한국인의 얼굴 표정 인식 기술 개발

팀명 EmoTex

착수 보고서

김영빈(201624441)

우사마칸(201724628)

라이스트배일라(201724626)

2020년 6월 9일

<목 차>

목차 1

1. 개요 및 목표 2

가. 개요 2

나. 필요성 2

다. 과제 목표 2

2. 대상 문제 및 요구사항 분석서 3

가. 대상 문제 3

나. 유사 시스템 분석 4

다. 요구사항 분석서 8

3. 현실적 제약 사항 및 대책 9

4. 설계문서 10

5. 역할분담 12

6. 추진일정 12

참고문헌 13

1. 개요 및 목표
   1. 개요

Youtube 영상을 시청하는 사용자의 감정을 인식하고 감정의 변화를 기록하는 얼굴 표정 인식 시스템을 개발한다.

본 졸업과제 팀의 목표는 Emotion Recognition분야의 기술을 적용한다. 영상을 보는 사용자의 감정의 변화를 데이터베이스에 저장하고 영상시청을 마친 후 그 영상에서 수집한 감정변화 데이터를 분석하여 사용자에게 제공한다. 또한 수집된 개별 유저의 감정변화 데이터를 이용하여 Youtube Contents Provider가 자신의 컨텐츠를 분석할 수 있는 Metric을 제공한다.

* 1. 필요성

쌍방향 인터넷 환경의 발전과 PC와 스마트폰의 확산으로 ‘1인 미디어’ 가 등장하고 발전하기 시작하였다. 개인의 단순한 취미가 돈이 된다는 인식이 차츰 확산되어 1인 크리에이터의 수가 급증하고 있다. 유튜브의 대다수 이용자는 즐거움을 위해서 영상을 시청하는데 실제로 영상을 보며 어떤 감정을 느끼고 어떤 종류의 컨텐츠를 선호하는지를 알지 못한채 유튜브의 알고리즘이 추천해주는 영상만을 시청한다. 또한 1인 크리에이터들은 영상의 흥행 실패를 조회수, 좋아요 수, 댓글 수 정도의 한정된 metric을 통해서만 분석할 수 있다.

본 졸업과제팀이 개발하려는 시스템은 영상을 시청하는 유저들의 감정을 분석하여 특정 영상에 대한 선호도를 설명하고 또한 개별유저의 감정데이터들을 모아서 특정 영상에 대한 선호도를 다양한 기준으로 분석할 수 있도록 한다

* 1. 과제 목표
     1. 감정 인식 딥러닝 모델 개발

휴대폰 혹은 웹카메라를 이용하여 영상에서 사용자의 얼굴을 인식해서 표정으로 감정을 분류하는 딥러닝 모델을 개발한다. 감정의 카테고리를 분류하는 것과 사용자의 순간 감정이 얼마나 긍정적인지 부정적인지를 –1(부정적)~1(긍정적) 사이값으로 예측하고 감정의 흥분정도를 –1(침착함)～1(흥분함)사이값으로 예측하는 모델을 개발한다.

* + 1. 사용자가 유튜브 영상을 시청하며 얼굴표정을 인식하는 앱 개발

사용자가 선택한 유튜브 영상을 시청할 때 얼굴표정을 인식하는 앱을 개발한다. 앱에서 사용자가 영상시청을 끝냈을 때 사용자 감정 정보는 서버에 전송되도록 한다. 그리고 분석된 감정 정보들을 그래프화해서 사용자가 영상을 보며 느꼈던 감정의 Report를 확인할 수 있도록 앱을 개발한다.

* + 1. 영상 분석 및 감정 정보 처리 서버 개발

영상시청이 끝난후 개별 안드로이드 폰에서 전송되는 감정 정보를 처리하고 데이터베이스 에 저장하는 백엔드 서버를 개발한다. 백엔드서버 데이터베이스에 저장된 개인의 감정 리 포트를 모아서 한 영상에 대한 감정 표현을 다양한 집단기준( 나이 , 성별, 문화 등）으로 　　filtering하는 기능을 개발한다

1. 대상 문제 및 요구사항 분석서
   1. 대상 문제
      1. 감정 인식 최신 기술의 정확도

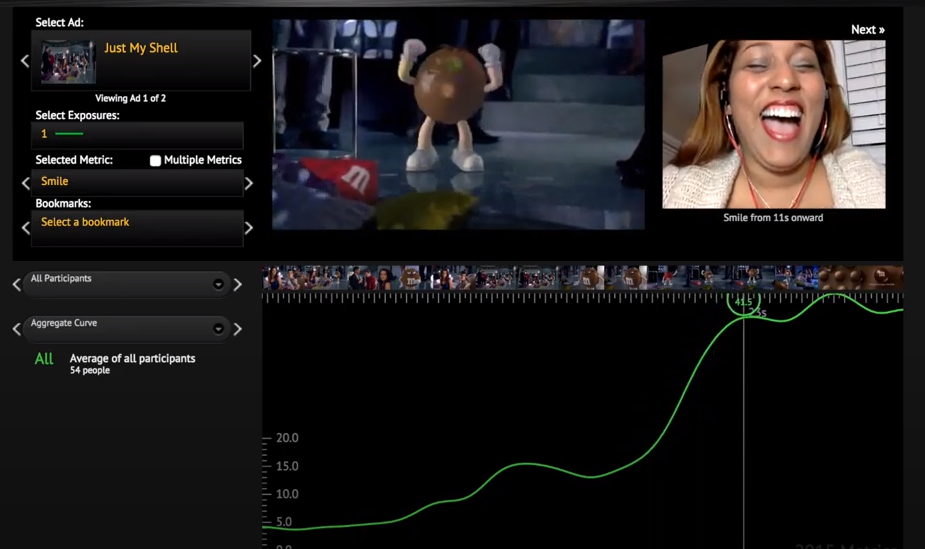
최신 모델의 감정인식 정확도와 관련된 표 논문 출처로 기입

* + 1. 척추 교정

장시간 컴퓨터 앞에 앉아서 작업에 집중하다 보면 바른 자세를 유지하기가 쉽지 않다. 바른 자세로 앉아있는 습관이 들어있지 않으면 무의식적으로 다리를 꼬아서 앉거나 한쪽으로 기대어 앉게 된다. 이러한 자세는 척추가 옆으로 휘거나 회전하는 척추변형상태인 측만증으로 발전할 수 있다.

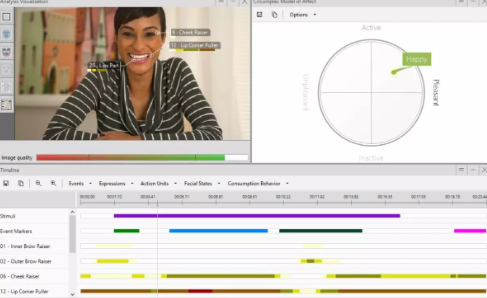
* 1. 유사 시스템 분석
     1. Affectiva

Affectiva는 Media를 시청하는 유저의 감정변화를 기록하여 비즈니스에 활용할 수 있도록 돕는 Solution입니다.Affectiva의 Solution은 고객들의 User Experience를 향상시키고 마케팅 및 영업에 활용할 수 있는 지표를 생성하도록 합니다. 다양한 Filter 기능(비슷한 연령대가 느끼는 평균 감정, 성별로 구분한 평균 감정)으로 감정을 다양한 관점에서 분석할 수 있습니다.

<그림 1> Affectiva Demo 동작 시스템

* + 1. FaceReader

Face Reader는 사진 혹은 영상을 동적으로 인식하는 툴입니다. 얼굴의 표현을 Action Unit(얼굴 근육 단위)으로 세밀하게 분석합니다. 또한 사용자가 느끼는 감정의 카테고리와 Action Unit의 움직임을 분석해줍니다.

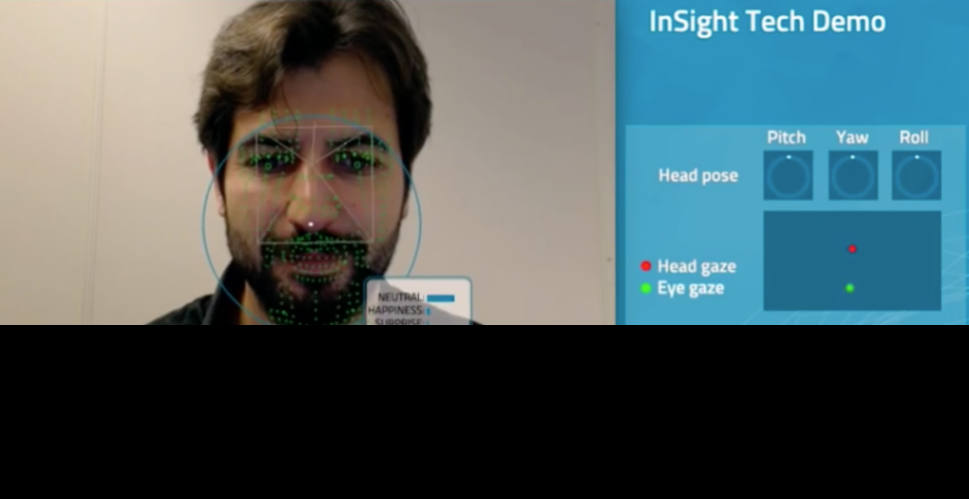


<그림 2> Face Reader 동작화면

* + 1. Sightcorp

Sightcorp는 감정 분석과 얼굴 인식 Library를 웹 Rest API 형태로 제공하는 사이트입니다.

University of Amsterdam(UaV)의 Vision Lab에서 발간한 Deep Learning Research를 기반으로 실시간 감정인식 기술의 상업적인 solution을 제공합니다. 얼굴의 위치 , 기울임 정도 그리도 시선까지도 분석해주는 것이 특징입니다.



<그림 4> Sightcorp 동작 확인

* + 1. 종합 및 결론

<표 1> 유사 시스템 분석 종합 및 결론

|  |  |  |  |
| --- | --- | --- | --- |
| **Application** | **기능** | **사용기기** | **비고** |
| Affectiva | 얼굴위치 detection | Microsoft Kinect | 카메라를 사용자의 옆에 둠 |
| Face Reader | 목 뒤에 장착하여 전방두부자세 여부를 판단후 감지시 진동을 울림 | Alex, NAMU.Inc 제작 | 25g 가량의 wearable device를 장착하고 다녀야 됨 |
| SightCorp | 사용자의 정면 자세를 얼굴과 어깨의 특징으로 판단. 기준 자세와 비교하여 경고 알림 | 노트북 혹은 모니터 상단에 내장되어 있는 웹 카메라 | 자리에 앉을 때 마다 기준 자세의 초기화 과정을 거침 |
|  |  |  |  |
|  |  |  | - |

기존의 유사논문은 카메라를 옆에 두고 인식하는 경우가 대부분인데 본 시스템에서는 카메라를 정면에 두고 자세를 인식하고자 한다. 기존의 논문들에서 소개된 시스템에서는 Depth 카메라를 위주로 측정하였는데 본 시스템에서는 RGB 카메라를 활용한 알고리즘을 추가해보자 한다. 또한 “단안 카메라를 이용한 자세교정유도 시스템”에서는 자리에 앉을 때 마다 기준 자세의 초기화 과정을 거치는데, 본 시스템에서는 기계학습을 통해 미리 기준자세를 정해놓고 사용자는 그냥 프로그램을 작동시키면 되는 방향으로 하려고 한다.

* 1. 요구사항 분석서
     1. 실시간 영상에서 감정인식

휴대폰 내장 RGB camera 혹은 web camera를 이용하여 사용자의 얼굴 영상을 실시간으로 감정을 인식하여 데이터를 저장한다.

* + 1. 분석한 결과를 그래프화

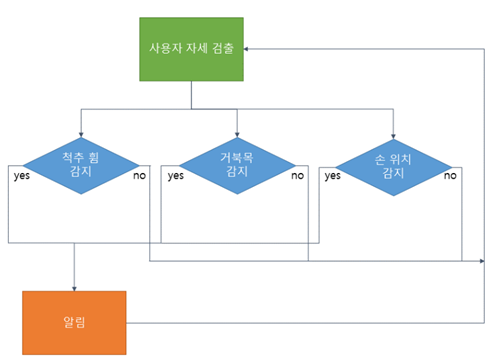
시스템은 입력 받은 얼굴 정보를 처리하여 사용자의 자세, 습관을 인식한 뒤 이를 반영하여 사용자에게 알림을 줄 수 있어야 한다.

1. 현실적 제약 사항 및 대책

<표 2> 현실적 제약 사항 및 대책

|  |  |  |
| --- | --- | --- |
| **제약사항** | **설명** | **대책** |
| 적외선으로 인한 노이즈 | 태양광에 포함되어 있는 적외선으로 인해서 적외선을 활용하여 깊이를 측정하는 depth카메라의 인식능력에 어려움이 생길 수 있다. | 실행환경을 태양광이 적게 드는 실내로 제한한다. |
| 범용성 문제 | 일반 카메라에 비해 비교적 고가인 depth카메라를 사용해야 한다. | RGB카메라를 활용한 알고리즘을 추가하여 정확도를 개선한다. |
| 조명 문제 | RGB카메라를 사용하는 알고리즘의 경우 조명의 정도에 따라 인식능력에 차이가 생길 수 있다. | 실행 환경을 밝은 조명으로 제한한다. |
| 다수의 사용자 문제 | 카메라 안에 2명 이상의 사람이 있는 경우 누구를 인식해야 할지 판단하기 어렵다. | 판단 기준을 카메라 기준 정 가운데 앉아 있는 사람으로 한다. |
| 카메라의 위치 | 카메라와 사용자 사이의 거리, 카메라와 사람 사이의 각도에 따라 인식능력에 차이가 생길 수 있다. | 카메라와 사용자 사이의 거리는 50 ~ 70cm로 제한. 카메라의 높이를 사용자 눈높이 기준 30cm이내로 제한하여 사용자 눈높이와의 각도를 20~30°로 유지한다. |
| 전방두부자세 판단 기준의 모호 | 해당 사용자 기준으로 전방 두부자세가 아닌데도 전방두부자세라고 판단하는 경우가 생길 수 있다. | 전방두부자세 판단 기준, 각도를 사용자가 설정할 수 있게 한다. |
| 스마트폰과 사용자와의 밀착 여부 | 5인치 이상의 스마트폰이 많아짐에 따라 스마트폰을 주머니에 넣고 앉는 경우가 드물게 되었다. | 습관 개선을 위해 스마트폰을 주머니에 넣어둬야 한다고 사용자 매뉴얼에 권고를 한다. |
| Emotion detection accuracy | In some cases, it would be hard to identify the user’s emotions accurately |  |
| Lack of module training samples | Our module training sample includes about 500,000 pictures only. Which is a relatively small data set for an initial training for the new made module |  |
| Determining the length | Determining if 80% or more of the video has been watched by the user. | Check if YouTube data API provide how much of the video has been watched or record the time the user has spent in the video page to determine that |
| Importing the content creators video list to the app | In order to show the content creators an analysis report about how the users feel about their videos, we need to import the list of the videos made by content creators. | Asking users to sign up/in by using their google account then import the content creator’s video library to the app using YouTube API |
| Avoiding Implementing a YouTube app from scratch | YouTube is a big application that would be hard to implement each feature of it from scratch. | Using WebView object to show the users an ordinary YouTube page |
| Stay “signed in” | The app shows YouTube webpage through the method WebView. Thus, when users sign in into their YouTube account and close the app, their log in information might be lost. | WebView might have a way to store cookies of the website that could be called every time the user opens the app |
| Linking users’ results into content creators’ reports. |  | Logging the user watched videos’ URLs in a database. As well as the content creators’ list of videos. Then comparing the content creator video URLs with the URLs in the users watched list database |

1. 설계문서



<그림 4> 사용자 감정 인식 흐름도

① 카메라를 통해 얼굴 위치를 검출한다.

② 인식된 얼굴을 전처리한다.

③ 전처리된 얼굴 이미지를 딥러닝 모델을 이용해서 예측한다.

1. 역할분담

|  |  |
| --- | --- |
| **이 름** | **역 할** |
| 김영빈 | 백엔드 서버 구축 및 개발  Youtube API 활용 |
| 우사마 칸 | 감정인식 딥러닝 모델 연구  감정인식 딥러닝 모델 서버 개발 |
| 라이스 트배일라 | Application UI 구현  안드로이드 앱 제작 |
| 공통 | 시스템 설계 및 보고서 작성  시스템 통합 및 디버깅 |

<표 3> 역할분담

1. 추진일정

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **월** | **6** | **7** | | | | **8** | | | | | **9** | | | | **10** | | | |
| **주차** | **5** | **1** | **2** | **3** | **4** | **1** | **2** | **3** | **4** | **5** | **1** | **2** | **3** | **4** | **1** | **2** | **3** | **4** |
| 관련지식학습 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 개발환경 구축 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 설계 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 구현 및 기계학습 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 통합 및 테스트 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 디버깅 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 결과 보고서 작성 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

<표 4> 추진일정

참고문헌