

깊이 에지 기반의 Curvature Scale Space Map을 이용한 손 제스처 인식

이창주*, 이준호**

*성균관대학교 컴퓨터공학과

**성균관대학교 컴퓨터공학과

e-mail:{itisfine,jhyi}@skku.edu

Hand Gesture Recognition Using Curvature Scale Space Map of Depth Edges

Chang-Ju Yi*, June-Ho Yi**

*Dept of Computer Engineering, SungKyunKwan University

**Dept of Computer Engineering, SungKyunKwan University

요 약

본 연구는 구조광 기반의 깊이 에지를 이용하여 조명의 변화와 복잡한 배경에 상관없이 손 제스처의 외곽선 영상을 안정적으로 획득하였고, 제스처 영상을 표현하기 위하여 Curvature Scale Space(CSS) map을 이용하였다. 기존의 CSS map은 외곽선 영상의 깊은 굴곡과 완만한 굴곡을 잘 구분하지 못하는 문제점이 있었으나, 본 연구에서는 이러한 문제점을 분석하고, 이를 개선하기 위해서 각도 좌표를 이용한 CSS map 생성 방법을 제안하였다. 실험을 통해서 제안한 방법이 기존의 CSS map보다 우수한 인식 성능이 있음을 보였다.

1. 서론

컴퓨터를 조작하는 방법에 있어서 기존의 한정된 입력 장치에 의존하는 사용자 인터페이스가 아닌, 인간이 기본적으로 의사 전달 수단으로 사용하는 방법을 이용한 인터페이스에 대한 요구가 매우 증가되고 있다. 그 중 제스처는 일상생활에서 사람과 대화를 하거나 감정을 전달하는 등 상호작용을 하는데 있어서 매우 중요한 요소 중에 하나이며, 또한 컴퓨터를 조작하는 인터페이스에 적용하는데 있어서도 쉽고 자연스러운 표현으로 상호작용이 가능하다.

제스처 인식은 손 영역을 획득하는 과정과 획득한 제스처 영상을 인식하는 과정으로 크게 구분된다. 일반적으로 손 영역을 획득하는 방법은 손의 피부 색상을 이용하여 색상 기반의 검출시스템이 사용된다. 이 방법은 조명의 변화에 매우 불안정하고 복잡한 배경 환경에서 성능이 크게 떨어지는 점 등의 단점이 있다. 본 연구에서는 안정적으로 손 영역을 획득하고, 획득한 영상을 이용하여 효율적이고 정확한 방법으로 표현하고 매칭하는 과정을 통해 제스처를

인식하는 기법을 제안한다.

2. 구조광 기반의 깊이 에지를 이용한 제스처 외곽선의 검출

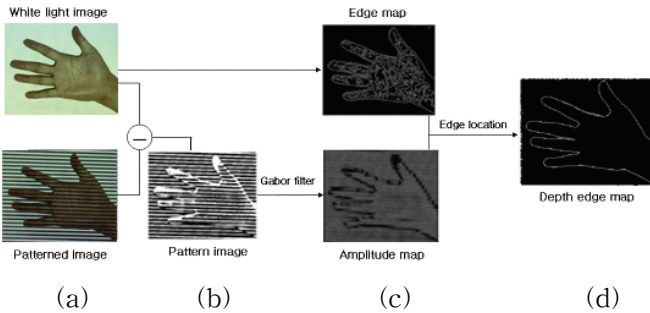
본 연구에서는 안정적인 제스처의 외곽선을 검출하기 위하여 깊이 에지 영상을 이용하였다. 물체에 패턴 광을 투영하고 카메라를 통해 영상을 획득해보면 깊이가 불연속적인 지점에서 패턴의 왜곡이 발생하며, 이러한 지점을 검출함으로써 깊이 에지를 찾을 수 있다[3]. 패턴 영상의 왜곡지점 검출에는 2D 가보 필터를 사용할 수 있는데 가보 필터는 복잡한 텍스처들을 효과적으로 분리하는데 매우 유용한 방법으로서 알려져 있다[4].

물체에 일정한 주기를 갖는 패턴을 투영하고, 그때 생기는 패턴의 왜곡 지점을 가보 필터로 검출함으로써 깊이 에지를 얻는 방법이다. 가보 필터는 다음과 같이 가우시안 함수에 의해 변조된 방향성을 갖는 사인 곡선의 함수로 정의 된다.

$$G_{\sigma,\phi,\theta}(x,y) = g_{\sigma}(x,y) \cdot \exp[2\pi j\phi(x\cos\theta + y\sin\theta)] \quad (1)$$

$g_\sigma = \frac{1}{2\pi\sigma^2} \exp[-(x^2 + y^2)/2\sigma^2]$, 가보함수의 인자 σ, ϕ , θ 는 각각 크기, 주파수, 방향을 의미한다. 가보 필터는 주기가 변하는 지점 즉, 패턴이 왜곡된 지점에서 낮은 앰플리튜드 값을 나타내는 특성이 있다[1].

제스처 영상의 깊이 에지를 검출하기 위해 (그림 1)(a)에서와 같이 검출하고자 하는 곳에 일반 프로젝터를 사용하여 연속적으로 백색광과 패턴광을 투영하고 그 영상을 획득한다. 그리고 카메라에 맺힌 백색광 영상과 패턴광 영상의 차영상을 계산함으로써 가로줄 패턴만을 추출한 영상을 얻게 된다((그림 1)(b) 참조). (그림 1)의 세 번째 단계인 (c)에서는 패턴영상과 백색광 영상의 에지 정보로부터 깊이 에지 맵을 생성하게 된다. 우선, 패턴영상에서 깊이 에지를 따라 가로줄 패턴이 어긋나는 왜곡 발생 지점을 찾게 된다. 패턴은 일정한 주파수를 갖는 주기 신호로 볼 수 있으며, 획득한 패턴영상에 2D 가보 필터(gabor filter)를 적용함으로써 필터 응답의 앰플리튜드가 달라지는 왜곡 발생 지점을 찾을 수 있다[4]. 마지막으로 백색광 영상의 에지맵을 이용하여 깊이 에지의 위치를 정확하게 결정하게 되면 (그림 1) (d)와 같은 깊이 에지 맵을 얻게 된다.



(그림 1) 패턴영상에서의 패턴 왜곡 발생위치와 백색광 영상의 에지정보로부터 깊이 에지 맵을 구하는 과정: (a) 백색광 영상과 패턴광 영상의 획득, (b) 가로줄 패턴만을 추출한 패턴영상, (c) 가보필터 앰플리튜드 맵과 백색광 영상의 에지 맵을 참조한 깊이 에지 검출, (d) 최종 깊이 에지 맵

3. Curvature Scale Space map

CSS map은 물체의 외곽선의 각 점에서 곡률의 부호가 바뀌는 점, 즉 극점을 가우시안 스케일 스페이스에 표현함으로써 물체의 외곽선에서 나타난 굴곡의 특징을 표현한다.

윤곽선 Γ 는 윤곽선의 길이를 나타내는 파라미터 u 에 의해 다음과 같이 표현 할 수 있다.

$$\Gamma(u) = (x(u), y(u)) \quad (2)$$

곡선의 각 점에 대한 곡률은 호의 길이에 대한 탄젠트 각도의 미분으로 정의 된다.

$$\kappa(u) = \lim_{h \rightarrow 0} \frac{\phi}{h} \quad (3)$$

$\vec{t}(u)$ 는 점 $\Gamma(u)$ 에서의 탄젠트 벡터, ϕ 는 $\vec{t}(u)$ 와 $\vec{t}(u+h)$ 사이의 각도를 나타낸다. 식 (3)의 정의를 이용하면 곡선 $\Gamma(u)$ 의 곡률은 다음과 같이 계산된다 [2].

$$\kappa(u) = \frac{\dot{x}(u)\ddot{y}(u) - \ddot{x}(u)\dot{y}(u)}{(\dot{x}^2(u) + \dot{y}^2(u))^{3/2}} \quad (4)$$

윤곽선 Γ 를 가우시안 함수 $g(u, \sigma)$ 를 사용하여 평탄화(smoothing)과정을 거친 윤곽선을 다음과 같이 표현 한다.

$$\Gamma_\sigma = ((X(u, \sigma), Y(u, \sigma))) \quad (5)$$

$$X(u, \sigma) = x(u) * g(u, \sigma) \quad (6)$$

$$Y(u, \sigma) = y(u) * g(u, \sigma) \quad (7)$$

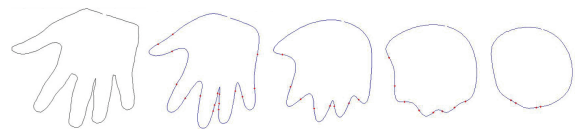
평탄화 된 윤곽선 Γ_σ 의 곡률은 식 (4)를 이용하여 다음과 같이 계산된다.

$$\kappa(u, \sigma) = \frac{X_u(u, \sigma)Y_{uu}(u, \sigma) - X_{uu}(u, \sigma)Y_u(u, \sigma)}{(X_u(u, \sigma)^2 + Y_u(u, \sigma)^2)^{3/2}} \quad (8)$$

CSS map $I(u, \sigma)$ 은 다음과 같이 정의된다.

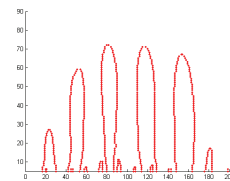
$$I(u, \sigma) = \{(u, \sigma) | \kappa(u, \sigma) = 0\} \quad (9)$$

σ 가 증가함에 따라, 외곽선 Γ_σ 의 모양은 (그림 2)(a)~(e)처럼 점점 평탄화된다. 이 변화의 과정에서 각각의 외곽선 Γ_σ 의 모든 점들의 곡률을 계산하고, 곡률의 부호가 바뀌는 지점 즉 극점의 위치를 가로축은 외곽선 호의 길이, 세로축은 σ 를 나타내는 평면의 (u, σ) 위치에 점으로 표시하여 나타내면 (그림 3)과 같은 CSS map을 얻는다.



(a)원영상 (b) $\sigma=10$ (c) $\sigma=30$ (d) $\sigma=50$ (e) $\sigma=70$

(그림 2) 각 σ 에서 평탄화된 윤곽선.



(그림 3) (그림 2)(a)영상의 CSS map.

(그림 2)에서 보듯 가우시안 함수의 σ 가 증가 할수록 물체의 외곽선은 평탄화 되어 원의 형태에 가까워 지게 되고 (그림 3)의 CSS map에서 두 개로

표현되던 곡률의 극점은 점점 가까워 지다 한 점으로 합쳐지는데 이 점을 극대점이라고 정의 한다. CSS map은 최종적으로 다음과 같이 극대점의 집합으로 나타내며, 이것을 이용해 외곽선의 모양을 기술한다.

$$\{(u_1, \sigma_1), (u_2, \sigma_2), \dots, (u_N, \sigma_N)\} \quad (10)$$

CSS map의 극대점 집합 A와 B에 대해 두 영상의 거리를 계산하는 방법은 다음과 같다

CSS map의 극대점 집합 A :

$$\{(u_1^A, \sigma_1^A), (u_2^A, \sigma_2^A), \dots, (u_N^A, \sigma_N^A)\} (\sigma_1^A > \sigma_2^A > \dots > \sigma_N^A) \quad (11)$$

CSS map의 극대점 집합 B :

$$\{(u_1^B, \sigma_1^B), (u_2^B, \sigma_2^B), \dots, (u_K^B, \sigma_K^B)\} (\sigma_1^B > \sigma_2^B > \dots > \sigma_K^B) \quad (12)$$

A의 극대점과 B의 극대점간의 대응쌍의 직선 거리가 일정 기준치 이하이면 서로 매칭된 극대점이라고 간주하고, CSS map 거리는 다음 식과 같이 매칭된 극대점 쌍의 유클리디안 거리와 매칭되지 않은 극대점의 σ 의 합으로 계산된다.

$$CSS_거리(A, B) = \sum_{i, j \in \text{매칭된 극대점}} \sqrt{(u_i^A - u_j^B)^2 + (\sigma_i^A - \sigma_j^B)^2} + \sum_{i \in \text{매칭안된 극대점}} \sigma_i^A + \sum_{j \in \text{매칭안된 극대점}} \sigma_j^B \quad i \in \{1, 2, \dots, N\}, j \in \{1, 2, \dots, K\} \quad (13)$$

4. CSSmap의 완만한굴곡표현의 문제점

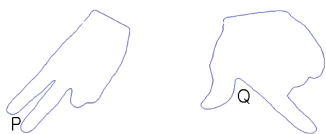
CSS map의 주요한 문제점 중의 하나는 곡률이 큰 깊은 굴곡과 곡률이 작은 완만한 굴곡의 구분이 어려운 점이다. 이번 장에서는 이런 현상이 발생하는 이유에 대해서 기술한다.

가우시안 함수의 평탄화 효과는 곡률변형의 일종인 기하학적 열평형 효과를 가진다[5].

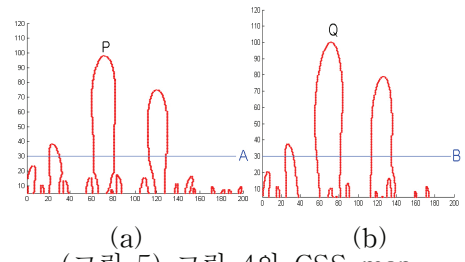
$$\frac{\partial \Gamma}{\partial (\sigma/2)} = \kappa N \quad (14)$$

이 식에서 Γ 는 물체의 윤곽선 벡터, N 은 노말 벡터를 나타내며, 가우시안 평탄화 효과에 의한 한 점의 위치변화는 곡률의 크기에 비례함을 보여준다.

그 결과 두개의 다른 모양의 제스처 영상인 (그림 4)는 (그림 5)의 CSS map에서 보듯 모두 비슷한 위치에서 극대점이 발생하게 된다. 이것은 굴곡이 발생한 위치에 대해서는 구분되지만 굴곡이 얼마나 깊고, 완만한지는 반영하지 못하므로 기존 CSS map의 극대점을 매칭하여 거리를 계산하는 방법으로는 구분하기 어려운 모호성이 발생하게 된다.



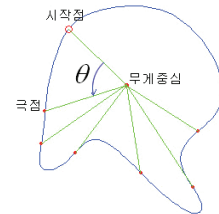
(그림 4) 제스처의 외곽선 영상



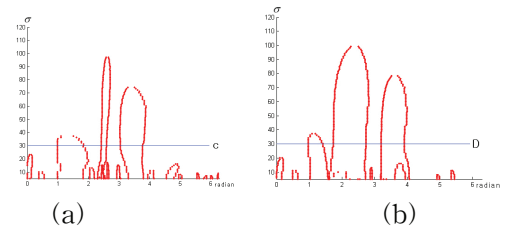
(그림 5) 그림 4의 CSS map

5. 각도좌표를 이용한 CSSmap 표현

이번 장에서는 깊은 굴곡과 얇은 굴곡을 잘 구분하지 못하는 기존의 CSS map의 문제점을 개선하기 위해 각도 좌표를 이용한 CSS map 생성 방법을 제안한다. 이것은 (그림 6)에서 보듯이, 물체의 무게중심을 중심으로 시작점과, 곡률의 극점이 발생하는 점이 이루는 각도(θ)를 계산하여 나타냄으로써 각도좌표를 이용한 CSS map을 표현하였다.



(그림 6) 외곽선에서 극점의 각도를 구하는 과정 (그림 5)에서 보듯 기존 CSS map 방법으로는 (그림 4)영상의 구분이 거의 힘들다. 그러나 (그림 7)에서 각도 좌표를 이용한 CSS map을 보면 극대점이 발생하는 위치는 비슷하지만, 극점으로 이루어진 곡선(이것을 포물선이라고 정의한다.)의 너비가 큰 차이를 보임을 확인할 수 있다. 이것은 $\sigma=30$ 일 때 곡률의 극점의 위치를 나타내는 (그림 5)의 A, B선분, (그림 7)의 C, D선분에서 확인할 수 있듯이, 외곽선상에서는 같은 위치에 극점이 발생하더라도 각도 좌표상에서는 굴곡의 정도에 따라 극점의 발생 위치가 영향을 받게 된다. 이러한 점은 기존 CSS map에서 같은 위치에서 극점이 발생함으로 생기는 모호성을 해결할 수 있는 단서가 된다.



(그림 7) (그림 4)의 각도좌표를 이용한 CSS map
각도좌표를 이용한 CSS map의 표현은 극대점의 좌표로 기술 되는 기존의 CSS map의 표현방법(식

(10)에 극대점의 포물선의 넓이(S)가 추가적으로 포함되어 아래식과 같이 기술된다.

$$\{(u_1, \sigma_1, S_1), (u_2, \sigma_2, S_2), \dots, (u_N, \sigma_N, S_N)\} \quad (15)$$

각도 좌표를 이용한 CSS map에서 거리차이를 계산하는 방법은 다음과 같다.

$$\text{각도좌표_CSS거리}(A,B)=$$

$$\sum_{i,j \in \text{매칭된극대점}} (\sqrt{((u_i^A - s) - u_j^B)^2 + (\sigma_i^A - \sigma_j^B)^2} + |S_i - S_j|) + \sum_{i \in \text{매칭안된극대점}} \sigma_i^A + \sum_{j \in \text{매칭안된극대점}} \sigma_j^B$$

$i \in \{1, 2, \dots, N\}, j \in \{1, 2, \dots, K\}$, s는 시프트매개변수(16) 각도좌표를 이용한 CSS map의 거리를 계산하는 방법에서는 매칭된 극대점의 포물선의 넓이 차 $|S_i - S_j|$ 를 추가적으로 더해 주는데, 이로 인해 CSS map에서 대응되는 극대점의 위치가 비슷하더라도 포물선의 넓이의 차가 크면 그만큼 굴곡의 깊이 차이가 크다고 볼 수 있으며, 기존의 CSS map의 모호성을 감소시키는 효과가 있다.

6. 실험 결과

본 연구에서는 깊이 예지 영상으로부터 얻어진 제스처 영상을 CSS map, 각도좌표를 이용한 CSS map, Zernike 모멘트, ART를 이용하여 인식 성능을 비교 평가하였다. 모델 영상은 (그림 8)의 9개의 제스처 영상을 사용하였으며 가로 200픽셀, 세로 200픽셀의 크기를 가진다.

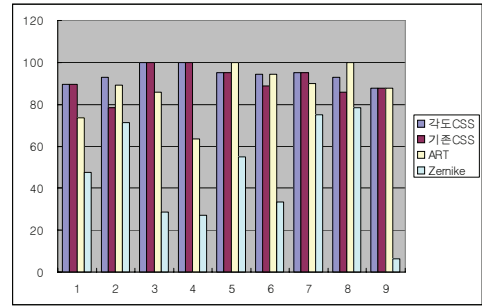


(그림 8) 모델 영상

입력 데이터 영상은 5명의 사용자로부터 9개의 다른 모양의 제스처 160장의 이미지를 얻었으며 영상의 크기와 회전된 각도는 각각 다르게 구성되어 있다. 실험을 수행한 결과 각각의 인식률은 <표 1>, (그림 9)와 같은 결과를 얻었다.

<표 1> 입력영상의 전체에 대한 인식률

방법	인식률
각도 좌표를 이용한 CSS map	93%
기존의 CSS map	90%
ART	88%
Zernike 모멘트	50%



(그림 9) 입력영상의 모델 별 인식률 그래프

실험 결과 각도 좌표를 이용한 CSS map의 성능이 기존의 CSS map과 Zernike 모멘트, ART를 이용한 방법보다 정확함을 알 수 있다. 특히 2번 모델에서 기존의 CSS map을 이용하였을 때 6번 모델과 잘못 매칭이 되는 경우가 발생하기도 하였으나, 제안한 각도좌표를 이용한 CSS map의 경우는 이러한 문제점을 보완하여 보다 높은 인식 성능을 보였다.

7. 결론

본 연구에서는 손의 모양을 추출하기 위해서 구조광을 기반으로 하는 깊이 예지 검출 방법을 이용하였으며, 검출된 외곽선은 CSS map을 이용하여 표현, 매칭하였다. CSS map은 곡률의 부호가 바뀌는 점 즉 외곽선의 굴곡의 특징을 이용하였으나 굴곡의 정도를 잘 표현 하지 못하므로 일부 영상을 잘 구분하지 못하는 문제점이 있었다. 본 연구에서는 각도 좌표를 이용한 CSS map을 이용함으로써 이 문제를 개선하였으며 실험 결과로 확인하였다.

참고문헌

[1] J. Park, C. Kim, J. Yi and M. Turk, "Efficient Depth Edge Detection Using Structured Light", ISVC, LNCS Vol. 2869, pp. 413-420, 2005.
 [2] F. Mokhtarian, "Silhouette-based isolated object recognition through curvature scale space", IEEE Trans. on PAMI, Vol. 17, No. 5, pp. 539-544, 1995.
 [3] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments", IEEE Trans. on PAMI, Vol.12, No.5, pp. 489-498, 1990.
 [4] A. C. Bovik, M. Clark and W. S. Geisler, "Multichannel Texture Analysis Using Localized Spatial Filters", IEEE Trans. on PAMI, Vol. 12, No. 1, pp. 55-73, 1990.
 [5] B. B. Kimia and K. Siddiqi, "Geometric heat equation and nonlinear diffusion of shapes and images", CVIU, Vol. 64, No. 3, pp. 305-332, 1996.