Cooperative Visual SLAMS by Homography

A. Danesi D. Fontanelli P. Murrieri V.G. Scordio A. Bicchi
Interdepartmental Research Center "Enrico Piaggio", University of Pisa
Via Diotisalvi 2
56100 Pisa
E–mail: bicchi,danesi,fontanelli,murrieri,scordio@piaggio.ccii.unipi.it

***Index Terms*— Simultaneous Localization and Map Building, Visual Servoing, Wheeled Vehicles, Collaborative Exploration, Nonholonomic Systems, Feature Based Vision, Homography.**

***Abstract*— The Simultaneous Localization And Map building for Servoing (SLAMS) problem with mobile vehicles is considered. In particular, the map construction of a structured indoor environment using data from infrared sensors and low-cost cameras xed on wheeled vehicles is addressed. Accurate and ef cient ready-to-use map building for human and robot navigation tasks is studied by using more than one vehicles in a collaborative paradigm. Two vehicles are adopted to provide different functionalities for a common explorative task: while one robot moves in the unknown environment, collecting visual data to construct a feature-based map, the other one keeps it in sight using a visual servoing approach and, contextually, extract salient geometric information of the environment using homography techniques. Environmental information collected, from low-cost cameras xed on the wheeled vehicles, are coherently fused online to build the map. In the proposed approach, the two autonomous vehicles closely collaborate, in a sort of "eye-to-hand" paradigm, where the moving robot plays the role of a hand probing the environment, and the second robot acts as the observing eye. This collaborative strategy allows the exploration of large areas, maintaining a low level of uncertainty in localization and mapping processes.**

Experimental results on laboratory vehicles are reported, showing the practicality and effectiveness of the proposed approach.

***Abstract*— In questo lavoro si affrontano le tematiche inerenti la costruzione di mappe e la localizzazione per l'asservimento (SLAMS) di veicoli mobili. In particolare è stato studiato il problema della costruzione della mappa di un ambiente interno strutturato tramite sensori infrarossi e telecamere di tipo economico, montati su veicoli con ruote. Particolare attenzione è stata posta sulla costruzione di mappe che fossero accurate, ef cienti e facilmente utilizzabili per la navigazione sia dagli esseri umani che dai robot, utlizzando un'architettura multi-robot di tipo collaborativo. Diverse funzionalità necessarie per l'esplorazione collaborativa sono espletate dai due robot utilizzati: mentre un robot si muove e raccoglie dati dalla telecamera per costruire una mappa costituita da punti caratteristici, l'altro robot lo mantiene nel proprio cono visivo utilizzando algoritmi di asservimento visuale e, contestualmente, estrae informazioni geometriche sull'ambiente utilizzando tecniche omogra che. Le informazioni, ottenute dai due robot sfruttando le telecamere ssate a bordo, sono coerentemente fuse insieme per costruire in tempo reale la mappa dell'ambiente. Nell'approccio proposto i due veicoli autonomi collaboreranno a stretto contatto in una sorta di paradigma "eye-to-hand", dove il robot in movimento gioca il ruolo della mano che sonda l'ambiente e il secondo assume il compito dell'occhio che osserva. Questa strategia di collaborazione consente di esplorare larghe aree mantenendo un basso livello di incertezza nei processi di localizzazione e di costruzione della mappa.**

Si riportano i risultati sperimentali su alcuni veicoli didattici, dimostrando la praticabilità e l'ef cacia dell'approccio proposto.

Manuscript submitted June 2003. Support from EC contract IST2001–37170 RECSYS, CNR contract C00E714–001, ASI contract I/R/124/02 "TEMA", MURST grants "Mistral" and "Matrics".

I. INTRODUCTION

Wheeled vehicles have a wide range of applications, both in indoor and outdoor environments, and represent one of the areas with larger potential for advanced robotics. A very important trend in research related to mobile robots is concerned with their capability of autonomously navigate in structured or unstructured environment. Typical vehicles tasks, such as path-planning, localization, parking and general reasoning about the working environment depend highly on an accurate and robust representation of the world. Unfortunately, any a priori information about the ambient surrounding the robot and its relative con guration is hardly ever available in many application elds.

Since the 1990s, the problem of map building has been dominated by probabilistic techniques [1]. Since then, the conjunction of the localization and mapping problems has commonly been referred to as SLAM [2], [3], or CML (short for Concurrent Mapping and Localization [4]). Our proposed approach can be regarded as Simultaneous Localization and Map building for Servoing (SLAMS), in which a team of collaborating vehicles explores the environment collecting visual information for map building and acts as a single multi-distributed sensor robot using recently developed visual schemes [5], [6].

One of the most important aspect to be taken into account in a classical SLAM context is the growth of positioning errors as the robots collects observation from the environment. Without outside information, the robot localization rapidly accumulates errors, corrupting the resulting map. This problem can be addressed by using more than one vehicle in a collaborative paradigm. A strong accent has been recently put on distributed robotics systems for coordinated localization, exploration, and mapping [7], [8], which can guarantee greater robustness and efficiency. Our proposed approach is based on a heterogeneous multi-vehicle system in which different kinds of information are used to provide robust vehicles localization, avoiding odometry errors, and consequently increasing the accuracy of obtained maps.

Due to the error growth rate of dead-reckoning, a very important problem to be addressed when localizing mobile robots in the environment is concerned with the necessity of sensorization. Different sensory apparatuses are available for mobile robots, such as ultra-sonic sonar, laser scanners, infrared systems and vision. In the literature the SLAM problem has been solved, among others, by [9] using sonar sensors and by [10] using laser scanners. This paper deals with the problem of using economic, off-the-shelf cameras to enable a cooperative team of autonomous robots to build an accurate map of a structured environment while navigating, without relying on a priori information or artificial landmarks, and representing the collected information in an human intelligible way. The amount of information contained in the collected environment images allows the use of algorithms to extract geometric constraints on the objects in sight. In this paper, homography techniques for plane extraction and feature-based techniques will be addressed for map building.

Another important feature of the SLAM problem is the practical representation of the map, both from a semantic and from an appearance point of view. In the literature, typical maps representation are used in [11] as occupancy grids or, in their 3D evolution, in [12] as digital elevation maps. Resulting maps are closely related to the used sensors, as can be seen in [9] using sonar sensors or in [10] using laser scanner sensors. Furthermore, map description is also correlated to the adopted algorithm to represent the information retrieved. [13] elaborates visual information with complex correlation algorithms off-line to produce a result that can be regarded as a 3D image of the environment. In this paper, we introduce real-time algorithms to produce maps that represents the environment as it is in natural 3D space, easily usable by human or by a navigating robot.

Within the paper, we present an effective solution to the SLAMS problem with low cost fixed cameras onboard the vehicles, by using a combination of previous

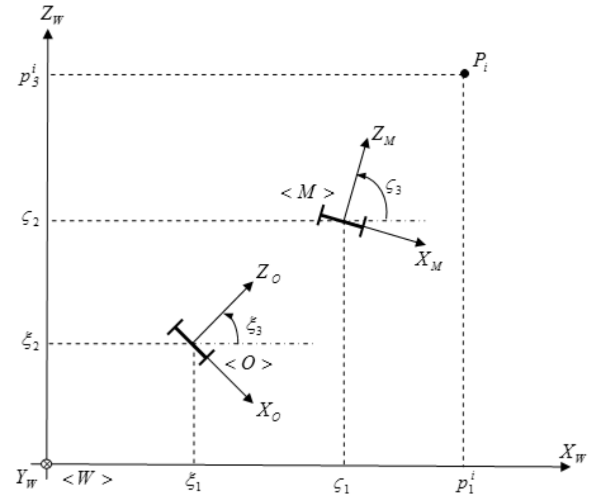

Fig. 1. Fixed frame $< W >$, Observing Robot and Moving Robot camera frames $< O >$ and $< M >$, and relative coordinates ${}^W\xi = {}^W[\xi_1, \xi_2, \xi_3]^T$ for the Observing Robot and ${}^W\zeta = {}^W[\zeta_1, \zeta_2, \zeta_3]^T$ for the Moving Robot.

results on visual servoing control techniques, computer vision and map building techniques. Experimental results on a laboratory vehicles are reported, showing the practicality of the proposed approach.

II. Problem description

Our approach is based on two cooperating vehicles to overcome common SLAM problems, mainly odometry growth errors and self-localization and mapping lack of accuracy during the explorative task.

The cooperative architecture is essentially defined as an eye-to-hand paradigm, detailed in what follows: a Moving Robot (MR) senses the working environment like a hand while the Observing Robot (OR) observes the surrounding ambient like an eye (see fig. 1). This collaborative strategy, joined with the resulting distributed sensing scheme, allows the exploration of large areas, maintaining a low level of uncertainty in localization and mapping processes and adding flexibility and robustness to the whole process.

More precisely, MR explores the environment collecting images used for building feature-based maps online. Such maps are based on features whose coordinates have been obtained with extended Kalman filter techniques, [1], [8]. Standard feature-based maps, used among others by [14], [9], [10], are usually retrieved from laser scanners or sonar sensors. Information contained in visual images of the surrounding ambient are richer of information than simple feature points. However, feature-based maps are sufficient for robot navigation once a visual servoing approach is adopted, as reported in [5]. Regarding cooperative visual localization constraints, MR has been equipped of a particular pattern – a chess board – easily detectable by OR vision system, while for homography plane extraction a wall detection and wall following algorithm are used, based on infrared sensors information.

Based on the known pattern xed on MR, OR provides an accurate localization of the Moving Robot with respect to a xed frame $< W >$ (see g. 1) using visual information [15]. As MR explores the environment by the wall following controller and collects information about the distance of the wall, OR keeps MR in its camera eld of view using a visual servoing control law [5]. As a straight regular wall is detected, the associated texture is extracted by OR using vision homography techniques. It is worthwhile to notice that the exploration strategy adopted by MR is independent from the localization methods used by OR.

To address economicity of applications, realistic assumptions on the nature and quality of the vehicles and of their sensorial equipment have been taken into account. Although the presence of two unicycle-like vehicles may be expensive, it is to be noticed that dynamic errors and inaccuracy in the wheel actuators are compensated by cooperation and sensor measurements. Furthermore, although different sensors (such as some models of laser range nders, or omidirectional cameras, or again pan-tilt heads) may not be affected by conventional cameras accuracy limitations, these are typically some orders of magnitude more expensive than the considered cameras, which are readily available even in the consumer market. In our paper, an accurate camera calibration has been employed to avoid camera distortion [16], in order to use a reliable a pinhole camera model of projection.

III. Cooperative localization

Let's consider two moving camera frames $< M >$ and $< O >$ xed on the mobile robots with the origin in the camera pinhole, with the Z_o axis directed along the camera optical axis and with the Y_o axis perpendicular to the plane of motion and passing through the middle point of the unicycle axles (see g. 1).

Consider now a xed frame $< W >$ whose origin is coincident with the origin of $< O >$ when the Observing Robot is in the initial con guration, and with $X_w = Z_o$ and $Y_w = Y_o$. Let ${}^W\xi = {}^W[\xi_1, \xi_2, \xi_3]^T \in \mathbb{R}^2 \times S$ denote the Observing Robot posture. Similarly, let ${}^W\zeta = {}^W[\zeta_1, \zeta_2, \zeta_3]^T$ denote the Moving Robot posture. More precisely, (ξ_1,ξ_2) and (ζ_1,ζ_2) are the cartesian coordinates of the middle point of the unicycle axles while ξ_3 and ζ_3 are the orientation of the unicycles between the Z_o and Z_m axis and the X_w axis, as represented in gure 1. The relative position between the described frames is represented in gure 2.

From its starting position, OR nds the position of MR's xed chess board pattern using the pattern

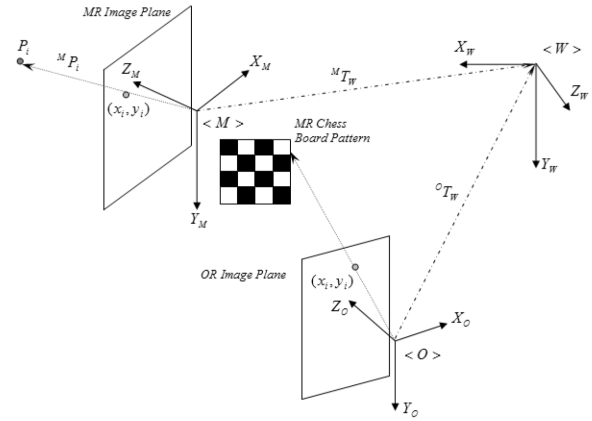

Fig. 2. Fixed frame $< W >$, Observing Robot camera frame $< O >$ and Moving Robot camera frame $< M >$. The coordinates P_i of the i-th feature is expressed with respect to $< M >$. The chess board pattern coordinates are expressed with respect to $< O >$. Notice that, with the cooperative localization, it is possible to estimates the ${}^M T_w$ and ${}^O T_w$ transformation matrixes and therefore expresses all the coordinates with respect to the xed frame $< W >$.

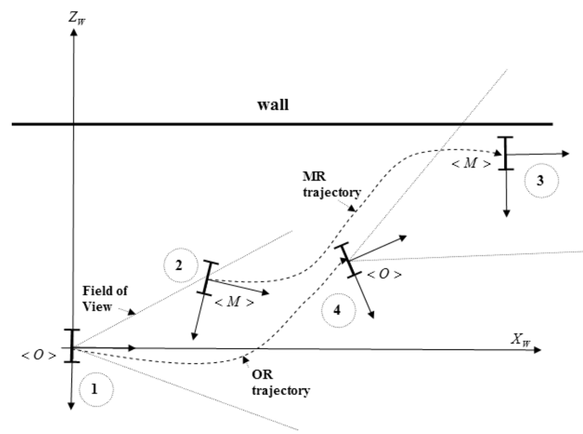

Fig. 3. OR(1) observes and localizes MR(2); MR(2) approaches the wall (3) reaching the border of OR(1) eld of view; OR moves from (1) to (4) by using visual servoing.

recognition algorithm; at this point, it grabs an image of the localized pattern and tells MR to nd and follow the desired wall (position 1 and 2 in gure 3). For simplicity's sake, we assume that the MR is placed at a distance detectable by infrared sensor distance far from the desired wall to map (the assumption will be avoided in future by e.g. using a more ef cient exploration algorithm). The visual servoing controller [5] is started and the Moving Robot is stopped when, during the wall following task, the chess board pattern placed on MR approaches the OR eld of view border by a threshold directly set on the image plane (position 3 in gure 3). The visual servoing controller parks the Observing Robot using the stored desired image as target, with the desired accuracy (position 4 in gure 3). When OR has parked, it grabs a new desired image of the chess board pattern needed for the servoing task and localizes itself with respect the xed frame $< W >$ using MR as a ducial landmark. The global localization of the robots is then obtained from the relative frame positions.

In the literature, the localization problem has been addressed with a cooperative approach among the others by [15], [17].

Within our approach, it is worthwhile to note that the error on the global localization, as the robots move in the unknown structured environment, is directly related to the measurement sensor accuracy, therefore to the camera calibration parameters.

The chess board pattern localization algorithm is explained in the following. From the pattern image, the set of characteristic points (features) are selected with coordinates in the Observing Robot camera frame ${}^{O}P_i = {}^{O}[p_1, p_2, p_3]_i^T$.

The position of each feature in the image plane is described by the perspective projection mapping $\Upsilon : \mathbb{R}^3 \rightarrow \mathbb{R}^2$

$$\Upsilon \,:\, {}^{O}P_i \;\rightarrow\; \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} \alpha_x \frac{{}^{O}p_1^i}{{}^{O}p_3^i} \\ \alpha_y \frac{{}^{O}p_2^i}{{}^{O}p_3^i} \end{bmatrix}. \tag{1}$$

where (x_i, y_i) are the feature coordinates in the image plane (see g. 2). α_x and α_y are camera calibration parameters that represent the focal length and the pixel dimension scale factor on the image.

Assuming that the 3D characteristics of the pattern are known, it is possible to invert equation (1) in a least-squares sense and express the current position ${}^{O}P_i$ from the measured positions of the features in the image plane and therefore the relative position of the two camera frames $< O >$ and $< M >$.

IV. Cooperative Map building

The map is built as a collection of planes with an appropriate texture extracted from the environment. Such texture is extrapolated with homographic techniques by the video stream of the OR which is observing MR. The wall position is known by OR thanks to the information exchanged with MR while it is sensing the wall.

Introducing the homographic techniques, we can here remember that, in homogeneous coordinate notation, the homography is a plane to plane transformation represented by the formulas:

$$\mathbf{H} = \begin{bmatrix} c_{00} & c_{01} & c_{02} \\ c_{10} & c_{11} & c_{12} \\ c_{20} & c_{21} & c_{22} \end{bmatrix} \tag{2}$$

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{H} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{3}$$

Where x and y denote the original pixel coordinates while x' and y' denote the pixel coordinates in the transformed image. In explicit notation the homography between x', y' and x, y is given by

$$\begin{aligned} x' &= \tfrac{c_{00}x + c_{01}y + c_{02}}{c_{20}x + c_{21}y + c_{22}} \\ y' &= \tfrac{c_{10}x + c_{11}y + c_{12}}{c_{20}x + c_{21}y + c_{22}} \end{aligned} \tag{4}$$

In this paper a particular homography is considered: the projection of a generic plane in the environment in the image plane. Notice that non degenerated homographies are invertible, such as the one we considered.

To compute such homography, let R_w, T_w be the matrix and the vector that characterize the af ne transformation of the chess board plane in to the wall plane. Let express a point on the wall with respect to the local wall coordinates. The relation between the point $\mathbf{X} = (X, Y, 0, 1)_{(R_w, T_w)}$ and its projection on the image plane is given by:

$$\begin{aligned} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} &= K(R|T)(R_w|T_w) \begin{pmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{pmatrix}; \\ &= K(R|T)(w_1 w_2 T_w) \begin{pmatrix} X_w \\ Y_w \\ 1 \end{pmatrix} \end{aligned} \tag{5}$$

Where $R_w = (w_1, w_2, w_3)$ and $K(R|T)$ is the matrix of the projection of a generic point in the image plane (typical of the pinhole camera model), K represents the camera intrinsic parameters and $(R|T)$ is the roto-translation between MR and OR. Relation described above is the same of equation 1 in homogeneous coordinates.

The homography matrix $\mathbf{H}$ is then computed as follows:

$$\mathbf{H} = K(R|T)\big(w_1 w_2 T_w\big) \tag{6}$$

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{H} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \tag{7}$$

In order to obtain the wall plane coordinates from the image coordinates it is suf cient to apply the inverse homography $\mathbf{H^{-1}}$.

Such plane transformation is correct if the point p of the image plane belongs to the wall plane (R_w, T_w). In gure 4 an example of such computation is reported.

Once the homography is computed, the texture to be applied to the wall is obtained.

In such textures some other object can appear if they lie between the camera and the wall also if they do not lie on the wall. Obviously, an example of such kind of object is the MR itself.

Fig. 4. Perspective projection inversion example: homography from image plane (top) to chess board plane (bottom). If all the elements of the top image would lie on the chess board plane, the bottom image coincides with a portion of the image that the camera would take when placed with focal axis perpendicular to the chess board. In fact, squares of the chess appear to be exactly squared.

Further work will provide techniques able to extract the elements of the image that really lie on the extracted plane. Such techniques will involve stereo vision from one or more of the Observing robots.

With the homography technique described above, a texture of the wall has been obtained. In order to reconstruct a 3D real time representation of the wall from the texture, an OpenGL-based tool has been implemented.

As the Moving Robot follows the room's walls a feature-based mapping of the environment can be contemporarily performed to produce a more informative map, collecting a video stream from its xed camera mounted as explained in gure 1. 3D feature position estimation is performed using an extended Kalman lter algorithm, widely adopted for real-time execution of SLAM [18], [2]. To overcome extended Kalman lter computational complexity, memory occupation and lack of guaranteed convergence that negatively affect its performance, much work has been done in scienti c community, such as in [8], [14]. Our implementation of the extended Kalman lter is based on the external localization of the robot and on a dynamic dimension of the estimated state space S, due to changes in tracked features number. More precisely, from the starting image of the MR optical ow, a set of n features are autonomously selected. The extended Kalman lter state S is then lled with the coordinates of the Moving Robot ζ, obtained by

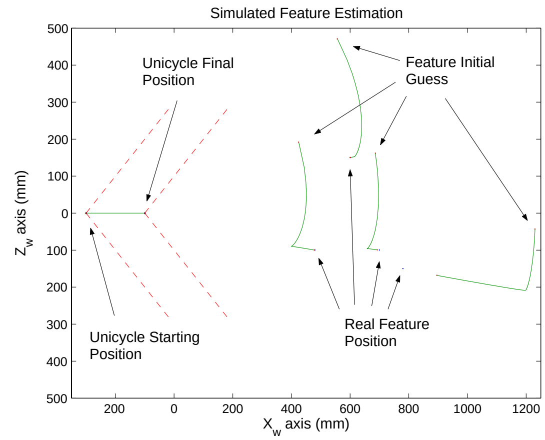

Fig. 5. Initial estimated feature position and related trajectories during the extended Kalman lter estimation. The simulated robot trajectory along the followed wall is also reported.

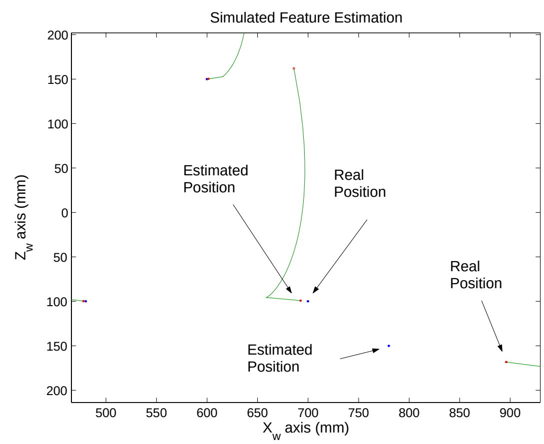

Fig. 6. Final estimated feature position. Notice that the most relevant position inaccuracy is on the Z_m axis, as expected from theory once the robot moves along a straight line towards the feature to estimate.

the external localization, and the 3D coordinates of the selected features referred to the $< W >$ frame: $S = \left[\zeta_1, \zeta_2, \zeta_3, p_1^1, p_2^1, p_3^1, p_1^2, \ldots, p_1^n, p_2^n, p_3^n\right]^T$, where p_j^i stands for the j-th coordinate of the i-th feature. Notice that the dynamical model of the proposed state space $\dot{S}$ is rather simple because the feature are motionless in $< W >$ and the localization of MR is obtained by an external process that is uncorrelated from the whole estimation process. Assuming an accurate external localization, the covariance matrix of the process is therefore a matrix lled with zeros. The covariance matrix of the measurement noise – the image plane coordinates noise of the feature – has been set to an identity matrix, assuming that the noise is gaussian with one pixel of uncertainty.

As the Moving Robot follows the wall, the selected features are tracked in the image by a standard feature

Fig. 7. Experimental set: MR follows the wall and brings the chess board pattern while OR localizes it and builds the map.

Fig. 8. Experimental set: snapshot of the environment to map.

Fig. 9. Experimental results: OpenGl 3D view of the mapped environment. The chess board pattern could be cut off the map without loss of information postprocessing the obtained data or using stereo vision.

tracker algorithm. A filter step is performed once the Moving Robot collects features image plane coordinates contextually with the localization retrieved from the Observing Robot. Initially set to an identity matrix, as the filter proceeds, the covariance matrix of the extended Kalman filter will be partitioned in sub-blocks of dimension 3 by 3, confirming that the estimation processes for each feature are completely uncorrelated.

As the i-th feature approaches the limited field of view of the camera fixed on MR, the state S and the covariance matrixes are coherently updated removing the i-th feature estimated values and fusing them in the map – 3D coordinates with relative confidence.

Simulation results, reported in figure 5 and 6, demonstrate the practicality and convergence of the proposed extended Kalman filter structure. It is worthwhile to note that the MR fixed camera has the relative $< Z_m >$ axis directed along the direction of the linear velocity. As the Moving Robot follows the desired wall, it moves towards the feature to estimate therefore a position inaccuracy along the direction of the motion has been noted, as expected from theory.

Finally, homographic information of OR can be coherently fused in a 3D map with the feature-based map produced by MR.

V. Experimental results

For the experiment a low-cost apparatus was employed, to highlight the robustness and applicability potential of the proposed technique. The experimental setup was comprised of two K-Team Koala vehicles [19], each one equipped with a cheap Kodak EZ200 web-cam [20] placed on the front part of the robot platform. The K-Team Koala vehicle has two symmetric rows of three wheels on its sides, each actuated by a single low-resolution stepper-motor actuator. Such conditions make it hard to use odometry for localization and control, and strongly motivates the use of vision apparatuses for sensing and servoing. The vehicles communicate each other by a wireless connection and an appropriate protocol. The controllers are implemented under Windows XP on two identical 1600MHz Pentium IV laptops mounted on-board of each vehicle. The Intel OpenCV [21] libraries are used for streaming video acquisition and features tracking. The hardware communication between the robots and the related laptops is performed by a RS-232 serial cable.

The robots explore a structured indoor environment (see fig. 7) building a visual map, based on textures of the walls. By comparing the picture reported in figure 8 and the map in figure 9, it is possible to appreciate the precision of the reconstruction: the textures are correctly placed thanks to the good localization of the MR provided by the OR. In figure 10 and 11 the top and global OpenGL 3D views of the map are reported, showing that two perpendicular walls of the room are correctly mapped. It is worthwhile to note that, during

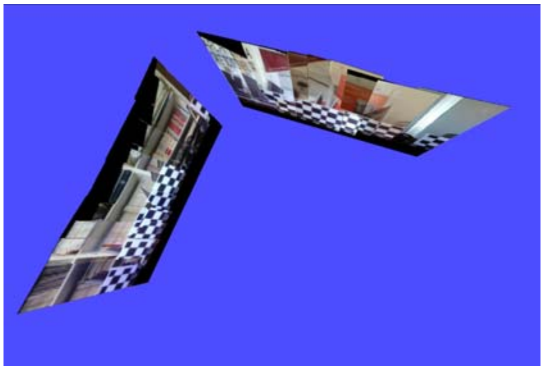

Fig. 10. Experimental results: OpenGl bird's eye 3D view of the mapped environment. The geometric constraint of the wall has been coherently represented in the resulting map.

Fig. 11. Experimental results: perspective OpenGl 3D view of the mapped environment. The image coincides with a portion of the image that a camera would have taken if placed in the environment at the corresponding position.

the manoeuvres to go from one wall to another, MR is not aligned with any wall and therefore OR reconstructs a map with a lacking part in correspondence of the room's angle.

VI. Conclusions

In this paper, a new collaborative control schemes able to solve the SLAM problem, named Simultaneous Localization and Map Building for Servoing (SLAMS), has been proposed. To address economicity of applications, realistic assumptions on the nature and quality of the vehicles and of their sensorial equipment have been considered. Accordingly, the control scheme uses exclusively information from proximity infrared sensors and conventional cameras xed on-board the vehicle, explicitly ignoring odometry data. Two autonomous vehicles closely collaborate to explore the structured environment, realizing an effective distributed sensing architecture. The collaborative strategy allows the exploration of large areas, maintaining a low level of uncertainty in localization and mapping processes and overcoming the sensors inaccuracy. Using visual sensors the adoption of sophisticated computer vision techniques is allowed, towards a robust and accurate solution to the SLAMS problem. Experiments on a low-cost platform has been executed, assessing both the practicality and effectiveness of the proposed approach. The collaborative control scheme joined with the visual servoing control techniques proposed will be extended. For example, distributed sensorial information could be fused, generating an accurate map of structured or unstructured environments.

References

[1] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehnicles*, G. T. Wilfong and I. J. Cox, Eds. Springer Verlag, 1990, vol. 5, pp. 167–193.

[2] J. J. Leonard, H. F. Durrant-White, and I. J. Cox, "Dynamic map building for an autonomous mobile robot," *Int. Jour. of Robotics Research*, vol. 11, no. 4, pp. 286–297, August 1992.

[3] S. Thrun, "Robotic mapping: A survey," in *Exploring Arti cial Intelligence in the New Millenium*, G. Lakemeyer and B. Nebel, Eds. Morgan Kaufmann, 2002.

[4] J. J. Leonard and H. J. S. Feder, "A computationally ef cient method for large-scale concurrent mapping and localization," in *Proceedings IEEE Int. Symposium on Robotics Research*, 2000, pp. 169–176.

[5] P. Murrieri, D. Fontanelli, and A. Bicchi, "Visual-servoed parking with limited view angle," in *Experimental Robotics VIII*, ser. Springer Tracts in Advanced Robotics (STAR), B. Siciliano and P. Dario, Eds. Springer Verlag, 2002, vol. 5, pp. 254–263.

[6] G. Chesi, E. Malis, and R. Cipolla, "Automatic segmentation and matching of planar contours for visual servoing," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 2000, pp. 2753–2758.

[7] S. Thrun, D. Fox, and W. Burgard, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Machine Learning*, vol. 31, pp. 29–53, 1998, also appeared in Autonomous Robots 5, 253–271 (joint issue).

[8] S. I. Roumeliotis and G. Bekey, "Distributed multirobot localization," *IEEE Trans. on Robotics and Automation*, vol. 18, no. 5, pp. 781–795, October 2002.

[9] J. D. Tardós, P. N. J. Neira, and J. J. Leonard, "Robust mapping and localization in indoor environments using sonar data," *Int. Jour. of Robotics Research*, vol. 21, no. 4, pp. 311–330, October 2002.

[10] J. Guivant, E. Nebot, and S. Baiker, "Autonomous navigation and map building using laser range sensors in outdoor applications," *Journal of Robotics Systems*, vol. 17, no. 10, pp. 565–583, October 2000.

[11] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *IEEE Computer*, vol. 22, no. 6, pp. 46–57, June 1989.

[12] I. S. Kweon and T. Kanade, "High resolution terrain map from multiple sensor data," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, vol. 14, no. 2, pp. 278–292, February 1992.

[13] Y. Liu, R. Emery, D. Chakrabarti, W. Burgard, and S. Thrun, "Using EM to learn 3D models with mobile robots," in *Proc. Int. Conf. on Machine Learning (ICML)*, 2001.

[14] J. Leonard, R. Rikoski, P. Newman, and M. Bosse, "Mapping partially observable features from multiple uncertain vantage points," *Int. Jour. of Robotics Research*, vol. 21, no. 11, pp. 943–975, November 2002.

[15] I. Rekleitis, R. Sim, G. Dudek, and E. Milios, "Collaborative exploration for the construction of visual maps," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, November 2001, pp. 1269–1274.

[16] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge University Press, 2000.

[17] R. Kurazume and S. Nagata, "Cooperative positioning with

multiple robots," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 1994, pp. 1250–1257.

[18] J. Borenstein, H. Everett, and L. Feng, *Navigating mobile robots*. A.K. Peters, 1996.

[19] *K-Team Koala Robot*, http://www.k-team.com/robots/koala/index.html.

[20] *Kodak EZ200 Web-Cam*, http://www.photo.net/ezshop/product?product_id=2371.

[21] *Intel Open Source Computer Vision Library*, http://www.intel.com/research/mrl/research/opencv/.