Journal of The Korea Society of Computer and Information
Vol. 19, No. 11, November 2014
www.ksci.re.kr
http://dx.doi.org/10.9708/jksci.2014.19.11.063

깊이 카메라와 SVM을 이용한 수화 인식 시스템

김 기 상*, 최 형 일*

Sign Language Recognition System Using SVM and Depth Camera

Ki-Sang Kim*, Hyung-Il Choi*

요 약

본 논문에서는 깊이 카메라를 이용한 사용자의 손 모양 인식 시스템을 제안한다. 특히, 본 시스템에서 이용된 손 모양 템플릿은 수화 언어 중 한국어를 채택 하였다. 손 모양 인식 시스템은 손의 특징 검출과 특징들을 이용한 손 인식으로 크게 2 단계의 작업으로 나눌 수 있다. 손의 특징으로는 손가락의 개수, 길이, 손바닥의 넓이 등이 있다. 특징을 추출하기 위해 본 논문에서는 거리 변환(Distance Transform)을 이용한 손의 뼈대 검출 방법을 제안한다. 이 방법을 사용하면 기존의 윤곽선(Contour)을 이용한 손가락 검출보다 정확도 측면에서 향상된다. 손 모양 인식으로 손의 특징을 이용하여 각 분기를 잘 나눌 수 있는 결정 트리(Decision Tree)를 사용한다. 사용자의 입력을 이용하면 분기값이 정확하게 나오지 못하므로 이 분기 값을 결정하기 위해 해당 분기마다 SVM을 이용하여 분기값을 결정하였다. 실험결과에서는 기존의 연구 방법보다 제안된 방법이 특징 추출과 인식하는데 있어 더욱 개선되었음을 보인다.

▸ Keywords : 수화, 손 모양 인식, SVM

Abstract

In this paper, we propose a sign language recognition system using SVM and depth camera. Especially, we focus on the Korean sign language. For the sign language system, we suggest two methods, one in hand feature extraction stage and the other in recognition stage. Hand features are consisted of the number of fingers, finger length, radius of palm, and direction of the hand. To extract hand features, we use Distance Transform and make hand skeleton. This method is more accurate than a traditional method which uses contours. To recognize hand posture, we develop the decision tree with the hand features. For

•제1저자 : 김기상 •교신저자 : 최형일
•투고일 : 2014. 9. 19, 심사일 : 2014. 9. 29, 게재확정일 : 2014. 10 28
* 숭실대학교 미디어학과(School of Media, Soongsil University)
※ 본 연구는 서울시 산학연 협력사업(SS110013)과 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.2013R1A1A2012012)

more accuracy, we use SVM to determine the threshold value in the decision tree. In the experimental results, we show that the suggested method is more accurate and faster when extracting hand features a recognizing hand postures.

▸ Keywords : Sign Language, Hand Posture Recognition, SVM

I. 서 론

최근 모바일과 스마트 TV 시장의 증가와 다양한 기기와 스마트 디바이스의 융합으로 인해 스마트 디바이스 장치는 다양한 장소에서 광범위 하게 활용되고 있다. 이들 중 특히 문자 입력 방법으로는 스마트 TV에 달린 카메라를 이용한 수화 인식[1] 시스템에 대해 많은 개발이 이루어지고 있다[1-5]. 기존의 수화 인식 시스템은 사용자가 칼라 장갑, 데이터 글러브 등의 특별한 입력 장치와 함께 인식되는 것이 대부분이었다. 하지만, 최근의 많은 연구는 사용자에게 칼라 장갑 등 특별한 매체를 장착하지 않고 인식하는 방법으로 진행되고 있다. 특별한 매체 정보를 이용하면 수화 인식 시스템은 손의 모양과 움직임 정보를 보다 쉽게 얻을 수 있지만, 사용자가 항상 장갑 등을 휴대하고 다녀야하는 불편함이 생길 수밖에 없다. 또한, 최근에는 적외선 센서를 이용한 깊이 카메라가 출시되어, 기존의 칼라 장갑 등 손을 인식하기 위한 매체들이 무의미 해지고 있다.

최근 연구에서는 손 모양을 인식하여 TV 등 스크린 상에 문자를 입력하는 방법이 화두가 되고 있다. 신경망(Neural Network)을 이용한 손 모양을 직접 학습시켜 인식하는 방법[6], 손바닥 영역을 제거하여 후보군을 추출시켜 인식하는 방법[7], 손의 특징을 추출하여 SVM(Support Vector Machine) 등을 이용하여 학습시키는 방법[8], 손의 테두리를 펼쳐 손가락을 인식하는 방법[9] 등 여러 가지 방안이 제안되었다. 신경망의 경우 손이 고정되어 움직인다는 전제하에 인식이 가능하므로, 회전 등 다양한 손 모양을 인식하기엔 부적합 하며, 사용자가 사용하기에도 많은 불편함이 발생하는 문제가 있다. 또한, 이러한 방식의 경우 인식 할 수 있는 손 모양의 개수에도 한계가 있어, 실제 문자 입력 시스템으로 적용하기에는 적합한 방법이 아니다. 이러한 문제점을 해결하기 위해 손바닥을 제거하여 인식하는 방법도 제안되었다. 하지만 손가락이 붙어 있는 경우, 이를 따로 떼어내기란 불가능하며, 이로 인해 오류가 발생하는 문제가 있다. 손의 테두리를 펼쳐 손가락을 인식하는 방법은, 손을 인식하는데 있어 좋은 알고리즘이나, 이 방법 또한 신경망 기반 인식 방법과 마찬가지로, 손의 회전에 대해 구분하기란 불가능에 가까워, 손 모양을 인식하는데 있어 인식 개수에 한계가 있다. 마지막으로 손의 특징을 추출해 SVM 등을 이용하여 학습시키는 방법은 적합한 방법이나, 손의 특징 추출 방법에 의존적이며, 개수에 한계가 있다.

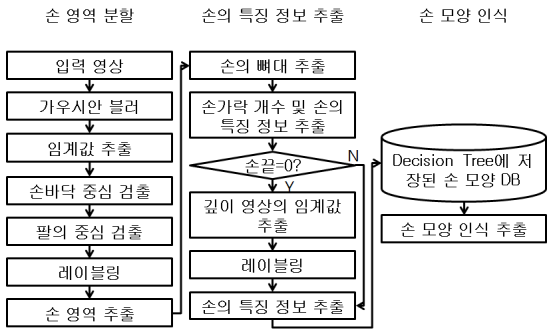


그림 1. 시스템 전체 흐름도
Fig. 1. Structure of Proposed Method

그림 1은 본 논문에서 제안하는 시스템의 제어 흐름도를 나타낸다. 본 논문에서는 자음 14자와 모음 14자, 총 28개의 템플릿을 인식하는 손 모양 인식 시스템에 대해 제안한다. 먼저 손을 검출하기 위해 손 영역은 항상 카메라로부터 가장 가까운 영역에 있다는 가정을 이용, 영상의 깊이 값을 이용하여 손 영역을 검출한다. 그 후 손 모양을 인식하기 위해 특징 추출을 한다. 특징 추출에 사용되는 방법은 손의 뼈대 모양 기반 방법과 깊이 값을 이용한 방법이 있다. 그림 1에서 손끝의

* 이 논문은 2014년 한국컴퓨터정보학회 제50차 하계학술대회에서 발표한 논문("깊이 영상 기반 정적 수화 인식 시스템")을 확장한 것임.

개수를 구분하는 부분이 이 방법을 구분하는 부분이다. 사용되는 특징으로는 손가락의 개수, 길이, 각도, 손바닥의 크기, 손의 각도 등을 추출하고 이를 인식하는데 이용하여 비교적 수화에 가깝게 인식하는 방법을 제안한다. 손 모양 인식으로는 결정 트리(Decision Tree)를 사용하였으며, 결정 트리를 만들기 위해 분기점 마다 분할되는 포인트는 SVM을 이용, 사용자가 직접 설정한 결과보다 정확한 분기점을 생성하였다. 이렇게 생성된 결정 트리를 이용해, 미리 입력된 정보와 비교하여 손 모양을 인식하였다. 본론에서는 손 모양 특징 추출방법과 인식에 대해 설명한다. 실험 결과에서는 기존의 방식에 비해 개선된 점을 비교하며, 실험 예제에 대해 제시한다. 결론에서는 개선해야 할 점 등에 대해 설명한다.

II. 본 론

1. 깊이 카메라를 이용한 손 영역 검출

깊이 카메라는 일반적으로 잡음이 많이 존재한다. 그러므로 이를 그대로 사용할 경우, 원하는 결과가 안 나오는 문제가 간혹 발생한다. 이를 미연에 방지하기 위해 스무딩 연산을 수행한다. 스무딩 연산 중 가우시안 커널을 이용한 스무딩 기법이 일반적으로 널리 사용되며, 노이즈 제거에 효과적이다. 가우시안 스무딩을 적용한 영상을 가지고 이진화 기법을 통해 손 후보 영역을 검출한다. 이 방법은 손은 항상 카메라로부터 가장 가까운 영역에 있다는 가정에서 시작하기 때문이다. 하지만 임의적으로 임계값(T)을 정하는 것이 아니라, 깊이 값과 해당하는 영역의 크기를 이용하여 검출하게 된다. 임계값을 계산하기 위해 우선 누적 영상(Integral Image, G)을 생성한다. 누적 영상을 통해 일정 영역 크기(ω)에서 그 평균 깊이 값이 최소가 되는 값을 찾는다. 이렇게 하는 이유는 픽셀 단위로 검증할 경우, 손 영역이 기울여 졌을 경우 손 영역 전체가 나오지 못하고 부분적으로만 나오는 문제가 발생하기 때문이다. 수식 (1)은 ω크기의 윈도우에 해당하는 깊이 값의 평균(S)을 구한다. ω값은 깊이 값에 따라 달라지지만, 최대 크기는 30으로 이용했다. 그러한 이유는 그 이상 커질 경우, 필요 없는 영역 포함되어, 제대로 된 검출을 못하기 때문이다.

$$S(x,y)=\frac{G(x+w,y+w)-G(x+w,y-w)-G(x-w,y+w)+G(x-w,y+w)}{(2w+1)^2} \quad (1)$$

수식 (1)의 S값들 중 최소가 되는 값을 임계값으로 설정한다. 임계값만을 사용할 경우 손 영역이 아슬아슬하게 검출되므로, 임계값과 α값을 더한 값이 최종 임계값(T)로 설정한다. 본 시스템에서 α값은 50으로 설정하였다.

임계값을 이용한 손 후보 영역을 검출하였다면, 정확한 손 영역을 설정하기 위해 필요 없는 부분, 팔 영역을 제거한다. 팔 영역을 제거하기 위해 우선 팔 영역의 중심을 계산한다. 팔 영역 중심은 T-30~T사이의 값을 가진 영역들의 중심점을 말한다. 이러한 방법이 가능한 것은 팔은 항상 손에 비해 카메라로부터 항상 뒤에 있기 때문이다. 팔 영역을 검출하면, 그 다음으로 손바닥 중심과 손바닥의 반지름을 거리 변환(Distance Transform)을 이용하여 검출한다. 손 영역에서 손바닥 중심은 손가락으로 인해 항상 영역에서 가장 깊숙한 곳에 있기 때문이다. 그림 2는 거리 변환 방법의 예제를 나타낸다.

0	0	1	0	0	0	0	0
0	1	1	1	0	0	0	0
1	1	1	1	1	0	0	0
1	1	1	1	1	1	0	0
1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1

(a)

0	0	1	0	0	0	0	0
0	1	2	1	0	0	0	0
1	2	3	2	1	0	0	0
1	2	3	3	2	1	0	0
1	2	3	3	3	2	1	0
1	2	2	2	2	2	2	1
0	1	1	1	1	1	1	1

(b)

(c)

그림 2. 거리 변환 예제. (a) 이진 영상, (b) (a)영상에 거리 변환이 적용된 영상, (c) 거리 변환이 적용된 손 영상
Fig. 2. Example of Distance Transform. (a) Binary Image, (b) Distance Transformed Image (c) Distance Transformed Hand Image

손바닥의 중심과 그 반지름, 그리고 팔 중심을 구했다면, 최종적으로 손 영역을 검출한다. 검출 방법으로는 손바닥 밑에 팔 영역이 있고, 손바닥 영역은 알고 있으므로, 손바닥 영역만 살리고 그 밑에 필요 없는 팔 영역은 제거하는 것이 이 방법의 주요 목적이다. 팔 중심으로부터 손바닥 중심까지의 거리를 반지름으로 하며, 팔을 중심으로 하는 원의 영역은 제거하대, 손바닥 중심을 기준으로 손바닥의 반지름 * 1.2안에 있는 영역은 그대로 남긴다. 그림 3의 (a)와 (b)는 해당 영역들의 범위들을 나타내며, (c)와 (d)는 손 영역 검출 결과를 나타낸다. 또한, (a)와 (c)는 손이 카메라에 가까워 팔 영역이 거의 없었을 경우에 검출 결과를 나타내며, (b)와 (d)는 손이 카메라와 멀어져서 팔 영역이 많이 나온 경우의 결과 영상을 나타낸다.

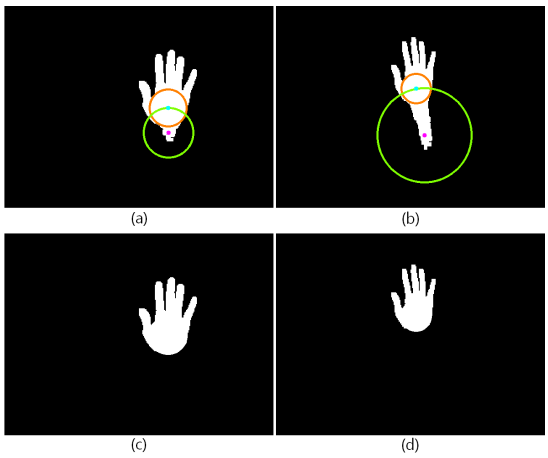


그림 3. 손 영역 검출 예제. (a) 손이 가까울 때 영역 검출을 위한 특징들, (b) 손이 멀 때 영역 검출을 위한 특징들, (c) (a)를 이용한 영역 검출 결과, (d) (b)를 이용한 영역 검출 결과
Fig. 3. Example of Hand Region Extraction. (a) Hand Region Features when the hand is near to camera, (b) Hand Region Features when the hand is far from camera, (c) Extracted Hand Region with (a), (d) Extracted Hand Region with (b)

2. 인식을 위한 손 특징 검출

본 시스템에서 사용되는 손 인식을 위한 특징 검출은 크게 2가지로 구분된다. 처음 방법으로는 손의 뼈대 검출을 통한 손의 특징 검출 방법이 있으며, 다른 방법으로는 손 특징을 위한 손 영역 분할과는 또 다른 임계값을 이용함으로써 특징을 추출하는 방법이 있다. 이와 같이 두 가지의 방법으로 나누게 된 계기는 그림 4와 같은 문제가 있기 때문에다. 그림 4 (a)의 경우 손바닥과 카메라가 수평 방향에 있다. 이와 같은 경우 손의 특징 검출하는데 있어 뼈대를 추출하여 검출하는 것이 효율적이다. 하지만 그림 4 (b)의 경우, 손바닥과 카메라가 수직 방향에 있어 뼈대를 검출하는데 큰 문제가 존재한다. 하지만 다행이도 수화 템플릿에서 이와 같은 패턴은 총 4개만 존재하며, 구분하기 어렵지 않게 되어 있다. 그러므로 이와 같은 것은 간단히 임계값 이용하여 손의 특징을 추출한다.

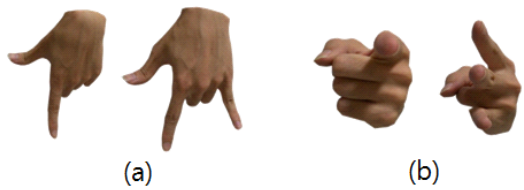


그림 4. 수화 예제 (a) ㄱ, ㅋ, (b) ㅓ, ㅖ
Fig. 4. Example of Sign Language

2.1 손의 뼈대 검출을 이용한 손 정보 검출

본 논문에서 손의 특징들, 특히 손가락을 구하기 위해 우선 손의 뼈대를 검출한다. 손의 뼈대를 검출하는 이유는 기존의 윤곽선(Contour)을 이용한 방식은 애매모호한 손끝을 검출하며, 이를 제거하기 위해 통합과정이 필요한 문제가 있다. 또한, 통합과정에 있어 설계를 잘못하면 손 사이가 붙어 있는 경우, 손가락을 구분하기 어렵다는 또 다른 문제를 야기한다. 그림 5에서는 엄지와 중지의 경우 여러 개의 점이 잡혀 손끝이 제대로 추출되지 못하는 문제점을 보여준다.

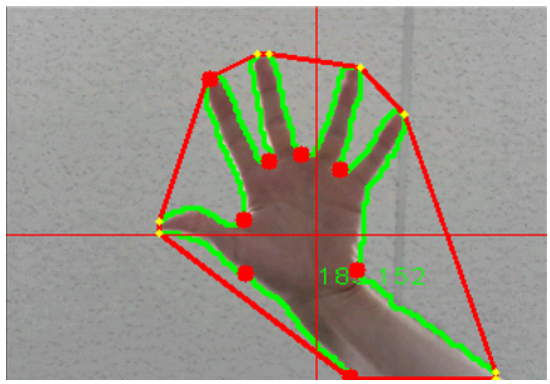


그림 5. 윤곽선 이용 방법의 문제점
Fig. 5. Problem of Contour-using Method

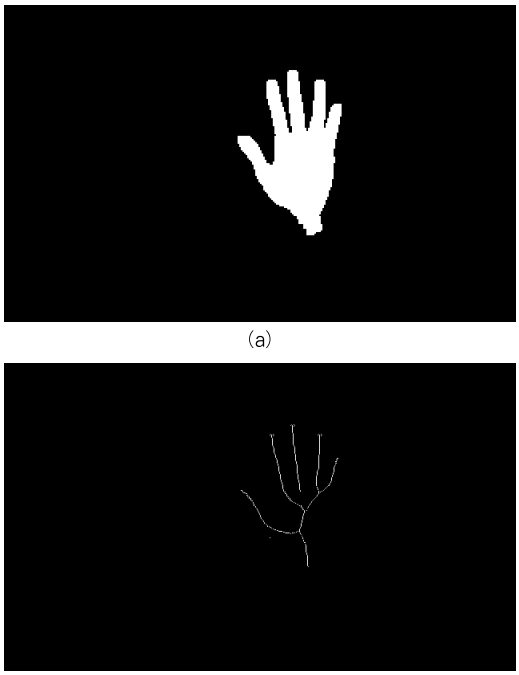
손의 뼈대를 검출하기 위해 Distance Transform의 특성을 이용한다. Distance Transform이 적용된 영상은 중심으로 갈수록 거리값이 높아지며, 배경으로 가까울수록 거리값이 작아지는 특성이 있다. 이런 특성을 이용, 수식 (2)~(4)의 방법을 통해 손의 뼈대를 추출할 수 있다.

$$Q_1(x,y)=\begin{cases}\text{if},D(x,y)\geqq L/10 & 1\\ else, & 0\end{cases} \quad (2)$$

$$Q_2(x,y)=\begin{cases}\text{if},c\leqq 2 & 1\\ else, & 0\end{cases} \quad (3)$$

$$Q(x,y)=\begin{cases}\text{if},Q_1(x,y)=1\cap Q_2(x,y)=1 & 1\\ else, & 0\end{cases} \quad (4)$$

수식 (2)는 손의 뼈대 여부가 필요 없는 부분을 제거한다. 너무 세밀한 부분까지 검출하게 될 경우, 손의 뼈대를 제대로 검출하지 못하고 노이즈가 발생하는 문제가 생기기 때문이다. L은 앞에서 거리 변환을 생성하면서 나온 손바닥의 반지름을 말하며 D(x,y)는 거리 변환을 이용해 생성된 영상을 말한다. 수식 (3)은 현재의 좌표 (x,y)를 기준으로 현재의 D값과 8방향의 주변 D값을 비교한다. 예를 들어 D(x,y)-D(x-1,y)를 했을 때 이 값이 음수 일 경우 c값을 하나 증가한다. 반대로 음수일 경우 c값은 변동이 없다. 여기서는 자신보다 주변의 거리값이 얼마나 큰지 비교하며 뼈대여부인지를 구분한다. 수식 (4)는 최종적으로 수식 (2)와 수식 (3)이 둘 다 참인 것을 최종적으로 뼈대에 해당되는 좌표라고 인식한다. 그림 6은 이 수식을 이용하여 검출한 손의 뼈대 영상이다.



(a)

(b)

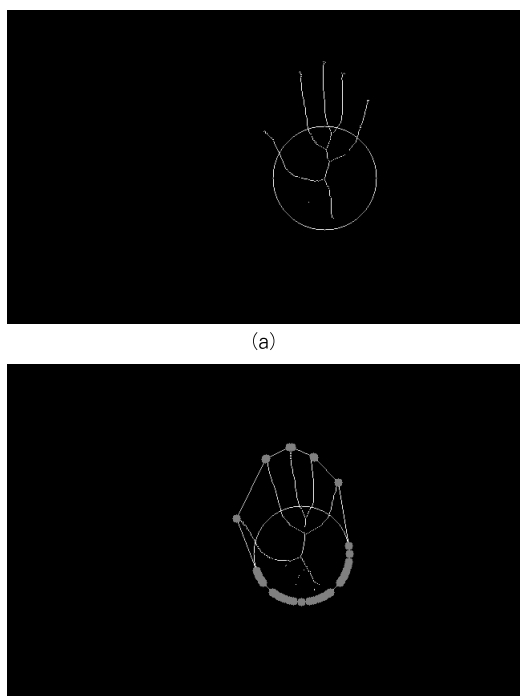
그림 6. 손의 뼈대 영역 검출 예제 (a) 손 영상 (b) 검출된 손의 뼈대 영상

Fig. 6. Example of Hand Skeleton Extraction (a) Hand Image (b) Extracted Hand Skeleton

손의 뼈대를 검출하였으면, 이를 이용해 손가락 검출을 시도한다. 우선 손가락 검출에 앞서, 손가락 검출의 시작지점, 손끝을 찾기 위해 볼록 껍질(Convex Hull) 알고리즘을 사용한다. 수식 (5)는 볼록 껍질 알고리즘을 나타낸다.

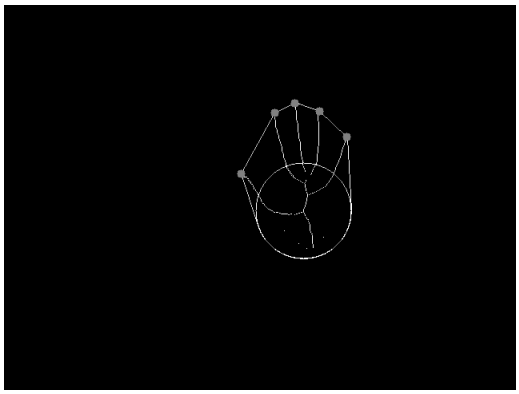
$$C\equiv\left\{\sum_{j=1}^{N}\lambda_j p_j:\lambda_j\geq \text{for } all\ j \text{ and } \sum_{j=1}^{N}\lambda_j=1\right\} \quad (5)$$

수식 (5)에서 $p_1,\ldots,p_N$은 수식4의 결과 Q(x,y)의 위치들을 나타내며 N은 Q(x,y)의 픽셀 개수를 나타낸다. 위의 볼록 껍질 알고리즘을 그림 6의 (b) 영상을 그대로 사용하여 손끝을 검출을 시도하면, 손바닥 영역이 제대로 표현되지 않아 제대로 된 손끝을 검출할 수 없다. 이와 같은 문제를 해결하기 위해 그림 7 (a)와 같이 손바닥 영역 자체를 원으로 표현하여 입력하게 된다. 이와 같은 영상을 이용하여 볼록 껍질 알고리즘을 돌리게 되면, 그림 7 (b)와 같이 손바닥 부근에도 손끝이라고 잘못 추출되는 결과가 나오게 된다. 이러한 결과는 손바닥 영역에 포함되는 손끝 후보군은 오류라고 판단, 이들을 제거하게 되면 그림 7(c)와 같은 손끝 추출 결과를 나타내게 된다.



(a)

(b)



(c)
그림 7. 손끝 검출 결과 예제 (a) 입력 영상, (b) 볼록 껍질 결과 영상, (c) 손끝 검출 결과 영상
Fig. 7. Example of Fingertips Extraction (a) Input image (b) Convex Hull (c) Extracted Fingertips

손끝을 검출하였다면 최종적으로 손가락을 검출하기 위해 손끝으로부터 손바닥 영역으로 탐색한다. 탐색으로는 재귀함수를 사용하였으며, 검출된 손끝을 시작으로 손바닥 영역 안으로 들어오거나, 뼈대가 이어지지 않을 때 까지 계속 탐색한다. 이러한 정보를 이용해, 각 손가락의 길이와 방향성을 결정한다. 또한 가장 긴 손가락의 방향을 현재 손의 대표 방향성으로 결정, 이러한 정보들을 손 모양 인식기에 입력하게 된다. 그림 8은 손의 특징 정보 검출 결과를 나타내며 그림 9의 경우 손이 붙어 있을 경우 다른 예외처리 없이, 제안한 방법으로 검출하면 관절 검출은 실패하지만, 손끝 검출은 문제없음을 보여준다.

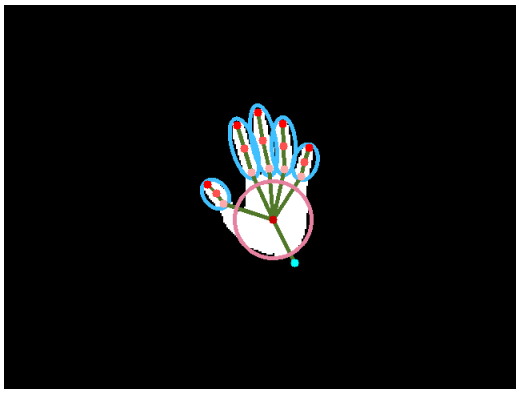


그림 8. 손의 뼈대를 이용한 손의 특징 정보 검출 결과
Fig. 8. Result of Hand Feature Extraction using Skeleton

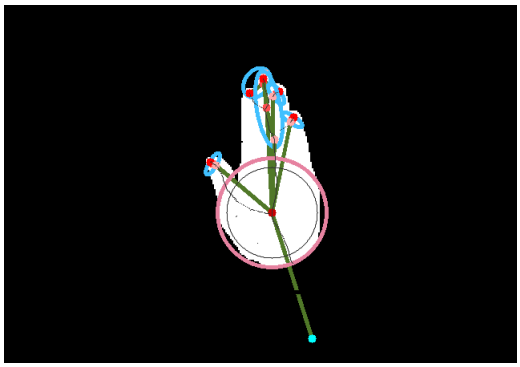
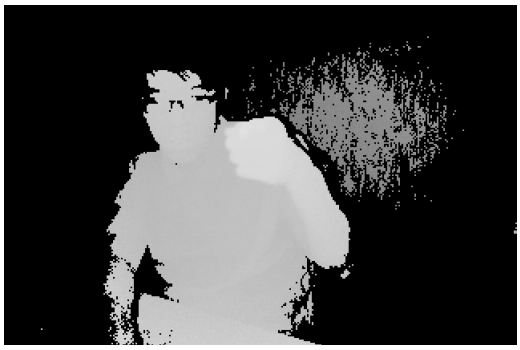


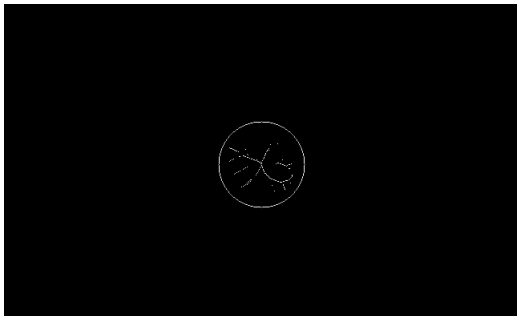
그림 9. 손가락이 붙어 있었을 경우에 손가락 검출 결과
Fig. 8. Result of Hand Feature Extraction under sticked fingers

2.2 깊이 값을 이용한 손 정보 검출

그림 10와 같은 경우 손바닥의 평면이 카메라와 직교함으로 손의 뼈대 영상을 추출할 수 없다. 이러한 경우, 깊이 값을 이용한 손 정보 검출에 시도한다. 깊이 값을 이용한 손 정보 검출은 뼈대 기반 손 정보 검출보다 구분해야 할 수화의 개수가 적어, 손가락 개수, 손가락 간의 거리로 구분하게 된다.



(a)



(b)
그림 10. 손의 뼈대 기반 특징 검출 실패 예제 (a) 원본 입력 영상, (b) 뼈대 검출 영상
Fig. 10. Failure Example of Extracting Hand Feature (a) Input Image, (b) Extracted Hand Skeleton

깊이 값을 이용한 손끝 검출은 우선 깊이 영상에서 최소가 되는 깊이값을 검출한다. 단, 검출된 최소 깊이 값은 손 영역 안에 해당되는 부분아 안에서 한정한다. 검출된 최소값이 있으면, 이 값에 55를 더하여 임계값을 만들고, 이진 영상을 만든다. 55라는 수치는 실험결과 일반적으로 사람의 손가락이 손으로부터 튀어나온 정도를 나누는데 적당한 값이다. 이진 영상을 이용, 레이블리을 통해 손가락의 개수를 검출한다. 손가락의 개수를 세는데 있어, 영역의 모양을 보고 판단한다. 특정 영역이 장축과 단축의 비율이 2:1 이상일 경우 해당 손가락은 붙어있는 것으로 판단, 2개의 손가락으로 구분하게 된다. 또한 손가락이 3개 있을 경우, 각각의 손가락 사이의 유클리디언 거리를 측정, 특징 정보로 사용된다. 그림 11은 깊이 값을 이용한 손 정보 검출에 대한 결과를 나타낸다.

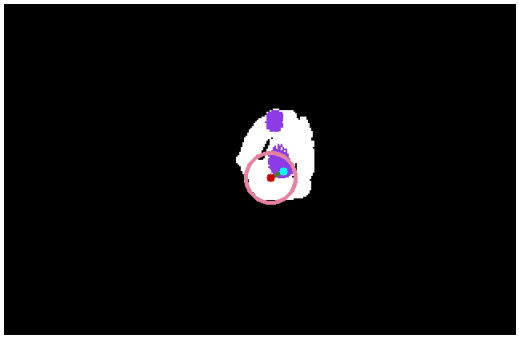


그림 11. 깊이 값을 이용한 손가락 특징 추출 결과
Fig. 11. Extracted Hand Feature using Depth Threshold Value

3. 손 모양 인식

손 모양 인식에 있어 결정 트리를 이용하는 것이 구현하기 쉬우며 빠른 결과를 도출 할 수 있다. 결정 트리는 그림 12에서와 같이 손 모양 인식에 사용되는 결정 트리는 뼈대 기반 특징은 네 가지 레이어로 분류 되며, 깊이 값을 이용한 손 정보를 이용한 인식은 두 가지 레이어로 분류 된다.

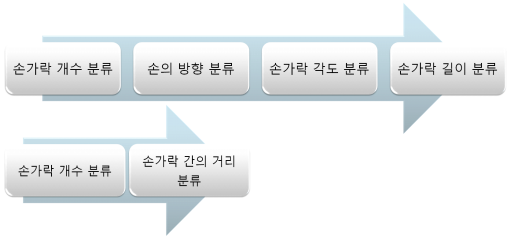


그림 12. 손 모양 인식 순서
Fig. 12. Steps of Hand Posture Recognition

하지만, 기존의 결정 트리는 분기부분을 사람이 직접 설정하여 축이 한쪽으로 쏠리는 문제가 생긴다. 표 1에서 'ㄱ', 'ㄴ', 'ㄹ', 등 특히 뼈대 기반 특징을 이용한 손 영역 검출에서 한쪽으로 쏠리는 문제가 유난히 들어난다. 이러한 문제점을 해결하기 위해, 본 시스템에서는 SVM을 이용하여, 분기를 설정한다. 전체에 대해 SVM을 적용하는 것은 아니며, 결정 트리에 있는 레이어들 중 손의 뼈대 기반 손 모양 인식에서 손의 각도 분류 단계와 손가락 각도 분류 단계에 적용하였다. SVM을 적용하고 난 후, 모든 제스처가 -10°~10° 사이에서는 100%의 인식률을 보였다.

표 1. 기존의 인식 방법을 사용하였을 때, 각도에 따른 제스처 인식 문제
Table 1. Problem of Hand Posture Recognition along Degrees using Traditional Method.

제스처	-10°	-5°	0°	5°	10°
ㄱ	100%	100%	100%	0%	0%
ㄴ	100%	100%	100%	0%	0%
ㄷ	0%	100%	100%	100%	0%
ㄹ	0%	100%	100%	0%	0%
ㅁ	100%	100%	100%	0%	0%
ㅂ	100%	100%	100%	100%	100%
ㅅ	0%	100%	100%	100%	0%
ㅇ	50%	100%	100%	50%	0%
ㅈ	100%	100%	100%	100%	100%
ㅊ	100%	100%	100%	100%	0%
ㅋ	100%	100%	100%	100%	100%
ㅌ	0%	100%	100%	100%	100%
ㅍ	100%	100%	100%	100%	100%
ㅎ	0%	100%	100%	100%	100%
ㅏ	100%	100%	100%	0%	0%
ㅑ	100%	100%	100%	0%	0%
ㅓ	100%	100%	100%	100%	100%
ㅕ	100%	100%	100%	100%	100%
ㅗ	0%	0%	100%	100%	100%
ㅛ	0%	0%	100%	100%	100%
ㅜ	0%	0%	100%	100%	100%
ㅠ	0%	0%	100%	100%	100%
ㅡ	0%	100%	100%	100%	100%
ㅣ	100%	100%	100%	100%	100%
ㅐ	100%	100%	100%	100%	100%
ㅒ	0%	100%	100%	100%	100%
ㅔ	100%	100%	100%	100%	100%
ㅖ	100%	100%	100%	100%	100%

III. 실험 결과

본 시스템을 개발하기 위한 환경으로는 운영체제는 Windows 7 Enterprise K 32Bit OS를 사용하였으며, 개발 툴은 Visual Studio 2010, MFC환경과 OpenCV 2.4.9

를 사용하였다. 하드웨어로는 SoftKinetic사의 DS325 단거리 적외선 조명 카메라를 사용하였다. DS325는 단거리 카메라라는 특성 때문에 주변 환경에 쉽게 영향을 타지 않는다. 단, 주위에 다른 적외선 카메라가 있다면 잡음이 발생하는 경우는 있다. 컴퓨터 사양으로는 Intel사의 i5-3470 CPU와 3.47GB 메모리에서 개발 되었다. 하드웨어 및 소프트웨어 사양으로 손 모양인식은 카메라로부터 최소거리는 11cm 부터 최대거리는 30cm까지 가능하고, 최적거리는 20cm로부터 ±5cm안에서 제스처 인식 성능이 가장 높았다.

실험으로는 그림 13에서와 같이 자음 14개의 템플릿, 모음 14개의 템플릿으로 총 28개의 템플릿에 대해 실험하였다.

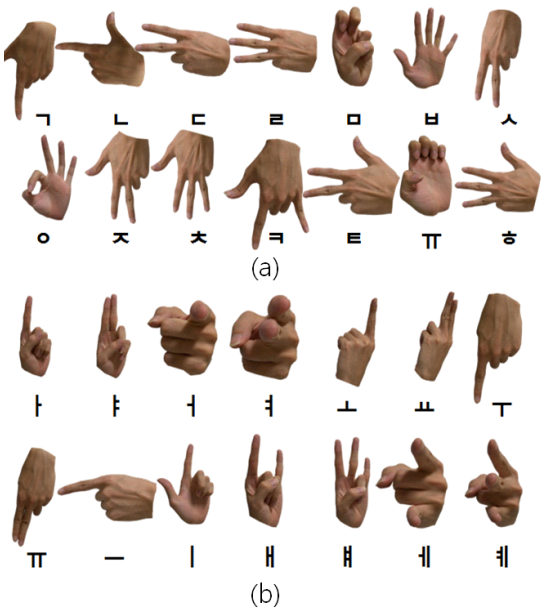


그림 13. 손 모양 인식에 사용된 템플릿. (a) 자음, (b) 모음
Fig. 13. Sample Templetes for Hand Posture Recognition. (a) Consonant, (b) Vowel

실험결과, 한 프레임을 인식하는데 걸리는 시간은 다음 프레임을 받아오기 전에 처리하기 때문에, 실시간으로 처리가 가능하다. 하지만, 현재까지 깊이 카메라는 잡음을 포함하여 들어오는 문제가 있다. 이러한 문제가 있기 때문에, 프레임 중간에 오 인식이 일어나기도 한다. 이러한 노이즈로 인한 오 인식률은 10% 이하이다. 이러한 문제를 해결하기 위해 본 시스템에서는 7프레임을 보고 5프레임 이상이 동일한 표현을 인식하게 되면, 그때 화면에 입력하게 된다. 이러한 시스템에서 사용자가 제대로 제스처를 취했을 경우, 제스처 인식률은 100%의 인식률을 보였다.

본 시스템에서는 기존의 방법과는 달리 인식하는데 있어 회전에 강건한 것을 보였다. 이전에 'ㄱ' 등 특정 제스처는 인식하는데 있어 한쪽으로 쏠리는 문제 발생하는데 반하여, 본 시스템에서는 정 중앙을 기준으로 -10°~10° 사이의 회전된 제스처에 대해서는 아무런 문제가 되지 않았다. 더욱 한계를 계산하기 위해 본 실험에서는 -25°~25° 사이의 회전을 측정하였으며, 회전 중심은 손의 회전 방향을 중심으로 측정하였고, X축과 일치 하였을 경우를 0도로 설정하고 테스트 하였다. 기존의 방법과 달리 SVM을 통해 나눈 결과, 각도가 한쪽으로 틀어지는 문제가 없어지는 것을 확인 할 수 있었다. 표 2는 이에 대한 인식 결과에 대해 나타낸다.

표 2. 제안하는 방법을 사용하였을 때, 각도에 따른 제스처 인식률
Table 2. Hand Posture Recognition along Degrees using Proposed Method.

제스처	-25°	-10°	0°	10°	25°
ㄱ	75%	100%	100%	100%	75%
ㄴ	75%	100%	100%	100%	75%
ㄷ	100%	100%	100%	100%	100%
ㄹ	100%	100%	100%	100%	100%
ㅁ	50%	100%	100%	100%	50%
ㅂ	100%	100%	100%	100%	100%
ㅅ	100%	100%	100%	100%	100%
ㅇ	25%	100%	100%	100%	25%
ㅈ	100%	100%	100%	100%	100%
ㅊ	100%	100%	100%	100%	100%
ㅋ	100%	100%	100%	100%	100%
ㅌ	25%	100%	100%	100%	25%
ㅍ	100%	100%	100%	100%	100%
ㅎ	100%	100%	100%	100%	100%
ㅏ	75%	100%	100%	100%	75%
ㅑ	75%	100%	100%	100%	75%
ㅓ	100%	100%	100%	100%	100%
ㅕ	100%	100%	100%	100%	100%
ㅗ	75%	100%	100%	100%	75%
ㅛ	75%	100%	100%	100%	75%
ㅜ	75%	100%	100%	100%	75%
ㅠ	75%	100%	100%	100%	75%
ㅡ	75%	100%	100%	100%	75%
ㅣ	75%	100%	100%	100%	75%
ㅐ	100%	100%	100%	100%	100%
ㅒ	100%	100%	100%	100%	100%
ㅔ	100%	100%	100%	100%	100%
ㅖ	100%	100%	100%	100%	100%

그림 14은 '안녕'를 입력하고 인식하여 글자를 직접 화면에 뿌려주는 실험 결과를 나타낸다.

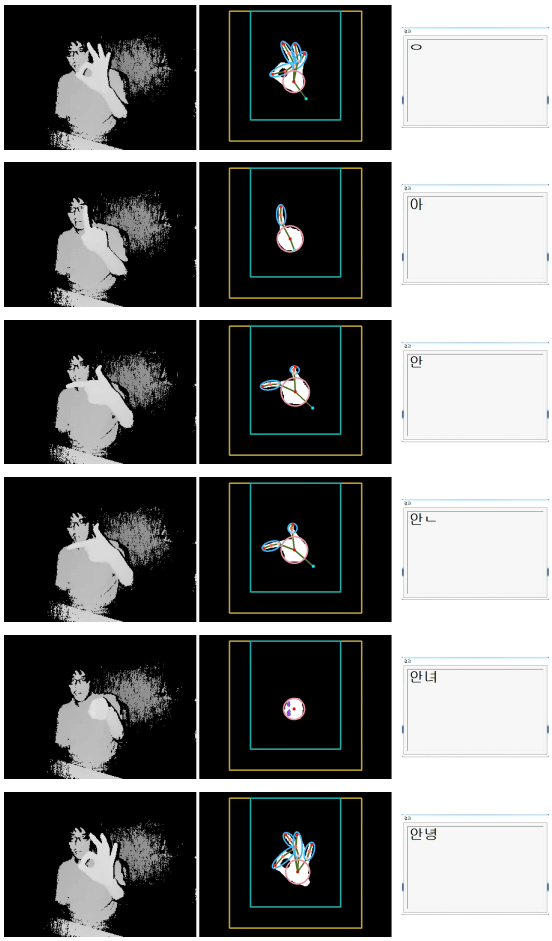


그림 14. 손 모양 인식 실험 결과 예제
Fig. 14. Example of Hand Gesture Recognition

IV. 결 론

본 논문에서는 손 모양 인식을 위해 두 가지 방법을 통해 손의 특징을 추출하고, 결정 트리를 통해 손 모양을 인식하는 방법을 제안하였다. 손의 특징을 추출하기 위해 손가락 방향이 카메라를 향하지 않을 경우, 세선화를 통해 특징 추출하는 방법을 제안하였다. 이 방법을 사용할 경우 기존의 테두리 기반 손가락 특징 검출에 비해 보다 정확하며, 손가락이 모여 있더라도 손끝 추출에는 문제없음을 보였다. 또 다른 특징 방법으로는 깊이값을 이용한 특징 추출 방법을 제안하였다. 이 방법은 손가락 방향이 카메라를 향할 경우, 세선화를 통해 특징을 추출 할 수 없으므로 그 대안으로 깊이 값을 사용하였다. 인식하기 위해서 결정 트리를 사용하였다. 기존의 방법인 사용자가 직접 값을 주어주면 정확한 결정 트리를 생성할 수 없어 SVM 학습을 통해 정확한 분기점을 추출하고 결정 트리를 생성하였다. 실험결과에서는 기존의 결정 트리에 비하여 SVM을 통한 결정 트리를 생성하였을 경우 보다 정확한 중심각을 통해 안정적인 인식 결과를 나타냄을 보였다. 향후 연구로 손 모양뿐만 아니라 동작을 통해 좀 더 수월한 입력 방법에 대한 연구가 필요하다고 생각된다.

참고문헌

[1] Zaki, Mahmoud M., and Samir I. Shaheen. "Sign language recognition using a combination of new vision based features." Pattern Recognition Letters, Vol. 32, No. 4 pp. 572-577, March 2011.

[2] Yang, Ruiduo, and Sudeep Sarkar. "Coupled grouping and matching for sign and gesture recognition." Computer Vision and Image Understanding, Vol. 113, No. 6, pp. 663-681, June 2009.

[3] Juang, Chia-Feng, Shih-Hsuan Chiu, and Shen-Jie Shiu. "Fuzzy system learned through fuzzy clustering and support vector machine for human skin color segmentation." IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Vol. 37, No.6, pp. 1077-1087, November 2007.

[4] Juang, Chia-Feng, and Ksuan-Chun Ku. "A recurrent fuzzy network for fuzzy temporal sequence processing and gesture recognition." IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 35, No. 4 pp. 646-658, August 2005.

[5] Yang, Ming-Hsuan, Narendra Ahuja, and Mark Tabb. "Extraction of 2d motion trajectories and its application to hand gesture recognition." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 8 pp. 1061-1074, August 2002.

[6] Nolker, Claudia, and Helge Ritter. "Visual recognition of continuous hand postures." IEEE Transactions on Neural Networks, Vol. 13, No.

4 pp. 983-994, July 2002.

〔7〕 Fang, Yikai, et al. "A real-time hand gesture recognition method." IEEE International Conference on Multimedia and Expo, pp. 995-998, July 2007.

〔8〕 Suryanarayan, Poonam, Anbumani Subramanian, and Dinesh Mandalapu. "Dynamic hand pose recognition using depth data." Pattern Recognition (ICPR), 20th International Conference on. IEEE, pp. 3105-3108, August 2010.

〔9〕 Ren, Zhou, Junsong Yuan, and Zhengyou Zhang. "Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera." Proceedings of the 19th ACM international conference on Multimedia. ACM, pp. 1093-1096, November 2011.

〔10〕 Zabulis, Xenophon, Haris Baltzakis, and Antonis Argyros. "Vision-based hand gesture recognition for human-computer interaction." The Universal Access Handbook. LEA, pp. 1-56, June 2009.

〔11〕 Zafrulla, Zahoor, et al. "A novel approach to american sign language (asl) phrase verification using reversed signing." IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 48-55, June 2010.

저 자 소 개

김 기 상

2007: 숭실대학교 컴퓨터학부 공학사.
2009: 숭실대학교 미디어학과 공학석사.
현 재: 숭실대학교 미디어학과 박사과정
관심분야: 컴퓨터비전, 영상처리
Email : illusion1004@gmail.com

최 형 일

1972: 연세대학교 전자공학과 공학사.
1982: 미시간대학교 전자공학과 공학석사.
1987: 미시간대학교 전자공학과 공학박사
현 재: 숭실대학교 미디어학과 교수
관심분야: 컴퓨터공학, 컴퓨터비전, 패턴인식, 영상처리, 기계학습
Email : hic@ssu.ac.kr