David Zhang
Hong Kong Polytechnic University, HONG KONG
Wangmeng Zuo
Harbin Institute of Technology, CHINA

Computational Intelligence-Based Biometric Technologies

Abstract: Computational intelligence (CI) technologies are robust, can be successfully applied to complex problems, are efficiently adaptive, and usually have a parallel computational architecture. For those reasons they have been proved to be effective and efficient in biometric feature extraction and biometric matching tasks, sometimes used in combination with traditional methods. In this article, we briefly survey two kinds of major applications of CI in biometric technologies, CI-based feature extraction and CI-based biometric matching. Varieties of evolutionary computation and neural networks techniques have been successfully applied to biometric data representation and dimensionality reduction. CI-based methods, including neural network and fuzzy technologies, have also been extensively investigated for biometric matching. CI-based biometric technologies are powerful when used in the representation and recognition of incomplete biometric data, discriminative feature extraction, biometric matching, and online template updating, and promise to have an important role in the future development of biometric technologies.

1556-603X/07/$25.00©2007IEEE

1. Introduction

Automatic identification and authentication systems that make use of biometric data, such as, distinctive anatomical and behavioral characteristics, are becoming ever more widely used for access control, surveillance, computer security, and in law enforcement. Several governments are now using or will soon be using biometric technology. The U.S. INSPASS immigration card and the Hong Kong ID card, for example, both store biometric features for authentication. Measured against alternative approaches, current automatic biometric systems are reliable, efficient, convenient and secure, yet at each module of automatic biometric authentication systems (Figure 1) there remains room for improvement. In the acquisition and pre-processing module, biometric data is sometimes noisy, partially occluded, or inaccurately located. In the feature extraction and recognition module, biometric authentication systems cannot yet completely eliminate or counter the adverse effects of limited training samples and within-class variations. In applications, current biometric systems have only a limited ability in adapting to different situations. All of these and other issues can be very effectively dealt with by making use of computational intelligence (CI) technologies including neural networks, fuzzy sets, and evolutionary computation, either alone or in combination with more traditional techniques.

CI techniques exhibit four particular features which make them useful for the purposes of biometric identification and authentication. First, they are *adaptive*. This is useful because, while it is a primary assumption of identification/authentication that features should be stable, in reality they do change over time. It is very complicated to model such changes but evolutionary computation and neural networks, for example, could solve this problem by allowing biometric feature extraction and matching to be updated adaptively. Second, CI approaches allow *uncertainty modeling*. Traditionally, complexity and uncertainty should be addressed in the modeling of the within-class variations between biometric data. Fuzzy technology, or probabilistic fuzzy technology, however, offers various ways to model uncertainty in these highly complex variations of biometric characteristics. Third, CI approaches are *robust*. Hopfield associative memory networks, which utilize local interactions to achieve a content-addressable memory function, can reliably retrieve patterns from memory even when incomplete or corrupted samples are presented. Finally, computational intelligence technologies usually have a *parallel computational architecture* and this means faster computation on highly complex systems.

It is beyond the scope of this article to describe all existing and potential applications of CI methodologies in biometric technologies. In the following, we will introduce and briefly describe two major CI applications, CI-based biometric feature extraction and CI-based biometric matching.

2. Computational Intelligence-Based Biometric Feature Extraction

Feature extraction, a basic part of any biometric system, usually has two main stages, biometric data representation and dimensionality reduction. Biometric data representation involves using signal and image processing technologies to extract a set of salient or discriminatory features. The feature dimensionality is then reduced and the redundancy and correlation between different features is eliminated using dimensionality reduction techniques.

2.1. Intelligent Biometric Data Representation

Biometric data is usually represented in one of three forms: one-dimensional waveforms, two-dimensional images, or three-dimensional geometric data. To date, various signal processing technologies, especially image processing techniques, have been extensively investigated and successfully applied to biometric data representation, however, currently, no biometric data representation scheme exists which can extract all invariance and discriminative information from biometric data.

Models of real-world biometric data are often highly nonlinear and very complex, yet most representation schemes are based on simplistic mathematical models that are linear, Gaussian, or least-square, and fail to model the richness and complexity of information in a biometric characteristic. Further, factors such as ageing and abrasion of the biometric characteristic, or changes in the sensor or capture environment,

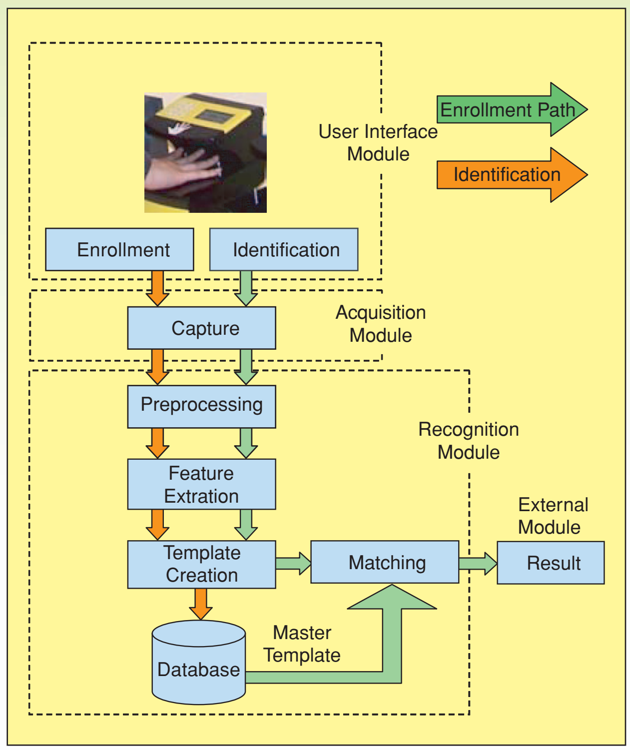

FIGURE 1 A generic biometrics system. In the enrollment stage, the biometric characteristic of an individual is first captured in the acquisition module, and then processed and stored in the prototype database. Similarly, in identification stage, biometric characteristic is captured and identified, and then the system makes a real-time respond according to recognition result.

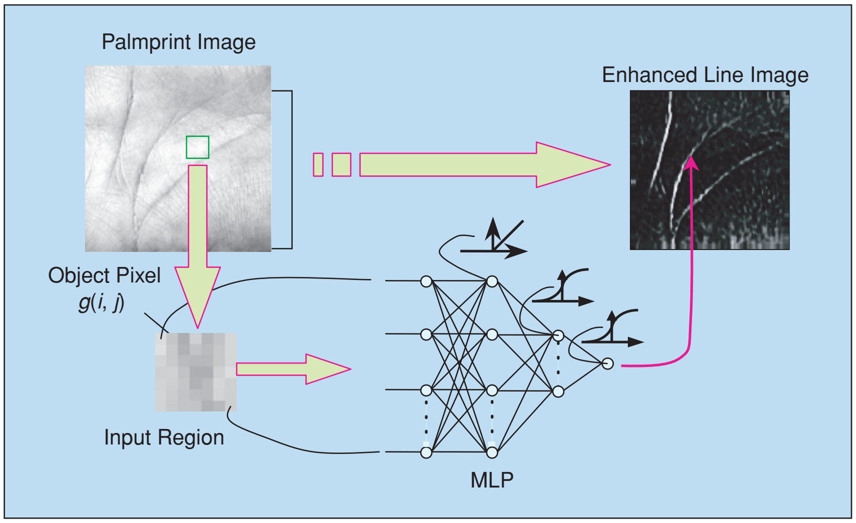

FIGURE 2 The architecture of the MLP-based neural palm-line enhancer.

can affect the quality of captured biometric data and thus in turn affect the performance of biometric systems.

In general, CI technologies can be used to represent biometric data by extracting as much discriminatory information as possible from the input biometric data. Unlike traditional signal processing techniques, CI-based methods can adaptively model a complex biometric characteristic, thereby obviating the need to make many assumptions about the precise mathematical model of the biometric data. In the following we present an overview of three major CI-based biometric data representation methods.

Multilayer Perceptron (MLP)

MLP is one of the most extensively used neural networks. Given two sets of data, input/output pairs, MLP is able to develop a specific nonlinear mapping by adjusting the network weights by using a learning algorithm. It has been demonstrated that a two-layer MLP will adequately approximate any nonlinear mapping. The most used MLP training method is the *backpropagation* (BP) algorithm, where a steepest descent gradient approach and a chain-rule are adopted for back-propagated error correction from the output layer.

Considerable efforts have been put into improving the speed of convergence, generalization performance, and the discriminative ability of MLP. To accelerate the BP algorithm, several heuristic rules have been proposed to adjust the learning rates or modify error functions [3]. Acceleration of training MLP can also be achieved by the use of other modifications to the standard BP algorithm, such as conjugate gradient BP, recursive least-square-based BP, and the Levenberg-Marquardt algorithm. To verify the generalization ability of MLP, the *independent validation* method can be used by dividing the available data set into a number of subsets for training, validation and testing [4]. To improve the discriminative capability of MLP when applied to a classification task, Juan and Katagiri proposed a discriminative MLP learning rule which is more suitable for pattern classification tasks [5].

Applications: MLP is capable of adaptively approximating the function of any linear or nonlinear filter, and thus is very competitive when applied to biometric data representation such as the extraction of efficient discriminative features, filtering, and enhancement of biometric signals. MLPs have been widely used in varieties of intelligent signal processing tasks, such as signal noise reduction [6], [7], edge enhancement [8], and singular point detection, which are very valuable in the pre-processing and feature extraction of biometric data. In biometric data presentation, MLP has been successfully used in skin segmentation, face detection and facial feature location [9]–[11]. For example, Figure 2 shows the architecture of an MLP-based neural palm-line enhancer. This enhancer has a number of advantages over classical line detectors in that it can be trained to be noise-robust and palm-line-specific, and thus is more suitable for palm-line enhancement.

Associative Memory Networks

Associative memory networks include *linear associative memory* and *Hopfield associative memory*. Linear associative memory is an effective single-layer network for the retrieval and reduction of information. Given a key input pattern $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K]$ and the corresponding output $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K]$, associative memory learns the memory matrix $\mathbf{W}$ to map the key input $\mathbf{x}_i$ to the memorized output $\hat{\mathbf{y}}_i$. There are a number of ways to estimate the memory matrix. One estimate of the memory matrix $\mathbf{W}$ is the sum of the outer product matrices from pairs of key input and memorized patterns

$$\mathbf{W} = \sum_{k=1}^{K} \mathbf{y}_k \mathbf{x}_k^T. \tag{1}$$

To further reduce the memorized error, an error correction approach has been introduced to minimize the error function

$$E(\mathbf{W}) = \frac{1}{2} \left\| \mathbf{y}_k - \mathbf{W}\mathbf{x}_k \right\|^2. \tag{2}$$

Hopfield associative memory is a nonlinear content-addressable memory for storing information in a dynamically stable environment [12]. The Hopfield network is a single-layer recurrent network which contains feedback paths from the output nodes back into their inputs. Given an input $\mathbf{x}(0)$, the Hopfield network iteratively updates the output vector by

$$\mathbf{x}(k+1) = f(\mathbf{W}\mathbf{x}(k) - \boldsymbol{\theta}), \tag{3}$$

until the output vector become constant, where $f(\cdot)$ is the activation function.

At each module of automatic biometric authentication systems there remains room for improvement, which can be very effectively dealt with by making use of computational intelligence technologies.

Applications: Associative memory is able to deduce and retrieve the memorized information from possibly incomplete or corrupted biometric data. This makes it very competitive in robust biometric data representation. In [13], [14], kernel associative memory (KAM) network has been used for the retrieval of the corresponding prototypes of input face images. Compared with some popular face recognition methods, KAM can achieve better recognition accuracy and is robust in recognizing incomplete face images.

Self-Organizing Neural Networks

Kohonen's *self-organizing map* (SOM) is an unsupervised neural network model which uses a competitive learning rule to project a structured, high-dimensional data manifold onto a low-dimensional, topologically ordered set of nodes. SOM preserves the topological relationship of the data. This makes SOM capable of simultaneously acquiring the topological ordering and good clusters of the input data.

Given a training vector $\mathbf{x}$ and a SOM with m nodes $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, a typical training algorithm usually include three major steps:

1. Determine the best matching node (*winner*) $\mathbf{w}_k$ according to the minimum distance between the nodes and the training vector $\mathbf{x}$,

$$k = \arg\min(d(\mathbf{x}, \mathbf{w}_i)), \tag{4}$$

2. Update the nodes of SOM using the learning rule

$$\mathbf{w}_i = \begin{cases} \mathbf{w}_i + \eta\Lambda(i,k)[\mathbf{x} - \mathbf{w}_i], & if\ \ i \in N(k) \\ \mathbf{w}_i, & if\ \ i \notin N(k) \end{cases}, \tag{5}$$

   where $N(k)$ denotes the neighborhood of the node k, $\Lambda(i,k)$ is the neighborhood function.

3. Modify the learning rate η, size of the neighborhood, and the neighborhood parameters σ.

A central problem with SOMs is that if the learning rate is not gradually reduced to zero, clusters formed by SOM may not be stable. This is known as the *stability/plasticity dilemma*. To overcome this dilemma, Carpenter and Grossberg proposed another self-organized neural network method, *adaptive resonance theory*, which involves accepting and adapting the nodes only when the input is sufficiently similar to it [15]. To accelerate the computation of SOM, Luttrell introduced a hierarchical SOM scheme [16], and Kohonen suggested using the batch formation of the SOM algorithm [17].

Applications: There are many examples of successful applications of self-organized neural networks in biometric data representation. In [18], a face image is first divided into a number of local image samples. A SOM is then applied to the samples, and the input is quantized into a number of topologically ordered values. SOMs have also been used for recognizing partially occluded and expression variant faces, where a multiple SOM scheme is adopted to train a single SOM for each class [19].

2.2. CI-based Dimensionality Reduction

Dimensionality reduction is an important task in biometric systems for three major reasons. First, biometric data is high dimensional and most current recognition approaches suffer from the "curse of dimensionality" problem. Second, original biometric data always contains information that is less discriminative or that is not useful for recognition. Dimensionality reduction allows this information to be efficiently suppressed without losing discriminative information. Third, dimensionality reduction reduces the system's memory and computational requirements. Dimensionality reduction techniques have been extensively applied to face [20], [21], ear [22], palmprint [23], [24], fingerprint [25], gait recognition [26], and even multimodal biometrics [22], [27]. Since dimensionality reduction usually can be formalized to an optimization problem, CI technologies have been very successful in biometric data dimensionality reduction. The following provides a survey of five of these CI-based dimensionality reduction techniques.

Adaptive Principal Component Analysis (PCA)/Linear Discriminant Analysis (LDA)

PCA and LDA are two powerful dimensionality reduction techniques for data representation/discrimination. The classical PCA and LDA algorithms are batch-based where all the training samples are already available. In many real-world applications, however, training data are received incrementally. In these cases, principal components (PCs) or discriminant vectors are best extracted using adaptive methods.

PCs have been adaptively extracted using various neural network approaches. Most adaptive PCA approaches can be derived from Oja's Hebbian learning rule, where a single linear neuron is used to learn the first principal component [28]. Sanger proposed a generalized Hebbian learning rule (GHA) to adaptively extract the first several PCs [29]. In many real-world applications, the GHA method may suffer from the overflow problem, where with the increment of the training samples, the first PC totally dominates and the influence of other PCs is diminished. To overcome this, Kung and Diamantaras introduced an anti-Hebbian learning rule, adaptive principal components extraction, which adds lateral

Models of real-world biometric data are often highly nonlinear and very complex. CI-based methods can adaptively model a complex biometric characteristic, thereby obviating the need to make many assumptions about the precise mathematical model.

connections at the output of the Hebbian-based PCA networks [30]. Other learning rules that have been investigated for adaptively extracting PCs include projection approximation subspace tracking [31] and candid covariance-free incremental PCA [32].

Neural network approaches have also been applied in adaptive LDA. Mao and Jain developed a two-layer PCA network architecture to perform LDA of training data [33]. In this network architecture, the training of the two layers cannot proceed simultaneously. Rather, the second layer should be trained after the convergence of the first layer. Chatterjee and Roychowdhury further proposed to use a $Q^{-1/2}$ algorithm to train the first layer and an adaptive eigenvector computation algorithm to train the second layer [34]. Subsequently Moghaddam and Zadeh introduced a modified version of a self-organizing neural network by using the steepest descent optimization method to compute $Q^{-1/2}$, and thus accelerate the convergence of adaptive LDA [35], [36].

Applications: When the biometric data are non-stationary or are received incrementally, adaptive PCA/LDA is a useful approach for online/dynamic extraction of PCs and discriminant vectors. In [32] and [37], adaptive PCA has been applied to face representation. In [38], Zhao et al. developed a novel adaptive PCA using the singular value decomposition updating algorithm and applied it to PCA- and PCA + LDA-based face recognition.

Independent Component Analysis (ICA)

ICA is an efficient extension of PCA and has been widely applied to blind signal separation and feature extraction. Unlike PCA, where the second-order dependence (correlation) is eliminated in the PCA-transformed space, ICA aims to make its projection coefficients mutually independent by eliminating both the second-order and the high-order dependence. There are a number of methods for performing ICA, such as the *infomax* algorithm [39] and the fixed-point (FastICA) algorithm [40]. The infomax algorithm is derived from a neural network point by maximizing the mutual information between the input data and the network output. Subsequently, Hyvärinen et al. developed a FastICA algorithm which uses a fixed-point iteration scheme to train the ICA projection matrix.

Applications: High-order dependence among biometric data may be very important for biometric verification and identification. Bartlett et al. [41], Yuen [42], and Liu [43] were among the first to apply ICA to face representation and recognition, and found that ICA outperforms PCA in face recognition. Subsequently, inconsistent results have been reported [44], [45]. Recently, Draper et al. [46] independently carried out a comprehensive comparison between the performances of PCA and ICA. They found the relative performance of ICA and PCA mainly depends on the ICA architecture (I and II) and the distance metric.

ICA can also be used for robust recognition of biometric data with local distortion and partial occlusion. In [47], Kim et al. proposed an efficient part-based local representation method, locally salient ICA (LS-ICA). LS-ICA employs only locally salient information from important facial parts in order to maximize the benefit of applying the idea of "recognition by parts," and thus creates part-based local basis images by imposing an additional localization constraint in the process of computing ICA architecture I basis images.

Evolutionary Feature Extraction (EFE)

Feature extraction is essentially a kind of optimization problem. Genetic algorithms (GAs) can contribute substantially to solving such problems. GA first encodes a set of candidate transform functions into individuals. These individuals are then iteratively evolved to generate a new generation of individuals according to the fitness function and the selection, crossover, and mutation rules. When the GA converges, the optimal transform function is satisfied by producing a final generation of the fittest individuals. Compared with other feature extraction algorithms, EFE has the advantage of potential parallelizability and are thus expected to be more applicable on large-scale and high-dimensional data.

When EFE is used for linear feature extraction or feature selection, the candidate transform function would be a projection matrix (*feature transform*) or a number of variables (*feature selection*). The projection matrix can be represented as a number of basis vectors. Some approaches have been proposed to encode and evolve the basis vectors. Liu and Wechsler presented an evolutionary pursuit method to search the optimal projection basis vectors by rotating the standard basis of the search space [48]. Most recently, Zhao et al. proposed to generate the projection basis vectors via linear combination of the basis vectors of the search space [50]. Zheng et al. introduced a two-step EFE framework, where GA is used for selecting principal components before the subsequent LDA in PCA subspace [49].

Applications: Face recognition is one of the major applications of evolutionary feature extraction. In [48], evolutionary pursuit was successfully applied to face recognition and achieved a better performance than the popular Eigenfaces and Fisherfaces methods. In GA-Fisher, GA is used to select principal components, and LDA is then performed in the GA-PCA subspace [49]. To overcome the computational and memory complexity of evolutionary pursuit, Zhao et al. presented a direct EFE method by introducing a novel basis vectors generation method and adopting a constrained search space strategy [50].

Multilayer perceptron is capable of adaptively approximating the function of any linear or nonlinear filter, and thus is very competitive when applied to biometric data representation.

Kernel Dimensionality Reduction Methods

Kernel dimensionality reduction (KDR) methods provide an efficient means of handling nonlinear feature extraction problems. One of the simplest ways to make the linear dimensionality reduction techniques nonlinear is to map the d-dimensional input data to an m-dimensional feature space. In kernel-based methods, data $\mathbf{x}$ are implicitly mapped into a higher dimensional or infinite dimensional *feature space* $\Phi : \mathbf{x} \rightarrow \Phi(\mathbf{x})$. The inner product in *feature space* can be easily computed using the kernel function

$$\mathrm{K}(\mathbf{x}, \mathbf{y}) = <\Phi(\mathbf{x}), \Phi(\mathbf{y})>. \tag{6}$$

From this, Schölkopf et al. pointed out that "*every (linear) algorithm that only use scalar (inner) products can implicitly be executed in* Φ *by using kernels, i.e., one can very elegantly construct a nonlinear version of a linear algorithm*." [51] Now a number of kernel-methods, such as kernel PCA (KPCA) [51], [52], kernel Fisher discriminant (KFD) [53], [54], complete kernel Fisher discriminant (CKFD) [55], and kernel direct discriminant analysis (KDDA) [56], and kernel canonical correlation analysis (KCCA) [57], have been developed from their corresponding linear dimensionality reduction algorithms.

Applications: When the variations of biometric data or features are complex and nonlinear, KDR outperform the classical linear methods. So far, varieties of KDR methods have been proposed and used in biometric data dimensionality reduction. Recent developments on KDR techniques can be grouped into two categories as either solutions to the singularity and poor estimation problem or as combinations with image transform techniques.

There are two popular strategies to avoid the singularity of the scatter matrix, the transform-based and the algorithm-based. The transform-based strategy, such as KFD [58] and kernel uncorrelated discriminant analysis [59], first reduces the feature dimensions and then uses LDA/CCA for feature extraction. Note however, that because some potential discriminatory information contained in some small PCs is lost in the KPCA step, these methods are only approximate. Different from the transform-based strategy, the algorithm-based, such as KDDA, finds an algorithm that can circumvent the singular case directly [56].

Another unfavorable effect of KDR is that limited sample size can cause the poor estimation of the scatter matrices, resulting in an increase in classification errors. Regularization technique, where a small perturbation is added to the within-class scatter matrix $\mathbf{S}_w$, can be used for estimating $\mathbf{S}_w^{\mathrm{R}}$ [60]. So far, regularization methods (e.g., 3-parameter and 1-parameter regularization), have been developed for estimating $\mathbf{S}_w^{\mathrm{R}}$ [61].

Image transforms have been used in combination with KDR techniques for dimensionality reduction. Figure 3 illustrates the procedure for transforming domain KDR techniques. After image transformation, the transform coefficients might be more robust against variations of landmark location and illumination. KDR has been found to be effective on the transform domains of the Gabor, wavelet, and discrete cosine transforms [62], [63].

The Iteratively Reweighted Least-Squares Method (IRLS)

IRLS is an efficient method for solving nonlinear optimal problems [64]–[66]. Assume that we are given a set of training data, $\mathbf{X} = \{\mathbf{x}_1, \cdots, \mathbf{x}_m\} \subseteq \mathbb{R}^d$ with the corresponding labels $\mathbf{y} = \{y_1, \cdots, y_m\}$. The IRLS algorithm can be used to fit an optimal linear mapping model between $\mathbf{x}$ and y, $y = \mathbf{w}^T\mathbf{x} + b$. Let $\beta = [b\ \mathbf{w}^T]^T$, $\mathbf{z}_i = [1\ \mathbf{x}_i^T]^T$, and $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_m]^T$. IRLS computes the optimal vector $\hat{\boldsymbol{\beta}}$ which minimizes the criterion

$$J(\boldsymbol{\beta}) = \sum_{i=1}^{m} \varphi\{(y_i - \boldsymbol{\beta}^T\mathbf{z}_i)^2\}, \tag{7}$$

where the function $\varphi(\cdot)$ should be defined according to the application tasks. The optimal vector $\boldsymbol{\beta}$ can be obtained by iteratively performing the next two steps until convergence:

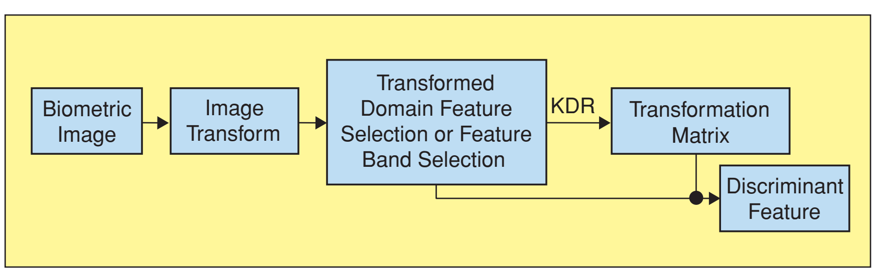

FIGURE 3 Procedure of the image transform domain kernel dimensionality reduction technique.

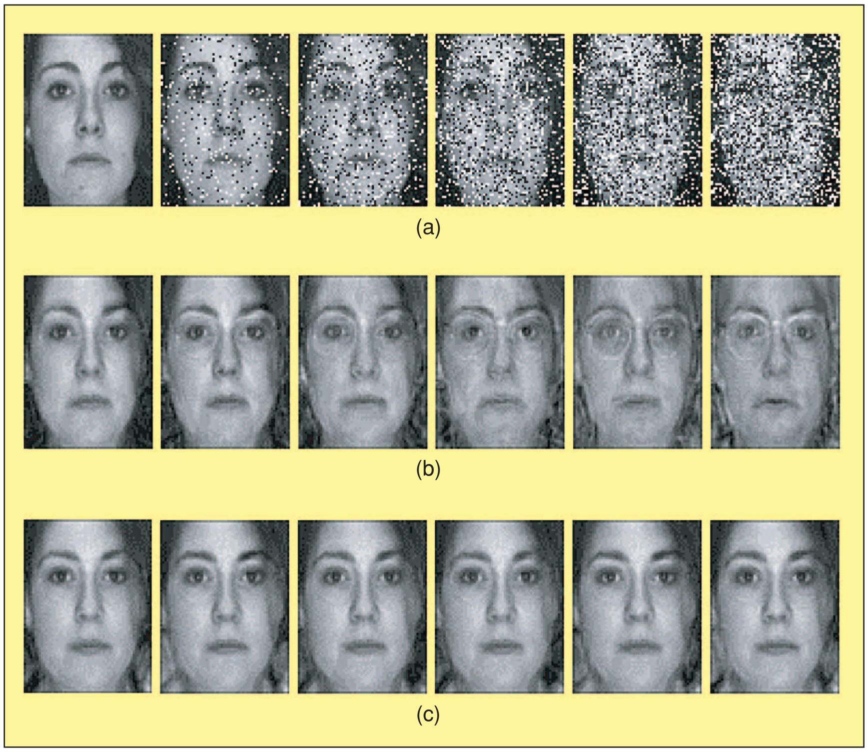

FIGURE 4 The reconstruction of face images with different degreess of salt and pepper noise: (a) original image; (b) reconstructed images using Eigenfaces; (c) reconstructed images using IRF-Eigenfaces.

1. Compute the weights $\mathbf{W}_{ii} = \frac{\partial\varphi\{(y_i - \boldsymbol{\beta}^T \mathbf{z}_i)^2\}}{\partial[(y_i - \boldsymbol{\beta}^T \mathbf{z}_i)^2]}$;
2. Update the vector $\boldsymbol{\beta} = (\mathbf{Z}^T\mathbf{W}\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{W}\mathbf{y}$.

Applications: IRLS has been successfully applied to the development of robust PCA, where outliers would cause significant deterioration in the performance of PCA. IRLS can also be combined with Eigenfaces to extract robust features from the noise and partial occluded facial images. For example, iteratively reweighted fitting of Eigenfaces (IRF-Eigenfaces) can address this by first defining a generalized objective function and then using the IRLS algorithm to compute the feature vector $\mathbf{y}$ by minimizing the objective function [67].

IRF-Eigenfaces is more robust when it comes to reconstructing or recognizing noisy and occluded facial images. Using the AR face database [68], the addition of salt and pepper noise has little effect on the reconstruction performance of IRF-Eigenfaces, as shown in Figure 4. IRF-Eigenfaces also has a robust reconstruction performance for partially occluded facial images, as shown in Figure 5. For other modalities of biometric image (e.g., palmprint), IRF-Eigenfaces also has achieve a robust reconstruction performance for noisy or partially occluded images, as shown in Figure 6 and Figure 7.

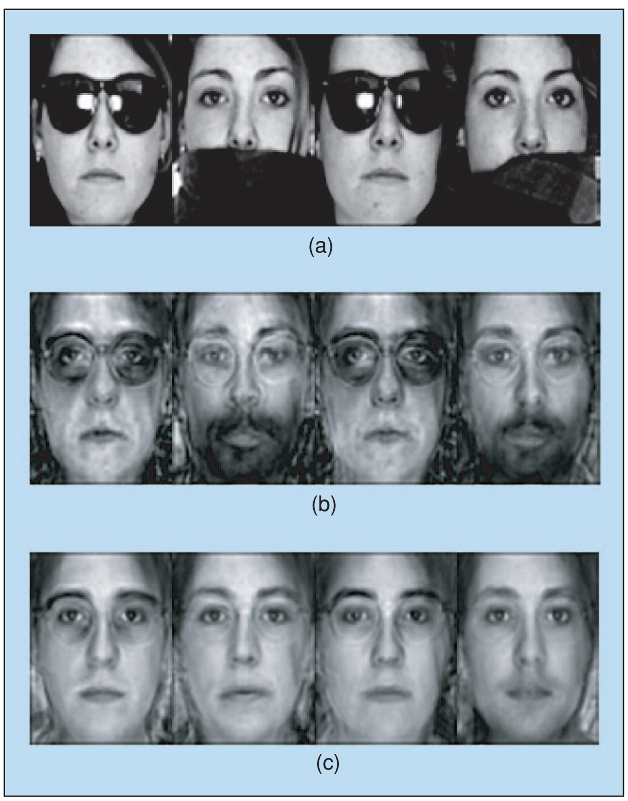

FIGURE 5 Reconstructed face images using Eigenfaces and IRF-Eigenfaces: (a) original partially occluded images; (b) reconstructed images using Eigenfaces; (c) reconstructed images using IRF-Eigenfaces.

3. Computational Intelligence-Based Biometric Matching

In the matching module, a feature vector extracted from biometric data is compared against the stored prototype templates and a decision is then made. In biometric verification, the matching issue is one of whether a person is "*who he claims to be*" whereas in biometric identification the matching issue is one of "*whose biometric data is this?*" based on a one-to-many comparison. Thus biometric matching can be formalized into a two-class or multi-class classification problem, and CI can be used to enhance the robustness, adaptivity, and recognition performance of the matching module. For example, neural networks can be used to train a neural biometric matcher and fuzzy technologies can be used as an uncertainty modeling tool to fuse the matching results obtained from different biometric systems. In the following we present an overview of three main CI-based biometric matching methods.

Radial Basis Function Neural Network (RBFNN)

RBFNN is a three-layer feed-forward neural network made up of an input layer, a hidden layer and an output layer. The output of a RBFNN with K hidden units is represented by

$$f(\mathbf{x}) = \sum_{k=1}^{K} w_k \phi_k(\mathbf{x}, \mathbf{c}_k) = \sum_{k=1}^{K} w_k \phi_k(\|\mathbf{x} - \mathbf{c}_k\|_2), \quad (8)$$

where $\mathbf{x}$ is an input vector, $\phi_k(\mathbf{x}, \mathbf{c}_k)$ is the RBF function, and $\mathbf{c}_k$ is the RBF centers in the input vector space. One common RBF function is the Gaussian basis function, $\phi_k(\mathbf{x}, \mathbf{c}_k) = \exp(-\|\mathbf{x} - \mathbf{c}_k\|_2^2/\sigma^2)$. The main task in training a RBFNN is to determine the centers $\mathbf{c}_k$, the weights w_k, and the parameter of Gaussian basis function σ. One simple algorithms is the stochastic gradient approach, which, however, is prone to getting stuck in a local minimum and cannot determine the number of centers. Usually a two-step strategy is adopted. In the first step, clustering is used to initialize the centers and the parameters. In the second step, the centers, σ, and weights are adjusted as an optimization procedure.

Applications: Being computationally simple and robustly generalizable, RBFNN has been applied to various tasks in face-processing including face detection, expression analysis, face recognition, and gender classification [69]–[79]. A number of RBFNN-based face recognition systems have been trained using PCA, LDA, pseudo Zernike moment, and discrete cosine transform features. In [69], [70], RBFNN is combined with a decision tree to implement a mixture of experts for classification of gender, ethnic origin, and facial poses. In [71], RBFNN is used to identify facial expression and motion. In [73], [74], RBFNN has been used to construct an integrated automatic face detection and recognition system.

Support Vector Machine (SVM)

SVMs are powerful tools for classification and regression with many desirable properties, including a good generalization performance and embedding a nonlinear decision function via kernel functions. The standard SVM approach is to design a linear optimal hyperplane as the decision boundary of a binary classification problem. Given a set of linearly separable data $\{\mathbf{x}_i, y_i\}, \mathbf{x}_i \in \mathbb{R}^d, y_i \in \{-1, 1\}, i = 1, \cdots, N$, the optimal hyperplane is chosen to maximize the margin between the two classes by minimizing the objective function

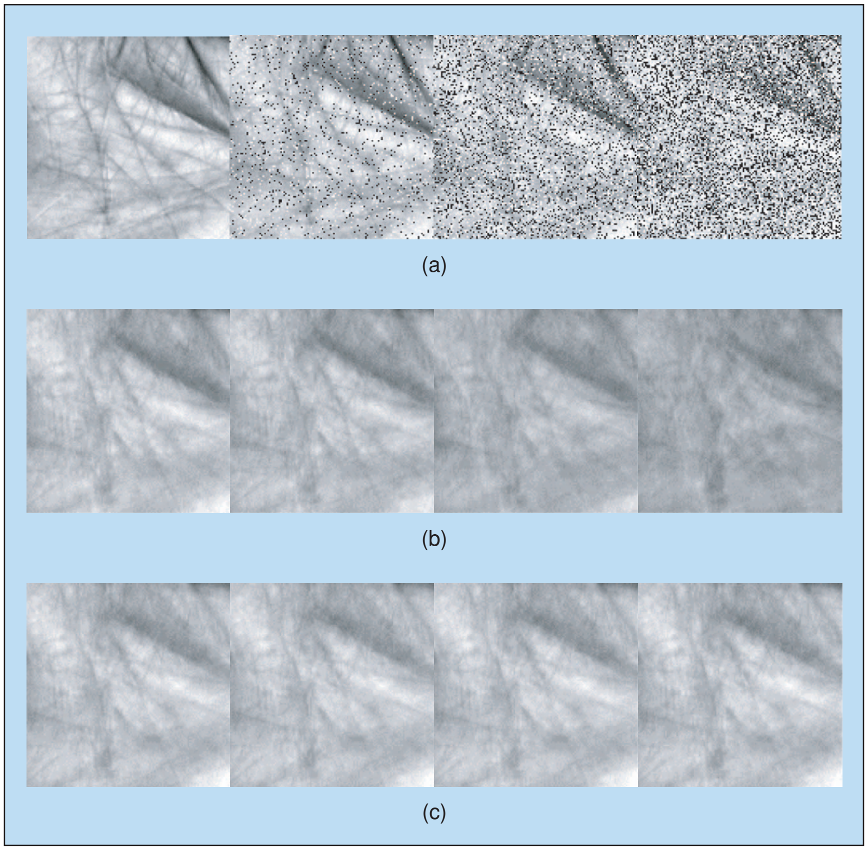

FIGURE 6 The reconstruction of palmprint images with different degrees of salt and pepper noise: (a) original image; (b) reconstructed images using Eigenfaces; (c) reconstructed images using IRF-Eigenfaces.

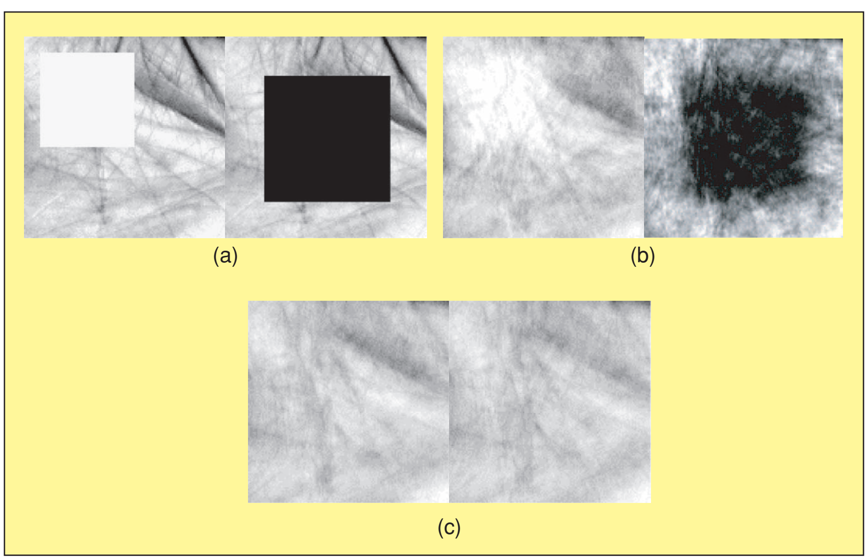

FIGURE 7 Reconstructed palmprint images using Eigenfaces and IRF-Eigenfaces: (a) original images; (b) reconstructed images using Eigenfaces; (c) reconstructed images using IRF-Eigenfaces.

Associative memory is powerful in the deduction and retrieval of the memorized information from possibly incomplete (noisy, partially occluded, or locally distorted) biometric data.

$$E = \|\mathbf{w}\|_2^2, \text{ s.t. } (\langle \mathbf{x}, \mathbf{w} \rangle + b)\, y_i \geq 1. \tag{9}$$

Generally, the set of training data is not linearly separable, and another objective function is defined by introducing "slack" variable ξ_i

$$E = \|\mathbf{w}\|_2^2 + C \sum_{i=1}^{N} \xi_i, \text{ s.t. } (\langle \mathbf{x}, \mathbf{w} \rangle + b)\, y_i \geq 1 - \xi_i, \tag{10}$$

where C is a hyper-parameter controlling the tradeoff between margin maximization and classification error. This optimization problem can be transformed into its dual formalization

$$\max_{\alpha} \sum_i \alpha_i + \sum_{i,j} \alpha_i \alpha_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle, \text{ s.t.}$$
$$0 \leq \alpha_i \leq C, \sum_i \alpha_i y_i = 0, \tag{11}$$

where α_i is the Largrange multiplier. The decision function of SVM can be represented as

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b = \sum_{i=1}^{N} \alpha_i y_i \langle \mathbf{x}_i, \mathbf{x} \rangle + b, \tag{12}$$

where $\mathbf{x}_i$ is a support vector if $\alpha_i > 0$.

Applications: SVM have been successfully applied to speaker and face recognition. When applied to speaker recognition, the regular SVM method would be inefficient when the number of training frames is large [80]. One strategy to overcome this is to classify sequences instead of frames by using sequence discriminative kernels [81], [82]. In face biometrics, SVM has been applied to various tasks in computational face-processing. In [83]–[85], better results have been reported using the SVM-based face detectors. Moghaddam and Yang carried out a comprehensive evaluation of classification methods for recognizing gender from facial images. The best classification performance was that of SVM [87]. In face recognition, SVM can be used to weigh the discriminatory power at the nodes of an elastic graph, thus enhancing the performance of elastic graph matching for face recognition [88].

Fuzzy Technology

Fuzzy sets are extensions of abstract sets by introducing appropriate membership functions. Fuzzy technology offers an effective, more flexible way to describe a complex system and has been successfully applied to artificial intelligence, information retrieval systems, pattern recognition, and image processing.

Applications: Fuzzy technologies have been successfully applied to many biometric recognition systems, such as face recognition, fingerprint recognition, and multimodal biometrics. In face recognition, fuzzy c-means has been applied to initialize RBFNNs and parallel neural networks [78], [89]. Kwak and Pedrycz proposed a fuzzy kNN rule to assign the class membership of each sample, and incorporate it with Fisherfaces, resulting in a fuzzy Fisherfaces method [90]. Wu et al. developed two fuzzy models to describe skin color and hair color for face detection [91]. In fingerprint recognition, a normalized fuzzy similarity measure has been proposed to match distorted fingerprints [92]. In palmprint recognition, Wu et al. proposed a fuzzy directional element energy feature approach [93].

Another important application of fuzzy technology is multimodal biometrics and multiple classifier systems. In [94], fuzzy clustering methods have been used for decision-level fusion of face and speaker biometrics, and achieve better performance than k-mean and other popular fusion algorithms. In [95], [96], the Choquet fuzzy integral, a fuzzy information fusion approach, is used to combine the outputs of individual classifiers for face recognition.

4. Conclusion and Discussion

Computational intelligence technologies, sometimes used in combination with traditional methods, have been proved to be effective and efficient in biometric feature extraction and matching tasks. Some of the key reasons that make CI competitive are:

1. Biometric recognition is one of the human capabilities. CI, characterized by imitating biological function, is a natural candidate technology in biometrics research.
2. CI-based technologies can provide a robust scheme for recognizing non-ideal and incomplete biometric data, which in the acquisition and pre-processing stage may sometimes be noisy, partially occluded, or inaccurately located.
3. The tasks of feature extraction and the matching of biometric data can usually be formalized as complex nonlinear optimization problems and CI has been very successful in solving highly complex problems.
4. CI-based technologies are efficiently adaptive. This capacity allows the development of real-time biometric systems capable of online learning and able to adjust to different times and environments.

This article presents a survey of several major CI-based applications in current biometric technologies. Varieties of evolutionary computation and neural network techniques have been successfully applied to biometric data representation and dimensionality reduction. CI-based methods,

When the biometric data are non-stationary or are received incrementally, adaptive PCA/LDA is a useful approach for online/dynamic extraction of principal components and discriminant vectors.

including neural networks and fuzzy technologies, have also been extensively investigated for biometric matching. It is our hope that this survey will encourage readers to further explore CI-based algorithms in the future development of biometric technologies.

Acknowledgments

The work is partially supported by the UGC/CRC fund from the HKSAR Government, the central fund from the Hong Kong Polytechnic University and the National Natural Science Foundation of China (NSFC) under the contract No. 60332010 and No. 60620160097.

References

[1] A.K. Jain, R. Bolle, and S. Pankanti, *Biometrics: Personal Identification in Networked Society*, Kluwer, Boston, 1999.

[2] D. Zhang, *Automated Biometrics: Technologies and Systems*, Kluwer, Boston, 2000.

[3] Y. Li, K. Wang, and D. Zhang, "Step acceleration based training algorithm for feedfoward neural networks," *Proc. Int'l Conf. Pattern Recognition 2002*, vol. 2, pp. 84–87, 2002.

[4] E. Levin, N. Tishby, and S.A. Solla, "A Statistical Approach to Learning and Generalization in Layered Neural Networks," *Proc. IEEE*, vol. 78, no. 10, pp. 1568–1574.

[5] B.H. Juang and S. Katagiri, "Discriminative Learning for Minimum Error Classification," *IEEE Trans. Signal Processing*, vol. 40, no. 12, pp. 3043–3054, 1992.

[6] Z.Z. Zhang, and N. Ansari, "Structure and Properties of Generalized Adaptive Neural Filters for Signal Enhancement," *IEEE Trans. Neural Netw.*, vol. 7, no. 4, pp. 857–868, July 1996.

[7] K. Suzuki, I. Horiba, and N. Sugie, "Efficient Approximation of Neural Filters for Removing Quantum Noise from Images," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1787–1799, 2002.

[8] K. Suzuki, I. Horiba, and N. Sugie, "Neural Edge Enhancer for Supervised Edge Enhancement from Noisy Images," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1582–1596, 2003.

[9] S.L. Phung, D. Chai, and A. Bouzerdoum, "A Universal and Robust Human Skin Color Model Using Neural Networks," *Proc. Int'l Joint Conf. Neural Networks*, vol. 4, pp. 2844–2849, July 2001.

[10] S.L. Phung, A. Bouzerdoum, and D. Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, no. 1, pp. 148–154, 2005.

[11] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 20, no. 1, pp. 23–38, 1998.

[12] J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. National Academy of Science*, USA, vol. 79, pp. 2554–2558, 1982.

[13] B. Zhang, H. Zhang, and S.S. Ge, "Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory," *IEEE Trans. Neural Netw.*, vol. 15, no. 1, pp. 166–177, 2004.

[14] H. Zhang, B. Zhang, W. Huang, and Q. Tian, "Gabor Wavelet Associative Memory for Face Recognition," *IEEE Trans. Neural Netw.*, vol. 16, no. 1, pp. 275–278, 2005.

[15] G.A. Carpenter, and S. Grossberg, "A Massively Parallel Architecture for Self-Organizing Neural Pattern Recognition Machine," *Computer Vison, Graphics, and Image Processing*, vol. 37, pp. 54–115, 1987.

[16] S.P. Luttrell, "Hierarchical Self-Organizing Networks," in *Proc. First IEE Conf. Artificial Neural Networks*. London: British Neural Network Soc., 1989, pp. 2–6.

[17] T. Kohonen, S. Kaski, K. Lagus, J. Salojarvi, J. Honkela, V. Paatero, and A. Saarela, "Self Organization of a Massive Document Collection," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, 2000.

[18] S. Lawrence, C.L. Giles, A.C. Tsoi, and A.D. Back, "Face Recognition: A Convolutional Neural-Network Approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, 1997.

[19] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Recognizing Partially Occluded, Expression Variant Faces From Single Training Image per Person with SOM and Soft k-NN Ensemble," *IEEE Trans. Neural Netw.*, vol. 16, no. 4, pp. 875–886, 2005.

[20] M.A. Turk, and A.P. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–96, 1991.

[21] P.N. Belhumeour, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces Versus Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, 1997.

[22] K. Chang, K.W. Bowyer, S. Sarkar, and B. Victor, "Comparison and Combination of Ear and Face Images in Appearance-based Biometrics," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1160–1165, 2003.

[23] G. Lu, D. Zhang, and K. Wang, "Palmprint Recognition Using Eigenpalms Features," *Pattern Recognit. Lett.*, vol. 24, nos. 9–10, pp. 1463–1467, 2003.

[24] X. Wu, D. Zhang and K. Wang, "Fisherpalms Based Palmprint Recognition," *Pattern Recognit. Lett.* vol. 24, pp. 2829–2838, 2003.

[25] C.H. Park, and H. Park, "Fingerprint Classification Using Fast Fourier Transform and Nonlinear Discriminant Analysis," *Pattern Recognit.*, vol. 38, no. 4, pp. 495–583, 2005.

[26] G. Zhao, R. Chen, G. Liu, and H. Li, "Amplitude Spectrum-Based Gait Recognition," *Proc. 6th IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 23–28, 2004.

[27] S. Ribaric, and I. Fratric, "A Biometric Identification System Based on Eigenpalm and Eigenfinger Features," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1698–1709, 2005.

[28] K.I. Diamantaras and S.Y. Kung, *Principal Component Neural Networks, Theory and Applications*. New York: Wiley, 1996.

[29] T.D. Sanger, "Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network," *Neural Networks*, vol. 2, pp. 459–473, 1989.

[30] S.Y. Kung, K.I. Diamantaras, and J.S. Taur, "Adaptive Principal Component EXtraction (APEX) and Applications," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1202–1217, 1994.

[31] S. Bannour, M.R. Azimi-Sadjadi, "Principal Component Extraction Using Recursive Least Squares Learning," *IEEE Trans. Neural Netw.*, vol. 6, no. 2, pp. 457–469, 1995.

[32] J. Weng, Y. Zhang, W.-S. Hwang, "Candid Covariance-Free Incremental Principal Component Analysis," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 8, pp. 1034–1040, 2003.

[33] J. Mao, A.K. Jain, "Discriminant Analysis Neural Networks," *IEEE Int'l Conf. Neural Networks 1993*, vol. 1, pp. 300–305, 1993.

[34] C. Chatterjee, V.P. Roychowdhury, "On Self-Organizing Algorithm and Networks for Class-Separability Features," *IEEE Trans. Neural Netw.*, vol. 8, no. 3, pp. 663–678, 1997.

[35] H.A. Moghaddam, K.A. Zadeh, "Fast Adaptive Algorithms and Networks for Class-Separability Features," *Pattern Recognit.*, vol. 36, no. 8, pp. 1695–1702, 2003.

[36] H.A. Moghaddam, M. Matinfar, S.M.S. Sadough, and K.A. Zadeh, "Algorithms and Networks for Accelerated Convergence of Adaptive LDA," *Pattern Recognit.*, vol. 38, no. 4, pp. 473–483, 2005.

[37] S. Yan, X. Tang, "Largest-Eigenvalue-Theory for Incremental Principal Component Analysis," *Proc. IEEE Int'l Conf. Image Processing 2005*, vol. 1, pp. 1181–1184, 2005.

[38] H. Zhao, P.C. Yuen, and J.T. Kwok, "A Novel Incremental Principal Component Analysis and Its Application for Face Recognition," *IEEE Trans. Systems, Man, Cybern.B*, vol. 36, no. 4, pp. 873–886, 2006.

[39] A.J. Bell, and T.J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[40] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Trans. on Neural Netw.*, vol. 10, no. 3, pp. 626–634, 1999.

[41] M.S. Bartlett, J.R. Movellan, and T.J. Sejnowski, "Face Recognition by Independent Component Analysis," *IEEE Trans. Neural Netw.*, vol. 13, no. 6, pp. 1450–1464, 2002.

[42] P.C. Yuen, and J.H. Lai, "Face Representation Using Independent Component Analysis," *Pattern Recognit.*, vol. 35, no. 6, pp. 1247–1257, 2002.

[43] C. Liu and H. Wechsler, "Independent Component Analysis of Gabor Features for Face Recognition," *IEEE Trans. Neural Netw.*, vol. 14, no. 4, pp. 919–928, 2003.

[44] B. Moghaddam, "Principal Manifolds and Probabilistic Subspaces for Visual Recogni-

tion," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, no. 6, pp. 780–788, 2002.

[45] D. Socolinsky and A. Selinger, "A Comparative Analysis of Face Recognition Performance with Visible and Thermal Infrared Imagery," *Proc. Int'l Conf. Pattern Recognition*, Quebec City, 2002.

[46] B.A. Draper, K. Baek, M.S. Bartlett, and J.R. Beveridge, "Recognizing Faces with PCA and ICA," *Computer Vision and Image Understanding*, vol. 91, no. 1/2, pp. 115–137, 2003.

[47] J. Kim, J. Choi, J. Yi, and M. Turk, "Effective Representation Using ICA for Face Recognition Robust to Local Distortion and Partial Occlusion," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, no. 12, pp. 1977–1981, 2005.

[48] C. Liu, and H. Wechsler, "Evolutionary Pursuit and Its Application to Face Recognition," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 22, no. 6, pp. 570–582.

[49] W.-S. Zheng, J.-H. Lai, and P.C. Yuen, "GA-Fisher: A New LDA-Based Face Recognition Algorithm with Selection of Principal Components," *IEEE Trans. Systems, Man, Cybern. B*, vol. 35, no. 5, pp. 1065–1078, 2005.

[50] Q. Zhao, D. Zhang, and H. Lu, "A Direct Evolutionary Feature Extraction Algorithm for Classifying High Dimensional Data," *American Association for Artificial Intelligence 2006*, pp. 561–566, 2006.

[51] B. Schölkopf, A. Smola, and K.R. Müller, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[52] K.I. Kim, M.O. Franz, and B. Schölkopf, "Iterative Kernel Principal Component Analysis for Image Modeling," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1351–1366, 2005.

[53] S. Mika, G. Rätsch, B. Schölkopf, A. Smola, J. Weston, and K.R. Müller, "Invariant Feature Extraction and Classification in Kernel Space," S.A. Solla, T.K. Leen, and K.R. Müller, Eds: *Advances in Neural Information Processing Systems 12*, pp. 568–574, MIT Press, Mass., 1999.

[54] G. Baudat, and F. Anouar, "Generalized Discriminant Analysis Using a Kernel Approach," *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.

[55] J. Yang, A.F. Frangi, J.Y. Yang, D. Zhang, and Z. Jin, "Kernel PCA plus LDA: A Complete Kernel Fisher Discriminant Framework for Feature Extraction and Recognition," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, no. 2, pp. 230–244, 2005.

[56] J. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, "Face Recognition Using Kernel Direct Discriminant Analysis Algorithms," *IEEE Trans. Neural Netw.*, vol. 14, no. 1, pp. 2829–2838, 2003.

[57] P.L. Lai, and C. Fyfe, "Kernel and Nonlinear Canonical Correlation Analysis," *International Journal of Neural Systems*, vol. 10, no. 5, pp. 365–377, 2000.

[58] M.H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: face recognition using Kernel methods," *Proc. 5th IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 215–220, 2002.

[59] T. Xiong, J. Ye, and V. Cherkassky, "Kernel Uncorrelated and Orthogonal Discriminant Analysis: A Unified Approach," *Proc. IEEE Conf. Computer Vision and Pattern Recognition 2006*, vol. 1, pp. 125–131, 2006.

[60] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, and J. Wang, "An Efficient Kernel Discriminant Analysis Method," *Pattern Recognit.*, vol. 38, no. 10, pp. 1788–1790, 2005.

[61] W.S. Chen, P.C. Yuen, J. Huang, and D.Q. Dai, "Kernel Machine-Based One-Parameter Regularized Fisher Discriminant Method for Face Recognition," *IEEE Trans. Systems, Man, Cybern. B*, vol. 35, no. 4, pp. 659–669, 2005.

[62] X. Jing, and D. Zhang, "A Face and Palmprint Recognition Approach Based on Discriminant DCT Feature Extraction," *IEEE Trans. Systems, Man, Cybern. B*, vol. 34, no. 6, pp. 2405–2415, 2004.

[63] C. Liu, "Gabor-Based Kernel PCA with Fractional Power Polynomial Models for Face Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 5, pp. 572–581, 2004.

[64] M.K. Kuntner, C.J. Nachtsheim, and J. Neter, *Applied Linear Regression Model*, 4th Edition, McGraw-Hill/Irwin, New York, NY, 2004.

[65] A. Navia-Vázquez, F. Pérez-Cruz, A. Artés-Rodríguez, and A.R. Figueiras-Vadal, "Weighted Least Squares Training of Support Vectors Classifiers which Leads to Compact and Adaptive Schemes," *IEEE Trans. Neural Netw.*, vol. 12, no. 5, pp. 1047–1059, 2001.

[66] K. Chen, L. Xu, and H. Chi, "Improved Learning Algorithms for Mixture of Experts in Multiclass Classification," *Neural Networks*, vol. 12, pp. 1229–1252, 1999.

[67] W. Zuo, K. Wang, and D. Zhang, "Robust Recognition of Noisy and Partially Occluded Faces Using Iteratively Reweighted Fitting of Eigenfaces," *Lecture Notes in Computer Science*, vol. 4261, pp. 844–851, Springer-Verlag, 2006.

[68] A.M. Martinez and R. Benavente. *The AR Face Database*. CVC Technical Report #24, June 1998.

[69] S. Gutta, H. Wechsler, and P.J. Phillips, "Gender and Ethnic Classification," *Proc. Third IEEE Int'l Autamatic Face and Gesture Recognition*, pp. 194–199, 1998.

[70] S. Gutta, J.R.J. Huang, P.J. Phillips, and H. Wechsler, "Mixture of Experts for Classification of Gender, Ethnic Origin, and Pose of Human Faces," *IEEE Trans. Neural Netw.*, vol. 11, no. 4, pp. 948–960, 2000.

[71] M. Rosenblum, Y. Yacoob, and L.S. Davis, "Human Expression Recognition from Motion Using a Radial Basis Function Network Architecture," *IEEE Trans. Neural Netw.*, vol. 7, no. 5, pp. 1121–1138, 1996.

[72] S. Ranganath and K. Arun, "Face Recognition Using Transform Features and Neural Networks," *Pattern Recognit.*, vol. 30, no. 10, pp. 1615–1622, 1997

[73] L.H. Koh, S. Ranganath, and Y.V. Venkatesh, "An Integrated Automatic Face Detection and Recognition System," *Pattern Recognit.*, vol. 35, pp. 1259–1273, 2002.

[74] F. Yang, and M. Paindavoine, "Implementation of an RBF Neural Network on Embedded Systems: Real-Time Face Tracking and Identity Verification," *IEEE Trans. Neural Netw.*, vol. 14, no. 5, pp. 1162–1175, 2003.

[75] D. Zhang, H. Peng, J. Zhou, and S.K. Pal, "A Novel Face Recognition System Using Hybrid Neural and Dual Eigenspaces Methods," *IEEE Trans. Systems, Man, Cybern. A*, vol. 32, no. 6, pp. 787–793, 2002.

[76] M.J. Er, S. Wu, J. Lu, and H.L. Toh, "Face Recognition with Radial Basis Function (RBF) Neural Networks," *IEEE Trans. Neural Netw.*, vol. 13, no. 3, pp. 697–710, 2002.

[77] M.J. Er, W. Chen, and S. Wu, "High-Speed Face Recognition Based on Discrete Cosine Transform and RBF Neural Networks," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 679–691, 2005.

[78] J. Haddadnia, K. Faez, and M. Ahmadi, "A Fuzzy Hybrid Learning Algorithm for Radial Basis Function Neural Network with Application in Human Face Recognition," *Pattern Recognit.*, vol. 36, no. 5, pp. 1187–1202, 2003.

[79] J. Haddadnia, and M. Ahmadi, "N-Feature Neural Network Human Face Recognition," *Image and Vision Computing*, vol. 22, no. 12, pp. 1071–1082, 2004.

[80] V. Wan, and W.M. Campbell, "Support Vector Machines for Speaker Verification and Identification," *Proc. Neural Networks for Signal Processing X*, pp. 775–784, 2000.

[81] V. Wan and S. Renals, "Evaluation of Kernel Methods for Speaker Verification and Identification," *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing*, vol. 1, 2002, pp. 669–672.

[82] V. Wan, and S. Renals, "Speaker Verification Using Sequence Discriminant Support Vector Machine," *IEEE Trans. Speech and Audio Processing*, vol. 13, no. 2, pp. 203–210, 2005.

[83] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An Application to Face Detection," *Proc. IEEE Computer Vision and Pattern Recognition 1997*, pp. 130–136, 1997.

[84] C.A. Waring, and X. Liu, "Face Detection Using Spectral Histogram and SVM," *IEEE Trans. Systems, Man, Cybern. B*, vol. 35, no. 3, pp. 467–476, 2005.

[85] R. Xiao, M.-J. Li, and H.-J. Zhang, "Robust Multipose Face Detection in Images," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 31–41, 2004.

[86] P.J. Phillips, "Support Vector Machines Applied to Face Recognition," *Advances in Neural Information Processing Systems 11*, M.S. Kearns, S. Solla, and D. Cohen, eds., MIT Press, vol. 11, pp. 803–809, 1998.

[87] B. Moghaddam, and M.-H. Yang, "Learning Gender with Support Faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 707–711, 2002.

[88] A. Tefas, C. Kotropoulos, and I. Pitas, "Using Support Vector Machine to Enhance the Performance of Elastic Graph Matching for Frontal Face Authentication," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 7, pp. 735–746, 2001.

[89] J. Lu, X. Yuan, and T. Yahagi, "A Method of Face Recognition Based on Fuzzy Clustering and Parallel Neural Networks," *Signal Processing*, vol. 86, no. 8, pp. 2026–2039, 2006.

[90] K.-C. Kwak, and W. Pedrycz, "Face Recognition Using a Fuzzy Fisherface Classifier," *Pattern Recognit.*, vol. 38, no. 10, pp. 1717–1732, 2005.

[91] H. Wu, Q. Chen, and M. Yachida, "Face Detection from Color Images Using a Fuzzy Pattern Matching Method," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 21, no. 6, pp. 557–563, 1999.

[92] X. Chen, J. Tian, and X. Yang, "A New Algorithm for Distorted Fingerprints Matching Based on Normalized Fuzzy Similarity Measure," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 767–776, 2006.

[93] X. Wu, K. Wang, and D. Zhang, "Fuzzy Directional Element Energy Feature (FDEEF) Based Palmprint Identification," *Proc. Int'l Conf Pattern Recognition 2002 (ICPR'02)*, pp. 95–98, 2002.

[94] V. Chatzis, A.G. Borş, and I. Pitas, "Multimodal Decision-Level Fusion for Person Authentication," *IEEE Trans. Systems, Man, Cybern. A*, vol. 29, no. 6, pp. 674–680, 1999.

[95] K.-C. Kwak, and W. Pedrycz, "Face Recognition Using Fuzzy Integral and Wavelet Decomposition Method," *IEEE Trans. Systems, Man, Cybern. B*, vol. 34, no. 4, pp. 1666–1675, 2004.

[96] K.-C. Kwak, and W. Pedrycz, "Face Recognition: A Study in Information Fusion Using Fuzzy Integral," *Pattern Recognit. Lett.*, vol. 26, no. 6, pp. 719–733, 2005.