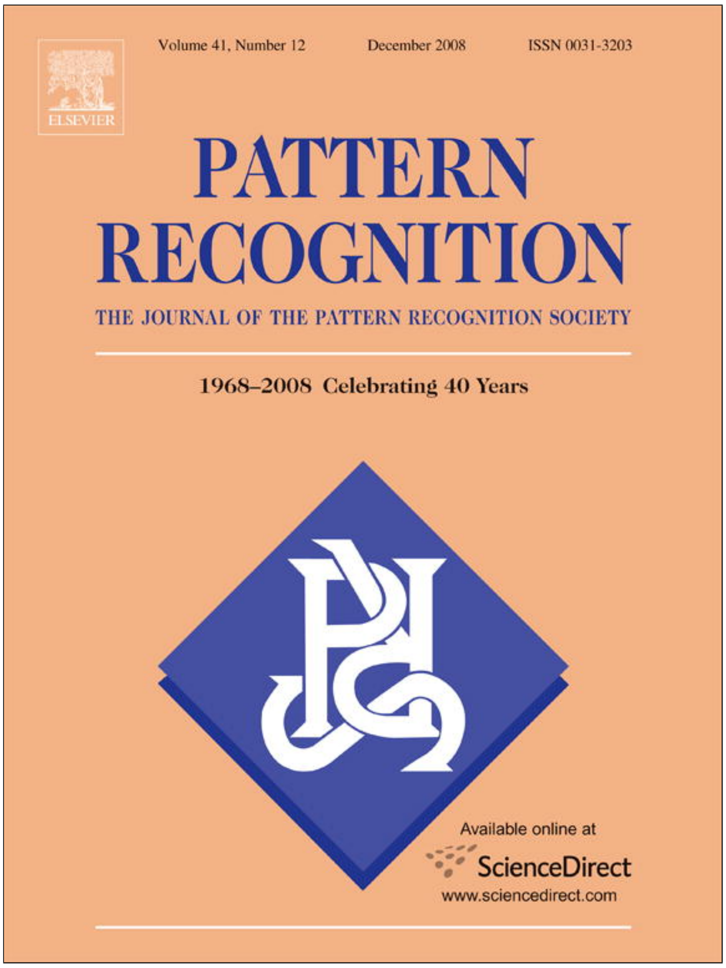
Volume 41, Number 12
December 2008
ISSN 0031-3203
ELSEVIER
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
1968–2008 Celebrating 40 Years
Available online at
ScienceDirect
www.sciencedirect.com

Pattern Recognition 41 (2008) 3834 – 3841

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

1D correlation filter based class-dependence feature analysis for face recognition

Yan Yan [a,b,*], Yu-Jin Zhang [a,b]

[a] Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China
[b] Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 14 December 2007
Received in revised form 5 April 2008
Accepted 28 May 2008

Keywords:
Correlation filter
Optimal extra-class origin output tradeoff filter (OEOTF)
Class-dependence feature analysis (CFA)
Linear subspace learning
Face recognition

ABSTRACT

In this paper, a novel one-dimensional correlation filter based class-dependence feature analysis (1D-CFA) method is presented for robust face recognition. Compared with original CFA that works in the two dimensional (2D) image space, 1D-CFA encodes the image data as vectors. In 1D-CFA, a new correlation filter called optimal extra-class origin output tradeoff filter (OEOTF), which is designed in the low-dimensional principal component analysis (PCA) subspace, is proposed for effective feature extraction. Experimental results on benchmark face databases, such as FERET, AR, and FRGC, show that OEOTF based 1D-CFA consistently outperforms other state-of-the-art face recognition methods. This demonstrates the effectiveness and robustness of the novel method.

1. Introduction

Over the past few decades, face recognition has become a popular area of research in pattern recognition and computer vision due to its wide range of commercial and law enforcement applications, such as biometric authentication, video surveillance, and information security [1].

Until now, a great number of face recognition methods have been developed and one of the most successful techniques is the appearance-based method. When using appearance-based methods, a face image is usually considered as a point in the high-dimensional space. Then, the statistical learning method is applied to derive an effective representation (a low-dimensional feature). Finally, a classifier is designed in the feature space. Linear subspace learning methods, such as Eigenface [2], Fisherface [3], LDA/FKT (linear discriminant analysis/Fukunaga–Koontz transform) [4], C-LDA (complete LDA) [5], MMSD (multiple maximum scatter difference) [6], and Laplacianface [7] are typical dimensionality reduction methods to find a low-dimensional feature space.

Turk and Pentland [2] proposed Eigenface algorithm for face recognition. The algorithm uses principal component analysis (PCA) which finds the principal components of the distribution of face images for dimensionality reduction. Note that PCA is optimal for representation, not necessarily for classification. Therefore, Fisherface algorithm [3] uses LDA to search a set of basis components which maximizes the ratio of between-class scatter to within-class scatter. Due to the 'small sample size' problem [8] in face recognition, the within-class scatter matrix is usually singular. Thus, the execution of LDA encounters computational difficulty. PCA is often used as a preprocessing step to reduce the dimensionality [3] and LDA is then performed in the low-dimensional PCA subspace where the within-class scatter matrix becomes nonsingular. However, this method may result in the loss of important discriminative information [4]. Many methods [4–6] have been developed to take full advantage of the discriminative information in the face space. LDA/FKT [4] obtains the discriminant subspace by applying FKT on the within-class scatter matrix and between-class scatter matrix while C-LDA [5] and MMSD [6] derive the discriminant features both in the range of the between-class scatter matrix and in the null space of the within-class scatter matrix. Unlike PCA and LDA which attempt to preserve the global Euclidean structure, Laplacianface algorithm [7] that is based on locality preserving projections (LPP) finds a face subspace to preserve the local structure of face manifold.

It can be seen that the projection matrices obtained by traditional linear subspace learning methods [2–7] are related to the statistical characteristics of all training samples. The projection axis tries to preserve (e.g. PCA) or discriminate (e.g. LDA) all classes.

Recently, Kumar et al. [9,10] proposed a novel linear subspace learning method called class-dependence feature analysis (CFA) for face recognition. Different from traditional linear subspace learning methods, the projection axis obtained by CFA tries to discriminate one specific class from all other classes. Different projection axes

* Corresponding author at: Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. Tel.: +86 1062781291; fax: +86 1062770317.
E-mail address: yan-y04@mails.tsinghua.edu.cn (Y. Yan).

0031-3203/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.05.028

concern different classes. In particular, CFA is based on the design of advanced correlation filter technique which emphasizes the outputs of one face class and suppresses the outputs of other face classes. According to different criterions, different correlation filters [11] can be designed.

Original CFA [9,10] designs correlation filters in the two-dimensional (2D) image space. For simplicity, we call original CFA based on the 2D correlation filter 2D-CFA. As a result, 2D-CFA cannot be applied to vector data or M-th order ($M \geqslant 3$) tensor data directly. In our previous work [12], a tensor correlation filter based CFA (TCF-CFA) method which generalizes 2D-CFA by encoding the image data as tensors was presented. It has been proved that TCF-CFA can be derived in a similar way as 2D-CFA, which is a special case of TCF-CFA when the image data are encoded as second-order tensors (i.e. image matrices) [12]. Moreover, commonly used correlation filters in TCF-CFA, such as MACE (minimum average correlation energy) filter [13], MVSDF (minimum variance synthetic discriminant function) filter [14], and OTF (optimal tradeoff filter) [15], have the same form as those in 2D-CFA.

In this paper, we mainly concentrate on one-dimensional correlation filter based CFA (1D-CFA), since traditional algorithms [2–7] show great superiority by encoding the face image data as vectors. As far as we know, few investigations concern the design of correlation filters in the 1D form for the face recognition problem.

It is worthwhile to highlight several aspects of the proposed approach here:

1. Correlation filters are designed in the low-dimensional PCA subspace. Compared with original CFA which designs correlation filters in the 2D image space [9,10], the correlation filters in 1D-CFA are designed in the 1D feature space. Designing correlation filters in the low-dimensional PCA subspace makes them less sensitive to noise.

Table 1
Summary of the notations used

Notations	Description
N	Number of training samples for all classes
N_l	Number of training samples for class l
L	Number of classes
p	Dimensionality of low-dimensional feature
$\vec{h}$	1D correlation filter in the space domain
$\vec{\mathbf{H}}$	1D correlation filter in the frequency domain
$\mathbf{Y}_l^{\mathrm{I}} = [\vec{\mathbf{Y}}_1^{\mathrm{I}}, \ldots, \vec{\mathbf{Y}}_{N_l}^{\mathrm{I}}]$	Intra-class transformed feature matrix, where $\vec{\mathbf{Y}}_i^{\mathrm{I}}$ is the 1D Fourier transform of intra-class low-dimensional feature $\vec{y}_i^{\mathrm{I}}$ for class l
$\mathbf{Y}_l^{\mathrm{E}} = [\vec{\mathbf{Y}}_1^{\mathrm{E}}, \ldots, \vec{\mathbf{Y}}_{N-N_l}^{\mathrm{E}}]$	Extra-class transformed feature matrix, where $\vec{\mathbf{Y}}_i^{\mathrm{E}}$ is the 1D Fourier transform of extra-class low-dimensional feature $\vec{y}_i^{\mathrm{E}}$for class l

2. A new correlation filter is proposed. A new correlation filter called optimal extra-class origin output tradeoff filter (OEOTF) which focuses on the origin correlation outputs is proposed. Two related correlation filters called minimum average extra-class origin correlation output energy (MAEOCE) filter and minimum extra-class origin variance synthetic discriminant function (MEOVSDF) filter are also presented. Extensive experimental results show that OEOTF is very effective for feature vector extraction.

The rest of the paper is organized as follows: Section 2 briefly reviews original CFA (2D-CFA) and widely used correlation filters. In Section 3, 1D correlation filter based CFA (1D-CFA) and OEOTF are discussed in detail. In Section 4, extensive experimental results on the FERET, AR, and FRGC (face recognition grand challenge) face databases are given. Comparisons between different linear subspace learning methods are also shown. Finally, conclusions are provided in Section 5.

For convenience, important notations used throughout the rest of the paper are listed in Table 1. Vectors are denoted by an arrow on top of the alphabet. Bold and upper case symbols refer to the frequency plane term while light and lower case symbols represent quantities in the space domain.

2. 2D correlation filter based class-dependence feature analysis

During the training stage, a set of 2D-CFA projection vectors (correlation filters) is generated and all of these correlation filters are used for feature vector extraction. More precisely, a specific correlation filter which discriminates one class from all other classes in the training set is designed for each face class. Overall a bank of class-dependence correlation filters is obtained [10]. And a new face image evaluated on all the correlation filters generates a feature vector in which each component represents the similarity between the new face image and a certain face class in the training set. See Fig. 1 for illustration.

The key of 2D-CFA is the design of correlation filters. Correlation filters work in the frequency domain (i.e. the 2D Fourier transforms of images). On the other hand, the phase spectrum is usually believed to contain more structural information in images that derive human perception than the magnitude spectrum [9]. Therefore, by going to the frequency domain, phase information is directly modeled by correlation filters.

The most simple correlation filter is known as the matched filter which is simply the complex conjugate of the 2D Fourier transform of the reference pattern. It has been shown that the matched filter is optimal for detecting a pattern which is the addition of the reference pattern and white noise [16]. However, in applications like face

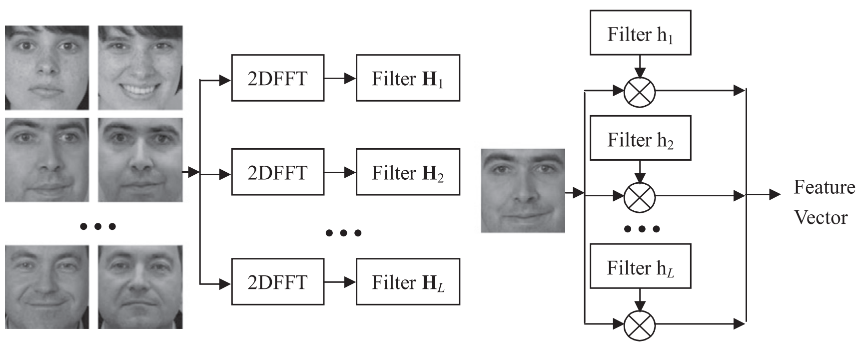

Fig. 1. Training of correlation filters (left) and feature vector extraction (right) in 2D-CFA. Note that 2D fast Fourier transform (2DFFT) is an efficient algorithm to compute 2D discrete Fourier transform (2DDFT).

recognition, due to variations of illumination, expression and age, etc., the probe face image is not simply the reference face image corrupted by additive white noise. Therefore, the matched filter is not suitable for face recognition problem.

Mahalanobis et al. [13] proposed the MACE filter. The objective of MACE filter is to minimize the average energy of the correlation outputs while satisfying correlation peak (origin correlation output) amplitude constraints.

More specifically, the average energy of the correlation outputs is $\vec{\mathbf{H}}^{+}\mathbf{D}\vec{\mathbf{H}}$, where $\vec{\mathbf{H}}$ represents the correlation filter in the frequency domain and $\mathbf{D}$ is a diagonal matrix whose diagonal entries are the average power spectrum of all N training images. '+' denotes conjugate transpose. And the linear constraint is $\mathbf{X}^{+}\vec{\mathbf{H}}=\vec{c}$, where $\mathbf{X}=[\vec{\mathbf{X}}_1,\vec{\mathbf{X}}_2,\ldots,\vec{\mathbf{X}}_N]$ and $\vec{\mathbf{X}}_i$ is the vector version of the 2D Fourier transform of the i-th training image. $\vec{c}=[c_1,c_2,\ldots,c_N]^{\mathrm{T}}$ is an $N\times 1$ vector and c_i denotes the origin correlation output of the i-th training image. Specifically, c_i is equal to 1 for intra-class training samples and 0 for extra-class training samples.

Based on the above optimization criterion, the optimum solution of MACE filter can be shown to be [13]

$$\vec{\mathbf{H}}_{\mathrm{MACE}}=\mathbf{D}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{D}^{-1}\mathbf{X})^{-1}\vec{c} \tag{1}$$

Kumar [14] proposed the MVSDF filter. MVSDF filter minimizes the correlation output noise variance $\vec{\mathbf{H}}^{+}\mathbf{C}\vec{\mathbf{H}}$, where $\mathbf{C}$ is a diagonal matrix whose diagonal elements represent the noise power spectral density while satisfying the correlation peak amplitude constraints. The solution of MVSDF filter is [14]

$$\vec{\mathbf{H}}_{\mathrm{MVSDF}}=\mathbf{C}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{C}^{-1}\mathbf{X})^{-1}\vec{c} \tag{2}$$

In order to produce sharp correlation peaks, MACE filter emphasizes high spatial frequencies which make MACE filter very susceptible to the input noise. On the other hand, MVSDF filter emphasizes low spatial frequencies to reduce noise. OTF [15] combines MACE filter and MVSDF filter together to produce sharp correlation peaks and suppress noise. The optimum solution of OTF is [15]

$$\vec{\mathbf{H}}_{\mathrm{OTF}}=\mathbf{T}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{T}^{-1}\mathbf{X})^{-1}\vec{c} \tag{3}$$

where $\mathbf{T}=\alpha\mathbf{D}+\sqrt{1-\alpha^2}\mathbf{C}$, $0\leqslant\alpha\leqslant 1$, is a parameter that controls the tradeoff. $\alpha=0$ leads to MVSDF filter and $\alpha=1$ leads to MACE filter. OTF has been widely used in the face recognition problem [9,10,12].

3. 1D correlation filter based class-dependence feature analysis

In this section, we first introduce the framework of 1D-CFA in Section 3.1. Then, the detailed derivation of the new correlation filter is presented in Section 3.2. Lastly, similarity measure and discussions are given in Sections 3.3 and 3.4, respectively.

3.1. 1D-CFA

On the whole, the framework of 1D-CFA is similar to 2D-CFA except that 1D-CFA designs 1D correlation filters in the low-dimensional PCA subspace.

After original feature extraction (original image matrix) and vectorization, the face image is represented as a high-dimensional vector. Then, PCA is performed to reduce the data dimensionality. Finally, in the low-dimensional subspace, correlation filters are designed by using the 1D Fourier transforms of the low-dimensional features (considered as 1D signals). See Fig. 2 for illustration.

Once correlation filters are designed for each face class in the training set, the feature vector of the face image can be extracted as shown in Fig. 3. To be specific, the feature vector is derived by the inner products of low-dimensional feature and all designed 1D correlation filters that are represented in terms of space domain.

It is worth noting the difference between 1D-CFA we propose in this study and TCF-CFA ($M=1$) [12], where correlation filters are designed without any dimensionality reduction step. Experimental results show that PCA is a necessary and effective step for 1D-CFA to obtain a good performance.

3.2. Optimal extra-class origin output tradeoff filter (OEOTF)

3.2.1. Background

1D-CFA needs to design a correlation filter for each face class in the training set. Suppose the correlation filter designed for the l-th class is $\vec{h}$. Let $\vec{g}_i(n)$ denote the correlation output produced by $\vec{h}$ in response to $\vec{y}_i$

$$\vec{g}_i(n)=\vec{y}_i(n)\odot\vec{h}(n) \tag{4}$$

where '$\odot$' stands for the correlation function of two 1D signals. $\vec{y}_i$ is the low-dimensional feature of the i-th training image in the PCA transformed subspace.

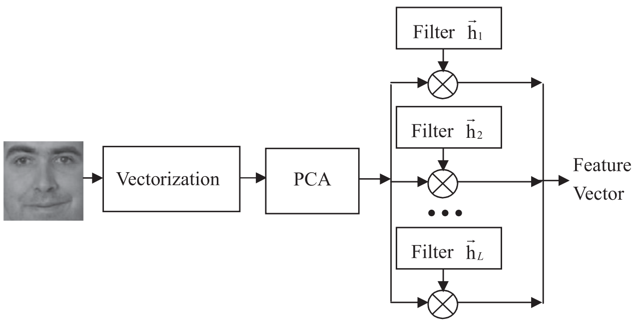

Fig. 3. Feature vector extraction in 1D-CFA.

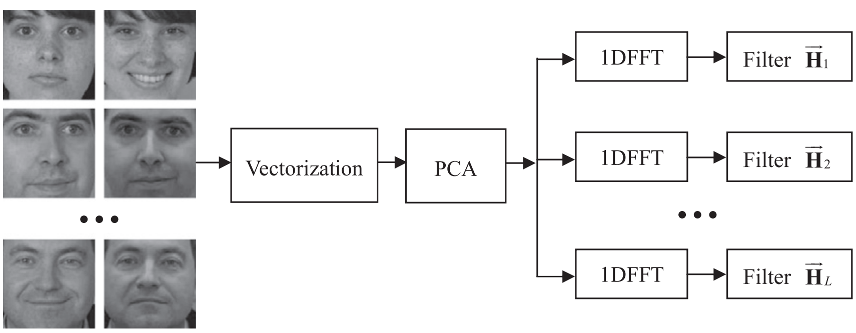

Fig. 2. Training of correlation filters in 1D-CFA. Note that 1D fast Fourier transform (1DFFT) is an efficient algorithm to compute 1D discrete Fourier transform (1DDFT).

The output can also be expressed using the frequency domain representations of $\vec{y}_i$ and $\vec{h}$:

$$\vec{g}_i(n) = \sum_{k=0}^{p-1} \vec{\mathbf{Y}}_i(k)^* \cdot \vec{\mathbf{H}}(k) e^{j2\pi kn/p} \tag{5}$$

where $\vec{\mathbf{Y}}_i(k)$, $\vec{\mathbf{H}}(k)$ are the 1D Fourier transforms of $\vec{y}_i$ and $\vec{h}$, respectively. '$*$' denotes the conjugate operation.

Notice that the point $\vec{g}_i(0)$ is often referred to the origin correlation output or correlation peak. According to Eq. (5), $\vec{g}_i(0)$ can be represented as the inner product of $\vec{\mathbf{Y}}_i$ and $\vec{\mathbf{H}}$. That is, $\vec{g}_i(0) = \sum_{k=0}^{p-1} \vec{\mathbf{Y}}_i(k)^* \cdot \vec{\mathbf{H}}(k) = \vec{\mathbf{Y}}_i^+ \vec{\mathbf{H}}$.

3.2.2. OEOTF

From Fig. 3, we see that only the origin correlation outputs (inner products of input feature and correlation filters) are used during the extraction process of feature vector. However, OTF [15] aims at the optimization of the whole correlation output plane. As a result, the optimization criterion of OTF is not consistent with the extraction of feature vector. Motivated by it, we develop a new correlation filter called OEOTF which only optimizes the origin correlation outputs.

The basic idea of OEOTF is to optimize two criterions of the origin correlation outputs for extra-class samples under the constraints on the origin correlation outputs for intra-class samples. To be specific, OEOTF is derived by combining MAEOCE filter and MEOVSDF filter.

The objective of MAEOCE filter is to minimize the origin correlation output energy for extra-class low-dimensional features, that is,

$$\begin{aligned}
\min_{\vec{\mathbf{H}}} \frac{1}{N-N_l} \sum_{i=1}^{N-N_l} |\vec{g}_i(0)|^2 &= \min_{\vec{\mathbf{H}}} \frac{1}{N-N_l} \sum_{i=1}^{N-N_l} \left| \sum_{k=0}^{p-1} \vec{\mathbf{Y}}_i^{\mathrm{E}}(k)^* \cdot \vec{\mathbf{H}}(k) \right|^2 \\
&= \min_{\vec{\mathbf{H}}} \frac{1}{N-N_l} \sum_{i=1}^{N-N_l} |\vec{\mathbf{Y}}_i^{\mathrm{E}+} \vec{\mathbf{H}}|^2 \\
&= \min_{\vec{\mathbf{H}}} \frac{1}{N-N_l} \sum_{i=1}^{N-N_l} \vec{\mathbf{H}}^+ \vec{\mathbf{Y}}_i^{\mathrm{E}} \vec{\mathbf{Y}}_i^{\mathrm{E}+} \vec{\mathbf{H}} \\
&= \min_{\vec{\mathbf{H}}} \vec{\mathbf{H}}^+ \left(\frac{1}{N-N_l} \sum_{i=1}^{N-N_l} \vec{\mathbf{Y}}_i^{\mathrm{E}} \vec{\mathbf{Y}}_i^{\mathrm{E}+} \right) \vec{\mathbf{H}} \\
&= \min_{\vec{\mathbf{H}}} \vec{\mathbf{H}}^+ \mathbf{R}_{\mathrm{Y}} \vec{\mathbf{H}}
\end{aligned} \tag{6}$$

where

$$\mathbf{R}_{\mathrm{Y}} = \frac{1}{N-N_l} \sum_{i=1}^{N-N_l} \vec{\mathbf{Y}}_i^{\mathrm{E}} \vec{\mathbf{Y}}_i^{\mathrm{E}+}$$

$\vec{\mathbf{Y}}_i^{\mathrm{E}}, i = 1, \ldots, N - N_l$ are the 1D Fourier transforms of extra-class low-dimensional features for class l.

For intra-class low-dimensional features in the l-th class, the constraints are that the origin correlation outputs are all equal to 1, that is,

$$\mathbf{Y}_l^{\mathrm{I}+} \vec{\mathbf{H}} = \vec{u} \tag{7}$$

where $\mathbf{Y}_l^{\mathrm{I}} = [\vec{\mathbf{Y}}_1^{\mathrm{I}}, \ldots, \vec{\mathbf{Y}}_{N_l}^{\mathrm{I}}]$. $\vec{\mathbf{Y}}_i^{\mathrm{I}}, i = 1, \ldots, N_l$, are the 1D Fourier transforms of intra-class low-dimensional features for class l. $\vec{u} = [1, \ldots, 1]^{\mathrm{T}}$ is an $N_l \times 1$ vector.

Therefore, the objective of MAEOCE filter is to

$$\min_{\vec{\mathbf{H}}} \vec{\mathbf{H}}^+ \mathbf{R}_{\mathrm{Y}} \vec{\mathbf{H}} \tag{8}$$

subject to the linear constraint $\mathbf{Y}_l^{\mathrm{I}+} \vec{\mathbf{H}} = \vec{u}$.

The objective of MEOVSDF filter is to minimize the origin correlation output noise variance for extra-class low-dimensional features while satisfying linear constraint for intra-class low-dimensional features. Therefore, the objective of MEOVSDF filter is to

$$\min_{\vec{\mathbf{H}}} \vec{\mathbf{H}}^+ \mathbf{C} \vec{\mathbf{H}} \tag{9}$$

subject to the linear constraint $\mathbf{Y}_l^{\mathrm{I}+} \vec{\mathbf{H}} = \vec{u}$, where $\mathbf{C}$ is a diagonal matrix whose diagonal elements represent the noise power spectral density of extra-class low-dimensional features. $\mathbf{C}$ is an identity matrix if the input noise is modeled as additive white noise.

MAEOCE filter emphasizes high spatial frequencies to produce sharp correlation peaks (since narrow functions in the correlation output should correspond to broad support in frequency domain [9]) while MEOVSDF filter suppresses high spatial frequencies to achieve noise tolerance. To trade off two criterions, OEOTF is derived by combining MAEOCE filter and MEOVSDF filter. Thus, the objective of MEOVSDF filter is to

$$\min_{\vec{\mathbf{H}}} \vec{\mathbf{H}}^+ \mathbf{T} \vec{\mathbf{H}} \tag{10}$$

subject to the linear constraint $\mathbf{Y}_l^{\mathrm{I}+} \vec{\mathbf{H}} = \vec{u}$, where $\mathbf{T} = \alpha \mathbf{R}_{\mathrm{Y}} + \sqrt{1-\alpha^2} \mathbf{C}$, $0 \leqslant \alpha \leqslant 1$. $\alpha = 0$ leads to MEOVSDF filter and $\alpha = 1$ leads to MAEOCE filter.

The solution to this problem can be founded by using the method of Lagrange multipliers. Similar derivation of the constrained optimization problem can be found in Refs. [13,14]. The optimum solution of Eq. (10) can be shown to be

$$\vec{\mathbf{H}}_{\mathrm{OEOTF}} = \mathbf{T}^{-1} \mathbf{Y}_l^{\mathrm{I}} (\mathbf{Y}_l^{\mathrm{I}+} \mathbf{T}^{-1} \mathbf{Y}_l^{\mathrm{I}})^{-1} \vec{u} \tag{11}$$

3.3. Similarity measure

After feature vector extraction, it needs to design a classifier. In this study, the simple nearest neighbor classifier is applied. There are four commonly used similarity measures: the L_1 norm, the Euclidean distance (L_2 norm), the Mahalanobis distance, and the whitened cosine distance. For CFA, the whitened cosine distance among these four is observed to provide the best performance [9,10]. We use whitened cosine distance in all the experiments. The computation of whitened cosine distance is shown

$$s(\vec{x}, \vec{y}) = \frac{-(\vec{x} \cdot \vec{y})}{\|\vec{x}\| \|\vec{y}\|} \tag{12}$$

where $\vec{x}$ and $\vec{y}$ are vectors. '$\|\cdot\|$' represents L_2 norm.

3.4. Discussions

Let us compare the differences between OTF and OEOTF. Based on the above analysis, both OTF and OEOTF try to produce sharp correlation peaks while achieving noise tolerance. However, the design criterions of the two filters are totally different. OTF optimizes the tradeoff of the average correlation output energy and output noise variance under the constraints on the origin correlation outputs for all training samples, whereas OEOTF minimizes the tradeoff of the average origin correlation output energy and origin output noise variance for extra-class samples, while the origin correlation outputs for intra-class samples are fixed. In contrast to OTF which concerns the entire correlation output plane, OEOTF just focuses on the origin correlation outputs. Furthermore, the constraints of OEOTF are easier than OTF, since only the origin correlation outputs for intra-class samples are considered. In addition, note that OEOTF can also be designed in the space domain, since only the origin correlation outputs are used in the optimization criterion. However, to compare with OTF, we mainly discuss OEOTF in the frequency domain.

From the perspective of pattern recognition, the design criterion of OEOTF is simpler than OTF (without considering all the outputs in the correlation output plane). The generalization capability of OEOTF may be better. Furthermore, the optimization criterion of OEOTF is more consistent with the extraction of feature vector than that of OTF. Therefore, OEOTF is more effective for feature vector extraction.

It should be pointed out that traditional correlation filters are designed in the 2D image space due to their shift-invariance property in the image space [9,10]. Hence, designing correlation filters in the 1D PCA subspace seems to violate the shift-invariance property. The main reason why we can do so lies in that the face images provided in the face database are well centered, rendering the shift-invariance advantages of correlation filter most irrelevant [9]. In fact, from Fig. 3, it can also be observed that only the origin correlation outputs are used for feature extraction. Therefore, the shift-invariance property is not important since the face images are well centered. In this study, all face images in the databases are normalized so that they are well aligned.

4. Experiments

In this section, the recognition accuracy of 1D-CFA is evaluated on three well-known benchmark face databases (FERET [17], AR [18], and FRGC [19]). We compare 1D-CFA with the performance of other state-of-the-art face recognition methods including Eigenface [2], Fisherface [3], LDA/FKT [4], C-LDA [5], MMSD[6], Laplacianface [7], TCF-CFA ($M = 1$) [12], and 2D-CFA [9,10].

All the facial images are normalized according to the eye coordinates for scaling, translation and rotation, such that the eye centers are in fixed position. All the images are cropped to the size of 64×64. And histogram equalization is applied to the face images for photometric normalization. No further preprocessing is done. The nearest neighbor classifier is employed for all methods. All methods use whitened cosine distance measure except Eigenface which employs the standard Mahalanobis distance measure.

4.1. Face databases and parameters setting

The FERET face database [17] has become a standard database for testing and evaluating state-of-the-art face recognition algorithms. We test on a subset of the FERET face database. This subset includes 800 images of 200 individuals (each one has four images). Several examples are given in Fig. 4. The AR face database [18] contains over 4000 face images of 126 people, including frontal view of faces with different facial expressions, lighting conditions and occlusions. The images of 120 individuals were taken in two sessions (separated by two weeks) and each session contains 13 color images. We select 14 face images (each session contains seven images) from each of these 120 individuals. Fig. 5 shows sample images of one person. We select 6000 images for 300 individuals (each one has 20 images) from FRGC version 2.0 face database [19]. The face images were captured in both controlled and uncontrolled conditions with harsh illumination and expression variations. Fig. 6 shows sample images of one person.

For all face databases, m different images per individual are randomly chosen to form the training set. The rest of the images in the database are used for testing. Totally, 20 experiments are performed. The final result is the average recognition rate over 20 random training sets. For FERET database, m is chosen as 2 and 3. For both AR and FRGC databases, m is chosen as 2, 4, and 6.

According to Ref. [20], the effectiveness of Fisherface is heavily dependent on the number of principal components used in the PCA stage. We have implemented Fisherface by using PCA to reduce the dimensionality to $\eta(N - L)$, where $0 < \eta \leqslant 1$. N is the number of all training samples and L is the number of classes. We present the best result of Fisherface with η varying from 0.8 to 1. For MMSD, the parameter c which can be used to adjust to balance between the range of the between-class scatter matrix and the null space of the within-class scatter matrix is empirically chosen to be 10 [6]. For Laplacianface, supervised mode with Gaussian kernel is applied. And the tradeoff parameter α in both OTF and OEOTF is set to 10^{-9} which shows good performances for all CFA methods. The reduced dimensionality p of PCA subspace in 1D-CFA is chosen as $N - 1$.

4.2. Experimental results on face databases

Table 2 shows the top mean recognition rate as well as the standard deviation achieved by each method and the corresponding dimensionality of reduced subspace. Table 3 compares 1D-CFA and 2D-CFA with different correlation filters on three face databases.

From Tables 2 and 3, the main observations from the performance comparisons include:

- PCA is an effective preprocessing step for 1D-CFA. Since some structure information is lost during the vectorization step, TCF-CFA ($M = 1$) with OTF obtains a worse performance than OTF based 2D-CFA. However, by taking PCA as a preprocessing step, OTF based 1D-CFA can achieve a comparable performance over OTF based 2D-CFA. The reason lies in that PCA is an effective method to reduce the noise and extract the most representational features [21].
- OEOTF based 1D-CFA outperforms other state-of-the-art face recognition methods. We can see that OEOTF based 1D-CFA consistently performs better than other methods, especially on the FRGC face database. Note that the performances of C-LDA and MMSD are comparable to OEOTF based 1D-CFA on the FERET and AR databases but more dimensions are needed. Although useful structure information is lost in 1D-CFA due to the vectorization of the 2D image matrix, OEOTF based 1D-CFA still outperforms OTF based 2D-CFA. More specifically, compared with OTF based 2D-CFA, the recognition rate of OEOTF based 1D-CFA for the AR database increases about 3%; the recognition rate increases about 10–18% for the FERET and FRGC databases. Hence, the generalization performance of OEOTF based 1D-CFA is significantly better than OTF based 2D-CFA.
- Joint optimization of 1D-CFA with MAEOCE filter and MEOVSDF filter can achieve much better recognition performance than

Fig. 4. Samples of the cropped images of two persons on the FERET face database.

Fig. 5. Samples of the cropped images of one person on the AR face database.

Fig. 6. Samples of the cropped images of one person on the FRGC face database.

Table 2
Top recognition rates (mean±std. dev.%) and corresponding reduced dimension of different algorithms on the FERET, AR, and FRGC face databases

Algorithm	FERET (2 train)	FERET (3 train)	AR (2 train)	AR (4 train)	AR (6 train)	FRGC (2 train)	FRGC (4 train)	FRGC (6 train)
Eigenface	63.90 ± 1.6 (397)	66.75 ± 3.0 (595)	73.90 ± 2.7 (236)	86.18 ± 1.1 (101)	90.98 ± 1.0 (101)	47.38 ± 0.8 (127)	64.35 ± 0.8 (305)	73.30 ± 0.7 (323)
Fisherface	73.58 ± 1.5 (73)	79.03 ± 2.1 (172)	80.46 ± 1.7 (74)	90.92 ± 1.0 (119)	92.96 ± 1.0 (119)	48.13 ± 1.1 (161)	57.31 ± 1.2 (251)	62.90 ± 1.1 (287)
Laplacianface	74.69 ± 1.7 (199)	82.15 ± 2.2 (199)	85.29 ± 1.3 (119)	92.17 ± 0.7 (119)	93.76 ± 0.6 (119)	53.31 ± 0.9 (299)	65.12 ± 1.0 (305)	69.05 ± 0.6 (305)
LDA/FKT	74.94 ± 1.5 (199)	82.15 ± 2.2 (199)	85.39 ± 1.4 (119)	92.15 ± 0.8 (119)	93.67 ± 0.6 (119)	51.63 ± 5.6 (299)	58.62 ± 0.9 (299)	66.37 ± 2.8 (299)
C-LDA	80.55 ± 1.1 (217)	91.57 ± 1.6 (226)	87.38 ± 1.4 (155)	93.94 ± 1.7 (236)	95.54 ± 0.6 (236)	59.81 ± 1.1 (359)	74.22 ± 0.9 (413)	78.73 ± 0.8 (521)
MMSD	80.40 ± 1.4 (235)	91.65 ± 1.4 (289)	86.86 ± 1.6 (173)	94.31 ± 0.8 (290)	95.34 ± 0.5 (416)	60.91 ± 0.9 (359)	73.52 ± 0.8 (449)	77.07 ± 0.8 (629)
2D-CFA (OTF)	74.95 ± 1.5 (200)	82.15 ± 2.2 (200)	85.40 ± 1.4 (120)	92.16 ± 0.8 (120)	93.69 ± 0.6 (120)	54.26 ± 0.8 (300)	66.01 ± 1.0 (300)	69.55 ± 0.6 (300)
TCF-CFA (M=1) (OTF)	72.08 ± 1.7 (200)	79.85 ± 1.9 (200)	84.35 ± 1.3 (120)	91.65 ± 0.9 (120)	93.31 ± 0.6 (120)	52.73 ± 0.8 (300)	64.62 ± 0.9 (300)	68.34 ± 0.8 (300)
1D-CFA (OTF)	75.07 ± 1.5 (200)	82.25 ± 2.2 (200)	85.38 ± 1.4 (120)	92.13 ± 0.7 (120)	93.68 ± 0.6 (120)	54.32 ± 0.8 (300)	65.99 ± 1.0 (300)	69.50 ± 0.6 (300)
1D-CFA (OEOTF)	83.65 ± 1.4 (200)	92.40 ± 1.5 (200)	87.83 ± 1.4 (120)	95.40 ± 0.6 (120)	96.84 ± 0.5 (120)	62.98 ± 0.9 (300)	81.10 ± 0.7 (300)	87.58 ± 0.6 (300)

Table 3
Top recognition rates (mean±std. dev.%) and corresponding reduced dimension of 2D-CFA and 1D-CFA with different correlation filters on the FERET, AR, and FRGC face databases

Algorithm	FERET (2 train)	FERET (3 train)	AR (2 train)	AR (4 train)	AR (6 train)	FRGC (2 train)	FRGC (4 train)	FRGC (6 train)
2D-CFA (MACE)	36.24 ± 1.4 (200)	44.48 ± 2.4 (200)	62.57 ± 1.1 (120)	77.85 ± 0.9 (120)	83.32 ± 1.0 (120)	34.19 ± 0.7 (300)	47.53 ± 0.8 (300)	53.74 ± 0.6 (300)
2D-CFA (MVSDF)	74.91 ± 1.7 (200)	82.10 ± 2.3 (200)	83.39 ± 1.4 (120)	91.15 ± 1.1 (120)	93.08 ± 0.7 (120)	53.35 ± 1.2 (300)	64.27 ± 0.9 (300)	67.52 ± 1.0 (300)
2D-CFA (OTF)	74.95 ± 1.5 (200)	82.15 ± 2.2 (200)	85.40 ± 1.4 (120)	92.16 ± 0.8 (120)	93.69 ± 0.6 (120)	54.26 ± 0.8 (300)	66.01 ± 1.0 (300)	69.55 ± 0.6 (300)
1D-CFA (MAEOCE)	75.02 ± 1.5 (200)	82.25 ± 2.3 (200)	85.39 ± 1.4 (120)	92.12 ± 0.7 (120)	93.15 ± 0.6 (120)	54.25 ± 0.9 (300)	65.97 ± 1.0 (300)	69.50 ± 0.6 (300)
1D-CFA (MEOVSDF)	14.55 ± 1.6 (200)	20.80 ± 1.9 (200)	25.73 ± 1.2 (120)	40.92 ± 1.0 (120)	51.76 ± 1.5 (120)	14.95 ± 0.7 (300)	24.81 ± 0.8 (300)	32.85 ± 0.8 (300)
1D-CFA (OEOTF)	83.65 ± 1.4 (200)	92.40 ± 1.5 (200)	87.83 ± 1.4 (120)	95.40 ± 0.6 (120)	96.84 ± 0.5 (120)	62.98 ± 0.9 (300)	81.10 ± 0.7 (300)	87.58 ± 0.6 (300)

with each correlation filter used separately. OEOTF based 1D-CFA is more effective for face recognition than MAEOCE filter and MEOVSDF filter based 1D-CFA. Experimental results also indicate that OTF based 2D-CFA still outperforms MACE filter and MVSDF filter based 2D-CFA. Therefore, joint optimization of two different correlation filters can improve the capability of CFA to reject false alarms.

Overall, OEOTF based 1D-CFA has a good performance for face recognition.

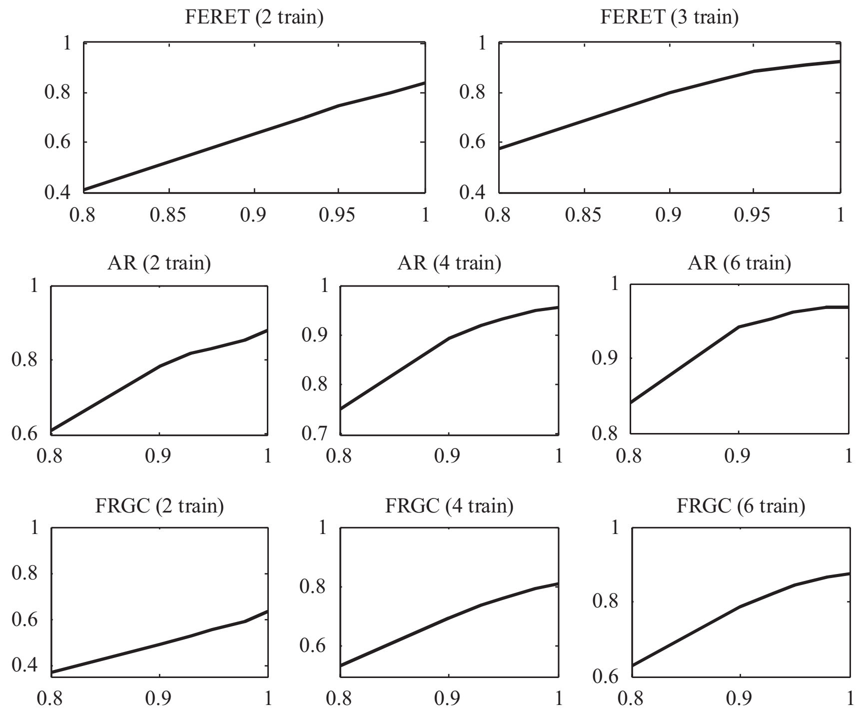

Fig. 7. Recognition rate vs. PCA energy ratio on the FERET, AR, and FRGC face databases.

Table 4
Recognition rate (mean%) vs. PCA energy ratio for OEOTF based 1D-CFA

PCA energy ratio (%)	FERET (2 train)	FERET (3 train)	AR (2 train)	AR (4 train)	AR (6 train)	FRGC (2 train)	FRGC (4 train)	FRGC (6 train)
80	41.04	57.42	60.97	75.01	84.01	36.80	53.70	62.77
90	63.38	79.58	78.28	89.40	94.09	49.06	69.55	78.74
93	69.75	85.45	81.59	91.92	95.37	52.76	73.71	82.42
95	74.30	88.22	83.00	93.37	96.08	55.29	76.17	84.38
98	80.06	90.77	85.41	94.83	96.74	59.10	79.20	86.66
100	83.65	92.40	87.83	95.40	96.84	62.98	81.10	87.58

4.3. Influence of PCA energy ratio on recognition performance

In the previous experiments, the PCA energy ratio that is based on the eigenvalue of Karhunen–Loeve transform is taken to be 1 (i.e. PCA subspace dimensionality is equal to $N-1$) for OEOTF based 1D-CFA. In this subsection, we investigate the influence of PCA energy ratio on the final recognition performance. Fig. 7 illustrates the plot of top recognition rate versus PCA energy ratio on three face databases. Some quantitative results are also given in Table 4.

From Fig. 7 and Table 4, it can be observed that OEOTF based CFA obtains the best recognition rate when the PCA energy ratio is equal to 1. Thus, all PCA dimensionality should be used to design a robust OEOTF.

4.4. Comparisons between different linear subspace learning methods

It is worth remarking upon the performance comparisons between different linear subspace learning methods including Eigenface [2], Fisherface [3], LDA/FKT [4], C-LDA [5], MMSD [6], Laplacianface [7], 2D-CFA [9,10], and 1D-CFA.

1. Eigenface is based on PCA which is optimal in sense of minimum mean squared error (MMSE) for reconstruction. Thus, PCA may not be optimal for classification problem. CFA emphasizes the outputs of intra-class samples while suppressing the outputs of extra-class samples. Therefore, CFA can extract discriminant features to distinguish different classes well.
2. Fisherface, LDA/FKT, C-LDA, MMSD, and Laplacianface methods are all based on the second order statistics of data while the information in the higher-order statistics is not considered. On the other hand, CFA directly models the phase information which captures the high-order statistics in the frequency domain.
3. One property of CFA is that when new classes are added in the training set, all previous trained correlation filters do not require retraining [10]. The property is very useful in the real face recognition applications.
4. While traditional linear subspace learning methods [2–7] need to determine the optimal reduced dimensionality (ORD), the ORD of CFA is equal to the number of classes in the training set.
5. Different from 2D-CFA which designs correlation filters in the 2D image space, correlation filters in 1D-CFA are designed in the low-dimensional PCA subspace and this makes them less sensitive to noise. Another advantage of designing correlation filters in the low-dimensional subspace is that the computational complexity can be reduced. Take OEOTF for example, assuming that OEOTF is designed in the 2D image space, the dimensionality of $\mathbf{T}$ will be $d^2 \times d^2$ (suppose the image size is $d \times d$). Note that $\mathbf{T}$ is not sparse. Therefore, the computational complexity of Eq. (11) is high when the face image is large. On the other hand, the dimensionality of $\mathbf{T}$ is just $(N-1) \times (N-1)$, where N is the number of training samples, with respect to OEOTF based 1D-CFA. Compared with

OEOTF based 2D-CFA, the computational complexity of Eq. (11) in OEOTF based 1D-CFA is greatly reduced for the small sample size problem in this study.

5. Conclusions

In this paper, we present a novel one-dimensional correlation filter based class-dependence feature analysis (1D-CFA) method for robust and effective face recognition. Furthermore, a new correlation filter called optimal extra-class origin output tradeoff filter (OEOTF) is developed in 1D-CFA. By focusing on the origin correlation outputs, OEOTF can extract discriminant features effectively. Moreover, we analyze the reason why correlation filters can be designed in the low-dimensional PCA subspace for the face recognition problem. Extensive experimental results on benchmark face databases demonstrate the superiority of OEOTF based 1D-CFA.

The approach presented in this study can also be applied to other biometric recognition problems, such as iris recognition [22] and palmprint recognition [23]. However, some limitations of this study are: (1) the basic assumption of 1D-CFA is that all face images are well centered. In other words, 1D-CFA is effective for frontal face recognition. For face recognition with pose variations, 1D-CFA may no longer be accurate. (2) In this paper, PCA is used to derive low-dimensional features. However, whether it can be generalized to other dimensionality reduction methods needs further research. (3) 1D-CFA needs to design a correlation filter for each face class in the training set. Consequently, the computational complexity of 1D-CFA is too high when there are thousands of face classes in the training set. Further studies about this problem [24] are still necessary. (4) Recent works [9,24] also show that nonlinear correlation filters (such as kernel extension of OTF) can attain a good performance. How to extend OEOTF in the kernel form needs further investigation.

Acknowledgments

The authors would like to thank the associate editor and three anonymous reviewers for their critical and constructive comments and suggestions to improve the quality of the paper. They would also like to thank Dr. Deng Cai at UIUC for providing the source codes of Laplacianface algorithm. This work was supported by National Natural Science Foundation (NNSF: 60573148) and Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP-20060003102).

References

[1] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, Face recognition: a literature survey, ACM Comput. Surv. 35 (4) (2003) 399–459.
[2] M. Turk, A. Pentland, Eigenfaces for recognition, Int. J. Cognitive Neurosci. 3 (1) (1991) 71–86.
[3] P.N. Belhumer, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.
[4] S. Zhang, T. Sim, Discriminant subspace analysis: a Fukunaga–Koontz approach, IEEE Trans. Pattern Anal. Mach. Intell. 29 (10) (2007) 1732–1745.
[5] J. Yang, J.Y. Yang, Why can LDA be performed in PCA transformed space?, Pattern Recognition 36 (2) (2003) 563–566.
[6] F. Song, D. Zhang, D. Mei, Z. Guo, A multiple maximum scatter difference discriminant criterion for facial feature extraction, IEEE Trans. Syst. Man Cybern. Part B Cybern. 37 (6) (2007) 1599–1606.
[7] X. He, S. Yan, Y. Hu, P. Niyogi, H. Zhang, Face recognition using Laplacianfaces, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (2005) 328–340.
[8] K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, New York, 1990.
[9] B.V.K.V. Kumar, M. Savvides, C. Xie, Correlation pattern recognition for face recognition, Proc. IEEE. 94 (11) (2006) 1963–1976.
[10] C. Xie, M. Savvides, B.V.K.V. Kumar, Redundant class-dependence feature analysis based on correlation filters using FRGC2.0 data, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2005, pp. 153–158.
[11] B.V.K.V. Kumar, A. Mahalanobis, Recent advances in composite correlation filter designs, Asian J. Phys. 8 (1999) 407–420.
[12] Y. Yan, Y.-J. Zhang, Tensor correlation filter based class-dependence feature analysis for face recognition, Neurocomputing 12 (2008).
[13] A. Mahalanobis, B.V.K.V. Kumar, D. Casasent, Minimum average correlation energy filters, Appl. Opt. 26 (1987) 3630–3640.
[14] B.V.K.V. Kumar, Minimum variance synthetic discriminant functions, J. Opt. Soc. Am. A. 3 (1986) 1579–1584.
[15] P. Refregier, Filter design for optical pattern recognition: multi-criteria optimization approach, Opt. Lett. 15 (1990) 854–856.
[16] A. VanderLugt, Signal detection by complex spatial filtering, Proc. IEEE. 51 (7) (1963) 1016–1027.
[17] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (12) (2000) 1090–1104.
[18] A.M. Martinez, R. Benavente, The AR face database, CVC Technical Report #24, 1998.
[19] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, Overview of the face recognition grand challenge, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2005, pp. 947–954.
[20] J. Yang, A.F. Frangi, J.Y. Yang, D. Zhang, Z. Jin, KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 27 (2) (2005) 230–244.
[21] X. Wang, X. Tang, A unified framework for subspace face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (9) (2004) 1222–1228.
[22] J. Thornton, M. Savvides, B.V.K.V. Kumar, A Bayesian approach to deformed pattern matching of iris images, IEEE Trans. Pattern Anal. Mach. Intell. 29 (4) (2007) 596–606.
[23] P.H. Hennings, B.V.K.V. Kumar, M. Savvides, Palmprint classification using multiple advanced correlation filters and palm-specific segmentation, IEEE Trans. Inf. Forensics Secur. 2 (3) (2007) 613–622.
[24] B.V.K.V. Kumar, C. Xie, M. Savvides, Correlation filters for large population face recognition, Proc. SPIE 6539 (2006) 65390F1–65390F12.

About the Author—YAN YAN received the B.S. degree in Electronic Engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2004. Currently, he is a Ph.D. candidate from the Department of Electronic Engineering at Tsinghua University, Beijing, China. His research interests include computer vision, face recognition and machine learning. He has now published several papers on face recognition and image analysis.

About the Author—YU-JIN ZHANG received the Ph.D. degree in Applied Science from the State University of Liège, Liège, Belgium, in 1989. From 1989 to 1993, he was post-doc fellow and research fellow with the Department of Applied Physics and Department of Electrical Engineering at the Delft University of Technology, Delft, the Netherlands. In 1993, he joined the Department of Electronic Engineering at Tsinghua University, Beijing, China, where he is a professor of Image Engineering since 1997. In 2003, he spent his sabbatical year as visiting professor in the School of Electrical and Electronic Engineering at Nanyang Technological University (NTU), Singapore.
His research interests are mainly in the area of Image Engineering that includes image processing, image analysis and image understanding, as well as their applications. He is an enthusiastic promoter of this new discipline. Since 1996, he has made (for 13 consecutive years) a survey series of yearly bibliographies on Image Engineering in China, the results are published both nationally and internationally. He is also an active researcher in Image Engineering, with current interests on object segmentation from images and video, segmentation evaluation and comparison, moving object detection and tracking, face recognition, facial expression detection/classification, content-based image and video retrieval, information fusion for high-level image understanding, etc. Recently, he is in charge of a number of projects from National Nature Science Foundation, National High Technology Projects, Ministry of Education, etc. He has published extensively in image engineering and related areas, with more than 300 research papers, two monographs: "Image Segmentation" and "Content-based Visual Information Retrieval" (Science Press), and two edited books: "Advances in Image and Video Segmentation" and "Semantic-Based Visual Information Retrieval" (IRM Press). http://www.ee.tsinghua.edu.cn/~zhangyujin/.