

---

# 망막 세포 특성에 의한 영상인식에 관한 연구

조재현\* · 김도현\*\* · 김광백\*\*\*

## A Study on Image Recognition based on the Characteristics of Retinal Cells

Jae-Hyun Cho\* · Do-Hyeon Kim\*\* · Kwang-Baek Kim\*\*\*

### 요 약

최근 시각 장애인을 위한 인공망막 모델 구현에 관한 연구 중 시피질 자극기 기술은 시각 자극 전달의 중간 단계를 생략하고 직접 뇌세포를 자극하는 것이다. 본 논문에서는 망막에서 시각 피질로 시각정보를 전달할 때 발생하는 시각 피질의 특성, 즉 방향성에 대한 반응 특성을 특징 데이터로 구성하여 인식함으로써 인간 시각 정보 처리와 유사한 영상 추출 및 인식 모델을 제안한다. 제안된 방법은 영상의 특징을 추출 한 후  $\Delta$  기반 오프 역전파 알고리즘을 적용하여 영상의 특징들을 인식한다. 제시된 방법의 성능을 분석하기 위하여 다양한 숫자 패턴들을 대상으로 실험한 결과, 제안된 망막 세포로부터 전달된 정보를 방향성에 대한 민감성을 고려하여 영상의 특징을 추출하여 인식하는 모델이 기존의 영상 추출 및 인식 모델보다 인식률에 있어서는 별 차이가 없지만 다양한 실험에서 확인할 수 있듯이 인간 시각과 같이 인식 성능이 민감하지 않는 것을 알 수 있었다.

### ABSTRACT

Visual Cortex Stimulator is among artificial retina prosthesis for blind man, is the method that stimulate the brain cell directly without processing the information from retina to visual cortex. In this paper, we propose image construction and recognition model that is similar to human visual processing by recognizing the feature data with orientation information, that is, the characteristics of visual cortex. Back propagation algorithm based on  $\Delta$  is used to recognize after extracting image feature by Kirsh edge detector. Various numerical patterns are used to analyze the performance of proposed method. In experiment, the proposed recognition model to extract image characteristics with the orientation of information from retinal cells to visual cortex makes a little difference in a recognition rate but shows that it is not sensitive in a variety of learning rates similar to human vision system.

### 키워드

Retina, Visual cortex, Neural network, Orientation,  $\Delta$

### I. 서 론

인공망막의 기술은 크게 4가지로 분류된다. 망막하 자극기(Sub-retinal Stimulator), 망막앞자극기(Epi-retinal

Stimulator), 시신경자극기(Optic Nerve Stimulator), 그리고 시피질자극기(Visual Cortex Stimulator)등이다. 망막하자극기는 망막 중 양극세포를 자극하여 망막의 신경망 처리과정을 활용함으로써 화상처리부가 생략되어

---

\* 부산가톨릭대학교 컴퓨터공학과  
\*\* 부산대학교 컴퓨터공학과  
\*\*\* 신라대학교 컴퓨터정보공학부

구성이 간단하지만 자극에 필요한 전류생성의 문제가 있다. 망막앞 자극기는 미국 MEEI에서 시도하는 방법으로 신경세포 자극기를 망막위에 붙이는 방법으로 전극의 삽입이 용이하지만 망막에 손상을 줄 우려가 있다. 그리고 시피질 자극기는 1968년 Brindley와 Lewin에 의해 첫 시도되었으며 망막손상과 시신경 손상환자에게 적용할 수 있도록 시피질을 자극한다. 이 방법은 시각 자극 전달의 중간단계를 생략하고 직접 뇌세포를 자극하는 것이다[1,2]. 본 논문에서는 시피질 자극 방법에 기반하여 기존의 시피질 자극 방법에서 배제되고 있는 중간단계를 망막과 시각피질의 특성을 고려하여 최적의 영상으로 표현하여 영상인식에 적용하고자한다. 인간의 시각 정보 처리과정은 여러 단계로 구성되어져 있으며, 그 일차적인 정보처리는 인간의 망막에서 이루어진다. 인간의 망막은 빛에너지를 전기 화학적 에너지로 변환하는 과정뿐만 아니라, 그 정보를 시각 경로(visual path)로 전달하는 역할을 하고 있다[3].

## II. 인간 시각의 정보처리 과정

### 2.1. 망막에서의 정보처리

인간의 시각 정보처리 흐름은 영상이 망막에 입력되어 광수용기에 상이 맺히게 된다. 망막으로부터 출력되는 전기화학적 신호를 신경절세포로 구성된 시신경(optic nerve)에 의해 초기시각피질로 이동된 이후 인지과정을 거치게 된다. 눈동자의 뒤에 위치한 망막은 그림 1과 같이 여러 개의 수직 층(vertical layer)으로 복잡하게 구성되어 있다.

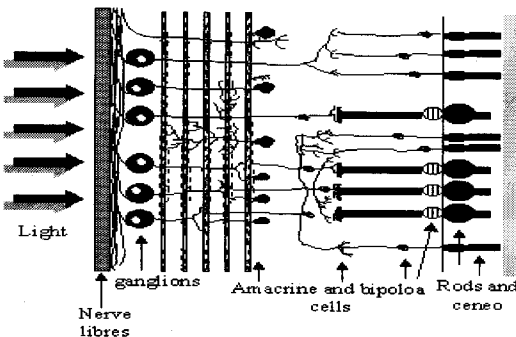


그림 1. 망막의 5 주요 세포들의 구성

Fig. 1 Cross section of the retina showing the five major cell types

그러나 신경절 세포는 한 개의 눈에 고작 100만 개뿐이므로 1억 2천 5백만 개의 수용기로부터 신경절 세포에 이르기까지 상당한 양의 시각정보의 압축이 일어난다. 실제적으로 약 1억 2천만 개의 간상세포와 6백만 개의 원추세포에서 샘플링 된 영상들이 망막의 기전에 의해 아날로그 형태의 전기적 신호로 변환되는데, 이 신호들은 망막의 시신경을 통하여 뇌의 대뇌피질의 시각영역으로 전달되기 위해 신경절 세포에서 펄스형태의 신호로 변환된다. 이런 신경절세포수가 약 100만개이므로, 약 130:1의 정보압축효과를 나타내면서 데이터압축을 가능케 한다. 즉, 인간은 100만개의신경절 세포에서 초기시각피질까지의 시각정보 전송으로도 물체를 인식하는데 별다른 문제를 보이지 않고 있다[5,6].

### 2.2. 시각피질에서의 정보처리

망막을 빠져 나가는 대부분의 영상정보는 외측생핵(LGN)의 P\_cell 과 M\_cell을 통해 시각피질로 전달된다. 그림 2는 시각피질의 정보 처리 특성을 나타내었으며 크게 단순피질, 복합피질, 그리고 끝맺음 피질로 구성되어 있다.

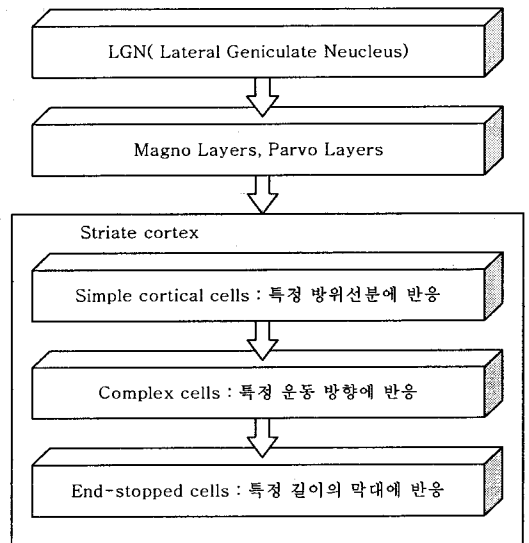


그림 2. 시각피질 특성을 고려한 영상인식

Fig. 2 Proposed image recognition based on visual cortex

Hubel과 Wiesel은 세 가지 주요유형의 뉴런들을 확인하고 가장 잘 반응하는 자극의 유형에 따라 분류를 하였

다. 단순 피질 세포는 중심-주변 수용장처럼 흥분성과 억제성을 갖는 수용장을 갖고 있지만 이들 영역들은 중심 주변의 모양보다는 측면을 따라 나란히 배열되어 있다.[5,8] 그림 3은 단순 피질세포의 수용장을 나타낸다.

단순피질세포의 특성은 특정한 방향으로 놓여 있는 빛의 막대에 가장 잘 반응하도록 되어 있다는 것이다. 즉 수용장의 길이를 따라 놓여있을 때 큰 반응이 발생한다는 것이다[5,8].

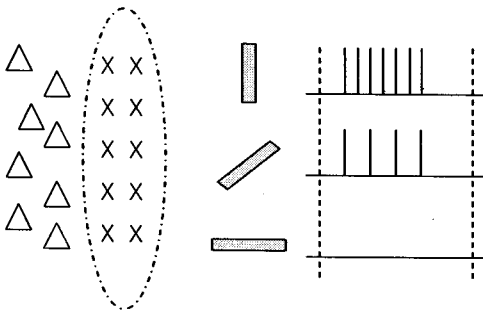


그림 3. 단순 피질 세포의 수용장  
Fig. 3 The receptive field of a simple cortical cell

### III. 시각피질의 특성을 고려한 영상인식

현재의 컴퓨터시각 이론은 획득한 영상을 그림 4와 같이 영상의 특징을 추출하여 물체를 인식하는 과정을 거치게 된다. 즉 인간 시각의 특징정보 추출과정을 모방하여 물체를 인식하고는 있으나 실제 그 응용 과정 중에 인간의 시각정보처리과정과는 상이한 형태로 발전되어지고 있다.

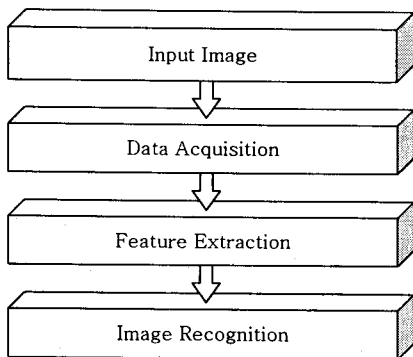


그림 4. 일반적인 영상 인식 모델  
Fig. 4. General image recognition model

최근 인간의 시각정보처리 과정에 관련한 연구가 진행되고 있으며, 점차 인간의 시각정보처리과정이 규명되고 있다. 현재 세계 각국 및 국내에서도 이러한 모델의 연구 시도가 활발히 일어나고 있다[1-4].

컴퓨터 시각은 앞장에서 살펴본 바와 같이 인간 시각정보처리과정의 개략적인 형태를 모방하고 있을 뿐 그 정보처리과정은 인간의 시각정보처리과정과는 상이한 형태를 보이고 있다.

본 논문에서는 시각피질의 특성인 방향성에 대한 반응 정도를 고려한 영상인식 방법을 그림 5에 나타내었다.

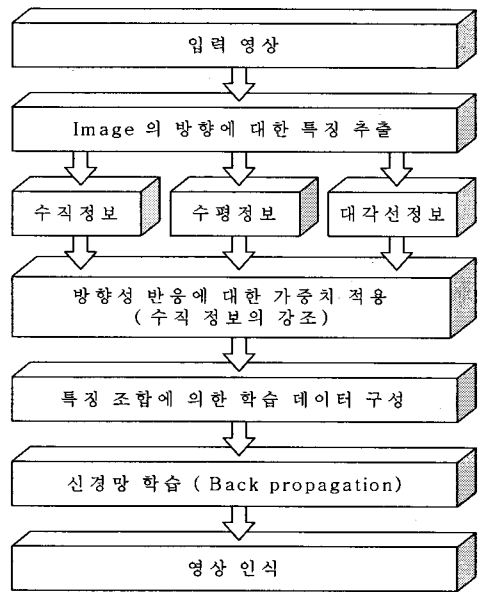


그림 5. 시각피질 특성을 고려한 영상인식  
Fig. 5 Proposed image recognition based on visual cortex

인간의 망막에서 전해지는 정보는 시각피질의 특성에 의해 수직방향에 매우 민감하게 반응한다. 따라서 제안된 방법은 먼저 4방향 경계선 성분을 추출하기 위하여 Kirsh edge detector를 사용하였으며[4,10], 추출된 정보에 대하여 가중치를 적용하여 학습데이터를 구성하였다. 수식은 다음과 같다.

$$K(i, j) = \max \{ 1, \max_{0 \leq k \leq 7} [|5S_k - 3T_k|] \} \quad (1)$$

여기서,  $S_k = A_k + A_{k+1}A_{k+2}$ ,

$$T_k = A_{k+3} + A_{k+4} + A_{k+5} + A_{k+6} + A_{k+7}$$

$A_0$	$A_1$	$A_2$
$A_7$	$(i, j)$	$A_3$
$A_6$	$A_5$	$A_4$

그림 6. 픽셀  $(i, j)$  의 8 이웃 픽셀  
Fig. 6 8 neighborhood pixels of pixel  $(i, j)$

그림 6은  $(i, j)$ 에서의 8 이웃 픽셀의 첨자를 정의한 것으로  $K(i, j)$ 는 픽셀  $(i, j)$ 에서의 그라디언트를 계산한 값이고,  $A$ 의 첨자는 modulo 8 연산 결과 값이다. 그리고 추출된 정보 중 수직정보를  $32 \times 32$ , 수평 정보를  $16 \times 32$ , 대각선 정보를  $16 \times 32$ 로 샘플링 하여 수직정보를 강조하였다. 여기서 대각선 정보는 역 대각선정보를 포함하였다. 인식기는 다양한 응용 분야에서 널리 사용되는 오류 역전과 알고리즘(BP)을 개선한 Delta-bar-delta 알고리즘을 사용하여 학습함으로써 생성한다. 오류 역전과 알고리즘(BP)은 순방향 다층 신경망의 학습에 널리 사용되는 대표적인 알고리즘이다. BP 알고리즘은 출력층 오차 신호를 이용하여 은닉층과 출력층간의 연결 강도를 변경하고, 이 오차 신호를 다시 은닉층에 역전파해서 은닉층과 입력층 사이의 연결 강도를 변경시키는 학습 방법을 사용한다.

Delta-bar-delta ( $\Delta - \bar{\Delta}$ ) 학습 알고리즘은 각 학습 단계마다 연결강도의 변화에 따라서 학습률을 적응적으로 변경함으로써 학습 단계를 단축하고 궁극적으로 학습 시간을 효과적으로 감소시키는 방법이다[7][9]. BP 알고리즘을 이용한 다층 신경망의 학습 절차는 다음과 같다[7].

**Step 1.** 입력층 패턴벡터  $x$ , 은닉층 출력벡터  $z$ , 출력층 출력벡터  $y$ 를 다음과 같이 표현한다.

$$\begin{aligned} x &= [x_1, x_2, \dots, x_n] \\ z &= [z_1, z_2, \dots, z_p] \\ y &= [y_1, y_2, \dots, y_m] \end{aligned} \quad (2)$$

입력층과 은닉층간의 연결강도  $v(p \times n)$ , 은닉층과 출력층간의 연결강도  $w(m \times p)$ 를 임의의 작은 값으로 초기화하고 임의의 학습률  $\alpha$ 를 설정한다.

**Step 2.** 학습 패턴 쌍을 차례로 입력하여 다음과 같이 은닉층의 가중합  $NET_z$  및 출력  $z$ , 출력층의 가중합  $NET_y$  및 최종 출력  $y$ 를 구한다.

$$NET_z = xv^T \quad (3)$$

$$z = f(NET_z) = \frac{1}{1 + \exp(-NET_z)} \quad (4)$$

$$NET_y = zw^T \quad (5)$$

$$y = f(NET_y) = \frac{1}{1 + \exp(-NET_y)} \quad (6)$$

**Step 3.** 목표치  $d$ 와 최종 출력  $y$ 를 비교하여 오차를 구하고 출력층의 오차 신호  $\delta_y$ 와 은닉층에 전파되는 오차 신호  $\delta_z$ 를 구한다.

$$E = \frac{1}{2} (d - y)^2 \quad (7)$$

$$\delta_y = (d - y)y(1 - y) \quad (8)$$

$$\delta_z = z(1 - z) \sum_{j=1}^m \delta_{y_j} w_{zj} \quad (9)$$

**Step 4.** 식 (9),(10)에 의해  $k$  학습 단계에서의 은닉층과 출력층간의 연결강도 변화량  $\Delta w^k$  및 입력층과 은닉층간의 연결강도 변화량  $\Delta v^k$ 를 구하고 각각의 연결강도를 수정한다.

$$w^{k+1} = w^k + \Delta w^k = w^k + \alpha \delta_y z \quad (10)$$

$$v^{k+1} = v^k + \Delta v^k = v^k + \alpha \delta_z x \quad (11)$$

**Step 5.** 학습 패턴 쌍을 반복 입력하여 연결강도를 변경하며, 오차  $E$ 가 특정 범위  $E_{max}$  보다 적어지거나 지정된 반복회수를 수행하면 학습을 종료한다. 학습률  $\alpha$ 와 오차신호  $\delta_z, \delta_y$ 에 의해서만 결정되는 연결강도 변화량  $\Delta v, \Delta w$ 에 이전 학습 단계에서의 연결강도 변화량을 보조적으로 활용하는 모멘텀(momentum) 알고리즘을 사용함으로써 학습 속도를 개선시킬 수 있다. 모멘텀 알고리즘을 사용한 연결강도의 변화량은 다음과

같다. 여기서  $\eta$ 는 모멘텀 상수이다.

$$w^{k+1} = w^k + \alpha \delta_y z + \eta \Delta w^{k-1} \quad (12)$$

$$v^{k+1} = v^k + \alpha \delta_z x + \eta \Delta v^{k-1} \quad (13)$$

Delta-bar-delta( $\Delta - \overline{\Delta}$ ) 학습 알고리즘은 각 학습 단계마다 연결강도의 변화에 따라서 학습률을 적응적으로 변경함으로써 학습 단계를 단축하고 궁극적으로 학습 시간을 효과적으로 감소시키는 방법이다. BP 알고리즘은 학습률  $\alpha$ 에 의존하여 연결강도가 변하는 데[8], 일반적으로 여러 학습 단계에 걸쳐 연결강도가 계속 증가하거나 혹은 증가와 감소를 반복하는 경우가 발생한다. 만약, 연결강도가 계속 증가한다면 이것은 학습률  $\alpha$ 가 너무 적기 때문에 연결강도를 적절히 변화시키는 데 상당히 많은 학습 단계가 요구될 것이라고 판단할 수 있으므로 학습률을 보다 큰 값으로 변경하여 학습 속도를 개선할 수 있다. 반면, 학습이 진행되면서 연결강도가 변화가 증가하거나 감소한다면 학습률이 너무 크기 때문에 연결강도가 적절히 변화되지 못한 것이라고 판단할 수 있으므로, 이 경우에는 반대로 학습률을 감소시킴으로써 학습 속도를 개선할 수 있다.

Delta-bar-Delta 알고리즘에서  $\Delta$ 와  $\overline{\Delta}$ 는 다음과 같이 정의된다.

$$\Delta_w^k \equiv -\delta_y z^k \quad (14)$$

$$\overline{\Delta}_w^k \equiv (1 - \beta)\Delta_w^k + \beta \overline{\Delta}_w^{k-1} \quad (15)$$

$$\Delta_v^k \equiv -\delta_z x^k \quad (16)$$

$$\overline{\Delta}_v^k \equiv (1 - \beta)\Delta_v^k + \beta \overline{\Delta}_v^{k-1} \quad (17)$$

위와 같이 정의한  $\Delta$ 와  $\overline{\Delta}$ 을 이용하여  $k+1$  단계에서의 새로운 학습률  $\alpha^{k+1}$ 은 다음과 같이 구한다.

$$\alpha^{k+1} = \begin{cases} \alpha^{k+\kappa} & ; \overline{\Delta}^{k-1} \cdot \Delta^k > 0 \\ (1-\gamma)\alpha^k & ; \overline{\Delta}^{k-1} \cdot \Delta^k < 0 \\ \alpha^k & ; \overline{\Delta}^{k-1} \cdot \Delta^k = 0 \end{cases} \quad (18)$$

여기서,  $\beta$ ,  $\kappa$ ,  $\gamma$ 는 임의의 상수이며 학습율은  $\Delta$ 와

$\overline{\Delta}$ 가 같은 부호를 가지면  $\kappa$  만큼 증가시키고 다른 부호를 가지면  $(1-\gamma)$ 만큼 감소시키는 방법으로 변경된다. 연결강도의 변화는 다음과 같다.

$$w^{k+1} = w^k + \alpha^k \delta_y z + \eta \Delta w^{k-1} \quad (19)$$

$$v^{k+1} = v^k + \alpha^k \delta_z x + \eta \Delta v^{k-1} \quad (20)$$

#### IV. 실험 및 결과

본 논문에서 제안한 방법 Pentium 3.4GHz, Window XP환경에서 Visual C++ 6.0으로 구현하였다. 실험에 적용된 영상은 32x32 크기를 가진 숫자 영상 340개 중에서 220개는 학습 데이터로 적용하였고 120개는 테스트 데이터로 적용하여 반복회수와 학습율에 따라 인식 성능의 민감도를 비교 분석하였다.

그림 7은 실험에 적용된 숫자 패턴의 일부를 나타내었다. 원영상의 학습 데이터는 64x64, 방향성 정보를 이용한 학습데이터는 64x32로 정규화 하여 원영상보다 적은 데이터로 구성하였다.

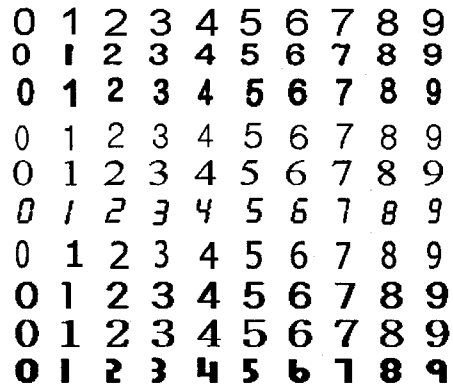


그림 7. 실험영상  
Fig. 7 Experiment images

실험에 적용된 Delta-bar-delta 기반 오류 역전파 알고리즘의 초기 학습율과 오류 한계 등은 모두 동일한 환경에서 실험하였다. 표 1에서 알 수 있듯이 원 영상에서 인식한 결과와 본 논문에서 제안한 방향성 정보의 특성에 의한 방법간의 인식율은 비슷하게 나타났으나 인식 성

능의 민감도에 있어서는 제시된 방법이 효율적인 것을 확인할 수 있다.

표 1. 초기 학습율과 반복횟수에 따른 인식율  
Table. 1 Recognition rate by initial learning rate and iteration

학습율	반복횟수	원영상 (%)	제안된 방법 (%)
0.3	1000	93.75	94.42
	3000	94.01	95.50
	5000	94.67	95.00
	10000	95.83	94.50
	15000	94.93	94.39
0.1	1000	95.25	94.67
	3000	94.83	94.75
	5000	95.58	94.67
	10000	95.75	95.08
	15000	95.83	94.40
0.01	1000	95.50	94.00
	3000	94.66	94.58
	5000	94.08	94.83
	10000	94.92	95.0
	15000	95.17	95.50

그러므로 원 영상을 그대로 적용하지 않고 시각피질의 특성을 이용하여 인식에 적용하는 것이 인간 시각과 같이 처리되는 것을 표 1에서 확인할 수 있다.

## V. 결론 및 향후 연구 계획

본 논문은 인공망막 구현 시 효율적 정보처리를 위하여 인공망막의 여러 방법 중 시각피질 자극 방법에 기반하여 인간시각처리와 유사하게 시각피질의 특성을 고려하여 좀더 인간시각의 정보처리에 의해 영상을 인식하는 방법을 제안하였다. 추후 시각피질과 망막의 특성을 결합하여 좀 더 인간시각의 정보 처리과정과 유사한 모델링을 제시하여 인식 성능을 개선하고자 한다.

## 참고문헌

- [ 1 ] L. Schwiebert, "A Biomedical Smart Sensor for the Visually Impaired," IEEE Sensors 2002, no. 62-2, 2002.
- [ 2 ] Martin fisxhler, Intelligence: The eye, the brain and the computer, Addison-Wesley, 1987.
- [ 3 ] 뇌과학연구소, "뇌정보처리에 기반한 인공 시정각 시스템 연구," 한국과학기술원 연구보고서, 과학기술부, 2001.
- [ 4 ] Rafael C. Gonzalez, Richard E. Woods, Digital image processing, Second edition, Prentice Hall, 2001.
- [ 5 ] 정찬섭의 공역, 감각과 지각, 시그마프레스, 1995.
- [ 6 ] 이인식, '사람과 컴퓨터', 까치 글방, 1992.
- [ 7 ] 김광백, 박충식, "퍼지 제어 시스템을 이용한 학습률 자동 조정 방법에 의한 개선된 역전파 알고리즘", 한국해양정보통신학회논문지, 제8권, 2호, pp.464-470, 2004.
- [ 8 ] D. Hubel and T.N Wiesel, "Receptive Fields Binocular Interaction and Functional Architecture in the Cat's Visual Cortec", Journal of Physiology, pp 106-184, 1962
- [ 9 ] 오창석, '뉴로컴퓨터', 내하출판사, 2000.
- [ 10 ] 김성우, "Support Vector machine을 이용한 문자 인식에 관한 연구", 부산대학교 석사논문, 2002.

## 저자소개



**조 재 현(Jae-Hyeon Cho)**

1998년 부산대학교 전자계산학과  
(이학박사)

2001년 3월~현재 부산가톨릭 대학교  
컴퓨터공학과 부교수

2005년~현재 해양정보통신학회 논문지 편집위원  
※ 관심분야: 신경회로망, 퍼지이론, 인간시각시스템.



**김 도 현(Do-Hyeon Kim)**

2001년 부산대학교 전자계산학과  
졸업

2003년 부산대학교 전자계산학과  
석사 졸업

2006년 부산대학교 컴퓨터공학과 박사 수료  
2007년 부산대학교 컴퓨터공학과 박사 재학 중.  
※ 관심분야: 패턴인식, 영상처리 및 컴퓨터비전, 퍼지  
및 신경망, 제어자동화시스템.



**김 광 백(Kwang-Baek Kim)**

1999년 부산대학교 전자계산학과  
(이학박사)

1997년~현재 신라대학교 컴퓨터정보  
공학부 부교수

2005년~현재 한국해양정보통신학회 이사 및 논문지  
편집부위원장  
※ 관심분야: Neural Networks, Image Processing, Fuzzy  
Logic, Medical Imaging