Towards Gauge Invariant Bundle Adjustment: A Solution Based on Gauge Dependent Damping

Adrien Bartoli

ICCV'03 — IN PROCEEDINGS OF THE IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION,
NICE, FRANCE, OCTOBER 2003.

Towards Gauge Invariant Bundle Adjustment: A Solution Based on Gauge Dependent Damping

Adrien Bartoli
INRIA Rhône-Alpes, 655, avenue de l'Europe
38334 Saint Ismier cedex, France. *Adrien.Bartoli@inria.fr*

Abstract

Bundle ajustment is used to obtain accurate visual reconstructions by minimizing the reprojection error. The coordinate frame ambiguity, or more generality the gauge freedoms, has been dealt with in different manners. It has often been reported that standard bundle adjustment algorithms were not gauge invariant: two iterations within different gauges can lead to geometrically very different results. Surprisingly, most algorithms do not exploit gauge freedoms to improve performances. We consider this issue. We analyze theoretically the impact of the gauge on standard algorithms. We show that a sufficiently general damping matrix in Levenberg-Marquardt iteration can be used to implicitly reproduce a gauge transformation. We show that if the damping matrix is chosen such that the decrease in the reprojection error is maximized, then the iteration is gauge invariant. Experimental results on simulated and real data show that our gauge invariant bundle adjustment algorithm outperforms existing ones in terms of stability.

1. Introduction

The recovery of accurate 3D structure and camera motion from images is a major research challenge in photogrammetry [1] and computer vision [13]. Bundle adjustment is a technique to refine a visual reconstruction to produce jointly optimal 3D structure and camera motion. Under certain assumptions on the noise on the observed features, bundle adjustment consists in minimizing the reprojection error, which is in general a non-linear procedure. Second-order non-linear least squares algorithms are usually employed, namely the Gauss-Newton and Levenberg-Marquardt methods. These methods iteratively improve sub-optimal parameter estimates by solving *normal equations*. Efficient solutions are possible thanks to the sparse block structure of the normal equations. The Levenberg-Marquardt method has proved the most successful due to the use of a trust region strategy, implemented via a *damping* of the normal equations, also called augmentation.

An inherent problem in bundle adjustment is the choice of the coordinate frame in which the reconstruction is expressed. This coordinate frame is called the *gauge*, or the datum in the photogrammetry community. A gauge is a subset of parameter vectors such that any two of them do not share the same underlying geometry. Parameter vectors corresponding to the same geometry are related by *gauge transformations*. It has been shown in [7] that these transformations have a group structure. The reprojection error is gauge invariant since it reflects the inherent merit of a reconstruction. These considerations hold whichever camera model is used and whichever calibration level is available. When ignored, gauge freedoms may induce numerical estimation problems, since they imply the rank deficiency of the normal equations.

It has been reported that most bundle adjustment algorithms [1, 3, 7, 8, 9, 10, 13] are *not gauge invariant*[1]. Carefully choosing the gauge can greatly improve the reliability. In practice, most algorithms converge to the same solution, but require different number of iterations. Surprisingly, gauge freedoms have rarely been exploited to improve on the reliability of algorithms.

Gauge freedoms imply that the normal equations to be solved at each iteration do not have a unique solution. Damping the normal equations consists in adding a symmetric positive definite matrix to the design matrix. This matrix is most of the time chosen as a diagonal one, which corresponds to an elliptical trust region, see e.g. [12] and §3 for more details.

The main contribution of this paper is an algorithm that maximizes the decrease in the reprojection error at each iteration, in a gauge invariant manner. In §4, we formally state the gauge invariance property of a bundle adjustment algorithm. In §5, we concentrate on Levenberg-Marquardt iterations and propose *implicit gauges*. In §6, we propose a gauge dependent damping matrix which (i) makes the underlying iteration gauge invariant, (ii) maximizes the decrease in the reprojection error up to first order and (iii) preserves the sparse block structure of the normal equations. Finally §7 reports some experimental results and §8 concludes. For simplicity of notation, we deal with projective reconstruction

[1] We employ 'gauge invariant' in a manner different from [7, 8, 9]. These authors propose algorithms based on using a pre-defined global gauge. Hence, no matter the gauge within which the initial solution is expressed, their algorithms will give the same result. However, the behaviour of these algorithms do depend upon the pre-defined global gauge. In this sense, these algorithms are *not* gauge invariant.

of points and cameras. Extension to other camera models, calibration levels, and types of features is straightforward, following [7, 9]. Note that an extended version of this paper is available.

Notation. We make no formal distinction between coordinate vectors and physical entities. Equality up to a non-null scale factor is denoted by $\sim$ and is sometimes written by introducing explicitly the scale factor: $(\mathbf{x} \sim \mathbf{x}') \Leftrightarrow (\exists \alpha | \mathbf{x} = \alpha \mathbf{x}')$. Transposition and transposed inverse are denoted by T and $^{-\mathsf{T}}$. Vectors are typeset using bold fonts ($\mathbf{q}$, $\mathbf{Q}$), matrices using sans-serif fonts (P, T) and scalars in italics. Indices are used to indicate the size of a matrix or vector ($\mathsf{P}_{(3\times4)}$, $\mathbf{q}_{(3\times1)}$) or to index a set of entities. The row-wise matrix vectorization is written vect.

Let P_i, $i = 1 \dots n$ denote the n reconstructed camera matrices and $\mathbf{Q}_j$, $j = 1 \dots m$ denote the m reconstructed 3D points. Structure and motion parameters are contained in a $(p \times 1)$ vector $\mathcal{X}$, partitionned into $12n$ motion parameters $\mathcal{M}$ and $4m$ structure parameters $\mathcal{S}$ as:

$$\mathcal{X}^{\mathsf{T}} = (\underbrace{\text{vect}^{\mathsf{T}}(\mathsf{P}_1) \ \dots \ \text{vect}^{\mathsf{T}}(\mathsf{P}_n)}_{\mathcal{M}^{\mathsf{T}}} \ \underbrace{\mathbf{Q}_1^{\mathsf{T}} \ \dots \ \mathbf{Q}_m^{\mathsf{T}}}_{\mathcal{S}^{\mathsf{T}}})$$

Changing the 15-degrees of freedom projective reconstruction basis and the individual scale factors of the camera matrices and point coordinate vectors leave the underlying geometry invariant. Hence there are $g = 15 + n + m$ degrees of gauge freedom.

2. Previous work

A survey of gauge freedom handling strategies is [13, §9]. Methods to deal with this problem are either *globally free*, i.e. they left the gauge free to drift, or *globally fixed*, i.e. they enforce a global gauge. Globally free methods are based on selecting a solution of the normal equations, using a pseudo-inverse [1], numerical damping [12, 13] or local gauge constraints, for example, the standard photogrammetric inner constraints. Globally fixed methods enforce a prechosen gauge. Some methods use reference elements (trivial gauge), either features [1, 4] (object-centered gauge) or cameras [2, 5, 6]. Another possibility for a globally fixed gauge is to use global gauge constraints [7, 8, 9], which raises the problem of enforcing these constraints during optimization. Besides the globally free photogrammetric inner constraints, all these methods are based on somehow arbitrary gauges which are not chosen to meet some efficiency criteria. A recent attempt in this direction is [10]. Based on gauge theory, the authors analyze, in the context of Euclidean reconstruction, which physical measure is the most likely to maximize the accuracy of the reconstruction while setting the scene scale.

3. Standard Algorithms

Bundle adjustment often consists in minimizing the reprojection error, given by the sum of squared differences between observed features $\mathbf{q}_{ij}$ and predicted features $\widehat{\mathbf{q}}_{ij}$:

$$\mathcal{C}(\mathcal{X}) = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \, d^2(\mathbf{q}_{ij}, \widehat{\mathbf{q}}_{ij}) = \mathbf{r}^{\mathsf{T}}\mathbf{r}, \qquad (1)$$

where w_{ij} is one if point j is visible in view i and zero otherwise. The $(r \times 1)$ vector $\mathbf{r}$ is the residual error vector. Predicted features are given by $\widehat{\mathbf{q}}_{ij} \sim \mathsf{P}_i \mathbf{Q}_j$.

Given a sub-optimal solution $\mathcal{X}_0$, the parameter vector $\mathcal{X}$ is iteratively updated by $\mathcal{X} \leftarrow \mathcal{X} + \boldsymbol{\delta}$, where the increment $\boldsymbol{\delta}$ is obtained as follows. Let $\mathsf{J} = \frac{\partial \mathbf{r}}{\partial \mathcal{X}}$ denotes the $(r \times p)$ Jacobian matrix of the residual error vector $\mathbf{r}$ with respect to structure and motion parameters $\mathcal{X}$, $\mathbf{g} = \mathsf{J}^{\mathsf{T}}\mathbf{r}$ be the $(p \times 1)$ gradient vector of $\mathcal{C}$ and let $\mathsf{N} = \mathsf{J}^{\mathsf{T}}\mathsf{J}$ be the Gauss-Newton approximation of the $(p \times p)$ Hessian matrix. The reprojection error is approximated by:

$$\mathcal{C}(\mathcal{X} + \boldsymbol{\delta}) \approx \mathcal{C}(\mathcal{X}) + \underbrace{\mathbf{g}^{\mathsf{T}}\boldsymbol{\delta} + \frac{1}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathsf{N}\boldsymbol{\delta}}_{e}. \qquad (2)$$

The minimum of this simple local quadratic model can be found by setting $\frac{\partial e}{\partial \boldsymbol{\delta}} = 0$, which gives the Gauss-Newton iteration, through the following normal equations:

$$\mathsf{N}\boldsymbol{\delta} = -\mathbf{g}. \qquad (3)$$

Note that N has a rank deficiency of g (the number of gauge freedoms). The Levenberg-Marquardt iteration is based on *damping* or augmenting the normal equations, as follows :

$$\underbrace{(\mathsf{N} + \mathsf{W}(\lambda))}_{\mathsf{N}^\star}\boldsymbol{\delta} = -\mathbf{g}, \qquad (4)$$

where $\lambda > 0$ is related to the trust region radius and $\mathsf{W}(\lambda)$ is some symmetric positive definite $(p \times p)$ weight matrix, called the *damping matrix*, often chosen as $\mathsf{W}(\lambda) = \lambda \mathrm{I}_{(p\times p)}$, which corresponds to a spherical trust region. This is the original strategy proposed by Levenberg and Marquardt and recommended in [13]. Another commonly used solution, due to [12] and recommended in [5, 11] is $\mathsf{W}(\lambda) = \lambda \mathsf{D}$ where D is a diagonal matrix containing the diagonal entries of N, which gives an elliptical trust region. The damping matrix must satisfy a normalization constraint so that the trust region radius is meaningful, e.g. $\|\mathrm{I}_{(p\times p)}\| = \sqrt{p}$. Note that the damping guarantees that $\mathsf{N}^\star$ is full-rank. Parameter λ is tuned as follows: if parameters $\mathcal{X} + \boldsymbol{\delta}$ decrease the error, i.e. if $\mathcal{C}(\mathcal{X} + \boldsymbol{\delta}) < \mathcal{C}(\mathcal{X})$, then the step is accepted and the value of λ is divided by some constant, often 10, else the step is rejected and λ is multiplied by the constant.

4. Influence of the Gauge

We formulate the gauge invariance property and examine under which conditions standard algorithms satisfy it.

4.1. Gauge Transformation and Invariance

Gauge transformations change the parameter vector without changing the underlying geometry, e.g. [7]. Let $\mathcal{X}^\mathsf{T} = (\mathcal{M}^\mathsf{T}\ \mathcal{S}^\mathsf{T})$ and $\check{\mathcal{X}}^\mathsf{T} = (\check{\mathcal{M}}^\mathsf{T}\ \check{\mathcal{S}}^\mathsf{T})$ be two parameter vectors of the same reconstruction expressed within two different gauges $\mathcal{G}$ and $\check{\mathcal{G}}$. Let T be the full-rank transformation relating the two underlying projective bases, defined such that $\check{\mathsf{P}}_i = \gamma_i \mathsf{P}_i \mathsf{T}$ and $\check{\mathbf{Q}}_j = \alpha_j \mathsf{T}^{-1}\mathbf{Q}_j$, where γ_i and α_j are unknown non-zero scale factors. Entities written with a $\check{}$ are expressed within the gauge $\check{\mathcal{G}}$. The structure and motion parameters transform as:

$$\begin{aligned}\check{\mathcal{S}} &= \mathrm{diag}(\underbrace{\alpha_1\mathsf{T}^{-1},\ldots,\alpha_m\mathsf{T}^{-1}}_{m})\,\mathcal{S}\\ \check{\mathcal{M}} &= \mathrm{diag}(\underbrace{\gamma_1\mathsf{T}^\mathsf{T},\ldots,\gamma_n\mathsf{T}^\mathsf{T}}_{3n})\,\mathcal{M},\end{aligned}$$

since $\mathrm{vect}(\gamma\mathsf{P}\mathsf{T}) = \mathrm{diag}(\gamma\mathsf{T}^\mathsf{T},\gamma\mathsf{T}^\mathsf{T},\gamma\mathsf{T}^\mathsf{T})\ \mathrm{vect}(\mathsf{P})$. We deduce that the parameter vectors are related by $\check{\mathcal{X}} = \widetilde{\mathsf{T}}\mathcal{X}$ where the *gauge transformation* $\widetilde{\mathsf{T}}$ is defined by:

$$\widetilde{\mathsf{T}} = \mathrm{diag}(\underbrace{\gamma_1\mathsf{T}^\mathsf{T},\ldots,\gamma_n\mathsf{T}^\mathsf{T}}_{3n},\underbrace{\alpha_1\mathsf{T}^{-1},\ldots,\alpha_m\mathsf{T}^{-1}}_{m}). \quad (5)$$

Note that $\det(\mathsf{T}) \neq 0 \Rightarrow \det(\widetilde{\mathsf{T}}) \neq 0$ since the γ_i and α_j are non-zero. In general, $\mathcal{X}$ and $\check{\mathcal{X}}$ are related by a unique gauge transformation.

The geometric equivalence of two reconstructions, denoted by $\cong$, is an egality 'up to the gauge' defined by:

$$(\mathcal{X} \cong \check{\mathcal{X}}) \Leftrightarrow \left(\exists\widetilde{\mathsf{T}} \mid \check{\mathcal{X}} = \widetilde{\mathsf{T}}\mathcal{X}\right). \quad (6)$$

with $\widetilde{\mathsf{T}}$ of the form (5). We can now formally state the gauge invariance property. A bundle adjustment algorithm is gauge invariant if it preserves geometric equivalence: $(\mathcal{X} \cong \check{\mathcal{X}}) \Rightarrow ((\mathcal{X}+\boldsymbol{\delta}) \cong (\check{\mathcal{X}}+\check{\boldsymbol{\delta}}))$. By writing explicitly this property using equation (6), expanding, and since $\mathcal{X}$ and $\check{\mathcal{X}}$ are related by a unique $\widetilde{\mathsf{T}}$, we obtain the following definition of gauge invariance:

$$\left(\exists\widetilde{\mathsf{T}} \mid (\check{\mathcal{X}} = \widetilde{\mathsf{T}}\mathcal{X})\right) \Rightarrow (\check{\boldsymbol{\delta}} = \widetilde{\mathsf{T}}\boldsymbol{\delta}). \quad (7)$$

Note that although derived in a different manner, this definition is similar to the one given in [9].

4.2. Gauss-Newton

Consider a Gauss-Newton iteration. The increment $\boldsymbol{\delta}$ within gauge $\mathcal{G}$ is given by solving the normal equations (3). Due to gauge freedoms, matrix N has a rank deficiency of g, and multiple solutions hold, corresponding to different gauges. Let G be a $(p \times g)$ full column-rank matrix defined by $\mathsf{N}\mathsf{G} = 0$, i.e. the columns of G span the nullspace of N. Denoting by † the Moore-Penrose pseudo-inverse, the possible solutions of the normal equations are parameterized by a $(g \times 1)$ vector $\mathbf{v}$ as $\boldsymbol{\delta} = -(\mathsf{J}^\mathsf{T}\mathsf{J})^\dagger\mathsf{J}^\mathsf{T}\mathbf{r} + \mathsf{G}\mathbf{v}$, where we substituted $\mathsf{N} = \mathsf{J}^\mathsf{T}\mathsf{J}$ and $\mathbf{g} = \mathsf{J}^\mathsf{T}\mathbf{r}$. Since $\mathsf{J} = \frac{\partial\mathbf{r}}{\partial\mathcal{X}} = \frac{\partial\mathbf{r}}{\partial\check{\mathcal{X}}}\frac{\partial\check{\mathcal{X}}}{\partial\mathcal{X}} = \check{\mathsf{J}}\widetilde{\mathsf{T}}$, we obtain: $\boldsymbol{\delta} = -(\widetilde{\mathsf{T}}^\mathsf{T}\check{\mathsf{J}}^\mathsf{T}\check{\mathsf{J}}\widetilde{\mathsf{T}})^\dagger\widetilde{\mathsf{T}}^\mathsf{T}\check{\mathsf{J}}\mathbf{r} + \widetilde{\mathsf{T}}^{-1}\check{\mathsf{G}}\mathbf{v}$. From equation (7), gauge invariance holds if and only if $(\widetilde{\mathsf{T}}^\mathsf{T}\check{\mathsf{N}}\widetilde{\mathsf{T}})^\dagger = \widetilde{\mathsf{T}}^{-1}\check{\mathsf{N}}^\dagger\widetilde{\mathsf{T}}^{-\mathsf{T}}$ and $\mathbf{v} = \check{\mathbf{v}}$, which is verified if and only if $\widetilde{\mathsf{T}}$ is orthonormal [9], i.e. $\widetilde{\mathsf{T}}^{-1} = \widetilde{\mathsf{T}}^\mathsf{T}$. Hence, Gauss-Newton-based bundle adjustment is *not* gauge invariant. The previously derived condition does not leave enough flexibility to be exploited in this direction.

4.3. Levenberg-Marquardt

Consider a Levenberg-Marquardt iteration. The increment $\boldsymbol{\delta}$ within gauge $\mathcal{G}$ is given by solving the damped normal equations (4) $\mathsf{N}^\star\boldsymbol{\delta} = -\mathbf{g}$. By substituting $\mathsf{N}^\star = \mathsf{N} + \mathsf{W}(\lambda) = \mathsf{J}^\mathsf{T}\mathsf{J} + \mathsf{W}(\lambda)$ and $\mathbf{g} = \mathsf{J}^\mathsf{T}\mathbf{r}$ in this equation and since $\det(\mathsf{N}^\star) \neq 0$, we obtain: $\boldsymbol{\delta} = -(\mathsf{J}^\mathsf{T}\mathsf{J} + \mathsf{W}(\lambda))^{-1}\mathsf{J}^\mathsf{T}\mathbf{r}$. By substituting $\mathsf{J} = \check{\mathsf{J}}\widetilde{\mathsf{T}}$ and expanding, we get:

$$\begin{aligned}\boldsymbol{\delta} &= -(\widetilde{\mathsf{T}}^\mathsf{T}\check{\mathsf{J}}^\mathsf{T}\check{\mathsf{J}}\widetilde{\mathsf{T}} + \mathsf{W}(\lambda))^{-1}\widetilde{\mathsf{T}}^\mathsf{T}\check{\mathsf{J}}^\mathsf{T}\mathbf{r}\\ &= -\widetilde{\mathsf{T}}^{-1}(\check{\mathsf{N}} + \widetilde{\mathsf{T}}^{-\mathsf{T}}\mathsf{W}(\lambda)\widetilde{\mathsf{T}}^{-1})^{-1}\check{\mathsf{J}}^\mathsf{T}\mathbf{r}.\end{aligned}$$

From equation (7), gauge invariance holds if and only if:

$$\check{\mathsf{W}}(\lambda) = \widetilde{\mathsf{T}}^{-\mathsf{T}}\mathsf{W}(\lambda)\widetilde{\mathsf{T}}^{-1}. \quad (8)$$

The usual choices for matrices $\mathsf{W}(\lambda)$ and $\check{\mathsf{W}}(\lambda)$, e.g. $\mathsf{W}(\lambda) = \check{\mathsf{W}}(\lambda) = \lambda\mathrm{I}$, do not fulfill the above-derived property[2]. Hence, standard implementations of Levenberg-Marquardt-based algorithms are *not* gauge invariant. Equation (8) can be verified if an appropriate gauge dependent choice is made for the damping matrices. This is the cornerstone for the gauge invariant method proposed in the next section.

5. Explicit and Implicit Gauges

Standard bundle adjustment algorithms are not gauge invariant. Hence, given a parameter estimate, there must exist an optimal local gauge within which the iteration reduces the reprojection error better than within the others. The problem of finding this gauge is dealt with in the next section. In this section, we examine the relationship between the damping matrix and gauge transformations.

A commonly used solution to globally enforce a gauge, e.g. [7, 8, 9], is what we call *explicit gauge* fixing. It consists in explicitly expressing the reconstruction within the desired gauge before each iteration.

We propose *implicit gauge* fixing. Consider equation (8). On the one hand, it tells us that the algorithm is not gauge invariant. On the other hand, it shows that there exists a close link between the gauge and the damping matrix. Hence, it

[2] It can be shown that equation (8) with $\mathsf{W}(\lambda) = \check{\mathsf{W}}(\lambda) = \lambda\mathrm{I}$ is verified as in the Gauss-Newton case, i.e. if $\widetilde{\mathsf{T}}$ is orthonormal.

can be used to perform bundle adjustment within gauge $\check{\mathcal{G}}$, while choosing $\check{\mathsf{W}}(\lambda)$ such that it behaves exactly as if is was conducted within gauge $\mathcal{G}$. For example if $\mathsf{W}(\lambda) = \lambda\mathsf{I}$ and $\check{\mathsf{W}}(\lambda) = \lambda\widetilde{\mathsf{T}}^{-\mathsf{T}}\widetilde{\mathsf{T}}^{-1}$, then running the algorithm within gauges $\mathcal{G}$ and $\check{\mathcal{G}}$ is strictly equivalent. We confirm in our experiments that explicit and implicit gauges give exactly the same results. Let $\mathsf{A} = \mathsf{T}^{\mathsf{T}}\mathsf{T}$, then $\check{\mathsf{W}}(\lambda) = \lambda\widetilde{\mathsf{T}}^{-\mathsf{T}}\widetilde{\mathsf{T}}^{-1}$ is given by:

$$\check{\mathsf{W}}(\lambda) = \lambda\,\mathrm{diag}(\underbrace{\gamma_1^2\mathsf{A}^{-1},\ldots,\gamma_n^2\mathsf{A}^{-1}}_{3n},\underbrace{\alpha_1^2\mathsf{A},\ldots,\alpha_m^2\mathsf{A}}_{m}). \quad (9)$$

Proposition 1 (Gauge invariance conditions) *Levenberg-Marquardt iterations can be made gauge invariant if the following two conditions are verified: (i) the form of the damping matrix is at least as general as the form given by equation (9) and (ii) the damping matrix is uniquely determined by a gauge dependent criterion defined such that equation (8) is verified.*

6. Gauge Dependent Damping

We propose a solution to choose a gauge dependent damping matrix $\mathsf{W}(\lambda) = \lambda\mathcal{W}$ which has the three following important properties. First, it makes the underlying iteration gauge invariant. Second, it maximizes the decrease in the reprojection error up to first order. Third, it preserves the sparse block structure of the normal equations.

In order to achieve this result, we begin by deriving an approximation of the reprojection error, depending upon $\mathsf{W}(\lambda)$. A quadratic approximation of $\mathcal{C}(\mathcal{X}+\boldsymbol{\delta})$ is given by equation (2). The incremental error e is given by $e = \mathbf{g}^{\mathsf{T}}\boldsymbol{\delta} + \frac{1}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathsf{N}\boldsymbol{\delta}$. The increment vector $\boldsymbol{\delta}$ is given by solving the augmented normal equations (4) $\boldsymbol{\delta} = -(\mathsf{N}+\mathsf{W}(\lambda))^{-1}\mathbf{g}$. Since $\det(\mathsf{W}(\lambda)) \neq 0$, we can formulate the following Taylor expansion $(\mathsf{N}+\mathsf{W}(\lambda))^{-1} = \sum_{n=0}^{\infty}(-1)^n\mathsf{Z}(n)$, where $\mathsf{Z}(n) = (\mathsf{W}(\lambda)^{-1}\mathsf{N})^n\mathsf{W}(\lambda)^{-1}$. This expression is valid provided $\|\mathsf{W}(\lambda)^{-1}\mathsf{N}\| < 1$, i.e. when N is small or λ is large. This leads to $\boldsymbol{\delta} = -\sum_{n=0}^{\infty}(-1)^n\mathsf{Z}(n)\mathbf{g}$. We substitute the above expression in equation (2). The first term is rewritten as $\mathbf{g}^{\mathsf{T}}\boldsymbol{\delta} = -\sum_{n=0}^{\infty}(-1)^n\mathsf{K}(n)$, where $\mathsf{K}(n) = \mathbf{g}^{\mathsf{T}}\mathsf{Z}(n)\mathbf{g}$. The second term becomes:

$$\frac{1}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathsf{N}\boldsymbol{\delta} = \frac{1}{2}\sum_{n=0}^{\infty}\left((-1)^n\mathbf{g}^{\mathsf{T}}\mathsf{Z}(n)\right)\mathsf{N}\sum_{n=0}^{\infty}\left((-1)^n\mathsf{Z}(n)\mathbf{g}\right).$$

After expansion and using the following property arising from the symmetry of $\mathsf{W}(\lambda)$: $\mathsf{Z}(n)\mathsf{N}\mathsf{Z}(n') = \mathsf{Z}(n+n'+1)$, we obtain $\frac{1}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathsf{N}\boldsymbol{\delta} = \frac{1}{2}\sum_{n=0}^{\infty}(-1)^n(n+1)\mathsf{K}(n+1)$, which gives the following Taylor expansion for e:

$$e = \sum_{n=0}^{\infty}(-1)^{n+1}\left(1+\frac{n}{2}\right)\mathsf{K}(n). \quad (10)$$

Proposition 2 (Gauge invariant iterations) *The damping matrix defined such that equation (10) is minimized satisfies proposition 1. Hence, the underlying iteration is gauge invariant.*

Proof. We drop the parameter λ for this proof. Equation (10) can be rewritten as:

$$e = \mathbf{g}^{\mathsf{T}}\left(-\mathsf{W}^{-1} + \frac{3}{2}\mathsf{W}^{-1}\mathsf{N}\mathsf{W}^{-1} - \ldots\right)\mathbf{g}.$$

Transfer the reconstruction from the gauge $\mathcal{G}$ to $\check{\mathcal{G}}$ by applying the gauge transformation $\widetilde{\mathsf{T}}$. By substituting $\mathbf{g} = \widetilde{\mathsf{T}}^{\mathsf{T}}\check{\mathbf{g}}$ and $\mathsf{N} = \widetilde{\mathsf{T}}^{\mathsf{T}}\check{\mathsf{N}}\widetilde{\mathsf{T}}$, we obtain:

$$e = \check{\mathbf{g}}^{\mathsf{T}}\widetilde{\mathsf{T}}\left(-\mathsf{W}^{-1} + \frac{3}{2}\mathsf{W}^{-1}\widetilde{\mathsf{T}}^{\mathsf{T}}\check{\mathsf{N}}\widetilde{\mathsf{T}}\mathsf{W}^{-1} - \ldots\right)\widetilde{\mathsf{T}}^{\mathsf{T}}\check{\mathbf{g}},$$

and hence, $\check{\mathsf{W}}^{-1} = \widetilde{\mathsf{T}}\mathsf{W}^{-1}\widetilde{\mathsf{T}}^{\mathsf{T}}$, which verifies the gauge invariance equation (8) and concludes the proof. ■

Finding the damping matrix that minimizes e, equation (10), is a complicated problem. We keep only the first-order term: $e \approx -\mathbf{g}^{\mathsf{T}}\mathsf{W}(\lambda)^{-1}\mathbf{g}$, which reduces the problem to finding a symmetric, positive definite, damping matrix $\lambda\mathcal{W}$ $\mathcal{W} = \arg\max_{\mathcal{W},\|\mathcal{W}\|^2=p}\mathbf{g}^{\mathsf{T}}(\lambda\mathcal{W})^{-1}\mathbf{g}$. The normalization condition $\|\mathcal{W}\|^2 = p$ is the same as the standard choice $\mathsf{W}(\lambda) = \lambda\mathrm{I}$ ($\|\mathrm{I}_{(p\times p)}\|^2 = p$). Since $\mathsf{W}(\lambda)$ is a symmetric matrix, $\mathsf{W}(\lambda)^{-1} = \frac{1}{\lambda}\mathsf{R}^{\mathsf{T}}\mathsf{R}$, where R is an upper-triangular matrix. With this parameterization, and since $\mathbf{g}^{\mathsf{T}}\frac{1}{\lambda}\mathsf{R}^{\mathsf{T}}\mathsf{R}\mathbf{g} = \frac{1}{\lambda}\|\mathsf{R}\mathbf{g}\|^2$, we obtain:

$$\mathsf{R} = \arg\max_{\mathsf{R},\|\mathsf{R}^{\mathsf{T}}\mathsf{R}\|^2=p}\frac{1}{\lambda}\|\mathsf{R}\mathbf{g}\|^2. \quad (11)$$

Hence, the coefficients of matrix R can be found by solving a simple linear least squares optimization problem.

From proposition 1 and imposing the fact that the damping should not spoil the sparse block structure of the normal equations, we are left with the following possible form:

$$\mathsf{R} = \mathrm{diag}(\underbrace{\mathsf{R}_{\mathcal{M},1},\ldots,\mathsf{R}_{\mathcal{M},n}}_{n},\underbrace{\mathsf{R}_{\mathcal{S},1},\ldots,\mathsf{R}_{\mathcal{S},m}}_{m}),$$

where the $\mathsf{R}_{\mathcal{M},i}$ and the $\mathsf{R}_{\mathcal{S},j}$ are respectively (12×12) and (4×4) upper-triangular matrices. Given that each of the p parameters gives one constraint on the damping matrix and since it must be uniquely determined (i.e. it must not have more than p parameters), we propose the following $10(n+1)$-parameter choice: $\mathsf{R}_{\mathcal{M},i} = \mathrm{diag}(\mathsf{M}_i,\mathsf{M}_i,\mathsf{M}_i)$ and $\mathsf{R}_{\mathcal{S},j} = \mathsf{S}$, where the M_i and S are (4×4) upper-triangular matrices. This choice means that the parameter block of each camera i has its own 10 parameter damping block $\mathsf{M}_i^{\mathsf{T}}\mathsf{M}_i$, while a unique 10 parameter damping block $\mathsf{S}^{\mathsf{T}}\mathsf{S}$ is defined for point parameters. This form is handy since each block can be found independently by solving a problem of the

1. Compute the damping matrix $\mathcal{W}(\lambda)^{-1} = \lambda \mathsf{R}^\mathsf{T}\mathsf{R}$ where:

$$\mathsf{R} = \mathrm{diag}(\underbrace{\mathsf{M}_1, \ldots, \mathsf{M}_n}_{3n}, \underbrace{\mathsf{S}, \ldots, \mathsf{S}}_{m}).$$

The (4×4) upper-triangular matrices M_i and S are formed by solving (see main text):

$$\mathsf{M}_i = \arg \max_{\mathsf{M}_i, \|\mathsf{M}_i^\mathsf{T}\mathsf{M}_i\|^2=12} \sum_{k=1}^{3} \|\mathsf{M}_i \mathbf{g}_{\mathcal{M},i}^k\|^2$$

$$\mathsf{S} = \arg \max_{\mathsf{S}, \|\mathsf{S}^\mathsf{T}\mathsf{S}\|^2=4} \sum_{j=1}^{m} \|\mathsf{S}\mathbf{g}_{\mathcal{S},j}\|^2,$$

where $\mathbf{g}_{\mathcal{S},j}$ and $\mathbf{g}_{\mathcal{M},i}^k$ are (4×1) gradient vectors for the j-th point and for the k-th row of the i-th camera matrix respectively.

2. Perform one Levenberg-Marquardt iteration with the damping matrix $\mathcal{W}(\lambda)$. This iteration is gauge invariant due to the above choice.
3. Optional: Project the estimate onto a global gauge.

Table 1. The gauge invariant iteration we propose. Periodical enforcement of global gauge constraints (step 3) such as renormalization of homogeneous coordinates is recommended. Gauge invariance holds in the sense that any global gauge can be enforced in step 3 without affecting the result.

form (11). More details are given in table 1. Below, we give details about solving the low-dimensional maximization problems of table 1. Let the problem to be solved be $\max_{\mathbf{x}, \|\mathbf{x}\|^2=b} \|\mathsf{C}\mathbf{x}\|^2$. The singular vector $\mathbf{v}$ associated to the largest singular value of matrix C gives the solution for $\mathbf{x}$ as $\mathbf{x} = \sqrt{b}\mathbf{v}$. Singular value decomposition of matrix C can be used to compute $\mathbf{v}$.

7. Experimental Results

We compare our gauge invariant algorithm, denoted GAUGE_INVARIANT, to some other ones. FREE directly optimizes the camera matrices and 3D points. Pseudo-inverse is used to solve the normal equations. HARTLEY [6] consists in using a reference camera to partially eliminate the gauge and BARTOLI [2] consists in using two reference cameras to eliminate the gauge. FAUGERAS [4] is based on fixing the coordinates of 5 reference points to enforce the gauge.MCLAUCHLAN [8] consists in enforcing a normalized basis after each iteration and in using first order gauge constraints in the normal equations and KANATANI [7] is similar to method FREE but periodically projects the estimate on a pre-chosen gauge to prevent it to drift too much. The initial solution, denoted by INIT, is computed by registering cameras in turn to a two-view reconstruction.

7.1. Simulated Data

We simulate $m = 100$ points lying in a cube with 1 meter side length, observed by $n = 5$ cameras with a focal length of 1000 pixels. The points are offset from a base plane lying inside the cube, with a mean offset denoted by d. Cameras are situated 10 meters away from the center of the cube. The baseline between consecutive cameras is 3 meters. All points are visible in all views. We add a centered Gaussian noise on true point positions with a 1 pixel variance. The results are averaged over 50 trials.

Figure 1 (left and middle) shows the results when the scene flatness, i.e. the mean offset d from the plane, is varied. We observe that the reprojection errors, i.e. the accuracy, are undistinguishable for all methods, except for methods FREE which converges to a different local minimum than the others, for weak geometry, i.e. when the mean offset from the plane d is small. Concerning the number of iterations, i.e. the computational cost, we observe that when the mean offset is large, i.e. when the geometry is strong, there is only slight differences between the different methods besides for method FREE, which takes clearly more iterations to converge. As the geometry becomes weaker, i.e. when the offset decreases, large discrepancies can be observed between the different methods. Method FAUGERAS based on reference points gives bad results, since using reference points to fix a projective basis can be very unstable. Methods HARTLEY and BARTOLI based on reference cameras and KANATANI based on periodical projection on a pre-chosen gauge give reliable results when the geometry is strong enough. Method MCLAUCHLAN performs better, while GAUGE_INVARIANT is the method the less sensitive to the instability of the scene.

We tested other strong to weak scene configurations, based on varying the number of points and cameras, the baseline between consecutive cameras and the visibility (not shown here due to lack of space). We observed similar results as above: strong geometries reduce the discrepancies between the different methods. Similarly, we tested the influence of the initialization. As expected, discrepancies between the different methods reduce as the initialization gets more accurate.

Figure 1 (right) shows the reprojection error as a function of the iterations for the scene setting with $d = 0.5$. As expected, we observe that the error is faster to decrease as the number of iterations gets low.

7.2. Real Data

We compare the algorithms on various image streams. For two of them, the office sequence and the hotel sequence[3], we show results, see table 2. For the office sequence, all methods give a reprojection error of 0.8392 pixel. For the hotel

[3]provided by the Modeling by Videotaping group in the Robotics Institute, Carnegie Mellon University.

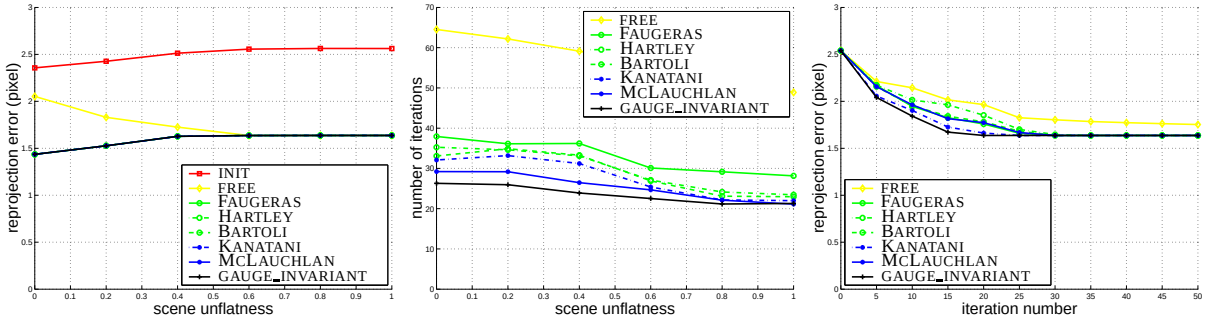

Figure 1. Reprojection error (left) and number of iterations (middle) when varying the scene unflatness from weak to strong geometry and reprojection error as a function of the iterations (right).

sequence, they all gave 0.9594 pixel, besides method FREE which gave 1.5268 pixels. Concerning the number of iterations, the same observations as for simulated data can be made. The hotel sequence gives a weak geometry since an affine camera model is well-adapted to these data. Hence, a full perspective projection model is not well-constrained, which tends to increase the discrepancies between the different methods.

Algorithm	office seq.	hotel seq.
FREE	12	28
HARTLEY	7	16
BARTOLI	8	19
FAUGERAS	10	22
MCLAUCHLAN	6	15
KANATANI	7	16
GAUGE_INVARIANT	7	13

Table 2. Number of iterations of the algorithms for the office and the hotel sequences.

8. Conclusions

We proposed a bundle adjustment algorithm which maximizes the decrease in the error at each iteration, regardless the gauge within which the reconstruction is expressed. We derived this algorithm based on a careful study of how the gauge influence the iterations of standard algorithms, which are not gauge invariant. In particular, concerning Levenberg-Marquardt iterations, we showed that the standard diagonal damping matrices, defining elliptical trust regions, transform to a more complex shaped trust region when changing the gauge. Based on this, we proposed a gauge dependent damping matrix and a practical algorithm to compute it, that allows gauge invariant iterations while maximizing the decrease in the reprojection error. The sparse block structure of the normal equations is preserved.

We compared our algorithm to existing ones using simulated and real data. We observed that it is more reliable in the sense that it is less sensitive to weak geometry and weak initialization. Moreover, the error decreases more intensely throughout the iterations. Most algorithms converge to the same solution, but require different number of iterations.

References

[1] K. Atkinson, editor. *Close Range Photogrammetry and Machine Vision*. Whittles Publishing, 1996.

[2] A. Bartoli. On the non-linear optimization of projective motion using minimal parameters. In *ECCV*, May 2002.

[3] A. Dermanis. The photogrammetric inner constraints. *J. of Photogrammetry and Remote Sensing*, 49(1):25–39, 1994.

[4] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *ECCV*, May 1992.

[5] R. Hartley. Euclidean reconstruction from uncalibrated views. In *Proceeding of the Workshop on Applications of Invariants in Computer Vision*, pages 187–202, October 1993.

[6] R. Hartley. Projective reconstruction and invariants from multiple images. *PAMI*, 16(10):1036–1041, October 1994.

[7] K. Kanatani and D. D. Morris. Gauges and gauge transformations for uncertainty description of geometric structure with indeterminacy. *Trans. on Inf. Theory*, 47(5), 2001.

[8] P. F. McLauchlan. Gauge invariance in projective 3D reconstruction. In *Proceedings of the Multi-View Workshop*, 1999.

[9] P. F. McLauchlan. Gauge independence in optimization algorithms for 3D vision. In *Proceedings of the Vision Algorithms Workshop, Dublin, Ireland*, 2000.

[10] D. D. Morris, K. Kanatani, and T. Kanade. Gauge fixing for accurate 3D estimation. In *CVPR*, December 2001.

[11] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C - The Art of Scientific Computing*. Cambridge University Press, 2nd edition, 1992.

[12] G. A. F. Seber and C. J. Wild. *Non-linear regression*. John Wiley & Sons, New York, 1989.

[13] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle ajustment — a modern synthesis. In *International Workshop on Vision Algorithms: Theory and Practice*, 2000.