Recent advances in visual and infrared face recognition—a review

Seong G. Kong,* Jingu Heo, Besma R. Abidi, Joonki Paik,
and Mongi A. Abidi

*Imaging, Robotics, and Intelligent Systems Laboratory, Department of Electrical and Computer Engineering,
The University of Tennessee, Knoxville, TN 37996-2100, USA*

Abstract

Face recognition is a rapidly growing research area due to increasing demands for security in commercial and law enforcement applications. This paper provides an up-to-date review of research efforts in face recognition techniques based on two-dimensional (2D) images in the visual and infrared (IR) spectra. Face recognition systems based on visual images have reached a significant level of maturity with some practical success. However, the performance of visual face recognition may degrade under poor illumination conditions or for subjects of various skin colors. IR imagery represents a viable alternative to visible imaging in the search for a robust and practical identification system. While visual face recognition systems perform relatively reliably under controlled illumination conditions, thermal IR face recognition systems are advantageous when there is no control over illumination or for detecting disguised faces. Face recognition using 3D images is another active area of face recognition, which provides robust face recognition with changes in pose. Recent research has also demonstrated that the fusion of different imaging modalities and spectral components can improve the overall performance of face recognition.
© 2004 Elsevier Inc. All rights reserved.

1. Introduction

Demands and research activities in machine recognition of human faces from still and video images have increased significantly over the past 30 years. Face

---

* Corresponding author.
  E-mail address: skong@utk.edu (S.G. Kong).

recognition has been a major research focus area since traditional means of security such as ID cards and passwords are not reliable or convenient enough [1,2]. Face recognition is essential for interpreting facial expressions, human emotions, intentions, and behavior, which is a key component for smart environments [3]. Face recognition offers a non-intrusive, and perhaps the most natural, way of identification. Although several biometric authentication methods based on other physiological characteristics (such as fingerprints, retina and iris patterns, hand geometry, and voice) can be used, such biometric identification systems mostly rely on the cooperation of the participants. Authentication using face recognition is intuitive and does not have to stop user activities. The analysis of frontal or profile face images is often effective without the participants' cooperation or knowledge. Application areas for face recognition technology are broad including identification for law enforcement, matching of photographs on passports or driver's licenses, access control to secure computer networks and facilities such as government buildings and courthouses, authentication for secure banking and financial transactions, automatic screening at airports for known terrorists, and video surveillance usage [4]. Such applications range from static matching of controlled format photographs to real-time matching of video image sequences. In the computer security area, a face recognition system can be used to continually re-verify the identity of the system's user, and to confirm authorization level prior to performing each action.

Face recognition addresses the problem of identifying or verifying one or more persons of interest in a scene by comparing input faces with face images stored in a database. The general procedure for face recognition can be formulated by face detection, feature extraction, and recognition. Face detection is used to segment out face-like objects from cluttered scenes. Face images are usually represented in terms of feature vectors in lower dimensional feature space for recognition. Face recognition tasks include both identification and verification. Face identification refers to the process that given an unknown face input, the system reports its identity by looking up a database of known individuals. In verification tasks, the system confirms or rejects the claimed identity of the input face. Additional information such as race, age, gender, and facial expression can be used to enhance recognition accuracy. Human faces look similar in structure with minor differences from person to person. Classical pattern recognition problems such as character recognition have a limited number of classes, typically less than 50, with a large number of training samples available for each category. In face recognition, on the other hand, a relatively small number of face images are available for training while there exist a large number of possible face classes.

While humans quickly and easily recognize faces under variable situations or even after several years of separation, the problem of machine face recognition is still a highly challenging task in pattern recognition and computer vision [5,6]. A face is inherently a 3D object illuminated by a variety of lighting sources from different directions and surrounded by arbitrary background objects. Therefore, the appearance of a face varies tremendously when projected onto a 2D image. Different pose angles also cause significant changes in 2D appearance. Robust face recognition requires the ability to recognize identity despite such variations in appearance that the face

can have in a scene. Simultaneously, the system must be robust to typical image acquisition problems such as noise, video–camera distortion, and image resolution.

Despite the success of automatic face recognition techniques in many practical applications, recognition based only on the visual spectrum has difficulties performing consistently in uncontrolled operating environments. Performance of visual face recognition is sensitive to variations in illumination conditions [7]. Variations between the images of the same face due to changes in illumination and viewing directions are typically larger than image variations raised from changes in face identity. Other factors such as facial expressions [8] and pose variations [9] further complicate the face recognition task. Visual face recognition techniques have difficulty in identifying individuals wearing disguises or makeup. Disguised face detection is of particular interest in high-end security applications. Simple disguises such as a fake nose or beard substantially change a person's visual appearance. Obviously, visual identification of identical twins or faces in which the appearance has been altered through plastic surgery is almost impossible.

Recognition of faces using different imaging modalities, in particular infrared (IR) imaging sensors has become an area of growing interest [10,11]. Electromagnetic spectral bands below the visible spectrum such as X-rays and ultraviolet radiation are harmful to the human body and cannot be employed for face recognition applications. Thermal IR imagery has been suggested as an alternative source of information for detection and recognition of faces. While visual cameras measure the electromagnetic energy in the visible spectrum range (0.4–0.7 μm), sensors in the IR camera respond to thermal radiation in the infrared spectrum range at 0.7–14.0 μm. The infrared spectrum comprises the reflected IR and the thermal IR wavebands. The reflected IR band (0.7–2.4 μm) is associated with reflected solar radiation that contains no information about the thermal properties of materials. The near-infrared (NIR) (0.7–0.9 μm) and the short-wave infrared (SWIR) (0.9–2.4 μm) spectra are reflective and differences in appearance between the visible and the reflective IR are due to the properties of the reflective materials. This radiation is for the most part invisible to the human eye. The thermal IR band is associated with thermal radiation emitted by the objects. The amount of emitted radiation depends upon both the temperature and the emissivity of the material. The thermal IR spectrum is divided into two primary bands: the mid-wave infrared (MWIR) of the spectral range 3.0–5.0 μm and long-wave infrared (LWIR) from 8.0–14.0 μm. There are strong atmospheric absorption bands at 2.4–3.0 μm between the SWIR and MWIR range and at 5.0–8.0 μm between the MWIR and LWIR band where imaging becomes extremely difficult. The human face and body emit thermal radiation in both bands of the thermal IR spectrum. Thermal IR cameras can sense temperature variations in the face at a distance, and produce thermograms in the form of 2D images. Face recognition in the thermal IR favors the LWIR due to much higher emissions in this band than in the MWIR.

While sacrificing color recognition, thermal IR face recognition techniques can be used to identify faces when there is little or no control over lighting conditions. One advantage of using thermal IR imaging over visible spectrum sensors arises from the fact that the light in the thermal IR range is emitted rather than

reflected [12]. Thermal emissions from skin are an intrinsic property, independent of illumination. Therefore, the face images captured using thermal IR sensors will be nearly invariant to changes in ambient illumination. IR energy can be viewed in any light condition and is less subject to scattering and absorption by smoke or dust than visible light. The within-class variability is also significantly lower than that observed in visible imagery. The infrared spectrum has been found to have advantages over the visible spectrum for face detection [13,14], detection of disguised faces [15], and face recognition under poor lighting conditions. Thermal IR imaging has been an effective diagnostic tool in the study of skin temperature distribution for breast cancer detection [16], heat source recognition for electronic parts inspection [17,18], and target detection [19] in military applications. Detecting suspects engaged in illegal and potentially harmful activities using thermal images are proposed in [20]. Symptoms such as alertness and anxiety can be used as a difficult to conceal biometric as redistribution of blood flow in blood vessels causes abrupt changes in local skin temperature. Facial expression recognition based on thermal imaging has been investigated in [21].

Visual cameras represent the reflectance information of a face object, while thermal IR sensors measure anatomical information. Fig. 1 shows visual and thermal image characteristics of faces with variations in illumination and facial expression. Although illumination and facial expression significantly change the visual appearance of the face, thermal characteristics of the face remain nearly invariant. In thermal IR images, background clutter is not visible and the tasks of face detection, location, and segmentation are relatively easier and more reliable than in visual images.

Face recognition under very low lighting is almost impossible from visual images. A face image taken under very little ambient light is unrecognizable as in Fig. 2A. In Fig. 2B, however, a thermal IR image taken in the same environment reveals robust thermal characteristics of the face regardless of ambient illumination.

Thermal imaging has limitations in situations such as recognition of a person wearing glasses or seated in a moving vehicle. Glass blocks a large portion of thermal energy resulting in a loss of information near the eyes. Figs. 3A and B illustrate the visual and thermal appearances of the face with eyeglasses. Variations in ambient or body temperature also significantly change the thermal characteristics of the object. In Figs. 3C and D, visual and thermal images are taken after the subject was involved in physical exercise. Increased body temperature changes the thermal characteristics of the face, while the visual image features do not show a significant difference.

Several excellent survey papers on face recognition techniques are available. A survey by Samal and Iyengar [22] covers early face recognition approaches. Valentin et al. [23] review techniques based on neural networks, where associative memory and classification neural networks are used for face recall and recognition. Chellappa et al. [24] provide a comprehensive survey on face recognition techniques over 20 years including psychophysics issues relevant to face recognition. Zhang et al. [25] reviewed face recognition techniques in the aspects of eigenface, elastic matching, and neural networks. Aspects of face and gesture recognition are covered in [26]. Efforts in face recognition using infrared sensors have been relatively limited com-
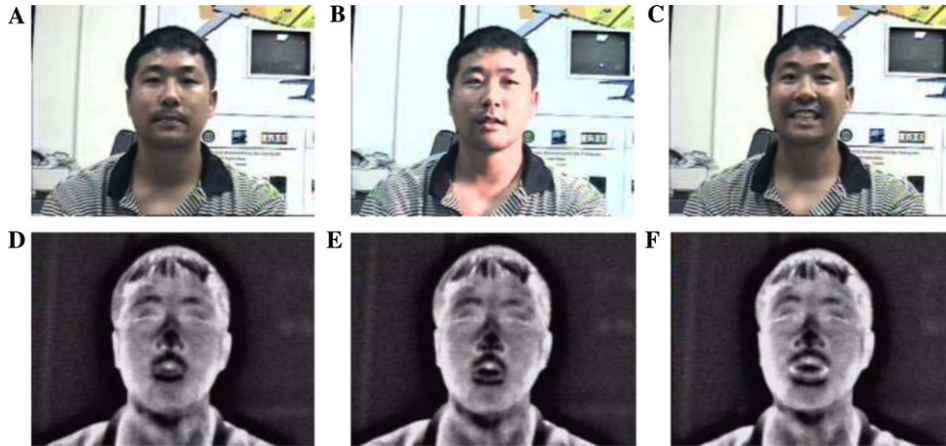
Fig. 1. Comparison of visual and thermal IR images under variations in illumination and facial expression. (A) and (B) visual face images with different illumination directions. (C) Different facial expression. (D) (E), and (F) are the corresponding thermal images to (A), (B), and (C).
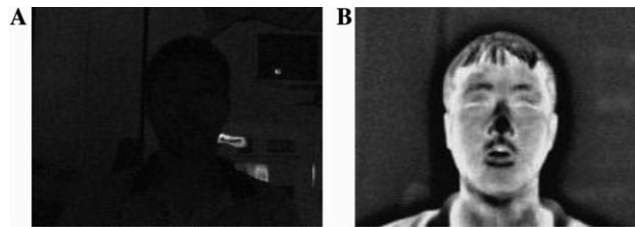

Fig. 2. Face images taken under very low lighting. (A) Visual image. (B) Corresponding thermal IR image.

Fig. 3. Changes in the thermal appearance of the face in the presence of eyeglasses and with body temperature changes. (A) and (B) wearing glasses. (C) and (D) after physical exercise.

pared to visual face recognition. A comparative performance study of multiple face recognition methodologies using visual and thermal IR imagery was conducted in [11,27].

This paper provides an up-to-date review of recent trends and major research efforts in face recognition techniques focusing on the two major imaging modalities: visible and infrared imaging sensors. We undertake an extensive survey of recent advances in face recognition algorithms and technologies since the previous major survey [24] published in 1995. This paper is organized as follows: Section 2 covers general approaches to face detection and feature extraction. Face region detection is an essential procedure for automatic face recognition. Feature extraction finds relevant information with good discriminating capability from the detected face region. Section 3 overviews evaluation methods for face recognition techniques based on the Face Recognition Technology (FERET) database and evaluation methods. The FERET program encouraged advances in the development of face recognition algorithms. A number of statistical and distribution-free classifiers have been employed for face recognition. In Section 4, major face recognition algorithms developed to date are reviewed. Section 5 describes face detection and recognition techniques using IR imaging sensors. Section 6 addresses fusion of multiple imaging modalities for enhancing recognition performance.

2. Face detection and feature extraction

2.1. Face detection

Detecting and tracking of face-like objects in cluttered scenes is an important pre-processing stage of an overall automatic face recognition system [28]. An automatic face recognition usually begin with the detection of the face pattern, and then proceeds to normalize the face images using information about the location and appearance of facial landmarks such as the eyes. The face region needs to be segmented out from a still image or video before recognition since most face recognition algorithms assume that the face location is known. The performance of face recognition software depends on how one controls the area where faces are captured. For applications like mug shot matching, segmentation is relatively easy due to a rather uniform background. Motion and skin color provide useful clues for face detection. For a video sequence acquired from a surveillance camera, segmentation of a person in motion can be more easily accomplished using motion as a cue [29–31]. Color information also provides a useful key for face detection [32–34] while color-based approaches may have difficulties in detecting faces in complex backgrounds and under different lighting conditions. Major face detection approaches are summarized in [33]. Recent survey papers on face detection techniques can be found in [35,36]. Detecting the eyes serves an important role in face normalization for the purpose of template matching and facilitates further localization of other facial landmarks. Most eye localization methods [37–39] are template-based approaches. Illumination variations and the objects such as glasses make eye detection and localization a

challenging problem. An image editing technique that automatically removes eyeglasses in the faces has been proposed in [40].

Face detection is a challenging machine vision task especially in outdoor or semi-outdoor environments where illumination varies greatly. Since faces are essentially 3D objects, lighting changes can cast significant shadows on a face. This is one of the primary reasons why current face recognition technology is constrained to indoor access control applications where illumination is well controlled. Light reflected from human faces also varies significantly from person to person. This variability, coupled with dynamic lighting conditions, causes a serious problem. The use of an artificial illuminator can reduce light variability, but it will distract the people in the scene and reveal the presence of a surveillance system.

Face detection can be viewed as a special case of face recognition, a two-class (face versus non-face) classification problem. Some face recognition techniques may be directly applicable to detect faces, but they are computationally very demanding and cannot handle large variations in face images. Conventional approaches for face detection include knowledge-based methods, feature invariant approaches, template matching, and appearance-based methods [41,42]. Knowledge-based methods encode human knowledge to capture the relationships between facial features. Feature invariant approaches find structural features that exist even when the pose, viewpoint, or lighting conditions vary [43]. Both knowledge-based and feature invariant methods are used mainly for face localization. In template matching methods, several standard patterns for a face are stored to describe the face as a whole or the facial features separately. The correlations between an input image and the stored patterns are computed for detection. The templates are allowed to translate, scale, and rotate. Segments obtained from the curvature discontinuities of the head outline can be used as templates. Appearance-based methods learn the models (or templates) using a set of training images to capture the representative variability of facial appearances. This category of methods includes neural network-based algorithms [44,45] that detect upright and frontal views of faces in gray-scale images. Face detection using the support vector machines is presented in [46]. Appearance-based methods are advantageous in finding small faces or faces in poor-quality images.

Viola and Jones [47] developed a cascade of boosting classifiers built on an overcomplete set of Haar-like features that integrates the feature selection and classifier design in the same framework for a rapid face detection. An image representation called integral image allows a very fast feature evaluation. The learning process focuses on a small number of important features using AdaBoost [48] to ensure fast classification. The algorithm successively combines a set of efficient classifiers in a cascade structure, which increases the speed of the detector by focusing attention on the regions of the image that might contain objects of interest. Faces are detected from 384×288 pixel images at 15 frame per second on a conventional 700 MHz Intel Pentium III computer. The face detection scheme is extended using a set of rotated Haar-like features [49]. A probabilistic method for detecting and tracking multiple faces in a video sequence [50] integrates the information of face probabilities provided by the detector and the temporal information provided by the tracker to produce a method superior to most available detection and tracking methods.

2.2. Feature extraction for face recognition

Face recognition involves feature matching through a database using similarity measures. The procedure compares an input image against a database and reports a match. Existing face recognition approaches can be classified into two broad categories: analytic and holistic methods [51]. The *analytic* or feature-based approaches compute a set of geometrical face features such as the eyes, nose, and the mouth. The use of geometric features for face recognition is popular in the earlier literature [52]. The positions of the different facial features and the face outline form a feature vector. The feature points are usually chosen in terms of significance for face representation and reliability for automatic extraction. The location of those points is used to compute geometrical relationships including the areas, distances, and angles among the 'fiducial' points. Since detection of feature points precedes the analysis, such a system is robust to position variations in the image. The geometrical features are used to search for a candidate from a face database [53].

The *holistic* or appearance-based methods consider the global properties of the human face pattern. The face is recognized as a whole without using only certain fiducial points obtained from different regions of the face. Holistic methods generally operate directly on pixel intensity array representation of faces without the detection of facial features. Since detection of geometric facial features is not required, this class of methods is usually more practical and easier to implement as compared to geometric feature-based methods. Holistic methods depend on techniques that transform the image into a low-dimensional feature space with enhanced discriminatory power. For a high dimensional feature space, the distances from a given probe to its nearest and farthest neighbors may become indistinguishable [54]. Like most natural signals, face images contain significant statistical regularities or redundancies. Several dimensionality reduction schemes have been developed to discover low-dimensional representations of human face images by relying on their statistical regularities [55]. Dimensionality reduction makes the face recognition problem computationally tractable. The holistic face recognition techniques are often sensitive to variations in position and scale. The faces presented to the algorithms usually need to be either segmented or surrounded by a simple background. The holistic techniques provide accurate recognition results with standard, well-illuminated frontal mug-shot images. This is due to the algorithms' dependence on fundamentally linear or quasi-linear analysis techniques. Performance often degrades rapidly with pose changes, non-uniform illumination, and background clutter.

A combination of analytic and holistic methods [56] combined 16-point features with regions of the eyes, nose, and the mouth and demonstrated success in the identification of the faces at different perspective variations using a database containing 40 frontal-view faces. The method is composed of two steps. The first step employs an analytic method to locate 15 feature points on a face: face boundary (6), eye corners (4), mouth corners (2), eyebrows (2), and the nose (1). Rotation of the face can be estimated using geometrical measurements and a head model. The positions of the feature points are adjusted so that their corresponding positions in the frontal view are approximated. These feature points are then compared with those of the faces in

a database. Only similar faces in the database are considered in the next step. In the second step, feature windows for the eyes, nose, and mouth are compared with the database by correlation. The two parts are combined to form a complete face recognition system. This approach achieved a high recognition rate under different perspective variations.

3. Evaluation of face recognition techniques

To encourage advances in face recognition technologies, the U.S. Defense Advanced Research Projects Agency and the U.S. Army Research Laboratory established the FERET program [57]. FERET is designed to measure the performance of face recognition algorithms on a large database in practical settings. The database must contain a large number of test images for adequate assessment. The sample must be statistically similar to the images that can be observed in real-world applications. The FERET program provides a large database of facial images taken from 1199 individuals and collected between August 1993 and July 1996 to support algorithm development and evaluation. Five to eleven images were collected from each individual under relatively unconstrained conditions. The FERET evaluation protocol offers methods for evaluating the performance of existing face recognition algorithms. The FERET database employs faces with variable positions, scales, and illumination in a manner consistent with mug shot or driver's license photography. Each set consists of two frontal views with different facial expressions (fa and fb). For 200 individuals, a third frontal image was taken using a different camera and different lighting (fc). The remaining images were non-frontal and included right (pr) and left (pl) profiles, right (qr) and left (ql) quarter profiles, right (hr) and left (hl) half profiles, and some arbitrary positions. Half profile face images were rotated from 40 to 70°. A duplicate set was collected to provide variations in scale, pose, expression, and illumination of the face. The FERET database consists of 14,126 images of 1564 sets (1199 original sets and 365 duplicate sets). For developmental purposes, 503 sets of images were released to researchers, and the remaining sets were sequestered for independent evaluation.

The first and second FERET evaluation tests were administered in August 1994 and March 1995, respectively. Design of the third FERET test performed in September 1996 and March 1997 was more complex than the first two evaluations. Two sets of images, a target set of 3323 images and a query set of 3816 images were presented to each algorithm to be tested. The FERET evaluation tested 10 algorithms and identified the three best algorithms that demonstrated the highest level of recognition accuracy: the probabilistic eigenface method from MIT [58], elastic graph matching (EGM) from the University of Southern California [59], and subspace linear discriminant analysis (LDA) from the University of Maryland [60]. The eigenface method finds a global representation of the face based on principal component analysis (PCA) [61]. The subspace LDA algorithm uses a linear discriminant for dimensionality reduction. EGM begins by computing Gabor jets from the image and then does a flexible template comparison of image descriptions using a graph-matching

algorithm. Both EGM and eigenface algorithms were able to detect and recognize faces with minimum constraints. Discriminant analysis systems require approximate eye locations to operate. Details of the three FERET evaluations can be found in [62,63]. Local feature analysis (LFA) [64–66] was an early contender, but withdrew from testing to form a commercial enterprise. With databases of fewer than 200 people and images taken under similar conditions, the algorithms perform nearly perfectly. The LFA system uses a sparse version of the eigenface transform followed by a discriminative neural network. The FERET evaluations did not systematically compare different implementations of the same representations. Different implementations of PCA-based face recognition algorithm and different distance metrics were compared in [67] using the FERET performance scores.

Based on the FERET evaluation protocol, the Face Recognition Vendor Test (FRVT) evaluated several commercial face recognition systems available on the market. Some of the leading face recognition products are based on the algorithms developed by the top contenders in the FERET competitions. Typical face recognition products include FaceIt developed by Identix which was the best performer in FRVT 2000 [68] and was also one of the three best commercial software packages in FRVT 2002. Based on the LFA algorithm, FaceIt represents facial images in terms of 12–40 local feature points derived statistically from a representative ensemble of faces. FaceIt automatically detects human presence, locates, extracts, and tracks faces, and performs identification by matching against a database of people. Ranked second in FRVT 2000, Viisage uses the eigenface-based recognition algorithm.

In FRVT 2002 [69], released in March 2003, ten commercial firms participated and were tested with a large dataset—121,589 operational facial images of 37,437 individuals. FRVT 2002 characterized identification and watch list performances as a function of database size, estimated the variability in performance for different groups of people, characterized performance as a function of elapsed time between enrolled and new images of a person, and investigated the effect of demographics on performance. Three main tasks were considered: *verification*, *identification*, and *watch—list matching*. For a verification task, a person presents his biometric and an identity claim to a face recognition system. The system then compares the presented biometric with a stored biometric of the claimed identity. Based on the results of comparing the new and stored biometrics, the system either accepts or rejects the claim. An identification task gives a ranked listing of the candidates in a database that best match an unknown person presented to the database. In the watch-list matching task, a face recognition system first detects if an individual is on the watch list. If the individual matches one of the individuals in the watch list, then the system identifies the individual. Cognitec Systems GmbH, Eyematic Interfaces, and Identix ranked the three best among the 10 participants. This report also provides demographic results showing that males and older people are relatively easier to recognize than females and younger people.

FRVT 2002 results show that normal changes in indoor lighting do not significantly affect the performance of the top systems. Approximately, the same performance results were obtained using two indoor data sets, with different lighting, while face recognition using outdoor images remains unsuccessful with only a 50%

recognition rate with a false acceptance rate of 1%. The database used in FRVT 2002 also consisted of images of the same person taken on different days. The performance results using indoor imagery in this case, showed improvement in the capabilities of the face recognition systems over the previous 2 years. Database size also affects the performance. The best system showed identification rates of 85, 83, and 73% with databases of 800, 1600, and 37,437 people, respectively. The performance decreases by two to three overall percentage points for every doubling of the database size. FRVT 2002 addressed several important face recognition topics and the impact of new techniques for improving face recognition: 3D morphable models [70,71], normalization of similarity scores, and face recognition from video sequences.

4. Face recognition algorithms

A number of earlier face recognition algorithms are based on feature-based methods [72–74] that detect a set of geometrical features on the face such as the eyes, eyebrows, nose, and mouth. Properties and relations such as areas, distances, and angles between the feature points are used as descriptors for face recognition. Typically, 35–45 feature points per face were generated. The performance of face recognition based on geometrical features depends on the accuracy of the feature location algorithm. However, there are no universal answers to the problem of how many points give the best performance, what the important features are, and how to extract them automatically. Face recognition based on geometrical feature matching is possible for face images at resolution as low as 8×6 pixels [75] when single facial features are hardly revealed. This implies that the overall geometrical configuration of the face features is sufficient for recognition.

Since the 1990s, appearance-based methods have been dominant approaches in face recognition. Appearance-based face recognition algorithms proceed by projecting an image into subspace and finding the closest pattern. PCA and LDA have been two approaches widely used for dimensionality reduction and feature extraction [76]. Several leading commercial face recognition products use face representation methods based on the PCA or Karhunen–Loeve (KL) expansion techniques, such as eigenface and LFA. Multispace KL was introduced as a new approach to unsupervised dimensionality reduction for pattern representation and face recognition [77], which outperform KL when the data distribution is far from a multidimensional Gaussian. Fisher's linear discriminant analysis [78] has been widely employed in face recognition. LDA determines a set of optimal discriminant basis vectors so that the ratio of the between- and within-class scatters is maximized. LDA finds the best projection direction in which training samples of different classes are best separated. In traditional LDA, separability criteria are not directly related to the classification accuracy in the output space. Object classes that are closer together in the output space are often weighted in the input space to reduce potential misclassification [79]. The LDA could be operated either on the raw face image to extract the Fisherface [80,81] or on the eigenface to obtain the discriminant eigenfeatures [82,83]. Feature representation methods that combine the strengths

of different realizations of LDA methods have also been recently proposed [84–87]. Kernel PCA [88,89] and generalized discriminant analysis using a kernel approach [90] have been successful in pattern regression and classification tasks. A new kernel direct discriminant analysis algorithm for face recognition has been proposed in [91].

Motivated by the fact that much of the important information may be contained in the high-order relationship, face recognition based on independent component analysis (ICA) was proposed as a generalization that is sensitive to higher-order statistics, not second-order relationships [92,93]. ICA provides a set of basis vectors that possess maximum statistical independence whereas PCA uses eigenvectors to determine basis vectors that capture maximum image variance. Face recognition techniques based on elastic graph matching [59] and neural networks [94,95] showed successful results. Support vector machines (SVMs) [96–98] find the optimal separating hyperplane that maximizes the margin of separation in order to minimize the risk of misclassification not only for the training samples, but also for the unseen data in the test set. SVM has applied to face recognition [99–102] and gender classification [103]. The line edge map approach [104] 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 [105] attempts to extend the capacity covering variations of pose, illumination, and expression for a face class by finding the candidate person owning the minimum distance between the feature point of a query face and the feature lines connecting any two prototype-feature points. A modified Hausdorff distance measure was used to compare face images for recognition [106,107].

This section provides an overview of some of the major face recognition techniques developed to date. Eigenfaces and local feature analysis are described in Sections 4.1 and 4.2, while ICA-based face recognition methods are covered in Section 4.3. Face recognition using a line edge map is reviewed in Section 4.4. Section 4.5 describes a popular face recognition algorithm using EGM [59]. In Section 4.6, face recognition using neural networks and support vector machines is presented.

4.1. Eigenface

Kirby and Sirovich [108] showed that any particular face can be represented along the eigenpictures coordinate space, and that any face can be approximately reconstructed by using a small collection of eigenpictures and the corresponding PCA coefficients. The basic idea of PCA is to construct a subspace that represents an input face image with lower dimensional feature vectors. The principal components, derived from an ensemble of the face images serving as feature vectors, span the significant variations among known face images. The PCA algorithm finds an optimal linear transformation that maps the original n-dimensional data space into an m-dimensional feature space ($m < n$) to achieve dimensionality reduction. Suppose a set of N sample face images $\{x_1, x_2, \ldots, x_N\}$ is given for training. Each face image is modeled as an n-dimensional vector formed via lexicographic ordering of a 2D pixel array. The total scatter matrix can be represented as the correlation of the 'centered' face images

$$S = \sum_{k=1}^{N} (\boldsymbol{x}_k - \bar{\boldsymbol{x}})(\boldsymbol{x}_k - \bar{\boldsymbol{x}})^t, \tag{1}$$

where $\bar{\boldsymbol{x}}$ denotes the mean of the N sample vectors in the training set

$$\bar{\boldsymbol{x}} = \frac{1}{N} \sum_{k=1}^{N} \boldsymbol{x}_k. \tag{2}$$

Consider a linear transform matrix: $\boldsymbol{E} = [\boldsymbol{e}_1|\boldsymbol{e}_2|\cdots|\boldsymbol{e}_m]$. The column vectors of $\boldsymbol{E}$ are the eigenvectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ of $\boldsymbol{S}$ associated with the first m largest eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_m$. Large eigenvalues account for large variance in the set of training images. The m eigenvectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ constitute an m-dimensional feature space. The linear transform matrix $\boldsymbol{E}^t$ maximizes the determinant of the total scatter matrix $\boldsymbol{E}^t \boldsymbol{S} \boldsymbol{E}$ of the transformed vectors. The set of basis images are "optimal" in that the image coordinates are uncorrelated in this new basis, i.e., they cannot be linearly predicted from each other. The eigenvectors have the same dimension as the original images and show face-like images, referred to as eigenfaces.

In practice, finding the eigenvectors of the scatter matrix $\boldsymbol{S} = \boldsymbol{X}\boldsymbol{X}^t$, where the input matrix is given by $\boldsymbol{X} = [\boldsymbol{x}_1|\cdots|\boldsymbol{x}_N]$, of the size $n \times n$ is an intractable task for typical image sizes. For images of size 128×128, for example, the dimension is $n = 128^2$ and the size of the scatter matrix $\boldsymbol{X}\boldsymbol{X}^t$ becomes $128^2 \times 128^2 \cong 2.7 \times 10^8$. Hence, a simplified method of calculation has to be adopted. The number of training images is usually much smaller than the number of pixels in an image ($N \ll n$). The eigenvectors $\boldsymbol{e}$ and associated eigenvalues λ of $\boldsymbol{X}\boldsymbol{X}^t$ can be found from the eigenvectors $\tilde{\boldsymbol{e}}$ and associated eigenvalues $\tilde{\lambda}$ of $\boldsymbol{X}^t \boldsymbol{X}$, which are mathematically more tractable and easier to obtain. The eigenvectors are $\boldsymbol{e} = \boldsymbol{X}\tilde{\boldsymbol{e}}$ and the eigenvalues remain the same ($\lambda = \tilde{\lambda}$). Fig. 4A shows a training set used to compute the eigenfaces in Fig. 4B. We computed a set of eigenfaces for 25 normalized face images of 100×100 size.

Given a face image $\boldsymbol{x}$ for testing, the PCA analysis expands the face in terms of m eigenfaces. The linear transformation $\boldsymbol{W}^t$ produces an m-dimensional feature vector $\boldsymbol{a} = (a_1, \ldots, a_m)^t$

$$\boldsymbol{a} = \boldsymbol{W}^t(\boldsymbol{x} - \bar{\boldsymbol{x}}). \tag{3}$$

The transform coefficients or weights $a_1, \ldots, a_m$ characterize the expansion of the given image in terms of eigenfaces. Each of the transform coefficients $a_i = \boldsymbol{e}_i^t(\boldsymbol{x} - \bar{\boldsymbol{x}})$, $i = 1, \ldots, m$ describes the contribution of each eigenface to that face. The transform coefficients serve as features for face recognition. To recognize an unknown test face, the feature vector is then compared to the feature vector of each face image in the database. This leads not only to computational efficiency, but also makes the recognition more general and robust. The face image can be approximately represented in terms of a linear combination of the m eigenfaces, or 'component' faces

$$\boldsymbol{x} \approx \bar{\boldsymbol{x}} + \boldsymbol{W}\boldsymbol{a} = \bar{\boldsymbol{x}} + \sum_{i=1}^{m} a_i \boldsymbol{e}_i. \tag{4}$$
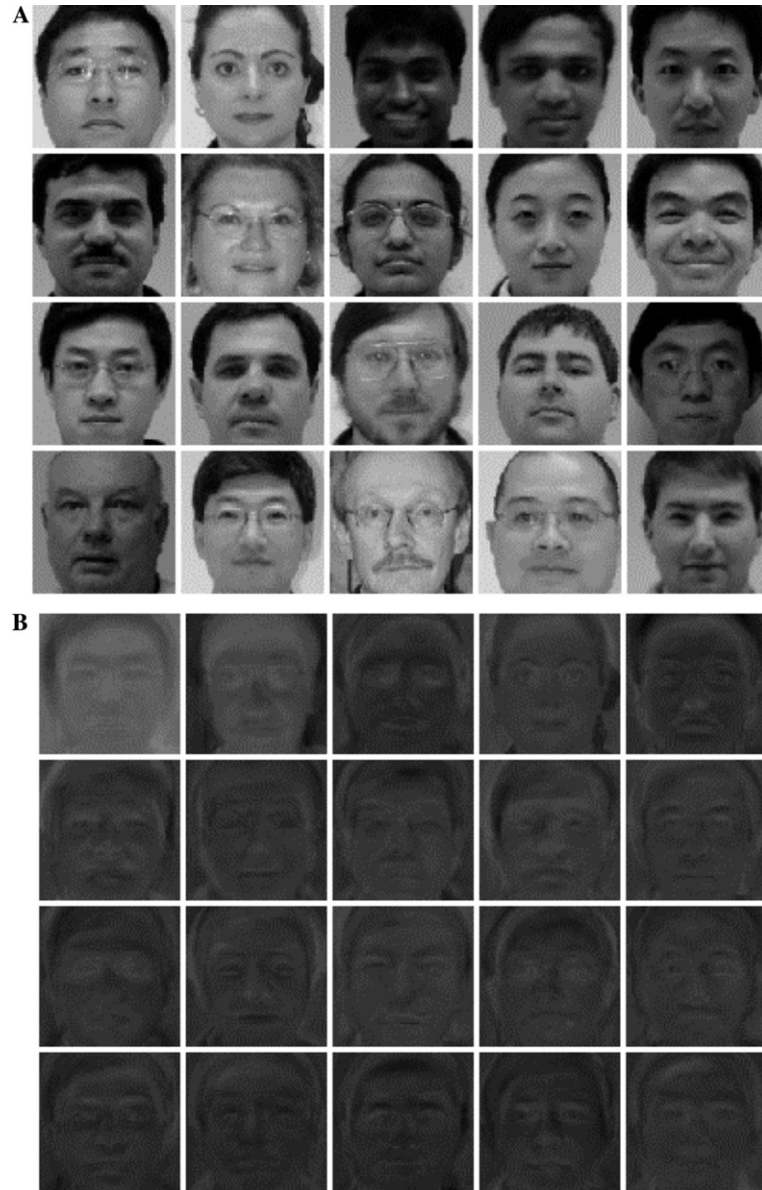
Fig. 4. Computation of the eigenfaces from a set of face images. (A) Sample training set. (B) Eigenfaces.

Eq. (4) provides the minimum mean square error among all possible approximations of x that use m orthonormal basis vectors. By using an increasing number of eigenvectors, one will get an improved approximation of the given image.

Turk and Pentland applied PCA further for detecting and recognizing faces in cluttered scenes, known as the eigenface method [109,110]. The eigenface method generates features that capture the holistic nature of faces through the PCA. They reported 96, 85, and 64% correct classifications averaged over lighting, orientation, and size variations, respectively for a database containing 2500 images of 16 individuals. Each face is often normalized using the location of the eyes in terms of rotation and scaling. Illumination normalization is usually necessary for the eigenface approach. Zhao and Yang [111] proposed a new method to compute the scatter matrix using three images each taken with different lighting conditions to account for arbitrary illumination effects. Pentland et al. [112] extended their early work on eigenfaces to modular eigenfeatures corresponding to face components, such as the eyes, nose, and mouth (referred to as eigeneyes, eigennose, and eigenmouth). This method would be less sensitive to appearance changes than the standard eigenface method. The system achieved a recognition rate of 95% on the FERET database of 7562 images of approximately 3000 individuals.

Applying discriminant analysis to the eigenface method improves the performance of face recognition [60]. The advantage of combining PCA and LDA is that they distinguish the different roles of within- and between-class scatter by applying discriminant analysis. A direct PCA approach does not distinguish the different roles of each principal component from the training samples and may lead to poor classification performance when the distributions of the face classes are not separated by the mean-difference. Fisherface methods [80] improve the performance of direct PCA approach by applying first PCA for dimensionality reduction and then Fisher's linear discriminant analysis. Fisherface algorithms are believed to outperform eigenface methods, since LDA extracts features more suitable for classification purposes, i.e., most discriminating features, while eigenfaces find the most expressive features, which are not necessarily helpful for recognition. As a drawback, Fisherfaces require multiple images for training for each person, which is not always available for some applications. An extended Fisherface algorithm using a single training image per person has been proposed in [113].

4.2. Local feature analysis

Local feature analysis [64] is a subspace representation scheme for facial data based on second order statistics results obtained by enforcing topographic indexing of the basis vectors, and minimizing their correlation. Locality and topography are desirable features in certain segmentation and pattern analysis tasks. The eigenface method provides optimal representation in reduced dimensionality, but is typically non-local and non-topographical. The LFA constructs a family of locally correlated feature detectors based on PCA decomposition. A selection, or sparsification, step produces a minimally correlated, and topographically indexed subset of features that define the subspace of interest. Local representations offer robustness against variability due to changes in localized regions of the objects. The features used in the LFA method are less sensitive to illumination changes, easier for estimating rotations, and have less computational burden than the eigenface method. The LFA

algorithm was used as a key algorithm in FaceIt, one of the successful commercial face recognition software packages.

The LFA algorithm selects kernels to represent local characteristics of the faces. The support of the kernels extends over the entire range of the data, but nearby values in the indexes do not possess any relationship to each other in contrast to nearby values of the grid variables that obey topography. Topography means that the kernels of the representation should be labeled with the grid variable x instead of the eigenvector index r. Given the eigenvectors $v_r(x)$ with eigenvalues λ_r, one can construct the topographic kernel that projects signals to the subspace spanned by the eigenvector

$$K(\boldsymbol{x}, \boldsymbol{y}) = \sum_{r=1}^{N} v_r(\boldsymbol{x}) \frac{1}{\sqrt{\lambda_r}} v_r(\boldsymbol{y}). \tag{5}$$

The rows of $K(x, y)$ contain kernels with spatially local properties. The kernel matrix $K(x, y)$ transforms the face set X into the LFA output O: $O = KX^t$. The original images can be reconstructed from O by $X^t = K^{-1}O$. The residual correlation of the outputs becomes

$$P(\boldsymbol{x}, \boldsymbol{y}) = \sum_{r=1}^{N} v_r(\boldsymbol{x}) v_r(\boldsymbol{y}). \tag{6}$$

The function $P(x, y)$ can be readily recognized as the projection operator onto the subspace spanned by the eigenvectors. The LFA reduces the dimensionality of the representation by choosing a subset of outputs that are as decorrelated as possible. The reconstruction error for the LFA representation is exactly equal to that of the PCA representation.

4.3. Independent component analysis

Independent component analysis for face recognition has been applied relatively recently [92,93]. ICA seeks non-orthogonal basis that are statistically independent [114,115], while PCA finds a set of orthogonal basis for face images of which the transformed features are uncorrelated. The basis images developed by PCA depend only on second-order image statistics. ICA generalizes the concept of PCA to higher-order image statistics relationships. The original motivation for this decomposition derived from the need to separate audio streams into independent sources without prior knowledge of the mixing process. Let $x = (x_1, x_2, \ldots, x_m)^t$ denote a zero-mean m-dimensional random variable observed, and let $s = (s_1, s_2, \ldots, s_n)^t$ be its n-dimensional transform. The ICA problem is to determine a constant weight matrix W so that the linear transformation of the observed variables

$$s = Wx, \tag{7}$$

yields components s_i that are statistically as independent from each other as possible. ICA is only possible if every independent component has a non-Gaussian distribution.

In face recognition, much of the important information may be contained in the high-order relationships among the image pixels. Higher order statistics of the images as the phase spectrum contain much of the information that perceptually distinguishes faces. Comparisons of ICA-based face representation with other methods can be found in [116,117]. Recently, face recognition performance using the ICA representations was benchmarked by comparing to performance using eigenfaces [93].

4.4. Line edge map

Humans can recognize rough line drawings quickly and almost as accurately as gray-level pictures. A line edge map (LEM) approach extracts lines from a face edge map as features, based on a combination of template matching and geometrical feature matching. Similarity of face images can be measured by a face feature representation scheme based on the LEM [104]. The faces are encoded into binary edge maps using the Sobel edge detection algorithm. The Hausdorff distance was chosen to measure the similarity of the two point sets, i.e., the edge maps of the two faces. The Hausdorff distance [106] is calculated without an explicit pairing of points in their respective data sets. The LEM method possesses the advantages of a feature-based approach, which is invariant to illumination, has low memory requirements, and shows high recognition performance using template matching.

Efficient coding of faces is an important aspect in a face recognition system. Face feature representation in the LEM integrates the structural information with the spatial information of a face image by grouping pixels of face edge maps into line segments. After thinning the edge map, a polygonal line fitting process is applied to generate the LEM of a face. The LEM representation records only the end points of line segments on curves, which further reduces the storage requirements. The LEM is also less sensitive to illumination changes due to the fact that it is an intermediate-level image representation derived from a low-level edge map representation. The basic unit of the LEM is the line segment grouped from pixels of the edge map. The line segment Hausdorff distance (LHD) measure is used to match LEMs of faces. LHD has better distinctive power because it uses additional structural attributes of line orientation, line-point association, and number disparity in obtaining the LEM. The LHD is a shape comparison measure based on LEMs, a distance defined between two line sets. In comparisons with the eigenface method, the LEM shows higher recognition rates.

4.5. Elastic graph matching

A face recognition system described in dynamic link architecture [118] represents individual faces by a rectangular graph, each node labeled with a set of complex Gabor wavelet coefficients, called a jet. A jet is used to represent the local features of the face images based on the Gabor wavelet transforms. Only the magnitudes of the coefficients are used for matching and recognition. For the recognition of a new face, each graph in the database is matched to the constructed graph of the new image separately and the best match indicates the recognized person. Rotation in depth is

compensated for by elastic deformation of the graphs. Wiskott and von der Malsburg [119] matched human faces against a gallery, or a collection of face images of known individuals of 112 neutral frontal view faces. Probe images were distorted due to rotation and change in facial expressions. Encouraging results using faces with large rotation angles were obtained. In general, dynamic link architecture is good in terms of invariance to rotation; however, the matching process is computationally expensive.

Elastic graph matching [59] extends the dynamic link architecture method in order to increase the matching accuracy for bigger databases and handle larger variations in poses. EGM uses the phase of the complex Gabor wavelet coefficients to achieve a more accurate location of the nodes and to disambiguate patterns, which would be similar in their coefficient magnitudes. EGM employs object adaptive graphs, so that nodes refer to specific facial landmarks, or fiducial points on the face such as the pupils, the corners of the mouth, and the tip of nose. The correct correspondences between two faces can then be found across large viewpoint changes. A new data structure called the bunch graph is introduced to serve as a generalized representation of faces by combining jets of a small set of individual faces. This allows the system to find the fiducial points in one matching process, which eliminates the need for matching each model graph individually. This reduces computational effort significantly. The goal of EGM on a test image is to find the fiducial points and thus extract from the image a graph that maximizes the similarity. Recognition experiments were conducted with galleries of 250 images from the FERET database, neutral frontal view (fa), frontal view with different facial expression (fb), half profile (hr and hl), and profile (pr and pl) images. For frontal against frontal images, the recognition rate was very high (98% for the first rank and 99% for the first 10 ranks) compared with matching of profile images. Morphological elastic graph matching [120] has been proposed for improvement. Use of support vector machines also enhances the performance of elastic graph matching for frontal face authentication [121]. EGM-based systems have good performance in general. However, they require a large-size image, e.g., 128×128. This restricts application to video-based surveillance, where the face image size is usually small.

4.6. Neural networks

Neural network approaches have been widely explored for feature representation and face recognition. An early face recognition technique using artificial neural networks is WISARD, a single-layer adaptive network containing a separate network for each stored individual [122]. Face recognition based on a hybrid neural network [123] has been proposed. Hybrid neural networks combine local image sampling, a self-organizing map (SOM) neural network, and a convolutional neural network. The SOM provides a quantization of the image samples into a topology preserving space where inputs located nearby in the original space also appear nearby in the output space. The SOM achieves dimensionality reduction and provides partial invariance to translation, rotation, scale, and deformation in the image sample. A convolution network extracts features successively in a hierarchical set of layers. The probabilistic decision-based neural network (PDBNN) [95] is effectively applied to

face detection and recognition. PDBNN has inherited the modular structure from its predecessor described in [94]. PDBNN-based identification systems have the merits of both neural networks and statistical approaches, and their distributed computing principle is relatively easy to implement on parallel computers. It was reported that face recognition using the PDBNN had the capability of recognizing up to 200 people and could achieve up to a 96% correct recognition rate in approximately 1 s [95]. However, when the number of people increases, the computation expenses become more demanding. In general, neural network approaches encounter problems when the number of classes (i.e., individuals) increases. Moreover, they are not suitable for a single model image recognition task because multiple model images per person are necessary in order to train the systems for 'optimal' parameter settings. Fusion of multiple neural networks classifiers improved the overall performance of face recognition [124]. A recent face recognition system using hybrid neural and dual eigenspace methods is proposed in [125].

Support vector machines have been proposed as a new technique for face recognition [100–102]. SVMs have been proposed by Vapnik [96,97] as an effective pattern classifier. SVMs find the optimal separating hyperplane that maximizes the margin of separation in order to minimize the risk of misclassification not only for the training samples, but also the unseen data in the test set. A SVM classifier is a linear classifier where the class separating hyperplane is chosen to minimize the expected classification defined by a weighted combination of a small subset of the training vectors, called support vectors. Estimating the optimal hyperplane is equivalent to solving a linearly constrained quadratic programming problem. SVMs can be considered as a new paradigm to train polynomial functions, neural networks, or radial basis function classifiers [126]. While most methods for training a classifier are based on minimizing the training error, i.e., empirical risk, SVMs operate on another induction principle called structural risk minimization. Structural risk minimization aims at minimizing an upper bound on the expected generalization error. An earlier study of SVMs in face recognition has been reported in [99]. Compared with a standard PCA method, the SVM verification system was found to be significantly better.

4.7. Hidden Markov models

Hidden Markov models (HMMs) have been successful in modeling temporal information in many speech, image, and video applications. When a traditional 1D Markov chain is used to model a 2D image signal, the signal has to be transformed into a 1D observation sequence. A high complexity of the 2D HMMs leads to models such as pseudo 2D HMMs [127], embedded HMMs [128], and embedded Bayesian networks [129]. Early work of HMM-based face recognition [127] models human faces with a vertical top-to-bottom 1D HMM structure composed of superstates. Each superstate contains a horizontal left-to-right 1D Markov chain. A similar 1D HMM [130] uses the discrete cosine transform of the observation strip as a feature vector to reduce the size of the feature vector by discarding the insignificant high-frequency coefficients. In [131], the image is scanned in a zigzag fashion to form a 1D observation sequence. In embedded HMMs, an image is scanned in a 2D manner where each

observation block retains vertical and horizontal indices for row and column, respectively. Embedded Bayesian networks, a generalized framework of embedded HMMs, shows a significant complexity reduction. More recently, a low-complexity 2D HMM structure [132] was derived based on the assumption of conditional independence among the neighboring observation blocks, which enables the separation of the 3D state transition matrix into two 2D vertical and horizontal state transition matrices. High recognition rates were obtained on the facial databases of the AT&T and Georgia Institute of Technology with reasonably low computational load. HMMs have been used in the temporal domain to perform face recognition in video signals, considering each frame in the video sequence as an observation [133].

5. Face recognition using infrared imagery

Thermal IR images or thermograms represent the heat patterns emitted from an object. Objects emit different amounts of IR energy according to their temperature and characteristics. The range of human face and body temperature is quite uniform, varying from 35.5 to 37.5 °C providing a consistent thermal signature. Skin temperature in a 21 °C ambient room temperature also has a small variable range between 26 and 28 °C. The thermal patterns of faces are derived primarily from the pattern of superficial blood vessels under the skin. The vessels transport warm blood throughout the body, and heat the skin above. Skin directly above a blood vessel is on the average 0.1 °C warmer than adjacent skin. The vein and tissue structure of the face is unique for each person, and therefore the IR images are also unique. It is known that even identical twins have different thermal patterns. The range and sensitivity are well within the specification of current IR imaging technology. The passive nature of the thermal IR systems lowers their complexity and increases their reliability.

5.1. Face and disguise detection using the infrared spectrum

Human faces from different ethnic groups have different reflectance characteristics. The problem of finding invariants related to skin color in the visible spectrum could be easily solved with a calibrated IR sensor. Infrared cameras can provide clear imaging signals during daytime and nighttime operation and even in certain adverse weather situations, such as hazy conditions. An external illumination source is not required since the face emits thermal energy. The visible spectrum sensor has certain disadvantages. It cannot easily see at night without the aid of an artificial illumination source, which is not applicable in many applications, as this would distract the human subject. In most cases the thermal signature of the face is distinct from that of the environment and facilitates robust segmentation. This is in contrast to the segmentation difficulties encountered in the visible spectrum due to physical diversity coupled with lighting, color, and shadow effects. While, light reflectivity of the skin varies significantly using visible sensors, emissivity values of the IR spectrum are relatively uniform among faces of different skin color. Face region detection and tracking methods using thermal IR sensors have been reported in [134].

The reflected IR spectral bands provide advantages for a solution to the face detection problem. Human skin shows unique reflectance characteristics and facial signatures are less variable in these spectral bands. A face detection system [13,14] is proposed based on NIR imagery and multi-band feature extraction. The reflected IR light is for the most part invisible to the human eye so the system can remain unobtrusive and covert. For IR imaging, the bulk of thermal radiation cannot transmit through glass as glass severely attenuates electromagnetic wave radiation beyond 2.4 μm. Unlike thermal IR bands, reflected IR light can easily penetrate a vehicle window. The use of the NIR spectrum has been successful in face detection for automatic detection and counting of vehicle occupants [135]. Vehicle occupant counting is one of crucial tasks for effective operation of freeway lanes reserved for car-pools or in high occupancy vehicle (HOV) lanes. Vehicle occupant counting facilitates law enforcement in the HOV lane and gathering of statistical data for road construction and planning.

Disguised face detection is critical for the deployment of face recognition systems in high-end security applications. Thermal face recognition is especially useful when the subject is wearing a disguise as well as under all lighting conditions including total darkness. Two types of disguises for altering facial characteristics are the use of artificial materials and surgical alterations. Artificial materials may include a fake nose, makeup, or wig. Surgical alterations modify facial appearance through plastic surgery. Visual identification of individuals with disguises or makeup is almost impossible without prior knowledge as the facial appearance of a person can change substantially through the use of a simple disguise. Disguises can be easily detected using the IR spectrum since various artificial materials used in a disguise change thermal signatures of the face [136]. The upper band of the SWIR at 1.4–2.4 μm may provide useful information for detecting disguised faces [15] due to unique and universal properties of the human skin in this subband. The truly unique advantage of thermal IR is its ability to uncover surgical alterations. Plastic surgery may add or subtract skin tissue, redistribute fat, add silicon, and create or remove scars. Surgical inclusions may cause alterations of blood vessel flow, which appears as distinct cold spots in the thermal imagery.

5.2. Face recognition using thermal infrared imagery

Face recognition based on the thermal IR spectrum utilizes anatomical information of the human face as features unique to each individual. Anatomical face features useful for identification can be measured at a distance using passive IR sensor technology with or without the cooperation of the subject. In addition to the currently available techniques for extracting features that depend only on external shape and surface reflectance, the thermal IR image offers new features that "uncover" thermal characteristics of the face. One advantage of the use of thermal IR imaging for boosting face recognition performance is apparent invariance to changing illumination. Changes in illumination appear to play less of a role in thermal infrared images. Thermal IR imagery is nearly invariant to changes in ambient illumination since the human face emits thermal energy, not reflected incident light.

Equinox Corporation collected an extensive database of face images using co-registered, broadband-visible/LWIR, MWIR, and SWIR camera sensors for experimentation and statistical performance evaluations. Equinox Corporation conducted validation to compare the performances of face recognition using visible and thermal IR imagery from their database [27]. Appearance-based face recognition algorithms applied to thermal infrared, particularly LWIR imaging, consistently performed better than when applied to visible imagery. Furthermore, face recognition algorithms applied to combined fusion of co-registered visible and LWIR using a fusion of experts' methodologies consistently demonstrated even better performance than when applied to either visible or LWIR imagery alone [137].

5.2.1. Thermal contour matching

Appearance-based approaches are commonly used for IR face recognition systems [137]. In contrast to visual face recognition algorithms that mostly rely on the eye location, thermal IR face recognition techniques present difficulties in locating the eyes. Initial research approaches to thermal face recognition extract and match thermal contours for identification. Such techniques include elemental shape matching and the eigenface method. Elemental shape matching techniques use the elemental shape of thermal face images. Several different closed thermal contours can be observed in each face. The sets of shapes are unique for each individual because they result from the underlying complex network of blood vessels. Variations in defining the thermal slices from one image to another has the effect of shrinking or enlarging the resulting shapes while keeping the centroid location and other features of the shapes constant. Perimeter, area, x and y coordinates of the centroid, minimum and maximum chord length through the centroid and between perimeter points, and standard deviation of that length are being considered. Automated face recognition using elemental shapes in real-time has reported 96% accuracy for cooperative access control applications.

The eigenface technique for face recognition can be successfully applied to thermal IR images. While the visible eigenfaces contain mostly low-frequency information and coding partly for variation in illumination, the corresponding thermal IR eigenfaces have fewer low-frequency components and many more high-frequency characteristics. Therefore, a majority of the variance of the data distribution is contained in a lower dimensional subspace for the thermal IR than for the visible imagery. Eigenfaces of thermal IR images are uniformly superior to eigenfaces of visible imagery. Classification performance is on average 17% higher for thermal IR imagery, and the best scenario always outperformed and yielded an improvement of 54% over visible imagery [138]. A 6-dimensional subspace is sufficient to capture over 95% of the variance of the thermal IR data, whereas a 36-dimensional subspace is needed to capture the same variance for the visible imagery. Fig. 5 illustrates a set of eigenfaces obtained from thermal IR images in the Equinox database.

Other approaches in this category are metrics matching and template matching. In metrics matching, feature points such as inner and outer eye corners, lowermost tip of the nose, and top and bottom points where the ears are connected to the head can be located from the visual and thermal images. Other feature points
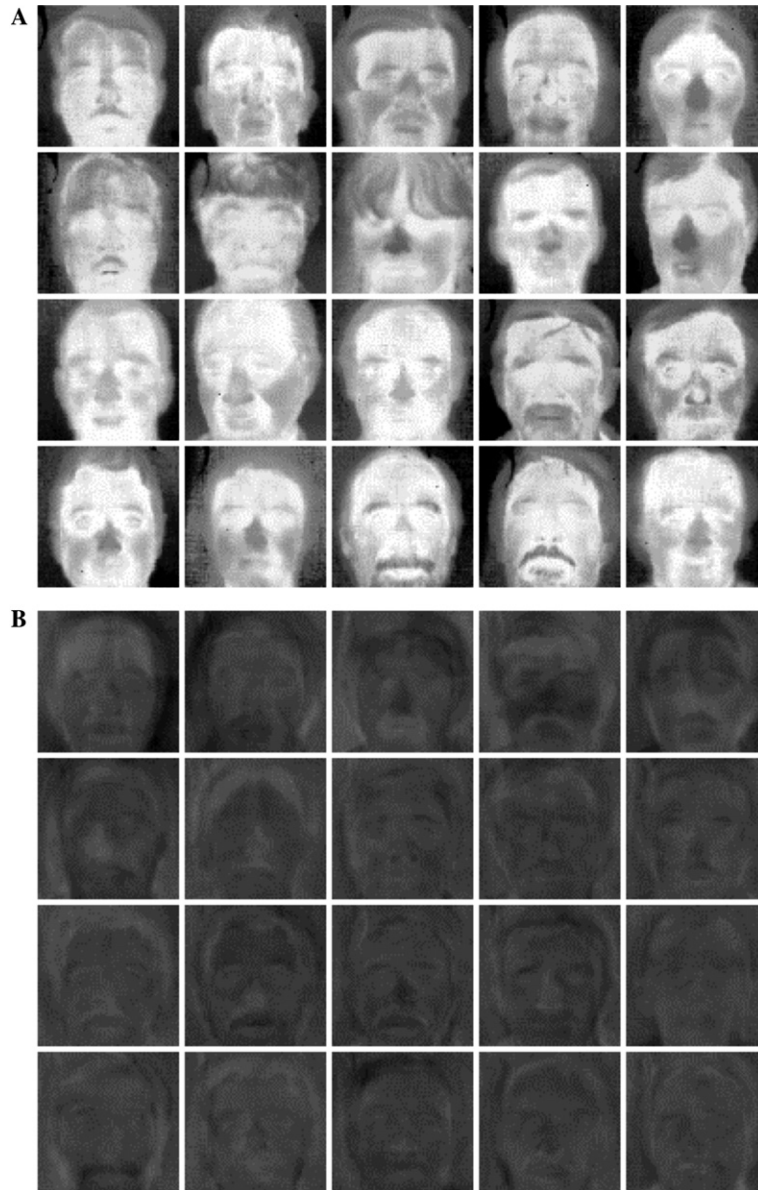
Fig. 5. Eigenfaces for thermal face images. (A) A set of thermal IR face images for training. (B) Eigenfaces created from the training set.

can also be obtained from thermal imagery such as the branching of particular blood vessels. Face metrics such as nose base width and face height above center are matched for identification, where the face center is defined as the midpoint of the line connecting the eyes. Template matching compares the areas in the

thermograms standardized in size in which the histogram is normalized. Templates often use the areas of the inner cheeks and between the eyes. An elemental shape refers to each nesting of thermal closed contours.

5.2.2. Anatomical structure

Another promising approaches include recognition of anatomical structures rather than thermal contours [139]. Some recent techniques in this category utilize symmetry waveforms and face codes. The IR vascular pattern provides a digital signature, which can be used for face recognition. Based on the hypothesis that the details of each person's asymmetries are unique, symmetry waveform technique exploits individual variations of asymmetries for identification. The image is first scaled and histogram normalized. To minimize head rotation effect, a vertical strip area within the outer corners of the eyes is analyzed. Thus these techniques are applicable only when both eyes can be seen. A value is assigned to the horizontal line of pixels in the face strip and the same values form the symmetry waveform. Symmetry analysis may be adequate for small databases. Symmetry waveforms can also be used for tracking of head rotation and for accurate alignment of images.

A face coding method utilizes 1D or 2D face bar codes derived from the degree of variation seen in different individual's symmetry waveforms. A bar code line is designated for each significant transition in the waveform. In the 2D bar coding technique, the face is divided into cells and each cell is compared to a library of face segments. The best matching segment is selected and its code used for that cell. Dividing the standardized face center into cells and indicating those in which minutiae are present can generate a simple binary code. The face-coding scheme will take account of head position, and allow for degraded accuracy of identification when only a partial image is seen. In the 2D bar coding scheme, a face is divided into cells and each cell is compared to a library of face segments. The best matching segment is selected and its code used for that cell. Dividing the standardized face center into cells can generate a simple binary code indicating the cells in which minutiae are present.

6. Fusion of imaging modalities for face recognition

Fusion exploits synergistic integration of images obtained from multiple sensors. Sensors are typically made sensitive to certain signals. For example, CCD cameras are designed for collecting visual signals, and IR sensors measures temperature distribution. Fusion of information from multiple sensors including visual, thermal, and 3D scanners can overcome the limitations of current face recognition techniques. Wilder et al. [140] compared the relative performances of the three face recognition algorithms for visible and IR images: matching pursuit filters [141,142], gray-scale projection, and eigenface techniques [109,143,144]. Fusion of global and local variables is discussed in [145]. Though visible and IR face recognition perform similarly across algorithms, the fusion of IR and visible imagery is a viable mean of enhancing performance. Correlation between thermal and visual facial imagery encouraged the security market to include uses, where no reference database of thermal images

exists. Fusing IR and visible imagery by linear pooling of the similarity scores from the individual modalities improved performance.

Face recognition from 3D range image data is another topic being actively studied by researchers. Sensitivity to variations in pose is another challenging problem in face recognition. As a face is inherently a 3D object, a good solution would be to use information about the 3D structure of a face. A range image contains the depth structure of the object. Range images can represent 3D shape explicitly and can compensate for the lack of depth information in a 2D image. The 3D shape is invariant to a change of color or reflectance properties due to changes in the ambient lighting. Creating a 3D face structure from multiple image views of a human face taken at different poses by appropriately morphing a generic 3D face is presented in [146–148]. In [149], the authors present a pose invariant face recognition system using a 3D deformable model. It is important to determine the value of the added information present in range data in terms of its effect on the accuracy of face recognition. A template-based recognition system involving descriptors based on curvature calculations made on range image data is presented in [150]. At each point on the surface, the magnitude and direction of the minimum and maximum normal curvatures are calculated. In the 3D domain, many researchers have handled the 3D face recognition problem using differential geometry tools for computing curvatures [151,152]. However, the computation of curvature is neither accurate nor reliable. In [153], the authors used point signature [154] to represent each face and treated the face recognition problem as a 3D recognition problem of non-rigid surfaces. The variety of gray-level information provided by different persons gives more detailed information for interpreting facial images, albeit its dependence on color and reflectance properties. Therefore, integrating 2D and 3D sensory information will be a key factor for achieving a significant improvement in performance over systems that rely solely on a single type of sensory data. Face recognition from both 2D and 3D facial images are studied in [155,156]. An efficient feature-based approach considering both shape and texture information is presented in [157].

Fusion of face images with other signal sources has been widely considered. A prototype integrated biometric system [158] performs personal identification by using both faces and fingerprints to overcome the limitations of face recognition and fingerprint verification systems. Fusion of face and speech data [159,160] enhances personal identification. The multi-sensor fusion problem for face recognition is discussed in [161].

7. Conclusion

Face recognition is an active research field due to its potential use in a wide variety of commercial and law enforcement applications including access control, security monitoring, and video surveillance. Unlike other biometric identification systems based on physiological characteristics, face recognition is a passive, non-intrusive system for verifying personal identity in a user-friendly way without having to interrupt user activity.

This paper reviews major efforts and advances in face recognition techniques focusing on two major sensing modalities: visual and thermal IR sensors. Various aspects of face recognition in the visible and infrared spectra are covered. The FERET evaluations had a significant impact on progress in the development of face recognition algorithms. Evaluation and benchmarking of numerous face recognition algorithms based on the FERET database and test methods are reviewed. Visual face recognition systems have demonstrated high performance under constrained conditions, such as frontal mug shot images and consistent lighting conditions. Performance of visual face recognition often degrades under uncontrolled illumination conditions as in outdoor surveillance applications. Visual face recognition has difficulties in detecting disguised faces, which can be critical in high-end security applications. IR face recognition techniques are useful for identifying faces under uncontrolled illumination conditions or for detecting disguises. Face recognition performance can be enhanced by the fusion of information obtained from different imaging sensors. Fusion of visual information obtained from reflectance intensity images and anatomical information from thermal IR images makes available information that cannot be obtained by processing visual or thermal images alone.

Table 1 summarizes the advantages and disadvantages of visual and thermal imaging methods as well as the fusion of visual and thermal imaging techniques for face recognition.

Table 1
Comparison of visual and thermal face recognition techniques

Imaging methods	Advantages	Disadvantages
Visual imaging	• Recognition algorithms well developed • Relatively easier to extract and locate facial features • Works well under controlled illumination conditions • Visual cameras are less expensive	• Poor performance with illumination variations and facial expressions • Difficult to segment out faces from cluttered scene • Useless in very low lighting • Unable to detect disguise
Thermal Imaging	• Face (and skin) detection, location, and segmentation are easier • Within-class variance smaller • Nearly invariant to illumination changes and facial expressions • Works even in total darkness • Useful for detecting disguises	• Glass blocks most of thermal energy • Not appropriate for recognition of vehicle occupants (speed, glass) • Thermal calibration is required Ambient temperature or activity level may change thermal characteristics • Low image resolution • Thermal cameras are expensive
Visual/Thermal Fusion	• Uses appearance and anatomical information of the face • Combines advantages of both imaging sensors • Wide application areas • Improves recognition accuracy	• Co-registration of visual and thermal images is required for data fusion • Imaging setup is more complicated • Higher computational requirements

Acknowledgment

References

[1] R.J. Baron, Mechanisms of human facial recognition, Int. J. Man Mach. Stud. 15 (2) (1981) 137–178.
[2] B. Miller, Vital signs of identity, IEEE Spectrum 31 (2) (1994) 22–30.
[3] A. Pentland, T. Choudhury, Face recognition for smart environments, IEEE Comp. 33 (2) (2000) 50–55.
[4] J.L. Center Jr., Practical application of facial recognition: automated facial recognition access control system, in: H. Wechsler, J.P. Phillips, V. Bruce, F.F. Soulie, T.S. Huang (Eds.), Face Recognition: From Theory to Applications, Springer-Verlag, Berlin, Germany, 1998, pp. 402–408.
[5] I. Craw, N. Costen, T. Kato, S. Akamatsu, How should we represent faces for automatic recognition? IEEE Trans. Patt. Anal. Mach. Intell. 21 (8) (1999) 725–736.
[6] A.M. Burton, V. Bruce, P.J.B. Hancock, From pixels to people: a model of familiar face recognition, Cognitive Sci. 23 (1999) 1–31.
[7] Y. Adini, Y. Moses, S. Ullman, Face recognition: the problem of compensating for changes in illumination direction, IEEE Trans. Patt. Anal. Mach. Intell. 19 (7) (1997) 721–732.
[8] Y.I. Tian, T. Kanade, J.F. Cohn, Recognizing action units for facial expression analysis, IEEE Trans. Patt. Anal. Mach. Intell. 23 (2) (2001) 97–115.
[9] J. Ben-Arie, D. Nandy, A volumetric/iconic frequency domain representation for objects with application for pose invariant face recognition, IEEE Trans. Patt. Anal. Mach. Intell. 20 (5) (1998) 449–457.
[10] Y. Yoshitomi, T. Miyaura, S. Tomita, S. Kimura, Face identification using thermal image processing, Proc. IEEE Int. Workshop Robot Hum. Commun. (1997) 374–379.
[11] D. Socolinsky, A. Selinger, J. Neuheisel, Face recognition with visible and thermal infrared imagery, Comput. Vis. Image Und. 91 (2003) 72–114.
[12] L. Wolff, D. Socolinsky, C. Eveland, Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery, Proc. IEEE Workshop Comput. Vis. Beyond Vis. Spectrum: Beyond Vis. Spectrum: Method Appl. (2001).
[13] J. Dowdall, I. Pavlidis, G. Bebis, A face detection method based on multi-band feature extraction in the near-IR spectrum, Proc. IEEE Workshop Comput. Vis. Beyond Vis. Spectrum: Method Appl. (2001).
[14] J. Dowdall, I. Pavlidis, G. Bebis, Face detection in the near-IR spectrum, Image Vis. Comput. 21 (7) (2003) 565–578.
[15] I. Pavlidis, P. Symosek, The imaging issue in an automatic face/disguise detection system, Proc. IEEE Workshop Comput. Vis. Beyond Vis. Spectrum: Methods Appl. (2000) 15–24.
[16] K.R. Foster, Thermographic detection of breast cancer, IEEE Eng. Med. Biol. Mag. 17 (6) (1998) 10–14.
[17] K.C. Lin, W.L. Chen, S.C. Chen, F.S. Wang, Diagnostic technique for classifying the quality of circuit boards using infrared thermal image, Proc. IFSA World Congr. NAFIPS Int. Conf. 1 (2001) 464–469.
[18] J. Wan, X. Li, PCB infrared thermal imaging diagnosis using support vector classifier, Proc. World Congr. Intell. Control Automat. 4 (2002) 2718–2722.
[19] M.V. Shirvaikar, M.M. Trivedi, A neural network filter to detect small targets in high clutter backgrounds, IEEE Trans. Neural Netw. 6 (1) (1995) 252–257.
[20] I. Pavlidis, J. Levine, P. Baukol, Thermal image analysis for anxiety detection, Proc. IEEE Int. Conf. Image Process. 2 (2001) 315–318.

[21] Y. Yoshitomi, N. Miyawaki, S. Tomita, S. Kimura, Facial expression recognition using thermal image processing and neural network, Proc. IEEE Int. Workshop Robot Hum. Commun. (1997) 380–385.

[22] A. Samal, P.A. Iyengar, Automatic recognition and analysis of human faces and facial expressions: a survey, Patt. Recogn. 25 (1) (1992) 65–77.

[23] D. Valentin, H. Abdi, A.J. O'Toole, G.W. Cottrell, Connectionist models of face processing: a survey, Patt. Recogn. 27 (9) (1994) 1209–1230.

[24] R. Chellappa, C.L. Wilson, S. Sirohey, Human and machine recognition of faces: a survey, Proc. IEEE 83 (5) (1995) 705–740.

[25] J. Zhang, Y. Yan, M. Lades, Face recognition: eigenface, elastic matching, and neural nets, Proc. IEEE 85 (9) (1997) 1423–1435.

[26] J. Daugman, Face and gesture recognition: overview, IEEE Trans. Patt. Anal. Mach. Intell. 19 (7) (1997) 675–676.

[27] D. Socolinsky, A. Selinger, A comparative analysis of face recognition performance with visible and thermal infrared imagery, Proc. Int. Conf. Patt. Recogn. 4 (2002) 217–222.

[28] K.K. Sung, T. Poggio, Learning human face detection in cluttered scenes, in: G. Goos, J. Hartmonis, J. van Leeuwen (Eds.), Computer Analysis of Images and Patterns, Springer-Verlag, New York, 1995, pp. 432–439.

[29] H. Wang, S.F. Chang, A highly efficient system for automatic face region detection in MPEG video, IEEE Trans. Circ. Syst. Video Technol. 7 (4) (1997) 615–628.

[30] S. Satoh, Y. Nakamura, T. Kanade, Name-It: naming and detecting faces in news videos, IEEE Multimedia 6 (1) (1999) 22–35.

[31] D. Chai, K.N. Ngan, Face segmentation using skin-color map in videophone applications, IEEE Trans. Circ. Syst. Video Technol. 9 (4) (1999) 551–564.

[32] C. Garcia, G. Tziritas, Face detection using quantized skin color regions merging and wavelet packet analysis, IEEE Trans. Multimedia 1 (3) (1999) 264–277.

[33] R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain, Face detection in color images, IEEE Trans. Patt. Anal. Mach. Intell. 24 (5) (2002) 696–706.

[34] H. Wu, Q. Chen, M. Yachida, Face detection from color images using a fuzzy pattern matching method, IEEE Trans. Patt. Anal. Mach. Intell. 21 (6) (1999) 557–563.

[35] E. Hjelmas, B.K. Low, Face detection: a survey, Comput. Vis. Image Und. 83 (3) (2001) 236–274.

[36] M.H. Yang, D.J. Kriegman, N. Ahuja, Detecting faces in images: a survey, IEEE Trans. Patt. Anal. Mach. Intell. 24 (1) (2002) 34–58.

[37] K.M. Lam, H. Yan, Locating and extracting the eye in human face images, Patt. Recogn. 29 (5) (1996) 771–779.

[38] W. Huang, Q. Sun, C.P. Lam, J.K. Wu, A robust approach to face and eyes detection from images with cluttered background, Proc. Int. Conf. Patt. Recogn. 1 (1998) 110–114.

[39] F. Smeraldi, O. Carmona, J. Bigün, Saccadic search with Gabor features applied to eye detection and real-time head tracking, Image Vis. Comput. 18 (4) (2000) 323–329.

[40] C. Wu, C. Liu, H.Y. Shum, Y.Q. Xu, Z. Zhang, Automatic eyeglasses removal from face images, IEEE Trans. Patt. Anal. Mach. Intell. 26 (3) (2004) 322–336.

[41] K.K. Sung, T. Poggio, Example-based learning for view-based human face detection, IEEE Trans. Patt. Anal. Mach. Intell. 20 (1) (1998) 39–51.

[42] A.J. Colmenarez, T.S. Huang, Face detection with information-based maximum discrimination, Proc. Conf. Comput. Vis. Patt. Recogn. (1997) 782–787.

[43] K.C. Yow, R. Cipolla, Feature-based human face detection, Image Vis. Comput. 15 (9) (1997) 713–735.

[44] H.A. Rowley, S. Baluja, T. Kanade, Neural network-based face detection, IEEE Trans. Patt. Anal. Mach. Intell. 20 (1) (1998) 23–38.

[45] R. Feraud, O.J. Bernier, J.E. Viallet, M. Collobert, A fast and accurate face detector based on neural networks, IEEE Trans. Patt. Anal. Mach. Intell. 23 (1) (2001) 42–53.

[46] E. Osuna, R. Freund, F. Girosi, Training support vector machines: an application to face detection, Proc. IEEE Conf. Comput. Vis. Patt. Recogn. (1997) 130–136.

[47] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, Proc. Conf. Comput. Vis. Patt. Recogn. 1 (2001) 511–518.

[48] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, in: Computational Learning Theory: Eurocolt'95, Springer-Verlag, Berlin, Germany, 1995, pp. 23–37.

[49] R. Lienhart, J. Maydt, An extended set of Haar-like features for rapid object detection, Proc. Int. Conf. Image Process. 1 (2002) 900–903.

[50] R.C. Verma, C. Schmid, K. Mikolajczyk, Face detection and tracking in a video by propagating detection probabilities, IEEE Trans. Patt. Anal. Mach. Intell. 25 (10) (2003) 1215–1228.

[51] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Patt. Anal. Mach. Intell. 15 (10) (1993) 1042–1052.

[52] A.J. Goldstein, L.D. Harmon, A.B. Lesk, Identification of human faces, Proc. IEEE 59 (5) (1971) 748–760.

[53] M.S. Kamel, H.C. Shen, A.K.C. Wong, T.M. Hong, R.I. Campeanu, Face recognition using perspective invariant features, Patt. Recogn. Lett. 15 (9) (1994) 877–883.

[54] K. Beyer, J. Goldstein, R. Ramakrishnan, U. Shaft, When is 'nearest neighbor' meaningful? Proc. Int. Conf. Database Theory (1999) 217–235.

[55] L. Sirovich, M. Kirby, Low-dimensional procedure for the characterization of human faces, J. Opt. Soc. Am. 4 (3) (1987) 519–524.

[56] K.M. Lam, H. Yan, An analytic-to-holistic approach for face recognition based on a single frontal view, IEEE Trans. Patt. Anal. Mach. Intell. 20 (7) (1998) 673–686.

[57] P.J. Phillips, P. Rauss, The face recognition technology (FERET) program, in: Proc. Office of National Drug Control Policy, CTAC Int'l Technology Symp. 1997, pp. 8–11.

[58] B. Moghaddam, C. Nastar, A. Pentland, A Bayesian similarity measure for direct image matching, Proc. Int. Conf. Patt. Recogn. (1996).

[59] L. Wiskott, J.M. Fellous, N. Krüger, C. von der Malsburg, Face recognition by elastic bunch graph matching, IEEE Trans. Patt. Anal. Mach. Intell. 19 (7) (1997) 775–779.

[60] W. Zhao, R. Chellappa, A. Krishnaswamy, Discriminant analysis of principal components for face recognition, Proc. IEEE Conf. Comput. Vis. Patt. Recogn. (1998) 336–341.

[61] I.T. Jolliffe, Principal Component Analysis, second ed., Springer-Verlag, Berlin, Germany, 2002.

[62] P.J. Phillips, H. Wechsler, J. Juang, P.J. Rauss, The FERET database and evaluation procedure for face recognition algorithms, Image Vis. Comput. 16 (5) (1998) 295–306.

[63] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Trans. Patt. Anal. Mach. Intell. 22 (10) (2000) 1090–1104.

[64] P.S. Penev, J.J. Atick, Local feature analysis: a general statistical theory for object representation, Netw. Comput. Neural Syst. 7 (3) (1996) 477–500.

[65] P.S. Penev, Local feature analysis: a statistical theory for information representation and transmission, Ph.D. Thesis, The Rockefeller University, 1998.

[66] P.S. Penev, Dimensionality reduction by sparsification in a local-features representation of human faces, Technical Report, The Rockefeller University, 1999.

[67] H. Moon, P.J. Phillips, Analysis of PCA-based face recognition algorithms, in: K.W. Bowyer, P.J. Phillips (Eds.), Empirical Evaluation Techniques in Computer Vision, IEEE Computer Society Press, Los Alamitos, CA, 1998, pp. 57–71.

[68] D.M. Blackburn, M. Bone, P.J. Phillips, Facial Recognition Vendor Test 2000: Evaluation Report, 2001.

[69] P.J. Phillips, P. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, M. Bone, Face Recognition Vendor Test 2002 : Evaluation Report, 2003.

[70] V. Blanz, T. Vetter, A morphable model for the synthesis of 3D faces, Proc. Int. Conf. Comput. Graphics Interact. Tech. (SIGGRAPH) (1999) 187–194.

[71] V. Blanz, S. Romdhami, T. Vetter, Face identification across different poses and illuminations with a 3D morphable model, Proc. Int. Conf. Autom. Face Gesture Recogn. (2002) 202–207.

[72] T. Kanade, Picture processing by computer complex and recognition of human faces, Ph.D. Dissertation, Kyoto University, 1973.

[73] I.J. Cox, J. Ghosn, P.N. Yianilos, Feature-based face recognition using mixture-distance, Proc. Int. Conf. Comput. Vis. Patt. Recogn. (1996) 209–216.

[74] B.S. Manjunath, R. Chellappa, C. von der Malsburg, A feature based approach to face recognition, Proc. IEEE Conf. Comput. Vis. Patt. Recogn. (1992) 373–378.

[75] S. Tamura, H. Kawa, H. Mitsumoto, Male/female identification from 8×6 very low resolution face images by neural network, Patt. Recogn. 29 (2) (1996) 331–335.

[76] A.M. Martinez, A.C. Kak, PCA versus LDA, IEEE Trans. Patt. Anal. Mach. Intell. 23 (2) (2001) 228–233.

[77] R. Cappelli, D. Maio, D. Maltoni, Multispace KL for pattern representation and classification, IEEE Trans. Patt. Anal. Mach. Intell. 23 (9) (2001) 977–996.

[78] R.A. Fisher, The use of multiple measures in taxonomic problems, Ann. Eugenic. 7 (1936) 179–188.

[79] R. Lotlikar, R. Kothari, Fractional-step dimensionality reduction, IEEE Trans. Patt. Anal. Mach. Intell. 22 (6) (2000) 623–627.

[80] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Trans. Patt. Anal. Mach. Intell. 19 (7) (1997) 711–720.

[81] C. Liu, H. Wechsler, A shape- and texture-based enhanced Fisher classifier for face recognition, IEEE Trans. Image Process. 10 (4) (2001) 598–608.

[82] D.L. Swets, J.J. Weng, Using discriminant eigenfeatures for image retrieval, IEEE Trans. Patt. Anal. Mach. Intell. 18 (8) (1996) 831–836.

[83] K. Etemad, R. Chellappa, Discriminant analysis for recognition of human face images, J. Opt. Soc. Am. A 14 (8) (1997) 1724–1733.

[84] L.F. Chen, H.Y.M. Liao, M.T. Ko, J.C. Lin, G.J. Yu, A new LDA-based face recognition system which can solve the small sample size problem, Patt. Recogn. 33 (10) (2000) 1713–1726.

[85] H. Yu, J. Yang, A direct LDA algorithm for high-dimensional data with application to face recognition, Patt. Recogn. 34 (10) (2001) 2067–2070.

[86] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Face recognition using LDA-based algorithms, IEEE Trans. Neural Netw. 14 (1) (2003) 195–200.

[87] C. Liu, H. Wechsler, Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition, IEEE Trans. Image Process. 11 (4) (2002) 467–476.

[88] B. Schölkopf, A. Smola, K.R. Müller, Nonlinear component analysis as a kernel eigenvalue problem, Neural Comput. 10 (1998) 1299–1319.

[89] K.I. Kim, K. Jung, H.J. Kim, Face recognition using kernel principal component analysis, IEEE Signal Process. Lett. 9 (2) (2002) 40–42.

[90] G. Baudat, F. Anouar, Generalized discriminant analysis using a kernel approach, Neural Comput. 12 (10) (2000) 2385–2404.

[91] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Face recognition using kernel direct discriminant analysis algorithms, IEEE Trans. Neural Netw. 14 (1) (2003) 117–126.

[92] M. Bartlett, T. Sejnowski, Independent components of face images: a representation for face recognition, Proc. 4th Annual J. Symp. Neural Computation, 1997.

[93] M. Bartlett, J. Movellan, T. Sejnowski, Face recognition by independent component analysis, IEEE Trans. Neural Netw. 13 (6) (2002) 1450–1464.

[94] S.Y. Kung, J.S. Taur, Decision-based neural networks with signal/image classification applications, IEEE Trans. Neural Netw. 6 (1) (1995) 170–181.

[95] S.H. Lin, S.Y. Kung, L.J. Lin, Face recognition/detection by probabilistic decision-based neural network, IEEE Trans. Neural Netw. 8 (1) (1997) 114–132.

[96] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.

[97] V.N. Vapnik, Statistical Learning Theory, John Wiley, New York, 1998.

[98] B. Schölkopf, Support Vector Learning, Oldenbourg-Verlag, Munich, Germany, 1997.

[99] P.J. Phillips, Support vector machines applied to face recognition, Proc. Conf. Adv. Neural Inf. Process. Syst. II (1998) 803–809.

[100] G. Guo, S.Z. Li, K. Chan, Face recognition by support vector machines, Proc. IEEE Int. Conf. Autom. Face Gesture Recogn. (2000) 196–201.

[101] B. Heisele, P. Ho, T. Poggio, Face recognition with support vector machines: global versus component-based approach, Proc. Int. Conf. Comput. Vis. 2 (2001) 688–694.

[102] K. Jonsson, J. Kittler, Y.P. Li, J. Matas, Support vector machines for face authentication, Image Vis. Comput. 20 (5–6) (2002) 369–375.

[103] B. Moghaddam, M. Yang, Gender classification with support vector machines, Proc. IEEE Int. Conf. Autom. Face Gesture Recogn. (2000) 306–311.

[104] Y. Gao, M.K.H. Leung, Face recognition using line edge map, IEEE Trans. Patt. Anal. Mach. Intell. 24 (6) (2002) 764–779.

[105] S.Z. Li, J. Lu, Face recognition using the nearest feature line method, IEEE Trans. Neural Netw. 10 (2) (1999) 439–443.

[106] B. Takacs, Comparing face images using the modified Hausdorff distance, Patt. Recogn. 31 (12) (1998) 1873–1881.

[107] K.H. Lin, K.M. Lam, W.C. Siu, Spatially eigen-weighted Hausdorff distances for human face recognition, Patt. Recogn. 36 (8) (2003) 1827–1834.

[108] M. Kirby, L. Sirovich, Application of the Karhunen–Loeve procedure for the characterization of human faces, IEEE Trans. Patt. Anal. Mach. Intell. 12 (1) (1990) 103–108.

[109] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cogn. Neurosci. 3 (1) (1991) 71–86.

[110] M. Turk, A. Pentland, Face recognition using eigenfaces, Proc. IEEE Conf. Comput. Vis. Patt. Recogn. (1991) 586–591.

[111] L. Zhao, Y.H. Yang, Theoretical analysis of illumination in PCA-based vision systems, Patt. Recogn. 32 (4) (1999) 547–564.

[112] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, Proc. IEEE Int. Conf. Comput. Vis. Patt. Recogn. (1994) 84–91.

[113] S. Shan, B. Cao, W. Gao, D. Zhao, Extended Fisherface for face recognition from a single example image per person, Proc. IEEE Int. Symp. Circ. Syst. 2 (2002) 81–84.

[114] M. Girolami, Advances in Independent Component Analysis, Springer-Verlag, Berlin, Germany, 2000.

[115] T.W. Lee, Independent Component Analysis: Theory and Applications, Kluwer, Boston, MA, 1998.

[116] C. Liu, H. Wechsler, Comparative assessment of independent component analysis (ICA) for face recognition, Proc. Int. Conf. Audio- and Video-based Biometric Person Authentication, 1999.

[117] M. Bartlett, in: Face Image Analysis by Unsupervised Learning, vol. 612, Kluwer, Boston, MA, 2001.

[118] M. Lades, J.C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Würtz, W. Konen, Distortion invariant object recognition in the dynamic link architecture, IEEE Trans. Comput. 42 (3) (1993) 300–311.

[119] L. Wiskott, C. von der Malsburg, Recognizing faces by dynamic link matching, Neuroimage 4 (3) (1996) S14–S18.

[120] C. Kotropoulos, A. Tefas, I. Pitas, Frontal face authentication using morphological elastic graph matching, IEEE Trans. Image Process. 9 (4) (2000) 555–560.

[121] A. Tefas, C. Kotropoulos, I. Pitas, Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication, IEEE Trans. Patt. Anal. Mach. Intell. 23 (7) (2001) 735–746.

[122] T.J. Stonham, Practical face recognition and verification with WISARD, in: H.D. Ellis, M.A. Jeeves, F. Newcombe, A. Young (Eds.), Aspects of Face Processing, Kluwer, Boston, MA, 1986, pp. 426–441.

[123] S. Lawrence, C.L. Giles, A.C. Tsoi, A.D. Back, Face recognition: a convolutional neural-network approach, IEEE Trans. Neural Netw. 8 (1) (1997) 98–113.

[124] S. Gutta, J.R.J. Huang, P. Jonathon, H. Wechsler, Mixture of experts for classification of gender, ethnic origin, and pose of human faces, IEEE Trans. Neural Netw. 11 (4) (2000) 948–960.

[125] D. Zhang, H. Peng, J. Zhou, S.K. Pal, A novel face recognition system using hybrid neural and dual eigenspaces methods, IEEE Trans. Syst. Man Cybernetics—Part A 32 (6) (2002) 787–793.

[126] M.J. Er, S. Wu, J. Lu, H.L. Toh, Face recognition with radial basis function (RBF) neural networks, IEEE Trans. Neural Netw. 13 (3) (2002) 697–710.

[127] F. Samaria, S. Young, HMM-based architecture for face identification, Image Vis. Comput. 12 (8) (1994).

[128] A. Nefian, M.H. Hayes III, An embedded HMM for face detection and recognition, Proc. Int. Conf. Acoustics, Speech, and Signal Process. 6 (1999).

[129] A. Nefian, Embedded Bayesian networks for face recognition, Proc. IEEE Int. Conf. Multimedia Expo. 2 (2002) 133–136.

[130] A. Nefian, M.H. Hayes III, Hidden Markov models for face recognition, speech, and signal processing, Proc. Int. Conf. Acoustics 5 (1998) 2721–2724.

[131] V.V. Kohir, U.B. Desai, Face recognition, Proc. IEEE Int. Symp. Circ. Syst. (2000) 309–312.

[132] H. Othman, T. Aboulnasr, A separable low complexity 2D HMM with application to face recognition, IEEE Trans. Patt. Anal. Mach. Intell. 25 (10) (2003) 1229–1238.

[133] X. Liu, T. Chen, Video-based face recognition using adaptive hidden Markov models, Proc. IEEE Conf. Comput. Vis. Patt. Recogn. 1 (2003) 340–345.

[134] C. Eveland, D. Socolinsky, L. Wolff, Tracking human faces in infrared video, Proc. Proc. IEEE Workshop Comput. Vis. Beyond Vis. Spectrum: Method Appl. (2001).

[135] I. Pavlidis, P. Symosek, B. Fritz, M. Bazakos, N. Papanikolopoulos, Automatic detection of vehicle occupants—the imaging problem and its solution, Mach. Vis. Appl. 11 (6) (2000) 313–320.

[136] F. Prokoski, Disguise detection and identification using infrared imagery, in: A.S. Hecht (Ed.), Proceedings of SPIE, Optics, and Images in Law Enforcement II, 1982, pp. 27–31.

[137] A. Selinger, D. Socolinsky, Appearance-based facial recognition using visible and thermal imagery: a comparative study, Technical Report 02-01, Equinox Corporation, 2002.

[138] D. Socolinsky, L. Wolff, J. Neuheisel, C. Eveland, Illumination invariant face recognition using thermal infrared imagery, Proc. IEEE Workshop Comput. Vis. Beyond Vis. Spectrum: Method Appl. 1 (2001) 527–534.

[139] F. Prokoski, History, current status, and future of infrared identification, Proc. IEEE Workshop Comput. Vis. Beyond Vis. Spectrum: Methods Appl. (2000) 5–14.

[140] J. Wilder, P.J. Phillips, C.H. Jiang, S. Wiener, Comparison of visible and infra-red imagery for face recognition, Proc. Int. Conf. Autom. Face Gesture Recogn. (1996) 182–187.

[141] P.J. Phillips, Matching pursuit filters applied to face identification, IEEE Trans. Image Process. 7 (8) (1998) 1150–1164.

[142] C. Liu, H. Wechsler, Evolutionary pursuit and its application to face recognition, IEEE Trans. Patt. Anal. Mach. Intell. 22 (6) (2000) 570–582.

[143] B. Moghaddam, T. Jebara, A. Pentland, Bayesian face recognition, Patt. Recogn. 33 (11) (2000) 1771–1782.

[144] B. Moghaddam, A. Pentland, Beyond Euclidean eigenspaces: Bayesian matching for visual recognition, in: Face Recognition: From Theories to Applications, Springer-Verlag, Berlin, Germany, 1998.

[145] Y. Fang, T. Tan, Y. Wang, Fusion of global and local features for face verification, Proc. Int. Conf. Patt. Recogn. (2002) 382–385.

[146] C. Zhang, F.S. Cohen, 3-D face structure extraction and recognition from images using 3-D morphing and distance mapping, IEEE Trans. Image Process. 11 (11) (2002) 1249–1259.

[147] D. Nandy, J. Ben-Arie, Shape from recognition: a novel approach for 3-D face shape recovery, IEEE Trans. Image Process. 10 (2) (2001) 206–217.

[148] M.J. Clarkson, D. Rueckert, D.L.G. Hill, D.J. Hawkes, Using photo-consistency to register 2D optical images of the human face to a 3D surface model, IEEE Trans. Patt. Anal. Mach. Intell. 23 (11) (2001) 1266–1280.

[149] M.W. Lee, S. Ranganath, Pose-invariant face recognition using a 3D deformable model, Patt. Recogn. 36 (8) (2003) 1835–1846.

[150] G. Gordon, Face recognition based on depth maps and surface curvature, in: SPIE proceedings, geometricm methods in computer vision, vol. 1570, 1991, pp. 234–247.

[151] H.T. Tanaka, M. Ikeda, H. Chiaki, Curvature-based face surface recognition using spherical correlation, Proc. IEEE Int. Conf. Autom. Face Gesture Recogn. (1998) 372–377.

[152] B. Achermann, X. Jiang, H. Bunke, Face recognition using range images, Proc. Int. Conf. Virtual Syst. Multimedia (1997) 129–136.

[153] C.S. Chua, F. Han, Y.K. Ho, 3D human face recognition using point signature, Proc. IEEE Int. Conf. Autom. Face Gesture Recogn. (2000) 233–238.

[154] C.S. Chua, R. Jarvis, Point signature: a new representation for 3D object recognition, Int. J. Comput. Vis. 25 (1) (1997) 63–85.

[155] S. Tsutsumi, S. Kikuchi, M. Nakajiama, Face identification using a 3D gray-scale image—a method for lessening restrictions on facial directions, Proc. IEEE Int. Conf. Autom. Face Gesture Recogn. (1998) 306–311.

[156] C. Beumier, M. Acheroy, Face verification from 3D and gray level clues, Patt. Recogn. Lett. 22 (12) (2001) 1321–1329.

[157] Y. Wang, C.S. Chua, Y.K. Ho, Facial feature detection and face recognition from, Patt. Recogn. Lett. 23 (10) (2002) 1191–1202.

[158] L. Hong, A. Jain, Integrating faces and fingerprints for personal identification, IEEE Trans. Patt. Anal. Mach. Intell. 20 (12) (1998) 1295–1307.

[159] S. Ben-Yacoub, Y. Abdeljaoued, E. Mayoraz, Fusion of face and speech data for person identity verification, IEEE Trans. Neural Netw. 10 (5) (1999) 1065–1074.

[160] N. Poh, J. Korczak, Hybrid biometric person authentication using face and voice features, Proc. Int. Conf. Audio- and Video-Based Biometric Person Authentication (2001) 348–351.

[161] F. Roli, J. Kittler, G. Fumera, D. Muntoni, An experimental comparison of classifier fusion methods for multimodal personal identity verification systems, Proc. Int. Workshop Mult. Classifier Syst. (2002) 325–336.