Distogram의 이동, 회전, 크기변화에 대한 불변특성 테스트

손미숙0　　이기혁
한국정보통신대학원대학교 공학부
(mssohn0, geehyuk)@icu.ac.kr

The Translation- and Rotation-Invariant and Scale Covariant property of Distogram

Misook Sohn0 Geehyuk Lee
Dept. of Engineering, Information & Communications University

요 약

패턴의 이동, 회전, 크기변화에 영향을 받지 않는 불변 패턴 인식은 컴퓨터 비젼 및 패턴인식에서 오랫동안 다루어온 문제이다. 본 논문에서는 픽셀간의 거리를 히스토그램으로 나타냄으로써 2차원 이진 패턴을 1차원 신호로 표현하는 방법을 제시한다. 제시되는 방법을 디스토그램(Distogram)이라 부르며 6개 로마문자의 이동, 회전, 크기 변화된 이미지를 테스트해봄으로써 디스토그램의 식별 능력을 테스트해본다.

1. 서 론

불변(invariant) 패턴 인식은 컴퓨터 비젼 및 패턴인식 분야에서 오랫동안 다루어온 문제이다[1]. 만족할만한 수학적 모델이 없었기 때문에 인간의 생물학적 두뇌작용을 모방한 신경회로망을 이용한 방법이 많이 사용되었다[2][3][4][5][6][7]. 그러나 많은 경우 구조가 너무 복잡하거나 계산 시간이 오래 걸리는 문제가 있다. 본 논문에서는 이미지의 픽셀간의 거리의 히스토그램을 이용해 2차원 이진 이미지를 1차원 신호로 나타내는 새로운 방법을 제안한다. 이 개념은 거리(distance)의 히스토그램(histogram)이라 하여 디스토그램(distogram)이라 부른다. 디스토그램은 구조가 단순하고 계산이 간단한 방법이다.

2. DISTOGRAM

먼저 2차원 이진 이미지는 (b,x,y)로 정의한다. 이때 b는 0, 1로 이진 값을 나타내고 x,y는 픽셀 좌표를 나타낸다. 전체 이미지는 픽셀의 집합 $\{(b_i,x_i,y_i): i = 1,2,\ldots,I\}$으로 나타낼 수 있으며 이때 I는 픽셀의 총 개수이다. 따라서 이진 이미지의 디스토그램은 전경(foreground) 이미지의 픽셀간 거리 r의 상대적인 빈도를 나타낸다. 이미지의 히스토그램을 나타내기 위해 디스토그램 도메인 $h(r)$을 d의 간격으로 나눈다. H_k는 $b_i = b_j = 1$인 픽셀 쌍(i,j)의 개수이고 픽셀간 거리는

$$d(i,j) \square \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

로 측정되고 k번째 간격은 $(kd,(k+1)d)$이다. 정규화된(nomalization) 형태는 다음과 같이 정의된다.

$$h_k \circ H_k \sum_{k=1}^{\infty} H_k^{-1}$$

연속화된 형태(continuous form)는

$$h(r) \circ h_{\lfloor r/d \rfloor}$$

으로 나타내어진다. 이때 $\lfloor r/d \rfloor$은 x를 초과하지 않는 가장 큰 정수이다. 디스토그램은 이미지의 픽셀의 상대적인 위치로 표현되기 때문에 이동이나 회전불변의 특성을 갖는다. 그리고 거리 $d(i,j)$가 이미지 크기변화와 함께 변하므로 디스토그램은 이미지 크기 변화에 대하여 covariant한 특성을 가진다. 그림 1은 디스토그램의 정의를 나타내기 위해 3개의 기본적인 패턴의 디스토그램을 나타낸다. 원과 십자 형태의 디스토그램을 수식으로 표현하면

$$h^{(\mathrm{circle})}(r) = \frac{1}{p\sqrt{1 - r^2/4}}, \quad h^{(\mathrm{circle})}(r) = \frac{1}{p\sqrt{1 - r^2/4}}$$

과 같다. 그림 1c의 $h^{(\mathrm{cross})}(r)$은 선의 두께가 가우시안 분포를 따른다고 가정한 경우 위와 같이 유도된다.

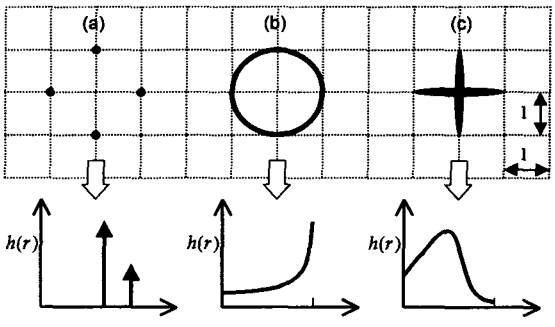

그림 1. 기본적인 패턴의 디스토그램 : (a) 코너에 있는 4개의 점 (b) 원 (c) 십자 모양

3. 실험

디스토그램의 식별 능력을 알아보기 위해서 이진 문자 G, E, H, Y, U, K 6개와 그것의 회전된 형태 G_r, E_r, H_r, Y_r, U_r, K_r와 크기 변화된 형태 G_s, E_s, H_s, Y_s, U_s, K_s의 18개 문자를 테스트 해보았다. 그림 2는 해당 테스트 문자의 디스토그램을 보여준다.

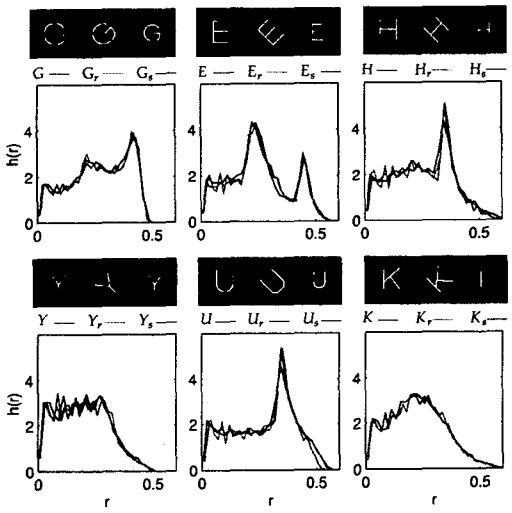

그림 2. 18개 테스트 문자의 디스토그램

그림 2에서 크기가 작아진 문자 패턴의 경우 디스토그램의 covariant한 특성을 보여주기 위해 크게 스케일링한다. 위의 테스트 결과는 디스토그램의 불변, covariant한 특성을 잘 보여주고 있으며 디스토그램이 이동, 회전, 크기 변화된 문자를 같은 문자로 식별함을 보여준다. 디스토그램의 식별능력을 평가하기 위해서는 좀더 많은 정량적인 실험이 필요하며 이동, 회전, 크기 변화된 디스토그램간의 거리를 측정해보는 것도 필요하다. 또한 문자의 임의 크기 변화에 상관없이 디스토그램을 재정규화(renormalize)하기 위해 다음과 같이 정의할 필요가 있다.

$$\hat{h}_p(s) = c_p h_p(c_p s)$$

c_p는 문자 p의 디스토그램의 center of mass이고 다음과 같이 정의된다.

$$c_p = \int_0^{*} r h_p(r)\,dr$$

재정규화된(renormalize) 디스토그램에서 문자 p와 q사이의 거리는 다음과 같이 계산된다.

$$d_{pq} = \int_0^{*} \left| \hat{h}_p(s) - h_q(s) \right| ds$$

그림 3a는 그림 2에서 보여 지는 6문자의 18가지 버전간의 거리 매트릭스(distance matrices)를 그레이 스케일로 표시한 것이다. 흰색은 문자와 문자사이의 거리의 최대 값을 나타내고 검은색은 최소 값을 나타낸다. 행과 열의 라벨은 문자의 인덱스를 나타낸다. 그림 3b는 $d = 0.17$에서 threshold한 후의 매트릭스를 표시한다. 그림 3b에서 검은색은 해당 문자 쌍을 같은 패턴으로 인식한 것을 의미한다.

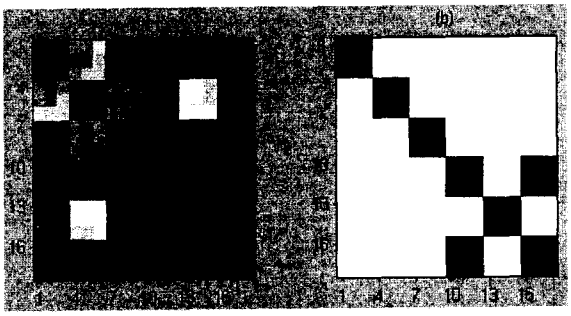

그림 3. (a) 거리 매트릭스 $d\left(h^{(i)}, h^{(j)}\right)$를 그레이 스케일로 표시한다. 흰색은 1 그리고 검은색은 0을 나타낸다. 위첨자 i, j는 테스트 문자의 순서를 나타내는 인덱스이며 순서는 다음과 같다. G, G_r, G_s, E, E_r, E_s, H, H_r, H_s, Y, Y_r, Y_s, U, U_r, U_s, K, K_r, K_s. (b) 는 (a)의 매트릭스를 $d = 0.17$에서 threshold 한 후의 매트릭스이다.

이 실험에서 디스토그램은 대부분의 경우 회전이나 크기변화에 상관없이 같은 문자로 인식했지만 Y 문자와 K문자는 구분하는데 실패했다. (매트릭스에서 3×3 의 크기만큼씩 검은색을 띄는 것은 원래 문자, 회전된 문자, 크기 변화된 문자들 사이의 거리차가 0에 가까우므로 같은 문자로 인식함을 보여준다. 그러나 Y문자와 K문자간의 경우에도 검은색 영역이 나타남으로써 두 문자를 구분함에 실패했음을 보여준다.) 디스토그램은 패턴의 모든 특징(feature)을 고려하는 것이 아니라 패턴의 구조적인 특징만을 고려한다. 따라서 구조적으로 닮은 형태인 Y와 K문자를 구분하는데 실패한 것이다.

대부분의 패턴인식에서는 불변 패턴 인식방법 한가지만을 사용하지 않고 두 가지 방법 혹은 여러 가지 방법을 복합적으로 사용한다. 디스토그램 역시 다른 불변 측정방법과 복합적으로 쓸 경우 모든 패턴을 정확하게 인식할 수 있을 것이다. 다음 실험에서는 디스토그램에 eccentricity를 추가하여 18개의 문자를 다시 테스트 해보았다[8].

패턴의 eccentricity는 패턴의 공분산행렬(covariance matrix)의 2개의 고유치(eigenvalue)의 비로 정의할 수 있다. 다음은 패턴의 eccentricity를 구하는 방법이다.

주어진 패턴 $\{(b_i, x_i, y_i): i = 1,2,...,I\}$의 center of mass를 다음과 같이 구한다.

$$\begin{bmatrix} m_x \\ m_y \end{bmatrix} = \frac{1}{I} \sum_{i=1..I, b=1} \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

패턴의 공분산 행렬은

$$C = \sum_{\substack{i=1,...,I \\ b=1}} \begin{bmatrix} (x_i - m_x)^2 & (x_i - m_x)(y_i - m_y) \\ (x_i - m_x)(y_i - m_y) & (y_i - m_y)^2 \end{bmatrix}$$

으로 나타낼 수 있다.

공분산 행렬의 2개의 고유치 값이 l_1과 l_2일 때 패턴의 eccentricity는 다음과 같이 나타내어진다.

$$e = \left| \frac{l_1}{l_2} \right|$$

이때 $|l_1| \le |l_2|$으로 가정하므로 eccentricity e는 $0 \# e \ 1$의 범위를 갖게 된다.

디스토그램과 eccentricity 측정법을 같이 사용해 패턴 p와 q사이의 거리를 d'_{pq}으로 재정의 한다.

$$d'_{pq} = \sqrt{d_{pq}^2 + e_{pq}^2}$$

이때 $e_{pq} = e_p - e_q$이고 e_p와 e_q는 패턴 p와 q의 eccentricity를 나타낸다. 그림 4는 재정의 된 거리 d'_{pq}를 사용해서 18개의 문자를 테스트한 결과를 보여준다. 예상한 바대로 디스토그램에 eccentricity를 추가한 경우 18개 테스트 문자 모두를 회전, 이동, 크기변화에 상관없이 정확하게 인식했다.

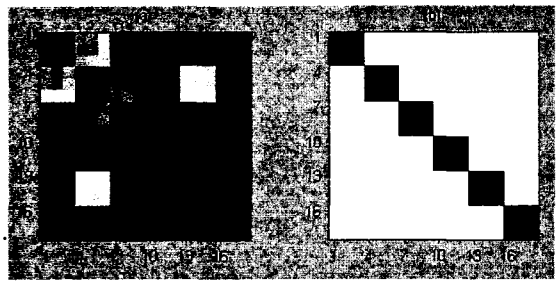

그림 4. d'을 0.18로 threshold한 후의 매트릭스 (b)의 경우 3×3 크기만큼 정확히 검은색으로 구분된다.

4. 결론

본 논문에서는 불변 패턴 인식 방법인 디스토그램에 대해 정의하고 디스토그램의 식별능력을 알아보기 위해 6개 로마문자의 회전, 이동, 크기 변화된 문자를 테스트해 보았다. 디스토그램은 패턴의 구조적인 특성을 고려한 방법으로 다른 불변 측정법과 복합적으로 사용 시 문자를 정확히 식별할 수 있었다. 디스토그램의 식별능력을 알아보기 위해서는 좀더 정량적인 실험이 필요하며, 지문인식이나 텍스쳐 식별, 도형의 가장자리 검출 등 다양한 패턴인식에로의 응용에 대한 연구가 진행 중이다.

5. 참고문헌

[1] R. Gonzalez and P. Wintz, Digital image processing, 2nd ed., Addison-Wesley, 1987.

[2] A. Khotanzad and J.-H. Lu, Classification of invariant image representations using a neural network, IEEE Trans. Acoustics, Speech, and Signal Processing, v. 38, pp. 1028-1038, 1990.

[3] S. J. Perantonis and P. J. G. Lisboa, Translation, rotation, and scale invariant pattern recognition by high-order neural networks and moment classifiers, IEEE Trans. Neural Networks, v. 3, pp. 241-251, 1992.

[4] B. A. Olshausen, C. H. Anderson, and D. C. van Essen, A multiscale dynamic routing circuit forforming size- and position-invariant object representations, J. Computational Neuroscience, v. 2, pp. 45-62, 1995.

[5] W.-G. Lin and S.-S. Wang, A new neural model for invariant pattern recognition, Neural Networks, v. 9, pp. 899-913, 1996.

[6] S. D. Kollias, A multiresolution neural network approach to invariant image recognition, Neurocomputing, v. 12, pp. 35-57, 1996.

[7] L. A. Torees-Mendez, J. C. Ruiz-Suarez, L. E. Sucar, and G. Gomez, Translation, rotation, and scale-invariant object recognition, IEEE Trans. Systems, Man, and Cybernetics, Part C, v. 30, pp. 125-130, 2000.

[8] C. Yuceer and K. Oflazer, A rotation, scaling, and translation invariant pattern classification system, Pattern Recognition, v. 26, pp. 687-710, 1993.