{M}\mathbf{U}^t, \tag{26}$$

where μ is a nonzero scale factor and **M** = $[\mathbf{m}]_\times$ is given by:

$$\mathbf{M} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

By substituting Eq.(26) into Eq.(10), a new expression for the Kruppa equations is obtained:

$$\mathbf{F}\mathbf{K}\mathbf{F}^t = \mu\mathbf{U}\mathbf{M}\mathbf{U}^t\mathbf{K}\mathbf{U}\mathbf{M}^t\mathbf{U}^t \tag{27}$$

Since **U** is an orthogonal matrix, left and right multiplication of Eq.(27) by $\mathbf{U}^{\,t}$ and **U** respectively, yields the following notably simple expression for the Kruppa equations:

$$\mathbf{D}\mathbf{V}^t\mathbf{K}\mathbf{V}\mathbf{D}^t = \mu\mathbf{M}\mathbf{U}^t\mathbf{K}\mathbf{U}\mathbf{M}^t \tag{28}$$

Because of the simple forms of the matrices **D** and **M**, relation (28) corresponds to three linearly dependent equations. Indeed, denoting by $\mathbf{u_1}$, $\mathbf{u_2}$, $\mathbf{u_3}$ the column vectors of **U** and by $\mathbf{v_1}$, $\mathbf{v_2}$, $\mathbf{v_3}$ the column vectors of **V**, the matrix equation (28) is equivalent to

$$\mathbf{D}\mathbf{V}^t\mathbf{K}\mathbf{V}\mathbf{D}^t = \begin{bmatrix} r^2\mathbf{v_1}^t\mathbf{K}\mathbf{v_1} & rs\mathbf{v_1}^t\mathbf{K}\mathbf{v_2} & 0 \\ sr\mathbf{v_2}^t\mathbf{K}\mathbf{v_1} & s^2\mathbf{v_2}^t\mathbf{K}\mathbf{v_2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{M}\mathbf{U}^t\mathbf{K}\mathbf{U}\mathbf{M}^t = \begin{bmatrix} \mathbf{u_2}^t\mathbf{K}\mathbf{u_2} & -\mathbf{u_2}^t\mathbf{K}\mathbf{u_1} & 0 \\ -\mathbf{u_1}^t\mathbf{K}\mathbf{u_2} & \mathbf{u_1}^t\mathbf{K}\mathbf{u_1} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The above expressions finally yield the following three linearly dependent equations for the matrix **K**:

$$\frac{r^2\mathbf{v_1}^t\mathbf{K}\mathbf{v_1}}{\mathbf{u_2}^t\mathbf{K}\mathbf{u_2}} = \frac{rs\mathbf{v_1}^t\mathbf{K}\mathbf{v_2}}{-\mathbf{u_2}^t\mathbf{K}\mathbf{u_1}} = \frac{s^2\mathbf{v_2}^t\mathbf{K}\mathbf{v_2}}{\mathbf{u_1}^t\mathbf{K}\mathbf{u_1}} \tag{29}$$

Only two of these three equations are linearly independent. They are the simplified Kruppa equations, derived in a particularly straightforward manner. Moreover, the use of the SVD has enabled us to deduce automatically which three out of the six equations present in the original formulation should be taken into account. Notice that in the case of a calibrated camera, the intrinsic calibration matrix can be assumed to be equal to the 3×3 identity matrix. Therefore, considering the essential matrix **E** instead of the fundamental matrix **F**, it is straightforward to show that Eqs.(29) reduce to

$$r^2 = \frac{0}{0} = s^2,$$

which implies that $r = s$, as shown by Huang and Faugeras [28]; and mentioned already in section 3.

It is worth noting that equations (29) are closely related to the generalization of the Kruppa equations that has been proposed by Luong [34]. Luong has used these equations for demonstrating the equivalence between the constraints of Trivedi [47] and those of Huang and Faugeras [28]. The same ideas can also be found at the origin of the recent article by Hartley [20], who directly derives the Kruppa equations from the fundamental matrix. Both Luong and Hartley base their developments on changes of rather astute reference points, which amount to generalizing the Kruppa equations by considering that the absolute conic can have two different images in each retina. In this work, a different approach is taken which remarkably simplifies the task of self-calibration. Using an algebraic method, a simpler variant of the Kruppa equations is derived without making use of the absolute conic. Subsequent sections describe experimental results from the calibration of synthetic as well as real image sequences with the aid the proposed method. To the best of our knowledge, this is the first time that detailed experimental results obtained from the application of the simplified Kruppa equations to imagery are being reported.

6 Self-Calibration Using the Simplified Kruppa Equations

In this section, an algorithm applying the simplified Kruppa equations to the problem of self-calibration is presented and related implementation issues are clarified. Following the approach of Zeller [51, 52], the equations derived in section 5 are embedded in a non-linear optimization framework and solved iteratively. We begin with a discussion regarding the choice of an appropriate initial solution that forms the starting point for the optimization stage. We then formulate the optimization problem and explain how it is solved to obtain the intrinsic calibration parameters.

6.1 Finding an Initial Solution

Let $\mathbf{S_F} = [r, s, \mathbf{u_1}^t, \mathbf{u_2}^t, \mathbf{u_3}^t, \mathbf{v_1}^t, \mathbf{v_2}^t, \mathbf{v_3}^t]^t$ be the 20×1 vector formed by the parameters of the SVD of $\mathbf{F}$. Let also $\frac{\rho_i(\mathbf{S_F},\mathbf{K})}{\phi_i(\mathbf{S_F},\mathbf{K})},\ i = 1 \ldots 3$ be the three ratios defined by Eq.(29). Each pair of images defines a fundamental matrix, which in turn yields the following two polynomial equations regarding the elements of $\mathbf{K}$:

$$\begin{aligned} \rho_1(\mathbf{S_F},\mathbf{K})\phi_2(\mathbf{S_F},\mathbf{K}) - \phi_1(\mathbf{S_F},\mathbf{K})\rho_2(\mathbf{S_F},\mathbf{K}) &= 0 \\ \rho_1(\mathbf{S_F},\mathbf{K})\phi_3(\mathbf{S_F},\mathbf{K}) - \phi_1(\mathbf{S_F},\mathbf{K})\rho_3(\mathbf{S_F},\mathbf{K}) &= 0 \end{aligned} \tag{30}$$

The above system of equations is of degree two in five unknowns defining the **K** matrix. A good initial approximation regarding the position of the principal point in the image, is to assume that it coincides with the image center. Additionally, if the skew angle θ is assumed to be equal to $\frac{\pi}{2}$, the number of unknowns in Eq.(30) is reduced to two, namely elements K_{11} and K_{22} of the **K** matrix which are related to the two focal lengths α_u and α_v. Therefore, the system of equations (30) becomes of degree two in two unknowns and thus it can be solved analytically. The system can have at most $2^2 = 4$ solutions, some of which might be meaningless. More specifically, every solution for K_{11} and K_{22} which is such that the associated **K** matrix is not real and positive definite, is discarded. Solutions are also discarded in the case that the related aspect ratio is very far from unity. Assuming the availability of M images that have been acquired with constant camera intrinsic parameters, a total of $N \leq \frac{M(M-1)}{2}$ fundamental matrices can be defined. These matrices give rise to N second order systems of the form of Eqs.(30) that have at most $4N$ solutions for the two focal lengths α_u and α_v. The following strategies for choosing among the available initial solutions have been examined:

- Use one of the solutions in random.
- Use the average of the available solutions.
- Use the median of the available solutions.

Although these strategies can produce considerably different starting points, our experiments have indicated that the choice of an initialization strategy from the above set is not crucial for the convergence of the non-linear optimization algorithm. In other words, the latter has generated very similar final results, starting from different starting points.

6.2 Non-Linear Optimization

Let $\pi_{ij}(\mathbf{S_F}, \mathbf{K})$ denote the difference of ratios $\frac{\rho_i(\mathbf{S_F},\mathbf{K})}{\phi_i(\mathbf{S_F},\mathbf{K})} - \frac{\rho_j(\mathbf{S_F},\mathbf{K})}{\phi_j(\mathbf{S_F},\mathbf{K})}$ and let $\sigma^2_{\pi_{ij}}(\mathbf{S_F},\ \mathbf{K})$ be its variance. This variance is approximated by[4]

$$\sigma^2_{\pi_{ij}}(\mathbf{S_F},\ \mathbf{K}) = \frac{\partial \pi_{ij}(\mathbf{S_F}, \mathbf{K})}{\partial \mathbf{S_F}} \Lambda_{\mathbf{S_F}} \frac{\partial \pi_{ij}(\mathbf{S_F}, \mathbf{K})}{\partial \mathbf{S_F}}^t , \tag{31}$$

where $\Lambda_{\mathbf{S_F}}$ is the 20×20 covariance matrix associated with $\mathbf{S_F}$ and $\frac{\partial \pi_{ij}(\mathbf{S_F},\mathbf{K})}{\partial \mathbf{S_F}}$ is the derivative of $\pi_{ij}(\mathbf{S_F}, \mathbf{K})$ at $\mathbf{S_F}$. Since $\mathbf{S_F}$ is a function of $\mathbf{F}$, its covariance matrix $\Lambda_{\mathbf{S_F}}$ is in turn computed from

$$\Lambda_{\mathbf{S_F}} = \frac{\partial \mathbf{S_F}}{\partial \mathbf{F}} \Lambda_{\mathbf{F}} \frac{\partial \mathbf{S_F}}{\partial \mathbf{F}}^t , \tag{32}$$

where $\Lambda_{\mathbf{F}}$ is the 9×9 covariance matrix of the fundamental matrix [5] and $\frac{\partial \mathbf{S_F}}{\partial \mathbf{F}}$ is the value of the jacobian of $\mathbf{S_F}$ at $\mathbf{F}$. This last step, i.e. the computation of the derivatives of the SVD components of a matrix with respect to that matrix, is explained in more detail in Appendix B. As will soon be clear, the variances $\sigma^2_{\pi_{ij}}(\mathbf{S_F},\ \mathbf{K})$ are used to automatically weight the residuals $\pi_{ij}(\mathbf{S_F}, \mathbf{K})$ according to their uncertainty. It should be noted, however, that the computation of these variances is achieved by two successive first order approximations, i.e. Eqs.(31) and (32). Therefore, this two step process might be more noise sensitive compared to a more direct computation of the variances. Alternative ways of computing better estimates of $\sigma^2_{\pi_{ij}}(\mathbf{S_F},\ \mathbf{K})$ are currently under consideration.

Matrix $\mathbf{K}$ is computed as the solution of a non-linear least squares problem, namely

$$\mathbf{K} = argmin_{\tilde{\mathbf{K}}} \sum_{i=1}^{N} \frac{\pi^2_{12}(\mathbf{S}_{\mathbf{F}_i}, \tilde{\mathbf{K}})}{\sigma^2_{\pi_{12}}(\mathbf{S}_{\mathbf{F}_i}, \tilde{\mathbf{K}})} + \frac{\pi^2_{13}(\mathbf{S}_{\mathbf{F}_i}, \tilde{\mathbf{K}})}{\sigma^2_{\pi_{13}}(\mathbf{S}_{\mathbf{F}_i}, \tilde{\mathbf{K}})} + \frac{\pi^2_{23}(\mathbf{S}_{\mathbf{F}_i}, \tilde{\mathbf{K}})}{\sigma^2_{\pi_{23}}(\mathbf{S}_{\mathbf{F}_i}, \tilde{\mathbf{K}})} \tag{33}$$

Recalling that each fundamental matrix yields two independent equations and $\mathbf{K}$ consists of five unknowns, the minimum number of required view pairs (i.e. N) is in the general case equal to three. Additional constraints provided by more than three fundamental matrices can improve the accuracy of the solution. The reason for minimizing the sum of the squared ratio differences $\pi_{ij}(\mathbf{S_F}, \mathbf{K})$ in Eq.(33) instead of the sum of the squared polynomials of Eqs.(30), is that the former formulation has proved to produce more accurate results. Slightly better quality results is also the reason for including the third simplified Kruppa equation (i.e. $\pi_{23}(\mathbf{S_F}, \mathbf{K})$) in Eq.(33), although it is dependent on the other two [33]. The minimization of Eq.(33) is done using a classical Levenberg-Marquardt algorithm [2], using the starting solution computed in the initialization stage. Apart from $\mathbf{K}$ itself, the minimization in Eq.(33) can also provide its associated covariance matrix [2]. In the case that a priori information in the form of angles or ratios of line segments in the scene is available, it can

[4] Assuming that $\mathbf{x}$ is a random vector with mean $\mathbf{x_0}$ and covariance $\Lambda_{\mathbf{x}}$, the covariance of vector $\mathbf{y} = \mathbf{f}(\mathbf{x})$ up to first order is equal to $\Lambda_{\mathbf{y}} = \frac{\partial \mathbf{f}(\mathbf{x_0})}{\partial \mathbf{x_0}} \Lambda_{\mathbf{x}} \frac{\partial \mathbf{f}(\mathbf{x_0})}{\partial \mathbf{x_0}}^t$; see [14] for details and proof.

[5] This covariance matrix is supplied as a by-product of the procedure for estimating $\mathbf{F}$ [7].

be incorporated in Eq.(33) as described in [51]. The matrix **A** is extracted from **K** in three steps. First, $\mathbf{A}^{-t}$ is computed by employing the Cholesky decomposition of $\mathbf{K}^{-1}$, then it is transposed and finally inverted to yield **A**.

7 Experimental Results

The proposed calibration method has been extensively validated with the aid of both synthetic and real image sequences. Representative results from several of these experiments are given in this section. To demonstrate the accuracy of the recovered calibration, the estimated **A** matrices have been employed to measure 3D angles and 3D length ratios from corresponding image line segments (see also appendix A). This computation, in addition to the calibration matrix **A**, requires the homography of the plane at infinity $\mathbf{H}_\infty$ to be known. To calculate $\mathbf{H}_\infty$, the essential matrix is first computed from the fundamental matrix using Eq.(8). Then, the essential matrix is decomposed using a linear method into a rotation matrix **R** and a translation vector **t**, such that $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$. More details concerning this decomposition can be found for example in [50, 19]. Finally, $\mathbf{H}_\infty$ is computed from **R** using Eq.(36). More details are available in appendix A.

7.1 Synthetic Experiments

To quantitatively study the effects of increasing amounts of noise on the recovered intrinsic calibration parameters as well as the Euclidean entities measured using them, a set of experiments using simulated data has been carried out. More specifically, a simulator has been constructed, which given appropriate values for the camera intrinsic parameters and the camera translational and rotational motion, simulates a series of rigid displacements of the camera and projects a set of randomly chosen 3D points on the simulated retina. Zero mean Gaussian noise is then added to the resulting retinal points, to account for the fact that in practice, feature extraction algorithms introduce some error when locating image interest points (i.e. corners). The experimental procedure and the related parameter values for the particular experiments reported here are as follows: The simulated retina is 640×480 pixels, the principal point is located at pixel (310, 270), the angle between the retinal axes is $\pi/2$ and the focal lengths are 840 and 770 in horizontal and vertical pixel units respectively. This implies that the field of view of the simulated camera roughly subtends 42 and 34 degrees in the horizontal and vertical directions respectively. After simulating a series of rigid displacements of the camera, a number of random 3D points are projected on the simulated retina. A non-linear method [54] is then employed to estimate from the noisy retinal points the fundamental matrices corresponding to the simulated displacements. The estimates of the fundamental matrices serve as the input to self-calibration. Subsequent subsections describe in detail the types of the experiments performed.

translation			rotation			rotation angle (deg)	epipole in the second view ($\mathbf{e}'$)	
320	-215	170	0.554	-0.832	0.028	8.0	1891.18	-703.824
550	755	125	0.707	0.707	0.035	9.0	4006	4920.8
650	655	150	-0.667	-0.333	-0.667	7.5	3950	3632.33

Table 1: The simulated 3D rigid motions and the retinal locations of their corresponding epipoles.

In the sequel, **ZF** will denote the self-calibration algorithm proposed by Zeller and Faugeras in [52], while **SVD** will refer to the algorithm proposed in this paper. The **ZF** algorithm is similar in spirit with **SVD** , with the major difference being that it employs the classical Kruppa equations (20) instead of the simplified ones. More specifically, in **ZF** two of the constraints (21) are arbitrarily chosen for each available pair of views and then the intrinsic parameters are estimated with the aid of a non-linear bundle adjustment algorithm. The covariances of the fundamental matrix estimates are used to compute weights for each constraint, in a manner similar to Eq.(33). The starting point for the bundle adjustment is found by assuming that the skew angle is equal to $\pi/2$ and the principal point is roughly at the image center and then analytically solving Eqs.(21) for the two focal lengths.

7.1.1 Recovering Intrinsic Calibration Under Varying Noise Levels

In this section, the behavior of the **SVD** algorithm in the case that its input is contaminated by different amounts of noise is studied. Moreover, the results produced by **SVD** are compared with these recovered by the **ZF** algorithm. In the series of experiments reported here, three camera motions and 300 random 3D points have been simulated. The exact motion parameters are shown in table 1. Notice that the simulated motions constitute a challenging sequence, since the translational components of motion along the optical (i.e. OZ) axis are much smaller compared to the components that are parallel to the retinal plane. This implies that the epipoles are outside the images, making their accurate computation very difficult [42]. The depths of the 3D points were uniformly distributed in the range [20, 100], measured in focal length units. The standard deviation of the Gaussian noise added to the retinal projections of the simulated 3D points ranged from 0 to 4.0 pixels. To ensure that the recovered intrinsic calibration parameters are independent of the exact location of the 3D points used to form 2D correspondences, each experiment was run 100 times, each time using a different random set of 3D points drawn from the uniform distribution described above. Each time, all six fundamental matrices defined by the three displacements were estimated. Then, the **SVD** and **ZF** algorithms were applied to recover the intrinsic calibration. Both algorithms were applied first by assuming an unknown skew angle (i.e. 5 intrinsic parameters) and then by assuming a skew angle equal to $\pi/2$ (i.e. 4 intrinsic parameters). The results of these experiments are shown in the graphs labeled “SVD with 5p” and “SVD with 4p” respectively for the **SVD** algorithm and “ZF with

5p” and “ZF with 4p” for the **ZF** algorithm. Both algorithms were also applied with and without considering the covariances for weighting each constraint. When the covariances are not employed, the weights $\sigma_{\pi_{ij}}(\mathbf{S_F},\ \mathbf{K})$ in Eq.(33) are all assumed to be equal to one. Figures 2 and 3 illustrate for all noise levels considered, the mean and standard deviation of the relative error in the recovered intrinsic parameters. Each point in the plots summarizes the error statistics computed from 100 runs. In the case that the covariances are taken into consideration, the results are shown in Figures 4 and 5 using labels following the naming conventions of Figs. 2 and 3.

As can be seen from Figs. 2, 3, 4 and 5, both methods perform roughly the same when the noise added to corresponding image points is small, i.e. its standard deviation is less than 1.5 pixels. For larger amounts of noise, however, it is evident that the **SVD** algorithm performs better than **ZF** , in both 5 and 4 intrinsic parameter experiments. Note that points corresponding to very large error values are not shown in the graphs. It is important to mention that almost always, both algorithms managed to converge to a solution. As expected, the difference in the accuracy of the two methods is larger in the case of minimizing over 5 instead of 4 parameters. Even in cases when the two methods yield similar mean errors, the error standard deviations are typically smaller for the **SVD** algorithm. Thus, the **SVD** algorithm is more stable with respect to noise, and therefore should be preferred over **ZF** since in practice there is no a priori information available regarding the amount of noise contaminating the employed image points. As already observed in other works, i.e. [52, 4], the focal lengths $\alpha_u,\ \alpha_v$ are recovered more accurately than the principal point $u_0,\ v_0$. It is also known, however, that the latter does not have to be accurately estimated in order to reconstruct metric information of satisfactory quality [45, 4].

Figures 4 and 5 give a quantitative assessment of the benefits gained by employing the covariance of the estimated fundamental matrices for computing the weights for the constraints considered. Clearly, using the covariances generally reduces the error in the computed estimates, although in the case of large amounts of noise, the inclusion of covariances can cause the **ZF** algorithm to oscillate wildly among large error values that are excluded from the plots. On the other hand, in simpler simulation scenarios (when, for example, the epipoles are within the simulated images), that are not reported here due to space considerations, it has been observed that the accuracy improvements gained by employing the covariance in the case of the **ZF** algorithm are larger compared to those for the **SVD** algorithm. This might be an indication that the method currently employed for estimating the constraint weights in the case of **SVD** is not accurate enough (see also section 6.2). Hence, the question regarding which is the best method for estimating these weights remains open and further investigation is needed in this direction.

7.1.2 3D Motion Recovery Using Noisy Intrinsic Calibration Parameters

The current section examines the error induced to the 3D motion estimates that are computed using the intrinsic calibration matrices recovered in subsection 7.1.1. Towards this, the recovered intrinsic calibration matrices are first combined with Eq.(8) to compute estimates of the essential matrix. 3D

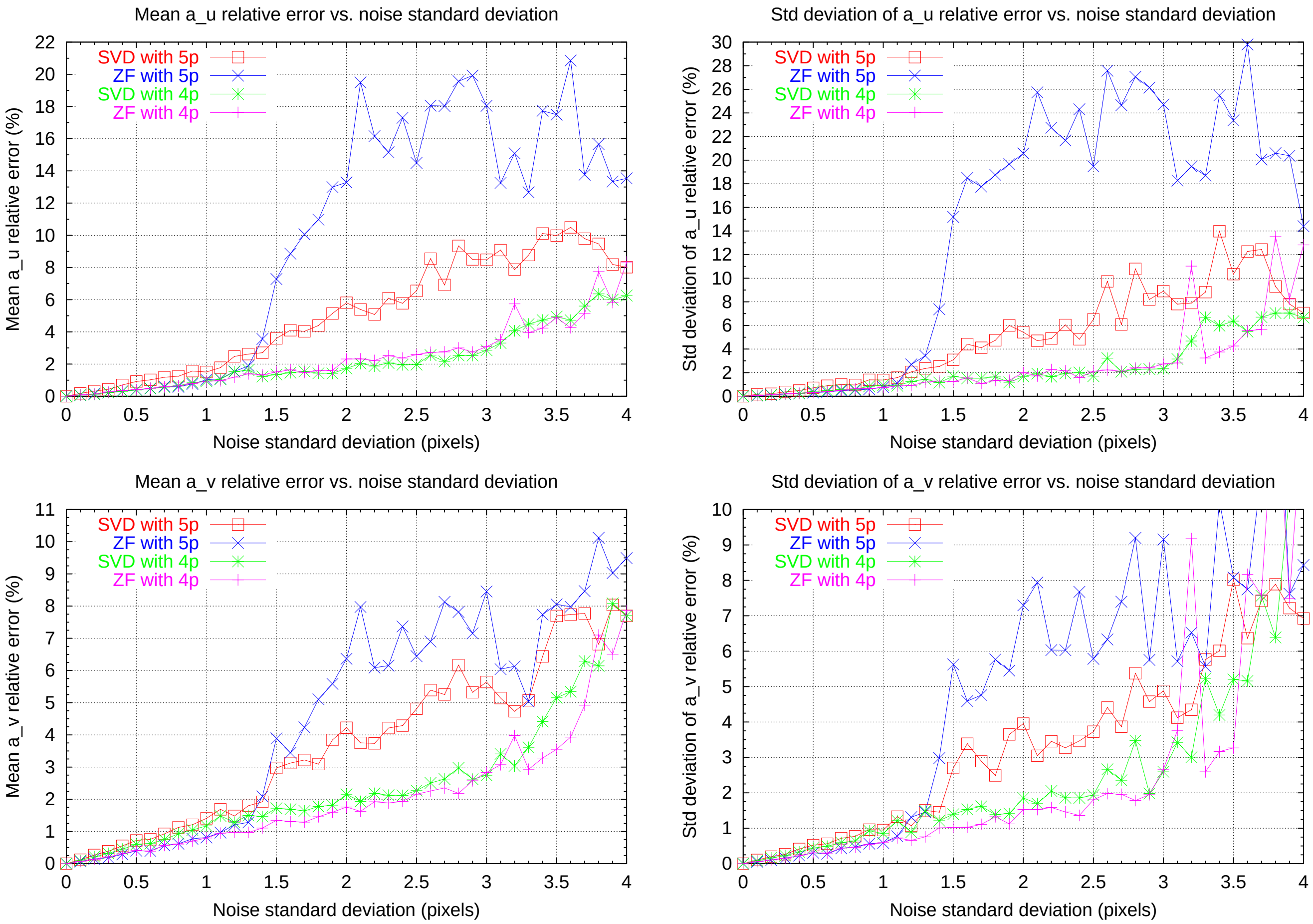


Figure 2: The error in the recovered focal lengths in the presence of noise, without employing the covariance. The first row corresponds to α_u, the second to α_v. Mean values are shown in the left column, standard deviations in the right.

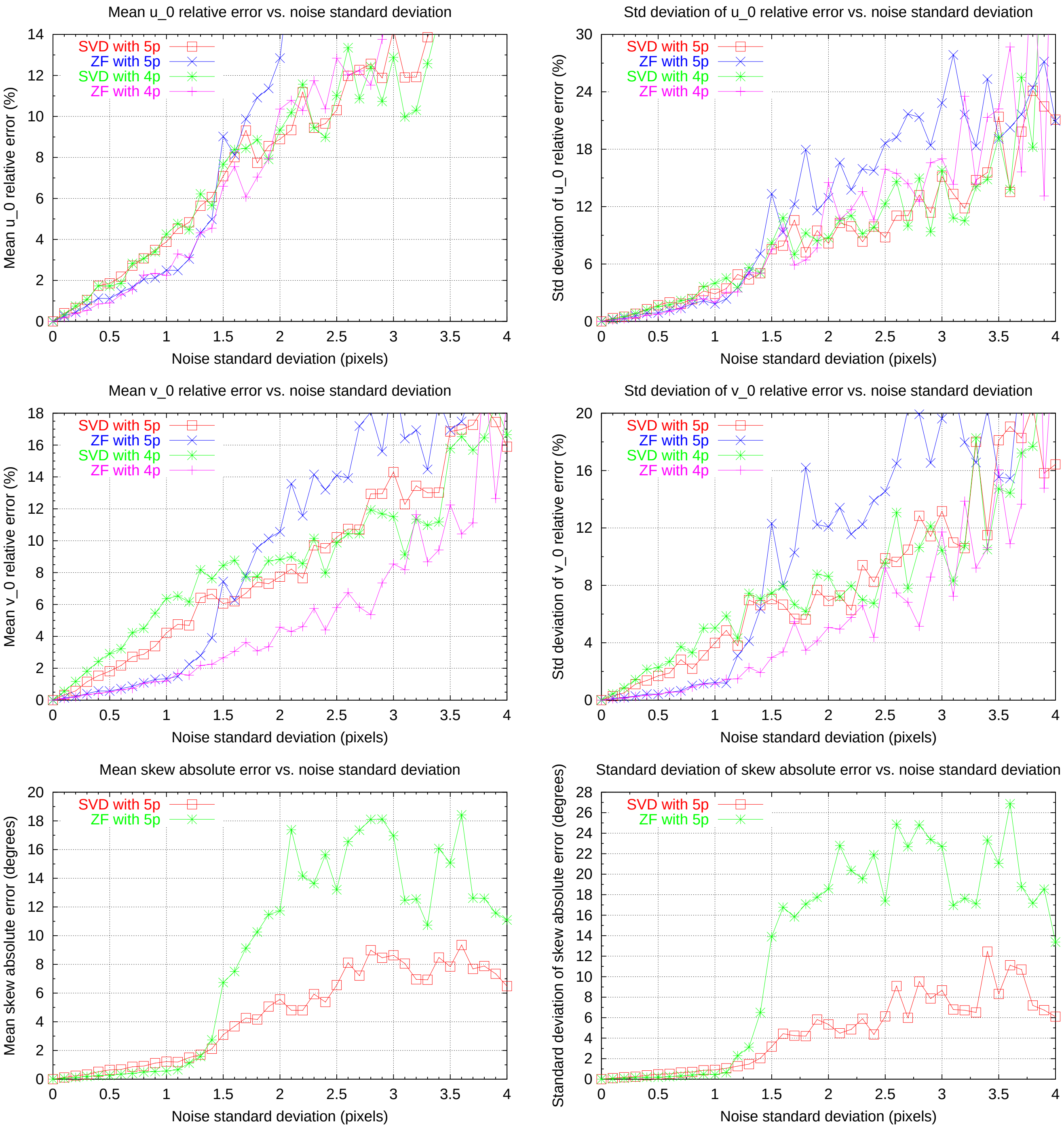


Figure 3: The error in the recovered principal points and skew angles in the presence of noise, when the covariance is not employed. The top row corresponds to u_0, the middle one to v_0 and the bottom one to the skew angle θ. Mean values are shown in the left column, standard deviations in the right.

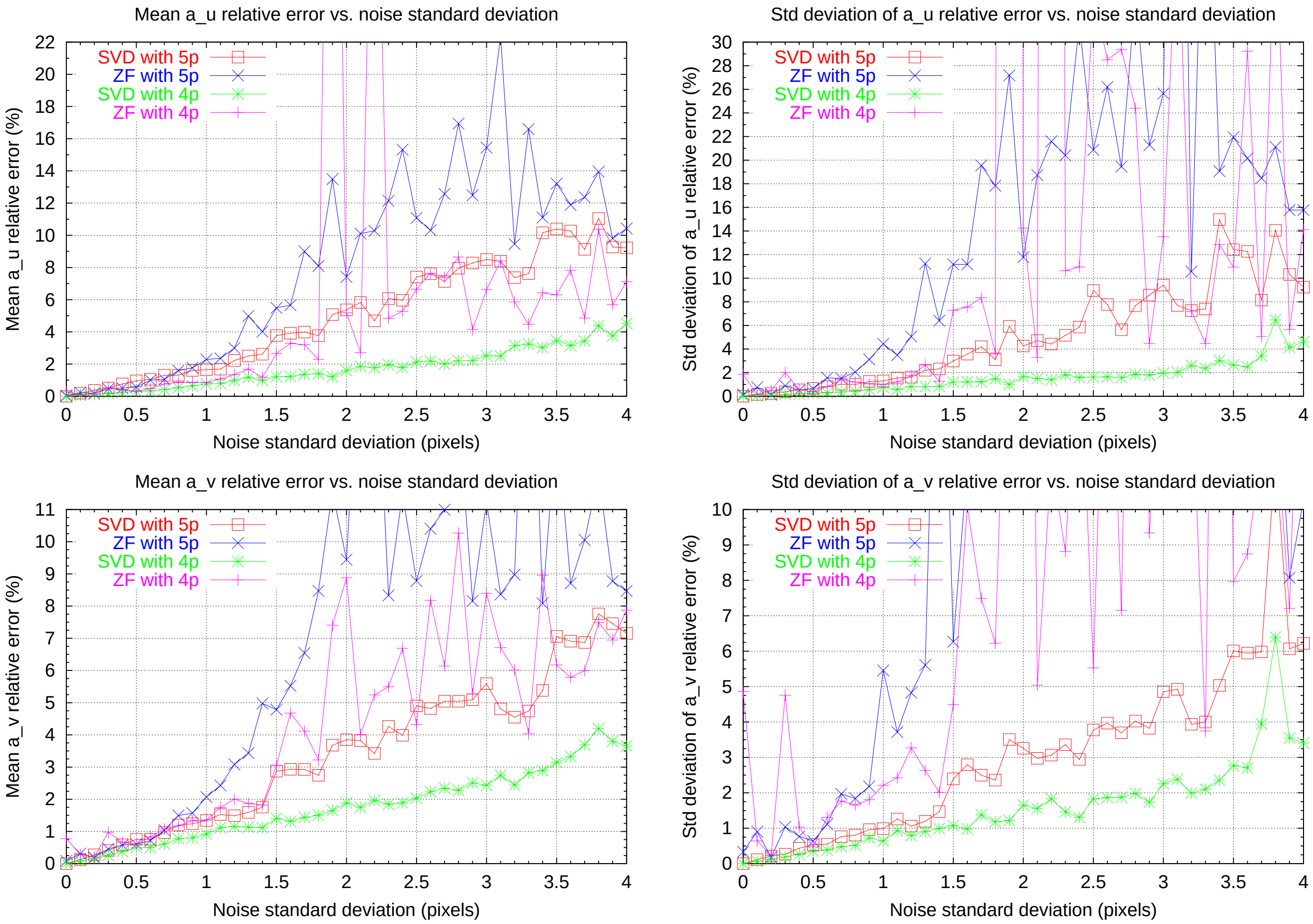


Figure 4: The error in the recovered focal lengths in the presence of noise when employing the covariance. The first row corresponds to α_u, the second to α_v. Mean values are shown in the left column, standard deviations in the right.

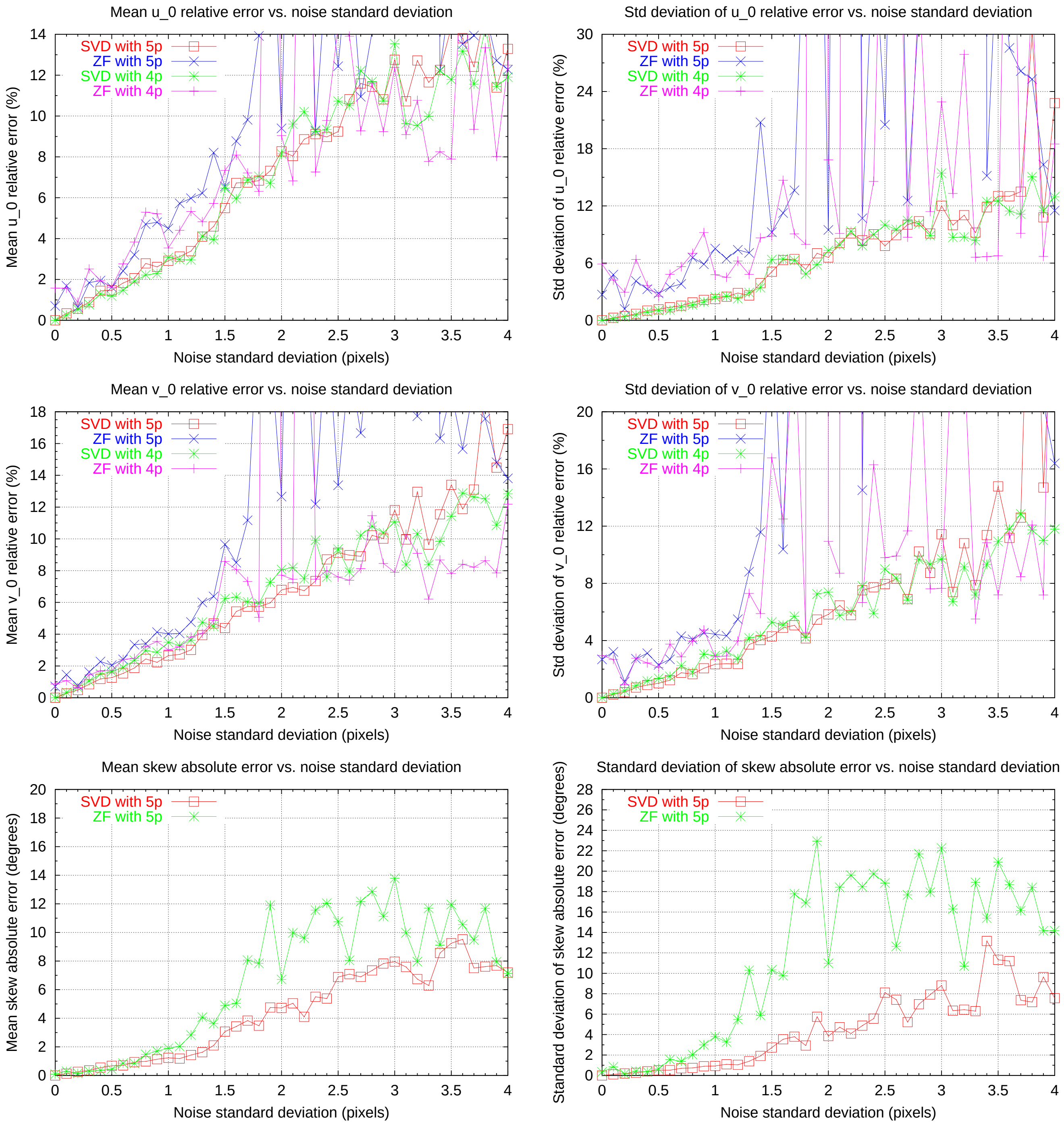


Figure 5: The error in the recovered principal points and skew angles in the presence of noise when employing the covariance. The top row corresponds to u_0, the middle one to v_0 and the bottom one to the skew angle θ. Mean values are shown in the left column, standard deviations in the right.

motion estimation using the essential matrix is a classical problem in computer vision. A detailed review of available methods is beyond the scope of this paper; the interested reader is referred to [29] and the references therein. In this work, the 3D motion is extracted from the estimates of the essential matrix using the technique proposed by Zhang [53]. The main idea behind this work is to estimate the 3D motion by gradually projecting the essential matrix parameters from a high dimensional space to lower dimensional ones. More specifically, the method starts by estimating an eight parameter essential matrix using a classical linear algorithm [32], then enforcing the rank-2 constraint on the resulting matrix for reducing the number of parameters to be estimated to seven and finally estimating the five motion parameters corresponding to the essential matrix which minimizes the reprojection error in the images. Due to the depth/scale ambiguity, translation can be estimated up to an unknown scale factor, i.e. only its direction can be recovered [14].

In the set of experiments described here, we attempted to recover the motion shown in the first row of table 1 with the aid of the estimated intrinsic calibration matrices. Figure 6 shows the mean and standard deviation of the error in the recovered direction of translation, the error in the direction of the recovered axis of rotation and the relative error in the recovered rotation angle. First, 3D motion was estimated by using the intrinsic calibration matrices computed in the experiments outlined in Figures 2 and 3, i.e. without employing the covariance of the estimated fundamental matrices. More specifically, the graphs labeled “SVD with 5p” and “SVD with 4p”, correspond to the error in the 3D motion that is estimated using the intrinsic parameter matrices (i.e. **A**) computed by the **SVD** algorithm. Similarly, when **A** is computed by the **ZF** algorithm, the graphs are labeled “ZF with 5p” and “ZF with 4p”. The graph labeled “True A” in Figure 6 also illustrates the mean and standard deviation of the error in the estimated 3D motion when the latter is recovered using the ground truth values for the intrinsic calibration. Analogous error statistics in the case that 3D motion is estimated using the intrinsic calibration parameters that result when covariances are taken into account, are shown in Figure 7. The naming conventions for the graph labels are as in Fig. 6. As before, very large error values are not included in the plots.

It is clear from Figs. 6 and 7 that the estimates of the translational component are highly erroneous. This, however, is an inherent difficulty associated with the simulated motion, since its direction is mainly parallel to the retinal plane. It is well known that this situation is very challenging for accurate motion recovery [8, 38]. That the main source of error in the translation estimates is not the error in the estimated intrinsic calibration, is evident from the fact that the translation estimates computed using the correct intrinsic parameters (plots labeled “True A”) are also very wrong and not much better that those computed with the estimated intrinsic calibration parameters. The rotation direction, on the other hand, is recovered pretty accurately, provided that the estimate of the **A** matrix used to recover it is of acceptable quality. The rotation angle estimates are less accurate, especially if they have been recovered using **A** matrices parameterized with 5 unknowns. At this point, however, it should be mentioned that in the case of less difficult simulated motions that are not shown here, the 3D motion parameters have been recovered with very good accuracy.

As expected, the use of **A** matrices estimated when employing the covariance leads to slightly better 3D motion estimates, since the former are more accurate. Nevertheless, the 3D motion estimates recovered are very similar, even when employing estimates of the $\mathbf{A}$ matrix computed by **ZF** which in certain cases are considerably worse than those computed by **SVD** . This is in agreement with the findings of previous works, e.g. [33, 45], which have concluded that even rough estimates of the intrinsic calibration parameters are sufficient for acceptable 3D motion estimation. This is because the primary source of error in 3D motion estimation is not the inaccuracy of intrinsic calibration but rather the error in the employed retinal displacements [8].

7.1.3 When is SVD Definitely Preferable Than ZF : A Case Study

This subsection is devoted to the study of a particular sequence of 3D motions which corresponds to a singular case for **ZF** , i.e. the latter fails to produce a solution. More specifically, three motions have been simulated with the third being the composition of the first two ones. The exact values used for translation and rotation can be found in table 2. Note that motions 1 and 2 have only one nonzero translational component, i.e. along the OX axis. The intrinsic calibration parameters were again $\alpha_u = 840,\ \alpha_v = 770,\ u_0 = 310,\ v_0 = 270,\ \theta = \pi/2$. No noise was added to the simulated 2D points. The **SVD** algorithm recovered the intrinsic calibration shown in the bottom part of table 2. On the contrary, **ZF** did not produce a solution. The reason for this failure is explained in the next paragraph.

As can be verified from the analytic expression for the fundamental matrix (i.e. Eq.(6)), a 3D translation of the form $\mathbf{t} = [t_x, 0, 0]^t$ yields a fundamental matrix with a zero first row. This is indeed the case for the fundamental matrices that were estimated for motions 1 and 2 and are shown in table 2. Moreover, it is straightforward to prove that in the case of a translation $\mathbf{t} = [t_{\,x}, 0, 0]^t$, the antisymmetric matrix of the epipole $[\mathbf{e}^{'}]_\times$ also has a zero first row. This implies that both $\mathbf{FKF}$ and $[\mathbf{e}^{'}]_\times \mathbf{K}([\mathbf{e}^{'}]_\times)^t$ matrices in the classical Kruppa equations formulation (i.e. Eq.(20)) also have a zero first row. The particular implementation of the **ZF** algorithm that we employed in our experiments, arbitrarily chooses the two constraints that are given by the top row of Eq.(21), i.e.

$$\frac{\mathbf{FKF}^t_{11}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{11}} - \frac{\mathbf{FKF}^t_{12}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{12}} = 0$$

$$\frac{\mathbf{FKF}^t_{11}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{11}} - \frac{\mathbf{FKF}^t_{22}}{([\mathbf{e}']_\times \mathbf{K}([\mathbf{e}']_\times)^t)_{22}} = 0$$

In the case of motions 1 and 2, however, those two constrains are meaningless. This is due to the fact that the first rows of the matrices involved are zero, which makes the coefficients of the unknown elements of matrix **K** zero as well. On the contrary, the **SVD** algorithm does not run into any difficulties arising from the equations derived from motions 1 and 2. As can be easily checked, for both motions the singular values r and s as well as the vectors $\mathbf{u_1}$ and $\mathbf{u_2}$ computed from the Singular Value Decomposition of the corresponding fundamental matrices, are nonzero. Thus, the

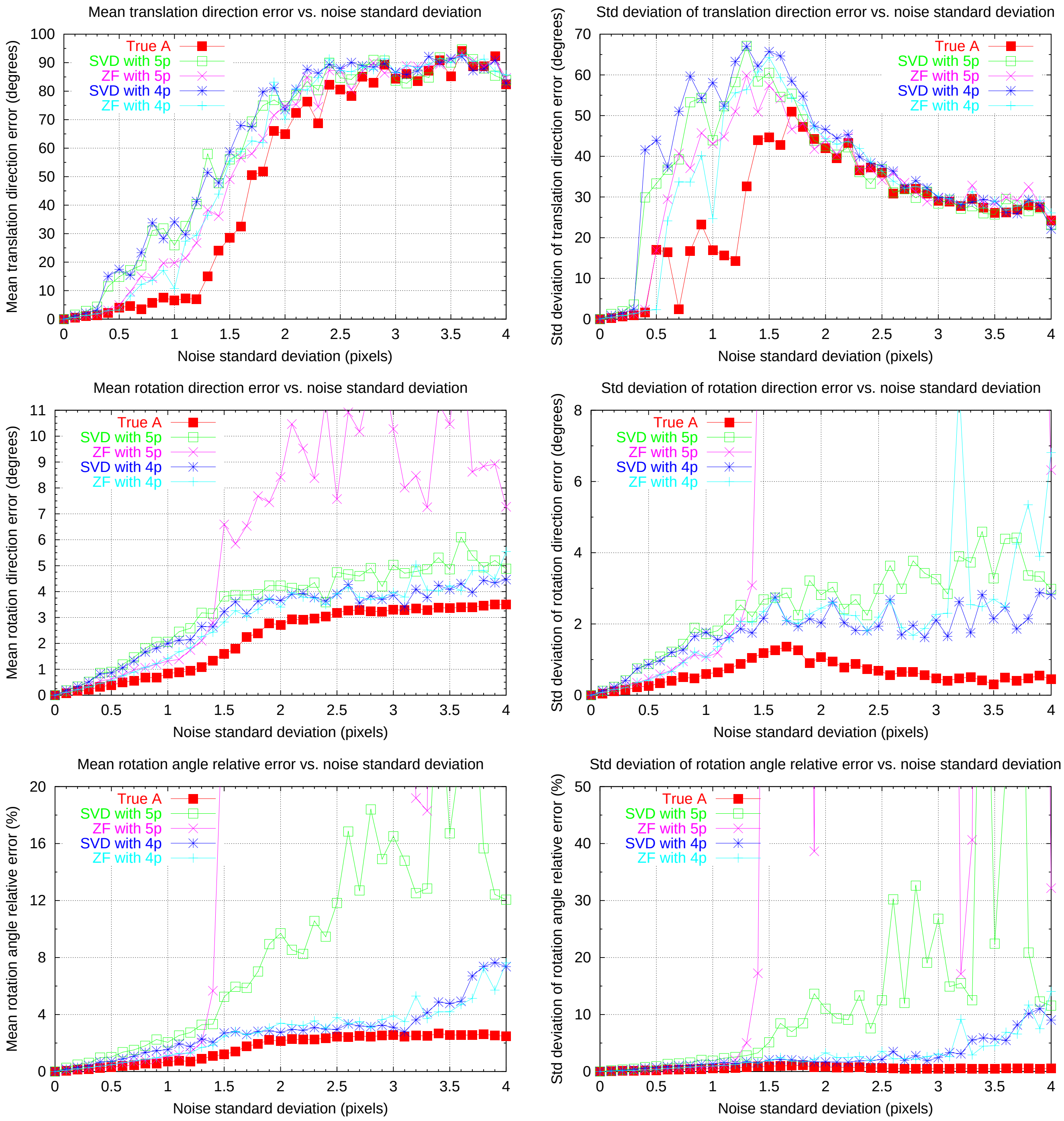


Figure 6: The error in the recovered 3D motion in the presence of noise, when the covariance is not employed. The top row shows the error in the recovered translation direction, the middle row corresponds to the error in the recovered rotation axis and the bottom row shows the error in the estimated rotation angle. Mean values are shown in the left column, standard deviations in the right.

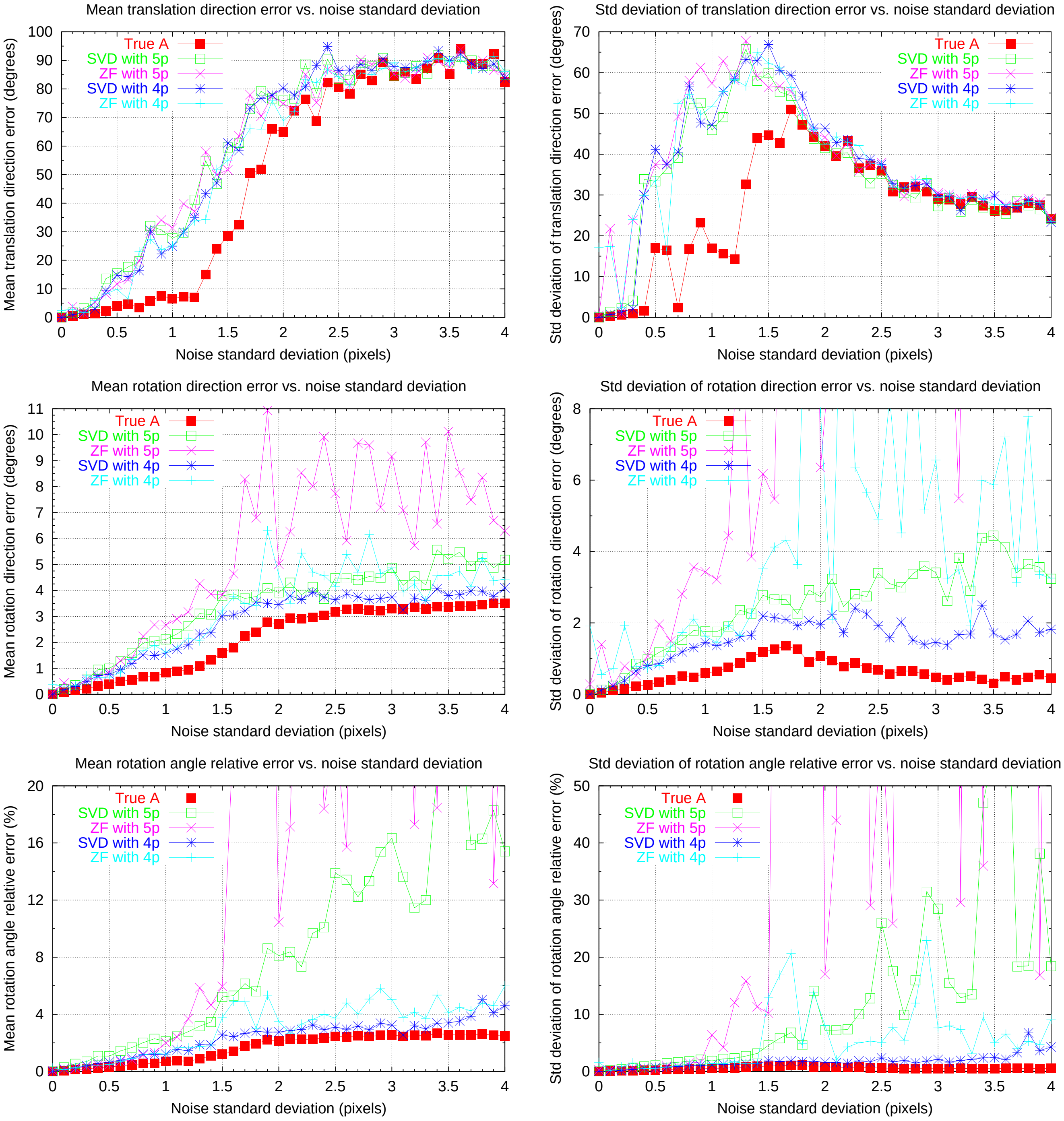


Figure 7: The error in the recovered 3D motion in the presence of noise when employing the covariance. The top row shows the error in the recovered translation direction, the middle row corresponds to the error in the recovered rotation axis and the bottom row shows the error in the estimated rotation angle. Mean values are shown in the left column, standard deviations in the right.

	translation			rotation			rotation	**F** matrix		
				axis			angle (deg)			
motion1	600	0	0	.409	.818	.409	7.0	0.0	0.0	0.0
								-2.465e-06	1.431e-06	0.021
								-0.0003	-0.021	0.999
motion2	700	0	0	.802	.267	.534	6.0	0.0	0.0	0.0
								-3.567e-07	1.288e-06	0.011
								-0.0005	-0.012	0.999
motion3	motion1 followed by motion2							1.441e-08	-1.370e-07	-0.0001
								-9.629e-07	1.337e-06	0.007
								-0.0004	-0.008	0.999

Estimated A matrix using SVD		
839.911	0.00559322	309.987
0	770.001	270
0	0	1

Table 2: A sequence of motions that is singular for **ZF** . The simulated 3D motions along with the corresponding fundamental matrices are shown in the top part of the table. The intrinsic parameters recovered using the **SVD** algorithm are shown in the bottom part of the table.

equations derived from (29) do not correspond to tautologies. This example clearly demonstrates one of the advantages of using the **SVD** method instead of **ZF** which arbitrarily selects two out of the six available constraints. It should be stressed at this point that even if the implementation of the **ZF** algorithm employed here selected a different pair of constraints, an appropriate 3D motion could be found that would make the selected constraints unsolvable for the elements of matrix **K**.

7.1.4 Metric Measurements

This subsection examines the noise sensitivity of angles and length ratios computed with the aid of the recovered intrinsic calibration; appendix A should be consulted for more details regarding this computation. Three rigid displacements of the camera have been simulated and 300 random 3D points have been projected to the corresponding four positions of the simulated retina. The standard deviation of the Gaussian noise added to retinal points was increased from 0 to 1.5 pixels. After estimating the six fundamental matrices defined by the three displacements, the **SVD** algorithm was applied to these estimates for recovering the intrinsic calibration matrix. The minimization of Eq.(33) has been performed using four unknowns, i.e. the skew angle has been assumed to be known and equal to $\pi/2$. The obtained estimates are shown in table 3. The leftmost column of this table indicates the standard deviation of the Gaussian noise added to retinal points, while the second left column consists of the estimated calibration matrices. To give an indication regarding the suitability of the recovered calibration matrices for metric measurements, they have been used to recover from the retinal projections a set of angles and length ratios formed by line segments defined by the simulated 3D points. Using the simulated 3D points, 100 random angles and 100 random segment pairs were formed. The corresponding angles and length ratios were then measured using the known 3D coordinates of the points defining the selected line segments. Subsequently, the recovered calibration matrices were combined with the 2D projections of the points defining the line segments to reestimate angles and length ratios using Eq.(34). The mean and standard deviation of the relative error between the actual values and the estimated ones are summarized in table 3. More specifically, the third column of this table refers to angle measurements while the fourth column to measurements of length ratios. The values in the parentheses are the error statistics computed when angles and length ratios were estimated using the true (i.e. not the estimated) calibration and fundamental matrices. These values represent lower bounds for the error in metric measurements from images, since in this case the **A** and **F** matrices are known with perfect accuracy and the only source of error is the noise corrupting the retinal projections. As can be verified from table 3, the introduction of the estimated matrices in the measurement process increases the error only slightly. This implies that both the calibration and the fundamental matrices have been obtained with satisfactory accuracy.

7.2 Experiments with Real Images

Three experiments performed with real images are reported in this section. In all of these experiments, point matches and the associated fundamental matrices have been computed with the aid of [5]. The line segments employed for 3D angle measurements have been recovered automatically with the aid of a line extractor; the endpoints of the line segments used for measuring 3D length ratios have been defined manually by clicking on the images with a mouse. Finally, throughout all experiments, the skew angle θ has been assumed known and equal to $\pi/2$; matrix **K** in Eq.(33) is thus parameterized using four unknowns.

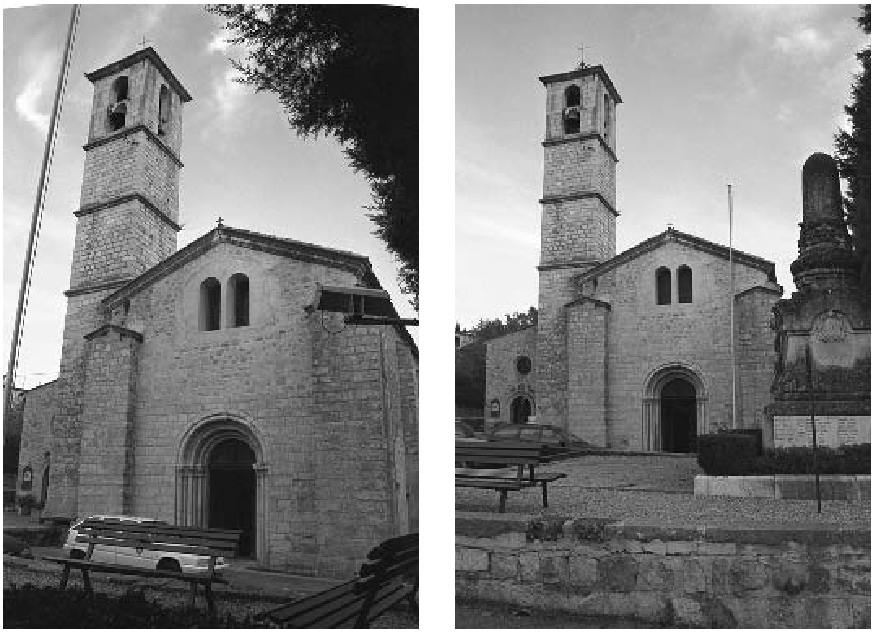
The first experiment is performed using five 512×768 images of the church of the village of Valbonne. Images 0 and 2 of this sequence are shown in Figs. 8(a) and (b). Self-calibration has been performed using the ten fundamental matrices defined by all possible image pairs. The estimated intrinsic calibration matrix is shown in the top row of table 4. Figures 9(a) and (b) illustrate the line segments used along with the recovered calibration to compute the angles reported in the left part of table 4. More specifically, the leftmost column of this table indicates the pairs of image line segments defining the angles, the second column from the left corresponds to the ground truth angle values, the third column from the left supplies the cosine of the angles as computed from the images using Eq. (34) and, finally, the fourth column shows the angles corresponding to the estimated cosines. In addition to estimating angles, the recovered intrinsic parameters have also been used for estimating

True A matrix		
840	0	310
0	770	270
0	0	1

Noise	Estimated **A** matrix			Angles rel. error		Length ratios rel. error	
std dev				mean	std dev	mean	std dev
	840.732	-1.44521e-14	310.058	0.0038	0.0064	0.0192	0.0705
0.0	0	770.606	270.606	(4.95e-06)	(7.95e-06)	(1.94e-05)	(6.32e-05)
	0	0	1				
	840.216	0	310.905	0.0076	0.0130	0.0190	0.0438
0.1	0	770.008	267.742	(0.0059)	(0.0099)	(0.0200)	(0.0505)
	0	0	1				
	828.684	1.39949e-14	320.451	0.0466	0.0726	0.0972	0.1748
0.5	0	769.914	262.45	(0.0449)	(0.0776)	(0.0924)	(0.1921)
	0	0	1				
	836.937	0	329.008	0.0971	0.2534	0.1290	0.2738
1.0	0	752.633	289.929	(0.0999)	(0.3480)	(0.1336)	(0.2864)
	0	0	1				
	830.949	2.81749e-14	327.062	0.1220	0.2895	0.1657	0.2789
1.5	0	770.784	261.244	(0.1034)	(0.2397)	(0.1716)	(0.3125)
	0	0	1				

Table 3: Synthetic experiments results. The left column shows the standard deviation (in pixels) of the noise added to image points, the second column shows the estimated **A** matrices, the third column contains the error statistics for the measured 3D angles and the fourth column illustrates the error statistics for the measured 3D segment length ratios. The true values for the **A** matrix are given in the top row of the table.

3D length ratios using the line segments shown in Figures 9(c) and (d). The computed length ratios are shown in the right part of table 4. The left column of this table indicates the pairs of image line segments defining the ratios, the middle column corresponds to the ground truth ratio values and the third column contains the ratios as computed from the images using Eq. (37).

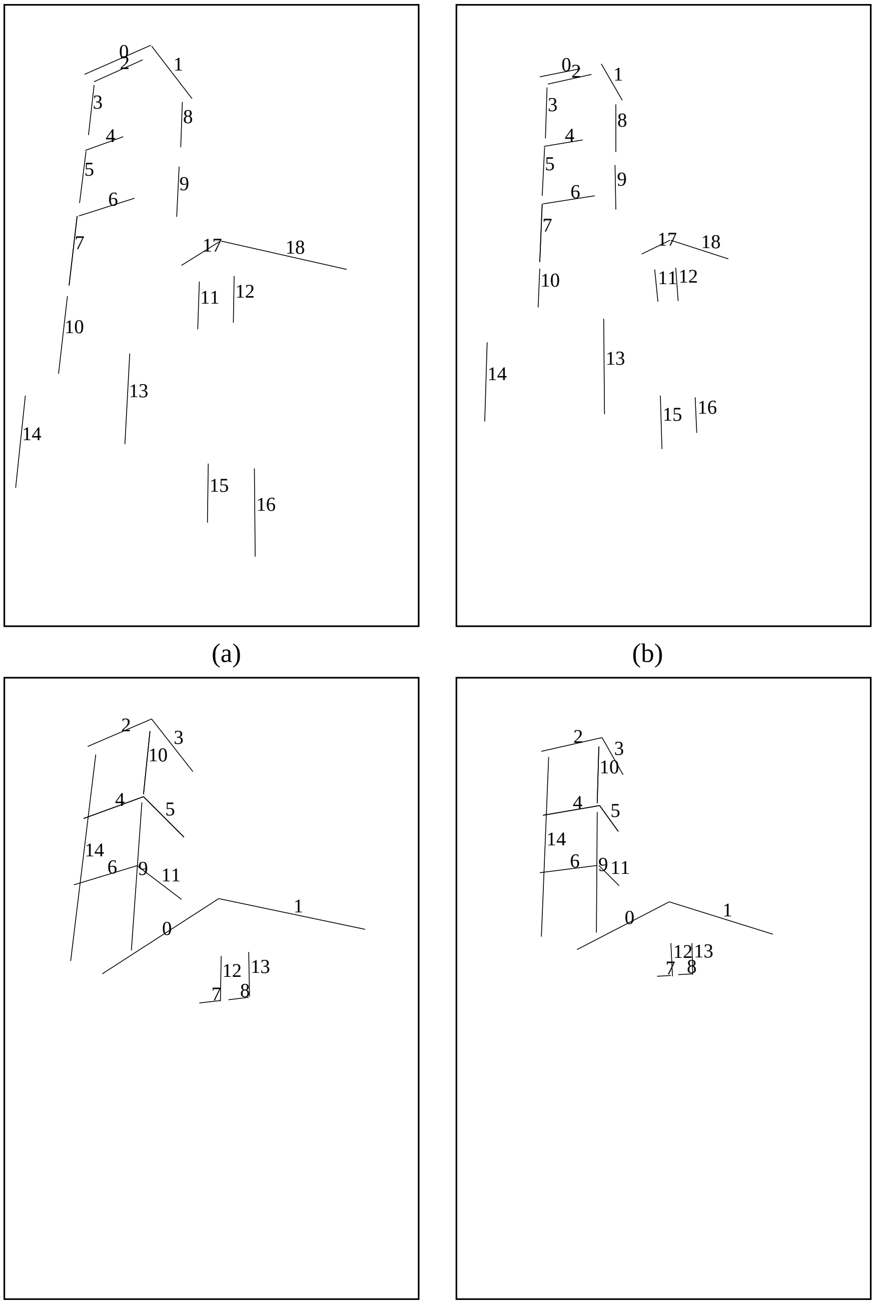


(a) (b)

Figure 8: Frames 0 and 2 of the Valbonne church sequence.

The second experiment refers to images 11 up to 15 of a long sequence depicting one of the INRIA buildings in Sophia-Antipolis. Images 11 and 12 of this sequence are shown in Figs. 10(a) and (b). Images are of dimension 765×509 pixels. Self-calibration has been performed using all ten fundamental matrices defined by images 11 to 15. The estimated intrinsic calibration matrix is shown in table 5. This table also illustrates the angles and the length ratios computed using the recovered intrinsic calibration and the line segments in Figs. 11(a)-(b) and (c)-(d) respectively.

The third experiment is based on a sequence depicting the Arcades Square in Valbonne. Images are of dimension 768×512 pixels; frames 3 and 4 of this sequence are shown in Figs. 12(a) and (b). Self-calibration has employed all ten fundamental matrices defined by the first five frames of the sequence. The estimated intrinsic calibration matrix is shown in table 6. This table also illustrates the angles and the length ratios computed using the recovered intrinsic calibration and the line segments in Figs. 12(a)-(b) and (c)-(d) respectively.



(a)

(b)

(c)

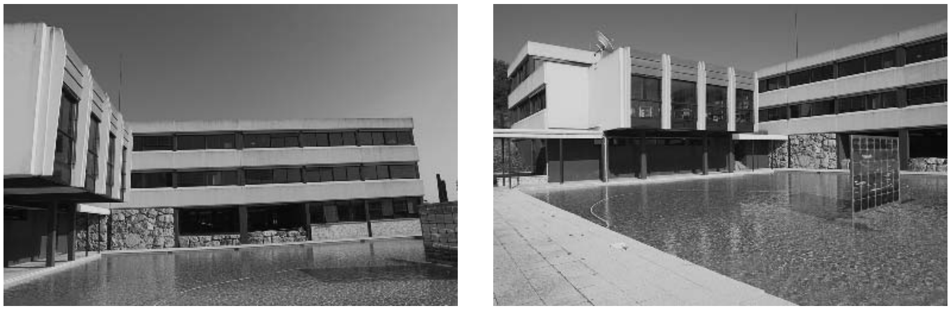
(d)

Figure 9: Line segments of the Valbonne church sequence: (a)-(b) the line segments that are used for measuring 3D angles, (c)-(d) the line segments that are used for measuring 3D length ratios.

Estimated intrinsic calibration matrix		
682.84	1.68554e-14	255.999
0	682.843	383.999
0	0	1

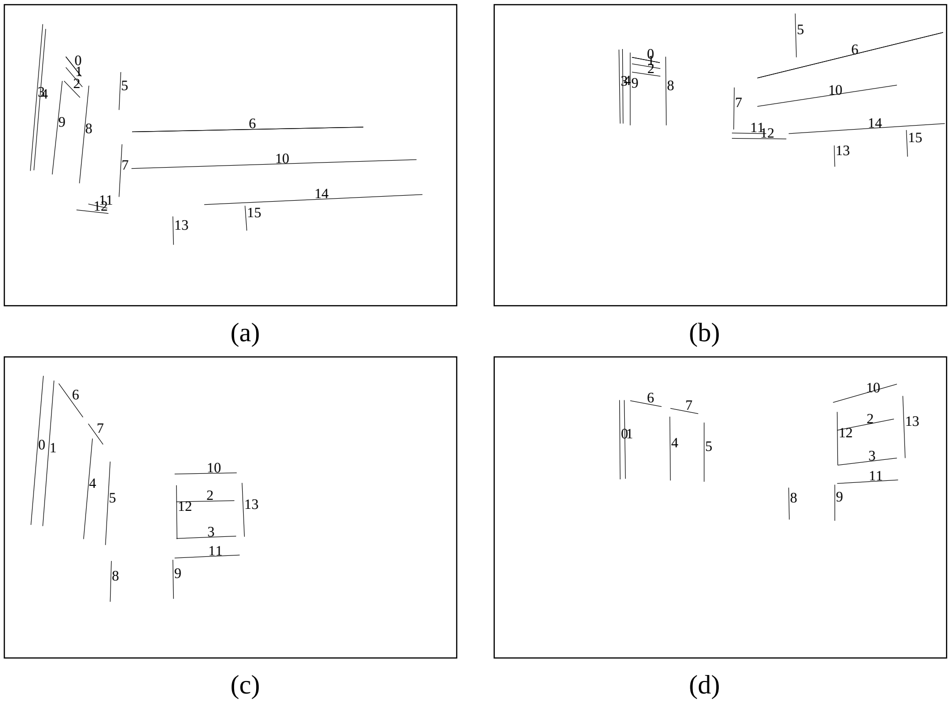
Line segments from Figs.9 (a)-(b)				**Line segments from Figs.9 (c)-(d)**		
Angle segments	Actual angle (deg)	Estimated cosine	Estimated angle (deg)	Line segments	Actual length ratio	Estimated length ratio
0 - 1	90	0.0020432	89.882933	0-1	1.0	0.977341
2 - 3	90	0.0207585	88.810540	2-3	1.0	0.833801
4 - 5	90	0.0293123	88.320287	4-5	1.0	0.851116
6 - 7	90	0.0287784	88.350891	4-6	1.0	1.04367
8 - 9	0	0.999115	2.411304	7-8	1.0	0.949896
7 - 10	0	0.999947	0.588870	9-10	2.0	1.9691
11 - 12	0	0.995942	5.163766	5-11	1.0	1.02813
13 -14	0	0.99997	0.446021	12-13	1.0	1.04544
15 - 16	0	0.999766	1.238219	14-10	3.0	3.07484
17 - 18	?	0.733482	42.820857			

Table 4: Valbonne church sequence: estimated intrinsic calibration matrix (top), ground truth and estimated angle values (left), ground truth and estimated length ratios (right).



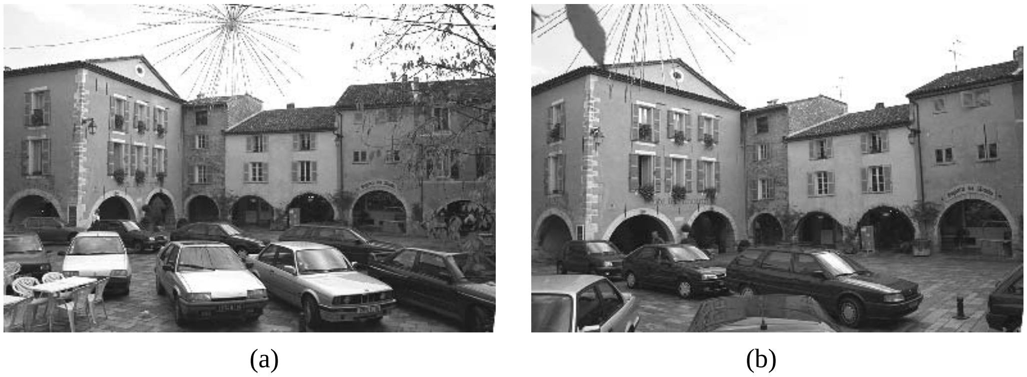
(a) (b)

Figure 10: Frames 11 and 12 of the INRIA building sequence.



(a) (b) (c) (d)

Figure 11: Line segments of the INRIA building sequence: (a)-(b) the line segments that are used for measuring 3D angles, (c)-(d) the line segments that are used for measuring 3D length ratios.



(a) (b)

Figure 12: Frames 3 and 4 of the Arcades Square sequence.

Estimated intrinsic calibration matrix		
578.351	-2.07997e-14	372.788
0	587.305	265.442
0	0	1

Line segments from Figs.11 (a)-(b)				**Line segments from Figs.11 (c)-(d)**		
Angle segments	Actual angle (deg)	Estimated cosine	Estimated angle (deg)	Line segments	Actual length ratio	Estimated length ratio
0 - 1	0	0.999988	0.279051	0-1	1.0	1.00816
0 - 2	0	0.999943	0.609908	2-3	1.0	0.978988
3 - 4	0	0.999972	0.427985	4-5	1.0	0.999802
5 - 6	90	0.110625	83.648654	6-7	1.0	1.04928
7 - 8	0	0.99956	1.699553	8-9	1.0	1.02679
9 - 10	90	0.157303	80.949621	10-11	1.0	0.976233
11 - 12	0	0.998336	3.305772	12-13	1.0	0.960852
13 - 6	90	0.106399	83.892238			
14 - 15	90	0.0171276	89.018615			

Table 5: INRIA building sequence: estimated intrinsic calibration matrix (top), ground truth and estimated angle values (left), ground truth and estimated length ratios (right).

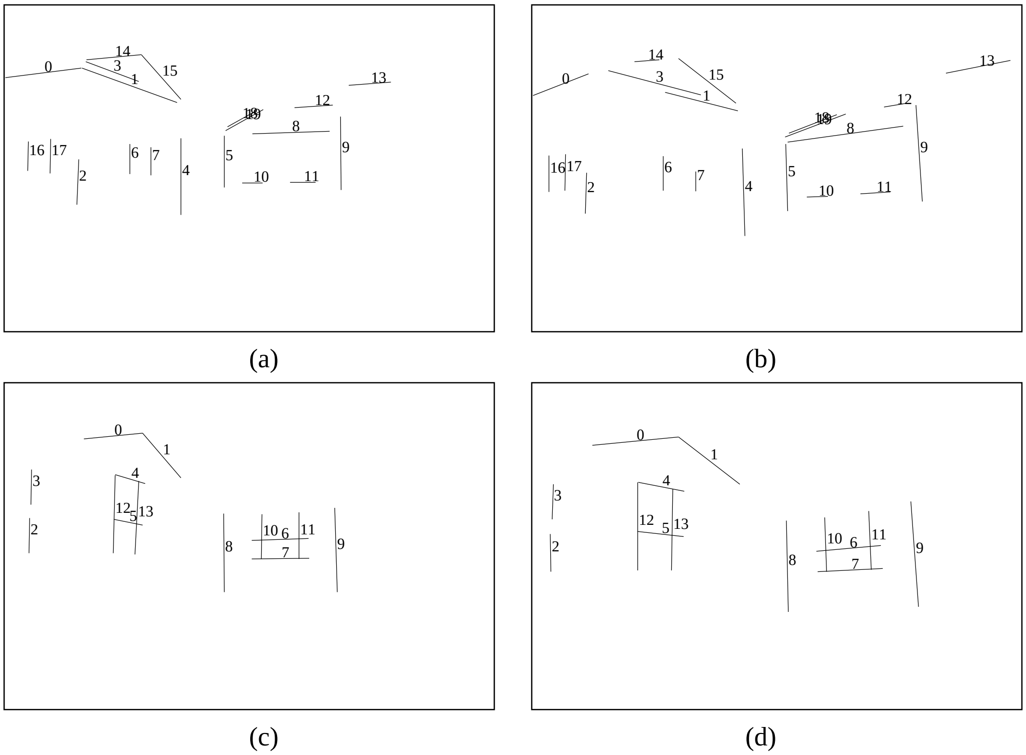


Figure 13: Line segments of the Arcades Square sequence: (a)-(b) the line segments that are used for measuring 3D angles, (c)-(d) the line segments that are used for measuring 3D length ratios.

Estimated intrinsic calibration matrix		
699.995	0	384
0	700.003	255.993
0	0	1

Line segments from Figs.13 (a)-(b)				**Line segments from Figs.13 (c)-(d)**		
Angle	Actual	Estimated	Estimated	Line	Actual	Estimated
segments	angle (deg)	cosine	angle (deg)	segments	length ratio	length ratio
0 - 1	90	0.0498395	87.143221	0-1	1.0	0.99806
2 - 3	90	0.0585627	86.642683	2-3	1.0	1.09528
4 - 5	0	0.999654	1.507388	4-5	1.0	0.947202
6 - 7	0	0.998966	2.606207	6-7	1.0	1.00159
8 - 9	90	0.00810136	89.535821	8-9	1.0	0.984137
10 - 11	0	0.99679	4.591879	10-11	1.0	0.981422
12 - 13	0	0.983164	10.528644	12-13	1.0	0.978607
14 - 15	?	0.808495	36.050881			
16 - 17	0	0.999652	1.510980			
18 - 19	0	0.999962	0.497828			

Table 6: Arcades Square sequence: estimated intrinsic calibration matrix (top), ground truth and estimated angle values (left), ground truth and estimated length ratios (right).

8 Conclusions

Camera calibration is a prerequisite to a wide variety of vision tasks. In this paper, a novel method for self-calibration has been proposed. The method employs a simplification of the Kruppa equations which relies solely on the SVD of the fundamental matrix and avoids recovering noise-sensitive quantities such as the epipoles. The simplified Kruppa equations form the basis for a non-linear minimization scheme that yields the intrinsic calibration parameters. A detailed performance evaluation of the proposed method has also been carried out and experimental results using synthetic data as well as real image sequences were reported. Additionally, the proposed method has been compared with a state of the art self-calibration method [52] and superior results have been obtained. As demonstrated by the Euclidean measurements performed with the aid of the recovered calibration matrices, the estimates of the latter are of acceptable accuracy for the purpose of metric reconstruction. Future work will investigate alternative ways for estimating the covariance of the SVD, so as to obtain better estimates of uncertainty that will lead to better weights for the constraints in the nonlinear minimization step, i.e. Eq.(33). Another possible direction for future study concerns the application of sparse elimination methods [13, 12] to systems of polynomial equations defined by Eqs. (30) (see also section 6.1). In this fashion, better initial solutions for the nonlinear minimization algorithm might be obtained.

A Direct Computation of Angles and Length Ratios from Images

Knowledge of the calibration matrix $\mathbf{A}$ enables certain 3D Euclidean measurements to be made directly from the images. More specifically, the angle between two 3D line segments L_1 and L_2 can be computed as follows [51, 52]: Let $\mathbf{l_1}$, $\mathbf{l_1'}$ and $\mathbf{l_2}$, $\mathbf{l_2'}$ be the projections of L_1 and L_2 in the two images. By employing Laguerre's formula, the angle between L_1 and L_2 is shown to be given by

$$cos(L_1, L_2) = \frac{|S(\mathbf{v_1}, \mathbf{v_2})|}{\sqrt{S(\mathbf{v_1}, \mathbf{v_1})S(\mathbf{v_2}, \mathbf{v_2})}}, \tag{34}$$

where $\mathbf{v_1}$ and $\mathbf{v_2}$ are the projections in the first image of the intersections of L_1 and L_2 with the *plane at infinity* and

$$S(\mathbf{m}, \mathbf{n}) = \mathbf{m}^t \mathbf{A}^{-t} \mathbf{A}^{-1} \mathbf{n} \tag{35}$$

Points $\mathbf{v_1}$ and $\mathbf{v_2}$ are determined by

$$\begin{aligned} \mathbf{v_1} &= \mathbf{l_1} \times \mathbf{H}_\infty{}^t \mathbf{l_1'} \\ \mathbf{v_2} &= \mathbf{l_2} \times \mathbf{H}_\infty{}^t \mathbf{l_2'}, \end{aligned}$$

where $\mathbf{H}_\infty$ is the *homography of the plane at infinity*, corresponding to the uncalibrated rotational component of motion and being defined by

$$\mathbf{H}_\infty = \mathbf{A}\mathbf{R}\mathbf{A}^{-1} \tag{36}$$

In addition to angle measurements, the intrinsic calibration parameters also permit line segment ratios to be directly calculated from pairs of images. Specifically, let $A,\ B,\ C$ and D be four 3D points giving rise to the following corresponding projections in the two images: $(a,\ a'),\ (b,\ b'),\ (c,\ c')$ and $(d,\ d')$. It can be shown [51] that the ratio of the 3D segments AB and CD is given by

$$\frac{AB}{CD} = \sqrt{\frac{S(\mathbf{v}_{ab}, \mathbf{v}_{ab})}{S(\mathbf{v}_{cd}, \mathbf{v}_{cd})} \cdot \frac{S(\mathbf{v}_{ac}, \mathbf{v}_{ac})S(\mathbf{v}_{bc}, \mathbf{v}_{bc}) - S^2(\mathbf{v}_{ac}, \mathbf{v}_{bc})}{S(\mathbf{v}_{ab}, \mathbf{v}_{ab})S(\mathbf{v}_{ac}, \mathbf{v}_{ac}) - S^2(\mathbf{v}_{ab}, \mathbf{v}_{ac})} \cdot \frac{S(\mathbf{v}_{bd}, \mathbf{v}_{bd})S(\mathbf{v}_{cd}, \mathbf{v}_{cd}) - S^2(\mathbf{v}_{bd}, \mathbf{v}_{cd})}{S(\mathbf{v}_{bc}, \mathbf{v}_{bc})S(\mathbf{v}_{bd}, \mathbf{v}_{bd}) - S^2(\mathbf{v}_{bc}, \mathbf{v}_{bd})}}, \tag{37}$$

where

$$\mathbf{v}_{ij} = \mathbf{l}_{ij} \times \mathbf{H}_\infty^t \mathbf{l}'_{ij}, \quad ij \in \{ab, ac, bc, bd, cd\}$$

is the projection in the first image of the point of intersection of the line defined by the 3D points I and J and the plane at infinity, while $\mathbf{l}_{ij}$ is the line defined by image points i and j, given by the cross product $i \times j$. Function $S(\cdot\ ,\ \cdot)$ is defined by Eq.(35).

B SVD Differentiation

This appendix shows how can the derivatives of the SVD decomposition of a matrix with respect to that matrix be computed. The derivation presented below is due to Théo Papadopoulo.

Let $\mathbf{A}$ be a $M \times N$ matrix and let $\mathbf{U}, \mathbf{D}$ and $\mathbf{V}$ be its SVD components, i.e. $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^t$ with $\mathbf{U}, \mathbf{V}$ being orthogonal ($\mathbf{U}\mathbf{U}^t = \mathbf{V}\mathbf{V}^t = \mathbf{I}$) and $\mathbf{D}$ being diagonal. In the following, we will use the notation m_{ij} to refer to the (i, j) element of matrix $\mathbf{M}$. With the above definitions, we are interested in computing $\frac{\partial}{\partial a_{ij}}\mathbf{U}$, $\frac{\partial}{\partial a_{ij}}\mathbf{V}$ and $\frac{\partial}{\partial a_{ij}}\mathbf{D}$ for every element a_{ij} of $\mathbf{A}$.

Taking the derivative of $\mathbf{A}$ with respect to a_{ij} yields the following equation

$$\frac{\partial}{\partial a_{ij}}\mathbf{A} = (\frac{\partial}{\partial a_{ij}}\mathbf{U})\mathbf{D}\mathbf{V}^t + \mathbf{U}(\frac{\partial}{\partial a_{ij}}\mathbf{D})\mathbf{V}^t + \mathbf{U}\mathbf{D}(\frac{\partial}{\partial a_{ij}}\mathbf{V}^t) \tag{38}$$

Clearly, $\frac{\partial a_{kl}}{\partial a_{ij}} = 0, \ \ \forall k \neq i$ and $l \neq j.$ By multiplying Eq. (38) by $\mathbf{U}^t$ and $\mathbf{V}$ from the left and right respectively, the following equation is obtained

$$\mathbf{U}^t(\frac{\partial}{\partial a_{ij}}\mathbf{A})\mathbf{V} = \Psi_{\mathbf{U}}\mathbf{D} + (\frac{\partial}{\partial a_{ij}}\mathbf{D}) + \mathbf{D}\Psi_{\mathbf{V}}, \tag{39}$$

where $\Psi_{\mathbf{U}}$ and $\Psi_{\mathbf{V}}$ are defined as

$$\begin{aligned}\Psi_{\mathbf{U}} &= \mathbf{U}^t(\frac{\partial}{\partial a_{ij}}\mathbf{U})\\ \Psi_{\mathbf{V}} &= \mathbf{V}^t(\frac{\partial}{\partial a_{ij}}\mathbf{V})\end{aligned}$$

Matrices $\Psi_{\mathbf{U}}$ and $\Psi_{\mathbf{V}}$ are antisymmetric, since, for example, for U the following holds:

$$\mathbf{U}^t\mathbf{U} = \mathbf{I} \implies (\frac{\partial}{\partial a_{ij}}\mathbf{U})^t\mathbf{U} + \mathbf{U}^t + (\frac{\partial}{\partial a_{ij}}\mathbf{U}) = {\Psi_{\mathbf{U}}}^t + \Psi_{\mathbf{U}} = \mathbf{0} \tag{40}$$

This implies that all the diagonal elements of $\Psi_{\mathbf{U}}$ and $\Psi_{\mathbf{V}}$ are zero, and therefore, recalling that $\mathbf{D}$ is diagonal, the same is true for the diagonal elements of $\Psi_{\mathbf{U}}\mathbf{D}$ and $\mathbf{D}\Psi_{\mathbf{V}}$. Thus, Eq.(39) gives that

$$\frac{\partial s_{kk}}{\partial a_{ij}} = u_{ik}\ v_{jk} \tag{41}$$

The elements of the matrices $\Psi_{\mathbf{U}}$ and $\Psi_{\mathbf{V}}$ can be computed by solving a set of 2×2 linear systems:

$$\begin{aligned}s_{ll}\ {\psi_{\mathbf{U}}}_{kl} + s_{kk}\ {\psi_{\mathbf{V}}}_{kl} &= u_{ik}\ v_{jl}\\ s_{kk}\ {\psi_{\mathbf{U}}}_{kl} + s_{ll}\ {\psi_{\mathbf{V}}}_{kl} &= -u_{il}\ v_{jk}\end{aligned}$$

where the range indices are $1 \ldots \min(M,N)-1$ for k and $i+1 \ldots \min(M,N)$ for l. After computing $\Psi_{\mathbf{U}}$ and $\Psi_{\mathbf{V}}$, $\frac{\partial}{\partial a_{ij}}\mathbf{U}$ and $\frac{\partial}{\partial a_{ij}}\mathbf{V}$ follow as

$$\begin{aligned} \frac{\partial}{\partial a_{ij}}\mathbf{U} &= \mathbf{U}\Psi_{\mathbf{U}} \\ \frac{\partial}{\partial a_{ij}}\mathbf{V}^t &= \Psi_{\mathbf{V}}\mathbf{V}^t \end{aligned} \tag{42}$$

In summary, the desired derivatives are supplied by Eqs. (41) and (42).

Acknowledgments

The authors thank Théo Papadopoulo for providing the code for the estimation of the uncertainty associated with the SVD of the fundamental matrix. He also provided valuable advise related to practical issues.

References

[1] J. Aloimonos, I. Weiss, and A. Bandopadhay. Active vision. *The International Journal of Computer Vision*, 1(4):333–356, 1988.

[2] E. Anderson, Z. Bai, C. Bishof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, 3600 University City Science Center, Philadelphia, PA 19104-2688, second edition, 1994.

[3] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In *Fourth European Conference on Computer Vision*, pages 3–16, April 1996.

[4] Sylvain Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *IEEE International Conference on Computer Vision*, pages 790–796, 1998.

[5] Sylvain Bougnoux and Luc Robert. Totalcalib: a fast and reliable system for off-line calibration of image sequences. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, June 1997. The Demo Session.

[6] M. J. Brooks, L. de Agapito, D. Q. Huyng, and L. Baumela. Direct methods for self-calibration of a moving stereo head. In *Fourth European Conference on Computer Vision*, volume II, pages 415–426, April 1996.

[7] Gabriella Csurka, Cyril Zeller, Zhengyou Zhang, and Olivier Faugeras. Characterizing the uncertainty of the fundamental matrix. *CVGIP: Image Understanding*, 68(1):18–36, October 1997.

[8] K. Daniilidis and M. Spetsakis. Understanding noise sensitivity in structure form motion. In Yiannis Aloimonos, editor, *Visual Navigation: From Biological Systems to Unmanned Ground Vehicles*, chapter 4, pages 61–88. Lawrence Erlbaum Associates, Hillsdale, NJ, 1997.

[9] L. de Agapito, R. Hartley, and E. Hayman. Linear calibration of a rotating and zooming camera. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 15–21, Fort Collins, Colorado, June 1999. IEEE Computer Society.

[10] L. de Agapito, E. Hayman, and I.L. Reid. Self-calibration of a rotating camera with varying intrinsic parameters. In *British Machine Vision Conference*, Southampton, UK, September 1998. BMVA Press.

[11] D. Demirdjian, G. Csurka, and Horaud R. Autocalibration in the presence of critical motions. In *British Machine Vision Conference*, pages 751–759., University of Southampton, UK, September 1998.

[12] I.Z. Emiris. A general solver based on sparse resultants: Numerical issues and kinematic applications. Technical Report 3110, INRIA/SAFIR, 1997.

[13] I.Z. Emiris and J.F. Canny. Efficient incremental algorithms for the sparse resultant and the mixed volume. *J. of Symbolic Computation*, 20(2):117–149, August 1995.

[14] O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.

[15] Olivier Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proceedings of the 2nd ECCV*, pages 563–578, may 1992.

[16] Olivier Faugeras. Stratification of 3-D vision: projective, affine, and metric representations. *Journal of the Optical Society of America A*, 12(3):465–484, March 1995.

[17] Olivier Faugeras and Steve Maybank. Motion from point matches: multiplicity of solutions. *The International Journal of Computer Vision*, 4(3):225–246, 1990.

[18] G.H. Golub and C.F. Van Loan. *Matrix computations*. The John Hopkins University Press, Baltimore, Maryland, 2 edition, 1989.

[19] R. I. Hartley. Estimation of relative camera positions for uncalibrated cameras. In G. Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision*, pages 579–587, Santa Margherita, Italy, May 1992. Springer-Verlag.

[20] R.I. Hartley. Kruppa's equations derived from the fundamental matrix. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):133–135, February 1997.

[21] Richard Hartley. Self-calibration of stationary cameras. *The International Journal of Computer Vision*, 22(1):5–24, February 1997.

[22] Richard Hartley, Rajiv Gupta, and Tom Chang. Stereo from uncalibrated cameras. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 761–764, Urbana Champaign, IL, June 1992. IEEE.

[23] A. Heyden and K. Aström. Algebraic varieties in multiple view geometry. In *Fourth European Conference on Computer Vision*, volume II, pages 671–682, 1996.

[24] Anders Heyden and Kalle Åström. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Comp. Vision and Pattern Rec.*, pages 438–443. IEEE Computer Society Press, 1997.

[25] R. Horaud, D. Demirdjian, and G. Csurka. Closed-form solutions for the euclidean calibration of a stereo rig. In *Proceedings of the 5th European Conference on Computer Vision*, Freibourg, Germany, June 98.

[26] R. Horaud, F. Dornaika, B. Boufama, and R. Mohr. Self calibration of a stereo head mounted onto a robot arm. In J.O. Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, pages 455–462. Springer-Verlag, 1994.

[27] Radu Horaud and Gabriella Csurka. Self-calibration and euclidean reconstruction using motions of a stereo rig. In *Proceedings of the 6th International Conference on Computer Vision*, Bombay, India, January 1998. IEEE Computer Society, IEEE Computer Society Press.

[28] Thomas S. Huang and Olivier D. Faugeras. Some properties of the E matrix in two-view motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12):1310–1312, December 1989.

[29] T.S. Huang and A.N. Netravali. Motion and structure from feature correspondences: A review. *Proc. IEEE*, 82(2):252–268, February 1994.

[30] K. Kanatani. *Geometric computation for machine vision*. Oxford university press, 1992.

[31] E. Kruppa. Zur Ermittlung eines Objektes aus zwei Perspektiven mit innerer Orientierung. *Sitz.-Ber. Akad. Wiss., Wien, math. naturw. Kl., Abt. IIa.*, 122:1939–1948, 1913.

[32] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.

[33] Q.-T. Luong and O. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. *The International Journal of Computer Vision*, 22(3):261–289, 1997.

[34] Quang-Tuan Luong. *Matrice Fondamentale et Calibration Visuelle sur l'Environnement-Vers une plus grande autonomie des systèmes robotiques*. PhD thesis, Université de Paris-Sud, Centre d'Orsay, December 1992.

[35] S. J. Maybank and O. D. Faugeras. A theory of self-calibration of a moving camera. *The International Journal of Computer Vision*, 8(2):123–152, August 1992.

[36] R. Mohr and B. Triggs. *Projective Geometry for Image Analysis*. ISPRS workshop tutorial, Vienna, Austria, July 1996.

[37] Joseph L. Mundy and Andrew Zisserman, editors. *Geometric Invariance in Computer Vision*. MIT Press, 1992.

[38] S. Negahdaripour. Direct Computation of the FOE with Confidence Measures. *Computer Vision and Image Understanding*, 64(3):323–350, 1996.

[39] M. Pollefeys and L. Van Gool. Self-calibration from the absolute conic on the plane at infinity. In *Proc. CAIP'97*, LNCS vol.1296, pages 175–182, Kiel, Germany, 1997. Springer-Verlag.

[40] M. Pollefeys, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *IEEE International Conference on Computer Vision*, pages 90–95, 1998.

[41] M. Pollefeys and L. Van Gool. A stratified approach to metric self-calibration. In *IEEE International Conference on Computer Vision and Pattern Recogition*, pages 407–412, 1997.

[42] Amnon Shashua. Projective structure from uncalibrated images: structure from motion and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):778–790, 1994.

[43] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Puerto Rico, USA*, pages 1100–1105, 1997.

[44] P. Sturm. Self-calibration of a moving zoom-lens camera by pre-calibration. *Image and Vision Computing*, 15(8):583–589, August 1997.

[45] T. Svoboda and P. Sturm. A badly calibrated camera in ego-motion estimation, propagation of uncertainty. In *Proc. of the 7th Int. Conf. on Computer Analysis of Images and Patterns (CAIP)*, Kiel, Germany, September 1997.

[46] B. Triggs. Autocalibration and the absolute quadric. In *IEEE International Conference on Computer Vision and Pattern Recogition*, pages 609–614, 1997.

[47] H. P. Trivedi. Can multiple views make up for lack of camera registration. *Image and Vision Computing*, 6(1):29–32, February 1988.

[48] R. Tsai. Synopsis of recent progress on camera calibration for 3D machine vision. In Oussama Khatib, John J. Craig, and Tomás Lozano-Pérez, editors, *The Robotics Review*, pages 147–159. MIT Press, 1989.

[49] Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, August 1987.

[50] R.Y. Tsai and T.S. Huang. Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(1):13–26, January 1984.

[51] Cyril Zeller. *Calibration Projective Affine et Euclidienne en Vision par Ordinateur*. PhD thesis, École Polytechnique, February 1996.

[52] Cyril Zeller and Olivier Faugeras. Camera self-calibration from video sequences: the Kruppa equations revisited. Research Report 2793, INRIA, February 1996.

[53] Z. Zhang. A new multistage approach to motion and structure estimation: From essential parameters to euclidean motion via fundamental matrix. Research Report 2910, INRIA Sophia-Antipolis, France, June 1996.

[54] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal*, 78:87–119, October 1995.

[55] Z. Zhang, Q.-T. Luong, and O. Faugeras. Motion of an uncalibrated stereo rig: Self-calibration and metric reconstruction. Research Report 2079, INRIA Sophia-Antipolis, 1993.

[56] Andrew Zisserman, David Liebowitz, and Martin Armstrong. Resolving ambiguities in auto-calibration. *Philosophical Transactions of the Royal Society of London, SERIES A*, 356(1740):1193–1211, 1998.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399