

# 한글의 미적 평가를 위한 특징 추출 및 유사도 함수 정의

## Feature Extraction and Similarity Measure Function Define For Beauty Evaluation of Korean Character

한군희

천안대학교 정보통신학부 조교수

오명관

해전대학 컴퓨터학과 조교수

이형우

천안대학교 정보통신학부 교수

전병민

충북대학교 컴퓨터공학과, 컴퓨터정보통신연구소 교수

Kun-Hee Han

Professor, Div. of Inf. & Com. Eng. , Cheonan University

Myoung-Kwan Oh

Professor, Dept. of Computer Science, Hyejeon College

Hyung-Woo Lee

Professor, Div. of Inf. & Com. Eng. , Cheonan University

Byung-Min Jun

Professor, Dept. of Com. Eng. , RICIC.

중심어 : 특징 추출, 유사도 함수, 미적평가

### 요약

본 논문에서는 입력의 자동화 및 교육을 위한 문자 인  
히기 시스템을 위하여 자소의 특징 추출과 유사도 함수를  
정의하여 한글에 대한 미적평가를 수행하였다. 이를 위해  
한글 문자의 자소에 대한 특징 추출 및 유사도 함수를 정  
의 한 후 표준 문자와 입력 문자가 얼마나 유사한지를 평  
가하는 방법을 제안하였다. 표준 문자와 입력 문자의 획  
에 대한 특징 추출 및 유사도 함수를 정의하였으며, 다양  
한 입력 문자 패턴에 대해 표준 문자 패턴과 얼마나 유사  
한지를 실험한 결과 예상한 값과 유사하게 일치하는 실험  
결과를 얻을 수 있었다. 또한 일반 사람들의 미적 평가  
결과와 제안한 방법의 실험 결과가 유사하게 일치한다는  
결과도 얻을 수 있었다.

### Abstract

This study pre-processed the characters, performed the  
feature extraction for the beauty evaluation, and then  
defined the similarity function. It suggested the definition of  
the similarity function, and the extraction of the features of  
character elements. it experimented how much the various  
input character patterns were similar with the standard  
character patterns, found their results were almost similar  
with the expected ones and the results of beauty evaluation  
on general people through the questionnaire with the results  
of the methods suggested here.

### 1. 서론

최근 컴퓨터의 급속한 보급으로 인해 자동화된 입력 데이  
터를 인식하기 위한 방법에 대해 지난 수년간 여러 방면에서  
연구가 이루어져왔다[1][2].

이러한 연구 분야로는 입력의 자동화를 위한 문자 인식  
과 생체 인식 시스템인 얼굴 인식, 홍채 인식, 지문 인식등이  
있다. 또한 의료 영상 처리나 ITS(Intelligent Transportation  
System)등도 입력 데이터를 처리하는 대표적인 응용 분야라  
할 수 있다[3-6].

이 같은 문자 인식 기술은 독일의 G.Tauscheck가 1928년  
패턴정합법을 이용한 문자 인식의 원리적 특허를 등록함으로  
써 시작되었다. 1950년대 중반에는 숫자 등 비교적 분류 대  
상이 적고 단순한 패턴의 문자 인식 연구가 수행되었으며,  
1960년대부터 영문자, 일본어, 한자 인식에 대한 연구가 시작  
되었고, 1980년대 영어 인식기, 일본어 인식기 등 자국 문자  
를 인식 대상으로 하는 문자 인식기가 상품화 되었다.

국내에서도 이를 위해 1970년대 초반부터 대학을 중심으로  
많은 연구가 수행되어 왔다. 그러나 문자 인식 시스템은 단순  
히 해당 국가의 문자 인식에만 초점이 맞추어져 있었지, 이를

확장한 서체 인식이나 문자 익히기 시스템의 연구가 미진하였다. 한글의 경우만을 예를 들더라도 외국인이나 유치원생 혹은 초등학교 저학년생들이 한글을 익히고자 할 때 문자 인식 시스템과 결합하여 사용할 수 있는 확장된 시스템이 없는 이유로 전적으로 교사에게만 의존하고 있다.

본 논문에서는 이를 위해 입력의 자동화와 문자 익히기를 위한 특징 추출 및 유사도 함수를 정의 한 후 표준 문자와 입력 문자가 얼마나 유사한지를 평가하는 방법을 제안한다. 이는 표준 문자 및 입력 문자에 대한 특징 추출 및 유사도 함수를 정의 한 후 미적 평가를 수행하기 위한 전처리 과정 알고리즘으로 사용하고자 제안한 방법이다. 또한 표준 문자와 입력 문자의 획에 대한 특징 추출 및 유사도 함수를 정의하여 미적 평가를 수행할 수 있는 유사도 함수를 제안한다.

우선 2장에서는 기존 방법에 의한 특징 추출과 유사도 함수에 대하여 다룬다. 즉, 경계선 추출 및 세선화 방법에 대하여 다루고, 3장에서는 본 논문에서 제안한 방법에 대해 기술한다. 그리고 4장에서는 표준 문자와 입력 문자를 전처리 한 후 이에 대한 특징값을 계산하고, 계산된 특징값을 유사도 함수에 적용시켜 표준문자 패턴과 다양한 입력 문자 패턴과의 비교를 통해서 미적 평가를 수행하고, 또한 설문을 통해서 정의한 유사도 함수의 타당성을 증명한다. 마지막으로 제5장에서는 결론과 함께 향후 연구 과제를 기술한다.

## II. 세선화 및 유사도 함수

전처리 과정에서 얻은 경계선에 2차원 히스토그램의 분포 특성을 이용하기 위해 세선화(thinning)를 수행한다. 세선화 된 결과에 히스토그램의 특징을 이용하여 획을 분리하며, 분리된 획에 대한 인식을 수행하게 된다. 그런 다음 획의 특징과 획과 획들 사이의 특징을 추출한다.

### 1. 세선화

세선화란 영상이나 문자에서 골격선(skeleton)을 구하는 작업으로 패턴을 가느다란 호와 곡선의 집합으로 표현하는데 사용되어 왔다. 패턴을 세선화하여 표현하는 이유는 형태 분석을 쉽게 할 수 있다는 사실뿐 아니라 데이터 량을 최소화 시켜야 할 필요성 때문이다. 문자의 경우, 세선화된 표현이 그 패턴에 대한 인간의 개념에 더 가깝기 때문에 보다 간단한 분석과 직관적인 인식 알고리즘을 설계 할 수 있게 해준다. 즉 추상적인 그래프 형태의 골격선은 문자의 관념적인 묘

사와 물리적 표현 사이를 연결시키는 것으로 볼 수 있다.

문자나 영상의 골격선을 구하는 작업으로 통상 Nagel의 알고리즘[7]과 Wakayama의 알고리즘[8]을 많이 사용된다. Nagel의 알고리즘은 문자나 물체의 윤곽에 내접하는 원을 이동시키면서 골격선을 추출하는 것이고, Wakayama의 알고리즘은 사각형을 이동시키면서 골격선을 찾는 방법이다. 본 논문에서는 Wakayama의 알고리즘을 이용하여 세선화를 수행한다.

### 2. 유사도 함수 정의를 위한 퍼지 이론

미적 평가를 위해서는 특징 추출을 수행하고 이에 대한 유사도 함수를 정의해야 한다. 본 논문에서는 유사도 함수를 정의하기 위해 퍼지 이론을 적용하였다. 퍼지 이론은 문자나 영상에서 발생하는 애매성을 제거하기에 가장 적합하며, 특히 유사도 함수를 정의하기에 적절한 이론으로 여겨진다.

또한 퍼지 집합의 소속도를 근사화하여 영상처리, 제어공학 등에 많이 적용하는데 이때 근사화된 소속도 함수를 퍼지 숫자라고 하며 통상 삼각 퍼지 숫자와 사다리꼴 퍼지 숫자들이 사용된다[9][10]. 우선 삼각퍼지 숫자에 대해 알아보도록 하자. 삼각퍼지 숫자는 세 개의 점으로 표현이 가능하기 때문에 사용이 간편하다. 예를 들어서 삼각퍼지 숫자  $A$ 를  $A = (a_1, a_2, a_3)$ 라고 정의하면 그림 1과 같은 퍼지 숫자가 되고 이때 각 원소(숫자)등의 소속 함수값은 식(1)과 같이 정의된다.

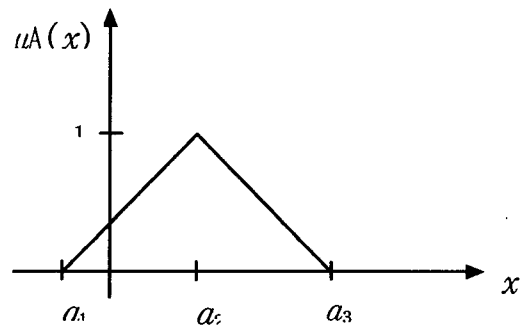


그림 1. 삼각 퍼지 숫자

$$\begin{aligned} \mu A(x) &= 0, & x &\leq a_1 \\ &= \frac{x - a_1}{a_2 - a_1}, & a_1 &\leq x \leq a_2 \\ &= \frac{a_3 - x}{a_3 - a_2}, & a_2 &\leq x \leq a_3 \\ &= 0, & x &> a_3 \end{aligned} \quad (1)$$

퍼지 숫자의 또 다른 하나인 특별한 형태의 사다리꼴 퍼지 숫자에 대해 알아보도록 하자. 이 형태의 숫자는 소속도가 최대가 되는 점이 여러 개가 되어 사다리꼴 모양이 된다. 즉, 다음과 같이 사다리꼴 퍼지 숫자  $A$ 를 정의하면  $A = (a_1, a_2, a_3, a_4)$ 가 되고 이를 그림으로 나타내면 그림 2와 같은 모양이 되고, 소속 함수값은 식(2)와 같이 정의한다.

$$\begin{aligned} \mu_A(x) &= 0 & , x < a_1 \\ &= \frac{x - a_1}{a_2 - a_1} & , a_1 \leq x \leq a_2 \\ &= 1 & , a_2 \leq x \leq a_3 \\ &= \frac{a_4 - x}{a_4 - a_3} & , a_3 \leq x \leq a_4 \\ &= 0 & , x > a_4 \end{aligned} \quad (2)$$

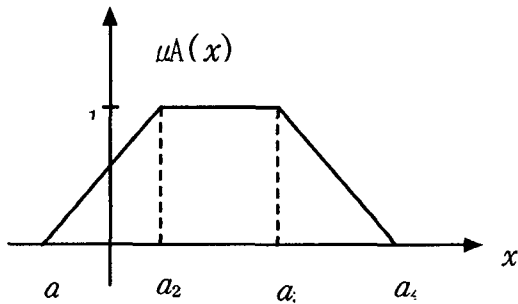


그림 2. 사다리꼴 퍼지 숫자

### III. 미적 평가를 위한 특징 추출과 유사도 함수

미적 평가를 위해서는 획에 대한 특징 추출을 수행하고, 이에 대한 각각의 유사도 함수를 정의해야만 한다. 미적 평가를 위한 유사도 함수는 퍼지수를 사용하였다. 이는 퍼지 이론의 적용이 가장 유리한 분야가 바로 일치도 연산이기 때문이다. 즉, 본 연구에서는 표준 문자 패턴과 입력 문자 패턴과의 일치도를 계산하여 미적 평가를 수행하기 때문에 퍼지수를 이용하여 유사도 함수를 정의하였다. 또한 퍼지 이론은 영상에서 발생하는 애매성을 제거하기에 가장 적합한 이론이며, 특별히 유사도 함수를 정의하기에 적절한 이론으로 여겨진다.

#### 1. 자음의 특징 추출 및 유사도 함수

획에 대한 특징과 획들 사이에 대해 특징 추출을 수행하여

야 하며 이를 바탕으로 유사도 함수를 정의해야 한다. 우선 획의 분리는 인식을 위해 기억된 히스토그램에 대해 최대값과 최소값의 좌표( $X, Y$ )를 구하면 획의 분리 및 획의 시작점과 끝점을 구할 수 있다. 이후 구해진 획과 획들 사이에 대한 특징을 추출하고, 구해진 특징값에 대해 유사도 함수를 정의해야 한다. 먼저 자음에 대해 살펴보기로 하자. 통상 획에 대한 특징 추출 요소는 크게 두 가지로 나누어진다. 첫째가 획에 대한 추출요소이며 또 다른 하나가 획들 사이에 대한 특징 추출 요소이다.

표 1. 자음의 특징 요소

특징		미적 평가 기준값	
		$p$	$q$
획 추출 요소	획의 각도	5	20
획 사이의 추출 요소	각도	5	15
	길이비	0.2	0.4

이 중 획에 대한 특징 추출 요소는 획의 시작점과 끝점사이의 각도로 정한다. 또한 획들 사이에 대한 특징 추출 요소는 획 들간의 이루는 각과 길이비로 한다. 이들에 대한 특징들은 표 1과 같이 나타낼 수 있다. 우선 획에 대한 특징 추출 요소에 대해 유사도 함수를 정의해야 하는데, 이는 식(3)과 같은 함수를 정의하여 수행한다.

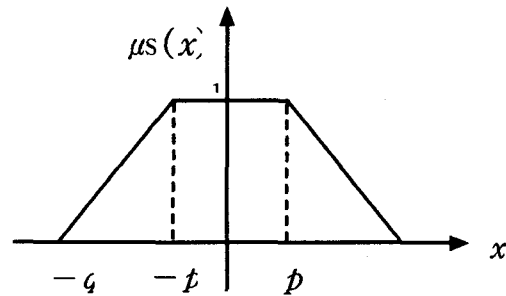


그림 3. 유사도 함수

$$\begin{aligned} \mu_S(x) &= 0 & \text{for } x \leq -q, x \geq q \\ &= \frac{x + q}{q - p} & \text{for } -q \leq x \leq -p \\ &= 1 & \text{for } -p \leq x \leq p \\ &= \frac{q - x}{q - p} & \text{for } p \leq x \leq q \end{aligned} \quad (3)$$

획에 대한 유사도 함수는 위의 그림 3에서  $p$ 는 원점들 중

심으로 5만큼,  $q$ 는 20만큼 떨어진 값을 취한다. 획들 사이에 미적 평가 요소로는 이루는 각과 길이비로 한다. 이중 이루는 각은 식(3)에서 원점을 중심으로  $p=5, q=15$ 의 값을 갖는다. 또한 길이비는 식(3)에서 원점을 중심으로  $p=0.2, q=0.4$ 의 값을 취하며 이때 원점은 표준 문자의 미적 평가값을 기준으로 한다. 이때 이루는 각은 식(4)와 같이 구한다.

$$\text{Intersection Angle} = \tan^{-1}\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \quad (4)$$

' $\alpha$ '와 ' $\theta$ '은 이 같은 방법으로 특징 추출과 이에 대한 유사도 함수를 정의할 수 없으므로 이를 따로 계산해야 한다. 우선 ' $\alpha$ '에 대한 특징 추출은 원의 형성 정도와 둥근 정도의 산술 평균을 통해 구한다.

최종적으로 ' $\theta$ '에 대한 미적 평가를 수행하여야 한다. 이는 획의 특징값과 획 사이의 특징값으로 나누어진다. 획의 특징값은 획의 각도와 ' $\alpha$ '에 대한 미적 평가 요소값과의 산술 평균을 통해 구한다. 획 사이의 특징값으로는 획을 이루는 요소들에 대해 이루는 각과 길이비 그리고 직선 성분의 획과 ' $\alpha$ ' 획 사이의 최소 거리값을 구하여 계산한다

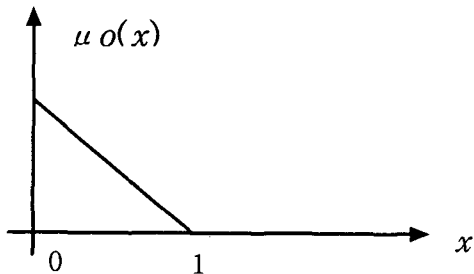


그림 4. ' $\alpha$ '에 대한 유사도 함수

2. 모음의 특징 추출 및 유사도 함수

모음에 대한 특징 추출은 크게 획과 획들 사이의 요소로 나누어진다. 우선 획은 획의 시작점과 끝점 사이의 각으로 정하며 이는 자음에 대해 적용했던 식을 반복 적용하면 된다. 획들 사이의 특징 요소는 크게 세가지로 구분할 수 있다.

첫째가 획들 사이의 이루는 각이고, 둘째가 획들 사이의 길이비이다. 마지막으로 셋째가 획과 획 사이의 교차점 위치이다. 이중 첫째 요소와 둘째 요소는 자음의 경우에 적용한 식을 사용하면 된다. 그러나 셋째 요소는 교점 위치에 대한 유사도 함수식을 새로이 정의해 주어야 한다. 획들 사이의 교점 위치도 크게 1교차점(예를 들어, '교', 'f' 등), 2교차점(예를 들

어, '교', 'f' 등) 그리고 4교차점(예를 들어, 'ff')등으로 나눌 수 있다. 예로서 1교차점에 대한 유사도 함수식을 정의하여 보자. 모음 교차점이란 세로의 긴 획과 가로의 획과의 만나는 위치를 뜻하는데 이는 모음의 미적 평가에 대단히 중요한 요소가 된다. 최종적으로 유사도 함수식은 그림 5와 같다. 그림 5에서  $p$ 와  $q$ 는 교차점에 있어 긴 획의 시작점과 끝점을 나타낸다. 먼저 1교차점은 중간점 즉,  $c = (p + q) / 2$ 이다. 2교차점에 대해서는 1교차점을 확대 적용하면 가능하다.

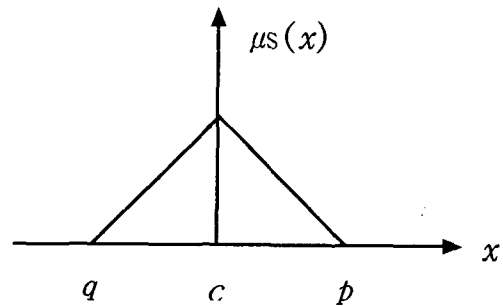


그림 5. 1교차점에 대한 유사도 함수

즉, 2교차점 중 위 교차점은 그림 5에서  $p = p, c = p / 3, q = 2p / 3$ 을 뜻하며 아래 교차점의 경우는  $p = 2p / 3, c = p / 3, q = q$ 가 된다. 2교차점 중 'ff'는 1교차점의 경우를 앞, 뒤 획에 대해 반복 적용하면 된다. 최종적으로 4교차점은 2교차점을 앞, 뒤 획에 대해 반복 적용하면 된다. 위의 내용을 정리하면 표 2와 같이 모음에 대한 특징 요소를 정리할 수 있다.

표 2 모음의 특징 요소

특징		미적 평가 기준값	
획 추출 요소	획의 각도	자음과 동일	
획 사이의 추출 요소	각도	자음과 동일	
	길이비	자음과 동일	
교차점	1교차점	$c = (p + q) / 2$	
	2교차점	위	$p = p$
			$c = p / 3$
	교차점	아래	$q = 2p / 3$
			$p = 2p / 3$
	4교차점		$c = p / 3$
		$q = q$	
	4교차점	2교차점의 앞·뒤 반복적용	

### 3. 히스토그램 분포 군집화

지금까지 한글 표준 문자 패턴에 대한 전처리 과정을 수행한 후 특징 추출 및 유사도 함수의 정의에 대해 제안하였다. 이후 모니터 상에 출력된 표준 문자 패턴에 대해 미적 평가를 원하는 사람이 쓴 문자를 정합시킴으로써 미적 평가를 수행할 수 있게 된다. 입력 문자에 대한 처리는 전처리 과정의 수행, 특징 추출 등의 과정을 통해 정의된 유사도 함수 값과의 정합을 수행함으로써 미적 평가를 수행할 수 있도록 한다. 이때 입력 문자의 경우 히스토그램 투사 프로파일시 누적 분포가 일정 범위 내에 있지 않게 되는 문제가 발생한다. 이를 위해 공간 좌표상의 일정 범위내에 투사 프로파일이 존재하면 이를 군집화를 수행하고, 군집화 된 요소들에 대해 중심점을 선택해 줌으로써 입력 문자에 대한 획의 분리와 인식을 수행할 수 있게 된다.

## IV. 실험 및 고찰

즉, 표준 문자를 선정하여 이에 대한 특징을 추출한 후 문자의 유사도를 나타낼 수 있는 함수를 정의한다. 입력 문자의 특징값과 표준 문자 패턴의 특징값을 정합할 수 있는 유사도를 측정하여 다양한 입력 문자에 대한 미적 평가를 실시할 수 있는 방법을 제안하였다. 정의한 유사도 함수가 객관적으로 유사한지를 평가하기 위해 대학생 100명을 대상으로 입력 문자에 대한 설문문을 통하여 미적인 평가를 실시한 후 본 논문의 실험에 의한 미적 평가값과 비교하여 결과를 나타내었다. 설문 방법은 신명조체 표준 문자를 표준 데이터로 하여, 다양한 필체로 쓴 입력 문자와 표준 문자를 비교하여 얼마나 유사하게 따라 쓰여졌는지를 점수화 한 후 평균값을 데이터로 하였다.

### 1. 표준 문자의 특징 추출과 유사도 함수

#### 1.1. 전처리 및 세선화

표준 문자의 패턴에 대한 전처리한 결과와 세선화한 결과를 각각 그림 6에 나타내었다. 전처리에 대한 결과가 잘 되었는지에 대한 평가는 모호한 문제이다. 복잡한 영상을 대상으로 한다면 주관적, 객관적 비교가 쉽게 되겠지만 문자 영상은 단조롭고 규칙성을 가지고 있기 때문에 전처리가 단조로운 편이다.

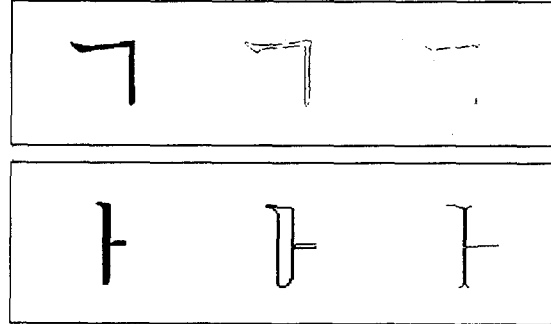


그림 6. 표준 문자, 전처리 및 세선화 결과

### 1.2. 특징 추출

표준 문자의 유사도 함수를 정의하기 위해서는 분리된 획에 대한 특징을 추출해야 한다. 표 3은 이에 대한 특징 추출 결과를 나타내었다. 'ㄱ' 문자의 경우 가로획과 세로획에 대한 각과 두획을 이루는 각 그리고 획에 대한 길이비를 특징으로 할 수 있다. 가로획 'ㅡ'의 경우 0도, 세로획 'ㅣ'은 90도를 나타낸다. 두 획의 각은 90도, 그리고 길이비는 1:1로 나타낼 수 있다. 'ㅏ' 경우는 같은 방법으로 정의하되 교차점의 위치까지 이용하여 미적 평가를 수행하기 때문에 교차점의 위치를 특징으로 한다.

표 3. 표준 문자에 대한 특징 추출 결과

대상문자	획의 특징추출	획들 사이의 특징 추출	
		각	길이비
ㄱ	0, 90	각	90
		길이비	1.0
ㅏ	90, 0	각	90
		길이비	0.1875
		교차점	0.5

### 1.3. 유사도 함수 정의

표준 문자에 대한 유사도 함수는 표 3에서 설정한 특징을 정의된 식에 적용했을 시 유사하게 나타남을 알 수 있었다. 또한 표 4와 5에 표준 문자 'ㄱ'과 'ㅏ'에 대한 유사도 함수를 나타내었다. 실험 결과 표준 문자 패턴에 대한 처리 결과가 예상 결과와 일치함을 확인할 수 있었다

표 4. 'ㄱ'에 대한 유사도 함수

획의 각	0 for $x \leq -20$ or $x \geq 20$ , $\frac{-20-x}{-10}$ for $-20 \leq x \leq -5$ , 1 for $-5 \leq x \leq 5$ , $\frac{x-20}{-15}$ for $5 \leq x \leq 20$
획의 각	0 for $x \leq 70$ or $x \geq 110$ , $\frac{70-x}{-15}$ for $70 \leq x \leq 85$ , 1 for $85 \leq x \leq 95$ , $\frac{x-110}{-15}$ for $95 \leq x \leq 110$
획사이 각	0 for $x \leq 75$ or $x \geq 105$ , $\frac{75-x}{-10}$ for $75 \leq x \leq 85$ , 1 for $85 \leq x \leq 95$ , $\frac{x-105}{-10}$ for $95 \leq x \leq 105$
길이비	0 for $x \leq 0.6$ or $x \geq 1.4$ , $\frac{0.6-x}{-0.2}$ for $0.6 \leq x \leq 0.8$ 1 for $0.8 \leq x \leq 1.2$ , $\frac{x-1.4}{-0.2}$ for $1.2 \leq x \leq 1.4$

표 5. 'ㄷ'에 대한 유사도 함수

획의 각	0 for $x \leq 70$ or $x \geq 110$ , $\frac{70-x}{-15}$ for $70 \leq x \leq 85$ , 1 for $85 \leq x \leq 95$ , $\frac{x-110}{-15}$ for $95 \leq x \leq 110$
획의 각	0 for $x \leq -20$ or $x \geq 20$ , $\frac{-20-x}{-10}$ for $-20 \leq x \leq -5$ , 1 for $-5 \leq x \leq 5$ , $\frac{x-20}{-15}$ for $5 \leq x \leq 20$
획사이 각	0 for $x \leq 75$ or $x \geq 105$ , $\frac{75-x}{-10}$ for $75 \leq x \leq 85$ , 1 for $85 \leq x \leq 95$ , $\frac{x-105}{-10}$ for $95 \leq x \leq 105$
길이비	1 for $x \leq 0.3875$ , $\frac{0.5875-x}{-0.2}$ for $0.3875 \leq x \leq 0.5785$ 1 for $x \geq 0.5785$
교차점	$\frac{x}{0.5}$ for $0 \leq x \leq 0.5$ , $\frac{1.0-x}{0.5}$ for $0.5 \leq x \leq 1.0$

2. 입력 문자에 대한 미적 평가

입력 문자 'ㄱ' 패턴에 대한 전처리 및 특징 추출을 한 후, 입력 문자 각각의 특징에 따라 유사도 함수를 적용시켜 미적 평가 값을 구하게 된다. 각각의 미적 평가값을 합산 한 후 그 값을 특징의 수로 나누어 문자 전체에 대한 미적 평가값을 계산하였다. 또한 객관적인 평가를 위해 설문에 의한 미적 평가값과 유사도 함수에 의한 미적 평가값을 비교하여 본 논문에서 정의한 유사도 함수가 얼마나 잘 정의되었는지를 판단할 수 있도록 하였다.

2.1. 문자 'ㄱ'에 대한 미적 평가

그림 7이 'ㄱ'에 대한 다양한 형태의 입력 문자이고, 그림 8이 이에 대한 전처리 결과를 그리고 그림 9가 세선화한 결과를 나타낸다. 또한 표 6은 세선화 입력 문자에 대하여 각각의 특징 추출 결과를 나타내었다. 입력 문자 'ㄱ'에 대한 특징 추출값의 요소는 가로획의 각과 세로획의 각, 가로획과 세로획 사이의 각 그리고 길이비이다. (a)에 대한 문자의 경우 가로획과 세로획은 직관상으로 봐도 거의神明조체 표준 문자와 거의 유사함을 알 수 있다. 실험 결과로 보아도 가로획의 각도가 -3.5(표준 문자 0도), 세로획의 각도는 92.5(표준 문자 90도)로 나타났다. 또한 가로획과 세로획 사이의 각도 또한 91.3(표준 문자 90도)로 나타남을 알 수 있다. 그리고 가로획과 세로획의 길이비도 0.8(표준 문자 1.0)로 유사 범위 안에 있음을 알 수 있다. 그러나 (b)에 대한 문자는 경우는 다른 결과가 나타났다.神明조체 표준 문자와 비교하지 않고 일반적으로 볼 경우 전반적으로 잘 쓰여진 문자로 보여진다. 세로획이 44.6도이고, 가로획과 세로획의 각은 43.5도로 특징 요소가 유사 범위를 벗어남을 알 수 있다.

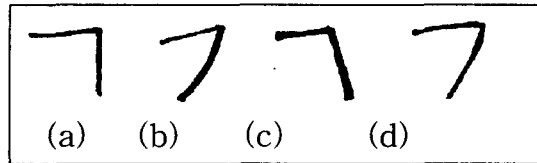


그림 7. 입력 문자(ㄱ)

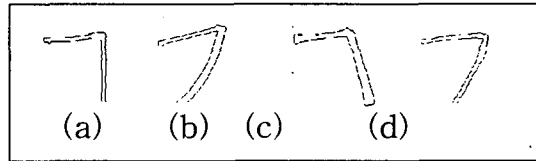


그림 8. 전처리과정 수행 결과(ㄱ)

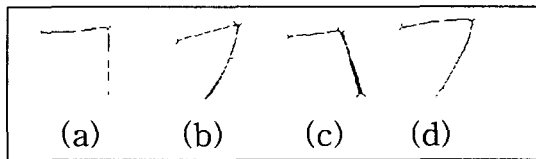


그림 9. 세선화 결과(ㄱ)

표 6. 입력 문자에 대한 특징 추출 결과(7)

대상문자	획의 특징추출	획들 사이의 특징 추출	
		각	길이비
ㄱ	-3.5, 92.5	각	91.3
		길이비	0.8
ㄴ	12.1, 44.6	각	43.5
		길이비	1.3
ㄷ	2.8, 124.2	각	124.2
		길이비	1.125
ㄹ	6.23, 40.82	각	55.41
		길이비	1.1

표 7의 미적 평가 결과에서 'ㄱ' 문자에 대한 평가 결과 값은 1.0, 0.25, 0.5, 0.4795로 각각 나타났다. 또한 그림 10의 비교 결과에서 알 수 있듯이 (b)의 문자를 제외하고는 전반적으로 유사도 함수에 의한 결과와 설문에 의한 결과가 거의 비슷한 값을 나타냄을 알 수 있다. 이상과 같은 결과를 판단해보면 획에 대한 각이 일정범위가 유사범위를 벗어나면 0으로 결과값을 판정했기 때문이다. 추가 연구시 유사 범위를 좀 더 세분화한다면 이 같은 결과는 많이 보완되리라 판단된다.

표 7. 미적 평가 결과(7)

대상 문자	유사도에 의한 미적평가	설문에 의한 미적 평가	미적 평가 결과		
			획	획들 사이	
ㄱ	1.0	0.92	1.0, 1.0	각	1.0
				길이비	1.0
ㄴ	0.25	0.42	0.553, 0	각	0.0
				길이비	0.5
ㄷ	0.5	0.54	1.0, 0.0	각	0.0
				길이비	1.0
ㄹ	0.4795	0.51	0.918, 0.0	각	0.0
				길이비	1.0

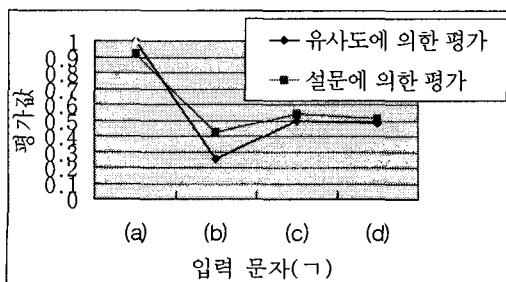


그림 10. 유사도 및 설문에 의한 평가 결과 비교

## 2.2 문자 'ㄹ'에 대한 미적 평가

모음에 대한 특징 추출은 자음과 유사하지만 추가적으로 교점 위치가 어디에 있느냐가 중요한 요소이다. 'ㄹ'의 경우 교점 위치는 세로획의 중앙 위치에 있으면 유사하다고 정의한다. 그림 11이 'ㄹ'에 대한 다양한 형태의 입력 문자이고, 그림 12가 이에 대한 전처리 결과를 그리고 그림 13이 세션화 한 결과이다.

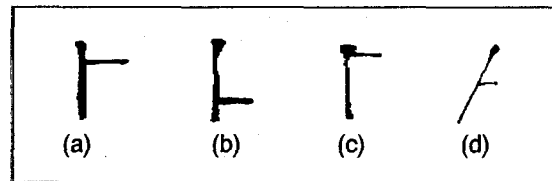


그림 11. 입력 문자(ㄹ)

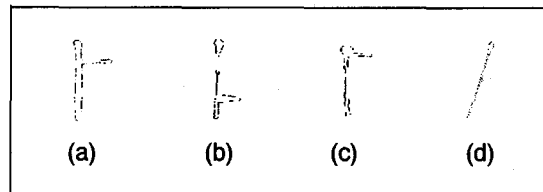


그림 12. 전처리과정 수행 결과(ㄹ)

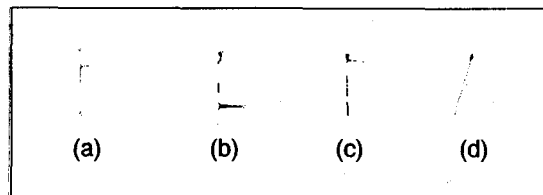


그림 13. 세션화 결과(ㄹ)

또한 표 8은 입력 문자에 대한 각각의 특징 추출 결과를 그리고 표 9는 미적 평가 결과와 설문에 의한 평가 결과를 나타내었다. 표 8의 특징 추출 결과에서 (a)의 경우 세로획의 각이나 길이비에서는 거의 표준문자와 유사하게 나타났으나 교차점 위치가 약간 유사 범위를 벗어났다. 또한 (b)와 (c)의 경우도 세로획은 일반적으로 양호하나 교차점 위치가 아래위로 치우쳤다. (d)의 경우는 세로획의 각도가 맞지 않고, 또한 세로와 가로의 각도 유사 범위를 벗어남을 알 수 있다. 그러나 교점 위치는 상당히 양호하게 나타났다.

표 8. 입력 문자에 대한 특징 추출 결과( )

대상문자	획의 특징 추출	획들 사이의 특징 추출	
		각	
ㄱ	89.8, 2.1	각	88.7
		길이비	0.263
		교차점	0.263
ㄴ	92.3, 1.6	각	90.7
		길이비	0.156
		교차점	0.75
ㄷ	91.7, -5.6	각	97.3
		길이비	0.256
		교차점	0.256
ㄹ	65.4, 6.3	각	59.1
		길이비	0.195
		교차점	0.536

표 9의 미적 평가 결과에서 (a)와 (b)의 유사도에 의한 평가 결과가 각각 0.905와 0.90으로 나타났다. 이 경우로 볼 때 각이나 길이비는 유사하였지만 교차점 위치에서 약간의 차이가 있었음을 알 수 있다. (c)의 경우 (a)와 (b)의 경우와 유사하지만 획의 각이나 교차점 위치에 있어서 약간의 유사 범위를 벗어나 결과가 0.848로 나타났다. (d)는 특징 추출에서도 알 수 있듯이 예상한 결과값인 0.569로 나타났다. 그림 13은 유사도 함수에 의한 미적 평가 결과와 설문에 의한 평가 결과를 나타낸 것인데 (a), (b), (d)는 비교값이 어느 정도 유사하게 나타났으나 (c)의 경우는 차이가 많이 나타남을 알 수 있다. 이같은 결과로 볼 때 유사도함수에서는 약간의 차이만 두었지만 사람이 보는 시각에서는 미적 평가를 위한 교점 위치가 중요하다는 것을 알 수 있다.

표 9. 미적 평가 결과( )

대상 문자	유사도에 의한 평가	설문에 의한 평가	미적 평가 결과		
			획	획들 사이	
ㄱ	0.905	0.91	1.0, 1.0	각	1.0
				길이비	1.0
				교차점	0.526
ㄴ	0.90	0.85	1.0, 1.0	각	1.0
				길이비	1.0
				교차점	0.5
ㄷ	0.848	0.56	1.0, 0.96	각	0.77
				길이비	1.0
				교차점	0.512
ㄹ	0.569	0.41	0.0, 0.913	각	0.0
				길이비	1.0
				교차점	0.932

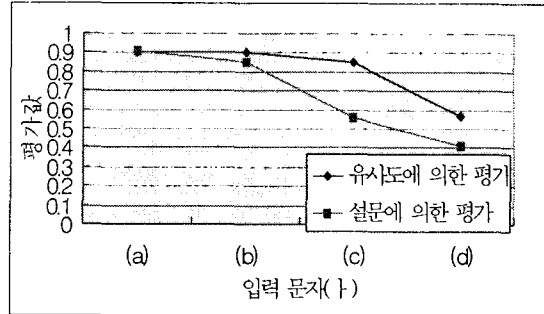


그림 14. 유사도 및 설문에 의한 평가 결과 비교( )

### V. 결 론

이상의 실험 결과에서 알 수 있듯이 본 논문에서 제시한 유사도 함수가 실험적으로 타당함을 입증할 수 있었으며, 또한 설문 데이터와의 비교에서도 유사함을 확인할 수 있었다. 이는 본 논문이 미적 평가용 방법으로 어느 정도 유용하게 사용할 수 있다는 것을 입증할 수 있는 자료가 됨을 알 수 있다. 그러나 현재의 실험 결과는 본 연구에서 제시한 미적 평가 요소 값만을 기준으로 한 결과이다. 이것을 언어학을 전공한 전문가에게 의뢰하여 어떤 것이 미적 평가값으로 더 추가되어야 할지 그리고 현재 정의한 유사도 함수의 수식 범위가 적절한지에 대한 추가 연구가 수행되어야 할 것으로 여겨진다. 아울러 획뿐만 아니라 문자 전체에 대한 전처리과정, 특징 추출과 유사도 함수 정의 및 정할 등에 대한 연구도 수행되어 문자 전체의 균형성에 대한 평가 요소의 연구가 지속적으로 이루어져야 할 것으로 여겨진다. 또한 현재까지 연구한 본 결과를 일반화가 가능하도록 프로그램 최적화와 입출력 기기와의 연결 작업등으로 실제 사용 가능토록 하는 것에 대한 실제적인 연구가 지속적으로 이루어져야 할 것으로 여겨진다. 아울러 한글 문자뿐만 아니라 동양권 문자(일본어, 중국어)와 서구권 문자 등에 대해서도 연구 범위를 확장하여 다기능의 문자 익히기 시스템을 구축하는 것에 대한 연구가 지속적으로 수행되어야 할 것으로 사료된다.

### 참 고 문 헌

[1] D. J. Burr, "Designing a Handwriting Reader", IEEE Trans. Pattern Anal. Machine Intel., Vol. 5, No. 5,



pp.554-559, Sep. 1983.

[2] L. A. Fletcher and R. Kasturi, "A Robust Algorithm for Text String Separation from Mixed Text/Graphics Images", *IEEE Trans. Pattern Anal. Machine Intel.*, Vol. 10, No. 6, pp.910-918, Nov. 1988.

[3] L. A. Fletcher & R. Kasturi, "A Robust Algorithm for Text String Separation from Mixed Text/Graphics Images", *IEEE Trans. on PAMI*, Vol. 10, No. 6, pp.910-918, 1988.

[4] S.Liang, M. Shridhar & M. Ahmadi, "Segmentation of Touching Characters in Printed Document Recognition", *Pattern Recognition*, Vol. 27, No. 6, pp.825-840, 1994.

[5] R. Chellappa et al, "Human and Machine Recognition of Faces : A Survey", *Proceedings of the IEEE*, Vol. 83, No. 5, 1995.

[6] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces", *Int'l Conference on Pattern Recognition*, 1991.

[7] Nagel & Blum, "A symmetric Axis Basis for Object Recognition and Description", *Proceedings of IEEE Meeting on Decision and Control*, 1976.

[8] Wakayama, "Skeleton Tracing Based on Maximal Square Moving", *IECE Technical Report PRL 78-87*, 1978.

[9] 오길록, 이광형, 퍼지 이론 및 응용, 1, 2권 흥릉과학출판사, 1991.

[10] Klir & Folgez, *Fuzzy Sets, Uncertainty and Information*, Prentice Hall, 1988.

한 군 희(Kun-Hee Han)

중신회원



1989년 2월 : 충북대학교 컴퓨터공학과 (공학사)  
 1994년 8월 : 경남대학교 컴퓨터공학과 (공학석사)  
 2000년 8월 : 충북대학교 컴퓨터공학과 (공학박사)

1989년 1월~1994년 12월 : 대우정보시스템 시스템개발

2001년 3월 ~ 현재 : 천안대학교 정보통신학부교수

<관심분야> : 영상처리, 패턴인식, 멀티미디어 및 Mobile 콘텐츠

오 명 관(Myoung-Kwan Oh)

정회원

1990년 2월 : 충북대학교 컴퓨터공학과 졸업 (공학사)

1993년 2월 : 충북대학교 대학원 컴퓨터공학과 졸업 (공학석사)

1997년 3월 ~ 현재 : 충북대학교 대학원 컴퓨터공학과 수료 (박사과정 수료)

1993년 9월 ~ 1999년 2월 : 고려정보테크(주)

1999년 3월 ~ 현재 : 해전대학 컴퓨터계열 교수

이 형 우(Hyung-Woo Lee)

정회원



1994년 2월 : 고려대학교 전산학과 졸업(이학사)

1996년 2월 : 고려대학교 전산학과 졸업(이학석사)

1999년 2월 : 고려대학교 전산학과 졸업(이학박사)

1999년 3월 ~ 현재 : 천안대학교 정보통신학부 교수

<관심분야> : 암호학, 정보보호, 전자상거래, 디지털 콘텐츠

전 병 민(Byung-Min Jun)

중신회원

1976년 2월 : 한국항공대 전자공학과(공학사)

1978년 2월 : 연세대학교 전자공학과(공학석사)

1988년 8월 : 연세대학교 전자공학과(공학박사)

1978년 8월 ~ 1982년 3월 : 공군사관학교 전자과 전임강사

1982년 4월 ~ 1986년 2월 : 동양공전 통신과 조교수

1986년 3월 ~ 현재 : 충북대학교 컴퓨터공학과 교수

1991년 1월 ~ 1992년 1월 : 미국 미시간대학교 교환교수

<관심분야> : 영상처리, 디지털 신호처리