

웨이브릿 변환과 모멘트를 이용한 문자인식에 관한 연구

조민환*

A Study on Character Recognition using Wavelet Transformation and Moment

Meen Hwan Cho *

요 약

본 논문에서는 웨이브릿 변환을 이용한 문자인식 방법 중 문자의 최소 단위인 자음과 모음을 분리시켜 문자의 모멘트를 분석하여 산출되는 정보를 사전에 컴퓨터에 입력시켜 문서화된 수기 문자를 컴퓨터에 저장하고 인식시키는 방법에 접근 하였다. 연구는 획득한 문장 이미지에서 잡음을 없애고 줄 단위로 분리, 분리된 줄 단위 문장은 한 문자 단위로 다시 분리된 후 자음과 모음으로 분리 하였다. 분리된 자소는 CVPtools를 사용하여 히스토그램 평활화와 침식 및 평균값 필터를 처리한 후 C++를 이용하여 세선화 처리하고 세선화된 자소는 팽창 및 크기 변환하여 모든 자소가 동일 크기, 크기 이미지로 만들었다. 표준화 이미지는 이진화 이미지로 변환하여 3단계 웨이브릿 변환을 이용하여 데이터의 양을 1/64로 줄인 후 해밍거리를 조사하였다. 연구 결과 다양한 'ㄱ'상호간 및 'ㅅ'상호간의 일치도는 매우 높게 나타났고, 서로 상이한 'ㄱ'과 'ㅅ'을 비교 했을 때 상호간 일치도가 매우 낮게 나옴을 알 수 있었다. 이 연구 결과로 더 많은 수기 자소들에 대한 해밍거리조사가 이루어지면 각각의 자음과 모음의 모멘트 구분하여 수기 문자 인식에 중요한 정보를 알 수 있을 것으로 판단된다.

Abstract

In this thesis, We studied on hand-written character recognition, that characters entered into a digital input device and remove noise and separating character elements using preprocessing. And processed character images has done thinning and 3-level wavelet transform for making normalized image and reducing image data. The structural method among the numerical Hanguk recognition methods are suitable for recognition of printed or hand-written characters because it is usefull method deal with distortion, so that method are applied to separating elements and analysing texture. The results show that recognition by analysing texture is easily distinguished with respect to consonants. But hand-written characters are tend to decreasing successful recognition rate for the difficulty of extraction process of the starting point, of interconnection of each elements, of mis-recognition from vanishing at the thinning process, and complexity of character combinations. Some characters associated with the separation process is more complicated and sometime impossible to separating elements. However, analysis texture of the proposed character recognition with the exception of the complex handwritten is aware of the character.

-
- 제1저자 : 조민환
 - 투고일 : 2010. 08. 11, 심사일 : 2010. 08. 30, 게재확정일 : 2010. 09. 01.
 - * 창원전문대학 메카트로닉스과 조교수

▶ Keyword : 문자인식(Character Recognition), 웨이블릿 변환(Wavelet Transform), 모멘트분석(Moment)

I. 서론

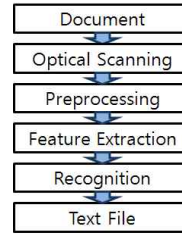
최근에는 고도 정보화 사회로 급진전하면서 매년 방대한 양의 정보가 처리되고, 이와 같은 정보의 신속한 활용 여부가 국가 경쟁력에 영향을 줄만큼 중요한 부분을 차지한다. 정보의 신속한 활용을 위해서는 전자결재체계가 구축되었지만 아직도 대부분 종이로 기록된 방대한 양의 정보가 신속하게 컴퓨터에 처리되어야 하고, 필요로 하는 사람에게 즉시 전달되어야 한다. 여기서 문자 인식의 기술은 수작업으로 컴퓨터에 입력시키는 작업을 자동으로 신속하게 입력시키는 역할을 한다[1].

1959년 BELL 연구소에서는 필기되는 문자를 인식하는 온라인 필기 문자 인식(On-line handwriting recognition)을 최초로 발표하였으며 1960년대 테블릿이 만들어지면서 시스템의 구성 요소는 매우 빠른 속도로 발전하기 시작하였다. 1970년대 말부터 인쇄체 한자 OCR 및 온라인 필기 한자의 인식에 대한 실용화 연구, 오프라인 필기체 한자 OCR의 최초 연구가 활발하게 진행되었다. 국내의 산업체에서도 상품화를 목표로 인쇄체 문서 인식 시스템을 개발하기 위한 연구가 진행된 것은 1980대 후반으로, 1989년 (주)삼보컴퓨터는 한국과학기술원 전산학과와 공동으로 초보적인 수준의 학습 기능을 갖춘 한글 문서 인식 시스템을 개발하였다. 이 문서 인식 시스템은 비록 상품화에는 성공하지 못하였지만 최초의 상업용 문서 인식 시스템이라고 볼 수 있다[2][3].

최근에 오프라인 인쇄체 및 필기체 문자를 인식하는 시스템이 상품화된 제품이 등장하고 있으며 우체국에서 우편물 자동 분류를 위한 우편 번호 인식, 사무실에서 문자 인식, 도면 분류 및 인식, 은행에서 수표 인식, 전표인식, 기업체에서 제품 분류 및 인식, 전자 도서관 등 여러 분야에서 이용하며 계속 확산될 전망이다. 오프라인 필기체 문자 인식에서는 아직도 해결해야 할 연구 과제가 많다[4].

II. 문자 인식 절차

문자 인식은 광학 스캐너로 부터 문자를 입력받아 전처리 과정을 거친 후 특징을 구하고 인식기를 통하여 정해진 문자 중의 한 문자를 분류하여 텍스트 파일로 만드는 작업이다. 문자 인식의 과정은 <그림 2-1>과 같이 전처리(Preprocessing), 특징 추출(Feature extraction), 인식(Recognition) 단계로 구분한다[5].



<그림 2-1> 문자 인식절차의 블록다이어그램

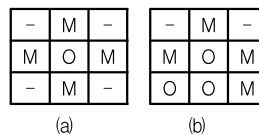
2.1 전처리

전처리(Preprocessing) 과정은 인식 대상 문서 영상으로부터 정보 영역을 분리 추출하고 잡음을 제거하여 추출된 특징의 변별력을 높이는 목적으로 많이 사용된다. 이 전처리과정은 특징 추출을 하기 이전의 단계로 다음과 같은 세부 기능을 갖는다.

2.1.1 평활화

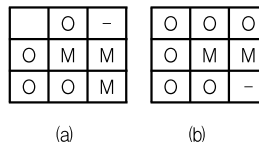
필기체 문서는 광학 스캐너를 통해 컴퓨터로 입력되는데, 스캐너의 기계적인 특성과 필기체의 문자 자체의 잡음으로 인하여 자획의 가장자리가 거칠게 입력되거나 가운데 부분에 구멍(Hole)이 존재하는 경우가 발생한다. 이러한 현상을 교정해 주는 과정이 평활화(Smoothing)이며 3×3의 윈도우를 사용하여 채우기(Filling)는 <그림 2-2(a)> 또는 <그림 2-2(b)>의 패턴과 정합이 되거나, 그들의 90°, 180°, 270° 회전으로 얻어지는 패턴과 정합이 되면 중심점을 채운다.

(O: Background, M: Stroke, -: don't care)



<그림 2-2> 3×3윈도우를 이용한 채우기

지우기(Deleting)는 <그림 2-3(a)> 또는 <그림 2-3(b)>의 패턴과 정합이 되거나, 그들의 90°, 180°, 270° 회전으로 얻어지는 패턴과 정합이 되면 중심점을 지운다.



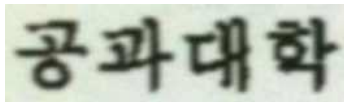
<그림 2-3> 3×3 윈도우를 이용한 지우기

2.1.2 정보 영역 추출

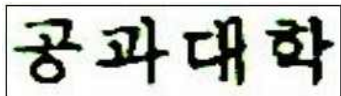
<그림 2-4>는 한글 문서를 디지털 카메라로 촬영하여 만든 디지털 영상 데이터를 보여준다. 이 때 이미지는 흑백 영상으로 각 화소의 명암도는 256이고 공간 해상도는 896×592화소로 구성되어 있다.

<그림 2-5>은 <그림 2-4>를 메디안(median)필터에 의해 잡음을 제거 후 문턱치 값을 200으로 설정하여 변환된 2진화 영상을 보여준다. <그림 2-4>의 경우 주변광이 혼합되어 배경(여백)의 밝기가 매우 어둡게 나타났으나, <그림 2-5>에서 문턱치 값을 적절하게 선택하여 배경과 글자의 비를 선명하게 할 수 있음을 보여준다.

이진화 영상 데이터에서 문자열을 찾아내는 방법으로는 허프 변환(hough-transform)과 히스토그램에 의한 문자열 방향검출, 문서 가장자리의 명암치를 이용한 문자열 방향 검색 등이 있다. 허프 변환에 의한 문자열과 문자열 사이의 공간 검사 방법은 처리 시간이 너무 길어 잘 사용하지 않으며, 문서 가장자리의 명암치를 이용한 문자열 방향 검색은 문서에 따라 문서 가장자리의 명암 변화가 없는 경우가 많이 있어 대부분 히스토그램을 이용한 문자열 검색을 많이 사용하고 있다[6][7].



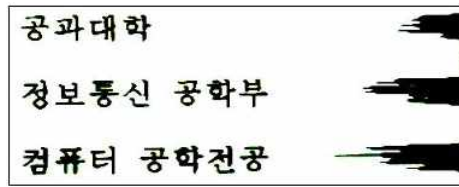
<그림 2-4> 원시 이미지 데이터



<그림 2-5> 잡음이 제거된 이진화 이미지 데이터

<그림 2-6>은 <그림 2-7>의 2진화 데이터 중 일부를 선택하여 조금씩 회전시키면서 y축에 투사 할 때 문자열 방향과 투사 방향이 일치 한 경우를 보여준다. y축의 스크린에 나타난 히스토그램에서 가장 밝은 부분(누적 값이 0인 경우)은 문자열과 문자열 사이의 줄 간격이 된다.

2진 데이터에서 한 줄을 검출 한 후 다시 한 문자를 분리하는 방법도 문자열 검색과 동일하게 히스토그램을 이용하여 문자 발체를 할 수 있다. 문자열은 각 문자와 문자 사이는 수직적으로 모두 띄어져 있으므로 x축에 투사 할 경우 띄어진 곳의 히스토그램 값이 0이 되어 쉽게 문자들은 분리 할 수 있다. <그림 2-7>은 한 문자열에서 히스토그램에 의한 문자 분리 과정을 보여준다.



<그림 2-6> 히스토그램에 의한 문자열 검색

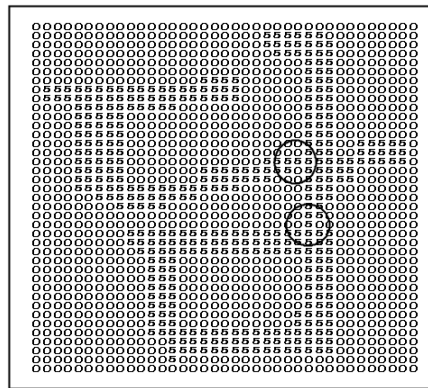


<그림 2-7> 문자열에서 문자 분리

2. 2 자소 분리

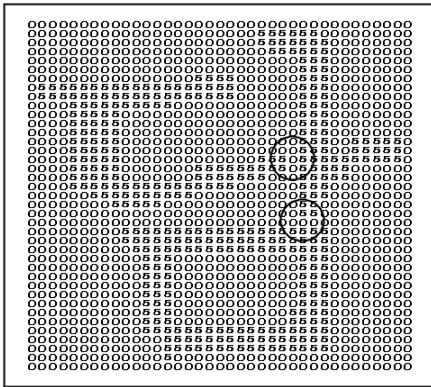
한 문자에 대한 자소 분리는 먼저 자음과 모음이 붙어 있는 경우 분리 하여야 한다. 이 경우 수직모음(V1)의 상단점은 문자의 좌단점과 일치하며, 수평모음의 좌단점 시작점은 문자의 좌단점과 일치하는 특성을 이용하여 붙어있는 자음과 모음을 분리한다. <그림 2-8>은 자음 'ㄷ' 과 'ㅁ'이 모음 'ㅏ'와 붙어있는 것을 보여주며 <그림 2-9>는 분리된 형태를 보여준다.

분리된 자소들은 영역 분할 방법을 사용하여 자소별 라벨을 하였다. <그림 2-10>은 영역 분할에 의해 라벨링된 결과를 보여준다. 영역 분할 알고리즘은 한 문자의 위에서 아래로, 좌에서 우로 탐색하여 화소가 0보다 크고 서로 4이웃 화소 중 한 화소라도 연결되면 동일한 라벨

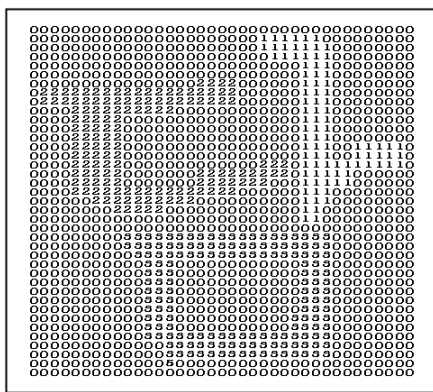


<그림 2-8> 자음과 모음이 연결된 문자

을 갖게 하였다. 인쇄체 한글 문자의 특성상 수직 모음의 끝점이 가장 먼저 라벨링 처리되므로 <그림 2-10>에서 'ㅏ' 화소가 제일 먼저 라벨링 된다.



<그림 2-9> 자음과 모음이 분리된 문자



<그림 2-10> 라벨링된 화소

2.3 특징 추출

필기체 문자 오프라인 인식에서는 필자에 따라 필체가 달라지고 동일한 사람이라도 환경에 따라 변화가 불규칙하며 필기 문자의 크기 및 기울어짐의 불일치 등 다양한 문자의 변형, 입력 장치의 특성으로 인한 잡음 그리고 정보 손실이 발생한다. 필기체 문자 인식 성능을 향상시키기 위해서는 이와 같은 변형을 흡수할 수 있는 특징을 찾는 일이 중요하며 실용적인 문서 인식 시스템을 위해서는 특징 추출 시간도 고려되어야 한다. 특징 추출 방법으로 윤곽선을 이용한 4방향 특징, 모포로지를 이용한 4방향 특징, 고리를 이용한 투영 및 망 특징, 교차거리+교차+투영+망 특징방법 중에서 모포로지를 이용한 4방향 특징을 사용하였다.

2.3.1 모포로지를 이용한 4방향 특징

기하학적 연산자들은 영상에서 기하학적 형태와 패턴으로 정의되어 있고 윤곽선 추출, 잡음 감쇄 필터링, Edge 검출, 세선화 등 영상 처리의 많은 분야에서 발전되고 있으며, 간단

한 논리적 연산과 병렬 처리 및 실시간 응용이 가능한 장점을 가지고 있다. 문자 영상의 집합 A 의 윤곽선을 $\beta(A)$ 라 하면 식(1)과 같이 윤곽선을 구할 수 있다.

$$\beta(A) = A - (A \oplus B) \dots\dots\dots (1)$$

B is circle structuring element

0	1	0
1	1	1
0	1	0

<그림 2-11> 구조화 요소 B

$\beta(A)$ 를 <그림 2-12>와 같은 커널의 B 로서 각각 Opening하면 4방향 성분을 구할 수 있다.

$$\beta(A) \circ B = (\beta(A) - B) \oplus B \dots\dots\dots (2)$$

0	0	0	1	0	0	0	1	0	0	0	1
1	1	1	0	1	0	0	1	0	0	1	0
0	0	0	0	0	1	0	1	0	1	0	0

(a) (b) (c) (d)

<그림 2-12> 4 방향의 구성 요소를 추출하기위한 구조화 요소

- (a) Horizontal structuring element
- (b) Reverse slant structuring element
- (c) Vertical structuring element
- (d) Slant structuring element

추출된 4(수직, 수평, 사선, 역사선) 방향 성분은 각각 7x5의 크기로 압축하면 140개 (7x5x4)의 특징 벡터가 된다.

2.4 세선화

세선화(Thinning) 과정은 화상 처리와 패턴 인식 분야에서 많이 사용되는 전처리과정으로써 정보량 감소, 전송 시간 감소, 그리고 구조적인 특징 추출의 용이성과 같은 기능을 갖는 것으로서 좋은 세선화 알고리즘에 대한 일반적인 사항으로는 다음과 같은 조건이어야 한다.

- 첫째. 입력 영상의 형태를 잘 반영하는 골격선을 추출해야 한다.
- 둘째. 패턴 및 패턴 배경의 연결성을 보존하는 골격선을 추출해야 한다.
- 셋째. 패턴의 중앙에 위치하는 골격선을 추출해야 한다.
- 넷째. 가능하면 동방성(isotropic)의 성질을 갖는 골격선을 추출해야 한다.
- 다섯째. 한 화소 두께를 갖는 골격선을 추출해야 한다.
- 여섯째. 입력 패턴의 끝점이 유지되도록 골격선을 추출해야 한다.
- 일곱째. 가장자리 부분의 잡음에 둔감해야 한다.

문자 영상을 압축하기 위해서는 데이터양을 최소화시켜야 할 필요성이 있다. 따라서 문자 영상을 압축하기 이전에 세선화가 전처리로 필요하다. 그러나 세선화 과정에서 필요한 정보를 잃어버리는 경우가 있으므로, 이에 대한 방안으로 세선화 알고리즘 연구가 많이 되어 왔다. 본 논문에서는 잔가지의 발생이 드물고 교차점과 분기점의 처리가 비교적 잘된다고 알려진 T. Y. Zhang이 개발한 알고리즘을 사용하였다[8].

Zhang의 알고리즘에서 제거될 수 있는지 검사되는 검은 화소를 p_1 이라 하면 <그림 2-13>과 같이 3×3의 윈도우가 있을 때 $N(p_1)$ 은 p_1 의 8인접(8-Neighborhood) 화소들 중에서 1인 화소의 수를 나타낸다.

$$N(p_1) = p_2 + p_3 + p_4 \dots + p_8 + p_9 \dots \dots \dots (3)$$

P9	P2	P3
P8	P1	P4
P7	P6	P5

<그림 2-13> 3×3 윈도우를 이용한 세선화

$S(p_1)$ 은 $p_2, p_3, p_4, \dots, p_8, p_9$ 의 순으로 화소값을 조사하여 화소값이 1로 변화하는 경우 조사하여 누적한 값이다. 골격의 연결성을 유지하기 위하여 두 단계의 실행 과정을 거친다.

<단계1> 다음 조건을 만족하면 화소 p_1 을 삭제한다.

- (a) $2 \leq N(p_1) \leq 6$
- (b) $S(p_1) = 1$
- (c) $p_2 \cdot p_4 \cdot p_6 = 0$
- (d) $p_4 \cdot p_6 \cdot p_8 = 0$

조건 (a), (b)는 4인접 영역의 연결성을 유지하고 (d)는 8인접성을 유지하기 위한 조건이다.

<단계2> 다음 조건을 만족하면 화소 p_1 을 삭제한다.

- (a) 상기 (a), (b)와 동일 조건
- (b) $p_2 \cdot p_4 \cdot p_8 = 0$
- (c) $p_2 \cdot p_6 \cdot p_8 = 0$

이상의 두 과정을 반복 수행하면서 더 이상 삭제되는 화소가 없으면 세선화 과정이 완료된다.

III. 웨이브릿 변환

웨이브릿 변환은 임의의 신호를 시간과 주파수 공간으로 표현할 수 있는데 주파수 공간은 다시 저주파와 고주파로 나누어진다. 전체 신호에 대해 저주파는 에너지를 90%이상 가

지고 있어 에너지를 원하는 영상에 집중시킬 수 있다[9]. 특히 문자 영상처럼 문자와 배경사이 그레이 값의 변화가 심한 영상에서는 에너지를 문자에 정확히 집중시키고 정보량도 줄일 수 있는 장점이 있다. 일반적으로 문자 인식에서 대상문자를 양자화 하는 전처리는 인식에 불필요한 잡음 제거 등 인식 성능을 좌우할 만큼 중요한 요인의 하나이다[10].

웨이브릿 변환 계수(wavelet transform coefficient)중 임계값보다 작은 웨이브릿 변환 유효 계수(WTS: Wavelet Transform Significant coefficient)는 다치 영상으로 무손실 압축(lossless compression)으로 원래의 영상으로 복원할 수 있는 중요한 정보를 가지고 있다[1].

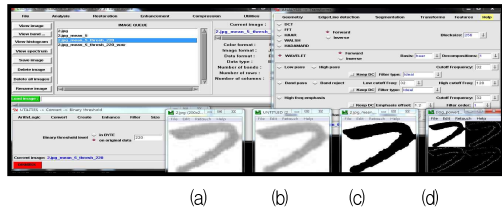
본 논문에서는 미국 Kentucky, Louisville University의 Computer Vision and Image processing Lab에서 개발한 CVIPtools 응용 S/W를 이용한다. CVIPtools를 이용하여 웨이브릿변환을 이용한 1차원 신호의 분해는 낮은 대역에 대해 계속 반복적인 분해를 하는 것으로 저주파 대역의 폭을 반감시키는 것이며 이는 2배의 주파수 해상도를 얻음을 의미한다. 또한 1/2의 감축(subsampling)을 통해 시간축의 해상도는 반으로 감소하게 된다. 저주파 대역 쪽으로 주파수 해상도를 2배씩 증가하는 대역 분할을 옥타브 웨이브릿(octave or dynamic wavelet) 대역 분할이라고 한다. 256×256 한글 문자 영상을 저주파 대역을 가지고 원하는 단계까지 서브샘플링을 한다. 마지막 단계의 최저주파 대역에서 임계값 보다 작은 값인 웨이브릿 유효 계수(wavelet transform significant coefficient: WTS)를 본 논문에서는 트렌드(Trend)이미지로 정의하였다.



<그림 3-1> '기'의 2단계 웨이브릿 변환

y축 투영에 의해 행을 발췌한 다음 x축 투영으로 문자 한자를 추출하여 y축 자소 분리 방법에 의해 분리된 자소 '기'에 대한 웨이브릿 변환을 CVIPtools를 사용하여 보여 준다<그림 3-1> '기'의 2단계 웨이브릿 변환. <그림 3-2(a)>는 분리된 자소 원래 이미지이며 <그림 3-2(b)>는 평균필터링을 한

이미지를 보여 준다. <그림 3-2(c)>는 <그림 3-2(b)> 이미지를 평균 밝기의 0.8배 문턱치 값으로 만든 이진화 영상이다. 이진화 영상을 CVIptools를 이용하여 1-Level Daub4 웨이블릿 변환시킨 이미지가 <그림 3-2(d)>이다.



<그림 3-2> '기' 자소의 단계별 변환

- (a) Separated original image
- (b) Mean filter mask size 5
- (c) Binary threshold level image
- (d) Wavelet Daub4 1-Level image

<표 4-1> 자소분리 및 '기' 평균값 이미지

소	자소 분리			웨이브릿변환 (초성)		
	초성	중성	종성	정규화	평활화	평균값 이미지
강	기	ㅏ	ㅇ	기	기	기
길	기	ㅣ	ㅇ	기	기	기
강	기	ㅏ	ㅇ	기	기	기
긴	기	ㅏ	ㅇ	기	기	기
김	기	ㅣ	ㅇ	기	기	기
김	기	ㅣ	ㅇ	기	기	기
김	기	ㅣ	ㅇ	기	기	기
김	기	ㅣ	ㅇ	기	기	기
김	기	ㅣ	ㅇ	기	기	기
김	기	ㅣ	ㅇ	기	기	기
김	기	ㅣ	ㅇ	기	기	기

IV. 분리된 한글자소 인식을 위한 모멘트 적용

4.1 모멘트 적용을 위한 패턴

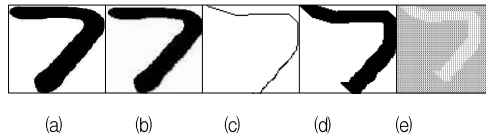
CVIptools를 이용하여 초성이 '기'와 'ㅏ'를 구분하여 비교하였다. <표 4-1> 및 <표 4-2>는 자소분리를 한 후 초성을 분리하였고 분리된 초성을 정규화(픽셀 150×150) 시켰으며 웨이블릿 변환을 통하여 평활화 시킨 후 초성의 평균값 이미지를 구하였다. 또한 평균값 이미지에 세선화한 후 정규화를 위해 일정한 굵기(9픽셀)의 자소를 생성시켜 <그림 4-1>의 결과를 만들었다.

<그림 4-1>에서 편의상 좌측 '기'부터 오른쪽으로 1번에서 10번까지 번호를 부여 했으며 'ㅏ'은 좌측부터 11번에서 20번 까지 번호를 부여하여 결과를 산출하였다.

<표 4-2> 자소분리 및 'ㅏ' 평균값 이미지

소	자소 분리			웨이브릿변환 (초성)		
	초성	중성	종성	정규화	평활화	평균값 이미지
승	ㅏ	ㅡ	ㅇ	ㅏ	ㅏ	ㅏ
수	ㅏ	ㅓ	ㅣ	ㅏ	ㅏ	ㅏ
신	ㅏ	ㅣ	ㅇ	ㅏ	ㅏ	ㅏ
석	ㅏ	ㅣ	ㅇ	ㅏ	ㅏ	ㅏ
수	ㅏ	ㅓ	ㅣ	ㅏ	ㅏ	ㅏ
승	ㅏ	ㅡ	ㅇ	ㅏ	ㅏ	ㅏ
송	ㅏ	ㅣ	ㅇ	ㅏ	ㅏ	ㅏ
성	ㅏ	ㅡ	ㅇ	ㅏ	ㅏ	ㅏ
성	ㅏ	ㅡ	ㅇ	ㅏ	ㅏ	ㅏ
선	ㅏ	ㅣ	ㅇ	ㅏ	ㅏ	ㅏ

변환된 평균값 이미지는 크기나 글자의 굵기 등이 제각각 이어서 인식을 위한 정규화 기법이 필요하다. 이를 위하여 본 논문에서 <그림 4-1>에서 첫 번째 ‘ㄱ’ 이미지의 결 분석을 <그림 4-2>의 예로 설명하고자 한다. <그림 4-2(a)>는 웨이브릿 변환된 이미지를 보여주며 세선화 하기 전 경계부분의 요철을 없애기 위하여 팽창 연산을 하였다. <그림 4-2(b)>는 팽창 연산을 한 이미지를 보여주며, <그림 4-2(b)>를 세선화하면 굵기에 관계없이 모두 한 픽셀 크기로 자소 이미지가 생성된다. 이렇게 생성된 이미지는 3×3 마스크를 이용하여 5회 팽창을 하면 자소의 굵기는



<그림 4-2> 문자 ‘ㄱ’의 결 분석 단계별 이미지
 (a) Trend image (b) Dilated image
 (c) Thinning (d) Normalized image
 (e) Labeling image

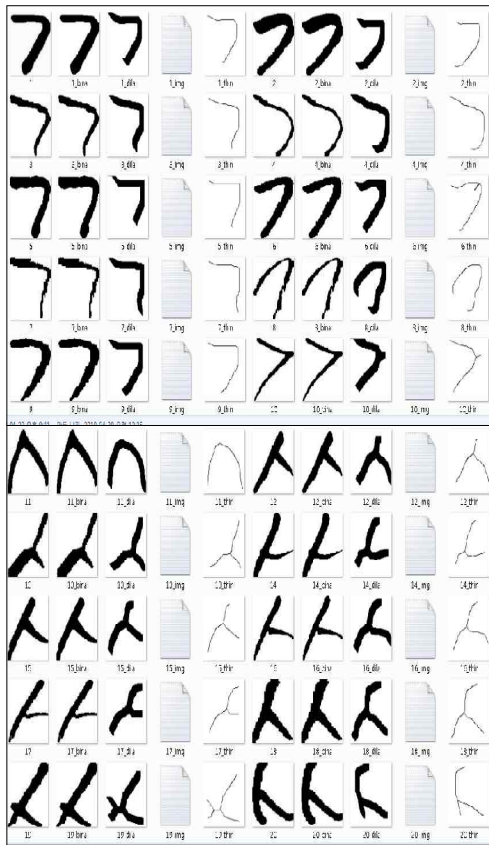
V. 실험 결과 및 분석

5.1 실험 환경

분석 시스템은 Windows 7 상에서 C++로 구현되었고, 입력 데이터는 수기로 작성되어 디지털카메라 와 스캐너를 이용하여 PC에 저장하였다. 입력된 데이터가 한글 인 점에서 초성, 중성, 종성을 Microsoft사의 그림판에서 편의상 분류하여 150×150 픽셀로 저장하였으며 CVIptools를 사용하여 분리된 문자의 이미지를 필터링하고 정형화하여 2진화 트렌드 이미지를 만들었다. C++를 이용하여 2진화 트렌드 된 이미지를 이용하여 세선화와 팽창을 이용한 표준화된 문자를 분석하였다.

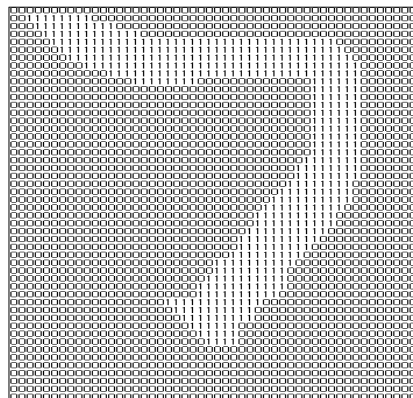
5.2 결과 및 분석

<그림 5-1>은 분리된 수기 ‘ㄱ’자소의 표준이미지를 보여주며 <그림 5-2>는 분리된 수기 ‘ㅅ’자소의 표준이미지를 보여준다. <그림 5-1>과



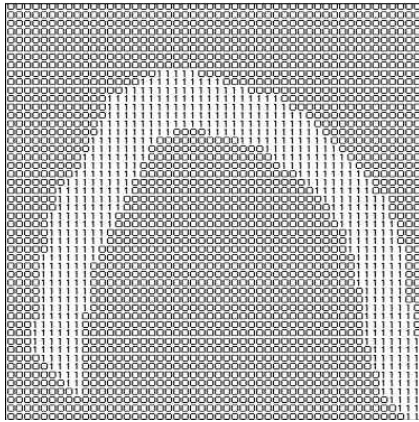
<그림 4-1> ‘ㄱ’과 ‘ㅅ’의 결 분석 결과

11픽셀로 동일 시킬 수가 있다. 동일된 굵기의 자소 이미지는 150×150 크기로 정규화하여 최종 <그림 4-2(e)>와 같이 라벨링을 한다.



<그림 5-1> 문자 ‘ㄱ’의 결 분석 예

<그림 5-2>는 모두 이미지 크기가 150×150 크기로 인쇄 이미지로 표현하기 위하여 50×50로 축소된 이미지이다.



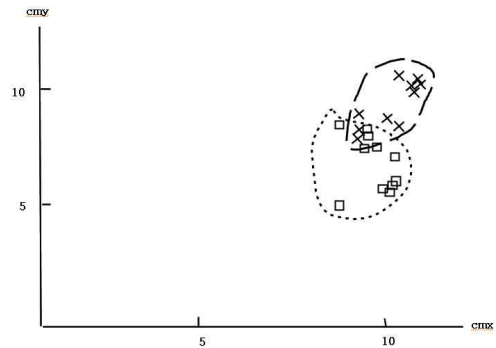
<그림 5-2> 문자 '시'의 결 분석 예

<표 5-1>는 'ㄱ'의 10개의 x-축 모멘트와 y-축 모멘트를 보여주며, <표 5-2>에서 '시'의 10개의 x-축 모멘트와 y-축 모멘트를 보여준다. <표 5-1>와 <표 5-2>에서 x-축 모멘트와 y-축 모멘트를 10000으로 나누어 정수화 한 x-y 벡터 공간에 나타내면 <그림 5-3>으로 나타난다.

<표 5-1> 수기문자 'ㄱ'상호간 비교 <표 5-2> 수기문자 '시'상호간 비교

cmx	cmx	cmx	cmx
109172	70253	114129	110728
107610	95799	114342	89784
112472	62830	124368	117795
127356	82251	88292	88927
112027	60500	87064	89116
96041	87632	119663	112069
118030	59649	98011	86544
100878	80400	113250	114186
115452	68650	109920	114211
87058	57451	82422	100842

<그림 5-3>에서 'ㄱ'과 '시'그룹의 클러스터들은 대체로 결정경계d로부터 잘 분리됨을 알 수 있다.



<그림 5-3> 벡터공간에서 수기 문자 'ㄱ' (□)와 '시' (×)의 분포

VI. 결 론

문자 인식에 관한 연구는 오래 전부터 서양에서 시작되어 많은 결과가 있으나 수기 문자에 관해서는 아직도 많은 연구가 진행 중이다. 특히 한글은 세계적으로 우수성을 자랑하고 있지만 영어나 숫자보다 복잡한 구조를 가지고 있고 이미 문서화된 데이터의 만여 가지 문장 구조에서 수기 문자를 인식하는 것은 아직도 많은 과제를 가지고 있다.

본 논문에서는 웨이브릿 변환을 이용한 연구가 여러 분야에서 응용 되어 지고 있는 점에서 착안하여 문자인식을 위한 방법 중 문자의 최소 단위인 자음과 모음을 분리시켜 문자의 모멘트를 분석하여 산출되는 정보를 사전에 컴퓨터에 입력시켜 문서화된 수기 문자를 컴퓨터에 저장하고 인식시키는 방법에 접근 하였다.

연구 방법으로 획득한 문장 이미지에서 잡음을 없애고 줄 단위로 분리, 분리된 줄 단위 문장은 한 문자 단위로 다시 분리된 후 자음과 모음으로 분리 하였다. 분리된 자소는 CVIPtools를 사용하여 히스토그램 평활화와 침식 및 평균값 필터를 처리한 후 C++를 이용하여 세선화 처리를 하고 세선화된 자소는 팽창 및 크기 변환을 하여 모든 자소가 동일 굵기, 크기 이미지로 만들었다. 표준화 이미지는 이진화 이미지로 변환하여 3 단계 웨이브릿 변환을 이용하여 데이터의 양을 1/64로 줄인 후 해밍거리를 조사하였다.

연구 결과 다양한 'ㄱ'상호간 및 '시'상호간의 일치도는 매우 높게 나타났고, 서로 상이한 'ㄱ'과 '시'을 비교 했을 때 상호간 일치도가 매우 낮게 나옴을 알 수 있었다. 이 연구 결과로 더 많은 수기 자소들에 대한 해밍거리조사가 이루어지면 각각의 자음과 모음의 모멘트 구분하여 수기 문자 인식에 중요한 정보를 알 수 있을 것으로 판단된다.

하지만, 수기 문자를 모멘트 분석함에 있어 난해한 형태의 수기 문자는 분석이 어렵다는 점이 아직도 많은 연구과제로 남아 있다. 수기의 경우 ‘ㄱ’과 ‘ㅇ’이 굴절 정도에 따라 모멘트 분석의 정확도가 떨어질 때도 있었다. 또한 더 높은 인식률을 위하여 향후 자음과 모음 연결부분의 분리에 대한 연구가 필요하다.

참고문헌

- [1] 김문호, 유기형, 곽훈성, “웨이브릿 변환을 이용한 필기체 한글 문자의 세션화 알고리즘,” 전자공학회 하계종합학술 논문집, 745-748쪽, 1998년.
- [2] 이성환, “문자인식 이론과 실제” 흥능과학출판사, 27-31쪽, 1993년.
- [3] 이영석, 김석광, “OCR의 기술 동향 및 인식 기술,” 기술 정보 시리즈, 제 21권, 1988년.
- [4] 박창순, “오프라인 필기체 한글 및 숫자 인식에 관한 연구,” 박사학위논문, 동아대학교, 1996년.
- [5] 조민환, “HMM을 이용한 수기숫자 인식에 관한연구,” 한국컴퓨터정보학회 논문집, 제 9권, 제 3호, 121-125쪽, 2004년 9월.
- [6] E. Trupin and Y. Lecourtier, “A modified contour following algorithm applied to document segmentation,” 11th LAPR, International Conference on Pattern Recognition, The Hague, The Netherlands, 1992년 8월.
- [7] Rafael C. Gonzalez, Richard E. Woods, “*Digital Image Processing*,” New Jersey: Pearson Books, 2008.
- [8] D. Wang and S. Srihari, “Classification of newspaper image blocks using texture analysis,” Computer Vision Graphics, and Image Processing, pp. 327-352, 1989.
- [9] S. G. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” IEEE trans. PAMI, vol.11, pp. 674-693, July 1989.
- [10] L. O. Gorman, “Binarization and multithresholding of document images using connectivity,” CVGIP: Graphical models and image processing, vol.56, no.6, pp. 494-506, 1994.

저자 소개



조민환

1989년 3월: 숭실대학교 산업대학원
전자계산학과(공학석사)
2001년 2월: 경남대학교 대학원
컴퓨터공학과(공학박사)
1994년 3월 ~ 현재: 창원전문대학
메카트로닉스과
조교수