

# Hand Tracking과 대화형 AI를 활용한 VR 실감형 수어 교육 콘텐츠 개발 연구

## Research on Development of VR Realistic Sign Language Education Content Using Hand Tracking and Conversational AI

천재성·문일영\*

Jae-Sung Chun · Il-Young Moon

Department of Computer Engineering Korea University of Technology and Education, Chungcheongnam-do, 31253, Korea

### [요 약]

본 연구는 청각장애인과 비장애인 모두를 위한 수어 교육의 접근성과 효율성을 개선하는 것을 목적으로 한다. 이를 위해 Hand Tracking 기술과 대화형 AI를 통합한 VR 실감형 수어 교육 콘텐츠를 개발하였다. 사용자는 이 콘텐츠를 통해 실시간으로 수어를 학습하며, 가상 환경에서의 직접적인 의사소통을 경험할 수 있다. 연구 결과, 이러한 통합 접근 방식이 수어 학습에 있어 몰입감을 크게 향상시키며, 학습자에게 더 깊은 이해를 제공함으로써 수어 학습의 장벽을 낮추는 데 기여한다는 것을 확인하였다. 이는 수어 교육의 새로운 패러다임을 제시하며, 기술이 교육의 접근성과 효과를 어떻게 변화시킬 수 있는지를 보여준다.

### [Abstract]

This study aims to improve the accessibility and efficiency of sign language education for both hearing impaired and non-deaf people. To this end, we developed VR realistic sign language education content that integrates hand tracking technology and conversational AI. Through this content, users can learn sign language in real time and experience direct communication in a virtual environment. As a result of the study, it was confirmed that this integrated approach significantly improves immersion in sign language learning and contributes to lowering the barriers to sign language learning by providing learners with a deeper understanding. This presents a new paradigm for sign language education and shows how technology can change the accessibility and effectiveness of education.

**Key word** : Sign language education, Hand Tracking technology, Chat GPT, Immersive VR content, OpenCV.

<http://dx.doi.org/10.12673/jant.2024.28.3.369>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 2 May 2024; Revised 24 June 2024

Accepted (Publication) 27 June 2024 (30 June 2024)

\*Corresponding Author: Il-Young Moon

Tel: +82-2-2209-3671

E-mail: [iymoon@koreatech.ac.kr](mailto:iymoon@koreatech.ac.kr)

## I. 서론

2017년 국립국어원에서 실시한 한국수어사용실태조사연구에 의하면 청각장애인들의 주된 의사소통 방법은 수어가 69.3%로 가장 높게 나타났다[1]. 수화(手話), 즉 수어(手語)는 손을 포함한 신체의 움직임을 이용해 대화하는 시각 언어로, 단어에 해당하는 동작이나, 해당 동작이 없을 경우 각 자음, 모음을 표현하는 지문자를 활용한다. 현대 사회에서 청각장애인들은 커뮤니케이션의 장벽에 직면해 있으며, 이는 주로 수어를 사용하는 사람의 부족과 수어 교육을 위한 실질적인 인프라의 제한 때문이다 [2]. 수어 통역 이용 경험이 있는 청각장애인들은 주변을 더 잘 이해할 수 있어 사회와의 소통의 폭이 넓어지기 때문에 수어 통역의 역할은 중요하다[3]. 현재 신조어와 같은 새로운 단어의 등장으로 일상생활에서 사용하는 수어의 약 80%가 지문자로 이루어져 있다고 한다. 수어 통역사들은 ‘코로나’, ‘AI’, ‘메타버스’처럼 신조어를 뜻하는 수어를 지정하는데 오랜 시간이 걸리고, 말의 속도도 따라가기엔 빠르다고 토로한다[4].

청각장애인이거나 비장애인들이 수어를 학습하는 대부분의 콘텐츠는 시청각 자료로 제공된다. 외국어를 학습하는 것과 유사하게, 시청각 자료만을 이용하는 것보다는 원어민과 직접 의사소통하며 배우는 것이 더 효율적일 수 있다. 수어 또한 실제로 의사소통하며 배우는 것이 효과적이지만, 오프라인에서 실시간으로 의사소통하며 수어를 배우는 것은 시간과 공간의 제약으로 인해 어려움이 있다.

이러한 배경 하에, 본 연구는 수어 교육의 접근성과 효율성을 개선하기 위한 새로운 접근 방식을 제시한다. 본 연구의 목적은 수어 데이터를 모델로 학습시키고 최적화된 알고리즘을 통해 수어를 문자로 정확하게 표현할 수 있는 기술을 개발하는 것이다. 더 나아가, 대화형 생성AI와의 접목을 통해, 사용자가 가상의 인물과 실시간으로 수어로 의사소통할 수 있는 교육 콘텐츠를 개발한다는 것이다. 이를 위해, Hand Tracking 기술을 활용하여 사용자의 수어 동작을 캡처하고, 대화형 AI 기술을 이용해 이를 인식하여 가상 인물이 애니메이션 효과와 함께 상호 작용하게 한다. 이 과정에서 사용자는 실시간 피드백을 받으며 수어를 학습하게 되어 학습의 몰입도와 효과를 극대화할 수 있다.

본 연구는 청각장애인뿐만 아니라 수어를 배우고자 하는 모든 사람들에게 수어 교육의 문턱을 낮추고, 보다 효율적이고 효과적인 학습 방법을 제공함으로써, 커뮤니케이션의 장벽을 해소하고 청각장애인 커뮤니티와 비장애인 커뮤니티 간의 상호 이해와 소통을 증진시키는 것을 목표로 한다.

## II. 이론적 배경

본 연구는 수어 교육의 개선을 목표로, 기존의 학습 방법을

기술적으로 향상시키려고 한다. 우선 수어 방식을 지문자와 지숫자[5]로 한정을 짓고 진행하기로 하였다. 이를 위해, Hand Tracking 기술과 대화형 AI, 그리고 가상 인물 애니메이션을 통한 교육 콘텐츠 개발에 초점을 맞춘다. 이 세 가지 기술의 공간을 이루는 이론적 배경과 알고리즘에 대해 아래에 자세히 설명한다.

### 2-1 Hand Tracking 기술

Hand Tracking 기술은 사람과 컴퓨터 간 상호작용에 초점을 둔 시스템으로, 손의 동작을 통해 직관적으로 컴퓨터를 작동한다[6]. 사용자의 손 움직임을 실시간으로 추적하고 인식하는 기술로, 컴퓨터 비전과 기계 학습 알고리즘을 기반으로 한다. 주로 OpenCV와 MediaPipe 같은 라이브러리를 활용하여 손의 랜드마크를 정확하게 검출한다. MediaPipe는 Google에서 주로 인체를 대상으로 하는 비전인식기능들을 AI모델 개발과 기계 학습까지 마친 상태로 제공하는 서비스이다. 다양한 프로그램 언어에서 사용하기 편하게 라이브러리 형태로 모듈화 되어 제공되며 사용방법 또한 풍부하게 제공되기 때문에 몇 가지 간단한 단계로 미디어파이프에서 제공하는 AI기능을 활용한 응용 프로그램개발이 가능하다[7]. 오픈소스 프레임워크로, 실시간으로 손의 랜드마크를 검출하는 데 탁월한 성능을 보인다. 이 데이터는 후에 수어 인식 모델의 학습에 활용된다.

### 2-2 대화형 AI

대화형 AI는 자연어 처리 (NLP; natural language processing)를 머신 러닝과 결합한다. 이러한 NLP 프로세스는 머신 러닝 프로세스와 함께 지속적인 피드백 루프로 흘러들어가 계속 AI 알고리즘을 향상한다. 대화형 AI에는 자연적인 방식으로 처리하고 이해하고 응답을 제공하도록 지원하는 주된 구성 요소가 있다[8]. 사용자와 자연스러운 대화를 가능하게 하는 인공 지능 시스템이다. 이 연구에서는 GPT-3 같은 고급 대화형 AI 모델을 활용하여 가상 인물과의 실시간 의사소통을 구현한다. 이 기술은 자연어 처리(NLP)의 최신 발전을 기반으로 하며, 사용자의 입력을 이해하고 적절한 반응을 생성한다. 대화형 AI의 목적은 문맥을 파악하고 사용자의 의도에 부합하는 응답을 제공하여 학습자가 실제와 유사한 상호작용을 경험하게 하는 것이다.

### 2-3 가상 인물 애니메이션

가상 인물 애니메이션은 사용자와 상호작용을 이루고 있다 [9]. 사용자와 상호작용하기 위해 디자인된 가상 캐릭터를 언리얼 엔진 같은 게임 엔진과 픽셀 스트리밍 기술을 사용해 구현된다. 언리얼 엔진은 고도의 리얼타임 3D 생성 툴을 제공하며, 픽셀 스트리밍은 웹 브라우저에서 고품질의 실시간 비디오 스트리밍을 가능하게 한다. 이를 통해 사용자는 별도의 고사양 장비 없이도 고품질의 가상 인물 애니메이션을 경험할 수 있다.



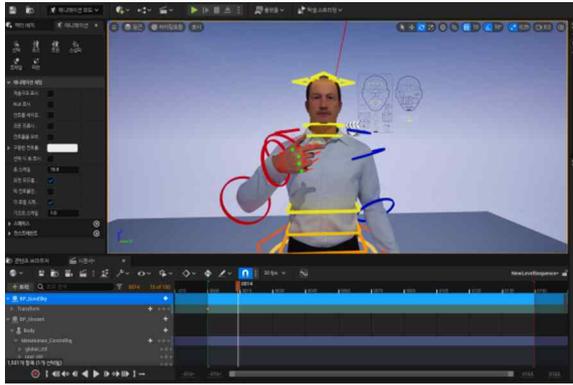


그림 4. 가상인물 애니메이션 픽셀 스트리밍 화면  
Fig. 4. Virtual character animation pixel streaming screen.

3) Flask 웹서버 및 대화형 AI 응답 시스템 구현

사용자가 웹 페이지를 통해 수어 동작을 인식하면, Flask 서버는 이를 처리하여 자모음을 조합하고 분리한 후, 해당 정보를 웹 페이지에 비디오 형태로 제공한다. 또한, chatGPT API를 활용하여 사용자의 수어 질문에 대한 응답을 생성하고, 이에 따른 가상 인물의 애니메이션을 동적으로 제공한다. 이 과정은 사용자와의 실시간 상호작용을 가능하게 하여 학습 경험을 풍부하게 한다.

4) 가상공간 및 가상인물 VR 콘텐츠 개발

그림 4와 같은 가상공간 및 가상인물 VR 콘텐츠의 개발은 학습자가 실시간으로 가상 인물과 수어로 의사소통할 수 있는 환경을 제공한다. 언리얼 엔진과 픽셀 스트리밍 플러그인을 활용하여 고품질의 가상 인물 애니메이션을 웹 페이지에 실시간으로 스트리밍 한다. WebRTC 기술을 통해 사용자는 별도의 소프트웨어 설치 없이 웹 브라우저를 통해 이러한 VR 콘텐츠에 접근할 수 있다.

IV. 시뮬레이션 및 연구결과

4-1 구성도

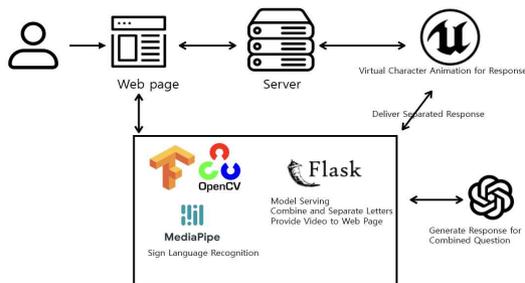


그림 5. 시스템 구성도  
Fig. 5. System configuration diagram.

그림 5의 구성도 처럼 사용자가 웹페이지를 통해 프로그램을 실행하고 웹캠 앞에서 수어 동작을 실행한다. 시스템에서 수어 동작을 인식하고 데이터를 생성한다. 그리고 난 뒤 생성된 데이터를 이용해 AI는 응답을 생성하고 이를 가상 인물 애니메이션으로 표현하여 웹페이지에 송출한다.

4-2 실행 화면

그림 6과 같이 왼쪽은 컴퓨터 웹캠 화면, 오른쪽은 가상 인물 애니메이션이다. 카메라 영역 내 자음, 모음 동작의 위치로 초성, 중성, 종성을 구분하며, 인식이 완료되면 웹캠 테두리에 초록색으로 표시된다.

그리고 난 뒤 VR 기기의 자체 Hand Tracking 기술을 통해 그림 7과 같이 화면에 사용자의 손 모양을 출력하고 가상 인물 애니메이션을 3D 공간에 출력한다.



Transmission of recognized character hand position values

```
data: L, (344, 278)
data: L, (346, 274)
data: L, (341, 271)
```

그림 6. 웹페이지 UI  
Fig. 6. Web page UI.



그림 7. 3D 공간 가상 인물 애니메이션  
Fig. 7. 3D space virtual character animation.

## V. 결 론

본 연구에서 개발된 수어 학습 플랫폼은 Hand Tracking 기술과 대화형 AI를 활용하여 수어 학습의 접근성과 효율성을 향상시키는 것을 목표로 했다. 특히, 이 플랫폼은 실시간으로 가상 인물과의 수어 의사소통을 가능하게 함으로써, 학습자의 몰입도와 효율성을 크게 증가시켰다. 연구 과정에서는 다양한 동작 인식 모델을 통합하여 수어의 다양한 표현을 정확하게 인식하고자 했다.

LSTM[12] 모델의 도입은 일상적인 수어의 동적인 동작을 정확히 인식하기 위한 전략의 일환으로, 기존의 랜덤 포레스트 분류 모델로 인식되는 정적인 지문자와 지숫자 인식을 넘어서, 연속된 동작을 포함하는 수어의 정확한 인식을 목표로 했다. 이는 LSTM 모델이 긴 시퀀스 데이터에 대해 높은 학습률을 보이기 때문에 선택되었다.

이번 연구에서 개발된 자모음 결합, 분리 알고리즘은 초성, 중성, 종성 위치 정보를 활용하여 쌍자음이나 겹치는 동작과 같은 경우에서 예외처리를 수행하는 알고리즘으로, 수어 인식의 정확도를 향상시키는 데 중요한 역할을 했다.

향후 연구 방향으로는, 사용자의 동작이 동적인지 정적인지를 판별하는 이진 분류 모델을 추가하여, 동작의 종류에 따라 적절한 모델(LSTM 모델 혹은 랜덤 포레스트 분류 모델)에 입력되도록 할 예정이며, 랜덤 포레스트 분류 모델에서 인식률이 떨어지는 비슷한 동작들(예: ㄷ, ㅌ, ㄹ 및 ㄱ, ㅋ, ㆁ, ㆁ)을 구분하기 위해 특정 키 포인트를 검출하여 이를 판별하는 추가적인 모델이 필요하다.

이러한 개선과 추가적인 모델 개발을 통해, 수어 학습 플랫폼이 수어 인식의 정밀도를 높이고, 수어 학습의 효과를 극대화할 수 있는 토대를 마련했다고 할 수 있다. 결론적으로, 본 연구는 수어 학습을 위한 새로운 기술적 접근법을 제시하며, 청각장애인과 수어를 배우고자 하는 사람들에게 새로운 학습 기회를 제공했다. 그러나 인식 모델의 정확도와 관련된 문제들을 해결하기 위한 지속적인 노력이 필요함을 인식하고, 이 분야의 연구는 수어 인식 기술의 정밀도를 높이고, 수어 학습의 효과를 극대화하기 위한 방법을 계속해서 모색해야 할 것이다.

## References

[1] National Institute of Korean Language, *Study on the Usage of Korean Sign Language* (2017-01-56), Seoul, Republic of Korea, 2017. Retrieved from [Seoul Hall] Storage (Request for

viewing, then available at the first floor loan desk), Call number: 401.9 -18-14. Available: <https://dl.nanet.go.kr/search/searchInnerDetail.do?controlNo=MONO1201834129#none>.

- [2] Deaf in Incheon Struggle Even with Sign Language Classes, Kyungin Daily [Internet]. Available: <http://m.kyeongin.com/view.php?key=20231208010000997>.
- [3] S. M. Koo, I. Jang, and Y. Son, "An open source hardware based sign language interpreter glove & situation awareness auxiliary IoT device for the hearing impaired," *Korean Institute of Information Scientists and Engineers Transactions on Computing Practices*, Vol. 24, No. 4, pp. 204-209, Apr. 2018. DOI: 10.5626/KTCP.2018.24.4.204.
- [4] H. J. Shin, There was no sign language for 'Corona'... information exclusion remains, Yonhap News Agency [Internet]. Available: <https://www.yna.co.kr/view/MYH20200307002300641>.
- [5] Monster, Expressing Korean alphabet and numbers in sign language-sign language communication, Naver Blog [Internet]. Available: <https://blog.naver.com/shin00512/220822640527>.
- [6] S. H. Jeon, "Technical term: Hand tracking," *Korea Broadcasting Engineers & Technicians Association Journal*, Retrieved from <http://journal.kobeta.com/%EA%B8%B0%EC%88%A0%EC%9A%A9%EC%96%B4-%ED%95%B8%EB%93%9C-%ED%8A%B8%EB%9E%98%ED%82%B9hand-tracking/>.
- [7] Nambo Gongbang, Developing AI python programs with MediaPipe, MakerNambo [Internet]. Available: <https://makernambo.com/154>.
- [8] IBM, What is conversational AI?, IBM [Internet]. Available: <https://www.ibm.com/kr-ko/topics/conversational-ai>.
- [9] Y. H. Son, H. J. Park, and M. H. Park, "Exploring key variables influencing group classification by reading literacy level using random forest: Focused on PISA 2018 data," *Asian Journal of Education*, Vol. 21, No. 1, pp. 191-215, Mar. 2020. DOI: <https://doi.org/10.15753/aje.2020.03.21.1.191>.
- [10] Yarisong, What is scikit-learn?, Dec. 19, 2023, Machine Learning/scikit-learn, Yarisong Tistory [Internet]. Available: <https://yarisong.tistory.com/72>
- [11] H. J. Yang, A study on tuning of noise covariance in localization filter for unmanned vehicle using LSTM, M.S. thesis, Korea University of Technology and Education, Republic of Korea, Feb. 2022. Retrieved from <https://lib.koreatech.ac.kr/pyxis-api/1/digital-files/6b6ca7af-4f4c-44ef-b618-68e3b65680db>



**천재성 (Jae-Sung Chun)**

2022년 8월 ~ 현재 : 한국기술교육대학교 컴퓨터공학과 석사과정  
※관심분야 : AI, 빅데이터, 가상현실, 웹



**문일영 (Il-Young Moon)**

2005년 2월 : 한국항공대학교 항공통신정보공학과 공학박사  
2005년 3월 ~ 현재 : 한국기술교육대학교 컴퓨터공학부 정교수  
※관심분야 : AI, 무선인터넷 응용, 무선 인터넷, 모바일 IP