

립모션을 이용한 수화 인식 툴 개발로 일반인과 청각 언어 장애인들간의 대화와 소통의 장 구성

조재현*, 이동훈*, 하대규*, 김지인*, 이상은**, 고석주*

*경북대학교 컴퓨터학부

**임베디드 소프트웨어 연구센터

e-mail : jyilm79@naver.com

A Development on Sign Language Recognition Tools for Communication with Dyslexic and Hearing-impaired People

Jae-Hyun Jo*, Dong-Hoon Lee*, Dae-gyu Ha*, Ji-In Kim*, Sang-Eun Lee**, Seok-Joo Koh*

*Dept. of Computer Science and Engineering, Kyung-pook National University

**Center for Embedded Software Technology (CEST)

요 약

본 논문은 최근 급변하는 스마트 시대에 발 빠르게 대처하기 힘든 장애인들과 일반인간의 단절을 해소하기 위해 도움을 주고자 한다. 사람들 간의 융합을 위해서는 기본적인 대화가 진행되어야 하는데 청각언어 장애인들과의 대화를 위해서는 수화를 사용하여야 한다. 하지만, 일반적으로 수화 언어에 대한 지식이 미흡하고 원활한 의사소통에 큰 문제점을 가지고 있다. 그러므로 본 논문에서는 립모션을 이용한 수화 인식 툴의 개발을 통해 일반인과 청각언어 장애인들간의 의사소통을 증진시켜주는 것이 목적이다.

1. 서론

최근 급변하는 스마트 시대에 발 빠르게 대처하기 힘든 장애인들과 일반인 간의 단절은 더욱 더 심해지고 있다. 특히, 기본적으로 사람들간의 융합을 위해서는 대화가 이루어져야 하는데 일반인들의 수화 언어 지식의 부족으로 인해 교류에 있어 더욱 큰 문제가 되고 있다. 현재 전국 장애인의 수는 2012년 기준 약 25만명이며, 그 중 청각 및 언어 장애인의 수는 장애인 그룹 중 가장 높은 26만명으로 전체의 11% 비중을 차지하고 있다 [1]. 그러므로 수화 인식 기술이 필요하며 효율적으로 사용될 것으로 전망된다.

하지만, 기존에 수화인식은 키넥트 또는 영상처리 기기를 사용하여 이루어졌다. 이는 세부적인 손동작 감지에 어려움이 있고, 인식률도 떨어지며, 사용자에게 있어 공간적인 제약을 요구한다.

본 논문에서는 손과 손가락 동작을 중점으로 인식하여 세밀한 움직임을 검출할 수 있는 립모션을 활용

한 수화 인식 도구에 대한 연구를 진행하였다. 특히, 립모션 기능 중 skeleton tracking 을 이용하여 손과 손가락의 좌표를 이용한 수화 정보 저장 및 번역하는 알고리즘을 구현하였다.

본 논문의 구성은 다음과 같다. 2절에서는 현재 나와 있는 수화 인식 도구에 대해서 분석하고, 3절에서 립모션을 이용한 수화 인식 도구를 제안한다. 4절에서는 구현한 립모션을 이용한 수화 인식 도구에 대한 실험을 수행하고, 마지막 5절에서 결론을 맺는다.

2. 기존 수화 인식 도구

기존의 수화 인식 도구로는 키넥트를 이용한 수화 인식 도구와 영상기기를 이용한 수화인식 도구가 있다.

1) 키넥트를 이용한 수화 인식 도구

키넥트를 이용한 수화 인식 도구를 살펴보면 키넥트의 skeleton tracking 에서 획득한 인체의 관절 좌표를

이용하여 손의 위치와 움직임의 특징을 검출한다 [1].

각 프레임에서는 손의 위치와 그에 따른 차이를 인식하고 수화에 필요한 동작을 구분하여 정의하는 방식으로 구현하였다.

하지만 이러한 키넥트를 이용한 수화 인식 방법은 전체적인 손 동작만 인식하므로 손가락을 포함한 세부적인 표현을 통한 수화의 인식에 있어서는 제한이 있다.

2) 영상 기기를 이용한 수화 인식 도구

영상 기기를 이용한 수화인식 도구로는 영상기기로 손의 이동에 따른 정보 추출하기가 비교적 힘들기 때문에 영문자와 숫자만 인식하는 알고리즘만이 현재 구현되어 있다 [2].

특히, 이 기법에서는 영문자 인식율을 높이기 위하여 회전변환을 이용하였고 영문자 지화를 다양한 조건에서 사진 촬영하여 윤곽선을 추출하는 방식으로 수화를 번역하였다. 이러한 방법은 정확도가 높을 수는 있지만 한글 수화에서는 사용할 수가 없다.

3. 립모션을 이용한 수화 인식 도구의 개발

1) 립모션

립모션은 높은 정밀도 및 트래킹 프레임률로 시야 내에 관찰되는 객체를 분석한다. 이는 손, 손가락들의 개별 위치, 제스처, 움직임을 모두를 인식한다. 이러한 립모션의 인식 범위는 장치의 중심을 기준으로 역 피라미드 모양이다. 인식 범위는 약 25~700mm 이다. 아래 그림 1 은 이러한 립모션과 그 인식 범위를 보여 준다.

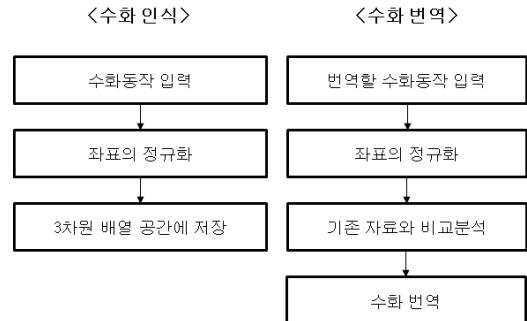


(그림 1) 립모션과 인식 범위

2) 수화 인식 및 번역 알고리즘

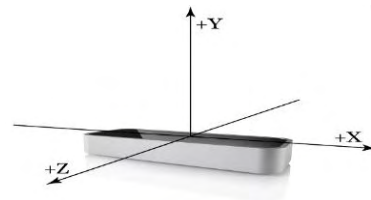
제안하는 수화 인식 알고리즘은 립모션으로 손과 손가락 및 방향 정보 좌표를 받아온다. 그리고 인식된 좌표를 정규화한 후 3 차원 배열 공간에 저장한다.

번역 알고리즘은 립모션 기기로 번역할 수화 동작을 입력 받은 후 기존에 저장된 자료와 비교 분석을 한다. 그리고 가장 유사한 수화동작으로 번역하는 순서를 따른다. 다음 그림 2 는 수화 인식 알고리즘의 흐름도이다.



(그림 2) 수화 인식 알고리즘 흐름도

한 동작의 정보에는 손의 3 차원 공간, 손가락의 3 차원 공간이 있으며, 왼손, 오른손이 구분되어 저장된다. 이동하는 손의 좌표 정보들이 3 차원 공간에 저장되며, 마찬가지로 손가락도 구분이 되어 이동되는 좌표의 정보들이 손가락의 3 차원 공간에 저장된다. 이러한 3 차원 공간의 기준은 그림 3 과 같다.



(그림 3) 립모션이 인식하는 3 차원 공간 기준

3 차원 배열 공간에 수화의 동작들을 저장하고 번역 할 때 가장 큰 문제점은 사람들마다 동작 범위와 위치가 달라질 수 있다는 점이다. 예를 들어, ‘ㄱ’ 이런 동작을 사용자가 한다고 가정했을 때, 한 사람은 크고 넓게 표현할 수도 있으며, 또 다른 사람은 작고 좁게 표현할 수도 있다. 그렇게 되면 서로 같은 뜻을 생각하고 동작을 하였지만, 결국 3 차원 배열 공간에서는 겹치는 부분이 없어 이를 다른 동작으로 인식하는 문제가 발생한다.

이러한 동작 인식에 있어 인식률을 높이기 위해 두 가지 기법의 알고리즘을 적용하였다. 립모션의 동작 인식 반경은 립모션 컨트롤러의 중심부터 반 구 형태

의 모양이다. 이를 그림 3 의 3 차원 공간 기준으로 봤을 때 X 축의 인식반경 범위는 약 70cm 이고 Z 축의 인식반경 범위는 약 60cm 이고 Y 축의 인식반경 범위는 45cm 이다. 이렇게 축마다 인식반경 범위가 다르므로 각 축마다 따로 정규화 하였으며 비교적 립모션 컨트롤러의 중심에서 많은 동작들이 행해지므로 중심에서 세부적이게 좌표들이 인식되도록 정규화 하였다.

정규화에 따라 X, Y, Z 축의 범위가 -1.0 ~ 1.0 으로 표현이 되며, 이 범위를 좌표로 나타내기 위해 0~199 범위로 변환하여 [200][200][200]의 3 차원 공간으로 나타내었다. 이러한 3 차원 공간을 [50][50][50] 크기로 압축하여 기존 인식 반경에서 주변의 좌표를 저장하여 인식률을 높였다.

이전까지의 알고리즘으로 수화의 동작을 인식하였을 때 인식률은 높아졌으나 또 하나의 문제점이 남아 있다. 예를 들면, 오른쪽으로 가는 일자 형태의 동작(→)과 왼쪽으로 가는 일자 형태의 동작(←)이 있을 때 엄연히 다른 뜻을 가진 두 동작이지만 기존의 알고리즘에서는 손과 손가락의 이동만 저장하므로 같은 좌표의 정보에 따라 같은 뜻으로 인식된다.

<표 1> 3 차원 공간에서의 방향성

-1	-1	-1	A
-1	-1	0	B
-1	-1	1	C
-1	0	-1	D
...			
1	1	-1	X
1	1	0	Y
1	1	1	Z

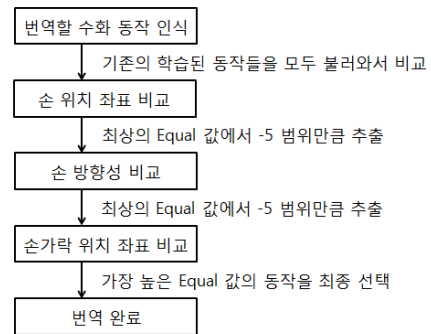
이를 해결하기 위해 동작의 방향성을 파악하는 알고리즘을 추가하였다. 표 1 을 살펴보면 3 차원 공간에서 이동되는 방향성은 총 26 가지이다. 즉, (-1, -1, -1) 위치에서 (0, 0, 0) 위치로 가면 방향성 'A'를 가지고, (-1, -1, 0) 위치에서 (0, 0, 0) 위치로 가면 방향성 'B'를 가진다. 이렇게 알파벳 문자의 개수에 맞게 총 26 가지의 방향성을 3 차원 공간을 새로이 추가하여 저장하였다.



(그림 5) 방향성 “V”동작 (그림 6) 방향성 “D”동작

그러므로 위의 알고리즘에 따라 앞서 그림 5 의 오른쪽으로 가는 일자 형태의 동작(→)은 (-1, 0, 0)에서 (0, 0, 0)으로 갈 경우 방향성은 ‘V’를 가지고, 그림 6 의 왼쪽으로 가는 일자 형태의 동작(←)은 (1, 0, 0)에서 (0, 0, 0)으로 갈 경우 방향성은 ‘D’를 가지게 되어 손의 이동 좌표는 같지만 서로 방향성은 달라 다른 뜻의 동작으로 인식된다.

수화의 번역을 위해서는 앞의 저장기법 알고리즘으로 몇 가지 수화의 동작을 저장(학습)한 후 번역 기능을 통해 원하는 동작을 새로 인식하여 저장된 동작들과 비교하여 가장 가까운 동작의 뜻을 찾아야 한다. 그림 7 은 이러한 수화 번역에 대한 알고리즘의 흐름도를 나타낸다.



(그림 7) 수화 번역 알고리즘 흐름도

위 그림과 같이 우선 립모션의 인식에서 손가락 개수에 따라 인식되는 범위와 오차가 크기 때문에 손가락보다는 비교적 손의 이동 인식률이 높다. 그러므로 손가락 보다는 우선적으로 손의 이동 정보를 활용한다. 손의 경우에도 수화 번역에 있어서 위치 정보가 방향성에 우선적으로 적용이 된다. 이는 위치 정보가 방향성 보다 수화 번역에 있어 효율이 좋기 때문이다.

첫째, 번역할 동작의 손의 3 차원 공간과 저장된 동작들의 손의 3 차원 공간을 비교하여 가장 유사한 동작을 찾아낸다. 하지만 상황에 따라 유사 값의 변화가 있을 수 있으므로 가장 높은 손의 유사 값의 동작에서 -5 만큼의 범위의 손의 유사 값의 동작들을 다시 모은다.

둘째, 모은 동작들을 다시 방향성 3 차원 공간 정보와 비교하여 그 중 가장 높은 방향성의 유사 값을 찾는다. 그리고 가장 높은 방향성의 유사 값의 동작에서 -5 만큼 범위의 방향성의 유사 값의 동작을 다시

모으게 되면 모아진 동작들은 손의 방향성 정보들이 번역을 원하는 동작의 정보와 가장 가까워진다.

마지막으로, 모아진 동작들 중 손가락의 3 차원 공간 정보를 비교하여 가장 높은 손가락의 유사 값을 찾으면 번역을 원하는 동작과 손, 방향성, 손가락 정보와 제일 가까운 동작으로 인식되어 뜻을 반환하게 된다.

4. 실험 및 결과

표 2 는 수화의 대표적인 동작 5 가지를 추출하여 인식하는 성공률을 나타낸다 [3].

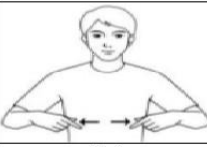
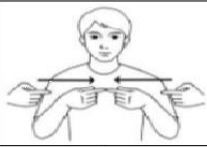
<표 2> 성능 평가를 위한 대표적인 수화의 성공률

	시도횟수	인식성공	인식실패	인식률
보다	100	97	3	97%
가다	100	93	7	93%
증가	100	96	4	96%
감소	100	97	3	97%
우리	100	91	9	91%

인식 성공률을 테스트하기 위해서 대표적인 수화로는 ‘보다’, ‘가다’, ‘증가’, ‘감소’, ‘우리’라는 수화를 채택하였으며, 이를 각각 100 번씩 시도하여 인식 성공 여부를 조사하였다. 테스트에서는 3 명의 사람이 같은 수화를 서로 다른 회수로 수행을 하였으며, 인식률은 ‘우리’라는 수화가 91%로 가장 낮은 인식 성공률을 보였고, ‘보다’와 ‘감소’수화가 최고 97%의 인식 성공률을 보였다. 다섯 가지의 대표 수화 모두를 봤을 때, 평균 94.8%로 높은 인식 성공률을 보였다.

또한, 구현한 수화 인식 알고리즘의 인식률 개선을 알아보기 위해 표 3 과 같은 실험을 하였다.

<표 3> 비슷한 수화 동작의 인식률 개선 과정

	손 위치 좌표 비교	손 방향성 비교	손가락 위치 좌표 비교
 증가	47 / 100 (47%)	95 / 100 (95%)	96 / 100 (96%)
 감소	52 / 100 (52%)	93 / 100 (93%)	97 / 100 (97%)

예를 들어 ‘증가’, ‘감소’ 동작의 경우 비교 동작으로써 손과 손가락의 위치 이동은 똑같으나 방향성만

다른 동작이다. 손의 위치만을 비교했을 경우 각각 47%와 52%의 인식률을 보였으나, 여기에 손의 방향성까지 고려를 하니 인식률이 각각 95%와 93%로 크게 향상되었다. 여기에 손가락의 위치 좌표까지 비교를 하게 되니 인식률이 96%, 97%까지 높아 지는 것을 알 수 있었다. 본 실험을 통해 구현한 손의 방향성을 고려하는 수화 인식 알고리즘이 높은 인식률로 개선 되는 것을 알 수 있다.

5. 결론

본 논문에서는 고가의 영상장비나 키넥트를 사용하는 기존의 수화인식 프로그램에서 벗어나 립모션을 이용한 수화인식 알고리즘 및 프로그램을 구현하였다. 단어의 동작을 모두 입력하는 방식이 아닌 양손의 움직임들 각각 인식함으로써 적은 동작의 인식만으로 많은 의미를 표현하는 수화인식 프로그램을 구현함으로써, 프로그램의 가독성을 높이고 입력된 동작과의 비교 시간을 줄이는 효과가 있다. 또한, 손의 좌표, 손가락 좌표, 손의 진행 방향에 필터링을 통해 반복되는 행동이나 비슷한 수화 동작에 오류를 줄여 인식률 향상에 기여하였다. 하지만 양손의 좌표가 겹쳐지는 경우의 인식률 저하와 원운동의 인식률 저하에 문제가 있어 이에 대한 개선이 필요하다.

그리고 제안된 키넥트를 이용한 수화 인식은 단순한 인체의 관절 정보만을 사용하기 때문에 수화가 아닌 지화의 검출에 한계가 있다. 추후 키넥트의 depth 이미지를 활용해 손가락의 외곽선의 특징을 비교하는 영상처리 프로그래밍을 통하여 지화를 검출하는 기능을 포함한다면 더욱 진보된 수화 인식이 가능할 것으로 보인다.

감사의 글

본 연구는 경북대학교 임베디드 SW 센터-컴퓨터학 부간 URP(Undergraduate R&D Program) 프로그램의 지원으로 수행되었음.

참고문헌

[1] Simon Lang, Marco Block, Raúl Rojas “Sign Language Recognition Using Kinect”, Artificial Intelligence and Soft Computing, Lecture Notes in Computer Science, Vol. 7297, pp.394-402, 2012
 [2] 최두현, 김연준, “손동작을 이용한 수화 인식 알고리즘 연구”, 과학고 R&E 결과보고서
 [3] 국립국어원, 한국수화사전, <http://222.122.196.111>