LOCAL FEATURE EXTRACTION AND MATCHING METHOD FOR REAL-TIME FACE RECOGNITION SYSTEM

Ho-Chul Shin, Hae Chul Choi and Seong-Dae Kim

Visual Communications Lab., Department of EECS
Korea Advanced Institute of Science and Technology, Daejeon, Korea
e-mail : horse@sdvision.kaist.ac.kr

ABSTRACT

This paper introduces a new local feature extraction and matching method for the localization and classification of frontal faces. Proposed method is based on the block-by-block projection onto directionally classified and quantized eigen-vertors in frequency domain, called as QCEBs (Quantized and Classified Eigen-Blocks). QCEB-based feature extraction is effective and reliable comparing to simple zigzag scanning of DCT-coefficients or PCA. Also proposed new block-based selective matching method is faster than exhaustive sliding window-based matching methods and robust to partial missing or occlusion. Performances of proposed methods are verified by the frontal face localization and classification test. Especially, the test result shows that proposed method is utilizable for the implementation of a real-time face recognition system.

1. INTRODUCTION

Feature extractor in recognition system must be designed to select discriminative characters of target patterns and extract them robustly to the various noise and distortion. When we have a priori knowledge about features - what and where the effective features are-, designing a feature extractor is easy and simply. But in many applications, knowledge about features is not available and also finding them is very hard work, especially in face recognition.

Many researchers have proposed the feature extraction method that can automatically find effective features and robustly extract them. Especially they have been concentrated on the methods which extract the features from the statistical information of training images by well known PCA or FLD.

PCA(Principal Component Analysis), FLD (Fisher Linear Discriminants) are statistical linear-mapping method most widely used as an generalized automatic feature extraction method. PCA extracts most expressive feature and FLD extracts most discriminative feature. FLD may seems to be more profitable than PCA in pattern recognition, but if the number of training images are not enough, FLD can't warrant better performance than that of PCA.

Like the eigenfaces[4] or fisherfaces[5] method, PCA or FLD are generally utilized to extract global features. But these globalized feature extraction cannot properly cope with the partial changing, partial distortions, missing part, occlusion and misalignment. Moreover PCA and FLD need seriously large computation loads so that they are not proper to fast real-time system.

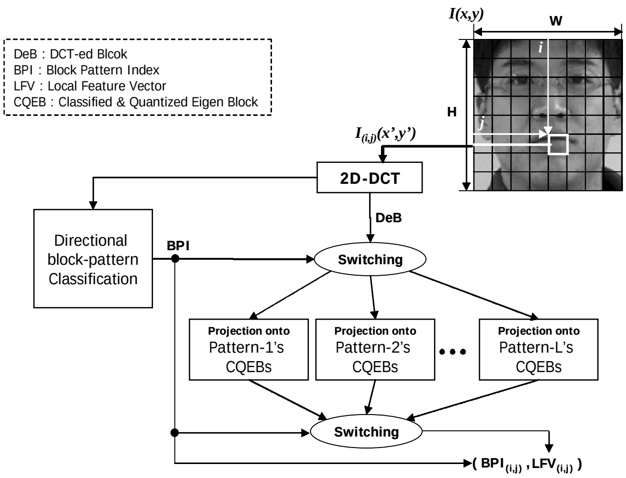


Fig. 1. Proposed block-based feature extraction method

2D-DCT(Discrete Cosine Transform)[2] and VQ(Vector Quantization)[3] - based block-by-block feature extraction methods have been proposed to solve the problem of partial distortions. In these methods, through the localized normalization, they can reject or reduce the partial distortions. Additionally they have many advantages in implementing the fast recognition system, because 2D-DCT or VQ needs comparatively lower computational load than PCA or FLD and have compatibilities with various image coding methods and hardware structures.

Main problem in designing a DCT or VQ-based feature extractor is how to arrange and select the coefficients for the higher performance. Most previous methods select low frequency coefficients with zig-zag scanning or use code-vectors associated with a few low frequency components. But for some localized region of images which contain more complex edges or textures, higher frequency coefficients are more valuable for the discrimination than lower frequency components. In these senses, coefficients selecting and arranging methods should be varied by the localized statistical character of target pattern.

In this paper we introduce a new 2D-DCT-based feature extraction method and block matching method. Showed in Fig. 1, proposed feature extraction method firstly classifies blocks by their directional characteristics and extracts feature vectors using proposed QCEB (Quantized and Classified Eigen-Block) from blocks. QCEB is a kind of basis vectors for projection that are consists of 3~10 weighting-values for DCT-coefficients. The weighting-values came from quantizing the DCT-coefficients-eigen-vectors of the set of gathered same directional blocks. Though it requires small number of calculation, with QCEB, a few weighted summing of coefficients is more effective than the simple zig-zag coefficient selection or block-level PCA[1] because of selective using of optimal basis as their directional characters.

Directional pattern classification of block is simple work using the ratio of first vertical and horizontal DCT-coefficients. But through the pre-classification of blocks, effective representation of block and fast block-based matching is available. This fast matching method is described in chapter 5.

Performances of proposed methods are verified by face-localization and recognition test with images polluted by various partial distortions like an missing, occlusion and illumination change, etc. Also, comparison with other global or local methods shows that our method is reliable and fast so that is applicable to real-time face recognition system.

2. 2D-DCT

2.1 Block-by-block N-point 2D-DCT

We divide a given raw image, ***I(x,y)***, into square blocks with size of $N \times N$ and apply N-point 2D-DCT to each block. $\boldsymbol{B_{(i,j)}(u',v')}$ and ***C(u,v)***, can be obtained by (1) and (4). ***(i,j)*** is the position index of a block in a given image. Operations - ***quo(a,b)*** and ***rem(a,b)*** means the quota and remainder resulted from dividing ***a*** by ***b***. ***W*** and ***H*** are the width and height of image.

for $0 \le u' \le N-1,\ 0 \le v' \le N-1$

and $0 \le i \le W/N-1,\ 0 \le j \le H/N-1$

$$B_{(i,j)}(u',v') = \alpha(u',v') \sum_{y=jN}^{jN} \sum_{x=iN}^{iN} I(x,y)\, \beta(x-iN, y-jN, u', v') \qquad (1)$$

$$\text{where } \boldsymbol{\alpha}(t_1,t_2) = \begin{cases} 1/N & \text{if } t_1=0 \text{ or } t_2=0 \\ 2/N & \text{otherwise} \end{cases} \qquad (2)$$

$$\text{and } \boldsymbol{\beta}(x_1,x_2,t_1,t_2) = \cos\left(\frac{(2x_1+1)\pi t_1}{2N}\right)\cos\left(\frac{(2x_2+1)\pi t_2}{2N}\right) \qquad (3)$$

for $0 \le u \le W-1$ and $0 \le v \le H-1$

$$C(u,v) = B_{(quo(u,N),quo(v,N))}(rem(u,N), rem(v,N)) \qquad (4)$$

2.2 DC-rejection and energy normalization

DC component rejection and AC component energy normalization can reduce the effects of global and local illumination changes caused by lighting conditions. Modified (1)' and (4)' equations show these processing.

$$B'_{(i,j)}(u',v') = \frac{\alpha(u',v') \sum_{y=jN}^{jN} \sum_{x=iN}^{iN} I(x,y)\, \beta(x-iN, y-jN, u', v')}{\sqrt{\sum_{u'=0}^{N-1} \sum_{v'=0}^{N-1} B_{(i,j)}(u',v')^2}} \qquad (1)'$$

$$C(u,v) = \begin{cases} 0 & \text{if } quo(u,N)=0 \text{ and } quo(v,N)=0 \\ B'_{(quo(u,N),quo(v,N))}(rem(u,N), rem(v,N)) & \text{if } quo(u,N)\ne 0 \text{ or } quo(v,N)\ne 0 \end{cases} \qquad (4)'$$

3. DIRECTIONAL CLASSIFICATION OF BLOCK

In a block region, directions of edges or textures are one of the most discriminative and noise-robust features. Especially, the ratio of first vertical and horizontal DCT-coefficients, ***B(1,0)/B(0,1)*** , directly matches with the direction of edgeness in a block. The block can be classified as their directional characters by estimated directions using *arctan* ***(B(1,0)/B(0,1))***.

Proposed directional block classification method in DCT-domain is presented by next equations and pseudo-code.

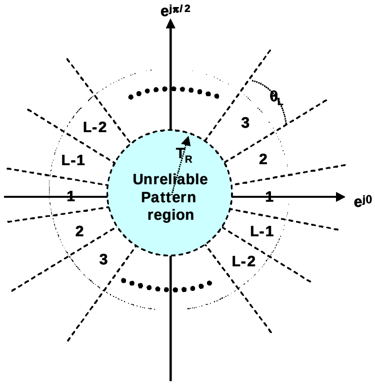


Fig. 2. Directional classification of block by angle of ***D***

$$\boldsymbol{D} = \boldsymbol{B}(1,0) + \boldsymbol{j}\cdot\boldsymbol{B}(0,1) \tag{4}$$

$$|\boldsymbol{D}| = \sqrt{\boldsymbol{B}(1,0)^2 + \boldsymbol{B}(0,1)^2} \tag{5}$$

$$\angle\boldsymbol{D} = \arctan\left((\boldsymbol{B}(0,1/\boldsymbol{B}(1,0))\right) \tag{6}$$

$$\mathrm{L} = \pi/\boldsymbol{\theta}_L + 2 \tag{7}$$

< Directional block classification algorithm>

DCB(* DCT-ed block *)

{

if ($|\boldsymbol{D}| \le T_R$) , $\boldsymbol{BPI} \Leftarrow 0$ (unreliable pattern)

else if ($0.5*|\boldsymbol{D}| < |\boldsymbol{B(1,1)}|$) , $\boldsymbol{BPI} \Leftarrow$ L (mixed pattern)

else {

if($\angle\boldsymbol{D}<0$) $\angle\boldsymbol{D} \mathrel{+}= \boldsymbol{\pi}$

if($0 \le \angle\boldsymbol{D} < \boldsymbol{\theta}_L/2$ or $\boldsymbol{\pi}-\boldsymbol{\theta}_L/2 \le \angle\boldsymbol{D} < \boldsymbol{\pi}$), $\boldsymbol{BPI} \Leftarrow 1$

else if ($\boldsymbol{\theta}_L/2 \le \angle\boldsymbol{D} < 3\boldsymbol{\theta}_L/2$) , $\boldsymbol{BPI} \Leftarrow 2$

…

else if ($(2\mathrm{L}-3)\boldsymbol{\theta}_L/2 \le \angle\boldsymbol{D} < (2\mathrm{L}-1)\boldsymbol{\theta}_L/2$) , $\boldsymbol{BPI} \Leftarrow$ L-1

}

}

return ***BPI***

Proposed directional pattern of block consists of $\pi/\boldsymbol{\theta}_L$ directional patterns, one mixed pattern and one unreliable pattern. In the process of classification, when the |B(1,1)| is comparatively large, the block will be classified as a mixed one which means that it has more complex pattern than a single edge. If |D| is too small, the block will be classified as unreliable pattern including block. They will be excluded or rejected during the feature extraction or matching process.

Blocks of mirrored patterns which have same directional characters but have inverted intensity distribution will be classified as same pattern. It's because that the localized intensity inversion is frequently occurred in images by lighting condition change.

4. QCEB

Proposed QCEB(Quantized and Classifed Eigen-Block) is new type of projection basis for the dimensionality reduction of image block. Original concept of block-by-block PCA for feature extraction has been introduced as Eigenpaxel method [1], but QCEB method is different in that it's applying PCA on the DCT domain and selective construction of projection basis according to the directional block pattern. Separated basis construction by DCT and PCA can compact energy of block at a few selected positions in DCT domain. By quantizing these separated basis vectors, we can reduce the computational load without loss of efficiency of energy compaction rates. This process is showed in Fig. 4 in detail.

4.1. PCA in DCT domain

Operation "$\rightarrow$" means raster scanning of block to make a matrix to a single column vector and "$\leftarrow$" is inverse one. ***DCB(…)***, function for directional classification of block introduced in previous chapter, returns the index of block pattern. ***S*** is set of DCT-ed blocks gathered through all training images and ***S₁* , *S₂* ,…, *S_L*** are the classified sets of training blocks by directional patterns 1 to ***L***.

$$\boldsymbol{S} = \{\vec{\boldsymbol{B}}_i \in \Re^{N^2\times 1} | \boldsymbol{i}=1,\ldots,\boldsymbol{M}\} \tag{11}$$

$$\boldsymbol{S} = \boldsymbol{S}_1 \cup \boldsymbol{S}_2 \cup \ldots \cup \boldsymbol{S}_L \text{ and} \tag{12}$$

$$\boldsymbol{S}_k = \{\vec{\boldsymbol{B}}_j^k \,|\, \boldsymbol{DCB}(\boldsymbol{B}_j^k) = \boldsymbol{k},\ \vec{\boldsymbol{B}}_j^k \in \boldsymbol{S},\ \boldsymbol{j}=1,\ldots,\boldsymbol{M}_k\} \tag{13}$$

Next, apply PCA to each pattern's training sets,

given $\boldsymbol{S_1}, \boldsymbol{S_2}, \ldots, \boldsymbol{S_L}$, for $\boldsymbol{k} = 1, \ldots, \boldsymbol{L}$,

$$\boldsymbol{m}_k = \frac{1}{\boldsymbol{M}_k}\sum_{j=1}^{M_k}\vec{\boldsymbol{B}}_j^k \quad , \quad \Xi_k = \frac{1}{\boldsymbol{M}_k}\sum_{j=1}^{M_k}(\vec{\boldsymbol{B}}_j^k - \boldsymbol{m}_k)(\vec{\boldsymbol{B}}_j^k - \boldsymbol{m}_k)^T \tag{14}$$

get $(\boldsymbol{u}_j^k, \boldsymbol{\lambda}_j^k)$ satisfies ,

$$\Xi_k \boldsymbol{u}_j^k = \boldsymbol{\lambda}_j^k \boldsymbol{u}_j^k \tag{15}$$

where $\boldsymbol{j} = 1,\ldots,\boldsymbol{N}^2$, $\boldsymbol{\lambda}_1^k \ge \boldsymbol{\lambda}_2^k \ge \ldots \ge \boldsymbol{\lambda}_{N^2}^k$

$, \boldsymbol{u}_j^k \in \Re^{N^2}$ and $\boldsymbol{\lambda}_j^k \in \Re$, $\boldsymbol{N}_k << \boldsymbol{N}^2$

$$\textbf{\textit{pattern-k's eigen-vectors}} = \{\vec{\boldsymbol{u}}_j^k | \boldsymbol{j}=1,\ldots,\boldsymbol{N}_k\} \tag{16}$$

Last, quantizes each pattern's eigen-vectors.

given $\boldsymbol{T_Q}$ = Quantization level, $\boldsymbol{T_s}$ = Decision level,

$$\boldsymbol{Q}(\vec{\boldsymbol{u}}_j^k = [\boldsymbol{u}_1\ \boldsymbol{u}_2 \ldots \boldsymbol{u}_{N_k}]^T, \boldsymbol{T_Q}) = \{(\boldsymbol{l},\boldsymbol{c}) | \boldsymbol{c} = \boldsymbol{quo}(\boldsymbol{u}_l, \boldsymbol{T}_Q) \ge \boldsymbol{T}_S\} \tag{17}$$

$$\textbf{pattern-k's } \boldsymbol{QCEBs} = \{\boldsymbol{Q}(\vec{\boldsymbol{u}}_j^k, \boldsymbol{T}_Q) | \boldsymbol{j}=1,\ldots,\boldsymbol{T}_k\} \tag{18}$$

Finally resulted QCEBs consist of quantized weight values matched at a few positions in DCT-domain and some examples of QCEBs are showed as the form of inverse-DCT-ed images in Fig. 3. In 1~5-th eigen-blocks of each directional block patterns, we can notify that their directional characters are dominantly revealed.

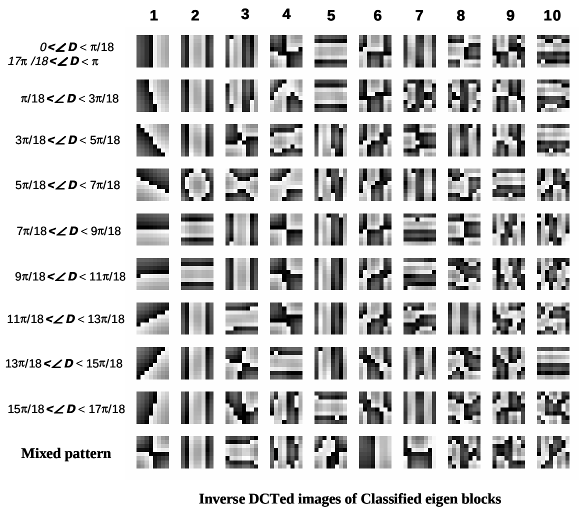


Fig. 3. Examples of inverse-DCT-ed QCEBs (8x8-block size, ***L***=11)

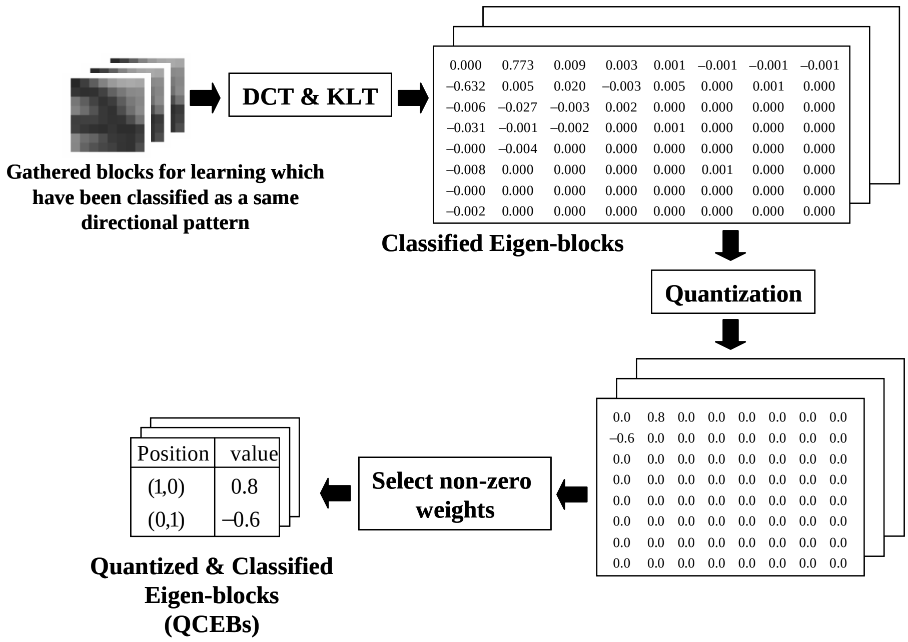


Fig. 4. QCEB (Quantized and Classified Eigen-Block) construction process

4.2 Local feature extraction using QCEB

If the pattern index of a given block, $\boldsymbol{B_{(i,j)}(u',v')}$ which comes from (i,j)-position in image, is $\boldsymbol{k}$, then local feature vector(***LFV***) is given as $\boldsymbol{z_{(i,j)}}$ by projection given block onto pattern-k's QCEBs.

$$\boldsymbol{U}_k^T = [\boldsymbol{u}_1^k\ \boldsymbol{u}_2^k \dots \boldsymbol{u}_{N_k}^k]^T, \quad \boldsymbol{z}_{(i,j)} \in \Re^{N_k}$$

$$\boldsymbol{LFV}_{(i,j)} = \boldsymbol{z}_{(i,j)} = \boldsymbol{U}_k^T(\vec{\boldsymbol{B}}_{(i,j)} - \overline{\boldsymbol{m}}_k) \qquad (19)$$

Conceptually, feature extraction by projection is (19), but real implementation of feature extraction using QCEB is just weighted summing of a few DCT-coefficients. Though QCEB needs smaller or almost same-level computational load comparing to PCA, FLD or 2D-DCT, but achieves higher representation and discrimination performance.

5. BLOCK-BASED MATCHING STRATEGY

Proposed image matching method can be used for finding translational position of templates in input image or classifying templates which are included in input images. Matched blocks between input image and templates are constrained as block-pairs which have same directional pattern. Matching measure (***MD***) is calculated for the all matched blocks and accumulated at the point of positional difference(***PD***) in accumulation map, ***A(m,n)***.

To avoid errors caused by misaligned block-matching, templates are analyzed by overlapped blocks. Since feature extraction from template is off-line process, doubled feature extraction process will not damage the performance of system seriously.

Resulted position of maximum value in ***A(m,n)*** means the translational position of template in input image. In the case of that many templates are included in input images, we can set threshold value and select multiple maximum position. Maximum value itself means localized similarity between input and template. Moreover, by comparing max-values of each template, the best matching template can be decided. It can be utilized as a classification process.

<Selective block-based matching algorithm >

Reset the accumulation map , ***A(m,n) =0***

where ***m***= …,-1,-1/2,0,1/2,1,… , ***n***=…,-1,-1/2,0,1/2,1,…

For ***i=0,…,W/N-1*** and for ***j=0,…,H/N-1***

Get $B_{(i,j)}$ from input image.

For ***i′=0,…,W′/N-1*** and for ***j′=0,…,H′/N-1***

Get $B'_{(i',j')}$ from template.

If ***DCB(*** $B_{(i,j)}$ ***)*** = ***DCB(*** $B'_{(i',j')}$ ***)***

$$MD = K_o e^{-\left\| LFV_{B(i,j)} - LFV_{B'(i',j')} \right\|^2 / \sigma^2}$$

$$PD = (i,j) - (i',j') = (i-i', j-j')$$

$$A(PD) \mathrel{+}= MD$$

End for ***i, i′***

Find maximum point ***(m*,n*)*** in ***A(m,n)***

Matchness ⇐ ***A(m*,n*)*** (20)

Translation ⇐ ***(m*,n*)*** (21)

6. COMPUTATIONAL LOADS

For the one block region of N×N size, feature extraction using proposed QCEB with fast-DCT algorithm [8] needs about $N^2(\log_2 N^2+3\sim10)$ calculation units. Though the fast-DCT and zigzag scanning method needs only $N^2 \cdot \log_2 N^2$ calculations, but because of its poor performance, comparatively large dimensions of feature vector will be needed. Large feature dimensions make worse the speed of block matching process. In the other hands, block-by-block PCA needs N^4 calculations and also produce the large dimensional feature vector for the same performance with QCEB.

For the block-based matching of two images which have K block regions, full search matching method needs K^2 mean square calculation between local feature vectors are needed. But proposed directional pattern-based matching method needs approximately K^2/L calculations. L is the number of directional classes. If the dimension of feature vector is smaller, calculations for block matching can be more reduced. The QCEB-based feature extraction can extract smaller feature vectors comparing to simple zigzag scanning or block-level PCA because of its good compaction abilities.

7. EXPERIMENTAL RESULT

7.1 Face localization test

All images used in test came from Yale Face Database [5]. Face localization means to find the exact locations of the face template in test images. We assumed to have only one template image and various input images got under different conditions were used for testing. In Fig. 7, test result shows that our method can find exact location of faces robust to changes of facial expression, background texture and lighting condition with only one template.

7.2 Face classification test

In our face classification test, only one template per class was used for classification and we assume that there is only translation between template and input image without rotation or pose variation of faces.

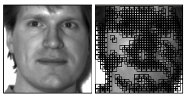
Followed 5 different methods were tested and compared. Manually gated one normal face image per one person was used as a template for face classification, showed in Fig. 5. Goal of classification is to find exact personal identification for the original 150 and generated 150 test images of which some parts are missed (Fig. 6). Identification of person is decided by comparing the matchness (eq.(20)) between each templates and input image.

Fig. 5. Examples of manually gated face templates

Fig. 6. Examples of generated test images

Miss-classification rate is showed in Fig. 8. Since the general PCA (called as eigenface[4]) method cannot adapt to missing parts or occlusions, almost all of the generated test images are miss-classified with method 1. With the block-based localized method; Method 2 & 3, including background and illumination changing damaged the classification results. Method 4 with localized distortion normalization processing showed comparatively higher performance but it cannot cope with the included noisy blocks came from background region.

Our proposed method showed best performance. It was well adapting to missing part by localized block matching and reject or minimize the effect of including background region by directional pattern-based selective matching. Moreover through the localized normalization processing in DCT domain, our method sowed good performance under the effect of lighting condition changes.



←Template and selected reliable blocks (black-boxed)

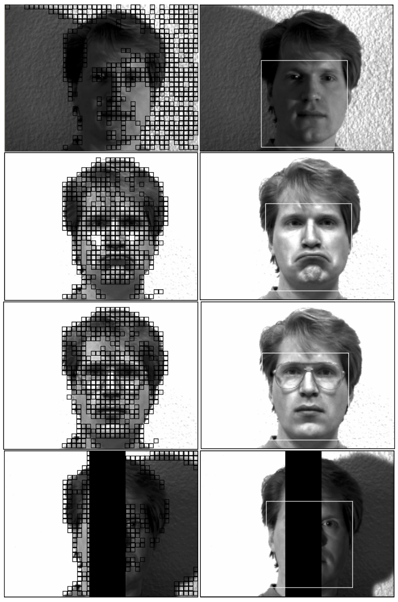


Fig. 7. Face localization test result(Left : selected reliable blocks (black-boxed) in tested input images , Right : Resulted locations of template(white-boxed)

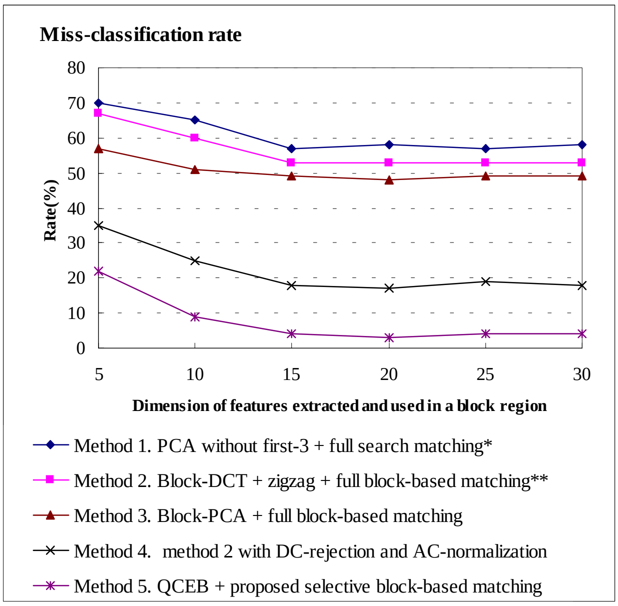


Fig. 8. Face classification test result - miss-classification rate of various feature extraction and matching method

	Feature extraction from one input image (320x240 , Gray-pixel)	Matching-based searcing for the one template and one input image
Method 1	60ms	1560ms (8-pixel step searching)
Method 2	12ms	1440ms
Method 3	64ms	1440ms
Method 4	24ms	1380ms
Method 5	36ms	80~100 ms (Selective matching)

Fig. 9. Processing time comparison

Observed processing time is listed in Fig. 9. All methods were implemented and run on the Intel Pentium-4 2.4GHz PC with Microsoft visual C++ 6.0. From the results of method 1 through 4, the exhaustive full searching matching or full block matching is not proper for the real-time system, but proposed selective matching method needs about 100ms and can be applicable to real-time system that can process about 10 frames per second.

The result of classification test and processing time observation shows that proposed QCEB-based feature extraction and selective block-based matching method can be the solution for the accurate and fast real-time face recognition system.

8. DISCUSSION

QCEB-based local feature extraction and selective block matching method is adequate for fast real-time system which can do localizing and classifying the occluded or partially polluted face images. Performance of proposed method was verified by test with various face images. In our method, we assumed that target images were already aligned for rotation or pose variation, but in real condition, aligning problem is very important issue. Advancing the method that can adapt to rotation, pose and scale changes is inspired as a further work.

9. REFERENCES

[1] P. McGuire and G. M. T. D'Eleuterio, "Eigenpaxels and a Neural-Network Approach to Image Classification," *IEEE Transactions on Neural Networks*, Vol. 12, No. 3, May 2001.

[2] C. Sanderson and K.K. Faliwal, "Polynomial features for robust face authentication," *in Proceeding of 2002 International Conference on Image Processing* , Vol. 3 , pp. 997 -1000, 2002.

---

* Full search matching means sliding templates for all of positions in input images and calculating similarities.

** Full block-based matching method matches a block in the template to the all of blocks in input image and accumulates similarities for all position in map.

[3] K. Kotani, C. Qiu and T. Ohmi, "Face recognition using vector quantization histogram method," *in Proceeding of 2002 International Conference on Image Processing*, Vol. 2 , pp. 105 -108, 2002.
[4] Zhujie, Y.L.Yu, "Face Recognition with Eigenfaces," *in Proceeding of the IEEE International Conference on Industrial Technology*, pp434-438, 1994.
[5] P. N. Belhumeur, J. P. Hespanha, and D.J.Kriegman, "Eigen-faces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No.7, pp. 713-714, July 1997.
[6] A. K. Jain, R.P.W. Duin and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, Jan. 2000.
[7] Baback Moghaddam and Alex Pentland, "Probabilistic Visual Learning for Object Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, 1997.
[8] Il Dong Yun and Sang Uk Lee, "On the Fixed-Point-Error Analysis of Several Fast DCT Algorithms," *IEEE Transactions on Circuits and Systems for Video Technology* , Vol. 3, NO. 1, Feb. 1993.
[9] D. S. Kim and S. U. Lee, "Image vector quantizer based on a classification in the DCT domain", *IEEE Transactions on Communications*, Vol. 39, No. 4, 1991.