Synthesis of supervised classification algorithm using intelligent and statistical tools

Supervised Classification

Ali DOUIK
U.R.: Automatique Traitement de Signal et Image (ATSI)
Ecole Nationale d'Ingénieurs de Monastir
Monastir – TUNISIE
Ali.douik@enim.rnu.tn

Mourad MOUSSA JLASSI
U.R.: Automatique Traitement de Signal et Image (ATSI)
Ecole Nationale d'Ingénieurs de Monastir
Monastir – TUNISIE
Mourad.enim@yahoo.fr

***Abstract*— A fundamental task in detecting foreground objects in both static and dynamic scenes is to take the best choice of color system representation and the efficient technique for background modeling. We propose in this paper a non-parametric algorithm dedicated to segment and to detect objects in color images issued from a football sports meeting. Indeed segmentation by pixel concern many applications and revealed how the method is robust to detect objects, even in presence of strong shadows and highlights.**

In the other hand to refine their playing strategy such as in football, handball, volley ball, Rugby..., the coach need to have a maximum of technical-tactics information about the on-going of the game and the players. We propose in this paper a range of algorithms allowing the resolution of many problems appearing in the automated process of team identification, where each player is affected to his corresponding team relying on visual data. The developed system was tested on a match of the Tunisian national competition. This work is prominent for many next computer vision studies as it's detailed in this study.

Keywords-component; Soccer Singular value decomposition; Classification; artificial intelligence; supervised algorithm; Moments Matri

I. Introduction

In the last ten years, motion detection and analysis have become very important for a wide range of applications, especially since complex algorithms can nowadays be processed real-time. Examples of applications that use motion segmentation techniques are gesture recognition, tracking applications [1, 2, 3, 4, 5], video surveillance systems [6, 7], industry, robotics [14], the medical field [15], aeronautics [17], Pattern Recognition [13] and recently, sports sector [18]. Although a lot of research has done in this field on objects segmentation, still a lot of difficulties have to be considered in this area, especially to produce good results in changing circumstances. The main purpose of this paper assignment is to present an overview of objects segmentation techniques and classification.

A large variety technique has developed and improved, K. Karman et al. used Kalman filter to model a dynamic background. Similarly K. Elgammal et al. [9] presented a non-parametric background model to model dynamic background. Toyama et. al. [10] used Wiener filter to make a linear prediction of the pixel intensity values, given the pixel historic. C. Wren et. al. [11], use a single Gaussian model per pixel and the parameters are updated by alpha blending. Unfortunately, these approaches fail in case the distribution of the background colour values do not fit into a single model. Ying Ming et al. [12] worked out a statistical algorithm inspired from the idea of Elgammel based on Cauchy distribution; they proved that ratios of intensity values between the background pixels and the current image pixels are adapted to Cauchy's distribution. In fact it is characterized by a little wide form covering the tails of the histogram; on the other hand Gaussian distribution has an exponential form.

Several works was done concerning classification, Pal et al. [21] proposed an SVM technique, their work reports the results of two experiments in which multi-class SVMs are compared with Maximum Likelihood (ML) and Artificial Neural Network (ANN) methods in terms of classification accuracy; SVM achieves a higher level of classification accuracy than either the ML or the ANN classifier.

Classification by artificial vision in soccer sector has been largely mediatized and became a significant research topic. The result of a match has serious consequences on the club life and its external environment (media, sponsors...). To refine their play strategy [20], coach and the leaders need to have technical-tactics and relevant information [19] about events of the play as well as of the players. Indeed the use of the color in computer vision application is yet very recent, musical field [22], metals classification [23], road scenes analysis and sensing domain [16, 29, 8]. In this paper various supervised classification techniques were applied. They are based on intelligent tools as fuzzy and neuronal classification on the one hand, statistic and hybrid classification based respectively on moments difference and determination of three significant color components on the other hand. A comparative study about player recognition rates was elaborated enabling us to conceive an adequate method for football players classification at the aim to classify each player in his suitable class automatically.

II. SEGMENTATION TECHNIQUES EASE OF USE

A. *Detection by histogram analysis*

In artificial vision field, colour images are taken by a video camera and then digitized by a computer. Since the soccer video is taken by static camera, the supporter and useless information can be removed by delimiting the playing zone with an affine function. Objects identification in colour images is a relevant stage in classification study; therefore we begin by the background subtraction to detect players and then to classify each of them in its suitable class. The algorithm is based on the following stages:

- Convert the original image to standards rgb levels removing the light reflections.
- Detect the high and low thresholds of each histogram then carry out the threshold on the three chromatic levels.
- Apply a logical operator "AND" on the three levels and the original images.

B. *Detection by statistical learning*

Detection by histogram analysis consists to segment colour images and to remove useless information that have no contribution in classification phase. This method isn't a good choice of segmentation for many applications as presented in figure 2. The major problem of this technique corresponds at the time when the background and foreground have the same characteristics, hence after thresholding histograms many false detection can be occurred: for example it can remove a large part of useful information hence we should develop an appropriate technique.

Because the parametric background model still lacks flexibility when dealing with non-static backgrounds, a highly flexible non-parametric technique is proposed for estimating background probabilities from many recent samples over time using Kernel density estimation.

In the non-parametric model all recently observed pixel values $x_1, x_2 \ldots x_N$ are modeled by probability density functions using a certain kernel estimator function, which is often chosen to be a Gaussian. The weighted sum of all these Gaussians results in the final probability density function of the pixel value x_t:

$$p(x_t) = \frac{1}{n}\sum_{i=1}^{N} K(x_t - x_i) \qquad (1)$$

The kernel estimator function K is chosen to be a Normal function $N(0,\Sigma)$ with Σ being the kernel function bandwidth. For simplicity reasons the color channels within Σ are assumed independent, but each with their own kernel bandwidth σ_j^2. Because of these assumptions the final density estimation can then be written as:

$$\Pr(x_t) = \frac{1}{K}\sum_{i=1}^{k}\prod_{j=1}^{d}\frac{1}{\sqrt{2\pi\sigma_j^2}}e^{-\frac{(x_{tj}-x_{ij})^2}{2\sigma_j^2}} \qquad (2)$$

When this probability is below a certain threshold, the pixel is classified as a foreground pixel. The threshold can be adjusted to achieve a desired proportion of false positives. In other kernel density estimation applications, the kernel bandwidth has to be dependent on the number of samples. When there are many samples the bandwidth should be smaller than if we have few one. However, in this case temporal properties are taken into account for determining the σ_j^2 of color channel j. For each color channel the median m of the absolute differences of each consecutive pairs of samples is calculated. We estimate σ by:

$$\sigma = \frac{m}{0,68\sqrt{2}} \qquad (3)$$

This method guaranties that the local deviation is large when there are many large jumps between consecutive samples and smaller when this is not the case.

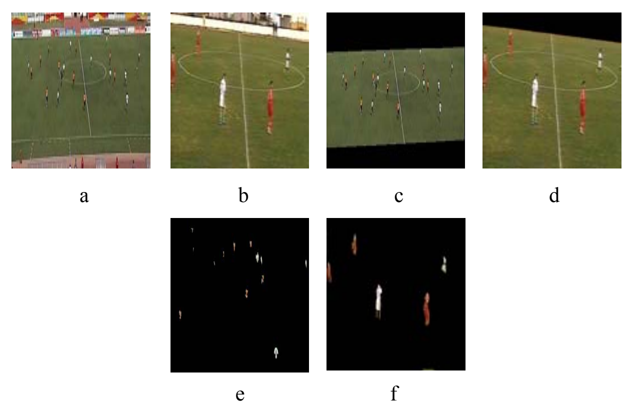


Figure 1. (a, b) Represent original images, (c, d) Represent delimited images, (e, f) Represent segmented images.

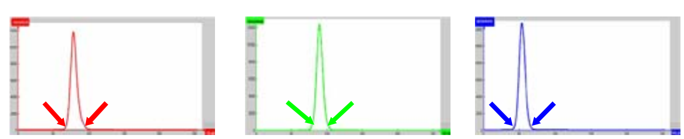


Figure 2. Histograms of standards RGB levels (r, g, b)

C. *Singular Value Decomposition Approach [24]*

1) *Introduction*

A non-parametric background modeling technique, has been applied on a soccer video images, the main problem that can be appears is the occurring of wrong detection pixels, indeed shadow pixels are detected as moving objects result to an over segmentation that will be damage, many later works where this paper it's registered, however this algorithm is extremely important because it's a part of players classification and tracking on a soccer video.

Two major issues in the SVD technique: (1) carrying out a mathematical approach and (2) explain main advantage of the method proposed here and showing influences of the singular values choice on a treated output image, besides we will see the prominent contribution using SVD theory to restore and eliminate shadow, highlights and noise from camera displacement and changed circumstances.

2) SVD approximation of an image

The main objective of background segmentation technique is to use singular value decomposition of a given image A represented by a matrix $A_p = [a_{ij}]$, when it can be decomposed into a product of three matrix $U_k S_k V_k^T$ as shown in figure 3. Where a_{ij} is the appearance frequency of chromaticity and intensity of background pixels (p= red, green, blue).

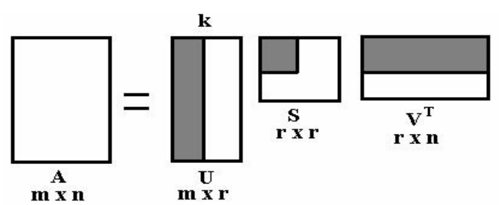


Figure 3. Singular values decomposition of the Matrix A.

SVD technique consists to reduce the size for each initial chromatic level from r to k rank by suppression of r-k column. Matrix S represents diagonally the singular values classified by decreasing orders. Low values have no influence on the total energy of A. U_k and V_k are orthogonal matrices issued from matrices U and V. The determined singular values for each plan were presented in frequency space. There representation proves that for each one corresponds a discrete frequency. The noise that can occur in the signal (in frequency space the amplitude of noise is constant) corresponds to a low amplitude of singular value whereas high amplitudes of these represents global signal energy.

3) Confidence intervals research

In this section, we describe the basic background model and the background subtraction process with singular value decomposition. It's both used in the restoration or the reconstruction of an image, to increase the compactness distribution of different classes and also to provide useful image information.

To evaluate mathematical contribution of SVD, a quantification of global signal energy distribution according to the weight of each singular value S_{kk} was done. The figure 4 illustrates the energy distribution E defined by:

$$E = \sum_{i=1}^{k} A_k^2 \tag{4}$$

The relative energy contained by each singular value K, noted p_k is defined by:

$$p_k = \frac{S_{kk}^2}{E} \tag{5}$$

Where the energy of the K singular value is equal to S_{kk}^2.

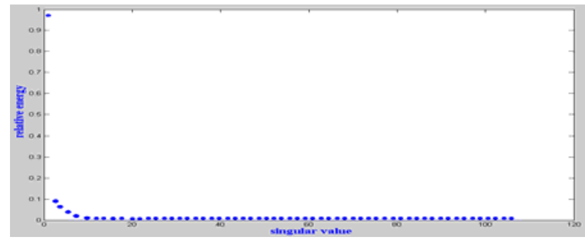
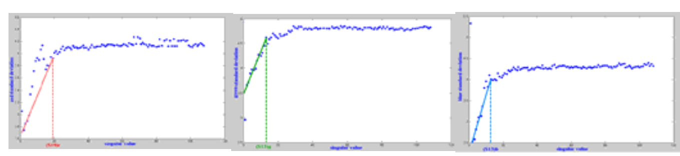


Figure 4. Weight distribution showing 7 dominant representing 99 % of the signal energy among 108 weights

As it's shown in figures 5a, 5b and 5c, the size of treated image will be deduced from the curves representing standard déviation of each colour levels according to the singular value decomposition. In fact a good choice of the size leads to reduce both compactness in different distributions and in computing time. According to figures 5a, 5b and 5c, we can denote two zones, the first one defined in the interval $\left[0,(S_{kkl})_i\right]$ where $(S_{kkl})_i$ is the singular value limits corresponding to the linear part of the curve (i = red, blue, green), in this zone the curve presents a slope, beyond $(S_{kkl})_i$ a second zone appears where the standard deviation varied slightly therefore the optimal singular value $(\hat{S}_{kkl})_i$ must necessarily belong to the first zone of each curve.



a b c

Figure 5. Evaluation of standard deviation according to singular values respectively the red, green and blue channel.

Table1 illustrates initial and improved standard deviations for three channels (RGB).

TABLE I. EVALUATION OF IMPROVEMENT PARAMETERS

R	4.2044	119.61	1156.9
G	4.7227	152.08	1073.3
B	4.313	88.988	1141.7
R_SVD	3.9274	119.07	1229.8
G_SVD	4.6204	152.05	1104.2
B_SVD	3.9954	88.407	1205.5

The choice of singular values will be kept depending on two issues: the first one is the energy curve evaluated by figure 4 and the second one is the standard deviation curves of each chromatic level shown in figures 5a, 5b and 5c. In fact we specify for each component the singular value limit previewed. The table 2 shows confidence intervals as well as the limit and the optimal singular values that vary from a level to another.

TABLE II. SPECIFICATION OF CONFIDENCE INTERVALS

	(S_{kkl})	$(\hat{S}_{kk})$	**confidence intervals**
Red plan	29	19	[0, 29]
Green plan	28	13	[0, 28]
Blue plan	19	13	[0, 19]

4) Foreground segmentation and shadow suppression

Using the probability $\Pr(x_t)$ calculated in equation 2, the pixel is considered as a foreground pixel if $\Pr(x_t) < th$. The threshold th is a global threshold over all the image that can be adjusted to achieve a desired percentage of false positives. The shadows detection as foreground regions is a source of confusion for subsequent phases of analysis. Color information [18] is useful for shadows suppression by separating color from lightness information. For a given three color variables, R, G and B, the chromaticity coordinates r, g and b(r=R/(R+G+B), g=G/(R+G+B), S=(R+G+B)/3). Consider the case where the background is completely static, and let the expected value for a pixel be (r_i, g_i, s_i) . Assume that this pixel is covered by shadow and let (r_t, g_t, s_t) be the observed value for this pixel at this frame. Then it is expected that $th_1 < s_t / s_i < th_2$.

To prove robustness of this algorithm, the figure 6 illustrates different player windows and shows how we can overcome segmentation (removing shadow pixels and pixels background having the same characteristics that those foreground pixels).

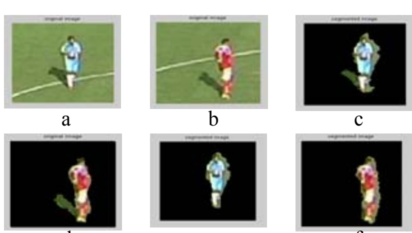


Figure 6. Effect of the singular value decomposition level. Where (a and b) are Original images, (c and d) Detection using chromaticity coordinates r, g and the lightness variable s. (e and f) Detection using chromaticity coordinates r, g and s with SVD.

III. CLASSIFICATION ALGORITHMS

A large variety of supervised classification algorithms was developed, ranged from statistical to intelligent tools and they operate on only color regions. Each one of them is expressed in adapted color system representation that will contribute to an optimal classification.

A. Hybrid classification

The advantage of this algorithm is that the color will be represented in a system of the three most discriminating levels, in order to be able to separate the colour nuance distributions corresponding to pixel players from each team.

1) Hybrid colour system

After extraction of useful information which represents the pixel players, we separate the two classes using colorimetric analysis. Nevertheless traditional RGB space cannot be the most discriminating representation space. Indeed, other colour systems, deduced from the RGB components, can be more suitable according to the considered case. For this reason, treatment of pixel players in various colour systems leads to a hybrid space represented by the three best colour components.

2) Method description

N.Vandenbroucke [29] considered a multidimensional space composed of the chromatic levels currently used as the following:

E = {R, V, B, r, v, b, X, Y, Z, I1, I2, I3, y, i, q, u, v, l, t, s}. In each level α (α∈E) the algorithm of discrimination is based on various phases:

- Select three training player windows J_{1A}, J_{2A} and J_{1B} in the RGB system.
- Convert in the α level player windows.
- Calculate the average of pixel coordinates (x, y) representing each player:

α (x, y): pixel value (x, y) in the plan α.

R_{1A} and S (R_{1A}) are respectively area and surface of the player J_{1A}.

R_{2A} and S (R_{2A}) are respectively area and surface of the player J_{2A}.

R_{1B} and S (R_{1B}) are respectively area and surface of the player J_{1B}.

Therefore we can evaluate for each level α:

α-average region for J_{1A}: $$m^{\alpha}_{1A} = \frac{\sum_{R_{1A}} \alpha(x,y)}{S(R_{1A})}$$

α-average region for J_{2A}: $$m^{\alpha}_{2A} = \frac{\sum_{R_{2A}} \alpha(x,y)}{S(R_{2A})}$$

α-average region for J_{1B}: $$m^{\alpha}_{1B} = \frac{\sum_{R_{1B}} \alpha(x,y)}{S(R_{1B})}$$

- Calculate the distance between J_{1A} and J_{2A} as well as the distance between J_{1A} and J_{1B}