# A Facial Image Segmentation for Video Coding and its Recognition Based on DWT

Chun-Hwan Lim*, Jong-An Park* *Regular Members*


## 요 약

이 논문에서는 잡음에 대해 유연성이 있는 신경망과 차영상법-DCT를 이용한 얼굴인식 알고리즘을 제안한다. 동일환경(조도의 세기, 얼굴에서 카메라 까지의 거리)에서 연속적으로 두 개의 영상을 캡쳐했다. 이 때 한 영상은 얼굴을 포함하지 않고 다른 영상은 얼굴을 포함하게된다. 차영상 방법을 이용하여 두 개의 이미지로부터 얼굴영상과 배경영상을 분리하고 그 다음에 분리된 얼굴영역에서 사각영역을 추출하여 이 영역을 얼굴의 특징영역으로 이용하였다. 이 사각 영역은 눈, 코, 입, 눈썹 등이 포함된다. 다음으로 이 영역에 대해 DWT 연산을 수행한 후 특징 벡터를 추출하였고 추출된 특징벡터는 정규화되어 신경망의 입력벡터로 사용되었다. 시뮬레이션 결과 학습된 얼굴영상에 대해서는 100% 인식률을 보였고 학습되지 않는 얼굴 영상에 대해서는 92.25%의 인식률을 보였다.


## I. INTRODUCTION

Edge image is mainly used in the early methods of discriminating whether there is face or not in the image.

Sakai applies oval mask to edge map extracted from input image, binds the approximate head area, checks the edge image of eyes and mouth within the head area and then extracts the final head area[1].

This method has the weakness that it is influenced greatly by the direction of lighting. Kelly also introduces the downward image interpretation method that extracts the outline of head and body automatically from input image and continues to extract the location of eyes, nose and mouth[2].

Craw and others proposed the method extracting head area by means of image and mask of hierarchical size in a give image[3].

Govindaraju used head outline composed of edge image as mask to search the position of face in the image of complex background[4].

Sirohey segments the face by using edge image and brightness image which are extracted with edge searcher from the image with background[5].

This method shows the precision of about 80% of 48 images without any restriction. But the recognition of the face is not easy with such a face segmentation when extraction algorithm is complex and the face is not segmented exactly.

In case of recognizing segmented images, feature extraction methods include local feature extraction of image outline or a specific feature and global feature extraction algorithm of feature parameters after blocking image.

Feature location detection methods include the method by geometrical symmetry, method using correlation of feature templates like eyes, nose and mouth and image, method detecting face with face candidate edges with snakelets, method detecting face features with self-organizing feature map and method extracting the feature from frequency areas like FFT, DCT and DWT.

The extraction and recognition of features use method using differential image and neural network. While the former has the strength that system implementation is easy, it has also weakness that its recognition rate is lowered when there are a lot of data.

* 조선대학교 전자정보통신공학부(japark@chosun.ac.kr)
논문번호 : 99270-0703, 접수일자 : 2000년 7월 3일
※ 이 논문은 1999년도 조선대학교 학술연구비의 지원을 받아 연구되었음.

In case of pattern recognition with the latter, since it is converted into spatial or frequency area depending on input object, feature parameter of an object is extracted and fractional input vector is used with neural network, the number of nodes and connection lines can be reduced and then the reduction of process time and system implementation are easy[6].

Therefore this study applies differential image when the use of two input images is possible, generates mask and proposes face segmentation methods.

In addition, it defines the characteristic areas of face and reduces the computational complexity in extracting the features.

It is also to extract elaborate feature vector by means of wavelet conversion, reduce the number of feature vectors, decrease learning data required in learning neural network and cope with noise and minute illumination change flexibly.

For these, chapter 2 suggests face segmentation algorithm based on differential image, chapter 3 face recognition algorithm based on wavelet and chapter 4 performs simulation and chapter 5 is conclusion.

## Ⅱ. Face Segmentation Algorithm Based on Differential Image

### 2.1 Face Segmentation Using Differential Image and Mask

The face is segmented by applying differential images in two input images including background image obtained in regular illumination condition and face.

For this image in the same distance is obtained with gray scale 256 level of the size of 256×256 and the noise within the image is rejected by Gaussian filtering.

And since the pixel value outside the face in input image including the face is not identical with that of background image which doesn't include the face exactly in obtaining the difference of input images including background image and face, threshold is given as shown in the following expression (1) and the pixel value of the image obtained with the same camera can get the desired image though it is changed a little.

$$\text{if } |Image1(x,y) - Image2(x,y)| < threshold \text{ then } Differ\_Image(x,y) = 0$$
$$\text{else } Differ\_Image(x,y) = |Image1(x,y) - Image2(x,y)| \quad (1)$$

for, $Image1$, $Image2$ : input image, $Differ\_Image$: differential image

The noise occurred by the change of illumination and the reflection of light occupies small area in the image because of low frequency of noise and it makes binary differential image and erodes the image.

Eroded binary image should be dilated to the size of face image.

Then pixel value is examined, mask is made, it is projected on the original image with the face and the face is segmented in the background. Fig. 1 shows the flow chart to segment the face.

### 2.2 Feature Detection of Face

Most feature information of human face is in eyebrow, eyes, nose, mouth and cheek. So the face features are detected after selecting local area including this area.

For the feature detection of segmented face image, edge is detected with sobel operator and eye area and the center of face are searched by using horizontal and vertical components of edge.

Further since the size of human faces in regular distance is similar, square area should be defined to include eyebrow, eyes, nose and cheek like Fig. 2 and it is considered as the feature area of face recognition.

After the conversion of discrete wavelet of extracted feature areas, extract the features. a(a=c+e+f) in Fig. 2 indicates horizontal distance of characteristic area and b the vertical distance.

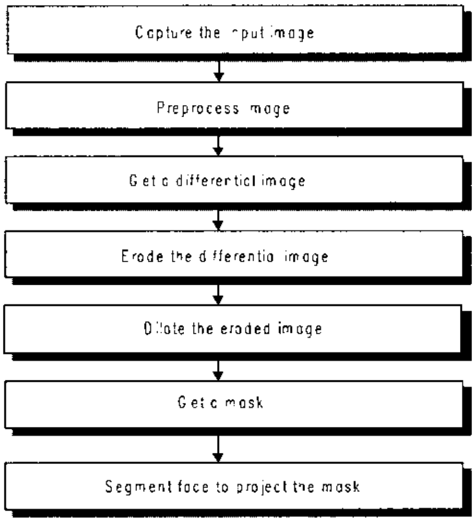



Fig. 1 Flow chart for facial image segmentation

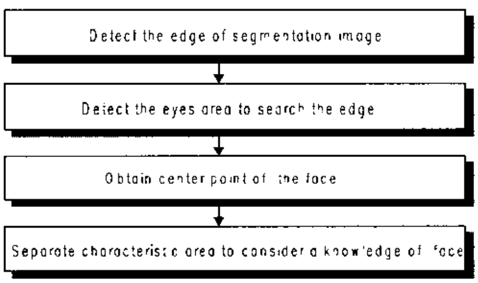



Fig. 2 Flow chart for characteristic detection

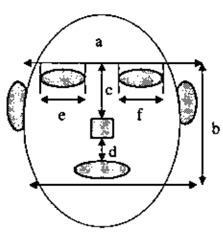



Fig. 3 The normalized distance for characteristic detection.

## Ⅲ. Face Recognition Algorithm Based on Wavelet Transform

This chapter proposes the method of recognizing the face image by using neural network after extracting the characteristics of face in wavelet conversion area. Fig. (3) shows the algorithm flow chart of face recognition. Wavelet is obtained by dilate/translating mother wavelet $\psi(x)$ and it is shown in Expression 2.

$$\Psi_{a,b}(x) = \frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right) \tag{2}$$

where, a is scaling parameter, b translation parameter and $\sqrt{a}$ normalization factor. Wavelet conversion Wf(s) of random signal f(x) is defined as convolution integration with original signal f(x) and wavelet $\psi(x)$ as shown in Expression 3.

$$Wf(x) = f * \psi(x) \tag{3}$$

In addition, when $\psi(x,y) = \psi(x)\psi(y)$ is two-dimensional scaling function in wavelet conversion of two-dimensional function f(x, y) and one-dimensional wavelet with one-dimensional scaling function $\psi(x)$ is $\psi(x)$, two-dimensional wavelets can be explained to separable multi-resolution approximation as shown in Expression (4).

Two-dimensional signal divided into normal orthogonal distance of Expression (4) is divided into frequency component of spatial direction. [8]

$$\begin{aligned} \phi^{A}(x,y) &= \phi(x)\ \phi(y) \\ \phi^{H}(x,y) &= \phi(x)\ \psi(y) \\ \phi^{V}(x,y) &= \psi(x)\ \phi(y) \\ \phi^{D}(x,y) &= \psi(x)\ \psi(y) \end{aligned} \tag{4}$$

The extraction process of feature parameters using wavelet conversion divides two input image signals with resolution of $[256 \times 256 \times 2^8]$ into characteristic areas with pixel of $91 \times 91$ using differential image and edge feature. Then DWT coefficient matrix is obtained with DWT. The following fig. 4 shows the distribution of coefficient matrix of 4 level DWT, where cA4 means coefficient matrix of 4 level low frequency, cH(i) horizontal high frequency coefficient matrix of (i) level, cV(i) vertical high frequency coefficient matrix of (i) level and cD(i) diagonal high frequency coefficient matrix of (i) level. Feature vectors are extracted by using such coefficient matrixes.

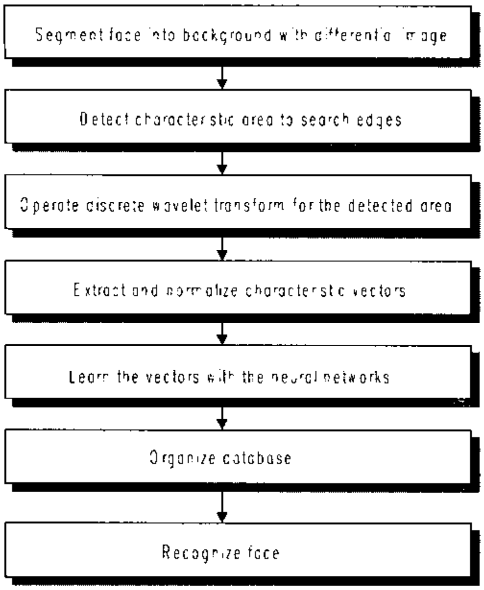



Fig. 4 Flow chart for face recognition algorithm

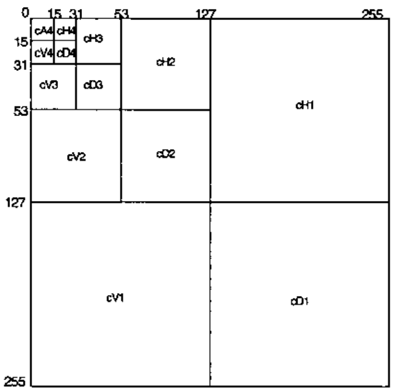



Fig. 4 Four Level DWT Coefficient Matrix

When four level DWT of segmented facial characteristic areas is performed, cA4, cH4, cV4 and cD4 which are coefficient matrixes with the size of 6×6 are obtained, where cA4 is four level low frequency coefficient matrix, cH4 horizontal high frequency coefficient matrix, cV4 vertical high frequency coefficient matrix and cD4 diagonal high frequency coefficient matrix, and after analyzing the distribution features of these matrixes, extract feature vectors.

First to examine the distribution of coefficient matrix of the same person's learning image, four level DWT is performed with four sample learning images. Then get the absolute values of coefficients of four coefficient matrixes with the size of 6×6, extract 36 feature vectors and normalize them between +1 and -1. Finally after calculating mean square error of normalization vector extracted from cA4, cH4, cV4 and cD4, utilize it as learning vector of neural network based on the size of error.

## Ⅳ. SIMULATION

### 4.1 Face Segmentation and characteristic Detection

This study acquires the input image in the same distance from CCD camera at regular illumination condition with gray scale 256 level of 256×256 and simulation is performed on PC(266MHz).

For this, the noise of input image like Fig. 5 and 6 is rejected by using gaussian filter, differential image is obtained and then the face is segmented in background image.

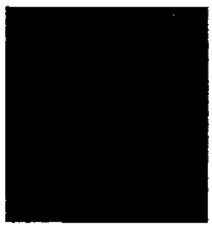


Fig. 5 Input Image(I)

Fig. 6 Input Image (II)

Then the differential image of two input images is obtained. Fig. 7 shows the resulting differential image and Fig. 8 is binary image.

Fig. 7 Differential Image

Fig. 8 Binary Image

Since the differential image cannot be immediately used as mask, it is binary as shown in Fig. 8 and the boundary is reduced. But since the reduction of boundary results in that of face area, the eroded(reduced) image is dilated as shown in Fig. 9. Fig. 10 is mask image generated by examining the pixel value from dilated image.

Fig. 9 Dilated Image

Fig. 10 Mask Image

And mask image is projected on original image with face and the face is segmented in the background as shown in Fig. 11.

Fig. 11 Segmentation Image

Fig. 12 Edge Image

After detecting edge using sobel operator in segmented face image, detect the position of eyes and eyebrow by horizontal distribution of edge components and segment the lower part of eyebrow. Then get the vertical distribution of edge and the center line of face.

Finally, the characteristic areas are detected on the basis of the knowledge of human face as shown in Fig. 13.

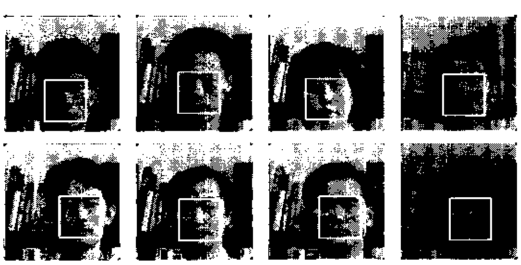
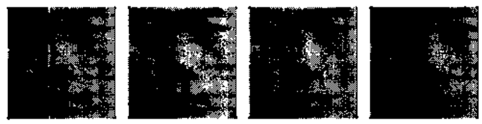


Fig. 13 Facial Characteristic Area

## 4.2 Face Recognition Experiment

Fig. 14(a) ~ 14(d) are characteristic area separated from four experimental images of the same figure and their size is 91×91 respectively.



(a) (b) (c) (d)

Fig. 14 Example of Separated Characteristic Image

Fig. 15 shows the image of coefficient matrixes of four level wavelet conversion cA4, cH4, cV4 and cD4 and the information of original image is concentrated on these four coefficient matrixes. The size of these coefficient matrixes is 6×6 respectively.

For feature extraction, four learning samples of the image in Fig. 14(a) ~ 14(d) are selected, they are converted into four level wavelet and then cA4, cH4, cV4 and cD4 are obtained. Then after analyzing normalization data distribution of the same coefficient matrixes, extract characteristic vector and it is the input vector of neural network by normalizing it between +1 and -1.

Fig. 16 shows the distribution condition of normalized characteristic vector. Normalization is performed after taking the absolute value of 6×6 coefficient matrix obtained from four sample images, where horizontal direction is the number of normalized characteristic vectors (36) and vertical direction is the range of normalization between +1 and -1.

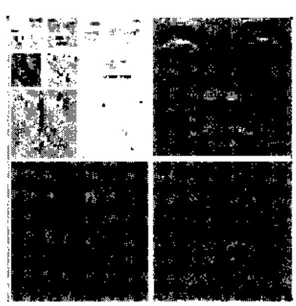


Fig. 15 Four Step Decomposition Image

Error among normalized vectors of the same coefficient matrix is obtained to select input vector of neural network of coefficient matrixes from four sample images like figure 14(a)~14(d). First, RMSE of normalized characteristic vector is obtained in cA4. Then RMSE of cH4, cV4 and cD4 is obtained through the same process. The following Table 1 shows RMSE of each coefficient matrix.

Total 108 of each 36 from cA4, cH4 and cV4 are extracted as characteristic vectors based on Table 1 and these are the input vectors of neural network. Fig. 17 shows the distribution of 108 neural network input vector extracted from four sample images of the same person.

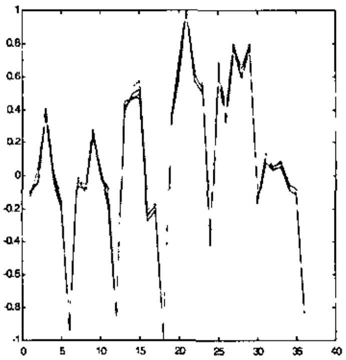


Fig. 16 Normalized Characteristic Vectors

Table 1. RMSE of normalized feature vectors

| Class | cA4 | cH4 | cV4 | cD4 |
|---|---|---|---|---|
| RMSE | 0.003575 | 0.004867 | 0.006381 | 0.006580 |

To confirm the validity of characteristic vector extraction, four images of three persons are obtained, wavelet conversion is conducted and the normalized results of coefficient matrix are shown.

For this, four sample images of experimental image like Fig. 17(a), 17(b) and 17(c) are selected, four level wavelet conversion of theses images is performed, 6×6 coefficient matrix is normalized and 36 characteristic vectors are extracted. And RMSE among the same coefficients is obtained.

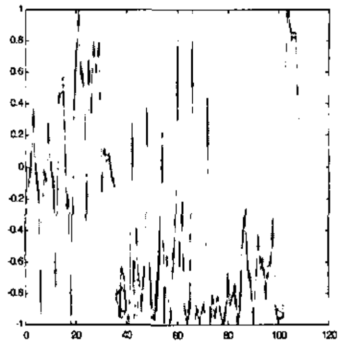


Fig. 17 Extracted Input Vectors

Table 2, 3 and 4 are RMSE of normalized vectors. The largest error is found in normalized vector of cD4.

Table 2. RMSE of normalized feature vectors

| Class | cA4 | cH4 | cV4 | cD4 |
|---|---|---|---|---|
| RMSE | 0.003076 | 0.005565 | 0.003042 | 0.006726 |

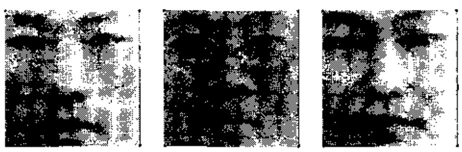
Table 3. RMSE of normalized feature vectors

| Class | cA4 | cH4 | cV4 | cD4 |
|---|---|---|---|---|
| RMSE | 0.001788 | 0.003830 | 0.002196 | 0.006224 |

Table 4. RMSE of normalized feature vectors

| class | cA4 | cH4 | cV4 | cD4 |
|---|---|---|---|---|
| RMSE | 0.000848 | 0.002928 | 0.001082 | 0.003476 |

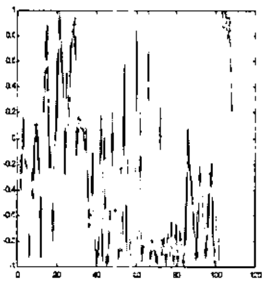
Fig. 19(a), 19(b) and 19(c) shows the neural network input vector extracted from Fig. 18(a), 18(b) and 18(c).



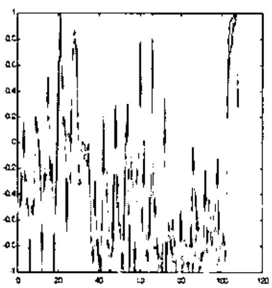
(a) (b) (c)

Fig. 18 Example of experimental images.

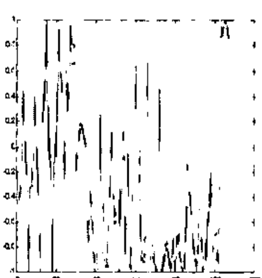
Such 108 normalized input vectors are multi-layer neural network. Learning algorithm of neural network uses error Backpropagation learning algorithm. The error of output layer is 0.005 and learning rate is 0.7. Weight of network generated after learning and feature vector of input comparative image are operated and the recognition is performed by comparing the error of output layer. Then when the error of output layer is less than 0.005, it is judged as the same person.



(a)



(b)



(c)

Fig. 19 Extracted input vectors

Test image (I) of Fig. 20 presents the type of recognition image and that (II) of Fig. 21 is the example of error recognition image, which shows inclined face.

Fig. 20 Test Image (I)

Fig. 21 Test Image (II)

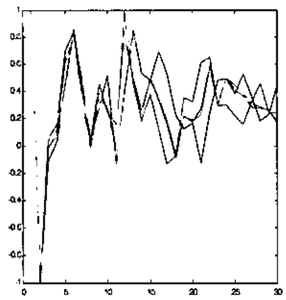
As a result of experiment with 4 learning images per person and 4 non-learning images of 30 images, 120 learning images show 100% of recognition rate and 120 non-learning images show 92% of recognition rate.

### 4.3 Comparison and Examination

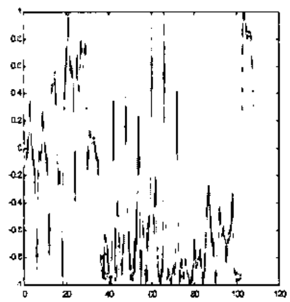
The simulation results of proposed algorithms are as follows.

1) Facial image can be segmented 100% in proposed experimental environment by using face segmentation algorithm based on differential image.

2) More elaborate feature vector can be obtained than by using DCT and the recognition rate increases 4%. Fig. 31 shows that the normalization characteristic of feature vectors extracted after wavelet conversion is more evenly distributed than the distribution characteristic of extracted after performing DCT of the same person.



(a) In case of DCT



(b) In case of 4 level DWT

Fig. 22 Distribution of Characteristic Vectors

3) Features are extracted by converting the data of spatial area into frequency area and data quantity can be reduced. Also data quantity for operation is reduced by handling characteristic area not the whole face in performing DWT.

4) Since the wrong recognition of included face image occurs, the compensation of inclination is required. And algorithm such as pass word input or fingerprint recognition is required to apply it to perfect complementary system.

## V. CONCLUSIONS

This study suggests how to segment the face when there is the man under complex environment, extracts the features from segmented image and proposes the effective recognition system applying the wavelet conversion.

First, get the differential image of two input images obtained in regular illumination condition and segment the face in background image. For this, get the image in the same distance from CCD camera with gray scale 256 level of 256× 256 size and reject the noise within the image by means of gaussian filter. And the difference between input images including background image and face is obtained, differential image is dilated and eroded, pixel value is examined after rejecting the noise within the face and mask is generated. Then mask is projected on input image and feature area is extracted from facial image after segmenting the face and the features are extracted from coefficient matrix obtained by applying wavelet conversion. Finally the feature vectors are normalized between +1 and -1 and learned by input to multi-layer neural network.

As a result of simulation, the face of input image is segmented completely and the recognition rate using DCT algorithm shows 100% of 120 learning images and 99% of 120 experimental images, but when the recognition is performed with wavelet conversion, high recognition rate is shown as 100% of 120 learning images and 92% of 120 experimental 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, 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

임 춘 환(Chun-Hwan Lim)

1991년 2월 : 조선대학교 전자공학과 졸업

1993년 8월 : 조선대학교 전자공학과(공학석사)

2000년 2월 : 조선대학교 전자공학과(공학박사)

1999년 3월 ~ 현재 : 동강대학 전자과 겸임교수

<주관심 분야> 영상통신, 의료영상, 패턴인식, 신경망 및 로보틱스

박 종 안(Jong-An Park)

1975년 2월 : 조선대학교 전자공학과(공학사)

1978년 2월 : 조선대학교 전자공학과(공학석사)

1986년 2월 : 조선대학교 전자공학과(공학박사)

1993 ~ 현재 : 조선대학교 전자정보통신공학부 교수

<주관심 분야> 디지털 신호처리, 패턴인식, 디지털 통신.