

# Developing an Intelligent Table Tennis Umpiring System: Identifying the ball from the scene

Dr. Patrick K. C. Wong  
Department of Information and Communication Technologies  
Open University  
Walton Hall, Milton Keynes, UK  
Email: k.c.p.wong@open.ac.uk

## Abstract

*This paper reports further development of an intelligent table tennis umpiring system, of which the idea and plan was previously published at this conference in 2007. Briefly, table tennis is a fast sport. A service usually takes a few seconds to complete but an umpire needs to make many observations and makes a judgment before or soon after the service is complete. This is a complex task and the author believes the employment of videography, image processing and artificial intelligence (AI) technologies could help evaluating the service. The aim of this research is to develop an intelligent system which is able to identify and track the location of the ball from live video images and evaluate the service according to the service rules.*

*In this paper, the techniques of identifying a table tennis ball from the scene is described and discussed. A number of image processing techniques have been employed to identify and measure the characteristics of the ball. Artificial neural networks have been applied as a classifier. It classifies whether the detected object is not-a- ball, a ball on the palm or a ball in mid air. The system has been tested on 21 still images which contain pictures of ball-like objects, balls on the palm and in mid air. The preliminary results are very promising. Out of 83 objects, 82 have been correctly classified. The system will be further tested on video images once the video is captured and processed.*

*This paper also discusses the idea of implementing the final system as a multi-agent system, which the author believes it is appropriate for this application because multiple cameras will have to be employed to obtain accurate results.*

## 1. Introduction

This is a pilot study regarding the development of an intelligent system in aiding table tennis umpires to make accurate judgment about services. The idea and plan was previously published at this conference in

2007[1]. Briefly, a table tennis service usually takes a few seconds to complete. However, the umpire needs to take over ten observations and make a judgment before or just after the service is complete. This is a very complex task and requires a lot of judgments, even for an experienced umpire. With the help of image processing and artificial intelligent techniques, a computer system may be able to analyze the service and make a recommendation for the umpire to consider. To turn the idea into a reality, four main tasks have to be accomplished and they are:

- to identify the ball from the scene and tracking the location of the ball;
- to take the necessary measurement (e.g., ball rise and deviations)
- to evaluate the service according to the service rules, which can be found from the International Table Tennis (ITTF) Handbook [2].
- to make recommendations

The goal of this research is to develop an intelligent system which is able to accomplish the above tasks from one or more live video links. A multi-agent system is to be developed to coordinate the processes. The prototype system is being developed using Matlab and its Image Process Toolbox [3] and Neural Networks Toolbox [4].

In this paper, the focus is on the identification of the table tennis ball from the scene.

## 2. Image Processing

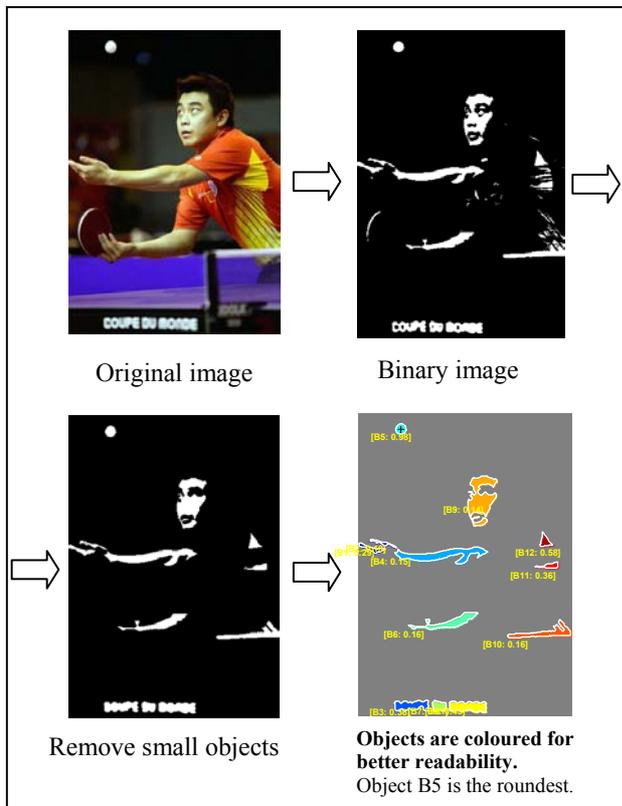
In the current developing phase, still images (rather than video images) are used to evaluate the system. These images were taken at real match scenes (Source: ITTF Photos Gallery [5]). Objects whose appearance similar to a ball, ball on the palm and ball in mid air can be found from these images.

### 2.1 First Attempt

To identify the table tennis ball from an image, one way is to conduct the following tasks:

1. Convert pixels that have similar colour of the ball to white and other pixels to black. This yields a binary image.
2. Connect the neighbouring white pixels together to form objects.
3. Remove irrelevant small objects and fill in holes.
4. Evaluate the properties of these objects and check if they are similar to a ball (e.g., size and roundness).
5. Classify whether it is an irrelevant object, a ball on the player's palm or a ball in mid air.

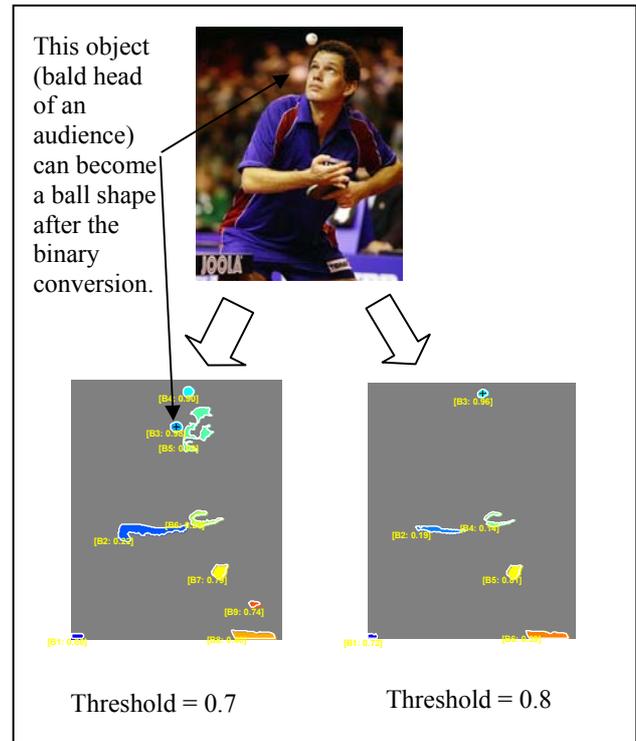
Figure 1 gives an example of the above process.



**Figure 1.** Identifying the ball from the scene

When converting the original image into a binary image, the aim is to turn pixels that have similar colour of the ball to white and other pixels to black. However, the degree of similarity (a threshold) is desired to define. In fact, the threshold must be set appropriately. Too low the threshold may result in too many irrelevant objects left in the binary image. Otherwise, part of or the whole ball may disappear from the binary image. Figure 2 illustrates an example, which shows that a small change in the threshold value resulted in different numbers of objects remained in the binary image. The shape of objects can be affected too.

Because of the sensitivity of the threshold value, it is impossible to find a threshold value that is appropriate for all match scenes. To combat this, a heuristic is needed to estimate an appropriate threshold for each image. Details of the technique will be described in Section 2.2.



**Figure 2.** An example showing how the threshold can affect converted binary image.

## 2.2 Two-Pass Method

One simple way to estimate the threshold may be to use the colour of a typical ball as a reference. However, the colour of the ball is not usually uniform, i.e., some parts are brighter and other parts are darker. Furthermore, it is still difficult to define the colour of a typical ball. The brightness of the ball varies at different locations, never mind at different match scenes. Statistical methods may be employed to estimate the average brightness of a typical ball by taking account of a large number of samples. However, in occasions such as the brightness of the scene is darker or brighter than an average scene, this threshold may still not be appropriate. This may lead to part of the ball is missing or many objects appear when converting the image in a binary picture.

To tackle this problem, a two-pass threshold method has been developed. In the first pass, a high threshold is applied so that all but those objects which have very

similar colour of the ball are eliminated. In the second pass, a low threshold is used but only applied to the objects detected in the first pass and their neighbourhoods. This way, most of the irrelevant objects are removed in the first pass. It does not only reduce the processing time required for analysing the image but also reduce the chance of having an object that has similar shape of the ball after the binary conversion (see the annotation in Figure 2). The remaining objects can be thoroughly investigated in the second pass as more details are revealed. Figure 3 shows an example of the two pass conversion.

### 2.3 Object Evaluation

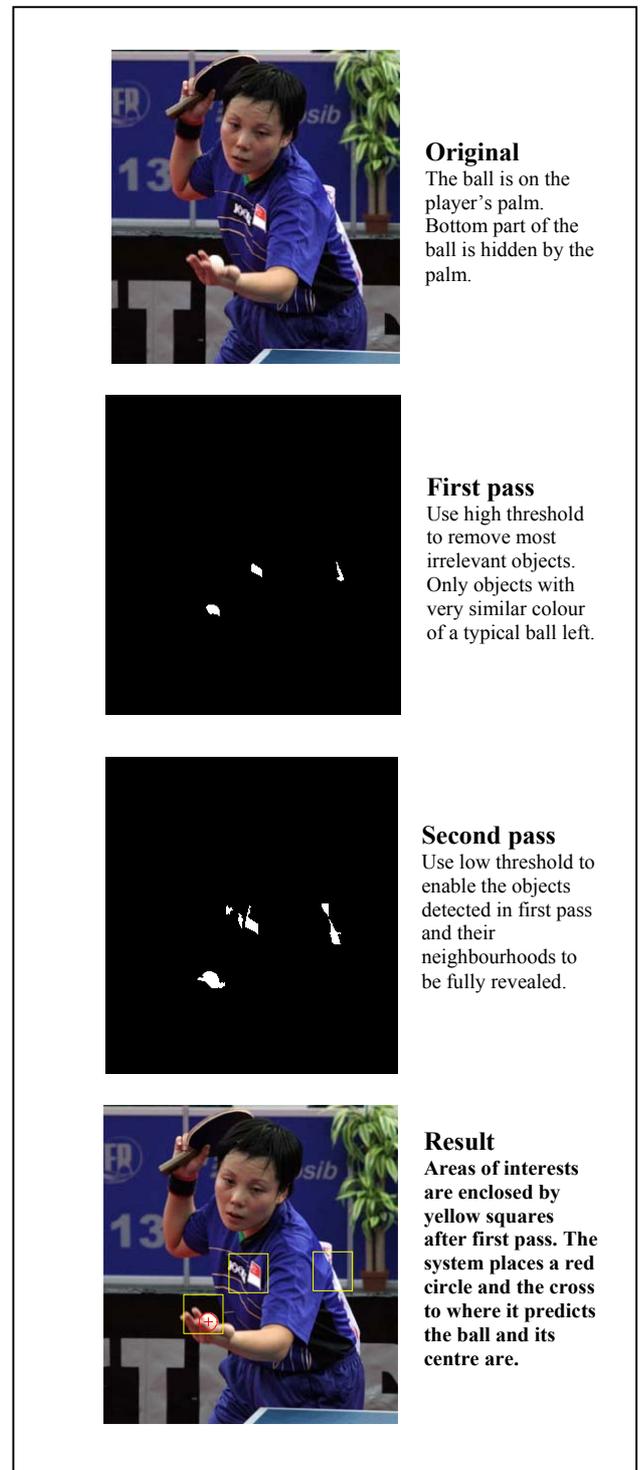
After the binary conversion, objects detected are to be evaluated. To check whether the detected objects are the ball, a number of features are to be evaluated and compared with a typical ball. Details of these features and how they are obtained are shown in Table 1.

**Table 1.** Features of a ball and how they are obtained

Features	Description
Area (A)	Number of pixels the object possess
Maximum width	The distance between the leftmost and rightmost pixels of the object
Maximum height	The distance between the top and bottom pixels of the object
Perimeter (P)	The length of the object's boundary
Roundness	A metric is employed to estimate the roundness of an object: $m = \frac{4\pi A}{P^2}$  The metric will be equal to 1 if the object is a circle because A will be $\pi r^2$ and P will be $2\pi r$ .
Round top	If object is a ball, the top of it will be round. To check this, the pixels on the upper boundary of the object are to be used to estimate the radius and the centre of a circle by substituting their co-ordinates to the equation below: $r^2 = (x - x_c)^2 + (y - y_c)^2$ Where r is the radius and (xc,yc) is the centre. Then a few points from the upper boundary are to be tested and checked whether they lie on the circle.

Ideally, when each of the features of the object matches with a typical ball, the object can then be considered as the ball. However, in most occasions some of the features do not match. This may be caused by factors such as poor lighting and interference by the ball coloured background. Furthermore, when the ball

is on the palm, it will not be round as the bottom part of it will be covered by the player's palm. In order to be able to detect and classify the objects, a point system has been devised.



**Figure 3.** An example showing the two-pass conversion

The system awards particular points to an object when it matches a particular feature. The object which obtains highest point and has a rounded top is considered to be the ball. Table 2 shows the features and their associated points. More points are awarded to objects that are circular and have a rounded top as these objects are more likely to be balls. To classify the object detected, some simple rules are used and the method is shown in Table 3.

**Table 2.** Features and Points

Features	Points
Area (A)	1
Maximum width (diameter)	1
Maximum height (diameter)	1
Perimeter (P)	1
Roundness	3
Rounded top	2

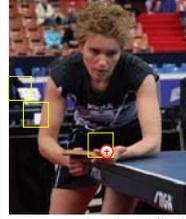
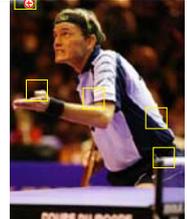
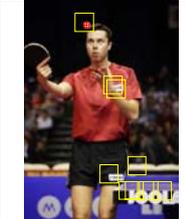
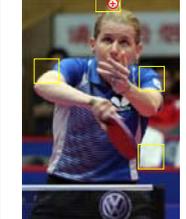
**Table 3.** Classification and rules

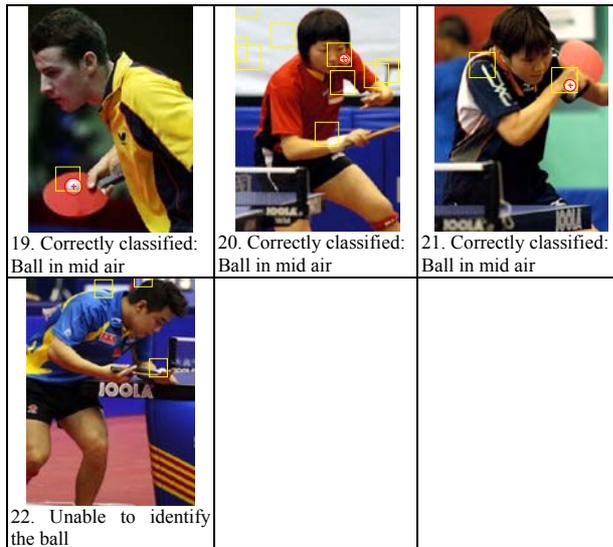
Class	Rule	Remark
Not a ball	Object do not have a round top and its size is much biggest or small than a typical ball	Most objects which are not a ball do not have a round top.
Ball on the palm	Object has a rounded top and has at least 2 points	Area, perimeter and roundness are unlikely to match as the bottom part of the ball is hidden
Ball in mid air	Object has a rounded top and has at least 7 points	Ideally all six features should match, but the system accepts a minor imperfection

## 2.4 Results

The ball identification system was put to test with 22 still images, which were captured in real match scenes and contain situations where the balls are on the palm and in mid air. In these 22 still images, the system extracted 83 objects which have similar colour and size of a typical ball. The test results are very encouraging. The system was able to correctly classify 80 out of 83 objects (2 non-classifications and 1 misclassification). The system can also highlight the ball on the image by using the radius and centre of the object which is considered to be the ball. Out of 20 occasions where a ball is thought to be detected, 19 occasions the system can accurately draw a circle around the ball and estimate its size. Table 4 shows the results in more details.

**Table 4.** Results of the tests

 1. Correctly classified: Ball in mid air	 2. Correctly classified: Ball on the palm	 3. Correctly classified: Ball in mid air
 4. Correctly classified: Ball in mid air	 5. Correctly classified: Ball on the palm	 6. Wrong object was identified
 7. Correctly classified: Ball in mid air	 8. Correctly classified: Ball in mid air	 9. Correctly classified: Ball in mid air
 10. Correctly classified: Ball in mid air	 11. Correctly classified: Ball in mid air	 12. Correctly classified: Ball on the palm
 13. Correctly classified: Ball in mid air	 14. Correctly classified: Ball in mid air	 15. Unable to identify the ball
 16. Correctly classified: Ball in mid air	 17. Correctly classified: Ball in mid air	 18. Correctly classified: Ball in mid air



### 3. Artificial Neural Networks

Artificial neural networks (ANN) are famous for their pattern recognition ability. Although the simple classification system described in Section 2.3 produced encouraging results, it is worth to investigate whether ANN can improve the results. In brief, ANN may be considered as a greatly simplified human brain. The network is usually implemented using electronic components or simulated in software on a computer. The massively parallel distributed structure and the ability to learn and generalise makes it possible to solve complex problems that otherwise are currently intractable.

ANNs are particularly good at classifying patterns. In this study, they have been employed to determine whether a detected object is a ball and whether the ball is on the palm or in mid-air. More information about ANN and their applications can be found in [6].

Two type of ANNs have been applied, namely the multi-layer perceptron (MLP) and the radial basis function network (RBF). MLP is a multi-layer feedforward network and employs supervised training, which is usually a gradient descent method such as backpropagation. Sigmoid functions are usually the activation functions for the hidden and output neurons. It is the most popular network architectures. It has been successfully deployed in solving many practical problems. Although it is well known that it can sometimes suffer premature saturation and difficult in designing the optimal structure, it is overall a robust type of network and easy to use and understand. On the other hand, RBF is a network that employs both unsupervised and supervised training. It is also a

feedforward network but always has 3 layers, namely input, prototype and output layers. Like MLP, the input neurons are not processing elements. They simply feed the input to the hidden layer. However, the prototype neurons work differently. Unlike MLP's hidden neurons, which pass weighted sum of the inputs to their activation functions, RBF's prototype neurons pass the Euclidean distance between the input and weight vectors to their activation functions, which are usually Gaussian functions. During unsupervised training, the network adjusts the weights between the input and prototype neurons in an attempt to minimise the Euclidean distance between input and weight vectors. When this is complete, a separate supervised training is conducted on the output neurons. The output neurons, which usually employ linear functions, are trained to associate each cluster with a particular class. The advantage of RBF is that it usually takes less time to design a workable network for a problem, especially when plenty of training patterns are available. However, it usually requires more processing neurons than MLP.

#### 3.1 Classifying the balls using ANN

ANNs have been employed in classifying objects detected using the system described in Section 2.2 and 2.3. The 83 objects which have similar colour and size of a typical ball were extracted from the 22 still images by the system described in Section 2.2. The six features of these objects, listed in Table 2, form the basis of the input part of the training patterns. To be precise, the inputs are:

- Difference between the object's area and a typical ball's area;
- Difference between the object's maximum width and a typical ball's diameter;
- Difference between the object's maximum height and a typical ball's diameter;
- Difference between the object's perimeter and a typical ball's perimeter;
- Roundness value
- Whether the object has a rounded top (Yes/No)

As for the desired outputs, which indicate the classes to be classified, there are 3 binary elements. When the desired outputs are "1 0 0", it represents the pattern belongs to class 1, which means the object is not a ball; "0 1 0" represents class 2 and so on. Full list is shown in Table 5. Table 6 shows an example training patterns.

**Table 5.** Desired outputs

Desired Outputs	Description
1 0 0	Object is not a ball
0 1 0	Ball is on the palm
0 0 1	Ball is in mid air

**Table 6.** Example training pattern – Ball is in mid air

Inputs						Desired output		
0.95	0.37	0.19	0.25	0.23	1	0	0	1

The whole data set consists of 83 patterns. 66 (80%) patterns were randomly selected as training patterns and the remaining 17 (20%) patterns were used as validating patterns during training. The training stopped when the network responses to the validating patterns was no longer improved. The ANNs were then tested with the whole data set.

### 3.2 Results from MLP and RBF Network

A number of MLP network structures were experimented and found that a MLP network with 6 input, 10 hidden and 3 output neurons was most successful. To reduce the chance of getting bad result because of premature saturation, the same MLP was re-initialised, re-trained and re-tested for more than 100 times. The best trained MLP can correctly classify 81 out of 83 objects (two class 2 objects were misclassified as class 1).

Similar to the procedure described above, a number of RBF network structures were trial and found that a RBF network with 6 input, 67 prototype and 3 output neurons produced the best result. The best trained RBF can correctly classify 82 out of 83 objects (one class 2 objects were misclassified as class 3).

### 4. Multi-agent system (MAS)

To obtain a more accurate and reliable measurements, multiple cameras should ideally be employed. The cameras should be situated at positions where the umpire and assistant umpire are. Additional cameras may be fixed high above the table to take aerial views. With many data feeding in from these cameras, it may be best to employ a MAS to coordinate these complicated processes. Each camera may be associated with an agent who can make independent decision based on the data it received. These local decisions can then be fed in to a higher hierarchy agent for consideration of the final decision. This higher hierarchy agent should have the ability to evaluate the decisions made by the camera agents. Factors such as the position and angle of the camera, agreement among agents and prior experiences can form the basis of a

heuristic that is applied to reason the final decision. This analogy is similar to what the real human umpiring system does, i.e. the umpire takes recommendation from the assistant umpire and makes the final decision based on both their judgments.

### 5. Conclusion

An intelligent system has been developed to identify table tennis balls from match scenes. The preliminary results are very encouraging. The system currently can identify the ball, pin-point its location and estimate its size from still image with a good accuracy. ANNs have been employed in aiding classifications. The results are also promising. The RBF network classifies the data slightly better than MLP but it requires over 6 times more hidden neurons. However, the time required to find a network structure that can produce a satisfactory result is much shorter.

The author strongly believes the final system has the potential to evaluate table tennis services from video in real time. The immediate next stage of the research will focus on identifying ball from video images. A technique for reducing processing time in analysing consecutive frame of video images has been investigated. Briefly, the principle of the technique is based on the fact that the location of the ball in next frame should be in the neighbourhood area of that in current frame, assuming the video is captured at a high enough rate. As a result, only the neighbourhood area of the subsequence frame needed to be analysed and hence the processing time and accuracy should be significantly improved.

### 6. References

- [1] Wong, Patrick K. C. "[Developing an intelligent assistant for table tennis umpires](#)", in First Asia International Conference on Modelling and Simulation, Phuket, Thailand, 27-30 March 2007.
- [2] "International Table Tennis Handbook 2006/2007", [http://www.ittf.com/ittf\\_handbook/ittf\\_hb.html](http://www.ittf.com/ittf_handbook/ittf_hb.html), accessed on 20/10/2006.
- [3] The MathWorks Inc, "User's Guide of Image Processing Toolbox For Use with Matlab", USA, 2005.
- [4] The MathWorks Inc, "User's Guide of Neural Networks Toolbox For Use with Matlab", USA, 2001.
- [5] "International Table Tennis Federation's Photo Gallery" [http://www.ittf.com/front\\_page/ittf1.asp?category=photo\\_gallery](http://www.ittf.com/front_page/ittf1.asp?category=photo_gallery), accessed on 25/10/2007.
- [6] Hopgood, Adrian A. "Intelligent Systems for Engineers and Scientists (Second Edition)", CRC Press, USA, 2001, ISBN 0-8493-0456-3.