

Article

Rotation Estimation: A Closed-Form Solution Using Spherical Moments [†]

Hicham Hadj-Abdelkader ^{1,*}, Omar Tahri ^{2,*}  and Housseem-Eddine Benseddik ³

¹ IBISC lab. EA 4526, University of Evry-Val-d'Essonne-Paris Saclay, 91000 Evry, France

² INSA Centre Val de Loire, PRISME lab. EA 4229, University of Orleans, 18000 Bourges, France

³ MIS lab. EA 4290, University of Picardie Jules Verne, 80039 Amiens, France; houssem.eln@gmail.com

* Correspondence: hicham.hadjabdelkader@univ-evry.fr (H.H.-A.); omar.tahri@insa-cvl.fr (O.T.)

[†] This paper is an extended version of our paper published in Hadj-Abdelkader, H.; Tahri, O.; Benseddik, H.E. Closed form solution for Rotation Estimation using Photometric Spherical Moments. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 627–634.

Received: 30 September 2019; Accepted: 24 October 2019; Published: 14 November 2019

Abstract: Photometric moments are global descriptors of an image that can be used to recover motion information. This paper uses spherical photometric moments for a closed form estimation of 3D rotations from images. Since the used descriptors are global and not of the geometrical kind, they allow to avoid image processing as features extraction, matching, and tracking. The proposed scheme based on spherical projection can be used for the different vision sensors obeying the central unified model: conventional, fisheye, and catadioptric. Experimental results using both synthetic data and real images in different scenarios are provided to show the efficiency of the proposed method.

Keywords: geometric moments; spherical image; motion estimation

1. Introduction

Rotation estimation from images is important for many application as for motion estimation and registration in image processing [1], pattern recognition [2], localization and control of ground/aerial vehicles [3], and computer vision [4]. Feature based and direct (appearance based) are the two main approach categories proposed in the literature. The first method type uses Epipolar geometry applied to geometrical features as as points [5], lines [6], or contours [7] to estimate the fundamental matrix (uncalibrated vision sensors), the essential matrix, or homography matrix for the calibrated case. Examples of such methods using conventional cameras are Agrawal et al. [8] and Malis et al. [9] and using omnidirectional vision sensors is [10–12]. The direct approaches, on the other hand, avoid geometric features extraction by using grey level information of the whole image to define the descriptors to be used. Those methods are referred to in the literature as dense, direct, or global. Methods within this category can avoid the need of some higher-level image processing steps such as feature extraction, matching, and tracking. For instance, Reference [13] exploits optical flow estimation. However, a consequence of using optical flow in References [13,14] is to be limited to small camera displacements. To deal with larger motion, in Reference [4], Makadia and Daniilidis proposed a method using correlations between the two sets of spherical harmonic coefficients.

To address even larger camera displacements, motion estimation techniques based on global image representation is utilized in the literature. Makadia and Daniilidis [4] represented the spherical omnidirectional images in the frequency domain to recover rotation using the maximum correlation between the two sets of spherical harmonic coefficients. In Reference [15], the authors proposed to estimate the rotation based on 3D-mesh representation of spherical image intensities. However,

these methods are not often computationally efficient due to the computation load of spherical harmonic coefficients.

The moments as descriptors provide useful characteristics such as the center of gravity of an object, orientation, disparity, and volume. As Fourier descriptors, they can be used as a compact representation of the image. In the case of 2D images, in addition to the classical geometrical moments, other kinds of moments have been defined in the literature as Zernike Moments [16] or complex moments [17]. Invariant characteristics to some transformations that an object can undergo, commonly called moment invariants, found many applications such as for pattern recognition. Thanks to their decoupling properties, they also have been used to define a minimal number of features for visual servoing and pose estimation as in References [18,19] and for visual servoing as in References [20,21].

Spherical projections can be obtained from images if the used camera obeys the unified central model. By using such projections, the authors in Reference [22] proposed a decoupled visual servoing scheme using spherical moments computed for a set of matched points. The proposed features allow to control the translational motions independently from the rotational ones. The current work exploits the spherical projection to estimate 3D rotations using spherical moments. The proposed scheme provides an analytical solution to rotation. An early version of this work was published in Reference [23]. In the next section, the central camera model and the definition of spherical moments are first recalled. The proposed approach is subsequently presented. In Section 3, experimental results of rotation estimation using synthetic data and real images acquired from a catadioptric camera mounted on a manipulator robot (Staubli) and a mobile robot (Pioneer) are presented to validate our approach.

2. Direct Estimation of Rotation Using Spherical Moments

2.1. Spherical Moments

The projection onto a unitary sphere $\mathbf{P}_s = [x_s \ y_s \ z_s]^\top$ of a 3D point \mathbf{P} is defined by the following:

$$\mathbf{P}_s = [x_s \ y_s \ z_s]^\top = \frac{\mathbf{P}}{\|\mathbf{P}\|}. \quad (1)$$

Actually, the projection onto the unit sphere can be obtained using the unified central model for conventional, fisheye, or catadioptric [24]. This makes the proposed scheme in this work valid for a wide range of cameras.

From a spherical image (or Gaussian sphere [25]) of the environment, the photometric moments can be defined by the following:

$$m_{ijk} = \iint_S x_s^i y_s^j z_s^k I(x_s, y_s, z_s, t) ds, \quad (2)$$

where S is defined as the sphere surface. Let \mathbf{R} be a rotation matrix and $\boldsymbol{\omega}$ be the instantaneous rotational speed applied to the sensor frame. First, it can be easily shown that a rotational motion defined by a matrix \mathbf{R} on 3D point \mathbf{P} ($\mathbf{P}' = \mathbf{R}\mathbf{P}$, where \mathbf{P}' is the same 3D point after rotation) induces the same motion on its projection onto the unit sphere ($\mathbf{P}'_s = \mathbf{R}\mathbf{P}_s$). In tangent space, the effect of a rotation speed $\boldsymbol{\omega} = [\omega_x \ \omega_y \ \omega_z]^\top$ on a projected point on the sphere is given by its time derivative $\dot{\mathbf{P}}_s$:

$$\dot{\mathbf{P}}_s = [\mathbf{P}_s]_\times \boldsymbol{\omega} = -[\boldsymbol{\omega}]_\times \mathbf{P}_s = \mathbf{L}_{\mathbf{P}_s} \boldsymbol{\omega} = \begin{bmatrix} 0 & -z_s & y_s \\ z_s & 0 & -x_s \\ -y_s & x_s & 0 \end{bmatrix} \boldsymbol{\omega}, \quad (3)$$

where \mathbf{L}_p is the interaction matrix and $[\]_\times$ defines the skew-symmetric matrix of a vector. On the other side, by taking the time derivative of Equation (2), the time variation of the moment m_{ijk} is linked to rotational speed ω by the following [26]:

$$\dot{m}_{ijk} = \begin{bmatrix} jm_{i,j-1,k+1} - km_{i,j+1,k-1} \\ km_{i+1,j,k-1} - im_{i-1,j,k+1} \\ im_{i-1,j+1,k} - jm_{i+1,j-1,k} \end{bmatrix}^\top \omega. \quad (4)$$

In the following, a scheme is proposed to define triples $\mathbf{x}_m = [x(m_{ijk}) \ y(m_{ijk}) \ z(m_{ijk})]^\top \in \mathbb{R}^{3 \times 1}$ that behave as a projected point onto the sphere with respect to rotations.

2.2. Closed-Form Solution of Rotation Estimation

From moments of order 1 and by using Equation (4), the interaction matrix that links time derivatives of the triple $\mathbf{x}_{m_1} = [m_{100} \ m_{010} \ m_{001}]^\top$ to the rotational velocities is obtained by the following:

$$\dot{\mathbf{x}}_{m_1} = \begin{bmatrix} 0 & -m_{001} & m_{010} \\ m_{001} & 0 & -m_{100} \\ -m_{010} & m_{100} & 0 \end{bmatrix} \omega = [\mathbf{x}_{m_1}]_\times \omega. \quad (5)$$

As it can be seen, the interaction matrix given by Equation (5) has an identical form to the one obtained for a projected point onto the sphere. This implies a 3D rotation results in the same motion on \mathbf{x}_{m_1} . Let us now consider triples \mathbf{x}_m defined from moments of orders larger than 1. In order to have $\mathbf{x}'_m = R \mathbf{x}_m$, the interaction matrix related to $\mathbf{x}_m = [x(m_{ijk}), y(m_{ijk}), z(m_{ijk})]^\top$ must have the same form as Equation (3), which means the following:

$$\dot{\mathbf{x}}_m = \begin{bmatrix} 0 & -z(m_{ijk}) & y(m_{ijk}) \\ z(m_{ijk}) & 0 & -x(m_{ijk}) \\ -y(m_{ijk}) & x(m_{ijk}) & 0 \end{bmatrix} \omega = -[\omega]_\times \mathbf{x}_m \quad (6)$$

Under the assumption of constant speed (i.e., $[\omega]_\times$ is a constant matrix), solving for Equation (6) leads to the following:

$$\mathbf{x}_m(t) = \exp(-[\omega]_\times t) \mathbf{x}_m(0) = \mathbf{R} \mathbf{x}_m(0) \quad (7)$$

where $\mathbf{x}_m(0)$ and $\mathbf{x}_m(t)$ are respectively the initial value and at the value time t of the triple (which defines the relation for a couple of images). Any rotational motion between two images can be obtained by applying a constant speed during time t . This means Equation (6) implies $\mathbf{x}_m(t) = \mathbf{R} \mathbf{x}_m(0)$ for any couple of images.

Let us now explain the procedure to obtain those triples using as an example: the moment vector $\mathbf{w}_{23} = [m_{300}m_{200} \ m_{300}m_{110} \ \cdots \ m_{003}m_{002}]^\top$. The latter is composed of all the possible products between moments of order 2 and those of order 3. There are 6 possible moments of order equal to 2 and 10 of order equal to 3. This implies that \mathbf{w}_{23} is of size equal to 60. Let us define triples $\mathbf{x}_{w_{23}} = [\alpha_x^\top \mathbf{w}_{23} \ \alpha_y^\top \mathbf{w}_{23} \ \alpha_z^\top \mathbf{w}_{23}]^\top$, where α_x , α_y , and α_z are vectors of coefficients that define the triple coordinates as linear combinations of the entries of \mathbf{w}_{23} . Our goal is to find the conditions on those vectors of coefficients such that the time derivative $\dot{\mathbf{x}}_{w_{23}}$ is of the same form as Equation (6).

As an example, let us consider the time derivative of the moment product $m_{030}m_{200}$, which is a component of \mathbf{w}_{23} . Using Equation (4), we have the following:

$$\begin{cases} \dot{m}_{200} = 0 \times \omega_x - 2m_{101}\omega_y + 2m_{110}\omega_z \\ \dot{m}_{030} = 3m_{021}\omega_x + 0 \times \omega_y - 3m_{120}\omega_z \end{cases}. \quad (8)$$

As it can be also concluded from Equation (4), the time variation of a moment is a function of moments of the same order and the rotational speed. From Equation (8), we obtain the time derivative of the product $m_{030}m_{200}$ by the following:

$$\begin{aligned} \frac{d(m_{030}m_{200})}{dt} &= m_{200}\dot{m}_{030} + m_{030}\dot{m}_{200} \\ &= 3m_{200}m_{021}\omega_x - 2m_{101}m_{030}\omega_y + (2m_{110}m_{030} - 3m_{200}m_{120})\omega_z. \end{aligned} \quad (9)$$

Note that the coefficients of ω_x , ω_y , and ω_z are nothing but linear combinations of products between moments of orders 2 and 3. Actually, the time derivative of any component of the vector \mathbf{w}_{23} will be a linear combination of the other components. Based on this fact, $\dot{\mathbf{w}}_{23}$ can be written under the following form:

$$\dot{\mathbf{w}}_{23} = (\mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w}_{23})\omega_x + (\mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w}_{23})\omega_y + (\mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w}_{23})\omega_z, \quad (10)$$

where $\mathbf{L}_{\mathbf{w}_{23}/\omega_x}$, $\mathbf{L}_{\mathbf{w}_{23}/\omega_y}$ and $\mathbf{L}_{\mathbf{w}_{23}/\omega_z}$ are 60×60 matrices. Using Equation (10), the time derivative of the triple $\mathbf{x}_{\mathbf{w}_{23}} = [\alpha_x^\top \mathbf{w}_{23} \quad \alpha_y^\top \mathbf{w}_{23} \quad \alpha_z^\top \mathbf{w}_{23}]^\top$ is obtained:

$$\dot{\mathbf{x}}_{\mathbf{w}_{23}} = \begin{bmatrix} \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w}_{23} & \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w}_{23} & \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w}_{23} \\ \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w}_{23} & \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w}_{23} & \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w}_{23} \\ \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w}_{23} & \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w}_{23} & \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w}_{23} \end{bmatrix} \omega. \quad (11)$$

If Equation (11) is of the same form as Equation (6), this leads to the 9 following conditions:

$$\begin{aligned} 1) \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w}_{23} &= 0; & 6) \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w}_{23} &= -\alpha_x^\top \mathbf{w}_{23}; \\ 2) \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w}_{23} &= -\alpha_z^\top \mathbf{w}_{23}; & 7) \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w}_{23} &= -\alpha_y^\top \mathbf{w}_{23}; \\ 3) \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w}_{23} &= \alpha_y^\top \mathbf{w}_{23}; & 8) \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w}_{23} &= \alpha_x^\top \mathbf{w}_{23}; \\ 4) \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w}_{23} &= \alpha_z^\top \mathbf{w}_{23}; & 9) \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w}_{23} &= 0. \\ 5) \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w}_{23} &= 0; \end{aligned} \quad (12)$$

The 9 constraints given by Equation (12) must be valid for any value of \mathbf{w}_{23} . Therefore, they can be simplified as follows:

$$\begin{aligned} 1) \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} &= \mathbf{0} & 4) \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} &= \alpha_z^\top & 7) \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_x} &= -\alpha_y^\top \\ 2) \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} &= -\alpha_z^\top & 5) \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} &= \mathbf{0} & 8) \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_y} &= \alpha_x^\top \\ 3) \alpha_x^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} &= \alpha_y^\top & 6) \alpha_y^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} &= -\alpha_x^\top & 9) \alpha_z^\top \mathbf{L}_{\mathbf{w}_{23}/\omega_z} &= \mathbf{0}. \end{aligned} \quad (13)$$

The constraint set of Equation (13) can be written under the following form:

$$\begin{bmatrix} \mathbf{L}_{\mathbf{w}_{23}/\omega_x}^\top & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{\mathbf{w}_{23}/\omega_y}^\top & \mathbf{0} & \mathbf{I} \\ \mathbf{L}_{\mathbf{w}_{23}/\omega_z}^\top & -\mathbf{I} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{L}_{\mathbf{w}_{23}/\omega_z}^\top \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \\ \alpha_z \end{bmatrix} = \mathbf{0}, \quad (14)$$

where \mathbf{I} and $\mathbf{0}$ are respectively the identity and zero matrices of size 60×60 . In practice, $\mathbf{L}_{\mathbf{w}_{23}/\omega_x}$, $\mathbf{L}_{\mathbf{w}_{23}/\omega_y}$ and $\mathbf{L}_{\mathbf{w}_{23}/\omega_z}$ are very sparse integer matrices. Solving Equation (14) for the case of the moment vector \mathbf{w}_{23} leads to three solutions for the triple $\mathbf{x}_{\mathbf{w}_{23}}$ (refer to Equations (A1)–(A3) in the Appendix A). Such conditions can be solved using a symbolic computation software as Matlab Symbolic Toolbox.

The procedure used to obtain triples from \mathbf{w}_{23} can be extended straightforwardly to a different moment vector as follows:

- Define the moment vector \mathbf{w} similarly as for \mathbf{w}_{23} . The vector can be built by moment products of two different orders or more. For instance, products such as $m_{200}m_{220}m_{300}$ which combine moments of order 2, 3, and 4 can be also used. the vector \mathbf{w} has to be built from moment products of the same nature and has to also include all of them.
- Using Equation (4), compute the matrices $\mathbf{L}_{\mathbf{w}_{23}/\omega_x}$, $\mathbf{L}_{\mathbf{w}_{23}/\omega_y}$, and $\mathbf{L}_{\mathbf{w}_{23}/\omega_z}$ such that we obtain the following:

$$\dot{\mathbf{w}} = (\mathbf{L}_{\mathbf{w}_{23}/\omega_x} \mathbf{w})\omega_x + (\mathbf{L}_{\mathbf{w}_{23}/\omega_y} \mathbf{w})\omega_y + (\mathbf{L}_{\mathbf{w}_{23}/\omega_z} \mathbf{w})\omega_z. \quad (15)$$

- Solve the following system:

$$\begin{bmatrix} \mathbf{L}_{\mathbf{w}_{23}/\omega_x}^\top & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{\mathbf{w}_{23}/\omega_y}^\top & \mathbf{0} & \mathbf{I} \\ \mathbf{L}_{\mathbf{w}_{23}/\omega_z}^\top & -\mathbf{I} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{L}_{\mathbf{w}_{23}/\omega_z}^\top \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \\ \alpha_z \end{bmatrix} = \mathbf{0}, \quad (16)$$

where \mathbf{I} and $\mathbf{0}$ are respectively the identity and zero matrices of the same size as $\mathbf{L}_{\mathbf{w}_{23}/\omega_x}$, $\mathbf{L}_{\mathbf{w}_{23}/\omega_y}$ and $\mathbf{L}_{\mathbf{w}_{23}/\omega_z}$.

In practice, two different triples \mathbf{x}_w are enough to obtain a closed-form solution of the rotational motion since they can define an orthonormal basis. Let us consider, for instance, the two triples \mathbf{P}_1 and \mathbf{P}_2 computed from the moments of a first image using Equations (A1) and (A2) and \mathbf{P}'_1 and \mathbf{P}'_2 , their corresponding ones, computed from the image after rotation. Therefore, we have $\mathbf{P}'_1 = \mathbf{R}\mathbf{P}_1$ and $\mathbf{P}'_2 = \mathbf{R}\mathbf{P}_2$, where \mathbf{R} is the rotation matrix. Let us define the following base:

$$\mathbf{v}_1 = \frac{\mathbf{P}_{n1} + \mathbf{P}_{n2}}{\|\mathbf{P}_{n1} + \mathbf{P}_{n2}\|}, \quad \mathbf{v}_2 = \frac{\mathbf{P}_{n1} - \mathbf{P}_{n2}}{\|\mathbf{P}_{n1} - \mathbf{P}_{n2}\|}, \quad \mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2 \quad (17)$$

where $\mathbf{P}_{n1} = \frac{\mathbf{P}_1}{\|\mathbf{P}_1\|}$ and $\mathbf{P}_{n2} = \frac{\mathbf{P}_2}{\|\mathbf{P}_2\|}$. First, it can be shown that the vectors \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 form an orthonormal base (i.e., $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 1$ and $\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_3 = \mathbf{v}_2 \cdot \mathbf{v}_3 = 0$). Second, let us now consider a second base \mathbf{v}'_1 , \mathbf{v}'_2 and \mathbf{v}'_3 computed from \mathbf{P}'_1 and \mathbf{P}'_2 using Equation (17). Since $\mathbf{P}'_1 = \mathbf{R}\mathbf{P}_1$ and $\mathbf{P}'_2 = \mathbf{R}\mathbf{P}_2$, it can be proven using Equation (17) that $\mathbf{v}'_1 = \mathbf{R}\mathbf{v}_1$, $\mathbf{v}'_2 = \mathbf{R}\mathbf{v}_2$ and $\mathbf{v}'_3 = \mathbf{R}\mathbf{v}_3$. The last equations can be written under the following form:

$$\mathbf{V} = \mathbf{R}\mathbf{V}', \quad (18)$$

where $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$ and $\mathbf{V}' = [\mathbf{v}'_1 \ \mathbf{v}'_2 \ \mathbf{v}'_3]$. Finally, we obtain:

$$\mathbf{R} = \mathbf{V} \mathbf{V}'^\top, \quad (19)$$

where the rotation matrix can be estimated straightforwardly. Solving Equation (16) ensures that all possible triples are obtained from a given moment vector. Obtaining a mathematical proof on the existence and their number is out of the the scope of this contribution.

2.3. Rotation Estimation and Scene Symmetry

As explained in the previous paragraph, an orthonormal base can be built from the triples \mathbf{P}_1 and \mathbf{P}_2 to estimate the rotation matrix. In practice, if the scene has no symmetry, extensive tests (not included in the paper for brevity) show that using the triples \mathbf{P}_1 and \mathbf{P}_2 could be enough. If the scene has a symmetry with respect to a plane (for instance xy plane), the z -coordinates of the triples \mathbf{P}_i are null. This means that all \mathbf{P}_i will belong to the xy plane. If the scene is symmetrical with respect to xy and xz planes, for instance, then the triples \mathbf{P}_i will form a line in the direction of the x -axis.

3. Validation Results

In the following section, simulation results and real experiments are shown. For all the validation results, the spherical moments are computed directly using the pixel's grey levels without warping the images to the sphere space. The used formulas are given in References [26,27] for the perspective and fisheye cases respectively:

$$m_{ijk} = \iint_{\text{Image}} x_s^i y_s^j z_s^k \frac{(\xi + z_s)^3}{1 + \xi z_s} I(x, y, t) dx dy, \quad (20)$$

where ξ is the distortion parameter and (x, y) are the coordinates of a point in the metric image. The formulas of $x_s, y_s,$ and z_s as functions of (x, y) can be found in Reference [28]. Using Matlab with a computer equipped with a processor Intel(R) Core(TM) i5-7600 CPU @ 3.50 GHz 3.50 GHz and 16,0 Go RAM, the whole process including the computation of spherical moments of an image of size 480×640 pixels until the rotation estimation takes around 20 ms. This leaves room for reducing the time cost using other programming tools than Matlab to few milliseconds.

3.1. Simulation Results

The objective of this part is threefold:

- To show the validity of the proposed method to cameras obeying unified model, two different kinds of camera model are used to compute the images. The first corresponds to a simulated fisheye camera with the parameters chosen as focal scaling factors $F_x = F_y = 960$ pixels, coordinates of the principal point $u_x = 240$ and $u_y = 320$ pixels, and distortion parameter $\xi = 1.6$. The second model corresponds a conventional camera with focal scaling factors $F_x = F_y = 600$ pixels and coordinates of the principal point $u_x = 240$ and $u_y = 320$;
- To Validate our approach for large rotational motion;
- To test the effect of translational motion on the accuracy of the estimated rotations.

The Matlab code used for those simulations results can be downloaded online (http://aramis.iup.univ-evry.fr:8080/~hadj-abdelkader/Sensors/Sim_ForSensors.zip).

In the first results, a fisheye camera model is used to generate the images. More precisely, from the reference image shown on Figure 1a, three sets of 100 new images are computed based on randomly generated displacements. An example of the used images is shown on Figure 1b. All the three image sets were obtained with the same random rotations shown on Figure 2b but with translational motion of norms equal to 0 for the first set and 10 cm and 20 cm for the two others. For the two last sets, the same random translation directions are used but with different norms equal to 10 cm and 20 cm, respectively. To assess the estimated rotations, the norm of the vector composed by the errors on those Euler angles is used. The obtained accuracy results are shown on Figure 2a. It can be seen that, for the case where the camera displacement does not include translational motion, the rotation motion is well estimated (refer to the red plot of the Figure). The results obtained in the case where the camera displacement involves translational motion of norm 10 cm and 20 cm are shown respectively by the blue and the green plots of the Figure 2a. From the obtained plots, it can be seen that a maximum error of 4° (respectively 8°) and an average error less than 3° (respectively 5°) are obtained in the case of translation norm equal to 10 cm (respectively 20 cm).

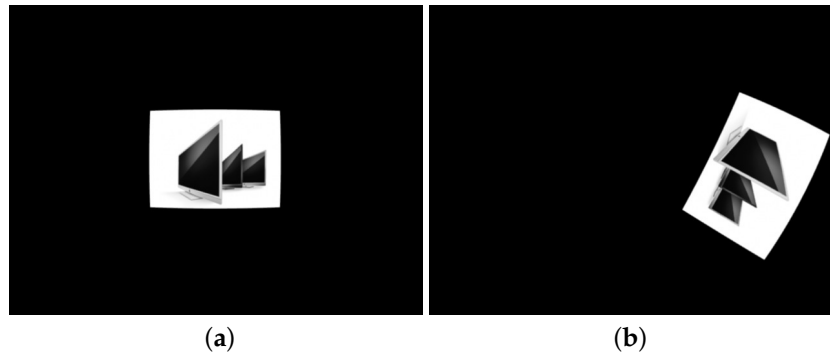


Figure 1. Results using a fisheye camera model: (a) the reference image and (b) example of rotated images.

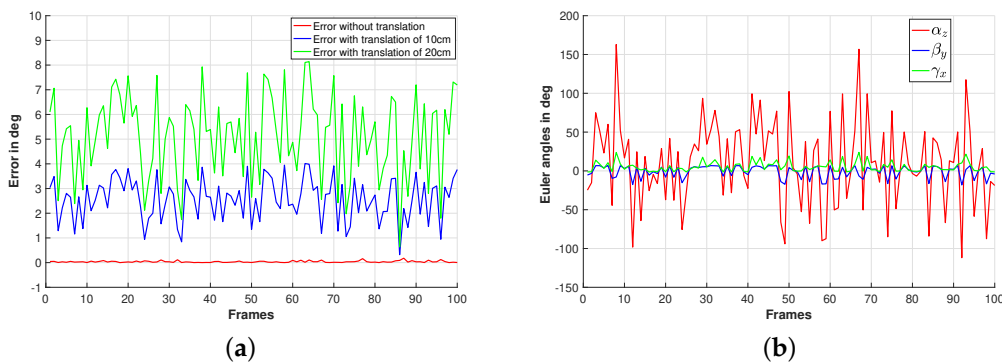


Figure 2. Simulation results using a fisheye camera model: (a) Estimation error of the rotation and (b) rotation ground truth expressed by Euler angles.

The second simulation results are based on the same previous setup, but this time, it is a conventional perspective camera that is used to generate the images. As for the first simulations, we consider three sets of 100 images corresponding to translational motions of different norms (0 cm, 10 cm, and 20 cm). Figure 3a,b shows the reference image and an example of the used rotated image for which the rotation has to be estimated. The ground truth of the rotational motions is shown on Figure 4b, while the accuracy of the rotation estimation is depicted on Figure 4a. From this Figure, it can be seen that the rotations are well estimated in the case where pure rotational motions are considered (refer to the red plot). It can also be noticed from the same figure (the blue and the green plots) that maximum rotation errors around 3.5° and 7.5° are observed respectively for the cases where the camera displacement involves a translational motion of norms equal to 10 cm and 20 cm, respectively.



Figure 3. Results using conventional camera model: (a) the reference image and (b) example of the rotated images.

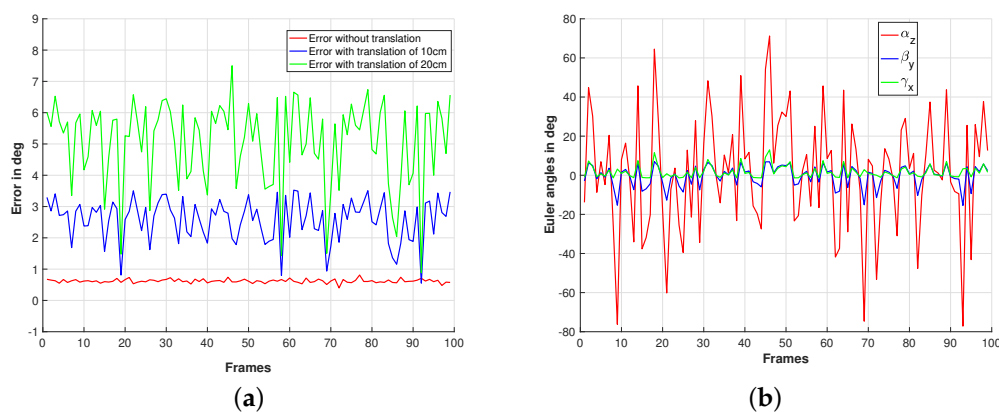


Figure 4. Simulation results using a perspective camera model: (a) Estimation error of the rotation and (b) rotation ground truth expressed.

3.2. Real Experiments

The following experimental results use the OVMIS (Omnidirectional Vision Dataset of the MIS laboratory) dataset (http://mis.u-picardie.fr/~g-caron/pub/data/OVMIS_dataset.zip). Based on the same database, the preliminary work [23] has provided several experimental results where the proposed method is compared to the SPHARM (SPherical HARmonic analyses) method [15] and to the phase correlation (COR) and photometric (PHO) methods [29]. In the following, the proposed method is validated on a larger set of scenarios from the dataset. Moreover, we compare the performance of the proposed method using the spherical moments computed from the image intensities and from its gradient. We also test the effect of introducing the weighting function on the spherical moment computation of image acquired by catadioptric camera. The used weighting is a function of the pixel coordinates (x, y) of the following form:

$$W(x, y) = e^{-\frac{(r-r_m)^4}{\sigma^4}} \quad (21)$$

where $r = \sqrt{(x - u_x)^2 + (y - u_y)^2}$ is the distance of the pixel to the principal point (u_x, u_y) and σ is a scalar that tunes the wide of the image zone to be used. Note that $W(x, y)$ has its maximum value 1 for $r = r_m$. Figure 5a shows the shape of the used weight function for this validation results. As can be seen from weighting function shape, the points close to the image border or to the blind zone in the mage center are given small weights to decrease the effect of appearance/disappearance of image parts. In all these experimental parts, the rotation motion is obtained by accumulating the estimated rotations between each two consecutive images.

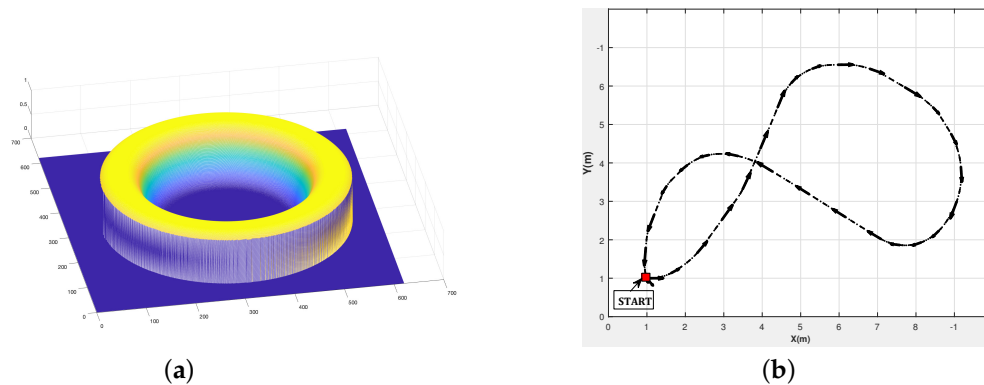


Figure 5. (a) Weight function and (b) Scenario 3: Eight trajectories of the pioneer robot.

The first validation result shown in the video (Supplementary Materials) is based on Dataset 1 of OVMIS with the Staubli robot arm. The considered sequence composed of 144 images has been acquired by applying regular pure rotations around the optical axis by a catadioptric camera mounted on the end effector of a Staubli in an indoor environment which corresponds to a rotation of 2.5° between two consecutive images. First, the proposed method is tested by considering all the image sequences. Then, one image from 10 and from 20 are considered, which corresponds to rotations of 25° and 50° , respectively, between two consecutive images. As it can be seen from the video that the obtained results show that the proposed method estimates accurately the rotations in the three considered cases.

The second validation result shown as well in the video (Supplementary Materials) is based on the Scenario 1 of the Pioneer database acquired using a catadioptric camera. A similar setup to the previous case is used: a pure rotational motion around the z-axis in an indoor environment. The proposed method is tested by considering all the image sequences, one image from 10 and one from 20. The obtained results show again that the proposed method estimates accurately the rotations in the three considered cases.

The last experiment tests the proposed method on the more challenging Scenario 3 of the OVMIS dataset of a pioneer robot moving according to the trajectory described in Figure 5b. The achieved path is of length 25.27 m, in the form of 8 motions that include both rotational and translation motions. It has been achieved by the mobile robot in 160 s, which corresponds to 0.187 m/s average speed. The ground truth is obtained by incorporating gyroscopic corrections into the odometry data based on wheel encoders. From the video (Supplementary Materials), it can be seen that the mobile robot has traversed uneven terrain, with changing illumination conditions due to moving cast shadows of the trees and with appearance/disappearance of some image parts.

Figure 6 compares the estimated rotation with/without using weighting function and based on moments computed from the image grey level or its gradient. From the obtained plots, it can be seen that using weighting function improves slightly the results obtained using the moment computed from the image intensities (compare the blue curve to the red one). On the other hand, it can be seen that the most accurate rotation estimation is obtained using the image gradient spherical moments and weighting function (refer to the green plot). The last experiments of the video (Supplementary Materials) describes the obtained results using the image gradient spherical moments by considering all the image sequences, one image from 10 and one from 20. From the obtained results, it can be seen that the rotation estimation is satisfactory for camera motion involving translational and rotational motions.

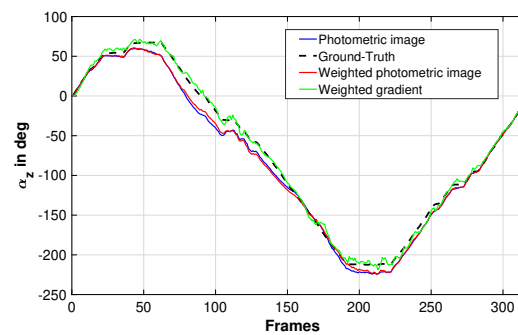


Figure 6. Estimated rotation of the pioneer robot around the Z-axis (Scenario 3).

3.3. Discussion

The triples computed from moment vectors provide a closed-form estimation to rotational motion. Compared to the classical method based on the alignment of principal axis, the proposed solution has no ambiguities on rotation formula. Recall that the alignment of principal axis method usually uses moments of higher orders to remove the ambiguities. If the estimated rotation is used as a feature in visual servoing or in active vision, the closed-form formula has a second advantage. For instance, as it has been proposed in References [30,31], the rotation vector is computed from the rotation matrix and used to control the rotational speed of a camera. For accurate visual servoing and active vision, most strategies are based on the analytical form of the visual features dynamics. Independently from the used method to estimate the rotation matrix, the relation between the time variation of a rotation vector and the rotational speed keeps the same form. On the other hand, its dynamic with respect to the translational speed becomes complex to obtain without a closed-form formula. Using the proposed solution in this paper, the dynamics of the rotation vector with respect to the whole degree can be obtained from those of the spherical moments [26,30,31].

A natural question arises about which are the best two triples to be used for building an orthonormal base. For instance, a base can be built from the two triples from P_1 to P_3 as well as by using two others from P_4 to P_7 . Although it is well accepted that the sensitivity to noise increases with the moment orders, there is not much difference between results obtained using moments of orders 2 and 3 and those obtained using moments of orders 4 and 3. Moreover, an orthonormal base can be obtained using more than only two triples as it is achieved in Reference [31]. Finally, recall also that the rotation matrix can be obtained using all the triples P_i as a generalized solution of the orthogonal Procrustes problem [32].

4. Conclusions

In this paper, a new method to estimate rotational motions using photometric moments has been provided. The obtained theoretical results have been validated using simulations and real experiments using onboard omnidirectional cameras and two different robots (robot arm and ground mobile robot). The proposed method has been also tested in different scenarios with appearance/disappearance of some image parts and lightening changes. Future works will be devoted to using photometric spherical moments for the control of the 6 degrees of freedom of a camera.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1424-8220/19/22/4958/s1>.

Author Contributions: Conceptualization, H.H.-A., O.T. and H.-E.B.; methodology, H.H.-A. and O.T.; validation, H.H.-A., O.T. and H.-E.B.; writing—original draft preparation, H.H.-A. and O.T.; writing—review and editing, H.H.-A. and O.T.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

- Triples computed from moments of order 2 and moments of order 3.

$$\mathbf{P}_1 = \begin{bmatrix} m_{003}m_{101} + m_{012}m_{110} + m_{021}m_{101} + m_{030}m_{110} + m_{101}m_{201} + m_{102}m_{200} + m_{110}m_{210} + m_{120}m_{200} + m_{200}m_{300} \\ m_{003}m_{011} + m_{011}m_{021} + m_{012}m_{020} + m_{020}m_{030} + m_{011}m_{201} + m_{102}m_{110} + m_{020}m_{210} + m_{110}m_{120} + m_{110}m_{300} \\ m_{002}m_{003} + m_{002}m_{021} + m_{011}m_{012} + m_{011}m_{030} + m_{002}m_{201} + m_{101}m_{102} + m_{011}m_{210} + m_{101}m_{120} + m_{101}m_{300} \end{bmatrix} \quad (\text{A1})$$

$$\mathbf{P}_2 = \begin{bmatrix} m_{002}m_{120} - 2m_{011}m_{111} + m_{020}m_{102} + m_{002}m_{300} - 2m_{101}m_{201} + m_{102}m_{200} + m_{020}m_{300} - 2m_{110}m_{210} + m_{120}m_{200} \\ m_{002}m_{030} - 2m_{011}m_{021} + m_{012}m_{020} + m_{002}m_{210} + m_{012}m_{200} - 2m_{101}m_{111} + m_{020}m_{210} + m_{030}m_{200} - 2m_{110}m_{120} \\ m_{002}m_{021} + m_{003}m_{020} - 2m_{011}m_{012} + m_{002}m_{201} + m_{003}m_{200} - 2m_{101}m_{102} + m_{020}m_{201} + m_{021}m_{200} - 2m_{110}m_{111} \end{bmatrix} \quad (\text{A2})$$

$$\mathbf{P}_3 = \begin{bmatrix} m_{002}m_{102} + 2m_{011}m_{111} + m_{020}m_{120} + 2m_{101}m_{201} + 2m_{110}m_{210} + m_{200}m_{300} \\ m_{002}m_{012} + 2m_{011}m_{021} + m_{020}m_{030} + 2m_{101}m_{111} + 2m_{110}m_{120} + m_{200}m_{210} \\ m_{002}m_{003} + 2m_{011}m_{012} + m_{020}m_{021} + 2m_{101}m_{102} + 2m_{110}m_{111} + m_{200}m_{201} \end{bmatrix} \quad (\text{A3})$$

References

- De Castro, E.; Morandi, C. Registration of Translated and Rotated Images Using Finite Fourier Transforms. *IEEE Trans. Pattern Anal. Mach. Intell.* **1987**, *PAMI-9*, 700–703. [[CrossRef](#)]
- Althloothi, S.; Mahoor, M.; Voyles, R. A Robust Method for Rotation Estimation Using Spherical Harmonics Representation. *IEEE Trans. Image Process.* **2013**, *22*, 2306–2316. [[CrossRef](#)] [[PubMed](#)]
- Bazin, J.C.; Demonceaux, C.; Vasseur, P.; Kweon, I. Rotation estimation and vanishing point extraction by omnidirectional vision in urban environment. *Int. J. Robot. Res.* **2012**, *31*, 63–81. [[CrossRef](#)]
- Makadia, A.; Daniilidis, K. Rotation recovery from spherical images without correspondences. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1170–1175. [[CrossRef](#)] [[PubMed](#)]
- Nister, D.; Naroditsky, O.; Bergen, J. Visual odometry. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; Volume 1, pp. I-652–I-659.
- Hadj-Abdelkader, H.; Mezouar, Y.; Martinet, P.; Chaumette, F. Catadioptric Visual Servoing from 3-D Straight Lines. *IEEE Trans. Robot.* **2008**, *24*, 652–665. [[CrossRef](#)]
- Safae-Rad, R.; Tchoukanov, I.; Smith, K.C.; Benhabib, B. Three-dimensional location estimation of circular features for machine vision. *IEEE Trans. Robot. Autom.* **1992**, *8*, 624–640. [[CrossRef](#)]
- Agrawal, M.; Konolige, K. Real-time Localization in Outdoor Environments using Stereo Vision and Inexpensive GPS. In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, 20–24 August 2006; Volume 3, pp. 1063–1068.
- Malis, E.; Marchand, E. Experiments with robust estimation techniques in real-time robot vision. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'06), Beijing, China, 9–15 October 2006; pp. 223–228.
- Micusik, B.; Pajdla, T. Structure from motion with wide circular field of view cameras. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1135–1149. [[CrossRef](#)] [[PubMed](#)]
- Hadj-Abdelkader, H.; Malis, E.; Rives, P. Spherical Image Processing for Accurate Visual Odometry with Omnidirectional Cameras. In Proceedings of the 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras—OMNIVIS, Marseille, France, 17 October 2008.
- Lim, J.; Barnes, N.; Li, H. Estimating Relative Camera Motion from the Antipodal-Epipolar Constraint. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1907–1914. [[CrossRef](#)] [[PubMed](#)]
- Horn, B.K.; Schunck, B.G. Determining Optical Flow. *Artif. Intell.* **1981**, *17*, 185–203. [[CrossRef](#)]
- Gluckman, J.; Nayar, S.K. Ego-motion and omnidirectional cameras. In Proceedings of the Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), Bombay, India, 4–7 January 1998; pp. 999–1005.
- Benseddik, H.E.; Hadj-Abdelkader, H.; Cherki, B.; Bouchafa, S. Direct method for rotation estimation from spherical images using 3D mesh surfaces with SPHARM representation. *J. Vis. Commun. Image Represent.* **2016**, *40*, 708–720. [[CrossRef](#)]

16. Teague, M.R. Image analysis via the general theory of moments. *JOSA* **1980**, *70*, 920–930. [[CrossRef](#)]
17. Abu-Mostafa, Y.S.; Psaltis, D. Image normalization by complex moments. *IEEE Trans. Pattern Anal. Mach. Intell.* **1985**, *PAMI-7*, 46–55. [[CrossRef](#)]
18. Tahri, O.; Araujo, H.; Mezouar, Y.; Chaumette, F. Efficient decoupled pose estimation from a set of points. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'2013), Tokyo, Japan, 3–7 November 2013; pp. 1608–1613.
19. Tahri, O.; Chaumette, F. Complex objects pose estimation based on image moment invariants. In Proceedings of the 2005 IEEE International Conference on Robotics and Automation, Barcelona, Spain, 18–22 April 2005; pp. 436–441.
20. Tahri, O.; Tamtsia, A.Y.; Mezouar, Y.; Demonceaux, C. Visual servoing based on shifted moments. *IEEE Trans. Robot.* **2015**, *31*, 798–804. [[CrossRef](#)]
21. Bakthavatchalam, M.; Tahri, O.; Chaumette, F. A Direct Dense Visual Servoing Approach using Photometric Moments. *IEEE Trans. Robot.* **2018**. [[CrossRef](#)]
22. Tahri, O.; Mezouar, Y.; Chaumette, F.; Corke, P. Decoupled image-based visual servoing for cameras obeying the unified projection model. *IEEE Trans. Robot.* **2010**, *26*, 684–697. [[CrossRef](#)]
23. Hadj-Abdelkader, H.; Tahri, O.; Benseddik, H.E. Closed form solution for Rotation Estimation using Photometric Spherical Moments. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 627–634.
24. Courbon, J.; Mezouar, Y.; Eckert, L.; Martinet, P. A generic fisheye camera model for robotic applications. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 29 October–2 November 2007; pp. 1683–1688.
25. Barnard, S.T. Interpreting Perspective Images. *Artif. Intell.* **1983**, *21*, 435–462. [[CrossRef](#)]
26. Tahri, O. Application des Moments à L'asservissement Visuel et au Calcul de Pose. Ph.D. Thesis, Université de Rennes, Rennes, France, 2004.
27. Tatsambon Fomena, R. Asservissement Visuel par Projection Sphérique. Ph.D. Thesis, University of Rennes 1, Rennes, France, 2008.
28. Hadj-Abdelkader, H.; Mezouar, Y.; Martinet, P. Decoupled Visual Servoing Based on the Spherical Projection of Points. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'09), Kobe, Japan, 12–17 May 2009; pp. 731–736.
29. Morbidi, F.; Caron, G. Phase Correlation for Dense Visual Compass from Omnidirectional Camera-Robot Images. *IEEE Robot. Autom. Lett.* **2017**, *2*, 688–695. [[CrossRef](#)]
30. Fomena, R.T.; Tahri, O.; Chaumette, F. Distance-based and orientation-based visual servoing from three points. *IEEE Trans. Robot.* **2011**, *27*, 256–267. [[CrossRef](#)]
31. Tahri, O.; Araujo, H.; Chaumette, F.; Mezouar, Y. Robust image-based visual servoing using invariant visual information. *Robot. Auton. Syst.* **2013**, *61*, 1588–1600. [[CrossRef](#)]
32. Schönemann, P.H. A generalized solution of the orthogonal Procrustes problem. *Psychometrika* **1966**, *31*, 1–10. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).