P-N 러닝을 이용한 실시간 축구공 검출 및 추적

황수걸[1], 이근[2], 이일병[1]
[1]연세대학교 컴퓨터과학과 인공지능연구실
[2]연세대학교 전기전자공학과 컴퓨터비젼연구실

e-mail :{handsomefox77@hotmail.com, leegeun@yonsei.ac.kr, cool2848@hotmail.com

Real-time Ball Detection and Tracking with P-N Learning in Soccer Game

Shuai-Jie Huang[1], Gen LI[2], Yill-Byung Lee[1]
[1]Artificial Intelligence Lab., Dept. of Computer Science, Yonsei University
[2]Computer Vision Lab., School of Electrical and Electronic Engineering, Yonsei University

Abstract

This paper shows the application of P-N Learning [4] method in the soccer ball detection and improvement for increasing the speed of processing. In the P-N learning, the learning process is guided by positive (P) and negative (N) constraints which restrict the labeling of the unlabeled data, identify examples that have been classified in contradiction with structural constraints and augment the training set with the corrected samples in an iterative process. But for the long-view in the soccer game, P-N learning will produce so many ferns that more time is spent than other methods. We propose that color histogram of each frame is constructed to delete the unnecessary details in order to decreasing the number of feature points. We use the mask to eliminate the gallery region and Line Hough Transform to remove the line and adjust the P-N learning's parameters to optimize accurate and speed.

1. Introduction

Sports video analysis, especially ball games like soccer, basketball, tennis, etc. has always received much attention from researchers. Despite a lot of research efforts in sports video, soccer video analysis still remains a challenging task. Analysis and summarization of soccer highlights are the active research topics. V.Pallavi.et.al [1] developed ball detection using static and dynamic features. They used the Hough circle to find the circle objects in the each frame and formed the candidates. By analyzing the feature, such as diameter, relative move of ball against the camera and so on, some of non-ball candidates are removed and by constructing acyclic graph and dynamic programming, the longest path is estimated for determining the real ball. Xiao-Feng Tong.et.al [2] processed the ball detection at coarse step with evaluation of color and shape and at fine step with examination for remained regions to determine the real ball. Afterwards, CONDENSATION algorithm was utilized to track ball. T.D'Orazio.et.al [3] developed the ball recognition using circle Hough transform and neural classifier. For processing the soccer ball, the procedure commonly consists of two aspects: detection and tracking. For detection, the feature extraction is the most important, including shape, color, size, position. Then for tracking, in general, the several methods would be used: markov chain, optical flow, Karman etc.

In this paper, the preprocessing decreases the data calculated by the P-N Learning [4] and adjusting parameters of the P-N Learning [4] makes it be fit for the soccer ball detection and tracking. The flow chart is illustrated as the Figure 1.

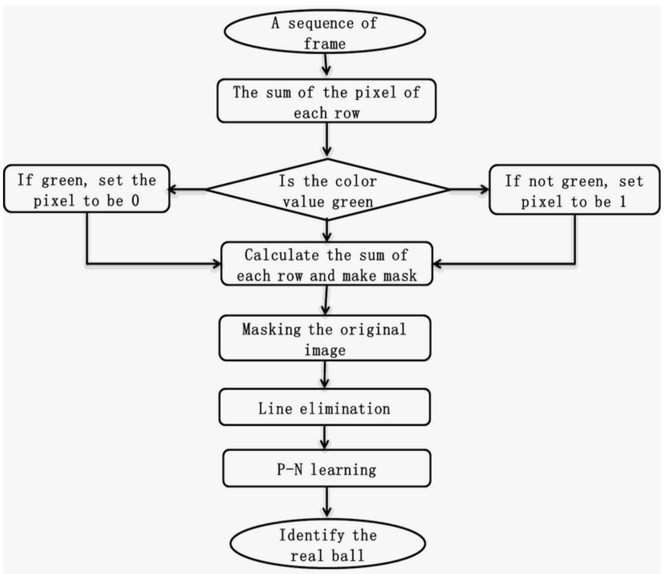

Figure 1. The framework of ball detection and tracking using L-N learning

2. Proposed method

For eliminating the unnecessary detail, we first consider to remove the gallery region and unnecessary objects to reduce the impact. For the two parts, we try to use masking and Line Hough Transform.

2.1 Masking

In the long view, the screen of soccer ball would be treated as 2 parts during the most of time which are gallery region and field. By making histogram this characteristics can be shown, which is illustrated in Figure 2.

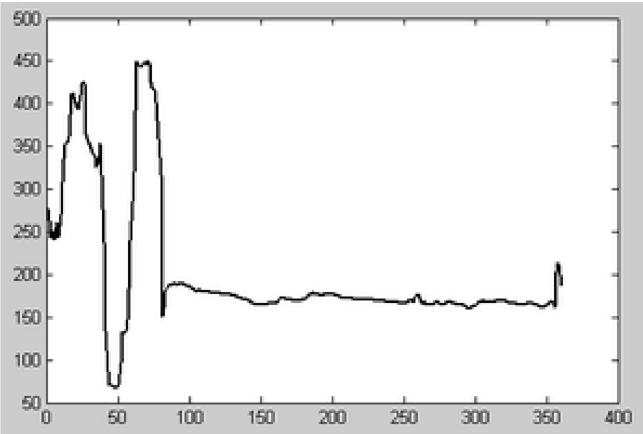

Figure 2. In HSV color space, the histogram of color distribution in image, where the X-axis means image column and Y-axis means image row.

We can observe the distinct variety between the first 100 columns and following 260 columns, which means the difference of between the field and gallery region is obvious in color. The part of color variety is the gallery region because in this region, the color variety is serious, the color value is obviously higher than the remainder because the value of green belongs to the low value region and the green is the dominant color in field result in that the sum of the color value in field is much lower than in gallery region.
We can use these characteristics to make mask. First, we change the RGB space into the HSV color space. After hue channel normalization, according to the experiment, the range of green is approximately from 0.2 to 0.4. We test the color value for each pixel and we set the pixel to be 1 if its color value is within the range, otherwise to be 0. After that, by calculating the sum of the each row and setting a threshold, if the sum is more than the threshold, we set the each pixel in this row to be 0, otherwise to be 1. Finally, we use this binary image to mask the original image and the gallery region will be removed.

2.2 Removing the white line in the field

Due to the P-N learning [4] using the 2bitBP [5] feature where the rectangle features produced by gray scale estimation are used to denote the features. It is considerable to remove the object which gray scale is similar to ball. With consideration about that the diameter of the ball is similar to the lines, we propose to use the Line Hough Transform to remove them. But before do it, we are supposed to use Canny edge detection and changing the contrast would be better for Canny operation. We set the gamma to be 0.5 so that the high grays scale could stand out. We use to Line Hough Transform for the line detection.
Line Hough Transform is the important method in the image processing. In the traditional X-Y plane, the line equation is:

$$y = kx + b \tag{1}$$

In this plane, all of point which is on this line will satisfy equation (1). If a point(x_0, y_0) is given, equation (1) can be rewritten:

$$y_0 = kx_0 + b \tag{2}$$

Then, the x and b will be perceived as unknown. So, in x-b plane, the equation (2) is a line either. All of the points which satisfy equation (1) in X-Y plane will be a set of lines in k-b plane and the line which satisfies equation (1) in X-Y plane will be a point of intersection of this set of lines in the k-b plane.

Based on this principle, we project points on the edge which be found by Canny into the k-b plane and set a threshold, if in the k-b plane, the number of the lines at a point of intersection is more than this threshold, it can be considered as a line in the image. But the found edges would not be consecutive or due to the width of the line, these lines would not intersect on a point exactly in k-b plane. We need to set a range in which all of point will be perceived as intersecting a point which is a point of intersection with the most number of lines. We call this method as non-maxima suppression. But the non-maxima suppression just focuses on the broader region in the k-b plane for anti-noise and repeating detection to a same bold line. To the narrow region in k-b plane, we propose to consider the set of lines as a line. The problem is that the lines in the field usually are long, so for the same line, the two methods would produce two different result namely detecting two lines. In the condition, we select the bisector of the intersection angle of the two lines as real line.

2.3 P-N Learning

In P-N learning [4], the input image will be input to the classifier and structural constraints at the same time, and the result from structural constraints go to the training set to train. Finally, deliver to the classifier to classify the input. The result would be output as result, as well as delivering to structural constraints to optimize constraints [4]. But, the classifier uses the scanning window strategy [5] and randomized forest classifier which would produce too many ferns, and the number of corner detection in the LK [6] tracker cannot be ignored, so preprocessing the image is necessary.

3. Experiment

We use the Visual Studio 2010 with OpenCV2.2 and Matlab 2010a to make our experiment. To test the performance of our method, we used two videos of

Netherland VS Uruguay and Netherland VS Brazil in the 2010 World Cup for analysis under the day and night respectively.

In experimentation, we first change the RGB color space into the HSV. The range of green value of hue can be obtained, which is between 0.2 and 0.4 after hue channel normalization. The size of the image is 640×320 and the threshold is set 200 for sum of the H channel in each row. Finally, mask is dot product by original image and we will obtain the image as following Figure 3.

Figure 3. The result of masking

Due to color and shape of the ball which is distinct in the field, we consider to make details of the ball stand out and eliminate other parts. Also, 2bitBP [5] and Line Hough Transform need to calculate the grayscale, so enhancement of contrast is considered, which is illustrated in figure 4.

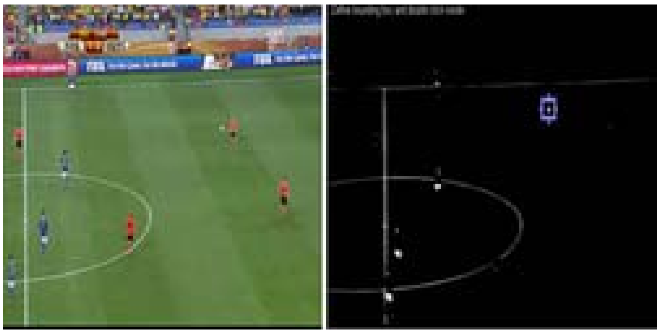

Figure 4. Show the original image and contrast image, the feature of high gray level is remained.

In the line elimination phase, we set the number of the peaks to be 10 due to the width of the line, and set the minimum length of the line to be 0.01 and connect them as a line. In the k-b plane, if the two lines are too close they would be perceived as on a same line which threshold of the distance is set to be 1000. If distance is less than this threshold, the line segments are merged into a single line segment which is illustrated in Figure 5.

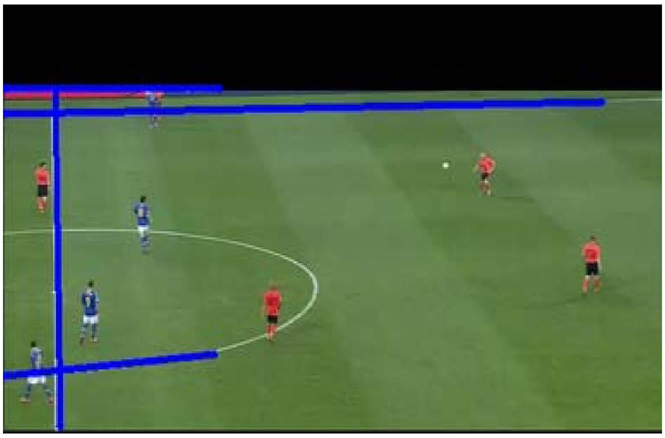

Figure 5. We can see majority of the lines can be detection

In experiment, the parameter of the P-N Learning also has to be reset. The small size of the ball results in the minimal size of the object's bounding box in the scanning grid which be set to be 5 pixels. And the number of the trees in the random forest is set to be 8 because we consider that the condition in the field, appearance and texture of soccer ball are less complicated than others. So it is acceptable to decrease processing time by scarifying a certain accurate.

Our objects are to reduce the processing time and improvement for accuracy of ball tracking. The difference of illumination will make obvious effect, so the discussion is necessary under the different illumination conditions. The result of experiment is shown as the following tables:

	P-N Learning		**Improved P-N Learning**	
	day	**night**	**day**	**night**
Frame number	**101**	**114**	**101**	**114**
Ball-frame	**101**	**112**	**101**	**112**
Ball detection correct	**92**	**103**	**99**	**109**
Accuracy	**92.0%**	**91.9%**	**99.0%**	**97.3%**

Table1. Accuracy of ball detection

	P-N Learning		**P-N Learning with preprocessing**	
	day	**night**	**day**	**night**
Frame number	**101**	**114**	**101**	**114**
Min time(fms)	**0.691**	**0.580**	**0.773**	**0.846**
Max time(fms)	**0.866**	**0.940**	**0.916**	**1.190**
Avg. time(fms)	**0.691**	**0.731**	**0.875**	**1.059**
Avg. elapsed time	**day**		**night**	
	22.6%		**44.9%**	

Table 2. Elapsed time

As shown the tables above, the accuracy of improvement P-N learning is advanced 7% in day and 3.6% in night. As long as the parameters such as minimal size of normalized patch in the object detector, the position and size of the bounding box is set correctly, the accuracy of P-N learning can be acceptable. The advancement of accuracy is attributed to the decrease of candidates. However, the process time is reduced dramatically attribute to the following reason: the number of candidates is reduced dramatically. Due to the small size, shape and color of the soccer ball, the great number of objects which have similar feature would occur. In addition, that P-N learning uses the 2bitBP feature, random decision forest [7] and forward-backward error [8] result in that time consuming which is considerable. So removing the gallery region and reducing the number of the useless information in field are important. For the 2bitPB feature, to change the contrast and to remove the white lines are necessary.

4. Conclusion and Future works

In this paper, we apply the P-N Learning [4] in the soccer ball detection and tracking. This method with compensation each other between the detection and tracking solves the obstacle and object lose after tracking failure in a way. But due to time consuming, in the preprocessing, filtering to candidate should be considerable.

In future, we will enhance the performance of application of P-N learning in soccer game. We will try to find a new classifier instead of the random decision forest to offer a nice combination of both of accuracy and speed.

Reference

[1] V.Pallavi, Jayanta Mukherjee, Arun K.Majumdar, and Shamik Sural.”Ball detection from broadcast soccer videos using static and dynamic features” J.Vis Commun. Image R 19(2008) 426-436

[2] Xiao-Feng Tong, Han- Qing Lu and Qing-Shan Liu.”An Effective and Fast Soccer Ball Detection and Tracking Method”. Proceedings of the 17th international Conference on Pattern Recognition(ICOR’04)

[3] T.D’Orazio, C.Guaragnella, M.Leo and A.Distante “A new algorithm for ball recognition using circle Hough transform and neural classifier”. 2003 Pattern Recognition Society.

[4] Zdenek Kalal, Jiri Matas and Kry Mikolajczyk.”P-N Learning: Bootstrapping Binary Classifier by Structural Constraints”. 2010 IEEE computer society conference on computer vision and pattern recognition

[5] Zdenek Kalal, Jiri Matas and Krystian Mikolajczyk, ”Online learning of robust object detectors during unstable tracking”.2009 IEEE 12th international conference on computer vision workshops, ICCV workshops

[6] Jianbo Shi and Carlo Tomasi,“Good Features to Track” computer vision and pattern recognition,1994.Proceeding s CVPR ’94., 1994 IEEE computer society conference on

[7] V.Lepetit, P. Lager, and P. Fua, “Randomized trees for real-time keypoint recognition.” CVPR, 2005.5

[8] Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas “Forward-Backward Error: Automatic Detection of Tracking Failures”. International conference on Pattern Recognition, 23-26 August, 2010, Istambul, Turkey