Planar Motion Estimation from 1D Homographies

Miguel Aranda, Gonzalo López-Nicolás and Carlos Sagüés

Instituto de Investigación en Ingeniería de Aragón (I3A)

Universidad de Zaragoza

Zaragoza, Spain

Email: {marandac,gonlopez,csagues}@unizar.es

Abstract—**This paper addresses the estimation of planar camera motion using 1D homographies. As contributions, we show analytically that, contrary to what occurs with the 2D homography, there is a family of infinite solutions to the 1D homography decomposition, and therefore infinite possible motion reconstructions. In addition, we propose a new method to compute the planar motion between two images from the information provided by two different 1D homographies, employing their associated homology transformations. Therefore, our approach computes a general planar camera motion from only two 1D views, when previous works needed three 1D views for this task. The use of 1D information makes the method particularly suitable for omnidirectional cameras, due to the wide field of view and precise angular information provided by this kind of sensors. The performance of our proposal is illustrated through simulations and experiments on real images.**

I. Introduction

Multiple view geometry [1] is a powerful tool suitable for a number of robotic tasks involving the use of visual information. In particular, the geometry of two views (known as epipolar geometry) has been classically used to reconstruct scene structure and camera motion. When computing the motion, if the scene is planar or the two camera centers are very close, the epipolar geometry fails to give a correct solution [2] and then a different model must be used. This alternative model may be the homography induced by a plane, or planar homography. In addition to being complementary to the epipolar geometry, the planar homography can be estimated from a smaller number of matched scene features, and is therefore faster to compute.

The motion between two views can be determined from the homography matrix through its decomposition [3]. It is typical in robotics to consider planar camera motion (for instance, when a robot moves on the floor), a scenario which adds constraints to the homography that can make the decomposition simpler [4], [5]. This case can be further particularized by considering vertical planes (which abound in man-made environments), exploiting the additional constraints this poses on the homography [6], [7].

The case of planar motion can be reduced to a representation in terms of 1D cameras observing a 2D scene [8]. 1D cameras provide a convenient and simplified description in this situation [9], and multi-view constraints can analogously

be defined on them. The 1D trifocal tensor [8] is a three-view constraint independent of scene structure that has been used to extract planar motion information [10]–[12]. Similarly to the 2D case, it is also possible to define a homography mapping between two 1D views, this time induced by a line in the 2D scene. The 1D homography has been employed for motion estimation purposes as well, by assuming certain simplifying approximations regarding the motion and the scene [6] or considering a circular motion [13].

In this paper, we address the use of 1D homography to perform planar camera motion estimation. As a first contribution, we prove analytically that the computation of camera motion using a single 1D homography provides an infinite family of valid solutions. We propose as further contribution a method which gets around this issue by using the information of two homographies through their associated homologies. Therefore, our method computes a general planar camera motion from only two 1D views, instead of the three 1D views required for this task in previous approaches.

We show the performance of our proposal applied on a robot equipped with an omnidirectional camera, moving on a planar ground. The omnidirectional vision sensor is particularly appropriate. The reasons are its wide field of view, which facilitates the detection of multiple scene planes, and the precise angular information it provides, which means it fits very well into a 1D imaging formulation considering only the angular coordinate. The advantages of our approach are the simplicity of the reduced 1D image representation, and the robustness and precision associated to the use of angular visual information.

II. Motion from the 2D Homography

This section is aimed at providing background on existing methods to perform motion estimation using the 2D homography.

A planar surface in a 3D scene induces a unique projective transformation, called homography, that relates the projections in two views of any given point belonging to the plane. As a mapping between a pair of image points in homogeneous coordinates, the planar homography is expressed by a 3×3 matrix. Since it is a general projective transformation defined up to scale (due to the scale ambiguity inherent to perspective projection), the homography matrix has 8 degrees of freedom. Expressing the homography mapping for one scene

point provides two independent equations. Therefore, from a minimum number of four matched points the homography can be estimated linearly by solving the resulting system of equations [1].

The planar homography mapping can be defined between image points expressed in pixels. Then, we obtain an uncalibrated homography $\mathbf{H_u^{2D}}$. If the calibration matrix $\mathbf{K}$ is known, a calibrated homography (i.e. a homography relating points in calibrated coordinates) can be computed as follows:

$$\mathbf{H_c^{2D}} = \mathbf{K}^{-1}\mathbf{H_u^{2D}}\mathbf{K}. \tag{1}$$

This calibrated homography matrix encapsulates the reconstruction of the relative location of the views (i.e. the camera motion) and the unit normal of the plane in the following way:

$$\mathbf{H_c^{2D}} = \lambda(\mathbf{R} + \mathbf{t}\mathbf{n}^T), \tag{2}$$

where $\mathbf{R} \in \mathbb{R}^{3\times3}$ is the rotation matrix, $\mathbf{t} \in \mathbb{R}^3$ is the translation vector (scaled by the distance to the plane), and $\mathbf{n} \in \mathbb{R}^3$ is the unit normal of the plane. This homography, which relates homogeneous coordinates of points in two camera frames, is defined up to a scalar factor λ. In order to extract the camera motion from the homography matrix, it is necessary to decompose it. This is addressed in the following section.

A. Decomposition of the 2D Homography

The objective of the planar 2D homography decomposition is to obtain the camera motion. The Euclidean homography matrix directly expresses the reconstruction of the relative motion between the views and the normal of the plane. As such, it is a unique matrix (up to sign), not defined up to a scalar factor [2]:

$$\mathbf{H_e^{2D}} = \mathbf{R} + \mathbf{t}\mathbf{n}^T. \tag{3}$$

However, when computing the homography from image feature correspondences in homogeneous calibrated coordinates, what we obtain is the Euclidean homography multiplied by a factor (2). As shown in [2], if a matrix has the form of a real 3×3 Euclidean homography matrix (3), then its second largest singular value is equal to one. Conversely, any matrix having its second largest singular value not equal to one is not a valid Euclidean homography. Taking into account that multiplying a matrix by a scale factor causes its singular values to be multiplied by the same factor, it can be concluded that for a given computed calibrated homography, we can obtain a unique Euclidean homography matrix (up to sign), by dividing the computed homography matrix by its second largest singular value. The sign ambiguity can be solved by employing the positive depth constraint.

When computing the camera motion ($\mathbf{R}$ and $\mathbf{t}$) from the Euclidean homography, four possible solutions are obtained. Two of them can be rejected by imposing the positive depth constraint in conventional cameras, leaving two physically valid solutions for the relative camera motion [2]. A complete procedure for the computation of the motion from a calibrated 2D homography is outlined in algorithm 1 [2], [14].

Algorithm 1 Computation of motion from 2D homography

1) Compute $\mathbf{H_c^{2D}} = \mathbf{K}^{-1}\,\mathbf{H_u^{2D}}\,\mathbf{K}$
2) Compute the SVD of $\mathbf{H_c^{2D}}$ such that $\mathbf{H_c^{2D}} \sim \mathbf{H_e^{2D}} = \mathbf{U}\,diag(\lambda_1, \lambda_2, \lambda_3)\,\mathbf{V}^T$ with $\lambda_2 = 1$
3) Let $\alpha = \sqrt{\frac{\lambda_1^2 - \lambda_2^2}{\lambda_3^2 - \lambda_1^2}}$, $\beta = \sqrt{\frac{\lambda_2^2 - \lambda_3^2}{\lambda_3^2 - \lambda_1^2}}$
4) Writing $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$, compute $\mathbf{v}_v = \alpha\,\mathbf{v}_1 \pm \beta\,\mathbf{v}_3$
5) Compute $\mathbf{R} = [\mathbf{H_e^{2D}}\,\mathbf{v}_v, \quad \mathbf{H_e^{2D}}\,\mathbf{v}_2, \quad \mathbf{H_e^{2D}}\,\mathbf{v}_v \times \mathbf{H_e^{2D}}\,\mathbf{v}_2][\mathbf{v}_v, \mathbf{v}_2, \mathbf{v}_v \times \mathbf{v}_2]^T$
6) Compute $\mathbf{t} = \mathbf{H_e^{2D}}\,\mathbf{n} - \mathbf{R}\,\mathbf{n}$, with $\mathbf{n} = \mathbf{v}_v \times \mathbf{v}_2$

B. Decomposition of the 2D homography with planar motion

When the camera motion is planar, there are additional constraints on the Euclidean 2D homography matrix that can be exploited to obtain its decomposition in a simpler way [4], [5]. This is an important case in practice, since it applies for example to the situation of a camera being carried by a vehicle moving on the ground plane. The planar camera motion is then given by a single angle ϕ expressing the rotation around the vertical axis, and a translation vector (scaled by the distance to the plane) lying in the plane of the cameras: $\mathbf{t} = (t_1, t_2, 0)^T$. Considering a unitary plane normal in an arbitrary direction, i.e. $\mathbf{n} = (n_1, n_2, n_3)^T$, the Euclidean homography matrix in this case turns out to be of the following form:

$$\mathbf{H_e^{2D}} = \begin{bmatrix} \cos\phi + t_1 n_1 & -\sin\phi + t_1 n_2 & t_1 n_3 \\ \sin\phi + t_2 n_1 & \cos\phi + t_2 n_2 & t_2 n_3 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

If we consider the further constraint that the plane inducing the homography is vertical (which is a reasonable hypothesis, as vertical planes abound in man-made scenarios), then $n_3 = 0$ and $\mathbf{H_e^{2D}}$ is as follows:

$$\mathbf{H_e^{2D}} = \begin{bmatrix} \cos\phi + t_1 n_1 & -\sin\phi + t_1 n_2 & 0 \\ \sin\phi + t_2 n_1 & \cos\phi + t_2 n_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5}$$

The computation of camera motion from the 2D homography in this latter case has been addressed in several works [7], [15]. In all cases, it is required as first step to compute a calibrated 2D homography $\mathbf{H_c^{2D}}$, employing a set of feature correspondences and the knowledge of the internal camera calibration. When the motion is planar, this homography is normalized in a straightforward manner, dividing by entry $\mathbf{H_c^{2D}}(3,3)$. This way, one obtains a Euclidean homography matrix of the form of (4), which can be directly decomposed to find the camera motion parameters. We provide in algorithm 2 a procedure based on the work [7] to compute the camera displacement in the case of planar motion.

When the motion between two views is planar, the description can be reduced to 1D cameras imaging a 2D space [8], [9]. This representation appears natural and may provide some advantages due to its simplicity. The computation of the camera motion can then be addressed using 1D cameras and 1D homographies. This is discussed in the next section.

Algorithm 2 Computation of planar motion from 2D homography

1) Compute a calibrated homography $\mathbf{H_c^{2D}} = \mathbf{K^{-1}} \mathbf{H_u^{2D}} \mathbf{K}$
2) Normalize $\mathbf{H_c^{2D}}$ to obtain the Euclidean homography $\mathbf{H_e^{2D}} = [(h_{11}, h_{21}, 0)^T, (h_{12}, h_{22}, 0)^T, (h_{13}, h_{23}, 1)^T]$
3) Compute $\mathbf{R}(\phi)$, with $\phi = atan2(h_{12} - h_{21}, -h_{11} - h_{22}) \pm \arccos \frac{h_{12}h_{21} - h_{11}h_{22} - 1}{\sqrt{(h_{12} - h_{21})^2 + (h_{11} + h_{22})^2}}$
4) For each of the two solutions for ϕ, compute the angle of translation in the motion plane, $\psi = \arctan(t_2/t_1)$, using either $\psi = \arctan \frac{h_{21} - \sin\phi}{h_{11} - \cos\phi}$ or $\psi = \arctan \frac{h_{22} - \cos\phi}{h_{12} + \sin\phi}$. $\psi + \pi$ is also a solution
5) Four solutions for (ϕ, ψ) result. The correct one can be chosen using a priori and/or additional information. This may include imposing the positive depth constraint, using a second plane and finding the compatible solution, or considering the properties of optical flow

III. Motion from the 1D homography

In the same way a 2D camera projects points in the 3D projective space to points in the 2D one, a 1D camera can be defined by considering the projection from a 2D to a 1D space [9]. A 1D camera is useful to model the situation where the camera moves on a planar surface [8]. In this case, the 2D retina can be reduced to a 1D retina observing the 2D scene generated by the orthogonal projection of the 3D scene on the motion plane of the cameras. For conventional cameras, considering an axis in the 2D image that is parallel to the motion plane, the 1D representation of a 2D image point is given by its coordinate in that axis: $(x, 1)^T$. In the case of an omnidirectional camera, assuming that the optical axis is perpendicular to the motion plane, the 1D coordinates of a given 2D image point are obtained from the angle to that point (α) measured from the center of the image. The point is then represented as $(\sin(\alpha), \cos(\alpha))^T$ in 1D homogeneous coordinates. Note that this generates an ambiguity, since 2D image points whose angles differ by π are mapped to the same 1D point.

It is possible to define in 1D a transformation analogous to the planar homography of 2D. This transformation, which we call 1D homography, defines a mapping between points in two 1D images induced by a line in the 2D scene. Thus, it is a general projective transformation between 1D spaces, and is therefore expressed by a 2×2 matrix $\mathbf{H}$ with three degrees of freedom:

$$\mathbf{x_2} = \mathbf{H}\mathbf{x_1}, \tag{6}$$

where $\mathbf{x_2}$ and $\mathbf{x_1}$ are the homogeneous 1D coordinates of the projections in two views of a point in the line that induces the homography. Since every two-view point correspondence in 1D provides one equation, the 1D homography can be estimated linearly from a minimum of three point matches.

In a situation where the camera motion occurs on a horizontal surface, vertical planes surrounding the sensor become lines in the 2D scene defined by the horizontal plane passing through the camera centers. Then, each of these planes, which may be real or virtual, induces a 1D homography between two views.

As is the case with the planar 2D homography, the 1D homography induced by a line encapsulates the camera motion information. It is therefore a tool that can be employed for planar motion estimation purposes. We note that the method proposed here to estimate the camera displacement can be used indistinctly for conventional cameras and for omnidirectional ones. However, when using only 1D information, omnidirectional cameras are better suited due to their unconstrained field of view and the precise angular information they provide.

A. Decomposition of the 1D homography

From a minimum number of three point matches between two 1D images corresponding to scene points in a line, we can estimate the homography induced by that line. If the image points are expressed in calibrated projective coordinates, the computed homography can be expressed as:

$$\mathbf{H_c} = \lambda(\mathbf{R} + \mathbf{t}\mathbf{n}^T), \tag{7}$$

where λ is a scale factor, $\mathbf{R} \in \mathbb{R}^{2 \times 2}$ is the rotation matrix, $\mathbf{t} \in \mathbb{R}^2$ is the translation vector (scaled by the distance to the line), and $\mathbf{n} \in \mathbb{R}^2$ is the unit normal of the line:

$$\mathbf{R} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \quad \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \quad \mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}. \tag{8}$$

We can define the Euclidean 1D homography in the same manner as in 2D, i.e. as a homography that expresses directly the camera motion parameters (which can be obtained through the decomposition of the matrix):

$$\mathbf{H_e} = \mathbf{R} + \mathbf{t}\mathbf{n}^T. \tag{9}$$

We will show next that, contrarily to what occurs with the 2D homography, there is in general a family of infinite valid Euclidean 1D homographies induced by a line between two 1D images, which means that there is a family of infinite valid motion reconstructions.

Proposition 1: Let $\mathbf{H_e} \in \mathbb{R}^{2 \times 2}$ be a Euclidean 1D homography matrix:

$$\mathbf{H_e} = \begin{bmatrix} \cos\phi + t_1 n_1 & -\sin\phi + t_1 n_2 \\ \sin\phi + t_2 n_1 & \cos\phi + t_2 n_2 \end{bmatrix}, \tag{10}$$

with ϕ, t_1, t_2, n_1 and $n_2 \in \mathbb{R}$. Then, the singular values d_1 and d_2, $(d_1 \geq d_2 \geq 0)$, of $\mathbf{H_e}$ are such that $d_1 \geq 1$ and $d_2 \leq 1$.

Proof: Given a real 2×2 matrix $\mathbf{H_e}$ that can be expressed as a Euclidean homography, i.e:

$$\mathbf{H_e} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} \cos\phi + t_1 n_1 & -\sin\phi + t_1 n_2 \\ \sin\phi + t_2 n_1 & \cos\phi + t_2 n_2 \end{bmatrix} \tag{11}$$

with ϕ, t_1, t_2, n_1, n_2 being real values, it can be readily shown that the following inequality holds:

$$(h_{12} - h_{21})^2 + (h_{11} + h_{22})^2 \geq (h_{12}h_{21} - h_{11}h_{22} - 1)^2. \quad (12)$$

This expression is equivalent to:

$$tr(\mathbf{H_e^T H_e}) - det(\mathbf{H_e^T H_e}) - 1 \geq 0. \quad (13)$$

Now, we take into account that $tr(\mathbf{H_e^T H_e}) = \lambda_1 + \lambda_2$ and $det(\mathbf{H_e^T H_e}) = \lambda_1 \cdot \lambda_2$, where λ_1 and λ_2 are the eigenvalues of $\mathbf{H_e^T H_e}$. Then, using the fact that the singular values d_1 and d_2 of $\mathbf{H_e}$ are equal to the square roots of the eigenvalues of $\mathbf{H_e^T H_e}$, the inequality above becomes:

$$d_1^2 + d_2^2 - d_1^2 \cdot d_2^2 - 1 \geq 0. \quad (14)$$

It is straightforward to see that if $d_1 \geq d_2 \geq 0$, this inequality is verified when $d_1 \geq 1$ and $d_2 \leq 1$. ■

Remark 1: Let us analyze the implications of the above proposition. Recall from section II-A that the Euclidean 2D homography (3) had its intermediate singular value equal to one. This meant that it was a unique matrix (up to sign) and could be obtained by multiplying any calibrated 2D homography (2) by the appropriate scale factor. In contrast, proposition 1 does not define any of the singular values of a Euclidean 1D homography as fixed. Rather, it states that in the 1D case, a valid Euclidean homography has its two singular values lying in certain ranges. This implies that when we compute a calibrated 1D homography (7), we cannot obtain from it a unique Euclidean 1D homography (9). Indeed, by multiplying the calibrated homography by any factor that puts its singular values in the defined ranges, one gets a valid Euclidean 1D homography. In particular, for a calibrated 1D homography $\mathbf{H_c}$ having singular values $d_{1c} \geq d_{2c} \geq 0$, from proposition 1 we can obtain the range of values for λ in (7) that give a valid Euclidean homography as follows:

$$d_{2c} \leq \pm \lambda \leq d_{1c}. \quad (15)$$

That is, for all these values of λ, $(1/\lambda) \cdot \mathbf{H_c}$ is a valid Euclidean homography. When this infinite family of homographies are decomposed, they result in an infinite family of valid reconstructions of the camera motion. It is therefore not possible in general to compute the real planar motion from the information given by one single 1D homography. We address this issue by using at least two homographies, associated to two different scene lines.

B. Planar motion from 1D homology

From two 1D homographies $\mathbf{H_1}$ and $\mathbf{H_2}$ between a pair of views, it is possible to compute a projective transformation that maps one of the images (for instance, image 1) to itself, called homology. We can then define two 1D homologies $\mathbf{G_1}$ and $\mathbf{G_2}$ as follows:

$$\mathbf{G_1} = \mathbf{H_2^{-1} H_1} \Rightarrow \mathbf{x_1} \sim \mathbf{G_1 x_1}, \quad (16)$$

$$\mathbf{G_2} = \mathbf{H_2 H_1^{-1}} \Rightarrow \mathbf{x_2} \sim \mathbf{G_2 x_2}, \quad (17)$$

where $\mathbf{x_1}$ is a point of image 1 and $\mathbf{x_2}$ is a point of image 2, both in homogeneous coordinates. Let us now consider $\mathbf{G_1}$ (the following discussion would be equivalent for $\mathbf{G_2}$). If the two lines inducing $\mathbf{H_1}$ and $\mathbf{H_2}$ are different, then one of the eigenvectors of the homology $\mathbf{G_1}$ is equal to the epipole of camera 2 in camera 1. This property (which is also true for a 2D homology) has been used in several works [10], [12], [14]. In the case of 1D images, the other eigenvector of the 2×2 homology matrix is equal to the projection in image 1 of the intersection of the two lines that induce $\mathbf{H_1}$ and $\mathbf{H_2}$. A degenerate situation occurs when the two homographies are induced by the same line. Then, the homology equals the identity and it is not possible to obtain the epipole. This situation can be detected in practice by imposing a lower threshold to the condition number of the homology matrix.

Thus, from two lines in a 2D scene, it is possible to obtain the epipoles in two 1D views by calculating the two homographies $\mathbf{G_1}$ and $\mathbf{G_2}$ and computing their eigenvectors. If the angle of the epipole in camera 1 is α_{12} and the angle of the epipole in camera 2 is α_{21}, then we can compute the camera motion (i.e. the rotation angle, ϕ, and the angle of the translation vector, ψ), considering the frame of camera 1 as the reference, as follows:

$$\phi = \alpha_{12} - \alpha_{21} + \pi \quad (18)$$

$$\psi = \arctan(t_2/t_1) = \alpha_{12}. \quad (19)$$

Notice that this method produces four possible reconstructions of the camera motion, which emerge from the combinations of the two eigenvectors of each of the two homologies. It is usually possible to reject some of these solutions directly, by imposing their coherence with the homography decomposition described in section III-A for $\mathbf{H_1}$ and $\mathbf{H_2}$. The range of values of the homography scale factor λ defined in (15) for the decomposition restricts the valid values for the camera rotation and translation.

The number of coherent camera motions we finally obtain depends on the locations of the cameras and the lines inducing the homographies. In general, there are at least two admissible solutions. Similarly to the 2D case, additional information must be used in order to determine the correct one. When working with omnidirectional images, applying the properties of optical flow to the matched image points can be helpful in order to choose the correct camera motion, and also allows to resolve the π ambiguity inherent to the computation of angles from 1D image data discussed in section III [12].

A particular case occurs when the two camera centers coincide (and the camera motion is therefore a pure rotation). Then, since the translation is null, it can be readily seen in (7) that the calibrated homography matrix is equal to a rotation matrix multiplied by a scale factor. In this situation, any line in the scene induces the same homography (up to scale), which means that we cannot use the method based on the epipoles extracted from the homologies to compute the camera motion. This is also in agreement with the well known fact that the epipoles cannot be computed reliably when the
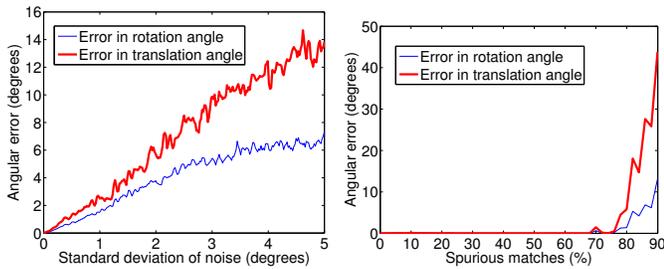
Fig. 1. Simulation results. Left: Errors in the estimated rotation and translation angles with respect to the added Gaussian noise. Right: Errors with respect to the percentage of outliers in the set of putative matches.
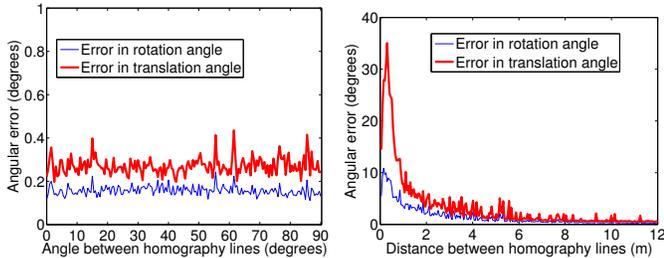
Fig. 2. Simulation results. Left: Errors in the estimated rotation and translation angles with respect to the angle between the lines inducing the two homographies. Right: Errors with respect to the distance between the lines inducing the two homographies.

baseline is short. However, the angle of rotation between the two cameras can be readily obtained in this case. We can detect the pure rotation case by checking how similar the homography matrices are to a scaled rotation matrix. If we divide them by their norm, we obtain rotation matrices. When two homographies induced by two different lines are available, the only degenerate case for our method appears when one of the camera centers belongs to one of these lines. In this situation, the camera motion cannot be estimated with our proposal, but this case is not feasible in practice.

IV. EXPERIMENTS

In this section, we present results from simulations and experiments with real images in order to illustrate the performance of the proposed method.

A. Simulations

For the simulations, a virtual set of points was generated and projected in two 1D cameras. From the set of point matches, we computed two homographies and used them to obtain the camera motion, following the approach described in section III-B. The homographies were calculated linearly using the DLT approach [1], and the RANSAC estimation method was employed in order to compute them robustly. The effect of adding Gaussian noise to the coordinates of the matched points is illustrated in Fig. 1. As expected, the error in the calculations of the angles grows with the level of noise. It can also be observed that the computation of the rotation angle is less sensitive to noise than the computation of the translation angle. In the same figure, we show how the errors behave for different

percentages of spurious matches. The estimated percentage of outliers for RANSAC was in all cases set to 50 %. As can be seen, the method can be successful even with large amounts of outliers. This is an important property in practice, since when computing a homography induced by a given plane, all the matched scene features not belonging to that plane are outliers.

We also study the influence in the method's performance of the relative geometry between the two scene lines inducing the homographies. Low-level Gaussian noise was added in these tests for better visualization of the behavior. It can be seen in Fig. 2 that the error is independent of the angle between the lines, which is depicted from $0°$ (parallel lines) to $90°$ (perpendicular lines). Therefore, the method works equally well when the two lines are parallel. This case occurs, for instance, when a robot travels along a corridor and sees only its two walls. Predictably, the error becomes larger as the distance between the two parallel lines decreases, as shown also in Fig. 2. In the limit, the two homographies are induced by one unique line, a case in which the camera motion cannot be obtained from the decomposition of the homography (Proposition 1).

B. Experiments on real images

We have used a publicly available dataset from the University of Amsterdam [16] to evaluate the performance of our method for planar camera motion estimation on real images. The omnidirectional images included in this set were acquired by a mobile robot traveling around an indoor environment, using a catadioptric vision sensor made up of a perspective camera and a hyperbolic mirror. The images were provided wrapped (i.e. as originally captured by the vision sensor) and had a size of 1024×768 pixels. Since we work with 1D coordinates (i.e. using only the angular coordinate of the image points), we did not require specific calibration information of the camera system other than the position of the projection center in the image. It was assumed that the camera and mirror axis were vertically aligned.

Our method was tested on three different sequences of images extracted from the set. We computed for every image the associated camera motion with respect to the first image of its corresponding sequence. In order to do this, we first extracted SIFT [17] keypoints for the two images and found matches between them. Note that only the angular coordinate of the points is used in our method. Then, the set of two-view correspondences was used to compute 1D homographies between the two images using the DLT and RANSAC algorithms. Usually, more than two homographies were found. From the set of available homographies, we chose the two that were obtained from larger sets of point matches while being at the same time judged as different by a test on the condition number of the homologies.

We selected from the image set three sequences with different characteristics. Sequence 1 was obtained as the robot moved along a corridor. Thus, there were two parallel dominant vertical planes, and the motion was practically rectilinear. The

Fig. 3. Results from the experiments with real images. Left column, rows 1 to 3: estimated rotation angle ϕ. Right column, rows 1 to 3: estimated translation angle ψ. The results are for sequences 1 (first row), 2 (second row) and 3 (third row). Fourth row: initial (reference) image for sequences 1 (left), 2 (center) and 3 (right). Bottom row: final image for sequences 1 (left), 2 (center) and 3 (right).

distance traveled by the robot was of around $4 \ m$. Sequence 2 was collected in a spacious room. In this case, the robot followed a curved path and its final position was $4.5 \ m$ away from the initial one. As for sequence 3, it is set in a smaller space near a staircase, and the motion performed by the robot has the shape of a loop, with its most distant point being around $1.5 \ m$ away from the initial location. The results of the planar camera motion estimations are displayed in Fig. 3 for the three sequences, along with the first and last images of each of them. As can be seen, the angle of rotation is computed with higher accuracy than the angle of translation. The performance of the motion estimation approach is good in the three cases.

V. Conclusion

We have shown that it is not possible to compute the motion between two 1D views on a plane from a single 1D homography induced by a line. Taking this fact into account, we have presented a method for planar camera motion estimation based on the computation of two 1D homographies between the two camera images. Our proposal uses the two homologies associated to the homographies to extract the epipoles and determine the motion parameters. The approach is adequate for use on a robot moving on a planar ground. It is particularly well suited to omnidirectional cameras, due to their wide field of view and accurate angular information. Still, the proposal can also be employed with conventional cameras, as well as with any sensor providing two-dimensional angular measurements. Some typical applications where our approach may be useful are localization, visual servoing, navigation and visual odometry.

References

[1] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.

[2] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry, *An invitation to 3D vision*. Springer Verlag, 2003.

[3] O. Faugeras and F. Lustman, "Motion and structure from motion in a piecewise planar environment," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 2, no. 3, pp. 485–508, 1988.

[4] J. Chen, W. Dixon, M. Dawson, and M. McIntyre, "Homography-based visual servo tracking control of a wheeled mobile robot," *IEEE Transactions on Robotics*, vol. 22, no. 2, pp. 407–416, 2006.

[5] J. Courbon, Y. Mezouar, and P. Martinet, "Indoor navigation of a non-holonomic mobile robot using a visual memory," *Autonomous Robots*, vol. 25, no. 3, pp. 253–266, 2008.

[6] C. Sagüés and J. J. Guerrero, "Visual correction for mobile robot homing," *Robotics and Autonomous Systems*, vol. 50, no. 1, pp. 41–49, 2005.

[7] G. López-Nicolás, J. J. Guerrero, and C. Sagüés, "Multiple homographies with omnidirectional vision for robot homing," *Robotics and Autonomous Systems*, vol. 58, no. 6, pp. 773 – 783, 2010.

[8] K. Åström and M. Oskarsson, "Solutions and ambiguities of the structure and motion problem for 1D retinal vision," *Journal of Mathematical Imaging and Vision*, vol. 12, no. 2, pp. 121–135, 2000.

[9] L. Quan and T. Kanade, "Affine structure from line correspondences with uncalibrated affine cameras," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 8, pp. 834–845, 1997.

[10] F. Dellaert and A. W. Stroupe, "Linear 2D localization and mapping for single and multiple robot scenarios," in *IEEE International Conference on Robotics and Automation*, 2002, pp. 688–694.

[11] J. J. Guerrero, A. C. Murillo, and C. Sagüés, "Localization and matching using the planar trifocal tensor with bearing-only data," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 494–501, 2008.

[12] M. Aranda, G. López-Nicolás, and C. Sagüés, "Omnidirectional visual homing using the 1D trifocal tensor," in *IEEE International Conference on Robotics and Automation*, 2010, pp. 2444–2450.

[13] G. Zhang, H. Zhang, and K.-Y. K. Wong, "1D camera geometry and its application to circular motion estimation," in *Proceedings of the British Machine Vision Conference, vol. I*, 2006, pp. 67–76.

[14] G. López-Nicolás, J. J. Guerrero, O. A. Pellejero, and C. Sagüés, "Computing homographies from three lines or points in an image pair," in *International Conference on Image Analysis and Processing, Lecture Notes in Computer Science 3617*, 2005, pp. 446–453.

[15] X. Zhang, Y. Fang, B. Ma, X. Liu, and M. Zhang, "Fast homography decomposition technique for visual servo of mobile robots," in *Chinese Control Conference*, 2008, pp. 404–409.

[16] Z. Zivkovic, O. Booij, and B. Krose, "From images to rooms," *Robotics and Autonomous Systems*, vol. 55, no. 5, pp. 411–418, 2007.

[17] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.