We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

4,900
Open access books available

124,000
International authors and editors

140M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

The Methodology for Facial Features Detection

Jacek Naruniec

Warsaw University of Technology, Institute of Radioelectronics
Poland

1. Introduction

Face detection is an important preprocessing task in biometric systems based on facial images. The result of the detection derives the localisation parameters and it could be required in various forms (Figure 1), for instance:

- a rectangle covering the central part of face,

- a larger rectangle including forehead and chin,

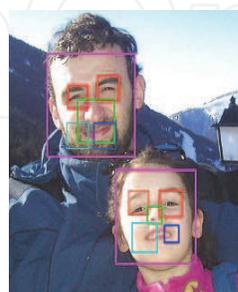- irregular mask of the face area,
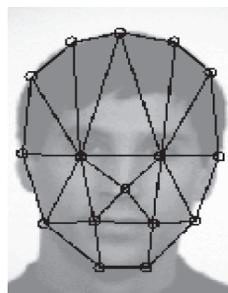
face graph Wiskott et al. (1997)

set of the face fiducial points Vukadinovic & Pantic (2005)

set of face fiducial points placed on face parts contours Cootes et al. (1998)

rectangles covering face parts and the face itself Erukhimov & Lee (2008)

Fig. 1. Different methods of representing face fiducial points and face parts.

- eyes centers,
- multiple face fiducial points,
- contours of the face parts,
- a set of rectangles covering individual parts of the face.

While from human point of view the area parameters are more convincing, for face recognition system, fiducial points are more important since they allow to perform facial image normalization – the crucial task before facial features extraction and face matching. Facial features localization algorithms are commonly divided into four groups (Ben Jemaa & Khanfir (2009); Celiktutan et al. (2008); Naruniec (2010)):

- appearance-based,
- geometry-based,
- knowledge-based,
- 3D vision-based.

This chapter aims at defining more general scheme for facial features localization, since all of the defined groups have common methodology. Moreover most of the efficient schemes doesn't rely on one of the methods, rather combining few different approaches. The goal of this work isn't concerned about giving the detailed information about each algorithm, but rather to develop the intuition of the approach to the described subject. Review of the given steps is presented in the following sections.

2. General facial features detection scheme

Typical image analysis and in particular facial features detection usually consists of several steps:

1. Preprocessing.
2. Defining regions of interest.
3. Features extraction.
4. Classification.
5. Postprocessing.

In many cases not all parts of the process are used, but their subset is always present. Example of the system consisting of all of the proposed steps may be observed in the face detection algorithm by Discrete Gabor Jets (Naruniec & Skarbek (2007)). Within this approach image is preprocessed by the Gaussian blurring to reduce the influence of the noise. Since all of the detected fiducial points are placed on the edges, the regions of interest are obtained by thresholding the magnitude of the Sobel filters response. Extraction is performed in the circular neighborhood of each edge pixel using FFT, integral image and simple min/max normalization. Classification is based on the modification of linear discriminant analysis adjusted for the two-class (fiducial point/non-fiducial point) problem. In the final steps all of the edges are assigned to one of the five categories: left eye corner, right eye corner, left nose corner, right nose corner or non-face point. Verification of the points - postprocessing step, is performed by fitting the defined points to the face graph.

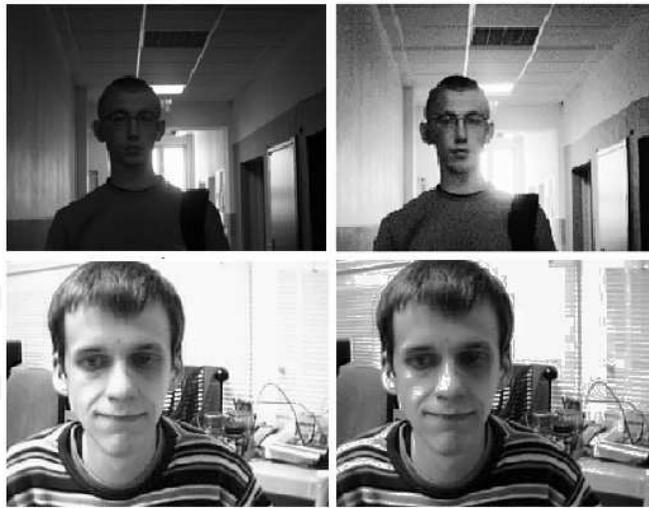Subsequent section describe each of the processing steps in more detail.

Fig. 2. Histogram equalization examples. On the left - original images, on the right - processed images. On the first image equalization significantly improves the quality of the face image, while on the latter the effect is opposite.

2.1 Preprocessing and defining ROI

Since facial features are placed on the specific regions, a procedure for removing most of the background can be defined. This stage has to be proceeded very carefully, because removing proper regions at this moment will result in a failure of the whole algorithm. Region of interest (ROI) is defined here as a facial feature candidate point or region.

Depending on the representation of the facial features, methods can be divided to region based and point based. Contour based ROI definition could be also defined, but these methods aren't usually used within this problem and thus will not be discussed here.

In some applications preprocessing may increase the accuracy of the localization. This applies mostly to the cases where the acquisition parameters are insufficient, for example poor lighting, noise or inadequate camera properties. Typical operations performed on the data includes:

- noise removal - Gaussian blurring, median, mean filters;
- lighting normalization - min/max normalization, histogram equalization, removing low-pass frequencies;
- removing camera distortion

However it must be noticed, that sometimes preprocessing can decrease efficiency of the detector. For example blurring could remove edges, that are crucial in many contour based methods. In good lighting conditions histogram equalization could decrease the contrast of the face (Figure 2).

Point based ROI detection can be performed in various ways. Most of the facial features, for example eye corners, mouth corners, nostrils, are placed on the edges. Therefore thresholding the responses of edge filters based for example on Prewitt, Sobel or Roberts operators can significantly reduce the number of analyzed pixels (Figure 3). Further reduction of the number of pixels can be achieved by using corner detectors. There are several methods for accomplishing this task. One of the simplest methods - Moravec corner detector (Moravec (1980)) in based on the assumption, that the sum of absolute differences (SAD) between intensity values of the window anchored in the corner position and windows anchored in

Input image

Sobel Gx   Prewitt Gx   Roberts I

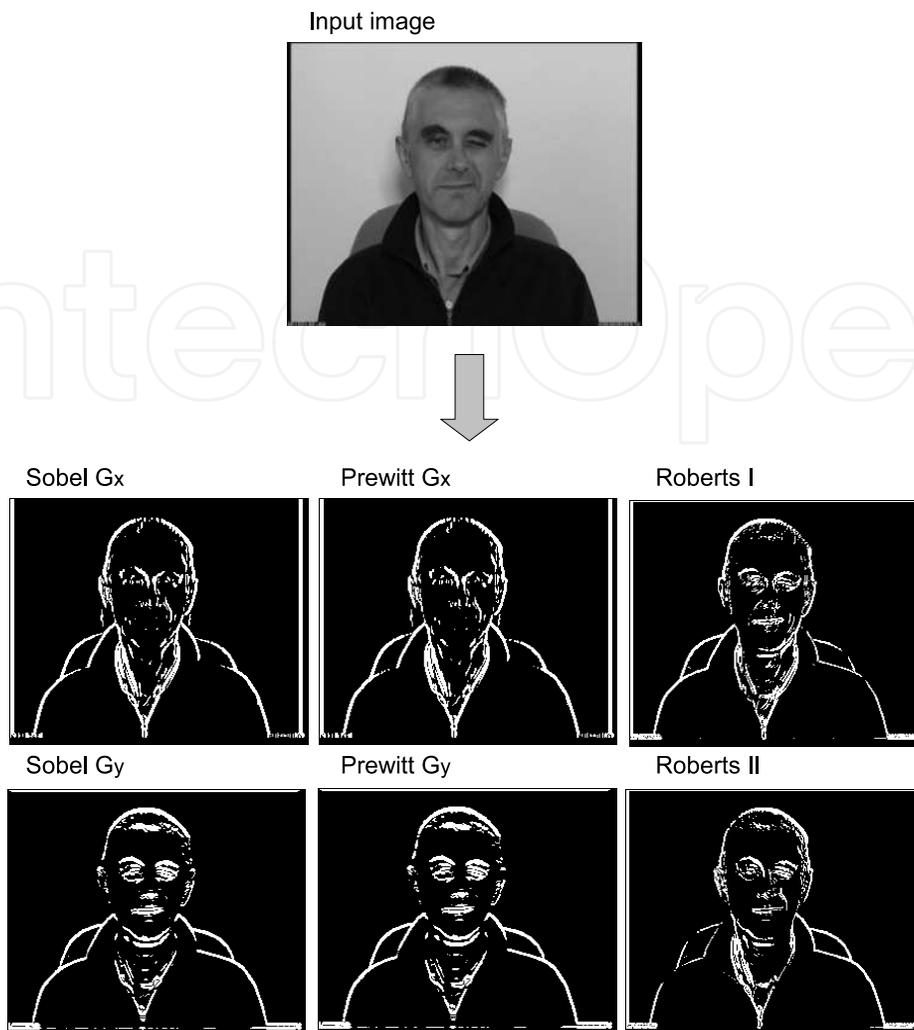Sobel Gy   Prewitt Gy   Roberts II

Fig. 3. Popular binarized edge filters responses for the face image.

the closest neighborhood of the analyzed point are high. Unfortunately algorithm doesn't consider directionality. Particularly the SAD can be low for the regions, that highly differ in the directionality of the edges and thus should be marked as corners. This disadvantage has been removed by the Harris corner detector (Harris & Stephens (1996)). Instead of taking intensities of the pixels directly, algorithm analyses following matrix of partial derivatives:

$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \tag{1}$$

where I_x and I_y defines gradients in x and y axes. Value of the eigenvalues of M define the variation in the edge direction. High value of both eigenvalues define corner.

Another interesting method for detecting corners is Features from Accelerated Segment Test (FAST) (Rosten & Drummond (2006)). Simple algorithm relies on the absolute difference of the analyzed pixel to the neighboring 16 pixels placed on defined circular vicinity. Advantage of such an approach is a very high speed of analysis, while achieving good results in particular applications.
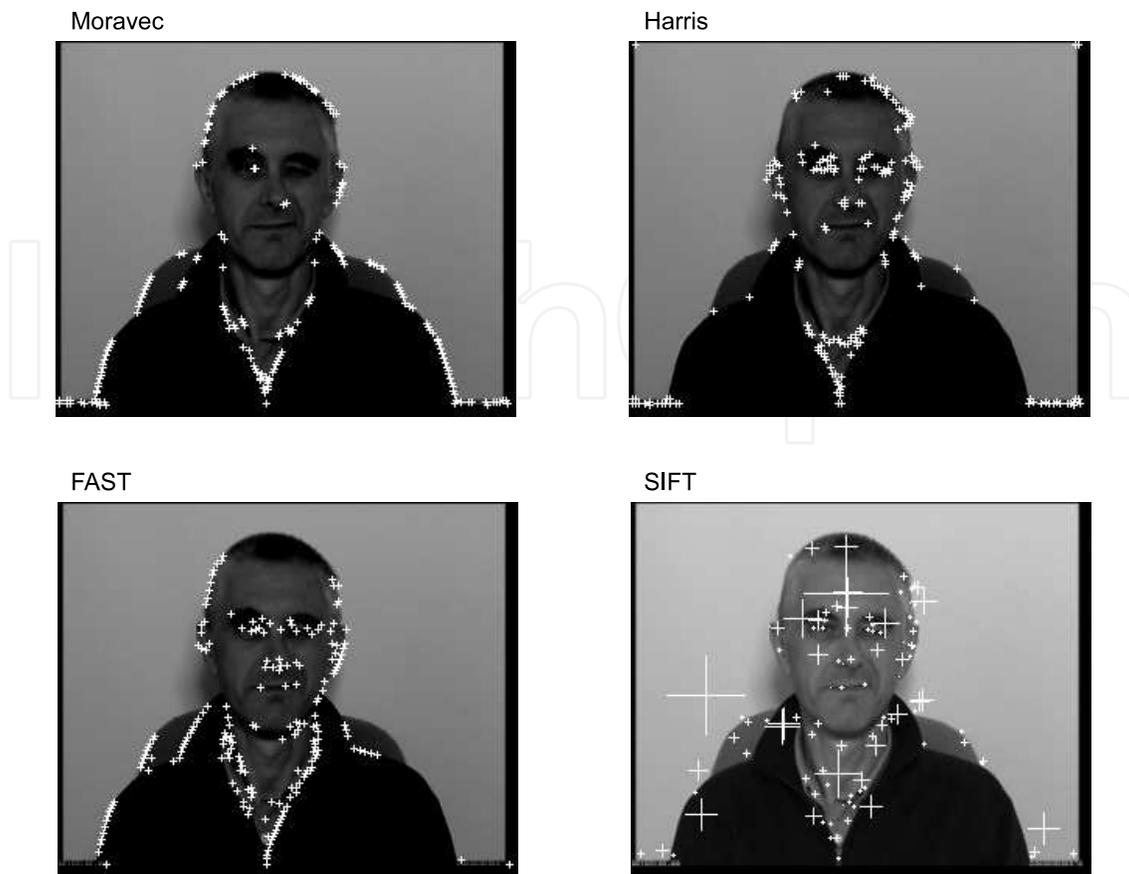
Fig. 4. Results of different corner detectors.

One of the most advanced interest point detection algorithm is Scale-Invariant Feature Transform (SIFT) Lowe (1999). Candidate points (keypoints) are detected using Differences of Gaussians (DoG) thresholding. In the second step non-maximum supression is applied in the 26 elements neighborhood - namely 8 pixels in the vicinity of the pixels and 18 points neighboring analyzed pixel in adjacent scales. In the end the edges are removed from the image by analyzing eigenvalues of the Hessian matrix. This step is explained by the fact, that in the common application of the SIFT algorithm - merging 3D clouds of points, position of the edges may differ at different viewpoints.

The choice of these methods should be adjusted to the particular detected set of points. For example if the goal of the algorithm is to detect corners of the eyes, the SIFT algorithm wouldn't be a good choice, in opposite to the Harris corner detector. Some results of the interest point detection are presented in Figure 4.

If the ROI is specified by the region, there are two most common approaches to this problem. The first one is color segmentation. Because the color of the skin is cumulated in the compact cluster of the RGB color space, skin and non-skin pixels can be distinguished from the image. On the other hand, general skin model, covering all the nationalities and races, is difficult to achieve. Another method for the ROI extraction is face detection. This topic is broadly described (for example in Hjelms & Low (2001); Naruniec (2010)), and thus it won't be analyzed here.
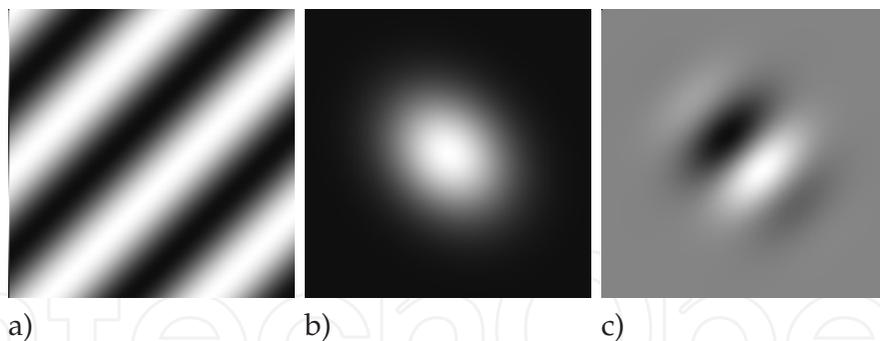
Fig. 5. Example of defining Gabor filter mask. Real part of a) Gabor sinusoidal carrier, b) gaussian envelope, c) resulting filter by multiplication a) and b).
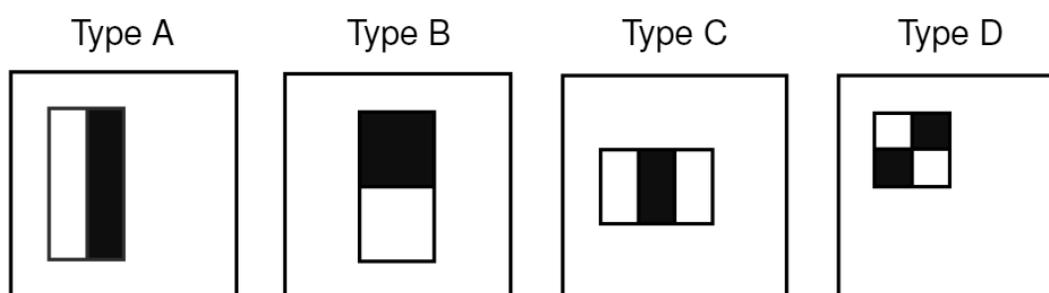

Fig. 6. Regions used in the Viola Jones AdaBoost detector.

2.2 Extraction

After defining initial regions or points of interest, a method for features extraction has to be given. The algorithms differ in the shape of the neighborhood and type of analysis.

One of the most known texture descriptors and facial features descriptor consists of the set of 40 Gabor filter responses (Wiskott et al. (1997)) and this set is called a "jet". Shape of the filter is defined by the two components: sinusoidal carrier and gaussian envelope (Figure 5). Set of functions with 8 orientations and 5 wavelengths form the Gabor jet.

Similarly descriptors formed by the Angular Radial Transform (ART) are computed by convolving the image with created base functions. The transformation is defined in the polar coordinates. Function consists of two components: modulation in angular direction (complex numbers) and sinusoidal function in radial direction (real numbers). In order to gain invariance to rotation, the absolute value of the complex function is taken into further consideration. Usually a set of 33 art coefficients are computed (3 wavelengths in radial direction and 11 in angular direction).

Gabor and ART coefficients are computationaly expensive and therefore inadequate for many real-time applications. Simple alternative to these methods are contrast features used in AdaBoost face detector (Viola & Jones (2001b), Figure 6). Set of such region contrasting filters is used for further AdaBoost classification. Integral image computed to speed up the algorithm provides result of summing any window in the image, in only 4 addition operations.

Another important issue in features extraction is reduction of dimensionality of the data. Simple projection of the data covariance matrix to the eigenvectors in many cases allow to represent vector in more compact form, while preserving most of the signal energy. Result of such principal components analysis (PCA) can be also achieved in a simpler way - by

performing SVD decomposition on the zero-mean data and choosing first left singular vectors corresponding to the largest singular values. Another decorrelation technique used for this task is independent components analysis (ICA) (Duda et al. (2000)). It's goal is usually interpeted in audio processing. Assuming, that a sound is produced from many source signals, ICA provides methods for the blind separation of these inputs.

Facial features extraction can be also performed by fitting actual face to the predefined model. One of the most commonly used method in this scenario is active shape proposed by Cootes et al. (1995). Because of the fact, that the points placed on the facial features are highly correlated, authors apply PCA to define parameters that control the shape of the whole face model. In this way changing one parameter results in the deformation of all points present in the grid. Matching is performed by deforming the model in such way, that it fits the edges present in the image. Extension of this work is called active apperance models (Cootes et al. (1998)). Within this approach, the texture information is given in addition to the shape parameters.

It is also worth mentioning, that extraction of the pixel intensities is often followed by transformations such as discrete cosinus transform (DCT), fast fourier transform (FFT) or the wavelets.

2.3 Classification

Classification of the facial features can be defined in several ways. In the simplest case, classification is based solely on the euclidean distance of the descriptor to the predefined models. Such approach is efficient only for the obvious cases, but in most of the algorithms, more advanced techniques are used.

Another basic classification algorithm is based on the Bayes theory for the conditional probability:

$$P(y_i|x) = \frac{P(x|y_i)P(y_i)}{P(x)} \tag{2}$$

where y_i denotes the class of the analysed object and x is a descriptor. The probability $P(x|y_i)$ is usually modeled in the training step by the mean and the covariance. The value of $P(x)$ is abbreviated at the step of descriptor matching and therefore doesn't have any influence for the classification result.

Better solution for the class separation can be achieved by discrimination methods such as linear discriminant analysis (LDA) Fisher (1936). This method takes in consideration the within class variance R_w and between class variance R_b (see figure 7). Minimalization term is defined as follows:

$$J_{LDA}(w) = \frac{w^t R_b w}{w^t R_w w} \tag{3}$$

Function can be minimized by finding the eigenvalues of the $(R_w)^{-1} R_b$ matrix. Other solutions of this problem are based on SVD decomposition. In this case data are projected firstly on the R_w matrix, and the eigenvectors corresponding to the lowest eigenvalues are taken. In the second step, projected data are maximized in the terms of the R_b matrix by choosing eigenvectors corresponding to the largest eigenvalues.It appears, that for some particular problems dual LDA (DLDA) problem formulation (Leszczynski & Skarbek (2007))
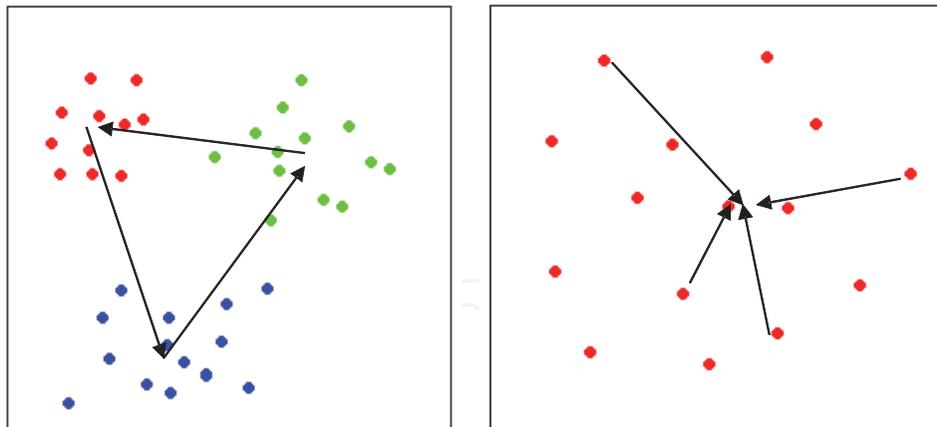
Fig. 7. Graphical illustration of between class (left side) and within class (right side) scatter. Different color correspond to different classes.
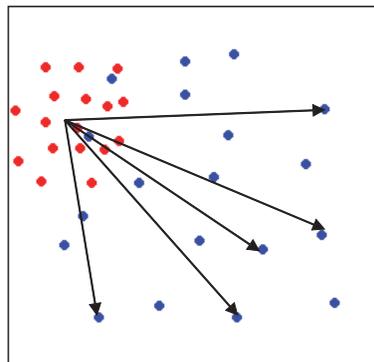

Fig. 8. Graphical illustration of between class variance in modified linear discriminant analysis. Red points correspond to the facial features, while the blue ones - the background.

can give better results. The maximalization function is defined as:

$$J_{DLDA}(w) = \frac{w^t R_w w}{w^t R_b w} \tag{4}$$

The differences in the approaches arise from orthogonality of the vectors minimizing R_w and maximizing R_b. To cope with the problem more generally, SDA analysis can be applied. The initial data is clustered to the moment, in which the angle between two optimization vectors exceeds specified threshold.

In the case of facial features/background two-class classification problem, other solution can give better results. Becouse it is very hard to define mean and variance of the background objects, therefore only facial feature within-class variance can be optimized. Moreover becouse the background class is very differentiated, treating every background example a separate class yields to better results. This assumptions are formulated in the modified linear discriminant analysis (MLDA) - (Naruniec & Skarbek (2007), see figure 8).

Another technique yelding very good separation results is support vector machine (SVM). Algorithm tends to separate two classes by providing two parallel linear hyperspaces leaning on some vectors of the data (support vectors) while retaining large margin between these hyperspaces (figure 9). Extension of this method introduces the error measure, that allow
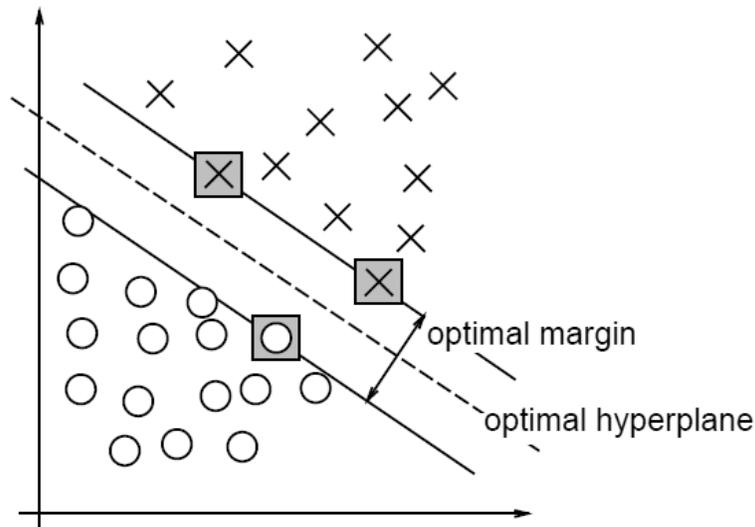
Fig. 9. SVM in two dimensional space. Crosses and circles are the two separated classes. Two lines form hyperplane diving these two distributions. (Cortes & Vapnik (1995)).

some vectors to be misclassified. Also kernel methods performing non-linear classification have been introduced. More informations about this subject can be found in the work of Cortes & Vapnik (1995).

AdaBoost is a method for combining many "weak classifiers" - having poor accuracy results to the very efficient "strong classifiers" Freund & Schapire (1995). In every iteration of the algorithm classifier with the lowest error according to the actual training examples weights, is added to the final classifier. Classification error is defined as:

$$\epsilon(\omega, \theta) = \frac{1}{2} \sum_{i=1}^{L} w_i |\delta_w(o_i) - v_i| \qquad (5)$$

where ω denotes weak classifier, θ is the detection threshold, L - number of the training examples, w_i is the weight of the i-th training example and v_i denotes the label of i-th example ("1" for the facial feature, "-1" for the background).

At every iteration weights are updated using following formula:

$$w_{i,t+1} = \frac{w_{i,t}e^{-\gamma_i(o_i)v_i}}{\sum_{i-1}^{L} w_{i,t}e^{-\gamma_i(o_i)v_i}} \qquad (6)$$

Costs of the positive or negative decision γ are computed using algorithm heuristics.

AdaBoost method have prooven to give fast and accurate results in the face detection (Viola & Jones (2001a;b)) and region based facial features detection scheme (Goldmann et al. (2006)).

2.4 Postprocessing

After classifying regions or points to the specific class, validation and refining of the selected facial features can be applied.
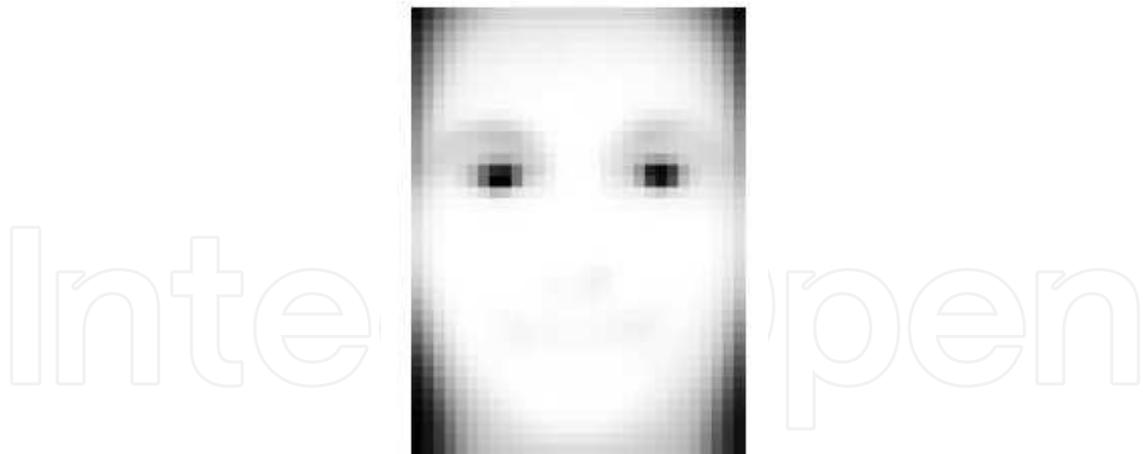
Fig. 10. Mixture coefficients of the model for face image patches. Bright pixels indicate a high probability for skin, dark pixels indicate a low probability for skin (Hoffmann et al. (2009)).

First remark concerns merging close results. Some detection algorithms may provide many results of the single facial features. In order to combine close fiducial points, simple clustering can be applied. In the case of regions, overlapping windows are merged in order to get single response.

Simplest approach for facial features validation is a geometrical matching. Relations between eyes, nose or mouth can be defined manually by the knowledge, or automatically by analysing specified data set.

Graph based methods convert each of the classified point to the graph node with assigned value, computed for example by the confidence of classification. In the next step possible graph values are compared to the trained mean face model. All the points that fits the model are marked as true faces, while the rest of the points are eliminated as false acceptances.

Validation can be also applied using color information, for example by convolving the face image with the color face patch (see Figure 10 - Hoffmann et al. (2009)).

Accuracy refinement is usually performed by moving facial features contours to the edges or by fitting to the specified model (for example created by the active shapes).

3. Conclusion

In this chapter a methodology for the facial features detection has been given. It describes all the basic semantic analyse stages - preprocessing, defining regions of interest, features extraction, classification and postprocessing.

4. References

Ben Jemaa, Y. & Khanfir, S. (2009). Automatic local Gabor Features extraction for face recognition, *ArXiv e-prints* .

Celiktutan, O., Akakin, H. & Sankur, B. (2008). Multi-attribute robust facial feature localization, *Automatic Face & Gesture Recognition, 2008. FG '08. 8th IEEE International Conference on*, pp. 1–6.

Cootes, T. F., Edwards, G. J. & J., T. C. (1998). Active appearance models, *Lecture Notes in Computer Science* 1407.

Cootes, T. F., Taylor, C. J., Cooper, D. H. & Graham, J. (1995). Active shape models-their training and application, *Computer Vision and Image Understanding* 61(1): 38–59.

Cortes, C. & Vapnik, V. (1995). Support-vector networks, *Machine Learning* 20: 273–297. 10.1023/A:1022627411411.
URL: *http://dx.doi.org/10.1023/A:1022627411411*

Duda, R. O., Hart, P. E. & Stork, D. G. (2000). *Pattern Classification (2nd Edition)*, 2 edn, Wiley-Interscience.

Erukhimov, V. & Lee, K. (2008). A bottom-up framework for robust facial feature detection, *FG*, pp. 1–6.

Fisher, R. (1936). The use of multiple measurements in taxonomic problems, *Annals of Eugenics* 7: 179–188.

Freund, Y. & Schapire, R. E. (1995). A decision-theoretic generalization of on-line learning and an application to boosting, *European Conference on Computational Learning Theory*, pp. 23–37.

Goldmann, L., Monich, h. U. & Sikora, T. (2006). Robust face detection based on components and their topology, Vol. 6077, SPIE, p. 60771V.

Harris, C. & Stephens, M. (1996). A combined corner and edge detector, pp. 147–151.

Hjelms, E. & Low, B. K. (2001). Face detection: A survey, *Computer Vision and Image Understanding* 83(3): 236 – 274.
URL: *http://www.sciencedirect.com/science/article/B6WCX-458P9XF-3/2/25c703bcdc7e964 39e46210c9c9fffc2*

Hoffmann, U., Naruniec, J., Yazdani, A. & Ebrahimi, T. (2009). Face Detection Using Discrete Gabor Jets and a Probabilistic Model of Colored Image Patches, *in* J. Filipe & M. S. Obaidat (ed.), *e-Business and Telecommunications, Communications in Computer and Information Science, Volume 48. ISBN 978-3-642-05196-8. Springer-Verlag Berlin Heidelberg, 2009, p. 331.*

Leszczynski, M. & Skarbek, W. (2007). Biometric verification by projections in error subspaces, *RSKT*, pp. 166–173.

Lowe, D. (1999). Object recognition from local scale-invariant features, pp. 1150–1157.

Moravec, H. (1980). Obstacle avoidance and navigation in the real world by a seeing robot rover, *tech. report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University doctoral dissertation, Stanford University*, number CMU-RI-TR-80-03.

Naruniec, J. (2010). A survey on facial features detection, *International Journal of Electronics and Telecommunications* .

Naruniec, J. & Skarbek, W. (2007). Face detection by discrete gabor jets and reference graph of fiducial points, *Rough Sets and Knowledge Technology*, Springer Berlin / Heidelberg, pp. 187–194.

Rosten, E. & Drummond, T. (2006). Machine learning for high-speed corner detection, *In European Conference on Computer Vision*, pp. 430–443.

Viola, P. & Jones, M. (2001a). Fast and robust classification using asymmetric adaboost and a detector cascade, *Advances in Neural Information Processing System 14*, MIT Press, pp. 1311–1318.

Viola, P. & Jones, M. (2001b). Robust real-time object detection, *International Journal of Computer Vision* .

Vukadinovic, D. & Pantic, M. (2005). Fully automatic facial feature point detection using gabor feature based boosted classifiers, *Proceedings of IEEE ICSMS* .

Wiskott, L., Fellous, J., Krï£¡ger, N. & von der Malsburg, C. (1997). Face recognition by elastic bunch graph matching, *IEEE Trans. Pattern Analysis and Machine Intelligence* 19: 775–779.
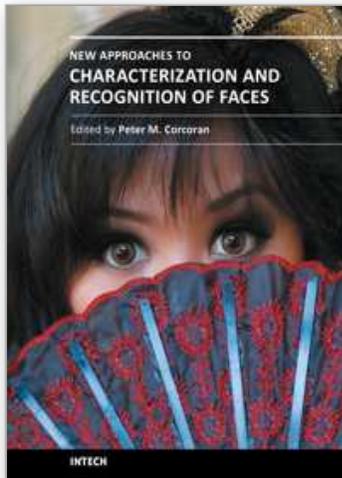
New Approaches to Characterization and Recognition of Faces

Edited by Dr. Peter Corcoran

As a baby, one of our earliest stimuli is that of human faces. We rapidly learn to identify, characterize and eventually distinguish those who are near and dear to us. We accept face recognition later as an everyday ability. We realize the complexity of the underlying problem only when we attempt to duplicate this skill in a computer vision system. This book is arranged around a number of clustered themes covering different aspects of face recognition. The first section presents an architecture for face recognition based on Hidden Markov Models; it is followed by an article on coding methods. The next section is devoted to 3D methods of face recognition and is followed by a section covering various aspects and techniques in video. Next short section is devoted to the characterization and detection of features in faces. Finally, you can find an article on the human perception of faces and how different neurological or psychological disorders can affect this.

INTECH

open science | open minds