行政院國家科學委員會專題研究計畫　成果報告

具備全景式視覺系統之行動機器人群隊的研究

計畫類別：個別型計畫
計畫編號：NSC91-2213-E-032-005-
執行期間：91 年 08 月 01 日至 92 年 10 月 31 日
執行單位：淡江大學機械與機電工程學系

計畫主持人：王銀添

計畫參與人員：劉秋豪,游智鈞

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 93 年 2 月 3 日

行政院國家科學委員會補助專題研究計畫 ☑ 成 果 報 告
□期中進度報告

具備全景式影像系統之行動機器人群隊的研究

計畫類別：☑ 個別型計畫　　□ 整合型計畫
計畫編號：NSC91－2213－E－032－005
執行期間：2002 年 8 月 1 日至 2003 年 10 月 31 日

計畫主持人：王銀添
共同主持人：
計畫參與人員：劉秋豪、游智鈞

成果報告類型(依經費核定清單規定繳交)：☑精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、
列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：淡江大學機械與機電工程學系

中　華　民　國 九十二 年 十一月 二十一 日

行政院國家科學委員會專題研究計畫成果精簡報告

具備全景式影像系統之行動機器人群隊的研究

Research on Formations of Mobile Robots with Panoramic Vision Systems

計畫編號：NSC91-2213-E-032-005

執行期限：91 年 8 月 1 日至 92 年 10 月 31 日

主持人：王銀添 淡江大學機械與機電工程學系副教授

計畫參與人員：劉秋豪、游智鈞 淡江大學機械工程學系碩士生

Abstract: A mobile robot with an on-board panoramic vision system is developed in this research, and utilized to play soccer games. The panoramic vision system includes a CCD facing upwards beneath the convex mirror and taking the picture reflected in the mirror. The mirror is tilted 23° to get a larger field of vision in front of the robot. We propose an image mapping method based on the fuzzy system theory to obtain a 3D ground coordinate from a 2D panoramic image. An effective procedure is constructed to set the color range of the object in the image. The integrated system is tested on a robot soccer field to evaluate the practical usage of the developed system.

Keywords：Mobile robot, Soccer robot, Robot vision, Panoramic vision

1. MOBILE ROBOT WITH A PANORAMIC VISION

A mobile robot with an on-board vision has the capability to capture and process the image locally. However, due to the limitation of the view-angle, an on-board vision system equipped in front of the robot cannot take the picture of the whole surrounding. A panoramic vision system is developed to overcome the limitation. Panoramic vision makes it possible to cover a 360° filed of vision, by analyzing only one image. The ideal of panoramic vision was firstly proposed by Rees (1970). Recently, Nayar (1999) has geometrically analyzed the complete class of single-lens single-mirror panoramic vision systems and developed an ideal vision system using a parabola mirror. In this paper, a panoramic vision with the mirror tilted 23° is implemented to get more image in front of a mobile robot and provide the mobile robot with a fast tracking capability.

The panoramic vision system is composite of two parts: a spherical mirror is utilized to reflect the image around the robot, and a CCD faces upwards beneath the mirror and take the picture reflected in the mirror, as shown in Figures 1 and 2. The image taken by the panoramic camera is distorted. Therefore, it has to develop a technique to correct the distortion and obtain an unwarped image. Gaspar and Victor (1999) solved the distortion problem using a look-up table. However, their method required a large amount of computer memory for mapping an asymmetric image. The proposed method in the paper is based on the fuzzy system theory and the concept of clustering methods.

Motion control system of the wheeled mobile robot can be divided into three levels, namely, self-localization, action selection and path planning, and contour control. This research investigates the problems including self-localization and contour control, and the issue of action selection and path planning will be discussed by the other paper. The algorithm proposed by Marques and Lima (2000) is applied to locate the position of the mobile robot. Their method used natural geometric landmarks of the environment.

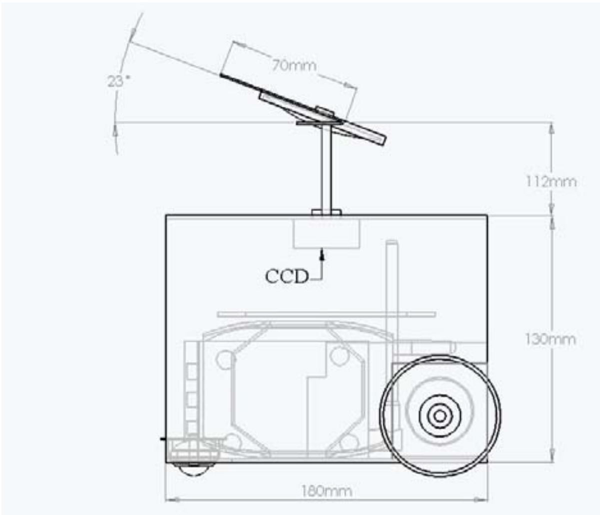

Figure 1 Design of mobile robot with on-board vision

Figure 2 Implementation of the Robot with a panoramic camera

2. TRANFORMATION OF PANORAMIC IMAGE

In the panoramic vision system, an object in 3D-space is reflected by a convex mirror and projected on a 2D image. The image is taken by a CCD and the digitized image is feedback to control system. In the robot servoing control process, the controller has to determine the object's coordinate in 3D-space from the 2D image. The inverse coordinate transform from 2D image to 3D coordinate can be derived by using the image formation theory of convex mirror. Firstly, determine the procedure of projecting a space point $P(x,y,z)$ in Cartesian coordinate onto a image point $P_i(\psi_i,r_i,z_i)$ in a cylindrical coordinate, and then derive the inverse transformation. The notations ψ, r, and z represent the angle, radius, and height of the cylindrical coordinates, respectively. The space point $P(x,y,z)$ can be expressed in cylindrical coordinate as

$$P=[\psi \quad r \quad z]^T=[\arctan(\frac{y}{x}) \quad \sqrt{x^2+y^2} \quad z]^T$$

Assume that a spherical mirror is utilized in the panoramic vision system, as shown in Figure 3, so the equation of image formation can be determined as the following expression.

$$z_m=-\tan(\frac{\pi}{2}-\beta)\cdot r_m+L \tag{1}$$

$$z_m^2+r_m^2=R^2 \tag{2}$$

$$2\arctan(\frac{r_m}{z_m})=\arctan(\frac{r-r_m}{z-z_m})-\beta \tag{3}$$

where r_m and z_m are the coordinates on the spherical mirror; R is the radius of the spherical mirror; L is the length from the center of the spherical mirror to the projection center; β is the angle of projected point. Solve the equations (1)-(3), and the expression of r_m, z_m, and β can be determined. The coordinates of a point on the image can be obtained by the perspective projection model,

$$r_i=\alpha_r f\tan\beta\cos\psi+r_{i0} \tag{4}$$

$$z_i=\alpha_z f\tan\beta\sin\psi+z_{i0} \tag{5}$$

$$\psi_i=\psi \tag{6}$$

where f is distance from the image plane to the projection center. α_r and α_z are the image scale factors; (r_{io},z_{io}) is the principal point of the image coordinate system. From the equations (1)-(6), the coordinate of the planar point $P_i(\psi_i,r_i,z_i)$ can be determined. Since these equations are nonlinear function, it is impossible to solve these equations on-line and in real-time. It is determined by a look-up table in the literature (Gaspar and Victor 1999). The distortion ratio ζ of the panoramic image in Figure 5 is a function of the radius r_i, therefore a table describing the function, $\zeta(r_i)$, can be used to determine the coordinate inversely.

In the paper, the spherical mirror is tilted 23° to get more image points in the front of the robot, and the image taken by the CCD is shown in Figure 5. In this case, the distortion ratio of the image is asymmetric generally. Therefore the coordinate can not be simply determined by a look-up table and the perspective projection model is more complex than that of a flat mirror.

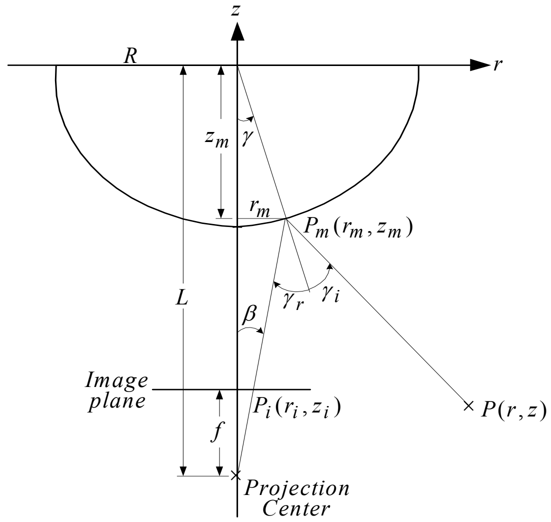

Figure 3 perspective projection model of a spherical mirror

3. FUZZY IMAGE MAPPING

In order to improve the disadvantage of complexity in the perspective projection model of a convex mirror, a Takagi-Sugeno (TS) fuzzy system model is proposed based on the clustering method (Wang 1997). Assume that there are a group of input-output data pairs, $x=(x(1), x(2), \ldots, x(p\text{-}1), x(p))$, in which $x(1), x(2), \ldots, x(p\text{-}1)$ represent $(p\text{-}1)$ input variables of the system, and $x(p)$ is the output variable. For a TS fuzzy system with c fuzzy rules, the rule base of the system can be described as the following expression (Wang 1997):
Rule m ∶ if x is $A_m(x)$ then y_m is

$$y_m=b_{m0}+\sum_{j=1}^{p-1}b_{mj}x(j) \quad m=1,2,\ldots,c \tag{7}$$

Where $A_m(x)$ is the antecedent proposition of the fuzzy rule; b_{mj} are coefficients; y_m is the consequence of the fuzzy rule. The proposition $A_m(x)$ can be viewed as the fuzzification from the original input data; while the consequence of the fuzzy rule, y_m, is set to be a linear combination of the input variables. If the fuzzy system has c fuzzy rules, the output of the TS fuzzy system can be obtained by the weighted average method (Wang 1997),

$$y=\frac{\sum_{m=1}^{c}A_m(x)\cdot y_m}{\sum_{m=1}^{c}A_m(x)}=\sum_{m=1}^{c}g_m\cdot\left[b_{m0}+\sum_{j=1}^{p-1}b_{mj}x(j)\right] \tag{8}$$

$$g_m=\frac{A_m(x)}{\sum_{j=1}^{c}A_j(x)}\ ,\ m=1,2,\ldots,c$$

As mention in the preceding section, the number of rules, c, in the rule base of the TS fuzzy model affects the performance of the system. If the fuzzy system contains too many rules than it needs, the system becomes a complex system. On the contrary, if the system has too few rules, it would not have enough information to model the real system. By the clustering method, the number of rules is decided by adjusting the relational grade of the clusters. It will be defined latterly, the relational grade as a Gaussian function, which is based on the distance between two vectors. In the procedure of clustering methods, the collected data is divided into several subsets or strings of data according to the relational grades between these data. Therefore, each string of data has its own characteristics that can be distinguished from other data strings. The procedures to determine the structure of the fuzzy system are described as follows.

Assume a set of n vectors in a p-dimensional space,

$$X = \{x_1, x_2, \cdots, x_n\}$$

where $x_i=(x_i(1), x_i(2),\ldots, x_i(p))$ is a vector with p variables, in which $(x_i(1), x_i(2),\ldots, x_i(p\text{-}1))$ and $x_i(p)$ are input and output variables of i^{th} data point, respectively. Among these n vectors, the vectors which have a high relational grade can be collected to be a string named cluster. In this paper, the concept of similarity proposed by Wong (Wong and Chen 1999) is utilized to determine the relational grade. According to the method, first select a data point as a reference vector, and then find a comparative vector that has high relational grade with the reference vector. Further, choose the comparative vector with high relational grade as a new reference vector, and repeat the procedure. By this recursion method, the reference vector can be replaced during each cycle of the procedure, and eventually converges to the center of a cluster. The procedure is summarized in five recursive steps (Wong and Chen 1999):

Step 1: define n movable vectors v_i $(i = 1,2,\ldots.,n)$ and let $v_i=x_i$, where x_i is the initial value of v_i;

Step 2: calculate the similarity by the following equation,

$$r_{ij} = \exp\left(-\frac{\left\|v_i - v_j\right\|^2}{2\sigma^2}\right), \quad i = 1,2,\ldots,n \;\; ; \;\; j = 1,2,\ldots.,n \quad (9)$$

where r_{ij} represents the relational grades between the reference vector v_i and the comparative vector v_j; $\left\|v_i - v_j\right\|$ is the Euclidean distance between v_i and v_j; and σ is the width of the Gaussian function in Equation (9);

Step 3: modify the relational grades between the reference vector v_i and the comparative vector v_j according to the following rule,

$$r_{ij} = \begin{cases} 0 & \text{if } r_{ij} < \xi \\ r_{ij} & \text{otherwise} \end{cases}$$

where ξ is a small constant set up to be 0.01 in this paper;

Step 4: calculate a new vector set $v_i' = (v_i'(1), v_i'(2), \ldots., v_i'(p))$, $i = 1,2,\ldots.,n$,

$$v_i'(k) = \frac{\sum_{j=1}^{n} r_{ij} v_j(k)}{\sum_{j=1}^{n} r_{ij}}, \quad k = 1,2,\ldots.,p$$

Step 5: if all the vectors v_i' are the same as v_i, i=1, 2, …, n, then stop; otherwise let $v_i = v_i'$, i=1, 2, …, n, and go to Step 2.

By this procedure, the data points with high relational grades are collected as a cluster. The relational grade is modified in Step 3 to prevent the movable vector from being affected by the vectors with low relational grades. The movable vectors will gradually converge to a vector of values. Therefore, the number of convergent vectors is the number of clusters, and the convergent vector is viewed as the center of the corresponding cluster. By the clustering method, n input-output data are divided into c clusters,

$$c_m = \{c_m(1), c_m(2), \ldots, c_m(p)\}, \quad m = 1,2,\ldots,c$$

Once the cluster centers are determined, the antecedent proposition, $A_m(x_i)$, of the i^{th} input, x_i, can be arranged according to the relationship between i^{th} input data and m^{th} cluster center. A Gauss function (Wang 1997) is chosen to represent the membership function,

$$A_m(x_i) = \exp\left(-\frac{\sum_{k=1}^{p-1}(x_i(k) - c_m(k))^2}{2\delta_m^2}\right) \quad (10)$$

$$\delta_m = \sqrt{\frac{-\sum_{k=1}^{p-1}\left(x_m^*(k) - c_m(k)\right)^2}{2\ln(\alpha)}}, \quad m = 1,2,\ldots,c$$

where $m = 1,2,\ldots,c$; $i = 1,2,\ldots,n$; c_m is the m^{th} clustering center; δ_m indicates the width of the Gaussian function in Equation (10); x_m^* is the most far away data point of the m^{th} clustering data; α is a constant between 0 and 1.

Using the procedure of the clustering method, n data points are distributed into c clusters. According to Equation (8), the output of the TS fuzzy system with n input-output data points and c cluster centers can be expressed as the following equations,

$$y_i = \sum_{m=1}^{c}\left[g_{im}b_{m0} + g_{im}\sum_{j=1}^{p-1} b_{mj}x_i(j)\right] \quad (11)$$

$$g_{im} = \frac{A_m(x_i)}{\sum_{j=1}^{c} A_j(x_i)}, \quad m = 1,2,\cdots,c; \quad i = 1,2,\cdots,n$$

Define the vectors and matrices containing input and output variables as the following equations,

$$Y = [y_1 \quad y_2 \quad \cdots \quad y_n]^T$$

$$h = \begin{bmatrix} b_{10} & b_{11} & \cdots & b_{1p-1} & \cdots & b_{c0} & b_{c1} & \cdots & b_{cp-1} \end{bmatrix}^T$$

$$W = \begin{bmatrix} g_{11} & g_{11}x_1(1) & \cdots & g_{11}x_1(p-1) & \cdots & g_{1c} & g_{1c}x_1(1) & \cdots & g_{1c}x_1(p-1) \\ g_{21} & g_{21}x_2(1) & \cdots & g_{21}x_2(p-1) & \cdots & g_{2c} & g_{2c}x_2(1) & \cdots & g_{2c}x_2(p-1) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n1}x_n(1) & \cdots & g_{n1}x_n(p-1) & \cdots & g_{nc} & g_{nc}x_n(1) & \cdots & g_{nc}x_n(p-1) \end{bmatrix}$$

The coefficient vector, h, can be determined using the pseudo-inverse matrix method (Strang 1980) or the RLSE method (Phillips and Nagle 1995). Once the coefficients are determined, the fuzzy system in Equation (11) can be used to model the perspective projection system.

The environment surrounding of the robot is arranged as Figure 4, and a TS fuzzy system is developed for the perspective projection model. In the figure, the black dots are utilized to mark the area of the soccer field. One image was taken by the panoramic vision system and shown in Figure 5. The developed TS fuzzy system is employed to represent the mapping of one pixel on the warped image to the unwarped image. The results are listed in Table 1. Several different widths of the Gaussian function, σ, are chosen for the fuzzy system, and result in different number of clusters for the fuzzy system. When the value of σ decreases, the Root-Mean-Squared-Error (RMSE) of the system output will decrease. However, the complexity of the system will increase. As an example, a square block on the soccer field is twisted as shown in Figure 6, and the mapped result by the fuzzy system is shown in Figure 7. The RMSE errors are listed in Table 2.

Table 1 Results of the fuzzy system model

	σ	Rules	RMSE	Matrix W
x	1.67	9	0.3912	$9p\times 9n$
	0.12	11	0.3642	$11p\times 11n$
	0.09	19	0.3044	$19p\times 19n$
y	2.21	7	0.9551	$7p\times 7n$
	0.17	13	0.4941	$13p\times 13n$
	0.16	16	0.4585	$16p\times 16n$

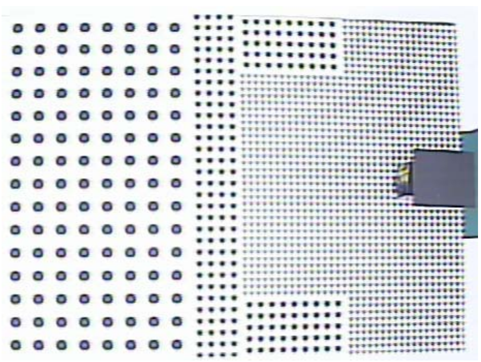

Figure 4 Black circle marks in soccer field

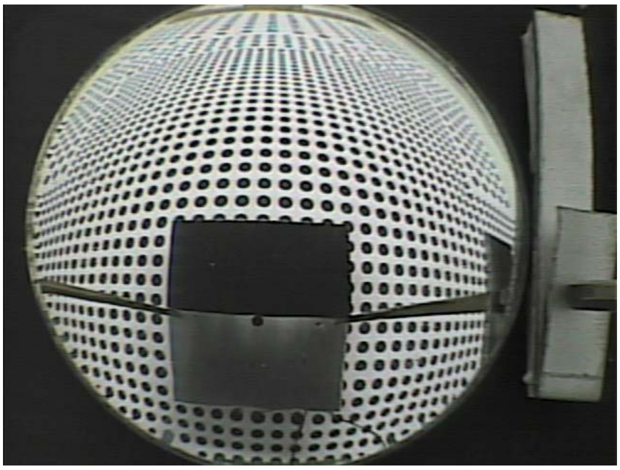

Figure 5 Image taken by a panoramic camera

Table 2 RMSE of the fuzzy system

	x (σ=0.12)	y (σ=0.17)
RMSE	0.5520	0.5768

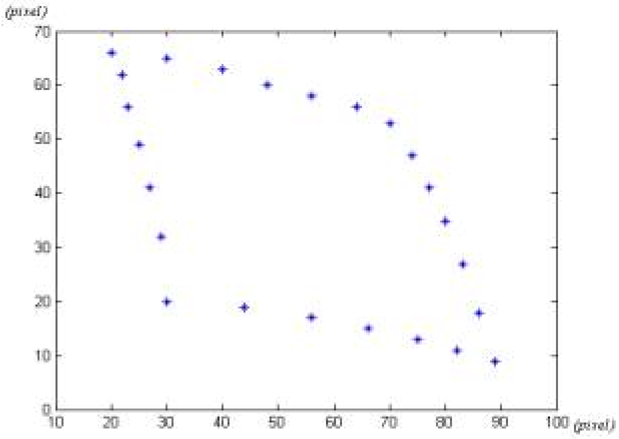

Figure 6 A square area on the panoramic image

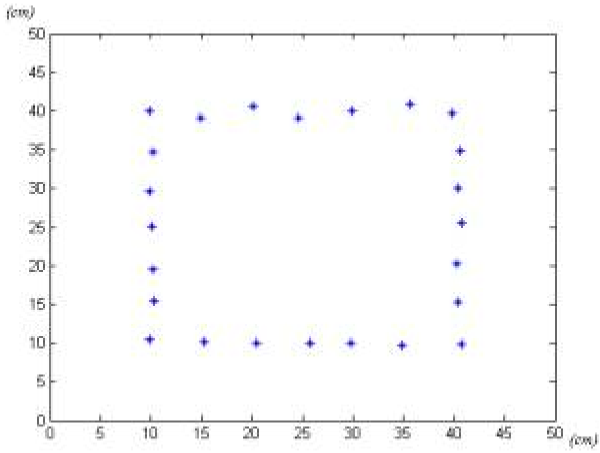

Figure 7 Unwarped image of the square area

4. COLOR MODEL

The purpose of a color model is to allow convenient specification of colors within some color gamut. The red, green, blue (RGB) color model is hardware- oriented and used with color CRT monitors. The RGB model employs a Cartesian coordinate system and the subset of interest is the cube shown in Figure 8. Unfortunately, the RGB color model is not easy to use, because it does not relate directly to intuitive color notions of hue, saturation, and brightness. Therefore, another class of models has been developed with ease of use as a goal (Foley *et al.* 1990). The HSV model is utilized in the paper. By contrast to the RGB model, the HSV (hue, saturation, value) model is user-oriented, being based on the intuitive appeal of the artist's tint, shade, and tone (Smith 1978, Foley *et al.* 1990). The HSV coordinate system is cylindrical, and the subset of the space within which the model is defined is a hexcone, or six-sided pyramid. Hue is measured by the angle around the vertical axis. The value of S is a ratio ranging from 0 on the center line to 1 on the triangle sides of the hexcone. Saturation is measured relative to the color gamut represented by the model. The RGB-to-HSV mapping is defined as follows (Smith 1978):

$$H = \cos^{-1}\left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} \right) \quad (12)$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \quad (13)$$

$$V = \frac{\max(R,G,B)}{255} \quad (14)$$

The algorithm of Figure 9 is an approximate calculation for Equation (12)-(14). Alternative method can be used to convert RGB to HSV by creating a look-up table. However, the total computer memory required for the table is about 48 megabytes (=256×256×256×3bytes).

For robot object and path tracking, it is needed to set a specific color range for each of the tracked object. A simple method for setting a color range is depicted in Figure 10.

5. OBJECT TRACKING CONTROL

The image mapping method developed in the preceding sections are integrated with a mobile robot and tested in a soccer filed to track a ball. The result is shown in Figures 11 and 12. The figure records the continuous motion of the robot approaching to the object.

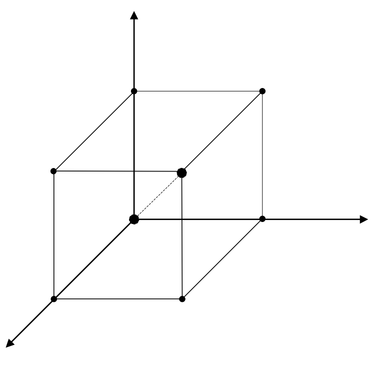

Figure 8 The RGB color model

```
Void RGB_To_HSV(int r, int g, int b, int &h, int &s, int &v)
//Given: r, g, b, each in [0,255].
//Desired: h, s, v in [0,255] except if s=0, then h=UNDEFINED, which
// is some constant with a value outside the interval [0,255].
{
    // r*=255; g*=255; b*=255;    If r, g, b are given in [0,1]
    int max = maximum(r,g,b);
    int min = minimum(r,g,b);
    v = max;   //This is the vale v.
    //Next calculate saturation, s.
    if (max <> 0)  s = (max – min)*255/max;        //s is the saturation.
    else s = 0;     //Saturation is 0 if red, green and blue are all 0.
    if (s == 0)    { h = UNDEFINED;}
    else   //Chromatic case: Saturation is not 0, so determine hue.
    {   delta = max – min;
        if (r == max)
        { h=(g–b)/delta; //Resulting color is between yellow and magenta. }
        else
        { if (g == max)
            {h=2+(b–r)/delta;//Resulting color is between cyan and yellow. }
            else
            { if (b == max)
                {h=4+(r–g)/delta;  //Resulting  color  is  between  magenta  and
cyan.}
            }
        }
        h = h * 60;                          //Convert hue to degrees.
        if (h < 0) {h = h +360;}             //Make sure hue is nonnegative.
        h*=255/360;                  //Convert hue to be in [0,255].
    }       // End of chromatic case.
}       //End of RGB_To_HSV.
```

Figure 9 RGB-to-HSV algorithm

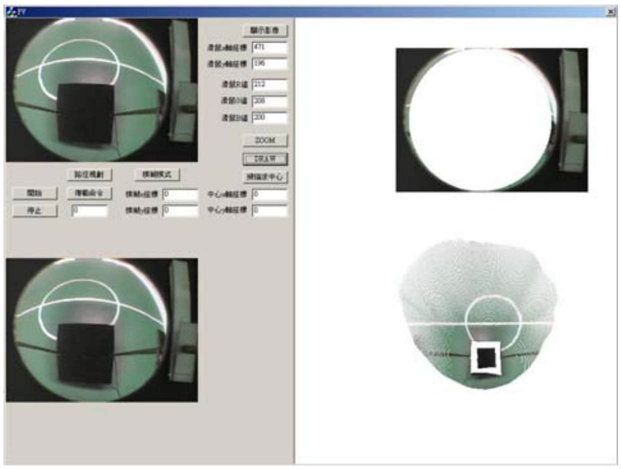

Figure 10 Color range setting

6. CONCLUSION

A mobile robot with an on-board panoramic vision system is developed in this paper. The complex problem of mapping a pixel on a 2D image to a 3D ground coordinate is solved in this paper by a simple and real-time procedure. The integrated system was tested on a robot soccer field, and the results indicate that the proposed method is effective and accurate.

ACKNOWLEDGEMENTS

This work is supported by the National Science Council in Taiwan under grant no. NSC91-2213-E-032-005.

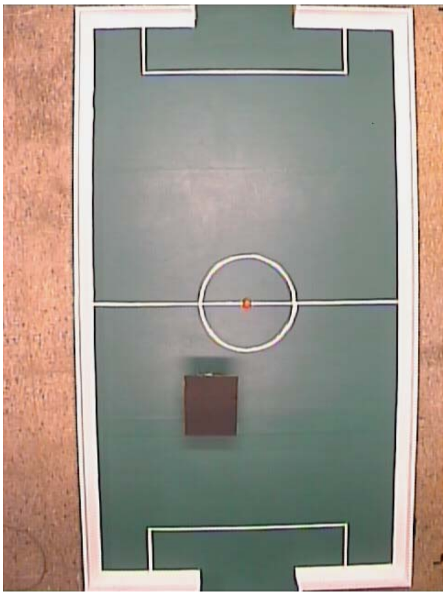

Figure 11 Robot in a reduced-size soccer field

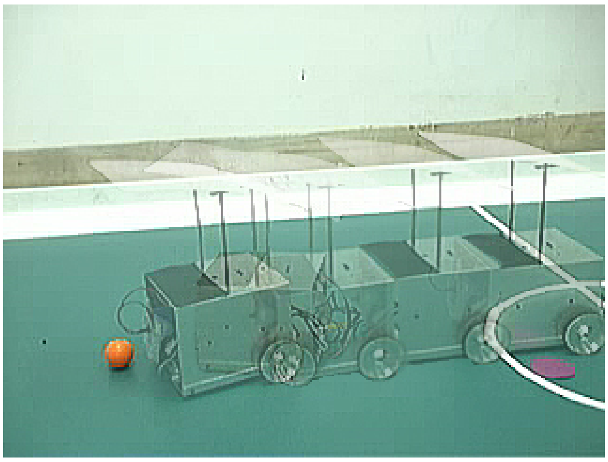

Figure 12 Robot tracking a ball

REFERENCES

Bräunl, T. (2002). Department of Electrical and Electronic Engineering, University of Western Australia, Perth, Australia, http://robotics.ee.uwa.edu.au/

Gaspar, J. and J.S. Victor (1999). Visual path following with a catadioptric panoramic camera, *International Symposium on Intelligent Robotic Systems*, pp.139-147.

Nayar, S.K. (1997). Catadioptric Omnidirectional Camera, *IEEE International Conference on Computer Vision and Pattern Recognition*, pp.482-488.

Phillips, C.L. and H.T. Nagle (1995). *Digital Control System*, 3rd ed., Prentice Hall.

Rees, D.W. (1970). Panoramic television viewing system, *United States Patent*, No. 3,505,465.

Strang, G. (1980). *Linear Algebra and Its Applications*, 2nd ed., Academic Press.

Sugeno, M., and G. T. Kang (1988). Structure identification of fuzzy model, *Fuzzy Sets and Systems*, vol.28, no.1, pp.15-33.

Takagi, G.T., and M. Sugeno (1985). Fuzzy identification of systems and its applications to modeling and control, *IEEE Transactions on Systems, Man and Cybernation*, vol.15, no.1, pp.116-132.

Wang, L.X. (1997). *A Course in Fuzzy System and Control*, Prentice-Hall.

Winters, N., J. Gaspar, G. Lacey and J.S. Victor (2000). *IEEE Workshop On Omnidirectional Vision*.

Wong, C.C. and C.C. Chen (1999). A Clustering- Based Method for Fuzzy Modeling, *IEICE Transaction on Information and System*, vol.E82-D, pp.1058-1065.