

# 인공지능 기반 시각장애인 정면 유도 시스템

정다빈, 김진수, 장유진, 강영명

성결대학교

ddaall@sungkyul.ac.kr, nsa06035@naver.com, syjin@sungkyul.ac.kr, ykang@sungkyul.ac.kr

## An artificial intelligence-based gaze correction system for visually impaired people

Dabin Jeong, Jinsu Kim, Yujin Jang, Young-myung Kang

Sungkyul Univ.

### 요약

시각장애인은 자기 모습이 카메라에 어떻게 나오는지 알지 못해 타인의 도움 없이는 정확한 사진 또는 영상 촬영에 어려움을 겪는다. 본 논문은 인공지능 기술을 활용하여 시각장애인이 타인의 도움 없이 정면을 응시하여 정확히 사진을 촬영할 수 있도록 도와주는 정면 유도 시스템을 소개한다. 정면 유도 시스템은 ML Kit의 얼굴 인식 API를 활용하여 얼굴 윤곽과 얼굴 랜드마크의 좌표값을 구하여 오일러 각을 계산한 다음 사용자에게 얼굴의 방향을 바꾸도록 안내한다. 얼굴 방향 보정 안내는 TTS 기술을 활용하여 음성 메시지로 출력한다. 논문에서 제안한 정면 유도 시스템은 Android Studio 앱 형태로 구현하여 정상적 동작을 확인하였다. 정면 유도 시스템을 통해 시각장애인은 스스로 자신의 정면 사진을 촬영할 수 있으며 증명사진이나 여권 사진 등 다양한 방법으로 활용할 수 있다.

### I. 서론

유튜브 시장이 확대됨에 따라 전문가를 비롯하여 일반인들도 유튜브 영상 제작을 활발히 하고 있다. 더욱이 자신의 일상을 공유하는 장애인 유튜버 또한 많이 볼 수 있다. 장애인 유튜버는 자신의 장애를 당당하게 밝히고 장애와 일상을 공유하며 비장애인들에게 장애인에 대한 편견을 깨고 더 넓은 시각을 가지게 한다. 같은 장애가 있는 사람들에게 그들은 힘이 되어주고 위로가 되어준다. 한편 그림 1과 같이 높은 구독자 수를 가지고 있어 영향력이 큰 시각장애인 유튜버가 시각장애인인 척 연기한다는 의심을 받아 시각장애인이 어떻게 카메라를 똑바로 볼 수 있는지에 대한 영상을 올린 사례도 있었다[1].

해당 영상에서 촬영을 도와주는 PD는 먼저 카메라 뒤에서 소리를 내어 시각장애인 유튜버가 화면을 똑바로 볼 수 있게 해준다. 그 이후 카메라 화면을 보며 예를 들어 1시, 11시 등의 시계 방향 가이드를 추가로 제공해서 시각장애인이 카메라에 초점을 맞출 수 있도록 하였다. 이런 방법은 영상을 촬영하는 시각장애인을 도와주는 사람이 없다면 불가능한 방법이다.

본 논문에서는 시각장애인이 혼자서 촬영하거나 다른 사람들과 함께 사진을 찍을 때, 남의 도움 없이 카메라를 정면으로 볼 수 있게 하는 방법을 고안해보았다. 우리가 도입한 방법은 안면인식 인공지능 기술을 이용하여 정면 유도 시스템을 구현하고 시각장애인의 얼굴을 카메라에 비추었을 때, 얼굴이 정면을 볼 때까지 얼굴이 정면에서 틀어진 정도를 오일러 각을 이용하여 계산하고 이를 음성안내(TTS, Text-To-Speech)를 통해 가이드를 제공하는 것이다. 오일러 각을 이용하여 계산하려면 카메라가 얼굴을 인식해야 하는데, 이를 위해 ML Kit[2]의 얼굴 인식 API를 활용하였다. 우리는 안드로이드 스튜디오를 이용해 본 시스템의 프로토타입을 제작하였고 직접 시연을 통해 우리 시스템이 효과적으로 동작하는 것을 확인하였다.



그림 1. 시각장애인 유튜버의 영상

### II. 정면 유도시스템

#### 2.1 활용 기술

정면 유도 시스템은 시선 보정과 보정 내용의 음성 안내가 필요하며 이를 위해 각각 구글 ML Kit 얼굴 인식 API와 TTS를 활용하였다.

#### 2.2 ML Kit

ML Kit은 Android 및 iOS 앱에 Google의 기기 내 머신러닝 기술을 제공하는 모바일 SDK이다. ML Kit을 통해 Vision과 Natural Language API를 사용할 수 있다. 본 논문에서는 ML Kit Vision 얼굴 인식 API를 활용한다. 눈, 코, 입 등 얼굴의 특징적인 것들을 얼굴의 랜드마크라고 표현하는데, ML Kit Vision API를 활용하면 얼굴의 랜드마크를 인식하고 이 좌표값을 가져올 수 있다. 즉, 그림 2와 같이 Vision API를 통해 얼굴을 인식하고 계산을 통해 얼굴 랜드마크 특징과 윤곽을 점 형태로 나타낼 수 있다[2]. 다음으로 얼굴의 윤곽 특징 계산을 통해 정면 응시 여부를 판별해야 하는데 우리는 Euler Angles를 적용하였다.

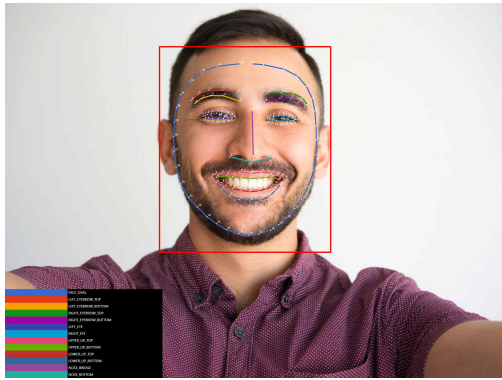


그림 2. 얼굴 윤곽 인식

### 2.3 Euler Angles

시각장애인의 얼굴이 정면인지 아닌지를 판단하기 위해 오일러 각을 활용한다. 오일러 각은 강체(사물)가 놓인 방향을 3차원 공간에 표시하기 위한 세 개의 각도이다. 그림 3에서  $x, y, z$ 는 고정 좌표계이고  $x', y', z'$ 은 회전 좌표계이다. 그림 4의 계산식은 각각  $x$ 축,  $y$ 축,  $z$ 축을 회전축으로 하여 회전할 때의 기본적인 회전 행렬이다. 고정 좌표계의 벡터에 회전 행렬을 곱하면 회전 좌표계의 벡터이다.

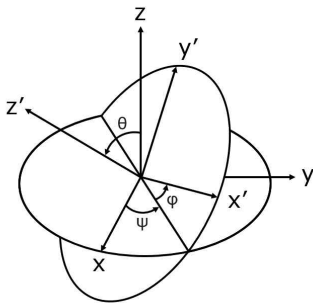


그림 3. 오일러 각

$$R_x(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

$$R_z(\phi) = \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

그림 4. Rotation matrices

오일러 각의 X값은 얼굴을 기준으로 위, 아래를 표현한다. X값이 양수이면 얼굴의 방향이 위를 향한다는 의미로, 고개를 치켜든 것이다. 이 경우에는 시각장애인에게 고개를 내릴 것을 안내한다. 오일러 각의 Y값은 얼굴을 기준으로 왼쪽과 오른쪽을 뜻한다. Y값이 양수이면 카메라 오른쪽을 바라보고 있는 것이고 음수이면 카메라 왼쪽을 바라보고 있는 것이다. 오일러 각의 Z값은 카메라를 기준으로 시계 반대 방향으로 회전하는 것을 표현한다. 시계 반대 방향으로 회전한다는 것은 귀가 어깨에 닿도록 얼굴을 돌린 것을 의미한다. 즉, 오일러 각의 X, Y, Z값이 모두 0에 가까울수록 시각장애인은 카메라의 정면을 응시하고 있는 것이다.

### 2.4 TTS(Text-to-Speech)

TTS 기술은 Text to Speech로 문자를 소리로 바꾸는 기술이다. 본 논문에서는 화면을 보지 못하는 시각장애인을 위해 오일러 각 좌표값에 따라 얼굴이 정면에 향하도록 얼굴을 움직일 것을 안내하는 메시지를 음성으로 출력한다. 한편 청력의 문제로 음성 메시지를 들을 수 없는 사람, 주변 소음으로 음성 메시지를 들을 수 없는 상황, 음성 메시지가 들리지 않게 하고 싶은데 이어폰이 없는 상황 등 다양한 상황에서도 메시지를 사용자에게 전달하고자 한다. 점자 스마트워치[4]를 본 시스템과 연동하여 ‘얼굴을 오른쪽으로 돌리세요.’와 같은 메시지를 간단하게 ‘오른쪽’, ‘왼쪽’ 등의 방향을 점자 스마트워치에 나타내어 음성 메시지를 대신할 수 있다.

### 2.5 시스템 구성

우리는 Android Studio를 이용해서 시각장애인을 위한 스마트폰 앱을 구현하였다. 안면인식은 스마트폰의 카메라 모듈을 활용하고, 앞서 설명한 것과 같이 ML Kit의 얼굴 인식 API를 활용하여 사용자의 얼굴을 인식하고 오일러 각을 계산한다. 계산한 오일러 각을 토대로 사용자의 얼굴이 어느 방향으로 움직여야 정면에 가까울지 계산한 후, TTS 기술을 통해 사용자에게 음성으로 안내한다. 그림 5에 시스템 구성도가 표현되어 있다.

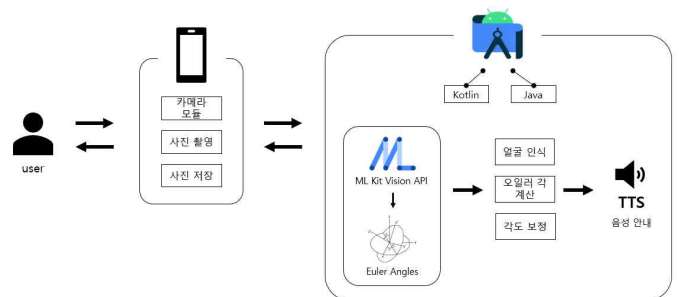


그림 5. 시스템 구성도

## III. 결론

본 논문에서는 ML Kit의 얼굴 인식 API와 TTS 기술을 활용하여 Android Studio로 시각장애인이 스스로 정면을 바라볼 수 있도록 도와주는 시선 보정 앱을 개발하고 시스템 구성을 소개하였다. 이 시스템은 비장애인 또한 시력이 좋지 않거나 전면카메라가 아닌 후면카메라 촬영으로 화면을 보기 어려운 경우에 활용할 수 있으며, 증명사진, 여권 사진 촬영 등 다양한 방향으로 확장 가능하다. 향후 연구에서는 점자 위치를 연동하여 시각과 청각 장애가 있는 경우에도 점자를 활용하여 원하는 동작을 안내할 수 있도록 할 계획이다.

## 참 고 문 헌

- [1] "시각장애인이 카메라를 어떻게 그렇게 똑바로 쳐다봐? 저는 사실...", 유튜브 비디오, 5:42. "원샷한술", 2020.11.26. ([https://youtu.be/sLjh\\_FKjbo](https://youtu.be/sLjh_FKjbo))
- [2] "ML Kit Face Detection." Google ML Kit. 2022, (<https://developers.google.com/ml-kit/vision/face-detection>)
- [3] Gregory G. Slabaugh. "Computing Euler angles from a rotation matrix." Queen Mary University of London, 1999.
- [4] "닷 워치", Dot, 2021, (<https://www.dotincorp.com/kr>)