INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Camera Self-Calibration Using the Kruppa Equations and the SVD of the Fundamental Matrix: The Case of Varying Intrinsic Parameters

Manolis I.A. LOURAKIS and Rachid DERICHE

N° 3911

Mars 2000

THÈME 3

ISSN 0249-6399

Camera Self-Calibration Using the Kruppa Equations and the SVD of the Fundamental Matrix: The Case of Varying Intrinsic Parameters

Manolis I.A. LOURAKIS and Rachid DERICHE

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

Rapport de recherche n° 3911 — Mars 2000 — 35 pages

Abstract: Estimation of the camera intrinsic calibration parameters is a prerequisite to a wide variety of vision tasks related to motion and stereo analysis. A major breakthrough related to the intrinsic calibration problem was the introduction in the early nineties of 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. Until recently, however, most research efforts have been focused on applying the autocalibration paradigm to estimating constant intrinsic calibration parameters. Therefore, such approaches are inapplicable to cases where the intrinsic parameters undergo continuous changes due to focusing and/or zooming. In this paper, our previous work for autocalibration in the case of constant camera intrinsic parameters is extended and a novel autocalibration method capable of handling variable intrinsic parameters is proposed. The 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, a purely algebraic derivation is proposed here which provides a straightforward answer to the problem of determining which constraints to employ among the set of available ones. Additionally, the new formulation does not employ the epipoles, which are known to be difficult to estimate accurately. The intrinsic calibration parameters are recovered from the

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

developed constraints through a nonlinear minimization scheme that explicitly takes into consideration the uncertainty associated with the estimates of the employed fundamental matrices. Detailed experimental results using both simulated and real image sequences demonstrate the feasibility of the approach.

Key-words: *Self-Calibration, Varying Intrinsic Parameters, Kruppa Equations, 3D Reconstruction, Motion Analysis, Stereo, Structure from Motion.*

Auto-Calibration par Équations de Kruppa et Décomposition en Valeurs Singulières de la Matrice Fondamentale : Le Cas d'une Caméra à Paramètres Intrinsèques Variables.

Résumé : Ce rapport fait suite à nos précédents travaux sur l'auto-calibration d'une caméra à partir d'un ensemble de points appariés entre différentes images en utilisant une méthode à base d'équations de Kruppa simplifiées. Alors que nos précédent travaux traitaient du cas où les paramétres intrinsèques de la caméra étaient considérés comme constants, ce rapport étend l'étude de l'autocalibration d'une caméra au cas où ses paramètres intrinsèques sont variables. L'estimation des paramètres intrinsèques s'effectue en prenant en compte les contraintes de Kruppa simplifiées par la décomposition en valeurs singulières de la matrice fondamentale ainsi que l'incertitude liée aux matrices fondamentales utilisées. Cette prise en compte de l'incertitude et sa propagation jusqu'à l'estimation des paramètres intrinsèques variables nous a permis d'améliorer sensiblement la qualité des résultats obtenus à partir de données synthétiques bruitées et de données de plusieures images réelles.

Mots-clés : *Auto-Calibration, Equations de Kruppa, Paramètres intrinsèques variables, Analyse du mouvement, Stéréovision, Structure à partir du mouvement.*

1 Introduction

Early research in uncalibrated vision has made clear that without any information regarding the camera intrinsic calibration parameters, a sequence of images can yield 3D structure up to an arbitrary projectivity of 3D space [13, 21]. In order to upgrade this projective reconstruction to an Euclidean (i.e. metric) one, knowledge of the camera intrinsic calibration parameters is required. Consequently, the problem of determining the latter is of paramount significance for numerous vision tasks. Early approaches for coping with this problem relied upon the presence of an artificial calibration object (i.e. a grid) in the set of captured images [52, 53]. 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. Recently, Zhang [58] alleviated some of the restrictions imposed by grid-based calibration methods by proposing a technique based on the observation of a planar calibration pattern. 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 change frequently during the image acquisition process due to mechanical and thermal variations or due to focusing and zooming. This is particularly true in the context of active vision, where the camera optical parameters are purposively controlled in order to simplify the vision tasks under consideration [1]. Evidently, in these cases it is unrealistic to assume that a calibration object always exists in the acquired images.

A major breakthrough related to the intrinsic calibration problem was the introduction by Maybank and Faugeras [34] of the autocalibration paradigm, which removes the need for a calibration object. By tracking a set of points among images of a rigid scene, captured by a moving camera 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 [28]. Since then, several researchers have investigated the application of the Kruppa equations for solving the camera calibration problem. For example, Zeller [55, 56] and Heyden and Åström [22] propose variants of the basic approach. Pollefeys and Van Gool [42, 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 [31] 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 aid of numerical continuation methods. 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 [47] proposes a scheme that captures the interdependence of the intrinsic parameters by employing 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 [46, 49], Zisserman et al [61] and Demirdjian et al [10]. However, it is worth mentioning that Hartley [19, 20] 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 by using three images, the intrinsic calibration parameters of a camera undergoing planar motion can be recovered up to a two fold ambiguity.

The applicability of the self-calibration methods cited in the previous paragraph is constrained by the requirement for constant intrinsic calibration parameters throughout the acquisition of the whole image sequence. Increased flexibility can be obtained by allowing the camera intrinsic parameters to change due to focusing and/or zooming. Such attempts for solving the self-calibration problem in the case of varying camera intrinsic parameters have started to appear only recently and are still relatively few. Pollefeys et al [43], for example, estimate variable focal lengths but their method has the limitation of requiring a pure translation for initializing the principal point estimates. A less restrictive method has been proposed by Heyden and Åström [24], which 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. However, although their method estimates intrinsic calibration through a non-linear bundle adjustment algorithm, no means of obtaining a suitable initial solution was proposed. Extending the work of [24], Pollefeys et al [41, 40] proved that the absence of skew in the image plane is sufficient to allow for self-calibration. They also proposed an optimization framework for recovering the intrinsic parameters. Recently, Heyden and Åström [23] have proven a more general result according to which the existence of one constant intrinsic parameter suffices for self-calibration. Techniques for achieving self-calibration assuming restricted types of camera motion and non constant intrinsic parameters have also appeared [9, 8]. In all of the aforementioned methods, variable intrinsic calibration parameters are estimated through the recovery of the *absolute quadric* [50]. The latter is a degenerate quadric consisting of planes tangent to the absolute conic and has the property that its image projection coincides with the dual image of the absolute conic. The first step in recovering the absolute quadric is to obtain a projective reconstruction, that is use bundle adjustment to compute for each image a projection matrix that is consistent with inter-image correspondences and is known up to a 3D projective transformation [4, 45]. Then, the motion rigidity assumption is utilized in some constrained optimization framework to yield the intrinsic parameters and rectify the projective reconstruction to an Euclidean one.

It is interesting to note that although the Kruppa equations can be extended to the case of non constant intrinsic calibration parameters, up to now only Bougnoux [5] has exploited them to achieve self-calibration in this case. Even Bougnoux, however, estimates the projective projection matrices corresponding to the available images and embeds them into a nonlinear minimization algorithm for finding the homography that maps the estimated projection matrices to Euclidean ones. The Kruppa equations are employed only for obtaining an initial solution for the nonlinear minimization process. Mendonça and Cipolla [35] take a different approach and exploit the well-known equality constraint regarding the two nonzero singular values of the essential matrix [25, 17]. Since the latter is a function of the fundamental matrix and the unknown intrinsic parameters, the equality constraint corresponds to a constraint involving the intrinsic parameters which are recovered by optimizing an appropriate objective function. A preliminary characterization of camera motions that are critical for the estimation of varying focal lengths when the remaining intrinsic parameters are known, can be found in [48].

The present paper is an extension of our previous work for simplifying the Kruppa equations and using them for self-calibration in the case of constant intrinsic parameters [29, 30]. More specifically, in this work we extend the simplification of the Kruppa equations to the case of varying intrinsic calibration parameters and show how the simplified equations can be used 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. The intrinsic calibration parameters are recovered using a nonlinear minimization scheme that explicitly takes into consideration the uncertainty that is associated with the estimates of the employed fundamental matrices. For all developments, only pairwise fundamental matrices need to be estimated, thus the need for a projective reconstruction is relaxed.

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. Using a purely algebraic scheme, the classical Kruppa equations are derived in section 3. The simplified Kruppa equations are derived in section 4. Section 5 describes in detail the proposed self-calibration method and discusses some implementation issues. Experimental results from a prototype implementation are presented in section 6. The paper concludes with a brief discussion in section 7.

2 Notation and Background

The projection model assumed for the camera(s) 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

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

tutorial by Mohr and Triggs [36] or the relevant chapters of the books by Faugeras [12], Kanatani [26] 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 two images acquired by 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:

$$\begin{aligned} s_1\hat{\mathbf{m}}_\mathbf{1} &= \mathbf{P_1}\hat{\mathbf{M}} \\ s_2\hat{\mathbf{m}}_\mathbf{2} &= \mathbf{P_2}\hat{\mathbf{M}} \end{aligned} \tag{2}$$

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

$$\begin{aligned} \mathbf{P_1} &= [\mathbf{A}|\mathbf{0}] \\ \mathbf{P_2} &= [\mathbf{A}'\mathbf{R}|\mathbf{A}'\mathbf{t}], \end{aligned} \tag{3}$$

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. Matrices **A** and $\mathbf{A}'$ are the 3×3 intrinsic parameters matrices of the two views, each depending on five parameters and having the following well-known form [14]:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \alpha_u & -\alpha_u \cot\theta & u_0 \\ 0 & \alpha_v / \sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \\ \mathbf{A}' &= \begin{bmatrix} \alpha_u' & -\alpha_u' \cot\theta' & u_0' \\ 0 & \alpha_v' / \sin\theta' & v_0' \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{4}$$

The parameters α_u (resp. α_u') and α_v (resp. α_v') correspond to the focal distances in pixels along the horizontal and vertical axes of the image, θ (resp. θ') is the angle between the two image coordinate axes and (u_0, v_0) (resp. (u_0', v_0')) are the coordinates of the image *principal point*, i.e. the

point defined by the intersection of the optical axis and the image plane. The ratio $\frac{\alpha_v}{\alpha_u}$ (resp. $\frac{\alpha_v{}'}{\alpha_u{}'}$) is known as the *aspect ratio*. When both images have been acquired by the same camera, $\theta = \theta'$ and $\frac{\alpha_v}{\alpha_u} = \frac{\alpha_v{}'}{\alpha_u{}'}$, even if the focal lengths are not identical. In practice, θ and θ' are very close to $\frac{\pi}{2}$ for real cameras [12]. Additionally, modern cameras have almost square pixels, therefore in this case the aspect ratio is very close to being equal to one.

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 [51], 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

$$\mathbf{F}\mathbf{K}\mathbf{F}^t = \mathbf{A}'^{*} [\mathbf{t}]_\times ([\mathbf{t}]_\times)^t \mathbf{A}'^{-1}, \tag{10}$$

where **K** is the symmetric matrix $\mathbf{A}\mathbf{A}^t$. This last equation will be employed in subsequent sections for algebraically deriving the Kruppa equations.

3 Deriving the Classical Kruppa Equations

In this section, the well-known Kruppa equations in the case of varying intrinsic parameters are derived in a simple and purely algebraic manner, i.e. without making use of the absolute conic [15, 34] or the plane at infinity [55, 39]. Part of this derivation will be later employed for obtaining the simplified equations proposed in this paper.

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

$$\mathbf{A}^{-t}\mathbf{R}^t([\mathbf{t}]_\times)^t\mathbf{A}'^{-1}\mathbf{e}' = \mathbf{0} \tag{11}$$

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

$$\mathbf{e}' = \lambda\mathbf{A}'\mathbf{t}, \tag{12}$$

where λ is a nonzero scalar. This equation also supplies the following expression for $\mathbf{t}$:

$$\mathbf{t} = \lambda'\mathbf{A}'^{-1}\mathbf{e}', \tag{13}$$

where $\lambda' = 1/\lambda$. Eq. (13) leads to the following relation for the matrix $[\mathbf{t}]_\times$[2]:

$$[\mathbf{t}]_\times = \lambda'\, det(\mathbf{A}'^{-1})\mathbf{A}'^{t}[\mathbf{e}']_\times\mathbf{A}' \tag{14}$$

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

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

where γ is an unknown, nonzero scalar and $\mathbf{K}'$ denotes the symmetric matrix $\mathbf{A}'\mathbf{A}'^t$. Since $\mathbf{F}\mathbf{K}\mathbf{F}^t$ is a symmetric matrix, Eq.(15) corresponds to the following scalar equations obtained by eliminating γ:

$$\begin{aligned}&\frac{\mathbf{F}\mathbf{K}\mathbf{F}^t_{11}}{([\mathbf{e}']_\times\mathbf{K}'([\mathbf{e}']_\times)^t)_{11}} = \frac{\mathbf{F}\mathbf{K}\mathbf{F}^t_{12}}{([\mathbf{e}']_\times\mathbf{K}'([\mathbf{e}']_\times)^t)_{12}} = \frac{\mathbf{F}\mathbf{K}\mathbf{F}^t_{22}}{([\mathbf{e}']_\times\mathbf{K}'([\mathbf{e}']_\times)^t)_{22}} = \\ &= \frac{\mathbf{F}\mathbf{K}\mathbf{F}^t_{13}}{([\mathbf{e}']_\times\mathbf{K}'([\mathbf{e}']_\times)^t)_{13}} = \frac{\mathbf{F}\mathbf{K}\mathbf{F}^t_{23}}{([\mathbf{e}']_\times\mathbf{K}'([\mathbf{e}']_\times)^t)_{23}} = \frac{\mathbf{F}\mathbf{K}\mathbf{F}^t_{33}}{([\mathbf{e}']_\times\mathbf{K}'([\mathbf{e}']_\times)^t)_{33}},\end{aligned} \tag{16}$$

where the subscripts specify indices of matrix elements. Equations (16), however, are linearly dependent since

$$(\mathbf{F}\mathbf{K}\mathbf{F}^t - \gamma[\mathbf{e}']_\times\mathbf{K}'([\mathbf{e}']_\times)^t)\mathbf{e}' = \mathbf{0} \tag{17}$$

As shown in [55, 5], there are only two independent equations among the set of the six equations given by Eq.(16). These equations are first order polynomials in the elements of $\mathbf{K}$ and $\mathbf{K}'$, and therefore of order two in the elements of $\mathbf{A}$ and $\mathbf{A}'$. When using the Kruppa equations for self-calibration,

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

it is common to start by estimating $\mathbf{K}$ (resp. $\mathbf{K}^{'}$) and then using Cholesky decomposition[3] to obtain $\mathbf{A}$ (resp. $\mathbf{A}^{'}$).

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 [15, 34, 31], 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 [56, 5]. In section 4, 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.

4 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 [32]. Towards this end, the Singular Value Decomposition (SVD) [16] of the matrix $\mathbf{F}$ is employed:

$$\mathbf{F} = \mathbf{U}\mathbf{D}\mathbf{V}^t \tag{18}$$

Recalling that $\mathbf{F}$ is of rank two, the diagonal matrix $\mathbf{D}$ has the following form:

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

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 Eq. (18), the epipole in the second image $\mathbf{e}^{'}$ can be deduced very simply. Specifically,

$$\mathbf{F}^t\mathbf{e}^{'} = \mathbf{V}\mathbf{D}^t\mathbf{U}^t\mathbf{e}^{'} = \mathbf{0} \tag{19}$$

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

$$\mathbf{e}^{'} = \delta\mathbf{U}\mathbf{m}, \quad \delta \neq 0 \tag{20}$$

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

$$[\mathbf{e}^{'}]_\times = \mu\mathbf{U}\mathbf{M}\mathbf{U}^t, \tag{21}$$

[3] The Cholesky decomposition of a positive definite matrix $\mathbf{B}$, is a matrix $\mathbf{C}$ s.t. $\mathbf{B} = \mathbf{C}^t\mathbf{C}$ [16].

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

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

Substitution of Eq.(21) into Eq.(14) and then into Eq.(10), yields a new expression for the Kruppa equations:

$$\mathbf{F}\mathbf{K}\mathbf{F}^t = \nu \mathbf{U}\mathbf{M}\mathbf{U}^t\mathbf{K}'\mathbf{U}\mathbf{M}^t\mathbf{U}^t, \;\; \nu \neq 0 \tag{22}$$

Since $\mathbf{U}$ is an orthogonal matrix, left and right multiplication of Eq.(22) by $\mathbf{U}^t$ and $\mathbf{U}$ respectively, yields the following notably simple expression for the Kruppa equations:

$$\mathbf{D}\mathbf{V}^t\mathbf{K}\mathbf{V}\mathbf{D}^t = \nu \mathbf{M}\mathbf{U}^t\mathbf{K}'\mathbf{U}\mathbf{M}^t \tag{23}$$

Because of the simple forms of the matrices $\mathbf{D}$ and $\mathbf{M}$, Eq. (23) corresponds to three linearly dependent equations. Indeed, denoting by $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ the column vectors of $\mathbf{U}$ and by $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ the column vectors of $\mathbf{V}$, the matrix equation (23) 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 matrices $\mathbf{K}$ and $\mathbf{K}'$:

$$\frac{r^2\mathbf{v}_1{}^t\mathbf{K}\mathbf{v}_1}{\mathbf{u}_2{}^t\mathbf{K}'\mathbf{u}_2} = \frac{rs\mathbf{v}_1{}^t\mathbf{K}\mathbf{v}_2}{-\mathbf{u}_2{}^t\mathbf{K}'\mathbf{u}_1} = \frac{s^2\mathbf{v}_2{}^t\mathbf{K}\mathbf{v}_2}{\mathbf{u}_1{}^t\mathbf{K}'\mathbf{u}_1} \tag{24}$$

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. At this point, it should be noted that the simplified Kruppa equations are not symmetric with respect to the pair of images used. In other words, since the fundamental matrix defined by reversing the role of the images in a pair is equal to $\mathbf{F}^t = \mathbf{V}\mathbf{D}\mathbf{U}^t$, the analogous of Eq. (23) becomes $\mathbf{D}\mathbf{U}^t\mathbf{K}'\mathbf{U}\mathbf{D}^t = \nu'\mathbf{M}\mathbf{V}^t\mathbf{K}\mathbf{V}\mathbf{M}^t$, which is different from Eq. (23). This issue is further discussed in section 5.2.

It is worth noting that Eqs. (24) are closely related to the generalization of the Kruppa equations that has been proposed by Luong [33]. Luong has used these equations for demonstrating the equivalence between the constraints of Trivedi [51] and those of Huang and Faugeras [25]. The same

ideas can also be found at the origin of the recent article by Hartley [18], who directly derives the Kruppa equations from the fundamental matrix. Both Luong and Hartley base their developments on changes of rather astute coordinate reference frames, 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 in the case of varying intrinsic parameters. Using an algebraic method, a simpler variant of the Kruppa equations is derived without making use of the absolute conic. As we have demonstrated in [29], the specialization of the simplified Kruppa equations to the case of constant intrinsic parameters yield better calibration results compared to those provided by the well-known algorithm of Zeller and Faugeras [56] which is based on the classical Kruppa equations. In the remainder of the paper, we supply more details regarding a practical algorithm that uses Eqs. (24) for achieving self-calibration and then present experimental results using both simulated and real data.

5 Using the Simplified Kruppa Equations In Practice: Self-Calibration Implementation Details

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 and Faugeras [55, 56], the equations derived in section 4 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. For clarity of notation, each image is identified by a positive index i, its corresponding intrinsic parameters matrix is denoted by $\mathbf{A}_i$ and the product $\mathbf{A}_i {\mathbf{A}_i}^t$ by $\mathbf{K}_i$. Finally, $\mathbf{F}_{ij}$ denotes the fundamental matrix between views i and j, taken in that order.

5.1 Determining an Initial Solution

Let $\mathbf{S}_{\mathbf{F}_{ij}} = [r_{ij}, s_{ij}, {\mathbf{u_1}_{ij}}^t, {\mathbf{u_2}_{ij}}^t, {\mathbf{v_1}_{ij}}^t, {\mathbf{v_2}_{ij}}^t]^t$ be the 14×1 vector formed by the parameters of the SVD of $\mathbf{F}_{ij}$. Let also $\frac{\rho_k(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_i)}{\phi_k(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_j)},\ k = 1 \ldots 3$ be the three ratios defined by Eq.(24). Each pair of images defines a fundamental matrix, which in turn yields the following two polynomial equations regarding the elements of $\mathbf{K}_i$ and $\mathbf{K}_j$:

$$\begin{aligned} \rho_1(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_i)\phi_2(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_j) - \phi_1(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_j)\rho_2(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_i) &= 0 \\ \rho_1(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_i)\phi_3(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_j) - \phi_1(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_j)\rho_3(\mathbf{S}_{\mathbf{F}_{ij}},\, \mathbf{K}_i) &= 0 \end{aligned} \tag{25}$$

The above system of equations is of degree one in ten unknowns defining the $\mathbf{K}_i$ and $\mathbf{K}_j$ matrices. A good initial approximation regarding the position of the principal points in the two images, is to assume that they coincide with the image centers. Additionally, if the aspect ratio and the skew angle θ are assumed to be equal to one and $\frac{\pi}{2}$ respectively, the number of unknowns in Eq.(25) is reduced to two, namely the first elements $K_{i_{11}}$ and $K_{j_{11}}$ of matrices $\mathbf{K}_i$ and $\mathbf{K}_j$, which are related to the two focal lengths α_{u_i} and α_{u_j} [4]. Therefore, the system of Eqs. (25) becomes of degree one in two unknowns and thus it can be solved analytically. The system can have at most two solutions, some of which might be meaningless. More specifically, every solution for $K_{i_{11}}$ and $K_{j_{11}}$ which is such that the associated $\mathbf{K}_i$ and $\mathbf{K}_j$ matrices are not real and positive definite, is discarded. Solutions are also discarded in the case that the related focal lengths are outside a predefined range. Similar closed-form solutions for a pair of focal lengths from the underlying fundamental matrix are described in [5, 27].

Assuming the availability of M images that have been acquired with varying camera intrinsic parameters, a total of $M-1$ fundamental matrices are defined by pairing the i-th image with each of the remaining $M-1$ ones. These matrices give rise to $M-1$ first order systems of the form of Eqs.(25) involving the focal length α_{u_i} of the i-th image and the focal lengths $\alpha_{u_j}, \quad i \neq j$ of all other images. Thus, for the focal length α_{u_i} associated with each image i, there exist at most $2(M-1)$ solutions, the median value of which is taken to be the initial estimate of α_{u_i}.

5.2 Non-Linear Refinement

Let $\pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$ denote the difference of ratios $\frac{\rho_k(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i)}{\phi_k(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_j)} - \frac{\rho_l(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i)}{\phi_l(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_j)}$ and let $\sigma^2_{\pi_{kl}}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$ be its variance. This variance is approximated by [5]

$$\sigma^2_{\pi_{kl}}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j) = \frac{\partial \pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)}{\partial \mathbf{S}_{\mathbf{F}_{ij}}} \mathbf{\Lambda}_{\mathbf{S}_{\mathbf{F}_{ij}}} \frac{\partial \pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)}{\partial \mathbf{S}_{\mathbf{F}_{ij}}}^t, \tag{26}$$

where $\mathbf{\Lambda}_{\mathbf{S}_{\mathbf{F}_{ij}}}$ is the 14×14 covariance matrix associated with $\mathbf{S}_{\mathbf{F}_{ij}}$ and $\frac{\partial \pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)}{\partial \mathbf{S}_{\mathbf{F}_{ij}}}$ is the Jacobian of $\pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$ at $\mathbf{S}_{\mathbf{F}_{ij}}$. This Jacobian is computed directly from the analytic expression for $\pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$. Since $\mathbf{S}_{\mathbf{F}_{ij}}$ is a function of $\mathbf{F}_{ij}$, its covariance matrix $\mathbf{\Lambda}_{\mathbf{S}_{\mathbf{F}_{ij}}}$ is in turn computed from

$$\mathbf{\Lambda}_{\mathbf{S}_{\mathbf{F}_{ij}}} = \frac{\partial \mathbf{S}_{\mathbf{F}_{ij}}}{\partial \mathbf{F}_{ij}} \mathbf{\Lambda}_{\mathbf{F}_{ij}} \frac{\partial \mathbf{S}_{\mathbf{F}_{ij}}}{\partial \mathbf{F}_{ij}}^t, \tag{27}$$

[4] If the aspect ratio and skew angle are known but not equal to one and $\frac{\pi}{2}$ respectively, they can be accounted for by appropriate coordinate systems changes.

[5] Assuming that $\mathbf{x}$ is a random vector with mean $\mathbf{x_0}$ and covariance $\mathbf{\Lambda_x}$, the covariance of vector $\mathbf{y} = \mathbf{f}(\mathbf{x})$ up to first order is equal to $\mathbf{\Lambda_y} = \frac{\partial \mathbf{f}(\mathbf{x_0})}{\partial \mathbf{x_0}} \mathbf{\Lambda_x} \frac{\partial \mathbf{f}(\mathbf{x_0})}{\partial \mathbf{x_0}}^t$; see [12] for details and proof.

where $\mathbf{\Lambda}_{\mathbf{F}_{ij}}$ is the 9×9 covariance matrix of the fundamental matrix $\mathbf{F}_{ij}$[6] and $\frac{\partial \mathbf{S}_{\mathbf{F}_{ij}}}{\partial \mathbf{S}_{\mathbf{F}_{ij}}}$ is the value of the Jacobian of $\mathbf{S}_{\mathbf{F}_{ij}}$ at $\mathbf{F}_{ij}$. 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 [38, 29]. As will soon be clear, the variances $\sigma^2_{\pi_{kl}}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$ are used to automatically weight the residuals $\pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$ according to their uncertainty.

The elements of matrices $\mathbf{K}_i$ are computed from the solution of a non-linear least squares problem, namely

$$\mathbf{K}_1 \dots \mathbf{K}_M = \tag{28}$$
$$argmin_{\tilde{\mathbf{K}}_1 \dots \tilde{\mathbf{K}}_\mathbf{M}} \sum_{i=1}^{M} \sum_{j=1,\ j\neq i}^{M} \frac{\pi^2_{12}(\mathbf{S}_{\mathbf{F}_{ij}},\ \tilde{\mathbf{K}}_\mathbf{i},\ \tilde{\mathbf{K}}_\mathbf{j})}{\sigma^2_{\pi_{12}}(\mathbf{S}_{\mathbf{F}_{ij}},\ \tilde{\mathbf{K}}_\mathbf{i},\ \tilde{\mathbf{K}}_\mathbf{j})} + \frac{\pi^2_{13}(\mathbf{S}_{\mathbf{F}_{ij}},\ \tilde{\mathbf{K}}_\mathbf{i},\ \tilde{\mathbf{K}}_\mathbf{j})}{\sigma^2_{\pi_{13}}(\mathbf{S}_{\mathbf{F}_{ij}},\ \tilde{\mathbf{K}}_\mathbf{i},\ \tilde{\mathbf{K}}_\mathbf{j})} + \frac{\pi^2_{23}(\mathbf{S}_{\mathbf{F}_{ij}},\ \tilde{\mathbf{K}}_\mathbf{i},\ \tilde{\mathbf{K}}_\mathbf{j})}{\sigma^2_{\pi_{23}}(\mathbf{S}_{\mathbf{F}_{ij}},\ \tilde{\mathbf{K}}_\mathbf{i},\ \tilde{\mathbf{K}}_\mathbf{j})}$$

Notice that in the above objective function, the lower value for the second sum's index (i.e. j) is equal to one, ensuring that no bias exists in the intra-order of image pairs, as would be the case if this index counted from $i+1$ (see also section 4). In the most general case, each $\mathbf{K}_i$ is parameterized with five unknowns (see Eq. (4)). However, when additional knowledge regarding the values of some of the intrinsic parameters of each camera is available, less than five unknowns can be used for parameterizing each $\mathbf{K}_i$. Note also that Eq. (28) employs all $M(M-1)/2$ fundamental matrices defined among M different images. Recalling that each fundamental matrix yields two independent constraints and letting N be the number of unknowns defining each $\mathbf{K}_i$, it is straightforward to see that M images yield $M(M-1)$ constraints and involve $N * M$ unknown intrinsic calibration parameters. Thus, in the general case where $N = 5$, the minimum number of images required for self-calibration is equal to six. Obviously, less than six images suffice if known or unknown but constant values for some of the intrinsic parameters can be assumed. Additional constraints provided by using more than the minimum number of required images can improve the accuracy of the solution. The reason for minimizing the sum of the squared ratio differences $\pi_{kl}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$ in Eq.(28) instead of the sum of the squared polynomials of Eqs.(25), is that the former formulation has proved to produce more accurate results (see also [31]). Slightly better quality results is also the reason for including the third of the simplified Kruppa equations (i.e. $\pi_{23}(\mathbf{S}_{\mathbf{F}_{ij}},\ \mathbf{K}_i,\ \mathbf{K}_j)$) in Eq.(28), although it is dependent on the other two [31]. It should also be noted that since Eq.(28) employs the fundamental matrices defined by all possible image pairs, it is not biased towards any of the images of the employed sequence.

The minimization of Eq.(28) is done using a classical Levenberg-Marquardt algorithm [2, 44], using the starting solution computed in the initialization stage. In the case that a priori information in the form of angles or ratios of line segments in the scene is available, it can be incorporated in Eq.(28) as described in [56]. Matrices $\mathbf{A}_i$ are extracted from the corresponding $\mathbf{K}_i$ in three steps. First, ${\mathbf{A}_i}^{-t}$ is computed by employing the Cholesky decomposition of ${\mathbf{K}_i}^{-1}$, then it is transposed and finally inverted to yield $\mathbf{A}_i$.

[6] This covariance matrix is supplied as a by-product of the procedure for estimating $\mathbf{F}$ [7, 60].

6 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. Before proceeding to the description of the conducted experiments, we note that the problem of estimating varying intrinsic calibration parameters is harder compared to that of estimating calibration parameters that are constant for all images involved. One reason for this difficulty is related to the fact that the imperfections of the optic and mechanical apparatus implementing a zoom lens result in more complex image formation, not fully accounted for by the simple pinhole camera model [54, 11]. An additional reason is the larger number of unknowns that need to be estimated. Considering that the skew angle and aspect ratio of a camera do not change with focusing/zooming, the total number of parameters that have to be estimated can be reduced by assuming a priori knowledge of the former pair of parameters for each camera. Estimates of these two parameters can be precomputed by applying the autocalibration method we describe in [29] to an image sequence captured with the camera intrinsic parameters being constant.

6.1 Synthetic Experiments

To quantitatively study the effects of increasing amounts of noise on the recovered intrinsic calibration parameters, 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 retinas. Zero mean Gaussian noise is then added to the coordinates of the resulting retinal points, to account for the fact that in practice, feature extraction algorithms introduce some error when locating image points (i.e. corners). The experimental procedure and the related parameter values for the particular experiments reported here were as follows: The simulated retina is 640×480 pixels, the principal point is different for each image, being within a circular disc of radius 45 pixels centered on the image center (320, 240), the angle between the retinal axes is $\pi/2$, the aspect ratio is equal to one and the focal length varies among images, ranging from 720 to 1440 pixels. A series of 19 rigid displacements of the camera have been simulated and 300 random 3D points have been projected on the corresponding 20 different locations of the simulated retina. The standard deviation of the Gaussian noise added to retinal points was increased from 0 to 2.0 pixels. A non-linear method [59] has then been employed to estimate from the noisy retinal points the 190 fundamental matrices defined by the pairs formed by the 20 simulated images. Following this, the proposed self-calibration method was applied to the estimates of the fundamental matrices to recover the intrinsic calibration matrix corresponding to each view. The minimization of Eq.(28) has been performed using 3 unknowns per image (i.e. the aspect ratio and skew angle have been assumed known), resulting in a total of 60 unknowns for the whole sequence. To ensure that the recovered intrinsic calibration parame-

ters 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. The graphs in Figures 1 show the mean and standard deviation of the relative error for the estimated intrinsic parameters for each image and each noise level. As is evident from the iso contours in the plots, the relative errors in the focal length estimates remain below 9% for noise levels up to 1.3 pixels. Notice also that although the relative error in the principal points estimates is larger than that of the focal lengths, it is acceptable even for the highest image noise levels. These results agree with the findings of several researchers (e.g. [31, 5, 41, 35]), who conclude that estimates of the principal points are more inaccurate compared to those of the focal lengths. It should be mentioned, however, that the accuracy of metric reconstruction does not depend critically on the accuracy of the principal point estimates [5]. Figure 2 illustrates the results obtained from the same simulated data in the restricted case where the principal points were fixed on the image centers and only the focal lengths have been assumed to vary. As can be seen from the iso contours in the plots, the error in the estimates of the focal lengths is in this case smaller, despite that the corresponding error surfaces are less smooth. Finally, it is worth mentioning that the incorporation of the covariances in the minimization process produces results of superior quality compared to those obtained when the uncertainty of the fundamental matrix estimates is not accounted for and all the weights in Eq. (28) have been assumed equal to one. More specifically, the relative error when the covariances are not employed is shown in Figures 3 and 4. Evidently, the relative error surfaces are less smooth in this case and correspond to larger error values compared to those shown in Figures 1 and 2.

6.2 Experiments with Real Images

Several experiments performed with real images are reported in this section. In all of these experiments, point matches have been obtained manually with the aid of [6] and the associated fundamental matrices have been computed with the aid of the technique described in [59]. Throughout all experiments, the skew angle θ and the aspect ratio have been assumed known and equal to $\pi/2$ and 1.0 respectively. In contrast to most published works dealing with self-calibration in the case of varying intrinsic parameters, we report experimental results obtained from the application of the proposed method to short (i.e. around ten frames long) image sequences for which the principal points have been assumed unknown and not fixed on the centers of images. The image sequences employed in the experiments as well as the VRML models recovered using the estimated calibration parameters can be found online at

`http://www-sop.inria.fr/robotvis/personnel/Manolis.Lourakis/autocaldemo.html`

As will soon be demonstrated, the proposed method performs satisfactorily on several real image sequences. If necessary, the intrinsic calibration parameters can be further improved by using the values computed by the proposed algorithm to initiate a global nonlinear minimization scheme which simultaneously estimates the intrinsic parameters and the underlying camera motions [31].

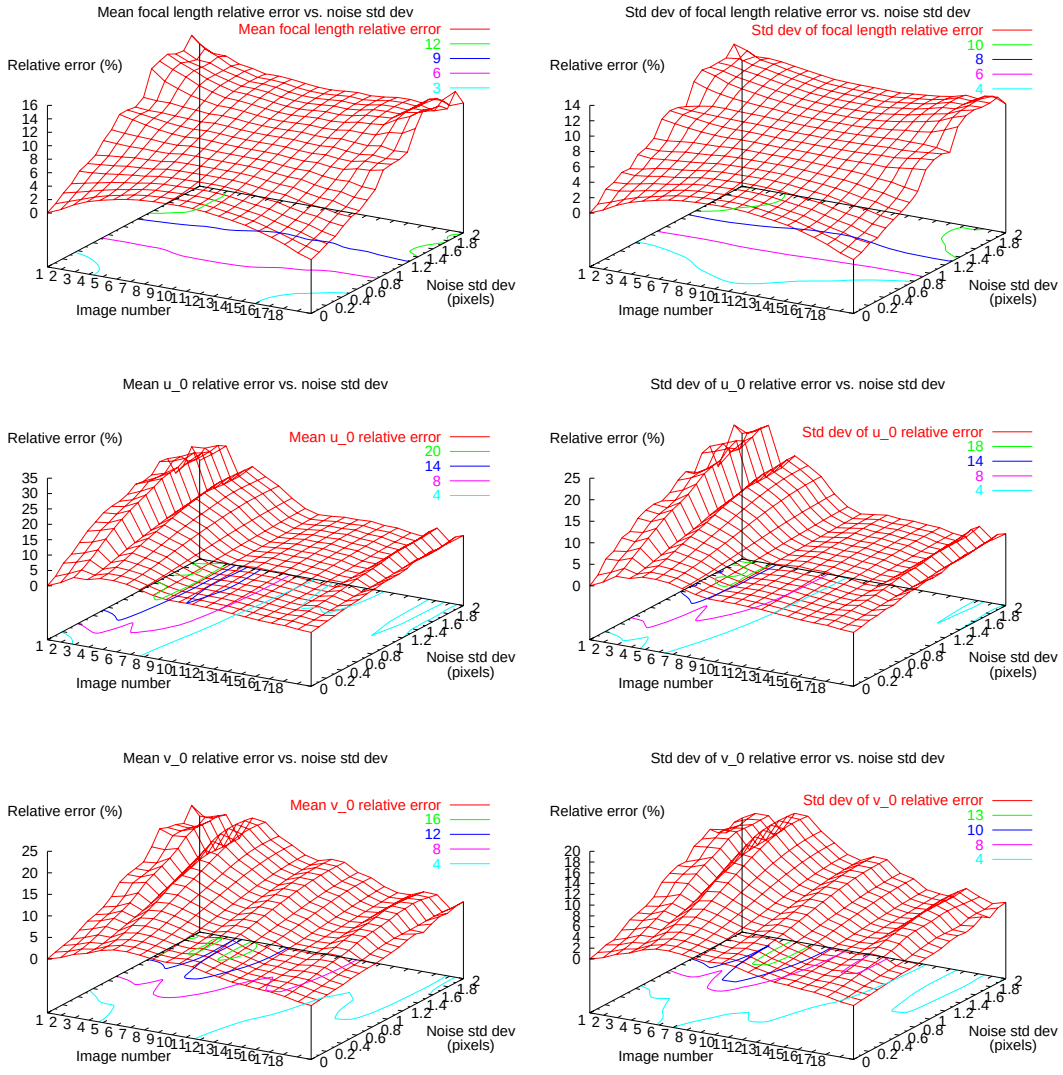


Figure 1: The relative error in the recovered focal lengths and principal points in the presence of noise when employing the covariances. The top row corresponds to the focal lengths, the middle to u_0 and the bottom to v_0. Mean values are shown in the left column, standard deviations in the right.

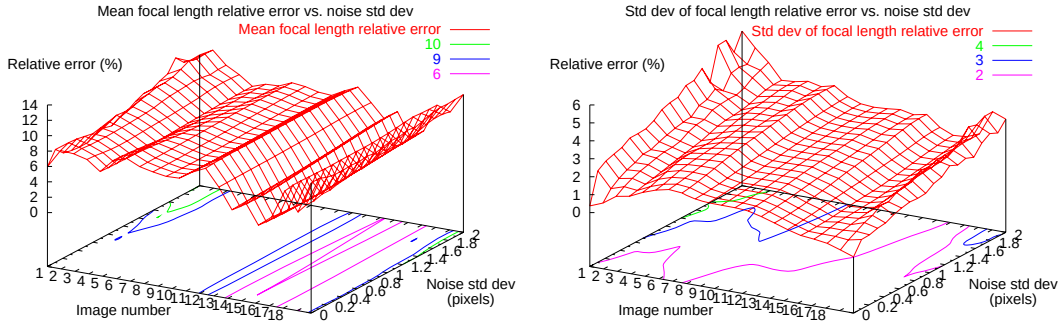


Figure 2: The relative error in the recovered focal lengths in the presence of noise when principal points are assumed fixed on image centers and covariances are employed. The left graph corresponds to mean values, the right to standard deviations.

The first experiment is performed using six 710×430 images of a house in Santorini, Greece, two of which are shown in Figs. 5(a) and (b). These images have been acquired with presumably identical camera intrinsic parameters using a simple fixed zoom pocket camera, therefore they constitute a good test case for checking whether the proposed calibration algorithm estimates constant intrinsic parameters for them. Figure 6 illustrates the focal lengths estimated for each frame in the following three cases: constant intrinsic parameters for all images, variable focal length for each image with the principal points fixed on the image centers and variable focal lengths and principal points for each image. It is clear that in both cases involving varying intrinsic parameters, the estimates of the corresponding focal lengths are almost constant and very close to the focal length estimate of 707 pixels computed by the algorithm we describe in [29], when all intrinsic parameters are assumed to be constant.

The second experiment employs ten 768×512 images depicting the library building at INRIA Sophia-Antipolis, taken with varying but unknown intrinsic calibration parameters. Three of these images are shown in Figure 7. All but the last image were taken from various viewpoints around the building, from human eye level. The last image has been taken from the top of a nearby hill using a large focal length so as to keep the apparent size of the building comparable to that in the other images. Since it is difficult to directly assess the accuracy of the estimated intrinsic calibration parameters, the quality of the latter is evaluated by reconstructing rough 3D wireframe models of the library using the intrinsic parameters computed in three cases [57]. The models consist of sets of adjacent polygons, each of which corresponds to a planar patch on the library's walls. Figures 8 (a) and (d) show top and side views of the model obtained using the intrinsic parameters that were estimated assuming that all images have been taken with constant intrinsic parameters. As can be seen from the reconstructed angle between the walls and the recovered camera viewpoints, the model

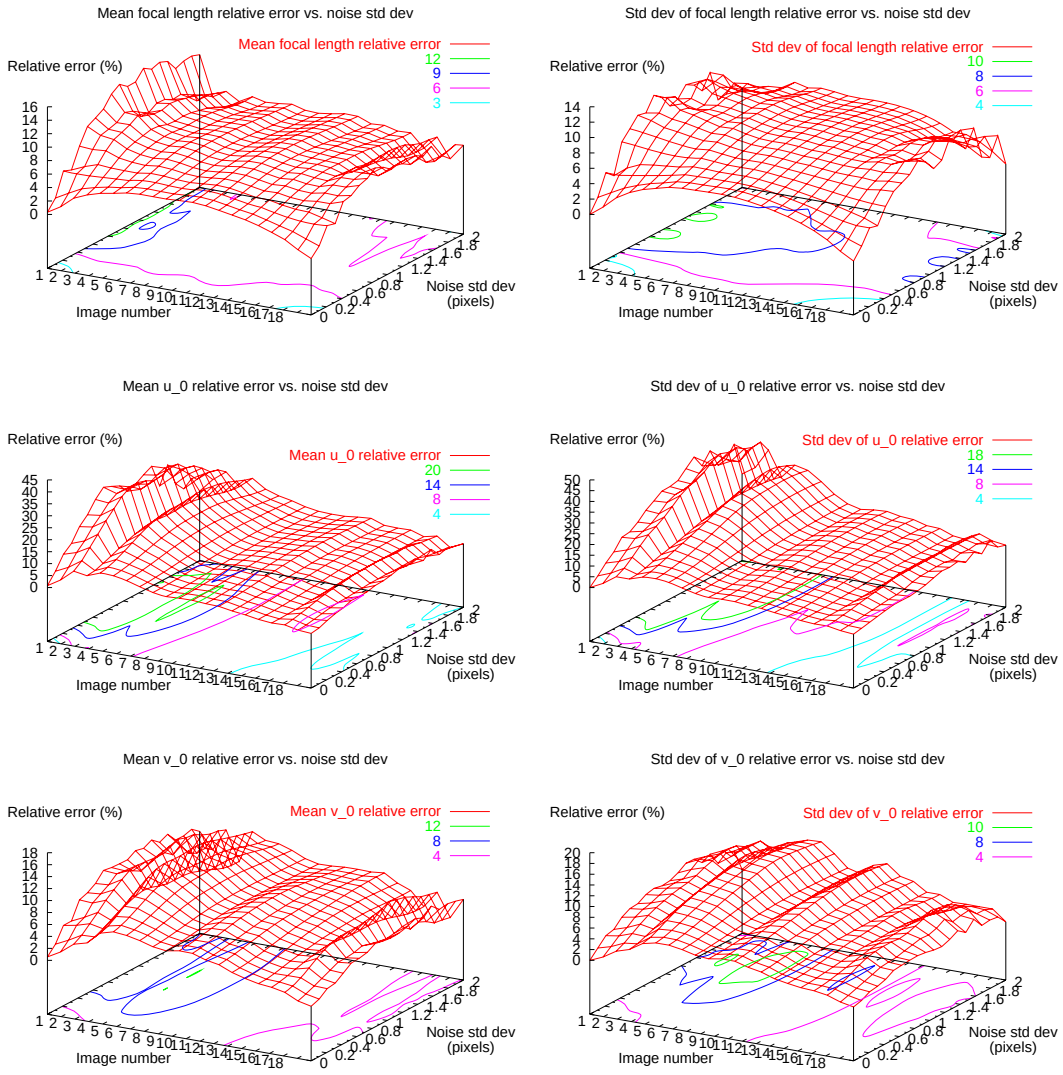


Figure 3: The relative error in the recovered focal lengths and principal points in the presence of noise when the covariances are not employed. The top row corresponds to the focal lengths, the middle to u_0 and the bottom to v_0. Mean values are shown in the left column, standard deviations in the right.

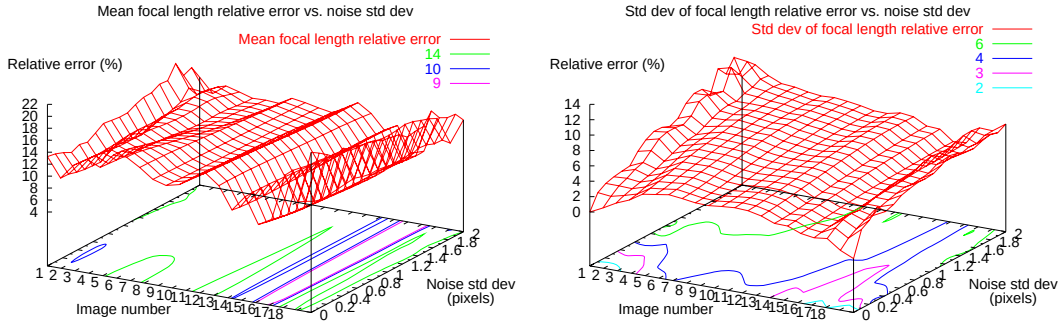


Figure 4: The relative error in the recovered focal lengths in the presence of noise when principal points are assumed fixed on image centers and the covariances are not employed. The left graph corresponds to mean values, the right to standard deviations.

(a) (b)

Figure 5: Frames 1 and 6 of the House sequence.

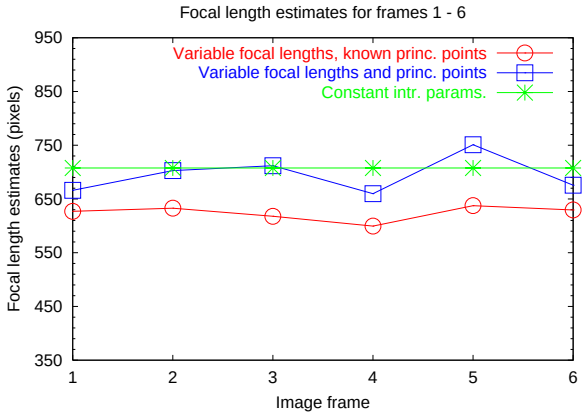


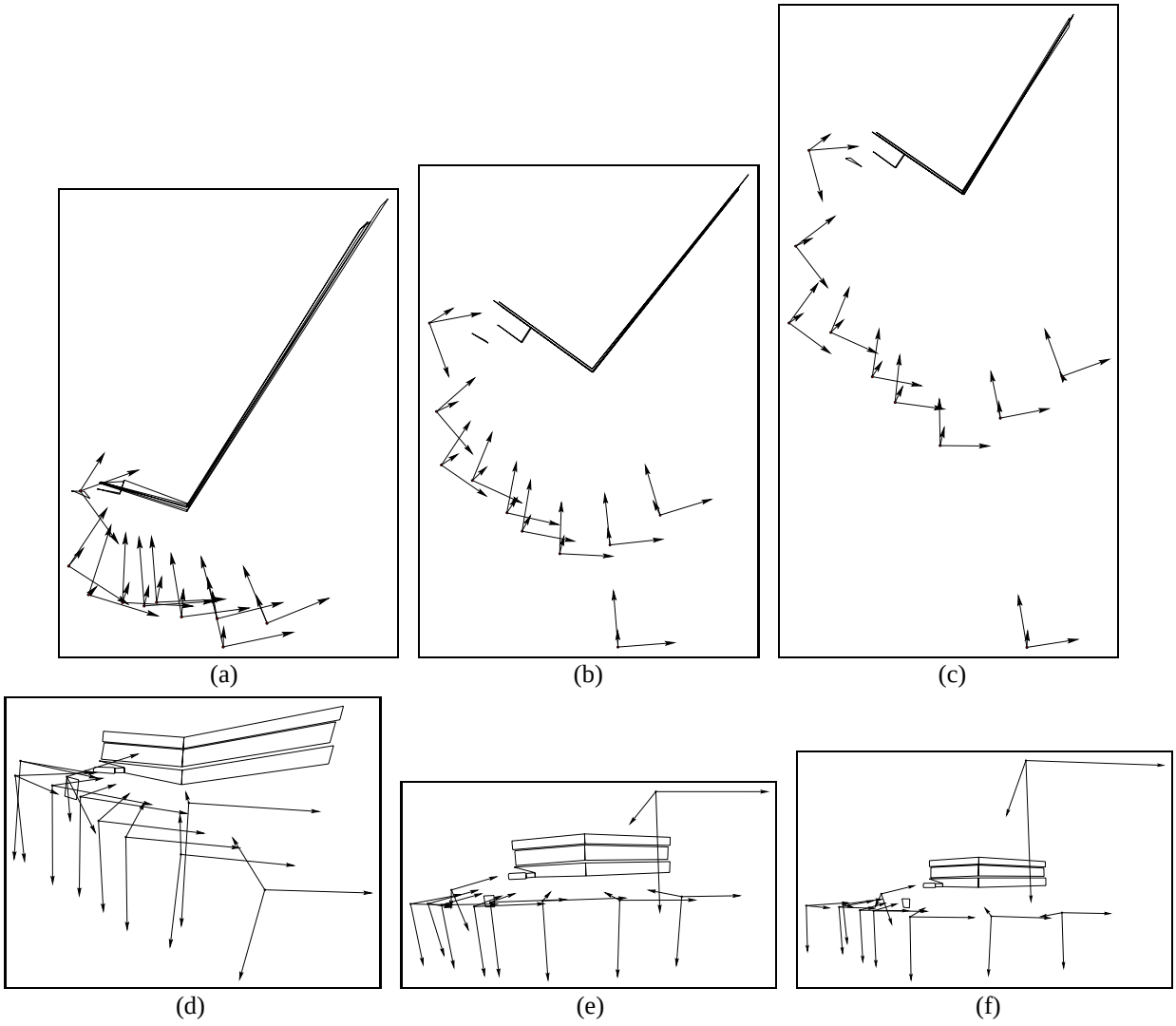
Figure 6: The focal lengths estimated for the house sequence when the intrinsic parameters have been assumed constant and varying with known/unknown principal points.

(a) (b) (c)

Figure 7: Three frames from the INRIA library sequence. Notice the zooming in image (c) which has been taken from the top of a hill.

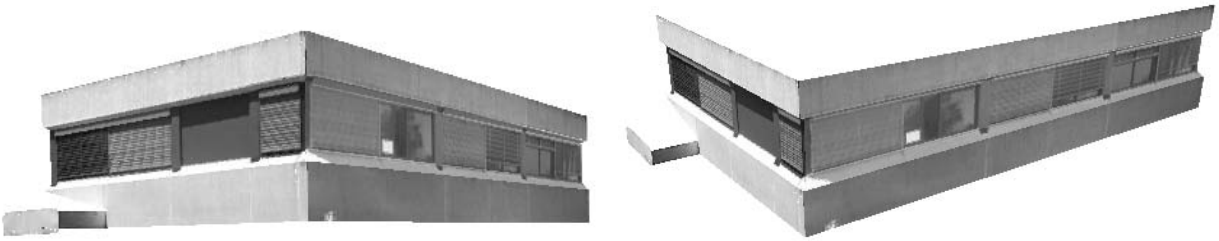
is of poor quality. The model obtained when the focal lengths have been assumed varying among images and the principal points fixed on the image centers is shown in Figs. 8 (b) and (e). Clearly, this model is of much better quality, although the ratio of the recovered wall lengths seems to be incorrect. Figures 8 (c) and (f) show the model obtained when all focal lengths and principal points have been assumed unknown. Based on our knowledge of the employed sequence, the length ratio of the recovered walls as well as the positions of the reconstructed camera viewpoints for this last reconstruction appear to be closer to their true values.

The third experiment is performed with the aid of twelve 886×579 images showing an archaeological site, namely the Minoan Palace of Knossos in Crete, Greece. The images have been acquired



First part

Figure 8: Results for the INRIA library sequence: The wireframe models and the corresponding camera viewpoints reconstructed in three different scenarios are illustrated. (a)-(c) correspond to top views, (d)-(f) to side views. The reconstruction of (a) and (d) has been obtained assuming that the intrinsic parameters are constant for all images, that of (b) and (e) assuming varying focal lengths and fixed principal points that coincide with the image centers and that of (c) and (f) assuming varying focal lengths and principal points. Note the difference in the reconstructed positions of the camera viewpoints, especially for the image taken from the hill's top.



(g) (h)

Second part

Figure 8: Results for the INRIA library sequence (con't): (g) and (h) are perspective views of the reconstruction shown in Figs.8(c) - (f), with texture mapping.

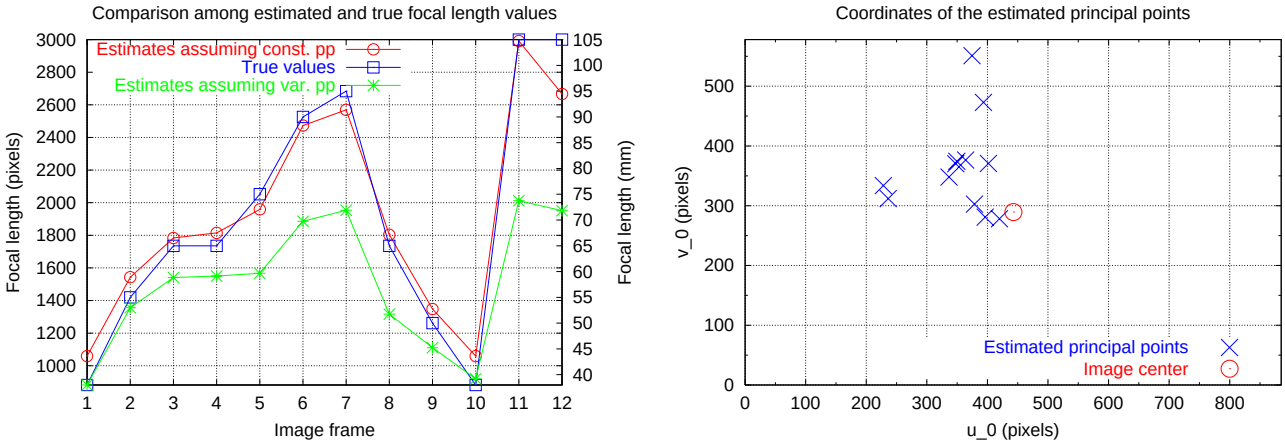
(a) (b) (c)

Figure 9: Three frames (2, 6 and 9) from the Palace of Knossos sequence. Notice the difference in zoom between (a) - (b) and (b) - (c).

from various locations around the building using a pocket camera with a 38-105mm variable zoom lens and then digitized using a scanner. Three of them are shown in Fig. 9. A comparison among the estimated focal lengths and those supplied by the camera during the image acquisition process is shown in Fig. 10(a). Although the latter are of questionable accuracy, they provide a good indication regarding the variation of the focal lengths among images. Fig. 10(a) includes the focal lengths estimated in two cases, namely when the principal points are assumed to be fixed, coinciding with the image centers and when they also are unknown. The graphs corresponding to the estimated focal lengths are plotted in units of pixels against the left vertical axis, while the graph for the true focal lengths is plotted in units of millimeters against the right vertical axis. Clearly, the congruence among the three graphs is fairly good. Fig. 10(b) illustrates the locations of the estimated principal points in the case that the later have been assumed to be unknown (i.e. not coinciding with the image centers). Several of the estimated principal points are as far as few hundreds pixels away from the image center, thus providing yet another example of the difficulty in obtaining stable and accurate principal point estimates. It should be mentioned, however, that the 3D model reconstructed with the focal lengths estimated along with these principal points looked correct, suggesting that 3D reconstruction of acceptable quality can be obtained even with not very accurate intrinsic parameters [5]. Using the intrinsic parameters estimated by the proposed algorithm when the principal points have been assumed fixed on the image centers, the relative camera displacements and the corresponding 3D model have been recovered. Figures 10(c)-(e) illustrate various views of this model. In particular, Fig. 10(c) also illustrates the estimated 3D camera locations.

The fourth experiment is based on twelve 886×579 images of the Koules fortress at the old port of Heraklion, Crete. These images were taken with the same camera used in the previous experiment and three of them are shown in Fig. 11. A comparison among the estimated and the true focal lengths is given in Fig. 12(a), while Fig. 12(b) shows the locations of the estimated principal points in the case that the latter have been assumed to be unknown. Figures 10(c)-(e) show various views of the 3D model reconstructed using the focal lengths estimated by assuming known principal points.

The last experiment is based on ten 886×579 images depicting the west side of FORTH's (Foundation for Research and Technology - Hellas) building in Heraklion, Crete. The camera used in the two previous experiments was again employed and three of the acquired images are shown in Fig. 13. A comparison among the estimated and the true focal lengths is given in Fig. 14(a). When the principal points were assumed to be unknown, the nonlinear minimization step did not converge to a meaningful solution for this sequence. Therefore, only results obtained with the principal points fixed on the image centers are reported. Figures 14(b), (c) show different views of the 3D model reconstructed using the estimated focal lengths. Notice that in this experiment, parts of the building are difficult to reconstruct accurately since they appear in very few images due to occlusions. These occlusions are mainly caused by the pillars in the foreground of the scene.



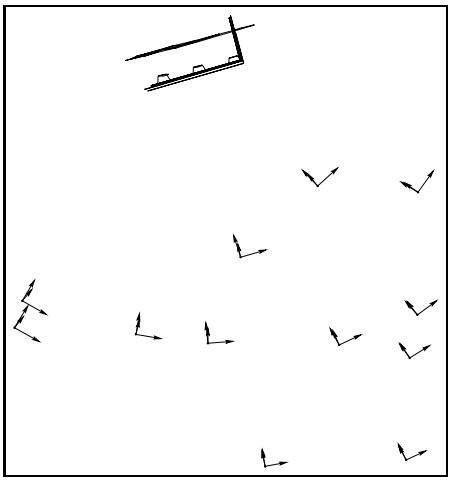
(a) (b)

First part

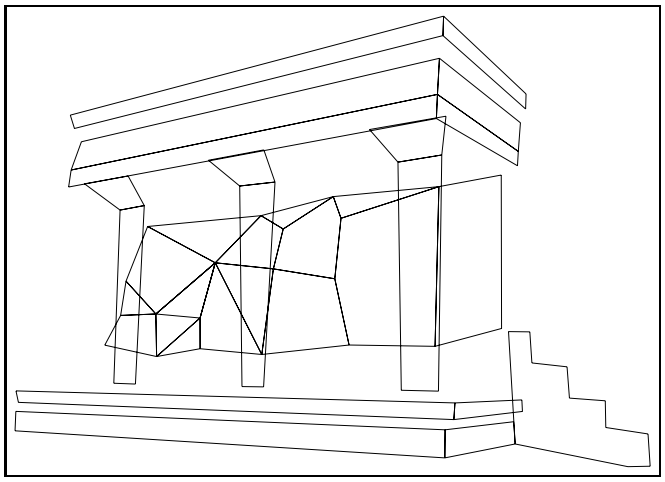
Figure 10: Results for the Knossos sequence: (a) Comparison among the true focal lengths and those estimated by assuming fixed and variable principal points and (b) the coordinates of the estimated principal points.

7 Conclusions

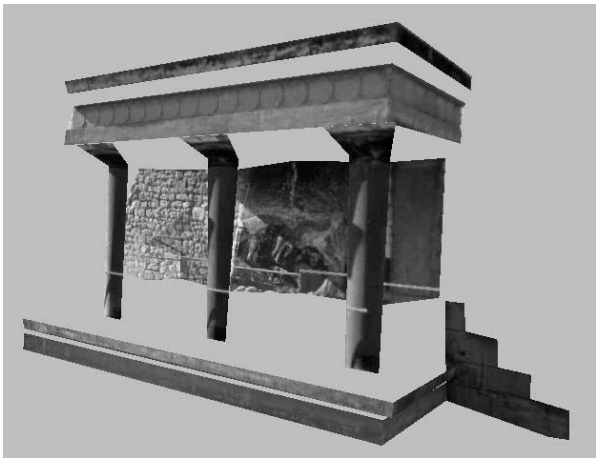
Camera calibration is essential for a wide variety of vision tasks. In this paper, a novel method for self-calibration assuming varying intrinsic parameters has been proposed. The method employs a simplification of the Kruppa equations which relies solely on the SVD of the fundamental matrix and avoids recovering noise-sensitive quantities such as the epipoles. The simplified Kruppa equations form the basis for a non-linear minimization scheme that explicitly takes into consideration the uncertainty associated with the employed fundamental matrices and yields the intrinsic calibration parameters for each image. Bias towards any particular image is avoided and no projective reconstruction is necessary. This last feature is very convenient when images cannot be registered in a single projective frame. Detailed experimental results from the application of the proposed method to synthetic data as well as real image sequences have been reported. Future work will focus on alternative ways for obtaining better initial estimates of the intrinsic parameters using the polynomial constraints defined by Eqs. (25), so as to initialize the non-linear minimization process closer to the correct solution (see also section 5.1). Another possible direction concerns the study of critical camera motions which render the simplified Kruppa equations degenerate.



(c)



(d)



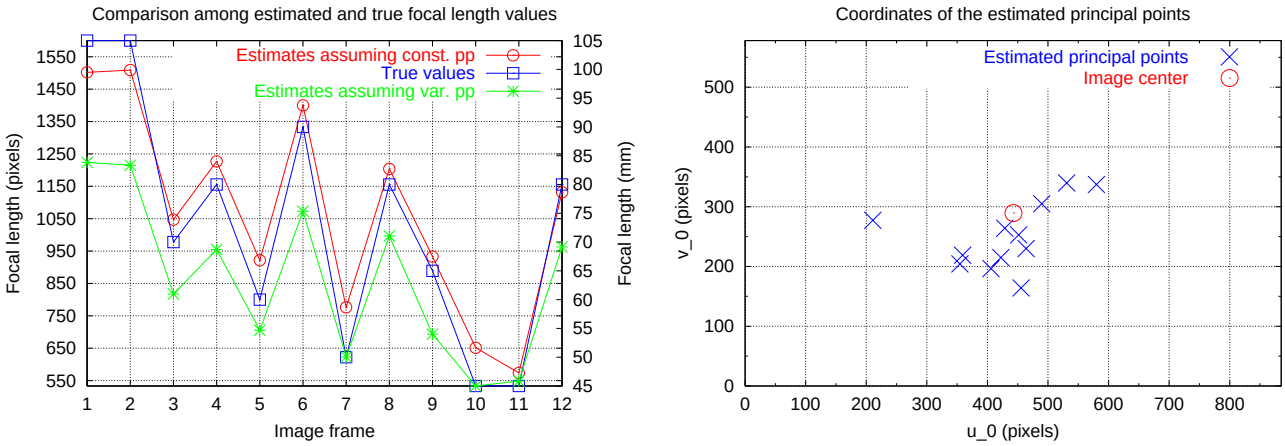
(e)

Second part

Figure 10: Results for the Knossos sequence (con't): (c), (d) top and side views of the reconstructed 3D model and (e) view of the reconstructed model with texture mapping.

(a) (b) (c)

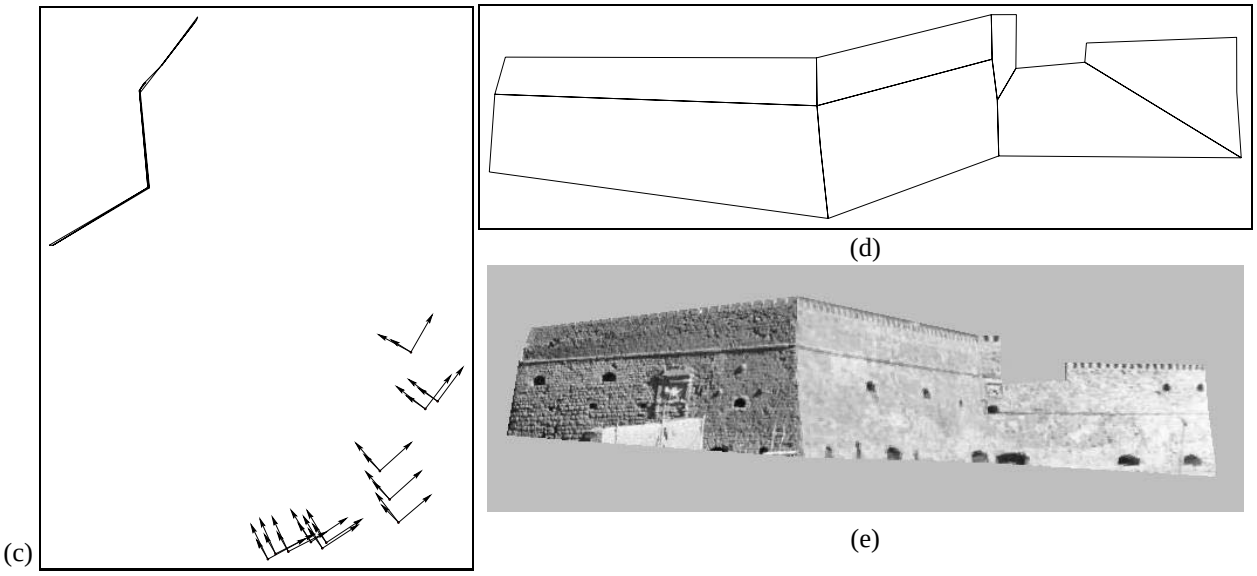
Figure 11: Three frames (1, 10 and 12) from the Koules sequence. Notice the different zooming among images.



(a) (b)

First part

Figure 12: Results for the Koules sequence: (a) Comparison among the true focal lengths and those estimated by assuming fixed and variable principal points and (b) the coordinates of the estimated principal points.



(c) (d) (e)

Second part

Figure 12: Results for the Koules sequence (con't): (c), (d) top and side views of the reconstructed 3D model and (e) view of the reconstructed model with texture mapping.

(a) (b) (c)

Figure 13: Three frames (2, 3 and 10) from the FORTH sequence.

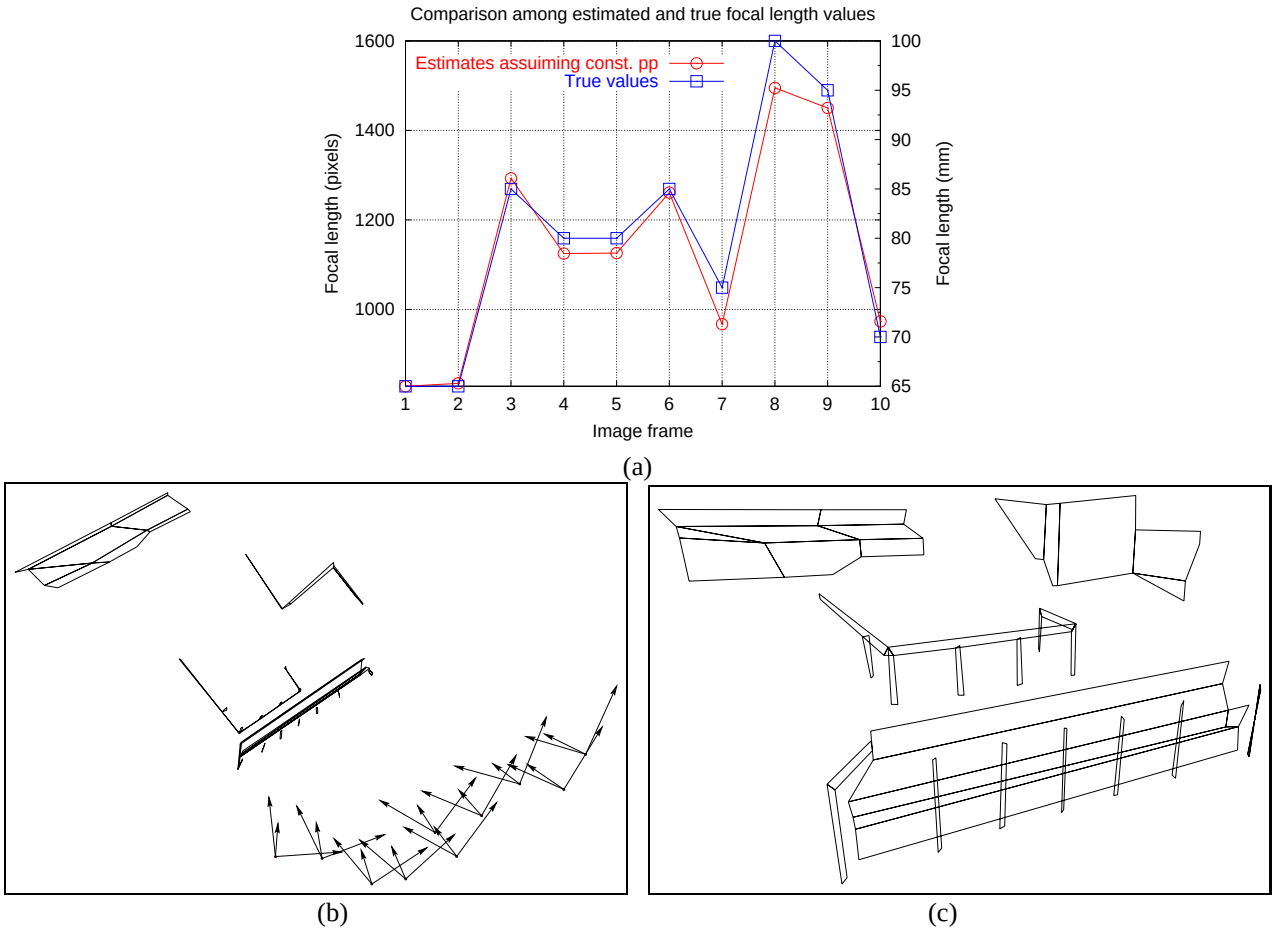


Figure 14: Results for the FORTH sequence: (a) Comparison among the true focal lengths and those estimated by assuming fixed principal points and (b), (c) top and side views of the reconstructed 3D model.

Acknowledgments

The authors thank Sylvain Bougnoux for providing them his 3D reconstruction software.

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] P. Beardsley, P.H.S. Torr, and A. Zisserman. 3D model acquisition from extended image sequences. In *In Proc. of ECCV'96*, pages 683–695, 1996.

[5] 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.

[6] 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.

[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] 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.

[9] 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.

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

[11] Reyes Enciso, Thierry Viéville, and Olivier Faugeras. Approximation du changement de focale et de mise au point par une transformation affine à trois paramètres. *Traitement du Signal*, 11(5):361–372, 1994.

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

[13] 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.

[14] 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.

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

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

[17] 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.

[18] 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.

[19] Richard Hartley. Self-calibration from multiple views with a rotating camera. In J-O. Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision*, volume 800-801 of *Lecture Notes in Computer Science*, pages 471–478, Stockholm, Sweden, May 1994. Springer-Verlag.

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

[21] 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.

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

[23] A. Heyden and K. Åström. Flexible calibration: Minimal cases for auto-calibration. In *Proceedings of the 7th International Conference on Computer Vision*, volume 1, pages 350–355, Kerkyra, Greece, 1999. IEEE Computer Society, IEEE Computer Society Press.

[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] 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.

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

[27] K. Kanatani and C. Matsunaga. Closed-form expression for focal lengths from the fundamental matrix. In *Proc. of the 4th Asian Conference on Computer Vision*, volume I, pages 128–133, January 2000.

[28] 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.

[29] Manolis I.A. Lourakis and Rachid Deriche. Camera self-calibration using the singular value decomposition of the fundamental matrix: From point correspondences to 3D measurements. Research Report 3748, INRIA Sophia-Antipolis, August 1999.

[30] Manolis I.A. Lourakis and Rachid Deriche. Camera self-calibration using the singular value decomposition of the fundamental matrix. In *Proc. of the 4th Asian Conference on Computer Vision*, volume I, pages 403–408, January 2000.

[31] 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.

[32] Q.-T. Luong and O.D. Faugeras. On the determination of epipoles using cross-ratios. *CVGIP: Image Understanding*, 71(1):1–18, July 1998.

[33] 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.

[34] 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.

[35] P.R.S. Mendonça and R. Cipolla. A simple techinique for self-calibration. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, volume I, pages 500–505, 1999.

[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] T. Papadopoulo and M.I.A. Lourakis. Estimating the jacobian of the singular value decomposition: Theory and applications. Research report, INRIA Sophia-Antipolis, 2000. In preparation.

[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, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. *The International Journal of Computer Vision*, 32(1):7–25, August 1999.

[42] 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.

[43] Marc Pollefeys, Luc Van Gool, and Marc Proesmans. Euclidean 3D reconstruction from stereo sequences with variable focal lenghts. In Bernard Buxton, editor, *Proceedings of the 4th European Conference on Computer Vision*, volume I, pages 31–42, Cambridge, UK, April 1996.

[44] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1988.

[45] C. Rothwell, O. Faugeras, and G. Csurka. A comparison of projective reconstruction methods for pairs of views. *Computer Vision and Image Understanding*, 68(1):37–58, October 1997.

[46] 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.

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

[48] P. Sturm. Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. In *10th British Machine Vision Conference*, pages 63–72, Nottingham, England, September 1999.

[49] Peter Sturm. *Vision 3D non calibrée. Contributions à la reconstruction projective et étude des mouvements critiques pour l'auto-calibrage.* PhD thesis, INPG, Grenoble, France, December 1997.

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

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

[52] 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.

[53] 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.

[54] Reg G. Willson. *Modeling and Calibration of Automated Zoom Lenses*. PhD thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, 1994. CMU-RI-TR-94-03.

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

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

[57] Z. Zhang. Motion and structure from two perspective views: From essential parameters to Euclidean motion via fundamental matrix. *Journal of the Optical Society of America A*, 14(11):2938–2950, 1997.

[58] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Proc. of ICCV'99*, pages 666–673, Kerkyra, Greece, sep. 1999.

[59] 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.

[60] Zhengyou Zhang. Determining the epipolar geometry and its uncertainty: a review. *The International Journal of Computer Vision*, 27(2):161–195, March 1998.

[61] 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