View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by HAL Clermont Université

Uncertainty Ellipsoids Calculations for Complex 3D Reconstruction

Maxime Lhuillier, Mathieu Perriollat

To cite this version:

Maxime Lhuillier, Mathieu Perriollat. Uncertainty Ellipsoids Calculations for Complex 3D Reconstruction. 2006, IEEE, pp.0, 2006. <hal-00091146>

HAL Id: hal-00091146

https://hal.archives-ouvertes.fr/hal-00091146

Submitted on 5 Sep 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Uncertainty ellipsoids calculations for complex 3D reconstructions

Maxime Lhuillier and Mathieu Perriollat
LASMEA-UMR 6602, Université Blaise Pascal/CNRS, 63177 Aubière Cedex, France
Maxime.Lhuillier@univ-bpclermont.fr, maxime.lhuillier.free.fr

***Abstract*— Many methods exist for the automatic and optimal 3D reconstruction of camera motion and scene structure from image sequence ('Structure from Motion' or SfM). The solution to this problem is not unique: we obtain an other solution by changing the coordinate system where points and cameras are defined. On the other hand, existing methods provide some measure of confidence or uncertainty for the estimation when the ground truth is not available, using a gauge constraint which settle the reconstruction coordinate system. Here we justify and describe a method which estimate the uncertainty ellipsoids, when previous methods are not straightforward to use due to a huge number of parameters to fit. Many examples are given and discussed for big reconstructions.**

I. Introduction

After a short description of the automatic 3D reconstruction, we present previous works about the uncertainty estimation.

Automatic and "optimal" 3D reconstruction: Robust and automatic estimation, from image sequence, of a calibrated (or not) camera and points of the scene has been an active research topic during the last decade [5], [6]. Several successful systems now exist [1], [8], [11], [12], [14].

First, image points are selected and matched between successive images. Then, geometry of sub-sequence of 2 or 3 images is computed using robust methods based on random process of these matched points.

Finally these geometries are joined and then the reprojection errors are minimized. These errors are the differences between images points m_{ij} and the projections "C_iX_j" of the n 3D points X_j by the m cameras C_i.

So, the score minimized by all "optimal" 3D reconstruction algorithms looks like:

$$E(C_1, \cdots, C_m, X_1, \cdots, X_n) = \sum_{i,j} ||C_iX_j - m_{ij}||^2,$$

with $||.||$ the Euclidean norm in the images. A 3D reconstruction which minimizes a such score is optimal in the statistical sense, as described in more details in section II.

The camera parameters to estimate are the extrinsic ones (a rotation matrix R_i representing the orientation and a vector for the center t_i) and sometimes few intrinsic parameters.

This problem has many solutions because a change of the 3D coordinate system has not effect on the cost function E.

So we have an over-parameterization which defines both the 3D reconstruction and its coordinate system (defined by 7 parameters: 3 for translation, 3 for rotation and 1 for scale factor).

Uncertainty ellipsoids: Although methods exist to evaluate uncertainty of "optimal" 3D reconstruction [6], [15], few publications in Robotic and Computer Vision really use it.

In a different context, the Kalman Filter applied to 3D reconstruction provides the uncertainty for an acceptable estimation of cameras and points. Although not optimal, this tool is so efficient to allow the simultaneous localization and mapping from a video sequence [16]. Our focus is the uncertainty calculation in the context of an optimal estimation, provided by bundle adjustment.

This paragraph refers to covariance matrices and confidence ellipsoids, which are described in detail in section II.

The covariance matrix may be computed as the inverse of the Hessian H of the final score E given by the minimization [13]. For this method, the Hessian must be invertible and the inverse must be computable even if the dimension of the Hessian is large.

In our case, H is not invertible because of over-parameterization. As suggested in [11] it is possible to make H be invertible by taking out 7 parameters to estimate, for example R_1, t_1 and $(t_m)_z$. In this case, a method is reported in [6] to compute some blocks of H^{-1}, taking advantage of the sparse block structure [2] of H. From diagonal block of H^{-1}, we compute confidence ellipsoids (for example for 3D point X_j and camera center t_i) corresponding to a given probability. A similar method is proposed by [7] to compute camera uncertainties. But in reality, [11] computes point uncertainties assuming that the result when several parameters are fixed is the same as in the case where all camera uncertainties are zero (with this assumption, the uncertainties are given by the inversion of the 3×3 diagonal blocks of H).

In the methods reported in [15] and [10], the coordinate system of the 3D reconstruction is fixed thanks to a constraint called "gauge constraint", so that the covariance matrix of the estimated parameters may be computed. This constraint is implicitly defined as $c(x) = 0$ with x the vector of parameters to estimate. This expression provides more global constraints in relation to the trivial ones suggested in [11]. The camera-based constraints (i.e. only based on camera parameters) are useful to evaluate camera uncertainties, especially for localization applications. The point-based constraints are useful to compute point uncertainties, suitable for scene reconstruction. The principles are summarized in [10], [15] and experimented on small 3D reconstructions in [10]. All covariance matrices given by this method (i.e. covariance matrices for different

gauge constraints) are linked together by multiplying on left and on right by linear projectors which depend on c and H: for small problem, it is easy to compute all covariance matrices from only one. Thus [10], [15] suggest methods to compute a covariance matrix for non trivial constraints from one covariance matrix which is easier to compute.

However there is a lack in the methods described in [10], [15], and they can't be applied directly to the real world high dimensional problems as those proposed in this paper.

Contributions: There are several contributions in this article. The section II is a survey about the covariance matrices and confidence ellipsoids. It is all-encompassing and nearly self-contained, and should be easier to understand than papers it is based on [6], [10], [15]. We choose to use here the perturbation analysis instead of a more rigorous presentation based on the Implicit Function Theorem as those of [4] and [3] in the case of invertible Hessian. The section III describes our method to compute the diagonal blocks of a covariance matrix (and thus the confidence ellipsoids) for any non-trivial global constraint as $c(x) = 0$ with x the set of estimated parameters. Its interest is to rigorously justify a method to deal with high dimensional problems (out of reach for the method used by [10]). At last, in the section IV we experiment our method with high 3D reconstructions and several constraints before concluding.

II. Confidence ellipsoids (survey)

This section presents the confidence ellipsoids for an optimal estimation of the parameters of an image sequence geometry (i.e. the cameras and 3D points parameters). We hope it will be easier to understand than other papers it is based on [6], [10], [15]. It may refer to some properties of the pseudo-inverse, given in the Appendix.

A. Notations and statistical modeling

2D Projections: m_{ij} represents the coordinate of the matching point in the image i which corresponds to the 3D point j. In other word, the 3D point j has been reconstructed by triangulation from matched 2D points m_{ij} detected in several images.

These 2D points are noisy measurements of their true unknown positions (denoted $\bar{m}_{ij}$) and are considered as random variables subject to Gaussian noise with mean $\bar{m}_{ij}$ and covariance matrix $\bar{\sigma}^2 I_2$ (the variance $\bar{\sigma}^2$ is unknown too).

3D points and cameras: From a set of m_{ij}, an automatic method has estimated the camera i parameters (denoted C_i) and those of the 3D point j, denoted X_j, such as the m_{ij} fits the projection of X_j by C_i at best. We note these projections as a matrix product "$C_i X_j$". More precisely, we used bundle adjustment to find the parameters minimizing a sum of square Euclidean norm $\sum_{(i,j)\in M} ||C_i X_j - m_{ij}||^2$. M is the set of possible pairs (i, j).

We assume that a true solution exists $\bar{C}_i$, $\bar{X}_j$, in the sense that $\bar{m}_{ij} = \bar{C}_i \bar{X}_j$. There are many solutions because we get an other solution from the first one by changing the Euclidean coordinate system and the scale factor.

Statistical model: The unknown parameters of the model are: $\bar{\sigma}, \bar{C}_1, \cdots, \bar{C}_m, \bar{X}_1, \cdots, \bar{X}_n$. The probability density functions of measurements m_{ij} are

$$f(m_{ij}, \bar{\sigma}, \bar{C}_i \bar{X}_j) = \frac{1}{2\pi\bar{\sigma}^2} e^{-\frac{1}{2\bar{\sigma}^2}||\bar{C}_i \bar{X}_j - m_{ij}||^2}$$

One deduces the probability density function of our model thanks to the independence hypothesis:

$$f(m_{ij}, (i,j) \in M) = \frac{1}{(2\pi\bar{\sigma}^2)^{|M|}} e^{-\frac{1}{2\bar{\sigma}^2}\sum_{(i,j)\in M}||\bar{C}_i \bar{X}_j - m_{ij}||^2}$$

where $|M|$ is the size of M.

Maximum likelihood estimation: The maximum likelihood estimation of the parameters $\bar{\sigma}, \bar{C}_1, \cdots, \bar{X}_1, \cdots$ implies to find $\hat{\sigma}, \hat{C}_1, \cdots, \hat{X}_1, \cdots$ that maximize the probability density function of the model. That's why $\hat{C}_i, \hat{X}_j$ minimize $\sum_{(i,j)\in M} ||C_i X_j - m_{ij}||^2$ and $\hat{\sigma}^2 = \frac{1}{2|M|}\sum_{(i,j)\in M} ||\hat{C}_i \hat{X}_j - m_{ij}||^2$. As we explain before, such $\hat{C}_i, \hat{X}_j$ are the result of the bundle adjustment method and many solutions exist as with $\bar{C}_i, \bar{X}_j$.

B. Error propagation

The aim of this paragraph is to answer the question: how is perturbed the value of the estimated parameters $\hat{C}_i, \hat{X}_i$ compared to true values $\bar{C}_i, \bar{X}_i$, when we perturb the true value of measurements $\bar{m}_{ij}$ to m_{ij} ? It is important because the best confidence of an estimation result will be obtained for the lowest perturbation on parameters. After introducing notations and summarizing the principle, the description becomes deductive.

Notations: $rk(A), tr(A), Ker(A), Im(A), A^+, dim(B)$ are the rank, the trace, the kernel, the image, the pseudo-inverse of matrix A, the dimension of vector space B.

Let x (respectively, y) be the vector concatenating the vectors $C_1, \cdots, X_1, \cdots$ (respectively, the vectors m_{ij}). The true values $\bar{x}, \bar{y}$ are defined by a similar way and $\hat{x}$ represents the estimation of x.

Let f be the function that computes 2D points from 3D points and cameras, so: $f(\bar{x}) = \bar{y}$ and $\hat{x}$ minimizes $x \mapsto ||y - f(x)||^2$. Then we define $dx = x - \bar{x}$ and $dy = y - \bar{y}$. With the chosen model, the mean of the perturbation dy is 0 and its covariance is $C_{dy} = \bar{\sigma}^2 I_{2|M|}$.

Summary: A summary of the problem is given by figures 1 and 2. The figure 1 shows the parameter space X and the measurement space Y.

The space X is too big to have uniqueness of the solution of the estimation problem because the value of $f(x)$ is the same for all euclidean basis and scale factor changes for x. Therefore we introduce a variety defined by a constraint $c(x) = 0$ (called "gauge constraint" [10], [15]), such that onto this variety, only one solution exists.

The figure 2 shows the approximations made in X and Y. In X, the variety defined by c is approximated by its affine tangent space in $\bar{x}$. The varieties where f is constant (called "gauge orbits" [10], [15]) are linearized too. In Y, the variety $f(X)$ is approximated by its affine tangent space in $f(\bar{x}) = \bar{y}$.

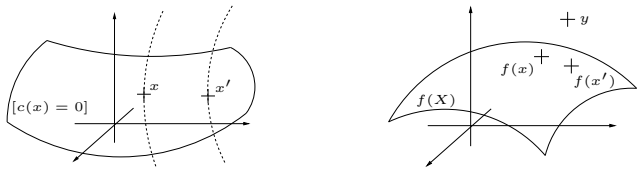


Fig. 1. **Left:** two points x and x' in the parameter space X. Each dotted "curve" represents a set of points where the function f is constant. Each one is a variety of dimension 7 and is parameterized by the group of similarities in R^3. In order to obtain uniqueness of the solution, we cut back the parameter space X on a shorter space implicitly defined by $\{x \in X, c(x) = 0\}$, and so each "curve" cuts it just in one point. **Right:** a point y in the measurement space Y, the image of X by f and of its points x and x' ($f(X)$, $f(x)$ and $f(x')$). f is one to one between $\{x \in X, c(x) = 0\}$ and $f(X)$.

Thanks to these approximations, the relation between dx and dy becomes linear, and so one deducts the distribution for dx, knowing these for dy.

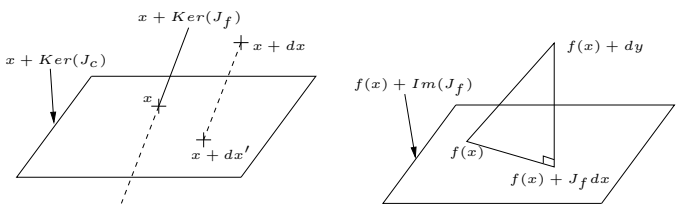


Fig. 2. **Left:** Approximation around x in X. The variety $\{x \in X, c(x) = 0\}$ is approximated by $x + Ker(J_c)$ with J_c the Jacobian matrix of c in x. The curves passing by x and $x+dx$ are approximated by $x+Ker(J_f)$ and $x+dx+Ker(J_f)$ with J_f the Jacobian matrix of f in x. The projection of a given dx on $Ker(J_c)$ parallel to $Ker(J_f)$ is dx'. **Right:** Approximation around $f(x)$ in Y. The variety $f(X)$ is approximated by $f(x)+Im(J_f)$. Given a dy, dx is such as $||dy - J_f dx||$ is minimum, therefore $dy - J_f dx$ and $Im(J_f)$ are orthogonal.

Approximation of the projection function: One approximates f in a neighborhood of a possible $\bar{x}$, and assumes dx is small:

$$f(x) = f(\bar{x} + dx) \approx f(\bar{x}) + J_f dx = \bar{y} + J_f dx$$

J_f is the Jacobian matrix of f in $\bar{x}$. So a dx solution (i.e. a dx that minimizes $x \mapsto ||y - f(x)||^2$) should minimize

$$dx \mapsto ||\bar{y} + dy - f(\bar{x} + dx)||^2 \approx ||dy - J_f dx||^2,$$

and so $dx \in (J_f^\top J_f)^+ J_f^\top dy + Ker(J_f)$.

Non uniqueness: We know that many exact parameters $\bar{x}$ and estimated ones $\hat{x}$ are possible, and they are linked by changing the coordinate system or the scale factor. More precisely, those coordinate system and scale factor changes can be written as parameterization $\bar{x}(\theta)$ and $\hat{x}(\theta)$ of class $\mathcal{C}^1$ with θ of dimension 7 (6 for euclidean coordinate system and 1 for scale factor).

We deduct that $f(\bar{x}(\theta))$ is constant, and then (by composed differentiations) that the dimension of $Ker(J_f)$ is at least 7 on any evaluating point. We suppose that $dim(Ker(J_f)) = 7$.

The choice problem: Without additional constraints on dx, we only know that

$$dx \in (J_f^\top J_f)^+ J_f^\top dy + Ker(J_f), \quad dim(Ker(J_f)) = 7.$$

With a well-defined dx, we could model dx by a random variable as dy.

A first possible choice is dx that minimizes $||dx||$. In this case, $dx = (J_f^\top J_f)^+ J_f^\top dy$.

An other choice is given by a constraint on the parameters x expressed as a function $c(x) = c(\bar{x}) = 0$: the resulting constraint on dx by differentiation is:

$$0 = c(\bar{x} + dx) \approx c(\bar{x}) + J_c dx = J_c dx$$

with J_c the Jacobian matrix of c in $\bar{x}$. So we can define dx if $Ker(J_c)$ is suitable: we suppose that $Ker(J_c)$ and $Ker(J_f)$ are supplementary in the parameter space, and so we have $dx = P_f^c (J_f^\top J_f)^+ J_f^\top dy$ with P_f^c the linear projector on $Ker(J_c)$ parallel to $Ker(J_f)$.

To get a suitable constraint [15], we must choose 7 scalar equations that fix a unique choice for $\bar{x}$ and $\hat{x}$ (i.e. a unique choice of euclidean coordinate system and scale factor).

Covariance matrix: As the perturbation dx is linearly defined by dy and knowing that dy obeys a zero-mean Gaussian distribution of covariance $C_{dy} = \bar{\sigma}^2 I_{2|M|}$, we deduct that dx obeys a zero-mean Gaussian distribution of covariance C_{dx}. We introduce:

$$\begin{aligned} C_{dx}^{\perp} &= (J_f^\top J_f)^+ J_f^\top C_{dy} J_f (J_f^\top J_f)^+ = \bar{\sigma}^2 (J_f^\top J_f)^+ \\ C_{dx}^{c} &= \bar{\sigma}^2 P_f^c (J_f^\top J_f)^+ (P_f^c)^\top. \end{aligned}$$

In the case where $\|dx\|$ is minimal, we have $C_{dx} = C_{dx}^{\perp}$, and when dx is defined with the constraint c, we have $C_{dx} = C_{dx}^{c}$.

Covariance matrix approximation: Theses covariance matrices have been defined according to $\bar{\sigma}$, P_f^c and J_f at an unknown point $\bar{x}$. So we should estimate C_{dx} by $\hat{\sigma}^2 (J_f^\top(\hat{x}) J_f(\hat{x}))^+$ or $P_f^c(\hat{x}) \hat{\sigma}^2 (J_f^\top(\hat{x}) J_f(\hat{x}))^+ (P_f^c(\hat{x}))^\top$ according to the constraint choice, using the fact that the pseudo-inverse is a continuous function on the set of symmetric positive matrices of fixed rank (it is not continuous anywhere).

The aim of the following paragraph is to improve the maximum likelihood estimator $\hat{\sigma}^2 = \frac{1}{2|M|} ||y - f(\hat{x})||^2$ of $\bar{\sigma}^2$ in the expressions of C_{dx}.

Improving the estimation of $\hat{\sigma}^2$: We have seen that the function that maps dy to $dx = \hat{x} - \bar{x}$ and that minimizes $dx \mapsto ||dy - J_f dx||^2$ is linear for all cases. We deduce that the application P that maps dy to $dy - J_f(\hat{x} - \bar{x})$ is linear too. P is an orthogonal projector onto $Im(J_f)^\perp$ parallel to $Im(J_f)$ because $J_f(\hat{x} - \bar{x})$ is the nearest point of dy onto $Im(J_f)$: we have $tr(P) = rk(P)$ and $P^\top = P = P^2$.

So $y - f(\hat{x}) \approx dy - J_f dx$ obeys a centered Gaussian distribution of covariance $P(\bar{\sigma}^2 I_{2|M|})P^\top = \bar{\sigma}^2 P$, and we have $tr(P) = rk(P) = 2|M| - rk(J_f) = 2|M| - (|x| - 7)$ with $|x|$ the number of parameters to estimate and 7 the dimension of $Ker(J_f)$. Denoting E the expected value of a random

vector, we get:

$$\begin{aligned} E(||y-f(\hat{x})||^2) &\approx E(||dy-J_f dx||^2) \\ &\approx tr(E((dy-J_f dx)(dy-J_f dx)^\top)) \\ &\approx tr(\bar{\sigma}^2 P) \\ &\approx (2|M|-(|x|-7))\bar{\sigma}^2, \end{aligned}$$

and so that $\tilde{\sigma}^2 = \frac{1}{2|M|-(|x|-7)}||y-f(\hat{x})||^2$ is a approximation of a non-biased estimator of $\bar{\sigma}^2$ (contrary to the maximum likelihood estimator $\hat{\sigma}^2$ which is biased). So we replace $\hat{\sigma}^2$ with $\tilde{\sigma}^2$ in the expressions of C_{dx}.

C. Confidence ellipsoids

With the choice of the perturbation dx and the approximations to estimate the resulting covariance, we can say that $\hat{x}$ obeys a Gaussian distribution of mean $\bar{x}$ and covariance C_{dx}. Now, the covariance matrix of a subset of parameters of $\hat{x}$ (for example, the 3 coordinates of a camera center or a 3D point) is easy to calculate.

Let $\bar{x}_i, \hat{x}_i, C_i$ be the exact value (unknown), the estimation (known) and the covariance matrix (a diagonal sub-block of C_{dx}, known) for a center of a camera i or a point i.

So $(\hat{x}_i-\bar{x}_i)^\top C_i^{-1}(\hat{x}_i-\bar{x}_i)$ satisfies a $\mathcal{X}_2$ distribution with 3 degrees of freedom. Let $E_i(x,p)$ be the ellipsoid of R^3 defined by: $E_i(x,p) = \{\tilde{x} \in R^3, (\tilde{x}-x)^\top C_i^{-1}(\tilde{x}-x) \leq F_3(p)\}$ with F_n the $\mathcal{X}_2$ cumulative distribution with n degrees of freedom.

That means that $\hat{x}_i \in E_i(\bar{x}_i, p)$ with a probability of p, and so $\bar{x}_i \in E_i(\hat{x}_i, p)$ with a probability of p.

Finally, we obtain an ellipsoid centered on $\hat{x}_i$ in which the exact unknown value $\bar{x}_i$ belongs with a probability of p, for a point i or a camera center i.

III. Computation for big dimensions

The previous section explains what is required to compute uncertainty ellipsoids, without providing a practical method. The aim of this section is to present and to justify a method that works even if it is impossible to compute (neither to stock) the pseudo-inverse of $J_f^\top(\hat{x})J_f(\hat{x})$ by SVD because the estimated vector $\hat{x}$ has a too large dimension.

A. Particular case

Assume that the constraint c is trivial: we fix 7 parameters as the pose of the camera 1 and an other parameter of a camera n. These 7 parameters are known, so one may take away them to the vectors $\bar{x}$, dx and $\hat{x}$. This constraint fixes the euclidean coordinate system and the scale factor, what greatly simplifies the presentation of the previous part.

In that case, we have only one $\bar{x}$ possible and $Ker(J_f) = 0$. Moreover, dx minimizing $dx \mapsto ||dy - J_f dx||^2$ is well defined: it is $dx = (J_f^\top J_f)^{-1} J_f^\top dy$. One deduces that $\hat{x}$ obeys a Gaussian distribution of mean $\bar{x}$ and of covariance $C_{dx} = \bar{\sigma}^2 (J_f^\top J_f)^{-1}$, and the confidence ellipsoids for points and cameras from diagonal blocks of C_{dx}.

A method [6] exists to compute only diagonal blocks of the inverse of matrix like $J_f^\top J_f$, so that the computation and the storage are possible for high dimension. The following method explain how to compute only the diagonal blocks of the pseudo-inverse of a matrix like $J_f^\top J_f$ for a minimal norm dx and for the general case of the constraint c, in the practical case of a very high dimension of x.

B. Factorization of covariance matrices

The two following lemmas lead to compute a simpler pseudo-inverse than that of $J_f^\top J_f$ (proofs in the Appendix).

Lemma 1: Let H be a symmetric positive matrix, G invertible and same dimension, and P the linear projector on $Im(H) = Ker(H)^\perp$ parallel to $Ker(H)$. So

$$H^+ = PG(G^\top HG)^+ G^\top P^\top.$$

Lemma 2: Let P_1 and P_2 be two linear projectors such as $Ker(P_1) = Ker(P_2)$. So $P_1 P_2 = P_1$.

Consequences: We apply the lemma 1 with $H = J_f^\top J_f$ and G invertible, then the lemma 2 with $P_f^\perp$ the projector on $Ker(J_f)^\perp$ parallel to $Ker(J_f)$ and P_f^c the projector on $Ker(J_c)$ parallel to $Ker(J_f)$. We obtain

$$\begin{aligned} C_{dx}^\perp &= \bar{\sigma}^2 P_f^\perp G(G^\top J_f^\top J_f G)^+ G^\top (P_f^\perp)^\top \\ C_{dx}^c &= \bar{\sigma}^2 P_f^c G(G^\top J_f^\top J_f G)^+ G^\top (P_f^c)^\top. \end{aligned}$$

The following paragraph shows that these factorizations lead to an individual computation of diagonal blocks of covariance if G is well chosen and we use the sparse structure of $J_f^\top J_f$.

C. Block computation of covariance matrix

The two following lemmas are useful to choose G and to explicit P_f^c and $P_f^\top$ (proofs in the Appendix).

Lemma 3: If $H = \begin{pmatrix} U & W \\ W^\top & V \end{pmatrix}$ with V invertible, then $G^\top HG = \begin{pmatrix} Z & 0 \\ 0 & V \end{pmatrix}$ with $G = \begin{pmatrix} I & 0 \\ -V^{-1}W^\top & I \end{pmatrix}$ and $Z = U - WV^{-1}W^\top$. Also we have the kernel relationship $Ker(H) = \begin{pmatrix} I \\ -V^{-1}W^\top \end{pmatrix} Ker(Z)$.

Lemma 4: Let K_f be a matrix whose columns define a vectorial basis of $Ker(J_f^\top J_f)$. We have

$$\begin{aligned} P_f^\perp &= I - K_f(K_f^\top K_f)^{-1} K_f^\top \\ P_f^c &= I - K_f(J_c K_f)^{-1} J_c. \end{aligned}$$

Sparse structure and consequences: One assumes that the m first parameters of x are those of cameras, and the n last ones, those of 3D points. We have the following block-structure [6]: $J_f^\top J_f = \begin{pmatrix} U & W \\ W^\top & V \end{pmatrix}$ with U a $m \times m$ matrix, W a rectangular one and V a $n \times n$ one which is diagonal by 3×3 invertible blocks. We assume that there are really more points than cameras, therefore $m \ll n$ (see Figure 3).

We consider the case where the computation of SVD is possible for a $m \times m$ matrix but not for a $(m+n) \times (m+n)$ one (due to a time or storage problem)

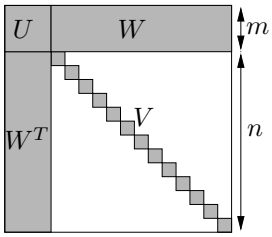


Fig. 3. Sparse structure of $J_f^\top J_f$ with $m \ll n$. Non-zeros entries in the matrix are shown in gray. All the diagonal blocks of V are 3×3.

Using the Lemma 3 on $H = J_f^\top J_f$ involves

$$\begin{aligned} C_{dx}^{\perp} &= \tilde{\sigma}^2 P_f^{\perp} G \begin{pmatrix} Z^+ & 0 \\ 0 & V^{-1} \end{pmatrix} G^\top (P_f^{\perp})^\top \\ C_{dx}^{c} &= \tilde{\sigma}^2 P_f^{c} G \begin{pmatrix} Z^+ & 0 \\ 0 & V^{-1} \end{pmatrix} G^\top (P_f^{c})^\top \end{aligned}$$

with $Z = U - WV^{-1}W^\top$ and $G = \begin{pmatrix} I & 0 \\ -V^{-1}W^\top & I \end{pmatrix}$. We obtain Z^+ and a basis of $Ker(Z)$ thanks to a SVD on Z which is a "small" $m \times m$ matrix. The inverse of V is the diagonal matrix of inverted 3×3 diagonal blocks of V. We also deduce from this lemma a basis of $Ker(H) = Ker(J_f)$ that we store in a matrix K_f, thanks to the relation between $Ker(H)$ and $Ker(Z)$.

According to the lemma 4, $C_{dx}^{\perp}$ and C_{dx}^{c} may be written as:

$$\begin{aligned} C &= (I - AB^\top)\tilde{H}(I - BA^\top) \\ &= \tilde{H} - A(\tilde{H}B)^\top - (\tilde{H}B)A^\top + A(B^\top \tilde{H} B)A^\top \end{aligned}$$

with $Y = WV^{-1}$, $\tilde{H}/\tilde{\sigma}^2 =$

$$G \begin{pmatrix} Z^+ & 0 \\ 0 & V^{-1} \end{pmatrix} G^\top = \begin{pmatrix} Z^+ & -Z^+Y \\ -Y^\top Z^+ & Y^\top Z^+ Y + V^{-1} \end{pmatrix},$$

and with A and B two known $(m+n) \times 7$ matrices.

Computation of diagonal blocks: A diagonal block of $\tilde{H}$ is obtained like this [6]: we pick up it in Z^+ for a camera, or we compute it with $Y^\top Z^+ Y + V^{-1}$ for a point. Then we compute $\tilde{H}B = \tilde{H}\begin{pmatrix} B_c^\top & B_p^\top \end{pmatrix}^\top$ without fully computing either storing $\tilde{H}$ thanks to the expression

$$\frac{1}{\tilde{\sigma}^2}\tilde{H}B = \begin{pmatrix} 0 \\ V^{-1}B_p \end{pmatrix} + \begin{pmatrix} I \\ -Y^\top \end{pmatrix} (Z^+ \begin{pmatrix} B_c & -YB_p \end{pmatrix}).$$

So it is easy to get the diagonal blocks of C (see Figure 4).

D. Time and storage complexities

We evaluate the time and storage complexities of the computation of 6×6 and 3×3 diagonal-blocks of the covariance matrix. These blocks correspond to the parameters of each camera and each 3D point.

Notations: Let c, p and i be the number of cameras, of 3D points and of 2D reprojections, respectively. We assume that a 3D points has been reconstructed from a maximum of r images. We note that

- $c \ll p < i < pr \leq pc$
- U and Z are $\Theta(c) \times \Theta(c)$ matrices.

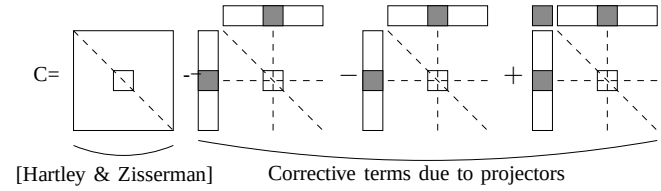


Fig. 4. The aim is to compute a diagonal block of the covariance matrix C from the expression $C = H - A(\tilde{H}B)^\top - (\tilde{H}B)A^\top + A(B^\top \tilde{H} B)A^\top$. We get it from the computation of a block from $\tilde{H}$ (on the left) according to the method described in [6], and from a corrective term (on the right) due to the projectors P_f^c or $P_f^\perp$ where only the gray part is useful during the computation.

- V is a $\Theta(p) \times \Theta(p)$ diagonal matrix of 3×3 invertible blocks.
- W and Y are $\Theta(c) \times \Theta(p)$ matrices and contain i non-zero 6×3 blocks at the same places.
- A, B and $\tilde{H}B$ are $\Theta(p) \times 7$ matrices.

Calculations: We compute successively U, W, V, V^{-1}, Y, Z then $Z^+, Ker(H), A, B, \tilde{H}B$, and at last the diagonal-blocks of the covariance matrix without computing neither storing the zero blocks of W, Y and V. First of all we can note that complexity in term of available space is the one of theses matrices, that is to say $\Theta(c^2 + i)$. Now we present the time complexity.

The time complexity of U, W, V, V^{-1}, Y is $\Theta(i)$, that of Z is $\Theta(ic)$ (each non-zero 6×3 block of Y is used one time in the computation of each of c columns of blocks of Z), that of Z^+ is $\Theta(c^3)$ [6], that of $Ker(H)$ is $\Theta(i)$, those of A and B are $\Theta(p)$, and that of $\tilde{H}B$ are $\Theta(i)$. For all computations above, the time complexity is $\Theta(ic + c^3)$.

As the individual computation of one coefficient of $Y^\top Z^+ Y$ is in $O(r^2)$, once Z^+ and Y are computed, we deduce that the additional time complexity for the computation of all diagonal blocks of $\tilde{H}$ is in $O(pr^2)$. The final time complexity is $O(ic + pr^2 + c^3)$.

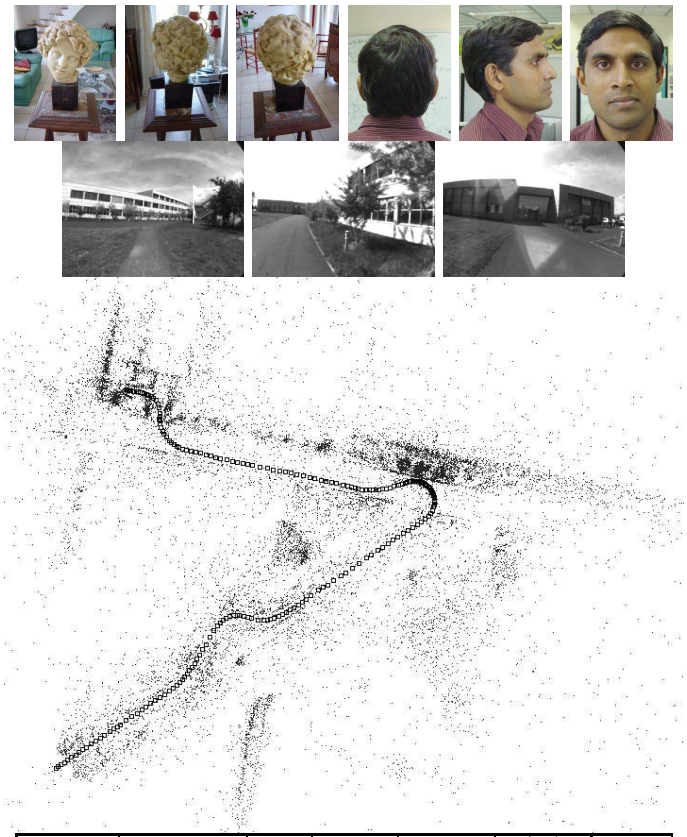
Discussions and comparisons: The time complexity $O(pc^2)$ is also correct, but cruder because we often have $r \ll c$ in practice.

Even if we use the sparse structure of H and the compact expressions (A, B) of projectors, the complexities would increase if we start to compute and store the whole $\tilde{H}$, and then applying projectors: the space complexity would be $\Theta(p^2)$, and the time one $\Theta(ip + c^3)$ because the computation of product $Y^\top(Z^+Y)$ is in $\Theta(ip)$ once Y and Z^+Y are computed. We find again here the complexity $O(cp^2 + c^3)$ of [10] with $i < pc$.

Without using the sparse structure of H, we would compute the pseudo-inverse of H by SVD, and then apply the projectors. In this case the time complexity would be $\Theta(p^3)$.

IV. Experimentation

The images sequences we used are described in Figure 5. First two non-trivial gauge constraints are presented, which will be experimented latter.



	$x \times y$	c	p	i	$\lvert H \rvert$	t
Bust1	480×640	26	885	3129	2811	2 s
Bust2	480×640	26	26497	80764	79647	45 s
Man	480×640	38	53406	159174	160446	86 s
Campus	512×384	198	22726	103607	69366	448 s

Fig. 5. **Top**: several images of our 3 sequences. **Middle**: top view of the reconstructed Campus sequence (a camera is represented by a square and a 3D point by a black point). **Bottom**: informations about the 3D reconstructions (Bust1 and Bust2 are two reconstructions of the bust sequence). We denote $x \times y$, c, p, i and t the dimension of images, the number of cameras, of 3D points, of matches that satisfies the geometry and the computation time (with a Pentium 4 2,4 GHz and 1024 Mb of RAM), respectively. So, the dimension $|H|$ of the Hessian is $6c + 3p$. The 3D reconstructions of Bust1, Bust2 and Man have been computed with the methods described in [8], and the Campus with the method explained in [14].

A. Two non-trivial gauge constraints

These constraints are referenced by [11]. Let A_i be the 3 parameters X_i, Y_i, Z_i of a 3D point or a camera center. Given the estimation A_i^0 provided by the bundle adjustment, we define the following constraints on the parameters A_i:

$$\sum A_i = 0 \tag{1}$$

$$\sum ||A_i||^2 = constant \tag{2}$$

$$\sum A_i^0 \wedge A_i = 0. \tag{3}$$

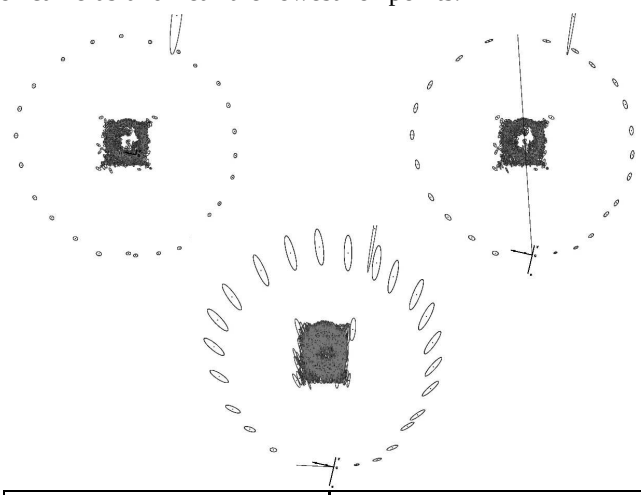
A camera-based (respectively, a point-based) gauge constraint is obtained if we chose A_i as the center of the ith camera (respectively, the ith 3D point). Equation (1) fixes the translation of the coordinate center, equation (2) fixes the scale factor, and equation (3) fixes the rotation.

We note that those constraints take advantage of being more symmetrical according to the parameters than the trivial ones, that make chosen points or cameras play a particular role. So we call a such gauge constraint "non-trivial symmetric gauge".

B. Gauge comparisons

Here we compare different gauges for a medium dimensional 3D reconstruction: Bust1. For these sequences, the "diameter" of cameras center is 1 (the camera centers are almost on a circle around the bust).

Camera-based gauge: The uncertainty ellipsoids of several camera-based gauge are presented in Figure 6 for Bust1 sequence. We note that the uncertainties decrease when the trivial constraints are applied on far cameras. On the other hand, the non-trivial symmetric constraint on camera centers gives the lowest uncertainties. Also it gives the better spread for cameras and near the lowest for points.

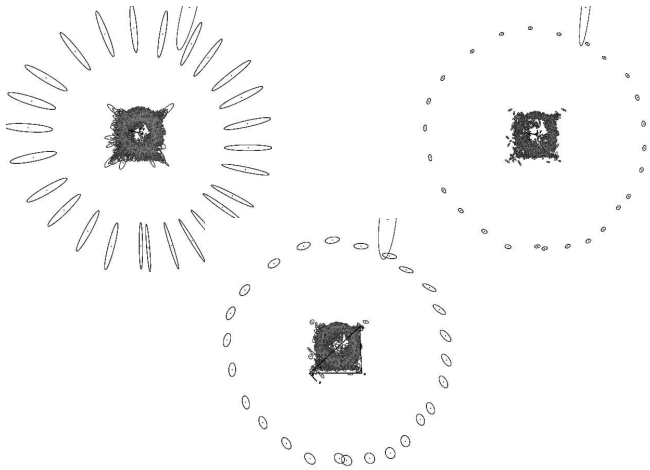


cameras			points		
1th quartile	mean	3th quartile	1th quartile	median	3th quartile
1.73e-3	1.99e-3	2.19e-3	1.58e-3	2.01e-3	2.83e-3
2.78e-3	3.31e-3	3.98e-3	1.63e-3	2.02e-3	2.81e-3
4.72e-3	9.86e-3	1.48e-2	8.89e-3	9.67e-3	1.03e-2

Fig. 6. 90% ellipsoids (5 times zoomed) for Bust1 and table of their big axis, for cameras based gauge. The large ellipsoid out of the "camera centers circle" corresponds to a 3D points of the remote background. **Top left**: symmetric non-trivial gauge, **Top right**: (respectively, **Bottom**) 2 distant cameras (respectively close cameras) fixed: the first one is entirely fixed and the second one is partial fixed by its center (the 2 picked out cameras are linked by a line).

Point-based gauge: The uncertainty ellipsoids of several point-based gauges are presented on Figure 7 for Bust1 sequence. Let A be the set of 3D points, on which we apply the symmetric non trivial constraint. We note that for a symmetric non trivial constraint, the uncertainty is spread on A. So, uncertainty on few background 3D points (we know it is large, and we do not see all of them in the figures) affects the bust points when A is the whole set of 3D points, and we obtain uncertainty for 3D points of bust around 3 or 4 times larger (top left) than in the case where A is reduced to the set of bust points (top right). Moreover this large uncertainty affecting the 3D points of the bust affects cameras too (the size of ellipsoid

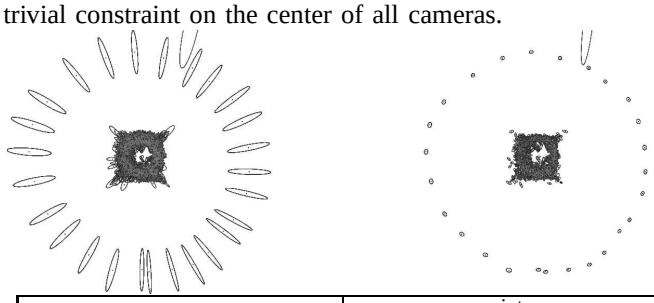
is multiplied by around 10), which observe the 3D points that belong almost to the bust. At last we note that ellipsoids are always larger for a trivial gauge on three remote points of the bust (on the bottom), than for a symmetric gauge with A reduced to the set of bust points.



cameras			points		
1th quartile	mean	3th quartile	1th quartile	median	3th quartile
2.36e-2	2.41e-2	2.44e-2	6.13e-3	7.08e-3	7.55e-3
2.08e-3	2.39e-3	2.58e-3	1.42e-3	1.87e-3	2.71e-3
6.13e-3	6.88e-3	7.21e-3	2.92e-3	3.31e-3	3.68e-2

Fig. 7. 90% ellipsoids (5 times zoomed) for Bust1 and table of their big axis, for points based gauge. **Top left**: symmetric non trivial for all points gauge, **Top right**: symmetric non trivial without remote background points gauge, **Bottom**: three fixed points (two entirely and one partially).

Particular gauges: The figure 8 shows the uncertainty ellipsoids when the chosen covariance matrix is the pseudo-inverse H^+ of the Hessian H (left), and $\tilde{H}$ (right). For this sequence and coordinate system, we observe that the ellipsoids of H^+ are similar to those obtained with the symmetric non-trivial constraint on all 3D points, and that those of $\tilde{H}$ are almost the same as those obtained with the symmetric non trivial constraint on the center of all cameras.



cameras			points		
1th quartile	mean	3th quartile	1th quartile	median	3th quartile
2.12e-2	2.16e-2	2.19e-2	5.39e-3	6.36e-3	6.91e-3
1.74e-3	2.00e-3	2.20e-3	1.56e-3	2.00e-3	2.81e-3

Fig. 8. 90% ellipsoids for Bust1 (5 times zoomed) and table of their big axis, for particular gauge. **Left**: pseudo-inverse, **Right**: $\tilde{H}$.

C. Two kind of reconstructions

We give here the ellipsoids for a 3D reconstruction Bust2 obtained from the same images sequence as Bust1 but computed with other 3D points: Bust1 is computed with 885 points of interest and Bust2 is computed from 26497 points located at corners of a regular grid in one image [8]. The ellipsoids are given in the figure 9 for the symmetric non trivial camera and 3D points without background based constraint, and we compare them with those we obtain for the same constraints with Bust1. We note that cameras uncertainty clearly decrease (near a factor 4) when the number of 3D points is larger. This result is intuitive. However, the magnitude of 3D points uncertainties increase by a factor 1.08 to 1.3. This is probably due to the fact that matching positions for regular grid points are less accurate than those of interest points detected in images.

cameras			points		
1th quartile	mean	3th quartile	1th quartile	median	3th quartile
4.60e-4	5.25e-4	5.61e-4	1.83e-3	2.37e-3	2.95e-3
5.12e-4	6.15e-4	6.77e-4	1.82e-3	2.37e-3	2.94e-3

Fig. 9. Table for the big axis of the 90% ellipsoids for Bust2. **Top**: symmetric non trivial based camera gauge. **Bottom**: symmetric non trivial based points gauge, without background remote points.

D. Complex 3D reconstructions

At last the figure 10 gives the ellipsoids for Man and Campus reconstructions that have high dimensions (in the sense of the parameter number), with the symmetric non-trivial camera-based constraint. The "diameter" of camera centers is set to 1 for Man, and the path of Campus measures 147 meters. The important number of remote points (that have high uncertainty) for Campus gives the ellipsoids schema illegible. So, only cameras ellipsoids are shown. It is clear that for such dimensions, the computation method and the memory management is important. For example, full storing W and Y for Campus should fill $2*198*22726*6*3*8$ bytes (8 bytes for a "double" real), that is to say 1236 Mb.

V. Conclusion

After a survey about covariance matrices and uncertainty ellipsoids within the framework of 3D optimal reconstruction, a method to compute them has been proposed, justified and experimented on complex real world examples, for several choices of constraints. The ellipsoids will be usefull for object reconstruction [8] and localization [14] applications.

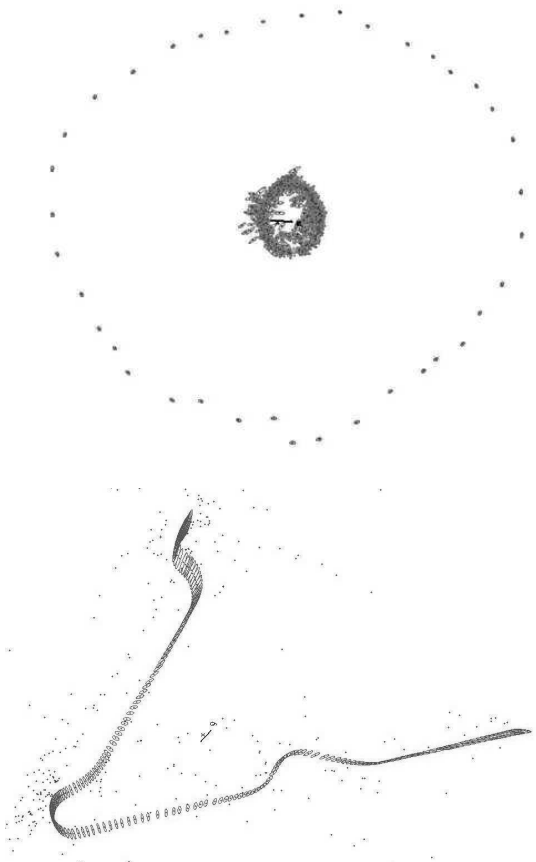
VI. Appendix

A. Pseudo-inverse properties

If A is a $m \times n$ matrix with $n \leq m$, we could compute its pseudo-inverse A^+ with SVD [13]: if $A = UDV^\top$ with U orthogonal, V orthogonal and square matrix, and D a positive diagonal matrix, then $A^+ = VD^+U^\top$ with D^+ the diagonal matrix getting by inverting non zero entries of D.

If A is symmetric positive, then we could take $U = V$ and take out the null diagonal entries of D and the corresponding column vectors of U, to get $A = UDU^\top, A^+ = UD^{-1}U^\top$.

Let y be a vector. The set of x that minimize $||Ax - y||^2$ is $(A^\top A)^+ A^\top y + Ker(A)$, denoting $x + E = \{x + y, y \in E\}$ with E a vector space and x a vector belong to a larger space.



cameras			points		
1th quartile	mean	3th quartile	1th quartile	median	3th quartile
8.65e-4	9.79e-4	1.06e-3	1.79e-3	2.03e-3	2.49e-3
0.156	0.220	0.272	0.348	1.059	12.43

Fig. 10. 90% ellipsoids (5 times zoomed) and table of their big axis for the non-trivial symmetric camera-based gauge. **Top:** Man. **Bottom:** Campus (values in meters).

B. Lemma proofs

Lemma 1 proof: Since H is symmetric (real) and positive, it is orthogonally similar to a diagonal positive matrix. We deduce the existence of a rectangular matrix Q such as $Q^\top Q = I$ and D diagonal strictly positive such as $H = QD^2Q^\top$. One checks that $H^+ = QD^{-2}Q^\top$ and $P = QQ^\top$. Moreover, $PG(G^\top HG)^+G^\top P^\top$ is successively equal to:

$$QQ^\top G(G^\top QD^2Q^\top G)^+G^\top QQ^\top$$
$$QD^{-1}A^\top(AA^\top)^+AD^{-1}Q^\top \text{ with } A = G^\top QD.$$

Let $A = U\tilde{D}V^\top$ be a SVD decomposition of A: we get $U^\top U = V^\top V = VV^\top = I$. Moreover, $\tilde{D}$ is invertible because $Ker(A) = Ker(QD) = D^{-1}Ker(Q) = 0$. This implies that $(AA^\top)^+ = (U\tilde{D}^2U^\top)^+ = U\tilde{D}^{-2}U^\top$. We obtain

$$A^\top(AA^\top)^+A = V\tilde{D}U^\top(U\tilde{D}^{-2}U^\top)U\tilde{D}V^\top = I,$$

and $PG(G^\top HG)^+G^\top P^\top = QD^{-1}ID^{-1}Q^\top = H^+$.

Lemma 2 proof: We get $x = (x - P_2x) + P_2x$ with $x - P_2x \in Ker(P_2) = Ker(P_1)$ and thus $P_1x = P_1P_2x$ for any x.

Lemma 3 proof: We check that

$$\begin{pmatrix} I & WV^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} Z & 0 \\ 0 & V \end{pmatrix}\begin{pmatrix} I & 0 \\ V^{-1}W^\top & I \end{pmatrix} = \begin{pmatrix} U & W \\ W^\top & V \end{pmatrix},$$

that is to say $G^{-\top}\begin{pmatrix} Z & 0 \\ 0 & V \end{pmatrix}G^{-1} = H$. This implies that

$$Ker(H) = Ker\left(\begin{pmatrix} Z & 0 \\ 0 & V \end{pmatrix}G^{-1}\right) = GKer\left(\begin{pmatrix} Z & 0 \\ 0 & V \end{pmatrix}\right)$$

and we deduce the expression of $Ker(H)$ thanks to the invertible V.

Lemma 4 proof: The width of K_f is the dimension of $Ker(J_f^\top J_f) = Ker(J_f)$, therefore 7. The constraint c is defined by 7 real independent equations, so the height of J_c is 7 too. Accordingly J_cK_f is a square matrix. Moreover $Ker(J_c)$ and $Ker(J_f) = Im(K_f)$ are supplementary and so their intersection is reduced to 0: $J_cK_fx = 0 \Rightarrow K_fx = 0$. Thus $Ker(J_cK_f) = Ker(K_f) = 0$ and J_cK_f is invertible.

Finally, we verify that $(I - K_f(J_cK_f)^{-1}J_c)x = 0$ if $x \in Ker(J_f)$ (i.e. $x = K_fy$) and $(I - K_f(J_cK_f)^{-1}J_c)x = x$ if $x \in Ker(J_c)$. This gives the P_f^c expression, and the $P_f^\perp$ expression with a similar proof.

References

[1] "Boujou," *2d3 Ltd*, http://www.2d3.com, 2000.

[2] D.C. Brown. The Bundle Adjustment–Progress and Prospects, *International Archives of Photogrammetry*, vol. 21, 1976.

[3] G. Csurka, C. Zeller, Z. Zhang and O.D. Faugeras. Characterizing the Uncertainty of the Fundamental Matrix, *Computer Vision and Image Understanding*, 68(1):18-36, 1997.

[4] O.D. Faugeras. *Three-Dimensional Computer Vision, A Geometric Viewpoint*, The MIT Press, 1993.

[5] O.D. Faugeras and Q.T. Luong, The Geometry of Multiple Images, *MIT Press*, 2001.

[6] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*, CU Press, 2000.

[7] V. Lepetit. Gestion des occultations en réalité augmentée, Thèse d'université, Universite Henri Pointcaré, Nancy, 2001.

[8] M. Lhuillier and Long Quan. A Quasi-Dense Approach to Surface Reconstruction from Uncalibrated Images, *IEEE TPAMI*, 27(3):418-433, 2005.

[9] P.F. McLauchlan. Gauge Independence in Optimization Algorithms for 3D Vision, *Vision Algorithms Workshop*, 2000.

[10] D. Morris. Gauge Freedoms and Uncertainty Modeling for 3D Computer Vision, PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburg, PA, 2001.

[11] D. Nister. Automatic dense reconstruction from uncalibrated video sequences, PhD Thesis, Royal Institute of Technology KTH, Stockholm, Sweden, 2001.

[12] M. Pollefeys, R. Koch and L. Van Gool. Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters, *ICCV'98*.

[13] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery. *Numerical Recipes in C*, Second Edition, CU Press, 1999.

[14] E. Royer, M. Lhuillier, M. Dhome and T. Chateau. Localization in urban environments: monocular vision compared to a differential GPS sensor, *CVPR'05*.

[15] B. Triggs, P.F. McLauchlan, R. Hartley and A. Fitzgibbon. Bundle Adjustment – a modern synthesis, vol. 1883 of *Lecture Notes in Computer Science*, pp. 298-372, Springer-Verlag, 2000.

[16] Stephen Se, David Lowe and Jim Little, Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks, *IJRR*, 21(8):735–758, 2002.