Cooperative Distributed Vision for Mobile Robots

Emanuele Menegatti, Enrico Pagello[†]

Intelligent Autonomous Systems Laboratory
Department of Informatics and Electronics
University of Padua, Italy
[†] *also with* Institute LADSEB of CNR
Padua, Italy
emg@dei.unipd.it

Abstract. Multiple robot systems in which every robot is equipped with a vision sensor are more and more frequent. Most of these systems simply distribute the sensors in the environment, but they do not create a real Cooperative Distributed Vision System. Distributed Vision System has been studied in the past, but not enough emphasis has been posed on mobile robots. In this paper we propose an approach to realize a Cooperative Distributed Vision System within a team of heterogeneous mobile robots. We present the two research streams which we are working on, along with theoretical and practical insights.

1 Introduction

Mobile robots are more and more fitted with vision systems. The popularity of such sensors arises from their capability of gathering a huge amount of information from the environment surrounding the robot. Nowadays, the relatively low cost of the required hardware allows to equip every robot of a mobile robot team with a vision system. An alternative approach is to control a robot team with a centralized vision system, i.e. a unique camera that monitors the whole environment where the robots move. This has been applied in well structured and relatively small environments [2], but it is unfeasible for large environments.

If a camera is mounted on every robot of the team, each robot can gather a more detailed information on its surrounding and the system is more versatile. In fact, fixed cameras positioned in *a priori* location in the environment limit the flexibility and robustness of the system. If something happens outside the field of view of the fixed cameras, the system cannot see this event. If we have cameras mounted on mobile robots, the system can send a robot to inspect the new location of interest.

Mounting a camera on each robot distributes the sensors in the environment, but this is not enough: we aim to the creation of a real Distributed Vision System. A Distributed Vision System requires not only a set of cameras scattered in the environment, but also the sharing of information between the different vision systems.

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.

Fig. 1. Our team of heterogeneous robots

2 Previous Works

Our work has been inspired by the work of Ishiguro [4]. He proposed an infrastructure called *Perceptual Information Infrastructure* (PII). In his paper, he proposed an implementation of the PII with a Distributed Vision System (DVS) composed by static Vision Agents, i.e. fixed cameras with a certain amount of computational power. The cameras, strategically placed in the environment, navigate a mobile robot. The robot is not autonomous, in the sense that it needs the DVS to navigate, but it has a certain amount of deliberative power, in the sense that it decides which Vision Agent provides him the more reliable information on the surroundings. The vision algorithms of the Vision Agents are really simple, because of the assumption that every Vision Agents is static.

A parallel but independent work is the one of Matsuyama [5]. Matsuyama explicitly introduced mobile robots in its Cooperative Vision System. In the experiments presented, he 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. As far as we know, no attempt has been tried to realize a DVS with truly mobile robots running robot vision algorithm.

3 The aim of our work

Our aim is to introduce a real Mobile Vision Agent in the DVS architecture, i.e. to apply the ideas and the concepts of Distributed Vision to a mobile robot equipped with a camera.

The domain in which we are testing our ideas is the RoboCup competitions. We are on the way to create a Distributed Vision System within a team of heterogeneous robots fitted with heterogeneous vision sensors. We want to create a dynamic model of the environment, which can be used by mobile robots or humans to monitor the environment or to navigate through it. The model of the environment is built fusing the data collected by every Vision Agent. The redundancy of observers (and observations) is a key issue for system robustness.

4 Implementation

4.1 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. Exploiting the knowledge acquired in our previous research [7], we want to create a Cooperative Vision System using an omnidirectional and a perspective vision system mounted on the same robot. The robot is our football player robot, called Nelson that we entirely built starting from an ActivMedia Pioneer2 base (see the web page www.dei.unipd.it/~robocup). The omnidirectional vision system is a catadioptric system composed by a standard colour camera and an omnidirectional mirror we designed [6].The omnidirectional camera is mounted on the top of the robot and offers a complete view of the surroundings of the robot [1]. 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].

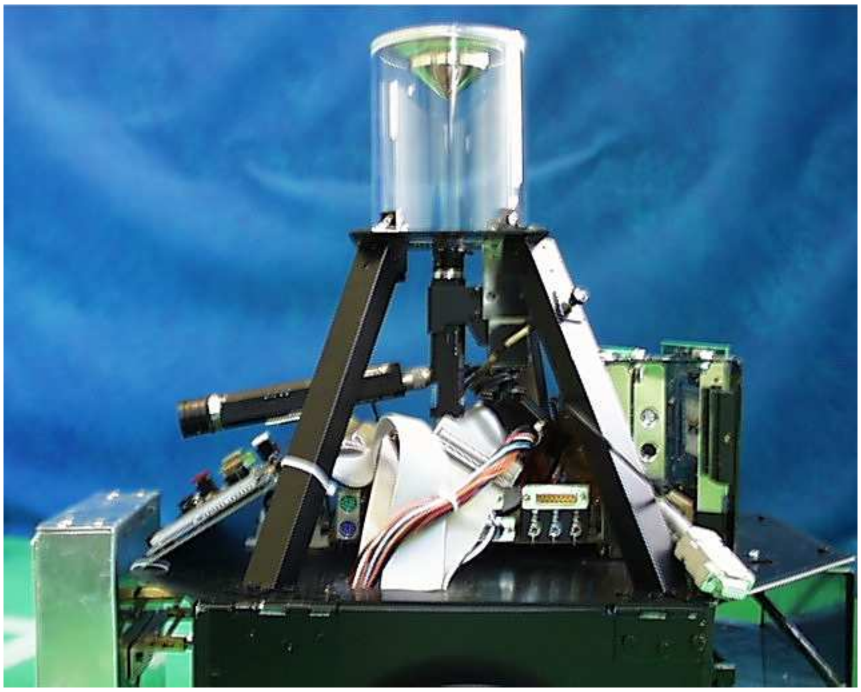

Fig. 2. A close view of the vision system of Nelson. On the left, the perspective camera. In the middle, pointed up-ward the omnidirectional camera

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

4.2 Coordination of several VAs mounted on different robots

Another stream of research is the creation of a Cooperative Distributed Vision System for our team of football player robots. Our aim is to implement the idea of the Cooperative Object Tracking Protocol proposed by Matsuyama [5]. 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. VAs exit the agency, if they are not able to see the tracked object anymore. 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 Agent is seen 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. All 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.

The described algorithm has been realised by Matsuyama with his fixed vied point cameras. His system was composed of four pan tilt zoom cameras mounted on a special active support in order to present a fixed view point. The system is able to track a radio controlled toy car in a small closed environment. As mentioned before, in such a system there is not a truly mobile agent. Moreover the vision algorithm used is typical of static Vision Agents. In fact, this is a smart adaptation of the background subtraction technique.

Our novel approach is to implement the Cooperative Object Tracking Protocol within a team of mobile robots equipped with Vision Agents. This requires a totally new vision approach. In fact, the point of view of the Vision Agent changes all the time. The changes in the image are due not only to the changes in the world (as in the Matsuyama testbed), but also to the change of position of the Vision Agent. Therefore, we need a vision algorithm able to identify the different objects of interest and not only to reveal the objects that are moving. Moreover, we have to introduce a certain amount of uncertainty in the estimation of the position of these objects, because the location of the Vision Agents is not known exactly anymore and there are errors in the determination of the relative distance between the objects the Vision Agents.

To explain these issues, let us come back to our RoboCup example. Above we said that if a new Vision Agent sees the ball, it sends a message to the master that checks if it has seen the correct ball. In a RoboCup match there is just one ball, but sometimes what a robot identifies as a ball is not the correct one. This can result either because the robot sees objects resembling the ball, and erroneously interprets them as a ball (like spectators hands or reflex of the ball on walls), or because it is not properly localized and so it reports the ball to be in a fallacious position.

To cope with the uncertainty in the objects position, every Vision Agent transmits to the master the calculated ball position with a confidence associated to this estimation. The master dispatches to the other robot a position calculated as an average of the different position estimations, weighted by the confidences reported by every Vision Agent (if there is more than one Vision Agents in the agency).

Especially in the described dynamic system, the master role is crucial in the correct functioning of the agency. The master role cannot be statically assigned. The ball is continuously moving during the game. The first robot that sees the ball will not have the best observational position for long. So, the master role must pass from robot to robot. The processes of swapping the master role is

Fig. 3. A close view of two of our robots. Note the different vision systems

critical. If the master role is passed to a robot that sees an *incorrect ball* all the agency will fail in the *ball tracking* task.

The simplest solution could be to pass the master role to the robot with the highest confidence on the ball position. This means to shift the problem to identify a reliable confidence function. This makes sense, because the confidence function will be used for two services that are two sides of the same coin. In fact, if a robot is correctly localized and correctly calculates the relative distance of the ball, it will have strong weight in the calculation of the ball position. Given this, it can reliably take the role of master.

The confidence function The confidence function ψ_{abs} associated to the reliability of the estimation of the absolute ball position is a combination of several factors. It has to account for the different aspects that contributes to a correct estimation of the ball position. In fact, the position of the ball in the field of play is calculated by a vectorial sum of the relative distance of the ball from the robot and the absolute position of the robot in the pitch. So, the confidence of the estimation of the absolute position of the ball is the sum of the confidences function associated to the self-localisation, ψ_{sl}, and of the confidence function associated to the estimation of the relative position of the ball with respect to the robot, ψ_{rel}.

$$\psi_{abs} = \psi_{sl} + \psi_{rel} \tag{1}$$

The self-localisation process uses the vision system to locate landmarks in the field of play. The process is run only by time to time and if the landmarks are visible. Between two of these process the position is calculated with the odometers. This means that the localisation information degrades with time. The confidence function associated with the self-localisation is the result of the following contribution:

– type of vision system (perspective, omnidirectional, etc.);
– a priori estimated absolute error made from the vision system in the calculation of the landmarks position;
– time passed after the last self-localisation process;

The relative position of the ball with respect to the robot is calculated as in [6]. The confidence function in this process presents the following contribution:

– type of vision system;
– distance from the ball;

At the moment the exact definition of the confidence function is under testing. The experiments will tell us how much every contribution should weight in the final function.

5 Conclusion

In this paper we presented the two research streams we are following to implement a Cooperative Distributed Vision System.

In this paper we proposed to realise the DVS with heterogeneous mobile Vision Agents. We suggested a way to fuse the information coming from two heterogeneous Vision Agents mounted on the same robot. Regarding the problems introduced by the mobile Vision Agents, we suggested a way to cope with the uncertainty introduced in the localisation of the objects of interest.

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

Acknowledgments

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. A. Bonarini. The body, the mind or the eye, first? In M. Veloso, E. Pagello, and H. Kitano, editors, *RoboCup99: Robot Soccer World Cup III*, volume 1856 pp. 210-221 of *LNCS*. Springer, 2000.
2. J. Bruce, T. Balch, and M. M. Veloso. Fast and inexpensive color image segmentation for interactive robots. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '00)*, volume 3, pages 2061 – 2066, October 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 International Joint Conference on Artificial Intelligence (IJCAI97)*, pages 36–43, 1997.
5. T. Matsuyama. Cooperative distributed vision: Dynamic integration of visual perception, action, and communication. In W. Burgard, T. Christaller, and A. B. Cremers, editors, *Proceedings of the 23rd Annual German Conference on Advances in Artificial Intelligence (KI-99)*, volume 1701 of *LNAI*, pages 75–88, Berlin, Sept. 13–15 1999. Springer.
6. E. Menegatti, F. Nori, E. Pagello, C. Pellizzari, and D. Spagnoli. Designing an omnidirectional vision system for a goalkeeper robot. In *Proceeding of RoboCup 2001 International Symposium*, 2001.
7. E. Menegatti, E. Pagello, and M. Wright. A new omnidirectional vision sensor for the spatial semantic hierarchy. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM '01)*, July 2001.