Error Propagations for Local Bundle Adjustment

Alexandre Eudes[1,2] and Maxime Lhuillier[1]

[1] LASMEA-UMR 6602 Université Blaise Pascal/CNRS, 63177 Aubière Cedex, France.

[2] CEA, LIST, Laboratoire Systèmes de Vision Embarqués, B.P. 94, Gif-sur-Yvette, F-91191 France

http://maxime.lhuillier.free.fr

Abstract

Local bundle adjustment (LBA) has recently been introduced to estimate the geometry of image sequences taken by a calibrated camera. Its advantage over standard (global) bundle adjustment is a great reduction of computational complexity, which allows real-time performances with a similar accuracy. However, no confidence measure on the LBA result such as uncertainty or covariance has yet been introduced. This paper introduces statistical models and estimation methods for uncertainty with two desirable properties: (1) uncertainty propagation along the sequence and (2) real-time calculation. We also explain why this problem is more complicated than it may appear at first glance, and we provide results on video sequences.

1. Introduction

The robust and automatic estimation of camera motion and scene points from an image sequence (Structure-from-Motion or **SfM**) is still today an active field of research. A real-time method based on local bundle adjustment (**LBA**) was introduced [10] for this problem three years ago. However, no confidence measure on the result such as uncertainty or covariance was introduced before. A such measure would be useful to give a quality information if ground truth is not available, or even to merge the reconstruction with data provided by other sensors such as GPS or odometer. This Section summarizes previous works and our contribution on this topic.

Bundle Adjustment Bundle adjustment is a well known iterative method [11] designed to solve non-linear least square problems for SfM. It estimates a vector $\mathbf{x}$ which minimizes a cost function $\mathbf{x} \mapsto ||\mathbf{y} - F(\mathbf{x})||^2$. Vector $\mathbf{x}$ usually concatenates 6 parameters for each camera pose and 3 parameters for each 3D point. Vector $\mathbf{y}$ concatenates detected features in images, F concatenates projection functions, and $||.||$ is the Euclidean norm. We call this method

global bundle adjustment (**GBA**) if $\mathbf{x}$ contains all parameters of the sequence.

The time complexity of one GBA iteration is $O(c^3 + cp)$ with c and p the numbers of camera poses and 2D points, respectively. Although this complexity benefits the sparse structure of the problem and the assumption $c \ll p$, real-time GBA is impossible for long sequences.

Error Propagation Error propagation provides confidence measures for bundle adjustment result [8]. Assume that the image vector $\mathbf{y}$ follows a known Gaussian noise. Then, the function which maps $\mathbf{y}$ to minimizer $\mathbf{x}$ is approximated by its linear Taylor expansion. Now, the Gaussian noise $\mathbf{y}$ is propagated to Gaussian noise $\mathbf{x}$ such that the $\mathbf{x}$ covariance matrix can be estimated. Thanks to this matrix, uncertainty ellipsoids for a given probability can be defined for the bundle adjustment result.

The time complexity to estimate covariances of all camera poses and 3D points is at least that of one GBA iteration.

Extended Kalman Filter Bundle adjustment is based on the Levenberg-Marquardt method [8], which is an improved (damped) version of the Gauss-Newton Method. On the other hand, the Extended Kalman Filter (EKF) may be viewed as an incremental version of the Gauss-Newton Method [3] and has been applied for real-time SfM [4]. EKF is known to provide less accurate results than bundle adjustment, but it is faster. It has is own covariance management for the estimated geometry (geometry and its covariance are jointly updated over time). EKF requires prior knowledge on the covariance of all estimated parameters. A recent improvement [2] cancels this drawback, but it still depends on linearization of constraints between frames.

Local Bundle Adjustment We recently introduced LBA for real-time SfM [10]. The incremental SfM method is summarized as follows. Camera poses and a sparse cloud of 3D points are reconstructed for video frames before time t, and a new video frame should be added to the reconstruction at time t. Interest points [7, 9] of frame t are detected and matched with those of time $t - 1$. Since many points

of frame $t-1$ are reconstructed, there are 2D-3D correspondences for points in frame t and the new pose can be estimated. This pose is initialized by a robust method (such as RANSAC [6]), and LBA is applied to refine the geometry (camera poses and 3D points) of the n most recent frames $t-n+1, \cdots, t-1, t$ by maintaining consistency with previous frames $t-N+1, \cdots, t-n$ $(n < N)$. LBA is a bundle adjustment such that (1) $\mathbf{x}$ concatenates all 3D parameters of the n most recent frames and (2) $\mathbf{y}$ contains detected features in the N last frames.

The time complexity of one LBA iteration is $O(n^3 + np)$ with p the numbers of 2D points in the N last images. Thanks to fixed and small values of n and N (e.g. $n = 3, N = 10$), the GBA complexity is greatly reduced.

Our Contribution We introduce statistical models and covariance estimation methods for the SfM parameters estimated by LBA.

Section 3 describes a first method derived from the original LBA definition using a strong independence hypothesis (between the poses of previous frames $t-N+1, \cdots, t-n$ and the detected features in the N last frames). However, the original LBA is not a Maximum Likelihood Estimator (MLE) for this hypothesis. Then Section 4 describes a second method derived from a new LBA which is MLE. These two methods satisfy both desirable properties (1) uncertainty propagation along the sequence and (2) real-time performance.

Section 5 introduces covariance methods for both original and new LBAs using a weak independence hypothesis, but the real-time performance is lost. Last, we provide covariance results on video sequences and compare them with the covariance derived from GBA (Section 6).

2. Notations and Definitions

Time and Frame Index The integer t is both a time index and a frame index. The LBA-based SfM method distinguishes key-frames and non key-frames of the given video sequence to ensure a stable estimation of 3D. Since LBA is only applied on key-frames, we ignore non key-frame in this paper and use t as a key-frame number and time.

Camera Poses As mentioned above, LBA is bundle adjustment such that estimated parameters $\mathbf{x}$ are all 3D parameters of the n most recent frames and data $\mathbf{y}$ contains detected points in the N most recent frames. We assume that $n = 3$ and $N = 10$ to simplify notations without loss of generality. Notice that the camera pose of a frame is estimated many times by many LBAs since the sliding window size n is greater than one. For this reason, a double index is used for camera poses. After the LBA of time t, the poses of frames $0, 1, \cdots t$ are defined by vectors $\mathbf{c}_t^0, \mathbf{c}_t^1, \cdots \mathbf{c}_t^t$ respectively. Vector $\mathbf{c}_t^{t'}$ does not exist if $t < t'$. The LBA-based

SfM method has two steps:

- $\exists t_0 > 0$ such that $\mathbf{c}_{t_0}^0, \mathbf{c}_{t_0}^1, \cdots \mathbf{c}_{t_0}^{t_0}$ are estimated by GBA.

- $\forall t > t_0, \mathbf{c}_t^{t-2}, \mathbf{c}_t^{t-1}, \mathbf{c}_t^t$ are estimated by LBA from $\mathbf{c}_{t-1}^{t-9}, \mathbf{c}_{t-1}^{t-8}, \cdots \mathbf{c}_{t-1}^{t-3}$.

The first step is initialization and fixes the coordinate frame of the whole reconstruction. The second step is incremental and does not modify the pose of frames $t-9, t-8, \cdots t-3$. In other words, $\forall t' \leq t-3, \mathbf{c}_t^{t'} = \mathbf{c}_{t-1}^{t'}$.

Local Bundle Adjustment LBA of time t also estimates the set of 3D points $\{\mathbf{s}_t^i\}$ which have at least one detected projection in frames $t-2, t-1$ and t. Let $\mathbf{m}_{t'}^i$ be the detection of point $\mathbf{s}_t^i$ in frame t' such that $t' \in \{t-9, t-8, \cdots t\}$ (if any). Vectors $\mathbf{c}_t^{t'}, \mathbf{s}_t^i, \mathbf{m}_{t'}^i$ have dimensions 6,3 and 2, respectively. If $\mathbf{x}$ and $\mathbf{y}$ are vectors, we define $[\mathbf{xy}] = [\mathbf{x}|\mathbf{y}] = [\mathbf{x}^T \mathbf{y}^T]^T$ the vector which concatenates $\mathbf{x}$ and $\mathbf{y}$. We define

$$\mathbf{x}_t = [\mathbf{c}_t^{t-2} \mathbf{c}_t^{t-1} \mathbf{c}_t^t | \cdots \mathbf{s}_t^i \cdots] \quad (1)$$

$$\mathbf{y}_t = [\cdots \mathbf{m}_{t'}^i \cdots] \quad (2)$$

$$\mathbf{p}_t = [\mathbf{c}_t^{t-9} \mathbf{c}_t^{t-8} \cdots \mathbf{c}_t^{t-3}] \quad (3)$$

such that LBA is concisely written as

$$\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t) \text{ with} \quad (4)$$

$$f_t(\mathbf{p}_t, \mathbf{y}_t) = argmin_{\tilde{\mathbf{x}}_t} ||\mathbf{y}_t - F_t(\mathbf{p}_t, \tilde{\mathbf{x}}_t)||^2. \quad (5)$$

Function F_t concatenates projections functions of 3D points $\{\mathbf{s}_t^i\}$ onto 2D points $\{\mathbf{m}_{t'}^i\}$ up to image noise. Different notations are used here for minimizer $\mathbf{x}_t$ and variable vector $\tilde{\mathbf{x}}_t$ of function F_t. Thus, f_t minimizes $[\tilde{\mathbf{c}}_t^{t-2} \tilde{\mathbf{c}}_t^{t-1} \tilde{\mathbf{c}}_t^t | \forall \tilde{\mathbf{s}}_t^i] \mapsto$

$$\sum_{\forall i, 9 \leq t-t' \leq 3} ||\mathbf{m}_{t'}^i - \mathbf{c}_t^{t'} \tilde{\mathbf{s}}_t^i||^2 + \sum_{\forall i, 2 \leq t-t' \leq 0} ||\mathbf{m}_{t'}^i - \tilde{\mathbf{c}}_t^{t'} \tilde{\mathbf{s}}_t^i||^2 \quad (6)$$

with $\mathbf{c}_t^{t'} \tilde{\mathbf{s}}_t^i, \tilde{\mathbf{c}}_t^{t'} \tilde{\mathbf{s}}_t^i$ projections of $\tilde{\mathbf{s}}_t^i$ by poses $\mathbf{c}_t^{t'}$ and $\tilde{\mathbf{c}}_t^{t'}$.

Other Notations Here we present our last definitions. A sub-vector of vector $\mathbf{x}$ is a vector obtained by removing one or many coordinates of $\mathbf{x}$. A sub-matrix of a square matrix $\mathtt{C}$ is a matrix obtained from $\mathtt{C}$ by removing a line and a column of the same index, and this operation may be repeated for many indexes. We introduce $\mathbf{x}_t^c = [\mathbf{c}_t^{t-2} \mathbf{c}_t^{t-1} \mathbf{c}_t^t]$ as a sub-vector of $\mathbf{x}_t$. Notation $\mathbf{z} \sim \mathcal{N}(\bar{\mathbf{z}}, \mathtt{C}_\mathbf{z})$ means that $\mathbf{z}$ is a Gaussian vector of mean $\bar{\mathbf{z}}$ and covariance $\mathtt{C}_\mathbf{z}$. Let $\mathbf{z} \mapsto g(\mathbf{z})$ be a $\mathcal{C}^1$ continuous function with Jacobian $\frac{\partial g}{\partial \mathbf{z}}$. Up to the first order, we have error propagation [8]

$$g(\mathbf{z}) \sim \mathcal{N}(g(\bar{\mathbf{z}}), \mathtt{C}_g) \text{ with } \mathtt{C}_g = \frac{\partial g}{\partial \mathbf{z}}(\bar{\mathbf{z}})\mathtt{C}_\mathbf{z}(\frac{\partial g}{\partial \mathbf{z}}(\bar{\mathbf{z}}))^T. \quad (7)$$

Recurrence Relation Now, the LBA-based SfM method may be written as a recurrence relation on camera poses. We estimate $[\mathbf{p}_t \mathbf{x}_t^c]$ from $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$ as follows:

1. $\mathbf{p}_t$ is a sub-vector of $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$

2. estimate $\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t)$ using LBA (Eq. 5)

3. $\mathbf{x}_t^c$ is a sub-vector of $\mathbf{x}_t$

Step 1 is obvious since $[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c] = [\mathbf{c}_{t-1}^{t-10} \cdots \mathbf{c}_{t-1}^{t-1}]$ and $\mathbf{p}_t = [\mathbf{c}_t^{t-9} \cdots \mathbf{c}_t^{t-3}] = [\mathbf{c}_{t-1}^{t-9} \cdots \mathbf{c}_{t-1}^{t-3}]$.

3. Original LBA and its Covariance

The goal of this Section is to link the covariance of $[\mathbf{p}_t\mathbf{x}_t^c]$ to the recurrence relation of Section 2: we assume that $[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c] \sim \mathcal{N}([\bar{\mathbf{p}}_{t-1}\bar{\mathbf{x}}_{t-1}^c], C_{[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]})$ with a known covariance $C_{[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$ and we would like to estimate $C_{[\mathbf{p}_t\mathbf{x}_t^c]}$ such that $[\mathbf{p}_t\mathbf{x}_t^c] \sim \mathcal{N}([\bar{\mathbf{p}}_t\bar{\mathbf{x}}_t^c], C_{[\mathbf{p}_t\mathbf{x}_t^c]})$.

We assume that $\mathbf{y}_t \sim \mathcal{N}(\bar{\mathbf{y}}_t, \sigma^2 I)$ with noise scale $\sigma > 0$ estimated by GBA during the SfM initialization [8].

At first glance, the covariance $C_{\mathbf{x}_t}$ of the parameters $\mathbf{x}_t$ estimated by LBA may be approximated by the inverse of approximated Hessian of LBA cost function at $\mathbf{x}_t$ [8]. However, this estimation does not propagate noise from the previous parameters $\mathbf{p}_t$ to the new parameters $\mathbf{x}_t$; it only propagates $\mathbf{y}_t$ noise to $\mathbf{x}_t$ noise. The practical consequence is that the uncertainty of camera poses will not grow with time, although this is the expected result. Our propagation methods do not have this problem, but they are more complicated.

3.1. Statistical Model of $[\mathbf{p}_t\mathbf{y}_t]$

Covariance $C_{\mathbf{p}_t}$ is a sub-matrix of $C_{[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$ since $\mathbf{p}_t$ is a sub-vector of $[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$. Furthermore, we have $\mathbf{y}_t \sim \mathcal{N}(\bar{\mathbf{y}}_t, \sigma^2 I)$. We assume that Gaussian vectors $\mathbf{p}_t$ and $\mathbf{y}_t$ are independent and obtain

$$[\mathbf{p}_t\mathbf{y}_t] \sim \mathcal{N}([\bar{\mathbf{p}}_t\bar{\mathbf{y}}_t], C_{[\mathbf{p}_t\mathbf{y}_t]}), \quad C_{[\mathbf{p}_t\mathbf{y}_t]} = \begin{pmatrix} C_{\mathbf{p}_t} & 0 \\ 0 & \sigma^2 I \end{pmatrix}. \quad (8)$$

3.2. First-Order Error Propagation

We approximate f_t by its linear Taylor expansion at point $[\bar{\mathbf{p}}_t\bar{\mathbf{y}}_t]$ and obtain

$$\begin{pmatrix} \mathbf{p} \\ f(\mathbf{p},\mathbf{y}) \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{p}} \\ f(\bar{\mathbf{p}},\bar{\mathbf{y}}) \end{pmatrix} + \begin{pmatrix} I & 0 \\ \frac{\partial f}{\partial \mathbf{p}} & \frac{\partial f}{\partial \mathbf{y}} \end{pmatrix} \begin{pmatrix} \mathbf{p}-\bar{\mathbf{p}} \\ \mathbf{y}-\bar{\mathbf{y}} \end{pmatrix}. \quad (9)$$

Index t is omitted in this expression. We deduce that $[\mathbf{p}_t\mathbf{x}_t] \sim \mathcal{N}([\bar{\mathbf{p}}_t\bar{\mathbf{x}}_t], C_{[\mathbf{p}_t\mathbf{x}_t]})$ with covariance

$$C_{[\mathbf{p}_t\mathbf{x}_t]} = \begin{pmatrix} I & 0 \\ \frac{\partial f_t}{\partial \mathbf{p}_t} & \frac{\partial f_t}{\partial \mathbf{y}_t} \end{pmatrix} C_{[\mathbf{p}_t\mathbf{y}_t]} \begin{pmatrix} I & \frac{\partial f_t}{\partial \mathbf{p}_t}^T \\ 0 & \frac{\partial f_t}{\partial \mathbf{y}_t}^T \end{pmatrix}. \quad (10)$$

Proposition 1 explains how to estimate f_t derivatives.

Proposition 1: If F is $\mathcal{C}^2$ continuous with full rank Jacobian $\frac{\partial F}{\partial \mathbf{x}}$ and $f(\mathbf{p},\mathbf{y}) = argmin_{\mathbf{x}}||\mathbf{y} - F(\mathbf{p},\mathbf{x})||^2$, then

$$\frac{\partial f}{\partial \mathbf{y}} \approx (\frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}})^{-1} \frac{\partial F^T}{\partial \mathbf{x}} \quad (11)$$

$$\frac{\partial f}{\partial \mathbf{p}} \approx -(\frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}})^{-1} \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{p}} \quad (12)$$

with derivatives of F taken at point $(\mathbf{p}, f(\mathbf{p},\mathbf{y}))$.

Proof is provided in Appendix A (index t is omitted). Note that Eq. 12 is a generalization of the well known Eq. 11. The target covariance matrix $C_{[\mathbf{p}_t\mathbf{x}_t^c]}$ is a top-left sub-matrix of $C_{[\mathbf{p}_t\mathbf{x}_t]}$.

3.3. Algorithm

The LBA recurrence relation is completed with covariance. We estimate $[\mathbf{p}_t\mathbf{x}_t^c]$ and $C_{[\mathbf{p}_t\mathbf{x}_t^c]}$ from $[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$ and $C_{[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$ as follows:

1. $\mathbf{p}_t$ is a sub-vector of $[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$

2. $C_{\mathbf{p}_t}$ is a sub-matrix of $C_{[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$

3. estimate $\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t)$ using LBA (Eq. 5)

4. estimate $C_{[\mathbf{p}_t\mathbf{x}_t]}$ using Eqs. 8, 10, 11, and 12

5. $C_{[\mathbf{p}_t\mathbf{x}_t^c]}$ is a top-left sub-matrix of $C_{[\mathbf{p}_t\mathbf{x}_t]}$.

We only estimate block $C_{[\mathbf{p}_t\mathbf{x}_t^c]}$ of $C_{[\mathbf{p}_t\mathbf{x}_t]}$ since the dimension of camera vector $[\mathbf{p}_t\mathbf{x}_t^c]$ is quite smaller than that of $[\mathbf{p}_t\mathbf{x}_t]$ which also contains 3D points parameters.

Now, technical details are given for this estimation. Combining Eq. 10 and 8 we obtain

$$C_{[\mathbf{p}_t\mathbf{x}_t]} = \begin{pmatrix} C_{\mathbf{p}_t} & C_{\mathbf{p}_t}\frac{\partial f_t}{\partial \mathbf{p}_t}^T \\ \frac{\partial f_t}{\partial \mathbf{p}_t}C_{\mathbf{p}_t} & \frac{\partial f_t}{\partial \mathbf{p}_t}C_{\mathbf{p}_t}\frac{\partial f_t}{\partial \mathbf{p}_t}^T + \sigma^2 \frac{\partial f_t}{\partial \mathbf{y}_t}\frac{\partial f_t}{\partial \mathbf{y}_t}^T \end{pmatrix} \quad (13)$$

The derivatives of f_t are given by Eq. 11 and 12. They are estimated at $(\mathbf{p}_t, \mathbf{x}_t)$ with approximation $(\mathbf{p}_t, \mathbf{x}_t) \approx (\bar{\mathbf{p}}_t, \bar{\mathbf{x}}_t)$. Let $H_t = \frac{\partial F_t}{\partial \mathbf{x}_t}^T \frac{\partial F_t}{\partial \mathbf{x}_t}$ be the Gauss-Newton approximation of the Hessian of LBA cost function. Then $C_{[\mathbf{p}_t\mathbf{x}_t^c]}$ is the top-left sub-matrix of $C_{[\mathbf{p}_t\mathbf{x}_t]}$ such that

$$C_{[\mathbf{p}_t\mathbf{x}_t^c]} = \begin{pmatrix} C_{\mathbf{p}_t} & A^T \\ A & B \end{pmatrix}, \quad B \text{ is a top-left sub-matrix}$$

of $\sigma^2 H_t^{-1} + H_t^{-1}\frac{\partial F_t}{\partial \mathbf{x}_t}^T \frac{\partial F_t}{\partial \mathbf{p}_t}C_{\mathbf{p}_t}\frac{\partial F_t}{\partial \mathbf{p}_t}^T \frac{\partial F_t}{\partial \mathbf{x}_t}H_t^{-1}$

and A is a top block of $-H_t^{-1}\frac{\partial F_t}{\partial \mathbf{x}_t}^T \frac{\partial F_t}{\partial \mathbf{p}_t}C_{\mathbf{p}_t}$.

We estimate A and B with block-wise inversion [11]

$$H_t = \begin{pmatrix} U & W \\ W^T & V \end{pmatrix} \Rightarrow H_t^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & V^{-1} \end{pmatrix} + Y^T Z^{-1} Y \quad (14)$$

with $Z = U - WV^{-1}W^T$ and $Y = \begin{pmatrix} I & -WV^{-1} \end{pmatrix}$. (15)

Matrices B, U and Z have the same dimensions; V is a block diagonal matrix with 3×3 invertible blocks thanks to Eq. 6. We successively estimate Z^{-1}, $D_0 = \frac{\partial F_t}{\partial \mathbf{x}_t}^T \frac{\partial F_t}{\partial \mathbf{p}_t}$, and $D_1 = H_t^{-1}D_0$ using

$$D_1 = \begin{pmatrix} 0 & 0 \\ 0 & V^{-1} \end{pmatrix} D_0 + (Y^T(Z^{-1}(YD_0))). \quad (16)$$

At this point, the estimation of A and B is straightforward: matrix A is a top block of $-D_1 C_{\mathbf{p}_t}$ and $B - \sigma^2 Z^{-1}$ is a top-left sub-matrix of $D_1 C_{\mathbf{p}_t} D_1^T$.

4. New LBA and its Covariance

In Section 3, the covariance for original LBA [10] is estimated assuming that $\mathbf{p}_t$ and $\mathbf{y}_t$ are independent. Section 4 introduces a new LBA function f_t^n which provides the maximum likelihood estimation (**MLE**) of $\mathbf{x}_t$ for this assumption. Section 4 also describes a recurrence relation to estimate geometry and its covariance: $[\mathbf{p}_t \mathbf{x}_t^c]$ and $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$ are estimated from $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$ and $C_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$.

4.1. Maximum Likelihood Estimator

Index t is omitted for $\mathbf{p}_t, \tilde{\mathbf{p}}_t, \mathbf{x}_t, \tilde{\mathbf{x}}_t, \mathbf{y}_t, F_t, f_t$ in this part. Assume that $\mathbf{p} \sim \mathcal{N}(\bar{\mathbf{p}}, C_{\mathbf{p}})$ and $\mathbf{y} \sim \mathcal{N}(\bar{\mathbf{y}}, \sigma^2 I)$ are independent with $\bar{\mathbf{y}} = F(\bar{\mathbf{p}}, \bar{\mathbf{x}})$. The unknown parameters of the statistical model are $\bar{\mathbf{p}}$ and $\bar{\mathbf{x}}$ (we assume that σ^2 and $C_{\mathbf{p}}$ are given as true values). The probability density function of Gaussian vector $[\mathbf{p}\mathbf{y}]$ is

$$d(\mathbf{p}, \mathbf{y}|\bar{\mathbf{p}}, \bar{\mathbf{x}}) = K e^{\frac{-1}{2}(\frac{1}{\sigma^2}||\mathbf{y} - F(\bar{\mathbf{p}}, \bar{\mathbf{x}})||^2 + (\mathbf{p} - \bar{\mathbf{p}})^T C_{\mathbf{p}}^{-1} (\mathbf{p} - \bar{\mathbf{p}}))} \quad (17)$$

with K a constant. Thus, the MLE $f_t^n(\mathbf{p}, \mathbf{y})$ of $[\bar{\mathbf{p}}\bar{\mathbf{x}}]$ is

$$argmin_{[\tilde{\mathbf{p}}\tilde{\mathbf{x}}]} \frac{1}{\sigma^2}||\mathbf{y} - F(\tilde{\mathbf{p}}, \tilde{\mathbf{x}})||^2 + (\mathbf{p} - \tilde{\mathbf{p}})^T C_{\mathbf{p}}^{-1}(\mathbf{p} - \tilde{\mathbf{p}}). \quad (18)$$

Function f_t^n defines estimations $\mathbf{p}_t^n$ and $\mathbf{x}_t^n$ of $\bar{\mathbf{p}}_t$ and $\bar{\mathbf{x}}_t$ (note that f_t does not provide an estimation of $\bar{\mathbf{p}}_t$).

4.2. First-Order Error Propagation

Function $f_t^n(\mathbf{p}_t, \mathbf{y}_t)$ in Eq. 18 may be rewritten as

$$f_t^n(\mathbf{y}_t^n) = argmin_{\tilde{\mathbf{x}}_t^n}||\mathbf{y}_t^n - F_t^n(\tilde{\mathbf{x}}_t^n)||^2 \quad (19)$$

with $\tilde{\mathbf{x}}_t^n = [\tilde{\mathbf{p}}_t|\tilde{\mathbf{x}}_t], \quad \mathbf{y}_t^n = [(C_{\mathbf{p}_t})^{-\frac{1}{2}}\mathbf{p}_t|\frac{1}{\sigma}\mathbf{y}_t] \quad (20)$

$$\text{and } F_t^n(\tilde{\mathbf{x}}_t^n) = [(C_{\mathbf{p}_t})^{-\frac{1}{2}}\tilde{\mathbf{p}}_t|\frac{1}{\sigma}F_t(\tilde{\mathbf{p}}_t, \tilde{\mathbf{x}}_t)]. \quad (21)$$

Now we see that the Jacobian $\frac{\partial f_t^n}{\partial \mathbf{y}_t^n}$ of f_t^n can be estimated using Eq. 11 of Proposition 1:

$$\frac{\partial f_t^n}{\partial \mathbf{y}_t^n} = (\frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n}^T \frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n})^{-1} \frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n}^T. \quad (22)$$

Furthermore, Eq. 20 implies $\mathbf{y}_t^n \sim \mathcal{N}(\bar{\mathbf{y}}_t^n, I)$. Eq. 22 and the linear Taylor expansion of f_t^n at $[\bar{\mathbf{p}}_t \bar{\mathbf{y}}_t]$ provides

$$C_{[\mathbf{p}_t^n \mathbf{x}_t^n]} = C_{f_t^n} = \frac{\partial f_t^n}{\partial \mathbf{y}_t^n} I \frac{\partial f_t^n}{\partial \mathbf{y}_t^n}^T = (\frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n}^T \frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n})^{-1}. \quad (23)$$

4.3. Algorithm

We may estimate $[\mathbf{p}_t \mathbf{x}_t^c]$ and $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$ from $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$ and $C_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$ as follows:

1. $\mathbf{p}_t$ is a sub-vector of $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$

2. $C_{\mathbf{p}_t}$ is a sub-matrix of $C_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$

3. estimate $[\mathbf{p}_t^n \mathbf{x}_t^n] = f_t^n(\mathbf{p}_t, \mathbf{y}_t)$ using LBA (Eq. 18)

4. estimate $C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$ using Eqs. 23 and 21

5. do $\mathbf{p}_t \leftarrow \mathbf{p}_t^n, \mathbf{x}_t \leftarrow \mathbf{x}_t^n$ and $C_{[\mathbf{p}_t \mathbf{x}_t]} \leftarrow C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$

6. $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$ is a top-left sub-matrix of $C_{[\mathbf{p}_t \mathbf{x}_t]}$.

The new LBA (Eq. 18) is more time consuming than the original LBA (Eq. 5) since camera poses of $\mathbf{p}_t^n$ should be estimated with those of $\mathbf{x}_t^n$. So we replace step 3 by $\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t)$ and remove step 5. Step 4 is unchanged: Eqs. 23 and 21 are still used with $\mathbf{p}_t^n = \mathbf{p}_t, \mathbf{x}_t^n = \mathbf{x}_t, C_{[\mathbf{p}_t \mathbf{x}_t]} = C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$. In other words, we approximate the result of the new LBA by the result of the original LBA (both results are exactly the same if there is no image noise). We only estimate block $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$ of $C_{[\mathbf{p}_t \mathbf{x}_t]}$ since the dimension of camera vector $[\mathbf{p}_t \mathbf{x}_t^c]$ is quite smaller than that of $[\mathbf{p}_t \mathbf{x}_t]$ which also contains 3D points parameters.

Now, technical details are given for this estimation. Eqs. 21 and 23 provide (t and $\tilde{}$ omitted)

$$\frac{\partial F^n}{\partial \mathbf{x}^n} = \begin{pmatrix} \frac{\partial F^n}{\partial \mathbf{p}} & \frac{\partial F^n}{\partial \mathbf{x}} \end{pmatrix} = \begin{pmatrix} C_{\mathbf{p}}^{-\frac{1}{2}} & 0 \\ \frac{1}{\sigma}\frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma}\frac{\partial F}{\partial \mathbf{x}} \end{pmatrix} \quad (24)$$

$$C_{[\mathbf{p}_t^n \mathbf{x}_t^n]} = \begin{pmatrix} C_{\mathbf{p}}^{-1} + \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{p}}^T\frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{p}}^T\frac{\partial F}{\partial \mathbf{x}} \\ \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{x}}^T\frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{x}}^T\frac{\partial F}{\partial \mathbf{x}} \end{pmatrix}^{-1} \quad (25)$$

We apply block-wise inversion Eqs. 14 and 15 to $C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$ with $\mathbf{x} = \begin{pmatrix} \mathbf{x}^c & \mathbf{x}^s \end{pmatrix}$,

$$U = \begin{pmatrix} C_{\mathbf{p}}^{-1} + \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{p}}^T\frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{p}}^T\frac{\partial F}{\partial \mathbf{x}^c} \\ \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{x}^c}^T\frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{x}^c}^T\frac{\partial F}{\partial \mathbf{x}^c} \end{pmatrix} \quad (26)$$

$$W = \frac{1}{\sigma^2}\begin{pmatrix} \frac{\partial F}{\partial \mathbf{p}} & \frac{\partial F}{\partial \mathbf{x}^c} \end{pmatrix}^T \frac{\partial F}{\partial \mathbf{x}^s}, \quad V = \frac{1}{\sigma^2}\frac{\partial F}{\partial \mathbf{x}^s}^T \frac{\partial F}{\partial \mathbf{x}^s} \quad (27)$$

and obtain $C_{[\mathbf{p}_t \mathbf{x}_t^c]} = (U - WV^{-1}W^T)^{-1}$. Thank to Eq. 6, V is a 3×3-block diagonal matrix and is easy to inverse.

5. Weak Hypothesis

Previous Sections 3 and 4 require that $\mathbf{p}_t \sim \mathcal{N}(\bar{\mathbf{p}}_t, C_{\mathbf{p}_t})$ and $\mathbf{y}_t \sim \mathcal{N}(\bar{\mathbf{y}}_t, \sigma^2 I)$ are independent. However, camera pose $\mathbf{c}_t^{t-3} = \mathbf{c}_{t-1}^{t-3}$ is a sub-vector of both $\mathbf{p}_t$ and $\mathbf{x}_{t-1}$ such that $\mathbf{x}_{t-1} = f_{t-1}(\mathbf{p}_{t-1}, \mathbf{y}_{t-1})$. Since $\mathbf{y}_t$ and $\mathbf{y}_{t-1}$ have common 2D points in frames $t-9, t-8, \cdots t-1$, we can not assert that $\mathbf{p}_t$ and $\mathbf{y}_t$ are independent.

In Section 5, recurrence relations are introduced to estimate covariances of original and new LBAs without this hypothesis: $[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]$ and $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]}$ are estimated from $[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$ and $\mathsf{C}_{[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$.

5.1. Statistical model of $[\mathbf{y}_t\mathbf{p}_t]$

New notations are needed here. Let $\mathbf{y}_t \cap \mathbf{y}_{t-1}$ (respectively, $\mathbf{y}_t \setminus \mathbf{y}_{t-1}$) be the sub-vector of $\mathbf{y}_t$ with 2D points which are (respectively, which are not) in $\mathbf{y}_{t-1}$. Covariance $\mathsf{C}_{[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$ is known, and covariance $\mathsf{C}_{[\mathbf{y}_t\cap\mathbf{y}_{t-1}|\mathbf{p}_t]}$ is a sub-matrix of $\mathsf{C}_{[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$ since $\mathbf{y}_t\cap\mathbf{y}_{t-1}$ is a sub-vector of $\mathbf{y}_{t-1}$ and $\mathbf{p}_t$ is a sub-vector of $[\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$.

We assume that $\mathbf{y}_t \setminus \mathbf{y}_{t-1}$ and $[\mathbf{y}_t \cap \mathbf{y}_{t-1}|\mathbf{p}_t]$ are independent. Thus $[\mathbf{y}_t\mathbf{p}_t]$ is Gaussian vector with covariance

$$\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t]} = \mathsf{C}_{[\mathbf{y}_t\setminus\mathbf{y}_{t-1}|\mathbf{y}_t\cap\mathbf{y}_{t-1}|\mathbf{p}_t]} = \begin{pmatrix} \sigma^2\mathtt{I} & 0 \\ 0 & \mathsf{C}_{[\mathbf{y}_t\cap\mathbf{y}_{t-1}|\mathbf{p}_t]} \end{pmatrix}. \quad (28)$$

5.2. First-Order Error Propagation (Original LBA)

We approximate f_t by its linear Taylor expansion at point $[\bar{\mathbf{y}}_t\bar{\mathbf{p}}_t]$ and obtain

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{p} \\ f(\mathbf{p},\mathbf{y}) \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{y}} \\ \bar{\mathbf{p}} \\ f(\bar{\mathbf{p}},\bar{\mathbf{y}}) \end{pmatrix} + \begin{pmatrix} \mathtt{I} & 0 \\ 0 & \mathtt{I} \\ \frac{\partial f}{\partial \mathbf{y}} & \frac{\partial f}{\partial \mathbf{p}} \end{pmatrix}\begin{pmatrix} \mathbf{y}-\bar{\mathbf{y}} \\ \mathbf{p}-\bar{\mathbf{p}} \end{pmatrix}. \quad (29)$$

Index t is omitted in this expression. We deduce that $[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t] \sim \mathcal{N}([\bar{\mathbf{y}}_t\bar{\mathbf{p}}_t\bar{\mathbf{x}}_t], \mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t]})$ with covariance

$$\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t]} = \begin{pmatrix} \mathtt{I} & 0 \\ 0 & \mathtt{I} \\ \frac{\partial f_t}{\partial \mathbf{y}_t} & \frac{\partial f_t}{\partial \mathbf{p}_t} \end{pmatrix} \mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t]} \begin{pmatrix} \mathtt{I} & 0 & \frac{\partial f_t}{\partial \mathbf{y}_t}^T \\ 0 & \mathtt{I} & \frac{\partial f_t}{\partial \mathbf{p}_t}^T \end{pmatrix}. \quad (30)$$

Derivatives of f_t are provided by Proposition 1.

5.3. Algorithm (Original LBA)

Sections 5.2 and 5.1 define the recurrence relation for the original LBA. We estimate $[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]$ and $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]}$ from $[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$ and $\mathsf{C}_{[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$ as follows:

1. $\mathbf{p}_t$ is a sub-vector of $[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$

2. $\mathsf{C}_{[\mathbf{y}_t\cap\mathbf{y}_{t-1}|\mathbf{p}_t]}$ is a sub-matrix of $\mathsf{C}_{[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$

3. estimate $\mathbf{x}_t = f_t(\mathbf{p}_t,\mathbf{y}_t)$ using LBA (Eq. 5)

4. estimate $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t]}$ using Eqs. 28, 30, 11 and 12

5. $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]}$ is a top-left sub-matrix of $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t]}$.

Now the full estimate of $\frac{\partial f_t}{\partial \mathbf{y}_t}$ by Eq. 11 is required for the top-left sub-matrix $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]}$ of $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t]}$ and the method is intractable for real-time.

5.4. Maximum Likelihood Estimator (New LBA)

Index t is omitted for $\mathbf{p}_t, \tilde{\mathbf{p}}_t, \mathbf{x}_t, \tilde{\mathbf{x}}_t, \mathbf{y}_t, F_t, f_t$ in this part. Assume that $[\mathbf{y}\mathbf{p}] \sim \mathcal{N}([\bar{\mathbf{y}}\bar{\mathbf{p}}], \mathsf{C}_{[\mathbf{y}\mathbf{p}]})$ with $\mathsf{C}_{[\mathbf{y}\mathbf{p}]}$ defined by Eq. 28 and $\bar{\mathbf{y}} = F(\bar{\mathbf{p}},\bar{\mathbf{x}})$. The unknown parameters of the statistical model are $\bar{\mathbf{p}}$ and $\bar{\mathbf{x}}$ (we assume that $\mathsf{C}_{[\mathbf{y}\mathbf{p}]}$ is given as true value). The probability density function of the Gaussian vector $[\mathbf{p}\mathbf{y}]$ is

$$d(\mathbf{p},\mathbf{y}|\bar{\mathbf{p}},\bar{\mathbf{x}}) = Ke^{-\frac{1}{2}\bar{\mathbf{z}}^T\mathsf{C}_{[\mathbf{y}\mathbf{p}]}^{-1}\bar{\mathbf{z}}}, \quad \bar{\mathbf{z}} = \begin{pmatrix} \mathbf{y}-F(\bar{\mathbf{p}},\bar{\mathbf{x}}) \\ \mathbf{p}-\bar{\mathbf{p}} \end{pmatrix} \quad (31)$$

with K a constant. Thus, the MLE $f_t^n(\mathbf{p},\mathbf{y})$ of $[\bar{\mathbf{p}}\bar{\mathbf{x}}]$ is

$$argmin_{[\tilde{\mathbf{p}}\tilde{\mathbf{x}}]}\tilde{\mathbf{z}}^T\mathsf{C}_{[\mathbf{y}\mathbf{p}]}^{-1}\tilde{\mathbf{z}} \text{ with } \tilde{\mathbf{z}} = \begin{pmatrix} \mathbf{y}-F(\tilde{\mathbf{p}},\tilde{\mathbf{x}}) \\ \mathbf{p}-\tilde{\mathbf{p}} \end{pmatrix}. \quad (32)$$

Function f_t^n defines estimations $\mathbf{p}_t^n$ and $\mathbf{x}_t^n$ of $\bar{\mathbf{p}}_t$ and $\bar{\mathbf{x}}_t$ (note that Eq. 18 is a special case of Eq. 32).

5.5. First-Order Error Propagation (New LBA)

Function $f_t^n(\mathbf{p}_t,\mathbf{y}_t)$ in Eq. 32 may be rewritten as

$$f_t^n(\mathbf{y}_t^n) = argmin_{\tilde{\mathbf{x}}_t^n}||\mathbf{y}_t^n - F_t^n(\tilde{\mathbf{x}}_t^n)||^2 \quad (33)$$

$$\text{with } \tilde{\mathbf{x}}_t^n = [\tilde{\mathbf{p}}_t|\tilde{\mathbf{x}}_t], \quad \mathbf{y}_t^n = \mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t]}^{-\frac{1}{2}}\left(\mathbf{y}_t^T \quad \mathbf{p}_t^T\right)^T \quad (34)$$

$$\text{and } F_t^n(\tilde{\mathbf{x}}_t^n) = \mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t]}^{-\frac{1}{2}}\left(F_t(\tilde{\mathbf{p}}_t,\tilde{\mathbf{x}}_t)^T \quad \tilde{\mathbf{p}}_t^T\right)^T. \quad (35)$$

Covariance $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t^n\mathbf{x}_t^n]}$ is needed for recurrence. Eq. 34 implies $\mathbf{y}_t^n \sim \mathcal{N}(\bar{\mathbf{y}}_t^n, \mathtt{I})$, $\mathbf{y}_t = \mathtt{K}\mathbf{y}_t^n$ and $\bar{\mathbf{y}}_t = \mathtt{K}\bar{\mathbf{y}}_t^n$ with $\mathtt{K}$ a top block of $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t]}^{\frac{1}{2}}$. We approximate f_t^n by its linear Taylor expansion at point $\bar{\mathbf{y}}_t^n$ and obtain

$$\begin{pmatrix} \mathbf{y}_t \\ f_t^n(\mathbf{y}_t^n) \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{y}}_t \\ f_t^n(\bar{\mathbf{y}}_t^n) \end{pmatrix} + \begin{pmatrix} \mathtt{K} \\ \frac{\partial f_t^n}{\partial \mathbf{y}_t^n} \end{pmatrix}(\mathbf{y}_t^n - \bar{\mathbf{y}}_t^n). \quad (36)$$

Thus vector $[\mathbf{y}_t f_t^n(\mathbf{y}_t^n)]$ is Gaussian with covariance

$$\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t^n\mathbf{x}_t^n]} = \mathsf{C}_{[\mathbf{y}_t f_t^n(\mathbf{y}_t^n)]} = \begin{pmatrix} \mathtt{K} \\ \frac{\partial f_t^n}{\partial \mathbf{y}_t^n} \end{pmatrix}\begin{pmatrix} \mathtt{K} \\ \frac{\partial f_t^n}{\partial \mathbf{y}_t^n} \end{pmatrix}^T. \quad (37)$$

and $\frac{\partial f_t^n}{\partial \mathbf{y}_t^n}$ is estimated with Eqs. 11 and 35.

5.6. Algorithm (New LBA)

Sections 5.4 and 5.5 define the recurrence relation for the new LBA (Eq. 32). We estimate $[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]$ and $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]}$ from $[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$ and $\mathsf{C}_{[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$ as follows:

1. $\mathbf{p}_t$ is a sub-vector of $[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]$

2. $\mathsf{C}_{[\mathbf{y}_t\cap\mathbf{y}_{t-1}|\mathbf{p}_t]}$ is a sub-matrix of $\mathsf{C}_{[\mathbf{y}_{t-1}\mathbf{p}_{t-1}\mathbf{x}_{t-1}^c]}$

3. estimate $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t]}$ using Eq. 28

4. estimate $[\mathbf{p}_t^n\mathbf{x}_t^n] = f_t^n(\mathbf{p}_t,\mathbf{y}_t)$ using LBA (Eq. 32)

5. estimate $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t^n\mathbf{x}_t^n]}$ using Eqs. 37, 11 and 35

6. do $\mathbf{p}_t \leftarrow \mathbf{p}_t^n$, $\mathbf{x}_t \leftarrow \mathbf{x}_t^n$ and $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t]} \leftarrow \mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t^n\mathbf{x}_t^n]}$

7. $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t^c]}$ is a top-left sub-matrix of $\mathsf{C}_{[\mathbf{y}_t\mathbf{p}_t\mathbf{x}_t]}$.

6. Experiments

6.1. Integrating Covariance to LBA-based SfM

Our real-time SfM system [10] has two steps: initialization and incremental reconstruction. The former estimates the camera poses and 3D points of the sequence beginning using standard GBA. The latter incrementally reconstructs the sequence (poses and points) using original LBA (Eq. 5).

Then we integrate our covariance methods in the incremental step. These methods also require covariance for the camera poses at the sequence beginning. This covariance is estimated by the standard method derived from GBA [8]: the inverse of approximated Hessian of the minimized cost function, multiplied by image noise σ^2.

A simple gauge is chosen to estimate the covariance of poses at the beginning: we fix the first frame pose (R_0, t_0) with rotation $R_0 = I$ and location $t_0 = 0$ and the largest coordinate of the t_0-th frame location t_{t_0} with $t_{t_0}^z = 1$ ($t_0 = 9$). This information should be given since it is known that the shape of uncertainty ellipsoids derived from covariance highly depends on the gauge choice [11]. Then we remove the columns and rows corresponding to these 7 parameters in the approximated Hessian before inversion.

6.2. How to Check LBA-based Covariance ?

LBA-based SfM method produces geometry estimations which are similar to those of GBA-based SfM [10]. So we expect to obtain the same result for geometry covariance: LBA-based covariance (our methods) should be similar to GBA-based covariance (standard method). We will compare both.

The last step of "GBA-based" SfM is GBA for the complete sequence: the vector $\mathbf{x}$ of all 3D parameters minimizes the cost function $\mathbf{x} \mapsto ||\mathbf{y} - F(\mathbf{x})||^2$, where $\mathbf{y}$ is the vector of all tracked 2D points along the whole sequence and F concatenates the corresponding projection functions. If $\mathbf{y} \sim \mathcal{N}(\bar{\mathbf{y}}, \sigma^2 I)$, we have the GBA-based covariance $C_{\mathbf{x}} = \sigma^2 (\frac{\partial F}{\partial \mathbf{x}}^T \frac{\partial F}{\partial \mathbf{x}})^{-1}$ [8]. This is the inverse of the approximated Hessian with the same gauge choice as the GBA for sequence beginning (Section 6.1). The value of σ is also the same. The complete calculation of Hessian inverse is not necessary: we only calculate the diagonal blocks we need thanks to Eqs. 14 and 15 (U and V are the sub-hessians of camera poses and 3D points [8, 11], respectively).

6.3. Results

Figure 1 shows three images of the sequence taken in urban area. The camera is calibrated, is mounted on a car and is pointing forward. The trajectory length is about 400m and the sequence has 2731 512×384 images. 384 key-frames are selected from the video. 16365 points are reconstructed by original LBA from 74236 Har-

Figure 1. Three images of the sequence.

ris points [7] matched using SURF descriptor [1] in keyframes (we slightly modify the original SURF method and re-implement it on GPU using CUDA for real-time performance). The means of points in a frame and track lengths are 193 and 4.5, respectively.

Figure 2 shows quantitative comparisons of GBA-based covariance with LBA-based covariance described in Section 5.3 (original LBA with weak independence hypothesis). The same comparison is made in Figure 3 for the LBA-based covariance described in Section 4 (new LBA with strong independence hypothesis). In both cases, we study the major axis of the uncertainty ellipsoid of the location of camera $\mathbf{c}_t^{t-2}$ with probability 90% (the t-2th camera is updated at times t-2,t-1,t due to our sliding windows size of LBA, and we choose the uncertainty at the last update). The x-axis is the key-frame number; the sequence beginning optimized by GBA is not considered in these figures.

The tops of Figures 2 and 3 show the ratio of major axis lengths between LBA and GBA. We see that the ratio is acceptable (close to 1.1) for the original LBA. Unfortunately, this method is not real-time. Furthermore, the ratio of new LBA is small (close to 0.6) due to the strong independence hypothesis between $\mathbf{p}_t$ and $\mathbf{y}_t$. The bottoms of Figures 2 and 3 show the angle between LBA and GBA major axes. The angles are acceptable (small) for both LBAs.

We have also experimented the original LBA-based covariance with strong hypothesis (Section 3). In this case, the ratio of major axis lengths diverges. Therefore the strong hypothesis should not be used with the original LBA.

At this point, the covariance of new LBA with strong hypothesis is the only choice in our real-time context (although its scale is too small). Let $\bar{e}$ and σ_e be the mean and standard deviation of ratios of major axis lengths between original LBA (weak hypothesis) and new LBA (strong hypothesis) for all key-frames of the sequence. We estimate $\bar{e} = 1.82$ and $\sigma_e = 0.13$. Since $\sigma_e/\bar{e}$ is low, we decide to improve new LBA covariances by multiplying them with $\bar{e}^2$. Now, the main axis lengths of new LBA ellipsoids are roughly the same as those of original LBA. Figure 4 shows a top view of the reconstructed sequence with uncertainty ellipsoids of camera locations of our amended covariances. We see that

1. ellipsoids shapes are similar for new LBA and GBA

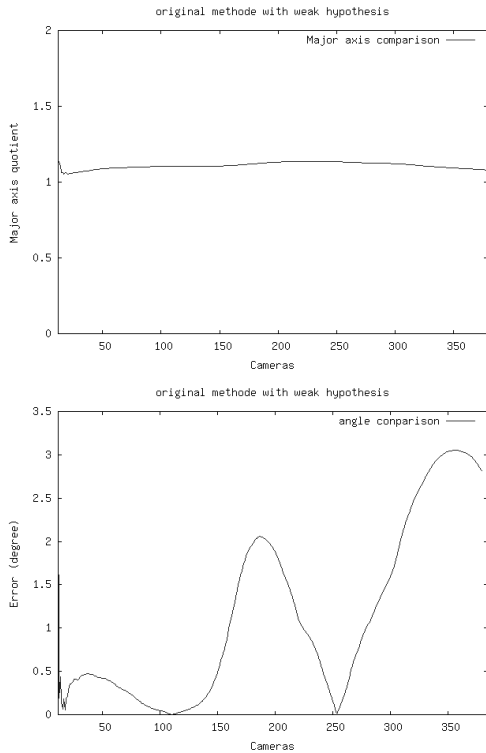2. the major axis length increases progressively with time

Figure 2. Top: ratio of major axis lengths between original LBA (weak hypothesis) and GBA ellipsoids. Bottom: angle between major axes of original LBA and GBA ellipsoids.
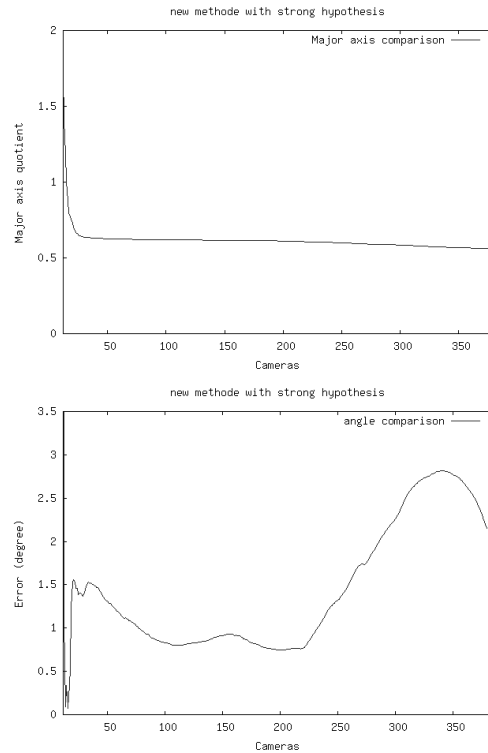

Figure 3. Top: ratio of major axis lengths between new LBA (strong hypothesis) and GBA ellipsoids. Bottom: angle between major axes of new LBA and GBA ellipsoids.

These are expected results.

Only 6.4 ms are needed by the new LBA-based covariance method for each key-frame, so our method is real-time. The total time of new LBA covariance is 2.4 s, which is (obviously) smaller than that of GBA covariance (145 s) for all camera poses.

Our experiments also includes Monte-Carlo simulations (to check our implementation of LBA covariances) and covariance estimations with other real sequences (similar results are obtained with similar values of $\bar{e}$ and σ_e).

7. Conclusion

This paper has introduced four covariance estimation methods for Structure-from-Motion (SfM) based on local bundle adjustment (LBA). They are derived from two noise hypotheses ("weak" and "strong") and two LBAs: our new LBA which provides Maximum Likelihood Estimation for these hypotheses, and the original LBA which does not. All methods propagate uncertainty along the sequence. Only two of them are real-time thanks to the strong hypothesis.

We must find a pair (statistical model, estimator) such that the estimated covariance is both physically plausible and real-time. On one side, realistic statistical models may require too much calculation to obtain real-time performance. On the other side, unrealistic statistical models may produces real-time but unrealistic uncertainty estimation. We have experimented our covariance methods on real sequences reconstructed by LBA-based SfM and have compared the results with standard covariance of global bundle adjustment. The original LBA with weak hypothesis provides acceptable covariance. This hypothesis is realistic but it does not allow real-time performance. Furthermore, the original LBA can not be used with the strong hypothesis. The new LBA with strong hypothesis provides acceptable covariance if empirical coefficient is introduced. This last method is the only choice in our real-time context.

Future work includes error propagation from key-frames to non key-frames, integration of calibration uncertainty in error propagation, experiments for many gauge choices, and fusion of our vision results with GPS or odometer data.

Appendix A

Proposition 1 (Section 3.2) is a particular case of Proposition 6.1 in [5], which also asserts that function f locally exists and is $\mathcal{C}^1$ continuous. In this appendix, the time index t is omitted. Furthermore, F_k, $\mathbf{x}_i$ and $\mathbf{p}_j$ are the k-th, i-th and j-th coordinates of vectors F, $\mathbf{x}$ and $\mathbf{p}$, respectively.

We drop $\mathbf{y}$ for Eq. 12 proof since $\mathbf{y}$ acts a constant embedded in F. Second order partial derivatives of function

$$g(\mathbf{x}, \mathbf{p}) = \frac{1}{2}||F(\mathbf{p}, \mathbf{x})||^2 \qquad (38)$$

are

$$\frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{p}_j} = \sum_k \{ \frac{\partial F_k}{\partial \mathbf{x}_i} \frac{\partial F_k}{\partial \mathbf{p}_j} + \frac{\partial^2 F_k}{\partial \mathbf{x}_i \partial \mathbf{p}_j} F_k \}. \qquad (39)$$

Gauss-Newton approximation of Eq. 39 is

$$\frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{p}_j} \approx \sum_k \frac{\partial F_k}{\partial \mathbf{x}_i} \frac{\partial F_k}{\partial \mathbf{p}_j} = (\frac{\partial F}{\partial \mathbf{x}}^T \frac{\partial F}{\partial \mathbf{p}})_{i,j}. \qquad (40)$$

A similar equation is

$$\frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{x}_j} \approx (\frac{\partial F}{\partial \mathbf{x}}^T \frac{\partial F}{\partial \mathbf{x}})_{i,j}. \qquad (41)$$

Since $f(\mathbf{p})$ is minimizer of $\mathbf{x} \to g(\mathbf{x}, \mathbf{p})$, we have

$$\forall i, \frac{\partial g}{\partial \mathbf{x}_i}(f(\mathbf{p}), \mathbf{p}) = 0 \qquad (42)$$

Thank to Eq. 42, 40 and 41, we deduce $\forall i, j$

$$0 = \frac{\partial}{\partial \mathbf{p}_j}(\mathbf{p} \to \frac{\partial g}{\partial \mathbf{x}_i}(f(\mathbf{p}), \mathbf{p})) \qquad (43)$$

$$= \frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{p}_j} + \sum_k (\frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{x}_k}) \frac{\partial f_k}{\partial \mathbf{p}_j} \qquad (44)$$

$$\approx (\frac{\partial F}{\partial \mathbf{x}}^T \frac{\partial F}{\partial \mathbf{p}})_{i,j} + \sum_k (\frac{\partial F}{\partial \mathbf{x}}^T \frac{\partial F}{\partial \mathbf{x}})_{i,k} (\frac{\partial f}{\partial \mathbf{p}})_{k,j} \qquad (45)$$

Last, Eq. 45 is equivalent to

$$(\frac{\partial F}{\partial \mathbf{x}}^T \frac{\partial F}{\partial \mathbf{x}}) \frac{\partial f}{\partial \mathbf{p}} \approx -\frac{\partial F}{\partial \mathbf{x}}^T \frac{\partial F}{\partial \mathbf{p}}. \qquad (46)$$

We obtain Eq. 12 since $\frac{\partial F}{\partial \mathbf{x}}$ has full rank. Eq. 11 is obtained here in the special case $F(\mathbf{p}, \mathbf{x}) = F(\mathbf{x}) - \mathbf{p}$ by swapping notations $\mathbf{y}$ and $\mathbf{p}$.

Vector $\mathbf{p}_t$ fixes the gauge (coordinate frame and scale) for LBA since $\mathbf{p}_t$ concatenates at least two camera poses [10]. In this context, $\frac{\partial F}{\partial \mathbf{x}}$ has full rank for general configurations of 3D points [8].
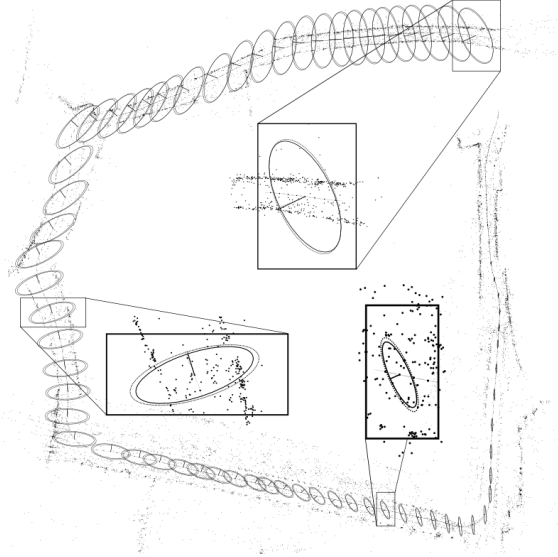
Figure 4. Top view of the reconstructed sequence by the LBA-based SfM. Camera locations and 3D reconstructed points are black dots. The 90% uncertainty ellipsoids of camera locations are also drawn for key-frames whose numbers are multiple of 5 (continuous line for standard GBA and dotted line for our new LBA with strong independence hypothesis).

References

[1] H. Bay, T. Tuytelaars, and L. Gool. Surf: Speeded up robust features. In *ECCV'06*.

[2] C. Bedder and R. Steffen. Incremental estimation without specifying a-priori covariance matrices for the novel parameters. In *VLMP Workshop'08*.

[3] D. P. Bertsekas. Incremental least squares methods and the extended kalman filter. *SIAM Journal of Optimization*, 6(3), 1996.

[4] A. J. Davison. Real-time simultaneous localization and mapping with a single camera. In *ICCV'03*.

[5] O. Faugeras, Q. Long and T. Papadopoulos. *Geometry of Multiple Images*. MIT Press, 2000.

[6] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM*, 24, 1981.

[7] C. Harris and M. Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, 1988.

[8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[9] D. G. Lowe. Distinctive image features from scales-invariant keypoints. *IJCV*, 60, 2004.

[10] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd. Real time localization and 3d reconstruction. In *CVPR'06*.

[11] B. Triggs, P. F. McLauchlan, R. Hartley, and A. W. Fitzgibbon. Bundle adjustment – a modern synthesis. In *LNCS*, 2000.