SFA: A Human Skin Image Database based on FERET and AR Facial Images

João Paulo Brognoni Casati*, Diego Rafael Moraes† and Evandro Luis Linhari Rodrigues‡
Department of Electrical and Computer Engineering
São Carlos School of Engineering
University of São Paulo
*jpcasati@usp.br, †diego.moraes@usp.br and ‡evandro@sc.usp.br

***Abstract*—Human skin color is a useful feature for several computer vision research, including face recognition. This paper presents a human skin image database called SFA. It aims to assist research that use skin color or texture as a feature. SFA was constructed based on face images of FERET (876 images) and AR (242 images) databases, from which skin and non-skin samples and the ground truths of skin detection were retrieved. The samples vary of dimension, from 1 pixel to 35x35 pixels. For each dimension, SFA has 3354 samples of skin and 5590 samples of non-skin. The samples validation was made by a pattern classification process using artificial neural networks, reaching around 93% accuracy. A comparison with UCI database concerning image segmentation was made, where SFA showed almost 4% of improvement. The SFA image database showed great potential to assist research which have skin color or texture as a relevant feature.**

I. Introduction

Face recognition and detection are widely studied subjects of computer vision, mainly for application in security biometric systems, robotics, intelligent user interfaces, face search in video databases and others [1]. A comprehensive survey of face recognition can be found in [2].

There is a discussion in the area of face recognition about the use or not of color as a relevant feature for identification of faces in an image. Karimi and Krzyzak [3] showed that color can raise the accuracy of several face recognition algorithms, achieving better results in traditional recognition methods, but improving even some new methods.

The focus of this paper is to develop a tool that would assist some research in skin color detection. Some papers use color or texture to segment human skin in images [4], [5], [6], [7], [8], [9], [10]. The skin color is an important feature for some methodologies, so it has to be detected with high accuracy or else the features will have much noise.

To assist research in the area of computer vision, specific to face recognition, there are some well-known color image databases, like The AR face database [11], FERET [12], UCD [13], and others.

The use of a skin sample database can assist research in human skin detection. The UCI skin segmentation data set [14] is based on images of two face images databases, FERET and PAL [15] databases. It consists of 245,057 samples of skin and non-skin RGB pixels, from which are 50,859 skin samples and 194,198 non-skin samples. There is a comparison of UCI and SFA in this paper.

Some images of FERET and AR databases were used by Severino Jr. and Gonzaga [7] to validate a methodology of skin detection in digital images. Aiming the validation of their methodology, they have manually made the ground truths (GT) of some images, which are the original image with the skin area segmented.

Aiming to assist research that uses human skin as useful feature it was created the SFA, a human skin face image database based on two classic face databases: The AR Face Database and FERET face database. It has images of people of different races (Asian, Caucasian and Negro), age and sex (male and female). Differently of UCI, the SFA has samples of various dimensions, and it's composed of four distinct kinds of images. SFA database is available for download in http://www.sel.eesc.usp.br/sfa. The details of SFA and of its construction are shown in next section.

II. Methodology

The methodology is separated in three parts, the first one presents how the database was created, the second one shows how it was validated and the third one shows the SFA being used to segment images, including a comparison with UCI [14] (another skin sample database).

The whole process was made using the Matlab programming environment, including its Neural Networks and Image Processing toolboxes, which has some functions that assists the development. [16].

1) Database Construction: SFA human skin image database was created as a tool to assist research in computer vision. It is composed of samples of skin, samples of non-skin, the original images and the ground truths. The sample images used to build the SFA database were retrieved from two classic and widely used in the literature face image databases, AR and FERET. These two databases consist of facial color images, as some examples are shown in Figure 1. Every image of SFA was saved in JPEG format, with 100% of quality level.

From the AR database were used 242 images. This database is more controlled with all images having white background and little skin color variations. From FERET were used 876 images, where the environment is less controlled than AR, with high skin color and background variation.

The skin and non-skin samples were randomly retrieved considering the GT mask to define which is a skin sample and which is a non-skin sample. From each image it was retrieved

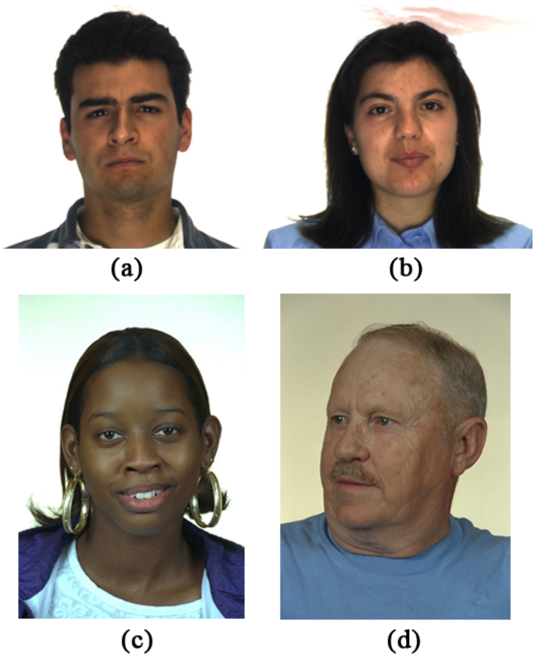

Figure 1. Figures (a) and (b) are examples of original images from AR. Figures (c) and (d) are examples of original images from FERET [11], [12].

three samples of skin and five of non-skin. Each sample has a central pixel, from which other sample sizes were created. The samples vary in two pixels, from a sample of 1 pixel to the bigger one, of 35x35 pixels dimension. In Figure 2 an example of skin sample creation is shown. The same process was used to non-skin samples.

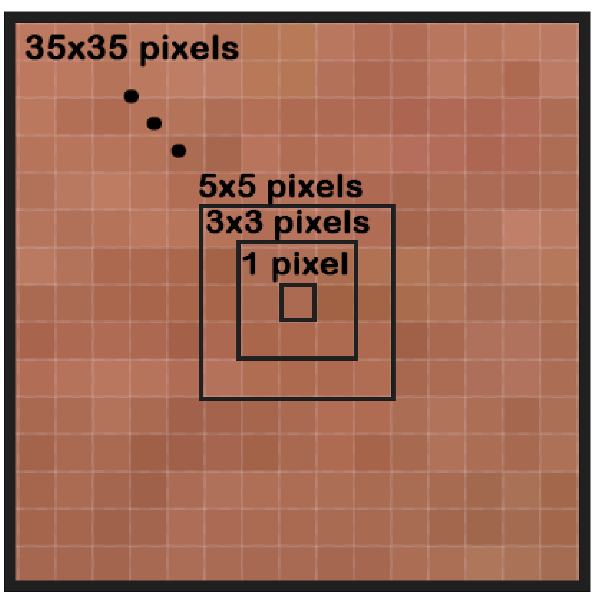

Figure 2. Example of samples creation.

The 1118 ground truths of the original images were provided by Severino Jr. and Gonzaga [7] for the use in SFA. In these images, every pixel but the skin were manually painted in black (RGB 0,0,0). Examples of these images are shown in Figure 3.

Figure 3. Original images followed of the ground truths (adapted from [7], [12]).

The images and sample images were organized in folders divided in four main divisions: Original Images (ORI), Ground Truths (GT), Samples of Skin (SKIN) and samples of non-skin (NS). In the folders SKIN and NS, there are divisions considering the dimension of the samples. The folder structure is better visualized in Figure 4.

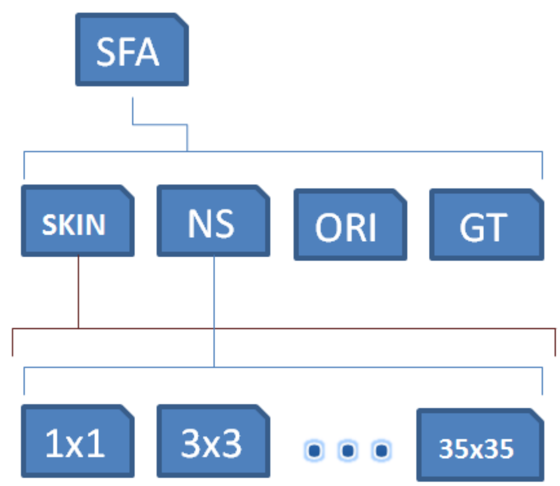

Figure 4. Folder structure of SFA.

To emphasize the quality of these generated samples, next subsection shows a validation step of skin and non-skin samples.

2) Samples Validation: The validation of the samples was made using Artificial Neural Networks (ANN). This methodology was chosen because it is an intelligent way to separate the sample space. The architecture used was Multi-Layer Perceptron (MLP), which is the most used architecture for pattern recognition [17]. MLP has supervised training, it means that for each sample, a desired output is needed (targets) [18]. The process of validation using ANN is based on [19].

This step of validation was made for every sample dimension, from 1 pixel to 35x35 pixels. For the samples bigger than 1 pixel, the ANN inputs are the mean of the RGB values in the sample, maintaining the number of three inputs of the neural networks.

Ten trainings for each one of the ten different topologies were made, aiming to choose the best one. The topologies varied from 2 to 20 neurons, varying in two neurons, in an unique hidden layer as shown in Figure 5. The training algorithm used was the Scaled Conjugate Gradient, which is the default algorithm for pattern recognition problems in the Matlab Neural Networks Toolbox.

The weights and bias were randomly set ranging between 0 and 1. The goal was defined 10^{-4}, epochs limit were 150, the activation function used in the hidden layer was hyperbolic tangent and the samples were divided 70% for training, 15% for validation and 15% for testing.

In the output layer, just one neuron was needed to classify the patterns, because there was only two possibilities, thus, the output of 0 represents non-skin and the output of 1 represents skin, and in this layer, it was used a pure linear activation function.

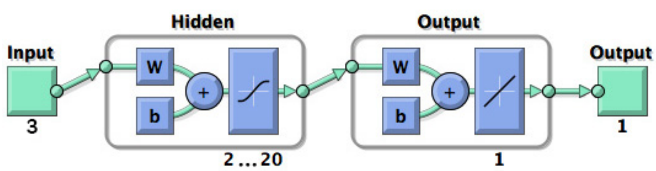

Figure 5. MLP neural network topologies (adapted from Matlab neural networks toolbox).

For the choose of the best topology, the correlation between the targets and outputs (R) of the test samples was considered the main parameter for the decision.

Other validation and example of use of the SFA database is shown in next subsection, where a comparison with UCI is made.

3) Image Segmentation: For this step of comparison, it was used ANN again. The same process used for validation of the samples for SFA database, was made to UCI database. Then, the best training was selected for each database, and the weights and bias of these trainings was used to segment the original images.

Although the SFA database has samples of several dimensions, the comparison was made using the one pixel samples, according to the UCI samples.

All the 1118 images were pixel-level segmented, and the output of the neural network was binarized resulting in a binary mask, where values with 1 represent the skin and 0 represent non-skin.

After the segmentation, the ground truths are binarized and then used as targets to benchmark the segmentation, thus, retrieving the accuracy for each one of the segmentations. Figure 6 shows an example of original image and its segmentation mask (GT).

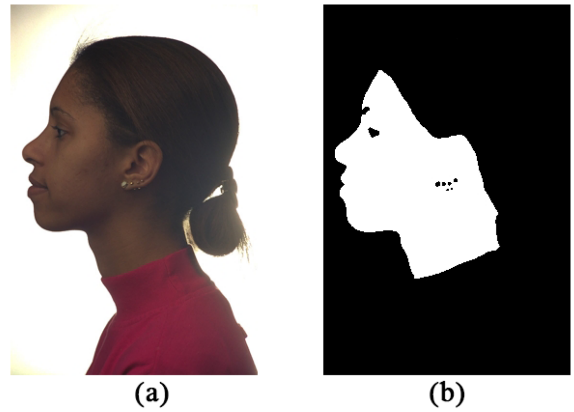

Figure 6. Original Image (a) and ground truth of skin segmentation (b).

The process was made for all the 1118 original images and after the ANN outputs the image, it was binarized with several threshold, from 0.05 to 0.95, ranging from 0.05 to 0.05. Then, the best thresholding for each database (SFA and UCI) was selected according to the ground truths.

III. Results

A. Database Construction

The total number of images, considering every sample, of every dimension and every image is of 163,228 images. Table I shows the division of the samples its quantity, and Table II shows some information about SFA.

Table I. Numbers of SFA database.

Type of Image	Quantity
Original Images	1,118
Ground Truths	1,118
Skin Samples*	3,354
Non-Skin Samples*	5,590
Total of Images	163,228

**For each one of the 18 different dimensions.*

Thinking about having more samples of just one pixel, it is possible to segment the 35x35 samples and then generate an amount of 4,108,650 skin samples and 6,847,750 non-skin samples.

B. Samples Validation

The validation of the samples generated is an important step. It shows if the samples are related with the targets (skin and non-skin). The results of the validation are presented just for the best training of the neural networks, for the samples of dimension of 1 pixel, 5x5 pixels, 15x15 pixels and 25x25 pixels.

The choice of the best training was made using the maximum correlation (R_{max}), aiming that the chosen training could generalize better.

Table II shows the information of (R_{max}) and the accuracy concerning the test group of samples (15%) which was not used in the training step.

Table II. ANN RESULTS FOR THE TEST GROUP OF SAMPLES VALIDATION.

Dimension	Neurons	R_{max}	Accuracy
1 pixel	20	0.9333	96.02%
5x5 pixels	20	0.9312	96.00%
15x15 pixels	20	0.9312	95.96%
25x25 pixels	10	0.9206	95.54%

The results showed high correlation and accuracy, with more than 95% in all the cases. It proves that the samples generated to SFA are correlated to the targets they have, skin and non-skin.

C. Image Segmentation

The results are shown concerning the best topology ANN training and the best threshold value for each database (SFA and UCI). The comparison with the GTs was made to gauge the accuracy.

Table III shows the results of the best threshold for both databases. The best threshold value for SFA and UCI was of 0.10. For the best threshold, SFA shows higher accuracy, overcoming UCI in 4%.

Table III. RESULTS OF ACCURACY FOR SFA AND UCI.

Database	Threshold	Accuracy
SFA	0.10	92.71%
UCI	0.10	88.74%

To understand how the threshold values worked for this process, Figure 7 shows the full range of thresholds and the accuracy for both databases.

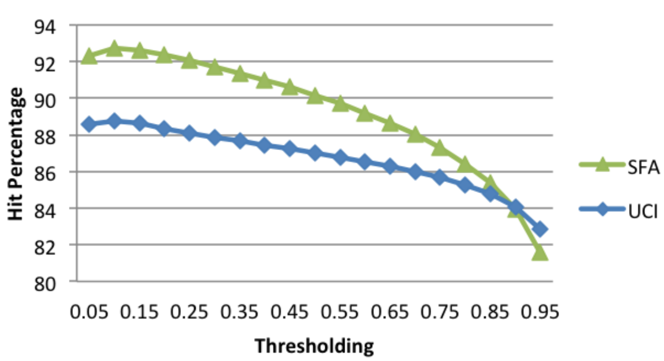

Figure 7. Accuracy and thresholding behavior for both databases (SFA and UCI).

As Table III and Figure 7 show that SFA database presented higher accuracy for the tested images than UCI, confirming that the samples generated to SFA human skin database could segment properly face images of FERET and AR face databases. Figure 8 shows some examples of images segmented in this test of validation, comparing the SFA with UCI segmentations.

As shown in Figure 8, the images segmented using SFA samples are visually closer to the ground truths, and the parameters showed in Table III confirmed it.

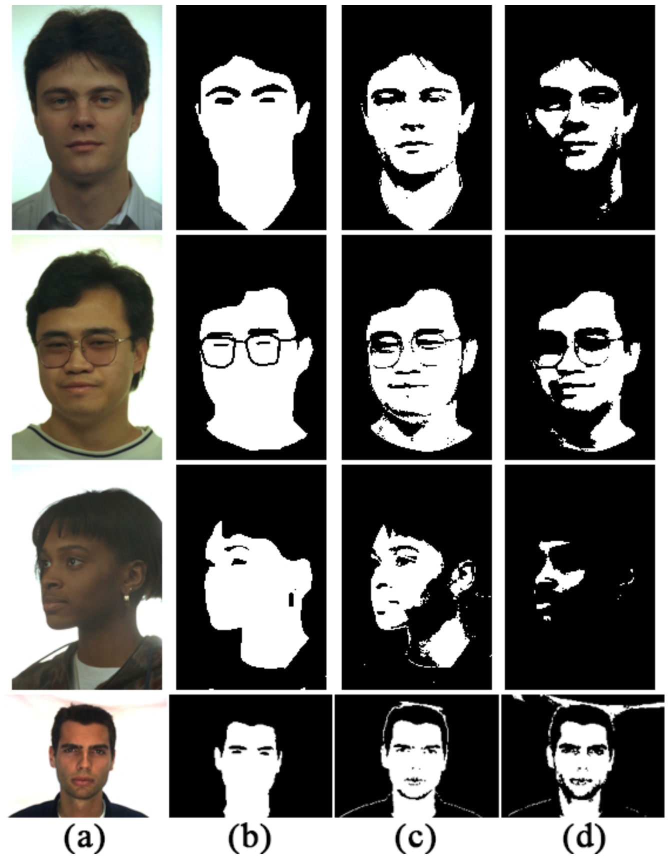

Figure 8. Comparison between SFA and UCI segmentation. The first column (a) has the original images, the second column (b) has the ground truths binarized, the third column (c) has the original images segmented using SFA samples and the fourth column (d) has the original images segmented using the UCI samples.

IV. CONCLUSION

This paper presents SFA, a human face skin image database based on images of FERET and AR face databases. SFA is composed of the original images, the ground truths for benchmarking the skin segmentation and the samples of skin and non-skin.

There was made a validation of these generated samples, which presented results that lead to conclude that the samples are high correlated to the targets, achieving more than 0.9 of correlation.

The segmentation of face images was made comparing with other database of skin samples (UCI) and the results were promising again. The SFA showed high accuracy for segmentation of the face images, achieving more than 92% in an universe of 1118 images of two different face databases. Furthermore, SFA samples improved the segmentation around 4% compared to UCI, opening great possibilities of its use for segmentation of human skin.

Great possibilities of assistance for research in the field of computer vision, specifically face recognition and skin detection can be aimed after the development of this work. SFA also could assist benchmarking, training for segmentation, classification and it can be useful in textures methodologies concerning human skin.

ACKNOWLEDGMENT

The authors would like to thank Dr. Osvaldo Severino Jr. and Dr. Adilson Gonzaga for the contribution and availability of the 1118 ground truth images used in SFA database.

REFERENCES

[1] K. Youssef and P.-Y. Woo, "A new method for face recognition based on color information and a neural network," in *Natural Computation, 2007. ICNC 2007. Third International Conference on*, vol. 2, aug. 2007, pp. 585 –589.

[2] W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face recognition: A literature survey," pp. 399–458, 2003.

[3] B. Karimi and L. Devroye, "A study on significance of color in face recognition using several eigenface algorithms," in *Electrical and Computer Engineering, 2007. CCECE 2007. Canadian Conference on*, 2007, pp. 1309–1312.

[4] S. Phung, A. Bouzerdoum, and S. Chai, D., "Skin segmentation using color pixel classification: analysis and comparison," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 1, pp. 148–154, 2005.

[5] Z. Jiang, M. Yao, and W. Jiang, "Skin detection using color, texture and space information," in *Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007. Fourth International Conference on*, vol. 3, 2007, pp. 366–370.

[6] X. Wang, X. Zhang, and J. Yao, "Skin color detection under complex background," in *Mechatronic Science, Electric Engineering and Computer (MEC), 2011 International Conference on*, 2011, pp. 1985–1988.

[7] O. Severino Jr and A. Gonzaga, "A new approach for color image segmentation based on color mixture," in *Machine Vision and Applications*, 2011.

[8] J. Y. Choi, Y.-M. Ro, and K. Plataniotis, "Boosting color feature selection for color face recognition," *Image Processing, IEEE Transactions on*, vol. 20, no. 5, pp. 1425–1434, 2011.

[9] J. Brand and J. Mason, "A comparative assessment of three approaches to pixel-level human skin-detection," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 1, 2000, pp. 1056–1059 vol.1.

[10] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern Recognition*, vol. 40, no. 3, pp. 1106 – 1122, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320306002767

[11] A. Martinez and R. Benavente, "The ar face database," in *CVC Technical Report Nº 24*, 1998.

[12] P. Phillips, H. Moon, P. Rauss, and S. Rizvi, "The feret evaluation methodology for face-recognition algorithms," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, 1997, pp. 137–143.

[13] P. Sharma and R. B. Reilly, "A colour face image database for benchmarking of automatic face detection algorithms," in *In Proceedings of EC-VIP-MC, the 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications*, 2003.

[14] A. Frank and A. Asuncion, "UCI machine learning repository," 2010. [Online]. Available: http://archive.ics.uci.edu/ml

[15] "Productive Aging Lab - Face Database." [Online]. Available: https://pal.utdallas.edu/facedb/

[16] R. Gonzalez, R. Woods, and S. Eddins, *Digital Image Processing Using Matlab*, 2nd ed. USA: Gatesmark Publishing, 2009.

[17] I. Silva, D. Spatti, and R. Flauzino, *Redes Neurais Artificiais para engenharia e ciências aplicadas: curso prático*, 1st ed. Artliber, 2010.

[18] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall, 1999.

[19] D. R. Moraes, J. P. B. Casati, and E. L. Linhari Rodrigues, "Analysis of polynomial behavior of the c3 cervical concavity to bone age estimation using artificial neural networks," in *Biosignals and Biorobotics Conference (BRC), 2013 ISSNIP*, 2013, pp. 1–6.