

PHODO: 머신러닝 기반 얼굴 인식 및 추적 시스템

윤예찬, 송근주, 박소희, 김혜경, 김수민*
한국산업기술대학교

E-mail : yyc837021@gmail.com, *suminkim@kpu.ac.kr

PHODO: a Machine Learning-Based Face Recognition and Tracking System

Ye Chan Yun, Keun Ju Song, So Hee Park, Hye Kyoung Kim, and Su Min Kim*
Dept. Electronics Engineering, Korea Polytechnic Univ.

요 약

본 논문에서는 최근 급증하는 온라인 개인방송을 위한 다양한 미디어 콘텐츠 제작을 손쉽게 지원할 수 있는 얼굴 인식 및 추적 시스템을 제안하고 구현한다. 제안 시스템은 ARM 프로세서 기반 임베디드 시스템으로서 기존의 삼각대에 손쉽게 설치할 수 있고 추가로 트래킹 레일도 조합하여 활용할 수 있으며 포터블한 사이즈로 사용자에게 휴대성을 제공한다. 학습 기반 얼굴 추적 기능은 스마트폰용 안드로이드 카메라 앱으로 구현되었고, 머신러닝 기술을 기반으로 다자 얼굴 감지와 개인 얼굴 학습 및 감지 기능을 지원한다. 추가적으로 3 분할 최적구도 결정 및 자동촬영, 음성인식 및 출력, 자동 확대/축소 및 레일 트래킹 이동 기능을 통해 다양한 자동 개인촬영 서비스를 제공한다.

I. 서론

최근 다중채널 네트워크(MCN) 기반 인플루언서 마케팅 시장 조사 결과, 2020년에는 11조원 규모로 성장할 것으로 예측된다 [1]. 또한, 유튜브 코리아에 따르면 유튜브에는 1분마다 400시간이 넘는 분량의 새 동영상들이 업로드되고 있다 [2]. 이 추세는 앞으로도 더욱 가속화될 것으로 예상되어 더욱 다양한 멀티미디어 콘텐츠들이 활발하게 제작될 것으로 전망된다. 이러한 영상 콘텐츠들은 언제 어디서나 손쉽게 제작되고 방송되는 방향으로 발전되고 있어 모바일 스트리밍 플랫폼에서 개개인의 미디어 제작의 중요성이 더욱 커질 것이다. 따라서 본 논문에서는 개개인의 미디어 콘텐츠 제작을 손쉽게 지원할 수 있는 머신러닝 기술을 이용한 학습기반의 개인맞춤형 얼굴인식 및 트래킹을 통한 자동 영상촬영 시스템을 제안한다.

II. 본론

그림 1에서 보는 바와 같이 제안 시스템은 크게 사용자의 얼굴을 인식하여 추적하도록 하는 ARM 기반 임베디드 시스템과 머신러닝 기반 얼굴인식 및 자동촬영 기능이 구현된 스마트폰용 카메라 촬영 안드로이드 애플리케이션으로 구성된다. 안드로이드 앱은 카메라 화면을 통해 사용자의 얼굴을 실시간으로 감지 및 추적하고 감지된 얼굴의 좌표값을 블루투스로 임베디드 시스템에 무선으로 전송한다. 임베디드 시스템은 수신한 얼굴 좌표값을 토대로 다수의 모터를 구동시켜 3차원으로 얼굴 추적이 가능하도록 한다.

그림 2는 제안 시스템의 하드웨어 설계모습을 보여준다. 앞서 소개한 바와 같이 삼각대에 설치하여 사용할 수 있고, 상하, 좌우로 회전하여 3차원 자유도를 갖는다. 상단에는 휴대폰 거치가 가능하도록 하여 스마트폰이 고정되도록 하고, 개발한 안드로이드 앱과 연동하여 얼굴 인식 및 트래킹을 통한 개인맞춤형 자동촬영이 가능하도록 설계하였다.

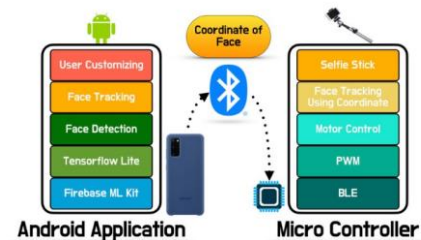


그림 1. 시스템 구성도

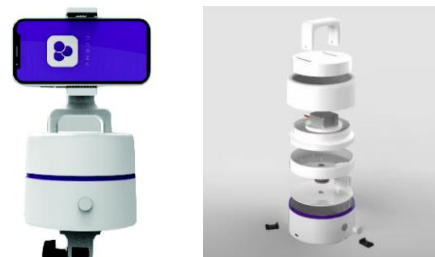


그림 2. 하드웨어 구성

그림 3은 제안 시스템의 실제 구현 결과를 보여준다. 그림에서 보는바와 같이 실제 구현 결과에서는 앞서 소개한 임베디드 시스템과 연동하여 추가적인 트래킹 자유도를 제공하는 카메라 레일이 추가되었다.



그림 3. 구현 결과 모습

추가된 카메라 레일은 끝단에 설치된 초음파 센서를 이용하여 x 좌표 위치를 제어하면서 보다 넓은 트래킹 및 촬영 범위를 제공한다.

그림 4 는 안드로이드 앱의 실행 화면을 보여준다. 앱은 자바 프로그래밍 언어로 Android Studio IDE 를 이용하여 개발하였다. 그림은 초기 실행화면, 얼굴 감지 영역, 학습 기반 사용자 얼굴인식의 실시간 카메라 화면을 보여준다. 얼굴인식은 Google ML Kit API 를 이용하여 구현하였다 [3]. 사용자 얼굴인식은 개인의 얼굴 특징을 파악해 다른 사람의 얼굴과 구분하여 등록된 사용자의 얼굴만 트래킹할 수 있도록 한다. 이 기능은 사전에 촬영하여 등록한 사용자의 사진을 이용하여 학습한 모델을 기반으로 한다. 사용자 얼굴 학습은 머신러닝 플랫폼 Tensorflow 를 이용하여 구현하였으며, 모바일 디바이스에 맞게 경량화 된 Lite 모델을 사용하였다 [4]. 구체적으로 학습 모델은 Google 에서 개발한 FaceNet 모델을 사용하였다. FaceNet 모델은 유클리디안 공간 기법을 기반으로 하는 머신러닝 모델이다 [5]. 이를 통해 초기 설정을 간단히 하여 처리 속도를 높였으며, LFW(Labeled Faces in the Wild) 데이터에서 99.63%의 높은 학습 성능을 보였다 [6].

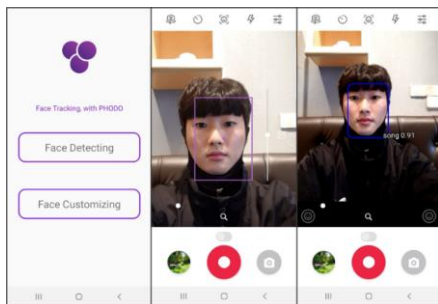


그림 4. 애플리케이션 실행 화면

그림 5 는 Android Studio IDE 에서 출력된 앱의 logcat 화면이다. logcat 에서 보여지는 값은 얼굴 좌표값이며 블루투스 무선통신을 통해 임베디드 시스템으로 전송된다. 사용자 개인맞춤형 인식 모드에서는 학습된 얼굴에 대한 유사도를 기준으로 판단하여 제어신호를 전송한다. 즉, 실시간으로 감지된 얼굴이 학습된 데이터와 90% 이상의 유사도 값을 가질 경우에만 좌표값을 전송한다.

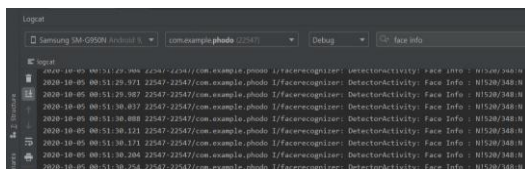


그림 5. Logcat 에서의 블루투스 좌표값 출력

리눅스 운영체제에서 블루투스 통신은 오픈소스 라이브러리인 pyserial 라이브러리를 선정하였다 [7]. 블루투스 스택은 bluez 패키지를 사용했으며 직렬통신 프로파일을 위한 RFCOMM 프로토콜을 통해 데이터를 송수신한다 [8]. 그림 6 에서 보는 바와 같이 블루투스 통신 환경은 사전에 관련 리눅스 명령어들을 묶어 셸 스크립트로 만들어 쉽고 빠르게 블루투스 통신 환경 구축을 마칠 수 있도록 하였다.

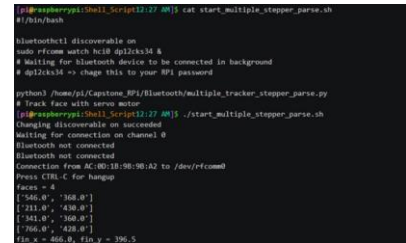


그림 6. 블루투스 데이터가 표시된 bash shell 화면

그림 7 은 최종적으로 구현된 사용자 트래킹 기능을 시연하는 모습을 보여준다. 1 인/다중 사용자 트래킹이 모두 가능하며 다중 트래킹은 도출된 다중 사용자의 얼굴좌표의 평균점을 최적의 구도에 위치하도록 트래킹하며 실시간 촬영하도록 구현되었다.

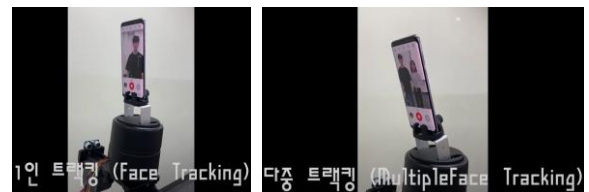


그림 7. 1 인/다중 트래킹 기능 시연 모습

III. 결론

본 논문에서는 머신러닝 기술을 적용한 얼굴 감지 및 추적을 통한 자동 촬영 시스템을 제안하였다. 제안 시스템은 최적구도 촬영, 음성인식 핸드프리 촬영, 자동 줌 기능 및 레일 트래킹 기능을 통해 사용자들에게 다양한 개인맞춤형 촬영기능을 제공한다. 이러한 학습 기반 사용자 인식 및 추적 시스템을 통해 1 인 방송 미디어 시대에 맞춰 개인 촬영의 편의성을 제공할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 논문은 2020 년 정부(과학기술정보통신부)의 재원으로 한국연구재단-현장맞춤형 이공계 인재양성 지원사업의 지원을 받아 수행된 연구임(NO. 2017H1D8A1032103).

참 고 문 헌

- [1] 돈줄 물리는 MCN 산업, 대형 VC 들의 선택은?, 국민일보, Apr. 2020. Online: <http://news.kmib.co.kr/article/view.asp?arcid=0014527039&code=61141411&cp=nv>
- [2] 1 분마다 400 시간 업로드되는 유튜브...구글은 어떻게 유튜브를 황금거위로 키웠나, 중앙일보, Jan. 2019. Online: <https://news.joins.com/article/23253498>
- [3] Google ML Kit, Online: <https://developers.google.com/ml-kit/vision/face-detection>
- [4] Tensorflow Lite, Online: www.tensorflow.org/lite?hl=ko
- [5] F. Schroff, D. Kalenichenko, J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering, Version 3," June 2015. Online: <https://arxiv.org/abs/1503.03832>
- [7] Document Page of the pyserial Library. Online: <https://pyserial.readthedocs.io/en/latest/pyserial.html>
- [8] Official Linux Bluetooth protocol stack - BlueZ. Online: <http://www.bluez.org/>