Visual Appearance Mapping for Optimal Vision Based Servoing *

Daniele Fontanelli[1], Paolo Salaris[1], Felipe A.W. Belo[1], and Antonio Bicchi[1]

Interdepartmental Research Center "E. Piaggio", University of Pisa, Italy
(daniele.fontanelli,bicchi)@ing.unipi.it
(felipebelo,salarispaolo)@gmail.com

Pure appearance based visual maps can be constructed without considering any 3–D spatial information, nevertheless they can be used for robot localization and navigation. This paper proposes a topological framework capable of coordinate multi-agent navigation through an appearance based topological map using only grabbed images as sensory data.

The main focus of the paper is on the navigation and parking problems for mobile robots moving in the mapped environment. Henca, a novel image based optimal planner is presented, that allows optimal vehicle trajectories. The planner uses a pure visual-based control with a calibrated visual sensor that takes into-account the field-of-view (FOV) constraint of low-cost monocular cameras in order to keep the tracked features in sight while the robot manoeuvres.

Controller simulations are presented for optimal visual scheme proof. The proposed architecture is then practically tested in an indoor environment, showing the effectiveness and the realization simplicity of the proposed approach.

1 INTRODUCTION

Viable environment interaction is one of the most challenging problem when autonomous robots are foreseen in a real application context. In the mobile robot field, interaction primarily means the agent ability to sense and navigate through its surroundings. As it may not be practical to provide robots in advance with models of the environment, they should be able to create such models navigating and gathering sensorial information of their vicinity. In Simultaneous Localization and Mapping (SLAM) a mobile robot or an autonomous vehicle builds a map of an unknown environment while keeping

* Work supported by E.C. under contracts IST-2004-004536 "RUNES"and IST-2004-511368 "HYCON"

track of its position and localizing itself ([1]). Conventionally, SLAM tracks the localization of a robot while building a map and fusing dead-reckoning information with landmark observations. In literature, the main way of addressing the problem is building a map of both the environment and the robot trajectory in global 3–D coordinates ([2, 3, 4, 5]), that leads to a perception-decision-action loop strongly related to the effectiveness of the localization step.

A quite different approach relies only on the acquired images, mapping the path as a set of key–frames ([6, 7, 8]) or organizing the acquired images in an appearance based graph that generates a topological map. In this paper, the appearance–based topological map has been chosen due to its flexibility (e.g. semantic labeling, as in [9]) and relative low processing cost (e.g. post-processing for detecting convex spaces, as in [10]).

In the so called Visual SLAM literature, different camera apparatuses are proposed, like monocular ([11, 5, 3]), omnidirectional ([7, 8]) and multiple ([12]) cameras. The use of single monocular cameras represents a particularly attractive solution for mobile agents navigation, mainly due to their low cost and their usage simplicity (no exogenous calibration is needed). For example, monocular cameras are suitable for the integration with dissimilar appearance based vision systems, carried by different agents, like heterogeneous robots or human beings, or preexistent in the environment, like surveillance cameras. In the presented architecture applied to a generic indoor environment, like a factory floor or an office, the appearance based map is divided into topologically described environmental zone (for instance, a room or a corridor), each of which is represented by a topological image based map. Each map or sub–map thus defined has a local *map manager* that controls the map sharing. Notice that the map manager can be either locally or physically distributed in the environment. Once a robot enters in an environmental zone it negotiates with the map manager the set of key–frames needed to reach a desired position (described with an image) or it collects new images and then send them back to the map manager. For, a simple *middleware* is needed between each agent and the image map, in order to obtain services (retrieve image maps) and to offer services (build image maps). Each local map manager is (wirelessly) connected to all the other map managers, in order to build the global topological map and to ensure the map connectivity of the overall environment. Given the presented framework, this paper contributes to the practical V-SLAM implementation proposing a monocular vision approach to appearance based topological mapping for heterogeneous robots.

In the rather extensive SLAM literature, exploration strategies represent a field of active research. Nevertheless, this paper does not explicitly focus on this aspect, delegating the exploration phase to a random walk. Instead, the paper is mainly related to the appearance localization and optimal navigation in the topological map. [7] proposes a similar approach using a similarity measure on panoramic images of an image graph map where shortest path strategies are computed in the appearance space. Unfortunately, shortest path in the appearance space increases the probability of correct and accurate localization/navigation but does not necessarily coincides with the shortest path in the metric space. Furthermore, in [7] the control law from one image to the subsequent is not specified. In our opinion, one of the main open problems in literature, nevertheless often neglected, seems to be the control part, i.e. the control law that steers the exploring agents through the map. Assuming an appearance based image map, that in fact allows a quite simple post-processing phase for detecting semantic areas like corridors or rooms, the control law plays a fundamental role since it has to be effective, practical and robust. Nonetheless, it can be solved by also planning optimal trajectory in the robot space. Usually this objective can be accomplished using a 3–D localization estimate. In this paper we solve the problem by planning feature trajectory in the image space, that, however, results in shortest path for the vehicle, and preserves the simplicity of the mapping architecture presented.

Since inexpensive calibrated monocular cameras are considered, the limited FOV constraint must be taken into account. In literature, the problem of keeping the features (e.g. points, edges or snakes) in view during the control task accomplishment has been solved using omnidirectional cameras ([13]), zooming cameras ([14]), pan–tilt heads ([15, 16]) or hybrid position based approaches ([17]). For feature path planning, extracted from a fixed on board camera, [18] recently proposes an optimal path planning for visually guided unicycle–like robots with limited FOV. Once the optimal trajectories are determined, a switched, homography–based, position based visual servoing ([19]) is developed. The solution provided in this paper is quite similar to [18, 19], even though it enlarges the set of optimal trajectories and, moreover, it applies the results of the 3–D robot working space analysis to the image based controller, using feature trajectories rather than robot trajectories. Hence, the camera limited FOV is more naturally considered in the image space (with the vehicle position errors computed directly in the feature space) rather than in the robot working space (where 3–D localization is needed) resulting in a more robust controller (see [20]). The solution is made possible by translating the basic robot manoeuvres into basic feature (point) motions.

It is worthwhile to note that this work is easily extendable to multiple robot's sharing the same map. This is possible because the architecture proposes local map managers that deal with map building, merging, processing and localization and solves loop–closing and kidnapped robot problems.

2 Monocular Appearance Based Topological Map

In the presented approach, no metric information are related to collected images, hence connections between images are obtained by visual appearance. The mapping phase is thus very flexible and completely relies on scene feature richness. However, simplicity in the appearance–based map results in a more difficult navigation phase, that is the central topic of the presented architecture, since localization, planning and control algorithms must be only image related. Therefore, the desired image with desired point feature positions must be correctly associated with the current image point features. Point features detection and association are implemented using *Scale Invariant Feature Transform* descriptors (SIFT, [21]), and are given for granted in the rest of the paper.

2.1 Appearance Based Topological Map

The appearance based topological map is not a novel concept with respect to literature (see, for example, [8, 7]), however, our approach proposes a generalization to low-cost monocular cameras with limited FOV. Briefly, the map is represented as a weighted graph $G = (F, S)$, where F_i is the node i (or the image I_i), and each link $S_{i,j}$ represents a similarity metric between nodes F_i and F_j. Since $S_{i,j}$ is related to the feature descriptor used, it is chosen as

$$S_{i,j} = \frac{\#(F_i \cap F_j)}{\min(n_i, n_j)}, \tag{1}$$

that is the number of SIFT features matched between nodes F_i and F_j normalized by the minimum of $n_i > 0$ (total number of features in image I_i, $\#F_i$) and $n_j > 0$ (total number of features in image I_j, $\#F_j$). The links $S_{i,j}$ may denote how similar are the images I_i and I_j, therefore is a measure of the probability of successful navigation between the two nodes. Notice that $S_{i,j} \in [0, 1]$, where $S_{i,j} = 0$ if there is no feature in common between images I_i and I_j and $S_{i,j} = 1$ if all the features of I_i (I_j) belong to the image I_j (I_i).

2.2 Map Building

The map building process is divided into two different steps. In the exploration phase, the robot randomly explores the environment and collects a sequence of images, sequentially connected as a description of the travelled robot trajectory. To relate the map size to the quantity of new information acquired, we adopt a similarity–based solution: the video stream is analyzed and the actual view is recorded as a new node if and only if its similarity with respect to the last added image is below a threshold. Then, the robot sends the acquired map to the map manager, which analyzes all the acquired images and constructs a connected image graph by enforcing new image links between similar nodes (and detecting possible closed loops) and, in case, deleting the images that are too similar to each other. In the presented architecture, the map building process is divided into two steps:

1. *intra-map* merging: the number of new mapped images are reduced on board the robot before they are transmitted to the map manager;
2. *inter-map* merging: the new entities are (off–line) added to the global map by the map manager.

Notice that this subdivision is not strictly needed by the V-SLAMS architecture, but it saves computational time and bandwidth consumption on the wireless communication.

New recording phases start when the image from the current robot view is not related to any prerecorded image, starting from the last recognized node. New explorations start also when the robot is in an unknown position (i.e. when the localization process does not complete successfully, problem known in literature as *kidnapped robot problem*).

2.3 Localization and Navigation

The similarity measure (1) plays a fundamental role in the architecture: it allows the robot localization on the image map (solving the kidnapped robot problem) by comparing the robot current view w.r.t. the stored map images and allows the computation of a heuristic estimate of the distance between nodes in the appearance space as

$$D_{i,j} = \frac{1 - S_{i,j}}{S_{i,j}}. \tag{2}$$

Indeed, given a controller that is able to stabilize the robot in the desired location using the current view and the desired view, navigation is performed by a planned sequence of images in the topological map. Therefore, as the robot enters in a mapped environment, it asks to the map manager the topological path to follow from its current view to the desired image. The optimal feature path through actual and desired nodes is calculated using the A* algorithm ([22]), this is used to compute the shortest path between two given images (the current image and the desired one). The algorithm demands an admissible heuristic estimate of the distance that was presented in (2), the sequence of selected images corresponds to the most robust path w.r.t. the A* algorithm, i.e. the sequence of images with the highest number of features in common. Notice that the heuristic estimate of the distance (2) is not a physical distance measure, since the triangle inequality is not verified. In fact, taking a sequence of images I_1, I_2 and I_3, it often happens that $\#(F_1 \cap F_2) > 0$ and $\#(F_2 \cap F_3) > 0$ but $\#(F_1 \cap F_3) = 0$ and hence $D_{1,3} = +\infty > D_{1,2} + D_{2,3}$.

It is worthwhile to note that our architecture allows interactions between heterogeneous robots controlled with whatever image–based controller, since only two constraints must be satisfied by the agents: their capability to explore and collect data and their capability to be controlled using only image information. Basically, these two services that allow to interact with the environmental map manager constitutes the components of our V-SLAM “middleware”.

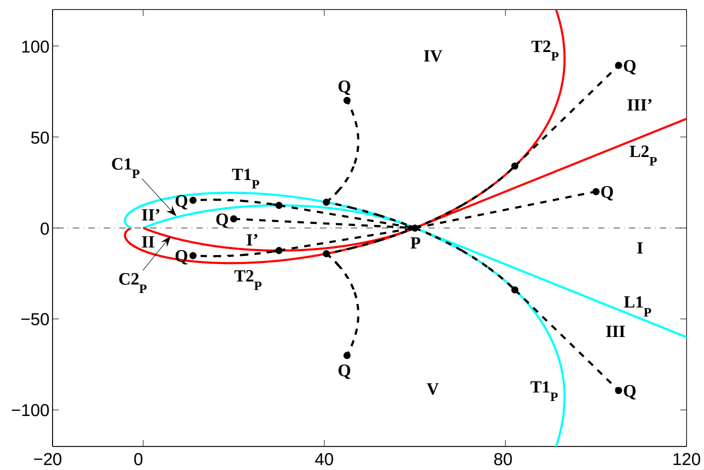


Fig. 1. Shortest paths (according to [18]).

3 Optimal Path Planning and Image Based Control

In this section we propose a set of robot trajectories that are optimal for a unicycle–like vehicle equipped with a rigidly fixed, limited aperture camera. Without loss of generality, consider a symmetric FOV, with characteristic angle ϕ, that is the angle between the camera optical axis and the boundary feature position.

A first solution to this problem has been proposed in [18], where the authors describe the optimal (i.e. shortest) trajectories for unicycle parking (see fig. 1). However, the taxonomy presented in [18] is only locally valid, i.e. near the desired configuration ([18] does not discuss the locality issue). A more accurate robot working space partition is determined using our approach which introduces a new type of optimal path, disregarded in [18], by which it is possible to obtain a global solution.

Using the Pontryagin's Maximum Principle, the shortest paths are described by words of three different kinds of manoeuvres: rotations on the spot, straight line (SL) and, due to the FOV constraint, logarithmic spirals (T–curves) (as in [18]). Using these three basic components, it is possible to split the plane of motion into disjoint regions in which the structure (or word) of shortest paths is invariant for any specified initial point Q.

Since a subset of optimal paths presented in this paper are also reported in [18], the common part description is omitted, referring the reader to the cited paper for completeness. Nevertheless, the fig. 1, that synthesizes the eight regions according to [18], is reported. Our approach comprises a more accurate partition of the *Regions II, II', III, III', IV* and *V*. For, let us consider

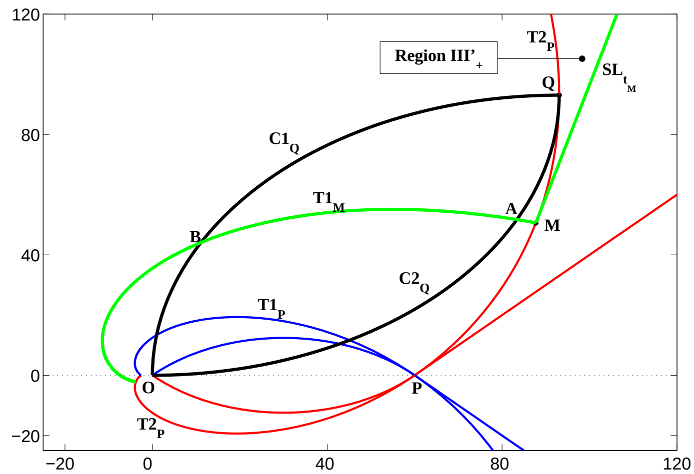


Fig. 2. New subdivision of the motion plane.

the *T Region*. According to [18], the shortest path from Q, initial position of the vehicle belonging to the logarithmic spiral $T2_P$ (or $T1_P$), to the desired position P, is the logarithmic spiral itself.

Consider a generic point Q on the spiral $T2_P$[2] and define a region similar to *Region I'* with desired position Q, bounded by the two circle arcs $C1_Q$ and $C2_Q$ (see figure 2). The points on the arcs can be reached by linear trajectories consistent with the FOV since the features reach the image boundary for points on the arcs. Furthermore, from each point on the arc the forward direction of the vehicle is tangent to some spiral with characteristic angle $\bar{\phi}$. Therefore, the following propositions hold:

Proposition 1. *Any trajectory starting from point Q and reaching subsequently two points Q' and Q'' on the arcs is longer than any direct path from Q to Q''.*

Proposition 2. *For any trajectory starting from a point on $C1_Q$ and going to the desired position P, there exists a point Q' on $C2_Q$ such that the minimal path starting from Q' to P is shorter.*

Proof. Since proposition 1 trivially holds, any minimal trajectory starting from $C1_Q$ follows two logarithmic spirals $T1 * T2_P$. Since this trajectory goes over the feature position and then intersects the symmetric arc of $C2_Q$ in Q'', the proposition holds noting that Q' is the symmetric of Q''.

From the previous propositions, the minimal path from Q to P must intersect the arc $C2_Q$. A particular optimal path from Q over $T2_P$ to P that

[2] For points on the spiral $T1_P$ the analysis is similar.

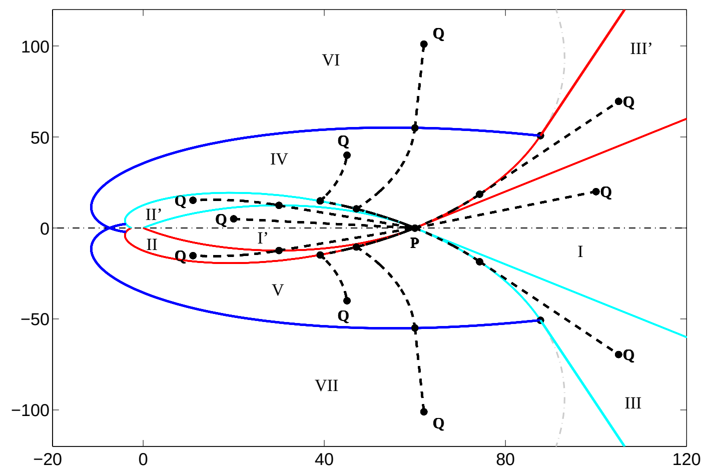


Fig. 3. Shortest Paths and new subdivision of the motion plane.

intersects $C2_Q$ is the logarithmic spiral $T2_P$ itself ([18]). However, if the point Q is on the intersection between the spiral $T2_P$ (or $T1_P$) and the horizontal axis for P with the feature between the starting point and the desired position, the minimal path is the rectilinear path $SL - SL$ over the horizontal axis. Therefore, there exists a point M on $T2_P$ such that starting from points an the spiral between P and M, the minimal path is as in [18], but for further points the minimal trajectories are $SL{-}T2{*}T1_P$ (or $SL{-}T1{*}T2_P$), disregard in [18]. The existence of the point M leads to a new partition of the motion plane, depicted in fig. 3. Notice that only within the proposed analysis it is also possible to analyze degenerated cases, like FOV of π (i.e. infinite–dimension image plane) or 0 radiants (i.e. the image plane is a column of pixels). In the first case the spirals degenerate in circles and the optimal paths are $SL - SL$, whereas in the second case the spirals degenerate in straight line and the optimal paths are $SL * SL$. This new kind of path cannot be explained in light of previously published results since there is no path that can degenerate to $SL - SL$.

The main objective of the proposed solution is to make a controlled parking of a nonholonomic mobile platform equipped with a limited FOV camera using only image information. Furthermore, the trajectories that guide the robot to the desired posture should be optimal, that is the shortest in the 3-D working space of the vehicle. Since the optimal paths are words of an alphabet, whose elements are the rotations on the spot, the straight line or the logarithmic spiral, an image–based control is then feasible once the words are translated from the 3-D world to the image space, i.e. when an equivalent alphabet and semantic rules are defined in the image space as well. Once the alphabet has

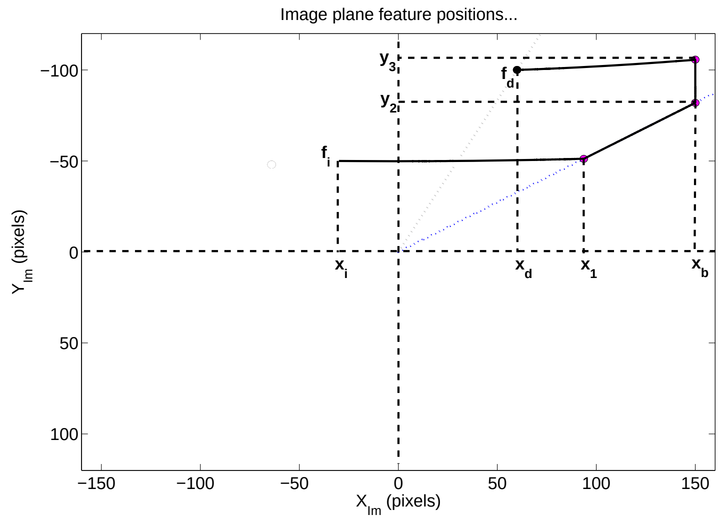


Fig. 4. Feature optimal path.

been defined, the rules to construct the optimal words and the choice of the correct path are given by a PMP problem applied to the current and desired images. More precisely, the rotation on the spot turns into a conic for the features on the image plane, the translation motion turns into a straight line passing through the principal point and the motion on the logarithmic spiral turns into a straight line parallel to image horizontal boundary. Moreover, it is possible to tie at each elementary trajectory executed from the feature on the image plane the variation of θ, i.e. the robot orientation, without using any 3–D information. Therefore, the selection of shortest path follows two principal rules:

1. the path must be entirely contained in the image plane;
2. the variation of θ related to the path must be equal to the angle between the current and desired images.

For example, the optimal path $SL - T1_P$ is reported in fig. 4, where the sequence of a counter–clockwise rotation on the spot (the image conic $x_i \rightarrow x_1$), a forward motion (the straight line $x_1 \rightarrow x_b$), a logarithmic spiral (the straight line $y_2 \rightarrow y_3$) and a clockwise rotation on the spot (the image conic $x_b \rightarrow x_d$) is needed. An experiment is instead depicted in fig. 5, where the red and yellow dots are respectively the feature desired and current positions. Notice that in the final position the feature position error is less than two pixels.

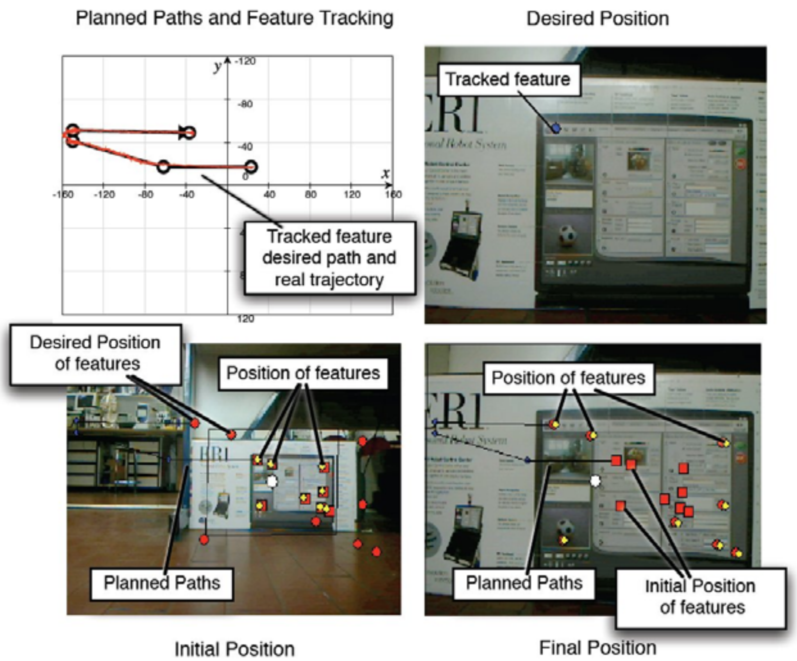


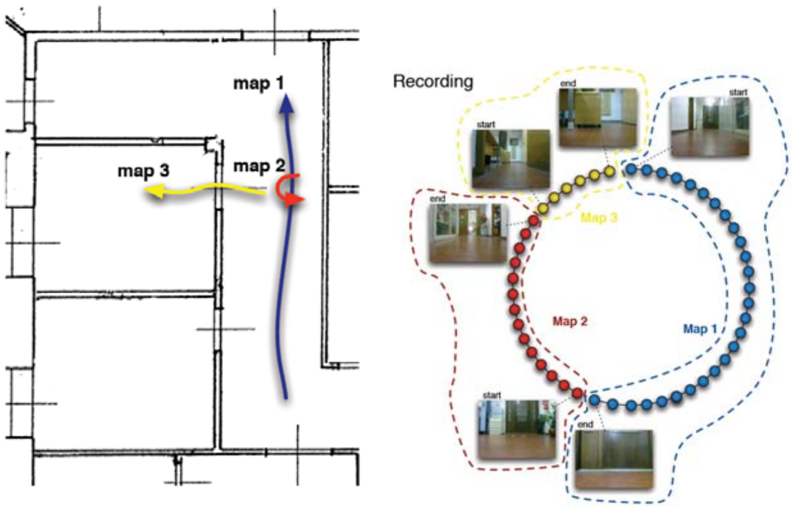
Fig. 5. Planned paths and trajectory of the tracked feature (up left). Initial (bottom left), final (bottom right) and desired (up right) images taken from the vehicle. The initial and desired positions of the features are plotted over the initial and final images taken from the vehicle. The actual positions of the features are also shown in these images.

4 Experimental Results

In this section, experimental results of the overall architecture are presented. Even though the adoption of heterogeneous robots is feasible for the presented solution, the experimental results are related to the same platform, a unicycle–like robot, equipped with a fixed monocular camera.

The experimental setup comprises of a Quickcam Ultravision camera, whose resolution is 320x240 pixels, mounted over the front-part of a K-team Koala vehicle. The ERSP vision library ([23]) is used to perform SIFT recognition. The controller bandwidth is almost 7Hz.

In the experiments, the same robot equipped with a single camera is used to collect data. The robot has followed three different random trajectories starting from three different unknown locations in the unknown indoor environment of the Interdepartmental Research Center "E. Piaggio" (see fig. 6-a). The maps, named *Map I, II* and *III*, are represented as raw collected data in fig. 6-b as three different collection of images, grabbed at different time instants.



a) b)

Fig. 6. Exploration trajectories travelled by the exploring vehicles for the map building process of the Interdepartmental Research Center "E. Piaggio" (a) and Raw sensed data, organized as sequences of the same exploration phase (b).

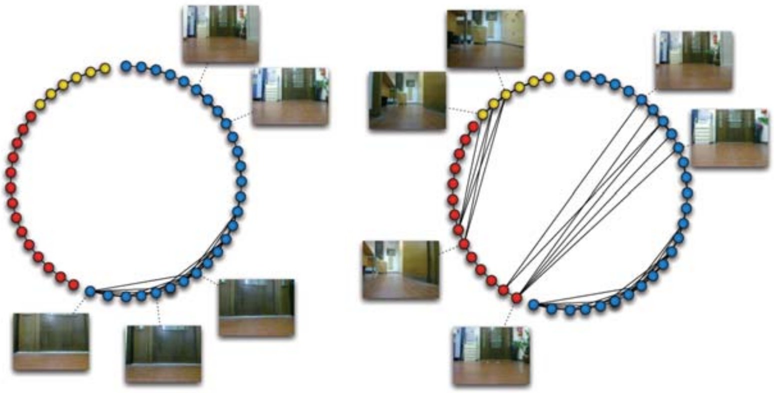
4.1 Map Building

The map building process is essentially the process that organizes and stores only useful information. The usefulness of the information to store is determined by the purpose of the map, in our case safe and accurate navigation through images. The *intra-map* merging is easily explained by means of enforced links between images of the same exploration phase, therefore performed on board the robot (fig. 7-a). Similarly, after each intra–map elaboration, the *inter-map* merging enforces links between images of different exploration phases (fig. 7-b). Recall that this process is performed by the map–manager, basically off-line, using the pre–elaborated maps given by each exploring agent.

4.2 Localization and Navigation

Since we have practically shown in the previous sections how it is possible to build a topological image map and to control an unicycle equipped with a rigidly fixed, monocular, limited FOV camera, the localization and navigation processes are finally reported.

The robot is firstly *kidnapped*, i.e. it is placed in an unknown position in the environment, and a desired image, that belongs to the map, is available (for example, an image of an object that the robot has to reach). In the presented experiment, the desired image is reported in fig. 8-a. The localization in the



a) b)

Fig. 7. The *intra-map* merging (a) and the *inter-map* merging (b).

a) b)

Fig. 8. Desired image to reach at the end of the navigation task (a) and image reached at the end of the navigation phase (b).

map is simply obtained by a rotation on the spot, grabbing the current image from the camera and passing it to the map–manager. If the map–manager does not find a "similar" image, the robot is not localized and, hence, it starts to explore (the same happens if the robot is lost during a navigation phase due to unpredictable faults). Otherwise, the robot is localized in the map and, therefore, the navigation planner, plans an image trajectory, based on similarity quality among images (fig. 9).

The image sequence is a set of ordered images that the robot has to travel through in order to reach the desired position. The visual guided robot motion between two successive images in the image sequence is governed by the optimal feature trajectory planner, showed for example in figure 4. The final experimental results, depicted in fig. 10, report the image sequence used by the robot as well as the map they belong.

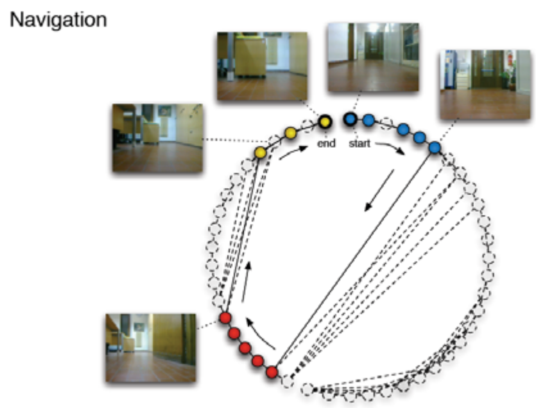


Fig. 9. Planned image trajectory in the map.

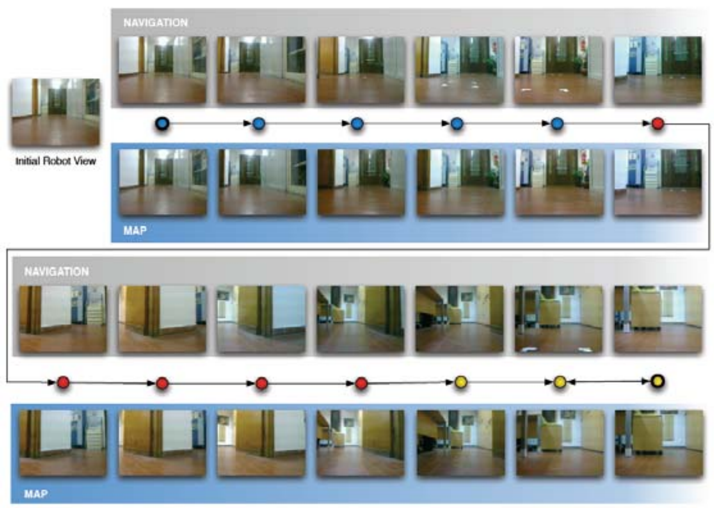


Fig. 10. Sequence of images travelled by the robot during the navigation phase.

Finally, the image reached at the end of the overall navigation task is reported in fig. 8-b, to be compared with the desired one reported in fig. 8-a.

5 Conclusions

In this paper, pure appearance based visual maps are constructed without considering any 3–dimensional spatial data. Although the lack of metric information may lead to robot control challengings, we propose the use of appearance based visual maps to control and navigate multiple, heterogeneous robots. The proposed V-SLAM architecture relies on the idea of a middleware between the map–manager, that can be fixed in the environment or placed on a moving agent, and the robots.

In order to solve the navigation and parking problem of mobile robots, an image based path planner is presented. Optimal trajectories in the robot working space are obtained building words on an alphabet of basic manoeuvres. Then, words are translated in the feature space for image based planning. The proposed solution enlarges the set of previously published optimal trajectories and applies the control law directly in the image space.

The proposed architecture is then practically tested in an indoor, static environment, showing the effectiveness of the proposed approach.

Therefore, the main contributions of this paper in respect to the literature can be summarized as follows: The use of monocular cameras in appearance based topological maps as a novel approach; A decentralized map manager is proposed; Monocular VSLAM systems like the one presented permit the sharing of information between non similar platforms; and, an optimal trajectory control using monocular view is presented for local paths navigation;

References

1. J. Leonard and H. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Trans. on Robotics and Automation*, vol. 7, no. 3, pp. 376–382, June 1991.
2. E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd, "Monocular vision based SLAM for mobile robots," in *Proc. Int. Conf. on Pattern Recognition*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 1027–1031.
3. E. Royer, J. Bom, M. Dhome, B. Thuilot, M. Lhuillier, and F. Marmoiton, "Outdoor autonomous navigation using monocular vision," in *Proc. of the IEEE/RSJ Conference on Intelligent Robots and Systems*, 2005, pp. 1253–1258.
4. A. Danesi, D. Fontanelli, and A. Bicchi, "Visual servoing on image maps," in *Experimental Robotics*, vol. 39. Springer Berlin / Heidelberg, January 2008, pp. 277–286.
5. ——, "Towards cooperative visual-based localization, mapping, and servoing," 2004, proc. IEEE Mediterranean Conf.
6. A. Remazeilles, F. Chaumette, and P. Gros, "Robot motion control from a visual memory," in *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 5, 2004, pp. 4695–4700.
7. H. Andreasson, T. Duckett, and A. Lilienthal, "Mini-SLAM: Minimalistic visual SLAM in large-scale environments based on a new interpretation of image similarit," in *Proc. IEEE Intern. Conf. on Robotics and Automation*, 2007, pp. 4096–4101.
8. O. Booij, B. Terwijn, Z. Zivkovic, and B. Krose, "Navigation using an appearance based topological map," in *Proc. IEEE Intern. Conf. on Robotics and Automation*, 2007, pp. 3927–3932.
9. T. Spexard, S. Li, B. Wrede, J. Fritsch, G. Sagerer, O. Booij, Z. Zivkovic, B. Terwijn, and B. Krose, "BIRON, where are you? Enabling a robot to learn new

places in a real home environment by integrating spoken dialog and visual localization," in *Proc. of the IEEE/RSJ Conference on Intelligent Robots and Systems*, 2006, pp. 934–940.
10. Z. Zivkovic, B. Bakker, and B. Krose, "Hierarchical map building using visual landmarks and geometric constraints," in *Proc. of the IEEE/RSJ Conference on Intelligent Robots and Systems*, 2005, pp. 2480–2485.
11. A. J. Davison, "Real–time simultaneous localization and mapping with a single camera," in *Proc. IEEE Int. Conf. on Computer Vision*, vol. 2, October 2003, pp. 1403–1410.
12. S. Se, D. Lowe, and J. Little, "Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks," *Int. Journal of Robotics Research*, vol. 21, no. 8, pp. 735–758, 2002.
13. H. Hadj-Abdelkader, Y. Mezouar, N. Aldreff, and P. Martinet, "Omnidirectional visual servoing from polar lines," in *Proc. IEEE Int. Conf. on Robotics and Automation*, May 2006, pp. 2385–2390.
14. S. Benhimane and E. Malis, "Vision–based control with respect to planar and non–planar objects using a zooming camera," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Taipei, Taiwan, 14–19 September 2003, pp. 991–996.
15. J. P. Hespanha, "Single camera visual servoing," in *Proc. IEEE Int. Conf. on Decision and Control*, 2000, pp. 2533–2538.
16. D. Tsakiris, P. Rives, and C. Samson, "Applying visual servoing techniques to control nonholonomic mobile robots," in *International Conference on Intelligent Robots and Systems*, September 1997, workshop on "New Trends in Image-based Robot Servoing".
17. P. Murrieri, D. Fontanelli, and A. Bicchi, "A hybrid-control approach to the parking problem of a wheeled vehicle using limited view-angle visual feedback," *Int. Jour. of Robotics Research*, vol. 23, no. 4–5, pp. 437–448, April–May 2004.
18. S. Bhattacharya, R. Murrieta-Cid, and S. Hutchinson, "Optimal paths for landmark–based navigation by differential–drive vehicles with field–of–view constraints," *IEEE Transactions on Robotics*, vol. 23, no. 1, pp. 47–59, February 2007.
19. G. López-Nicolás, S. Bhattacharya, J. Guerrero, C. Sagüés, and S. Hutchinson, "Switched homography-based visual control of differential drive vehicles with field-of-view constraints," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Rome, April 2007, pp. 4238–4244.
20. F. Chaumette and S. Hutchinson, "Visual servo control, Part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, December 2006.
21. S. Se, D. Lowe, and J. Little, "Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks," *International Journal of Robotics Research*, vol. 21, pp. 735–758, 2002.
22. P. Hart, N. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Tran. on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, July 1968.
23. L. Goncalves, E. DiBernardo, D. Benson, M. Svedman, J. Ostrowski, N. Karlsson, and P. Pirjanian, "A visual front–end for simultaneous localization and mapping," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Barcelona, Spain, April 2005, pp. 44–49.