The Goalkeeper Strategy of RoboCup MSL Based on Dual Image Source

Xueyan Wang[1,*], Peng Zhao[2,*], Shiyu Mou[1], Di Zhu[7], Jieming Zhou[8], Dongbiao Sun[1],Song Chen[3,*], Zhe Zhu[4,*], Ye Tian[5], Zongyi Zhang[5], Lv Ye[6], Xinxin Xu[6], Wanjie Zhang[6],Kang Wu[6], Binbin Li[5], Yuan Li[1], Abudouyimujiang Kader[1], Yiheng Su[1], Chenyu Wang[6]

[1] Beijing Information Science & Technology University, No.12, East Qinghexiaoying Rd, 100192, China.
[2] NO.21 Research Institute of China Electronics Technology Group Corp, 200233, Shanghai.
[3] Beijing Aerospace Guanghua Elctronics Technologies Limited Corporation,100854, Beijing.
[4] China Telecom, 100032, Beijing.
[5] Shanghai Yingchuang. Ltd, 201100, Shanghai.
[6] Beijing Waterplus. Ltd, 100000, China.
[7] University of Science & Technology Beijing, 100083, Beijing.
[8] Austalian National University, 00120C, Canberra.
* Corresponding author:
Xueyan Wang, 342429193@qq.com
Song Chen, 25127601@qq.com
Zhe Zhu, luke.zhezhu@yahoo.com
Peng Zhao, peng.zhao@ia.ac.cn

Abstract. Gatekeeping is one of the most important tasks in MSL, especially defending aerial shots. This paper proposes a goalkeeper strategy based on combining two image sources. The goalkeeper uses the Omni-vision system to adjust heading orientation of the Kinect sensor for capturing both depth and RGB images of the ball. The depth information is used to detect the spatial position of the ball and the RGB image ensures the validity of the recognition. Once opponents shoot, the proposed system uses the first few detected positions of the ball to predict the interception point by the goalkeeper based on the least square method. The accuracy and effectiveness of the proposed strategy are verified by experimentation.

Keywords: Goalkeeper strategy, Dual image source, Kinect sensor, Depth camera, RGB camera, Defensive position.

1   Introduction

RoboCup Soccer has been one of the competitions of RoboCup ever since. The rules used in the middle size league (MSL) are gradually approaching the rules of human soccer competition of FIFA. The standard ball has been introduced into the middle size league and the field has been extended to 12m*18m. The current technical level in MSL makes it a very exciting competition in RoboCup, attracting large audiences and even investors given its high commercial potential [1,2,3]. Moreover, MSL covers several scientific and technological domains from computer vision [4] to localization and navigation [5], multi-agents cooperation [6], communications, etc, making it an important research platform[1]. Team Water from Beijing Information Science &Technology University is a regular team in RoboCup MSL and has performed well in the competition in recent years [7]. Based on the Omni-vision system and local map matching technology, the team Water keeps improving the

precision and robustness of robot self-localization, navigation and detection of targets and obstacles. Last year, the Microsoft Kinect depth camera [8] was introduced to improve the performance of objects detection, which played an important role for goalkeeping in RoboCup2015.

The goalkeeper ability to prevent the opponent team to score is called interception. The performance of the interception behavior has a significant impact on the competition. Based on the ball coordinates got from the robot's Omni-vision system, two main strategies are used to intercept the ball. One is keeping the goalkeeper standing on the line connecting the ball to the center of the goal. Another is fitting the ball trajectory according to the records of recent past ball coordinates and then predicting the point of interception. Both methods perform well when the ball rolls on the floor. Facing high balls, however, these strategies fail due to the distortion introduced by the convex mirror of the Omni-vision system and the impossibility to fit the flying trajectory causing the goalkeeper to miss the actual interception point.

To improve the ball interception with aerial shots we propose a dual image source strategy for the goalkeeper. We use the omnidirectional mirror image and the RGB primary color and depth images from Kinect. Then we reduce the dimension of the ball trajectory to that of a top view straight line using the least squares method and then we predict the closest position to intercept the ball.

The structure of this paper is as follows. The next section introduces the structure of the goalkeeper. Section 3 establishes the coordinates conversion. Section 4 details the strategy of the dual image sources. Section 5 shows the experiments designed to test the strategy proposed in Section 4. Section 6 presents the conclusion.

2   Structure of the Goalkeeper Robot

Figure 1 shows the detection system of goalkeeper, which includes an Omni-vision system, an IMU sensor and a Kinect. The Omni-vision system collects the image from all around the robot using a convex Omni-mirror facing down, projecting the 360º image on a vertically mounted camera in the same axis. The Kinect sensor is installed on the front of the robot. This sensor has an RGB camera and a depth camera. The RGB camera can collect the color image and the depth camera can acquire the depth information associated to each pixel.
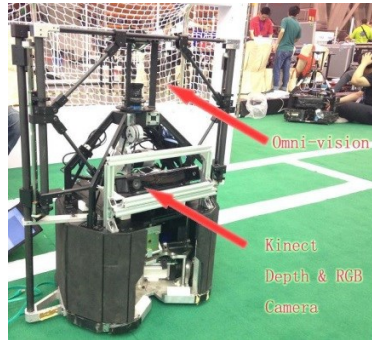
**Fig. 1.** The goalkeeper physical structure.

As shown in Fig 2, the robot system mainly includes a laptop, a control board and servo drivers. The laptop receives the commands such as "Start", "Pause" and "Stop" from the Referee Box based on the TCP/IP protocol through the laptop wireless LAN connection. Then, combining the information of the Omni-vision system, the Kinect and an electronic compass, the robot can locate itself using a local matching method, detect the ball and distinguish barriers. Based on this information, we compute the speed and target coordinates for the robot so that it intercepts the ball and we send such set-points to the robot control board. This board converts the received commands to speed set-points for the wheels motor servo drives. In addition, the power unit consists of Li-batteries and a voltage monitor module for power management.
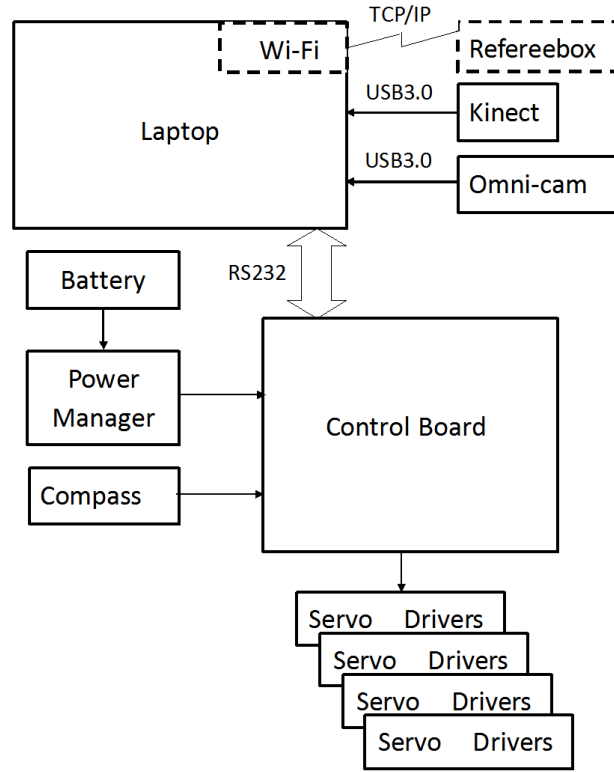
Fig. 2. The architecture of robot control system.

3   Coordinates Transformation

This section presents the notation used and the conversion between different coordinate systems, namely the World $\{W\}$, the Robot $\{R\}$ and the Kinect $\{K\}$ coordinate systems (Fig. 3). The world coordinate system $\{W\}$ takes the center of the field as the origin, the X-axis is parallel to the field midline and pointing left when facing the opponent's half field. The Y-axis is perpendicular to the midline and points to the goal. The Z-axis is perpendicular to the ground and pointing upward.
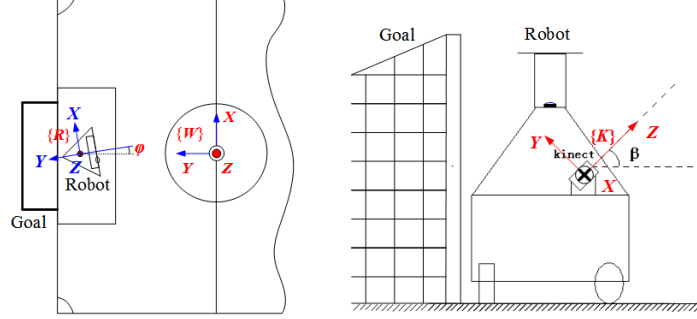
Fig. 3. The World Coordinate System $\{W\}$, Robot Coordinate System $\{R\}$ and the Kinect Coordinate Systems $\{K\}$.

The robot coordinate system $\{R\}$ has the origin on the center of the projection of the robot on the field. The axis toward the right side of the robot, when facing it, is the X-axis. The axis toward the back of the robot is the Y-axis. The Z-axis points upwards, vertically. Based on the position of the field lines using the Omni-vision system and a map matching method, the robot gets its location information in the world coordinate system $\{W\}$. The location information includes the robot coordinates $(x_R, y_R, 0)$ and heading angle φ. The conversion from robot coordinates in $\{R\}$ to the world coordinate system $\{W\}$ is achieved with the transformation matrix ^WT_R (Eq. 1).

$$^WT_R = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 & x_R \\ \sin\varphi & \cos\varphi & 0 & y_R \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

The RGB and depth cameras of Kinect share the same lens. Hence, they share the coordinate system $\{K\}$ whose origin corresponds to the camera focal point, the X-axis is parallel to the ground and points to the right side of the robot, when facing it, the Z-axis points outwards along the lens axis and the Y-axis follows the right-hand rule. The conversion from the Kinect coordinate system $\{K\}$ to robot coordinates $\{R\}$ is achieved with the transformation matrix ^RT_K (Eq. 2) where (x_K, y_K, h) is the position of the Kinect origin in the robot coordinate system $\{R\}$ and β represents its tilt angle with respect to the horizontal plane.

$$^RT_K = \begin{bmatrix} 1 & 0 & 0 & x_K \\ 0 & \sin\beta & -\cos\beta & y_K \\ 0 & \cos\beta & \sin\beta & h \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

Therefore, a generic point $^K P_i$ ($^K x_i$, $^K y_i$, $^K z_i$) in the Kinect system is represented in the world coordinate system with $^W P_i$ (x_i, y_i, z_i) as given by Eq. 3.

$$^W P_i = \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} = {}^W T_R * {}^R T_K * \begin{bmatrix} ^K x_i \\ ^K y_i \\ ^K z_i \\ 1 \end{bmatrix} \tag{3}$$

4   Goalkeeper Strategy Based on Dual Image Source

Following the dual image source strategy the goalkeeper keeps the ball within the Kinect field of view using the images captured by the Omni-vision system. The ball 3D coordinates in $\{W\}$ are then obtained from the depth information captured by the Kinect while the RGB information is used to remove false identifications. Once the goalkeeper finds that the ball is flying towards to the goal, it fits its trajectory projection on the ground into a straight line, determines the intersection of that line with the goal line and moves to that defensive position.

4.1   Adjust Goalkeeper Heading with Omni-vision System

The image acquired by the Omni-vision system is an ordinary RGB image, thus the pixels of the ball in the image can be conventionally recognized by the threshold segmentation and the region-growing method [9]. Firstly, the threshold range parameters of the ball in HSV space are calibrated by the calibration toolbox. Based on the range parameters, the image is processed with the banalization method. Then, the scattered pixels in the binary image are clustered together into different connected regions by the region-growing method. The biggest connected region in the image is taken as the ball.
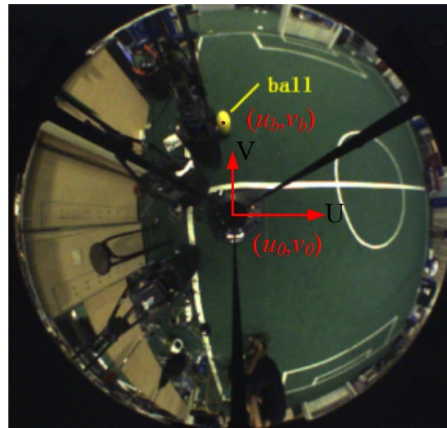

Fig.4. The ball in the image captured by the Omni-vision system.

Fig. 4 shows an image captured by the Omni-vision system. The U-axis points to the right of the goalkeeper while the V-axis points towards the front. The central pixel of the ball image (u_b, v_b) is used to get the angle of the ball relative to the robot (θ_b), as described by Eq. 4, where (u_0, v_0) is the fiducial point of the Omni-vision system camera. Knowing θ_b the goalkeeper can adjust its orientation to track the ball angle, thus keeping it within the field of view of the Kinect.

$$\theta_b = \tan^{-1} \frac{v_b - v_0}{u_b - u_0} \qquad (4)$$

4.2 Recognize the Ball's Imaging Blob in the Depth Image

The pixels value ^Hz in the depth image represent the distance from the camera to the target surface along the Z-axis of the Kinect coordinate system. However, the ball is a regular small sphere resulting in different depth values along its surface that are close to each other. Thus, these ball pixels in the depth image can be taken as a blob with similar depth value. The whole ball detection process is described in Fig. 5.
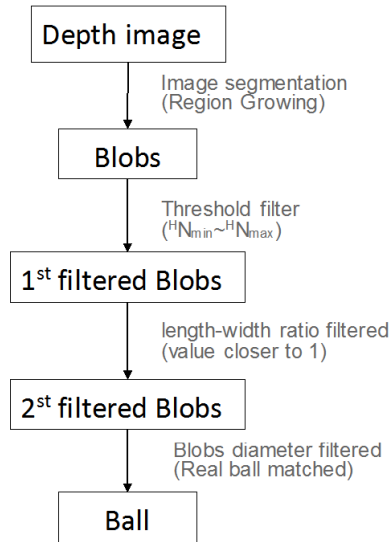
Fig.5. The ball detection process with depth information.

Firstly, the blobs from the depth image with different depth values are segmented. Then, the blobs of the wall and surrounding robots that are oversize are filtered out. At the same time, the blobs of image noise that are undersize are also removed. The rest of the blobs are referred as available blobs.

Secondly, the blobs whose ratio between length and width is close to 1 are extracted. Thirdly, using Eq. 5 we estimate the corresponding diameters in the world coordinates system (D_b) and we take the blob that has a diameter closer to that of the real ball. Note that $\overline{^HZ_b}$ is the blob mean depth value, ^HD_b is the diameter of the blob and ^Hk is the amplification factor of the depth camera.

$$D_b = \overline{^HZ_b} * \frac{^HD_b}{^Hk} \tag{5}$$

4.3 Check the Imaging Blob and Measure the Ball's Spatial Location

The result of the detection process described in the previous section is the blob barycenter pixel coordinates (^Hu_b, ^Hv_b). Based on these coordinates, the corresponding RGB image is used to check whether the surrounding pixels belong to the calibrated threshold range in HSV space. A match means that the detection is valid and the ball coordinates in the Kinect coordinate system ^KP_b can be computed with Eq. 6. Note that ($^Hu_0$, $^Hv_0$) is the fiducial point of depth camera. Finally, the ball position in the world coordinate system (^WP_b) can be determined using Eq. 3.

$$^KP_b = \begin{bmatrix} ^Kx_b \\ ^Ky_b \\ ^Kz_b \end{bmatrix} = \overline{^HZ_b} \begin{bmatrix} ^Hk & 0 & ^Hu_0 \\ 0 & ^Hk & ^Hv_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ^Hu_b \\ ^Hv_b \\ 1 \end{bmatrix} \tag{6}$$

4.4 Determining the Goalkeeper Defensive Position

Once the goalkeeper detects that the distance from the ball to the goal is decreasing fast, which means the ball is flying towards to the goal, it tracks the ball trajectory with just the first two dimensions of the 3D coordinates ^WP_b (its ground projection or top view) [8]. Then, it applies the Least Squares Method to the ground projected ball path with n (x_i, y_i) points to derive a straight line. The line parameters {a, b} are computed as in Eq. 7 where $\bar{x}$ and $\bar{y}$ are the average values of the respective coordinates.

$$\begin{cases} b = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} \\ a = \bar{y} - b\bar{x} \end{cases} \tag{7}$$

Finally, we determine the intersection (x_d, y_d) between the linearized ground projected ball path and the goal line using Eq. 8, where L is the length of the field. This intersection is our defensive point to where we drive the goalkeeper.

$$\begin{cases} y_d = \frac{L}{2} \\ x_d = \frac{L/2 - a}{b} \end{cases} \tag{8}$$

4.5 Two Measurement Methods Based on Single Image Source

To show the advantages of the proposed dual image based approach, we define here the two single image source based methods which we use for comparison. The first method uses the RGB camera of the Kinect, only. It recognizes the ball region in the image based on calibrated threshold range in HSV color space and estimates the distance between the ball and camera according to the size of the region. The spatial position is calculated with the camera pin-point model. The second method uses the

depth camera of the Kinect, only, which essentially does the recognition process described in Section 4.3 without cross checking with RGB information.

5   Experiments

In order to verify the performance of our proposed strategy we carried out several experiments that are described next. Firstly, the goalkeeper is placed in a fixed location. Then, the ball is thrown at the goal from three different distances with 20 times at each spot. For each spot we also determine the straight line between the throwing point and the landing point, and we record the intersection between this line and the goal line as the desired defensive point. At the same time, we also determine the defensive point using the Kinect system according to the strategy proposed in Section 4. Finally, we record the errors between the two defensive points, the observed and the computed one.
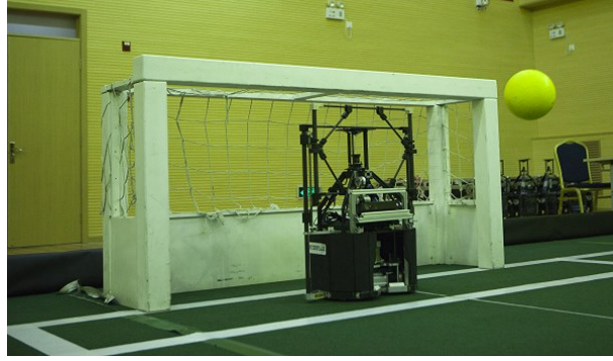
Fig.6. The experimental scenario

The above experimental process was repeated for the dual image source method (RGB-D) as well as for the methods that separately use RGB and depth images. The statistics of the measured errors (average / standard deviation) for these three methods are shown in Table 1, for the three different distances.

Table 1. Error statistics for 3 methods (cm)

Method	Distance		
	2.5m	4.5m	6.5m
RGB-D	10.42/5.88	17.65/12.81	38.74/10.37
RGB	27.83/14.37	42.50/27.94	112.07/82.75
Depth	13.42/6.75	18.65/13.09	28.47/7.279

In order to ensure the image quality captured by the Kinect, the goalkeeper will remain still for a while once the ball leaps up from the field. Additionally, the goalkeeper needs time to move to the interception point, thus, it will not observe the ball for a long time. Therefore, in the experiments the goalkeeper captures only a few frames of information from the Kinect to compute the defensive point for every shot.

The experimental results show that the proposed dual image source method is better than the other two methods for short distances. For longer distances, just using depth seems to yield good results. Conversely, the defensive points measured by the RGB camera alone show significantly larger errors resulting in a bad performance in handling aerial shots. In general, as expected, the results also show that the error increases as the shooting distance increases.

6 Conclusion

This paper proposed a goalkeeper strategy based on dual image sources combining the Omni-vision system and the Kinect sensor. It uses the Omni-vision system to get orientation of the ball and direct the Kinect to it. The Kinect utilizes the depth information to measure the spatial positions of the ball and the RGB image to ensure the validity of the recognition. The Least Square Method is applied to the ground projection of the ball path to derive a straight line whose intersection with the goal line determines the defensive position of the goalkeeper. The experimental results show the accuracy and effectiveness of this strategy.

Acknowledgment

References

1. Min T, Shuo W, Zhiqiang C.: Multi-robot Systems. Beijing:Tsinghua University Press (2005)
2. Shi L, Xuming X, Qin Y, Zengqi S.: International Robot Soccer Tournament and Correlative Technique. Robot 05 (5):420-426 (2000)
3. Stone P.: Layered Learning in Multi-agent Systems: A Winning Approach to Robotic Soccer. MIT Press (2000)
4. Fei L.: Research on Key Issues of Color Omni-directional Vision for Soccer Robots. Doctor of Engineering in Control Science and Engineering, National University of Defense Technology (2007)
5. Yonghui H.: A Method for the Middle - sized Mobile Robot's Self - location Based on Mark Lines in RoboCup Enviroment. Microcomputer Application (10):71-75 (2011)
6. Fangyi S.: Research on Behavior Control of Soccer Robot with Reinforcement Learning. Master's Degree, National University of Defense Technology(2008)
7. Chen S, Zhu Z, Tian Y, Lynch C, Zhu D, Lv Y, Xu X, Zhang W, Zhao L, Liu X, Wang M, Zhao P, Wang C, Liu W, Yang F, Li B, Zhou J, Zhang Z, Wang X.: Team Water: The Champion of the RoboCup Middle Size League Competition 2013. In: Behnke S, Veloso M, Visser A, Xiong R (eds) RoboCup 2013: Robot World Cup

XVII, vol 8371. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp 37-48. doi:10.1007/978-3-662-44468-9_4 (2014)

8. Ribeiro Neves AJ, Trifan A, Dias P, Azevedo JL.: Detection of Aerial Balls in Robotic Soccer Using a Mixture of Color and Depth Information. In: Autonomous Robot Systems and Competitions (ICARSC), 2015 IEEE International Conference on, 8-10 April 2015. pp 227-232. doi:10.1109/ICARSC.2015.13 (2015)

9. Huimin L.: The Research on the Omnidirectional Vision System for Autonomous Mobile Robots. Doctor of Engineering in Control Science and Engineering, National University of Defense Technology (2010)