

시각장애인들을 위한 얼굴 표정 인식 및 피드백 시스템

이신우, 신수용*

국립금오공과대학교

nakgongs@kumoh.ac.kr, *wdragon@kumoh.ac.kr

Facial expression recognition and feedback system for visually impaired people

Shin Woo Lee, Soo Young Shin*

Kumoh National Institute of Technology

요약

의사소통의 종류 중 비언어적 요소의 비중은 55%를 차지한다. 본 논문은 시각장애인이 대화를 하는 도중 흐름을 끊지 않고 상대방의 표정 정보를 인식하여 제공하는 것을 목적으로 한다. 대화를 하는 상대방의 얼굴을 인식하여 얼굴 그물망으로 표정을 분석하여 표정 정보를 진동 패드로 제공하는 시각장애인을 위한 얼굴 표정 인식 시스템을 제안한다.

I. 서론

장애인은 신체, 정신적으로 오랫동안 일상생활과 사회생활에서 상당한 제약을 받는 사람들을 말한다.[1] 이런 제약 중 의사소통 문제는 장애인이 겪게 되는 많은 어려움 중 하나이며, 현대사회에서 해결되어야 할 중요한 문제이다. 의사소통의 종류로는 언어적 의사소통과 비언어적 의사소통으로 나눌 수 있으며, 시각장애인은 대화에서 말의 내용과 표현력, 음성의 크기와 속도 등 언어적 의사소통에 대한 정보는 인식할 수 있으나 표정, 용모, 복장 등의 비언어적 요소에 대한 정보는 인식할 수 없다는 문제를 가지고 있다. 기존 이런 문제점을 해결하기 위하여 사운드 피드백 등을 사용하여 이런 불편함을 개선하고자 했다.[2] 하지만, 이런 방법의 경우 의사소통 도중 흐름이 끊기거나, 청취 해상도에 문제가 생길 수 있다.

본 시스템은 기존 방법의 문제점을 해결하면서 시각장애인에게도 비언어적 요소를 전달하여 의사소통의 폭을 넓히고 싶다는 사회적 목적을 가지고 있다.

II. 본론

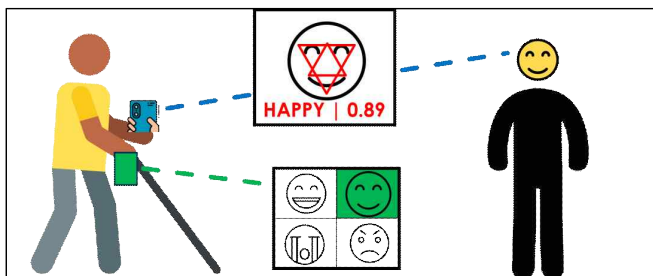


그림 1 제안 시스템 예시

본 논문에서 제안하는 시스템은 시각장애인들을 위한 얼굴 표정 인식 및 피드백 시스템이며, 그림 1은 본 논문의 시스템을 대략적으로 나타낸다.

먼저, 사용자의 안경 또는 스마트폰에 영상 수집장치(카메라)가 장착된다. 영상 수집 장치는 MediaPipe의 얼굴 인식 프로그램을 활용하여 표정

을 판별하게 되며, 이를 사용하여 사용자에게 진동 패드로 피드백을 전달하는 구조이다. 이 방법은 기존의 사운드 피드백 방식과는 달리 진동 패드를 사용하여 대화에 끊임이 없게 만드는 것이 특징이다.

MediaPipe는 구글에서 제공하는 AI Framework이며, 비디오 형식 데이터를 이용하여 다양한 AI 기능을 파이프라인 형태로 손쉽게 사용할 수 있다.



그림 2 MediaPipe Face Mesh Module 예시

본 시스템에서는 그림 2와 같은 얼굴 그물망 모듈(Face Mesh Module)을 사용하여 사용자에게 상대방의 얼굴 표정 정보를 제공한다. 얼굴 그물망 모듈은 실시간으로 468개의 3D 얼굴 랜드마크를 추정하며, 전용 센서 없이 일반 카메라만으로 사용할 수 있는 것이 장점이다. 본 시스템에서는 총 12개의 랜드마크(입술-4, 눈-4, 눈썹-4)를 사용하며, 행복, 웃음, 울기, 화남 4가지 표정의 랜드마크를 스냅샷으로 저장하여 입력 이미지로 사용한다.

사용 딥러닝 모델은 접근성이 용이한 Yolo v8 을 사용하게 되며, 커스텀 데이터셋 트레이닝 기능을 사용하여 입력 데이터가 어떤 표정인지 판단하게 된다.[3]

판단된 표정은 진동 패드의 입력으로 사용되어 사용자가 표정을 판단하는데 주요한 정보로 사용되어, 사용자의 의사소통을 보조하게 된다.

III. 결론

본 논문에서는 시각장애인이 대화를 하는 도중 흐름을 끊지 않고 상대방의 표정 정보를 인식할 수 있도록 표정 진동 페드로 표정 정보를 제공하는 시각장애인을 위한 얼굴 표정 인식 시스템을 제안한다. 해당 시스템을 통해 시각장애인들이 일상 생활에서 다른 사람들과 소통하고 상호작용하는 데 더 적극적으로 참여할 수 있어 자립적으로 활동할 수 있는 능력이 향상되도록 도울 수 있다. 추후 1대1 이상의 대화 상황에서도 각 대상마다의 얼굴 표정 정보를 시각장애인에게 제공하는 방식에 대한 연구가 진행될 예정이다.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2018R1A6A1A03024003) This work was supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (IITP-2024-2020-0-01612)

참 고 문 헌

- [1] 서울특별시남부 지역장애인보건의료센터” https://www.ssrhmc.org/information/inform_guide.html.
- [2] 임수연, and 반영환. "시각장애인의 커뮤니케이션 만족도 향상을 위한 방안연구-사운드 피드백을 중심으로." 커뮤니케이션 디자인학연구 81 (2022): 340-349.
- [3] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," Computer vision and Image understanding, vol. 106, no. 1, pp. 59 - 70, 2007.