Journal of Harbin Institute of Technology (New Series), Vol. 11, No. 5, 2004

An object tracking and global localization method using the cylindrical projection of omnidirectional image

SUN Ying-jie[1], CAO Qi-xin[1], HONG Bing-rong[2]
孙英杰, 曹其新, 洪炳熔
(1. Robotics Research Institute, Shanghai Jiaotong University, Shanghai 200030, China;
2. School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

Abstract: We present an omnidirectional vision system we have implemented to provide our mobile robot with a fast tracking and robust localization capability. An algorithm is proposed to do reconstruction of the environment from the omnidirectional image and global localization of the robot in the context of the Middle Size League RoboCup field. This is accomplished by learning a set of visual landmarks such as the goals and the corner posts. Due to the dynamic changing environment and the partially observable landmarks, four localization cases are discussed in order to get robust localization performance. Localization is performed using a method that matches the observed landmarks, i.e. color blobs, which are extracted from the environment. The advantages of the cylindrical projection are discussed giving special consideration to the characteristics of the visual landmark and the meaning of the blob extraction. The analysis is established based on real time experiments with our omnidirectional vision system and the actual mobile robot. The comparative studies are presented and the feasibility of the method is shown.

Key words: omnidirectional vision system; cylindrical projection; object tracking; global localization; robust
CLC number: TP242.6 **Document code:** A **Article ID:** 1005-9113(2004)05-0474-07

In mobile robotics it is essential to the performance of many tasks, such as autonomous navigation and map exploration, to have complete information of the environment. For example, in order to navigate an autonomous mobile robot in an unknown environment, it is very useful to have sensors capable of seeing in all directions. Recently, there has been increased interest in omnidirectional vision for applications in autonomous mobile robotics[1,2].

Omnidirectional vision system covers a 360° field of view by analyzing only one image. This makes it possible to implement fast vision sensors suitable for a wide range of applications, such as: autonomous navigation, scene reconstruction and multi-robot cooperation[24]. Omnidirectional vision offers a number of significant benefits. Specifically, it is much easier to deal with rotation of the camera mounted on the robot because the objects will not disappear from the omnidirectional view.

Generally there are two kinds of localization methods: absolute localization and relative localization[5,6]. However, most of the proposed solutions do not take into account some real constraints encountered in practical applications and are not suitable for robust localization under a dynamic and partially observable environment. During the robot soccer game, the robot motion planning requires no human intervention. The real robot soccer field is dynamic and adversarial. Assuming that no collisions occur is unacceptable since the dead reckoning may generate unbounded error due to wheel slippage and collision. Since other robots may occlude landmarks during the soccer game so the robot must consider the collisions and occlusions in order to get real time robust performance. Therefore, active robust localization is necessary for the mobile robot in a partially observable and adverse environment.

One of the most important activities for a soccer robot is searching and tracking the ball while looking for goals and posts for self-localization and perceiving other teammates in order to interact with them and avoid opponents. For the Middle Size League RoboCup teams, global information is not available. Most of the robots had a single, fixed camera on board which was pointed forward. 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 purposefully, for instance advancing the ball or trying to reach a position, the vision direction (usually the heading of the robot) might not be optimal. For example, if the robot dribbles the ball, it has to check the presence of the ball in front of it as well as the presence of the goal or opponents in other

Received 2003-12-09
Sponsored by the National Natural Science Foundation of China (Grant No. 50128504 and 50390063).

directions. To address this problem, some teams mounted a camera on a pan-and-tilt system[7] to decouple the vision direction from the robot movement, but the pan-and-tilt system was relatively slow and did not allow effective tracking. The pan-and-tilt system just improved the performance of the normal perspective vision system. It did not achieve a wide field of view.

However, problems exist for the omnidirectional vision system. One of them is the shape distortion of the object in the panoramic image. Although the goals are rectangular in the Middle Size League RoboCup field, they are fan-shaped in the original image of the omnidirectional vision system. If we extract the blob of the fan-shaped object, i. e. the center, width, height and area of the object, it will result in a system error and a loss of many useful visual features.

In this paper we propose a robust global localization method for the soccer robots in the Middle Size League RoboCup field using an omnidirectional camera. First, we get the original omnidirectional image and transform it into a panoramic image using Cylindrical Projection. Secondly, we automatically or manually learn the RoboCup field landmarks from the training images and extract the color blob of the goals, the posts and the ball in HSL color space from the background of the rectangular image. Then we calibrate the omnidirectional vision system, track the ball, select the landmarks and localize the robot.

The paper is organized as follows: In Section 1 the omnidirectional vision system is described. The Cylindrical Projection and its advantages and disadvantages are presented in Section 2. Section 3 demonstrates the robust landmark extraction and global localization method. The validity is shown by real time experiments. Conclusions are presented in Section 4.

1 Omnidirectional Vision System

For the RoboCup Middle Size League, we have developed an autonomous mobile robot Jiaolong which is equipped with an omnidirectional camera as shown in Fig. 1. The hardware of the omnidirectional vision system consists of two major components: a mirror which is symmetrical on rotation and an apparatus which supports the mirror.

Technically, an omnidirectional vision system can be achieved in various ways. Generally, there are four types of mirrors: spherical mirror, conical mirror, hyperboloidal mirror and parabola mirror[8]. The omnidirectional system with the spherical mirror does not have a single center of projection and cannot be transformed into normal perspective images. The hyperboloidal mirror is best for an optic system using a normal CCD camera and the original image can be transformed to a normal perspective, cylindrical image. However, it is very difficult to design since the focal point of the hyperboloid needs to be set on the camera center. The system with a parabola mirror is very expensive and the size of a telecentric lens is not small.

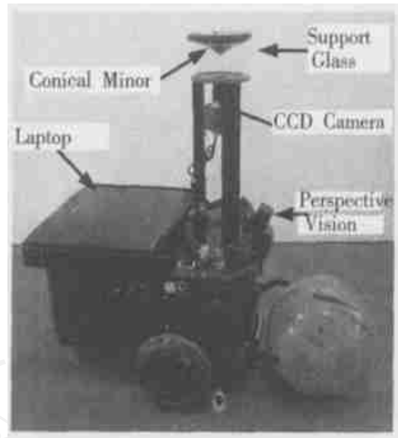


Fig. 1 Mobile robot with omnidirectional vision system

In this paper the omnidirectional vision system consists of catadioptric systems with conical mirrors and ordinary perspective cameras (as shown in Figs. 1 and 2). This omnidirectional vision system has the following advantages over the perspective vision systems:

1) The catadioptric systems capture the entire scene in a single instant, making them well-suited to analysis of dynamic environments.

2) The image resolution and distortion of the cone are favorable in comparison with other shapes of mirrors.

3) The conical mirror is relatively inexpensive and standard CCD cameras are widely available.

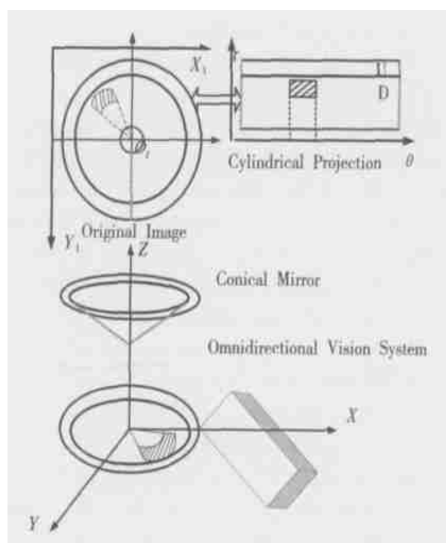


Fig. 2 Omnidirectional mirror and the image

2 Omnidirectional Image and Cylinder Projection

2.1 Omnidirectional Image

Robot control in a highly dynamic environment, e.g. the Middle Size League of RoboCup, is a real-time sensory data acquisition and processing task. The surrounding scene of the robot is rapidly changing and there is a need to continuously track the ball, the goals and the corner posts to accomplish the soccer robot's tasks such as global localization, map building and teammate cooperation.

The input image of our omnidirectional vision system is captured from a standard CCD camera which observes through the conical mirror. The camera is connected through a USB-PCI adapter with a laptop. The frame grab rate is 15 frames per second and the resolution is 320 × 240 pixels (24 bit true color) as shown in Fig. 3. The source image shows how the camera sees the robot body, other robots, the field, the goals, the ball and a part of the ceiling.

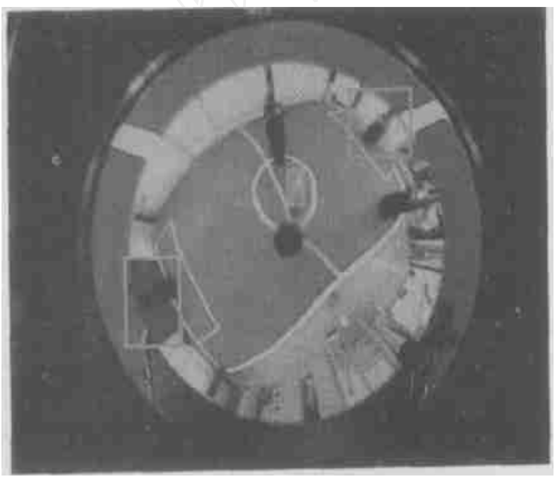


Fig. 3 Omnidirectional source image

2.2 Cylindrical Projection

To increase the robustness for occlusions and to gain additional rotational independence, we performed Cylinder Projection on the original source image. The omnidirectional source image is warped onto the cylinder. In Fig. 4 the cylindrical panoramic image contains only the scene, discarding the robot body and the ceiling. As shown in Figs. 3 and 4, the cylindrical projection which projects the images onto a cylindrical surface, yields significant data reduction. After the Cylinder Projection, the image data is reduced from 320 × 240 to 360 × 104 (48.75 % of the original image size) or from 640 × 480 to 360 × 210 (24.6 %).

For each omnidirectional vision system, we should first find the centre of each image with pixel accuracy, i.e. the coordinate (x_0, y_0) and the valid radius r_0 of the mirror center. Once we know the above parameters, we can transform the "circular image" (as shown in Fig. 3) polar coordinate representation $I(\theta_i, r_j)$ that is obtained directly from the usual Cartesian coordinates format $I_R(x_i, y_j)$ or $I_{CP}(\theta_i, r_j)$ where $\theta_i \in [-180, 180]$, $r_j \in [1, 104]$ (as shown in Fig. 4).

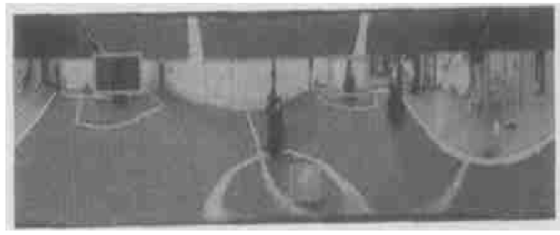


Fig. 4 Cylindrical projection of the source image

Considering the storage format of the image in the computer, the Cylinder Projection algorithm is as follows:

```
Algorithm: CylindricalProjection()
{
for i = 0; i < W_Dest
  for j = 0; j < H_Dest
    x_1 = j + r_0 · sin i
    y_1 = j + r_0 · cos i
    Dest[W_Dest · j + i] = Source[W(H - y_0 - y_1) + x_0 + x_1]
    Dest[W_Dest · j + i] = Source[W_Source(H_Source - y_0 - y_1) + x_0 + x_1]
  end
end
}
```

Source[·] and *Dest*[·] are the original source image and the rectangular image respectively. For our omnidirectional vision system, the parameters are $W_{\text{Source}} = 320$, $H_{\text{Source}} = 240$; $x_0 = 167$, $y_0 = 127$; $W_{\text{Dest}} = 360$, $H_{\text{Dest}} = 104$.

Fig. 5 shows the comparison of the blob area ratio K_{ball}, K_{blue}, K_{yellow} between the original image and the Cylinder Projection image K_{CP}.

$$K = \frac{A}{(B_{\text{right}} - B_{\text{left}}) \cdot (B_{\text{bottom}} - B_{\text{top}})},$$

where B_{left}, B_{right}, B_{top} and B_{bottom} describe the boundary of the blob and A is the blob area.

K_{CP} is mostly invariant to change in the heading θ_R of the robot when the robot rotates at a given position (X_R, Y_R). In Fig. 5 the parameters are normalized.

$$K\theta = 0.5 + \frac{\theta_R}{360}, \text{ where } \theta_R \in [-180, 180].$$

θ_R is calculated by the pose estimation method which is proposed in Section 4.2.1. In the experiment, we let the robot rotate at a given position $[v, \omega] = [0, 10°/s]$ where two goals are available. Fig. 5 shows that θ_R changes uniformly in [-180, 180], which proves that the proposed robot pose estimation method is very stable and precise.

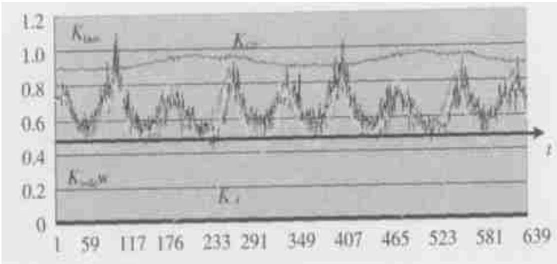


Fig. 5 Rotation dependence comparison

The omnidirectional image is stored in matrix format and it cannot directly and continually process the "circular image". It is easier for humans to view them in the rectangular format, i.e. perspective format. The most important concern is that all the landmarks in the RoboCup field are rectangular. In order to get valid blob information we cannot simply extract the blob in the original "circular image"; therefore, the blob in rectangular format is more meaningful for comparing with the original image (as shown in Figs. 3 and 4). Considering the partial occlusions of the visual landmarks and rotational differences between the robot's pose at the current position compared to the robot's pose at the next reference position of the training phase, we cannot use only the coarse center position of the goals to localize the robot in the field.

The advantages of Cylinder Projection are as follows: image data reduction, high resolution, rotation independent of self-localization and robust ball tracking. If we use the original omnidirectional image to localize the robot, the result is dependent upon the rotation (as shown in Fig. 5). In the original image the ball is a very small blob when it's distance to the robot is from 80 mm to 1 500 mm while within the cylindrical image the ball is big enough for blob extraction.

2.3 Property of Cylindrical Projection

The Cylinder Projection maps the sector/ellipse on the original source image to the rectangle/circle making it possible to extract the objects' geometrical features. For example, we can extract the edges of the objects directly from the blob. The Omnidirectional vision system can get precise azimuth information of the object and Cylinder Projection makes full use of this advantage of the panoramic image. It can get the precise angle of the goal's two edges and the corner posts.

We define the eigen area of the landmarks as follows:

$$E = r \cdot \delta\theta = \pi r \cdot (\theta_r - \theta_l)/180.$$

Where θ_l and θ_r ($\theta_r > \theta_l$) are the left and right edge of the blob in Fig. 4 respectively and r is the bottom of the blob. Generally E is constant for a given visual landmark. E is proportional to the width of the landmark especially for the cylindrical landmarks, e.g. the corner posts, which enable us to distinguish the corner posts and the goal very easily.

3 Global Localization

3.1 Blob Extraction and Omnidirectional Vision Calibration

In order to increase the robustness and reduce the computational burden, the object tracking algorithm is based on the blob extraction[7,9], which supplies the pixel coordinate of the center (X_c, Y_c), top left corner (X_l, Y_t) and bottom right corner (X_r, Y_b) of the blob (i.e. the width and height) and area A. This estimation model is used to map the image pixel $b^I(X_c, Y_c)$ (Pixels) to the field coordinates $b^C(X_b^C, Y_b^C)$ (mm).

The camera was calibrated by placing a ball at known distances in the field and reading the pixel value of the ball blob. The estimate can be used by the localization of the ball, the opponents and teammates. This mapping is learned by training the correspondence between pixel coordinates and Cartesian coordinates for a set of well-chosen positions and using Lagrangian interpolation for other pixels (as shown in Fig. 6). By keeping a history list of positions for all objects, the heading and velocities can be determined in order to predict the ball's position.

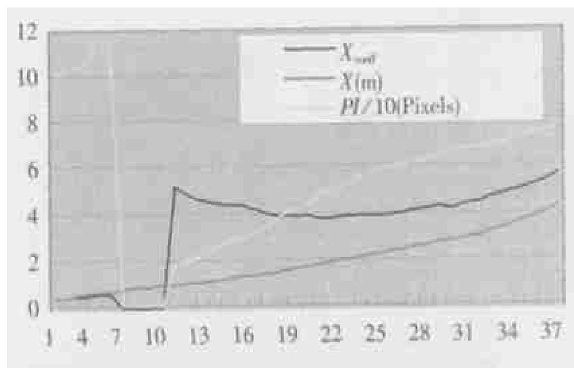


Fig. 6 Omnidirectional vision calibration

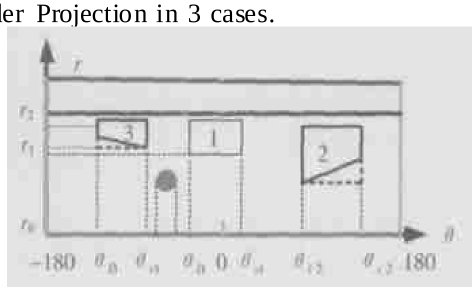
Our omnidirectional vision software permits interactive calibration of the omnidirectional vision system. The user simply specifies the center and the radius of the omnidirectional image using a graphical interface.

3.2 Global Localization of the Robot

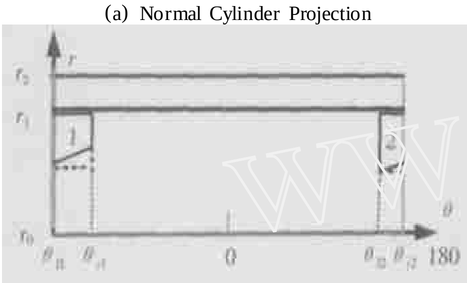
The robot pose is represented by $R(X_R, Y_R, \theta_R)$ in the working space $W: XOY$. The more precise and rich information the omnidirectional vision system provides, the more robust the localization will be. For a soccer robot in the Middle Size League RoboCup field there are a total of 6 landmarks (2 goals and 4 posts). However, at a given position, there are a maximum 4 landmarks available. During the game there are 8 robots on the field. Occlusion is inevitable for the soccer robot, especially for the forward player.

There are 3 cases of Cylinder Projection for the localization of the robot and ball tracking (as shown in Fig. 7). The drawback of Cylinder Projection is that it may divide a whole object in the source image into two blobs as shown in Fig. 7 (a) and (b). The two blobs must first be merged before ball tracking and self-localization are accomplished.

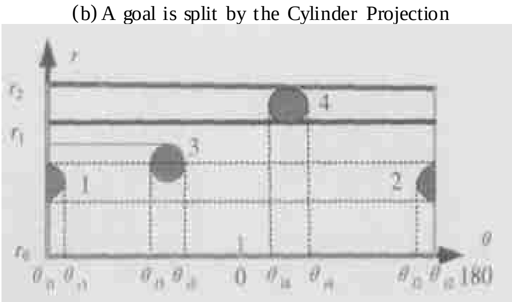
Fig. 8 demonstrates the experiment results of Cylinder Projection in 3 cases.



(a) Normal Cylinder Projection

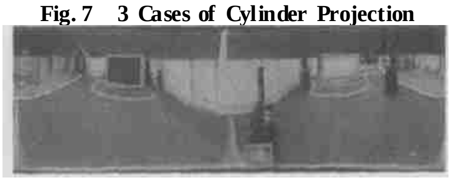


(b) A goal is split by the Cylinder Projection

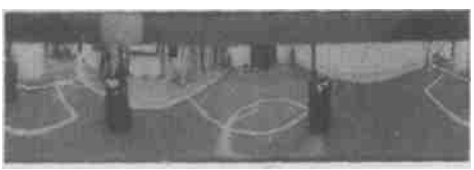


(c) The ball is split by the Cylinder Projection

Fig. 7 3 Cases of Cylinder Projection



(a) Normal result



(b) Split one goal is divided

(c) Ball is divided

Fig. 8 Experiment results of Cylinder Projection

3.2.1 Case I: both of the goals are available

If both of the goals are available for the robot, localization performance is most reliable and the localization algorithm is simple. The soccer robot with only a forward perspective camera or a pan-and-tilt camera cannot observe both of the goals through only one image. However, an omnidirectional vision system makes it possible to use two goals to localize the robot and track the ball in a wide field of view range.

In Fig. 9 the two goals are represented in the local robot coordinate and in the field coordinate by (Unit: cm):

LCS: OA: (a, θ_{mygoal}), OB: $(b, \theta_{oppgoal})$, $|AB| = c$

WCS: OA: $(a, -90 - \mathrm{sgn}(X_R) \cdot \alpha)$,

OB: $(b, 90 + \mathrm{sgn}(X_R) \cdot \beta)$, AB: $(100, 90)$.

The heading of the robot plays an important role in the localization and object tracking system because it decides how to interpret the omnidirectional vision data. Therefore, we should first calculate the orientation θ_R.

$$\delta\theta = \theta_{mygoal} - \theta_{oppgoal},$$

$$\mathrm{sgn}(X_R) = \begin{cases} 1 & |\ \delta\theta \in [-\infty, -180] \vee [180, \infty] \\ -1 & |\ \delta\theta \in [-180, 180] \end{cases},$$

where $\mathrm{sgn}(X_R)$ represents the sign of X_R. If $\mathrm{sgn}(X_R)$ equals 1, X_R is positive and vice versa.

$$\alpha = \cos^{-1}\left(\frac{a^2 + c^2 - b^2}{2ac}\right),$$

$$\beta = \cos^{-1}\left(\frac{c^2 + b^2 - a^2}{2bc}\right),$$

$$\theta_1 = 90 + \mathrm{sgn}(X_R) \cdot \beta - \theta_{oppgoal};$$

$$\theta_2 = 90 - \mathrm{sgn}(X_R) \cdot \alpha - \theta_{mygoal};$$

$$\theta_R = \frac{\theta_1 + \theta_2}{2}.$$

After determining the robot heading, the position X_R and Y_R can be calculated:

$$d = \frac{ab\sin\theta}{c},\quad X_R = \mathrm{sgn}(X_R)\, d,$$

$$YR = [100 - a \cdot \sin(\theta_{oppgoal} - \theta_R) + b \cdot \sin(\theta_R + \theta_{mygoal})]/2.$$

3.2.2 Case II: only one goal is available

If only one goal and one post are visible for the robot at a given position, we must first get the goal information $(\theta_l, \theta_r, r_{\min})$ i.e. $(\theta_a, \theta_b, r_a)$, and then we can get two candidate poses which are on the same circle in the World Coordinate System as shown in Figs. 10 and 11.

$$\delta\theta = |\theta_a - \theta_b|, \alpha_1 = 90 - \delta\theta,$$

$$|AC| = \frac{|AB|}{2 \cdot \sin\delta\theta}, \alpha_2 = \cos^{-1}\left(\frac{r_{\min}}{2|AC|}\right),$$

$$\theta_R = \begin{cases} \theta_a + \alpha_1 + \alpha_2 - 180, \text{Type I} \\ \theta_b + \alpha_1 + \alpha_2 - 180, \text{Type II.} \end{cases}$$

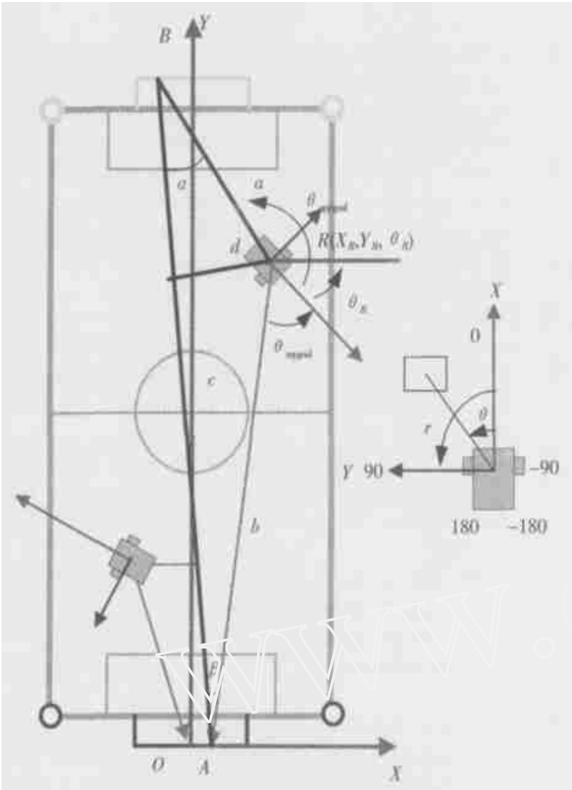


Fig. 9 Localization using two goals

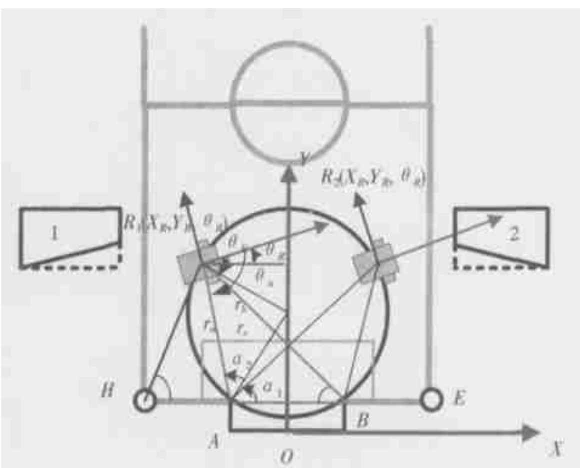


Fig. 10 Localization with only one goal

Then we select the unique pose of the robot from the type of the goal. If the shape of the goal is type I, then the robot pose is $R_1(X_R, Y_R, \theta_R)$, otherwise $R_2(X_R, Y_R, \theta_R)$.

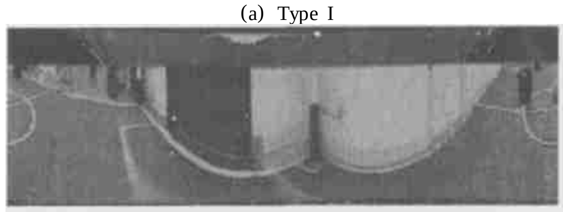
$$R = \begin{cases} \text{Type I}: R_1(X_R, Y_R, \theta_R) \text{ if } \dfrac{\theta_a + \theta_b}{2} < \theta_c \\ \text{Type II}: R_2(X_R, Y_R, \theta_R) \text{ if } \dfrac{\theta_a + \theta_b}{2} \geq \theta_c, \end{cases}$$

where θ_c is the center of the blob.

3.2.3 Case III: one goal and one post

Using A and H to localize the robot is similar but simpler than the above. Then we can fuse the pose calculate from A, H and from A, B together.

(a) Type I



(a) Type II

Fig. 11 Localization experiments with only one goal

3.2.4 Case IV: no visual landmark is available

If there are no visual landmarks available, we have to employ the odometer or ask the off field server for the global map. Actually, there are at most 8 robots in the field and this case is very infrequent or temporary. By recording the history list of the robot's pose and updating the odometer, we can get satisfactory results.

The global localization and object tracking method may be summarized as follows. First, we get the original omnidirectional image and perform Cylindrical Projection on it. Secondly, we train the omnidirectional system to learn the visual landmarks and extract the color blob of the ball, the goals and the posts in HSL color space. Thirdly, the split blobs are merged into one object and the omnidirectional vision system is calibrated. Finally, after the landmark's geometrical feature is extracted, we consider 4 localization cases which appear in real robot soccer games.

4 Conclusion

This paper deals with the essential problems of Middle Size League RoboCup and an analysis is established based on experiments which have practical application for real competitions. Absolute localization of robots in the RoboCup field using an omnidirectional camera is presented in this paper. First we discuss the advantages of the omnidirectional vision system. After learning the RoboCup field landmarks from training images, the omnidirectional vision extracts the reliable color blobs of the ball, the goals and the posts. This paper presents theoretical and experimental comparisons between the Cylindrical Projection and the original image. The advantages of Cylindrical Projection are obvious through the experiment result and analysis and it not only greatly improves the computing efficiency but improves the precision and robustness of object tracking and global localization for our Jiaolong Middle Size League robot soccer. Robustness and stability are most significant for the design of robot soccer and this paper has taken all the possible localization cases into consideration in order to get reliable localization performance.

In future work we will implement a distributed omnidirectional vision system using Multi-Sensor Data Fusion (MSDF) which is essential for Simultaneous Localization and Map Building (SLAM) and the cooperation of the whole robot soccer team.

Acknowledgement

The first author is grateful to Wang Lei and Zhang Zheming for their previous work.

References :

[1] DROCOURT C, DELAHOCHE L , PEGARD C , *et al*. Mobile robot localization based on an omnidirectional stereoscopic vision perception system[A]. IEEE International Conference on Robotics and Automation[C]. [s.l.]: [s. n.] , 1999. 1329 - 1334.

[2] BUNSCHOTEN R , KROSE B. Robust scene reconstruction from an omnidirectional vision system [J]. IEEE Transactions on Robotics and Automation , 2003 , 19 (2) : 351 - 357.

[3] MATSUMOTO Y, IKEDA K, INABA M , *et al*. Visual navigation using omnidirectional view sequence[A]. International Conference on Intelligent Robots and Systems[C]. [s.l.]:[s. n.] , 1999. 317 - 322.

[4] KATO K, ISHIGURO H, BARTH M. Identifying and localizing robots in a multi-robot system environment[A]. International Conference on Intelligent Robots and Systems [C]. [s.l.]:[s. n.] , 1999. 966 - 971.

[5] YATA T, OHYA A , YUTA S. Fusion of omni-directional sonar and omni-directional vision for environment recognition of mobile robots[A]. IEEE International Conference on Robotics and Automation [C]. [s. l.]: [s. n.] , 2000. 3925 - 3930.

[6] MARHIC B , MOUADDIB E M , PEGARD C. A localisation method with an omnidirectional vision sensor using projective invariant[A]. International Conference on Intelligent Robots and Systems [C]. [s. l.] : [s. n.] , 1998. 1078 - 1083.

[7] WEIGEL T, GUTMANN J S, DIETL M , *et al*. CS Freiburg: coordinating robots for successful soccer playing [J]. IEEE Transactions on Robotics and Automation , 2002 , 18(5) : 685 - 699.

[8] ISHIGURO H. Development of low-cost compact omnidirectional vision sensors and their applications[A]. Proceedings of Int Conference on Information Systems , Analysis and Synthesis[C]. [s.l.]:[s. n.] , 1998. 433 - 439.

[9] BRUCE J , BALCH T, VELOSO M. Fast and inexpensive color image segmentation for interactive robots[J]. International Conference on Intelligent Robots and Systems , 2000 ,3 : 2061 - 2066.