An omnidirectional vision sensor for fast tracking for mobile robots

Andrea Bonarini, Paolo Aliverti, Michele Lucioni
Dipartimento di Elettronica e Informazione
Politecnico di Milano
Piazza Leonardo da Vinci 32, Milano, Italy
e-mail: bonarini@elet.polimi.it

Abstract

We present an omnidirectional vision system we have implemented to provide our mobile robot (Polyphemus - the one with only one eye) with a fast tracking capability. Polyphemus has to recognize elements in a semi-structured environment (the soccer field of Robocup) at a rate close to 20 Hz. We have designed a multi-shaped mirror to optimize the resolution of the image and its vertical coverage. The interpretation system is based on focused, multi-level, and opportunistic procedures.

1 Introduction

We present an omnidirectional vision system we have implemented to provide our mobile robot Polyphemus with a fast tracking capability. Polyphemus is one of the members of ART (Azzurra Robocup Team[5]) the Italian, national team of robots participating in the Robocup Initiative [1]. In section 2 we present Robocup, and the experimental setting.

Our vision system is based on a camera facing upwards beneath a mirror obtained by the intersection of a truncated cone and a sphere. The environment surrounding the robot reflects in the mirror and the camera can take an image that contains information about what is surrounding the robot. The shape of the mirror optimizes the resolution on the border of the image, making it possible to recognize the smallest object in the field (the ball), and the amplitude of the visual field, covering about 6 meters around the robot, and more than 50 cm in height on the border of the image. Section 3 is devoted to the description of the mirror and the other physical features of the sensor. In section 4 we present the image interpretation system, and in section 5 we discuss the performance of the overall system.

2 Robocup, the experimental environment

Within the Robocup middle size league, each robot should have a maximum width and length of about 50 cm. It should be able to play soccer interacting with 3 teammates and 4 competitors in a field sized 4x8 meters. The ground is a green carpet surrounded by white walls, the ball is reddish, all robots should have a "mostly" black body and a cyan or purple marker 10 cm high, visible from any direction and positioned between 30 and 60 cm above the ground. It was clear at the last Robocup World Championship (Paris-98 [3]) that a fast vision system, coupled with fast movements are compulsory to obtain satisfactory performance. One the most important activities a soccer robot should perform is looking for the ball and tracking it, but also looking for goals and other environmental elements for self-localization, and other robots to interact with them. The ball can move at a maximum speed of up to 80 cm/sec, and the robots may run at 50 cm/sec.

Most of the robots had a single, fixed camera on board, pointed towards. This vision system covers only a portion of the field and requires fast movements of the robot to track the ball and other moving robots. Moreover, when navigating purposingly,for instance bringing the ball or trying to reach a position, the vision direction (usually the heading of the robot) might not be optimal. For instance if the robot brings the ball It has to check the presence of the ball in front of it, but also the presence of the goal, or of competitors in other directions. To face this problem, some teams mounted two wide range cameras back to back, but this leaves some blind regions, end requires double image processing. Other teams mounted a camera on a pan and tilt system, to decouple the vision direction from the robot movement, but the pan and tilt system is relatively slow and did not allow effective tracking.

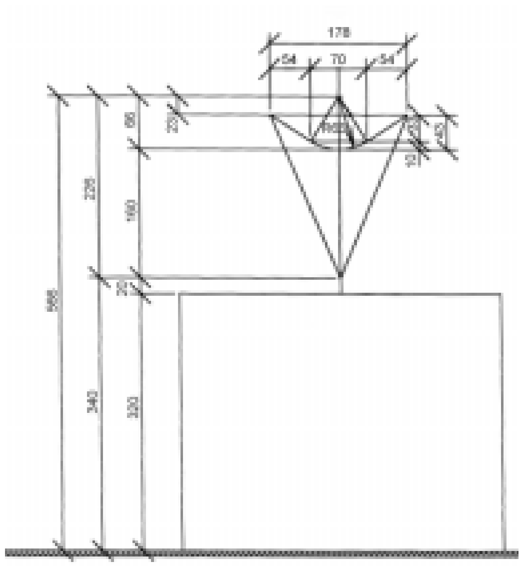

Figure 1. The designed mirror over a box representing the robot body. Measures are in [mm].

The sensor we are proposing (see figure 1) can take information about all the surrounding environment from only one image, thus requiring only the time to process one image. Moreover, the robot can take any heading while tracking interesting environmental elements all around. In the next section we discuss the shape of the mirror and the physical settings. Section 4 is devoted to the algorithms for fast image interpretation.

3 The mirror and the camera

We have designed the sensor with the specific aims of covering the widest area around the robot, while maintaining enough resolution on the far area to be able to identify relibly the smallest object in the environment, which is the ball. The ball diameter is 20 cm, and we we needed at least a resolution of 1 pixel/cm to reliably distinguish it in the image. Moreover, we want to be able to see a significant part of the ball when it is in contact with the robot body, in order to be able to control it.

These requirements cannot be matched by the classical shapes used in for sensors of this kind [7, 4, 2, 6]: cone, sphere, hyperbole, parabola. We have started by considering a cone-shaped mirror, which was immediately discarded because of the strong distorsions that introduces on the image. A benefit of this mirror lies in the possibility to select the pan angle to obtain a wide field of vision at will. Working with a spheric-shaped mirror we have noticed that it doesn't modify heavily the shape of objects, but a large percentage of the image is useless for our goals because it has been taken up by the chassis of the underneath robot. The spheric-shaped mirror we adopted had a bend radius of 0.18 m centered at 0.67 m height from ground. The camera was at 0.32 m from ground. With this system we obtained a vision field ranging more than 5 m but the resolution (in pixel/cm) of the image obtained in this way was not sufficient for the other image processing activities.

To give an idea of the resolution and the properties of the image reflected in this mirror, we give some data in the following table:

Dist (m)	PN	DO (cm)	R (pixel/cm)
1.6	20	25	0.8
2	15	20	0.75
4	13	50	0.26

where $Dist$ is thedistance from the mirror in meters, PN is the number of pixels in the image coresponding to a dimension of the object DO given in the third column. R is the resolution in pixel/cm. Therefore, it was necessary to find a compromise between the bend radius of the mirror and the height where it was put, avoiding an excessive reduction of the vision field, or its exagerate opening with drastic reductions of the resolution.

We have designed a mirror composed by a sphere intersecting a reversed cone, as depicted in figure 1. This shape allows to exploit the characteristics of the spherical mirror to a distance of 2 m, so the objects which fall in this range are not much deformated and have a satisfying resolution. The objects at a distance above 2 m are reflected in the conical part of the mirror, designed to allow the identification of objects at a distance up to 5 m with sufficient resolution. Although the conic part of the mirror can show more distorsions, this allows to obtain higher resolution than the spherical mirror. In the following table, some data taken from images produced by the last mirror.

Dist (m)	PN	DO (cm)	R (pixel/cm)
1.6	16	15	1.07
2	32	16	2.00
4	33	30	1.1

When designing the miror we also had to take into consideration the dimension limits imposed by RoboCup regulations. So we have obtained a mirror

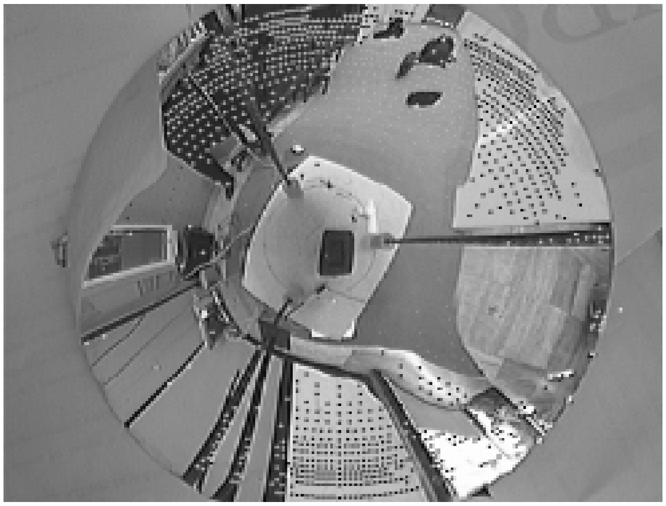

Figure 2. A tipical image taken by the sensor. The marks on it corespond to the receptors corresponding to the different classi• ed targets (see below).

with reduced dimension in comparison with the previous, sperical one without reducing the range of vision and succeeding in positioning the mirror at a lower height above the camera. The spherical mirror employed before had a diameter of 270 mm and a bend radius of 180 mm. It was put at a height of 330 mm from the camera, which was 320 mm from the ground. The new mirror has a diameter of 178 mm and the spherical part has a bend radius of 66 mm; it is 160 mm from the camera and 480 mm from the ground.

With the spheric-shaped mirror we succeeded in obtaining resolutions of 0.26 pixel/cm at a distance of 4 m, with the new mirror the resolution is doubled, achieving 0.5 pixel/cm. Indeed the maximum range of the mirror is above 5 m to include in the supplied image the walls limiting the field, which are 0.5 m high.

The camera is a low cost card camera, having 512x582 sensible elements, with a view angle of about 60 degrees.

A typical image taken from the sensor we have designed is shown in black and white in figure 2. In white, at the center of the image the body of the robot, on the left a computer video, on the top part of the image abll and other robots. The black and whithe marks come from the elaboration discussed in the next section.

4 Image processing

In this context, image processing should give to the robot control system, in real time information about the position of different identified objects and self-localization. Working on images of 640 per 640 pixel where each colour can change on 255 levels we immediately noticed that in applications which require a steady images flow at a rate of at least of 10 frames per second, we had to process a considerable data flow.

To achieve this speed by employing the common image processing and computer vision techniques, it is necessary to add knowledge and provide smart algorithms. Taking hints from biology, we have designed an image processing system based on the concept of *receptor*. The generical term "receptor" is used for any byological unit able to receive stimuls from the outside sourrondings and to transform them in nervous impulses which are trasmitted to the central nervous system. In the most inner coat of the eyeball globe we find the retina, where, together with the different kind of nervous cells, there are receptors sensitive to light. These receptors with a peculiar shape are named *cones* and *rods*. The first are specialized in the day colour vision; the second are dedicated to the twilight, black and white vision.

The vision system we have designed firstly estimates on the image the likely position of the possibly interesting objects, and, in a second phase, processes the pixels surronding these points to recognize the nature of the objects and to detect their correct location. The idea of receptor is applied during the first processing stage. To determine approximately the location of the different objects, it is useless to work on all the pixels in the image. We have devised a receptor mask so that no object can fall between two adjacent receptors without influencing at least one of them (see figure **??**). This is possible, since we know a priori, in this semi-structured environment, the dimension of each object we are interested in.

During a setup phase the system we have implemented receives by the user the parameters concerning the mirror shape, the typical dimension of the desiderable objects and the colours to consider as reference. With these data, our system creates the receptor map which will be used for all the following processing activities. A "pre-scanning" of the image with receptors gives an effect similar to that we would have when looking at the image through a sheet of paper on which has been punched a matrix of holes: possibly, we are not able to identify the objects and their dimension, but we could recognize their colour and location. The "pre-scanning" based on the receptors is a kind of image sampling. Acting on the location of the receptors we can obtain a sort of low-filtering which allows us to discard the undesirable objects and increase the processing speed.

Each receptor is composed by a 3 per 3 pixel matrix and it is characterised by 3 colour parameters: qR, qG, qB. These are computed as follows:

$$qR = \frac{vmR}{vmI}; qG = \frac{vmG}{vmI}; qB = \frac{vmB}{vmI};$$

where vmR, vmG e vmB are the average values of the Red, Green and Blue components of all the 9 pixels; vmI is the average intensity of the receptor:

$$vmI = \frac{(vmR + vmG + vmB)}{3}.$$

The receptor identifies a colour if its 3 colour parameters are close to those defined for that colour. We have observed that carefully choosing the values of the sampling parameters, the system can recognize with a low error rate the different colours in different light conditions. In other terms, it is possible to performorm a reliable abstraction process, by classifying all the receptors as belonging to classes corresponding to the colors we are interested in. In figure 2 the classified receptors are drawn in black and white, with different shapes (dots, crosses, etc.).

After having classified the receptors, we aggregate them in structures named *target*. The creation of a target is possible through a sequential scanning of receptors, which become part of a target if they have expected characteristics of colour and position, exploiting blob-growing techniques. A target is a part of the image where it could be present one of the interesting objects. Onece identied targets, it is possible to operate on the part of image defined by each of them by adopting the classic image processing and recognition techniques. In these final stages of the processing the receptors are not used because they give too poor of information: we work on single pixels.

The technique adopted is blob growing. With this technique, we segment the target into groups of connected pixels. Each object silhouette forms a group, as does the background. In a second phase, we calculate the features of the groups (centre of area, moments of area), and, by matching these to reference parameters, we recognize the objects.

The application of these techniques on the image portions identifiied in the first phase, and not on the whole image gives a considerable slimming of the quantity of information to handle and increases the processing rate. It is evident that our receptor system has high benefits on the processing time.

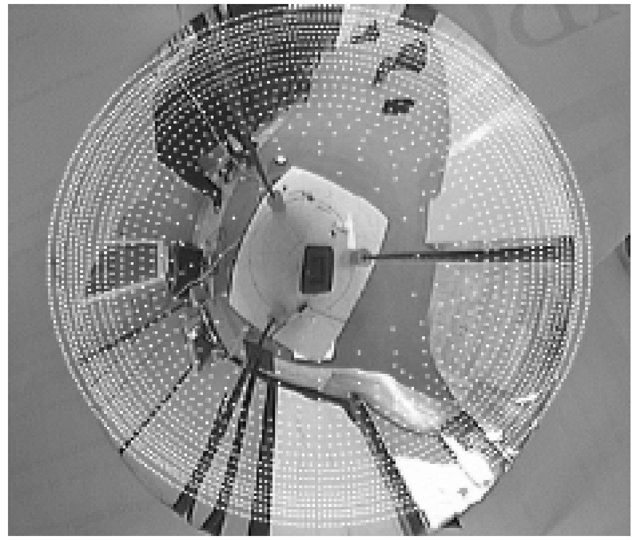

Figure 3. All the receptors for the RoboCup environment.

5 System performances

Working directly on an image of 640 x 640 pixel, there would be 409600 pixels to process. Employing a "pre-scanning" system with 2500 receptors of 9 pixels each, there are only 22500 pixels to process. The number of receptors to be adopted depends on different factors such as the shape of the mirror and the minimum dimension of the objects to be recognized. For an effective setting a careful study of the mirror shape and the distribution of the receptors is needed. We have adopted a circular disposition of the receptors: the mask we have employed is composed by 20 rims of receptors, as shown in figure **??**.

The distance between the receptors and between the rims of receptors is optimized by the setup program on the basis of the information about the dimensions and the distances of the object. The time needed to control these receptors and to create the target is about 1 ms , with about ten objects of different colours in the field. This result has been obtained by using the PC on the robot, based on 200 Mhz K6 cpu, 32 mb of ram, a Matrox Meteor frame grabber and working with Linux RedHad 5.1 (Kernel version 2.0.34). These performances leave the bound to the image processing rate to the acquisition rate of the image, done at 25 frame/sec. Moreover, since the image processing activity shares computation resources with control, planning and communication with other robots, the burden given by our algorithm is really light.

6 Conclusions

We have presented the hardware and software design of an omnidirectional sensor giving a mobile robot the possibility to track environmental elements present in any direction, with a frame rate higher than 20 frame/sec. The application that has pushed us to implement this sensor has been RoboCup, where the robot operates in a semi-structured environment with strong real-time constraints. Our system satisfies the requirements and will show its strength at the next RoboCup World Championship in Stockholm, August 1999.

The system we have implemented may be used also in other environments and for other tasks having similar requirements. The image processing approach is fast and reliable and may be succesfully adopted also for surveillance robots and any other applications where it is important to track markers of a given dimension and with characteristic colourings. The mirror shaping approach is also interesting to design revolutionary mirrors to match user requirements.

References

[1] M. Asada, P. Stone, H. Kitano, A. Drogoul, D. Duhaut, M. Veloso, H. Asama, and S. Suzuki. The robocup physical challenge: goals and protocols for phase i. In H. Kitano, editor, *RoboCup 97: Robot Soccer World Cup I*, pages 42–61. Springer-Verlag, Berlin, D, 1997.

[2] J. B. Gregersen. *A system for* 360° *vision in robot soccer.* PhD thesis, RMIT, Melbourne, AU, 1998.

[3] H. Kitano, editor. *RoboCup 98: Robot Soccer World Cup II*, Paris, F, 1998. Eurobot Project - CEC. Proceedings of the second RoboCup Workshop.

[4] E. Mouaddib and C. Pegard. Localization using omnidirectional vision. In *ICAR-95*, pages 133–138, Piscataway, NJ, 1995. IEEE Computer Press.

[5] D. Nardi, G. Clemente, and E. Pagello. Art azzurra robot team. In H. Kitano, editor, *RoboCup 98: Robot Soccer World Cup II*, pages 467–474, Paris, F, 1998. Eurobot Project - CEC.

[6] S. Suzuki, T. Kato, H. Ishizuka, Y. Takahashi, E. Uchibe, and M. Asada. An application of vision-based learning for a real robot in robocup - a goal keeping behavior for a robot with omnidirectional vision and an embedded servoing. In H. Kitano, editor, *RoboCup 98: Robot Soccer World Cup II*, pages 467–474. Eurobot Project - CEC, Paris, F, 1998.

[7] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional image sensor (copis) for vision-guided navigation. *IEEE Transactions on Robotics and Automation*, 10(1):11–22, 1994.