Omnidirectional Distributed Vision for Multi-Robot Mapping

Emanuele Menegatti[1] and Enrico Pagello[1 2]

[1] **Intelligent Autonomous Systems Laboratory**
Department of Informatics and Electronics
University of Padua, Italy
[2] Institute LADSEB of CNR Padua, Italy

Abstract. In this paper, we propose our ideas for realising a Distributed Vision System for building a map of a large environment with a team of mobile robots equipped only with vision sensors. The vision sensors used in this work are catadioptric omnidirectional vision systems. The mirrors of these catadioptric systems have profiles expressly designed for the robot's body and for the robot's task. We report preliminary experiments on the interactions between the vision systems of the different robots.

1 Introduction

The exploration and the map building problems attracted the attention of robotics researchers since the beginning. These problems have been widely studied in the past, but the attention has been mainly focused on single robot exploration. Lately, the attention of some researchers moved to multi-robot systems able to explore an unknown environment in a parallel fashion. This technique is appealing not only because permits to reduce the time needed to cover the whole environment, but also because using a multi-robot system, the redundancy of the observation and of the observers permits to increase the accuracy of the map built by the robots. As an example, refer to the work of Burgard, Thrun *et al.* [12] [2] that are porting on multi-robot systems the expertises and techniques successfully developed on single robot systems like Minerva [13] and Rhino [1]. The sensor used by the robots in these works is a laser range finder. Every robot builds a local map of the environment with the data collected by its own range finder. It uses this local map to pre-process successive scans that are sent to an "*external central mapper*" that create a global map of the environment. The global map is not created from the local maps of the different robot but directly from their preprocessed readings. The cited works focus on the issue of a coordination strategy to maximise the information gain and then minimise the exploration time.

In this paper we want to report our efforts to build a distributed vision system intended for exploration and mapping using mobile robots. In this work we do not focus on the coordination of the robots in order to maximise the information gain or the coverage, we are more interested on the processes of fusion and matching of the visual information gathered from the different robots. We apply a very simple coordination strategy that we called **misrobot**, i.e. the robots choose the exploring direction that increases the distance from the other robots present in their field of view. The exploration is carried on without an explicit strategy, every robot moves according to simple reactive control laws, like wall-following or corridor following. The only sensors used by these robots are the on-board vision systems. The robot team is composed of heterogeneous robots fitted with heterogeneous vision sensor, like standard perspective cameras and catadioptric omnidirectional vision system.

The mapping strategy we use is the same developed for a single robot in the Caboto Project [11]. A single robot equipped with an omnidirectional vision sensor build a map using the Spatial Semantic Hierarchy of Kuipers [6]. The robots then fuse the local maps to generate a global map. In the future we want to be able to use the redundancy of the observation, i.e. the fact that two robots observe the same objects, in order to improve the accuracy of the map.

In Section 2, we describe the mapping strategy used by the single robots and the catadioptric omnidirectional sensors used by the robots. In Section 3, we introduce the idea of Vision Agent, i.e. a vision system with communication capabilities that is embodied in the robot but for some aspects transcends the single robot and we report the two stream of research we are following to develop the prototypes of Distributed Vision Systems. In Section 4, we will present our final target: a community of Vision Agents that realises a Distributed Vision System able to explore and realise a map of an unknown environment. The Distributed Vision System is something more than simply scattering the sensors in the environment.

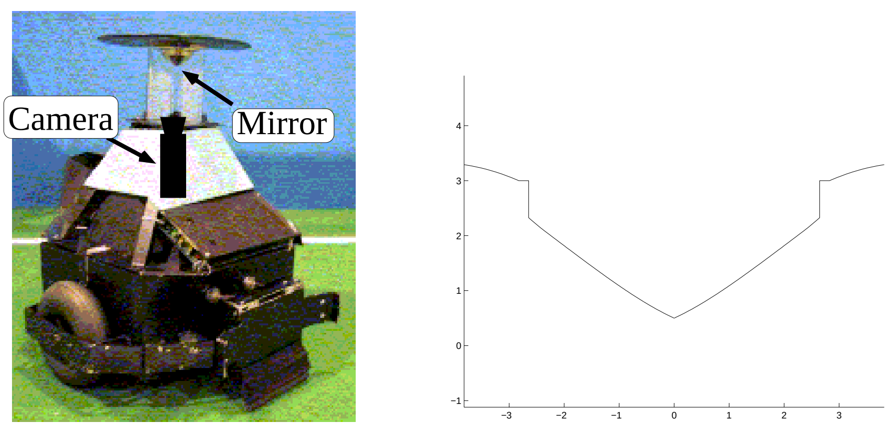


Fig. 1. (a) The robot with its omnidirectional sensor. (b) The actual mirror profile of the new omnidirectional mirror designed for this application.

2 The Caboto project

In the *Caboto Project* we proposed a new approach to the map building task: the fusion of a well understood method, the Spatial Semantic Hierarchy (**SSH**) [5] [6] with a relatively new sensor, an omnidirectional vision system [15]. The Caboto Project was the first project that created a link between the Spatial Semantic Hierarchy and the omnidirectional images. We proposed a set of topological events, that it is possible to identify in the sequence of omnidirectional images acquired while the robot moves. These topological events can be used to pose a discrete set of places that will be the nodes of the topological map [11]. We designed a mirror profile that permits to simply identify topological events while the robots moves [9]. In the next section we will discuss the omnidirectional sensor used, but first let us see the basic structure of the Spatial Semantic Hierarchy (**SSH**).

2.1 Spatial Semantic Hierarchy

The **SSH** (Spatial Semantic Hierarchy) is a model of the knowledge of large-scale spaces of humans, intended to serve as a "method for robot exploration and map building" [5]. The SSH is made up of several layers. Every layer can be implemented independently, even if they strongly interact. Let us see briefly what each layer is about:

- The *Sensory Level* is the interface with the agent's sensory system. It extracts the useful environmental clues from the continuous flux of information it receives from the robots' sensors.
- The *Control Level* describes the world in terms of continuous actions called control laws. A control law is a function which relates the sensory input with the motor output. A control law is retained until a transition of state is detected. A transition of state can be detected with a function called a **distinctiveness measure**.
- The *Causal Level* abstracts a discrete model of the environment from the continuous world. In other words, it is at the causal level that a discrete set of experiences (actions and perceptions) is extracted from the continuous world. These places will be the nodes of the topological map. The discrete model of the environment is composed of **views** , **actions** and the causal relations between them. A **view** is defined as the sensor's reading at a **distinct** place. A **distinct place** is a place where a transition of state is detected. An **action** is defined as the application of a sequence of control laws.
- The *Topological Level* represents an environment with geographical features in the world, such as places, paths and regions connected or contained one in the other.
- The *Metrical Level* augments the topological representation of the environment by including metrical properties such as distance, direction, shape, etc. This may be useful, but is seldom essential.

So far, the SSH has only been implemented either on simulated robots or on real robots with very simple sensors (as sonars). As far as we know, no attempt to use an omnidirectional vision sensor has been made. In the following, we will present the omnidirectional sensor used and we will show why an omnidirectional sensor is a good sensor for building a topological map within the Spatial Semantic Hierarchy frame.

2.2 The Omnidirectional Sensor

In this work we used a mobile robot fitted with an omnidirectional sensor. The robot is depicted in Figure 1(a). Its omnidirectional sensor is composed of a perspective camera pointed upwards at the vertex of a multi-part omnidirectional mirror.

The optical axis of the camera and the geometrical axis of the mirror are aligned. The mirror is supported by a transparent Perspex cylinder. The shape of the mirror is designed in order to maximise the resolution in the regions of interest[1]. This shape permits to exploit in a better way, with respect to the mirror we used in [11], the information is possible to gather from the environment, as we detailed in [9].

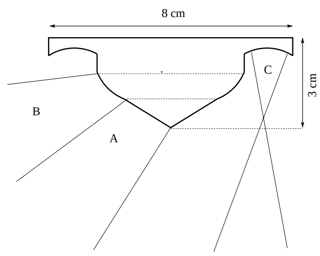
The mirror used is a multi-part mirror, where each segment is designed to view a specific region of space. The design of this mirror was inspired by the work of Marchese and Sorrenti [7]. To understand the mirror profile consider the rough sketch in Fig. 2. The inner part of the mirror (Part A in Fig. 2) is designed to view objects from 60 cm around the robot up to six meters, without displaying the body of the robot . This part produce the main part of the image. The middle ring (Part B) permits to view very distant objects and can be used for a better planning of the exploration movements, using the ideas about the catastrophe theory exposed in [14]. The external ring (Part C) displays at higher resolution (compared to the resolutions attainable in the other two sections) the area closer to the robot. This is useful for the design of control laws like *corridor following* and *wall following*. The actual mirror profile is displayed in Figure 1 (b) . The height of the mirror tip from the floor is 48 cm. The pin hole of the camera is at 32 cm over the floor.

2.3 Omnidirectional vision is good for the Spatial Semantic Hierarchy.

As we demonstrated in [11], the omnidirectional images can be strictly correlated with the **views**[2] introduced in the causal level of the SSH. A **view** is the sensor reading at a **distinct** place, the omnidirectional image is a global

[1] For details on the procedure we used to design the custom profile of the mirror, please refer to[10].

[2] In the following, the bold font is used to indicate we are using the SSH meaning of the words.



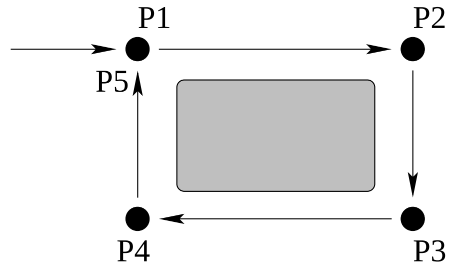


Fig. 2. (a)A rough sketch of the mirror profile where the curvatures of the different sections are exaggerated for sake of clearness. (b) The "exploring around the block" problem. The problem of recognizing the same place under different state labels.

reading of the surrounding at a certain place. Associating **view**s with omnidirectional images simplifies the data interpretation and then the map building process.

For instance, the rotational invariance and the link between **views** and **actions** permit a straightforward solution to the problem of *exploring around the block*, i.e. the problem of recognising the same place under different state labels, see Figure 2. Here the robot is moving around the block following the arrows. When it reaches Place 5 from Place 4, it is very difficult to recognize it as the previously visited Place 1, unless it is equipped with an omnidirectional camera and it makes use of the rotational invariance. Using the SSH terminology, it is easy to spot whether the current **view** is the same which has been experienced before and therefore to consider this view not as a different **place** but as the same **place** reached from a different direction.

In the next sections, we will outline how this mapping technique can be scaled to a team of heterogeneous robots performing a mapping task with their vision systems as only sensors.

3 The Vision Agents

First of all, let us precise a terminology issue: in the following we will prefer the term *Vision Agent* instead of "vision system". The term Vision Agent emphasizes that the vision system is not just one of the several sensors of a single robot, but that it interacts with the other vision systems to create an intelligent distributed system, as in the definition of Ishiguro [4] that is one of the most active researcher in distributed vision. At the moment, we are working on two research streams: interaction of two VAs mounted on the same robot, coordination of several VAs mounted on different robots.

3.1 Two VAs mounted on the same robot

The first implemental step is to realise a Cooperative behavior between two heterogeneous vision agents embodied in the same robot. We want to create a Cooperative Vision System using an omnidirectional and a perspective vision system mounted on the same robot. The omnidirectional vision sensor is a catadioptric system composed of a standard colour camera pointing up-ward and a customly designed omnidirectional mirror on top of it, see Fig. 3 (a).

The omnidirectional camera is mounted on the top of the robot and offers a complete view of the surroundings of the robot, but with a low resolution. The perspective camera is mounted in the front of the robot and offers a more accurate view of objects in front of it. These two cameras mimic the relationship between the peripheral vision and the foveal vision in humans. The peripheral vision gives a general, and less accurate, information on what is going on around the observer. The foveal vision determines the focus of attention and provides more accurate information on a narrow field of view. So, the omnidirectional vision is used to monitors the surroundings of the robot to detect the occurrence of particular events. Once one of these events occurs, the Omnidirectional Vision Agent (OVA) send a message to the Perspective Vision Agent (PVA). If the PVA is not already focused on a task, it will move the robot in order to put the event in the field of view of the perspective camera. This approach was suggested by our previous researches presented in [3].

Experiments on such a system are running and they will provide more insight on the cooperation of the two heterogeneous vision agents.

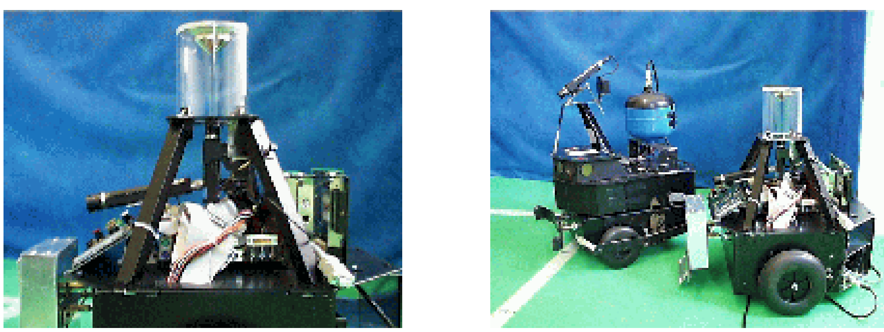


Fig. 3. (a) A close view of the vision system of Nelson. On the left, the perspective camera. In the middle, pointed up-ward the omnidirectional camera. (b) A close view of two of our robots. Note the different vision systems.

3.2 Coordination of several VAs mounted on different robots

The second stream of research is the creation of a Cooperative Distributed Vision System for a multi-robot team. The practical implementation we decided to tackle is the realisation of the Cooperative Object Tracking Protocol proposed by Matsuyama. In the experiments presented in [8], Matsuyama used active cameras mounted on a special tripod. The active cameras were pan-tilt-zoom cameras modified in order to have a fix view point. This allowed the use of a simple vision algorithm, not very different from the case of static cameras. We want to introduce mobile robots and robot vision algorithms in such a system and realise a cooperative distributed tracking of the ball within our team of football player robots.

In the work of Matsuyama the central notion is the concept of *agency*. An **agency**, in the definition of Matsuyama, is *the group of the VAs that see the objects to be tracked and keeps an history of the tracking*. This group is neither fixed nor static. In fact, a VA exits the agency, if it is not able to see the tracked object anymore and a new VA can joint the agency as soon as the tracked object comes in its field of view. To reflect the dynamics of the agency we need a dynamic data structure with a dynamic role assignment.

Let us sketch how the agency works, using an example draw from our application field: the RoboCup domain. Suppose to have a team of robots in the field of play. Each robot is fitted with a Vision Agent. None of the Vision Agents is seeing the ball. In such a situation no agency exists. As soon as a Vision Agent see the ball, it creates the agency sending a broadcast message to inform the other Vision Agents the agency has been created and it is the master of the agency. After this message a second message follows, telling the other Vision Agents the estimated position of the ball. The other Vision Agents maneuver the robots in order to see the ball. Once a Vision Agent has the ball in its field of view, it asks permission to joint the agency and send to the master its estimation of the ball position. If this information is compatible with the information of the master, i.e. if the new Vision Agent has seen the *correct ball*, it is allowed to joint the agency. Also this experiment is in progress and it is a preliminary testbed for the realisation of the more complex system described in the next section.

4 The Distributed Vision System

Our main target is the realisation of a Distributed Vision System mounted on a team of mobile robots able to explore and map an unknown environment.

The only sensors mounted on the fleet of robots are vision systems. Some robots mounts only omnidirectional vision sensor while other mounts omnidirectional and perspective vision sensor. The vision sensor are part of the previously introduced Vision Agents that in addition to the image processing capabilities have also some communication capabilities, i.e. they can send and receive messages over the network.

The basic idea is that every robot perform a local map of the environment while it moves through it. When two of the robots sees each other, they perform a matching on the overlapping regions of the two local maps and if this matching is successful they merge their maps, so every robot has a bigger map composed of two local maps.

The first experiment we are settling is the cooperative exploration and mapping of a simple environment with three robots mounting omnidirectional vision agents. Consider Fig. 4 (a), the robots start from a common location. Every robot moves using one of the designed control laws (wall-following, corridor-following) and moves in a direction chosen in order to increase its distance from the other robots[3]. While the robots move they build a local map of the surrounding using the vertical edges of the environment as described in Section 2. The local map of every robot is represented in Fig. 4 (b) by the lines beside the objects boundary. The robots are able to detect the teammates and to recognise them thanks to a markers placed on top of the robots. If a robot spots in its field of view one or more teammates, it creates a link with them and they start a matching process on the local features they see in the omnidirectional image. The matching/merging process is repeated every time the robot spots a new vertical edge in the environment, in order not to merge maps that contains non-new information again and again. To simplify the matching process we decided to use the colours in the image. We built the maze with boxes with a different colour for every side, so if the colours and the dimensions of the boxes match, it is possible to create a correspondence between the two local maps built by each robot and merge them in a single map. At the end of the merging process, every robot has the single map.

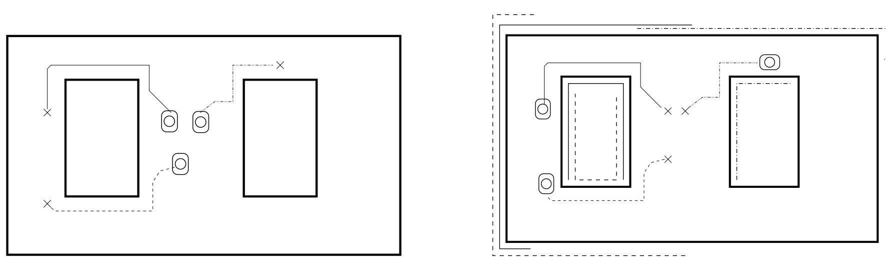


Fig. 4. (a) The robots at the starting point with the path each robot will follow. (b) The robots during the exploration and for each robot the local map it built. Note the two robots on the left, they can see each other, so they will perform a matching on the features they can see and merge their local maps.

This project is at a preliminary state. At the moment we did not implemented a system to automatically stop the exploration when the map is complete and we are not concerned on the fact that every robot has a complete map of the environment. In the end every robot will have a map that

[3] As we stated in the introduction in this work we are not focusing on coordination issues.

contains only the portion of the environment it traversed and the portions explored by teammates it found on its way.

5 Future works

In the future we want to be able to use the redundancy of the observation, i.e. the fact that two robots observe the same objects can be used to improve the accuracy of the map.

Another issue we want to investigate is the exploitation of the heterogeneity of the Vision Agents fitted on the robots. Our idea is to use the multi-robot system to perform a search and inspect task. The task will be to locate an object in the maze and observe it with the high resolution vision sensor: the perspective camera. Therefore, if the object is discovered by one of the robots equipped only with the omnidirectional sensor, this robot should be able to request the intervention of the robot equipped with the perspective camera.This means that the robot must be able to store the path it traveled and to communicate it to the other robots. Therefore, a compact representation of the map and of the traveled path is required. If the two robots already merged their maps, the problem is quite trivial. If this is not the case, the called robot should first return to the starting position using its own map and then reach the calling robot using the map build by the latter. The starting location is the bridge between the two maps, in fact this is represented in both maps because the robots started from this common position.

6 Conclusions and Acknowledgments

In this paper we presented the research stream we are following to realise a map of an unknown environment with a Distributed Vision System using only omnidirectional vision sensor. The approach we used is to port the work done on the single robot within the Caboto Project on to a multi-robot system. The mapping strategy chosen for Caboto allows this porting without major modifications. The interaction between the Vision Agents is under study with the two experiments presented in the text. The exploitation of omnidirectional mirrors with a profile tailored for the particular application showed to be effective on the field.

At the time of writing experiments are running on such a systems providing theoretical and practical insight.

We wish to thanks the student of the ART-PD and Artisti Veneti Robocup teams who built the robots. This research has been partially supported by: the EC TMR Network SMART2, the Italian Ministry for the Education and Research (MURST), the Italian National Council of Research (CNR) and by the Parallel Computing Project of the Italian Energy Agency (ENEA).

References

1. W. Burgard, A. Cremers, D. Fox, D. Hhnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence (AI)*, 114((1-2)), 1999.
2. W. Burgard, D. Fox, M. Moors, R. Simmons, and S. Thrun. Collaborative multi-robot exploration. In *In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2000.
3. S. Carpin, C. Ferrari, E. Pagello, and P. Patuelli. Bridging deliberation and reactivity in cooperative multi-robot systems through map focus. In M.Hannebauer, J. Wendler, and E. Pagello, editors, *Balancing Reactivity and Social Deliberation in Multi-Agent Systems,*, LNCS. Springer, 2001.
4. H. Ishiguro. Distributed vision system: A perceptual information infrastructure for robot navigation. In *Proceedings of the Int. Joint Conf. on Artificial Intelligence (IJCAI97)*, pages 36–43, 1997.
5. B. Kuipers. The spatial semantic hierarchy. *Artificial Intelligence*, 119:191–233, February 2000.
6. B. J. Kuipers and T. Levitt. Navigation and mapping in large scale space. *AI Magazine*, 9(2):25–43, 1988. Reprinted in *Advances in Spatial Reasoning, Volume 2*, Su-shing Chen (Ed.), Norwood NJ: Ablex Publishing, 1990, pages 207–251.
7. F. Marchese and D. G. Sorrenti. Omni-directional vision with a multi-part mirror. *The fourth international workshop on robocup*, pages pp. 289–298, 2000.
8. T. Matsuyama. Cooperative distributed vision: Dynamic integration of visual perception, action, and communication. In W. Burgard, T. Christaller, and A. B. Cremers, editors, *Proc. of the 23rd German Conf. on Advances in AI (KI-99)*, volume 1701 of *LNAI*, pages 75–88, Berlin, Sept. 1999. Springer.
9. E. Menegatti, , E. Pagello, and M. Wright. Using omnidirectional vision sensor within the spatial semantic hierarchy. In *(Submitted to ICRA 2002)*.
10. E. Menegatti, F. Nori, E. Pagello, C. Pellizzari, and D. Spagnoli. Designing an omnidirectional vision system for a goalkeeper robot. In A. Birk, S. Coradeschi, and P. L. (Eds.), editors, *Proceeding of RoboCup 2001 Int. Symposium (to appear in RoboCup-2001: Robot Soccer World Cup V.)*. Springer, 2001.
11. E. Menegatti, E. Pagello, and M. Wright. A new omnidirectional vision sensor for the spatial semantic hierarchy. In *IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics (AIM '01)*, pages pp. 93–98, July 2001.
12. R. G. Simmons, D. Apfelbaum, W. Burgard, D. Fox, M. Moors, S. Thrun, and H. Younes. Coordination for multi-robot exploration and mapping. In *Proc. of the AAAI Nat. Conf. on AI AAAI/IAAI*, pages 852–858, 2000.
13. S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. Cremers, F. D. Fox, D. Haehnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva. In *Int. Journal of Robotics Research*, volume Vol. 19, pages pp. 972–999, November 2000.
14. M. Wright and G. Deacon. A catastrophe theory of planar orientation. *Int. Journal of Robotics Research*, 19(6):531–565, June 2000.
15. Y. Yagi, Y. Nishizawa, and M. Yachida. Map-based navigation for a mobile robot with omnidirectionalimage sensor copis. *IEEE Trans. on Robotics and automation*, VOL. 11(NO. 5):pp. 634–648, October 1995.