

논문 2005-42SP-5-15

SVM기반의 선택적 주의집중을 이용한 중첩 패턴 인식

(Recognition of Superimposed Patterns with Selective Attention based on SVM)

배 규 찬*, 박 형 민*, 오 상 훈**, 최 용 선***, 이 수 영***

(Kyu-Chan Bae, Hyung-Min Park, Sang-Hoon Oh, Youg-Sum Choi, and Soo-Young Lee)

요 약

본 논문에서는 신경회로망보다 우수한 성능을 보이는 학습 이론인 SVM을 기반으로, 인간의 인지 과학에서 많은 연구가 이루어지고 있는 선택적 주의집중을 응용한 중첩 패턴 인식 시스템을 제안한다. 제안된 선택적 주의집중 모델은 SVM의 입력단에 주의집중층을 추가하여 SVM의 입력을 직접 변화시키는 학습을 하며 선택적 필터의 기능을 수행한다. 주의집중의 핵심은 학습을 멈추는 적절한 시점을 찾는 것과 그 시점에서 결과를 판단하는 주의집중 척도를 정의하는 것이다. 지지벡터는 주변에 존재하는 패턴들을 대표하는 표본이므로 입력 패턴이 초기상태일 때 주의집중을 하고자 하는 클래스의 가장 가까운 지지벡터를 기준으로 그 지지벡터와의 거리가 최소가 되었을 때 주의집중을 멈추는 것이 적절하다. 일반적인 주의집중을 적용하면 주의집중 척도를 정의하기가 난해해지기 때문에 변형된 입력이 원래 입력의 범위를 넘지 않는다는 제약조건을 추가하여 사용할 수 있는 정보의 폭을 넓히고 새로운 척도를 정의하였다. 이때 사용한 정보는 변형된 입력과 원래 입력의 유클리드 거리, SVM의 출력, 초기상태에 가장 가까웠던 히든뉴런의 출력값이다. 인식 실험을 위해 USPS 숫자 데이터를 사용하여 45개의 조합으로 중첩시켰으며, 주의집중을 적용시켰을 때 단일 SVM보다 인식 성능이 월등히 우수함을 확인하였고, 또한 제한된 주의집중을 사용하였을 때 일반적 주의집중을 이용하는 것 보다 성능이 더 뛰어났음을 확인하였다.

Abstract

We propose a recognition system for superimposed patterns based on selective attention model and SVM which produces better performance than artificial neural network. The proposed selective attention model includes attention layer prior to SVM which affects SVM's input parameters. It also behaves as selective filter. The philosophy behind selective attention model is to find the stopping criteria to stop training and also defines the confidence measure of the selective attention's outcome. Support vector represents the other surrounding sample vectors. The support vector closest to the initial input vector in consideration is chosen. Minimal euclidean distance between the modified input vector based on selective attention and the chosen support vector defines the stopping criteria. It is difficult to define the confidence measure of selective attention if we apply common selective attention model, A new way of defining the confidence measure can be set under the constraint that each modified input pixel does not cross over the boundary of original input pixel, thus the range of applicable information get increased. This method uses the following information; the Euclidean distance between an input pattern and modified pattern, the output of SVM, the support vector output of hidden neuron that is the closest to the initial input pattern. For the recognition experiment, 45 different combinations of USPS digit data are used. Better recognition performance is seen when selective attention is applied along with SVM than SVM only. Also, the proposed selective attention shows better performance than common selective attention.

Keywords : SVM, 선택적 주의집중, 정지점, 주의집중 척도

* 정희원, 한국과학기술원 전자전산학과 및 뇌과학연구소
(Department of Electrical Engineering & Computer Science and Brain Science Research Center, Korea Advanced Institute of Science and Technology)

** 정희원, 목원대학교 정보통신전파공학부
(Division of Information Communication and Radio Engineering, Mokwon University)

*** 정희원, 한국과학기술원 바이오시스템학과 및 뇌과학연구소
(Department of Biosystems and Brain Science Research Center, Korea Advanced Institute of Science and Technology)

접수일자: 2005년2월25일, 수정완료일: 2005년4월28일

I. 서론

컴퓨터의 발달로 인해 복잡한 기계들이 등장하면서 인간이 직접 할 수 없는 많은 일들이 가능하게 되었다. 이와 더불어 기계와 친숙하지 않은 사람들도 좀 더 편리하게 기계를 조작할 수 있게 하기 위해 인간과 기계 사이에 보다 간단한 인터페이스(interface)가 필요하게 되었고, 이에 대한 해결법으로 음성인식, 문자인식 등의 이용이 간편하면서도 성능이 우수한 기술이 등장하였는데, 그 방법 중의 하나로 인간의 시청각 시스템을 모방하여 접목하려는 시도가 있었다. 인간처럼 기계를 학습시키는 학습 이론이 그 중 하나이며, 대표적으로 신경회로망을 들 수 있다^[1]. 하지만 이 알고리즘은 학습 과정에서 국소 최소값에 빠지는 위험 등이 존재하여, 안정적인 학습을 언제나 보장하지 못한다는 단점이 존재하였다. 이에 V.Vapnik는 1995년 Support Vector Machine(SVM)이라 불리는 새로운 학습 방법을 제안하였고^[2], 이것은 커널 공간 해석기법과의 접목을 통해 커널 머신이라는 강력한 알고리즘으로 발전되었다^{[3][4]}. SVM은 이후 전통적인 신경회로망처럼 주어진 문제를 근사를 통해 학습하는 것이 아니라 주어진 데이터를 이용하여 2차 계획 문제를 세우고 이를 푸는 것으로, 주어진 문제에 대해 항상 최적의 해를 보장하는 알고리즘으로 성장했다. 그러나 일반적인 알고리즘의 발달과는 별개로, 실생활에서 흔히 일어나는 노이즈가 섞인 패턴이나 두 개 이상이 중첩되어있는 패턴들에 대해서는 높은 인식률을 얻을 수가 없었으며, 이 문제는 패턴 인식에서 최우선적으로 해결해야 할 문제로 남아있다.

이러한 문제를 해결하는 접근법 중 하나로 인간의 인지능력을 모방하여 보다 좋은 성능을 얻으려는 연구가 시도되었다. 인간에게는 여러 개의 음원이 섞여있는 시끄러운 환경에서도 원하는 사람의 목소리에 집중을 할 수 있고, 여러 개의 그림이 중첩된 가운데에서 원하는 그림을 정확히 짚어낼 수 있는 능력이 있다. 이를 선택적 주의집중이라고 하며 심리학에서 1950년대 이후로 많은 연구가 이루어져 왔다. 그 예로 Cherry는 피실험자의 양쪽 귀에 다른 신호를 인가하고 그 반응을 조사함으로써, 이후의 인간 청각 시스템에서의 선택적 주의집중을 묘사하는 가설에 공헌하였으며^[5], 또한 Triesman은 인간의 시각 시스템에서의 모델을 제안하고, 선택적 주의집중 메카니즘은 여러 다른 차원에서 존재하는 정보들을 결합하는 접착제의 역할을 한다고 설명하였다^[6]. 최근에 들어서는 이러한 인간의 주의집중

능력과 학습 이론을 접목시킨 새로운 인식 시스템으로써 신경회로망을 이용한 주의집중 모델이 발표되었고 실험을 통하여 그 성능이 향상됨이 입증되었다^[7].

그러나 지금까지는 선택적 주의집중과 전통적인 신경회로망과의 접목만이 시도되었고, SVM과의 접목은 아직 발표되지 않고 있다. 따라서 본 논문에서는 신경회로망보다 강력한 SVM과 선택적 주의집중을 접목시키는 새로운 시도를 하여 중첩 패턴 인식에 보다 강력한 인식 시스템을 제안하고자 한다. 본 논문의 II장에서는 본 연구의 기본 배경지식이 되는 SVM과 선택적 주의집중에 대하여 간략하게 설명하고, III장에서는 본 연구가 제안하고 있는 중첩패턴 인식을 위한 주의집중 네트워크의 개념 및 구조에 대한 제안을 한다. 또한 다중 클래스로의 확장 방법 및 각각의 세부 동작에 대한 상세한 설명을 한다. IV장에서는 제안된 개념 및 구조에 대한 실제 실험에 대하여 설명하고, 그 결과를 토대로 인식성능이 향상됨을 검증하고, V장에서 결론을 맺는다.

II. 배경지식

1. Support Vector Machine

SVM은 벡터공간에서 임의의 비선형 경계를 찾아 두 개의 집합을 분류하는 알고리즘이다. 이 때, 두 집합의 경계가 되는 공간을 연속적으로 연결하면 하나의 초월평면이 완성된다. 두 집합을 이상적으로 나눌 수 있는 초월평면의 수는 무수히 많이 존재할 수 있으나, 수학적으로 최적화된 초월평면은 오직 하나만이 결정되며, SVM은 이러한 최적의 초월평면을 찾는 방법이다. 패턴인식에서 전통적으로 많이 사용되어 오던 신경회로망(Artificial Neural Network)의 경우, MLP(Multi Layer Perceptron), RBF(Radius Basis Function) 등의 방법을 이용하여 비슷한 문제를 풀 수 있지만, 문자인식이나 화자인식 등에서 SVM은 신경회로망보다 안정적이고 나은 성능을 보이고 있음이 실험에 의해 보여졌다^[8]. 여기에서는 일반적인 SVM의 구조 및 방법, 선형분리의 가능성, SVM을 멀티클래스에 적용하는 방법에 대하여 간략하게 설명한다.

N 개의 벡터 \mathbf{x}_i 의 분류색인이 $y_i = \pm 1$ 인 학습 패턴들이 있다고 하자. 모든 패턴들이 $y_i(\mathbf{x}_i \cdot \mathbf{w} + b) \geq 1, i = 1, \dots, N$ 를 만족시키는 \mathbf{w} 와 b 가 존재할 때 '선형적 분할 가능'이라고 한다. 이 때,

최적초월평면은 $\mathbf{x}_i \cdot \mathbf{w} + b = 0$ 을 만족하고, 학습자료 중에서 최적초월평면에 가장 가까운 점들을 지지벡터(support vector)라 한다. 최적초월평면을 표현하는 \mathbf{w} 와 b 의 계수 α 는 식(1)의 조건을 만족하는 범위내에서 2차 계획식(2)을 세워 $W(\alpha)$ 최소화함으로써 구할 수 있다.

$$0 \leq \alpha_i \leq C$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \tag{1}$$

$$W(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) \tag{2}$$

여기서 상수 C 는 허용오류의 범위이다. 최적의 초월평면을 $\mathbf{x}_i \cdot \mathbf{w}_0 + b_0$ 라고 하면, $y_i = 1$ 일때 \mathbf{w}_0 와 b_0 는 식(3)과 식(4)처럼 표현된다. 계수 α 는 N 개의 패턴중 지지벡터에만 0이 아닌 값으로 대응된다.

$$\mathbf{w}_0 = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i \tag{3}$$

$$b_0 = 1 - \mathbf{w}_0 \cdot \mathbf{x}_i, \quad i = 1, \dots, N \tag{4}$$

이 경우 클래스를 결정짓는 결정함수는 식(5)와 같이 표현된다.

$$f(\mathbf{x}_i) = \text{sign}(\mathbf{w}_0 \cdot \mathbf{x}_i + b_0), \quad i = 1, \dots, N \tag{5}$$

일반적인 경우의 벡터 패턴 분류를 위해서는 비선형 분류평면이 필요하다. 이러한 비선형 분류 평면을 위해서는 원래의 벡터공간을 다차원 공간으로 확장하여, 확장된 공간의 선형분류가 그것에 해당되도록 할 수 있지만, 일반적으로 확장을 하게 되면 차원이 증가함에 따라 매우 복잡해지고 계산량이 많아진다는 단점이 있다. 그러므로 커널함수를 이용한 방법이 널리 이용된다. 2차식 $W(\alpha)$ 와 $f(\mathbf{x}_i)$ 는 지지벡터와 입력 벡터간의 내적 $(\mathbf{x}_i \cdot \mathbf{x}_j)$ 만으로 계산이 가능하므로 Gaussian RBF, Sigmoid 등의 커널 함수를 사용하여 고차공간의 연산을 원공간 데이터만으로 쉽게 계산할 수 있다. $\Phi(\mathbf{x}_i)$ 를 \mathbf{x}_i 의 고차공간으로 변환된 벡터, k 를 커널 함수라 하면 식(6)처럼,

$$\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) = k(\mathbf{x}_i, \mathbf{x}_j) \tag{6}$$

로 간단히 표현되며 이로써 비선형 경계 분류를 선형

경계 분류문제로 쉽게 해결할 수 있다. 이 때 새로운 $W(\alpha)$ 와 $f(\mathbf{x}_i)$ 를 표현하면 식(7), 식(8)과 같이 표현된다.

$$W(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j \alpha_i \alpha_j k(\mathbf{x}_i, \mathbf{x}_j) \tag{7}$$

$$f(\mathbf{x}_i) = \text{sign}(\sum_{j=1}^N \alpha_j y_j k(\mathbf{x}_i, \mathbf{x}_j) + b_0), \quad i = 1, \dots, N \tag{8}$$

그러나 기본적으로 SVM은 이진 분류만이 가능한 형태이다. 이 경우 방법의 변환을 통하여 멀티클래스 분류 방법으로 이용할 수 있다. 기존의 제안방법 중 본 연구가 다루고 있는 선택적 주의집중과의 접목에 가장 적당한 것은 일대다수대응분류 기법^[9]이다.

일대 다수 대응 분류기법은 주어진 C 개의 클래스를 클래스 A와 클래스 A가 아닌 두 부류로 나누어 문제를 재 정의한다. 그러면 만들 수 있는 쌍의 개수는 다시 C 개가 되고, 각각의 짝은 독립적이기 때문에, SVM을 병렬적으로 적용할 수 있으며 주의집중을 바로 적용할 수 있다. 각 짝에서 C 개의 결정함수가 구해지면 적절한 척도를 통해 가장 신뢰성이 높은 패턴의 클래스로 인식 가능하게 된다. 이 때 우리가 얻게 되는 C 개의 함수의 r 번째 표현식은 식(9)와 같다.

$$z_r(\mathbf{x}_i) = \sum_{j=1}^N \alpha_j^r y_j k(\mathbf{x}_i, \mathbf{x}_j) + b_0^r \tag{9}$$

입력의 입력에 대하여 위의 C 개의 결정함수 값을 계산하고, 그 값이 가장 큰 경우를 찾으면, 입력이 해당하는 클래스를 알 수 있지만, 실제의 경우에는 해당 클래스에서만 이 값은 양수가 되므로 다음과 같이 클래스를 알 수 있다.

$$\begin{aligned} \text{Class} &= \text{argmax} \{z_r : r = 1, \dots, C\} \\ &= \text{arg} \{z_r(\mathbf{x}_i) > 0\} \end{aligned} \tag{10}$$

2. 선택적 주의집중

여러 사람이 동시에 말하는 환경 속에서 인간은 그 중 한사람의 목소리를 쉽게 구분하여 그 사람이 하는 말을 놓치지 않고 들을 수 있다. 이러한 '칵테일파티현상'을 가능하게 하는 물리적인 요소로는 음원의 위치차이와, 서로 미묘한 차이를 가지고 있는 주파수 특징 등을 들 수 있다. 어떠한 특정 정보를 바탕으로 입력되는 정보들 중 필요한 부분에만 집중하는 이러한 현상은 청각뿐만 아니라 시각에서도 찾아 볼 수 있다. 서로 다른

자극이 경쟁적으로 존재하는 가운데 인간은 하나의 사물에만 관심을 두고 집중할 수 있으며, 이러한 현상은 인간이 아닌 다른 동물들에서도 관찰된다. 이러한 현상을 주의집중이라 하며 심리학에서는 오래된 주제중의 하나이다.

주의집중 메커니즘은 지각 능력의 한계를 극복하기 위해 반드시 필요한 수단이다. Willam James는 주의집중이란 명확한 형태로써 동시에 일어나는 가능한 몇 개의 물체나 생각 중 하나가 명확한 형태로써 마음을 소유하는 것이고 의식의 집중(Focalization, Concentration)이 그 요체라고 정의하였으며, 이것은 또한 어느 하나를 효과적으로 다루기 위해서는 다른 것에는 소홀하게 된다는 것을 의미한다고 서술하였다^[10]. 인간이 환경으로부터 정보를 받아들이고 처리하는 경우, 크게 시간 제약, 공간제약의 두 가지 관점에서 능력이 제한되는데, 특히 공간제약을 극복하는 데에는 주의집중이 필수적으로 필요하다. 인간이 선택적 주의집중 메커니즘 없이 자극을 모두 받아들여 두뇌에서 처리하게 되면 두뇌에 과부하가 걸리기 때문이다.

인간의 인지과정은 외부의 자극을 받아들이는 상향식(Bottom-Up)과정과 적극적 주의집중에 의하여 두뇌 내부에서 정보를 만들어 내는 하향식(Top-Down)과정의 조합으로 볼 수 있는데^[11], 상향식과정과 하향식 과정의 협력은 선택적 주의집중에서 중요한 역할을 하는 것으로 알려져 있다.

이러한 현상을 공학적으로 모델링하여 선택적 주의집중의 모델을 만들면 그림 1과 같다. 입력되는 데이터가 있을 때, 그 데이터는 이전에 학습된 지식을 바탕으로 한 필터를 통과한 후 처리된다. 즉, 인가되는 정보는 바로 처리되기 이전에, 미리 어떠한 형태로 추측되고, 그에 따라 도출된 결론의 신뢰도에 따라서 최종적인 결론이 나게 된다.

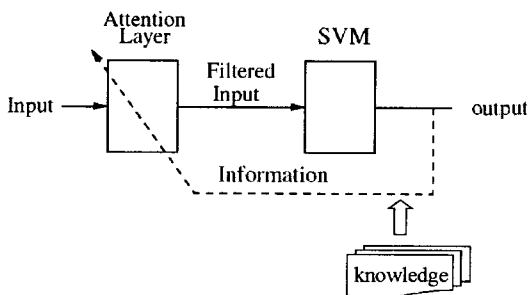


그림 1. 선택적 주의집중의 공학적 모델
Fig. 1. Engineering model of selective attention.

III. 중첩패턴 인식을 위한 선택적 주의집중 네트워크

1. 선택적 주의집중 네트워크의 개요

II장에서 그림 1과 같은 선택적 주의집중의 공학적 모델에 따르면, 주의집중 층이 입력단에 존재하여 원하는 신호는 받아들이며 원하지 않는 신호는 억압하는 선택적 필터의 역할을 수행하고, 이 필터는 출력층에서 역전파되는 어떠한 정보와 입력 패턴에 따라서 네트워크가 자동적으로 구성되어야 한다. 이 역전파는 미리 학습된 지식의 하향식 전달 과정에 대응된다.

이 때, SVM의 출력값은 입력 패턴들이 초월평면으로부터 떨어진 거리를 나타낸다. 하지만 실질적으로 패턴의 클래스를 결정하는 것은 떨어진 정도가 아닌, 단지 초월평면을 기준으로 입력 패턴들이 존재하는 방향 즉, 부호'만'이기 때문에 출력값의 절대적 수치는 큰 의미를 나타내지 않는다. 하지만 기하학적으로 어떠한 패턴이 초월평면으로부터 멀리 떨어져 있으면 그 클래스의 특성을 질게 나타내고 상대적으로 반대편 클래스의 특성은 적게 나타낸다고 말할 수 있다. 선택적 주의집중은 주의집중을 하고자 하는 클래스로 입력 패턴을 변화시켜 보는 것이기 때문에 주의집중 층을 학습하기 위해 필요한 역전파 정보는 SVM 출력의 부호가 아닌 그 값 자체가 된다. 따라서 만일 주의집중 클래스를 초월평면으로부터 양의 방향으로 설정을 했다면, 주의집중 층을 학습시킬 때 입력의 변화량은 입력에 따른 SVM 출력의 증가량이 되고, 반대의 경우에는 감소량이 된다.

2. 선택적 주의집중의 수학적 모델

본 연구에서는 커널 함수로 가우시안 RBF를 사용하였다. 이 함수를 사용하면 학습이 종료된 SVM은 RBF와 동일한 구조를 갖게 된다(그림 2). RBF에서 센터에 해당하는 값은 지지벡터이고 가중치는 $\alpha_i y_i$ 이다. 이 가중치는 SVM을 학습시킴으로써 이미 구해진 값이다.

입력단에서는 외부의 입력 \mathbf{x} 에 SVM이 어느 정도 주의집중을 기울이는지 조절하는 역할이 필요하다. 이 때 두 가지의 가능성을 생각해 볼 수 있다. 첫 번째는 주의집중이득 a 를 정의하여 외부입력 \mathbf{x} 와 SVM의 입력 $\hat{\mathbf{x}}$ 를 일대일로 연결하여 조절하는 방법으로

$$\hat{x}_t = a_t x_t, \quad 0 \leq t \leq M \tag{11}$$

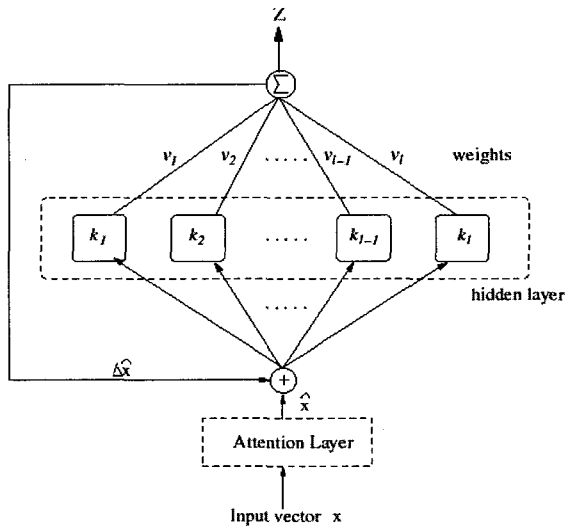


그림 2. SVM의 구조
Fig. 2. Architecture of SVM.

와 같이 표현되며, $a_t = 1 \forall t$ 이면 이 층은 아무런 역할도 하지 않게 되어 전체 네트워크는 순수한 SVM과 동일한 형태가 된다.

두 번째는 어떠한 이득을 정의하지 않고 SVM입력을 바로 변형시키는 방법으로 출력단으로부터 얻은 정보에 의해 입력의 변화량 $\Delta \hat{x}$ 를 계산하여 \hat{x} 로부터 더하거나 빼주는 형식으로 표현된다. 두 가지 방법 모두 입력을 변형시킨다는 점에서는 동일하나 첫 번째 방법의 경우 입력요소 중 0의 값을 갖는 요소가 존재할 경우 이득값이 영향을 전혀 줄 수 없기 때문에 경우에 따라 정보의 손실을 가져올 수 있다는 단점이 있다. 따라서 두 번째 방법으로 입력을 변형시키는 것이 바람직한 선택이다.

즉, 주의집중은 출력값이 원하는 클래스의 영역으로 점점 증가하는 방향으로 이루어지고 주의집중 층에서는 SVM 입력 자체를 변형시킨다. 역전파되는 정보는 출력값의 부호가 아니라 출력값 자체이기 때문에 sign 함수를 거치지 않은 출력함수를 사용해야한다. 가중치에 해당하는 $\alpha_i y_i$ 를 v_i 로, 가우시안 커널함수를 k_i 로 치환하여 SVM 출력함수를 다시 표현하면 식(12)처럼

$$z = \sum_{i=1}^l v_i k_i + b_0, \quad i \in \text{hidden neuron} \leq l \quad (12)$$

where,

$$k_i = \exp(-\gamma \|\hat{x} - s_i\|^2) = \exp(-\gamma \sum_{t=1}^M |\hat{x}_t - s_{it}|^2), \quad t \in \text{input dimension} \leq M \quad (13)$$

가 된다. s_i 는 지지벡터, γ 는 $1/2\sigma^2$ 을 나타낸다.

주의집중 과정은 gradient ascent 알고리즘을 응용한다.

$$\hat{x}_{new^t} = \hat{x}_{old^t} + \eta \frac{\partial z}{\partial \hat{x}_{old^t}} \quad (14)$$

여기서 η 는 학습율을 의미하는 상수이다. 입력에 따른 출력의 변화량은 다음과 같다.

$$\frac{\partial z}{\partial \hat{x}_t} = \sum_{i=1}^l \left(\frac{\partial z}{\partial k_i} \cdot \frac{\partial k_i}{\partial \hat{x}_t} \right) \quad (15)$$

이 때, 식(12)과 식(13)로부터

$$\frac{\partial z}{\partial k_i} = v_i \quad (16)$$

$$\frac{\partial k_i}{\partial \hat{x}_t} = -2\gamma(\hat{x}_t - s_{it}) \exp\left(-\gamma \sum_{t=1}^M |\hat{x}_t - s_{it}|^2\right) \quad (17)$$

와 같이 각각의 미분항을 구할 수 있다. 위 식들을 정리하면 SVM의 변형된 입력의 학습식은

$$\hat{x}_{new} = \hat{x}_{old} - 2\eta \gamma \sum_{i=1}^l v_i (\hat{x}_{old} - s_{it}) \exp\left(-\gamma \sum_{t=1}^M |\hat{x}_{old} - s_{it}|^2\right) \quad (18)$$

로 최종적인 형태를 갖추게 된다.

3. 중첩 패턴 인식을 위한 네트워크의 동작

그림 3은 일대다수 대응기법으로 확장시킨 주의집중 네트워크의 블록도이다. N개의 클래스가 있을 때 N개의

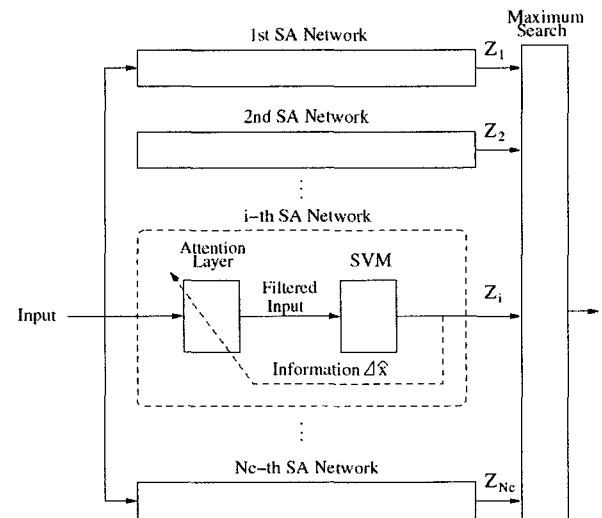


그림 3. 일대다수 대응기법으로 확장시킨 주의집중 네트워크
Fig. 3. Selective Attention network extended by one-against-rest method.

SVM을 학습시키고, 주의집중을 할때도 최대 N개의 SVM에서 병렬적으로 진행할 수 있는 가장 적합한 확장방법이다. 이 네트워크의 동작은 다음과 같은 순서로 이루어진다.

가. SVM의 학습

선택적 주의집중은 SVM의 학습이 끝난 후에 이루어진다. 학습시 지지벡터의 개수와 인식률 사이에 적절한 trade off가 필요하다. 지지벡터 개수가 너무 많으면 네트워크의 계산량이 많아져서 주의집중을 통해 얻는 인식률의 이득이 비효율적이 될 가능성이 생기고, 인식률을 너무 낮게 잡으면 주의집중 효과를 증대시키기 위한 인위적인 조작이 되기 때문이다. 하지만 중첩 패턴의 분리가 주목적이라면 인식성능의 신뢰도가 중요하므로 가급적이면 인식률이 최대가 되도록 학습시키는 것이 바람직하다. SVM 학습이 종료되면, 선택적 주의집중 학습으로 이어지고 이때 구해진 변수들은 값을 고정시킨다.

나. 1차 주의집중

두 가지의 패턴이 중첩되어있고 이를 인식하기 위해서는 두 패턴을 모두 인식해야한다. 첫 번째 주의집중에서는 두 패턴 중에서 어느 하나만을 인식한다. 이것을 1차 주의집중이라 부른다. 이때 적당한 시점에서 주의집중을 멈추는 것이 중요하다. 주의집중을 거의 하지 않게 되면 그 의미가 없어지고 반대로 너무 많이 하게 되면 전혀 엉뚱한 입력 패턴에 대해서도 오인식을 하는 경우가 생기기 때문이다. SVM은 다중 클래스로 확장을 하더라도 이진 분류의 특성을 그대로 지니기 때문에 주의집중 학습이 진행되면 출력값은 초월평면의 수직방향으로 어떠한 목적이 없이 계속 커지기만 한다. 따라서 모든 패턴에 대해 논리적으로 공평하게 적용할 수 있고, 높은 인식률도 얻을 수 있는 정지점 척도가 필요하다.

(1) 정지점 척도

그림 4는 고차원 커널 공간을 2차원적으로 알아보기 쉽게 표현한 것이다. 지지벡터들이 놓여있는 경계는 SVM 학습 단계에서 어떤 클래스로 분류되기 위한, 초월평면으로부터 떨어진 최소한의 위치를 나타낸다. 그리고 그 지지벡터들은 학습패턴의 일부로써 그 주변의 입력패턴들을 대표하는 일종의 견본이라고 볼 수 있다.

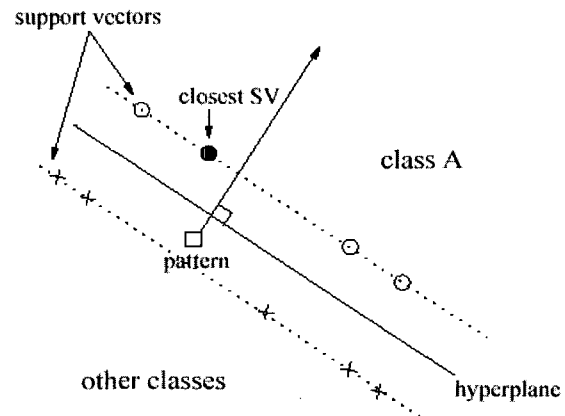


그림 4. 주의집중 과정에서의 입력패턴의 이동
Fig. 4. movement of input pattern through attention.

이 지지벡터들을 중심으로 가우시안 함수가 씌어져 있기 때문에 어떤 지지벡터와 가까운 거리에 있는 패턴들은 그 지지벡터의 특성을 많이 가지고 있다고 볼 수 있다. 따라서 초기상태의 입력패턴에 주의집중을 하면 초월평면과 수직한 방향으로 이동을 하게 되고 초기에 자신과 가장 비슷하였던 지지벡터와의 거리가 최소가 되는 시점, 즉 지지벡터들이 놓여있는 경계에 다다르게 되면 그 지지벡터와 가장 흡사한 모양으로 변화하게 된다. 따라서 SVM에서의 선택적 주의집중을 위한 ‘정지점 척도’를 다음과 같이 정의할 수 있다.

- 임의의 입력패턴을 클래스 A로 주의집중 할 때, 목표벡터는 초기상태에서 그 입력패턴과 가장 가까운 곳에 위치하는 클래스 A 영역의 지지벡터이며, 주의집중을 멈추는 시점은 그 지지벡터와 가장 가까워졌을 때 이다.

이러한 정지점 척도를 사용하면, 초기상태에서 지지벡터 경계보다 해당 클래스 쪽으로 위치했던 패턴들에 대해서는 굳이 주의집중을 할 필요가 없으므로 효율적인 계산을 할 수 있다.

(2) 주의집중 척도

주의집중이 종료되면 임의의 입력패턴에 대해 각각의 클래스에 대한 신뢰도를 나타내는 주의집중 척도를 정의하여 클래스를 결정한다. 이 척도를 어떻게 정의하느냐가 인식률에 결정적인 영향을 주게 된다. 주의집중이 종료된 시점에서 주의집중 척도를 위해 사용할 수 있는 정보는 다음과 같다.

- ㄱ. Y_0^c : 주의집중 이전의 SVM 출력값
- ㄴ. Y_A^c : 주의집중 이후의 SVM 출력값
- ㄷ. D_A : 원래의 입력과 변형된 입력사이의 유클리드 거리 (Euclidean distance)
- ㄹ. H_p^c : 초기상태에 가장 가까웠던 지지벡터에 해당하는 히든뉴런의 최대값

입력패턴이 c 클래스에 속해있을때 ㄱ, ㄴ, ㄷ의 정보는 상대적으로 큰 값을 가지고, ㄹ의 정보는 상대적으로 작은 값을 가지게 된다. 주의집중 척도로 사용할 수 있는 정보는 위의 네 가지 외에는 없다. 하지만 이 중에서 ㄴ과 ㄹ의 정보를 사용하는 데는 논리적인 모순이 존재한다.

입력패턴은 수직방향으로 이동을 하고 초기에 가장 가까운 지지벡터와의 거리가 최소가 되는 지점은 지지벡터들이 놓여있는 경계이기 때문에 ㄴ 정보의 경우 모든 패턴들에 대해 같은 값을 갖게 되어 변별력이 없다. 그리고 ㄹ 정보는 그림 5에서 볼 수 있듯이 어떠한 패턴이 클래스 c에 속한다고 해서 반드시 다른 클래스에 비해 가장 큰 값을 갖는다는 보장이 없다. 이러한 모순으로 인해 결국 논리적으로 문제가 없는 정보는 ㄱ과 ㄷ 뿐이다. 이 정보들만으로 주의집중 척도를 식(19)와 같이 정의하였다.

$$CM = \frac{Y_0}{offset + (D_A)^\alpha} \tag{19}$$

초기상태에 지지벡터의 경계보다 초월평면으로부터 더 멀리 떨어져있던 패턴들은 주의집중을 하지 않기 때

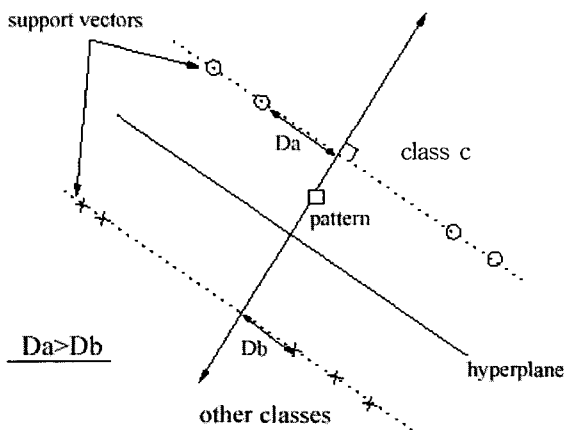


그림 5. 정보사용의 모순
Fig. 5. contradiction of use of information.

문에 $D_A = 0$ 이 된다. 이를 방지하기 위해 offset 상수를 이용한다. 이는 CM값에 거의 영향을 끼치지 않는 범위 내에서 적당히 작은 값을 사용하면 된다. α 는 Y_0 와 D_A 의 상대적 중요도를 나타내는 것으로써 0일 경우에는 Y_0 만을 고려하여 주의집중을 하지 않는 일반적인 SVM의 경우이며, α 가 커질수록 주의집중에 보다 더 많이 의존하게 된다. 최종적으로는 CM값이 가장 큰 클래스를 인식 결과로 사용한다. 하지만 Y_0 값은 신뢰도가 매우 중요하다. SVM만을 사용했을 때 인식률이 충분히 높다면, 이 정보를 사용할만한 가치가 있지만 복잡한 형태의 중첩패턴 인식 문제에서는 신뢰도가 높지 않기 때문에 사용하는데 다소 무리가 있다. 또한 이 문제로 인해 D_A 값 만을 척도로 사용할 수도 없다. 이것은 IV절에서 실험적으로 증명되었다. 사용가능한 정보는 앞서 언급한 네 가지 정보가 전부이기 때문에, 이 문제를 해결하기 위해서는 네가지 정보들을 효율적으로 사용할 수 있어야 한다.

(3) 선택적 주의집중의 제한

일반적인 선택적 주의집중은 입력패턴을 원하는 클래스로 변형을 시키는 것인데, 중첩패턴을 인식하는 경우 가장 이상적인 상황은 입력패턴을 주의집중 하고자 하는 클래스로 변형하는 것 보다는 원치 않는 클래스의 정보를 제거하여 원하는 클래스의 정보만을 남기는 것이다. 두개의 패턴이 중첩되어 충분한 정보를 지니고 있는 입력에 굳이 주의집중 하고자 하는 클래스의 정보를 더 추가시킬 필요가 없는 것이다. 극단적인 예를 들면 정보가 전혀 없는 블랙 이미지를 입력으로 인가할 경우, 매우 많은 주의집중을 거치면 모든 클래스에 대해 이미지를 생성하는 결과가 나온다. 이러한 방식의 주의집중은 입력 패턴이 가지고 있는 원래의 특성을 없애버린다.

이러한 문제를 해결하기 위해 선택적 주의집중에 제약조건을 걸어줄 수 있다. 이는 단순히 기존의 중첩패턴이 가지고 있던 정보의 범위를 변형패턴이 넘지 않도록 하는 것이다. 어떠한 입력패턴을 클래스 c로 주의집중 할 때 원래 속해있던 패턴은 제약조건으로 인한 영향을 거의 받지 않지만 다른 클래스의 패턴이 들어오면 원래 가지고 있던 정보의 범위를 넘어서지 못하기 때문에 겹치는 부분만 남고 나머지는 모두 제거된다. 따라서 일반적인 주의집중에서는 전혀 다른 형태의 패턴이

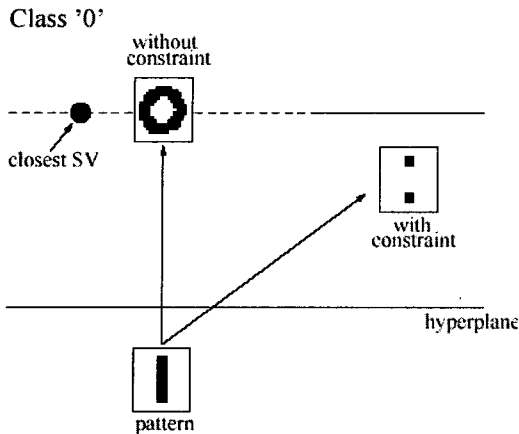


그림 6. 제약조건의 유무에 따른 입력패턴의 변화
 Fig. 6. change of input pattern depends on existence of constraint.

들어와도 주의집중 클래스의 지지벡터 경계까지 도달할 수 있었지만, 제약조건을 걸어주면 패널티를 받게 되어 주의집중 클래스의 경계조차도 넘지 못하는 경우가 생긴다. 이러한 점은 Y_A 를 사용할 수 있는 변별력을 만들어 준다.

또한 제한된 선택적 주의집중은 입력패턴의 학습 방향에도 영향을 준다.

그림 6에서는 '0'과 '1'을 예로 들어 제약조건의 유무에 따른 입력패턴의 변화를 나타내고 있다. 일반적인 주의집중을 할 경우에는 입력패턴이 초월평면으로부터 수직방향으로 이동을 하며 '1'이라는 클래스의 특성이 무시된 채 완전한 '0'의 모양을 만들어낸다. 반면에 제한된 주의집중을 사용한 경우에는 '0'과 '1'의 공통부분만을 남기고 나머지는 제거된다. 이럴 경우 변형된 입력은 건본이 되는 지지벡터와도 가까워질 수 없기 때문에 학습방향이 수직방향에서 비껴난다. 이 경우 건본벡터와 가장 가까워지는 시점은 지지벡터들이 놓여있는 경계가 아닌 오히려 초월평면에 가까운 위치가 되고 이때의 건본벡터와의 거리는 상대적으로 매우 멀어지게 된다. 따라서 H_p 의 필요성을 알 수 있다. 원래 주의집중 클래스에 속하는 패턴들은 제약조건의 영향이 거의 없고, 속하지 않는 패턴들에 대해서만 강력한 패널티를 적용하기 때문에 인식성능 면에서도 더 뛰어나다.

이러한 이유로 Y_A 와 H_p 를 사용하여 새로운 주의집중 척도를 식(20)과 같이 정의하였다.

$$CM = \frac{\text{sign}(Y_A) \cdot |Y_A|^\alpha \cdot (H_p)^\beta}{\text{offset} + (D_A)^\gamma} \quad (20)$$

Y_A 의 부호를 고려한 이유는 주의집중을 기울인 클래스의 영역으로 미처 들어오지 못한 패턴들에 대해 패널티를 적용하기 위함이다.

다. 2차 주의집중

(1) 주의집중 천이

1차 주의집중이 종료되면 SVM의 입력벡터는 인식한 클래스의 형태로 변형되어 있다. 두 클래스 중 어느 한 쪽을 인식하는데 성공하면 다른 하나의 클래스를 인식해야 하는데 이때의 주의집중을 2차 주의집중이라고 한다. 2차 주의집중의 입력패턴을 만들기 위해 주의집중 천이 기법을 사용한다. 즉, 변형된 입력이 가리키는 클래스로의 주의집중을 해제하고 나머지 클래스의 정보를 남긴다는 것인데, 이때 중요한 점은 중첩된 두 패턴의 공통부분은 천이가 일어나지 않아야 한다는 것이다. 두 패턴이 중첩된 부분이 많으면 주의집중을 해제하면서 2차 인식에 필요한 정보까지 제거되기 때문에 오인식을 유발할 가능성이 높아지기 때문이다.

여기서 다시 하향식 주의집중의 원리를 짚어보면, 입력을 통해 결과를 얻는 것이 아니라 가정된 결과를 토대로 입력의 변화하는 방향을 결정짓는 것이다. 이러한 사실은 우리에게 해결책을 제시해준다.

중첩패턴 \mathbf{x} 가 있고 1차 주의집중을 통해 클래스 c 로 인식을 했다고 할 때, 변형된 패턴을 $\hat{\mathbf{x}}$ 라고 정의하자. 다중 클래스 주의집중 네트워크의 각각의 작은 주의집중 클래스가 이미 정해져있는 상태이다. 따라서 각 클래스의 표본이 되는 특징벡터를 정의하여 원래의 중첩패턴으로부터 변형된 패턴을 제거할 때, 특징벡터와 중첩패턴의 공통부분을 열외시킴으로써 이 문제를 해결할 수 있다. 이 과정은 우선 중첩 패턴에서 변형된 패턴을 빼고, 특징벡터를 더하되 중첩패턴의 크기를 넘지 못하도록 제한함으로써 이루어진다. 이 특징벡터를 $\mathbf{x}_{feature}$, 2차 주의집중의 입력패턴을 \mathbf{x}_{2nd} 라고 하면, 식(21)과 같은 주의집중 천이 알고리즘을 정의할 수 있다.

$$\mathbf{x}_{2nd} = \mathbf{x} - \hat{\mathbf{x}}_c + a \cdot \mathbf{x}_{feature} \quad (21)$$

$$\text{if } \mathbf{x}_{2nd} > \mathbf{x}^t, \mathbf{x}_{2nd} = \mathbf{x}^t$$

'a'는 특징벡터의 크기를 조절해 주는 상수이다. 각 클래스로 주의집중을 할 때 입력패턴이 특징벡터의 성분을 너무 많이 포함하면, 인식해야할 클래스의 순위권이 밀려나는 경우가 생길 수 있으므로, 1보다 작은 값

을 선택한다.

(2) 2차 주의집중

이 과정은 주의집중 방법 면에서는 1차 주의집중 방법과 동일하다. 다만 1차 주의집중에서 N개 중 2개를 고르는 방법과는 달리, 2차 주의집중에서는 1차에서 올바르게 인식한 경우에 대해서만 나머지 하나를 찾는 방법 이므로 N-1 개 중 1개를 고르는 방법이 된다. 실제로 이 과정은 반드시 필요한 과정은 아니다. 1차 주의집중 을 거친 후, 주의집중 천이를 거친 새로운 입력을 주의 집중 없이 SVM만 통과시켜서 높은 인식률을 얻는다면, 굳이 계산량이 많은 주의집중을 통과시킬 필요는 없다. 하지만 주의집중 천이 과정에서 공통부분을 효과적으로 살리지 못한 경우나 낮은 인식률을 얻었을 경우는 2차 주의집중을 통해 이를 보완할 수 있다.

IV. 인식 실험

1. 데이터베이스 및 실험 구성

본 연구의 실험에서서 이용한 데이터베이스(DB)는 SVM 알고리즘 검증 실험으로 널리 사용되는 ‘USPS handwritten dataset’이다. 이 DB는 ‘US Postal Service’에서 제작하였으며, ‘0~9’까지 10개의 클래스로 이루어져 있다. 학습을 위해서는 7291개의 패턴이 있고, 테스트를 위해서는 2007개의 패턴이 있다. 각각의 이미지는 16×16의 크기를 가지며, [-1 1] 영역의 그레이 스케일로 이루어져 있다. USPS DB 구성패턴의 예를 그림 7에, 구성성분을 표 1에 각각 나타내었다.

USPS DB의 원본데이터는 선택적 주의집중을 하기 전, SVM을 학습시킬 용도로 이용되었다. 그리고 이 연구의 핵심이 되는 중첩된 이미지는 원본데이터 중에서 임의로 두개의 패턴을 골라 중첩시킨 이미지를 사용하였다. 10개의 클래스가 존재하므로 총 45개의 조합이

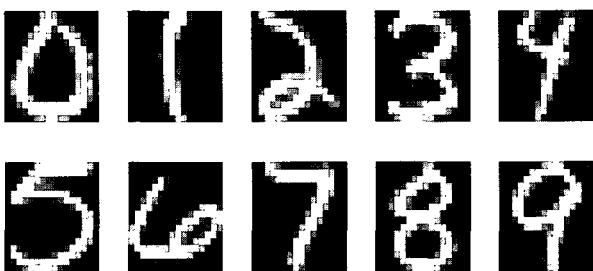


그림 7. USPS DB 구성패턴의 예
Fig. 7. Examples of USPS DB patterns.

표 1. USPS Dataset의 구성
Table 1. composition of USPS Dataset.

USPS DATASET		
Class	Train	Test
0	1194	359
1	1105	264
2	731	198
3	658	166
4	652	200
5	556	160
6	664	170
7	645	147
8	542	166
9	644	177



그림 8. USPS DB로 조합한 중첩패턴의 예
Fig. 8. Examples of superimposed patterns.

나오고, 각각 50개로 구성되어 총 2250개의 중첩 패턴을 만들었다. 그림 8에 몇 가지 예가 나타나 있다. 중첩되는 부분이 많고 복잡한 형태를 띠기 때문에 인간이 직접 구분을 하기도 어려운 패턴들이 많이 존재함을 알 수 있다.

2. SVM 학습

멀티클래스 인식실험을 위하여 그림 3과 같은 10개의 독립적인 SVM 시스템을 설정하였다. 이러한 일대다수 대응기법의 SVM은 각각 ‘i-th class & others’의 형태를 띠고 있다. 일반적인 SVM을 이용한 USPS DB의 인식률은 기존의 연구에서 4~45%의 오차로 발표된 바 있으므로^[8], 본 연구의 실험에서도 객관적인 비교가 용이하도록, 전체 인식률을 기존의 결과와 비슷하게 설정하였다. 즉, 전체 지지벡터의 수를 조절하고 학습을 시켜, 최종적으로 95.6%의 인식률을 갖는 SVM의 계수를 얻었다. 각 SVM의 지지벡터 개수와 총 인식률을 표 2에 나타내었다.

표 2. 각 네트워크에서의 지지벡터 개수와 총 인식률
Table 2. # of support vectors of each network and total recognition ratel.

class	0	1	2	3	4	5	6	7	8	9
#SV	464	114	678	590	563	685	430	373	662	493
R.R	95.6%									

3. 단일패턴에 대한 주의집중 인식실험

중첩패턴에 대한 주의집중 실험에 앞서, 단일패턴에 대해서 주의집중이 얼마만큼의 영향을 끼치는지 알아보기 위해 인식 실험을 해보았다. 이 때, SVM은 학습이 종료되었기 때문에 가중치가 고정되었으며, 오직 선택적 주의집중 네트워크만 더 부가되어 그 효과를 검토하였다. 이때의 인식률은 95.8%로 나타났다. baseline 보다 0.2% 인식률의 향상이 있었으며, 실제로 거의 인식률이 상승하지 않은 것으로 간주할 수 있다. 기존의 연구에 의하면 주의집중을 인식기와 접목하였을 때 노이즈가 없을 경우에는 인식률에 거의 영향을 끼치지 않지만 노이즈가 많을수록 그 성능을 발휘하는 것으로 발표된 바 있다^[12]. 노이즈가 섞이지 않은 패턴들은 굳이 주의집중을 할 필요 없이 인식기만으로도 충분히 신뢰성이 있는 인식률을 얻을 수 있음을 알 수 있다.

4. 1차 주의집중 인식실험

가. SVM만을 이용한 중첩패턴의 인식실험

2250개의 중첩패턴에 대하여 SVM의 가중치를 고정시킨 채, 주의집중 과정 없이 인식실험을 수행하여 보았다. 중첩패턴에는 두가지 클래스의 정보가 섞여있으므로, 두 클래스중 하나의 클래스를 선택하였을 때, 올바르게 인식한 것으로 간주하였다. 클래스를 선택하는 기준은 SVM의 출력값들을 비교해 가장 큰 값에 해당하는 클래스를 인식하는 것으로 하였으며, 한 가지 제한조건은 SVM을 학습할 때 주의집중 클래스를 초월평면으로부터 양의 방향으로 설정해주었기 때문에 최소한 양의 값을 가지는 출력들에 대해서만 인식을 하였다. 이는 오입력 거절의 역할을 자체적으로 수행하는 효과를 갖고 있다. 이 경우의 인식률은 46.7%로 나타났다. 이러한 매우 저조한 인식률을 바탕으로, 단일 인식기만으로는 복잡한 중첩패턴을 인식하는 것이 불가능하다는 사실이 실험적으로 증명되었다.

나. 일반적인 주의집중을 적용한 중첩 패턴 인식 실험

식(19)의 주의집중 척도를 이용하여 실험을 반복하였다. SVM만을 이용한 인식실험에서처럼 오입력 거절을 위해 초월평면으로부터 양의 영역으로 이동한 패턴들에 대해서만 인식을 인정하였다. 이 경우 인식률은 45.7%로 나타났다. 즉, SVM만을 사용했을 때와 비교하여 인식률의 상승이 없음을 알 수 있다. 이는 중첩패턴들의 경우 SVM만을 통과시킨 출력값의 대부분이 음의 값을

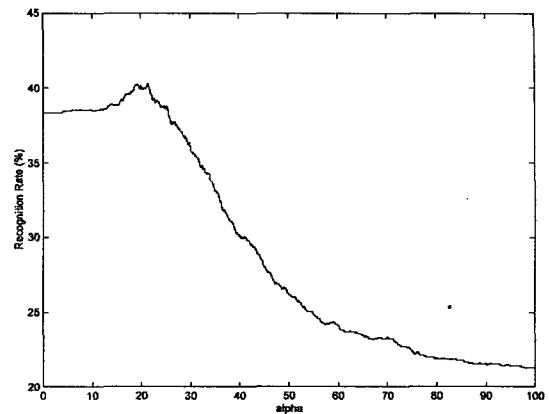


그림 9. α 의 변화에 따른 인식률
Fig. 9. recognition rates with various α values.

가지기 때문에, 주의집중 척도에서 Y_0 의 신뢰도가 매우 낮기 때문에 발생한 현상이다.

주의집중 척도에서 Y_0 가 인식을 방해하는 요소로 작용하므로, $Y_0 = 1$ 로 고정하고 D_A 만의 정보로는 어떠한 결과를 나타내는지 실험을 반복하였다. 파라미터로 α 의 값을 설정하여야 하는데, α 의 값을 변화시키며 인식률을 구해보았다. 그 결과를 그림 9에 나타내었다.

$Y_0 = 1$ 로 고정하고 실험한 결과, 최대 인식률이 약 40%정도인 것을 알 수 있다. 따라서 D_A 도 논리적인 모순은 없으나 단일 정보로 사용하기에는 무리가 있음을 알 수 있다. 그러므로 일반적인 주의집중 척도는 이러한 중첩패턴에 대하여 적용하는데에 무리가 있으며, 새로운 제약조건을 가진 선택적 주의집중 방법이 필요함을 알 수 있다.

다. 제한된 주의집중을 적용한 중첩패턴의 인식실험

(1) 제한된 주의집중에 대한 검증

일반적이 주의집중이 인식 시스템의 성능향상을 가져오지 못하므로 III절에서 제안한 제한된 주의집중을 적용할 필요가 있다. 실제의 제한된 주의집중의 효과를

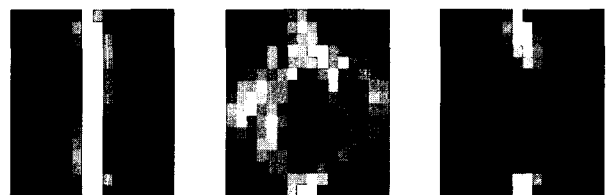


그림 10. 제약조건에 따른 입력패턴의 변화
Fig. 10. changes of input pattern with existence of constraint.

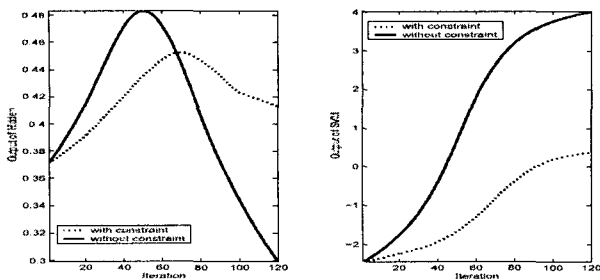


그림 11. 제약조건에 따른 히든뉴런의 출력과 SVM 출력
Fig. 11. output graphs of hidden neuron and SVM with existence of constraint.

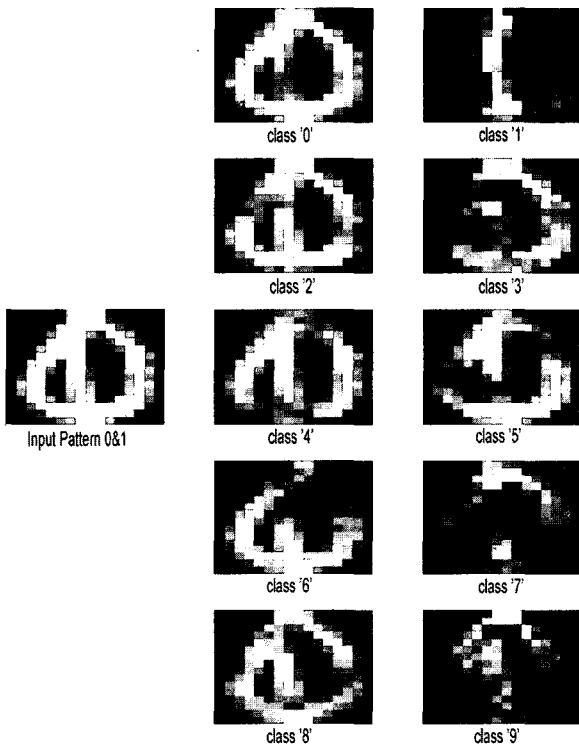


그림 12. 제한된 주의집중을 이용한 입력패턴의 각 클래스로의 변화
Fig. 12. change of input pattern to each class based on constrained selective attention.

‘1’의 클래스에 속하는 패턴을 ‘0’으로 주의집중 시켜보았다. 제약조건의 유무에 따른 주의집중의 효과를 그림 10에 나타내었다.

왼쪽의 그림은 원래의 입력패턴, 가운데 그림은 제약조건이 없을 때 ‘0’으로 주의집중한 결과, 오른쪽 그림은 제약조건이 있을 때의 ‘0’으로 주의집중 한 결과이다. 제약조건이 없을 때에는 불필요한 정보를 많이 생성하였으며, 실제의 정보와는 전혀 다른 정보로 변모했음을 알 수 있다. 이러한 현상은 실제의 정보는 ‘1’클래스임에도 불구하고 너무 많은 주의집중의 효과에 따라 실제 인식결과가 ‘0’으로 나오게 되는 결과를 초래한다. 이러

한 현상을 자세히 알아보기 위하여, ‘1’의 클래스에 해당하는 패턴을 ‘0’으로 주의집중 하는 경우, 제약조건에 따른 SVM의 히든뉴런값 및 출력값을 그림 11에 나타내었다.

각 그래프의 가로축은 주의집중 시 iteration의 횟수이며, 세로축은 그때의 값을 나타낸다. 왼쪽의 그래프는 SVM 내부의 히든뉴런값이고, 오른쪽의 그래프는 SVM의 출력값이다. 제약조건이 없을 경우 실선으로 표현하였고, 제약조건이 있을 경우 점선으로 표현하였다. 제약조건이 없을 경우 출력값이 매우 커짐을 알 수 있다. 이것은 실제로 이 패턴이 ‘0’으로 인식된다는 것을 의미한다. 그러나 제약조건이 있을 경우 출력값이 적절한 패널티를 받아 어느값 이상으로 커지지 않는다는 것을 알 수 있다. 이러한 실험 결과에서 Y_A 와 H_P 의 정보는 인식기에서 변별력을 갖고 있다는 논리가 성립한다.

다음으로, ‘0’과 ‘1’이 중첩된 패턴이 입력으로 인가될 때, 제한된 선택적 주의집중을 사용하여 입력을 여러 클래스로 변화시켜보았다. 그 결과를 그림 12에 나타내었다. 변형패턴이 원래 입력패턴의 범위를 넘지 못하기 때문에 ‘0’이나 ‘1’이외의 클래스로 주의집중을 할 때는 형태가 잘 갖추어지지 않음을 확인할 수 있다.

(2) 제한된 주의집중을 이용한 인식실험

앞서 검증된 제한된 주의집중을 이용하여 실험을 반복하였다. 인식결과는 새로이 제안한 식(20)의 주의집중척도를 이용하여 계산되었다. 이 때, 주의집중 이전의 정보를 사용하지 않기 때문에 각 정보의 상대적 중요도를 나타내는 지수항 α, β, γ 의 값은 인식 성능에 직접적인 영향을 준다. 각 변수에 대해 높은 인식률을 찾기 위하여 변수마다 일정한 범위내의 가능한 모든 조합으로 인식률을 계산하였다. 하지만 변수가 많아지면 조합이 기하급수적으로 많아지므로 불필요한 계산이 생기게 된다. 주의집중 척도 식(20)에서 각각의 지수항은 서로 상대적인 관계를 가지고 있기 때문에 하나의 변수를 고정시키고 실험하였다. 본 실험에서는 $\gamma = 1$ 로 고정하였고, α, β 에 대해서 인식률의 변화를 측정해보았다. α, β 를 각각 0~10까지 변화시키면서 인식률을 구한 그래프가 그림 13에 나타나있다.

Y_A 와 H_P 두 정보 모두 중요도가 높지만 Y_A 가 좀더 인식률에 민감하게 작용하는 것으로 나타났다. 이 때 구해진 최대 인식률은 $\alpha = 1.5, \beta = 1.5$ 에서 83.3%였다. 주의집중을 하기 전 SVM만을 이용한 인

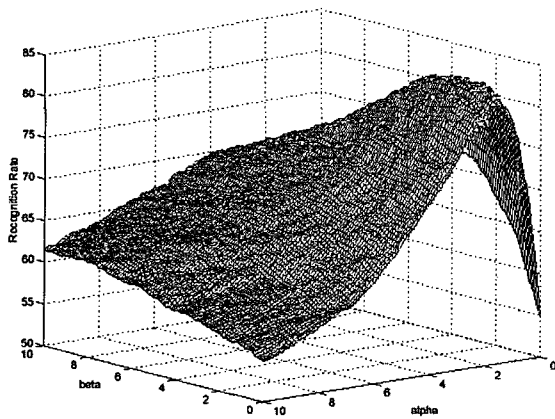


그림 13. α, β 에 따른 인식률의 변화

Fig. 13. recognition rates with various α, β .

식실험이 인식률이 46.7%이었던 것이 비하여 매우 높은 성능향상이 있었다.

라. 적용방식에 따른 인식률 비교

세 가지 방식에 따른 인식률의 변화를 표 3에 나타내었다. 중첩패턴에 대하여 SVM만을 이용하여 얻은 인식률 부분을 'SVM only'라고 표현하였고, 일반적인 주의집중을 이용하여 얻은 인식률 부분을 'GSA'라고 표현하였으며, 제한된 주의집중을 이용하여 얻은 인식률 부분을 'RSA'라고 표현하였다. 이 결과에서 제한된 주의집중 방법은 매우 뛰어난 성능을 가지고 있는 것이 입증되었다.

표 3. 적용방식에 따른 인식률 비교

Table 3. comparison of recognition rates among various methods.

	SVM only	GSA	RSA
인식률	46.7 %	46.7 %	83.3 %

5. 2차 주의집중 인식실험

가. 주의집중 천이

두 가지 클래스의 정보가 섞인 중첩패턴에 대하여 두 가지 클래스를 모두 알아내기 위해서는, 1차 주의집중이 끝난 뒤에 주의집중 천이과정이 필요하다. 이 때 이용되는 식은 식(21)이며, 'a'의 값이 중요한 파라미터가 된다. x_{2nd} 가 더 이상의 주의집중이 없을 경우 나머지 하나의 패턴을 찾아내는지 알아보기 위하여 SVM만을 이용한 인식실험을 하였다. 'a'값을 변화시켜 가며 얻은 결과를 표 4에 정리하였다.

표 4. a 값에 따른 인식률의 변화

Table 4. change of recognition rate with various 'a'.

a	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3
인식률(%)	39.9	39.9	42.1	42.5	42.6	42.7	42	40.3

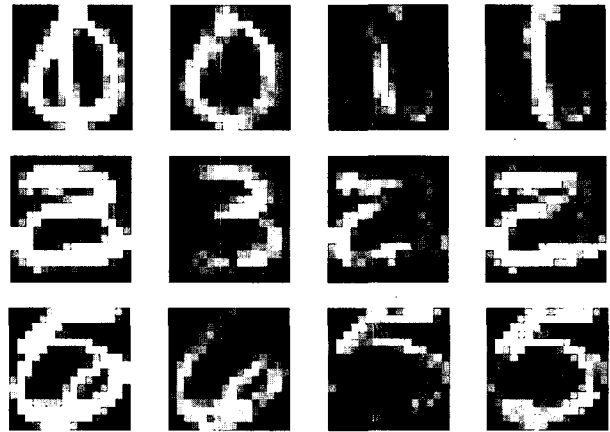


그림 14. 특징벡터의 사용에 따른 2차 주의집중의 입력 패턴

Fig. 14. Input patterns of 2nd selective attention with using the feature vector.

이 결과로부터 a=0.5인 경우 인식률은 42.7%로 가장 높음을 알 수 있다. 여기에 2차 주의집중을 인가하면 더 높은 인식률을 얻을 수 있다. 이러한 특징벡터를 이용하여 2차 주의집중 입력패턴을 만드는 과정을 몇 가지 예를 들어 그림 14에서 나타내었다.

첫 번째 열은 원래의 중첩패턴을 나타내고 두 번째 열은 1차 주의집중이 끝난 후 변형된 패턴, 세 번째 열은 중첩패턴에서 변형된 패턴을 뺀 결과, 네 번째 열은 특징벡터를 이용하여 중첩부분을 보완한 결과이다. 세 번째 열에서 손상이 많이 된 중첩부분이 특징벡터를 더해줌으로써 많이 보완된 것을 확인할 수 있다.

나. 2차 주의집중 인식실험

주의집중 천이로 인해 만들어진 새로운 입력 패턴은 인식기만을 사용했을 때 인식률이 저조하기 때문에 2차 주의집중이 필요하다. 2차 주의집중은 1차에서 이미 인식한 클래스는 제외하고, 나머지 9개의 클래스 중에서 1개의 클래스를 찾는 주의집중 방법이다. 인식률은 두 가지 형태로 계산하였다. 하나는 전체 패턴에 대해 두 가지 클래스를 모두 올바르게 인식한 경우의 인식률, 다른 하나는 1차에서 올바르게 인식된 패턴(P_{right})들에 대해서만 실험을 반복한 경우 나머지 클래스의 인식률이다. 그 결과를 표 5에 나타내었다.

2차 주의집중으로 얻어진 인식률의 상승은 6.7%이다.

표 5. 단일 인식기와 주의집중 네트워크의 인식을 비교

Table 5. Comparison of recognition rates between SVM only and SA network.

인식대상	SVM only	with SA
for all	42.7%	49.4%
for (P_{right})	51.2%	59.3%

1차 주의집중이 36.6% 향상된데 비해 2차 주의집중에서 인식률의 상승이 상대적으로 낮은 원인을 다음과 같이 짚어볼 수 있다. 먼저, 중첩된 부분이 많아질수록 2차 주의집중문제의 난이도가 높아진다는 데에 있다. 두 클래스의 중첩된 정보손실을 막기 위해 특징 벡터를 사용하였지만 이 방법은 올바르게 인식하기를 원하는 패턴에만 적용되는 것이 아니라, 전체 패턴에 대해 적용되는 문제이기 때문에 중첩이 많은 패턴이 나타나면 다른 어떠한 클래스로도 주의집중이 일어날 수 있다는 것을 의미한다. 또한 입력패턴의 글씨체가 너무 두껍다는 것을 지적할 수 있다. 이는 중첩되는 정보의 비율이 높아진다는 것을 의미하여, 예를 들어 두개의 숫자를 겹쳐서 쓸 때 매직으로 썼을 경우 얇은 볼펜으로 썼을 경우에 비하여 겹치는 부분에 대한 비중이 높아진다는 것을 생각해볼 수 있다. 정보를 너무 많이 포함하고 있는 두 패턴이 중첩이 되면, 1차 주의집중에서 선택되지 않은 나머지 클래스의 특징벡터조차 처음 클래스의 부분 집합이 될 가능성이 높아지기 때문에 낮은 인식을 초래할 수 있다.

V. 결 론

본 연구에서는 SVM와 선택적 주의집중을 접목시키는 새로운 시도를 하였다. 주의집중 학습이 멈추는 시점을 결정하기 위해 입력패턴이 초기상태일 때 주의집중 클래스의 영역에서 가장 가깝게 존재하는 지지벡터를 기준으로 삼고, 그 벡터와 가장 가까워지는 시점에서 학습을 멈추는 제약조건을 정의하였다. 일반적인 주의집중을 그대로 사용하기에는 주의집중 척도를 정의하는데 문제가 되므로, 제한된 주의집중을 제안하여 새로운 주의집중 척도를 제시하였다. 이는 주의집중 클래스에 속해있던 패턴에는 영향이 없으나 속해있지 않은 패턴들에 대해서는 강력한 패널티를 적용함으로써 좀 더 효과적인 패턴 분리가 가능하다.

제안된 알고리즘은 USPS DB로 제작한 중첩패턴을 이용하여 그 인식 성능을 확인하였다. 단일 패턴의 경

우 주의집중 이전의 정보를 사용하는 것이 좀 더 효과적이거나 그 차이가 미미하고, 중첩 패턴의 경우에는 제한된 주의집중 하에서 새롭게 제안한 척도가 매우 효과적임을 증명하였다. 1차 주의집중 실험에서는 인간의 눈으로도 구분하기 어려운 중첩 패턴에 대해 인식 실험을 시도하여 50% 미만의 인식률을 80% 이상까지 상승시켰다. 또한 주의집중 천이와 2차 주의집중 부분을 이용하여 주의집중이 없을 경우에 비하여 향상된 인식률을 얻어내었다. 중첩된 패턴을 분리할 때 1차 주의집중이 종료된 후 2차 주의집중을 위해 천이를 거치는 과정에서 이미 인식된 클래스의 정보만을 제거하고 나머지 클래스의 정보를 효과적으로 살리는 방법을 찾는 것이 추후 연구 과제로 제안될 수 있다.

본 연구에서는 SVM과 선택적 주의집중을 접목시키기 위한 새로운 정지점 척도와 주의집중 척도를 제안함으로써 복잡한 중첩 패턴들에 대해 인식률의 높은 향상을 가져올 수 있다는 가능성을 제시하였다.

참 고 문 헌

- [1] C. Bishop. Neural Networks for Pattern Recognition. Clarendon Press, London, 1995.
- [2] Vladimir N. Vapnik, "The Nature of Statistical Learning Theory", Wiley, N.Y., pp. 131-170, 1995.
- [3] Christopher J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, vol2, pp. 121-167, Kluwer Academic Publishers, Boston, 1998.
- [4] Corinna Cortes and Vladimir Vapnik, Support-Vector Networks, Machine Learning, vol 20, No 3, p273-297, 1995.
- [5] E. Cherry. Some experiments on the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America, 25:975-979, 1953.
- [6] A. Treisman. Monitoring and storage of irrelevant messages in selective attention. Journal of Verbal Learning and Verbal Behavior, 3:449-459, 1964.
- [7] K. Fukushima. "Neural network model for selective attention in visual pattern recognition and associative recall", Applied Optics, 26[23], pp. 4985-4992 Dec. 1987.
- [8] B. Schölkopf, K.-K. Sung, C. J. C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik. Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function

Classifiers. IEEE Transactions on Signal Processing, 45:2758-2765, 1997.

[9] J. Weston and C. Watkins. Multi-class support vector machines. Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, Egham, TW20 0EX, UK, 1998.

[10] M. Posner. Attention in cognitive neuroscience: An overview. In M. Gazzaniga et al., editor, The Cognitive Neurosciences, number V, pp. 615-624,

MIT Press, 1966.

[11] M. Posner and M. Raichle. Images of Mind, Scientific American Library, New York, 1994.

[12] Ki-Young Park and Soo-Young Lee, Selective Attention for Robust Speech Recognition in Noisy Environments, International Joint Conference on Neural Networks, Washington, USA, July, 1999.

저 자 소 개



배 규 찬(학생회원)
 2001년 2월 한국과학기술원 전자 전산학과 졸업(공학사)
 2005년 2월 한국과학기술원 전자 전산학과 졸업(공학석사)
 <주관심분야 : 패턴인식, DSP control 등>



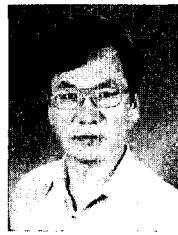
최 용 선(학생회원)
 2000년 2월 한국과학기술원 전자 전산학과 졸업 (공학사)
 2002년 2월 한국과학기술원 전자 전산학과 졸업 (공학석사)
 2002년 3월~현재 한국과학기술원 바이오시스템학과 박사과정 재학중

2002년 7월~2003년 6월 미국 University of Connecticut Health Center 방문 연구원
 <주관심분야 음성인식, 신호처리, 청각모델등>



박 형 민(정회원)
 1997년 2월 한국과학기술원 전기 및 전자공학과 학사
 1999년 2월 한국과학기술원 전기 및 전자공학과 석사
 2003년 8월 한국과학기술원 전자 전산학과 박사

2003년 9월~2005년 2월 한국과학기술원 바이오 시스템학과 박사후 연수과정
 2005년 3월~현재 미국 카네기멜론대학교 Electrical and Computer Engineering학과 박사후 연수과정
 <주관심분야: ICA, 음성인식, 음성향상 등>



이 수 영(정회원)
 1975년 서울대학교 전자공학과 졸업(공학사)
 1977년 한국과학기술원 전기전자 공학과 졸업(공학석사)
 1984년 Polytechnic Institute of New York 졸업(공학박사)

1977년~1980년 대한엔지니어링(주) (대리)
 1982년~1985년 미국 General Physics Corp. (Staff / Senior Scientist)
 1986년~2002년 한국과학기술원 전자전산학과 전기 및 전자공학전공 교수
 2002년~현재 한국과학기술원 바이오시스템학과 교수
 1998년~현재 뇌과학연구센터 소장
 <주관심분야: 음성인식, 신경회로망, 인공시스템, 청각모델 등>



오 상 훈(정회원)
 1986년 2월 부산대학교 전자공학과 학사
 1988년 2월 부산대학교 전자공학과 석사
 1999년 8월 한국과학기술원 전기 및 전자공학과 박사
 1988년 1월~1989년 12월 금성반도체(주) 사원
 1990년 1월~1998년 6월 ETRI 선임연구원
 2000년 4월~2000년 10월 Brain Science Institute, RIKEN, Research Scientist

2000년10월~2001년 10월 (주)엑스텔테크놀로지 연구소장
 2001년11월~2002년 2월 KAIST 초빙교수
 2002년 3월~현재 목원대학교 정보통신공학부 조교수
 <주관심분야: ICA, BSS, 지능적 신호처리, 패턴인식 등>