AN EFFICIENT WAVELET/NEURAL NETWORK-BASED FACE DETECTION ALGORITHM

Bardia Mohabbati, Mohammad Shiri

Amirkabir University of Technology
Department of Computer Science
P.O.Box 15875-4412, Tehran,Iran
b_mohabbati@aut.ac.ir

Shohre Kasaei

Sharif University of Technolog
Department of Computer Engineering
P.O.Box 11365-8639, Tehran,Iran
skasaei@sharif.edu

ABSTRACT

In this paper, we proposed an efficient method to address the problem of face detection that is based on neural network and wavelet representation. We utilized a multilayer perceptron (MLP) so as to classify skin and non-skin pixel in *YCrCb* plan. Skin samples in images with varying lighting conditions are used for obtaining a skin color distribution, and the training data were generated consisting of positive and negative training patterns in *Cb-Cr* planes. Subsequently, training set is fed to a multylayer perceptron, trained using Levenberg-Marquardt algorithm with the skin samples. We apply the above neural network skin classifier to chrominance values corresponding to coarsest level lowpass chrominance subimages obtained from wavelet transform to classify candidate face pixels. Furthermore, we have proposed a subspace approach in the space-frequency domain for the fast detection of face utilizing wavelet representation.

1. INTRODUCTION

Face detection in complex environments is a challenging problem which has fundamental importance to model-based video coding, image-database, content-based image retrieval, face recognition systems. Face detection in real-time processing is an important and preliminary step of a variety of applications requiring intelligent human-computer interaction. Since the human face is an intensely dynamic object, the main challenge in face detection is the amount of variation in visual appearance, such as size, color, shape, surrounding condition,light condition, shadows, pose as well as position and orientation. A general statement of face detection problem could be described as the determination of the location and size of the presence of human face in images. Several methods for face detection are discussed in the literature (for details see [1, 2]). In accordance with face prior knowledge, face detection algorithms could be classified in three main categories. The appearance-based approaches take advantage of the current advances in pattern recognition theory. These approaches build detection models directly from the image data, and the classification of face group is carried out by using training schemes and machine learning techniques [3, 4, 5].

Feature-based algorithms detect face by first detecting distinct components of the face, such as skin color ,measuring the face geometric relations,motion and visual features derived from the images [6, 7, 8]. The main trouble of these approaches is that it is very difficult to translate human face knowledge to computer representation, and reliable facial feature detection is still an instance of research problem. The multi-classification based approaches utilize both feature-based and appearance-based models to classify face and non-face objects to fulfill efficient face detection tasks [5, 9].

In this paper, we propose an efficient scheme to detect face by combining neural network and wavelet representation properties. The reminder of this paper is organized as follows: Section 2 reviews some techniques related to our method. Section 3 introduces color representation and skin-color model. In section 4,we describe neural networks trained with skin and non-skin samples so as to estimate the probable skin regions generated by the neural network classifier. We elaborate the concept of energy distribution of wavelet transform coefficients to verify the face detection result and remove false alarms in section 5. Section 6 provides an analysis of the proposed model. Finally, section 7 comes up with both the conclusion and an insight into the issues of human face tracking and future works.

2. RELATED WORKS

The detection of faces and facial features in images and video sequences has been regarded as a challenging problem in the field of computer; hence, many approaches have been proposed to accomplish this task. Categorizing face detection methods based on representation used reveals that detection algorithms using holistic representation have the advantage of finding small face or faces in poor-quality images, while those using geometrical facial features provide a good solution for detecting faces in different poses. A combination of appearance (ie holistic) and feature-based approaches [10, 11, 3] is a promising approach to face detection, tracking, and also face recognition systems. Recently, several color-based systems have been proposed to perform face detection and tracking in images. Processing color is much faster than processing other facial features; furthermore, color is an orientation invariant under certain lighting conditions. This property makes motion estimation much easier since only a translation model is needed. Hsu, Abdel-Moualeb, and Jain [12] proposed a method based on skin-color model using a parametric ellipse in a two-dimensional transformed color space. This approach is able to handle a wide range of variations in static color images, based on a lighting compensation technique and a nonlinear color transformation. The work of Yang and Waibel [13] presents an adaptive statistical skin-color model, which is invariant to people of different races. Moreover, this model was applied to a real-time face tracker [14]. The task of face detection is one of the general object detection problem ;by the same token, Schneiderman and Kanade [15] proposed a complex statistical classifier to detect objects. They used a Bayesian method for detection which represented the statistics of both face and non-face appearances taking advantage of a product of histograms that is used in respect to the joint statistics of a subset of wavelet coefficient capturing local feature in space, frequency and orientation and the position of the aforementioned appearances. Neural networks have also been extensively used for pattern recognition problems, including face detection. Rowley et al.'s [4] proposed a connected neural network which incorporates face knowledge. The neural network is designed to look at 20×20 pixel windows. One hidden layer with 26 units looks at different regions based on facial feature knowledge. The dominant factor in the running time of the Rowley system is the number of 20×20 pixel windows which the neural networks must process. Viola and Jones [9] proposed a face detection system based on a multi-classifier cascade and Ada-Boosted perceptron. They used a set of rectangular features to represent face objects which are captured by Harr basis functions. Their method utilizes AdaBoost which is an effective learning algorithm and strong bounds on generalization performance [16] to both select critical features and construct a classifier by selecting a small number of important features. Eventually, they combine classifiers in a cascade,similar to the decision tree, which allows background regions of the image to be quickly discarded while spending more computation on promising face-like regions.

Our approach is the combination of feature and appearance-based methods by the utilization of neural networks skin classifier and wavelet representation to establish a face tracking system.

3. SKIN COLOR MODEL

Human skin tones form a special category of colors, distinctive from the colors of most other natural objects; in consequence, color is the fundamental cue that can used as the first step in the process of face detection in complex images; furthermore, color image segmentation is computationally fast while being relatively robust to change in illumination, scale, viewpoint, shading and complex background as compared to the segmentation of the gray-level image. Human skin color segmentation strongly banks on the selected color space, because the skin color distribution depends on the specific color subspace. Skin color detection rate is affected by illumination significantly; consequently, large variations of illuminance need to be eliminated . The normalized color spaces are shown to yield the best segmentation results because normalization eliminates a large part of illuminance variation. It has been observed that different human skin colors give rise to compact clusters in color space, such as a normalized *RGB*(red,green,blue) [17], *YIQ*(luminance,r-cyan,magenta-green) [18], *YES* [19], *YUV* [20], *YCbCr* [12] and *HIS-HSV*(hue,saturation,intensity value) [21], even when faces of different races are the subjects at hand. Terrillon et al.'s comparison of nine different color spaces for face detection reveals that the *TSL* (tint-saturation,luma) spaces provide the best results for two kinds of Gaussian models. Also, for each color subspace, the detection efficiency is critically dependent on the appropriateness of the fit of the skin color distribution to the proposed model, and to a lesser extent on the discriminability between skin and non-skin distributions [22].

In our proposed method,the *YCbCr* space which is broadly used in video compression standards and still-images is adopted. The separation of luminance and chrominance is *YCbCr* proper to image compression schemes that use psychovisual redundancy. Furthermore, the luminance separability of this space is similar to the *TSP*. The segmentation of skin colored regions becomes robust only if the chrominance component is used in analysis. Therefore, we eliminate the variation of luminance component as much as possible by choosing the *Cb-Cr* plane (chrominance) of the *YCbCr* color space to build the model. Skin colors from various races of the world are collected from the World Wide Web in the form of 32×32 pixels per skin sample for each individual from each image. The distribution of the training skin pixels in the *Cb-Cr* plane is given in Figure 1. The figure shows that the color of human skin pixels is confined to a very small region in the chrominance space.

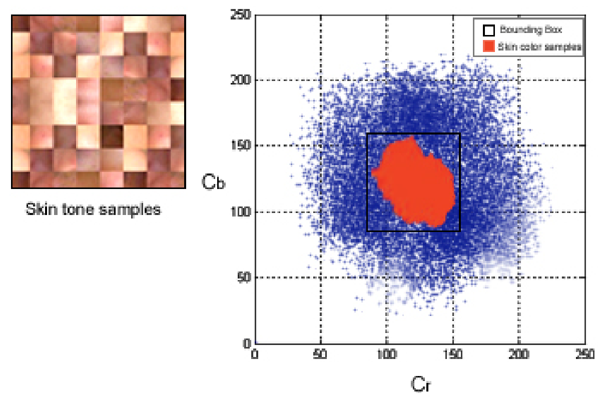

Figure 1: Skin color pixels in *Cb-Cr* plane. Red area represents skin color samples and the bounding box determines the minimum and maximum of *Cb* and *Cr* thresholds approximated by NN.

The methods of modelling as regards skin color in *YCbCr* color space fall into the following categories: parametric, non-parametric, and semi-parametric [23]. A parametric skin-color model has a specific functional form with adjustable parameters chosen to fit the model to the data set such as Gaussian mixture models. Non-parametric model does not assume any particular form, for example histogram thresholding. The key concept of the non-parametric skin modelling methods is to estimate skin color distribution from the training data regardless of the derivation of an explicit model of the skin color. A semi-parametric approach applies a very general form with adaptive parameters systematically varied in number as well as in values in order to create flexible models such as utilizing neural networks. To fulfill this aim, in this paper, our proposed method uses neural networks as a semi-parametric method for the sake of classifying skin-color pixel and creating face region candidates.

4. NEURAL NETWORKS APPROACH

An artificial neural network (ANN) is a generic parametric model which learns to represent a specific input-output relation. An ANN is composed of a set of non-linear processing units operating in parallel and arranged in a specific topology. Many different neural net topologies exist (see [24] for a comprehensive foundation). In this section, we utilize artificial neural networks so that we extract the human's skin regions from the *YCbCr* plane and interpolate them so as to provide an optimum decision boundary and subsequently the positive skin samples for the skin classification and labelling face candidates.

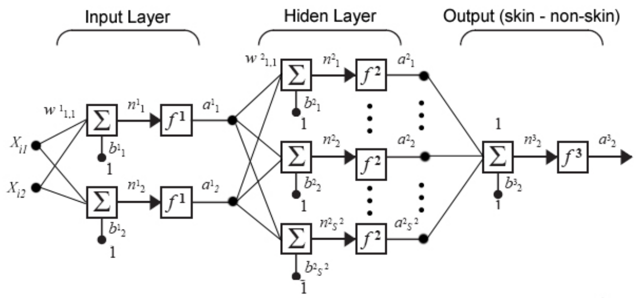

Figure 2: Neural networks architecture used for skin color learning.

Skin color modelling in essence can be considered to be a classification problem. The aim of skin color pixel classification is to determine whether a color pixel is a skin color or a non-skin one. Good skin color pixel classification should provide a coverage of all different skin types (blackish, yellowish, brownish, whitish, etc.) This type of problem is well-suited to artificial neural networks, which have been proven as an effective tool for pattern classification tasks where decision rules are hidden in highly complex data and can be learnt only from examples. A quadratic or more generally a non-linear function such as one hidden layer of a neural network is a good choice for a satisfactory approximation of skin color distribution. The neural network used in our work is the multilayer perceptron (MLP) which is a feed-forward neural network that has been used extensively in classification and regression. The MLP is capable of producing more complex decision boundaries. We use a neural network with the MLP architecture and feedforward topology to classify skin and non-skin pixels in the *Cb-Cr* plane (Figure2. The employed multilayer feedforward neural network consists of neurons with a sigmoidal activation function. The employed neural networks are used in two modes. In classification mode, an unknown input, or feature vector, is presented at the input layer and is propagated forward through the network to compute the activation value for each output neuron. The second mode is called the training or learning mode. Learning in ANN's involves the adjustment of the weights in order to achieve the desired processing for a set of learning skin-tone samples. More specifically, the second mode includes feeding a neural network with a number of training pairs, each of which consists of a feature vector and a corresponding class indicator, skin and non-skin class. Then the networks parameters are adjusted through a supervised training algorithm so that it produces the expected class indicators for the given feature vectors. The previous section proves that chrominance values $[Cb \;\; Cr]^T$ are appropriate choices for the feature vectors; as a result, the training set of our neural networks consists of transferred skin-color pixels in *YCbCr* spaces.

4.1. Skin classification

Let X be an ensemble of skin color samples: $X = \{x_1, x_2, \ldots, x_N\}$ where $x_i = [Cb \;\; Cr]^T$, and $Cb \in \mathbf{R_{Cb}}$, $Cr \in \mathbf{R_{Cr}}$. The neural networks skin-color classification can be expressed in the following way:

$$y_i = \sigma(\sum_{i=1}^{N} w_{ij}^{(l)} x_i + b) \qquad (1)$$

where $w_{ij}^{(l)}$ signifies the weight on connection between the i^{th} unit in layer $(l-1)$ to j^{th}layers unit in layer l. We often consider the threshold b to be another weight $w_0 = -b$ which is attached to the neuron with a constant input, $x_0 = 1$, and σ is the activation function, a sigmoid type, on the weighed sum to generate a single output y.

$$\sigma = \frac{1}{(1+e^{-\gamma})} \qquad (2)$$

In a two-class, skin and non-skin, classification problem, we assign skin-color as input patterns to skin class if $y_1 = 1$, and to the non-skin class if $y_0 = 0$; accordingly, the decision boundary divides the space into two halves.

5. FACE REGION VERIFICATION

The previous stage of the proposed scheme consists of locating the potential face areas in the image, using skin chrominance information which is generated by neural networks, given that such information strongly reduces the search space.The main purpose of this section is to reduce data and verify the face region candidates and remove false alarms caused by objects with colors similar to skin tones by performing discrete wavelet transform (DWT).

There are many discontinuities in the intensity level because of the existence of facial features such as eyes and moustache, among others. These give rise to high frequency wavelet coefficients in the luminance component *Y*. We elaborate the concept of wavelet block proposed for grouping wavelet coefficients in *Y* based on zerotree hypothesis [25, 26].

5.1. Wavelet transform and energy distribution of coefficients of luminance component

Wavelet theory has proved to be one of the most promising approaches to image processing. The main characteristic of wavelets in comparison with the other transformations is the possibility to provide a multiresolution analysis of the image in the form of coefficient matrices with a spatial and a frequential decomposition of the image at the same time; in the word, multiresolution techniques intend to transform images into a representation in which both spatial and frequency information is present.

Wavelet gives an orthonormal linear transformation of image data that has the property which the property of data is concentrated in only a few coefficients. A complete mathematical framework has been recently built [27, 28, 29] in particular for what concerns the construction of wavelet bases and efficient algorithms for its computation. This leads to an efficient real-space implementation of the wavelet transform using quadrature mirror filters. In the 2-D case, the wavelet transform is usually performed by applying a separable filter bank to the image. Typically, a low filter and a bandpass filter are used. The convolution with the low pass filter results in a so-called approximation image and the convolutions with the bandpass filter in specific directions result in so-called details images [21]. The wavelet decomposition of an 2-D image can be obtained by performing the filtering consecutively along horizontal and vertical directions. Wavelet coefficients are organized

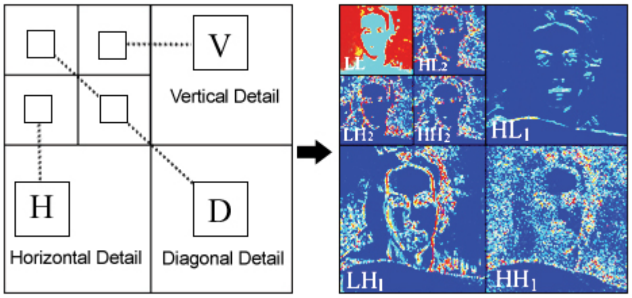

Figure 3: The subband labelling scheme for a two-level, 2-D wavelet. H, V, and D represents the horizontal, vertical, and diagonal subimages respectively.

into wavelet blocks as shown in Figure 3, where H, V, and D correspond to horizontal, vertical, and diagonal edge subimages respectively, while the upper most left subimage corresponds to coarsest level low pass subimage (L). The concept of wavelet block provides an association between wavelet coefficients and what they represent spatially in the frame.

Given a candidate face region of size $M \times N$ pixels with respect to the masked image which is generated by neural networks classifier, the energy of the corresponding luminance blocks in the DC and H and V areas are calculated as below by pruning the wavelet transform coefficients:

$$E = L(x,y)^2 + \sum_{l=0}^{3} \sum_{m=0}^{2^l-1} \sum_{n=0}^{2^l-1} [H_{4-l}(m+2^l x, n+2^l y)^2 + \quad (3)$$

$$V_{4-l}(m+2^l x, n+2^l y)^2 + D_{4-l}(m+2^l x, n+2^l y)^2]$$

$$E_{DC} = [L(x,y)^2] \quad (4)$$

$$E_V = \sum_{l=0}^{3} \sum_{m=0}^{2^l-1} \sum_{n=0}^{2^l-1} [V_{4-l}(m+2^l x, n+2^l y)^2 \quad (5)$$

$$E_H = \sum_{l=0}^{3} \sum_{m=0}^{2^l-1} \sum_{n=0}^{2^l-1} [H_{4-l}(m+2^l x, n+2^l y)^2 \quad (6)$$

where L corresponds to coarsest level low pass subimage. E, E_{DC}, E_H, and E_V are the total, DC, horizontal, and vertical energies of a single wavelet block. For face region of size $M \times N$, these energies are obtained as follows:

$$E_{MN} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E_{i,j} \quad (7)$$

$$E_{DC_{MN}} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E_{DC_{i,j}} \quad (8)$$

$$E_{H_{MN}} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E_{H_{i,j}}, \quad E_{V_{MN}} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} E_{V_{i,j}} \quad (9)$$

Eqn (7) gives the total energy of all the wavelet blocks in the candidate face region. It equals the energy of the pixel values of this face region, because of the wavelet transforms energy-conserving property. $E_{MN}, E_{DC_{MN}}, EH_{MN}$,and $E_{V_{MN}}$ are the energies of all the wavelet coefficients in the candidate region of size $M \times N$ corresponding to average DC, horizontal and vertical information, respectively. Face verification can be achieved if $\frac{E_{DC_{MN}}}{E_{MN}} < Th_{DC}, \frac{E_{H_{MN}}}{E_{MN}} > Th_H, \frac{E_{V_{MN}}}{E_{MN}} > Th_V$ then only the candidate block is declared as face block, where Th_{DC}, Th_H, and Th_V are the threshold values. The reason is that the face region should not have near 100% energy in DC coefficients. Also, the energy corresponding to horizontal and vertical details should be large enough. Using these thresholds, each candidate face region declared by neural network skin-color classifier is verified.

6. EXPERIMENTAL RESULTS

A multilayer perceptron (MLP) is trained to classify pixels into skin-tone and non-skin-tone. The input vector to the network consists of *(Cb,Cr)* values corresponding to chrominance. Image database we used consists of 240 images of positive and negative training patterns. Also, 150 images containing group photos were collected. Skin colors from various races of the world are collected in the form of 32×32 pixels per skin sample for each individual from each image. 120 such samples were collected *Cb-Cr* planes. As a result, there are 122,880 skin pixels having different illuminations in our skin color database used for training the neural network. The neural networks were trained using Levenberg-Marquardt(LM) method [30] in order to generate binary outputs for skin and non-skin. The Levenberg-Marquardt method will have the fastest convergence compared with other methods such as conjugate gradients or gradient. In general,the LM algorithm will have the fastest convergence for networks that contain up to a few hundred weights, like the topology being brought up by the authors of this article. The scalar output of MLP is converted into binary output, 0 as non-skin and 1 as skin, using a fixed threshold $\tau = 0.35$ and sigmoid as activation function. The MLP classifiers that we trained have one hidden layer ranging from 9 to 25 neurons; furthermore, different network sizes were investigated but we only report the performance of the most efficient network. The best result was 89.6 correct classification achieved with the neural networks of size (2-25-1) (one hidden layer of 25 neurons),and the false detection and false dismissal rate were 4.6% and 4.5% , respectively . We investigate that the appropriate skin-color decision boundary is generated by training neural networks with 1000 to 5000 samples.

The algorithm starts at the LL subimage, a lower resolution version of the image obtained from the wavelet transform, so that the amount of data to be processed is greatly reduced. The lower resolution image is sufficient for the detection of face regions rather than detailed low-level features. We apply the above skin classifier to *(Cb,Cr)* values (Figure 4.(c)) corresponding to LL subimages of chrominance to check for candidate face pixels. As we are using four levels of wavelet transform, each pixel in LL subimage corresponds to 16×16 pixels in the original image. So, whenever any pair of *(Cb,Cr)* gets classified as skin pixel, it means that the corresponding area of 16×16 pixels with respect to this pair is a face block. After the classification, a binary mask image is obtained for each image (Figure 4.(d)), but with a reduced resolution. Each value in the mask image indicates the classification results of the corresponding block of size 16×16 in the original image. A

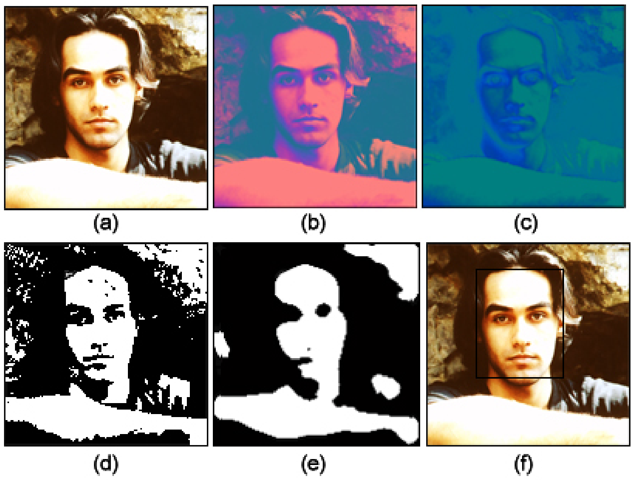

Figure 4: (a) Original image. (b) Shows the image in *YCbCr* space. (c) The chrominance components of original image.(d) Shows the scaled binary mask image after the skin-color classification.(e) The result of applied median filter. (f) Face detection result.

median filter is applied to the above generated binary mask image to remove noise and fill in the holes (Figure 4.(e)). Eventually ,face verification task is accomplished by the aforementioned energy distribution of wavelet transform coefficients of luminance components (Figure 4.(f), 5) .

Figure 5: Face detection results on still images.

7. CONCLUSIONS

The proposed method has incorporated the concept of the skin-color model, neural networks and a wavelet representation based face detection technique to provide an efficient face detection algorithm.A skin and non-skin color classifier using multilayer perceptron is presented. The major advantage of the new method is that accurate approximation of the decision boundary for skin colors in *Cb-Cr* planes is achievable with small-sized networks. The neural network based model has been shown to provide remarkable coverage of all human skins. The utilized MLP classifier is a good candidate if low memory usage is also a requirement. This also provides a promising direction for the efficiently and accurately extracting skin irrespective of color of the skin as is evident from the results. One of the most remarkable merits of our proposed algorithm is that, unlike a plethora of others, it attenuates the problem of containing exhaustive searches. The computation time has been reduced considerably. However, it is to be noted that the training neural networks is performed off-line ;hence, neural networks computational cost is not substantial ,and computation time needed to calculate is the forward wavelet transform of the image in real-time processing. By adjusting the thresholds in all stages, face detection rate can be controlled depending upon the application. Nevertheless, it is not entirely full-blown and suffers from a couple of shortcomings and cannot be taken into account as generalized. Moreover,it can only be applied to color images, because of the use of chrominance information. The algorithm gives false alarms under non-uniform lighting conditions which is seemingly inevitable in such algorithms. False dismissals cannot be totally avoided, especially in very cluttered scenes with many small faces.

In the final analysis, despite its restrictions, the proposed face detection is efficient and can be applied to large image databases for indexing and recognition. Once these face regions are detected, they can be further used for face tracking, and face recognition using more sophisticated techniques. Further work is in progress to develop a real-time face tracking and recognition system and index individuals for surveillance purposes.

Acknowledgement

We would like to express our gratitude and acknowledgement to Mr. Reza Bahrami for his feedback on this article.

8. REFERENCES

[1] N. Ahuja M. Yang, D. Kriegman, "Detecting faces in images: A survey," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24 (1), pp. 34–58, 2002.

[2] C.L. Wilson R. Chellappa and S. Sirohey, "Recognition of faces: A survey," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 83(5), pp. 705–740, 1995.

[3] K.K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 23(1), pp. 39–51, 1998.

[4] S. Baluja H.A. Rowley and T. Kanade, "Neural network-based face detection," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 20(1), pp. 23–28, 1998.

[5] J.E. Viallet and M.Collobert R. Fraud, O.J. Bernier, "A fast and accurate face detection based om neural networks," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 23(1), pp. 42–53, 2001.

[6] H. Ellis I. Craw and J. R. Lishman, "Automatic extraction of facefeature," *Pattern Recognition Letters*, pp. 183–187, 1987.

[7] J. S. Kim C. H. Lee and K. H. Park, "Automatic human face location in a complex background," *Pattern Recognition Letters*, vol. 29, pp. 1883–1889, 1996.

[8] J. Dugelay F. Perronnin and K. Rose, "Deformable face mapping for person identification," in *in Proc. ICIP, Barcelona, Spain*, 2003.

[9] P. Viola and M. Jones, "Robust real-time object detection," in *in 2nd International Workshop on Statistical and Computational Theories of Vision-Modelling, Learning, Computing, and Sampling, Vancouver*, 2001.

[10] M. Grudin, "On internal representation in face recognition systems," *Pattern Recognition*, vol. 33, pp. 1161–1177, 2000.

[11] K. M Lam and H. Yan, "An analytic-to-holistic approach for face recognition based on single frontal view," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 20(7), pp. 673–686, 1998.

[12] M.Abdel-Moualeb R.L.Hsu and A.K.Jain, "Face detection in color images," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24(2), pp. 696–706, 2002.

[13] W. Lu J. Yang and A. Waibel, "Skin-color modeling and adaptation," Tech. Rep., Carnegie Mellon University, 1997.

[14] J. Yang and A. Waibel, "A real-time face tracker," in *IEEE Workshop Applications of Computer Vision*, 1996, pp. 142–147.

[15] H. Schneiderman and T. Kanade, "A statistical model for 3d object detection system applied to faces and cars," in *Proc. of IEEE Conf. Computor Vision and Pattern Recognition*, 2002.

[16] P. Bartlett R. E. Schapire, Y. Freund and W. S. Lee, "Boosting the margin: a new explanation for the effectiveness of voting methods," *Ann. Stat*, vol. 26(5), pp. 16511686, 1998.

[17] S.C. Ahn S.H. Kim, N.K. Kim and H.G. Kim, "Object oriented face detection using range and color information," in *in Proc. of the Third International Conference on Automatic Face and Gesture Recognition, Nara, Japan*, 1998, pp. 76–81.

[18] Y. Dai and Y. Nakano, "Face-texture model based sgld and its application," *Pattern Recognition Letters*, vol. 29, pp. 1007–1017, 1996.

[19] E. Saber and A. M. Tekalp, "Frontal-view face detection and facial feature extraction using color, shape, and symmetry based cost functions," *Pattern Recognition Letters*, p. 19, 1998.

[20] M. Abdel-Mottaleb and A. Elgammal, "Face detection in complex environments," in *in Proceedings International Conference on Image Processing*, 1999.

[21] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions, merging and wavelet packet analysis," *In IEEE Transaction on multimedia*, vol. 1(3), pp. 264–277, 1999.

[22] H. F. M. N. Shirazi J. C. Terrillon and S. Akamatsu, "Comparative prefromance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," in *in Proceedings International Conference on on Face and Gesture Recognition*, 2000, pp. 54–61.

[23] Bishop, *Neural Network for Pattern Recognition*, Clarendon Press, Oxford, June 1995.

[24] S. Haykin, *Neural Networks: A Comprehensive Foundation*, N. Y. Macmillan, 1994.

[25] J. M. Shapiro, "Embedded image coding using zerotree of wavelet coefficients," *IEEE Transaction on Signal Processing*, p. 3462, 1993.

[26] P. G. Poonacha J. Karlekar and U. B. Desai, "Image compression using zerotree of wavelet coefficients and multistage vector quantization," in *in Proc. of International Confference on Image Processing. (ICIP '97), Santa Barbara, USA*, 1997.

[27] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 11(7), pp. 674–693, 1989.

[28] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36(5), pp. 9611005, 1990.

[29] S. G. Mallat, "Multiresolution channel decomposition of images and wavelet models," *In IEEE Transaction on Acoustic, Speech, and Signal Processing*, vol. 37(12), pp. 2091–2110, 1989.

[30] G.P. Drago and S. Ridella, "Statistically controlled activation weight initialization(scawi)," *IEEE Transaction on Neural Networks*, vol. 3, pp. 899905, 1992.