

Uncertainty analysis of 3D reconstruction from uncalibrated views

E. Grossmann*, J. Santos-Victor

Instituto de Sistemas e Robótica, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal

Received 9 December 1998; received in revised form 23 July 1999; accepted 25 October 1999

Abstract

We consider reconstruction algorithms using points tracked over a sequence of (at least three) images, to estimate the positions of the cameras (*motion* parameters), the 3D coordinates (*structure* parameters), and the calibration matrix of the cameras (*calibration* parameters). Many algorithms have been reported in literature, and there is a need to know how well they may perform. We show how the choice of assumptions on the camera intrinsic parameters (either fixed, or with a probabilistic prior) influences the precision of the estimator. We associate a Maximum Likelihood estimator to each type of assumptions, and derive analytically their covariance matrices, independently of any specific implementation. We verify that the obtained covariance matrices are realistic, and compare the relative performance of each type of estimator. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Uncalibrated 3D reconstruction; Maximum-likelihood estimation; Covariance matrix

1. Introduction

The problem of 3D reconstruction from images has drawn considerable attention. We focus on the problem of reconstruction from *matched points* (corners). The parameters of interest are the *structure parameters*, i.e. the 3D coordinates of the points, the *motion parameters* that describe the positions of the cameras; and the *calibration parameters* that describe the intrinsic characteristics of the used sensors. The case of known intrinsic parameters has been thoroughly studied in photogrammetry [1]. Work on uncalibrated reconstruction progressed dramatically in recent years with the works of Hartley [2], Faugeras [3], Maybank [4], Pollefeys et al. [5], who showed how to obtain projective, affine, and, finally, euclidean reconstructions from uncalibrated views. We are interested in euclidean reconstruction. Many algorithms have been proposed, differing, e.g. on the assumptions concerning the calibration parameters and/or motion [6]. Studies of the precision of the estimation of the “fundamental matrix” [7] and “trifocal tensor” [8], which represent multilinear constraints that tracked 2D features must verify can be found in Refs. [9–11]. A study of critical (pathological) cases for self-calibration can be found in Ref. [12], and the achievable precision in the calibrated case is addressed in Ref. [13].

In this paper, we study the precision with which 3D points, camera orientation, position and calibration are estimated. In some studies [14,15] some intrinsic parameters are fixed to nominal values. We want to compare, in terms of precision, the effect of these assumptions and the precision achieved in the calibrated case. One contribution of this paper is to compare the precisions of calibrated and uncalibrated reconstruction. Although the former always performs better, experimentation shows that when more than ten images are available uncalibrated reconstruction performs honorably.

Errors in the localization of image features introduce errors in the reconstruction. Some algorithms are numerically unstable, intrinsically, or in conjunction to particular setups of points and/or of cameras. However, an in-depth study of the precision of these algorithms has not been presented. The issue of the accuracy of uncalibrated reconstruction has been raised and studied repeatedly, but always associated to a particular algorithm. Our aim is to give a more general treatment to the question, while remaining as independent as possible of any particular implementation.

1.1. Scope of the paper

Most algorithms combine an “algebraic” part, and an optimization part that solves for a Maximum Likelihood [2] (or related [16]) estimate. Maximum Likelihood (ML) and related estimators are often reported [16] to converge to the solution only if started close from it. It is the purpose of

* Corresponding author.

E-mail addresses: etienne.jasv@isr.ist.utl.pt (E. Grossmann); jasv@isr.ist.utl.pt (J. Santos-Victor).

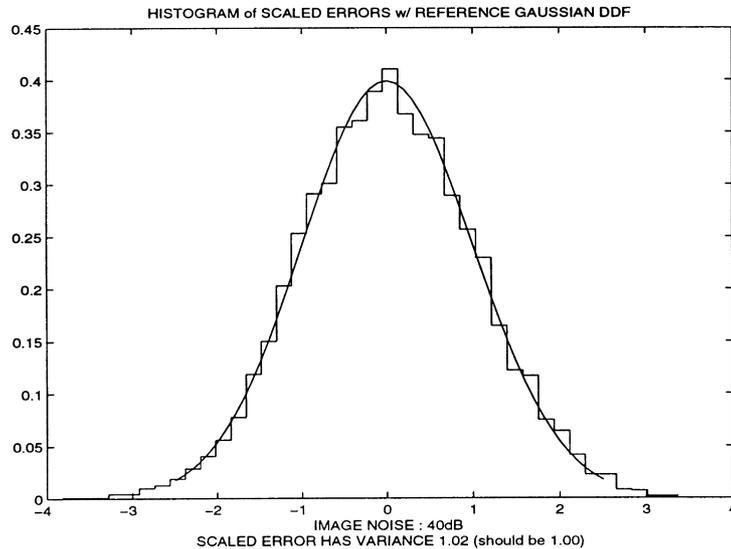


Fig. 1. Histogram of scaled errors of the ML estimator. In abscissa is the error, divided by the theoretical standard deviation. The Gaussian density function is superposed for comparison. The observed variance is 1.02, while parameters \mathcal{X} , \mathcal{W} , \mathcal{T} and \mathbf{K} have variances in [0.98,1.09].

the “algebraic” algorithm to provide the starting position. In this paper, we study the precision of the ML-like estimator, *not* that of the algebraic algorithm. The true parameters are considered as random variables with a distribution that is defined from the observations. The estimator is defined by the observation model, independently from any specific algorithm; we derive analytically its covariance matrix in various cases of interest: we distinguish the cases in which only the observations are available (ML estimation) and those where some knowledge of the estimated quantities is available a priori. Amongst the later, we further distinguish the cases of probabilistic knowledge (maximum a posteriori estimation), and that of “exact” knowledge, where some parameters are fixed.

When estimating all parameters from only the observations, the estimation is often numerically ill posed. For example, in Ref. [14] some intrinsic parameters are highly correlated with some of the motion parameters, and the focal length is correlated to the depth (cinema uses the fact that zooming is almost indistinguishable from forward motion).

If some calibration parameters are fixed, they may be removed from the estimated vector. This simplifies the study and implementation of the estimator, and—presumably—ameliorate the numerical stability. Typical assumptions are that pixels are rectangular or square, or that the principal point coincides with the image center [5,15]. We verify in Section 5.1 the effect on precision of fixing the intrinsic parameters, either to values obtained from a pre-calibration step or to nominal values (corresponding to square pixels and centered principal point).

Finally, the likelihood function may be modified to take into account a priori knowledge expressed probabilistically, e.g. assuming that structure or calibration follow a known distribution. One then performs *maximum a posteriori*

(MAP) estimation. A prior on structure serves most often to retrieve precisely the intrinsic parameters, and is then called *calibration from a known object*.

A prior on the calibration parameters, may come either from a previous calibration step, or from assuming that the camera parameters follow a “nominal” distribution, e.g. the expected value of the principal point is the center of the image, and that its standard deviation is approximately 10% of the image size.¹ This is the probabilistic counterpart of fixing the principal point to image center, in Section 2.2. In terms of the theoretical precision, priors are preferable to fixed parameters.

We will write analytically the covariance matrices corresponding to the studied cases in Eqs. (16)–(18). The diagonal terms correspond to the variances of the individual estimated parameters. The validity of our analytical expressions is verified by comparing the theoretical and the observed behavior of a reconstruction algorithm, in Section 5.1. One important contribution of this paper lies in showing how big the variances of the considered estimators are in practice.

2. The estimation problem

2.1. Notations

We now define the notations used throughout the paper. We consider that a set of P points has been tracked over a sequence of N images. The following notation is adopted:

- $p \in \{1, 2, \dots, P\}$ and $n \in \{1, 2, \dots, N\}$ are the indices used for numbering points and images, respectively.
- $\mathbf{x}_p \in \mathbb{R}^3$ is the vector of the coordinates, in the world

¹ Lenz and Tsai [18] cite values of this order of magnitude.

frame, of the p th point. Its components are $i \in \{1, 2, 3\}$. The symbol \mathcal{X} shall denote the 3D coordinates of all the points $\mathbf{x}_1, \dots, \mathbf{x}_P$.

The projection of these 3D points in the image depends on the relative orientation and position of the camera. Let

- $\mathbf{A}_n = [\mathbf{a}_{n1} \mathbf{a}_{n2} \mathbf{a}_{n3}]^T$ be the rotation matrix relating world coordinates to coordinates in n th image frame. It can be uniquely defined by three parameters \mathcal{W} will represent the orientation of all the camera frames, $\mathbf{w}_1, \dots, \mathbf{w}_N$.
- \mathbf{T}_n be the coordinates of the world frame origin, expressed in the n th camera frame. \mathcal{T} will represent the positions of all the coordinates $\mathbf{T}_1 \dots \mathbf{T}_N$.

Assuming that the camera has unit focal length, square pixels and a centered principal point, the p th point \mathbf{x}_p , produces the (noiseless) observations $\tilde{\mathbf{u}}_{np} = (\tilde{u}_{np1}, \tilde{u}_{np2})$:

$$\tilde{u}_{npk} = \frac{\mathbf{a}_{nk} \cdot \mathbf{x}_p + t_{nk}}{\mathbf{a}_{n3} \cdot \mathbf{x}_p + t_{n3}} \quad k \in \{1, 2\} \quad (1)$$

taking into account the intrinsic parameters and noise yields:

$$\mathbf{u}_{np} = \mathbf{B} \tilde{\mathbf{u}}_{np} + \mathbf{C} + \boldsymbol{\epsilon}_{np} \quad (2)$$

with

- $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2]^T$ the 2×2 matrix that models the skew, pixel size and camera focal length.
- $\mathbf{C} = [c_1, c_2]^T$ the pixel coordinate of the principal point.
- $\boldsymbol{\epsilon}_{np} = [\epsilon_{np1}, \epsilon_{np2}]^T$ the observation noise, which is assumed to have Gaussian, independent and identically distributed terms, with known variance σ^2 .

Let \mathcal{U} denote all the observations u_{npk} , for $k \in \{1, 2\}$, $p \in \{1 \dots P\}$ and $n \in \{1 \dots N\}$, and $\mathcal{U}^* = \mathcal{U} - \boldsymbol{\epsilon}$ the noiseless observations. The intrinsic parameters, \mathbf{B} and \mathbf{C} , will be noted \mathbf{K} . An asterisk denotes the true values of the parameters, \mathcal{X}^* , \mathcal{W}^* , \mathcal{T}^* and \mathbf{K}^* . The problem is defined as estimating the structure, camera orientation and position, and intrinsic parameters, from the observations \mathcal{U} . We write as a single vector, all the parameters: $\boldsymbol{\Theta} = (\mathcal{X}, \mathcal{W}, \mathcal{T}, \mathbf{K})$. For a given $\boldsymbol{\Theta}$, the *prediction* of the (n, p, k) th observation is defined as:

$$v_{npk}(\boldsymbol{\Theta}) = \mathbf{b}_k \tilde{\mathbf{u}}_{np}(\boldsymbol{\Theta}) + c_k \quad \text{where} \quad \tilde{u}_{npk}(\boldsymbol{\Theta}) = \frac{\mathbf{a}_{nk} \cdot \mathbf{x}_p + t_{nk}}{\mathbf{a}_{n3} \cdot \mathbf{x}_p + t_{n3}} \quad (3)$$

2.2. Maximum likelihood estimator

With the observation model defined in Eq. (2), the probability density of observing \mathbf{u}_{np} , for a parameter vector $\boldsymbol{\Theta}$ is

$$P(\mathbf{u}_{np} | \boldsymbol{\Theta}) = \frac{1}{2\pi\sigma^2} e^{-\|\mathbf{v}_{np}(\boldsymbol{\Theta}) - \mathbf{u}_{np}\|^2 / 2\sigma^2}$$

and the conditional probability of observing \mathcal{U} given the parameters $\boldsymbol{\Theta}$, is $P(\mathcal{U} | \boldsymbol{\Theta}) = \prod_{np} P(\mathbf{u}_{np}, \boldsymbol{\Theta})$. The maximum

likelihood (ML) estimator is defined as the function that associates to the observations \mathcal{U} , the parameter $\hat{\boldsymbol{\Theta}}$ that maximizes $P(\mathcal{U} | \boldsymbol{\Theta})$, or, equivalently, which minimizes $Q(\mathcal{U}, \boldsymbol{\Theta}) = -\log(P(\mathcal{U} | \boldsymbol{\Theta}))$. The function Q has the more convenient form:

$$Q(\mathcal{U}, \boldsymbol{\Theta}) = \sum_{npk} \frac{1}{2\sigma^2} (u_{npk} - v_{npk}(\boldsymbol{\Theta}))^2 + \text{Constant} \quad (4)$$

where n ranges from 1 to N , p ranges from 1 to P and k ranges from 1 to 2; these ranges are used for all subsequent sums and products over n, p and k .

In our case, this function does *not* have a unique minimum: it is well known that the reconstruction is defined only up to a similarity transformation. A way of resolving the ambiguity is to constrain the structure parameters to be centered ($\sum_p \mathbf{x}_p = \mathbf{0}_3$) and have unit mean norm ($\sum_p \|\mathbf{x}_p\|^2 = 3P$), the camera matrix \mathbf{B} to be lower triangular, and the first camera frame to coincide with the world frame ($\mathbf{A}_1 = \mathbf{I}_3$). The restricted parameter set is defined as the zeros of the function:

$$S(\boldsymbol{\Theta}) = \left[\sum_p \|\mathbf{x}_p\|^2 - 3P, \sum_p x_{p1}, \sum_p x_{p2}, \sum_p x_{p3}, a_{12}, a_{13}, a_{23} \right]^T \quad (5)$$

There are still some critical setups yielding a continuum of ML estimates, even within the set $S^{-1}(\{0_4\})$. Uniqueness conditions have been studied in Ref. [12]. In this article, we consider that the minima of Q that verify $S(\boldsymbol{\Theta}) = 0$ are isolated. Note that Eq. (5) does not impose that $\mathbf{A}_1 = \mathbf{I}_3$, but rather that $\mathbf{A}_1 \in \{\text{diag}(1, 1, 1), \text{diag}(1, -1, -1), \text{diag}(-1, -1, 1), \text{diag}(-1, 1, -1)\}$. Because this set contains only isolated points, the analysis that follows is not influenced. The maximum likelihood estimate is defined by:

$$\hat{\boldsymbol{\Theta}} = \arg \min_{\boldsymbol{\Theta}} Q(\mathcal{U}, \boldsymbol{\Theta}) \quad \text{subject to } S(\boldsymbol{\Theta}) = 0$$

2.3. Maximum a posteriori estimator

In some situations, one may have some prior knowledge concerning the value of the parameters we are estimating. In what follows, this knowledge is expressed by assuming that the true parameter vector $\boldsymbol{\Theta}^*$ is a Gaussian random variable, independent of the observation noise, of known mean $\bar{\boldsymbol{\Theta}}$ and covariance $\boldsymbol{\Sigma}$, which we write as $\boldsymbol{\Theta}^* \sim N(\bar{\boldsymbol{\Theta}}, \boldsymbol{\Sigma})$. Although using a Gaussian prior is not always realistic, the quantities we estimate in this paper will be parametrized in such a way that a Gaussian prior is reasonable (Section 2.5).

Under these assumptions, using the Bayes rule [17], we can write the posterior probability $P_{\text{post}}(\boldsymbol{\Theta} | \mathcal{U})$ of observing

\mathcal{U} when the parameters vector is Θ :

$$P_{\text{post}}(\Theta|\mathcal{U}) \propto \prod_{np} \left(\frac{1}{2\pi\sigma^2} e^{-\|v_{np}(\Theta) - u_{np}\|^2/2\sigma^2} \right) \\ \times \frac{1}{\sqrt{(2\pi)^M |\Sigma|}} e^{-1/2(\Theta - \bar{\Theta})^T \Sigma^{-1} (\Theta - \bar{\Theta})},$$

the inverted logarithm of this function becomes

$$Q_{\text{post}}(\mathcal{U}, \Theta) = \sum_{npk} \frac{1}{2\sigma^2} (u_{npk} - v_{npk}(\Theta))^2 \\ + \frac{1}{2} (\Theta - \bar{\Theta})^T \Sigma^{-1} (\Theta - \bar{\Theta}) + \text{Constant} \quad (6)$$

In practice, one most often has a prior on a subset of the parameters only. We can consider that Θ is split in $[\Theta_1, \Theta_2]$, that no knowledge on Θ_1 is available, but that we have a prior such as $\Theta_2 \sim N(\bar{\Theta}_2, \Sigma_2)$. One then obtains a cost function similar to Eq. (6), but with Σ replaced by $\text{diag}(0, \Sigma_2)$ and Σ^{-1} by $\text{diag}(0, \Sigma_2^{-1})$.

It is important to note that the use of a probabilistic prior may suppress the need for constraints (5), since Q_{post} (taken as a function of Θ), contrarily to Q , may have an isolated minima. For example, a prior on the structure parameters \mathcal{X} “fixes the scale and orientation” and removes the need for the constraints. A prior on calibration parameters \mathbf{K} , on the contrary, does not ensure that Q_{post} has isolated minima.

2.4. Estimator with fixed parameters

One sometimes fixes some of the parameters, that otherwise could be estimated, for example if “good” values are known beforehand. Doing so may improve the numerical aspects of the estimation problem if one removes parameters that are redundant. For example, it is known that if the observations come from a distant object, the focal length and the distance of the object are difficult to estimate simultaneously.

This “restricted” estimation problem can be seen either as a ML problem with a smaller vector of parameters, or as a MAP problem with a zero-covariance prior probability on a subset of the parameters.

2.5. Choice of parameterization

If the estimated quantities have very different orders of magnitude, their estimators may become numerically unstable, and the theoretical covariances irrelevant. The parameterization is chosen to avoid these pitfalls, by having $E(\|\mathbf{K}\|^2) \approx 1$, since this is the module of the x_{pi} parameters. Note that this requirement on $E(\|\mathbf{K}\|^2)$ cannot be exactly enforced, since these parameters, unlike the x_{pi} cannot be arbitrarily scaled and one does not know their distribution precisely.

Neither the rotation parameters \mathcal{W} nor the translation parameters \mathcal{T} are normalized in the present work, but

their order of magnitude is reasonable. Alternatively, we could consider a parameterization of \mathcal{T} that takes into account the fact that t_{n3} is strictly positive. For example, one could estimate an affine function of $\log(t_{n3})$, chosen so that the estimated quantity has zero expectancy and unit variance. This assumes that one knows a priori the expectation and variance of the scene-to-camera distance. In practice, a rough guess would be used. The mapping $\mathbf{w}_n \rightarrow \mathbf{A}_n$ is “centered” on some rotation matrix: \mathbf{A}_n^* if it is known, or $\hat{\mathbf{A}}_n$. One takes, e.g. $\mathbf{A}_n = \mathbf{A}_n^* R(\mathbf{w}_n)$, where $R(\mathbf{w}_n)$ is the matrix of the rotation by $\|\mathbf{w}_n\|$ radians around the 3-vector \mathbf{w}_n . Using this parameterization, the standard deviation of the estimators is easily related to an angle.

For \mathbf{K} , based on our experience, and on remarks by Lenz and Tsai [18], we assumed that the parameters b_{21}/b_{11} (skew), $b_{22}/b_{11} - 1$ (aspect ratio), c_1 and c_2 (principal point) all have an expected absolute value of approximately 0.1. This leads us to the parameterization:

$$\mathbf{K} = [10b_{21}/b_{11}, 10b_{22}/b_{11} - 1, 10c_1, 10c_2, \log b_{11}]. \quad (7)$$

The focal length b_{11} could be better encoded using a different affine function of $\log b_{11}$.

3. Covariance of estimators

We derive the covariance matrices of estimators for three possible cases: the “plain” maximum likelihood (ML) estimator defined from the observations only, the maximum a posteriori (MAP) estimator obtained when a probabilistic prior is available for a subset of the estimated parameters and finally for the “restricted” ML estimator, in which a subset of the estimated parameters is fixed to given values. The obtained expressions, some of which being identical to those in Refs. [19,20], only involve Q and its derivatives; they can be applied to any problem of estimation just by specializing to the particular probability density function at hand. The lengthy parts of the derivation are placed in Appendix A.

We shall denote the “true” parameters Θ^* , our estimate $\hat{\Theta}$ (whether ML, MAP or “restricted”), and the error, $\Delta\Theta = \hat{\Theta} - \Theta^*$. Also, we will write $Q^* = Q(\mathcal{U}^*, \Theta^*) = 0$, $\hat{Q} = Q(\mathcal{U}, \hat{\Theta})$ and likewise $\hat{S} = S(\hat{\Theta})$. In general, an asterisk will denote a function evaluated in Θ^* or $(\mathcal{U}^*, \Theta^*)$, and a hat will denote evaluation at $\hat{\Theta}$ or $(\mathcal{U}, \hat{\Theta})$. We will write D_{Θ} and $D_{\Theta\Theta}^2$ the operations of first and second differentiation with respect to Θ . Likewise, $D_{\Theta \mathcal{U}}^2$ denotes differentiation with respect to Θ and \mathcal{U} . One thus has:

$$D_{\Theta} Q^* = \frac{\partial Q}{\partial \Theta}(\mathcal{U}^*, \Theta^*)$$

$$D_{\Theta\Theta}^2 Q^* = \frac{\partial^2 Q}{\partial \Theta^2}(\mathcal{U}^*, \Theta^*)$$

$$D_{\Theta \mathcal{U}}^2 Q^* = \frac{\partial^2 Q}{\partial \Theta \partial \mathcal{U}}(\mathcal{U}^*, \Theta^*) \text{ etc...}$$

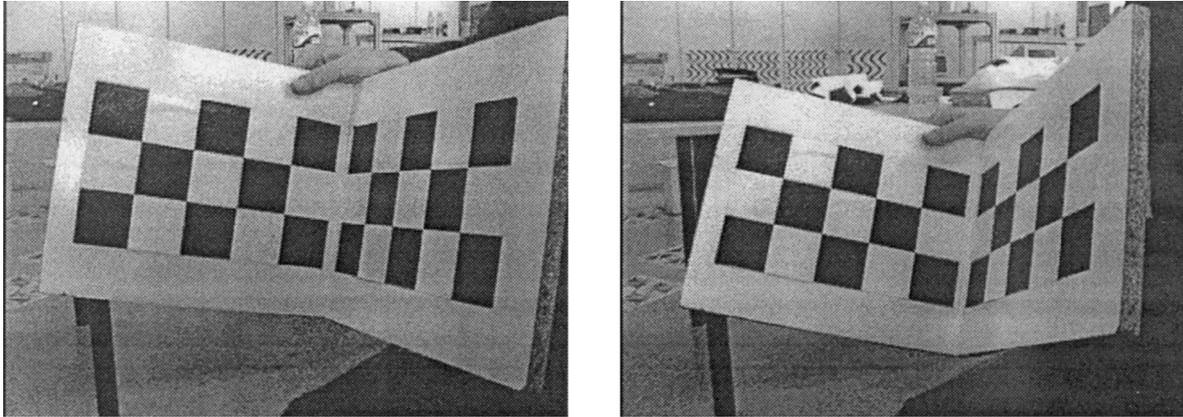


Fig. 2. Close range: First and fifth images of the sequence.

3.1. Derivation of the covariance of the ML estimator

A well-known property is that, at the minimum $\hat{\Theta}$, the derivative of Q (defined in Eq. (4)) is a linear combination of the derivatives of constraints; that is, there exist a (row) vector Λ of Lagrange multipliers such that:

$$D_{\Theta}\hat{Q} + \Lambda D_{\Theta}\hat{S} = 0_{1 \times \text{size}(\Theta)} \quad (8)$$

These are the so-called *normal equations*. The first-order Taylor series of $D_{\Theta}Q$ at $(\hat{\mathcal{U}}, \hat{\Theta})$, yields the following approximation:

$$D_{\Theta}Q^* \approx D_{\Theta}\hat{Q} - D_{\Theta\Theta}^2\hat{Q}\Delta\Theta - D_{\Theta\mathcal{U}}^2\hat{Q}\epsilon \quad (9)$$

It is easy to see that $D_{\Theta}Q^* = 0$. Since $Q^* = 0$, and, for all (\mathcal{U}, Θ) , $Q(\mathcal{U}, \Theta) \geq 0$, then $(\mathcal{U}^*, \Theta^*)$ is a global minimum of Q (regardless of constraints), which implies that $D_{\Theta}Q^* = 0$. Thus one has:

$$D_{\Theta\Theta}^2\hat{Q}\Delta\Theta + D_{\Theta\mathcal{U}}^2\hat{Q}\epsilon + \Lambda D_{\Theta}\hat{S} \approx 0$$

Likewise, since $S^* = 0 = \hat{S}$, and $S^* \approx \hat{S} - D_{\Theta}\hat{S}\Delta\Theta$, we obtain: $D_{\Theta}\hat{S}\Delta\Theta \approx 0$. Using this approximation and writing in matrix form, one has:

$$\begin{bmatrix} D_{\Theta\Theta}^2\hat{Q} & D_{\Theta}\hat{S}^T \\ D_{\Theta}\hat{S} & 0 \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Lambda^T \end{bmatrix} = \begin{bmatrix} -D_{\Theta\mathcal{U}}^2\hat{Q}\epsilon \\ 0 \end{bmatrix}.$$

Taking $H \equiv D_{\Theta\Theta}^2\hat{Q}$, $F \equiv D_{\Theta\mathcal{U}}^2\hat{Q}$ and $G \equiv D_{\Theta}\hat{S}$, one may write in a shorter way:

$$\begin{bmatrix} H & G^T \\ G & 0 \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Lambda^T \end{bmatrix} = \begin{bmatrix} -F\epsilon \\ 0 \end{bmatrix}. \quad (10)$$

The vector $[\Delta\Theta^T, \Lambda]$ is thus a linear combination of ϵ . Its covariance is:

$$\text{cov} \begin{bmatrix} \Delta\Theta \\ \Lambda^T \end{bmatrix} = \begin{bmatrix} H & G^T \\ G & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 FF^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H & G^T \\ G & 0 \end{bmatrix}^{-T} \quad (11)$$

It should be noted that in Eq. (10), one could eliminate Λ and write a similar expression in which $\Delta\Theta$ is a linear

combination of ϵ :

$$\begin{aligned} & [(I - G^T(GG^T)^{-1}G)H + GG^T]\Delta\Theta \\ & = -(I - G^T(GG^T)^{-1}G)F\epsilon. \end{aligned} \quad (12)$$

The square matrix in the left-hand side may be invertible even if H is not (as in our case). We could use this expression rather than Eq. (10) to derive an expression for $\text{cov}(\Delta\Theta)$, but the resulting expressions are not more intuitive.

3.2. Covariance of the MAP estimator

When there is some prior knowledge on the estimated parameters, the likelihood function is modified by the addition of terms (see Eq. (6)). The covariance matrix of $[\Delta\Theta^T, \Lambda]$ (the detailed derivation is in Section A.1) takes the form:

$$\begin{aligned} \text{cov} \begin{bmatrix} \Delta\Theta \\ \Lambda^T \end{bmatrix} &= \begin{bmatrix} H_{\text{post}} & G^T \\ G & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 FF^T + \Sigma^{-1}\Sigma^*\Sigma^{-T} & 0 \\ 0 & 0 \end{bmatrix} \\ &\times \begin{bmatrix} H_{\text{post}} & G^T \\ G & 0 \end{bmatrix}^{-T} \end{aligned} \quad (13)$$

where Σ^* is the true covariance of $\Theta^* - \bar{\Theta}$ and H_{post} is the Hessian of Q_{post} in $(\mathcal{U}, \hat{\Theta})$.

3.3. Covariance of the ML estimator with fixed parameters

Another common situation arises when a subset of the parameters is known, or we assume that its true value can be replaced by “nominal values”, e.g. to achieve numerical stability. We split $\Theta = [\Theta_1, \Theta_2]$ where Θ_1 is known and are interested in $\hat{\Theta}_2$, such that, for a given $\bar{\Theta}_1$, $[\bar{\Theta}_1, \hat{\Theta}_2]$ is the minimum of the function that associates $Q(\mathcal{U}, [\Theta_1, \Theta_2])$, to $(\mathcal{U}, \Theta_1, \Theta_2)$. The various differentials of this function are written

$$G_i = \frac{\partial Q}{\partial \Theta_i}(\mathcal{U}, [\Theta_1, \Theta_2]) \quad i \in \{1, 2\}$$

Table 1

Expected standard deviation of the error on the estimated parameters. The error on \mathbf{x} and \mathbf{t} is “metric”; the mean squared value of \mathbf{x} being 1, an error of 0.1 denotes 10% of error. The error on \mathbf{w} is the standard deviations of the angle between the true and the estimated camera frames. The error on \mathbf{K} is on the calibration parameters, as defined in Eq. (7)

Close range, $N = 5, P = 48$				
	\mathbf{x}	\mathbf{w}	\mathbf{t}	\mathbf{K}
Calib	0.0005	0.1947	0.0778	0.0518
ML	0.734	2.62	4.47	3.88
TP	0.156	0.57	1.074	0.734
CP	0.068	0.346	0.155	0.0517
TF	0.551	2.314	6.195	1.0
SF	0.256	0.860	1.695	1.188
CF	0.0685	0.346	0.155	0.0518

$$H_{ij} = \frac{\partial^2 Q}{\partial \Theta_i \partial \Theta_j}(\mathcal{U}, [\Theta_1, \Theta_2]) \quad i, j \in \{1, 2\}$$

$$F_i = \frac{\partial^2 Q}{\partial \Theta_i \partial \mathcal{U}}(\mathcal{U}, [\Theta_1, \Theta_2]) \quad i \in \{1, 2\}.$$

We show in Section A.2 that the covariance of $[\Delta \Theta_2^T, \Lambda]$ is:

$$\begin{aligned} \text{cov} \begin{bmatrix} \Delta \Theta_2 \\ \Lambda^T \end{bmatrix} &= \begin{bmatrix} H_{22} & G_2^T \\ G_2 & 0 \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} \sigma^2 F_2 F_2^T + H_{21} \Sigma_1^* H_{21}^T & H_{21} \Sigma_1^* G_1^T \\ G_1 \Sigma_1^* H_{21}^T & G_1 \Sigma_1^* G_1^T \end{bmatrix} \begin{bmatrix} H_{22} & G_2^T \\ G_2 & 0 \end{bmatrix}^{-T}, \end{aligned} \quad (14)$$

where $\Sigma_1^* = \text{cov}(\Theta_1 - \bar{\Theta}_1)$. This matrix is not usually known in practice, just like the matrix Σ^* in the previous section.

4. Specialization to the problem of reconstruction

The above formulas hold for any estimator of the considered types (ML, MAP or “restricted” ML). We now specialize them to the case of Gaussian noise, when the log-likelihood is a sum of squared differences between observations and predictions

$$D_{\Theta_i} Q = \sum_{npk} D_{\Theta_i v_{npk}} (v_{npk} - u_{npk}) / \sigma^2 \quad (15)$$

$$D_{\Theta_i \Theta_j}^2 Q = \frac{1}{\sigma^2} \sum_{npk} D_{\Theta_i v_{npk}} D_{\Theta_j v_{npk}} + D_{\Theta_i \Theta_j}^2 v_{npk} (v_{npk} - u_{npk})$$

$$D_{\Theta_i u_{npk}}^2 Q = D_{\Theta_i v_{npk}} / \sigma^2$$

A first practical consideration: notice that in the previous section, one may perform the expansions in Taylor series around Θ^* rather than in $\hat{\Theta}$. One would then obtain expressions like Eqs. (11)–(14), but with H replaced by $D_{\Theta \Theta}^2 Q^*$, and likewise for G and F . Thus, if Θ^* is known, e.g. as in

simulations, one may compute (an approximation of) the covariance matrix of the ML estimator, without even needing to know how to implement it.

Second practical consideration: one can eliminate the need of knowing the second derivatives $D_{\Theta_i \Theta_j}^2 v_{npk}$ when computing covariance matrices, because, at $(\Theta^*, \mathcal{U}^*)$, one has $v_{npk} = u_{npk}$, and thus the second order terms in $D_{\Theta_i \Theta_j}^2 Q$ are eliminated in Eq. (15). Even when Θ^* is unknown, we estimate the covariance of an estimator at $\hat{\Theta}$, without using the second derivatives. Justification for such practice may be found in of [21]. In what follows, the matrices H will not include the second derivative terms.

Covariance of the ML estimator: as we consider that noise terms ϵ_{npj} are independent and have same variance σ^2 , one has: $\sigma^2 F F^T = \sigma^2 D_{\Theta}^2 \mathcal{U} Q^T \cdot D_{\Theta}^2 \mathcal{U} Q = H$. Replacing in Eq. (11) yields.

$$\text{Cov} \begin{bmatrix} \Delta \Theta \\ \Lambda \end{bmatrix} = \begin{bmatrix} H & G^T \\ G & 0 \end{bmatrix}^{-1} \begin{bmatrix} H & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H & G^T \\ G & 0 \end{bmatrix}^{-T} \quad (16)$$

The same simplification can be carried out in Eqs. (13) and (14).

A prior on the structure, $\mathcal{X}^* \simeq N(\mathcal{X}_0, \Sigma_2)$ makes the constraint defined in Eq. (5) irrelevant: all the parameters can be uniquely determined without having to restrict the parameter set. The normal equations are:

$$H^+ \Delta \Theta = F \epsilon + \begin{bmatrix} \Sigma_2^{-1} (\mathcal{X}_0 - \mathcal{X}^*) \\ 0_{6N+5} \end{bmatrix},$$

$$\text{where } H^+ = H + \begin{bmatrix} \Sigma_2^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

is the modified matrix of second derivatives (assuming that the \mathcal{X} parameters are stored at the beginning of Θ). The covariance of the estimate, assuming that $\Sigma_2^* = \Sigma_2$ is:

$$\text{cov} \Delta \Theta = (H^+)^{-1} H^+ (H^+)^{-T} = (H^+)^{-T} \quad (17)$$

A prior on the calibration parameters, $\mathbf{K}^* \simeq N(\mathbf{K}_0, \Sigma_2)$, is treated likewise, but keeping the constraints \mathbf{S} , and the matrix \mathbf{G} of its derivatives. If \mathbf{K} is stored at the end of Θ , and defining:

$$H^+ = H + \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_2^{-1} \end{bmatrix},$$

the covariance matrix is:

$$\text{Cov} \begin{bmatrix} \Delta \Theta_2 \\ \Lambda \end{bmatrix} = \begin{bmatrix} H^+ & G^T \\ G & 0 \end{bmatrix}^{-1} \begin{bmatrix} H^+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H^+ & G^T \\ G & 0 \end{bmatrix}^{-T} \quad (18)$$

Fixed parameters: fixing $\Theta_1 = \mathbf{K}$ to some value $\mathbf{K}_0 \neq \mathbf{K}^*$, and assuming that $\mathbf{K}^* \sim N(\mathbf{K}_0, \Sigma_1)$, the covariance

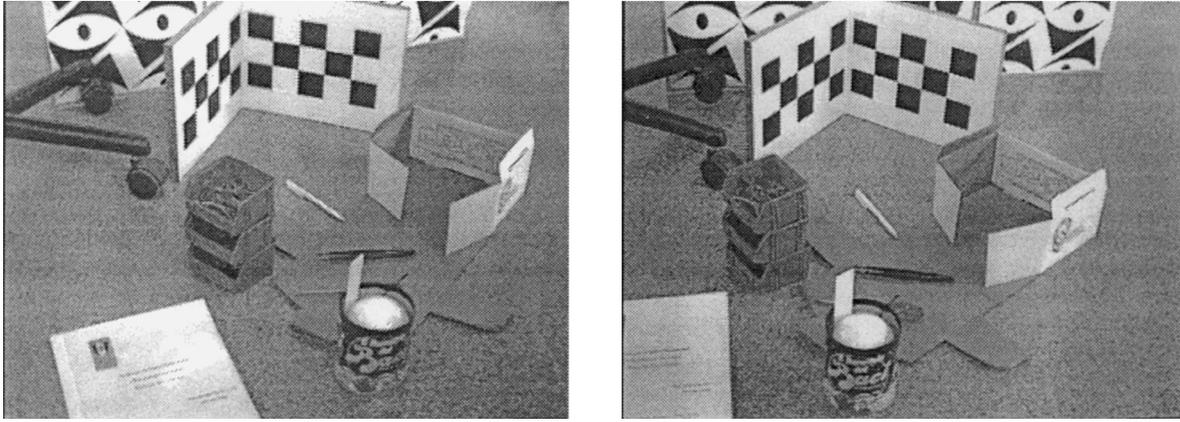


Fig. 3. Long range: First and tenth images of the sequence.

matrix of the estimator of $\hat{\Theta}_2 = (\mathcal{X}, \mathcal{W}, \mathcal{T})$, takes the form:

$$\text{Cov} \begin{bmatrix} \Delta \Theta_2 \\ \Lambda \end{bmatrix} = \begin{bmatrix} H_{22} & G_2^T \\ G_2 & 0_{7 \times 7} \end{bmatrix}^{-1} \begin{bmatrix} H_{22} + H_{21} \Sigma_1 H_{21}^T & 0 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} H_{22} & G_2^T \\ G_2 & 0 \end{bmatrix}^{-T} \quad (19)$$

Here, like in Eq. (14), H_{22} is the matrix of Q derived twice with respect to $\Theta_2 = (\mathcal{X}, \mathcal{W}, \mathcal{T})$, at $(\mathcal{U}, \hat{\Theta}_2)$; H_{21} is the matrix of the second derivatives of Q with respect to $\Theta_1 = \mathbf{K}$ (columns) and Θ_2 (rows). And G_2 is the matrix of derivatives of S with respect to Θ_2 . The matrix G_1 that appears in Eq. (14) is zero, and the corresponding terms in Eq. (19) have been removed. Note the same form of covariance matrix is obtained when only a sub-vector of \mathbf{K} is fixed.

The variance of each estimated parameter appears on the diagonal terms of the matrices in Eqs. (16)–(19). As they are, these expressions do not tell the actual values that one will observe in practice. However, once one has obtained the output from one of the considered estimators, it is possible to determine the precision of this estimate.

5. Experimental results

We performed various experiments (real and simulated) to study the performance of the estimators. The errors on the parameters \mathcal{X} , \mathcal{W} , \mathcal{T} and \mathbf{K} are studied separately. For \mathcal{X} and \mathbf{K} , which are normalized for having $E(\|\mathbf{x}_p\|^2) = 1$, and $E(\|\mathbf{K}\|^2) \simeq 1$, the error measures are the standard deviations of $\|\mathbf{x}_p - \mathbf{x}_p^*\|$ and $\|\mathbf{K} - \mathbf{K}^*\|$. For \mathcal{T} , the standard deviation of $\|\mathbf{t} - \mathbf{t}^*\|$ is used too. We saw in Section 2.5 that \mathcal{T} is expressed in the same unit as \mathcal{X} . For \mathcal{W} , the measure is the standard deviation of the angle formed between axes of the true and the estimated camera frames, $\|\mathbf{w}_n - \mathbf{w}_n^*\|$.

We begin by validating experimentally the expressions for the covariance matrices in Eqs. (11), (16) and (17) to the case of 3D reconstructions. We then show the precision

obtained by estimators in two real-world situations, one a close-range sequence of five images, and the other a long-range sequence of 12 images. Finally, we use simulation to evaluate the effect on precision of the number of images in a sequence and the relative disposition of the cameras.

5.1. Validation of the analytical expressions of covariances

The covariance matrices in Eqs. (16)–(19) are obtained using the approximations (9) and (15). We must verify that they are valid in practice. This is carried out by implementing the considered estimator, and verifying that the error committed is consistent with the predictions. We have built 100 “general position” setups of 10 points seen in five images. For each setup, the corresponding theoretical covariance matrices Σ_θ , are computed. The observations $\mathcal{U}(\Theta)$ are contaminated by i.i.d. Gaussian noise, at 40 dB,² and a ML estimate $\hat{\Theta}$ is determined. The error committed on each individual parameter of Θ is scaled by the corresponding theoretical standard deviation. These values should follow a law $N(0,1)$ if the theoretical variances were correct. The histogram of the resulting values is shown in Fig. 1 together with a reference Gaussian density curve. For that noise level, the theoretical and true covariances are very similar, and we conclude that the theoretical variances are realistic.

5.2. Variance of estimators

In the next sections, we will compare the precision of various estimators. In the following tables, each row displays results concerning a given estimator. To allow easy identification, we use the following labels:

1. Full reconstruction based on observations only.
ML: maximum-Likelihood estimator of the whole vector Θ , with covariance defined in Eq. (16).
2. Full reconstruction with a prior on structure (MAP estimation).

² Noise level, in decibels is defined as $\text{dB} = -10 \log_{10}(\text{var}(\epsilon)/\text{var}(u))$.

Table 2

Expected standard deviation of the error on either structure \mathbf{x} , orientation \mathbf{w} , position \mathbf{t} or calibration parameters \mathbf{K} . \mathbf{w} is expressed in degrees

Long range, $N = 12$, $P = 48$				
	\mathbf{x}	\mathbf{w}	\mathbf{t}	\mathbf{K}
Calib	0.00029	0.21517	0.19428	0.1248
ML	0.0432	0.1866	0.6983	0.4944
TP	0.0424	0.1754	0.5285	0.3462
CP	0.0364	0.1525	0.1919	0.0461
TF	2.1625	10.7319	6.3661	1.0
SF	0.0404	0.1608	0.2995	0.1005
CF	0.0407	0.1763	0.2068	0.05175

Calib contains results obtained when calibrating from known object, i.e. when one uses a prior on the structure. The covariance Σ_{Calib} is determined by Eq. (17). The diagonal terms of Σ_{Calib} from the diagonal of matrix Σ_1 in Eqs. (18) and (19), all non-diagonal terms being 0.

3. Full reconstruction with a probabilistic knowledge on the intrinsic parameters \mathbf{K} . This corresponds to MAP estimation, and the covariance is given by Eq. (18). We further consider the following two cases:

TP: since we parametrized \mathbf{K} in such a way that the individual parameters have (approximately) zero mean, and unit variance, one can say that $\mathbf{K}_0 = [0 \ 0 \ 0 \ 0 \ 0]$ constitutes a (nominal) estimator of \mathbf{K}^* , with covariance $\Sigma_1 = \mathbf{I}_5$.

CP: we assume that a prior estimate on \mathbf{K}^* is available, whose covariance Σ_1 is equal to Σ_{Calib} on the diagonal terms, and zero otherwise.

4. In case of some intrinsic parameters being fixed, the covariance of the estimator is given by Eq. (19), and we distinguish the following three cases:

TF: $\Theta_1 = \mathbf{K}$ is fixed, e.g. to $[0 \ 0 \ 0 \ 0 \ 0]$. This constitutes a (nominal) estimate of \mathbf{K}^* , with covariance $\Sigma_1 = \mathbf{I}_5$. In the notation of Section 3, one has $\Theta_1 = \mathbf{K} = 0_5$.

SF: in real-world situations, the skew and aspect-ratio (k_1 and k_2) do not change as much as the other intrinsic parameters and they may be given by the camera manufacturer, or by a previous calibration step. In the SF estimator, $[k_1, k_2]$ are fixed to values $[\bar{k}_1, \bar{k}_2]$ given a priori. The precision of this estimate is assumed to be comparable to that of calibration from known-object (the Calib estimator). In the notation of Section 3, the subvector Θ_1 is $[\bar{k}_1, \bar{k}_2]$, and Σ_1 is formed by the diagonal elements of upper-left 2×2 subblock of Σ_{Calib} .

CF: we assume $\Theta_1 = \mathbf{K}$ is known with the same precision as in Calib. For Σ_1 , we take the diagonal part of Σ_{Calib} .

5.2.1. Short-range

We compare the relative precisions using a real-world

sequence of 5 images of a static scene, with fixed intrinsic parameters and 48 hand-matched points on a calibration grid. This setup is shown in Fig. 2. Total rotation is ≈ 30 degrees. Hand-measurements provided us (a priori on) the 3D point positions, with an accuracy that we estimate at 1% (≈ 2 mm). We can thus perform calibration from a known object, and compute $\hat{\mathcal{X}}, \hat{\mathcal{W}}, \hat{\mathcal{T}}$ and $\hat{\mathbf{K}}$. Table 1 shows the standard deviations obtained for each type of estimator, using Eqs. (16)–(19). The value of σ that we used is the standard deviation of the residuals $\mathbf{u}_{npi} - \mathbf{v}_{npi}(\hat{\Theta})$, which corresponds to 48 dB.

The most important features apparent from this table are:

- The ML and TF estimators have very high covariance, for which the validity of Eqs. (16) and (19) should be verified.
- The precision obtained with full calibration information (CP and CF) is much better than that obtained without (lines ML, TP, TF), or with only skew and aspect ratio information (line SF).
- Without pre-calibration, the use of a nominal prior (TP, third line) provides much better estimates than either the ML or the nominally fixed parameter (TF) estimators (second and fifth lines).

5.2.2. Long-range

The grid in Fig. 3 is seen along 12 images, from 1.5–2.5 m, and the maximum camera–camera distance is ≈ 1.3 m. The standard deviations of the five tested estimators are displayed in Table 2. For the CP and CF (fourth and sixth lines) estimators, the covariance of the prior is that of short-range calibration. The ML and TP estimator (second and third lines) perform nearly as well as the estimators that use prior calibration (fourth and sixth lines). The estimator with fixed nominal parameters, however, appears to behave relatively poorly. The first point appears to be due to the increased number of images used and to the greater baseline between the cameras, as will be shown in the following experiments.

5.2.3. Influence of the number of images

Fig. 4 plots the base 10 logarithm of the variance in structure and in intrinsic parameters, as function of the number of images used. In this experiments, thirty synthetic setups were generated, each with twelve 3D points, generated as Gaussian white noise, and then normalized. The camera orientations have Euler angles independently uniformly distributed in $[\pm\pi/4] \times [\pm\pi/4] \times [\pm\pi/8]$. The scene–camera distance is 6–12 times the size of the scene. The precision of the camera calibration is that of line Calib in Table 1, and the observations noise has variance 10^{-4} .

Fig. 4 shows that long sequences of uncalibrated images allow as good 3D reconstruction as short calibrated sequences, whereas short uncalibrated sequences give

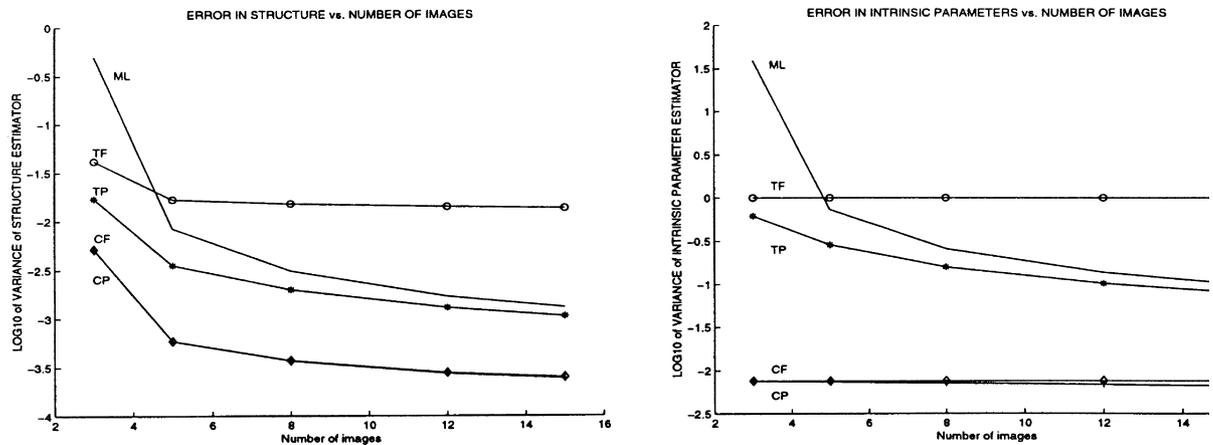


Fig. 4. The log (base 10) of the error on the structure parameters \mathcal{X} and calibration parameters, \mathbf{K} . The curves are tagged ML, TP, CP, TF and CF as explained in the text.

poor results. In all cases, it is better to use a nominal prior than to fix the intrinsic parameters to nominal values.

Sections 5.2.4 and 5.2.5 present studies of the effect of the positioning of three cameras on the precision of various estimators.

5.2.4. Importance of orientation change

In Ref. [12] it is shown that if the rotations between the camera frames all have a common axis, then the problem of self-calibration is ill-posed. To investigate this, we fix two cameras: the orientation of the first is the 3×3 identity matrix, that of the second is a rotation around the \mathbf{y} axis by $\pi/6$ radians (bottom cameras in Fig. 5(b)). The third camera is rotated by a variable angle around the \mathbf{x} axis. If the angle is zero, the first and third cameras are equal and it is impossible to estimate all the parameters. Fig. 5(a) shows how the error in estimated structure parameters varies with the rotation of the third camera. We generated 40 random setups, each consisting of 20 points. For each point setup, and for each tested angle value, we use Eqs. (16)–(19) to compute the covariance of the considered estimators, at the true value Θ^* . Fig. 5 plots the base-10 logarithm of the standard deviation of an estimated structure parameter $\hat{\mathbf{x}}_{pi}$. The main observations are the following:

- Three of the curves go down sharply, in the right neighborhood of zero: the ML (with asterisks), SF (circles) and TP (plain curve). This last curve, if continued towards zero, would not become unbounded, whereas ML would. The error of these estimators is quite high ($>3\%$) when the angle is smaller than $0.4 \approx \pi/8$.
- The TF (top curve) displays very high error ($\approx 30\%$), which may be outside the domain of validity of Eq. (19). In practice, the implementation of that estimator will not produce estimates with independent error terms of high variance, but rather, will produce a much-distorted estimate of the structure. The curve is bounded near zero.
- The two bottom curves, CF (with circles), and CP (plain)

are bounded near zero too, as the estimation problem is well posed.

5.2.5. Angle between viewpoints (baseline)

To analyze the influence of varying the viewing angle, we again use three cameras. This time, the second and third cameras are rotated by an equal amount, around the \mathbf{x} and \mathbf{y} axes, respectively. When the angle between viewpoints is small, the three cameras are nearly equal, and the estimation problem is singular. When the angle is nearly π radians, the second and third cameras, are nearly equal, and are facing the first. This situation is also singular. Fig. 6 shows the effect of the rotation angle on the error. The error is computed in the same manner as in the previous experiment.

The general tendencies of the error on structure (Fig. 6a) and calibration (Fig. 6b) are the same:

- The TF curve (with circles) which is almost systematically above all others, both for structure and calibration (Fig. 6a and b).
- The three curves of the ML, TP (plain) and SF (circles) estimators are joined together on the left side, and split after $2\pi/3$. The estimator with fixed skew and aspect-ratio (SF) performs clearly better for bigger angles. Using a nominal prior (TP) improves the accuracy for angle near 0 and π . In all three cases, the precision is best for angles in $[\pi/3, 5\pi/6]$. Next to $2\pi/3$, the precision is nearly as good as in the calibrated case.
- The curves of the CF (with circles) and CP (plain) are always below the others.

6. Conclusions

We have analyzed the problem of the precision that is achievable in 3D reconstruction from uncalibrated views. Although a lot of work has been carried out on various

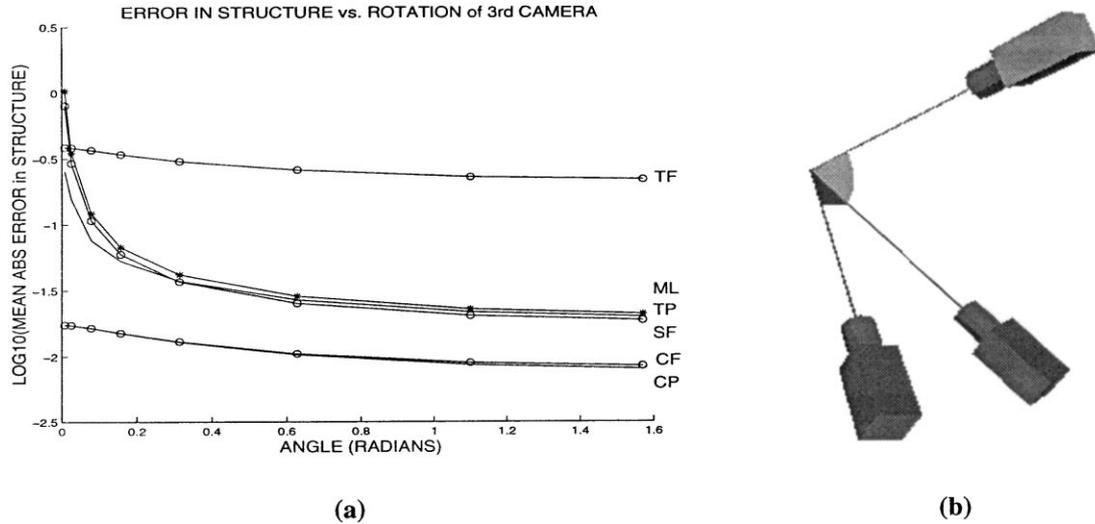


Fig. 5. (a) Error vs. Angle. The first and second cameras are related by a $\pi/6$ radians rotation around the y axis. The error in the structure parameters \mathcal{X} is plotted vs. the angle relating the first and third (top) camera. (b) Setup. The angle between the two first camera (at bottom) is fixed to $\pi/6$, whereas the angle of the third (top) camera is made to vary.

forms of reconstruction, the problem of precision evaluation is seldom addressed in a systematic way.

We have formulated the problem in a probabilistic framework. We further considered that various types of prior information may be available and defined the corresponding estimators.

One contribution of this work is the analytical derivation of the covariance matrices for the set of estimators. These derivations are immediately applicable to other estimation problems in which the noise of the observations is i.i.d, and may easily be generalized to non-i.i.d. noise. This analysis, applied to the case of 3D reconstruction, provides insight relative to what precision can be expected in each circumstance, and *does not* depend on a particular implementation. We validated experimentally the obtained expressions.

Finally, we compared experimentally the precisions of various estimators. For each, using both synthetic and real image data, we analyzed the influence on precision of the number of images, the camera disposition etc. The main conclusions of our experimental work can be summarized as:

- Pre-calibration, if one may assume that the intrinsic parameters do not vary, greatly improves the precision of reconstruction. When realistic calibration parameters are available, they can be fixed: compare the “CP” and “CF” lines in the tables and graphs above.
- Long sequences, or good positioning of cameras, greatly improve the quality of uncalibrated reconstruction. Therefore it shows the potential quality of euclidean reconstruction obtained from long uncalibrated sequences.

We presently are further analyzing the influence of sequence length, number of points and noise in image measurements. On the analytical side, we plan to extend the present work to variable intrinsic parameters.

Acknowledgements

This work has been supported by projects INCO COPERNICUS Proj. 960174-VIRTUOUS and PRAXIS 2/2.1/TPAR/2074/95.

Appendix A. Derivation of the covariance matrices

A.1. Covariance of an estimator, when a prior is used

The derivation is very similar to that of the ML estimator, in Section 3. In what follows, we assume that our prior may be inaccurate. The *true* matrix of the covariance of $\Theta^* - \bar{\Theta}$, Σ^* , may be different from the *assumed* covariance matrix, Σ . In practice, Σ^* is most often not known, and it is replaced by Σ in numerical computations. However, for theoretical considerations, we assume that $\Sigma \neq \Sigma^*$. The cost function (proportional to the inverted logarithm of the posterior probability density of Θ when \mathcal{U} is given) is now:

$$Q_{\text{post}}(\mathcal{U}, \Theta) = Q(\mathcal{U}, \Theta) + (\Theta^* - \bar{\Theta})^T \Sigma^{-1} (\Theta^* - \bar{\Theta})$$

At the minimum $\hat{\Theta}$:

$$D_{\Theta} \hat{Q}_{\text{post}} + \Delta D_{\Theta} \hat{S} = 0$$

Since (expanding as first-order Taylor series)

$$D_{\Theta} Q_{\text{post}} = (\Theta^* - \bar{\Theta})^T \Sigma^{-1} \approx D_{\Theta} \hat{Q}_{\text{post}} - D_{\Theta\Theta}^2 \hat{Q}_{\text{post}} \Delta \Theta - D_{\Theta\mathcal{U}}^2 \hat{Q}_{\text{post}} \epsilon,$$

one has

$$D_{\Theta\Theta}^2 \hat{Q}_{\text{post}} \Delta \Theta + D_{\Theta\mathcal{U}}^2 \hat{Q}_{\text{post}} \epsilon + \Delta D_{\Theta} \hat{S} \approx (-\Theta^* + \bar{\Theta})^T \Sigma^{-1}.$$

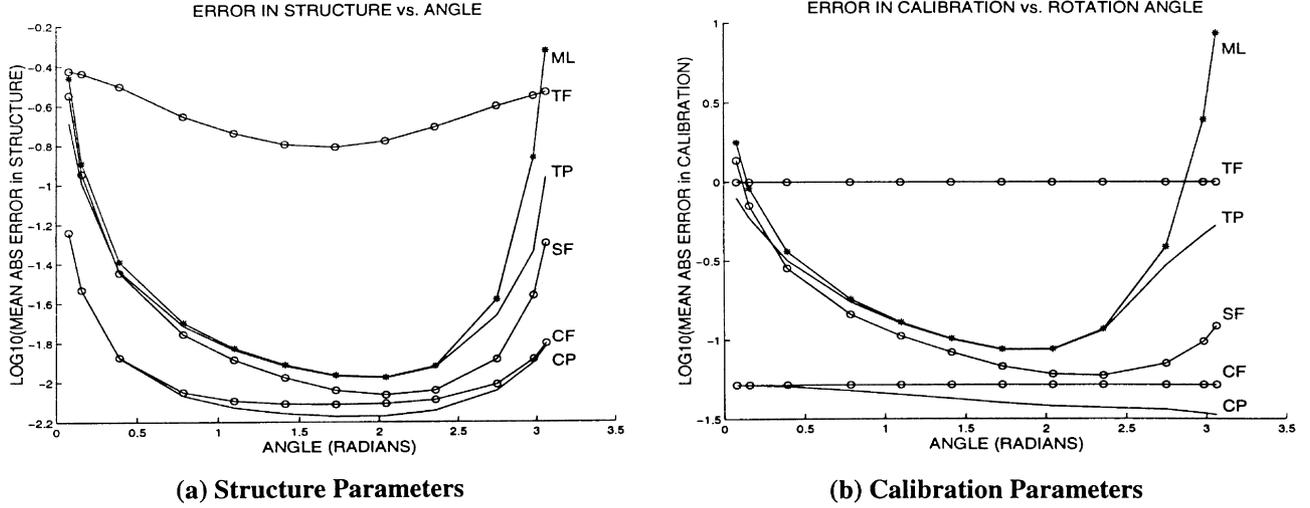


Fig. 6. Error vs. Angle. The second and third cameras are rotated with respect to the first camera around the y and x axes, respectively. The error in the structure parameters \mathcal{X} , and in the calibration parameters \mathbf{K} are plotted vs. the angle, in (a) and (b), respectively.

Likewise, having $S^* = 0 = \hat{S}$ and $S^* \approx \hat{S} - D_{\theta} \hat{S} \Delta \theta$, implies that $D_{\theta} \hat{S} \Delta \theta \approx 0$. Altogether (neglecting the higher order terms), it yields:

$$\begin{bmatrix} D_{\theta\theta}^2 \hat{Q}_{\text{post}} & D_{\theta} \hat{S}^T \\ D_{\theta} \hat{S} & 0 \end{bmatrix} \begin{bmatrix} \Delta \theta \\ \Lambda^T \end{bmatrix} = \begin{bmatrix} -D_{\theta\theta}^2 \hat{Q} \epsilon + \Sigma^{-1} (\theta^* - \bar{\theta}) \\ 0 \end{bmatrix},$$

and the covariance of $[\Delta \theta^T \Lambda]$ is:

$$\text{cov} \begin{bmatrix} \Delta \theta \\ \Lambda^T \end{bmatrix} = \begin{bmatrix} H_{\text{post}} & G^T \\ G & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 FF + \Sigma^{-1} \Sigma^* \Sigma^{-T} & 0 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} H_{\text{post}} & G^T \\ G & 0 \end{bmatrix}^{-T}, \quad (\text{A1})$$

where Σ^* is the true covariance of $\theta^* - \bar{\theta}$, which may be different from Σ , and $H_{\text{post}} = D_{\theta\theta}^2 \hat{Q} + \Sigma^{-1}$ is the Hessian of Q_{post} in $(\mathcal{U}, \hat{\theta})$.

A common situation is when a prior is available for some parameters only, for example when one is performing calibration from a known object, or when one has a prior on the intrinsic parameters only. If we assume the parameter vector is split like $\theta = [\theta_1, \theta_2]$, with a prior $\theta_2 \sim N(\bar{\theta}_2, \Sigma_2)$, one obtains a result similar to Eq. (13), where Σ is replaced by $\text{diag}(0, \Sigma_2)$ and Σ^{-1} by $\text{diag}(0, \Sigma_2^{-1})$, and likewise, Σ^* .

A.2. Covariance of the ML estimator, when some of the parameters are fixed

We split $\theta = [\theta_1, \theta_2]$, and are interested in $\hat{\theta}_2$, such that, for a given $\bar{\theta}_1, [\bar{\theta}_1, \hat{\theta}_2]$ is the minimum of $Q(\mathcal{U}, \cdot)$

within the set $S^{-1}(\{0\})$. We note

$$D_i Q = \frac{\partial Q}{\partial \theta_i} (\mathcal{U}, [\theta_1, \theta_2]) \quad i \in \{1, 2\}$$

$$D_{ij}^2 Q = \frac{\partial^2 Q}{\partial \theta_i \partial \theta_j} (\mathcal{U}, [\theta_1, \theta_2]) \quad i, j \in \{1, 2\}$$

At the minimum $\hat{\theta}_2$:

$$\underbrace{D_2 Q([\bar{\theta}_1, \hat{\theta}_2])}_{D_2 \hat{Q}} + \Lambda \underbrace{D_2 S([\bar{\theta}_1, \hat{\theta}_2])}_{D_2 \hat{S}} = 0$$

Writing $\Delta \theta_2 = \hat{\theta}_2 - \theta_2^*$, the first-order Taylor expansion is:

$$D_2 Q^* = 0 \approx D_2 \hat{Q} - \underbrace{D_{21}^2 \hat{Q}}_{H_{21}} (\bar{\theta}_1 - \theta_1^*) - \underbrace{D_{22}^2 \hat{Q}}_{H_{22}} \Delta \theta_2 - \underbrace{D_{2U}^2 \hat{Q}}_{F_2} \hat{Q} \epsilon,$$

one has

$$H_{22} \Delta \theta_2 + \Lambda D_2 \hat{S} \approx H_{21} (\bar{\theta}_1 - \theta_1^*) - F_2 \epsilon.$$

Now, since $S([\theta_1^*, \theta_2^*]) = 0 = S([\bar{\theta}_1, \hat{\theta}_2])$, and

$$S([\theta_1^*, \theta_2^*]) \approx S([\bar{\theta}_1, \hat{\theta}_2]) - D_1 \hat{S} \cdot (\bar{\theta}_1 - \theta_1^*) - D_2 \hat{S} \Delta \theta_2,$$

one has:

$$\underbrace{D_2 \hat{S}}_{G_2} \Delta \theta_2 \approx - \underbrace{D_1 \hat{S}}_{G_1} \cdot (\bar{\theta}_1 - \theta_1^*).$$

Neglecting the higher order terms, and writing in matrix form, we obtain:

$$\begin{bmatrix} H_{22} & G_2^T \\ G_2 & 0 \end{bmatrix} \begin{bmatrix} \Delta \theta_2 \\ \Lambda^T \end{bmatrix} = \begin{bmatrix} -F_2 \epsilon + H_{21} (\theta_1^* - \bar{\theta}_1) \\ G_1 (\bar{\theta}_1 - \theta_1^*) \end{bmatrix},$$

and the covariance of $[\Delta\Theta_2^T, \Lambda]$ is:

$$\begin{aligned} \text{Cov} \begin{bmatrix} \Delta\Theta_2 \\ \Lambda^T \end{bmatrix} &= \begin{bmatrix} H_{22} & G_2^T \\ G_2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 F_2 F_2^T + H_{21} \Sigma_1^* H_{21}^T & H_{21} \Sigma_1^* G_1^T \\ & G_1 \Sigma_1^* H_{21}^T & G_1 \Sigma_1^* G_1^T \end{bmatrix} \\ &\times \begin{bmatrix} H_{22} & G_2^T \\ G_2 & 0 \end{bmatrix}^{-T}, \end{aligned} \quad (\text{A2})$$

where $\Sigma_1^* = \text{cov}(\Theta_1^* - \bar{\Theta}_1)$. This matrix is not usually known in practice, just like the matrix Σ^* in the previous section.

References

- [1] P.R. Wolf, Elements of photogrammetry, with air photo interpretation and remote sensing, 2, McGraw-Hill, New York, 1983.
- [2] R.I. Hartley, Euclidean reconstruction from uncalibrated views, In in 2nd Proc. Europe-U.S. Workshop on Invariance (1993) 237–256.
- [3] O.D. Faugeras, What can be seen in three dimensions with an uncalibrated stereo rig? in: G. Sandini (Ed.), Proceedings of the ECCV, Springer, Berlin, 1992, pp. 563–578.
- [4] S.J. Maybank, O.D. Faugeras, Theory of self-calibration of a moving camera, International Journal Computer Vision 8 (2) (1992) 123–151.
- [5] M. Pollefeys, R. Koch, L. Van Gool, Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, Proceedings of the Sixth ICCV (1998) 90–95.
- [6] A.W. Fitzgibbon, G. Cross, A. Zisserman, Automatic 3D model construction for turn-table sequences, in: R. Koch, L. Van Gool (Eds.), 3D Structure from Multiple Images of Large-Scale Environments, Lecture Notes in Computer Science 1506 Springer, Berlin, 1998, pp. 155–170.
- [7] Gang Xu, Zhengyou Zhang, Epipolar geometry in Stereo, Motion and Object Recognition, A Unified Approach, Kluwer Academic, Dordrecht, 1996.
- [8] A. Shashua, Algebraic functions for recognition, IEEE Transactions on PAMI 17 (8) (1994) 779–789.
- [9] G. Csurka, C. Zeller, Z. Zhang, O. Faugeras, Characterizing the uncertainty of the fundamental matrix, Computer Vision and Image Understanding 68 (1) (1997) 18–35.
- [10] R.I. Hartley, Minimizing algebraic error in geometric estimation problems. Proceedings of the Sixth International Conference on Computer Vision (ICCV), Bombay, India, January 1998, pp. 469–476.
- [11] P.H.S. Torr, A. Zisserman, S. Maybank, Robust detection of degenerate configurations for the fundamental matrix, Computer Vision and Image Understanding 71 (3) (1998) 312–333.
- [12] P. Sturm, Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction, Proceedings of CVPR (1997) 1100–1105.
- [13] J. Weng, N. Ahuja, T.S. Huang, Optimal motion and structure estimation, IEEE Transactions on PAMI 15 (9) (1993) 864–884.
- [14] S. Bougnoux, From projective to euclidean space under any practical situation, a criticism of self-calibration, Proceedings of the Sixth ICCV (1998) 790–796.
- [15] R. Szeliski, Sing Bing Kang, Shape ambiguities in structure from motion, IEEE Transactions PAMI 19 (5) (1997) 506–512.
- [16] R. Mohr, F. Veillon, L. Quan, Relative 3d reconstruction using multiple uncalibrated images, Proceedings of the IEEE CVPR (1993) 543–548.
- [17] J.V. Beck, K.J. Arnold, Estimation in Engineering and Science, Wiley, New York, 1977.
- [18] R.K. Lenz, R.Y. Tsai, Techniques for calibration of the scale factor and image center for high-accuracy 3D machine vision metrology, IEEE Transactions on PAMI 10 (5) (1988) 713–720.
- [19] R.M. Haralick, Propagating covariance in computer vision, Proceedings of the Workshop on Performance Characteristics of Vision Algorithms (1996) 1–12.
- [20] R.M. Haralick, Propagating covariance in computer vision. Technical Report CITR-TR-29, Centre for Image Technology and Robotics, University of Auckland, <http://www.tcs.auckland.ac.nz/research/tech-reports/CITR-TR-29.pdf>, September 1998.
- [21] W.H. Press, S.A. Teutolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes, The Art of Scientific Computing, 2, Cambridge University Press, Cambridge, UK, 1992.