Proceedings of the 2005 IEEE
International Conference on Mechatronics
July 10-12, 2005, Taipei, Taiwan

Robust Color Classification for Global Soccer Vision

Yen-Hsun Wu[1] and Han-Pang Huang[2]
Department of Mechanical Engineering, National Taiwan University, Taipei, Taiwan R.O. C.
[1]r91522831@ntu.edu.tw and [2]hanpang@ntu.edu.tw
[1]Graduate student [2]Professor

Abstract—The task of global vision module is to extract meaningful data for the strategy decision module. According to these data, the decision-making module estimates the field condition, and then plans strategies to offense or defense. The data extracted should be reliable and accurate for strategy decision module so as to plan efficient tactics. Robust color classification plays a dramatic role in analyzing the scene based on pre-defined color classes. In addition, appropriate color classification can reduce the computational time and improve the reliability of extracted data by eliminating the uninterested background information. In this paper, Principal Component Analysis (PCA) is adopted to seek for a color subspace. In this color space, a color classification model can be constructed straightforward. By using this model, colors slightly varied can be robustly classified.

I. INTRODUCTION

Micro Robot World Cup Soccer Tournament (MiroSot) is one of the soccer robot competitions held by Federation of International Robot-soccer Association (FIRA). Reference [10], the regulated system architecture is shown in Fig. 1 taken from the rules of MiroSot. Each team has one global vision camera, one host computer, communication modules, and soccer robots. In each match, the global vision module grabs the field and analyzes the scene; the host computer decides the most efficient strategy; the communication modules deliver instructions to the robots; the robots receive the comments and move automatically. When the competition starts, no human contact, no matter direct or indirect, is allowed.

The MiroSot game provides a platform on which a wild range of techniques could be integrated and experimented. For example, the objects on the soccer field should be tracked by the global vision module; the strategy module should plan an ideal path to goal; the instructions should be transmitted reliably by the communication module; the mechanism of robots should be well designed to play the soccer game.

In order to handle the rapidly-changed situation on the soccer field, the shorter latency of each module gives the more chance for the team to win the game. Moreover, to deal with the dynamic environment, the reliability of each module also plays an important role. This paper is focused on the global vision module employed by the NTU-Formosa soccer robots.

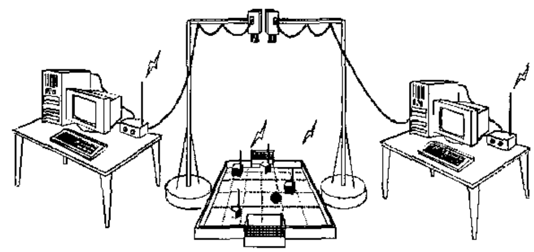

Fig. 1 The overall system architecture of MiroSot. A match is played by two teams. Each team consists of robots, host computer, global vision module and communication module.

A. *Global Vision in MiroSot*

The field and lighting conditions of MiroSot are also regulated in the rules of MiroSot. The field is a black (non-reflective) wooden rectangular playground. The marks on the field are all painted in different intensity of gray color. The lighting condition in the competition site should be fixed around 1,000 Lux. However, the lighting condition level would slightly vary in different competition duration, and may be non-uniform at different regions on the soccer field.

The global vision module is the most significant part to sense the soccer field. The global vision system acts as the eye of the robots on the playground. According to the data extracted from the vision module, the decision mechanism will plan the most suitable offense or defense strategy. The robots on the playground all obey the instructions delivered from the strategy mechanism. Hence, the accuracy and precision of the playground information are the most important factor to form efficient strategy and to win the game.

The following section describes the complete architecture of NTU-Formosa in further detail. Section II.C describes the approaches adopted in the global vision module of NTU-Formosa. Section IV shows the results that illustrate the performance of the global vision module. Section V draws the conclusions to illustrate the application of this approach in other robotic domains.

II. ARCHITECTURE OF NTU-FORMOSA

The soccer robot system of NTU-Formosa is composed of four subsystems: soccer robots, global vision module, strategy decision mechanism, and communication modules. The global vision module first shoots the soccer field and analyzes this scene. After useful information has been extracted, the strategy decision module plans an ideal path for each teammate robots. Once the strategy has been decided, the communication module transmits the instructions to each teammate robot. These soccer robots receive their own instructions and move to corresponding locations on the soccer filed. This procedure from the vision module to robots constitutes one operation cycle.

A. Global Vision Module

The global vision module provides the strategy decision module with object information on the soccer field. The basic information should include the position of ball; the position of opponent robots; the position and orientation of teammate robots. It is important that the information delivered to the strategy decision module is reliable, accurate, and with minimal latency.

While the rules of MiroSot regulate the lighting condition around 1,000 Lux, the lighting level can be slightly varied for different competition in practice. Moreover, variations could possibly take place as well for different regions of the soccer field. Different lighting condition can enormously effect the appearance of colors in most color space, such as RGB or HIS.

The rules also define some preserved color types on the field: orange, blue and yellow. Orange color belongs to the ball. The blue and yellow colors are used to mark the robots in different teams. However, the object colors at different locations on the field may be nonuniform because of the slightly varying lighting conditions on different field regions.

The vision module of NTU-Formosa consists of three parts. The first part, including a CCD camera and a image grabber, shoots the field and delivers samples to the memory of the host computer over PCI bus. The off-line processor, which creates the color classes for on-line recognition, forms the second part. The on-line processor, which analyzes the grabbed images and throws data to the decision-making module, forms the third part.

B. Input Specification of Vision Module

The global vision system of NTU-Formosa contains one JAI progressive CCD color camera, CVM-70, and one Metero-II frame grabber card. The camera shoots a frame by progressive scan, and then the progressive video signal is translated to RGB values. The RGB values are transmitted to the frame grabber card on the host computer by cable.

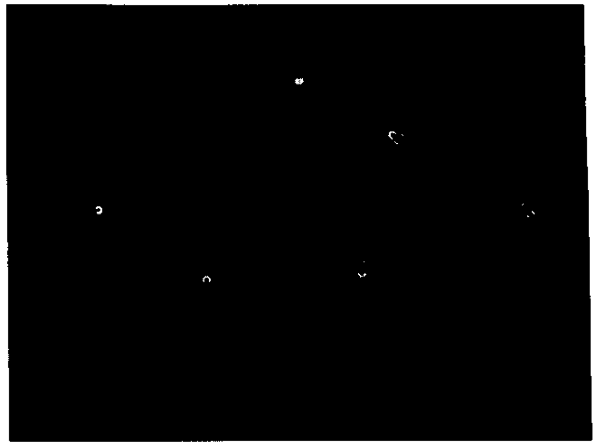

Fig. 2 A sample frame of soccer field is grabbed by NTU-Formosa vision module. Each team has three robots. The robots with blue marks belong to the blue team, and with yellow marks belong to the yellow team. The orange pattern presents the ball. The field size is 150 x 130 cm^2, and the top view of robots is 7.5 x 7.5 cm^2.

Table 1 The output data of Fig. 2. The origin is at top left corner of the grabbed image, and The image size is 640 x 480 pixels. In Fig. 2, the yellow team is opponent, and the blue team is our team. The positions of all objects are recorded in pixel, and angles of teammate robots are in degree.

	Ball		
X (pixel)	506.929		
Y (pixel)	315.500		
	Oppnent 1	Opponent 2	Opponent 3
X (pixel)	433.231	569.214	387.000
Y (pixel)	148.000	224.500	286.000
	Teammate 1	Teammate 2	Teammate 3
X (pixel)	92.154	219.250	322.727
Y (pixel)	227.000	311.500	74.000
Angle (degree)	36.1158	-46.4851	130.3498

The card provides 640 pixels for the width of the field, and 480 pixels for the height of the field. The field size is 150 x 130 cm^2, and the top view of robots is restricted in 7.5 x 7.5 cm^2. Each pixel is about 2.3 x 2.3 mm^2, and each robot size is about 900 pixels. Fig. 2 shows a sample image grabbed in regulated lighting condition. Three blue blobs belong to blue team robots, and three yellow blobs belong to yellow team robots. The orange blob is the soccer ball. The field is non-reflective black, and marks are white lines or gray points.

C. Output Specification of Vision Module

The output data of global vision module are fed to the decision-making module. To plan an ideal path for each teammate robot, the decision-making module should monitor the situation on the field as detailed as possible. Color objects in the 3 vs. 3 MiroSot are one orange golf ball and six soccer robots, three for each team. The robots have two attributes, position and orientation. The attribute of ball is only position.

Since the calculated path is based on the information output from the vision module, the accuracy and reliability of output data are quite demanded. The output data are displayed in Table 1.

III. COLOR-BASED IMAGE ANALYSIS

The key to analyze color images is to choose a color space suitable for extracting the data from the images. How to choose a color space, which provides enough distinguishable characteristics, depends on the properties of the input images. Many color spaces have been developed and each of them is adopted in different applications. Reference [5], conversions between each color space also had been derived. After a suitable subspace has been chosen, a color model is required to classify the colors.

Many approaches of color classification model construction have been developed. Reference [3], Reference [4], Reference [6], Reference [7], and Reference [8], some of them set several color cube models in the color space, or set a specified prototype vector. It is often difficult to construct a color model for non-uniform soccer field when setting color cubes in the color space. Reference [9], one method is to construct the color classification model by analyzing the color histograms. The threshold is calculated independently for each axis of the color space.

In MiroSot, the input images are formed by the background field, the ball, and the robots. The background field and marks on the field are, respectively, regulated as black and gray in different intensity. The ball is colored in orange, and the robots are marked by blue and yellow in different teams.

The color space used to extract the information of the MiroSot field should at least provide the basic color clusters: black and gray for background and marks; orange for ball; blue and yellow for robots. Obviously, the black and gray cluster is the largest one because of its proportion of the overall soccer field. From the RGB input image, the color value of each pixel is plotted on RGB color space, as shown in Fig. 3.Black, orange, blue and yellow points indicate, respectively, the pixels of background, ball, and robots.

A. *Color Classification Space Determination*

Reference 0, Due to plenty of quantities, about 300,000 color points in each grabbed image, of the input color pixels, the approach—Principal Component Analysis (PCA)—will be adopted to filter out background pixels. PCA looks for principal axes which span the space best representing the data of all samples in a least-square sense.

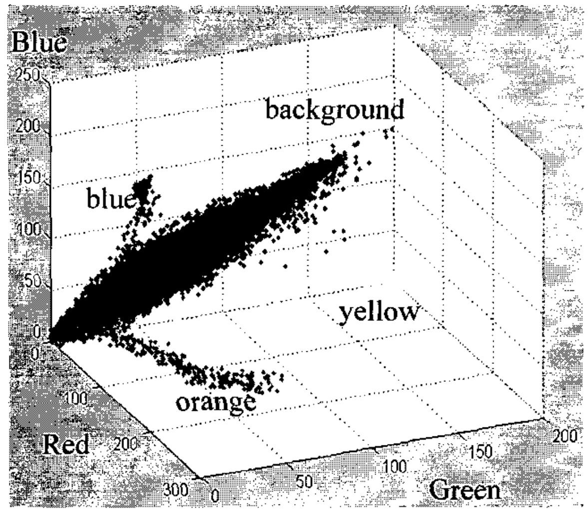

Fig. 3 The RGB values of every pixel in the input image are plotted in the RGB color space. The background cluster is marked by black points. The ball cluster is marked by orange points. The yellow and blue points are the robot clusters.

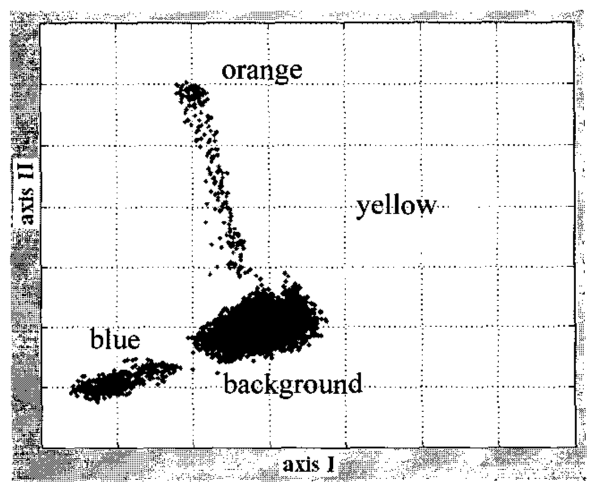

Fig. 4 The distribution of the background and pre-defined colors in the subspace. The black points represent the background color. The other color points denote the pre-defined colors: orange, yellow, and blue. The direction of axis I and axis II are the two eigenvectors that correspond to the smallest and the middle eigenvalues, respectively.

$$Sw = bw \qquad \text{eq. III-1}$$

For removing the background pixels which provide nothing but static field conditions, a scene with no color objects on the playground is grabbed first. The pixels in this frame are treated as the background samples, and are used to construct a subspace for color classification. These background samples, denoted by $x_1 \ldots x_n$, are all 3-dimensional points in RGB space. The mean vector m of the background samples is defined as eq. III-2. n is the total number of samples. The covariance matrix of the samples is defined as eq. III-3.

To find the principal vectors of the background field, the eigen-value equation eq. III-1 of the covariance matrix should be solved. b denotes the eigne-value, and w denotes the eigen-vector. The direction of the principal axis in which the samples most widely distribute is the same as the eigenvector corresponding to the largest eigenvalue.

$$m = \frac{1}{n}\sum_{k=1}^{n} x_k \qquad \text{eq. III-2}$$

$$S = \sum_{k=1}^{n} (x_k - m)(x_k - m)^T \qquad \text{eq. III-3}$$

The best space to filter out the background points is the one in which these points distribute most closely. The space is spanned by the two eigenvectors that correspond to the smallest and the middle eigenvalues. After being projected onto the space, the points can be easily enclosed with a simple geometry model, like a circle or ellipse. The samples of background and pre-defined colors are projected onto the subspace. The distribution is shown in Fig. 4.

B. Color Classification

The clusters of the color objects are classified by different discriminant functions. Fig. 5 shows the histogram of the background and the color samples of objects on the soccer field. The color clusters without the background can also be enclosed by simple models. The discriminant functions can be chosen as the contour of these simple models. According to these discriminant functions, the classified samples in Fig. 6 are marked as different clusters by different color points.

C. Color Look-Up-Table

In order to speed up the processing time, the color look-up-table (LUT) for color recognition is constructed. After the most suitable subspace for filtering out the background points is obtained and the models for classifying the color clusters are defined, RGB valued points in each grabbed frame will be projected on the color plane found by PCA. According to the discriminant function of different clusters, each point, i.e. all colors, in the RGB space will be classified into different color clusters. During the operation cycle, the color LUT can shorten the computational time of color object tracking and recognition.

D. Blob Analysis

Since the color table has been constructed, pixels belonged to the same color class can be congregated. Collection of each classified color is an array which stores the pixel coordination. By connected component procedure, properties of the same color blobs can be obtained. Reference [1], this morphology approach is similar to the connected regions method. In order to merge different row of the same blob, the first step is to mark components in the same row by blob numbers. Second, to check components of the adjacent row with those of the previous row can merge the neighbor components by the same blob number. If overlap is detected, all blob numbers will be rearranged to proper components.

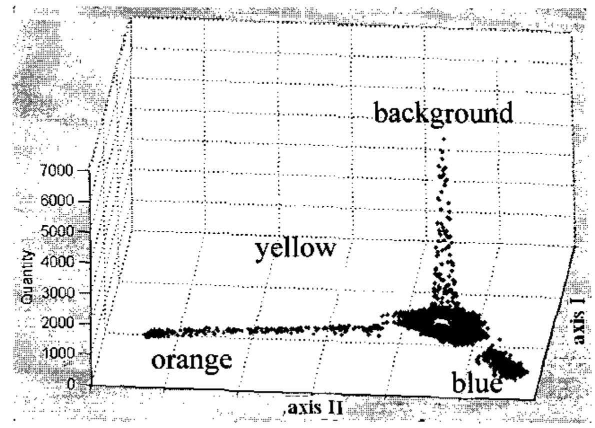

Fig. 5 The histogram of the background and the color samples of objects on the soccer field. The z-axis denotes the quantity. The direction of axis I and axis II are the same as Fig. 4.

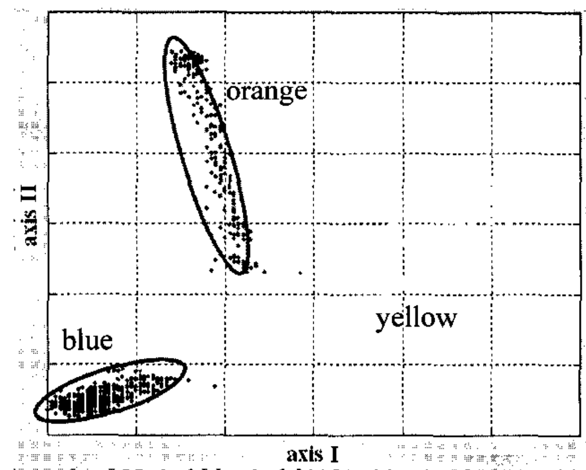

Fig. 6 The classified samples marked by different color dots. The discriminant functions are also marked by different color contours. The direction of axis I and axis II are the same as Fig. 4.

This procedure can divide the color pixels collection into several color blobs. These color blobs are checked against minimum mass constraints to eliminate any noise blobs. The centroids of these color blobs are recorded for soccer objects detection.

E. Information of Ball and Robots

The color blobs contain the information on the overall soccer field. The orange, blue, and yellow blobs denote, respectively, the position of ball, teammate robots, and opponent robots. In the vision module of the NTU-Formosa, two other colors, green and magenta, are used to represent the orientation and member identification. Each magenta blob is assigned to the closest teammate color blob. The number of magenta blobs denotes the identification number of each teammate color blob. Each teammate blob centroid is considered as the identification number robot position. The green blobs are also assigned to proper teammate color blobs, and the direction from each teammate color blob centroid to its own green blob centroid represents the orientation of this teammate. The robot patterns are shown in Fig. 7.

IV. Experimental Results

The global vision module of NTU-Formosa consists of three parts: a Meteor-II frame grabber, a JAI CVM70 CCD camera, and a host computer. The frame grabber converts input signal data of the camera into RGB value data structure. The frequency of CCD camera is operated at 30 frames per second, and the resolution is 640 x 480. The CPU of host computer is Pentium 4 2.8GHz.

The goal of vision module should feed precise field information to strategy mechanism as fast as possible during the soccer competition. Based on the information, the strategy decision mechanism plans the overall strategy. Therefore, the accuracy of the information dominates the efficiency of the offense or defense strategy. The performance can be determined in terms of accuracy, and latency.

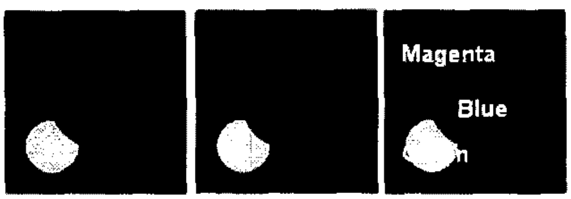

Fig. 7 The patterns of NTU-Formosa soccer robot. The center circle is marked by blue or yellow as teammate color. Each green circle denotes the orientation of corresponding robot. The magenta circles represent the identification number of the robots.

A. Accuracy

All information on the playground is recognized by the vision module, and transmitted to the strategy decision mechanism. The module accurately seeks for the position of the color objects on the field and the orientation of each teammate robot. Once the color pixels have been classified Fig. 8, the vision module can extract all useful information straightforward. The result of the processed image is shown in Fig. 9. Each pixel on the scene is painted by different colors. The black pixels represent the background; the red pixels denote the ball color; the yellow and blue pixels mark the team colors; the green and magenta pixels display the assistant color of the NTU-Formosa vision module. The information of position and orientation is also marked as white circle and arrow, respectively.

B. Latency

The algorithm of the module used to classify the colors and construct the look-up-table (LUT) is only off-line operated. After constructing the LUT, the vision module can reduce the computational time by filtering out the uninterested background pixels in each frame. This procedure will reduce the quantity of original data, about 300,000 (640 x 480) pixels, down to the number about 3,000 pixels.

Table 2 Time consuming table. The most time consuming step is globally color classification of every pixels in one grabbed image.

	Consuming time (ms)
Meaningful color classification	10.5524
Blob information analysis	1.7551
Soccer object detection	0.0400
Total	12.3475

Fig. 8 The color classification result of grabbed image. Each pixel on the scene is painted by different colors. The black pixels represent the background; the red pixels denote the ball color; the yellow and blue pixels mark the team colors; the green and magenta pixels display the assistant color of the NTU-Formosa vision module.

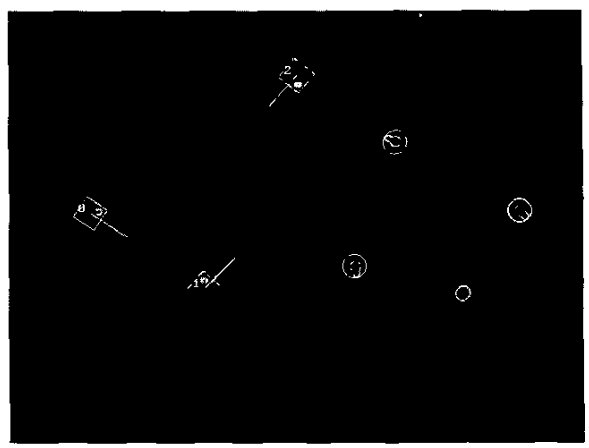

Fig. 9 The result of processed image. The information of position and orientation is also marked as white circle and arrow, respectively.

The module processes one frame while the frame grabber grabs the next one, and the strategy mechanism receives the extracted data form the last one. The first step of the processing procedure is to classify overall image pixels. According to the color classification model constructed off-line, coordination of pixels with the similar color values are collected into the same array. Blobs are connected by comparing collected pixels. Finally, soccer objects can be detected from the blob information. The total latency between the event and the reception of the data on the strategy mechanism is in the order of 12 ms.

V. Conclusion

Choosing a suitable color space is a critical stage in color-based image processing. Colors on the soccer field will change under different lighting condition, therefore to analyze the color distribution of the grabbed images can provide the most appropriate color space for meaningful color clusters classification. According to this subspace, the color clusters can be separated by simple discriminant functions, and it is straightforward to detect the color objects on the field.

The global vision module described here performs all requirements in the robot soccer competition with excellent results. The proposed algorithm can find the most appropriate color space for color classification. Due to the robust color classification, the soccer objects detection is also robust and adaptive to slightly varying lighting condition. The only desired improvement is on the latency of the pixel aggregation algorithm. The latency can potentially be decreased by adopting some local pixel aggregation approaches.

VI. References

[1] J. Bruce, T. Balch, and M. Veloso, "Fast and inexpensive color image segmentation for interactive robots," in *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '00)*, vol. 3, pp. 2061–2066, October 2000.

[2] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed., John Wiley and Sons,: 2001, pp.114-124.

[3] C. He, R. Xiong, and L. K. Dai, "Fast Segmentation and Identification in Vision System for Soccer Robots", in *Proceedings of the IEEE 4th World Congress on Intelligent Control and Automation*, vol.1, pp. 532-536, June 2002.

[4] C. Y. Lai, *System Integration and Development for Soccer Robots on Virtual Reality and Visual Servoing*, Master Thesis, Department of Mechanical Engineering, National Taiwan University, Taiwan: 2002, pp.25-29.

[5] W. K. Pratt, *Digital Image Processing*, third edition, John Wiley and Sons, 2001, pp. 63-87.

[6] G. Sen Gupta, C. H. Messom, S. Demidenko, and Lim Yuen Siong, "Identification and Prediction of a Moving Object Using Real-Time Global Vision Sensing", in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, pp 1402-1406, May 2003.

[7] G. F. Tong, W. D. Xiao, and X. H. Xu, "Design of RoboCup robot soccer global vision system", in *Proceedings of the IEEE 3rd World Congress on Intelligent Control and Automation*, vol.1, pp. 212-214, June 2000.

[8] G. Wyeth, B. Browning, and A. D. Tews, "UQ RoboRoos: Preliminary Design of a Robot Soccer Team", *Lecture Notes in AI: RoboCup '98 vol. 1604*, Springer, 1999, pp. 428- 434.

[9] G. Wyeth, and B. Brown, *Mechatronics and Machine Vision:A Research Studies Press*, John Billingsley, 2000, pp. 41-48.

[10] http://www.fira.net