Color based soft biometry for hooligans detection

Angela D'Angelo
Eurecom
Multimedia Communications Department
Sophia Antipolis Cedex
FRANCE
Email: angela.dangelo@eurecom.fr

Jean-Luc Dugelay
Eurecom
Multimedia Communications Department
Sophia Antipolis Cedex
FRANCE
Email: Jean-Luc.Dugelay@eurecom.fr

***Abstract*—Biometric systems based on human traits like fingerprinting, face, iris, etc., even widely explored in literaure, cannot be used in scenarios like video surveillance of crowded envinroments. A different class of biometry, usually referred as *soft biometry*, including individual's height, weight, skin color, clothes color, trajectory, etc., can be more easily extracted from cameras in a large network to provide some useful information about the users. In this work we focus our attention on the color of human clothes. We apply color based soft biometric to a specific scenario, i.e. to prevent hooligans fights, based on the color of the clothes they wear. A video surveillance system able to quickly identify the presence of rival teams fans in a specific location would be helpful in avoiding vandalism and destructive behaviour.**

I. Introduction

In the last years, biometrics has been rapidly gaining acceptance as the technology that can meet the ever increasing need for security in critical applications. It is an authentication mechanism that relies on the automated identification or verification of an individual based on unique physiological or behavioral characteristics. The fundamental requirement of any biometric recognition system is the use of human traits having several desirable properties like universality, distinctiveness, permanence, acceptability, and resistance to circumvention. A wide variety of biometric systems have been developed for automatic recognition of individuals that make use of a single or a combination of traits like fingerprint [1], face [2], hand-geometry, iris, retina, gait, voice, signature, etc. These systems have been widely deployed in forensic, government and commercial applications.

A biometric system has to face with several problems affecting the overall perfomance and the applicability of the system itself, such as noisy sensor data, non-universality and lack of distinctiveness of the chosen biometric trait. Moreover, existing automatic recognition systems based on these biometric identifiers are not able to deal with poor quality images and hence they do not satisfy the high accuracy requirements of critical applications. Classical biometric traits are more applicable in a controlled envinroment with a slow speed and to a relatively small number of cooperative subjects.

A different class of biometry, including height, weight, skin color, hair color, clothes color, motion trajectory, etc., can be, instead, more easily extracted from video cameras in a large network to provide some information about the users. They are, in fact, not expensive to compute, they do not require the cooperation of the subjects and they can be sensed at a distance in a crowded envinronment. These characteristics are usually referred in literature as *soft biometry* [3], because they do not have the distinctiveness and permanence to uniquely identify an individual over a period of time and they are not "legally" accepted identification signatures like fingerprint or DNA.

Among soft biometric signatures, in this work we focus our attention on color, specifically on the color of human clothes. The goal is to develop a color detection framework that is able to robustly identify colors in video sensors in real illumination conditions. A so defined tool could be useful for several security applications. It can be used to track a suspicious person, based on his/her clothing colors, across the fields-of-view (FOV) of multiple cameras. Another application could be to search a person in a crowded enviroment, such as to localize missing children in a amusement park. In this paper we apply the proposed color detection framework to a specific scenario, i.e. to prevent hooligans fights. Soccer hooliganism is unruly and destructive behaviour by association football club fans. Hooligans fights, or, more generally, fights between supporters of rival teams, may take place before or after football matches at pre-arranged locations away from stadiums or they can erupt spontaneously at the stadium or in the surrounding streets. A video surveillance system able to quickly identify the presence of rival teams fans in a specific location would be helpful in avoiding vandalism and destructive behaviour. Based on the idea that sport fans usually wear clothes of the same color of the team they support, in this paper we provide a video analysis framework for hooligans detection. The proposed framework can be also useful to draw statistical considerations about the number of fans at the stadium or their distribution in the stadium.

II. Color detection algorithm

As explained in the introduction, the goal of this work is to design a color detection framework to identify the presence of rival teams fans in a specific location by looking for the colors of the corresponding teams. The proposed framework could be integrated, for example, in the video surveillance system of the stadium. If during football matches the system reveals the presence of rival fans in a specific area of the stadium, the number of policemen or guards in that area should be increased to prevent fights.

From a more technical point of view, the goal is to design a robust color detection framework to search for a specific color in a large camera network. Identifying or matching the surface color of a moving object in a video surveillance system, however, is very critical. Traditional color models provide only little help, since the surface of an object is usually not flat, the object motion can alter the surface's orientation, and the lighting conditions can vary when the object moves. Many papers in different research fields can be found in literaure dealing with this problem, but most of them only try to find a solution for a specific color or application, like skin color identification, or they aim at capture the salient color characteristics like mean color or dominant color, as for image retrieval [4].

To partially overcome the problem of changes in illumination conditions, we performed an extensive analysis on the selection of the more suitable color space and in the meantime we exploited the advantages of color constancy algorithms, leading to the general framework described in Fig.1. In the following a description of the proposed system is provided. For sake of brevity, only a high level description is introduced.

The main blocks of the algorithm are: *color constancy*; *colorspace transformation*; *color matching*.

A. Color constancy

Color constancy is the ability of a vision system to accurately describe the color of an object in spite of variations in illumination condition. Color constancy algorithms [5] have been developed to reduce color variation from fluctuation in source illumination. The main limitation of these algorithms is that they usually assume scene illumination to be uniform in the region of interest or changing gradually. Such an assumption almost always fails in a surveillance scenario. Moreover, most color-constancy algorithms depend on reliable estimates of parameters such as angles between light sources and the object, reflection angles, and surface materials. These parameters can be unknown or difficult to estimate in real-time when the object being observed is in motion.

Even if these algorithms are not robust to reliably identify colors in motion, they are useful to normalize the illumination in the scene and thus to facilitate the following steps. We analyzed some of the most popular color constancy algorithms (Grey-World, Max-RGB, general Grey-World, Shades-of-Gray and Grey-Edge algorithm) as described in [5] and, based on the experimental results, we decided to use the Grey-World algorithm for our color detection framework.

B. Color transformation

Once the illumination in the scene has been normalized through the use of a Grey-World algorithm, it is necessary to design the transformation T by which the new color space is reached from the RGB color space. This is the most important step in the overall system and the main contribution of this work. In fact, in order to facilitate the color matching step, each color should be as much as possible visually distinguishable from the others.

In this context, many works dealing with color recognition are based on the use of the HS plane of the HSV color space thanks to its robustness towards changes in illumination and shadows. Other works exploit the way the Human Visual System (HSV) perceives colors by adopting perceptually uniform color space like the CIE 1976 (Lab) color space. These and many other color spaces have been widely used in the scientific literature [6]. However, quite surprisingly, most of the papers does not provide strict justification of their color space choice neither a comparison of the existing color spaces. Even if it is generally agreed that there is no single color system which is suitable for all color images, an analysis of the performance of the most popular color spaces in identifying colors in motion in video sensors would be an invaluable help for our color identification task.

We proposed a new approach aiming at providing a comparison of the most widely used color spaces in order to understand which one is the more suitable to identify colors in video sensors. We employed a statistical approach by conducting extensive data mining on video clips collected under various lighting conditions and distances from several video-cameras. The goal was to learn how individual colors can drift in different illumination conditions and with different color spaces.

In order to provide a very general approach, we need to collect pixels describing different colors. Thus we quantized the entire color space into eleven bins: black, white, red, yellow, green, blue, brown, purple, pink, orange, and grey. These colors are usually referred to as culture colors, which have been used in literature of different cultures in the past years to refer to colors [7]. The second step consisted in collecting pixels based on the above described color quantization, under various lighting conditions, from different cameras and from several distances from the cameras. Our idea to obtain a so diversified dataset of colors was to collect from the web video clips of eleven teams with the color of the uniforms corresponding to the eleven culture colors, and to obtain the sample colors from them. The video clips of the selected sport teams were chosen randomly from the web. This procedure allowed us to obtain a great number of samples in real illumination conditions and, with very high probability, taken from different cameras.

We collected about 1200 samples for the training and about 1350 samples for the testing and we analyzed five of the most popular and widely used color spaces (RGB, normalize RGB, HSV, Lab, YUV) using a fuzzy k-nearest neighbor classification algorithm [8]. The obtained results showed that the Lab color space is the one that provided the best correct classification rate (94.1%[1]) among the analyzed ones and thus it has been adopted for the proposed color detection framework. Even if the Lab color space is already widely

[1]the other detection rates are: RGB$=$ 91.3%, normalize RGB$=$ 89.9%, HSV$=$ 86.8%, YUV$=$ 91.4%

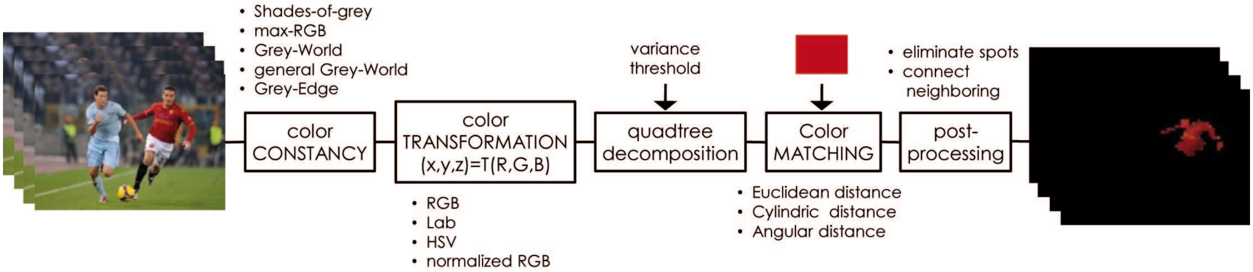


Fig. 1. General scheme of the proposed color detection algorithm.

adopted in image processing since it is a perceptually uniform color space, the performed analysis assures us the robustness of the system for a generic color and in generic illumination conditions.

C. Color matching

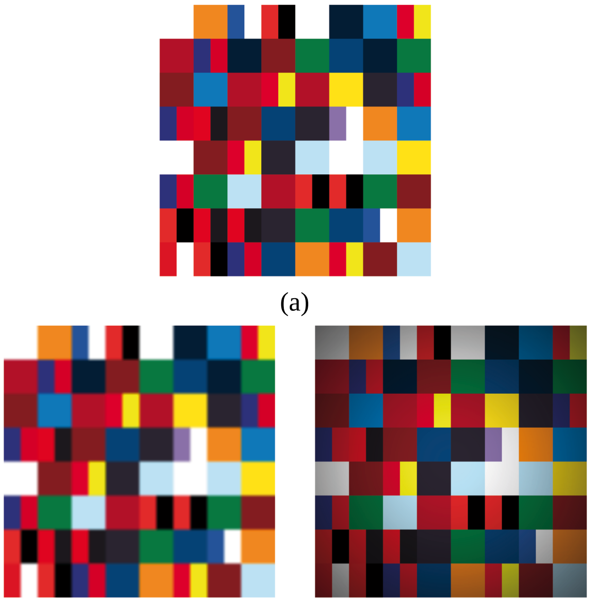
After the color transformation, a *quadtree decomposition* is applied. Each video frame is divided into four equal-sized square blocks, and then each block is tested to see if it meets some criterion of homogeneity. If a block meets the criterion, it is not divided any further. If it does not meet the criterion, it is subdivided again into four blocks, and the test criterion is applied to those blocks. This process is repeated iteratively until each block meets the criterion. The used criterion is the local variance of the block. In this way each smooth region of the image will be approximated with an unique block whose value is the mean color of the area. This step speeds up the overall process and in the same time reduces the possibility of false detection due to isolated pixels.

In order to find the similarity between colors we need to use a distance metric, that is, for each block of the decomposed image, we have to evaluate the difference between the color of the block and the color we are looking for in the video. Many distance metrics have been proposed based on the existing color models taking into account the way the HVS perceives colors. The Euclidean distance is frequently used in perceptually uniform color spaces, thus we decided to use this metric where the threshold of the volume bounding the color is set experimentally.

In football (as well as in other sports), it is common to find teams with two or more official colors. The proposed system is designed to deal with the problem of multiple colors pattern. In this case, the described framework is firstly applied for each color present in the pattern, after that only areas with the official colors close to each other are considered as detected regions.

III. Experimental results

A statistical analysis of the performances of the described system may be difficult due to the lack of videos recording football fans of different teams. Thus we chose to use synthetic images to test the proposed approach. We created a database



(a)

(b)

(c)

Fig. 2. Synthetic pattern generation: (a) original pattern; (b) blur; (c) change in illumination.

with all the colors of the football teams of the national French Ligue 1. For those teams who have two colors as official colors, we saved the corresponding bicolor patterns. We then generated 100 random patterns consisting of 16×16 blocks each describing a team of the French Ligue. In order to simulate the distortions that affect videos recorded in real conditions we then apply to each of the generated patterns different kinds of distortions: addition of Gaussian noise, blur, jpeg compression, changes in illumination. The parameters of the distortions (i.e: the strength noise, the blur factor, the jpeg quality factor) are each time randomly selected in a predefined range in order to have different levels of distortions ranging from high quality to very annoying distortions. This mechanism allows us to introduce randomness in the system and thus to obtain artefacts simulating as much as possible real conditions. An example of a generated pattern and two corresponding distorted versions are shown in Fig.2.

We tested the proposed approach by applying the color

	TPR	FPR	Accuracy
original	0.7395	0.0150	0.9721
blur	0.3448	0.0152	0.9512
noise	0.6836	0.0151	0.9691
illumination	0.3733	0.0207	0.9474
jpeg	0.5750	0.0179	0.9607

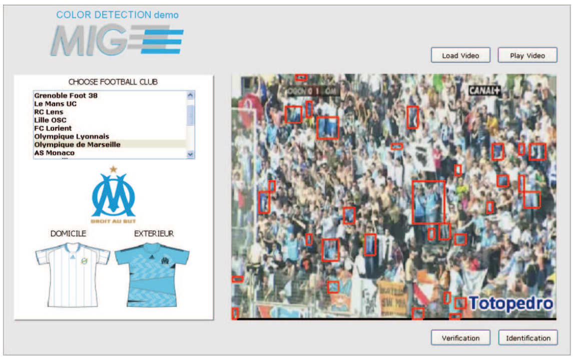
TABLE I
PERFORMANCES OF THE PROPOSED APPROACH.

detection algorithm to all the generated images for all the teams. The overall performances are obtained by averaging the results we obtained for all the teams. Table I shows the values of the true positive rate, the false positive rate and the accuracy of the system. We can observe that the accuracy of the proposed framework is high even in case of very annoying distorions. As expected, the distortions, and in particular the blurring of the images, drastically decrease the values of the true positive rate while maintaining a low false positive rate. It should be pointed out that the generate images are much more complex than in reality. All the colors of the teams are in fact mixed together and the closeness of the colored patterns may create new colors leading to errors in the system. A decreasing of the performances is also due to the use of the same official colors for different teams (such as Boulogne and Nice). While it is not possible to distinguish between two teams with the same colors, in a real application for which the proposed system is intended for, maximum two teams at time are compared reducing the probability of mismatches.

An example of a real application is shown in Fig.3. The figure shows a frame of a video clip realized during a football match Nice versus Marseille and recording the supporters of the Olympique de Marseille. In the screenshot on the top (Fig.3(a)) we apply the proposed color detection algorithm to look for the color of the Marseille shirt (as appears in the demo realized with a Graphical User Interface of Matlab). The red boxes in the frame show the presence of the fans in that area of the stadium due to the colors of their shirts or accessories (like hats, flags, etc.). Even if we can observe some miss detections, the presence of the Marseille fans is well detected. Fig.3(a)), instead, shows the proposed framework applied to the same video clip to look for the colors of the Nice shirt, the bicolor pattern black/red. As expected, the system does not reveal the presence of any Nice fans in that area of the stadium, excluding the possibility of a potential conflicts. We tested the system with several video clips with the colors of different clubs of the French Ligue 1, for both single color and bi-colors pattern, and the system was always able to identify the presence of the football fans.

IV. CONCLUSIONS

In this paper we exploit the use of color based soft biometry to identify the presence of rival sport fans in a specific location in order to avoid vandalism and destructive behaviour. The described system is robust in presence of different illumination condition and different video sensors. The proposed work is an efficient tool for hooligans detection and, more generally, can be seen as a step for the development of a robust video surveillance system for security applications.



(a)

(b)

Fig. 3. Color detection algorithm applied to reveal the presence of the Marseille fans (a) and the Nice fans (b).

ACKNOWLEDGMENT

The research described in this paper has been partially funded by the FR OSEO BIORAFALE project.

REFERENCES

[1] R. Cappelli, D. Maio, D. Maltoni, J. Wayman, and A. Jain, "Performance evaluation of fingerprint verification systems," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 1, pp. 3–18, Jan. 2006.
[2] F. Perronnin, J.-L. Dugelay, and K. Rose, "A probabilistic model of face mapping with local transformations and its application to person recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 7, pp. 1157–1171, July 2005.
[3] A. Jain, S. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," *Lecture notes in computer science*, pp. 731–738, 2004.
[4] R. Schettini, G. Ciocca, and S. Zuffi, "A survey of methods for colour image indexing and retrieval in image databases," *Color Imaging Science: Exploiting Digital Media*, pp. 183–211, 2001.
[5] D. Hilbert, "Color constancy and the complexity of color," *Philosophical Topics*, vol. 33, no. 1, p. 141, 2005.
[6] G. Wyszecki and W. Stiles, "Color science: concepts and methods, quantitative data and formulae," 2000.
[7] B. Berlin, F. Birren, J. Gage, J. Gatto, T. Hoban, and H. Kuppers, "Basic color terms; their universality and evolution [by] Brent Berlin and Paul Kay," 1993.
[8] J. Keller, "A Fuzzyˆ-Nearest Neighbor Algorithm JAMES M. KELLER, MICHAEL R. GRAY, and JAMES A. GIVENS, JR." *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 4, p. 581, 1985.