

# 자음의 미적 평가 추출 요소를 이용한 유사도 함수 정의 A Definition of Similarity Measuring Function using Beauty Evaluation Extraction Factor of the Consonant

한군희\*, 백순화\*\*, 연승호\*\*\*, 전병민\*\*\*\*

Kun-Hee Han\*, Soon-Hwa Back\*\*, Seung-Ho Baek\*\*\*, Byoung-Min Jun\*\*\*\*

### <Abstract>

This paper proposes on the Hanguel character CAI system using image processing. For this, firstly, the characters written by elementary school students or foreigners are captured by CCD camera. Secondly, Recognition is accomplished by pre-processing, thinning and recognition processes. Thirdly, strokes are separated and beauty evaluation is done by matching feature value of the input image from the similarity measure function. In particular, this paper describe to define the similarity measuring function using extracted factor values after getting the beauty evaluation factor values of the consonant in the entire CAI system. Finally, the effectiveness of the proposed system is demonstrated by experiments.

**Key words** : Consonant, Beauty Evaluation, CAI

## 1. 서론

최근 컴퓨터의 급속한 보급 및 자동화된 입력 데이터를 위한 패턴인식 방법에 대해, 지난 수년간 여러 방면에서 연구가 이루어졌다[1,2]. 특히 패턴인식 분야 중 문자 인식분야는 실용적으로는 컴퓨터와 인간의 보다 원활한 인터페이스를 추구하려는 목적에서 출발하였으며, 학문적으로는 인간의 우수한 능력중의 하나인 패턴인식의 능력을 컴퓨터에게 부여하여 인간처럼 사고하고 판단할 수 있는 인공지능컴퓨터 실현

에 목적이 있다. 문자 인식의 방법은 가장 단순한 원형정합 방법에서 시작하여 신경망을 이용하는 방법에 이르기까지 다양한 접근 방법들이 동서양을 떠나 폭넓게 연구되고 있다[3].

국내의 경우 1970년대 초반부터 몇몇 대학에서 한글 문자 인식을 중심으로 연구를 시작하여 1980년대는 보다 많은 대학과 업체 연구소 등에서 집중적으로 연구되었으며, 1990년대 초반부터 ReadEx(1992), NeuroOct(1993), 글눈(1993), 아르미(1994), Speed Reader(1994), EyeQ(1995), SuperReader(1997), AnyPage(1999)

\* 정희원, 대전대학 전기전산학부 교수·工博  
355-830, 보령시 주포면 관산리  
E-mail hankh@dcc.ac.kr

\*\* 천안외국어대학 컴퓨터 정보과 교수

\*\*\* 한국통신 멀티미디어연구소

\*\*\*\* 충북대학교 컴퓨터공학과 교수·工博

\* Assistant Professor, Dept. of Electric Computer Eng.  
Jupomyun, Boryungsi, Chungnam, 355-830,



등의 문자 인식기가 상품화되었다. 1999년의 기술을 중심으로 한글, 한자(1만 6천자), 영어, 일어, 독어, 불어 등 14개 국어를 인식할 수 있는 시스템이 탄생되었고 도표 인식, 사진합성 등이 가능한 시스템이 출현되고 있다.

그러나 거의 모든 시스템이 인쇄체 문자 인식에만 머물러 있으며, 인식 시스템을 확장하여 문자 익히기까지 행하는 시스템은 전무한 형편이다. 또한 한글은 일본의 가나 문자나 중국의 한자에 비해 상대적으로 뛰어난 문자이지만 문자를 익히기가 쉽지 않다는 문제가 존재한다.

따라서 국내의 유치원생이나 초등학생 그리고 외국인들이 한글을 익힐 때 이에 대해 많은 시간을 투자하고 있으며 이에 따른 적절한 교수법의 개발이 요구되고 있다. 현재까지는 이러한 부분에 대한 연구가 활발하게 이루어지지 않고 있지만 앞으로는 활발이 진행되리라 판단된다. 이를 위해 본 논문에서는 자동으로 문자를 인식하고 이에 대한 미적 평가 (beauty evaluation)를 통해 한글문자를 익히는 CAI (Computer Aided Instruction) 시스템을 개발하고자 한다. 이는 다음과 같은 과정을 거친다.

첫째, 학생들이 자신의 개인용 컴퓨터(PC)에 앉아서 모니터상에 나오는 문자를 따라서 쓴다. 둘째, 학생들이 따라 쓴 문자를 CCD카메라가 입력받아 문자에 대한 인식을 행한다. 셋째, 인식된 문자에 대해 표준 패턴과의 유사도를 계산해 줌으로써 문자 익히기와 올바른 서체를 익히도록 해 준다.

또한 기존의 문자인식[3,4,5]이 입력의 자동화에만 머물러 있었던 것을 확장하여 문자 익히기와 서체 익히기까지 행하고자 하는 시도이며, 실시간 처리가 가능토록 하기 위해 히스토그램의 특성을 이용하여 자소를 인식하고 획을 분리하여, 자소에 대한 특징 추출 및 이를 입력 영상과 얼마나 정확하게 일치하는 가를 계산하기 위한 유사도 함수를 정의하고자 한다.

이상과 같이 본 논문에서는 전체 시스템 중 표준 문자 자소의 특징 추출 및 유사도 함수의 정의에 대한 방법론을 다루고자 하며 이를 신명조체 40포인트를 중심으로 행하고자 한다. 신명조체를 표준 문자 자소로 선정한 이유는 저학년 학생들이나 외국인들이 처음 문자를 익히

기에 가장 적합한 글씨체로 여겨지기 때문이다.

또한 전체 시스템을 표준문자 자소의 특징 추출 및 유사도 함수의 정의와 입력 문자의 처리 및 정합과 같이 크게 두 가지로 나누었을 시 이중 표준문자 자소의 특징 추출 및 유사도 함수의 정의에 대해 다루고자 하며, 우선적으로 자음을 중심으로 이를 다루고자 한다. 본 방법은 특정 필체에 대해 표준 문자 패턴을 형성한 후 이에 대한 유사도 함수를 정의하면 한글 문자 익히기 CAI 시스템을 필자 인식(writer's recognition) 시스템으로도 확장 가능한 방법으로 여겨진다.

## 2 시스템의 구성 및 개념

개발하고자 하는 시스템에 대한 전체 구성도는 Fig.1과 같다. 우선 표준문자에 대해 자음, 모음, 밑받침의 인식과 미적 평가 요소값의 추출을 행한다.

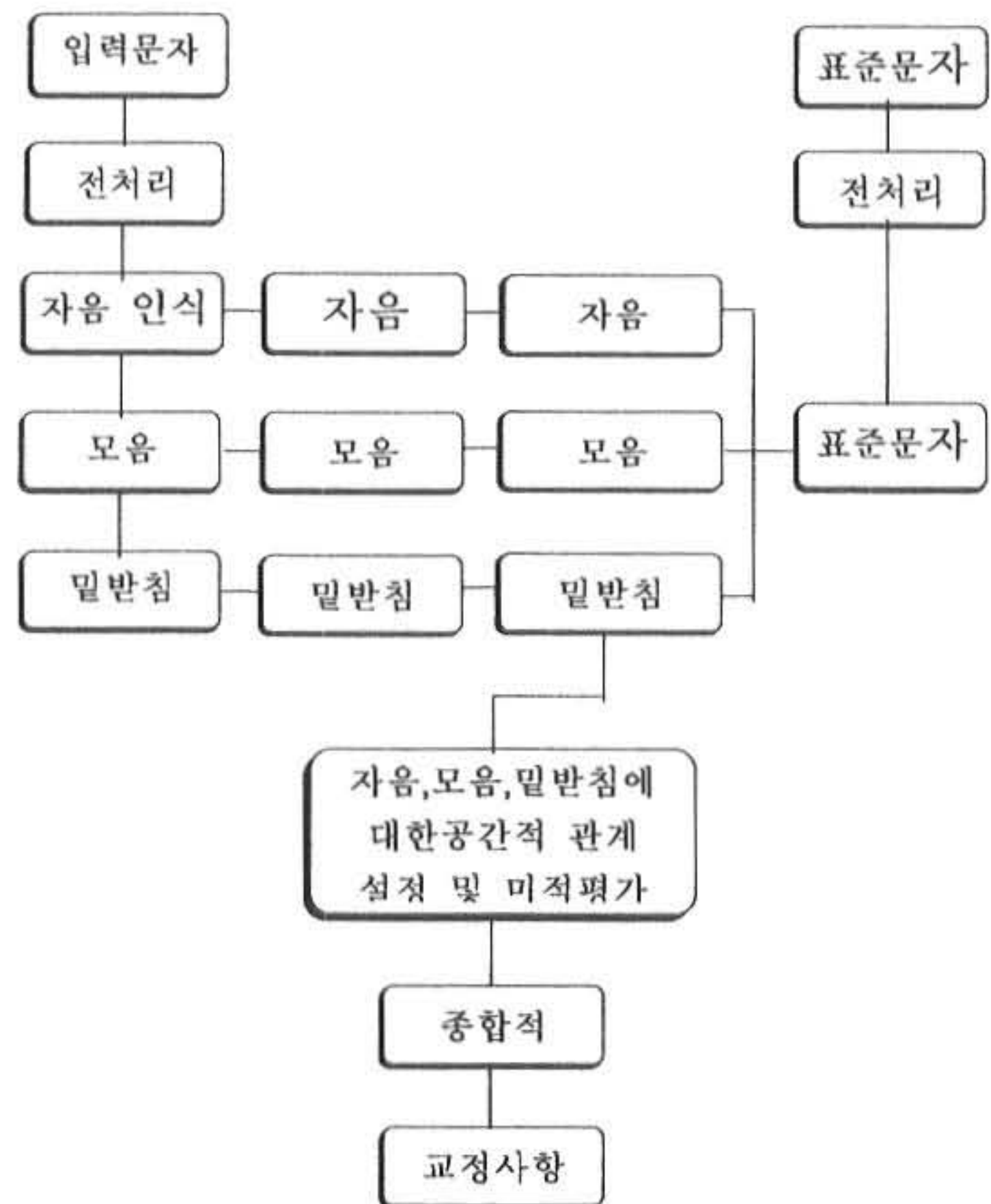


Fig.1 Block diagram of the overall system

차후 입력 문자에 대해 인식과정과 표준문자와의 유사도를 계산함으로써 미적 평가를 행한다. 이는 자음, 모음, 밑받침에 대해 차례로 적용하



며 자음, 모음, 밀반침간의 공간적 관계를 설정하여 전체문자에 대한 미적 평가도 행한다. 즉, 표준문자 패턴과 입력문자 패턴이 얼마나 잘 일치하는가를 유사도 함수 계산에 의해 미적 평가를 행함으로써 문자 익히기를 행하는 방법을 제안하고자 한다. 특히 전체 시스템 중 본 논문은 자음의 표준문자에 대해 먼저 전처리 과정을 행한 후 미적 평가를 행하는 방법론을 제안하고자 한다.

본 논문에서 다루고자 하는 Fig.2의 구성도를 보면, 먼저 표준 문자를 입력 받은 후 전처리 과정을 통해 잡음 제거 및 경계선 추출을 하고 세선화를 행한다. 그런 후 히스토그램 특성과 원형 정합을 이용해서 인식을 수행한다. 인식된 결과를 이용하여 자소로부터 획을 분리하고, 분리된 획과 획들간의 관계를 설정지어 자음에 대한 미적 평가를 위한 특징을 추출한다. 마지막으로 입력문자와 표준문자의 유사도로 자음에 대한 미적평가를 행하는 과정을 나타내었다.

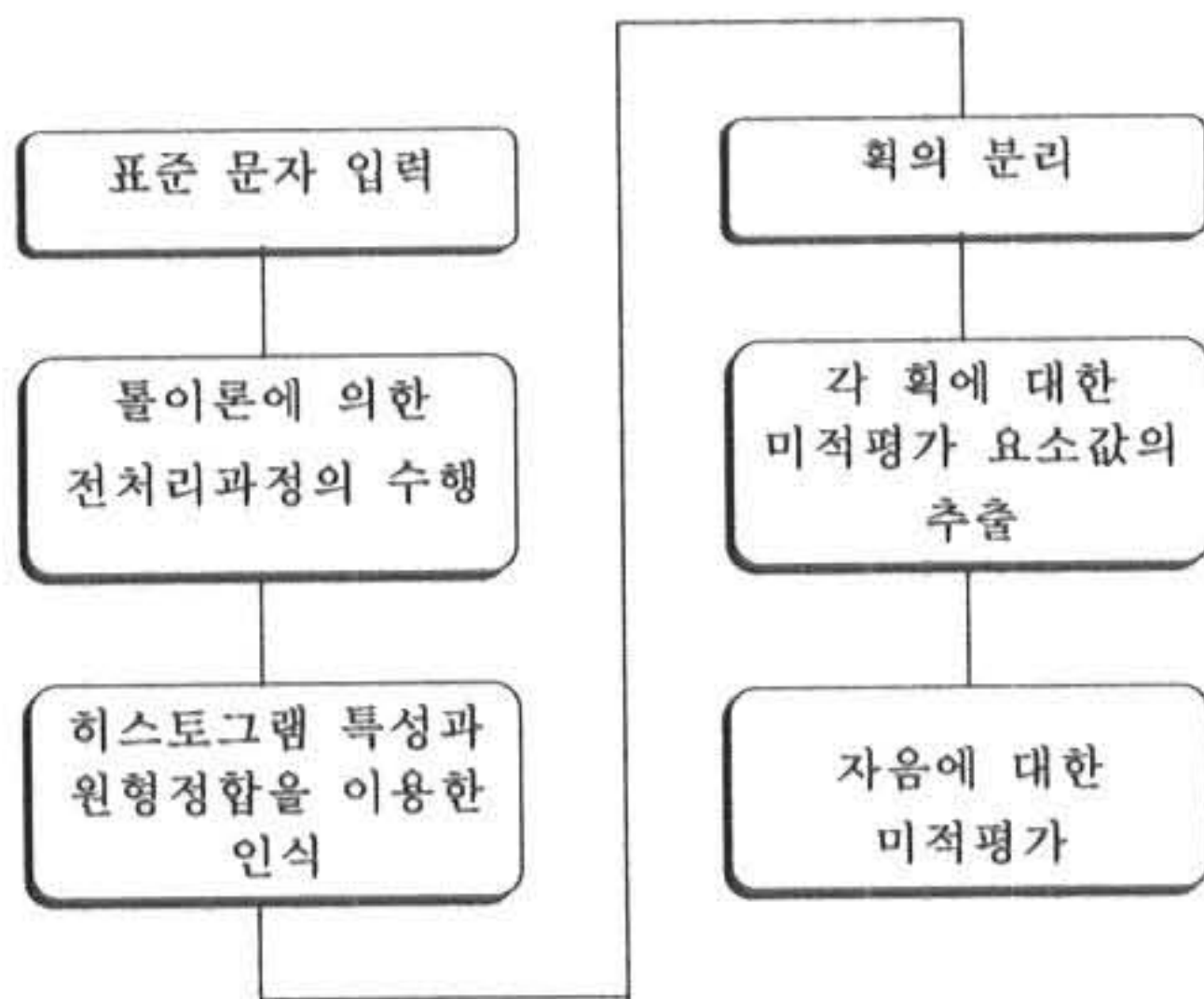


Fig.2 Block diagram of beauty evaluation system for consonant

### 3. 전처리 및 인식

#### 3.1 전처리 과정

입력 영상에서 전처리과정(pre-processing)을 통해 잡음 제거 및 경계선 추출을 행하여야 한다. 이를 위해 틀이론[6,7]을 적용하여 영상의 구조를 파악하여 전처리과정을 행한다. Fig.3에 영상 구조의 예를 나타내었으며, 영상 구조를 파악하기 위해 틀 멤버쉽 함수를 식(1)과 식(2)

와 같이 정의한다. 이를 통해 가장 적은 비용을 갖는 영상구조를 선택함으로써 경계선 추출을 행한다.

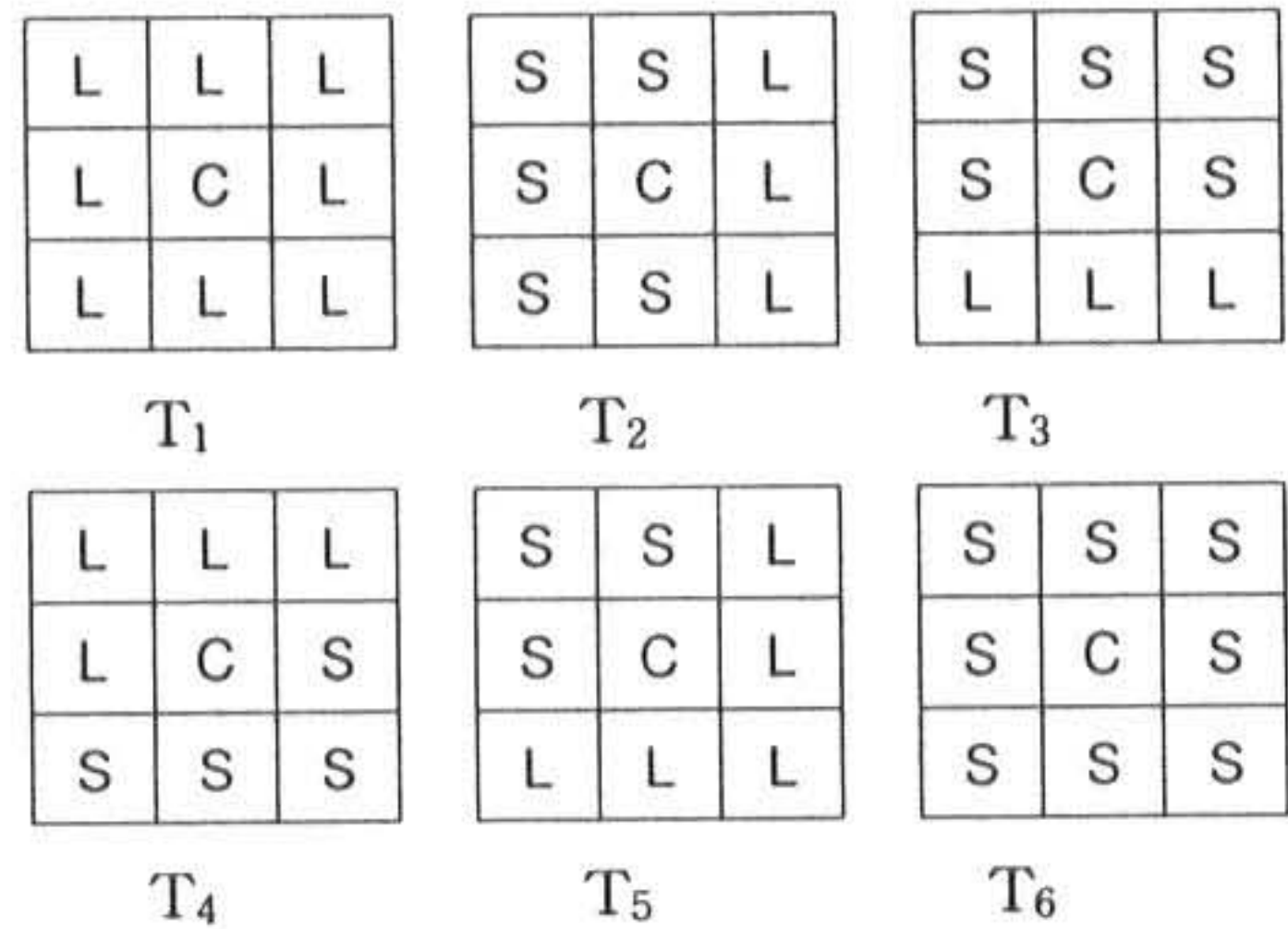


Fig.3 Examples of image structures

Fig.3에서 L은 중심화소 'C'를 중심으로 명암도의 차가 '크다 (Large : L)'는 의미를 S는 '작다(Small : S)'는 것을 뜻한다.

$$\psi_L(x) = -\log [x/255] \quad (1)$$

여기서  $x = |C - L|$  나타내고,

$$\psi_S(x) = -\log [-(x-255)/255] \quad (2)$$

또한  $x = |C - S|$  나타낸다.

#### 3.2 인식

문자 인식 방법은 크게 원형 비교 방법 (Template Matching Method), 통계적 방법 (Statistical Method), 구조적 방법 (Structural Method), 인공신경망 (Artificial Neural Network) 을 이용한 방법 및 조합에 의한 방법 등으로 이루어진다[8].

본 논문에서는 자음의 인식을 2차원 히스토그램 특성을 이용한 원형 정합 방법을 통하여 수행한다. 히스토그램이란 명암도의 분포나 이치 화상에서 검은 화소나 경계선의 분포를 일컫는다[9]. 즉 문자는 항상 가로 또는 세로의 규칙성을 갖고 있어 수평과 수직 히스토그램으로 분리된다[1,2,10].

이러한 특성을 이용해 전처리 과정에서 추출된 경계선에 세선화과정[11]을 수행하여 2차원 히스토그램의 분포를 구한다.

세선화는 Wakayama의 알고리즘으로써 문자 영상의 내부에 인접한 원을 그려 이의 중심점



을 추적함으로써 세선화를 행하였다.

즉, Wakayama의 알고리즘은 문자를 구성하고 있는 두터운 경계선을 따라 경계선을 채우는 내접원을 작성하고 이 내접원의 중심을 이어 나가는 방식으로 세선화를 행하는 방식이다.

이제 히스토그램에 대해 알아보면, 아래에 Y축 히스토그램과 X축 히스토그램의 수식을 나타내었다.

$$H_x(i) = \sum_{j=1}^n Y(j) \quad (3)$$

$$H_y(j) = \sum_{i=1}^n X(i) \quad (4)$$

여기서 X(i)와 Y(j)는 각각 X축과 Y축을 따라 누적되는 이치 화상의 수를 뜻한다. 즉, X축과 Y축의 히스토그램으로부터 일정범위내의 히스토그램 분포를 찾고, 이 분포식의 최대값을 히스토그램 분포수로 정하면 Table1과 같이 자음에 따른 히스토그램의 분포수를 얻을 수 있게 된다.

Table 1. Recognition of consonant

X 히스토그램	Y 히스토그램	대상문자
0	0	ㅅ, ㅇ
0	1	ㅈ
0	2	ㅎ, ㅊ
1	1	ㄱ, ㄴ
1	2	ㄷ, ㅋ
1	3	ㅌ
2	2	ㅍ, ㅂ, ㅅ
2	3	ㄹ

이는 표준문자 패턴으로 신명조체를 기준으로 하였으며 학습자도 신명조체에 대해 얼마나 문자 패턴을 잘 익히는가를 입력문자 패턴으로 작성하기 때문에 아래 표와 같은 방법으로 효과적으로 자음을 인식할 수 있다. 만약 다른 문자 패턴(예를들어 흘림체) 등을 익히고자 한다면 해당 글씨체의 특성에 따른 수정.보완이 뒤따라야 한다.

이때 히스토그램의 분포수가 같은 자음일 경

우 원형 정합에 의해 인식을 수행한다. 현재 논문에서 다루고자 하는 데이터는 표준문자 영상을 구성하기 위한 것으로 신명조체 40포인트를 기준으로 하였기 때문에 이같이 히스토그램 분포 특성에 따라 쉽게 인식을 행할 수 있었다. 그러나 필자 인식이나 특정 필기체에 대한 인식을 행할시는 보다 복잡한 경우를 다루어야 하므로 이에 대해서도 차후 보다 종합적인 시스템 구성시 인식에 대한 알고리즘 보완이 이루어져야 하겠다.

#### 4. 미적 평가 요소의 선정과 값의 추출

##### 4.1 획의 분리

자소로부터 획을 분리하고, 분리된 획과 획들간의 관계를 설정지어 자음에 대한 미적평가요소값을 추출해야 한다. 우선 획은 X,Y 히스토그램으로부터 임계값 TH<sub>1</sub>을 넘고 균집화한 요소중 최대값을 찾아 해당(X,Y)좌표를 구해 획의 분리와 분리된 획으로부터 시작점과 끝점을 구한다. 이 같은 방법으로 획의 분리가 안되는 자음은 'ㅅ', 'ㅇ', 'ㅎ', 'ㅈ', 'ㅊ', 'ㅋ'인데 이것은 다음과 같은 방법으로 획의 분리를 행한다.

첫째, 'ㅅ'의 경우 3×3창 내에서 세선화된 경계선에서 추적을 행하여 분기점을 기억시켜 놓고 헤어(hair)여부를 확인 후 헤어가 아니면 분기점으로부터 '/'과 '\'부분을 분리한 후 획의 양 끝점을 결정한다.

둘째, 'ㅋ'의 경우 히스토그램분포로부터 '/'부분의 획만이 특징점으로 추출이 안되므로 남은 획에 대해 (Xmin, Ymin)과 (Xmax, Ymax)를 찾아 획의 시작점과 끝점을 검출한다.

셋째, 'ㅇ'의 경우 시작점과 끝점 사이의 거리 d가 임계치 TH<sub>2</sub>이하이면 'ㅇ'의 획으로 처리한다. ㅎ, ㅈ, ㅊ의 경우는 분리가 되는 부분(즉, -)는 분리하여 처리하고 분리가 안되는 부분은 위에서 언급한 세가지 경우 중 해당되는 알고리즘을 적용하여 결합시키면 ㅎ, ㅈ, ㅊ의 경우에 대한 획의 분리를 행할 수 있게 된다.

##### 4.2 미적 평가 요소값의 추출

미적 평가요소는 크게 획에 대한 요소와 획들 사이의 요소로 나뉘어진다. 획은 시작점과



끝점 사이의 각으로 획들 사이의 요소는 획들 사이의 이루는 각(intersection angle)과 길이비(length ratio)를 미적 평가 요소 값으로 선정한다. 우선 획에 대한 요소인 획의 시작점과 끝점 사이의 각도는 다음과 같은 함수를 정의하여 미적 평가를 행한다. 미적 평가를 위한 유사도 함수를 Fig.4와 같은 퍼지수를 사용하였다.

이는 퍼지 이론의 적용이 가장 유리한 분야가 바로 일치도(similarity measure) 연산이기 때문이다. 즉, 본 시스템은 표준 문자 패턴과 입력 문자 패턴과의 일치도를 계산하여 미적 평가를 행하기 때문에 퍼지수를 이용하여 유사도 함수를 정의하였다.

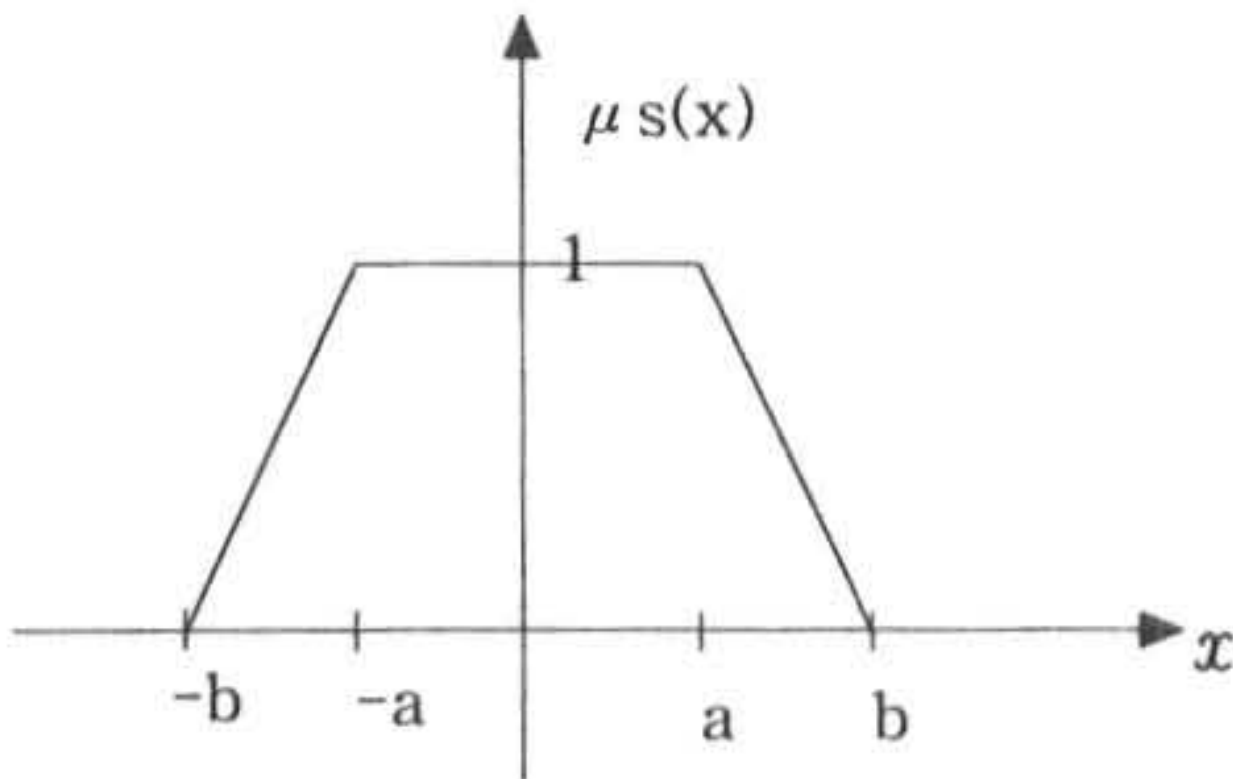


Fig.4 Function of beauty evaluation

$$\begin{aligned} \mu_s(x) &= 0 && \text{for } x \leq -b, \quad x \geq b \\ &= \frac{b-x}{b-a} && \text{for } -b \leq x \leq -a \\ &= 1 && \text{for } -a \leq x \leq a \\ &= \frac{x-b}{a-b} && \text{for } a \leq x \leq b \end{aligned} \quad (5)$$

획에 대한 미적 평가 요소값은 Fig.4에서 a는 원점을 중심으로 5만큼, b는 20만큼 떨어진 값을 취한다. 획들 사이의 미적 평가는 이루는 각과 길이비로 한다. 이중 이루는 각은 식(5)에서 원점을 중심으로 a=5, b=15의 값을 갖는다.

또한 길이비(length ratio)는 식(5)에서 원점을 중심으로 a=0.2, b=0.4의 값을 취하며 이때 원점은 표준 영상의 미적 평가값을 기준으로 한다. 'ㅇ'과 'ㅎ'은 이같은 방법으로 미적 평가 요소값을 추출할 수 없으므로 이에 대한 미적 평가 요소값을 따로 계산해야 한다. 우선 'ㅇ'에 대한 미적 평가 요소값은 원의 형성 정도와 둥근 정도(degree of circleness)의 산술평균을 통

해 구한다. 우선 원의 형성 정도는 하식에 의해 구한다.

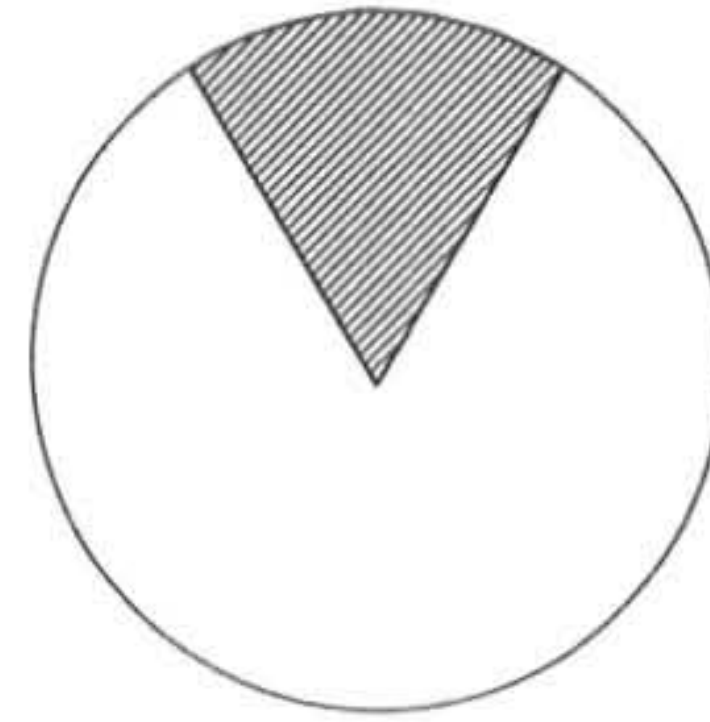


Fig.5 Area degree of circle

$$1 - \frac{\text{빛금친 원의 면적}}{\text{전체 원의 면적}} \quad (6)$$

이 때 원의 면적을 구하기 위해서는 반지름 R을 구해야 한다. 이는 다음과 같은 방법으로 구한다. 원을 이루는 두 점을 A(Xi, Yj), B(Xp, Yq)라고 하면 식(7)과 같은 현의 수직 이등분선을 구하여 이를 누적시켜 누적 분포가 최대가 되는 Index(M, N)이 원의 중심이 되며, 중심으로부터 각 점에 이르는 거리를 계산함으로써 원의 반지름 R을 구하게 된다.

$$Y - (Yj + Yq) / 2 = (Xp - Xi) / (Yq - Yj) / [X - (Xi + Xp) / 2] \quad (7)$$

원의 반지름 R을 구하는 수식은 아래와 같다.

$$\text{SQRT} [(Xi - M)^2 + (Yj - N)^2] = R \quad (8)$$

둥근 정도는 원의 반지름 R에 대한 전체 누적 배열에서 최대 누적 배열을 갖는 반지름 R\*에 대한 비율로 정한다. 즉,

$$\frac{\text{최대 누적 배열을 갖는 R의 개수}(R^*)}{\text{전체 R의 갯수}}$$

전체 R의 갯수

최종적으로 'ㅎ'에 대한 미적 평가를 행하여야 한다. 이는 획의 특징값과 획 사이의 특징값으로 나뉘어 진다. 획의 특징값은 획의 각도와 'ㅇ'에 대한 미적 평가 요소값과의 산술 평균을 통해 구한다. 획 사이의 특징값으로는 획을 이



루는 요소들에 대해 이루는 각과 길이비 그리고 직선 성분의 획과 'o' 획 사이의 최소 거리 값을 구하여 계산한다

### 5. 실험 및 고찰

본 논문에서의 실험은 IBM-PC상에서 C언어를 사용하여 구현하였다. 다음의 그림은 실험한 전체 자음 중에서 'ㄴ', 'ㄹ', 'ㅁ'에 대해서만 나타내었는데, Fig.6은 40포인트 표준 문자 영상이며, Fig.7은 표준문자 영상에 대해 경계선을 추출한 결과이다. 40포인트를 기준으로 한 이유는 출력 모니터 상에 문자가 출력되고 이를 보고 학습자가 학습을 행해야 하는 관계로 40포인트는 되어야 이 같은 작업이 가능하기 때문이다.

그리고 Fig.8은 추출한 경계선에 세선화를 행한 결과를 나타낸다. 아울러 Table 2에 표준 문자영상에 대한 특징 추출결과를 나타내었다. 또한 Table 3, Table 4와 Table 5는 입력 영상의 미적 평가값을 계산하기 위해 유사도 함수를 정의한 결과를 나타내었다.

본 논문에서는 문자의 특징을 추출하여 미적 평가를 행하기 위한 유사도 함수의 정의하였다. 문자 특징은 획과 획들 사이 특징값으로 나누어지며, 이는 크게 이루는 각과 길이비로 나누어질 수 있다. 이를 특정 벡터로 선정함으로써 빠른 처리 속도로 유용한 특징 벡터를 선정할 수 있었다. 또한 유사도 함수를 정의함으로써 향후 서체 인식 등에도 활용할 수 있는 한 방법을 제시하였다. 또한 입력의 자동화 분야에 적용하기 위한 문자 인식을 문자 익히기까지 확대 적용할 수 있는 가능성의 제시와 이에 대한 방법론의 제안을 행하였다.

이상의 제안한 방법에 대해 실험을 행한 결과 효율적으로 신명조체 40포인트 크기에 대한 표준문자 영상을 구성할 수 있었다. 그러나 현재까지 행해진 실험은 자음에 대해서만 이루어져 있다. 차후로는 모음과 밑받침에 대해 특징 추출을 행하는 연구가 지속적으로 이루어져야 하리라 여겨진다. 또한 자음, 모음 그리고 밑받침에 대해 공간적 위치 관계를 설정하기 위한 연구도 계속 행해져야 하리라 사려된다.

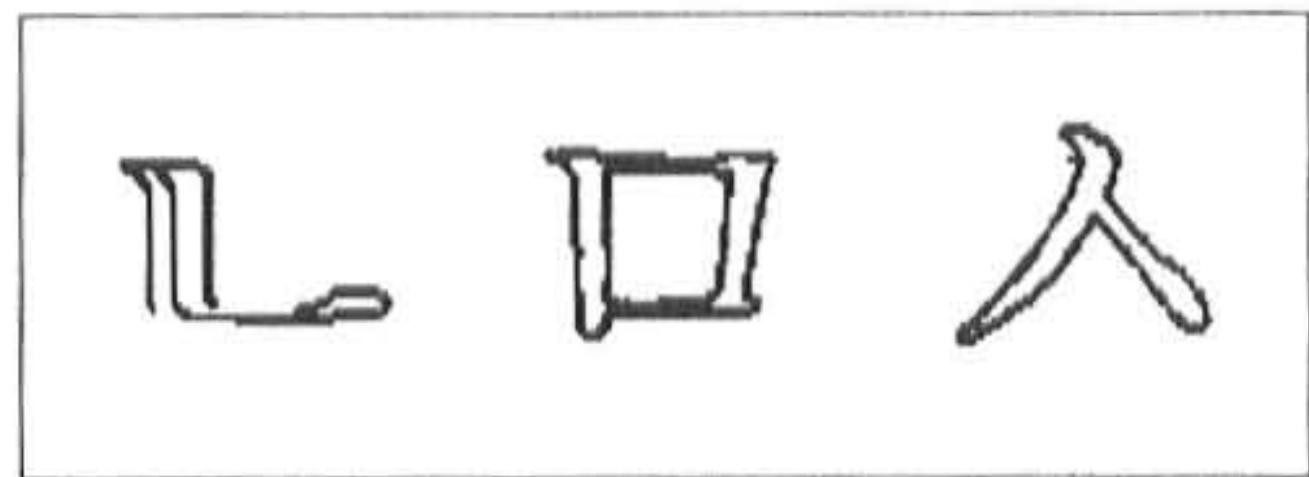


Fig.6 Standard characters image

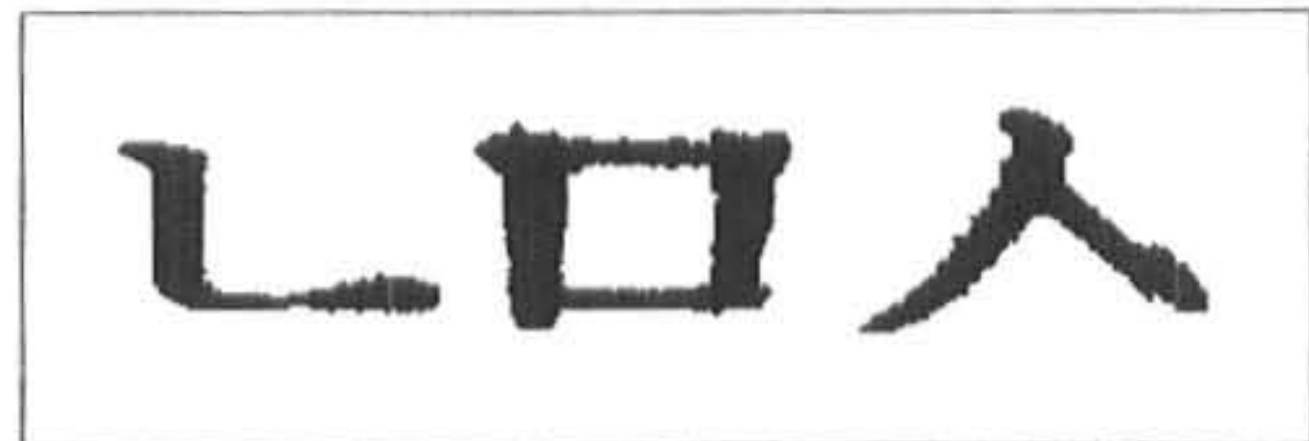


Fig.7 Results of edge detection

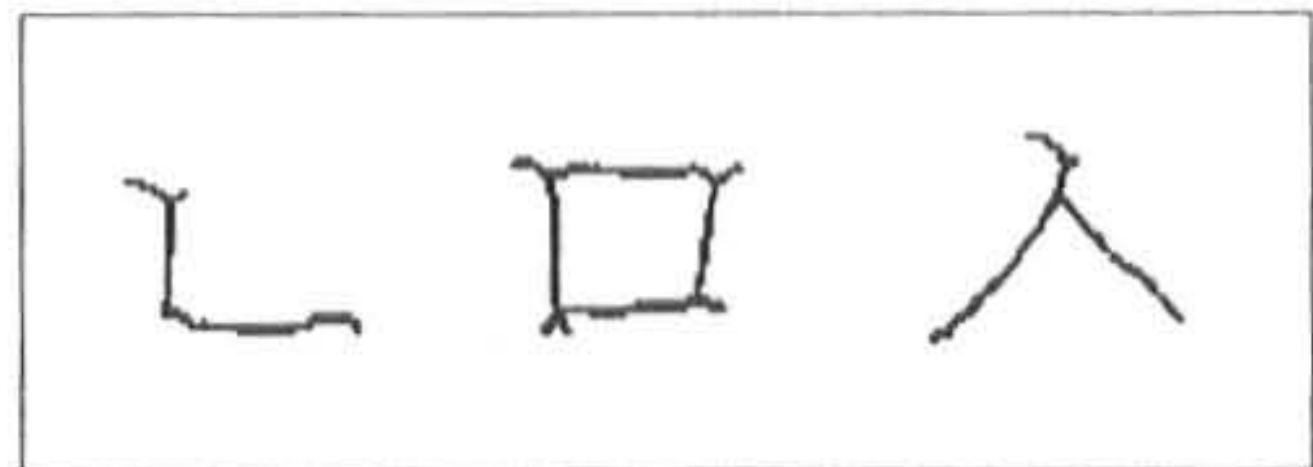


Fig.8 Thinned results

Table2. Feature extracted results for standard characters image

대상문자	획의 특징 추출	획들 사이의 특징 추출	
		Intersection Angle	Length Ratio
ㄴ	90, 0	90	0.9
		90	0.9
ㄹ	90, 0 90, 0	90, 180, 90	0.615, 1, 0.8
		90, 180, 90	1.625, 1.3, 0.8
		90	
ㅁ	45, 135	90	1.3
		90	

그리고 Table 2~4에서의 유사도 함수 정의는 model based 방법과 CAD 방법을 결합하여 미적 평가를 위한 정의 식으로 규정하였다. 특히 앞으로 중점적으로 연구가 되어야 할 부분은 입력 문자 패턴에 대한 처리와 미적 평가 계산



**Table3. Similarity measuring function for calculating beauty evaluation of 'ㄴ'**

1	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 $95 \leq x \leq 110$
2	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 $5 \leq x \leq 20$
3	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 $95 \leq x \leq 105$
4	0 for $x \leq 0.5$ or $x \geq 1.3$ , $\frac{0.5-x}{-0.2}$ for $0.5 \leq x \leq 0.7$ 1 for $0.7 \leq x \leq 1.1$ , $\frac{x-1.3}{-0.2}$ for $1.1 \leq x \leq 1.3$

**Table4. Similarity measuring function for calculating beauty evaluation of 'ㄹ'**

1	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 $95 \leq x \leq 110$
2	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 $5 \leq x \leq 20$
3	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 $95 \leq x \leq 110$
4	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 $5 \leq x \leq 20$
5	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 $95 \leq x \leq 105$
6	0 for $x \leq 160$ or $x \geq 200$ , $\frac{160-x}{-15}$ for $160 \leq x \leq 175$ , 1 for $175 \leq x \leq 185$ , $\frac{x-200}{-15}$ for $185 \leq x \leq 200$
7	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 $95 \leq x \leq 105$
8	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 $95 \leq x \leq 105$

9	0 for $x \leq 165$ or $x \geq 195$ , $\frac{165-x}{-10}$ for $165 \leq x \leq 185$ , 1 for $175 \leq x \leq 185$ , $\frac{x-195}{-10}$ for $185 \leq x \leq 195$
10	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 $95 \leq x \leq 105$
11	0 for $x \leq 0.215$ or $x \geq 1.015$ , $\frac{0.215-x}{-0.2}$ for $0.215 \leq x \leq 0.415$ , 1 for $0.415 \leq x \leq 0.815$ , $\frac{x-1.015}{-0.2}$ for $0.815 \leq x \leq 1.015$
12	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 $1.2 \leq x \leq 1.4$
13	0 for $x \leq 0.4$ or $x \geq 1.2$ , $\frac{0.4-x}{-0.2}$ for $0.4 \leq x \leq 0.6$ , 1 for $0.6 \leq x \leq 1.0$ , $\frac{x-1.2}{-0.2}$ for $1.0 \leq x \leq 1.2$
14	0 for $x \leq 1.225$ or $x \geq 2.025$ , $\frac{1.225-x}{-0.2}$ for $1.225 \leq x \leq 1.425$ , 1 for $1.425 \leq x \leq 1.825$ , $\frac{x-2.025}{-0.2}$ for $1.825 \leq x \leq 2.025$
15	0 for $x \leq 0.9$ or $x \geq 1.7$ , $\frac{0.9-x}{-0.2}$ for $0.9 \leq x \leq 1.1$ , 1 for $1.1 \leq x \leq 1.5$ , $\frac{x-1.7}{-0.2}$ for $1.5 \leq x \leq 1.7$
16	0 for $x \leq 0.4$ or $x \geq 1.2$ , $\frac{0.4-x}{-0.2}$ for $0.4 \leq x \leq 0.6$ , 1 for $0.6 \leq x \leq 1.0$ , $\frac{x-1.2}{-0.2}$ for $1.0 \leq x \leq 1.2$

**Table5. Similarity measuring function for calculating beauty evaluation of 'ㄷ'**

1	0 for $x \leq 25$ or $x \geq 65$ , $\frac{40-x}{-15}$ for $25 \leq x \leq 40$ , 1 for $40 \leq x \leq 50$ , $\frac{x-65}{-15}$ for $50 \leq x \leq 60$
2	0 for $x \leq 115$ or $x \geq 155$ , $\frac{115-x}{-15}$ for $115 \leq x \leq 130$ , 1 for $130 \leq x \leq 140$ , $\frac{x-155}{-15}$ for $140 \leq x \leq 155$
3	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 $95 \leq x \leq 105$
4	0 for $x \leq 0.9$ or $x \geq 1.7$ , $\frac{0.9-x}{-0.2}$ for $0.9 \leq x \leq 1.1$ 1 for $1.1 \leq x \leq 1.5$ , $\frac{x-1.7}{-0.2}$ for $1.5 \leq x \leq 1.7$



그리고 본 시스템을 확장하여 필자 인식이나 특정한 서체 인식으로까지의 확장에 대한 연구가 지속적으로 행해져야만 한다. 아울러 개발된 시스템의 상품화에 대한 노력도 이루어져야 하리라 여겨진다.

## 5. 결 론

본 논문에서는 기존의 문자 인식 시스템이 가지고 특성을 이용하여 한글 문자 익히기 및 서체 익히기까지의 CAI 시스템으로 확장하는 방법에 대해 제안하였다.

즉, 문자 익히기를 행할 수 있는 시스템을 제안함으로써 기존 문자 인식 시스템의 적용성 확장을 기하였다. 이를 위해 신명조체 40포인트 자음에 대해 환경에 불변인 전처리 과정과 히스토그램의 특성을 이용한 자소의 인식 및 획의 분리 그리고 특징 추출하여 입력 영상의 특징값과 표준 문자 패턴의 특징값을 정합 할 수 있는 즉, 유사도를 측정할 수 있는 함수에 대해 정의하였다.

현재는 전체 시스템 중 표준 문자 영상에 대해서만 다루었으며, 이중 자음 부분의 처리에 대한 실험을 하였다. 표준문자 자소의 특징추출과 유사도 함수 정의까지 실험을 하였으며, 제안한 방법이 효과적임을 확인할 수 있었다.

따라서 앞으로 모음부 및 밑받침 부분에 대한 처리 및 전체 문자에 대한 각 자소의 분리와 이들간의 특징벡터 선정에 대한 알고리즘 개발과 실험 수행 등이 지속적으로 행해져야 하리라 여겨진다.

입력 문자 처리에 대한 연구도 지속적으로 수행되어 상용화가 가능한 전체 시스템을 구축하는 것이 향후 과제로 생각된다.

또한 현재의 시스템은 신명조체에 대해 표준 문자 패턴을 형성하는 방법을 다루었지만 차후 다른 폰트에 대해서도 실험을 수행하여 특정 폰트 익히기까지 가능한 시스템이 되도록 해야 한다.

아울러 이를 특정인의 서체에 대한 표준 문자 패턴을 생성하여 필자를 인식(writer's recognition)하는 시스템으로 확장하는 방법에 대해서도 지속적인 연구가 있어야 하리라 여겨진다.

## 참고문헌

- 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) 이성환, "문자 인식 - 이론과 실제 I, II 권", 홍릉과학출판사, 1994, 4
- 4) 이상범, "한글처리 - 문자 중심 인식 기술 고찰", 한국정보처리학회, Vol.5, No.5, 1998
- 5) 美濃導彦, "文書畫像處理의 現狀과 動向", 日本電子情報通信學會誌, Vol.76, No.5, pp. 502~509, 1993
- 6) D. Dubois and H. Prade, "Toll Sets," Proc, IFSA'91, Brussels, Artificial Intelligence, pp. 21~24, 1991.
- 7) D. U. Cho, "Noise Removal and Edge Detection of Image by Image Structure Understanding," Transactions of the Korea Information Processing Society, Vol. 4, No. 7, 1997.
- 8) K. S. Fu and A. Rosenfeld, "Pattern Recognition and Computer Vision," IEEE Computer, pp. 274~282, Oct. 1984.
- 9) Ballard & Brown, "Computer Vision", Prentice-Hall, 1982.
- 10) Tsuji, et, al., "Document image analysis based on split detection method, paper of the Technical Group on pattern Recognition and Learning," IEICE, PRL85-17, pp. 63~70, 1985
- 11) Wakayama, "Skeleton Tracing Based on Maximal Square Moving," IEICE Technical Report, PRL 78~87, 1978.

(2000년 3월 23일 접수, 2000년 8월 18일 채택)