

인식기 간의 상호정보를 이용한 인식기 선택

(Selecting Classifiers using Mutual Information between Classifiers)

강 희 중 ^{*}

(Hee Joong Kang)

요 약 패턴인식 문제에 있어서 다수의 인식기를 사용하는 연구는 주로, 선택된 다수 인식기를 어떻게 결합할 것인가에 중점을 두어 왔으나, 최근에는 인식기 풀로부터 다수 인식기를 선택하려는 연구로 점차 진행되고 있다. 실제로 다수 인식기 시스템의 성능은 인식기들의 결합 방법은 물론, 선택되는 인식기에 의존한다. 따라서, 우수한 성능을 보이는 인식기 집합을 선택하는 것이 필요하며, 다수의 인식기를 선택하는데 있어서 정보이론에 기초한 접근 방법이 시도되었다. 본 논문에서는 인식기 간의 상호정보를 기반으로 인식기를 선택하여 인식기 집합을 구성하고, 다른 인식기 선택 방법들에 의해 구성된 인식기 집합과 그 성능을 비교해 보고자 한다.

키워드 : 다수 인식기, 상호정보, 인식기 선택 방법, 인식기 집합

Abstract The study on combining multiple classifiers in the field of pattern recognition has mainly focused on how to combine multiple classifiers, but it has gradually turned to the study on how to select multiple classifiers from a classifier pool recently. Actually, the performance of multiple classifier system depends on the selected classifiers as well as the combination method of classifiers. Therefore, it is necessary to select a classifier set showing good performance, and an approach based on information theory has been tried to select the classifier

· 본 연구는 2006년도 한성대학교 교내연구비 지원과제 임
· 이 논문은 2007 한국컴퓨터종합학술대회에서 '인식기 간의 상호정보를 이용한 인식기 선택 방법'의 제목으로 발표된 논문을 확장한 것임

^{*} 종신회원 : 한성대학교 컴퓨터공학과 교수

hjkang@hansung.ac.kr

논문접수 : 2007년 9월 27일

심사완료 : 2008년 2월 14일

Copyright©2008 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨팅의 실제 및 레터 제14권 제3호(2008.5)

set. In this paper, a classifier set candidate is made by the selection of classifiers, on the basis of mutual information between classifiers, and the classifier set candidate is compared with the other classifier sets chosen by the different selection methods in experiments.

Key words : multiple classifiers, mutual information, selection of classifiers, classifier set

1. 서론

패턴인식 문제에 있어서 다수의 인식기를 결합함으로써 인식 성능의 향상을 보여준 연구 결과는 10여년 동안 보고되어 왔다[1,2]. 이러한 다수 인식기 시스템의 성능은 결합 방법은 물론, 선택되는 인식기에 의존한다. 그렇지만, 인식기 풀로부터 다수 인식기를 선택하려는 연구는 최근에 수행되고 있다[3]. 다수 인식기의 선택에 관한 연구, 즉, 어떻게 선택해야 하는가에 관한 문제와 얼마나 많이 선택해야 하는가에 관한 문제는 여전히 중요한 문제로 남아 있다. Woods 등은 다수 인식기의 혼합체를 선택하는데 있어서 선택 전략이 필요한 이유를 보여 주었다[2]. 그 이유는 적은 수의 인식기들이 더 많은 수의 인식기들 보다 더 우수한 성능을 보이는 예를 관찰했기 때문이다. 다수 인식기의 선택에 관한 초기 연구 결과도 소개되고 있다[3]. 그리고, 인식기에 사용되는 특징을 선택하기 위해 제안된 다양한 유전 알고리즘들[4]이 인식기 선택 방법에 활용되기도 한다.

본 논문에서는, 다수의 인식기를 선택하는데 있어서 정보이론에 기초한 접근 방법을 검토하여, 이를 바탕으로 인식기 간의 상호정보를 기반으로 인식기를 선택하여 인식기 집합을 구성하고, 다른 인식기 선택 방법들에 의해 구성된 인식기 집합과 그 성능을 비교해 보고자 한다. 다수 인식기의 선택 문제를 완화하기 위하여 선택되는 인식기의 개수는 미리 정해져 있다고 가정한다. 단순한 인식기의 선택 방법으로는 인식율(forced recognition rate) 또는 신뢰율(reliability rate)의 순위에 따라 인식기를 나열한 다음, 정해진 인식기의 개수 만큼 선택하는 방법이 있다.

정보이론에 기초한 인식기 선택 방법으로는 인식기들로 구성된 확률 분포를 활용하여 인식기들 간의 의존관계를 고려하는 유사도(measure of closeness)에 기반한 선택 방법[5]과 조건부 엔트로피(conditional entropy)에 기반한 선택 방법[6]이 있다. 그리고, 인식기 간의 상호정보(mutual information)를 기반으로 상호정보의 합이 최소가 되도록 인식기를 선택하는 mMI(minimization of mutual information) 선택 방법과 최대가 되도록 인식기를 선택하는 MMI(maximization of mutual information) 선택 방법을 사용하고자 한다. mMI 선택 방법

은 선택되는 인식기들이 가급적 서로 상호 보완적이라도 하기 위함에서 제안된 방법이고, MMI 선택 방법은 가급적 서로 상관관계가 높도록 하기 위함에서 제안된 방법이다.

이들 인식기 선택 방법들은 무제약 필기 숫자를 인식 하는 인식기를 대상으로 적용되어, 인식기 풀로부터 정해진 수 만큼의 인식기를 선택하여 인식기 집합을 구성 하고, 구성된 인식기 집합 후보 중에서 최선의 인식기 집합을 선택한다. 선택된 인식기 집합은 다양한 인식기 결합 방법에 의해서 그 성능을 비교해 본다. 사용된 무 제약 필기 숫자 데이터는 UCI(University of California, Irvine) 대학[7]으로부터 얻은 것이다.

본 논문의 구성은 다음과 같다. 2장에서는 정보이론에 기초한 다양한 인식기 선택 방법에 대해서 소개하고, 3 장에서는 실험을 통해서 얻은 인식기 집합의 선택과 성능 결과에 대하여 설명한다. 4장에서는 제안된 방법에 대한 토의와 향후 과제로 끝맺는다.

2. 정보이론에 기초한 인식기 선택 방법

2.1 유사도 기반의 선택 기준

유사도는 식 (1)과 같이 인식 대상이 되는 레이블 클래스와 인식기들로 구성된 고차 확률 분포 $P(C)$ 가 있을 때, 고차 확률 분포를 직접 사용하기에 적절하지 않은 경우, 최적의 근사 확률 분포 $P_a(C)$ 를 구하기 위한 기준으로 사용되었다[5]. 즉, 확률 분포의 요소인 레이블 클래스를 포함한 인식기들 간의 1차 또는 2차 또는 d 차 의존관계를 고려하여, 각 의존관계 차수에 따라 $P_a(C)$ 를 결정한다. 아래 수식은 유사도 정의 $I(P(C), P_a(C))$ 로부터 d 차 의존관계에 따라 최적의 근사 확률 분포를 구하기 위한 유도식이며, 확률 분포의 요소 변수에 사용된 아래 첨자는 1부터 $(K+1)$ 까지의 정수에 대한 미지의 순열을 나타낸다. 사용되는 인식기의 개수는 K 라고 가정한다.

$$\begin{aligned} I(P(C), P_a(C)) &= \sum_c P(c) \log \frac{P(c)}{P_a(c)} \\ &= \sum_c P(c) \log P(c) - \sum_{j=1}^{K+1} \sum_c P(c) \log P(C_{n_j} | C_{n_{id(j)}}, \dots, C_{n_{il(j)}}) \\ &= - \sum_{j=1}^{K+1} M(C_{n_j}; C_{n_{id(j)}}, \dots, C_{n_{il(j)}}) + \sum_{j=1}^{K+1} H(C_{n_j}) - H(C) \\ P_a(C_1, \dots, C_{K+1}) &= \prod_{j=1}^{K+1} P(C_{n_j} | C_{n_{id(j)}}, \dots, C_{n_{il(j)}}), (0 \leq id(j), \dots, il(j) < j) \\ H(C) &= - \sum_c P(c) \log P(c) \\ M(C_{n_j}; C_{n_{id(j)}}, \dots, C_{n_{il(j)}}) &= \sum_c P(c) \log \frac{P(C_{n_j} | C_{n_{id(j)}}, \dots, C_{n_{il(j)}})}{P(C_{n_j})} \end{aligned} \quad (2)$$

$$P(C_{n_j} | C_0, C_{n_{i \cdot (j)}}) \equiv P(C_{n_j}, C_{n_{i \cdot (j)}}) \quad (3)$$

위의 식 (1)로부터 유사도를 최소로 낮추는 것이 바람직 한데, 이를 위해서는 식 (2)와 같이 정의된 d 차 의존관계에 따른 상호정보 $M()$ 의 합이 최대가 되도록 아래 첨자에 대한 미지의 순열 (n_1, \dots, n_{K+1}) 을 결정하면 된다. 확실적인 성질을 만족하도록 식 (3)을 정의한다. 유사도 기반의 인식기 선택 기준에서는 이러한 맥락에 따라, 각 인식기 집합에 대하여, 구할 수 있는 d 차 의존관계에 따른 상호정보 $M()$ 의 합이 높을수록 해당 인식기 집합이 더 우수할 것이라 가정하고, 그러한 인식기 집합을 다수 인식기 시스템에 포함될 인식기 집합으로 결정한다.

2.2 조건부 엔트로피 기반의 선택 기준

조건부 엔트로피는 식 (4)와 같이 베이스 에러율 P_e 의 상한과 관련이 있으며, 인식 대상인 레이블 클래스 L 과 인식기 그룹 E 간의 조건부 엔트로피 $H(L|E)$ 를 최소화 함으로써 베이스 에러율의 상한을 낮추려는 의도에서 활용되었다. 조건부 엔트로피의 정의로부터 클래스-결정 $(C-D)$ 상호정보 $U(L;E)$ 가 아래 식 (5)와 같이 유도되고, 베이스 에러율의 상한을 낮추는 것은 곧 클래스-결정 상호정보 $U()$ 를 최대화하는 것에 상응된다[6]. 이와 같이, 클래스-결정 상호정보는 레이블 클래스와 인식기들로 구성된 고차 확률 분포 $P(E, L)$ 와 인식기들만으로 구성된 고차 확률 분포 $P(E)$ 가 있을 때, 이들 고차 확률 분포를 직접 사용하기에 적절하지 않은 경우, 최적의 근사 확률 분포 $P_a()$ 를 각각 구하기 위한 기준으로 사용되었다. 아래의 수식들은 클래스-결정 상호정보 $U()$ 로부터 d 차 의존관계에 따라 최적의 근사 확률 분포를 구하기 위한 유도 과정을 보여 준다. 확률 분포의 요소 변수에 사용된 아래 첨자는 1부터 K 까지의 정수에 대한 미지의 순열을 나타낸다.

$$P_e \leq \frac{1}{2} H(L|E) = \frac{1}{2} [H(L) - U(L;E)] \quad (4)$$

$$\begin{aligned} U(L;E) &= \sum_e \sum_l P(e, l) \log \frac{P(e|l)}{P(e)} \\ &= \sum_e \sum_l P(e, l) \log \left[\frac{1}{P(l)} \prod_{j=1}^K P(E_{n_j} | E_{n_{id(j)}}, \dots, E_{n_{il(j)}}, l) \right] \\ &\quad - \sum_e P(e) \log \prod_{j=1}^K P(E_{n_j} | E_{n_{id(j)}}, \dots, E_{n_{il(j)}}) \\ &= H(L) + \sum_{j=1}^K \Delta D(E_{n_j}; E_{n_{id(j)}}, \dots, E_{n_{il(j)}}, L) \end{aligned} \quad (5)$$

$$\begin{aligned} P_a(E_1, \dots, E_K, L) &= \prod_{j=1}^K P(E_{n_j} | E_{n_{id(j)}}, \dots, E_{n_{il(j)}}, L), \\ &\quad (0 \leq id(j), \dots, il(j) < j) \\ P_a(E_1, \dots, E_K) &= \prod_{j=1}^K P(E_{n_j} | E_{n_{id(j)}}, \dots, E_{n_{il(j)}}), \\ &\quad (0 \leq id(j), \dots, il(j) < j) \end{aligned}$$

$$H(L) = - \sum_l P(l) \log P(l)$$

$$\Delta D(E_{n_j}; E_{n_{a(j)}}, \dots, E_{n_{a(j)}}, L)$$

$$= D(E_{n_j}; E_{n_{a(j)}}, \dots, E_{n_{a(j)}}, L) - D(E_{n_j}; E_{n_{a(j)}}, \dots, E_{n_{a(j)}}) \quad (6)$$

$$D(E_{n_j}; E_{n_{a(j)}}, \dots, E_{n_{a(j)}}, L)$$

$$= \sum_e \sum_l P(e, l) \log \frac{P(E_{n_j} | E_{n_{a(j)}}, \dots, E_{n_{a(j)}}, l)}{P(E_{n_j})}$$

$$D(E_{n_j}; E_{n_{a(j)}}, \dots, E_{n_{a(j)}})$$

$$= \sum_e P(e) \log \frac{P(E_{n_j} | E_{n_{a(j)}}, \dots, E_{n_{a(j)}})}{P(E_{n_j})}$$

$$P(E_{n_j} | E_0, E_0, L) \equiv P(E_{n_j}, L) \quad (7)$$

$$P(E_{n_j} | E_0, E_{n_{i(j)}}, L) \equiv P(E_{n_j} | E_{n_{i(j)}}, L) \quad (8)$$

$$P(E_{n_j} | E_0, E_{n_{i(j)}}) \equiv P(E_{n_j}, E_{n_{i(j)}}) \quad (9)$$

위의 식 (4)로부터 베이스 엔트로피의 상한을 최저로 낮추는 것이 바람직 한데, 이를 위해서는 클래스-결정 상호정보 $U(L;E)$ 를 최대화 시켜야 한다. 즉, 식 (5)에서와 같이 d차 의존관계에 따른 상호정보 $\Delta D()$ 의 합이 최대가 되도록 미지의 순열(n_1, \dots, n_k)을 결정한다. 확실적인 성질을 만족하도록 식 (7)-(9)를 정의한다. 이와 같이, 조건부 엔트로피 기반의 인식기 선택 기준에서는 각 인식기 집합에 대하여, 구할 수 있는 d차 의존관계에 따른 상호정보 $\Delta D()$ 의 합이 높을수록 해당 인식기 집합이 더 우수할 것이라 가정하고, 그러한 인식기 집합을 다수 인식기 시스템에 포함될 인식기 집합으로 결정한다.

2.3 mMI 선택 기준

mMI 선택 기준은 인식기 풀로부터 선택되는 인식기들이 가급적 서로 상호 보완적 이었으면 하는 바램에서 사용되었다. 레이블 클래스와 인식기, 또는 인식기 간의 d차 의존관계 정도를 결정하는데 있어서 2.1절의 식 (2)와 같은 상호정보 $M()$ 의 정의를 사용하였으며, 이러한 상호정보의 정의는 이들 간의 통계적인 상호 보완의 정도를 상대적으로 측정한다고 가정하였다. 즉, 상호정보의 값이 낮으면 낮을수록 상호 보완의 정도는 더 높다고 가정한다. mMI 선택 기준은 초기에 인식기 집합 S가 공집합이라고 가정하고, 가능한 레이블 클래스와 인식기 간의 d차 의존관계에 의한 상호정보와 인식기들 간의 d차 의존관계에 의한 상호정보를 모두 각각 계산한다. 그 다음 아래와 같은 순서에 따라 인식기를 선택하여 인식기 집합을 결정한다.

- (1) 레이블 클래스와 인식기 간의 상호정보 값이 최대인 인식기를 선택하여 인식기 집합 S에 넣는다.
- (2) 인식기 집합 S의 인식기들과 가급적 상호 보완적인 인식기를 선택하도록 인식기 풀로부터 S에 있는 모든 인식기와의 계산된 상호정보 값을 합하여, 그 합이 최소가 되는 인식기를 선택하고 S에 추가한다.

- (3) 미리 정의된 인식기의 개수 만큼 선택될 때 까지, (2) 단계를 반복 수행한다.
- (4) (3) 단계를 마친 상태에서 S에 포함되어 있는 인식기들을 다수 인식기 시스템에 포함시킨다.

2.4 MMI 선택 기준

MMI 선택 기준은 인식기 풀로부터 선택되는 인식기들이 가장 우수한 인식기와 가급적 서로 높은 상관관계를 지녔으면 하는 바램에서 사용되었다. 2.3절과 마찬가지로 2.1절의 식 (2)와 같은 상호정보 $M()$ 의 정의를 사용하였으며, 통계적인 상관관계의 정도를 상대적으로 측정한다고 가정하였다. 즉, 상호정보의 값이 높으면 높을수록 상관관계의 정도는 더 높다고 가정한다. MMI 선택 기준도 초기에 인식기 집합 S가 공집합이라고 가정하고, 가능한 레이블 클래스와 인식기 간의 d차 의존관계에 의한 상호정보와 인식기들 간의 d차 의존관계에 의한 상호정보를 모두 각각 계산한다. 그 다음 아래와 같은 순서에 따라 인식기를 선택하여 인식기 집합을 결정한다.

- (1) 레이블 클래스와 인식기 간의 상호정보 값이 최대인 인식기를 선택하여 인식기 집합 S에 넣는다.
- (2) 인식기 집합 S의 인식기들과는 가급적 상관관계가 높은 인식기를 선택하도록 인식기 풀로부터 S에 있는 모든 인식기와의 계산된 상호정보 값을 합하여, 그 합이 최대가 되는 인식기를 선택하고 S에 추가한다.
- (3) 미리 정의된 인식기의 개수 만큼 선택될 때 까지, (2) 단계를 반복 수행한다.
- (4) (3) 단계를 마친 상태에서 S에 포함되어 있는 인식기들을 다수 인식기 시스템에 포함시킨다.

3. 인식기 집합의 선택 및 성능 실험

다양한 인식기 선택 방법을 비교하기 위하여, 실험에서 E1, E2, E3, E4, E5, E6 등 6 개의 숫자 인식기를 사용한다. 이들 인식기는 KAIST와 전북대학교에서 필기 숫자 인식을 위해 숫자의 통계적인 특징 정보와 구조적인 지식을 사용하여 개발되었으며, 개략적인 인식기의 특징과 테스트용 데이터 집합에 대한 인식 성능은 표 1과 같다. UCI 대학의 필기 숫자 데이터 집합은 학습용으로 tra, cv, wdep 등이 있고, 테스트용으로 windep가 있다[7]. 인식기 E4와 E5에 사용된 구조적 지식들은 주로 Concordia 대학의 숫자 데이터를 참조해서 만들어 졌기 때문에, UCI 대학의 숫자 데이터(windep)에 대해서는 기각율이 높아 인식을 성능이 상대적으로 낮았다.

각 신경망 기반의 인식기들은 학습용 데이터 집합인 tra를 이용하여 학습되었으며, 인식기 선택 방법을 위한

표 1 개별 인식기의 개요와 성능

인식기	구조	인식 방법	거리 함수	인식율	신뢰율
E1	단일	신경망	픽셀 거리 함수	93.77	93.77
E2	다중	신경망	방향성 거리 분포	97.11	97.11
E3	단일	신경망	메쉬 특징	91.82	96.95
E4	다중	규칙 기반	신 구조적 지식	67.67	93.11
E5	다중	규칙 기반	구조적 지식	70.01	94.80
E6	단일	신경망	칸투어 특징	96.66	96.66

표 2 인식기 선택 기준

기호	선택 기준	의존관계 차수
FRR	인식율의 순위 기반	해당 없음
RR	신뢰율의 순위 기반	해당 없음
MC1	유사도 기반	1차
CMC1	유사도 기반	조건부 1차
MC2	유사도 기반	2차
CE1	조건부 엔트로피 기반	1차
CE2	조건부 엔트로피 기반	2차
mMI	인식기 간의 상호정보	1차
MMI	인식기 간의 상호정보	1차

인식기 간의 상호정보와 의존관계를 고려한 베이지안 결합 방법을 적용하기 위한 최적의 저차 확률 분포의 값을 구하기 위해서 tra, cv, wdep가 사용되었다. 인식기 선택 기준은 정해진 인식기 개수에서 가능한 인식기 집합에 적용되어, 각 기준에 따라 가장 유망한 인식기 집합을 선택한다. 특히, 유사도(MC) 기반의 선택 기준과 조건부 엔트로피(CE) 기반의 선택 기준은 고려하는 의존관계의 차수에 따라 다양한 선택 기준으로 적용될 수 있다. 다만, mMI 선택 기준과 MMI 선택 기준의 경우에는 1차 의존관계만을 고려하였다. 인식기 선택 기준을 간단한 기호로 표기하면 표 2와 같다.

인식기 선택 기준에 의하여 다수 인식기 시스템에 포함될 인식기 집합이 결정되면, 다양한 인식기 결합 방법으로 그 성능을 비교한다. 사용된 결합 방법으로는 투표 방법, 보다(Borda) 카운트 방법, 의존관계를 고려한 베이지안 결합 방법 등이 있으며[1,5,6], 이들 결합 방법도 역시 기호로 표기하면 표 3과 같다.

테스트용 데이터 집합에 대하여 3개의 인식기로 이루어진 다수 인식기 시스템을 구축한다면 총 20개의 가능한 인식기 집합이 구성되어 후보로 고려되고, 4개의 인식기로 이루어진다면 총 15개의 가능한 인식기 집합이, 5개의 인식기로 이루어진다면 총 6개의 가능한 인식기 집합이 각각 구성되어 후보로 고려된다. 그 중에서, 각 인식기 선택 기준에 따라 결정된 인식기 집합을 보면 표 4와 같고, 이들 결정된 인식기 집합이 다수 인식기 시스템에 사용되어 다양한 결합 방법으로 결합되었을 때의 인식 성능은 표 5 - 표 7과 같다. 인식 성능을 보

표 3 인식기 결합 방법

기호	결합 방법의 개요
voting	절대 다수 투표 기법
Borda	보다 카운트 기법
CIAB	조건부 독립을 가정한 베이지안 기법
ODB1	1차 의존관계에 의한 베이지안 기법
CODB1	조건부 1차 의존관계에 의한 베이지안 기법
ODB2	2차 의존관계에 의한 베이지안 기법
DODB1	인식기들만의 1차 의존관계에 의한 베이지안 기법
DODB2	인식기들만의 2차 의존관계에 의한 베이지안 기법

표 4 인식기 개수에 대하여 선택된 인식기 집합

인식기 개수	인식기 선택 기준	인식기 집합
3	FRR,RR,MC1,CMC1,MC2,MMI	E2,E3,E6
	CE1,CE2	E2,E4,E6
	mMI	E1,E2,E4
4	FRR,RR,MC1,CMC1,MC2,MMI	E1,E2,E3,E6
	CE1	E2,E3,E4,E6
	CE2	E3,E4,E5,E6
	mMI	E1,E2,E3,E4
5	FRR,RR,MC1,CMC1,MC2,MMI	E1,E2,E3,E5,E6
	CE1,mMI	E1,E2,E3,E4,E6
	CE2	E2,E3,E4,E5,E6

표 5 인식기 선택 기준에 의한 3개 인식기 집합의 성능(%)

결합 방법	FRR,RR, MC,MMI	CE	mMI	best
voting	96.77	97.38	95.33	97.38
Borda	96.44	96.16	95.99	98.33
CIAB	96.88	97.22	97.38	97.77
ODB1	96.88	97.22	97.38	97.83
CODB1	96.99	97.55	96.77	97.77
ODB2	97.05	97.05	96.38	97.77
DODB1	97.16	97.94	97.61	97.94
DODB2	97.33	97.89	96.77	97.89

표 6 인식기 선택 기준에 의한 4개 인식기 집합의 성능(%)

결합 방법	FRR,RR,MC,MMI	CE1	CE2	mMI	best
voting	97.27	97.77	95.83	97.16	97.77
Borda	96.66	98.05	96.83	97.55	98.05
CIAB	96.83	97.44	96.94	97.44	97.66
ODB1	96.83	97.38	97.38	97.44	97.77
CODB1	97.11	97.83	97.33	97.89	98.16
ODB2	97.38	97.83	97.33	97.89	98.00
DODB1	97.44	97.89	96.88	97.94	98.05
DODB2	97.38	97.66	97.44	97.94	98.05

인 표들에서 MC 기호는 MC1, CMC1, MC2의 선택 기준을 모두 지칭한다고 가정하고, CE 기호는 CE1, CE2의 선택 기준을 모두 지칭한다고 가정한다. 그리고, 인식 성능표에 추가로 나타난 best 선택 기준이라는 것은

표 7 인식기 선택 기준에 의한 5개 인식기 집합의 성능(%)

결합 방법	FRR,RR,MC,MMI	CE1,mMI	CE2	best
voting	97.72	97.94	98.16	98.16
Borda	97.66	97.55	97.22	97.83
CIAB	97.05	97.22	97.33	97.33
ODB1	97.05	97.22	97.44	97.44
CODB1	97.61	97.77	97.38	98.00
ODB2	97.77	98.11	98.11	98.11
DODB1	97.89	98.44	98.27	98.44
DODB2	98.05	98.11	98.22	98.33

주어진 인식기 결합 방법에 있어서 가장 높은 인식율을 보인 인식기 집합의 성능을 의미한다고 가정한다. 즉, 특정의 한 인식기 집합을 지칭하는 것이 아니다. 표 5로부터 CE 선택 기준이 우수한 결과를 보임을 알 수 있었고, 표 6으로부터 CE1 선택 기준과 mMI 선택 기준이 우수한 결과를 보임을 알 수 있었으며, 표 7로부터 CE 선택 기준과 mMI 선택 기준이 우수한 결과를 보임을 알 수 있었다.

4. 토 의

위 실험을 통하여 단순히 인식율이나 신뢰율 측면에서 우수한 인식기를 순위에 따라 차례로 선택하는 방법보다는 정보이론 또는 상호정보에 기반한 인식기 선택 방법이 양호한 결과를 보임을 알 수 있었다. 특히, 조건부 엔트로피(CE) 또는 mMI 기반의 선택 기준은 다수 인식기 시스템을 위한 인식기 집합 후보를 결정하는데 유효했다고 볼 수 있다. 그러나, 후보로 결정된 인식기 집합이 항상 가장 좋은 인식율을 보인다고 보증할 수 없음도 알 수 있었다. 또한, 선택하려는 인식기의 개수가 미리 정해져 있다는 한계도 지니고 있다. 따라서, 상호정보에 기반한 인식기 선택 방법, 특히 유용한 mMI 선택 기준은 고려할 인식기 간의 의존관계 차수를 더 높은 차수까지 높이는 방법을 검토할 만 하다고 본다. 그리고, 좀 더 다양하고 많은 수의 데이터 집합들에 대해 실험을 하거나, 선택하려는 인식기 개수에 상관없이 가장 우수한 성능을 보이는 임의 수의 인식기 집합을 선택하는 방법이라던가, 또는 후보 인식기 집합 간의 우열을 제대로 구분해 낼 수 있는 체계적인 선택 기준의 고안 등과 같은 향후 연구가 필요하다고 본다.

참 고 문 헌

- [1] Kittler, J., Hatef, M., Duin, R.P.W., Matas, J., "On Combining Classifiers," IEEE TPAMI, Vol.20, pp. 226-239, 1998.
- [2] Woods, K., Kegelmeyer Jr., W.P., Bowyer, K., "Combination of Multiple Classifiers Using Local

Accuracy Estimates," IEEE TPAMI, Vol.19, pp. 405-410, 1997.

- [3] Kang, H.-J., Lee, S.-W., "Experimental Results on the Construction of Multiple Classifiers Recognizing Handwritten Numerals," In Proc. of the 6th ICDAR, pp. 1026-1030, 2001.
- [4] 오일석, 이진선, 문병로, "특징 선택을 위한 혼합형 유전 알고리즘과 분류 성능 비교", 정보과학회논문지:소프트웨어및응용, 제31권, 제8호, pp. 1113-1120, 2004.
- [5] Kang, H.-J., Lee, S.-W., "A Dependency-based Framework of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals," In Proc. of IEEE Computer Soc. Conf. on CVPR, Vol.2, pp. 124-129, 1999.
- [6] Kang, H.-J., Lee, S.-W., "Combining Classifiers based on Minimization of a Bayes Error Rate," In Proc. of the 5th ICDAR, pp. 398-401, 1999.
- [7] Blake, C., Merz, C., UCI repository of machine learning databases [http://www.ics.uci.edu/~mllearn/mlrepository.html], Irvine, CA, Dept. of Info. and Comp. Sciences, 1998.