Accurate Playfield Detection Using Area-of-Coverage

Viet Anh Ngo, Wenxian Yang, Jianfei Cai
Centre for Multimedia and Network Technology, Nanyang Technological University
Email: {ngovietanh, wxyang, asjfcai}@ntu.edu.sg

Abstract—Playfield detection is an important task in sports video content analysis, as it provides the ground for further operations such as object detection, object tracking or semantic event highlight and summarization. Traditional approaches to playfield detection usually rely on the homogeneity cue of the field pixels’ color to separate it from other areas within the image frame. Although encouraging results have been achieved using this methodology, the accuracy of the detection still largely varies depending on the grass pattern of the field. In this paper, we propose a novel approach to detect the playfield region using the concept of Area-of-Coverage. First, the system uses Gaussian Mixture Model (GMM) to obtain a rough segmentation of the playfield area. Based on this initial segmentation, feature points of the field are detected for calibration using homography and then the Area of Coverage (AoC) of the camera view is calculated. Finally, this AoC is back-projected to the image frame to provide a refined detection result. Experiments on real videos show that the proposed system achieves higher accuracy as well as stability compared to existing approaches.

I. INTRODUCTION

Sports video analysis, especially soccer, has received much of the attention due to its popularity as well as potential commercial value. A broad range of applications such as: content retrieval and indexing, semantic event highlight and summarization or objects recognition and tracking make building sports video analysis system a hot research area. In those video analysis systems, playfield (or landmark) detection is usually the very first step to provide the ground for further detection operations. By identifying the whereabouts of the playfield region, noise from unimportant areas could be eliminated and the detection job become much simpler.

While playfield detection is essential to sports video analysis, there are obstacles that make a robust and accurate detection system challenging. The variance of information such as grass color pattern or camera illumination make playfield detection difficult. Moreover, the interference of noise from the audience area also introduces complexity to the process.

There exists two main approaches to playfield detection: non-parametric [1] [2] and parametric approaches [3] [4] [5]. For non-parametric approaches, in [1], the dominant color of the playfield is calculated based on image histogram and the cylindrical distance to this mean color is used to segmented the playfield. In [2], color thresholding and spatial coherence (connected component) are combined in the playfield detection. For parametric approaches, the Gaussian Mixture Model (GMM) is the most widely used technique. Most currently, in [5], Incremental Expected Maximization (IEM) is introduced to adapt the GMM to the variation of the playfield with time. Although encouraging results have been achieved by the existing playfield detection methods, they do not perform very well when the grass or illumination pattern varies greatly between different stadiums. Also, the detected results often contain abundant regions from outside the playfield area.

In this paper, we propose a novel system to improve the accuracy of playfield detection by combining color and homography information. Figure 1 shows the system diagram. It can be seen that the proposed system consists of three main parts: field detection, homography estimation and AoC segmentation. The field detection part is to use color cue to obtain a rough estimation of the playfield area, while the parts of homography estimation and AoC segmentation are to refine the result by using the homography information. The details of each part are elaborated on the subsequent sections.

II. FIELD ESTIMATION

As shown in Figure 1, the field estimation part contains three steps: dominant color region detection, GMM estimation and region-growing refinement. The objective of the dominant color region detection step is to filter out the image pixels that belong to the dominant color and then fit them as inputs to the GMM model. The GMM estimation step then analyzes these pixels and classifies them into corresponding clusters, where the detected playfield often contains many fragments among grass pixels due to noise or mis-classification. The
region-growing refinement step is to eliminate these fragments to provide a cleaner result.

### A. Dominant Color Region Detection

In this step, color histogram of the image is used to analyze the statistical information of the pixels’ color. Note that the image is converted to YCbCr space for processing as this color space is less sensitive to noise than other color spaces (such as RGB or HSV). The dominant color region is determined in the image histogram as the connected bins with the largest or "peak" number of pixels. Detail of the detection procedure is outlined in Algorithm 1.

#### Algorithm 1 Dominant Color Region Detection

**Procedure**

1. Build 2D histogram based on the CbCr of the image pixel values.
2. Determine the main peak $P_1$ of the 2D histogram.
3. Find 4-connected region around $P_1$ with threshold $T_1 = Value(P_1)$ and compute the sum of these connected bins, denoted as $sum_1$.
4. If $sum_1$ is larger than a threshold $T$ (default value $T = 0.5$), goto 7.
5. Determine main peak $P_2$ for the rest of the bins, find 4-connected region around $P_2$ and denote the sum of those bins as $sum_2$.
6. Return the connected region in the histogram corresponding to the larger of $sum_1$ and $sum_2$.
7. The colors whose CbCr values fall in the bins is the selected connected region are consider as the playfield colors.

Figure 2 shows a sample image frame and its detected dominant color region. It can be seen that this region can roughly represent the playfield area; however, there are still a lot of noise and pixels from outside areas wrongly detected in the result.

#### B. GMM Estimation

After the dominant color region has been identified in the image histogram, adaptive GMM is used to filter the pixels within the selected bins into color-related clusters. On the first frame, the batch version of Expectation Maximization (EM) algorithm is used to initialize the GMM parameters. And on subsequent frames the incremental version of EM is used to update the GMM, as proposed in [5].

In our detection scenario, we use three gaussian components for the GMM: two for modeling the field grass and one for modeling the non-grass component within the field region. The reason for this is because we want to clearly distinguish the grass and non-grass elements and at the same time able to handle the illumination variance of the grass pattern. Typically the pixels classified in the two larger gaussian components are detected as field pixels, and those in the gaussian component with the least size are non-field pixels.

Figure 3 shows the result of the GMM clustering process, in which component 1 corresponding to the light grass region, component 2 corresponding to the dark grass region and component 3 corresponding the non-grass region (including the field court lines and noise from the dominant color region detection step).

#### Algorithm 2 Region-Growing Refinement

**Procedure**

1. Input the binary image resulting from playfield pixel detection, where 0 represents a non-playfield pixel and 1 represents a playfield pixel.
2. In the input image, find the connected components of the playfield pixels.
3. For each playfield components, if its size smaller than a threshold, set all the pixels in this component as non-playfield pixels.
4. In the input image, find the connected components of the non-playfield pixels.
5. For each non-playfield components, if its size smaller than a threshold, set all the pixels in this component as playfield pixels.

Figure 4 shows a sample of the GMM estimation and the region-growing refinement results. It can be seen that the results excluded the court lines and a lot of noise around the player regions are removed, compared with the dominant color region detection.

#### III. Homography Estimation

The objective of the homography estimation is to derive the transformation between the camera pixel coordinates and its corresponding playfield position. As shown in Figure 1 the homography estimation consists of two steps: feature point detection and camera calibration.
A. Feature Point Detection

The feature point detection step is to identify a collective set of points belonging to the playfield, whose coordinates are well defined with regard to the standard court model (such as corners, or special points on the court lines). This is done by analyzing the GMM component of the playfield which contains the court line pixels.

First, the system performs Hough line detection on the binary image of the GMM court-line component. The result of this process is a set of detected line segments which are represented in the parameter space as pairs of \((\theta, d)\), where \(\theta\) is the angle between the normal and the horizontal axis and \(d\) is the distance of the line segment to the origin.

To identify the court lines within the detected set, the angle \(\theta\) and distance \(d\) of each line segments are matched against those belong to the court lines detected on the previous frame. Note that the user intervention is required for the first frame to identify the initial court-line segments; and for subsequent frames the angles and distances of those court-line segments are tracked and predicted using Kalman filter. Also, after the homography estimation of each image frame is completed, the list of court-line segments is updated by back projecting the court-line segments from the field model to the image frame and recording those lying within the estimated playfield region. Finally, the feature points are detected as the intersections of the identified court-line segments.

Figure 5 shows a sample of the Hough line detection and feature point detection result.

![Feature Point Detection](image)

B. Camera Calibration

After the feature point detection step, the system maintains a list of corner points whose 3D coordinates on the playfield is well-defined with respect to the court model. The camera calibration is the next step to calculate the transformation, basically a \(3 \times 3\) matrix, between the image frame coordinates and the playfield coordinates. A direct calculation of the camera matrix can be done with four pairs of image frame and playfield coordinates. Details on this process are described in our previous work [6].

IV. AoC SEGMENTATION

The AoC segmentation part is to obtain an accurate segmentation of the playfield. There are two steps in the AoC segmentation: AoC calculation and AoC back projection.

A. AoC Calculation

The Area-of-Coverage (AoC) of a camera view is defined as the area on the playfield which the camera captures. As described in Algorithm 3, there are five steps in calculating the AoC: coordinates projection, polygon type checking, intersection checking, vertices checking and finally AoC generation.

In the first step, the corners of the image frame are projected onto the playfield. Depending on the relative positions and the viewing angle of the cameras, there are two cases for view projection, it can either be a convex polygon or a self-intersecting polygon. Fig. 6 shows the illustration of the two cases, in which \((P_{i1}, P_{i2}, P_{i3}, P_{i4})\) denote the image corner points in clockwise order for view \(V_i\), and \((P'_{i1}, P'_{i2}, P'_{i3}, P'_{i4})\) denote their projection onto the playfield.

In the second step, the polygon check process distinguishes the type of the projectedpolygon. In the case that \((P'_{i1}, P'_{i2}, P'_{i3}, P'_{i4})\) represent a convex polygon, \(A_i\) is the overlapping area of this polygon and the playground. Otherwise, \(A_i\) is the overlapping area of the extension of the edges with the playground. The coverage areas are represented by the gray regions as shown in Fig. 6 for both cases.

In the third step, intersections points of the projected polygon \((P'_{i1}, P'_{i2}, P'_{i3}, P'_{i4})\) with the boundary of the playground \((P_{T1}, P_{T2}, P_{T3}, P_{T4})\) is computed. In Fig. 6 the intersection points are identified as \(\{Q_1, Q_2\}\) for both cases.

The vertices checking is the fourth step of the AoC Calculation process. In this step, two checkings are performed. First, the module searches for vertices of the projected polygon that lie in the playground, e.g., \(\{P'_{i2}, P'_{i3}, P'_{i4}\}\) for Case 1 and \(\{P'_{i3}, P'_{i4}\}\) for Case 2. Note that \(P'_{i2}\) in Case 2 is excluded by the polygon check. Second, it searches for vertices of the playground that lie in the projected polygon, e.g., \(\emptyset\) for Case 1 and \(P_{T3}\) for Case 2.

![Algorithm 3 Area-of-Coverage Calculation](image)
After the vertices checking step, the vertices of the AoC are defined and the final step is to rearrange those vertices so that every pair of consecutive vertices forms an edge of the AoC. For example, the rearranged vertices are \( \{Q_1, P_{T2}', P_{T3}', Q_2\} \) for Case 1 and \( \{Q_1, P_{T3}', Q_2, P_{T4}', P_{T4}'\} \) for Case 2. The Quickhull algorithm [7] is used to rearrange the points.

### B. AoC Back Projection

The AoC back projection is the final step of our segmentation system. In this step, the AoC generated from the previous part is back projected to the image frame to obtain the segmentation of the playfield. Figure 7 shows the computed AoC as well as the final refined playfield segmentation.

![AoC Calculation](image1)
![AoC BackProjection](image2)

Fig. 7. AoC Segmentation

### V. EXPERIMENTAL RESULT

![Input Images](image3)
![Results of Ekin [1]](image4)
![Results of Liu [5]](image5)
![Our Results](image6)

Fig. 8. Comparison of Playfield Detection Approaches

To verify the accuracy of our proposed system, experiments were carried out with video clips from three different soccer matches. We compare our method with Ekin [1] (dominant color segmentation) and Liu [5] (adaptive GMM segmentation) which are the representative non-parametric and parametric playfield segmentation methods respectively.

Figure 8 shows the playfield detection results. From the visual inspection, it can be seen that our system detects the playfield area more accurately compared to Ekin’s and Liu’s approaches. This is due to the ability of the system to remove abundant outside-field grass areas based on the homography information. Note that, the dominant color segmentation achieves good results when the grass colors of the field are homogenous, but fails to deal with the variance when shades are introduced to the field. The GMM handles this issue better by using multiple gaussians to model the field grass patterns.

Table I shows the quantitative measure on the detection accuracy, which is calculated based on the ratio of the total number of detected pixels minus the number of wrongly detected pixels over the number of the ground-truth playfield pixels (manually annotated). It can be seen that our system improves the accuracy significantly.

### VI. CONCLUSION

In this paper, we proposed a novel system to detect playfield region from soccer videos. The system performs two-stage operations: first it uses GMM segmentation to obtain a rough estimation of the playfield and subsequently applies the concept of Area-of-Coverage to extract the refined playfield region. Experiments carried out with videos from three different matches have demonstrated that the system can achieve very accurate detection results. The proposed approach can be easily applied to other sports such as basket ball, table tennis or extended to other areas such as video surveillance or 3D image registration.

### REFERENCES


