

Object Motion Estimation Based on Hybrid Vision for Soccer Robots in 3D Space

Huimin Lu^(✉), Qinghua Yu, Dan Xiong, Junhao Xiao,
and Zhiqiang Zheng

College of Mechatronics and Automation,
National University of Defense Technology, Changsha, China
{lhmnew,xiongdan,zqzheng}@nudt.edu.cn, yuqinghua@163.com,
junhao.xiao@ieee.org

Abstract. Effective object motion estimation is significant to improve the performance of soccer robots in RoboCup Middle Size League. In this paper, a hybrid vision system is constructed by combining omnidirectional vision and stereo vision for ball recognition and motion estimation in three-dimensional (3D) space. When the ball is located on the ground field, a novel algorithm based on RANSAC and Kalman filter is proposed to estimate the ball velocity using the omnidirectional vision. When the ball is kicked up into the air, an iterative and coarse-to-fine method is proposed to fit the moving trace of the ball with parabolic curve and predict the touchdown-point in 3D space using the stereo vision. Experimental results show that the robot can effectively estimate ball motion in 3D space using the hybrid vision system and the proposed algorithms, furthermore, the advantages of the 360° field of view of the omnidirectional vision and the high object localization accuracy of the stereo vision in 3D space can be combined.

1 Introduction

In the highly dynamic RoboCup Middle Size League (MSL) competition, accurate estimation of object motion states, such as the velocity and the shooting touchdown-point of the ball, is the basis of ball passing and intercepting for regular robots, and shoot defending for goalie robots. Furthermore, because the ball is often lifted by the robots' high kicks during MSL competition, the ball motion should be estimated in three-dimensional (3D) space to improve the performance of soccer robots.

In [1], a Kalman Filter was used to detect whether the ball is moving or stationary. In the corrector part of the Kalman Filter, a multilayer perceptron artificial neural network was integrated to reduce the affection of image noises caused by the motion vibration of the robot, so the robustness of the state detection could be improved. In [2], Lauer et al. assumed that, during a small piece of time, the motion of a ball rolling on the field is a linear movement with a constant velocity. Therefore, the ball velocity estimation could be modelled as a standard linear regression problem, which could be solved by ridge regression — a least squares matching method (LSM). In [3], Kalman Filter was employed for

ball position estimation; based on the estimated positions, an algorithm similar as that in [2] was utilized to evaluate the ball velocity, resulting improved ball velocity estimation accuracy.

However, the methods mentioned above can only be used when the ball is located on the ground field. In [4, 5], Taiana et al. applied particle filter to track the ball using omnidirectional vision in 3D space, where the 3D shape of the ball was considered and the colour histograms of the inner and outer boundary on the panoramic image projected by the ball were used to construct the observation model in particle filter. The experimental results show that it can precisely track the ball and acquire the ball position in 3D space, but other motion states like velocity were not estimated.

In [6–8], a mixed stereo camera sensor was constructed based on omnidirectional vision and perspective camera, which was employed to recognize and localize the ball in 3D space. As a result, the advantages of the 360° field of view of the omnidirectional vision and the long field of view of the perspective camera can be combined. In [6, 7], triangulation was used to calculate the ball position in 3D space. In [8], Käppeler et al. found that better results could be achieved when calculating the 3D ball position using the angle to the ball determined by the omnidirectional vision and the distance to the ball derived from the size in the image of the perspective camera in comparison with the triangulation method, but they did not discuss ball velocity estimation in 3D space. In [6], Voigtländer et al. extended the approach in [2] from 2D space to 3D space to estimate the ball velocity. In [7], Lauer et al. developed a maximum likelihood estimator based on the ECM approach and a Bayesian approach based on Gibbs sampling to estimate the ball position and velocity in 3D space. However, the field of view of this kind of mixed stereo camera sensor is quite limited for motion estimation in 3D space. Particularly, when the ball is kicked up higher than the omnidirectional vision, the omnidirectional vision can not work any more.

Tech United Eindhoven team used a Laser Range Finder (LRF) attached to the highest point of the goalie robot to detect lob balls, when the ball disappears out of the view of the omnidirectional vision [9]. After the ball is detected, the LRF is tilted further upwards with a servo to measure additional positions of the ball in the air. With these points, the complete path of the ball can be calculated using a parabolic fit. Recently, they use the RGB-D camera Kinect to detect and track the ball in 3D space [10]. The main drawbacks are limitations in resolution and field of view, and only balls within approximately six meters can be recognized, which is not enough for goalie robots to intercept lob balls successfully.

In this paper, we present a hybrid vision system combining omnidirectional vision and stereo vision, which is employed to recognize the ball and estimate the ball motion states in 3D space including the velocity and the shooting touchdown-point. The following sections are organized as follows: the system overview is introduced in Sect. 2; an algorithm to estimate the ball velocity based on RANSAC and Kalman filter is proposed in Sect. 3, using the omnidirectional vision to deal with the situation that the ball is located on the ground

field; an object motion estimation algorithm including moving trace fitting and touchdown point prediction is presented in Sect. 4, which deals with the situation that the ball is kicked up into the air using the stereo vision; experimental results are presented in Sect. 5; Sect. 6 concludes this paper.

2 The System Overview

Omnidirectional vision can provide a 360° view of the robot's surrounding environment in a single image, thus is quite suitable for ball recognition and motion estimation when the ball is located on the ground field [11]. In the mixed stereo vision proposed in [6–8], the overlapped field of view of omnidirectional vision and perspective camera is quite small. Especially when the ball is higher than the robot, it is beyond the field of view of the stereo vision. Furthermore, since the imaging resolution of the omnidirectional vision is limited, the accuracy of ball localization can not be high in 3D space. Therefore, it is hard to deal with object motion estimation in 3D space well with such a mixed stereo vision. In this paper, we add Bumblebee2, a stereo vision developed by Point Grey Research, to construct the hybrid vision system for our soccer robots, as shown in Fig. 1. Typical images acquired by the omnidirectional vision and stereo vision, and the corresponding ball recognition results are also illustrated.

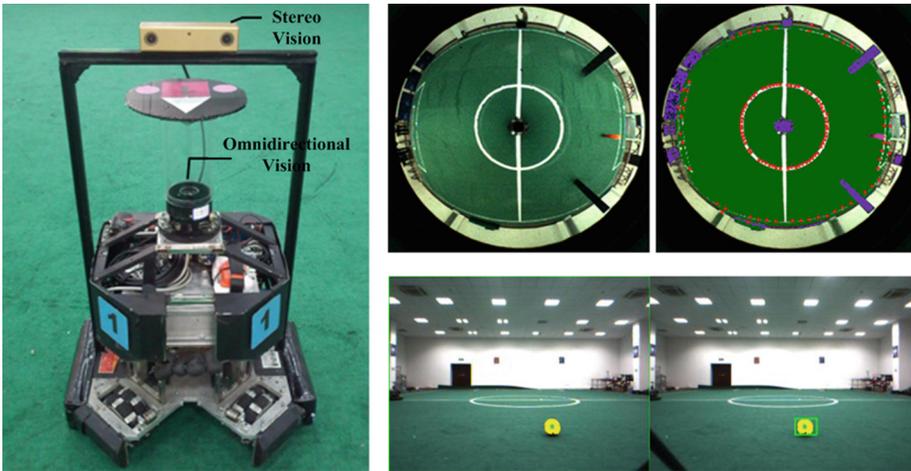


Fig. 1. The NuBot soccer robot equipped with the hybrid vision system constructed with omnidirectional vision and stereo vision. Typical images acquired by each vision and the ball recognition results are also shown.

The working architecture of the hybrid vision system is depicted in Fig. 2. Firstly, the omnidirectional vision searches the ball in the 360° view. If the ball is detected, the robot will turn to the ball. Afterwards, the stereo vision is employed

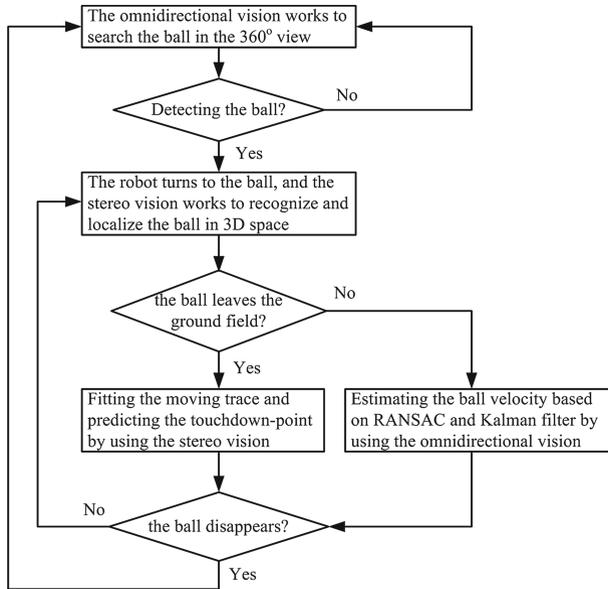


Fig. 2. The working architecture of the hybrid vision system to estimate the ball motion in 3D space.

to recognize and localize the ball in 3D space based on triangulation, by which the robot can determine whether the ball leaves the ground field. If the ball is located on the ground field, an algorithm based on RANSAC and Kalman filter is used to estimate the ball velocity with the omnidirectional vision. If the ball is kicked up into the air, a ball motion estimation algorithm is used to fit the moving trace and predict the touchdown-point where the robot should move to intercept the ball using the stereo vision. When the ball disappears from the view of the robot, the robot has to search the ball again using the omnidirectional vision. In such an architecture, the advantages of the 360° field of view of the omnidirectional vision and the high object localization accuracy of the stereo vision in 3D space can be combined.

3 Ball Velocity Estimation Based on RANSAC and Kalman Filter

When the ball is located on the ground field, omnidirectional vision is the best choice for ball recognition and motion estimation. In the highly dynamic RoboCup MSL competition, outliers and noises always exist in the information of ball positions due to the limited imaging resolution of omnidirectional vision, image noises, occlusions, motion blurs, etc., which will decrease the ball velocity estimation accuracy. RANdom SAMple Consensus (RANSAC) is an iterative method to estimate parameters of a mathematical model from a set of observed

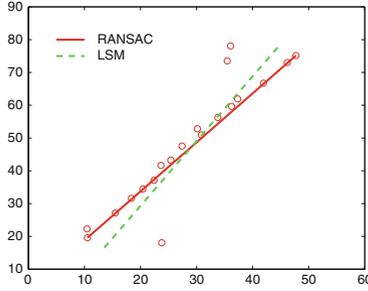


Fig. 3. Fitting points with outliers into a line by using RANSAC and LSM.

data which contains outliers [12]. A typical example to fit a group of points with outliers into a line is shown in Fig. 3, using RANSAC and least squares matching (LSM), which demonstrates that better robustness to outliers can be achieved by RANSAC. Therefore, in this paper, we apply RANSAC for ball velocity estimation when the ball is located on the ground field.

We assume that the ball velocity is constant between cycles in a short time such as several hundred milliseconds, therefore we use RANSAC to estimate the ball velocity with ball positions X_i and corresponding timestamp $t_i, i = 1, \dots, n$. We can acquire the velocity value V_m as follows:

$$V_m = \frac{X_i - X_j}{t_i - t_j}, (i \neq j, i, j = 1, \dots, n, m = 1, \dots, n(n-1)/2) \quad (1)$$

Then we can choose k values from $\{V_m, m = 1, \dots, n(n-1)/2\}$ randomly as hypothetical inliers, and calculate the mean value $M_l = \sum_{i=1}^k V_i/k$ as a candidate model with a counter $C_l = 0$. All the other velocity values are tested against the model by comparing the distance between the value and the model. If the distance is less than a threshold, $C_l = C_l + 1$, and the value is added into hypothetical inliers as the consensus set. Once this test finishes, a better model M_l is updated by calculating the mean value from the consensus set. After performing this operation iteratively by L times, $\{M_l, C_l, l = 1, \dots, L\}$ can be achieved. The M_l with the largest C_l is considered as the estimated ball velocity.

To further improve the robustness and accuracy, we also use Kalman filter to optimize the ball positions before estimating the ball velocity similar as that in [3]. As the ball velocity is assumed to be constant between cycles in a short time, we use the prediction model as follows:

$$\begin{pmatrix} X_{k+1} \\ V_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_k \\ V_k \end{pmatrix} \quad (2)$$

where X_k is the ball position, and V_k is the ball velocity. To model the measurement variance in the Kalman filter, we place the ball on the field with different distances to the robot, and then the robot recognizes and localizes the ball using the omnidirectional vision. As a result, a group of measurement variances is

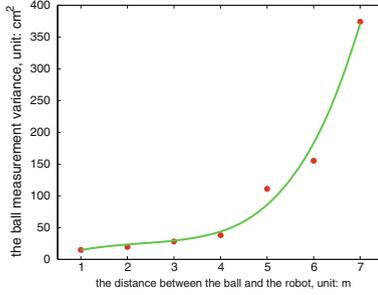


Fig. 4. The ball measurement variance when the ball is located on different distances to the robot.

acquired shown as the red points in Fig. 4. Afterwards, we fit these variances by using a four order polynomial, shown as the green curve in Fig. 4. The result is also useful in the simulation experiments in Sect. 5.1.

With Kalman filter, ball collision with other objects can also be detected, which happens frequently during the competition. We consider a collision happens when the distance between the ball measurement and the filtered result is larger than a threshold in five consecutive cycles. In that case, the Kalman filter will be restarted, and the ball measurements of the latest five cycles will be used as the initial data for ball velocity estimation.

4 Fitting the Moving Trace and Predicting of the Touchdown-Point in 3D Space

4.1 Fitting the Moving Trace in 3D Space

After obtaining the 3D position of the ball by triangulation using stereo vision, the robot can determine whether the ball leaves the ground field. If the 3D positions are higher than the ground field by 10 cm in consecutive three cycles, the ball is considered to be kicked up into the air. We use parabolic curve to fit the moving trace of the ball. The ball motion can be modelled as follows:

$$\begin{cases} x_i = a_0 * t_i + a_1 \\ y_i = a_2 * t_i + a_3 \\ z_i + g * t_i^2 / 2 = a_4 * t_i + a_5 \end{cases} \quad (3)$$

where x_i , y_i and z_i are 3D coordinates obtained by stereo vision in the timestamp t_i , and g is the gravity acceleration. The three equations are simple and linear. When a group of $\{x_i, y_i, z_i, t_i\}$ is obtained, a_0, \dots, a_5 can be calculated by the least squares method in three coordinates respectively. Therefore, the moving trace of the ball can be acquired using the parabolic fit.

4.2 Predicting of the Touchdown-Point

After fitting the moving trace of the ball, the robot can predict the touchdown-point where the ball falls on the ground field or the ball passes through the goal plane. The touchdown-point can be calculated easily by intersecting the moving trace with the ground field and the goal plane. Then the robot can use the touchdown-point and the ball velocity a_0 , a_2 to decide how to intercept or defend the ball.

To speed up the robot's response to the ball, the touchdown-point should be predicted as early as possible, requiring the ball positions used in fitting and predicting should be as few as possible. However, to improve the accuracy of ball intercepting or defending, more ball positions should be used in fitting and predicting to acquire more accurate prediction results. As a compromise, we use an iterative method to deal with this contradiction. Once the robot obtains five ball positions in the air, the first fitting and prediction is performed to acquire a coarse result of the predicted touchdown-point, thus the robot can respond to the ball very quickly. Then the robot will go on fitting and predicting with more data after obtaining new ball positions to update the prediction results. So the robot can achieve coarse-to-fine fitting and prediction results iteratively during ball intercepting or defending.

5 Experimental Results

In this section, we evaluate the two algorithms proposed in Sect. 3 and Sect. 4 respectively.

5.1 Ball Velocity Estimation Results by Omnidirectional Vision

Because no ground truth about ball positions and velocities can be provided in the actual experiments, we firstly perform simulation experiments to evaluate three algorithms: the proposed Kalman filter+RANSAC, Kalman filter+LSM, and Kalman filter. We add Gaussian noises into the simulated ball positions according to the variances in Fig. 4. Three different situations are considered including no collision, one collision, and multiple collisions of the ball. The ball positions and the estimated ball velocities in one experiment are shown in Fig. 5. We perform ten such experiments, and the statistics are shown in Table 1, where \bar{E} is the mean estimation error, and \bar{P} the mean ratio between the error and the real velocity. From Table 1, we see that when no collision happens, all the algorithms work well, and Kalman filter is the best algorithm to estimate the ball velocity, because the added noises are Gaussian in this simulation experiment. When collision happens, the motion direction of the ball changes, and better performance can be achieved by using Kalman filter+RANSAC and Kalman filter+LSM than Kalman filter. Furthermore, in comparison with Kalman filter+LSM, the estimation accuracy can be improved by 15%~40% when using Kalman filter+RANSAC.

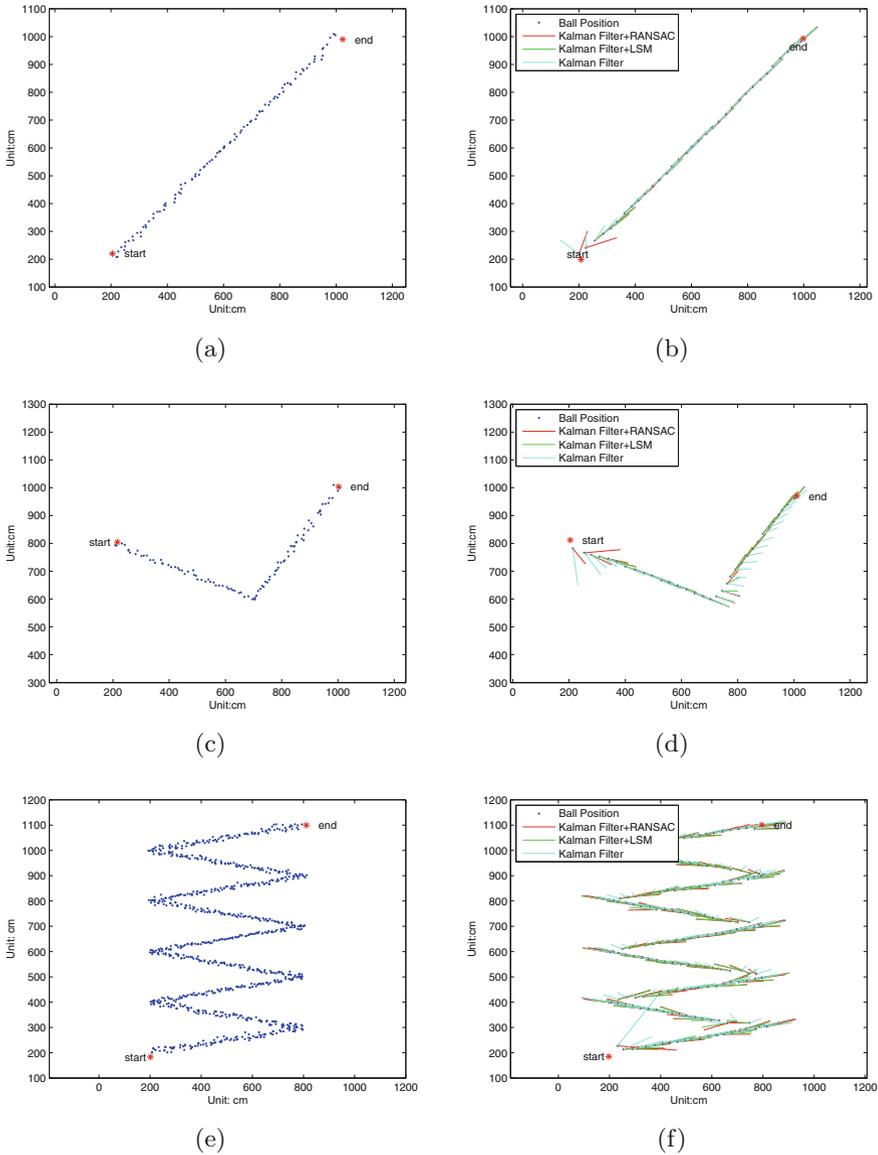


Fig. 5. The ball positions (a)(c)(e) and the ball velocities estimated by three different algorithms (b)(d)(f) in simulation experiments.

We also test Kalman filter+RANSAC and Kalman filter+LSM in the actual experiments using our NuBot soccer robot. The ball velocity estimation results are shown in Fig. 6 when the robot is stationary or moving. Because no ground truth can be provided, we only can say both algorithms work well. A video

Table 1. The statistics about ball velocity estimation by three different algorithms in simulation experiments. The unit of \bar{E} is cm/s.

	No collision		One collision		Multi-collision	
	\bar{E}	\bar{P}	\bar{E}	\bar{P}	\bar{E}	\bar{P}
KalmanFilter+LSM	10.84	4.01 %	32.78	9.47 %	27.22	11.68 %
KalmanFilter+RANSAC	10.97	4.05 %	20.20	5.83 %	22.64	9.710 %
Kalman Filter	4.780	1.80 %	91.86	26.6 %	59.50	25.70 %

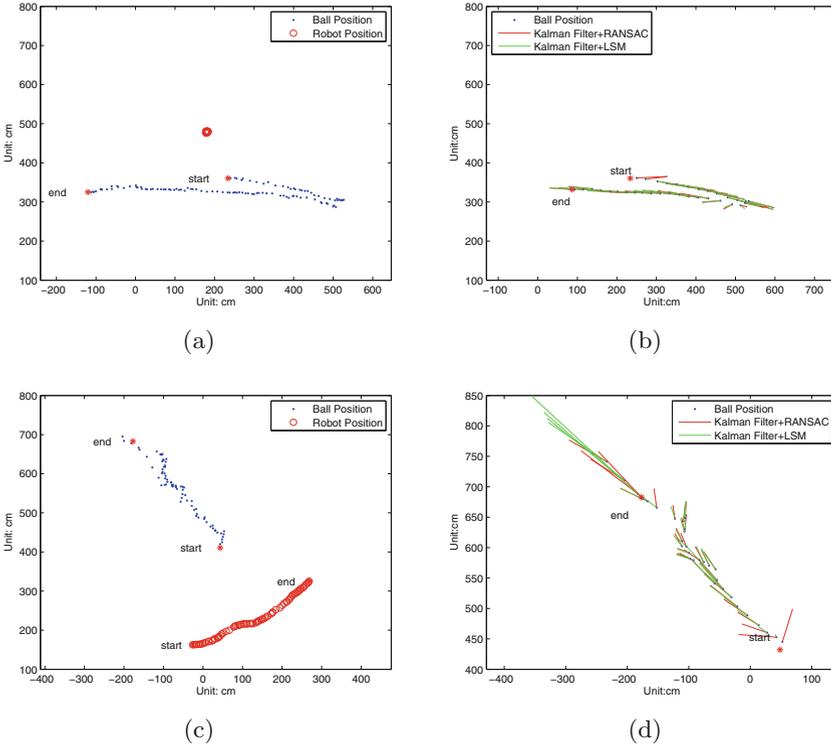


Fig. 6. The robot and ball positions (a)(c) and the ball velocities estimated by Kalman filter+RANSAC and Kalman filter+LSM (b)(d) in actual experiments when the robot is stationary (a)(b) or moving (c)(d).

showing that our robot estimating the ball velocity by using the proposed Kalman filter+RANSAC algorithm can be found on:

http://v.youku.com/v_show/id_XNzI2MDM5MzAw.html

Because the computation complexity of Kalman filter+RANSAC is quite low and it can be performed within several milliseconds, the discussion about the real-time performance is not necessary in this paper.

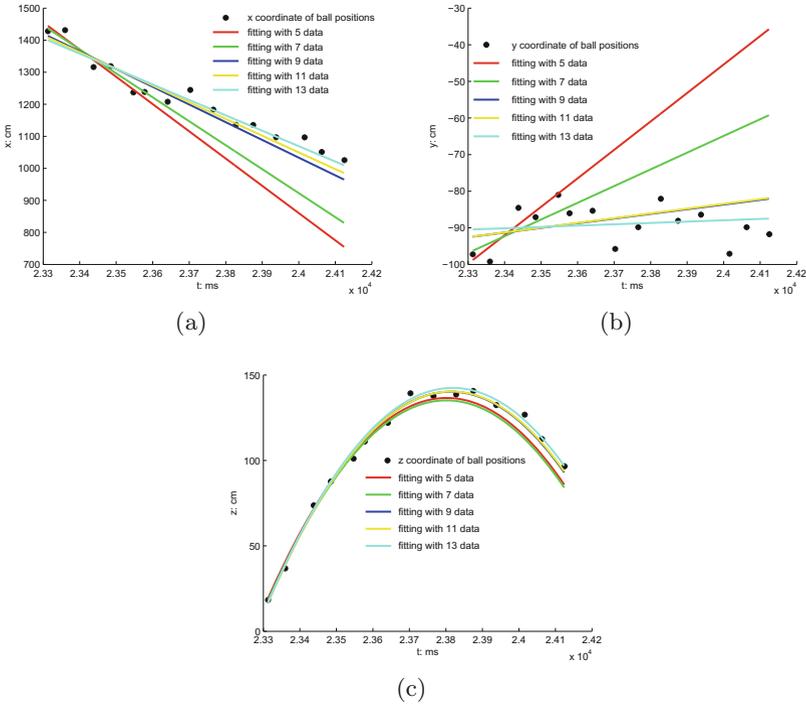


Fig. 7. The moving trace of the ball is fitted by the proposed algorithm in x (a), y (b), z (c) coordinate respectively when using different numbers of ball positions.

5.2 Ball Motion Estimation Results by Stereo Vision

The moving trace of the ball is fitted by the proposed algorithm in x , y , z coordinate respectively. A typical fitting process is shown in Fig. 7, where the first fitting with five ball positions, and the subsequent fitting with seven, nine, eleven, and thirteen ball positions are demonstrated. We see that the fitting results converge during the iterative process, which verifies that the fitting accuracy becomes higher as more data of ball positions are used. Although the accuracy of the first fitting is low, it is significant for the robot’s quick response.

We use our goalie robot to test the fitting and prediction results by defending high shooting. During the experiment, a person simulates the high shooting by throwing the ball to the goal from different positions with different distances to the goal about 5~9 meters. The results show that the goalie robot can defend the high shooting with a successful rate of 80 %, which also verifies that the touchdown-point can be predicted effectively using the proposed algorithm. A typical successful defending is shown in Fig. 8. The video about this experiment can be found on: http://v.youku.com/v_show/id_XNzI2MDQwNDcy.html.

Again, the computation complexity of the proposed algorithm is quite low and it can be performed in less than one millisecond, so the discussion about the

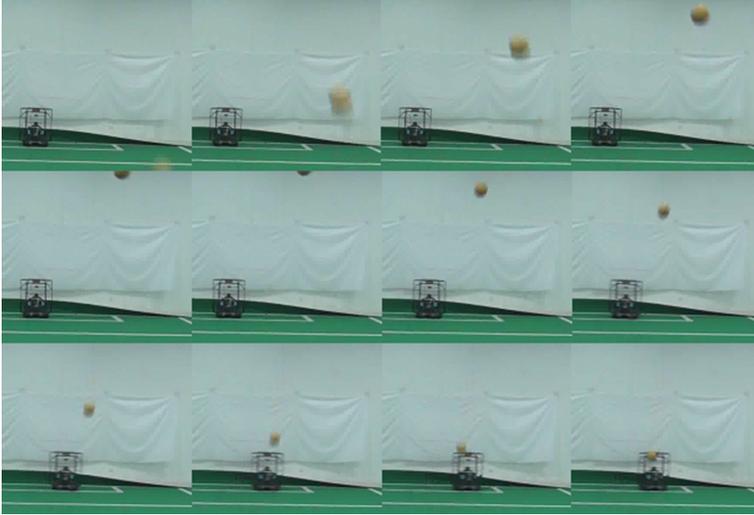


Fig. 8. A typical successful defending of our goalie robot to the high shooting.

real-time performance of the algorithm itself is not necessary in this paper. However, in our hybrid vision system, the omnidirectional vision and stereo vision often work simultaneously, therefore we should evaluate the real-time performance of the hybrid vision system as a whole. The robot's computer is equipped with a 1.66G CPU and 1.0G memory. When the robot works in the competition state, the omnidirectional vision can work on a frame rate of 25 fps, and the stereo vision can work on a frame rate of 20 fps, which meets the real-time requirement of RoboCup MSL competition.

6 Conclusion

In this paper, a hybrid vision system is constructed with omnidirectional vision and stereo vision to realize ball recognition and motion estimation in 3D space. To the best of our knowledge, it is the first time that this kind of hybrid vision is used in RoboCup MSL. When the ball is located on the ground field, a novel algorithm based on RANSAC and Kalman filter is proposed to estimate the ball velocity using the information from the omnidirectional vision. When the ball is kicked up into the air, an iterative and coarse-to-fine method is proposed to fit the moving trace of the ball with parabolic curve and predict the touchdown-point in 3D space using the information from the stereo vision. As a result, the quick response and accuracy can be met simultaneously for soccer robots to intercept or defend the ball. Experimental results show that the robot can realize effective estimation of ball motion in 3D space using the hybrid vision system and the proposed algorithms.

In the future, we would like to compare our ball motion estimation with other approaches existing in RoboCup MSL after trying to acquire the ground

truth. We also want to realize better recognition and motion estimation of other objects like obstacles employing our hybrid vision system.

References

1. Taleghani, S., Aslani, S., Shiry, S.: Robust moving object detection from a moving video camera using neural network and Kalman filter. In: Iocchi, L., Matsubara, H., Weitzenfeld, A., Zhou, C. (eds.) *RoboCup 2008*. LNCS, vol. 5399, pp. 638–648. Springer, Heidelberg (2009)
2. Lauer, M., Lange, S., Riedmiller, M.: Modeling moving objects in a dynamically changing robot application. In: Furbach, U. (ed.) *KI 2005*. LNCS (LNAI), vol. 3698, pp. 291–303. Springer, Heidelberg (2005)
3. Silva, J., Lau, N., Rodrigues, J., Azevedo, J.L., Neves, A.J.R.: Sensor and information fusion applied to a robotic soccer team. In: Baltés, J., Lagoudakis, M.G., Naruse, T., Ghidary, S.S. (eds.) *RoboCup 2009*. LNCS, vol. 5949, pp. 366–377. Springer, Heidelberg (2010)
4. Taiana, M., Gaspar, J.A., Nascimento, J.C., Bernardino, A., Lima, P.: 3D Tracking by catadioptric vision based on particle filters. In: Visser, U., Ribeiro, F., Ohashi, T., Dellaert, F. (eds.) *RoboCup 2007: Robot Soccer World Cup XI*. LNCS (LNAI), vol. 5001, pp. 77–88. Springer, Heidelberg (2008)
5. Taiana, M., Santos, J., Gasper, J., et al.: Tracking objects with generic calibrated sensors: an algorithm based on color and 3D shape features. *Robot. Auton. Syst.* **58**(6), 784–795 (2010)
6. Voigtländer, A., Lange, S., Lauer, M., et al.: Real-time 3D ball recognition using perspective and catadioptric cameras. In: *Proceedings of 2007 European Conference on Mobile Robots (ECMR07)* (2007)
7. Lauer, M., Schönbein, M., Lange, S., Welker, S.: 3D-objecttracking with a mixed omnidirectional stereo camera system. *Mechatronics* **21**(2), 390–398 (2011)
8. Käppeler, U., Höferlin, M., Levi, P.: 3D object localization via stereo vision using an omnidirectional and a perspective camera. In: *Proceedings of the 2nd Workshop on Omnidirectional Robot Vision*, pp. 7–12 (2010)
9. Kanters, F.M.W., Hoogendijk, R., et al.: Tech united eindhoven team description 2011. In: *RoboCup 2011 Istanbul*, CD-ROM (2011)
10. Schoenmakers, F.B.F., Koudijs, G., et al.: Tech united Eindhoven team description 2013 middle size league. In: *RoboCup 2013 Eindhoven*, CD-ROM (2013)
11. Lu, H., Yang, S., Zhang, H., Zheng, Z.: A robust omnidirectional vision sensor for soccer robots. *Mechatronics* **21**(2), 373–389 (2011)
12. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM* **24**, 381–395 (1981)