Classification of Human Skin Color and its Application to Face Recognition

Marwa Jmal, Wided Souidene Mseddi, Rabah Attia
Electronic Systems and Communication Network Dept,
TUNISIA Polytechnic School
marwa.jmal@telnet-consulting.com,
wided.souidene@ept.rnu.tn, rabah.attia@enit.rnu.tn

Anis Youssef
Head of Innovation Activities
TELNET Group, Tunisia
Technological Park El Ghazala,Tunisia
anis.youssef@groupe-telnet.net

***Abstract*—This paper presents a novel human skin color classification into skin color tones: White and Black. This is performed by developing a skin color classifier based on pixel-based classification using RGB model. Our proposed method is classified under the category of an explicitly defined skin region model. The skin classifier divides our database formed by some images from the FERET set of faces into two sub-databases according to the skin color. The skin color classification method is then applied on a face recognition technique by reducing the number of trained images in the matching process. The performance of the proposed human skin color classifier is evaluated perceptually. Experimental results showed that our proposed skin color classifier is able to classify a face into its possible skin color tone and reaches 87% as hit rate.**

Keywords-Skin color classification; Face recognition; Pixel-based classification; RGB model.

I. INTRODUCTION

Face recognition is a biometric technique, which aims to identify a person from a digital image by comparing its extracted facial features with the ones of images in database. This field has presented for the past decades, the center of extensive research and it was mainly used for security and access control. However, human face image is vulnerable to a lot of variations caused by aging, illumination changes, facial expressions and low resolution, which make it harder for face recognition techniques to acquire interesting discrimination results.

Several approaches of face recognition were developed as a solution to this problem. In this paper, we employ Lowe's Scale Invariant Feature Transform (SIFT) descriptors [6] [13] to detect facial features.

SIFT descriptors are known as the most local invariant feature descriptors. First, they were developed for object recognition systems and have become, recently, the core of many algorithms in computer vision applications. This method transforms an image into local feature vectors, which are invariant to image translation, scaling and rotation, and partially invariant to illumination changes and 3D projective transform. However, SIFT descriptors were designed only for gray images [9]. Thus, the color component in an image grants weighty information for object classification (animals, flowers, faces, etc.). For some sort of applications, such as face recognition, color may be an important distinction tool for discrimination and it has been proven that it is very salutary and robust for applications applied on faces (detection, tracking and recognition).

Human skin color classification finds out to which color tone the skin belongs. The simplest and most employed technique for skin modeling is to explicitly define skin region [12]. The advantage of this method is the simplicity of detection rules which leads to building a very fast classifier. Other skin modeling techniques employing statistical based approaches are involved such as neural networks [7], k-means clustering [3] and Bayesian networks [8].

Unlike it seems to be, skin modeling is complex and quite challenging. In fact, skin color in an image depends mostly on illumination conditions which affect the distribution model of the skin color. Other problems facing skin color classification are shade and shadow occlusions, resolution as well as skin tone variation between races.

The purpose of this study is to develop a skin color classifier into skin color tones in order to improve the face recognition results based on SIFT descriptors.

The rest of the paper is organized as follows: Section 2 is dedicated to the human skin color classification in which our proposed method is detailed. The application of our classifier to face recognition using SIFT descriptors is presented in Section 3. Experimental results are covered in Section 4. In Section 5, the conclusion is drawn.

II. HUMAN SKIN COLOR CLASSIFICATION

In general, the purpose of this study is to enhance a human skin classifier that is able to effectively classify skin color tones. To reach this objective, this section will present the pursued methodology which is divided into two steps: skin segmentation and skin color modeling.

A. Skin segmentation

Precise skin segmentation aims to remove all "non-skin" pixels in order to acquire good results in skin classification. Each image in the database was segmented according to the method proposed in [1]. In this article, a skin segmentation scheme based on RGB (Red, Green, Blue) pixels' color is developed. The model presented in [1] is divided into two rules. A pixel is considered as skin if:

$$0.0 \leq \frac{R-G}{R+G} \leq 0.5 \qquad (1)$$

 ISBN: 978-1-61208-320-9

and

$$\frac{B}{R+G} \leq 0.5 \tag{2}$$

The RGB values of skin pixels were preserved as they are, while the RGB values of non-skin pixels were mapped to [0 0 0]. Fig. 1 exemplifies the skin segmentation process to preserve only skin pixels.

(a) (b)

Figure 1. Example of skin segmentation: (a) Original image, (b) Mask image.

This process excludes different parts like eyes, hair and accessories. Thus, it is not totally effective since it confuses, for example, skin with bright hair such as in Fig. 1 where some parts of the hair were not removed.

This method outputted a mask image (Fig. 1 (b)) and stored it for further applications.

B. Skin color modeling

The fundamental goal of skin color modeling is to create a decision rule that will distinguish between skin color tones. To achieve this objective, this section will describe the methodology which is divided into three main steps as follows:

- Selection of the color space,
- Choice of skin color tones,
- Creation of the decision rule.

1) Selection of the color space: The most important step is to select the color space in order to acquire more accurate classification results. Many color spaces have been involved in the problem of skin color representation and recognition such as RGB, YCrCb (Luminance, Chroma: Red, Chroma: Blue) and HSV (Hue, Saturation, Value). Shin et al. [4] suggest a comparison of the performance of eight color spaces for skin detection. As a conclusion, RGB and YCrCb were the best classified color spaces when dealing with separability. Based on this result, the RGB color space will be used for skin color mapping.

2) Choice of skin color tones: In general, skin color tones are white, yellow, brown and black. In [10], the skin tone set was classified as white, brown and black since the yellow had its tone too close to the white one. However, experiment results had shown that the highest incidence of error was found in the brown skin color set: almost half of brown colored skin was misclassified. From this result, in the current paper, we present a new method for skin color classification into white and black tones based on the RGB model. Brown tone is classified under the black set.

3) Creation of the decision rule: The human skin color classification technique is derived from histograms. In fact, histogram-based segmentation approach is an efficient method for image segmentation given its rapidity in training. Moreover, Vezhnevets, Sazona and Andreeva [2] revealed that this approach is independent from the shape of skin distribution. A new developed RGB ratio histogram was plotted to elicit new threshold for skin color tones. This ratio was formed by mixing RGB values in order to define new colors. It is defined as follows:

$$\text{Ratio:} \qquad \frac{B-G}{R+\ B+G} \tag{3}$$

The skin tone classification will be based on computing the distance between two histograms, histogram of a reference skin tone and histogram of a query skin tone, and comparing it to some thresholds.

In literature, two methodologies exist in histogram distance measure: probalistic and vector. Probalistic based approach measures the distance between probability density functions. Examples of distances used in this approach are the Bhattacharyya distance or B-distance [14]. However, vector based measures between fixed histograms are more used in image indexing and retrieval [15] [11] such as city block, euclidean, correlation or intersection. In this paper, correlation has been used as a distance measure between two histograms h_1 and $h2$. The range of values of this measure is always included between 0 and 1. The closer the distance to 1, more similarity is detected. This measure is given by:

$$d(h_1, h_2) = \frac{\sum_{i=1}^{N} \overline{h_1}(i)\overline{h_2}(i)}{\sqrt{\sum_{i=1}^{N} \overline{h_1}^2(i)\ \sum_{i=1}^{N} \overline{h_2}^2(i)}} \tag{4}$$

Where

$$\bar{h}(i) = h(i) - \frac{1}{N}\sum_{i=1}^{N} h(i) \tag{5}$$

In order to improve the classification results, we have added a second rule, which is based on the interpretation of the color distribution of an image as a probability distribution. In fact, the color distribution can be

 ISBN: 978-1-61208-320-9

characterized by three moments on each channel: mean, variance and skewness. Skewness measures the asymmetry of the probability distribution. If the value of the c color channel at the (x, y) image pixel is $f_c(x,y)$ and the number of pixels in the image is M × N, then skewness is given by:

$$(skewness)_c = \left(\frac{1}{M \times N}\sum_{x=1}^{M}\sum_{y=1}^{N}(f_c(x,y) - m_c)^3\right)^{\frac{1}{3}} \quad (6)$$

Where

$$m_c = \frac{1}{M \times N}\sum_{x=1}^{M}\sum_{y=1}^{N} f_c(x,y) \quad (7)$$

We computed the mean and skewness of red component values from the RGB color space of each image and compare them to some boundary. This choice was based on the fact that a majority of skin colors cluster in red channel.

C. *Human skin color classification process*

Skin color classification is composed from a pre-processing step and classification step. The entire process is introduced in Fig. 2.

The pre-processing consists first in computing ratio given in equation (3) of the reference image, which represents black or white skin color tone and it is segmented so that the only parts left are skin regions. Then, histogram $hist_{ref}$ is plotted.

The classification phase starts by first processing the query image: face detection, using the Viola-Jones face detector [17], and segmentation using the method described in paragraph II.A. After that, ratio (3), mean and skewness of red component values are computed.

Then, histogram $hist_{query}$ is plotted and distance between histogram of query image and the one of reference image is computed. Let d_{hist} the distance between $hist_{ref}$ of reference image and $hist_{query}$ of query image.

Finally, obtained values are compared to some defined thresholds in order to classify the skin color into its corresponding tone where at least two rules should be satisfied.

Experiments of this proposed method have led to a new rule. Skin color tone is classified as black if at least two out of the following three conditions are satisfied:

$$d_{hist} > T_1 \quad (8)$$

$$Mean < T_2 \quad (9)$$

$$Skewness > 0 \quad (10)$$

Where T_1 and T_2 are thresholds experimentally determined.

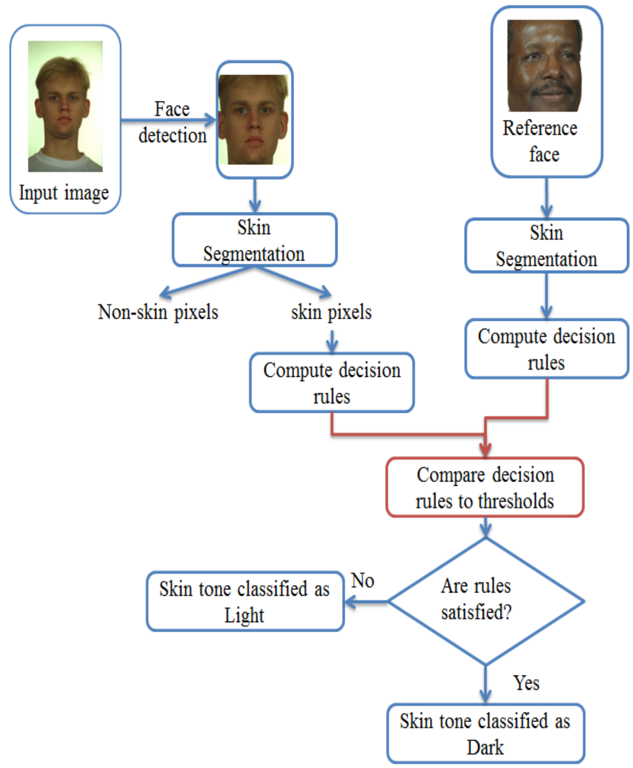

Figure 2. Human skin color classification process

III. Application of the Human Skin Classifier to Face Recognition

Human skin color classification may be considered as a preprocessing operation to many applications. In our proper application, we applied it on face recognition using SIFT descriptors.

A. *SIFT descriptors definition*

SIFT is an algorithm developed by David Lowe [6]. It aims to detect and identify similarities between extracted features of different digital images. SIFT features extraction consists mainly in four steps:

- First step is to detect points of interest which correspond to the extreme points in an image. Those points are calculated from plane subsets of Difference of Gaussian (DoG) filters applied to the image at different scales.
- Then, points of interest with low extreme of DoG are discarded.
- After that, one or many orientations are given to the relevant points of interest.
- Finally, digital descriptors derived from these orientations are modeled with a set of 128-length feature vectors.

SIFT descriptor outputs large number of features with different scales and locations that cover the whole image.

 ISBN: 978-1-61208-320-9

Once extracted, feature vectors of an image may be compared to the query ones to find the most relevant. This is called matching process.

B. Matching features

Feature matching has presented a major concern in computer vision and pattern recognition for several decades. For image matching and recognition, extracted features from the input image are compared with ones extracted from training images in database in order to identify the most similar image to the input one. This process is based on an image similarity measure between two images. Many factors can affect the performance of the matching such as the matching measure criterion and the type of used features. In this paper, we employ the fast approximate near neighbors measure [5].

In [5], Muja and Lowe compare many algorithms for Fast Approximate Nearest Neighbor (FANN) search. As a result, two algorithms showed the best performance. This algorithm used either the hierarchical k-means tree or multiple randomized kd-trees.

C. Face recognition

The human skin color classification process is first applied on our database in order to divide it into two sub-databases and to compute and save SIFT descriptors for each image. In the recognition phase, the skin color tone of the query image is determined, SIFT descriptor is computed and finally, the feature matching distance between SIFT descriptor of query image and ones of each of trained images in the appropriate database is evaluated. The resulting distance matches the query image to the nearest ones in database.

IV. RESULTS AND DISCUSSION

Our experiments were carried out with a set of FERET images [16]. This database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images.
In this paper, we selected a set of 200 (near) frontal FERET faces with different skin color tones: 100 samples of faces classified as black and 100 samples of faces classified as white. We first present the evaluation results of our skin color classifier followed by the face recognition results before and after classification.

1) *Human skin color classifier:* The obtained results of our classifier applied on each skin tone are shown in table I.

TABLE I. HIT AND ERROR RATES

Skin color	white	black	Total
Hit percentage (%)	90	84	87
Error percentage (%)	10	16	13

For the 100 white tone sample images, 10 were classified incorrectly, while for the 100 black tone sample images, 16 were classified incorrectly. In fact, misclassification is caused mainly by bad illumination conditions. Fig. 3 shows some of misclassified skin color tones. The two faces belong to white tone but due to dark illumination they were classified as black tone.

Moreover, when considering the miscalssified black skin tones, we note that most of them belong to brown skin tones.

Figure 3. Examples of incorrectly classified black skin color tone

Figure 4. Examples of incorrectly classified skin color tones

Our developed classifier is able to classify 83% of faces successfully. This result is higher than the hit rate reached in [10] where only 70% of skin color tones were well classified. In that study, faces were classified into black, brown and white skin tones using also a pixel-based classification based on the RGB model.

The time elapsed for the classification of our database is around 100sec (0.5sec for each image). This elapsed time includes: reading image from database, face detection, face segmentation, average of red component computing, histograms computing and comparison to thresholds.

2) *Face recognition using SIFT descriptors:* An example of the performance of the employed face recognition technique before and after skin classification is presented. Fig. 5 presents the query image. In our database, five images belong to this person.

Figure 5. Query image

 ISBN: 978-1-61208-320-9

The face recognition results before and after classification are illustrated in Fig. 6 and Fig. 7 where the first most similar images to the query one are displayed.

Figure 6. The first similar images to the query before skin classification

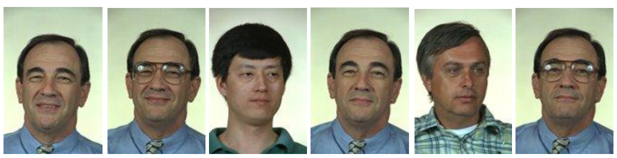

Figure 7. The first similar images to the query after skin classification

According to the qualitative evaluation of the retrieved images, retrieval results are ameliorated and become more accurate. In fact, in the presented example, 90% of faces belonging to the same person appeared in the first ten retrieved images.

V. CONCLUSION AND FUTURE WORK

Human skin color classification into skin tones is a hard operation since skin color may be easily affected by environmental effects especially illumination (light, shade, etc.). Moreover, it is considered as a delicate operation since it is employed as a preprocessing step in many systems such as face recognition. As a consequence, those systems' performance is highly related to the results obtained in the classification step. In spite of those facts, our proposed human skin color classifier based on RGB model succeeded to reach a hit rate of 87%. Also, its high speed and accuracy makes it appropriate for real time applications. Hence, classification reduces the processing time, but can degrade the recognition performance.

In the future, this work should focus on overcoming the effect of illumination in skin color classification. In fact the luminance histogram skewness is correlated with surface brightness. When the image of a surface has positively skewed statistics, it tends to appear darker than a similar surface with lower skewness. Thereby, image illumination can be enhanced basing on the skewness value.

ACKNOWLEDGMENT

This research and innovation work is carried out within a MOBIDOC thesis funded by the EU under the PASRI project.

REFERENCES

[1] G. Osman, M. S. Hitam and M. N. Ismail, "Enhanced skin colour classifier using RGB Ratio model," arXiv preprint arXiv:1212.2692, 2012.

[2] V. Vezhnevets, V. Sazonov and A. Andreeva, "A survey on pixel-based skin color detection techniques," Proc. Graphicon, 2003, pp. 85-92.

[3] K. Ravichandran and B. Ananthi, "Color skin segmentation using k-means cluster," International Journal of Computational and Applied Mathematics, vol. 4, no. 2, 2009, pp. 153-157.

[4] M. C. Shin, K. I. Chang and L. V. Tsap, "Does colorspace transformation make any difference on skin detection?," in Applications of Computer Vision, 2002. (WACV 2002). Proceedings. Sixth IEEE Workshop on, IEEE, 2002, pp. 275-279.

[5] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," In VISAPP(1), 2009, pp. 331-340.

[6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, no. 2, 2004, pp. 91-110.

[7] N. Bourbakis, P. Kakumanu, S. Makrogiannis, R. Bryll and S. Panchanathan, "Neural network approach for image chromatic adaptation for skin color detection," International journal of neural systems, vol. 17, no. 01, 2007, pp. 1-12.

[8] D. Chai, S. L. Phung and A. Bouzerdoum, "A Bayesian skin/non-skin color classifier using non-parametric density estimation," in Circuits and Systems, 2003. ISCAS'03. Proceedings of the 2003 International Symposium on, pp. II-464, 2003.

[9] A. E. Abdel-Hakim and A. A. Farag, "Csift: A sift descriptor with color invariant characteristics.," In Computer Vision and Pattern Recognition,2006 IEEE Computer Society Conference on, vol. 2, 2006, pp. 1978-1983.

[10] I. Boaventura, V. Volpe, I. da Silva and . A. Gonzaga, "Fuzzy Classification of Human Skin Color in Color Images," IEEE Int. Conf. on Systems, Man and Cybernetics, vol. 6, 2007, pp. 5071-5075.

[11] M. J. Swain and D. H. Ballard, "Color indexing," International journal of computer vision, vol. 7, no. 1, 1991, pp. 11-32.

[12] J. Kovac, P. Peer and F. Solina, "Human skin color clustering for face detection," vol. 2, 2003.

[13] D. G. Lowe, "Object recognition from local scale-invariant features," Computer vision, 1999. The proceedings of the seventh IEEE international conference on, vol. 2, 1999, pp. 1150-1157.

[14] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," Communication Technology, IEEE Transactions on, vol. 15, no. 1, 1967, pp. 52-60.

[15] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic and others, "Query by image and video content: The QBIC system," Computer, vol. 28, no. 9, 1995, pp. 23-32.

[16] P. J. Phillips, H. Wechsler, J. Huang and P. J. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," Image and vision computing, vol. 16, no. 5, 1998, pp. 295-306.

[17] P. Viola et M. Jones , "Rapid object detection using a boosted cascade of simple features," In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, IEEE, 2001, pp. I-511.

 ISBN: 978-1-61208-320-9