

대용량 필기체 문자 인식을 위한 비선형 형태 정규화 방법의 정량적 평가

(Quantitative Evaluation of Nonlinear Shape Normalization Methods for the Recognition of Large-Set Handwritten Characters)

李晟瓌*, 朴貞善**

(Seong Whan Lee and Jeong Seon Park)

要約

최근 들어, 필기체 문자에서의 형태 변형을 보상하기 위하여 다양한 비선형 형태 정규화 방법들이 제안되었다. 본 논문에서는 지금까지 발표된 비선형 형태 정규화 방법들을 특징 투영과 특징 밀도 균일화의 두 가지 관점에서 재검토하는데, 전자는 입력 영상의 각 점에서 특정 특징을 수평축 또는 수직축으로 투영하여 특징 투영 히스토그램을 구하는 단계이고, 후자는 특징 투영 히스토그램을 재 표본화함으로써 입력 영상의 특징 밀도를 균일화하는 단계이다. 체계적인 평가를 위하여 인식률, 처리 속도, 계산 복잡도, 변형량을 평가 기준으로 설정하였으며, 이 기준을 바탕으로 각 방법을 다양한 필기체 한글 데이터에 대하여 정량적으로 평가한 결과를 제시한다.

Abstract

Recently, several nonlinear shape normalization methods have been proposed in order to compensate for the shape distortions in handwritten characters. In this paper, we review these nonlinear shape normalization methods from the two points of view: feature projection and feature density equalization. The former makes feature projection histogram by projecting a certain feature at each point of input image into horizontal- or vertical-axis and the latter equalizes the feature densities of input image by re-sampling the feature projection histogram. A systematic comparison of these methods has been made based on the following criteria: recognition rate, processing speed, computational complexity and measure of variation. Then, we present the result of quantitative evaluation of each method based on these criteria for a large variety of handwritten Hangul syllables.

1. 서론

* 正會員, ** 學生會員, 忠北大學校 컴퓨터科學科
(Dept. of Computer Science, Chungbuk Nat'l Univ.)

(* 이 논문은 1992년도 한국과학재단의 목적
기초 연구비에 의하여 연구되었음)

接受日字: 1993年 3月 6日

필기체 문자에서 발생하는 변형은 크게 선형 형태 변형(linear shape variation)과 비선형 형태 변형(nonlinear shape variation)으로 분류할 수 있다. 선형 형태 변형은 이동(translation), 신축(scale), 회전(rotation), 밀림(shear) 등을 포함한 것으로 일반적인 선형 형태 정규화 기법으로 해결할 수 있지

만 불규칙적이면서 부분적으로 발생하는 비선형 형태 변형, 특히 형태 왜곡(shape distortion)은 변형의 양상을 예측할 수 없으므로 해결하기가 매우 어려운 문제이다.^[1] 그러므로, 효과적인 필기체 문자 인식을 위하여 해결해야 할 가장 중요한 문제 중의 하나는 형태 왜곡을 어떻게 보상하느냐 하는 것이라고 할 수 있다.

주로 영숫자를 인식하는데 사용되어 온 구문론적인 방법에서는 형태 왜곡에 의해 발생하는 변형을 보상하기 위한 새로운 규칙을 추가함으로써 변형을 해결하고자 하였다. 그러나, 중국의 한자, 일본의 Kanji, 우리나라의 한글 등과 같이 문자의 형태가 복잡하고, 인식해야 할 문자 부류의 수가 방대한 대용량의 필기체 문자 인식에 구문론적인 방법을 적용하기에는 한계가 따른다. 다른 인식 방법인 패턴 정합 방법^[2]에서는 변형을 보상하기 위하여 방향 패턴의 추출^[3, 4] 또는 흐름 기법^[4] 등과 같은 고차원의 특징 추출 기법을 적용한다. 이와 같은 패턴 정합 방법이 구문론적 방법에 비하여 대용량의 필기체 문자 인식에 보다 효과적인 것으로 알려져 최근 들어서는 필기체 한자나 Kanji의 인식에 패턴 정합 방법을 널리 사용하고 있는 추세이다.

상관 방법(correlation method)은 패턴 정합 방법의 하나로써 수학적인 모델링이 쉽고 빠른 처리가 가능하기 때문에, 필기체 문자 인식에 이 방법을 적용한 논문이 다수 발표되고 있다.^[1, 4] 이러한 상관 방법에서는 패턴의 형태 정규화가 정합 과정 이전에 수행되어야 한다. 왜냐하면, 상관 방법에서는 각각의 표본화 점을 미리 알고 있다는 가정하에 표본화된 각 점들의 차이를 측정함으로써 유사도를 계산한 다음, 유사도가 높은 문자를 인식 문자로 결정하기 때문이다. 따라서, 상관 방법에서는 정합 과정을 수행하기 전에 미리 표본화할 점의 위치를 고정시키기 위한 정규화 과정이 매우 중요하다.^[1]

일반적으로 정규화의 목적은 외접 다각형 또는 모멘트(moment)^[5] 등을 이용하여 영상을 일정한 위치, 크기, 기울기 등을 갖는 영상으로 변환하는데 있다. 이러한 변환은 식 (1)을 사용하여 영상의 좌표 (x, y)를 새로운 좌표 (x', y')로 변환함으로써 가능하게 된다.

$$\begin{aligned} x' &= a_1x + a_2y + a_3 \\ y' &= a_4x + a_5y + a_6 \end{aligned} \quad (1)$$

단, $a_i, i = 1, 2, \dots, 6$ 은 상수임.

이와 같은 선형 변환은 선형성을 보존하는 특성이

있으며, 수학적인 표현이 쉬운 반면^[6], 불규칙적이면서 부분적으로 발생하는 필기체 문자의 형태 왜곡을 보상하기에는 역부족이기 때문에, 최근 들어서는 필기체 문자에서 발생하는 형태 왜곡을 보상하기 위하여 다양한 비선형 형태 정규화 방법들이 제안되고 있다.^[1, 2, 4, 7] 그러나, 제안된 다양한 비선형 형태 정규화 방법들에 대한 객관적인 성능 평가 및 최근의 연구 동향을 파악할 수 있는 체계적인 사례 연구는 이루어져 있지 않은 실정이다. 본 논문에서는 기존의 비선형 형태 정규화 방법들을 특징 투영(feature projection)과 특징 밀도 균일화(feature density equalization)의 두 가지 관점에서 재검토하고, 다양한 평가 기준을 바탕으로 각 방법에 대한 정량적 평가 결과를 제시함으로써, 추후 체계적인 필기체 문자 인식 연구에 기여함을 목적으로 한다.

본 논문의 구성은 다음과 같다. II장에서는 선형 형태 정규화 방법과 4 종류의 비선형 형태 정규화 방법들을 특징 투영과 특징 밀도 균일화의 두 가지 관점에서 재검토한다. III장에서는 각 방법을 체계적으로 평가하기 위한 평가 기준을 설정한 다음, 한글 사용 빈도순 상위 520자의 필기체 문자 데이터에 대하여 각 방법을 정량적으로 비교 분석한 결과를 소개한다. 마지막으로, IV장에서 결론 및 앞으로의 연구 방향을 제시한다.

II. 형태 정규화 방법의 재검토

본 장에서는 지금까지 발표된 여러 가지 형태 정규화 방법들을 입력 영상의 각 점에서 특징 특징을 수평축 또는 수직축으로 투영하여 특징 투영 히스토그램을 구하는 특징 투영과 특징 투영 히스토그램을 재표본화함으로써 입력 영상의 특징 밀도를 균일화하는 특징 밀도 균일화의 두 가지 관점에서 재검토한다. 형태 정규화 방법은 선형 형태 정규화 방법과 비선형 형태 정규화 방법으로 나누어지는데, 비선형 형태 정규화 방법은 다시 투영하는 특징의 종류에 따라 점 밀도^[4]를 이용한 방법과 획 밀도를 이용한 방법으로 나눌 수 있다. 획 밀도를 이용한 방법은 교차 횟수(crossing count)에 바탕을 둔 획 밀도^[7]를 이용한 방법, 획 간격(line interval)에 바탕을 둔 획 밀도^[2]를 이용한 방법, 그리고 내접원(inscribed circle)에 바탕을 둔 획 밀도^[1]를 이용한 방법으로 구분된다.

각 방법의 비교를 보다 명확히 하기 위하여 기호를 아래와 같이 정의하기로 한다.

$$f(i, j): I \times J \text{ 크기의 입력 이진 영상}(i = 1, 2, \dots, I, j = 1, 2, \dots, J)$$

$G(m, n)$: $M \times N$ 크기의 정규화 이진 영상($m = 1, 2, \dots, M, n = 1, 2, \dots, N$)

$H(i)$: 입력 영상의 점 (i, j) 에서 특정 특징을 수평축으로 투영하는 함수($j = 1, 2, \dots, J$)

$V(j)$: 입력 영상의 점 (i, j) 에서 특정 특징을 수직축으로 투영하는 함수($i = 1, 2, \dots, I$)

(m, n) : 입력 영상 $f(i, j)$ 의 좌표 (i, j) 에 대응하는 정규화 영상 $G(m, n)$ 의 좌표

이와 같은 기호 정의에 따르면 특징 투영은 함수 $H(i)$ 와 $V(j)$ 를 계산하는 과정이라고 할 수 있고, 특징 밀도 균일화는 입력 영상 $f(i, j)$ 의 좌표 (i, j) 에 대응하는 정규화 영상 $G(m, n)$ 의 좌표 (m, n) 을 구하는 과정이라고 할 수 있다.

1. 선형 형태 정규화(방법 1)

선형 형태 정규화(Linear Shape Normalization : LSN)는 문자 인식에 널리 사용되어 왔던 방법으로 일반적으로 식 (1)의 형태로 정의된다. 식 (1)은 이동, 신축, 회전 등의 다양한 선형 형태 정규화를 포함한 방법이지만, 본 논문에서는 신축만을 고려하기로 한다.

1) 특징 투영

신축은 입력 영상의 크기와 정규화 영상의 크기에만 의존하므로, 특징 투영 함수를 단순히 다음과 같이 정의할 수 있다.

$$H(i) = 1 \tag{2}$$

$$V(j) = 1$$

여기에서 우리는 이 함수가 입력 영상의 특징에는 영향을 받지 않는다는 것을 쉽게 알 수 있다.

2) 특징 밀도 균일화

특징 밀도 균일화는 특징 투영 히스토그램을 재 표본화함으로써 입력 영상의 특징 밀도를 균일화하는 단계로써 정규화 영상의 좌표 (m, n) 을 다음과 같이 계산할 수 있다.

$$m = \sum_{k=1}^i H(k) \times \frac{M}{\sum_{k=1}^i H(k)}$$

$$n = \sum_{l=1}^j V(l) \times \frac{N}{\sum_{l=1}^j V(l)} \tag{3}$$

식 (2)를 식 (3)에 대입하면, 식 (3)은 일반적으로 널리 알려진 다음과 같은 단순한 형태의 크기 정규화 식이 된다.

$$m = i \times \frac{M}{I}$$

$$n = j \times \frac{N}{J} \tag{4}$$

2. 비선형 형태 정규화

본 절에서는 여러 가지 비선형 형태 정규화(Nonlinear Shape Normalization : NSN) 방법들을 각 방법에서 사용한 입력 영상의 특징에 따라 구분하여 재검토한다.

2.1. 점 밀도를 이용한 NSN(방법 2)

이 방법은 Yamashita 등^[4]이 제안한 방법으로 입력 영상에서의 검은 화소의 수, 다시 말해서 점 밀도를 특징으로 사용한다.

1) 특징 투영

점 밀도를 이용한 NSN 방법의 특징 투영 함수는 입력 영상을 왼쪽에서 오른쪽 또는 위에서 아래로 주사하면서, 만나는 검은 화소를 수평축과 수직축으로 투영함으로써 얻어진다.

$$H(i) = \sum_{j=1}^J f(i, j) + \alpha_H$$

$$V(j) = \sum_{i=1}^I f(i, j) + \alpha_V \tag{5}$$

여기서, α_H 와 α_V 는 상수로써 이 값의 의미는 III장에서 설명될 것이며, 일단 0이라고 가정하자.

2) 특징 밀도 균일화

이 방법에서 정규화 영상의 좌표 (m, n) 은 선형 형태 정규화에서 정의된 식 (3)과 같은 방법으로 계산된다.

2.2. 교차 횟수에 바탕을 둔 획 밀도를 이용한 NSN(방법 3)

이 방법은 Yamada 등^[7]에 의하여 제안되었으며, 입력 영상을 왼쪽에서 오른쪽 또는 위에서 아래로 주사하면서 흰 화소에서 검은 화소로 바뀌는 횟수인 교차 횟수로 정의되는 획 밀도를 특징으로 사용한다.

1) 특징 투영

교차 횟수에 바탕을 둔 획 밀도를 이용한 NSN 방법에서는 입력 영상을 주사하면서 흰 화소에서 검은 화소로 바뀌는 횟수를 수평축과 수직축으로 투영함으로써 다음과 같은 특징 투영 함수가 계산된다.

$$H(i) = \sum_{j=1}^J \overline{f(i, j-1)} \cdot f(i, j) + \alpha_H$$

$$V(j) = \sum_{i=1}^I \overline{f(i-1, j)} \cdot f(i, j) + \alpha_V \tag{6}$$

여기서, $f(i, 0) = 0, f(0, j) = 0$ 이며, $\overline{f(i, j)}$ 는 f

(i, j)의 논리역으로써, f(i, j)가 흰 화소일 때 1이 된다. 상수 α_H 와 α_V 는 방법 2에서와 마찬가지로 일단 0이라고 가정하자.

2) 특징 밀도 균일화

특징 밀도 균일화는 특징 투영 히스토그램의 연속 시스템으로부터 등 간격으로 표본화하는 과정으로 생각할 수 있다. 그러나, H(i)와 V(j)는 이산 함수이므로, 다음 식과 같은 방법으로 표본화하여 정규화 영상의 좌표 (m, n)을 계산한다.

$$m = \left\lfloor \frac{\sum_{k=1}^i H(k) \times \frac{M}{\sum_{k=1}^M H(k)} \right\rfloor \quad (7)$$

$$n = \left\lfloor \frac{\sum_{l=1}^j V(l) \times \frac{N}{\sum_{l=1}^N V(l)} \right\rfloor$$

여기서, m은 입력 영상의 i번째 x축 좌표에 대해 서 정규화 영상의 m번째 x축 좌표와 대응함을 의미한다. n과 j의 관계도 이와 같은 의미이다.

2.3. 획 간격에 바탕을 둔 획 밀도를 이용한 NSN (방법 4)

Tsukumo와 Tanaka [2] 에 의해 제안된 이 방법은 단순히 획의 수만을 이용하는 교차 횟수에 바탕을 둔 획 밀도를 이용한 NSN 방법 [7] 을 개선하기 위하여 획 간격(line interval)을 사용하여 획 밀도를 정의한다.

1) 특징 투영

획 간격에 바탕을 둔 획 밀도를 이용한 NSN 방법에서는 입력 영상의 점 (i, j)에서의 수평축 특징 함수 $F_H(i, j)$ 와 수직축 특징 함수 $F_V(i, j)$ 를 사용하였다. 점 (i, j)가 배경 영역의 한 점이면, 특징 함수 $F_H(i, j)$ 와 $F_V(i, j)$ 는 각각 이웃하는 두 획 사이의 간격에 대한 역수로 정의된다(그림 1).

$$F_H(i, j) = \frac{1}{h(i, j)} \quad (8)$$

$$F_V(i, j) = \frac{1}{v(i, j)}$$

여기서, h(i, j)는 점 (i, j)의 오른쪽과 왼쪽에 이웃하는 두 획 사이의 수평 간격을 나타내며, v(i, j)는 점 (i, j)의 위와 아래에 이웃하는 두 획 사이의 수직 간격을 나타낸다.

이 때, 점 (i, j)가 패턴 영역의 한 점이면, $F_H(i,$

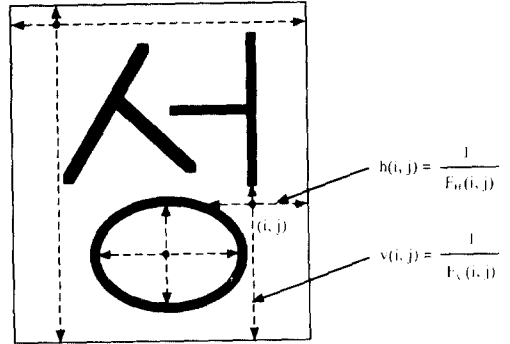


그림 1. 방법 4의 특징 함수 $F_H(i, j)$ 와 $F_V(i, j)$
Fig. 1. Feature functions, $F_H(i, j)$ and $F_V(i, j)$, for Method 4.

j)와 $F_V(i, j)$ 는 매우 작은 값을 갖게 된다.

이와 같은 특징 함수 $F_H(i, j)$ 와 $F_V(i, j)$ 를 수평축과 수직축으로 투영하는 특징 투영 함수는 다음 식과 같다.

$$H(i) = \sum_{j=1}^J F_H(i, j) \quad (9)$$

$$V(j) = \sum_{i=1}^I F_V(i, j)$$

이 방법의 정규화 목적은 획들을 정규화 영상의 패턴 영역에 균등하게 재배열하는 것이다.

2) 특징 밀도 균일화

Tsukumo와 Tanaka [2] 는 특징 투영 함수를 누적하는 것으로 정의되는 뒤틀림 함수(warping function) $\phi(i)$ 와 $\psi(j)$ 를 제안하고, 이 함수를 이용하여 정규화 영상의 좌표 (m, n)을 계산하였다.

$$m = \phi(i) = A \sum_{k=1}^i H(k) + i_0 \quad (10)$$

$$n = \psi(j) = B \sum_{l=1}^j V(l) + j_0$$

여기서, i_0 와 j_0 는 정규화 영상의 시작 위치이고, A와 B는 뒤틀림 함수 $\phi(i)$ 와 $\psi(j)$ 를 정규화하기 위한 상수이다.

이 때, i_0 와 j_0 를 각각 0이라고 하고, 상수 A와 B를 다음과 같이 정의하면, 식 (10)은 방법 1의 특징 밀도 균일화 단계에서 정의된 식 (3)과 같은 형태를 이룬다.

$$A = \frac{M}{\sum_{k=1}^M H(k)} \quad (11)$$

$$B = \frac{N}{\sum_{l=1}^N V(l)}$$

이와 같은 방법으로 방법 1과 방법 2에서 정의된 특징 밀도 균일화는 획 간격에 바탕을 둔 획 밀도를 이용한 방법 4의 특징 밀도 균일화의 특별한 경우임을 유도해 낼 수 있다.

2.4. 내접원에 바탕을 둔 획 밀도를 이용한 NSN (방법 5)

이 방법은 Yamada 등^[1]이 자신들이 제안한 방법 3^[7]을 개선한 것으로, 내접원(inscribed circle)을 사용하여 획 밀도를 정의한다. 방법 3을 개선한 것으로 앞 절에서 소개한 Tsukumo와 Tanaka^[2]의 획 간격에 바탕을 둔 획 밀도를 이용한 방법(방법 4)이 있으나, 이 방법은 방법 4와는 독립적으로 연구되어 발표되었다.

1) 특징 투영

이 방법에서는 각 점에서의 획 밀도 함수 $\rho(i, j)$ 를 정의한다. 먼저, 4 종류의 경계 E_1, E_2, E_3, E_4 를 다음과 같이 정의한다.

$$\begin{aligned} E_1 &= \max\{i' | i' \langle i, f(i', j) \cdot \overline{f(i'+1, j)} = 1 \} \\ E_2 &= \min\{i' | i' \geq i, f(i', j) \cdot \overline{f(i'+1, j)} = 1 \} \\ E_3 &= \max\{i' | i' \langle i, \overline{f(i'-1, j)} \cdot f(i', j) = 1 \} \\ E_4 &= \min\{i' | i' \geq i, \overline{f(i'-1, j)} \cdot f(i', j) = 1 \} \end{aligned} \tag{12}$$

이 때, E_1 과 E_3 는 (i, j) 의 왼쪽에 위치하며, E_2 와 E_4 는 (i, j) 의 오른쪽에 위치한다. 그리고, i' 이 정의되지 않는 경우도 존재한다.

다음으로, 점 (i, j) 에서의 수평축 획 간격 L_H 를 4 종류의 경계를 이용해서 계산한다. 그림 2는 점 (i, j) 가 배경 영역의 한 화소일 때, 수평축 획 간격 L_H 의 정의를 보여준다. 그림에서, 검은 점은 입력 영상의 한 점 (i, j) 를 나타내고, 빗금친 영역은 영상의 검은 부분을 나타내고, 여백은 흰 영역을 나타낸다. 그리고, 점선은 두 경계의 중앙을 표시한다. 점 (i, j) 가 패턴 영역의 한 화소일 경우의 수평축 획 간격도 같은 방법으로 정의된다. 그림에서, 상수 W 는 입력 영상의 폭이며, (d)부터 (i)의 경우에 $2W$ 또는 $4W$ 를 취하는 이유는 검은 획의 사이에 위치하는 점의 획 간격보다 문자 영역의 외곽에 위치하는 점의 획 간격을 크게 하기 위해서이다. 수직축으로의 획 간격 L_V 도 같은 방법으로 계산한다. 이와 같은 방법으로 계산된 수평축의 획 간격 L_H 와 수직축의 획 간격 L_V 중에서 작은 값이 내접원의 지름에 대한 근사값이라고 생각할 수 있다.

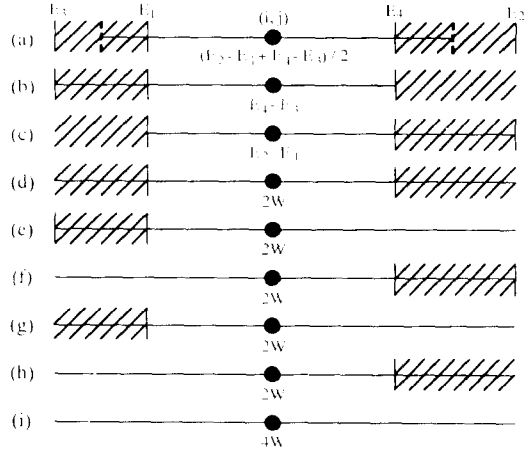


그림 2. 방법 5의 획 간격 L_H 의 정의
Fig. 2. Definition of the line interval L_H for Method 5.

다음으로, 획 밀도 함수 $\rho(i, j)$ 를 다음과 같이 정의한다.

$$\rho(i, j) = \begin{cases} \max(W/L_H, W/L_V) & \text{if } L_H + L_V < 6W \\ 0 & \text{if } L_H + L_V \geq 6W \end{cases} \tag{13}$$

각 문자를 둘러싸는 외접 사각형을 고려할 때, $L_H + L_V < 6W$ 인 경우는 외접 사각형 내에 위치하는 화소의 획 밀도를 의미하고, $L_H + L_V \geq 6W$ 인 경우는 외접 사각형의 외부에 위치하는 화소의 획 밀도를 의미한다.

이와 같은 획 밀도 함수를 수평축과 수직축으로 투영하는 특징 투영 함수는 다음 식과 같이 정의된다.

$$\begin{aligned} H(i) &= \sum_{j=1}^J \rho(i, j) \\ V(i) &= \sum_{j=1}^J \rho(i, j) \end{aligned} \tag{14}$$

이 방법은 국부적이며, 2-차원적인 특성을 갖고 있다. 왜냐하면, 획 밀도를 이웃하는 획의 분포에 따라 정의하였기 때문에 국부적인 특성이 있다고 할 수 있으며, 획 밀도 함수 $\rho(i, j)$ 를 계산하는 과정에서 수평축 획 간격과 수직축 획 간격을 하나로 결합하였으므로, 2-차원적인 특성을 가진다고 할 수 있다.

2) 특징 밀도 균일화

교차 횟수에 바탕을 둔 획 밀도를 이용한 방법 3에서 정의된 식 (7)과 같은 방법으로 정규화 영상의 좌표 (m, n) 를 계산한다.

Ⅲ. 정량적 평가

Ⅱ장에서 재검토된 기존의 비선형 형태 정규화 방법들을 PC 486(33MHz) 상에서 Turbo-C 언어를 사용하여 구현하였으며, 7mm x 8mm 사각형내에 필기된 문자를 Microtek MSF-300C 영상 스캐너를 통하여 200 DPI의 해상도로 입력받은 필기체 한글 데이터를 테스트에 사용하였다. 테스트 대상 문자로는 한글 기계화 연구소에서 조사한 한글 사용 빈도순 상위 520자^[6]를 선택하였으며, 10명의 필기자가 한 문자에 대하여 각각 10번씩 필기함으로써, 총 52,000 자의 필기체 한글 데이터를 사용하였다.

각 방법에 대한 객관적인 평가를 위해서는 크게 다음의 두가지 사항에 유의하여야 한다. 첫째로, 각 방법이 기술된 논문을 바탕으로 하여 그 방법을 구현할 때, 자칫하면 구현 대상 방법의 중요한 아이디어나 매개 변수를 놓치는 경우가 발생할 수 있으므로, 이러한 실수를 하지 않으려면 최소한 그 논문에서 예시된 데이터를 사용하여 구현한 프로그램을 실행한 다음, 그 결과가 논문에서 제시된 결과와 일치하는지를 확인하여야 한다. 둘째로, 각 방법이 동일한 입력력 구조 및 자료 구조를 사용함은 물론, 동일한 프로그래밍 언어로 구현되어야 한다.

각 방법의 차이를 쉽게 비교할 수 있도록 각 방법을 특징 투영과 특징 밀도 균일화로 나누어 표 1에 나타내었으며, 한자 "星"에 대한 각 방법의 결과를 그림 3에 나타내었다. 그림 3에서 각 NSN 방법의 특징 밀도 균일화 결과가 항상 균일하지는 않음을 볼 수 있다. 그 이유는 각 방법에서 히스토그램 균일화가 공간적으로 수행되었기 때문이다.^[13]

표 1. 형태 정규화 방법들의 요약
Table 1. Summary of different shape normalization methods

단계 방법	특징 투영		특징 밀도 균일화	
	H(i)	V(i)	m	n
방법 1	1	1	$\frac{1}{M} \times \sum_{i=1}^M$	$\frac{1}{N} \times \sum_{i=1}^N$
방법 2	$\sum_{i=1}^M (h_i) \times \alpha_i$	$\sum_{i=1}^N (v_i) \times \alpha_i$	$\sum_{i=1}^M (h_i) \times \frac{M}{\sum_{i=1}^M h_i}$	$\sum_{i=1}^N (v_i) \times \frac{N}{\sum_{i=1}^N v_i}$
방법 3	$\sum_{i=1}^M (h_i) \times (h_i) \times \alpha_i$	$\sum_{i=1}^N (v_i) \times (v_i) \times \alpha_i$	$\left[\sum_{i=1}^M (h_i) \times \frac{M}{\sum_{i=1}^M h_i} \right]$	$\left[\sum_{i=1}^N (v_i) \times \frac{N}{\sum_{i=1}^N v_i} \right]$
방법 4	$\sum_{i=1}^M f_i(h_i)$	$\sum_{i=1}^N f_i(v_i)$	$\psi(h) = A \times \sum_{i=1}^M h_i \times \dots$	$\phi(v) = B \times \sum_{i=1}^N v_i \times \dots$
방법 5	$\sum_{i=1}^M (h_i)$	$\sum_{i=1}^N (v_i)$	$\left[\sum_{i=1}^M (h_i) \times \frac{M}{\sum_{i=1}^M h_i} \right]$	$\left[\sum_{i=1}^N (v_i) \times \frac{N}{\sum_{i=1}^N v_i} \right]$

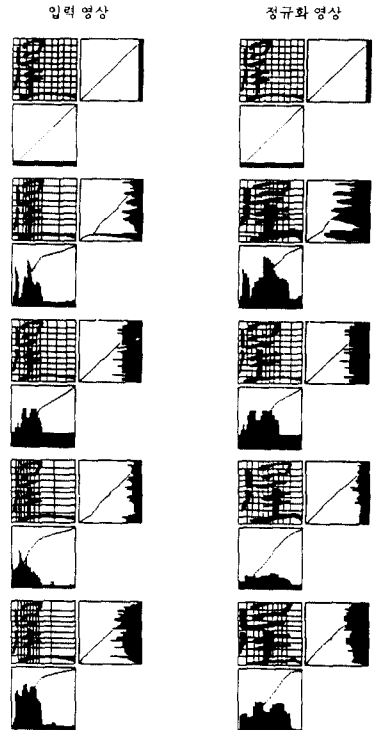


그림 3. 한자 "星"에 대한 특징 투영과 특징 밀도 균일화의 예

Fig. 3. Examples of feature projection and feature density equalization of Chinese character "星".

각 방법의 성능 및 장단점을 체계적으로 비교하기 위하여 인식률, 처리 속도, 특징 투영 단계에서의 계산 복잡도, 정규화된 영상의 변형량을 평가 기준으로 설정하고, 이 기준을 바탕으로 각 방법을 평가한다.

1. 평가 기준

본 절에서는 각 알고리즘의 성능 및 장단점을 정량적으로 평가하기 위한 기준에 대하여 자세히 설명한다.

■ 인식률

비선형 형태 정규화의 목적은 필기체 문자 데이터의 품질을 개선함으로써^[9] 결과적으로는 인식 성능을 향상시키기 위한 것이므로 인식률을 첫번째 평가 기준으로 설정하였다. 이 때, 필기체 문자 데이터의 반은 학습 데이터로 사용하였으며 나머지는 테스트 데이터로 사용하였다.

인식 실험은 단순한 상관 정합 방법(단순 정합)과 보다 복잡한 상관 정합 방법(동적 정합)의 두 종류를 수행하였다. 첫번째 방법에서는 변형이 적은 인쇄체 문자 인식에 흔히 사용되는 방법으로 각 방법에 의해 정규화된 영상으로부터 그물 특징 벡터를 추출한 다음 단순한 최소 거리 분류기를 사용하여 입력 영상을 분류하였다. 이와 같은 인식 방법을 선택한 이유는 실험에 의해 얻어진 인식 결과가 형태 정규화의 성능에 영향을 받는 것이 아니라, 인식 방법의 성능에 영향을 받는 현상을 초래하지 않기 위한 것으로 이 인식 실험에 의해 얻어진 인식 성능이 우수하다면 해당 비선형 형태 정규화 방법이 변형이 심한 필기체 문자 데이터 집합을 변형이 적은 문자 데이터 집합으로 변환할 수 있음을 시사한다.

두번째 인식 실험은 보다 성능이 우수한 인식 방법을 사용하면 인식률이 향상됨을 보이기 위한 것으로 이 실험에 의해 얻어진 인식 결과는 인식 방법의 성능에 다소 영향을 받을 수 있다. 이 인식 방법을 간단히 설명하면 다음과 같다. 대부분이 직선 성분으로 구성된 한글의 특성을 이용하기 위하여 각 방법에 의해 정규화된 영상으로부터 수평, 수직, 사선, 역사선의 방향 성분을 추출하며, 획 간의 접촉 변형에 무관한 특징 벡터를 추출하기 위하여, 검은 화소의 밀도에 따라 동적으로 그물을 결정하였다. 다음으로, 추출된 그물 특징 벡터에 가중치 마스크를 사용하여 흐름 단계를 적용함으로써, 획 간의 위치 변형을 흡수하고자 하였다.^[10] 분류기로는 첫번째 실험과 마찬가지로 최소 거리 분류기를 사용하였다. 이 인식 방법은 특징 추출 단계에서 점 밀도에 따라 동적으로 그물을 결정하기 때문에 점 밀도를 이용한 비선형 형태 정규화 과정을 포함한다는 것을 유의하기 바란다.

■ 처리 속도

실용적인 실시간 필기체 문자 인식 시스템에서 고려해야 할 중요한 성능으로써 처리 속도를 들 수 있다. 그러므로, 각 방법을 사용하여 비선형 형태 정규화를 수행하는데 소요되는 문자당 평균 CPU 시간을 측정하여 비교하였다.

■ 계산 복잡도

처리 속도는 사용중인 시스템의 성능에 다소 영향을 받는 기준이지만 알고리즘의 계산 복잡도는 시스템의 성능에 영향을 받지 않는 요소로써 서로 다른 방법들의 복잡도를 분석적으로 비교할 수 있는 유용한 평가 기준이다. 표 1을 보면 각 방법의 특징 밀도 균일화 단계는 계산 방법에만 차이가 있을 뿐 계산 복잡도에는 차이가 없다는 것을 쉽게 알 수 있으므로, 특징 투영 단계에서 필요한 계산 복잡도만을 비

교 대상으로 고려하였다.

계산 복잡도를 구하는 방법을 예를 들어 설명하면 다음과 같다. $I \times J$ 크기의 영상 $f(i, j)$ 를 교차 횟수에 바탕을 둔 획 밀도를 이용한 NSN 방법(방법 3)으로 정규화한다고 가정하자. 그러면, 식 (6)으로 표현된 특징 투영 함수 $H(i)$ 와 $V(j)$ 를 계산하기 위해서는 각 점 (i, j) 에서 4번의 비교가 필요하다. 즉, 함수 $H(i)$ 를 계산하기 위해 $f(i, j-1)$ 이 흰 점인가와 $f(i, j)$ 가 검은 점인가를 비교해야 하고, 마찬가지로 $V(j)$ 를 계산하기 위해서는 $f(i-1, j)$ 와 $f(i, j)$ 에 대해 2번의 비교가 필요하다. 이런 방법으로 $I \times J$ 크기의 영상에 대해 총 $4 \times (I \times J)$ 번의 비교가 필요하며, 이것이 계산 복잡도가 된다.

다른 방법들의 특징 투영 단계에 필요한 계산 복잡도도 이와 유사한 방법으로 구한다.

■ 변형량

변형량은 많은 문자 데이터 집합이 있을 때, 그 문자 집합이 얼마나 많은 변형을 포함하고 있는지를 평가하기 위한 객관적이고 정량적인 척도이다. 필기체 문자에서 발생하는 변형을 보상하기 위해 형태 정규화를 수행하는 것이므로, 우수한 형태 정규화 알고리즘에 의해 정규화된 문자 데이터 집합의 변형량은 작아져야 할 것이다. 변형량은 평균 엔트로피(average entropy), 분포 엔트로피(distribution entropy), 분산 계수(dispersion factor), 평균 엔트로피 변화량(average entropy difference) 등 다양하지만, 본 논문에서는 우수한 변형량이 만족해야 할 조건을 만족한다고 발표된 바 있는 평균 엔트로피 변화량^[11]만을 고려하였다.

평균 엔트로피 변화량은 다음 과정을 통해 계산된다. $I \times J$ 크기의 이진 영상 $f(i, j)$ 에 대해, 점 (i, j) 의 엔트로피는 다음과 같이 정의된다.

$$h(i, j) = -p(i, j) \times \log_2 p(i, j) - q(i, j) \times \log_2 q(i, j) \quad (15)$$

단, $q(i, j) = 1 - p(i, j)$

이 때, $p(i, j)$ 는 점 (i, j) 가 검은 화소일 확률로써 다음과 같다.

$$p(i, j) = \frac{\sum_{k=1}^K f^k(i, j)}{K}, \quad 0 \leq p(i, j) \leq 1 \quad (16)$$

여기서, K 는 같은 부류에 속한 문자 영상의 수이며, $f^k(i, j)$ 는 k -번째 영상의 점 (i, j) 를 나타낸다.

다음으로, 점 (i, j) 의 엔트로피 변화량은 다음과 같이 계산한다.

$$v(i, j) = \frac{h(i, j)}{\omega \times \max\{|h(i+1, j) - h(i, j)|, |h(i, j+1) - h(i, j)|\} + 1} \quad (17)$$

여기서, 상수 ω 는 $v(i, j)$ 의 범위를 결정하는 확장 계수이다.

마지막으로, 평균 엔트로피 변화량은 다음과 같다.

$$v_D = \frac{1}{|C|} \sum_{i=1}^L \sum_{j=1}^L v(i, j), \quad |C| \neq 0 \quad (18)$$

여기서, $C = \{(i, j) \mid h(i, j) \neq 0\}$ 이고, $|C|$ 는 집합 C 에 속하는 원소의 수이다.

2. 결과 및 고찰

본 절에서는 상수 α_H 와 α_V 의 의미를 자세히 설명하고, 각 방법의 특성 및 장단점을 논의한다.

2.1. 상수 α_H 와 α_V 의 의미

Ⅱ장에서는 상수 α_H 와 α_V 의 값을 0으로 가정하였다. 하지만, $\alpha_H = 0$ 인 상태에서 점 밀도를 이용한 NSN 방법을 사용하여 문자 "이"를 수평축으로 정규화할 경우, 문자 "이"의 가운데 공백 부분은 투영된 특징 함수의 값이 0이 되므로, 정규화된 영상에서는 가운데 공백 간격이 없어져 접촉하는 현상이 발생한다. 이와 같은 비선형적인 정규화는 너무 강하고 자연스럽지 못한 제약이다.

상수 α_H 는 이런 효과를 줄이기 위해 적용되는 선형 가중치로써, 문자를 둘러싸는 외접 사각형의 외부에서는 α_H 의 값을 0으로 하고, 내부는 양의 상수로 하면, 문자 영역 내부의 공백 영역은 점 밀도가 1인 경우에 비해 $\alpha_H / (\alpha_H + 1)$ 의 비율로 정규화되며, 문자 영역 외부의 공백 영역은 무시된다. 수직축 정규화에 대한 상수 α_V 도 α_H 와 같은 의미를 갖는다. 이때, 상수 α_H 와 α_V 의 값을 매우 크게 하면 투영된 특징의 효과가 줄어들기 때문에, 직사각형을 정사각형으로 정규화하는 선형 정규화와 같게 된다.

표 2. 상수 α_H 와 α_V 의 효과

Table 2. Effects of constant α_H and α_V .

$\alpha_H (= \alpha_V)$ 방법	방법 2	방법 3
0	93.27%	90.77%
1	95.96%	91.54%
2	96.35%	94.62%
3	96.15%	93.85%
10	92.88%	88.46%
$\infty (= LSN)$	85.26%	84.92%

상수 α_H 와 α_V 의 값에 따른 인식 성능의 결과를 살펴 보기 위해 단순한 상관 정합 방법(단순 정합)을 사용하여 인식 실험을 수행하였으며 결과를 표 2에 나타내었다. 이 때 상위 5등 까지의 누적 인식률을 비교하였다. 실험 결과, 상수 α_H 와 α_V 의 값에 따라 성능의 차이가 있으므로, 이 값을 결정하는 것 또한 매우 중요한 문제임을 알 수 있다.

2.2. 고찰

그림 4는 다양한 필기체 한글 데이터 "를"에 대해서 각 방법을 사용하여 형태 정규화한 예를 보여준다. 이 결과를 바탕으로 주관적으로 평가한 각 방법의 특성 및 장단점을 살펴 보자.

입력 영상을 보면, 필기자의 습관에 따라 다양하게 획을 길게 늘어 쓰는 경향이 있음을 알 수 있다. 이러한 경향에 의해 발생하는 불규칙적인 형태 왜곡은 선형 형태 정규화(방법 1)로는 해결할 수 없는 변형이다.

점 밀도를 이용한 방법 2는 긴 획을 포함한 부분은 점 밀도가 높아지는 특성에 따라 긴 획을 포함하는 부분 영역은 확대되고, 그렇지 않은 부분 영역은 축소된다. 이 방법에 의해 문자 형태의 불균형을 어느 정도는 수정할 수 있는 장점이 있지만, 획의 두께가 획의 길이에 민감하게 변하는 단점이 있다.

유사한 방법으로, 방법 3에 의해 교차 횟수가 큰 즉, 획 밀도가 높은 부분 영역은 확대되고 획 밀도가 낮은 부분 영역은 축소된다. 이 방법에 의해 정규화된 영상은 획의 두께가 거의 일정하며, 문자의 형태도 대략적으로 균형을 이루지만 내부 획들의 위치 정규화에 있어서는 여전히 문제점을 안고 있다.

방법 4는 획 사이의 간격이 좁은 부분 영역은 확대되고, 그렇지 않은 부분 영역은 축소되는 특성이 있다. 이 방법의 장점은 정규화된 영상의 획의 두께가 거의 일정하며, 문자의 형태 불균형을 수정하고 획 사이의 간격을 균등하게 정규화한다는 것이다. 그러나, 이 방법은 문자 영역의 외부에 약간의 돌출 부분이 있는 경우에 이 돌출 부분 영역이 확대되므로 이런 효과를 없애기 위해서 문자 영역의 외부에 부가적인 공백 영역을 추가하여 처리하였다.

내접원에 바탕을 둔 방법 5는 문자의 형태가 복잡한 부분 영역은 확대되고 그렇지 않은 부분 영역은 축소되는 특성이 있으며, 사람이 습관적으로 마지막 획의 오른쪽 아랫 부분을 길게 늘어 쓰는 경향에 의해 발생하는 획을 축소하는데 유용하다. 또한 이 방법으로 정규화된 영상은 획의 두께가 거의 일정하고, 문자의 형태도 균형을 이룬다. 그러나, 이 방법은 다른 방법에 비해 매우 많은 계산량을 필요로 하므로

실시간 문자 인식에 사용되기에는 한계가 있다.

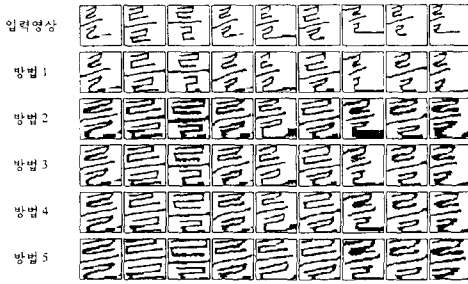


그림 4. 한글 “를”에 대한 형태 정규화의 예
Fig. 4. Examples of shape normalization for Hangul syllable “를”.

각 방법의 개관적인 평가를 위하여 다양한 평가 기준을 바탕으로 하여 각 방법을 정량적으로 평가한 결과(표 3), 전반적으로 평가해 볼 때, 비선형 형태 정규화 방법들이 선형 형태 정규화 방법보다는 다소 복잡하고 처리 속도도 늦지만 인식 성능을 대폭 향상시키고 다양한 형태 왜곡을 포함한 필기체 문자 데이터 집합의 변형량도 감소시킴을 알 수 있다.

다음으로 기존의 비선형 형태 정규화 방법들의 성능을 각 기준별로 살펴 보면 다음과 같다. 먼저, 인식률의 측면에서 본다면, 계산 과정보다 복잡하긴한데 내접원에 바탕을 둔 획 밀도를 이용한 방법 5가 인식 성능이 가장 우수하며, 처리 속도 및 계산 복잡도의 측면에서 본다면, 점 밀도를 이용한 방법 2가 비교적 우수함을 알 수 있다. 변형량을 기준으로 본다면, 직간격에 바탕을 둔 획 밀도를 이용한 방법 4와 내접원에 바탕을 둔 획 밀도를 이용한 방법 5가 많은 변형을 포함한 필기체 문자 집합의 변형량을 줄이는데 유용함을 알 수 있다. 참고로 점 밀도를 이용한 방법 2가 빠르고 계산 복잡도가 적은 것에 반해 비교적 우수한 인식 성능을 보였는데, 이것은 인식 실험에 적

표 3. 정량적 평가 결과
Table 3. Result of quantitative evaluation.

기준 방법	누적인식률(5등까지)		처리속도 (초/문자)	계산 복잡도 (특징 추출 단계)	변형량
	단순점합	동적점합			
방법 1	84.67%	95.22%	0.005	1+J	0.27445
방법 2	96.35%	98.17%	0.010	2(LJ)	0.25970
방법 3	94.62%	97.31%	0.012	4(LJ)	0.26259
방법 4	96.73%	97.50%	0.162	(1+J+2)(LJ)	0.24517
방법 5	96.92%	98.27%	0.757	(2x+2xJ+3)(LJ)	0.24536

용한 두가지 인식 방법 모두 각 그물 영역별로 검은 화소의 수를 특징으로 사용하기 때문인 것으로 판단된다.

IV. 결론

본 논문에서는 지금까지 발표된 여러 종류의 형태 정규화 방법들을 특징 투영과 특징 밀도 균일화의 두가지 관점에서 재검토한 다음, 각 방법에 대한 정량적 평가 결과를 제시하였다. 다양한 필기체 문자 데이터에 대한 테스트를 통하여 선형 형태 정규화 방법에 비해 비선형 형태 정규화 방법이 변형이 심한 필기체 문자 데이터의 품질을 개선하는데 매우 효율적임을 확인할 수 있었다. 그러나, 인식률과 처리 속도 등 다양한 측면에서 우수한 성능을 갖는 비선형 형태 정규화 방법은 찾아볼 수 없었다. 따라서, 비선형 형태 정규화 방법을 사용하여 필기체 문자 인식을 시도할 경우에는 각 방법의 장단점을 충분히 파악하여 각자의 목적에 가장 적합한 방법을 선택하는 것이 효과적일 것으로 판단된다.

추후, 본 연구 결과에 바탕을 둔 효과적인 비선형 형태 정규화 알고리즘의 개발에 관한 연구가 필요하며, 궁극적으로 필기체 문자에서 발생하는 형태 왜곡에 무관하게 효과적으로 입력 패턴을 인식할 수 있는 비선형 패턴 정합^[3]에 관한 연구도 필요하리라 사료된다. 또한 3차원 물체의 표면 위에 쓰여진 영상을 카메라로 입력하거나 또는 평면 위에 쓰여진 문자라고 하더라도 카메라의 위치에 따라 다양한 형태로 발생하는 변형을 bilinear, quadratic, cubic, Coons, harmonic 등의 수학적인 모델을 사용하여 근사시킴으로써, 원래 형태로 복원할 수 있는 비선형 영상 복원^[2]에 관한 연구도 문자 인식 기술의 실용화를 위해서는 절실한 실정이다.

參考文獻

[1] H. Yamada, K. Yamamoto and T. Saito, "A Nonlinear Normalization for Handprinted Kanji Character Recognition - Line Density Equalization," Pattern Recognition, vol. 23, no. 9, 1990, pp. 1023-1029.
[2] J. Tsukumo and H. Tanaka, "Classification of Handprinted Chinese Characters Using Nonlinear Normalization Methods," Proc. of 9th Int. Conf.

- on Pattern Recognition, Rome, Italy, November 1988, pp. 168-171.
- [3] J. Tsukumo, "Handprinted Kanji Character Recognition Based on Flexible Template Matching," Proc. of 11th Int. Conf. on Pattern Recognition, The Hague, The Netherlands, September 1992, pp. 483-486.
- [4] Y. Yamashita, K. Higuchi, Y. Yamada and Y. Haga, "Classification of Handprinted Kanji Characters by the Structured Segment Matching Method," Pattern Recognition Letters, vol. 1, no. 5, 1983, pp. 475-479.
- [5] R. G. Casey, "Moment Normalization of Handprinted Characters," IBM J. Research and Development, vol. 14, September 1970, pp. 548-557.
- [6] Z. C. Li, T. D. Bui, Y. Y. Tang and C. Y. Suen, *Computer Transformation of Digital Images and Patterns*, World Scientific Press, 1989.
- [7] H. Yamada, T. Saito and K. Yamamoto, "Line Density Equalization - A Nonlinear Normalization for Correlation Method," Trans. IECE of Japan, vol. J67-D, no. 11, November 1984, pp. 1379-1383(in Japanese).
- [8] 한글 기계화 연구소, 한글 기계화 연구, 1975.
- [9] A. Gudesen, "Quantitative Analysis of Preprocessing Techniques for the Recognition of Handprinted Characters," Pattern Recognition, vol. 8, no. 4, 1976, pp. 219-227.
- [10] 이 성환, 박 정선, "샘물체 필기 한글의 오프라인 인식," 한국정보과학회 논문지, 제 20권 제 10호, 1993년 10월(게재 예정).
- [11] 김 대환, "필기체 문자 영상 데이터의 변형량", 석사 학위 논문, 포항공과대학 전자계산학과, 1993년 2월.
- [12] 김 은순, 이 성환, "Coons 모델을 이용한 비선형 변환된 문서 영상의 효율적인 복원," 제 1 회 문자 인식 워크샵 발표 논문집, 청주, 1993년 5월, pp. 81-88

著 者 紹 介



李 晟 煥(正會員)

1962年 6月 2日生. 1984年 서울대학교 계산통계학과 졸업. 1986年 한국과학기술원 전산학과(석사). 1989年 한국과학기술원 전산학과(박사). 1987年 네덜란드 Delft 공과대학교 패턴인식 연구센터 방문 연구원. 1989年 캐나다 Concordia 대학교 패턴인식 및 기계지능 연구센터 방문 과학자. 1989年 ~ 현재 충북대학교 컴퓨터과학과 조교수로 근무 중. 주관심 분야는 패턴 인식, 컴퓨터 시각, 신경망 등임.



朴 貞 善(準會員)

1968年 12月 30日生. 1992年 충북대학교 전자계산학과 졸업. 1993年 현재 충북대학교 컴퓨터과학과 석사과정 재학중. 주 관심 분야는 필기체 문자 인식, 패턴 정합, 비선형 형태 정규화 등임.