3차원 얼굴 랜드마크 기반 시선 추적 시스템 구현

남성현, 송하영, 손은진, 김영원*

한국전자기술연구원

(sadguest, shy1230, ejson, *kimforever920)@keti.re.kr

Implementation of a 3D face landmark-based gaze tracking system

Nam Seong Hyeon, Song Ha young, Son Eun Jin, Kim Young Won*

Korea Electronics Technology Institue

요 약

본 논문에서는 단일 RGB 카메라를 통해 획득한 2차원 영상으로부터 3DDFA를 통해 얼굴 영역과 얼굴 랜드마크를 검출하여, 시선 특징을 추출하고, 시선을 추적하는 시스템을 구현하였다. 시선 추적 모델의 학습을 위하여 9개의 시선 특징을 실시간으로 취득하여 약 4만개의 데이터를 통해 딥러닝 기반 회귀모델을 통해서 모델 학습을 검증한 결과 x축 오차율 8%, y축 오차율 10%를 달성하였다.

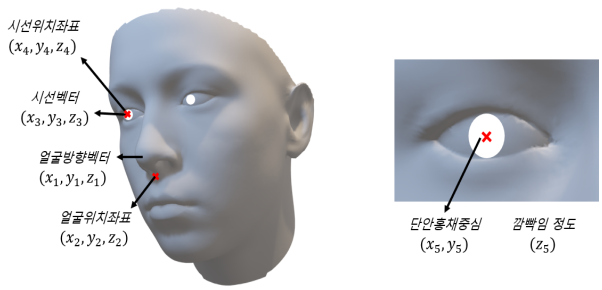
Ⅰ. 서 론

시선 추적 기술은 인간의 눈동자에 대한 데이터를 취득하여, 어느 부분을 응시하고 있는지 판단하는 기술을 말한다. 이러한 시선 추적 기술을 활용하여 심리학적 연구도 진행되고 있을만큼 인간의 시선은 인간 행동에 대한 많은 정보들을 가지고 있다[1]. 시선 추적을 위해서 근적외선 기반, 컴퓨터 비전 기반 등 다양한 방법들이 연구되고 있다. 컴퓨터 비전 기반의 시선 추적 기술은 HMD(Head Mount Display)와 결합되어 사용자의 시선을 추적, 상호작용 분야에도 활용되고 있다. 하지만, 별도의 시선추적용 센서를 사용하거나 HMD와 같은 별도의 장비를 사용하므로 초기 비용이 발생할 뿐만 아니라 HMD를 착용하는 경우 사용자에 따라서 인터랙션시 몰입감 저하를 가지고 올 수 있다는 단점이 존재한다. 몰입감 저하는 단순히 사용자의 몰입도의 저하 뿐만 아니라 정확한 시선 특징 데이터의 수집을 어렵게 하여 이를 활용한 심리 분석, 콘텐츠에도 부정적인 영향을 끼칠 수 있다.

본 논문에서는 이러한 몰입도 저하를 최소화하기 위하여 단일 RGB 카메라를 사용한 2차원 영상으로부터 얼굴 영역과 얼굴 랜드마크(특징점)를 검출하여 시선 특징들을 추출하고, 시선을 추적하는 시스템을 구현한다.

Ⅱ. 본론

본 논문에서는 정면을 기준으로 상하좌우 및 대각 방향을 포함하여 총 9방향에 대한 시선 위치를 판단하는 시선 추적 시스템을 구현하였다.

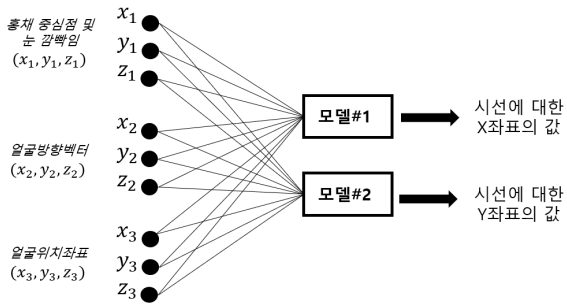


[그림 1] 얼굴 랜드마크 기반 측정 가능한 시선 추적 데이터

먼저, 3DDFA(3D Dense Face Alignment)[2] 모델을 통해 68개의 얼굴 랜드마크를 추출하였다. 추출된 얼굴 랜드마크를 기반으로 측정 가능한 시선 추적 관련 데이터는 [그림1]과 같으며 총 15개의 3차원 벡터 및 좌표값을 가진다. 시선 벡터는 우측 단안의 홍채 중심과 안구의 회전 중심축을 잇는 벡터로 시선 방향에 대한 주요 측정 지표로 사용되며, 홍채 중심의 경우 얼굴 랜드마크를 기반으로 취득된 눈 영역의 ROI(Region Of Interest)를 통해서 산출된다. 홍채 중심을 검출하기 위하여 원 형태의 홍채 특성 활성도가 높은 곳을 홍채 중심으로 검출하는 방사형 기울기 기반 알고리즘 사용하였으나 높은 해상도의 눈 영역 이미지에서 홍채 중심을 검출할 경우 기울기 측정을 위한 2D 컨볼루션 필터의 크기가 이미지 크기와 비례하여 연산량이 증가하여 실시간 검출 속도 저하가 발생하였다. 이러한 문제점을 해결하기 위하여 높은 해상도의 눈 이미지에서 홍채 중심점 검출을 위한 딥러닝 기반 홍채 중심점 검출 알고리즘을 개발하였다. 효율적인 연산과 실시간 처리를 위해 경량화 백본인 MobileNetV2[3] 모델을 사용하였고, UBIRIS.v2[4] 데이터베이스를 홍채 중심점 검출 모델 학습하였다.

시선 추적을 위해서는 안구분석을 통해 측정되는 시선 축의 방향뿐만 아니라 얼굴의 방향 또한 주요 특징으로 사용된다. 앞서 3DDFA로 추출한 랜드마크를 사용하여 얼굴의 위치 좌표와 얼굴 방향 벡터를 계산한다.

[그림 2]는 시선 특징(얼굴 방향 벡터, 얼굴 위치 좌표, 우측 단안 홍채 중심 좌표 및 눈 깜빡임)을 입력으로하여 응시점의 X좌표와 Y좌표를 출력하는 시선 추적 시스템의 개요를 나타낸다.



[그림 2] 시선 특징 기반 시선 추적 시스템 개요

Ⅲ. 실험

눈 깜빡임 정도는 HSV 채널 기반 피부 영역 검출을 통해 계산되며 눈 영역에서 피부가 아닌 영역의 비율을 통해 계산되며, [그림 3]은 차례대로 눈을 100(a), 60(b), 30(c), 0(d)% 뜬 상태의 HSV 채널 기반 피부 영역 검출 결과이다.

[그림 3] 눈 깜빡임 측정을 위한 눈영역 이미지 분할 결과

시선 추적 모델의 학습을 위하여 9개의 시선 특징을 실시간으로 취득하는 데이터 수집 프로그램을 개발하여 모델의 학습 데이터로 사용하였다. 수집 프로그램은 총 12개의 모니터 평면 포인트를 사용하며, 우측 상단 포인트부터 좌측 하단 포인트까지 순차적으로 진행된다. 수집되는 데이터는 9개의 시선 특징값이며 5명으로부터 약 4만개의 시선 데이터를 수집하였다. 딥러닝 기반 회귀모델을 통해서 시선추적모델을 학습한 결과 x축 오차율 8%, y축 오차율 10%를 확인하였다.

모니터 평면상에 응시하는 지점을 추론하는 시선추적모델을 적용한 소프트웨어는 아래 [그림4]와 같으며 9개의 시선 특징과 1포인트 캘리브레이션 과정을 통해서 대상이 응시하는 곳을 실시간으로 추론한다. 소프트웨어에는 대상자의 얼굴 랜드마크와 우측 단안에 적색점이 표시된다. 시선 추적을 위하여 얼굴이 정면을 바라볼 수 있도록 얼굴 각도 범위를 -5도에서 5도까지로 기준값을 지정하였다. 얼굴 방향 각도가 기준값을 충족하였을 때 [그림 4-(b)]와 같이 좌측 상단에 응시점이 나타난다.

(a) 얼굴 방향 벡터가 기준값에 맞추어지지 않은 경우

(b) 얼굴 방향 벡터가 기준값에 맞추어진 경우

[그림 4] 시선 추적 시스템 소프트웨어 구동 화면

Ⅳ. 결론

본 논문에서는 3DDFA를 통해 추출한 68개의 얼굴 랜드마크를 기반으로 홍채 중심점, 눈 깜빡임, 얼굴 방향 등 시선 추적에 사용되는 시선 특징을 검출하는 알고리즘을 개발하였다. 시선 특징들을 학습 데이터로 사용하여 시선 추적 모델 검증 결과, 시선 추적 응시점에 대한 x축 오차율 8%, y축 오차율 10%를 보였다. 시선 추적 오차율 개선을 위하여 추후 연구에서는 학습 데이터의 추가 수집 및 기존의 1포인트 방식의 시선 추적 알고리즘이 가지는 모니터의 크기, 사용자 위치 및 카메라 각도 등 환경 변화에 따라 정확도 및 시선 추적 안정성 저하 문제를 해결하기 위하여, 모니터 각 코너에 보정을 위한 포인트를 사용하여 시선 특징 공간을 모니터 공간으로 변환하는 perspective transform matric을 산출하여 시선 추적 알고리즘을 개선하려한다.

ACKNOWLEDGMENT

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2020년도 문화체육관광 연구개발사업으로 수행되었음(과제명: 소통약자의 소통능력 및 사회성 증진을 위한 음악활동 기반의 기능성 콘텐츠 플랫폼 개발, 과제번호: R2020040186, 기여율: 100%)

참 고 문 헌

[1] Karatekin, C., “Eye tracking studies of normative and atypical development.Developmental Review”, 27(3), pp. 283-348, 2007.

[2] Guo, Jianzhu, et al. “Towards fast, accurate and stable 3d dense face alignment.” European Conference on Computer Vision(ECCV). Springer, Cham, 2020.

[3] Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." Proceedings of the IEEE conference on computer vision and pattern recognition(CVPR). 2018.

[4] H. Proenca, S. Filipe, R. Santos, J. Oliveira and L. A. Alexandre, "The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-the-Move and At-a-Distance," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(8), pp. 1529-1535, 2010.