



HAL
open science

Resolvability of Visual-Inertial Structure from Motion in Closed-form

Agostino Martinelli

► **To cite this version:**

Agostino Martinelli. Resolvability of Visual-Inertial Structure from Motion in Closed-form. [Research Report] RR-8076, 2012. hal-00735826v1

HAL Id: hal-00735826

<https://inria.hal.science/hal-00735826v1>

Submitted on 26 Sep 2012 (v1), last revised 24 Mar 2013 (v2)

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.



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

***Resolvability of Visual-Inertial Structure from
Motion in Closed-form***

Agostino Martinelli

N° 8076

Septembre 2012

Thèmes COG et NUM

 *Rapport
de recherche*

Resolvability of Visual-Inertial Structure from Motion in Closed-form

Agostino Martinelli

Thèmes COG et NUM — Systèmes cognitifs et Systèmes numériques
Équipe-Projet Emotion

Rapport de recherche n° 8076 — Septembre 2012 — 22 pages

Abstract: This paper investigates the visual-inertial structure from motion problem. A simple closed form solution to this problem is introduced. Special attention is devoted to identify the conditions under which the problem has a finite number of solutions. Specifically, it is shown that the problem can have a unique solution, two distinct solutions and infinite solutions depending on the trajectory, on the number of point-features and on their layout and on the number of camera images. The investigation is also performed in the case when the inertial data are biased, showing that, in this latter case, more images and more restrictive conditions on the trajectory are required for the problem resolvability.

Key-words: Sensor Fusion, Inertial Sensors, Vision, Structure from Motion

Résumé : Cet article étudie le problème *visual inertial structure from motion*. Une solution analytique est proposée. Une attention particulière est consacrée à identifier les conditions dans lesquelles le problème a un nombre fini de solutions. Plus précisément, il est montré que, en fonction de la trajectoire, du nombre de points et du nombre d'images de caméra, le problème peut avoir une solution unique, deux solutions distinctes et infinies solutions. L'analyse est également effectuée dans le cas où les données inertielles sont biaisées, montrant que, dans ce dernier cas, plus d'images et des conditions plus restrictives sur la trajectoire sont nécessaires pour la résolubilité du problème.

Mots-clés : Fusion Sensoriel, Capteurs inertiels, Vision, Structure from Motion

1 Introduction

The structure from motion problem (SfM) consists of determining the three-dimensional structure of the scene by using the measurements provided by one or more sensors over time (e.g. vision sensors, ego-motion sensors, range sensors). In the case of visual measurements only, the SfM problem has been solved up to a scale [3, 4, 8, 12, 16] and a closed form solution has also been derived [8, 12, 16], allowing the determination of the three-dimensional structure of the scene, without the need for any prior knowledge.

The case of inertial and visual measurements, i.e., the visual-inertial structure from motion problem (from now on the Vi-SfM problem), has particular interest and has been investigated by many disciplines, both in the framework of computer science [2, 10, 11, 13, 17] and in the framework of neuroscience (e.g., [1, 5, 7]). Prior work has answered the question of which are the observable modes, i.e. the states that can be determined by fusing visual and inertial measurements [2, 10, 11, 13]. Specifically, it has been shown that the velocity, the absolute scale, the gravity vector in the local frame and the bias-vectors which affect the inertial measurements, are observable modes. On the other hand, the vast majority of existing techniques for navigation using visual and IMU measurements employ either a recursive Bayesian estimation approach [2, 10, 11], or a smoothing formulation [17]. In [9] the authors investigated the estimator inconsistency in the Vi-SfM problem. They found that standard EKF-based estimators lead to spurious information gain along unobservable directions. They also suggested a modification on the basic estimator in order to enforce the unobservable directions and thus to reduce inconsistency. On the other hand, both a recursive Bayesian estimation approach and a smoothing formulation require a prior estimate for the initial state: since both rely on linearization of the nonlinear measurement models, lack of a good initial estimate can lead to divergence. This important limitation would be eliminated by introducing a deterministic solution, i.e., by analytically expressing all the observable modes in terms of the measurements provided by the sensors during a short time-interval. Closed form solutions have been introduced very recently in [13]. Specifically, in [13] an observability analysis allowed us to quantify the information resulting when combining visual and inertial measurements. This allowed us to analytically derive the observable modes. Then, starting from the differential equations which characterize a generic 3D-motion and from the analytical expression of the visual observations, closed form expressions of the observable modes in terms of the sensor measurements were derived. On the other hand, these derivations did not allow us to detect the conditions under which the Vi-SfM can be solved. This important issue was very marginally investigated in [13]. Specifically, the observability analysis carried out in [13] only allowed us to detect a very limited number of singular cases where the sensor information does not allow us to determine the observable modes.

Here we derive a new simple and intuitive closed solution to the Vi-SfM problem. Compared with the solutions proposed in [13], this new solution allows us to investigate the intrinsic properties of the Vi-SfM problem and to identify the conditions under which the problem can be solved in closed form. In particular, these conditions regard the trajectory, the number of point-features and their layout and the number of monocular images where the same point-features are seen. Additionally, minimal cases have been fully investigated, i.e., necessary

and sufficient conditions on the trajectory and on the feature layout have been provided for the cases when the number of features and the number of camera images is the minimum required for the Vi-SfM problem resolvability. The investigation is also performed in the case when the inertial data are biased, showing that, in this latter case, more images and more restrictive conditions on the trajectory are required in order to have a finite number of solutions.

The paper is organized as follows. The system is defined in section 2. In section 3 the Vi-SfM problem is reduced to a polynomial equation system, whose resolvability is investigated in section 4, both in the unbiased (4.1) and biased (4.2) case. All the possible cases are then summarized in section 5 and concluding remarks are provided in 6.

2 The Considered System

We consider a system (from now on we call it the *platform*) consisting of a monocular camera and an Inertial Measurement Unit (*IMU*). The IMU consists of three orthogonal accelerometers and three orthogonal gyrometers. We introduce a global frame in order to characterize the motion of the platform moving in a 3D environment. Its z -axis points vertically upwards. As we will see, for the next derivation we do not need to better define this global frame. We will adopt lower-case letters to denote vectors in this frame (e.g. the gravity is $\mathbf{g} = [0, 0, -g]^T$, where $g \simeq 9.8 \text{ ms}^{-2}$). We assume that the transformations among the camera frame and the IMU frame are known (we assume that the platform frame coincides with the camera frame and we call it the local frame). We will adopt upper-case letters to denote vectors in this frame. Since this local frame is time dependent, we adopt the following notation: $\mathbf{W}_t(\tau)$ will be the vector with global coordinates $\mathbf{w}(\tau)$ in the local frame at time t . Additionally, we will denote with $C_{t_1}^{t_2}$ the matrix which characterizes the rotation occurred during the time interval (t_1, t_2) . We have: $\mathbf{W}_{t_1}(\tau) = C_{t_1}^{t_2} \mathbf{W}_{t_2}(\tau)$ and $(C_{t_1}^{t_2})^{-1} = C_{t_2}^{t_1}$. Finally, C^t will denote the rotation matrix between the global frame and the local frame at time t , i.e., $\mathbf{w}(\tau) = C^t \mathbf{W}_t(\tau)$.

The *IMU* provides the platform angular speed and acceleration. Regarding the acceleration, the one perceived by the accelerometer (\mathbf{A}) is not simply the inertial acceleration ($\mathbf{A}^{inertial}$). It also includes the gravitational acceleration (\mathbf{G}).

We assume that the camera is observing one or more point features during the time interval $[T_{in}, T_{fin}]$. The platform and one of these observed features are displayed in fig 1.

3 The closed form solution

Prior work has answered the question of which are the observable modes, i.e. the states that can be determined by fusing visual and inertial measurements [2, 10, 11, 13]. The observable modes are: the platform velocity, the absolute scale, the gravity vector in the local frame and the bias-vectors which affect the inertial measurements. Note that the knowledge of the gravity in the local frame is equivalent to the knowledge of its magnitude together with the roll and pitch angle, i.e., the orientation of the platform with respect to the horizontal

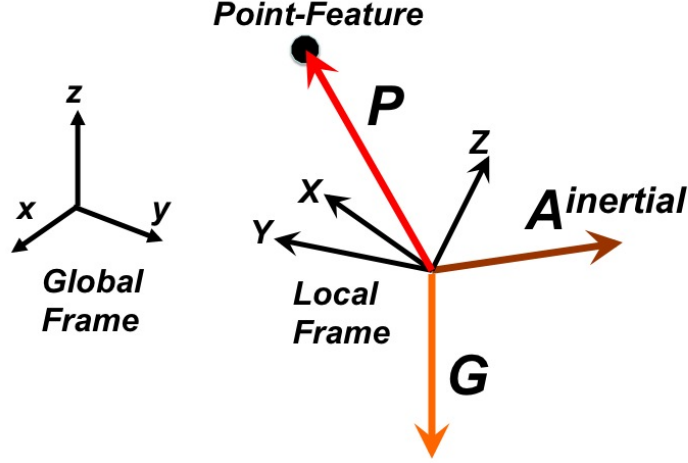


Figure 1: Global and local frame with the feature position (P), the platform acceleration ($\mathbf{A}^{inertial}$) and the gravitational acceleration (\mathbf{G}).

plane. Our goal is to express in closed-form all the observable modes at a given time T_{in} only in terms of the visual and inertial measurements obtained during the time interval $[T_{in}, T_{fin}]$.

The position of the platform \mathbf{r} at any time $t \in [T_{in}, T_{fin}]$ satisfies the equation:

$$\mathbf{r}(t) = \mathbf{r}(T_{in}) + \mathbf{v}(T_{in})\Delta t + \int_{T_{in}}^t \int_{T_{in}}^{\tau} \mathbf{a}(\xi) d\xi d\tau \quad (1)$$

where $\mathbf{v} \equiv \frac{d\mathbf{r}}{dt}$, $\mathbf{a} \equiv \frac{d\mathbf{v}}{dt}$ and $\Delta t \equiv t - T_{in}$. The accelerometer does not provide the vector $\mathbf{a}(\xi)$. It provides the acceleration in the local frame and it also perceives the gravitational component. Additionally, its data are usually biased [6], i.e., they are corrupted by a constant term (\mathbf{B})¹. In other words, the accelerometer provides the vector: $\mathbf{A}_{\xi}(\xi) \equiv \mathbf{A}_{\xi}^{inertial}(\xi) - \mathbf{G}_{\xi} + \mathbf{B}$. Note that the gravity comes with a minus since, when the platform does not accelerate (i.e. $\mathbf{A}_{\xi}^{inertial}(\xi)$ is zero), the accelerometer perceives an acceleration which is the same of an object accelerated upward in absence of gravity. Note also that the vector \mathbf{G}_{ξ} depends on time only because the local frame can rotate.

We write equation (1) by highlighting the vector $\mathbf{A}_{\xi}(\xi)$ provided by the accelerometer:

$$\mathbf{r}(t) = \mathbf{r}(T_{in}) + \mathbf{v}(T_{in})\Delta t + \mathbf{g} \frac{\Delta t^2}{2} + C^{T_{in}} [\mathbf{S}_{T_{in}}(t) - \Gamma \mathbf{B}] \quad (2)$$

where:

¹Actually, the accelerometer bias slightly changes with time, i.e., it would be more appropriate to write $\mathbf{B}(\xi)$. However, as we will show in the next section, few camera images allow us to determine this component and we can assume that the bias is constant during the time interval needed to collect few camera images.

$$\begin{aligned}\mathbf{S}_{T_{in}}(t) &\equiv \int_{T_{in}}^t \int_{T_{in}}^{\tau} C_{T_{in}}^{\xi} \mathbf{A}_{\xi}(\xi) d\xi d\tau; \\ \Gamma &\equiv \int_{T_{in}}^t \int_{T_{in}}^{\tau} C_{T_{in}}^{\xi} d\xi d\tau\end{aligned}$$

The matrix $C_{T_{in}}^{\xi}$ can be obtained from the angular speed during the interval $[T_{in}, \xi]$ provided by the gyroscopes [6]. Hence, also the matrix $\Gamma = \Gamma(t)$ can be obtained by directly integrating the gyroscope data during the interval $[T_{in}, t]$. Finally, the vector $\mathbf{S}_{T_{in}}(t)$ can be obtained by integrating the data provided by the gyroscopes and the accelerometers delivered during the interval $[T_{in}, t]$.

Let us suppose that N point-features are observed, simultaneously. Let us denote their position in the physical world with \mathbf{p}^i , $i = 1, \dots, N$. According to our notation, $\mathbf{P}_t^i(t)$ will denote their position at time t in the local frame at time t . We have: $\mathbf{p}^i = \mathbf{r}(t) + C^{T_{in}} C_{T_{in}}^t \mathbf{P}_t^i(t)$ and, by using (2), we obtain the following equation:

$$\begin{aligned}\mathbf{P}_t^i(t) &= C_t^{T_{in}} \left[\mathbf{P}_{T_{in}}^i(T_{in}) - \mathbf{V}_{T_{in}}(T_{in})\Delta t - \mathbf{G}_{T_{in}} \frac{\Delta t^2}{2} + \right. \\ &\quad \left. + \Gamma(t)\mathbf{B} - \mathbf{S}_{T_{in}}(t) \right]; \quad i = 1, 2, \dots, N\end{aligned}\quad (3)$$

A single image provides the bearing angles of all the point-features in the local frame. In other words, an image taken at time t provides all the vectors $\mathbf{P}_t^i(t)$ up to a scale. Since the data provided by the gyroscopes during the interval (T_{in}, T_{fin}) allow us to build the matrix $C_{T_{in}}^t$, having the vectors $\mathbf{P}_t^i(t)$ up to a scale, allows us to also know the vectors $\mathbf{P}_{T_{in}}^i(t) = C_{T_{in}}^t \mathbf{P}_t^i(t)$ up to a scale. From equation (3) we easily obtain:

$$\begin{aligned}\mathbf{P}_{T_{in}}^i(t) &= \mathbf{P}_{T_{in}}^i(T_{in}) - \mathbf{V}_{T_{in}}(T_{in})\Delta t - \mathbf{G}_{T_{in}} \frac{\Delta t^2}{2} + \\ &\quad + \Gamma(t)\mathbf{B} - \mathbf{S}_{T_{in}}(t); \quad i = 1, 2, \dots, N\end{aligned}\quad (4)$$

We assume that the camera provides n_i images of the same N point features at the consecutive times: $t_1 = T_{in} < t_2 < \dots < t_{n_i} = T_{fin}$. From now on, for the sake of simplicity, we adopt the following notation:

- $\mathbf{P}_j^i \equiv \mathbf{P}_{T_{in}}^i(t_j)$, $i = 1, 2, \dots, N$; $j = 1, 2, \dots, n_i$
- $\mathbf{P}^i \equiv \mathbf{P}_{T_{in}}^i(T_{in})$, $i = 1, 2, \dots, N$
- $\mathbf{V} \equiv \mathbf{V}_{T_{in}}(T_{in})$
- $\mathbf{G} \equiv \mathbf{G}_{T_{in}}$
- $\Gamma_j \equiv \Gamma(t_j)$, $j = 1, 2, \dots, n_i$
- $\mathbf{S}_j \equiv \mathbf{S}_{T_{in}}(t_j)$, $j = 1, 2, \dots, n_i$

Additionally, we will denote with $\boldsymbol{\mu}_j^i$ the unit vector with the same direction of \mathbf{P}_j^i and we introduce the unknowns λ_j^i such that $\mathbf{P}_j^i = \lambda_j^i \boldsymbol{\mu}_j^i$. Finally, without

loss of generality, we can set $T_{in} = 0$, i.e., $\Delta t = t$. Our sensors provide $\boldsymbol{\mu}_j^i$ and \mathbf{S}_j for $i = 1, 2, \dots, N; j = 1, 2, \dots, n_i$. Equation (4) can be written as follows:

$$\mathbf{P}^i - \mathbf{V}t_j - \mathbf{G}\frac{t_j^2}{2} + \Gamma_j\mathbf{B} - \lambda_j^i\boldsymbol{\mu}_j^i = \mathbf{S}_j \quad (5)$$

The Vi-SfM problem is the determination of the vectors: \mathbf{P}^i , ($i = 1, 2, \dots, N$), \mathbf{V} , \mathbf{G} . In the case with biased accelerometer data, we also need to determine the vector \mathbf{B} . We can use the equations in (5) to determine these vectors. On the other hand, the use of (5) requires to also determine the quantities λ_j^i . By considering $j = 1$ in (5), i.e. $t_j = t_1 = T_{in} = 0$, we easily obtain: $\mathbf{P}^i = \lambda_1^i\boldsymbol{\mu}_1^i$. Then, we write the linear system in (5) as follows:

$$\begin{cases} -\mathbf{G}\frac{t_j^2}{2} - \mathbf{V}t_j + \Gamma_j\mathbf{B} + \lambda_1^1\boldsymbol{\mu}_1^1 - \lambda_j^1\boldsymbol{\mu}_j^1 & = \mathbf{S}_j \\ \lambda_1^1\boldsymbol{\mu}_1^1 - \lambda_j^1\boldsymbol{\mu}_j^1 - \lambda_1^i\boldsymbol{\mu}_1^i + \lambda_j^i\boldsymbol{\mu}_j^i & = \mathbf{0}_3 \end{cases} \quad (6)$$

where $j = 2, \dots, n_i$, $i = 2, \dots, N$ and $\mathbf{0}_3$ is the 3×1 zero vector. This linear system consists of $3(n_i - 1)N$ equations in $Nn_i + 6$ unknowns (or $Nn_i + 9$ in the biased case). Let us define the two column vectors \mathbf{X} and \mathbf{S} :

$$\mathbf{X} \equiv [\mathbf{G}^T, \mathbf{V}^T, \mathbf{B}^T, \lambda_1^1, \dots, \lambda_1^N, \dots, \lambda_{n_i}^1, \dots, \lambda_{n_i}^N]^T$$

(or

$$\mathbf{X} \equiv [\mathbf{G}^T, \mathbf{V}^T, \lambda_1^1, \dots, \lambda_1^N, \dots, \lambda_{n_i}^1, \dots, \lambda_{n_i}^N]^T$$

in absence of bias), and

$$\mathbf{S} \equiv [\mathbf{S}_2^T, \mathbf{0}_3, \dots, \mathbf{0}_3, \mathbf{S}_3^T, \mathbf{0}_3, \dots, \mathbf{0}_3, \dots, \mathbf{S}_{n_i}^T, \mathbf{0}_3, \dots, \mathbf{0}_3]^T$$

and the matrix:

$$\Xi \equiv \begin{bmatrix} T_2 & S_2 & \Gamma_2 & \boldsymbol{\mu}_1^1 & \mathbf{0}_3 & \mathbf{0}_3 & -\boldsymbol{\mu}_2^1 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_{33} & \mathbf{0}_{33} & \mathbf{0}_{33} & \boldsymbol{\mu}_1^1 & -\boldsymbol{\mu}_1^2 & \mathbf{0}_3 & -\boldsymbol{\mu}_2^1 & \boldsymbol{\mu}_2^2 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}_{33} & \mathbf{0}_{33} & \mathbf{0}_{33} & \boldsymbol{\mu}_1^1 & \mathbf{0}_3 & -\boldsymbol{\mu}_1^N & -\boldsymbol{\mu}_2^1 & \mathbf{0}_3 & \boldsymbol{\mu}_2^N & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ T_{n_i} & S_{n_i} & \Gamma_{n_i} & \boldsymbol{\mu}_1^1 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & -\boldsymbol{\mu}_{n_i}^1 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_{33} & \mathbf{0}_{33} & \mathbf{0}_{33} & \boldsymbol{\mu}_1^1 & -\boldsymbol{\mu}_1^2 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & -\boldsymbol{\mu}_{n_i}^1 & \boldsymbol{\mu}_{n_i}^2 & \mathbf{0}_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}_{33} & \mathbf{0}_{33} & \mathbf{0}_{33} & \boldsymbol{\mu}_1^1 & \mathbf{0}_3 & -\boldsymbol{\mu}_1^N & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & -\boldsymbol{\mu}_{n_i}^1 & \mathbf{0}_3 & \boldsymbol{\mu}_{n_i}^N \end{bmatrix} \quad (7)$$

where $T_j \equiv -\frac{t_j^2}{2}I_3$, $S_j \equiv -t_jI_3$ and I_3 is the identity 3×3 matrix; $\mathbf{0}_{33}$ is the 3×3 zero matrix (note that the third set of columns disappear in absence of bias). The linear system in (6) can be written in the following compact format:

$$\Xi\mathbf{X} = \mathbf{S} \quad (8)$$

The sensor information is completely contained in the above linear system. Additionally, we assume that the magnitude of the gravitational acceleration is a

priori known. This extra information is obtained by adding to our linear system the following quadratic equation: $|\mathbf{G}| = g$. By introducing the following $3 \times (Nn_i + 6)$ matrix (or $3 \times (Nn_i + 9)$ in the biased case), $\mathbf{\Pi} \equiv [I_3, \mathbf{0}_3 \dots \mathbf{0}_3]$, this quadratic constraint can be written in terms of \mathbf{X} as follows:

$$|\mathbf{\Pi X}|^2 = g^2 \quad (9)$$

The Vi-SfM problem can be solved by finding the vector \mathbf{X} , which satisfies (8) and (9).

4 Existence and number of distinct solutions

We are interested in understanding how the existence and the number of solutions of the Vi-SfM problem depend on the motion, on the number of observed point-features, on the point-features layout and on the number of camera images. The resolvability of the Vi-SfM problem can be investigated by computing the null space of the matrix $\mathbf{\Xi}$ in (7). Let us denote with $\mathcal{N}(\mathbf{\Xi})$ this space. The following theorem holds:

Theorem 1 (Number of Solutions) *The Vi-SfM problem has a unique solution if and only if $\mathcal{N}(\mathbf{\Xi})$ is empty. It has two solutions, if and only if $\mathcal{N}(\mathbf{\Xi})$ has dimension 1 and, for any $\mathbf{n} \in \mathcal{N}(\mathbf{\Xi})$, $|\mathbf{\Pi n}| \neq 0$. It has infinite solutions in all the other cases.*

Proof: : The first part of this theorem is a trivial consequence of the theory of linear systems. Indeed, the vector \mathbf{X} can be uniquely obtained by inverting the matrix $\mathbf{\Xi}$. Let us consider the case when the dimension of $\mathcal{N}(\mathbf{\Xi})$ is 1. The linear system in (8) has infinite solutions with the following structure: $\mathbf{X}(\gamma) = \mathbf{\Xi}^* \mathbf{S} + \gamma \mathbf{n}$, where $\mathbf{\Xi}^*$ is a pseudoinverse of $\mathbf{\Xi}$, \mathbf{n} is a vector belonging to $\mathcal{N}(\mathbf{\Xi})$ and γ is an unknown scalar value [15]. We use (9) to obtain γ . We have: $|\mathbf{\Pi X}(\gamma)|^2 = g^2$, which is a second order polynomial equation in γ if and only if $|\mathbf{\Pi n}| \neq 0$. Hence, we have two solutions for γ , γ_1 and γ_2 , and two solutions for \mathbf{X} , $\mathbf{X}_1 \equiv \mathbf{X}(\gamma_1)$ and $\mathbf{X}_2 \equiv \mathbf{X}(\gamma_2)$. When $|\mathbf{\Pi n}| = 0$ equation $|\mathbf{\Pi X}(\gamma)|^2 = g^2$ is independent of γ . Hence, this equation is automatically satisfied, independently of γ . This means that the Vi-SfM problem has infinite solutions. However, it also means that the vector \mathbf{G} can be uniquely determined. ■

The previous theorem allows us to obtain all the properties of the Vi-SfM problem by investigating the null space of $\mathbf{\Xi}$. The dimension of this null space does not change by multiplying the columns of $\mathbf{\Xi}$ by any value different from zero. Hence, we will refer to the following matrix:

$$\mathbf{\Xi}' \equiv \left[\begin{array}{c|c|c|c|c|c} \mathcal{M}_2 & \mathcal{P}_1 & \mathcal{P}_2 & \mathbf{0}_{3N \ N} & \dots & \mathbf{0}_{3N \ N} \\ \mathcal{M}_3 & \mathcal{P}_1 & \mathbf{0}_{3N \ N} & \mathcal{P}_3 & \dots & \mathbf{0}_{3N \ N} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathcal{M}_{n_i} & \mathcal{P}_1 & \mathbf{0}_{3N \ N} & \dots & \mathbf{0}_{3N \ N} & \mathcal{P}_{n_i} \end{array} \right] \quad (10)$$

where $\mathbf{0}_{3N \ N}$ denotes the $3N \times N$ zero matrix and:

$$\mathcal{M}_j \equiv \left[\begin{array}{c|c|c} T_j & S_j & \Gamma_j \\ \hline 0_{33} & 0_{33} & 0_{33} \\ \hline \dots & \dots & \dots \\ \hline 0_{33} & 0_{33} & 0_{33} \end{array} \right], \quad \mathcal{P}_j \equiv \left[\begin{array}{c|c|c|c|c} \mathbf{P}_j^1 & 0_3 & 0_3 & \dots & 0_3 \\ \hline \mathbf{P}_j^1 & \mathbf{P}_j^2 & 0_3 & \dots & 0_3 \\ \hline \mathbf{P}_j^1 & 0_3 & \mathbf{P}_j^3 & \dots & 0_3 \\ \hline \dots & \dots & \dots & \dots & \dots \\ \hline \mathbf{P}_j^1 & 0_3 & \dots & 0_3 & \mathbf{P}_j^N \end{array} \right]$$

(note that the last three columns in the matrix \mathcal{M}_j disappear in absence of bias). In the following, theorem 1 will be applied by using Ξ' instead of Ξ . We remark that the difference $\mathbf{P}_j^i - \mathbf{P}_1^i$, $i = 1, 2, \dots, N$, $j = 2, \dots, n_i$, is independent of i (see equation (4)). Hence, we will set $\chi_j \equiv \mathbf{P}_j^i - \mathbf{P}_1^i$. This quantity characterizes the motion of the platform.

We will make the following assumption:

Assumption 1 For any $i = 1, 2, \dots, N$, $j = 2, \dots, n_i$, $\mathbf{P}_j^i \neq 0_3$ (or equivalently, $\chi_j \neq -\mathbf{P}_1^i$).

This assumption means that during the platform motion, no point feature can be on the origin of the camera frame. It ensures that no column of Ξ' vanishes.

4.1 Unbiased case

Let us denote a vector belonging to $\mathcal{N}(\Xi')$ as follows:

$$\mathbf{n} \equiv [\boldsymbol{\alpha}^T, \boldsymbol{\nu}^T, n_1^1, \dots, n_1^N, n_2^1, \dots, n_2^N, \dots, n_{n_i}^1, \dots, n_{n_i}^N]^T \quad (11)$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\nu}$ are 3D column vectors. \mathbf{n} must satisfy:

$$\Xi' \mathbf{n} = 0_{3(n_i-1)N} \quad (12)$$

where $0_{3(n_i-1)N}$ is the zero $3(n_i-1)N \times 1$ column vector. We can write this system as follows ($j = 2, \dots, n_i$, $i = 2, \dots, N$):

$$-\frac{t_j^2}{2} \boldsymbol{\alpha} - t_j \boldsymbol{\nu} + (n_1^1 + n_j^1) \mathbf{P}_1^1 + n_j^1 \chi_j = 0_3 \quad (13)$$

$$(n_1^1 + n_j^1) \mathbf{P}_1^1 + (n_1^i + n_j^i) \mathbf{P}_1^i + (n_1^j + n_j^j) \chi_j = 0_3 \quad (14)$$

We start our analysis by investigating two very special cases: the *planar* case and the *linear* case.

4.1.1 Planar case

Let us suppose that all the vectors \mathbf{P}_j^i , $i = 1, \dots, N$, $j = 2, \dots, n_i$, belong to a plane². This means that it exists a frame such that all these vectors have the last component equal to zero. In this new frame the linear system in (12) can be separated in two parts: the former corresponds to the first two lines of (13) and the first two lines of (14) for $j = 2, \dots, n_i$; the latter corresponds to the third line of (13) for $j = 2, \dots, n_i$, which only involves the third component of

²This is equivalent to say that the position of any feature and the position of the platform at any time t_j ($j = 1, \dots, n_i$), are coplanar.

α and ν . Let us denote with Ξ_1^{plane} and Ξ_2^{plane} the matrices which characterize these two systems. Their size is $2(n_i - 1)N \times (Nn_i + 4)$ and $(n_i - 1) \times 2$, respectively. When $n_i \leq 2$, the dimension of $\mathcal{N}(\Xi_1^{plane})$ is at least 4. Hence, from theorem 1, we obtain that a necessary condition in order to have a finite number of solutions (one or two) is that $n_i \geq 3$. The null space of Ξ_2^{plane} has dimension 0 as $n_i \geq 3$. Let us consider the case when $n_i = 3$. The size of Ξ_1^{plane} is $4N \times (3N + 4)$. Hence, in order to have the dimension of $\mathcal{N}(\Xi_1^{plane})$ not larger than 1 it is necessary to have $N \geq 3$. Let us consider the case when $n_i = 4$. The size of Ξ_1^{plane} is $6N \times (4N + 4)$. Hence, in order to have the dimension of $\mathcal{N}(\Xi_1^{plane})$ not larger than 1 it is necessary to have $N \geq 2$. Finally, when $n_i \geq 5$ no necessary condition constrains N .

We summarize the results of this subsection with the following property:

Property 1 (Unbiased: Planar Layout) *When all the observed features and the platform positions are coplanar, a necessary condition to have a finite number of solutions is that $n_i \geq 3$. Specifically, if $n_i = 3$, $N \geq 3$, if $n_i = 4$, $N \geq 2$.*

4.1.2 Linear case

When all the vectors P_j^i , $i = 1, \dots, N$, $j = 2, \dots, n_i$, belong to a line it exists a frame such that all these vectors have the last two components equal to zero. In this new frame the linear system in (12) can be separated in two parts: the former corresponds to the first line of (13) and the first line of (14) for $j = 2, \dots, n_i$; the latter corresponds to the second and third line of (13) for $j = 2, \dots, n_i$, which only involve the last two components of α and ν . Let us denote with Ξ_1^{line} and Ξ_2^{line} the matrices which characterize these two systems. Their size is $(n_i - 1)N \times (Nn_i + 4)$ and $2(n_i - 1) \times 4$, respectively. The null space of Ξ_1^{line} has dimension at least $N + 4$. Hence, the Vi-SfM problem has always infinite solutions. This result is obvious and could be derived in a simpler manner. When the platform motion is on a straight line, any feature belonging to this line provides the same bearing data independently of its distance from the platform.

We summarize the results of this subsection with the following property:

Property 2 (Unbiased: Linear Layout) *When all the observed features and the platform positions are collinear, the Vi-SfM problem has always infinite solutions. Additionally, when the platform motion is on a straight line, it is not possible to determine the distance of all the features belonging to this line even if there are other features outside the line.*

Let us consider now the general 3D case and let us focus our attention on the following subsystem:

$$-\frac{t_j^2}{2}\alpha - t_j\nu = -\chi_j, \quad j = 2, \dots, n_i \quad (15)$$

This linear system can always be solved if $n_i = 3$, independently of the platform motion (i.e., for any set of vectors χ_j). In particular, the equations in (15) for $j = 2, 3$ form a linear square system, which has a unique solution, (α_0, ν_0) .

Additionally, the system in (15) can be solved for any $n_i \geq 4$ if and only if $\chi_j = \nu_0 t_j + \alpha_0 \frac{t_j^2}{2}$. This situation corresponds to a platform motion with constant acceleration α_0 and initial speed ν_0 .

From (13-15) we obtain that the vector in (11) with $\alpha = \alpha_0$, $\nu = \nu_0$, $n_1^1 = \bar{n}_1^1 \equiv -1$, $n_j^1 = \bar{n}_j^1 \equiv 1$, $n_1^i = \bar{n}_1^i \equiv 1$, $n_j^i = \bar{n}_j^i \equiv -1$ ($j = 2, \dots, n_i$, $i = 2, \dots, N$) belongs to $\mathcal{N}(\Xi')$ when $n_i = 3$ and/or when the platform moves with constant acceleration. We will denote this vector with \mathbf{n}_0 :

$$\mathbf{n}_0 \equiv [\alpha_0, \nu_0, \bar{n}_1^1, \dots, \bar{n}_1^{i_1}, \dots, \bar{n}_j^1, \dots, \bar{n}_j^{i_j}]^T \quad (16)$$

We proved the following property:

Property 3 $\mathbf{n}_0 \in \mathcal{N}(\Xi')$ in the following two cases:

- when $n_i = 3$, for any motion;
- when the platform moves with constant acceleration.

In order to apply theorem 1, we need to understand if \mathbf{n}_0 is the only generator of $\mathcal{N}(\Xi')$, i.e., if $\mathcal{N}(\Xi')$ has dimension equal or larger than 1.

4.1.3 $n_i = 2$

The system in (15) has not a unique solution. Specifically, in this case the linear system is characterized by a matrix whose rank is 3. Hence, its null space has dimension 3 and the dimension of $\mathcal{N}(\Xi')$ is at least 3.

4.1.4 $n_i = 3$

From property 3 we know that the dimension of $\mathcal{N}(\Xi')$ is at least 1, independently of the number of features. When $N = 1$, Ξ' is a 6×9 matrix. Hence, the dimension of $\mathcal{N}(\Xi')$ is at least 3. Let us consider the case when $N = 2$. In this case Ξ' is a 12×12 matrix. We have the following property:

Property 4 (Minimal case: $n_i = 3$, $N = 2$) The dimension of $\mathcal{N}(\Xi')$ is 1 if and only if the following two conditions are met:

- (i) for a given j (e.g., for $j = 2$), the three vectors \mathbf{P}_1^1 , \mathbf{P}_1^2 and χ_j span the entire 3D-space;
- (ii) for the other value of j (e.g., for $j = 3$) \mathbf{P}_j^i is not proportional to \mathbf{P}_j^k , $\forall i, k = 1, 2, \dots, N$.

Otherwise, the dimension of $\mathcal{N}(\Xi')$ is larger than 1.

Proof: : If (i) is not true, all the vectors \mathbf{P}_j^i , $i = 1, 2$, $j = 2, 3$, belong to a plane. Since $N = 2$, the dimension of $\mathcal{N}(\Xi')$ is larger than 1 (see property 1). Let us suppose now that the condition (i) is met for $j = 2$. From (14) with $j = 2$ we obtain: $n_1^1 = n_2^2 = -n_2^1 = -n_1^2$. From (13) with $j = 2$ we obtain the same equation in (15) with $n_2^1 \chi_2$ instead of χ_2 . From (14) with $j = 3$ we obtain: $(-n_2^1 + n_3^1) \mathbf{P}_3^1 = (n_2^1 + n_3^2) \mathbf{P}_3^2$. If the condition (ii) is met, we have: $n_3^1 = -n_3^2 = n_2^1$ and from (13) with $j = 3$ we obtain the same equation in (15) with $n_2^1 \chi_3$ instead of χ_3 . In other words, when the condition (ii) is met we

have the same equations as in (15) for $j = 2, 3$, with $n_2^1 \chi_j$ instead of χ_j . As previously mentioned, this system has a unique solution and $\mathcal{N}(\Xi')$ is generated by \mathbf{n}_0 . If the condition (ii) is not met, equation $(-n_2^1 + n_3^1) \mathbf{P}_3^1 = (n_2^1 + n_3^2) \mathbf{P}_3^2$ has further solutions and consequently \mathbf{n}_0 is not the only generator of $\mathcal{N}(\Xi')$ ■
From now on, we will say that a condition is satisfied *in general* when the probability that it is not satisfied is zero. We remark that both conditions (i) and (ii) are met in general.

Also for $N > 2$ there are still conditions, which occur with zero probability, under which the dimension of $\mathcal{N}(\Xi')$ is larger than 1. We summarize the results of this subsection with the following property:

Property 5 (Unbiased with $n_i = 3$, $N \geq 2$) *When $n_i = 3$ and $N \geq 2$ the Vi-SfM problem has in general two distinct solutions. In some special cases it has infinite solutions.*

4.1.5 $n_i \geq 4$

When $n_i \geq 4$ the number of equations is larger than the number of unknowns, except when $n_i = 4$ and $N = 1$. In this case the matrix Ξ' is 9×10 and the dimension of its null space is at least 1. We have the following property:

Property 6 (Minimal case: $n_i = 4$, $N = 1$) *The dimension of $\mathcal{N}(\Xi')$ is 1 if and only if the four vectors \mathbf{P}_1^1 , χ_2 , χ_3 and χ_4 span the entire 3D-space.*

Proof: If the vectors \mathbf{P}_1^1 and χ_j , $j = 2, 3, 4$, are coplanar, since $N = 1$, the dimension of $\mathcal{N}(\Xi')$ is larger than 1 (see property 1). Let us suppose now that the vectors \mathbf{P}_1^1 and χ_j , $j = 2, 3, 4$ span the entire 3D-space. From the first 6 equations in (12) (i.e., the equation (13) for $j = 2, 3$) we obtain α and ν as linear functions of \mathbf{P}_1^1 , χ_2 and χ_3 . By substituting the expressions of α and ν in the last three equations (i.e., in (13) with $j = 4$) we obtain the following equation: $a_1 \mathbf{P}_1^1 + a_2 \chi_2 + a_3 \chi_3 + a_4 \chi_4 = 0_3$, where a_1, a_2, a_3, a_4 are linear expressions of $n_1^1, n_2^1, n_3^1, n_4^1$. Since the four vectors span the entire 3D-space, the null space of the 3×4 matrix $[\mathbf{P}_1^1, \chi_2, \chi_3, \chi_4]$ has dimension 1. Let us denote with $[a_1^*, a_2^*, a_3^*, a_4^*]^T$ a generator of this null space. We consider the linear system $a_k(n_1^1, n_2^1, n_3^1, n_4^1) = a_k^*$, $k = 1, 2, 3, 4$. We analytically compute the determinant of the 4×4 matrix, which characterizes this linear system. We obtain:

$$\frac{(t_4 - t_2)(t_4 - t_3)t_4^2(t_4 t_2 - t_2 t_3 - t_4^2 + t_3 t_4)}{(t_2 - t_3)^2 t_3^2 t_2^2}$$

This determinant is always different from 0 (note that $0 < t_2 < t_3 < t_4$). Hence, the previous linear system provides a unique solution and the dimension of $\mathcal{N}(\Xi')$ is 1 ■

We do not derive necessary and sufficient conditions for any value of n_i and N . We summarize the results with the following property:

Property 7 (Unbiased with $n_i \geq 4$) *When $n_i = 4$ and $N = 1$ the Vi-SfM problem has in general two distinct solutions. If $n_i = 4$, $N \geq 2$ or if $n_i \geq 5$, $\forall N$ it has in general a unique solution.*

4.1.6 Constant acceleration

Let us consider the case when the platform moves with constant acceleration, i.e. when $\chi_j = \nu_0 t_j + \alpha_0 \frac{t_j^2}{2}$, $j = 2, \dots, n_i$, where ν_0 and α_0 are two $3D$ -vectors. We already know from property 3 that the dimension of $\mathcal{N}(\Xi')$ is at least 1. Specifically, the vector \mathbf{n}_0 in (16) belongs to the null space of Ξ' . In order to use theorem 1, we need to understand when $\mathcal{N}(\Xi')$ has dimension equal or larger than 1. The following property provides a sufficient condition which ensures the Vi-SfM resolvability.

Property 8 (Unbiased with constant acceleration) *Let us suppose that the platform moves with constant acceleration, i.e., $\chi_j = \nu_0 t_j + \alpha_0 \frac{t_j^2}{2}$, $j = 2, \dots, n_i$. When for a given feature k the vectors ν_0 , α_0 and \mathbf{P}_1^k span the entire $3D$ -space the dimension of $\mathcal{N}(\Xi')$ is 1.*

Proof: Without loss of generality, let us set $k = 1$. From the first 6 equations in (12) (i.e., the equation (13) for $j = 2, 3$) we obtain α and ν as linear functions of \mathbf{P}_1^1 , α_0 and ν_0 . By substituting the expressions of α and ν in (13) with $j = 4$ we obtain the following equation: $a_1 \mathbf{P}_1^1 + a_2 \alpha_0 + a_3 \nu_0 = 0_3$, where a_1, a_2, a_3 are linear expressions of $n_1^1, n_2^1, n_3^1, n_4^1$. Since the three vectors span the entire $3D$ -space, we must have $a_k(n_1^1, n_2^1, n_3^1, n_4^1) = 0$, $k = 1, 2, 3$. This linear system is characterized by a 3×4 matrix. Hence, it has at least one non trivial solution. By a direct computation, it is possible to see that the dimension of the null space of this matrix is 1. The non trivial solution is $n_1^1 = -1, n_2^1 = n_3^1 = n_4^1 = 1$. Now, let us consider the equation (14). We obtain ($i \neq 1$):

$$(n_1^i + n_j^i) \mathbf{P}_1^i + (1 + n_j^i) (\nu_0 t_j + \alpha_0 \frac{t_j^2}{2}) = 0_3 \quad (17)$$

On the other hand, a further consequence of the fact that ν_0 , α_0 and \mathbf{P}_1^1 span the entire $3D$ -space, is that the two vectors ν_0 and α_0 cannot be collinear. Hence, it exists a value of $j = j^*$, such that \mathbf{P}_1^i is not proportional to $\nu_0 t_{j^*} + \alpha_0 \frac{t_{j^*}^2}{2}$. From (17) we immediately obtain $n_{j^*}^i = -1$ and $n_1^i = 1$. For the other $j \neq j^*$ we obtain: $(1 + n_j^i) (\mathbf{P}_1^i + \nu_0 t_j + \alpha_0 \frac{t_j^2}{2}) = 0_3$. If $n_j^i \neq -1$ $\mathbf{P}_1^i = -\nu_0 t_j - \alpha_0 \frac{t_j^2}{2} = -\chi_j$, which is not possible because of the assumption 1 ■

A special case of constant acceleration occurs when the vector α_0 vanishes, i.e., when the platform moves with constant speed. Since $|\Pi \mathbf{n}_0| = |\alpha_0| = 0$, according to theorem 1, the Vi-SfM has infinite solutions. However, as it has been proven at the end of the proof of that theorem, in this case the local gravity \mathbf{G} can be uniquely determined. Hence, the orientation of the platform with respect to the horizontal plane can be uniquely determined. We proved the following property:

Property 9 (Unbiased with constant speed) *Let us suppose that the platform moves with constant speed. The Vi-SfM has infinite solutions. Additionally, the orientation of the platform with respect to the horizontal plane can be uniquely determined.*

4.2 Biased case

We investigate now the resolvability of the Vi-SfM problem when the accelerometers are affected by a bias. Obviously, all the necessary conditions derived in 4.1 are still necessary in this harder case. On the other hand, there are cases where conditions which ensure resolvability in the unbiased case, are no longer sufficient in this case. By proceeding as in the unbiased case (see (11)), we will denote a vector belonging to $\mathcal{N}(\Xi')$ as follows:

$$\mathbf{n} \equiv [\boldsymbol{\alpha}^T, \boldsymbol{\nu}^T, \mathbf{b}^T, n_1^1, \dots, n_1^N, n_2^1, \dots, n_2^N, \dots, n_{n_i}^1, \dots, n_{n_i}^N]^T \quad (18)$$

where \mathbf{b} is a $3D$ -vectors (as $\boldsymbol{\alpha}$ and $\boldsymbol{\nu}$). \mathbf{n} must satisfy (12) where now Ξ' also includes the third set of columns. We can write this system as in (13-14). In this case, the first equation must be replaced with:

$$-\frac{t_j^2}{2}\boldsymbol{\alpha} - t_j\boldsymbol{\nu} + \Gamma_j\mathbf{b} + (n_1^1 + n_j^1)\mathbf{P}_1^1 + n_j^1\boldsymbol{\chi}_j = \mathbf{0}_3 \quad (19)$$

We start by remarking that, in the unbiased case, the platform rotations do not affect the problem resolvability. Indeed, in the matrix Ξ' , only the third set of columns are affected by the platform rotations. In the biased case we have the following property:

Property 10 (Biased: impact of rotations) *When the platform does not rotate the dimension of $\mathcal{N}(\Xi')$ is at least 3. When the platform rotates always around the same axis the dimension of $\mathcal{N}(\Xi')$ is at least 1.*

Proof: : When the platform does not rotate $\Gamma_j = \frac{t_j^2}{2}I_3$ (see the definition of Γ_j in (2)). Hence, the third set of columns coincides with the first set up to a sign. This means that the dimension of $\mathcal{N}(\Xi')$ is at least 3. Let us consider the case when the rotations only occur around the same axis. We can assume without loss of generality that it is the z -axis (indeed, we can change the camera frame in such a way that its new z -axis is aligned with the axis of rotation). From the definition of Γ in (2) we remark that, in this case, the third column of Γ_j coincides with the third column of T_j in (7) up to a sign. Hence, the vector in (18) with all the entries zero with the exception of the third and the ninth entry equal one each other, belongs to $\mathcal{N}(\Xi')$ ■

Regarding the planar and linear cases, properties 1 and 2 still hold since they only provide necessary conditions. However, regarding the planar case, more restrictive conditions can be derived, which even hold in the $3D$ -case³.

Let us consider the general $3D$ case. By proceeding as in the unbiased case, i.e., by considering the system in (15), we obtain the same property 3 with:

$$\mathbf{n}_0 \equiv [\boldsymbol{\alpha}_0^T, \boldsymbol{\nu}_0^T, 0, 0, 0, \bar{n}_1^1, \dots, \bar{n}_1^{n_i}, \dots, \bar{n}_j^1, \dots, \bar{n}_j^{n_i}]^T \quad (20)$$

Property 3 holds for any motion also in this case. Let us focus now our attention on the following subsystem:

$$-\frac{t_j^2}{2}\boldsymbol{\alpha} - t_j\boldsymbol{\nu} + \Gamma_j\mathbf{b} = -\boldsymbol{\chi}_j, \quad j = 2, \dots, n_i \quad (21)$$

³Note that it is not possible to proceed as in the unbiased case since it is not possible to separate the linear system in (12) in two parts because of the bias.

Let us assume that the platform accomplishes rotations in such a way that this linear system can be uniquely solved by using $j = 2, 3, 4$ (according with property 10, this means that the platform must rotate at least around two independent axes). The solution will be denoted with $(\boldsymbol{\alpha}_0, \boldsymbol{\nu}_0, \mathbf{b}_0)$. The system in (21) can be solved for any n_i if and only if $\boldsymbol{\chi}_j = \boldsymbol{\nu}_0 t_j + \boldsymbol{\alpha}_0 \frac{t_j^2}{2} - \Gamma_j \mathbf{b}_0$ for $j \geq 5$. Hence, we also have the following vector:

$$\mathbf{n}_0^b \equiv [\boldsymbol{\alpha}_0, \boldsymbol{\nu}_0, \mathbf{b}_0, \bar{n}_1^1, \dots, \bar{n}_1^{n_i}, \dots, \bar{n}_j^1, \dots, \bar{n}_j^{n_i}]^T, \quad (22)$$

which belongs to the null space of Ξ' under the conditions stated in the following property:

Property 11 *The vector \mathbf{n}_0^b belongs to the null space of Ξ' in the following cases:*

- when $n_i = 4$ and the linear system in (21) can be uniquely solved (by using $j = 2, 3, 4$);
- when the platform moves with the motion characterized by $\boldsymbol{\chi}_j = \boldsymbol{\nu}_0 t_j + \boldsymbol{\alpha}_0 \frac{t_j^2}{2} - \Gamma_j \mathbf{b}_0$.

On the other hand, when the linear system in (21) cannot be uniquely solved, its matrix has a non trivial null space. In other words, it is possible to find at least three vectors, $\mathbf{n}_\alpha, \mathbf{n}_\nu, \mathbf{n}_b$ such that:

$$-\frac{t_j^2}{2} \mathbf{n}_\alpha - t_j \mathbf{n}_\nu + \Gamma_j \mathbf{n}_b = 0_3, \quad j = 2, \dots, n_i$$

and the vector

$$\tilde{\mathbf{n}}_0^b \equiv [\mathbf{n}_\alpha, \mathbf{n}_\nu, \mathbf{n}_b, 0, \dots, 0]^T \quad (23)$$

belongs to $\mathcal{N}(\Xi')$. We have the further property:

Property 12 *When the linear system in (21) cannot be uniquely solved, the dimension of $\mathcal{N}(\Xi')$ is at least 1 and $\tilde{\mathbf{n}}_0^b \in \mathcal{N}(\Xi')$.*

4.2.1 $n_i \leq 3$

The system in (21) cannot be uniquely solved. When $n_i = 2$ the dimension of the null space of the matrix which characterizes this system is 6 and when $n_i = 3$ this dimension is 3. By using theorem 1 we obtain the following property:

Property 13 (Biased case, $n_i \leq 3$) *The Vi-SfM problem has always infinite solutions in the biased case when $n_i \leq 3$.*

4.2.2 $n_i = 4$

By using properties 11 and 12 either $\mathbf{n}_0^b \in \mathcal{N}(\Xi')$ or $\tilde{\mathbf{n}}_0^b \in \mathcal{N}(\Xi')$. When $N = 1$ the number of unknowns in (8) is 13 and the number of equations is 9. When $N = 2$ the number of unknowns is 17 and the equations are 18. Hence, we have the following property:

Cases	Number of Solutions
Varying Acceleration $n_i = 4, N \geq 2 ; n_i \geq 5, \forall N$	Unique Solution
Varying Acceleration $n_i = 3, N \geq 2; n_i = 4, N = 1$	Two Solutions
Constant and non null Acceleration $n_i = 3, N \geq 2; n_i \geq 4, \forall N$	Two Solutions
Null Acceleration $\forall n_i, \forall N$	Infinite Solutions
Any Motion $n_i \leq 2, \forall N; n_i = 3, N = 1$	Infinite Solutions

Table 1: Number of distinct solutions for the Vi-SfM problem in the unbiased case

Property 14 (Biased case, $n_i = 4$) *In the biased case, when $n_i = 4$ the Vi-SfM problem has always infinite solutions if $N = 1$ and in general two distinct solutions if $N \geq 2$.*

4.2.3 $n_i \geq 5$

When $n_i = 5$ and $N = 1$ the number of unknowns in (8) is 14 and the number of equations is 12. When $n_i = 5$ and $N = 2$ the number of unknowns is 19 and the equations are 24. When $n_i = 6$ the number of unknowns is $6N + 9$ and the equations are $15N$. Hence, we have the following property:

Property 15 (Biased case, $n_i \geq 5$) *In the biased case, when $n_i = 5$ the Vi-SfM problem has always infinite solutions if $N = 1$ and in general a unique solution if $N \geq 2$. When $n_i \geq 6$ the Vi-SfM problem has in general a unique solution for any N .*

5 Discussion

Tables 1 and 2 summarize our results by providing the number of solutions case by case, respectively in the case without bias (table 1) and with bias (table 2). Note that these tables do not account the features layout. Specifically, the motion and the features are not supposed to be neither coplanar nor collinear. Regarding these cases, necessary conditions are provided in properties 1 and 2.

Figures 1 and 2 illustrate two cases when the Vi-SfM problem has two distinct solutions. The platform configurations and the position of the point-features in the global frame are shown. The two solutions are in blue and red. The platform configuration at the initial time is the same for both the solutions and it is in black. Both figures show unbiased cases. Fig 1 regards the case of one feature seen in four images. Fig 2 displays the case of constant acceleration: the case of three point-features in seven images has been considered and the seven poses of the platform at the time when the images are taken are shown in the figure together with the position of the features.

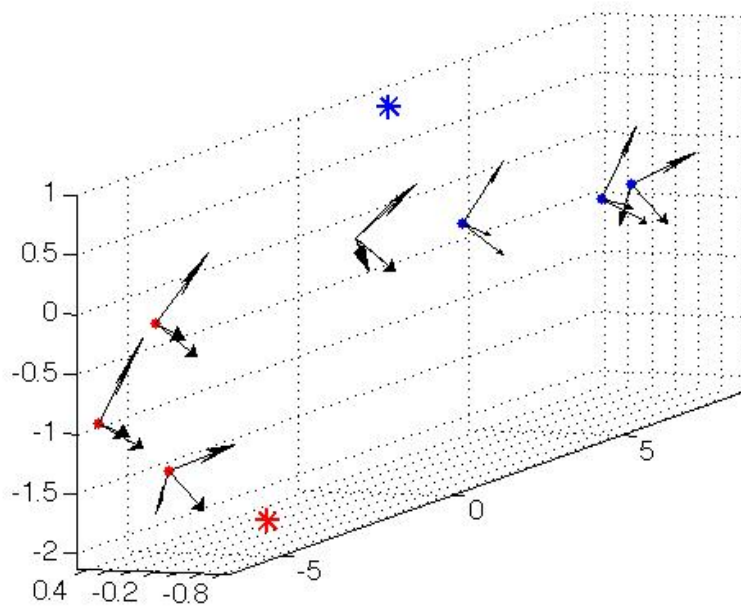


Figure 2: Illustration of two distinct solutions for the unbiased case with $n_i = 4$, $N = 1$ (star symbols indicate the position of the point-feature respectively for the two solutions).

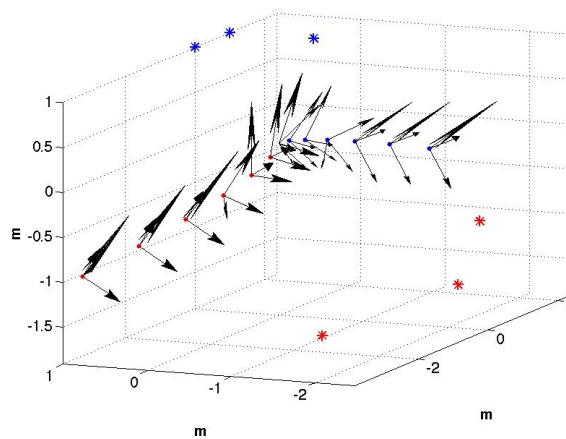


Figure 3: Illustration of two distinct solutions for the unbiased case with constant acceleration (star symbols indicate the position of the point-features respectively for the two solutions).

Cases	Number of Solutions
Rotation around more than 2 axes Varying Acceleration $n_i = 5, N \geq 2 ; n_i \geq 6, \forall N$	Unique Solution
Rotation around a single axis Varying Acceleration $n_i = 5, N \geq 2 ; n_i \geq 6, \forall N$	Two Solutions
Rotation around 1 or 2 axes Varying Acceleration $n_i = 4, N \geq 2$	Two Solutions
Rotation around more than 2 axes Constant and non null Acceleration $n_i = 4, 5, N \geq 2; n_i \geq 6, \forall N$	Two Solutions
Rotation around a single axis Constant and non null Acceleration	Infinite Solutions
No rotation $\forall n_i, \forall N$	Infinite Solutions
Null Acceleration $\forall n_i, \forall N$	Infinite Solutions
Any Motion $n_i \leq 3, \forall N; n_i = 4, 5, N = 1$	Infinite Solutions

Table 2: Number of distinct solutions for the Vi-SfM problem in the biased case

6 Conclusion

In this paper we derived a simple and intuitive closed solution to the visual-inertial structure from motion problem. We used this derivation to investigate the intrinsic properties of the Vi-SfM problem and to identify the conditions under which the problem can be solved in closed form. In particular, we showed that the problem can have a unique solution or two distinct solutions or infinite solutions depending on the trajectory, on the number of point-features and their layout and on the number of monocular images where the same point-features are seen. The investigation was also performed in the case when the inertial data are biased, showing that, in this latter case, more images and more restrictive conditions on the trajectory are required in order to have a finite number of solutions. The most useful applications of the closed-form solution here derived will be in all the applicative domains which need to solve the SfM problem with low-cost sensors and which do not demand any infrastructure (e.g., in GPS denied environment). In these contexts, there is often the need to perform the estimation without any prior knowledge. Typical examples of applicative domains are the emergent fields of humanoid robotics and unmanned aerial navigation in urban-like environments [18], where the use of the GPS is often forbidden. Additionally, our results could also play an important role in the framework of neuroscience by providing a new insight on the process of vestibular and visual integration for depth perception and self-motion perception. The influence of extra retinal cues in depth perception has extensively been investigated in the last decades. In particular, a very recent study investigates this problem by per-

forming trials with passive head movements [5]. The conclusion of this study is that the combination of retinal image with vestibular signals can provide rudimentary ability to depth perception. Our findings could provide a new insight to this integration mechanism for depth perception since, according to the closed-solution here derived, by combining retinal image with vestibular signals it is possible to determine the scale factor even without any knowledge about the initial speed. Our findings also show that it is possible to easily distinguish linear acceleration from gravity. Specifically, the closed form solution performs this determination by a very simple matrix inversion. This problem has also been investigated in neuroscience [14]. Our results could provide a new insight to this mechanism since they clearly characterize the conditions (type of motion, features layout) under which this determination can be performed.

References

- [1] A. Berthoz, B. Pavard and L.R. Young, Perception of Linear Horizontal Self-Motion Induced by Peripheral Vision (Linearvection) Basic Characteristics and Visual-Vestibular Interactions, *Exp. Brain Res.* 23, 471–489 (1975).
- [2] M. Bryson and S. Sukkarieh, Observability Analysis and Active Control for Airbone SLAM, *IEEE Transaction on Aerospace and Electronic Systems*, vol. 44, no. 1, 261–280, 2008
- [3] Alessandro Chiuso, Paolo Favaro, Hailin Jin and Stefano Soatto, "Structure from Motion Causally Integrated Over Time", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4), pp 523–535, 2002
- [4] Andrew J. Davison, Ian D. Reid, Nicholas D. Molton and Olivier Stasse, "MonoSLAM: Real-Time Single Camera SLAM", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), pp 1052–1067, 2007
- [5] Dokka K., MacNeilage P. R., De Angelis G. C. and Angelaki D. E., Estimating distance during self-motion: a role for visual-vestibular interactions, *Journal of Vision* (2011) 11(13):2, 1-16
- [6] J. A. Farrell, *Aided Navigation: GPS and High Rate Sensors*. McGraw- Hill, 2008.
- [7] C. R. Fetsch, G. C. DeAngelis and D. E. Angelaki, Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory, *Eur J Neurosci.* 2010 May ; 31(10): 1721-1729
- [8] Richard I. Hartley (June 1997). "In Defense of the Eight-Point Algorithm". *IEEE Transaction on Pattern Recognition and Machine Intelligence* 19 (6): 580–593.
- [9] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, Towards consistent vision-aided inertial navigation, in *Proc. of the Int. Workshop on the Algorithmic Foundations of Robotics*, Cambridge, MA, Jun. 1315, 2012.
- [10] E. Jones and S. Soatto, "Visual-inertial navigation, mapping and localization: A scalable real-time causal approach", *The International Journal of Robotics Research*, vol. 30, no. 4, pp. 407–430, Apr. 2011.
- [11] J. Kelly and G. Sukhatme, Visual-inertial simultaneous localization, mapping and sensor-to-sensor self-calibration, *Int. Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.
- [12] H. Christopher Longuet-Higgins (September 1981). "A computer algorithm for reconstructing a scene from two projections". *Nature* 293: 133–135.
- [13] A. Martinelli, Vision and IMU Data Fusion: Closed-Form Solutions for Attitude, Speed, Absolute Scale and Bias Determination, *Transaction on Robotics*, Volume 28 (2012), Issue 1 (February), pp 44–60.
- [14] Merfeld D. M., Zupan L. and Peterka R. J., Humans use internal models to estimate gravity and linear acceleration, *Nature*, 398, pp 615–618, 1999

- [15] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000
- [16] D. Nistér, An efficient solution to the five-point relative pose problem, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 26(6):756-770, June 2004
- [17] D. Strelow and S. Singh, Motion estimation from image and inertial measurements, *International Journal of Robotics Research*, 23(12), 2004
- [18] Weiss., S., Scaramuzza, D., Siegwart, R., Monocular-SLAM-Based Navigation for Autonomous Micro Helicopters in GPS-Denied Environments, *Journal of Field Robotics*, Volume 28, issue 6, 2011

Contents

1	Introduction	3
2	The Considered System	4
3	The closed form solution	4
4	Existence and number of distinct solutions	8
4.1	Unbiased case	9
4.1.1	Planar case	9
4.1.2	Linear case	10
4.1.3	$n_i = 2$	11
4.1.4	$n_i = 3$	11
4.1.5	$n_i \geq 4$	12
4.1.6	Constant acceleration	13
4.2	Biased case	14
4.2.1	$n_i \leq 3$	15
4.2.2	$n_i = 4$	15
4.2.3	$n_i \geq 5$	16
5	Discussion	16
6	Conclusion	18



Centre de recherche INRIA Grenoble – Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399