IMAGE-BASED AND INTRINSIC-FREE VISUAL NAVIGATION OF A MOBILE ROBOT DEFINED AS A GLOBAL VISUAL SERVOING TASK

C. Pérez, N. García-Aracil, J.M. Azorín, J.M. Sabater, L. Navarro[2]
R. Saltarén[1]
[1]*Departamento de Automática, Electrónica e Informática Industrial Universidad Politécnica de Madrid.*
[2]*Dept. Ingeniería de Sistemas Industriales. Universidad Miguel Hernández.*
Avd. de la Universidad s/n. Edif. Torreblanca. 03202 Elche, Spain

Keywords: Visual servoing, mobile robot navigation, continuous path control.

Abstract: The new contribution of this paper is the definition of the visual navigation as a global visual control task which implies continuity problems produced by the changes of visibility of image features during the navigation. A new smooth task function is proposed and a continuous control law is obtained by imposing the exponential decrease of this task function to zero. Finally, the visual servoing techniques used to carry out the navigation are the image-based and the intrinsic-free approaches. Both are independent of calibration errors which is very useful since it is so difficult to get a good calibration in this kind of systems. Also, the second technique allows us to control the camera in spite of the variation of its intrinsic parameters. So, it is possible to modify the zoom of the camera, for instance to get more details, and drive the camera to its reference position at the same time. An exhaustive number of experiments using virtual reality worlds to simulate a typical indoor environment have been carried out.

1 INTRODUCTION

Image-based visual servoing approach is now a well known control framework (Hutchinson et al., 1996). A new visual servoing approach, which allows to control a camera with changes in its intrinsic parameters, has been published in the last years (Malis and Cipolla, 2000; Malis, 2002c). In both approaches, the reference image corresponding to a desired position of the robot is generally acquired first (during an off-line step), and some image features extracted. Features extracted from the initial image or invariant features calculated from them are used with those obtained from the desired one to drive back the robot to its reference position.

The framework for robot navigation proposed is based on pre-recorded image features obtained during a training walk. Then, we want that the mobile robot repeat the same walk by means of image-based and intrinsic-free visual servoing techniques. The main contribution of this paper are the definition of the visual navigation as a global visual control task. It implies continuity problems produced by the changes of visibility of image features during the navigation and the computing of a continuous control law associated to it.

According to our knowledge, the approximation proposed to the navigation is totally different and new in the way of dealing with the features which go in/out of the image plane during the path and similar to some references (Matsumoto et al., 1996) in the way of specifying the path to be followed by the robot.

2 AUTONOMOUS NAVIGATION USING VISUAL SERVOING TECHNIQUES

The strategy of the navigation method used in this paper is shown in Figure 1. The key idea of this method is to divide the autonomous navigation in two stages: the first one is the *training step* and the second one is the *autonomous navigation step*. During the *training step*, the robot is human commanded via radio link or whatever interface and every sample time the robot acquires an image, computes the features and stores them in memory. Then, from near its initial position, the robot repeat the same walk using the reference features acquired during the *training step*.

Pérez C., García-Aracil N., M. Azorín J., M. Sabater J., Navarro L. and Saltarén R. (2005).
IMAGE-BASED AND INTRINSIC-FREE VISUAL NAVIGATION OF A MOBILE ROBOT DEFINED AS A GLOBAL VISUAL SERVOING TASK.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics - Robotics and Automation*, pages 189-195
DOI: 10.5220/0001155301890195
Copyright © SciTePress

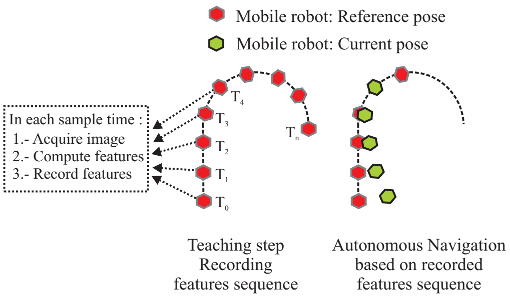


Figure 1: The strategy of the navigation method implemented. First a training step and the autonomous navigation.

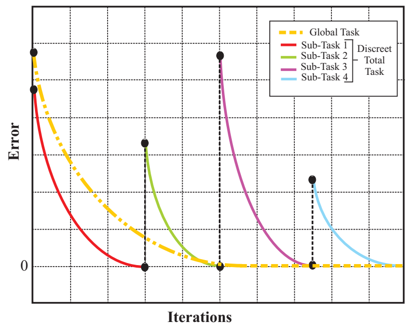


Figure 2: Navigation as a global task vs discretization of the navigation task

2.1 Control law

As it was mentioned in Section 1, image-based and intrinsic-free visual servoing approaches (Hutchinson et al., 1996; Malis and Cipolla, 2000) was used to develop the autonomous navigation of the robot. Both approaches are based on the selection of a set $\mathbf{s}$ of visual features or a set $\mathbf{q}$ of invariant features that has to reach a desired value $\mathbf{s}^*$ or $\mathbf{q}^*$. Usually, $\mathbf{s}$ is composed of the image coordinates of several points belonging to the considered target and $\mathbf{q}$ is computed as the projection of $\mathbf{s}$ in the invariant space calculated previously. In the case of our navigation method, $\mathbf{s}^*$ or $\mathbf{q}^*$ is variable with time since in each sample time the reference features is updated with the desired trajectory of $\mathbf{s}$ or $\mathbf{q}$ stored in the robot memory in order to indicate the path to be followed by the robot.

To simplify in this section, the formulation presented is only referred to image-based visual servoing. All the formulation of this section can be applied directly to the invariant visual servoing approach changing $\mathbf{s}$ by $\mathbf{q}$. The visual task function (Samson et al., 1991) is defined as the regulation of an global error function instead of a set of discrete error functions (Figure 2):

$$\mathbf{e} = \mathbf{C}(\mathbf{s} - \mathbf{s}^*(t)) \tag{1}$$

The derivative of the task function,considering $\mathbf{C}$ constant, will be:

$$\dot{\mathbf{e}} = \mathbf{C}(\dot{\mathbf{s}} - \dot{\mathbf{s}}^*) \tag{2}$$

It is well known that the interaction matrix $\mathbf{L}$, also called image jacobian, plays a crucial role in the design of the possible control laws. $\mathbf{L}$ is defined as:

$$\dot{\mathbf{s}} = \mathbf{L}\,\mathbf{v} \tag{3}$$

where $\mathbf{v} = (\mathbf{V^T}, \omega^\mathbf{T})$ is the camera velocity screw ($\mathbf{V}$ and ω represent its translational and rotational component respectively).

Plugging the equation (3) in (2) we obtain:

$$\dot{\mathbf{e}} = \mathbf{CLv} - \mathbf{C}\,\dot{\mathbf{s}}^* \tag{4}$$

A simple control law can be obtained by imposing the exponential convergence of the task function to zero:

$$\dot{\mathbf{e}} = -\lambda\,\mathbf{e} \quad \text{so} \quad \mathbf{CLv} = -\lambda\,\mathbf{e} + \mathbf{C}\,\dot{\mathbf{s}}^* \tag{5}$$

where λ is a positive scalar factor which tunes the speed of convergence:

$$\mathbf{v} = -\lambda\,(\mathbf{CL})^{-1}\mathbf{e} + (\mathbf{CL})^{-1}\mathbf{C}\,\dot{\mathbf{s}}^* \tag{6}$$

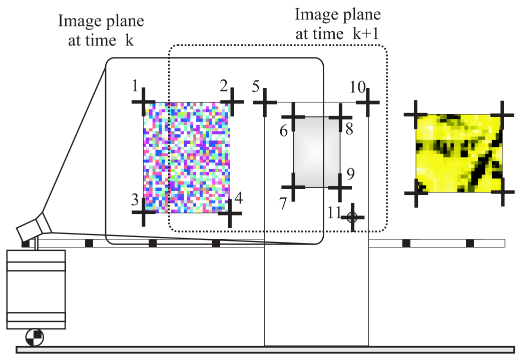
if $\mathbf{C}$ is setting to $\mathbf{L}^+$, then $(\mathbf{CL}) > 0$ and the task function converge to zero and, in the absence of local minima and singularities, so does the error $s - s^*$. Finally substituting $\mathbf{C}$ by $\mathbf{L}^+$ in equation (6), we obtain the expression of the camera velocity that is sent to the robot controller:

$$\mathbf{v} = -\lambda\,\mathbf{L}^+\,(\mathbf{s} - \mathbf{s}^*(t)) + \mathbf{L}^+\,\dot{\mathbf{s}}^* \tag{7}$$

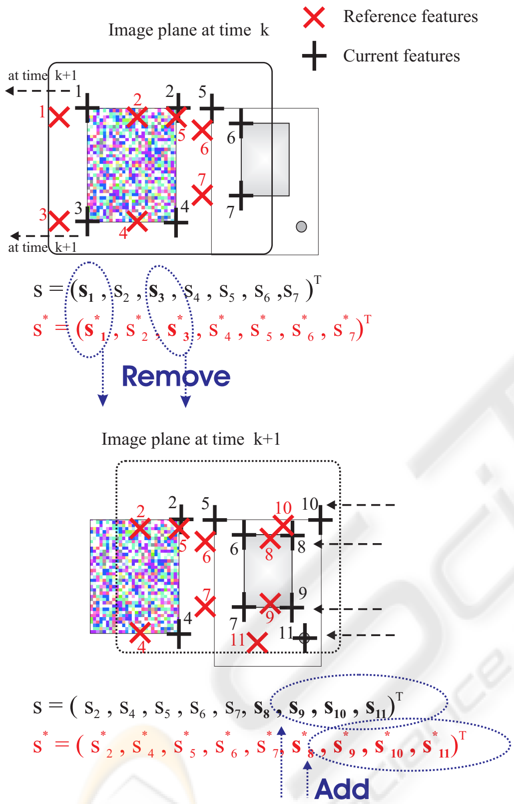
Remember that the whole formulation is directly applicable to the invariant visual servoing approach changing $\mathbf{s}$ by $\mathbf{q}$. As will be shown in the next section, discontinuities in the control law will be produced by the appearance/disappearance of image features during the navigation of the robot. The answer to the question: *why are these discontinuities produced* and their solution are presented in the following section.

3 DISCONTINUITIES IN VISUAL NAVIGATION

In this section, we describe more in details the discontinuity problem that occurs when some features go in/out of the image during the vision-based control. A simple simulation illustrate the effects of the discontinuities on the control law and on the performances of the visual servoing. The navigation of a mobile robot is a typical example where this kind of problems are produced.



(a) Croquis of the autonomous navigation of the robot



(b) Appearance/Disappearance of image features from time k to k+1

Figure 3: Navigation of a mobile robot controlled by visual servoing techniques.

3.1 What happens when features appear or disappear from the image plane?

During autonomous navigation of the robot, some features appear or disappear from the image plane so they will must be added to or removed from the visual error vector (Figure 3). This change in the error vector produces a jump discontinuity in the control law. The magnitude of the discontinuity in control law depends on the number of the features that go in or go out of the image plane at the same time, the distance between the current and reference features, and the pseudoinverse of interaction matrix.

In the case of using the invariant visual servoing approach to control the robot, the effect produced by the appearance/disappearance of features could be more important since the invariant space $\mathcal{Q}$ used to compute the current and the reference invariant points$(\mathbf{q}, \mathbf{q}^*)$ changes with features(Malis, 2002c).

4 CONTINUOUS CONTROL LAW FOR NAVIGATION

In the previous section, the continuity problem of the control law due to the appearance/disappearance of features has been shown. In this section a solution is presented. The section is organized as follows. First, the concept of weighted features is defined. Then, the definition of a smooth task function is presented. Finally, the reason to reformulate the invariant visual servoing approach and its development is explained.

4.1 Weighted features

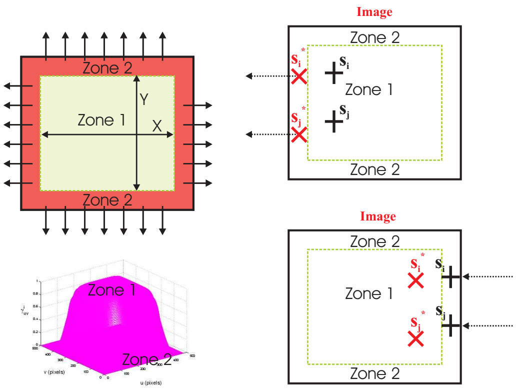
The key idea in this formulation is that every feature (points, lines, moments, etc) has its own weight which may be a function of image coordinates(u,v) and/or a function of the distance between feature points and an object which would be able to occlude them, etc (García et al., 2004). To compute the weights, three possible situations must be taking into account:

4.1.1 Situation 1: Changes of visibility through the border of the image (Zone 2 in Figure 4 a)

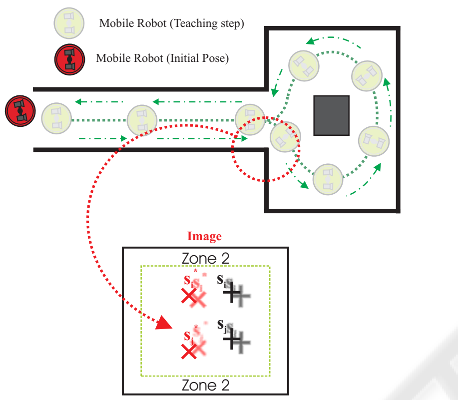
To anticipate the changes of visibility of features through the border, a total weight Φ_{uv} is computed as the product of the weights of the current and reference features which are function of their position in the image $(\gamma^i_{uv}, {\gamma^i_{uv}}^*)$. The weight $\gamma^i_{uv} = \gamma^i_u \cdot \gamma^i_v$ is computed using the definition of the function $\gamma_y(x)$ $(\gamma^i_u = \gamma_y(u_i)$ and $\gamma^i_v = \gamma_y(v_i))$ respectively) (García et al., 2004).

4.1.2 Situation 2: The sudden appearance of features on the center of the image (Zone 1 in Figure 4 b)

To take into account this possible situation, every new features (current and reference) must be checked pre-



(a) Border of the image



(b) Center of the image

Figure 4: Appearance/Disappearance of image features through the border and the center of the image during the navigation of the mobile robot.

viously to known if they are in Zone 2 or Zone 1. If the new features are in Zone 2, the appearance of the features are considered in Situation 1. If they are in Zone 1, a new weight function must be defined to add these new features in a continuous way.

The weight function proposed Φ_a^i is a exponential function that tends to 1, reaching its maximum after a certain number of steps which can be modified with the α and β parameters (Figure 5):

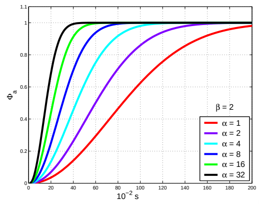
$$\Phi_a^i(t) = 1 - e^{-\alpha \cdot t^{\beta}} \quad \alpha,\ \beta > 0 \tag{8}$$

4.1.3 Situation 3: The sudden disappearance of features on the center of the image because of a temporal or definitive occlusion (Zone 1 in Figure 4 b)

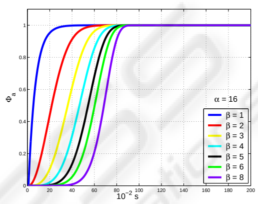
In this situation, the occlusions produced in the teaching step on the Zone 1 are only considered since they can be easily anticipated by the observation of reference features vector prerecorded previously.

To take into account this possible situation, a new weight function must be defined to remove these features from the current and reference vector in a continuous way. The weight function proposed Φ_o^i is a exponential function that tends to 0, reaching its minimum after a certain number of steps which can be modified with the ν and σ parameters:

$$\Phi_o^i(t) = e^{-\nu \cdot t^{\sigma}} \quad \nu,\ \sigma > 0 \tag{9}$$



(a) $\Phi_a(\alpha)$



(b) $\Phi_a(\beta)$

Figure 5: Plotting Φ_a in function of α and β.

4.1.4 Global weight function Φ^i

In this section, a global weight function Φ^i which includes the three possible situations commented before is presented. This function is defined as the product of the three weight functions (Φ_{uv}^i, Φ_a^i y Φ_o^i) which takes into account the three possible situations:

$$\Phi^i = \Phi_{uv}^i \cdot \Phi_a^i \cdot \Phi_o^i \quad \text{where} \quad \Phi^i \in [0, 1] \tag{10}$$

4.2 Smooth Task function

Suppose that n matched points are available in the current image and in the reference features stored. Everyone of these points(current and reference) will have a weight Φ^i which can be computed as it's shown in the previous subsection 4.1. With them and their weights, a task function can be built (Samson et al., 1991):

$$\mathbf{e} = \mathbf{CW}\ (\mathbf{s} - \mathbf{s}^*(t)) \tag{11}$$

where $\mathbf{W}$ is a $(2n \times 2n)$ diagonal matrix where its elements are the weights Φ^i of the current and reference features multiplied by the weights of the reference features.

The derivate of the task function will be:

$$\dot{\mathbf{e}} = \mathbf{CW}\ (\dot{\mathbf{s}} - \dot{\mathbf{s}}^*) + (\mathbf{C}\dot{\mathbf{W}} + \dot{\mathbf{C}}\mathbf{W})(\mathbf{s} - \mathbf{s}^*(t)) \tag{12}$$

Plugging the equation ($\dot{\mathbf{s}} = \mathbf{L}\ \mathbf{v}$) in (12) we obtain:

$$\dot{\mathbf{e}} = \mathbf{CW}\ (\mathbf{Lv} - \dot{\mathbf{s}}^*) + (\mathbf{C}\dot{\mathbf{W}} + \dot{\mathbf{C}}\mathbf{W})(\mathbf{s} - \mathbf{s}^*(t)) \tag{13}$$

A simple control law can be obtained by imposing the exponential convergence of the task function to zero

($\dot{\mathbf{e}} = -\lambda \mathbf{e}$), where λ is a positive scalar factor which tunes the speed of convergence:

$$\begin{aligned}\mathbf{v} &= -\lambda\,(\mathbf{CWL})^{-1}\mathbf{e} + (\mathbf{CWL})^{-1}\mathbf{CW}\dot{\mathbf{s}}^* + \\ &- (\mathbf{CWL})^{-1}\,(\mathbf{C}\dot{\mathbf{W}} + \dot{\mathbf{C}}\mathbf{W})\,(\mathbf{s} - \mathbf{s}^*(t))\end{aligned} \quad (14)$$

if $\mathbf{C}$ is setting to $(\mathbf{W}^*\mathbf{L}^*)^+$,then $(\mathbf{CWL}) > 0$ and the task function converge to zero and, in the absence of local minima and singularities, so does the error $\mathbf{s}-\mathbf{s}^*$. In this case, $\mathbf{C}$ is constant and therefore $\dot{\mathbf{C}} = 0$. Finally substituting $\mathbf{C}$ by $(\mathbf{W}^*\mathbf{L}^*)^+$ in equation (14), we obtain the expression of the camera velocity that is sent to the robot controller:

$$\begin{aligned}\mathbf{v} &= -(\mathbf{W}^*\mathbf{L}^*)^+\,(\lambda\,\mathbf{W} + \dot{\mathbf{W}})\,(\mathbf{s} - \mathbf{s}^*(t)) + \\ &+ (\mathbf{W}^*\mathbf{L}^*)^+\mathbf{W}\dot{\mathbf{s}}^*\end{aligned} \quad (15)$$

4.3 Visual servoing techniques

The visual servoing techniques used to carry out the navigation are the image-based and the intrinsic-free approaches. In the case of image-based visual servoing approach, the control law (15) is directly applicable to assure a continuous navigation of a mobile robot. On the other hand, when the intrinsic-free approach is used, this technique must be reformulated to take into account the weighted features.

4.3.1 Intrinsic-free approach

The theoretical background about invariant visual servoing can be extensively found in (Malis, 2002b; Malis, 2002c). In this section, we modify the approach in order to take into account weighted features (García et al., 2004).

Basically, the weights Φ^i defined in the previous subsection must be *redistributed*(γ_i) in order to be able to build the invariant projective space $\mathcal{Q}^{\gamma_i}$ where the control will be defined.

Similarly to the standard intrinsic-free visual servoing, the control of the camera is achieved by stacking all the reference points of space $\mathcal{Q}^{\gamma_i}$ in a ($3n\times1$) vector $\mathbf{s}^*(\boldsymbol{\xi}^*) = (\mathbf{q}_1^*(t), \mathbf{q}_2^*(t), \cdots, \mathbf{q}_n^*(t))$. Similarly, the points measured in the current camera frame are stacked in the ($3n\times1$) vector $\boldsymbol{s}(\boldsymbol{\xi}) = (\boldsymbol{q}_1(t), \boldsymbol{q}_2(t), \cdots, \boldsymbol{q}_n(t))$. If $\boldsymbol{s}(\boldsymbol{\xi}) = \boldsymbol{s}^*(\boldsymbol{\xi}^*)$ then $\boldsymbol{\xi} = \boldsymbol{\xi}^*$ and the camera is back to the reference position whatever the camera intrinsic parameters.

In order to control the movement of the camera, we use the control law (15) where $\mathbf{W}$ depends on the weights previously defined and $\mathbf{L}$ is the interaction matrix. The interaction matrix depends on current normalized points $\boldsymbol{m}_i(\boldsymbol{\xi}) \in \mathcal{M}$ ($\boldsymbol{m}_i$ can be computed from image points $\boldsymbol{m}_i = \boldsymbol{K}^{-1}\,\boldsymbol{p}_i$), on the invariant points $\boldsymbol{q}_i(\boldsymbol{\xi}) \in \mathcal{Q}^{\gamma}$, on the current depth distribution $\boldsymbol{z}(\boldsymbol{\xi}) = (Z_1, Z_2, ..., Z_n)$ and on the current redistributed weights γ_i. The interaction matrix in the weighted invariant space ($\boldsymbol{L}_{qi}^{\gamma_i} = \boldsymbol{T}_{mi}^{\gamma_i}\,(\boldsymbol{L}_{mi} - \boldsymbol{C}_i^{\gamma_i})$)is obtained like in (Malis, 2002a) but the term $\boldsymbol{C}_{\boldsymbol{i}}^{\gamma_i}$ must be recomputed in order to take into account the redistributed weights γ_i.

5 EXPERIMENTS IN A VIRTUAL INDOOR ENVIRONMENT

Exhaustive experiments have been carried out using a virtual reality tool for modeling an indoor environment. To make more realistic simulation, errors in intrinsic and extrinsic parameters of the camera mounted in the robot and noise in the extraction of image features have been considered. An estimation $\widehat{\mathbf{K}}$ of the real matrix $\mathbf{K}$ has been used with an error of 25% in focal length and a deviation of 50 pixels in the position of the optical center. Also an estimation $\widehat{\mathbf{T}}_{RC}$ of the camera pose respect to the robot frame has been computed with a rotation error of $u\theta = [3.75\ 3.75\ 3.75]^T\ degrees$ and translation error of $t = [2\ 2\ 0]^T\ cm$. An error in the extraction of current image features has been considered by adding a normal distribution noise to the accurate image features extracted.

In Figure 7, the control signals sent to the robot controller using the classical image-based approach and the image-based approach with weighted features are shown. In Figure 7 (a,b,c), details of the control law using the classical image-based approach, where the discontinuities can be clearly appreciated, are presented. To show the improvements of the new formulation presented in this paper, the control law using the image-based with weighted features can be seen in Figure 7 (d,e,f).

The same experiment, but in this case using the intrinsic-free visual servoing approach, is performed. In Figure 8 , the control signals sent to the robot controller using the intrinsic-free approach and some details, where the discontinuities can be clearly appreciated, are shown. The improvements of the new formulation of the intrinsic-free approach with weighted features are presented in Figure 9. The same details of the control law shown in Figure 8 are presented in Figure 9. Comparing both figures and their details, the continuity of the control law is self-evident despite of the noise in the extraction of image features.

Also in (García et al., 2004), a comparison between this method and a simple filtration of the control law was presented. The results presented in that paper corroborate that the new approach with weighted features to the problem works better than a simple filter of the control signals.

6 CONCLUSIONS

In this paper the originally definition of the visual navigation as a global visual control task is presented. It implies continuity problems produced by the changes of visibility of image features during the navigation which have been solved by the definition of a smooth task function and a continuous control law obtained from it. The results presented corroborate that the new approach is continuous, stable and works better than a simple filter of the control signals. The validation of this results with a real robot is on the way by using a B21r mobile robot from iRobot company.

REFERENCES

García, N., Malis, E., Reinoso, O., and Aracil, R. (2004). Visual servoing techniques for continuous navigation of a mobile robot. In *1st International Conference on Informatics in Control, Automation and Robotics*, volume 1, pages 343–348, Setubal, Portugal.

Hutchinson, S. A., Hager, G. D., and Corke, P. I. (1996). A tutorial on visual servo control. *IEEE Trans. Robotics and Automation*, 12(5):651–670.

Malis, E. (2002a). Stability analysis of invariant visual servoing and robustness to parametric uncertainties. In *Second Joint CSS/RAS International Workshop on Control Problems in Robotics and Automation*, Las Vegas, Nevada.

Malis, E. (2002b). A unified approach to model-based and model-free visual servoing. In *European Conference on Computer Vision*, volume 2, pages 433–447, Copenhagen, Denmark.

Malis, E. (2002c). Vision-based control invariant to camera intrinsic parameters: stability analysis and path tracking. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 217–222, Washington, USA.

Malis, E. and Cipolla, R. (2000). Self-calibration of zooming cameras observing an unknown planar structure. In *Int. Conf. on Pattern Recognition*, volume 1, pages 85–88, Barcelona, Spain.

Matsumoto, Y., Inaba, M., and Inoue, H. (1996). Visual navigation using view-sequenced route representation. In *IEEE International Conference on Robotics and Automation (ICRA'96)*, volume 1, pages 83–88, Minneapolis, USA.

Samson, C., Le Borgne, M., and Espiau, B. (1991). *Robot Control: the Task Function Approach.* volume 22 of Oxford Engineering Science Series.Clarendon Press., Oxford, UK, 1st edition.

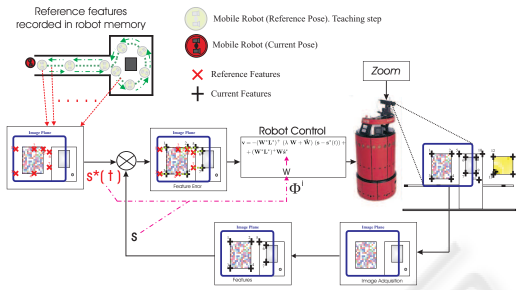
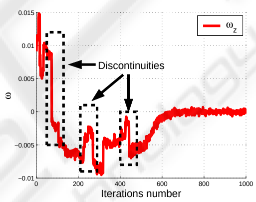
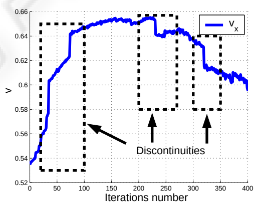


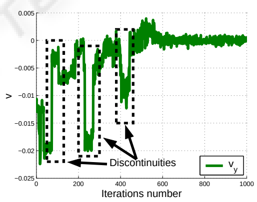
Figure 6: Block diagram of the controller proposed



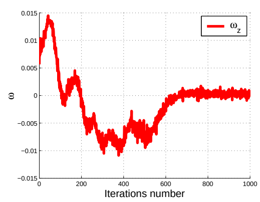
(a) ω_z



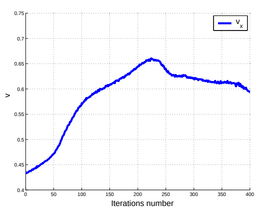
(b) v_x



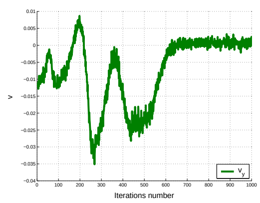
(c) v_y



(d) ω_z

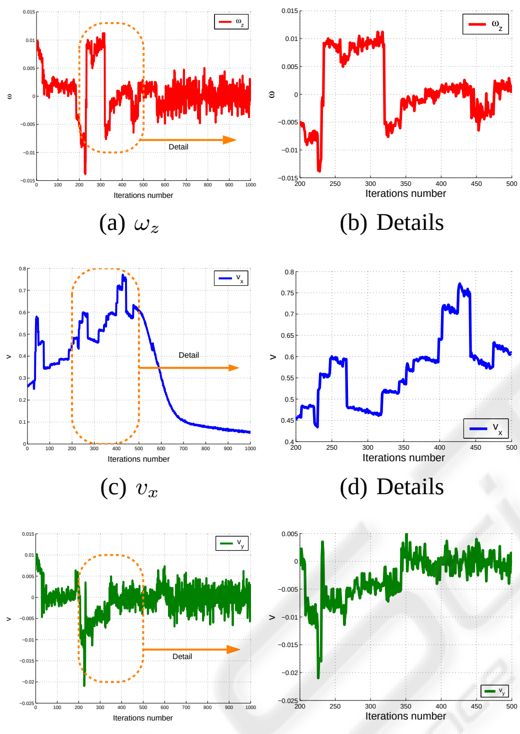


(e) v_x



(f) v_y

Figure 7: Control law: Classical image-based approach (a-b-c) and image-based approach with weighted features (d-e-f).The translation and rotation speeds are measured respectively in $\frac{m}{s}$ and $\frac{deg}{s}$

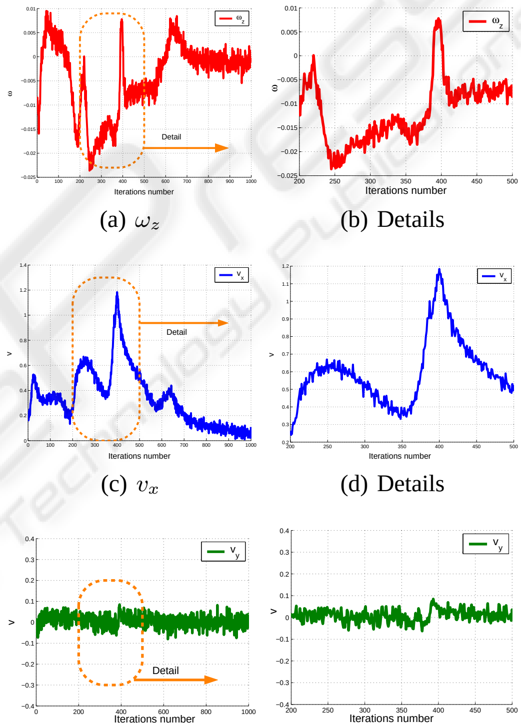


(a) ω_z (b) Details

(c) v_x (d) Details

(e) v_y (f) Details

Figure 8: Discontinuities in the control law: Intrinsic-free approach.



(a) ω_z (b) Details

(c) v_x (d) Details

(e) v_y (f) Details

Figure 9: Continuous control law: Intrinsic-free approach with weighted features.