

Face Detection in Color Images using Wavelet Packet Analysis

C. Garcia, G. Zikos, G. Tziritas

*Institute of Computer Science
Foundation for Research and Technology-Hellas
P.O.Box 1385, GR 711 10 Heraklion, Crete, Greece
E-mail: {cgarcia,gzikos,tziritas}@csi.forth.gr*

Abstract

In this paper, we propose a novel scheme for automatic and fast detection of human faces in color images where the number, the location, the orientation and the size of the faces are unknown, under non-constrained scene conditions such as complex background and uncontrolled illumination. First, each frame is segmented using skin chrominance values, providing face area candidates. Then, shape analysis and wavelet packet decomposition are performed on the face area candidates in order to detect human faces. Each face area candidate is described by a subset of band filtered images containing wavelet coefficients. These coefficients characterize the face texture and a set of simple statistical data is extracted in order to form compact and meaningful feature vectors. Then, an efficient and reliable probabilistic metric derived from the Bhattacharyya distance is used in order to classify the face area candidate feature vectors into face or non-face areas.

1. Introduction

Face detection may be viewed as a very important tool for a content-based indexing video retrieval system. Such a system is currently developed within the Esprit project DIVAN [1] which aims at building and evaluating a distributed audio-visual archives network providing a community of users with facilities to store video raw material, and access it in a coherent way, on top of high-speed wide area communication networks. The video raw data is first automatically segmented into shots and from the content-related image segments and keyframes, salient features such as region shape, intensity, color, texture and motion descriptors are extracted and used for indexing and retrieving information. In order to allow queries at a higher semantic level, some particular pictorial objects have to be detected and exploited for indexing. We focus on human faces detection, given that such data are of great interest for users queries.

Face recognition algorithms have received most of the attention in the academic literature compared to face detection

algorithms. In recent years, considerable progress has been made on the problem of face recognition, especially under stable conditions such as small variations in lighting, facial expression and pose. A detailed survey is presented in [2].

However, detecting and locating a face is equally important and is as difficult as recognizing it. In most of the approaches for face recognition, the existence and location of human faces in the processed images are known *a priori*, so there is little need to detect and locate faces. In image and video databases, there is generally no constraint on the number, location, size, and orientation of human faces and the background is generally complex.

Some techniques have been developed recently for detecting faces in “non-mugshot” images. A good survey may be found in [5]. These methods can be roughly divided into three broad categories: local facial features detection, template matching and image invariants. In the first case, low level computer vision algorithms are used to detect facial features such as eyes, mouth, nose and chin and statistical models of human face are used like in [6, 7, 8, 9]. In the second case, several correlation templates are used to detect local sub-features which can be considered as rigid in appearance (view-based eigenspaces [3]) or deformable (deformable templates [4]). In the last case, image-invariants schemes assume that there are certain spatial image relationships common and possibly unique to all face patterns, even under different imaging conditions [10]. Instead of detecting faces by following a set of human-designed rules, alternative approaches are based on neural networks [11, 12] which have the advantage of learning the underlying rules from a given collection of representative examples of face images, but have the major drawback of being very computationally expensive and challenging to train because of the difficulty in characterizing “non-face” representative images.

In this paper, we propose a novel algorithm for automatic and fast detection of human faces in both digital still color images or in MPEG compressed video frames. The proposed scheme starts by performing a chrominance-based segmentation using two different color models, the YCbCr

model being naturally related to MPEG streams and JPEG images and the HSV model being used mainly in computer graphics. The proposed algorithm has been designed to index a huge amount of video and digital images and has to cope with high-speed requirements. In the case of MPEG streams, only I frames are analyzed as we want to avoid costly decompression. Color segmentation of these I frames is performed at MPEG macro-block level (16×16 pixels) following the approach of Wang and Chang in [5]. The same strategy is applied in the case of still images, by analyzing them at the same coarser resolution of 16×16 pixels. First, each frame is segmented according to skin chrominance analysis, providing face area candidates. Then, shape analysis and wavelet packet decomposition are performed on each face area candidates in order to detect human faces. Each face area candidate is described by a subset of band filtered images containing wavelet coefficients. These coefficients characterize the face texture and a set of simple statistical moments is extracted in order to form compact and meaningful feature vectors. Then, an efficient and reliable probabilistic metric derived from the Bhattacharyya distance is used in order to classify the face area candidate feature vectors into face or non-face areas, using some prototype face area vectors, acquired in a previous training stage.

2. YCbCr and HSV Skin color sub-spaces

The first stage of the proposed scheme consists in locating the potential face areas in the image, using chrominance information only. Two color models have been evaluated and used. The first one is the YCbCr model used in MPEG and JPEG coding. The second one is the HSV (Hue, Saturation, Value) model which is considered by many to be more intuitive to use, closer to how an artist actually mixes colors. Skin color patches are used in order to approximate the color sub-space characterizing skin color, in both models. The training data set is composed of 950 skin colors samples which have been extracted from various still images and video frames, covering a large range of skin color appearance (different races, different lighting conditions). We observed that skin color samples form a single and quite compact cluster in both YCbCr and HSV spaces. Our purpose is to approximate the boundaries of these two clusters.

In the YCbCr case, it can be noticed that the intensity value Y has little influence on the distribution over the CbCr plane and that sample skin colors form a small and very compact cluster over the CbCr plane. Wang and Chang in [5] performed skin colors classification directly in the chrominance plane (CbCr) without taking into account the intensity value Y . We decided to estimate the shape of the skin color sub-space in YCbCr, using the intensity (Y) value to cope with strong lighting variations, as we noticed that the distribution turns out to be different for the extrema of Y values corresponding to dark (for values of Y around 50)

and light (for values of Y around 240) lighting conditions.

Given that quite a large amount of noise is contained in the skin color samples (dark areas of the face, facial hair, etc...), we found efficient to determine the sub-space borders using linear approximations that can be easily adjusted. A set of planes equations have been found by successive adjustments according to segmentation results. In Fig. 1.a, we plot the intersections of the adjusted bounding planes with the CbCr plane for $Y = 160$. One may notice that some sample points are not contained in the skin color areas. These points have not been included during the successive adjustments given that they correspond to non-representative samples of the skin colors and cause errors in the face detection process.

We give hereafter the equations defining the bounding planes that have been found. As one can notice, there are two sets of eight equations depending on two areas of the color space, separated by the horizontal plane $Y = 128$, in order to approximate the distribution borders in the light and dark extreme cases.

$$\begin{cases} \text{if } (Y \leq 128) & \theta_1 = -2 + \frac{256-Y}{16}; \theta_2 = 20 - \frac{256-Y}{16}; \theta_3 = 6; \theta_4 = -8 \\ \text{if } (Y > 128) & \theta_1 = 6; \theta_2 = 12; \theta_3 = 2 + \frac{Y}{32}; \theta_4 = -16 + \frac{Y}{16} \end{cases}$$

$$\begin{cases} Cr \geq -2(Cb+24); Cr \geq -(Cb+17); Cr \geq -4(Cb+32); \\ Cr \geq 2.5(Cb+\theta_1); Cr \geq \theta_3; Cr \geq -0.5(Cb-\theta_4); \\ Cr \leq -\frac{Cb-220}{6}; Cr \leq -1.34(Cb-\theta_2); \end{cases}$$

We noticed that the cluster of skin color is less compact in HSV space than in YCbCr space. The projection onto the HS plane only is used by some authors like in [15] where skin color classification is performed by setting appropriate thresholds to Hue and Saturation. Using these thresholds, we found that the segmentation results are affected by variations in lighting conditions.

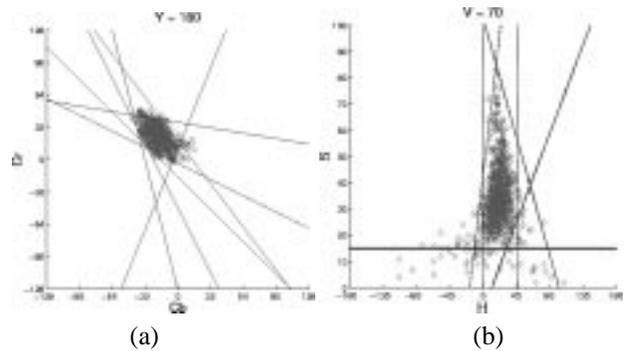


Figure 1. Skin color bounded areas

Like in the YCbCr case, we directly estimated the shape of the skin color sub-space in HSV. A set of planes equations have been found by successive adjustments according to segmentation results. In Fig. 1.b, we plot the intersections of the adjusted bounding planes with the HS plane for $V = 70$. We give hereafter the equations defining the six bounding

planes that have been found in the HSV color space case.

$$\begin{cases} H \geq 0; S \geq 15; S \geq 0.75H + 0.3V - 30 \\ S \leq -H - 0.1V + 110; H \leq -0.4V + 75; S \leq 0.08(100 - V)H + 0.6V \end{cases}$$

Segmentation results are quite equivalent using both colors models. We noticed that even if the skin color cluster is more compact over the CbCr plane than over the HS plane, bounding planes are more easily adjusted using the HSV model, because of a direct access to H (Hue) which mainly encodes skin colors.

3. Detection of skin color regions

Using one of these two models, if the mean color value of a macro-block belongs to the corresponding skin color sub-space, this macro-block is a possible candidate to be part of a face. After classification, a binary mask image that is denoted by BM is obtained for each frame. In this macro-block mask image, a macro-block set to "1" corresponds to a macro-block whose average color is a skin color. Then, a filter is applied to the binary mask image to remove noise and smooth the image, given that faces correspond to connected regions in the mask and are homogeneous in chrominance. Each macro-block mb_k is updated to "1" or "0" according to the values of its neighbors macro-blocks and a color similarity distance.

Then, we propose to scan through BM in order to detect and extract the candidate face regions. Our goal is to find continuous regions in the mask image that may be bounded by rectangles that are denoted by bm_i . To respect speed requirements, a simple method of integral projection is applied, like in [5]. Each macro-block mask image is segmented into non-overlapping rectangular regions that contain either no "1" or a contiguous "1" region, analyzing the zero-runs and the nonzero-runs of simple horizontal and vertical projections. Four different examples of chrominance-



Figure 2. Four different examples of the chrominance segmentation process

based segmentation are given in Fig. 2. In each column, we show the binary mask image BM and the non-overlapping bm_i areas (first line) and the final detection results (second line). In the first example, one can notice that four different

bm_i non-overlapping areas have been selected, in which face detection will be achieved. In the second example, one area has been selected. In that case, it can be seen that facial parts have been excluded from the resulting BM , because the corresponding colors are not included in the approximated skin color sub-space. The third and fourth examples show that, in some cases, the background have a color similar to skin color.

4. Detection of candidate face regions

Our goal is to locate candidate face regions that we denote by CF_{ij} in each bm_i areas. Given that we don't know *a priori* the size of the face regions contained in the frame, we first search for the largest possible CF_{ij} areas and we iteratively reduce their size and run over the possible aspects ratios and the possible positions in each bm_i area. Constraints related to size range, shape (aspect ratio) and homogeneity are applied to detect these CF_{ij} regions. The allowed aspect ratio range is set to $[1.0, 1.8]$, which was chosen quite large in order to cope with different orientations and poses. Moreover, constraints related to allowed sizes are applied. The range of allowed size is lower-bounded by 5×3 macro-blocks, i.e., 80×48 pixels. It is naturally upper-bounded by the size of the whole frame image. The constraint of color homogeneity is applied by computing the density of skin color macro-blocks in the candidate face area. This is done easily by checking the density of "1" into the candidate face area binary mask. Two areas, respectively the *outer* and the *inner* areas are defined into each CF_{ij} , the *outer* corresponding to the border area of CF_{ij} (whose width is a percentage of the bounding box width, typically 15%) and the *inner* area corresponding to the remaining central part of CF_{ij} . A CF_{ij} area is accepted as a potential face area if the percentage of "1" in the *outer* mask area is above a threshold set to 65% and the percentage of "1" in the *inner* mask area is above a threshold p described by (1). Homogeneity in the *inner* area has to be high given that this area is the center of the face with no background whereas in the *outer* area, hair and parts of background have to be taken into consideration, and therefore the threshold that we use has to be less restrictive.

$$\begin{cases} p = 1 - 0.3 \frac{(w-2\delta)(h-2\delta)}{w h} & \text{if } w > 2\delta \\ p = 0.80 & \text{otherwise} \end{cases} \quad (1)$$

where w and h are respectively the width and the height of CF_{ij} (in macro-blocks units) and $\delta = 2$.

5. Classification of candidate face areas based on wavelet packet analysis

The main purpose of the last stage of the proposed algorithm is to verify the face detection results obtained after CF_{ij} segmentation and remove false alarms caused by objects with color similar to skin colors and similar aspect ratios such as exposed parts of the body or part of background.

For this purpose, a wavelet packet analysis scheme is applied using a method which is similar to the one we described in [13, 14] for face recognition. For a candidate face region, we search for a face in every position inside the region and for a number of possible dimensions for its bounding rectangle. This search is based on the classification of feature vectors which are extracted from a set of wavelet packet coefficients in the corresponding regions. This decomposition is performed on the intensity image of each region. At this stage of development, we applied the wavelet packet analysis to Y images only. In the case of HSV images, a transform from HSV to YCbCr is performed. These intensity images convey information about the faces texture which will be kept into the wavelet coefficients.

5.1. Wavelet packet analysis

The main characteristic of wavelets (if compared to other transformations) is the possibility to provide a multi-resolution analysis of the image in the form of coefficient matrices with a spatial and a frequential decomposition of the image at the same time. A complete framework has been recently built [16, 17] 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 (H and G respectively) 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*.

In classical wavelet decomposition, the image is divided into an approximation and details images. The approximation is then divided itself into a second-level approximation and details. The *wavelet packet decomposition*, that we perform in the proposed scheme, is a generalization of the classical wavelet decomposition that offers a richer signal analysis (discontinuity in higher derivatives, self-similarity,...). In that case, details as well as approximations can be split. This results in a wavelet decomposition tree. The filters that we applied have been chosen to be Conjugate Quadrature Filters, according to [18] and series of experiments. H is a symmetrical filter of support size 10 with the following left part: $[0.0378, -0.0238, -0.1106, 0.3774, 0.8526]$ and $G_{n+1} = (-1)^n H_n$ with $n \in \{1, 10\}$.

Usually, an entropy-based criterion is used to select the deepest level of the tree, while keeping the meaningful information. In the present case, the depth of the wavelet packet decomposition tree is ruled by the size of the CF_{ij} areas which are processed. We formed a database of 50 manually extracted face areas which contains a large number of different cases (size, chrominance, intensity, position and orientation). The processed frames have a size of 288×352

pixels. These samples are used in order to estimate prototype vectors. These samples have been classified into two categories, according to their size. A face is classified as *medium* if its height is smaller than 128 and as *large* if its height is bigger than 128. Based on lots of experiments, we decided to decompose the candidate faces intensity images with the discrete wavelet packet analysis until level 3 for *large* areas and until level 2 for *medium* areas. In both cases, a deeper decomposition will provide no more valuable information, the coefficient images becoming too small. According to its size, a candidate face image is described by a set of n wavelet coefficient matrices belonging to the deepest level of decomposition, ($n = 16$, i.e., one approximation image and 15 details images for *medium* areas and $n = 64$, i.e., one approximation image and 63 details images for *large* areas) which, in all cases, represent quite a huge amount of information (equal to the size of the input image). Dimensionality reduction has to be performed by extracting discriminatory information from these images.

5.2. Feature vectors extraction

In our approach for face recognition [14], we reported good results obtained by extracting wavelet coefficients in some different specific areas of the face. We first located the face bounding box and then we divided it into two areas, the top part and the bottom part, separated by the nose baseline. In the present case, we follow a slightly different approach by considering each CF_{ij} as a bounding box, and by dividing it into 4 parts: a left top part (top_1), a right top part (top_2), a left bottom part ($bottom_1$) and a right bottom part ($bottom_2$), all of equal size, where the wavelet packet analysis is performed. By extracting moments from wavelet coefficients in these areas, we obtain information about the face texture, related to different facial parts, like the eyes, the nose and the mouth, including facial hair.

In [14], we extracted means and variances (μ_i and σ_i^2) of each wavelet coefficients matrices as statistical moments describing the face texture. Variances were extracted from the wavelet coefficients images to get a global measurement of the face texture, according to different filtering directions. Means extracted from the details images were null (like in the present case), according to the design of the bank filters that we applied. Means extracted from the approximation images were used to get an average intensity value of the face to be recognized. These faces came from the FERET database, presenting small variations of lighting conditions. In the present case, means values are not used because faces to be detected can appear in a lot of different lighting conditions. Only variances are used as face descriptors coefficients.

Therefore, from the top and bottom areas, we extract the corresponding variances $\sigma_{top1}^2, \sigma_{top2}^2, \sigma_{bottom1}^2$ and $\sigma_{bottom2}^2$ of the wavelet coefficients contained in the approximation image of the selected level of decomposition. From the m detail images ($m \in \{15, 63\}$), the corresponding variances

σ_i ($i=4, \dots, n$, $n = m + 3$) are extracted from the CF_{ij} area.

Thus, extracted feature vectors contain a maximum of $4 + m$ components (4 variances for the approximation image and m variances for the details images) and are described as follows: $\mathcal{V} = \bigcup_{i=0}^{m+3} \{\sigma_i^2\}$ where indices $i = 0, 1, 2, 3$ stand respectively for the top_1 , top_2 , $bottom_1$ and $bottom_2$ variances. Finally, a set of vectors of size 19 and 67 for *medium* and *large* CF_{ij} areas, respectively, is built.

5.3. Feature vectors classification

A CF_{ij} feature vector has to be classified into two possible classes: face or non-face. From the database of manually extracted face areas, that we previously classified into two size categories, features vectors have been extracted and an average prototype feature vector has been retained for each of these two categories. Therefore, classification is performed by evaluating the following distance \mathcal{D} from each CF_{ij} feature vector \mathcal{V}_k to the prototype feature vector \mathcal{V}_l of the corresponding size category:

$$\begin{aligned} \mathcal{D}(\mathcal{V}_k, \mathcal{V}_l) &= \frac{1}{2} \sum_{i=0}^{m+3} \ln \left[\frac{\frac{1}{2}(\sigma_{ik}^2 + \sigma_{il}^2)}{\sqrt{\sigma_{ik}^2 \sigma_{il}^2}} \right] \\ &+ \frac{1}{2} \sum_{C \in \{Cb, Cr\}} \ln \left[\frac{\frac{1}{2}(\sigma_C^2 + V_C)}{\sqrt{\sigma_C^2 V_C}} \right] \end{aligned} \quad (2)$$

where σ_{Cb}^2 and σ_{Cr}^2 are the variances of the color components Cb and Cr inside CF_{ij} and V_{Cb} and V_{Cr} are pre-computed values corresponding to the average Cb and Cr variances values over the set of reference face patches. The mathematical form of both terms in \mathcal{D} is derived from the probabilistic Bhattacharyya distance. The first term aims at comparing the two feature vectors, while the second term acts as a constraint on color homogeneity over the candidate face areas.

Each CF_{ij} feature vector has to be classified as face or non-face according to distance \mathcal{D} to the average prototype vector of the corresponding size category. CF_{ij} is classified as a face area if \mathcal{D} is below a threshold T_{HD} and rejected otherwise. A T_{HD} value of 7.0 has been found to be optimal after a large number of experiments for both size categories.

Finally, problems of overlapping selected CF_{ij} areas are solved as follows. The set of overlapping face areas is sorted in ascending order according to normalized distances $\frac{\mathcal{D}}{h^2 \sqrt{hw}}$, related to the size of CF_{ij} . Then, the first ranked face area is selected and the other ones are rejected. This criterion is used to retain the biggest face regions when overlaps are considered. Extension to this rule may be necessary when two face areas share a small intersecting surface, then both of them may be kept.

6. Experimental results and discussion

The proposed algorithm has been evaluated using a test data set which contains images that have been extracted as key-frames from various MPEG videos and specifically

from the test videos used in the DiVAN project evaluation phase [1]. The video material has been kindly provided by the *Institut National Audiovisuel, France* and by *ERT television, Greece*. The test data set contains 100 images, most of them being extracted from advertisements, news, movies and external shots. This set of 100 images contains 104 faces (with sizes above the minimal one) and 10 images do not contain faces. They cover most of the cases that the algorithm has to deal with (48 frontal views, 25 semi-frontal views, 18 side-views, 13 tilted poses).



Figure 3. Some results for the test data set

In Fig. 3, we present some results of the proposed face detection scheme applied on the test data set. These examples include frames with multiple faces of different sizes, different colors, different positions and frames with no face. False alarms and false dismissals examples are presented as well. The algorithm detects 97 of the 104 faces which means a successful detection rate of 93.27%. 78.85% of the detected faces have been framed correctly while 14.42% have been framed with less accuracy. This is mainly due to the macro-block approach which provides a rough approximation of the CF_{ij} areas. In some cases, moderate framed faces are obtained because of occlusion or extreme lighting conditions. 23 false alarms and 7 false dismissals (6.73%) are obtained. False alarms appear because some regions have colors and textures of faces, without corresponding to human faces. False dismissals cannot be totally avoided,

especially in scenes with many partially occluded faces or under extreme lighting conditions. Among the 7 false dismissals cases, 4 faces were either very light or very dark and 3 faces were small, close to the lower bound allowed size. On a SUN ULTRA-2 workstation, the average run time is 5.400 secs. On a SGI O2 workstation the average run time is 2.440 secs. The reported time of the algorithm presented in [5] was 32.6 ms on a SPARC 5 workstation. It has to be noticed that wavelet decomposition is the main time consuming part of the proposed scheme. Special hardware may be used and we believe that a similar processing time will be obtained, with better results.

7. Conclusion

Our experiments have shown that an appropriate chrominance segmentation using YCbCr or HSV color models at macro-block level followed by a wavelet packet analysis provide a robust scheme for face detection even if no constraint is imposed on the faces to be detected (except a minimal size). For a data set of 100 images with 104 faces covering most of the cases of human faces appearance, a 93.27% good detection rate, 23 false alarms and a 6.73% false dismissals rate were obtained. Thus, the wavelet transform proved to provide an excellent image decomposition and texture description. In addition to this, very fast implementations of wavelet decomposition are available in hardware form. We show that even very simple statistical features such as variances provide an excellent basis for face detection, if an appropriate distance is used. As an extension of this work, we believe that it would be interesting first to enhance the chrominance segmentation stage by performing skin colors clustering in the bm_i areas, especially when a face is surrounded by a skin color background. The particular color of the face may be distinguished from the background by clustering these two sets of skin colors. Second, a slightly different wavelet decomposition may be performed, by considering only the details images which convey most information about faces texture.

Acknowledgements

This work was funded in part under the DiVAN Esprit Project EP 24956. We would like to thank the INA's Innovation Department and ERT television for having provided test video materials.

References

[1] DiVAN Esprit Project EP24956, "Distributed audioVisual Archives Network", <http://divan.intranet.gr/info>, 1997.
 [2] C. L. Wilson, C. S. Barnes, R. Chellappa, S. A. Sirohey, "Face Recognition Technology for Law Enforcement Applications", *NISTIR 5465, U.S. Department of Commerce*, 1994.

[3] A. Pentland, R.W. Picard, S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases", *Proceedings of the SPIE Storage and Retrieval and Video Databases II*, 2185, San Jose, CA, 1994.
 [4] L. Wiskott, JM. Fellous, N. Kruger, C. Von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 1997, pp. 775-779.
 [5] H. Wang and S.-F. Chang, "A Highly Efficient System for Automatic Face Region Detection in MPEG Video", *IEEE Transactions on Circuits and Systems For Video Technology*, 7(4), 1997, pp. 615-628.
 [6] A. Eleftheridis and A. Jacquin, "Model-assisted coding of video teleconferencing sequences at low bit rates", *Proceedings of IEEE Int. Symp. Circuits and Systems*, May/June 1994, pp. 3.177-3.180.
 [7] G. Yang and T.S. Huang, "Human face detection in a complex background", *Pattern Recognition*, 27(1), 1994, pp. 55-63.
 [8] S.-H. Jeng, H. Y. M. Yao, C. C. Han, M. Y. Chern and Y. T. Liu, "Facial Feature Detection Using Geometrical Face Model: An Efficient Approach", *Pattern Recognition*, 31(3), 1998, pp. 273-282.
 [9] K. C. Yow, C. Cipolla, "Feature-based human face detection", *Image and Vision Computing*, 15, 1997, pp. 713-735.
 [10] P. Sinha, "Object Recognition via Image Invariants: A Case Study", *Investigative Ophthalmology and Visual Science* 35, 1994, pp. 1.735-1.740.
 [11] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1998, pp. 23-28.
 [12] S.-H. Lin, S.-Y. Kung and L.-J. Lin, "Face Recognition/Detection by Probabilistic Decision-Based Neural Network", *IEEE Transactions on Neural Networks*, 8(1), 1997, pp. 114-131.
 [13] C. Garcia, G. Zikos, G. Tziritas, "A Wavelet-Based Framework for Face Recognition", Presented at *Int. Workshop on Advances in Facial Image Analysis and Recognition Technology*, 5th European Conference on Computer Vision Freiburg, Germany, June 2 - 6, 1998.
 [14] C. Garcia, G. Zikos, G. Tziritas, "Wavelet Packet Analysis for Face Recognition", To appear in *Image and Vision Computing*, 1999.
 [15] S. Tsekeridou and I. Pitas, "Facial Features Extraction in Frontal Views Using Biometric Analogies" *Proceedings of the IX European Signal Processing Conference*, I, Rhodes, Greece, 1998, pp. 315-318.
 [16] J.S. Mallat, "A theory of multiresolution signal decomposition. The wavelet representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 1989, pp. 674-693.
 [17] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis", *IEEE Transactions on Information Theory*, 36(5), 1990, pp. 961-1005.
 [18] M. Smith and T. Barnwell, "Exact reconstruction techniques for tree structured subband coders", *IEEE Transactions on Acoust. Speech, Signal Processing*, ASSAP-34(3), 1986, pp. 434-441.