

방사형 기저함수 신경회로망 기반 숫자 인식 시스템의 설계 : 전처리 알고리즘을 이용한 인식성능의 비교연구

Design of Digits Recognition System Based on RBFNNs : A Comparative Study of Pre-processing Algorithms

김 은 후* · 김 봉 연* · 오 성 권†

(Eun-Hu Kim · Bong-Youn Kim · Sung-Kwun Oh)

Abstract - In this study, we propose a design of digits recognition system based on RBFNNs through a comparative study of pre-processing algorithms in order to recognize digits in handwritten. Histogram of Oriented Gradient(HOG) is used to get the features of digits in the proposed digits recognition system. In the pre-processing part, a dimensional reduction is executed by using Principal Component Analysis(PCA) and (2D)2PCA which are widely adopted methods in order to minimize a loss of the information during the reduction process of feature space. Also, The architecture of radial basis function neural networks consists of three functional modules such as condition, conclusion, and inference part. In the condition part, the input space is partitioned with the use of fuzzy clustering realized by means of the Fuzzy C-Means algorithm. Also, it is used instead of gaussian function to consider the characteristic of input data. In the conclusion part, the connection weights are used as the extended type of polynomial expression such as constant, linear, quadratic and modified quadratic. By using MNIST handwritten digit benchmarking database, experimental results show the effectiveness and efficiency of proposed digit recognition system when compared with other studies.

Key Words : Histogram of Oriented Gradient(HOG), Principal Component Analysis(PCA), pRBFNNs, Fuzzy C-Means(FCM), Least Square Estimation(LSE)

1. 서 론

생체인식 기술이란 지문, 손, 얼굴, 홍채, 망막, 정맥 등의 신체적인 특성과 글씨체, 음성, 걸음걸이 등의 행동학적 특성 등 개별적인 특성을 이용해 보안시스템에 활용하는 기술이다. 현재 사용되고 있는 대부분의 인식 기술의 경우 암기를 해야 하거나 분실 및 도용 등 노출이 될 수 있다는 단점이 있다. 따라서 인식 기술은 최신의 네트워크 환경과 정보 발달에 의한 보다 확실하고 정확한 사용자 인식 기술을 필요로 한다. 인간이 가지는 고유한 특징을 이용하는 생체인식 기술은 분실 및 도난 등의 문제가 발생하지 않아 다양한 분야에서 사용되고 있으며, 인간의 뇌신경세포와 구조를 모방하여 처리하는 신경회로망과 같은 연구 또한 활발히 진행되고 있다. 그 중에서 수기로 작성된 필기체 문자 및 숫자 인식은 수표, 서명, 우편물 등과 같은 다양한 곳에 적용이 가능하여 패턴인식 분야에서 중요성이 더욱 강조되고 있다.

최근 Deep Learning 기법을 이용하여 인식을 개선하기 위한 다

양한 방법들이 연구되어 좋은 결과를 얻고 있지만, 구조가 복잡하고 고성능의 하드웨어를 필요로 한다는 단점이 있다. 또한 문자 및 숫자인식을 위한 데이터는 고차원으로서는 많은 노이즈를 포함하고 있기 때문에 인식률을 저하 시킬 수 있다.

본 연구에서는 이러한 단점을 보완하며 인식을 저하를 방지하기 위해 Edge성분의 벡터 기울기 및 크기와 같은 특징을 Histogram of Oriented Gradient(HOG) 알고리즘[1]으로 추출한다. 그리고 차원 축소 알고리즘인 주성분 분석법(Principal Component Analysis: PCA)[2,3]과 행 방향과 열방향을 각각 (2D)PCA를 이용하여 2차원 이미지의 특징을 양쪽 방향에서 축소하는 (2D)²PCA[4]를 이용하여 고차원의 숫자 이미지를 저차원으로 차원 축소하고, 이를 제안된 방사형 기저함수 신경회로망(Radial Basis Function Neural Networks: RBFNN) 패턴 분류기의 입력으로 이용한다. 일반적인 RBFNNs 은 입력층 은닉층 출력층으로 구성되어 Multi-layer 신경회로망보다 간결한 구조로 이루어져 있으며 최소자승법을 통해 연결가중치의 학습을 수행하여 비교적 짧은 학습시간을 필요로 한다. 따라서 다양한 학습방법 및 다른 알고리즘과의 결합이 용이하여 적절한 구조 선택 및 다양한 알고리즘과의 결합을 통해 시너지 효과를 기대할 수 있다. 본 연구에서는 전처리 및 클러스터링 알고리즘을 RBFNNs 에 결합하여 hybrid RBFNNs을 설계하였으며 MNIST 데이터를 이용해 제안된 분류기의 성능평가를 수행하였다. 기존의 RBFNNs 패턴분류

† Corresponding Author : Dept. of Electrical Engineering, The University of Suwon, Korea

E-mail: ohsk@suwon.ac.kr

* Dept. of Electrical Engineering, The University of Suwon, Korea.

Received : March 2, 2016; Accepted : January 10, 2017

기를 확장한 다항식 기반 RBFNNs 패턴분류기[5,6]는 FCM (Fuzzy C-Means) 알고리즘[7,8]을 통해 은닉층 활성화수의 형태를 표현하여 패턴분류기의 전반부 구조 동정을 수행한다. 상수항이었던 연결가중치의 값은 일차식, 이차식, 변형된 이차식과 같은 3가지 다항식의 형태로 확장하여 후반부 구조 동정을 수행하며, 이와 같은 다항식 형태의 연결가중치를 사용함으로써 퍼지 규칙과 같은 언어적 관점에서의 해석이 가능해졌다. 후반부 파라미터 동정은 각 규칙에 대한 파라미터를 Least Square Estimation(LSE) 방법[9]을 사용하였다. 마지막으로 제안된 패턴 분류기의 성능평가를 위해 숫자 데이터인 MNIST 데이터를 사용하였으며, 기존의 제안된 Neural Networks, Linear Classifier, Quadratic Classifier, Convolutional Net 등의 분류기와 비교한다[10-13].

본 논문의 구성은 다음과 같다. 2장에서는 데이터의 전처리 과정으로서 HOG를 이용한 특징벡터의 검출과 PCA 및 (2D)²PCA를 이용한 숫자 이미지의 차원축소에 대해 설명한다. 3장에서는 인식을 위한 pRBFNNs 패턴 분류기의 구조 설계를 위해 전반적인 구조를 알아보고 전반부와 후반부 동정에 대해 설명한다. 4장에서는 시뮬레이션 및 결과 고찰로 실험 전체의 개요와 연구에 사용되는 MNIST 숫자 데이터에 대해 설명하고, 제안된 패턴분류기의 파라미터 설정을 통하여 성능을 확인 및 평가한다. 그리고 5장에서 결론에 대하여 설명하며 마무리를 한다. 아래의 그림 1은 본 논문에서 제안하는 숫자 인식 시스템[14,15]의 전반적인 흐름도를 나타낸다.

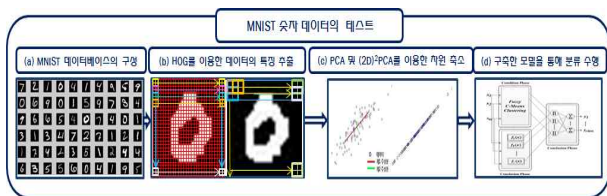


그림 1 숫자인식 시스템의 전체 구성
Fig. 1 Overall flow of digits recognition system

2. 숫자 이미지를 위한 전처리 방법

본 논문에서는 숫자 인식을 위한 전처리 과정으로서 입력 이미지로부터 HOG 알고리즘을 사용하여 특징을 추출하는 방법에 대하여 소개하고, 획득한 고차원의 특징을 저차원으로 축소하기 위한 차원축소 알고리즘 PCA와 (2D)²PCA를 설명한다.

2.1 숫자 이미지의 HOG 특징 추출

Histogram of Oriented Gradients(HOG) 알고리즘은 2005년 Dalal과 Triggs가 보행자를 검출하기 위해 제안된 특징으로 국소 영역에서 근접화소들 간의 밝기 차이를 이용하여 기울기와 방향성을 구하고 이를 히스토그램으로 구성하여 벡터를 나타낸다. 이러한 HOG 특징은 히스토그램으로 구성하기 때문에 조명의 영향에 민감하지 않고, 기하학적 변화에 강한 특징을 가진다.

[Step 1] 입력 영상을 Cell과 Block 단위로 분할
기울기의 방향성과 크기를 계산 후 히스토그램을 구성하기 전에 Cell과 Block을 정의하여 입력영상을 분할한다. 그림 2는 입력영상에서의 $C \times C$ 의 Cell 크기와 $B \times B$ 의 Block으로 구성된 영상분할을 나타내며, 본 논문에서는 Cell과 Block을 각각 2×2 로 동일하게 구성하였다.

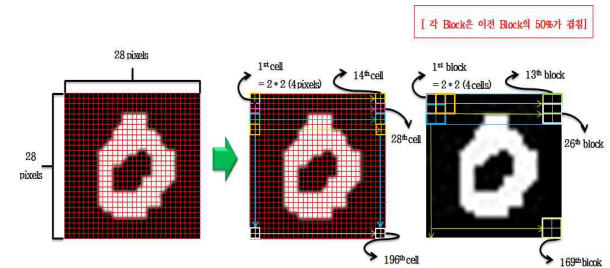


그림 2 Cell과 Block 단위 영상 분할
Fig. 2 Image segmentation of cell and block units

[Step 2] Block 내 Gradient의 크기와 방향 연산 및 히스토그램 생성

근접화소들 간의 기울기와 방향성을 이용한 특징 추출에 앞서 입력영상에서 식(1)을 이용하여 x축과 y축의 기울기 f_x, f_y 를 구하고, 기울기의 크기 S 는 식 (2)를 통해 구할 수 있다. 그리고 식 (3)은 각각의 축 f_x, f_y 와 크기 S 를 이용하여 기울기의 방향성 θ 를 구한다.

$$f_x(x,y) = I(x+1,y) - I(x-1,y),$$

$$f_y(x,y) = I(x,y+1) - I(x,y-1) \tag{1}$$

$$S(x,y) = \sqrt{f_x(x,y)^2 + f_y(x,y)^2} \tag{2}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{f_y(x,y)}{f_x(x,y)}\right) \tag{3}$$

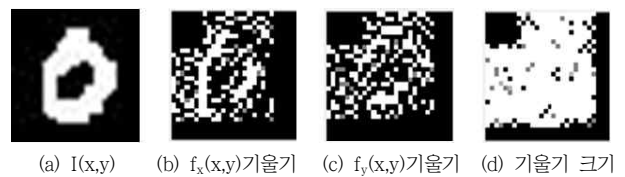


그림 3 HOG를 이용한 기울기와 기울기의 크기
Fig. 3 Gradients and gradient magnitude using HOG

그림 3의 (a)는 28×28 로 구성된 입력영상을 나타내며, (b)는 입력영상 $I(x, y)$ 에 대한 각 축의 기울기 $f_x(x, y), f_y(x, y)$ 로 나타낸다. 그리고 (c)는 기울기의 크기 $S(x, y)$ 를 의미한다.

히스토그램을 구성할 때 기울기의 방향 성분을 이용하여 bin을 나타내는데, 기울기의 방향성분은 크게 $0 \sim 180^\circ, 0 \sim 360^\circ$ 로 나타내며, 보통 $0 \sim 180^\circ$ 를 20° 만큼 나눈 9개의 bin으로 구성한다. 본 논문에서 히스토그램은 그림 4의 (a)와 같이 한 Cell에서

하나씩 생성하고, 이 히스토그램은 0~180°의 기울기의 방향 성분을 이용한 9개의 bin을 생성한다. 그리고 그림 4 (a)와 같이 이미지 안에서 셀 내부의 bin에 해당하는 기울기 크기를 누적시킨다. 또한, 크기를 누적 시키는 과정에서 Aliasing에 의한 정보 손실을 줄이기 위해 선형 보간 방식을 사용하였고 그 예로 그림 4 (b)의 95° 방향성을 가진 기울기 크기에 인접하고 있는 bin 90°과 110°에 나누어서 누적한다.

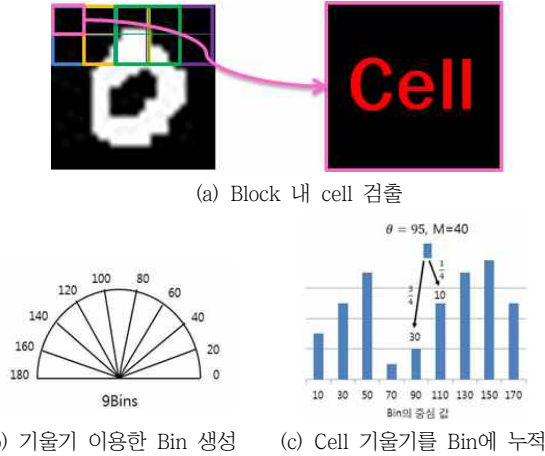


그림 4 Cell의 HOG 특징과 히스토그램
Fig. 4 HOG feature of cell and histogram

[Step 3] 전체 히스토그램 병합

각 Block에 대하여 정규화 된 히스토그램을 하나로 병합하여 HOG특징 벡터를 생성한다. 여기서 각 블록에 대한 정규화 방법은 norm, L1-norm, L2-norm, L1Sqrt, L2Hys 그리고 이 밖에도 많은 방법이 있다. 본 연구에서는 L2-norm을 이용하여 정규화를 수행하였으며 이는 아래의 식(5)와 같다. 여기서, v_i 는 각 블록별로 정규화 된 히스토그램을 하나로 병합한 HOG 특징벡터이며, vb 는 block내의 특징벡터, vc 는 cell내부의 특징벡터, n 은 영상 내 block의 수, k 는 cell의 bin 수, i 는 block 내 cell의 수를 나타낸다. Cell 하나당 9개의 bin이 있으므로 특징벡터는 식 (6)과 같으며, block은 2×2 cell로 구성되어 있어 식 (4)와 같이 나타낼 수 있다.

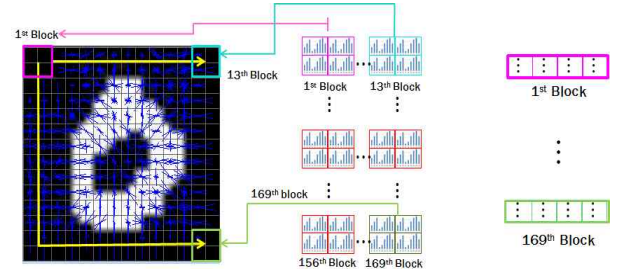
$$H = [\nu_1, \nu_2, \dots, \nu_n], (n: \text{블록의 수}) \quad (4)$$

$$\nu_i = \frac{vb_i}{\sqrt{\|vb_i\|^2 + \epsilon^2}} (\epsilon = 1) (i = 1, 2, \dots, n) \quad (5)$$

$$vb = [vc_1, vc_2, vc_3, vc_4], \quad vc = [v_1, v_2, \dots, v_k], (k = 9) \quad (6)$$

2.2 데이터의 차원축소를 위한 PCA 기법

영상에서 추출된 이미지는 2차원의 영상 이미지이고 고차원의 데이터를 가지고 있기 때문에 학습성능 및 속도를 향상시키기 위해서 대표적인 차원축소 알고리즘인 주성분분석법(PCA)을 사용하



(a) 기울기 크기 및 방향 (b) 기울기 히스토그램 (c) 히스토그램 병합

그림 5 기울기를 이용한 히스토그램 병합
Fig. 5 Merger of histogram using gradient

여 차원을 축소한다. PCA는 고차원의 특징벡터를 저차원의 특징 벡터로 정보의 손실이 최소화 되도록 차원을 축소시키는 대표적인 알고리즘이다. PCA알고리즘을 이용한 숫자인식은 영상의 크기가 $N \times N$ 이고 숫자인식을 위한 이미지의 개수가 M 개일 때 각 숫자영상을 M 개의 $N^2 \times 1$ 행 벡터로 인식영상의 숫자벡터 집합을 구성한다. 인식영상들의 숫자벡터 집합을 사용하여 평균영상을 구할 수 있다. 공분산 행렬을 구하기 위해 학습영상 벡터와 평균 얼굴영상 벡터의 차 영상벡터를 구한다. 이 결과는 $N^2 \times M$ 행렬이 되며 공분산행렬은 $N^2 \times N^2$ 의 차원을 갖는다. 공분산 행렬로부터 M 개의 고유 값 λ_i 와 이에 대응하는 고유벡터를 계산한다. 여기서 고유 값에서 얻어진 고유벡터를 $N \times N$ 로 재배열하면 숫자 모양과 닮은 고유 이미지를 얻을 수 있다. 고유 이미지는 다음과 같은 식으로 표현되며 고유 이미지가 완성되면 숫자인식을 위한 학습단계가 끝나게 된다. 새로운 인식 후보가 들어오게 되었을 때 인식후보들의 평균 벡터를 사용하여 차 벡터를 만들고 그것을 학습 단계에서 만든 고유 이미지에 투영하여 각 인식후보들과의 유클리디언 거리가 가장 작게 나오는 숫자를 인식 대상으로 판정한다. 본 논문에서는 PCA를 숫자인식을 위한 전처리 과정에 사용하였으며, 주성분 분석법의 수행단계는 아래와 같다.

[Step 1] 인식을 위한 숫자 벡터 집합 구성

$$S = \Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M \quad (7)$$

[Step 2] 평균과 분산을 기준으로 이미지 정규화 수행

$$\Gamma_{ij} = (\Gamma_{ij} - \mu_i) \times \frac{\sigma}{\sigma_i} + \mu \quad (8)$$

[Step 3] 평균 숫자 벡터를 계산

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \quad (9)$$

[Step 4] 인식후보 숫자 벡터와 평균 숫자 벡터와의 차이 벡터를 계산

$$\Phi_i = \Gamma_i - \Psi \quad (10)$$

[Step 5] 인식후보 숫자에서 공분산 행렬을 계산

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \cdot \Phi = A \cdot A^T \quad (11)$$

[Step 6] M개의 고유벡터에서 가장 큰 고유값을 갖는 M개만을 선택

$$C = U \Lambda U^T \quad (12)$$

[Step 7] 각 인식후보 숫자와 고유벡터와의 사영을 통해 가중치를 획득

$$\Omega^T = [w_1, w_2, w_3, \dots, w_M] \quad (13)$$

2.2 데이터의 차원축소를 위한 PCA 기법

기존의 PCA방법은 입력받은 2차원의 이미지를 1차원으로 변환하여 특징 추출 및 차원 축소하는 과정을 의미하였다. 그러나 2-Dimensional 2-Directional PCA((2D)²PCA)는 2방향 2차원이라는 뜻으로 입력받은 2차원 데이터를 1차원으로 변환하는 과정 없이 황에서 한번 열에서 한번 특징 추출 및 차원을 축소하는 과정을 의미한다. 이론적으로는 PCA와 같은 알고리즘이나 차원을 축소하는 크기에 있어 큰 차이를 보인다. 예를 들어, 100×100의 이미지를 벡터로 표현하면 10000×1이 된다. 따라서 공분산 행렬의 크기는 10000×10000이 되어 계산량이 증가하고 공분산의 고유벡터를 정확히 구하는데 어려움이 따른다. 하지만 (2D)²PCA는 2차원 이미지 자체를 행렬로 두고 공분산 행렬을 구하기 때문에 행렬의 크기가 행과 열 각각 100×100, 100×100으로 구성되어 데이터 처리에 있어 기존 PCA보다 차원 수가 낮다.

[Step 1] 전체 학습이미지의 데이터 개수는 M개, n번째 학습이미지를 $A_n^{R \times C}$ 행렬($n=1, \dots, M$), 학습이미지의 샘플에 대해 평균 이미지를 $\bar{A}^{R \times C}$ 로 표현할 때 학습이미지의 열에 대한 공분산 행렬 RG_t 는 다음의 식으로 계산

$$RG_t = \frac{1}{M} \sum_{n=1}^M (A_n - \bar{A})^T (A_n - \bar{A}) \quad (14)$$

[Step 2] RG_t 와 같이 행의 공분산 행렬 LG_t 는 다음의 식으로 계산

$$LG_t = \frac{1}{M} \sum_{n=1}^M (A_n - \bar{A})(A_n - \bar{A})^T \quad (15)$$

공분산 행렬은 $RG_t^{R \times R}$ 와 $LG_t^{C \times C}$ 행렬로 구성

[Step 3] 공분산 행렬 $RG_t^{R \times R}$ 의 고유값 $\lambda_R^{R \times R}$ 과 이에 대응하는 고유벡터 $U_R^{R \times R}$ 을 계산

[Step 4] 위의 [Step 3]과 같이 공분산 행렬 $LG_t^{C \times C}$ 의 고유값 $\lambda_L^{C \times C}$ 과 이에 대응하는 고유벡터 $U_L^{C \times C}$ 을 계산

[Step 5] [Step 3]과 [Step 4]에서 나온 고유값 $\lambda_R^{R \times R}$ 과 $\lambda_L^{C \times C}$ 에 대해 고유값이 큰 순서대로 재배열하고, 재배열된 고유값에 대해 고유 벡터도 재배열을 수행

[Step 6] 재배열된 고유벡터를 줄이고 싶은 차원의 수 d만큼 축소하게 되면 $U_{dR}^{R \times d}$ 과 $U_{dL}^{C \times d}$ 로 구성

[Step 7] d만큼 차원을 줄인 벡터 $U_{dR}^{R \times d}$ 와 $U_{dL}^{C \times d}$ 을 가지고 실제 전체 이미지를 획득하기 위해 아래의 식을 수행

$$\tilde{A}_j = \sum_{i=1}^M U'_{dR} \times A_j \times U_{dL}, \tilde{A}^{d \times d} = [\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_M] \quad (16)$$

3. 인식을 위한 pRBFNNs 패턴분류기의 설계

3.1 pRBFNNs 패턴 분류기의 구조

본 장에서는 제안된 얼굴인식 시스템의 인식 성능을 확인하기 위해 다항식 기반 RBFNNs 패턴분류기[1][2]에 대하여 설명한다. 기본적인 신경회로망은 인간의 두뇌를 구현하기 위한 알고리즘으로, RBFNNs 패턴분류기는 신경회로망의 구조를 기반으로 하여 입력층, 은닉층, 출력층이 3가지로 구분할 수 있다. 입력층에서는 각 노드마다 입력변수의 개수만큼 활성화함수가 존재하며, 활성화함수는 방사형 기저함수의 형태를 하고 있고, 주로 가우시안 형태를 사용하고 있다. 입력 데이터는 전처리 과정을 수행하여 인식에 사용하기 적합한 저차원 데이터로 축소되어, 은닉층의 입력으로 들어가게 된다. 은닉층에서 활성화함수로 변환된 출력은 은닉층과 출력층 사이의 연결가중치 값과 곱해서 출력층의 최종 출력으로 얻어지며, 이때 연결가중치의 값은 상수를 사용한다. 일반적인 RBFNNs의 은닉층 활성화함수는 가우시안 함수를 사용한다. 따라서 가우시안 함수의 중심점 및 분포상수의 값에 의해 모델의

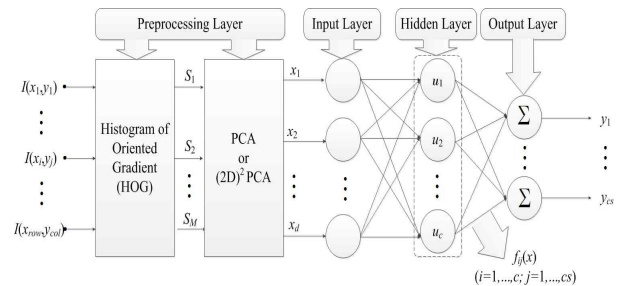


그림 6 다항식기반 RBFNNs 패턴 분류기 구조

Fig. 6 Overall architecture of polynomial-based RBFNNs pattern classifier

성능 차이가 심하다. 또한 연결가중치의 학습은 최소자승법을 사용하며 2차원 이미지와 같은 고차원 입력변수를 전처리없이 사용하면 학습량 증대 및 적절한 연결가중치를 추정할 수 없다.

본 연구에서는 hybrid RBFNNs을 제안한다. 전처리 과정으로써 두 단계를 RBFNNs 입력층 앞에 추가하였다. 첫 번째는 Histogram of Oriented Gradient(HOG)를 사용하여 0~255사이의 이미지 값을 기울기 값($0^\circ \sim 180^\circ$)으로 변환하였으며, 기울기 값으로 변환된 고차원 입력변수를 주성분 분석법 또는 2차원 2방향 주성분 분석법을 통해 차원을 축소하여 RBFNNs의 입력층으로 연결된다. RBFNNs의 은닉층은 기존의 가우시안 함수 대신에 은닉층 자체를 Fuzzy C-Means(FCM) 클러스터링으로 대체하였다. FCM 클러스터링의 소속행렬 분포는 가우시안 함수와 유사하여 클러스터의 중심점과 데이터의 소속정도를 표현하는 소속 값을 은닉층 각 노드의 출력으로 사용하였다. 따라서 FCM의 각 클러스터가 은닉층의 노드가 된다. 상수항에서 일차식과 이차식, 그리고 변형된 이차식과 같은 다항식의 형태로 확장하여 사용한다는 특징을 가지고 있다.

[Type 1] 1차 선형 추론 (Linear)

$$f_j(x_1, \dots, x_d) = a_{j0} + \sum_{i=1}^d a_{ji} x_i \quad (17)$$

[Type 2] 2차 선형 추론 (Quadratic)

$$f_j(x_1, \dots, x_d) = a_{j0} + \sum_{i=1}^d a_{ji} x_i + \sum_{i=1}^k a_{j(d+1)} x_i^2 + a_{(2d+1)} x_1 x_2 + \dots + a_{d(d+3)/2} x_{(d-1)} x_d \quad (18)$$

[Type 3] 변형된 2차 선형 추론 (Modified Quadratic)

$$f_j(x_1, \dots, x_d) = a_{j0} + \sum_{i=1}^d a_{ji} x_i + a_{(2d+1)} x_1 x_2 + \dots + a_{d(d+3)/2} x_{(d-1)} x_d \quad (19)$$

여기서 $x = [x_1, x_2, \dots, x_d]$, d 는 입력변수의 수, R^j 는 j 번째 퍼지 규칙이며, $f_j = (x_1, \dots, x_d)$ 는 j 번째 규칙에 대한 후반부로서 j 번째 퍼지 규칙에 대한 로컬 모델이다.

모델의 최종출력은 아래 식처럼 구해진다.

$$\hat{y} = \sum_{j=1}^c u_j f_j(x_1, \dots, x_d) \quad (20)$$

3.2 FCM(Fuzzy C-Means) 클러스터링

FCM(Fuzzy C-Means) 클러스터링 알고리즘은 비슷한 패턴, 속성, 형태 등의 기준을 통해 데이터를 분류하는 알고리즘으로, 데이터와 각 클러스터와의 거리를 기준으로 소속정도를 측정하여 데이터를 분류한다. 이를 이용하여 다항식 기반 RBFNNs 패턴분

류기의 은닉층 활성화함수 형태를 표현하였으며, 아래 단계를 통해 수행된다.

[Step 1] 클러스터의 개수, 퍼지화 계수를 선택하고 소속함수 ($U^{(0)}$)를 초기화 한다.

$$U^{(0)} = \left\{ u_{ik} \in [0,1], \sum_{i=1}^c u_{ik} = 1 \forall k, 0 < \sum_{k=1}^n u_{ik} < n \forall i \right\} \quad (21)$$

[Step 2] 각 클러스터에 대한 중심 벡터를 구한다.

$$v_{ij} = \frac{\sum_{k=1}^n (u_{ik})^m x_{kj}}{\sum_{k=1}^n (u_{ik})^m} \quad (22)$$

[Step 3] 중심과 데이터와의 거리를 계산하며, 이를 통해 새로운 소속함수($U^{(1)}$)를 계산한다.

$$d_{ik} = d(x_k - v_i^{(r)}) = \left[\sum_{j=1}^l (x_{kj} - v_{ij}^{(r)})^2 \right]^{1/2} \quad (23)$$

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}} \quad (24)$$

[Step 4] 오차가 허용범위 안에 도달하면 알고리즘을 종료하고, 그렇지 않으면 [Step 2]로 돌아간다.

$$\| U^{(r+1)} - U^{(r)} \| \leq \epsilon \quad (25)$$

3.3 최소자승법

후반부 다항식의 파라미터 계수는 최소자승법을 사용하여 모델을 설계한다. LSE는 오차제곱의 합이 최소가 되도록 계수를 추정하며 후반부 다항식들의 계수를 한 번에 구하기 때문에 전역 모델의 학습에 수행 가능하다. LSE에서의 성능평가함수 Q_L 을 행렬 형식으로 표현하면 식 (26)과 같다.

$$Q_L = (Y - XA)^T (Y - XA) \quad (26)$$

여기서 A 는 후반부 다항식 파라미터 계수, Y 는 출력 데이터, X 는 FCM의 소속행렬이 곱해진 입력데이터의 행렬로 x 는 입력행렬, u 는 소속행렬을 의미한다. 후반부 구조가 Linear일 경우 X, Y, A 행렬은 다음과 같은 식 (27)로 표현 되며, 여기서 m 은 데이터의 수이며, k 는 입력수의 수, c 는 규칙 수(클러스터 수)가 된다.

$$A = [a_{10}, a_{11}, \dots, a_{1d}, \dots, a_{c0}, a_{c1}, \dots, a_{cd}]^T, \quad Y = [y_1, y_2, \dots, y_n]^T,$$

$$X = \begin{bmatrix} u_{11} \cdots u_{c1} & x_{11}u_{11} \cdots x_{11}u_{c1} & \cdots & x_{d1}u_{11} \cdots x_{d1}u_{c1} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} \cdots u_{cn} & x_{1n}u_{1n} \cdots x_{1n}u_{cn} & \cdots & x_{dn}u_{1n} \cdots x_{dn}u_{cn} \end{bmatrix} \quad (27)$$

$$A = (X^T X)^{-1} X^T Y \quad (28)$$

4. 실험 결과

4.1 실험의 전체 개요

제안된 숫자 인식 시스템의 과정은 아래의 그림 7과 같이 학습과 테스트를 위한 두 개의 파트로 나뉘어져 있다. MNIST 숫자 데이터를 위한 학습 과정에서는 구성된 MNIST 학습 데이터를 HOG 알고리즘을 통하여 특징 벡터를 추출하게 된다. HOG 특징은 조명변화와 이미지의 색상과 형태에 대한 영향이 적기 때문에 널리 이용되고 있다. 하지만 HOG 특징은 입력 영상에서 Cell과 Block의 크기에 따라 특징 차원 수는 기하급수적으로 증가하기 때문에 연산이 많고 불필요한 특징 차원에 대해서도 연산하기 때문에 처리속도가 느리다는 단점이 있다. 이는 차원이 증가함에 따라 생기는 차원의 저주와 같은 맥락이며 이러한 문제 해결을 위해 제안하는 분류기의 학습을 수행하기 이전에 주성분 분석법(PCA) 및 (2D)²PCA를 이용하여 추출된 고차원의 특징을 저차원으로 축소한다. 여기서 PCA와 (2D)²PCA를 이용하는 가장 큰 목적은 추출된 특징 데이터가 차원축소 이후에 이전의 실제 데이터 정보를 손실 없이 최대한 유지하도록 하는 것으로서 적용이 된다. 그리고 차원 축소된 입력 영상은 제안하는 패턴분류기의 입력으로 사용이 되어 학습을 수행하게 된다.

패턴분류기를 통한 학습이 되면 구성된 테스트 데이터를 아래 그림 7에서와 같은 학습 방법과 동일하게 Step (a)부터 Step (c)까지 수행을 한다. 마지막으로 차원 축소된 테스트 데이터의 특징벡터를 구축한 모델을 통해 분류를 수행하여 인식성능을 확인하게 된다.

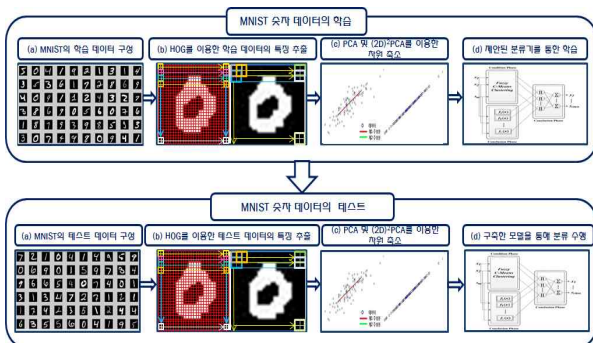


그림 7 MNIST 데이터를 이용한 숫자인식 시스템 수행과정
 Fig. 7 Performance procedure of digits recognition system using MNIST digits database

4.2 MNIST 실험 데이터의 구성 및 초기 파라미터의 설정

본 논문에서는 HOG 특징을 이용한 다항식 방사형 기저함수 신경회로망 기반 숫자 인식 방법을 설계를 위해 인식성능 비교 연구를 하였다. 성능을 평가하기 위해 사용된 숫자 이미지는 필기체로 구성된 MNIST 숫자 데이터베이스를 사용하였으며, 0부터 9까지의 숫자들을 모아 놓은 데이터의 집합으로 구성 되어 있다. 하나의 이미지는 가로 28, 세로 28의 픽셀로 이루어져 있으며 데이터의 수는 총 70000개로 각각 학습 데이터 60000개와 테스트 데이터 10000개로 구성되어 있다. 이는 아래의 표 1로도 실험데이터의 구성을 확인할 수 있다.

표 1 실험 데이터의 구성

Table 1 Configuration of experimental dataset

Data	MNIST
Image size	28×28
No. of inputs	784
No. of training data	60000
No. of testing data	10000
No. of classes	10

제안된 패턴 분류기의 성능을 평가하기 위한 평가 지수는 식 (29)과 같이 오분류율(Misclassification Rate)을 이용한다. 또한, 객관적인 평가를 위해 기존에 제안되었던 Neural Networks, Linear Classifier, Quadratic Classifier, Convolutional net 등 다른 패턴 분류기들과 성능을 비교 분석한다.

$$\text{오분류율} = \frac{\sum_{i=1}^N f(y_i - \hat{y}_i)}{N} \times 100, \quad f(y_i - \hat{y}_i) = \begin{cases} 1, & y_i \neq \hat{y}_i \\ 0, & y_i = \hat{y}_i \end{cases} \quad (29)$$

특징 추출을 위해 사용된 HOG알고리즘에서 Cell 과 Block의 크기는 각각 2×2로 지정하였으며 총 70000개의 이미지에서 학습 데이터로 60000개, 테스트 데이터로 10000개씩 각각 사용된다. 클래스의 개수는 0부터 9까지 총 10개의 클래스로 이루어진다. 조건부에서 사용된 Fuzzy C-Means 클러스터링 알고리즘의 퍼지화 계수는 일반적으로 사용되는 2.0으로 정의하였으며 결론부의 다항식 형태는 1차선형식(Linear)으로 구성하였다. 표 2는 제안된 패턴분류기의 파라미터를 나타낸 표이다.

표 2 초기 파라미터의 설정

Table 2 Setting of initial parameters

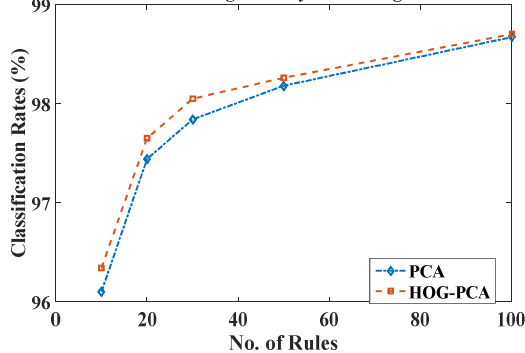
Parameters	Values
Size of cell	2×2
Size of block	2×2
Fuzzification coefficient	2.0
Polynomial Type	Linear

4.3 실험 결과

본 논문에서는 HOG 특징을 이용한 다항식 방사형 기저함수 신경회로망 기반 숫자 인식 방법을 설계하기 위한 인식 성능 비교 연구를 진행하였다. 그림 8에서는 PCA, (2D)2PCA, HOG-PCA 그리고 HOG-(2D)2PCA를 이용한 전처리과정을 수행함에 있어서 방사형 기저함수 신경회로망에서의 규칙 수를 증가함에 따라 어떤 성능변화를 나타내는지 보이기 위한 그래프이다. 먼저 아래의 그림 8(a)는 PCA와 HOG-PCA 숫자 인식 시스템의 인식 성능을 나타낸다. 전처리과정으로 PCA만을 사용하여 차원축소를 하였을 경우보다 HOG-PCA를 사용하였을 때 조금 더 우수한 성능을 보이는 것을 확인하였다. 제안된 인식시스템의 노드 수 변화에 따른 인식 성능은 아래의 그림8(b)에서 보이는 바와 같이 노드의 수가 100개 까지 증가하여도 분류율이 좋아지는 것을 확인할 수 있다. 규칙의 수도 전자와 같이 10-100까지 증가하였을 때 어떠한 성능을 나타내는지 확인하였다. (2D)²PCA를 이용하여 차원 축소하였을 때보다 HOG-(2D)²PCA를 이용하여 전처리과정을 수행하였을 때 보다 좋은 성능을 나타내었으며 수치상으로는 PCA와 HOG-PCA를 사용하였을 때보다 조금 더 우수한 성능을 보인다.

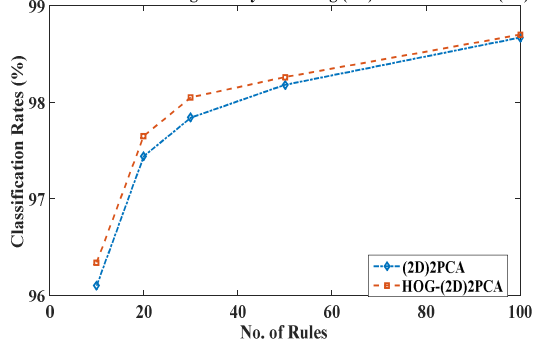
마지막으로 그림 9와 표 3은 MNIST 숫자 데이터의 성능평가

Classification results of recognition system Using PCA and HOG-PCA



(a) PCA와 HOG-PCA 숫자 인식 시스템의 성능 비교

Classification results of recognition system Using (2D)2PCA and HOG-(2D)2PCA



(b) (2D)²PCA와 HOG-(2D)²PCA 숫자 인식 시스템의 성능 비교

그림 8 제안된 숫자 인식 시스템의 분류율 비교

Fig. 8 Comparison of classification rates for the proposed recognition system

를 위해 본 논문에서 제안한 패턴 분류기와 기존 논문에서 나타난 인식 성능을 비교하기 위한 그래프와 표이다. 그림 9에서 보이는 바와 같이 제안된 분류기의 성능이 대부분의 다른 분류기보다 우수한 성능을 출력하는 것을 확인할 수 있다. 특히 PCA와 (2D)²PCA기반 RBFNNs 패턴 분류기는 기존의 PCA기반 Quadratic classifier 보다 더 우수한 성능을 나타냄을 확인할 수 있었으며, 이를 개선한 HOG-PCA와 HOG-(2D)²PCA의 성능은 이보다 조금 더 우수한 성능을 나타낸다. 그러나 Convolutional net의 경우에는 본 논문에서 사용한 패턴분류기에 비해 인식 성능보다 우수하지만, Convolutional net은 상대적으로 RBFNN 패턴분류기에 비해 구조가 복잡하고 계산량이 많다는 단점과 추가적으로 고성능의 하드웨어(GPU)를 요구하나 본 논문에서 사용된

표 3 MNIST 숫자 데이터를 이용한 인식 시스템의 성능 평가

Table 3 Performance evaluation of recognition system using MNIST digits database

Classifier	Preprocessing	MR(%)
Linear Classifier(1-layer NN)[10]	Deskewing	8.4
K-nearest neighbors, Euclidean[10]	Deskewing	2.4
40 PCA + Quadratic classifier[10]	PCA	3.3
Convolutional net LeNet-1[10]	16 × 16 pixels	1.7
Convolutional net LeNet-4[10]	None	1.1
2-Layer MLP(CE)[11]	None	1.6
Support Vector Machine(SVM)[12]	Affine	1.4
PCA+100 RBFNNs[13]	PCA	1.33
Proposed recognition system		
(No. of dimensional reduction : 64, No. of rules : 100)		
(2D) ² PCA + 100 RBFNNs	(2D) ² PCA	1.28
HOG_PCA + 100 RBFNNs	HOG_PCA	1.30
HOG_(2D) ² PCA + 100 RBFNNs	HOG_(2D) ² PCA	1.26

MR: Missclassification rate

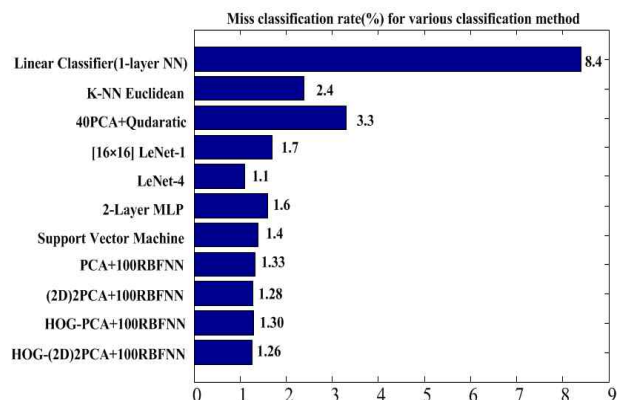


그림 9 다양한 분류기 방법에 대한 인식 성능의 비교

Fig. 9 Comparison of missclassification rates for various classification methods

분류기는 구조가 더 간단하고 계산량이 적다는 장점이 있다. 마지막으로 RBFNNs 패턴 분류기에 PCA를 적용하였을 때보다 (2D)²PCA를 사용하였을 때 개선된 성능을 출력함을 확인하였다. 또한 이러한 숫자 인식시스템의 전처리과정에 HOG 알고리즘의 결합하였을 때 단독적으로 PCA 및 (2D)²PCA를 사용한 경우보다 조금 더 우수한 성능을 출력함을 확인하였다.

5. 결 론

본 논문에서는 HOG 특징을 이용한 다항식 방사형 기저함수 신경회로망 기반 숫자 인식 방법을 설계하고 전처리 알고리즘을 이용하여 인식성능의 비교연구를 제안하였다. RBFNNs 패턴 분류기에서 HOG 알고리즘을 사용하여 MNIST 숫자 이미지의 특징을 추출하였으며, 이를 통하여 추출된 숫자이미지의 특징은 높은 차원의 입력을 가진다. 따라서 고차원의 입력데이터를 사용하고 연산량을 줄이기 위해 PCA 및 (2D)²PCA알고리즘을 이용하여 특징벡터를 저차원으로 축소하고 정보의 손실이 최소화되도록 하였다. 조건부 활성화함수에 가우시안 함수 대신 Fuzzy C-Means 알고리즘을 사용하여 입력 데이터간의 특성을 반영한 패턴 분류기를 구성하였다. 본 논문에서 제안한 숫자 인식시스템과 기존에 제안되었던 분류기들의 성능을 비교하였을 때, HOG-PCA 및 HOG-(2D)²PCA알고리즘을 통해 입력 차원을 효율적으로 축소하고 비교적 간단한 구조의 패턴 분류기를 사용하여 다른 인식시스템보다 상대적으로 우수한 성능을 출력하는 것을 확인할 수 있었다. 향후 연구에서는 제안된 패턴 분류기의 성능을 개선하는 방법으로 HOG 및 PCA알고리즘 이외의 다른 전처리 알고리즘을 사용하여 숫자 이미지의 정보를 보다 효율적으로 사용할 수 있는 방법에 대해 연구하며 제안된 분류기를 딥 구조로 확장 및 최적화 방법에 대해 연구하고자 한다.

감사의 글

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning [grant number: NRF-2015R1A2A1A15055365] and also supported by the GRRC program of Gyeonggi province [grant number: GRRC Suwon 2016-B2, Centre for U-city Security & Surveillance Technology].

References

- [1] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 886-893, 2005.
- [2] H. Addi and L. J. Williams, "Principal Component Analysis," Wiley Interdisciplinary Reviews: Computational Statistics, Vol. 2, no. 4, pp. 433-459, 2010.
- [3] Y. Ke, and R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors", IEEE Computer Society Conference on Computer Vision and Pattern Recognition Vol. 2, pp. 506-513. 2004.
- [4] J.-Y. Kim, Y.-S. Kim " Face Tracking and Recognition in Video with PCA-based Pose-Classification and (2D)²PCA recognition algorithm." Journal of Korean Institute of Intelligent Systems, Vol. 23, No. 5, pp. 423-430, 2013.
- [5] S.-K. Oh, W. Pedrycz, B.-J. Park, "Polynomial-based Radial Basis Function Neural Networks realized with the Aid of Particle Swarm Optimization," Fuzzy Sets and Systems, Vol. 163, pp. 54-77, 2011.
- [6] S.-B. Roh, S.-C. Joo, W. Pedrycz, and S.-K. Oh, "The Development of Fuzzy Radial Basis Function Neural Networks Based on the Concept of Information Ambiguity," Neurocomputing, Vol. 73, No. 13-15, pp. 2464-2477. 2010.
- [7] W. Pedrycz, "Conditional Fuzzy Clustering in the Design of Radial Basis Function Neural Networks", IEEE Trans. Neural Networks, Vol. 9, pp. 601-612, July 1998.
- [8] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
- [9] S. P. Lloyd, "Least Squares Quantization in PCM," IEEE Trans. on Information Theory, Vol. 28, No. 2, pp. 129-137, 1992.
- [10] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner "Gradient-Based Learning Applied to Document Recognition", IEEE, Vol. 86, pp. 2278-2324, 1998.
- [11] P. Y. Simard, D. Steinkraus, and J. C Platt, "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis," Seventh International Conference on Document Analysis and Recognition, Vol. 1 pp.958-963, 2003.
- [12] D. Decoste and B. Scholkopf, ""Training Invariant Support Vector Machines"", Machine Learning Journal, vol 46, No 1-3, 2002.
- [13] S.-C. Lee, S.-K. Oh and H.-K. Kim "Design of PCA-based pRBFNNs Pattern Classifier for Digit Recognition" Journal of Korean Institute of Intelligent Systems, Vol. 25, No. 4, pp. 355-360, 2015.
- [14] S. Knerr and L. Personnaz and G. Dreyfus, "Handwritten Digit Recognition by Neural networks with single-layer training," IEEE Trans. Neural Networks, Vol. 3, No. 6, pp. 962-968, 1992.
- [15] S. W. Lee, "Off-line recognition of totally unconstrained handwritten numerals using multi layer cluster neural network," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No. 6, pp. 648-652, 1996.

저 자 소 개



김 은 후 (Eun-Hu Kim)

2009년 수원대학교 전기공학과 공학사, 동 대학원 석사(2011), 박사(2016), 2016~현재 수원대학교 전기공학과 Post-Doc. 관심분야는 퍼지 이론, 신경 회로망, 클러스터링, 최적화 알고리즘, 컴퓨터 지능 및 지능 제어 등.

Tel: 031-222-6544

Fax: 031-220-2667

E-mail: wdkim@suwon.ac.kr



김 봉 연 (Bong-Youn Kim)

2014년 수원대학교 전기공학과 공학사, 동 대학원 석사(2016).

관심분야는 퍼지추론시스템, 신경회로망, 패턴인식, 얼굴인식, 진화 최적화 등.

Tel: 031-222-6544

Fax: 031-220-2667

E-mail: zillot07@naver.com



오 성 권 (Sung-Kwun Oh)

1981년 연세대학교 전기공학과 공학사, 동 대학원 석사(1983), 박사(1993), 1983~1989년 금성산전연구소(선임연구원). 1996~1997년 캐나다 Manitoba 대학 전기 및 컴퓨터 공학과 Post-Doc. 1993~2004년 원광대학교 전기전자 및 정보공학부 교수, 2005~현재

수원대학교 전기공학과 교수. 2002~현재 대한전기학회, 퍼지및 지능시스템학회 편집위원, 2013년~현재 Information Sciences 편집위원. 관심분야는 퍼지 시스템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 computational intelligence, 지능 제어 등.

Tel: 031-229-8162

Fax: 031-220-2667

E-mail: ohsk@suwon.ac.kr