In *British Machine Vision Conference, Southampton 1998*

The precision of 3D Reconstruction from Uncalibrated Views

Etienne Grossmann*and José Santos Victor
Instituto de Sistemas e Robótica
Instituto Superior Tecnico
Av. Rovisco Pais, 1
P1096 Lisboa Codex Portugal
{etienne,jasv}@isr.ist.utl.pt

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.

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 [12]. Work on uncalibrated reconstruction progressed dramatically in recent years with the works of Hartley [4], Faugeras [2], Maybank [8], Pollefeys et al [7], 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. In some studies [1, 11] some intrinsic parameters are fixed to trivial values. We want to compare, in terms of precision, the effect of these assumptions and the precision achieved in the calibrated case. A study of critical (pathological) cases for self-calibration can be found in [10], and the achievable precision in the calibrated case is addressed in [5]. One contribution

*This work has been supported by the project INCO COPERNICUS Proj. 960174-VIRTUOUS FCT-PBIC/TPR/2550/95

of this paper is to compare the precisions of calibrated and un-calibrated reconstruction. Although the former always performs better, experimentation shows that when more than ten images are available un-calibrated 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 ML [4] (or related [9]) estimate. ML estimators are often reported [9] to converge to the solution only if started close from it. It is the purpose of the "algebraic" algorithm to provide the starting position. In this paper, we study the precision of the ML 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 ML estimator is defined by the observation model, independently from any specific algorithm; we derive analytically its covariance matrix in various cases of interest :

1.1.1 Full reconstruction from the observations only:

This is the most general case, but the estimation is often numerically ill-posed. For example, in [1] 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).

1.1.2 Determination of a reduced set of parameters :

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 [11, 7]. We verify in Section 4.1 the effect on precision of fixing the intrinsic parameters, either to values obtained from a pre-calibration step or to trivial values (corresponding to square pixels and centered principal point).

1.1.3 Full reconstruction with probabilistic prior knowledge

The likelihood function may be modified to take into account a-priori knowledge expressed probabilistically, e.g. assuming that either structure or calibration follow a known distribution. 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 "trivial" distribution, e.g. the expected value of the principal point is the center of the image, and that its standard deviation is approximately 10 percent of the image size [1]. This is the probabilistic counterpart of fixing the principal point to image center, as in Section 1.1.2. 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. (11) to (13). 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 4.1. One important contribution of this paper lies in showing how big the variances of the considered estimators are in practice.

2 Observation Model

2.1 Notations

We consider that a set of P points has been tracked over a sequence of N images. The following notation is adopted :

- $p \in \{1, 2, \cdots, P\}$ and $n \in \{1, 2, \cdots, N\}$ are the indices used for numbering points and images, respectively.
- $\mathbf{x}_p \in R^3$ is the vector of the coordinates, in the world frame, of the p^{th}point. Its components are x_{pi}, for $i \in \{1, 2, 3\}$. The symbol $\mathcal{X}$ shall denote all the x_{pi}, for $i \in 1..3$, and $p \in 1..P$.

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 the n^{th} image frame. It can be uniquely defined by 3 parameters $\mathbf{w}_n$. $\mathcal{W}$ will represent all the $\mathbf{w}_n$, for $n \in 1..N$.
- $\mathbf{T}_n$ be the coordinates of the world frame origin, expressed in the n^{th} camera frame. $\mathcal{T}$ will represent $\mathbf{T}_1 ... \mathbf{T}_n$.

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 $\tilde{\mathbf{u}}_{np} = (\tilde{u}_{np1}, \tilde{u}_{np2})$:

$$\tilde{u}_{npi} = \frac{\mathbf{a}_{ni}\mathbf{x}_p + t_i}{\mathbf{a}_{n3}\mathbf{x}_p + t_3} \quad i \in \{1, 2\} \tag{1}$$

Taking into account the intrinsic parameters and noise yields :

$$\mathbf{u}_{np} = \mathbf{B}\tilde{\mathbf{u}}_{np}\mathbf{C} + \varepsilon_{np} \tag{2}$$

with

[1] Lenz and Tsai [6] cite values of this order of magnitude.

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

Let $\mathcal{U}$ denote all the observations u_{npi}, for $i \in \{1, 2\}$, $p \in \{1 \cdots P\}$ and $n \in \{1 \cdots N\}$; 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 : $\Theta = (\mathcal{X}, \mathcal{W}, \mathcal{T}, K)$. For a given Θ, the *prediction* of the $(n, p, i)^{\text{th}}$ observation is defined as :

$$v_{npi}(\Theta) \quad = \quad \mathbf{b}_i \tilde{\mathbf{u}}_{np}(\Theta) + c_i \quad \text{where} \quad \tilde{u}_{npi}(\Theta) = \frac{\mathbf{a}_{ni}\mathbf{x}_p + t_{ni}}{\mathbf{a}_{n3}\mathbf{x}_p + t_{n3}} \tag{3}$$

2.2 Likelihood function :

An estimator of the parameters is a function $\mathcal{E} : \mathcal{U} \to \Theta$, that associates a parameter vector Θ to a data set $\mathcal{U}$. The Maximum Likelihood estimator is defined by a global minimum of the (inverted) log-likelihood, taken as a function of Θ :

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

This function does *not* have a unique minimum : it is well-known that the reconstruction is defined only up to a similarity. A way of resolving the ambiguity is to constrain the structure parameters to be centered ($\sum_p \mathbf{x}_p = 0_3$) 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 ($\mathbf{A}_1 = \mathbf{I}_3$). After removing the first rotation matrix, and the upper right coefficient in $\mathbf{B}$ from the parameter vector Θ, the restricted parameter set is defined as the zeros of the function :

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

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 [10]. In this article, we consider that the minima of Q that verify $S(\Theta) = 0$ are isolated. Note that constraints are not needed when using a prior on the structure.

Probabilistic prior on a subset of the parameters : Assuming prior knowledge on the parameters leads to adding a term to the log-likelihood function $Q(\mathcal{U}, \Theta)$: a prior on the structure, $\mathcal{X}^* \sim N(\mathcal{X}_0, \Sigma_X)$ adds the following term :

$$(\mathcal{X} - \mathcal{X}_0)^T \Sigma_X^{-1} (\mathcal{X} - \mathcal{X}_0)$$

A prior on the calibration is treated likewise.

3 Covariance of estimators

We show how the covariance matrix of Maximum-Likelihood estimator derives from its definition. This derivation, which can be found, e.g. in [3], is then applied to the problem of estimating structure, motion and calibration.

3.1 Definition of the estimators :

The maximum likelihood estimate is defined by :

$$\hat{\Theta} = \arg_{\Theta} \min Q(\mathcal{U}, \Theta) \quad \text{subject to } S(\Theta) = 0$$

At the minimum, the derivative of Q is a linear combination of the derivatives of constraints.

$$D_{\Theta} Q(\mathcal{U}, \hat{\Theta}) + \Lambda D_{\Theta} S(\hat{\Theta}) = 0_{1 \times \text{size}(\Theta)} \tag{6}$$

where D_{Θ} is the differentiation operator. Assuming that the observation noise, and $\Delta\Theta = \hat{\Theta} - \Theta^*$ are not too large, the first order development of S and Q in $(\Theta^*, \mathcal{U}^*)$ is a good approximation. The following holds:

$$\begin{array}{rcl} 0 = S(\Theta^*) & \simeq & S(\hat{\Theta}) + D_{\Theta} S(\hat{\Theta}) \Delta\Theta = D_{\Theta} S(\hat{\Theta}) \Delta\Theta, \\ D_{\Theta} Q(\mathcal{U}, \hat{\Theta}) & \simeq & D_{\Theta} Q(\mathcal{U}^*, \Theta^*) + D^2_{\Theta\Theta} Q(\mathcal{U}, \hat{\Theta}) \Delta\Theta + D^2_{\Theta U} Q(\mathcal{U}, \hat{\Theta}) \varepsilon \end{array} \tag{7}$$

Note that the linear development can be done either at $(\mathcal{U}^*, \Theta^*)$ or at $(\mathcal{U}, \hat{\Theta})$. eq.s (6) and (7) may be written in matrix form :

$$\begin{bmatrix} H & G^T \\ G & 0_{4\times 4} \end{bmatrix} \begin{bmatrix} \Delta\Theta \\ \Lambda \end{bmatrix} = - \begin{bmatrix} F\varepsilon \\ 0_4 \end{bmatrix}, \quad \text{where} \quad \begin{array}{rcl} H & = & D^2_{\Theta\Theta} Q(\mathcal{U}^*, \Theta^*) \\ F & = & D^2_{\Theta U} Q(\mathcal{U}^*, \Theta^*) \\ G & = & D_{\Theta} S(\Theta^*), \end{array} \tag{8}$$

This equation uniquely defines $\Delta\Theta$ -if the matrix is invertible- as a linear transformation of the noise ε. The covariance is :

$$\text{Cov} \begin{bmatrix} \Delta\Theta \\ \Lambda \end{bmatrix} \simeq \begin{bmatrix} H & G^T \\ G & 0 \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2 F F^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H & G^T \\ G & 0 \end{bmatrix}^{-1}, \tag{9}$$

We now specialize the above formulas to our case. The definition of Q gives :

$$\begin{array}{lcl} D_{\Theta_i} Q & = & \sum_{npk} D_{\Theta_i} v_{npk} (v_{npk} - u_{npk}) / \sigma^2 \\ D^2_{\Theta_i \Theta_j} Q & = & \frac{1}{\sigma^2} \sum_{npk} D_{\Theta_i} v_{npk} D_{\Theta_j} v_{npk} + D^2_{\Theta_i \Theta j} v_{npk} (v_{npk} - u_{npk}) \\ D^2_{\Theta_i u_{npk}} Q & = & D_{\Theta_i} v_{npk} / \sigma^2 \end{array} \tag{10}$$

At $(\Theta^*, \mathcal{U}^*)$, one has $v_{npk} = u_{npk}$, and thus the second order terms in $D^2_{\Theta_i \Theta_j} Q$ are eliminated. In what follows, these terms are systematically eliminated. Noting that $D_{\Theta} Q^T \cdot D_{\Theta} Q = D^2_{\Theta\Theta} Q / \sigma^2$, and replacing in eqs. (9) yields.

$$\text{Cov}\begin{bmatrix}\Delta\Theta\\ \Lambda\end{bmatrix}=\begin{bmatrix}H & G^T\\ G & 0\end{bmatrix}^{-1}\begin{bmatrix}H & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}H & G^T\\ G & 0\end{bmatrix}^{-1} \tag{11}$$

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 taking the expected unit square module of the parameters to be $\simeq 1$, since this is the module of the x_{pi} parameters. For **K**, based on our experience, and on remarks by Lenz and Tsai [6], we assumed that the parameters b_{21}, $b_{22}-1$, C_1 and C_2 all have approximately an expected absolute value of 0.1, which leads us to the parameterization $\mathbf{K}=10[b_{21}/b_{11}, b_{22}/b_{11}-1, C_1, C_2, \log f]$, where f is the focal length. 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.

3.1.1 Covariance when a prior is used :

A prior on the structure, $\mathcal{X}^*\simeq N(\mathcal{X}_0,\Sigma_X)$, modifies the likelihood function, as discussed before. The term $(\mathcal{X}_0-\mathcal{X})\Sigma_X^{-1}$ is added to the differential of Q with respect to $\mathcal{X}$, D_XQ, and $\Sigma_X^{-1}$is added to $D^2_{XX}Q$, the second differential of Q with respect to $\mathcal{X}$. This prior renders the constraint defined in eq. (5) irrelevant : all the parameters can be uniquely determined without having to restrict the parameter set. Furthermore, the differential of Q in $(\Theta^*,\mathcal{U}^*)$ is not (in general) zero anymore. Altogether, this yields the normal equations :

$$H^{+x}\Delta\Theta=F\varepsilon+\begin{bmatrix}\Sigma_X^{-1}(\mathcal{X}_0-\mathcal{X}^*)\\ 0_{6N+5}\end{bmatrix},\quad \text{where}\quad H^{+x}=H+\begin{bmatrix}\Sigma_X^{-1} & 0\\ 0 & 0\end{bmatrix}$$

is the modified matrix of second derivatives. The covariance of the estimate is then :

$$\text{Cov}\,\Delta\Theta=(H^{+x})^{-1}H^{+x}(H^{+x})^{-1}=\left(H^{+x}\right)^{-1} \tag{12}$$

A prior on the calibration parameters is treated likewise, but one keeps the constraints **S**, and the matrix of **G** of derivative :

$$\text{Cov}\begin{bmatrix}\Delta\Theta\\ \Lambda\end{bmatrix}=\begin{bmatrix}H^{+k} & G^T\\ G & 0\end{bmatrix}^{-1}\begin{bmatrix}H^{+k} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}H^{+k} & G^T\\ G & 0\end{bmatrix}^{-1} \tag{13}$$

3.1.2 Fixed parameters :

Fixing $\mathbf{K}=\mathbf{K}_0$ for some value $\mathbf{K}_0\neq\mathbf{K}^*$, and assuming that $\mathbf{K}^*\sim N(\mathbf{K}_0,\Sigma_K)$, the covariance matrix of the estimate $\hat{\Theta}_1=(\mathcal{X},\mathcal{W},\mathcal{T})$, takes the form:

$$\text{Cov}\begin{bmatrix}\Delta\Theta_1\\ \Lambda\end{bmatrix}=\begin{bmatrix}H_{11} & G_1^T\\ G_1 & 0_{4\times4}\end{bmatrix}^{-1}\begin{bmatrix}H^T+H_{1K}\Sigma_K H_{1K}^T & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}H_{11} & G_1^T\\ G_1 & 0\end{bmatrix}^{-1} \tag{14}$$

Where and H_{11},G_1 are the appropriate sub-blocks of H and G.

4 Experimental Results

Measure of error : We study separately the errors on the parameters $\mathcal{X}$, $\mathcal{W}$, $\mathcal{T}$ and $\mathbf{K}$. 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 $\sqrt{E(||\mathbf{x}_p - \mathbf{x}_p^*||^2)}$ and $\sqrt{E(||\mathbf{K} - \mathbf{K}^*||^2)}$. For $\mathcal{T}$, $\sqrt{E(||\mathbf{t} - \mathbf{t}^*||^2)}$ is used. For $\mathcal{W}$, the measure is the standard deviation of the angle formed between axes of the true and the estimated camera frames, $\sqrt{E(||\mathbf{w}_n - \mathbf{w}_n^*||^2)}$. When a prior on the structure is used, and the first camera may be different from the identity matrix, $\mathbf{w}_n$ is taken as the difference between the first camera and the n^{th}camera.

4.1 Validation of the analytical expressions of covariances:

The covariance matrices (11)-(14) are obtained using the approximations (7) and (10); we must verify that they are valid in practice. This is done 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 5 images. The noise is 40dB[2]. 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 40dB, 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 Figure 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.

4.2 Variance of estimators:

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 a close-up of Figure 1 , taken from $\simeq .75 - 1$m. Total rotation is $\simeq 30$ degrees. Since the 3D point positions are known, one may retrieve $\mathcal{W}, \mathcal{T}$and $\mathbf{K}$ with precision, and later use these values as ground-truth. The precision of this calibration step, as determined by eq. (12), is shown on the first line in Table 1. We assumed that the observation noise had variance $1e-4$, based on the residuals $v_{npi}(\hat{\Theta}) - u_{npi}$ which have variance $7e-5$. We further assumed that the standard deviation of the error on the ground-truth was of 1%, corresponding to $\simeq 2mm$. We label the lines of the table in the following manner :

- **Calib** contains results obtained when calibrating.
- **ML** : Maximum-Likelihood estimator, with covariance defined in eq. (11).
- **TP** (Trivial Prior) and **CP** (Calibration Prior) are for estimators with a prior on $\mathbf{K}$.
- **TF** and **CF** denote the estimators with fixed intrinsic parameters, to trivial values (**TF**), or values obtained by a previous calibration step (**CF**).

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

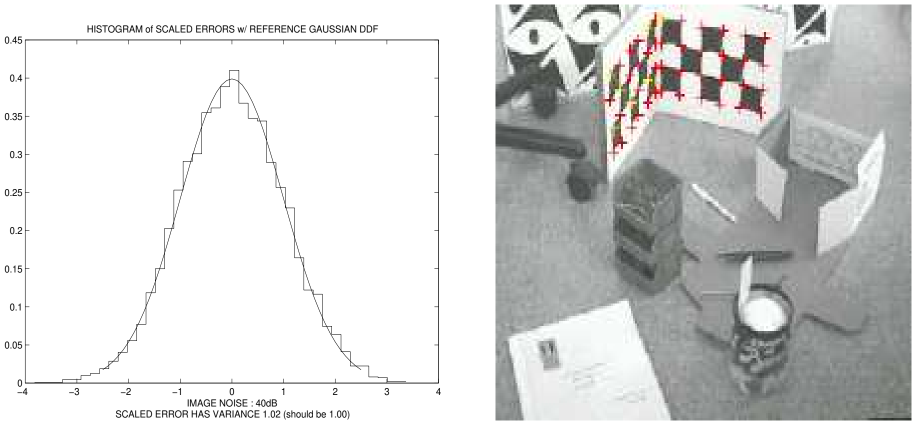

Figure 1: .
Left : Histogram of scaled errors of the ML estimator. In abscissa is the error, divided by the theoretical standard deviation. The Gaussian density function is superposed for comparison. The observed variance is 1.02, while parameters $\mathcal{X}$, $\mathcal{W}$, $\mathcal{T}$ and $\mathbf{K}$ have variances in $[0.98, 1.09]$.
Right : An image from the Long-Range sequence.

The most important features apparent from this table are :

- The ML estimator (second line) gives totally wrong the intrinsic parameters.
- The precision obtained with calibration information (fourth and sixth lines) is much better than that obtained without (second, third and fifth lines).
- Without pre-calibration, the use of a trivial prior (third line) provides much better estimates than either the ML or the trivially-fixed-parameter (TF) estimators (second and fifth lines).

Long-Range : the grid (shown on right, in Figure 1) is seen along 12 images, from $1.5 - 2.5$m, and the maximum camera-camera distance is $\simeq 1.3$m. The variances of the five tested estimators are displayed in Table 1. For the **CP** and **CF** (fourth and sixth lines) estimators, the covariance of the prior is that of short-range calibration. The ML estimator and the estimator with trivial prior (second and third lines) perform nearly as well as the estimators that use prior calibration (fourth and sixth lines). The estimator with fixed trivial parameters, however, appears to behave relatively poorly. The first point appears to be due to the increased number of images used.

4.3 Influence of the number of images :

Figure 2 plots the base 10 logarithm of the variance in structure and in intrinsic parameters, as a function of the number of images used. Long sequences of uncalibrated images

	Short Range, $N = 5, P = 48$				**Long Range**, $N = 12, P = 48$			
	x	**w**	**t**	**K**	**x**	**w**	**t**	**K**
Calib	0.0079	0.20	0.0350	0.0872	0.008	0.34	0.556	0.239
ML	0.429	2.66	1.08	3.30	0.0816	0.68	1.44	0.440
TP	0.134	1.00	0.437	0.767	0.0774	0.65	1.35	0.373
CP	0.084	0.61	0.0705	0.0867	0.0646	0.54	0.329	0.0844
TF	0.299	2.54	1.63	1.00	0.148	1.83	7.10	1.00
CF	0.0840	0.61	0.0706	0.0872	0.0646	0.54	0.330	0.0872

Table 1: Each column contains the **expected standard deviation** of either the structure **x**, orientation **w**, position **t** or calibration parameters **K**. **w** is expressed in degrees.

allow as good 3D reconstruction as short calibrated sequences, whereas short uncalibrated sequences give poor results. In all cases, it is better to use a trivial prior than to fix the intrinsic parameters to trivial values.

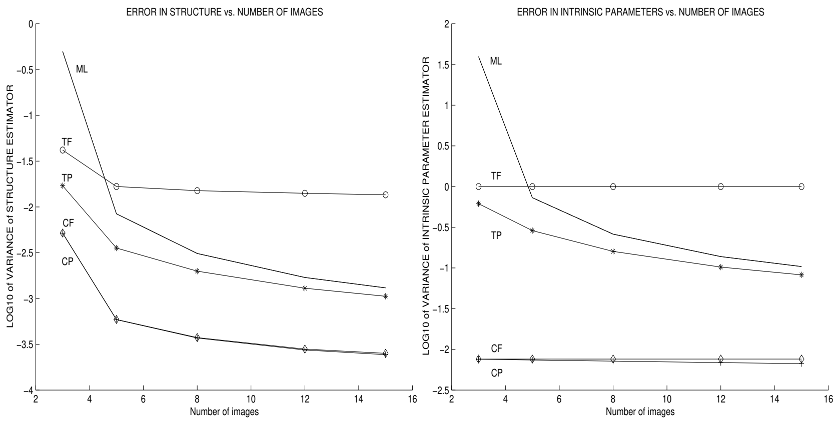

Figure 2: Log (base 10) of the error on the structure parameters $\mathcal{X}$ and calibration parameters, **K**. The curves are tagged **ML**, **TP**, **CP**, **TF** and **CF** as explained in the text. Thirty setups were generated, each with twelve 3D points, generated as Gaussian white noise, and then normalized. The camera orientations have Euler angles independently uniformly distributed in $[\pm\pi/4, \pm\pi/4, \pm\pi/8]$. The scene-camera distance is 6 to 12 times the size of the scene. The precision of the camera calibration is that of Table 1, and the observations noise has variance $1e-4$.

5 Conclusions:

We have presented analytical expressions of the covariance matrices of the estimators, and verified experimentally their validity. We compared the precision of various 3D-from-matched-points algorithms, and showed how it depends on the physical setup. To summarize :

- 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 above.
- Long sequences of uncalibrated images allow as good 3D reconstruction as short calibrated sequences. Therefore it shows the potential quality of euclidean reconstruction obtained from long uncalibrated sequences.

We are presently working to better analyze the influence of the sequence length, number of points and noise in image measurements. On the analytical side, the present work could be extended to variable intrinsic parameters .

References

[1] Sylvain Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *Proc. of the 6th ICCV*, 98.

[2] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In G.Sandini, editor, *Proc. ECCV*, pages 563–578, 92.

[3] R.M. Haralick. Propagating covariance in computer vision. In *Proc. Workshop on Performance Characteristics of Vision Algorithms*, pages 1–12, 96.

[4] R.I. Hartley. Euclidean reconstruction from uncalibrated views. In *in 2nd Proc. Europe-U.S. Workshop on Invariance*, pages 237–256, 93.

[5] N. Ahuja J. Weng and T.S. Huang. Optimal motion and structure estimation. *IEEE Trans. PAMI*, 15(9):864–884, 93.

[6] Reimar K. Lenz and Roger Y. Tsai. Techniques for calibration of the scale factor and image center for high-accuracy 3d machine vision metrology. *IEEE Trans PAMI*, 10(5):713–720, Sept 88.

[7] Reinhard Koch Marc Pollefeys and Luc Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proc. of the 6th ICCV*, pages 90–95, Jan 98.

[8] S.J. Maybank and O.D. Faugeras. A theory of self-calibration of a moving camera. *Intl. J. Computer Vision*, 8(2):123–151, 92.

[9] L.Quan R.Mohr, F.Veillon. Relative 3d reconstruction using multiple uncalibrated images. In *Proc. of the IEEE CVPR*, pages 543–548, 93.

[10] Peter Sturm. Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction. In *Proc. CVPR*, 97.

[11] Richard Szeliski and Sing Bing Kang. Shape ambiguities in structure from motion. *IEEE Trans. PAMI*, 19(5):506–512, May 97.

[12] Paul R. Wolf. *Elements of photogrammetry, with air photo interpretation and remote sensing.* McGraw-Hill, 83. 2nd ed.