

이동과 축척과 회전에 불변인 실용적인 패턴 인식 시스템¹⁾

金 會 律

漢陽大學校 工科大学 電子工學科

개 요

본 논문에서는 물체들의 이동(translation), 축척(scale) 그리고 회전방향(orientation)에 무관하게 물체를 인식하는 실용적인 패턴 인식 시스템을 소개한다. 이 시스템은 2진영상으로 변환하는데 필요한 임계치(threshold)의 큰 변화에도 덜 민감하다. 특징 벡터(feature vector)로서는 Zernike 모멘트를 사용하였는데 지금까지 잘 알려진 Hu가 제안한 7개의 모멘트 불변수(moment invariants)와 비교한다. 또한, 실용적인 기계 시각(machine vision) 시스템에 대해 세 가지 중요한 문제로서 패턴 정규화(pattern normalization), Zernike 모멘트의 신속한 계산, 그리고 k-NN 규칙을 이용한 분류 등을 논의하였다. 실험에서는 임의의 회전 방향에서 문자들의 크기가 10×10 화소(pixel)에서 512×512 화소까지 변하는 서로 다른 크기를 가진 인쇄된 62개의 문자와 숫자 그리고 기호들을 서로 다른 임계치에서 인식하는 것을 보여준다.

1. 서 론

미지의 크기와 임의의 회전방향에 있는 기호와 패턴을 인식할 수 있는 기계 시각 시스템을 사용하면 생산공정을 자동화하는데 융통성이 커진다. 인식속도가 중요한 생산공정에서는 일반적으로 이진영상을 많이 사용하는데 흔히 이용하는 특징 벡터로서는 Fourier descriptor 와 모멘트 특징을 들 수 있다. 첫번째 방법은 패턴의 윤곽선을 2차원 복소함수로 표시한 다음

그 함수를 주파수로 바꿔서 모델의 주파수와 비교하는 것이다. 이 방법은 속도 면에 있어서는 중요한 장점이 되지만, 결점으로는 물체를 외부윤곽으로만 표현하기 때문에, 외부윤곽상의 잡음에 매우 민감하다는 것이다. 예를 들면 물체의 영상이 잡음에 의하여 부분적으로 금이 간 것처럼 보일 때 윤곽선이 크게 달라지는 경우이다. 모멘트에 근거한 방법은 물체가 표시되는 영역전체를 2차원 함수로 간주하여 여러 차수 모멘트를 모델의 그것들과 비교하여 물체를 인식하는 방법이다. 이 방법은 Fourier descriptor가 갖는 단점은 없지만 모멘트를 구하기 위하여 물체가 차지하는 영역전체에 있는 화소를 전부 포함하여 계산해야 하기 때문에 특징벡터를 계산하는데 시간이 많이 걸리는 단점이 있다. 그러한 단점을 극복하기 위해 광학적으로 실시간에 모멘트를 계산하는 방법도 고안되었으며, 여러 가지 특수목적의 구조와 프로 세서들이 개발되었다.^[1,2] Prokop와 Reeves는 모멘트에 근거한 패턴 인식 시스템을 개괄적으로 소개하였다.^[3]

본 논문에서는 전용의 하드웨어를 사용하지 않고 물체의 이동, 축척 및 회전방향에 상관 없이 실시간으로 동작하는 실용적인 패턴 인식 시스템을 소개한다. 이 시스템은 또한 다진 영상으로 나타난 물체를 이진화하는데 필요한 임계치의 변화에도 다른 기종의 시스템처럼 민감하지 않다. 실험 예에서 이 시스템이 상이한 크기와 임의의 회전방향에 있는 문자들을 인식하는 것을 보여준다. 특징 벡터(feature vector)로

1) 본 논문의 영문 판은 1995년 6월 21일부터 23일까지 미국 Washington에 있는 Seattle에서 IEEE Computer Society Conference on Computer Vision and Pattern Recognition에서 같은 제목으로 발표되었다.

서 이진 영상의 Zernike 모멘트가 사용되었다. 실시간 환경에서의 실용적인 기계 시각 시스템에 대해 다음과 같은 3가지 중요한 문제 즉, (1) 정규화 과정 (normalization process), (2) Zernike 모멘트 계산, (3) 패턴 분류를 위한 k-NN 규칙 등이 논의된다. 특히 직교 좌표계에 근거한 lookup-table을 사용하여 Zernike 모멘트의 계산 속도를 향상시킨다.

II. 모멘트 (Moment)

2차원 함수 $f(x,y)$ 가 주어질 때, 차수 $p+q$ 인 정규 모멘트는 다음과 같이 정의된다.^[4,5]

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy$$

디지털 영상에서 정규 모멘트는 로 표시되는데 그 값은 이동, 회전 방향이나 축척에 따라 달라지지만, 다음과 같이 표시되는 중심 모멘트는 이동에 불변이다.

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x,y)$$

여기서 $(\bar{x} = \sum_x x, \bar{y} = \sum_y y)$ 는 물체의 중심을 나타내는데, 디지털 영상 $f(x,y)$ 의 좌표 원점은 물체의 중심으로 변환된다.

1. 모멘트 불변치 (Moment invariants)

위에 언급된 중심 모멘트는 회전에 민감하다. 회전 불변이기 위해서 Hu는 3차 이하의 중심모멘트를 조합하여 7개의 모멘트 불변치(moment invariants: HMI)를 소개하였다.^[6] 이들 불변치들은 항공기나 배의 형상 그리고 많은 다른 물체들을 인식하는데 응용되었다.^[7,8] 또한, 축척 불변을 위해서는 먼저 물체의 축척인자(scale factor) 또는 배율을 계산하고 그 값을 1 이나 상수 β 로 정규화 시키는 것이다.^[3,9,10] 여기서 축척인자라면 x, y 좌표상의 크기뿐만 아니라 그레이 스케일의 크기까지 포함하지만 본 논문에서는 x, y 좌표상의 배율 s_x, s_y 만 고려 한다.

$$\begin{aligned} \mu'_{pq} &= \int_x \int_y x^p y^q f\left(\frac{x}{s_x}, \frac{y}{s_y}\right) dx dy \\ &= \int_x \int_y (s_x x)^p (s_y y)^q f(x,y) dx dy \\ &= s_x^{p+1} s_y^{q+1} \mu_{pq} \end{aligned}$$

여기서 일반적으로 $s_x = s_y = s$ 이므로 중심 모멘트를 사용하여 상수로 정규화 시킨다면 정규화된 영역은 $\mu'_{00} = s^2 \mu_{00}$ 로 계산되고 배율 s 는 $s = \frac{\sqrt{\beta}}{\sqrt{\mu_{00}}}$ 로 계산된다. 따라서 영역 μ_0 가 변함에 따라 배율 s 가 쉽게 계산되고 특징 벡터 값에 반영된다. 이때 μ_0 를 0차 모멘트 (zeroth order moment) 혹은 ZOM이라 한다.

2. Zernike 모멘트

Zernike와 pseudo-zernike 모멘트는 정보의 중복 (information redundancy)과 잡음감도 (noise sensitivity)에서 다른 모멘트에 비해 가장 우수하다.^[5] 디지털 영상에 대한 $(m+n)$ 차의 Zernike모멘트는 정규 모멘트와 비슷하게 다음과 같이 정의된다.^[11,12]

$$A_{nm} = \frac{(n+1)}{\pi} \sum_x \sum_y f(x,y) V_{nm}^*(x,y), x^2 + y^2 \leq 1,$$

여기에서 $V_{nm}(x,y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) e^{jm\theta}, |m| \leq n, n-|m| = 2k$ 는 다음과 같이 정의된다.:

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|} (-1)^s \frac{(n-s)!}{s! \binom{n+m-s}{2} \binom{n-m-s}{2}!} \rho^{n-2s}$$

실제 영상함수 $f(x,y)$ 에 있어서는 오직 $V_{nm}(\rho, \theta)$ 만이 복소수이므로, 다음 식이 성립한다.

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) [VR_{nm}(x,y) + jVI_{nm}(x,y)], x^2 + y^2 \leq 1$$

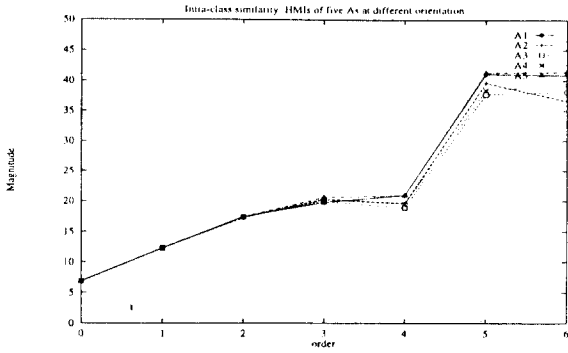
여기서 VI, VR 는 기본함수 $V_{nm}^*(\rho, \theta)$ 가 실수 및 허수부분을 나타낸다. Zernike 모멘트의 크기는 회전

표 1. 본 시스템에서 사용한 8차 이하의 Zernike moment 크기 벡터들

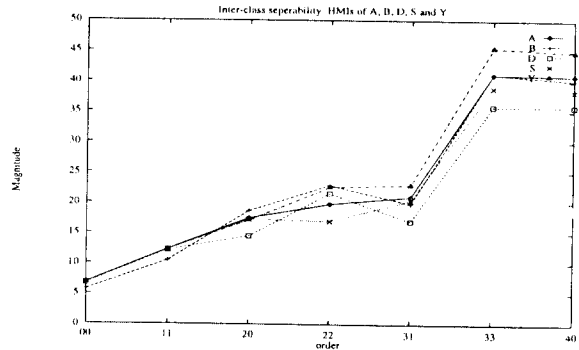
- |A₀₀|,
- |A₁₁|,
- |A₂₀|, |A₂₂|,
- |A₃₁|, |A₃₃|,
- |A₄₀|, |A₄₂|, |A₄₄|,
- |A₅₁|, |A₅₃|, |A₅₅|,
- |A₆₀|, |A₆₂|, |A₆₄|, |A₆₆|,
- |A₇₁|, |A₇₃|, |A₇₅|, |A₇₇|,
- |A₈₀|, |A₈₂|, |A₈₄|, |A₈₆|, |A₈₈|,

에 불변이다.^[4,5,9] 특징벡터에 대한 본 시스템에서는 사용한 Zernike 모멘트크기 (moment magnitude : ZMM) 벡터의 수는 n 과 m 이 8이하로 아래 표 1에 보여진 바와 같이 25개이다. 실제로는 첫번째와 두번째의 모멘트를 제외한 23개의 벡터가 쓰인다. 왜냐하면 첫 번째의 모멘트는 정규화 하는데 쓰였고, 두번째 모멘트인 A_{11} 은 정의에 따라서 0이 되기 때문이다.

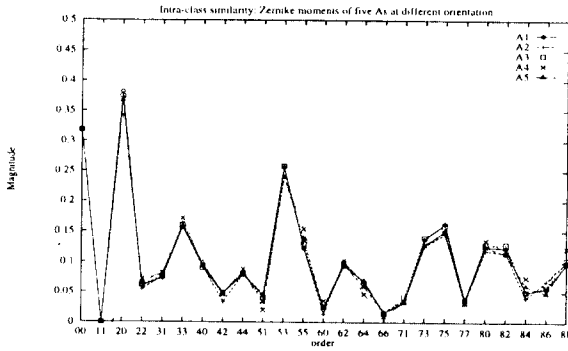
같은 부류 안에서는 두 가지 특징벡터 모두 각도의 변화에 둔감함을 보여주고 있는데 그것은 회전 불변 인식에 적합한 것이다. 한편, 그림 2에서는 'A', 'B', 'D', 'S'와 'Y' 문자들에서 추출된 HMI들과 ZMM들을 보여준다. 이들 특징들은 가능한 한 뚜렷한 차이가 있어야한다. 그림 1과 비교해 보면 HMI의 경우 그 차이가 비슷한 반면 ZMM의 경우에는 매우 다름을 알 수 있다. 다시 말하면 ZMM의 내부 종류간의 가분성이 HMI의 그것보다 훨씬 더 크다.



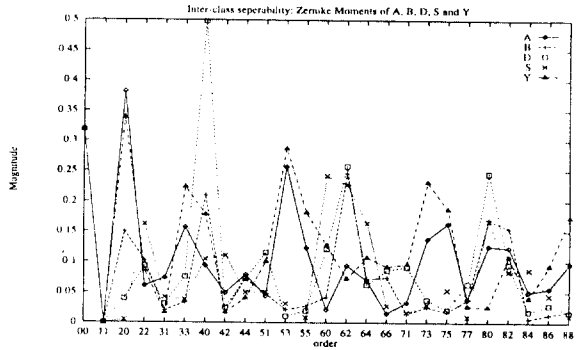
(a)



(a)



(b)



(b)

그림 1. 같은 글자 'A'에 대해 상이한 각도 0°, 30°, 60°, 90°에서 HMI(a)와 ZMM(b)의 특성비교

그림 2. 상이한 글자 'A', 'B', 'D', 'S', 'Y'에 대해 0°에서 HMI(a)와 ZMM(b)의 특성비교

3. 특징비교: HMI 와 ZMM

많은 패턴 인식 시스템에서 같은 부류의 패턴을 인식할 때는 패턴의 조그만 차이에는 둔감해야하지만 다른 부류의 패턴일 경우에는 패턴사이의 차이에 민감해야 한다. 그림 1은 서로 다른 각도에 있는 문자 'A'로부터 추출한 7개의 HMI 와 8차까지인 25개의 ZMM를 보여 주고 있다. 그림에서 보이는 바와 같이

Ⅲ. 전체적인 시스템 동작

그림 3에 예시된 시스템의 전체적인 제어는 다섯 가지 연속적인 모듈 즉 임계화 모듈 (thresholding

module), 표식 모듈(labeling module), 정규화 모듈(normalization module), ZMM 계산 모듈 그리고 분류 모듈(classifier module)로 이루어져 있다. 이 장에서는 각 모듈에 대하여 간단히 설명한다.

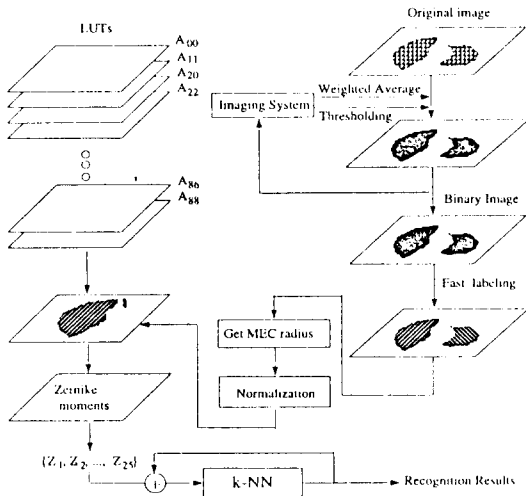


그림 3. 전체 시스템의 구성도와 흐름

1. 임계화(Thresholding)

임계 모듈은 다음의 두 가지 기본 모듈로 이루어졌다. 첫째는 간단히 전체 영상을 임계화하는 것이고,

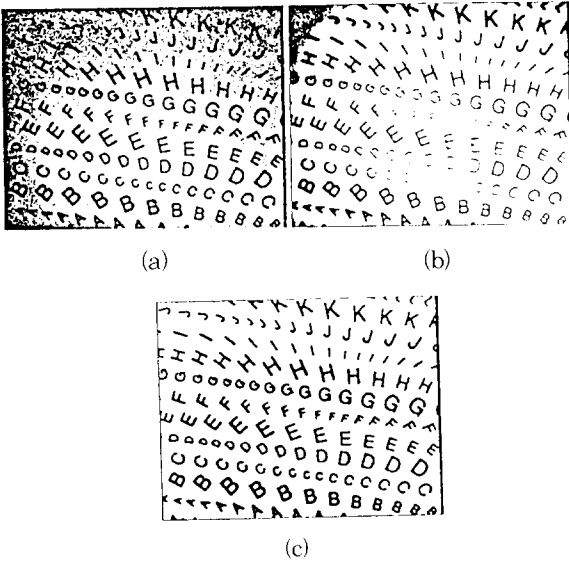


그림 4. 부적절한 조광에 의한 영상(a)과 전체적인 임계화 결과(b)와 LIG 알고리즘에 의한

다른 하나는 적응 임계화를 하는 것이다. 그림 4에서 나타난 것과 같이 영상의 조도가 고르지 못할 때에는 전체적인 임계화를 적용할 수 없기 때문에, 적응 임계화 알고리즘(LIG 알고리즘)을 사용하여 임계지도를 작성해서 부적절한 조광효과를 줄이는데 사용된다.^[11] 적응 임계화 공정은 알고리즘의 복잡성에 따른 시간 때문에 off-line으로 동작한다. 그러나 산업체에 적용할 때 조도 환경은 잘 변하지 않으므로 적응 임계화 모듈은 조명 환경이 변할 때만 사용된다. 임계지도는 LIG 알고리즘과 함께 생성되고 영상 시스템에 저장된다. 그리고 상이 포착될 때마다 최적의 이진영상을 만들기 위해 그 상으로부터 지도값을 뺀다.

2. 표식 (Labeling)

몇 가지 구성성분 표식 알고리즘(component labeling algorithm) 중에서 윤곽선 추적방법에 기초한 빠른 알고리즘을 구현하였다.^[12] 알고리즘은 먼저 물체들의 윤곽선을 추적하고, 그것들에 표식을 붙이고, 영상전체를 한번에 주사함으로써 물체의 위상 특성을 찾아낸다.

그 개념은 다음과 같이 기술된다. 영상의 모든 점들을 왼쪽 위부터 오른쪽으로 차례로 주사를 하여 화소가 물체에 해당하는지를 조사한다. 처음으로 물체를 이루고 있는 화소에 도달하게 되면 그 점은 물체의 제일 왼쪽 상단에 있게 되는데 이를 초기점(initial point:IP)이라고 부른다. 이때 일단 주사를 중단하고, 그 점에서 시작하여 이번에는 물체의 윤곽선에 있는 모든 점들을 차례로 따라가며 같은 영역으로 표시한다. 이 작업은 다시 IP에 도달 할 때까지 계속된다. 다시 IP에 도달하게 되면 중단하였던 주사를 계속하는데 이때 화소의 영역은 현재 위치에서 왼쪽에 있는 화소와 같게 한다. 다시 말하면 각 수평선을 따라서 같은 물체에 해당 하는 모든 화소들은 그 선에 있는 가장 왼쪽 점과 같은 영역에 해당하게 된다. 같은 방법으로 그 다음 선에서도 계속 반복되는데 물체의 왼쪽 외각에 있는 모든 점들은 같은 물체의 윤곽선에 해당되므로 결국은 물체 위에 있는 모든 화소가 같은 영역을 이루게 된다. 이 방법은 기존의 두 번 주사 방법에 비해 memory도 적게 차지할 뿐 아니라 속도 또한 실용적인 시스템에 쓸 수 있을 만큼 빠르다.

3. 校正 (Calibration)

Zernike 모멘트는 원안에서 정의되므로 회전불변

특성을 위해서는 화상의 종횡비(aspect ratio)가 시스템 성능에 많은 영향을 미치므로 자세히 추정될 필요가 있다. 이를 위해서 일련의 원들을 영상조절 목적에 사용하였다. 카메라의 종횡비 때문에 원이 digitize되었을 때 타원형으로 기록이 되기 때문에 먼저 타원들의 외곽선에 적합한 타원값들을 결정한다. 그로부터 실제 영상의 종횡비는 이전 비율들의 평균값으로부터 결정된다.^[13]

4. 정규화 (Normalization)

물체를 이진 영상으로 축척에 무관하게 인식하려면 물체의 상을 먼저 정규화해야 한다. 인쇄된 문자와 같은 물체의 경우 문자획의 폭은 사용된 임계치에 따라 변하며 여러 개의 화소만큼 넓어지거나 가늘어진다. 문자의 전체 면적의 변화는 문자가 전체 크기에 비해 상대적으로 가늘 때 심각해지고, ZOM에 의해 정규화된 결과 모멘트 특성은 신뢰성을 상실한다. 이러한 팽창 혹은 부식의 영향은 물체를 정확하게 초점이 맞지 않을 때 더욱 심각해진다. 이러한 결점을 보완하기 위해 본 시스템에서는 축척 불변을 위해 물체 주위에서 최대 확장원 (maximum enclosing circle: MEC)의 반지름이 계산된다.^[13] 그 후 물체는 정규화되고 직교 좌표계의 lookup-table을 생성하기 위하여 101×101 크기의 배열로 사상된다. 이러한 정규화 방법은 ZOM에 의한 방법^[9]보다 신뢰성 있는 축척 불변 Zernike모멘트를 생성한다. 또한 상이한 임계치에 의하여 계산된 모멘트 값에서도 변화가 최소화되기 때문에, 유사 임계 불변 특성을 나타낸다.

5. Zernike 모멘트의 계산

Zernike 모멘트를 계산하는 가장 간단한 방법중의 하나는 식(2)로부터 계산된다. 하지만 그 방법은 Hu의 모멘트 불변치들을 계산하는 것 보다 많은 계산을 해야 하는데 그 이유는 식(2)에 있는 인자들 때문이다. 그러나 본 시스템에서는 모멘트를 lookup-table을 이용한 영역에 근거한 방법에 따라 계산할 수 있다.

이 방법에서는 먼저 lookup-table의 크기를 결정한다. 101x101 화소 크기의 2개의 2차원 배열을 사용한다. 다음에 lookup-table안에 꼭 맞는 원의 반지름은 51 화소가 되고 그 원의 내부에 있는 모든 (x,y)좌표에 대해 상응하는 (R, θ 좌표를 계산한 다음 식(2)에 의하여 Zernike 근간함수 $V_{nm}(x,y)$ 을 계산하는데 실수부분과 허수부분을 각각 2개의 배열 상에

저장한다. 물체의 크기를 결정하는데 필요한 축척요소는 MEC 방법에 의해 결정된다. 여기서 MEC 방법이란 물체의 중심으로부터 가장 멀리 떨어진 윤곽점으로써 물체의 축척을 결정하는 것을 말한다. 물체의 중심을 lookup-table의 원점(중심)에 맞춘 후 물체의 모든 (x,y)좌표는 축척요소에 비례하여 계산되고, 양자화되어 lookup-table상으로 다시 그려진다. 식(2)을 만족시키는 m과 n의 값은 8차까지인 23개가 있으므로 전체 lookup-table의 수는 23쌍의 실수와 허수부 배열이 필요하다. 그리고 나서 Zernike 모멘트 A_{nm} 은 다음의 LUT 방법으로 계산된다.

```
AR = AI = 0.0
for(x = 0; x < MaxX; x++)
for(y = 0; y < MaxY; y++)
  if(ff[x][y] == ObjectRegion){
    AR = AR + f[x][y]*VR[x][y];
    AI = AI + f[x][y]*VI[x][y];
  }
A = sqrt(AR*AR + AI*AI)*(n+1)/3.1415926;
```

$f[x][y]$ 가 이진 영상일 때 곱셈은 생략되므로 Zernike 모멘트를 계산하는데 단순 함으로 되어 LUT연산은 계산속도를 크게 향상시킨다.

6. 패턴 분류

마지막 단계로서는 물체를 표시하는데 상응하는 일련의 Zernike 모멘트 크기가 미지의 표본벡터의 분류기 입력으로 다루어진다. k-최근방 이웃값(k-nearest neighbor:k-NN) 패턴 분류 방법은 분류된 표본들 사이에서 일련의 미지 표본의 모멘트를 분류하는데 이용된다.

미지의 표본과 미리 분류된 훈련 표본간 거리가 계산된다. 그 거리로는 Euclidean distance를 이용하는데 k개의 최근방 이웃값 중에서 다수를 차지하는 부류가 그 벡터군의 부류로 할당 된다. 즉, X의 k개 최근방 이웃값들중 $k_i(i=1, \dots, n; n>2)$ 개의 표본들이 i번째 부류에 속한다 하자. 그때 만약 $k_i = \max(k_1, k_2, \dots, k_n)$ 이라면 미지의 표본 X는 i번째 부류로 분류된다.

k-NN 알고리즘은 이웃값들중 다수 갯수에 의하여 미지의 표본을 분류하지만, 가중 k-NN 알고리즘^[6]에서는 그 거리에 따라 상이한 가중치가 i번째 부류에 할당된다. 다시 말하면 X에 가까운 표본일수록 일정

한 양의 가중치를 그 표본이 해당하는 부류에 더해준다. 이렇게 해서 차지하는 값이 최대가 되는 부류가 미지의 표본 X의 부류로 분류된다.

IV. 실험 결과

모든 공정이 그림 3에 예시되어 있고, 200개의 물체를 인식하는 과정에서 Sun4/370을 사용하였으며 실행시간은 약 10.0초가 걸렸다. 그림 5에서는 전형적인 실험 결과들을 나타낸다.

1. 표본 준비

본 시스템에서 사용한 방법들을 평가하기 위하여 다음의 실험들을 실행하였다. 첫째로 대문자와 소문자 alphabet 및 10개의 숫자와 기호(symbol)를 포함한 62개의 글자를 PostScript 언어를 사용하여 레이저 프린터에 인쇄한 후 512 512 화소의 해상도를 갖는 CCD 카메라를 사용하여 digitize하였다. 문자들 간의 각도는 15°이고 각 문자의 크기는 10에서 32 까지의 화소에 걸쳐 점진적으로 증가시키거나 감소시켰다.

2. 훈련

본 시스템을 훈련시키기 위하여 따로 표본을 준비하였다. 각 기호당 8개의 문자를 각각 30도 간격으로 회전시키고 크기도 32에서 64 화소까지 다양한 크기의 기호 별로 8개의 문자를 관찰하였다. 각 문자에 대하여 Zernike 모멘트가 추출되었다. 벡터들의 8개 집합은 k-NN 분류를 위하여 부류별로 한 덩어리를 이룬다.

3. 인식 결과

그림 5에서 영상은 상이한 각도와 크기의 200개 이상의 다른 문자들을 포함하고 있다. 모든 문자들을 인식하는 데 걸리는 총 실행 시간은 각 문자가 인식될 때 파일의 입출력과 모니터로의 전시를 포함하여 10초 이하가 소요되었다. 부분적으로 보이지 않는 경계면에 있는 몇 문자들을 제외하고는 모두 정확하게 인식되었다. 많은 표본 중이들을 임의의 방향에서 100번 이상 실험하였다. 굵힌 자국이 있는 작은 문자들에 대하여 몇 개의 잘못된 인식이 이루어졌지만 99.36%의 인식률을 나타내었다.

상이한 k-NN 알고리즘에 대한 결과가 비교되고

이를 표 2에 나타내었다. 이 표에서 k-NN 과정은 $k > 1$ 일 때 분류착오가 1/3로 줄었으며, $k = 1$ 인 경우에는 가장 간단한 최근방 이웃값 방법을 나타낸다. 본 실험을 통하여 가중치 있는 k-NN 방법의 성능은 오차율의 견지에서는 가중치 없는 k-NN 방법과 매우 비슷하였다.

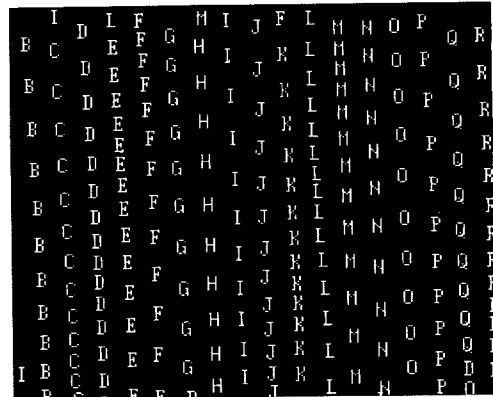


그림 5. Helvetica타입 문자에 대한 실험 결과: 원래의 영상(검은색 문자)과 인식결과(흰색 문자)를 동시에 나타내기 위해 바탕색을 회색으로 처리하였다.

표 2. 다른 k값에 대한 k-NN 분류기의 평균 오차

k	k = 1	k = 5	k = 9
Error Rate	1.89%	0.64%	0.64%

4. 임계 불변

본 실험에서는 시스템의 임계 불변 특성을 평가하였다. 그림 5에 나타난 것과 비슷한 표본에 대하여 세 개의 상이한 임계값을 수동적으로 골랐다. 최저 임계치는 130이고, 최적의 임계치는 165이고, 가장 큰 값은 200이다. 상이한 임계치에서 각 문자의 평균 면적은 표 3에 나타내었다. 최저 임계치에서 면적 변화율은 31%이고, 최고 임계치에서의 변화율은 40%

표 3. 임계치의 변화에 따른 문자의 영상에서의 화소의 평균수와 표준편차

Threshold	200	165	130
Average area	275.34	400.45	562.82
STD	17.74	18.84	21.84

인데 이것은 반지름의 변화가 최적 값의 1.2배에서 0.84배가 되는 것을 의미한다. MEC 방법을 이용하였을 때 인식율은 99% 이상이었고, ZOM 방법을 썼을 때에는 85%에 불과하였다.

5. 지도 해석

실제 영상에서 임계 불변을 예시하기 위하여 그림 6에서는 달라스 지역에서의 지도 영상인식의 결과를 나타내고 있다. 이 실험에서는 그림 5와는 달리 지도가 2진영상이 아닌 다진 영상이므로 임계치가 증가함

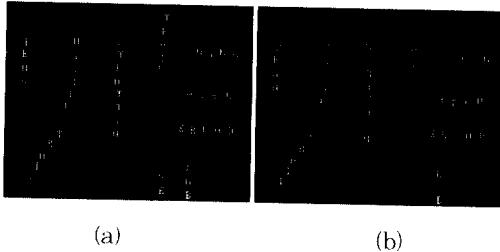


그림 6. 다진영상에 대한 상이한 임계치 162(a)와 102(b)에서 지도 인식 결과. 그림 5와 마찬가지로 원래의 영상(흑색 문자)과 인식 결과(흰색 문자)를 동시에 나타나기 위해 바탕색을 회색으로 나타냄

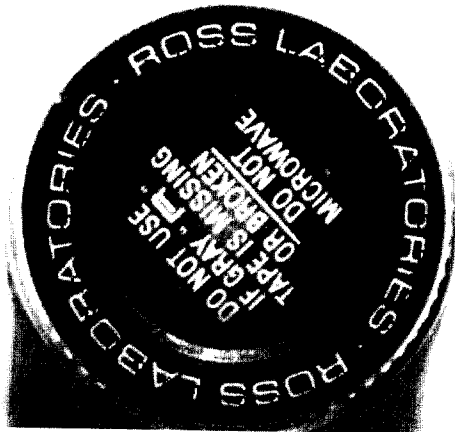


그림 7. 병 뚜껑에 대한 실험결과: 영상의 크기는 512x512이고 병 뚜껑이 임의의 각도로 놓여있다. 인식된 결과는 "ROSS LABORATORIES"로 원주위를 따라 검은 색으로 원래의 영상 위에 덮어 쓰여 있다.

에 따라 문자들은 더 두꺼워지고 몇 개는 도로와 연결되었다. 도로와 연결된 문자들은 도로의 부분으로 간주되었고, 시스템에 의하여 무시되었다. 두 그림에 나타난 백색 문자체를 사용하여 원래의 문자들에 대한 컴퓨터 해석 결과를 화면에 보여 주었고 두개의 인식율은 충분히 높다는 것을 보여준다.

6. 뚜껑 검색

본 실험에서 표지가 원둘레를 따라 쓰여진 병 뚜껑을 읽는 것을 시도하였다. 중앙에 있는 난잡하게 쓰여진 글씨 때문에 원 둘레에 쓰여진 표식만을 본 시스템이 읽을 수 있었다. 본 시스템은 사용된 문자체와 같은 형태의 문자체로 훈련이 되었다. 인식된 결과는 검은 색으로 덧씌워져 있다. 뚜껑이 어떤 각도로 회전하더라도 본 시스템은 모든 회전 방향에서 기호를 모두 인식한다.

V. 결론

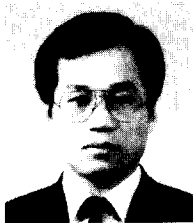
제조산업에 있어서 융통성이 있고 자동적인 조작과 검사를 위하여 기호나 패턴의 크기와 회전 방향에 대해 불변인 기호나 패턴을 인식하는 실용적인 기계 시각 시스템이 개발되었다. Zernike 모멘트는 정합 과정에서 초기 특징으로서 사용된다. 인식 과정의 속도를 높이기 위해 세 가지 공정으로서 적응 임계 지도를 가진 임계화, one-pass 물체 표식, Zernike 모멘트를 계산하기 위한 lookup-table등을 조사하였다. MEC 방법이 축적 불변에 대한 정규 과정에 사용될 경우, 인식율은 최악의 경우를 제외하고는 다양한 임계치에서 고정적으로 99% 이상 유지되었다.

參考文獻

- [1] M. Hatamian, "A real-time two-dimensional moment generating algorithm and its single chip implementation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, June 1986.

- [2] D. Casasent, L. Cheatham, and D. Fetterly, "Hybrid optical/digital moment-based robotic pattern recognition system," *Proc. SPIE*, vol. 360, pp. 105-111, 1982.
- [3] R. J. Prokop and A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP: Graphical Models and Image Processing*, vol. 54, pp. 438-460, September 1992.
- [4] M. R. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Am.*, vol. 70, August 1980.
- [5] C.-H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 10, July 1988.
- [6] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, pp. 179-187, February 1962.
- [7] R. Y. Wong, "Scene matching with invariant moments," *Computer Graphics and Image Processing*, no. 8, 1978.
- [8] S. A. Dudani, K. J. Breeding, and R. B. McGhee, "Aircraft identification by moment invariants," *IEEE Trans. on Computers*, vol. C-26, pp. 39-46, January 1977.
- [9] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 489-498, May 1990.
- [10] A. P. Reeves and R. W. Taylor, "Identification of three-dimensional objects using range information," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, April 1989.
- [11] J. Parker, "Grey level thresholding in badly illuminated images," *IEEE Transactions on PAMI*, vol. 13, August 1991.
- [12] H. Takahashi and F. Tomita, "Fast region labeling with boundary tracking," in *IEEE International Conference on Image Processing*, Sept. 1989.
- [13] W. Kim and P. Yuan, "Rotation and scale invariant recognition system using zernike moment," Tech. Rep. EE93-17, School of Engineering and Computer Science, University of Texas at Dallas, Richardson, TX 75083-0688, 1993. ㉔

筆者紹介



金 會 律

1956年 5月 27日生

1980年 한양대학교 전자공학과 졸업 (학사)

1983年 Pennsylvania 주립대학교 대학원 전기공학과 졸업 (M.S)

1989年 Purdue대학교 대학원 전기공학과 졸업 (Ph. D.)

1989年 ~ 1994年

Texas 대학교 조교수

1994年 ~ 현재

한양대학교 전자공학과 조교수

주관심 분야 : 영상압축, 2D 및 3D 컴퓨터 시각 시스템, 광영상처리등