

인간의 정보처리 방법에 기반한 특징추출 및 필기체 문자인식에의 응용

Feature extraction motivated by human information processing method and application to handwritten character recognition

윤 성 수* 변 혜 란* 이 일 병*
Sungsoo Yoon, Hyeran Byun, Yillbyung Lee

요 약 본 논문에서는 인간의 정보처리 과정에 관한 심리학적 실험에 바탕을 두고 인간이 사용하고 있는 것으로 생각되는 특징을 이용하여 이를 문자 인식에 적용하였다. 인간의 경우 화소단위의 정보뿐만 아니라 일정지역의 정보를 함께 처리하는 경향이 있다. 그러므로 일정지역에 대한 정보를 표시하는 영역 특징을 정의하고 정의된 이 영역 특징과 기존의 화소단위 특징들을 결합하였다. 사용한 특징으로는 영역 특징에 기반한 초등적 분석결과, 영역특징을 포함한 망 특징, 교차거리차 특징 그리고 기울기 특징들이다. 성능 평가 실험은 필기 한글자모, 숫자 그리고 대소영문자를 대상으로 하였으며, 인식기는 역전파 학습 방법을 이용한 신경망 인식기를 사용하였다. 각각의 인식 결과는 90.27~93.25%, 98.00% 그리고 79.73~85.75%였다. 영역 특징과 유사한 UDLRH 특징을 대상으로 비교한 결과 전체적으로 1~2%정도 인식을 향상이 있었으며 인간이 판단하기에 보다 납득하기 쉬운 오인식 성향을 보였다.

주제어 문자인식, 특징추출

Abstract In this paper, the features which are thought to be used by humans based on the psychological experiment of human information processing are applied to character recognition problem. Man will deal with a little large area information as well as pixel by pixel information. Therefore we define the feature that represents a little wide region information called region feature, and combine the features derived from region feature and pixel by pixel features that have been used by now. The features we used are the result of region feature based preanalysis, mesh with region attributes, cross distance difference and gradient. The training and test data in the experiment are handwritten Korean alphabets, digits and English alphabets, which are trained on neural network using back propagation algorithm and recognition results are 90.27~93.25%, 98.00% and 79.73~85.75%, respectively. Experimental results show that the feature we are suggesting in this paper is 1~2% better than UDLRH feature similar in attribute to region feature, and the tendency of misrecognition is more easily acceptable by humans

1. 들어가면서

인간의 정보처리능력과 컴퓨터의 정보처리능력을 비교해보면 받아들인 정보를 효율적으로 사용하는 것이 매우 중요함을 분명히 알 수 있다. 컴퓨터의 경우 수치에 관한 처리능력은 내부적으로 극히

* 연세대학교 컴퓨터과학과 인공지능연구소
주 소 : 서울 서대문구 신촌동 134 연세대학교 공과
대학 컴퓨터과학과 인공지능연구소 120-749
TEL : 365 - 4598
FAX : 365 - 2579
E-mail : gomdoli@csai.yonsei.ac.kr

자연스럽게 표현될 수 있으며, 이로 인해 인간의 처리능력과 비교도 할 수 없이 일듯하다. 그러나 입력정보가 내부적으로 정확히 어떻게 표현되어야 좋을지 모르는 다른 여러 경우에는 아직도 인간의 정보처리능력을 따라오지 못하고 있다. 문자 인식의 경우에서도 예외는 아니다. 그러므로 입력된 영상정보를 처리하기에 알맞게 가공하는 특징추출 과정은 매우 중요하다. 그러나 아직 인간같이 입력 정보를 효과적으로 표현할 수 있는 방법은 개발되지 못하였고 지금까지의 많은 시도는 극히 수학적인 방법에 의존해오고 있다. 이러한 상황에서 간과할 수 없는 사실이 있는데 그것은 인식결과이다. 인간이 지각하기 힘든 것을 인식할 수 있는 것은 분명 바람직한 일이다. 그러나 그보다 더 중요한 것은 인간이 납득할 수 있는 오류성향을 갖는 것이다. 그래야만 신뢰성을 얻을 수 있기 때문이다. 그러므로 인간의 방법을 고찰해볼 필요가 있다.

간략히 인간의 정보처리과정을 살펴보면, 대략적인 정보 분석이 먼저 이루어진 후, 필요한 정보에 주의를 기울이게 되며, 그런 다음 보다 더 고차적인 과정, 예를 들면 분석이나 인식 같은 과정이 이루어지게 된다[1][2]. 주의작용 전에 사용되는 특징은 물리적인, 즉 주파수 특성이나 물체의 윤곽 같은 기초적 정보가 사용되고 이후에는 의미나 문맥 같은 보다 복잡한 정보가 이용된다. 그러나 이러한 과정들은 단순한 하나의 순차적 과정들이 아니라 서로 서로가 긴밀한 상호작용을 하고 있기 때문에 이런 과정들을 단순화시킨 모형에서는 여러 가지 구현하기 어려운 특성들이 존재하게 된다. 이들 중 특징 추출이란 관점에서 볼 때 크게 두 가지 어려운 점이 있는데, 첫째가 인간의 감각이 갖는 정확성이다. 인간의 대략적 정보 분석은 지금의 컴퓨터 성능과 비교해 볼 때 상대적으로 상당히 정확하며, 만일 오류가 있다 하더라도 아주 빠르게 수정이 일어날 수 있다. 둘째로는 인간의 처리들이 단순 순차적 과정들이 아니라는 것이다. 주의 작용이 일어난 후의 단계는 좀더 복잡적으로 일어나 추가적인 정보들을 받아들여 사용할 뿐만 아니라 인식 결과도 함께 이용된다. 그러므로 본 논문에서는 인간의 정보처리과정 전체를 모형화하지는 못하고 특징추출 관점에서 심리학실험에서 얻어진 인간이 이용하는 특징을 문자 인식에 적용하여 보고자 한다.

일단 주의가 일어날 수 있도록 하는 대략적인

정보를 위하여 '영역 특징'을 정의한다. 영역 특징은 전체적인 윤곽을 표현하는 특징으로 오목한 부분의 형태에 따라 다른 특성값을 갖게 된다. 또 이를 바탕으로 대략적인 분류에 의한 초등적인 인식의 결과를 특징으로 사용한다. 이러한 과정은 특정 인식대상집단에 주의를 환기시키는 기재로 사용된다. 또한 국부적 특성을 갖는 특징들을 함께 사용함으로써 대략적 정보에 기반한 특징들과 서로 보완을 이룰 수 있도록 하였다. 사용된 국부적 특징들로는 영상의 외형의 특징을 나타내는 교차거리차 특징, 영상의 국부적 부분의 기울기를 나타내는 기울기 특징들이다.

2장에서는 각각의 보다 자세한 특징에 대한 설명을 하고, 3장에서는 이 특징들을 구성하는 방법 그리고 4장에서는 특징의 성능평가를 한다. 마지막으로 5장에서 결론을 맺는다.

2. 영역 특징과 추가적 특징

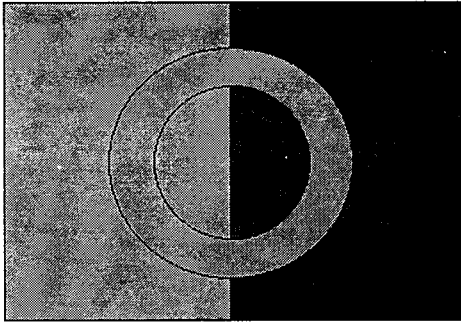
2.1 인간이 사용하는 특징

인간의 정보처리 초기단계에서 일어나는 낮은 수준의 분석은 기초적인 물리적 특징들을 사용하고 있는 것으로 생각되는데, 예를 들면 주파수 특성이나 물체의 윤곽 같은 것이다[1][3]. 그러나 이런 심리학적 실험 결과들은 주의과정의 실체나 사용되는 특징들을 정확히 보여주는 못하는 실정이다. 단지 기초적인 정보를 이용한 대략적인 분석을 통해 주의 작용이 일어난다는 정도이다.

(그림 1)은 우리의 인식 체계 안에서 명암 대비와 형태 지각 상호 간에 어떻게 영향을 미치는지를 보여주는 심리학 실험이다. 왼쪽 사각형이 가장 밝고, 그 다음이 원, 그 다음이 오른쪽 사각형 순으로 명암이 변화하고 있다. 지금은 가운데 있는 원이 모두 같은 밝기로 보이지만 만일 두 직사각형의 경계 위로 연필을 놓으면, 조금 전까지만 해도 같은 밝기를 갖고 있었던 원의 반쪽들이 이제는 서로 다른 명암을 나타내는 것처럼 지각된다.

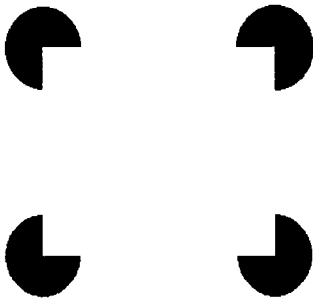
또 (그림 2)는 가운데 직사각형을 만드는 변이 존재하지 않지만 우리는 마치 직사각형이 있는 것 같은 착각을 일으킨다.

위와 같은 여러 심리학 실험들을 보면 인간의 정보 습득 과정과 그 처리가 단순히 수학에서 제시하는 단계 즉, 점을 얻어서 선을 만들고 다시 면을, 그리고 공간을 형성하는 그런 단계를 거친 후



(그림 1) 영역 대비와 형태 지각의 상호작용

이를 종합하여 마침내 처리하는 그런 방식이 아님을 알 수 있다. 생리학적으로 기본적인 단위¹⁾를 얻는 시스템이야 존재하겠지만 이들의 단순한 결합은 아니다. 왜냐하면 (그림 1)의 실험은 가운데 원의 어떤 점이나 선이 변화한 것이 아니라 주의 환경을 변화시켰을 뿐이며 (그림 2)에서도 직접 사각형의 변에 대한 어떠한 조작도 없이 일어나는 결



(그림 2) 주관적 윤곽

과이기 때문이다. 결국 망막 수준의 정보뿐만 아니라 인식 대상이 될 수 있는 근접 공간내의 정보 모두가 함께 통합된 후 기존의 정보와 복합되어 처리되고 있음을 알 수 있다. 이런 결과를 좀 다른 각도에서 바라보면 형태를 나타내는 윤곽에 의해서 만들어진 외부적 또는 내부적 근접 공간을 인식의 기본 특징으로 사용할 수 있음을 알 수 있다. 또 다른 심리학적 결과를 보면 큰 소리, 갑작스러운 움직임, 밝은 색에 대한 처리 등은 기본적으로 주의가 쏠리기 쉽도록 되어 있으며, 물체에서 곡률의

변화가 큰 부분이 인식에 중요한 역할을 담당한다 (4).

이러한 모든 사실들을 종합적으로 고려해 보면 하나의 획에 의하여 큰 곡률을 보이는 넓은 지역의 특성을 나타내는 특징을 생각해 볼 수 있으며, 이를 지금부터는 '영역 특징'이라 정의한다.

2.1 영역 특징

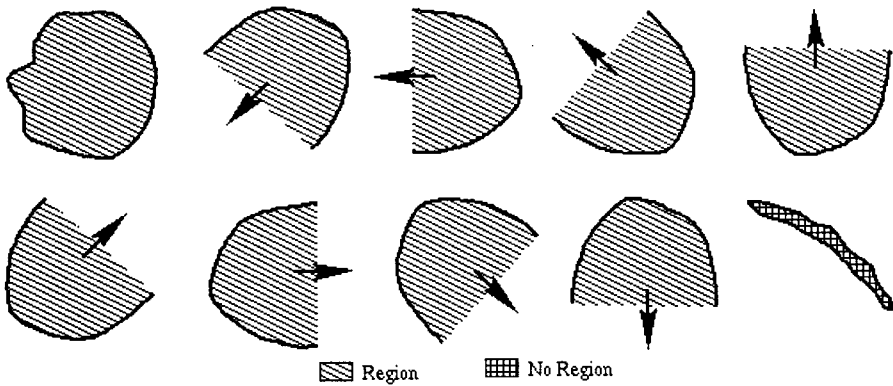
영역 특징은 기본적으로 주의가 쏠리기 쉽고 중요한 정보를 보다 더 많이 갖고 있는 곡률이 큰 부분의 정보를 표현하는 특징이다. 그러나 일반적으로 이용하는 획 자체에 대한 정보가 아니라 앞서 설명한 심리학 실험 결과를 바탕으로 곡률이 큰 부분이 갖는 지역을 정보로 사용한다. 그럼으로써 미약하지만 영역 특징 자체에 주의가 포함되게 된다. 다시 말하면, 인간의 경우는 대략적인 물리적 특징을 이용하여 주의를 가할 부분을 찾게 되지만 본 방법에서는 주의가 가해지는 부분을 미리 곡률이 큰 부분이라 가정하고 이 부분에서 대략적인 정보를 추출하여 사용한다. 간단한 단일 획에 의해서 공간이 나뉘어지는 형태는 폐곡선에 의해 공간이 완전히 두개로 나뉘어지는 경우, 곡선에 의해 반쯤 둘러싸인 하나의 지역이 생기는 경우와 단순히 직선에 의해 공간이 분리되는 경우이다(그림 3).

이때 단일 획에 의한 지역만을 특징으로 사용하는 이유는 변형에는 좀 덜 수용적이지만 단일한 의미의 특징을 얻기 위해서이다. 또 지역 크기가 전체 입력 영상 크기에 비례해서 어느 정도 이상의 것만을 사용하는데 이 또한 특징으로서 분명히 가치 있는 것만을 선택하기 위함이다. 이렇게 만들어진 지역을 변형에 덜 민감한 대략적 특징으로 만들기 위해서 지역을 형성할 때 갖는 방향성에 따라 몇 가지로 구분한다. 즉, 반쯤 둘러싸인 지역을 터져 있는 방향에 따라 8가지 특성으로 나누고, 닫혀 있는 지역에는 구멍 특성을 주고, 어떤 특성도 존재하지 않는 입력에 대해서는 무영역 특성을 갖도록 한다(그림 3).

2.2 영역 특징에 기반한 초등적 분석 결과

정보가 처리되는 과정을 보면 기초적 정보를 분석하여 주의가 가해질 부분을 먼저 찾는다. 그러나 본 방법에서는 미리 주의가 가해질 부분을 정의하고 이곳으로부터 기초적인 정보를 추출하여 이를 바탕으로 간단한 분석을 한다. 원래 간단한 분석은

1) 양서류에 속하는 개구리는 망막 수준에서 빛의 산란, 대조, 어두워짐, 움직임, 경계, 오목성 등을 감지하는 기초적인 세포들을 갖고 있다.



(그림 3) 영역특징의 종류

주의를 두기 위하여 행해지는 처리일 뿐이나 인간의 경우와 달리 이 분석 결과에 전적으로 의존할 수 없기 때문에 보다 세밀한 정보를 위한 특징과 함께 이 분석 결과도 특징으로 사용한다. 이렇게 주의가 가해지는 부분의 정보에 의한 대략적인 분류 내용을 특징에 포함시키는 이유는 주의가 일어나는 과정을 모방하는 한편, 대략적이지만 분명한 특징들에 의한 분류 결과이므로 인식에 도움을 줄 수 있을 것으로 생각하기 때문이다. 초등적 분석 결과를 얻는 방법으로는 이상적인 인식 대상 집단에 대하여 영역 특징을 미리 추출하여 원형 벡터화시킨 후 입력된 영상에서 얻은 영역 특징 벡터와 비교하는 방법을 사용한다. 완전히 일치하는 벡터가 존재하는 경우는 입력 영상이 일치된 원형 벡터의 글자이고 그렇지 않은 경우는 가장 많이 일치하는 벡터의 글자들을 선택한다. 선택된 글자의 특징값은 '1', 그렇지 못한 것은 '0'의 특징값을 갖는다. 영역 특징에서 특징 벡터를 구성하는 방법과 비교 방법 등에 대한 자세한 내용은 3장 주의과정을 이용한 특징구성에서 서술한다.

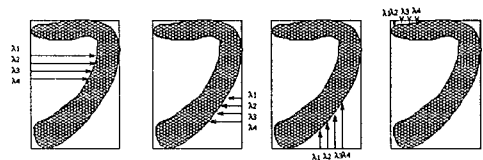
2.3 영역 특징을 포함한 망 특징

망 특징은 대표적인 통계적 특징으로서 각각의 망에 있는 검은 화소의 수를 합하여 특징값으로 이용한다. 그러므로 약간의 잡영에는 상대적으로 강하다[5][6]. 그러나 본 논문에서는 검은 화소뿐 아니라 영역 특징의 특성도 함께 사용한다. 영역

특징의 특성은 구멍, 무영역 특징 그리고 방향에 따른 8가지의 영역 특성들로 이루어지므로 입력된 영상의 검은 화소까지 사용하면 한 화소가 가질 수 있는 특성의 수는 모두 11가지가 된다. 입력 영상을 5×5 로 구분한 후 한 망에서 존재하는 각각의 11가지의 특성을 센 후 망 크기로 나누어 표준화한다. 특징의 수는 $5 \times 5 \times 11 = 275$ 개가 된다.

2.4 교차거리차 특징

교차거리차 특징은 Shridhar[7]와 Heute[8]가 사용하였으며, Shridhar는 왼쪽과 오른쪽의 영상 윤곽의 차만을 사용하였고 Heute는 이를 확장하여 위쪽과 아래쪽 윤곽차도 함께 사용하였다. 본 논문에서는 Heute가 사용한 4방향 모두를 이용한 특징을 사용한다. 교차거리란 영상의 한 면에서 수직/수평선을 그어 검은 화소를 만날 때까지의 거리 $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \dots$ 를 말하며 바로 전 교차거리와의 차 $\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \lambda_3 - \lambda_4 \dots$ 를 교차거리차라 하고, 이런 교차거리차를 $[0,1]$ 사이의 값으로 정규화시킨 값을 특징값으로 사용한다(그림 4).



(그림 4) 교차거리차 특징

이렇게 구해진 교차거리값들을 수평 또는 수직

으로 3등분하고 각 등분의 교차거리값들을 평균하여 특징값으로 사용한다. Heute 논문에서는 입력 영상의 1/5, 1/2, 4/5 되는 지점의 값만을 사용하였으나, 본 논문에서는 변형이 심한 경우에도 수용성을 갖기 위하여 위에서와 같이 평균값을 사용한다. 이 경우에 특징 수는 $3 \times 4 = 12$ 개가 된다. 교차거리차 특징은 입력 영상의 평활성을 나타내는 특징으로 '0'에 가까운 값은 윤곽의 큰 변화가 없음을 의미하고 '1'에 가까운 값일수록 끊어진 윤곽을 나타낸다.

교차거리차 특징은 글자의 윤곽을 추출하는 특징으로 영역 특징의 경우는 곡률이 나타나는 획의 안쪽만을 특징의 대상으로 하기 때문에 바깥쪽은 물론 위쪽과 아래쪽의 윤곽에 대한 정보도 함께 고려할 필요가 있다.

2.5 기울기 특징

기울기는 (x, y) 지점에서 방향 성분의 변화율을 가장 잘 나타내는 값으로 알려져 있으며 국부적 성격이 강하다[9]. x 축, y 축 각각의 방향을 얻는 마스크를 사용하면 x , y 축 각각의 방향 벡터 성분 값을 계산할 수 있고 이 벡터값을 이용해서 기울기 방향과 크기를 모두 얻을 수 있다. 본 논문에서는 방향 성분만을 사용한다. 마스크는 일반적으로 많이 사용하는 3×3 Sobel 연산자를 사용하지 않고 5×5 마스크를 사용한다. 3×3 Sobel 연산자에 비하여 계산량은 많으나 보다 균일한 값을 얻을 수 있다[10]. 이렇게 얻어진 값을 8개의 방향 성분으로 샘플링한 후 입력 영상을 $5 \times 5 = 25$ 망으로 나누어 각 망에서 나타나는 8방향 성분의 개수를 누적한다. 각 성분을 망 크기로 나누어 $[0, 1]$ 사이로 정규화한 값을 특징값으로 사용한다. 특징의 수는 $5 \times 5 \times 8 = 200$ 개가 된다.

기울기 특징의 경우 개개의 화소에 대한 결과 값을 이용하므로 국부적 성격이 매우 강하며 영역 특징이 획 자체에 대한 특성을 반영하기보다는 획에 의해 만들어지는 지역의 모양을 특징으로 삼기 때문에 획 자체에 대한 정보도 포함할 필요가 있어서 사용했다.

3. 주의 과정을 이용한 특징 구성

인간의 초기 정보 처리 과정에 대한 여러 상이한 학설이 있지만 이를 간추리면 다음과 같다. 즉,

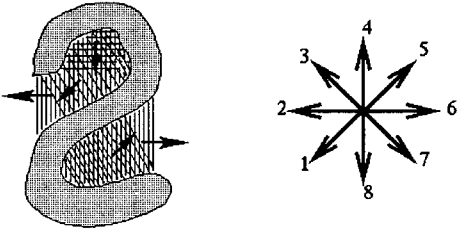
외부에서 받아들인 자극은 아주 짧은 기억 시간을 갖는 감각 레지스터에 일단 저장된 후 보다 고차적인 처리를 위한 단계로 이동되기 전에 정보의 선별이 이루어지는데, 이때 사용되는 특징들은 기초적인 물리적 특징들을 이용하는 것으로 알려져 있다. 이런 선별 과정을 소위 주의 작용이라 부르며, 이때 자극을 어떻게 선별하느냐에 따라서 병목이론, 약화이론, 가용이론 등 많은 상이한 이론들이 존재한다. 그러나 아직까지는 모든 현상을 두루 잘 설명할 수 있는 이론은 없다[1]. 그러므로 완전한 인간의 주의 과정의 모형화는 어려운 실정이며, 또한 인간의 정보처리 과정이 단순 순차적 과정이 아니라 병렬적이고 비순차적인 특징을 갖고 있으므로 특징 추출이란 하나의 과정에 이 모든 과정을 포함시키는 일은 쉬운 일이 아니다. 그러므로 본 논문에서 제안하는 방법에서는 대략적 특징인 영역 특징에 기초한 초등적 분석 과정을 통하여 유사한 특징 집합을 포함한 한 개의 그룹 또는 여러 개의 그룹들을 미리 선별하여 이를 특징에 사용함으로써 주의를 환기시키는 한편, 대략적인 영역 특징을 망 특징에 포함시켜 사용하므로써 주의 작용이 일어나는데 필요한 기초적 정보를 제공한다. 또한 국부적 특징인 기울기 특징과 교차거리차 특징을 사용하므로써 보다 세세한 정보를 제공하는 역할을 한다.

3.1 영역 특징에 기반한 초등적 분석 방법

3.1.1 영역 특징 추출

앞에서 정의한 의미 있는 영역 특징을 추출하기 위해서는 하나의 획에 의해서 만들어진 영역이어야 함으로 4방향 즉, 수직, 수평, 왼쪽 대각선 그리고 오른쪽 대각선 순으로 영역의 시작 위치를 찾는다. 영역의 시작 위치는 각 화소에서 각 방향으로 연장선을 그었을 때 양 끝점에서 검은 화소가 존재하는 곳을 각 방향의 영역 특징의 시작점으로 한다. 4방향 중 어떤 방향의 영역 시작점을 찾은 후에는 시작점에서의 연장선과 평행한 직선을 계속 만들어 가면서 영역의 크기와 위치를 알아낸다. 하나의 영역에서 하나의 영역 특성의 조사가 끝난 후에는 그 특성의 위치, 영역이 터져 있는 방향, 깊이 그리고 면적에 대한 정보를 저장하여 후에 그 영역의 대표 특성을 정하는데 사용한다. 하나의 특성이 찾아진 영역은 통일된 정보를 얻기 위해서 다른 방향에 대한 검사를 행하여 존재하는 특성들을 모두 찾는다.

이렇게 해서 입력 패턴에서 존재하는 영역들을 모두 검출한 이후에, 각각의 영역에서 가장 큰 면적과 깊이를 갖는 영역을 각 영역의 대표 영역 특징으로 정한다. 예를 들면 (그림 5)의 왼쪽에서 보인 것과 같이 각각 영역은 하나 이상의 영역 특징이 존재하고 있고 이들 중 가장 큰 것 즉, ←과 →이 각 영역의 대표 영역 특징이 된다.



(그림 5) 영역특징과 영역특징 배열 순서

3.1.2 초등적 분석을 위한 벡터 구성 방법

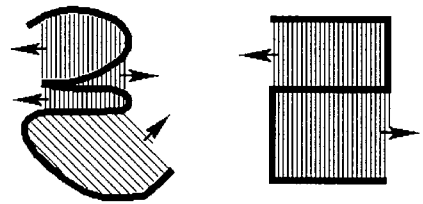
앞에서 얻어진 영역 특징을 초등적 분석에 이용하기 위하여 위치 정보를 고려해서 특징 벡터를 만든다. 영역 특징은 그 특성상 하나의 획 주위에 나타나게 되므로 (그림 5)의 오른쪽에서 보인 방향 순서를 기준으로 찾아진 영역 특징을 특징 벡터화시킨다. 이렇게 영역 특징을 벡터화시키면 왜곡에 의해서 다른 영역 특징이 생기더라도 왜곡 전의 특징들의 순서에는 무관하게 된다. 예를 들어 (그림 5)의 '2'자의 경우에는 두 개의 영역이 존재하고 각 영역은 ←, ↙, ↓과 ↘, →의 특성을 갖게 된다. 이들 중에서 가장 큰 영역은 ←과 →이 되므로 이것들이 각 영역의 대표 특징이 되고 특징 벡터열은 (그림 5)의 화살표 순서에 따라 (← →)이 된다. 만일 많이 왜곡된 '2'자가 입력되더라도 ←과 →의 순서는 뒤섞이지 않는다.

이때 구멍과 무영역 특징은 방향성이 없으므로 구멍의 경우는 그것을 둘러싸고 있는 같은 획에 존재하는 영역 특징들 배열 중 마지막에 삽입하고 무영역의 경우는 가장 가까운 영역 특징들 사이에 끼워 넣는다. 만일 복자음과 같이 독립된 영역들을 갖는 경우를 위하여 각각의 영역 특징은 블록번호를 갖게 된다. 블록번호란 각각의 단일 획에 주어지는 표식을 말한다. 필기체의 경우 그 변형이 매우 다양하므로 단순히 같은 영역 특징이 있다. 없다면으로는 변별력 있는 올바른 판단을 기대하기

힘들다. 그러므로 영역 특징들의 상대적 위치 관계를 위의 예와 같이 고려하여 특징 벡터를 구성함으로써 후에 특징들의 비교에 중요한 정보로 이용한다.

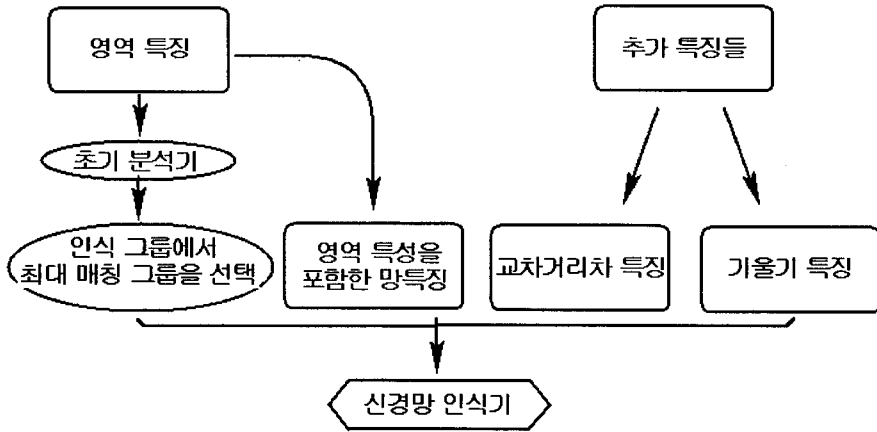
3.1.3 초등적 분석 방법

이상적인 인식 대상 집합에서 영역 특징 벡터를 미리 만들어 놓고 이 특징 벡터들과의 비교를 통해서 가장 유사한 한 개 그룹 또는 여러 개의 그룹들을 선택하는 방법을 사용한다. 이때 미리 추출된 특징 벡터를 앞으로 원형 특징 벡터로 부른다. 원형 특징 벡터는 필기 형태에 따른 변형을 수용하기 위해서 1개에서 많게는 3개까지 원형 벡터를 갖도록 하였다. 이때 원형 특징과의 비교 방법으로는 순서를 고려하면서 같은 특징이 가장 많은 그룹을 선택한다. 이때 같은 특징이라함은 각 영역 특징 벡터안에서 영역 특징이 나타내는 방향이 서로 45도 차이 이내일 경우를 의미한다. 정확히 같은 방향만으로 구분할 수 없는 이유는 필기체의 특성상 어느 정도의 기울기 변화는 수용해야 하기 때문이다. 또한 순서를 고려함이란 두 벡터를 비교할 때 벡터 중간에 불일치 하는 특징값이 있다하더라도 순서가 뒤바뀌지 않는 범위내에서 가장 많이 일치한 특징값을 얻을 수 있도록 비교해 나간다는 것이다. 그 이유는 변형이나 잠영에 의해서 원형과 다른 영역 특징이 존재할 수는 있으나 그 순서들은 변화되지 않기 때문이다. 입력 영상의 변형이나 잠영 정도에 따라서 원형 집합과 완전히 일치하는 경우는 하나의 그룹이 선택되고 그렇지 않은 경우는 가장 많은 특징의 일치를 보이는 한 개의 그룹 또는 여러 개의 그룹들이 선택된다. 예를 들어 들어온 입력 패턴과 이 패턴에 대한 영역 특징이 (그림 6)에서와 같이 주어진다면 입력 패턴은 4개의 영



(그림 6) 입력 패턴과 원형 패턴의 영역특징

역 특징을 갖는 벡터를 구성하게 될 것이고, 이를 원형 특징 벡터와 비교하면 완전히 일치하는 원형



(그림 7) 전체적인 특징 구성도

(표 1) 입력·원형 특징벡터

입력문자	특징벡터
ㄹ	← ← ↗ →

원형 문자	특징벡터
ㄹ	← →
ㄷ	→ →
ㅋ	↙ ←
ㅌ	→ →

은 없고 4개의 비슷한 유형의 원형을 찾게 된다 (표 1).

3.2 전체적인 특징 구성

입력 영상에 대하여 우선 최소로 둘러싸는 직사각형을 찾은 후 각각의 특징들을 추출한다. 우선 영역 특징을 추출하고 이 것에 기반을 둔 초등적 분석 결과와 망 특징을 얻는다. 초등적 분석 결과는 입력 패턴에서 얻은 영역 특징 벡터와 인식하고자 하는 대상 문자 집합의 원형 특징 벡터와의 비교들 통해서 가장 비슷한 벡터(들)를 선택하게 된다. 선택된 벡터의 문자(들)를 나타내는 특징 벡터 위치에 '1'값을 주고 다른 경우는 모두 '0'을 준다. 그런 다음 영역 특징까지 고려한 망 특징으로부터 275개의 특징값을 얻어 특징 벡터에 추가한다. 또한 교차거리차 특징과 기울기 특징을 각각 추출하여 특징 벡터에 함께 포함시켜서 하나의 복합 특징

벡터를 만들어 낸다.

대부분의 특징 추출 방법의 경우에 잡영 제거, 세선화, 크기 정규화 같은 전처리 과정이 추가된다. 그러나 전처리 과정 또한 매우 어려운 주제로 아직 뚜렷하게 만족할 만한 방법이 존재하지 않고 인간의 경우도 전처리를 한다는 분명한 증거도 발견되지 않고 있으므로 본 방법에서는 전처리를 사용하지 않기로 한다. 각각 추출된 특징값들은 영상 자체의 크기에 따라 [0,1] 사이로 정규화 과정을 거쳐서 사용한다. 영역 특징에 기초해 얻은 초등적 분석 결과는 인식 대상 집합에 따라서 특징 벡터의 크기가 변하게 되며 그 밖의 특징들은 모두 벡터 크기가 고정되어 있다. (그림 7)에서 앞서 기술한 내용을 도표로 설명하고 있다. 그림 7에서 초기 분석기란 초등적 분석을 담당하는 모듈이고 인식 그룹이란 인식하고자 하는 대상 문자 집합을 의미한다. 또한 최대 매칭 그룹의 선택이란 벡터 원소들 간의 순서는 유지하면서 두 벡터가 가장 많은 원소들 간의 일치율을 보이도록 하는 매칭 방법에 의해 선택된 문자 집합을 의미한다. 사용된 4개의 특징 집합 중에서 영역 특징에 바탕을 둔 것과 그 이외의 특징을 영역 특징과 추가 특징들로 표시하여 나타내고 있다

4. 실험 및 결과

4.1 실험 범위와 환경

좋은 특징추출 방법은 중요한 여러 다양한 정보

를 잘 찾아 낼 수 있어야 한다. 그러나 하나의 문자 집합에서는 대개 비슷한 유형의 정보가 많이 사용된다. 예를 들면 한글의 경우는 직선적 특징이 많이 나타나고, 영문자의 경우는 곡선적 특징이 복잡하게 결합되어 있으며, 숫자의 경우는 곡선과 직선적 성격이 비교적 간단히 결합되어 있다. 그러므로 제시한 특징 추출 방법의 효용성을 실험하기 위해서는 여러 문자 집합을 대상으로 하는 것이 바람직하다. 또한 필기 문자의 경우가 기울어짐, 형태 왜곡, 크기 변화 등이 보다 다양하므로 성능 평가에 보다 적합하다고 생각되어 성능 평가는 필기체 한글 자모와 숫자 그리고 영문자에 대하여 실험했다. 평가 실험에 사용한 한글 자모는 포항공과대학에서 구축한 필기체 한글 문자 데이터베이스 PE92[11]를 사용했다. 원래 PE92는 음절 단위의 데이터베이스이므로 이들 중 자소를 인위적으로 분리하여 초성 4109자, 수평 중성 770자, 수직 중성 1312자 그리고 중성 3376자를 사용했다. 이들 중 3분의 2를 학습에 사용하였고 3분의 1을 테스트에 사용했다. 숫자 데이터는 캐나다 Concordia 대학의 CENPARMI 데이터베이스를 사용했다. 이 데이터베이스는 전체 A, B, C 세 그룹으로 구성되어 있으며 각 그룹은 0에서 9까지 200벌로 구성되어 있다. 이 중 A그룹과 B그룹을 학습에 사용하고 마지막 C그룹을 테스트에 사용하였다. 영문자의 경우는 NIST CD-ROM의 by_class 디렉토리 밑의 각 영문자 디렉토리에 있는 Special 19번 데이터베이스를 사용했다. 학습 데이터는 이 곳에 있는 학습용 데이터 중 대문자 A부터 Z까지 각각 230자씩, 그리고 소문자 a부터 z까지 각각 230자씩을 학습에 사용하였다. 테스트 데이터도 같은 수를 사용하였다. 실험에 사용된 인식기는 다층 퍼셉트론 구조의 신경망 인식기를 사용하였으며, 은닉층의 수는 180개로 설정했다. 연구는 Linux를 OS로 하는 펜티엄 166Mhz 컴퓨터에서 C언어를 사용하였다.

4.2 실험 내용 및 분석

본 논문에서 제안하는 특징의 성능을 비교하기 위해서 CEDAR에서 사용하였던 특징인 UDLRH 특 [12][13]과 교차거리차 특징 그리고 기울기 특징을 혼합하여 사용한 특징을 비교 기준으로 삼았다. UDLRH 특징은 모든 흰 화소에서 상하좌우 방향으로 주사하여 검은 화소를 만나는 경우에 따라, 각각 위, 아래

왼쪽, 오른쪽, 구멍의 특징값을 주게 된다. 이러한 특징은 영역 특징의 성격과 유사한 점이 많으므로 좋은 비교 대상이라 생각된다. 또한 초등적 분석 결과가 인식 성능에 어떻게 작용하는지를 보이기 위하여 초등적인식결과를 뺀 실험도 함께 했다. 영역 특징을 이용한 초기 분석 결과의 이용에 있어서 보다 정확하게 인간의 정보처리 과정을 흉내내어 실험하지 못하고 단지 하나의 특징으로 나타내어 성능을 평가했다. 그러나 이러한 분석 결과 역시 일정한 규칙을 갖는 형태라면 특징으로서의 역할을 할 수 있으며 실험한 결과 특징으로 사용할 때가 보다 우수함을 알 수 있다. 또한 영역 특징과 유사한 특성을 갖고 있으나 화소 단위의 특징인 UDLRH를 이용하여 비교 실험

(표 2) 실험 내용

실험	내용	특징 크기
UD	UDLRH + 교차거리차 + 기울기	337
NA	영역특징을 포함한 망 + 교차거리차 + 기울기	487
AT	초등적 분석 결과 + NA	문자수 +487

(표 3) 실험에 사용한 한글자소, 숫자 및 영문자 수

자소	학습데이터	테스트데이터
초성	2740	1369
수평중성	513	257
수직중성	875	437
중성	2250	1126
숫자	4000	2000
영문	5980	5980

을 함으로써 결합영향에 의한 특징간의 성능 항상 정도를 수치적으로 제시하였다. 다음 (표 2)는 실험 내용을 요약한 것이고 (표 3)에서는 실험에 사용한 데이터의 수를 보이고 있다.

4.3 필기 한글 자모에 대한 실험

다음 (표 4)에서 실험에 사용한 한글 자모에 대한 인식 성능 평가 결과를 보이고 있다.

전체적으로 보면 UD 특징의 경우보다 영역 특징을 이용한 특징들이 보다 좋은 결과를 보이고 있다. 또한 초등적 분석 결과를 첨가한 특징이 뺀 특징보다 1% 정도 나은 결과를 보이고 있으나, 중성에서는 초등적 분석 결과를 뺀 특징이 약간 더 좋은 결과를 보이고 있는데 이것은 초성보다는 중성

〈표 4〉 한글자소에 대한 인식 성능 평가결과

자소	학습 데이터			테스트 데이터		
	UD	NA	AT	UD	NA	AT
초성	99.96	98.69	100	90.72	91.02	92.33
수평중성	99.81	100	100	88.33	87.94	90.27
수직중성	97.83	96.64	98.51	89.47	90.39	90.62
종성	99.96	100	100	91.92	93.43	93.25

이 왜곡되는 경향이 두드러져 간단한 분석결과가 효율적으로 적용되지 못한 까닭이다. 오인식된 결과를 분석해 보면 영역 특징을 사용한 경우가 보다 납득하기 쉬운 경우가 많았다. 예를 들면 UD 특징의 경우 π 을 α 으로 4개나 오인식하는 반면 AT 특징의 경우는 그런 경우가 없었다. 중성의 경우도 AT 특징에서는 대부분 α 와 β 또는 γ 와 π 등을 구별하지 못한 경우가 대부분인 반면 UD의 경우는 그렇지 않은 경우도 많았다. 자소들간의 결과를 보면, 중성이 초성이나 종성에 비해서 좋지 않은 결과를 보이고 있다. 이러한 이유로는 먼저 학습데이터의 수가 너무 부족한 때문으로 분석되며, 또 다른 이유로는 자음보다 모음이 분명한 영역 특징을 만들기 어렵기 때문으로 생각된다. 예를 들어 수평 모음의 경우 α 나 γ 를 α 로 오인식하는 경우가 상당히 많았으며, 수직 모음의 경우는 α 를 β 로 오인식 하는 경우가 많았다. 이는 평가에 사용된 데이터가 음절 단위 문자데이터베이스에서 추출한 영상을 사용하였기 때문에 위/아래/왼쪽/오른쪽으로 빼친 획의 크기가 작아서 영역 특징을 바로 만들지 못했기 때문이다. 또한 α 와 β , γ 와 π , δ 와 β , η 와 β 쌍에 대한 혼동되는 경향도 역시 두 획간의 간격이 좁고 크기가 작아 바르게 영역 특징을 찾지 못한 것이 오인식의 주요 원인이다.

4.4 필기 숫자에 대한 실험

다음 〈표 5〉에서 실험에 사용한 숫자에 대한 인식 성능 평가 결과를 보이고 있다.

〈표 5〉 숫자에 대한 인식 성능 평가 결과

문자	학습 데이터			테스트 데이터		
	UD	NA	AT	UD	NA	AT
숫자	99.88	100	99.90	97.80	97.75	98.00

숫자의 경우가 3개의 문자 집합 중에서 가장 좋은 결과를 얻을 수 있었다. 이러한 원인은 숫자의 문자 구성이 대부분 하나의 획으로 형성되어 있으며 이러한 특성으로 영역 특징이 만들어지기에 좋

은 조건이 제공되었기 때문이다. 또한 다른 문자 집합과 비교하여 상대적으로 획에 의한 복잡도가 가장 작은 것도 또 다른 원인이다. UD 특징의 인식 결과와 비교해 보면 2를 3으로, 4를 9로 오인식하는 경향이 좀더 많았으나 다른 경우에 대해서는 모두 나은 결과를 얻을 수 있었다. 특히 8자의 경우 UD 특징은 0, 3, 4, 5로 오인식하는 경향을 보인 반면 AT에서는 3과 9로 오인식하는 경향을 보여 보다 납득하기 쉬운 결과를 얻었다. 전체적인 결과를 비교해 보면 AT 특징의 경우는 오인식된 문자 모두가 약간의 변형이지만 큰 영역을 만들만큼 중요한 부분이 변형된 경우의 문자들이 많았으나 상대적으로 UD 특징의 경우는 오인식된 글자와 바른 글자사이에 연관성이 많지 않은 경우가 많았다. 이런 결과는 영역 특징이 국부적 성격의 특징이 아니기 때문에 문자 변형에 의해서 잘못된 특징을 만들더라도 그 전체적인 유사성이 존재하기 때문이다.

4.5 필기 영문자에 대한 실험

〈표 6〉 영문 대소문자에 대한 인식 성능평가결과

문자	학습 데이터			테스트 데이터		
	UD	NA	AT	UD	NA	AT
영문 소문자	99.13	96.86	97.94	79.03	79.82	79.78
영문 대문자	99.60	99.43	99.05	84.10	86.19	85.92

다음 〈표 6〉에서 실험에 사용한 영문자에 대한 인식 성능 평가 결과를 보이고 있다.

영문자의 경우가 가장 좋지 못한 결과를 보이고 있다. 이는 영문자의 특징들이 지금까지의 문자 집합의 것보다 훨씬 다양하고 변형이 많이 일어 날 수 있기 때문이다. 이중 소문자의 인식률이 더욱 좋지 못한데 이런 이유는 문자 구성 방법을 보면 명확히 그 이유가 드러난다. 즉 소문자의 경우는 굴곡이 많으며 획의 복잡도가 대문자에 비해 상대적으로 높고 유사한 형태가 보다 많다. 예를 들면 t와 f, i와 l, g와 q 등이다. 전체적인 인식 평가 결과를 보면 영역 특징을 사용한 특징들이 UD 특징 보다 우수한 결과를 내고 있다. 그러나 어떤 문자에 대해서는 그렇지 못한 경우도 있었는데, 예를 들면 i나 g와 같은 문자이다. 이런 결과는 영역 특징이 단일한 획에 의해 만들어진 넓은 오목한 부분을 찾아 특징으로 사용하는데 반해

서 UDLRH 특징은 화소단위의 특징이기 때문이다. 그러므로 넓은 영역을 제대로 찾을 수 없을 정도로 작은 부분으로 이루어져 있는, 예를 들어 i에서 윗첨, g에서 밑부분의 구멍, m의 밑 부분 같은 부분은 특징을 만들지 못하게 되어 유사한 l, q 그리고 n으로 오인식 되었다. 다른 오인식된 문자들을 살펴보면 한글 모음에서 나타난 문제점처럼 V와 Y의 빠친 획을 구별하지 못하거나 I와 Z에서처럼 기울어진 변형에 너무 둔감한 경우이다. 영문자의 경우 초등적 분석 결과를 첨가시킨 특징이 오히려 인식률이 약간 더 나쁘다. 이러한 이유는 몇 가지 간단한 원형 집합으로 영문자와 같이 복잡하고 변형이 심한 문자 집합을 분류하기가 쉽지 않기 때문이다. 그러므로 원형 특징의 수를 증가시킨다면 좀더 나은 결과를 얻을 수 있을 것이다.

5. 나오면서

본 논문에서는 인간의 정보처리 과정을 흉내내어 보다 효율적인 특징을 제안하였다. 물론 주의 작용을 정확히 흉내내는 것에는 한계가 있었으며, 사용한 영역 특징 또한 인간이 사용하는 특징과 유사하다는 확실한 증거는 없다. 그러나 현재 사용되는 특징들이 극히 수학적·통계적 특징으로 넓은 지역의 특성을 반영하지 못하는 경향이 있으나 인간의 경우는 국부적 특징이외에 전체적 특징을 사용하는 것으로 알려져 있으므로 특징으로 전체적인 특성을 반영하는 특징이 반드시 필요하다. 제안된 특징의 성능을 확인하기 위하여 PE92에서 추출한 한글 자모 3189자와 CENPARMI데이터베이스의 숫자 6000자 그리고 NIST 영문 대소문자 데이터 23920자를 사용하였다. 한글 자모의 경우, 각각 초성 4009자, 수평 중성 770자, 수직 중성 1312자, 종성 2250자중 2/3를 학습에 1/3을 평가에 사용하였으며 각각의 인식률은 92.33%, 90.27%, 90.62% 그리고 93.25%였다. 숫자의 경우, A그룹과 B그룹 4000자를 학습에 사용하고 C그룹 2000자를 평가에 사용하였으며 인식 결과는 98.00%였다. 영문자의 경우 대소문자 각각 5980자를 이용하여 학습하고 테스트하였으며 인식 결과는 대문자 85.92%, 소문자 79.78%였다. 실험결과 영역 특징을 포함시켜 만든 특징이 전체적으로 우수한 성능을 보이고 있을 뿐만 아니라 오인식성향도 국부적 특징에만 의존하는 경우보다 인간이 판단하기에 납득하기 쉬운 특성을 보였다.

참고문헌

- [1] John B. Best(1995), Cognitive Psychology, West publishing company, pp33-73.
- [2] Michael W. Eysenck, Mark T. Keane(1995), Cognitive Psychology A Student's Handbook, Lawrence Erlbaum Associates Ltd., Publishers Vol.19 No.6, pp 142-146.
- [3] 김재갑(1994), 한글 글자 맥락에서의 자모지각, 박사 학위논문, 서울대학교.
- [4] F. Attavee (1971), The multistability of perception, Scientific America 225, 62-71.
- [5] 류강수, 김우태, 진성일(1995), "다중 특징과 모듈화된 신경회로망을 이용한 인쇄체 및 필기체 혼용 숫자 인식", 전자공학회논문지, pp1347-1357.
- [6] 정규식(1995), 문자인식 기반기술연구 권소시업 제1차년도 중간결과 보고서, 한국과학기술원 인공지능연구센터.
- [7] M. Shridhar and A. Badreldin(1984) "Recognition of isolated and connected handwritten numerals", Proc. IEEE International Conference on Systems, Man and Cybernetics, pp 142-146.
- [8] L. Heutte, J. V. Morean, T. Paquet, Y. Lecourtier and C. Oliver(1996), "Combining Structural and Statistical Features for the Recognition of Handwritten Characters", IEEE Proc. ICPR '96 Vol II, pp210-214
- [9] R. G. Gonzalez(1992), R. E. Woods, Digital Image Processing, Addition-Wesley.
- [10] Henning Bässman, Philipp W. Besslich(1995), Ad Oculos Digital Image Processing, Thomson publishing, pp 153-194
- [11] 김대환, 방승양(1992), "한글 필기체 영상 데이터베이스 PE92의 소개", 제 4회 한글 및 한국어 정보처리학술발표논문집, pp567-575.
- [12] J. T. Favata, G. Srikanan, S. N. Srihari(1994), "Handprinted Character/Digit Recognition using a Multiple Feature/Resolution Philosophy", Proc. IWFHR-IV(Taiwan), pp57-66.
- [13] 윤종민, 고대석, 김종렬, 정규식(1996), "필기체 문자 인식에 있어서 특징성능 비교 및 상호보완

적인 특징 선택”, 1996년도 인공지능연구회 춘계
학술발표 논문집, pp39-48.