

적외선 센서를 사용한 손 동작 인식

*안준영 **이상화 ***조남익

서울대학교 INMC 전기정보공학부

*sohwa360@ispl.snu.ac.kr

Hand Pose and Gesture Recognition Using Infrared Sensor

*Ahn, Joon-young **Lee, Sang-hwa ***Cho, Nam-ik

INMC, School of Electrical and Computer Engineering

Seoul National University

요약

최근 IT기술 영역에서 미래기술로 촉망받는 증강현실(AR)과 가상현실(VR)환경을 구축함에 있어서, 마우스나 키보드 등의 별도 장치 없이 기기에 원하는 동작을 입력 하도록 하는 NUI(Natural User Interface)기술이 각광받고 있다. 또한 NUI를 구현 하는데 중요한 기술 중 하나로 손동작 인식 기술, 얼굴 인식 기술 등이 대두되고 있다.

이에 본 논문은 적외선 센서의 일종인 Leapmotion 센서를 사용하여 손동작 인식을 구현하고자 하였다. 첫 번째로 우선 거리변환 행렬을 사용하여 손바닥의 중심을 찾았다. 이후 각각의 손가락을 convex hull 알고리즘을 사용하여 추출한다. 제안한 알고리즘에서는 손가락, 손바닥 부분의 optical flow를 구한 후, 두 optical flow의 특성을 사용하여 손의 이동, 정지, 클릭 동작 을 구분 할 수 있도록 하였다.

1. 서론

손 자세 인식 및 손동작 인식은 마우스나 키보드 등 일반적인 입력 장치 없이 기기와 사용자 간 상호작용을 가능하게 하는 NUI(Natural User Interface)의 중요한 기술로 주목받고 있으며, NUI 환경이 꼭 필요한 증강현실(AR)과 가상현실(VR)세계를 구축하는 데 있어서 빼놓을 수 없는 기술이다. 특정 센서(적외선 센서, RGBD 카메라 등)로부터 여러 가지 구조의 입력 영상을 받아 손 영역을 추출하고, 손가락이나 손바닥의 위치와 모양을 통해 다양한 동작을 인식하여 사용자 인터페이스와 기기-사용자 간 상호작용에 적용시키려는 연구가 많이 진행되어 왔다. 그러나 아직도 실제 환경에서 사용하기에는 센서의 연동 문제나 동작 속도 등에서 다소 미흡한 수준이다.

손 자세 및 손동작 인식 시스템을 구현하는 전통적인 방법에는 크게 3가지 방법이 있다. 첫 번째는 입력 영상에서 피부색을 검출하여 손 영역을 찾아 낸 후, 해당 손 영역에서 손가락과 손바닥을 추출 하는 방법이다[3,4]. 두 번째는 손가락 끝 혹은 손 전체에 색깔이 있는 장갑을 끼거나 고무처럼 생긴 장비를 착용하여 손 영역과 손가락을 인식하는 방법이 있다[5]. 하지만 이런 방법을 사용 할 경우 밝기 변화나 그림자에 취약하고, 특정 장비를 사용해야하기 때문에 손동작 인식을 할 수 있는 상황이 상당히 제한 될 수밖에 없다는 단점이 있다.

마지막 세 번째는 RGBD(RGB-Depth)카메라의 영상을 입력 영상

으로 받아서 RGBD 카메라의 깊이(depth) 정보를 사용하는 방법이다. 깊이 이미지로부터 가까운 거리에 있는 영역을 손 후보 영역으로 추출한 후, convex decomposition 알고리즘을 사용하여 손바닥 영역과 손가락 영역을 추출한다. 이렇게 손 영역을 추출 한 뒤에는 손가락의 구조를 추정하여 손동작과 손 자세를 인식한다[1].

본 논문에서는 RGBD 카메라를 사용하는 방법과 유사한 손 추출 방법을 사용하였지만, 일반적으로 사용하는 RGBD 카메라 대신 적외선 센서인 Leapmotion을 입력으로 사용하였다. 적외선 센서는 일반적으로 RGBD 카메라에 비해 크기가 작고 가벼우며, 입력 영상의 해상도가 높고 설치가 용이하며 다른 기기와의 연동이 수월하다는 장점이 있다. 또한 이러한 특성 때문에 증강현실 및 가상현실 시스템 구현에 많이 쓰이는 HMD(Head Mount Display)나 스마트폰에 내장하여 사용할 수 있어서, 손 영역 추출 시 입력 기기로 RGBD 카메라 대신 적외선 센서를 사용하였다.

본 논문에서 제안하는 손동작 인식 알고리즘은 손 영역 추출, 손바닥 추출, 손가락 추출, 손동작 인식의 4가지 과정으로 나눌 수 있다. 거리변환(distance transform)을 사용하여 손바닥의 중심좌표와 손바닥의 크기를 알아내고, convex hull method와 convexity defect를 계산하여 손가락을 추출한다. 그 후, 손가락과 손바닥 부분의 optical flow를 계산하여 클릭, 이동, 정지의 세 가지 손동작을 인식하였다. 인식된 손동작을 활용하여 간단한 사용자 인터페이스 및 상호작용을 구현하

여 향후 증강현실(AR) 및 가상현실(VR)시스템에 활용 할 수 있도록 하였다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문에서 제안하는 손동작 인식 알고리즘에 대하여 자세히 기술하였다. 3절에서는 손동작 인식 결과와 간단한 인터페이스 구현 내용을 소개하였고, 마지막으로 4절에서는 본 논문에 대한 결론과 향후 추가로 연구 가능한 주제를 정리하였다.

2. 손동작 인식 알고리즘

2.1 손 영역 추출

적외선 센서로 취득하는 이미지는 센서 의 특성상 가까운 거리에 위치하는 물체가 먼 거리에 있는 위치에 비해 밝은 값을 가진다. 이러한 특성을 사용하여 밝기 값을 기준으로 최대 밝기 값의 0.1에 해당하는 값을 임계점으로 잡고 임계점을 넘는 영역을 대략적인 손 영역으로 추출한 후, 제 2의 임계점을 최대 밝기 값의 0.3에 해당하는 값으로 적용하여 더 밝은 손 부분과 상대적으로 어두운 부분인 팔뚝 영역을 분리하였다. 이 때 추출된 경계선의 길이가 짧은 부분들은 손 영역에서 제외함으로써 불필요한 잡음을 제거하였다. 이를 수식으로 나타내면 다음과 같다

$$I_{rough}(x,y) = \begin{cases} 1, & \text{if } I(x,y) > 0.1 \times \max(I) \\ 0, & \text{else} \end{cases}$$

$$I_{hand}(x,y) = \begin{cases} 1, & \text{if } I(x,y) > 0.3 \times \max(I) \\ 0, & \text{else} \end{cases}$$

$$I_{forearm}(x,y) = \begin{cases} 1, & \text{if } 0.3 \times \max(I) > I(x,y) > 0.1 \times \max(I) \\ 0, & \text{else} \end{cases}$$

그림 1은 취득한 이미지와 손 영역, 팔뚝 영역을 추출해 낸 결과이다.

2.2 손바닥 추출

손바닥 인식은 추출한 손 영역에 거리변환(distance transform)을 이용하여 가장 큰 블랍(blob)을 찾아 낸 후, 그 중심을 손바닥 중심으로 정하였다. 그 후 중심점의 거리변환(distance transform) 값을 반지름으로 하는 원을 추출하여 손바닥 영역으로 정하였다. 팔뚝 중심 역시 추출한 팔뚝 영역에 거리변환(distance transform)을 이용하여 팔뚝의 중심점을 찾아 주었다.

2.3 손가락 추출

손가락 추출 시에 사용한 방법은 convex hull 알고리즘이다. convex hull 알고리즘은 주어진 이진 화 이미지상의 여러 개 점이나 곡선, 도형에 대해, 모든 영역을 포함하는 convex polygon을 찾아주는 알고리즘이다.

자세한 방법은 다음과 같다. 우선 그림 2처럼 convex hull algorithm을 통해 추출한 손 영역을 감싸는 convex polygon을 구한다. 그 후, convexity defect가 생기는 부분과 convex polygon의 꼭지 점

을 모두 연결한 도형을 만든다. 마지막으로 해당 도형에서 두 개의 defect point와 그 사이의 convex polygon 꼭지 점이 이루는 각도와 defect값을 측정하여, 임계값을 넘는 경우 해당 convex polygon의 꼭지 점을 손가락 끝 영역으로 추출 한다

$$x_{finger} = \begin{cases} 1, & \text{if } degree(x_i) > 70^\circ \cap defect(x_i) < \sqrt{700} \\ 0, & \text{else} \end{cases}$$

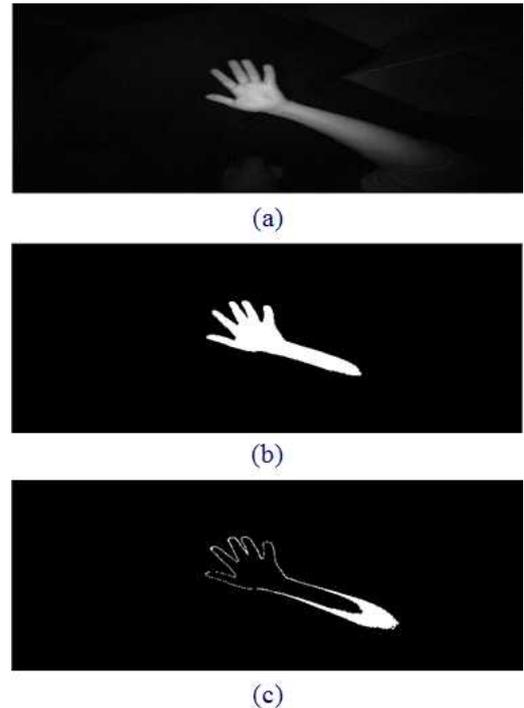


그림 1. 입력 영상과 각 영역 추출영상. (a) 입력영상, (b) 손 영역 I_{hand} , (c) 팔뚝 영역 $I_{forearm}$

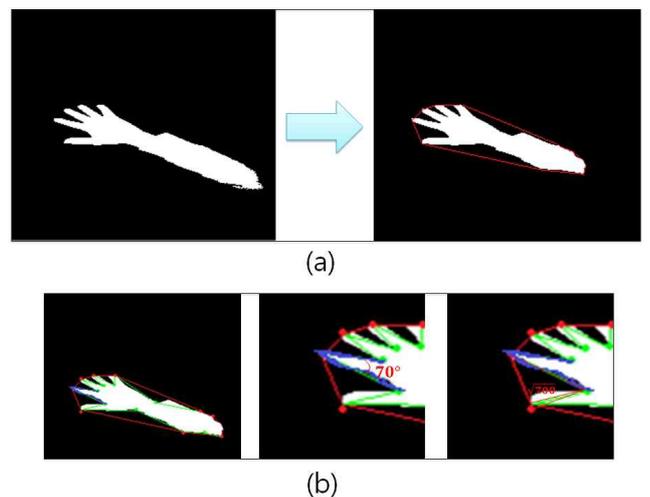


그림 2. 손가락 추출 과정. (a) 손 영역에서 convex polygon 추출, (b) 추출한 convex polygon과 defect 꼭지 점으로 손가락 추출

2.4 손동작 인식

손바닥과 손가락 인식이 끝난 후에 추출된 손가락과 손바닥 영역

의 optical flow를 구해 손동작 인식을 구현하였다. 그림 3과 같이 손 영역의 optical flow를 계산 해 보면 클릭 시와 정지, 이동 시 optical flow의 형태가 각각 다른 것을 알 수 있다. 클릭 시에는 손바닥 영역과 손가락 영역의 optical flow 방향과 크기가 서로 다른 것을 알 수 있고, 정지 시에는 두 군데 모두 optical flow가 거의 없으며, 이동시에는 두 영역에서 optical flow의 방향과 크기가 거의 비슷한 것을 알 수 있다.

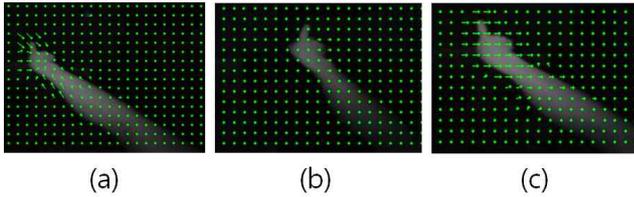


그림 3. 클릭, 정지, 이동시 optical flow. (a) 클릭 시 optical flow, (b) 정지 시 optical flow, (c) 이동 시 optical flow

각각 손가락과 손바닥 중심에서 19*19영역의 영상에 대해 farnback method[6]를 통해 optical flow를 계산한다. 이후 각 영역의 optical flow의 평균을 취한 후 손가락 영역과 손바닥 영역의 평균 optical flow를 아래와 같이 비교하여 손동작 인식을 구현하였다

$$x_{act} = \begin{cases} 0, & \text{mean}(F_{finger}) \approx 0 \cap \text{mean}(F_{palm}) \approx 0 \\ 1 & \text{mean}(F_{finger}) > \theta_{mov} \cap \text{mean}(F_{palm}) > \theta_{mov} \\ 2 & \text{mean}(|F_{finger} - F_{palm}|) > \theta_{mov} \\ & \cap \text{mean}(F_{finger}) > \theta_{mov} \end{cases}$$

위 식에서 x_{act} 가 0일 경우 정지, 1일 경우 이동, 2일 경우 클릭에 해당하고 θ_{mov} 값은 5로 하였다.

3. 실험 결과

본 논문에서는 손동작 인식을 구현하기 위하여 기본적인 이동, 클릭, 정지의 3가지 동작을 구분하도록 하였다. optical flow를 계산하여 평균을 구한 후, 손가락과 손바닥 부분의 optical flow 차이를 계산하여 현재 동작이 어떤 동작인지 구분 할 수 있었다. 이를 통해서 가상현실(VR)혹은 증강현실(AR)환경에서 시스템의 메뉴 팝업 및 선택, 물체 이동 등의 조작이 가능하다. 적외선 센서를 사용했기 때문에 실험 환경에 따라서 인식률의 편차가 크지만, 평균적으로 92.3%이상 클릭 동작을 인식하는 것으로 나타났다.



그림 4. 클릭 시 동작 인식 영상

그림 4는 클릭 동작 인식 화면이다. 이 외에도 다양한 방향과 각

도에서의 클릭 동작에 대해 안정적으로 동작 인식을 하는 것을 알 수 있었다.

표 1에서는 본 논문에서 구현한 알고리즘과 기존의 알고리즘 간 속도를 비교하였다. 기존의 방식에 비하여 연산속도가 크게 개선되어, 거의 실시간으로 영상 처리가 가능함을 알 수 있다.

표 1 각 알고리즘 별 손 인식 소요시간

방법	수행시간(s)
Thresholding Decomposition + FEMD in [2]	0.5004s
Near-convex Decomposition + FEMD in [2]	4.0012s
제안 알고리즘	0.073s

4. 결론 및 향후 연구 방향

최근에 미래 사용자 인터페이스 기술로 대두된 가상현실(VR)과 증강현실(AR) 시스템을 구현하는데 있어서 손동작 인식은 필수적으로 들어가야 할 기술이다. 본 논문에서는 거리변환 방법과 convex hull 알고리즘, farnback optical flow 알고리즘을 사용하여 손동작을 인식 하도록 구현함으로써 거의 실시간으로 동작하는 빠르고 정확한 손동작 인식 방법을 제안하였다. HMD(head mount display)나 모바일 환경에서도 동작 할 수 있도록 RGBD 카메라 대신 적외선 센서를 사용하는 시스템을 구성하였고, 동작을 확인하였다.

향후에는 손가락 하나로 동작을 인식하는 것이 아니라 멀티터치와 양손 손동작 인식 등 다양한 상호작용을 구현하여 실제 가상현실(VR)과 증강현실(AR)시스템에서 그 활용도를 높이는 연구를 수행하고자 한다.

참고문헌

[1] S. Qin, X. Zhu, and Y. Yang, "Real-time Hand Gesture Recognition from Depth Images Using Convex Shape Deomposition Method", Journal of Signal Processing Systems. 2013.
 [2] Ren, Z., Yuan, J., Zhang, Z., "Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera," Proc. of ACM International Conference on Multimedia, pp. 1093 - 1096, 2011.
 [3] Argyros, A., Lourakis, M., "Real-time tracking of multiple skin-colored objects with a possibly moving camera," Proc. of ECCV, pp. 368-379. 2004.
 [4] Argyros, A., Lourakis, M., "Vision-based interpretation of hand gestures for remote control of a computer mouse," Proc. of

Computer Vision in Human-Computer Interaction, pp. 40-51, 2006.

[5] Wang, R. Y., Popovic, J., "Real-time hand-tracking with a color glove," ACM Transactions on Graphics, 28(63), 1-8 2009.

[6] G. Farneback, "Very high accuracy velocity estimation using orientation tensors, parametric motion, and simultaneous segmentation of the motion field," ICCV 2001