# Euclidean Distance Versus Manhattan Distance for New Representative SFA Skin Samples for Human Skin Segmentation

Ouarda Soltani*, Souad Benabdelkader

Electronics Department, Batna2 University, Batna 05000, Algeria

Corresponding Author Email: ouarda.soltani05@gmail.com

## ABSTRACT

The human color skin image database called SFA, specifically designed to assist research in the area of face recognition, constitutes a very important means particularly for the challenging task of skin detection. It has showed high performances comparing to other existing databases. SFA database provides multiple skin and non-skin samples, which in various combinations with each other allow creating new samples that could be useful and more effective. This particular aspect will be investigated, in the present paper, by creating four new representative skin samples according to the four rules of minimum, maximum, mean and median. The obtained samples will be exploited for the purpose of skin segmentation on the basis of the well-known Euclidean and Manhattan distance metrics. Thereafter, performances of the new representative skin samples versus performances of those skin samples, originally provided by SFA, will be illustrated. Simulation results in both SFA and UTD (University of Texas at Dallas) color face databases indicate that detection rates higher than 92% can be achieved with either measure.

## 1. INTRODUCTION

Advances in information and communication technologies over the past years have expanded the scope of application of human skin detection: face detection [1-3], visual tracking for surveillance [4], gesture recognition [5], image retrieval and filtering image contents on the web [6], face recognition system [7, 8], to mention just a few examples.

Face detection represents the first step in a fully automatic face recognition process that could be performed using a wide range of methods including those based on skin color analysis. These latter are considered to be the simplest ones. In addition, color processing is much faster and robust in nature.

Color is a very important human face parameter. Much academic work has been done to investigate human skin color features. Results indicate that skin color has a limited range of hues and is not deeply saturated making it clustered within a small area in the color space [9]. Also under certain lighting conditions, it is robust towards changes in orientation and scaling and can tolerate occlusion well. However, skin detection remains very challenging because of differences in illumination intensity, photos taken using an assortment of cameras with their own characteristics, range of skin colors due to different ethnicities, and many other variations.

Several algorithms for skin detection have been developed over the last years [10, 11]. These may be divided into two major groups: pixel-based methods [12-14] and region-based methods [15-18]. While the former classifies each pixel as skin or non-skin without considering its neighborhood, the latter take advantage of pixels neighborhood to improve the color segmentation process. The presence of skin or non-skin region in a digital image can be determined by manipulating tone and/or texture of pixels.

Among important contributing tools to developing new methodologies more effective and accurate for skin color detection we could easily identify color human skin databases specifically designed to assist research in the area of face recognition [19-22]. One of these is the recently built SFA database [19] that showed high accuracy for segmentation of face images.

SFA database is distinguished by its multiple skin and non-skin samples, which in various combinations with each other allow creating new samples that could be useful and more effective. This particular aspect will be investigated, in the present paper, by constituting out four new representative (min, max, mean, median) skin samples according to the four rules of minimum, maximum, mean and median. The obtained new samples will be exploited for the purpose of skin segmentation. In more details, the scheme proposed here consists, in a first step, in creating four new representative skin samples from original skin samples in SFA, using four different combinations or rules. The second step consists of using the resulting new skin samples together with original non-skin samples from SFA to achieve more effective skin segmentation of any color facial image within and outside SFA database. The well-known Euclidean distance (ED) and Manhattan distance (MD) will act as skin color similarity measures between the facial image to be segmented and the new considered SFA skin sample. In the end, skin detection performances generated by the use of new representative (min, max, mean, median) samples are compared to those obtained by means of original SFA skin samples.

The paper is organized as follows: Section 2 presents a short overview of SFA database. Section 3 describes the new representative skin samples formation. Section 4 gives an overview of the ED and MD based skin segmentation strategy. Section 5 shows the experimental results. Finally, the conclusion is given in section 6.

# 2. SFA DATABASE OVERVIEW

SFA database is built upon two classical facial color images databases, AR (created in 1998 by Martinez and Benavente) [21] and the Facial Recognition Technology (FERET) database [22]. The former contributed with 242 images having all white background, whereas the latter brought 876 images with high skin color and background variation. Every image of SFA was saved in JPEG format, with 100% of quality level.

By design, SFA is composed of four distinct sets of images structured in folders as shown in Figure 1. Each of the 1118 original images (**ORI** folder) has its ground truth (**GT** folder) version, in which every pixel but the skin was manually painted in black (RGB 0, 0, 0). Figure 2-c illustrates an example of such images. With the help of ground truth images to determine which a skin (**SKIN**) sample is and which is a non-skin (**NS**), a total of three skin and five non-skin samples were randomly retrieved from each original image. Every sample consists of a set of square sample masks collected from the same central pixel, and varying in sizes by two pixels, from the smaller of 1 pixel to the largest of 35x35 pixels (Figure 2-a).
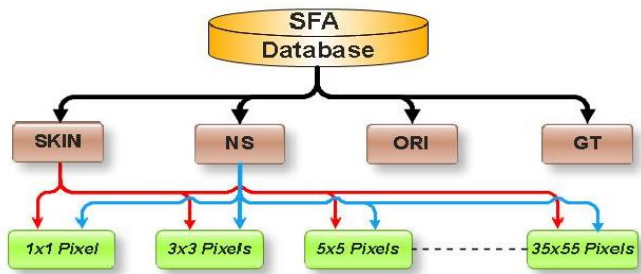


**Figure 1.** The folder-like structure of SFA
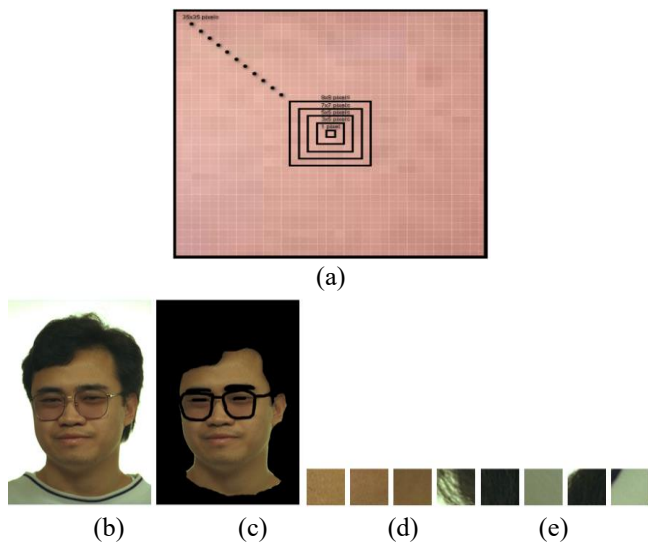


(a)



(b)　　　(c)　　　(d)　　　(e)

**Figure 2.** A typical example of SFA database content. (a) Example of skin samples creation. (b) ORI: img(163). (c) GT: img(163). (d) Three SKIN samples of size 35×35: skin-163-1; skin-163-2; skin-163-3. (e) Five NS samples of size 35×35: ns-163-1; ns-163-2; ns-163-3; ns-163-4; ns-163-5

Figure 2 illustrates a typical example of SFA face database content. The original image is saved in the **ORI** folder as **img(163)**. Its ground truth version is saved in the **GT** folder as **img(163)**. The corresponding **skin** samples of size **35×35** are saved in the subdirectory **SKIN\35** as **skin-163-1**; **skin-163-2** and **skin-163-3**. Finally, the corresponding **ns** samples of size **35×35** are saved in the subdirectory **NS\35** as **ns-163-1**; **ns-163-2**; **ns-163-3**; **ns-163-4** and **ns-163-5**.

# 3. NEW REPRESENTATIVE SKIN SAMPLES

For any given sample size, a new representative skin sample is constituted out of the three original SFA skin samples according to the four rules below:



**Figure 3.** Original SFA skin samples of size **3×3**



**Figure 4.** New SFA representative skin samples according to: (a) Min rule; (b) Max rule; (c) Mean rule; (d) Median rule

- **The min rule:** each new skin pixel takes the lowest value of all three intensities at the same position.
- **The max rule:** each new skin pixel takes the highest value of all three intensities at the same position.
- **The mean rule:** each new skin pixel takes the mean value of all three intensities at the same position.
- **The median rule:** each new skin pixel takes the median value of all three intensities at the same position.

The idea behind the concept is to address three main issues that affect the skin detection process performance, namely: 1) the sensitivity to illumination intensity, 2) the color difference caused by cameras characteristics, 3) the variation of skin color from person to person that occurs not just between different ethnicities, but within the same ethnicity.

In the following, we provide a numerical example to further illustrate the proposed design. Figure 3 represents three original SFA skin samples, from which the new representative skin samples are carried out. Note that each skin sample is an RGB color image. Figure 4 shows the new constituted min, max, mean and median skin samples.

## 4. METHODOLOGY

In mathematics, ED is the ordinary distance between two points in the Euclidean space. The associated norm is called Euclidean norm whose generalized term is the L2 distance.

In image processing, various image distances have been proposed to evaluate similarities and differences between two images or sub-regions of two images. Of all these, ED in addition to its different variants have been the most suitable for a range of computer vision tasks including human skin detection [23-25].

### 4.1 Euclidean distance

For two vectors in an n dimensional vector space, $U=(x_1, x_2, ... , x_n)$ and $V = (y_1, y_2, ... ,y_n)$, the ED between $U$ and $V$, $d(U,V)$, is given by:

$$d(\mathrm{U},\mathrm{V}) = \sqrt{\left(x_1 - y_1\right)^2 + \left(x_2 - y_2\right)^2 + ... + \left(x_n - y_n\right)^2}$$
$$= \sqrt{\sum_{i=1}^{n}\left(x_i - y_i\right)^2} \qquad (1)$$

In our approach, we define the ED between two color pixels $X(R_x, G_x, B_x)$ and $Y(R_y, G_y, B_y)$, in the three planes Red, Green and Blue as:

$$ED(X,Y) = \sqrt{\left(R_x - R_y\right)^2 + \left(G_x - G_y\right)^2 + \left(B_x - B_y\right)^2} \qquad (2)$$

$(R_x, G_x, B_x)$ and $Y(R_y, G_y, B_y)$ stand for pixels intensities in the RGB color space.

### 4.2 Manhattan distance

Similarly, we define the MD between two color pixels $X(R_x, G_x, B_x)$ and $Y(R_y, G_y, B_y)$ in the RGB color space as:

$$MD(\mathrm{X},\mathrm{Y}) = \left| R_x - R_y \right| + \left| G_x - G_y \right| + \left| B_x - B_y \right| \qquad (3)$$

### 4.3 Similarity measures-based method

The proposed approach is based on the principle of skin color similarity measure, by means of either ED or MD, between each facial image to be segmented and one of the new representative skin samples, on the one hand; and one of the non-skin (ns) original SFA sample, on the other.

Pixels classification to skin and non-skin is therefore made depending on the color similarity measurements made above.

The steps of either ED-based or MD-based skin detection algorithm are displayed in the flowchart of Figure 5. The following describes the methodology for the ED metric.
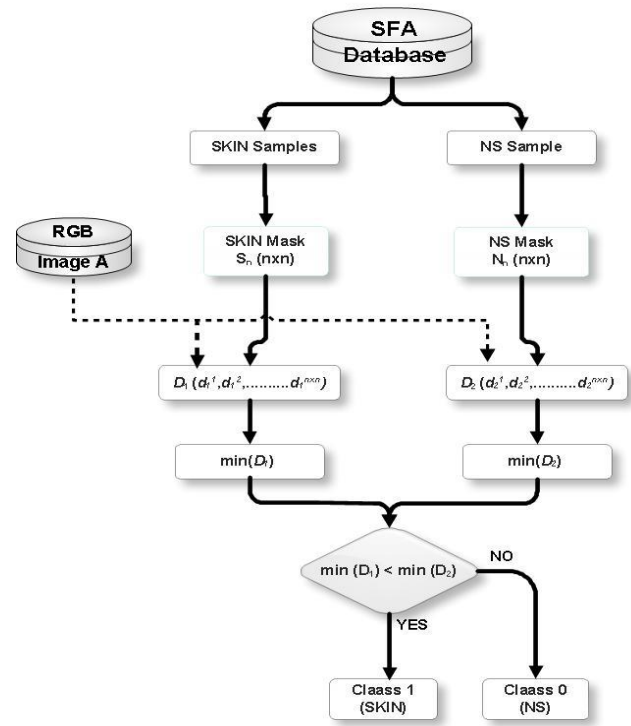


**Figure 5.** Skin color similarity measure based-skin detection algorithm

i. For each test facial color image A to be segmented into skin and non-skin classes, chose two sample masks, $S_n$ and $N_n$, of size $n \times n$ ($n = 1,3,5, ..., 35$), with $S_n$ belonging to the new-SKIN folder containing the new (min, max, mean, median) skin samples, and $N_n$ to the NS SFA folder mentioned in Figure 1.

ii. For every pixel $A_{ij}(R_{ij}, G_{ij}, B_{ij})$ in image A defined on a lattice $L=\{(i, j)/ 1 \leq i \leq M ; 1 \leq j \leq K\}$:

a. Perform ED, $d_1^r$ ($r = 1, ..., n \times n$), between $A_{ij}(R_{ij}, G_{ij}, B_{ij})$ and every skin pixel, $S_n^r$ ($r = 1, ... , n \times n)$, in the new skin mask $S_n$, using Eq. (2), so that gives a total of ($n \times n$) different ED values forming the components of vector $D_1$ such as: $D_1 = [d_1^1, d_1^2, ..., d_1^{n \times n}]$.

b. Perform ED, $d_2^r$ ($r = 1, ..., n \times n$), between $A_{ij}(R_{ij}, G_{ij}, B_{ij})$ and every non-skin pixel, $N_n^r$ ($r=1,..., n \times n)$, in the non-skin mask $N_n$, using Eq. (2), which gives a total of ($n \times n$) different ED values forming the components of vector $D_2$ such as: $D_2 = [d_2^1, d_2^2, ..., d_2^{n \times n}]$.

c. Find the minimum ED values, $\min(D_1)$ and $\min(D_2)$, in both $D_1$ and $D_2$ respectively.

d. Finally, the pixels classification to produce a skin segmented image is carried out according to the following rule:

If     $\min(D1) < \min(D2)$ then Class 1  (skin)
Else If   $\min(D1) > \min(D2)$ then Class 0 (non-skin)    (4)

Following the classification, a binary image emerges in which skin color areas are white while non-skin areas take the black. The same process steps apply to the MD metric, but with Eq. (3) instead of Eq. (2).
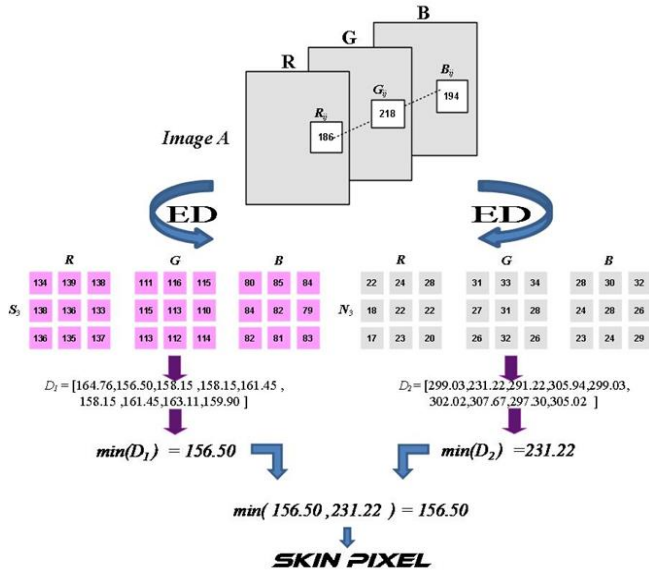
**Figure 6.** Numerical example of the ED-based skin detection method using 3x3 sample masks

Figure 6 shows a numerical example of ED-based skin detection method. In this specific case, ED was evaluated between the RGB pixel, at *ijth* position, $A_{ij}(R_{ij}, G_{ij}, B_{ij})=$ (186,218,194) and every pixel in both RGB skin and non-skin masks of size (3×3) $S_3$ and $N_3$, respectively. The corresponding nine different ED values obtained are retained in vectors $D_1$ and $D_2$, minimum ED values in $D_1$ and $D_2$ are determined, giving the pair $(d_1^2, d_2^2) = (156.50, 231.22)$. According to the classification rule expressed in (2), it goes without saying that $A_{ij}$ (186,218,194) is a skin pixel because $min(D_1) < min(D_2)$ (156.50 < 231.22).

**4.4 Performance assessment**

Both ED-based and MD-based skin detection methods have been assessed by means of the accuracy measure formulated as:

$$Accuracy = \frac{Number\ of\ detected\ segments}{Total\ number\ of\ image's\ segments}(\%) \quad (5)$$

**4.5 Comparison results**

The method comparison, for the SFA database, was conducted by comparing the mean accuracy rates and the standard deviations as well.

The arithmetic mean, $\bar{X}$, of a set of N values can be expressed by:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} x_i \quad (6)$$

The standard deviation, σ, makes it possible to evaluate the dispersion of the measurements around the mean value. It can be expressed by:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(x_i - \bar{X}\right)^2} \quad (7)$$

## 5. EXPERIMENTAL RESULTS

Extensive simulations of the proposed method have been carried out for performance assessment of both ED and MD metrics together with the new representative (min, max, mean, median) skin samples in the six various sample sizes of 1×1, 3×3, 5×5, 7×7, 19×19 and 35×35.

As mentioned earlier, skin detection performance generated by the use of new representative skin samples are compared to those obtained with original SFA skin samples. In this paper, all the three skin samples together with the first of the five ns samples have been considered for ED and MD calculation. For instance, all the three skin samples of size 35×35 (skin-163-1, skin-163-2, skin-163-3) illustrated in Figure 2-d have been used. However, only the ns sample ns-163-1, in Figure 2-e, has been involved in the processing.

**5.1 Data set**
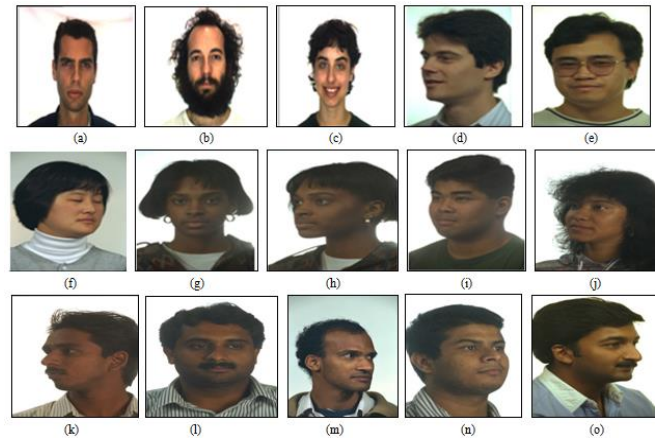


**Figure 7.** Original RGB images from SFA database.
(a) Image1: White skin. (b) Image2: White skin with beard. (c) Image3: White skin with smile. (d) Image4: White skin with deviation to the right. (e) Image5: Yellow skin with glasses. (f) Image6: Yellow skin with deviation to the left and eyes closed. (g) Image7: Black skin. (h) Image8: Black skin with deviation to the right. (i) Image9: Black skin with deviation to the right. (j) Image10: Black skin with deviation to the right. (k) Image11: Brown skin with deviation to the right. (l) Image12: Brown skin. (m), (n) and (o) Images 13, 14 and 15: Brown skin with different deviations to the left



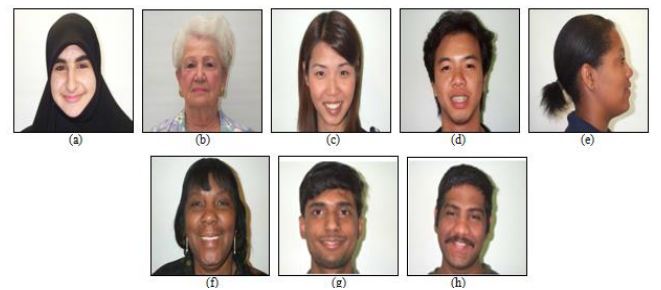**Figure 8.** Original RGB images from UTD database.
(a) Image1: White skin with Islamic headscarf. (b) Image2: White skin with makeup and white hair. (c) Image3: Female yellow skin with smile. (d) Image4: Male yellow skin with smile. (e) Image5: Black skin with profile position. (f) Image6: Black skin with smile. (g) Image7: Brown skin with smile. (h) Image8: Brown skin with smile, mustache and a head tilt

The data used in this study are taken from both SFA and UTD [19, 20] face databases, representing facial images within different skin colors, positions and lightings. Figures 7 and 8 illustrate original RGB images corresponding to each one of the two databases, respectively.

## 5.2 Simulations with SFA facial images

### 5.2.1 SFA original skin samples simulations

The first simulation experiments aim at assessing the skin detection accuracy for a given image from SFA database, on the basis of either ED or MD color similarity measure. Results are reported in terms of mean accuracies. Therefore, Table 1 shows the mean accuracies per sample size (**MASS**) averaged over the six different SFA sample sizes considered. The corresponding standard deviation ($\sigma$) results are reported in Table 2. Finally, the mean accuracies per image (**MAI**) averaged over the fifteen considered SFA images are reported in Table 3.

**Table 1.** Mean accuracies per sample size (**MASS**) with SFA database

| Metric | Sample | MASS (%) | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | 1×1 | 3×3 | 5×5 | 7×7 | 19×19 | 35×35 |
| ED | Skin-1 | 88.89 | 90.96 | 91.77 | 91.63 | 90.89 | 90.58 |
| | Skin-2 | 89.86 | 92.19 | 92.78 | 92.78 | 92.21 | 90.76 |
| | Skin-3 | 91.16 | 92.44 | 92.63 | 92.55 | 92.45 | 91.04 |
| MD | Skin-1 | 88.43 | 90.95 | 91.75 | 91.82 | 91.06 | 89.65 |
| | Skin-2 | 89.58 | 92.15 | 92.81 | 92.91 | 92.37 | 91.42 |
| | Skin-3 | 90.79 | 92.41 | 92.45 | 92.58 | 92.52 | 91.22 |

**Table 2.** Standard deviation ($\sigma$) results per sample size, with SFA database

| Metric | Sample | σ (%) | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | 1×1 | 3×3 | 5×5 | 7×7 | 19×19 | 35×35 |
| ED | Skin-1 | 5.58 | 4.12 | 3.88 | 3.61 | 3.73 | 3.71 |
| | Skin-2 | 4.37 | 3.64 | 3.24 | 3.15 | 3.42 | 3.93 |
| | Skin-3 | 4.27 | 3.92 | 3.78 | 3.73 | 3.27 | 3.80 |
| MD | Skin-1 | 5.77 | 4.15 | 3.87 | 3.73 | 3.75 | 4.58 |
| | Skin-2 | 4.64 | 3.71 | 3.26 | 3.19 | 3.46 | 3.92 |
| | Skin-3 | 4.48 | 3.92 | 4.00 | 3.76 | 3.54 | 3.97 |

**Table 3.** Mean accuracies per image (**MAI**) with SFA database

| Images | ED: MAI (%) | | | MD: MAI (%) | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Skin-1 | Skin-2 | Skin-3 | Skin-1 | Skin-2 | Skin-3 |
| 1 | 96.06 | 96.45 | 96.15 | 95.95 | 96.49 | 96.14 |
| 2 | 94.24 | 94.32 | 94.92 | 94.44 | 94.41 | 94.92 |
| 3 | 97.04 | 97.28 | 97.21 | 97.12 | 97.35 | 97.30 |
| 4 | 92.81 | 93.62 | 94.24 | 92.59 | 93.69 | 94.26 |
| 5 | 91.60 | 92.23 | 92.39 | 91.57 | 92,20 | 92.76 |
| 6 | 95.32 | 95.05 | 96.37 | 95.49 | 95.56 | 96.61 |
| 7 | 86.08 | 89.68 | 89.36 | 86.11 | 89.75 | 89.62 |
| 8 | 88.59 | 90.38 | 89.52 | 88.48 | 90.52 | 89.51 |
| 9 | 90.03 | 92.82 | 93.48 | 89.81 | 93.04 | 93.32 |
| 10 | 93.96 | 94.07 | 94.45 | 93.84 | 94.20 | 94.43 |
| 11 | 88.03 | 87.70 | 85.38 | 88.20 | 87.64 | 85.24 |
| 12 | 84.38 | 86.22 | 87.61 | 83.89 | 86.24 | 87.22 |
| 13 | 90.14 | 90.64 | 91.04 | 90.025 | 90.67 | 91.01 |
| 14 | 87.25 | 90.08 | 90.97 | 85.91 | 90.24 | 90.87 |
| 15 | 86.31 | 86.04 | 87.57 | 85.80 | 86.12 | 86.74 |

Whatever skin samples are used, two general remarks can be made:

- Firstly, it may be observed that both ED and MD perform

well with the possibility of having skin detection accuracies higher than 90% irrespective the skin color. Second, accuracy rates vary from one samples size to another and from one color skin to another.

- By comparing the results obtained with both used metrics, ED achieves higher detection rates with the lowest sample sizes (from 1×1 to 5×5) and lower rates with the highest sizes (7×7, 19×19 and 35×35).

### 5.2.2 SFA representative skin samples simulations

These simulations were conducted to see any major improvement in the process of segmentation into skin and non-skin classes of any SFA facial image, using jointly multiple SFA skin samples within the specific (min, max, mean, median) combinations.

So, Table 4 illustrates the **MASS** values averaged over the fifteen considered images, Tables 5 and 6 exhibit the **MAI** values averaged over the six different sample sizes, related to ED and MD metrics respectively and Table 7 illustrates those of the standard deviation ($\sigma$) results for each sample size.

**Table 4.** Mean accuracies per sample size (**MASS**) related to the new representative samples with SFA database

| Metric | Sample rule | MASS (%) | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | 1×1 | 3×3 | 5×5 | 7×7 | 19×19 | 35×35 |
| ED | Min | 87.52 | 87.07 | 86.65 | 86.64 | 85.17 | 84.12 |
| | Max | 85.93 | 89.43 | 89.84 | 90.14 | 89.51 | 87.90 |
| | Mean | 89.65 | 90.56 | 90.79 | 90.75 | 89.55 | 88.39 |
| | Median | 89.93 | 91.86 | 92.35 | 92.32 | 91.12 | 89.34 |
| MD | Min | 86.55 | 87.07 | 86.86 | 87.16 | 86.52 | 84.98 |
| | Max | 86.21 | 89.26 | 89.85 | 90.27 | 89.62 | 88.54 |
| | Mean | 89.99 | 90.62 | 90.87 | 90.90 | 89.93 | 89.00 |
| | Median | 89.13 | 91.67 | 86.55 | 92.34 | 91.34 | 89.99 |

**Table 5.** Mean accuracies per image (**MAI**) related to the four rules with SFA database: ED metric

| Images | ED: MAI (%) | | | |
| --- | --- | --- | --- | --- |
| | Min | Max | Mean | Median |
| 1 | 95.59 | 94.96 | 95.64 | 95.95 |
| 2 | 91.73 | 94.37 | 94.81 | 94.76 |
| 3 | 95.62 | 96.46 | 96.87 | 97.24 |
| 4 | 88.15 | 90.33 | 91.52 | 93.07 |
| 5 | 87.47 | 88.64 | 91.11 | 91.14 |
| 6 | 75.08 | 95.94 | 96.81 | 96.57 |
| 7 | 83.43 | 82.07 | 84.58 | 87.24 |
| 8 | 86.73 | 85.61 | 86.73 | 89.24 |
| 9 | 91.28 | 86.38 | 88.11 | 91.16 |
| 10 | 88.41 | 93.16 | 94.31 | 93.76 |
| 11 | 84.80 | 87.61 | 88.24 | 88.13 |
| 12 | 80.39 | 80.00 | 80.98 | 84.23 |
| 13 | 80.48 | 89.68 | 89.74 | 90.31 |
| 14 | 87.45 | 84.82 | 86.47 | 89.32 |
| 15 | 76.32 | 81.84 | 83.32 | 85.18 |

An examination of the results in Tables 4, 6 and 7 reveals no significant improvement in the detection accuracies. In more details, the min rule for both ED and MD metric favors the lowest size making representative samples size 1×1 lead performance in terms of average accuracy rates. Nevertheless, the three other rules appear to provide samples sizes ranking consistent with those of each original SFA skin sample. Also, with respect to individual image average rates grading, results of the max, mean and median rules comply with those of the three skin samples separately. The min rule, however, provides a slightly different ranking. Finally, it is important to note that

the median rule sample performance, for all considered sizes, is far better compared to those of min, max and mean rule samples.

By comparing the results obtained with ED and MD metrics, both achieve higher detection rates at the lowest sample sizes (from 1×1 to 5×5) and lower rates at the highest sizes (7×7, 19×19 and 35×35).

**Table 6.** Mean accuracies per image (**MAI**) related to the four rules with SFA database: MD metric

| | MD: MAI (%) | | | |
|---|---|---|---|---|
| Images | Min | Max | Mean | Median |
| 1 | 96.53 | 95.17 | 95.82 | 96.00 |
| 2 | 93.31 | 93.61 | 94.82 | 94.70 |
| 3 | 96.78 | 96.63 | 97.09 | 97.38 |
| 4 | 88.60 | 92.63 | 91.76 | 93.10 |
| 5 | 86.64 | 82.17 | 84.85 | 87.18 |
| 6 | 87.80 | 85.96 | 87.07 | 89.33 |
| 7 | 91.67 | 86.45 | 88.26 | 91.10 |
| 8 | 89.44 | 92.94 | 94.12 | 93.61 |
| 9 | 85.26 | 87.40 | 88.40 | 88.28 |
| 10 | 80.18 | 79.88 | 81.16 | 83.96 |
| 11 | 79.49 | 89.92 | 90.09 | 90.44 |
| 12 | 88.23 | 84.83 | 86.71 | 89.11 |
| 13 | 76.34 | 81.13 | 83.26 | 84.78 |
| 14 | 87.88 | 88.67 | 91.58 | 91.14 |
| 15 | 84.72 | 96.10 | 96.96 | 96.86 |

**Table 7.** Standard deviation (σ) results per sample size related to the new representative samples, with SFA database

| Metric | Sample rule | σ (%) | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1×1 | 3×3 | 5×5 | 7×7 | 19×19 | 35×35 |
| ED | Min | 6.18 | 6.47 | 6.23 | 6.12 | 6.41 | 6.55 |
| | Max | 7.22 | 5.41 | 5.26 | 5.00 | 4.79 | 5.12 |
| | Mean | 5.71 | 5.11 | 4.89 | 4.85 | 4.95 | 5.07 |
| | Median | 5.20 | 4.03 | 3.70 | 3.56 | 3.87 | 4.52 |
| MD | Min | 7.88 | 6.70 | 6.70 | 6.98 | 7.10 | 8.41 |
| | Max | 7.79 | 5.69 | 5.37 | 5.02 | 5.00 | 5.29 |
| | Mean | 5.61 | 5.14 | 4.92 | 4.92 | 4.93 | 4.91 |
| | Median | 5.71 | 4.16 | 3.73 | 3.58 | 3.84 | 4.54 |

## 5.3 Experiments with UTD facial images

### 5.3.1 Original SFA skin samples simulations

Both ED and MD metric-based skin detection have been applied to eight images belonging to UTD face database, representing various human colour skins, ages and poses as well. The specific purpose of these simulations is assessing the feasibility of accurate segmentation of any color face image outside SFA by means of skin samples belonging to SFA.

**Table 8.** Mean accuracies per sample size (**MASS**) with UTD database

| Metric | Sample | MASS (%) | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1×1 | 3×3 | 5×5 | 7×7 | 19×19 | 35×35 |
| ED | Skin-1 | 90.68 | 89.84 | 88.59 | 88.99 | 88.14 | 89.40 |
| | Skin-2 | 90.16 | 90.86 | 90.78 | 90.91 | 90.94 | 90.86 |
| | Skin-3 | 90.83 | 90.46 | 89.79 | 89.37 | 89.16 | 89.80 |
| MD | Skin-1 | 88.27 | 88.58 | 89.00 | 88.43 | 88.26 | 89.61 |
| | Skin-2 | 90.59 | 90.81 | 90.75 | 90.87 | 90.77 | 91.28 |
| | Skin-3 | 90.72 | 90.45 | 89.91 | 89.49 | 89.11 | 89.95 |

Similarly, results in terms of **MASS** together with their corresponding standard deviation (σ) are illustrated in Tables 8 and 9, respectively. The **MAI** are shown in Table 10. Again,

we can state that ED and MD have continued to function well when comparing colour skin data belonging to different facial databases with average detection rates, in general, higher than 88%, but can exceed 90% according to the skin sample used in addition to the size and the skin colour considered. In this case, however, sample sizes performance in terms of mean accuracy rates are close to each other with a slight advantage for smaller sizes 1×1 and 3×3. By comparing the results obtained with both used metrics, ED and MD achievements are almost equivalent.

**Table 9.** Standard deviation (σ) results per sample size with UTD database

| Metric | Sample | σ (%) | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1×1 | 3×3 | 5×5 | 7×7 | 19×19 | 35×35 |
| ED | Skin-1 | 5.79 | 6.42 | 6.26 | 6.73 | 7.00 | 5.15 |
| | Skin-2 | 5.78 | 5.28 | 5.17 | 5.12 | 4.91 | 4.04 |
| | Skin-3 | 5.20 | 5.56 | 5.36 | 5.42 | 5.22 | 4.28 |
| MD | Skin-1 | 10.63 | 9.04 | 7.93 | 7.92 | 7.11 | 4.93 |
| | Skin-2 | 5.34 | 5.28 | 5.14 | 5.09 | 4.61 | 3.65 |
| | Skin-3 | 5.11 | 5.48 | 5.36 | 5.36 | 5.18 | 4.65 |

**Table 10.** Mean accuracies per image (**MAI**) with UTD database

| | ED: MAI (%) | | | MD: MAI (%) | | |
|---|---|---|---|---|---|---|
| Images | Skin-1 | Skin-2 | Skin-3 | Skin-1 | Skin-2 | Skin-3 |
| 1 | 93.22 | 91.21 | 91.35 | 93.17 | 90.94 | 91.43 |
| 2 | 80.01 | 89.25 | 91.92 | 73.66 | 89.86 | 92.09 |
| 3 | 87.47 | 87.12 | 83.36 | 87.59 | 87.41 | 85.48 |
| 4 | 93.39 | 94.69 | 94.52 | 94.74 | 94.89 | 92.62 |
| 5 | 91.19 | 92.51 | 89.86 | 91.56 | 90.47 | 89.79 |
| 6 | 80.31 | 80.91 | 80.78 | 80.13 | 83.19 | 81.04 |
| 7 | 94.17 | 95.17 | 93.50 | 94.09 | 95.04 | 93.44 |
| 8 | 94.43 | 95.14 | 93.90 | 94.60 | 94.98 | 93.61 |

### 5.3.2 Representative skin samples simulations

This is the most important set of simulations. In Table 11 are reported the **MASS** values averaged over the six different sample sizes considered for the new representative (min, max, mean, median), whereas Tables 12 and 13 present the **MAI** averaged over the eight considered facial images from UTD, related to Ed and MD respectively. The corresponding standard deviation values (σ) are listed in Table 14.

Here too, with the exception of the mean rule, the lowest sizes representative samples 1×1 and 3×3 yield the best results on average. But in all cases, the samples sizes ranking remains consistent with that obtained with each individual skin sample, that is to say close to each other. Also, with respect to individual image grading, results provided by all rules comply with those of the three skin samples separately.

**Table 11.** Mean accuracies per sample size (**MASS**) related to the four rules with UTD database

| Metric | Sample rule | MASS (%) | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1×1 | 3×3 | 5×5 | 7×7 | 19×19 | 35×35 |
| ED | Min | 82.33 | 82.44 | 81.67 | 81.62 | 82.08 | 81.68 |
| | Max | 88.38 | 89.68 | 91.17 | 90.55 | 90.62 | 90.87 |
| | Mean | 92.01 | 92.19 | 91.97 | 91.70 | 90.92 | 91.46 |
| | Median | 91.47 | 91.47 | 90.68 | 90.40 | 91.39 | 90.85 |
| MD | Min | 85.72 | 86.13 | 85.42 | 85.61 | 85.10 | 84.98 |
| | Max | 88.43 | 89.35 | 90.92 | 90.10 | 90.84 | 90.77 |
| | Mean | 91.93 | 92.08 | 91.92 | 91.78 | 91.47 | 91.43 |
| | Median | 92.21 | 91.48 | 90.68 | 90.49 | 91.37 | 91.14 |

**Table 12.** Mean accuracies per image (**MAI**) related to the four rules with UTD database: ED metric

| | ED: MAI (%) | | | |
|---|---|---|---|---|
| **Images** | **Min** | **Max** | **Mean** | **Median** |
| 1 | 91.19 | 91.76 | 92.72 | 93.81 |
| 2 | 69.80 | 84.83 | 92.59 | 86.86 |
| 3 | 79.83 | 84.83 | 90.90 | 89.19 |
| 4 | 70.90 | 93.62 | 94.81 | 95.27 |
| 5 | 83.53 | 93.18 | 93.77 | 92.83 |
| 6 | 75.85 | 79.50 | 79.82 | 81.00 |
| 7 | 92.60 | 93.56 | 94.34 | 94.51 |
| 8 | 92.05 | 93.57 | 94.73 | 94.87 |

**Table 13.** Mean accuracies per image (**MAI**) related to the four rules with UTD database: MD metric

| | MD: MAI (%) | | | |
|---|---|---|---|---|
| **Images** | **Min** | **Max** | **Mean** | **Median** |
| 1 | 92.10 | 91.34 | 92.85 | 93.78 |
| 2 | 82.55 | 84.13 | 92.68 | 87.66 |
| 3 | 80.74 | 91.53 | 90.84 | 89.92 |
| 4 | 82.82 | 93.57 | 94.91 | 95.43 |
| 5 | 84.48 | 93.25 | 93.74 | 92.68 |
| 6 | 76.21 | 79.54 | 80.15 | 80.89 |
| 7 | 92.70 | 93.57 | 94.20 | 94.65 |
| 8 | 92.34 | 93.61 | 94.78 | 94.82 |

By comparing, the median rule has worked best in ED and MD-based method for SFA providing mean rates over 92%, whilst the mean rule showed more effective for UTD allowing

an increase in the mean detection rates up to 92%.

Finally, the skin detection suggests to highlight the following points: 1) skin segmentation of face color images outside SFA based on SFA skin samples, by means of both ED and MD has showed efficiency and accuracy as well, especially with representative mean and median skin samples, 2) as regards the stability, both approaches yield almost similar low standard deviations (except for the min rule), 3) unlike the results obtained with SFA, the lowest samples sizes 1×1 and 3×3 achieve the best detection rates in UTD database.

Finally, Figures 9 and 10 display examples of face image segmentation in SFA and UTD databases, respectively. Both ED and MD-based methods have been used together with the new median skin sample of size 3 × 3. As we can see, the proposed scheme provides facial segmented images, inside and outside SFA, with an appropriate quality.

**Table 14.** Standard deviation (σ) per sample size related to the four rules with UTD database

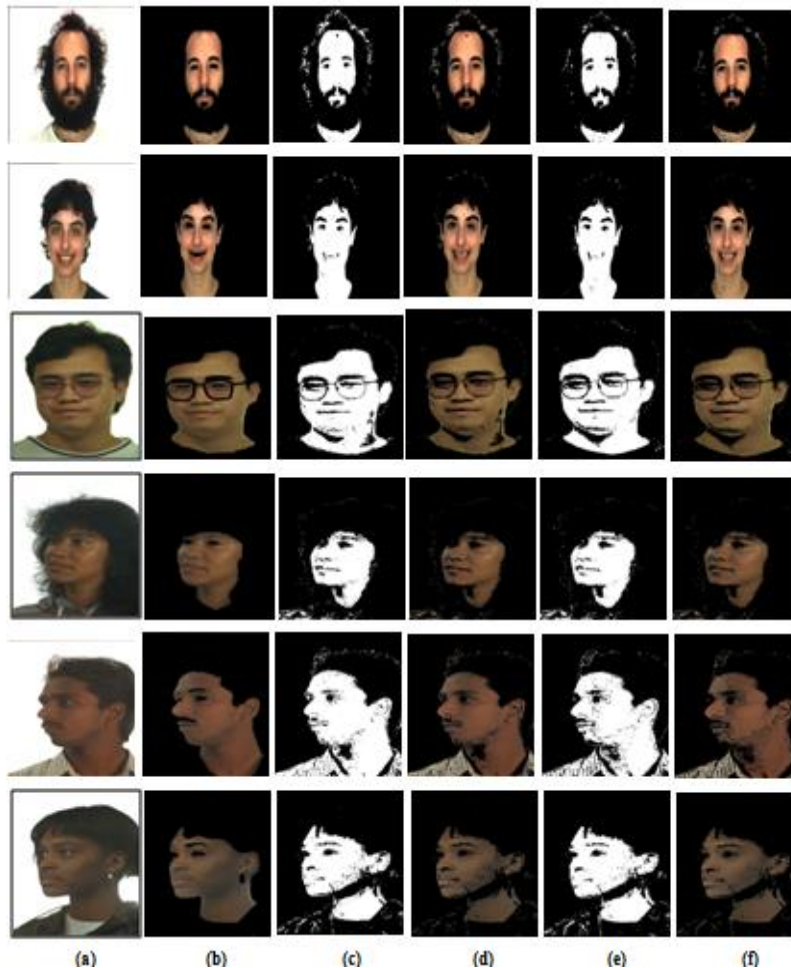| Metric | Sample rule | σ (%) | | | | | |
|---|---|---|---|---|---|---|---|
| | | **1×1** | **3×3** | **5×5** | **7×7** | **19×19** | **35×35** |
| ED | Min | 12.19 | 9.82 | 10.25 | 9.80 | 8.45 | 7.54 |
| | Max | 6.90 | 6.72 | 4.84 | 5.38 | 5.09 | 3.94 |
| | Mean | 5.26 | 5.18 | 5.12 | 5.09 | 5.07 | 4.43 |
| | Median | 5.22 | 5.22 | 5.72 | 5.84 | 4.30 | 4.67 |
| MD | Min | 8.04 | 6.59 | 6.79 | 6.51 | 6.95 | 5.20 |
| | Max | 6.60 | 7.03 | 4.93 | 6.02 | 5.02 | 3.88 |
| | Mean | 5.21 | 5.16 | 5.08 | 4.99 | 4.78 | 4.17 |
| | Median | 5.25 | 5.21 | 5.69 | 5.83 | 4.51 | 4.34 |



**Figure 9.** Comparison between ED and MD metrics in SFA database. (a) Original images (ORG). (b) Ground truths (GT). MD-based method: (c) ORG image and (d) GT image. ED-based method: (e) ORG image and (f) GT image
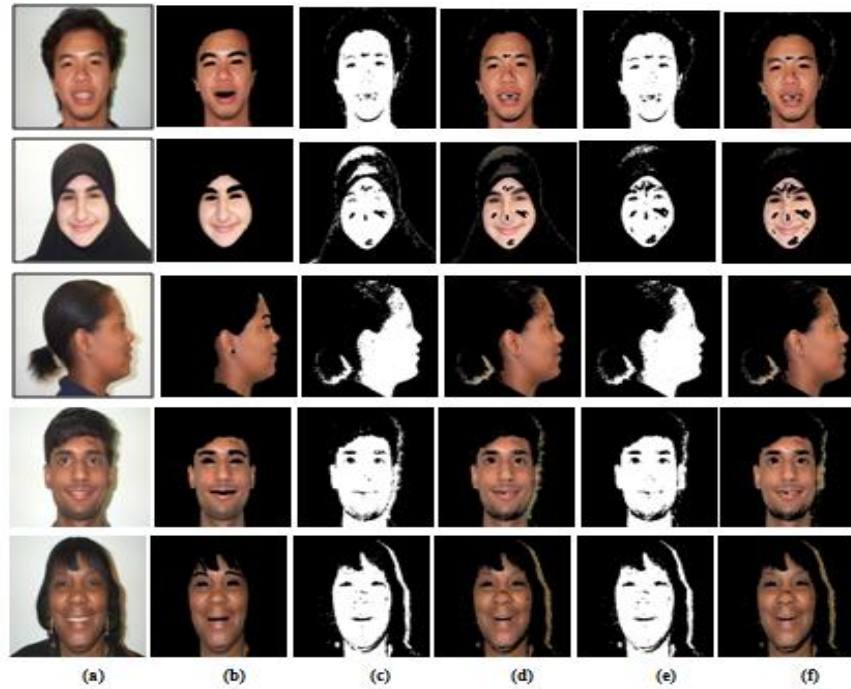
**Figure 10.** Comparison between ED and MD metrics in UTD database. (a) Original images (ORG). (b) Ground truths (GT). MD-based method: (c) ORG image and (d) GT image.
ED-based method: (e) ORG image and (f) GT image

## 6. CONCLUSION

Skin detection is very challenging because of the differences in illumination or cameras characteristics, or the range of skin colors due to different ethnicities, and many other variations. A skin detection approach has been implemented with the use of SFA facial database in conjunction with the two well known ED and MD metrics. Both metrics performances have been assessed within another facial database, here UTD, with an effective contribution of SFA.

The most interesting aspect of SFA database resides in the multiple samples provided, which in various combinations with each other allow creating new samples that could be useful and more effective. This particular aspect has been addressed by creating four new representative skin samples according to the four rules of minimum, maximum, mean and median. The main idea is to give new universal skin samples that can address the issues of illumination sensitivity and skin color variation even within the same ethnicity. Indeed, the mean and median representative skin samples results outperform those of individual skin samples, allowing a significant increase of mean detection rates in the UTD database, up to 92% for both ED and MD metrics. ED and MD metrics to skin segmentation of SFA images with individual skin samples has showed efficient, accurate and stable, yielding mean accuracies over than 92%.

These satisfactory results open room to further improvements to face segmentation, based on SFA samples, with the use of more precise metrics and more robust models like MRFs, HOG, etc.

## REFERENCES

[1] Chaves-González, J.M., Vega-Rodríguez, M.A., Gómez-Pulido, J.A., Sánchez-Pérez, J.M. (2010). Detecting skin in face recognition systems: A colour spaces study. Digital Signal Processing, 20(3): 806-823. https://doi.org/10.1016/j.dsp.2009.10.008

[2] Yang, X.Y., Liang, N.N., Zhou, W., Lu, H.M. (2020). A face detection method based on skin color model and improved AdaBoost algorithm. Traitement du Signal, 37(6): 929-937. https://doi.org/10.18280/ts.370606

[3] Hasan, M.K., Ahsan, M., Newaz, S.H., Lee, G.M. (2021). Human face detection techniques: A comprehensive review and future research directions. Electronics, 10(19): 2354. https://doi.org/10.3390/electronics10192354

[4] Fawzi, L.M., Ameen, S.Y., Alqaraawi, S.M., Dawwd, S.A. (2018). Embedded real-time video surveillance system based on multi-sensor and visual tracking. Appl. Math. Infor. Sci, 12: 345-359. https://doi.org/10.1007/s11263-016-0957-7

[5] Wang, Y.N., Yang, Y.M., Zhang, P.Y. (2020). Gesture feature extraction and recognition based on image processing. Traitement du Signal, 37(5): 873-880. https://doi.org/10.18280/ts.370521

[6] Bosco, J., Jayakumar, S.K.V. (2018). A study on web images retrieval using content based image retrieval methods. Journal. of Computer Science and Mobile Computing, 6(7): 128-137.

[7] Meng, W.L., Mao, C.Z., Zhang, J., Wen, J., Wu, D.H. (2019). A fast recognition algorithm of online social network images based on deep learning. Traitement du Signal, 36(6): 575-580. https://doi.org/10.18280/ts.360613

[8] Adjabi, I., Ouahabi, A., Benzaoui, A., Taleb-Ahmed, A. (2020). Past, present, and future of face recognition: A review. Electronics, 9(8): 1188. https://doi.org/10.3390/electronics9081188

[9] Kakumanu, P., Makrogiannis, S., Bourbakis, N. (2007). A survey of skin-color modeling and detection methods.

Pattern Recognition, 40(3): 1106-1122. https://doi.org/10.1016/j.patcog.2006.06.010

[10] Al Naffakh, H.A.H., Ghazali, R., El Abbadi, N.K. (2021). Statistical survey and comprehensive review on human skin detection. Bulletin of Electrical Engineering and Informatics, 10(1): 118-128. https://doi.org/10.11591/eei.v10i1.2486

[11] Naji, S., Jalab, H.A., Kareem, S.A. (2019). A survey on skin detection in colored images. Artificial Intelligence Review, 52(2): 1041-1087. https://doi.org/10.1007/s10462-018-9664-9

[12] Kumar, A., Malhotra, S. (2015). Pixel-based skin color classifier: A review. International Journal of Signal Processing, Image Processing and Pattern Recognition, 8(7): 283-290. http://dx.doi.org/10.14257/ijsip.2015.8.7.27

[13] Kolkur, S., Kalbande, D., Shimpi, P., Bapat, C., Jatakia, J. (2017). Human skin detection using RGB, HSV and YCbCr color models. Proceedings of the International Conference on Communication and Signal Processing 2016 (ICCASP 2016). https://doi.org/10.2991/iccasp-16.2017.51

[14] Mohammed, S.G., Majeed, A.H., Aldujaili, A., Hassan, E.K., Abdul-Jabbar, S.S. (2020). Image segmentation for skin detection. Journal of Southwest Jiaotong University, 55(1): 1-11. https://doi.org/10.35741/issn.0258-2724.55.1.17

[15] Poudel, R.P., Nait-Charif, H., Zhang, J.J., Liu, D. (2012). Region-based skin color detection. In VISAPP (1) VISAPP 2012-Proceedings of the International Conference on Computer Vision Theory and Applications, 1: 301-306. VISAPP.

[16] Bilal, S., Akmeliawati, R., Salami, M.J.E., Shafie, A.A. (2015). Dynamic approach for real-time skin detection. Journal of Real-Time Image Processing, 10(2): 371-385. https://doi.org/10.1007/s11554-012-0305-2

[17] Mahmoodi, M.R. (2017). Fast and efficient skin detection for facial detection. arXiv preprint arXiv:1701.05595.

[18] Mortazavi, M. (2019). An improved human skin detection and localization by using machine learning techniques in RGB and YCbCr color spaces (No. e27488v1). PeerJ Preprints. https://doi.org/10.7287/peerj.preprints.27488v1

[19] Casati, J.P.B., Moraes, D.R., Rodrigues, E.L.L. (2013). SFA: A human skin image database based on FERET and AR facial images. In IX workshop de Visao Computational, Rio de Janeiro.

[20] Parker, A. (2016) The University of Texas Dallas, 22 April [online]. https://utdallas.app.box.com/v/facedatabase, accessed on 3 May 2018.

[21] Martinez, A., Benavente, R. (1998). The AR Face Database: CVC Technical Report, 24.

[22] Philips, P.J., Moon, H., Risvi, S.A., Rauss, P.J. (2000). The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(10): 1090-1104. https://doi.org/10.1109/34.879790

[23] Hasnat, A., Halder, S., Bhattacharjee, D., Nasipuri, M., Basu, D.K. (2013). Comparative study of distance metrics for finding skin color similarity of two color facial images. ACER: New Taipei City, Taiwan, 99-108. https://doi.org/10.5121/CSIT.2013.3210

[24] Sharma, M.K., Maurya, R., Shukla, A.S., Gupta, P.R. (2012). Skin infection identification using color and Euclidean distance algorithm. In International Conference on Contemporary Computing, CCIS, 306: 471-480. https://doi.org/10.1007/978-3-642-32129-0_47

[25] Indra, D. (2019). Skin detection using color distance measurement and thresholding. International Journal of Engineering and Advanced Technology (IJEAT), 8(5): 1441-1443.