A Probabilistic Neural Network for Human Face Identification based on Fuzzy Logic Chromatic Rules

Anagnostopoulos I., Student *Member IEEE*, Anagnostopoulos C., Student *Member IEEE, Vergados D., Member IEEE,* Loumos V., *Member IEEE* and Kayafas E., *Member IEEE*

Abstract-- This paper deals with the problem of face detection in color images. Unlike in face recognition, where the classes to be discriminated are different faces, in face detection, the two respective classes are the "face area" and the "non-face area". A novel approach to face detection is presented, based on fuzzy logic rules especially defined for skin area detection within the image frame. A Probabilistic Neural Network (PNN) was trained for the identification of the facial areas, which were extracted using the fuzzy logic rules.

The performance of the whole system was tested using 317 color images with different illumination conditions containing human faces. The images were taken using a digital camera, having a resolution analysis of 1280x960 pixels.

The proposed artificial neural network, joint with the fuzzy logic rules, became the basis for developing a computer-based face detection system, whose overall identification performance was measured to be 83%. However, this performance level is achieved for frontal-parallel faces, since the classification performance deteriorates when extended to different views of a human face.

Index terms— **Neural Networks, Fuzzy Logic, Image Processing, Skin Detection, Face Recognition**

I. Introduction

Nowadays, intelligent extraction of information is becoming exceptionally important due to the fact that image and video databases are widespread and expansive. It is also commonly accepted, that human faces are one of the most common and very specific objects, attempted to be solved in video sequences and color images.

A system capable of performing automatic detection and recognition of human faces represents an example of such an intelligent information retrieval. Some possible applications for automatic face detection and identification can be summarized on supervision and security applications, videoconferences, animation of facial expressions, as well as remote control camera applications. However, it must be emphasized that automatic face detection, like most other automatic object-detection methods, is a very elaborate task, due to significant sample variations that cannot be analytically described with simple parameters.

Many applications require the automatic detection of human faces in images with complex background. In the recent past there was a growing interest in image processing algorithms joint with Artificial Neural Network (ANN) technology and acting as valuable assistance tools, which can render feasible advanced solutions in the problem of face detection [1], [2]. The most commonly used family of neural networks for the problem addressed, are the feed-forward networks, a category which includes Multilayer Perceptrons (MLPs), Radial-Basis Function (RBF) networks and Generalized Regression Neural Networks (GRNN).

Multilayer perceptrons (MLPs) are layered feed-forward networks typically trained with static backpropagation. These networks have found their way into countless applications that require static pattern classification. Their main advantage is that they are easy to use and that they can approximate any input/output map. The key disadvantages are that they train slowly and require a large amount of training data, which is typically three times more training samples than network weights. On the other hand, Radial Basis Function (RBF) networks are nonlinear hybrid networks, typically containing a single hidden layer of processing elements (PEs). This layer uses Gaussian transfer functions, rather than the standard sigmoid functions employed by the MLPs. The centers and widths of the Gaussians are set by unsupervised learning rules and supervised learning is applied to the output layer. Thus, these networks tend to learn much faster than MLPs. If a generalized regression (GRNN) or probabilistic (PNN) net is chosen, all the weights of the network are calculated analytically. In this case, the number of cluster centers is by definition equal to the number of exemplars, and they are all set to the same variance.

In this paper, a neural network based algorithm is presented, assisted with fuzzy logic rules to detect human faces in color images. The proposed classifier belongs to the RDF networks and, in particular, to the Probabilistic Neural Networks, since these networks excel in case the number of the training exemplars is quite small or so much dispersed that clustering is ill defined [3]. Moreover, training a neural network for the face detection task is challenging due to the difficulty of the problem, the diversity of the illumination

Contact authors: Anagnostopoulos I., Anagnostopoulos C.,
National Technical University of Athens, Heroon Polytechneiou 9 Str., 15773, ATHENS, GREECE
e-mail: {janag,canag@telecom.ece.ntua.gr}

conditions and the variety of positions of the face in the image. Unlike face recognition, where the classes to be discriminated are different faces, the two classes assigned in face detection are “images with faces” and “images without faces”.
In the following section the Fuzzy Logic rules (FL rules) applied for skin color discrimination are presented, and the technical details are briefly discussed.

II. The Fuzzy Logic Rules

The color of human skin is distinctive from the color of many other objects and, therefore, the statistical measurements of this attribute are of great importance for face detection [4], [5]. Evaluating the skin tone statistics, it is expected that the face color tones will be distributed over a discriminate space in the RGB color plane. So, the first step of the proposed system is the location of potential skin areas in the image, using RGB color information. Many approaches in the literature used similar detection procedures, either based on the RGB, chrominance (CbCr) [6] or Hue and Saturation (HSV) space [7].
The basic concept in fuzzy logic, which plays a significant role in most of its applications, is that of a fuzzy if-then rule or, simply, the fuzzy rule. In the proposed schema, the skin-masking algorithm presented in [8], is partially used along with RGB cluster groups that represent skin color extracted from experimental tests in a large database of human face images. The above measurements and the skin-masking algorithm, formed the basis for the definition of the fuzzy logic rules. The aforementioned if-then rule statements are used to formulate the conditional statements that comprise the fuzzy logic-based skin color detector.
A basic rule, of significant importance to the application of the proposed system, is resulted from the experimental in [9]. In this method, the chromatic field YCbCr is used, since it was proved to be more representative for the choice of regions that suit human skin [10,][11], [12], [13], [14].
In the YCbCr field, the brightness (luminosity) is stored as a simple element (Y) and the value of chrominance as two different elements (Cb and Cr). The values Cb and Cr represent the difference between light blue and the current calculated value, as well as the difference between red and the current calculated value, respectively. The following equations perform the transformation in the RGB field.

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$
$$C_b = -0.169 * R - 0.331 * G + 0.5 * B + 128$$
$$C_r = 0.5 * R - 0.419 * G - 0.081 * B + 128$$

Through the application of Fuzzy Logic rules, the proposed system decides whether a specified window in the inspected image contains a potential skin region. However, a skin region does not represent always a face, and therefore the candidate area should be further normalised and checked in order to discern whether it represent a face or not. Figure 1, presents the chromatic distribution of the human skin area in respect to Cb and Cr values.

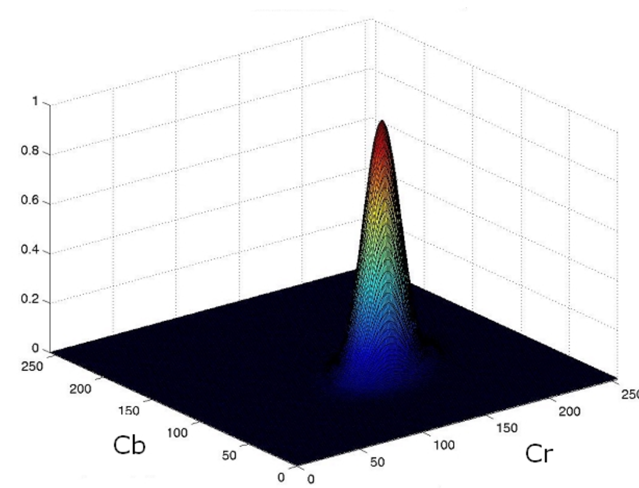


Figure 1. Human skin chromatic distribution

The fuzzy logic rules applied for skin area discrimination are depicted below (R=Red, G=Green, B=Blue).

1. If R<100 then no skin (shadow)
2. If R<G then no skin
3. If R<B then no skin
4. If R/G >1.3 and R/B >1.4 then possible skin
5. If R/G <1.3 or R/B <1.4 then no skin
6. If R/G >1.3 and G/B >1.5 then possible skin
7. If R/G <1.3 or R/B <1.4 then no skin
8. If 77<G<127 and 133<B<173 then possible skin

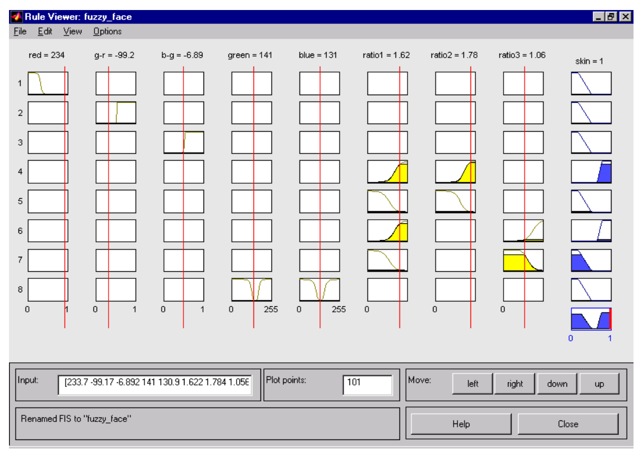


Figure 2. Definition of the Fuzzy Logic Rules

The first step is to take the inputs and determine the degree to which they belong to each of the appropriate fuzzy sets through membership functions. Once the inputs have been fuzzified, the fuzzy logical operations must be implemented. For this application the OR operator was used. The weights in every rule were set equal to one and the aggregation method for the rules is the maximum value. Finally, the defuzzification method is the middle of maximum (the average of the maximum value) of the output set.
The fuzzy rules were successfully applied to a Fuzzy Inference System (FIS), using the Fuzzy Logic Toolbox of Matlab 5.2 by MathWorks Inc. The inputs of the FIS are the RGB values of the tested image. It is evident that, as the size of the inspected image grows, the processing time

increases. In a Pentium IV at 1.5 MHz with 512 MB RAM, the required time for skin area detection varied from 2 to 7 seconds. A screenshot of the application is shown in Figure 2. Input and output images are presented in Figures 2 and 3 respectively. Figure 3 depicts the possible skin areas according to the defined Fuzzy Logic Rules. The following section, describes analytically the proposed image processing operations used in order to define the Region of Interest (RoI) of possible skin areas in a tested image, as well as to extract the input vector of the artificial neural network.

Figure 3. Input image

Figure 4. Output image according the FL rules

III. Image Processing

Digital image analysis is the manipulation of information within a visual representation of an object or group of objects, performed with a computer. Lately, digital image analysis is widely used in various industrial applications combined with ANNs or with Fuzzy Logic systems that are capable of interpreting data generated by this type of image processing and drawing useful conclusions. The image-processing operations proposed here, consist of four distinct parts.

Part 1: Definition of the Region of Interest (RoI) for all the possible skin areas.

During this step, potential skin areas are clustered to form the Region of Interest (RoI), roughly describing its shape, on the basis of the FL output (Figure 3). Firstly, every image is transformed in gray scale and in the specific size of 100x100 pixels. The next step consists of two morphological operations, which help to eliminate some of the noise in the tested image. In particular, simple erosion with a 10x10 matrix of ones is performed followed by dilation. Further on, the created image is parsed through a skeletonisation technique, removing simultaneously all the areas that are considered as ‘holes’. As a result of the previously described image processing steps, the RoIs of all the possible skin areas are depicted in Figure 5.

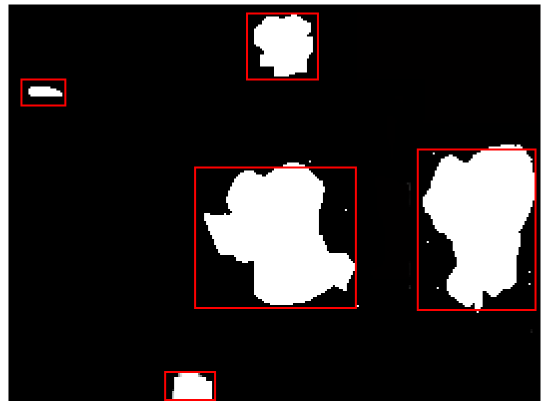


Figure 5. Definition of the RoIs

Part 2: Transformation of each skin color sub–image to gray scale.

Having defined the RoI in the previous part, the algorithm is applied to the initial tested image, merging objects that belong to one defect, performing a simple dilation once again, with a structural element, which is a 5x5 matrix of ones. With this technique, segmented pixels in the same neighbourhood, are merged in one region. All the image parts that are included in the defined RoIs, are then transformed to gray scale as shown in Figure 6.

Part 3: Resize all the segmented images to a specific size of 225x225 pixels.

Part 4: Division of the segmented images to non-overlapping sub-images and histogram modification.

At this point, the 225x225 pixel images are divided into non-overlapping sub-images of size 15x15 and the mean value for each is calculated, followed by histogram equalization, which expands the range of intensities in the window [15]. During this procedure, a lower resolution image in respect to the RoI is created, forming in parallel a descriptor vector that consists of 225 gray scale values from 0 to 255. Figure 7 presents the input for the proposed neural network. The advantages of this operation and the architecture of the PNN are explained in detail in the next section, along with the experimental results.

Figure 6. Extracted images according the defined RoIs

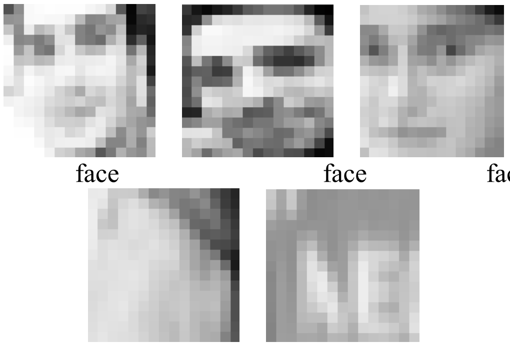


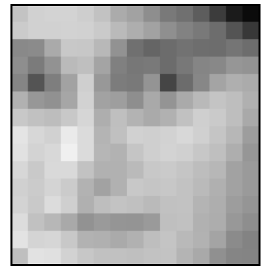
Figure 7. Candidate inputs for classification

IV. THE PROPOSED PROBABILISTIC NEURAL NETWORK

The proposed ANN system is trained to identify which of the skin regions detected from the FL system represent faces. The training set of the ANN consists of a large group of images of the size 15x15, representing face regions or other skin areas. The idea of this approach was motivated by the observation that human faces present a high degree of resemblance when they are sampled in low-resolution [2]. This is quite natural, since all faces have darker areas, which represent the eyes and the mouth. It is undoubtedly easier for an ANN to recognize the presence or absence of a face, judging from a low quality image. Additionally, the numbers of the computational units are significantly smaller for a low quality image.

The ANN is a two layer Probabilistic Neural Network with biases and Radial Basis Neurons in the first layer and Competitive Neurons in the second one. Probabilistic Neural Networks (PNN) are a class of neural networks that combine some of the best attributes of statistical pattern recognition and feed-forward neural networks [16], [17]. PNNs feature very fast training times and produce outputs with Bayes posterior probabilities. Compared to conventional neural networks, the useful features of PPNs, come at the expense of larger memory requirements and slower execution speed for prediction of unknown patterns. Thus, PNNs can be used for high performance classification problems.

The main difficulty of ANN-based face detection is the great diversity of facial skin tones that vary according to illumination conditions. For this application the ANN's training set consists of representative face patterns of specific size (15x15 pixels) extracted from sample images without any restrictions imposed on them. The variety of these samples should ensure the successful implementation of the Probabilistic Neural Network system, as its performance is straightforward and does not depend on time-consuming training, like the backpropagation scheme. Training a neural network for the face detection task is quite challenging due to the difficulty in characterizing prototypical "non-face" images. Unlike in face recognition, where the classes to be discriminated are different faces, in face detection, the two classes to be discriminated are "face area" and "non-face area". Figure 9 depicts a face image as it is transformed in a vector form, consisting of 225 gray scale values. This transformation is needed in order to be recognised and processed by the neural network.



195	209	211	210	205	194	171	161	142	121	113	94	61	28	11
203	209	209	210	209	202	186	174	162	143	130	121	109	84	56
138	139	169	199	199	184	146	114	102	108	114	110	109	117	109
142	123	146	186	191	164	128	119	121	108	120	132	136	150	147
133	84	126	162	207	152	124	121	147	69	108	135	164	175	171
176	151	145	166	210	163	158	142	169	148	170	184	196	190	176
200	192	189	199	214	174	191	178	170	177	193	207	202	190	168
228	215	209	233	219	179	192	209	206	210	216	208	197	183	157
225	210	216	239	218	182	177	196	206	210	207	198	190	174	150
215	200	216	216	194	177	176	186	200	201	197	189	182	162	154
208	203	212	202	178	163	177	200	197	199	194	185	176	162	154
211	201	219	207	176	190	195	192	186	193	193	180	173	157	149
220	193	182	169	146	154	149	148	154	191	193	176	173	152	142
226	212	213	199	178	176	170	176	190	196	185	178	164	139	131
190	225	224	210	199	193	193	191	194	195	180	169	144	133	132

Figure 9. The input vector of the PNN

A sample of 129 frontal view face images was used as training set for the class 'Face', as well as a large sample of 296 images corresponding to other correct or erroneously detected skin areas, such as hands, legs and other objects. Table 1 presents the confusion matrix percentages in terms of the learning ability during the training epoch. The training set consists of 425 sub-images of size 15x15 in a vector form, as these were extracted from 103 color images according the proposed image processing steps. In other words, the neural network 'learned' to identify 128 from the 129 sub-images corresponding to human faces as well as 293 from the 296 sub-images corresponding to other skin areas and objects.

The time needed for the completion of one training epoch in a Pentium IV at 1.5 MHz with 512 MB RAM, was 22 seconds. The neural network was implemented in C++ and its performance was evaluated afterwards with the help of Matlab 6 R12 Neural Network Toolbox. The following equations outline the Akaike's Information Criterion (AIC) as well as the Rissanen's Minimum Description Length (MDL) as extracted measurements during the training period.

$$AIC(k) = N * \ln(MSE) + 2 * K$$

$$MDL(k) = N * \ln(MSE) + 0.5 * K * \ln(N)$$

where:

k : number of network weights

N: number of exemplars in the training set

MSE: Mean Square Error

AIC measure the tradeoff between training performance and network size, while similarly MDL combines the model's error with the number of degrees of freedom for determining the level of generalization. Thus, requiring a

neural architecture with the best possible generalization ability, the minimum achieved values for the problem addressed are depicted in table 1.
Furthermore, the topology of the proposed neural network is 225-425-2. This means that the PNN has a 225-input vector (the 15x15 input image) and a 2-output vector corresponding to the decision of the system (whether it is a face or not). Finally, the system has 425 nodes in the middle layer corresponding to the total training set.

	Face	Other skin area – object
Face	99.22% (128/129)	0.88% (1/129)
Other skin area - Object	1.01% (3/296)	98.99% (293/296)
AIC / MDL	190847.24 / 270366.58	

Table 1. The Training Confusion Matrix

The performance of the whole system was tested using 317 colour images with various illumination conditions containing human faces. The pictures were taken with a common digital camera, at a resolution of 1280x960 pixels. Our sample of 317 colour images contained 482 faces. The system implementing the fuzzy logic rules segmented totally 841 skin areas. However, 30 faces were not selected and therefore the performance of this system is 93.77% (452/482). Following the fuzzy logic system, the ANN received the 841 skin areas and decided that 397 of them represent faces. Thus, the performance of the ANN is 87.83% (397/452). Finally, the overall system performance is 82.36%, since 397 from a total of 482 faces were identified. All the results are shown analytically in Table 2.

Training Set		
Total color images	**Number of faces**	**Skin areas - Other objects**
103	**129**	**296**
Testing Set		
Total images		**Number of faces**
317		**482**
FL rules		
1) Segmented areas	**841**	**452 faces + 389 possible skin areas**
2) FL Rules performance	**452/482**	**93.77%**
Artificial Neural Network (ANN)		
Faces	**397**	
No faces	**444**	
ANN Performance	**397/452**	**87.83%**
3) Total System Performance	**397/482**	**82.36%**

Table 2. Performance measurements of the proposed system

V. CONCLUSIONS

This paper presents in detail the application of Fuzzy Logic rules along with the necessary image processing procedures. The anticipated outcome is to properly feed a neural network, which targets to identify the presence of human faces in color images. The necessary steps are described in order to emphasise the correlation between the fuzzy rules and the image processing. A cooperation of two different kinds of intelligent information retrieval systems outlines the basis of a system, which ensures high performance in human frontal face detection. Despite the limitations confronted in terms of different image illumination conditions, the method is effective enough if the implemented architecture would not require fast operation. Thus, it could be easily serve as a front part of a face recognition system, where the falsely detected faces could be eliminated using correspondence to a database of already known faces. An other possible and interesting extension is the expansion of the Neural Network ability to recognize faces in different angle positions and not only frontal or slightly rotated as it is taken under consideration in this paper. Furthermore the proposed application could possibly encapsulate geometric techniques for highlighting human face features, not only for face detection but for face recognition as well.

VI. REFERENCES

[1] Neural Network based Face Detection, Henry A. Rowley, Shumeet Baluja, Takeo Kanade, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, pp. 23-28, 1998.

[2] Recognition of facial images with low resolution using a Hopfield memory model, Y. Dai and Y. Nakano, Pattern Recognition, vol. 31, no. 2, pp. 159-167, 1998.

[3] "Statistical Pattern Recognition: A Review", Anil K. Jain, Robert P.W. Duin, and Jianchang Mao, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 22, NO. 1, JANUARY 2000.

[4] Color –and texture- based image segmentation using EM and its application to content-based image retrieval, S. Belongie, C. Carson, H. Greenspan and J. Malik, Proceedings of the 6th IEEE International Conference in Computer Vision, January 1998.

[5] Face recognition using 3D distance maps and principal component analysis, H. Grecu, V. Buzuloiu, R. Beuran, E. Podarou, Proceedings of Optim 2000, Brasov, May 2000.

[6] Face detection using quantized skin color regions merging and wavelet packet analysis, C. Garcia and G. Tziritas, IEEE Trans. on Multimedia, vol. 1, no. 3, pp. 264-277, 1999.

[7] Facial features extraction in frontal views using biometric analogies, S. Tsekeridou and I. Pitas, Proceedings of 9th European Signal Processing Conference, vol. 1, 1998.

[8] Computer Vision and Image Processing, Scott E. Umbaugh, pp. 334, Prentice Hall International, NJ, 1998.

[9] Chai D. and K. N. Ngan, "Locating facial region of a head-and-shoulders color image," Third IEEE International Conference on Automatic Face and Gesture Recognition (FG'98), Nara, Japan, pp. 124-129, Apr. 1998.

[10] Menser Bernd and Michael Brünig. Locating human faces in color images with complex background. In Proc. IEEE Int. Symposium on Intelligent Signal Processing and Communication Systems ISPACS '99, pages 533-536, Phuket, Thailand, December 1999.

[11] Saber A. and A. M. Tekalp, "Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions," Pattern Recognition Letters, vol. 19, pp. 669–680, June 1998.

[12] K. Sobottka, I. Pitas, A novel method for automatic face segmentation, facial feature extraction and tracking, Signal processing: Image communication, 12, pp. 263-281, 1998.

[13] Chai D. and K. N. Ngan, “Face segmentation using skincolor map in videophone applications,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 9, pp. 551–564, June 1999.

[14] Wang H. and S. F. Chang, “A highly efficient system for automatic face region detection in MPEG video,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 7, pp. 615–628, Aug. 1997.

[15] Example-based learning for view-based human face detection., K.K Sung and T. Poggio, A.I. Memo 1521, CBCL Paper 112, MIT, December 1994.

[16] Face recognition/detection by probabilistic decision-based neural network, S.-H. Lin, S.-Y. Kung, and L.-J. Lin, IEEE Trans. on Neural Networks, vol. 8, pp. 114-131, January 1997.

[17] Probabilistic visual learning for object detection, B. Moghaddam and A.P. Pentland, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 17, pp. 696-710, July 1997.