

한글 수화의 실시간 인식 시스템의 구현

Implementation of Real-time Recognition System for Korean Sign Language

한영환

상지대학교 컴퓨터정보공학부

Young-Hwan Han(yhhan@sangji.ac.kr)

요약

본 논문은 복잡한 배경에서 표식없는 손을 추적하여 한글 수화를 인식하는 시스템에 관한 것이다. 제안한 방법은 먼저 인접한 프레임간의 차 영상에 대하여 엔트로피를 측정한다. 큰 값을 갖는 영역에 대해 피부색에 가까운 분포를 갖는 색상 정보를 추출하여 배경 영상으로부터 손 영역만을 추출한다. 추출된 손 영역에 대해 윤곽선을 검출하고 개선된 무게중심 프로파일(centroidal profile) 방법을 적용하여 수화를 인식한다.

6가지 모양의 수화를 사용한 실험결과 기존의 방법들과 달리 표식을 사용하지 않고도 복잡한 배경과 조명의 변화에서 안정적으로 수화를 인식할 수 있다. 또한 초당 15프레임 정도의 처리속도로도 각 사람별로 95%이상, 각 수화별로 90~100%의 인식률을 얻을 수 있다.

■ 중심어 : | 수화 | 엔트로피 | 프로파일 | 제스처 인식 |

Abstract

In this paper, we propose recognition system which tracks the unmarked hand of a person performing sign language in complex background. First of all, we measure entropy for the difference image between continuous frames. Using a color information that is similar to a skin color in candidate region which has high value, we extract hand region only from background image. On the extracted hand region, we detect a contour and recognize sign language by applying improved centroidal profile method.

In the experimental results for 6 kinds of sing language movement, unlike existing methods, we can stably recognize sign language in complex background and illumination changes without marker. Also, it shows the recognition rate with more than 95% for person and 90~100% for each movement at 15 frames/second.

■ Keyword : | Sign Language | Entropy | Profile | Gesture Recognition |

1. 서론

인간의 사고나 인지, 생활의 측면에서 언어의 역할은

절대적이다. 그래서 모든 사람은 언어를 본능적으로 습득해 나간다. 사람은 언어를 습득해 사용하면서 해당 문화나 국가에 대한 정체성을 확립해 나간다. 구어를 통한

언어 습득이 어려운 경우 농아인은 농아인만의 독특한 언어인 수화를 통해 사고하고 학습한다[1].

농아인이 의사소통에 사용하는 수화는 한손 또는 양손을 다양하게 사용해서 자기의 의사를 표현하며 수화와 지화가 있다. 수화는 동적인 손 제스처로 손의 움직임과 손의 모양에 의하여 의미를 전달하며, 지화는 손의 형태에 의하여 한글 자모를 표현하는 지문자의 조합에 의하여 글자를 표현하는 것이다[2]. 다양한 의사 전달을 위하여 농아인은 수화와 지화를 섞어서 사용하며, 한글 표준 수화에서는 명확한 의사 전달을 위하여 조사나 어미 등을 지문자를 이용하여 표현한다. 각 나라마다 자기 나라의 생활 관습에 기초한 표준화된 수화를 정해서 사용하고 있다. 그러므로 농아인의 의사소통 수단으로서 수화의 실시간 인식 시스템이 요구된다.

수화 분류과정을 위한 데이터 수집에는 기구화된 글러브를 기반으로 하는 방법과 비디오를 기반으로 하는 방법이 있다. 비디오를 기반으로 하는 접근 방법은 사용자가 컴퓨터에 연결되지 않아도 되므로 더 자연스러운 인터페이스가 가능하다. Vogler와 Metaxas[3]는 전자기 추적 시스템과 함께 3대의 카메라를 사용하여 미국 수화(ASL)를 위한 시스템을 제안하였다. 53개의 수화 단어를 97개의 시험 문장으로 사용하여 92.1%에서 95.8%의 인식률을 얻었다. Staner[4] 등은 하나의 카메라를 사용하여 40개의 부호로 구성된 단어를 인식하였다. 이 방법은 은닉 마코프 모델(HMM)을 사용하여 384개의 문장에 대해 74.8%에서 97.8%의 인식률을 얻었다. McGuire[5] 등은 카메라와 글러브에 기반한 시스템과 은닉 마코프 모델을 사용하여 미국 수화 번역기에 대해 연구하였다. 5개의 소형 2축 가속도계를 사용하여 관절의 굽힘 정보를 얻어 141개의 수화 단어로 94%의 인식 결과를 얻었다. Brashear[6] 등은 복합 센서를 사용한 이동 보조 기술에 대해 연구하였다. 모자에 부착된 카메라와 손목과 몸통에 부착된 가속도계를 사용하여 90.5%에서 95.3%의 수화를 인식하였다.

본 논문에서는 시각을 기반으로 하면서도 배경에 대한 제약이나 표식을 사용하지 않고 수화를 인식하는 시스템을 개발하고자 한다. 제안한 방법은 인접한 프레임간의 차 영상에 대하여 엔트로피를 측정하여 복잡한 배경 영

상으로부터 손 영역을 분리하는 방법을 제안하고 개선된 무게중심 프로필 방법을 사용하여 수화를 인식한다.

본 논문의 구성은 다음과 같다. 2장에서는 연속적으로 입력되는 컬러 영상에서 엔트로피 분석을 기반으로 움직임 검출과 손 영역을 추출하는 방법에 대해 설명하고, 3장에서는 추출된 손 영역에서 수화를 인식하기 위해 사용한 무게중심 프로필방법에 대해 기술한다. 4장에서는 제안한 수화 인식방법에 대한 실험 및 결과를 보이고 마지막으로 5장에서 결론을 맺는다.

II. 움직임 검출과 손 영역 추출

본 논문에서 제안한 수화 인식을 위한 블록도는 [그림 1]과 같다. 수화를 인식하기 위해서 먼저 수행 되어야 할 것은 인식의 대상이 되는 손 영역을 배경으로부터 분리하고 추적하는 것이다.

손 영역을 검출하는 방법은 크게 다음의 네 가지 방법들로 대별된다. 첫 번째는 컬러영상을 이용하는 방법으로[7-8], HSI, YIQ, 정규화된 RGB 컬러 모델 등에 의해 색상과 유사한 색을 영상에서 찾아 손 영역을 추출한다. 그러나 이방법은 보통의 그레이 영상보다 3배 이상의 데이터가 사용되어 처리량이 많아지게 된다. 두 번째는 광류(optical flow)[9], 차(difference)영상[10]과 같은 프레임간의 차이를 이용하는 방법이다. 그레이 영상만을 이용하므로 처리해야하는 데이터양이 적으나 움직임이 없는 경우 손 영역을 검출하기 어려운 단점이 있다. 세 번째는 모델링을 이용하는 방법[11]이다. 그레이 영상만으로 충분하고 손 영역의 움직임이 없어도 손 영역을 검출할 수 있으나 손 모양에 따라 여러 가지 모델이 필요하며 처리 시간이 길다는 단점이 있다. 네 번째는 위의 세 가지 방법을 복합적으로 사용하는 방법으로 컬러 정보와 동영상상을 복합해서 사용한다[12-13].

본 논문에서 제안한 방법은 복잡한 배경을 가진 영상으로부터 엔트로피를 이용하여 움직임 영역을 추출하고 움직임이 검출된 영역내에서 피부색에 해당하는 컬러 정보를 얻어 손 영역을 추출한다.

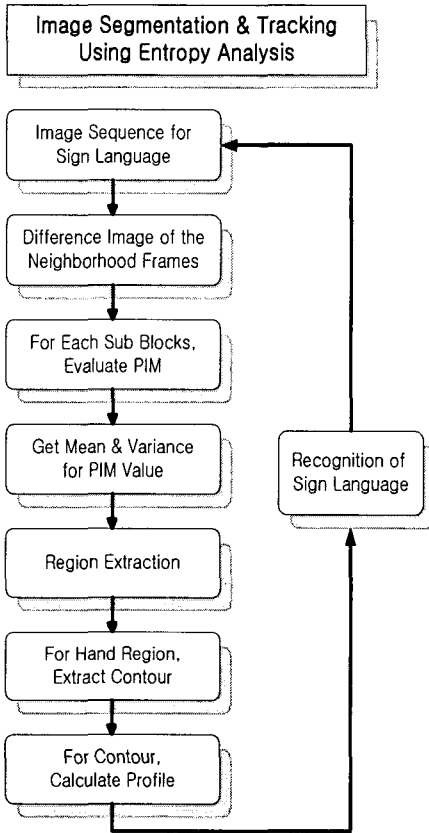


그림 1. 수화 인식 방법의 블록도

먼저, 실시간으로 입력되는 연속적인 영상들 간에 차영상을 구하여 이 차영상을 8×8의 작은 블록으로 분할한다. 분할된 각 블록에 대해 엔트로피를 측정하여 큰 값을 갖는 영역을 움직임이 존재하는 후보영역으로 간주한다. 후보영역 내에서 피부색에 가까운 분포를 갖는 색상 정보를 추출하여 배경 영상으로부터 손 영역만을 추출한다. 움직임이 존재하는 후보영역 내에서 가변적인 피부색 정보를 얻으므로 조명의 변화나 개인의 피부색 차이가 있는 경우에도 안정적으로 사용될 수 있다. 추출된 손 영역에 대해 윤곽선을 검출하고, 개선된 무게중심 프로필(centroidal profile) 방법을 사용하여 수화를 인식한다.

1. PIM에 기초한 움직임 검출

본 논문에서는 영상 데이터의 특징을 얻기 위해서 엔

트로피를 이용하고, 영상에서 엔트로피를 정상화하기 위해 Chang이 제안한 PIM(Picture Information Measure)[14]을 도입하여 손 부위를 검출한다. PIM을 수식으로 표현하면 다음 식 (1)과 같다.

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

여기서 $h(i)$ 는 각 영상 또는 해당 블록의 히스토그램(histogram)을 의미한다. L 은 그레이 레벨을 의미하며 본 논문에서는 256레벨을 사용하였다. $Max_j h(i)$ 은 가장 큰 히스토그램 값이며 j 는 그대의 값을 나타낸다. PIM값은 해당 블록의 전체 화소수와 가장 큰 도수의 히스토그램 값과의 차이로 계산된다.

블록 내의 화소값들이 모두 같은 경우, 즉 블록의 엔트로피가 '0'인 경우에는 식 (2)와 같이 되어 PIM = 0의 최소값을 갖는다.

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

또한 블록 내에서 화소의 각 레벨 값들이 균일하게 분포하는 경우 즉, 엔트로피가 큰 경우에는 $Max_j h(i)$ 가 작은 값을 갖기 때문에 큰 PIM값을 갖게 된다. 즉, PIM은 해당 블록이 많은 정보를 갖는 경우에는 큰 값을 가지며, 적은 정보량을 갖는 경우에는 작은 값을 갖는다. 또한 정규화 PIM(Normalized PIM)은 식 (3)과 같이 정의된다. 일반적인 형태의 PIM_k 과 $NPIM_k$ 은 식(4), 식(5)와 같이 정의된다.

$$NPIM = 1 - Max_j [P(h(i))] \quad (3)$$

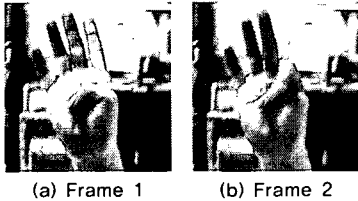
$$PIM_k = \sum_{i=0}^{L-1} h(i) - \sum_{k \text{ largest } h(i)} h(i) \quad (4)$$

$$NPIM_k = 1 - \sum_{k \text{ largest } h(i)} P(i) \quad (5)$$

여기서 P 는 확률값을 나타내며 식(4)와 식(5)에서 k 는 그레이 레벨을 의미한다.

[그림 2]는 복잡한 배경을 가진 두 입력 영상에 대하여

차영상을 구하고 이 차영상에 PIM을 적용하여 손 영역만을 분리(segmentation)한 결과를 보여준다.



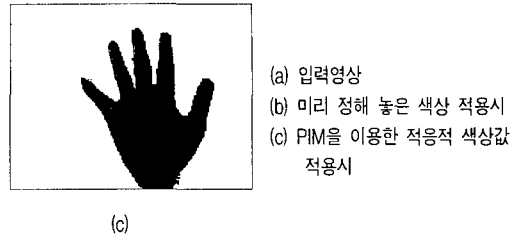
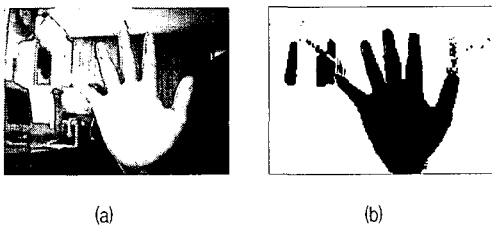
(a) Frame 1 (b) Frame 2
(c) 손 영역 추출 결과
그림 2. PIM을 이용한 손영역 추출

2. 손 영역 추출

엔트로피가 큰 영역은 움직임이 존재하는 후보영역에 대한 정보를 제공한다. 따라서 엔트로피를 측정하여 큰 값을 갖는 영역에서 피부색에 관한 가변적인 색상 정보를 추출한다. 이 정보에 의해 배경 영상으로부터 손 영역만을 추출한다. 엔트로피가 큰 영역은 가변적인 색상 정보를 제공해 주므로 조명의 변화나 개인의 피부색 차이가 있는 경우에도 안정적으로 사용될 수 있다.

살색에 대한 정보를 이용하기 위해 YIQ, 정규화된 RGB, HSI 컬러 모델 등 다양한 컬러 모델이 사용되고 있다. 본 논문에서는 영상처리 알고리즘을 개발하는데 유용한 HSI 컬러 모델을 사용한다.

PIM을 이용하여 PIM 값이 큰 영역에 대해 색상(hue) 값을 적응적으로 적용하므로 좀 더 정교한 손 영역을 추출할 수 있다. [그림 3]은 PIM을 이용한 적응적 색상값을 사용하여 손 영역을 추출한 결과를 보여주고 있다.

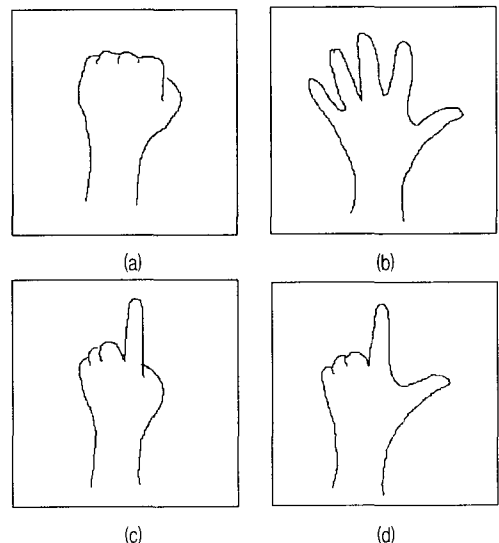


(a) 입력영상 (b) 미리 정해 놓은 색상 적용시 (c) PIM을 이용한 적응적 색상값 적용시
그림 3. 엔트로피를 사용한 적응적 색상값 적용 결과

III. 수화 인식을 위한 무게중심 프로파일

수화를 인식하기 위해 손 영역을 추출한 후 그 영역에 대한 윤곽선을 검출하고 개선된 무게중심 프로파일(centroidal profile) 방법을 사용하였다. 서로 다른 6가지 모양의 수화를 대상으로 인식을 수행하였으며 모양은 [그림 4]와 같다.

[그림 4]의 6가지 손 모양은 한국 수화에서 사용하는 것 중에서 몇 가지를 선택한 것이다. (a)는 알파벳 S와 익다(얼매맷다), (b)는 닳다(모방하다) (c)는 한글 모음 ㅏ와 내일을 각각 나타내는데 사용한다. (d)는 알파벳 L, 아니요(부정)와 해(태양), (e)는 무지개와 배구, (f)는 한글 자음 ㅂ과 대학교를 표시하는데 각각 사용한다[15]. 같은 손 모양이 상황에 따라 알파벳, 동사, 한글 자음, 한글 모음, 명사 등으로 사용되고 있음을 알 수 있다.



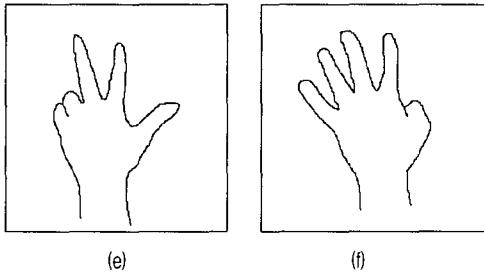
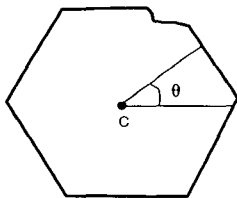


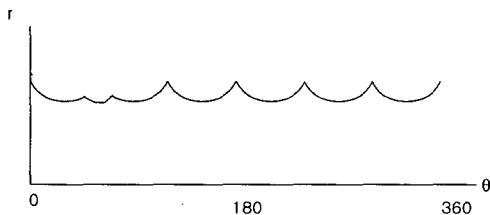
그림 4. 6가지의 손 모양

1. 무게중심 프로파일(centroidal profile)

무게중심 프로파일[16]은 검출된 물체 영역에 대해 물체의 외곽을 각도로 표현함으로써 물체의 형태를 판단하는데 사용된다. 대부분 이러한 형태의 알고리즘은 물체의 무게중심을 구하고 무게중심으로부터 각도에 따른 거리(r) 값으로 표시한다. [그림 5]는 물체 인식을 위한 무게중심 프로필을 보여준다. 물체의 형태는 무게중심 프로필을 구한 후 이상적인 물체에 대한 프로필과 비교하여 파악된다.



(a) 한 코너가 손상된 육각형 모양



(b) 무게중심 프로파일

그림 5. 물체 인식을 위한 무게중심 프로파일

2. 수정된 무게중심 프로파일

물체의 외곽선이 [그림 6]과 같은 경우에는 한 각도에 대해 여러 가지 값을 가지므로 정합시 문제가 된다. 이러한 문제를 해결하기 위한 방법으로 주어진 각도(θ)에 대해 무게중심으로부터 제일 가까운 값을 취한다. 그러나

손의 경우는 이러한 형태로 표현할 수 없다. 따라서 각도를 사용하지 않고 윤곽선 상에 있는 모든 화소들을 대상으로 한 무게중심 프로필을 제안한다.

본 논문에서는 윤곽선 검출을 위해 8 방향 코드를 사용하였다. 윤곽선 위에 존재하는 모든 점들 (x_i, y_i)에 대해 무게중심 (\bar{x}, \bar{y})으로부터 거리(r)를 (식 6)에 의해 구한다. 이진 영상에서 각 화소에 대한 밝기값(intensity)을 그 점에 대한 무게로 간주하면 면적의 중심은 무게중심과 같게 된다. 그러므로 무게중심은 (식 7)에 의해 구해진다.

$$r = \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2} \tag{6}$$

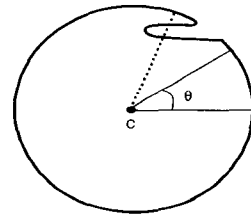


그림 6. 무게중심 프로필이 다중값을 갖는 경우

$$\bar{x} = \frac{\sum_{i=1}^N \sum_{j=1}^M j I(i, j)}{A}, \quad \bar{y} = \frac{\sum_{i=1}^N \sum_{j=1}^M i I(i, j)}{A} \tag{7}$$

여기서 $I(i, j)$ 는 화소의 밝기값을 의미한다. A 는 영상에서 물체의 전체 화소수를 뜻하며 (식 8)과 같이 표현된다.

$$A = \sum_{i=1}^N \sum_{j=1}^M I(i, j) \tag{8}$$

[그림 3] (c)의 손 모양에 대해 무게중심 프로필을 구해 보면 [그림 7]과 같다.

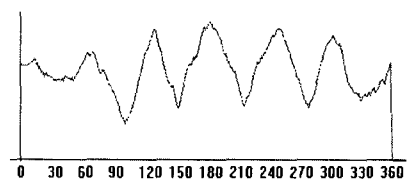


그림 7. 그림 3(c)의 무게중심 프로파일

IV. 실험 및 결과

본 논문에서의 실험은 PC 상에서 C/C++ 언어를 이용하여 구현하였다. 영상 입력 장치로는 저가형의 USB 컬러 카메라를 사용하였으며 영상의 크기는 320×240이다.

[그림 8]은 PC상에서 구현한 수화 인식 시스템의 화면이다. 연속적으로 입력되는 영상으로부터 엔트로피를 분석하여 손 영역을 추출하고 이 추출된 영역에 대해 8방향 코드를 적용하여 윤곽선 검출한다. 검출된 윤곽선을 기반으로 무게중심 프로필을 적용하여 수화를 인식하는 것을 보여주고 있다.

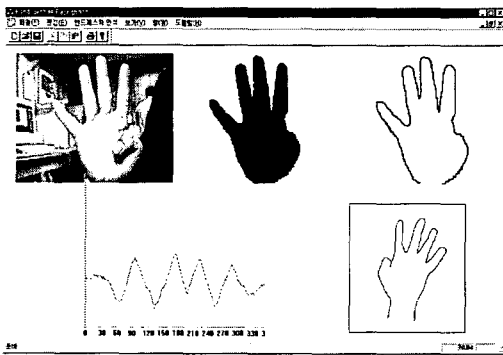
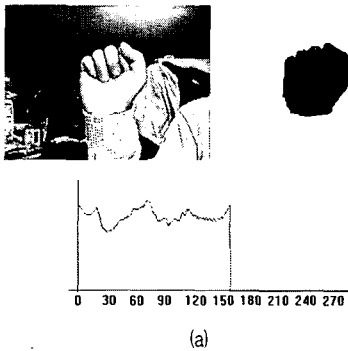
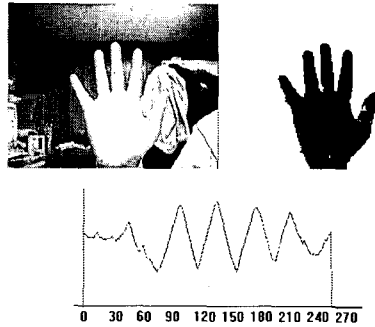


그림 8. 수화인식 시스템

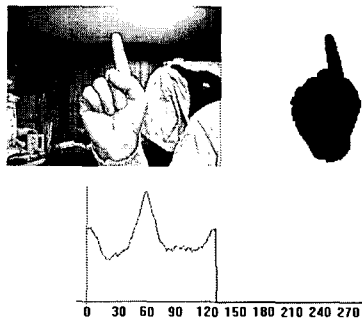
6가지의 수화동작에 대해 제안한 방법을 적용하여 손 영역 추출 및 무게 중심 프로필을 구한 결과는 [그림 9]와 같다. [그림 9]의 결과로부터 무게중심 프로필의 최대값 분포를 이용하면 손 모양을 쉽게 인식할 수 있음을 알 수 있다.



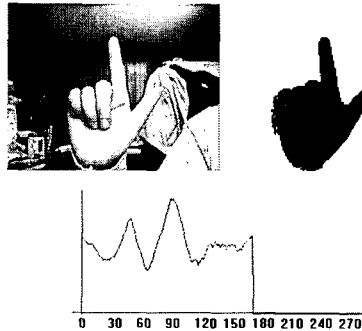
(a)



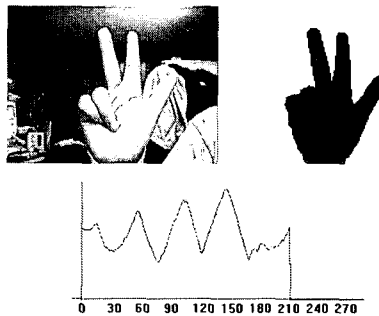
(b)



(c)



(d)



(e)

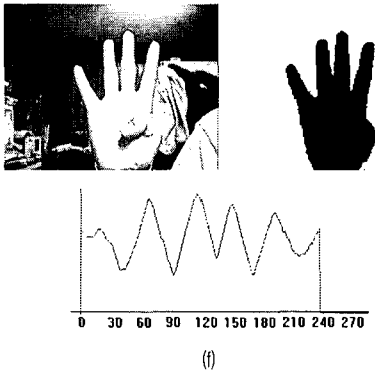


그림 9. 6가지 수화동작에 대한 영역 추출 및 무게중심 프로파일

사람마다 손동작에 있어 다소간의 차이가 존재하며 피부색도 차이가 나므로 PIM을 이용한 적응적 알고리즘의 타당성을 보이기 위해 실험을 행하였다. 6명의 실험자를 대상으로 6가지 손 모양 각각에 대하여 20개의 입력 영상을 대상으로 한 인식결과는 [그림 10]과 같다.

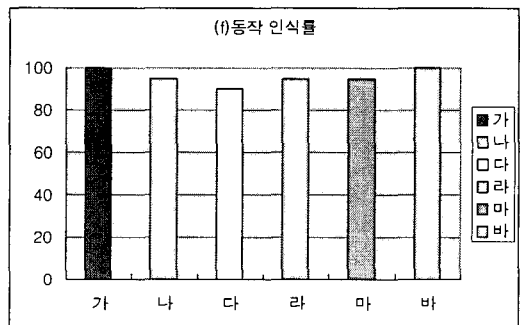
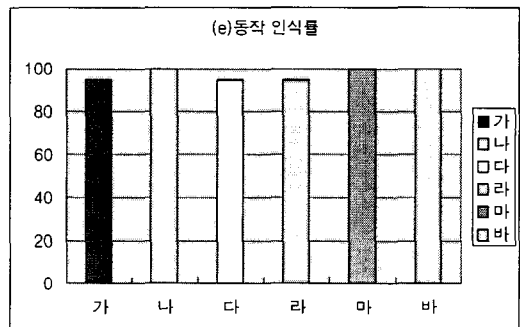
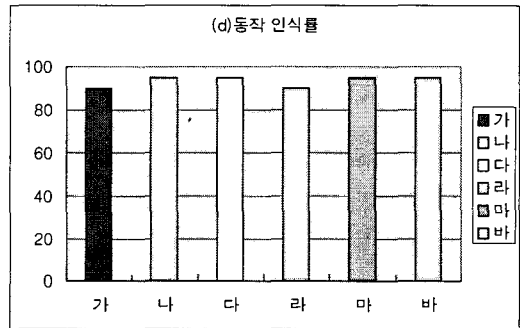
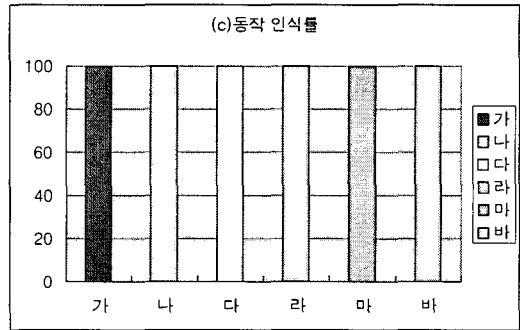
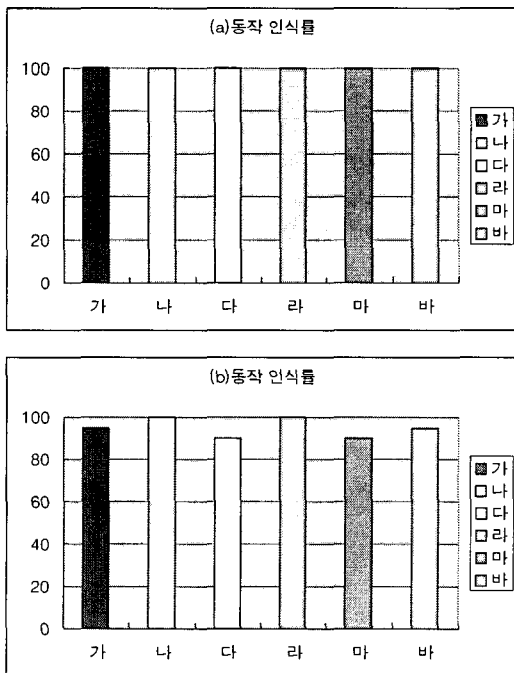


그림 10. 6가지 동작의 인식률

본 논문에서 제시한 방법에 의한 실험환경은 배경에 대한 특별한 제한이 없으며 별도의 표식도 사용하지 않았다. 또한 손동작의 크기와 방향에 따른 영향을 거의 받지 않고 수화동작을 인식할 수 있다. 제안한 알고리즘을 이용한 전체적인 수화동작의 인식률은 [표 1]과 같다. 오인식은 표에서 알 수 있듯이 주로 인접한 손 모양에 의해 발생하게 된다. 이는 인접한 손가락 사이의 잡음 발생으로 인해 정확하게 손 영역을 검출하지 못한 데 기인한다.

을 얻었으며 수화에 사용되는 모양별로 90~100%의 인식률을 얻을 수 있다.

이 시스템은 기본적인 6가지 수화뿐만 아니라 더 많은 수화를 추가할 수 있는 충분한 확장성을 가지고 있어 농아인과 일반인 사이에 의사소통을 위한 수화 통역 시스템으로 개발될 수 있을 것이다. 또한 가상현실 환경을 위한 제스처 인식과 컴퓨터와의 자연스러운 인터페이스 개발 등에도 활용 가능할 것으로 예상된다.

표 1. 제안한 알고리즘을 이용한 수화동작의 인식률

손동작 실험자	(a)	(b)	(c)	(d)	(e)	(f)	전체
가	20/20 (100%)	19/20 (95%)	20/20 (100%)	18/20 (90%)	19/20 (95%)	20/20 (100%)	116/120 (97%)
나	20/20 (100%)	20/20 (100%)	20/20 (100%)	19/20 (95%)	20/20 (100%)	19/20 (95%)	118/120 (98%)
다	20/20 (100%)	18/20 (90%)	20/20 (100%)	19/20 (95%)	19/20 (95%)	18/20 (90%)	114/120 (95%)
라	20/20 (100%)	20/20 (100%)	20/20 (100%)	18/20 (90%)	19/20 (95%)	19/20 (95%)	116/120 (97%)
마	20/20 (100%)	18/20 (90%)	20/20 (100%)	19/20 (95%)	20/20 (100%)	19/20 (95%)	116/120 (97%)
바	20/20 (100%)	19/20 (95%)	20/20 (100%)	20/20 (95%)	20/20 (100%)	20/20 (100%)	119/120 (99%)
비고	■	(f)로 오인식	■	(c), (e)로 인식	(d), (f)로 인식	(e)로 인식	■

참고 문헌

[1] 최상배, 안성우, 한국수어의 이론, 서현사. 2003.
 [2] 이찬수 외, "지문자를 포함한 연속된 한글 수화의 실시간 인식 시스템 구현", 대한전자공학회논문지, 98-35C-6-9, pp.76-87, 1998.
 [3] C. Vogler, D. Metaxas, "Adaptive Hidden Markov Model for ASL Recognition by using 3D Computer Vision method", In Proc. of IEEE International Conference on System, Man and Cybernetics, pp.156-161, 1997.
 [4] T. Staner, and A. Pentland et al, "Real Time American Sign Language Recognition using Desk and Wearable Computer Based Video," IEEE Trans. on PAMI, Vol.20, No.12 pp.1371-1375, 1998.
 [5] R. M. McGuire, J. Hernandez-Rebollar et al, "Towards a One-Way American Sign Language Translator," Proc. of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, 2004.
 [6] H. Brashear, and T. Starner et al, "Using Multiple Sensors for Mobile Sign Language Recognition," Proc. of the 7th IEEE International Symposium on Wearable Computers, 2003.
 [7] X. Zhu, J. Yang, and A. Waibel, "Segmenting Hands of Arbitrary Color," Proc Int'l Workshop on Automatic Face and Gesture Recognition,

V. 결론

본 논문에서는 복잡한 영상에서 추출해낸 손 영역으로 부터 수화를 인식하는 새로운 방법을 제안하였다.

제안한 방법은 카메라로 입력되는 영상의 엔트로피 분석을 통하여 자연스러운 형태로 인식을 수행하였다. 기존의 방법들과는 달리 배경의 제약이나 표식을 사용하지 않고 복잡한 배경에서 손 영역만을 추출해 인식에 사용하였다.

개선한 무게중심 프로필을 이용하여 6가지 모양의 수화를 대상으로 실험을 행하였다. 실험결과 초당 15프레임 정도의 인식속도로 각 사람별로 95% 이상의 인식률

pp.446-453, 2000.

[8] 윤호섭, 소정, 배창석, 전병태, 민병우, “손동작 제스처 인식을 이용한 윈도우 관리기 제어”, 제9회 영상처리 및 이해에 관한 워크샵 발표논문집, pp. 264-269, 1997.

[9] 송효섭, 양윤모, “Optical Flow를 이용한 손의 움직임 및 영역 검출”, 한국정보과학회 '98 봄 학술 발표논문집(B), pp.642-644, 1998.

[10] 이형지, 이우규, 정재호, “수화용 손 제스처 인식을 위한 특징 벡터 추출”, 제11회 신호처리합동학술대회, pp.895-898, 1998.

[11] P. Hong, M. Turk and T. Huang, “Gesture Modeling and Recognition using Finite State Machines,” Proc, Int'l Workshop on Automatic Face and Gesture Recognition, pp.544-549, 2000.

[12] 조오용, 김형곤, 고성제, 안상철, “인터랙티브 가상환경을 위한 손 제스처 인식 시스템”, 대한전자공학회 36-s(4), pp.70-82, 1999.

[13] K. Imagawa, S. Lu, and S. Igi. “Color-Based Hands Tracking System for Sign Language Recognition,” in Proc. of IEEE International Conference on Automatic Face and Gesture Recognition, Killington, pp.151-156, 1996.

[14] S. K. Chang, Principles of Pictorial Information Systems Design, Prentice-Hall, 1989.

[15] 강옥, 가장 아름다운 언어 수화, 혜성출판사, 2001.

[16] R. C. Gonzalez, and R. E. Woods, Digital Image Processing, 2nd Edition, Prentice-Hall, pp.295-302, 2002.

저자 소개

한영환(Young-Hwan Han)

정회원



- 1989년 2월 : 인하대학교 전자공학과(공학사)
- 1991년 2월 : 인하대학교 전자공학과(공학석사)
- 1995년 8월 : 인하대학교 전자공학과(공학박사)

• 1996년 3월~현재 : 상지대학교 컴퓨터정보공학부 교수

<관심분야> : 영상처리, 컴퓨터비전, IT 시스템