웨이브릿 국부 최대-최소값을 이용한 영상 정합

박철진°, 김주영, 고광식
경북대학교 전자공학과
전화 : 053) 950-5540

Image matching by Wavelet Local Extrema

Cheol-Jin Park°, Joo-Young Kim, Kwangsik Koh
Dept. of Electronics, Kyungpook National University
E-mail : dsd@ee.kyungpook.ac.kr

Abstract

Matching is a key problem in computer vision, image analysis and pattern recognition. In this paper a multiscale image matching algorithm by wavelet local extrema is proposed. This algorithm is based on the multiscale wavelet transform of the curvature which can utilize both the information of local extrema positions and magnitudes of transform results. This method has advantages in computational cost to a single scale image matching. It is also rotation-, translation-, and scale-independent image matching method. This matching can be used for the recognition of occluded objects.

I.서 론

영상 정합은 컴퓨터 시각 인식, 영상 해석, 형태 인식 등의 분야에서 많이 이용되고 있다. 영상 정합 방법은 크게 세 개의 계층으로 나누어지는데, 첫째, 영상에서의 픽셀 값(pixel value)의 상관 관계를 이용하는 방법, 둘째, 에지나 코너점 같은 특징점을 이용하는 방법, 셋째, 목적 영상의 일부를 이용하는 방법 등이 있다.

영상의 픽셀 값을 이용하는 방법은 영상간의 미세한 차이에도 아주 민감하므로, 영상 정합에서 처리량이 많은 단점이 있다. 목적 영상의 일부를 사용하는 방법에서는 이러한 변화에는 영향을 받지 않지만, 특징점을 찾는 자체가 쉽지 않다는 단점이 있다[1].

본 논문에서는 다층 스케일을 이용한 영상 정합 방법을 제안하였다. 제안한 방법은 영상의 윤곽선으로부터 곡률(curvature)을 추출하고, 웨이브릿 변환된 곡률로부터 스케일 맵을 구성하여 각 스케일 맵에서의 국부 최대-최소값(local extrema)의 크기 변화와 국부 최대-최소값간의 유클리드 거리(Euclidean distance)를 이용하여 영상 정합을 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 웨이브릿 변환에 대한 기본적인 이론과 국부 최대-최소값의 정의를 나타내었다. 3장에서는 제안한 영상 정합 방법에서 이용한 비용 함수(cost function)을 정의하였다. 4장에서는 정의한 비용함수를 이용하여 영상의 이동, 회전, 크기에 따른 정합 결과 및 겹쳐진 물체의 인식에 대한 결과를 나타내었다.

II.웨이브릿 변환

1.Wavelet Transform

웨이브릿 변환은 푸리에 변환(Fourier transform)과 같이 기저 함수(basis function)들의 집합에 의한 신호의 분해로 이해될 수 있다. 웨이브릿 변환에서의 기저 함수는 모 웨이브릿(mother wavelet)이라 하며, 웨이브릿 변환은 모 웨이브릿을 확장 또는 수축에 의해 얻어지는 웨이브릿 집합들을 이용하여 신호를 해석하게 된다[2].

각 스케일에서 웨이브릿은 모 웨이브릿 $\Psi(x)$을 스케일 변수 2^j만큼 확장한 형태로

$$\Psi_{2^j}(x) = \Psi(\frac{x}{2^j}) \ , \ \Phi_{2^j}(x) = \Phi(\frac{x}{2^j}) \qquad (1)$$

와 같이 표현한다.

$\Psi(x)$를 가우시안 함수와 같은 스무딩 함수의 1차 도함수로 보면

$$\Psi(x) = \frac{d\theta(x)}{dx} \qquad (2)$$

이고, s 스케일에서의 함수 $f(x)$의 웨이브릿 변환은

$$W_s f(x) = f * \Psi_s(x) = f * s\frac{d\theta_s(x)}{dx} = s\frac{d}{dx}(f * \theta_s)(x) \quad (3)$$

이다[2][3].

본 논문에서는 물체의 윤곽선을 방향공간에서 나타낸 각도 함수 f 는 다음과 같이 정의한다.

$$f = \phi(i) = \tan^{-1}\frac{y_{i+k} - y_{i-k}}{x_{i+k} - x_{i-k}} \quad (4)$$

여기에서 (x_i, y_i) 는 영상에서의 i 번째에서의 가로축 및 세로축의 위치이며, k는 스무딩 레벨로 정의한다.

각도 함수 $\phi(x)$를 웨이브릿 함수와 상승 적분을 취하면

$$W_s^a\phi(x) = \phi * \Psi_s^a(x) = \phi * (s\frac{d\theta_s}{dx})(x) = s\frac{d}{dx}(\phi * \theta_s)(x) \quad (5)$$

과 같이 나타낸다. 식 (5)에서는 각도 함수를 s 스케일에서의 스케일링 함수와 상승 적분하여 스무딩된 각도 함수를 미분한 것이 되므로, 결국 각도 함수의 변화율인 곡률을 구하게 된다[4].

2.웨이브릿 국부 최대-최소값

영상의 특징점은 일반적으로 영상 내에서 변화의 정도가 급격한 점으로 정의한다. 영상의 윤곽선으로부터 취한 곡률을 웨이브릿 변환한 경우에는 임의의 스케일에서의 웨이브릿 국부 최대-최소값이 변화가 가장 급격한 점이 되므로 주어진 영상에 대한 특징점으로 정의할 수 있다.

물체의 윤곽선을 최대한 분해 가능한 스케일까지 분해하면 $\{S_{2^J}^d\phi, (W_{2^j}^d\phi)_{, 1 \le j \le J}\}$와 같은 이산 집합들로 구성된다. 여기에서 J 는 분해 가능한 최대의 스케일이다.

임의의 j 스케일에서의 웨이브릿 국부 최대-최소값을 각각 $M_a(x)$, $M_i(x)$ 라 하면,

$$M_a(W_{2^j}^d\phi(x)) = \{ |W_{2^j}^d\phi(x+1)| \le |W_{2^j}^d\phi(x)|, \; |W_{2^j}^d\phi(x-1)| \le |W_{2^j}^d\phi(x)| \} \quad (6)$$

$$M_i(W_{2^j}^d\phi(x)) = \{ |W_{2^j}^d\phi(x)| \le |W_{2^j}^d\phi(x+1)|, \; |W_{2^j}^d\phi(x)| \le |W_{2^j}^d\phi(x-1)| \} \quad (7)$$

과 같이 정의된다. 여기에서 d 는 이산 신호를 의미한다.

3.특징점의 정보

스케일 맵에서의 특징점에 관한 정보는 인접한 특징점간의 거리와 특징점의 크기로 나눌 수 있다.

먼저 인접한 특징점간의 거리에 대해서 웨이브릿 계수의 개수를 c 라 정의한다. 스케일 맵 상에서 임의의 j 스케일에서 나타나는 픽셀에 영향을 주는 $j-1$ 스케일에서의 픽셀의 개수 d 는

$$d = c \cdot 2^{j-1} - (2^{j-1} - 1) \quad (8)$$

이다. 식 (7)은 j 스케일에서 나타나는 픽셀에 대하여 $j-1$ 스케일에서 나타나는 픽셀에 관한 관계식일 뿐만 아니라, 인접한 특징점간의 거리에 관한 식이기도 하다. 즉, 특징점간의 거리가 d 보다 크다면, j 스케일에서 특징점의 성질을 유지하므로 특징점으로 판단한다. 만약 특징점간의 거리가 d보다 작거나 같다면, $j-1$ 스케일에서는 인접한 두 특징점을 각각의 특징점으로 판단하지만, j 스케일에서는 인접한 두 특징점을 동일한 특징점으로 판단하게 된다.

다음으로 스케일 맵에서의 특징점의 크기 정보와 스케일과의 관계는 다음과 같다. 스케일이 증가할수록 웨이브릿 변환된 곡률의 스케일 맵은 점점 더 스무드하게 된다. 따라서 스케일이 증가할수록 웨이브릿 변환된 곡률의 값의 크기는 점점 감소하게 된다. 그러므로, 임의의 스케일 이상에서는 국부 최대-최소값이 존재하더라도 유용한 정보라고 판단하기 어렵다.

III.영상 정합

제안한 방법에서는 임의의 스케일에서의 웨이브릿 국부 최대-최소값에 대한 다음의 두 가지 정보를 가지고 두 영상의 정합 정도를 판단한다.

첫번째 정보는 웨이브릿 국부 최대-최소값의 크기이다. 최대-최소값의 크기는 영상의 형태에 대한 상세한 정보를 알려 준다.

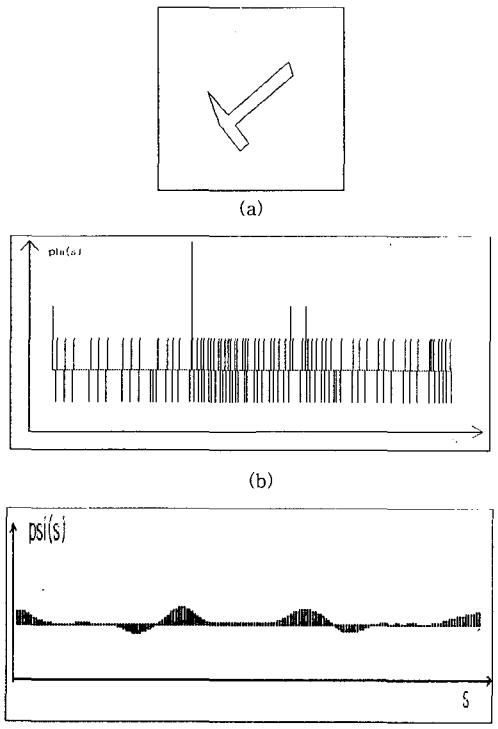
두번째 정보는 국부 최대-최소값간의 유클리드 거리(Euclidean distance)로서, 영상의 형태에 대한 개략적인 정보를 가지고 있다. 영상의 크기가 다른 경우에도 거리를 정규화함으로써 좀 더 간단하게 정합을 구할 수 있다.

제안한 방법에서의 국부 최대-최소값의 크기와 거리 정보를 이용한 정합 과정은 다음과 같다.

첫째, 입력 데이터에서 영상의 윤곽선을 추출한 다음, 윤곽선을 이용하여 곡률을 구하여, 웨이브릿 변환된 결과에서 스케일 맵을 구한다.

둘째, 스케일 맵 상에서의 국부 최대-최소값을 구한다.

셋째, 국부 최대-최소값의 정보를 가지고 비용 함수를 이용하여 정합을 판단한다.



(a)

(b)

(c)

그림 1. 영상 정합 과정
(a) 원 영상(256×256) (b) 곡률
(c) 스케일 맵(스케일 = 5)
Fig. 1. The matching procedure
(a) Original image(256×256)
(b) Curvature
(c) Scale map(scale = 5)

여기에서 사용한 비용함수는

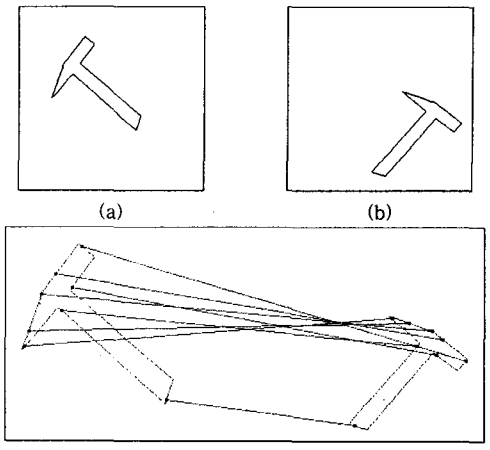
$$C_s(P_r,P_o) = \min_{all\ i} \sum_{k=0}^{k=N_o} c_s(i,J(k))$$

$$\begin{aligned} where \quad P_r^s &= \{ p_i^s \mid i = 0,1,\ldots, N_r^s \}, \\ P_o^s &= \{ p_k^s \mid k = 0,1,\ldots, N_o^s \}, \\ c(i,J(k)) &= d(i,J(k)) * v(i,J(k)) \end{aligned} \tag{9}$$

와 같다. 여기에서 $J(k)$는 p_i 에 대하여 p_k 를 대응시키는 함수이고, $d(i,J(k))$는 매핑되는 국부 최대-최소값에서의 정규화된 거리의 차이 함수이고, $v(i,J(k))$는 매핑되는 국부 최대-최소값의 차이 함수이다. 그림 1은 제안한 영상 정합 과정을 나타낸다.

IV.실 험

제안한 방법의 성능을 알아 보기 위하여 다음과 같이 실험하였다. 먼저 이동, 회전 및 크기가 다른 영상에 대한 정합 실험을 한다. 여기에서, 목적 영상은 크기가 원 영상의 70%이고, 시계 방향으로 90° 회전시켰으며, 이동된 정도는 가로축으로 50픽셀, 세로축으로 50픽셀 단위로 이동시켰다. 그리고, 스케일은 5를 사용하였다. 그림 2를 통해 정합이 잘 이루어짐을 알 수 있다.



(a) (b)

(c)

그림 2. 이동, 회전 및 크기가 다른 물체와의 정합
(a) 기준 영상(256×256) (b) 목적 영상
(c) 실제 영상에서의 정합
Fig. 2. Matching with translated, rotated and different scaled image
(a) reference image(256×256) (b) object image
(c) Matching between reference image and object image

다음으로, 겹쳐진 물체의 경우, 실험에서는 20 % 정도 겹친 물체를 가지고, 스케일 5에서의 정합 실험을 한다. 그림 3을 통해 겹쳐진 물체의 경우에도 비교적 정확하게 인식함을 알 수 있다.

V.결 론

본 논문에서는 웨이브릿 국부 최대-최소값을 이용한 영상 정합 방법을 제안하였다. 제안한 방법은 입력 영상의 윤곽선으로부터 곡률을 추출하고, 웨이브릿 변환된 곡률로부터 스케일 맵을 구성하여 각 스케일 맵에서의 최대-최소값의 크기 변화 및 유클리드 거리를 이용한다. 또한, 영상의 곡률 정보를 이용하기 때문에 이동, 회전 및 크기에 무관하며, 다층 스케일에 의해 정합점을 줄일 수 있는 장점이 있다.

따라서 겹쳐진 물체의 인식이나 고속 정합 등에 유용하게 사용될 것으로 생각된다.

참고문헌

[1] G. Borgefors, "Hiearchical chamfer matching: a parametric edge matching algorithm", IEEE Trans. *Pattern Anal. Machine Intell.*, vol. PAMI-10, no. 6, pp. 849-865, 1988.

[2] S. Mallat, S. Zhong, "Characterization of Signals from Multiscale Edges", IEEE Trans. *Pattern Anal. Machine Intell.*, vol. PAMI-14, no. 7, pp. 710-732, 1992.

[3] S. Mallat, W.L. Hwang, "Singularity Detection and Processing with Wavelets", IEEE *Inform Theory.*, vol. IT-38, no. 2, pp. 617-643, 1992.

[4] Jian-Shu Lee, Yung-Nien Sun, Chin-Hsing Chen and Ching-Tsorng Tsai, "Wavelet based corner detection", *Pattern Recogn.*, vol. 26, no. 6 pp. 853-865, 1993

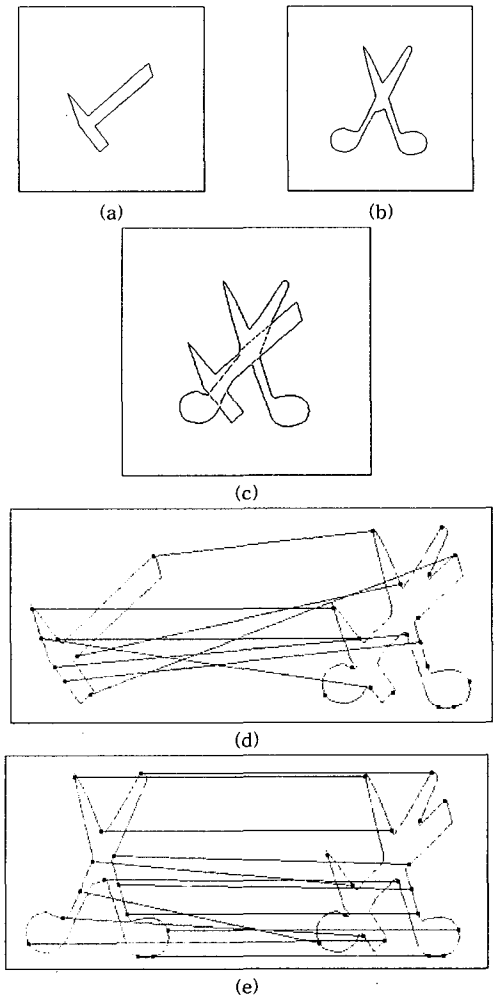


그림 3. 겹쳐진 물체의 인식
(a) 기준 영상 1 (b) 기준 영상 2 (c) 겹쳐진 물체
(d) 기준 영상 1과의 정합
(e) 기준 영상 2와의 정합

Fig. 3. Recognition of the occluded image
(a) reference image 1 (b) reference image 2
(c) occluded objects
(d) Matching reference image 1 with occluded objects
(e) Matching reference image 2 with occluded objects