

Sfax-Miracl Hand Database for Contactless Hand Biometrics Applications

Salma Ben Jemaa¹, Mayssa Frikha², Imen Moalla²,
Mohamed Hammami², and Hanene Ben-Abdallah¹

¹ MIRACL-FSEG, Sfax University, Road Aeroport Km 4, 3018 Sfax, Tunisia

² MIRACL-FS, Sfax University, Road Sokra Km 3 BP 802, 3018 Sfax, Tunisia
benjemaa.salma@hotmail.com, {frikha.mayssa, imen.moalla}@hotmail.fr,
mohamed.hammami@fss.rnu.tn, hanene.benabdallah@fsegs.rnu.tn

Abstract. A new branch of biometrics, hand recognition, has attracted increasing amount of attention in recent years. In this paper, we propose an approach of hand detection based skin color pixel for biometric applications using multi layer perceptron (MLP) neural network. This later has the ability to classify skin pixels belonging to people with different skin tones and captured in different lighting conditions and complex background environments. To improve the achieved results, a succession of post-processing was proposed. The choice of the database is an important step in testing a biometric process. For this, we build a database named “Sfax-Miracl hand database”. This database contains a total of 1080 images having the advantage of being captured from freely posed hands in contact free settings. Various conducted experiments on this database show promising results and demonstrate the effectiveness of the proposed approach.

Keywords: Hand biometric, Contactless hand detection, Skin color pixel, Multi layer perceptron (MLP).

1 Introduction

Automatic recognition using biometric characteristics is becoming more and more popular in the current e-world. As an important member of the biometric family, hand has merits such as robustness, user-friendliness, high accuracy and cost-effectiveness. Hand biometric has many modality characteristics such as palmprint, fingerprint, hand shape, etc. All these modalities need a detection step of the hand that represents the first stage of biometric recognition process. In the literature, three approaches have been proposed to solve the problem of skin detection. The color based approach [1–11], the shape based approach [12, 13] and the hybrid approach [14]. Color is a powerful fundamental feature for skin detection. In fact, it is invariant to the change of position or scale of a given hand and it had shown satisfactory results in the literature. For these reasons, we interested in the color based approach. Regarding this approach, we find three categories of skin modeling named parametric [2, 6, 7], non-parametric [1, 4, 5, 9–11] and explicit models [3, 8]. The importance of non-parametric skin modeling

methods is to estimate skin color distribution from the training data without deriving an explicit model of the skin color. In addition, they are characterized by their rapidity in both training and classification which is very interesting in real-time applications. Therefore, they could be used efficiently for skin detection. The most known methods in the non-parametric skin modeling methods are the normalized lookup table [4, 5], the histogram based segmentation [1], the induction graph [11] and the neural network [9, 10]. The neural network has the ability of a machine learning which produce a generic prediction model from a simple input data. Thus, we opted for the neural network to construct our prediction model.

In the most existing hand biometric systems in the literature, the hand is placed on a glass plate which implies the presence of noise that may provide diseases transfer. Our contribution consists mainly to establish a robust approach for contactless hand detection. A wide variety of color spaces have been applied to the problem of skin color modeling. In the most previous works, researchers opted for a specific color space to construct their skin color model. Our work consists of studying the most widely used color spaces in the problem of skin color modeling in order to determine the adequate color space for our prediction model. In our knowledge, the most existing hand databases in the literature are linked to several constraints like grayscale images with a simple and familiar environment or in other case, markers are placed on the fingertips which disturb the users. For these reasons, another novelty of our work is to build a database which contains hand images detected from freely posed hands in contact free settings that we have called “Sfax-Miracl hand database”.

The remainder of this paper is organized as follows: Section 2 presents the construction details of our database “Sfax-Miracl hand database”. Section 3 describes our proposed hand detection approach. Section 4 discusses the experimental results. Finally, section 5 draws conclusions and futures work.

2 Data Collection: “Sfax-Miracl Hand Database”

Data acquisition is the first stage of every biometric system. Between February 2011 and April 2011, we collected a database of over 1080 hand images captured from 54 peoples where 29 of them are woman and 25 of them are men. This contactless hand images, named “Sfax-Miracl hand database”, consists mainly of hand images collected from volunteers include students, graduate students, workers, retired, etc. 48 of them are less than 30 years old, 3 of them are between 30 and 60 and 2 of them are more than 60 years old. Each one of them is asked to provide about 10 images for their left hand and 10 images for their right hand in one session. In total, each subject provides about 20 images. So that, our database contains a total of 1080 images captured from 108 different hands. This database is built by using Samsung ES75 digital camera. All the images are available in JPEG format with a resolution of 1024*768 pixels. In the process of acquisition, the user places his hand opened in front of the perpendicular axe of the camera with a distance of about 30 cm. The hand images are

captured by taking into account several problems such as the variation of the background (complex, simple, etc.), the variation of the skin tones according to the person and to the races (white, black, brown, red, etc.), the variation of the position and the orientation of the hand, the variation of the hand size which vary according to the age and to the sex, the presence of hand's accessories (ring, bracelet, watch, etc.) and the variation of the lighting conditions (indoor and outdoor environments). We note the presence of parts or the totality of the face for the concerned person or others for some images. The main objective in building this database is to have the unsupervised conditions for the imaging which attempts to represent more realistic application environment. This database is publically available in the contact of the authors. Fig. 1 shows an example of typical acquisitions.

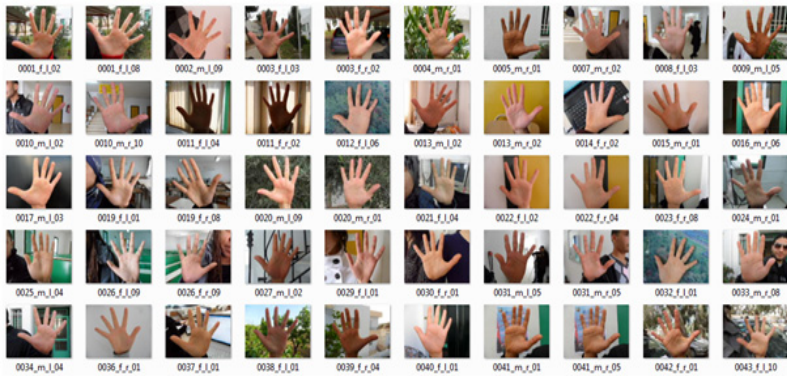


Fig. 1. Extract of images from the database “Sfax-Miracl hand database”

Besides, our database has the advantage to contain masks in which the skin regions were preserved and the background was replaced by black color. These masks are created manually for each images of our database using Adobe Photoshop 7 platform and it were stored in independent files with a name corresponding to its origin images. The creation of these masks served us to extract automatically the skin color and the non-skin color pixels from the original image to be used for the preparation of the training data set and test data set. Fig. 2 shows the mask corresponding to the acquisitions of Fig. 1.

3 The Proposed Hand Detection Approach

Our proposed approach involves two steps: (1) the learning step to construct the adequate prediction model to discriminate the pixels of skin from those of non-skin and (2) the detection step to only conserve the hand region. The total process of our proposed approach is shown in Fig. 3.



Fig. 2. Corresponding masks to images acquisition for “Sfax-Miracl hand database”

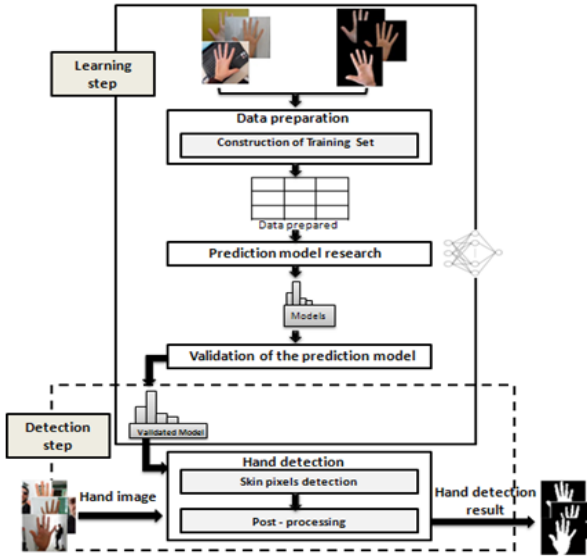


Fig. 3. The total process of our proposed approach

3.1 Learning Step

The learning step is composed of three major phases: (1) a data preparation phase for the construction of the training data set, (2) a prediction model research phase which looks for a generalizable prediction model and (3) a validation phase which consists on assessing the quality of the learned prediction model.

Data Preparation. In this phase, our purpose is to build a two-dimensional table from our training corpus. Each column in the table represents a color space axes and each rows represents a skin pixel value in each color space axes. In the

most previous works, researchers have proposed many prediction models based on skin color, but everyone has opted for a specific color space to present his model. In our work, we try to fill this vacuum by studying the most widely used color spaces in the problem of skin color modeling in order to determinate the adequate color space for our prediction model. The set of training pixels are extracted automatically from the training images and their corresponding binary masks. Thereafter, for each pixel we computed its representation in the used color spaces: RGB, YCrCb, and HSV. This leads to a features vector composed of 9 color space axes. With each pixel features vector is associated its class label denoted as 1 for skin color and 0 for non-skin color.

Prediction Model Research. In this phase, we propose to build a prediction model based on a supervised learning approach. In the literature, there are several techniques of supervised learning, each having its advantages and disadvantages. In our approach, we opted for the use of the MLP as a prediction model method. The MLP [15] is characterized by their capacity of memorization and generalization of the data in addition to their ability to solve not linearly separable problem. A typical MLP network consists of a set of source nodes forming the input layer, one or more hidden layers of computation nodes, and an output layer of nodes. The input signal propagates through the network layer-by-layer. The computations performed by such a feedforward network with a single hidden layer with nonlinear activation functions and a linear output layer can be written as:

$$y = f(x) = B\alpha(Ax + a) + b. \quad (1)$$

where x and y are respectively the vector of inputs and outputs and α is the activation function. A is the matrix of weights of the first layer and a is the bias vector of the first layer. B and b are, respectively, the weight matrix and the bias vector of the second layer.

One of the important aspects in designing an MLP neural network is how to determine the network topology. The input size is dictated by the number of available inputs features. Each color component of the used color space is treated as an input neuron; in our case we need three neurons. The output layer will have one neuron. Thus, the two decisions that must be made regarding the hidden units are to determine the number of hidden layers and the number of neurons in each hidden layer. Fu in 1994 [16] stated that using only one hidden layer is sufficient to solve many practical problems, so in our work we used one hidden layer MLP neural network. The determination of the number of neurons in the hidden layer will be discussed in section 4. The network is trained using error back propagation training algorithm.

Validation of the Prediction Model. Several possible metrics have been proposed in the literature for assessing the quality of a prediction model for skin color pixel detection. Among this metrics, four types of rates were used in our work: (1) The Correct Detection Rate (CDR): represent the probability of the correct classified pixels, (2) The False Detection Rate (FDR): represent the

probability of the wrongly classified pixels, (3) The A Priori Error Rate (APER): represent the probability of the skin pixels which are not assigned to this class, (4) The A POsteriori Error Rate (APOER): represent the probability of pixels assigned to the skin class and which do not belong to this class.

3.2 Detection Step

The detection procedure is done by scanning all the pixels in the image and for each pixel, the MLP neural network takes as input three neurons, each one takes the value of each color component of the chosen color space. Then the MLP neural network computes the probability that each pixel is a skin pixel. So, the pixels whose probability is higher to the threshold value will considered as skin color and takes the value of “1” otherwise it will be considered as background tone and will be set to “0”. This process produces a binary image highlighting the skin color pixels with white color and the others with the black color as shown in Fig. 4(b).

The result of detection produced by the MLP neural network may contain hand region that is corrupted by false detection (Fig. 4(b)). Therefore, a succession of post-processing was proposed allowing us to remove this false detection. We started by applying morphological operator: opening with a circular structuring element. The basic effect for this operation is to remove some false skin pixels detected in the background. Given that in some hand images, the presence of parts or the totality of the face for the concerned person or others, so finally, we eliminate the regions having a surface lower than 500 to preserve only the larger white objects in the image e.g. the hand. The result of post-processing phase is shown in Fig. 4(c).

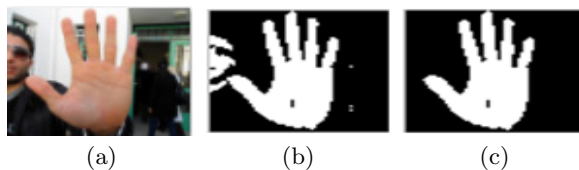


Fig. 4. Hand detection results: (a) original image (b) detection results of skin color (c) hand image after post-processing phase

4 Experimental Results and Evaluation

In the herein reported experiments, we try to validate our contribution and to evaluate the effectiveness and robustness of our approach. All the images used in the following experiments are taken from our database and resized with a size of 64*64 pixels. For the construction of our training data set, we try to choose images from different individuals and having different variations. This training data set is composed of 245760 pixels from which 83349 are skin color pixels and 162411 are non-skin color pixels. Independently of the training data set, we

construct our test data set. This later is composed of 4177920 pixels from which 1376772 are skin color pixels and 2801148 are non-skin color pixels.

Since the transfer function used is a sigmoid function, the MLP neural network produces an output between 0 and 1. Therefore, the output of the neural network needs to be modified so that it is either 0 or 1. To achieve this, we need a threshold value. To choose the threshold value, we calculate the CDR on the training data set using a single neuron in the hidden layer (Table 1).

Table 1. Correct detection rate for different thresholds value

Correct Detection Rate(%)	0.5	0.6	0.7	0.8	0.9
RGB	33.9148	89.7587	89.1207	88.3309	87.168
YCrCb	33.9148	89.9512	89.4775	88.798	86.8542
HSV	33.9148	85.4822	84.8039	80.3259	66.0852

As we have seen from these results, our approach achieved its best performance with the threshold 0.6 for the various used color spaces.

The choice of the adequate neurons number is determined by the calculation of the CDR on the training data set by modifying the neuron number in the hidden layer from 1 to 25 with a step of 2 neurons (Table 2).

Table 2. Correct detection rate for the different neurons number in the hidden layer for the three color spaces: RGB, YCrCb and HSV

Correct Detection Rate(%)	1	3	5	7	9	11	13	15	17	19	21	23	25
RGB	89.76	90.52	66.09	90.35	90.5	90.53	90.41	90.46	90.5	90.4	90.37	90.5	66.09
YCrCb	89.95	90.49	90.48	90.63	90.69	66.08	90.67	90.78	90.66	90.58	90.74	90.73	90.76
HSV	85.48	90.13	89.9	90.11	90.31	90.26	90.37	90.34	90.31	90.38	66.28	90.46	66.08

From the results of previous experiment, we can observe that the best result of correct detection rate is achieved using 15 neurons in the hidden layer.

In order to determinate the adequate color space for our prediction model, we conducted a comparison between the obtained results by the RGB, the YCrCb, and the HSV color space. This comparison does not only concern the obtained CDR but also the FDR, the APER, and the APOER (Table 3).

From this experiment, since a compromise between the CDR, the FDR, the APER, and the APOER, the best result is obtained with the RGB color space. The execution time is another interesting factor in a detection skin process. Table 4 illustrates the average computing time of our proposed approach.

As illustrated from the previous result, our approach is not only effective but also very fast. Therefore, we have once again shown the performance of our proposed approach.

Table 3. Comparison between the CDR, the FDR, the APER, and the APOER using the RGB, the YCrCb, and the HSV color space

(%)	CDR	FDR	APER	APOER
RGB	94.53	5.46	15.57	2.15
YCrCb	94.52	5.47	15.13	2.02
HSV	94.41	5.58	15.67	2.12

Table 4. Average computing time of our proposed hand detection approach

Color space	Average computing time (s)
RGB	0.0625 s

5 Conclusions and Futures Works

We have introduced in this paper a method for hand detection captured without contact and without constraints on the capture environment for hand biometric applications. This method is based on color based approach adopting the MLP for the skin color modeling. First of all, we begin with presenting all the details of the construction process of our database “Sfax-Miracl hand database”. Then, we describe the various steps of our approach. This later involves two steps: the first is the learning step to construct the adequate prediction model and the second is the detection step to only conserve the hand region. Finally, we conducted several experiments in our database. These experiments showed that our approach provide an efficient detection results with a CDR of 94.53% in the RGB color space.

Our future orientations consist of studying the possibility of generating an hybrid color space to better discriminate the skin pixels. Given that the run-time is important for such a biometric application, it would be interesting to explore other learning techniques such as the induction graphs.

References

1. Saxe, D., Foulds, R.: Toward robust skin identification in video images. In: 2nd International Face and Gesture Recognition Conference, pp. 379–384. IEEE Computer Society Press, USA (1996)
2. Yang, M., Ahuja, N.: Detecting human faces in color images. In: IEEE International Conference on Image Processing, Chicago, pp. 127–130 (1998)
3. Garsia, C., Tziritas, G.: Face detection using quantized skin color region merging and wavelet packet analysis. J. IEEE Trans. on Multimedia 1, 264–277 (1999)
4. Berard, F.: Vision par ordinateur pour l’interaction homme-machine fortement couplee. PhD thesis, Joseph Fourier University, France (2000)

5. Wu, A., Shah, M., da Vitoria Lobo, N.: A virtual 3d blackboard: 3d finger tracking using a single camera. In: Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 536–543. IEEE Press, France (2000)
6. Greenspan, H., Goldberger, J., Eshet, I.: Mixture model for face-color modeling and segmentation. *J. Patt. Recogn. Lett.* 22, 1525–1535 (2001)
7. Lee, J.Y., Yoo, S.I.: An elliptical boundary model for skin color detection. In: International Conference on Imaging Science Systems and Technology, Las Vegas, pp. 579–584 (2002)
8. Chiang, C.C., Tai, W.K., Yang, M.T., Huang, Y.T., Huang, C.J.: A novel method for detecting lips, eyes and faces in real time. *J. Real-Time Imaging* 9, 277–287 (2003)
9. Bhoyar, K.K., Kakde, O.G.: Skin color detection model using neural networks and its performance evaluation. *J. of Computer Science* 6, 963–968 (2010)
10. Doukim, C.A., Dargham, J.A., Chekima, A., Omatu, S.: Combining neural networks for skin detection. *J. Signal and Image Processing* 1, 1–11 (2010)
11. Hammami, M., Tsishkou, D., Chen, L.: Data-mining based Skin-color Modeling and Applications. In: Third International Workshop on Content-Based Multimedia Indexing, pp. 157–162. SuviSoft Oy Ltd., France (2003)
12. Cootes, T.F., Taylor, C.J.: Statistical models of appearance for computer vision, Technical report, University of Manchester, United Kingdom (1998)
13. Liu, N., Lovell, B.C.: Hand gesture extraction by active shape models. In: Digital Imaging Computing: Techniques and Applications Conference, pp. 59–64. IEEE CS Press, Australia (2005)
14. Doublet, J., Lepetit, O., Revenu, M.: Hand Detection for Contact less Biometrics Identification. In: International Conference Cognitive System with Interactive Sensors, France (2006)
15. Krose, B., Smagt, P.V.D.: An introduction to neural networks, Amsterdam (1996)
16. Fu, L.M.: Neural networks in computer intelligence. McGraw-Hill, India (1994)