A Fuzzy Approach to Skin Color Detection

Abstract. Skin detection is a useful tool for many subsequent image processing tasks, such as face detection, face tracking or human-computer interfaces. In many cases, using a skin color detection scheme is related to a proper representation of the color space chosen to interpret image information. In this work, we propose a fuzzy system for detecting skin in the RGB color space, so that each color plane is modeled using fuzzy sets. As a consequence, the development of the membership functions, the inference system and the defuzzifying process are described throughout this paper. The experiments have been performed with the XM2VTS and VALID face databases and the results show that our system has high detection rates and low false positive rates compared to other similar skin color detectors.

1 Introduction

One of the most active research areas in computer vision is face recognition. Before a recognition algorithm is applied, faces must have been segmented and located using some detection method. Therefore, detecting a face is the first step for the recognition process to be efficiently completed. From this detection, one can utilize different methods to recognize users using facial features.

There are many approaches to face detection, such as knowledge- and appearance-based methods, feature invariant algorithms or template matching techniques (for further details, see [1], [2], [3]). In this work, we shall focus on a pre-processing stage for detecting faces in color images. To do this, we assume that the problem consists of detecting skin regions in the image, and then the segmented skin clusters will be further processed to validate they belong to face regions.

Many recent proposals are based on the underlying idea of representing the skin color in an optimal color space (such as RGB, YIQ or HSV) by means of the so-called skin cluster (see [2], [3], [4]). Color information is an efficient tool for identifying facial areas if the skin color model can be properly adapted for different lighting environments. This fact leads to avoid the use of RGB in practical systems, since the red, green and blue components are highly correlated and dependent on lighting conditions. However, the RGB space corresponds most closely with the physical sensors for colored light such as the cones in the human eye or red, green and blue filters in most color CCD sensors. In addition, using a RGB model should simplify the design of any algorithm, since there is no need to transform the color spaces.

As a consequence, our proposal is based on the RGB color space, using fuzzy sets instead of traditional color spaces; thus, input color variables are fuzzified and, as

a result, a pixel is classified after the defuzzifying process is completed. To do this, in Section 2 we will revise some skin color detection algorithms and then the designing of the fuzzy system for detecting skin is described in Section 3. After that, the experiments are outlined and discussed in Section 4 and, finally, we shall conclude with some remarks to our work in Section 5.

2   Skin Color Detection

As described before, the main goal of skin color modeling is to define a decision rule that accurately separates the pixels of an image that correspond with skin from any other elements in a scene. There are a wide variety of methods to perform this task (see [1], [2] for a complete review); in this section some of them are shown.

In general, a method to develop a skin classifier consists of defining the bounding limits of the regions corresponding to a color which belongs to the skin, by means of some numerical (and often empirical) rules, i.e., defining explicitly the skin regions. It is clear that the efficiency of the color segmentation stage highly depends on the color space chosen. Color spaces with separated luminance and chrominance components, like HSV or YCbCr, seem to be appropriate for face detection (see, for instance, [1], [3], [4]). However, Albiol et al. [2] shown that if an optimum skin detector is designed for every color space, their performance will be almost the same.

The main advantage of this method is the simplicity of the skin detection rules, which permit a fast classification of pixels. The major drawback is the difficulty in finding both robust color spaces and reliable decision rules for a convenient skin detection. Nevertheless, for any color space, the problem to be solved can be summarized as: finding the optimal bounding limits for robust skin color detection.

On the other hand, there is a set of techniques which estimate the distribution of skin color by means of a training phase, without deriving an explicit model. These methods are often referred as non-parametric skin models. Generally, they create a skin probability map (SPM) [5], which consists of assigning a probability to each point in different color spaces. Finally, other methods include parametric skin distribution models, such as the Gaussian skin color model or the elliptic boundary model [1], [2].

Non parametric methods achieve low computation times for both training and classifying skin pixels, and they do not depend on the skin distribution shape or on the selection of the color space, either. However, they require a huge storage capacity and that the training must be performed with a representative dataset.

Parametric methods can also be fast, as they make it possible to interpolate data and to generalize the training phase for incomplete data. In spite of this, their performance highly depends on the skin distribution shape and many parametric models ignore the static non-skin color. Subsequently, this fact leads to high false positive rates compared with non-parametric methods.

As a consequence, our proposal consists of taking the main advantages of explicitly defined skin models, i.e., the simplicity of their decision rules, and combining them with the main benefits of non-parametric skin models, that is to say, low computation times for training and classifying. To do this, our method includes a

fuzzy model for detecting skin in color images. This model is explained in the following section.

3   A Fuzzy System for Detecting Skin

In any image segmentation scheme, there is a high level of uncertainty for a classifier to automatically obtain an optimum segmentation. This fact can be also extended to face detection and, in particular, to skin color segmentation. Thus, we find that applying fuzzy theory can be a convenient way to obtain good detection rates, since a fuzzy set-theoretic model provides a mechanism to represent and manipulate uncertainty within an image.

Color image segmentation using fuzzy classification is a pixel-based segmentation method. This method assigns a color class to each pixel of an input image by applying a set of fuzzy rules on it. We can use this approach to achieve our goal: a pixel can be classified as 'skin' or 'non-skin' according to a set of fuzzy rules extracted from a training stage using different color spaces. To do this, each color plane will be considered as a fuzzy set, so that the skin detection is performed through fuzzy functions representing the membership degree of each pixel to the different classes. Let us define some terms first.

Definition 1. A color vector $c \in C$, where C is a color space, of a pixel p is defined as a l-tuple of color components $c(p) = \{c_1(p), c_2(p),\ldots,c_l(p)\}$, where $c_i(p)$, for $i = 1,2,\ldots l$, may have N different values.

For instance, for the RGB color space, $l = 3$ and $c(p) = \{r_p, g_p, b_p\}$, where $r_p, g_p, b_p \in [0, 255]$.

Definition 2. Given a color image I of size $W = n \times m$ pixels, where each pixel is defined by a color vector c in a color space C, so that $c(p) = \{c_1(p), c_2(p),\ldots,c_l(p)\}$, $\forall\, p \in I$, the histogram of C, $H(C)$, is defined as a $q \times l$ array $H(C) = \{f_1, f_2, \ldots, f_l\}$, such that each f_i is the frequency vector, using q bins, of the color component c_i, for $i = 1,2,\ldots l$, on the image I.

As a result, the value of each bin is the number of pixels in image I having the color c_i. If $H(C)$ is normalized by W, then $H(C)$ takes the color space C into the interval $[0, 1]$; that is, $H(C)$ represents the probability distribution of each colour c_i to be present in image I. According to Zadeh's theory [6], a fuzzy set is a pair (A, m) where A is a set and $m: A \rightarrow [0, 1]$. This can be applied to the color histogram, where the fuzzy set can be defined as the pair (C, H), where C is the color space and $H: C \rightarrow [0, 1]$ is the normalized histogram. For each $c \in C$, $H(c)$ is the grade of membership of c, so that $c \in (C, H) \leftrightarrow c \in C\ AND\ H(c) \neq 0$.

Some previous works have modeled color spaces by means of fuzzy sets and relations. For the HSV space and its variants, the hue component has been defined through a fuzzy representation to take into account the non-uniformity of colors

distribution. Thus, Truck et al. proposed to represent colors with trapezoidal or triangular fuzzy subsets, associating colors with fuzzy sets [7]; Herrera and Martinez use fuzzy linguistic hierarchies with different number of labels, depending on the desired granularity [8]. Some other approaches, such as [9], [10], use other common spaces for color-based classification, such as YCbCr or CIELab. However, there are little works dealing with fuzzy RGB classification and a technique for classifying skin color pixels in the RGB color space using a fuzzy approach has not been reported yet.

In order to use this fuzzy approach, we must calculate first the normalized histogram for the RGB color space for a training set, $H(C) = \{f_R, f_G, f_B\}$. Since the skin detection will be used a pre-processing task for detecting a face in an image, for the training of the system let us consider two different face databases:

− The XM2VTS face database [11] contains 8 recordings of 295 subjects each, acquired over a period of 4 months. The background in this set of images is homogeneous, with a good contrast for detecting skin.

− The VALID database [12] consists of five recording sessions of 106 subjects over a period of one month. One session is recorded in a studio with controlled lighting and no background noise, the other 4 sessions are recorded in office type scenarios.

The training has been performed using a set of 100 images from both databases, extracting only the skin information, and using different ethnic groups and changing lighting conditions.

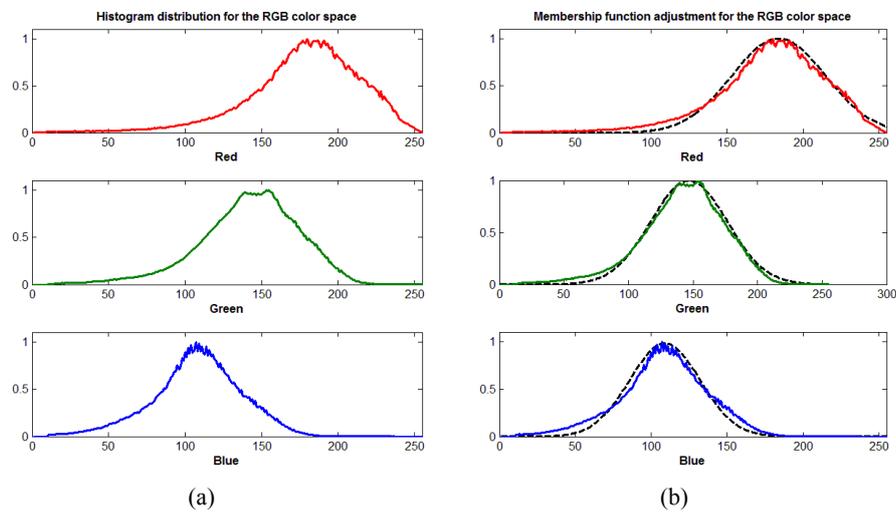The results after obtaining $H(C)$ are shown in Fig. 1(a).

(a)   (b)

Fig. 1. Creation of a fuzzy system for the RGB model. (a) Computation of $H(C)$ for the training set. (b) Gaussian approximation of $H(C)$ for each color plane.

From these results, the membership functions for the skin color in each plane can be modeled using a Gaussian function, such that:

$$\mu_{SKIN_i}(c_i) = \beta_i \, e^{-\frac{(c_i - \alpha_i)^2}{2\sigma_i^2}}. \tag{1}$$

where $i = \{R, G, B\}$; $\{c_R, c_G, c_B\} \in [0, 255]$; $\beta_i = \max(H(c_i))$, $\alpha_i = \underset{c_i}{\arg\max} H(c_i)$, and σ_i^2 is the variance of each fuzzy set c_i. The results of the model for the skin pixels are shown on Fig. 1(b).

Finally, for the background pixels, i.e., the non-skin pixels in the image, let us consider a variation of the model introduced by Murthy and Pal [13], which identifies the fuzziness in the transition region between the object (in this case, the skin) and the background classes. Thus, the membership value of a point to the object is determined by applying an S-function and a Z-function to the each color plane, so that:

$$\mu^S_{NON-SKIN_i}(c_i) = \begin{cases} 0 & \text{for } c_i \leq a_{Si} \\[2mm] 2\left(\dfrac{c_i - a_{Si}}{\gamma_{Si} - a_{Si}}\right)^2 & \text{for } a_{Si} \leq c_i \leq b_{Si} \\[2mm] 1 - 2\left(\dfrac{c_i - \gamma_{Si}}{\gamma_{Si} - a_{Si}}\right)^2 & \text{for } b_{Si} \leq c_i \leq \gamma_{Si} \\[2mm] 1 & \text{for } \gamma_{Si} \leq c_i \end{cases}. \tag{2}$$

$$\mu^Z_{NON-SKIN_i}(c_i) = \begin{cases} 1 & \text{for } x \leq a_{Zi} \\[2mm] 1 - 2\left(\dfrac{x - a_{Zi}}{\gamma_{Zi} - a_{Zi}}\right)^2 & \text{for } a_{Zi} \leq x \leq b_{Zi} \\[2mm] 2\left(\dfrac{x - \gamma_{Zi}}{\gamma_{Zi} - a_{Zi}}\right)^2 & \text{for } b_{Zi} \leq x \leq \gamma_{Zi} \\[2mm] 0 & \text{for } \gamma_{Zi} \leq x \end{cases}. \tag{3}$$

where the values b_{Si}, b_{Zi} are the cross-over points of the fuzzy sets defined by $\mu^S_{NON-SKIN_i}(c_i)$, $\mu^Z_{NON-SKIN_i}(c_i)$, respectively, for $i = \{R, G, B\}$; i.e., the membership value of $c_i = b_{Si}$ (or b_{Zi}) that is equal to 0.5; $\{c_R, c_G, c_B\} \in [0, 255]$; $a_{Si} = \gamma_{Zi} = \underset{c_i}{\arg\max} H(c_i)$; $a_{Zi} = 0.5 \cdot \gamma_{Zi}$; and $\gamma_{Si} = 1.5 \cdot a_{Si}$.

The results for the models of the fuzzy skin and non-skin classes are shown in Fig. 2; we have three classes for each color fuzzy set.
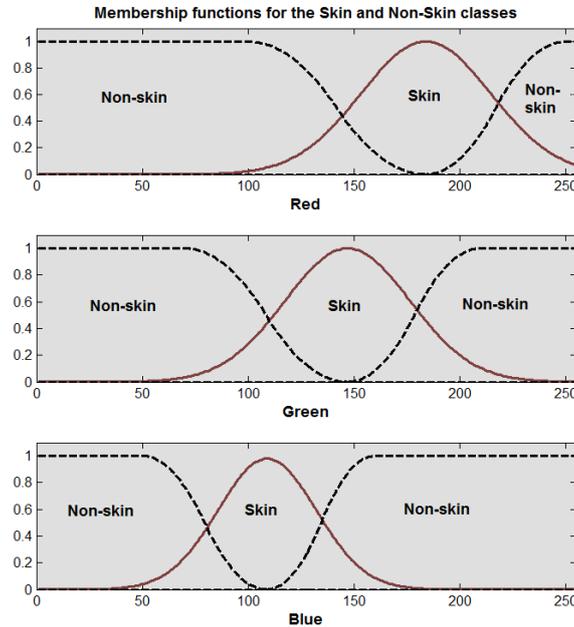
Fig. 2. Modeling the Skin and Non-skin Classes for the RGB color space.

Now, given an input image I, for any pixel $p \in I$, its R, G, B components are fuzzified, according to the parameters defined in Eqs. (1)-(3). Then, the inferencing system processes each pixel and, by using the available knowledge in the form of IF-THEN rules, it identifies and classifies skin color pixels in the output image. This procedure results in the assignment of one output fuzzy set for each rule. A total of 12 rules were extracted for our system. The min–max inferencing technique was used, where the output membership function of each rule is clipped off at a height corresponding to the rule premise's computed degree of truth. The combined fuzzy output membership function is built by combining the results of all the fuzzy rules. If an output fuzzy set is activated by more than one rule, the maximum of all activations is considered in the construction of the combined output membership function.

The Mamdani method was chosen as the defuzzification procedure, which means that the fuzzy sets obtained by applying each inference rule to the input data were joined through the add function; the output of the system was then computed as the centroid of the resulting membership function.

After the fuzzy inference system has been defined, in the following section we describe the results of the experiments performed for the different databases.

4   Experimental Results

Let us show now the results of some experiments completed for our model and compare it with some other existing methods. The tests have been performed with the

XM2VTS (homogeneous background) and the VALID (homogeneous and more complex backgrounds) face databases, as described in Section 3.

According to the fuzzy skin detector defined before, Fig. 3 shows the results for different images in both databases. The images include different gender and ethnic groups, and –for the VALID database– complex backgrounds and changing lighting conditions.

Fig. 3. Results of the fuzzy skin detector for different images. First row: some images from the XM2VTS database. Second row: some images from the VALID database.

From these results, we must point out that the proposed method leads to very accurate results for the XM2VTS database, which has both good illumination conditions and high contrast between the foreground and the background. In the case of the VALID database, the results of the skin detector are also quite precise, in spite of the unfavorable conditions for achieving good quality results. Some errors appear, mainly due to background colors that are very similar (or even identical) to skin colors; in addition, the lighting conditions will affect the final result. However, with a post-processing stage for detecting a face from the skin segmentation results, most of these errors will be corrected.

Finally, let us compare our system with some other existing algorithms for skin color detection. First, an explicitly defined skin region algorithm for the RGB color space is used; in [14], a pixel (R_0, G_0, B_0) is classified as *skin* if:

$$\begin{aligned}
&\text{(i)}\ \ R_0 > 95 \wedge G_0 > 40 \wedge B_0 > 20,\ \text{and} \\
&\text{(ii)}\ \max(R_0, G_0, B_0) - \min(R_0, G_0, B_0) > 15,\ \text{and} \\
&\text{(iii)}\ \left| R_0 - G_0 \right| > 15 \wedge R_0 > G_0 \wedge R_0 > B_0.
\end{aligned} \tag{4}$$

On the other hand, a skin probability map (SPM) approach to the RGB space is also taken into account. For this method, given skin and non-skin histogram models, a skin pixel classifier is defined using, for instance, the Bayes' theorem. Then, the inequality $P(skin|c) \geq T$ can be used as a skin detection rule, where T is a selected threshold value and $P(skin|c)$ is the probability of observing skin, given a concrete c color value (see [1], [5] for further details). For completion, we have included SPMs for the HSV and YCbCr spaces, as well.

Figs. 4-6 show the results for some images for the proposed and existing segmentation methods. In every case, for the SPM models, we show the threshold T that minimizes the number of false positives and false negatives.
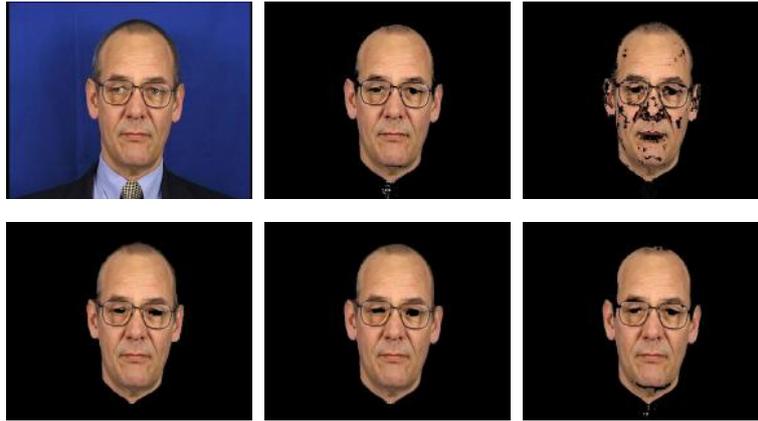

Fig. 4. Results from the XM2VTS database. From left to right, top to bottom: original image; results for explicit RGB rules; results for SPM in RGB, $T = 0.12$; results for SPM in HSV, $T = 0.06$; results for SPM in YCbCr, $T = 0.04$; results for our model.

Fig. 5. Results from the VALID database. From left to right, top to bottom: original image; results for explicit RGB rules; results for SPM in RGB, $T = 0.25$; results for SPM in HSV, $T = 0.37$; results for SPM in YCbCr, $T = 0.43$; results for our model.

Fig. 6. Results from the VALID database with a complex background. From left to right, top to bottom: original image; results for explicit RGB rules; results for SPM in RGB, $T = 0.15$; results for SPM in HSV, $T = 0.05$; results for SPM in YCbCr, $T = 0.03$; results for our model.

Finally, in Table 1 we show the results of the false positive (FP) and false negative (FN) rates, and correct skin detection rate for the images in the experiments, for all the models considered in our implementation.

Table 1. Comparison between algorithms

Algorithm	XM2VTS			VALID			Global Results		
	FP(%)	FN(%)	Correct(%)	FP(%)	FN(%)	Correct(%)	FP(%)	FN(%)	Correct(%)
Explicit RGB rules	2.13	3.07	94.80	4.67	9.28	86.05	2.97	5.14	91.89
SPM RGB	3.67	3.20	93.13	8.74	13.55	77.71	5.36	6.65	87.99
SPM HSV	1.62	3.09	95.29	4.38	8.07	87.55	2.54	4.75	92.71
SPM YCbCr	2.51	2.61	94.88	3.46	5.74	90.80	2.82	3.65	93.53
Fuzzy RGB	1.54	2.73	95.73	2.11	5.57	92.32	1.73	3.67	94.60

The results confirm the validity of the fuzzy RGB skin detector; the global results achieve a detection rate of 94.60% (the highest one from the compared methods), with only a 1.73% of false positives (the lowest one from the compared methods). The performance is better than the other compared RGB proposals; when there is a homogeneous background, with good lighting conditions, almost any skin detector gives precise results, but when we use unconstrained environments, it is more difficult to get a good skin segmentation. This fact can be clearly noticed for the VALID database and the RGB space for both the explicitly defined regions and the SPM methods, where many false positives and negatives appear and the correct detection rates are far from being optimal, especially for the SPM approach. Unlike the other

RGB proposals, our method behaves well for complex backgrounds and changing illumination, giving very accurate results.

On the other hand, the FP and FN rates show that the fuzzy RGB skin detector has a comparable behavior than other methods in color spaces traditionally used for skin color detection, such as HSV or YCbCr. From Fig. 7 it can be extracted that when there is a strong presence of a red object, almost any skin segmentation method gives false positives for the pixels in the object; although the HSV and YCbCr spaces provide a good segmentation result, they can't remove many of the red pixels from the objects in the background, whereas the fuzzy RGB achieves a more appropriate result. However, there are some situations with unconstrained backgrounds where it is rather difficult for any skin detector to generate a proper segmentation. Thus, the performance of the last three algorithms in Table 1 is comparable, so that any of them can be used as a pre-processing stage for more complex algorithms, such as a face recognition system or a human-machine interface.

To sum up, the fuzzy system has high correct detection rates and, as a conclusion, our method is a suitable technique for segmenting skin in different environment conditions, such as complex backgrounds or changing lighting conditions.

5   Conclusions

Face detection is object of intensive research in the last few years and this tendency is increasing, since it is the initial and, perhaps, the critical process for an integral face recognition system. One of the methods used for detect a face is to use a skin segmentation scheme, which has been proved to be highly effective in many applications.

In this work, a fuzzy skin color detector using the RGB color space has been proposed. Each color plane has been modeled using different fuzzy sets for the skin and non-skin classes and the inference system and defuzzifying process have been exposed. Experiments show that our method leads to very accurate skin detection results with low false positive and false negative rates.

As a future work, we propose to extend the fuzzy system to other color spaces, such as HSV or YCbCr. Then, some algorithm for detecting a face in an image must be developed, as well. It would be also desirable to use our algorithm with other databases and in a real environment to test the validity of our fuzzy system proposal.

References

1. Yang, M. H., Kriegman, D. J., Ahuja, N.: Detecting faces in images: A survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1), pp. 34—58 (2002)
2. Vezhnevets, V., Sazonov, V., Andreeva, A.: A Survey on Pixel-Based Skin Color Detection Techniques. In: Proc. Graphicon-2003, Moscow, Russia, September 2003, pp. 85—92 (2003)

3. Albiol, A. Torres, L., Delp, E. J.: Optimum color spaces for skin detection. In: Proc. ICIP-2001, Thessaloniki, Greece, 2001, pp. 122—124 (2001)

4. Yang, J., Fu, Z., Tan, T., Hu, W.: Skin Color Detection Using Multiple Cues. In: Proc. ICPR'04, Cambridge, UK, 23-26 Aug. 2004, vol. I, pp. 632—635 (2004)

5. Brand, J., Mason, J. S., Roach, M., Pawlewski, M.: Enhancing Face Detection in Colour Images using a Skin Probability Map. In: Proc. of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, China, 2001, pp. 344—347 (2001)

6. Zadeh, L. A.: Fuzzy Sets, Information and Control, 8(3), pp. 338—353 (1965)

7. Truck, I., Akdag, H., Borgi, A.: A Symbolic Approach for Colorimetric Alterations. In: Proc. of EUSFLAT 2001, Leicester, England, pp. 105—108 (2001)

8. Herrera, F., Martínez, L.: A Model based on Linguistic two-tuples for Dealing with Multigranularity Hierarchical Linguistic Contexts in Multiexpert Decision-Making, IEEE Transactions on Systems, Man and Cybernetics, 31(2), pp. 227—234 (2001)

9. Vertan, C., Boujemaa, N.: Using Fuzzy Histograms and Distances for Color Image Retrieval. In: Proc. of CIR'2000, Brighton, United Kingdom (2000)

10. Guironnet, M., Pellerin, D., Rombaut, M.: Video Classification Based on Low-Level Feature Fusion Model. In: Proc. of EUSIPCO '05, Antalya, Turkey (2005)

11. Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G.: XM2VTSDB: The Extended M2VTS Database. In: Proc. Second International Conference on Audio and Video-based Biometric Person Authentication, Washington D.C., March 1999, pp. 72—77 (1999)

12. Fox, N.A., O'Mullane, B.A., Reilly, R.B.: VALID: A New Practical Audio-Visual Database, and Comparative Results. In Takeo Kanade, Anil K. Jain, Nalini K. Ratha (Eds.) AVBPA 2005. LNCS, vol. 3546, pp. 777—786. Springer, Heidelberg (2005)

13. Murthy, C. A., Pal, S. K.: Fuzzy Thresholding : Mathematical Framework, Bound Functions and Weighted Moving Average Technique, Pattern Recognition Letters 11, pp. 197—206 (1990)

14. Kovač, J., Peer, P., Solina, F.: Human Skin Colour Clustering for Face Detection. In: Proc. EUROCON 2003, Ljubljana, Slovenia, 22-24 Sept. 2003, vol. 2, pp. 144—148 (2003)