



OPEN

## A fully automatic method for segmentation of soccer playing fields

Carlos Cuevas<sup>1,2</sup>, Daniel Berjón<sup>1,2</sup> & Narciso García<sup>1,2</sup>

This paper proposes a strategy to segment the playing field in soccer images, suitable for integration in many soccer image analysis applications. The combination of a green chromaticity-based analysis and an analysis of the chromatic distortion using full-color information, both at the pixel-level, allows segmenting the green areas of the images. Then, a fully automatic post-processing block at the region-level discards the green areas that do not belong to the playing field. The strategy has been evaluated with hundreds of annotated images from matches in several stadiums with different grass shades and light conditions. The results obtained have been of great quality in all the images, even in those with the most complex lighting conditions (e.g., high contrast between sunlit and shadowed areas). In addition, these results have improved those obtained with leading state-of-the-art playing field segmentation strategies.

Association football, more commonly known as football or soccer, is the most popular sport worldwide<sup>1,2</sup>. Since its foundation, in 1863, it has spread throughout the planet, reaching almost all countries<sup>3</sup>. It is the sport with the largest television audience with nearly 4 billion followers<sup>4</sup>, it has millions of practitioners<sup>5</sup>, and it is also the most studied sport<sup>6</sup>.

As a consequence of this great popularity and thanks to the technological advances produced in the last decade, the demand for artificial vision applications to automatically analyze soccer matches has grown enormously in recent years<sup>7</sup>. On the consumer end, spectators demand applications capable of enriching the content of live broadcasts<sup>8</sup>. Meanwhile on the professional side, clubs and players request applications aimed at a better understanding of the game, studying team tactics, or creating training sessions to improve player performance<sup>9–11</sup>. Additionally, the Video Assistant Referee (VAR), introduced in 2018 into the Laws of the Game of the International Football Association Board (IFAB) to help referees in reviewing decisions by means of video footage, has given rise to a notable increase in the technology used in the stadiums<sup>12</sup>. Besides, during recent years there has been a surge of interest in applications focused on accurately predicting soccer games<sup>4</sup>.

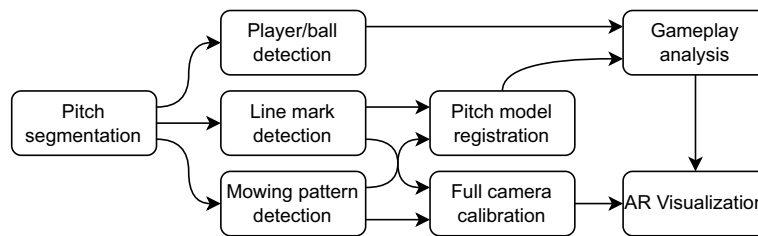
**Motivation.** Many soccer-related applications require detecting different kinds of elements that appear on the playing field: players, ball, line marks, and grass bands. The detection of the players<sup>13</sup> and/or the ball<sup>14</sup> is used in applications that try to identify relevant events that take place during the matches (e.g., goal scoring, shots on goal, or corner kicks), or in applications that perform high-level analyses about the matches (e.g., possession statistics, offside detection, or team tactics). Regarding the detection of the line marks and the grass bands<sup>15</sup>, it is used in applications that require to register the images in a model of the playing field and/or estimate the pose of the cameras for high-level purposes, such as including augmented reality in images or obtaining some player statistics (e.g., distance traveled, speed, etc.).

The pipeline in Fig. 1 illustrates how the proposed segmentation strategy could be used to complement the analysis strategies described in the previous paragraph. The segmentation of the playing field allows discarding areas of the images that are irrelevant for such strategies (e.g., the stands, the billboards or the sky)<sup>7</sup>. Therefore, included as a first stage in any of them, it has the potential to prevent false detections outside the playing field, thus improving their results.

It is important to note that all of these applications are intended for use on video streams. Consequently, the segmentation strategies that are proposed to complement them must be fully automatic.

Typical segmentation strategies assume green as the dominant color in the playing field and try to isolate such color by analyzing the histogram of the hue component in the HSV (hue, saturation, value) color space or

<sup>1</sup>Grupo de Tratamiento de Imágenes (GTI), Information Processing and Telecommunications Center (IPTC), Universidad Politécnica de Madrid (UPM), 28040 Madrid, Spain. <sup>2</sup>These authors contributed equally: Carlos Cuevas, Daniel Berjón, Narciso García. ✉email: carlos.cuevas@upm.es



**Figure 1.** Possible system pipeline including the proposed method as a first step.

by establishing simple rules in the RGB (red, green, blue) color space. Although these methods are capable of providing high-quality results on many images, they fail in images containing certain colors in the stands or in the billboards, and also in images featuring heavy contrast between sunlit and shadowed areas. Furthermore, all of these strategies include region-level post-processing steps that depend on thresholds which, in turn, depend on the characteristics of the images (e.g., the zoom and the spatial resolution). This scene-dependent thresholds make it difficult for strategies to be fully automatic.

**Contribution.** We propose a new playing field segmentation strategy that is capable of providing very high quality results in images taken from different points of view and in complex situations (e.g., lawn striping resulting from bending grass in different directions, or images with a high contrast between sunlit and shadowed areas). Two public databases that contain over 450 manually annotated images from 25 matches in several stadiums with different shades of grass and light conditions have been used to assess the quality of our strategy. This evaluation has shown that our results are significantly better than those of most outstanding playing field segmentation methods.

The specific contributions of this paper are:

1. An analysis based on normalized green is combined with another analysis based on chromatic distortion in the RGB color space to segment, at the pixel level, the green areas of images. The combination of these two analyses allows obtaining high-quality results regardless of the shades of green of the grass bands and the lighting of the stadiums.
2. A post-processing stage at the region level that allows discarding areas of the image that do not belong to the playing field but that show a color similar to that of grass. Unlike the post-processing included in other methods, the one we propose does not depend on scene-dependent thresholds, which greatly simplifies its application regardless of the characteristics of the images (e.g., their resolution) and their content (e.g., the size of players).

**Organization.** The paper is organized as follows. First, “[Related work](#)” section reviews the main image segmentation strategies in the literature, paying special attention to those used in soccer. In “[System overview](#)” section an overview of the proposed strategy is provided, which includes three main parts: a pre-processing stage (“[Pre-processing](#)” section), a pixel-level analysis (“[Pixel-level analysis](#)” section), and a region-level analysis (“[Post-processing](#)” section). Experiment results are reported in “[Results](#)” section and, finally, “[Conclusions](#)” section presents the conclusions of the paper.

## Related work

This section summarizes the main types of strategies that can be found in the literature to segment images. First, general purpose strategies are mentioned. Secondly, those used in the scope of soccer are described. Finally, we focus on the strategies that, like the proposed one, have the aim of segmenting the playing field in soccer images.

**Image segmentation.** The goal of image segmentation is to segment an image into consistent objects or regions of interest (ROI)<sup>16</sup>. It is a pre-processing stage in many image-based applications like biometric identification, medical imaging, object detection and classification, and pattern recognition<sup>17</sup>.

In the literature, a wide range of image segmentation algorithms have been developed, starting with the simplest ones like thresholding<sup>18</sup>, region-growing<sup>19</sup>, k-means clustering<sup>20</sup>, or watershed methods<sup>21</sup>, and progressing to more complex ones like active contours<sup>22</sup>, graph cuts<sup>23</sup>, conditional and Markov Random Fields<sup>24</sup>, or deep learning methods<sup>25</sup>.

In addition, there are many different applications of video analysis that require segmenting the moving objects in the scene (called foreground), separating them from the static data (called background)<sup>26</sup>, such as video surveillance<sup>27</sup>, Human-Machine Interaction (HMI)<sup>28</sup>, or object tracking<sup>29</sup>. To achieve this, a variety of moving object segmentation algorithms have been presented in the literature, ranging from unsupervised techniques (e.g., background subtraction) to semi-supervised (e.g., spatio-temporal graphs) or interactive (e.g., graph partitioning models) techniques<sup>30</sup>.

**Segmentation in soccer.** In the case of image analysis of soccer matches, we can find in the literature algorithms designed to carry out the following three types of segmentation:

- Playing field segmentation: They are focused on segmenting the playing field and are typically based on modeling the greenness of the grass (see “[Soccer playing field segmentation](#)” section).
- Line mark segmentation: They focus on the segmentation of the lines that delimit the playing field, since these lines are required to register the images in a model of the playing field and/or estimate the pose of the cameras, which in turn allows addressing different high-level tasks, such as including virtual elements in the images<sup>31</sup>.
- Players and/or ball segmentation: They try to segment the moving objects on the playing field (i.e., players and ball). These objects are typically taken as input data by different types of high-level analysis applications (e.g., analysis of the ball possession or analysis of team tactics)<sup>7</sup>.

**Soccer playing field segmentation.** Although it is technically possible to segment the playing field using many of the segmentation algorithms mentioned above (e.g., watershed, region-growing, or deep learning), due to their simplicity and because the green color of the grass is the most significant feature of the playing field, the most outstanding strategies in the literature use thresholding algorithms that are applied to histograms of different color components. These strategies can be classified into the following three types:

- Hue-based methods: Those based on the hue component of the HSV color space.
- RGB-based methods: Those based on the analysis of the RGB color space.
- *g*-based methods: Those that analyze the green chromaticity (*g*) of the image.

Hue-based methods analyze the histogram of the hue component in search of the dominant mode, which is presumed to correspond to the green color of the playing field. Then, they separate this mode from the rest of the histogram data. In some works<sup>32,33</sup>, fixed thresholds around the dominant mode in the hue histogram are used. In<sup>34</sup>, the thresholds are selected by considering a fixed width around the dominant mode of the hue histogram. In<sup>35,36</sup>, training sets of images are used to determine the average position and width of the dominant peak in the hue histograms. Although hue-based methods are able of providing successful results in simple cases, automatically selecting the appropriate thresholds can be very complex in some images, as hue histograms are often very multimodal. Moreover, these methods are very sensitive to the size of the histogram bins and fail in images where the dominant mode does not correspond to the color of the playing field, but to the sky or the stands<sup>31</sup>.

RGB-based methods, assuming that the color of the playing field is primarily green, try to segment the playing field by determining what pixels satisfy certain relationships between the red (*R*), green (*G*), and blue (*B*) components of the RGB color space. In<sup>37–39</sup>, it is assumed that the pixels of the playing field must comply with the rule  $G > R > B$ . Alternatively, in<sup>40</sup> it is assumed that the playing field pixels are those satisfying  $G > R$  and  $G > B$ . In<sup>41</sup>, 2-dimensional histograms in a normalized RG color space are used to deal with grass color variations, where the peaks determine the playing field data. In<sup>42,43</sup>, RGB 3-dimensional histogram techniques are proposed. The main advantage of these methods is their simplicity. However, they fail in images with areas saturated by sunlight or with strong shadows.

*g*-based methods have been proposed recently as an alternative to the two types of previous methods. They start from the fact that *g* is highly invariant to changes in illumination and, additionally, provides a very simple criterion for classifying colors in terms of their closeness to green (a purer green as *g* increases). In<sup>44</sup>, the probability density function (pdf) of *g* is approximated using a Gaussian Mixture Model (GMM). Then, the playing field mask is obtained by selecting the modes of the pdf that denote data with mainly green information and applying a region based analysis. In<sup>45</sup>, a similar analysis of the pdf of *g* is applied to perform a segmentation that is robust to both floodlighting and natural lighting. Although these methods yield good-quality results in playing fields with different shades of grass and variations in lighting, they generally include areas of the billboards or stands with colors such as cyan or yellow, since the criterion they use to segment does not allow separating green from those colors.

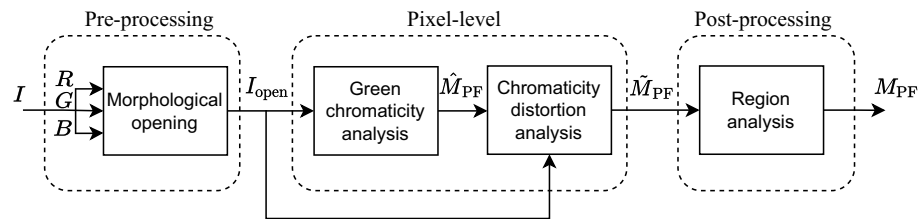
It is important to point out that all these strategies are based on the analysis of the color characteristics of the images at the pixel level. Consequently, they require applying different kinds of region-based post-processing (e.g., morphological operations, convex hulling, etc.) to: i) discard areas of the image that do not belong to the playing field but that show a color similar to that of grass and; ii) deliver a single connected region. These post-processing stages include several thresholds that are highly dependent on the size and shape of the non-green objects that appear on the playing field (the white lines and the players), which in turn are dependent on the size of the images, the location of the camera, and the zoom used. This ad-hoc tailoring significantly reduces the usability of all these strategies.

## System overview

The proposed strategy comprises a pre-processing block, two pixel-level processing steps, and one region-level post-processing step, as shown in Fig. 2.

In the pre-processing block (“[Pre-processing](#)” section) an RGB image  $I_{\text{open}}$ , in which the white line marks in the playing field are integrated in the grass, is derived from the original image,  $I$ .

Then, a pixel-level segmentation that is based on the analysis of the green chromaticity of  $I_{\text{open}}$  is performed (“[Pixel-level analysis](#)” section), which yields an initial binary mask,  $\hat{M}_{\text{PF}}$ , indicating the pixels in which the green component is the dominant one. RGB-based chromaticity distortion analysis is then performed to remove pixels that, despite having a significant contribution from the green channel, are not part of the playing field (“[Pixel-level analysis](#)” section). As a result of this second analysis, the binary mask  $\tilde{M}_{\text{PF}}$  is obtained.



**Figure 2.** Block diagram of the proposed strategy.

The final mask of the playing field,  $M_{PF}$ , is obtained by applying a region-level post-processing module that discards green regions that, despite having a color compatible with the grass of a soccer playing field, are not part of it (“Post-processing” section).

### Pre-processing

In many images, the presence of white line marks results in the playing field being represented by multiple connected regions after segmentation at the pixel level. Since the number and characteristics of these regions are unknown, complex threshold-based rules are typically required at the region-level post-processing stage. The purpose of this pre-processing is to prepare the original image in such a way that after pixel-level segmentation the playing field is represented by a single connected region, thus making much easier the post-processing at the region level.

Let  $I$  be an original RGB image of a soccer match (see Fig. 3.a). To integrate the white line marks in the playing field (i.e., so that the lines adopt a color similar to that of the grass that surrounds them), a morphological grayscale opening operation is applied to each channel to obtain a new RGB image,  $I_{open}$  (Fig. 3.b). For this, a circular flat structuring element is used, whose diameter  $d_e$  must be slightly greater than the width of the line marks. If this diameter is too small, the lines will not blend well enough into the playing field. On the other hand, if it is too large, some elements of the playing field perimeter will be wrongly integrated into the grass, which will reduce the accuracy of the results. Since the thickness of the white line marks is proportional to the size of the images, we have decided to use a diameter proportional to the size of the diagonal of the image:

$$d_e = \frac{\alpha_e}{100} \sqrt{H^2 + W^2}, \quad (1)$$

where  $H$  and  $W$  are, respectively, the height and width of the image in pixels. The choice of the appropriate value of  $\alpha_e$  and its influence on the quality of the proposed strategy are discussed in the “Results” section.

### Pixel-level analysis

**Green chromaticity analysis.** This analysis makes it possible to identify, at the pixel level, the areas of the image with a greater contribution of green than other primary colors. In addition, it allows these areas to be separated based on the shade of green that characterizes them. To do this, the normalized green component of  $I_{open}$  is analyzed, which is obtained as:

$$g(r, c) = \frac{G_{open}(r, c)}{R_{open}(r, c) + G_{open}(r, c) + B_{open}(r, c)}, \quad (2)$$

where  $R_{open}$ ,  $G_{open}$ , and  $B_{open}$  are, respectively, the red, green, and blue components of  $I_{open}$ , and  $(r, c)$  are the coordinates of the pixels (row and column, respectively).

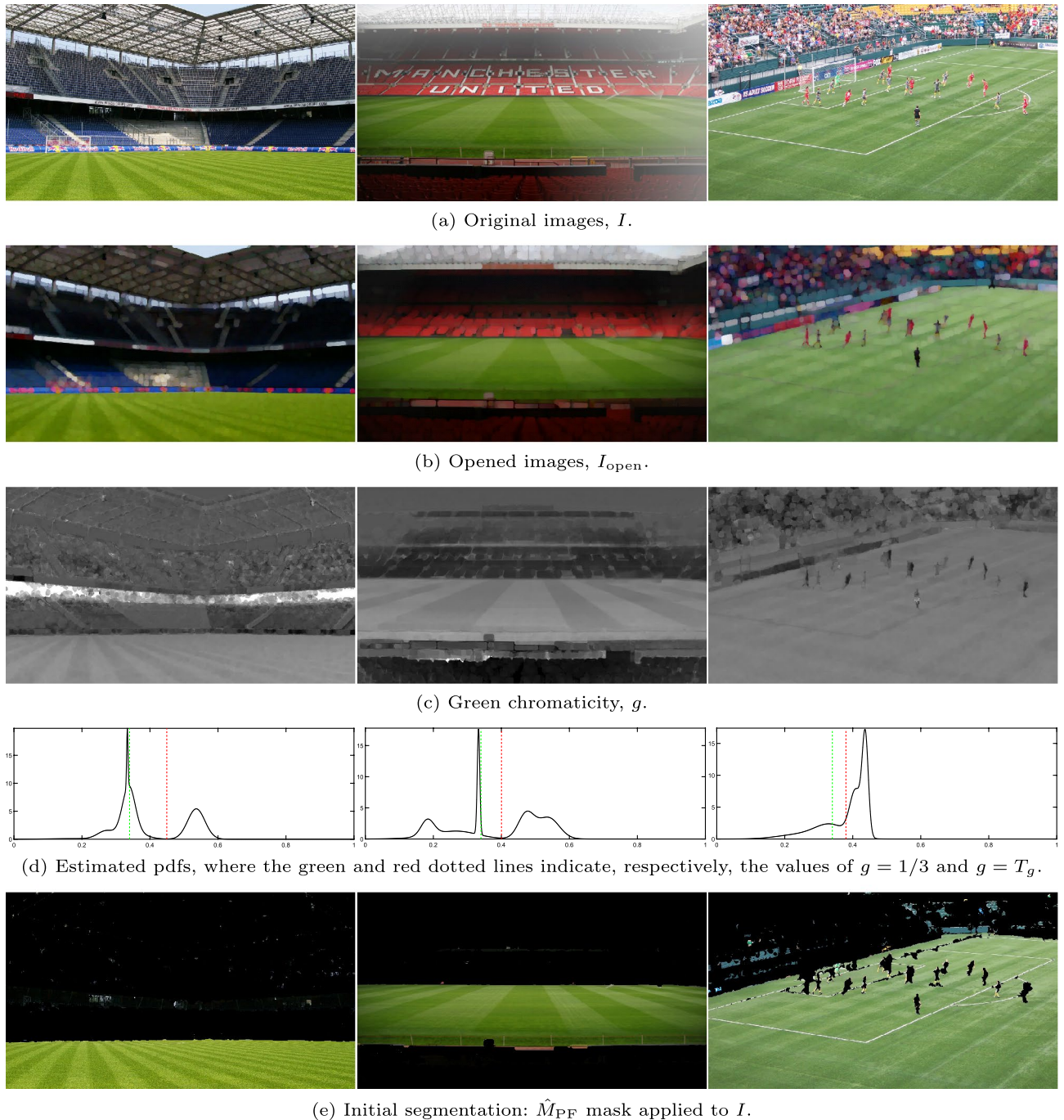
The pdf of  $g$  is expected to exhibit one (or several) modes corresponding to the different shades of green at high values, and several spurious modes corresponding to other colors at lower values<sup>44</sup>. To identify the different groups of pixels with different shades of green, first, the pdf of  $g$  is approximated using the Expectation-Maximization (E-M) algorithm from a different number of Gaussian distributions (from 1 to  $N_G$ ) initialized with equal weights and with means uniformly distributed over the data range. Then, the Akaike information criterion (AIC)<sup>49</sup> is used to select the set of Gaussians that provide the best fitting. The experiments carried out have shown that  $N_G = 6$  is enough to obtain good fits in all the images and that, in addition, the quality obtained hardly changes if other values close to the selected one are used. The results section analyzes in detail the influence of this parameter on the strategy.

Assuming that green is the dominant color of the playing field (i.e.,  $G_{open}(r, c) > R_{open}(r, c)$  and  $G_{open}(r, c) > B_{open}(r, c)$ ), the pixels that constitute it must meet  $g > 1/3$ . Taking this into account, an initial mask of the playing field,  $\tilde{M}_{PF}$ , can be easily obtained by selecting the pixels with values of  $g$  greater than the threshold

$$T_g = \max\left(\frac{1}{3}, m_0\right), \quad (3)$$

where  $m_0$  is the first local minimum of the pdf that is below the first of the peaks above  $g = 1/3$ . In this way, only the data represented by modes of the pdf whose peaks are located above  $g = 1/3$  are selected.

Figure 3 shows, for some representative images, their green chromaticity components (Fig. 3c), the pdf resulting from the adjustment with the E-M algorithm (Fig. 3d) and the images segmented using the mask



**Figure 3.** Results obtained on three images after applying the pre-processing and the analysis at the pixel level. Original images from<sup>46,47</sup>, and<sup>48</sup>.

$\hat{M}_{\text{PF}}$  (Fig. 3e). In the case of the first two examples (images in the left and central columns), it can be seen that the result of the segmentation is of high quality, having correctly identified the green areas of the playing field. However, in these results it can be seen that some small regions in the stands have also been segmented. These regions will be removed by applying the post-processing described in “Post-processing” section. Regarding the third of the examples in the figure (images in the right column), the playing field has also been segmented correctly. However, much of the billboards also appear incorrectly segmented as part of the grass. This is because the green chromaticity of some colors like yellow or cyan is also greater than  $1/3$ . Therefore, green chromaticity analysis cannot discriminate between green and these other colors. To deal with this problem, the chromatic distortion analysis described below is applied.

**Chromatic distortion analysis.** In the pdf obtained in the previous analysis, each of the modes above the selected threshold  $T_g$  represents data with a similar level of green. However, as just discussed, there are some

pixels that do not have the color of grass and yet are also represented by these modes (e.g., those with cyan or yellow colors). By analyzing the chromatic distortion between the RGB value of the data of each mode and the average RGB value of said data it is possible to identify and discard pixels that, having values of  $g$  greater than  $1/3$ , are nevertheless of a color very different from the green of the grass.

This analysis is carried out as follows:

1. Within the envelope of the set of Gaussian distributions used to estimate the pdf of  $g$ , the set of  $N_L$  local minima above  $T_g$  are selected,  $L = \{m_g\}_{i=1}^{N_L}$ .
2.  $N_v = N_L + 2$  sets of RGB data are obtained by selecting the RGB values of the pixels of  $I_{open}$  with values of  $g$  being between:
  - $T_g$  and the first element in  $L$ .
  - Consecutive pairs of local minima in  $L$ .
  - The last local minimum in  $L$  and 1.
3. For each of these sets, each containing RGB data with a similar green chromaticity, the chromatic distortion is analyzed to discard colors that, although similar in their green chromaticity, can be readily told apart using the full color information. Let  $\mathbf{v} \in \mathbb{R}^3$  be the vector with components R, G and B, representing the reference color resulting from averaging all RGB colors in the data set. All the vectors of the form  $\alpha\mathbf{v}$  represent the same color with varying luminance proportional to their magnitude, while different colors are represented by vectors running in different directions: the greater the angle between two such vectors, the more dissimilar they are. Let  $\mathbf{u} \in \mathbb{R}^3$  be the vector corresponding to the RGB color of a pixel we want to compare to the reference color  $\mathbf{v}$ . We can uniquely write  $\mathbf{u}$  as the sum of two orthogonal components:  $\mathbf{u}_v = (\mathbf{u} \cdot \mathbf{v})\mathbf{v}/\|\mathbf{v}\|^2$ , parallel to  $\mathbf{v}$ , and  $\mathbf{u}_{v^\perp} = \mathbf{u} - \mathbf{u}_v$ , perpendicular to  $\mathbf{v}$ . Thus, we can define the tangent of the angle between  $\mathbf{u}$  and  $\mathbf{v}$  as

$$cd = \frac{\|\mathbf{u}_{v^\perp}\|}{\|\mathbf{u}_v\|} \tag{4}$$

The greater  $cd$ , the more dissimilar  $\mathbf{u}$  and  $\mathbf{v}$ , discounting the effect of lightness due to illumination changes.

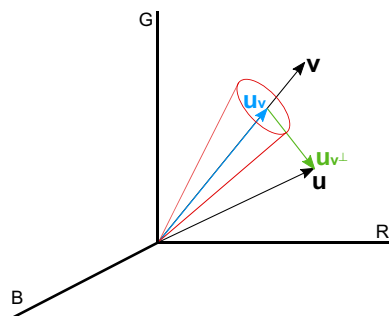
4. A new mask of the playing field is obtained as:

$$\tilde{M}_{PF}(r, c) = \begin{cases} 1, & \hat{M}_{PF}(r, c) = 1 \wedge cd(r, c) < T_c \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

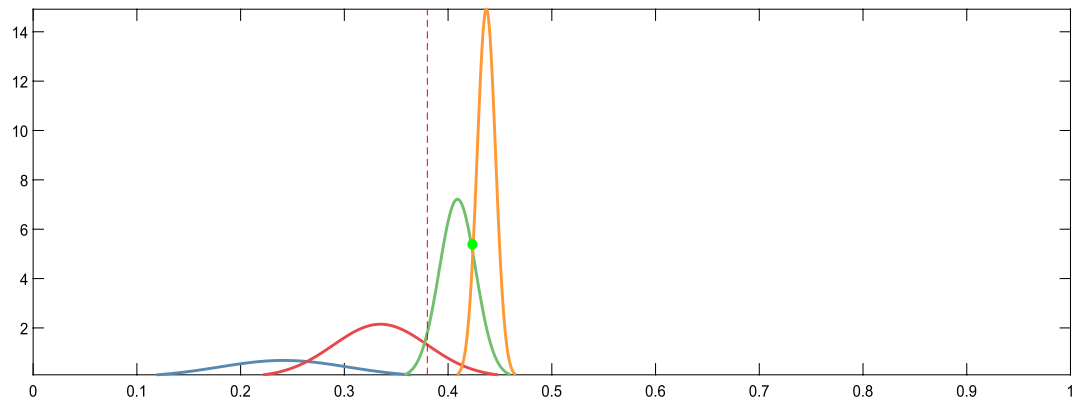
where  $T_c$  is a threshold that effectively restricts acceptable colors to those inside a cone whose axis is the reference color, as illustrated in Fig. 4. The range of adequate values for this threshold is analyzed in the “Results” section.

In the example of Fig. 5, which corresponds to the third case illustrated in Fig. 3 (image in which the chromaticity analysis of green is not sufficient to correctly segment the playing field), the pdf of  $g$  (Fig. 5a) has been obtained by mixing 4 Gaussian distributions, two of them representing modes above  $T_g$ . The local minima separating these two modes has been depicted with a green dot and the dotted red line indicate the position of  $T_g$ . The images in Fig. 5b,c correspond, respectively, to the RGB data and their ratios in each mode. It can be seen that billboards have given rise to the highest  $cd$  values. Consequently, as can be seen in Fig. 5d, where the result of the segmentation is compared before and after the chromatic distortion analysis, billboards have been removed by applying Eq. (5).

In images where the predominant color in the stands is yellow or cyan, some of the  $N_v$  sets of data analyzed could be composed entirely of non-pitch pixels. An example of this can be seen in Fig. 6, where the people in the stands wear predominantly yellow, and the first set of data analyzed is representing mainly pixels in the stands. To deal with this issue, before applying Eq. (5), the histogram of  $cd$  values is analyzed. While the histograms of the data sets that are part of the playing field have a very narrow main mode that is close to zero (i.e., the RGB color of



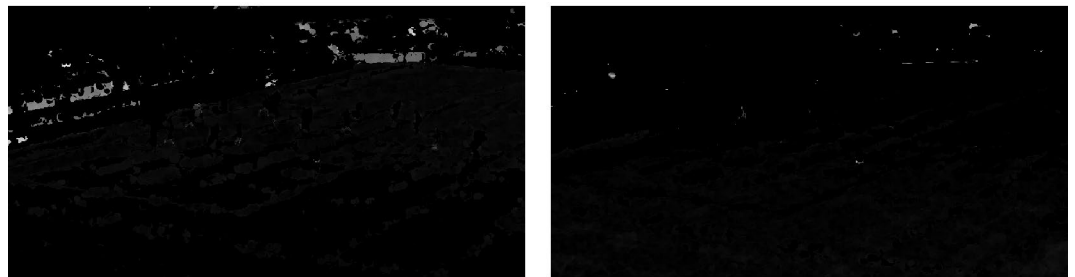
**Figure 4.** Three-dimensional representation of an example of the cone that restricts the acceptable colors given the vectors  $\mathbf{v}$  and  $\mathbf{u}$ .



(a) Set of Gaussians used to estimate the pdf of  $g$ , where the red dotted line indicates the position of  $T_g$  and the green dot indicates the local minimum obtained from the envelope of the Gaussians.



(b) RGB data corresponding to the two modes above  $T_g$ .



(c) Values of  $cd$  corresponding to the two modes above  $T_g$ .



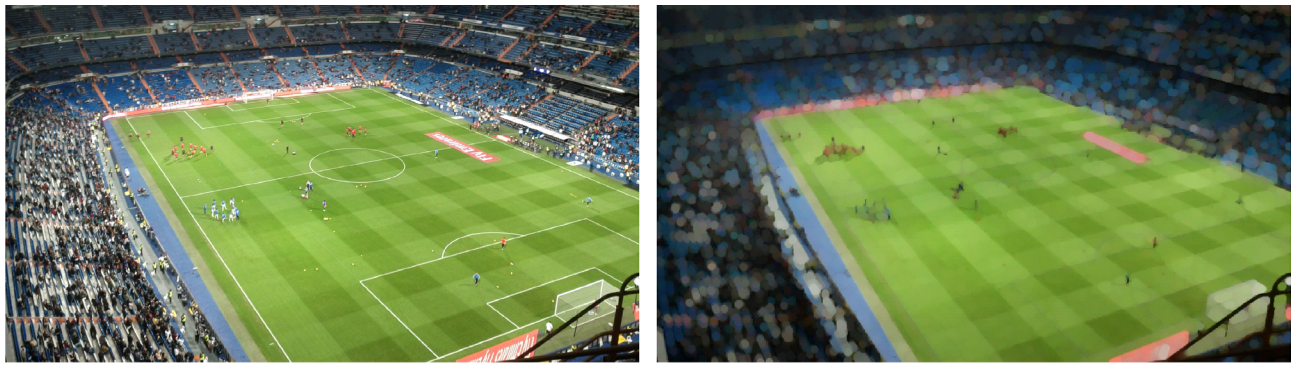
(d)  $\tilde{M}_{PF}$  mask applied to  $I$  (left image) and  $\tilde{M}_{PF}$  applied to  $I$  (right image).

**Figure 5.** Example of the quality improvement obtained with the chromatic distortion analysis. Results corresponding to the original image from<sup>48</sup>.

the vast majority of this data is very uniform), histograms of data sets that do not represent the playing field have a much wider main mode that, sometimes, is placed significantly away from zero (i.e. the RGB colors of the set differ considerably). Therefore, these unwanted data sets can be easily identified and discarded by setting as valid only data sets that have a histogram of  $cd$  with a main mode below  $T_c$ . In the example in Fig. 6 the first two data sets analyzed have histograms that do not meet this condition. Therefore, these two data sets will be discarded.

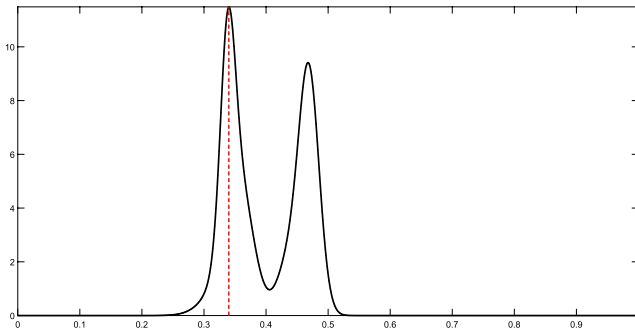
### Post-processing

By means of the analyses described in “Pixel-level analysis” section it is possible to correctly segment the grass of the playing field. However, as these analyses are at the pixel level, the results obtained include some false detections due to elements that do not belong to the playing field but have a very similar color. Furthermore,

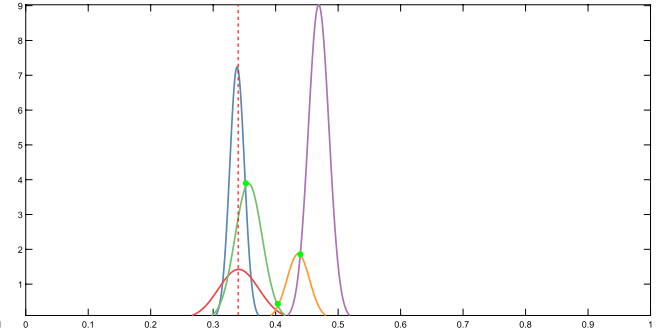


(a) Original RGB image,  $I$ .

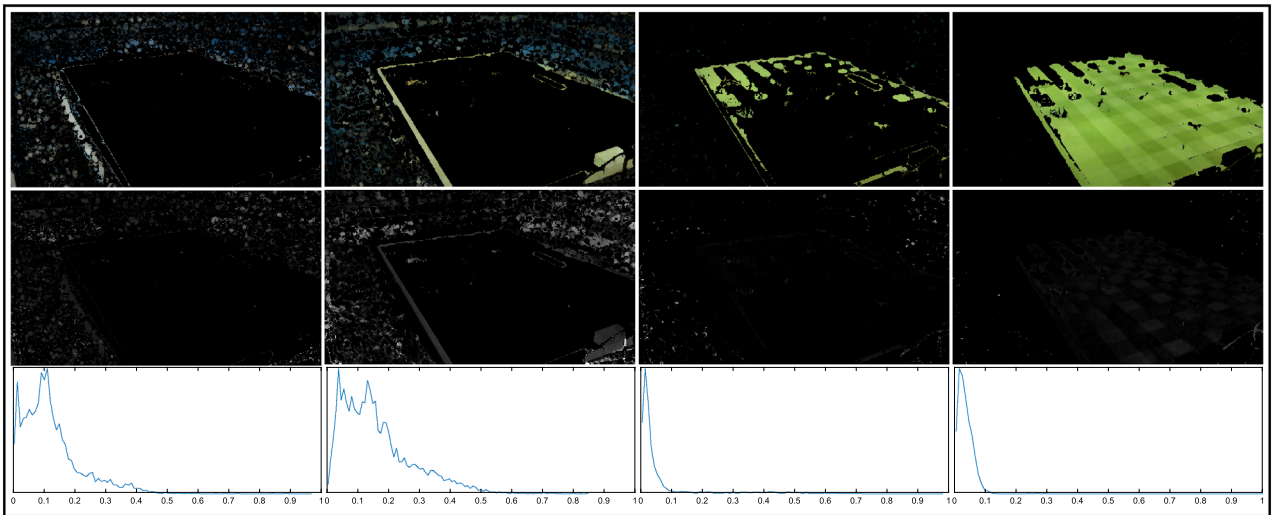
(b) Opened image,  $I_{open}$ .



(c) Estimated pdf, where the red dotted lines indicate the position of  $T_g$ .



(d) Gaussians used to estimate the pdf of  $g$ , position of  $T_g$  (red dotted line), and local minima from the envelope of the Gaussians (green dots).



(e) Masks corresponding to each of the  $N_v$  sets of RGB data (upper row), their values of  $cd$  (center row), and their histograms of  $cd$  (lower row).

**Figure 6.** Analysis of the histograms of the chromatic distortion values. Original image from<sup>46</sup>.

players were also not included in the segmentation result, appearing as holes in the playing field mask. To solve these problems, the proposed strategy includes a final analysis at the region level. This analysis, unlike other segmentation strategies, is completely automatic and avoids the use of morphological operations, which largely depend on the size of the images, the location of the camera, and the zoom used.

The final binary mask,  $M_{PF}$ , resulting from this region-level post-processing is obtained as follows:

1. To include the pixels occupied by the players in the final segmentation, a flood-fill operation<sup>50</sup> is applied to the mask  $\tilde{M}_{PF}$ .
2. Finally, to eliminate false detections, the largest region from the resulting ones is selected.



It is important to note that this simple mechanism to get rid of false detections is made possible by the pre-processing stage included in our strategy. If this pre-processing is not included, the mask  $\hat{M}_{PF}$  could contain multiple regions associated with the playing field and, therefore, it would be necessary to analyze how many regions to select.

## Results

This section summarizes the results provided by the proposed playing field segmentation strategy. First, “[Analysis of the results obtained](#)” section analyzes the quality of the results obtained, as well as the influence on this quality of each of the stages that make up the strategy. Then, “[Parameter analysis](#)” section discusses parameter selection for optimal results. “[Limitations](#)” section discusses the limitations of the strategy. Finally, “[Comparison with other strategies](#)” section compares our results with those obtained using other segmentation strategies.

To analyze the quality of our strategy, images from the following two public databases have been used, which, to our knowledge, are the only ones that provide ground truth files that include binary masks indicating what areas of the images correspond to the playing field:

- LaSoDa: The Labeled Soccer Database (LaSoDa) consists of 60 annotated Full HD images ( $1920 \times 1080$  pixels) corresponding to five matches played in stadiums with different characteristics (different camera positions and different shades of grass). These images show different zoom levels (from images that show only the goal area to images that show more than half of the pitch) and have been acquired with four different types of cameras (master camera, side camera, end camera, and aerial camera). Additionally, it includes challenging lighting conditions (day and night matches and strong contrast between sunlit and shaded areas). This dataset is available at <https://www.gti.ssr.upm.es/data/lasoda>.
- Homayounfar’s database: The database proposed in<sup>51</sup> is composed by 395 HD images ( $1280 \times 720$  pixels) from twenty matches in stadiums with different grass textures and lighting conditions. Unlike LaSoDa, all of its images have been acquired with the master camera (the one used most of the time in soccer broadcasting, placed approximately on the extension of the halfway line) and show similar zoom levels. However, they are more varied than the LaSoDa images in terms of shades of grass and presence of shadows.

Quality has been measured at the pixel level by the recall (*rec*), precision (*pre*), and F-score (*f*) as follows:

$$\text{rec} = \frac{\text{tp}}{\text{tp} + \text{fn}}, \text{pre} = \frac{\text{tp}}{\text{tp} + \text{fp}}, f = \frac{2\text{tp}}{2\text{tp} + \text{fp} + \text{fn}}, \quad (6)$$

where *tp*, *fn*, and *fp* are, respectively, the amounts of true positives, false negatives and false positives. Note that the F-score is also known as F1-score or Dice Similarity Coefficient (DSC).

Regarding the computational cost of the strategy, the most costly step, by far, is the well-known E-M algorithm that is used to approximate the pdf of *g*. However, the literature reports that it is feasible to run E-M on a problem of our scale (histograms made of just a few hundred data points) within very few milliseconds<sup>52</sup>. Consequently, we consider it feasible to make our system work in real time on video sequences.

**Analysis of the results obtained.** Table 1 summarizes the results obtained for each of the 25 matches in which the 455 analyzed test images are distributed (the images corresponding to all these results are available at <https://www.gti.ssr.upm.es/data/playing-field-segmentation>). These results correspond to the following cases:

- Case 1: Results from the mask  $\hat{M}_{PF}$  (after performing the green chromaticity analysis).
- Case 2: Results from the mask  $\tilde{M}_{PF}$  (after performing the chromatic distortion analysis).
- Case 3: Results from the mask  $M_{PF}$  (final results).

In addition, Fig. 7 shows some representative results obtained in images with different lighting conditions, shades of grass, zoom levels and colors on billboards and stands.

The high recall values obtained after applying the green chromaticity analysis (Case 1) shows that this first stage of analysis correctly identifies the vast majority of the pixels that make up the playing field. However, false detections due to the presence of cyan or yellow regions have resulted in significantly lower precision values, especially in the case of the images of some matches (e.g., Match 3) in Homayounfar’s database in which the predominant color in the stands is yellow or cyan.

Most of these false detections disappear after applying the chromatic distortion analysis (Case 2), which results in a significant increase in precision.

The final results (Case 3) show that, after applying the analysis at the region level, an improvement in both recall and precision is achieved. This is because the gaps due to the presence of players on the playing field have been filled in and, in addition, the false detections caused by small regions in the stands with colors similar to those of the grass have been eliminated.

**Parameter analysis.** We had previously stated that the strategy depends on three parameters that must be configured manually (one in the pre-processing stage and two in the pixel-level analysis stage). In this subsection, the influence of these parameters on the quality of the results is analyzed.

The results in Table 1 have been obtained with the combination of parameters that has resulted in the highest overall F-score. These parameters are summarized in Table 2, whereas the graphs in Fig. 8 report the variations



**Figure 7.** Some representative results obtained with the proposed strategy. Original images from<sup>46</sup> and<sup>53</sup>.

in quality of the results when any one of them is modified. The following conclusions can be obtained from these graphs:

- Proportionality factor that determines the diameter of the structuring element used in the pre-processing,  $\alpha_e$ : Although the best results are obtained with  $\alpha_e = 0.5$ , for higher values of this parameter the quality is only very slightly reduced. On the other hand, if  $\alpha_e$  is too low (e.g.,  $\alpha_e = 0.25$ ) or the pre-processing is not applied (i.e.,  $\alpha_e = 0$ ), the quality reduction is very noticeable, since the white lines are not well integrated into the grass.
- Maximum number of Gaussian distributions in the estimation of the pdf of the green chromaticity,  $N_G$ : For values above 2 the quality is very similar, being slightly better in the case of  $N_G = 6$ .
- Maximum allowed chromatic distortion,  $T_c$ : The quality of the results is very high with values of  $T_c$  in a relatively wide band ( $T_c \in [0.15, 0.4]$ ). Outside this range the quality is noticeably reduced.

This analysis shows that none of the parameters is especially critical for the strategy, since all of them have significantly wide ranges of values in which the quality of the results is very similar.

**Limitations.** It should be noted that the proposed segmentation strategy is based on the assumption that the playing field is the largest green element in the image.

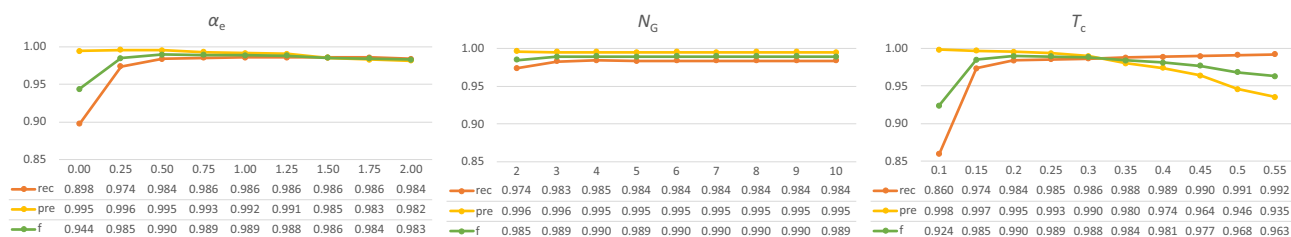
Consequently, it can fail in scenarios where the playing field is surrounded by large regions that are also green. Although these situations are not common in professional stadiums (there is usually a wide variety of colors in the stands due to the amount of spectators that occupies them), they can occur in stadiums with green stands

Match	Frames	Case 1			Case 2			Case 3		
		rec	pre	f	rec	pre	f	rec	pre	f
LaSoDa										
1	12	0.984	0.865	0.920	0.970	0.937	0.953	0.988	0.990	0.989
2	12	0.993	0.974	0.984	0.978	0.985	0.982	0.992	0.997	0.994
3	12	0.986	0.964	0.975	0.983	0.973	0.978	0.994	0.988	0.991
4	12	0.991	0.923	0.956	0.975	0.978	0.976	0.989	0.995	0.992
5	12	0.980	0.935	0.957	0.963	0.981	0.972	0.982	0.987	0.984
Homayounfar										
1	22	0.996	0.661	0.795	0.959	0.971	0.965	0.983	0.996	0.989
2	14	0.988	0.893	0.938	0.959	0.976	0.967	0.977	0.993	0.985
3	15	0.986	0.852	0.914	0.969	0.987	0.978	0.986	0.998	0.992
4	17	0.998	0.756	0.860	0.963	0.987	0.975	0.981	0.997	0.989
5	16	0.980	0.942	0.960	0.970	0.985	0.977	0.987	0.988	0.987
6	17	0.961	0.929	0.944	0.951	0.980	0.966	0.976	0.999	0.987
7	11	0.978	0.960	0.969	0.973	0.994	0.984	0.983	1.000	0.991
8	29	0.997	0.722	0.837	0.964	0.988	0.976	0.981	0.996	0.988
9	33	0.996	0.760	0.862	0.955	0.987	0.971	0.971	0.999	0.985
10	12	0.976	0.966	0.971	0.966	0.992	0.979	0.986	0.998	0.992
11	15	0.986	0.905	0.944	0.973	0.993	0.983	0.989	0.998	0.994
12	23	0.990	0.875	0.929	0.970	0.993	0.982	0.985	0.999	0.992
13	15	0.990	0.923	0.955	0.971	0.991	0.981	0.983	0.994	0.988
14	15	0.983	0.941	0.962	0.973	0.990	0.981	0.988	0.998	0.993
15	20	0.995	0.848	0.916	0.971	0.986	0.978	0.988	0.995	0.992
16	42	0.987	0.833	0.903	0.958	0.978	0.968	0.978	0.998	0.988
17	22	0.983	0.956	0.969	0.971	0.991	0.981	0.988	0.999	0.993
18	18	0.975	0.938	0.956	0.962	0.990	0.976	0.986	0.997	0.991
19	22	0.997	0.735	0.846	0.954	0.985	0.969	0.975	0.999	0.987
20	17	0.989	0.866	0.923	0.969	0.982	0.975	0.989	0.993	0.991
Overall	455	0.988	0.860	0.919	0.967	0.982	0.974	0.984	0.995	0.990

**Table 1.** Summary of results obtained with the proposed strategy.

$\alpha_e$	$N_G$	$T_c$
0.5	6	0.2

**Table 2.** Set of parameters used in the reported results.

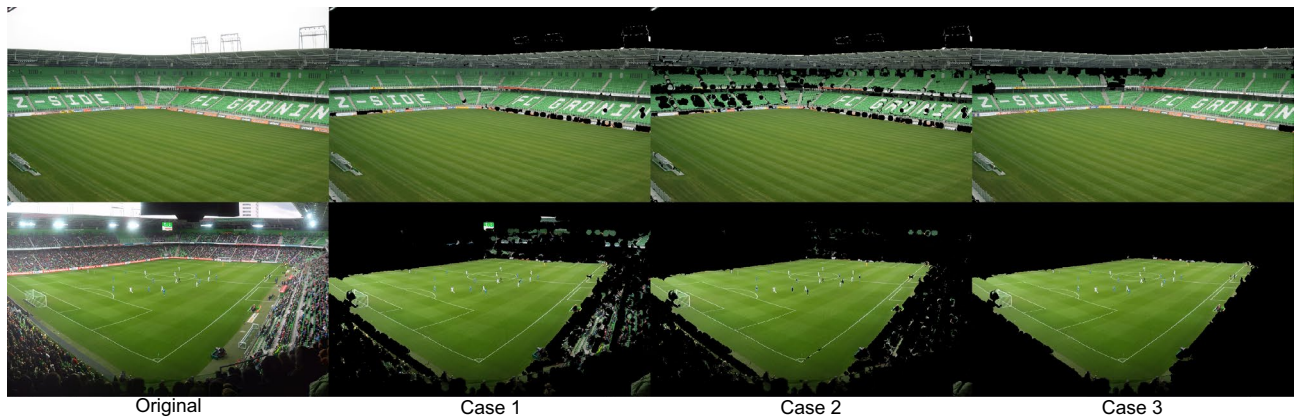


**Figure 8.** Quality of the results based on the values of the parameters on which the proposed strategy depends.

and with little or no fans, or in non-professional playing fields that, instead of being surrounded by stands, are surrounded by vegetation.

An example of this limitation is illustrated in Fig. 9, where the top row of images show the results obtained in a stadium with empty green stands and the bottom row of images shows the results in the same stadium but with a large number of spectators in the stands.

**Comparison with other strategies.** The proposed strategy has been compared to four playing field segmentation methods that are representative of the three types of strategies described in “Related work” section:



**Figure 9.** Results obtained in a stadium with green stands: without an audience (top row of images) and with some audience (bottom row of images). Original images from<sup>54</sup> and<sup>55</sup>.

- M1: RGB-based method used in<sup>37–39</sup>, which is based in the rule  $G > R > B$ .
- M2: RGB-based method recently proposed in<sup>40</sup>, which uses the rules  $G > R$  and  $G > B$ .
- M3: Hue-based method used in the strategies in<sup>32–36</sup>, which is based on separating the dominant mode from the rest of the data in the histogram of the hue component.
- M4:  $g$ -based method used in<sup>31,44</sup>, which is based on the analysis of the pdf of the green chromaticity.

The comparisons have been made both in the case of not applying any post-processing at the region level (called Case 2 in “Analysis of the results obtained” section) and applying post-processing (Case 3). As stated in “Related work” section, the strategies in which these methods are used apply different region-based post-processing stages. To make a fair comparison between methods and given that the post-processing in the proposed strategy is the only one that does not depend on pre-established thresholds, our post-processing has been applied to all methods in the evaluation of the Case 3.

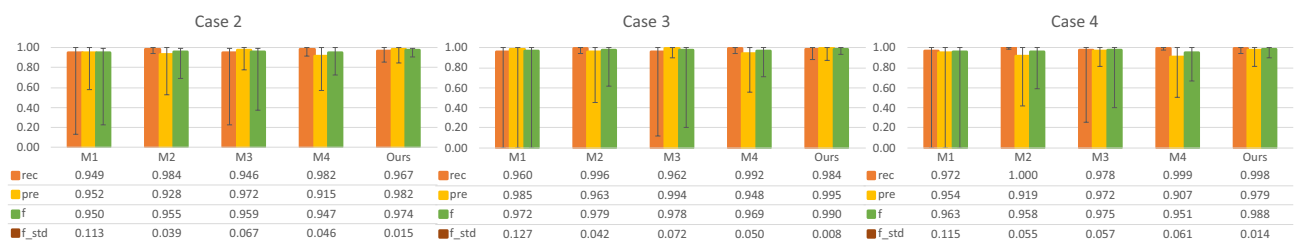
Since the area of the playing field visible in any image is always a single convex region, many of the strategies we compare against apply a convex hull as the last stage of their region-level post-processing. For this reason, we have decided to include in the comparisons a fourth case (Case 4) in which convex hulling is applied as the last stage of the post-processing.

The graphs in Fig. 10 compare the global quality obtained with our strategy and with the 4 previously described methods in the three cases mentioned. In these graphs, in addition to the values of *rec*, *pre*, and *f*, the range of values of each of these variables has also been included, as well as the standard deviation of the values of  $f$  ( $f_{std}$ ).

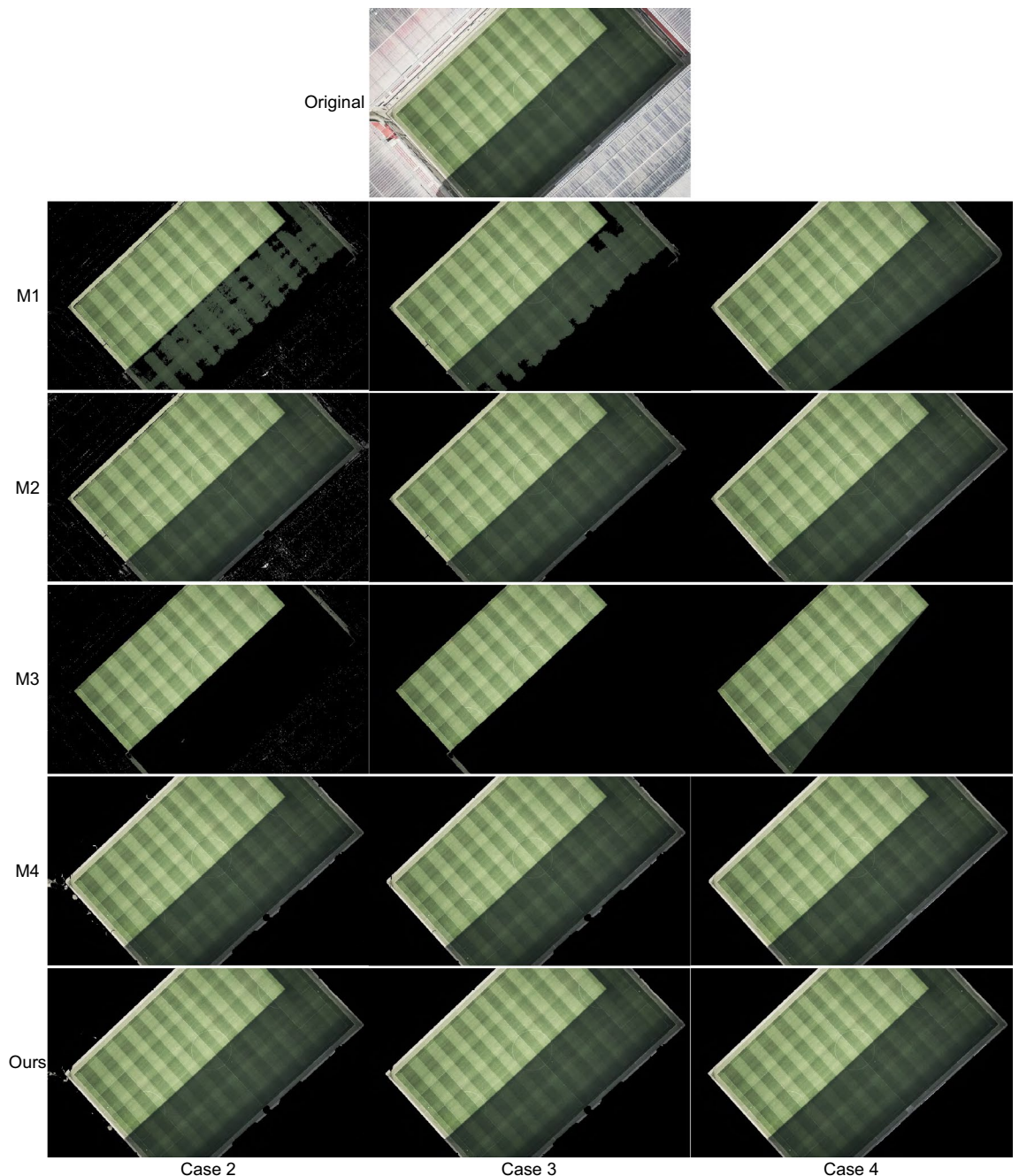
The results before applying the post-processing (Case 2) show that our strategy is the one that obtains the best results overall. The methods M1 and M3 result in many false negatives in images with areas of the playing field with strong shadows (see Fig. 11). In images where the stands include areas with poor color information (i.e. the red, green, and blue channels are very similar) the methods M1 and M2 result in several false positives (see Fig. 12). Regarding, the method M4, it fails in images with areas with colors that cannot be correctly filtered in the green chromaticity color space (e.g., the sky in Fig. 12 or the billboards in Fig. 13).

The graphs in Fig. 10 also show that after including post-processing (Case 3) the quality of the results of the five compared strategies is improved (with our strategy still obtaining the best results).

Regarding the Case 4, as expected, by including the convex hulling in the post-processing the recall of all methods is improved. However, this improvement does not compensate for the worsening of the precision values (i.e., the F-score values get worse). The method least affected by this quality reduction is the one proposed.



**Figure 10.** Summary of the quality obtained with our strategy and 4 other segmentation methods. The error bars report the best and worst result for each metric across the whole dataset, showing the robustness of our proposal.

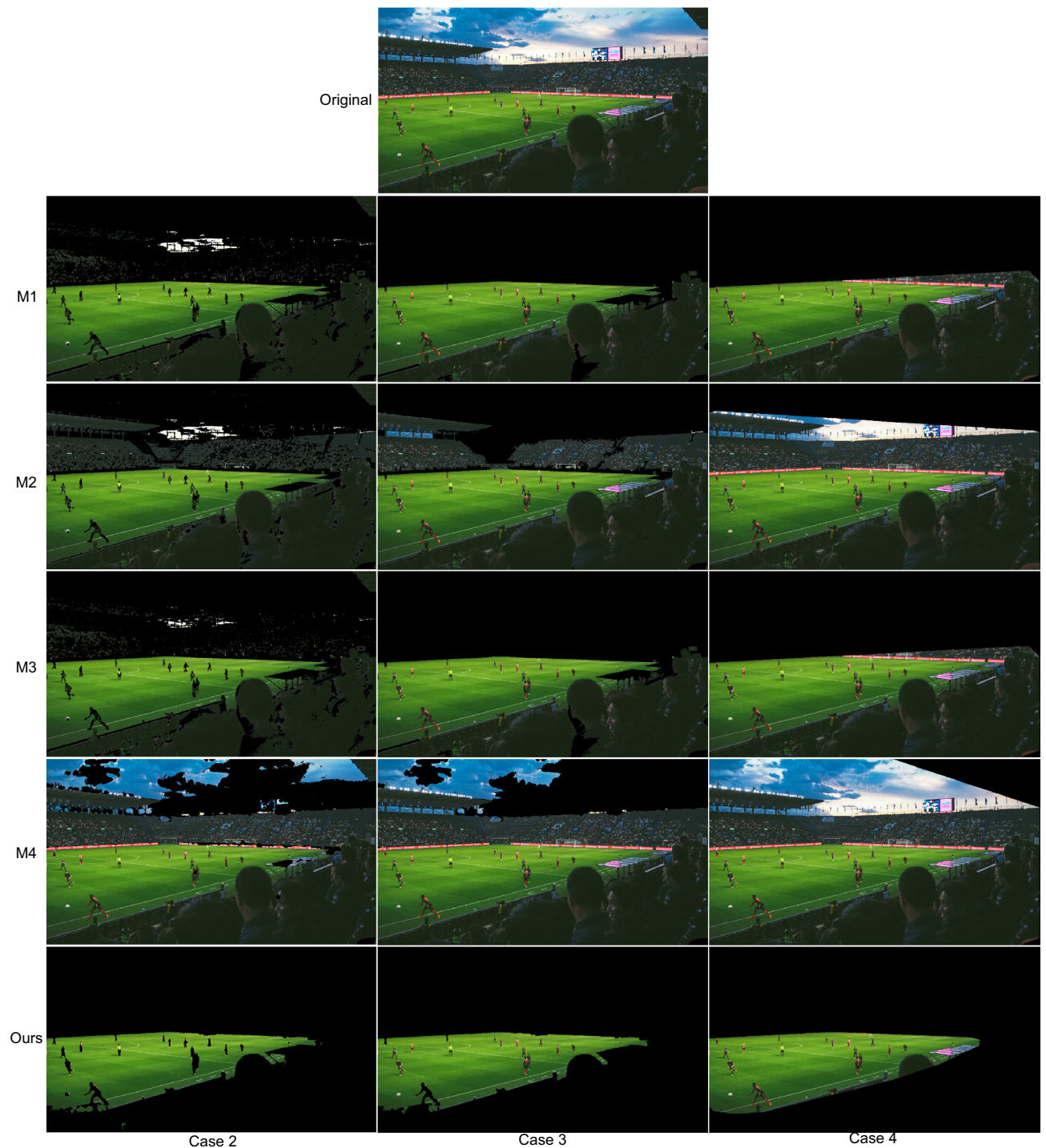


**Figure 11.** Results obtained in an image with strong shadows. Original image from<sup>46</sup>.

Finally, we must take into account that our strategy is not only the one that provides the best overall results (the highest F-score), but it is also the one that provides the lowest value of  $f_{std}$ , which shows that our results are the most consistent.

### Conclusions

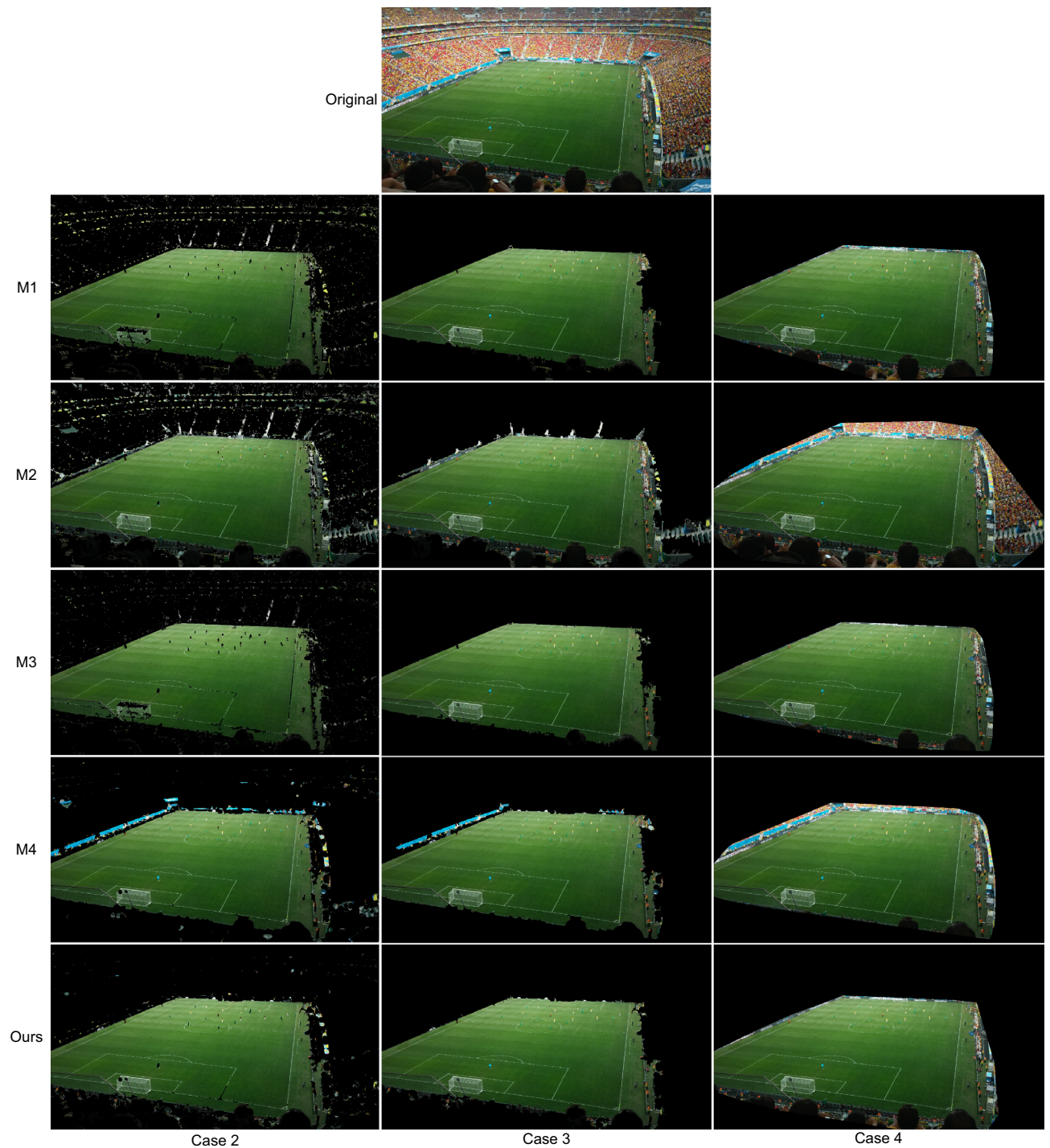
A novel strategy to efficiently segment the playing field in soccer images has been presented. By approximating the pdf of the green chromaticity using the E-M algorithm, the pixels of the image that mainly include green data are segmented. Then, to remove the pixels that are not part of the grass but have colors with an important contribution of green (e.g., yellow), a chromatic distortion analysis is performed. Finally, a region-level stage



**Figure 12.** Results obtained in an image with poor color information in the stands. Original image from<sup>46</sup>.

that does not depend on any user-defined parameter allows removing regions that do not belong to the playing field but that show a color similar to that of grass.

This strategy has been validated on two databases that feature a wide variety of stadiums with different shades of grass and illumination conditions, yielding excellent results with a single set of parameters across both databases and clearly outperforming existing state-of-the-art approaches.



**Figure 13.** Results obtained in an image with cyan billboards. Original image from<sup>46</sup>.

### Data availability

All the original images and their corresponding results are available at <https://www.gti.ssr.upm.es/data/playing-field-segmentation>.

Received: 25 August 2022; Accepted: 23 January 2023

Published online: 26 January 2023

### References

1. Khodaei, M. & Mathern, S.A. Soccer. In: *Sports-Related Fractures, Dislocations and Trauma* 951–953. <https://doi.org/10.1007/978-3-030-36790-9> (2020).

2. Lucas, G. M. *et al.* Goallll!: Using sentiment in the world cup to explore theories of emotion. *Image Vis. Comput.* **65**, 58–65. <https://doi.org/10.1109/acii.2015.7344681> (2017).
3. Tovar, J. Soccer, world war II and coronavirus: A comparative analysis of how the sport shut down. *Soccer Soc.* **22**(1–2), 66–74. <https://doi.org/10.1080/14660970.2020.1755270> (2021).
4. Eryarsoy, E. & Delen, D. Predicting the outcome of a football game: A comparative analysis of single and ensemble analytics methods. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences* 1107–1115. <https://doi.org/10.24251/HICSS.2019.136> (2019).
5. Félix, L. G. S., Barbosa, C. M., Carvalho, I. A., da F. Vieira, V. & Xavier, C. R. Forecasting soccer market tendencies using link prediction. In: *Computational Science and Its Applications—ICCSA 2020*, 663–675. [https://doi.org/10.1007/978-3-030-58799-4\\_48](https://doi.org/10.1007/978-3-030-58799-4_48) (2020).
6. Kirkendall, D. T. Evolution of soccer as a research topic. *Prog. Cardiovasc. Dis.* **63**(6), 723–729. <https://doi.org/10.1016/j.pcad.2020.06.011> (2020).
7. Cuevas, C., Quilon, D. & Garcia, N. Techniques and applications for soccer video analysis: A survey. *Multimed. Tools Appl.* **79**(39), 29685–29721. <https://doi.org/10.1007/s11042-020-09409-0> (2020).
8. Goebert, C. & Greenhalgh, G. P. A new reality: Fan perceptions of augmented reality readiness in sport marketing. *Comput. Hum. Behav.* **106**, 106231 (2020).
9. Andrienko, G. *et al.* Constructing spaces and times for tactical analysis in football. *IEEE Trans. Visual Comput. Graph.* **27**(4), 2280–2297. <https://doi.org/10.1109/TVCG.2019.2952129> (2021).
10. Narizuka, T., Yamazaki, Y. & Takizawa, K. Space evaluation in football games via field weighting based on tracking data. *Sci. Rep.* **11**(1), 1–8. <https://doi.org/10.1038/s41598-021-84939-7> (2021).
11. Pons, E. *et al.* Integrating video tracking and GPS to quantify accelerations and decelerations in elite soccer. *Sci. Rep.* **11**(1), 1–10. <https://doi.org/10.1038/s41598-021-97903-2> (2021).
12. Armenteros, M., Benitez, A. J. & Betancor, M. A. *The Use of Video Technologies in Refereeing Football and Other Sports*. <https://doi.org/10.4324/9780429455551-8> (2019).
13. Caetano, F. G. *et al.* Football player dominant region determined by a novel model based on instantaneous kinematics variables. *Sci. Rep.* **11**(1), 1–10. <https://doi.org/10.1038/s41598-021-97537-4> (2021).
14. Rezaei, A. & Wu, L. C. Automated soccer head impact exposure tracking using video and deep learning. *Sci. Rep.* **12**(1), 1–12. <https://doi.org/10.1038/s41598-022-13220-2> (2022).
15. Cuevas, C., Berjón, D. & García, N. Grass band detection in soccer images for improved image registration. *Signal Process. Image Commun.* **109**, 116837. <https://doi.org/10.1016/j.image.2022.116837> (2022).
16. Mittal, H. *et al.* A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. *Multimed. Tools Appl.* **81**(24), 35001–35026. <https://doi.org/10.1007/s11042-021-10594-9> (2022).
17. Zaitoun, N. M. & Aqel, M. J. Survey on image segmentation techniques. *Procedia Comput. Sci.* **65**, 797–806. <https://doi.org/10.1016/j.procs.2015.09.027> (2015).
18. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66. <https://doi.org/10.1109/TSMC.1979.4310076> (1979).
19. Nock, R. & Nielsen, F. Statistical region merging. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(11), 1452–1458. <https://doi.org/10.1109/TPAMI.2004.110> (2004).
20. Dhanachandra, N., Manglem, K. & Chanu, Y. J. Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. *Procedia Comput. Sci.* **54**, 764–771. <https://doi.org/10.1016/j.procs.2015.06.090> (2015).
21. Najman, L. & Schmitt, M. Watershed of a continuous function. *Signal Process.* **38**(1), 99–112. [https://doi.org/10.1016/0165-1684\(94\)90059-0](https://doi.org/10.1016/0165-1684(94)90059-0) (1994).
22. Kass, M., Witkin, A. & Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1**(4), 321–331. <https://doi.org/10.1007/BF00133570> (1988).
23. Boykov, Y., Veksler, O. & Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239. <https://doi.org/10.1109/34.969114> (2001).
24. Plath, N., Toussaint, M. & Nakajima, S. Multi-class image segmentation using conditional random fields and global classification. In: *Proceedings of the 26th Annual International Conference on Machine Learning* 817–824. <https://doi.org/10.1145/1553374.1553479> (2009).
25. Minaee, S. *et al.* Image segmentation using deep learning: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(7), 3523–3542. <https://doi.org/10.1109/TPAMI.2021.3059968> (2022).
26. Cuevas, C., Martínez, R. & García, N. Detection of stationary foreground objects: A survey. *Comput. Vis. Image Underst.* **152**, 41–57. <https://doi.org/10.1016/j.patcog.2017.09.009> (2016).
27. Berjón, D., Cuevas, C., Morán, F. & García, N. Real-time nonparametric background subtraction with tracking-based foreground update. *Pattern Recognit.* **74**, 156–170. <https://doi.org/10.1016/j.patcog.2017.09.009> (2018).
28. Dulayatrakul, J., Prasertsakul, P., Kondo, T. & Nilkhamhang, I. Robust implementation of hand gesture recognition for remote human–machine interaction. In: *2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE)* 247–252. <https://doi.org/10.1109/icitee.2015.7408950> (IEEE, 2015).
29. Mandellos, N. A., Keramitsoglou, I. & Kiranoudis, C. T. A background subtraction algorithm for detecting and tracking vehicles. *Expert Syst. Appl.* **38**(3), 1619–1631. <https://doi.org/10.1016/j.eswa.2010.07.083> (2011).
30. Yao, R., Lin, G., Xia, S., Zhao, J. & Zhou, Y. Video object segmentation and tracking: A survey. *ACM Trans. Intell. Syst. Technol. (TIST)* **11**(4), 1–47. <https://doi.org/10.1145/3391743> (2020).
31. Cuevas, C., Quilon, D. & García, N. Automatic soccer field of play registration. *Pattern Recognit.* **103**, 107278. <https://doi.org/10.1016/j.patcog.2020.107278> (2020).
32. Kataoka, H., Hashimoto, K. & Aoki, Y. Player position estimation by monocular camera for soccer video analysis. In: *SICE Annual Conference 2011 1985–1990* (2011).
33. Hoernig, M., Herrmann, M. & Radig, B. Real time soccer field analysis from monocular TV video data. In: *11th International Conference on Pattern Recognition and Image Analysis (PRIA-11-2013)*, vol. 2 567–570 (2013).
34. Cioppa, A., Deliege, A. & Van Droogenbroeck, M. A bottom-up approach based on semantics for the interpretation of the main camera stream in soccer games. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* 1765–1774. <https://doi.org/10.1109/cvprw.2018.00229> (2018).
35. Hoernig, M., Herrmann, M. & Radig, B. Real-time segmentation methods for monocular soccer videos. *Pattern Recognit. Image Anal.* **25**(2), 327–337. <https://doi.org/10.1134/s105466181502011x> (2015).
36. Javed, A., Malik, K. M., Irtaza, A. & Malik, H. A decision tree framework for shot classification of field sports videos. *J. Supercomput.* <https://doi.org/10.1007/s11227-020-03155-8> (2020).
37. Ali, M. M. N., Abdullah-Al-Wadud, M. & Lee, S.-L. An efficient algorithm for detection of soccer ball and players. In: *Proceedings of 16th ASTL Control and Networking*, vol. 16 39–46 (2012).
38. Rao, U. & Pati, U. C. A novel algorithm for detection of soccer ball and player. In: *IEEE International Conference on Communications and Signal Processing (ICCS)* 0344–0348. <https://doi.org/10.1109/iccs.2015.7322903> (2015).



39. Sarkar, S., Chakrabarti, A. & Mukherjee, D. P. Generation of ball possession statistics in soccer using minimum-cost flow network. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* 2515–2523. <https://doi.org/10.1109/cvprw.2019.00307> (2019).
40. Rianthong, T., Thewsuan, S., Charoenpong, T. & Pattanaworapan, K. A method for detecting lines on soccer field by color of grass variation. In: *IEEE International Conference on Knowledge and Smart Technology (KST)* 131–134. <https://doi.org/10.1109/kst48564.2020.9059550> (2020).
41. Choroś, K. Automatic playing field detection and dominant color extraction in sports video shots of different view types. In: *Multimedia and Network Information Systems* 39–48. [https://doi.org/10.1007/978-3-319-43982-2\\_4](https://doi.org/10.1007/978-3-319-43982-2_4) (2017).
42. Huang, Y., Llach, J. & Bhagavathy, S. Players and ball detection in soccer videos based on color segmentation and shape analysis. In: *International Workshop on Multimedia Content Analysis and Mining* 416–425. [https://doi.org/10.1007/978-3-540-73417-8\\_50](https://doi.org/10.1007/978-3-540-73417-8_50) (2007).
43. Khatoonabadi, S. H. & Rahmati, M. Automatic soccer players tracking in goal scenes by camera motion elimination. *Image Vis. Comput.* **27**(4), 469–479. <https://doi.org/10.1016/j.imavis.2008.06.015> (2009).
44. Quilón, D., Mohedano, R., Cuevas, C. & García, N. Unsupervised high-quality soccer field segmentation. In: *IEEE International Symposium on Consumer Electronics (ISCE)* 1–2. <https://doi.org/10.1109/ISCE.2015.7177808> (2015).
45. Qian, Y. & Lee, D. D. Adaptive field detection and localization in robot soccer. In: *Robot World Cup* 218–229. [https://doi.org/10.1007/978-3-319-68792-6\\_18](https://doi.org/10.1007/978-3-319-68792-6_18) (2016).
46. <https://pxhere.com/>. [Images under CC0 license].
47. Zahn, A. *Manchester United Match Postponed Over Dummy Bomb*. [https://media.socastarm.com/wordpress/wp-content/blogs.dir/1930/files/2019/02/Old\\_Trafford\\_inside\\_20060726\\_1.jpg](https://media.socastarm.com/wordpress/wp-content/blogs.dir/1930/files/2019/02/Old_Trafford_inside_20060726_1.jpg). [Under CC-BY license] (2016).
48. Wikipedia user LtPowers: Women's Professional Soccer—2011 Championship—Flash vs Independence. [https://commons.wikimedia.org/wiki/File:Women%27s\\_Professional\\_Soccer\\_-\\_2011\\_Championship\\_-\\_Flash\\_vs\\_Independence.jpg](https://commons.wikimedia.org/wiki/File:Women%27s_Professional_Soccer_-_2011_Championship_-_Flash_vs_Independence.jpg). [Under CC-BY license] (2011).
49. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**(6), 716–723. [https://doi.org/10.1007/978-1-4612-1694-0\\_16](https://doi.org/10.1007/978-1-4612-1694-0_16) (1974).
50. Soille, P. *Morphological Image Analysis: Principles and Applications* (Springer, 2004). <https://doi.org/10.1007/978-3-662-05088-0>.
51. Homayounfar, N., Fidler, S. & Urtasun, R. Sports field localization via deep structured models. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 5212–5220. <https://doi.org/10.1109/cvpr.2017.427> (2017).
52. Kumar, N. P., Satoor, S. & Buck, I. Fast parallel expectation maximization for Gaussian mixture models on gpus using CUDA. In: *IEEE International Conference on High Performance Computing and Communications* 103–109. <https://doi.org/10.1109/HPCC.2009.45> (2009).
53. Wikipedia user Voltmetro: Lionel Messi missed a penalty kick. [https://en.wikipedia.org/wiki/2018\\_FIFA\\_World\\_Cup\\_Group\\_D#/media/File:FWC\\_2018\\_-\\_Group\\_D\\_-\\_ARG\\_v\\_ISL\\_-\\_Messi\\_penalty\\_kick.jpg](https://en.wikipedia.org/wiki/2018_FIFA_World_Cup_Group_D#/media/File:FWC_2018_-_Group_D_-_ARG_v_ISL_-_Messi_penalty_kick.jpg). [Under CC-BY license] (2018).
54. Wikipedia user Kevster: De Euroborg, het stadion van FC Groningen. <https://commons.wikimedia.org/wiki/File:Euroborg.jpg>. [Under CC-BY license] (2006).
55. Wikipedia user Rickazio: Stadium of FC Groningen. <https://commons.wikimedia.org/wiki/File:EuroborgGroningen2015.jpg>. [Under CC-BY license] (2015).

## Acknowledgements

This work has been partially supported by project PID2020-115132RB (SARAOS) funded by MCIN/AEI/10.13039/501100011033 of the Spanish Government.

## Author contributions

The authors contributed equally to this work.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023