

# Environment Modelling with an Autonomous Mobile Robot for Cultural Heritage Preservation and Remote Access

Grazia Cicirelli and Annalisa Milella  
*Institute of Intelligent Systems for Automation (ISSIA)*  
*National Research Council (CNR)*  
*Italy*

## 1. Introduction

As awareness of cultural heritage raised, much effort was devoted, in the last decade, to improve accessibility and preservation of cultural assets. At present, several methods are available that generally make use of laser scanners and cameras to construct 3D photo-realistic models of different-sized items, ranging from small objects, like statues, up to large buildings and archaeological sites. These methods provide effective technological solutions for cultural heritage preservation, while guaranteeing, at the same time, their accessibility to as much people as possible. Nevertheless, the generation of models may turn into a time-consuming procedure if data acquisition and processing is done by hand, since, in order for models to be sufficiently accurate, an extremely painstaking work is required.

Contributions to implement automated model building techniques and remote access systems for cultural heritage applications may derive from experience in the mobile robotics field. Recently, several research projects have attempted to develop mobile robotic agents in museums (Burgard et al., 1999; Thrun et al., 2000; Trahanias et al., 2005), with different tasks, such as to supply remote access to distant users, accommodate and guide people in the museum, survey those areas where access is not permitted. Equipped with sensors, like cameras and laser rangefinders, mobile platforms provide a variety of viewpoints and may supply the user with dedicated tours of the exhibition and personalized tele-presence, which result in greater interaction capabilities than fixed or even remotely controllable cameras (Trahanias et al., 2005). Generally, in order for a mobile robot to perform its tasks, the knowledge of a map of the environment is needed. Hence, a number of methods for efficient environment modeling, based on information from onboard robot sensors, have been also developed. Yet, relatively little work has been done to extend these techniques to cultural heritage applications, such as modeling of historical and archaeological sites.

In this chapter, we describe our research concerning the development of methods for environment exploration and modelling by a multisensor mobile platform, in the context of cultural heritage access and preservation. Our goal is to have a system able to navigate in the environment and acquire sensorial data, in order to either construct global or local models of the site, or send information to a remote console.

Source: Robotics, Automation and Control, Book edited by: Pavla Pecherková, Miroslav Flidr and Jindřich Duník,  
 ISBN 978-953-7619-18-3, pp. 494, October 2008, I-Tech, Vienna, Austria



Fig. 1. The robotic mobile platform in the configuration used in the polygonal environment.

Usually, when exploring an unknown environment for the first time, a mobile robot is remotely controlled by a joystick or other tele-operation devices. Recently, a novel scenario is receiving considerable attention, which relates to the possibility of teaching the robot its environment by human interaction. This concept, introduced in (Topp & Christensen, 2005), is known as Human Augmented Mapping (HAM). Pursuing this trend, we propose a novel laser-based leg detection and tracking algorithm (Milella et al., 2007) that enables a mobile platform to follow a human user in a tour of the environment, in order to explore the surroundings, acquire sensorial data for map building, and learn particular regions or locations specified by the user.

We present two case studies. The first one deals with the problem of modelling a polygonal environment, such as a museum. The second case study is concerned with the use of a mobile robot for exploration and mapping of a pre-historical underground cave in Southern Italy, named "Grotta dei Cervi", rich in ancient wall paintings of historical and artistic relevance.

The mobile robot employed to carry out this research is shown in Fig.1. It consists of a Pioneer-P3AT by ActivMedia Robotics, equipped with a SICK LMS 200 laser range finder, sixteen forward and rear sonar sensors, encoders, gyroscope, and a monocular pan-tilt-zoom camera. The laser range finder is mounted at a height of about 30cm from the floor and is able to sense objects at a distance of up to 80m with a resolution of 0.5°. Sonar sensors can detect obstacles up to 7m away. Its four tractor wheels can scale a 45° gradient and sills of 9cm. The robot case contains four motors, a local processor, and the batteries. For experimentation in the cave, a 1m high aluminium support was added to the platform to carry the illumination system and the camera (see Fig. 2). Such a configuration allowed us to acquire the wall paintings from an appropriate perspective, since they are mainly located approximately between 1m and 2m above the ground. ARIA C++ libraries by ActivMedia Robotics were used for communication between sensors and the robot controller.



Fig. 2. The mobile robotic platform in the configuration used in the cave.

The results of the experimental sessions realized in both case studies show that the proposed methods are feasible and accurate. They allow us to produce detailed 2D and 3D representations that can be usefully employed to support the study of relevant historical treasures, guaranteeing, at the same time, their safety.

The remainder of this chapter is organized as follows. After a review of the literature related to this work in Section 2, the people following method for Human Augmented Mapping (HAM) is described in Section 3. Section 4 illustrates our approach to polygonal environment modelling. Section 5 describes the study conducted in the pre-historical cave. Finally, Section 6 draws some conclusions.

## 2. Related works

Several works in literature have shown mobile robots to be useful in environment modelling tasks, since, equipped with sensors, such as cameras and laser rangefinders, they can acquire and process sensorial data while navigating in the environment, with minimum human intervention. In (Biber et al., 2004), a 3D modeling method using data from a laser range finder and an omnidirectional camera mounted on a mobile robot is proposed. The method consists of manual, semi-automatic, and automatic parts. Data collection and sensor calibration is carried out manually by teleoperating the robot; wall extraction is done semi-automatically with a user interface; the rest of the processing is fully automatic. In (Nevado et al., 2004), 3D models of the inside of buildings are obtained by using points measured by a laser scanner on-board a mobile robot. The most likely orientation of the surface normal is first calculated at every point by considering also the neighbouring regions to avoid measurement noise. Similar planes are merged and each point is projected onto the planes they belong to. Finally, the contours of the points in the corresponding planes are obtained by a triangulation procedure producing a simple representation of the environment. In this case, the lack of video information limits the final 3D model to a topological representation of the environment without texture. In (Leiva et al., 2001), a 3D model of the environment is

built combining sonar and video information. First, sonar sensors and odometers are used to estimate the distance of the robot from the objects in the environment. Odometric errors are corrected when image data is suitable. Then, a 2D probabilistic map of the environment is divided into segments which become planes in 3D. Finally, the texture from snapshots is assigned to each plane. A rough 3D representation of the environment is obtained, as the aim of the authors is to construct a basic 3D model by an easy and fast method using cheap sensors. Simultaneous Localization and Mapping (SLAM) approaches for concurrent robot localization and modelling of unknown environments can also be found in (Gutmann & Konolige, 1999; Se et al., 2002; Grisetti & Iocchi, 2004; Thrun et al., 2004; Ip & Rad, 2004; Stachniss et al., 2005; Wolf & Sukhatme, 2005).

Despite this high number of model building methods using mobile robots, a few authors have explicitly suggested the use of mobile robotics techniques for applications in the domain of cultural heritage, and realized field tests. Examples can be found in (Allen et al., 2001; Allen et al., 2003; Hirzinger et al., 2005). Specifically, in (Hirzinger et al., 2005), methods for 3D modelling in robotic environments are applied to digitization of cultural heritage from small to large scale objects. In (Allen et al., 2001; Allen et al., 2003), instead, the use of a mobile robot is proposed to build models of urban environments and historic sites.

In our work, we employ a multisensor mobile platform for data acquisition and processing in the context of cultural heritage, and present the results of two practical implementations of the approach, one in a typical indoor polygonal environment and the other in a pre-historical cave. In both cases, the constructed environment model is employed not only for in site navigation of the robot, but also for remote access to cultural assets. For the polygonal environment, our work is related to (Biber et al., 2004); however, we employ a monocular camera and processing is completely automatic. Furthermore, we suggest the use of a Human Augmented Mapping (HAM) (Topp & Christensen, 2005) approach to perform laser and video data acquisition, in place of usual robot teleoperation. First, an accurate 2D map of the environment is generated, based on a SLAM algorithm using laser data (Gutmann & Schlegel, 1996). Starting from this map, the wireframe model of the environment is constructed. Then, images are processed to extract the texture to be added to the wireframe in order to obtain the complete 3D model. For exploration and mapping of the cave, instead, we present an integration of two different algorithms. Specifically, the robot constructs the 2D map of the environment using laser data, and then builds the 3D model of some zones of particular interest using computer vision techniques. The combination of these two approaches makes it possible to obtain a complete knowledge of the environment in an automatic way. The proposed solution allows the access to the cave without damaging it, thus providing an effective system to monitor and preserve its relevant treasures.

### **3. People-following for Human Augmented Mapping (HAM)**

There are several works in literature that concentrate on either people-tracking or following for interaction. They rely on the use of laser range finders or vision or both. (Kleinehagenbrock et al., 2002) integrate vision and laser range data to track a human user, based on a multi-modal anchoring technique. The legs of the user are extracted from laser range data, while his face is from camera images. In (Fod et al., 2002) a laser-based method for real-time tracking of multiple objects is presented, as a step towards the wider objective of identifying people and their activities. Range measurements are grouped into entities for an abstract representation of objects. A Kalman Filter is, then, associated to each object to

address occlusion and sensor noise problems. (Feyrer & Zell, 2000) present a system that detects and pursues people by using both vision and laser data. Color information is used to extract faces in images, whereas convex intervals are extracted in laser scans to detect human legs. In (Pineau et al., 2003) efficient particle filter techniques are used to detect and track people.

In our work, only laser data are employed for detecting people, based on typical human leg shape and motion characteristics. Due to safety reasons, laser range sensors have to be attached near the bottom of the mobile robot; hence, laser information is merely available in a horizontal plane at leg height. In this case, legs constitute, therefore, the only part of the human body that can be used for laser-based people-tracking.

The objective of the proposed approach is to develop a Human Augmented Mapping (HAM) system, that is, to implement a human-robot interaction approach to mapping. The people detection and following method consists of two main modules:

- the **Leg Detection and Tracking (LDT)** module: this module allows the robot to detect and track people using range data based on typical shape and motion characteristics of human legs;
- the **People-Following (PF)** module: this module enables the mobile platform to navigate safely in a real indoor environment while following a human user. During the tour the robot can acquire data for environment mapping tasks.

Details about both modules are provided in the remainder of this section, along with the results of some tests performed in a real context.

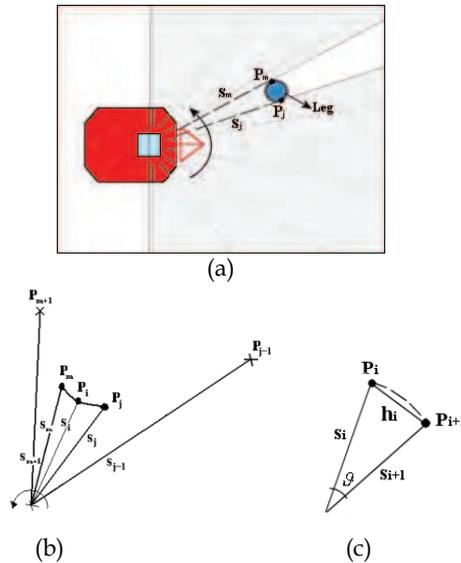


Fig. 3. (a) The robot detecting a leg-shaped object; (b) geometrical representation of a scan interval; (c) geometrical representation of two subsequent scan points.

### 3.1 Leg-detection and tracking

The Leg Detection and Tracking (LDT) method allows to detect and track legs, based on typical human leg shape and motion characteristics. The algorithm starts by acquiring a raw

scan covering a  $180^\circ$  field of view. Laser data are analyzed to look for scan intervals with significant differences in depth at their edges (Feyrer & Zell, 2000). Specifically, let us denote

$$S = [s_1, s_2, \dots, s_k, \dots, s_n] \quad (1)$$

a raw scan, being  $s_k$ , for  $k = 1, 2, \dots, n$ , the laser readings ordered according to the rotation sense of the laser beam (counterclockwise sense in Fig. 3(a)). A scan interval, delimited by scan points  $P_j$  and  $P_m$  (see Fig. 3(b)), i.e.

$$S_{jm} = [s_j, s_i, s_{i+1}, s_{i+2}, \dots, s_m] \quad (2)$$

with  $S_{jm} \in S$ , is selected, if it satisfies the conditions

$$s_{j-1} - s_j > \tau_j \quad (3)$$

$$|s_i - s_{i+1}| \leq \tau_i \quad \text{for } j < i < m \quad (4)$$

$$s_{m+1} - s_m > \tau_m \quad (5)$$

The thresholds  $\tau_i$ , for  $j \leq i \leq m$ , are dynamically computed at each step based on the following considerations. Let us denote  $P_i$  and  $P_{i+1}$  two subsequent point scan readings (see Fig. 3(c)). Since the rotational resolution  $\mathcal{G}$  of the laser beam is small (approximately  $0.5^\circ$ ), it yields

$$h_i \cong s_i \mathcal{G} \quad (6)$$

and

$$\mathcal{G} \cong \sin \mathcal{G} \quad (7)$$

so that we have

$$h_i \cong s_i \sin \mathcal{G} \quad (8)$$

If  $P_i$  and  $P_{i+1}$  belong to a continuous region, then a small difference between the corresponding readings will be observed, which is

$$|s_i - s_{i+1}| \ll h_i \quad (9)$$

Conversely, a discontinuity between  $P_i$  and  $P_{i+1}$  will generate the condition

$$|s_i - s_{i+1}| \gg h_i \quad (10)$$

The value of  $h_i$  can be, therefore, usefully employed to define a threshold  $\tau_i$ , establishing whether  $P_i$  and  $P_{i+1}$  lie on a continuous surface or belong to different objects. Specifically, the threshold  $\tau_i$  can be expressed as

$$\tau_i = \beta \cdot h_i \quad (11)$$

where  $\beta > 1$  is an empirically determined coefficient that takes into account laser noise effects.

Once a set of scan intervals has been selected, a criterion to differentiate between human legs and other similar objects, such as legs of chairs and tables and protruding door frames, must be defined. To achieve this aim, first, the width of each pattern is calculated as the Euclidean distance between its end-points and is compared with the typical diameter of a human leg ( $0.1m$  to  $0.25m$ ). Then, a Region of Interest (ROI) is fixed in the vicinity of each candidate pattern. A leg-shaped region detected within each ROI at the next scan reading is classified as a human leg if the displacement of the pattern relative to its previous position has occurred with a velocity compatible to a typical human leg velocity ( $0.2m/s$  to  $1m/s$ ). Note that if the robot is moving and thus so is the scanner, the effect of ego-motion must first be accounted for. This can be done employing the information provided by the on-board odometers or by the laser scanner.

### 3.2 People-following

The people following algorithm consists of the following steps:

1. detect human legs using the LDT module;
2. choose the closest moving person within a certain distance and angular position relative to the robot;
3. keep track and follow the target person until he/she stops or disappears from the scene.

A control loop is employed, which sets speed and turn rate of the robot, based on the distance from the person and from other objects present in the environment. An obstacle avoidance routine that uses sonar information is also implemented.

As will be shown below, the People Following (PF) module can be effectively employed in the context of Human Augmented Mapping (HAM).

### 3.3 Some tests

In order to test both the performance of the people detection and tracking algorithm and the effectiveness of the people-following method, some tests were performed in our institute. Specifically, three different test scenarios were analysed: 1) robot still, one person present; 2) robot still, two people present; 3) robot following one person for Human Augmented Mapping (HAM). Each experimental setup is discussed in the rest of this subsection.

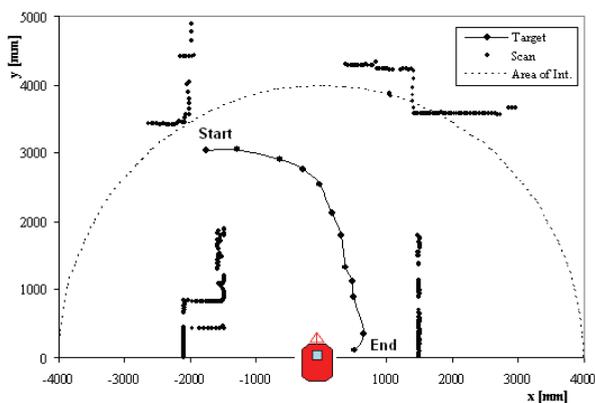


Fig. 4. The trajectory of a person moving in the area surveyed by the robot.

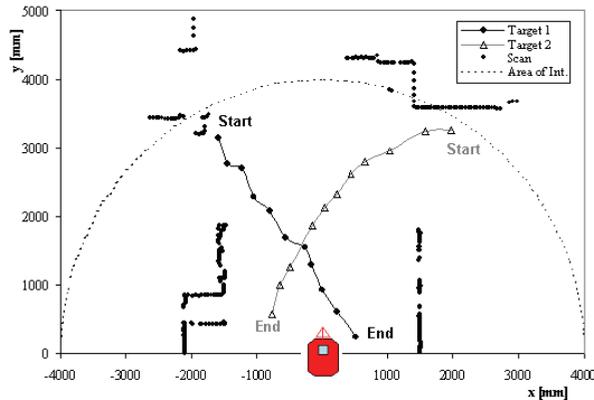


Fig. 5. The trajectories of two people crossing the area surveyed by the robot.

1. **Robot still, one person** - In this case, the robot is not moving. Only one person is present and crosses the field of interest at varying speed. Fig. 4 shows the trajectory of the target within the inspected area (i.e. the semi circumference marked by the dashed line) obtained by the tracking system during one experiment. Black dots represent laser scan readings. The target was classified at all times as a moving person.
2. **Robot still, two people** - Two people cross the scene surveyed by the robot which is not moving. Assuming that the motion direction of each person does not vary significantly, the system is able to keep track of the two trajectories separately, as is shown in Fig. 5.



Fig. 6. Two image sequences taken during the pursuit of a person.

3. **Robot following one person for Human Augmented Mapping (HAM)** - In this case, the robot follows the user in a tour of the environment, maintaining a safety distance from him. The user is identified as the closest person in a range of  $1m$ . Fig. 6 shows two short image sequences taken during the experiments. While following the user, the

robot is able to acquire laser data that allows to concurrently reconstruct its trajectory and build a 2D map of the environment. The simultaneous localization and mapping process is shown in Fig. 7 for two different tests. In this figure, the continuous lines indicate the robot path, while small arcs represent the user's legs and black points the reconstructed map.

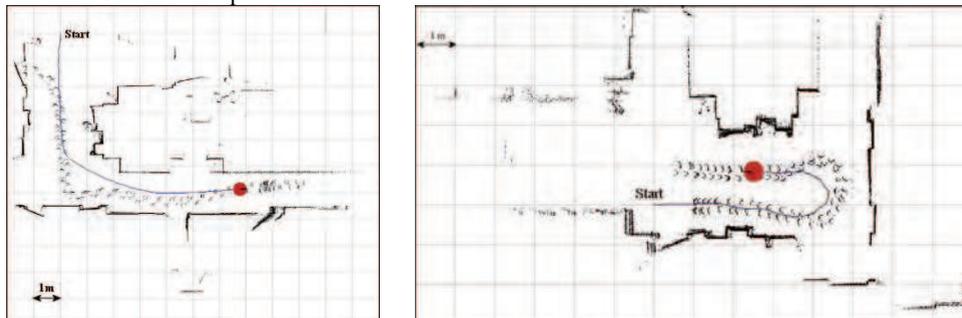


Fig. 7. Two people following experiments for Human Augmented Mapping (HAM).

In the next section, details about the 2D mapping approach will be provided. It will be also shown how, assuming the environment to be polygonal and adding visual information, the 2D map can be used to recover a full 3D model of the environment.

#### 4. Modelling a polygonal environment

In this section, we illustrate our approach to build the 3D model of a polygonal environment, such as a museum, based on laser and video data acquired by a mobile platform. Experiments were carried out in a corridor of our institute. This does not lose generality, since an office environment is very similar to a museum one. First, data acquisition was performed taking the robot on a tour of the environment, and a 2D map was constructed using a laser-based Simultaneous Localization and Mapping (SLAM) algorithm. Then, a 3D wireframe model was generated, adding vertical planes to the contour defined by the planar map. Finally, the 3D model was completed, applying the texture recovered from the acquired images to the wireframe model. Note that the camera was mounted on board the robot above the laser range finder, and its optical axis was oriented  $90^\circ$  to the right with respect to the forward direction of the robot in order to acquire images of the lateral walls.

##### 4.1 2D Map of a polygonal environment

The construction of a map of the environment is a basic step both for robot navigation and accurate knowledge of the environment. The robot can use the map to localize itself as well as to recognize places already explored. This turns the problem of building a map into a problem of on-line Simultaneous Localization and Mapping (SLAM).

In this work, the Combined Scan Matcher (CSM) method proposed in (Gutmann & Schlegel, 1996) is applied to construct a laser-based 2D map and simultaneously estimate the robot trajectory. This approach integrates the IDC algorithm by Lu and Milios (Lu & Milios, 1997) with the method proposed by Cox (Cox, 1991). Fig. 8 shows the point map, with overlaid the robot trajectory, obtained applying the CSM algorithm followed by a statistical filtering for

noise removal. This map is composed of 112 frames scanned by the laser and covering an area of approximately  $20 \times 2m^2$ . The next step is to fit the points with line segments. The result of the line fitting procedure is a map to which a gap removing algorithm has been applied. The final map is shown in Fig. 9. Each couple of consecutive line segments are intersected to obtain the corners in the map. These points will be useful for building the 3D model of the environment.

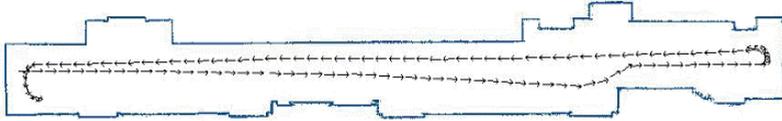


Fig. 8. Planar point map. The path of the vehicle is also drawn.



Fig. 9. Final planar map characterized by line segments.

#### 4.2 Building the 3D Model of a polygonal environment

Once the planar map of the polygonal environment has been obtained, a 3D model can be constructed. The first step is to build the wireframe of the 3D model. The wireframe is generated by adding one vertical plane to each line segment in the 2D map. This operation is possible since the explored environment is polygonal. Each segment in the 2D map refers to the contour of walls, cupboards, doors, and other objects with a polygonal shape. Fig. 10 shows the resulting wireframe.

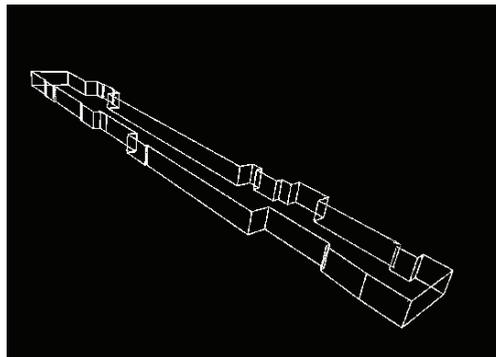


Fig. 10. Wireframe model.

The next step to obtain the complete 3D model is the addition of texture to improve the visual appearance of the model. The texture is obtained by using the images acquired by the camera, on-board the vehicle, during the acquisition phase. Note that images were taken at a

fixed horizontal translation distance from each other and the positions of the vehicle, from which the images were acquired, are known. The process for adding the texture can be divided into the following fundamental steps: 1) generation of the entire image covering the whole contour of the environment, i.e. image registration or mosaicking; 2) correct positioning of the registered image to the wireframe.

1. **Image registration.** The fundamental step in image mosaicking is to find correspondences between different views, and estimate the homographies between the reference image and all other images. Each pair of consecutive images is processed as follows: feature extraction; feature matching; selection of correct matches; final image generation. Specifically, the Harris corner detector is, first, applied to extract corners. Fig. 11 shows the extracted corners on two adjacent images. A simple cross-correlation algorithm is applied to find matches. False matches are then removed using RANSAC (Hartley & Zisserman, 2003). Fig. 12 shows the correct correspondences after applying the RANSAC algorithm.

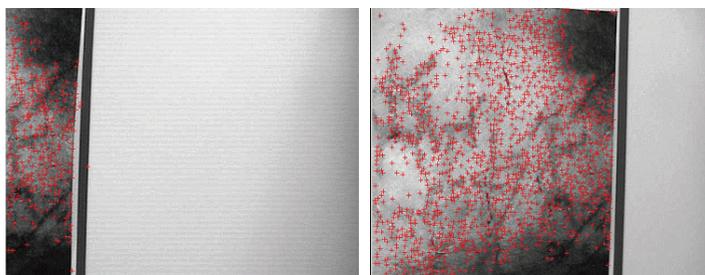


Fig. 11. Corner points extracted by the Harris operator on two adjacent images.

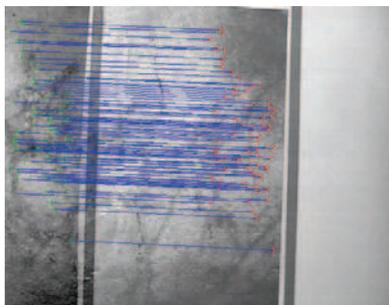


Fig. 12. Correct correspondences of the extracted corners after the application of RANSAC algorithm.

Successively, homographies are estimated by solving overdetermined systems (number of points  $> 4$ ) using the linear-least squares method. By using the homographies, it is possible to estimate the horizontal translation between two consecutive images and then the images can be aligned correctly. The resulting image, of  $38981 \times 480 \text{ pixel}^2$ , is obtained aligning 123 images of  $640 \times 480 \text{ pixel}^2$ . Fig. 13 shows a portion of the registered image.

The described procedure works properly when images present sufficient texture. In our experiment, regions with pictures on the walls have enough texture. Problems arise, instead, in homogenous areas corresponding to walls, doors, etc. In these cases, feature

extraction is difficult and, therefore, homographies cannot be estimated automatically. To solve this problem, the horizontal translations estimated between textured images are used. They are correlated to the known translation distances, performed by the robot during the acquisition phase and estimated by using the odometers. The median of the translations, estimated by using RANSAC and homographies, among textured images is then used to align those images without texture.



Fig. 13. Portion of the mosaic image.

2. **Adding texture to the wireframe.** The final step in building the 3D model is to add the texture, defined by the mosaic image, to the wireframe model. To achieve this aim, the idea is to match correctly the corners of the wireframe model to those in the mosaic image. The corners in the wireframe model are easily detected since they are known on the 2D map. The detection of the corners in the mosaic image needs some additional elaborations. For the sake of simplicity, processing is done on the single images forming the mosaic. First, the Sobel operator is applied to extract edges. Then, the vertical lines in the images are determined by using the Hough Transform (Hough, 1962). Vertical lines can be relative to door edges, picture edges, cupboard edges and so on. Only the ones relative to the corners of the environment are useful for the correct mapping of the texture on the wireframe. The knowledge of the robot positions corrected by using the scan-matching algorithm and the knowledge of the field of view of the camera allow us to find which corner falls in which image. In this way, it is possible to connect each corner to each image acquired during the acquisition phase. The result of the described procedure is the 3D model complete with texture shown in Fig. 14. For a more detailed illustration, Fig. 15 shows some portions of the same model from different points of view. The Virtual Reality Modelling Language (VRML) was used to obtain the model.

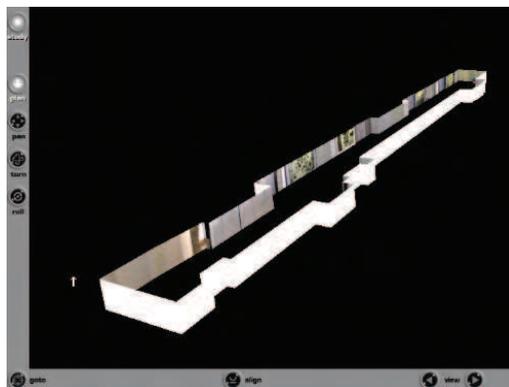


Fig. 14. Textured 3D model.

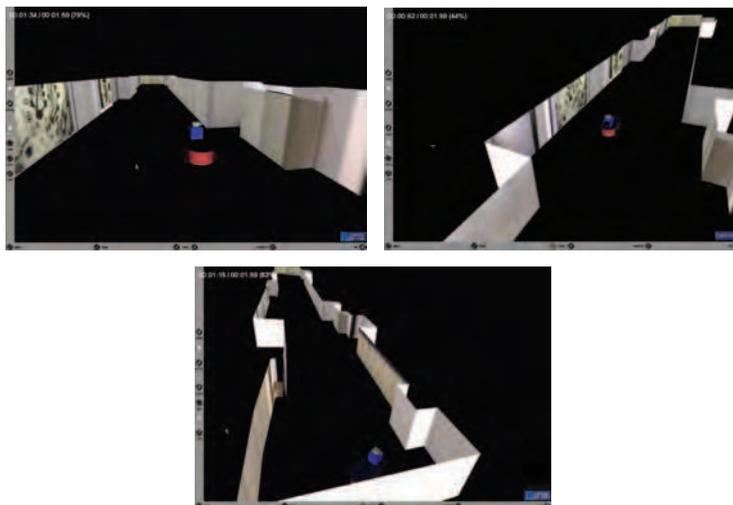


Fig. 15. Portions of the 3D model.

### 5. Reconstruction of a pre-historical cave

In the south of Italy, along the Adriatic coast, a cave, named “Grotta dei Cervi”, holds a pre-historical treasure remarkable for its complexity, and artistic and historical relevance. The cave has on its walls a huge collection of paintings of hunting scenes, stags, men, and small animal groups, realized with red ochre and bat guano, dated to the Middle Neolithic period (see Fig. 16). The access to the cave is restricted to a few authorized people. Care must be taken to guarantee their safety and to prevent polluting elements from being introduced in this particular and valuable environment. The application of a technological solution seems to be the best way to allow remote access to the archaeological site, thus satisfying the need for cave preservation and safety.



Fig. 16. Some paintings present on the walls of the cave.

In this section, we describe the study conducted in the cave, using a mobile robot able to navigate throughout the cave and acquire useful data by means of its on-board sensors. This solution reduces the risk of damaging the cave, as it does not require the installation of invasive infrastructures. The only hand-drawn map of the cave available up to now is illustrated in Fig. 17. The cave is formed by a series of narrow and twisting corridors. The area inspected by the robot is the corridor highlighted in the figure. First, a planar map of the corridor was constructed; then, the 3D model of particularly interesting areas, rich in paintings, was generated. Details for each phase are provided in the rest of this section.

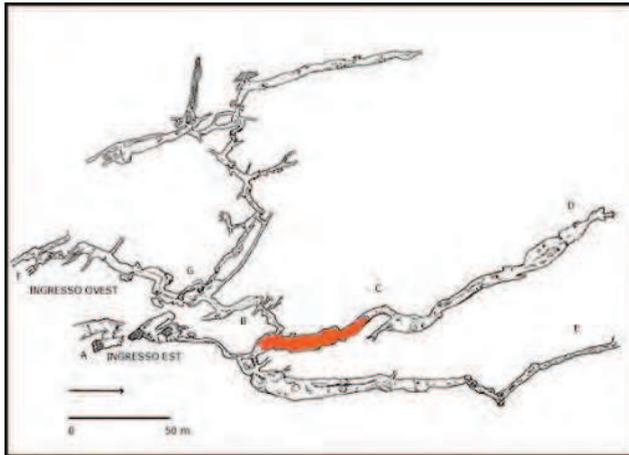


Fig. 17. The handmade map of the cave "Grotta dei Cervi". The highlighted area represents the corridor explored by the mobile robot.

### 5.1 2D map of the cave

Generally, the high accuracy of laser data allows to build accurate planar maps especially when the robot moves in a plane. The cave environment, instead, presents a rough terrain characterized by depressions and bumps (see Fig. 2). In this case, laser data, that supply planar information, must be integrated with data from an inclinometer in order to obtain accurate information about the scanned environment.

Fig. 18(a) shows the robot path and laser data during the scanning procedure. This map is composed of 189 frames scanned by the laser and covering an area of approximately  $15 \times 40 m^2$ . The CSM algorithm was applied to reconstruct both the robot trajectory and the map of the corridor. Fig. 18(b) shows the point map obtained after the application of the CSM algorithm and noise removal. Four zones are also indicated. They are of particular interest for the presence of pre-historical paintings. In order to detect these zones, four artificial, non-invasive landmarks, distinguishable on the map by the laser, have been placed near those areas. Using these landmarks the robot can plan the path to these regions enabling the acquisition of images from the onboard camera of the pre-historical paintings, and the subsequent construction of a 3D model of the observed area.

The robot positions estimated on the map after the CSM application were compared to those provided by the odometers. The estimated errors on the  $x$ ,  $y$ , and  $\theta$  components of the robot pose are plotted in Fig. 19(a), (b), and (c) respectively. The errors on the  $x$ ,  $y$  coordinates are

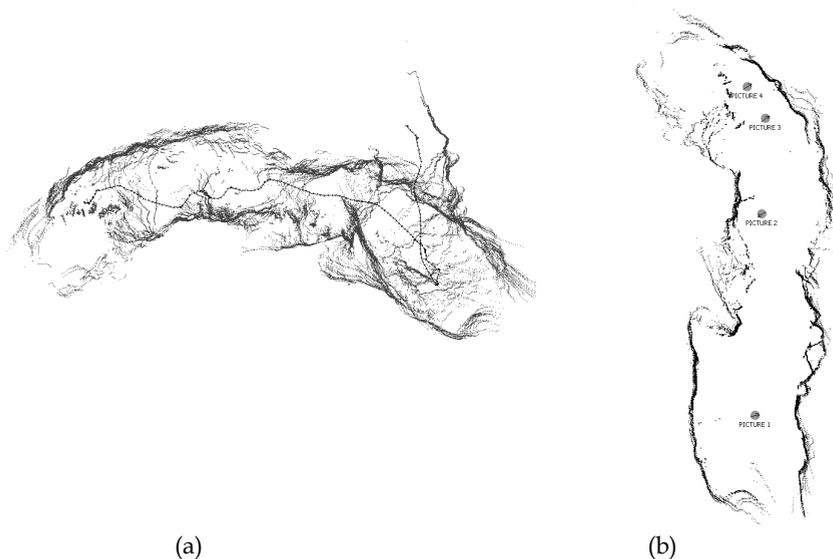


Fig. 18 (a) Laser readings and odometer data after the scanning procedure. The odometer data are represented by little arrows. (b) 2D map of the explored corridor of the cave after the application of the scan-matching algorithm. Four zones rich in interesting paintings are also highlighted.

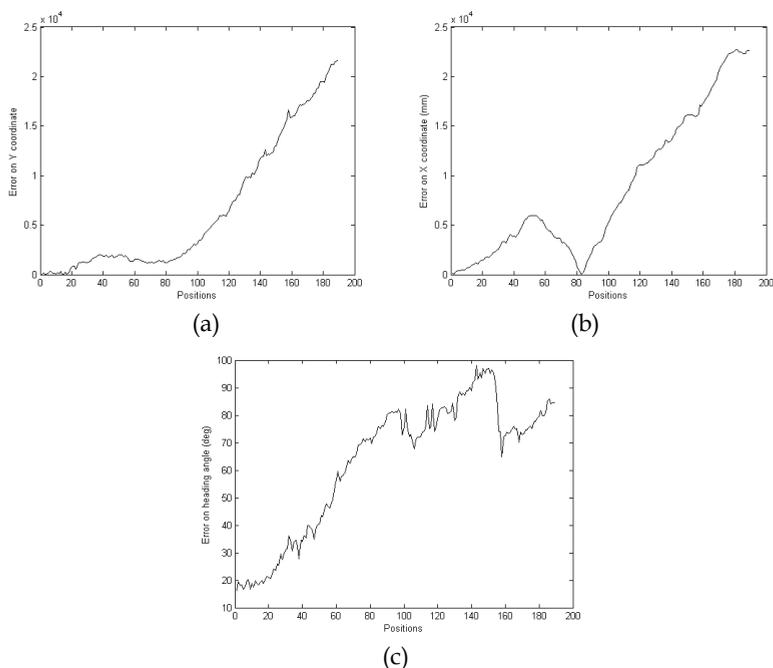


Fig. 19. Estimated errors on (a)  $x$ , (b)  $y$ , and (c)  $\theta$  components of the robot pose.

expressed in millimetres, whereas the error on the robot orientation  $\theta$  is expressed in angle degrees. As it can be noticed, errors are considerable because of the wheel slipping on the rough terrain causing high inaccuracy in odometer position estimates. Nevertheless, the CSM algorithm is able to correct them producing an accurate map. The map obtained using the laser scanner supplies new and useful, although still approximate, information about the structure of the cave: such information was not available before our visit. It is important to note that the structure of the cave, supplied by the map, is very important for the knowledge and the study of the archaeological site, as it describes the morphology of the whole environment placing each painting inside its context and facilitating a better understanding of its role and meaning. Furthermore, the planar map is useful for robot navigation inside the cave.

## 5.2 Building the 3D model of particular areas

Reconstructing 3D models using computer vision techniques generally requires to extract the features (points, lines, target objects) and match them (Aggarwal & Vemuri, 1986; Biber et al., 2004; Gramegna et al., 2005). Moreover, it is important to determine the correspondence between features in different images, since the accuracy of the resulting model depends directly on the accuracy of the feature correspondence. The method described here uses as the only geometrical constraint the correspondence between corners in different images. A complex 3D scene is reconstructed using a set of three images acquired from three different viewpoints of the same scene. The only requirement is that images must be acquired by the same camera with a fixed focal length.

After image acquisition, feature points that correspond to high curvature points (corners), are extracted in each image using the Harris corner detector (Sequeira et al., 1999). The maximum number of corners to be extracted in each image is fixed a priori. A matching procedure is then applied to each couple of images. A classical correlation technique is first used to establish the matching candidates between two images by determining a correlation score for each couple of points. If the correlation score is higher than a given threshold, the related couple of points is considered as a candidate match.

In order to verify the candidate matches, a parameter counts the number of similar candidate matches found in the neighbourhood of each candidate matched point. The sum of these parameters for all candidate matches defines an energy function. The minimization of the energy function through a relaxation technique solves the ambiguity problem (Zhang et al., 1994). After the determination of the corner correspondence for each couple of images, the set of correct matches for all the three images is determined.



Fig. 20. Three images of an area of the wall rich in paintings. White crosses represent the correct point matches.

Fig. 20 shows three images with the matched points. Knowing the corresponding corners it is possible to determine the Fundamental Matrix and the intrinsic parameters of the camera (Hartley, 1997). At this point, all the necessary data to reconstruct the 3D scene are known. The 3D model is reconstructed through the application of the polygonal mesh technique. The 3D model of the scene was made by using the VRML. Fig. 21 shows the 3D model of the scene acquired in the cave.



Fig. 21. 3D model of an area of the cave.

## 6. Conclusions

In order to promote preservation of cultural heritage though guaranteeing accessibility to as much people as possible, novel technological solutions need to be researched. Methods from the mobile robotics field supply effective contributions to the development of environment modelling techniques that can be potentially used to either support the study of historically and artistically relevant assets or provide remote access to museums and archaeological sites, thus satisfying the need for cultural heritage conservation and accessibility.

In this chapter, we presented the results of our research in the field of cultural heritage, concerning the use of a multisensor mobile platform for data acquisition and processing. First, we described a laser-based people-following approach that enables a mobile robot to keep track of and pursue a human user in a tour of the environment. During the tour, the robot can acquire sensorial data to be used for environment modelling. This generates what is usually referred to as Human Augmented Mapping (HAM). Then, we presented the results of two case studies. The first one was related to the problem of constructing a model of a polygonal environment, such as a museum. Data acquisition was performed using a mobile robot, equipped with a 2D laser rangefinder and a CCD camera. Specifically, laser information was employed to simultaneously reconstruct the robot trajectory and build a planar map of the environment. From this map, a wireframe model was recovered. Finally, images were used to generate the texture to be added to the wireframe. Experimental results obtained for tests performed in our institute demonstrated the effectiveness of the proposed methods.

The second case study presented in the chapter focused on the application of a technological solution for remote access and mapping of a pre-historical cave in Southern Italy, named "Grotta dei Cervi". A multisensor mobile robot platform was used to explore the cave and send useful information to a remote console. Based on sensor data, the two-dimensional

map of the site was reconstructed, along with the 3D model of zones of particular interest. Despite the structural complexity of the site, the proposed technological solution proved to be effective for making the archaeological cave available without damaging it. It was shown that the use of such a solution allows the growth of knowledge of this kind of sites, and improves the capability of monitoring and preserving their relevant archaeological treasures.

## 7. References

- Aggarwal, J. K. & Vemuri, B. C. (1986), 3-D model construction from multiple views and intensity data, *Proceedings of IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 435-437, Miami Beach, June 1986, IEEE Computer Society, Washington
- Allen, P., Stamos, I., Gueorguiev, A., Gold, E. & Blaer, P. (2001), AVENUE: Automated Site Modeling in Urban Environments, *Proceedings of IEEE International Conference on 3-D Digital Imaging and Modeling*, pp. 357-364, ISBN 0-7695-0984-3, Québec City, Canada, 28 May-1 June 2001, IEEE Computer Society, Washington
- Allen, P. K., Stamos, I., Troccoli, A., Smith, B., Leordeanu, M. & Hsu, Y. C. (2003), 3D Modeling of Historic Sites using Range and Image Data, *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 145-150, ISBN 0-7803-7736-2, Taipei, Taiwan, 14-19 September 2003, IEEE Computer Society, Washington
- Biber, P.; Andreasson, H., Duckett, T. & Schilling, A. (2004), 3D Modeling of indoor environments by a mobile robot with a laser scanner and panoramic camera, *Proceedings of IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, pp. 3430-3435, ISBN 0-7803-8464-4, Sendai, Japan, September 2004, IEEE Computer Society, Los Alamitos
- Burgard, W.; Cremers, A. B., Fox, D., Hahnel, D., Lakemeyer, G., Schulz, D., Steiner, W. & Thrun, S. (1999), Experiences with an interactive museum tour-guide robot, *Artificial Intelligence*, Vol. 114, No. 1-2, (October 1999), (3-55), ISSN 0004-3702
- Cox, I.J. (1991), Blanche-an experiment in guidance and navigation of an autonomous robot vehicle, *IEEE Transactions on Robotics and Automation*, Vol.7, No. 2, (April 1991), (193-204), ISSN 1042-296
- Feyrer, S. & Zell, A. (2000), Robust Real-Time Pursuit of Persons with a Mobile Robot using Multisensor Fusion, *Proceedings of Int. Conference on Intelligent Autonomous Systems (IAS6)*, pp. 710-715, ISBN 1-58603-078-7, Venice, Italy, July 2000, IOS Press, Amsterdam
- Fod A.; Howard, A. & Mataric, M. (2002), Laser-Based People Tracking, *Proceedings of IEEE Int. Conference on Robotics and Automation*, pp. 3024-3029, ISBN 0-7803-7273-5, Washington, DC, USA, May 2002, IEEE Computer Society, Los Alamitos
- Gramegna, T., Attolico, G., Distante, A. (2005), Different approaches to improve the construction of 3D models using computer vision techniques, *Proceedings of AI\*IA Workshop for Cultural Heritage*, ISBN 88-900910-0-2, Milan, Italy
- Grisetti, G. & Iocchi, L. (2004), Map building in planar and non planar environments, *Proceedings of The 2<sup>nd</sup> International Workshop on Synthetic Simulation and Robotics to Mitigate Earthquake Disaster*, Lisbon, Portugal, June 2004
- Gutmann, J. S. & Schlegel, C. (1996), AMOS: Comparison of scan matching approaches for self-localization in indoor environments, *Proceedings of First Euromicro Workshop on*

- Advanced Mobile Robots*, pp. 61–67, Kaiserlautern, Germany, ISBN 0-8186-7695-7, October 1996, IEEE Computer Society, Los Alamitos
- Gutmann, J. S. & Konolige, K. (1999), Incremental mapping of large cyclic environments, *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pp. 318-325, Monterey, CA, ISBN 0-7803-5806-6, November 1999, IEEE Computer Society, Los Alamitos
- Hartley, R., (1997), Kruppa's equations derived from the fundamental matrix, *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 19, No. 2, (February 1997), (133-135), ISSN 0162-8828
- Hartley, R. & Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*, 2nd edition, Cambridge University Press, ISBN 0521540518
- Hirzinger, G., Bodenmüller, T., Hirschmüller, H., Liu, R., Sepp, W., Suppa, M., Abmayr, T. & Strackebrock, B. (2005), Photo-realistic 3D modelling - From robotics perception towards cultural heritage. *Proceedings of International Workshop on Recording, Modeling and Visualization of Cultural Heritage*, Ascona, Switzerland, May 2005
- Hough, P. V. C. (1962), Method and means for recognizing complex patterns, U.S. Patent 3069654
- Ip Y. L. & Rad A. B. (2004), Incorporation of feature tracking into Simultaneous Localization and Mapping building via sonar data, *Journal of Intelligent and Robotic Systems*, Vol. 39, No. 2, (February 2004), (149-172), ISSN 0921-0296
- Kleinhagenbrock, M.; Lang S., Fritsch J., Lomker F., Fink G. & Sagerer G. (2002), Person tracking with a mobile robot based on multi-modal anchoring, *Proceedings of IEEE Int. Workshop on Robot and Human Interactive Communication*, pp. 423-429, ISBN 0-7803-7545-9, Berlin, Germany, September 2002, IEEE Computer Society, Los Alamitos
- Leiva, J. M.; Martinez, P., Perez, E. J., Urdiales, C. & Sandoval, F. (2001), 3D Reconstruction of static indoor environment by fusion of sonar and video data, *Proceedings of Int. Symposium on Intelligent Robotics Systems*, pp.179-188, ISBN 2-907801-01-5, Toulouse, France, July 2001, LAAS-CNRS, Toulouse
- Lu, F. & Milios, E. (1997), Robot pose estimation in unknown environments by matching 2D range scans, *Journal of Intelligent and Robotic Systems*, Vol. 18, No. 3, (March 1997), (249–275), ISSN 0921-0296
- Milella, A., Dimiccoli, C., Cicirelli, G. & Distante, A. (2007), A., Laser-based people-following for human-augmented mapping of indoor environments, *Proceedings of the 25th IASTED International Multi-Conference: artificial intelligence and applications*, pp. 151-155, ISBN 978-0-88986-631-7, Innsbruck, Austria, February 12-14, 2007, ACTA Press Anaheim, CA, USA
- Nevado, M. M; Garcia-Bermejo, J. G., Casanova, E. Z. (2004), Obtaining 3D models of indoor environments with a mobile robot by estimating local surface directions, *Robotics and Autonomous Systems*, Vol. 48, No. 2-3, (September 2004), (131–143), ISSN 0921-8890
- Pineau J.; Montemerlo, M., Pollack, M., Roy, N. & Thrun, S. (2003), Towards robotic assistants in nursing homes: challenges and results, *Robotics and Autonomous Systems*, Vol. 42, No. 3-4, (March 2003), (271-281), ISSN 0921-8890

- Se, S., Lowe, D. G. & Little, J. (2002), Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks, *International Journal of Robotics Research*, Vol. 21, No. 8, (August 2002), (735-758), ISSN 0278-3649
- Sequeira, V., Ng, K., Wolfart, E., Gonçalves, J. G. M., Hogg, D. C. (1999), Automated Reconstruction of 3D Models from Real Environments, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 54, No. 1, (February 1999), (1-22), ISSN 0924-2716
- Stachniss, C; Hanhel, D., Burgard W. & Grisetti, G. (2005), On actively closing loops in grid-based fast-slam, *Advanced Robotics*, Vol. 19, No. 10, (1059-1079), ISSN 0169-1864
- Thrun, S.; Beetz, M. Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Hahnel, D., Rosenberg, C., Roy, N., Schulte, J. & Schulz, D. (2000), Probabilistic algorithms and the interactive museum tour-guide robot minerva, *The International Journal of Robotics Research*, Vol. 19, No. 11, (November 2000), (972-999), ISSN 0278-3649
- Thrun, S; Liu, Y., Koller D., Ng, A. Y., Ghahramani, Z., Durrant-Whyte, H., (2004), Simultaneous Mapping and Localization with Sparse Extended Information Filters: Theory and Initial Results, *International Journal of Robotics Research*, Vol. 23, No. 7-8 (July-August 2004), (693-716), ISSN 0278-3649
- Topp, E. A. & Christensen, H. I. (2005). Tracking for Following and Passing Persons, *Proceedings of IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*, pp. 2321-2327, ISBN 0-7803-8913-1, Edmonton, Alberta, Canada, August 2005, IEEE Computer Society, Los Alamitos
- Trahanias, P.; Burgard, W., Argyros A., Hahnel, D., Baltzakis, H., Pfaff, P. & Stachniss, C. (2005), TOURBOT and WebFAIR: Web-operated mobile robots for telepresence in populated exhibitions, *IEEE Robotics & Automation Magazine*, Vol. 12, No. 2, (June 2005), (77-89), ISSN 1070-98932
- Wolf, D.F. & Sukhatme, G. S. (2005), Mobile robot simultaneous localization and mapping in dynamic environment", *Autonomous Robots*, Vol. 19, No. 1, (July 2005), (53-65), ISSN 0929-5593
- Zhang, Z., Deriche, R., Faucher, O., Luong, Q. (1994), A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, *Technical report N° 2273*, Institut national de recherche en informatique et en automatique.



## **Robotics Automation and Control**

Edited by Pavla Pecherkova, Miroslav Flidr and Jindrich Dunik

ISBN 978-953-7619-18-3

Hard cover, 494 pages

**Publisher** InTech

**Published online** 01, October, 2008

**Published in print edition** October, 2008

This book was conceived as a gathering place of new ideas from academia, industry, research and practice in the fields of robotics, automation and control. The aim of the book was to point out interactions among various fields of interests in spite of diversity and narrow specializations which prevail in the current research. The common denominator of all included chapters appears to be a synergy of various specializations. This synergy yields deeper understanding of the treated problems. Each new approach applied to a particular problem can enrich and inspire improvements of already established approaches to the problem.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Grazia Cicirelli and Annalisa Milella (2008). Environment Modelling with an Autonomous Mobile Robot for Cultural Heritage Preservation and Remote Access, Robotics Automation and Control, Pavla Pecherkova, Miroslav Flidr and Jindrich Dunik (Ed.), ISBN: 978-953-7619-18-3, InTech, Available from:  
[http://www.intechopen.com/books/robotics\\_automation\\_and\\_control/environment\\_modelling\\_with\\_an\\_autonomous\\_mobile\\_robot\\_for\\_cultural\\_heritage\\_preservation\\_and\\_remote\\_](http://www.intechopen.com/books/robotics_automation_and_control/environment_modelling_with_an_autonomous_mobile_robot_for_cultural_heritage_preservation_and_remote_)

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.