Efficient Optical Character Recognition System for Automatic Soccer Player's Identification

A. Alsuqayhi and J. I. Olszewska

School of Computing and Engineering, University of Huddersfield
Queensgate, Huddersfield, HD1 3DH, United Kingdom
j.olszewska@hud.ac.uk

Abstract. In this paper, we present a new system of automatic character recognition in order to quickly and reliably identify soccer on the basis of their team number for sport's scene understanding. First, the proposed approach extracts the number on the player's jersey by using a chromaticity-based segmentation method. Then, the extracted character recognition is performed by template matching. Hence, the innovative combination of these two techniques leads to a more computationally efficient system for the player identification purpose than the state-of-the-art ones as demonstrated on real-world image large datasets.

Keywords: Scene Understanding; Automated Sport Analysis; Object Detection; Optical Character Recognition; Feature Extraction; HSV Color Space; Boundary Tracing; Template Matching.

1 Introduction

Nowadays, automatic scene understanding of team sports [1] is essential for sport events' refereeing and analysis and involves vision-based technologies such as object detection [2], object recognition [3], tracking [4], or spatio-temporal reasoning [5].

In particular, the automatic identification of team players is of prime importance to support both sport comments production and media archiving.

For that, face recognition techniques such as [6] have been applied to process soccer games [7]. However, this biometric approach is intrinsically not adapted to identify a player whose back is turned to the camera, in which case his face is poorly or not visible at all.

As a result, optical character recognition (OCR) methods have been developed to recognize numbers on team player's uniform. Most of them exploit the temporal redundancy of a character across several frames and thus are only limited to video analysis [8], [9], [10], [11], [12], [13] and not suited for tasks such as still image dataset retrieval. Other works use both facial and textual cues [14], but their computational speed is low.

Hence, in this work, we focus on the sport scene analysis based on the automatic player identification in images of any type, relying on the detection and recognition of the player's jersey number, and therefore, on the development of a full, efficient optical character recognition (OCR) system for this purpose.

OCR major phases are (i) character extraction and (ii) character recognition. In the first step, the system localizes and extracts the character by detecting its geometrical features like edges or color features, or both [15]. In the second step, character recognition is usually performed by matching [16] or by using classifiers, e.g. AdaBoost [17]. However, these existing OCR systems are mainly applied to recognize license plate numbers or handwritten characters, whereas player number recognition presents additional challenges. Indeed, the foreground, i.e. the character, could be highly skewed with respect to the camera, or the background, i.e. the jersey, could be folded so that part of the number could be hidden. Moreover, sport images are often blurred, since cameras or players or both are quickly moving.

In this paper, we propose to automatically extract the characters from the images based on their pixel chromaticity properties, while we use a digit template to recognize the extracted characters, leading to an OCR system robustly dealing with sport applications, while being computationally effective.

No temporal redundancy assumption is made in our method, which is thus valid not only for video frames, but also for still images such as those contained in sport datasets or on Internet.

In our approach, players could be identify even in back profile, since our OCR system detects and recognize characters which could be anywhere on the team player's clothes.

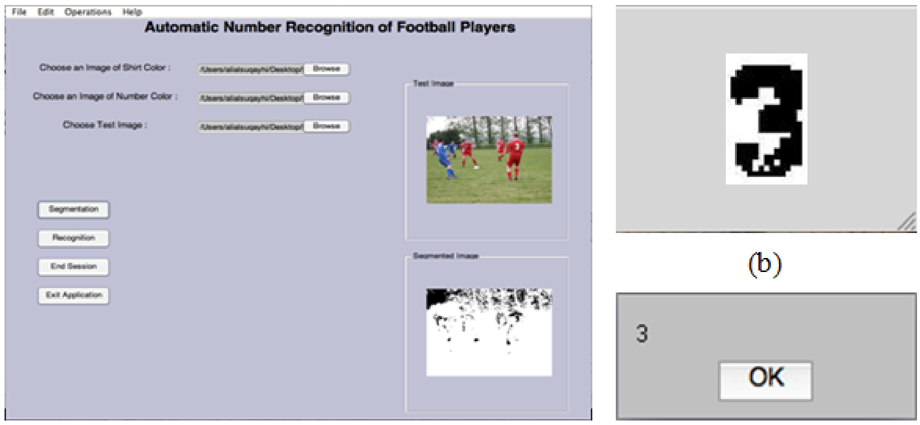
Hence, the contributions of this paper are twofold:

- the association of the chromatic/achromatic segmentation approach with our new adapted boundary tracing algorithm for the robust character extraction, in context of sport event analysis;
- the development of a new powerful OCR system based on the innovative combination of the aforementioned character detection with the template matching-based technique for the fast character recognition, in order to automatically and efficiently identify players in soccer image datasets to understand team sports scenes.

The paper is structured as follows. In Section 2, we describe our optical character recognition approach for fast and effective number extraction and identification. Our method has been successfully applied to soccer players' real-world image datasets as reported and discussed in Section 3. Conclusions are presented in Section 4.

2 Character Recognition and Identification System

In this section, we present our optical character recognition approach for the reliable identification of soccer player's numbers present in real-world images. Firstly, the studied image is chromatically segmented (Fig. 1(a)), and then, the character is extracted (Fig. 1(b)) as explained in Section 2.1. Finally, the



(a) (b) (c)

Fig. 1. Our OCR system's graphic user interface: (a) main system window displaying the initial parameters settings, the input image and the automatically segmented binarized image; (b) pop-up window with the extracted character; (c) pop-up window with the recognized character.

extracted character is recognized (Fig. 1(c)) by means of template matching described in Section 2.2. The implementation of the resulting OCR system is depicted in Section 2.3.

2.1 Character Extraction

Character extraction is performed in two stages, namely, image segmentation and character detection. The first stage consists in the binarization of the image based on chromaticity properties of the foreground and background pixels as described in Section 2.1, while the second stage relies on the characters' inner boundary tracing in order to extract the numbers as presented in Section 2.1.

Image Segmentation Let us consider a color image I, where M and N are its width and height, respectively. The first step to extract numbers or foregrounds of this still image is to separate them from their background. In fact, in football, players' number color is chosen by the football league to be in contrast with players' kit (shirt and sweater), in order to allow visibility of the number in diverse conditions. The study of [18] has found that this contrast is the most important in the hue, saturation and value (HSV) color space when looking at the saturation of the number pixels and the jersey pixels. Consequently, the image I could be segmented based on the low and high saturated pixels, i.e. objects' achromatic and chromatic colors, respectively, leading to a binary image I_B. In particular, a color pixel under investigation $\mathbf{P} = [P_H, P_S, P_V]$ is considered as achromatic if its saturation (P_S) is below the saturation threshold (Y_S) or if

Algorithm 1 Achromatic-color Number & Achromatic-color Jersey

```
if ((N_S < Y_S or N_V < Y_V) and (J_S < Y_S or J_V < Y_V)) then
    if J_V > V_thresh then
        for all P do
            if P_V < V_thresh then
                I_B(P) = 0                              ▷ set pixel as black
            else
                I_B(P) = 1                              ▷ set pixel as white
            end if
        end for
    else
        for all P do
            if P_V < V_thresh then
                I_B(P) = 1                              ▷ set pixel as white
            else
                I_B(P) = 0                              ▷ set pixel as black
            end if
        end for
    end if
end if
return I_B
```

Algorithm 2 Achromatic-color Number & Chromatic-color Jersey

```
if ((N_S < Y_S or N_V < Y_V) and (J_S > Y_S and J_V > Y_V)) then
    for all P do
        if ((P_S < Y_S) and (P_V < Y_V)) then
            I_B(P) = 0                                  ▷ set pixel as black
        else
            if (h_diff(J_H, P_H) < H_thresh) then
                I_B(P) = 1                              ▷ set pixel as white
            else
                I_B(P) = 0                              ▷ set pixel as black
            end if
        end if
    end for
end if
return I_B
```

Algorithm 3 Chromatic-color Number & Achromatic-color Jersey

```
if ((J_S < Y_S or J_V < Y_V) and (N_S > Y_S and N_V > Y_V)) then
    if J_V > V_thresh then
        for all P do
            if ((P_S < Y_S) and (P_V < Y_V)) then
                if P_V < V_thresh then
                    I_B(P) = 0                          ▷ set pixel as black
                else
                    I_B(P) = 1                          ▷ set pixel as white
                end if
            else
                I_B(P) = 0                              ▷ set pixel as black
            end if
        end for
    else
        for all P do
            if (P_S < Y_S and P_V < Y_V) then
                if P_V > V_thresh then
                    I_B(P) = 0                          ▷ set pixel as black
                else
                    I_B(P) = 1                          ▷ set pixel as white
                end if
            else
                I_B(P) = 0                              ▷ set pixel as black
            end if
        end for
    end if
end if
return I_B
```

Algorithm 4 Chromatic-color Number & Chromatic-color Jersey

```
if ((N_S > Y_S and N_V > Y_V) and (J_S > Y_S and J_V > Y_V)) then
    for all P do
        if (h_diff(N_H, P_H) < H_thresh) then
            I_B(P) = 0                                  ▷ set pixel as black
        else
            I_B(P) = 1                                  ▷ set pixel as white
        end if
    end for
end if
return I_B
```

its intensity (P_V) is below intensity threshold (Y_V). If the pixel saturation and intensity are above these thresholds, then it is considered as chromatic.

The segmentation is initialized by defining the mean color vector of the jersey $\mathbf{J} = [J_H, J_S, J_V]$ and the mean color vector for the number $\mathbf{N} = [N_H, N_S, N_V]$, based on provided image samples. Next, the image I is processed depending if the number color is chromatic or achromatic and if the jersey color is chromatic or achromatic, leading to four cases, i.e. to four Algorithms 1-4. The segmentation is based on the hue threshold H_{thresh} and the hue difference in the case of a chromatic-color jersey, whereas the intensity difference and the intensity threshold V_{thresh} are used in the case of an achromatic-color jersey [18]. In the case where the number has an achromatic color and the jersey color is chromatic (Algorithm 2), the hue difference h_{diff} is defined as follows:

$$h_{diff}(J_H, P_H) = \begin{cases} \Delta(J_H, P_H) & \text{if} \Delta(J_H, P_H) < 180°, \\ 360° - \Delta(J_H, P_H) & \text{otherwise,} \end{cases} \tag{1}$$

with

$$\Delta(J_H, P_H) = \mid J_H - P_H \mid . \tag{2}$$

When both the jersey and the number have chromatic colors, the image is segmented as described in Algorithm 4, using the hue difference h_{diff} defined as follows:

$$h_{diff}(N_H, P_H) = \begin{cases} \Delta(N_H, P_H) & \text{if} \Delta(N_H, P_H) < 180°, \\ 360° - \Delta(N_H, P_H) & \text{otherwise,} \end{cases} \tag{3}$$

with

$$\Delta(N_H, P_H) = \mid N_H - P_H \mid . \tag{4}$$

Character Detection In the binarized image I_B computed by the process explained in Section 2.1, jerseys appear as white objects, while numbers as black ones. Based on that fact, tracing internal boundaries of these objects is an efficient method for number region localization and extraction. For this purpose, we have adapted the Boundary Tracing approach [19]. Hence, our process presented in Algorithm 5 initiates by tracing all the boundaries B_i within the segmented binary image, and then, in relation to the specific area aspect ratio F characterizing the number region, the boundaries are filtered, in order to select only those containing numbers. Once this process is completed, the binary image I_B is cropped and the cropped image I_C is transferred to the recognition stage which then identifies the numbers as detailed in Section 2.2.

This section has presented the single digit case. The identification of two-digit numbers is as follows. If two cropped images are of the same size and are in adjacent bounding rectangles, they are flagged as forming a two-digit number.

Algorithm 5 Boundary Tracing

Step 1
Find boundaries $B = \{B_i\}$ of all objects

Step 2
for all B_i **do**
    if B_i of black object **then**
        if B_i dimensions $= F$ dimensions **then**
            $x_1 = \min(B_i[1])$
            $y_1 = \min(B_i[2])$
            $x_2 = \min(B_i[1])$
            $y_2 = \max(B_i[2])$
            $I_C = I_B[x_1 : x_2][y_1 : y_2]$
        else
            Ignore boundary
            Go the next boundary
        end if
    end if
end for
return I_C

2.2 Character Recognition

For this phase, we have adopted a template matching approach. Indeed, this pattern classification method is well suited in the identification of small regions [20], which is the case in our application.

The basis of template matching is that a processed image is compared to each of the images stored within a template. In many instances, the extracted number region has smaller or larger dimensions compared to the template dimensions, or has not the same orientation. Thus, the extracted number image has first to be rotated and rescaled to fit the template orientation and size, respectively. Then, the correlation coefficient r between the two compared images is computed as follows:

$$r = \frac{\sum_m \sum_n (T_{mn} - \bar{T})(S_{mn} - \bar{S})}{\sqrt{\left[\sum_m \sum_n (T_{mn} - \bar{T})^2\right]\left[\sum_m \sum_n (S_{mn} - \bar{S})^2\right]}}, \quad (5)$$

where T_{mn} are the values of the pixels of the template image with an $m \times n$ size and a mean $\bar{T}$; S_{mn} are the values of the pixels of the processed image, i.e. the rescaled cropped binarized image, with a mean $\bar{S}$.

When the structure of the processed image is greatly similar to the structure of one of the template images, then the correlation coefficient value is high and this means the number is identified.

To recognize two-digit numbers, single numbers flagged as constituting a two-digit number in Section 2.1 are recognized individually by matching each of them with the template. The two-digit number is then formed based on that information.

We can notice that the use of the template matching technique is well suited for our system of automatic number recognition of soccer players. On one hand, template matching is particularly fast when used in context of our system, because it requires only the recognition of numerical characters, rather than a wider range of alphanumerical characters as in other applications, such as license plate recognition (LPR). Indeed, our template stores in total only 10 images of one-digit numbers (0 to 9). Hence, the matching is performed against a maximum of ten stored images, in order to recognize the extracted character, which is computationally very efficient. Moreover, the scale sensitivity of the template matching technique is used in our work as an advantage, since smaller dimensions of the template dimensions lead to a faster matching. On the other hand, the recognition rate obtained by our implementation of this method in our system is much higher than those presented in the literature as discussed in Section 3.

2.3 Our OCR System

Our overall system such as presented in Fig. 1(a) has been developed using MatLab software for both the implementation of the algorithms presented in Sections 2.1-2.2 and for the design of the graphical user interface (GUI) to interact with our OCR system.

The system is able to support different types of image formats such as jpeg, tiff, bmp, and png.

Our OCR system's main steps are as follows. At first, an image dataset is provided as input. Then, the user could enter the initial parameters such as jersey and number color samples. The second phase is the image segmentation as detailed in Section 2.1. Next, the system processes the extraction of the character (Section 2.1) which is then displayed in a pop-up window as illustrated in Fig. 1(b). Finally, the recognized character (Section 2.2) appears in another window as in Fig. 1(c).

3 Results Evaluation and Discussion

To validate our method, we have carried out experiments which consist in automatically recognizing numbers from the soccer players' jerseys within a database containing data images with soccer-related content, as such illustrated in Figs. 1-2.

For this purpose, our system has been applied on a dataset containing 4500 football images whose average resolution is of 230x330 pixels and which were captured in outdoor environment. This database owns challenges of quantity, pose and scale variations of the players. Moreover, the colors of the teams' uniforms have various colors and the fonts on the players' jerseys could vary strongly.

All the experiments have been run on a computer with an Intel(R) Core(TM)2 Duo 2.53 GHz processor, 4 Gb RAM, and using our OCR software.

In order to assess the performance of our OCR system, we use the following criteria:

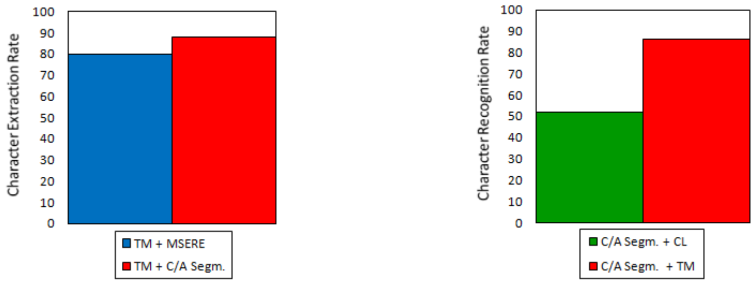
$$extraction\ rate = \frac{CL \times 100}{TT}, \tag{6}$$

Fig. 2. Examples of results obtained with our OCR system. First row: input image. Second row: segmented image. Third row: extracted character. Fourth row: recognized character.

$$recognition\ rate = \frac{CR \times 100}{TT}, \tag{7}$$

with CL, the number of correctly localized characters, CR, the number of correctly recognized characters, and TT, the total number of tested characters.

Some examples of the results of our OCR system are presented in Fig. 2. These samples present difficult situations such as variability of the jerseys' colors, i.e. different pixels' chromaticity properties of the foregrounds and the backgrounds; numbers' changing characteristics, i.e. different characters' geometrical



(a) (b)

Fig. 3. Comparison of performance in (a) character extraction and in (b) character recognition. In each phase, our approach (right bar in red) outperforms the other methods (left bars).

Table 1. Average rates of the automatic character extraction and the automatic character recognition obtained for all the dataset using different approaches.

Rate	[14]	[18]	**our**
character extraction rate	80.0%	83.0%	**88.0%**
character recognition rate	67.5%	52.0%	**86.0%**

and spatial properties; scale effects such as zoom out (Fig. 1 (a) - Fig. 2 (a)) or close-up (Fig. 2 (b)); and varying number of objects of interest.

We can observe that using our approach, characters are correctly extracted as displayed in Figs. 2(e)-(f) and correctly recognized as in Figs. 2(g)-(h), despite their geometrical and chromatical differences. Hence, our OCR system is robust towards changes in numbers and colors of the foregrounds and the backgrounds as well as towards variations of fonts, size, and orientation of the characters.

In Table 1, we have reported the extraction and recognition rates of our OCR method combining chromatic/achromatic segmentation and template matching *(C/A Segm. + TM)* against the rates achieved by approaches such as *MSERE + TM* [14] and *C/A Segm. + CL* [18]. Indeed, [14] uses matching to recognize characters, whereas its strategy to detect them involves maximally stable extremal region extraction. On the other hand, the character extraction in the work of [18] relies on pixels' chromaticity property, but it applies a classifier to recognize the detected characters.

We can see in Table 1 that our OCR method relying on the combination of chromatic/achromatic segmentation and matching-based recognition outperforms the state-of-art approaches for soccer player's number identification. In particular, we can notice in Fig. 3 (a) than the extraction rate is improved when using the chromatic/achromatic segmentation instead of other state-of-the-art techniques such as maximally stable extremal region extraction, while from Fig. 3 (b), we can observe the positive effect of our template matching approach on the recognition rate compared to other classification methods.

Moreover for all the dataset, the average computational speed of our combined OCR method is in the range of few seconds, and thus, our developed system could be used in context of online scene analysis.

4 Conclusions

To automatically identify players in images of different types, we have adopted an OCR approach rather than a face recognition one, in order to recognize players independently of their orientation to the camera. Whereas a lot of research has been done to develop effective OCR systems for handwriting and license plate character recognition, only few works have tackled with OCR systems for player numbers' recognition. In this paper, we have proposed a new OCR system whose performance are greater than the ones found in the literature in both extraction and recognition of soccer players' numbers. Our OCR system is based on the novel combination of the chromatic-contrast segmentation and template-based character recognition for a robust and computationally efficient player identification in soccer image datasets. For all these reasons, our OCR approach is well suited for automatic online sport image retrieval and team sport's scene understanding.

References

1. Olszewska, J. I., McCluskey, T. L.: Ontology-coupled active contours for dynamic video scene understanding. In: Proceedings of the IEEE International Conference on Intelligent Engineering Systems, pp. 369–374 (2011)
2. Alqaisi, T., Gledhill, D., Olszewska, J. I.: Embedded double matching of local descriptors for a fast automatic recognition of real-world objects. In: Proceedings of the IEEE International Conference on Image Processing (ICIP'12), pp. 2385–2388 (2012)
3. Olszewska, J. I.: A new approach for automatic object labeling. In: Proceedings of the EPSRC Workshop on Vision and Language, Sheffield, UK (2012)
4. Olszewska, J. I.: Multi-target parametric active contours to support ontological domain representation. In: Proceedings of RFIA, pp. 779–784 (2012)
5. Olszewska, J. I.: Spatio-temporal visual ontology. In: Proceedings of the EPSRC Workshop on Vision and Language, Brighton, UK (2011)
6. Wood, R., Olszewska, J. I.: Lighting-variable AdaBoost based-on system for robust face detection. In: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, pp. 494–497 (2012)
7. Bertini, M., Del Bimbo, A., Nunziati, W.: Player identification in soccer videos. In: Proceedings of the ACM SIGMM International Workshop on Multimedia Information, pp. 25–32 (2005)

8. Andrade, E. L., Khan, E., Woods, J. C., Ghanbari, M.: Player classification in interactive sport scenes using prior information region space analysis and number recognition. In: Proceedings of the IEEE International Conference on Image Processing, pp. III.129–III.132 (2003)
9. Kokaram, A., Rea, N., Dahyot, R., Tekalp, A. M., Bouthemy, P., Gros, P., Sezan, I.: Browsing sports video: Trends in sports-related indexing and retrieval work. IEEE Signal Processing Magazine. **23**, 47–58 (2006)
10. D'Orazio, T., Leo, M.: A review of vision-based systems for soccer video analysis. Pattern Recognition. **43**, 2911–2926 (2010)
11. Huang, C.-L., Shih, H.-C., Chao, C.-Y.: Semantic analysis of soccer video using dynamic Bayesian network. IEEE Transactions on Multimedia. **8**, 749–760 (2006)
12. Ekin, A., Tekalp, M., Mehrotra, R.: Automatic soccer video analysis and summarization. IEEE Transactions on Image Processing. **12**, 796–806 (2003)
13. Niu, Z., Gao, X., Tao, D., Li, X.: Semantic video shot segmentation based on color ratio feature and SVM. In: Proceedings of the IEEE International Conference on Cyberworlds, pp. 157–162 (2008)
14. Bertini, M., Del Bimbo, A., Nunziati, W.: Matching faces with textual cues in soccer videos. In: Proceedings of the IEEE International Conference on Multimedia and Expo, pp. 537–540 (2006)
15. Lin, C.-C., Huang, W.-H.: Locating license plate based on edge features of intensity and saturation. In: Proceedings of the IEEE International Conference on Innovative Computing, Information and Control, pp. 227–230 (2007)
16. Guanglin, H., Yali, G.: A simple and fast method of recognizing license plate number. In: Proceedings of the IEEE International Forum on Information Technology and Applications, pp. II.23–II.26 (2010)
17. Chen, X., Yuille, A. L.: Detecting and reading text in natural scenes. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. II.366–II.373 (2004)
18. Saric, M., Dujmic, H., Papic, V., Rozic, N.: Player number localization and recognition in soccer video using HSV color space and internal contours. In: Proceedings of the World Academy of Science, Engineering and Technology, pp. 531–535 (2008)
19. Ren, M., Yang, J., Sun, H.: Tracing boundary contours in a binary image. Image and Vision Computing. **20**, 125–131 (2002)
20. Brunelli, R.: Template Matching Techniques in Computer Vision: Theory and Practice. John Wiley & Sons (2009)