Chapter 2
Face Recognition in Unconstrained Environment

Przemysław Kocjan and Khalid Saeed

Abstract This chapter addresses the problem of face recognition from images with lighting problems such as shadows or brightness level. Authors describe face recognition processing, including major components such as face detection, tracking, alignment, and feature extraction. Technical challenges of building a face recognition system are pointed out. The chapter emphasizes the importance of subspace analysis and learning, providing not only an understanding of the challenges therein but also the most successful solutions developed to date. In the following sections, authors present brief history of face recognition systems, show problems that affect results of these systems, and present their own approach based on finding fiducial points in face image and their further use for face recognition.

2.1 Introduction

Face recognition is a task that humans perform routinely and effortlessly in their daily lives. Our brains perform this task remarkably easily and accurately, although this apparent simplicity is dangerously misleading. The automatic face recognition is a problem that is still far from being solved. Wide availability of powerful and low-cost desktop and embedded computing systems has created an enormous interest in automatic processing of digital images and videos in a number of applications, including biometric authentication, surveillance, human-computer interaction, and multimedia management. In spite of more than 20 years of

P. Kocjan (✉)
AGH, Kraków, Poland
e-mail: przemyslaw.kocjan@gmail.com

K. Saeed
AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Al. Mickiewicza 30, Kraków PL-30059, Poland
e-mail: saeed@agh.edu.pl

K. Saeed and T. Nagashima (eds.), *Biometrics and Kansei Engineering*,
DOI 10.1007/978-1-4614-5608-7_2,

extensive research and large number of papers published in journals and conferences dedicated to this area, we still cannot claim that artificial systems are comparable to human performance.

Automatic face recognition is intricate primarily because of differences in conditions like lighting and viewpoint changes induced by body movement during image acquisition. Aging, facial expressions, occlusions etc., also make the problem more difficult. Researchers from the areas of computer vision, image analysis and processing, pattern recognition, machine learning, and many other are working cooperatively, motivated not only by the fundamental challenges this recognition problem generates but also by numerous practical applications in which human identification is needed. The interest of scientists is also increased by the fact that with the rising public concern for security, the need for identity verification such as face recognition is more apparent. Also, advances in technology, such as in digital cameras and mobile devices, made face recognition more important and easier to approach.

Face recognition has an important advantage over many other biometric technologies – it is a nonintrusive, noninvasive, and easy-to-use method. Because of this, it became one of three identification methods used in e-passports and the biometric of choice for many other security applications. Hietmeyer [1] considered six biometric attributes. From all of them, face scored the highest compatibility in a Machine Readable Travel Documents (MRTD) [2].

A face recognition system is expected to identify faces in images and videos automatically. This system may operate in either or both of two modes: face verification (authentication) and face identification (recognition). Face verification compares a query face image of supposedly known person against this person's template face image stored by the system (one-to-one match). Face identification compares a query face image of unknown identity against all the template images in the database to determine the identity of the query face (one-to-many matches). Another face recognition scenario involves a watch-list check, where a query face is matched to a list of suspects (one-to-few matches).

The performance of face recognition systems has improved significantly since the first automatic face recognition system developed by Kanade [3]. Currently face detection, facial feature extraction, and recognition can be performed in real time for images captured under constrained conditions.

2.2 Early Approaches to the Face Recognition Problem

The need of face recognition rises from the moment when machines become more "intelligent" and powerful and gained the ability to improve, supplement, or substitute human abilities and senses.

The subject of face recognition is as old as computer vision because of not only the practical importance of the topic but also the theoretical interest from cognitive scientists. Clearly, using a face to recognize people is not the only method of differentiation between people. Humans also use different senses (e.g., hearing) in order to recognize each other. Machines may utilize a wider range of recognition techniques using, for example, fingerprint images or iris scans. Despite the fact that other methods of identification can be more accurate, face recognition, because of its noninvasive nature and because it is human's primary method of person identification, remains a major area of research.

Since the beginning of the research in that field of technology, there were two main approaches to face recognition: feature based (geometrical) and appearance based (pictorial).

The geometrical approach uses the spatial configuration of facial features. It means that the main geometrical features of the face such as the positions of eyes, nose, and mouth are first located and then faces are classified on the basis of various geometrical distances and angles between these features. On the other hand, the pictorial approach uses templates of the facial features. That method is using the templates of the major facial features and entire face to perform recognition on the frontal views of the faces. Many of the projects based on those two approaches have some similar common extensions that handle different poses and backgrounds.

Apart from these two techniques, there are other more recent template-based approaches. In one of the methods, templates are formed from the image gradient. The other one is principal component analysis approach, which can be interpreted as a suboptimal template approach. Finally there is the deformable template approach that combines elements of both the pictorial and feature geometry approaches and can be applied to faces at varying poses and expressions.

Perhaps the most famous early example of a face recognition system is the one developed by Kohonen in 1989 [4]. It was demonstrated there that a simple neural network could perform face recognition for aligned and normalized face images. The type of network he employed computed face description by approximating the eigenvectors of the face image autocorrelation matrix. These eigenvectors are now known as eigenfaces.

Kohonen's system was, however, not a practical success, because of the need for precise alignment and normalization. In the following years many researchers tried face recognition schemes based on edges, inter-feature distances, and other neural network approaches. While several methods were successful on small databases of aligned images, none successfully addressed the more realistic problem, where database is large and the location and scale of the face are unknown.

A year later, Kirby and Sirovich [5] introduced an algebraic manipulation technique which made it easy to directly calculate the eigenfaces and showed that less than 100 of them were required to accurately describe carefully aligned and normalized face images. Turk and Pentland [6] demonstrated in 1991 that the residual error when coding using the eigenfaces can be used both to detect faces in cluttered natural imagery and to determine the precise location and scale of faces in an image. They then proved that by coupling this method for detecting and localizing faces with the eigenface recognition method, one could achieve reliable real-time recognition of faces in a minimally constrained environment. This demonstration that simple, real-time pattern recognition techniques could be combined to create a useful system sparked an explosion of interest in the topic of face recognition.

With the rapid evolution of the technology and the commercialization of technological achievements, face recognition became more and more popular not only as research subject but also for the use in security systems.

This fact gave the motive to many researchers and also companies to develop techniques for automatic recognition of faces. These products have many applications, also in security and human-computer interaction. For instance, a face-recognizing machine could allow automated access control for buildings or enable a computer to recognize the person using it at the moment. Most existing face recognition systems, however, can recognize only frontal or nearly frontal images of faces. By recognizing faces under varying pose, one would make the conditions under which face recognition systems operate less rigid.

2.3 Face Recognition in a Changing Environment

The progress in face recognition has been promising over the years. However, the same task for unconstrained environments – where we have to take into account changes of viewpoint, illumination, expression, occlusion, accessories, and so on – is still far from being solved.

2.3.1 Problem

Subspace analysis techniques for face recognition are based on the fact that a class of patterns of interest, such as the face, resides in a subspace of the input image space. If we consider a small grayscale image of size 64×64 which has 4,096 pixels, this picture can express a large number of pattern classes, such as trees, houses, and faces. However, among the 256^{4096}, which is more than 10^{9864} possible configurations of pixels, only a few correspond to faces. Because of this, the original image representation is highly redundant. The dimensionality of this representation could be greatly reduced when only the face patterns are of interest.

With the eigenface or principal component analysis (PCA) [6, 7] approach, a small number of eigenfaces [8] are derived from a set of training face images by using the Karhunen-Loeve transform or PCA. These modeling techniques allowed to efficiently represent face image as a feature vector of low dimensionality. The features in such subspace provide more valuable and richer information for recognition than the raw image.

The manifold or distribution of all faces accounts for variation in face appearance, whereas the nonface manifold accounts for everything else. Closer look into manifolds in the image space shows that they are highly nonlinear and nonconvex [9, 10]. Figure 2.1a illustrates face versus nonface manifolds and (b) illustrates the manifolds of two individuals contained in one face manifold. Distinguishing between the face and nonface manifolds in the image space is the task of face

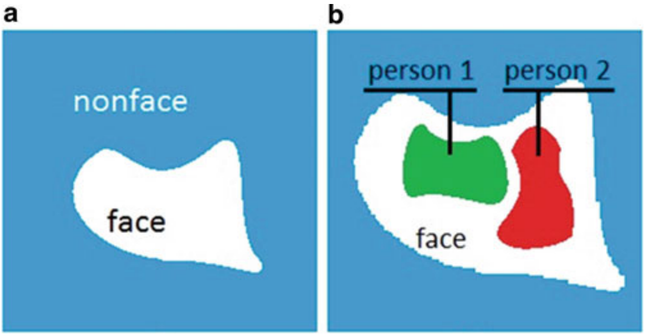

Fig. 2.1 (**a**) Face versus nonface manifolds, (**b**) face of different individuals

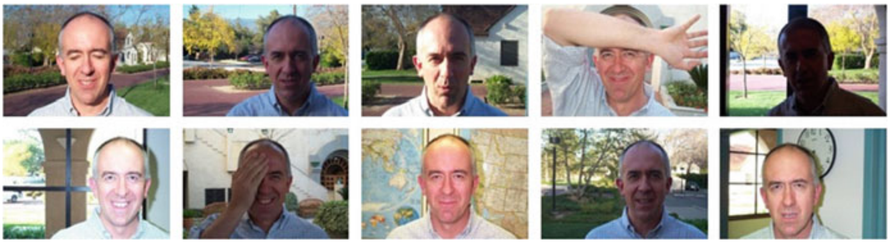

Fig. 2.2 Sample intrasubject variations in pose, illumination, expression, occlusion, and brightness [16]

detection. Face recognition, however, is concerned with distinguishing multiple individuals in the single face manifold.

As shown in Fig. 2.1, the classification problem associated with face detection is highly nonlinear and nonconvex, even more so for face matching. Face recognition evaluation reports (e.g., [11, 12]) and other independent studies indicate that the performance of many state-of-the-art face recognition methods deteriorates with changes in lighting, pose, and other factors [13–15].

Whereas shape and reflectance are intrinsic properties of a face object, the appearance (i.e., the texture look) of a face is also subject to several other factors, including the facial pose (or, equivalently, camera viewpoint), illumination, and facial expression. Figure 2.2 shows an example of significant intrasubject variations caused by some of these factors.

To complicate the problem, we could add various imaging parameters, such as aperture, exposure time, lens aberrations, and sensor spectral response which also increases intrasubject variations. Face-based person identification is even more difficult with possible small intersubject variations (Fig. 2.3). All these factors are present in the image data, so "the variations between the images of the same face due to illumination and viewing direction are almost always larger than the image variation due to change in face identity" [17]. This variability makes it

Fig. 2.3 Similarity of frontal faces between son and father (a) [81, 82], twins (b) [81]

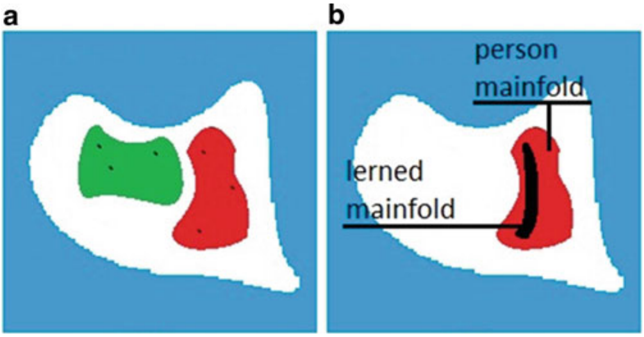

Fig. 2.4 Challenges in face recognition from subspace viewpoints. (a) Euclidean distance is unable to differentiate between individuals. (b) The learned manifold of classifier is unable to characterize unseen images of the same individual face

difficult to extract the consistent intrinsic information of the face objects from their respective images.

As illustrated above, the entire face manifold is highly nonconvex, and so is the face manifold of any individual under various circumstances. Linear methods such as PCA [6, 18], independent component analysis (ICA) [19], and linear discriminant analysis (LDA) [20] project the data linearly from a high-dimensional space (e.g., the image space) to a low-dimensional subspace. Because of this fact, they are unable to preserve the nonconvex variations of face manifolds necessary to differentiate among individuals. In a linear subspace, Euclidean and Mahalanobis distances are normally used for template matching. Unfortunately they do not perform well for distinction between face and nonface manifolds and between manifolds of individuals (Fig. 2.4a). This crucial fact limits the possibility of linear methods to achieve highly accurate face detection and recognition.

Another problem is the ability to generalize, illustrated by Fig. 2.4b. A canonical face image of 112×92 resides in a 10,304-dimensional feature space. The number of examples per person available for learning the manifold is usually much smaller than the dimensionality of the image space. A system trained on so few examples may not generalize well to unseen instances of the face.

2.3.2 Proposed Solutions

To deal with difficulties mentioned above, it is possible to choose one out of two strategies. First method is to construct a "good" feature space in which the face manifolds become simpler, that is, less nonlinear and nonconvex than those in the other spaces. This method needs two things to be done. First thing is the normalization of face images geometrically and photometrically, such as using morphing and histogram equalization. Second thing is the extraction of features from the normalized images which are stable with respect to such variations as ones based on Gabor wavelets.

The second strategy is to construct classification engines able to solve difficult nonlinear classification and regression problems in the feature space and to generalize better. Since the first option reduces the nonlinearity and nonconvexity, it does not solve the problems completely, and classification engines able to deal with such difficulties are still necessary to achieve high performance. A successful algorithm should combine both strategies.

When thinking about first strategy, it is necessary to mention the geometric feature-based approach used in the early days of face recognition [21–24], where facial features such as eyes, nose, mouth, and chin are detected. Properties of and relations between these features are used as descriptors for face recognition. Advantages of this approach include efficiency when achieving data reduction and insensitivity to variations in illumination and viewpoint. However, facial feature detection and measurement techniques developed to date are not reliable enough, as mentioned before, for the geometric feature-based recognition [25]. Such geometric properties alone are inadequate for face recognition because the rich information contained in the facial texture or appearance is discarded. These are reasons why early techniques are not effective. The statistical learning approach learns from training data, like appearance images or features extracted from appearance, to extract good features and construct classification engines. During the learning, both prior knowledge about faces and variations seen in the training data are taken into consideration.

The appearance-based approach, such as PCA- [6] and LDA [20]-based methods, incorporates more advanced face recognition techniques. Such an approach generally operates directly on an image-based representation. It extracts features in a subspace derived from training images. Using PCA, a face subspace is constructed to represent optimally only the face object. Using LDA, a discriminant subspace is constructed to optimally distinguish faces of different persons. Comparative reports (e.g., [20]) show that LDA-based methods generally yield better results than PCA-based ones.

Although these linear, holistic, appearance-based methods avoid instability of the early geometric feature-based algorithms, they are not accurate enough to describe subtleties of original manifolds in the original image space. This is due to their limitations in handling nonlinearity in face recognition. Fortunately such linear methods can be extended using nonlinear kernel techniques (kernel PCA [26] and kernel LDA [27]) to deal with this nonlinearity [28–31]. In these approaches, a nonlinear projection (dimension reduction) from the image space to a feature space is performed. The manifolds in the resulting feature space become simple, yet with subtleties preserved. The kernel methods may achieve good performance on the training data; however, it may not be so for unseen data because of their much greater flexibility than the linear methods and overfitting thereof.

Another approach to handle the nonlinearity is to construct a local appearance-based feature space using appropriate image filters, so the distributions of faces are less affected by various changes. Local feature analysis (LFA) [32], Gabor wavelet-based features (such as elastic graph bunch matching, EGBM) [33–35], and local binary pattern (LBP) [36] have been used for this purpose.

Some of these algorithms may be considered as combining geometric (or structural) feature detection and local appearance feature extraction to increase stability of recognition performance when viewpoint, illumination, or expression changes. Face recognition algorithms can be divided based on pose dependency into pose dependent and pose invariant. One can distinguish three types of pose-dependent algorithms (viewer-centered images). The first are feature based (analytic) which detect a set of geometrical features on the face such as nose, mouth, chin, and eyes. The second are appearance based (holistic) such as PCA and LCA, and the third are hybrid, such as LFA or EGBM, which are the combination of the two previous. In the pose-invariant algorithms, 3D face models are utilized to reduce the variations in pose and illumination. Gordon and Lewis [37] proposed an identification system based on 3D face recognition. The 3D model used by them is represented by a number of 3D points associated with their corresponding texture features. This method requires an accurate estimate of the face pose. Lengagne et al. [38] proposed a 3D face reconstruction scheme using a pair of stereo images for recognition and modeling. However, they did not implement the recognition module. Atick et al. [39] proposed a reconstruction method of 3D face surfaces based on the Karhonen-Loeve (KL) transform and the shape-from-shading approach. They also discussed the possibility of using eigenhead surface face recognition applications. Yan and Zhang [40] proposed a 3D reconstruction method to improve the performance of face recognition by making reconstruction method, introduced by Atick et al., rotation invariant. Zhao and Chellappa [41] proposed a method to adapt a 3D model from a generic range map to the shape obtained from shading for enhancing face recognition performance in different lighting and viewing conditions.

A large number of local features can be produced with varying parameters in the position, scale, and orientation of the filters. For example, more than 100,000 local appearance features can be produced when an image of 100×100 is filtered with Gabor filters of five scales and eight orientations for all pixel positions, causing increased dimensionality. Some of these features are effective and important for the classification task, whereas the others may not be so. AdaBoost methods have been used successfully to tackle the feature selection and nonlinear classification problems [42–44].

Researchers put much effort in creating recognition system resistant to shadows, low brightness level, or flashes. Some of these problems can be dealt with using mathematical tools. One of the difficulties that researchers still deal with is images with unusual lighting. In the situation when the intensity variability in input images changes, we have to deal with different types of variability in the area of the face and its background. Local shadows change the form of individual parts of the face, that is, nose, mouth, and eyes, and distort the boundaries of the face area. Global shadows significantly reduce efficiency of discrimination of various face areas against general background or completely hide them.

The results of face recognition systems are very sensitive to variability in the face area. The analysis of the recent literature devoted to face recognition from images with lighting problems [45–53] leads to the observation that there exist some methods to solve this problem, such as:

– Processing of images in order to equalize brightness variation (intensity equalization)
– Reduction of intensity gradient (gamma correction, logarithmic transformation), invariant (in respect to intensity) image representation using the LBP (local binary patterns) and LTV (logarithmic total variation)
– Representation of face images with lighting problems using the eigenbase decompositions and corresponding models based on eigenfaces
– Representation of face images with lighting problems with spectral features using the wavelets and cosine transformation with elimination of low-frequency components
– Extension of face recognition system database with new patterns having all distortions related to lighting problem of face images

The eigenbasis approaches are usually obtained using PCA and LDA [45–47] and more recently also using CCA (Canonical Correlation Analysis [50]), at the same time using DCT (Discrete Cosine Transform) as face images with lighting problems preprocessing step.

For example, in [45] it was shown that in order to solve recognition task using the PCA and LDA, face images should be transformed into spectral features using 2DDCT. At the same processing step, the low-frequency spectral components are removed, as corresponding to “shadow” components.

2.3.3 Authors’ Approach

Authors of this chapter also tried to solve the aforementioned problem. We tried to overcome the difficulties caused by shadows using Toeplitz matrices [54] where different ways of calculating coefficients in matrices are presented. Attempt to use this approach in face recognition [55] was performed with 25 points marked on each face form database.

Also different types of classifiers were used to determine usefulness of proposed matrices. Although the points could be determined manually, there are many algorithms which could perform this task automatically, that is, ASM or AAM. Authors propose a way of determining some of fiducial face points by performing a

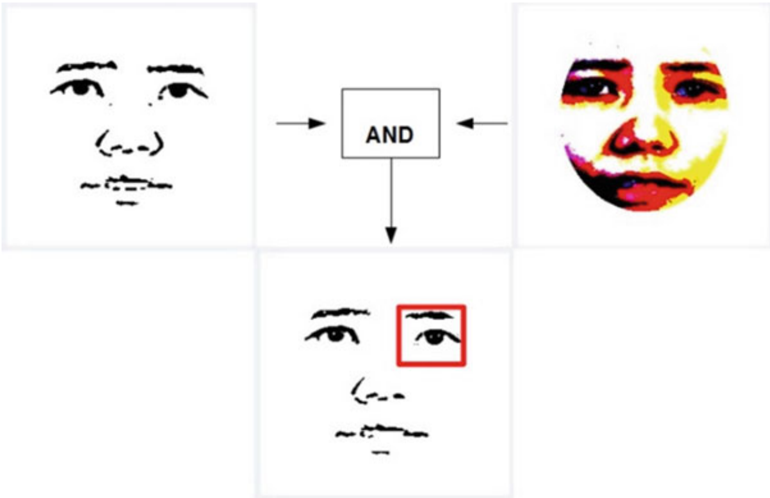

Fig. 2.5 Part of the process of determining the second coordinate of the eye [56]

couple of morphological operations like adaptive threshold and binarization in each of RGB channel (Fig. 2.5). Research shows that Toeplitz matrices perform successfully on small databases, although increasing the size of the database causes the results to deteriorate.

Figure 2.6 represents the scheme of the authors' algorithm. A process of selection and extraction of characteristic points is performed step by step. After successful feature point localization, feature points are used in Toeplitz matrix. The steps of the algorithm are easy to perform and implement.

The algorithm works with color images. After loading image, the first step that has to be done is correct face localization.

To detect faces in images, P. Viola and M. Jones algorithm is used. Characteristic feature of this approach is high efficiency. Speed of the algorithm is 15 fps on Pentium III 700 MHz with resolution 384 $\times$ 288 pixels. The algorithm works with grayscale images, so color images have to be transformed. It is estimated that rate of correct face detection is 93.7%.

After conversion to grayscale, an image is searched for faces. When it is localized, it is being cut off from image and rescaled to size of 240 $\times$ 240 pixels.

An image is transformed into the HSV color space to perform skin color classification. Acronym HSV stands for hue, saturation, and value. Hue has values from 0° to 360°, saturation from 0 to 1, and value from 0 to 1.

As a result we obtain black and white image with some noise that can be easily removed using median filter. On computer equipped with Intel Pentium Dual CPU T3400, processing time from beginning of the program to obtaining black and white image took less than 200 ms. Medium values from test showed 192 ms. Result can be further reduced with optimization. Final image with detected skin is presented in Fig. 2.7b.

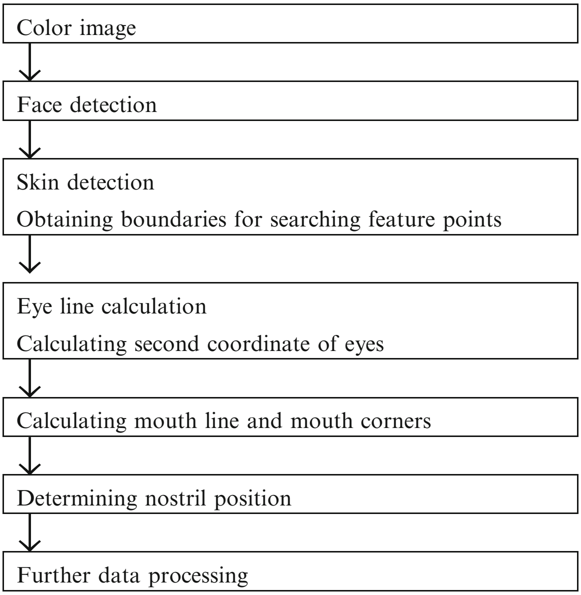

Fig. 2.6 The main stages of the algorithm

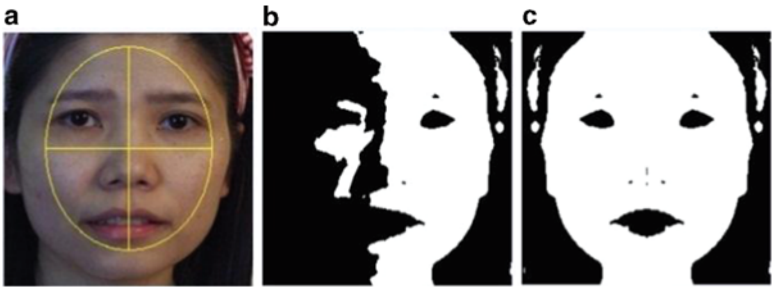

Fig. 2.7 (**a**) Input image with boundaries, (**b**) the result of skin detection, (**c**) mirror image of the detected skin

Skin detection does not always give a correct result. It may be caused by not equally illuminated face or skin color differences caused by makeup or illness. As can be seen in Fig. 2.7b, the skin was not detected on the whole face. Using the symmetry of the face, we can approximate that the detected skin is more or less equally distributed on both sides.

Creating mirror image allows to set boundaries more correctly. Calculating boundaries that are visible in Fig. 2.7a is relatively simple. To obtain them, we are using the idea of image moments. We need to calculate the zero moment (M_{00})

and the first moments for y (M_{10}) and x (M_{01}) from binary mask in Fig. 2.7c. Then using equations below, it is possible to set a center of the object in image:

$$Yc = \frac{M_{10}}{M_{00}}, \tag{2.1}$$

$$Xc = \frac{M_{01}}{M_{00}}, \tag{2.2}$$

Calculation of the second moments (M_{11}, M_{02}, M_{20}) and (2.3), (2.4), allowing for creation of the ellipse

$$L = 1,\, 5\sqrt{\left[(a+c) + \sqrt{\left(bb + (a+c)^2\right)}\right]/2}, \tag{2.3}$$

$$W = 1,\, 5\sqrt{\left[(a+c) - \sqrt{\left(bb + (a+c)^2\right)}\right]/2}, \tag{2.4}$$

Parameters a, b, and c are:

$$a = \left(\frac{M_{02}}{M_{00}}\right) - XcXc, \tag{2.5}$$

$$b = \left(\frac{M_{11}}{M_{00}}\right) - XcXc, \tag{2.6}$$

$$c = \left(\frac{M_{20}}{M_{00}}\right) - YcYc, \tag{2.7}$$

When boundaries were successfully set, we are now able to search for the first coordinate of the eye. In order to do so, we are using Sobel filter in two dimensions OX and OY.

The face is divided in two parts, left and right. For each side of the face, we are creating a projection. Obtained function *H(y)* (see Fig. 2.8a) is multiplied by function *W(y)* (see Fig. 2.8b) in order to minimize the influence of the mouth and nostrils.

After multiplication of the *H(y)* by *W(y)*, we obtain the function shown in Fig. 2.8c. Maximum value of this function denotes the eye line and the first coordinate of the eye.

The second coordinate is calculated by performing adaptive binarization. As a result, we obtain a binary image with the eyes, nostrils, and mouth remaining on it. Additionally, the threshold is performed in the color image (RGB) on every channel. Then logical operation AND is executed for every pixel in the image.

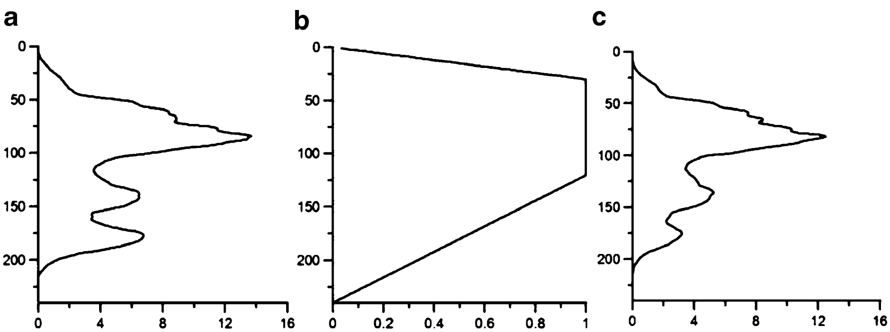

Fig. 2.8 (**a**) Projection *H(y)*, (**b**) function *W(y)*, (**c**) function *F(y)*

When the pixel value is "black" in both images, we obtain a "black" pixel. Otherwise, the pixel is white. The result of this operation is shown in Fig. 2.5.

Vertical projection for eyes is performed in the last step after thresholding and the AND operation. The maximum value denotes the second coordinate of the eye position.

The mouth line is calculated in a similar way to the eye line. The projection is multiplied by the function *W(y)*. This function was modified in a way that minimizes the influence of the eyes.

The position of the mouth corners and width of the jaw and the nostrils are also acquired using projection approach. The maximum value of the projection result denotes the coordinates of features we are looking for.

The obtained feature points will be used for the image description. These points are used to build Toeplitz matrices and calculate their minimal eigenvalues [54].

This simple algorithm is able to work under different lighting conditions (part of the shadows is being removed), and it is robust to elements of environment because feature points are not being searched for outside face area. Skin detection process works also in case of people with dark skin carnation.

2.4 Examples of Face Recognition Algorithms

A number of face recognition algorithms are based on feature-based methods that detect a set of geometrical features on the face such as the positions of the eyes, eyebrows, nose, and mouth [57]. Geometric properties and relations such as areas, distances, and angles between the features are selected as the descriptors of faces. Therefore, the geometric attributes provide benefits in data reduction and make the algorithm less sensitive to variations in illumination, viewpoint, and expressions. Typically, 30–40 feature points per face are generated. The performance of face recognition based on geometrical features depends on the accuracy of the feature location algorithm. However, there are no perfect answers to the problem of how many feature points shall be acquired for the best performance, what the important features are, or how to extract them automatically. Face recognition based on geometrical feature matching is possible for face images at very low resolution. However, precision will suffer, and on large databases this method does not satisfy accuracy demands.

Ivancevic et al. [58] stated that there are about 80 landmark points on a human face and the number of points chosen is application dependent. However, some authors used more than 80 facial points in their algorithms. One example of such work is Cootes et al. [59] where 122 landmark points are used. On the other hand many authors base their algorithms on much smaller number of points – for example, Huang and Huang [60] used 90 facial feature points, Kobayashi and Hara [61] used 30 facial characteristic points, Pantic and Rothkrantz [62] used 19 facial fiducial points, Valstar and Pantic [63] used 20 facial fiducial points, Cohn et al. [64] used 46 fiducial points, and Zhang et al. [65] used 34 fiducial points.

Also a lot of effort was put in the works describing feature point tracking [64, 66], action unit recognition for facial expression analysis [62, 63, 67–69], review papers in facial expression analysis [70–72], and many others.

Appearance-based face recognition algorithms are alternative group of methods which proceed by projecting an image into the subspace and finding the closest point in such subspace [73]. Two well-known linear transformation methods that have been most widely used for dimensionality reduction and feature extraction are the principal component analysis (PCA) [20] and linear discriminant analysis (LDA) [73]. Object classes that are closer together in the output space are often weighted in the input space to reduce potential misclassification. The PCA could be operated either on the raw face image to extract the fisherface or on the eigenface to obtain the discriminant eigenfeatures [20]. Feature representation methods that combine the strengths of different realizations of PCA methods can be found in [74]. Kernel PCA [75] and the generalized discriminant analysis (GDA) using Kernel approach [76] have proved they are successful in pattern regression and classification tasks. Independent component analysis (ICA) provides a set of basis vectors that possess maximum statistical independence [77]. Face recognition techniques based on elastic graph matching [34] and support vector machines (SVMs) [78] also showed successful results. Line edge map approach [57] extracts lines from a face edge map as features, based on a combination of template matching and geometrical feature matching. The nearest feature line classifier [79] attempts to extend the capacity covering variations of pose, illumination, and expression for a face class by finding the candidate person that has the minimum distance between the feature point of query face and the feature lines connecting any two prototype feature points. A modified Hausdorff distance measure was also used to compare face images for recognition [79].

The environment proposed by the authors in [83] (called FaReS-Mod) only gives an option to design system based on two algorithms – PCA and LDA. These algorithms lack robustness to shadows and changing environment. Pictures used in the environment were all nearly frontal images and were captured under more or less constrained environment. The system created by this environment is not very

Fig. 2.9 Seven eigenfaces calculated from input images [6]

robust and cannot be used for images in less rigid environment; however, it does have an educational aspect. It allows to track every stage of face recognition process as it works.

We will briefly describe the basic idea of a face recognition algorithm using eigenface recognition described by Turk and Pentland [6] and face recognition based on elastic bunch graph matching [34] by Wisskot et al.

The main idea of PCA approach is to extract the relevant information from a face image, encode it as efficiently as possible, and compare one encoded face with a database of models encoded similarly. The approach includes extracting the information contained in an image of a face to somehow capture the variation in a collection of face images, independent of any judgment on features, and use this information to encode and compare individual face images. In mathematical terms, one has to find the principal components of the distribution of faces or the eigenvectors of the covariance matrix of the set of face images. These eigenvectors can be thought of as a set of features which together characterize the variation between face images. Each image location contributes to each eigenvector so that we can display the eigenvector as a sort of ghostly face which we call an eigenface. Some of these faces are shown in Fig. 2.9.

Each face image in the training set can be represented exactly in terms of a linear combination of the eigenfaces. The number of possible eigenfaces is equal to the number of face images in the training set. However, the faces can also be approximated using only the "best" eigenfaces – those that have the largest eigenvalues and which therefore account for the most variance within the set of face images. Computational efficiency is the primary reason for using a smaller number of eigenfaces.

If we consider the training set of face images $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_M$, the average face of the set is defined by $\Psi = \frac{1}{M} \sum_{n=1}^{M} \Gamma_n$. Each face differs from the average by the vector $\Phi_i = \Gamma - \Psi$. Set of very large vectors is then subject to principal component analysis which seeks a set of M orthonormal vectors u_n and their associated eigenvalues x_k which best describes the distribution of the data.

Since a complete algorithm can be found in many publications, we did not find it necessary to present it here.

Eigenfaces seem adequate for describing face images under controlled conditions. To perform identification task, new face image (Γ) is transformed into eigenface components by simple operation $\omega_k = u_k^T(\Gamma - \Psi)$ for $k = 1 \ldots M'$. This describes a set of point-by-point image multiplications and summations, operations performed at approximately frame rate on image processing hardware.

Since PCA approach cannot deal with nonlinearity as mentioned earlier, we think that presenting hybrid approach like EBGM in contrast would be more interesting.

Without going into details how individual faces and general knowledge about faces are represented by respectively labeled graphs and the face bunch graph (FBG) [34], we are now going to explain how these graphs are generated. Authors have used a method to generate initial graphs for the system, one graph for each pose, together with pointers to indicate which pairs of nodes in graphs for different poses correspond to each other. Once the system has an FBG (possibly consisting of only one manually defined model), graphs for new images can be generated automatically by elastic bunch graph matching. Initially, when the FBG contains only a few faces, it is necessary to review and correct the resulting matches, but once the FBG is big enough (approximately 70 graphs), one can rely on the matching and generate large collections of model graphs automatically.

Manual definition of graphs is done in three steps. At first the authors mark a set of fiducial points for a given image. Most of these points are positioned at well-defined features which are easy to locate, such as the left and right pupil, the corners of the mouth, the tip of the nose, the top and bottom of the ears, the top of the head, and the tip of the chin. These points were selected to make manual positioning easy and reliable. Additional fiducial points are positioned at the center of gravity of certain easy-to-locate fiducial points. This allows an automatic selection of fiducial points in the regions where well-defined features are missing, for example, at the cheeks or the forehead. Then, edges are drawn between fiducial points, and edge labels are automatically computed as the differences between node positions. Finally, the Gabor wavelet transform provides the jets for the nodes. In general, the set of fiducial points should cover the face evenly.

A key role in elastic bunch graph matching is played by the function evaluating the graph similarity between an image graph and the FBG of identical pose. It depends on the jet similarities and the distortion of the image grid relative to

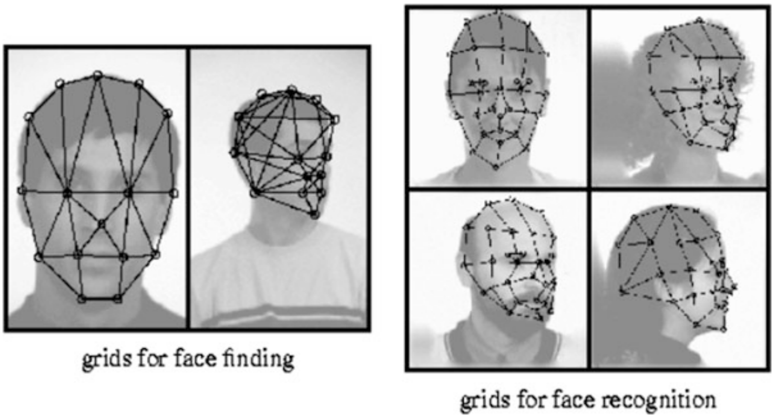

Fig. 2.10 Object-adapted grids for different poses. The nodes are positioned automatically by elastic graph matching against the corresponding face bunch graphs [34]

the FBG grid. Since the FBG provides several jets for each fiducial point, the best one is selected and used for comparison. These best fitting jets serve as local experts for the image face.

The goal of elastic bunch graph matching on a query image is to find the fiducial points and thus to extract the graph which maximizes the similarity with the FBG from the image. In practice, one has to apply a heuristic algorithm to find near-optimum results within a reasonable time. Authors use a coarse-to-fine approach in which they introduce the degrees of freedom of the FBG progressively: translation, scale, aspect ratio, and finally local distortions. Authors introduce phase information and increase the focus of displacement estimation in the similar way: no phase, phase with focus 1, and then phase with focus 1 up to 5. The matching schedule described here assumes faces of known pose and approximately standard size so that only one FBG is required.

The resulting graph is called the image graph and is stored as a representation of the individual face of the image (Fig. 2.10).

To minimize computing effort and to optimize reliability, authors extract a face representation in two stages. The first stage called the normalization stage is described in greater detail in [80]. Its purpose is to estimate the position and size of the face in the original image so that the image can be scaled and cut to standard size. The second stage takes this standardized image as an input and extracts a precise image graph appropriate for face recognition purposes. In the experiments on the face database, original images had a format of 256×384 pixels, and the faces varied in size by a factor of 3. The poses were known and did not need to be determined. The normalization stage used three FBGs of appropriate pose which differed in face size. Authors arbitrarily picked approximately 30 images to form each FBG. More careful selection of images to cover a wider range of variations can only improve system performance. The grids used in the construction of the FBGs put little emphasis, that is, few nodes, on the interior of the face and have fewer nodes than those used for the second stage (see Fig. 2.10 for two examples). The smaller number of nodes speeds up the process of face finding. Using a matching scheme similar to the one described earlier, authors match each of the three FBGs to the input image. Authors select the graph that matches best, cut a frame of appropriate size around it from the image, and resize it to 128 × 128 pixels. The poses could be determined analogously [80], although here they are assumed to be known. During the experiments, normalization took approximately 20 s on a SPARCStation 10–512 with a 50 MHz processor, and the system identified face position and scale correctly in approximately 99% of the images.

The simple algorithm proposed by the authors achieves 90% success recognition rate. However, this result can be reached only on databases smaller than 30 individuals. This limitation is caused by Toeplitz matrices, since they do not perform well on large databases. Algorithm such as PCA, ICA, and LDA can achieve 83% success rate [84], but tests performed with these algorithms were performed on databases that contained data for more than 1,000 individuals. If we compare PCA and EBGM algorithms [85] on similar database, we can see that results are quite similar. These results may vary based on different variables used in algorithms.

Authors' algorithm based on this size of database performed much worse. Because of that, further research on Toeplitz matrices must be performed to achieve good results on large databases.

2.5 Conclusions

Rapid progress and development of new technologies which increased the computational power of computers created possibility to build systems more complex and adjusting to the environment. During the past 20 years of constant development of new algorithms, researchers were able to create systems able to detect faces in images and recognize them. However, these systems are far from being perfect. They are still vulnerable to unconstrained environment, changes in facial expressions, or possibility of stealing a biometric key. These challenges show direction in which researchers should follow.

The algorithm for feature points detection presented by authors is very simple. However simple, it is able to work under different lighting conditions and it is robust to elements of environment because feature points are not being searched for outside the face area. Unfortunately the efficiency of a face descriptor based on Toeplitz matrices is not very high, which is not surprising. Based on the results we can notice that as the number of classes increases, the recognition rate drops. Currently, research on Toeplitz matrices focuses on maintaining recognition rate while increasing the size of the database.

Acknowledgement This work was partially supported by AGH University of Science and Technology in Cracow, grant no. 11.11.220.01. The authors are indeed indebted to Marcin Rogowski for his constructive remarks and thorough proofreading of the chapter.

References

1. Hietmeyer R (2000) Biometric identification promises fast and secure processing of airline passengers. Int Civ Aviat Organ J 55(9):10–11
2. Machine Readable Travel Documents (MRTD). http://www.icao.int/Security/mrtd/Pages/default.aspx. Accessed 23 May 2012
3. Kanade T (1973) Picture processing by computer complex and recognition of human faces. Ph.D. thesis, Kyoto University
4. Kohonen T (1989) Self-organization and associative memory. Springer, Berlin
5. Kirby M, Sirovich L (1990) Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Pattern Anal Mach Intell 12(1):103–108
6. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cog Neurosci 3(1):71–86
7. Fukunaga K (1990) Introduction to statistical pattern recognition, 2nd edn. Academic, Boston
8. Sirovich L, Kirby M (1987) Low-dimensional procedure for the characterization of human faces. J Opt Soc Am A 4(3):519–524
9. Bichsel M, Pentland A (1994) Human face recognition and the face image set's topology. CVGIP Image Understand 59:254–261
10. Turk M (2001) A random walk through eigenspace. IEICE Trans Inf Syst E84-D (12):1586–1695
11. Face Recognition Vendor Tests (FRVT). http://www.nist.gov/itl/iad/ig/frvt-home.cfm. Accessed 25 May 2012
12. Phillips PJ, Moon H, Rizvi SA, Rauss PJ (2000) The FERET evaluation methodology for face-recognition algorithms. IEEE Trans Pattern Anal Mach Intell 22(10):1090–1104
13. Chellappa R, Wilson C, Sirohey S (1995) Human and machine recognition of faces: a survey. Proc IEEE 83:705–740
14. Valentin D, Abdi H, O'Toole AJ, Cottrell GW (1994) Connectionist models of face processing: a survey. Pattern Recogn 27(9):1209–1230
15. Zhao W, Chellappa R, Phillips P, Rosenfeld A (2003) Face recognition: a literature survey. ACM Comput Surv 35(4):399–458. doi:10.1145/954339.954342
16. Caltech database. http://www.vision.caltech.edu. Accessed 10 Dec 2011
17. Moses Y, Adini Y, Ullman S (1994) Face recognition: the problem of compensating for changes in illumination direction. In: Proceedings of the European conference on computer vision, Stockholm, Sweden, vol A, pp 286–296
18. Sirovich L, Kirby M (1987) Low-dimensional procedure for the characterization of human faces. J Opt Soc Am A 4(3):519–524
19. Bartlett MS, Lades HM, Sejnowski TJ (1998) Independent component representations for face recognition. In: Proceedings of the SPIE, conference on human vision and electronic imaging III, San Jose, California, USA, vol 3299, pp 528–539
20. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19(7):711–720
21. Brunelli R, Poggio T (1993) Face recognition: features versus templates. IEEE Trans Pattern Anal Mach Intell 15(10):1042–1052
22. Goldstein AJ, Harmon LD, Lesk AB (1971) Identification of human faces. Proc IEEE 59(5):748–760
23. Kanade T (1973) Picture processing by computer complex and recognition of human faces. Ph.D. thesis, Kyoto University
24. Samal A, Iyengar PA (1992) Automatic recognition and analysis of human faces and facial expressions: a survey. Pattern Recogn 25:65–77
25. Cox IJ, Ghosn J, Yianilos P (1996) Feature-based face recognition using mixture-distance. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition, San Francisco, USA, pp 209–216
26. Scholkopf B, Smola A, Muller KR (1999) Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput 10:1299–1319

27. Mika S, Ratsch G, Weston J, Scholkopf B, Mller KR (1999) Fisher discriminant analysis with kernels. Neural Netw Signal Process IX:41–48
28. Guo GD, Li SZ, Chan KL (2000) Face recognition by support vector machines. In: Proceedings of fourth IEEE international conference on automatic face and gesture recognition, Grenoble, France, pp 196–201
29. Li Y, Gong S, Liddell H (2001) Recognising trajectories of facial identities using kernel discriminant analysis. In: Proceedings of British machine vision conference, Manchester, UK, pp 613–622
30. Moghaddam B (1999) Principal manifolds and Bayesian subspaces for visual recognition. In: International conference on computer vision (ICCV'99), Corfu, Greece, pp 1131–1136
31. Yang MH, Ahuja N, Kriegman D (2000) Face recognition using kernel eigenfaces. In: Proceedings of the IEEE international conference on image processing, Vancouver, BC, Canada, vol 1, pp 37–40
32. Penev P, Atick J (1996) Local feature analysis: a general statistical theory for object representation. Neural Syst 7(3):477–500
33. Lades M, Vorbruggen J, Buhmann J, Lange J, Malsburg C, Wurtz RP, Konen W (1993) Distortion invariant object recognition in the dynamic link architecture. IEEE Trans Comput 42:300–311
34. Wiskott L, Fellous J, Kruger N, Malsburg C (1997) Face recognition by elastic bunch graph matching. IEEE Trans Pattern Anal Mach Intell 19(7):775–779
35. Liu C, Wechsler H (2002) Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Trans Image Process 11(4):467–476
36. Ahonen T, Hadid A, Pietikainen M (2004) Face recognition with local binary patterns. In: Proceedings of the European conference on computer vision, Prague, Czech, pp 469–481
37. Gordon GG, Lewis ME (1995) Face recognition using video clips and mug shots. In: Proceedings of office of national drug control policy (ONDCP) international technical symposium, Nashua, NH
38. Lengagne R, Tarel JP, Monga O (1996) From 2d images to 3d face geometry. In: Proceedings of IEEE international conference automatic face and gesture recognition, Killington, USA, pp 301–306
39. Atick JJ, Griffin PA, Redlich AN (1996) Statistical approach to shape from shading: reconstruction of 3Dface surfaces from single 2Dimages. Neural Comput 8(6):1321–1340
40. Yan Y, Zhang J (1998) Rotation-invariant 3D recognition for face recognition. In: Proceedings of IEEE international conference image processing, Prague, Czech, vol 1, pp 156–160
41. Zhao WY, Chellappa R (2000) 3D model enhanced face recognition. In: Proceedings of IEEE international conference on image processing, Nashua, NH
42. Yang P, Shan S, Gao W, Li SZ, Zhang D (2004) Face recognition using ada-boosted gabor features. In: Proceedings of international conference on automatic face and gesture recognition, Killington, USA
43. Zhang G, Huang X, Li SZ, Wang Y (2004) Boosting local binary pattern (LBP)-based face recognition. In: Li SZ, Lai J, Tan T, Feng G, Wang Y (eds) Advances in biometric personal authentication, vol 3338, Lecture notes in computer science. Springer, Berlin, pp 180–187
44. Zhang L, Li SZ, Qu Z, Huang X (2004) Boosting local feature based classifiers for face recognition. In: Proceedings of first IEEE workshop on face processing in video, Washington, DC
45. Chen W, Er M, Wu S (2005) PCA and LDA in DCT domain. Pattern Recogn Lett 26:2474–2482
46. Chen W, Meng JE, Shiqian W (2006) Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Trans Syst Man Cybern B: Cybern 36(2):458–466
47. Tan X, Triggs B (2007) Preprocessing and feature sets for robust face recognition. In: IEEE conference on computer vision and pattern recognition, CVPR '07, Minneapolis, Minnesota, USA, pp 1–8

48. Xie X, Zheng W, Lai J, Yuen PC (2008) Face illumination normalization on large and small scale features. In: IEEE conference on computer vision and pattern recognition, CVPR '08, Anchorage, AK, USA, pp 1–8
49. Abbas A, Khalil MI, Abdel HS, Fahmy HMA (2009) Illumination invariant face recognition in logarithm discrete cosine transform domain. IEEE ICIP'2009, Cairo, Egypt, pp 4157–4160
50. Shao M, Wang Y (2009) Joint features for face recognition under variable illuminations. In: Fifth international conference on image and graphics, ICIG'09, Xi'an, Shanxi, China, pp 922–927
51. Liau HF, Isa D (2010) New illumination compensation method for face recognition. Int J Comput Netw Secur 2(3):308–321
52. Han H, Shan S, Qing L, Chen X, Gao W (2010) Lighting aware preprocessing for face recognition across varying illumination. LNCS 6312/ECCV 2010, Crete, Greece, pp 308–321
53. Goel T, Nehra V, Vishwakarma VP (2010) Comparative analysis of various illumination normalization techniques for face recognition. Int J Comput Appl 28(9):1–7
54. Saeed K (2004) Image analysis for object recognition. Bialystok Technical University Press, Bialystok
55. Kocjan P, Saeed K (2011) A feature based algorithm for face image description. In: Proceedings of IEEE-ICBAKE, IEEE CS Press–CD, Takamatsu, 19–21 Sep 2011, Japan, pp 175–178
56. Kocjan P, Saeed K (2011) Algorithm for extraction feature points from human face and their use in Toeplitz matrices. Faculty of Biomedical Engineering, Silesian University of Technology, Gliwice
57. Gao Y, Leung MKH (2002) Face recognition using line edge map. IEEE Trans Pattern Anal Mach Intell 24(6):764–779
58. Ivanevic V, Kaine AK, Mclindin BA, Sunde J (2003) Factor analysis of essential facial features. In: 25th international conference on information technology interfaces, Cavtat, Croatia, pp 187–191
59. Cootes TF, Edwards GJ, Taylor CJ (1998) Active appearance models. In: Computer vision – ECCV'98, Freiburg, Germany, vol 2, pp 484–498
60. Huang CL, Huang YM (1997) Facial expression recognition using model-based feature extraction and action parameters classification. J Vis Commun Image Represent 8:278–290
61. Kobayashi H, Hara F (1992) Recognition of six basic facial expression and their strength by neural network. In: IEEE international workshop on robot and human communication, Tokyo, Japan, pp 381–386
62. Pantic M, Rothkrantz LJM (2004) Facial action recognition for facial expression analysis from static face images. IEEE Trans Syst Man Cybern 34(3):1449–1461
63. Valstar M, Pantic M (2006) Fully automatic facial action unit detection and temporal analysis. In: Proceedings of the 2006 conference on computer vision and pattern recognition workshop (CVPRW'06) NY, USA
64. Cohn JF, Zlochower AJ, Lien JJ, Kanade T (1998) Feature-point tracking by optical flow discriminates subtle differences in facial expression. In: Proceedings of third IEEE FG, Nara, Japan, pp 396–401
65. Zhang Z, Lyons M, Schuster M, Akamatsu S (1998) Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron. In: Proceedings of third IEEE FG, Nara, Japan, pp 354–459
66. Cohn J, Zlochower A, Lien JJ, Kanade T (1999) Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. Psychophysiology 36:35–43
67. Tian YI, Kanade T, Cohn JF (2001) Recognizing action units for facial expression analysis. IEEE Trans Pattern Anal Mach Intell 23:97–115
68. Donato G, Bartelett MS, Hager JC, Ekman P, Sejnowski TJ (1999) Classifying facial actions. IEEE Trans Pattern Anal Mach Intell 21:974–989
69. Essa IA, Pentland AP (1997) Coding, analysis, interpretation, and recognition of facial expressions. IEEE Trans Pattern Anal Mach Intell 19:757–763

70. Fasel B, Luettin J (2002) Automatic facial expression analysis: a survey. Pattern Recogn 36:259–275
71. Pantic M, Rothkrantz LJM (2000) Expert system for automatic analysis of facial expressions. Image Vision Comput 18:881–905
72. Pantic M, Rothkrantz LJM (2000) Automatic analysis of facial expressions: the state of the art. IEEE Trans Pattern Anal Mach Intell 22:1424–1445
73. Bartlett MS, Movellan JR, Sejnowski TJ (2002) Face recognition by independent component analysis. IEEE Trans Neural Netw 13(6):1450–1464
74. Lu J, Plataniotis KN, Venetsanopoulos AN (2002) Face recognition using LDA-based algorithms. IEEE Trans Neural Netw 14(1):195–200
75. Kim KI, Jung K, Kim HJ (2002) Face recognition using kernel principal component analysis. IEEE Signal Process Lett 9(2):40–42
76. Baudat G, Anouar F (2000) Generalized discriminant analysis using a Kernel approach. Neural Comput 12(10):2385–2404
77. Bell AJ, Sejnowski TJ (1995) A non-linear information maximization algorithm that performs blind separation. Adv Neural Inf Process Syst 7:467–474
78. Phillips PJ (1999) Support vector machines applied to face recognition. Adv Neural Inf Process Syst 11:113–123
79. Li SZ, Lu J (1999) Face recognition using the nearest feature line method. IEEE Trans Neural Netw 10(2):439–443
80. Krüger N, Pötzsch M, Malsburg C (1997) Determination of face position and pose with a learned representation based on labeled graphs. Image Vision Comput 15(8):665–673
81. Images from public domain. http://www.totallyfreeimages.com. Accessed 10 May 2012
82. Images from public domain. http://www.publicdomainpictures.net. Accessed 10 May 2012
83. Kukharev G, Kuzminski A (2002) Biometric techniques: methods of face recognition. Szczecin Technical University Press, Szczecin, Poland (in Polish: Techniki Biometryczne: Metody rozpoznawania twarzy)
84. Delac K, Grgic M, Grgic S (2006) Independent comparative study of PCA, ICA, and LDA on the FERET data set. Int J Imag Syst Technol 15(5):252–260
85. Katadound S (2004) Face recognition: study and comparison of PCA and EBGM algorithms. Master thesis, Western Kentucky University

Springer
http://www.springer.com/978-1-4614-5607-0

Biometrics and Kansei Engineering
Saeed, K.; Nagashima, T. (Eds.)
2012, XII, 276 p., Hardcover
ISBN: 978-1-4614-5607-0