Applied Bionics and Biomechanics 9 (2012) 399–407
DOI 10.3233/ABB-2011-0025
IOS Press

A fast vision system for soccer robot

Tianwu Yang*, Changjiu Zhou and Mohan Rajesh
Advanced Robotics and Intelligent Control Centre, Singapore Polytechnic, Singapore

Abstract. This paper proposes a fast colour-based object recognition and localization for soccer robots. The traditional HSL colour model is modified for better colour segmentation and edge detection in a colour coded environment. The object recognition is based on only the edge pixels to speed up the computation. The edge pixels are detected by intelligently scanning a small part of whole image pixels which is distributed over the image. A fast method for line and circle centre detection is also discussed. For object localization, 26 key points are defined on the soccer field. While two or more key points can be seen from the robot camera view, the three rotation angles are adjusted to achieve a precise localization of robots and other objects. If no key point is detected, the robot position is estimated according to the history of robot movement and the feedback from the motors and sensors. The experiments on NAO and RoboErectus teen-size humanoid robots show that the proposed vision system is robust and accurate under different lighting conditions and can effectively and precisely locate robots and other objects.

Keywords: Vision system, soccer robot, edge detection, object recognition, object localization

1. Introduction

All the RoboCup real robot soccer competitions of different leagues, such as humanoid, standard platform and middle size, take place in a colour coded environment. Generally one or two cameras are used for object detection and localization. Besides cameras, many other types of sensors, such as sonar, infrared, and laser range finder can be used to improve the accuracy of object detection and localization. But the main focus of this paper is how to use camera(s) to detect and locate robots and objects. Our vision system has been implemented in RoboErectus (RE) [5] and Aldebaran NAO humanoid robot [4]. RE humanoid robot has one camera while NAO humanoid robot has two cameras for object recognition, but only one object can be seen by one camera at a time. The soccer robot vision is a typical active vision system in which the viewpoint of the camera(s) is manipulated to investigate the robot soccer field. The applications of active vision include automatic surveillance, SLAM (Simultaneous Localization and Mapping), route planning, human-machine interaction, etc. [2, 6, 11]. In order to reduce the computational complexity, active vision selectively processes the object of interest in the view. The objects of interest in soccer robot vision system include an orange ball, blue and yellow goals, white lines, robots, poles, etc.

Several distinguished features developed in recent years, for example, SIFT, KLT and features inspired by visual cortex, can be used for the object recognition and tracking [7, 9, 10], but a colour-based method is more effective and robust in a colour coded environment. The colour-based object recognition is composed of two steps, i.e., colour segmentation and shape recognition. In colour segmentation, all colours are mapped into several colour classes which are generally defined by a colour cube. The camera grabs image frames in RGB or YUV format, but it is very difficult to find out a set of colour classes in RGB or YUV colour model without any overlap. Several other colour models, such as YIQ, HSV and HSL [1] have been investigated. Based on our observation; the hue components in HSL give the best

*Corresponding author: Tianwu Yang, Advanced Robotics and Intelligent Control Centre, Singapore Polytechnic, Singapore 139651, Singapore. E-mail: yangtw@sp.edu.sg.

segmentation of colour while the environment is not too dark or too bright. Hence, the HSL colour space is adopted, but the saturation component is changed into the colourfulness [8] to achieve a better distinction among light yellow, orange and white colour. Based on the width and length represented by one pixel in the image, the whole image pixels are selectively scanned to find out some of the edge pixels. Other edge pixels are detected by searching around these edge pixels. By using only the edge pixels, all colour blobs, lines and circles in the image frame are found, and the objects are detected according to the feature of each colour blob.

While 26 key points are pre-defined on the soccer field, only the key points seen by the soccer robot are used to calculate the robot and object position in the robot coordinate system (RCS) and also the world coordinate system (WCS). If the soccer robot is equipped with a gyro or tilt sensor, the feedback from these sensors could be integrated with the object localization to reduce the localization error. It is however assumed that the soccer robot is not equipped with these types of sensor. The coordinate of key points in robot camera view is first estimated based on the movement history of robot and the feedback from motors. If two or more key points are seen in the robot camera view, the space relationship of these key points is employed to adjust the rotation angles used in object localization.

The rest of this paper is organized as follows. In Section 2 the object recognition in the colour coded scenario is introduced. The object localization strategy is discussed in Section 3. The performance of the vision system is shown in Section 4. Finally, the conclusion is drawn in Section 5.

2. Colour-based object recognition

The vision system has been implemented in our RE teen-size and NAO humanoid robot, and will be applied in our new designed RE kid-size humanoid robot (Fig. 1). These robots can grab the CIF or QSIF format image at 15–25 frame/s. Due to the limitation of computational power, the complexity level of object recognition is crucial. The computational complexity can be significantly reduced by only scanning the selected pixels. The robustness to light condition can be achieved by using a suitable colour model.

2.1. Modified HSL colour model

As shown in Fig. 2, the colour at the edge of white lines often changes into the light yellow or light orange which has a similar hue and saturation with the goal and pole yellow or the ball orange. In order to distinguish the white lines with the goals, poles and ball,

a. RE Teen-size b. RE Kid-size c. NAO

Fig. 1. RoboErectus and NAO humanoid robot.

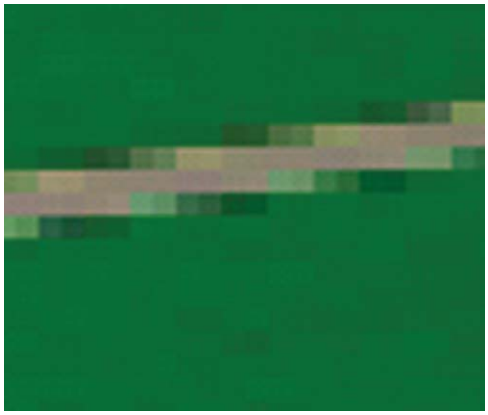


Fig. 2. Colour changes at the edge of white line.

Fig. 3. The scanned pixels on the image.

the saturation in HSL colour model is replaced by the colourfulness.

$$C = \frac{3}{2}\sqrt{(R-A)^2 + (G-A)^2 + (G-B)^2} \qquad (1)$$

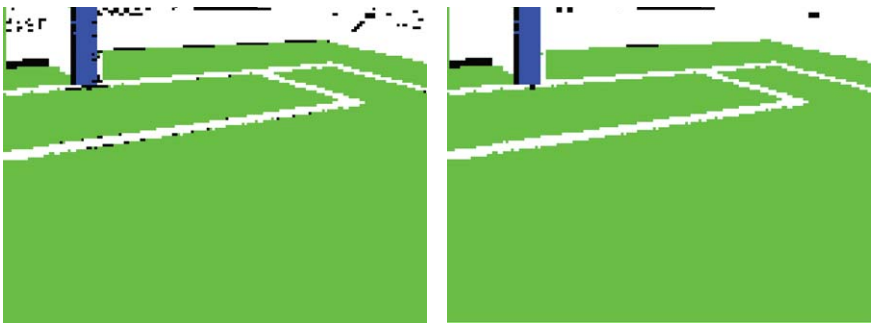
where

$$A = \frac{R+G+B}{3}$$

The colourfulness is directly relative to the distance of one colour to the grey axis in the RGB colour space [8]. The colourfulness of light yellow or light orange is smaller than that of the yellow colour of goal and pole and the orange colour of the ball as well, thus they can easily be distinguished. Using the modified HSL model, a set of colour classes without any overlap can easily be found out. The colour classes are trained offline. A look up table is constructed for fast indexing of a colour in RGB.

2.2. *Object recognition*

All the information of a colour coded image is located in the edges. Considering that all edge pixels are only a tiny part of the whole image, therefore the object recognition based on the edge information can significantly reduce the computational complexity. In edge detection, only small part of whole image pixels is selected to scan. As shown in Fig. 3, the distance of two scanned pixels is always less than 4 cm such that all edge pixels could not be missed. Once an edge pixel is detected, all the pixels (the red segment in Fig. 3) between the two scanned pixels are further searched to obtain the exact edge position. After the first scanning, only small parts of the edge pixels are detected. By searching around these edge pixels, all the other edge pixels are found out. In fact, the colour of some edge pixels frequently changes into other colour which does not belong to any colour classes. These pixels are called isolating pixels, which should be eliminated to speed



a. Before b. After

Fig. 4. The elimination of the isolating pixels. a) Before, b) After.

up the processing; Fig. 4a and b show the edges before and after eliminating the isolating pixels respectively.

With the edge information, the line segments can be easily found. After merging the line segments of the same colour, all the colour blobs are found out. During the merging of the line segments, some features, such as the colour, size and coordinates of the centre, upper right and lower left of colour blob, are collected for each colour blob. The object is then determined based on these features and the posture of the camera. For example, the biggest orange blob is regarded as the ball.

2.3. Lines and ball centre

The performance of a soccer robot is strongly influenced by the accuracy of object localization. Lines are detected mainly for the robot and object localization. All white lines, pole and goal line are detected in our vision system. In general, Hough transform is applied to estimate the straight lines and other shapes. However it is time-consuming and impractical for robot soccer vision system. A fast method is proposed to detect all straight lines. The following part explains how to detect all white lines in an image in detail.

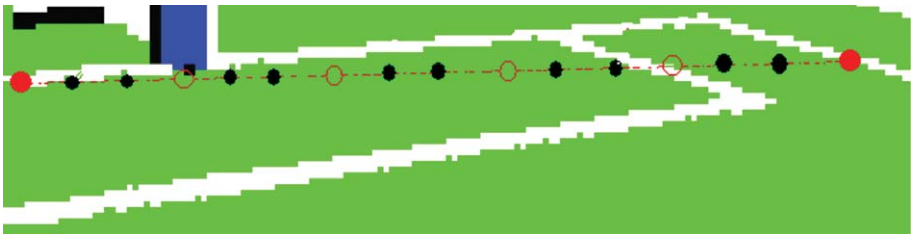


Fig. 5. The checking points of supposed line.

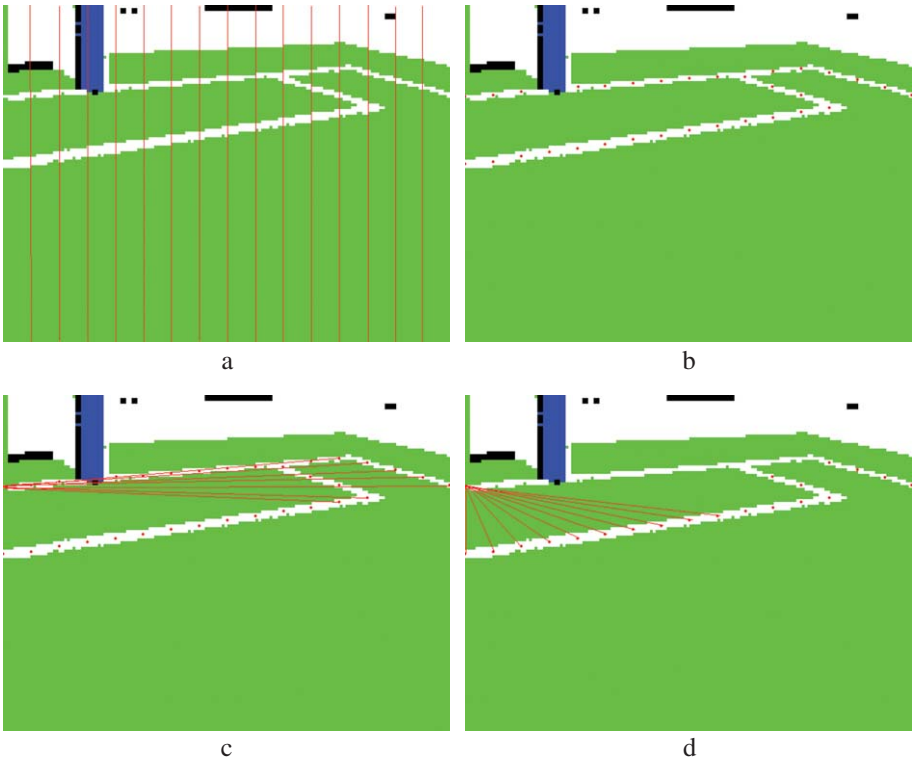


Fig. 6. Illustration of line detection.

The main procedure of the proposed white line detection is to check whether a supposed line is a white line. As shown in Fig. 5, the two red points fix a supposed white line. Two points sets are predefined, i.e. point set $\mathbf{S_1}$ (four hollow red points) and $\mathbf{S_2}$ (ten solid black points) respectively. The four points in $\mathbf{S_1}$ are checked firstly. If less than two points are on the white line, the supposed line is not a white line. Otherwise ten points in $\mathbf{S_2}$ is checked. By considering the effects of partial occlusion, the supposed line is regarded as a white line if more than half of the points in $\mathbf{S_1}$ and $\mathbf{S_2}$ are on the white line.

The proposed white line detection selects some vertical or horizontal lines and scans them from top to bottom and left to right for finding out a set ($\mathbf{S_0}$) of pixels on the lines (Fig. 6a and b).The first and last pixels in $\mathbf{S_0}$ is chosen as the start and end point of the supposed line. If the supposed line is checked and is not a white line, the end point is moved backward to another pixel until the supposed line is regarded as a white line (Fig. 6c). All points on the white lines except the start point are removed from the $\mathbf{S_0}$ and the end point is continually moved backward over the rest points in $\mathbf{S_0}$ until all the points are checked (Fig. 6d). Finally, the start point is moved forward and the checking is repeated until all points are checked. The detailed algorithm of white line detection is given below.

The circle checking method is similar to the line checking. The line segment connecting the two chosen edge pixels in the ball area is assumed to be a diameter (Fig. 7) and a supposed circle is drawn. After checking all the supposed circles, the diameter and ball centre are detected.

Algorithm 1: Detect the white lines

```
Scan the selected lines;
S0 = All intersections of the scanned lines with white lines;
num = Total number of points in S0;
visit[0 . .num-1] = fa lse;
 for ( i = 0 ; i < num ; i++)
 {
     if(visit[i]) cont inue;
    for(j = num -1 ; j > i ; j--)
    {
        if(visit[j]) cont inue;
        L=the supposed line connecting S0[i] and S0 [j];
        if( !IsWhiteLine(L) ) cont inue;
        for( k = i+1 ;k < num ; k++)
            if(S0 [k] is on L) visit [k] =true;
        Output the white line L.
    }
 }
```

Fig. 7. The checking points of the supposed circle.

3. Object localization

The target of object localization is to estimate the position and posture of our robots and the position of ball, opponent robots and other objects in the camera, robot and world coordinate system which are denoted as $\mathbf{X_C Y_C Z_C}$, $\mathbf{X_R Y_R Z_R}$, and $\mathbf{X_W Y_W Z_W}$ coordinate systems respectively (Fig. 8). The origin of $\mathbf{X_C Y_C Z_C}$ and $\mathbf{X_R Y_R Z_R}$ is the camera focus. The axis $\mathbf{Y_R}$ is perpendicular to the ground; and the axis $\mathbf{Y_C}$ is the optical axis. Twenty six key points, marked as the red dots and black dots in Fig. 8, are pre-defined. The red key points in the image can be directly recognized based on the object recognition results, but the recognition of the black key points which are the intersections of white lines needs a reference to the history of robot position and posture (see Fig. 9).

As shown in Fig. 10, point P in the image plane is the projection of point Q in the soccer field. The coordinate $\begin{bmatrix} P_{x,R} & P_{y,R} & P_{z,R} \end{bmatrix}^T$ of point P relative to the robot coordinate system is given by

$$\begin{bmatrix} P_{x,R} \\ P_{y,R} \\ P_{z,R} \end{bmatrix} = R_C \begin{bmatrix} P_{x,C} \\ P_{y,C} \\ P_{z,C} \end{bmatrix} \tag{2}$$

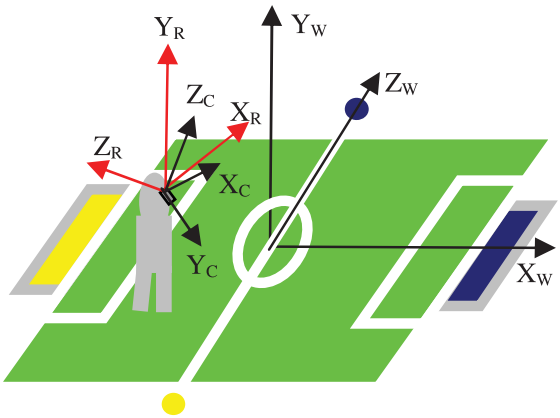


Fig. 8. The coordinate systems used in the object localization.

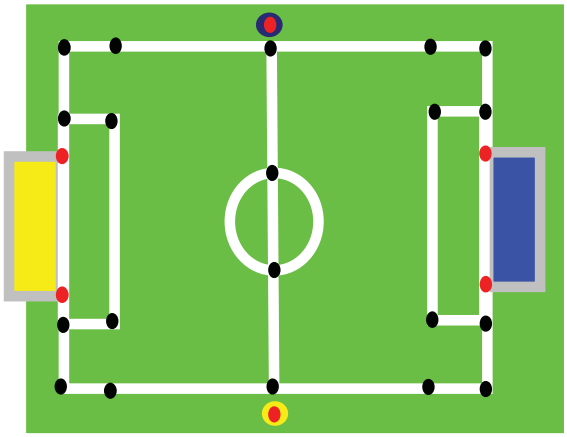


Fig. 9. The position of key points.

where the rotation matrix $\boldsymbol{R_C}$ is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix}$$

$\begin{bmatrix} P_{x,C} & P_{y,C} & P_{z,C} \end{bmatrix}^T$ is the coordinate of point P in relation to the camera coordinate system. α and β are the rotation angle around the axis $\mathbf{Y_C}$ and $\mathbf{X_C}$ respectively. Based on the pinhole camera model, the coordinate of pixels $\boldsymbol{Q}$ in the robot coordinate system is calculated by

$$\begin{bmatrix} Q_{x,R} \\ Q_{y,R} \\ Q_{z,R} \end{bmatrix} = -\frac{R_h}{P_{y,C}} \cdot \begin{bmatrix} P_{x,C} \\ P_{y,C} \\ P_{z,C} \end{bmatrix} \tag{3}$$

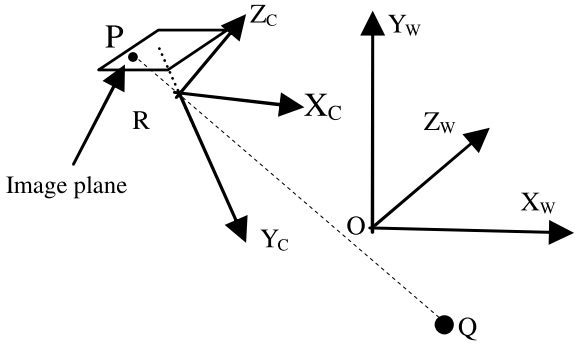


Fig. 10. The projection of the robot camera.

where R_h is the distance from the ground to camera focus. The coordinate of point Q is estimated by

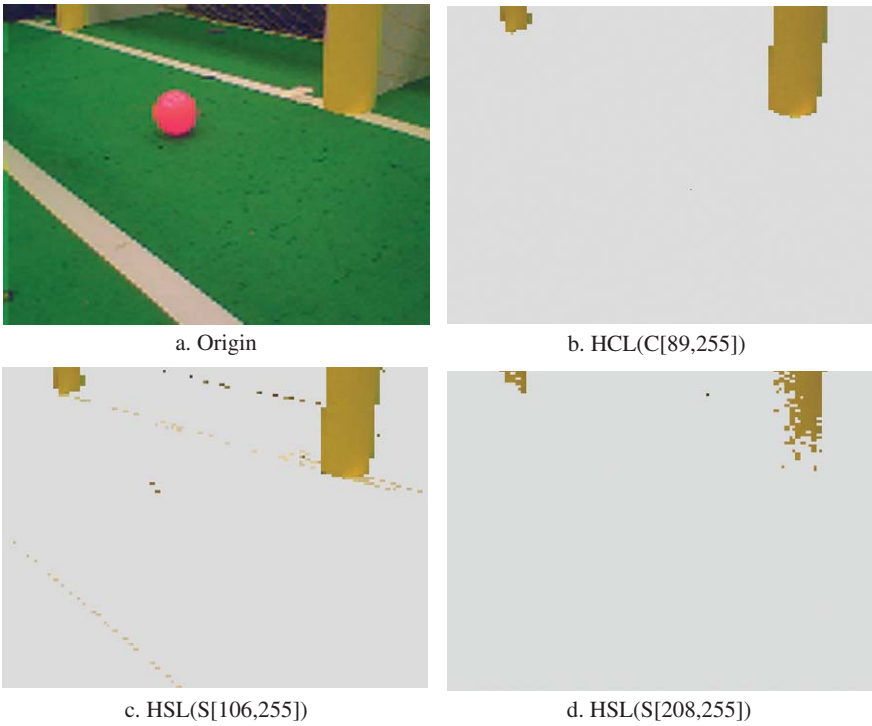
$$\begin{bmatrix} Q_{x,W} \\ Q_{y,W} \\ Q_{z,W} \end{bmatrix} = \begin{bmatrix} \cos\gamma & 0 & \sin\gamma \\ 0 & 1 & 0 \\ -\sin\gamma & 0 & \cos\gamma \end{bmatrix} \begin{bmatrix} P_{x,I} \\ P_{y,I} \\ P_{z,I} \end{bmatrix} + \tau \tag{4}$$

where β is the rotation angle around the axis Y_R. τ is a translation vector form the camera focus to the origin of the world coordinate system. If α, β, and γ is known and pixels Q is a key point, then τ can be obtained, namely the robot location in the world coordinate system.

Generally, α, β and γ are estimated based on the movement history and feedback of the motors and sensors. After several frames, the estimation of α, β and γ becomes inaccurate and will obviously affect the accuracy of the object localization, and thus the adjustment of rotation angles using the robot camera is necessary. In order to do the adjustment, at least two key points must appear in robot view. If only two key points are seen, the rotation angle β is adjusted. The adjustment angle $\Delta\beta$ is repeatedly estimated by

$$\Delta\beta = \arctan\left(\frac{\tilde{d}_{mid}}{R_h}\right) - \arctan\left(\frac{\hat{d} \cdot \tilde{d}_{mid}}{R_h \cdot \tilde{d}}\right) \tag{5}$$

until the estimated distance of two points is very near to its real value. $\tilde{d}_{mid}$ is the estimated distance from the centre of two key points to the origin of the world coordinate system. $\hat{d}$ and $\tilde{d}$ are the real and estimated distance between two key points respectively. If three or more key points are seen, the rotation angle β is adjusted first. Then some pairs of the rotation angle α and γ are evenly selected within the possible range of α and γ. For each rotation angle pair, each key point is used to estimate a translation vector, and several



a. Origin b. HCL(C[89,255])

c. HSL(S[106,255]) d. HSL(S[208,255])

Fig. 11. Colour segmentation results of yellow colour.

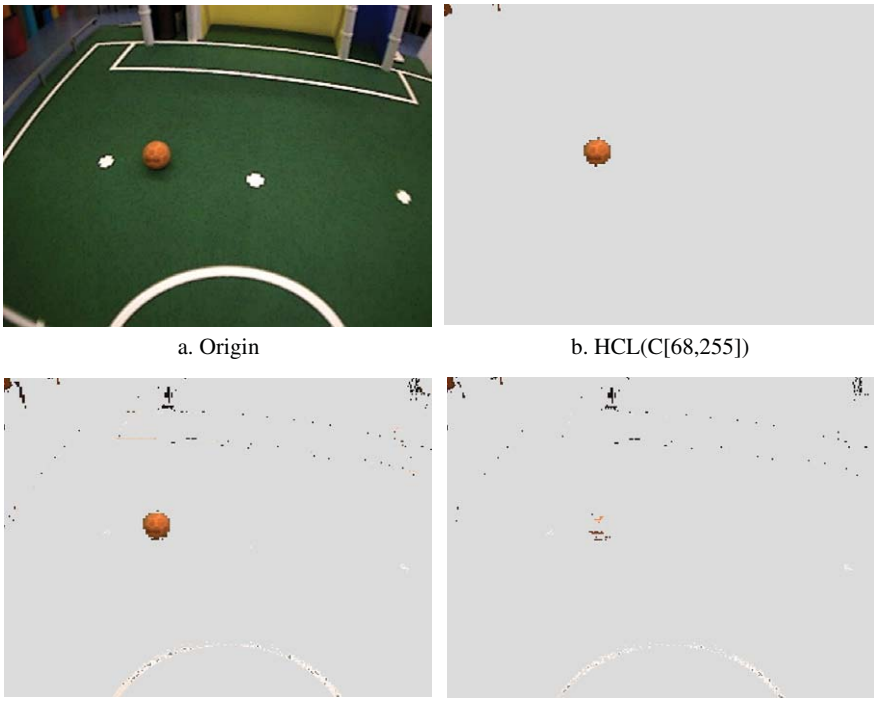
different translation vectors are obtained. The variance of translation vectors is calculated. The best rotation angle pair which yields the minimum variance is the new estimation of rotation angle α and γ.

4. Experimental results

The experiments have been carried out on the NAO robot and RE teen-size humanoids. The NAO uses an on-board AMD Geode 500 MHz processor. It has two cameras, but only one camera can be used at a time. Each camera can grab the QCIF image up to 25 frame/s. The RE teen-size uses a Sony VAIO 1.33 GHz ultra portable PC for robot strategy and vision processing. Its camera has a field of view of 77° by 56° and captures the SIF image at 25 frame/s.

4.1. HCL colour space

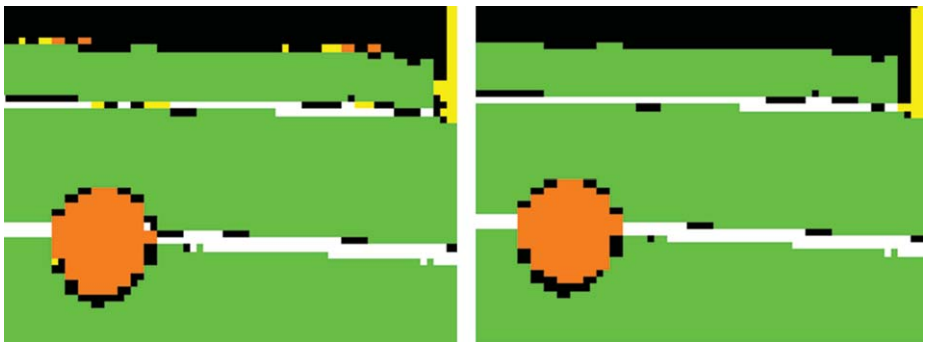
The performance of colour segmentation in HSL colour model outperforms the RGB, YUV and many other colour models. The proposed HCL colour model modifies the saturation in HSL to colourfulness in order to achieve a better colour segmentation. The range of each component in the HSL and HCL is set to [0,255]. Due to the different lighting-condition and shadow, the luminance changes dramatically. The comparison between the HCL and HSL is compared by setting the range of luminance to [0,255]. The best range of hue is then tuned for two colour spaces. The range of saturation in HSL and the colourfulness in HCL is adjusted to obtain the best colour segmentation. Figure 11 shows the colour segmentation results of the yellow colour. The yellow colour of door can be easily separated in HCL (Fig. 11b) where the range of colourfulness is [89,255]. The Fig. 11c shows the colour segmentation results in HSL where the range of the saturation is [106...255]. Many pixels on white line and net are classified into yellow colour. If the minimum saturation of the yellow colour increases, some pixels on the yellow door begin to disappear. Until it reaches 209, all pixels on white line and net can not be classified into yellow colour, but only less than 50% pixels on yellow door are correctly classified. The Fig. 12 demonstrates that the colour segmentation results of the orange ball in



a. Origin b. HCL(C[68,255])

c. HSL(S[86,255]) d. HSL(S[220,255])

Fig. 12. Colour segmentation results of orange colour.



a. HSL b. HCL

Fig. 13. Colour segmentation results in the HCL and HSL colour model.

the HCL are much better than that in HSL. A suitable range of the orange colour class can not be found in the HSL.

Although the performance of the HSL could be improved if the luminance is considered in the colour tuning, the problem can not be solved totally and it is very difficult to tune the luminance. The HCL always give the better colour segmentation and the colour tuning can easily be done. The best colour segmentation with the consideration of luminance is shown in Fig. 13. Many pixels on the white lines and net are classified into yellow or orange colour class in HSL, these pixels may be detected as the ball or other objects. At the same time, the total number of the detected colour blobs will increase, thus the computational complexity will increase too. The vision processing time for one

Table 1
The vision processing time of one frame in the HCL and HSL colour space

Colour space	Min (ms)	Average (ms)	Max (ms)
HSL	7.6	9.5	12.1
HCL	6.8	8.6	11.2

Table 2
The performance of the object localization

Object	Distance to the object (m)	Avg. error in distance (m)	Max. error in distance (m)
Ball	1.0	0.036	0.052
	2.0	0.052	0.069
	3.0	0.069	0.089
Pole	1.0	0.066	0.092
	2.0	0.089	0.113
	3.0	0.116	0.142

frame with the two colour spaces in RE teen-size robot is listed in Table 1.

4.2. Object localization

The performance of object localization in robot coordinate system is evaluated. The robot is placed in several random positions around the soccer field with a fixed distance to the objects, such as the ball and pole. The results of object localization performance are listed in Table 2. The results of object localization in robot coordinate system can be further improved by employing the Monte Carlo localization to locate a robot in the walking situation. The maximum error in distance is about 0.163 meter [3]. The estimation accuracy of the camera position significantly impacts the object localization performance. After the adjustment of the rotation angles, the accuracy of object localization in RCS is significantly improved. The average error in distance is within 0.15 meter range. The error depends on the distance to the object and also the object itself.

5. Conclusion

In this paper, a fast vision system for soccer robot is presented. The vision system is able to detect all the objects on the soccer field at high speed and is robust under different lighting conditions. The rotation angles are adjusted based on the key points detected in the image frames, thus the accuracy of object localization is improved and is accurate enough for RoboCup soccer competition.

References

[1] M.K. Agoston, *Computer Graphics and Geometric Modeling: Implementation and Algorithms*, Springer, London, 2005, pp. 300–306.
[2] C. Breazeal, A. Edsinger, P. Fitzpatrick, et al., Active vision for sociable robots, *IEEE Trans on Man, Cybernetics, and Systems, Part A: Systems and Humans* **31**(5) (2001), 443–453.
[3] W. Hong, C. Zhou, and Y. Tian, Robust monte carlo localization for humanoid soccer robot, *Proceedings of International Conference on Advanced Intelligent Mechatronics*, Singapore, July 14–17, 2009, pp. 934–939.
[4] (2011) [Online]. Available: http://www.aldebaran-robotics.com/en/
[5] (2011) [Online]. Available: http://www.robo-erectus.org/robots.php
[6] V. Ila, J.M. Porta and J. Andrade-Cetto, Information-based compact pose SLAM, *IEEE Trans Robotics* **26**(1) (February 2010), 78–93.
[7] D.G. Lowe, Object recognition from local scale-invariant features, *Proceedings of the International Conference on Computer Vision* **2** (1999), 1150–1157.
[8] H. Palus, Colourfulness of the image and its application in image filtering, *Proceedings of International Symposium on Signal Processing and Information Technology*, Athens, Greece, December 21, 2005, pp. 884–889.
[9] T. Serre, L. Wolf, S. Bileschi, et al., Robust object recognition with cortex-like mechanism, *IEEE Trans Pattern Analysis and Machine intelligence* **3** (March 2008), 411–426.
[10] J. Shi and C. Tomasi, Good features to track, *IEEE Conference on Computer Vision and Pattern Recognition* (1994), 593–600.
[11] J. Solà, A. Monin, Michel, et al., Fusing monocular information in multicamera SLAM, *IEEE Trans Robotics* **24**(5) (October 2008), 958–968.