Directionally classified eigenblocks for localized feature analysis in face recognition

Ho-Chul Shin  
Korea Advanced Institute of Science and Technology (KAIST)  
Division of Electrical Engineering  
Department of Electrical Engineering and Computer Science  
373-1 Guseong-dong, Yuseong-gu  
Daejeon 305-701, Republic of Korea  
E-mail: shc@kaist.ac.kr

Hae-Chul Choi  
Electronics and Telecommunications Research Institute (ETRI)  
Broadcasting Media Research Group  
161 Gajeong-dong, Yuseong-gu  
Daejeon, 305-701, Republic of Korea

Seong-Dae Kim  
Korea Advanced Institute of Science and Technology (KAIST)  
Division of Electrical Engineering  
Department of Electrical Engineering and Computer Science  
373-1 Guseong-dong, Yuseong-gu  
Daejeon 305-701, Republic of Korea

Abstract. A new local feature extraction method is introduced. The directionality of local facial regions is regarded as essential information for discriminating faces in our approach, which is motivated by the directional selectivity of the Gabor wavelet transformation, which has been preferred to others for face recognition. The discriminative directional information is forced to be compacted in a few coefficients by applying principal-component analysis with the support of directional classification in the discrete cosine transform domain. The local features extracted by our method are better at discriminating face patterns than previous ones, as was verified by comparison of class-separability results. Also, in face recognition simulations using rigid and flexible face matching strategies based on locally extracted features, our proposed method showed outstanding performance. © 2006 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2227000]

Subject terms: face recognition; local feature extraction; directional classification; eigenblock.

1 Introduction

During the last two decades, numerous algorithms have been proposed for face recognition and used in various applications: bankcard identification, access control, security monitoring, and surveillance. Recently, in addition to the well-known appearance-feature-based approaches, 1–14 local-feature-based approaches 5–14 have been widely studied for face recognition.

Dividing a face into several local regions and utilizing low-dimensional features extracted from each region in a localized observation is the conventional approach to face recognition, based on the hidden Markov model (HMM), neural networks, and elastic graph matching. But the generally used 2-D discrete cosine transform (DCT), localized principal component analysis (PCA), and Gabor wavelet transformation are not optimized for pattern discrimination, and a proficient method for pattern discrimination is required to achieve more reliable recognition performance.

For the local feature extraction, the 2-D DCT has been favored in many approaches 10,13,14 because of its small computational load, simplicity of implementation, and compatibility with still-image coding standards such as JPEG. Eickeler et al. 10 proposed a face recognition system based on 2-D DCT features and pseudo-2-D HMM with consideration of the recognition of faces that are included in an incompletely uncompressed image from a JPEG stream. Sanderson and Faliwal 13 introduced the DCT-mod2 facial feature extraction technique, which utilizes polynomial coefficients derived from 2-D DCT coefficients of spatially neighboring blocks, and Kohir and Desai 14 introduced the DCT-HMM approach which is intended to combine the advantages of DCT and HMM. Despite many practical advantages, the 2-D DCT coefficients themselves are not the best features for recognition, due to some limitations:

1. The unmodified DCT coefficients are necessarily sensitive to illumination variations, through the compensating dc component.
2. The DCT is good for compact representation of a local image pattern, but DCT coefficients themselves are not always good in pattern-discrimination-sense.
3. Fixed basis functions for 2D DCT cannot adapt to the statistical variation of target patterns.

The well-known Karhunen-Loève transform (KLT), based on PCA, is better than the DCT in data compression because the KLT uses statistically optimized eigenvectors for data representation. 15 For robust face recognition, McGuire and D’Eleuterio 9 used the KLT for analyzing features in a local scope, and an error-correcting neural network for classification tasks. They found eigenvectors of the set of image blocks randomly sampled from face images and called the eigenvectors eigenpixels, which may be thought of as localized principal components. On the other hand, modular PCA, which was proposed by Mohgaddam and
Pentland, is a method using localized eigenvectors estimated in specific regions of facial features. They tried to achieve recognition performance robust to partial occlusion through matching local facial features based on features extracted by localized eigenvectors around facial features. Alexi has done similar work for occlusion-robust face recognition, but he used a probabilistic approach rather than a voting space. Though the KLT provides higher data compression performance than the DCT, it imposes a heavier computational load, and as the target scope becomes smaller, its performance becomes similar to that of the DCT.

The Gabor wavelet transformation (GWT) is the most widely used local feature extraction method in face recognition. Generally, approaches based on elastic graph matching (EGM) adopt the GWT for extracting node-labeling features, because the Gabor features have good locality and provide observations with several resolutions. Liu utilized the Gabor wavelets for preliminary dimensionality reduction before applying kernel PCA. The good discrimination power of the Gabor wavelets is thought to originate from the directional selectivity in the kernels, while the variability in resolution is helpful for hierarchical coarse-to-fine matching. But in spite of these advantages of the Gabor wavelets, their high computational complexity and high sensitivity to feature positions make them difficult to apply in practice.

The most valuable local features for face recognition should be the most compressed representation of the most discriminative information in a local facial region. When concentrating on a small region in a human face, there is no complex information. Simple edges, intensity gradients, or spots are almost all that can be found in a local scope. Previous approaches based on the localized DCT or KLT tried to represent these simple local patterns compactly and achieve discrimination between faces at the level of a full pattern that consists of several local features. But in fact, discriminative information is definable and extractable even in a local scope, as is proved by the higher recognition performance of Gabor wavelets in previous approaches. Suppression of the illumination-change-sensitive dc component and detailed interpretation of discriminative directional patterns are the reasons for the higher performance of the Gabor wavelets. On the other hand, the Gabor wavelets are not optimal for compact representation of information, which leads to high dimensionality of Gabor-wavelet-based local features.

In our approach, localized transitions, such as edges or gradients, were assumed to contain the most discriminative information in a local facial region, and the orientation of transitions was thought to be a key to discriminating local patterns, in view of the directionally distinct shapes of the Gabor wavelets. Therefore the compact representation of discriminative directional information is the objective of our proposed local feature extraction method in this paper. Firstly, the proposed method estimates the strength of directionality of a local pattern, and selects patterns with strong directionality. Here, directionality means the concentrated disposition of transitions in a local region to a direction. Secondly, discriminative directional information is represented as a weight vector by using our proposed eigenkernels. In contrast with the conventional KLT, the proposed eigenkernels are estimated from sets of same-direction patterns, and each direction has one orthogonal eigenkernel set that is optimal for describing a pattern of that direction. Classified eigen blocks (CEBs) is the name for the proposed eigenkernels. The methods for generating CEBs—(1) normalization processing, (2) directional pattern classification in the DCT domain, and (3) directionally separated eigenanalysis—are introduced in Sec. 2 and 3. It is proved by the simulation result presented in Sec. 4.1 that the local features extracted by the proposed method provide better discriminativeness than the same number of features extracted by the DCT, the KLT, or even Gabor wavelets.

Requiring additional computations for directional classification and several projections of a pattern onto the CEBs is one of the disadvantages of our method. For reduction of the computational load, we have estimated the proposed directional eigenkernels in the DCT domain, which can be approximately represented as tables of quantized weights, called quantized CEBs (QCEBs). A few weighted summations based on the QCEBs are enough to replace inner-product operations between a local pattern and the CEBs. Much computational load has been saved, and the results of computation-time comparison presented in Sec. 4.3 show the computational effectiveness of feature extraction using the proposed QCEBs.

2 Normalization and Directional Classification in the DCT Domain

Our approach (Fig. 1) to localized feature extraction for the face recognition starts from the 2-D DCT, which has several advantages: (1) there are several fast DCT algorithms, and thus very fast image analysis in frequency domain is possible by using the 2-D DCT; (2) most of the energy is compressed in a few low-frequency coefficients by the 2-D DCT, and therefore, some complicated calculations in the image domain can be replaced by simple ones based on a

![Fig. 1 Proposed local feature extraction method based on directional classification and transformation based on CEBs (or QCEBs) in the DCT domain.](image-url)
few energy-compacted coefficients in the DCT domain; (3) today, since the most of images are stored and transmitted as a compressed stream by the block-by-block 2-D DCT, half-reconstructed images on the DCT-coefficient level can be used directly without additional transformation for work in the frequency domain.

To sum up, the 2-D DCT is the fundamental transformer in our method for local feature analysis. Specifically, we do two preliminary jobs in the DCT domain: (1) pattern normalization for enhancing the robustness of local features to global illumination changes caused by variation of lighting, and (2) estimation of pattern directionality for extracting discriminative information selectively.

### 2.1 Local Pattern Normalization in the DCT Domain

Illumination change over a facial region, caused by variation of lighting, seriously degrades the performance of intensity-feature-based recognition systems. We effectively remove the illumination change effect by a simple normalization method in the localized DCT domain.

To minimize the effect of illumination changes, the dc component is set to zero and the energy of the ac components in the DCT domain is set to one. Smoothly changing illumination over the entire facial region is approximately thought to be flat in a local scope, so the rejection of the dc component and normalization of the energy of ac components is enough to remove the local effect of global illumination change. When a sampled image block is given as $I$, then $B$ is the result of transforming it with the $N$-point 2-D DCT,

$$B(u,v) = \alpha(u,v) \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} I(x,y) \beta(x,y,u,v),$$

where

$$0 \leq u < N, \ \ 0 \leq v < N, \ \ \alpha(t_1,t_2) = \begin{cases} \frac{1}{N} & \text{if } t_1 = 0 \text{ or } t_2 = 0, \\ \frac{2}{N} & \text{otherwise,} \end{cases}$$

and

$$\beta(x_1,x_2,t_1,t_2) = \cos \left( \frac{(2x_1 + 1) \pi t_1}{2N} \right) \cos \left( \frac{(2x_2 + 1) \pi t_2}{2N} \right).$$

The normalized block, $B_n$, in Eq. (2), contains only the spatial transitional pattern of normalized ac coefficients, which is discriminative information robust to illumination change:

$$B_n(u,v) = \begin{cases} 0 & \text{if } (u,v) = (0,0), \\ \frac{B(u,v)}{\xi} & \text{if } (u,v) \neq (0,0), \end{cases}$$

where

$$\xi = \left[ \sum_{(u,v) \neq (0,0)} B(u,v)^2 \right]^{1/2}.\tag{2}$$

As shown in Fig. 2, the patterns of unitized ac components are not seriously changed even though the directional lighting globally changed the intensity distribution in images.

**Fig. 2** Compensation of global illumination change by dc rejection and ac normalization in localized DCT domain. For convenience in display, the dc component of every block is fixed at 1024.0, and the energy of the ac coefficients at 300.0 (image size: 168×224; block size: 8×8).

### 2.2 Determination of the Transition Direction in the DCT Domain

Commonly, in a human or a machine, a transition of intensity or color in an image is the most important feature for segmenting and recognizing an object included in the image. When the concentration scope is scaled down, the position and orientation of the transition are the whole of the discriminative information. Since the position of a transition is treated independently as important information about a local feature for recognition, the transition orientation (or direction) of a local pattern is an essential feature for discriminating local image patterns.

The directional characteristics of local patterns are easily detectable in the DCT domain. Simple but powerful directional classification methods based on the 2-D DCT have been proposed for the classified vector quantization (CVQ), in which the first one or two horizontal and vertical coefficients were commonly used to estimate the orientation of a block. The ratio of the horizontal to the vertical coefficient has an approximately linear relationship with the real direction of a single edge included in a block. The arctangent of the ratio was used to find the real direction of an edge in the DCT domain.

In contrast with the conventional approaches, our goal is not to find the direction of single edge, but to estimate the directional trend of texture in a local region. When there is only clear single edge in a local scope, the ratio of first- to second-order coefficients is enough to estimate the direction of the edge, but when there is an edge of more complex shape with some noise, the simple ratio cannot work properly. For the reliable estimation of the averaged directional trend of a texture in a local pattern, we use three special energy terms, viz.,
where $2 \leq K \leq N$. Here $E_H$ represents the transitional energy in the horizontal direction, and $E_V$ that in the vertical direction; $E_M$ is the energy of mixed transitions or nondirectional transitions. Because high-frequency coefficients are too sensitive to sensor noise, we calculate the energy terms by summing the energy up to a bound $K$. Determination of the proper $K$ should depend on a heuristic method, but for a $100 \times 100$ face image, $K=4$ is adequate for an $8 \times 8$ local scope, and $K=8$ is adequate for a $16 \times 16$ local scope.

The directional classification algorithm shown in Fig. 3 first checks the strength of directionality, $R$, which is used for deciding whether the local pattern is valuable as a discriminative one. The parameter $T_R$ in Fig. 3, which should be selected in $[0, 1]$, is the threshold to filter out nondirectional patterns, which will be treated as trivial ones in the recognition process. We have

$$R = \frac{E_H + E_V}{E_H + E_V + E_M}. \quad (4)$$

In the next step, the averaged transitional orientation is estimated—not the direction of a simple edge, but the directional trend of transitions, represented by the ratio of directional energies, i.e.,

$$\angle D = \angle \left( \frac{\text{sign}(B_k(0,1))}{\text{sign}(B_k(1,0))} \frac{E_V}{E_H} \right) \text{ for } E_H \neq 0,$$

$$\angle \left( \frac{\pi}{2} \right) \text{ for } E_H = 0,$$

where $-\pi/2 \leq \angle D \leq \pi/2$ and

$$\text{sign}(x) = \begin{cases} 1.0 & \text{if } x \geq 0, \\ -1.0 & \text{if } x < 0. \end{cases}$$

Then, the signs of the first DCT coefficients are used to distinguish the directions around $\pi/4$ from the directions around $3\pi/4$, and $-3\pi/4$ from $-\pi/4$, which cannot be distinguished by the ratio of the energy terms. Directions, of opposite sense, such as $\pi/4$ and $-3\pi/4$, have to be classified as the same direction, because we want to know not the direction of an edge but the trend of directional transitions. As a result, the range of $\angle D$ is bounded in $[-\pi/2, \pi/2]$, and we have

$$f_{\pi}(B_k) = \begin{cases} 1 & \text{if } -\frac{\pi}{2} \leq \angle D < -\frac{\pi}{2} + \frac{\theta_k}{2} \text{ or } \frac{\pi}{2} - \frac{\theta_k}{2} \leq \angle D < \frac{\pi}{2}, \\ k & \text{if } \frac{\pi}{2} - \frac{(2k-3)\theta_k}{2} \leq \angle D < \frac{\pi}{2} - \frac{(2k-1)\theta_k}{2}, \\ L & \text{if } -\frac{\pi}{2} - \frac{\theta_k}{2} \leq \angle D < -\frac{\pi}{2} - \frac{(2L-1)\theta_k}{2}. \end{cases} \quad (6)$$

Finally, the orientation index $f_{\pi}$ of a local pattern is determined as one of $L$ directions uniformly distributed in $[-\pi/2, \pi/2]$ with interval $\theta_k (=\pi/L)$. The classification rule is given in Eq. (6). The identified orientations of local patterns (Fig. 4) are used for clustering same-direction patterns, which is needed for learning directionally optimized kernels, introduced in the next section.

3 Directionally Classified Kernels

The orientations of local patterns are effective and robust features for face detection or localization, but the amount of information in them is too small to discriminate individual persons. More detailed multidimensional features representing local facial regions are required for higher recognition performance. A brand-new kernel function for local feature extraction is introduced in this section. Directionally classified eigenvectors are estimated from the directionally clustered set of local patterns with strong directionality and used for extracting discriminative local pattern descriptors for labeling local facial regions.
In general PCA, eigenvectors are orthogonal basis vectors which are best in energy compaction, but not good in discriminativeness. On the other hand, the Gabor-wavelet kernels are good in catching discriminative features, but cannot make compact features. Our proposed kernel function for the local feature extraction (or localized dimensionality reduction) is designed to achieve both compactness and discriminativeness. Through a separate construction of the kernels according to the directional characteristic of a local region, the discriminative potentials of extracted features are improved over those for the Gabor kernels, while the compactness of the features is also preserved, because the kernels are the eigenvectors generated by applying PCA.

3.1 Classified Eigenblocks

The proposed kernel function for local feature extraction is constructed by applying PCA to the directionally classified sets of sampled local patterns. First, we sample enough local patterns with strong directionality from face images in a database and make the learning set $S$. Then, the set is divided into the $L$ directional sets, $\{S_1, S_2, \ldots, S_L\}$, by the proposed directional classification algorithm shown in Fig. 3. Let $\tilde{B}$ denote a vector made by zigzag-scanning the DCTed and normalized $N \times N$ block $B_n$. Let $M$ denote the total number of gathered samples, and $M_k$ the number of samples clustered into the $k$'th directional set. For all $k$, $M_k$ should be larger than $N^2$. We have

$$S = \{ \tilde{B}_i = [0, B_n(1,0), \ldots, B_n(N-1,N-1)]^T | i = 1, \ldots, M \},$$

where $\tilde{B}_i \in \mathbb{R}^{N^2}$, and

$$S_k = \{ \tilde{B}_j | \tilde{B}_j \in S \text{ and } f_\pi(\tilde{B}_j) = k \}, \text{ where } j = 1, \ldots, M_k,$$

where $S = S_1 \cup S_2 \cup \cdots \cup S_L$ and $S_k \cap S_l = \phi \ (k \neq l)$. In the next step, for the eigenanalysis, we get the mean vector and covariance matrix of each set, i.e.,

$$\bar{m}_k = \frac{1}{M_k} \sum_{j=1}^{M_k} \tilde{B}_j^k \quad \text{and} \quad \Xi_k = \frac{1}{M_k} \sum_{j=1}^{M_k} (\tilde{B}_j^k - \bar{m}_k)(\tilde{B}_j^k - \bar{m}_k)^T,$$

where $\tilde{B}_j^k \in S_k$. Then, the eigenvectors and eigenvalues of each set are acquired by solving the equation

$$\Xi_k \bar{u}_j^k = \lambda_j^k \bar{u}_j^k,$$

where $1 \leq k \leq L$, $1 \leq j < N^2$, $\bar{u}_j^k \in \mathbb{R}^{N^2}$, and $\lambda_j^k \in \mathbb{R}$. The CEBs are a set of eigenblocks that are inverse-zigzag-scanned eigenvectors acquired from Eq. (10). The CEBs of the $k$'th orientation constitute the set of $C$ eigenvectors of $\Xi_k$ selected and sorted according to the decreasing order of the eigenvalues. Since the eigenvectors actually correspond to block patterns, we call them eigenblocks and can store them as block type, i.e.,
\{\text{CEBs of } k\text{th orientation}\} = \left\{ \begin{bmatrix} 0 & \tilde{u}_{\theta_k}^j(0) & \ldots & \tilde{u}_{\theta_k}^j(1) & \ldots & \tilde{u}_{\theta_k}^j(N^2 - 1) \end{bmatrix} \right\}_{j = 1, \ldots, C}, \tag{11}

where \( C < N^2 \), \( \lambda_1^k \geq \lambda_2^k \geq \cdots \geq \lambda_{N^2}^k \), and \( \tilde{u}_{\theta_k}^j \in \mathbb{R}^{N \times N} \). Examples of generated CEBs are shown in Fig. 5. For convenience in observation, they are shown as an inverse-DCTed and scaled image of each original eigenblock. The first through fourth eigenblocks of each direction clearly reveal their directional characteristics, and the fifth through eighth eigenblocks contain detailed nondirectional texture patterns. Therefore, after a local region is transformed with the CEBs, the discriminative directional information of the region will be compacted in the first through fourth positions of extracted feature vectors, and the remaining nondirectional texture pattern will be described in the fifth position.

Therefore it is expected that higher discriminativeness of locally extracted features will be achievable with lower dimensionality. Some experimental results proving good performance of the proposed CEBs for local feature extraction are presented in Secs. 4.1 and 4.2.

### 3.2 Quantized and Classified Eigenblocks

Since we have generated CEBs from a set of DCTed blocks, most of the energy is packed in a few positions in an eigenblock due to the energy compaction property of the DCT. As shown in Fig. 6, the components of estimated eigenblocks from DCTed blocks are definitely divided into a few dominant and several trivial values. Therefore, the eigenvectors acquired in the DCT domain can be effectively approximated by a few selected dominant values and their positions in the eigenblocks. The approximated representation of eigenvectors is the key point for the fast feature extraction introduced in the next section.

The approximated representation of an eigenblock is...
Thus, the computation load will increase, and the accuracy will also increase. Local Feature Extraction Using CEBs

3.3 Local Feature Extraction Using CEBs and QCEBs

Two approaches are considered for extracting local features using CEBs. The first is a simple projection of the transformed and normalized local input pattern onto the subspace spanned by all of the CEBs:

\[
\{u^k_j\}_{q} = \{(u^v,v^v),c(a^a, a^a)\}^c(a^a, a^a) \]

where \( u^k_j \) denotes the \( k \)-th eigenvector of the \( j \)-th CEB, and \( c(a^a, a^a) \) is the index of the input pattern and use the eigenvectors of the selected direction. A more compact representation is available with the second method, but it requires a specialized distance-measuring strategy, because

\[
\bar{z} = U^T \cdot \bar{B}_n, 
\]

where \( U = [\bar{u}_1^1, \ldots, \bar{u}_L^1, \bar{u}_1^2, \ldots, \bar{u}_L^2, \ldots, \bar{u}_1^C, \ldots, \bar{u}_L^C] \) and \( \bar{z} \in \mathbb{R}^L \). Here \( \bar{u}_j^k \) denotes the \( j \)-th eigenvector of the \( k \)-th direction, and \( U \) is a columnwise eigenvector matrix. The eigenvectors are ordered by the rule that vectors of lower eigenvalues (larger eigenvalues) take earlier positions in \( U \). If there are six eigenvectors for each of eight directions, the dimension of the feature vector is determined as 48. This kind of method is compatible with the previous methods based on DCT, localized PCA, and Gabor wavelets. Without any special modification, the local feature extraction part of previous recognition systems can be replaced by this strategy based on CEBs.

In contrast with the general usage of PCA-based feature extraction, the subtraction of the mean vector is omitted in our method. Since we have already removed the dc component of the pattern in the DCT domain and we have classified patterns of opposite directionalities as having the same direction, most of the mean vectors of collected sample patterns are almost zero. By ignoring almost zero mean vectors, much computational complexity can be saved in feature extraction.

The second method is based on the determined directionality of an input pattern. First, we check the directional index of the input pattern and use the eigenvectors of the selected direction, i.e., when \( f_n(\bar{B}_n) = k \),

\[
\bar{z}_k = U^T \cdot \bar{B}_n, 
\]

where \( U_k = [\bar{u}_1^k, \ldots, \bar{u}_L^k, \ldots, \bar{u}_1^C, \ldots, \bar{u}_L^C] \) and \( \bar{z}_k \in \mathbb{R}^L \). The local pattern, of which the directionality is already determined, is represented by a \( C \)-dimensional feature vector extracted using the eigenvectors of the selected direction. A more compact representation is available with the second method, but it requires a specialized distance-measuring strategy, because
the feature vectors extracted by eigenvector sets of different directions cannot be directly compared. The feature vectors have to be transformed into a common space and then compared in that space. As an example, if the cosine distance is used as the distance metric for the comparison between locally extracted feature vectors, matrices relating the different directions must be evaluated as

\[ \mathbf{M}_{k,l} = \mathbf{U}^T_k \mathbf{U}_l, \]

where \( \mathbf{M}_{k,l} \in \mathbb{R}^{C \times C} \). Then the cosine distance is given as

\[ d_{\text{cosine}}(\mathbf{z}_k, \mathbf{z}_l) = \begin{cases} \frac{\mathbf{z}_k^T \mathbf{M}_{k,l} \mathbf{z}_l}{\| \mathbf{z}_k \| \| \mathbf{z}_l \|} & \text{if } k \neq l, \\ \frac{\mathbf{z}_k^T \mathbf{z}_l}{\| \mathbf{z}_k \| \| \mathbf{z}_l \|} & \text{if } k = l. \end{cases} \]  

(17)

On the other hand, the Euclidean distance is

\[ d_{\text{Euclidean}}(\mathbf{z}_k, \mathbf{z}_l) = \begin{cases} \left( \| \mathbf{z}_k \|^2 + \| \mathbf{z}_l \|^2 - 2 \mathbf{z}_k^T \mathbf{M}_{k,l} \mathbf{z}_l \right)^{1/2} & \text{if } k \neq l, \\ \| \mathbf{z}_k - \mathbf{z}_l \| & \text{if } k = l. \end{cases} \]  

(18)

Finally, one of the two previously introduced feature extraction methods can be chosen and used in various systems according to their working environment and specification.

Lastly, feature extraction using the CEBs can be done with reduced computational load by using QCEBs. The two feature extraction methods introduced commonly involve several inner products between eigenvectors and normalized pattern vectors:

\[ \mathbf{z} = \mathbf{U}^T \mathbf{B}_n = \left[ (\mathbf{u}_1^T \cdot \mathbf{B}_n, \mathbf{u}_2^T \cdot \mathbf{B}_n, \ldots, \mathbf{u}_C^T \cdot \mathbf{B}_n) \right]^T, \]

(19)

and each inner product can be replaced by a few weighted sums based on the tables defined in Eqs. (12) and (13), i.e.,

\[ \tilde{\mathbf{u}}_j^T \cdot \mathbf{B}_n = \frac{1}{T_Q} \sum_{(u^*, v^*) \in \mathbf{Q}_n} c_{(u^*, v^*)} \mathbf{B}_n(u^*, v^*). \]

(20)

The approximated inner product requires fewer computations than the original one.

As an experimental example, using an \( 8 \times 8 \) eigenblock, only six values survived the quantization and were stored in the table for the QCEB when \( T_Q = 16 \), and only about 9.4% of the original number of computations were needed for the same inner-product calculation between the eigenblock (or -vector) and a local pattern.

Generally, computational complexity has been a minor issue in pattern recognition, because commercial applications were not envisaged. But recently, pattern recognition technology has been applied to an increasing number of commercial systems, and the reduction of computational complexity is becoming important for effective implementation based on a standalone digital signal-processing (DSP) platform. When taking a local-feature-based object recognition approach, the 2-D DCT is the most attractive choice with regard to computation complexity, but its performance in recognition is inferior. On the other hand, the Gabor wavelet transformation is good in performance, but its high complexity in implementation prevents applying it to more cases. QCEB-based local feature extraction is a reasonable counterproposal not only to the 2-D DCT but also to Gabor wavelets.

4 Experiments and Results

Three experiments for evaluating the performance of our method have been done, and the results are presented in this section. The first experiment is on the class separability observation, which is for comparing the discriminativeness of various types of local features and confirming whether the proposed kernels for local feature extraction is valuable for enhancing recognition performance. The second experiment is on face recognition simulation. To evaluate the practical performance of the proposed method and compare it with that of conventional methods, we have done a face recognition simulation using the recognition system based on nearest neighbor classification in the local feature do-
main and observed the recognition success rate at each type of local feature on increasing the dimensionality. Finally, the third experiment is a computational complexity comparison. The processing time required for each local feature extraction method was observed. Specifically, the variation of computational load for the QCEBs was observed with variation of the quantization level.

### 4.1 Class Separability Comparison

In pattern recognition, the fundamental method for evaluating the discriminative performance of a transformation is observing the ratio of the traces of the within-class and between-class scatter matrices in a transformed and dimensionality-reduced domain. There are many criteria for measuring class separability, but they have similar physical meaning, and therefore we choose a representative criterion, viz.,

$$ R = \frac{\text{tr}(S_w)}{\text{tr}(S_b)} $$

where $S_w$ is a within-class scatter matrix and $S_b$ is a between-class scatter matrix. The two scatter matrices were made up in transformed and reduced space by various local feature analysis methods. For a face image, the image was divided into several local regions, and each region was represented by low-dimensional feature vectors extracted using the methods listed in Table 1. Then, the entire face image was represented by a combined vector of localized feature vectors, and scatter matrices were calculated from these combined vectors acquired from all face images for testing. For the simulation, the Yale face database was used, which consists of 15 classes, each consisting of 11 different conditional face images. Only facial regions of size $100 \times 100$ were manually segmented from the full face image, and 100 sampling positions were selected with the fixed interval of 10 pixels along the $x$ and $y$ axes. The local features were extracted from the $16 \times 16$ region centered at those sampling positions.

Through the experiment in this section, we wanted to examine the discriminativeness of kernels of each transformation for a local region. Therefore, the fifth method in Table 1 was ruled out from the simulation for fair comparison. The fifth method uses directionally selected kernels, and although other methods also can be supported by some kernel selection rules, giving a chance at kernel selection only to our method is unfair, so all kernels of the CEBs were used according to the rule (14).

The graph of Fig. 7(a) shows the observation results on the class separability measure $R$. The local features extracted by the kernels of the proposed CEBs showed the best class separability performance through all dimensions, while the performance was slightly better than that of Gabor wavelets, but considerably better that of the 2-D DCT and eigenpaxels. As a conclusion, the directional specialization and compensation of illumination change are more important conditions on transform kernels for local feature extraction than is high energy compaction.

The graph of Fig. 7(b) shows the variation of the ratio at
various values of \( T_Q \) for QCEBs. A smaller \( T_Q \) means a better approximation of the CEBs. Although the computation load is dramatically reduced by decreasing \( T_Q \), the performance also becomes worse. But the worsening rate of the performance is slower than the lowering rate of the computational load, as is seen from the results in Sec. 4.3. Specifically, good performance is maintained at dimensionalities of local features around 6 to 12, which is a high enough dimensionality to achieve reliable recognition performance. Therefore QCEB-based feature extraction is valuable as an approximation to the CEB-based method.

4.2 Face Recognition Simulation

The class separability measure is not adequate to evaluate the reliability of features used in a real face recognition system. We have tested various local feature extraction methods through face matching and recognition simulation, and tried to examine the suitability of the proposed CEBs and QCEBs for a face recognition system.

The face recognition system based on a nearest neighbor classifier working in the local feature domain was used for the simulations. To observe the pure performance variation caused by variation of the type of local feature, the simplest approach was used for the simulation. A combination of localized distances based on various local features was used to determine the nearest prototype to the input. The recognition success rates were observed while we varied the distance measure in local feature domain. Four different distance measures were used; two for rigid versus flexible local feature matching times two for the Euclidean versus the cosine distance metric.

A distance between input and prototype faces is defined as a sum of localized distances between local features. The rigid approach is a computationally inexpensive and linear process that can easily adapt to various classification tools. The flexible approach in Fig. 8 is the case for the classification based on elastic graph matching (EGM). The flexible approach is computationally ex-

---

**Fig. 9** Recognition success rates with various types of local features and various local feature matching strategies: (a) rigid, Euclidean; (b) rigid, cosine; (c) flexible, Euclidean; (d) flexible, cosine.
pensive but has advantages in locally deformable face recognition. Because the two approaches each have various uses, we have tested them both.

For the localized distance measure, there is also a large difference between using the Euclidean metric and the cosine metric. As an example, if two faces are substantially different in a local region, that difference will be exactly reflected to the face-to-face distance when using the Euclidean metric, but when using the cosine one, the localized difference will be reflected as a lesser one according to the share assigned to that local region. Indeed, using the Euclidean distance metric can maximize interclass distances by exactly measuring localized differences, but at the expense of maximizing intraclass distances, which should be minimized for good recognition performance. On the other hand, using the cosine distance metric can minimize intraclass distances by normalizing localized distortions, while also minimizing interclass distances. Since the two methods commonly have both advantages and disadvantages, they should be used selectively according to the condition of the recognition target and the environment in which the system works. Therefore, we tested two methods in our face recognition simulation.

The manually segmented Yale face database was also used for face recognition simulations. From a total of 11 different-condition images of an individual’s face, one normal image was used as prototype, and the remaining ten, including changes in lighting condition, facial expression, and wearing glasses, were used for testing. The simulation results, presented in Fig. 9 and Table 2, indicate that:

1. The flexible approaches were better than the rigid ones because they fully utilized the advantages of the local-feature-based approach by compensating local deformations caused by facial expression changes.
2. The 2-D DCT and eigenpaxel methods performed well with the Euclidean distance metric, while the Gabor wavelets and proposed CEBs performed well with the cosine distance metric. Moreover, the Gabor wavelets and the CEBs were better in performance than the 2-D DCT and eigenpaxels, which proves that the directionally specified kernels are effective in robust local feature extraction for face recognition.
3. For most cases, the recognition rates increased with local dimensionality up to dimensionalities of 10 to 14, and then saturated. This shows that a reasonable local feature dimensionality for our simulation setting is about 12, but the selection should be varied according to image size, local scope size, and transformation method.
4. The proposed CEBs and directionally selected CEBs commonly showed good and consistent performance over all simulations, which proves that our proposed kernels for local feature extraction may be compatible with various classification methodologies. Specifically, the combination of flexible matching, cosine distance metric, and directionally selected CEBs achieved the highest rate, about 94.7%.
5. With regard to the feature extraction method based on our CEBs, the method based on Eq. (14) achieved the highest recognition rate with a rigid matching strategy, and the directionally biased method based on Eq. (15) achieved the highest rate with a flexible strategy. In contrast with the rigid approach, the greater locality of local features guarantees higher performance in flexible matching, because the sharp local response possible by using a directionally selected kernel set is very effective for the flexible strategy, which searches locally for the best matching point.

As a conclusion, the face recognition simulation results afford proof of effectiveness of the proposed kernels for local feature extraction in face recognition applications. In particular, the superiority of the proposed method is maximized with the use of the cosine distance metric and a flexible matching strategy.

### 4.3 Computational Complexity

Table 3 contains the processing times taken to extract local feature vectors of the same dimension using the 2-D DCT, eigenpaxels, Gabor wavelets, proposed CEBs, and QCEBs.

As shown in the table, the 2-D DCT is a very inexpensive method with regard to computational complexity, due
to the existence of fast algorithms, but its performance as a feature extractor is limited in pattern recognition applications. The Gabor wavelets, eigenpaxels, and CEBs with the general usage of Eq. (14) would impose the same huge computational load if they tried to extract a feature vector of the same dimension. The QCEB-based feature extraction imposed a comparatively small computational load, as small as 25% of that of the original CEBs. Therefore, if the minor performance worsening can be accepted, major reduction of the computational load is available by using a QCEB-based implementation of the CEBs.

5 Conclusion

In this paper, we have proposed a new method for localized feature analysis adequate for face recognition applications. The proposed method utilizes the well-known DCT and PCA with the support of the directional classification of localized patterns. Energy normalization for compensating effects of illumination changes, and directional classification for concentrating on discriminative information, are performed in the frequency domain realized by a 2-D DCT. The concept of directional factorization of kernels for transformation is motivated by the shape of Gabor wavelets, and the proposed kernels are generated by PCA as eigenkernels, which are effective for generating energy-compacted features. Specialization of the kernels to each orientation by directional classification enhances the discriminativeness of resulted local features and statistical optimization using PCA improves the compactness and reliability of features. Through various simulations, we verified the superior performance of our method, which is expected to be applied in realistic face recognition applications based on localized features.

Additionally, we proposed a method that can reduce the large computational load for local feature extraction with little sacrifice in performance. Since the proposed kernels were eigenkernels estimated in the DCT domain, the successful approximation of the kernels via quantization is possible. Representation of kernels by integer values and the trade-off between computational complexity and performance are very useful properties when the face recognition system is implemented on a standalone platform using DSP, which is very sensitive to complexity of an algorithm and can handle only fixed-point processing.

Developing an automated method for determining the proper number of orientations of kernels, the sampling interval between local features, and the size of the local scopes is strongly required for enhancing the proposed method. Future research on using the proposed CEBs with complicated classification techniques (HMM, neural networks, kernel PCA, and kernel LDA), is also desired.

References

Ho-Chul Shin received his BS in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2000, and his MS in 2002. Now he is a PhD candidate in the Department of Electrical Engineering and Computer Science at KAIST. His research interests are face recognition, pattern recognition, and image processing.

Hae-Chul Choi received his BS in electrical engineering from Kyungpook National University, Daegu, Korea, in 1997. He received his MS in electrical engineering from KAIST, Daejeon, Korea, in 1999, and his PhD in 2004. He is now a researcher at Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. His research interests are object segmentation, pattern recognition, and computer vision.

Seong-Dae Kim received his BS in electrical engineering from Seoul National University, Seoul, Korea, in 1977, and received his MS in electrical engineering from KAIST, Daejeon, Korea, in 1979. He obtained his PhD in electrical engineering from the ENSEEIHT, INPT, Toulouse, France, in 1983. Since 1984, he has been a professor in the Department of Electrical Engineering at KAIST and in the Department of Electrical Engineering and Computer Science at KAIST. His research interests are image processing, computer vision, pattern recognition, and image coding.