Playfield Detection Algorithm Based on AGMM in Sports Video

Ding Bin and Ma Yan
Military Department of Heilongjiang Ba Yi Agricultural University, Heilongjiang, Daqing, China
Wu Yiwei
School of Computer, Valparaiso University, Chicago, USA

Abstract—Because under the condition of multiple factors, the sports video content analysis of the traditional playfield detection algorithm can cause the non-homogeneity of playfield color, the playfield detection algorithm based on AGMM is proposed in this paper. Firstly, extract the image at random from the sports video and make automatic analysis of the image color and find the approximate values of color distribution. Then, use GMM fitting color distribution, the playfield test results and incremental expectation maximum algorithm to update model parameters so as to get more accurate parameter estimation which will be used to make classification for pixels of playfield and non-playfield in subsequent image. Finally, the test of algorithm performance is carried on and the proposed algorithm in this paper is used to classify sports scene, experiments. The experiment shows that compared with the traditional playfield detection algorithm, a playfield detection algorithm based on AGMM model is greatly improved on the recall ratio and accuracy.

Index Terms—Video Extraction, Moving Target, Color Area Detection, Parameter Drift

I. INTRODUCTION

Based on the single Gaussian background model, mixture Gaussian background model is widely used. The single Gaussian model uses a single Gaussian distribution to represent background, which can handle simple scenario with tiny and slow changes, while it is not applicable for the scenario that the background has great changes or mutation and the background pixel values are multimodal distribution [1]. Because when the background changes faster, its distribution doesn’t gradually transit from a relatively stable unimodal distribution to another unimodal distribution [2-3].

Because the distribution of background pixels is multimodal, according to the ideas of the single mode method, multiple single mode set is used to describe the changes of pixel value in a complex scene. GMM just uses multiple single Gaussian functions to describe the background scenario with multimodal. It better describes the distribution of background model in complex scenarios and has better effect on the processing of mutation background, background disturbance and the other situations.

Many researchers have devoted to the research field of sports video content analysis. The playfield detection plays an important role in sports video content analysis, such as the classification of sports video scene, object detection (including ball, players and marking, etc.), motion video structure analysis and so on [4-6]. Therefore, the attention of researchers is brought to playfield detection.

M Luo and others used the predetermined green stadium to detect stadium in HSI space distribution area [7]. Although this method can well detect the stadium area, it also has significant limitations. Because different sports stadiums have different color, even if they are all football matches, the stadium's color is also different [8]. A Ekin and others made introduction on the pitch detection algorithm in detail for the first time. In their work, the author used the two color space (the space can be chosen from HSI, RGB and YCbCr color space, etc). One is taken on as the basic color space and another is the control space which has the effect of control and supplement on the former. In basic color space, the average color of the playfield and threshold are used to detect playfield pixel [9]. As the prephase work of this paper, S Q Jiang and L Wang et al. adopted GMM to detect stadium area, but they can’t adaptively adjust the model parameters [10].

Playfield detection plays an important role in sports video analysis. L X Xie et al. use main color information to analyze football video structure [11-13]. A Ekin et al. uses the playfield test results to make classification for the camera in sports video. According to the playfield test results, O Utsumi et al make further detection for moving target in video image, including the players and the ball. Based on playfield detection, J Assfalg et al. make semantic hierarchical analysis for football video and automatically extract exciting sections [14-17].

This paper adopts GMM to model the playfield color. The reasons that use this method are based on the following consideration: a) A playfield may contain more than one color. For example, football playfield may have stripes and the goal area in some playfields has presented land color because of the frequent trample. b) The same playfields show inconsistent color in different cameras, different illumination and in different shooting angles. Especially, the broadcast video which is synthesized from...
multiple camera data often presents this phenomenon. Therefore, the appropriate adjustment for model parameters should be made in the process of operation.

Based on the above-mentioned factors, the GMM is proposed to module playfield color and adaptively adjust parameter according to IEM algorithm. The main characteristics of this method are as follows: a) GMM can more accurately describe the playfield color. Especially, when the color is not consistent on the playfield or there exists a lot of noise, the method has stronger probabilistic description ability. b) Because of the use of IEM algorithm, the model can constantly adapt to the change of playfield color in video. 3) IEM is a kind of online learning method that the system doesn’t have to keep large numbers of sample points.

The playfield detection algorithm is applied to the football video analysis in this paper and according to the test results of playfield, the lens in the video can be divided into inside and outside scenario. For inside scenario, this paper proposes the idea that according to the fitting of the above and under edge points of playfield area in image, the inside scenario is further divided into features and medium long shot scenario and use the slope of fitting straight line of above edge to detect the area near the goal.

This paper makes developing and innovative work from the following aspects.

Aiming at the inconsistence of playfield color which is caused by the traditional playfield detection algorithm of sports video content analysis under the condition of multiple factors, the detection algorithm based on AMG is proposed. According to the initialization of the sports video samples, extract the color of pictures of video sequence as the data and calculate the color histogram of all pixels in CbCr color space. Then, adaptively update the various parameters of the AGMM so as to get more accurate parameter estimation, which is used to make classification for the pixels of playfield and non-playfield in subsequent image.

In order to further verify the correctness and effectiveness of playfield detection algorithm based on AGMM, the experiment is made to classify inside and outside scenario and football video scenario and the recall ratio and accuracy are considered as the test standard. The experiment shows that compared with the traditional playfield detection algorithm, the playfield detection algorithm based on AGMM is greatly improved on the recall ratio and accuracy. The sports video scenario classification method is effective and it can be easily extended to the video analysis of hockey and other sports.

II. PROPOSED SCHEME

Firstly, introduce the system framework of the playfield detection algorithm based on GMM. Then, introduce the framework of each part in detail.

A. System Framework

Figure 1 shows the framework of the playfield detection algorithm based on AGMM. Its process is as the follows. First, according to detect the primary color in the histogram area, initialize the parameters of GMM. Then, the pixels of playfield area are obtained by post-processing. Finally, through incremental EM algorithm, modify the parameters of GMM. The following sections introduce each step in the system process in detail.

B. Sample Initialization

Through observation, it can be found that in sports program with playfield, the color of a period of video image mainly distributes in a few more dense area in the color space. Based on this characteristic, this algorithm looks for a dense area with larger number in color histogram for as the probable distribution area of the playfield color.

This algorithm extracts the first few minutes’ picture color data of video sequence and calculates color histogram $H_1$ of all these pixels in CbCr color space. Choosing CbCr color space is based on the following two points. a) The current popular video compression standards (such as MPEG and H26L) all adopt this color space so as to reduce computing resources consumption brought by the color space conversion. b) The CbCr that removes the brightness information can better describe the color characteristics of the playfield area and reduce the effect of shadow, light and the other factors that are sensitive to luminance on test results.

Algorithm 1 describes the process detecting the main color area in CbCr space. It should be pointed out that this algorithm looks for the connected area near two peaks in CbCr color space histogram. The aim is to avoid the condition that the value of a certain color is very big in histogram of, but the number of covered pixel of the connected area is very small. Because in sports video, playfield pixel inevitably has a large proportion and the pixel located in the area with few pixels can’t be playfield pixel, ignore the corresponding area with the fewer pixels in Sum1 Sum2 (The phenomenon is often caused by the video decoding process).

Algorithm1. The detection of the main color area in color space

(a) Ensure the location of the largest peak point in $H_1$ called $P_1$.
(b) Look for the bin that is connected with $P_1$ and its value is greater than $s \times \text{Value}(P_1)$. Calculate the sum1 of all bins in connected area. Then, abandon this connected area from the original histogram and the rest
area is set to H2. Among them, the s is proportional coefficient which is set to 0105 in this paper.

c) Make (a) and (b) operation in H2 and the sum of bins in the area that is connected with the largest peak point P2 in H2 is set to Sum2.

d) Return the connected area with more pixels in Sum1 and Sum2, which is taken on as the general distribution area of playfield color.

The returned area in 4) of algorithm 1 is considered as the general area of playfield color distribution, but this area is often inaccurate and there are also many playfield pixel colors that aren’t included. Based on this consideration, the next section proposes using IEM to learn model parameters on line.

C. The Parameters Adjustment of Model

As mentioned in section 1, the playfield color may be inconsistent (such as the stripe lawn in a football playfield) and contain noise. In addition, due to the different shooting angles, light and different cameras, the playfield color will be constantly changing in radio show, which requires the color mode used by the algorithm should be able to describe complicated probability distribution and at the same time, it can continuously adjust its parameters, with the change of playfield color in video. The GMM using IEM algorithm to update the parameters has such characteristics.

a) GMM

The GMM p is made up of multiple Gaussian functions, which is as the format (1).

\[ p = \sum_{i=1}^{k} \pi_i p_i \]  

\[ p_i(x, q_i) = \frac{1}{(2\pi)^{d/2} |\Sigma|} \exp\left(-\frac{1}{2}(x - \mu_i)(1-x)^t \right) \]  

\[ \int_{x_i}^{m} \frac{dx^{(i)}(m)}{dm} + a \int_{x_i}^{m} \frac{dx^{(i)}(m)}{dm} = \sum_{i=1}^{m} dt \]

Each gaussian function is determined by the parameter set \( \theta \), including mean vector \( \mu_i \) and covariance matrix \( \Sigma \), which are the dimensions of the sample points X. \( \pi_i \) is the weight of each gaussian function and k is the number of gaussian items. The collection \( \Omega = \{ \pi, \theta \} \) contains all of the unknown parameters, which belong to a certain parameter space. Generally, people use BEM to estimate the parameters \( \Omega \) of this group. In this paper, there are three GMMs and two of them are use to fit field color, another gauss items used to fit the noise in playfield (such as the land color in bad turf of playfield).

b) AGMM

In order to more accurately estimate the color distribution of the playfield in video stream, this paper uses IEM algorithm to adjust the GMM parameters on line. Just as the idea proposed by R M Neal and the others, IEM is more suitable to deal with the arrival in succession of sample points which is often encountered in the video analysis. By statistical point of view, the E and M in EM algorithm can be regarded as common maximum of a certain function. The function is described as format (4) and the IEM algorithm and BEM algorithm can be derived from this function.

\[ F(p, \ell) = \sum_{x \in \Omega} I_{\ell}(y, x)[\theta] \]  

\[ \sum_{j=1}^{k} I_{\ell}(y, x) + H(p) \]

\[ F(p, \theta) = \text{Exp}[I_{\ell}(y, x)[\theta]] + H(p) \]  

\[ F(p, \theta) = \text{Inx}(y) - \text{Inx}(y) \]

In format (6), \( y \) is the value of hidden random variable \( Y \), which corresponds to the weights of every normal distribution in GMM. \( x \) is the value of observed random variable X, which corresponds to the pixel values. \( \Omega \) is the model parameters.

For BEM, its E step and M step can be expressed as the follows through the format (6) and t is the time step.

\[ E: \text{select the } p^{(i)} \text{, and so } F(p^{(i)}, \theta^{(i)}) \text{most;} \]

\[ M: \text{Set } \theta^{(i)} \text{ to } \theta, \text{ and so } F(p^{(i)}, \theta) \text{most.} \]

Set some independent observed samples \( x = (x_1, x_2, x_3, x_4) \) and theirs hidden information \( y = (y_1, y_2, y_3, y_4) \). The function F can be written as the follows.

\[ F^{(i)}(p_i) = (x_i^{(i)}(R_j) - x_i^{(i)}(R_{j-1}))\Delta H(p_i) \]

\[ \text{Inx}(y) - \text{Inx}(y) \]

\[ F(p, \theta) = \sum_{i=1}^{m} F_i(R_i, \theta) \]

\[ F(p, \Theta) = \sum_{i=1}^{m} F_i(p_i, \Theta) \]

\[ F_i(p_i, \Theta) = E_{p_i}[I_{\ell}(y_i, x_i[\Theta]) + H(p_i)] \]

IEM adoptes the structure of format (9). It is unnecessary to wait the involvement of all data in E step and then perform M step, namely when every sample executes one E step, one M step is also executed. That is to say, the effect of every sample on distribution P will be soon showed in the M step so as to realize the online update parameters. If the joint distribution of X and \( Y \) is exponential form, a sufficient statistic can be calculated in E step and M step calculates the maximum of \( \Omega \) in the statistics. Use \( si \) si (yi,xi) to represent the statistics of the ith sample points in IEM and its E step and M step can be expressed as follows.

E: select the \( i \) sample points

Use \( s_j = s_i \), if \( j \neq i \).
Use $s_j = s_j \iff t \neq i$.

Use

$$s_j = E_{ij}[s_i(y_i, x_j)] \text{ inside } p_i(y_i | x_i, C^{j-i}),$$

according to IEM, the iterative process of adaptive Gauss mixture model is as follows:

$$k = \begin{bmatrix} w & n & 0 \\ w^2 & n^2 & \vdots \\ n^3 & w^3 & \vdots \\ 0 & w^{n-1} & \vdots \end{bmatrix}$$

(10)

$$\pi_k^{(n+1)} = \pi_k^{(n)} + \frac{1}{n+1} (p^c(w_k | x_k) + 1 - \pi_k^{(n)}),$$

$$\hat{\pi}_k^{(n+1)} = u_k^{(n)} + \frac{p^c(w_k | x_k) + 1}{\sum_{i=1}^n p^c(w_i | x_i)} (x_k + 1 - u_k^{(n)})$$

(11)

$$(x_k + 1 - u_k^{(n)} \sum_{i=1}^n p^c(w_i | x_i)),$$

$$\text{inside } k = 1, 2, 3, p^c(w_i | x_i) = \frac{p^c(x_i, q_k)}{p^c(x_i)}.$$

From format (11), it can be seen that the join of each new sample will have an effect on model parameters so as to make the model constantly close to the real distribution in the run. The algorithm itself doesn’t much increase computation, because in the process of pixel labeling, $\hat{p}(x_i, \hat{q}_i)$ and $\hat{p}(x_i)$ are both calculated and the algorithm doesn’t need calculating again. $\sum_{i=1}^{n+1} p(w_i | x_i)$ can be noted and accumulated. The only increased calculation is $(x_n + 1 - u_n^{(n)})(x_n + 1 - u_n^{(n)})$. For 2 d CbCr pixel values, there are four multiplication.

**D. Choose the Incremental Training Data**

![The incremental training set](image)

Figure 2. The incremental training set.

(a) The circle represents noise. After morphological filtering treatment, they are removed.

(b) The image after processed. The pixel within circle is considered as incremental training set.

According to format (6), the key problem of the online update model parameters is the selection of samples. First of all, use the current parameter $C^{j-i}$ transform binary image pixels into the playfield and non-playfield pixels and use the morphology filter to get playfield area. The pixel of playfield area will be used as incremental training data and join the update operations of model parameter so that the model can gradually learn the color distribution of all pixels in different cameras, different perspectives and different color of the light. Figure 2 shows the schematic diagram.

**E. Process Parameter Drift**

As all of the adaptive system, in the process of parameter adjustment, it may cause the estimated parameters drift away the real value. This algorithm adopts the following mechanism to determine whether it produces drift, because it uses initial parameters and current parameters at the same time every 30 s to make the playfield detection. Set the playfield color pixel ratios detected by the parameters of two groups of models in the same frame separately are Ri and Rc. If they are greater than the threshold value Td and $|Ri- Rc|$ is greater than threshold Tr, at this time, the current parameters of model are back to the initial parameters (i. Td = 0.5, Tr = 0.1). The former ensures a frame contains enough playfield pixels so as to judge whether model parameter drifts. The latter requires initial parameters of the model and the current parameters should have few differences on the test results.

**F. The Application of Playfield Detection in Football Video Analysis**

Playfield detection algorithm plays an important role in football video analysis. This section uses the testing results of the football playfield to classify the scenes in football match programs, including that football video sequence can be divided into inside and outside scenario; playfield scenario is divided into features and global scenario. For global scenario, it further detects the area near goal and provides clues for further semantic event detection (such as the test shot). Figure 3 depicts the classification process of scenario.

![The classification process of scenario](image)

Figure 3. Depicts the classification process of scenario.

**G. The Judgement of The Inside and Outside of Playfield**

When playfield area is smaller than a certain percentage in an image, it should be judged as outside area. Generally, this type of lens is not only used to show the audience and the bench, but also corresponds to the fragments of non-game.
H. The Classification of Inside Scenario of Playfield

For the inside scenario of playfield which is judged by the last step, it is further divided into features containing playfield area and global scenario. The latter is usually the main part of video stream with a large number. Because the exciting sections in football video tend to occur near the goal area, finding playfield area from the global scenario can compress a lot of game process, which provides help for the generation of summary.

This paper proposes using straight line to fit playfield area on the upper edge and bottom edge in image. According to the ratio of non-playfield area inside fitting lines and the slope of fitting lines classification, classify features, goal area, and non-goal area. The process is as follows.

a) The edge detection and fitting of playfield

Figure 4(a) and figure 4(b) separately describe the fitting process of playfield edge in medium long shot scenario and features. The first column is original image and the second column is the binary image after the playfield detection and the third column is the detected edge points. The process is that first of all, use rays which are perpendicular to image scan binary image from top to bottom edge points. The first playfield pixel point that meets in each column is considered as an upper edge point $e_{1}^{up}$ and then the upper edge point set $\{e_{1}^{up}, \cdots, e_{m}^{up}, m \leq \text{image width}\}$ is gotten. Similarly, if the rays scan images from bottom to top, you can get the bottom edge point set $\{e_{1}^{bottom}, \cdots, e_{n}^{bottom}, n \leq \text{image width}\}$. At this time, two straight lines are respectively used to fit the two collections and the two straight lines $y_{i} = K_{i}x + B_{i}$ and $i = 1, 2$ are gotten. The forth column is the fitting line.

b) The classification of features of inside scenario and global scenario

From the figures in the last column of the figure 4(a) and figure 4(b), it can be seen that the two fitting straight lines and edges of the image form a closed area. Assume that $MA$ is the biggest connected area of non-playfield area in the closed area and $A$ is the area of this closed area. If $\frac{MA}{A} > T_{s}$, it is considered as features of inside scenario. If not, it is global scenario, because in global scenario, the occupied area of the players and ball in image is very small. In this paper, $T_{s}$ is set to 0.25.

c) Goal area detection

For global scenario, because the goal area shows the model of figure 5 in image which is shot by main camera image, according to the fitting straight line slope $K_{1}$ of its upper edge, judge the right and left goal area and the judgment criteria is as follows.

$$S_{q} = \begin{cases} \text{The left goal area, } K_{1} \leq T_{\text{left}} \\ \text{The right goal area, } K_{1} \geq T_{\text{right}} \end{cases}$$

In this formt, $T_{\text{left}}$ and $T_{\text{right}}$ are two around two thresholds that judge the left and right goal area which are considered to be the middle area between them. This paper uses the value of ±0.1. Figure 5 shows several patterns of the goal area and figure 6 shows some results gotten from the experiment.

III. THE EXPERIMENT RESULTS AND ANALYSIS

In order to test the performance of the proposed algorithm, a large number of tests are conducted based on the actual data. The data recorded from the central TV sports channel and Changsha Television station sports channel whose time is more than 20h, including football, badminton, basketball, tennis and other games.

A. The Detaction Results of AGMM

In order to make quantitative evaluation for algorithm, the author manually tagged a part of video clips, including three sections of basketball, one sections of football, one sections of tennis and one sections of badminton. Each section of the video takes samples of 20 frames. Table 1 shows the accuracy of GMM, AGMM and histogram method in playfield detection. It can be seen that using the AGMM to make playfield detection can achieve better results. The precision is defined as:

$$\text{accuracy} = \frac{\# \text{PLP and PSSP}}{\# \text{All the pixel number}}$$
**TABLE I. THE COMPARISON OF THE PRECISION OF DETECTION RESULTS**

<table>
<thead>
<tr>
<th>Video sequence</th>
<th>GMM</th>
<th>histogram</th>
<th>Adaptive GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soccer-1</td>
<td>80.02</td>
<td>78.35</td>
<td>90.74</td>
</tr>
<tr>
<td>Soccer-2</td>
<td>95.78</td>
<td>90.89</td>
<td>97.89</td>
</tr>
<tr>
<td>Soccer-3</td>
<td>96.65</td>
<td>91.35</td>
<td>99.02</td>
</tr>
<tr>
<td>badminton</td>
<td>97.37</td>
<td>92.03</td>
<td>98.02</td>
</tr>
<tr>
<td>tennis</td>
<td>95.48</td>
<td>93.32</td>
<td>96.93</td>
</tr>
<tr>
<td>basketball</td>
<td>94.68</td>
<td>90.12</td>
<td>98.56</td>
</tr>
</tbody>
</table>

**TABLE II. THE CLASSIFICATION RESULTS OF FOOTBALL VIDEO SCENARIO (%)**

<table>
<thead>
<tr>
<th>match</th>
<th>Close up</th>
<th>Scene without Playfield(r/p)</th>
<th>Global view(r/p)</th>
<th>Goal area(r/p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>England:</td>
<td>94.8/90</td>
<td>100/100</td>
<td>100/89.9</td>
<td>96.5/91.1</td>
</tr>
<tr>
<td>france:</td>
<td>.84</td>
<td>100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>Portugal:</td>
<td>95.84/9</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>spain:</td>
<td>1.2</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>Greece:</td>
<td>96.8/89</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>ussia:</td>
<td>.02</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>Italy:swe</td>
<td>94.2/89</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>den:</td>
<td>.4</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>Germany:</td>
<td>95.9/92</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>-holland:</td>
<td>.3</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>Denmark:</td>
<td>92.4/93</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
<tr>
<td>-bulgaria:</td>
<td>.6</td>
<td>100/100</td>
<td>98.7/91.9</td>
<td>93.5/89.9</td>
</tr>
</tbody>
</table>

**B. The Classification Experiments of Football Video Scenario**

For the results of playfield detection based on AGMM, the performance test is made for the classification method of football video scenario in the third section. The test data is from 6 games in the 2012 European Championships, including nearly 810 minute’s video programs. Table 2 shows the recall ratio and accuracy of classification results in 6 games and they are respectively defined as:

\[
\text{precision ratio} = \frac{\#\text{The number of correct score scene}}{\#\text{The number of correct score scene} + \#\text{The number of incorrect score scene}}
\]

\[
\text{recall ratio} = \frac{\#\text{The number of correct score scene}}{\#\text{The number of correct score scene} + \#\text{The number of incorrect score scene}}
\]

From experiment results, it can be seen that the classification method of football video scenario proposed in the third section in this paper is effective and it can be easily extended to the video analysis of hockey and other sports.

**IV. CONCLUSION**

At first, this paper proposes a playfield detection algorithm based on AGMM. Its biggest characteristic is to constantly adjust parameters of GMM in the process of system operation according to IEM algorithm. This method brings a lot of advantages. For example, it can perfect model on line so that it adapts the change of footfield color in video stream; in the learning process, it doesn’t have to keep the training sample, which saves the storage space. The experiment shows the effectiveness and feasibility of the playfield detection algorithm based on AGMM. Besides, compared with the traditional method, it can get better results in the recall ratio and accuracy. Secondly, the classification method of football scenario proposed in this paper makes effective use of the distribution of the playfield area in the image. The method is simple and effective and it is easy to spread to video analysis of other sports.

**REFERENCES**


