Kuwait J. Sci. **45** (**1**) pp 39-53, 2018

Design of an efficient multi-objective recognition approach for 8-ball billiards vision system

Jiaying Gao[1,*], Qiuyang He[2], Hong Gao[3], Zhixin Zhan[4] , Zhe Wu[5]

[1]*School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China - gaojiaying_1988@163.com*

[2]*School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China - heqiuyang926@163.com*

[3]*China Zhongyuan Engineering Corporation, Beijing 100191, China - gaohong@czec.com.cn*

[4]*School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China - zzxupc@163.com*

[5]*School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China - wuzhe@dgut.edu.cn*

**Corresponding author: gaojiaying_1988@163.com*

Abstract

In this paper, some key technologies based on colour image processing for 8-ball billiards robot vision system are discussed and an efficient approach for multi-objective recognition is proposed. This approach is divided into two parts, i.e. multi-objective detection and ball pattern recognition. In image pre-processing, the normalized RGB colour space and histogram statistics are adopted for segmentation of background (table cover) and foregrounds. In order to accurately locate and isolate the single ball in a local region, the improved Hough Transform (HT) algorithm and the Least Squares (LS) method are adopted in combination. The improved HT algorithm is used for the purpose of eliminating the noise concentrated at edge points, and the LS method is used for fitting the circle center accurately with the least mean square error. Based on single ball detection in a local region, the multi-ball detection approach has been worked out to locate the position of each ball on the table. In the experiment, the proposed approach has been proved to complete the detection with an accuracy of 99.4% in 0.65s in average, and the performance is better than the traditional Circular Hough Transform (CHT) algorithm and the K-means cluster method. In addition, the Convolution Neural Network (CNN) method is adopted for pattern recognition of each target ball being segmented, i.e. identification of a solid ball or a striped ball. In order to improve the quality of CNN training: the colour segmentation and morphologic operation are applied for the segmented ball image pre-processing; the training set images are rotated for augmentation; pre-training stage is introduced in for optimizing the initial weight matrices. The calibrated image blocks are imported to the network for training. In the verification test, the trained CNN model shows a recognition rate of over 98.5%, and outperforms the other three classic methods. The introduction of CNN method has been proved to be correct and effective, and is an innovative and significant step for the design process of the 8-ball billiards robot vision system.

Keywords: Billiards recognition; convolution neural network (CNN); improved Hough Transform (HT) algorithm; least squares (LS) method; normalized RGB colour space.

1. Introduction

The motion law of a ball in the billiards game is a significant subject of research attracting the interest of many professionals engaged in mechanics, mathematics, and computation. With the development of artificial intelligence, quite a number of billiards robots have been developed to imitate the skill of human beings (Ho *et al*., 2007; Cheng *et al*., 2004), in which computer vision technology plays an important role. In the sport game, visual support systems provide abundant useful information for referees and players to pursue the fairness of the game and avoid wrong judgements (Shih & Chu, 2007). In terms of the billiards robot, the function of its vision system is to obtain essential data, which serves as the basic link of upper intelligence algorithm implementation and game situation analysis (Höferlin *et al*., 2010; Gao *et al*., 2015; Gao *et al*., 2017). In general, the pocket billiards include Snooker, 9-ball and 8-ball: Snooker is played with one cue ball, 15 red balls and 6 balls with other colours (yellow, green, brown, blue, pink and black). The player must hit the cue ball to shoot one red into the pocket before potting another colourful ball, and the player with a relative higher total score will win the game. 9-ball game is played with ten balls: The cue ball, and other nine object balls. The winner needs to pot the ball No.9 into pocket according to the game rule. 8-ball is played with 16 balls: a cue ball, and 15 object balls. Each object ball on table is distinctly painted with white, colour fringe and number (from No.1 to No.15). Before the game starts, all the 15 object balls are arrayed in the fixed initial triangular, and the black ball (No.8) is placed in the middle of the triangular. After hitting the cue ball with a break shot, object balls will be scattered. The player starting the game will then assign either the group of solid balls or striped balls; automatically, the other player will use those balls not being selected by his opponent. According to the game rules, after potting all the called balls, the player will finally shoot the black ball (No.8) into the called pocket to win the game. This paper is focused on some key

technologies adopted in the research of 8-ball billiards robot vision system, and the multi-objective recognition approach is proposed accordingly for the purpose of locating the accurate position and recognizing the pattern of each ball on table.

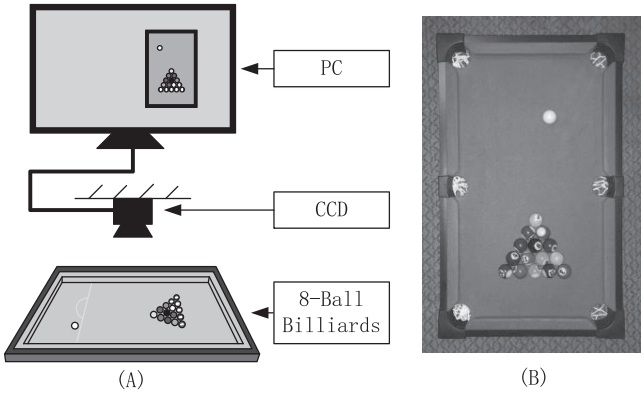


Fig.1. The 8-ball billiards robot vision system

The 8-ball billiards robot vision system is composed of a billiards table, balls, a CCD camera, and a PC as shown in Fig.1 (A). The CCD camera is fixed directly over the table through proper adjustment, and a 2D top-down view will be obtained for convenience of analysis as shown in Fig.1 (B). The solid and striped balls have different colour fringes and numbers on surface. Comparing with the single pattern of a solid ball, the fringes on a striped ball is relatively complex. The experiments for image processing and analysis were implemented over computer Lenovo T420, core i5, 2.30GHz. The relevant programs and algorithms were built by C++ language and openCV functions.

The ultimate object of 8-ball billiards robot vision system is to detect the location and recognize the pattern of each ball on table. In this paper, we combine the improved Hough Transform (HT) algorithm with the Least Squares (LS) method to detect single ball centre point position in a local region, and the multi-ball detection process is regarded as a set of finite single ball detection iterations. In the experiment, the proposed method has been proved to complete the detection with an accuracy of 99.4% in 0.65s in average. The performance is better than the traditional Circular Hough Transform (CHT) algorithm and the K-means clustering method. The combination algorithm (the improved HT algorithm and the LS method) effectively takes advantages of two methods and overcomes their drawbacks, it provides a new technical solution for the billiards ball detection process. In addition, we first introduce the Convolution Neural Network (CNN) method to the 8-ball billiards robot vision system. The training quality has been effectively improved after pre-processing, data augmentation and pre-training. The trained CNN model works with good robustness and is able to accurately identify each ball pattern. In terms of recognition rate, the CNN method shows a 2% of betterment than the random Forest method and outperforms the other two classic methods. The introduction of the CNN method is an innovative and significant step in the process of the robot recognition design, which is more consistent with the human visual perception.

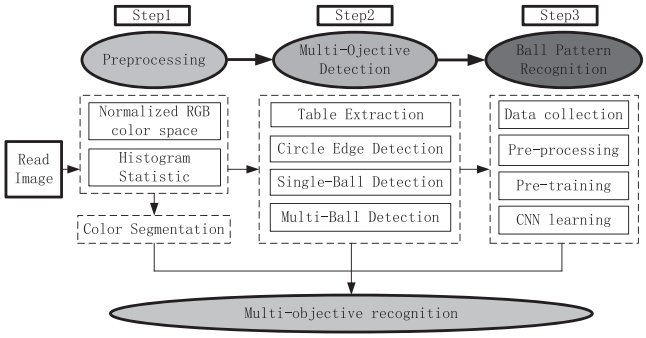


Fig.2. The block diagram of the multi-objective recognition methodology for 8-ball billiards robot vision system.

The block diagram of the methodology proposed in this paper is shown in Fig.2, and the rest of paper is organised into the corresponding 5 sections. In Section 2, the related works and methods about billiards robot vision system has been summarized. In Section 3, the normalized RGB colour space for image pre-processing is described. In Section 4, the multi-objective detection processes 4-step including table extraction, edge detection, the single ball detection in a local region and the multi-ball detection in the whole image. Based on the multi-objective detection, the CNN method is used for ball pattern recognition in Section 5. Data collection, image pre-processing (including greying and morphological operating), pre-training and CNN learning are the four parts for the ball pattern recognition model. In a word, the multi-objective detection and ball pattern recognition are the core modules in the 8 ball billiards robot vision system. The key technical details are discussed in the following sections.

2. Related research

In the relevant research, some image processing algorithms have been proposed to solve the multi-objective detection problem in Snooker game. Ho *et al*. (2007) summarises the technical details of a Snooker robot, in which the AV4 vision system is able to locate twenty three colour balls on the table by using its own built-in image segmentation algorithm, and the average gray scale value of the ball pixels is assumed to represent the colour. The main drawback of his image processing technique is that it is very sensitive to the surrounding light variation. Shen & Wu (2010) developed an algorithm to extract all the billiard objects from Snooker videos. Y. Ling *et al*. (2012) proposed an effective multi-ball detection method to locate

all the balls automatically in Snooker game. Since the green ball has a similar colour to the table background, he divides the method into two parts, i.e. non-green balls detection by marked watershed algorithm and the green ball detection by picking up the reflection point. Legg *et al*. (2011) developed a system for 3D reconstruction based on Snooker video data.

The object of the billiards robot vision system is to detect the location and recognize the pattern of each ball on the table. The solution process of this task can be divided into the following three stages: (1) Image pre-processing; (2) Multi-objective detection; (3) Ball pattern recognition. The details of the design process depend on the specific requirements of the game type.

Background detection is an important step in image pre-processing stage of billiards robot vision system. Snooker is played on the table covered with green baize, while the 9-ball usually adopts the blue table cover. As for 8-ball, either green or blue colour can be used for background table cover (Americans usually use blue, while Chinese often choose green), and the green colour is adopted in this paper. In the relevant studies, template subtraction (Ling *et al*., 2012), histogram statistic (Takahashi *et al*., 2008) and colour segmentation have been applied for background detection. These techniques are combined with different colour spaces (HSV, YUV) and work well under a constant light condition, but are sensitive to the ambient illumination variation (Ho *et al*., 2007). This paper takes advantage of the normalized RGB for background detection because this colour space is simple, fast and can effectively filter out the influence of light (Subban & Mishra, 2012; Dargham & Chekima, 2006).

The purpose of multi-objective detection is to accurately locate the position of each ball on a table. The issue of multi-ball detection is identical to multi-circle shape detection. Actually, the problem of detecting circular features plays an important role in several engineering applications and has been commonly solved by introducing HT algorithm.

The HT algorithm is a strong tool widely used all over the world in areas of computer vision and digital image processing. The purpose of HT algorithm is to detect the corresponding shape by transforming the points from the Cartesian space to parameter space and then performing statistical analysis. The classical HT algorithm is used for identification of lines, and its application has been extended to detection of arbitrary shapes, such as circles or ellipses (Cuevas *et al*., 2012; Chen *et al*., 2012).

The HT algorithm is widely used in circle detection due to its ideal robustness to partial deformation and tolerance to noise, especially when the target being detected has flaws, or when the target is partially occluded. However, there are two disadvantages in the use of HT algorithm for shape detection. Firstly, a considerable amount of computation and a large memory capacity will be consumed during the voting procedure. Secondly, in case a high accuracy for fitting is preferred, the HT algorithm may not meet the requirement, as the detection accuracy is subject to the selection of the threshold. On the other hand, the LS method can obtain the absolute accuracy curve of the given data set with the least mean square error, but the HT algorithm cannot. However, the LS method is very easily affected by the so-called outlier points, that is to say, when there are errors in the data set, the fitted result of the LS method may be absolutely wrong. Considering the advantages and disadvantages of the two methods, a new method is proposed in this paper. An improved HT algorithm combined with the LS method for accurately detecting and fitting a single circle in a local region is proposed in this paper. On the basis of single ball detection, an effective multi-ball detection approach for locating positions of all the balls in the whole image is presented. The multi-ball detection approach has been proved to be effective and demonstrated as a better performance than other relevant multi-circle cluster methods.

In addition to development of the multi-ball segmentation and detection method, it is worthy of paying attention to the point that the billiards robot must be able to recognize the pattern of each ball on table for the game situation analysis. In a Snooker game, such identification is simple as there is nothing but a single colour on the ball surface. Histogram statistics and average gray scale calculation of ball region pixels can be straightforwardly determined by the ball colour and pattern (Wu *et al*., 2010). However, in the 8-ball billiards game, this ball pattern recognition problem becomes complicated. Considering the fact that the colour of marked numbers and their fringes painted on the ball surface will be identical to other balls, e.g. the white plus yellow colour on ball No. 1 and ball No. 9, it is difficult to effectively extract the implicit features on ball image. In this paper, the CNN method is adopted to solve this ball pattern recognition problem.

The CNN is a type of feed-forward artificial neural network with multiple layers based on deep learning, which has been widely used in the areas of image and video processing (Lau *et al*., 2015; Verma & Vig, 2014). By introducing the CNN in image recognition and classification, instead of developing complicated algorithms to extract and learn image characteristics manually, what needs to be done is to simply train the system by adding a little pre-processing. Compared with the traditional Neural Network (NN) and other classification methods (Kumar *et al*., 2015; Nallamuth & Palanichamy, 2015; Gao *et al*., 2016), using the CNN for

image recognition has two advantages. Firstly, the architecture of the CNN model is able to learn complex invariance in 2D image (such as scale, rotation and translation invariance). Secondly, the use of local receptive field and tied weights can significantly reduce the number of parameters to be obtained from learning. In this paper, such a CNN model is proposed for ball pattern recognition in 8-ball billiards robot vision system.

3. Image pre-processing

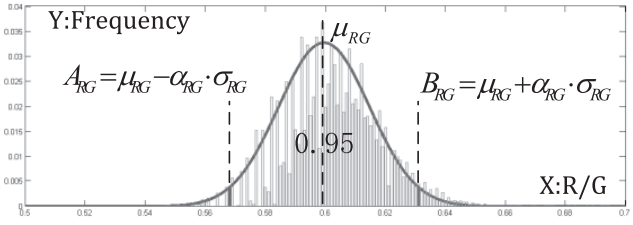
The segmentation of foreground and background is an important pre-processing step in the multi-objective detection. In this paper, the table covered with green baize is taken into consideration as shown in Fig.1 (B). Green baize colour has been used as an important feature for table detection and ball detection. The colour space used for background identification can be roughly divided into two categories: hardware equipment oriented space (such as RGB) and perception oriented space (such as HSV). HSV space is more consistent with human colour feeling, and is generally used for image post-processing and reconstruction. In the pre-processing stage, the original bitmap data read by PC are generally RGB pixels. The transition from RGB space to HSV space is nonlinear, and the efficiency is relative low so as to need introducing other optimization algorithm. Hence the RGB space is considered to be adopted because the space is simple, fast and meeting the subsequent processing requirements. In order to have a better adaption to the ambient light variation, the normalized RGB colour space is used for the segmentation of background pixels and foreground pixels.

Assuming a given pixel ***A*** on the table coloured with R, G, B, where R, G, B are the intensity of red, green and blue. In the normalized RGB colour space, this colour can be converted into (r, g, b), where r, g, b represents the proportion of red, green and blue in the original colour. This kind of conversion can effectively remove light intensity variations. The value of r, g, b can be calculated by the normalization procedure in Equation (1):

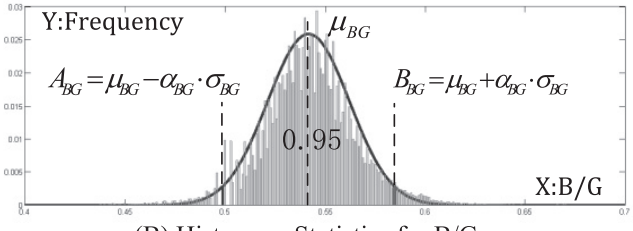
$$\begin{aligned} r &= R/(R+G+B); \\ g &= G/(R+G+B); \\ b &= B/(R+G+B); \\ r+g+b &= 1 \end{aligned}, \tag{1}$$

In terms of an image, the illumination intensity will cast an influence on the value of R, G, B, but the percentage of each colour component will not change dramatically. The sum of r, g, b is equal to 1, and the g dimension can be truncated without causing any information loss. For any point ***P*** of the green background colour, the ratio of the colour components of the R/G and B/G should be in a stable range. Similar to the processing in gray image, the histogram statistics for the R/G and B/G value of each pixel in the whole image are shown in Fig.3 (A) and Fig.3 (B). The height of the gray bar represents the frequency of the corresponding colour ratio, and the relation between the frequency and the ratio is able to approximately satisfy the Gauss distribution. The red curve in Fig.3 (A) and the blue curve in Fig.3 (B) respectively stand for the effects of Gauss fitting. Then the mean and the standard deviation of the value R/G and B/G are estimated and denoted as: μ_{RG} , μ_{BG} , σ_{RG} , σ_{BG} The coefficients α_{RG} and α_{BG} represent the accuracy of the segmentation between background and foreground, in which the lower values means a more strict judgment of the background point, refer to Equation (2). When the confidence level is set at 95%, the confidence interval for R/G and B/G can be calculated and expressed as: (A_{RG}, B_{RG}) and (A_{BG}, B_{BG}). Then the values of α_{RG} and α_{BG} are obtained: $\alpha_{RG} = (\mu_{RG} - A_{RG})/(\sigma_{RG})$; $\alpha_{BG} = (\mu_{BG} - A_{BG})/(\sigma_{BG})$. If the pixel P coloured with R, G, B satisfy the two inequalities below in Equation (2), then the ***P*** is considered as a background colour pixel, denoted as ***B*** type. Conversely, the ***P*** is regarded as a foreground point, denoted as ***F*** type.

$$\begin{aligned} \mu_{RG} - \alpha_{RG}\cdot\sigma_{RG} < \frac{R}{G} < \mu_{RG} + \alpha_{RG}\cdot\sigma_{RG} \\ \mu_{BG} - \alpha_{BG}\cdot\sigma_{BG} < \frac{B}{G} < \mu_{BG} + \alpha_{BG}\cdot\sigma_{BG} \end{aligned}, \tag{2}$$



(A) Histogram Statistics for R/G



(B) Histogram Statistics for B/G

Fig.3. Histogram Statistics based on normalized RGB colour space: (A) Histogram statistics for R/G. (B) Histogram statistics for B/G.

4. Multi-objective detection

The aim of multi-objective detection is to accurately locate the position of all the targets on table. This image segmentation processing can be divided into following three steps:

(1) Table detection.

(2) Single ball detection in a local region.

(3) Multi-ball detection in the whole image.

4.1 Table detection

Table detection is mainly for filtering out the interference factors outside the table region. In the top-down view, the movement of each ball is confined in the rectangular area covered with green baize, as shown in Fig.4 (A). The key of the rectangle detection is to find out four straight edge lines by introducing the HT algorithm. The edge points will be extracted by the threshold segmentation, as shown in Fig.4 (B).

The four edge lines of billiards table and their intersection points will be correctly detected by the HT algorithm as shown in Fig.4 (C). After table detection, the target area for image processing will be segmented effectively, as shown in Fig.4 (D).

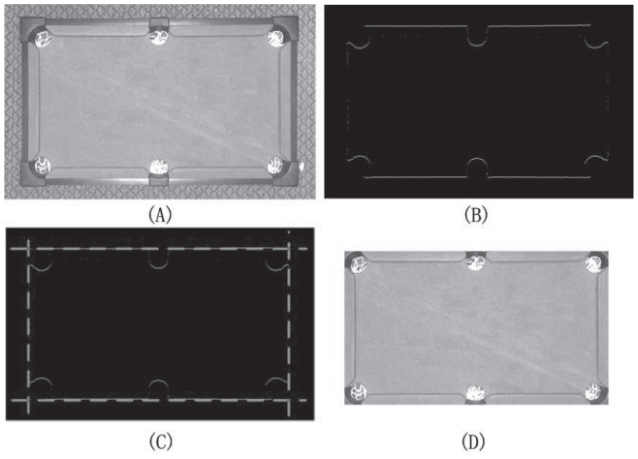


Fig.4. Billiards table extraction process by the HT algorithm

4.2 Single ball detection in a local region

The flowchart of the single ball detection in a local region is shown in Fig.5. Edge extraction, improved HT algorithm and LS method are the three steps. There are colour fringes and the marked numbers painted on the surface just as the yellow ball in Fig.5. These factors do not affect the process of ball detection and will be discussed in the Section 5. In addition, the ball diameter D in the image is an important parameter that must be calculated prior to the detection process. Assuming the distance between the table and the CCD camera is denoted as l_{CCD}. When the value of l_{CCD} is given, the D can be always expressed as: $D = K / l_{CCD}$, in which K is a constant parameter which can be measured in experiments and its unit is (pixel·m).

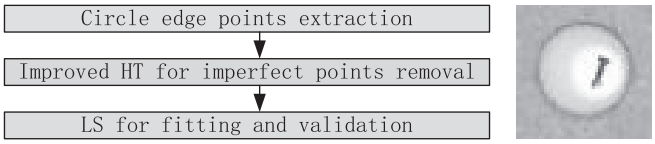


Fig.5. The flowchart of single ball detection in a local region

4.2.1 Circle edge points extraction

When scanning the whole image from top left corner to bottom right corner, we may see pixel P_i and its 4-neighbor pixels N_1, N_2, N_3, N_4 as shown in Fig.6 (A). If P_i belongs to $\boldsymbol{F}$ type and its two arbitrary adjacent neighbour pixels are of $\boldsymbol{B}$ type, then the pixel P_i can be defined as an edge point, denoted as $P_i \in \boldsymbol{E}$. The circle edge point set $\boldsymbol{S}$ will be obtained after scanning, in which P_1, $P_2 \ldots P_{num}$ are the *num* edge points being detected. If P_i is detected as an edge point, then there may be at least one complete ball in a local region centered with P_i. Assuming the coordinate position of P_i is (x_i, y_i), then the range of horizontal and vertical coordinates of the local region can be expressed as $(x_i - D, y_i - D)$ and $(x_i + D, y_i + D)$ respectively.

Based on the original image shown in Fig.6 (B), Fig.6 (C) displays the effect after binary processing in which the pixels belonging to set $\boldsymbol{S}$ are denoted as 1. In a local region shown in Fig.6 (D), although most of circle edge points have been effectively extracted, there are still quite some noises identified as edge points mistakenly and mixed with the correct ones. In order to accurately locate the center point of circle, it is necessary to eliminate the noises from the set **S**. The improved HT algorithm for imperfect points eliminating is presented.

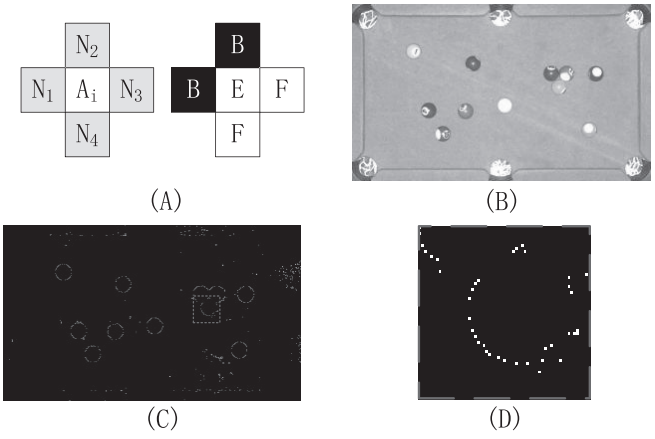


Fig.6. Circle edge points detection

4.2.2 The improved HT algorithm for noise removal

Although there may be multiple balls, only the circle with the most complete edge points and features in a local region will be eventually found in the single ball detection stage. The purpose of introducing the improved HT algorithm is to find out and

remove noises from the assumed circle in the area. In terms of num1 edge points i.e. $P_1 \ldots P_i \ldots P_{num1}$, most of them are correct, while the error ones among them must be eliminated. Since the ball diameter is calculated and denoted as D in advance, for arbitrary three different points P_i, P_j, P_k in the set $\boldsymbol{S}$, assuming that $D'(P_i, P_j, P_k)$ is the circumcircle diameter between the three points, and β is a suitable threshold value. If the absolute value of difference between $D'(P_i, P_j, P_k)$ and D exceeds the threshold β, then there must be at least one point among P_i, P_j, P_k is wrong.

However, the error point among the above three ones cannot be accurately identified, and the three points P_i, P_j, P_k should be under suspicion of being the one in error. Based on this fact, a new property named ***error*** is introduced to represent the suspiciousness of each point. The principle of the improved HT algorithm is that the point with the highest error will be eliminated from the set $\boldsymbol{S}$ by voting. Finally only the correct edge points are left in the $\boldsymbol{S}$ after a finite iteration. The ***error*** of each point is equal to zero in the initial iteration, which can be expressed in Equation (3):

$$\begin{aligned} &P_1.error = 0; \\ &P_2.error = 0; \\ &\ldots \\ &P_{num_t}.error = 0. \end{aligned} \quad , \tag{3}$$

$$\begin{aligned} &if: |D'(P_i,P_j,P_k) - D| > \beta, then: \\ &P_i.error = P_i.error + 1; \\ &P_j.error = P_j.error + 1; \\ &P_k.error = P_k.error + 1; \end{aligned} \quad , \tag{4}$$

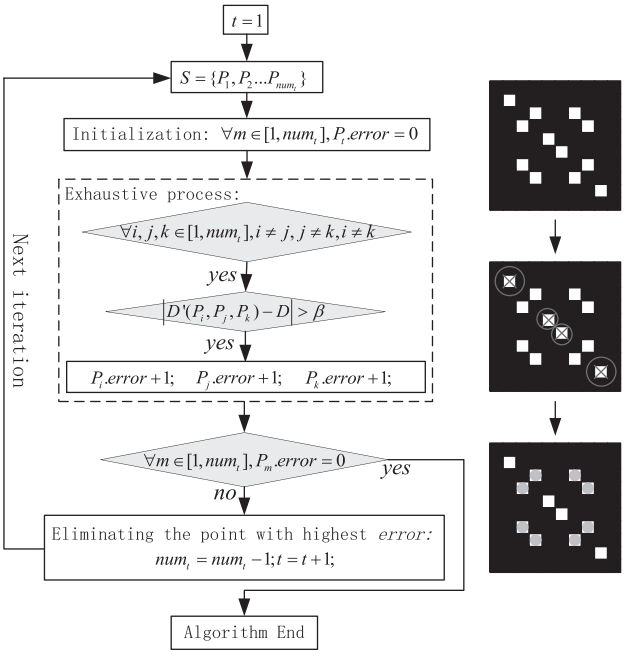


Fig.7. The flowchart of the improved HT algorithm for noise removal

The flowchart of the algorithm is shown in Fig.7. Assuming that the set $\boldsymbol{S}$ containing num_t points is confronted the t-th iteration, all possible 3-different-point combination can be found out by the exhaustive method. For each combination composed of P_i, P_j, P_k, if the absolute difference between $D'(P_i, P_j, P_k)$ and D exceeds the threshold β, then the ***error*** of 3 points are added by 1 respectively as shown in Equation (4). The point with the highest ***error*** will be voted and eliminated from set $\boldsymbol{S}$ in the exhaustive process. Then the left (num_t *-1*) points in the updated set $\boldsymbol{S}$ will then be confronted by the next (*t+1*)-th iteration. The algorithm process will call an end when the ***error*** of every point is equal to zero after several iterations and eliminations. The points survived to the last are considered as the correct edges points.

4.2.3 The LS fitting and verification

After noise removal, the LS method is adopted for fitting a single circle in a local region. Assuming there are num_E points in set $\boldsymbol{S}$ after eliminating: $P_1, P_2 \ldots P_{numE}$, and the circle can be expressed as Equation (5), where (A, B) and R stands for the center position and the fitting radius. Introducing intermediate variables a, b, c, where $a = -2A$, $b = -2B$, $c = A^2 + B^2 - R^2$. Then the circle will be expressed as Equation (6):

$$(x - A)^2 + (y - B)^2 = R^2, \tag{5}$$

$$x^2 + y^2 + ax + by + c = 0, \tag{6}$$

For an arbitrary sample point P_i (x_i,y_i) on circle, d_i is the distance between P_i and center (A, B), q is the difference between d_i^2 and R^2. Assuming $Q(a, b, c) = q^2 = x_i^2 + y_i^2 + a\,x_i + b\,y_i + c$, then the aim of LS method is to find an optimal curve to obtain the minimum value for $Q(a, b, c)$. The lagrange multiplier rule is adopted for calculating and the derivation process is given in Equations (7), (8), (9):

$$\frac{\partial Q(a,b,c)}{\partial a} = \Sigma 2(x_i^2 + y_i^2 + ax_i + by_i + c)x_i = 0, \tag{7}$$

$$\frac{\partial Q(a,b,c)}{\partial b} = \Sigma 2(x_i^2 + y_i^2 + ax_i + by_i + c)y_i = 0, \tag{8}$$

$$\frac{\partial Q(a,b,c)}{\partial c} = \Sigma 2(x_i^2 + y_i^2 + ax_i + by_i + c) = 0, \tag{9}$$

Introducing variables C, F, E, G, H, where:

$$C = num_E \Sigma x_i^2 - \Sigma x_i \Sigma y_i, \tag{10}$$

$$F = num_E \Sigma x_i y_i - \Sigma x_i \Sigma y_i, \tag{11}$$

$$E = num_E \Sigma x_i^3 + \Sigma x_i \Sigma y_i^2 - \Sigma(x_i^2 + y_i^2)\Sigma x_i, \tag{12}$$

$$G = num_E \Sigma y_i^2 - \Sigma x_i \Sigma y_i, \tag{13}$$

$$H = num_E \Sigma x_i^2 y_i + num_E \Sigma y_i^3 - \Sigma(x_i^2 + y_i^2)\Sigma y_i, \tag{14}$$

The parameters a, b, A, B can be calculated by Equations (10)-(14), as shown in Equations (15)-(19):

$$A = -\frac{a}{2} = -\frac{(HF - EG)}{2 \cdot (CG - F^2)}, \tag{15}$$

$$B = -\frac{b}{2} = -\frac{(HC - EF)}{2 \cdot (F^2 - CG)}, \tag{16}$$

$$c = -((\Sigma x_i^2 + \Sigma y_i^2) + a \cdot \Sigma x_i + b \cdot \Sigma y_i) / (num_E), \tag{17}$$

$$D_{LS} = \sqrt{a^2 + b^2 - 4 \cdot c}, \tag{18}$$

$$validation: | D_{LS} - D | < \delta, \tag{19}$$

(A, B) represents the fitted location of circle center, D_{LS} is the fitting diameter. As the ball diameter *D* has been calculated in advance, if the absolute difference between D_{LS} and *D* is less than the threshold δ as shown in Equation (19), then the single ball is considered to be detected correctly, otherwise, the detection result will be discarded.

4.3 Multi-ball detection in whole image

Multi-ball detection process can be regarded as a set of finite single circle detection iterations. The variables temp is used to record the number of balls being discovered, $\boldsymbol{C_1, C_2 \ldots C_{temp}}$ are the centers being detected, and dis (P_i, P_j) represents the distance between the two pixels P_i and P_j, then the flowchart of multi-ball detection is shown in Fig.8.

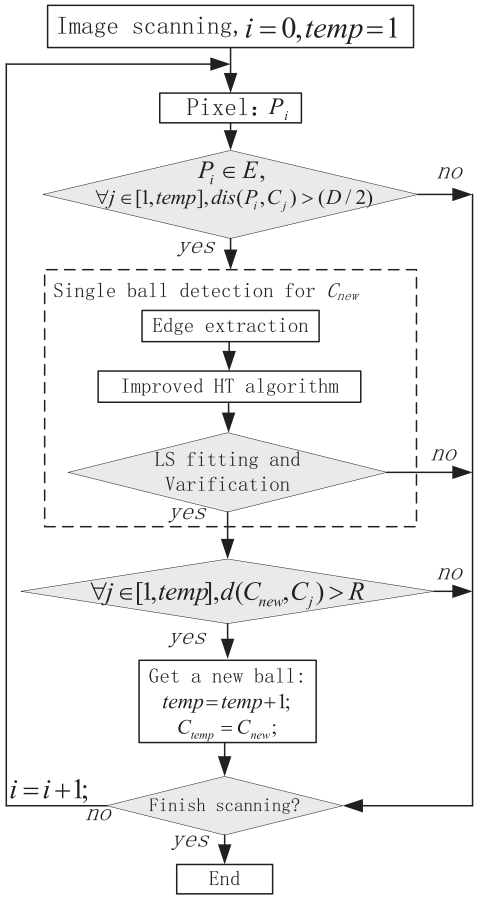


Fig.8. The flowchart of Multi-ball detection algorithm

Each pixel in the whole image is scanned from the up left corner to the bottom right corner. If one pixel P_i is detected belongs to ***E*** type, and the distances between the P_i and the all circle centers have been found are larger than radius (*D*/2), P_i might be the undiscovered edge point of a new ball. Then every pixel in the local square region sized *D*×*D* of P_i will be scanned, and the edge points set ***S*** in this local region can be found. The improved HT algorithm will eliminate the imperfect points from set ***S*** and the LS method will be used for fitting and verification. If the distances between the new fitted circle center and other discovered circle centers are all larger than *D*, then the new ball is detected. After single iteration of whole image scanning, the location of all balls on billiards table will be detected.

4.4 Experiments and performance comparison

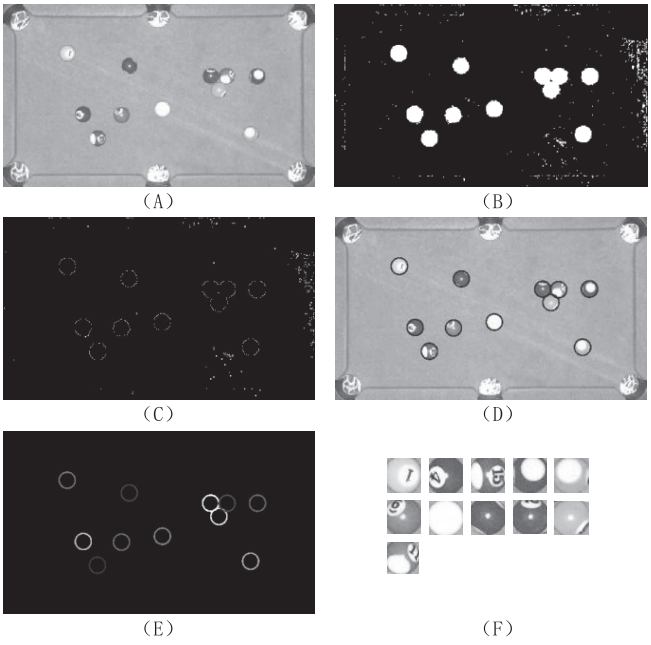


Fig.9. Multi-ball detection effect

Fig.9 (A) is the original image extracted from the table. Fig.9 (B) shows the binary image of background and foreground segmentation based on the normalized RGB colour space. Fig.9 (C) displays the circle edge point detected in the whole image. Fig.9 (D) and Fig.9 (E) show the final multi-ball detection effect. Fig.9 (F) shows the images of balls being segmented. The experimental results indicate that each ball on the table has been correctly fitted and extracted, the normalized RGB colour space is appropriate for the 8-ball billiards robot vision system and the multi-ball detection method is effective. Fig.10 gives the effect of single ball detection in a local region by using the improved HT algorithm and the LS method. The edge points not belonging to the ball have been neglected, and the detected single ball is fitted and painted in pink red circle

as shown in Fig.10. The improved HT algorithm is able to effectively eliminate the imperfect points from the edge points set **S**, and the reliability and robustness has been proved. The LS method ensures the sufficient accuracy of the fitted circle.

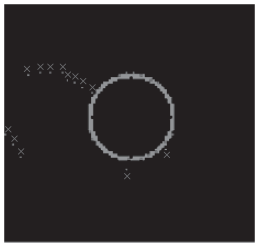


Fig.10. The effect of single ball detection in a local region after being processed by the improved HT algorithm and LS fitting

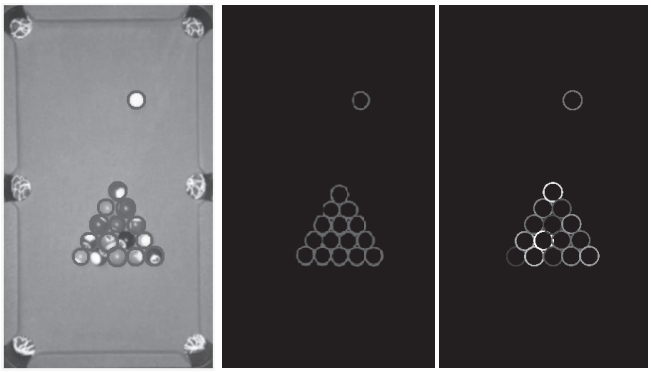


Fig.11. The multi-ball detection effect through the three methods: (A) the CHT method; (B) the K-means method; (C) the approach presented in this paper.

Fig.11 shows the comparison of multi-ball detection effect for game beginning situation between the approach proposed in this paper and other classical methods for multi-circle cluster. Fig.11 (A) shows the detection effect by the CHT method. Although most circles are detected correctly, there are still some circles being missed or inaccurately detected. This is mainly because the detection accuracy of CHT method is depended on the selection of the threshold. Fig.11 (B) is the effect by the K-means cluster method. Using K-means method can obtain a better segmentation result than CHT, but its efficiency is relatively low. K-means method needs well-done pre-processing step for noise removal and an accuracy estimation of the cluster number before the method is applied (Shih *et al*., 2007). Fig.11 (C) is the multi-circle detection effect by the approach presented in this paper. Table 1 shows the corresponding average execution time and detection rate for the 3 methods performed in the 8-ball billiards robot vision system. The comparison results indicate the effectiveness, robustness and reliability of the multi-ball detection approach proposed in this paper.

Table 1. The perform statistic comparison between the 3 multi-circle cluster methods

Method	Execution time(s)	Success rate(%)
The approach in this paper	0.65	99.4%
CHT	0.6	89.4%
K-means	1.4	95.2%

5. Ball pattern recognition based on CNN

In the above discussion, the method proposed in the paper has been proved to be effective to locate the position of all the balls on the table. However, simply introducing multi-ball detection cannot solve all the problems in the 8-ball billiards robot vision system. The system should have the ability to recognize the target ball pattern automatically so as to give correct judgment and avoid foul shooting. The recognition of solid ball and striped ball lies in the characteristics of the fringes distribution (as shown in Fig.12), rather than the colour, that accounts for difference between 8-ball billiards game and Snooker game. In the visual perception of the human eye, the solid ball is of the single colour while the striped ball seems to be of the mixed colour. From a specific point of view, both solid and striped balls may be mistakenly recognized by human eye. It is difficult to develop an accurate mathematical model to sort out patterns of the solid and striped balls. This paper takes advantages of the CNN method to extract the implicit feature. The block diagram of the ball pattern recognition procedure based on CNN is shown in Fig.13. Data collection, data pre-processing, CNN training and verification test are the four stages for the solution of this problem.

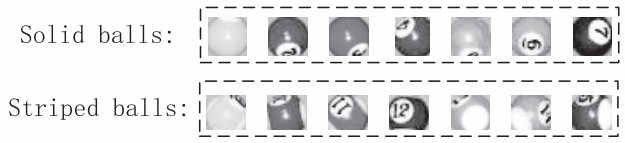


Fig.12. The comparison between solid and striped ball patterns.

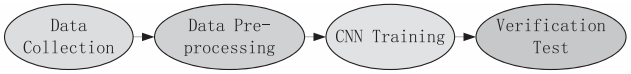


Fig.13. The block diagram of ball pattern recognition.

5.1 Data collection

Using the multi-ball detection method proposed in this paper, the data of ball images with pattern label (solid or striped balls) in different game situations will be segmented and recorded as shown in Fig.14. The original images of segmented balls are all adjusted to 24X24 pixels for calibration. There are 2000 image recordings in the original database. The training set includes 800 solid balls and 800 striped balls, and the remaining 400 images are for testing. Since the difference between solid

balls and striped balls lies in the characteristics of the fringes distribution, and the colour image contains a lot of irrelevant information that will prolong the training time and reduce the performance of CNN model, the colour images will be greyed by colour segmentation and morphological pre-processing will be done prior to being inserted into the CNN model.

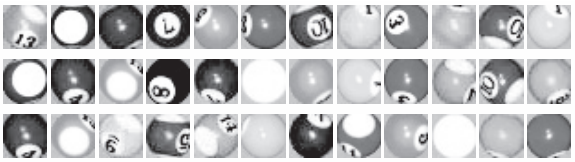


Fig.14. Original Data Collection for CNN training

5.2 Colour segmentation and morphologic processing

For an arbitrary target ball after being extracted as shown in Fig.16 (A), there are about three types of colours on the ball surface i.e. white area, colour fringe and black number. The distribution of the colour fringe and white area are the critical factors for ball pattern recognition. While the number on the ball surface is the noise to be filtered out. In order to remove the influence of the marked numbers, the black colour and the white area are merged into grey. Then an arbitrary pixel P(R, G, B) on the ball surface will be classified into grey or non-grey colour by the threshold segmentation based on the normalized RGB space, defined as below in Fig.15:

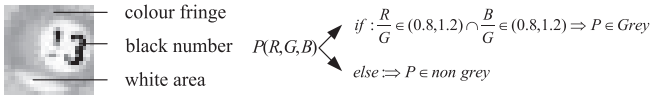


Fig.15. Classification of grey area and non-grey area by threshold segmentation based on normalized RGB colour space

After colour segmentation, the gray image of the ball is obtained. The value of 255 and 0 represent the grey pixels and the non-grey pixels respectively, and the pixels outside circle area is valued with 125, as shown in Fig.16 (B). There are still some irrelevant areas in the grey image after colour segmentation, which are mainly caused by the threshold segmentation error and the light reflection. Morphological pre-processing is used to remove the noise in the grey image. Fig.16 (C) gives the effect of the image processing after closing and opening operation.

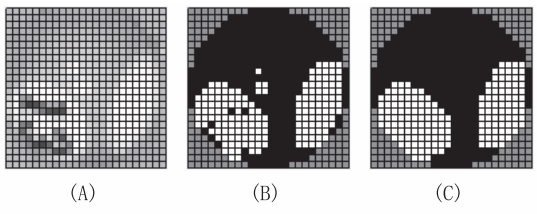


Fig.16. The effect of morphological pre-processing for grey image

5.3 Structure of the CNN model for ball pattern recognition

After being processed to the calibrated grey images, the CNN model is introduced for ball pattern recognition to the target balls in database. The CNN is an artificial neural network with multi-hidden layers, which provides good robustness for the recognizing variation of translation, scaling and rotation of images. The local pixels with similar feature and the same weight values in convolution layers are adopted for saving memory and improve efficiency. These advantages of the CNN make it very convenient for image recognition problem solving. The structure of the CNN model for ball pattern recognition in this paper is shown in Fig.17.

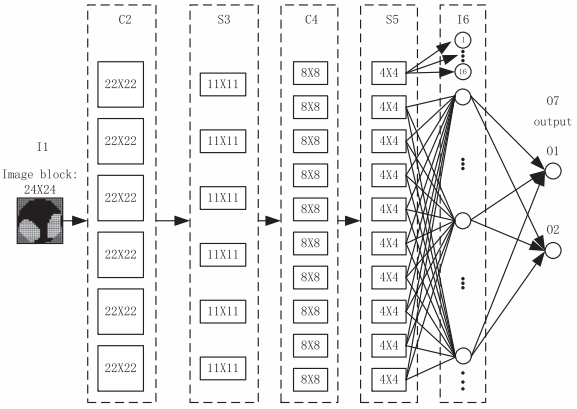


Fig.17. Structure of the CNN model for ball pattern recognition

The network architecture consists of seven layers, i.e. *I1, C2, S3, C4, S5, I6,* and *O7*. In the input layer *I1*, the data matrices of calibrated grey-scale images are inserted into the network. *C2* and *C4* are the two convolution layers, S3 and S5 are the two subsampling layers. After automatic feature map extraction and pooling processing in convolution and subsampling layers, the data are inserted into a single layer perceptron model. *I6*, *O7* respectively represents the fully connected input layer and the output layer of the single layer perceptron.

The dimension of original input image is 24X24. In the convolution layer *C2*, the image from *I1* is convoluted by 3X3 kernels, and the four feature maps sized 22X22 are generated and fed into the subsampling layer *S3*. In S3, the four feature maps are subsampled into four 11X11 feature maps by the pooling process. In the followed *C4* layer, the four 11X11 feature maps from *S3* are again patched by eight 4X4 convolution kernels, and the eight 8X8 feature maps are generated and inserted into *S5* layer. After being pooled in *S5*, eight 4X4 feature maps are generated. The pixels of these feature maps are expanded into 192 input neurons in the layer *I6*. In the single layer perceptron model, the neurons between *I6* and the output layer *O7* are fully connected. Finally, the values of the two output neurons (O1 and O2) in *O7* represent the pattern of the target ball.

Assuming the values of the two output neurons in O7 are O_1 and O_2 respectively, and $\boldsymbol{thr_1}$ and $\boldsymbol{thr_2}$ are introduced in as two thresholds, in which $\boldsymbol{thr_1}$=0.85 and $\boldsymbol{thr_2}$=0.15, then the relation between O_1, O_2 and the classification result can be defined as below:

If $O_1 > \boldsymbol{thr_1}$ and $O_2 < \boldsymbol{thr_2}$, then the ball pattern is considered as the solid ball.

If $O_1 < \boldsymbol{thr_2}$ and $O_2 > \boldsymbol{thr_1}$, then the ball pattern is considered as the striped ball.

Otherwise, the ball belongs to an unknown pattern.

5.4 Improvement on introducing pre-training stage

CNN is a powerful neural network model for deep learning. The training process of the network needs a lot of sample data acquisition. However, due to the actual experimental conditions, the mode of manual data collection is adopted. There are only 1600 images for training. We have proposed two points in the improvement:

(1) Training set augmentation

All the training set images were clockwise rotated 90 degrees, 180 degrees, 270 degrees and added to the training set as shown in the Fig.18. After this augmentation processing, the image number for the training set will be increased to 6400. While the test set is not modified and kept the same number (400).

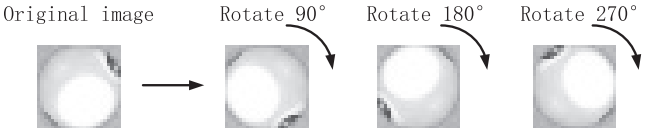
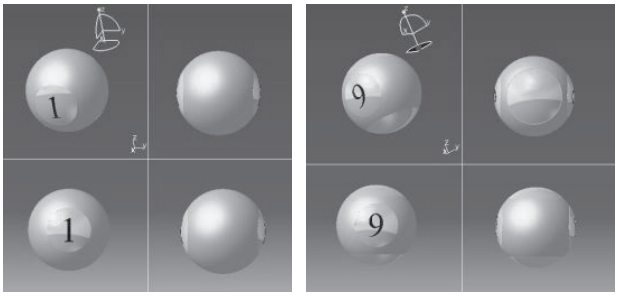


Fig.18. The augmentation of an image in training set

(2) Pre-training

In spite of the augmentation processing, the training set quantity is still insufficient for the deep CNN learning. In addition, when the CNN model is initialized, the weight matrices of the convolution kernel in the C2 and C4 layers are selected at random into the interval (0, 1), which will make the training effect not stable and difficult to achieve the local optimum. In order to improve the quality of CNN learning model, we have built the 3D model of the balls in the simulation software, and a considerable amount of data collected from the 3D models will be imported into the CNN model for pre-training. The pre-trained CNN model has the same structure as the actual model shown in Fig.17. After the pre-training is completed, the weight matrices of each convolution kernel will be saved and used for initializing the actual CNN training.

The software CATIA is adopted to build the 3D model for solid balls and striped balls. In the modelling process, the geometric features and the colour fringes are consistent with the actual balls. Fig.19 shows the orthographic views and the axonometric drawings of the 3D models for a solid ball and a striped ball.



(A) 3D modeling for solid ball (B) 3D modeling for striped ball

Fig.19. The orthographic views and the axonometric drawings of the 3D models by CATIA

On the basis of 3D modelling, a program for automatic collection of image data is developed. The program will be carried out to take pictures of the ball in various angles in the 3D work space. Each captured image will be reshaped into 24X24 pixels for calibration, and will be pre-processed in the same way as shown in Section 5.2. Finally, the image matrices will be recorded into the database for pre-training.

More than 40000 images are collected automatically for pre-training, in which the sample number ratio of solid to striped balls is close to 1:1. In addition, about 2000 images (excluding in the training set) are used in verification test. The pre-trained CNN model has the same 7-layer structure as the actual CNN model (i.e. I1, C2, S3, C4, S5, I6, O7), and the numbers of the feature maps and the kernels in each layer are also the same as the actual CNN model. The error descent curve of pre-training is shown in Fig.20, after 800 iterations, the training error was reduced and converged to about 0.002. In the verification test for pre-trained CNN model, the recognition rate for solid and striped balls all exceeded 99.8%. The pre-trained CNN model will be saved, and each weight matrix of the convolution kernels in the layer C2 and C4 will be extracted and prepared for the actual CNN training.

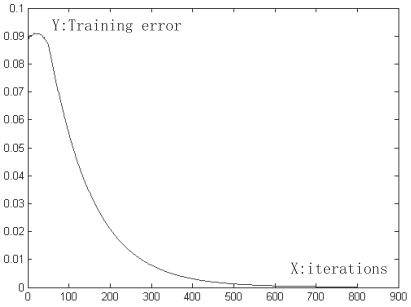


Fig.20. The error descent curve of the pre-training

5.5 Validation result and analysis

6400 images in the dataset were imported into CNN for establishment of the ball pattern recognition classifier model. After being trained by the CNN, 400 images were taken out for verification test. There is no intersection between the images in the training set and the test set. During the initialization of the network training, the weights obtained from pre-trained model are used as the initial weights of the actual CNN training, and the training error curve is given in the red path as shown in the Fig.21. After being pre-trained, the error of actual training was converged to 0.009 after about 5000 iterations. Compared with the original CNN training process (the blue curve in Fig.21), the CNN model after pre-training shows a faster convergence rate and a lower training error. In the verification test, the recognition rate of the solid and the striped balls have increased to 98.5% and 98.87% respectively as shown in Table 2. It has been proved that the CNN model is effective and reliable, and the introducing of pre-training process will improve the quality of the network.

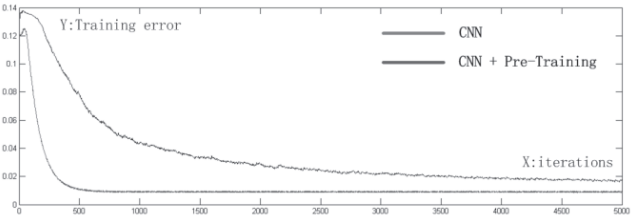


Fig.21. The error descent curve of the actual CNN training

Table 2. Statistic of recognition rate

	Solid ball	Striped ball
Scale of training set	800	800
Scale of test set	200	200
Recognition rate of the original CNN	97.00%	96.50%
Scale of training set after data augmentation	3200	3200
Scale of test set	200	200
Recognition rate after pre-training and data augmentation	98.50%	98.87%

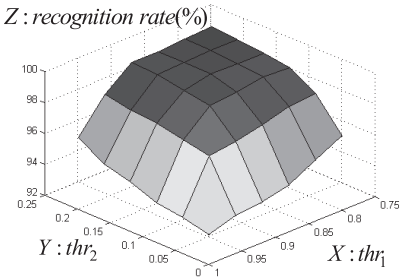


Fig.22. The robustness relation between the recognition rate and the classifier thresholds

In addition, Fig.22 shows the relation between the average recognition rate and the selection of the threshold th_1 and thr_2. The X axis and the Y axis respectively represent the value of thr_1 and thr_2. The Z axis stands for the corresponding average recognition rate. The statistic result indicates that a decrease of the recognition rate of the classifier can be found when $thr_1 > 0.95$ and $thr_2 < 0.05$ due to the harsh conditions. However, in the range of $0.75<thr_1<0.92$ and $0.09<thr_2<0.25$, the recognition rate always exceeds 98%. As long as the threshold value thr_1 and thr_2 are in a suitable range, the CNN model will show its good robustness.

On the bases of adopting the same database, this paper presents comparison of the CNN method with other methods used in image classification. Table 3 gives the average correct rate for ball identification through the traditional back propagation neural network (BPNN), the support vector machine (SVM) and the Rand Forest method. The BPNN uses the classical 3-layer structure, the hidden dimension and the learning rate is set as 71 and 0.01. The SVM uses the linear kernel function and the Rand Forest runs with 50 classification trees. The comparison results show that the CNN model presented in this paper has the higher recognition rate than other classical methods.

Table 3. The comparison between different methods

Method	Recognition rate
BPNN(hidden dimension = 71,learning rate = 0.01)	94.30%
SVM(linear kernel)	96.20%
Random Forest(50 trees)	96.80%
CNN(after pre-training)	98.70%

As for a single ball (a solid or a striped ball) being detected, its pattern recognition process takes about 20ms, in which the pre-processing function and the CNN recognition function respectively take 8ms and 12ms. The total time consuming of the pattern recognition is proportional to the number of detected balls. Taking the game start situation in Fig.11 into consideration, the number of ball on the table is 16 at most, then the total time consuming for multi-objective recognition is about 0.97s, the details are shown in Table 4. In some 8-ball game competitions, the players have to follow the time-limit rule, that is to say, the player must complete each of his shot within a specified time. The shooting time-limit set for the American and Chinese 8-ball billiards games are respectively 60s and 45s. As the multi-objective recognition method proposed in this paper is able to obtain the whole game situation in about 1s, the billiards robot will have plenty of time to work out his tactics before completing the shot action. In terms of the vision support system, the referee is able to obtain the game situation timely and give correct judgements.

In a word, the multi-objective recognition rate given in this paper is sufficient to satisfy the requirement of the billiards robot and other intended applications.

Table 4. The timing consuming details for the overall recognition pipeline

	Multi-ball detection	CNN recognition for single ball		Ball number
		pre-processing	CNN function	
Execution time(s)	0.65	0.008	0.012	16

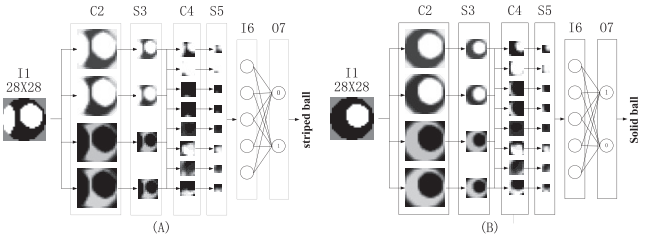


Fig.23. Feature maps in the convolution and subsampling layers

Fig.23 displays the sample feature maps of the tested ball images (one solid ball in Fig.23 (A) and one striped ball in Fig.23 (B)) being extracted from the convolution and subsampling layers of the well-trained CNN model. The convolution layer shows smooth image effect and signal characteristics enhancement. After being pooled, the dimensions of feature maps of subsampling layers will be reduced while the key information will be filed as well. Fig.24 shows the ball pattern recognition process in a specified game situation. After multi-ball detection shown in Fig.24 (A) and Fig.24 (B), the segmented ball image blocks will be greyed and processed by the morphologic operation as shown in Fig.24 (C), in which the cue ball and black ball (No.8) are expressed by white and black square respectively. Then the data matrix of the target ball image blocks of 24X24 pixels are insorted into the trained CNN model, and each ball pattern has been recognized correctlly. In the Fig.24 (D), the grey and black circle respectively represent the cue ball and the black ball, the circles coloured with all blue and blue-white represent the solid and the striped balls being recognized. Fig.25 shows some other image recognition effects in different game situations. The verification result has proved the correctness and effectiveness of the multi-objective recognition approach for 8-ball billiards robot vision system.

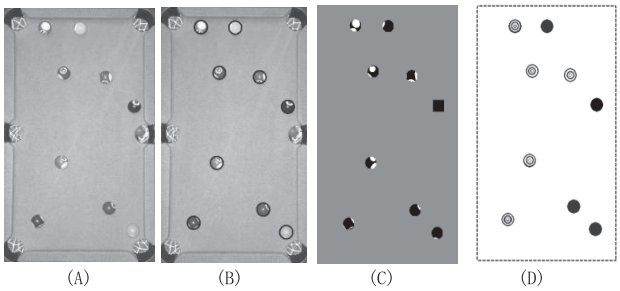


Fig.24. The ball pattern recognition process

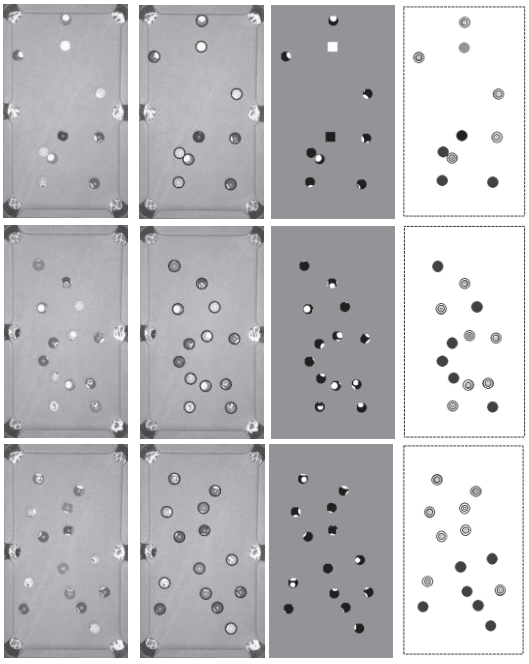


Fig.25. Ball pattern recognition effects in arbitrary game situations

6. Conclusions

This paper focuses on the research of the key technologies in the 8-ball billiards robot vision system and proposes an effective multi-objective recognition method for locating and identifying the targets on the table. The method is mainly divided into two parts, i.e. multi-objective detection and ball pattern recognition. The important conclusions are summarized as follows:

1. The normalized RGB colour space is suitable for the segmentation of foregrounds and background of 8-ball billiards robot vision system, which is able to effectively filter out the influence of illumination variation. Making use of the normalized RGB colour space and histogram statistics can effectively detect the edge points of balls and the table.

2. The combination of the improved HT algorithm and the LS method can detect the single ball position accurately. The improved HT algorithm can effectively eliminate the imperfect edge points, while the LS method can fit and obtain the most accurate circle with the least mean square error. The combination of the two methods is able to overcome the disadvantage of the LS method in noise sensitivity, and improve the accuracy of the HT algorithm.

3. Based on the above conclusions of 1 and 2, the multi-ball detection method proposed in this paper is able to effectively locate and segment of each ball on the table. The HT algorithm and LS method will ensure the correctness and robustness of the detection process.

4. The introducing of pre-training can effectively improve the training quality. The pre-trained CNN model will provide better groups of weight matrices for the initialization of actual CNN training.

5. With each target ball being isolated, the well-trained CNN model will finally recognize the ball pattern, and the pre-processing includes greying and morphology operation of the ball image block can make the training process more efficient. The trained CNN model provides good robustness for recognizing the variation of translation, and rotation in ball images. CNN has been proved to be a suitable and effective solution for the ball pattern recognition problem in 8-ball billiards robot vision system.

In our next step of research works, we will further improve the relevant algorithm and update the efficiency of multi-ball detection. As the billiards robot vision system proposed in this paper is based on the 2D up-down view, the camera will be installed at the upper side of the table with a certain angle due to the limit of competition conditions like the insufficient ceiling height and the position overlap of camera and lighting equipment. 3D reconstruction technology development is an aspect worthy of our consideration in future.

References

Ho, K.H.L., Martin, T. & Baldwin, J. (2007). Snooker Robot Player - 20 Years on. IEEE Symposium on Computational Intelligence and Games (CIG), pp. 1-8.

Ling, Y., Li, S., Xu, P. & Zhou, B. (2012). The detection of multi-objective billiards in snooker game video. IEEE International Conference on Intelligent Control & Information Processing (ICICIP), pp. 594-596.

Shen, W. & Wu, L. (2010). A method of billiard objects detection based on Snooker game video. IEEE International Conference on Future Computer and Communication (ICFCC), V2-251 - V2-255.

Wu, L., Liu, J., Cheng, Z., Wang, H. & Liu, Q. (2010). An effective multi-object detection approach. IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), pp. 1 -4.

Legg, P., Parry, M.L., Chung, D.H.S. & Jiang, R. M. (2011). Intelligent filtering by semantic importance for single-view 3D reconstruction from Snooker video. IEEE International Conference on Image Processing (ICIP), pp. 2385-2388.

Shih, C. & Chu, W.C. (2007). A Vision Based Interactive Billiard Ball Entertainment System. Proceedings of the First IEEE International Workshop on Digital Game and Intelligent Toy Enhanced Learning, pp. 200-202.

Takahashi, M., Kasai, T. & Suzuki, Y. (2008). Support system for pocket billiards. SICE Annual Conference, IEEE, pp. 3233–3236.

Höferlin, M., Grundy, E., Borgo, R., Weiskopf, D. & Chen, M. (2010). Video Visualization for Snooker Skill Training. Computer Graphics Forum **29**(3): pp. 1053–1062.

Gao, J., He, Q., Zhu, M., Liang, H. & Guo, X. (2015). Design of the multiple Neural Network compensator for a billiard robot. IEEE International Conference on Networking, Sensing, and Control (ICNSC), pp. 17-22.

Gao, J., He, Q. & Zhan, Z. (2017). Design of neural network controller for a billiard robot. Journal of Beijing University of Aeronautics & Astronautics, **43**(3): 533-543.

Cheng, B.R., Li, J.T. & Yang, J.S. (2004). Design of the neural-fuzzy compensator for a billiard robot. IEEE International Conference on Networking, Sensing, and Control (ICNSC), Vol.2, pp. 909-913.

Subban, R. & Mishra, R. (2012) Rule-based face detection in colour images using normalized RGB colour space— A comparative study. IEEE International Conference on Computational Intelligence & Computing Research (ICCIC), pp.1-5.

Dargham, J.A. & Chekima, A. (2006). Lips Detection in the Normalised RGB Colour Scheme. IEEE International Conference on Information & Communication Technologies, pp.1546-1551.

Cuevas, E., Wario, F., Osuna-Enciso, V. & Zaldivar , D. (2012). Fast algorithm for multiple-circle detection on images using learning automata. IET Image Processing **6**(8): 1124-1135.

Chen, X., Lu, L. & Gao, Y. (2012). A new concentric circle detection method based on Hough transform. IEEE International Conference on Computer Science & Education (ICCSE), pp. 753758-.

Lau, M.M., Lim, K.H. & Gopalai, A.A. (2015). Malaysia traffic sign recognition with convolutional neural network. IEEE International Conference on Digital Signal Processing (DSP), pp. 1006-1010.

Verma, A. & Vig, L. (2014). Using Convolutional Neural Networks to Discover Cogntively Validated Features for Gender Classification. IEEE International Conference on Soft Computing & Machine Intelligence (ISCMI), pp.33-37.

Kumar, V. D. A., Ramakrishnan, M., Malathi, S. & Kumar, V. D. A. (2015). Performance improvement using an automation system for recognition of multiple parametric features based on human footprint. Kuwait Journal of Science, **42**.1(2015):109-132.

Nallamuth, R. & Palanichamy, J. (2015). Optimized construction of various classification models for the diagnosis of thyroid problems in human beings. Kuwait Journal of Science, **42** (2):189-205.

Gao, J., He, Q., Zhan, Z. & Gao, H. (2016). Dynamic modeling based on fuzzy Neural Network for a billiard robot. IEEE International Conference on Networking, Sensing, and Control (ICNSC), pp. 1-4.

Submitted: 25/12/2015
Revised : 28/04/2016
Accepted : 17/05/2016

تصميم كُفوء لطريقة متعددة الأغراض للتعرف على نظام لوحة بلياردو من ثمانِ كور

جيابنغ جاو[1،*] ، كيويانغ هي[2] ، هونغ جاو[3] ، زيكسين زان[4] ، زهي وو[5]

[1،2،4،5] مدرسة علوم الفضاء والهندسة، جامعة بيهانغ، بكين 100191، الصين

[3] مؤسسة الصين زونجيان للهندسة، بكين 100191، الصين

gaojiaying—1988@163.com *

خـلاصـة

نناقش في هذا البحث بعض التقنيات الهامة التي تعتمد على استخدام الألوان لمعالجة الصور لنظام رؤية آلي للوحة بلياردو ذات ثمانِ كور، وسوف نقدم نظام تعَرُف كُفوء متعدد الأغراض. ينقسم هذا النظام إلى جزئين، نظام كشف متعدد الأغراض ونظام تَعَرُف لنمط الكرات. في المعالجة الأولية للصور يتم تطويع فضاء ألوان RGB ومدرج تكراري إحصائي للفصل بين الخلفية (لون غطاء الطاولة) ومقدمة الصور. تم تطويع طريقة تحويل هوج المحسن (HT) وطريقة المربعات الصغرى لإيجاد وعزل الكرة الواحدة في منطقة محلية. استُخدمت طريقة HT بغرض التخلص من الضوضاء المتركزة على نقاط الحواف. استُخدمت طريقة المربعات الصغرى للحصول بدقة على مركز الدائرة الذي يحقق أصغر متوسط لمربع الأخطاء. للكشف عن كرة واحدة في منطقة محلية، تم تشغيل طريقة الكشف المتعددة الكرات لإيجاد موقع كل كرة على الطاولة. تمكنت الطريقة المقترحة من أداء الكشف بدرجة دقة 99.4% في 0.65 ثانية في المتوسط. ودرجة الأداء أفضل من طريقة HT الدائرية وطريقة K-means العنقودية. إضافة إلى ذلك، تم تطويع طريقة الشبكة العصبية الملتفة (CNN) للتعرف على نمط كل كرة مستهدفة للتقسيم لتحديد إذا كانت من لون واحد أو مُخططة. من أجل تحسين درجة جودة تدريب CNN تم تطبيق فصل الألوان وعملية التكافوء لصورة الكرة المحدد لونها في المعالجة الأولية، ثم تم تدوير الصور لزيادة الدقة وتم إدخال مرحلة قبل التدريب للحصول على أمثل مصفوفة للأوزان الابتدائية. تم إرسال قطاع الصور التي تمت معايرتها للشبكة بغرض التدريب. في مرحلة التحقيق، اتضح أن CNN المُدرب أنتج معدل تعرف يزيد عن 98.5% ويتفوق على الطرق الكلاسيكية الثلاثة الأخرى. تم اثبات أن طريقة CNN صحيحة وكفوء وهي خطوة مبتكرة وهامة لعملية تصميم نظام رؤية آلي للعبة البلياردو من ثمانِ كرات.