❏     2828

# Face Detection in a Mixed-Subject Document

**L houssaine Bouhou[1], Rachid El Ayachi[2], Mohamed Baslam[3], Mohamed Oukessou[4]**
[1,2,3]Computer Sciences Departement, Faculty of Science and Technology, Sultan Moulay Slimane University, Morocco
[4]Mathematics Departement, Faculty of Science and Technology, Sultan Moulay Slimane University, Morocco

## Article Info

## ABSTRACT

Before you recognize anyone, it is essential to identify various characteristics variations from one person to another. Among of this characteristics, we have those relating to the face. Nowadays the detection of skin regions in an image has become an important research topic for the location of a face in the image. In this research study, unlike previous research studies related to this topic which have focused on images inputs data faces, we are more interested to the fields face detection in mixed-subject documents (text+images). The face detection system developed is based on the hybrid method to distinguish two categories of objects from the mixed document. The first category is all that is text or images containing figures having no skin color, and the second category is any figure with the same color as the skin. In the second phase the detection system is based on Template Matching method to distinguish among the figures of the second category only those that contain faces to detect them. To validate this study, the system developed is tested on the various documents which including text and image.

*Corresponding Author:*

Rachid El Ayachi,
Computer Sciences Departement,
Faculty of Science and Technology,
Sultan Moulay Slimane University, Beni Mellal, Morocco.
Email: rachid.elayachi@usms.ma

## 1.  INTRODUCTION

Among the current problems, the knowledge to determine the identity of a person automatically. Passwords and ID cards cannot provide vital authentication functions such as detection of multiple entries. For example, users can easily deny the use of a service by claiming that their password has been stolen or guessed. Individuals can also hide their true identity by presenting falsified identity documents duplicates.

Therefore, it becomes increasingly evident that these mechanisms are not sufficient to determine reliably the identity of a person and a more robust mechanism for identification based on something that you are, ie biometrics, is more than necessary.Biometric features are an alternative to former identity verification means. The advantage of these biometric is to be present in all people to identify. Face detection in a hypothesis without image is a very difficult problem because of the great variability of the shape and color to detect (Orientation and any size, variable lighting, different skin color).

## 2.  BIOMETRICS

Biometrics is an alternative to the old methods of identification. It consists of identifying a person from physical or behavioral characteristics [1]. Face, fingerprint, iris, etc. are examples of physical characteristics. Voice, writing, typing rhythm on a keyboard, etc. are behavioral characteristics. These features, as they are innate or acquired digital fingerprint as signature, are attached to each individual and therefore does not suffer from the weaknesses of the methods based on knowledge or possession [2].

## 2.1. Architecture Of A Biometric System

A biometric system is often composed of three modules: learning, recognition and adaptation [3-4]. During learning, the system will acquire one or more biometrics will be used to build a model of the individual. This reference model will serve as a point of comparison during recognition. The model will be reevaluated after each use with the adapter module.

### 2.1.1. Learning Module

During the phase of learning, biometric characteristic is measured by a sensor; there is talk of acquisition or capture. In general, this capture is not directly stored and transformations are applied to it. Indeed, the signal contains useless information to the recognition and only the relevant parameters are extracted. The model is a compact representation of the signal which facilitates the recognition phase, but also to reduce the amount of data to store.

### 2.1.2. Recognition Module

During the phase of recognition, the biometric characteristic is measured and a set of parameters is extracted as when learning. The sensor used must have properties as close as possible the sensor used during the training phase. If both sensors have different properties too, it will generally apply a series of additional pre-treatments to reduce performance degradation. Following the recognition will be different according to the procedure of the system.

### 2.1.3. Adaptation Module

During the learning phase, the biometric system often catch a few instances of the same attribute so as to limit the inconvenience for the user. It is difficult enough to build a general model to describe all the possible variations of this attribute. In addition, the characteristics of this biometrics and its acquisition conditions may vary. Adaptation is therefore necessary to maintain and improve the performance of a system use after use. The adjustment can be done in supervised mode or unsupervised but the second mode is by far the most useful in practice [5].

Among the biometric techniques, there are facial recognition. In everyday life everyone identifies us throughout the day different faces. So when we encounter a person, our brain will search our memory and see if that person is listed or not.

In recognition of 2D face, several methods have been developed. However, they have a number of limitations related to the orientation of the face or laying, lighting, facial expression, occlusions, etc. In recent years, we talk more and more 3D face recognition techniques as an alternative solution to solve the problems mentioned above. Indeed, the wealth of information provided by 3D measurements allows to reconstruct the three-dimensional shape of the face. This type of facial representation is invariant to changes in illumination and pose [6].

Facial recognition is a common and popular technique. It is the most acceptable because it corresponds to what humans use in visual interaction; and comparing it to other methods, facial recognition is more advantageous, on the one hand it is a non-intrusive method, that is to say it does not require the cooperation of the subject (observing individuals remote), and on the other hand the sensors used are inexpensive (single camera) unlike the fingerprint and iris where the subject will be very close to the sensor and will cooperate in the development of picture without forgetting the cost of the equipment necessary for the acquisition (expensive special equipment).

## 2.2. Face Detection: State Of The Art

An automatic face recognition system consists of three subsystems: face detection, feature extraction and face recognition. Face detection is the problem of locating faces in a input image. The interest of facial localization goes beyond the implementation of the present work. Its utility is manifested in various fields ranging from video surveillance to interactive game. The first difficulties encountered by harnessing methods to detect faces are the variations in pose (profile view, front), expression, face rotation, age and illumination. This type of difficulty can be overcome by pretreatment standards and illumination compensation.

In a controlled environment, settings such as the background, the direction and intensity of the light source, the angle of the shot, are parameters that can be controlled.

Yang et al., [7] proposed a classification of methods of facial location:

a. "Knowledge-based methods". Knowledge of the various elements that constitute a face and relations between them are of these methods basing. Thus, the relative positions of the various key elements such as the mouth, nose and eyes are then used to measures the classification 'face' and 'non-face' in Chiang et

al., [8]. The problem with this method is that it is difficult to define uniquely a face. If the definition is too detailed, some faces will be rates while if the description is too general the false positive rate soar.

b.    "Feature invariant approaches". For detection, these approaches use the elements invariant to changes in illumination, orientation or expression, such as the texture or the signature color of the skin.

c.    "Template matching methods". The characteristic patterns of an entire face or face of subpart (mouth, eye, nose) are created. The location is then made based on the correlation of these models with the candidates.

d.    "Appearance-based methods". These methods use the same principle as the one presented in the previous paragraph but are based on specific models. These methods have the advantage of running very quickly but require a long training time. The methods in this class have shown good results against 3 other methods [9].

The face detection methods can be divided into four categories [6]:

a.    Approches Based on knowledge: the main parts to what this method is interested in their characteristics as facial features are the nose, mouth and eyes. These methods are designed primarily for face localization.

b.    Approches Based "Template-matching": These methods to calculate the correlation between the candidate image and the template. They use an algorithm that calculates and luminance ratios between the areas of the face and retains leadership of these reports (eg, is the region 1 lighter or darker than the region 2). Figure 1 shows a predefined template corresponding to 23 relationships.
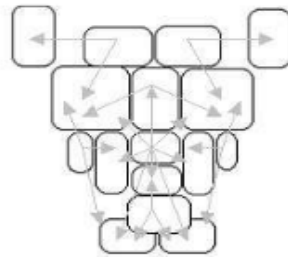


Figure 1. Face model consists of 16 regions (rectangles) associated with relationships 23 (arrows) [10]

d.    Approches Based on appearance: Methods typically use machine learning techniques. They are used for detection. The main idea of these methods is to consider that the problem of face detection is a classification problem (face, non-face). One of the most known approaches face detection is the Eigen face, it involves projecting the image in space and calculate the Euclidean distance between the image and its projection.

e.    Approches Based on invariant features: These approaches are used primarily for face localization. Developed algorithms aim to find the existing structural characteristics even if the pose, the viewpoint, or the lighting condition change. Then they use these invariable features to locate faces. We can mention two families of methods belonging to this approach: Methods based on skin color. In fact, they reduce the search space of the face region in the image. In addition, the color of the skin is a robust information face rotations, scale changes and partial occlusions. Several color spaces can be used for detecting in the image, the pixels which have the color of the skin. The detection efficiency depends mainly on the color space selected. Among the color spaces used, the RGB color space, it is to represent the color space from three monochromatic radiation of colors: Red-Green-Blue. For example, the skin is classified into the RGB color space by using the following rules [11]:

$$R > 95 \; et \; G > 40 \; et \; B > 20 \; et \; |R - G| > 15 \; et \; R > G \; et \; R > B \qquad (1)$$

We present here a comparative study on skin color detection by thresholding according to the RGB color spaces. Step thresholding provide us binary images whose skin sections take the highest gray scale or 1 (white) while other sections (not skin) take the lowest gray level is 0 (black).

### 2.3. Problematic

In general, a face recognition system consist of three parts: Detection, Extraction and Recognition. In this paper, our good is to make a contribution in the first part. This contribution allows the development of the first part of the system for detecting faces in mixed document. For information, a mixed document contains several items, including: text, figure and image.

## 3.     SCENARIO

### 3.1. Defining the Objective Enumeration and its Consequences

The aim of the application is to allow the user which to detect the faces of people in a mixed document containing text and image. A scenario like the use of this system in three phases using two methods.

*1st phase segmentation skin areas.

Based on the Hybrid Skin color method, this phase involves cutting skin color representative pixels from a mixed input document and filters independent skin color pixels.

* 2nd phase: segmentation regions containing face

Based on Template Matching method this phase focuses on removing skin color regions not presenting a face.

*3rd phase: Face Detection

In this phase face detection is established by matching the template "Template Matching" who is interested in comparing the intensity of pixels between a predefined template and several sub regions of the image to be analyzed. This process is in practice to perform multiple scans covering the entire area of the image. The most favorable to the presence of faces places will therefore be easily identified by minimum distance between the template and the underlying image.

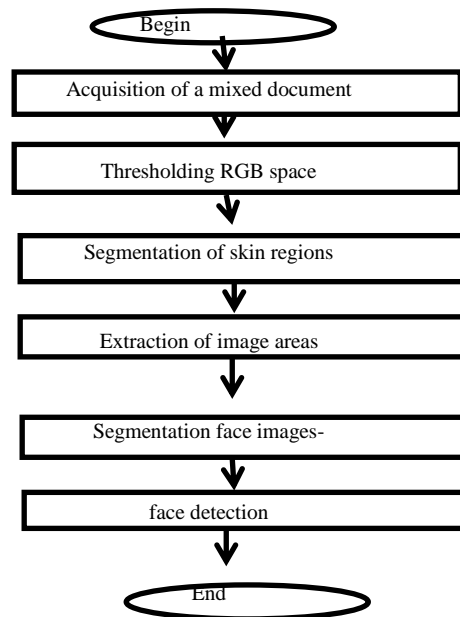Figure 2 shows the detailed process of the various phases listed disappointed:



Figure 2. Process Reflects the General Objective of the Developed System.

### 3.2. Detection of Skin Color Sections

Several color schemes that can be applied to the skin detection this variability depends on the color space adopted for the representation of pixels. The most commonly used models are: RGB space, normalized RGB, HSV, YCbCr. In this work we have chosen RGB space that reflects the physiology of the eye. Indeed the human eye distinguish colors through receptors called cones. There are all in all three types of cones, which means that any color discernible by man is characterized by a point in a three-dimensional space.

Specifically, the values measured by the three types of cones essentially represent the intensities of red, green and blue, RGB space being defined from the three primary colors. His coding was developed in 1931 by the International Commission on Illumination (CIE) [12].

The skin is shown in the majority of cases, by a portion of an area of particular colors. To isolate the areas likely to contain faces, using the borders of these regions as threshold values on an image, it is possible to extract pixels whose color can be matched to that of the skin [13]. Indeed, for detecting skin regions, a threshold S characterizing the color of the skin being defined, and to detect skin color pixels is the rule:

$$\text{if } x \text{ (pixels)} < S \rightarrow x = 1 \quad \text{else} \quad x = 0 \tag{2}$$

Figure 3 shows an example set by our system that reflects the conversion of a mixed color document (text+images) into two classes of pixels or the white pixels which represent the color of the skin and the black pixels representing the color not skin.
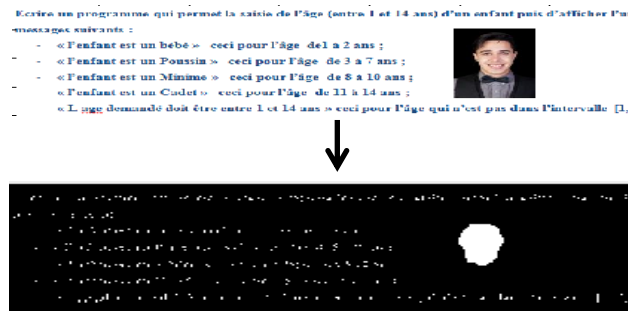


Figure 3. Sample Binary Converts Mixed Document Developed by our System,
White Pixels=Skin Color-Black Pixels=Not Skin Color.

### 3.3. Filtering Independent Skin Color Regions

During this stage, treatment is applied to the document resulting from the previous phase, then used a method of segmentation of the white areas which is to browse through a window, all pixels of this document. Along the way, each window is calculated the report $\frac{PB}{PN}$ (PN: The number of black pixels and PB: number of white pixels) to be compared to a threshold s to deduct if the content of the window is or is not a party an image

$$\frac{PN}{PB} \leq s \quad \rightarrow \quad converting\ pixels\ of\ the\ window, to\ black \tag{3}$$
$$\frac{PN}{PB} > s \quad \rightarrow \quad do\ nothing\ for\ all\ pixels\ of\ the\ window$$

Thus the pixels not checking rule (3) are forced to take the black despite their initial white color, consequently only pixels, white contained in a skin color image are white. The figure 4 shown below illustrates the result obtained by applieng the filtring processing on Figure 4.



Figure 4. Example of Filtered Document by the Method Described in this paragraph

### 3.4. Extracting Images Skin Color and Face Detection Regions by Template Matching

During this step, a scan to locate the areas with white pixels. The difficulty of this phase lies with the fact that the input acquired document can contain multiple skin color images (Figure 5). In this case, the process consists in detecting the first image whose pixels are white, undergo her treatment involving template matching method to detect whether it contains or not a face, then iterate the same processing to detect the presence or absence face in the other regions whose pixels are white.
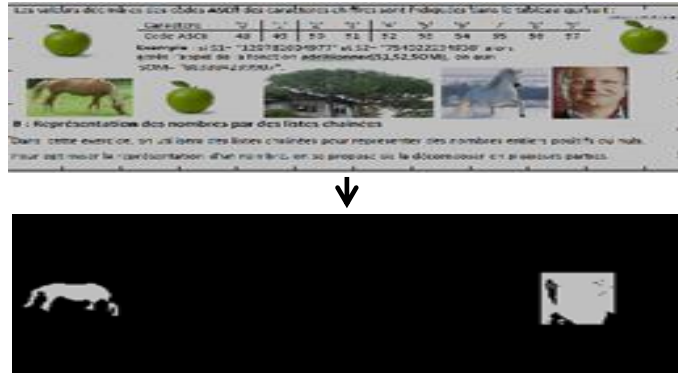
Figure 5. Example of Location of Both Skin Color Images in a Mixed Document

The good of Template Matching method is to scan each area among the detected areas, and comparing the intensity of pixels between a model template (template) predefined (in our case Figure 6) and several sub-regions of the area to be analyzed. This process is in practice to perform multiple scans covering the entire area of the zone. The most favorable to the presence of faces places will therefore be easily identified by minimum distance between the template and the underlying area. Among the possible distances, there are particularly the norm L1 (Manahan distance) and the L2 norm (Euclidean distance), cross-correlation, etc.

$$\text{Norme L1:} \quad \sum_{(i,j)} |a_{(i,j)-} \ b_{(i,j)}| \tag{4}$$

$$\text{Norme L2:} \ \sqrt{\sum_{(i,j)} a^2_{(i,j)} - \ b^2_{(i,j)}} \tag{5}$$

where $a_{(i,j)}$ and $b_{(i,j)}$ are the value of the pixel $(i, j)$ in the two images to compare. It therefore follows that the more these values are closer, the more images are similar [14].

We present the following (Figure 6), the template that was used to detect, for our system, if an image contains a face or not.



Figure 6. Model Template Used by our System to Detect Whether or not Image Contains a Face
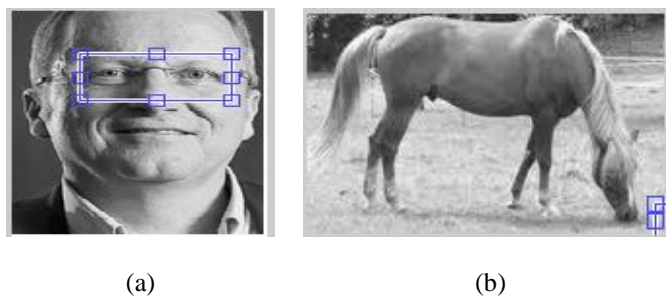


(a)                                            (b)

Figure 7. Face Detection Method by Template Matching

By setting a threshold for the detected position of the segment that is most similar to the template, we can judge this segment contains a face or not. Figure 7 is an example that illustrates this process, in fact,

part (a) of this figure reflects the position of the template on the face by cons in part (b), the position of the template is beyond the threshold consequently the image contains no face.

The pixels of the image that contains no face will then be delivered to the lowest gray level is 0 (black), consequently only the pixels of the image containing a face (Figure 8) which will be in gray scale 1 (White).

Figure 8. Segmentation of the Image Containing a Face

To detect face, first, we need to extract the binary image of the document. We then try to locate the face area, based on the position of the template identified in the preceding paragraph as an initial benchmark. The result is a gray scale image that represents the face to detect in question (Figure 9).

Figure 9. Extraction of the Face

## 4.    RESULTS AND DISCUSSIONS

We conducted some tests on text documents which were inserted images we have downloaded from the Net. Tests are performed on images that come in different colors and presenting figures that may or may not be people's faces. We note in the majority of the tests (85% of tests) positive results for both face images that the images not faces. Among the false alarms, testing on an image containing a face of a child or a baby (Figure 10). This false alarm is due to the deviation of the similarity between the template and each window under the image of the input document when scanning in search of the segment of the image most similar to the Template.

Figure 10. Example False Alert Upon Detection of the Face

## 5.    CONCLUSION

The detection system that we have developed and introduced in this article is a preliminary stage and of great importance in any face recognition system as before launching the system to recognize individuals, verifying the presence or not of a face in the scene can be very beneficial. This step will allow the system to divert the recognition operation if the scene contains no face.

Our application is based, firstly, on the hybrid method that performs segmentation of two classes of zones, the first classes for areas, present in the scene, the color being that of the skin and second categories for areas, present in the scene, the color of which differs from that of the skin, the color representation space selected is the RGB space. And secondly, our application is based on Template Matching method to detect the presence or not in the stage of a face.

The results are satisfactory because they allowed us to increase the performance of the method color detection and the matching of the template used separately. These results are significantly effective in comparison with those made during the face detection phase in our own previous research [15]. However, In order to better get around the problem of face detection, we propose, as a short term perspective, to try, in face detection phase, other color spaces such as RGB and normalized space HSV space and compare their results with the results of the RGB space and try, in face recognition phase, to improve the methods used in [15] tennant account new method of extraction and analysis non-linear characteristics (EANF) [16].

## REFERENCES

[1]    A.K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition". *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 4 – 20, Jan 2004.
[2]    Florent PERRONNIN, Jean -Luc DUGELAY, "Introduction à la Biométrie Authentification des Individus par Traitement Audio-Vidéo", *Revue Traitement du Signal*, Volume 19 –no 4, 2002.
[3]    S. Liu, M. Silveman, "A pratical Guide to Biometric Security Technology", *IEEE Computer Society*, IT Pro-Security, Janvier-Février 2001.
[4]    A.K. Jain, L. Hong, S. Pankanti, "Biometrics: Promising Frontiers for Emerging Identification Market", *Communications of the ACM*, pp. 91-98, February 2000.
[5]    C. Fredouille, J. Mariethoz, C. Jaboulet, J. Hennebert, J.-F. Bonastre, C. Mokbel, F.Bimbot, "Behavior of a Bayesian Adaptation Method for Incremental Enrollment in Speaker Verification", *International Conference on Acoustics, Speech, and Signal Processing*, pp. 1197-1200, Istanbul, Turquie, 5-9 Jun 2000.
[6]    S. Guerfi Ababsa, "Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D", Doctoral thesis, UNIVERSITE D'EVRY VAL D'ESSONNE, october 2008.
[7]    Ming-Hsuan Yang, David J. Kriegman, and Narendra Ahuja. Detecting faces in images: A survey. *IEEE Transactions on pattern analysis and machine intelligence*, 24(1): 1746 -1762, 2002.
[8]    Cheng-Chin Chiang, Wen-Kai Tai, Mau-Tsuen Yang, Yi-Ting Huang, and Chi-Jaung Huang. A novel method for detecting lips, eyes and faces in real time. *Real-Time Imaging*, 9(4): 277-287, 2003.
[9]    Wenlong Zheng and Suchendra M. Bhandarkar. Face detection and tracking using a boosted adaptive particle filter. *Journal of Visual Communication and Image Representation*, 20(1): 9 - 27, 2009.
[10]   P. Sinha, "Object Recognition via Image Invariants: A Case Study", *Investigative Ophthalmology and Visual Science*, vol. 35, no. 4, pp. 1735-1740, 1994.
[11]   P. Peer, J. Kovac, F. Solina. Human skin colour clustering for face detection. In submitted to *EUROCON 2003 – International Conference on Computer as a Tool*, 2003.
[12]   Ch. Bencheriet, A.H. Boualleg & H. Tebbikh  & B. Guerzize & W. Belguidoum, "Détection de Visages par Méthode Hybride  Couleur de Peau et Template Matching", 4th *International Conference: Sciences of Electronic, Technologies of Information and Telecommunications*, March 25-29, 2007 – TUNISIA.
[13]   Garcia C., Tziritas G., "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis", *IEEE Transactions on Multimedia*, 1(3), September 1999, p. 264-277.
[14]   G. Heusch: "Détection  automatique  de visages dans une séquence vidéo", SSC semestre 5 - 18 Février 2002.
[15]   L. Bouhou, R. El Ayachi, M. Fakir, M. Oukessou, "Recognition of a Face in a Mixed Document", *TELKOMNIKA Indonesian Journal of Electrical Engineering*, Vol 15 No 2, 2015 pages 301-312.
[16]   A. Mahdavi Hormat, K. Faez, Z. Shokoohi, M. Zaher Karimi, "The new method of Extraction and Analysis of Nonlinear Features for face recognition", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol 2 No 6, 2012 pages 766-773.