Proceedings of the 5th World Congress on Intelligent Control
and Automation, June 15-19, 2004, Hangzhou, P.R. China

第五届全球智能控制与自动化大会
2004 年 6 月 15 - 19 日, 中国杭州

An Object Tracking and Global Localization Method using Omnidirectional Vision System*

Yingjie Sun and Qixin Cao
Research Institute of Robotics
Shanghai Jiao Tong University
Shanghai, 200030, China
{yjsun & qxcao}@sjtu.edu.cn

Weidong Chen
Institute of Automation
Shanghai Jiao Tong University
Shanghai, 200030, China
wdchen@sjtu.edu.cn

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 partial observable of the 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 especially considering the characteristic of the visual landmark and the meaning of the blob extraction. The analysis is established on real time experiment with our omnidirectional vision system and the actual mobile robot, the comparative studies are presented and show the feasibility of the method.**

Index Terms - Omnidirectional vision system; Cylindrical projection; Object tracking; Global localization.

I. INTRODUCTION

In mobile robotics it is essential for 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 is an increased interest in omnidirectional vision for applications in autonomous mobile robotics [1, 2, 3].

Omnidirectional vision system covers a 360° field of view by analysing 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 [4, 5, 6]. Omnidirectional vision offers a number of significant benefits. Specifically, it is much easier to deal with rotation of the camera which is mounted on the robot, because the objects will not disappear from the omnidirectional view.

One of the most important activities for a soccer robot is searching and tracking the ball, but also looking for goals and posts for self-localization, and perceive other team mates to interact with them and avoid the opponents. For the Middle Size League RoboCup teams, the global information is not available. Most of the robots had a single, fixed camera on board, 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 bringing the ball or trying to reach a position, the vision system orientation (usually the heading of the robot) might not be optimal. For instance, if the robot dribbles the ball, it has to check the presence of the ball in front of it, but also the presence of the goal, or of opponents. To deal 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 is relatively slow and does not allow effective tracking. The pan and tilt system just improve the performance of the normal perspective vision system, although it does not achieve wide field of view.

Omnidirectional vision system is prevailing in Middle Size League RoboCup [8, 9, 10, 11] due to its wide field of view. However there still exist some problems for the omnidirectional vision system, and one of them is the shape distortion of the object in the panoramic image. Although the goals are rectangle in 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, the width, the height and area of the object, it will cause system error and lost many useful visual features.

In this paper we propose a robust global localization method for the soccer robots in Middle Size League RoboCup field using an omnidirectional vision system. 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 rectangle 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 II the omnidirectional vision system is described. The Cylindrical Projection, its advantages and disadvantages are presented in Section III. Section IV demonstrates the robust landmark

* This work is partially supported by NNSF of China Grant #50128504 (Young Oversea Investigator Award) to Qixin Cao.

extraction and global localization method. The validity of the proposed method is shown by real time experiments. Comparative experiment is shown in Section V. The conclusions are present in Section VI.

II. Omnidirectional Vision System

For the Middle Size League RoboCup, we have developed 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, 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 [12]. 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 the best one for optic system using normal CCD camera, and the original image can be transformed to normal perspective image, cylindrical image. However it is very difficult to design since the focal point of the hyperboloid need to be set on the camera center. The system with parabola mirror is very expensive and the size of telecentric lens is not small.

In this paper the omnidirectional vision system consists of catadioptric systems with conical mirrors and ordinary perspective cameras (see Fig. 1 and Fig. 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.

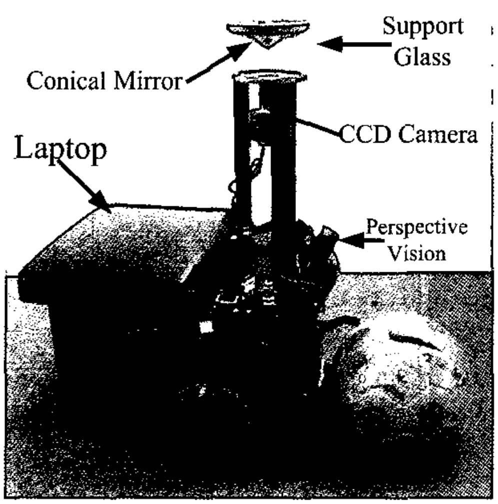


Fig. 1 The mobile robot with omnidirectional vision system

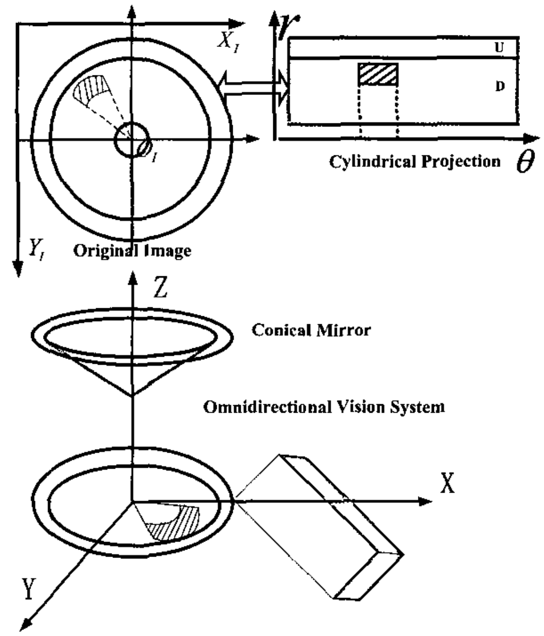


Fig. 2 Principle of the omnidirectional mirror and the image

III. Omnidirectional Image and Cylinder Projection

A. 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 regarding the robot soccer's task such as global localization, map building and teammates cooperation. 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 partial observable and adverse environment.

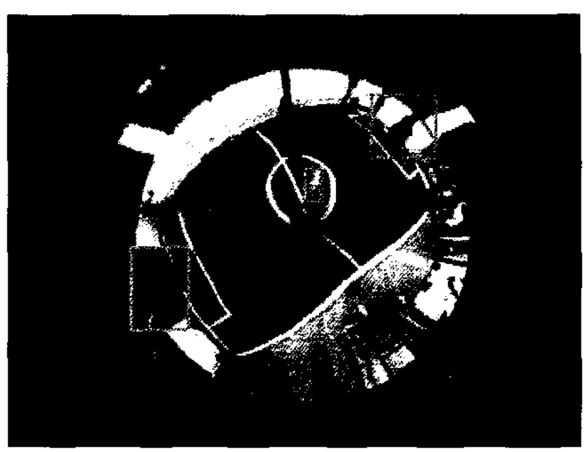


Fig. 3 The omnidirectional source image

The input image of our omnidirectional vision system is captured from standard CCD camera which observes the conical mirror. The camera is connected through USB adapter with the 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.

B. Cylindrical Projection

To increase the robustness for occlusions and to additionally gain rotational independence, we performed Cylinder Projection on the original source image. The omnidirectional source image is warped onto the cylinder. In Fig. 4 cylindrical panoramic image contains only the scene, discarding the robot body and the ceiling. As shown in Fig. 3 and Fig. 4 the cylindrical projection which projects the images onto a cylindrical surface, yielding a significant data reduction. After the Cylinder Projection the image data is reduced from 320x240 to 360x104 (48.75% of the original image size) or from 640x480 to 360x210(24.6%).

Fig. 4 Cylindrical projection of the source image

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" (see Fig. 3) polar coordinate representation $I(\theta_i, r_j)$ that they are obtained in directly to the usual Cartesian coordinates format $I_R(x_i, y_j)$ or $I_{CP}(\theta_i, r_j)$ where $\theta_i \in [-180, 180]$, $r_j \in [1, 104]$ (see Fig. 4).

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

```
Algorithm: CylindricalProjection()
{
for i = 0; i < W_Dest
  for j = 0; j < H_Dest
    x1 = j + r0 · sin i
    y1 = j + r0 · cos i
    Dest[W_Dest · j + i] = Source[W(H − y0 − y1) + x0 + x1]
    Dest[W_Dest · j + i] = Source[W_Source(H_Source − y0 − y1) + x0 + x1]
  end
end
}
```

Fig. 5 Cylindrical projection algorithm

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

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 RoboCup field are rectangle. In order to get valid blob information we cannot simply extract the blob in the original "circular image", therefore the blob in rectangle format is more meaningful comparing with the original image (See Fig. 3 and Fig. 4). Considering the partial occlusions of the visual landmarks and orientation 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 using only the coarse blob center position of the goals in the original image to localize the robot in the field.

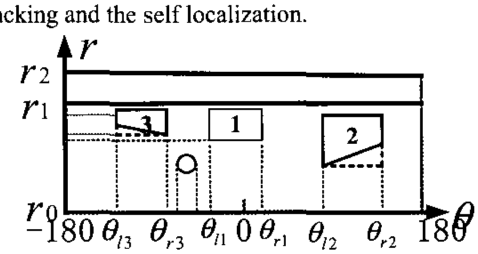
C. Advantages of Cylindrical Projection

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

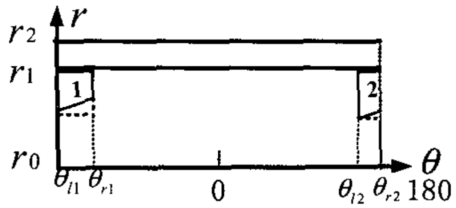
The advantages of the Cylinder Projection are as follows: image data reduction, high resolution, rotation independent for self localization and robust ball tracking. If we use the original omnidirectional image to localize the robot the result is dependent with the rotation. In the original image the ball is a very small blob when it's distance to the robot is from 80mm to 1500mm, while in the cylindrical image the ball is big enough for blob extraction.

D. Disadvantages of Cylindrical Projection

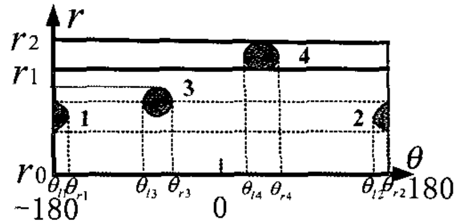
There are 3 cases for the localization of the robot and the ball tracking after Cylinder Projection (See Fig. 6). The drawback of Cylinder Projection is that it may split a whole object in the source image into two blobs as shown in Fig. 6 (b) and (c), and we have to merge the two blobs first before the ball tracking and the self localization.



(a) Normal Cylinder Projection



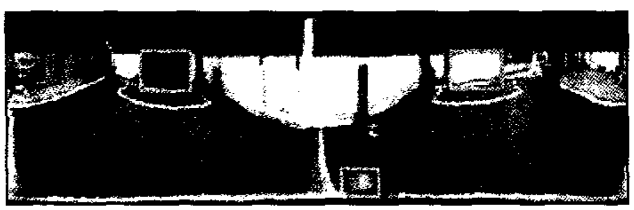
(b) Goal is split by the Cylinder Projection.



(c) Ball is split by the Cylinder Projection.

Fig. 6 Three Cases of Cylinder Projection

Fig. 7 demonstrates the experiment results of Cylinder Projection in 3 cases.



(a) Normal Result

(b) Split goal

(c) Split Ball

Fig. 7 Experiment Results of Cylinder Projection

IV. Global Localization

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

The object tracking algorithm is based on the blob extraction, which supply 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, 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). A fast and inexpensive color image segmentation method for blob extraction is proposed in [13].

A. Case 1: both of the goals are available

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

In Fig. 8 the two goals are represented in the local robot coordinate and in the field coordinate by: (Unit: cm)

$$LCS: \overrightarrow{OA}:(a, \theta_{mygoal}),\ \overrightarrow{OB}:(b, \theta_{oppgoal}),\ |\overrightarrow{AB}|=c \quad (1)$$

$$WCS: \overrightarrow{OA}:(a, -90-\mathrm{sgn}(X_R)\cdot\alpha), \quad (2)$$
$$\overrightarrow{OB}:(b, 90+\mathrm{sgn}(X_R)\cdot\beta),\ \overrightarrow{AB}:(100,90)$$

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.

$$\delta\theta = \theta_{mygoal} - \theta_{oppgoal}, \quad (3)$$

$$\mathrm{sgn}(X_R) = \{ \begin{matrix} 1 \mid \delta\theta \in [-\infty, -180] \vee [180, \infty] \\ -1 \mid \delta\theta \in [-180, 180] \end{matrix} \}, \quad (4)$$

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.

$$\alpha = \cos^{-1}\left(\frac{a^2+c^2-b^2}{2ac}\right), \quad (5)$$

$$\beta = \cos^{-1}\left(\frac{c^2+b^2-a^2}{2bc}\right), \quad (6)$$

$$\theta_1 = 90 + \mathrm{sgn}(X_R)\cdot\beta - \theta_{oppgoal}; \quad (7)$$

$$\theta_2 = 90 - \mathrm{sgn}(X_R)\cdot\alpha - \theta_{mygoal}; \quad (8)$$

$$\theta_R = \frac{\theta_1+\theta_2}{2}. \quad (9)$$

Once we get the robot heading, the position X_R and Y_R could be calculated:

$$d = \frac{ab\sin\theta}{c}, \tag{10}$$

$$X_R = \mathrm{sgn}(X_R)\cdot d, \tag{11}$$

$$Y_R = \frac{100 - a\cdot\sin(\theta_{oppgoal} - \theta_R) + b\cdot\sin(\theta_R + \theta_{mygoal})}{2} \tag{12}$$

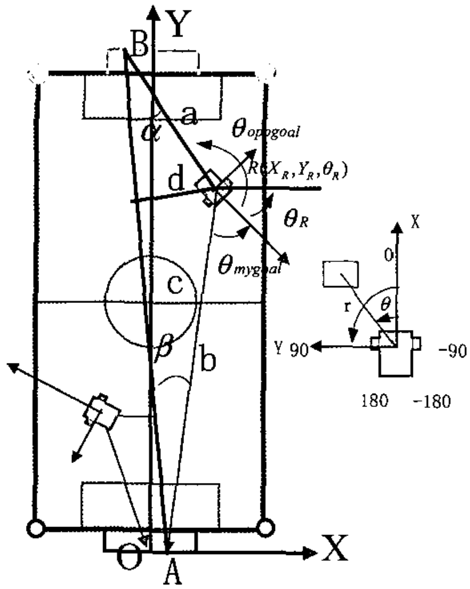


Fig. 8 Localization using two goals

B. Case II: only one goal is available

If only one goal and one post are visible for the robot at a given position. First we 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 Fig. 9.

$$\delta\theta = \theta_a - \theta_b, \tag{13}$$

$$\alpha_1 = 90 - \delta\theta, \tag{14}$$

$$|AC| = \frac{|AB|}{2\cdot\sin\delta\theta}, \tag{15}$$

$$\alpha_2 = \cos^{-1}\left(\frac{r_{\min}}{2|AC|}\right), \tag{16}$$

$$\theta_R = \begin{cases} \theta_a + \alpha_1 + \alpha_2 - 180, \text{Type I} \\ \theta_b + \alpha_1 + \alpha_2 - 180, \text{Type II} \end{cases} \tag{17}$$

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 $R_2(X_R, Y_R, \theta_R)$.

$$R = \begin{cases} \text{Type I: } R_1(X_R, Y_R, \theta_R) \text{ if } \frac{\theta_a + \theta_b}{2} < \theta_c \\ \text{Type II: } R_2(X_R, Y_R, \theta_R) \text{ if } \frac{\theta_a + \theta_b}{2} \geq \theta_c \end{cases} \tag{18}$$

where θ_c is the center of the blob.

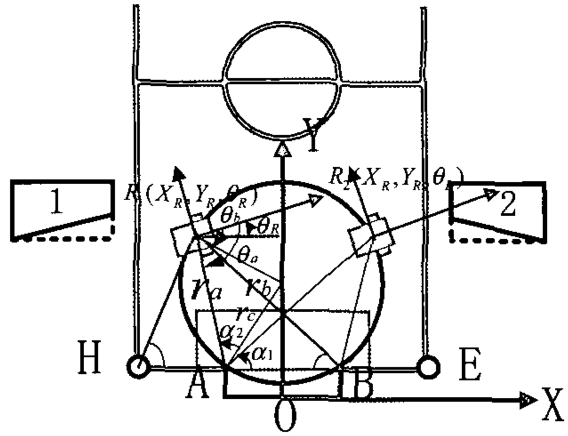


Fig. 9 Localization with only one goal

(a) Type I

(a) Type II

Fig. 10 Localization experiments with only one goal

C. 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.

D. Case IV: no visual landmark is available

If there is no visual landmarks available, we have to employ the odometer or rely on the off field server for the global map and global pose update. Actually there are at most 8 robots in the field, and this case is very occasional or temporal. We record the history list of the robot's pose and update the odometer can get satisfied result.

We summarize the global localization and object tracking method 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. Then merge the split blobs into one object and calibrate the omnidirectional vision system. After the landmark's geometrical feature is extracted, we consider 4 localization cases which will appear in real robot soccer games.

V. Comparative Experiment

In order to validate the advantages of Cylindrical Projection using in RoboCup, Fig. 11 shows the comparison of blob area ratio $K_{ball}, K_{Blue}, K_{yellow}$ between the original image and the Cylinder Projection image K_{CP}.

K is defined as follows:

$$K = \frac{A}{(B_{right} - B_{left}) \cdot (B_{bottom} - B_{top})}. \qquad (19)$$

where B_{left}, B_{right}, B_{top} and B_{bottom} describe the boundary of the blob, 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. 11 the parameters are normalized.

$$K_\theta = 0.5 + \frac{\theta_R}{360^\circ}. \qquad (20)$$

where $\theta_R \in [-180, 180]$. θ_R is calculated by the pose estimation method which is proposed above. In the experiment we let the robot rotate at a given position with $[v, \omega] = [0mm/s, 10^\circ/s]$, where two goals are available. Fig. 11 shows that θ_R changes uniformly in $[-180, 180]$, which proves that the proposed robot pose estimation method is very stable and precise.

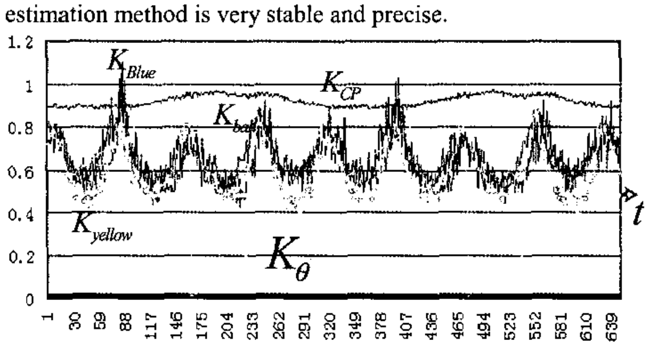


Fig. 11 Rotation dependence comparison

VI. Conclusion

This paper deals with the essential problems of Middle Size League RoboCup, and the analysis is established on experiments which have practical sense for the real competitions. Absolute localization of robots in the RoboCup field using an omnidirectional vision system is presented in this paper. First we discuss the advantages as well as the disadvantages 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. It not only improves the computing efficiency greatly but improves the precision and the robustness of object tracking and global localization for our Jiaolong Middle Size League robot soccer. The robustness and stability are the most significant for the design of robot soccer, and this paper has taken four possible localization cases into consideration in order to get reliable localization performance.

In the future work we will implement distributed omnidirectional vision system using multi sensor data fusion (MSDF) which is essential for the Simultaneous Localization and Map Building (SLAM) and the cooperation of the whole robot soccer team.

Acknowledgment

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

References

[1] C. Drocourt; L. Delahoche, C. Pegard, and A. Clerenti, "Mobile robot localization based on an omnidirectional stereoscopic vision perception system," IEEE International Conference on Robotics and Automation, vol. 2, pp.1329 -1334, 1999.

[2] R. Bunschoten, and B. Krose, "Robust scene reconstruction from an omnidirectional vision system," IEEE Transactions on Robotics and Automation, vol. 19, no. 2, pp. 351 -357, 2003.

[3] Y. Matsumoto, K. Ikeda, M. Inaba, and H. Inoue, "Visual navigation using omnidirectional view sequence," International Conference on Intelligent Robots and Systems, vol. 1, pp.317 -322, 1999.

[4] K. Kato, H. Ishiguro, and M. Barth, "Identifying and localizing robots in a multi-robot system environment," International Conference on Intelligent Robots and Systems, vol. 2, pp. 966 -971, 1999.

[5] T. Yata, A. Ohya, and S. Yuta, "Fusion of omni-directional sonar and omni-directional vision for environment recognition of mobile robots," IEEE International Conference on Robotics and Automation, vol. 4, pp. 3925 -3930, April 2000.

[6] B. Marhic, El.M. Mouaddib, and C.Pegard, "A localisation method with an omnidirectional vision sensor using projective invariant," International Conference on Intelligent Robots and Systems, vol. 2, pp. 1078 -1083, 1998.

[7] T. Weigel, J.-S. Gutmann, M. Dietl, A. Kleiner, and B. Nebel, "CS Freiburg: coordinating robots for successful soccer playing," IEEE Transactions on Robotics and Automation, vol. 18, no. 5, pp. 685 -699, 2002.

[8] Takeshi Matsuoka, and Motoki Katoh, "Description of Team 'Fusion'," RobCup Symposium: Robot Soccer World Cup VII, 2003.

[9] Kazuhiro Furuta, Shingo Takehana, Takatugu Matumura, etc, "WinKIT 2003 Team Description," RobCup Symposium: Robot Soccer World Cup VII, 2003.

[10] Yasutake Takahashi, Kazuhiro Edazawa, etc. "Osaka University 'Trackies 2003'," RobCup Symposium: Robot Soccer World Cup VII, 2003.

[11] Alireza Fadaei Tehrani, Hamid Reza Moballegh, and Peiman Amini, "Team Description: Persia 2003," RobCup Symposium: Robot Soccer World Cup VII, 2003.

[12] H. Ishiguro, "Development of low-cost compact omnidirectional vision sensors and their applications," In proceedings of Int. Conference on Information Systems, Analysis and Synthesis, pp. 433-439, 1998.

[13] J. Bruce, T. Balch, and M. Veloso, "Fast and inexpensive color image segmentation for interactive robots," International Conference on Intelligent Robots and Systems, vol.3, pp. 2061 – 2066, Nov. 2000.