

# 원형 마스크 팽창법에 의한 문자인식

박영석\* 최철용\*\*

경남대학교 전자공학과

## The Character Classifier Using Circular Mask Dilation Method

Young Seak Park\* Cheol Yong Choi\*

Dept. of Electronic Eng., Kyungnam Univ.

\*yspark@hanma.kyungnam.ac.kr, yspark@iname.com

### <ABSTRACT>

In this paper, to provide the robustness of character recognition, we propose a recognition method using the dilated boundary curve feature which has the invariance characteristics for the shift, scale, and rotation changes of character pattern. And its some characteristics and effectiveness are evaluated through the experiments for both the english alphabets and the numeral digits. The feature vector is represented by the Fourier descriptor for a boundary curve of the dilated character pattern which is generated by the circular mask dilation method, and is used for a nearest neighbor classifier(NNC) or a nearest neighbor mean classifier(NNMC). These classifiers provide a satisfiable result in view point of the processing time and the recognition rate, and take also the robustness of recognition for both some internal noise and partial corruption of an image pattern.

### 1. 서론

문자인식은 문자패턴의 특징을 추출하여 특징벡터를 생성하고 다양한 분류기(classifier)를 통해 대응하는 부류(class)로 식별하는 과정이다. 이러한 인식기의 구현에서 문자의 특징 추출은 인식기의 성능을 좌우하는 기본 요소이며, 입력된 동일 문자에 대해서는 유사한 특징 벡터를 얻을 수 있어야 한다.

최근 패턴인식에 관한 꾸준한 연구성과에 힘입어 제한된 영역에서의 패턴인식 시스템의 성능은 매우 높은 수준에 이르렀음에도 불구하고 패턴의 위치가 이동하거나

크기가 변하거나 또는 기울어짐으로 인하여 패턴의 형태가 변형되었을 때에는 인식 시스템의 성능이 급격히 저하됨은 주지의 사실이다. 위치, 크기, 회전 변형에 무관한 패턴인식은 특히 도형이 혼용된 일반적인 분식의 인식에 유용한데, 이러한 영상변형에 무관한 패턴인식에 관해서는 최근 들어 제한적이나마 개선책들이 발표되고 있다.

본 논문에서는 원형 마스크 팽창법(circular mask dilation method)<sup>1)</sup>을 사용함으로써 문자 패턴의 이동, 크기, 그리고 회전 변형에 불변특성을 지니는 문자 인식 방법을 제안하고 그 유효성을 영문 알파벳과 숫자에 대한 인식 실험을 통해서 평가한다.

## 2. 문자의 팽창

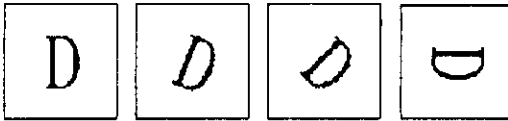
### 2.1 회전 불변 팽창법

팽창 윤곽선에 의한 인식의 착안점은 문자 영상의 내부의 정보량이 문자의 가장자리 부분에 존재한다는 사실에 근거한다<sup>2)</sup>. 우선 문자의 팽창을 통한 외각 형상을 구하고 외각 윤곽선의 변화를 편각함수로 변환한다. 이 편각함수는 주기함수가 되며 이것을 푸리에 급수로 전개하여 푸리에 기술자(Fourier descriptor)<sup>2)</sup> 즉, 계수를 구한다. 이 푸리에 기술자는 문자영상의 회전, 크기, 그리고 이동에 무관한 불변특성을 제공하며 더구나 문자의 팽창을 통해 영상의 열화 혹은 잡음의 영향도 충분히 무시할 수 있는 특성을 얻을 수 있다. 이 때 회전된 문자에 대해서 4 방향 팽창(4-neighbor dilation)이나 8 방향 팽창(8-neighbor dilation)과 같은 기존의 2진 팽창법(binary dilation)<sup>3)</sup>을 적용하면 회전 각도에 따라 추출된 패턴의 윤곽 형상이 매우 달라져 상이한 특징벡터를 얻게 되므로 문자의 회전에 독립적인 원형 마스크 팽창

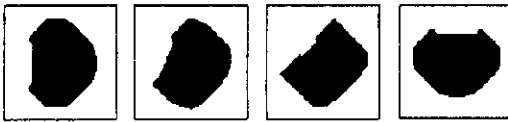
이 제안되었다<sup>1)</sup>.

### 2.1 원형 마스크 팽창

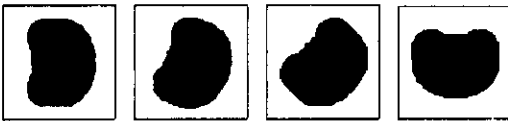
원형 마스크 팽창은 패턴을 임의의 팽창비율에 따라 정해진 원형 마스크 즉, 구조화 요소를 적용하는 것으로, 이 때 원형 마스크 패턴은 원의 공식을 양자화 된 디지털 화상에 대하여 적용하고 팽창처리속도의 개선을 위하여 4 연결 윤곽 화소만을 추출하여 얻는다<sup>3)</sup>. 그림 2-1은 4방향 팽창의 회전 의존성과 원형 마스크 팽창의 회전 독립성을 보여 주고 있으며, 단일 처리 알고리즘이므로 기존의 방법에 비해 처리속도도 월등히 뛰어나다<sup>1)</sup>.



(a) 문자 "D"의 0°, 22.5°, 45°, 90° 회전된 이미지



(b) 4방향 팽창의 결과



(c) 원형 마스크 팽창의 결과  
그림 2-1 팽창 방법에 따른 결과

### 3. 특징벡터와 문자인식

팽창되어진 문자의 윤곽선은 하나의 폐곡선을 이루므로 그 곡선을 표현하는 하나의 주기 함수를 구할 수 있고, 그것을 푸리에 급수를 사용하여 푸리에 기술자로 나타내어 그것을 문자의 특징벡터로 사용할 수 있다.

#### 3.1 주기 편각 함수

$\gamma$  를  $(X(t), Y(t)) = Z(t)$ 의 파라메트릭 표현을 가지는 시계방향 단순 폐곡선(simple closed planar curve)이라 가정하자. 여기서  $l$ 은 폐곡선 원호(arc)의 길이 그리고  $L$ 은 총길이이며,  $0 \leq l \leq L$ 이다. 그림 3-1에서처럼 폐곡선 상의 한점  $X(0), Y(0)$ 를 시점으로 하고 곡선 위를  $l$ 만큼 나아간 점에서의 접선의 편각(angular direction)을 함수  $\theta(l)$ 로 나타낸다. 여기서 시점  $l=0$ 과 한점  $l$ 간의 편각의 총량(net amount) 즉, 편각의 차를 누적 편각 함수(cumulative angular direction function)  $\phi(l)$ 이라 정의하면 정규화된 하나의 누적편각함수  $\phi^*$ 를 다음처럼 정의할 수 있다.

$$\phi^*(l) = \phi\left(\frac{Ll}{2\pi}\right) + l$$

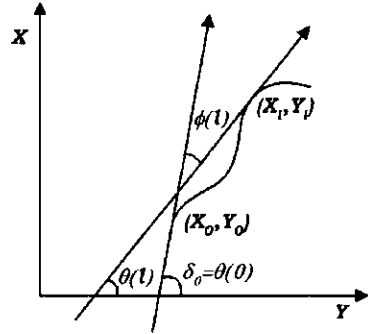


그림 3-1 편각 함수

$\phi^*(l)$ 는 이동, 회전, 그리고 길이  $L$ 의 변화(혹은 크기)에 불변인 주기함수이다. 그래서  $\phi^*$ 를 푸리에 급수로 전개하면 다음과 같다.

$$\begin{aligned} \phi^*(l) &= \mu_0 + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt) \\ &= \mu_0 + \sum_{n=1}^{\infty} (A_n \cos(nt - \alpha_n)) \end{aligned}$$

여기서,  $(A_n, \alpha_n)$ 은  $(a_n, b_n)$ 의 극좌표이고 성분 각각은  $n$ 차 조화진폭( $n$ th harmonic amplitude) 그리고 위상각으로 알려진 폐곡선  $\gamma$ 의 푸리에 기술자이다.

동일 문자에 대한 팽창 윤곽선의 푸리에 기술자는 문자영상의 미소한 변형에 따라 저주파 성분에서는 거의 변화가 없거나 고주파 성분에서는 상당한 차이를 초래할 수 있다. 이것은 결국 분류기에서 오분류의 가능성을 크게 한다. 따라서 푸리에 기술자의 차수를 적정한 선으로 한정하여 특징벡터로 할 필요가 있다.

#### 3.2 문자의 인식

각각의 문자에서 추출된 특징벡터는 사실상 다소 차이가 존재한다. 따라서 입력된 패턴을 문자의 표준 특징벡터와 비교해서 가장 유사한 부류로 분류 즉, 인식하여야 한다.

한 표준패턴의 특징벡터  $Z = \{z_1, z_2, \dots, z_N\}$ 와 임의의 입력패턴 특징벡터  $X = \{x_1, x_2, \dots, x_N\}$  간의 유클리드 거리(Euclidean distance)는 다음과 같다<sup>3)</sup>.

$$D = \left[ \sum_{i=1}^N (x_i - z_i)^2 \right]^{1/2} = \sum_{i=1}^N |x_i - z_i|$$

최소 거리 분류기(minimum distance classifier: MDC)<sup>3)</sup>는 미지의 문자 패턴 특징  $X$ 와 각 문자 표준 특징  $Z$  간의 거리를 계산하여 가장 가까운 거리를 갖는 문자로 패턴  $X$ 를 분류한다. 같은 거리를 갖는 부류가 여러 개일 경우는 임의로 배정한다.

MDC를 보다 일반화한 것이 최근접 이웃 분류기(nearest neighbor classifier: NNC)<sup>3)</sup>이다. 주어진 한 부류의 표준패턴 집합  $S = \{s_1, s_2, \dots, s_L\}$ 을 고려해보자. 여

기서, S의 모든 원소는 가능한 부류집합  $w = (\omega_1, \omega_2, \dots, \omega_M)$  중의 단지 한 부류  $\omega_j$ 에 속한다고 가정한다.  $s_i \in S$ 가 S의 모든 원소중 최소거리를 가진다면, 거리 개념을 확장하여 근접거리는 다음과 같다.

$$D(s_i, x) = \min_l \{D(s_i, x_l)\}, \quad l = 1, 2, \dots, L$$

NNC는 미지의 패턴 특징 X와 각 문자의 표본특징 집합 S 간의 근접거리를 구하고 그들 중 최소의 값을 가지는 부류  $\omega_j$ 로 패턴 X를 분류한다. 즉, X와  $\omega_j$ 간의 근접거리가 최근접 거리가 된다.

그러나 NNC는 한 부류에 대한 표본 패턴 집합 S의 원소수가 L이면 문자인식 처리시간이 MDC의 L배가 소요되는 단점이 있다. 이것을 극복하기 위해  $S = \{s_1, s_2, \dots, s_L\}$ 의 모든 원소의 평균값으로 이뤄지는 대표 표본특징벡터들의 집합  $Z = \{\hat{z}_1, \hat{z}_2, \dots, \hat{z}_M\}$ 와 임의의 입력 특징벡터 X에 대해 MDC를 적용하는 최근접 이웃 평균 분류기(nearest neighbor mean classifier: NNMC)라 정의한다. NNMC는 NNC와 유사한 인식특성을 지니면서 처리시간은 MDC와 같다.

팽창 문자에 대한 편각함수 푸리에 기술자를 이용하는 NNC 및 NNMC는 일반적인 문자인식 기법에서 요구되는 대부분의 불변특성(invariance characteristics)를 지닌다. 문자 영상의 잡음이나 열화 특성은 문자팽창처리를 통한 윤곽 폐곡선 특징에 의해 완화되고 소멸되어 문자 인식에 영향을 미치지 못하게 된다. 문자의 이동 특성은 문자의 추출과정을 통해 무시된다. 그리고 회전 및 크기의 특성은 편각함수의 푸리에 기술자의 불변특성에 의해 특징화된다. 결과적으로 NNC 및 NNMC에 의한 문자 인식 처리는 문자 영상의 잡음, 열화, 이동, 회전, 그리고 크기에 불변인 특성을 가진다고 할 수 있다.

## 4. 실험 및 고찰

### 4.1 실험과정

본 논문의 실험은 16M 메모리의 Pentium MMX 200MHz 시스템 환경하에서 Matlab을 사용하여 수행하였으며, 실험에 사용된 데이터는 명조체의 폰트 크기 24, 36, 48의 영어 대문자와 숫자들로써 회전에 대한 인식을 비교하기 위해서 각각 0, 22, 45, 90도로 회전시켜 총 432개의 문자 표본을 사용하였다.(그림.2-1(a)참조)

이러한 인식 실험에서 인식률을 개선시킬 수 있는 다양한 요인들로 푸리에 기술자의 차수, 팽창 비율 및 분류기의 선택에 따른 인식률을 조사하였다. 또한 팽창 윤곽선 방법의 단점인 팽창시 유사 윤곽선을 가지는 '0'과 'O', '6'과 '9', 'A'와 'V', 'H'와 'M'을 제외한 인식률을 조사한다.

### 4.2 실험결과

#### 4.2.1 특정 기술자의 차수 의존 인식률

그림 4-1은 기준 폰트 크기 36의 정상 상태 문자들을 1/4 비율의 원형 마스크(9 bit 팽창 마스크)로 팽창시킨 후 특징 벡터를 추출하여 기준 벡터로 사용했을 때 일정 계수 이상의 주파수 값들을 제외시킴에 따른 인식률의 변화를 보이고 있다. 차수의 1차에서 5차까지의 절단은 인식률의 급격한 증가를 수반하지만, 그 이후로는 완만해지고 10차에서 최고의 인식률 79.86%을 보이며, 그 이상의 고차항이 포함됨에 따라 인식률은 오히려 떨어짐을 보이고 있다. 따라서, 인식에 적절한 계수는 10차 전후임을 알 수 있다.

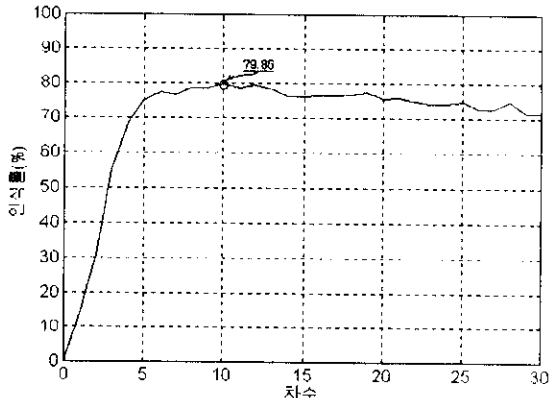


그림 4-1 특정 기술자의 차수 의존 인식률(1/4 팽창, MDC)

#### 4.2.2 팽창비율과 분류기 의존 인식률

팽창 윤곽선 방법에 있어 문자의 크기에 대한 어느 정도 비율로 몇 bit 씩 팽창할 것인지 결정할 필요가 있다.

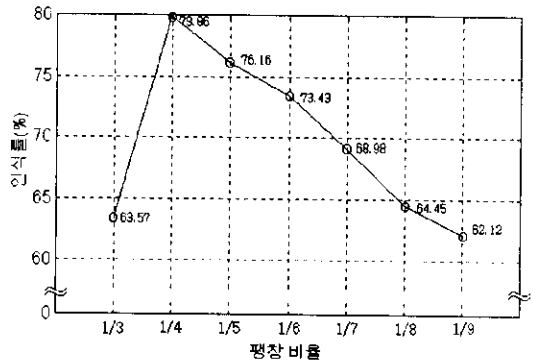


그림 4-2 팽창 비율 의존 인식률(10차, MDC)

그림 4-2에 나타난 결과에서 적절한 팽창 비율은 폰트 크기의 1/4이 적절한 수준임을 알 수 있다.

인식을 수행함에 있어, 3.2절에서 언급한 3가지 분류기를 적용하여 MDC에서는 폰트 크기 36의 0도 회전된 문자의 특성 벡터를 기준 벡터로 삼고 NNC는 폰트 크기 36의 0, 22, 45, 90도 회전된 4 문자부류의 특성 벡터를 기준 벡터로 삼았다. NNMC는 NNC의 4 문자 부류의

특성 벡터를 평균하여 기준 벡터로 삼고 입력 벡터와의 최소거리로서 문자를 분류한다. 이들에 대한 결과가 그림 4-3에 나타나 있으며 NNMC 방법이 MDC 방법의 인식률에 비해 10 % 이상 증가함을 알 수 있으며 NNC 방법과는 거의 유사함을 알 수 있다.

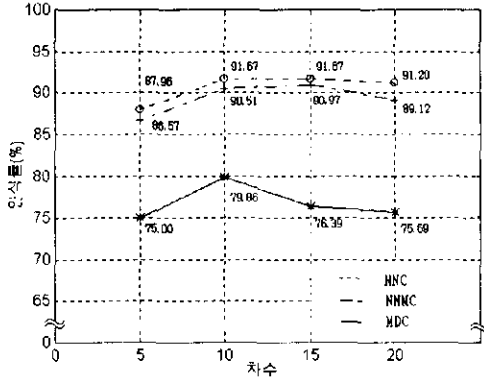


그림 4-3 분류기에 따른 인식률(1/4 팽창)

### 4.2.3 유사 문자를 제외한 인식률

위의 실험 결과에서 나타난 원형 마스크 방식의 장점에도 불구하고 팽창 윤곽선 방식은 윤곽선만을 사용한다는 단순성 때문에 인식의 한계가 있다. 팽창된 외부 윤곽선이 유사한 문자 "O"와 "0", "H"와 "M"은 일정 팽창 후에는 윤곽이 유사하게 되며, "6"과 "9", "V"와 "A"는 윤곽선의 푸리에 기술자를 추출하는 과정에서 방향 정보를 잃어버리므로 식별이 곤란하게 된다. 그림 4-4는 문자 "H"와 "M"의 팽창 근사에 의한 기술자의 유사도와 "A"와 "V"의 회전에 의한 유사도에 보여주고 있다.

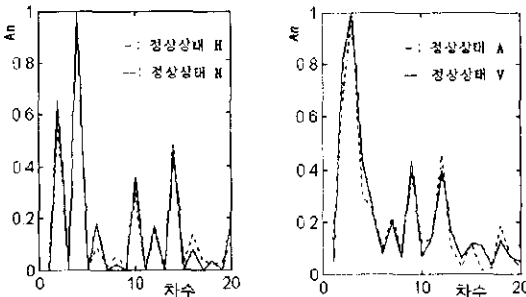


그림 4-4 팽창된 ("H"- "M"), ("A"- "V")에 대한 벡터의 비교

그림 4-5는 문자 전체와 유사문자를 제외한 후 각 분류기에 대한 인식률을 조사한 것이다. 이때 NNMC에 의한 최고의 인식률은 NNC에 의한 인식률과 같은 97.62%를 보이고 있으나 계산 속도는 MDC와 같이 4배 정도 빠르다.

## 5. 결론

본 연구에서는 문자인식의 강인성을 제공하는 문자 패턴의 팽창 윤곽 특성에 의한 문자 인식 방법을 제안하고 그 유효성을 영문 알파벳과 숫자에 대한 인식 실험을 통

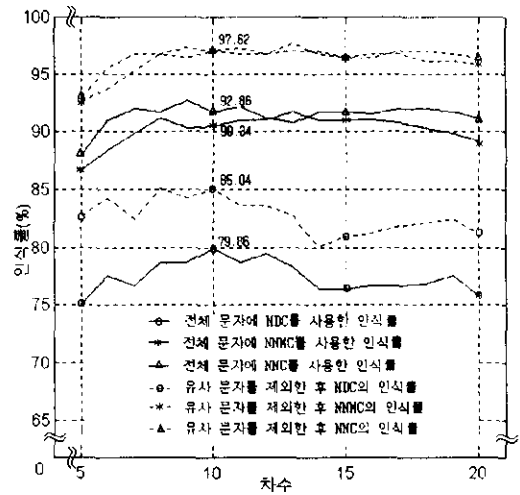


그림 4-5 각 분류기에서 유사 문자("O"- "0", "6" "9", "A"- "V", "H"- "M")를 제외한 인식률과의 비교

해서 확인하였다. 제안된 인식 방법은 문자 패턴의 회전 및 크기 변형에 무관한 인식성을 제공하며 문자 패턴의 부분적 열화나 내부 잡음에도 강인성을 지닌다.

동일한 문자에 대한 팽창 윤곽선의 푸리에 기술자는 문자 패턴의 미소한 변형에도 고차 성분이 상당히 달라질 수 있다. 따라서 인식률 개선을 위해서는 적절히 고차 성분을 배제할 필요가 있고 실험적으로는 10차에서 15차이하의 범위로 차수를 제한하는 것이 바람직함을 확인하였다.

본 접근법은 불행히도 팽창윤곽선의 특성상 팽창윤곽이 유사한 4개 문자 쌍: (O, 0), (H, M), (6, 9), (V, A)에 대한 분류 능력이 없다. 실험적인 문자의 인식률은 이동, 크기 그리고 회전 변형에도 불구하고 91.67%였다. 그러나 유사 문자를 제외한 인식률은 97.62%로 매우 높다. 이 사실은 본 접근법과 기존의 접근법을 함께 병행하면 강인하고 보다 고정도의 문자인식이 가능함을 시사한다.

## 참고 문헌

1. 박영석, 최철용, "회전 불변특성을 가지는 패턴 팽창법," 한국정보과학회 봄학술발표논문집 Vol.25, No. 1, 1998, pp. 678-680.
2. C. T. Zahn, & R. Z. Roskies, "Fourier Descriptors for Plane Curves," IEEE Trans. on Computer, Vol. 21, No. 1, 1972, pp.269-281.
3. 박영석, "문자인식을 위한 팽창 윤곽선 특성," 경남대 부설 신소재 연구소 논문집, 제 10권, 1998.
4. S. Moharir, Pattern-Recognition Transforms, John Wiley & Sons Inc. 1992.