International Journal of Innovative Computing, Information and Control
Volume **9**, Number **3**, March **2013**

ICIC International ©2013 ISSN 1349-4198
pp. 991–1012

ADAPTIVE VISION-BASED SELF-LOCALIZATION SYSTEM FOR HUMANOID ROBOT OF ROBOCUP

Chih-Hsien Hsia[1], Jen-Shiun Chiang[2] and Shih-Hung Chang[2]

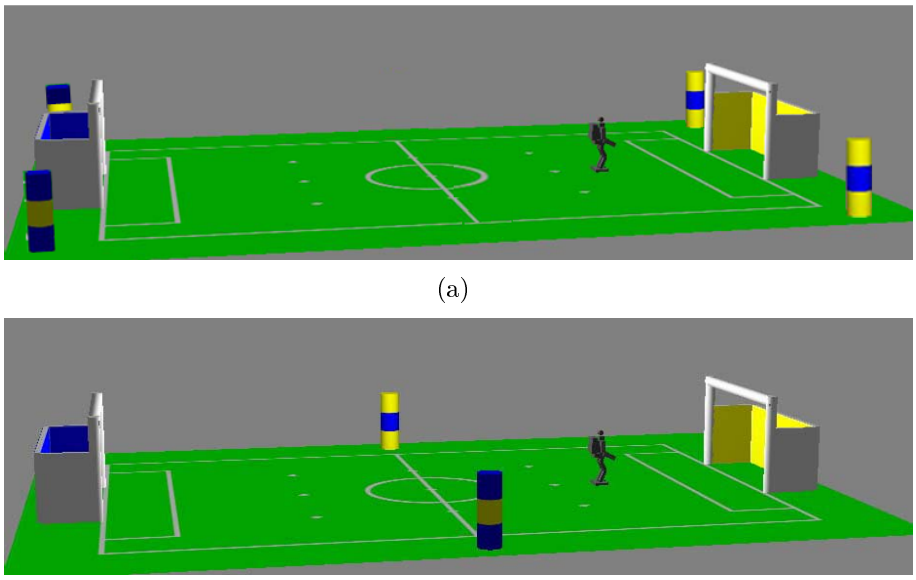
[1]Department of Electrical Engineering
National Taiwan University of Science and Technology
No. 43, Sec. 4, Keelung Road, Taipei 106, Taiwan
chhsia@ee.tku.edu.tw

[2]Department of Electrical Engineering
Tamkang University
No. 151, Yingjhuan Road, Danshui District, New Taipei 25137, Taiwan
696450682@s96.tku.edu.tw; chiang@ee.tku.edu.tw

Received December 2011; revised April 2012

Abstract. *Robotic soccer games represent the most significant form of research in artificial intelligence. Using the humanoid soccer robot's basic movement and strategic actions, the robot takes part in a dynamic and unpredictable contest and must recognize its own position on the field at all times. Therefore, the localization system for the soccer robot represents the key technology for improving its performance. This work proposes a new approach for self-localization, an Adaptive Vision-Based Self-Localization System (AVBSLS), which allows the humanoid robot to integrate the information from the pan/tilt motors and a single camera to achieve self-localization. The proposed approach uses a trigonometric function to find the approximate location of the robot and uses a measuring artificial neural network technique to adjust the position of the humanoid robot. A systematic method to measure the intrinsic parameters is proposed for the CCD camera adjustment. Using this approach, any type of CCD camera can be used to precisely calculate the robot's position. The experimental results indicate that the average accuracy of the localization is 92.3% for a frame rate of 15 frames per second (FPS).*
Keywords: Soccer robot, Adaptive vision-based self-localization system (AVBSLS), Humanoid robot

1. **Introduction.** Robotic soccer games represent one of the most important areas of research, in recent years. With the advances in robotics and artificial intelligence, autonomous robots handle not only simple and monotonous problems, but are also capable of independent thought, when dealing with complex states in unpredictable and dynamic environments. All of these functions rely on a powerful vision system, so the vision system is one of the most critical parts of an autonomous robotic system. Therefore, the ability to sense the distance and the corresponding angle between the robot and the interesting target, or the self-localization system, represents the key to improving the performance, so that the data for the strategic actions of the humanoid robot are more robust and the strategic actions follow the most appropriate decision. A good self-localization system for a humanoid soccer robot not only allows the robot to acquire information about the entire field quickly and accurately, but also makes appropriate decisions. To allow easy manipulation all of the locations in the field are preset as a Cartesian coordinate system and the robot self-localizes itself using the coordinate system. In recent years, the competition fields for RoboCup [1] and FIRA Cup [2] have become more similar to human environments. Figure 1 shows the RoboCup soccer fields for humanoids from kid-size of



(a)

(b)

FIGURE 1. The configuration of RoboCup soccer fields for humanoid kid-size: (a) for 2007, (b) for 2008 and 2009

2007, 2008 and 2009, respectively. The number of landmarks decreases from four to two (2007-2009) [3-5]. In other words, the reference checkmarks for self-localization become less and less, and how to use fewer landmarks and increase the degree of accuracy become important issues [6,7].

Zhong *et al.* [8] proposed three techniques for robotic localization. The first approach uses stereo vision. This approach gathers a lot of information, but the problems of matching characteristic or images between the left and right camera lead to inaccurate distance perception [9] and reduce the accuracy of localization. The second method uses omni-directional vision. Although this method obtains much more features, the omni-directional device causes geometrical distortions in the perceived scene [10], so the recognition of a target of interest is difficult. The third method uses monocular vision. It must have required robust features within a specific region [11], so it is unsuitable for a dynamic contest environment. SLAM [12] is an algorithm for simulating localization and mapping. Although this algorithm performs well, it requires very complex image processing procedures. These complex image processing procedures may affect the real-time operation in a competition, especially in a robotic soccer contest. Actually for real time consideration in the RoboCup competition, the localization should be simple and efficient, and SLAM may not be suitable for this application.

The uses of stereo vision, omni-directional vision, and SLAM for robot localization are not suitable for operating in a dynamic contest environment such as the RoboCup competition, because their operation is more complicated than that for monocular vision. Therefore, for the RoboCup soccer competition a visual self-localization method, an Adaptive Vision-Based Self-Localization System (AVBSLS), is proposed, which uses a single CCD camera and pan/tilt motors on the robot head to identify robust features and to analyze the environmental information for the RoboCup soccer field of 2009.

Firstly, the proposed AVBSLS measures the distance between the robot and the target of interest, using the image. It then establishes the patterns for the distance, using a Back Propagation Neural (BPN) network to improve the accuracy of the distance estimation. It then uses statistics to accurately ascertain the direction of the robot in the competition

field. Lastly, the distance between the robot and the target of interest is combined with its direction. This results in an adaptive vision-based self-localization system for a humanoid robot in RoboCup that is suitable for use with monocular vision. The proposed AVBSLS can also complete self-localization system within a frame that contains the target of interest, so that the robot can execute an appropriate strategy.

The rest of this paper is organized as follows. Section 2 presents the general localization methods and suffered problems. The proposed AVBSLS mechanism is described in Section 3, and the experimental results for the self-localization system are shown in Section 4. Section 5 compares and analyzes the self-localization for humanoid robot by AVBSLS. Finally, Section 6 gives a brief conclusion.

2. **Robotic Vision-Based Localization.** For a humanoid robot localization involves the robot analyzing its probable position on the field. The key technology of the self-localization allows the robot to take the advantage of the information from various sensors to match the position of the robot. Because the robot's perception is restricted and the ambient environment is with enormous interferences, this is difficult to make the robot have efficient and robust localization. During localization, because of the restrictions imposed by the performance of various sensors and interferences of the outside environment, it may have uncertainties for the orientation. The main problems are: 1) the dynamic variance of the outside environment; 2) undependable information from the intrinsic parameters of CCD camera. These non-ideal elements lead to reducing the precision of localization. To solve these problems, many researches have examined better ways to model the environment and mathematical tools for simulation [13,14]. This paper proposes an efficient mechanism that overcomes the problem that many machine vision systems require intrinsic parameters (focal distance or horizontal and vertical viewpoints, etc.), by allowing self-localization for a humanoid robot without the intrinsic parameters provided by the makers provided. Therefore, the humanoid robot can recognize its position explicitly on the field, and can track a soccer ball tracking and make strategic plans.

3. **Adaptive Vision-Based Self-Localization System.** Self-localization is one of the most important problems for a humanoid soccer robot. This study uses the landmarks of the 2009 RoboCup soccer field for humanoid kid-size soccer robot as a reference. The humanoid soccer robot has a single camera vision system on its head with pan/tilt motors. As soon as the humanoid soccer robot detects one of the landmarks, the self-localization system begins to localize itself. During the self-localization process, the robot measures the distance between the landmark and the robot itself, using the IBDMS (image-based distance measuring system) [15,16] technique that can be applied to measure the distance. However, there are distortions in the CCD camera and it is necessary to adjust the coarse data to distance.

The Image-Based Distance Measuring System (IBDMS) [15,16] uses two fixed lasers to project laser beams an object, in order to measure the distance between the CCD camera and the object. Figure 2 shows the relationship of the CCD camera and the object. The unit of D_{SH} is cm that is the actual length of the distance between the two lasers. F_H is the pixel numbers that the laser beams project to the object in a frame. If the maximum horizontal pixels that a CCD camera can capture in an image frame is $F_{H(\max)}$, the maximum width (in cm) that a CCD camera can reach can be calculated as (1):

$$D_H(\max) = \frac{F_H(\max)}{F_H} \times D_{SH} \tag{1}$$

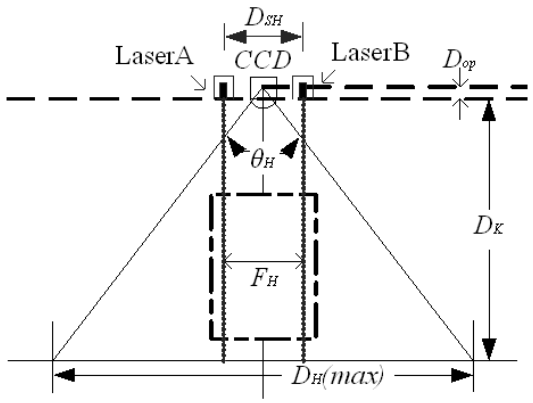


FIGURE 2. The location relationship of the CCD camera and the object

Figure 2 shows that the distance, D_K (in cm), between the CCD camera and the object is calculated, using the intrinsic CCD parameters, as (2)

$$D_K = \frac{1}{2} \times D_H(\max) \times D_{SH} \times \cot\left(\frac{\theta_H}{2}\right) - D_{op} \tag{2}$$

In the RoboCup 2009 rule, the landmark poles have a diagram of 20cm. They consist of three segments of 20cm height, placed above each other. The lowest and the highest segments are colored in the same color as the goal at its left side. Because the two lasers of the IBDMS are not suitable for a wide and dynamic field and one of the lasers is covered by something, the IBDMS applied to measure the distance between the CCD camera and the landmark pole will be ineffective.

Based on the above concepts, this paper proposes a self-localization system that uses an adaptive vision-based self-localization system (AVBSLS) without other assistant devices (such as two lasers) and the exclusive CCD camera, for a humanoid soccer robot for the 2009 RoboCup contest. The proposed AVBSLS consists of eight steps, and the operational flow chart of the self-localization mechanism is shown in Figure 3. The details of the self-localization operations are described in the following subsections.

3.1. **Establishment of the coordinate system.** If the coordinate of a geometric map is available, it will be convenient to retain a lot of information in the whole field. For easy manipulation of the self-localization of a robot, the coordinate system of the field must be established in advance. Before processing the localization, we must establish two appropriate coordinate systems. One is called "absolute coordinate system" on the field, and the other is called "relative coordinate system" in the image. There are four steps to establish the absolute coordinate system: 1) to estimate the sizes of the field and robot; 2) to find the interested position in the soccer field; 3) to divide the field into several blocks with the same size and assign the interested position as the center block; 4) according to the proportion of the robot in the field to adjust the value in each block. The relative coordinate system will store the interested information of the objects. Through these coordinate systems, the locations of the robot, landmark, and goal can be located explicitly.

3.2. **Landmark search.** Since the landmarks are the fixed features on the 2009 RoboCup soccer field, we treat the landmark as the reference for localization. In the initialization of the orientation, the robot keeps searching one of the landmarks until finding it. After finding the landmark, the system will take the interested feature by converting the image

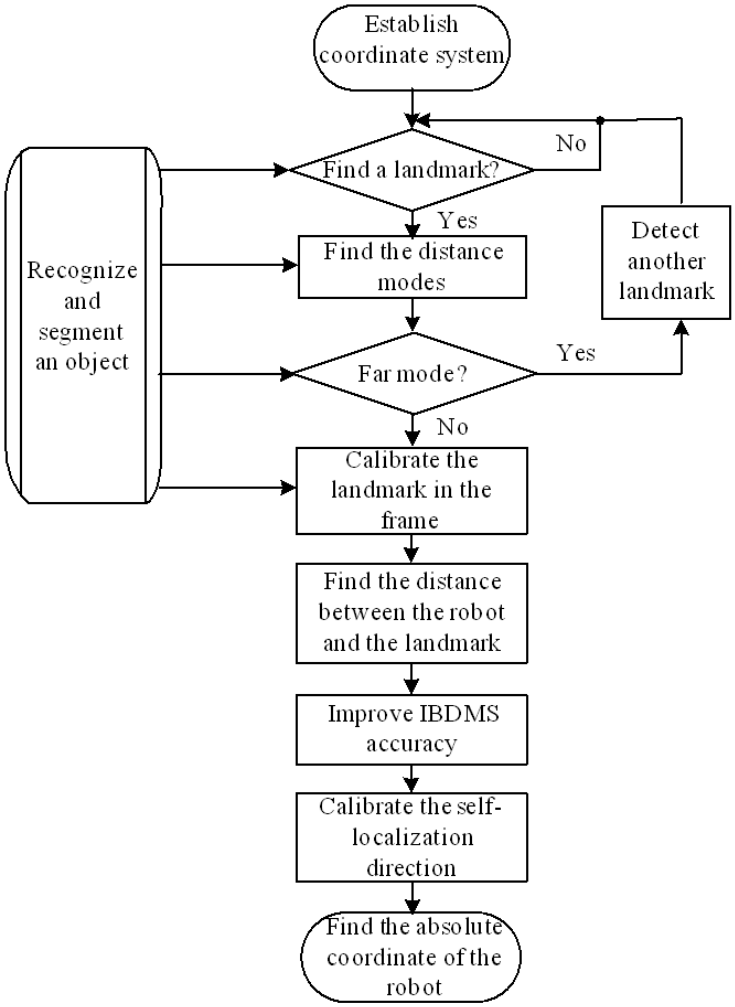


FIGURE 3. The AVBSLS flowchart for the humanoid robot

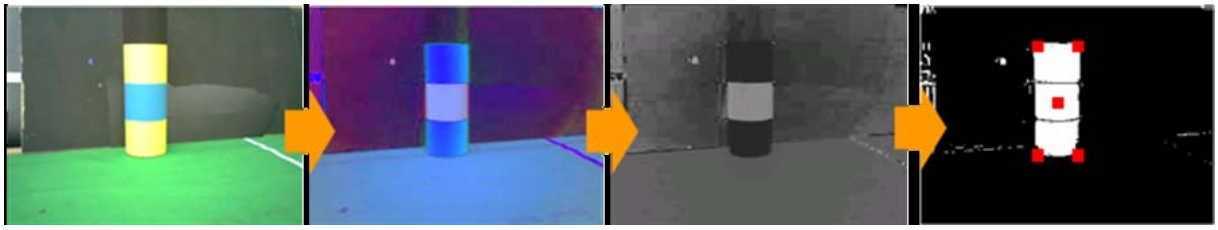
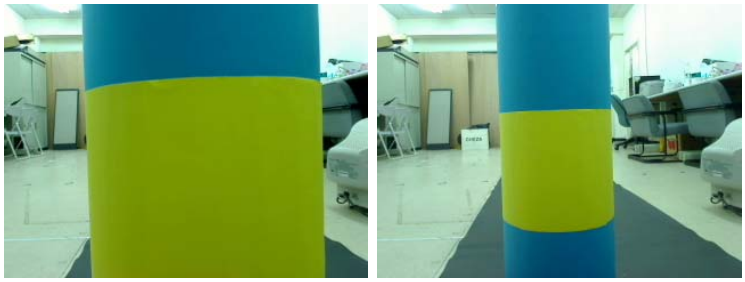


FIGURE 4. The process to mark the upper left, upper right, lower left, lower right, and center of the landmark

from RGB to HSI space. In order to remove the influence of brightness of light, it takes the H-, S-space only [17]. Finally, it will mark five feature points, upper left (X_1, Y_1), upper right (X_2, Y_2), lower left (X_3, Y_3), lower right (X_4, Y_4), and center (X_C, Y_C), for the landmark in the image, as shown in Figure 4. According to the five feature points, the horizontal length F_H and vertical length F_V for the landmark in the frame can be found. To get robust F_H and F_V, the landmark shape must be completely displayed in the frame, as shown in Figures 5(b) and 5(c) show the incomplete landmarks with partial landmark displayed in the frame. During the landmark searching if the robot visual system finds the incomplete landmark in the frame or the landmark is occluded by other robots, the robot system will command the robot to make a suitable movement [18]. In order to insure the

(a) (b) (c)

FIGURE 5. Landmark searching: (a) incomplete landmark, (b) complete landmark, (c) incomplete landmark



(a) (b)

FIGURE 6. The landmark images with distances of: (a) 30cm and (b) 50cm. (F_H, F_V) are equal to (a) (104, 238), (b) (155, 238) pixels in the frame.

completeness of the landmark, the pixel values between the edge of the landmark and the edge of the frame must be greater than an appropriate value.

3.3. **Distance mode selection.** The distance between the robot itself and the landmark is measured by calculating of trigonometric functions together with the lengths of F_H and F_V. The lengths of F_H and F_V are determined by the resolution of the CCD camera. If the object (landmark) is too far away from the robot itself, the lengths of F_H and F_V cannot be accurately determined. The distance between the robot itself and the landmark, it is classified into three modes, near mode, mid mode, and far mode. During the self-localization process the pan motor can move randomly but the tilt motor is fixed in the center position. Figure 6 shows the landmark images when the robot is very close to the landmark. In this situation, both F_V's of the landmarks are indistinguishable and F_H can be used to find the distance between the robot itself and the landmark, and this range is classified as the near mode.

If the distance between the robot itself and the landmark is far excessively (more than 400cm for example), F_H's and F_V's for different landmarks are indistinguishable and Figure 7 shows this situation. Under this situation, it is impossible to measure the distance between the robot itself and the landmark by the proposed self-localization mechanism, and therefore the robot has to give up this landmark and find the other landmark for self-localization. This distance range is classified as the far mode.

If the distance between the robot itself and the landmark is between the near mode and the far mode, it is classified as the mid mode. Figure 8 shows images of landmarks in the mid mode. In the mid mode situation, F_H's are indistinguishable but F_V's are

(a) (b)

FIGURE 7. The landmark images with distances of: (a) 410cm and (b) 450cm. (F_H, F_V) are equal to (a) (14, 38), (b) (14, 38) pixels in the frame.

(a) (b) (c)

FIGURE 8. The images of the landmark with distances of: (a) 200cm, (b) 210cm and (c) 250cm. (F_H, F_V) are equal to (a) (28, 80), (b) (28, 77) and (c) (27, 71) pixels in the frame.

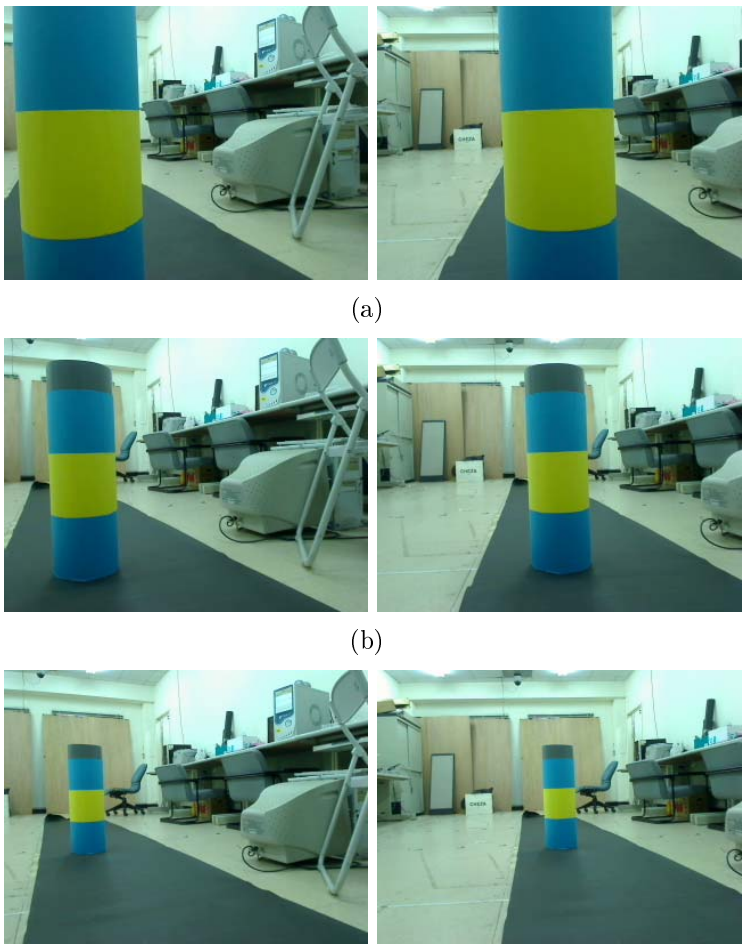
distinguishable, and therefore F_V can be used to find the distance between the robot itself and the landmark.

3.4. **Landmark calibration in the frame.** Because the pan motor can move arbitrarily, the landmark may appear in any position of the frame. However, nonlinear factors, such as lens distortion in the CCD camera and the influence of the brightness of light, may produce inconsistencies in the captured images. With the same distance between the robot itself and the landmark, but with the landmark in different positions of the frame, the pixel sizes of the landmarks are different and Figures 9 and 10 show the features. For finding a more accurate location, the captured landmark by the CCD camera must be calibrated. Here we have the statistical approach to calibrate the size of the captured landmark.

3.5. **The distance between the robot and the landmark.** The measurement of the distance between the robot itself and the landmark is accomplished using the proposed modified IBDMS. The details of distance measurement are described in the following subsections.

3.5.1. *Intrinsic parameters of the CCD.* The vision-based distance measurement needs the intrinsic parameters of the CCD camera, such as focal distance and horizontal and vertical viewpoints. Generally camera manufacturers rarely provide such parameters and, if they do, the mechanical mechanism, processing circuits, and lens distortion may cause

FIGURE 9. With the same distance from the robot to the landmark, (F_H, F_V)'s of the landmark at left, middle, and right positions of the frame are (86, 227), (80, 219), (87, 223) pixels, respectively.



(a)

(b)

(c)

FIGURE 10. The landmarks at the same position in the frame with different distances of: (a) 50cm, (b) 100cm and (c) 200cm have variations of F_H between the left and right frames of (a) 9 pixels, (b) 5 pixels and (c) 1 pixel.

deviations in the parameters [19-21]. This study proposes some simple procedures that allow an independent determination of the intrinsic parameters of a CCD camera.

The relationship of the distance between the CCD camera and the object and some parameters are shown in Figure 11. OP is the optical origin, θ_H and θ_V are the horizontal

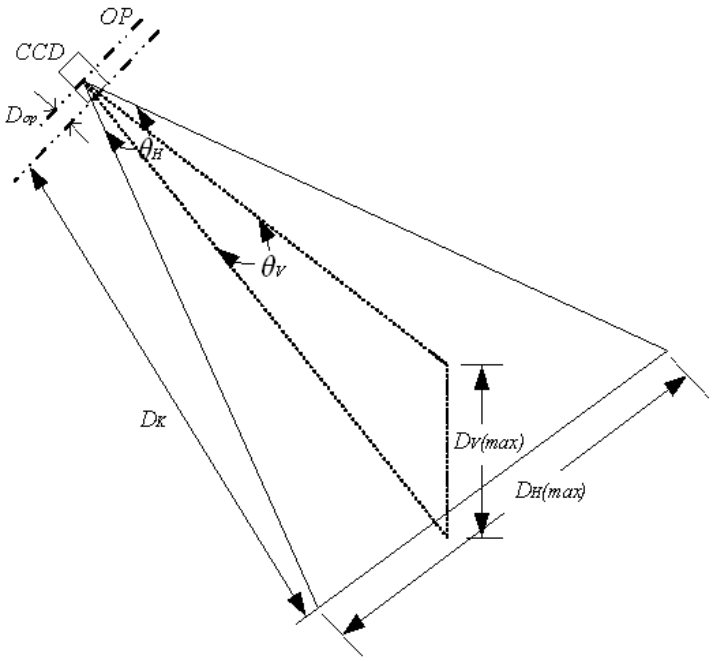


FIGURE 11. The diagram of the internal parameters of a CCD camera

and vertical viewpoints, respectively, and D_{OP} is the distance between OP and the edge of the CCD camera lens.

In order to determine the distance between the CCD center and the object, θ_H, θ_V, and D_{OP} must be determined. The width of an object, $D_{H1}(\max)$, is measured at a pre-defined position, D_A, and the width of the same object, $D_{H2}(\max)$, at a pre-defined position, D_B. The two frames are combined to form Figure 12(a). Similarly, we can capture $D_{V1}(\max)$ and $D_{V2}(\max)$ with known pre-defined positions, D_A and D_B, in two frames. These two captured frames are combined to form Figure 12(b). Due to the characteristics of the CCD camera, there is a different value for D_{OP} at D_A and D_{K1}, and there is also a D_{OP} difference in between D_B and D_{K2}. The same thing happens in the vertical viewpoint in Figure 12(b). Together with Figures 12(a) and 12(b) and the trigonometric functions we can derive θ_H and θ_V are derived in (3) and (4).

$$\theta_H = 2\cot^{-1}\left(2 \times \frac{D_{K2} - D_{K1}}{D_{H2}(\max) - D_{H1}(\max)}\right) \tag{3}$$

$$\theta_V = 2\tan^{-1}\left(\frac{1}{2} \times \frac{D_{V2}(\max) - D_{V1}(\max)}{D_{K2} - D_{K1}}\right) \tag{4}$$

By the theorem of similar triangles and Figures 12(a) and 12(b), the horizontal and vertical D_{OP}'s can be found calculated as shown in (5) and (6).

$$\frac{D_{OP(H)} + D_{K1}}{D_{OP(H)} + D_{K2}} = \frac{D_{H1}(\max)}{D_{H2}(\max)} \tag{5}$$

$$\frac{D_{OP(V)} + D_{K1}}{D_{OP(V)} + D_{K2}} = \frac{D_{V1}(\max)}{D_{V2}(\max)} \tag{6}$$

D_{OP} can be found by averaging $D_{OP(V)}$ and $D_{OP(H)}$ as shown in (7):

$$D_{op} = \frac{1}{2} \times \left(\frac{D_{K2}D_{H1}(\max) - D_{K1}D_{H2}(\max)}{D_{H2} - D_{H1}} + \frac{D_{K2}D_{V1}(\max) - D_{K1}D_{V2}(\max)}{D_{V2} - D_{V2}}\right) \tag{7}$$

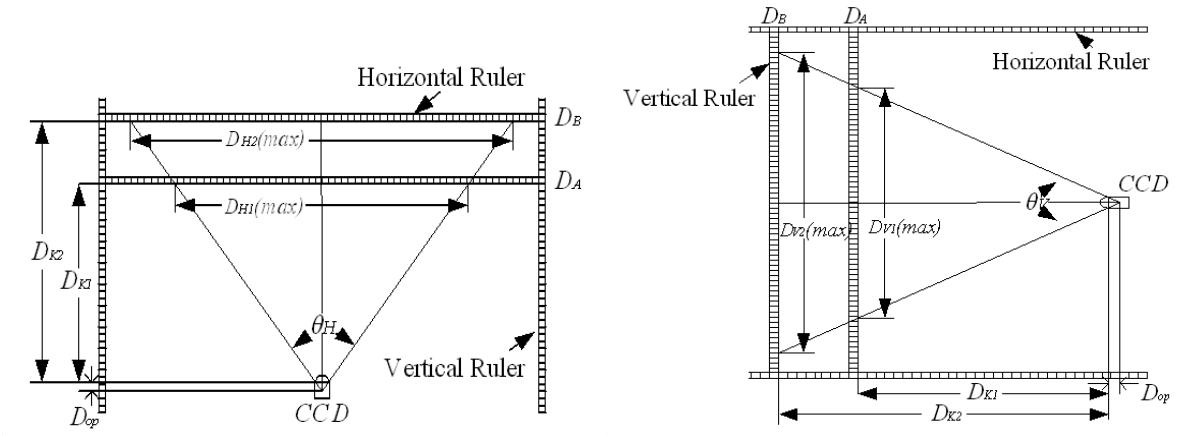


FIGURE 12. The intrinsic parameters measuring: (a) horizontal viewpoint, (b) vertical viewpoint

This approach can be to determine the intrinsic parameters for any kind of CCD cameras.

3.5.2. *Image-based distance measurement.* As mentioned in the previous subsection, when the CCD camera is close enough to the landmark (near mode), F_V is indistinguishable and F_H is used to calculate the distance between the CCD camera and the landmark. On the other hand, when the CCD camera keeps a medium range from the landmark, F_H is indistinguishable and F_V to calculate the distance between the CCD camera and the landmark. The IBDMS method to calculate the distance between the CCD camera and landmark, F_H and F_V must be converted to $D_H(\max)$ (the maximum horizontal width) and $D_V(\max)$ (the maximum vertical width). The relationship between F_H and $D_H(\max)$ and F_V and $D_V(\max)$ is shown in Figure 13, and the conversion equations are shown as (8) and (9).

$$D_H(\max) = \frac{F_H(\max)}{F_H} \times D_{SH} \tag{8}$$

$$D_V(\max) = \frac{F_V(\max)}{F_V} \times D_{SV} \tag{9}$$

For a 320×240 image frame, $F_H(\max) = 320$, $F_V(\max) = 240$, D_{SH} = 20cm and D_{SV} = 60cm. F_H and F_V are the lengths, in pixels, of the horizontal and vertical lengths of the landmark in the image frame, as shown in Figure 14.

Using the IBDMS method, the horizontal view distance, D_K (photo-distance), between the CCD camera and the landmark can be determined from (10) and that of the vertical view can be determined from (11).

$$D_K = \frac{1}{2} \times D_H(\max) \times \cot\left(\frac{\theta_H}{2}\right) - D_{OP} \tag{10}$$

$$D_K = \frac{1}{2} \times D_V(\max) \times \cot\left(\frac{\theta_V}{2}\right) - D_{OP} \tag{11}$$

Because the CCD lens and the brightness are not ideal, the measurement of D_K may not be sufficiently accurate. It can be fine-tuned using an artificial neural network. The neural network technique is described in the following subsections.

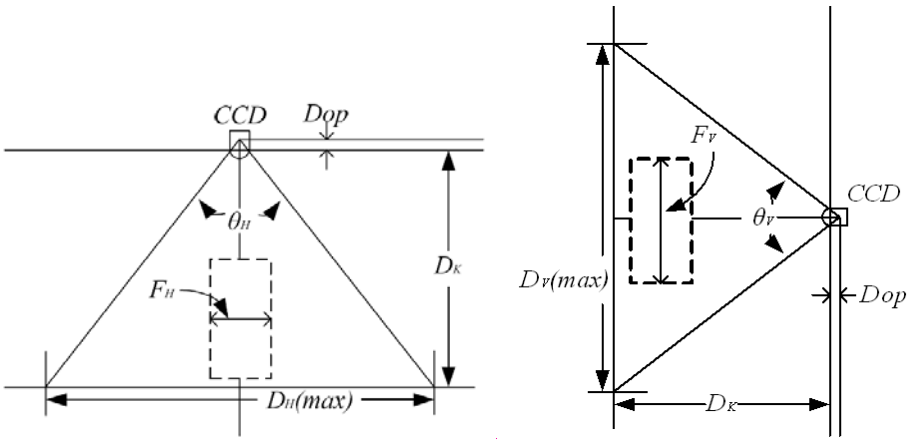


FIGURE 13. Three-dimension distance measurement: (a) horizontal measurement, (b) vertical measurement

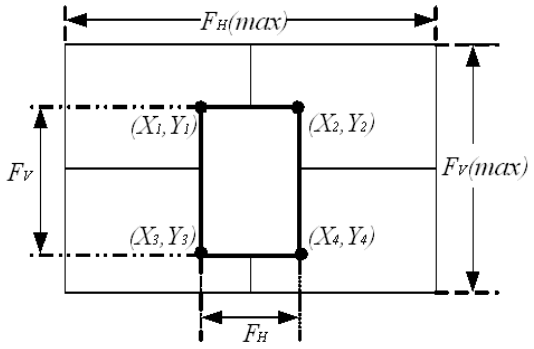


FIGURE 14. The horizontal and vertical information of the landmark in the frame

3.6. **The improved IBDMS.** So far several neural network methods have been proposed, such as the Back Propagation Neural (BPN) network and the Self-Organizing Neural network. Because the BPN network learns more precisely and has a fast recall speed [22], the BPN network is used to focus on the known environment. Depending on the features and the goal distance, the relative parameters for different distances can be trained, in order to acquire the data in advance. If there is a need to measure the distance in the actual competition, the preset relative parameters can be used directly in the BPN formula. This approach allows more accurate measurement of the distance between the robot and the landmark. Using the BPN rules suggested by Chang and Chang [23] a distance measurement is proposed which provides an improved IBDMS. This method has seven steps to improve the distance precision. The procedures are indicated in Figure 15.

Step 1: Prepare for the solid information, including the interesting features of X_C, Y_C and size. In the image frame, the expectable distance value is set as the objective function and these data are then normalized to the appropriate values. The expected distance value is D_K (photo-distance) which is between the CCD camera and the landmark with different distances. The appropriate normalization is referred to the activation function, f, as follows:

$$y_j^n = f(net_j^n) \tag{12}$$

where y_j^n is the output value of the nth layer, and it is also the input value of the first layer. net_j^n is the weight accumulative value for the output value of the $(n-1)$th layer

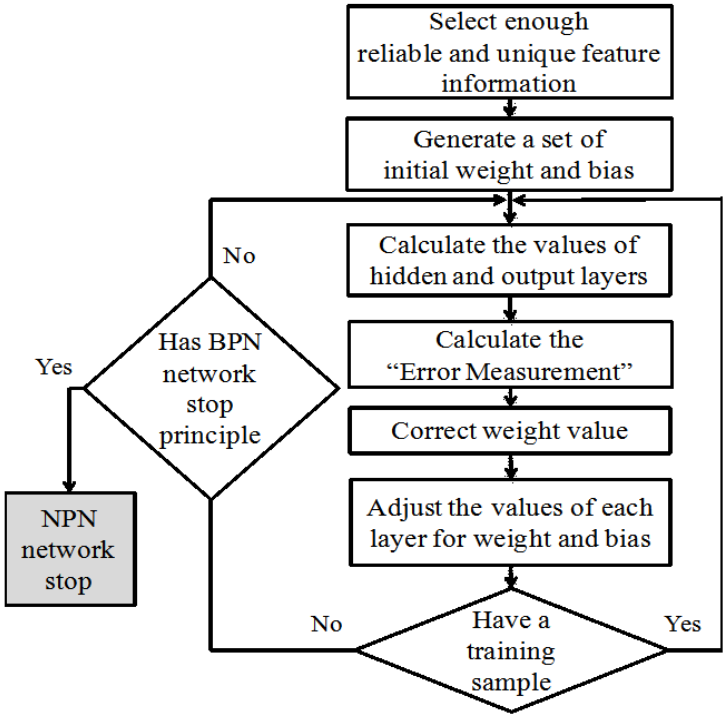


FIGURE 15. The procedure for improving precision

and is represented as follows:

$$net_j^n = \sum_i^n w_{ji}^n y_i^{n-1} + b_j^n \tag{13}$$

where w_{ji}^n is the weighted connections between the jth neuron in the nth layer and the ith neuron in the $(n-1)$th layer and b_j^n is the bias of the jth neuron in the nth layer.

Step 2: Initialize W_{ji} and W_{kj}, using random values.

Step 3: Select a suitable activation function from Figure 16 and input the trained data to the selected activation function. Then, calculate the output value, y, from the hidden layer, and output the value of y_k from the output layer.

Step 4: Calculate the error function, E. In order to find the optimum solution of E, the steepest descent method is used, as shown in (14):

$$E = \frac{1}{2} \sum_k (d_k - y_k)^2 \tag{14}$$

where d_k is the kth neuron objective output value and y_k is the output value of the kth neuron at the output layer. This step reduces the difference between the input and output values.

Step 5: Calculate δ_k^n, $k = 1, \ldots, K$, in the output layer, using (15), and δ_j^n, $j = 1, \ldots, L$, in the hidden layer, using (16).

$$\delta_j^n = (d_j - y_j^n) f'(net_j^n) \tag{15}$$

$$\delta_j^n = \left[\sum_k \delta_k^{n+1} w_{kj} \right] f'(net_j^n) \tag{16}$$

Step 6: Correct the weight $\left(W_{kj}(p+1) = W_{kj}(p) + \eta \delta_k^n(p) y_j^{n-1}(p)\right)$ in the output layer and the weight $\left(W_{ji}(p+1) = W_{ji}(p) + \eta \delta_j^n(p) y_i^{n-1}(p)\right)$ in the hidden layer, where p is the

module of the group, p (the training module includes input and output values) and η is the learning rate, with a general value between 0 and 1.

Step 7: Go back to Step 3 and then repeat the calculation and correction procedures until the objective function reaches the stop standard, or the largest training times.

The neural network method allows calculation of a more accurate distance between the robot and landmark.

3.7. **Calibration of the self-localization direction.** This study has so far been concerned with the photo-distance, but self-localization must also include the establishment of the direction of the robot, to determine the absolute coordinates of the robot itself on the soccer field. The rotation angle of the pan motor is used to determine the robot direction. The rotation angle, α', of the pan motor is assumed to be 0 degrees, when the center of the landmark is at the center of the frame. The pan motor can rotate at most left and right 25 degrees, respectively. In our self-localization method, as soon as the robot sees a complete landmark, it begins to calculate the photo-distance and measures the pan motor angle. However, the pan motor angle is "α", instead "α' ", as shown in Figure 17. In order to determine α', the horizontal pixel numbers, X_C, of the center of the landmark must be measured. Together with the photo-distance and X_C Table 1 is used to find α' of the pan motor angle. Figure 17 shows that when α' is determined, β can thus be calculated and then the absolute coordinates of x' and y' can be calculated,

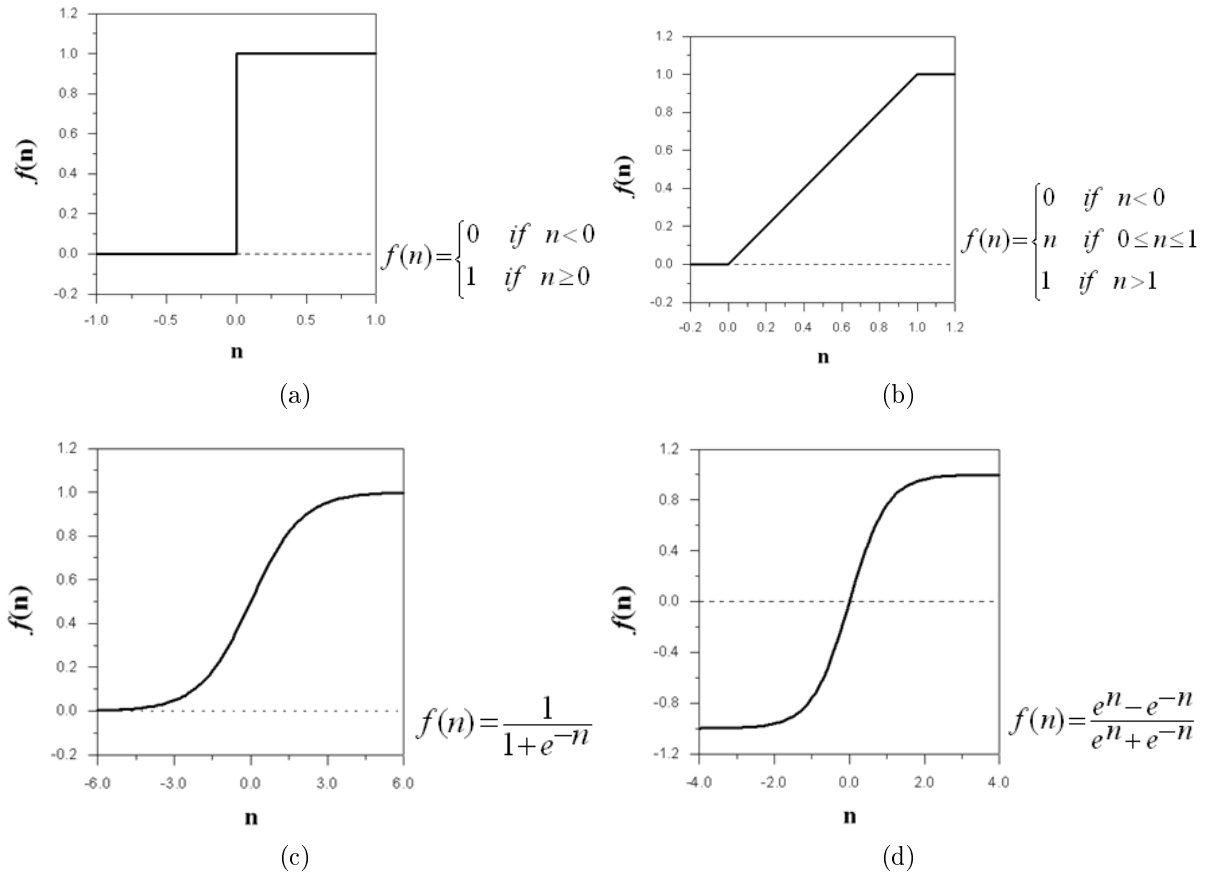


FIGURE 16. Four activation functions: (a) step function, (b) saturating linear function, (c) sigmoid function, (d) hyperbolic function

TABLE 1. The pixel numbers of X_C with given D_K and α

α \ D_k	–25	–20	–15	–10	–5	0	5	10	15	20	25
50	35	59	77	105	128	152	174	197	222	251	264
100	33	60	77	104	127	150	169	197	219	248	261
200	33	58	77	101	126	148	168	191	219	244	262
300	30	58	76	101	127	144	162	190	215	241	263
400	30	58	76	102	127	145	161	189	215	241	253

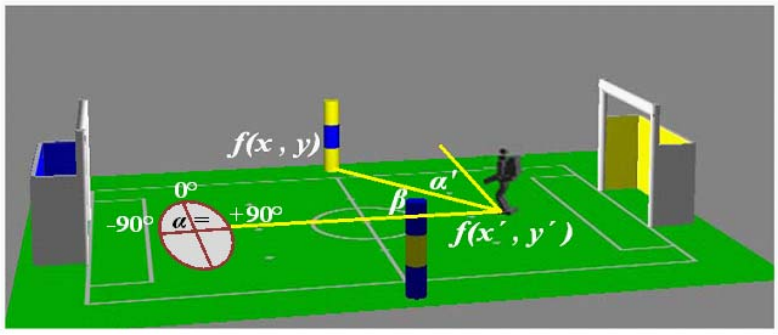


FIGURE 17. The direction of the robot in the soccer field

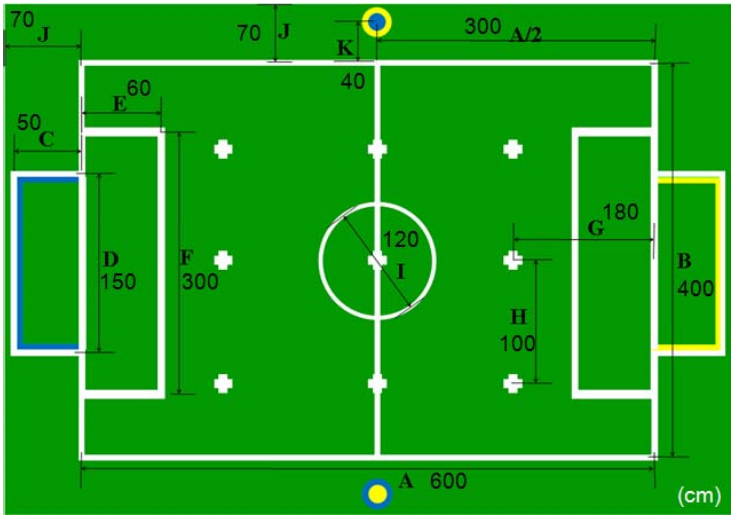


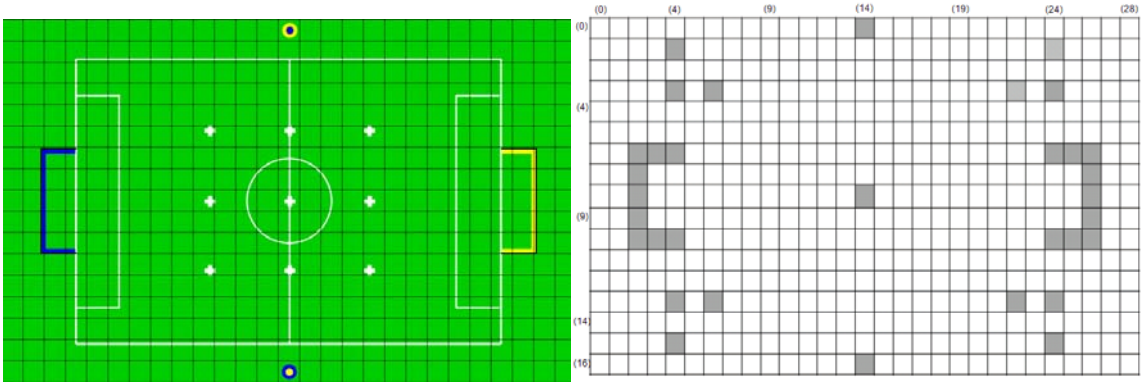
FIGURE 18. Configuration of the RoboCup soccer field for humanoid kid-size in 2009

as follows:

$$\begin{cases} x' = x + r\cos\beta \\ y' = y - r\sin\beta \end{cases}, \quad -90^\circ \le \beta \le 90^\circ \tag{17}$$

4. Experimental Results.

4.1. **The experimental environment and the robotic vision module.** This experiment uses the features of the competition field for the 2009 RoboCup soccer humanoid league. The field contains two goals and two landmark poles, and has representative colors for the interesting features. The border strip width, J, is 70cm, as shown in Figure 18. Because the width of the robot shoulder is 26cm, the unit length of the coordinate is set as 30cm in length and the field is divided into 29×17 blocks, as shown in Figure 19.



(a) (b)

FIGURE 19. The RoboCup soccer field: (a) the original field with 29×17 blocks, (b) the coordinate of the soccer field (2009)

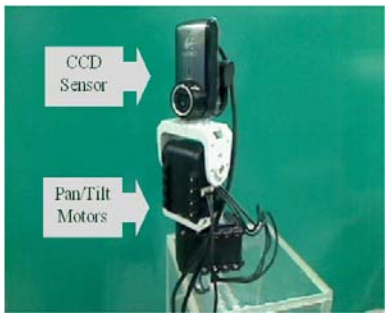


FIGURE 20. The robot vision module

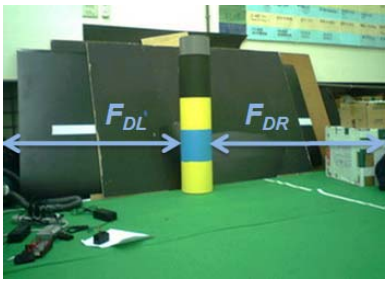


FIGURE 21. F_{DL} and F_{DR}

The experimental robotic vision module comprises a single CCD camera and pan/tilt motors, as shown in Figure 20. The CCD camera is a Logitech QuickCam® Professor [24] for Notebooks and the pan/tilt motors are ROBOTIS Dynamixel AX-12 [25]. The input for the robot vision module is RGB 24 bits, with a resolution of 320×240. The output is the absolute coordinates of the humanoid robot in the field.

4.2. **Complete landmark determination.** In order to determine the completeness of the landmark from 30cm to 400cm distances, the pan motor rotates five degrees, each time, to measure the difference between the edge of F_H and the frame. In the following step a pixel value is chosen which can determine whether the landmark is complete in the frame, from 30cm to 400cm distance. However, if the restricted landmark is suitable for the near mode, the other modes (such as mid mode and far mode) can be used. If the humanoid robot faces the landmark correctly, F_{DL} and F_{DR} are measured at the different distances and horizontal angles and the smaller values are listed in Table 2 if the humanoid robot faces the landmark right, wherein "X" shows that the landmark has exceeded the frame and the F_{DL} and F_{DR} are as shown in Figure 21.

Using the information in Table 2, the pan motor angle is rotated by ±30 degrees, so the complete landmark can be calculated in any distance. Thus, the pan motor angle can be rotated by ±25 degrees at most. Beside from that, if the difference between the edge of F_H and the edge of the frame is smaller than 32 pixels, the landmark is treated as an incomplete landmark.

TABLE 2. $Min(F_{DL}, F_{DR})$ for the different distances (cm) and angles (degree) in the near mode

D_k \ α	–30	–25	–20	–15	–10	–5	0	5	10	15	20	25	30
30	X	X	X	X	55	83	85	84	55	X	X	X	X
40	X	X	X	40	70	98	100	96	70	41	X	X	X
50	X	X	32	54	82	110	112	111	82	55	33	X	X
60	X	X	34	62	89	118	116	116	90	63	35	X	X
70	X	X	36	70	93	122	124	122	94	72	37	X	X
80	X	X	42	76	100	124	127	126	100	76	42	X	X
90	X	X	49	80	103	129	130	130	104	81	48	X	X
100	X	X	53	84	106	131	132	134	106	84	52	X	X
110	X	X	57	86	110	135	135	136	110	86	56	X	X
120	X	X	58	88	112	137	137	138	113	88	58	X	X
130	X	32	60	90	114	139	139	139	115	90	59	X	X
140	X	33	60	92	116	140	140	140	116	92	60	34	X
150	X	36	63	94	117	141	141	141	118	95	62	36	X
160	X	37	66	96	119	142	142	142	120	97	66	38	X
170	X	39	68	98	123	143	143	144	121	99	68	39	X
180	X	42	68	100	126	145	144	145	126	101	69	42	X
190	X	44	72	103	127	147	145	147	126	103	72	45	X

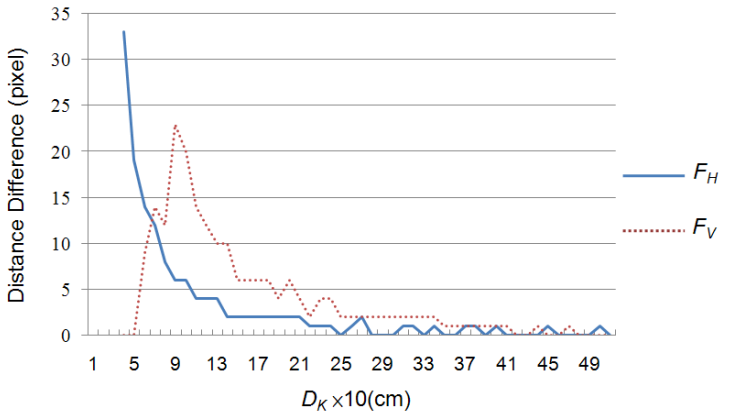


FIGURE 22. The distance difference for F_H and F_V with various photo-distances

4.3. **Localization distance analysis.** The photo-distances, D_K, are classified into three modes. If D_K is less than 30cm, F_H and F_V exceed the 320×240 image frame to loss the analysis and the robot has located in the outside field. The greatest distance between the robot and the landmark is 500cm, in the competition field. Thus, F_H and F_V are captured starting with D_K = 30cm and the distance is increased in 10cm steps, until D_K = 500cm. Figure 22 shows the difference in pixels of F_H and F_V for each distance step. Figure 22 shows that when $D_K \leqq 190$cm, F_H is distinguishable, and hence it is set as the near mode. When $D_K \geqq 400$cm, F_H and F_V are indistinguishable, and it is set as the far mode. While 190cm $\leqq D_K \leqq$ 400cm, it is set as the mid mode.

4.4. **Calibration of the landmark and the robot direction in the frame.** Because the CCD camera lens and the brightness are not ideal, there may be a deviation between the measured photo-distance and the pan motor angle and this must be calibrated. Since the far mode is out of the processing capability, only the near mode and mid mode are

TABLE 3. The pixel numbers of F_H with different D_K (cm) and X_C (pixels)

X_C \ D_k	30	55	80	105	130	150	170	195	220	245	270
100	62	60	58	56	54	54	54	56	58	60	62
150	44	42	40	38	38	38	38	38	40	42	44
200	34	32	30	28	28	28	28	28	30	32	34
250	74	70	68	68	64	64	64	68	68	70	74
300	60	58	56	56	54	54	54	56	56	58	60
350	50	48	46	46	45	45	45	46	46	48	50
400	43	42	41	41	40	40	40	41	41	42	43

calibrated. To calibrate the pixel of F_H in the frame, Table 3 shows the parts of the pixel numbers of F_H at different D_K and X_C. From Table 3 shows that at the same photo-distance, D_K the F_H value is symmetrical to the center value of X_C (150 pixels). Because of the restrictions of the pan motor angle, the value of X_C is in between 30 and 270 pixels and the unit of measurement for the pixel numbers is set as 15 pixels. X_C of the mid pixel number of the measured X_C's is selected under the restrictions of Equation (18). To obtain a correct value for F_H, F_H must be calibrated to the center value of X_C (150 pixels) by X_C and F_H. Therefore, we statistically derive (18) and (19) to calibrate F_H and F_V under different X_C's.

$$\left\{\begin{array}{l} X_C < 43 | X_C \geq 256 \left\{\begin{array}{l} F_H \geq 72, F_H = F_H - 10 \\ F_H < 72 \& F_H \geq 59, F_H = F_H - 8 \\ F_H < 59, F_H = F_H - 6 \end{array}\right\} \\ (X_C \geq 43 \& X_C < 67) | (X_C \geq 234 \& X_C < 256) \left\{\begin{array}{l} F_H \geq 70, F_H = F_H - 10 \\ F_H < 70 \& F_H \geq 49, F_H = F_H - 6 \\ F_H < 49, F_H = F_H - 4 \end{array}\right\} \\ (X_C \geq 67 \& X_C < 90) | (X_C \geq 208 \& X_C < 234) \left\{\begin{array}{l} F_H \geq 117, F_H = F_H - 10 \\ F_H < 117 \& F_H \geq 75, F_H = F_H - 6 \\ F_H < 75 \& F_H \geq 55, F_H = F_H - 4 \\ F_H < 55, F_H = F_H - 2 \end{array}\right\} \\ (X_C \geq 90 \& X_C < 116) | (X_C \geq 183 \& X_C < 208) \left\{\begin{array}{l} F_H \geq 113, F_H = F_H - 4 \\ F_H < 113 \& F_H \geq 49, F_H = F_H - 2 \\ F_H < 49, F_H = F_H - 0 \end{array}\right\} \\ X_C < 183 | X_C \geq 116 \{F_H = F_H\} \end{array}\right\} \tag{18}$$

$$\left\{\begin{array}{l} X_C < 44 | X_C \geq 250 \left\{\begin{array}{l} F_V \geq 84, F_V = F_V - 10 \\ F_V < 84 \& F_V \geq 62, F_V = F_V - 8 \\ F_V < 62 \& F_V \geq 49, F_V = F_V - 6 \\ F_V < 49, F_H = F_V - 4 \end{array}\right\} \\ (X_C \geq 44 \& X_C < 68) | (X_C \geq 228 \& X_C < 250) \left\{\begin{array}{l} F_V \geq 82, F_V = F_V - 8 \\ F_V < 82 \& F_V \geq 70, F_V = F_V - 6 \\ F_V < 70 \& F_V \geq 47, F_V = F_V - 4 \\ F_V < 47, F_V = F_V - 2 \end{array}\right\} \\ (X_C \geq 68 \& X_C < 90) | (X_C \geq 201 \& X_C < 228) \left\{\begin{array}{l} F_V \geq 58, F_V = F_V - 4 \\ F_V < 58 \& F_V \geq 47, F_V = F_V - 2 \\ F_V < 47, F_V = F_V - 1 \end{array}\right\} \\ (X_C \geq 90 \& X_C < 115) | (X_C \geq 177 \& X_C < 201) \left\{\begin{array}{l} F_V \geq 58, F_H = F_V - 4 \\ F_V < 58 \& F_V \geq 47, F_V = F_V - 2 \\ F_V < 47, F_V = F_V - 1 \end{array}\right\} \\ X_C < 177 | X_C \geq 115 \{F_V = F_V\} \end{array}\right\} \tag{19}$$

In Section 3.7, it is shown that it is statistically possible to find a better α' for different X_C's, as shown in Equation (20).

$$\left\{\begin{array}{ll} X_C < 43, & \alpha' = \alpha - 25 \\ X_C \geq 256, & \alpha' = \alpha + 25 \\ X_C \geq 43 \& X_C < 67, & \alpha' = \alpha - 20 \\ X_C \geq 234 \& X_C < 256, & \alpha' = \alpha + 20 \\ X_C \geq 67 \& X_C < 90, & \alpha' = \alpha - 15 \\ X_C \geq 208 \& X_C < 234, & \alpha' = \alpha + 15 \\ X_C \geq 90 \& X_C < 116, & \alpha' = \alpha - 10 \\ X_C \geq 183 \& X_C < 208, & \alpha' = \alpha + 10 \\ X_C \geq 116 \& X_C < 138, & \alpha' = \alpha - 5 \\ X_C \geq 160 \& X_C < 183, & \alpha' = \alpha + 5 \\ X_C \geq 138 \& X_C < 160, & \alpha' = \alpha \end{array}\right\} \tag{20}$$

4.5. **Obtaining the camera intrinsic parameters.** According to Section 3.5, the distance step is set as 10cm, from $D_K = 30$cm until $D_K = 400$cm, to measure $D_H(\max)$ and $D_V(\max)$. The results for $D_H(\max)$ and $D_V(\max)$ with various D_K's are shown in Tables 4 and 5, respectively.

TABLE 4. The information of $D_H(\max)$ with D_K from 30cm to 190cm

D_K	30	40	50	60	70	80	90	100	110
$D_H(\max)$	46	58	69	81	94	105	116	128	139
D_K	120	130	140	150	160	170	180	190	\
$D_H(\max)$	151	165	173	182	193	204	217	231	\

TABLE 5. The information of $D_V(\max)$ with D_K from 190cm to 400cm

D_K	190	200	210	220	230	240	250	260
$D_V(\max)$	172	180	189	195	206	218	225	232
D_K	270	280	290	300	310	320	330	340
$D_V(\max)$	240	248	257	267	277	288	300	313
D_K	350	360	370	380	390	400	\	\
$D_V(\max)$	320	327	335	343	351	360	\	\

The values of θ_H, θ_V and D_{OP} are then obtained. The average intrinsic parameters, θ_H, θ_V and D_{OP} are calculated using (21)-(23), respectively.

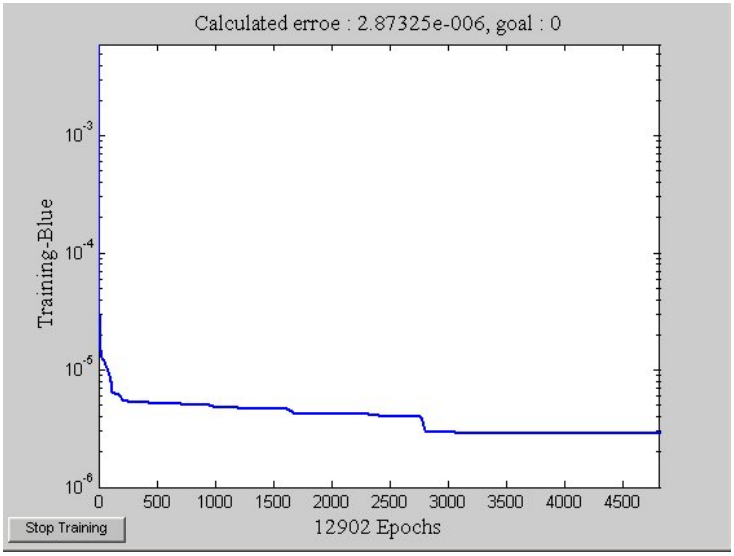
$$\cot\frac{\theta_H}{2} = \frac{1}{N-1}\sum_{i=1}^{N-1}\cot\frac{\theta_H}{2}(i), \quad i = 1 \sim (N-1) \Rightarrow \cot\frac{\theta_H}{2} = 1.709 \tag{21}$$

$$\tan\frac{\theta_V}{2} = \frac{1}{N-1}\sum_{i=1}^{N-1}\tan\frac{\theta_V}{2}(i), \quad i = 1 \sim (N-1) \Rightarrow \tan\frac{\theta_V}{2} = 2.29 \tag{22}$$

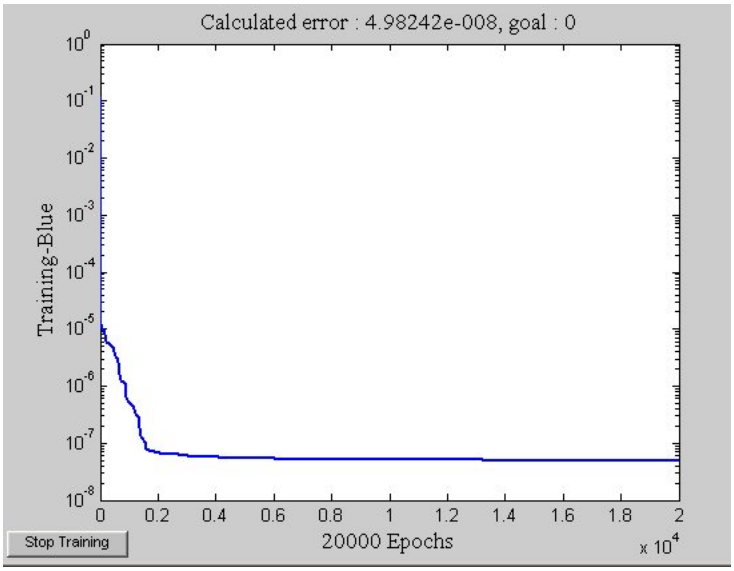
$$D_{op} = \frac{1}{N-1}\sum_{i=1}^{N-1} D_{op}(i), \quad i = 1 \sim (N-1) \Rightarrow D_{op} = 7.87 \tag{23}$$

The parameter, D_{OP}, is 7.87, which means that the distance of D_{OP} is 7.87cm. This distance will cause an error of the final localization result.

4.6. **Precision simulation of the distance measurement.** Using the BPN network, data is required for the two neurons in the input layer (F_H and the approximate distance, D_K) for the near mode and the input layer (F_V and the approximate distance, D_K) for the mid mode. These training data focus on the right side of the field to measure the distance at each block in advance. The total training sample is 130 and the hidden layer is one. Because the absolute coordinate system of the soccer field is invariable, the on-line data can be trained beforehand. The simulation result indicates that the most suitable number of neurons is ten; the learning rate is 0.1 and the output layer is one. After completing the on-line training, the relationship between the information of the frame and the distance is determined. The off-line process is then used for the invariable parameters (relationships) to improve the precision of the distance measurement. The precision can reach 0.12cm in the near mode and 0.005cm in the mid mode is possible. The simulation results are shown in Figure 23.



(a)



(b)

FIGURE 23. The errors between the simulated and real distances are about: (a) 0.12cm in the near mode, (b) 0.005cm in the mid mode

5. **Comparisons and Analysis.** There follows a brief comparison of the results for the IBDMS and the improved IBDMS and the self-localization results, using AVBSLS, for the IBDMS and the improved IBDMS.

5.1. **Analysis of the actual measuring distance.** This study uses IBDMS and improved IBDMS techniques to measure the distance between a robot and a landmark, from 30cm to 400cm. Figure 24 shows the errors between the measured distance and actual distance; the results are listed in Table 6. For distances from 30cm to 400cm, the average and maximum errors for the IBDMS are 7.08cm and 15.0cm, respectively. However, those for the improved IBDMS are 0.82cm and 6.0cm, respectively. The proposed improved IBDMS method significantly improves the accuracy of the distance measurement.

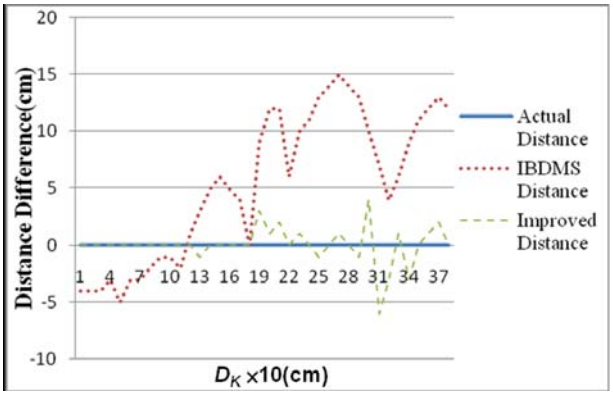


FIGURE 24. The distance differences for the IBDMS, improved IBDMS, and actual distance

TABLE 6. Comparisons of the maximum and average error for different methods

Methods	Maximum error	Average error
IBDMS	15.0cm	7.08cm
Improved IBDMS	6cm	0.82cm

5.2. **Results for the self-localization in the field.** The robot position's was operated by AVBSLS for different measurement methods in the actual field. Since the left and right sides of the field are with the same situation, as shown in Figure 18, there is no loss of generality if the experiment focuses on the right side of the field. Figure 25 shows the measurement results for various locations when the robot uses AVBSLS with the improved IBDMS technique. The stars indicate the various locations of the robot. Table 7 shows a comparison of the robot position using AVBSLS at different measurement distance methods. The accuracy rate for the AVBSLS with the improved IBDMS technique is 92.3%; that for localization using normal IBDMS is only 70.77%. The main reason for this lower accuracy is the inaccuracy in the measurement of distance results.

TABLE 7. Comparisons of the correct rates for different localization methods

Total experimental points = 130			
Method	Correct	Incorrect	Accuracy rate
AVBSLS with normal IBDMS	92	38	70.77%
AVBSLS with improved IBDMS	120	10	92.3%

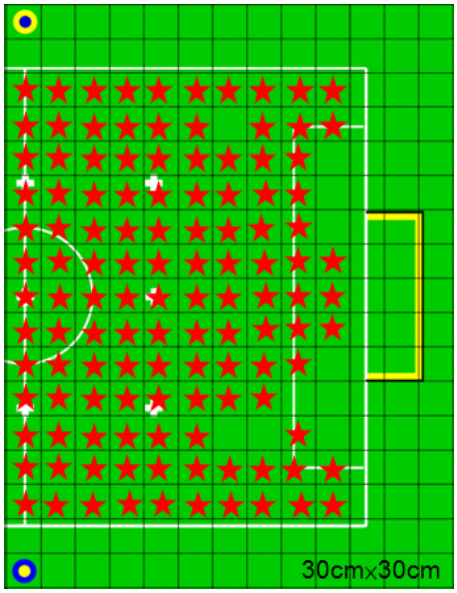


FIGURE 25. The positions operated by AVBSLS with the improved IBDMS technique (the stars are the robot locations on the correct positions at that moment)

6. **Conclusions.** Self-localization is one of the most important problems for a humanoid soccer robot. This research proposes a very efficient vision approach, using an adaptive vision-based self-localization system (AVBSLS), to self-localize the soccer robot during the soccer contest for the 2009 RoboCup rule. The mechanism of the AVBSLS consists of only a CCD camera and a set of pan/tilt motors on the head of the humanoid soccer robot. One of the landmark poles on the soccer field was used as the reference point to correctly self-localize the humanoid robot. The distance between the robot and the landmark was measured, using AVBSLS. The IBDMS (image-based distance measuring system) method was modified to measure the intrinsic parameters of the CCD camera and further to calculate the distance between the robot itself and the landmark. A BPN (back propagation neural network) technique was used to more accurately refine the calculated distance. Using the AVBSLS method with the modified IBDMS technique, the accuracy rate of the self-localization of the humanoid soccer robot was 92.3%. The AVBSLS method is simple to operate, and the processing speed is as high as 15 FPS.

Acknowledgment. This work was supported by the National Science Council of Taiwan under grant number: NSC 98-2218-E-032-003.

REFERENCES

[1] H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda and E. Osawa, Robocup: The robot world cup initiative, *IJCAI-95 Workshop on Entertainment and AI/ALife*, pp.19-24, 1995.
[2] *FIRA RoboWorld Congress*, http://www.fira2009.org.
[3] *RoboCup Soccer Humanoid League Rules and Setup for the 2007 Competition*, http://waziwazi.com/robocup, 2007.
[4] *RoboCup Soccer Humanoid League Rules and Setup for the 2008 Competition*, http://www.robocup-cn.org/, 2008.
[5] *RoboCup Soccer Humanoid League Rules and Setup for the 2009 Competition*, http://www.robocup2009.org/, 2009.
[6] I. Shimshoni, On mobile robot localization from landmark bearings, *IEEE Transactions on Robotics and Automation*, vol.18, no.6, pp.971-976, 2002.

[7] M. Betke and L. Gurvits, Mobile robot localization using landmarks, *IEEE Transactions on Robotics and Automation*, vol.13, no.2, pp.251-263, 1997.
[8] Z.-G. Zhong, J.-Q. Yi, D.-B. Zhao, Y.-P. Hong and X.-Z. Li, Motion vision for mobile robot localization, *IEEE International Conference on Control, Automation, Robotics and Vision*, vol.1, pp.261-266, 2004.
[9] D.-J. Kriegman, E. Triendl and T.-O. Binford, Stereo vision and navigation in buildings for mobile robots, *IEEE Transactions on Robotics and Automation*, vol.5, no.6, pp.792-802, 1989.
[10] S.-K. Choi, J. Yuh and G.-Y. Takashiqe, Development of the omni-directional intelligent navigator, *IEEE Robotics & Automation Magazine*, vol.2, no.1, pp.44-53, 1995.
[11] P.-R. Liu, M.-Q. Meng and P.-X. Liu, Moving object segmentation and detection for monocular robot based on active contour model, *Electronics Letters*, vol.41, no.24, 2005.
[12] J. Kim, K.-J. Yoon, J.-S. Kim and I. Kweon, Visual SLAM by single-camera catadioptric stereo, *SICE-ICASE International Joint Conference*, pp.2005-2009, 2006.
[13] C.-J. Zhang, S.-J. Ji and X.-N. Fan, Study on distance measurement based on monocular vision technique, *Journal of Shandong University of Science and Technology*, vol.26, no.4, pp.65-68, 2007.
[14] Y. Xie and Y.-M. Yang, A self-localization method with monocular vision for autonomous soccer robot, *Computer Science and Information Engineering*, vol.22, no.10, pp.129-132, 2005.
[15] C.-C. Chen, M.-C. Lu, C.-T. Chuang and C.-P. Tsai, Image-based distance and area measurement system, *IEEE Sensors Journal*, vol.6, no.2, pp.495-503, 2006.
[16] C.-C. Hsu, M.-C. Lu, W.-Y. Wang and Y.-Y. Lu, Distance measurement based on pixel variation of CCD images, *ISA Transactions*, vol.48, no.4, pp.389-395, 2009.
[17] W.-H. Chang, C.-H. Hsia, Y.-C. Tai and J.-S. Chiang, An efficient object recognition system for humanoid robot vision, *IEEE International Conference on Ubi-Media Computing*, 2009.
[18] C.-C. Wong, C.-T. Cheng, K.-H. Huang, Y.-T. Yang and H.-M. Chan, Humanoid robot design and implementation for obstacle run competition of FIRA Cup, *CACS International Automatic Control Conference*, pp.876-881, 2007.
[19] Y.-J.-C. Bizais, R.-W. Rowe, I.-G. Zubal, G.-W. Bennett and A.-B. Brill, New approach to 2D linear interpolation for geometric distortion correction of images, *International Conference on Medical Computer Science/Computational Medicine*, pp.170-173, 1982.
[20] J.-T. Dijak, A method for correcting geometric distortion in video cameras, *Proc. of IEEE National Aerospace and Electronics Conference*, vol.2, pp.1382-1388, 1985.
[21] M. Rebiai, S. Mansouri, F. Pinson and B.-B. Tichit, Image distortion form zoom lenses: Modeling and digital correction, *International Conference on Broadcasting Convention*, pp.438-441, 1992.
[22] L.-H. Chiang, *Neural Network – Application of MATLAB*, 7th Edition, Gau-Lih Publish, 2005.
[23] F.-J. Chang and L.-C. Chang, *Artificial Neural Network*, 3rd Edition, Tun-Ghua Publish, 2007.
[24] *Logitech QuickCam® Pro for Notebooks*, http://www.logitech.com/.
[25] *ROBOTIS Dynamixel AX-12*, http://www.robotis.com/zbxe/software_en.
[26] C.-C. Wong, C.-T. Cheng, K.-H. Huang, Y.-T. Yang, Y.-Y. Hu, H.-M. Chan and H.-C. Chen, Humanoid soccer robot: TWNHR-IV, *Journal of Harbin Institute of Technology*, vol.15, no.2, pp.27-30, 2008.