

구조적 특징 분석기를 이용한 무제약 필기 숫자 인식기의 결합

A Combination Method of Unconstrained Handwritten Numerals
Recognizers Using Structural Feature Analyzer

김원우[†], 백종현^{††}, 이관용^{†††}, 변혜란^{††††}, 이일병^{†††††}
Wonwoo Kim, Jonghyun Paik, Kwanyong Lee, Hyeran Byun, Yillbyung Lee

요 약

본 논문에서는 구조적 특징 분석을 이용한 무제약 필기 숫자의 검증기를 개발하고 이를 다중 인식기의 결합 알고리즘으로 사용하는 방법을 제안한다. 일반적인 다중 인식기 결합은 학습이나 확률적 방법을 주로 사용하고 입력 영상의 구조적 특징에 대해서는 전혀 고려하지 않기 때문에 인간이 명백히 판단할 수 있는 숫자임에도 불구하고 인식기

† 연세대학교 컴퓨터과학과 석사과정
Dept. of Computer Science, Yonsei Univ.,
134 Shinchon-dong, Seodaemoon-ku, Seoul
e-mail:wonwoo@csai.yonsei.ac.kr

†† 연세대학교 컴퓨터과학과 박사과정
Dept. of Computer Science, Yonsei Univ.,
134 Shinchon-dong, Seodaemoon-ku, Seoul
e-mail:baegjh@csai.yonsei.ac.kr

††† 연세대학교 컴퓨터과학과 연구원
Dept. of Computer Science, Yonsei Univ.,
134 Shinchon-dong, Seodaemoon-ku, Seoul
e-mail:kylee@csai.yonsei.ac.kr

†††† 연세대학교 컴퓨터과학과 조교수
Dept. of Computer Science, Yonsei Univ.,
134 Shinchon-dong, Seodaemoon-ku, Seoul
e-mail:hrbyun@csai.yonsei.ac.kr

††††† 연세대학교 컴퓨터과학과 교수
Dept. of Computer Science, Yonsei Univ.,
134 Shinchon-dong, Seodaemoon-ku, Seoul
e-mail:yblee@csai.yonsei.ac.kr

*)본 연구는 한국과학재단으로부터 연구비를 지원받은 "지능형 문서 처리 시스템"에 관한 연구 중의 일부입니다.

의 특성에 따라 오인식을 할 수 있다. 이런 약점을 보완하기 위하여 자주 혼동되는 숫자쌍에 대하여 구조적 특징을 비교 분석하여 판단하는 일대일 검증기를 구현하고 이를 인식기의 결합에 적용한다. 검증을 위한 구조적 특징으로는 윤곽선, 방향 코드, 다각형 근사와 수직/수평 영교차 횡수 등이 있다. 제안하는 방법의 성능 평가를 위한 실험은 CENPARMI 숫자 데이터를 사용하였으며, 실험 결과 전체 신뢰도는 97.95%를 얻었고 또한 일반적인 결합 알고리즘에서 발생할 수 있는 오인식 요소들이 제거됨을 확인할 수 있었다.

주제어 숫자인식, 구조적 특징 분석기, 다중 인식기 결합 알고리즘

ABSTRACT

In this paper, we design a verifier for unconstrained handwritten numerals using structural feature analysis, and use it as a combination algorithm for multiple recognizers. The existing combination algorithms mainly use learnings, statistical methods, or probabilistic methods without considering structural features of numerals. That is why they cannot recognize some numerals which human can identify clearly. To overcome the shortcomings, we design one-to-one verifiers which compare and analyze the relative structural features between frequently confused numeral pairs, and apply them to combine multiple recognizers. Structural features for verification consist of contour, directional chain code, polygonal approximation, and zero crossing number of horizontal/vertical projections.

We gained a 97.95% reliability with CENPARMI numeral data, and showed that some misconceived factors generated from typical combination algorithms can be removed.

Keywords Numeral Recognition, Structural Feature Analyzer, Multiple Recognizer Combination Algorithm

1. 서론

오늘날 대부분의 문서들은 컴퓨터를 이용하여 작성되며 이런 컴퓨터의 활용은 자료들의 전자 문서화를 지향하여 검색, 수정 등 전반적

인 업무 처리 속도를 매우 향상시키고 있다. 그러나 과거에 수기로 작성된 문서, 혹은 컴퓨터를 이용하기 힘든 곳에서 작성되는 문서의 경우 전자 문서화하기 위하여 재입력을 해야만 한다. 따라서 여기에 투입되는 시간과 비용

을 절감하기 위한 방법으로 문서 자동 입력에 대한 연구가 진행되어 왔다. 특히 숫자는 금액이나 검색키로써 활용되므로 다른 문자들보다 더 중요하다고 할 수 있다. 이런 배경에서 여러 종류의 필기숫자 인식기가 개발되었으며 그들 각각은 독특한 인식방법을 가지고 있다. 초기의 필기숫자 인식 시스템은 하나의 특징 벡터를 이용하는 인식기가 주를 이루었으나, 이런 메커니즘은 필요한 특징이 잘 드러나지 않는 형태를 가진 필기숫자에 대하여는 충분한 정보를 얻을 수 없기 때문에 오인식이 많이 발생할 수 있다. 그래서 더 좋은 성능을 갖는 인식 시스템을 개발하기 위한 연구가 진행되었다. 인식 시스템 성능을 향상시키기 위한 대표적인 연구로서, 둘 이상의 인식기들을 결합한 다중 인식기 시스템이 제안되었다.

다중 인식기 시스템은 여러 개의 인식기를 결합하여 최종 결과를 얻는 것으로 인식 후보가 생성될 수 있으며 인식기 하나에 많은 노력을 들이지 않고 각 인식기 모듈들을 따로 관리하는 등의 효율성 및 성능 향상이 용이하다는 장점을 가지고 있다. 이런 다중 인식기 시스템에서 인식기로부터 얻어진 결과를 조합하는 방법에는 여러 가지가 있으며, 다수결 방식 [1], 베이시안 방식 [1], 보다 함수, 인공 신경망 응용, BKS [2], 그리고 퍼지 적분 방식 [3] 등이 대표적인 예이다. 그러나 이런 방법들은 입력 영상보다는 각 인식기가 제시하는 결과

에 민감하므로 다수의 인식기들이 틀린 결과를 생성하거나 틀린 결과에 대한 출력 값이 옳은 결과에 대한 것보다 높게 나타난다면 오인식할 가능성이 있다. 무제약 필기 숫자의 경우 그런 가능성은 더 높게 나타나는데, 예를 들어 '5'가 입력 데이터일 때 '5'와 '6'은 비슷한 모양을 가지고 있기 때문에 인식기들의 결과에는 '5'와 '6'이 동시에 나타날 수 있다. 만일 다수의 인식기가 '6'을 지지한다면 확률적, 또는 통계적 결합방법에 의한 경우 최종결과는 '6'이 될 것이다.

본 논문에서는, 입력 영상의 구조적 특징을 이용하여 다중 인식기의 인식결과를 결합하는 검증기를 제안, 구현함으로써 인간이 명백히 판단할 수 있는 형태임에도 불구하고 발생할 수 있는 오류를 방지하도록 한다. 검증기는 각 인식기로부터 얻은 결과를 토대로 혼동될 수 있는 숫자쌍으로 분류하고, 입력 영상에 대하여 구조적 특징 분석을 이용한 검증 작업을 통해 올바른 인식결과를 얻어낸다. 이런 과정을 통하여 인간이 보기에 명백하지만 오인식될 수 있는 입력 영상에 대하여 올바른 결과를 제공함으로써 인식 시스템의 성능 향상 및 인간적 요소의 결합을 시도한다.

2. 검증기

2.1 검증기의 종류

검증 방법에는 일반적으로 세 가지 종류가 존재한다. 각 숫자의 고유한 특징을 이용하는 절대 검증, 혼동되기 쉬운 숫자쌍에 대하여 지역적 특징의 차이를 이용하는 일대일 검증, 혼동되기 쉬운 숫자집합에 대하여 행해지는 집단 검증이 있다.[4]

절대 검증은 10 개의 숫자 각각의 고유한 특징만을 고려하므로 인식기들로부터의 결과들을 조합하여 얻은 결과에 대하여 그 결과가 정말 맞는 지를 검증한다. 이 경우 구현해야할 검증 대상 집합의 크기는 10으로 고정되어 있어서 구현이 쉬운 듯 보이지만, 특징이 불분명한 무제약 필기숫자의 고유한 특징을 찾는다는 것은 쉽지 않다. 신뢰도 측면에서 볼 때 고유한 특징이 잘 탐지되지 않는다면 최종결과는 상당 부분 기각으로 처리되어야만 한다.

일대일 검증기는 혼동될 수 있다고 보이는 두 숫자 입력 영상에 대한 상대적인 지역적 특징을 고려하므로[5] 구현이 비교적 쉽다. 상대적인 지역적 특징만의 고려로 검증이 가능하다고 보는 것은, 인간이 혼동되는 숫자를 구별할 때 그들 사이에 나타나는 뚜렷한 특징의 차이를 먼저 고려한다고 보기 때문이다. 그러므로 일대일 검증기는 인간이 혼동되는 대상을 구별하는 메커니즘을 반영한다고 할 수 있다. 또한 여러 인식기들이 서로 다른 결과를 보일 경우 각 인식기들의 결과가 모두 다른 경우는 드물기 때문에 일대일 검증기를 이용하는 것

이 절대 검증기를 사용하는 것 보다 더 적합하다고 볼 수 있다. 그러나 일대일 검증의 단점으로는 혼동될 수 있는 숫자쌍의 가능한 모든 경우가 $({}_{10}C_2) = 45$ 로 구현해야할 경우가 절대 검증에 비해 4.5 배나 많다는 것이다. 하지만 실제 자주 혼동되는 인식 결과는 대략 15 개 정도이며, 혼동되는 경우가 극히 드문 경우들도 상당수 되므로 구현해야할 숫자쌍의 종류는 그리 많지 않다고 볼 수 있다.

집단 검증은, 각 인식기가 동시에 생성하는 결과들의 집합에 대한 정보를 미리 수집하고 같은 집합의 원소 사이에서 발생하는 모든 상대적인 지역적 차이를 고려하여 문제를 해결하기 때문에 다중 인식기의 결과들을 조합함에 있어서 가장 좋은 해결책이라 할 수 있다. 하지만 이것은 검증 대상인 집합 각각의 크기가 불규칙하고 발생할 수 있는 경우를 모두 알아야만 하므로 복잡한 내부 구조를 가지게 되어 구현이 매우 어렵다. 다행히 이런 단점을 극복하기 위한 방법으로서 여러 개의 일대일 검증기와 선형 토너먼트 알고리즘[4]을 사용하여 집단 검증기를 쉽게 구현할 수 있다. 예를 들어 (3, 5, 8)이 혼동되는 집합이라면, 먼저 (3, 5)를 일대일 검증기로 보내서 승자를 선택하고, 만일 '5'가 승자라면 (5, 8)이 다시 검증되는 과정을 통해 크기가 3인 혼동 숫자 집합 (3, 5, 8)에서 올바른 결과를 선택할 수 있다. 본 논문에서는 다중 인식기 시스템에 가장 적합

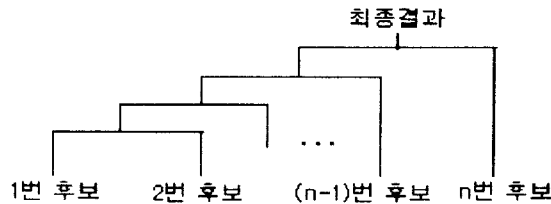
하다고 보이는 일대일 검증기의 확장된 형태인 선형 토너먼트 방식의 집단 검증기를 적용하였다.

2.2. 본 방법에 사용된 구조적 특징

평활화는 이진 입력 영상의 의미 없는 구멍을 메우거나 잡영을 제거하여 골격선을 부드럽게 해주기 때문에 구조적 특징 분석을 용이

때 그림 2 (a)처럼 8-이웃에 7 개 이상의 화소가 있다면 중심을 메운다. 평활화 과정은 그림 2의 (b), (c), (d), (e) 처럼 주로 공간을 메우는 방향으로 진행되는데, 이것은 획이 너무 가늘어서 내부/외부 윤곽선의 경계가 잘 드러나지 않는 경우 윤곽선이 잘못 추출되어 이후의 처리에서 어려움이 생길 수도 있기 때문이다.

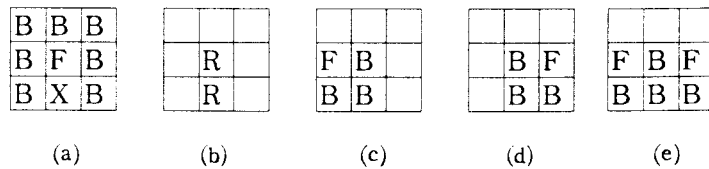
평활화를 수행한 후 몇 가지 구조적 특징을



(그림 1) 선형 토너먼트 알고리즘

하게 해준다. 본 논문에서는 그림 2와 같은 3x3 마스크 윈도우를 회전시켜 가면서 고립화소를 제거하고, 고립 공간을 메워주는 평활화를 수행한다. 만일 윈도우의 중심이 비어 있을

추출하며 이것들로는 윤곽선[6], 방향 코딩[6]을 통한 다각형 근사[7, 8], 수직/수평 주사선의 영교차 횟수[9]와 4-방향 주사선을 이용한 기울기 변화 탐지 등이 있다. 윤곽선 추출은



(그림 2) 평활화를 위한 3x3 마스크 윈도우

X : Don't care, B: 화소, F: 채우기, R: 지울 화소

이진 입력 영상으로부터 4-연결 페리미터[10] 혹은 8-연결 페리미터를 얻는 방법을 이용한다. 4-연결 페리미터의 경우 영상의 각 화소들의 8-이웃이 모두 검은 화소라면 해당 화소를 제거하며, 8-연결 페리미터의 경우는 모든 4-이웃이 모두 검은 화소라면 해당 화소를 제거한다. 이 외의 경우에는 그 화소를 보존한다.

$$R : \text{영역}, N_k(x, y) : \text{화소 } (x, y) \text{의 } k\text{-이웃}$$

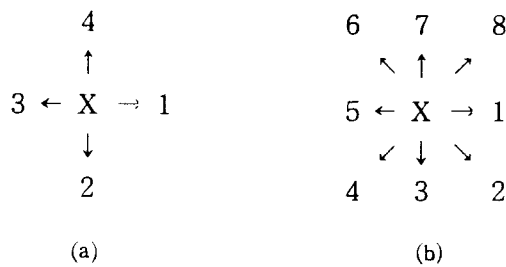
$$P_4 = \{ (x, y) \in R \mid N_4(x, y) - R \neq \phi \} \quad (1)$$

$$P_8 = \{ (x, y) \in R \mid N_8(x, y) - R \neq \phi \} \quad (2)$$

방향 코딩은 추출된 윤곽선의 각 점에 방향 코드를 부여하는 것이다. 만일 윤곽선이 4-연결 요소들이라면 1부터 4까지로 구성된 방향 코드를 이용하며, 8-연결 요소들이라면 1부터

8까지로 구성된 방향 코드를 이용한다(그림 3). 하나의 윤곽선과 마주 보는 반대편 윤곽선과의 거리가 짧아서 발생할 수 있는 방향 코딩의 오류는 에러 코드를 부여하여 사용상의 오류가 없도록 조정한다.

방향 코딩이 된 영상에 대한 다각형 근사는 이진 입력 영상의 형태 윤곽을 구성하고 있는 몇 개의 연결된 획으로 변환해 준다. 다각형 근사를 통해 현재 처리하고 있는 획에 대한 몇 가지 정보와 — 시작점, 끝점, 이전에 연결되었던 획과의 각도 — 오목한 부분과 볼록한 부분의 탐지로 획의 변화를 구체적으로 알 수 있다. 본 연구에서 새로운 획을 결정하는 임계치로 두 획의 사잇각이 $\frac{4}{5}\pi$ 보다 작은 경우를 이용하였다. 때때로, 각도의 변화는 크지만 잡영 등에 의해 획의 진행이 다시 원래 방향으로 돌아오는 경우도 있으므로 획의 분류는 획의 길



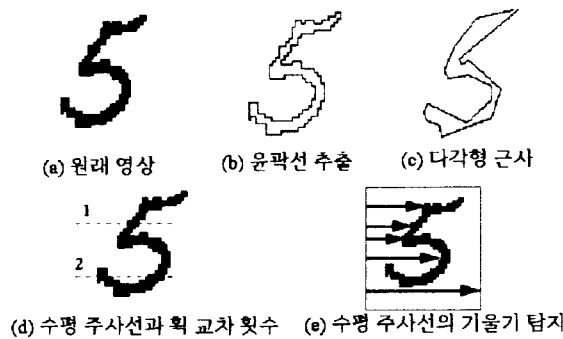
(그림 3) (a) 4-방향 코드와 (b) 8-방향 코드

이가 임계치를 넘어가는 경우에 한한다.

수직/수평 주사선의 영교차 횡수의 탐지는, 초기 입력 영상에 대하여 수직/수평으로 발사된 주사선이 반대편 경계까지 가는 동안 몇 번이나 획과 교차하는지를 알아내는 것이다. 이것은 '1'의 경우 수평 주사선에 대하여 대부분 그 교차 횡수가 1로 나타나지만 '8'은 대부분이 2로 나타난다는 것을 이용하는 것이다. 이런 특징들은 전체 영상에 대하여 행해질 수도

하고, 그 거리의 변화도를 측정하는 것이다. 이것은 단순히 이전에 구한 거리와 현재 구한 거리의 차이가 큰 경우를 탐지 해내는 경우도 있고, 혹은 지금까지 유지되고 있던 거리의 변화보다 더 큰 변화가 있는 경우를 미분을 통하여 탐지할 수도 있다.

다각형 근사와 기울기 탐지 방법이 단순한 지역적 차이를 가지고 있는 숫자쌍에 적용시킬 수 있다고 한다면, 다각형 근사는 복잡한



(그림 4) 제안한 방법에서 사용한 구조적 특징

있고 중요한 특징이 있는 부분에 대하여 행해질 수도 있다. 이것은 (1, 8)이나 (0, 1) 등의 혼동 숫자쌍들을 검증하는데 유용하다.

4-방향 주사선을 이용한 기울기 탐지는, 영상 주위 틀의 상하좌우 각 경계에서 발사된 주사선이 처음 만나게 되는 획까지의 거리를 구

차이를 가진 숫자들을 구분하는데 효과적이다. 다각형 근사를 통해 얻은 획에 대한 정보는, 획의 진행 방향을 추적하면서 혼동될 수 있는 숫자쌍의 지역적 차이로부터 탐지된다.

다각형으로 근사된 획은 연결된 순환 그래프이기 때문에 현재 진행 방향에 비추어 시계 방

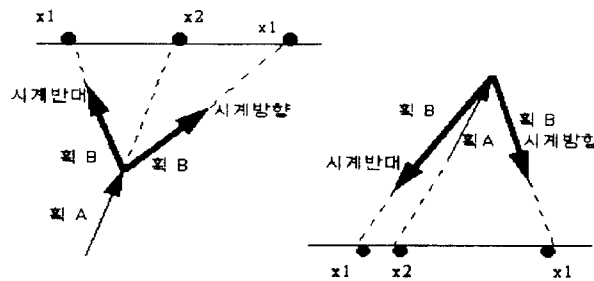
향 혹은 시계 반대방향 중 어느 쪽으로 진행하게 될 지를 아래와 같은 알고리즘(그림5)으로 쉽게 판정할 수 있다.

- 획 A의 끝점은 획 B의 시작점과 같다.
- 획 A와 획 B가 수직이라면 방향 변화는 쉽게 판정될 수 있다.
- 획의 연장선을 긋는다. 연장선상에서 같은 y 좌표에 대해 x 좌표를 비교하여 방향 변화를 판정한다.

재하는 잡영이나 끊어진 획의 구조적 문제점을 모두 포함하지 않을 수 있으므로 특징 탐지에서 일어날 수 있는 오류를 최소화할 수 있다.

언급한 구조적 특징들을 이용하여 빈번히 발생할 수 있는 혼동 숫자쌍인 (7, 9), (3, 8)과 (0, 6)에 대하여 검증 예를 들겠다.

- (1) (7, 9)의 경우 두 숫자에서 턱을 찾아내고 턱 위쪽의 서로 다른 특징을 비교하여



(그림 5) 획의 방향 변화 탐지 알고리즘

다각형 근사는 전체 영상에 대해 행해지는 것이 아니라 지역적 차이를 탐지할 수 있는 부분에 대해서만 적용된다. 왜냐하면 전체 영상에 대하여 다각형 근사를 한다는 것은, 고려되는 대상이 실제로 구별될 수 있는 특징을 가진 곳은 일부 지역이라는 점을 고려할 때 시간의 낭비이다. 또한 특징 지역과 상관없는 곳에 존

검증한다(그림6). 턱의 존재를 찾아내는 방법은 아래쪽에서부터 위쪽으로 올라가면서 수평 주사선을 발사하여 주사선이 검은 화소에 부딪힐 때까지의 길이가 급격히 변하는 부분을 턱으로 정한다. 이 때 '7'의 경우 턱이 상당히 위쪽에서 나타날 수 있는데 이 경우는 이후의 과정을 거치지

않고 곧 '7'로서 검증 결과를 제시한다. 턱을 찾았다면 그 아랫부분을 제거하고 남은 부분에 대하여 수직 주사선을 발사하여 획과의 교차 횟수가 대부분 1이면 '7'로, 대부분 2이면 '9'로 판정한다.

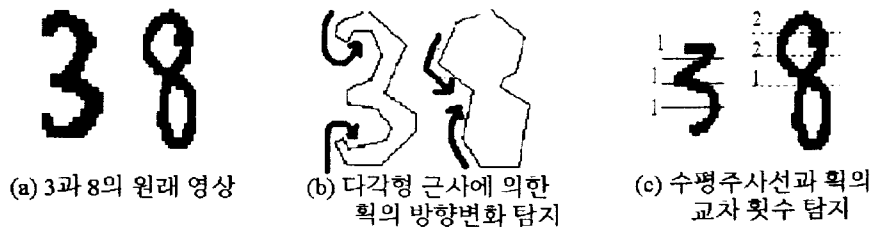
획과 좌하단에서 올라오는 획의 방향이 급격히 변하지만 '8'에선 왼쪽이 닫힌 폐곡선 형태가 강하게 나타나므로 좌상단과 좌하단의 두 획이 그대로 자신들의 진행 방향을 유지한다는 것이 검증의 단서가 될



(그림 6) (7, 9)의 턱 아래를 제거한 이후의 검증

(2) (3, 8)은 입력 영상의 왼쪽이 개방되어 있는지를 조사한다. '3'의 경우 왼쪽이 개방되어 있기 때문에 좌상단에서 내려오는

수 있다(그림7(b)). '3'에서 획 방향 변경이 제대로 탐지되지 않는 경우는 그림 7(c)의 경우이므로 수평 주사선과 획과의 영



(그림 7) (3, 8)의 검증

교차 횡수를 비교하여 그 차이를 탐지한다. 다른 숫자쌍들에게도 그것들이 가지고 있는 비교 조건이 될 수 있는 상대적 차이를 고려하면서 위 방법론을 조금씩 변형시켜 적용하면 숫자쌍들의 구조적 특징 차이를 이용한 검증을 할 수 있다.

(3) (0, 6)은, '6'의 경우 허리 부분에 오른쪽에서 발사된 주사선이 우상단에서 내려오는 회과 만날 때의 거리가 급격히 변하지만 '0'은 그 거리의 변화가 작다는 차이를 이용하여 검증할 수 있다.

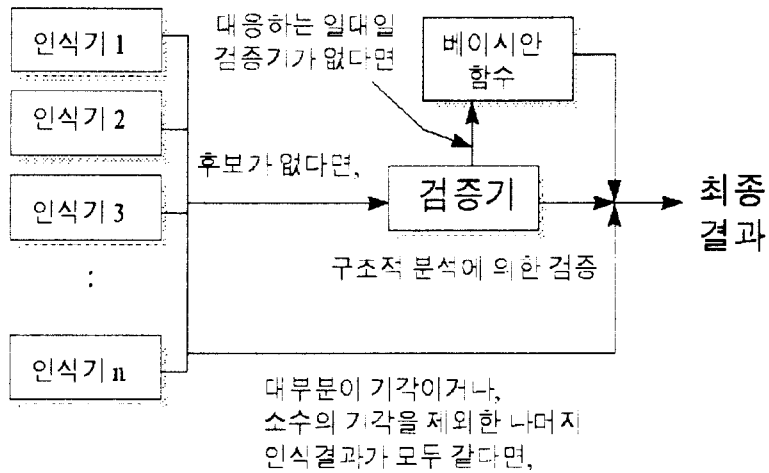
〈표 1〉 혼동될 수 있는 숫자쌍에 대한 검증 방법

숫자쌍	검증 방법
(0, 1) (1, 8)	수평 주사선과 회과의 교차 횡수를 이용
(1, 7)	수직 주사선과 회과의 교차 횡수를 이용
(0, 6)	오른쪽 경계와 영상까지의 거리의 급격한 변화
(1, 4)	① 수평주사선과 회과의 교차 횡수 ② '4'는 아랫쪽에서 위로 올라가면서 왼쪽으로 꺾이는 진행방향을 보임
(2, 3)	우상단에서 아래로 내려오면서 오목점을 만나게 되고 이후의 진행방향을 조사. '2'는 오른쪽으로 진행하려 하지만, '3'은 왼쪽으로 방향을 바꾸려는 성질을 가진다.
(5, 6) (2, 6)	① 좌상단에서 아래로 내려오는 '5'와 '2'는 중간에 오른쪽으로 진행하지만, '6'은 계속 방향을 유지하면서 바닥에서 오른쪽으로 진행. ② '5'는 '6'의 성질을 유지하다가 바닥에서 왼쪽으로 꺾여지는 경우도 있다.
(5, 8) (2, 8)	① 좌하단에서 올라오면서 '5'는 오른쪽으로 갑자기 방향을 바꾸지만 '8'은 그대로 방향 유지. '2'는 좌상단에서 내려오면서 같은 변화를 보인다. ② 허리 아래쪽 영상과 수평 주사선과의 회 교차 횡수 이용
(7, 8) (8, 9)	허리 아래쪽 영상과 수평 주사선과의 회 교차 횡수 이용
(1, 7)	① '7'은 좌상단에서 내려오던 회이 중간 정도 높이에서 갑자기 위로 방향을 바꾼다. ② 좌상단에서 내려오면서 최소한의 진행 각도 변화가 발생했을 때 그 변화가 크다면 '7'로 간주한다.
(3, 5)	① '3'은 좌상단에서 내려오면서 회이 진행 방향이 갑자기 위쪽으로 변경된다. ② '5'는 오른쪽 상단이 개방되어 있기 때문에 오른쪽 경계부터의 거리의 변화가 급하게 나타난다.
(1, 2)	'2'는 상단과 하단에서 회이 진행 방향에 변화가 있다.
(7, 9)	턱을 찾아서 아래부분을 제거한 후 수직 주사선과 회이 교차 횡수를 이용한다.
(3, 8)	좌상단에서 내려오는 회과 좌하단에서 올라오는 회이 방향이 유지상태를 비교한다.

3. 제안하는 결합 방법

제안하는 결합 방법은 n 개의 인식기들의 병렬 조합과 구조적 특징 분석을 하는 검증기, 그리고 베이지안 결합 모듈로 구성된다. 제안하는 방법의 블록 다이어그램은 그림 8과 같다.

을 이용하여 높은 성능을 얻을 수 있기 때문이다. 인식기로 사용한 신경망은 입력 층, 1 개의 은닉층, 출력 층으로 구성되어 있다. 본 연구에서 사용하는 인식기는 모두 5 가지로 필요에 따라 가감될 수 있다. 너무 적은 갯수의 인식기를 이용할 경우 얻는 효과는 적을 것이며 너무 많은 갯수의 인식기를 사용할 경우 각 인식



(그림 8) 제안하는 결합 방법의 블록 다이어그램

3.1 인식기 모듈

본 연구에서 사용한 인식기들은 각기 다른 특징 벡터를 입력으로 갖는 다층 퍼셉트론 분류기로 이루어진다. 이것은 결합 시에 각 인식기의 출력 결과의 일관성을 유지할 수 있을 뿐만 아니라 신경망이 가지는 고유한 학습 기능

기로부터 얻어진 결과들의 조합에 대한 부하가 커지므로 적절한 갯수의 인식기를 사용하는 것도 중요한 문제일 것이다.

3.1.1 N-인식기

입력 영상을 16x16의 크기로 정규화한 영상

을 특징 벡터로 사용한다. 입력 영상의 전역적인 특징을 잘 표현할 수 있으면서 가장 쉽게 얻을 수 있는 특징이다. 신경망의 구조는 256-77-10 이다.

3.1.2 K-인식기

입력 영상의 윤곽선과 윤곽선의 방향 성분은 문자 인식에 가장 높은 효율을 보이는 특징이다. K-인식기는 이처럼 윤곽선의 방향 성분을 얻는데 효율적인 방법인 키르쉬 마스크를 이용한 특징들을 인식에 사용한다. 먼저 입력 영상을 16x16의 크기로 정규화 한 후, 이 영상에 대해 수평, 수직, 좌우 대각선에 대한 방향 성분을 검출한다. 이렇게 형성된 4 개의 16x16의 특정 배열을 4x4의 크기로 압축하고, 정규화된 크기의 초기 영상도 압축하여 5 개의 4x4의 특징 배열을 신경망 인식기의 입력으로 사용한다.[11] K-인식기에서 사용된 신경망의 구조는 80-41-10이다.

3.1.3 G-인식기

그라디언트는 (x, y) 지점에서의 방향 성분의 변화율을 가장 잘 나타내는 성분으로 알려져 있다.[10] 여기에는 두 가지의 중요한 성분이 있는데, 그 하나는 벡터의 절대값이고 다른 하나는 방향 성분이다. G-인식기는 이들 중에서 방향 성분만을 특징으로 취하였다. (x, y) 지점에서 그라디언트 벡터의 방향 성분을 $\alpha(x,$

$y)$ 라고 하면 이는 다음과 같이 나타낼 수 있다.

$$\alpha(x, y) = \tan^{-1}(G_y/G_x) \quad (3)$$

여기서 G_x 와 G_y 는 두 개의 소벨 오퍼레이터 마스크를 사용하여 쉽게 구할 수 있다. 이렇게 주어진 영상에 대한 그라디언트 벡터가 정해지면 12 개의 방향 성분 단위로 나누고, 또한 전체 영상도 16 개의 정방형 영역으로 나누어 각 영역 안에서의 방향 성분의 히스토그램을 구하여 이를 특징 벡터로 사용한다. 이 때 사용된 신경망은 192-77-10의 구조를 갖는다.

3.1.4 4C-인식기와 8C-인식기

방향 코드는 연결되어 있는 패턴의 경계 정보를 표현하는데 널리 쓰인다. 특히 문자 인식에서 방향 성분을 나타내는 좋은 특징으로 알려져 있다. 일반적으로 4-방향 또는 8-방향의 코드가 널리 쓰인다. 본 연구에서는 이들 두 가지의 방향 코드를 사용하는 인식기를 각각 만들어 사용하였다.[6] 방향 코드를 생성한 후 전체 영상을 4x4의 영역으로 분할한 후에 각 영역에서의 방향 코드 히스토그램을 구하여 특징 벡터화 한다. 이를 인식하기 위한 신경망의 구조는 각각 64-35-10, 128-71-10 이다.

3.1.5 인식기의 결과에 대한 처리

검증기를 가진 인식 시스템은 각 인식기로부터 얻은 결과를 검증기로 보낼 것인지에 대해

결정해야 한다. 만일 인식기의 대부분이 동일한 결과를 낸다면 검증 대상이 될 수 있는 후보가 존재하지 않으므로 검증기를 거칠 필요는 없다. 이것에 대한 규칙은 다음과 같다.

- 만일 $(n/2-1)$ 개나 그 이하의 인식기들만이 기각을 하고 나머지 인식기들이 모두 동일한 인식결과를 준다면, 그것을 최종결과로 선택하고 검증기의 처리를 거치지 않는다.
- 만일 $n/2$ 개나 그 이상의 인식기들이 기각을 한다면 최종결과는 기각이며 검증기의 처리를 거치지 않는다.
- 이 외의 경우 각 인식기의 결과를 검증기로 보낸다.

3.2 검증기 모듈

각 인식기의 결과가 동일하지는 않더라도 결과 모두가 서로 다른 경우는 많지 않다. 입력 영상이 동시에 혼동될 수 있는 경우는 대개 2~3 개 정도이기 때문이다. 왜냐하면 입력 영상과 비슷한 특징을 가진 숫자가 혼동되는 대상으로 제시되는 것은 혼동되는 대상과의 공통적인 특징을 가지기 때문이다. 즉, 공통된 특징을 가지지 않는 숫자는 동시에 혼동될 수 있는 다른 대상으로서 나타나기 힘들다는 것을 말해준다. 특히 숫자의 경우 그 모양이 다른 숫자와 구별될 수 있는 특징이 비교적 다른 문자 집합에 비해 강한 편이라서 그런 경향이 더욱 뚜렷하다.

예를 들어 입력 영상이 '5'인 경우 '6'과 혼동되는 경우는 있다. 하지만 '6'과 혼동될 수는 있지만 '5'와 혼동될 소지가 별로 없는 숫자인 '0'이 ('5', '6')과 동시에 나타나는 경우는 거의 없다. 이런 특성 때문에 인식기로부터 얻어진 각 결과들은 몇 개의 그룹으로 크게 나누어질 수 있다. 그룹화된 각 결과를 대응되는 일대일 검증기로 넘기고 일대일 검증기에서 얻어진 결과를 선형 토너먼트 방법으로 계속 조합해가면서 최종결과를 얻는다. 혼동 숫자쌍을 선택한 후 다음 규칙을 적용하여 검증 결과를 얻는다.

- 토너먼트 중간 결과가 기각이라면 최종결과도 기각이다.
- 혼동된 숫자쌍이 자주 발생하는 것이 아니라면 베이시안 방법에 의해 최종결과를 얻는다.
- 이 외의 경우는 검증 과정을 거쳐서 최종결과를 얻는다.

토너먼트 도중 일대일 검증기를 통한 중간 검증 결과가 기각이라면, 올바른 인식 결과가 현재 검증 중인 숫자쌍이 아닌 다른 혼동 숫자쌍에 존재할지라도 최종 결과를 기각으로 처리한다. 때때로 입력 영상은 검증기가 구분하기 힘들 정도로 두 숫자 사이의 상대적인 구조적 특징 차이가 약한 경우가 있다. 올바른 인

식 결과가 앞으로 검증될 숫자쌍에 있을지라도 입력 영상의 구조적 특징이 약한 경우 올바른 인식 결과가 도출되기 어렵기 때문에 검증기가 기각이라고 말하는 것이 당연하다. 다시 말해, 검증 과정에서 만일 기각이 발생하였을 경우, 현재 올바른 인식결과가 현재 검증 중인 숫자쌍에 들어 있을 수 있음에도 불구하고 올바른 결과와 함께 혼동되는 숫자로 간주하여 기각 발생을 무시하고 계속 토너먼트를 진행한다면 더 나쁜 결과를 얻게 될 것이다. 즉 토너먼트 중간에 얻은 중간 검증 결과에서 기각이 발생하였다면 최종결과를 기각하는 편이 안전하다.

혼동될 가능성이 많지 않다고 생각된 숫자쌍, 즉 검증기가 제공하지 않는 숫자쌍에 대해서는 검증기를 거치지 않고 베이지안 결합 방식을 이용하여 최종 결과를 얻는다. 자주 혼동되지 않는 숫자쌍이 인식 결과로 나온다는 것은, 인식기가 입력 영상을 오인식하여 만들 수도 있는 경우이므로 그런 숫자쌍에 대해 구조적 차이를 따질 경우 오인식할 가능성이 더 높아진다. 본 연구에서 제안하는 검증기는 입력 영상의 전역적 특징을 고려하지 않기 때문에 입력 영상과 같지 않은 인식 후보들일지라도 입력 숫자의 구조적 특징을 가지고 있을 수도 있으므로 최종결과에서 오류가 발생할 수 있다. 이런 경우 상대적인 구조적 특징의 고려는 위험하므로 그것을 고려하지 않는 기존의 확률

적 결합 방법에 맡기는 것이 더 나을 수 있다. 자주 혼동되지 않는 숫자쌍이 발생하는 것이 어떤 인식기의 오류였다고 하여도 여러 인식기에서 모두 그런 오인식이 발생하기는 힘들다. 인식기 다수가 올바른 결과를 지지하거나 아니면 각 인식기가 모호한 입력에 대하여 서로 다른 결과를 제시할 것이다. 그러므로 이럴 때 확률적/통계적 결합 방법의 지원을 받으면 구조적 특징에 대한 분석을 하지 않고도 올바른 최종결과를 얻어낼 수 있을 것이다. 본 연구에서는 다중 인식기 시스템의 인식기 조합에서 좋은 성능을 나타내는 베이지안 방식을 이용하여 검증기가 처리하지 못하는 인식기의 결과 조합에 사용하고 있다.

3.3 베이지안 결합 모듈

베이지안 방식은 각 인식기의 오류 가능성이 최종결과에 영향을 미치도록 한다. 즉, 인식기의 가중치를 달리하는 결합 방식이라 할 수 있다. 각 인식기의 오류 가능성은 혼동 행렬로 표현되며, 이를 이용하여 조건부 확률 $P(x \in C_i | e_k(x) = j)$ 를 끌어낸다. 이것은 k -번째 인식기 e_k 가 입력 x 를 j -번째 클래스에 소속시켰을 때 입력 x 가 클래스 i 에 속할 확률을 의미한다. 각 인식기가 상호 독립적이라는 가정 하에 베이지안 방법은 다음과 같이 정의할 수 있다.

$$F_A(e_k(x)) = \begin{cases} j & \text{if } BEL(j) = \max_{j \in A} BEL(i) \geq a \quad (0 < a \leq 1) \\ M+1 & \text{otherwise} \end{cases} \quad (4)$$

여기서(식(4)) a 는 신뢰도의 선택 여부를 결정하는 임계치이다. 이 때 신뢰도 함수는 식(5)와 같이 계산된다.

안된 방법의 성능을 평가하기 위하여 캐나다 Concordia 대학의 CENPARMI 숫자 데이터를 이용하여 인식 실험을 수행하였다.

$$BEL(i) = \eta \prod_{k=1}^K P(x \in C_i | e_k(x) = j_k), \quad \text{for } i=1, \dots, M \quad (5)$$

이 때 η 는 $\sum_{i=1}^M BEL(i) = 1$ 을 만족시키기 위한 정규화 계수이다.

CENPARMI 데이터는 A 그룹, B 그룹, C 그룹 각 2,000 자로 구성되어 있으며, A 그룹은 인식기의 학습 데이터로, B 그룹은 검증기의 구현을 위한 데이터로 사용하였고 C 그룹은 최종 테스트 데이터로 사용하였다. A 그룹 데이터로 학습한 후 C 그룹 데이터로 실험한 각 인식기의 성능은 표 2와 같다.

4. 실험 결과 및 분석

4.1. 실험 환경

본 시스템은 Sun SPARC Classic, Sun OS 4.1.3에서 C 언어를 이용하여 구현되었다. 제

〈표 2〉 CENPARMI 데이터에 대한 인식기의 성능

	인식률	오인식률	기각률	신뢰도
N-인식기	89.05	7.00	3.95	93.00
K-인식기	95.40	3.75	0.85	96.25
G-인식기	95.55	3.10	1.35	96.90
C4-인식기	93.95	4.10	1.95	95.90
C8-인식기	93.80	4.15	2.05	95.85

4.2 검증기 모듈 구성

A 그룹에서 학습된 인식기를 이용하여 B 그룹에 대하여 인식 결과를 얻고, 그 결과 중에서 동시에 두 가지 이상의 다른 결과를 제시하는 경우만을 모은다. 그 경우들의 빈도수를 조사하여 자주 혼동되는 숫자쌍을 수집하고, 각 숫자쌍의 상대적인 지역적 차이를 고려할 수 있도록 고안된 일대일 검증기를 구현한다. 이 실험에서는 모두 15 가지 숫자쌍에 대하여 빈

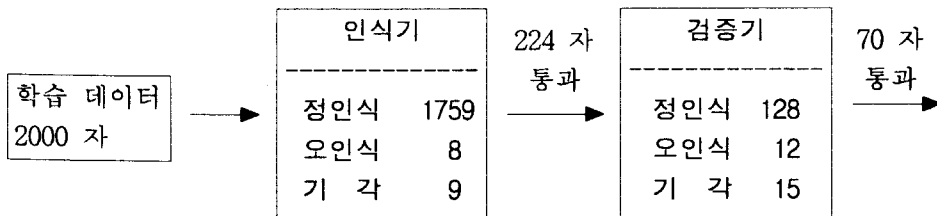
도수가 높은 것부터 우선적으로 구현하였다. 빈도수가 그리 많지 않은 것들 중 빈번히 발생하는 것과 같이 나타나서 자주 혼동될 수 있는 숫자쌍에 영향을 줄 수 있다고 판단되는 낮은 빈도수의 숫자쌍에 대하여도 구현하였다.

4.3 베이시안 모듈 구성

베이시안 모듈도 검증기 모듈과 마찬가지로 A 그룹에서 학습된 5 개의 인식기를 B 그룹에

〈표 3〉 본 방법에서 구현된 혼동 숫자쌍

숫자쌍	빈도수	숫자쌍	빈도수	숫자쌍	빈도수
(3, 8)	10.3	(3, 5)	4.6	(1, 5)	2.3
(7, 9)	9.9	(2, 6)	3.8	(0, 6)	1.9
(8, 9)	8.0	(2, 8)	3.8	(0, 3)	1.5
(2, 3)	7.6	(5, 6)	3.4	(1, 7)	1.1
(5, 8)	6.1	(1, 8)	3.0	(7, 8)	0.4



(그림 9) B 그룹에 대한 검증기 수행 결과

적용하여 얻은 결과를 토대로 구성되며 학습된 베이지안 모듈을 이용하여 C 그룹에 적용시켜 혼동된다고 보이는 숫자쌍을 결합 하였을 때 얻은 결과는 표 4와 같다.

4.4 결과 및 분석

최종 실험은 검증기와 베이지안 모듈을 갖춘

식기들이 오인식된 결과를 제시할 때 올바른 결과를 제시하는 경향이 있어서 그대로 사용하였고 입력 영상의 형태에 의존적이기 때문에 다른 숫자 데이터에서는 더 향상된 성능을 보여줄 수 있다고 본다. 현재 인식기에서 넘겨진 데이터에 대한 검증기의 처리율은 약 68% 정도로 베이지안 모듈로 넘겨지는 숫자쌍이

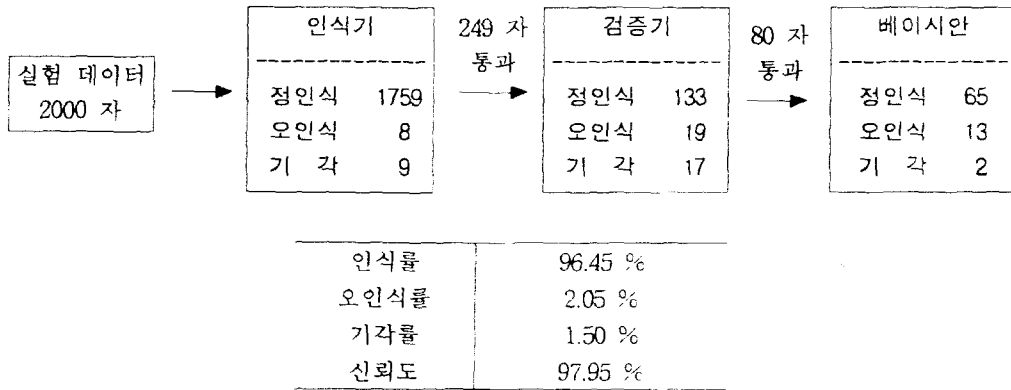
〈표 4〉 CENPARMI 데이터에 대하여 베이지안 방식만을 이용했을 때의 결과

인식률	97.45 %
오인식률	2.35 %
기각률	0.20 %
신뢰도	97.65 %

다중 인식기 시스템에서 C 그룹 데이터를 이용하여 행해졌다. 인식기 모듈에서의 인식 결과는 검증기로 전해지고 구현된 숫자쌍이 아닌 경우 베이지안 함수 모듈로 이동된다. 최종 실험 결과는 그림 10과 같다. 표 2에서 나타나는 것처럼 N-인식기의 성능이 상대적으로 다른 인식기들에 비해 떨어지기 때문에 인식기 모듈의 평균 인식률이 낮아져서 인식기에서 검증기로 넘어오는 데이터의 갯수가 비교적 많은 편에 속한다. 하지만, N-인식기가 다른 인식기보다 낮은 성능을 보일지라도 다른 인

적지 않다고 볼 수 있다.

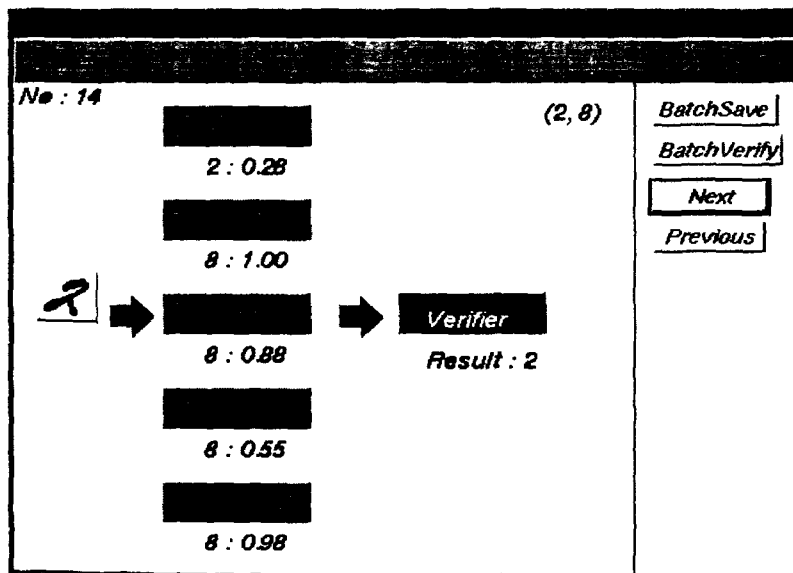
구조적 특징 분석을 통하여 검증기는, 다중 인식기의 확률적 결합 방법의 약점을 보완한다는 것을 보여주었다. 그림 11에서 보이는 것처럼 입력 영상이 비교적 명확하게 '2'로서 나타났음에도 불구하고 5 개의 인식기 중 입력 영상에 대하여 단지 N-인식기만이 '2'라는 옳은 답을 주고있으며 나머지 인식기들은 모두 '8'이라는 오인식된 결과를 제시하고 있다. 또한 N-인식기에 비해 나머지 인식기들이 제시한 자신들의 결과에 대해 0과 1 사이로 정규



(그림 10) 제안한 방법을 통해 얻은 최종 결과

화된 출력값도 매우 높은 편이다. 만일 기존의 확률적/통계적 결합 방법을 사용한다면, 어떤 방법을 이용하여도 최종 결과는 '8'이 될 것이

지만 구조적 특징 분석을 이용한 검증기를 통하여 얻어진 결과는 '2'라고 정확한 결과를 제시하고 있다. 이것은 비슷한 모양을 가진 숫자



(그림 11) 검증기의 실행 화면

에 대하여 인간이 보기엔 명백하게 판단할 수 있음에도 불구하고 틀린 답을 줄 수도 있는 확률적 결합 방법의 약점을 개선하였음을 보여 준다.

5. 결론 및 향후 과제

문자 인식에 대한 연구의 진행에 따라 다양한 특성의 인식기와 인식 시스템이 제안되고 있으며 이 중 대표적인 연구는 다중 인식기를 사용하여 각 인식기로부터의 결과를 조합하여 최종 인식 결과를 얻어내는 것이다. 본 논문에서는 새로운 다중 인식기의 결합 방법으로, 입력 영상의 구조적 특징을 사용한 검증기를 제안하고 이를 무제약 필기 숫자의 인식에 적용하였다. 일반적인 검증기는 하나의 결과에 대해서 참/거짓을 판별하는 것이 목적이지만, 본 논문에서 적용한 검증기는 둘 이상의 후보들에 대해서 가장 가까운 결과를 판정해주는 것을 목적으로 한다. 이는 기존에 사용되고 있는 인식기 결합 알고리즘과 동일한 역할을 한다고 볼 수 있다. 기존의 결합 알고리즘은 주로 학습 또는 확률/통계에만 의존할 뿐, 입력 영상의 구조적 특징은 전혀 고려하지 않는다. 이로 인하여 인간이 명백하게 판단하는 영상에 대해서도 오인식을 하는 예가 존재한다. 이와 같은 경우를 방지하기 위하여 검증기를 이용한 결합 알고리즘을 제시하게 되었다. 검증을

위한 구조적 특징들로는 윤곽선, 방향코드, 다각형 근사, 영교차 횡수등이 사용되었으며, 15개의 혼동되는 숫자쌍에 대한 검증기를 구현하였다. 5개의 인식기가 인식 결과를 내어주면, 일대일 검증과 선형 토너먼트 방식을 결합하여 최종 결과를 얻도록 하였다. 이렇게 구성된 시스템을 CENPARMI 필기 숫자 데이터에 적용하여 97.95%의 신뢰도를 얻었다. 이는 다른 결합알고리즘과 비교해볼 때 매우 높은 수준의 결과는 아니지만, 명확하게 판단 가능한 숫자에 대한 오인식을 최소로 줄였다는 데 큰 의미를 갖는다.

참고 문헌

- [1] L. Xu, A. Krzyzak and C. Y. Suen (1992), "Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition", *IEEE Transactions on Systems and Cybernetics*, Vol. 22, No. 3, pp. 418-435
- [2] Y. S. Huang and C. Y. Suen (1993), "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
- [3] M. Tanahi and J.M.Keller (1990), "Information Fusion in Computer Vision Using the Fuzzy Integral", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 20, No. 3, pp.733-741
- [4] H. Takahashi (1993), "Recognition Enhancement by Linear Tournament Verification", *ICDAR 93*, pp. 585-588
- [5] C. C. Tappert, H.H. Jeanty (1982), "A Study of Several Accuracy Improvement Methods for a Handwriting System", *IBM Journal of Research and Development*, Vol. 26, No. 6, November
- [6] Toshihiro Matsui, Toshio Tsutsumida, Sargur N.Srihari (1994), "Combination of Stroke/Background Contour Direction Features in Handprinted Alphanumeric Recognition", *IWFHR-IV*, December 7-9
- [7] Karin Wall, Peer-Erik Danielsson (1984), "A Fast Sequential Method for Polygonal Approximation of Digitized Curves", *Computer Vision, Graphics, and Image Processing*, Vol. 28, pp.220-227
- [8] Louisa Ram (1988), "Structural Classification and Relaxation Matching of Totally Unconstrained Handwritten ZIP-Code Numbers", *Pattern Recognition*, Vol. 21, No. 1, pp.19-31
- [9] Mitsuru Ohkura, Yasuhiro Shimada, Mitsuru Shiono, Reiji Hashimoto(1993), "On Discrimination of Handwritten Similar KANJI characters by Subspace Method Using Several Features", *ICDAR, 93*, pp.589-592
- [10] Robert M. Haralick (1992), "Computer and Robot Vision", Vol. I, *Addison-Wesley*, pp. 60-61
- [11] Y. J. Kim and S.W Lee (1994), "Offline Digital Image Processing Recognition of Unconstrained Handwrittern Numerals Using Multilayer Cluster Neural Network Combined with Genetic Algorithm", *The Journal of Korea Information Science Society*, pp.1468-1479