

Indoor Navigation of a Wheeled Mobile Robot along Visual Routes

Guillaume Blanc, Y. Mezouar, Philippe Martinet

▶ **To cite this version:**

Guillaume Blanc, Y. Mezouar, Philippe Martinet. Indoor Navigation of a Wheeled Mobile Robot along Visual Routes. 2005 IEEE International Conference on Robotics and Automation, Apr 2005, Barcelona, Spain. pp.3354-3359, 10.1109/ROBOT.2005.1570628 . hal-02467192

HAL Id: hal-02467192
https://inria.hal.science/hal-02467192

Submitted on 5 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Indoor Navigation of a Wheeled Mobile Robot along Visual Routes

Guillaume Blanc, Youcef Mezouar and Philippe Martinet
LASMEA
UBP Clermont II, CNRS - UMR6602
24 Avenue des Landais, 63177 AUBIERE, FRANCE
firstname.lastname@lasmea.univ-bpclermont.fr

***Abstract*— When navigating in an unknown environment for the first time, a natural behavior consists in memorizing some key views along the performed path, in order to use these references as checkpoints for a future navigation mission taking a similar path. This assumption is used in this paper as the basis of a navigation framework for wheeled mobile robots in indoor environments. During a human-guided teleoperated learning step, the robot performs paths which are sampled and stored as a set of ordered key images, acquired by a standard embedded camera. The set of these obtained visual paths is topologically organized and provides a visual memory of the environment. Given an image of one of the visual paths as a target, the robot navigation mission is defined as a concatenation of visual path subsets, called visual route. When running autonomously, the robot is controlled by a visual servoing law adapted to its nonholonomic constraint. Based on the regulation of successive homographies, this control guides the robot along the reference visual route without explicitly planning any trajectory. Real experiment results illustrate the validity of the presented framework.**

***Index Terms*— Visual Navigation, Visual Memory, Visual Servoing for Mobile Robots**

I. INTRODUCTION

Vision is a central clue of most of recent mobile robots navigation frameworks. The authors of [4] accounts of twenty years of works at the meeting point of mobile robotics and computer vision communities. Often used among more "traditional" embedded sensors - proprioceptive sensors like odometers as exteroceptive ones like sonars - it provides accurate localization methods. In many works, and especially those dealing with indoor navigation as in [8], computer vision techniques are used in a landmark-based framework. Identifying extracted landmarks to known references allows to update the results of the localization algorithm. These methods are based on some knowledges about the environment, such as a given 3D model or a map built online. They generally rely on a complete or partial 3D reconstruction of the observed environment through the analysis of data collected from disparate sensors. The mobile robot can thus be localized in an absolute reference frame. Both motion planning and robot control can then be designed in this space. The results obtained by the authors of [15] leave to be forcasted that such a framework will be reachable using a single camera. However, although an accurate global localization is unquestionably useful, our aim is to build a complete vision-based framework without recovering a the position of the mobile robot with respect to a reference frame. The authors of [4] call this kind of framework a qualitative approach of navigation.

The principle of this approach is to represent the robot environment with a bounded quantity of images gathered in a set called visual memory. In [14], this concept is exploited to control the 6 *dof* of a robotic arm under large displacements. A set of images is extracted from a previously learnt database which describes successive targets for a global visual servoing task. The authors do not consider the kinematic constraints of a mobile robot. In the context of mobile robotics, [12] also proposes to use a sequence of images, but recorded during a human teleoperated motion, and called View-Sequenced Route Reference. This concept underlines the close link between a human-guided learning and the performed paths during an autonomous run. However, the automatic control of the robot in [12] is not formulated as a visual servoing task.

In this paper, we propose a complete image-based framework (*i.e* from environment learning to control) for mobile robots navigation. A sequence of images, acquired during a human-guided learning, allows to derive paths driving the robot from its initial to its goal locations. In order to reduce the complexity of the image sequences, only key views are stored and indexed on a visual path. The set of visual paths can be interpreted as a visual memory of the environment. The visual memory is structured as a graph which takes into account the environment topology. A navigation task consists then in performing autonomously a visual route which is a concatenation of visual paths. The visual route connects thus in the sensor space the initial and goal configurations. Section II details more precisely this point.

The Section III deals with the vision-based control scheme designed to control the robot motions along a visual route. The nonholonomic constraints of most current wheeled mobile robots makes the classical visual servoing methods unexploitable since the camera is fixed on the robot [18]. However, motivated by the development of 2D 1/2 visual-servoing method proposed by Malis *et al* (see [11]), some authors have investigated the use of homography and epipolar geometry to stabilize mobile robots [5], [3]. In this paper, because the notions of visual route and path are very closed, we turn the nonholonomic visual-servoing issue into a path following one. The designed control law

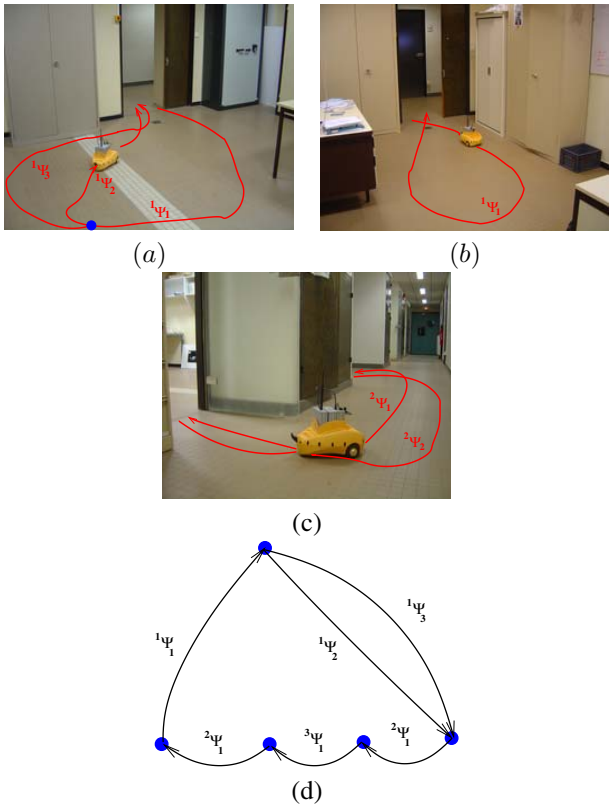


Fig. 1. Building a visual memory: Into the rooms (a) and (b) and the corridor (c), the paths ${}^{r}\Psi_p$ have been learnt by teleoperating the robot. As a result, the graph (d) represents the topological organization of the visual memory. The blue circles show the vertices

does not need any explicit off-line path planning step.
In Section IV, experiments on a small indoor wheeled mobile robot illustrate the implementation of the proposed framework.

II. Visual memory and routes

In [4], approaches using a "memorization" of images of the environment taken from an embedded camera are ranked among mapless navigation systems. Indeed, as proposed in [12] or in [9], any notion of map nor topology of the environment appears, neither to build the reference set of images, nor for the automatic guidance of the robot. The first step of our framework consists on a learning stage to build the visual memory. The visual memory is structured according to the environment topology to reduce the computational cost. We assume that the features observed by the camera are quasi-static. However, the remainder of the environment is not restricted to staticity, since both perception and control algorithms are robust to dynamic constraints such as occlusions.

A. Structure of the visual memory

The learning stage relies on the human experience. The user guides the mobile robot into its workspace. In the case of an urban vehicle for instance, as proposed in [15], the user manually drives the robot along all the streets where it will go down in an autonomous way. To each street is associated a visual path, which is a set of key images ordered from the beginning to the end of the street. We use a similar learning approach in an indoor environment. The user guides the mobile robot along one or several paths into each room where the robot is authorized to go (see Figure 1 (a),(b) and (c)). A visual path ${}^{r}\Psi_p$ is then stored and indexed as the p^{th} learnt path in the r^{th} room.

1) Visual paths: A visual path ${}^{r}\Psi_p$ is composed of n key images:

$$ {}^{r}\Psi_p = \{{}^{r}\mathcal{I}_i^p | i = \{1, 2, \ldots, n\}\} $$

For control purpose (refer to Section III), the authorized motions during the learning stage are assumed to be limited to those of a car-like vehicle, which only goes forward. The following Hypothesis 2.1 formalizes these constraints.

Hypothesis 2.1: Given two frames ${}^{R}\mathcal{F}_i$ and ${}^{R}\mathcal{F}_{i+1}$, respectively associated to the mobile robot when two successive key images $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ of a visual path Ψ were acquired, there exists an admissible path ψ from ${}^{R}\mathcal{F}_i$ to ${}^{R}\mathcal{F}_{i+1}$ for a car-like vehicle whose turn radius is bounded, and which only moves forward.

Moreover, because the controller is vision-based, the robot is controllable from ${}^{r}\mathcal{I}_i^p$ to ${}^{r}\mathcal{I}_{i+1}^p$ only if the hereunder Hypothesis 2.2 is respected.

Hypothesis 2.2: Two successive key images $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ contain a set $\mathcal{P}_i$ of matched visual features, which can be tracked along a path performed between ${}^{R}\mathcal{F}_i$ and ${}^{R}\mathcal{F}_{i+1}$.

During the acquisition of a visual path, the Hypothesis 2.2 constrains the choice of the key images. Two main approaches have been implemented in our laboratory. The first one, described in [1], uses planar patterns tracking. A new key image $\mathcal{I}_{i+1}$ is stored when a pattern, which has been tracked since $\mathcal{I}_i$ was acquired, is likely to leave the image. In the second approach, proposed in [15], $\mathcal{I}_{i+1}$ is stored if the camera motion between $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ is sufficient to ensure a good 3D reconstruction of matched interest points, and if the number of interest points is upper than a fixed threshold. In this case, $\mathcal{P}_i$ is the set of all correctly matched points between $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$. As a consequence of Hypothesis 2.1 and 2.2, each visual path ${}^{r}\Psi_p$ corresponds to an oriented edge which connects two configurations of the robot's workspace. Moreover, the number of key images of a visual path is directly linked to the human-guided path complexity. According to this parameter, we define the *value of a visual path* as its cardinal.

2) Visual memory vertices: In order to connect two visual paths, the terminal extremity of one of them and the initial extremity of the other one must be constrained as two consecutive key images of a visual path. The paths are then connected by a vertex, and two adjacent vertices of the visual memory are connected by a visual path (see Figure 1 (d)).

Proposition 2.1: Given two visual paths $\Psi_{p_1} = \{\mathcal{I}_i^{p_1} | i = \{1, 2, \ldots, n_1\}\}$ and $\Psi_{p_2} = \{\mathcal{I}_i^{p_2} | i = \{1, 2, \ldots, n_2\}\}$, if the two key images $\mathcal{I}_{n_1}^{p_1}$

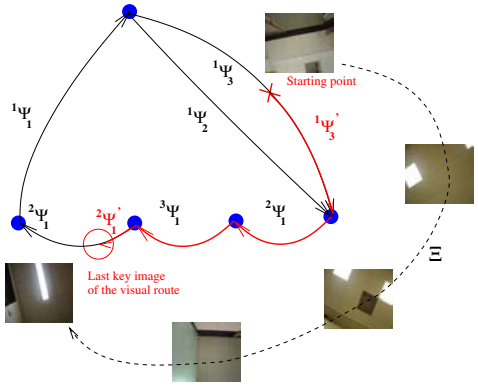


Fig. 2. A visual route: $\Xi = {}^1\Psi'_3 \oplus^2 \Psi_1 \oplus^3 \Psi_1 \oplus^2 \Psi'_1$. ${}^1\Psi'_3$ and ${}^2\Psi'_1$ are subsets of respectively ${}^1\Psi_3$ and ${}^2\Psi_1$. ${}^1\Psi_3$ is splitted at the closest key image to the initial $\mathcal{I}_c$, while the last key image of ${}^2\Psi'_1$ is a desired image to reach by navigating onto the visual memory.

and $\mathcal{I}_1^{p_2}$ abide by both Hypothesis 2.1 and 2.2, then a vertex connects Ψ_{p_1} to Ψ_{p_2}.

We also assume this Proposition 2.1 in the particular case where the terminal extremity of a visual path Ψ_{p_1} is the same key image as the initial extremity of another visual path Ψ_{p_2}. This is useful in practice, when building the visual memory.

3) A strongly connected digraph: According to Sections II-A.1 and II-A.2, the visual memory structure is defined as a digraph, whose arcs are the visual paths. It is yet necessary that this digraph is strongly connected. Indeed, this condition warrants that any vertex of the visual memory is attainable from every others, through a set of visual path.

B. Visual route

A visual route describes the robot's mission in the sensor space. Given two key images of the visual memory $\mathcal{I}_c$ and $\mathcal{I}_g$, corresponding respectively to the current and goal locations of the robot, a visual route is a set of key images which describes a path from $\mathcal{I}_c$ to $\mathcal{I}_g$, as presented in Figure 2. In practice, $\mathcal{I}_c$ is not exactly a key image. It is then necessary to determine which stored key image is the closest one to $\mathcal{I}_c$. This can be done in an off-line stage, as Remazeilles *et al* propose in [14], by comparing the photometric invariants of the request image with those of the images store onto the visual memory. Assuming that the user is able to point out which is the room r where the robot is staying, the search of the closest image is bounded to the learnt visual paths in r. The visual route is the minimum length path of the visual memory connecting two vertices associated to $\mathcal{I}_c$ and $\mathcal{I}_g$. According to the definition of the value of a visual path, the length of a path is the sum the values of its arcs. The minimum length path is obtained in a very simple way, using Dijkstra's algorithm. Consequently, the visual route results from the concatenation of indexed visual paths. Given two visual paths Ψ_{p_1} and Ψ_{p_2}, respectively containing n_1 and n_2 indexed key images, the concatenation operation of Ψ_{p_1} and Ψ_{p_2} is defined as follows:

$$\Psi_{p_1} \oplus \Psi_{p_2} = \left\{\mathcal{I}_j^{p_{1,2}} | j = \{1, \ldots, n_1, n_1+1, \ldots, n_1+n_2\}\right\}$$

$$\mathcal{I}_j^{p_{1,2}} = \begin{cases} \mathcal{I}_j^{p_1} \text{ if } j \leqslant n_1 \\ \mathcal{I}_{j-n_1}^{p_2} \text{ if } n_1 \leqslant j \leqslant n_1+n_2 \end{cases}$$

The visual route describes a set of consecutive states that the image has to reach in order that the robot joins the goal configuration from the initial one. The robot motions are controlled along the visual route using the data provided by the embedded camera. The next section deals with this issue.

III. Visual route following

Visual-servoing is often considered as a way to achieve positioning tasks. Classical methods, based on the task function formalism, are based on the existence of a diffeomorphism between the sensor space and the robot's configuration space. Due to the nonholomic constraints of most of wheeled mobile robots, under the condition of rolling without slipping, such a diffeomorphism does not exist if the camera is rigidly fixed to the robot. In [18], the authors add extra degrees of freedom to the camera. The camera pose can then be regulated in a closed loop.

In the case of an embedded and fixed camera, the control of the camera is generally based on wheeled mobile robots control theory [16]. In [10], a car-like robot is controlled with respect to the projection of a ground curve in the image plane. The control law is formalized as a path following problem. More recently, in [5] and [3], a partial estimation of the camera displacement between the current and desired views has been exploited to design vision-based control laws. The camera displacement is estimated by uncoupling translation and rotation components of an homography matrix. In [5], a time-varying control allows an asymptotical stabilization on a desired image. In [3], a trajectory following task is achieved. The trajectory to follow is defined by a prerecorded video and the control law is proved stable using Lyapunov-based analysis. In our case, unlike a whole video sequence, we deal with a set of relay images which have been acquired from geometrically spaced out points of view.

A visual route following can be considered as a sequence of visual-servoing tasks. A stabilization approach could thus be used to control the camera motions from a key image to the next one. However, a visual route is fundamentally a path. In [13], the authors propose to plan the trajectories of image features directly in the sensor space from the first image to the last one. These trajectories are then used as references to control a robotic arm.

To design the controller, described in the sequel, the key images of the reference visual route are considered as consecutive checkpoints to reach in the sensor space. The control problem is formulated as a path following to guide the nonholonomic mobile robot along the visual route.

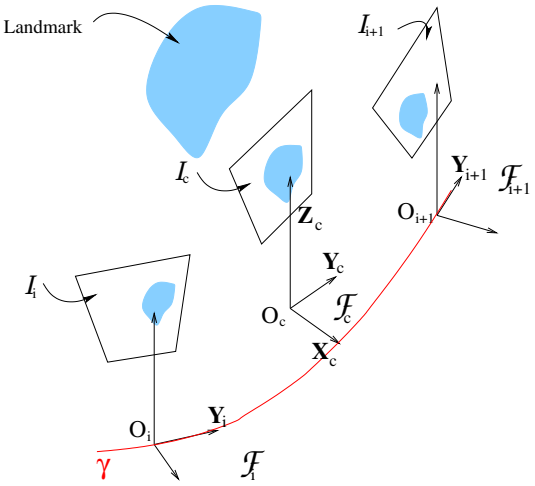


Fig. 3. Frames and images: $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ are two consecutive key images, acquired along a teleoperated path γ

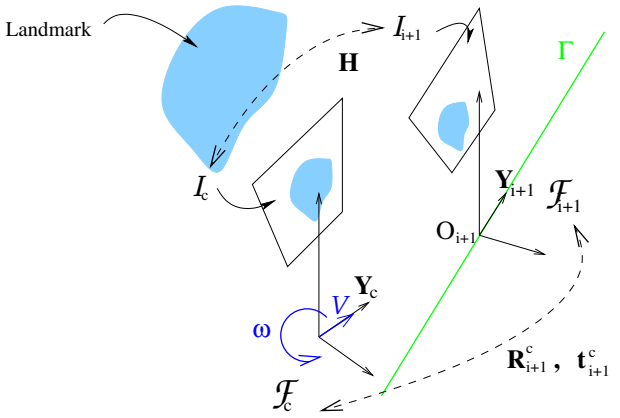


Fig. 4. Control strategy

A. Assumptions and models

Let $\mathcal{I}_i$, $\mathcal{I}_{i+1}$ be two consecutive key images of a given visual route to follow and $\mathcal{I}_c$ be the current image. Let us note $\mathcal{F}_i = (O_i, \mathbf{X_i}, \mathbf{Y_i}, \mathbf{Z_i})$ and $\mathcal{F}_{i+1} = (O_{i+1}, \mathbf{X_{i+1}}, \mathbf{Y_{i+1}}, \mathbf{Z_{i+1}})$ the frames attached to the robot when $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$ were stored and $\mathcal{F}_c = (O_c, \mathbf{X_c}, \mathbf{Y_c}, \mathbf{Z_c})$ a frame attached to the robot in its current location. Figure 3 illustrates this setup. The origin O_c of $\mathcal{F}_c$ is on the axle midpoint of a cart-like robot, which evolutes on a perfect ground plane.

The control vector of the considered cart-like robot is $\mathbf{u} = [V, \omega]^T$ where V is the longitudinal velocity along the axle $\mathbf{Y_c}$ of $\mathcal{F}_c$, and ω is the rotational velocity around $\mathbf{Z_c}$. The hand-eye parameters (*i. e.* the rigid transformation between $\mathcal{F}_c$ and the frame attached to the camera) are supposed to be known.

According to Hypothesis 2.2, the state of a set of visual features $\mathcal{P}_i$ is known in the images $\mathcal{I}_i$ and $\mathcal{I}_{i+1}$. Moreover $\mathcal{P}_i$ has been tracked during the learning step along the path ψ between $\mathcal{F}_i$ and $\mathcal{F}_{i+1}$. The state of $\mathcal{P}_i$ is also assumed available in $\mathcal{I}_c$ (*i.e* $\mathcal{P}_i$ is in the camera field of view). The task to achieve is to drive the state of $\mathcal{P}_i$ from its current value to its value in $\mathcal{I}_{i+1}$.

B. Principle

Consider the straight line $\Gamma = (O_{i+1}, \mathbf{Y}_{i+1})$ (see Figure 4). The control strategy consists in guiding $\mathcal{I}_c$ to $\mathcal{I}_{i+1}$ by regulating asymptotically the axle $\mathbf{Y}_c$ on Γ. The control objective is achieved if $\mathbf{Y}_c$ is regulated to Γ before the origin of $\mathcal{F}_c$ reaches the origin of $\mathcal{F}_{i+1}$. This can be done using chained systems. Indeed in this case chained system properties are very interesting. A chained system results from a conversion of a mobile robot non linear model into an almost linear one [16]. As long as the robot longitudinal velocity V is non zero, the performances of path following can be determined in terms of settling distance [17]. The settling distance has to be chosen with respect to robot and perception algorithm performances.

The lateral and angular deviations of $\mathcal{F}_c$ with respect to Γ to regulate can be obtained through partial Euclidean reconstructions as described in the next section.

C. Evaluating Euclidean state

In the sequel, we suppose that $\mathcal{P}_i = \{\mathbf{p}_{ik},\ k = 1 \cdots n\}$ is a set of image points of $\mathcal{P}_{i+1}$. These points are matched with the set of image points $\mathcal{P}_c = \{\mathbf{p}_{ck},\ k = 1 \cdots n\}$ of $\mathcal{I}_c$. Let Π be a 3D reference plane defined by three 3D points whose projections onto the image plane belong to $\mathcal{P}_i$ (and $\mathcal{P}_c$). The plane Π is given by the vector $\pi^T = [\mathbf{n}^*\ d^*]$ in the frame $\mathcal{F}_{i+1}$, where $\mathbf{n}^*$ is the unitary normal of Π in $\mathcal{F}_{i+1}$ and d^* is the distance from Π to the origin of $\mathcal{F}_{i+1}$. It is well known that there is a projective homography matrix $\mathbf{G}$, relating the image points of $\mathcal{P}_i$ and $\mathcal{P}_c$ [7]:

$$\alpha_k \mathbf{p}_{ik} = \mathbf{G}\mathbf{p}_{ck} + \beta_k \mathbf{e}$$

where α_k is a positive scaling factor and β_k is a scaling factor, null if the target point belongs to Π. Given at least four matched points belonging to Π, $\mathbf{G}$ can be estimated by solving a linear system. If the plane Π is defined by 3 points, at least five supplementary points are necessary to estimate the homography matrix [7]. Assuming that the camera calibration $\mathbf{K}$ is known, the Euclidean homography of plane Π is estimated as $\mathbf{H} = \mathbf{K}^{-1}\mathbf{G}\mathbf{K}$ and it can be decomposed into a rotation matrix and a rank 1 matrix:

$$\mathbf{H} = \mathbf{R}^c_{i+1} + \mathbf{t}^c_{i+1} \frac{\mathbf{n}^{*\top}}{d^*} \tag{1}$$

As exposed in [6], it is possible from $\mathbf{H}$ to determine the camera motion parameters, that is $\mathbf{R}^c_{i+1}$ and $\frac{\mathbf{t}^c_{i+1}}{d^*}$. The normal vector $\mathbf{n}^*$ can also be determined, but the results are better if $\mathbf{n}^*$ has been previously well estimated (note that it is the case in indoor navigation with a camera looking at the ceiling for instance). In our case, the mobile robot is supposed to move on a perfect ground plane. Then an estimation of the angular deviation θ between $\mathcal{F}_c$ and $\mathcal{F}_{i+1}$ can be directly extracted from $\mathbf{R}^c_{i+1}$. Furthermore, we can get out from $\frac{\mathbf{t}^c_{i+1}}{d^*}$ the lateral deviation y up to a scale factor between the origin of $\mathcal{F}_c$ and a straight line Γ.
As a consequence, the control problem can be formulated

as following Γ in regulating to zero y and θ before the origin of $\mathcal{F}_c$ reaches the origin of $\mathcal{F}_{i+1}$

D. Control law

Exact linearization of nonlinear models of wheeled mobile robot under the assumption of rolling without slipping is a well known theory, which has already been applied in many vehicle guidance applications, as in [17] for a car-like vehicle, and in our previous works (see [2]). The used state vector of the robot is $\mathbf{X} = \begin{bmatrix} s & y & \theta \end{bmatrix}^\top$, where s is the curvilinear coordinate of a point $\mathbf{M}$, which is the orthogonal projection of the origin of $\mathcal{F}_c$ on Γ. The derivative of this state give the following state space model:

$$\begin{cases} \dot{s} = V\cos\theta \\ \dot{y} = V\sin\theta \\ \dot{\theta} = \omega_c \end{cases} \tag{2}$$

The state space model (2) is converted into a chained system of dimension 3 $\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}^\top$. Deriving this system with respect to a_1 gives an almost linear system. By choosing $a_1 = s$ and $a_2 = y$, and thanks to classical linear automatics, it is then possible to design an asymptotically stable guidance control law, which performances are theoretically independent to the longitudinal velocity V:

$$\omega(y,\theta) = -V\cos^3\theta K_p y - |V\cos^3\theta| K_d \tan\theta \tag{3}$$

K_p and K_d are gains which set the performances of the control law. They must be positive for the control law stability. Their choice determine a settling distance for the control, *i. e.* the impulse response of y with respect to the covered distance by the point $\mathbf{M}$ on Γ. However, as y is estimated up to a scale factor $\frac{1}{d^*}$, this modifies the performances of the control law (3). In practice, to alleviate this difficulty, we choose K_p and K_d for a given d^* that we though maximum. As said before, the control performances are independent to V. Then V can be fixed or tuned by a supervisor. However, V has to be non-null to allow the regulation. Then, in practice, V is controlled in an open loop at the first and the last image of the visual route.

IV. Experiments



Fig. 5. Following a visual route: the previously learnt visual path, about $10m$ long, is materialized on the ground. The pictures were taken during an autonomous run

The proposed framework is implemented on a Pekee$^{\text{TM}}$ robot which is controlled from an external PC. A small 1/3" CMOS camera is embedded on the robot and looks at the ceiling. Ceiling images are used for two main purposes. Firstly, extracted features from an image of the ceiling or from the top of the walls during a visual path learning ought to be recovered in an autonomous run under similar conditions.

Secondly, the ceiling is generally a plane parallel to the ground plane. Then, it is quite easy to give a good approximation of the normal vector $\mathbf{n}^*$ to the reference plane Π in order to evaluate an homography. If the camera frame is confounded with the robot frame $\mathcal{F}_c$, we can assume that $\mathbf{n}^* = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$. The Euclidean homography matrix $\mathbf{H}$ has thus a very simple expression. Indeed, as the displacement between $\mathcal{F}_{i+1}$ and $\mathcal{F}_c$ only consists of one rotation $\theta\mathbf{Z}_c$ and two translations $t_{Xc}\mathbf{X}_c$ and $t_{Yc}\mathbf{Y}_c$, $\mathbf{H}$ is given by (refer to equation (1)):

$$\mathbf{H} = \begin{bmatrix} \cos\theta & -\sin\theta & t_{Xc} \\ \sin\theta & \cos\theta & t_{Yc} \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Therefore, $\mathbf{H}$ has only three degrees of freedom. Only two points lying on Π and matched in $\mathcal{I}_{i+1}$ and $\mathcal{I}_c$ are theoretically necessary to estimate $\mathbf{H}$. The angular deviation θ with respect to Γ can be estimated directly from the computation of $\mathbf{H}$. The lateral deviation y can also be estimated since it is the $\mathbf{X}_{i+1}$ coordinate of the origin of $\mathcal{F}_c$:

$$y = -t_{Xc}\cos\theta - t_{Yc}\sin\theta \tag{5}$$

Thus, from the computation of the Euclidean homography $\mathbf{H}$, the state of $\mathcal{F}_c$ with respect to the straight line Γ can easily be determined.

The Figure 6 illustrates the evolution of planar patterns

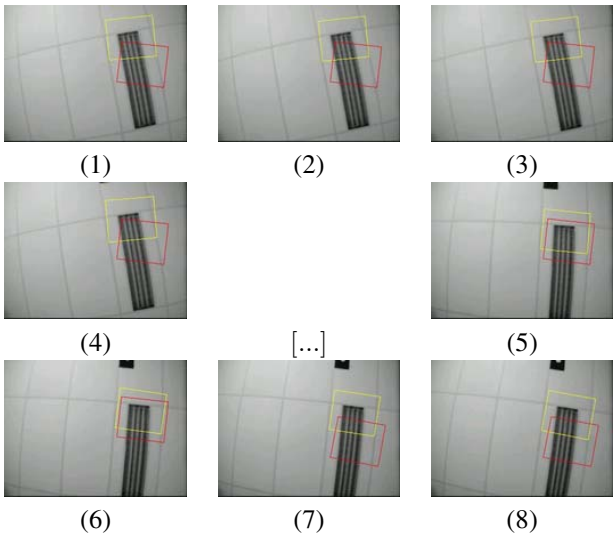


Fig. 6. Evolution of the image space when the robot is regulated between two consecutive key image: in each image, the yellow square is the current state of the tracker, the red one is the state to reach. At image (7), a new reference state is given for the tracker. The image (6) is thus considered close to the previous reference key image: the control has succeeded.

tracked during the robot motion along a given visual route. These tracked planar patterns have been extracted while the user was creating a visual path which is included

into the visual route to be followed. They result from the extraction of a set of interest points, regarding oriented gradient maxima into a rectangular area selected by the user in the current image. An algorithm, based on particle filtering, provides a robust tracking of planar patterns with respect to partial occlusion and to illumination changes. Details on both pattern extraction and tracking can be found in [1].

To create a visual path, the user specifies the beginning of the learning. Then he teleoperates the robot and selects, into the current image, areas which contain possible interesting features. The remainder of the key image selection is then automatic (see [1]). The user indicates when the path to learn has been performed. Then, the new visual path is stored.

At the first step of an autonomous run, the current camera image has to be located into the visual memory. The tracking of learnt planar pattern in this image is then automatic. As a consequence, the user must have chosen at least one reference attitude of the robot which has to be associated with one key image. If the robot has ever achieved a mission since it has been started up, the current image is already supposed to be closed to a key image. At each frame, the tracker provides the coordinates of a current tracked planar pattern. $\mathbf{H}$ is then computed thanks to the knowledge of this pattern in the key image to reach $\mathcal{I}_{i+1}$. A key image is assumed to be reached when a distance between the current points coordinates and the desired one goes under a fixed threshold. The reference path, which is represented on the Figure 5 by the white squares which are lying on the ground, has been acquired as a visual route of fifteen key images. The corresponding path length is about $10m$. The longitudinal velocity V was fixed to $0.2m.s^{-1}$. When the robot stops at the end of the visual route, the final errors in the image corresponds to a positioning error around $5cm$ and an angular error about 3°. Nevertheless, note that the robot has been stopped roughly, by setting V to zero since the last key image of the visual route has been detected. Moreover, both camera intrinsic and hand-eye parameters has been roughly determined. The positioning accuracy depends above all on the threshold which determines if a key image is reached. Our future works will improve that point.

V. Conclusion

This paper presents an original image-based navigation framework dedicated to nonholonomic mobile robots. The approach is illustrated in the context of indoor navigation. We propose to learn the environment as a graph of visual paths, called visual memory. A navigation mission into this visual memory is a visual route. A visual route is made of a sequence of key images of the environment which describes, in the sensor space, an admissible path for the robot. This visual route can be performed thanks to a visual-servoing control law, which is adapted to the robot nonholonomy and does not require any absolute geometrical localization of the robot.

Future works will be devoted to relax the staticity constraint of the environment. We will try to analyse and to take into account environment modifications, which may occur between learning steps and autonomous runs, in both visual route building and following.

Acknowledgment

These works are supported by a grand of the WACIF RNTL (2002-2004) project of the French government.

References

[1] O. Ait Aider, T. Chateau, and J. T. Lapresté. Indoor autonomous navigation using visual memory and pattern tracking. In *British Machine Vision Conference*, volume 2, pages 657–665, Kingston, England, September 2004.

[2] G. Blanc, Y. Mezouar, and P. Martinet. An image driven vision-based control framework for wheeled mobile robots. In *International Symposium on Robotics and Automation*, volume 1, pages 507–512, Querétaro, México, August 2004.

[3] J. Chen, W. E. Dixon, D. M. Dawson, and M. McIntire. Homography-based visual servo tracking control of a wheeled mobile robot. In *Proceeding of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*, pages 1814–1819, Las Vegas, Nevada, October 2003.

[4] G. N. DeSouza and A. C. Kak. Vision for mobile robot navigation: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 24(2), february 2002.

[5] Y. Fang, D.M. Dawson, W.E. Dixon, and M.S. de Queiroz. 2.5d visual servoing of wheeled mobile robots. In *Conference on Decision and Control*, pages 2866–2871, Las Vegas, NV, December 2002.

[6] O. Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. *Int. Journal of Pattern Recognition and Artificial Intelligence*, 2(3):485–508, 1988.

[7] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[8] J.B. Hayet, F. Lerasle, and M. Devy. A Visual Landmark Framework for Indoor Mobile Robot Navigation. In *Proc. Int. Conf. on Robotics and Automation (ICRA'02)*, pages 3942–3947, Washington DC, USA, 2002.

[9] S. D. Jones, C. Andersen, and J. L. Crowley. Appearance based processes for visual navigation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Grenoble, France, September 1997.

[10] Y. Ma, J. Kosecka, and S. S. Sastry. Vision guided navigation for a nonholonomic mobile robot. *IEEE Transactions on Robotics and Automation*, pages 521–37, June 1999.

[11] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Transactions on Robotics and Automation*, 15(2):238–250, April 1999.

[12] Y. Matsumoto, M. Inaba, and H. Inoue. Visual navigation using view-sequenced route representation. In *Proc. of the IEEE International Conference on Robotics and Automation*, volume 1, pages 83–88, Minneapolis, Minnesota, April 1996.

[13] Y. Mezouar, A. Remazeilles, P. Gros, and F. Chaumette. Images interpolation for image-based control under large displacement. In *IEEE Int. Conf. on Robotics and Automation, ICRA'2002*, volume 3, pages 3787–3794, Washington DC, USA, May 2002.

[14] A. Remazeilles, F. Chaumette, and P. Gros. Robot motion control from a visual memory. In *IEEE Int. Conf. on Robotics and Automation, ICRA'04*, New Orleans, April 2004. To Appear.

[15] E. Royer, M. Lhullier, M. Dhome, and T. Chateau. Towards an alternative gps sensor in dense urban environment from visual memory. In *British Machine Vision Conference*, volume 1, pages 197–206, Kingston, England, September 2004.

[16] C. Samson. Control of chained systems. application to path following and time-varying stabilization of mobile robots. *IEEE Transactions on Automatic Control*, 40(1):64–77, 1995.

[17] B. Thuilot, C. Cariou, P. Martinet, and M. Berducat. Automatic guidance of a farm tractor relying on a single cp-dgps. *Autonomous Robots*, 13:53–71, 2002.

[18] D. Tsakiris, P. Rives, and C. Samson. Extending visual servoing techniques to nonholonomic mobile robots. In G. Hager D. Kriegman and A. Morse, editors, *The Confluence of Vision and Control*, volume 237 of *LNCIS*, pages 106–117. Springer Verlag, 1998.