

MRF를 이용한 수화 동영상에서의 효율적인 손 형상 추출

송 효 섭⁰, 양 윤 모

고 려 대 학 교 전 자 정 보 공 학 과

{hssong, ymyang}@hard.korea.ac.kr

The Extraction of the Shape of Hands in the Sign Language Sequence by using MRF Model

Hyo-Seob Song⁰, Yun-Mo Yang

Dept. of Electronics & Information Eng., Korea University

요 약

영상 처리를 통한 수화(手話)의 인식에 있어 가장 중요한 정보는 손의 형상, 위치, 이동방향 등을 들 수 있다. 이 중 손의 형상은 세가지 정보 중 가장 중요하며, 실제로 자음과 모음, 숫자 등을 나타내는 지문자의 경우 손의 형상만으로도 인식될 수 있다. 본 논문에서는 선 처리 모델(Line Process Model)을 3차원으로 확장하여 적용한 Markov Random Field(MRF)를 사용하여 효율적으로 손의 형상을 추출하였다.

1. 서론

영상 분할을 위해 사용되는 방법에는 여러 가지가 있지만, 확률론적으로 접근하는 방법인 MRF를 이용하는 방법은 다른 방법에 비해 많은 계산량과 구현하기가 어려운 단점에도 불구하고, 프로세스 모델을 결합하여 동시에 여러 가지 파라미터를 추출할 수 있다는 장점에 영상복원과 분할, 움직임 예측 등에 많이 사용되고 있다.^[1,2,3,5] 기존에 연구되었던 optical flow에 의한 손의 형상 추출 방법^[4]과 키프레임의 시간 방향 에지 정보를 이용한 손의 형상 추출 방법^[6]은 손 영역에 대한 윤곽을 추출하는 데에는 뛰어났지만 손의 정확한 형상을 추출하는 데에는 미흡했기 때문에, 본 논문에서는 MRF 모델에 Konrad와 Dubois가 제안한 선 처리 모델^[1]을 3차원으로 확장하여 한국어 수화 동영상 인식의 전처리 단계로서 보다 효율적으로 정확한 손의 형상을 추출하였다. 3차원으로 확장된 선 처리 모델은 움직임을 파악할 수 있는 정보를 담고 있기 때문에 동영상에서 움직이는 영역을 추출하기에 적합하다. 실험은 한국어 수화 동영상 중 자음과 모음에 대한 동영상과 양손 수화에 기본이 되는 x, y, z 방향 양손 움직임에 대한 동영상을 대상으로 하여 손의 형상을 추출하였다.

2. Markov Random Field

MRF는 1차원 causal Markov chain의 2차원 확장이며, 영상의 모델링이나 처리 등에 유용하게 사용되고 있다.^[1,2,3,5] MRF는 local conditional pdf를 통해 정의할 수 있다.

랜덤 필드 $z = \{z(x), x \in A\}$ 가 다음 식 (1)과 식 (2)를 만족하면 이웃 N_x 에 대한 MRF라 한다.

$$p(z) > 0, \text{ for all } z \quad (1)$$

$$\begin{aligned} p(z(x_i) | z(x_j), \forall x_j \neq x_i) \\ = p(z(x_i) | z(x_j), x_j \in N_x) \end{aligned} \quad (2)$$

그러나 local conditional pdf를 사용하여 MRF를 정의하는 것은 어렵기 때문에 실제로는 Gibbs 분포를 이용하여 MRF와 GRF(Gibbs Random Field)의 동가성을 설명하여 정의한다. N을 이웃 시스템이라 가정할 때, N에 대해서 $z(x)$ 가 MRF이면 N에 대한 $p(z)$ 는 Gibbs 분포를 따르며, 역도 성립한다는 Hammersley-Clifford 이론^[2]에 의해 GRF와 MRF는 동가가 된다.

3. 선 처리 모델^[1,2,3,5]

선 처리 모델은 영상 내의 불연속성과 영상 간의 불연속성을 표현하는데 유용하게 사용되며, joint probability

분포로 모델화될 수 있다.

$$p(l | g_{k-1}) = \frac{1}{Q_l} \exp\{-U_l(l | g_{k-1}) / \beta_l\} \quad (3)$$

여기서 l 은 선 처리 모델, g_{k-1} 는 입력 영상, β 는 threshold, Q_l 은 분할 함수이다.

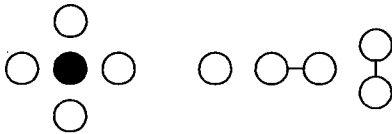
또한 입력 영상에 대한 선 처리 모델의 Gibbs potential $U_l(l | g_{k-1})$ 은 식 (4)와 같다.

$$U_l(l | g_{k-1}) = \sum_{c \in C_l} V_l^c(l | g_{k-1}) \quad (4)$$

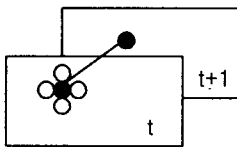
여기서 C_l 은 line field 를 위한 Clique 의 모든 집합이며, $V_l^c(\cdot)$ 는 $c \in C_l$ 인 Clique 함수를 나타내고 있다.

Clique 함수 $V_l^c(\cdot)$ 를 이용하여 움직임의 불연속성을 ON과 OFF 두 가지 상태로 분류하여 불연속이 되는 지점을 구분할 수 있다.

다음 [그림 1] 은 실험에 사용된 선 처리 모델에 적용하기 위한 Clique 설정이다. Clique 는 1 차 시스템(4근방)을 사용하였고, 동영상에 적용시키기 위해 clique potential은 시간 t 와 $t+1$ 과 같이 시간차가 나는 두 장의 영상에서 clique 내부에 존재하는 이웃하고 있는 픽셀 간의 차를 사용하여 계산하였다.



(가) Clique set C_l



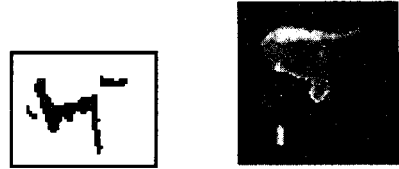
(나) clique potential 계산

[그림 1] Clique set C_l 과 clique potential

4. 실험 결과

실험은 영상 획득 장치로 Matrox Meteor-II와 PULNIX TMC-7DSP CCD 카메라를 사용하여 30fps로 샘플링된 흑백 동영상을 사용하였으며, 얻어진 한국어 수화 지문자 동영상과 양손이 x, y, z 방향으로 움직인 시험 영상에 MRF를 적용하였다. 우선 기존에 연구되었던 키프레임의 시간 방향 에지 정보를 이용한 손의 윤곽 추출 결과와

optical flow를 이용한 손의 형상 추출 결과를 살펴보면 [그림 2]와 같다.



(가) 시간방향에지 방법 (나) optical flow 방법

[그림 2] 기존의 손 영역 추출에 대한 연구 결과

시간 방향 에지를 이용하는 방법은 차영상에 sobel 연산자와 같은 역할을 하는 mask를 적용하여 얻은 결과로서 손의 움직임 윤곽은 정확히 나타나지만, 윤곽만으로 형상을 추출하기에는 부족하다. Optical flow를 이용하는 방법은 경계면에서 flow 벡터의 불연속성을 이용한 방법으로 움직이는 물체에 대한 대강의 영역을 추출해 낼 수 있었지만 정확한 손의 형상을 추출하지는 못했다.

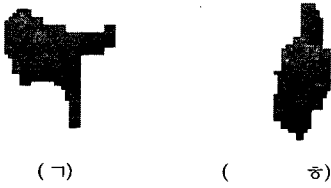
본 논문에서는 기존에 연구 결과를 개선시키기 위해 MRF 모델과 선 처리 모델을 결합한 MRF 모델의 두 가지 방법으로 실험하였다. 우선 선 처리 모델을 결합하지 않은 MRF를 적용하였을 경우 추출된 손의 영상을 살펴보면 다음 [그림 3]과 같다.



[그림 3] MRF 적용 후 추출된 손의 형상(ㅎ)

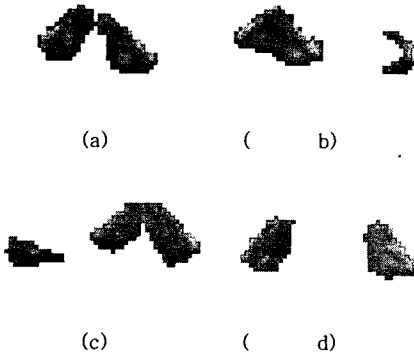
[그림 3]에서 볼 수 있듯이 손의 형상은 잘 나타나지만 움직임이 있는 다른 부분들이 같이 추출되는 문제가 발생하였다. 이러한 문제를 해결하기 위해 단순한 clique potential 만이 아닌 선 처리 모델을 결합하여 손의 형상을 추출하였다.

다음 [그림 4]는 한국어 수화 지문자 동영상에 대해 선 처리 모델을 적용시킨 후 움직임이 많이 발생된 영역에 대해 레이블링하여 1위 영역만 추출한 결과이다.

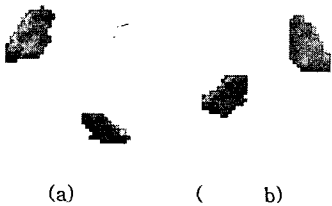


[그림 4] 손 형상 추출 결과

다음 [그림 5]는 양손 기본 움직임 동영상에 선 처리 모델을 적용시킨 후 움직임이 많이 발생하는 부분에 대해 레이블링하여 2위 영역까지 추출한 손의 형상이다.



(가) x 방향(좌우) 움직임 (a, b, c, d 순)



(나) y 방향(상하) 움직임 (a, b 순)



(다) z 방향(전후) 움직임 (a, b, c, d 순)

[그림 5] 손 추출 형상 결과(양손 수화 기본 움직임)

[그림 5]에서 x 방향 움직임 결과를 살펴보면 양손이 서로 떨어져 있는 경우는 문제없이 레이블링이 잘 되었다. 그러나, 양손이 겹쳐있는 경우에는 하나로 인식되어 손 이외의 영역이 같이 추출되는 결과가 나왔다.

참고로, [그림 4] 및 [그림 5]의 결과 영상이 모자이크와 같이 추출된 것은 원 영상의 작은 형상을 확대하였기 때문이다.

5. 결론 및 향후 방향

본 논문에서는 흑백 한국어 수화 지문자 동영상과 양손 수화에 기본이 될 수 있는 x, y, z 방향 움직임 동영상에 대상으로 하여 선 처리 모델과 결합된 MRF를 적용하여 손의 형상을 추출하였다. 지문자와 같이 한 손 움직임의 경우 그 형상이 잘 추출되었다. 양손 추출의 경우 손이 겹치는 경우(occlusion)에는 손 이외의 다른 영역이 추출되는 문제에 대한 해결책과 겹쳐있는 손을 분리시키는 방법에 대한 연구가 필요하다. 또한 이동 방향을 검출하기 위한 움직임 처리 모델(motion process model)에 대한 연구와 나아가 수화를 인식하기 위한 방법에 대한 연구도 필요하다.

6. 참고 문헌

- [1] Janusz Konrad and Eric Dubois, "Bayesian Estimation of Motion Vector Fields", *IEEE trans. PAMI*, vol. 14, pp. 910~927, Sep. 1992.
- [2] A. Murat Tekalp, *Digital Video Processing*, Prentice Hall, 1995.
- [3] S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Springer, 1995
- [4] 송효섭, 양윤모, "Optical Flow를 이용한 손의 움직임 및 영역 추출", *98 정보과학회 봄 학술발표논문집*, 25권, 1호, pp. 642~644, 1998년 4월
- [5] André Redert, Emile Hendriks, and Jan Biemond, "Correspondence Estimation in Image Pairs", *IEEE Signal Processing Magazine*, pp. 29~46, May 1999
- [6] 송효섭, 장세진, 신병주, 양윤모, "손의 형상과 움직임 방향 정보를 이용한 수화 인식", *정보과학회논문지(B)*, 26권, 6호, pp. 804~810, 1999년 6월