Global Feature Extraction and Recognition from Matrices of Gabor Feature Faces

Wilfred O. Odoyo, Beom-Joon Cho*, *Member, KIMICS*

Abstract— **This paper presents a method for facial feature representation and recognition from the Covariance Matrices of the Gabor-filtered images. Gabor filters are a very powerful tool for processing images that respond to different local orientations and wave numbers around points of interest, especially on the local features on the face. This is a very unique attribute needed to extract special features around the facial components like eyebrows, eyes, mouth and nose. The Covariance matrices computed on Gabor filtered faces are adopted as the feature representation for face recognition. Geodesic distance measure is used as a matching measure and is preferred for its global consistency over other methods. Geodesic measure takes into consideration the position of the data points in addition to the geometric structure of given face images. The proposed method is invariant and robust under rotation, pose, or boundary distortion. Tests run on random images and also on publicly available JAFFE and FRAV3D face recognition databases provide impressively high percentage of recognition.**

Keywords — **covariance matrices; gabor filters (kernels) ; eigenvectors; facial component features**

I. INTRODUCTION

FEATURE extraction and representation of faces is a crucial task in the many steps involved in object recognition especially facial identification process in the wider field of biometrics. Decades of research have shown that this task is achievable and have succeeded to a level of commercial use in various applications. Among the various methods used, geometric feature-based methods of feature extraction have performed much better compared with their counterpart template-based methods. Many researchers have applied successfully low-dimensional feature extraction methods from intensity of face images which include Eigenface (the core of Principal Component Analysis) [1][2], Fisherfaces (a combination of PCA and Linear Discriminant Analysis) [3]. The property of LDA in discriminating the within-class scatter and the between-class scatter merits Fisherfaces as a better method in this category. There are other non-linear methods that have been suggested that map the input space to a high dimensional space. These include Kernel PCA [4], and Kernel LDA [5].

We propose a combination of Gabor Filters and Covariance Matrices with a Geodesic distance measure selection as the discriminatory factor between the gallery image and the probe. Gabor filters have been widely appreciated for the strong characteristics of spatial locality, scale and orientation selectivity that it exhibits [6]. More of its attributes that work to our advantage is the fact that they can expressly show and mimic the functions of simple cells in the visual cortex [7], a property critically needed in the imagery field. Gabor functions have different parameter values that can be used to efficiently detect image edges, facial contour detection, and some visual perception effects can also be explained in the resulting Gabor images. Covariance matrices we have used have significantly contributed to various state-of-art technologies, especially in tracking devices in videos, surveillance cameras at passenger terminals, and in medical imagery, to mention but a few. It mainly generates object-based representations embodying both spatial and statistical properties of faces. We also reckon here that it provides a perfect solution to fusion of multiple features contained in the target objects [8]. The metric measure adapted here is based on geodesic distance which compares the similarities between the covariance matrices of the probe image and the gallery model faces for classification.

We will explain in detail Gabor Filters in Section II as part of the image preprocessing and normalization. Section III further explains the Covariance Matrices and its execution on the Gabor filtered images for feature extraction. Geodesic distance measure occupies Part IV. Part V and VI contain experimental results and conclusion of the paper, respectively.

II. GABOR FILTERS

Mathematically, Gabor wavelets are explained as functions that hold up data into different frequency components. This separation provides us with the capability to study each individual component with a resolution of its own scale. We chose Gabor wavelets in this research for its excellent spatial locality and orientation selectivity [12,13]. Gabor wavelets extract

Manuscript received January 31, 2011; revised March 10, 2011; accepted March 24, 2011.

Beom-Joon Cho* (Corresponding author) is with the Department of Computer Engineering, Chosun University, Gwangju, 501-759, Korea (Email: bjcho@chosun.ac.kr)

Wilfred O. Odoyo is with the Department of Computer Engineering, Chosun University, Gwangju, 501-759, Korea (Email: wilfody2004@yahoo.com)

spatial frequencies and local structural characteristics within the local area of images at multiple directions. As an added advantage, these kinds of wavelet have been proven to have high tolerance on the changes in motion, deformation, scaling, rotation, and illumination. Gabor wavelets are defined as in the following formula:

$$\varphi_{\mu,v}(z) = \frac{\| k_{\mu,v} \|^2}{\sigma^2} e^{\frac{-\|k_{\mu,v}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{\mu,v}z} - e^{\frac{-\sigma^2}{2}} \right] \quad (1)$$

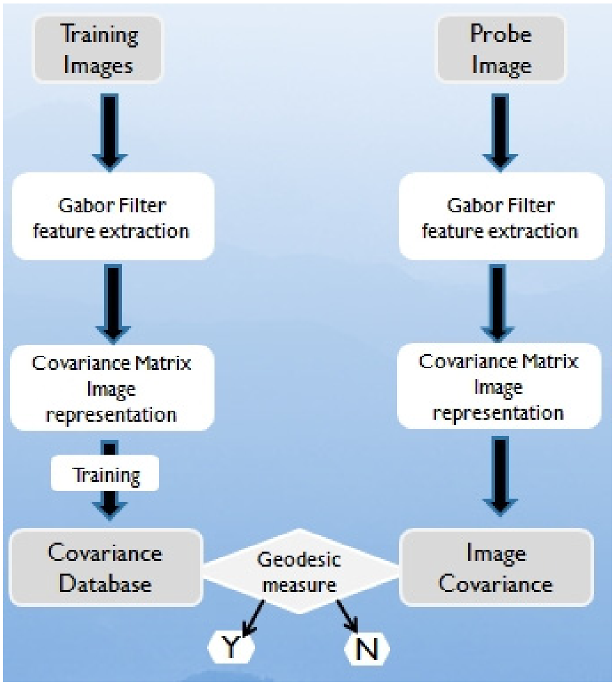

Fig. 1. The process of training and recognition.

In formula (1) above, Gabor filter uses a gabor function with different defined parameters like orientation, wavelength, phase offset, aspect ratio, and bandwidth. These parameters have values which are calculated and displayed as an intensity map image. In (1)

$$k_{\mu,v} = k_v e^{i\Phi_\mu}, k_v = k_{\max}/f^v, \Phi_\mu = \pi\mu/8, z = (x, y)$$

$\|.\|$ represents the norm operation while i is complex operator. δ defines the bandwidth of the wavelet filter, μ is the direction (orientation) and v is the scale of the Gabor filters. The application of the above filter onto a distribution of gray scale image $f(x, y)$ is done by convolution as shown below.

$$G(x, y, \mu, v) = f(x, y) * \Psi_{\mu,v}(z) \quad (2)$$

where $f(x, y)$ refers to all pixels in the image, * denotes the convolution operator, and $G(x, y, \mu, v)$ is the convolution result corresponding to the Gabor kernel at various orientations and scales given. It is worth noting that wavelength and orientation are two very important parameters that define the center location of the filter. *Wavelength*: is the scale and its values specified in pixels. It is set at more than one-fifth of the input image to get desirable effects at image borders. The pixel size increases with higher values of wavelength. *Orientation*: specifies the direction of the normal to the parallel stripes of a Gabor function and it is in degrees. The final output of a Gabor filter is a combination of the convolutions based on the number of orientations called for. That is, for a set integer value $N, N \geq 1$, then N convolutions will be computed.

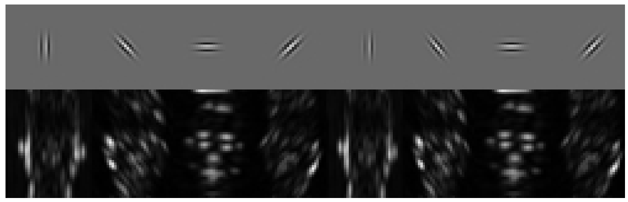

Fig. 2. Sections of real part of Gabor kernels at five scales and eight orientations on the upper row while the lower row displays the magnitude.

The set $S = \{\varphi_{\mu,v}(z) : \mu \in \{0,...,7\}, v \in \{0,...,4\}\}$ depicting different orientations and scale forms the Gabor filter representation of the image in figure 3b below. Different orientations set the filters to be able to capture most of the texture characteristics in the image thus the production of more accurate results. All the representation results have been concatenated encompassing the different spatial frequencies, spatial localities, and orientation as selected by the Gabor function.

Fig. 3a. shows sample images from FRAV3D.

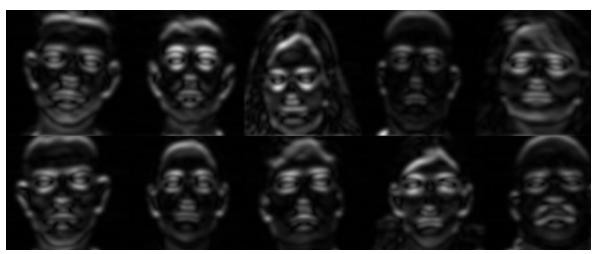

Fig. 3b. is the Gabor filtered feature extraction of the sampled FRAv3D images above in 3a.

III. COVARIANCE MATRICES REPRESENTATION

In order to develop a more accurate facial feature representation, we needed to compute the covariance matrix [8], from the result of the Gabor filtered images above. The computed covariance enhances robustness against severe illumination changes, noise, and any erratic motion detected that would contribute to the low performance during feature extraction. These matrices embody both spatial and statistical properties of objects, thereby providing an elegant solution of multiple features. Covariance Matrices is an essential measure of how much the deviation of two or more variables (image matrices) match. For a given *d*-dimensional feature points inside a particular region ***R*** are represented by $\{z_k\}, k = 1,...,n$. The entire image can be represented by a $d \times d$ covariance matrix of feature points computed as [8];

$$C_R = \frac{1}{n-1}\sum_{k=1}^{n}(z_k - \mu)(z_k - \mu)^T \tag{3}$$

where μ is the mean value of z_k and is computed as;

$$\mu = \frac{1}{n}\sum_{k=1}^{n} z_k \tag{4}$$

C_R is a matrix-form which encapsulates the facial feature extracted and stored to be used in the classification process. This matrix C_R captures important information embodied in both the histograms and appearance models of the processed images. The covariance of different features is constructed on the image from a given window. The features captured could be in many forms like gradient, edge, coordinates, color, motion and texture. Static images used in this work require us to capture gradient, edge, color, and texture features.

IV. GEODESIC DISTANCE MEASURE FOR MATCHING AND CLASSIFICATION

To detect particular face images presented by the probe, we first train our classifier. This training of the classifier is done offline by applying the covariance matrices to target images in the database as descriptors. We then evaluate our classifier online on each new probe presented to determine whether they belong to the target images or not. We chose to apply Forstner distance measure to compute the minimum distance between covariance matrix representations of both the target image and the probe. The algorithm used here requires computation of distances between feature points, a property that Euclidean distance method lacks. The formula explained in detailed in [9], is stated as below. Note that C_{R1} and C_{R2} represent covariance matrix of the first and second image, respectively.

$$\rho(C_{R1}, C_{R2}) = \sqrt{\sum_{k=1}^{d} \ln^2 \lambda_k (C_{R1}, C_{R2})} \tag{5}$$

where $\lambda_k(C_{R1}, C_{R2})$ are the generalizes eigenvalues of C_{R1} and C_{R2}. The eigenvalues are computed from

$$\lambda_k C_{R1}\mathbf{x}_k - C_{R2}\mathbf{x}_k = 0 \quad k = 1 \dots d \tag{6}$$

while $X_k \neq 0$ are the generalized eigenvectors. The distance measure ρ satisfies the metric axioms for positive definite symmetric matrices C_{R1} and C_{R2} of positivity, symmetry, and triangle inequality

1. $\rho(C_{R1}, C_{R2}) \geq 0$ and $\rho(C_{R1}, C_{R2}) = 0$ only if $C_{R1} = C_{R2}$,
2. $\rho(C_{R1}, C_{R2}) = \rho(C_{R2}, C_{R1})$,
3. $\rho(C_{R1}, C_{R2}) + \rho(C_{R1}, C_{R3}) \geq \rho(C_{R2}, C_{R3})$.

To find the best match, we compare the two covariance matrices of the input image and the database images to see which image in the database has the least distance with the input. The smallest distance from the current input is returned as the best match. Several experiments have been done and the results shown below.

V. EXPERIMENTAL RESULTS

In order to evaluate the performance of our method, JAFFE [11] publicly available database has been used. The database contains 213 images of facial expressions. We chose the neutral faces to represent individual faces in our training for the gallery. We have also conducted experiments using the 2D images of the multimodal database provided online from Face Recognition and Artificial Vision Group (FRAV3D) website [10] for extended evaluations. The images have been resized to 256 x 256 and are converted into grayscale images as the process advances. Parts of FRAV3D with 380 images from 20 subjects have been sampled for both training and testing process. The process passes through Gabor filter enhancing the features before a 7 x 7 covariance matrix is computed. Our gallery contains 2 images from each of the 10 individuals represented in [11] in case of the JAFFE database. 20 representative covariance matrices for each represented image is computed and stored in the database for comparison with the covariance computed from the input image. Below are the 10 images used for training the JAFFE database.

Fig. 4a. shows JAFFE images used for training database.

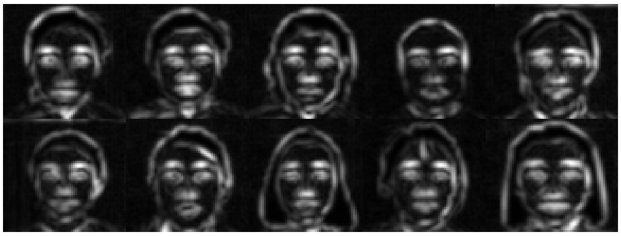

Fig. 4b. shows the Gabor filters for the database images.

These images go through Gabor filter and are transformed as in figure 4b. The concatenation of the different spatial frequencies, localities, and orientation localities form the image features and is represented by the covariance matrices in the database.

1.3575	0.7364	-0.1889	0.1919	1.3777	-0.0227	2.7225
0.7364	0.5447	-0.1386	0.0163	0.8433	-0.0170	1.6674
-0.1889	-0.1386	0.0706	0.0420	-0.2298	0.0397	-0.4460
0.1919	0.0163	0.0420	0.2181	0.1237	-0.0010	0.2613
1.3777	0.8433	-0.2298	0.1237	1.4995	-0.0088	2.9657
-0.0227	-0.0170	0.0397	-0.0010	-0.0088	0.8485	0.0481
2.7225	1.6674	-0.4460	0.2613	2.9657	0.0481	5.8742

Fig. 5a. is an example of a 7 by 7 covariance matrix of a grayscale image

1.3575	0.7364	0.2435	0.0092	0.0890	-0.0021	0.1741
0.7364	0.5447	0.1558	-0.0026	0.0466	-0.0006	0.0904
0.2435	0.1558	0.0554	0.0003	0.0148	0.0004	0.0304
0.0092	-0.0026	0.0003	0.0027	0.0009	0.0000	0.0021
0.0890	0.0466	0.0148	0.0009	0.0065	-0.0006	0.0123
-0.0021	-0.0006	0.0004	0.0000	-0.0006	0.0100	0.0006
0.1741	0.0904	0.0304	0.0021	0.0123	0.0006	0.0244

Fig. 5b. is an example of a 7 by 7 covariance matrix of a Gabor filtered image

Figures 5a and 5b above show the covariance of grayscale image and for a Gabor filtered image, respectively. The latter used as feature representation is less in value which decreases computation time. In Table 1 below, we show random image distances compared to each other during the validation process. The distance values are rounded off to the tenths place for simplicity and clarity of the table. U stands for user and D for distance. Now we can interpret the table as User1 (U1), compared against its own to give the value at U1, D1. The other image distance values in D2 to D9 against particular users can be understood the same way. A particular incoming image can be compared against the whole images in the dataset. Recognition rate on the JAFFE database is 97.4 percent. On selected images from FRAV3D, the recognition rate was 99 percent. This fact is attributed to the constant background texture in all the images used.

TABLE I
SAMPLED COVARIANCE DISTANCE VALUES

User /dist	*D1*	*D2*	*D3*	*D4*	*D5*	*D6*	*D7*	*D8*	*D9*
U 1	1.0	1.3	4.3	3.5	1.5	3.8	3.4	4.6	7.9
U 2	2.8	0.5	4.1	3.2	1.8	3.7	2.7	5.5	8.1
U3	6.0	6.4	2.6	5.6	4.9	3.6	4.4	4.5	4.0

V. CONCLUSIONS

We have computed the covariance matrices of Gabor filtered images for similarity measurement between them. The property of these matrices in finding global optimum solution has been used to our advantage for feature extraction and representation. Gabor filtering process used to magnify and expose the trivial but very important features on the facial images. The dimensionality of the covariance matrix of Gabor images is exponentially reduced. Főrstner distance measure is employed for classification and recognition. A very impressive recognition rate is observed on the two databases used. We reckon that applying this method on more specific target region of interest would improve the method and accelerate recognition process.

REFERENCES

[1] M. Turk, A. Pentland,"Eigenfaces for Recognition", Journal Cognitive Neuro-science, vol. 3, pp. 71-86. 1991.

[2] M. Kirby and M. Sirovich," Application of the Karhunen-Loeve procedure for the characterization of human faces", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12(1): pp. 103-108, 1990.

[3] W. Zhao, R. Chellapa, A. Rosenfeld, P. J. Phillips," Face Recognition: A Literature Survey", Computing Surveys, vol. 35(4), pp. 399-458, 2003.

[4] J. Yang, Z. Jin, et al, "Essence of Kernel Fisher Discriminant: KPCA Plus LDA", Pattern Recognition Letter., vol. 37, pp. 2097-2100, 2004.

[5] Jian Huang, Pong C Yuen, Wen-Sheng Chen, J H Lai, "Kernel Subspace LDA with Optimized Kernel Parameters on Face Recognition," fg, pp.327, Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FG'04), 2004.

[6] Tai Sing Lee,"Image Representation Using 2D Gabor Wavelets", IEEE Transactions on Analysis and Intelligence. Vol 18. No. 10, October 1996.

[7] D.A. Pollen, S.F. Ronner," Visual cortical neurons as localized spatial frequency filters", IEEE Trans. Syst., Man. Cybern. SMC-13(5), pp. 907-916, 1983.

[8] Tuzel, O. Porikli, F., Meer, P.,"Region Covariance: A fast Descriptor for Detection and Classification", European Conference on Computer Vision (ECCV, 2006TR2005-111), May 2006.

[9] W. Forstner and B. Moonen," A metric for covariance matrices", Technical report, Dept. of Geodesy and Geoinformatics, Struttgart University, 1999.

[10] Face Recognition and Artificial Vision Group, http://www.frav.es/research/facerecognition/FRAV3D/.

[11] M. J. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba,"Coding Faial Expresions with Gabor Wavelets", Proceedings, Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara japan, IEEE Computer Society, pp. 200-205, April 1998.

[12] C. Grigorescu, N. Petkov, and M.A. Westenberg."Contour detection based on nonclassical receptive field inhibition", IEEE Trans. Image Processing, 12(7): 729 – 739, 2003.

[13] N. Petkov, and E. Subramanian,"Motion detection, noise reduction, texture suppression, and contour enhancement by spatiotemporal Gabor filters with sorround inhibition", Journal:Biological Cybernetics, vol. 97, no. 5-6, pp 423 – 439, 2007.

Wilfred O. Odoyo Received B.S and M.S degrees in Computer Engineering from Chosun University in 2005 and 2007, respectively. He is currently working toward the Ph.D. degree in the department of Computer Engineering at Chosun University. His research interests include Pattern Recognition, Image Segmentation, Wavelets, and Computer Vision.

Beon-Joon Cho Received B.S. and M.S. degrees in Electrical Engineering from Chosun University in 1980 and 1982, respectively. He received his Ph.D degrees in electrical Engineering from Hanyang University in 1988 and Electrical Engineering & Computer Science from KAIST in 2004. He is currently a professor in the Department of Computer Engineering at Chosun University. His research interests include the Pattern Recognition, Neural Network, Character Recognition, and Computer Vision.