Catadioptric Stereo with Planar Mirrors: Multiple-View Geometry and Camera Localization

Gian Luca Mariottini, Stefano Scheggi, Fabio Morbidi and Domenico Prattichizzo

Abstract *Planar catadioptric stereo* (PCS) vision sensors consist of a pinhole camera and two or more planar mirrors. PCS systems have recently received an increasing attention in computer vision and have a promising applicability in robotics, since the use of mirrors allows to obtain a stereo view without the need of exact multi-camera synchronization and stereo calibration. The paper presents a rigorous analytical treatment of the imaging geometry of PCS sensors and introduce new multiple view properties that are instrumental in addressing the camera localization problem. Original results on mirror calibration are also provided. Extensive simulation and real-data experiments conducted with an eye-in-hand robot illustrate the theory and show the effectiveness of the proposed designs.

1 Introduction

1.1 Motivation and related works

Stereoscopic vision has been playing a key role in many fields, such as, e.g, 3-D imaging for entertainment and mapping, medicine, industrial robotics, augmented reality [13, 11]. Standard stereo sensors are composed of two coupled pinhole cameras and are generally expensive and difficult to calibrate. In addition, the limited field of view considerably limits their application range. To alleviate this problem, several works in the robotics and computer vision literature have recently proposed the use of *catadioptric* vision sensors [18, 4, 21, 15], which, combining

Gian L. Mariottini
Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA, e-mail: gianluca@cs.umn.edu

Stefano Scheggi, Fabio Morbidi and Domenico Prattichizzo
Department of Information Engineering, University of Siena, 53100 Siena, Italy,
e-mail: {scheggi, morbidi, prattichizzo}@dii.unisi.it

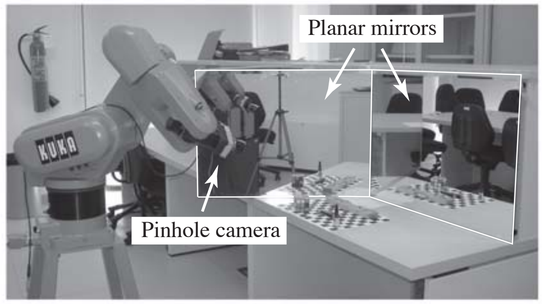


Fig. 1 PCS experimental setup composed of a pinhole camera mounted on the end-effector of a manipulator and two planar mirrors.

both refracting (lens) and reflecting (mirrors) elements, naturally enlarge the field of view of classic pinhole cameras.

By using mirror reflections of a scene, stereo images can be captured with a single camera (catadioptric stereo). This paper focuses on a special class of single camera stereo systems called PCS sensors, consisting of a pinhole camera and two, or more, planar mirrors [8, 19]. Single camera stereo has several advantages over traditional two-camera stereo: first of all, only a single set of internal camera calibration parameters needs to be determined. Second, no additional software or hardware is needed for exact camera synchronization [10]. The necessity of possibly large and unwieldy planar mirrors makes PCS systems not suited for mobile robotics. However, these sensors find a natural application in industrial robotics: in fact multiple planar mirrors disposed in the workspace can be used to improve the pose accuracy of an eye-in-hand robotic arm performing high precision assembling and close manipulation tasks.

Several PCS sensors have been designed and implemented in the last few years (see, e.g., [20, 3]) and preliminary results on the geometry, calibration and 3-D scene reconstruction from reflective symmetry have been presented in [5, 6, 17]. The set of constraints to be satisfied in order to obtain rectified stereo images from a catadioptric system has been studied in [7]. An affine approximation for epipolar-based PCS rectification has been recently proposed in [22].

From the previous overview, we see that although some research has been done in this field, the specific properties of PCS systems have been addressed only sparsely in the literature and a systematic theory on PCS sensors has yet to be established. In particular, the case of a moving camera in a fixed mirrors scenario and the multiple view geometry associated with it, have not been explored so far.

1.2 Contributions

The original contribution of this paper is threefold:

- We provide a rigorous analytical treatment of the imaging geometry of PCS sensors composed of a pinhole camera and two planar mirrors (see Fig. 1) and introduce new algebraic results on the multiple view geometry for the case of static and moving cameras (multi-view PCS).

- On the grounds of these novel results, we address the *image-based camera localization* problem and present original methods for *mirror calibration* (i.e., computation of the angle between the mirrors).
- We present extensive simulation and real-data experiments conducted with an eye-in-hand robot, in order to illustrate the theory and show the effectiveness of the proposed designs in real scenarios.

A preliminary version of this paper appeared in [16], compared to which we provide here new theoretical results as well as a more extensive experimental validation.

1.3 Organization

Section 2 reviews the basic theory related to perspective projection through planar mirrors and introduces the reflective epipolar geometry. Section 3 deals with the single and multiple view geometry for PCS sensors. Section 4 addresses a solution to the mirror calibration and image-based camera localization problems. Simulation and real-data experiments are reported in section 5. In section 6, the main contributions of the paper are summarized and possible avenues of future research are highlighted.

2 Planar mirrors and perspective projection

In this section we review the imaging properties of catadioptric systems with a single planar mirror [5]. The basic concepts of *virtual point*, *reflection transformation* and *virtual camera projection equivalence* are introduced. The original notion of *reflective epipolar geometry* is presented at the end of the section.

2.1 The virtual point and the reflection transformation

Let us consider the setup reported in Fig. 2(a) where a perspective camera $\langle c \rangle$ is in front of a planar mirror Π with normal vector $\mathbf{n}_\pi$. A 3-D point $\mathbf{X} \triangleq [x\ y\ z]^T$ is supposed to be in front of the mirror as well ($\widetilde{\mathbf{X}}$ indicates its extension in homogeneous coordinates). For the sake of clearness, hereafter we will refer to the simplified setup in Fig. 2(a): however, the results of this section are valid for generic camera-mirror arrangements. Note that the perspective image $\mathbf{u}$ (pixels) of $\mathbf{X}$ after its reflection by the planar mirror Π can be calculated as the direct projection on $\langle c \rangle$ of the so-called *virtual point* $\mathbf{X}^{[\pi]}$.

Proposition 1 (Perspective projection). *Let us consider the setup of Fig. 2(a) in which a planar mirror with normal vector $\mathbf{n}_\pi$ is distant d_π from the camera $\langle c \rangle$. Then the perspective projection $\widetilde{\mathbf{u}} \triangleq [u\ v\ 1]^T$ (pixels) of a generic 3-D point $\mathbf{X}$ that*

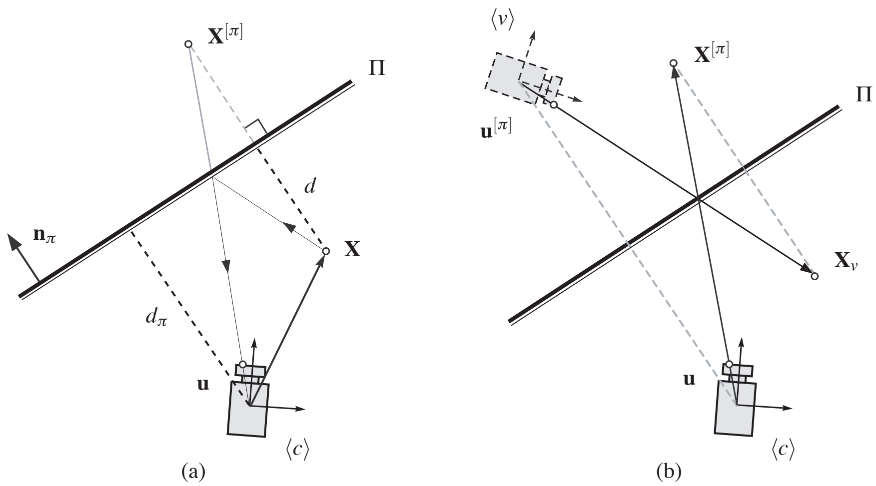


Fig. 2 (a) Reflection through the planar mirror Π and the virtual point $\mathbf{X}^{[\pi]}$; and (b) the virtual camera $\langle v \rangle$.

is mirrored by Π *is given by,*

$$\lambda\, \widetilde{\mathbf{u}} = \mathbf{K}\,[\mathbf{I}\ \ \mathbf{0}]\,\mathbf{D}^{[\pi]}\,\widetilde{\mathbf{X}}\ \ \textit{with}\ \ \lambda \in \mathbb{R}^+, \tag{1}$$

where $\mathbf{K}$ *is the calibration matrix of the camera and* $\mathbf{D}^{[\pi]}$ *is the* reflection transformation *about the mirror* $\mathbf{n}_\pi$*, defined as,*

$$\mathbf{D}^{[\pi]} = \begin{bmatrix} \mathbf{S}^{[\pi]} & 2\,d_\pi\,\mathbf{n}_\pi \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{2}$$

where $\mathbf{S}^{[\pi]} = \mathbf{I} - 2\,\mathbf{n}_\pi \mathbf{n}_\pi^T$.

Proof. The perspective projection $\widetilde{\mathbf{u}}$ of the virtual point $\widetilde{\mathbf{X}}^{[\pi]}$ is given by,

$$\lambda\, \widetilde{\mathbf{u}} = \mathbf{K}\,[\mathbf{I}\ \ \mathbf{0}]\,\widetilde{\mathbf{X}}^{[\pi]}. \tag{3}$$

From geometrical considerations in Fig. 2(a), we see that,

$$\mathbf{X}^{[\pi]} = \mathbf{X} + 2\,d\,\mathbf{n}_\pi. \tag{4}$$

Since $d = d_\pi + \mathbf{X}^T\mathbf{n}_\pi$, then (4) can be rewritten as,

$$\mathbf{X}^{[\pi]} = (\mathbf{I} - 2\,\mathbf{n}_\pi \mathbf{n}_\pi^T)\mathbf{X} + 2\,d_\pi\,\mathbf{n}_\pi. \tag{5}$$

By introducing the matrix $\mathbf{D}^{[\pi]}$ defined in (2), Equation 5 becomes,

$$\widetilde{\mathbf{X}}^{[\pi]} = \mathbf{D}^{[\pi]}\,\widetilde{\mathbf{X}}. \tag{6}$$

Substituting (6) in (3), we obtain Equation 1. □

Note that $\mathbf{S}^{[\pi]} \in \mathrm{O}(3)$, $\det(\mathbf{S}^{[\pi]}) = -1$ and that $(\mathbf{D}^{[\pi]})^{-1} = \mathbf{D}^{[\pi]}$.

2.2 The virtual camera and the projection equivalence

Proposition 1 of section 2.1 shows how the 3-D point $\mathbf{X}$ is mirrored by Π at $\mathbf{u}$ onto the image plane of the camera $\langle c \rangle$. Analogously to the concept of virtual point, we can introduce the geometrically intuitive notion of *virtual camera* $\langle v \rangle$ (dashed in Fig. 2(b)), whose reference frame is simply reflected with respect to $\langle c \rangle$. The proof of the next proposition follows directly from the observation that,

$$\mathbf{X}_v = \mathbf{X}^{[\pi]}, \tag{7}$$

where $\mathbf{X}_v$ is the point $\mathbf{X}$ in $\langle v \rangle$.

Proposition 2 (Projection equivalence). *Let* $\mathbf{u}$ *be the perspective projection in* $\langle c \rangle$ *of a 3-D point* $\mathbf{X}$ *after its reflection by* Π. *Then,*

$$\mathbf{u} = \mathbf{u}^{[\pi]}, \tag{8}$$

being $\mathbf{u}^{[\pi]}$ *the perspective projection of* $\mathbf{X}$ *onto the image plane of the virtual camera* $\langle v \rangle$.

Proposition 2 states that the perspective projection $\mathbf{u}$ of $\mathbf{X}^{[\pi]}$ coincides with the perspective projection $\mathbf{u}^{[\pi]}$ of $\mathbf{X}_v$. In other words, the camera projections of the reflected points correspond to the virtual camera projections of the real points.

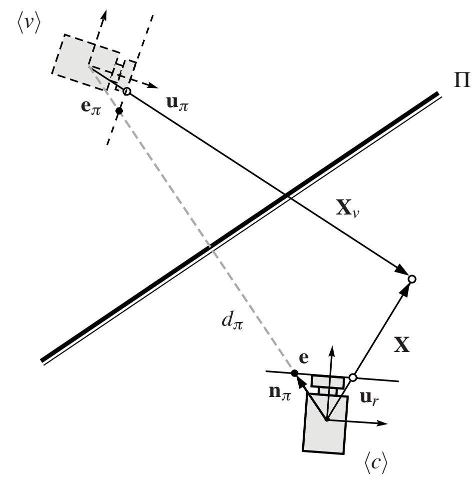
2.3 Reflective epipolar geometry

In this section we study the imaging geometry relating cameras $\langle c \rangle$ and $\langle v \rangle$. Note that this is different from [5, section 3.1], where the epipolar geometry between the virtual cameras is investigated.

Proposition 3 (Reflective epipolar constraint). *Let us consider the setup in Fig. 3 and let* d_π *and* $\mathbf{n}_\pi$ *be the distance and the normal of the mirror* Π *measured from* $\langle c \rangle$, *respectively. Let* $\widetilde{\mathbf{u}}_r$, $\widetilde{\mathbf{u}}_\pi \in \mathbb{R}^3$ *be the homogeneous representation of the projection of a 3-D point in the image plane of the views* $\langle c \rangle$ *and* $\langle v \rangle$, *respectively. Then, the* reflective epipolar constraint *is given by,*

$$\widetilde{\mathbf{u}}_\pi^T\,\mathbf{E}^{[\pi]}\,\widetilde{\mathbf{u}}_r = 0,$$

Fig. 3 Reflective epipolar geometry.



where the reflective essential matrix*,*

$$\mathbf{E}^{[\pi]} = 2\, d_\pi\, [\mathbf{n}_\pi]_\times,$$

being $[\mathbf{n}_\pi]_\times$ *the skew-symmetric matrix associated with the vector* $\mathbf{n}_\pi$*.*

Proof. Let $\mathbf{X}$ and $\mathbf{X}_v$ be the 3-D coordinates of a point in the camera frames $\langle c \rangle$ and $\langle v \rangle$, (see Fig. 3). $\mathbf{X}$ and $\mathbf{X}_v$ are related by the following rigid-body transformation,

$$\mathbf{X}_v = \mathbf{S}^{[\pi]}\,\mathbf{X} + 2\, d_\pi \mathbf{n}_\pi, \tag{9}$$

readily derived by substituting (7) into (5). If the internal camera calibration matrix $\mathbf{K}$ is the identity, then $\mathbf{X}_v = \lambda_\pi\, \widetilde{\mathbf{u}}_\pi$, $\mathbf{X} = \lambda_r\, \widetilde{\mathbf{u}}_r$ and (9) can be rewritten as follows,

$$\lambda_\pi \widetilde{\mathbf{u}}_\pi = \lambda_r\, \mathbf{S}^{[\pi]}\, \widetilde{\mathbf{u}}_r + 2\, d_\pi \mathbf{n}_\pi, \tag{10}$$

where $\lambda_\pi, \lambda_r \in \mathbb{R}^+$ are unknown depths. Simple matrix manipulations on (10), lead directly to the epipolar constraint,

$$\widetilde{\mathbf{u}}_\pi^T\, (2\, d_\pi\, [\mathbf{n}_\pi]_\times\, \mathbf{S}^{[\pi]})\, \widetilde{\mathbf{u}}_r = 0.$$

By definition,

$$\mathbf{E}^{[\pi]} \triangleq 2\, d_\pi [\mathbf{n}_\pi]_\times\, \mathbf{S}^{[\pi]} = 2\, d_\pi [\mathbf{n}_\pi]_\times (\mathbf{I} - 2\, \mathbf{n}_\pi\, \mathbf{n}_\pi^T) = 2\, d_\pi [\mathbf{n}_\pi]_\times.$$

□

Note that the vector $\mathbf{n}_\pi$ can be readily recovered (up to a scale factor), from the right null-space of $\mathbf{E}^{[\pi]}$. If the camera calibration matrix

$$\mathbf{K} = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

where f_x, f_y (pixels) denote the focal length of the camera along the x and y directions, s is the skew factor and (u_0, v_0) (pixels) are the CCD's principal point coordinates, we can introduce the *reflective fundamental matrix* $\mathbf{F}^{[\pi]}$,

$$\mathbf{F}^{[\pi]} \triangleq \mathbf{K}^{-T}\,\mathbf{E}^{[\pi]}\,\mathbf{K}^{-1}. \tag{11}$$

Since $\mathbf{F}^{[\pi]}$ is skew-symmetric (in fact we have, $\mathbf{F}^{[\pi]} + (\mathbf{F}^{[\pi]})^T = 2d_\pi(\mathbf{K}^{-T}[\mathbf{n}_\pi]_\times \mathbf{K}^{-1} - \mathbf{K}^{-T}[\mathbf{n}_\pi]_\times \mathbf{K}^{-1}) = \mathbf{0}$), the left and right null-space of $\mathbf{F}^{[\pi]}$ are equal. As such, the epipoles are equal,

$$\widetilde{\mathbf{e}} = \widetilde{\mathbf{e}}_\pi.$$

Note that since $\mathbf{F}^{[\pi]}$ has only 2 DOF (that correspond to the position of the epipole $\widetilde{\mathbf{e}}$), at least 2 pairs of corresponding points are necessary to determine $\mathbf{F}^{[\pi]}$. Since $\widetilde{\mathbf{e}} = \widetilde{\mathbf{e}}_\pi$ the epipolar lines in the real and virtual view coincide and the epipole can be regarded as a *vanishing point*, being the intersection of 3-D parallel lines (the lines joining the real and virtual points) projected onto the image plane.

Remark 1. Note that the imaging geometry relating cameras $\langle c\rangle$ and $\langle v\rangle$ corresponds to that existing between two cameras undergoing a pure translational motion.

3 Single and multiple view geometry for PCS sensors

3.1 Single view geometry

In this section we assume that a camera $\langle c\rangle$ observes a set of 3-D points reflected by two planar mirrors (see Fig. 4). In this case two corresponding virtual cameras $\langle v_1\rangle$ and $\langle v_2\rangle$ exist and suitable geometries relating $\langle c\rangle$ with both $\langle v_1\rangle$ and $\langle v_2\rangle$ can be defined. The generalization to multiple mirrors is straightforward and it will be not discussed herein. Let be given the image points $\mathbf{u}_i^{[1]}$, $\mathbf{u}_i^{[2]}$, $i \in \{1,\ldots,n\}$ in $\langle c\rangle$, projections of a set of $n \geq 8$ 3-D points $\mathbf{X}_i$ reflected onto the mirrors $\mathbf{n}_1$ and $\mathbf{n}_2$, respectively. Note that while the subscript i is the point index, the superscript inside the brackets will always refer to the mirror number through which that vector is reflected (for the sake of simplicity, we will henceforth neglect the subscript π in the mirrors parameters). Given the two-mirror setup reported in Fig. 4, let $\mathbf{D}^{[1]}$ and $\mathbf{D}^{[2]}$ be the corresponding reflection transformations. The following expression holds true,

$$\mathbf{D}^{[2]}\,\mathbf{D}^{[1]} = \begin{bmatrix} \mathbf{R}_D & \mathbf{t}_D \\ \mathbf{0}^T & 1 \end{bmatrix},$$

where,

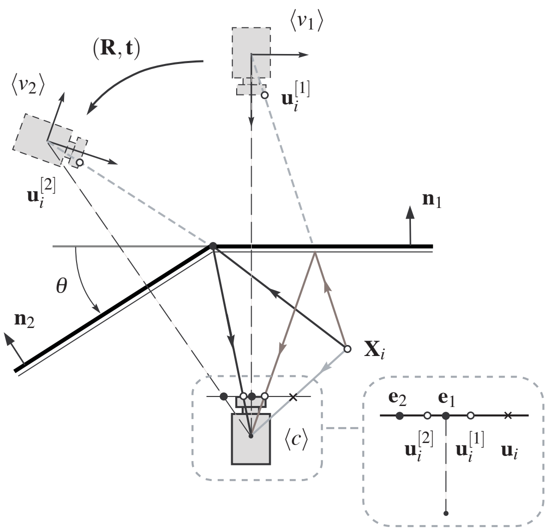


Fig. 4 The camera $\langle c \rangle$ and the virtual cameras $\langle v_1 \rangle$ and $\langle v_2 \rangle$.

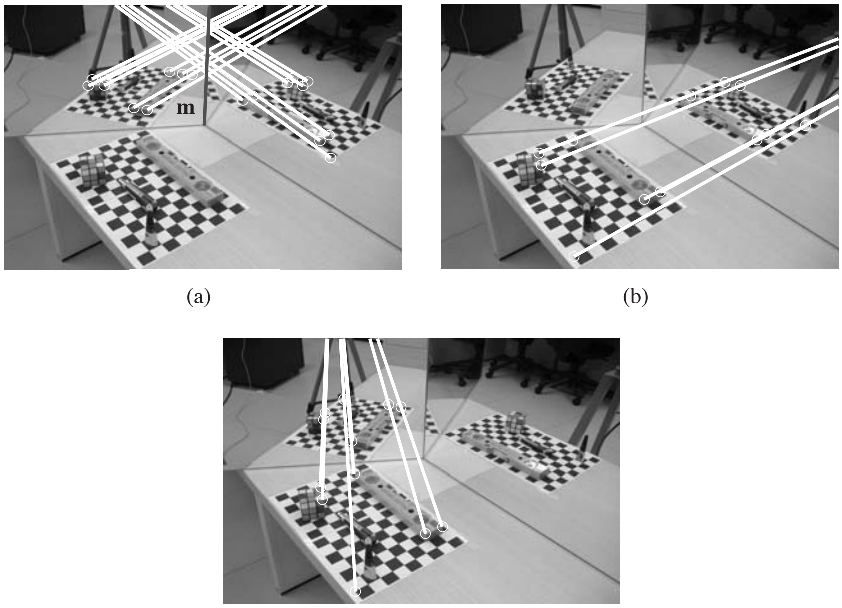
$$\mathbf{R}_D \triangleq \mathbf{I} + 4(\mathbf{n}_1^T\mathbf{n}_2)\,\mathbf{n}_1\mathbf{n}_2^T - 2\,\mathbf{n}_1\mathbf{n}_1^T - 2\,\mathbf{n}_2\mathbf{n}_2^T,$$
$$\mathbf{t}_D \triangleq 2\,d_1\mathbf{n}_1 - 2(d_1(\mathbf{n}_1^T\mathbf{n}_2) + d_2)\mathbf{n}_2.$$

Note that $(\mathbf{R}_D, \mathbf{t}_D)$ only depends on the mirror setup (i.e., $\mathbf{n}_1$, $\mathbf{n}_2$ and d_1, d_2), and not on the observed scene. On the other hand, let $\mathbf{H}_{v_1}^{v_2}$ be the homogeneous transformation matrix representing the rigid body motion between the frames $\langle v_1 \rangle$ and $\langle v_2 \rangle$. It is easily found that (see Fig. 4),

$$\mathbf{H}_{v_1}^{v_2} \triangleq \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} = \mathbf{D}^{[1]}\,\mathbf{D}^{[2]}, \tag{12}$$

where $(\mathbf{R}, \mathbf{t})$ is the rigid body motion between $\langle v_1 \rangle$ and $\langle v_2 \rangle$. Owing to proposition 2, the points $\mathbf{u}_i^{[1]}$ and $\mathbf{u}_i^{[2]}$ in $\langle c \rangle$ are corresponding in both the virtual cameras $\langle v_1 \rangle$ and $\langle v_2 \rangle$, (see Fig. 4). This implies the existence of the epipolar geometry relating $\langle v_1 \rangle$ and $\langle v_2 \rangle$, i.e., $\widetilde{\mathbf{u}}_i^{[2]\,T}\mathbf{F}\,\widetilde{\mathbf{u}}_i^{[1]} = 0$. The fundamental matrix $\mathbf{F}$ can be estimated from a set of (at least) 8 image points and the epipoles $\mathbf{e}_1$ and $\mathbf{e}_2$ are obtained as the right and left null-spaces of $\mathbf{F}$ [9]. Moreover, given the camera calibration matrix $\mathbf{K}$, from $\mathbf{F}$ we can compute the essential matrix $\mathbf{E} \triangleq [\mathbf{t}]_\times\,\mathbf{R}$. Once $\mathbf{E}$ is known, a decomposition [9] can be carried out to finally compute the matrix $\mathbf{R}$ and the vector $\mathbf{t}$ (up to a scale factor).

Fig. 5 shows the epipolar lines (white) relative to pairs of corresponding points in the real and virtual views, on a sample image. Fig. 5(a) reports the epipolar lines between the virtual cameras $\langle v_1 \rangle$ and $\langle v_2 \rangle$: the baseline lies far above the edge of the image. As shown in [6], all corresponding epipolar lines intersect at the image projection $\mathbf{m}$ of the mirrors screw axis (i.e., the 3-D line of intersection between the mirrors). Figs. 5(b), 5(c) show the epipolar lines between $\langle v_1 \rangle$ and $\langle c \rangle$, and between $\langle v_2 \rangle$ and $\langle c \rangle$, respectively.



(a) (b) (c)

Fig. 5 Epipolar lines (white) on a sample image: (a) between $\langle v_1 \rangle$ and $\langle v_2 \rangle$; (b) between $\langle v_1 \rangle$ and $\langle c \rangle$; and (c) between $\langle v_2 \rangle$ and $\langle c \rangle$.

3.2 Multiple view geometry

In this section we address the case of a moving camera that observes a set of 3-D points $\mathbf{X}_i$ reflected by two mirrors, from two views $\langle c_1 \rangle$ and $\langle c_2 \rangle$ (see Fig. 6). Let $\mathbf{H}_R \triangleq \mathbf{H}_{c_1}^{c_2}$ be the homogeneous transformation matrix relating $\langle c_1 \rangle$ and $\langle c_2 \rangle$, and $\mathbf{H}_{v_1^{[2]}}^{v_2^{[2]}}$, $\mathbf{H}_{v_2^{[1]}}^{v_2^{[2]}}$, $\mathbf{H}_{v_1^{[1]}}^{v_2^{[1]}}$, $\mathbf{H}_{v_1^{[1]}}^{v_1^{[2]}}$ (with a slight abuse of notation since differently from section 3.1 the subscript of v refers herein to the camera number), the four homogeneous matrices encoding the relative pose between the virtual views. Finally, let $\mathbf{D}_{c_1}^{[1]}$ and $\mathbf{D}_{c_1}^{[2]}$ denote the reflection transformations about the two mirrors, written in $\langle c_1 \rangle$. We are now ready to introduce the following proposition, which establishes a link between $\mathbf{H}_R$ and $\mathbf{H}_{v_2^{[1]}}^{v_2^{[2]}}$.

Proposition 4 (Multiple cameras mirroring). *The following equation holds true,*

$$\mathbf{H}_{v_2^{[1]}}^{v_2^{[2]}} = \mathbf{H}_R^{-1}\, \mathbf{D}_{c_1}^{[1]}\, \mathbf{D}_{c_1}^{[2]}\, \mathbf{H}_R \tag{13}$$

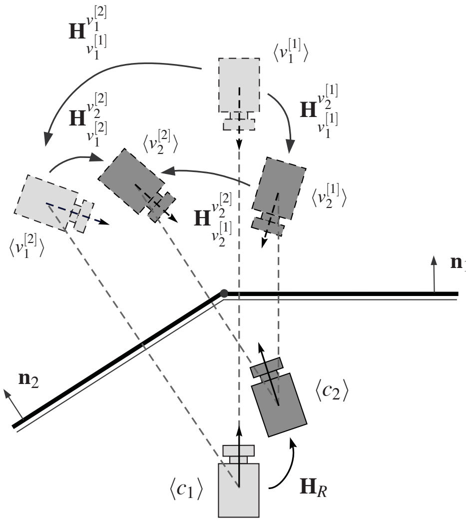


Fig. 6 Multiple view geometry for the cameras $\langle c_1 \rangle$ and $\langle c_2 \rangle$.

Proof. From (12) we have that,

$$\mathbf{D}_{c_1}^{[1]}\,\mathbf{D}_{c_1}^{[2]} = \mathbf{H}_{v_1^{[1]}}^{v_1^{[2]}}. \tag{14}$$

Premultiplying and postmultiplying (14) by $\mathbf{H}_R^{-1}$ and $\mathbf{H}_R$, respectively, we obtain,

$$\mathbf{H}_R^{-1}\,\mathbf{H}_{v_1^{[1]}}^{v_1^{[2]}}\,\mathbf{H}_R = \mathbf{H}_R^{-1}\,\mathbf{D}_{c_1}^{[1]}\,\mathbf{D}_{c_1}^{[2]}\,\mathbf{H}_R. \tag{15}$$

From an inspection of Fig. 6, it is easy to verify that the following equalities hold true,

$$\mathbf{H}_R = \mathbf{H}_{v_1^{[1]}}^{v_2^{[1]}} = \mathbf{H}_{v_1^{[2]}}^{v_2^{[2]}}, \tag{16}$$

i.e., the rigid motion between $\langle c_1 \rangle$ and $\langle c_2 \rangle$ is equal to the rigid motion between $\langle v_1^{[1]} \rangle$ and $\langle v_2^{[1]} \rangle$, and between $\langle v_1^{[2]} \rangle$ and $\langle v_2^{[2]} \rangle$. Using (16) and (14) into (15), we obtain,

$$\mathbf{H}_R^{-1}\,\mathbf{H}_{v_1^{[1]}}^{v_1^{[2]}}\,\mathbf{H}_R = \mathbf{H}_{v_2^{[2]}}^{v_1^{[2]}}\,\mathbf{H}_{v_1^{[1]}}^{v_1^{[2]}}\,\mathbf{H}_{v_1^{[2]}}^{v_2^{[2]}} = \mathbf{H}_{v_2^{[2]}}^{v_1^{[2]}}\,\mathbf{H}_{v_1^{[1]}}^{v_2^{[1]}}\,\mathbf{H}_{v_2^{[1]}}^{v_2^{[2]}} = \mathbf{H}_R^{-1}\,\mathbf{H}_R\,\mathbf{H}_{v_2^{[1]}}^{v_2^{[2]}} = \mathbf{H}_{v_2^{[1]}}^{v_2^{[2]}},$$

and (13) is thus proved. □

Remark 2. Note that Equation 13 allows one to estimate the rigid motion $\mathbf{H}_R$ also when the epipolar geometry between $\langle c_1 \rangle$ and $\langle c_2 \rangle$ is not well-defined (*small baseline* case). In fact the epipolar geometry between the virtual cameras is always well-defined by construction.

4 Mirror calibration and image-based camera localization

In this section we address the mirror calibration and image-based camera localization problems using the PCS properties presented in sections 2 and 3. Proposition 5 in the next section will be instrumental for proposition 8 in section 4.2.

4.1 Mirror calibration

Consider the setup in Fig. 4 and assume that at least $n \geq 2$ points can be directly observed by $\langle c \rangle$ at $\mathbf{u}_i$, $i \in \{1,\dots,n\}$. If the same set of points is also reflected by the mirrors $\mathbf{n}_1$ and $\mathbf{n}_2$ at $\mathbf{u}_i^{[1]}$ and $\mathbf{u}_i^{[2]}$ (white dots), then two reflective fundamental matrices $\mathbf{F}^{[1]}$ and $\mathbf{F}^{[2]}$ do exist and can be computed from the corresponding pairs $(\mathbf{u}_i^{[1]}, \mathbf{u}_i)$ and $(\mathbf{u}_i^{[2]}, \mathbf{u}_i)$, respectively (cf. proposition 2). From $\mathbf{F}^{[1]}$ and $\mathbf{F}^{[2]}$, we can then determine the epipoles $\mathbf{e}_1$ and $\mathbf{e}_2$ as their right null-spaces (black dots). Since the direction of each epipole $\mathbf{e}_j$ is always parallel to $\mathbf{n}_j$, $j \in \{1,2\}$, we then obtain the following:

Proposition 5 (Mirror calibration with the epipoles). *The angle θ between the mirrors $\mathbf{n}_1$ and $\mathbf{n}_2$ is given by,*

$$\theta = \arccos\left(\frac{(\mathbf{K}^{-1}\widetilde{\mathbf{e}}_1)^T(\mathbf{K}^{-1}\widetilde{\mathbf{e}}_2)}{\|\mathbf{K}^{-1}\widetilde{\mathbf{e}}_1\|\,\|\mathbf{K}^{-1}\widetilde{\mathbf{e}}_2\|}\right).$$

In the next proposition the epipoles between the virtual cameras are used to solve the mirror calibration problem (see Fig. 7(a)). Let $\mathbf{F}^{[21]}$ be the fundamental matrix computed from the corresponding points $(\mathbf{u}_i^{[1]}, \mathbf{u}_i^{[2]})$ and let γ be the angle between the virtual epipoles $\mathbf{e}^{[12]}$ and $\mathbf{e}^{[21]}$. It is easy to verify that,

$$\gamma = \arccos\left(\frac{(\mathbf{K}^{-1}\widetilde{\mathbf{e}}^{[12]})^T(\mathbf{K}^{-1}\widetilde{\mathbf{e}}^{[21]})}{\|\mathbf{K}^{-1}\widetilde{\mathbf{e}}^{[12]}\|\,\|\mathbf{K}^{-1}\widetilde{\mathbf{e}}^{[21]}\|}\right).$$

Proposition 6 (Mirror calibration with the virtual epipoles). *The angle θ between the mirrors $\mathbf{n}_1$ and $\mathbf{n}_2$ is given by,*

$$\theta = \frac{\pi - \gamma}{2}.$$

Remark 3. It is worth emphasizing here that our calibration notion is different from that considered in previous works (and notably in [6]). In fact, with "mirror calibration" we mean the estimation of the *angle* between the mirrors, while in [6] the authors mean the estimation the *focal length* of the camera and the orientation of the mirrors *screw axis*.

4.2 Image-based camera localization

This section deals with the estimation of the rigid motion $(\mathbf{R}_w^c, \mathbf{t}_w^c)$ of the camera $\langle c \rangle$ with respect to a world frame $\langle w \rangle$. Without loss of generality, we will assume that the z-axis of $\langle w \rangle$ coincides with the mirrors screw axis and the x-axis lies on mirror $\mathbf{n}_1$. Proposition 7 provides a method to estimate the matrix $\mathbf{R}_w^c$, while proposition 8 presents two methods for computing the projection of $\mathbf{t}_w^c$ on the plane defined by the camera centers.

Consider the setup reported in Figs. 7(a) and 7(b). If at least two corresponding points exist between $\langle c \rangle$ and $\langle v_1 \rangle$ and between $\langle c \rangle$ and $\langle v_2 \rangle$, then the fundamental matrices $\mathbf{F}^{[1]}$ and $\mathbf{F}^{[2]}$ exist. Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be the epipoles having unitary norm considered in proposition 5, $\mathbf{z}_{w(c)} \triangleq \mathbf{e}_1 \times \mathbf{e}_2$ and $\mathbf{x}_{w(c)} \triangleq \mathbf{e}_1 \times \mathbf{z}_{w(c)}$.

Proposition 7 ($\mathbf{R}_w^c$ estimation). *For every rigid-motion* $(\mathbf{R}_w^c, \mathbf{t}_w^c)$*, the following equation holds true,*

$$\mathbf{R}_w^c = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{n}_{\Gamma(c)} & \mathbf{e}_1 & \mathbf{a}_{(c)} \end{bmatrix}^{-1} \tag{17}$$

Proof. Due to the assumption that the x-axis of $\langle w \rangle$ lies on $\mathbf{n}_1$, then for every pose of $\langle c \rangle$ the world y-axis expressed in the camera frame $\langle c \rangle$ corresponds to $\mathbf{e}_1$. Consequently, for every pose of $\langle c \rangle$, $\mathbf{e}_1$ and $\mathbf{e}_2$ lie on the same plane Γ defined by the three camera centers having normal vector $\mathbf{n}_{\Gamma_{(w)}} = [0\ 0\ 1]^T$ in $\langle w \rangle$ (see Fig. 7(a)), then,

$$\mathbf{z}_{w(c)} \triangleq \mathbf{e}_1 \times \mathbf{e}_2,$$

where $\mathbf{z}_{w(c)}$ is the z-axis of the world reference frame expressed in the camera frame $\langle c \rangle$. The world frame x-axis can be easily obtained as the cross product of $\mathbf{e}_1$ and $\mathbf{z}_{w(c)}$,

$$\mathbf{x}_{w(c)} \triangleq \mathbf{e}_1 \times \mathbf{z}_{w(c)}.$$

Finally we have that,

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \mathbf{R}_w^c \begin{bmatrix} \mathbf{x}_{w(c)} & \mathbf{e}_1 & \mathbf{z}_{w(c)} \end{bmatrix},$$

from which (17) follows. □

Note that, in absence of noise on the image points, Equation 17 provides us with the *exact* $\mathbf{R}_w^c$. In the case of noisy data, the estimated $\mathbf{R}_w^c$ will not be, in general, a rotation matrix. To overcome this problem, the correct rotation matrix should be computed as $\mathbf{U}\mathbf{V}^T$, where matrices $\mathbf{U}$ and $\mathbf{V}$ are obtained from the singular value decomposition of the initial estimate of $\mathbf{R}_w^c$.

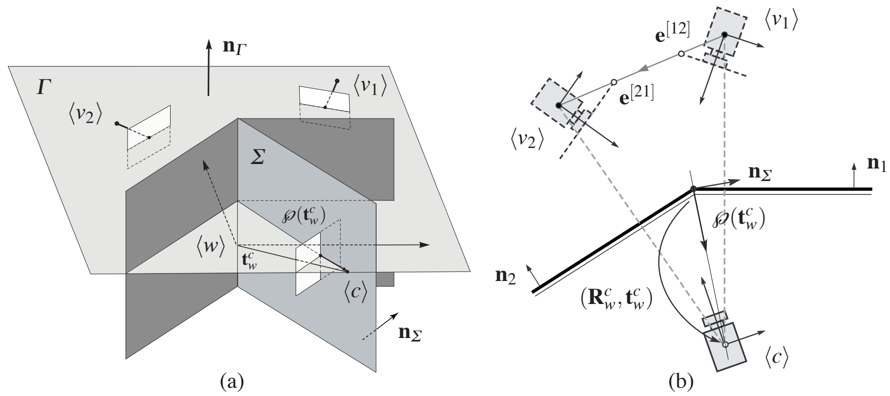


Fig. 7 (a) $\wp(\mathbf{t}_w^c)$ is the projection of $\mathbf{t}_w^c$ on the plane Γ; and (b) computation of $\wp(\mathbf{t}_w^c)$ using the epipolar lines and Σ.

Remark 4. In section 3.2 we have seen that Equation 13 relates the pose of two cameras $\langle c_1 \rangle$ and $\langle c_2 \rangle$ with the pose of their virtual counterparts. An analogous equation, that will be instrumental in proving the second statement in proposition 8, relates the pose of $\langle c \rangle$ and $\langle w \rangle$ with that of the virtual cameras $\langle v_1 \rangle$ and $\langle v_2 \rangle$. In fact, as from Fig. 6, assuming that the camera frame $\langle c_1 \rangle$ is coincident with $\langle w \rangle$ and $\langle c_2 \rangle$ is coincident with $\langle c \rangle$, we get,

$$\mathbf{H}_{v_1}^{v_2} = \mathbf{H}_c^w \, \mathbf{D}_w^{[1]} \, \mathbf{D}_w^{[2]} \, \mathbf{H}_w^c. \tag{18}$$

By inverting and then premultiplying (18) by $\mathbf{H}_w^c$, we obtain the *Sylvester equation* (with unknown $\mathbf{H}_w^c$),

$$\mathbf{H}_w^c \, \mathbf{H}_{v_2}^{v_1} = \mathbf{D}_w^{[2]} \, \mathbf{D}_w^{[1]} \, \mathbf{H}_w^c, \tag{19}$$

where $\mathbf{D}_w^{[1]}$ and $\mathbf{D}_w^{[2]}$ are the reflection transformations about the two mirrors written in $\langle w \rangle$,

$$\mathbf{D}_w^{[1]} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{D}_w^{[2]} = \begin{bmatrix} \mathbf{S}_w^{[2]} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix},$$

being $\mathbf{n}_1 = [0 \;\; 1 \;\; 0]^T$ and $\mathbf{n}_2 = [-\sin\theta \;\; \cos\theta \;\; 0]^T$ in $\langle w \rangle$.

Consider the setup in Fig. 7(a) and let $\wp(\mathbf{t}_w^c) \in \mathbb{R}^2$ be the vector containing the first two components of the projection of $\mathbf{t}_w^c$ on the plane Γ. In the next proposition we present two methods for estimating the direction of $\wp(\mathbf{t}_w^c)$, i.e., $\frac{\wp(\mathbf{t}_w^c)}{\|\wp(\mathbf{t}_w^c)\|}$. The first one uses the fundamental matrix $\mathbf{F}^{[12]}$ between the virtual cameras $\langle v_1 \rangle$, $\langle v_2 \rangle$ and the second one the Equation 19. Let Σ be the plane defined by the mirrors screw axis and the center of $\langle c\rangle$, and let $\mathbf{n}_\Sigma$ be the normal to Σ (see Fig. 7(b)). In the next proposition $\mathbf{R}_z(\gamma)$ denotes a rotation about the z-axis by an angle γ.

Proposition 8 (Estimation of $\frac{\wp(\mathbf{t}_w^c)}{\|\wp(\mathbf{t}_w^c)\|}$). *The direction of $\wp(\mathbf{t}_w^c)$ is given by,*

$$\frac{\wp(\mathbf{t}_w^c)}{\|\wp(\mathbf{t}_w^c)\|} = \frac{\mathbf{t}_w^{c*}}{\|\mathbf{t}_w^{c*}\|}, \tag{20}$$

where,
(i)

$$\mathbf{t}_w^{c*} = [\mathbf{a}(1)\ \ \mathbf{a}(2)]^T, \tag{21}$$

with $\mathbf{a} = \mathbf{R}_z(-90^\circ)\,\mathbf{n}_\Sigma$, $\mathbf{n}_\Sigma(1) > 0$*, or alternatively,*
(ii)

$$\mathbf{t}_w^{c*} = \left[\frac{\mathbf{a}(1)(1-\mathbf{n}_2^2(2)) + \mathbf{n}_2(1)\mathbf{n}_2(2)\mathbf{a}(2)}{2\,\mathbf{n}_2^2(1)} \quad \frac{\mathbf{a}(2)\mathbf{n}_2(1) - \mathbf{a}(1)\mathbf{n}_2(2)}{2\,\mathbf{n}_2(1)}\right]^T, \tag{22}$$

with $\mathbf{a} = -\mathbf{R}_w^c\mathbf{t}_{v_2}^{v_1}$*, where* $\mathbf{n}_2$ *is the normal vector to the second mirror, written in* $\langle w\rangle$.

Proof. Let us start by proving the part (i) of the statement. $\mathbf{t}_w^c$ and $\wp(\mathbf{t}_w^c)$ lie on the plane Σ with normal vector $\mathbf{n}_\Sigma = [\mathbf{n}_\Sigma(1)\ \ \mathbf{n}_\Sigma(2)\ \ 0]^T$ in $\langle w\rangle$. Since the corresponding epipolar lines (computed from the fundamental matrix $\mathbf{F}^{[12]}$) all intersect at a single image line $\mathbf{m}$, projection of the screw axis, we have that,

$$\mathbf{n}_\Sigma = \mathbf{R}_w^c\,(\mathbf{K}^T\mathbf{m}),$$

where $\mathbf{K}^T\mathbf{m}$ is the normal vector to Σ in $\langle c\rangle$. Since $\mathbf{n}_\Sigma \perp \wp(\mathbf{t}_w^c)$, it is then sufficient to rotate $\mathbf{n}_\Sigma$ of -90° around the z-axis in order to obtain $\mathbf{t}_w^{c*}$ in (21).

To prove the second part of the statement, consider the Equation 19. Let $\mathbf{R}_w^c$ be given (computed, for example, using proposition 7) and $\mathbf{D}_w^{[2]}$, $\mathbf{n}_2$ be estimated using one of the algorithms in section 4.1. Let $\mathbf{H}_{v_2}^{v_1}$ also be given (see section 3.1). From (19) we have that $\mathbf{R}_w^c\mathbf{t}_{v_2}^{v_1} + \mathbf{t}_w^c = \mathbf{S}_w^{[2]}\,\mathbf{S}_w^{[1]}\,\mathbf{t}_w^c$. Collecting $\mathbf{t}_w^c$ on the left-hand side of the equation, we get $(\mathbf{I} - \mathbf{S}_w^{[2]}\,\mathbf{S}_w^{[1]})\,\mathbf{t}_w^c = -\mathbf{R}_w^c\mathbf{t}_{v_2}^{v_1}$, that is,

$$\begin{bmatrix} 2\,\mathbf{n}_2^2(1) & -2\,\mathbf{n}_2(1)\,\mathbf{n}_2(2) & 0 \\ 2\,\mathbf{n}_2(1)\,\mathbf{n}_2(2) & 2\,(1-\mathbf{n}_2(2))^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{t}_w^c \;=\; -\mathbf{R}_w^c\mathbf{t}_{v_2}^{v_1},$$

from which, after few manipulations, we obtain (22). Note that the normalization of $\mathbf{t}_w^{c*}$ in (20), removes the ambiguity due to the up to scale estimation of $\mathbf{t}_{v_2}^{v_1}$. □

5 Simulation and experimental results

5.1 Simulations

Simulation experiments have been conducted with the *epipolar geometry toolbox* [14], in order to test the effectiveness of the algorithms presented in the previous sections. The setup is composed of a pinhole camera with calibration matrix,

$$\mathbf{K} = \begin{bmatrix} 951.8 & 0 & 640.66 \\ 0 & 951.8 & 605.11 \\ 0 & 0 & 1 \end{bmatrix},$$

and two planar mirrors with normal vectors $\mathbf{n}_1 = [0.6\ 0.55\ -0.57]^T$ and $\mathbf{n}_2 = [-0.32\ 0.87\ -0.37]^T$, corresponding to an angular displacement $\theta = 60^\circ$. The camera observes directly a total of 20 random points. For the sake of simplicity, we will henceforth suppose that the correspondence matching problem is exact. In order to solve the correspondence problem in practice, one might first use SIFT (*scale-invariant feature transform*) [12] and, in a second phase, assign the points to each mirror which has been uniquely identified by a colored landmark.

The rigid motion $(\mathbf{R}_w^c, \mathbf{t}_w^c)$ between the camera and world frame is,

$$\mathbf{R}_w^c = \mathbf{R}_z(40^\circ)\,\mathbf{R}_y(-20^\circ)\,\mathbf{R}_x(30^\circ),\ \ \mathbf{t}_w^c = [0.6\ \ -2\ \ -0.2]^T\ \text{m}.$$

Four well-known fundamental matrix estimators [1], *normalized 8-point algorithm*, *iterative Sampson method*, *robust Torr's M-estimator* and *LMedS*, have been compared in the simulation experiments. The M-estimator revealed the best compromise between estimation accuracy, computation load and robustness to uncorrect correspondence matching due to outliers.

In Fig. 8, we ran 200 iterations of the proposed algorithms for increasing value of the image noise standard deviation. Fig. 8(a) reports the mirror angle estimation error $|\widehat{\theta} - \theta|$ for the algorithms in propositions 5 and 6. The algorithm that uses the virtual epipoles is the most accurate with a maximum mean error of 1.5°. Fig. 8(b) shows the $\mathbf{R}_w^c$ estimation error computed as the roll-pitch-yaw angle errors (cf. proposition 7). We experienced a good robustness to noise of the proposed method, with a maximum mean error around 2°. Fig. 8(c) finally reports the estimation error,

$$\varepsilon(\hat{\mathbf{t}}_w^c) = \left\| \frac{\wp(\hat{\mathbf{t}}_w^c)}{\|\wp(\hat{\mathbf{t}}_w^c)\|} - \frac{\wp(\mathbf{t}_w^c)}{\|\wp(\mathbf{t}_w^c)\|} \right\|,$$

for the two algorithms in proposition 8: (*i*) *screw axis* and (*ii*) *Sylvester equation*. From an inspection of Fig. 8(c), we observe that the second method is less sensitive to noise since it does not rely on the direct computation of the epipoles. A maximum mean error of 5 cm is achieved.

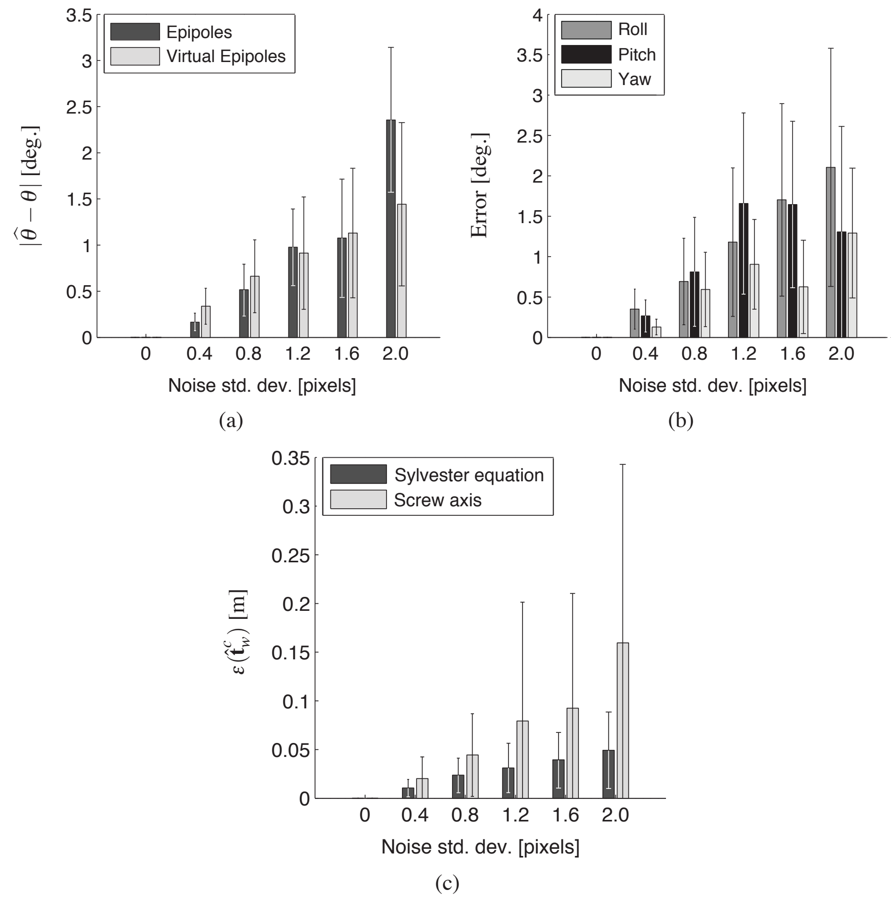


Fig. 8 Simulations: (a) mirror angle estimation error $|\widehat{\theta}-\theta|$ for the algorithms in propositions 5 and 6; (b) $\mathbf{R}_w^c$ estimation error for the algorithm in proposition 7; and (c) error $\varepsilon(\hat{\mathbf{t}}_w^c)$ for the algorithms in proposition 8.

5.2 Experiments

In order to test the robustness of proposed algorithms in real scenarios we performed a series of experiments using the setup shown in Fig. 9. A Lumenera® LU071C camera mounted on the end-effector of a 6 axes KUKA® KR 3 manipulator, observes a structured scene directly and reflected through two planar mirrors. The mirrored objects lie in a box of size 0.42×0.3 m^2.

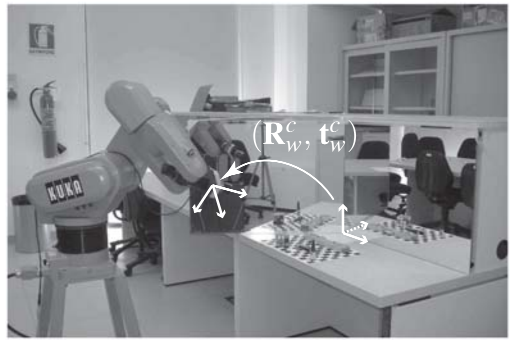


Fig. 9 Experimental setup: an eye-in-hand robot observes a structured 3-D scene directly and reflected through two planar mirrors.

We exploited the high positioning accuracy of the KR 3 manipulator in order to have precise ground truth reference measurements of position and orientation of the camera. The calibration matrix $\mathbf{K}$ of the camera is the same as in the simulations[1].

Fig. 10(a) shows the mirror angle estimation error $|\widehat{\theta} - \theta|$ for $\theta \in \{30°, 45°, 60°\}$. The values reported in Fig. 10(a) are the average of 5 experiments conducted for each θ. In confirmation of the simulation results, the algorithm that exploits the virtual cameras epipoles is the most accurate, with a maximum mean error of about 6°. Fig. 10(b) reports the $\mathbf{R}_w^c$ estimation error computed as the roll-pitch-yaw angle errors, for the six rotational displacements reported in Table 1, ($\theta = 60°$). As already pointed out in section 5.1, the proposed algorithm has a reduced sensitivity to noise: the maximum error is less than 6°. Finally, Fig. 10(c) depicts the estimation error $\varepsilon(\widehat{\mathbf{t}}_w^c)$, for the six translational displacements reported in Table 1. As in the simulation experiments, the algorithm based on the Sylvester equation exhibits the best performance, with a maximum error of about 7 cm.

Table 1 Rotational and translational displacements considered in Figs. 10(b) and 10(c). The mirror angle $\theta = 60°$.

	$\mathbf{R}_w^c$			$\mathbf{t}_w^c$
	Roll [deg.]	*Pitch* [deg.]	*Yaw* [deg.]	[m]
a	−4.00	−3.03	−37.07	$[\,0.147\ \ -0.824\ \ 0.377\,]^T$
b	8.16	−3.49	−43.13	$[\,0.157\ \ -0.824\ \ 0.610\,]^T$
c	13.99	−20.87	−40.57	$[\,0.432\ \ -0.960\ \ 0.532\,]^T$
d	−6.38	−9.01	−51.60	$[\,0.106\ \ -0.613\ \ 0.408\,]^T$
e	−3.63	−3.46	−43.73	$[\,0.215\ \ -0.824\ \ 0.552\,]^T$
f	7.74	−13.55	−35.60	$[\,0.313\ \ -0.960\ \ 0.535\,]^T$

[1] The calibration matrix has been estimated with the *camera calibration toolbox* [2].

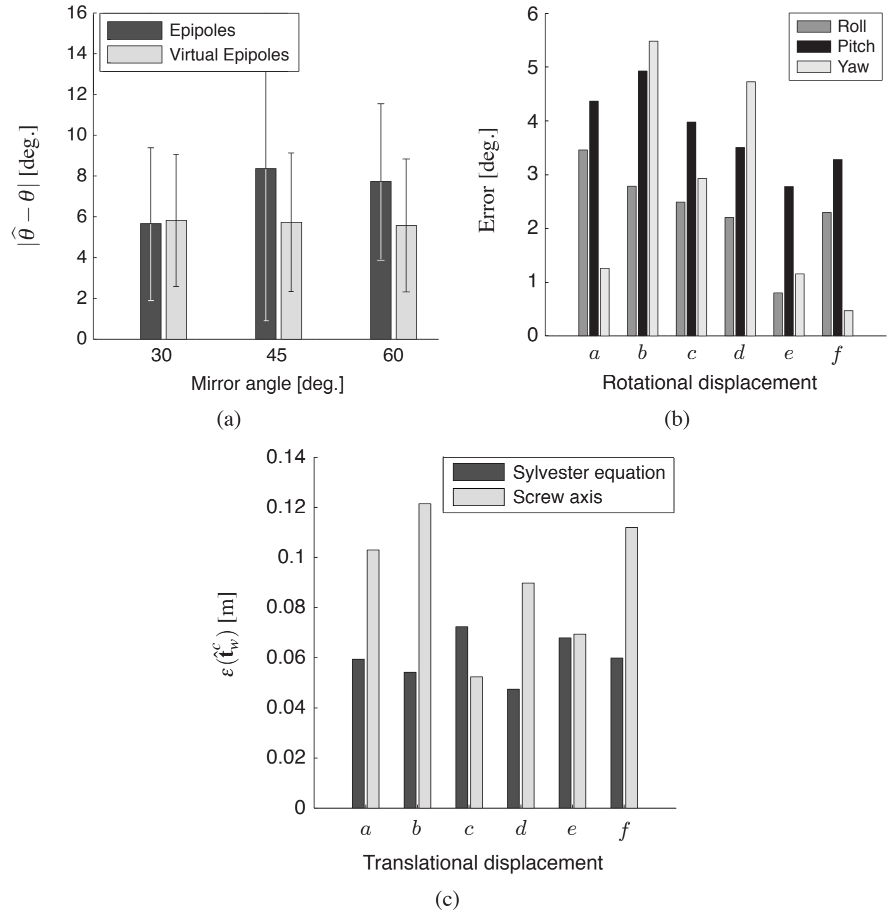


Fig. 10 Experiments: (a) mirror angle estimation error $|\widehat{\theta} - \theta|$ for $\theta \in \{30°, 45°, 60°\}$; (b) $\mathbf{R}_w^c$ estimation error for the rotational displacements reported in Table 1; and (c) error $\varepsilon(\hat{\mathbf{t}}_w^c)$ for the translational displacements reported in Table 1.

6 Conclusions and future work

The paper presents a systematic study of the imaging geometry of *planar catadioptric stereo* vision sensors. New algebraic results on the multiple view geometry for the case of static and moving cameras (multi-view PCS) are introduced. On the basis of these new results, we addressed both the image-based camera localization and mirror calibration problems. Extensive simulation and experimental results have been presented in order to illustrate the theory and show the applicability of the proposed algorithms in real scenarios.

The extension of our results to multiple camera networks is a subject of future research. Future investigations will also focus on the 3-D scene reconstruction problem and will show the connection between PCS sensors and the trifocal geometry [9].

Acknowledgements The authors are grateful to Francesco Chinello for his help in the preparation of the experimental setup.

References

1. X. Armangué and J. Salvi. Overall View Regarding Fundamental Matrix Estimation. *Image Vision Comput.*, 21(2):205–220, 2003.
2. J.Y. Bouguet. *Camera Calibration Toolbox for Matlab*. Available from the page: http://www.vision.caltech.edu/bouguetj, 2007.
3. T.E. Boult and S.K. Nayar. *A Handy Stereo Camera*. University of Columbia and Colorado, http://vast.uccs.edu/ tboult/VSAM.html#stereo, 2000.
4. C. Geyer and K. Daniilidis. Catadioptric Camera Calibration. In *Proc. 7th IEEE Int. Conf. Comp. Vis.*, volume 1, pages 398–404, 1999.
5. J. Gluckman and S.K. Nayar. Planar Catadioptric Stereo: Geometry and Calibration. In *Proc. IEEE Conf. Comp. Vis. and Pattern Rec.*, volume 1, pages 22–28, 1999.
6. J. Gluckman and S.K. Nayar. Catadioptric Stereo using Planar Mirrors. *Int. J. Comput. Vision*, 44(1):65–79, 2001.
7. J. Gluckman and S.K. Nayar. Rectified Catadioptric Stereo Sensors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(2):224–236, 2002.
8. A. Goshtasby and W. Gruver. Design of a Single-Lens Stereo Camera System. *Pattern Recogn.*, 26(6):923–937, 1993.
9. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004.
10. W. Hong, A. Yang Yang, K. Huang, and Y. Ma. On Symmetry and Multiple-View Geometry: Structure, Pose, and Calibration from a Single Image. *Int. J. Comput. Vision*, 60(3):241–265, 2004.
11. J. Konrad. View Reconstruction for 3-D Video Entertainment: Issues, Algorithms and Applications. In *Proc. IEEE Int. Conf. Image Proc. and Appl.*, volume 1, pages 8–12, 1999.
12. D.G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision*, 20(2):91–110, 2004.
13. Y. Ma, S. Soatto, J. Kosecká, and S.S. Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*. Interdisciplinary Applied Mathematics. Springer, 2003.
14. G.L. Mariottini and D. Prattichizzo. EGT for Multiple View Geometry and Visual Servoing: Robotics Vision with Pinhole and Panoramic Cameras. *IEEE Robot. Autom. Mag.*, 12(4):26–39, 2005.
15. G.L. Mariottini and D. Prattichizzo. Image-based Visual Servoing with Central Catadioptric Camera. *Int. J. Robot. Res.*, 27(1):41–57, 2008.
16. G.L. Mariottini, S. Scheggi, F. Morbidi, and D. Prattichizzo. Planar Catadioptric Stereo: Single and Multi-View Geometry for Calibration and Localization. In *Proc. IEEE Int. Conf. Robot. Automat*, pages 1510–1515, 2009.
17. H. Mitsumoto, S. Tamura, K. Okazaki, N. Kajimi, and Y. Fukui. 3-D Reconstruction using Mirror Images based on a Plane Symmetry Recovering Method. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(9):941–946, 1992.
18. S.K. Nayar. Catadioptric Omnidirectional Camera. In *Proc. IEEE Conf. Comp. Vis. and Pattern Rec.*, pages 482–488, 1997.

19. S.A. Nene and S.K. Nayar. Stereo with Mirrors. In *Proc. 6th IEEE Int. Conf. Comp. Vis.*, pages 1087–1094, 1998.
20. T.P. Pachidis and J. Lygouras. A Pseudo Stereo Vision System as a Sensor for Real-Time Path Control of a Robot. In *Proc. 19th IEEE Conf. Instr. and Measurem. Tech.*, volume 2, pages 1589–1594, 2002.
21. T. Svoboda and T. Pajdla. Epipolar Geometry for Central Catadioptric Cameras. *Int. J. Comput. Vision*, 49(1):23–37, 2002.
22. H.-H.P. Wu. Rectification of Stereoscopic Video for Planar Catadioptric Stereo Systems. *IEEE Trans. Circ. Syst. Video Tech.*, 17(6):686–698, 2007.