Unsupervised high-quality soccer field segmentation

Daniel Quilón, Raúl Mohedano, Carlos Cuevas and Narciso García

[1]Abstract—**Field segmentation is a fundamental step in many soccer applications. However, despite its importance, the existing segmentation algorithms are not able to provide successful results in complex scenarios. Moreover, they require the manual selection of several parameters, hindering their usability. Here, an unsupervised field segmentation strategy based on the estimation of the probability density function of the green chromacity of the image is proposed. Results show its ability to provide high-quality results in a wide variety of scenarios.**

Keywords—soccer; segmentation; chromaticity; automatic; unsupervised; field

I. INTRODUCTION

Soccer popularity has motivated the proliferation of many computer vision applications for game analysis [1], which include field segmentation as a fundamental starting step [2]. Most segmentation strategies assume green as the dominant color in the field and analyze the H component of the Hue-Saturation-Value (HSV) space. Some algorithms [3] apply predefined thresholds around the ideal H value for green, resulting in segmentation inaccuracies that depend on the field tone and the image color balance. Other works [4][5] identify the dominant playfield color from the analysis of the H image histogram and infer thresholds to isolate it. The fact that H is cyclically defined complicates the selection of the thresholds and the discrimination of statistical color modes, making these approaches prone to failure in fields with alternated dark and light grass bands or strongly shadowed areas. In addition, these methods are dependent on the histogram bin size.

Instead, a real-time high-quality soccer field segmentation strategy based on the analysis of the probability density function (PDF) of the green chromaticity of the image is presented. This strategy completely avoids manual parameter selection (*e.g.* bin-size of histogram-based algorithms, idealized green value in the chosen color space). Moreover, since no dominant color mode in the field is assumed, segmentation results are successful in stadiums with multiple green colors and in presence of shadows on the playfield.

II. PROPOSED METHOD

This proposal extracts the main statistics of the green chromaticity channel of the image by approximating its empirical probability distribution with a Gaussian Mixture
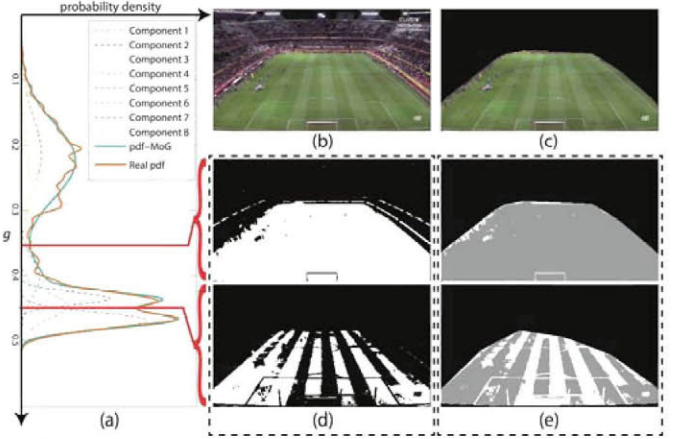

Fig. 1: (a) Real PDF (orange curve), estimated GMM PDF (blue curve) and candidate thresholds (red lines). (b) Original image. (c) Final segmentation. (d) Masks Σ corresponding to the data above the thresholds automatically selected in the estimated PDF. (e) Masks Γ after applying morphological operations and the convex hull (shadowed parts represent intersection $\Sigma \cap \Gamma$).

Model (GMM) estimated using the EM (Expectation-Maximization) algorithm. This approximation is used to infer candidate thresholds for segmentation, whose convenience is evaluated by analyzing the shape and density of the resulting segmentation. These steps are illustrated in Fig. 1.

A. Green chromacity as discriminant feature

The choice of the green chromaticity $g = G/(R+G+B)$, $0 \leq g \leq 1$, shows two main advantages. First, it is highly invariant to illumination changes, which is essential to effectively analyze field color statistics where illumination can range from natural to artificial and where strong shadows can be present. Second, it provides a univocal criterion for comparing colors in terms of their closeness to green.

B. EM-based analysis for candidate threshold extraction

Unlike most works, analyzing color histograms to infer the features of the color mode corresponding to the green field, this proposal estimates the underlying continuous PDF of the g channel from the observed image samples. This PDF is expected to show one mode (or several) at high values, corresponding to the field itself, and several spurious modes centered at lower g values, corresponding to the rest of structures and areas of the image. Such PDF profile (Fig. 1(a)) can be accurately approximated by a mixture of an a priori unknown number of Gaussians with the appropriate parameters: those parameters are estimated from the image data using EM. A GMM approximation allows to easily calculate a limited number of local minima of the PDF: these minima indicate the potential boundaries between the different

modes of the color PDF and are therefore considered as candidate threshold values for the final field segmentation (the selection of the right final threshold is addressed in the next subsection).

This work assumes a fixed number of 8 Gaussians for such approximation. This number is a good tradeoff between PDF representation accuracy (using more Gaussians than necessary has no detrimental effect on the resulting approximation) and threshold extraction complexity. Gaussians are initialized uniformly centered in the range 0-1 and with standard deviation 1/8. Such initialization ensures that EM always captures the features of the green field modes in detail.

C. Spatial consistency and final threshold selection

The candidate thresholds need to be evaluated to decide which one delimits best the color of the field. To that end, a simple color criterion is first used to filter candidates from its color: they must lie in the area which can be considered green, that is, $G>(R+B)$ (or alternatively $g>1/3$).

Second, and more importantly, the spatial consistency of the segmentation derived from each retained threshold is analyzed. To that end, morphological operations are applied so as to remove the fragmented noise of the segmentation, and the convex hull of the result is extracted (since the view of the soccer field from any unobstructed view must correspond to a convex polygon). The fit of such convex hull Γ and the raw segmentation mask Σ is evaluated through the evaluation of the density $d = A(\Sigma \cap \Gamma)/A(\Gamma)$, where $A(\cdot)$ represents the area (in pixels) of a given binary mask. The hull of the densest segmentation provides the final field segmentation.

Fig. 1 depicts the complete proposal for a specific test field composed of bands in different tones of green. Fig. 1(a) shows how the multimodal PDF of the g channel of Fig. 1(b) is accurately approximated by our 8 component GMM, and how two candidate thresholds are extracted. The consistency analysis of the results of both thresholds (Fig. 1(d) and (e)) indicates that the lowest is the best threshold, providing the accurate final segmentation result shown in Fig. 1(c).

III. RESULTS

The proposed strategy has been tested on more than 30 images taken from multiple soccer stadiums and from different view-points. These images contain many complex situations such as fields with two grass colors or fields partially occluded by shadows. The obtained results have been compared with two outstanding segmentation methods: the first one [3] applies two fixed thresholds around the green on the H histogram, while the second one [5] identifies the dominant mode of the HSV histograms and applies two thresholds to separate them from the rest of data. Fig. 2 shows some representative results. It can be observed that the strategy in [3] fails in images with fields composed by different tones of green (central image in Fig. 2), since it is based on fixed thresholds. Additionally, it wrongly selects green regions not belonging to the field (right image of the figure). On the other hand, since the algorithm in [5] is based on a unique peak histogram, it fails in fields partially covered by shadows (left image in Fig. 2) and also in images where the
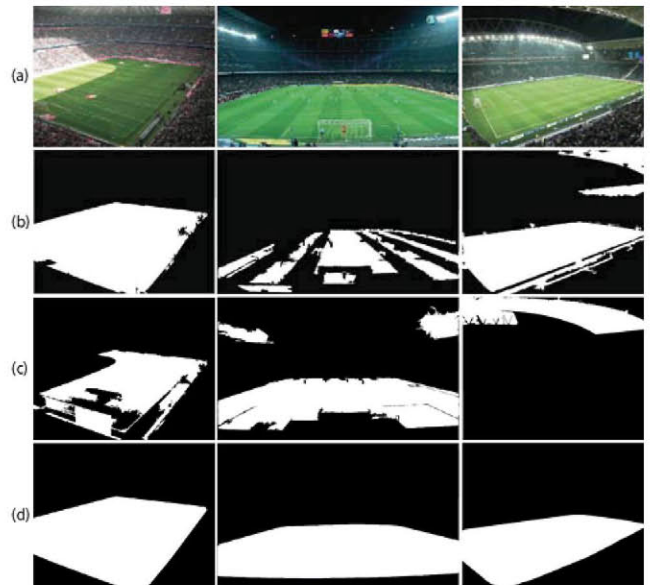

Fig. 2. (a) Original images. (b) Results with the strategy in [3]. (c) Results with the strategy in [5]. (d) Results with the proposed strategy.

most representative region is not the playfield (for example, in the right image of Fig. 2, where the most representative histogram peak corresponds to the sky). Finally, the last row of images in Fig. 2 shows that the proposed strategy provides successful results in all the conditions, improving the quality of previous approaches.

IV. CONCLUSIONS

A fully autonomous method for accurately segmenting soccer fields from shots taking from an arbitrary point of view is proposed. It combines the choice of green chromaticity (as the discriminant feature for distinguishing field from the rest of image elements) with an automatic selection of the potential threshold for the segmentation. These potential thresholds are finally evaluated following geometric considerations, which leads to the selection of the best threshold and its associated mask as the final result. The proposed segmentation strategy has been tested on a battery of test images composed of challenging examples Obtained results show the satisfactory performance of the proposal and its outperformance of the comparable works in the literature.

REFERENCES

[1] T. DOrazio and M. Leo, "A review of vision-based systems for soccer video analysis," Pattern Recognition, vol. 43, no. 8, pp. 2911–2926, 2010.

[2] J. Bu, S. Lao, and L. Bai, "Automatic line mark recognition and its application in camera calibration in soccer video," IEEE Int. Conf. Multimedia and Expo, pp. 1-6, 2011.

[3] M. Hoernig, M. Herrmann, and B. Radig, "Real time soccer field analysis from monocular TV video data," Int. Conf. Pattern Recognition and Image Analysis, vol. 2, pp. 567-570, 2013.

[4] A. Ekin, A. M. Tekalp, and R. Mehrotra, "Automatic soccer video analysis and summarization," IEEE Trans. Image Processing, vol. 12, no. 7, pp. 796-807, 2003.

[5] V. Ngo and J. Cai, "Converting 2D soccer video to 3D cartoon," IEEE Int. Conf. Control, Automation, Robotics and Vision, pp. 103-107, 2008.