

## 퍼지소속함수를 이용한 온라인 한글 인식

○  
심영철, 오경환

서강대학교 전자계산학과 인공지능연구실

On-line Recognition in Korean Character Using Fuzzy Membership Function

○  
Young Chul Shim, Kyung Whan Oh

Dept. of Computer Science, Sogang Univ

### 요 약

본 논문에서는 온라인 한글 인식을 위하여 퍼지소속함수를 사용하였다. 획의 오인식으로 인한 문자의 오인식 문제를 해결하기 위하여 인식 시스템 내에서 획을 퍼지집합으로 표현하며, 자모를 인식하는데 사용되는 획의 중심점들 간의 방향은 퍼지소속함수로 정의하여 추론한다.

본 논문에서 제시하는 퍼지추론 방법은 같은 획으로 시작되는 모든 자소에 대하여 적용되며, 애매모호한 상황하에서도 인식을 수행한다.

따라서 퍼지소속함수를 사용한 시스템은 종래에 오인식 되었던 애매한 글자들을 정확하게 인식할 수 있었다.

### I. 서론

온라인(on-line) 문자인식은 스캐너(scanner)에 의한 오프라인(off-line) 문자인식과는 달리 태블릿(tablet), 전자펜 혹은 마우스(mouse)등에 의하여 손으로 씌여진 필기문자를 인식하는 것을 말하며, 오프라인에 비하여 우수한 인식률과 오인식의 발견 및 수정을 즉석에서 행할 수 있는 잇점등 때문에 널리 연구되고 있다.[1,2,4]

입력되는 문자는 문자인식의 기본단위인 획(stroke) 단위로 구분되어 처리되며, 획이란 태블릿 버튼을 누르고 있는 동안 받아들인 점들의 집합을 말한다.[7]

입력되는 획은 태블릿의 위치를 표시하는 X, Y축상의 좌표점들로 표현되며, 이것들은 중복 데이터와 잡음을 많이 포함한다. 따라서 전처리 과정을 통해 인식에 불필요한 데이터들을 제거해야 한다.

문자는 획의 종류, 획의 중심점들 간의 방향에 의하여 인식되며 애매한 데이터의

경우에도 올바른 판단을 하기 위하여 인식 시스템 내에서 획은 퍼지집합으로 표현하고, 획의 중심점들 간의 방향은 퍼지소속함수로 표현하여 추론한다.

추론하는 방법은 같은 획으로 시작하는 모든 자음 또는 모음에 대하여 병렬적으로 연산을 수행하고 그들중에서 최대값(maximum)을 가지는 것을 인식된 것으로 판단한다.

최종적으로 출력되는 문자는 각각 이러한 방법으로 추론된 초성, 중성, 종성을 조합하여 얻어진다.

## II. 획 인식

한글은 조합문자이므로 문자수가 대단히 많고 문자를 구성하는 기본자소의 수도 24개나 되지만 이러한 자소를 구성하는 획은 그림 2.1과 같이 7개 밖에 되지 않으므로 획을 인식하는 과정은 그리 복잡하지가 않다.[3]

획은 일정한 임계치내에 존재하는 좌표들을 소거하여 중복 데이터와 잡음을 제거한 뒤, 정량화하여 인식에 필요한 선분들의 특징벡터(feature vector)를 만들어 인식한다.

이렇게 얻어진 특징벡터를 판단하려고 할 경우 애매한 상황이 많이 발생하며, 이러한 문제를 해결하기 위하여 퍼지소속함수로 표현된 특징벡터를 사용하여 획을 인식한다.[7]

이 때 사용되는 특징벡터는 X좌표 길이의 비, Y좌표 길이의 비, 시작점과 끝점 사이의 길이의 비, 시작선분의 방향정보, 종료선분의 방향정보이다.

## III. 문자인식

문자를 인식하기 위하여 사용되는 특징벡터는 획의 종류와 획의 중심점들 간의 방향이며, 애매한 상황하에서도 효과적인 결론을 얻기 위하여 각각 퍼지집합과 퍼지소속함수로 표현하여 추론하였다.

### 3-1. 퍼지집합으로 표현된 획

입력되는 문자는 획 단위로 구분되어 처리되며, 획을 판단하기에 애매한 경우가 많이 발생한다. 특히 그림 2.1에서 볼 수 있는 바와 같이 a, b, c, d는 각도만 조금씩 다르며, 나머지 특징들은 거의 비슷하기 때문에 서로 오인식되는 경우들이 많이 발생한다. 그리고 획을 쓰는 사람이 잘못 입력하는 경우들도 자주 발생하기 때문에 한글 인식 시스템의 인식률을 높이기 위해서는 획 인식 방법에 치중하기 보다는 어느정도의 애매함과 획의 오인식을 허용하면서 문자를 인식하는 방법이 필요하다.

따라서 본 연구에서는 한글을 구성하는 기본 획들을 각각 퍼지집합으로 표현하고

이러한 집합들 사이의 퍼지추론을 통하여 문자를 인식하고 있다.

표 3.1은 한글의 기본 획들을 퍼지집합으로 표현한 예이다.

|  |
|--|
| $f_a = \{ a/1.0, c/0.8 \}$               |
| $f_b = \{ b/1.0, c/0.8, d/0.8 \}$        |
| $f_c = \{ a/0.8, b/0.8, c/1.0 \}$        |
| $f_d = \{ b/0.8, d/1.0 \}$               |
| $f_e = \{ e/1.0 \}$                      |
| $f_f = \{ f/1.0 \}$                      |
| $f_g = \{ g/1.0 \}$                      |
| $f_t = \{ a/0.8, b/0.8, c/1.0, d/0.7 \}$ |

표 3.1 퍼지집합으로 표현된 한글의 기본획

### 3-2. 퍼지소속함수로 표현된 획의 중심점들 간의 방향

획의 순서만으로는 문자를 인식하기가 힘들다. 예를 들면 '고'와 '가'는 획의 순서는 같지만 다른 글자로 인식되어야 한다. 이러한 문제는 획들간의 위치 정보를 이용하면 쉽게 해결될 수 있다.

하나의 획이 그어지면 그림 3.1과 같이 그 획에 해당하는 영역을 얻을 수 있으며, 이 영역의 중심점을 쉽게 구할 수 있다.

임의의 두 중심점 사이의 방향은 그림 3.2와 같이 8방향으로 나타내어 지며, 각 방향은 그림 3.3과 같이 퍼지소속함수로 표현된다.

### 3-3. 인식 방법

먼저 초성을 인식하기 위하여 같은 획으로 시작하는 모든 자음에 대해 추론을 한다. 이 때 획의 중심점들 간의 방향이 임계치내에 존재하는 획에 대해서만 추론을 하며, 이렇게 계산된 값들중 최대값(maximum)을 가지는 자음을 인식된 초성으로 판단한다.

그 다음 중성과 종성도 각각 위와 같은 방법으로 인식한 뒤, 최종적으로 초성, 중성, 종성을 조합하여 인식된 문자를 얻는다.

## IV. 실험 결과

본 논문에서는 온라인 한글 인식 시스템을 개발하는데 있어서 IBM 386PC와 SummaSketch Tablet을 사용하고 있다.

퍼지소속함수를 사용하지 않은 인식 시스템은 판단을 위한 각 조건들이 TRUE,

FALSE에 의해서만 판단되므로 여러 조건들중 하나만 FALSE가 되어도 글자를 인식하지 못하게 된다. [5, 6]

반면 퍼지소속함수를 사용한 인식 시스템의 경우 각 조건들은 TRUE와 FALSE사이의 퍼지소속정도에 의해 판단되므로 이와같은 문제가 발생하지 않으며, 애매한 상황하에서도 비교적 정확한 판단을 내릴 수 있다.

그림 4.1은 퍼지소속함수를 사용하지 않은 경우 인식할 수 없었던 것들을 본 시스템에서 인식하게 된 예이다.

## V. 결론

본 연구에서는 온라인 한글 인식에 대하여 퍼지소속함수를 적용하였다.

일반적으로 한글 인식 시스템의 인식률을 높이기 위하여 획 인식에 많은 비중을 두고 있지만, 본 시스템에서는 획의 인식이 잘못되더라도 이것을 어느정도 극복할 수 있도록 획을 퍼지집합으로 정의하였고 획의 중심점들 간의 방향을 퍼지소속함수로 정의하여 추론하였다.

따라서 기존의 온라인 한글 인식 시스템들에서 발생되었던 오인식 패턴들을 정확하게 인식할 수 있었다.

## 참고문헌

- [1] 전병환, 구본석, 김성훈, 김재희 “퍼지이론을 응용한 펜 컴퓨터에서의 On-line 획인식 기법”, 한국 퍼지시스템 연구회 춘계 학술회의 논문집, p168-176
- [2] 한국정보과학회, “온라인 필기 인식 기술의 현황”, 정보 과학회지, VOL.9, NO.1, P54-64
- [3] 손영우, “속성에 의한 한글 패턴의 온라인 인식에 관한 연구”, 광운공과대학교 대학원, 1982년
- [4] 장덕문, “속성 관계 그래프를 이용한 한글의 구조적 인식”, 서강대학교 대학원, 1989년
- [5] 변중남, 이광형, “퍼지이론을 이용한 자동제어기술 및 전문가 시스템”, 한국과학기술원 산학협동 공개강좌
- [6] 박민용, 최항식, “획지 시스템의 응용입문”, 대영사
- [7] 심영철, 오경환, “퍼지소속함수를 이용한 온라인 한글 획 인식”, 91' 추계 학술 발표 논문집, 한국정보과학회, 1991, 10, 홍익대

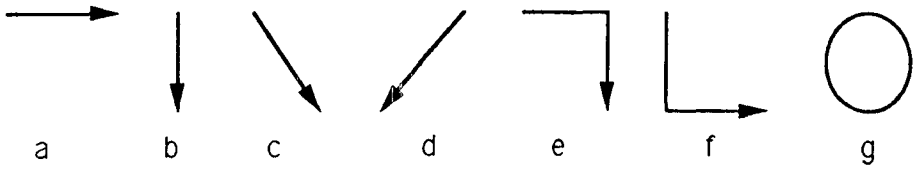


그림 2.1 한글 자소를 구성하는 획의 종류

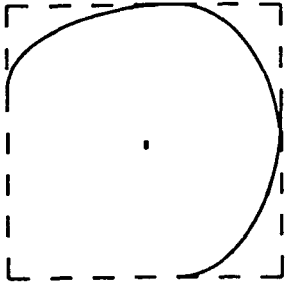


그림 3.1 획의 영역과 중심점

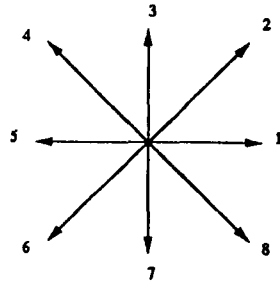
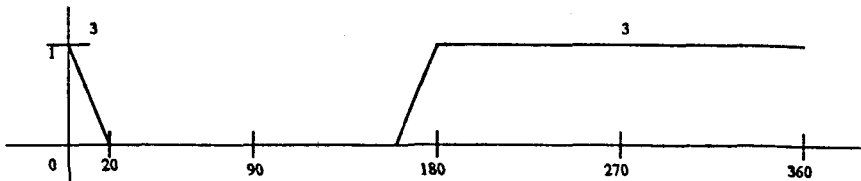
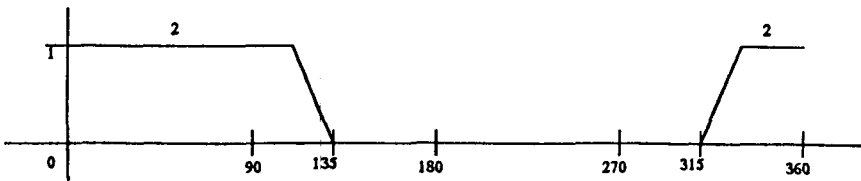
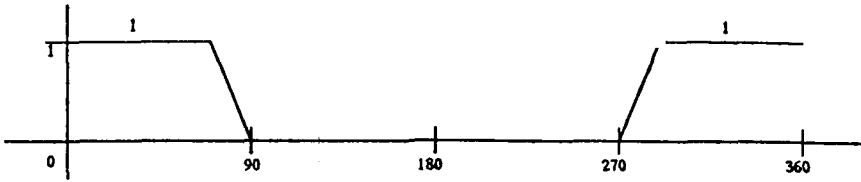


그림 3.2 임의의 두 중심점 사이의 방향



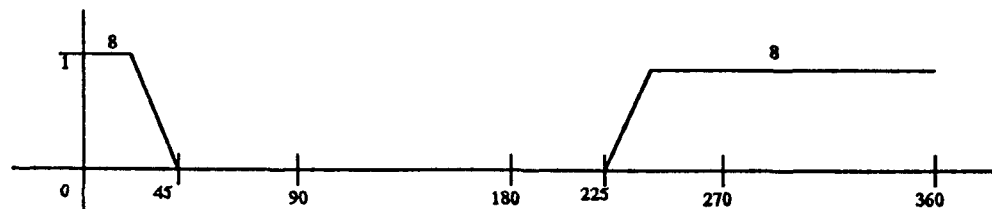
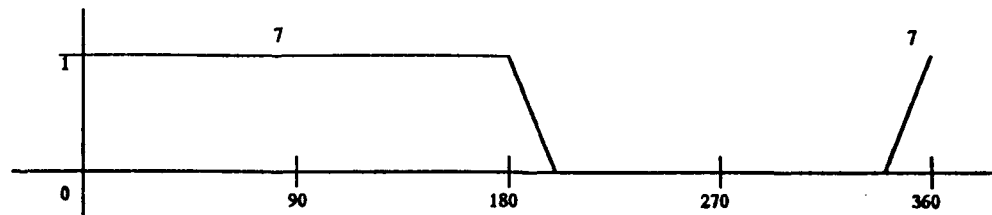
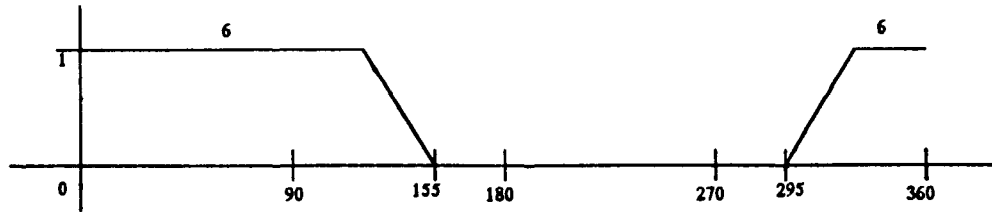
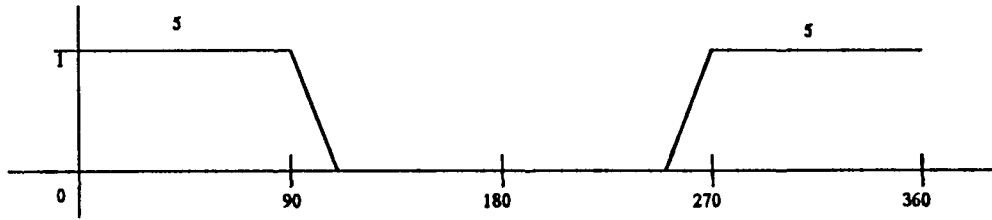
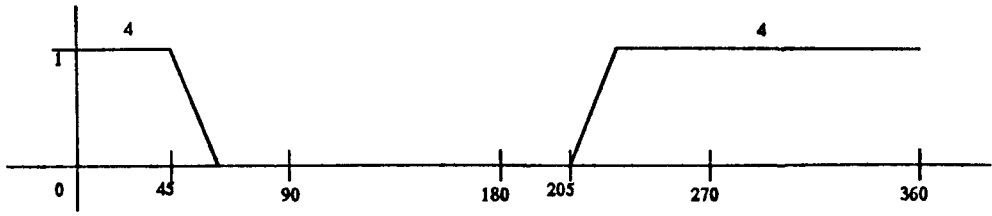


그림 3.3 퍼지소속함수로 표현된 임의의 두 중심점 사이의 방향

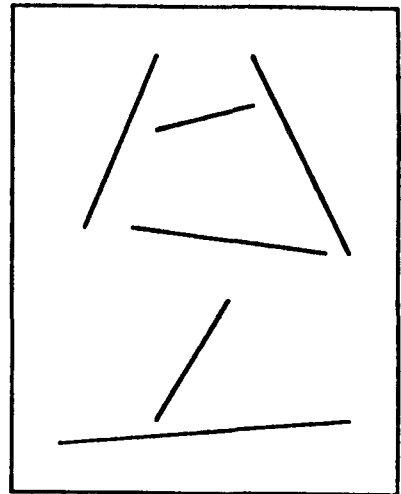
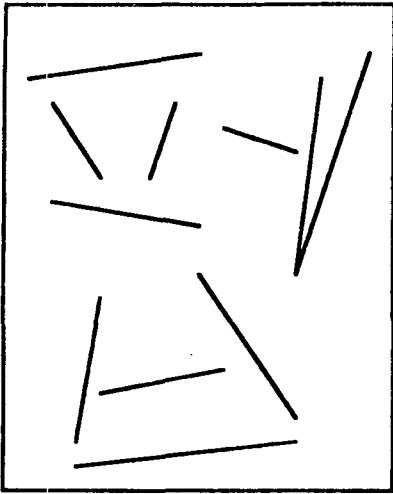


그림 4.1 퍼지소속함수를 사용하여 인식이 가능해진 글자들의 예