Citation: 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

Multiple Homographies with Omnidirectional Vision for Robot Homing

G. López-Nicolás, J.J. Guerrero and C. Sagüés

Dept. Informática e Ingeniería de Sistemas - Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza, 50018 Zaragoza, Spain
{gonlopez, jguerrer, csagues}@unizar.es

Abstract

Two relevant issues in vision-based navigation are the field-of-view constraints of conventional cameras and the model and structure dependency of standard approaches. A good solution of these problems is the use of the homography model with omnidirectional vision. However, a plane of the scene will cover only a small part of the omnidirectional image, missing relevant information across the wide range field of view, which is the main advantage of omnidirectional sensors. The interest of this paper is on the new approach for computing multiple homographies from virtual planes using omnidirectional images and its application in an omnidirectional vision-based homing control scheme. The multiple homographies are robustly computed from a set of point matches across two omnidirectional views relying in virtual planes independently of the structure of the scene. The method takes advantage of the planar motion constraint of the platform and computes virtual vertical planes from the scene. The family of homographies is also constrained to be embedded in a three dimensional linear subspace to improve numerical consistency. Simulations and real experiments are provided to evaluate our approach.

Key words: Visual control, homographies, omnidirectional images, mobile robots.

1 Introduction

In order to perform navigation, mobile robots need to interact with the environment, and for this purpose many different types of sensors are available. From all of them, vision systems stand out because they provide very rich information and because of their versatility and low cost. For the last years the use of omnidirectional cameras in robotics is also growing because of their effectiveness due to the panoramic view from a single image. Camera motion between two views can be obtained from rigidity constraints across the views. In general, problems like guided matching, motion computation, or visual control, are directly tackled from

multi-view tensors like the fundamental matrix, which assume a general model with respect to the unknown scene structure. The problem of robust estimation of the epipolar geometry and autocalibration from point correspondences using omnidirectional vision has been studied in [1], [2], [3], [4]. However, there are common situations in which the fundamental matrix fails, reducing the applicability of the model [5] [6] [7] [8]. In particular, the fundamental matrix is ill conditioned with short baseline and undefined with pure rotations. Moreover, its estimation with general algorithms degenerates when all the points lie on a plane. However, this degeneracy with coplanar scene points can be overcome if the camera is calibrated by using the five-point algorithm [9]. Different approaches for vision-based robot navigation are [10] [11] [12]. The Fourier components of the omnidirectional images stored in a visual memory are used in [10]; a trifocal tensor-based approach is presented in [11]; and the homing task is solved in [12] using angular information of a minimal set of three landmark correspondences with isotropic distribution.

A good solution with planar scenes is the homography model. The homography has been used with omnidirectional vision in recent works for visual control [13], visual odometry [14] or navigation using a visual memory of the environment following a sequence of subgoals [15] [16]. However, if no plane is detected, a homography-based approach would fail. This problem can be solved using a virtual plane [17], but in general, estimation based on a virtual plane with wide baseline is not robust to mismatches, noise or occlusions, and a dominant plane is preferred to get better accuracy. Additionally, a plane of the scene will cover only a small part of the omnidirectional image, missing relevant information across the wide range field of view, which is the main advantage using omnidirectional sensors. A proposal of using several homographies defined by real planes is presented in [18] to improve reliability in the visual control task. Another issue is the precision with almost planar scene structures: A precise homography would reject good point matches close to the plane and, on the other hand, joining closer points to support the homography will reduce its precision.

In this paper, we present a new approach for computing multiple homographies from virtual planes and its application in an omnidirectional image-based control scheme for mobile robot homing. The multiple homographies are robustly computed from a set of matches relying in virtual planes independently of the scene structure. Thus, the approach is not compromised by the non planarity of the scene. Additionally, the use of a set of virtual planes can be supported by any inlier across the image, taking advantage of the wide field of view of omnidirectional images. The method considers the planar motion constraint of the platform and, additionally, the virtual planes are constrained to be vertical with respect to the motion plane of the robot. These constraints allow to parameterize the homographies to be computed from only two point correspondences. This minimal sample size improves the performance of the robust estimation with respect to outliers. An additional constraint used for improving the numerical consistency of the family of homographies obtained is that they are embedded in a three dimensional linear subspace. The per-

formance of the method is tested and compared with classical approaches [6] in terms of percentage of spurious, planarity of the scene and image noise showing the advantages of our proposal.

We also propose a new visual control approach taking advantage of the multiple homography estimation method. The system consists of a calibrated omnidirectional camera mounted on top of a differential drive robot. The visual control task uses the classical idea of homing, where the target location is defined by an image previously taken at that location. The current and target images are related by the set of homographies, which are used to design our controller. This visual homing takes advantage of the wide field of view of the omnidirectional camera being independent of the scene structure.

The paper is organized as follows. Section 2 describes the camera model and the homography matrix within our framework. Section 3 presents the computation of a homography considering motion and plane constraints and the multiple homography estimation method is proposed in Section 4. Section 5 presents the control scheme for omnidirectional vision-based homing. Simulations and real experiments given in Section 6 show the performance of the proposed scheme.

2 Imaging Geometry

In this section we define the geometry of the imaging system and describe the omnidirectional camera model using the unified sphere model [19], [20]. The homography matrix is also introduced within our framework.

2.1 Camera model

The considered imaging system consists of an omnidirectional camera mounted on top of a differential drive robot (Fig. 1). The motion of the robot occurs in the $x - y$ plane, and the z-axis of the robot frame is aligned with the camera optical axis, which is also the rotation axis of the robot. The origin of the coordinate frame is set in the optical center of the omnidirectional camera. Its model consists of a projection onto the unit sphere followed by a perspective projection. This unified model is briefly described here and refer to [19], [20] for more details. The camera with the hyperbolic mirror is shown in Fig. 1 with an example of a world point projection (middle) and some system parameters (Table 1). The projection model for central catadioptric cameras is depicted in Fig. 1 (right). A world point $\mathcal{X}$ is projected in a unit sphere centered in $\mathbf{c}$ as $\mathcal{X}_s$. The omnidirectional images can be described by a perspective camera, involving the mirror parameters, related with the sphere by a simple translation of ξ along z-axis.

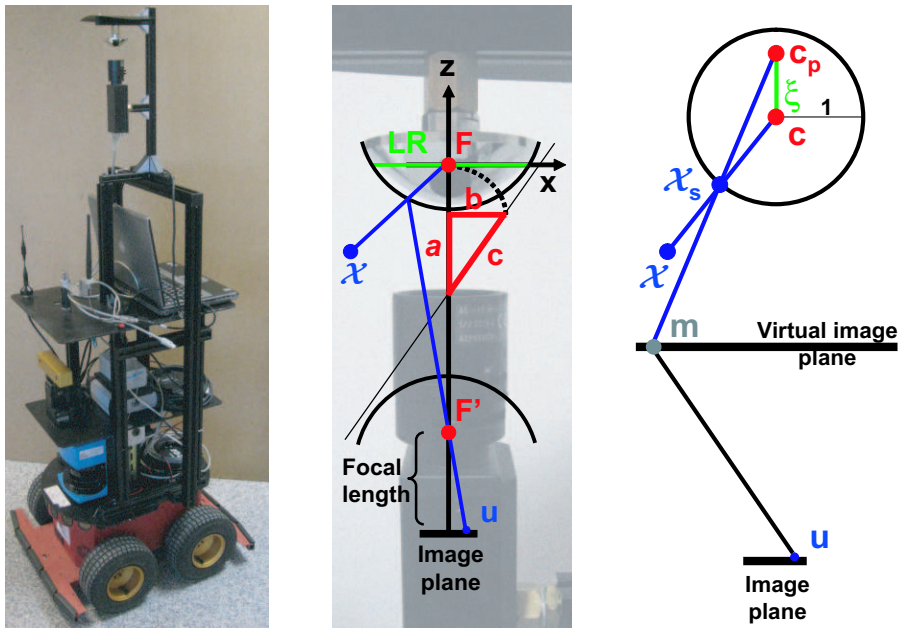

Fig. 1. The robot platform, a Pioneer from ActivMedia with the camera on top (left). Omnidirectional system with the camera (Sony xcd-x7101cr) and the mirror (Neovision H3S) with an example of point projection (middle). Unified model for the camera system (right).

Let us suppose that $\mathbf{u} = (u, v, 1)^T$ is a point of the image acquired by the omnidirectional camera. The point is transformed to a virtual plane using $\mathbf{m} = \mathbf{K}^{-1}\mathbf{u}$, being $\mathbf{K}$ a matrix containing the conventional intrinsic parameters of the camera coupled with mirror parameters (Table 1). Finally, the projection of the point $\mathbf{m} = (x, y, 1)^T$ in the unit sphere can be computed with $\mathcal{X}_s = (\lambda x, \lambda y, \lambda - \xi)^T$. Hereafter, these points are denoted with $\mathbf{p}$.

2.2 The Homography

Two perspective images can be geometrically linked through a plane by a homography $\mathbf{H} \in \mathbb{R}^{3\times 3}$. This projective transformation $\mathbf{H}$ relates points of a planar scene projected in both images. Pairs of corresponding points from the omnidirectional images $(\mathbf{p}, \mathbf{p}')$ are then related up to scale by $\mathbf{p}' = \mathbf{H}\,\mathbf{p}$. The homography between the two images can be computed from a minimal set of four point correspondences solving a linear system [6]. The calibrated homography can be related to camera motion (rotation $\mathbf{R}$ and translation $\mathbf{c}$) and plane location as follows

$$\mathbf{H} = \mathbf{R}\,(\mathbf{I} + \mathbf{c}\,\mathbf{n}^T/d)\ , \tag{1}$$

where $\mathbf{n} = (n_x,\, n_y,\, n_z)^T$ is the unit normal of the plane with respect to the reference frame and d is the distance along $\mathbf{n}$ between the plane and the reference position. We consider the mobile robot constrained to planar motions in $x - y$ (and

Table 1
Calibration parameters of the imaging system using the unified sphere model. The mirror parameters determine the geometry of the system mirror-camera and are depicted in Fig. 1 (middle). The size of the images acquired is 1024×768 pixels.

Mirror parameters
$a = 28.0950$ mm
$b = 23.4125$ mm
$c = \sqrt{a^2 + b^2} = 36.5715$ mm
$d = \overline{FF'} = 2c = 73.1430$ mm
$LR = 2b^2/a = 39.0208$ mm

Model parameters	
$\xi = 0.9685$ mm	$\mathbf{K} = \begin{bmatrix} \gamma & \gamma s & u_0 \\ 0 & \gamma & v_0 \\ 0 & 0 & 1 \end{bmatrix}$
$\gamma = -399.1505$ pixels	
$u_0 = 513.9324$ pixels	
$v_0 = 400.7654$ pixels	$\lambda = \frac{\xi + \sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1}$

rotation ϕ in z-axis). This planar motion constraint yields

$$\mathbf{R} = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{c} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \text{ and } \mathbf{t} = \begin{pmatrix} -x\cos\phi + y\sin\phi \\ -x\sin\phi - y\cos\phi \\ 0 \end{pmatrix}, \quad (2)$$

where $\mathbf{t} = (t_x,\, t_y,\, t_z)^T$ is defined as $\mathbf{t} = -\mathbf{R}\mathbf{c}$.

3 Estimation of a Vertical Homography from Two Points

In this section we present the estimation method of a homography considering not only the motion constraints of the platform but also an additional constraint on the plane that define the homography. Let us consider a set of coplanar points in the world belonging to a vertical plane, and then $\mathbf{n} = (n_x,\, n_y,\, 0)^T$. Therefore, the plane is aligned with the z-axis of the omnidirectional camera. Thus, the homography matrix corresponding to a planar motion scheme and considering a vertical

plane is

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\phi - t_x \frac{n_x}{d} & -\sin\phi - t_x \frac{n_y}{d} & 0 \\ \sin\phi - t_y \frac{n_x}{d} & \cos\phi - t_y \frac{n_y}{d} & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{3}$$

Elements h_{31} and h_{32} are zero because of the planar motion constraint, and h_{13} and h_{23} are zero because only vertical planes are considered. The homography will be estimated up to scale and given that h_{33} is never zero, due to the constraints, we can always normalize the homography by this entry.

The point coordinates are given by $\mathbf{p} = (x,\, y,\, z)$, or $\mathbf{p} = (x/z,\, y/z,\, 1)$ in homogeneous coordinates. A point correspondence $(\mathbf{p}, \mathbf{p}')$ in the omnidirectional images is related up to scale by the homography as $\mathbf{p}' = \mathbf{H}\mathbf{p}$, which can be expressed in terms of the vector cross product as $\mathbf{p}' \times \mathbf{H}\mathbf{p} = \mathbf{0}$ [6]. From this expression two linearly independent equations in the entries of $\mathbf{H}$ are obtained:

$$\begin{bmatrix} xz' & yz' & 0 & 0 & -zx' \\ 0 & 0 & xz' & yz' & -zy' \end{bmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{21} \\ h_{22} \\ h_{33} \end{pmatrix} = 0 . \tag{4}$$

Each point correspondence gives two independent equations. Given that $\mathbf{H}$ is defined by five unknown entries, a set of two point correspondences allows to determine the homography up to a scale factor by solving a linear system. In general, more than two point correspondences are available and the presence of image noise or mismatches is assumed. Then, we use the RANSAC method [21]. In the general case, four point correspondences are needed to determine $\mathbf{H}$, however, four points in the world are not necessarily coplanar. Then, the estimation of $\mathbf{H}$ requires to reject not only outliers but also, good correspondences out of the plane. In our case, a vertical plane is defined by two world points, and also $\mathbf{H}$ is determined by two point correspondences (except the degenerate case of vertically aligned points). Therefore, it is guaranteed that any two inlier point correspondences give a homography. Then, RANSAC method can be used to reject only outliers, and not points out of the plane, because any two correspondences belong to a vertical virtual plane. This feature is used in the next section to estimate multiple homographies from virtual planes.

There is an additional advantage of the two point correspondences minimal set. The RANSAC method proceeds by repeatedly generating hypothesis from a minimal

set of points. Reducing the size of the minimal set of points, the probability of outliers in the samples is reduced and the method performance is greatly improved. In particular, the number N of samples required in RANSAC to ensure that at least one of the samples of size s has no outliers with a probability of p is $N = \log(1 - p)/\log(1 - (1 - \epsilon)^s)$. where $p = 0.99$ is usually chosen and ϵ is the probability that any selected matched point is an outlier. For example, assuming that half of the point matches are outliers $\epsilon = 0.5$, the number of samples required to estimate the general homography ($s = 4$) is $N = 72$ and taking into account the homography constraints considered here ($s = 2$), we have $N = 17$, which implies a considerable improvement.

The estimation method of multiple homographies presented also requires the estimation of the relative omnidirectional camera orientation ϕ, corresponding to each homography. Next, we describe a method for computing ϕ from $\mathbf{H}$. Equation (3) can be rearranged in the following way

$$\begin{bmatrix} h_{11} - \cos\phi & h_{12} + \sin\phi & 0 \\ h_{21} - \sin\phi & h_{22} - \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -t_x \frac{n_x}{d} & -t_x \frac{n_y}{d} & 0 \\ -t_y \frac{n_x}{d} & -t_y \frac{n_y}{d} & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{5}$$

The first two rows of the matrix on the right are linearly dependent, therefore the determinant of this matrix is zero; and also the determinant of the matrix on the left, and then

$$(h_{11} - \cos\phi)(h_{22} - \cos\phi) - (h_{12} + \sin\phi)(h_{21} - \sin\phi) = 0 \ , \tag{6}$$

and reorganizing terms we have

$$(h_{12} - h_{21}) \sin\phi - (h_{11} + h_{22}) \cos\phi = h_{12}h_{21} - h_{11}h_{22} - 1 \ . \tag{7}$$

This previous expression is a trigonometric function which is known to have solution if the next inequality holds

$$(h_{12} - h_{21})^2 + (h_{11} + h_{22})^2 \geq (h_{12}h_{21} - h_{11}h_{22} - 1)^2 \ . \tag{8}$$

Using the values of the homography entries (3) and simplifying we obtain the next expression

$$\left(y\frac{n_x}{d} - x\frac{n_y}{d} \right)^2 \geq 0 \ , \tag{9}$$

which is always true. Therefore (7) has solution and it is given by

$$\phi = \text{atan2}\,(h_{12} - h_{21}, -h_{11} - h_{22}) \pm \text{acos}\frac{h_{12}h_{21} - h_{11}h_{22} - 1}{\sqrt{(h_{12} - h_{21})^2 + (h_{11} + h_{22})^2}} \quad (10)$$

Two solutions are obtained for ϕ. A way for solving this ambiguity problem is to use another plane. Then, we obtain two pair of solutions, and the solution is the compatible pair. The values of ϕ obtained from each homography are used in the multiple homography estimation method proposed in the next section. An advantage of this estimation method is that the correct solution of ϕ is additionally obtained as output data.

4 Estimation of Multiple Homographies

In this section we present the algorithm to compute multiple homographies from a set of point correspondences obtained from the omnidirectional images. A constraint on the rank is also used to enforce numerical consistency across the family of homographies.

In the framework considered here, the minimal set for vertical homography estimation is two point correspondences. Assuming the presence of mismatches, an estimation method robust to outliers is required. We present the Algorithm 1, which is based in the RANSAC robust estimation algorithm [21], [6]. The algorithm uses the distance threshold t, which represents the maximum camera orientation error to consider the homography (and its corresponding point matches) as inlier. Another threshold T is used to define the number of inliers required for minimum consensus. In practice, T is selected considering the assumed percentage of outliers ϵ as $T = (1-\epsilon)M$, with M point correspondences. A degenerate situation occurs when two point correspondences selected are vertically aligned. Then, there will exist infinite vertical virtual planes defined across these two points. This degeneracy is similar to the case of the general homography from four points, with three of them collinear. Points resulting in degenerated situations are rejected as outliers by the estimation method.

Next, we analyze the rank of the multiple homographies obtained. It has been proved that the space of all the homographies between two views is embedded in a four dimensional linear subspace of $\mathcal{P}^8$ [22] [23]. The homography depends on the relative location between the two views; since it also depends on the orientation and location of a plane, we obtain a family of homographies by considering all possible planes. Let us consider a set of homographies normalized by h_{33i}:

$$\mathbf{H}_i = \mathbf{R} + \mathbf{t}\,\mathbf{n}_i^T/d_i \ , \quad i = 1..n \ . \quad (11)$$

Algorithm 1 Robust estimation of multiple vertical homographies with RANSAC.

(I) Given a set of M point matches. Randomly select two correspondences of points $(\mathbf{p}_i, \mathbf{p'}_i)$, $(\mathbf{p}_j, \mathbf{p'}_j)$ obtained from two omnidirectional views and compute the homography by solving (4). Compute the two solutions of the orientation angle (ϕ_{ij}, ϕ'_{ij}) using (10).

(II) For $k = 1$ to M, with $k \neq i, j$, compute the corresponding homography from matches $(\mathbf{p}_i, \mathbf{p'}_i)$, $(\mathbf{p}_k, \mathbf{p'}_k)$. Compute each pair of solutions (ϕ_{ik}, ϕ'_{ik}). The set of homographies with ϕ_{ik} or ϕ'_{ik} within a distance threshold t with respect to ϕ_{ij} or ϕ'_{ij} are considered as inliers.

(III) A threshold T is defined with the number of inliers required for minimum consensus. If the number of inliers is greater than T, the algorithm finishes and the set of homographies computed from the inliers is returned with the angle ϕ of largest consensus.

(IV) If the number of inliers is lesser than T, repeat from (I).

(V) After N trials, the set of homographies and the angle ϕ computed from the largest consensus set is returned.

In our case, we have added two additional constraints: the robot moves on the floor plane and only virtual vertical planes are considered. In order to analyze the rank of the system, we write the set of homographies $\mathbf{H}_i$ in vector form as $9 \times n$ matrix

$$\begin{bmatrix} & \mathbf{h}_{1i} & \\ \cdots & \mathbf{h}_{2i} & \cdots \\ & \mathbf{h}_{3i} & \end{bmatrix}_{9\times n} = \begin{bmatrix} \mathbf{r}_1 & & \mathbf{r}_1 \\ \mathbf{r}_2 & \cdots & \mathbf{r}_2 \\ \mathbf{r}_3 & & \mathbf{r}_3 \end{bmatrix}_{9\times n} + \begin{bmatrix} \mathbf{t} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{t} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{t} \end{bmatrix}_{9\times 3} \begin{bmatrix} & n_{xi}/d_i & \\ \cdots & n_{yi}/d_i & \cdots \\ & 0 & \end{bmatrix}_{3\times n} \quad ,(12)$$

where $\mathbf{H} = [\mathbf{h}_{1i},\ \mathbf{h}_{2i},\ \mathbf{h}_{3i}]$ and $\mathbf{R} = [\mathbf{r}_1,\ \mathbf{r}_2,\ \mathbf{r}_3]$. The null rows corresponding to entries h_{13}, h_{23}, h_{31} and h_{32} can be removed, and rearranging the previous expression we obtain

$$\begin{bmatrix} & h_{11i} & \\ & h_{12i} & \\ \cdots & h_{21i} & \cdots \\ & h_{22i} & \\ & 1 & \end{bmatrix}_{5\times n} = \begin{bmatrix} c\phi & t_x & 0 \\ -s\phi & t_y & 0 \\ s\phi & 0 & t_x \\ c\phi & 0 & t_y \\ 1 & 0 & 0 \end{bmatrix}_{5\times 3} \begin{bmatrix} & 1 & \\ \cdots & n_{xi}/d_i & \cdots \\ & n_{yi}/d_i & \end{bmatrix}_{3\times n} . \quad (13)$$

Thus, the rank of the system is 3 and the family of homographies is embedded in a 3 dimensional linear subspace of $\mathcal{P}^4$. The maximum rank is 3 and, if the scene is purely planar, the rank is 1. This condition on the rank is an additional constraint and it allows enforcing numerical consistency across the set of homographies. We use this property in the estimation method of multiple homographies. The matrix

$5 \times n$ of the family of homographies in (13) estimated from real data will not have rank 3 in general. This constraint can be easily enforced by the solution obtained using the Singular Value Decomposition (SVD). Let the matrix $\mathbf{H}_n \in \mathbb{R}^{n\times 5}$ be the set of homographies and $\mathbf{H}_n = \mathbf{U}\mathbf{S}\mathbf{V}^T$ its singular value decomposition, where $\mathbf{S}$ is a diagonal matrix $\mathbf{S} = diag(s_1, s_2, s_3, s_4, s_5)$ with singular values in decreasing order. Then, we replace the matrix $\mathbf{H}_n$ by

$$\mathbf{H}_n = \mathbf{U}\, diag(s_1, s_2, s_3, 0, 0)\, \mathbf{V}^T \ , \tag{14}$$

which is the closest singular matrix to the initially estimated $\mathbf{H}_n$ under Frobenius norm.

We have obtained a family of homographies of rank 3 defined from virtual vertical planes. All of them have a common relative translation and rotation ($\mathbf{R}$, $\mathbf{t}$) compatible with the omnidirectional views and different plane parameters ($\mathbf{n}_i$, d_i). From the application point of view, it can be useful to group all the homographies together. We now describe the linear combination of these homographies to obtain a sole homography from a virtual plane to be used in the omnidirectional visual control scheme proposed in the next section. Consider a set of homographies as defined in (11) where for all i they are normalized with h_{33i} (i.e. $h_{33i} = 1$). The linear combination of the family of homographies gives

$$\mathbf{H}_V = \sum_i (\lambda_i \mathbf{H}_i) = \sum_i (\lambda_i)\, \mathbf{R} + \mathbf{t} \sum_i (\lambda_i \mathbf{n}_i^T / d_i) \ , \tag{15}$$

where the parameters λ_i are non-negative real numbers. The entry used for normalization of the homography matrix obtained is given by $h_{33} = \sum_i (\lambda_i)$. The resulting homography can also be normalized by h_{33} yielding

$$\mathbf{H}_V = \frac{\sum_i (\lambda_i \mathbf{H}_i)}{\sum_i (\lambda_i)} = \mathbf{R} + \mathbf{t}\, \frac{\sum_i (\lambda_i \mathbf{n}_i^T / d_i)}{\sum_i (\lambda_i)} \ , \tag{16}$$

with the corresponding entry $h_{33} = 1$. This previous homography corresponds to a homography generated from a virtual plane ($\mathbf{n}_V, d_V$) where the plane parameters can be defined as

$$\mathbf{n}_V = \frac{\sum_i (\lambda_i \mathbf{n}_i^T / d_i)}{||\sum_i (\lambda_i \mathbf{n}_i^T / d_i)||} \quad \text{and} \quad d_V = \frac{\sum_i (\lambda_i)}{||\sum_i (\lambda_i \mathbf{n}_i^T / d_i)||} \ . \tag{17}$$

Thus, any linear combination of the family of homographies results in a new homography. Outlier correspondences are rejected in the robust estimation of the homographies following Algorithm 1. However, the accuracy of the individual homographies is always affected in practice by uncertainties like image noise. Given that

a linear combination of the homographies is also a homography, a simple arithmetic mean of them leads to a new homography with an average accuracy:

$$\mathbf{H}_V = \sum_i (\lambda_i \mathbf{H}_i), \quad i = 1..n \ , \tag{18}$$

where $\lambda_i = 1/n$ are the weights, being n the number of homographies. In this equation, noisy homographies score the same as accurate homographies. This issue reminds the case of the RANSAC algorithm, in which outliers score nothing (they are rejected) but all the inliers (either good or poor estimations) weight the same. The use of a cost function where inliers score according on how well they fit the data has been previously presented [24]. Different weight functions were also used in [18] based on re-projection error, homography determinant or area of the plane to compute a reliable homography from four real planes. Hence the idea of defining the weights λ_i in such a way that homographies with less residual error receive higher scores while poor estimations have less influence.

The absolute residual error of the orientation angles are defined for each homography $\mathbf{H}_i$ as $\varepsilon_i = |\phi_{ij} - \phi_{ik}|$ with $k \neq i, j$ from Algorithm 1. And the maximum residual error is given by $\varepsilon_{max} = max(\varepsilon_i) = max(|\phi_{ij} - \phi_{ik}|)$. Then, we propose a weighted mean redefining λ_i where the minimum error gives the maximum weight, namely $\lambda_i = 1$, and maximum error is assigned with $\lambda_i = 0$ as follows

$$\lambda_i = 1 - \frac{\varepsilon_i}{\varepsilon_{max}} \ , \quad i = 1..n \ . \tag{19}$$

Therefore, the family of homographies is combined linearly (18) with parameters λ_i defined as (19). The resultant homography can be normalized from $h_{33} = \sum_i \lambda_i$ to $h_{33} = 1$. Actually, this normalization is equivalent to define normalized weights in (19). In summary, with (18) and the assignment of weights (19) normalized we obtain as output a homography defined from a virtual plane that merges together the family of homographies as

$$\mathbf{H}_V = \frac{\varepsilon_{max} \sum_i \mathbf{H}_i - \sum_i (\varepsilon_i \mathbf{H}_i)}{n\varepsilon_{max} - \sum_i \varepsilon_i}, \quad i = 1..n \ . \tag{20}$$

5 Visual Control Scheme

In this section the multiple homography estimation method previously described is used in a new visual control scheme. The goal is to design a control law for autonomous robot navigation based on omnidirectional visual information using the classical idea of homing, where the desired location is defined by an image taken previously at that location. Then, a current image together with a reference image

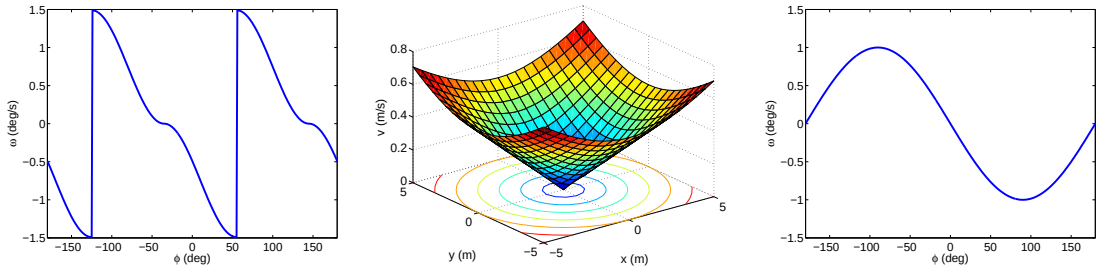

Fig. 2. Example of the velocities given by the control laws in each step. Left: Rotational velocity ω as a function of ϕ at step 1. Middle: Velocity v as a function of x and y with $\phi = \psi$ at step 2. Right: Rotational velocity ω as a function of ϕ at step 3.

is used to compute the set of multiple homographies. The family of homographies can be merged using (20), obtaining a robust homography $\mathbf{H}_V$ independent of the scene structure. The individual entries of this homography are used in the design of the control law proposed next.

We consider the standard motion model for unicycle robots [25], in which the motion occurs in the x-y plane and the z-axis of the robot frame coincides with the camera optical axis. Then, the robot location is given by $\mathbf{x} = (x, y, \phi)^T$. The robot is steered with the linear velocity v and angular velocity ω, where v is in the direction of the robot y-axis and ω is around the robot z-axis.

5.1 Control Laws

The motion strategy consists of three sequential steps. During step 1, the robot with the omnidirectional camera is rotated until its orientation points to the target location. During step 2 the robot drives toward the target following a straight-line path maintaining the orientation. Finally, in step 3 the robot performs a pure rotation until the target orientation is reached. This kind of strategy is especially feasible when omnidirectional sensors are used. This strategy was also used in [26] with a conventional camera and a dominant plane for the homography estimation to avoid discontinuities in the control output. Here, the multiple homography method, not only is independent of the scene structure but also allows to exploit the omnidirectional image information. Next, the control law for each step to achieve the trajectory of this motion strategy is defined in terms of the entries of $\mathbf{H}_V$ (20).

The desired orientation in the first step is $\phi = \psi$ where $\psi = \arctan(x/y)$. When this orientation is reached we also have that $t_x = 0$ (2). From (5) we can compute $t_x^2 = (h_{11} - \cos\phi)^2 + (h_{12} + \sin\phi)^2$ and the rotational velocity can be defined using

this value:

$$\text{Step 1} \begin{cases} v = 0 \\ \omega = k_1 \sqrt{(h_{11} - \cos\phi)^2 + (h_{12} + \sin\phi)^2} \end{cases} \tag{21}$$

where k_1 is defined as $k_1 = k_w sign((h_{12} + \sin\phi)/(h_{22} - \cos\phi))$ being k_w a scalar control gain. This function has two minimum values and at the end of this step the robot is pointing to the target location with $\phi = \psi$ or $\phi = \psi + \pi$. An example of the behaviour of this controller is given in Fig. 2 (left). In this example $\psi = -35°$ or $145°$ and it can be seen that these values coincide with the minimal angular velocities of the controller.

The second step requires that the robot moves toward the target (forward or backward) in a straight-line path. The velocity in this step can be related with t_y, and when $t_y = 0$ the robot is in the target location up to a rotation. From (5) we can compute $t_y^2 = (h_{21} - \sin\phi)^2 + (h_{22} - \cos\phi)^2$ and controller in step 2 is defined as

$$\text{Step 2} \begin{cases} v = k_2 \sqrt{(h_{21} - \sin\phi)^2 + (h_{22} - \cos\phi)^2} \\ \omega = k_1 \sqrt{(h_{11} - \cos\phi)^2 + (h_{12} + \sin\phi)^2} \end{cases} \tag{22}$$

where $k_2 = k_v sign((\cos\phi - h_{22})n_y)$ being k_v a scalar control gain and the sign of n_y is required. The resultant velocity v as a function of x and y with $\phi = \psi$ is depicted in Fig. 2 (middle) and it can be seen that the gradient of velocities converge to zero at the target location. The velocity v when $\phi = \psi + \pi$ is analogue but with opposite sign.

In step 3 the robot orientation is corrected. There we have $h_{21} = \sin\phi$ and the controller is defined using this homography matrix entry as

$$\text{Step 3} \begin{cases} v = 0 \\ \omega = -k_w \, h_{21} \end{cases} \tag{23}$$

being k_w a scalar control gain. The output velocity ω of this controller as a function of ϕ is depicted in Fig. 2 (right). It can be seen that an stable equilibrium point appears in $\phi = 0$. The design of the previous control laws takes into account that the homographies are based on virtual planes. Then, it is assumed that during the navigation the number of correspondences obtained from the omnidirectional images will vary, new correspondences appear and other will be lost. Therefore, the virtual plane is different each time. The control design has taken into account this issue avoiding plane dependency in the control variables. Thus, the entries used in

the control output computation have been combined properly to remove the plane parameters dependency.

5.2 Stability Analysis

In the following, the stability of the control scheme is analyzed by means of the *Lyapunov's Direct Method* [27]. We define a candidate Lyapunov function for the control law of each step as $V_1 = (t_x^2/2)$ for step 1, $V_2 = (t_y^2/2)$ for step 2 and $V_2 = (\phi^2/2)$ for step 3. Next, we show that these candidate functions are Lyapunov functions when using the proposed control laws. These functions are positive definite and radially unbounded, and next we show that their derivatives are negative definite:

$$\begin{aligned}
\dot{V_1} &= t_x\,\dot{t_x} = -\omega\,t_x\,t_y \\
&= -k_w sign\left(\frac{h_{12}+\sin\phi}{h_{22}-\cos\phi}\right)\sqrt{(h_{11}-\cos\phi)^2+(h_{12}+\sin\phi)^2}\,t_x\,t_y \\
&= -k_w\ sign\left(\frac{t_x}{t_y}\right)\ |t_x|\ t_x\ t_y < 0 \qquad (24)\\
\dot{V_2} &= t_y\,\dot{t_y} = -v\,t_y \\
&= -k_v sign((\cos\phi - h_{22})n_y)\sqrt{(h_{21}-\sin\phi)^2+(h_{22}-\cos\phi)^2}\,t_y \\
&= -k_v\ sign(t_y)\ |t_y|\ t_y < 0 \qquad (25)\\
\dot{V_3} &= \phi\,\dot{\phi} = \phi\,\omega \\
&= -k_w\,\phi\,h_{21} = -k_w\,\phi\,\sin\phi < 0 \qquad (26)
\end{aligned}$$

In the previous development we have used expressions of (2), (5), the kinematic robot model equations $\dot{x} = v\sin\phi$, $\dot{y} = v\cos\phi$ and the control law expressions (21), (22) and (23). Then, we conclude that $(\dot{V_1}, \dot{V_2}, \dot{V_3}) < 0$ and the system is stable with each control law in the Lyapunov sense. The stability for slow switching without a common Lyapunov function is studied in [28]. If the individual controllers are asymptotically stable, the switched linear system is globally stable if a dwell time between switches is large enough. In our control scheme, the switching of the control laws is performed sequentially and enough dwell time between switches can be assumed.

6 Experiments

In this section we present experimental validation of the approach presented. First, we test the multiple homography estimation method with simulated and real data.

And second we evaluate the omnidirectional visual control scheme with simulations and real experiments.

6.1 Evaluation of the Estimation Method of Multiple Homographies

Different simulations are presented to show the validity of the estimation method of multiple homographies. We test and compare the performance of the method presented (2 points) with the classical computation of the homography (4 points) [6] taking into account the percentage of outliers, planarity of the scene or image noise. The method performance is also compared with two general scene structure solutions using the eight-point algorithm [6] and five-point algorithm [9], since the camera is calibrated. The virtual scene consists of a random set of 3D points which are projected to the image through a hyperbolic mirror and camera system using the unified sphere model described in Section 2 [19], [20]. The method presented is then applied using the image point correspondences in calibrated coordinates.

The error of the estimated orientation ϕ using our approach is tested with respect to the percentage of outliers and compared with the orientation obtained from the decomposition of the homography computed with the classical approach [6] (Fig. 3). The estimated percentage of outliers parameter used in RANSAC in both methods is always set to 50% ($\epsilon = 0.5$). The 3D points of the virtual scene used in Fig. 3 (left) are randomly distributed in a plane. Both methods reject correctly outliers under 60% of spurious matches. When the percentage of outliers increases over ϵ, the probability of RANSAC success decreases and outliers can be considered as inliers. So, the behaviour of both methods is similar with the limitations inherent to RANSAC. The same experiment is repeated in Fig. 3 (middle) with a non-planar scene (random general 3D points). The performance of the multiple homographies is similar as having a planar scene, but the general homography has to deal with outliers plus non-coplanar points. Therefore, the general homography has less possibility of large consensus, and this situation is especially relevant when using omnidirectional images. The performance of the approach as a function of the scene structure is studied in Fig. 3 (right). The error of the estimated orientation ϕ is tested with respect to the planarity of the scene, i.e. the horizontal axis represents the planarity from 0% (3D points are randomly distributed in a plane) to 100% (3D points are randomly distributed in a cube). The point correspondences are perfect data and it can be seen that the multiple homography method performs correctly independently of the scene. On the other hand, the general homography estimation is worse with a less planar scene.

The next experiment evaluates the performance of the approach in the presence of image noise (Fig. 4). Gaussian image noise is added to the point correspondences of the omnidirectional images with standard deviation σ from 0 to 3 pixels. In the case of a planar scene, both our approach and the general homography methods

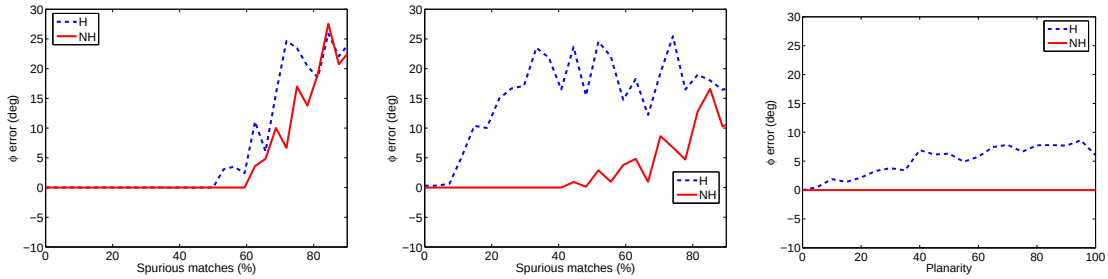

Fig. 3. Error of the estimated orientation ϕ with respect to the percentage of outliers when the scene is a plane (left) or non-planar (middle). ϕ error with respect to the planarity of the scene (right). The multiple homography method (NH-solid line) is compared with the classical homography estimation (H-dashed line). Each value is the mean of ten repetitions.

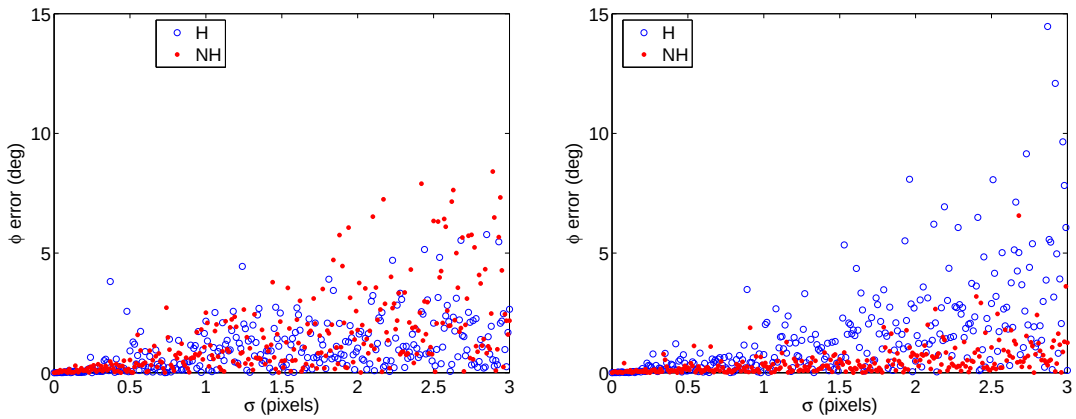

Fig. 4. Error of the estimated orientation ϕ with respect to image noise when the scene is planar (left) or non-planar (right). The multiple homography method (blue '•') is compared with the classical homography estimation (red '∘').

are quite similar although the general homography has a little less error. This is because its computation involves minimizing a cost function, which results in good performance given perfect matching and Gaussian image noise. On the other hand, our approach is based on multiple homography consensus with the rank constraint. Both methods show greater differences when the scene is non-planar. In this case the multiple homography method presents much better performance.

Experimental evaluation has been carried out to compare the performance of the proposed approach (NH) and the two-view epipolar geometry estimated by means of the eight-point ($F8$) and five-point ($E5$) algorithms [6] [9]. In the following experiments, we have computed and compared the error of the orientation (ϕ) using these three methods (Fig. 5). The first test compares the performance of these approaches with respect to the planarity of the scene. The results show that both NH and $E5$ perform correctly in planar and non planar scenes whereas $F8$ fails with coplanar scene points, Fig. 5 (left). Notice that $E5$ plot is hardly seen because it is under NH plot and they are nearly the same. The next experiment evaluates the influence of the baseline in the performance of the methods. The error of the

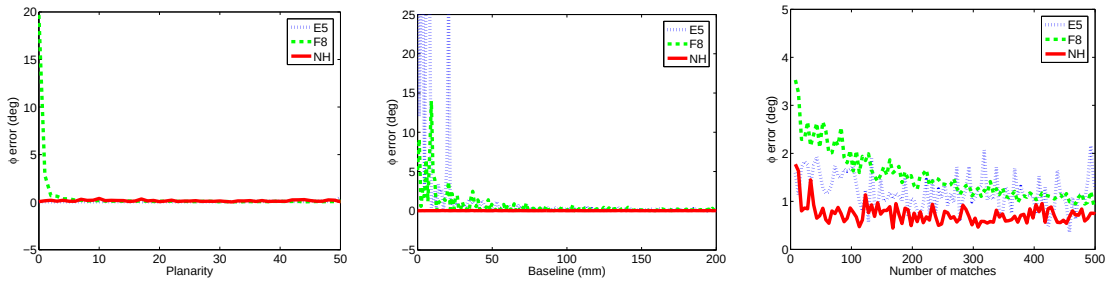

Fig. 5. Error of the estimated orientation ϕ with respect to the planarity of the scene (left), baseline (middle) and number of matches (right). The multiple homography method (NH-solid line) is compared with the fundamental matrix ($F8$-dashed line) and the essential matrix ($E5$-dotted line).

estimated orientation with respect to the distance between the images is given in Fig. 5 (middle), showing the problem of short baseline with the epipolar geometry ($E5$, $F8$). This experiment specially shows the difference on performance with short baseline. We have also tested the behavior of the methods according to the number of matches Fig. 5 (right) in the presence of image noise. The results show that NH and $E5$ perform better than $F8$ with few matches while $F8$ performance is considerably improved with the number of matches.

The multiple homography estimation method is also tested with real images. The images are acquired with the camera system described in Section 2 and image size 1024×768 pixels. An example of the images used is given in Fig. 6 (left) together with the putative SIFT (Scale Invariant Feature Transform) matches found [29]. The SIFT point correspondences present great repeatability, distinctiveness, and robustness. However, this descriptor is not designed for omnidirectional cameras, which presents a great challenge because of image deformation and varying resolution. Thus, many mismatches can be expected in the putative correspondences. The set of images taken follows a pure rotation from $0°$ to $90°$ by steps of $15°$. Another set follows the same sequence of rotation plus a translation. The orientation estimated from the multiple homographies computed with respect to the real orientation is depicted in Fig. 6 (right). It can be seen that very good results have been obtained. The example in Fig. 7 shows a pair of images, their SIFT putative matches and the point correspondences left after the standard homography and multiple homography estimation. The standard homography cover a wall of the room while the multiple homography approach correspondences are more distributed across the images.

The next experiment has been carried out using an omnidirectional image database from the University of Amsterdam and available via web [30]. The omnidirectional images were taken by a camera with a hyperbolic mirror on a mobile robot driving through an indoor environment. The calibration of the omnidirectional camera and data from laser range scanner, sonar range scanners and odometry sensors are also available. From the more than one thousand omnidirectional images available in the database we have selected a sequence of 160, in particular images

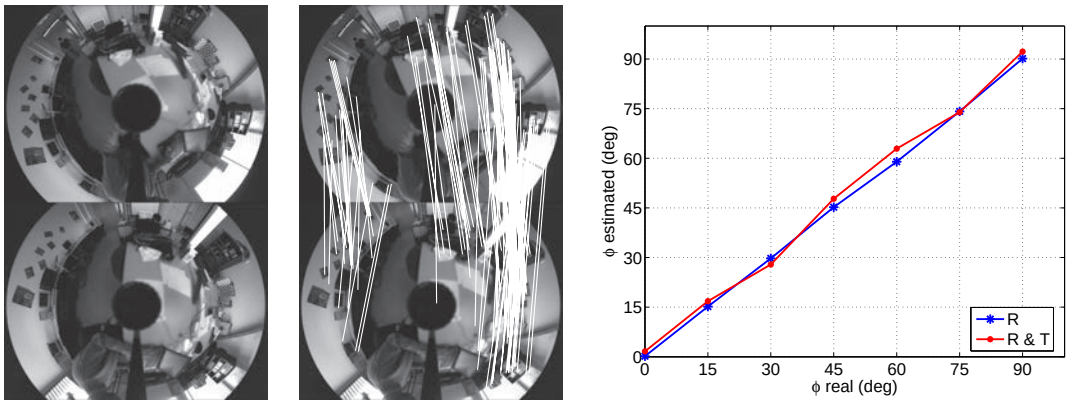

Fig. 6. Images presenting a pure rotation of $30°$ and putative matches (left). Estimated orientation ϕ with respect to the real orientation (right). The sequence of real data varies from $0°$ to $90°$ with pure rotation (blue '$*$') and the same with translation of 1 m (red '$\bullet$').

Fig. 7. Example of two images (1st column). Putative SIFT matches (2nd column). Point correspondences left after homography estimation with standard method (3rd column) and using the multiple homography approach (4th column).

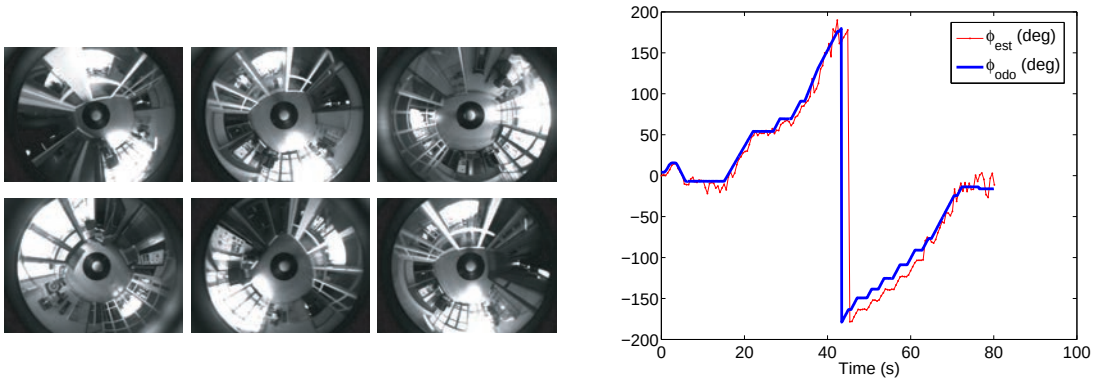

Fig. 8. Examples from the sequence of 160 images taken from the database (left). Estimated orientation ϕ_{est} and orientation given by the robot odometry ϕ_{odo}(right).

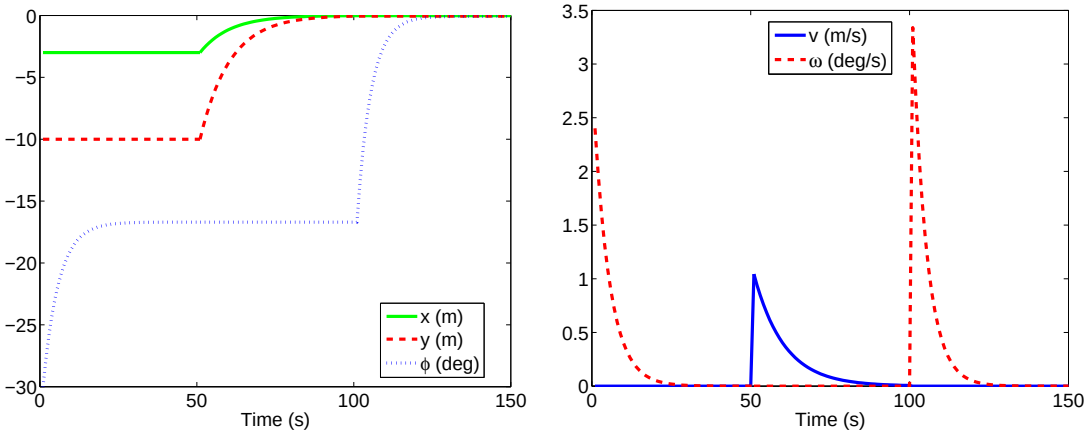

Fig. 9. Simulation using the control scheme proposed. Evolution of the robot pose (x, y, ϕ) (left). Velocities (v, ω) given by the controllers (right).

from 100 to 260. Some examples of the images used are shown in Fig. 8 (left). The orientation estimated with our approach is depicted with the orientation given by the robot odometry in Fig. 8 (right) showing good behaviour. SIFT points give very few matches or none with these images mainly because of light conditions. Then, the putative matches have been obtained using the Lucas and Kanade tracking algorithm [31], [32]. The experimental evaluation shows the effectiveness of the approach presented supporting the benefits of considering the planar motion constraint. Besides, our proposal of multiple homographies makes also possible to compute the scene structure or the robot localization, although this is not addressed in this work because we focus on the robot control. So, without loss of generality, the multiple homographies estimation approach is also suitable for different applications like visual odometry, Structure and Motion or SLAM. In the next section we evaluate the visual control scheme, which is the application addressed here, using the combination of the family of homographies as previously presented.

6.2 Evaluation of the Omnidirectional Visual Control Scheme

Simulated and real experiments of the omnidirectional visual control scheme are presented in this section. Simulation showing the behavior of the control scheme following the motion strategy is given in Fig. 9. The initial location is (x, y, ϕ) = (-3 m, -10 m, $-30°$) and the goal position is $(0, 0, 0)$. The evolution of the robot coordinates x, y, and orientation ϕ along the time and the velocities given by the control laws are depicted in Fig. 9.

The real experiments have been carried out with the robot platform shown in Fig. 1 (left). The robot is equipped with a camera plus mirror system mounted on top Fig. 1 (middle). The camera is connected through Firewire to a laptop onboard the robot and plugged into a battery. The laptop is an Intel® Core^{TM} 2 Duo CPU at 2.50 GHz with operating system Debian Linux. The goal position of the robot is

Fig. 10. Sequence of some images taken by the robot camera during an experiment. The first is the target image, the second is the initial and the last is the image at the end of the motion.

defined by a target image taken previously at the desired position. In each controller loop, a current image is taken and used to estimate the multiples homographies between the current and target images. The images are acquired using Player [1] , and the image data is processed using the OpenCV library [2] . The interest points of the current and target images used to estimate the homographies are initially extracted and matched. During the motion, the points are tracked using Lucas and Kanade algorithm [31], [32]. The homography estimated (20) is the input of the control law and the velocities are the output. These velocities computed by the control scheme are sent to the robot through a serial cable using again Player. When the control finishes, the robot is in the goal position and the homography is the identity matrix. The control loop of the implementation runs at 150 ms.

Experimental results obtained are presented in Fig. 10 and Fig. 11. A sequence of some of the images taken during an experiment are shown in Fig. 10. The robot covers around 1 m in x and 2.5 m in y. The evolution of the robot pose given by the odometry with respect to the initial pose is shown in Fig. 11 (left), the estimated orientation is also depicted. The relative initial orientation with respect to the target is approximately zero degrees. The velocities computed by the control law and sent to the robot are depicted in Fig. 11 (right).

In the next experiment the robot goes round a square covering a distance around 10 meters. The desired trajectory is specified with four target images taken at the corners of the square path (Fig. 12). When each target is reached, the control changes to the next target image of the sequence. The resultant path followed by the robot and the evolution of the homography entries h_{11}, h_{12}, h_{21} and h_{33} are shown (Fig. 12). It can be seen that the robot navigates successfully following the square path with good performance and without cumulative errors. Experiments show that the multiple homography based control performs correctly and that the motion strategy results in a robust navigation.

[1] The Player Project creates free software tools for robot and sensor applications. `http://playerstage.sourceforge.net/`

[2] The Open Computer Vision Library is a collection of algorithms for computer vision. `http://www.intel.com/technology/computing/opencv/`

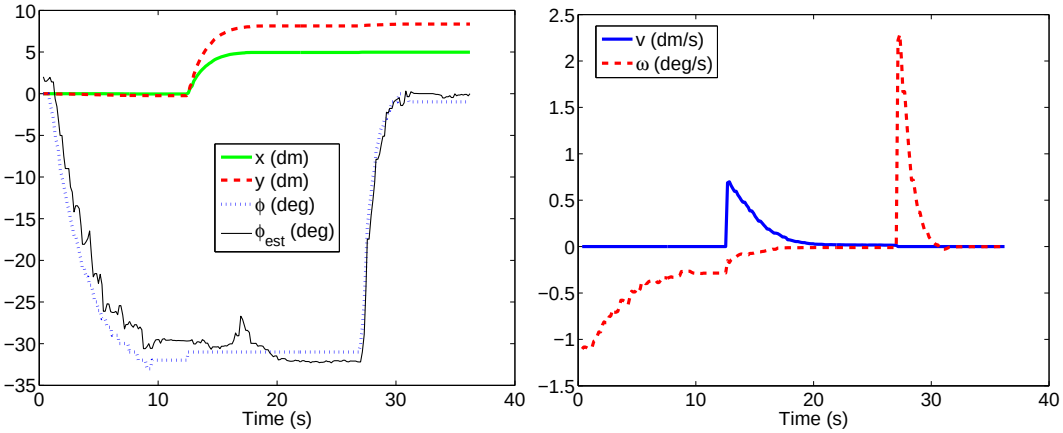

Fig. 11. Experimental results with the robot platform. Evolution of the robot pose (x, y, ϕ) referred to the initial position given by robot odometry (left) and estimated orientation (ϕ_{est}). Velocities (v, ω) given by the controllers (right).

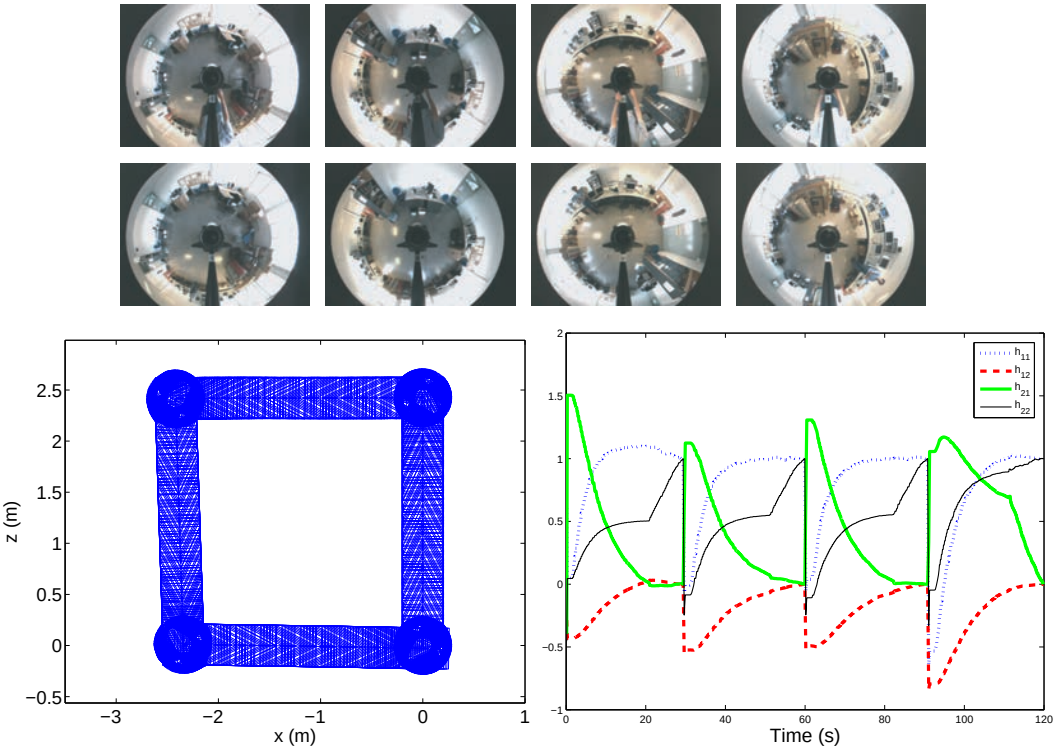

Fig. 12. Experimental results commanding the robot to go round a square. The first row is the sequence of target images taken at the corners of a square path. The second row are the final images acquired when each respective target is reached. Resultant path of the robot (left) and evolution of the homography entries used in the control (right).

7 Conclusion

We have presented a new approach for estimation of multiple homographies defined from virtual vertical planes with omnidirectional image pairs. The minimal

set of point correspondences required for homography estimation is reduced to two considering constraints on the camera motion and on the virtual planes. As output, the method gives a family of homographies of rank 3 and the relative orientation of the cameras. The classical methods based on the epipolar geometry fails with planar scenes, but this problem is overcome using the five-point algorithm for calibrated cameras or using the homography model. However only the homography solves the problem of short baseline. The experiments carried out show robust behavior with respect to outliers and good performance independently of the scene structure. This approach is especially useful in omnidirectional vision, where multiples homographies allow the use of information from the complete scene given by the omnidirectional images. We have also presented an omnidirectional visual control scheme defined in terms of the multiple homographies independent of the plane parameters. The motion strategy consists of three control laws sequentially switched which are intended for omnidirectional vision-based navigation. The stability analysis of the control scheme is also presented. The feasibility of our approach has been demonstrated through simulations and experiments with a mobile robot.

Acknowledgment

The authors would like to thank Jesús Bermúdez for his contribution in the experiments. This work was supported by projects DPI2009-08126 and DPI2009-14664.

References

[1] T. Svoboda and T. Pajdla, "Epipolar geometry for central catadioptric cameras," *International Journal on Computer Vision*, vol. 49, no. 1, pp. 23–37, 2002.

[2] B. Micusik and T. Pajdla, "Estimation of omnidirectional camera model from epipolar geometry," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 485–490, June 2003.

[3] C. Geyer and K. Daniilidis, "Mirrors in motion: epipolar geometry and motion estimation," *IEEE International Conference on Computer Vision*, pp. 766–773 vol.2, Oct. 2003.

[4] B. Micusik and T. Pajdla, "Structure from motion with wide circular field of view cameras," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, pp. 1135–1149, July 2006.

[5] P. H. S. Torr, A. Fitzgibbon, and A. Zisserman, "The problem of degeneracy in structure and motion recovery from uncalibrated image sequences," *International Journal of Computer Vision*, vol. 32, no. 1, pp. 27–45, 1999.

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

[7] G. López-Nicolás, C. Sagüés, and J. J. Guerrero, "Automatic matching and motion estimation from two views of a multiplane scene," in *Pattern Recognition and Image Analysis, LNCS 3522*, June 2005, pp. 69–76.

[8] G. López-Nicolás, C. Sagüés, J. J. Guerrero, D. Kragic, and P. Jensfelt, "Switching visual control based on epipoles for mobile robots," *Robotics and Autonomous Systems*, vol. 56, no. 7, pp. 592–603, 2008.

[9] D. Nister, "An efficient solution to the five-point relative pose problem," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, June 2003, pp. 195–202.

[10] E. Menegatti, T. Maeda, and H. Ishiguro, "Image-based memory for robot navigation using properties of omnidirectional images," *Robotics and Autonomous Systems*, vol. 47, no. 4, pp. 251 – 267, 2004.

[11] G. López-Nicolás, J. J. Guerrero, and C. Sagüés, "Visual control through the trifocal tensor for nonholonomic robots," *Robotics and Autonomous Systems (2009), doi:10.1016/j.robot.2009.09.005*.

[12] A. A. Argyros, K. E. Bekris, S. C. Orphanoudakis, and L. E. Kavraki, "Robot homing by exploiting panoramic vision," *Autonomous Robots*, vol. 19, no. 1, pp. 7–25, 2005.

[13] S. Benhimane and E. Malis, "A new approach to vision-based robot control with omni-directional cameras," in *IEEE International Conference on Robotics and Automation*, 2006, pp. 526–531.

[14] D. Scaramuzza and R. Siegwart, "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1015–1026, Oct. 2008.

[15] H. Hadj-Abdelkader, Y. Mezouar, and P. Martinet, "Path planning for image based control with omnidirectional cameras," *IEEE Conference on Decision and Control*, pp. 1764–1769, Dec. 2006.

[16] 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.

[17] E. Malis and F. Chaumette, "2 1/2 D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement," *International Journal of Computer Vision*, vol. 37, no. 1, pp. 79–97, 2000.

[18] D. S. Kumar and C. V. Jawahar, "Robust homography-based control for camera positioning in piecewise planar environments," in *ICVGIP*, ser. LNCS, P. Kalra and S. Peleg, Eds., 2006, vol. 4338, pp. 906–918.

[19] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical applications," in *Proceedings of the 6th European Conference on Computer Vision, ECCV-Part II*. Springer-Verlag, 2000, pp. 445–461.

[20] J. Barreto and H. Araujo, “Issues on the geometry of central catadioptric image formation,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR.*, vol. 2, pp. 422–427, 2001.

[21] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[22] A. Shashua and S. Avidan, “The rank 4 constraint in multiple (≥ 3) view geometry,” *Lecture Notes In Computer Science*, vol. 1065, pp. 196–206, 1996.

[23] L. Zelnik-Manor and M. Irani, “Multiview constraints on homographies,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 214–223, 2002.

[24] P. H. S. Torr, “Bayesian model estimation and selection for epipolar geometry and generic manifold fitting,” *International Journal of Computer Vision*, vol. 50, no. 1, pp. 35–61, 2002.

[25] J.-P. Laumond, *Robot Motion Planning and Control*. Springer Verlag, 1998.

[26] G. López-Nicolás, C. Sagüés, and J. J. Guerrero, “Shortest path homography-based visual control for differential drive robots,” G. Obinata and A. Dutta, Eds., Vision Systems. I-Tech Education and Publishing, Austria, 2007, pp. 583–596.

[27] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2001.

[28] D. Liberzon, *Switching in Systems and Control*. Birkhauser, 2003.

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

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

[31] J. Shi and C. Tomasi, “Good features to track,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR.*, pp. 593–600, Jun 1994.

[32] B. D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in *Proc. of 7th International Joint Conference on Artificial Intelligence (IJCAI)*, 1981, pp. 674–679.