

## 엔트로피 분석을 이용한 수화 영상에서의 손의 추적

이종실, 한영환\*, 민홍기\*\*, 홍승홍

인하대학교 전자공학과, \*상지대학교 전산학과, \*\*인천대학교 정보통신과

### Tracking of Human Hands from a Sign-Language Image Sequence using Entropy Analysis

Jongshill Lee, Younghwan Han\*, Hongki Min\*\*, Seunghong Hong

Dept. of Electronic Eng., Inha Univ., \*Dept. of Computer Science, Sangji Univ.

\*\*Dept. of Info. and Telecomm. Eng., Incheon Univ.

<http://biselab.ee.inha.ac.kr>, E-mail : netlee@netian.com

#### 요약

본 논문은 연속적인 수화 동영상에 있어서 가장 의미가 있는 손 영역을 검출하는 방법에 대한 연구이다. 부자연스러운 센서 등을 피하고 사용자 친화성과 범용성 및 실시간 구현을 도모하기 위하여 흑백 동영상의 처리가 요구된다. 농아자의 수화를 인식하기 위해 진처리로서, 복잡한 배경으로부터 손 영역의 검출은 매우 중요하다. 정확한 손 영역의 검출로 수화 인식률의 개선과 영상처리의 고속화가 가능하다.

제안한 방법은 인접한 프레임간의 차 영상에 대하여 엔트로피를 측정하여 배경영상과 분리를 행한 후, 손 영역에 대한 추적을 수행하였다.

#### 서론

의사전달에 있어서 언어, 즉 문자와 말의 중요성과 더불어 사람의 손짓, 몸짓, 얼굴 표정 등을 이해하여 상대방의 의사를 인식하려는 지적 인터페이스 연구가 활발히 진행되고 있으며, 구체적으로 얼굴인식, 표정인식, 시선인식, 입술의 움직임, 몸짓 및 손짓인식 등이 있다[1].

사람의 의사표현 수단으로서 손짓은, 그 표현의 다양성으로 인하여 손짓만으로도 언어를 이

루고 있다. 손짓만으로 이루어진 언어 즉, 수화는 농아자들의 중요한 의사전달체계로, 특히 선천적인 농아자의 경우에는 높은 문맹률로 인하여, 수화가 유일한 의사전달 수단이 되기도 한다. 국내 인구 중 수화사용자는 35만 농아자와 그 가족, 친지를 포함하여 100만인으로 추산하고 있다. 또한 병원이나 관공시 등에서 농아자가 개인적인 문제를 다룰 때는 수화 통역자를 통하지 않고 직접해야하는 경우가 많으며, 긴급시 의사전달, 컴퓨터를 통한 수화교육 및 장기적으로 수화에 의한 전화통화 등의 이유로 인해 수화인식의 필요성이 높아지고 있다[2].

지금까지의 제스처와 수화 등의 인식에 관한 연구는 대개 Data Glove<sup>TM</sup>를 사용하였다[3-5]. 데이터 글러브를 사용한 방법은 장갑에 센서를 부착하여 손동작의 3차원 공간성의 정보를 실시간으로 입력받아 사용하므로 비교적 용이하게 손짓의 분석과 인식에 이용하고 있다. 그러나 이 기구는 인식 대상자 또는 장애자에게 부자유스러움과 심리적 부담을 주게 되므로 사용자 친화성(親和性)을 지향하는 지적 인터페이스와는 거리가 멀다.

이에 반해, 영상처리에 의한 손동작의 이해는 인식 대상자의 동작을 카메라 영상을 통하여 인식하기 때문에 부자유스러운 제약을 요구하지 않는다. 그러나 지금까지의 영상처리에 의한 몸

짓 및 손짓인식 연구는, 표식(Marker)에 의하여 손가락이나 손의 형상을 추출하거나 칼라에 의하여 손을 추출하는 방법을 사용함으로써 옷의 색깔이나 피부색에 강한 제약을 두고 있다.

이들 방법들은 표식의 모양이나 색깔, 배경의 명암과 색깔에 따른 제약점이 생길 수 있다. 그리고, RGB 칼라 카메라를 이용하여 화면의 영상내 피부색이나 특정색깔의 장갑을 착용하여 장갑의 색깔만을 추적함으로써 손의 이동 궤적이나 손의 모양을 인식하는 연구가 활발하게 진행되어 오고 있다[6]. 그러나, 배경에 특정색깔이 중복되거나 동일한 피부색인 얼굴과 손이 겹쳐질 경우 손동작의 추적이 어려워지는 문제점이 있다.

따라서 본 연구에서는 부자연스러운 센서 등을 피하고 사용자 친화성과 범용성 및 실시간 구현을 도모하기 위하여 흑백 수화 동영상을 사용하며, 엔트로피를 분석하여 복잡한 배경영상으로부터 손영역을 분리(segmentation)하는 방법을 제안하며, 손영역을 추적한다.

본론

일반적인 흑백 카메라에서도 수화 동영상을 안정적으로 인식하기 위하여 프레임간의 움직임을 추출하기 위하여 연속되는 동영상에 대하여 차 영상을 구하며, 주로 DB 검색에 있어서 응용되던 PIM(Picture Information Measure)방법 중 Chang's PIM 방법을 개선하여 차영상에 적용하여 가장 중요시되는 손 부위를 추적한다.

이때, PIM의 평균값과 분산을 이용하여 손 부위를 검출해 낸다.

손 영역을 검출 및 추적을 위한 알고리즘은 그림1과 같다.

1. PIM에 기초한 움직임 검출

화상 정보 시스템(pictorial information system)을 설계할 때 흔히 주어진 영상의 정보 내용을 측정할 필요가 생긴다. 영상에 정보량이 적으면 보통 이 정보는 저장되지 않거나 압축되어 저장되어야 한다. 화상 통신의 경우를 생각해 보면, 정보가 적은 영상은 원래 형태로 전송해서는 안되며 압축을 해서 전송이 이루어져야 한다. 따라서 영상의 정보량을 결정해주는 측정

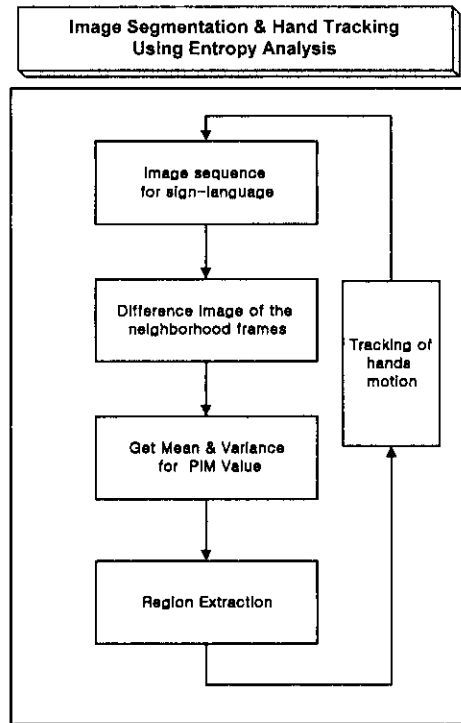


그림 1. 수화 영상에 대한 손 영역 검출 및 추적

Fig. 1. Segmentation and Tracking of hand region for gesture of sign-language

기구를 갖는 것이 중요한 것이다.

본 연구에서는 영상 데이터의 특징을 얻기 위해서 화소간의 엔트로피를 이용하고, 영상에서 엔트로피를 정량화 하기 위해 Chang이 제안한 PIM[ref1]을 도입한다. 본 논문에서 사용한 PIM(Picture Information Measure)을 수식으로 표현하면 다음과 같다.

$$PIM = \sum_{i=0}^{L-1} h(i) - \text{Max}_i h(i) \quad \text{----- (1)}$$

여기서 h(i)는 각 영상 또는 해당 블록의 히스토그램(histogram)을 의미한다. PIM값은 해당 블록의 전체 화소수와 히스토그램 중 가장 큰 도수의 히스토그램 값과의 차이로 계산된다. 블록 내의 화소값들이 모두 같은 경우, 즉 블록의 엔트로피가 '0'인 경우에는 식 (2)와 같이

$$\sum_{i=0}^{L-1} h(i) = \text{Max}_i h(i) \quad \text{----- (2)}$$

되어 PIM=0의 최소값을 갖는다. 또한 블록 내에서 화소의 각 레벨 값들이 균일하게 분포하는 경우 즉, 엔트로피가 큰 경우에는 Max h(i)이 작은 값을 갖기 때문에 큰 PIM값을 갖게 된다. PIM의 최대값은 알고리즘 1에 의해 구해진다.

For M×N size, L-gray level image  
 if M×N≤L,  
 then Max(PIM)=M×N-1  
 if M×N>L,  
 then Max(PIM)=M×N-[M×N/L]  
 ([n] : n을 넘지 않는 최대의 정수)

[ 알고리즘 1 ]

즉, PIM은 해당 블록이 많은 정보를 갖는 경우에는 큰 값을 얻을 수 있고, 적은 정보량을 갖는 경우에는 작은 값을 얻을 수 있다. 또한 정규화 PIM(Normalized PIM)은 식(3)와 같이 정의되고, 일반적으로 PIM<sub>k</sub>, NPIM<sub>k</sub>는 각각 식(4), 식(5)과 같이 정의된다.

$$\text{NPIM} = 1 - \text{Max}_p(h(i)) \quad \text{----- (3)}$$

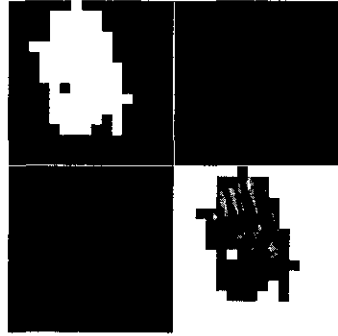
$$\text{PIM}_k = \sum_{i=0}^{L-1} h(i) - \sum_{i=0}^k h(i) \quad \text{----- (4)}$$

$$\text{NPIM}_k = 1 - \sum_{i=0}^k p(i) \quad \text{----- (5)}$$

그림 2는 복잡한 배경영상이 존재하는 영상에 대하여 차영상을 구하여 이 차영상에 PIM을 적용하여 손 영역만을 복잡한 배경으로부터 분리(segmentation)를 보여준다.



(a) Frame 1 (b) Frame 2



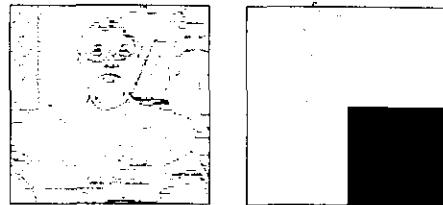
(c) 결과

그림 2. PIM을 이용한 손영역 추출

Fig. 2. Extraction of Hand Region using PIM



(a) 시간 t에서의 (b) t+Δt에서의



(c) (a),(b)두 frame간의 차영상

(d) 결과

그림 3. 노이즈에 강함을 보여주는 예  
 Fig. 3. Examples of Robustness to Noise

그림 3은 움직임이 작은 연속된 두 프레임간의 차에 대해 제안한 알고리즘을 적용한 예이다. 그림 3-(d)에서 보면 3-(c) 차영상에 대해 어떠한 움직임도 검출하지 않음을 알 수 있다. 즉, 제안한 알고리즘은 노이즈에 강함을 알 수 있다.

2. 손 영역 추적

본 연구에서는 그림 1에서 제안한 알고리즘을 실제 연속적인 손 움직임 영상에 대해 손의 위치를 추적해 가는 방법을 사용해 수화인식에서 사용 가능성을 보인다.

그림 4-(a)는 입력되는 영상을 나타내며 (b)는 입력된 영상에 대해 제안한 알고리즘을 이용하여 검출된 손 부위를 보여주고 있다.

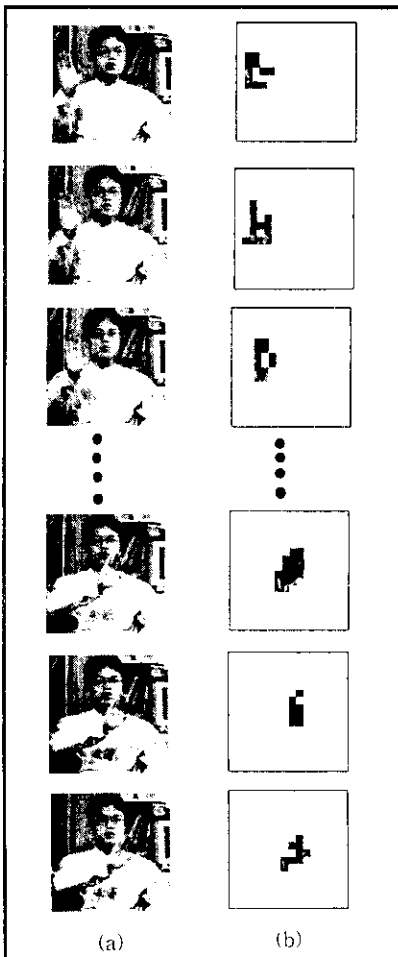


그림 4. 제안한 알고리즘을 이용한 손 영역 추적  
Fig. 4. Tracking of hand region using the proposed method.

결론 및 추후방향

제안한 알고리즘은 실제 차영상만을 이용하여 이진화를 행하는 것보다 손 영역 검출에 있어 손 내부의 노이즈 제거가 가장 큰 특징이다. 실험에서 보면 복잡한 배경영상은 사라짐을 알 수 있다. 배경만을 제거하고 중요시되는 부위를 추출할 수 있으므로 검출된 영역 주위류 기반으로 하여 정교한 알고리즘이 요구된다.

실험에서는 8\*8블록으로 분할을 행하였으나 좀 더 경계부분을 심세하게 분할하기 위해서는 가변 블록 적용이 요구된다. 그리고 검출된 영역에 대해 정교한 검출을 거친 후 HMM을 이용하여 수화 단어를 인식할 것이다.

참고문헌

[1] T.S. Hunang and V.I. Pavloic, "Hand Gesture Modeling, Analysis, and Synthesis, Proc. of International Workshop on Automatic Face-and Gesture-Recognition, pp.73-79, Zurich, June 1995.

[2] 최창우, 최운영, "수화 인식 및 동영상 생성의 연구동향", 전자공학회지, 제 23권, 제 6호, 1996.6, pp.672-681

[3] Jong-Sung Kim, Gyntae Park, Zeungnam Bien, Wonjang, "A Dynamic Pattern Recognition System for the Korean Sign Language(KSL)", Proc. of the Asian Control Conference, Tokyo, pp.713-716, July 27-30, 1994

[4] Hirochiko Sagawa et al., "Sign Language Translation System Using Continuous DP Matching", Tech. Report of IEICE SP 92-38, 1992

[5] Tomoichi Takahashi et al., "A Hand Gesture Recognition Method and Its Application", IEICE D-II, Vol. J73-D-II, No. 12, pp.1985-1992, Dec. 1990

[6] 이현규, 김호연, 김진형, "은닉 마르코프 모델을 이용한 연속 제스처 Spotting", 제 9 회 영상처리 및 이해에 관한 워크샵, pp. 114-119, 1997

[7] Shi-Kuo Chang, *Principles of Pictorial Information Systems Design*, Prentice-Hall, 1989.