DIIS - I3A
C/ María de Luna num. 1
E-50018 Zaragoza
Spain

Nonholonomic Epipolar Visual Servoing

G. López-Nicolás, C. Sagüés and J.J. Guerrero, D. Kragic and P. Jensfelt

Nonholonomic Epipolar Visual Servoing

G. López-Nicolás, C. Sagüés and J.J. Guerrero
DIIS - I3A, Universidad de Zaragoza,
C/ María de Luna 1, E-50018 Zaragoza, Spain
{gonlopez, csagues, jguerrer}@unizar.es

D. Kragic and P. Jensfelt
CAS - CVAP, Royal Institute of Technology,
SE100 44 Stockholm, Sweden
{danik, patric}@nada.kth.se

Abstract— A significant amount of work has been reported in the area of visual servoing during the last decade. However, most of the contributions are applied in cases of holonomic robots. More recently, the use of visual feedback for control of nonholonomic vehicles has been reported. Some of the examples are docking and parallel parking maneuvers of cars or vision-based stabilization of a mobile manipulator to a desired pose with respect to a target of interest. Still, many of the approaches are mostly interested in the control part of visual servoing loop considering very simple vision algorithms based on artificial markers. In this paper, we present an approach for nonholonomic visual servoing based on epipolar geometry. The method facilitates a classical teach-by-showing approach where a reference image is used to define the desired pose (position and orientation) of the robot. The major contribution of the paper is the design of the control law that considers nonholonomic constraints of the robot as well as the robust feature detection and matching process based on scale and rotation invariant image features. An extensive experimental evaluation has been performed in a realistic indoor setting and the results are summarized in the paper.

I. INTRODUCTION

The field of service robotics is a fast growing one. Apart from building reliable platforms, mobile manipulation represents one of the biggest challenges since it encompasses research problems such as e.g. navigation, obstacle avoidance, visual servoing, object grasping and manipulation. Although there have been a significant amount of work reported in each of the above areas, there are still no systems that can robustly and completely safely move in the environment and manipulate objects at the same time.

Some of the key problems to perform mobile manipulation tasks are the level of cooperation between different subsystems and the amount of sensors required. It is generally accepted that machine vision is one of the most important sensory modalities for navigation, object manipulation and grasping purposes. Hence, one of the major building blocks of a mobile manipulation systems is visual servoing [1]. During the last few years, an enormous amount of research has been reported in visual servoing, [2]–[6]. However, many of the above references concentrate mainly on problems and differences between 2 1/2 D, image and position based visual servoing assuming that there are no constraints on the robot motion itself.

In terms of mobile manipulation it is very important to take into account the motion of the base that carries a manipulator or a pan-tilt unit. Such platforms are commonly nonholonomic and most of the above approaches are unsuitable in this case. Some researches have concentrated on control problem related to mobile platforms, [7]–[10]. These will be presented in more detail in the next section where we compare them to our approach.

Our method facilitates a classical teach-by-showing visual servo approach based on epipolar geometry estimation. The proposed approach is suited for a nonholonomic mobile platform and does not require complete camera calibration or any specific knowledge of the scene geometry. An extensive experimental evaluation performed in a realistic indoor setting shows the effectiveness of our method.

This paper is organized as follows. In Section II we review the related work and outline the contributions of our approach. In Section III we present the estimation of epipolar geometry based on scale invariant image features. Visual and motion models are described in Section IV followed by the design of the control law in Section V. Experimental evaluation is given in Section VI. Conclusion and avenues for future research are given in Section VII.

II. RELATED WORK

The basic idea of teach-by-showing visual servoing is to control a robot to a specific pose in the environment by regulating to zero an error term which is estimated by matching image data between the current and the reference position. This approach has for sometime now been the most common way of controlling robots since it does not have to deal with hard vision problems such as scene segmentation or place recognition.

An early example of this approach is presented in [11]. Here, image-based navigation is considered where the motion control is performed by recovering the relative pose of the robot with respect to the desired one. This relative pose is estimated through the essential matrix, assuming that the internal parameters of the camera are precisely known. In comparison, our method does not require perfect knowledge of the camera calibration parameters. In addition, our approach does not require the computation of any of the 3D distances since the control is based directly on the trajectory of the epipoles. In [11] the distance to the target is determined as a number of steps by using two consecutive images, being

each step of the same length as the one performed between the two images used. In our method the motion at every step is independent of the previous steps. This allows to perform auxiliary tasks or to avoid obstacles while navigating without affecting the convergence to the target.

Another approach presented by [9] is based on the epipolar geometry and exploits the auto-epipolar property. It does not require any information about the internal camera parameters. However the algorithm developed is designed only for holonomic robots. The approach presented in [10] extends the auto-epipolar visual servoing method of [9] to cope with nonholonomic constraints. The motion is performed in three sequential steps to reach the target position. The second step, which deals with the problem of nonholonomic constraints, is based on the epipolar geometry and consists in an input-output feedback linearizing control law. However, the motion performed in the second step drives the robot from and not to the target (just moving away) before turning back to it in the final step, which is not the intuitive behaviour we might expect. This is one of the basis of our work, which is focused in obtaining a direct motion towards the target.

Classically the problem of homing is solved by using epipolar geometry, but it is ill conditioned with planar scenes, which occur frequently in man-made environments. In addition, images provided during visual servoing can result in small baselines, where the estimation of the fundamental matrix also gives bad results. It has been shown that working with homographies eliminates this problem [8], [12].

III. FROM IMAGES TO EPIPOLES

In this section, we shortly describe the computation of epipolar geometry as well as visual features used in the process.

A. Feature Extraction and Matching

In a recent study, Mikolajczyk and Schmid [13] analyzed a large number of interest point descriptors and their behaviors under changes, such as scale and illumination change. The descriptor that turned out to be most robust in this study was the Scale Invariant Feature Transform - SIFT descriptor originally proposed in [14]. It was also concluded that the point detector used was less significant. In the current implementation, we use SIFT features as originally developed by [14] where, primarily for reasons of low computational cost, the descriptor uses feature points determined by the peaks of a series of difference of Gaussians on varying scales. In [15], the so called Harris-Laplace features that respond to regions of high curvature, instead of blob-like image structures as in original SIFT has been presented. This is also the method used in our current work. Unlike ordinary Harris' features, peaks are found spatially as well as in scale, thus making scale invariance possible. This leads to features accurately localized spatially, which is essential if features are used for pose estimation, instead of just matching.

As it can be seen in Fig. 1, the descriptor corresponds to highly distinctive image locations and is robustly invariant to
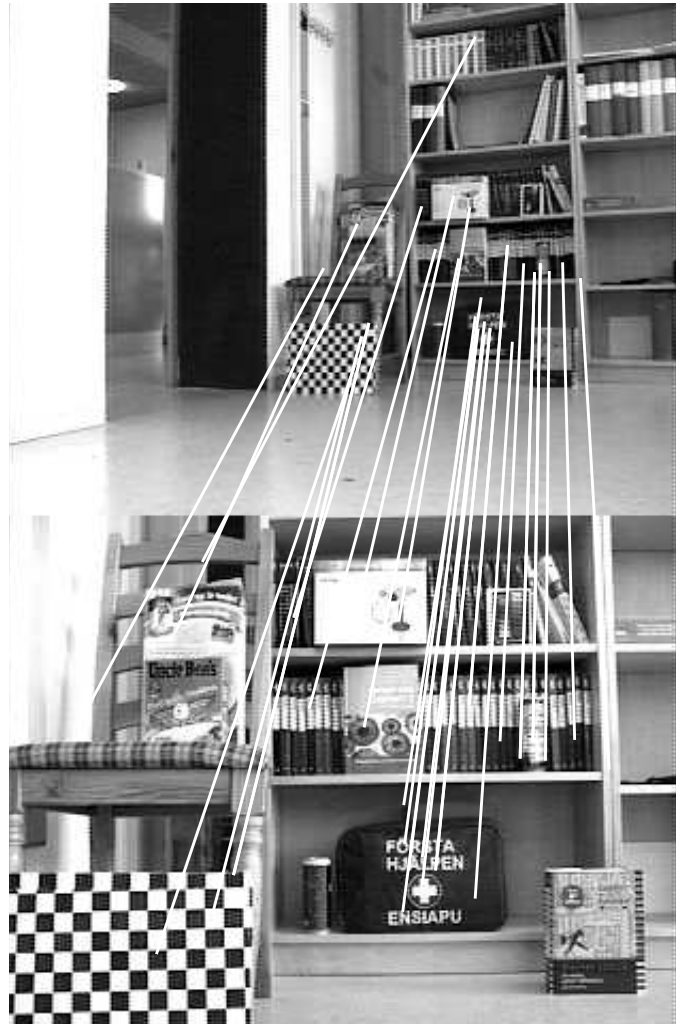

Fig. 1. Total 31 matches have been found. A robust estimation process is used to remove the mismatches as explained in the text.

image plane transformations such as translation, rotation and scaling. Matching between images is performed using a simple squared distance measure between descriptors. To make the matching process more robust, we require that for each pair of matched points, the match is the best one regarding both first-second and second-first image pair directions.

B. Epipolar Geometry Estimation

Once matches are available, the position of epipoles can be estimated. We first start by estimating the fundamental matrix using the robust approach proposed by Torr [16] called Maximum a posteriori SAmple Consensus (MAPSAC) which is based on a well-known 7-point algorithm, [17]. Similar to RANSAC, the method proceeds by repeatedly calculating putative solutions from a minimal seven point correspondences and minimizes an error term for a predefined set of point combinations. Fig. 2 shows the estimated epipoles and corresponding epipolar lines for one of the current-target image pairs.
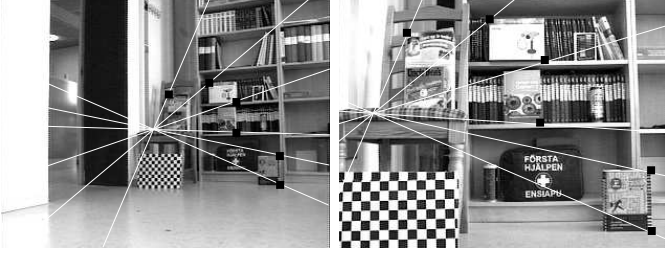
Fig. 2. Estimated epipoles and corresponding epipolar lines: current image (left) and target image (right).

Fig. 3. Geometric relations for e_{cx} (left) and e_{tx} (right). C_c and C_t are the current and target camera positions respectively.

IV. VISUAL AND MOTION MODELS

The general pinhole camera model considers a calibration matrix defined as

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix},$$

where α_x and α_y are the focal length of the camera in pixel dimensions in the x and y directions respectively; s is the skew parameter and (x_0, y_0) are the coordinates of the principal point. We have that $\alpha_x = f m_x$ and $\alpha_y = f m_y$, where f is the focal length and m_x, m_y are the pixels per distance unit.

For the approach proposed in this paper, neither a complete nor an accurate knowledge of the camera calibration parameters is necessary. The important step is to define the desired trajectories of the epipoles used in the input of the control law. In practice we can suppose that the principal point is in the center of the image ($x_0 = 0$, $y_0 = 0$), that there is no skew ($s = 0$) and that the camera has squared pixels.

Let us now suppose that the state of the robot is given by its position and orientation coordinates $\mathbf{x} = (x, z, \theta)^T$. From the perspective projection of Fig. 3, the x-coordinates of the epipole in the current image (e_{cx}) can be expressed as a function of the state of the robot as

$$e_{cx} = \alpha_x \frac{x \cos\theta - z \sin\theta}{z \cos\theta + x \sin\theta} \quad . \tag{1}$$

In a similar way (Fig. 3) we have the epipole x-coordinates in the target image (e_{tx}) as

$$e_{tx} = \alpha_x \frac{x}{z} \quad . \tag{2}$$

Let us also suppose that the nonholonomic differential kinematics to be expressed in a general way as

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= h(\mathbf{x}) \end{aligned}$$

and let us consider the problem of visual servoing as a tracking problem in a nonlinear system. The particular nonholonomic differential kinematics of the robot expressed in state space representation as a function of the translation and rotation speeds (v, ω) is as follows
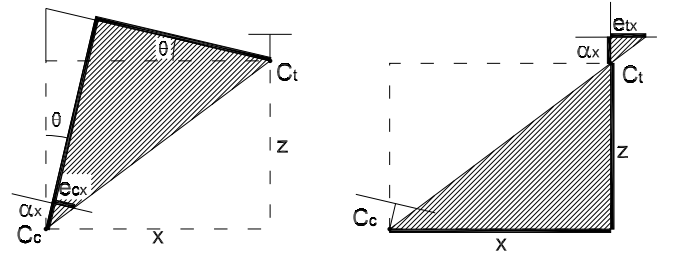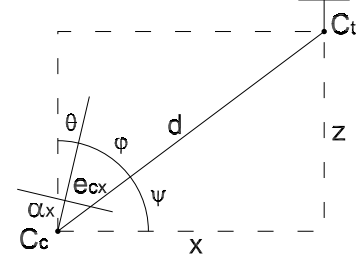

Fig. 4. Geometric relations for θ and ψ.

$$\begin{bmatrix} \dot{x} \\ \dot{z} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} v \sin\theta \\ v \cos\theta \\ \omega \end{bmatrix} \tag{3}$$

$$\mathbf{y} = [e_{cx} \ e_{tx}]^T$$

From Fig. 4 we can deduce the next geometric relations

$$\begin{aligned} z &= d \sin\psi \\ x &= d \cos\psi \\ d^2 &= x^2 + z^2 \end{aligned} \tag{4}$$

V. CONTROL LAW

The objective is to carry out the control of the system by using only the epipoles coordinates as input. Then the visual servoing is transformed in a tracking problem where the desired trajectories of the epipoles are defined. With the appropriately defined trajectories of the epipoles we propose a control law which allows a direct navigation towards the target pose. This is also one of the strongest contributions of our work compared to previously reported approaches.

As the system is nonlinear, an input-output linearization is made. This linearization is carried out by differentiating the system outputs until the inputs appear explicitly, and then solving for the control inputs. So, for the epipole in the current position according to (1), we have

$$\dot{e}_{cx} = \frac{\partial e_{cx}}{\partial t} = \frac{\partial}{\partial t} \left(\alpha_x \frac{x \cos\theta - z \sin\theta}{z \cos\theta + x \sin\theta} \right) \quad .$$

Using (3) and (4) after the differentiation, it follows

$$\dot{e}_{cx} = -v \frac{\alpha_x \cos(\theta + \psi)}{d \sin^2(\theta + \psi)} - \omega \frac{\alpha_x}{\sin^2(\theta + \psi)} \quad . \tag{5}$$

Similarly for the epipole in the target image (2) we have

$$\dot{e}_{tx} = \frac{\partial e_{tx}}{\partial t} = \frac{\partial}{\partial t}\left(\alpha_x \frac{x}{z}\right) = \alpha_x \frac{\dot{x}z - x\dot{z}}{z^2} \quad .$$

Again, using (3) and (4) it follows that

$$\dot{e}_{tx} = -v\frac{\alpha_x \cos(\theta + \psi)}{d\sin^2\psi} \quad . \tag{6}$$

With the expressions just deduced, we can write a linear relation between the linearized input and the output as

$$\left[\begin{array}{c} v_c \\ v_t \end{array}\right] = E\left[\begin{array}{c} v \\ \omega \end{array}\right] \quad ,$$

where v_c and v_t, which are defined later, are functions depending on $\dot{e}_{cx}$ and $\dot{e}_{tx}$ respectively. From (5) and (6) the decoupling matrix obtained is

$$E = \left[\begin{array}{cc} -\frac{\alpha_x \cos(\theta+\psi)}{d\sin^2(\theta+\psi)} & -\frac{\alpha_x}{\sin^2(\theta+\psi)} \\ -\frac{\alpha_x \cos(\theta+\psi)}{d\sin^2\psi} & 0 \end{array}\right] \quad ,$$

with $\psi = \arctan(\alpha_x/e_{tx})$ and $\theta = \arctan(\alpha_x/e_{cx}) - \psi$, (Fig. 4).

Therefore, the input of the system will be obtained as

$$\left[\begin{array}{c} v \\ \omega \end{array}\right] = E^{-1}\left[\begin{array}{c} v_c \\ v_t \end{array}\right] \quad ,$$

where (v_c, v_t) are the new inputs to be determined. Assuming the control objective to be the output to track the set point output $\dot{e}_{cx}^{des}$, $\dot{e}_{tx}^{des}$ we can make

$$\left[\begin{array}{c} v_c \\ v_t \end{array}\right] = \left[\begin{array}{c} \dot{e}_{cx}^{des} - k_c(e_{cx} - e_{cx}^{des}) \\ \dot{e}_{tx}^{des} - k_t(e_{tx} - e_{tx}^{des}) \end{array}\right] \quad ,$$

with k_c and k_t the controller gains. In a similar way as presented in [10] this results in a exponentially stable error dynamics.

This control law needs the matrix E to be invertible. From $det(E) = -\alpha_x^2\cos(\theta+\psi)/(d\sin^2\psi\sin^2(\theta+\psi))$ we have that the matrix E is singular if $(\theta+\psi) = 90°$ (which is equivalent to $e_{cx} = 0$). Our objective is to perform a motion with the robot aligned with the baseline, only when it is not aligned the control law needs to correct the orientation; otherwise the robot moves forward with a constant velocity, avoiding the singularity of the decoupling matrix. The distance between the current and target camera (d) is unknown, but it can be replaced with a constant parameter without affecting the convergence of the system.

The behavior of the robot fully depends on the desired trajectories of the epipoles. Now, different possibilities can be followed in order to select the trajectories depending on the desired robot motion. The trajectory of each epipole is implemented as a set of trajectories which make the epipoles to evolve accordingly (Fig. 5)

$$e_{cx}^{des}(t) = \left\{\begin{array}{ll} (e_{cx}(0) - x_0)(\frac{t^2}{T_c^2} - \frac{2t}{T_c} + 1) + x_0 & if \quad 0 \le t < T_c \\ x_0 & if \quad t \ge T_c \end{array}\right.$$

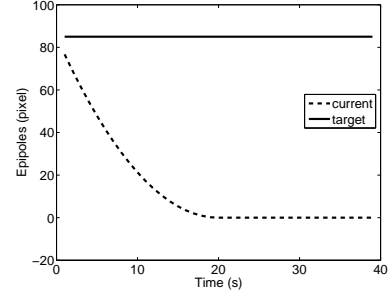$$e_{tx}^{des}(t) = e_{tx}(0) \qquad \forall t \quad . \tag{7}$$

Fig. 5. Desired trajectories of the epipoles.

The idea is first to perform a rotation until $e_{cx} = x_0$ while $e_{tx} = constant$. After this step, when the time reaches T_c, the robot is aligned with the baseline and the current camera is pointing to the target. Next, the robot moves in a straight trajectory until the target is reached. More detail can be found in our previous work [18].

When the robot is close to the target, the epipolar geometry is not well defined. In our approach, the robot reaches the target without the desired orientation and we cannot correct for this rotation error by using a control law based on epipolar geometry. Therefore, an additional procedure is needed to correct the final error in the target position and in this case, a homography based control can be used [8]. This is a viable solution since close to the target, one can assume that the environment is partially planar. The next section provides an example of how this problem was solved in the current work.

VI. EXPERIMENTAL EVALUATION

Simulated and real experiments have been carried out to show the validity and the performance of the approach. Simulation experiments demonstrate the response of the control law under different conditions and the real experiments demonstrate the good performance of the method in a real, domestic environment.

Observing the navigation sequence, the resulting motion can be divided in three phases. The first one consists in the rotation of the robot at the initial position until it points to the target. Next, the robot moves straight to the target position. Finally, it rotates to reach the desired orientation. The first and second phase are carried out together by the control law. However, the last phase cannot be accomplished by the control law since in this position the epipolar geometry is not well defined. Therefore, a correlation scheme based on the image features is used in this phase instead. The transition between the control law and the correlation scheme occurs when the values of the epipoles change suddenly over a threshold, pointing out that the epipolar geometry is no more well defined. The value of the threshold is determined experimentally taking into account that the epipoles will be computed with noise.

A. Simulation Results

The simulated data consists in a synthetic scene of random points which are generated and projected on the image planes
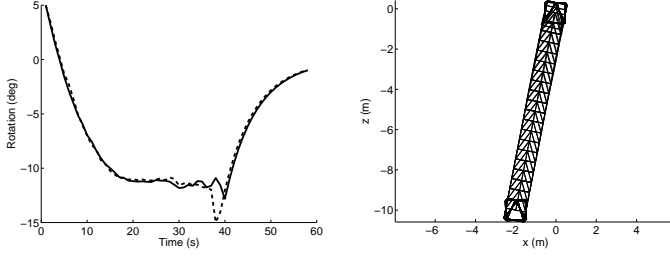
Fig. 6. Robot rotation and path with image white noise of $\sigma = 0.1$ (solid line) and 2 (dashed line) pixels.
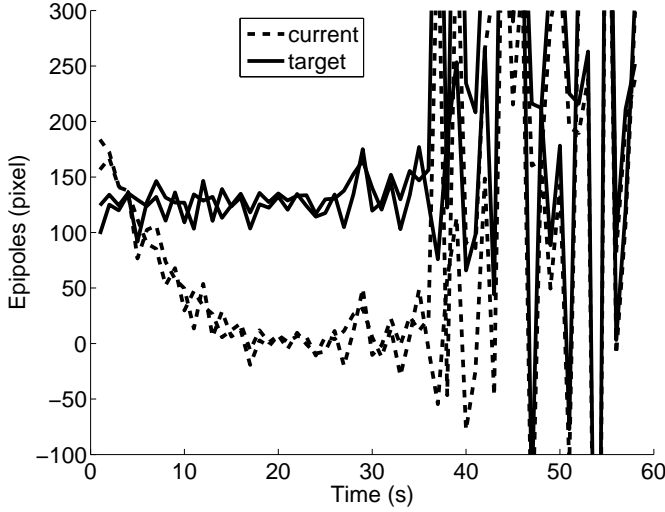

Fig. 8. Robot rotation and path with $f = 6$, 16 and 26 mm.

Fig. 7. Epipoles evolution with image white noise of $\sigma = 0.1$ and 2 pixels.
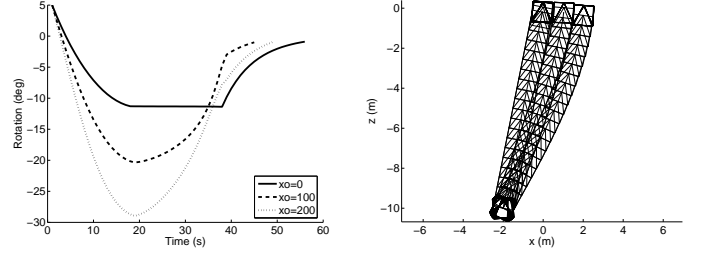
Fig. 9. Robot rotation and path with $x_0 = 0$, 100 and 200 pixels.

in each iteration of the navigation. The size of the images produced by the virtual camera is 320×240 pixels. Then, the epipolar geometry is computed from the point matches. For the following simulations the start position considered is $(x, z, \theta) = (-2, -10, 5°)$ and the target position is $(0, 0, 0°)$.

In the first experiment, white noise has been added to the image points and the experiment has been repeated with different values of standard deviation (σ). The rotation, the path and the epipoles are represented in Fig. 6 and Fig. 7 with the different results superposed. The results show that the navigation with the control law is robust to image noise. As said, the epipolar geometry is ill-conditioned when the baseline is too short; it can be seen in Fig. 7, where the fundamental matrix does not give good values for the epipoles in the last phase of the motion (from 40 seconds on).

In the second experiment the real focal length is modified with different values while the control law holds $f = 6$ mm. The objective is to show the robustness to calibration errors. In fact, the results show that is not necessary to know the focal length. The path obtained is the same in all the cases (Fig. 8) and the only effect is a proportional variation in the output of the control law. This can be solved with the appropriate setting of the controller gains.

The third experiment is carried out with the assumption that the x coordinate of the principal point is $x_0 = 0$ while its
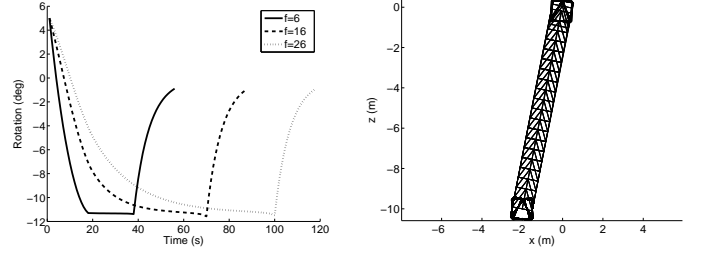
real value is changed. Fig. 9 reveals that there is a lateral error in the position without any rotation error depending on the inaccuracy of x_0. This is because the expression for the desired trajectories of the epipoles in the input of the control law according to (7) depends on x_0. Hence, if the final desired position of the current epipole is the principal point, its inaccuracy is propagated to the control, resulting in the lateral error. On the other hand, the target epipole is set to remain constant independently of the principal point, helping to decrease the final lateral error.

B. Experimental Platform

The experimental platform is a PowerBot from ActivMedia (Fig. 10). It has a nonholonomic differential drive base with two rear caster wheels. The robot is equipped with a SICK LMS200 laser scanner placed low in the front, a sonar ring with 28 Polaroid sensors, a Canon VC-C4 pan-tilt-zoom CCD camera mounted on top of the laser scanner and a firewire camera on the last joint of a 6 DOF arm.

In the experiments the platform is controlled by sending translation and rotational speeds (v, ω). The vision control loop has not yet been optimized and it currently runs at 0.5Hz. Therefore, the velocities have to be quite low. The maximum translation velocity is set to 0.04m/s and the maximum rotation velocity is set to 0.02rad/s.

C. Real World Experiments

The real world experiments are performed with the above platform in a domestic setting. It has to be mentioned that no particular knowledge of the environment, other than the image from the desired pose, is provided.

The camera calibration parameters are completely unknown. The values of the calibration parameters are set freely without performing calibration and these are $f = 6$ mm for the focal
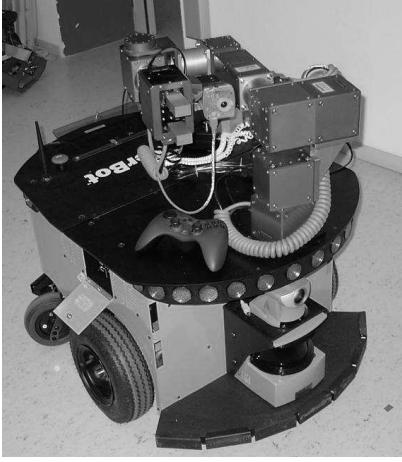
Fig. 10. The experimental platform: ActiveMedia's PowerBot.

length and $x_0 = 0$ for the principal point. The size of the images is 320×240 pixels.

Fig. 11 shows the evolution of an experiment with the images taken during the navigation, where (a) is the image taken in the initial position, (b-i) some navigation images and (j) the target image previously taken. The final image is shown in Fig. 11(i) comparing it with the target image in Fig. 11(j).

Results from one of the experimental runs are shown in Fig. 12. The lateral distance to the target (a) and depth distance (b) shows the robot position evolution and (c) shows the rotation. As expected the robot moves directly toward the target along the shortest path (d). The position data have been obtained from the robot odometry. The final position error obtained is 8 cm in x-coordinate, 11 cm in z-coordinate and negligible rotation error. This deviation is due to the fact that the camera is not mounted on the rotation axis of the robot. Thus, when the robot rotates in the final position around the rotation axis of the robot instead of rotating around the center of the camera the translational error is added. The evolution of the epipoles is shown in (e). The number of SIFT matches found in each step is shown in (f). As expected the number of matches increase with the similarity of the current and target image as the robot advance to the target. Note here that there is a checkerboard pattern put in the scene, in order to show that the feature matching process is robust even with respect to repetitive patterns.

VII. Conclusion

The benefits of visual servoing in mobile manipulation settings have been widely acknowledged. It is our belief that for this purpose both holonomic and nonholonomic constraints have to be considered. In addition, methods that do not require any special modelling of the environment and those that do not consider the use of fiducial markers are of significant importance.

In this paper, we have presented a visual servoing method for nonholonomic mobile platforms based on epipolar geometry. The control law implemented is obtained from the

(a) Initial image   (b)

(c)   (d)

(e)   (f)

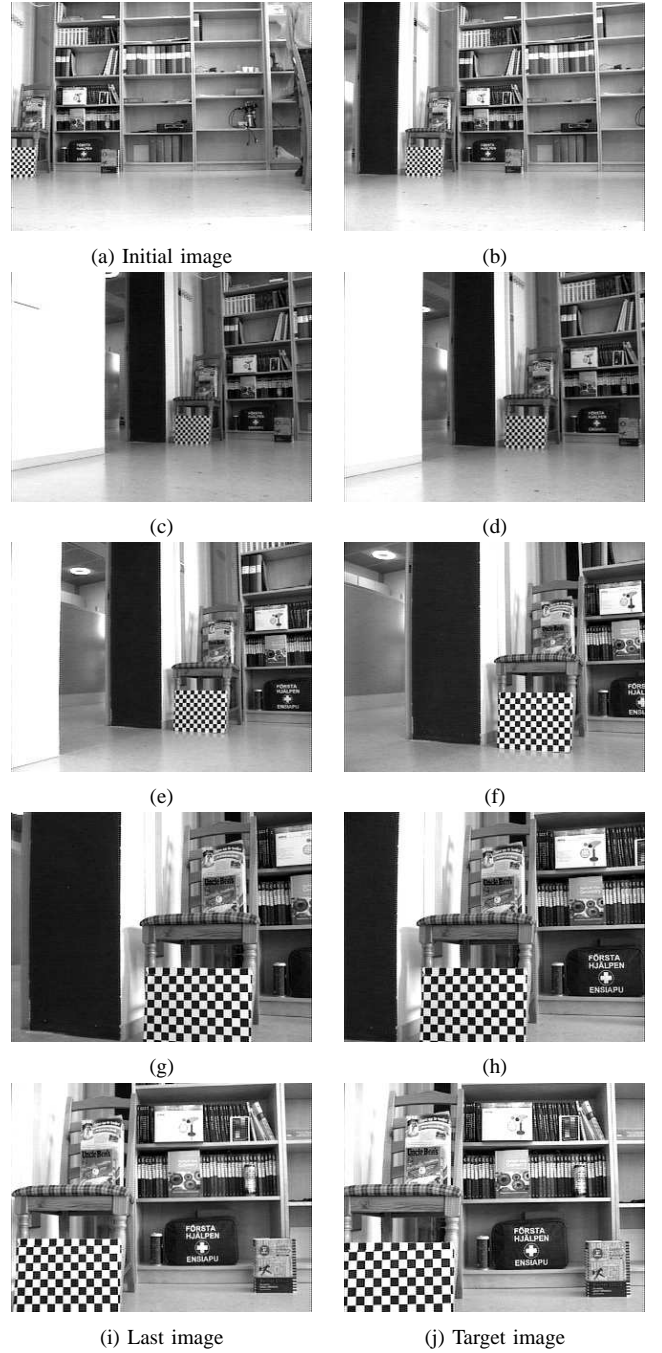(g)   (h)

(i) Last image   (j) Target image

Fig. 11. A sequence of images taken during an experiment.

input-output linearization of the system. The basic idea is to servo the robot to the target position by simultaneously tracking the desired trajectories of the epipoles in the current and target image. An automatic robust feature detection and matching process is performed for the estimation of the fundamental matrix to obtain the epipoles. This process is based on matching scale and rotation invariant feature points in order to allow successful control under significant scale changes.

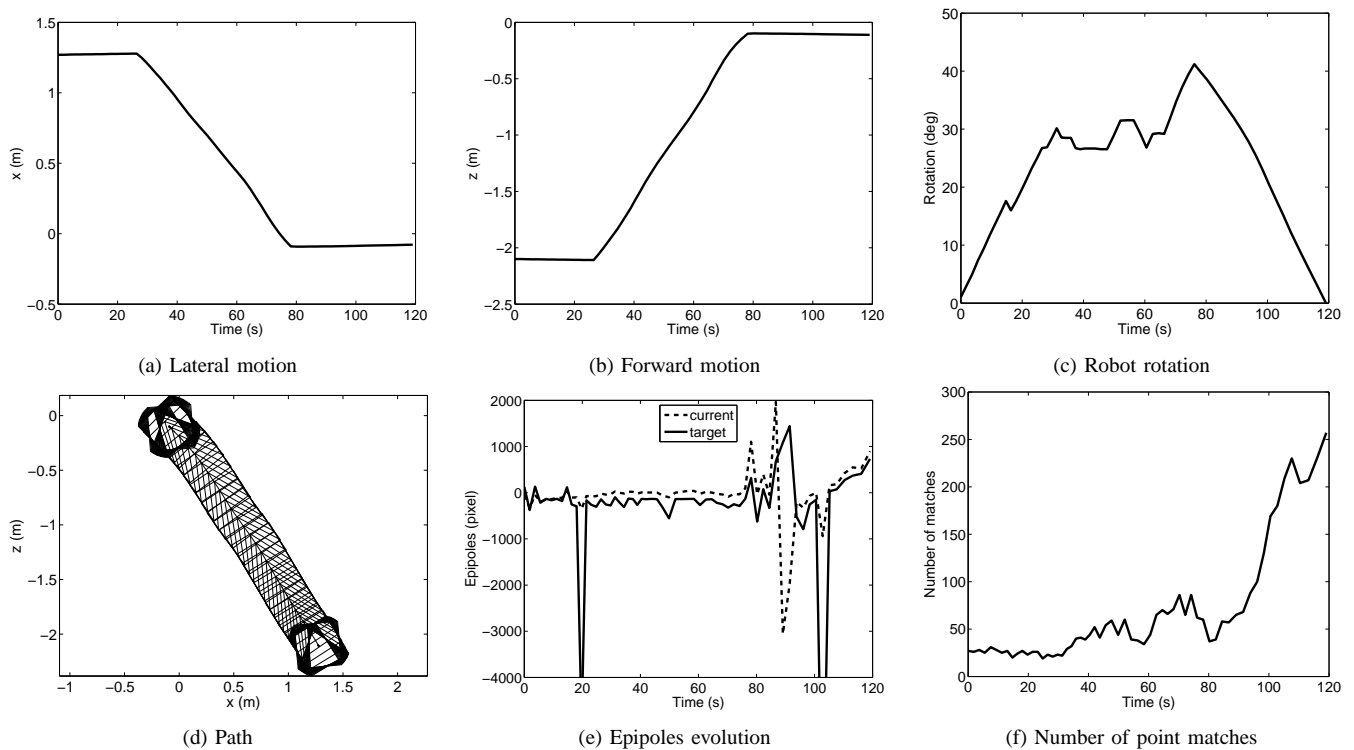It has been shown that the presented approach does not

Fig. 12.   A real experiment with start position at $(1.27, -2.1, 1°)$ and target position at $(0, 0, 0°)$.

require complete camera calibration or any particular knowledge about the environment. With only a target image taken at the desired position the controller executes a straight robot motion directly towards the target. Real experiments have been performed in a realistic indoor setting to show the validity of the approach. Simulations and real experiments have proven robustness of the method to image noise.

One of the current research issues that we are pursuing is the development of a controller able to switch between the epipolar geometry based control and a homography based control for the cases where the fundamental matrix becomes ill conditioned. In this case the same visual feature detection algorithm can be used.

REFERENCES

[1] S. Hutchinson, G. Hager, and P. Corke, "A tutorial on visual servo control," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 651–670, 1996.

[2] K. Deguchi, "Optimal motion control for image-based visual servoing by decoupling translation and rotation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Victoria, B.C., Canada, Oct. 1998, pp. 705–711.

[3] F. Chaumette and E. Malis, "2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoings," in *IEEE International Conference on Robotics and Automation*, vol. 1, San Francisco, USA, April 2000, pp. 630–635.

[4] A. Comport, M. Pressigout, Éric Marchand, and F. Chaumette, "A visual servoing control law that is robust to image outliers," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, Nevada, Oct. 2003, pp. 492–497.

[5] N. R. Gans and S. A. Hutchinson, "An asymptotically stable switched system visual controller for eye in hand robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, Nevada, Oct. 2003.

[6] V. Kyrki, D. Kragic, and H. I. Christensen, "New shortest-path approaches to visual servoing," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, Sept. 2004, pp. 349–354.

[7] S. Benhimane, E. Malis, P. Rives, and J. R. Azinheira, "Vision-based control for car platooning using homography decomposition," in *IEEE International Conference on Robotics and Automation, Barcelona, Spain*, April 2005.

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

[9] J. Piazzi and D. Prattichizzo, "An auto-epipolar strategy for mobile robot visual servoing," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2, 27–31 Oct. 2003, pp. 1802–1807.

[10] G. L. Mariottini, D. Prattichizzo, and G. Oriolo, "Epipole-based visual servoing for nonholonomic mobile robots," *IEEE International Conference on Robotics and Automation*, 2004.

[11] R. Basri, E. Rivlin, and I. Shimshoni, "Image-based robot navigation under the perspective model," in *IEEE International Conference on Robotics and Automation*, vol. 4, 10–15 May 1999, pp. 2578–2583.

[12] B. Liang and N. Pears, "Visual navigation using planar homographies," in *IEEE Conference on Robotics and Automation*, 2002, pp. 205–210.

[13] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, June 2003, pp. 257–263.

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

[15] M. Bjorkman and D. Kragic, "Combination of foveal and peripheral vision for object recognition and pose estimation," in *IEEE International Conference on Robotics and Automation*, vol. 5, 26 April–1 May 2004, pp. 5135–5140.

[16] P. Torr, *Motion Segmentation and Outlier Detection*.   University of Oxford, Department of Engineering Science, 1995.

[17] R. Hartley and A. Zisserman, Eds., *Multiple View Geometry in Computer Vision*.   New York, NY: Cambridge University Press, 2000.

[18] C. Sagüés, G. López-Nicolás, and J. Guerrero, "Visual servoing based on epipolar geometry," DIIS - I3A, Universidad de Zaragoza, Tech. Rep. V05, 2005.