

의존관계를 기반으로 다수 인식기를 결합하는 방법과 그 응용

(Combining Multiple Classifiers based on Dependency and
Its Application)

강 회 중 [†] 김 진 형 ^{††}

(Hee-Joong Kang) (Jin H. Kim)

요약 다수 인식기를 결합하는 여러가지 방법들이 제안되었으나, 대부분이 다수 인식기를 결합하는데 있어서 인식기 간의 의존관계를 고려하지 않았다 이것은 의존관계가 매우 높은 인식기가 추가될 때 다수 인식기를 결합하는 방법의 인식 성능을 저하시키거나 결합된 결과가 편향되도록 할 수도 있다 본 논문에서는 이러한 단점을 극복하고 안정적인 인식 성능을 얻기 위하여 의존관계를 기반으로 다수 인식기를 결합하는 방법을 제안한다 다수 인식기의 인식 결과로부터 의존관계를 결정하기 위하여 1차 의존관계로 근사하였으며, 두 가지 방법을 사용하였다. 하나는 상호 정보의 개념을 사용하는 것이고, 다른 하나는 통계적으로 측정된 결합도의 개념을 사용하는 것이다 최종으로 결정된 1차 의존관계는 베이저안 공식을 사용하여 다수 인식기의 인식 결과를 결합하는데 사용된다 부제약 온라인 숫자, 영문 일파벳 인식을 위한 문자 인식기를 사용하였다. 실험한 결과, 다수 문자 인식기를 결합한 인식 성능이 대체로 개별 문자 인식기의 성능보다 우수하였으며, 특히 의존관계가 매우 높은 문자 인식기가 추가되었을 때 의존관계를 기반으로 결합하는 방법이 다른 방법보다 더 우수한 성능을 보여 주었다.

Abstract Although many methods for combining multiple classifiers have been proposed, most of them did not focus on dependency among classifiers in combining decisions from multiple classifiers. That makes classification performance of combining multiple classifiers be degraded and biased, in case of adding highly dependent classifiers To overcome such weaknesses and obtain robust classification performance, the present study used the dependency for better combining multiple classifiers. In order to identify the dependency by observing the outputs of multiple classifiers, two methods are used on the basis of first-order dependency. One is to use the concept of mutual information, and the other is to use the concept of statistically measured association. The optimally identified first-order dependencies are used to combine decisions from multiple classifiers, using Bayesian formalism A number of classifiers for on-line recognition of totally unconstrained handwritten numerals and English alphabet are used Our experimental results show that the classification performance of combining multiple classifiers is superior to that of individual ones, and in particular, considering the dependency outperforms others in recognition accuracy, when the highly dependent classifiers are added.

1. 서 론

이제까지 문자 및 패턴 인식 분야에서 인식 성능을

향상시키기 위한 연구는 주로 인식기의 개발에 초점을 두어 왔다 그리하여 많은 인식기가 개발되었지만, 예상만큼 만족스럽지 않았다 또한, 인식 과정을 위하여 다양한 형태의 많은 특징점이 필요하지만, 그들을 하나의 인식기에 모두 포함시킨다는 것도 결코 쉬운 일이 아니었다 최근에는 인식기 자체의 인식 성능을 향상시키려고 노력하는 대신 다수의 서로 다른 인식기를 동시에 사용하고, 그들의 인식 결과를 결합함으로써 인식 성능을

* 본 연구는 노트패드 컨소시움과 문자인식 기반기술 연구 컨소시움으로부터 지원을 받았다

† 정 회 원 한국과학기술원 전산학과

†† 종 신 회 원 한국과학기술원 전산학과 교수

논문접수 1995년 6월 1일

심사완료 1995년 9월 16일

향상시키려는 다수 인식기 시스템에 대한 연구가 시작되었다[1,2] 다수 인식기 시스템은 다수 인식기의 집합체와 이들의 인식 결과 (또는 결정)를 결합하는 결합기로 구성되는데, 다수의 상호 보완적인 특징점이나 인식 알고리즘이 동시에 사용된다면 인식 성능을 향상시킬 수 있다는 가정에서 연구되었다. 인식 성능의 향상을 위하여 다수 인식기를 결합하는 연구 아이디어는 낙관적으로 보인다

다수 인식기 시스템에서의 많은 연구는 인식기가 다수 전문가 그룹의 전문가에 상응한다는 사실에 근간을 두었으며, 다수 인식기로부터 다수의 인식 결과를 결합하는 방법론에 대한 연구가 수행되었다[1,2,3,4,5,6,7,8]. 이러한 연구는 다수의 인식 결과를 결합하는 방법의 향상을 위하여 수행되었으나, 대부분이 인식기 간의 의존관계를 고려하지 않았다[9,10]. 그 대신, 다수 인식기가 독립이라고 가정하고 인식 결과를 결합하였다 만일 매우 높은 의존관계를 지닌 열등한 인식기가 기존의 다수 인식기 시스템에 추가되고, 이전의 결합 방법으로 다수의 인식 결과를 결합한다면 인식 성능이 저하되거나 결합된 최종 인식 결과가 편향될 수도 있다 본 논문의 연구 동기는 이러한 기존 결합 방법의 단점을 극복하려는 것이다

인식 성능을 향상시키기 위하여 다수의 인식기가 상호 보완적 이도록 다수 인식기 시스템을 구성하고, 필요하다면 의존관계를 기반으로 다수의 인식 결과를 결합하는 것이 바람직하다 본 논문에서는 의존관계를 결정하는 방법과 결정된 의존관계를 기반으로 베이지안 공식을 이용하여 다수의 인식 결과를 결합하는 방법을 제안한다 이러한 의존관계는 인식기의 인식 결과만으로부터 측정될 수 있다고 본다.

K개의 인식기로부터, 한 인식기의 의존관계를 결정하기 위하여 Lewis가 정의한 곱 근사 (Product Approximation) 정의에 의하여 가능한 모든 저차원의 분포를 고려해야 한다[11]. 그것은 많은 계산량과 공간을 필요로 한다 따라서, 본 논문에서는 n차원 이진 변수 분포를 최적인 n-1개의 2차 확률 분포로 근사한 Chow 등의 연구 결과를 응용하여 K차원의 인식기 분포를 최적인 K-1개의 2차 분포로 근사하여 최적인 1차 의존관계의 집합을 결정하는 방법을 제안한다 두 인식기 간의 1차 의존관계를 결정하기 위하여 두 가지 방법을 사용하였다. 하나는 Chow 등이 정의한 상호 정보 (Mutual Information)을 사용하는 것이고[12], 다른 하나는 통계적으로 측정된 결합도를 사용하는 것이다[18] 계산된 이들 1차 의존관계들로부터 의존관계의 합

이 최대가 되는 최적인 1차 의존관계를 결정하기 위하여 Kruskal이 개발한 최대 가중치 스캐닝 트리 (Maximum Weight Spanning Tree) 알고리즘을 적용하였다 최적으로 결정된 1차 의존관계는 베이지안 공식을 이용하여 다수의 인식 결과를 결합하는데 사용된다

본 논문에서 사용된 인식기는 은닉 마르코프 모델을 기반으로 만든 문자 인식기이며, 무제약 온라인 숫자, 영문 알파벳 인식 문제를 해결하기 위하여 개발되었다. 은닉 마르코프 모델은 온라인 홀릴체 문자를 인식하기 위한 방법론으로 널리 사용되며[13], 시간적 정보와 공간적 정보의 변화를 잘 모델링한다 이들 문자 인식기는 다수 문자 인식기 시스템의 구성원으로 사용된다. 매우 높은 의존관계를 지닌 문자 인식기가 결합 방법에 미치는 효과를 보이기 위하여 여러가지 다수 문자 인식기 시스템을 구성하고 실험하였으며, 의존관계를 기반으로 결합하는 제안된 방법과 다른 결합 방법을 비교하였다

2. 관련 연구

인식 성능을 향상시키기 위하여 다수의 인식기를 결합하려는 기존의 관련 연구를 살펴보고, 교차원의 이진 변수 분포를 저차원의 분포로 근사하는 방법에 관한 관련 연구를 간략히 기술하고자 한다 후자의 관련 연구는 인식기가 독립이라는 가정을 극복하는데 기여한다.

2.1 인식 결과에 따른 결합 방법

다수 인식기를 결합하려는 연구는 인식기가 출력하는 인식 결과의 유형에 따라 구분할 수 있다. 이들 유형은 인식 대상 클래스에 대한 인식 점수, 클래스의 순위, 그리고 단 하나의 우승 클래스 등이 있다. 클래스에 대한 인식 점수로부터 순위를 결정하므로써 각 클래스의 순위를 표현할 수 있다 난 하나의 우승 클래스는 가장 우수한 인식 점수를 지닌 클래스 또는 1등 순위의 클래스만을 선택하므로써 결정된다

인식기의 인식 점수를 이용하여 다수의 인식 결과를 결합하는 연구로는 평균된 Bayes 인식기에 대한 연구[4]와 퍼지 적분[14] 또는 퍼지 논리를 이용하여 결합하는 연구 등이 있다 단 하나의 우승 클래스를 이용하여 결합하는 연구에는 투표 기법[2,4,5], 인식기가 독립이라는 가정하에서 사용된 베이지안 공식 기법[4], Dempster-Shafer 공식을 이용하는 기법[3,4,5], 그리고 Behavior-Knowledge Space (BKS) 기법[6] 등이 있다 특히 Xu 등이 제안한 베이지안 공식 기법에서[4] 그들은 인식기가 독립된 특징점을 사용하거나 독립적인 데이터로부터 훈련된 독립이라고 가정하였다 이 방법은 실제로 인식기가 독립이 아닐 경우에 문제가 발생할

수 있다. BKS 기법은 이미 본 K개의 인식 결과에 대한 입력 클래스의 빈도수 정보를 입력 정보 형태로 테이블에 저장했다가 최대 빈도수를 지닌 클래스를 결합한 우승 클래스로 결정하는 방식이다. 그러나, 보지 않은 인식 결과에 대해서는 각각하는 문제점을 지니고 있다. 또한, K가 커짐에 따라 이론적으로 기하학적인 계산량과 공간을 요구하는 문제점이 있다.

Ho 등은 인식 결과를 클래스의 순위로 표현하길 주장했으며, 다수 인식기로부터 그러한 순위를 이용하여 결합하는 방법에 관한 연구를 수행하였다[7,8]. 이들 방법에는 최고 순위를 이용하여 결합하는 방법(The highest rank method), Borda 카운트 방법, 그리고 Logistic Regression을 이용하는 방법 등이 있다. Borda 카운트 방법은 사회적 선택 함수 기법 중의 하나이며, Borda 함수라고도 불린다[15] 강희중 등은 Ho 등이 주장한 순위 표현의 장점을 살리기 위하여 인식기가 인식 결과를 클래스의 순위로 표현할 때 다수의 인식 결과를 결합할 수 있는 사회적 선택 함수 기법의 응용에 관한 연구를 수행하였다[15,16] 위 연구에서 사회적 선택 함수 기법은 투표 기법과 마찬가지로 훈련 단계가 필요없으며, 대체로 투표 기법보다는 나은 인식 성능을 보인다고 하였다. 본 논문에서는 순위 표현의 장점을 취하고 범용성 있는 다수 인식기의 결합 방법을 제시하기 위하여 인식기의 인식 결과가 인식 대상 클래스의 순위로 표현된다고 가정한다.

2.2 고차원 분포의 근사 방법

Lewis는 Hartmanis의 확장(extension) 아이디어를 이용하여 n차원의 이진 분포를 다수의 저차원의 분포에 의한 곱(product)로서 근사하는 문제를 다루었다[11] 그는 이러한 곱 근사가 적절한 조건하에서 최소 정보의 성질을 지님을 보였다. 제안한 근사 방법은 두 분포의 일치도(Closeness)를 측정하는 정보 측정치와 최대 엔트로피 원칙을 근거로 하고 있다. 이들에 근거하여 둘 이상의 제한된 근사 결과는 실제 분포에 대한 지식이 없이도 비교되어, 최적인 근사 결과가 선택된다 즉, 비교 프로세스는 실제 분포와 가장 큰 상관성을 지닌 근사 결과를 선택한다. 그러나, 최적인 근사를 구성하는 저차원 분포의 집합을 결정하는 문제는 남아 있다.

Chow 등은 Lewis가 해결하지 못한 문제를 연구하여 n차원 이진 분포를 최적인 n-1개의 2차원 분포의 곱으로서 잘 근사하는 방법을 제안하였다[12]. 즉, 제안된 방법은 n차원 이산 확률 분포를 최적인 2차원 확률 분포의 곱 또는 1차 의존관계 트리로 근사하는 것이다. n개 변수에서 최적인 n-1개의 1차 의존관계의 집합을

결정하기 위하여 실제 정보와의 차이가 최소로 되는 근사 결과를 결정하는 프로시듀어가 제안되었다 이러한 프로시듀어는 임의의 두 변수 간의 상호 정보의 합을 최대화하는 것(즉, 일치도를 최소화하는 것)과 Kruskal에 의한 최대 가중치 스패닝 트리(Maximum Weight Spanning Tree) 알고리즘을 기반으로 한 것이다.

3. 의존관계를 기반으로 다수 인식기의 결합

다수의 인식기가 독립이라는 가정하에서 사용된 베이저안 공식 기법은 Xu 등에 의해서 제안되었다[4] 본 논문에서는 의존관계가 높은 인식기가 다수 인식기 시스템에 추가될 때 인식기가 독립이라는 가정이 적절하지 않아서 다수의 인식 결과를 결합한 최종 결과가 편향될 수도 있다는 문제점을 지닌다고 본다 또한, 독립이라는 가정을 필요치 않는 BKS 기법은 기하학적인 계산량과 공간을 필요로 할 수 있고, 보지 않은 인식 결과에 대해서 각각하는 문제점을 지닌다고 보았다. 따라서, 이러한 문제점을 극복하기 위하여 인식기 간의 의존관계에 초점을 둔 결합 방법을 제안하고자 한다. 의존관계를 기반으로 다수의 인식기를 결합하는 방법은 두 단계로 구성된다. 첫째 단계는 인식기 간의 의존관계를 측정하여 최적인 의존관계의 집합을 결정하는 것이다 둘째 단계는 결정된 최적의 의존관계의 집합을 기반으로 베이저안 공식을 이용하여 다수의 인식 결과를 결합하는 것이다. 위와 같은 방법을 기술하기 위하여 인식할 대상 클래스를 (M_1, \dots, M_M) 으로 표현하고, 인식기 집합은 $C=(C_1, \dots, C_K)$ 으로, 그리고 입력 패턴은 x 로 표현한다

3.1 의존관계를 결정하는 방법

K개의 인식기에 동일한 입력 x 를 적용하면, K개의 인식 결과(또는 출력)이 발생한다. 즉, $C_k(x) = M_j, k=1, \dots, K, j=1, \dots, M$ 이러한 K개의 인식 결과로부터 BKS 기법에서와 같이 K차원의 결합 확률 분포를 직접 결정할 수 있다. 그러나, 그것은 K가 커짐에 따라서 많은 계산량과 M^K 정도의 공간을 요구하기 때문에 비실용적이라 할 수 있다 또한, 모든 경우의 수를 포함하는 이상적인 K차원의 결합 확률 분포를 구할 수 없는 경우가 종종 발생한다. 이러한 문제점을 극복하는 다른 방법은 확률의 체인 규칙을 기반으로 결합 확률 분포를 아래 수식과 같이 저차원의 확률 분포의 곱으로 변환하여 표현하는 것이다.

$$P(C_1, \dots, C_K) = P(C_1)P(C_2|C_1) \cdot P(C_K|C_1, \dots, C_{K-1})$$

K개의 인식기로부터 정확한 의존관계는 위와 같은

체인 규칙에 따라서 최대 $K - 1$ 차원의 의존관계를 고려해야 한다. 본 논문에서 이러한 의존관계는 동일한 입력에 대한 개별 인식기의 출력을 관찰하므로써 결정될 수 있다고 가정한다. 그러나, 정확한 의존관계를 결정하기 위하여 $K - 1$ 차원의 조건부 결합 확률 분포를 계산하는 것도 마찬가지로 기하학적인 계산량과 공간을 필요로 한다. 그리하여, 대부분의 연구 결과는 인식이 독립이라는 가정을 전제로 1차 확률 분포만을 고려하여 수행되었다. 본 논문에서는 인식이 독립이라고 가정하는 것보다는 1차 의존관계라도 고려하는 것이 바람직하다는 사실을 근간으로 하여, K 차원의 결합 확률 분포를 아래 수식과 같이 1차 의존관계를 기반으로 하는 2차 확률 분포의 곱으로 근사하고자 한다:

$$P(C_1, \dots, C_K) = P(C_1)P(C_2|C_{(1)})P(C_{(K)}|C_{(K)}).$$

위 수식에서 $C_{(k)}$ 는 C_j 와 의존관계를 이루는 부모 인식을 의미한다. 그러면, 문제는 "의존관계를 어떻게 정의하여 표현하는가" 하는 것과 "주어진 인식기의 인식 결과로부터 어떻게 최적인 1차 의존관계의 집합을 결정하는가" 하는 것이 된다. 본 논문에서는 전자의 문제를 해결하기 위하여 두 가지 방법을 사용하였다. 하나는 인식기 간의 상호 정보를 사용하는 것이고, 다른 하나는 통계학에서 사용되는 결합도를 사용하는 것이다. 정의된 의존관계의 측정치를 기반으로 최적인 1차 의존관계의 집합을 결정하는 후자의 문제는 두 분포에 일치도 정의를 적용하여 최적으로 근사되는 인식기 간의 2차 확률 분포를 계산하는 것이다. 이것은 인식기를 노드로 지정하고, 훈련 단계에서 구축된 모든 쌍의 인식기에 대한 2차 결합 확률 분포로부터 계산된 의존관계 측정치를 노드 간의 링크에 대한 가중치로 지정한 그래프에서 Kruskal이 개발한 최대 가중치 스패닝 트리 알고리즘을 적용하여 최대 가중치를 지닌 의존관계 트리를 계산하는 것이다.

3.1.1 상호 정보 (Mutual Information)

K 차원의 확률 분포를 $K - 1$ 개의 2차 확률 분포의 곱으로 근사할 때 미지의 실제 확률 분포인 $P(C)$ 와 아래 수식과 같이 1차 의존관계로부터 근사될 확률 분포 $P(C)$ 의 일치도를 측정하는 상대 엔트로피 $K(P(C), P(C))$ 가 가능한 최소값을 지니도록 $P(C)$ 를 결정해야 한다[11,12]. 이 상대 엔트로피가 0 이면, 관련된 두 분포가 동일하다고 볼 수 있다.

$$P(C) = \prod_{j=1}^K P(C_j|C_{(j)})$$

$$K(P(C), P(C)) = \sum_C P(C) \log \frac{P(C)}{P(C)}$$

$$= \sum_C P(C) \log P(C) - \sum_C P(C) \sum_{j=1}^K \log P(C_j|C_{(j)})$$

$$= \sum_C P(C) \log P(C) - \sum_C P(C) \sum_{j=1}^K \log P(C_j) - \sum_C P(C)$$

$$\sum_{j=1, (j,b) \neq 0}^K \log \frac{P(C_j, C_{(j)})}{P(C_j)P(C_{(j)})}$$

$$= -H(C) + \sum_{j=1}^K H(C_j) - \sum_{j=1}^K M(C_j, C_{(j)})$$

$$H(C) = -\sum_C P(C) \log P(C)$$

$$M(C_j, C_{(j)}) = \sum_{C_j, C_{(j)}} P(C_j, C_{(j)}) \log \frac{P(C_j, C_{(j)})}{P(C_j)P(C_{(j)})}$$

상대 엔트로피 $K(P(C), P(C))$ 을 계산하는 위 수식으로 부터, $K(P(C), P(C))$ 을 최소화하는 것은 모든 쌍의 인식기 C_j 와 인식기 $C_{(j)}$ 에 대해서 상호 정보의 합, 즉 $\sum_{j=1}^K M(C_j, C_{(j)})$ 을 최대화하는 것과 동일함을 알 수 있다. 인식기 $C_{(j)}$ 은 1차 의존관계 트리에서 인식기 C_j 노드의 부모 노드임을 의미한다. 단, 의존관계 트리에서 루트 노드는 부모 노드를 지니지 않는다. $H(C)$ 은 확률 분포 C 에 대한 엔트로피를 의미하며, C_j 은 C 의 구성원이므로 $H(C)$ 와 $H(C_j)$ 은 동일하다고 볼 수 있다. 상호 정보의 정의는 [12,17]에 기술되어 있다. Gallager가 정의한 평균된 상호 정보[17]는 본 논문에서 사용하는 Chow 등이 정의한 상호 정보와 동일한 개념이다. $\sum_{j=1}^K M(C_j, C_{(j)})$ 을 최대화하여 최적인 1차 의존관계의 집합을 결정하고, 곱 근사에 사용되는 2차 확률 분포를 계산하는 절차는 다음과 같다.

1. 동일한 입력에 의한 모든 인식기의 출력 데이터로부터 모든 쌍의 인식기에 대한 2차 결합 확률 분포를 계산한다.
2. 위 결합 확률 분포로부터 모든 쌍의 인식기에 대한 상호 정보의 값을 계산한다.
3. 모든 인식기를 노드로 지정하고, 위에서 계산한 상호 정보의 값을 노드 간의 링크에 대한 가중치로 지정한 그래프를 정의한다.
4. 위 그래프에서 가중치의 합이 최대가 되는 스패닝 트리를 Kruskal의 최대 가중치 스패닝 트리 알고리즘을 적용하여 계산한다.
5. 계산된 최대 가중치 스패닝 트리로부터 최적인 1차 의존관계의 집합을 결정하고, 곱 근사에 사용되는 2차 확률 분포를 계산한다.

3.1.2 결합도 (Measures of Association)

상호 정보가 두 확률 분포 간에 해소될 불확실성의 평균적인 양을 표현한 반면, 결합도는 두 확률 분포 사이의 상관성에 대한 결합의 강도를 표현한다[18]. 상관 계수가 수적인 데이터 분포 사이의 선형 의존관계를 표

현하는 척도인 반면, 결합도는 비수적인 명목상의 데이터 분포 사이의 의존관계를 표현하기 때문에 본 논문에서 사용되는 인식기 간의 의존관계를 측정하는 기준으로 사용하고자 제안한다.

통계학에서 결합도를 측정하는 방법으로는, Pearson χ^2 통계치로부터 계산될 수 있는 Cramer 값 V와 분할(Contingency) 계수 CC 외에, 두 인식기 간의 관계를 예측하는데 있어서 불확실성을 줄이려는 목적으로 고안된 엔트로피 대칭 측정치 E_{sym} , 그리고 예측적 결합도의 계수인 람다 대칭 측정치 λ_{sym} 등이 있다[18] 이들은 모두 0 또는 양의 실수 값을 지닌다. 본 논문에서는 이들 결합도의 측정치가 인식기 간의 의존관계의 강도를 표현한다고 본다 예를 들어, 분할 계수를 제외하고는 최대 결합 강도의 측정치가 1 이 된다. 이러한 결합도를 계산하는 식을 기술하는데 있어서, N은 입력 표본의 갯수, I는 인식기 i에 의한 출력 클래스 표본의 갯수, J는 인식기 j에 의한 출력 클래스 표본의 갯수, O_{ij} 또는 n_{ij} 는 인식기 i와 j에 의해서 관찰되는 결합 출력 표본의 갯수를 표현한다. 그리고, n_{i+} 는 $\sum_j n_{ij}$ 으로, n_{+j} 는 $\sum_i n_{ij}$ 으로, 그리고 E_{ij} 는 인식기 i와 j에 의해서 기대되는 결합 출력 표본의 갯수, 즉 $\frac{n_{i+}n_{+j}}{N}$ 를 의미한다.

$$\chi^2 = \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$V = \sqrt{\frac{\chi^2}{N \cdot \min(I-1, J-1)}}$$

$$CC = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

$$H(i) = - \sum_j \frac{n_{ij}}{N} \log \frac{n_{ij}}{N}$$

$$H(j) = - \sum_i \frac{n_{ij}}{N} \log \frac{n_{ij}}{N}$$

$$H(i, j) = - \sum_i \sum_j \frac{n_{ij}}{N} \log \frac{n_{ij}}{N}$$

$$E_{sym} = 2 \left[\frac{H(i) + H(j) - H(i, j)}{H(i) + H(j)} \right]$$

$$\lambda_{sym} = \frac{(\sum \max_i n_{ij} + \sum \max_j n_{ij} - \max_i n_{i+} - \max_j n_{+j})}{(2 \cdot N - \max_i n_{i+} - \max_j n_{+j})}$$

본 논문에서는 이러한 결합도를 상호 정보 대신에 이 용하여 최적인 1차 의존관계의 집합을 결정할 수 있다고 본다. 즉, 앞서 기술한 절차에서 상호 정보의 값을

결합도의 측정치로 대체하고 가중치가 최대로 되는 의존관계 트리를 계산하면 된다 이들 결합도의 측정치에 대한 크기는 각각 다르지만, 실험에서 계산된 최적의 1차 의존관계의 집합은 대체로 일치함을 보여 주었다

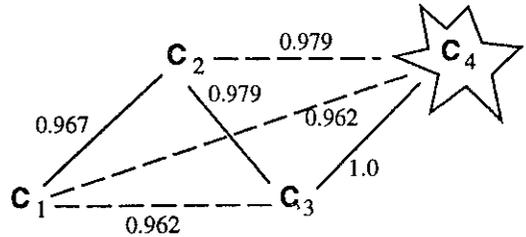


그림 1 의존관계 트리의 예

위의 그림 1은 초기에 3개의 인식기 C_1, C_2, C_3 에서 1차 의존관계를 λ_{sym} 의 측정치로 표현한 다음에, 실선으로 표현된 최적인 의존관계 트리를 구한 것이다 따라서, C_1, C_2, C_3 의 결합 확률 분포는 아래와 같이 근사될 수 있다

$$P(C_1, C_2, C_3) = P(C_2)P(C_3|C_2)P(C_1|C_2)$$

그 다음은 인식기 C_3 와 매우 의존관계가 높은 인식기 C_1 가 추가된 경우에 1차 의존관계를 다시 측정하고 최적인 의존관계 트리를 구한 예이다. 따라서, C_1, C_2, C_3, C_4 의 결합 확률 분포는 아래와 같이 근사될 수 있다

$$P(C_1, C_2, C_3, C_4) = P(C_2)P(C_3|C_2)P(C_1|C_2)P(C_4|C_3)$$

위 식에서 $P(C_1|C_2) = 1$ 이므로 위의 두 결합 확률 분포는 동일하다고 볼 수 있다.

3.2 베이저안 공식의 응용

이상과 같은 방법으로 결정된 최적인 1차 의존관계의 집합과 훈련 단계에서 계산된 1,2차 혼동 확률 분포로부터 하나의 출력 클래스를 기준으로 다른 모든 클래스에 대한 그룹 혼동 확률을 테스트 단계에서 계산할 수 있게 된다 출력 클래스 M_i 에 대한 신뢰도의 값 $Bel(M_i)$ 은 베이저안 정리와 K차 확률 분포를 최적인 1차 의존관계의 확률 분포의 곱으로 근사하는 $P(C) = \prod_{k=1}^K P(C_k|C_{(k)})$ 를 적용하면 아래와 같은 식을 얻게 된다.

$$Bel(M_i) = P(x \in M_i | C_1(x) = M_1, \dots, C_K(x) = M_K)$$

$$= P(x \in M_i) \frac{P(C_1(x) = M_1, \dots, C_K(x) = M_K | x \in M_i)}{P(C_1(x) = M_1, \dots, C_K(x) = M_K)}$$

$$\begin{aligned}
 &= P(x \in M_j) \frac{\prod_{i=1}^K P(C_i(x) = M_j | C_{iD}(x) = M_{iD})}{\prod_{i=1}^K P(C_i(x) = M_j | C_{iD}(x) = M_{iD})} \\
 &= P(x \in M_j) \frac{\prod_{i=1}^K P(x \in M_j | C_i(x) = M_j, C_{iD}(x) = M_{iD})}{\prod_{i=1}^K P(x \in M_j | C_{iD}(x) = M_{iD})}
 \end{aligned}$$

위 수식으로부터, 상수 항을 제거하면,

$$Bel(M_j) = \eta \frac{\prod_{i=1}^K P(x \in M_j | C_i(x) = M_j, C_{iD}(x) = M_{iD})}{\prod_{i=1}^K P(x \in M_j | C_{iD}(x) = M_{iD})}$$

여기서 η 는 $\sum_{i=1}^M Bel(M_i) = 1$ 이기 위한 정규화 계수이다. 만일, 인식기 C_{iD} 이 없으면, 인식기 C_i 은 의존관계 트리의 루트 노드임을 의미한다. 이 경우에는 $P(x \in M_j | C_i(x) = M_j, C_{iD}(x) = M_{iD})$ 을 $P(x \in M_j | C_i(x) = M_j)$ 으로 근사하여 사용한다. 끝으로, 위와 같이 계산된 $Bel(M_i)$ 값을 근거로 입력 x 를 아래와 같은 결정 규칙 $E(x)$ 에 따라 한 클래스로 분류하게 된다.

$$E(x) = \begin{cases} i, & \text{if } Bel(i) = \max_{M \in M} Bel(M) \\ M+1, & \text{otherwise.} \end{cases}$$

4. 실험 결과와 분석

무제약으로 필기된 숫자와 영문 알파벳을 온라인으로 인식하기 위하여 개발된 다수의 문자 인식기를 사용하여 다수 문자 인식기 시스템을 구성하였다. 이러한 문자 인식기는 주어진 입력에 대해서 각 클래스에 대한 일치 정도를 가능성 점수로서 출력한다. 따라서, 이들 점수를 기준으로 클래스의 순위를 결정할 다음에 구현된 다수의 결합 방법을 적용하였다.

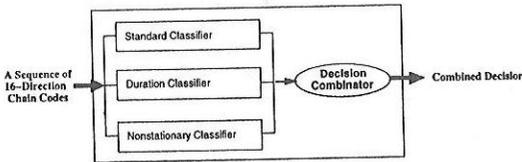


그림 2 Base 다수 문자 인식기 시스템

Base 다수 문자 인식기 시스템 (그림 2)는 최초의 3개 문자 인식기로 구성되는데, 각각 별도의 통계적 훈련 방법 [19]에 따라 만들어 졌다. 사용된 통계적 훈련 방법은 기본 (Standard) 모델링 방법, 지속 (Duration) 모

델링 방법, 비정체 (Nonstationary) 모델링 방법 등이 있다. 동일한 은닉 마르코프 모델 구조를 지닌다 하더라도 다른 통계적 모델링 방법으로 훈련되었다면 다른 문자 인식기라고 가정한다.

의존관계가 매우 높은 문자 인식기의 존재가 다수의 문자 인식기를 결합하는데 끼치는 영향을 살펴보기 위하여, 현존하는 문자 인식기 중의 하나를 위조하여 만들고 이것을 Base 다수 문자 인식기 시스템에 추가하는 방식으로 여러 가지 다수 문자 인식기 시스템을 구축하였다. 예를 들면, 우수 복제형 다수 문자 인식기 시스템은 최초의 3개 문자 인식기와 1등을 기준으로 우수한 문자 인식기를 위조, 복사하여 만든 문자 인식기로 구성되고, 반대로 열등 복제형 다수 문자 인식기 시스템은 열등한 문자 인식기를 위조, 복사하여 만든 문자 인식기로 구성된다.

다수 문자 인식기 시스템에서 다수의 인식 결과는 Decision Combinator에 지정된 결합 방법에 따라 결합되어 우수 클래스가 결정되거나 그렇지 못하게 된다. 모든 문자 인식기의 입력은 일련의 16-방향의 체인 코드를 공통으로 사용한다. 개별 문자 인식기의 출력은 클래스의 순위로 변환되어 Decision Combinator의 입력으로 사용된다. 클래스의 순위를 출력하는 어느 문자 인식기라도 쉽게 본 논문에서 구축한 다수 문자 인식기 시스템의 구성원이 될 수 있다.

유효한 인식 결과를 출력한 문자 인식기만이 Decision Combinator 프로세스에 참가한다. 또한, 유효한 순위를 지닌 클래스만 다수의 인식 결과를 결합하는 과정에서 참조된다. 이러한 제한하에서, Decision Combinator는 많은 결합 방법을 지닐 수 있다. 예를 들면, 투표 기법, 사회적 선택 함수 기법, BKS 기법, 독립을 가정하여 베이지안 공식을 사용한 기법, 본 논문에서 제안한 1차 의존관계를 기반으로 베이지안 공식을 사용한 기법 등이 있다.

투표 기법과 사회적 선택 함수 기법을 적용하는 데에는 훈련 단계가 불필요하지만, BKS 기법과 베이지안 공식을 이용한 기법을 적용하는 데에는 필요하다. 훈련 데이터로는 13명이 작성한 4088개의 숫자, 19명이 작성한 3749개의 소문자와 2464개의 대문자를 사용하였다. 이들 훈련 데이터로부터 BKS 정보와 1.2차 혼동 확률 분포 및 1차 의존관계를 예측하기 위한 2차 결합 확률 분포를 구축하였다. 테스트 데이터로는 10명이 작성한 988개의 숫자, 9명이 작성한 1684개의 소문자와 1169개의 대문자를 사용하였다. 훈련 데이터의 작성자와 테스트 데이터의 작성자는 서로 다른 사람으로 독립적인

테스트를 하고자 했다. 이들 데이터는 작성자가 자유롭게 펼기한 것으로 한국과학기술원 인공지능연구센터가 산학협동과제로 수행한 노드패드 컨소시움에서 온라인 인식을 개발하기 위하여 동 대학, 대학원생들로부터 수집하여 사용한 것이다. 영문 데이터의 경우에는 인쇄체와 필기체 유형이 섞여 있는데, 펼기된 단어에서 문자 단위로 분리된 알파벳도 다수 포함되어 있다. 인식률은 1등에 대해서만 계산하였으며, 테스트 데이터에 대한 각 문자 인식기의 인식률은 아래 표 1과 같다. 이러한 인식 성능을 보유한 문자 인식기로부터 다수 문자 인식기 시스템을 구성하고 테스트용 숫자, 영문 알파벳에 적용하여 실험하였다.

표 1 테스트 데이터에 대한 문자 인식기의 인식률

문자 영역	숫자		소문자		대문자	
	1등	기각률	1등	기각률	1등	기각률
문자 인식기	93.09	0.61	78.92	1.25	88.37	1.80
Standard Classifier	92.28	0.10	82.30	0.53	91.02	0.43
Duration Classifier	94.39	1.32	86.05	2.73	87.51	4.23
Nonstationary Classifier						

표 2 숫자에 대한 다수 문자 인식기 시스템의 인식률

결합 방법	시스템		Base 시스템		열등 복제형		우수 복제형	
	1등	기각률	1등	기각률	1등	기각률	1등	기각률
단순 다수 투표	94.99	0.71	91.98	6.28	94.29	1.92		
절대 다수 투표	94.99	0.71	91.98	6.38	94.29	1.92		
Condorcet 합수	94.99	0.10	91.98	0.10	94.29	0.10		
Borda 합수	95.09	0.10	91.98	0.10	94.59	0.10		
가중화 Borda 합수	95.29	0.10	95.29	0.10	95.19	0.10		
Nanson 합수	94.99	0.10	91.98	0.10	94.29	0.10		
역 Nanson 합수	95.09	0.10	91.98	0.10	94.59	0.10		
BKS 기법	94.09	2.61	94.09	2.61	94.09	2.61		
독립 가정 베이지안	95.29	0.10	95.49	0.10	95.19	0.10		
상호 정보 베이지안	95.39	0.10	95.59	0.10	95.39	0.10		
V 베이지안	95.39	0.10	95.59	0.10	95.39	0.10		
CC 베이지안	95.39	0.10	95.59	0.10	95.39	0.10		
E_{vm} 베이지안	95.39	0.10	95.59	0.10	95.39	0.10		
λ_{ym} 베이지안	95.39	0.10	95.59	0.10	95.39	0.10		

숫자의 실험 결과 (표 1과 표 2를 참조)로부터, 열등하거나 우수한 문자 인식기가 위조, 복사되어 추가된 다수 문자 인식기 시스템에서는 투표 기법과 사회적 선택 합수 기법의 인식률이 저하됨을 보여 준다. 가중화 Borda 합수 방법은 특이하게 열등 복제형 다수 문자 인식기 시스템에서 인식률의 저하가 거의 없음을 보여 준다. 인식률에 있어서 저하되는 정도는 열등한 문자 인식기가 추가된 경우에 더 크게 나타난다. 그러나, 베이지안 공식을 이용하는 기법 중에는 인식 성능의 차이가 거의 없다. BKS 기법의 인식률도 변화가 없음을 보여 준다.

다. 의존관계를 기반으로 다수의 인식 결과를 결합하는 방법이 독립이라 가정하는 방법보다는 인식률이 약간 높으나, t-테스트에 의하면 유의 수준 0.1에서 두 방법에 우열이 있음을 보여 준다. 개별 문자 인식기의 인식 결과로부터 얻을 수 있는 최적의 인식률은 98.10%이다. 최적의 인식률은 1등 클래스를 기준으로 다수의 인식 결과에 입력 클래스가 존재하는 가를 고려하여 계산한 결과이다.

표 3 소문자에 대한 다수 문자 인식기 시스템의 인식률

결합 방법	시스템		Base 시스템		열등 복제형		우수 복제형	
	1등	기각률	1등	기각률	1등	기각률	1등	기각률
단순 다수 투표	84.80	3.80	78.86	8.85	84.62	4.45		
절대 다수 투표	84.80	3.80	78.86	11.40	84.14	7.00		
Condorcet 합수	85.21	0.42	78.74	0.42	84.56	0.42		
Borda 합수	85.27	0.42	82.60	0.42	85.10	0.42		
가중화 Borda 합수	85.39	0.42	85.33	0.42	86.46	0.42		
Nanson 합수	85.33	0.42	80.70	0.42	84.68	0.42		
역 Nanson 합수	85.27	0.42	82.60	0.42	85.10	0.42		
BKS 기법	86.05	7.36	86.05	7.36	86.05	7.36		
독립 가정 베이지안	87.00	0.42	86.34	0.42	87.23	0.42		
상호 정보 베이지안	87.77	0.42	87.77	0.42	87.77	0.42		
V 베이지안	87.77	0.42	87.77	0.42	87.77	0.42		
CC 베이지안	87.77	0.42	87.77	0.42	87.77	0.42		
E_{vm} 베이지안	87.77	0.42	87.77	0.42	87.77	0.42		
λ_{ym} 베이지안	87.77	0.42	87.77	0.42	87.77	0.42		

소문자의 실험 결과 (표 1과 표 3을 참조)로부터, 열등하거나 우수한 문자 인식기가 위조, 복사되어 추가된 다수 문자 인식기 시스템에서는 숫자의 경우와 같이 투표 기법과 사회적 선택 합수 기법의 인식률이 저하됨을 보여 준다. 가중화 Borda 합수 방법은 특이하게 열등 복제형 다수 문자 인식기 시스템에서 인식률의 저하가 거의 없고, 우수 복제형 다수 문자 인식기 시스템에서는 인식률이 향상됨을 보여 준다. 역시, 인식률에 있어서 저하되는 정도는 열등한 문자 인식기가 추가된 경우에 더 크게 나타난다. 의존관계를 기반으로 다수의 인식 결과를 결합하는 방법이 독립이라 가정하여 결합하는 방법보다는 인식률이 약 0.5 - 1.4% 정도 더 높게 나타났다. Base 다수 문자 인식기 시스템에서 의존관계를 고려함으로써 향상된 인식 성능은 유의 수준 0.05로 수행한 t-테스트에서 통계적으로 의미가 있음을 보여 주고, 열등 복제형 다수 문자 인식기 시스템에서는 유의 수준 0.005로 수행한 t-테스트에서, 우수 복제형 다수 문자 인식기 시스템에서는 유의 수준 0.1로 수행한 t-테스트에서, 각각 인식 성능의 향상이 통계적으로 의미가 있음을 보여 준다. 의존관계에 의한 인식률의 향상이 통계적으로 의미가 있음을 검증하기 위하여 열등 복제형 다수

문자 인식기 시스템에서 수행한 t-테스트의 예를 아래에 기술한다

- H_0 : 의존관계를 기반으로 결합하는 방법 (즉, 상호 정보 베이지안)의 평균 인식률이 독립이라 가정하여 결합하는 방법 (즉, 독립 가정 베이지안)의 것보다 같거나 열등하다. (즉, $\mu_D \leq 0$ 여기서 D는 인식률 차이)
- H_a : H_0 에 대한 대안 가설. (즉, $\mu_D > 0$)
- n을 작성자의 수 라고 하면,

$$T = \frac{\bar{D} - 0}{S_D / \sqrt{n}} = 3.51180$$

where

$$\bar{D} = \frac{1}{n} \sum_{i=1}^n D_i = \frac{1}{9} (13.16) = 1.46222$$

and

$$S_D^2 = \frac{1}{n-1} \sum_{i=1}^n (D_i - \bar{D})^2 = 1.56031$$

- $T = 3.51180 > t_{0.05} = 3.355$ 이기 때문에 자유도 8에서 가설 H_0 을 유의 수준 0.005에서 기각할 수 있다. 이것은 의존관계에 의한 인식률의 향상을 99.5%의 신뢰도로 인정할 수 있음을 의미한다.

숫자의 경우와 같이, BKS 기법의 인식률은 변화가 없다 개별 문자 인식기의 인식 결과로부터 얻을 수 있는 최적의 인식률은 90.02% 이다.

표 4 대문자에 대한 다수 문자 인식기 시스템의 인식률

결합 방법	시스템		열등 복제형		우수 복제형	
	Base	사시스템	1등	기각률	1등	기각률
단순 다수 투표	90.33	2.74	88.02	4.45	89.39	4.28
절대 다수 투표	90.33	2.74	87.85	5.56	88.79	5.39
Condorcet 함수	90.59	0.43	87.94	0.43	88.96	0.43
Borda 함수	90.08	0.43	89.14	0.43	89.14	0.43
가중화 Borda 함수	90.16	0.43	89.73	0.43	89.82	0.43
Nanson 함수	90.33	0.43	88.11	0.43	89.39	0.43
역 Nanson 함수	90.08	0.43	89.14	0.43	89.14	0.43
BKS 기법	89.14	6.24	89.14	6.24	89.14	6.24
독립 가정 베이지안	90.59	0.43	89.48	0.43	91.27	0.43
상호 정보 베이지안	90.59	0.43	90.68	0.43	90.59	0.43
V 베이지안	90.59	0.43	90.68	0.43	90.59	0.43
CC 베이지안	90.59	0.43	90.68	0.43	90.59	0.43
E_{nm} 베이지안	90.59	0.43	90.68	0.43	90.59	0.43
λ_{nm} 베이지안	90.59	0.43	90.68	0.43	90.59	0.43

대문자의 실험 결과 (표 1과 표 4를 참조)로부터, 열등하거나 우수한 문자 인식기가 위조, 복사되어 추가된 문자 인식기 시스템에서는 투표 기법과 사회적 선택 함수 기법의 인식률이 약간 저하됨을 보여 준다. 가중화 Borda 함수 방법은 특이하게 열등 복제형 다수 문자 인

식기 시스템에서 인식률의 저하가 거의 없음을 보여 준다. 역시, 인식률에 있어서 저하되는 정도는 열등한 문자 인식기가 추가된 경우에 더 크게 나타난다. 열등 복제형 다수 문자 인식기 시스템에서 의존관계를 고려함으로써 향상된 인식 성능은 유의 수준 0.05로 수행한 t-테스트에서 통계적으로 의미가 있음을 보여 준다. 한편, 우수 복제형 다수 문자 인식기 시스템의 경우에는 독립이라 가정한 결합 방법이 더 높은 인식률을 보이지만, 이것은 유의 수준 0.1로 수행한 t-테스트에서 의미가 없음을 보여 준다. 역시, BKS 기법의 인식률은 변화가 없다. 개별 문자 인식기의 인식 결과로부터 얻을 수 있는 최적의 인식률은 94.70% 이다.

이상과 같은 결과는 앞서 언급한 문제점의 명확한 예를 보인 것이다 따라서, 현존하는 다수 문자 인식기 시스템에 간단히 임의의 문자 인식기를 추가하고, 투표 기법 또는 사회적 선택 함수 기법을 적용하여 다수의 인식 결과를 결합하는 것은 문제가 있다고 본다. 이상과 같은 실험 결과를 분석하면, 다음과 같은 몇 가지 결론을 얻게 된다.

- 다수 문자 인식기 시스템의 인식 성능은 대체로 개별 문자 인식기의 것보다 우수한 편이다.
- 투표 기법과 사회적 선택 함수 기법은 다수의 인식 결과를 결합하는데 있어서 비슷한 성능을 보여 준다.
- 사회적 선택 함수 기법 중에서 가중화 Borda 함수 방법은 열등한 문자 인식기의 추가에도 불구하고 인식률의 저하가 Borda 함수보다 미약함을 보였다. 이것은 가중치가 기여한 결과라고 본다. 즉, 적절한 가중치가 계산되면 Borda 함수보다는 가중화 Borda 함수가 더욱 효과적임을 보이는 예라고 볼 수 있다.
- 베이지안 공식을 사용하는 결합 방법은 대체로 다른 결합 방법보다 우수한 성능을 보였다.
- 의존관계를 고려한 베이지안 결합 방법은 다수 문자 인식기 시스템의 인식 성능을 향상시키는데 기여한다. 특히, 의존관계가 매우 높은 문자 인식기가 추가되는 경우에 그렇다. 영문 소문자에 대한 실험 결과의 표에 대하여 t-테스트 결과가 통계적으로 의미가 있음을 보이므로써 본 논문의 주장을 뒷받침한다고 본다.
- 문자 인식기 간의 의존관계를 고려하지 않고 단순히 문자 인식기를 추가하여 다수 문자 인식기 시스템을 구축하는 것은 문제의 소지가 있다. 왜냐하면, 의존관계가 매우 높은 문자 인식기가 추가되는 경우에 투표 기법이나 사회적 선택 함수 기법을 적용하면 인식 성능이 저하될 수도 있기 때문이다.

- BKS 기법의 인식 성능은 문자 인식이 위조, 복사된 경우에 BKS 정보에 변화를 주지 않으므로 인식률의 변화가 없었다

5. 결론 및 향후 과제

이상과 같은 연구 결과를 통해서 다수의 문자 인식기로부터 인식 결과를 결합하고 다수 문자 인식기 시스템을 구성하는데 있어서 의존관계에 대한 연구의 필요성을 예시해 보였다. 본 논문에서는 표본 데이터로부터 최적인 1차 의존관계의 집합을 결정하는 두 가지 방법을 제안하였다. 하나는 상호 정보를 사용하는 것이고, 다른 하나는 결합도를 사용하는 것이다. 그리고, 최적으로 결정된 1차 의존관계의 집합을 기반으로 베이지안 공식을 이용하여 다수의 인식 결과를 결합하는 방법을 제안하였다 제안된 결합 방법을 무제약으로 펼기된 숫자와 영문자 알파벳의 온라인 인식 문제에 적용하고, 다른 결합 방법과 비교해 보았다. 그 결과, 의존관계가 다수의 문자 인식기를 결합하는 과정에는 물론, 다수 문자 인식기 시스템을 구성하는 과정에서도 고려되어야 함을 보였다

향후 과제는 의존관계를 보다 정확히 결정하기 위한 기준을 연구하고, 1차가 아닌 고차 의존관계를 필요로 하는 상황에서 문자 인식기 간의 고차 의존관계를 잘 근사하는 방법을 연구하는 것이다. 또한, 주어진 다수의 인식 결과로부터 이론적으로 최적인 인식률을 얻기 위한 방법을 연구하여 결합 방법에 의한 인식률을 향상시키는 것이다

참 고 문 헌

- [1] J. J. Hull, A. Commike, and T.-K. Ho, "Multiple Algorithm for Handwritten Character Recognition," in *Proceedings of the 1st International Workshop on Frontiers in Handwriting Recognition*, pp 117-129, 1990
- [2] C. Y. Suen, C. Nadal, T. A. Mai, R. Legault, and L. Lam, "Recognition of Totally Unconstrained Handwritten Numerals Based on the Concept of Multiple Experts," in *Proceedings of the 1st International Workshop on Frontiers in Handwriting Recognition*, pp. 131-143, 1990
- [3] E. Mandler and J. Schuermann, "Combining the classification results of independent classifiers based on the Dempster/Shافر theory of evidence," in *Uncertainty in Artificial Intelligence* (E. S. Gelsema and L. N. Kanal, eds), pp. 381-393, Elsevier Science Publishers, 1988.
- [4] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no 3, pp. 418-435, 1992
- [5] J. Franke and E. Mandler, "A Comparison of Two Approaches for Combining the Votes of Cooperating Classifiers," in *Proceedings of the 11th IAPR International Conference on Pattern Recognition*, vol. 2, pp 611-614, 1992.
- [6] Y. S. Huang and C. Y. Suen, "An Optimal Method of Combining Multiple Classifiers for Unconstrained Handwritten Numeral Recognition," in *Proceedings of the 3rd International Workshop on Frontiers in Handwriting Recognition*, pp. 11-20, 1993.
- [7] T. K. Ho, *A Theory of Multiple Classifier Systems and Its Application to Visual Word Recognition*, PhD thesis, Dept. of Computer Science, SUNY at Buffalo, 1992.
- [8] T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision Combination of Multiple Classifier Systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 1, pp. 66-75, 1994.
- [9] 강희중, 김진영, "의존관계를 기반으로 다수 결정을 결합하는 방법," 한국인지과학회 봄 학술발표 논문집, pp 108-118, 1995.
- [10] H.-J. Kang and J. H. Kim, "Dependency Relationship Based Decision Combination in Multiple Classifier Systems," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pp. 1130-1136, 1995.
- [11] P. M. Lewis, "Approximating Probability Distributions to Reduce Storage Requirement," *Information and Control*, vol 2, pp. 214-225, 1959.
- [12] C. K. Chow and C. N. Liu, "Approximating Discrete Probability Distributions with Dependence Trees," *IEEE Transactions on Information Technology*, vol. 14, no. 3, pp 462-467, 1968
- [13] B. Sin, J.-Y. Ha, S.-C. Oh, and J. H. Kim, "Network-Based Approach to On-line Cursive Script Recognition," Submitted to *IEEE Transactions on Systems, Man, and Cybernetics*, 1994
- [14] H. Tanahi and J. M. Keller, "Information Fusion in Computer Vision Using the Fuzzy Integral," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 3, pp. 733-741, 1990.
- [15] C.-L. Hwang and M.-J. Lin, *Group Decision Making under Multiple Criteria*, Lecture Notes in Economics and Mathematical Systems,

Springer-Verlag, 1987.

- [16] 강희중, 김진형, "다수의 순위를 결정하는 방법과 온라인 문자 인식에의 응용." 한국정보과학회 봄 학술발표 논문집, pp. 213--216, 1995.
- [17] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons, Inc., 1968.
- [18] W. L. Hays and R. L. Winkler, *Statistics: probability, inference, and decision*, Holt, Rinehart and Winston, Inc., 1971.
- [19] B. Sin and J. H. Kim, "Nonstationary Hidden Markov Model," Accepted by Signal Processing, 1995.



강희중

1986년 서울대학교 컴퓨터 공학과(학사).
 1988년 한국과학기술원 전산학과(석사).
 1986년 ~ 현재 삼성전자주식회사 근무.
 1992년 ~ 현재 한국과학기술원 전산학과
 박사과정. 관심분야는 인공지능, 패턴인식,
 집단 의사결정, Groupware, 전문가시스템
 임.



김진형

1971년 서울대학교 공과대학 졸업. 1979
 년 UCLA 전산학 석사학위 취득, 1983년
 UCLA 전산학 박사학위 취득. 1973년~
 1976년 KIST 전산실 연구원, 1976년~
 1977년 미국 캘리포니아 도로국 연구원.
 1981년~1988년 한국정보과학회 산하 인
 공지능 연구회 위원장. 1989년~1990년 미국 IBM 와트슨 연
 구소 초빙 과학자. 1985년~현재 한국과학기술원 전산학과 교
 수. 1991년~현재 인공지능 연구센터 부소장, 1995년~현재
 한국과학기술연구원 연구개발정보센터 소장. 관심분야는 인식
 시스템, 지능형 인터페이스, Computer Aided Education,
 Groupware, 지식기반 시스템임.