Bayesian Sensor Fusion for Cooperative Object Localization and World Modeling

Pedro Pinheiro, Pedro Lima
Instituto de Sistemas e Robótica, Instituto Superior Técnico
Av Rovisco Pais 1, 1049-001 Lisboa, Portugal
{pmpin@mega.ist.utl.pt, pal@isr.ist.utl.pt}

Abstract. This paper introduces a method for representing, communicating and fusing distributed, noisy and partial observations of an object by multiple robots. This technique describes how to model sensors and the information they acquire. Each sensor is considered as a team member making decisions locally to achieve a local estimate. The local estimates of a robot are then fused with the other robots local estimates to achieve a global fusion estimate of the objects surrounding the team, creating a much more reliable and accurate world model. This method was implemented and tested in RoboCup Middle Size League robots.

Keywords : Sensor Fusion, Object Localization, RoboCup

Introduction

Individual robots typically obtain partial and noisy data from the surrounding environment. This data is often erroneous, leading to miscalculations and wrong behaviors, and to the conclusion that there are fundamental limitations on the reconstruction of environment descriptions using only a single source of sensor information. If robot systems are ever to achieve a degree of intelligence and autonomy, they must be capable of using many different sources of sensory information in an active and dynamic manner.
In dynamic environments, information previously collected about currently unobservable parts of the environment can quickly become inaccurate. Sharing information among robots increases the effective instantaneous visibility of the environment, allowing for more accurate modeling and more appropriate response. Information collected from multiple points of view can provide reduced uncertainty, improved accuracy and increased tolerance to single point failures in estimating the location of observed objects. By combining information from many different sources, it would be possible to reduce the uncertainty and ambiguity inherent in making decisions based only in a single information source. Our goal is to apply the sensor fusion method introduced by Durrant-Whyte [1] to a team of real robots. Durrant-Whyte's method is summarized in Sections 2 and 3. In Section 4 we discuss our experimental setup and results obtained with real soccer robots. Finally, in Section 5 some conclusions of this work are drawn.

1. Background and Related Work

Even though one can find related applications in other areas, we refer only to related work on the specific application described in this paper: soccer robots. Most soccer robot team approaches use vision and/or sonar to localize the robots and vision to locate objects in the environment. Some teams share information for planning and dynamic role modeling with shared data [9]. Other distributed sensing approaches include merging independent gridcell occupancy probabilities measured by multiple robots [2], others use Kalman filters to track objects [3], or a combination of Kalman filters with Markov localization [6]. Others yet use a probabilistic state estimator [7]. The task addressed in [1] is distinct from the others described above, since the author focuses on fusing multiple simultaneous observations of the same object from distributed vantage points (as opposed to observations from the same vantage point over multiple instants in time). Our goal is to provide more accurate instantaneous estimations of the location of dynamic objects that are simultaneously visible by multiple robots, without relying on historical data.

2. Fusing Gaussing Distributions

In order to cooperatively use sensor fusion, team members must exchange sensor information. This information exchange provides a basis through which individual sensors can cooperate with each other, resolve conflicts or disagreements, and/or complement each other's view of the environment. Our goal is to analyze the efficiency of a previously introduced sensor fusion method [1], where uncertainties in the sensor state and observation are modeled by Gaussian distributions. This approach takes in to account the last known position of the object and tests if the readings obtained from several sensors are close enough, using the Mahalanobis distance, in order to fuse them. When this test fails, no fusion is made and the sensor reading which has less variance (more confidence) is chosen. The conditions under which this test fails or succeeds are presented in Section 3. The remainder of this section provides the necessary mathematical background to understand how the Gaussian distributions are fused.

Durrant-Whyte[1] considers a sequence of observations $z^P = \{z_1, \ldots, z_n\}$, of an environment feature $p \in P$, which are assumed to derive from a sensor modeled by a contaminated Gaussian Density, so that the i^{th} observation is given by:

$$f_i(z_i \mid p) = \left[(1-e)N\left(p, \Lambda_i^1\right) + eN\left(p, \Lambda_i^2\right)\right] \tag{1}$$

where $0.05 < e < 0.1$ and $\Lambda_i^2 >> \Lambda_i^1$.

It is well known that if the prior distribution $p(p)$ and the conditional observation distribution $f(z \mid p)$ are modeled as independent Gaussian random vectors $p \sim N(\hat{p}, \Lambda_0)$ and $z_1 \sim N(\hat{p}, \Lambda_1)$ respectively, then the posterior distribution $p(p \mid z)$ after taking a single observation z_1 can be derived using Bayes' law and is also jointly Gaussian with mean vector

$$\hat{p}' = \left[\Lambda_0^{-1} + \Lambda_1^{-1}\right]^{-1}\left[\Lambda_1^{-1} z_1 + \Lambda_0^{-1} \hat{p}\right] \tag{2}$$

and covariance matrix

$$\Lambda' = \left[\Lambda_0^{-1} + \Lambda_1^{-1}\right]^{-1} \tag{3}$$

This method can be extended to n independent observations, as explained in [1].

3. The Multi-Bayesian Team

In the Multi-Bayesian system, each team member individual utility function is given by the posterior likelihood for each observation z_i:

$$u_i\left(\overline{p} = d_i(z_i), p\right) = p\left(p \mid z_i\right) \approx N\left(\hat{p}, d_i\right) \qquad i = 1,2 \tag{4}$$

A sensor or team member will be considered *rational* if, for each observation z_i of some prior feature $d_i(z_i) \in P$, it chooses the estimate that maximizes its individual utility $u_i(d_i(z_i), p) \in \Re$. In this sense, utility is just a metric for constructing a complete lattice of decisions, allowing any two decisions to be compared in a common framework. For a two-member team, the team utility function is given by the joint posterior likelihood:

$$U(\overline{p}(z_1, z_2), p) = F(p \mid z_1, z_2) = f_1(p \mid z_1) f_2(p \mid z_2) \quad . \tag{5}$$

The advantage of considering the team problem in this framework is that both individual and team utilities are normalized so that comparisons can be performed easily, supplying a simple and transparent interpretation to the group rationality problem. The team itself will be considered group rational if together the team members choose to estimate $\overline{p} \in P$ (environment feature), which maximizes the joint posterior density.

$$\overline{p} = \arg\max F(p \mid z_1, z_2) = \arg\max \; f_1(p \mid z_1) f_2(p \mid z_2). \tag{6}$$

There are two possible results for (6)

- $F(p \mid z_1, z_2)$ has a unique mode equal to the estimate $\overline{p}$;

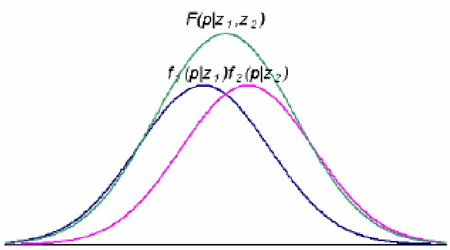

Fig. 1 - Two Bayesian observers with joint posterior likelihood indicating agreement

- $F(p \mid z_1, z_2)$ is bimodal and no unique group rational consensus estimate exists.

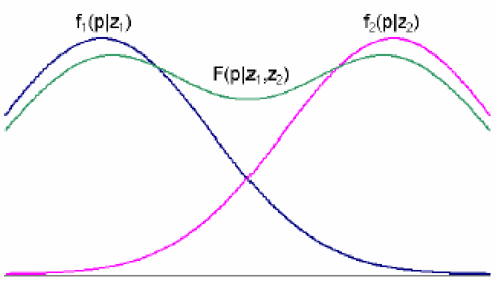

Fig. 2 - Two Bayesian observers with joint posterior likelihood indicating disagreement

If $F(p \mid z_1, z_2)$ has a unique mode, as displayed in Fig. 1, it will satisfy:

$$\max \; F(\mathrm{p} \mid \mathrm{z}_1, \mathrm{z}_2) \geq \max \; f_i(\mathrm{p} \mid \mathrm{z}_i) \qquad i = 1,2 \tag{7}$$

Conversely, if $F(p \mid z_1, z_2)$ is bimodal, as displayed in Fig. 2, then:

$$\max f_i(\mathrm{p} \mid \mathrm{z_i}) \geq \max F(\mathrm{p} \mid \mathrm{z}_1, z_2) \qquad i = 1,2 \qquad (8)$$

A rational team member will maximize utility by choosing to either agree or disagree with the team consensus. If a team member satisfies (8), then it will not cooperate with the team estimate. Thus the decision made by a team member based of its observations z_i is:

$$\overline{p} = d(z_i) = \arg\max\{f_i(p \mid z_i), F(p \mid z_1, z_2)\} \qquad i = 1,2 \qquad (9)$$

Whether or not the individual team members will arrive at a consensus, the team estimate will depend on some measure of how much they disagree $| z_1 - z_2 |$. If z_1 and z_2 are close enough, then the posterior density $F(p \mid z_1, z_2)$ will be unimodal and satisfy (7), with the consensus estimate given by (6). As $| z_1 - z_2 |$ increases, $F(p \mid z_1, z_2)$ becomes flatter and eventually bimodal. At this point, the joint density will satisfy (8), and no consensus team decision will be reached. To find the point at which this space is no longer convex and disagreement occurs, one must ensure that the second derivative of the function $F(p \mid z_1, z_2)$ is positive. Differentiating leads to:

$$\frac{\partial^2 F}{\partial p^2} = \frac{1}{f_1}\frac{d^2 f_1}{dp^2} + \frac{1}{f_2}\frac{d^2 f_2}{dp^2} + \frac{2}{f_1 f_2}\frac{df_1}{dp}\frac{df_2}{dp} = \left(s_1^{-2} + s_2^{-2}\right) - \left[s_1^{-2}(p - z_1) + s_2^{-2}(p - z_2)\right]^2 \qquad (10)$$

For this to be positive and hence $F(p \mid z_1, z_2)$ to be convex, we are required to find a consensus over the feature of the environment p which satisfies

$$\left[s_1^{-2}(p - z_1) + s_2^{-2}(p - z_2)\right]^2 \left(s_1^{-2} + s_2^{-2}\right)^{-1} \leq 1 \qquad (11)$$

Notice that (11), is a normalized weighted sum, a scalar equivalent to the Kalman gain matrix. The consensus $\overline{p}$ which maximizes F is therefore given by

$$\overline{p} = \frac{\left(s_1^{-2} z_1 + s_2^{-2} z_2\right)}{\left(s_1^{-2} + s_2^{-2}\right)} \qquad (12)$$

Replacing (12) into (11), we obtain

$$\frac{(z_1 - z_2)(z_1 - z_2)}{\left(s_1^{-2} + s_2^{-2}\right)} = D_{12}(z_1, z_2) \qquad (13)$$

where $D_{12} \leq 1$. The disagreement measure $D_{12}(z_1, z_2)$ is called the Mahalanobis distance.

4. Experimental Test and Results

4.1 Experimental Setup

An experiment was conducted on the ISocRob soccer team of RoboCup Middle Size League [8]. The software architecture of the robots is composed by several micro-agents that acquire and process sensor data, so a new micro-agent was created to handle the sensor fusion. The data is acquired by two micro-agents that handle each of the two on-board cameras, here designated as the *up* camera and the *front* camera. This data is then processed

and sent to the team blackboard, a distributed memory that stores all the relevant information about the robot status, micro-agents behaviors and acquired sensor data. The *up* camera produces a 360° view of part of the field, while the *front* camera produces an image of what is in front of the robot. Since both players and ball are moving, their positions are dynamically evolving during a game. In order to deal with this, we had to model the observed data in both cameras to reflect the uncertainty of the objects position, based on its distance to the player who was collecting the data. This was done by placing the ball in pre-defined areas around the player and calculating the observed variance of the ball position. After several observations, an observation model was built for the cameras. The goal is to obtain a better estimate for all player positions and the ball position. The results were obtained and visualized on a game interface client, which allowed the visualization of the location of the relevant objects as determined by the sensor fusion method.
Using Durrant-Whyte's fusion algorithm, the decision process to determine the ball position is made by first determining if both observed ball positions from the two cameras can be merged locally through the Mahalanobis distance. This is accomplished by putting a time stamp in each camera observation, and using the time difference between stamps to modify the variance of each observation, in order to synchronize the fusion. When this synchronization is possible, the ball position will be the result of the fusion; otherwise, the observation with the smallest variance is chosen, meaning that the observation with the highest confidence is used to determine the ball position. After the local ball position estimate has been determined, the estimation of the global ball position is attempted, by fusing all local estimates of each robot, to get a global sensor fusion, as shown in Figure 3. Each player acts as a sensor, taking observations from its two cameras, modifying the variance based on the difference of the observation time stamps, fusing and reporting them to the other team members.

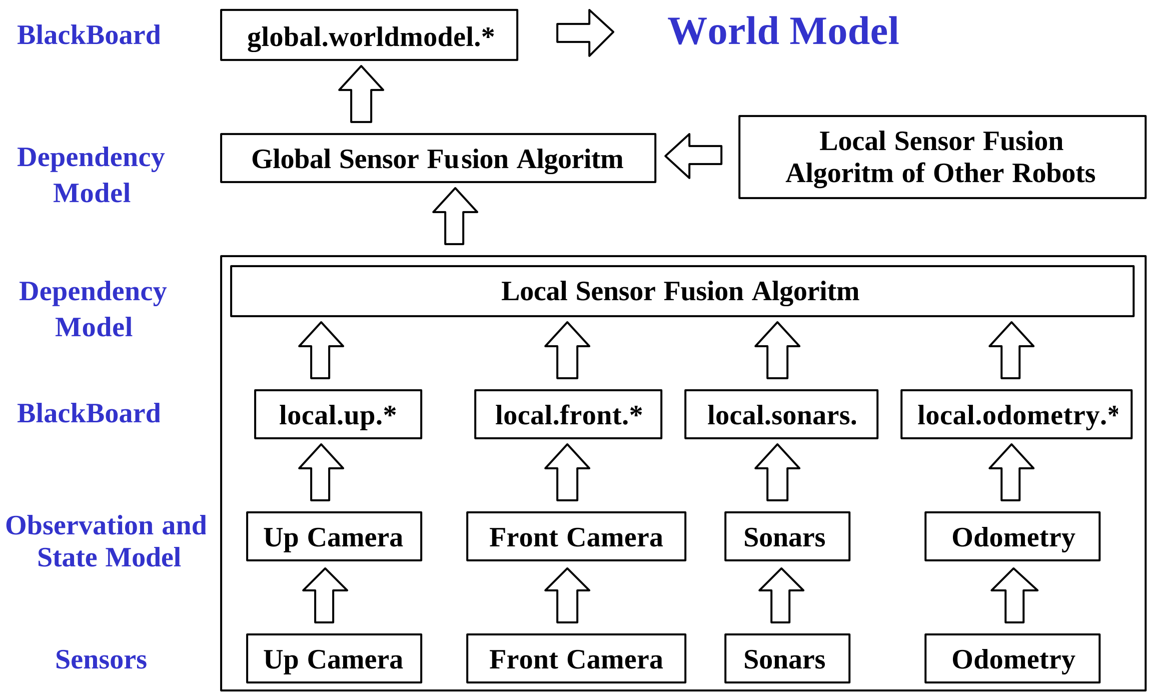

Fig. 3 – Diagram of the Local and Global Sensor Fusion

4.2 Experimental Results

To test the local sensor fusion of both cameras of the robot, the ball was positioned on the field in different positions along a straight line, making several observations with both cameras. Figure 4 shows both the observations made by the two cameras, the real position of the ball and the estimated position, calculated by Durrant-Whyte's method. The result of the fusion improved the estimation of the ball position, converging to the real position of the ball. It can be seen from the plot in the figure that the *front* camera provides better estimates for distances below two meters, while the *up* camera gives better results for long distances. Fusing the information from the two cameras, the estimated position is much closer to the real position and has reduced estimate noise, improving the certainty of the estimation.

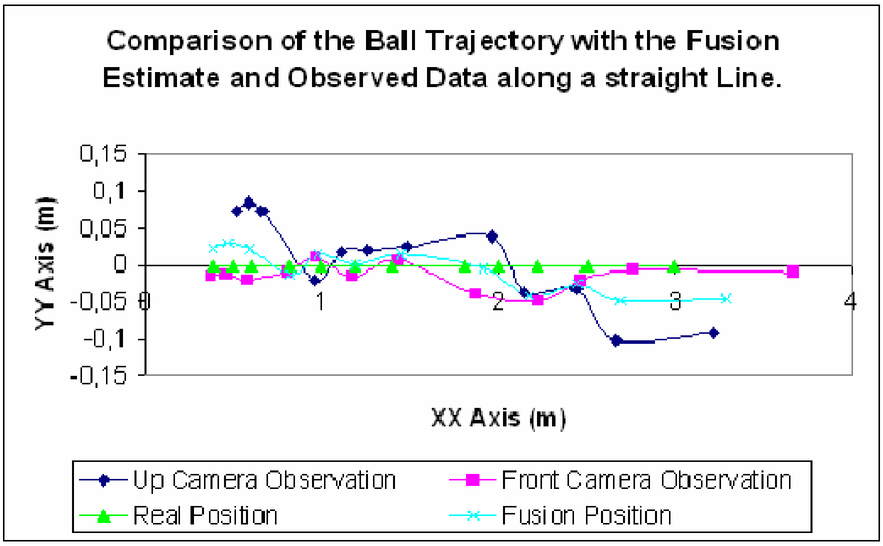

Fig.4 - Comparison of the Ball Trajectory with the Fusion Estimate and Observed Data along a straight line.

To test the global sensor fusion, three robots were placed on the field. We then ran the algorithm in each robot with only local sensor fusion working (Figure 5) and then with both local and global sensor fusion working (Figure 6). Each robot has a measure of quality (local fusion variance) of its local sensor fusion, using it to decide who has priority in the global sensor fusion. The robot with the best measure of quality has priority over the others.

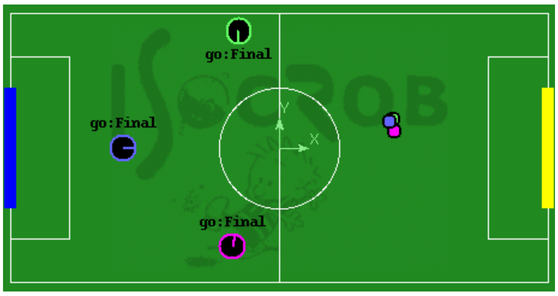

Fig. 5 – Local Sensor Fusion Enabled and Global Sensor Fusion Disabled.

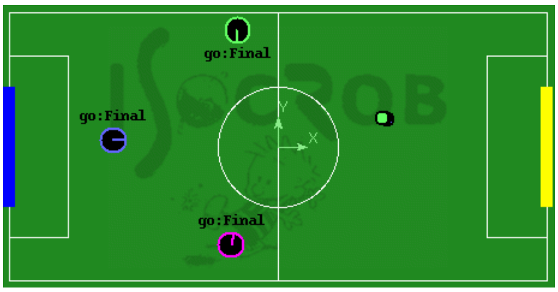

Fig. 6 – Both Local Sensor Fusion and Global Sensor Fusion Enabled.

As seen in Figures 5 and 6, the global sensor fusion improved the ball estimate, and since the new ball position information is shared by all robots, a more robust world model results, allowing the team robots to share information and create more complex and interesting actions. One of these actions is illustrated in Figure 7 where, although one of the robots cannot see the ball with its own cameras, because it is too far away, it knows where the ball is, since all robots share the same world information. This is the result of communicating all the features that each robot extracts from the environment to all the other teammates, and then using sensor fusion to validate those observations. Testing the agreement among all the team sensors eliminates sporadic and erroneous observations. In Figure 8, the robot in the bottom part of the field cannot see the ball, so it gets the ball position from the global fusion of the other robot observations. Since the other two robots disagree with each other, the global fusion becomes equal to the local fusion of the robot with the best variance among the two.

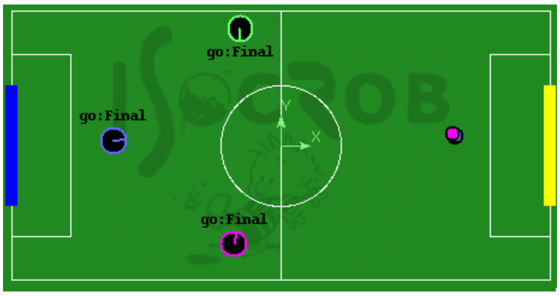

Fig. 7 – Leftmost robot receives ball position information of the other two.

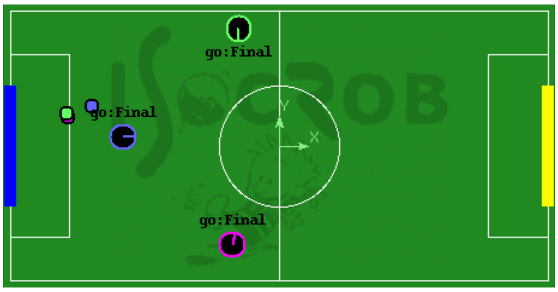

Fig. 8 – Bottom robot receives ball position from top robot, while top and leftmost robot disagree.

In Figure 9 we see two robots showing disagreement. This happened in this case because there were two balls in the field and each robot was detecting a ball in different positions. Although each robot has its own local sensor fusion estimate, they cannot reach an agreement about the global sensor fusion. When this happens, the robot makes its global sensor fusion estimate equal to its local sensor fusion estimate. In Figure 10 we see the same two robots showing agreement. Although they have slightly different local sensor fusion estimates, they have the same global sensor fusion estimate of the ball, which is a result of the fusion of their local estimates.

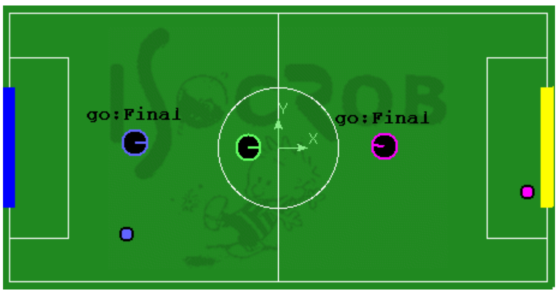

Fig. 9 – Two robots showing disagreement.

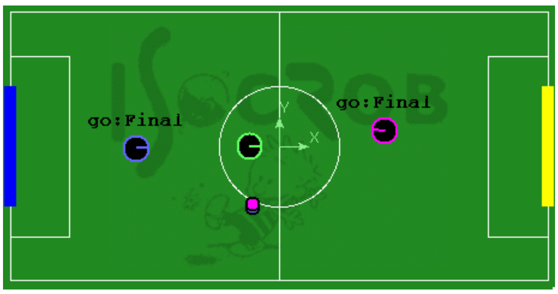

Fig. 10 – Two robots showing agreement.

Before each local fusion is made, each sensor observation and the local sensor estimate at the previous step are fused, with an increase in the variance of the latter, to reflect the time that has passed since the fusion was made. This helps to validate the new observation, because if fusion is successful then the new observation is a valid one and we are predicting the same feature as in the previous fusion operation. Otherwise, this means that the latest observation was probably a bad one and that we could not predict the feature evolution.

5. Conclusions

As shown in this work, our modifications and application of Durrant-Whyte's approach is very effective to improve the position estimation of relevant world features, because it considers the previous ball position in the estimation of its new value, and uses a measure

of the quality of the observations and fusions to make decisions, thus eliminating bad observations and producing a more robust world model.
Also, if sensor outputs vary to such degree that the Mahalanobis distance increases and becomes greater than one, then we have two disagreeing sensors. In this case, our method chooses the sensor output with the best quality (smallest variance). This will eliminate any erroneous and spurious data that might appear, giving a much more accurate view of the world and of its state. These are the reasons why we have chosen to use this method for real robots.
Using multi-sensor Bayesian techniques for the team decision problem, techniques for control and coordination of the information acquisition were developed and implemented in each individual team member. These techniques will allow more frequent cooperation between team members, so as to solve conflict situations and achieve a team consensus faster.
Although in this case study we were only concerned with the ball position, this method can be extended to any other feature that the robots can extract from the environment., particularly to improve the estimation of teammate locations. This is the work we are currently carrying out.

Acknowledgements

Work partially supported by Programa Operacional Sociedade de Informação (POSI) in the frame of QCA III and the FCT Project Rescue - Cooperative Navigation for Rescue Robots (SRI/32546/99-00).

References

[1] Hugh F. Durrant-Whyte, *Integration, Coordination and Control of Multi-Sensor Robot Systems*, Kluwer Academic Publishers, 1988.
[2] A. Cai, T. Fukuda and F. Arai, “Information Sharing among Multiple Robots for Cooperation in Cellular Robotic System”, In: *Proc. IROS 97*.
[3] A.W. Stroupe, M.C. Martin, and T. Balch. “Merging Probabilistic Observations for Mobile Distributed Sensing”. Carnegie Mellon University Technical Report CMU-RI-TR-00-30, 2000.
[4] Gonçalo Neto, Hugo Costelha, Pedro Lima, "Topological Navigation in Configuration Space Applied to Soccer Robots", *Proceedings of RoboCup 2003 International Symposium*, Padova, Italy, 2003.
[5] Miguel Arroz, Vasco Pires, Luis Custódio, "Logic based Distribution Decision System for a Multi-Robot Team", *Actas do Encontro Cientifico do Robotica 2003 - Festival Nacional de Robótica*, 2003.
[6] Markus Dietl, Jens-Steffen Gutmann and Bernhard Nebel, “CS Freiburg: Global View by Cooperative Sensing”, in *Robocup 2001: Robot Soccer World Cup V*, Silvia Coradeschi, Satoshi Tadokoro, Rainer P. S. Noske, Andreas Birk (editors).
[7] Thorsten Schmitt, Robert Hanek, Sebastian Buck and Michael Beetz, “Cooperative Probabilistic State Estimators for Vision-Based Autonomous Soccer Robots”, in *Robocup 2001: Robot Soccer World Cup V*, Silvia Coradeschi, Satoshi Tadokoro, Rainer P. S. Noske, Andreas Birk (editors).
[8] Pedro Lima, Luis Custódio “IsocRob 2003 Team Description”, *Proceedings of RoboCup 2003 International Symposium*, Padova, Italy, 2003.
[9] Daniele Nardi, Giovanni Adorni, Andrea Bonarini, Antonio Chella, Giorgio Clemente, Enrico Pagello, and Maurizio Piaggio, “ART-99 Azzurra Robot Team”, in *Robocup 1999: Robot Soccer World Cup III*, M. Veloso, E. Pagello and H. Kitano H (editors)