

다중 분류기 통합을 위한 퍼지 행위지식 공간 Fuzzy Behavior Knowledge Space for Integration of Multiple Classifiers

김봉근[†], 최형일^{††}
Bongkeun Kim, Hyungill Choi

요 약

본 논문에서는 다중 분류기의 통합을 위해 퍼지 행위지식 공간을 구성하고 이를 이용하는 방법을 제안한다. 기존의 행위지식 공간은 각 분류기들이 서로 독립적인 필요가 없고 적응적 학습이 가능한 것으로 단지 하나의 클래스 레이블만을 출력하는 분류기들의 통합에 가장 최적의 방법으로 알려졌다. 그러나 행위지식 공간은 각 분류기가 출력하는 클래스 레이블에 대한 측정값과 경험적 지식을 통합과정에 반영하기 어렵다는 문제점을 갖고 있다. 이러한 행위지식 공간의 문제점을 해결하기 위해 본 논문에서는 퍼지개념을 이용한 퍼지 행위지식 공간을 정의하고 이를 다중 분류기의 통합에 적용하기 위한 방법을 기술한다. 또한, 퍼지 행위지식 공간의 유용성을 증명하기 위해 각 분류기로부터 얻어진 클래스 레이블들과 이에 관련된 측정값을 포함하는 분류결과들의 통합에 적용된 실험결과를 기술한다.

주제어 인공지능, 패턴인식, 퍼지이론

† 국립 청주전문대학 전자계산과 조교수
Department of Computer Science,
Chong-Ju National Junior College
213-1, Sachang-dong, Chongju,
Chungbuk, Korea, 361-280
e-mail: bkkim@soback.kornet.nm.kr

†† 송실대학교 대학원 전자계산과 부교수
Department of Computer Science,
Soongsil University
1-1, Sangdo 5-dong, Kwanak,
Seoul, Korea, 156-743
e-mail: hic@stone.soongsil.ac.kr

*) 본 논문은 1994년도 대학부설연구소 연구비지원(학술진흥재단)에 의한 결과임.

ABSTRACT

In this paper, we suggest the "Fuzzy Behavior Knowledge Space(FBKS)" and explain how to utilize the FBKS when aggregating decisions of individual classifiers. The concept of "Behavior Knowledge Space(BKS)" is known to be the best method in the context that each classifier offers only one class label as its decision. However, the BKS does not consider measurement value of class label. Furthermore, it does not allow the heuristic knowledge of human experts to be embedded when combining multiple decisions. The FBKS eliminates such drawbacks of the BKS by adapting the fuzzy concepts. Our method applies to the classification results that contain both class labels and associated measurement values. Experimental results confirm that the FBKS could be a very promising tool in pattern recognition areas.

Keyword Behavior knowledge space, classifier, combination of multiple classifiers, adaptive learning

1. 서론

패턴인식과 같은 분야에서 높은 성능을 가진 분류기의 설계는 매우 중요하며 이를 위해 많은 연구자들이 약 20년 이상 지속적인 연구를 하고 있다. 이러한 결과로 최근 비교적 높은 성능을 갖는 분류기가 많이 개발되었다. 그러나 이들 중 어느 것도 단일 분류기만으로는 만족할 만한 결과를 가져다 주지 못하고 있다. 이러한 문제를 해결하기 위한 방법으로 최근에 다중전문가 통합(CME : Combination of Multiple Experts)이라고 불리는 새로운 경향이 나타나게 되었다. 이것은 서로 다른 방법론 또는 특징

을 갖는 분류기들이 서로 상호보완적으로 작용하도록 함으로써 분류기의 성능을 향상시키고자 하는 것이다.

CME의 개념에 근거한 분류기 통합방법의 대표적인 것에는 행위지식 공간(BKS : Behavior Knowledge Space)을 이용한 방법 [1][2], 다득표(majority voting) 방법, 보더 카운트(Borda count) 방법, 논리적 회귀(logistic regression) 및 클래스 집합 축소 및 재배열(class set reduction and reordering)을 이용한 방법 [3], 선형 확신도 누적(LCA : Linear Confidence Accumulation) 방법 [4], 퍼지융합(fuzzy fusion) 방법 [5] 등이 있다.

이들 통합 방법들은 통합에 사용될 분류기가 출력하는 결과의 유형에 따라 크게 세가지로 분류될 수 있다. 첫째는 측정값을 갖지않은 하나의 클래스 레이블(class label)만을 출력하는 것(추상화 레벨 정보)이고, 두번째는 측정값을 갖지 않은 클래스 레이블들의 순서화된 리스트(ordred list)를 출력하는 것(순위레벨 정보)이며, 세번째는 확신도(confidence vlaue) 또는 측정값을 가진 클래스 레이블들의 순서화된 리스트를 출력하는 것(측정값 레벨 정보)이다. 첫번째 유형의 통합 방법에는 행위지식 공간, 다득표, 보더 카운트와 같은 방법들이 있고, 두번째 유형을 위한 통합 방법에는 논리적 회귀 및 클래스 집합 축소 및 재배열을 이용한 방법이 있다. 첫번째와 두번째 유형을 위한 통합 방법들은 분류기들의 결과를 통합하기 쉽고 단순하다는 장점을 갖는다. 그러나 분류기가 출력하는 정보는 단지 클래스 레이블 뿐이므로 세번째 유형의 분류기를 통합하기 위한 방법에 비해 근본적으로 최적의 결정(optimal solution)을 유도하기 어렵다는 데에 문제가 있다(4).

세번째 유형의 분류기를 통합하기 위한 방법으로 선형 확신도 누적 방법과 퍼지융합을 이용한 방법이 제안되었다. 선형 확신도 누적 방법은 통합을 위해 각 측정값을 확신도로 변환하기 위한 변환함수를 제안하고, 각 클래스에 대한 누적 확신도들 중에서 최대의 누적 확신도를 갖는 클래스로 최종 결정을 내리고 있다. 그러나 이 방법은 분류기가 다른 분류기들과 독

립적이어야 하고, 누적된 확신도는 개별 확신도의 선형적인 합(linear summation)으로 정의된다는 가정하에서 성립되고 있다(4). 퍼지융합을 이용한 방법은 통합을 위한 경험적 지식(heuristic knowledge)을 기존의 수학적인 모델링에 기초한 방법으로는 통합 과정에 반영하기 어렵다는 문제를 제기하고 이를 해결하기 위해 퍼지이론을 이용한 방법을 제시하고 있다. 이 방법에서는 확신도의 범위를 정해진 수의 순위(level)로, 확신도 차이를 정해진 수의 간격(interval)으로 정규화하고 이들로부터 각 분류기에 대한 순위-간격 귀속함수를 정의하며, 각 분류기에 대한 퍼지관계(fuzzy relation)을 고려하여 최종적인 결정을 내리는데 이용하고 있다(5).

위의 방법들 중 BKS 방법은 단지 하나의 클래스 레이블만을 출력하는 분류기들의 통합에 가장 최적의 방법으로 알려져 있다. 이 방법은 분류기가 다른 분류기들과 반드시 독립적일 필요가 없고, 단순한 구조를 갖고 있지만 분류기의 행위지식을 비교적 효과적으로 표현하고 있으며, 적응적 학습(adaptive learning)이 가능하다는 것 이외에도 많은 장점을 가진 방법으로 생각된다. 또한, 선형 확신도 누적 방법은 측정값 레벨 정보가 분류기 통합 시스템의 성능을 향상시키는 데 매우 중요한 역할을 수행한다는 것을 주장하고 있으며, 퍼지융합을 이용한 방법에서는 확률개념에 기초한 방법의 문제점을 지적하고 이를 해결하기 위해 퍼지개념을 이용하여 통합과정에 경험적 지식을 반영하고자 했다.

본 논문에서는 BKS 방법의 장점을 유지하면서도 세번째 유형의 분류기들을 통합하는 경우에 효과적으로 사용될 수 있는 퍼지 행위지식 공간(FBKS : Fuzzy BKS) 방법을 제안한다. FBKS 방법은 크게 두가지 기능을 갖는다고 할 수 있다. 하나는 입력된 샘플 패턴에 대해 각 분류기가 출력하는 결과로부터 FBKS의 내용 즉, 행위지식을 갱신(update)하는 것이고, 다른 하나는 FBKS내의 행위지식을 이용하여 최적의 결정(optimal solution)을 내리는 것이다. BKS방법에서는 행위지식의 갱신과정에 단지 하나의 상황만이 고려되는 것에 비해 FBKS에서는 하나 이상의 상황들을 고려하여 갱신과정을 수행하게 된다. 또한, 최적의 결정과정에서도 역시 여러개의 상황들이 결정과정에 참여하게 된다. 따라서 선택된 다수의 상황들로부터 최적의 결정을 유도하기 위해 FBKS는 측정값들에 대한 정규화 과정을 통해 그 크기와 의미를 일치시키고 결정과정에 인간 전문가의 경험적 지식을 반영함으로써 보다 정확한 결정을 유도하고 있다.

본 논문의 2장에서는 먼저 BKS와 FBKS를 개념적으로 비교하고, 3장에서는 FBKS의 개념과 구조, 그리고 적응적 학습과정과 결정규칙에 관해 기술한다. 그리고 4장에서는 제안된 퍼지 행위지식 공간을 이용한 분류기 통합 방법의 실험결과를 보인다.

II. BKS와 FBKS

이 장에서는 먼저 BKS에 대한 간단한 예[2]를 통해 BKS의 기본 개념과 단점, 그리고 BKS에 퍼지개념을 추가함으로써 FBKS를 제안하게된 동기를 설명한다. <표1>은 BKS방법을 설명하기 위한 한 예를 보이고 있다.

<표 1> 주어진 400문제에 대한 A와 B의 누적된 결과

A \ B	yes	no
yes	(90, 10)	(40, 10)
no	(60, 20)	(10, 160)

퀴즈대회에 참석한 두 사람 A와 B는 주어진 400개의 문제에 "yes" 또는 "no"로만 답할 수 있다고 하자. 그리고 400개의 문제 중 200개는 정답이 "yes"이고 나머지 200문제는 정답이 "no"이지만 A와 B는 정답이 "yes"인 문제의 수와 "no"인 문제의 수를 알지 못한다고 가정할 때, <표 1>은 주어진 400개의 문제에 대해 A와 B의 답을 누적한 결과를 보이고 있다. <표 1>에서는 A와 B의 대답에 따라 4가지 상황(situation)이 존재하는 것을 보여 주고 있다. 첫째는 A와 B가 모두 "yes"로 답한 경우이고, 둘째는 A는 "yes"로 답하고 B는 "no"로 답한 경우이며, 셋째는 A는 "no"로 답하고 B는

“yes”로 답한 경우, 그리고 A와 B가 모두 “no”로 답한 경우이다. <표1>에서 결과를 누적하는 방법은 먼저 A와 B의 답에 따라 네가지 상황중 하나를 선택하고 선택된 상황에 대해 값을 누적하는 것이다. 예를 들어, (40, 10)은 두번째 상황에 대한 누적결과로서 40은 주어진 문제의 정답이 “yes”인 문제의 수를 나타내며, 10은 주어진 문제의 정답이 “no”인 문제의 수를 나타내고 있다. 나머지 세가지 상황에 대해서도 같은 방법으로 결과를 누적한 것이다.

그러면 <표 1>에서 A는 첫번째 상황에서 10개, 두번째 상황에서 10개, 세번째 상황에서 60개, 네번째 상황에서 10개로서 총 90개 문제의 정답을 맞추지 못하였고 마찬가지로 B는 80개 문제의 정답을 맞추지 못한 것을 알 수 있다. 그러나 만일 A와 B의 행위에 대한 지식을 통합한다면 오답은 줄어들 것이다. 예를 들어, 주어진 어떤 문제에 대해 A는 “yes”로 답하고 B는 “no”로 답한 경우에는 <표1>의 두번째 상황 (40, 10)을 선택하여 상황내의 정보에 나타난 것과 같이 주어진 문제의 정답이 “yes”일 확률이 80%이고 “no”일 확률이 20%이므로 “yes”라고 답하는 것이 타당할 것이다. 따라서 A와 B의 결과를 통합한 경우에 발생할 수 있는 오답의 수는 첫번째 상황에서 10개, 두번째 상황에서 10개, 세번째 상황에서 20개, 네번째 상황에서 10개로서 총 50개이므로 전체적인 오답의 수를 줄일 수 있다는 것을 알 수 있다[2]. 이러한 단순하고 이해하기 쉬운 개념을 이용하여 다중 분류기 통합을 수행하고자 하는 것이

BKS이다.

그러나 앞의 예에서 A와 B는 단순히 “yes”로 답할 가능성과 “no”로 답할 가능성 중에서 가장 가능성이 높다고 생각하는 것으로 답한 것이다. 즉, A 또는 B가 주어진 문제에 대한 답이 “yes”일 가능성과 “no”일 가능성에 대한 확신이 없는 경우에도 확신이 없는 최종적인 답이 <표 1>을 구성하는데 기여하게 된다. 따라서 BKS 방법에서는 각 전문가 또는 분류기의 결과에 대한 확신도를 무시함으로써 분류기의 행위지식을 보다 정확히 표현하지는 못하고 있는 것으로 생각된다. 또한, BKS방법은 기본적으로 최종결정을 유도하기 위해 확률개념을 이용하고 있다. 그러나 확률개념만으로는 통합과정에 전문가의 경험적 지식을 효과적으로 반영하기 어렵다[5]. 따라서 본 논문에서는 측정값 레벨 정보를 출력하는 분류기들의 통합을 위해 BKS의 기본 개념을 이용하면서 분류기의 행위지식을 보다 정확히 표현하고 결정과정에 전문가의 경험적 지식을 효과적으로 반영하기 위해 퍼지개념을 도입한 FBKS 방법을 제안한다.

III. FBKS 방법

FBKS는 각 분류기로부터 출력된 결과를 이용하여 각 분류기의 행위지식을 효과적으로 표현함으로써 다중 분류기 통합 시스템이 FBKS를 이용하여 최적의 결정을 내릴 수 있도록 지원하는 데에 그 목적이 있다. 따라서 FBKS 방

$$L^{(k)} = \{ t, a, c; (c_1^{(k)}, v_1^{(k)}), (c_2^{(k)}, v_2^{(k)}), \dots, (c_m^{(k)}, v_m^{(k)}) \} \quad (1)$$

where,

t : type of measurement values(similarity or dissimilarity)

a, c : range [a, c] of measurement values

c_i : class label i

v_i : measurement value for class label i

법은 BKS방법에서와 마찬가지로 크게 두가지 기능을 갖는다고 할 수 있다. 하나는 입력된 샘플 패턴에 대해 각 분류기가 출력하는 결과로부터 FBKS의 내용 즉, 행위지식을 갱신(update)하는 것이고(학습단계), 다른 하나는 FBKS내의 행위지식을 이용하여 최적의 결정(optimal solution)을 내리는 것(결정단계)이다. FBKS를 구성하는 모든 상황은 주어진 샘플패턴에 대한 각 분류기의 결과에 의해 적응적으로 학습되며, 알려지지 않은 분류문제에 대한 각 분류기들의 결과는 FBKS를 이용한 결정과정을 통해 최적의 결정으로 유도된다..

3.1 측정값 레벨 정보와 FBKS

FBKS 방법은 분류기가 클래스 레이블들의

순서화된 리스트 뿐만 아니라 각 클래스 레이블에 대한 측정값을 포함하고 있는 경우에 적용될 것이다. 이것은 측정값 레벨 정보라고 부르며, 우리는 분류기 k 에 의해 출력되는 측정값 레벨 정보 $L^{(k)}$ 가 (식1)과 같은 형식을 갖는다고 정의한다.

측정값은 분류기의 성격에 따라 거리(distance), 유사성(similarity) 또는 확신도(confidence value)와 같은 유형이 있다. 일반적으로 유사성과 확신도값의 범위는 0와 1사이의 값이 되지만, 거리와 같은 경우에는 분류기가 출력하는 측정값의 크기에 따라 여러 가지가 있을 수 있다. 예를 들면, [0, 2], [1, 5], [3, 20]과 같은 것들이 있을 수 있다. 따라서 서로 다른 유형의 측정값들을 통합에 사용하기 위해서는 이들 측정값들에 일관된 의미를 부여하여야만 한다. 이를 위해서는 정규화(normalization)가 가장 널리 사용되고 있는 방법이며 FBKS를 구성하기 전에 $L^{(k)}$ 로부터 $L'^{(k)}$ 로의 정규화과정이 선행되어야만 한다.

(식 1)에서 t 는 측정값의 유형을 나타내고, a 와 c 는 측정값의 최소-최대 범위를 나타내며, t

는 각 분류기가 출력 가능한 클래스 레이블의

FBKS가 어떻게 구성되고 있는지 그 예를 보여 주고 있다.

$$FBKS = \{ E, C, S \} \quad (2)$$

$$\text{where, } E = \{ e_1, e_2, \dots, e_M \}$$

$$C = \{ c_1, c_2, \dots, c_m \}$$

$$S = \{ s_{(q, \dots, r)} \mid 1 \leq q, \dots, r \leq m \}$$

집합을 의미한다. S는 FBKS내에 존재하는 모든 상황의 집합을 의미하는데, (q, \dots, r) 은 어떤 상황 s의 색인(index)이 된다. 색인에서 q는 첫번째 분류기가 출력한 클래스 레이블, r은 M번째 분류기가 출력한 클래스 레이블을 의미한다. (그림 2)는 두개의 분류기 e_1, e_2 에 의해

(그림 2)에서는 각 분류기 e_1, e_2 가 내린 결정에 의해 m^2 개의 상황이 존재하고 있다는 것을 알 수 있다. 만일 클래스의 수 m 이 10이라면 e_1 과 e_2 가 출력할 수 있는 가능한 클래스의 수는 10이다. 따라서 (그림 2)와 같은 경우에는 모두 100개의 상황이 존재하게 된다. 상황은 각 분류기들의 결과로부터 통합에 필요한 정보를 제공하기 위한 목적을 가진다. 필요한 정보란 어떤 상황에서 참클래스일 가능성이 가장 큰 클래스는 무엇이고 그 결정에 대해 얼마만큼 신뢰할 수 있는가 하는 것이다.

BKS 방법에서는 측정값을 갖지 않는 단 하나의 클래스 레이블만을 출력하는 분류기를 통합하기 위한 것이므로, 상황에 들어오는 샘플패턴의 갯수를 누적함으로써 이러한 정보를 표현

$e_2 \backslash e_1$	c_1	c_2	...	c_i	...	c_m
c_1	$S(1,1)$	$S(1,2)$...	$S(1,i)$...	$S(1,m)$
c_2	$S(2,1)$	$S(2,2)$...	$S(2,i)$...	$S(2,m)$
\vdots
c_i	$S(i,1)$	$S(i,2)$...	$S(i,i)$...	$S(i,m)$
\vdots
c_m	$S(m,1)$	$S(m,2)$...	$S(m,i)$...	$S(m,m)$

(그림 2) 두개의 분류기 e_1, e_2 에 대한 FBKS의 구조

하고 있다. 예를 들어, 참클래스가 2인 샘플패턴 α 에 대해, 분류기 A와 B가 각각 클래스 2와 클래스 3이라는 서로 상이한 결과를 출력했다면, BKS로 부터 (식 3)과 같이 표현되는 상황 $s_{(2,3)}$ 을 선택한 후, 상황 $s_{(2,3)}$ 의 클래스 2에 해당하는 누적수를 1증가시킨다. 만일 충분히 많은 샘플패턴이 BKS에 입력되어 $s_{(2,3)} = \{246, 84\}$ 가 되었다면 이 상황에서 참클래스로 가장 유력한 것은 누적된 수가 가장 큰 클래스 2가 될 것이라는 것이다. 따라서 만일 알려지지 않은 패턴 α 에 대해 분류기 A와 B가 각각 클래스 2, 클래스 3으로 결과를 출력했다면, 선택된 상황 $s_{(2,3)}$ 에서 α 는 클래스 2에 속한다고 최종적인 결정을 내리게 된다[2].

그러나 FBKS 방법에서는 주어진 분류문제에 대해, 각 분류기들은 각 클래스에 대한 측정값의 리스트를 결과로 출력하는 것으로 가정하므로 이들로 부터 의미있는 클래스들의 조합에 의해 다수의 상황이 선택될 수 있다. 따라서 상황의 표현은 다수의 상황에 포함된 정보들

로부터 최종적으로 하나의 결정을 유도하기 용이하도록 표현되어야 한다. 또한, 각 클래스 레이블에 대한 측정값은 FBKS를 구성하는데 반영되어 최적의 결정을 내리는 데 도움이 될 수 있어야 한다.

(식 3)은 이러한 요구사항을 만족시키기 위해 FBKS의 상황이 어떻게 표현되고 있는지를 보여주고 있다. B 는 $s(q, \dots, r)$ 내에 표현되어 있는 (N_i, P_i) , $i = q, \dots, r$ 의 쌍으로부터 얻어질 수 있고, R 은 B 에 대한 신뢰도를 의미하고 있다.

3.2 적응적 학습

학습은 대부분의 분류기에 대해 요구되는 매우 중요한 기능이며, 적응적 학습이란 분류기의 행위지식을 주어진 특정 분류문제에 적응하도록 지속적인 갱신을 수행하는 것을 말한다. 적응적 학습과정은 크게 두 단계를 거치게 되

$$s_{(q, \dots, r)} = \{ B, R; (N_q, P_q), \dots, (N_r, P_r) \} \tag{3}$$

where, B : best class on $s_{(q, \dots, r)}$

R : reliability for best class

N_i : accumulated number of incoming sample for class i

P_i : possibility that class i to be true class.

는데, 첫번째 단계는 FBKS로 부터 적절한 상황을 선택하는 것이고, 두번째 단계는 선택된 상황들을 갱신하는 것이다.

주어진 입력 패턴을 x 라 하자. 그러면 M 개의 분류기들은 x 에 대해 (식 1)과 같은 $L^{(1)}, L^{(2)}, \dots, L^{(M)}$ 들을 출력할 것이다. 각 분류 결과에는 클래스 레이블과 측정값의 쌍이 m 개가 존재하며, 각 측정값들은 정규화를 거쳐 정규화된 확신도로 변환된다. 정규화된 결과 리스트를 $L^{(1)}, L^{(2)}, \dots, L^{(M)}$ 라 하자. 그러면 각 분류결과내의 클래스 레이블들에 대한 가능한 모든 조합은 FBKS내의 어떤 상황으로 사상(mapping)될 수 있다. 예를 들어, M 이 2이고, m 이 10인 경우 분류기 A와 B의 분류결과들이 정규화 결과로부터 정규화된 확신도가 임계치 λ 보다 큰 것이 (식 4)와 같이 되었다면, 총 100개의 상황이 존재하는 FBKS로 부터 $s_{(2,5)}, s_{(2,7)}, s_{(4,5)}, s_{(4,7)}, s_{(7,5)}, s_{(7,7)}$ 6개 상황이 선택될 수 있다.

따라서 상황을 선택하기 위한 규칙은 (식 5)와 같이 표현될 수 있다.

임계치 λ ($0 < \lambda < 1$)는 선택될 상황의 수를 조절하기 위해 사용되는 파라미터로서 정규화된 확신도가 λ 보다 큰 것만을 상황선택에 고려한다는 의미를 가진다. 만일 λ 가 작게 결정되면, 선택될 상황의 수는 많아지게 되어 FBKS를 갱신하는 과정에 보다 많은 클래스들이 참여하게 됨으로써 비교적 정확한 지식의 표현이 가능하지만 계산상의 낭비가 심하게 된다는 문제가 있다. 반대로 λ 가 크게 결정되면, 선택될 상황의 수는 작아지게 되어 계산 시간을 절약할 수는 있지만 정확한 상황의 표현이 어렵게 된다.

FBKS내의 상황들이 선택되고 나면, 선택된 상황들에 대한 갱신이 요구되는데, 상황 $s_{(q, \dots, r)}$ 에 대한 갱신과정은 다음과 같이 설명될 수 있다.

$$L^{(1)} = \{(2, 0.2), (4, 0.1), (7, 0.7)\} \tag{4}$$

$$L^{(2)} = \{(5, 0.6), (7, 0.4)\}$$

$$\begin{cases} \text{select } s_{(q, \dots, r)} & , \text{ if } (c_q^{(1)} \in L^{(1)}) \text{ and } \dots \text{ and } (c_r^{(M)} \in L^{(M)}) \\ & \text{and } \forall v_i^{(k)} \geq \lambda, i = q, \dots, r \\ \text{ignore} & , \text{ otherwise} \end{cases} \tag{5}$$

$$n = N_c, \quad m = N_c + 1, \quad , \text{ if } c = y \quad (6)$$

$$N_c = N_c + 1, \quad , \text{ if } c = y \quad (7)$$

$$P_c = P_c + \left(\frac{n}{m} \cdot P_c + \frac{1}{m} \cdot v_c^{(k)} \right), \text{ if } c = y \quad (8)$$

where, c : class label contained in situation $s_{(q, \dots, r)}$
 y : true class which contains input pattern x

$$B = \{ j \mid V_j = \max_{i=q, \dots, r} \{ V_i \} \} \quad (9)$$

$$V_i = \alpha \cdot N_i \cdot P_i \quad (9-1)$$

where, α is number of class i in $s_{(q, \dots, r)}$

$$R = \xi \cdot T + \omega \cdot I \quad (10)$$

$$\xi + \omega = 1 \quad (10-1)$$

where, ξ and ω are some thresholding values

$$T = f\left(\sum_{i=q}^r N_i\right) \quad (10-2)$$

$$f(x) = \begin{cases} 2 \left(\frac{x}{t}\right)^2 & , \text{ if } 0 \leq x \leq \frac{t}{2} \\ 1 - 2 \left(\frac{x-t}{t}\right)^2 & , \text{ if } \frac{t}{2} \leq x \leq t \end{cases} \quad (10-3)$$

where, t is total number of sample patterns

$$I = \frac{V_B - \max_{i=q, \dots, r, i \neq B} \{ V_i \}}{V_B} \quad (10-4)$$

상황 $s(q, \dots, r)$ 에 대한 갱신은 q 에서 r 까지의 클래스들 중에서 입력 패턴 x 의 참클래스 y 와 같은 클래스에 대해서만 일어나게 된다. (식 6)과 (식 7) 그리고 (식 8)이 적용되고 난 후, 선택된 상황 $s(q, \dots, r)$ 에 포함된 B 와 R 은 (식

9)와 (식 10)을 이용하여 갱신된다.

(식 10)에서는 정규화된 확신도의 누적값들 중에서 최대값(maximum value)과 이차-최대값(second-maximum value)사이의 차이를 이용하여 어떤 상황 $s(q, \dots, r)$ 에서 결정된 B 의 신

뢰도를 표현하고 있다. 만일 차이가 크다면 B 의 신뢰도는 높고, 차이가 작다면 B 의 신뢰도는 낮다고 할 수 있다. 또한, $s(q, \dots, r)$ 내에 들어온 샘플패턴의 수가 매우 작다면, 결정된 B 의 신뢰도는 작아져야 하고, 만일 샘플패턴의 수가 충분히 많다면 결정된 B 의 신뢰도는 높아져야 한다.

R 은 다중 분류기 결과의 통합에 결정적인 역할을 담당하게 되는데, 효과적인 파라미터 R 을 구하기 위해서는 인간 전문가의 지식이 필요하다. 인간 전문가의 지식을 확률개념으로 정확히 표현하는 것은 어려운 일이지만 퍼지이론은 이러한 전문가의 지식을 용이하게 표현한다.

3.3 최적의 결정

알려지지 않은 입력 데이터 x 에 대해 각 분류기로부터 생성된 분류결과들이 정규화과정을 거쳐 FBKS에 입력되면, 입력된 분류결과들로부터 최적의 결정을 유도하는 과정은 학습과정과 마찬가지로 크게 두 가지 단계로 구분할 수 있다. 첫번째 단계는 적절한 상황을 FBKS로부터 선택하는 것이고, 두번째 단계는 선택된 상황들과 입력된 분류결과를 이용하여 최종적인 결정을 내리는 것이다. 상황의 선택 방법은 (식 11)과 같다.

$$\left\{ \begin{array}{ll} \text{select } s_{(q, \dots, r)} & , \text{ if } (c_q^{(1)} \in L'^{(1)}) \text{ and } \dots \text{ and } (c_r^{(M)} \in L'^{(M)}) \\ & \text{and } \forall v_i^{(k)} \geq \eta, \quad (\lambda \leq \eta < 1) \\ \text{ignore} & , \text{ otherwise} \end{array} \right. \quad (11)$$

따라서 우리는 귀속함수를 (식 10-2)와 (식 10-4)로 정의함으로써 R 을 구할 수 있다. 이러한 상황의 갱신과정은 매우 단순하며, 따라서 계산상의 효율성도 보장된다고 할 수 있다. 우리는 이러한 갱신과정이 학습에 사용되는 샘플의 수가 증가할수록 상황을 보다 정확히 묘사할 수 있을 것으로 기대한다.

(식 11)은 (식 9)와 거의 동일한 방법으로 수행된다. 그러나 임계치 η 의 값에 따라 선택될 상황의 수가 달라질 수 있으므로, η 는 통합 결과에 많은 영향을 주게 된다.

선택된 상황들의 집합 S 로부터 최종적인 결정을 내리기 위해 우리는 먼저 (식 12)와 같은 $rList$ 를 정의한다.

$$rList = \{(r_1, a_1), (r_2, a_2), \dots, (r_m, a_m)\} \quad (12)$$

$rList$ 는 각 클래스에 대한 누적 신뢰도(r)와 입력된 클래스 레이블에 대한 정규화된 확신도(a)를 나타내며, $rList$ 는 선택된 모든 상황 S 와 정규화된 분류결과 $L^{(1)}, L^{(2)}, \dots, L^{(M)}$ 에 의해 결정되는데, 누적 신뢰도는 각 상황에 정의되어 있는 B 와 R 에 의해 결정된다. (식 13)은 $rList$ 의 i 번째 원소가 어떻게 결정되는지를 표현하고 있다.

$$r_i = r_i + \sum_{s_{(a, \dots, r)} \in \tilde{S}} \{R\}$$

where, $R \in S_{(a, \dots, r)}$ and $i = (B \in S_{(a, \dots, r)})$

(13)

$$a_i = \left(\sum_{k=1}^M v_i^{(k)} \right) / M$$

누적 신뢰도(r_i)는 FBKS로부터 얻어진 정보이고, 정규화된 확신도(a_i)는 각 분류기들의 동의(agreement) 정도를 의미하고 있다. 따라서 선택된 모든 상황들로부터 최적의 결정 $d(x)$ 를 내리기 위한 결정규칙은 (식 14)와 같이 표현될 수 있다.

$$d(x) = \{j \mid V_j = \max_{i \in rList} \{r_i \cdot a_i\}\}$$

(14)

결국, 선택된 모든 상황에 대해 $rList$ 가 형성되면, 이 $rList$ 로 부터 최대값 V_j 를 갖는 클래스 j 를 입력 데이터 x 에 대한 최종 결정으로 한다.

IV. 실험 및 결과

제안된 FBKS방법에 대한 실험은 필기체 숫

자 인식을 위해 고안된 세가지 분류기 즉, Euclidian거리 분류기와 Mahalanobis거리 분류기, 그리고 역전파 알고리즘(backpropagation algorithm)을 이용한 신경망 분류기로부터 분류결과를 획득하고 이를 이용하여 FBKS와 BKS를 비교 분석하였

다.

Euclidian거리 분류기와 Mahalanobis거리 분류기는 학습을 통해 결정트리(decision tree)를 생산하는 분류기로서 인식을 위해 사용된 거리척도(distance measure)에 따라 구분하였으며, 이들 분류기는 클래스 레이블에 대한 측정값으로서 거리값을 출력하는 유형의 분

Fourier 변환을 수행하고 이들로 부터 각각 64개의 특징(feature)을 추출하여 개별 분류기의 분류결과를 얻는 데 사용하였다. 테스트 샘플에 대한 개별 분류기들의 분류성능은 <표2>에 나타난 것과 같다.

<표 2>에서 볼 수 있는 것과 같이 2000개의 테스트 샘플에 대한 개별 분류기의 성능이 다소

<표 2> 테스트 샘플에 대한 개별 분류기의 성능

분류기	샘플의 수	정확히 분류된 샘플의 수	분류율(%)
신경망 분류기	2,000	1734	86.70
Euclidian-distance 분류기	2,000	1825	91.25
Mahalanobis-distance 분류기	2,000	1851	92.55

류기이다. 실험을 위해 거리값의 범위 [a, c]는 실제로 분류기가 출력할 수 있는 최소거리와 최대거리로 결정하였다. 신경망 분류기는 클래스 레이블에 대한 가중치(weight)를 출력하는 유형의 분류기로서 출력되는 측정값의 범위는 0에서 1사이의 값이된다.

개별 분류기의 분류결과를 획득하기 위해 사용된 실험 데이터는 CEDAR의 CDROM-I 데이터베이스[8]의 digits디렉토리에 존재하는 서브 디렉토리 br로 부터 4000개의 샘플, good-bs 디렉토리로 부터 2000개의 샘플을 사용하였다. 또한, 각 샘플들로 부터 2차원

낮아 보인다. 그러나 FBKS와 BKS에 대한 실험에서 중요한 것은 개별 분류기의 성능이 아니라 개별 분류기의 결과를 어떻게 효과적으로 통합함으로써 분류성능의 향상을 가져다 줄 수 있는가에 있다.

제안된 FBKS방법을 기존의 BKS방법과 비교하기 위해 이 두가지 방법을 IBM-PC/486 상에서 Microsoft사의 Visual C++언어를 이용하여 구현하였으며, 실험에 사용될 개별 분류기들의 분류결과는 파일에 기록하여 사용하였다. <표 3>은 실제 실험에 사용된 개별 분류기들의 분류결과들 중 일부를 보이고 있다.

〈표 3〉 개별 분류기들의 분류결과 리스트 예

0	0.000000	1.000000	0.999044	0.000867	0.000000	0.041231	0.206398	// 샘플 1
			0.708520	0.099443	0.000323	0.000000	0.034505	
1	0.000000	0.708520	0.053852	0.197484	0.708520	0.041231	0.708520	
			0.708520	0.030000	0.708520	0.708520	0.708520	
1	0.000000	0.708520	0.093808	0.708520	0.708520	0.999420	0.000448	
			0.400456	0.057446	0.708520	0.708520	0.203675	
0	0.000000	1.000000	0.000006	0.465469	0.000000	0.708520	0.065574	// 샘플 2
			0.065574	0.000025	0.098976	0.000000	0.000000	
1	0.000000	0.708520	0.708520	0.026458	0.076811	0.708520	0.065574	
			0.708520	0.708520	0.086603	0.051962	0.708520	
1	0.000000	0.708520	0.708520	0.026458	0.708520	0.055579	0.999939	
			0.000000	0.708520	0.708520	0.086023	0.103482	
0	0.000000	1.000000	0.000097	0.000472	0.999297	0.708520	0.074162	// 샘플 3
			0.017321	0.041589	0.000000	0.000000	0.000000	
1	0.000000	0.708520	0.708520	0.708520	0.066333	0.708520	0.044721	
			0.017321	0.708520	0.084261	0.036056	0.708520	
1	0.000000	0.708520	0.708520	0.708520	0.066333	0.708520	0.044721	
			0.017321	0.708520	0.061644	0.036056	0.708520	

〈표 3〉에 보이고 있는 예는 주어진 샘플에 대한 각 분류기의 출력결과를 나타낸 것이며, 이것은 (식 1)에 보이고 있는 것과 같은 형태를 가지게 된다. 즉, 첫 번째 항목은 분류기가 출력하는 측정값이 거리값인지 확신도인지를 구분하기 위한 것이고, 두 번째와 세 번째의 값

들은 각 분류기가 출력하는 측정값의 범위를 나타내고 있다. 단, 측정값들은 클래스 레이블의 순서(0에서 9까지)와 같은 순서를 가지고 있다.

제안된 FBKS방법의 유용성을 증명하기 위해 기존에 제안된 BKS방법과 비교 실험을 수행하였다. BKS방법을 구현하기 위해 참고문

〈표 4〉 BKS방법과 FBKS방법의 성능 비교

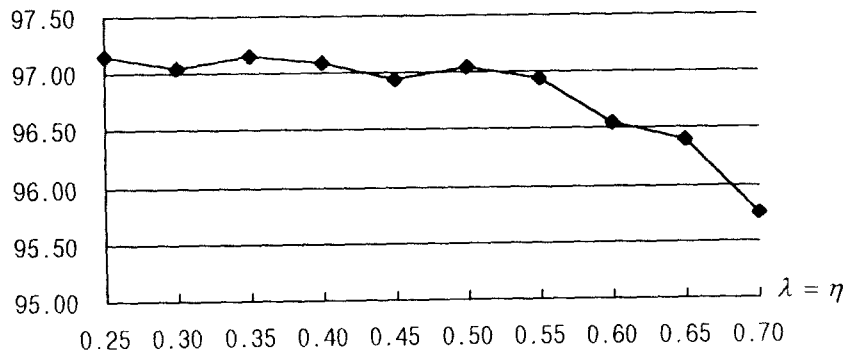
통합 분류기	분류율(%)
BKS	96.20
FBKS	97.05

헌 [2]를 참고하였으며, α 는 0.9로 하였다. BKS 방법을 실험하기 위해 사용된 데이터는 각 분류기의 분류결과로 부터 가장 큰 측정값 (신뢰도 또는 유사성)을 가진 클래스 레이블을 BKS에 대한 개별 분류기의 분류결과로 가정하였다. 이것은 BKS가 단일 클래스 레이블을 출력하는 경우에만 적용될 수 있기 때문이다. <표 4>는 BKS방법과 FBKS방법에 대한 분류성능을 보이고 있다.

<표4>에서 알 수 있듯이 제안된 FBKS방법이 BKS방법에 비해 비교적 좋은 성능을 가진 것으로 확인된다. FBKS방법은 개별 분류기들의 평균 분류율 90.17%에 비해 약 6.88%의 분류율 향상을 보이고 있다. 특히, BKS방법에

0.05씩 증가시켜 가면서 실험하였고 λ 와 η 의 값은 같도록 하였다. (그림3)은 FBKS 방법에 대해 임계치 $\lambda(=\eta)$ 의 변화에 따른 분류율을 보여주고 있다.

(그림 3)은 임계치 λ 의 변화에 따른 분류율의 변화를 그림으로 보여 주고 있다. λ 는 선택될 수 있는 상황의 수를 제한하는 역할을 수행하게 되는데, 실험에서 알 수 있는 것과 같이 λ 의 값이 0.5이하일 때 비교적 분류율이 높게 된다. 이것은 개별 분류기의 통합에 많은 상황이 선택됨으로써 분류율의 향상을 가져올 수 있다는 것을 의미한다고 할 수 있다. 그러나 만일 λ 의 값이 매우 작게 결정되면 매번 매우 많은 수의 상황이 선택되므로 FBKS의 학습과 결정에 소



(그림 3) 임계치 $\lambda(=\eta)$ 의 변화에 따른 분류율의 변화

비해 약 0.85%의 성능향상이 있음은 CME개념에 기초한 유사한 방법들간에 있어서는 매우 중요한 발전이라고 할 수 있다.

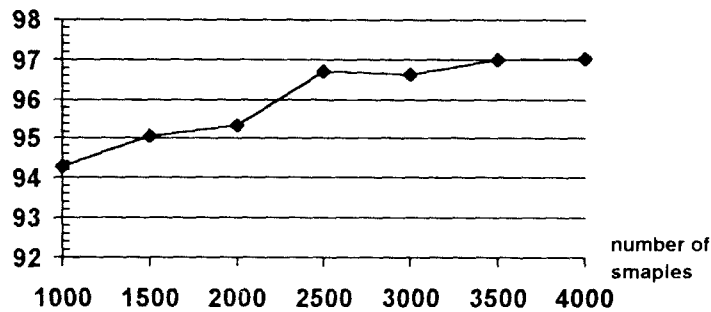
FBKS방법은 상황의 다중 선택을 위해 사용되는 임계치 λ 와 η 의 값을 0.25로부터 0.7까지

요되는 시간이 길어지게 된다는 단점이 있다. 따라서 수행시간을 단축하고 분류율의 향상을 기하기 위해 λ 을 0.5 정도로 하는 것이 바람직한 것으로 보인다.

FBKS를 이용한 방법에서 분류율은 FBKS

를 구성하는 데 기여한 학습샘플의 수에 따라 달라질 수 있다. 즉, 학습을 위한 샘플의 수가 많으면 많을수록 FBKS에 포함되어 있는 분류기들의 행위지식은 보다 안정적으로 표현될 수 있다. (그림4)는 FBKS 구성에 사용된 샘플의 수에 따른 분류율의 변화를 그림으로 보여주고 있다.

기대되며, λ 와 FBKS구성에 적정한 샘플의 수는 응용분야나 분류기의 특성, 실험환경등의 변화에 따라 다르게 결정될 수 있다. 본 논문의 실험에서는 λ 를 0.5로 하고 FBKS의 구성에 약 3500개 이상의 샘플이 사용되었을 때 만족할 만한 결과를 얻을 수 있었다. 그러나 보다 다양한 실험 데이터와 개별 분류기에 대한 실험이



(그림 4) FBKS의 학습에 사용된 샘플의 수에 따른 분류율의 변화

(그림 4)는 λ 와 η 를 0.5로 고정시켜 놓고, FBKS를 구성하는데 사용되는 학습샘플의 수를 변화시켰을 때의 분류율이며 학습샘플의 수가 증가될수록 분류율도 향상되는 것을 알 수 있다. 그러나 어느정도 샘플의 수가 증가되면 더 이상 샘플의 수가 증가되더라도 분류율의 증가는 완만하게 될 것이다. 따라서 λ 를 적절히 결정하고 FBKS구성에 충분한 샘플이 사용되었다면 분류율의 많은 향상을 가져올 것으로

필요할 것으로 생각된다.

실험결과에서 나타난 것과 같이 FBKS는 기존의 BKS에 비해 여러 가지 특성을 가지고 있다. 첫째, BKS는 단지 하나의 클래스 레이블만을 출력하는 분류기들의 통합에 이용될 수 있지만, FBKS는 클래스 레이블들의 순서화된 리스트와 관련된 측정값을 출력하는 분류기의 통합에 이용될 수 있다. 그러나 BKS에 비해 다소 복잡한 결정과정을 갖고 있다. 두 번째로

BKS 방법에서는 학습과 결정과정에서 단지 하나의 상황만을 고려하지만, FBKS 방법에서는 다수의 상황을 선택할 수 있다. 이러한 특징은 결정과정에 다수의 상황을 고려함으로써 보다 정확한 결정을 유도할 수 있도록 지원해 준다. 세 번째로 각 상황내에서 가장 가능성이 높은 클래스를 결정하기 위해 인간 전문가의 경험적 지식을 효과적으로 반영할 수 있다는 점이다. 이밖에도 많은 특징을 가지고 있는 FBKS 방법은 기존의 BKS에 퍼지개념을 추가함으로써 측정값 레벨 정보를 출력하는 분류기의 통합에 이용될 수 있다는 장점을 갖고 있다. 그러나 개별 분류기의 수가 많아질수록 FBKS는 고차원 공간으로 구성되므로 기억장소의 요구량이 매우 방대하며, FBKS를 구성하는 데 비교적 많은 시간을 요구하게 된다. 따라서 구현시 요구되는 기억장소의 양을 최소화 시키면서 소요되는 시간을 단축할 방법이 강구되어야 할 것으로 생각되며, 임계치의 자동결정 방법등에 대한 연구가 이루어져야 할 것으로 생각된다. 또한, 분류결과 의의와 크기를 효과적으로 일치시키기 위한 정규화방법에 관한 연구가 필요할 것으로 생각된다.

V. 결론

본 논문에서는 다중 분류기의 통합을 위해 퍼지 행위지식 공간을 구성하고 이를 이용하는 방법을 제안하였다. 기존의 행위지식 공간은 각 분류기들이 서로 독립적일 필요가 없고 적응적

학습이 가능한 것으로 단지 하나의 클래스 레이블만을 출력하는 분류기들의 통합에 가장 최적의 방법으로 알려졌다. 그러나 행위지식 공간은 각 분류기가 출력하는 클래스 레이블에 대한 측정값과 경험적 지식을 통합과정에 반영하기 어렵다는 문제점을 갖고 있다. 이러한 행위지식 공간의 문제점을 해결하기 위해 본 논문에서는 퍼지개념을 이용한 퍼지 행위지식 공간을 정의하고 이를 다중 분류기의 통합에 적용하기 위한 방법을 기술하였다.

제안된 방법을 각 개별 분류기로 부터 얻어진 클래스 레이블들과 관련된 측정값을 포함하는 분류결과들의 통합에 적용하여 봄으로써 퍼지 행위지식 공간 방법의 효과를 설명하였다. 또한, 기존의 행위지식 공간 방법과 실험을 통해 비교하여 봄으로써 제안된 방법의 유용성을 증명하였다. 그러나 구현시 요구되는 기억장소의 양을 최소화 시키면서 소요되는 시간을 단축할 방법과 임계치의 자동결정 방법등에 대한 연구가 심도있게 이루어져야 할 것으로 생각되며, 측정값의 의미와 크기를 효과적으로 일치시키기 위한 정규화방법에 관한 연구가 필요할 것으로 생각된다.

참고문헌

- [1] Y.S. Huang and C.Y.Suen, "The Behavior-Knowledge Space Method for Combination of Multiple Classifiers,"

Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 347-352, New York, 1993.

[2] Y.S. Huang and C.Y. Suen, "An Optimal Method of Combining Multiple Experts for Handwritten Numerical Recognition," *the Third International Workshop on Frontiers in Handwriting Recognition*, pp. 11-20, Buffalo, New York, USA, 1993.

[3] T.K. Ho, J.J. Hull and S.N. Srihari, "Decision Combination in Multiple Classifier Systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 1, pp. 66-75, January, 1994.

[4] Y.S. Huang and C.Y. Suen, "Combination of Multiple Classifiers with Measurement values," *Proc. 2nd International Conference on Document Analysis and Recognition*, pp. 598-601, 1993.

[5] Fumiaki Yamaoka, Yi Lu, Adnan Shaout and M. Shridhar, "Fuzzy Integration of Classification Results in a Handwritten Digit Recognition System," *the 4th International Workshop on Frontiers in Handwriting Recognition*, pp. 255-264, 1994.

[6] F. Kimura and M. Shridhar, "Handwritten Numerical Recognition Based On Multiple Algorithms," *Pattern Recognition*, 1991, vol.24, no.10, pp.969-983, 1991.

[7] Y.S. Huang and C.Y.Suen, "A Method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 1, Jan., pp. 347-352, 1995.

[8] Richard Fenrich and Jonathan J. Hull, "Concerns in Creation of Image Databases," *the Third International Workshop on Frontiers in Handwriting Recognition*, pp. 112-121, Buffalo, New York, USA, 1993.

[9] H.J. Zimmermann, *Fuzzy Set Theory- and Its Applications*, 2nd Edition.