Facial Recognition Algorithm Based on Edge Detection and Discrete Wavelet Transform

Min-Hyuk Chang, Mi-Suk Oh, Chun-Hwan Lim, Muhammad Bilal Ahmad, and Jong-An Park

Abstract : In this paper, we proposed a method for extracting facial characteristics of human being in an image. Given a pair of gray level sample images taken with and without human being, the face of human being is segmented from the image. Noise in the input images is removed with the help of Gaussian filters. Edge maps are found of the two input images. The binary edge differential image is obtained from the difference of the two input edge maps. A mask for face detection is made from the process of erosion followed by dilation on the resulting binary edge differential image. This mask is used to extract the human being from the two input image sequences. Features of face are extracted from the segmented image. An effective recognition system using the discrete wavelet transform (DWT) is used for recognition. For extracting the facial features, such as eyebrows, eyes, nose and mouth, edge detector is applied on the segmented face image. The area of eye and the center of face are found from horizontal and vertical components of the edge map of the segmented image. Other facial features are obtained from edge information of the image. The characteristic vectors are extracted from DWT of the segmented face image. These characteristic vectors are normalized between +1 and −1, and are used as input vectors for the neural network. Simulation results show recognition rate of 100 % on the learned system, and about 92% on the test images.

Keywords: mask, differential image, discrete wavelet transform, neural network, face recognition

I. Introduction

One area in computer vision that has grown significantly in importance over the past decade is that of computer face processing in visual scenes. Researchers attempt to teach the computer to detect, localize and recognize humane faces in a scene. The face is our primary focus of attention in social intercourse, playing a major role in conveying identity and emotion. Face detection and recognition from images is an active research area with numerous applications in user identification or verification, law enforcement and human-computer interaction. In the daunting task of human face processing, face detection is one of the most important problem to be solved. It is a prerequisite for automatic face recognition and expression analysis. In a recent survey, it is pointed out that segmentation of face regions from images is an important problem, which has received surprisingly little attention. Why is face detection an interesting problem? In terms of applications, face detection has direct relevance to the face recognition problem, because the first important step of a fully automatic human face recognizer is usually one of detecting and locating faces in an unknown image. For real-time applications, it is very important to locate faces in an image fast and accurately.

Many face detection approaches utilizes edge information of the image for face detection. Sakai [1] applied oval mask to edge map extracted from input image. He set the approximate head area, checked the edge image of eyes and mouth within the head area and then extracted the final head area. This method has a drawback that it is greatly affected by the direction of lighting. Kelly [2] also produced the downward image interpretation method that extracts the outline of head and body automatically from input image and continuously extracts the location of eyes, nose and mouth. Craw and others proposed method for extracting head area using mask of hierarchical size in a given image. Using the outline of head, which is composed of edge image, Govindaraju [4] searched the face in the complex background image. Govindaraju model the face as a set of three curves, namely, top, left, and right. Face detection is performed by detecting these curves and grouping them based on their relative position. Sirohey segmented the face by using edge image and brightness image, which are extracted with Canny edge searcher in the image with background. This method showed the precision of about 80% on 48 images without any restriction.

There are a variety of methods used in feature-based approaches and likewise a large variety of "features" which are considered suitable as a cue for the hypothesis of a face. These features are usually part of the human face anatomy that will give rise to a consistent image pattern across individual. Examples of such features could be the iris of the eye, the corners of the eyes, the whole eye itself, nostrils, nose tip, and the mouth. The principle of a feature-based approach is that in a human face, these features have a fixed relative position from each other and this geometrical relationship is more invariant to changes in face pose and identity (to a small extent) than other properties such as intensity or shape. For recognizing segmented images, feature extraction methods consist of local feature extraction and global feature extraction algorithms [6]. Local feature extraction algorithms include the method by geometrical symmetry, method using correlation of feature templates like eyes, nose and mouth and image, method detecting face to search face candidate edges with snake lets, method detecting features of face with self-organizing feature map and method extracting the feature from frequency domain like FFT, DCT. After extracting the features, generally, Euclidean distance and neural network is used to recognize the face. Euclidean distance implementation is easy, but its recognition rate is lowered when there are a lot of data. In neural network, input images are converted into spatial or frequency domain. Then, feature parameters of the image are extracted

Manuscript received: Mar. 10, 2001., Accepted: Aug. 21, 2001.
Min-Hyuk Chang, Mi-Suk Oh, Chun-Hwan Lim, Muhammad Bilal Ahmad, Jong-An Park: Chosun University(japark@chosun.ac.kr)
※This study was supported(in part)by research funds from chosun university, 2000.

and it is used with input vectors of neural network. This method has the properties that the number of nodes and connection lines can be reduced and system implementation is easy [7].

In this paper, edge detection is applied on the two input images. One input image contains only background, while the second image contains human being in the same background. A differential image of the two binary edge maps is obtained to form an oval shaped mask. The process of erosion and then followed by dilation removes noisy edges in the resulting binary edge differential image. This oval shaped mask is used for segmenting the face from the pair of input images. After getting the segmented face from the images, we applied edge detection to find the characteristic areas of face to reduce the computational complexity in extracting the characteristic. We extract the exact characteristic vectors using discrete wavelet conversion. This method reduces the number of characteristic vectors and decreases the required learning data with neural network. In section 2, we describe face segmentation algorithm based on binary edge differential image. Section 3 describes

II. Face segmentation using differential edge image

1. Face segmentation

A fast and near real-time system capable of detecting faces is one exploiting motion information. Motion-based algorithms assume that the background is stationary and the face is always in motion, and therefore the location of the face can be easily estimated using simple techniques such as frame differencing. Based on this assumption, the search space for the face can be largely reduced, thereby improving performance and processing time. In addition, detection can be made more reliable by checking for consistency between image frames. Though motion is a strong cue, it is nevertheless insufficient to base detection on motion alone. Other information such as skin color, edges are also helpful for detecting face. We used edge information of the images to get the human face, and human facial features. So the face of human being is extracted from the input images as follows. From the fixed CCD camera, two images are taken. One image is taken of the background, and the second is taken with the human being in the image. Noise in the input images are rejected using Gaussian filtering. The pixel values of the face in input image are different than that of background part of the image, so the difference image almost accurately detects face from the two input images. As we are mostly interested in the contour of human being face, we took the difference image of the binary edge maps of the two input images. Sobel edge detector is applied on the two input images. We applied the Sobel edge detector, because it gives both horizontal and vertical gradients. Other edge detectors can also be applied, but Sobel edge detector is simple in the operation and widely used in the applications. For removing the background and extracting the face of human being, we find the edge maps of the two input images, and find the binary difference image of the resulting two input edge maps, as:

$$D(x,y) = ABS(E_2(x,y) - E_1(x,y)), \tag{1}$$

where $E_1(x,y)$, is the binary edge map of the only background input image, while $E_2(x,y)$ is the binary edge map of the input image of human being with the same background. $D(x,y)$ is the resulting binary edge difference image of the two edge maps of the input image sequence. The resulting binary difference image $D(x,y)$ gives us the possible location of human being. But there are also some unwanted edges in the difference image due to change of illumination, or slight change of camera position. These unwanted edges are removed with the process of binary erosion, and then followed by the binary dilation. To remove the noise edges, the morphological erosion is applied on the difference image $D(x,y)$ as :

$$ER(x,y) = erode(D(x,y)), \tag{2}$$

where 'erode' is the morphological erosion, and $ER(x,y)$ is the eroded image. After erosion, we need dilation operation to get the mask for face of the human being as :

$$DL(x,y) = dilate(ER(x,y)), \tag{3}$$

where $DL(x,y)$ is the dilated image. An oval shaped mask is made from the dilated binary differential image. This mask is projected on the original input image that contains human being, and the face is segmented from the background. Fig. 1. shows the flow chart for segmenting the face.

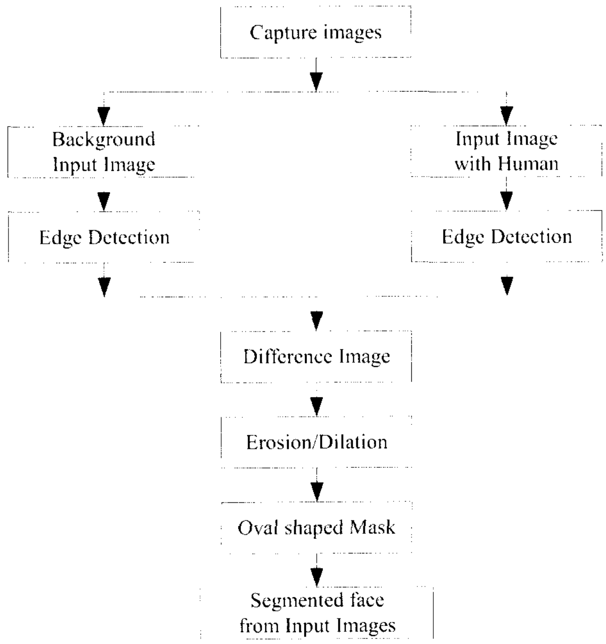


Fig. 1. Face segmentation from input image sequence.

2. Localization of facial features

For higher accuracy of face recognition, the characteristic facial area should include only distinguishing facial features, such as eyes, eyebrows, nose and lips. Head, hairs, and ears play very little role in human face recognition. In the 2-D front view of human being face, it is difficult to get the shape of ears. Second, human being's head hairs are not the constant feature of human being. Human being can cut or change the color of hairs. So, in short, eyes, eyebrows, nose and lips play the important role for face recognition. After segmenting the human being face from the input image sequence (see section 2.1), we need to find the locations of distinguishing characteristic features of face, such as locations of eyes, eyebrows, nose, and lips. We want to reduce the area of face recognition to increase the reliability of the face recognition system. We apply Sobel edge detector on the segmented face, and find the horizontal and vertical edges. The vertical edges are used to find the nose and hence find the center of the face. Horizontal edges are used to find the areas of eyes, eyebrows, and lips. We know from the contextual knowledge of human being face that both eyes are located on either side of the nose. Similarly, the location of lips is below the nose, and the position of nose is almost above the center of the lips. Since the size of human face at a constant distance from the camera is approximately same, so a constant square area (in experiments it is 91 x 91 pixels) is defined to include eyebrow, eyes, nose and lips and it is considered as the characteristic area of face recognition. The process of finding the approximate location of the characteristic facial features highly reduces the computation of wavelets, and also increases the accuracy of human being face recognition. Discrete wavelet is applied on the square region of the segmented face, the characteristic vectors of the face are found. These characteristic vectors are then used as input vectors for neural network. The flow diagram of finding the location of characteristic facial features is shown in Fig. 2.

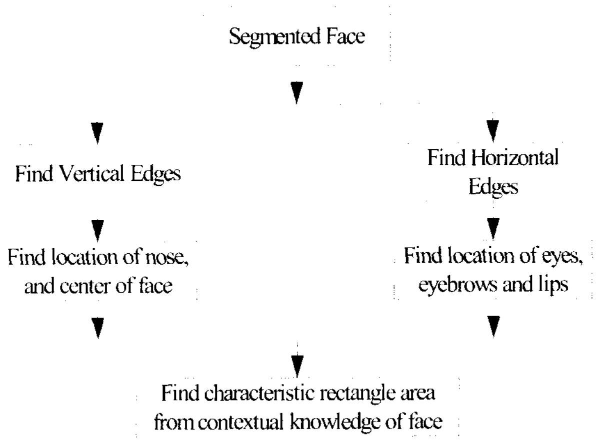


Fig. 2. Flow chart for characteristic area detection.

III. Face recognition algorithm based on wavelet transform

We now describe the method of recognizing the face image by using neural network. The input vectors for neural network are extracted by two level wavelet transform of the rectangle of characteristic features of the segmented face. Fig. 3. is flow chart for face recognition algorithm. The extraction process of feature parameters using wavelet conversion divides two input image signals with resolution of [256×256] into characteristic areas with block of 91×91 pixels using differential image and edge feature. The image size is 256 x 256 pixels, and the segmented face image rectangle size is 91 x 91 pixels. Then DWT coefficient matrix is obtained with discrete wavelet transform about 91×91 pixels. Feature vectors are extracted by using wavelet coefficient matrices.

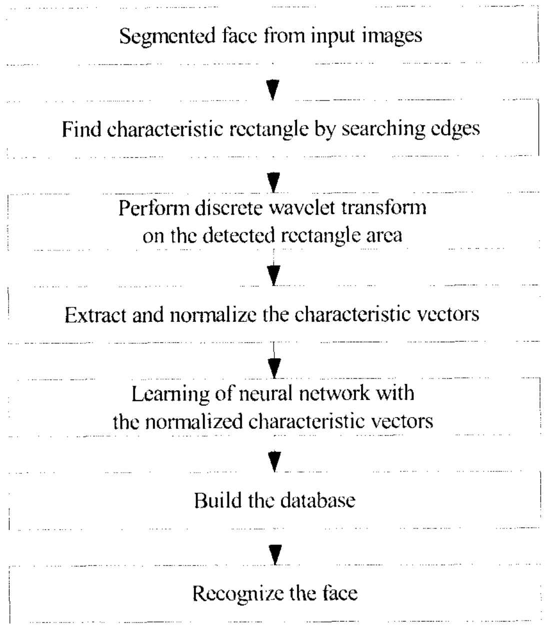


Fig. 3. Face recognition algorithm.

Calculating two-level DWT of the segmented facial characteristic areas, coefficient matrix with size 23x23 is obtained, which contains coefficients CA2, CH2, CV2 and CD2. The CA2 is two level low frequency coefficient matrix, CH2 is vertical direction high frequency coefficient matrix, CV2 is horizontal direction coefficient matrix and CD2 is diagonal high frequency coefficient matrix. After analyzing the distribution features of these matrixes, we can extract feature vectors. To obtain characteristic vectors using coefficient matrix of the same person's learning image, we get the sums of absolute values of four coefficient matrixes with size 23 by 23 and vertical direction, horizontal direction and diagonal direction as shown in Fig. 4.

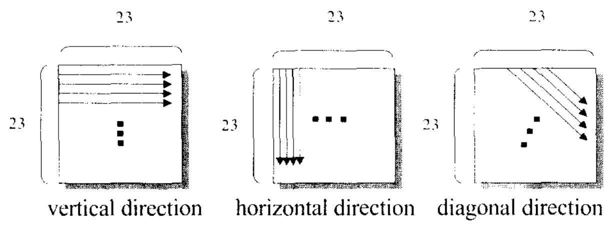
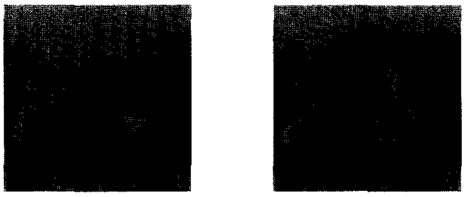


Fig. 4. Characteristic extraction using coefficient matrix.

IV. Simulation

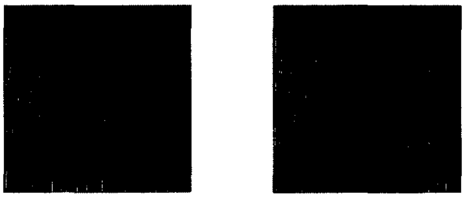
1. Face segmentation and characteristic detection

We simulated our algorithm on various test images. Two gray scale input images of 256 x 256 pixels size are taken from CCD camera at equal distance with regular illumination conditions. Noise of input images is rejected by Gaussian filter. A differential image is obtained from the two edge images, and then the face is segmented from the background image. Fig. 5. shows the two test input images, while Fig. 6. shows the corresponding edge maps of the input images. Fig. 7. shows the differential edge image, and the segmented face of the human being.



(a) (b)

Fig. 5. Input image.



(a) (b)

Fig. 6. Edge images.

(a) (b)

Fig. 7(a). Differential edge image, (b) Segmented face image of human being.

After segmenting the human face, we find the rectangle, which contains the distinguishing facial characteristic features, suc as, eyes, eyebrows, nose and lips. The position of eyes and eyebrow and lips are detected from horizontal distribution of edge components. The vertical distribution of edges finds center of the face, and position of the nose. Finally, the characteristic area is marked with a rectangle on the basis of the contextual knowledge of human face. Figure 8 shows marked rectangles on the input image.

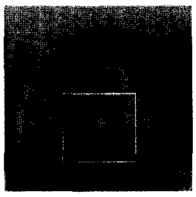


Fig.8. Facial characteristic areas marked by rectangles.

2. Face recognition

After calculating cumulative sum, from CA2, CH2, CV2 and CD2, the characteristic feature vectors are extracted; 46 characteristic vectors extracted from CA2 by an accumulative sum of diagonal direction, 23 from CH2 in horizontal direction, and 23 from CV2 in vertical direction, and 46 in diagonal direction. However, these characteristic vectors can not be used directly as input vector of neural network because of the sigmoid function characteristic. It requires normalization. So we normalize the vectors. The characteristic vector magnitudes are normalized between +1 and -1. Finally after calculating the mean square error of normalization vector extracted from CA2, CH2, CV2 and CD2, we utilize it as the learning vector of neural network based on the size of error.

We find the two-level DWT of the segmented face, and extracts 92 characteristic vectors (23 horizontal, 23 vertical and 46 diagonal). These 92 characteristic vectors are then normalized and used as input vectors for neural network. We took four pictures of a same person at different times, so that we can learn our network. The segmented faces of four sample images are shown in Fig.9. We analyzed the coefficient matrices, CA2, CH2, CV2, and CD2, and extracted 92 characteristic vectors for each of the sample images. Fig. 10. shows an image of size 23x23 of each wavelet coefficient matrices CA2, CH2, CV2, and CD2.

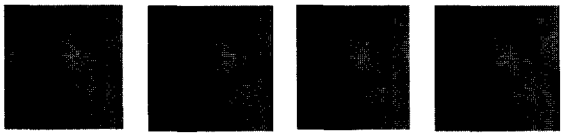


Fig. 9. Sample images.

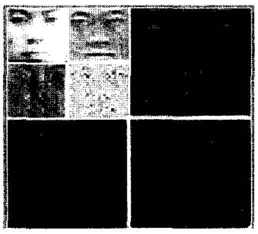


Fig. 10. Two step decomposition image.

Fig. 11. shows the extracted neural network input vectors form normalized characteristic vectors after DWT conversion of four sample images. The value of RMSE used between the vectors is 0.0030108.

The extracted 92 normalized vectors are input to back propagation learning algorithm. The number of layers is 2 and output layer is 0.005 and the learning rate is 0.7. The number of the learning image is 1801 to 2086. The target vector of the neural network is (1,-1,-1,-1,1,1,1,1). The weight of the generated network after learning and feature vector of input comparative image are operated. Comparing the error of the output layer performed the face recognition. When the error of output is smaller then 0.005 it is judged to be the same person and if it is more than 0.005 it is judged as a different person. The experiment of extracting and learning four people's characteristic vectors is performed in order to prove the learning recognition rate of the neural network.

Fig. 12(a) shows the various test segmented face images to test the performance of the algorithm. Table 1. shows results of recognizing sample images. Now to test the recognition rate of our algorithm, a random input image of one of four persons in the database is given to the face recognition system. The characteristic vectors of the input image are found, and the error between the output vector and target vector is calculated. If error is less than 0.005, the face is recognized as matched person, otherwise the face is unmatched.

From experimental results, it is found that the recognition rate is 100 % for the four learned images of the four images, while four non-learning images of 30 images gives 92% of recognition rate.

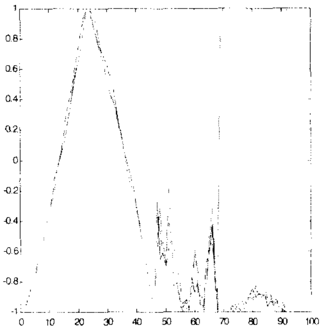
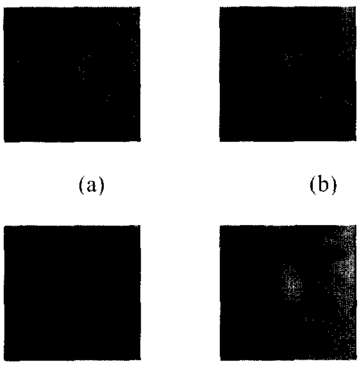


Fig. 11. Normalized characteristic vector of CA2,CH2 and CV2.



(a) (b)

(c) (d)

Fig. 12. Example of input images for face recognition.

Table 1. Recognition results of separated comparison image.

Output		error
Fig. 12(a)	+0.995166838	+0.004833162
	-0.995007694	-0.004992306
	-0.995122492	-0.004877508
	-0.995208919	-0.004791081
	+0.995001256	+0.004998744
	+0.995147288	+0.004852712
	+0.995214224	+0.004785776
	+0.995098591	+0.004901409
Fig.12(b)	+0.995122492	+0.004877508
	-0.994648218	-0.005351782
	-0.994692087	-0.005307913
	-0.994943976	-0.005056024
	+0.994760394	+0.005239606
	-0.995094240	+0.004905760
	-0.995186865	+0.004813135
	-0.994715452	+0.005284548
Fig. 12(c)	+0.995095313	+0.004904687
	-0.994617522	-0.005382478
	-0.994681001	-0.005318999
	-0.994945109	-0.005054891
	+0.994720876	+0.005279124
	+0.995042801	+0.004957199
	+0.995160282	+0.004839718
	+0.994695306	+0.005304694
Fig. 12(d)	+0.995150685	+0.004849315
	-0.994889736	-0.005110264
	-0.994985163	-0.005014837
	-0.995125353	-0.004874647
	+0.994912982	+0.005087018
	+0.995108247	+0.004891753
	+0.995193005	+0.004806995
	+0.994973660	+0.005026340

V. Conclusions

A new method of face recognition based on differential edge image and discrete wavelet transform is presented. In the algorithm, two gray-level images were captured in constant illumination. We used a Gaussian filter to remove noise of input images. We find edge maps of the input images, and then find differential edge image between background and input image. A mask was made from erosion and dilation process of the binary differential edge image. From projection of mask on the input image, we segmented human face from the input images. The distinguishing characteristic features of human face are in eyebrows, eyes, nose and mouth. Edge detection of the segmented face gives the location of characteristic facial features. The location of eyes, eyebrows, and lips are found from the horizontal components of edges, while the location of nose and the center of face were extracted from the vertical components of edges. A rectangular area is found on the human's face that includes the distinguishing facial features. Finally, characteristic vectors were extracted by performing DWT on the rectangular segmented region of the face and are normalized between +1 and -1. Normalized vectors were used with input vector of neural network. From simulation results, we have shown a recognition rate of 100 % on learned images and 92% on test images.

References

[1] T. Sakai, M. Nagao, and Fukibayashi : Line extraction and pattern recognition in a photograph : *Pattern Recognition*, vol. 1, pp. 233-248, 1969.

[2] M.D. Kelly : Visual identification of people by computer : Tech. Rep. AI-130, Stanford AI Proj., Stanford, CA, 1970.

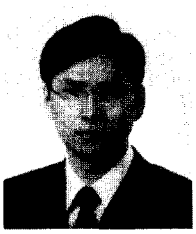
[3] I. Craw, H. Ellis, and J. Lishman : Automatic extraction of face feature : *Pattern Recognition Lett.*, vol. 5, pp. 183-187, 1987.

[4] V. Govindaraju, S. N. Srihari, and D. B. Sher : A computational model for face location : *Proc. 3rd Int. Conf. on Computer Vision*, pp. 718-721, 1990.

[5] S. A. Sirohey : Human face segmentation and identification : Tech. R. CAR-TR-695 : Center for Automation Research, Univ. of Maryland, 1993.

[6] Myung-Kil Lee, Joo-Sin Lee, "A study on pattern recognition using DCT and neural network" *The Journal of the Korean Institute of Communication Science*, vol. 22 no. 3, pp. 481-492, 1997.

[7] Myung-Kil Lee and Joo-Sin Lee, "Pattern recognition of SMD IC using wavelet transform and neural network" *Journal of the Korea Institute of Telematics and Electronics*, vol. 34-B no. 7, pp. 768-777, 1997.



Min-Hyuk Chang
Received the B.S. and M.S. degrees in Electronic Eng., from Chosun University in 1995 and 1997. He is currently proceeding towards Ph.D program in the Department of Electronic Eng. at Chosun Univ., Kwangju, Korea. His Research interests are in the field of Digital Signal Processing, Vision System motion estimation, motion tracking, and Pattern Recognition.

Chun-Hwan Lim
Received the B.S., M.S., and Ph.D. degrees from Chosun University in 1991, 1993, and 2000. Recently, He is a lecturer at Dong-Kang collage and Chosun Univ., Kwangju, Korea. His Research interests are in the field of Digital Signal Processing, Vision System and Pattern Recognition

Jong-An Park
Received the B.S., M.S., and Ph.D. degrees from Chosun University in 1975, 1978, and 1986. Since March 1978, he has been a Professor at the Division of Electronic Information & Commuications Eng., Chosun Univ., Kwangju, Korea. He was a visiting fellow at the Univ. of Mass., U.S.A. in 1983~1984 and the Univ. of Surrey, England in 1990~1991. His Research interests are in the field of Digital Signal Processing, Vision System motion estimation, motion tracking, and Pattern Recognition.

Mi-Suk Oh
Received the B.S. and M.S. degrees in Electronic Eng., from Chosun University in 1983 and 1999. He is currently proceeding towards Ph.D program in the Department of Electronic Eng. at Chosun Univ., Kwangju, Korea. Her Research interests are in the field of Digital Signal Processing, Vision System and Pattern Recognition.

Muhammad Bilal Ahmad
Was born in D.I.Khan, Pakistan in 1970. He received his M.S. degree in Mechatronics, from Kwangju Institute of Science and Technology, Kwangju, Korea in 1999. He received his B.S. degree in electrical engineering from University of Engineering and Technology, Peshawar, Pakistan, in 1993. He also received a post graduate diploma in computer systems software and hardware. He is currently proceeding towards Ph.D program in the Department of Mechatronics at Kwangju Institute of Science and Technology, (KJIST) Kwangju, Korea. His research interests include image processing, digital signal processing, machine/robot vision, data communications, motion estimation, motion tracking, and image retrieval.