

Mellin 변환 방식과 BPEJTC를 이용한 영상 문자 인식

(Image Character Recognition using the Mellin Transform and BPEJTC)

서준원* · 고성원 · 이병선

(Choon - Weon Seo · Sung - Won Ko · Byung - Sun Lee)

요 약

자연계에서 다양한 형태로 입력되는 물체 영상을 효과적으로 인식하려면, 물체의 위치, 회전, 크기 변화에 관계 없이 인식할 수 있는 왜곡 불변 특성의 추출이 반드시 요구된다. 이러한 왜곡 불변 특성은 동일한 영상의 변화에 대하여 인식 특성이 같고, 서로 다른 영상의 변화에 대해서는 분리 식별이 용이해야 한다. 이러한 인식 특성을 얻기 위해 다각도로 많은 연구가 진행되고 있으며, 특히 회전 및 크기에 불변 특성을 동시에 얻을 수 있는 Mellin 변환을 이용한 방법 등이 영상 인식에 많이 이용되고 있다[1][2][3].

따라서, 본 논문에서는 Mellin 변환 방법에 의한 크기 및 회전에 대한 불변 특성을 얻을 수 있는 문자 인식 시스템을 위한 문자 특징 추출 방법을 제시하고자 하였으며, 영문자 26 문자의 입력 영상에 대하여 무게 중심법에 의한 문자 이동과 Mellin 변환 방법에 의한 특징 추출 방법에 보간법을 이용하여 특징을 추출하였으며, 추출된 특징에 대하여 특징의 이질도를 검사하여, 각 특징의 이질도가 약 50% 이상의 결과를 얻었다. 또한, Mellin 변환 방법에 의해 추출된 특징을 기준 영상으로 하는 BPEJTC(Binary Phase Extraction Joint Transform Correlator)를 이용하여 크기, 회전 및 이동에 따른 입력 문자의 인식이 가능한 BPEJTC 시스템을 구현하였으며, 이에 따라 본 논문에서는 약 90%의 인식률을 얻을 수 있었다.

따라서 본 논문에서 제시하는 Mellin 변환 방법에 따라 추출된 문자의 특징과 BPEJTC를 이용하여 영상 문자를 인식할 수 있는 영상 문자 인식 시스템의 가능성을 제시하였다.

Abstract

For the recognizing system to be classified the same or different images in the nature, the rotation, scale and transition invariant features is to be necessary. There are many investigations to get the feature for the recognition system and the log-polar transform which is to be get the invariant feature for the scale and rotation is used.

In this paper, we suggested the character recognition methods which are used the centroid method and the log-polar transform with the interpolation to get invariant features for the character recognition system and obtained the results of the above 50% differential ratio for the character features.

And we obtained the about 90% recognition ratio from the suggested character recognition system using the BPEJTC which is used the invariant feature from the Mellin transform method for the reference image. and can be recognized the scaled and rotated input character.

Therefore, we suggested the image character recognition system using the Mellin transform method and the BPEJTC is possible to recognize with the invariant feature for rotation, scale and transition.

Key Words : Mellin transform, character recognition, invariant feature, BPEJTC

* 주저자 : 김포대학 전자정보계열 조교수
Tel : 031-999-4652, Fax : 031-999-4775
E-mail : cwseo@kimpo.ac.kr
접수일자 : 2003년 5월 27일
1차심사 : 2003년 6월 5일
심사완료 : 2003년 6월 24일

I. 서 론

과학기술의 발전이 급속하게 이루어짐에 따라 현대 사회에서는 수많은 양의 정보를 신속하게 처리할 수 있는 자동화 시스템의 개발이 절실히 요구되고

있으며, 첨단 반도체 기술 및 컴퓨터 과학의 발달로 컴퓨터를 이용한 신호 처리 기술은 급속도로 연구 발전되고 있다. 이미 이러한 시도들은 여러 각도에서 가시적인 효과를 나타내고 있으며 영상 인식, 음성 인식 등의 분야에서 수년간 많은 연구가 활발하게 진행되고 있다. 특히, 로봇트 시각 시스템, 자동 감시 시스템, 자동 표적 인식 시스템과 같이 인간의 시각을 모방한 영상 정보 처리 시스템이 매우 중요한 위치를 차지하고 있으며 활발히 연구되고 있다. 일반적으로, 물체 인식이란 규격화된 영상이 아닌 자연 현상에서 입력된 랜덤한 정보로부터 인간과 같이 필요한 영상을 자연스럽게 인식하는 것이다. 여기서 입력 정보의 랜덤성이란 자연계에서 발생하는 외부 잡음 이외에도 물체의 위치 이동 및 회전과 같은 상태 변화 및 관측 거리에 따른 크기 변화도 포함된다.

회전 및 크기 불변 특성은 CHF(circular harmonic filter)[1], Mellin 변환[2]등에서 각각을 얻을 수 있으며, 이러한 불변 알고리즘들은 주파수 영역에서 형성되는 고조파 성분을 이용하기 때문에 입력 영상 신호의 푸리에 변환이 필수적이다. 한편, 회전 및 크기 불변 특성을 동시에 갖고 있는 알고리즘으로 Mellin 변환 방법이 있는데,[3] Mellin 변환[4-7]은 푸리에 변환 없이 단순히 영상의 좌표계만을 변환하는 방법으로 시스템 구현 시 수학적 연산처리가 아니라 메모리 매핑(memory mapping) 방법을 이용할 수 있으므로 고속처리가 가능하다. 그러나, Mellin 변환 방법은 모든 입력영상에 대해 변환이 시작되는 중심점을 검출해야 하는 제약점이 있어 전처리 과정을 요구하게 된다. 이에 대하여 무게중심을 구하는 방법 및 푸리에 변환에 의한 이동 불변 등의 방법이 이용되고 있지만, 문자의 경우 고정된 형태를 갖고 있으므로 무게중심에 의한 방법이 많이 사용되고 있다. 이는 위치에 불변인 성질을 나타내며, 무게 중심에 의한 문자 중심을 찾고 이에 대하여 Mellin 변환 방법으로 회전, 크기 및 이동에 불변인 특징을 추출할 수 있다.

이에 따라 Mellin 변환 방식을 사용함으로써 디지털적인 문자 인식의 특징 추출에 비하여 고유한 특성을 갖는 특징들을 추출할 수 있으며, 이에 따른 회

전 및 크기의 불변인 문자 인식이 가능하다.

따라서, 본 논문에서는 위치, 회전, 크기 변화에 대한 불변 특성을 얻기 위하여 무게 중심법을 이용하여 이동 불변 특성을 얻음과 동시에 회전과 크기 변화에 무관한 변환의 중심점을 얻어 Mellin 변환을 적용하고자 한다.

II. Mellin 좌표 변환

1. Mellin 변환 알고리즘

일반적으로, $x-y$ 좌표 평면상의 임의의 함수 $w(x, y)$ 가 존재하는 위치에 회전 및 크기에 무관한 불변 특성을 얻기 위해서는 극좌표계로 이동된 함수 $p(\rho, \theta)$ 로 변환하여 얻을 수 있으며, 다음과 같다.

$$\rho = \sqrt{(x^2 + y^2)}, \quad \theta = \arctan(y/x) \quad (1)$$

여기서, x, y 는 직각 좌표계의 x 축 좌표 값과 y 축 좌표 값, ρ 와 θ 는 직각 좌표계에서 변환된 극 좌표계의 값, ρ 는 직각 좌표계의 한 점과 원점의 길이 즉, 반지름을 나타내며, θ 는 직각 좌표계와 원점이 이루는 선분이 1사분면과 이루는 각도를 나타낸다.

식 (1)에서 보는 바와 같이 직각 좌표계의 한 점 p 가 임의의 각도로 회전하는 경우 θ 값만이 변하게 되며, 이는 회전에 대하여 불변인 특성을 나타내며, 또한 p 가 원점에서 멀어지거나 작아지는 경우 θ 는 고정되고 ρ 만 커지거나 작아지게 되는 값을 갖는다. 이는 역시 크기에 대하여 불변인 특성을 나타낸다.

이러한 극 좌표계로의 변환에 의하여 나타내는 $p(\rho, \theta)$ 점은 그림 1과 같이 실수축과 허수축을 갖는 복소평면의 실수축 θ 와 허수축 ρ 를 갖는 복소평면을 가정하면, $x-y$ 평면상의 한 점인 w 는 식 (2)에 의하여 또 다른 좌표평면의 한 점으로 변환될 수 있다[1].

$$w = e^z \quad (2)$$

여기서, z 는 복소평면으로 $z = \rho + j\theta$ 로 주어진다. 이에 따라, 다음과 같이 $x-y$ 영상평면내의 화소들에 대수변환을 취하면, 식 (3)와 같이 설정이 가능

하다.

$$\ln(x + jy) = \ln[e^{\rho + i\theta}] = \ln(r) + i\theta = \rho + j\theta \quad (3)$$

물체 영상의 크기가 a배만큼 변화된 경우 Mellin 변환을 취하게 되면, $u = \ln(ar) = \ln(a) + \ln(r)$ 에 의하여 ρ 축 방향으로의 이동만이 나타난다. 또한, 영상이 회전한 경우는 ρ 성분은 무관하게 θ 성분은 대해서만 변화가 발생하게 된다[8].

결과적으로, 입력영상 $r(x, y)$ 에 대하여 입력영상 $r(x-x_0, y-y_0)$ 의 크기가 변하고 회전이 발생하더라도 $r(x, y)$ 에 의한 $w(\rho, \theta)$ 는 $w(\rho + \alpha, \theta + \beta)$ 만을 가질 뿐 변환된 특징의 모양은 크기와 회전에 대하여 불변인 형태로 이루어지게 된다. 이와 같은 관계로 인하여 그림 1에 따라 Mellin 변환을 이용하면, 크기 변화량은 $\ln(r)$ 성분 그리고 회전량은 θ 성분에서 의하여 이동 성분으로 변화되어 나타난다. 특히, Mellin 변환을 위하여 주어진 입력이 무게 중심법에 의하여 생성된 문자 영상임을 고려하면, 위치 불변 특성이 함께 나타남을 알 수 있다.

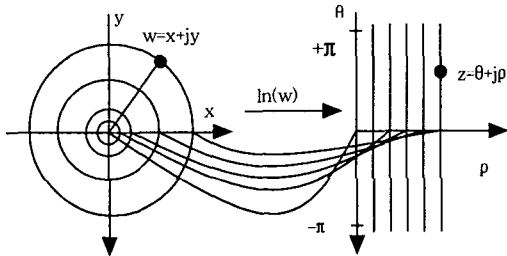
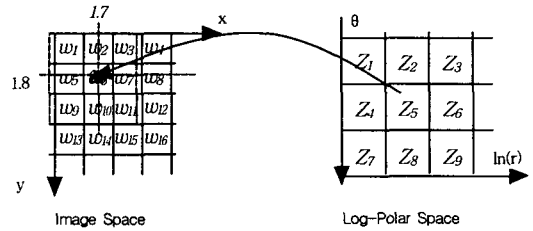


그림 1. Mellin 변환을 위한 메모리 매핑
Fig. 1. Memory mapping for the Mellin Transform

2. Mellin 변환의 보간법

그림 1에서 좌표변환평면의 공간좌표 (ρ, θ) 를 갖는 화소에 대하여 ρ 는 x - y 평면상에서 $\ln\sqrt{x^2 + y^2}$ 로부터 $\ln\sqrt{(x + \alpha)^2 + (y + \beta)^2}$ 까지 걸쳐 있으며, θ 좌표는 $\arctan[y/(x + \alpha)]$ 에서 $\arctan[(y + \beta)/x]$ 에 걸쳐 있게 된다.



$$z_6 = [(0.7)(0.8)w_1 + (1.)(0.8)w_2 + \dots + (0.3)(0.2)w_{16}] / 9$$

그림 2. Mellin 변환 오차를 보정하기 위한 보간법
Fig. 2. Interpolating Method for the Error Correction of the Mellin Transform

이러한 범위는 변환의 중심 ($\rho=0$)에 가까운 점일수록 α 와 β 의 값이 작게 되며, 중심에서 먼 점에서는 매우 큰 값을 지닌다. 이러한 이유에서 x - y 평면에 존재하는 화소는 중심에 보다 가까운 화소가 보다 먼 화소에 비하여 대응 영역이 보다 크게 분포하게 되는데 이것은 인간의 시각계가 보다 가까운 물체를 보다 세밀히 보려는 것과 같은 것이다.

따라서, 본 논문에서는 영상의 각 화소가 대응하는 변환 영역내에 범위를 계산하고 이 범위에 대응하는 각 화소의 밝기 정보를 위치시키는 방법을 도입하였다[9]. 즉, 좌표 변환된 영역내 각각의 화소에 상응하는 영상의 한 지점이 정의되고, 이 정의된 점 주변 화소들의 밝기를 이용한 보간법을 사용하게 된다.

3. Centroid에 의한 무게 중심법

입력된 영상에서 물체의 중심을 구하는 것은 Centroid 공식으로부터 쉽게 구할 수 있으며 다음 식과 같이 구해진다. 임의의 영상을 $f(x, y)$ 라고 할 때 x 축의 Centroid와 y 축의 Centroid는 다음과 같다 [10].

$$H(x) = \sum_{y=0}^N f(x, y) \quad H(y) = \sum_{x=0}^N f(x, y) \quad (4)$$

$$\text{Centroid}_x = \frac{\sum_{x=0}^N xH(x)}{\sum_{x=0}^N H(x)} \quad (5)$$

$$\text{Centroid}_y = \frac{\sum_{y=0}^N yH(y)}{\sum_{y=0}^N H(y)}$$

영상에서 물체의 위치가 구해지면 원하는 위치와
의 차이가 모터 제어기의 입력으로 전달되는 데 물
체의 위치는 항상 영상의 중심에 놓이도록 제어한다
면 그 입력 값은 물체의 현재 위치와 영상 중심과의
차이가 된다. 즉, 영상을 256×256로 샘플링 한다면
영상의 중심은 (x, y)가 (128, 128)이 된다.

III. BPEJTC 시스템

JTC 시스템은 푸리에 입력평면을 상하단으로 분
리하여 하단에 기준영상을 위치시키고, 상단에는 시
스템의 입력영상을 위치하여 구성되는데 식(6)과 같
은 다중입력영상에 대한 JTC 시스템을 분석하면 다
음과 같다.

$$r(x-x_0, y-y_0), \tag{6}$$

$$s(x, y) = \sum_{i=1}^n s_i(x-x_i, y-y_i)$$

r(x,y)와 si(x,y)를 동일한 영상으로 가정하고, 식
(6)의 기준영상과 다중입력영상을 동시에 푸리에 변
환하고 제곱함으로써 JTCS를 얻을 수 있으며 이를
다시 역푸리에 변환을 통하여 식(9)와 같은 최종 상
관 결과를 얻게 된다.

여기서, ⊗와 *는 상관함수(correlation function)
와 상승적분(convolution)을 나타내고, 첫번째항과
두번째항은 기준함수와 입력영상 자체에서 발생하
는 자기상관 성분으로 상관첨두치는 DC를 중심으로
이동없이 나타나게 된다. 그리고, 두번째항과 세번째
항은 기준영상과 입력영상간에 발생하는 상호상관
결과이고, 네번째와 마지막항은 입력영상간에 발생
하는 이중상관성분을 나타내며, 상관첨두치는 DC를
중심으로 영상간의 분리간격에 따라 원점대칭으로
나타나게 된다.

$$C(x, y) = r(x, y) \otimes r(x, y) + \sum_{i=1}^n [s_i(x, y) \otimes s_i(x, y)] + \sum_{n=1}^n [r(x, y) \otimes s_n(x, y)] * \delta(x_0 - x_i, y_0 - y_i) + \sum_{n=1}^n [r(x, y) \otimes s_n(x, y)] * \delta(-x_0 + x_i, -y_0 + y_i) + \sum_{i=1}^{n-1} \sum_{k=i+1}^n [s_i(x, y) \otimes s_k(x, y)] * \delta(x_i - x_k + y_i - y_k) + \sum_{i=1}^{n-1} \sum_{k=i+1}^n [s_i(x, y) \otimes s_k(x, y)] * \delta(-x_i + x_k - y_i + y_k) \tag{7}$$

BPEJTC 시스템[11]은 기존의 광 JTC 및 광
BJTC 시스템을 다중표적에 적용하는 경우에 발생
하는 과도한 DC, 상관오류등을 제거하고 분리조건
에 보다 탄력적으로 대처하기 위하여 광세기 분포로
나타나는 JTCS에서 자기상관 성분과 동일영상 평면
에서 나타나는 상호상관을 효과적으로 위상함수만
을 추출함으로써 상관판별력을 보다 개선하고, 광효
율을 극대화하였으며 또한, 공간 광 변조기를 이용
한 광학적 구현을 용이하게 하였다.

$$E_{bny}(u, v) = \begin{cases} 1 & \cos(\phi) \geq 0 \\ -1 & \cos(\phi) < 0 \end{cases} \tag{8}$$

식 (8)에서 보여주고 있는 (u, v)는 주파수 좌표를
나타내며, E_{bny}(u, v)는 이진화된 파워 스펙트럼을
나타내며, ϕ=π 혹은 ϕ=-π의 값을 나타내므로써 단지
위상에만 영향을 받게 된다. 또한, cosine 값은 입력
과 기준영상의 위상을 동시에 지니고 있으므로 상관
평면에서 얻는 결과는 POC와 유사하게 된다. 따라
서, 상관결과는 식(9)와 같이 얻을 수 있다.

$$C(x, y) = \sum_{n=1}^n \{ \text{edge}[r(x, y)] \otimes \text{edge}[s_i(x, y)] \} * \delta(x_0 - x_i, y_0 - y_i) + \sum_{n=1}^n \{ \text{edge}[r(x, y)] \otimes \text{edge}[s_i(x, y)] \} * \delta(-x_0 + x_i, -y_0 + y_i) \tag{9}$$

위식에서 edge[·]는 영상의 이상적인 경계를 나
타내며, ⊗는 convolution을 그리고 *는 상관을 나타
낸다.

IV. 실험 및 결과

1. 문자특징 추출

문자 영상에 대하여 회전 및 크기에 불변인 특징을 추출하기 위하여 그림 3과 같은 순서에 의하여 특징을 구하였으며, 각 문자에 대하여 이질성 여부를 검사하였다.

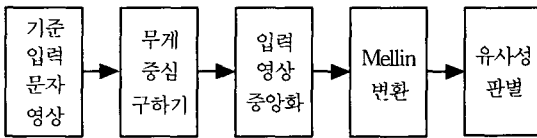


그림 3. Mellin에 의한 특징 추출 및 이질성 판별 블록도
Fig. 3. Block Diagram for the Feature Extraction and the differential ratio with Mellin transform

그림 3에서 보는 바와 같이 기준 입력 문자 영상에 대하여 Mellin 변환에 의한 특징을 추출하며, 이와 같은 방법으로 다른 임의의 입력 문자 영상에 대하여서도 같은 방법으로 특징을 추출하여 구한 Mellin 변환 영상과 이질성을 비교한다.

본 논문의 실험에 사용된 컴퓨터는 인텔 펜티엄-III 이고, 폰트는 페이트 샵 프로 5.0을 사용하였으며, 사용된 폰트는 Baskerville Old Face이며, 사용된 폰트의 크기는 72 포인트가 된다. 사용된 폰트의 문자는 영문 대문자를 사용하였으며, 크기 변화를 보기 위하여 72, 48 포인트의 폰트를 사용하였다. 또한 회전된 폰트는 90°각도마다 회전된 문자 영상을 사용하였다. 폰트는 A부터 Z까지 사용하였으며, 본 논문에서는 A부터 D까지의 폰트만을 나타내었다. 사용된 프로그램 언어는 마이크로소프트 비주얼 C++ 6을 이용하였다.

그림 4에서는 본 논문에서 사용된 Mellin 변환 방법을 도식화하였다. 입력 문자 256×256 크기에 대하여 (128, 128)을 중심으로 하여 변환하며, 변환 범위는 θ 에 대하여 0에서부터 4π 까지의 경우를 변환하였다.

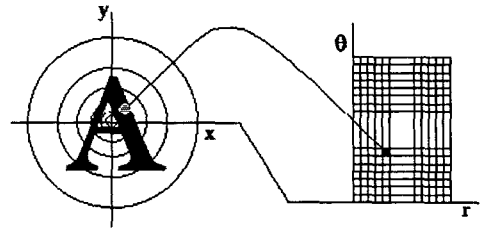


그림 4. Mellin 변환의 매핑에 의한 직각 좌표계 좌표 변환
Fig. 4. Mapping Method of the Mellin transform

그림 5에서는 임의의 위치에 입력되는 문자에 대하여 Centroid 기법을 사용하여 문자의 중심을 알아낼 수 있다.

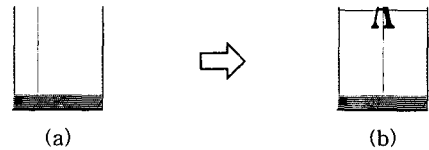


그림 5. Centroid 방법에 의한 문자 이동
Fig. 5. Moving Character using the Centroid Method

Centroid 기법에 의하여 알아낸 문자의 중심을 Mellin 변환을 위하여 입력 영역의 중심으로 좌표를 변환한다. 이로서 Mellin 변환에서 문제가 되고 있는 중심점 찾기를 해결할 수 있다. 그림 5의 (b)는 256×256 크기의 입력 영역에 'A' 문자의 중심을 (128, 128) 위치로 이동한 모습을 나타낸다.

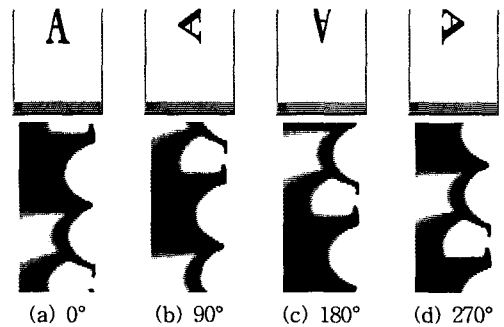


그림 6. 문자 'A'의 회전에 의한 특징(4 π)
Fig. 6. Feature Extraction for the rotated character 'A' (4 π)

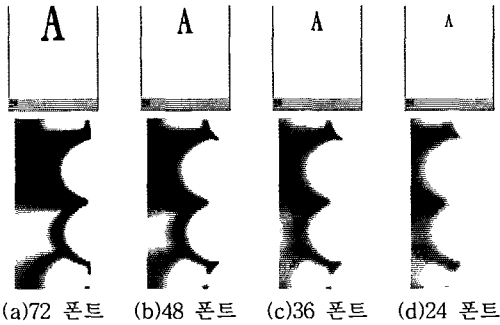


그림 7. 문자 'A'의 크기에 의한 특징(4 π)
 Fig. 7. Feature Extraction for the rotated character 'A' (4 π)

그림 6에서는 문자 A에 대해서 Mellin 변환에 의한 특징 추출 결과를 보여주고 있다.

문자 'A'를 90°마다 회전시킴으로서 특징도 y축으로 이동이 되는 것을 알 수 있다. 그러나 전반적인 특징의 형태는 변화하지 않음을 알 수 있다.

문자의 360° 회전에 따라 고정된 특징 패턴을 구하기 위하여 8에 대하여 2배 이상의 주기를 사용하면 같은 패턴이 나타남을 알 수 있다. 즉, 8를 0부터 4 π 이상의 주기를 변환하면 된다. 이와 같은 방법에 의하여 변환한 결과가 그림 6와 그림 7에서 보여주고 있다.

그림 7에서는 크기에 대하여 같은 패턴이 나타남을 알 수 있다. 이와 더불어 회전에 대하여 같은 패턴이 나타나는 형태가 그림 6에서 보여주고 있는데, 이 때 같은 패턴은 y축으로 이동만이 있을 뿐 그 형태뿐만 아니라 패턴의 위치도 유지함을 알 수 있다.

표 1. 각 문자들의 Mellin 변환 특징의 이질도
 Table 1. Differential ratio of the log-polar transformed features for characters

| | | | | | | | | | |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 문 자 | A | B | C | D | E | F | G | H | I |
| 이질도 | 64.15 | 61.23 | 47.12 | 48.46 | 60.92 | 62.08 | 48.54 | 59.35 | 59.50 |
| 문 자 | J | K | L | M | N | O | P | Q | R |
| 이질도 | 55.95 | 63.69 | 48.23 | 60.58 | 62.38 | 47.96 | 61.00 | 49.88 | 63.54 |
| 문 자 | S | T | U | V | W | X | Y | Z | |
| 이질도 | 61.42 | 59.27 | 47.31 | 60.08 | 65.19 | 62.04 | 61.12 | 61.46 | |

식 (10)에서는 i 번째 문자에 대해서 A부터 Z의 문자에 대한 이질도를 나타내고 있다. 이식에 의하여 한 문자에 대한 다른 문자와의 이질도를 검사 할

수 있으며, 이를 표 1에서 나타내고 있다.

$$D_i = \frac{\sum_{j=A}^Z [(-\frac{DPN_{ij}}{TPN_i}) \times 100]}{TCCN} \quad (10)$$

식 (10)에서 DPN_{ij} 는 i 번째 문자와 j 번째 문자의 다른 픽셀의 숫자를 나타내며, TPN_i 는 i 번째 문자의 전체 픽셀수를 나타내고, $TCCN$ 은 비교되는 문자의 전체 수를 나타낸다.

Mellin 변환에 의하여 얻어진 문자의 특징은 픽셀 매핑에 의하여 특징이 추출됨으로서 픽셀의 유사성이 다들수록 이질도가 높아지게 되어 있다. 이를 바탕으로 얻은 이질도는 표에서 보는 바와 같이 각 문자들은 다른 문자들과의 이질 정도가 약 50%정도에 이르는 것을 알 수 있다.

이는 다른 문자들과의 영상 특성이 상이하게 다를 수 있으며, 이는 패턴 인식 시스템에 기준 입력 패턴으로 사용이 가능하며, 패턴 매칭 등의 방법으로 인식이 가능함을 알 수 있다.

2. 문자 인식 시스템

Mellin 변환 방법에 의하여 문자 영상에 대한 회전 및 크기의 불변인 특징을 추출한 후에 그림과 같은 순서에 의하여 해당 문자의 인식에 대한 실험을 하였다.

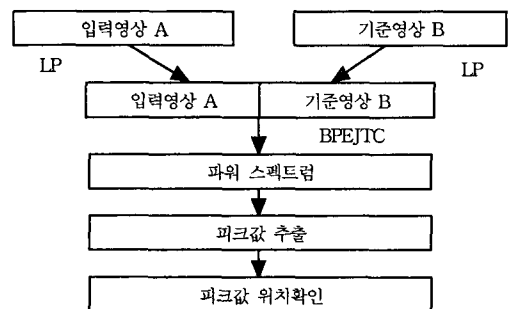


그림 8. BPEJTC에 의한 문자 인식 블록도
 Fig 8. Block Diagram for character recognition with the BPEJTC

그림 8에서 보는 바와 같이 입력 영상과 기준 영상을 각각 Mellin 변환을 한 다음 각각의 문자 영상을 한 프레임으로 설정한다. 설정된 한 프레임의 영

Mellin 변환 방식과 BPEJTC를 이용한 영상 문자 인식

상은 본 논문에서 사용하고 있는 BPEJTC를 사용하여 변환한 후 주파수 영역의 파워 스펙트럼을 확인한다. 파워 스펙트럼에서 가장 높은 값을 갖는 피크값이 Correlator의 결과로서 문자의 영상에 대한 인식을 나타낸다.

본 논문에서 사용한 입력 영상과 기준 영상은 256×256 이며, 입력 영상은 72, 48 포인트의 폰트를 사용하였으며, 기준 영상은 72 포인트의 영상을 사용하여 각각 Mellin 변환을 하였다. 이후 BPEJTC를 사용하기 위하여 Mellin 변환된 입력 영상과 기준 영상을 한 프레임에 설정한다. 한 프레임의 크기는 256×256 으로 하였으며, Mellin 변환된 영상은 256×256 프레임 영상내에 각각 32×64 크기로 입력 영상과 기준 영상을 한 프레임의 상단 중앙과 하단 중앙에 위치하도록 하였다. 다시 256×256 의 프레임 영상을 512×512 크기의 영상으로 변환하였으며, 이때 256×256 의 프레임 크기는 변환하지 않고, 현 크기를 512×512 크기의 중앙에 위치하도록 하였다. BPEJTC에 의하여 변환된 주파수 영역에서의 파워 스펙트럼은 512×512 의 크기로 결과가 나오며, 파워 스펙트럼의 결과에서 최대치의 피크값을 선택한다.

한 문자의 Mellin 입력 영상에 대하여 A부터 Z까지의 Mellin 변환된 기준 영상에 대한 각각의 피크값들 중에서 가장 높은 피크값을 선택하여 해당하는 기준 영상을 인식 문자로 결정한다.

1) 입력 프레임 구성하기

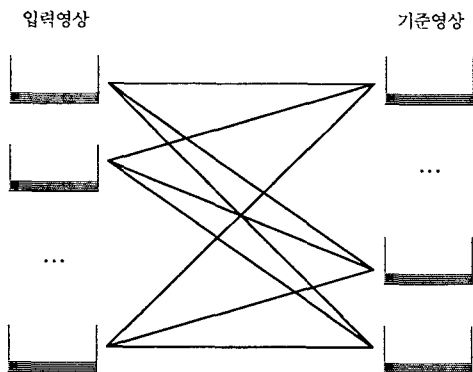


그림 9. 입력 영상과 기준 영상의 프레임 구성
Fig. 9. Composed Frames with input and reference images

입력 프레임은 그림 9와 10과 같이 구성하며, 이때 입력 프레임의 입력 영상은 A에서부터 Z까지 구성되며, 기준 영상 또한 A에서 Z까지 구성된다.

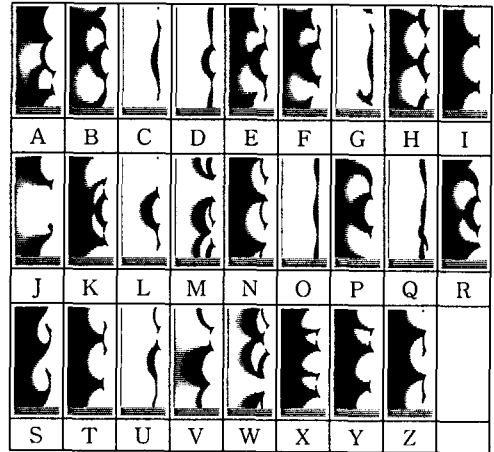


그림 10. 72포인트, 0° 회전된 영문자(A~Z)의 Mellin 변환된 결과 ($0^\circ \sim 720^\circ$)
Fig. 10. Mellin-transformed results for 72pt., 0° characters ($0^\circ \sim 720^\circ$)

그림 10에서는 'A'에서부터 'Z'까지의 문자 영상에 대한 Mellin 변환된 결과를 보여주고 있다. 본 논문에서 사용하고자 한 Mellin 변환 범위가 0° 에서부터 720° 의 결과이다.

2) BPEJTC 구현

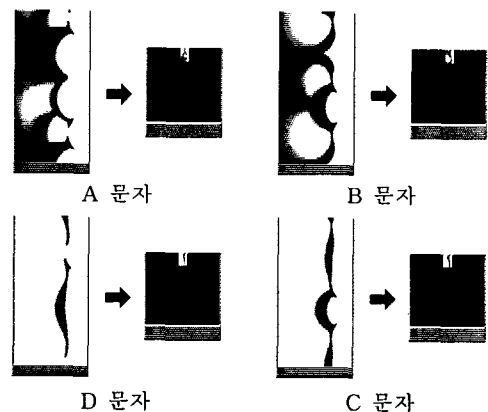


그림 11. Mellin 변환된 결과의 입력 프레임에 매핑
Fig. 11. Mapping for the input frame of the Mellin-transformed results

그림 11에서는 문자 'A', 'B', 'C', 'D'의 Mellin 변환된 영상을 BPEJTC를 위하여 256×256 평면의 중앙에 32×64 크기로 매핑시키는 것을 보여주고 있다. 본 논문에서는 입력 프레임에서 입력 영상과 기준 영상에 대하여 같은 방식을 사용하였다.

그림 12의 좌측에서는 BPEJTC의 주파수 도메인에서 알 수 있는 파워 스펙트럼을 보여주고 있다. 파워 스펙트럼에서 알 수 있듯이 피크는 두가지 이상이 검출되고 있다. 피크의 종류는 자기 상관과 이중 상관으로 검출될 수 있으며, 원점에 대하여 대칭의 모습을 보여주고 있다. 자기 상관과 이중 상관을 나타내는 값은 같은 값으로 구성이 되며, 거리차만 존재한다. 본 논문에서는 y축 원점을 대칭으로 입력 영상과 기준 영상을 배치하였으므로, 상관 피크는 y축을 중심으로 발생하게 된다. BPEJTC를 수행한 결과 얻고자 하는 피크는 상관 피크 중 가장 높게 나타나는 피크로서 유사도가 가장 높음을 나타낸다.

그림 12는 입력 문자 영상의 문자 크기가 72 포인트의 문자로 구성되어 있음을 보여주고 있으며, 각 포인트에 대하여 0°에 대하여 입력을 구성하였다. 기준 영상에 대하여 문자 'A'부터 'Z'까지 동일하게 72 포인트, 0°의 문자 영상을 사용하였으며, 각각의 Mellin 변환된 입력 문자 영상과 기준 문자 영상을 BPEJTC의 입력 프레임으로 구성하였다.

그림 12의 좌측에서는 입력 문자와 기준 문자 중에서 입력 문자와 기준 문자가 같은 경우에는 피크 값이 매우 높게 나타나고 있다. 그러나, 그림 14에서 보는 바와 같이 문자가 서로 틀리거나 같은 입력 문자와 영상 문자라 할지라도 입력 문자가 회전이 되었을 경우에는 피크값이 점점 낮아지는 것을 알 수 있다. 그러나, 같은 문자에 대해서는 Mellin 변환된 영상의 명암만 차이가 있으므로 해서 같은 문자의 입력 영상과 기준 영상에 대해서는 기준 영상이 'A'에서 'Z'까지 변화하는 동안 입력 영상에 대하여 같은 기준 영상이 구성되었을 경우에는 26가지의 파워 스펙트럼에서 구할 수 있는 각 피크값 중에서 가장 높은 것을 알 수 있다.

그림 13에서는 입력 문자 26개의 모든 경우와 기준 문자 26개의 모든 경우에 대한 BPEJTC의 피크 값을 보여주고 있으며, 72 포인트의 0° 회전, 72 포인트의 90° 회전, 48 포인트의 0° 회전, 48 포인트의 90°

회전에 대하여 인식률은 각각 100%, 92.3%, 92.3%, 76.9%로서 평균 90.4%의 인식률을 나타냈으며, 9.6%의 오인식률을 갖는 경우가 발생하였다.

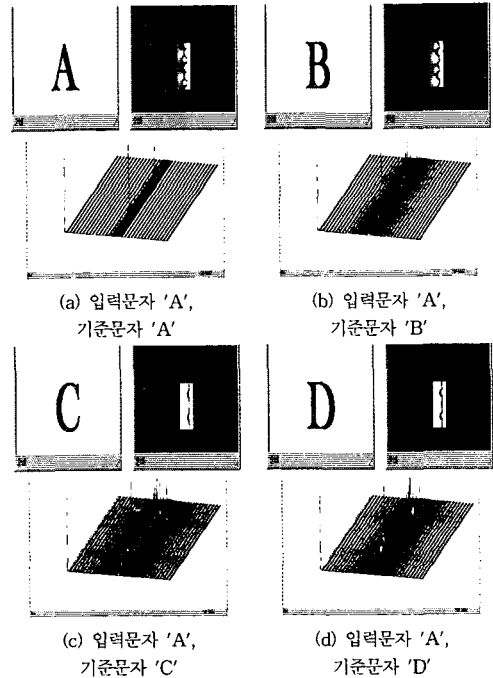


그림 12. 입력 문자 크기가 72포인트, 0° 회전된 'A' 문자의 BPEJTC 변환

Fig. 12. BPEJTC transform of input characters with 72pt., 0°

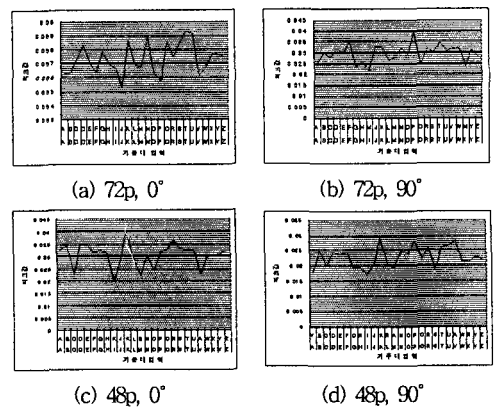


그림 13. 입력 문자 26개에 대한 BPEJTC의 최고 피크값
Fig. 13. Maximum peak value of the BPEJTC for 26 input characters

Mellin 변환 방식과 BPEJTC를 이용한 영상 문자 인식

위의 결과에서 알 수 있듯이 입력되는 문자에 대하여 Mellin 변환 방법에 의한 특징을 추출하고 BPEJTC를 이용하여 상관값을 알아본 결과 약 90.4%의 인식률을 갖으며, 9.6%의 오인식률은 입력 문자의 크기가 기준 문자의 크기에 비해 절반 정도가 작으며, 본 논문에서 실시한 BPEJTC의 입력단으로서 256×512의 특징 영상을 32×64의 크기로 축소함으로써 오인식의 결과가 발생할 수 있다고 사려된다. 이러한 결과는 실제로 가장 많이 사용되는 문자의 크기인 10에서 12 포인트에 맞추어 특징을 추출하며, BPEJTC의 입력단에는 얻어진 특징의 크기를 그대로 사용함으로써 극복될 수 있다.

3. 실험 결과 및 고찰

본 논문에서 언급한 Mellin 변환 방식과 BPEJTC를 이용한 영상 문자 인식 시스템은 근래에 많이 나와 있는 디지털 영상 처리 기술을 이용한 문자 인식 기술과는 다르게 광학적 구현을 근거로 하였으며, 이에 따라 본 논문에서 언급하고 있는 Mellin 변환 방식의 좌표 변환과 함께 BPEJTC를 이용한 상관기의 구현은 광학적 시스템을 겸비할 수 있으며, 이를 바탕으로 실시간 처리가 가능하다는 장점이 있다. 이와 더불어 기존에 사용하고 있는 디지털적인 방법에서 많이 사용하고 있는 매칭 기술은 패턴의 근본적이고 많은 데이터베이스가 형성되어야 한다는 전제 조건이 부가되는 한편 본 논문의 실험 결과로 알 수 있듯이 패턴의 이동이나 회전, 크기의 변화에 대한 불변 특징을 추출함으로써 적은 패턴의 특징 수에 따라 이를 바탕으로 한 상관기의 광학적 구현의 가능성을 한층 높였다고 할 수 있다.

본 논문의 실험으로는 72, 48 포인트 크기의 문자에 대하여 수행하였으며, 이는 본 연구의 차 모델인 광학적 구현의 기법에서 입력 영상의 입력 단에 해당하는 장치가 카메라에 의한 입력을 전제로 하였으며, 카메라의 입력에 의한 문자 입력 패턴의 크기는 실제 폰트의 크기에 크게 좌우되지 않으며, 또한 본 논문에서 보여주고 있는 크기 불변의 Mellin 변환에 의하여 카메라 입력 장치의 입력단에 의하여 실제 포인트는 크게 좌우되지 않음을 알 수 있다.

이에 따라 본 논문에서 제시하고 있는 Mellin 변

환 방식과 BPEJTC를 이용한 영상 문자 인식은 크기 및 회전 그리고 이동에 불변인 영상 문자 인식 시스템으로서 그 효율성이 높음을 제시할 수 있다.

V. 결론

본 논문에서는 26개의 영문 대문자에 대하여 고유한 특징을 추출하고, 얻어진 특징을 이용하여 다른 형태의 영문자에 대한 인식을 가능하게 하는 시스템을 개발하였다.

문자 인식을 위하여 인식하고자 하는 문자의 특징을 추출하기 위하여 특징 추출 방법으로는 Mellin 변환 방법을 사용하였으며, 이 변환 방법의 특징은 크기와 회전에 불변인 영문자 특징을 추출할 수 있었다.

Mellin 변환 방법의 단점인 크기와 회전에 고정된 이미지의 중심점을 사용한다는 점을 극복하기 위하여 무게 중심법을 사용하였으며, 임의의 위치에 대한 문자의 고유한 특징을 얻을 수 있는 근거를 마련하였다.

무게 중심법과 Mellin 변환 방법에 의한 영문자 특징 추출은 회전에 대하여 고유한 패턴을 사용하기 위해 2주기의 결과를 얻었으며, 각 문자의 변환 특징의 이질도를 비교한 결과 약 50% 이상의 패턴 이질성을 얻을 수 있었다.

다음으로 BPEJTC를 이용하여 Mellin 변환 방법에 의해 추출된 특징에 따라 영문자의 고유 특징이 형성되며, 이 고유 특징을 이용하여 다른 형태의 영문자 입력에도 인식이 가능한 BPEJTC 시스템을 구현하였으며, 이에 따라 본 논문에서는 약 90%의 인식률을 얻을 수 있었다.

본 논문의 결과에 따른 문자의 특징 추출에 의하여 기존의 방식에서 문제가 되고 있는 인식 기준의 특징들에 대한 과도한 데이터베이스 구축의 문제점을 단일 문자에 대한 단일 특징을 추출함으로써 적은 데이터베이스를 구축하였으며, 많은 문자의 크기와 회전에 따른 다양한 입력 문자의 형태에도 적응적인 인식을 가능하게 하는 시스템으로서의 개발 가능성이 매우 높다고 판단된다.

이에 따라 본 논문에서는 기존의 많은 양의 데이

터 베이스에 의하여 인식의 처리가 지연되며, 기존의 인식 기술에 의한 오인식률이 많이 나타나 인식의 정도가 떨어지는 것으로 나타나고 있는 디지털적인 인식 방법에 비해 회전과 크기에 불변인 특징을 이용하여 문자 데이터베이스의 최소화를 이룰 수 있는 Mellin 변환 방식과 BPEJTC 기술을 이용하여 문자의 다양성을 극복하고 인식률을 높일 수 있는 시스템을 제시하였다.

본 논문에서 제시한 Mellin 변환 방법에 의한 문자 특징 추출 방법과 문자 인식을 위한 BPEJTC 시스템의 구현은 영상 문자 인식뿐만 아니라 다양한 시각 패턴 인식 등에 그 응용 가능성이 있음을 알 수 있었다.

References

- [1] Ramesh C. Jain, "Segmentation of Frame Sequences Obtained by a Moving Observer", IEEE vol.PAMI-6, no.5, pp.624-629, September, 1984.
- [2] Saburo Tsuji, Michiharu Osada and Masahiko Yachida, "Tracking and Segmentation of Moving Objects in Dynamic Line Images", IEEE vol.PAMI-2, no.6, pp.516-522, November, 1980.
- [3] Jong-Kwon Won, Sang-Yi Yi, Chung-Sang Ryu, Seung-Hyun Lee, Eun-Soo Kim, "Moving Target Segmentation using the Sequential Correlation", Proc. of KICS, vol.14, no.2, pp.947-950, 1995.
- [4] James H. Duncan and Tsai-Chia Chou, "On the Detection of Motion and the Computation of Optical Flow", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.14, no.3, pp.346-352, March, 1992.
- [5] Y. N. Hsu and H. H. Arsenault, "Optical Pattern Recognition using the Circular Harmonic Expansion", Appl. Opt., vol.21, pp.4016-4025, 1982.
- [6] Anqi Ye and David Casasent, "Morphological Wavelet Transform for Distortion-Invariant Object Detection in Clutter", SPIE Wavelet Applications, vol.2242, pp.525-537, 1994.
- [7] D. J. Gregoris, S. K. W. Yu, and S. Tritchew, "Wavelet transform-based filtering for the enhancement of dim targets in FLIR images", SPIE Wavelet Applications, vol. 2242, pp.573-583, 1994.
- [8] R. Jain, S. Bartlett, and N. O'Brien, "Some Experiments in Ego-motion Complex Logarithmic Mapping", Advances in Computer Vision and Image Processing, vol.3, pp.145-177, 1988.
- [9] D. Casasent, S. F. Xia, A. J. Lee and J. Z. Jung, "Real-time Deformation Invariant Optical Pattern Recognition using Coordinate Transformations", Appl. Opt., vol.26, no.9, pp.938-942, 1987.

- [10] Skolnick MM, Brown, RH., Bhagvati, C., Wolf, B.R., "Morphological algorithms for centroid normalization in relational matching", IEEE International Symposium on Circuits and Systems 1989, vol.2, 987-990, 1989.
- [11] E. S. Kim, S. Y. Yi, and J. H. Lee, "Real-Time Tracking System based on Joint Transform Correlator and Neural Network Algorithm", Proc. SPIE, vol.1812, 1992.

이 논문은 2003학년도 김포대학의 연구비 지원에 의하여 연구되었음.

◇ 저자소개 ◇

서춘원 (徐春源)

1964년 7월 4일생. 1988년 2월 광운대학교 전자공학사. 1990년 2월 광운대학교 대학원 공학석사. 1997년 2월 광운대학교 대학원 공학박사. 1998-2000 서울정보기능대학 전자과 전임강사. 2000-현재 김포대학 전자정보계열 조교수.

고성원 (高成元)

1960년 8월 6일생. 1983년 2월 한양대학교 전자 공학사. 1995년 2월 한양대학교 대학원 공학석사. 1986-1995 한국통신 통신망 연구소. 1995-1997 영일전문대학교 전자과 전임강사. 현재 김포대학 전자정보계열 조교수.

이병선 (李秉善)

1958년 8월 5일생. 1985년 2월 서울산업대학 전자 공학사. 1992년 3월 건국대학교 대학원 공학석사. 2000년 2월 단국대학교 대학원 공학박사. 1994-1997 경북전문대학 전자과 전임강사. 현재 김포대학 전자정보계열 조교수.