



## 신경망을 이용한 한국수어 번역 알고리즘 구현

Korean Sign Language Translation using Neural Network

---

저자 (Authors)	강소희, 김용겸, 김유신, 배승용, 박지현, 이정환, 정동화, 임완수 Sohee Kang, Younggyeom Kim, Yousin Kim, Seungyong Bae, Jihyun Park, Jeonghwan Lee, Donghwa Jeong, Wansu Lim
출처 (Source)	<a href="#">대한전자공학회 학술대회</a> , 2018.11, 765-768(4 pages)
발행처 (Publisher)	<a href="#">대한전자공학회</a> The Institute of Electronics and Information Engineers
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07624999">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07624999</a>
APA Style	강소희, 김용겸, 김유신, 배승용, 박지현, 이정환, 정동화, 임완수 (2018). 신경망을 이용한 한국수어 번역 알고리즘 구현. 대한전자공학회 학술대회, 765-768
이용정보 (Accessed)	동신대학교 220.95.42.*** 2020/08/18 22:27 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

# 신경망을 이용한 한국수어 번역 알고리즘 구현

강소희<sup>1</sup>, 김용겸<sup>2</sup>, 김유신<sup>2</sup>, 배승용<sup>2</sup>, 박지현<sup>2</sup>, 이정환<sup>2</sup>, 정동화<sup>2</sup>, 임완수<sup>1</sup>

<sup>1</sup>금오공과대학교 전자공학과

<sup>2</sup>금오공과대학교 컴퓨터소프트웨어공학과

*e-mail : wansu.lim@kumoh.ac.kr*

## Korean Sign Language Translation using Neural Network

Sohee Kang<sup>1</sup>, Younggyeom Kim<sup>2</sup>, Yousin Kim<sup>2</sup>, Seungyong Bae<sup>2</sup>, Jihyun Park<sup>2</sup>, Jeonghwan Lee<sup>2</sup>, Donghwa Jeong<sup>2</sup>, Wansu Lim<sup>1</sup>

<sup>1</sup>Department of Electronic Engineering

<sup>2</sup>Department of Computer Software Engineering

Kumoh National Inst. of Tech.

### Abstract

In this paper, we introduce sign recognition algorithms to help facilitate communication between hearing-impaired using sign language and speech-language users. The algorithm is developed to use Microsoft's depth sensor, kinect, to receive human joint information and to analyze sign language by learning joint information in neural networks. This neural networks is formed models that analyze arm movement and finger, about 88% accuracy was achieved in the test set.

권리를 보장받는다 해도, 음성언어와의 장벽으로 인해 여러 가지 제약을 받고 있는 상황이다. 기존 수어번역 알고리즘은 영어수화언어(American Sign Language, ASL)에 국한되어 있고 한국수어를 문자 및 음성으로 번역하는 알고리즘은 아직 개발되지 않았다 [3,4].

이에 농아인과 음성언어사용자 간의 원활한 의사소통을 위해 본 논문은 한국수어를 문자로 변환해 주는 알고리즘을 구현하고자 한다.

### I. 서론

문화체육관광부의 한국수어(Korean Sign Language, KSL)법에 대한 보도 자료에 따르면, 2014년 후반에 대한민국의 청각장애인은 27만 명을 넘었다[1]. 또한, 한국수화언어법이 2016년 8월부터 시행됨으로서 한국수어는 청각장애인들의 공식 언어로서, 한국어와 동등한 지위를 얻게 되었다[2]. 하지만, 한국수어를 사용하는 청각장애인이 대한민국 국민으로서 언어의 동등한

### II. 시스템 구성

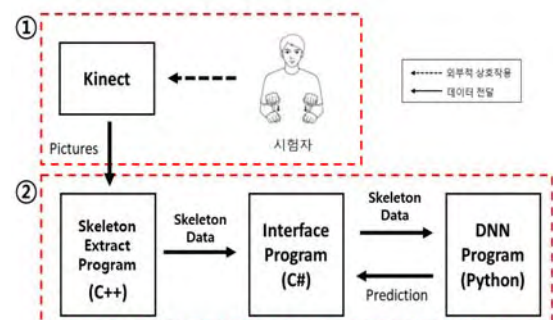


그림 1. 한국수어 번역 시스템 구성

그림 1은 한국수어 번역을 위한 시스템 구성도이며, 수어동작을 입력하는 부분과 수어를 학습하는 부분으로 구성된다. 수어를 입력하기 위해서 마이크로소프트사의 키넥트를 사용하였고(그림1의 ①), 수어의 학습은 특징추출과정 및 DNN을 이용한 머신러닝 기법(그림1의 ②)을 포함한다. 그림 1의 Skeleton Position Extract (관절 위치 추출) 모듈은 수어를 촬영한 영상에서 관절과 손의 위치를 실시간으로 추적하여 저장한다. AI 모듈은 Skeleton Position Extract 모듈에서 받은 관절과 손의 위치 정보를 기반으로 수어 동작을 학습한 후, 실시간으로 입력되는 수어 동작을 예측한다. Interface 모듈은 Skeleton Position Extract 모듈과 AI 모듈 간 데이터를 서로 전달한다.

### III. 제안한 수어번역 알고리즘

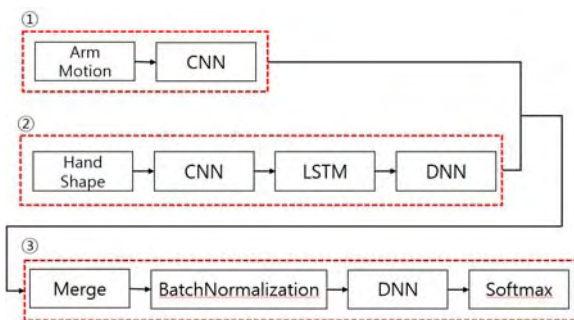


그림 2. 수어번역 알고리즘 흐름도

그림 2는 제안한 수어번역 알고리즘이며, 한국수어는 손 모양과 팔의 움직임의 차이가 중요하기 때문에, 수어의 팔 움직임(Arm motion, 그림 2의 ①)과 손 모양(Hand shape, 그림 2의 ②)을 각각 학습한 후, 두 모델을 합(Merge, 그림 2의 ③)하여 다시 학습하는 알고리즘으로 구성하였다.

#### 3.1 팔 동작 학습 모델

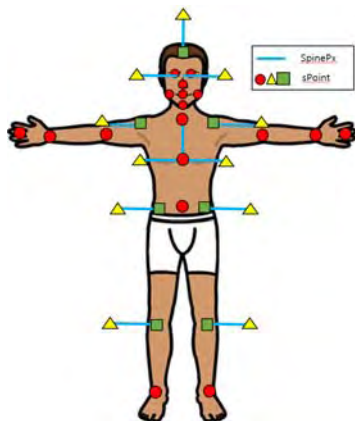
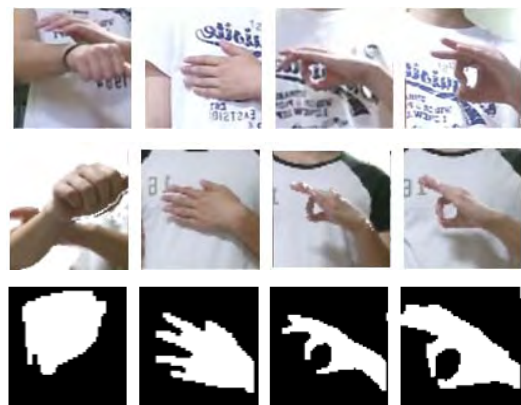


그림 3. 키넥트에서 제공하는 관절 및 얼굴 위치 정보

그림 3은 키넥트에서 제공하는 관절, 손 및 얼굴의 위치 정보를 세모, 동그라미, 네모 모양으로 표시했으며, 이러한 위치 정보를 기반으로 손과 팔의 움직임을 추적한다. 척추 상단부분과 척추 중간부분간의 거리를 측정하여 SpinePx라는 길이로 지정하여 특정 관절에서 SpinePx만큼 떨어진 위치정보 또한 저장한다. 저장한 37개의 모든 점들을 sPoint로 지칭하며, 키넥트에서 기본적으로 제공하는 왼손과 오른손의 위치와 각 sPoint위치 사이의 거리를 측정, 저장해서 손의 동작을 추적함으로써 팔 동작을 추출하도록 한다. 손을 허리 위로 올림으로서 수어의 시작을 알리고 손을 허리 밑으로 내림으로서 수어의 끝을 알려 수어의 전체 동작을 150개의 프레임으로 촬영한다.

저장한 팔의 동작 데이터를 Convolution 연산 레이어에 통과시켜 특징을 추출한 후, MaxPooling 레이어로 추출한 특징 중 가장 큰 값을 찾아 강조한다. MaxPooling 레이어를 통과한 값들을 Flatten 레이어를 통해 이어주고, Dense 레이어로 각 특징에 가중치를 곱해서 중요도를 학습시키는 과정을 거쳐 팔의 동작을 훈련시킨다.

#### 3.2 손 모양 학습 모델



안녕하세요 나 입금 출금

그림 4. 손 모양 전처리 과정

그림 4는 손의 위치를 추적하여 손과 주변 영역의 이미지를 추출한 후, 이미지 전처리를 통해 손의 모양을 뒷배경과 잘 구분할 수 있게 하는 과정이다. 키넥트에서 기본적으로 제공하는 양손의 위치를 중심으로 가로, 세로가 20픽셀인 사각형 영역에서 깊이 값이 최소가 되는 점을 찾는다 [5]. 해당 사각형 범위는 손의 이동속도가 빨라서 손의 중심점을 키넥트가 제대로 추적할 수 없는 경우를 방지하기 위해 설정한 영역이다.

각 픽셀의 깊이 값이 시험자의 신체깊이값을 초과할 경우에는 흰색(255)으로 채워 뒤의 배경을 제거한다. 그 후, 가로, 세로의 픽셀에 해당하는 깊이 값이 임계값을 초과하는 경우, 검은색(0)으로 채우고, 범위 안에 들어올 경우 흰색(255)으로 채워 손의 윤곽을 추출한다. 수어의 전체 동작에서 손의 모양만 35개의 프레임으로 촬영한 후, 촬영을 시작한 후 11개 프레임과 종료전의 5개 프레임을 삭제해서 총 19개의 이미지를 저장한다.

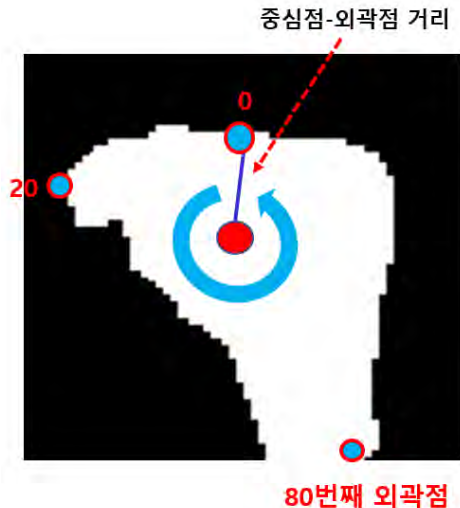


그림 5. a) 손 모양 전처리후 외곽선 거리 측정

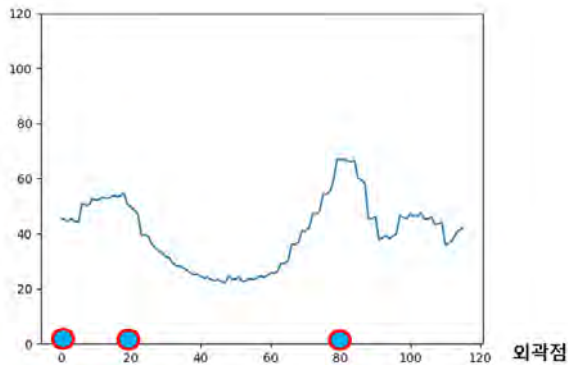


그림 5. b) 손 모양 전처리후 외곽선 거리 측정

추출한 손의 윤곽에서 손의 중심점, 외곽 점의 거리를 측정한다(그림 5 a, b). 손을 허리 밑으로 내린 경우는 수어동작을 하지 않는 것으로 간주하고, 외곽선을 추출하지 않는다. 그림 5의 과정을 통해 손 모양 이미지를 그래프로 도식화 한다. 저장한 손의 모양 데이터를 Convolution 연산 레이어에 통과시켜 이미지의 특징을 추출한 후, Dense 레이어를 통해 각 특징의 중요도를 학습시킨다. 그 후, 시간에 따른 손모양의 연관성을 파악하기 위해 시계열

신경망의 일종인 Long-Short Term Memory(LSTM) 레이어를 통해 시간에 따른 프레임간의 특징을 학습시키고, 다시 Dense레이어를 통해 중요도를 학습시킨다.

### 3.3 특징 학습 모델

Merge레이어를 통해 3.1에서 제안한 모델과 3.2에서 제안한 모델에서 학습된 특징들을 이어준다. 배치 정규화(BatchNormalization)레이어를 통해 가중치가 편향되는 현상을 막아 과적합(Overfitting)을 방지한다. DNN을 통해 처리한 특징들을 훈련시키고, Softmax레이어를 통해 활성화 한다.

## IV. 실험 결과

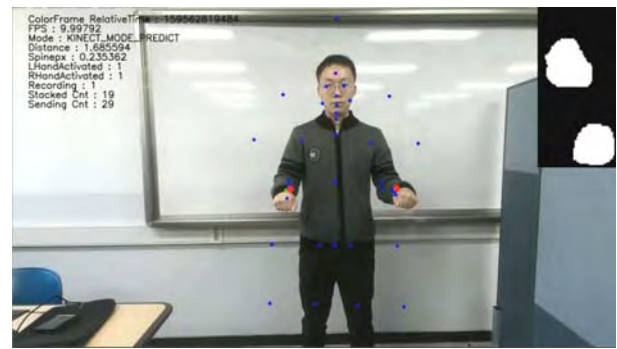


그림 6. a) 실제 수어 시연 장면(안녕하세요)



그림 6. b) 실제 수어 인식 장면(안녕하세요)

그림 6은 실제 수어를 실시간으로 신경망으로 예측한 결과를 보여준다. 붉은색 점은 손의 위치를 표현하며 공간상의 손의 위치를 추적하는 점이다. 푸른색 점은 시험자의 신체나 주변에 고정되어 서로간의 거리를 측정하여 팔의 동작을 예측할 수 있게 하는 점이다. 그림 6의 b는 실제 수어를 인식해서 예측한 결과를 나타내는 장면으로, 그림 1에 언급된 인터페이스 프로그램

이다. 실시간 예측은 약 75%의 정확도를 보이며, 빠른 수어에서도 높은 정확도를 보였다.

## V. 결론

본 논문에서는 신경망을 이용하여 수어를 실시간으로 인식하고 번역하는 알고리즘을 구현하였다. 관절의 정보를 용이하게 수집하기 위하여 마이크로소프트사의 깊이 센서 키넥트를 사용하였고, 더 나아가 관절의 정보를 바탕으로 팔 동작과 손의 위치를 추론하는 방법을 구현하였다. 또한, 손의 외곽선을 추출해서 손의 모양을 인식하는 알고리즘을 이용하여 효과적으로 신경망에 추출한 특징들을 학습시켰다.

## 사사문구

본 논문은 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 사회맞춤형 산학협력 선도대학(LINC+) 육성사업의 연구결과입니다.

이 성과는 2017년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2017R1C1B5016837).

본 논문은 산업통상자원부 글로벌전문기술개발사업(10063078)으로 지원된 연구결과입니다

## 참고문헌

- [1][https://www.mcst.go.kr/web/s\\_notice/press/pressView.jsp?pSeq=14849](https://www.mcst.go.kr/web/s_notice/press/pressView.jsp?pSeq=14849)
- [2][http://www.law.go.kr/법령/한국수화언어법/\(13978,20160203\)](http://www.law.go.kr/법령/한국수화언어법/(13978,20160203))
- [3]<http://www.romanakozak.com/sign-language-translator/>
- [4]<https://youtu.be/DpcI5h1EuqI>
- [5]<https://hal.inria.fr/hal-00725654/document>
- [6] 이호성(2018), CNN 및 Edge detection 기반 고속 손 동작 인식