

논문 2020-1-6 <http://dx.doi.org/10.29056/jsav.2020.06.06>

사용자 상호작용 기반의 시선 검출을 위한 비강압식 캘리브레이션

이태균*, 유장희**†

Non-intrusive Calibration for User Interaction based Gaze Estimation

Tae-Gyun Lee*, Jang-Hee Yoo**†

요 약

본 논문에서는 웹 페이지 탐색 시 지속해서 발생하는 사용자 상호작용 과정을 이용하여 시선 검출을 위한 캘리브레이션 데이터를 획득하고, 사용자의 시선을 검출하는 동안 자연스럽게 캘리브레이션을 수행하는 방법에 관하여 기술하였다. 제안된 비강압식 캘리브레이션은 획득한 캘리브레이션 데이터를 이용하여 미리 학습된 시선 검출 CNN 모델을 새로운 사용자에게 적응하도록 보정하는 과정이다. 이를 위해 훈련을 통해서 시선을 검출하는 일반화된 모델을 만들고 캘리브레이션에서는 온라인 학습 과정을 통해 빠르게 새로운 사용자에게 적응하도록 하였다. 실험을 통하여 다양한 사용자 상호작용의 조합으로 시선 검출 모델을 캘리브레이션 하여 성능을 비교하였으며, 기존 방법 대비 개선된 정확도를 얻을 수 있었다.

Abstract

In this paper, we describe a new method for acquiring calibration data using a user interaction process, which occurs continuously during web browsing in gaze estimation, and for performing calibration naturally while estimating the user's gaze. The proposed non-intrusive calibration is a tuning process over the pre-trained gaze estimation model to adapt to a new user using the obtained data. To achieve this, a generalized CNN model for estimating gaze is trained, then the non-intrusive calibration is employed to adapt quickly to new users through online learning. In experiments, the gaze estimation model is calibrated with a combination of various user interactions to compare the performance, and improved accuracy is achieved compared to existing methods.

한글키워드 : 시선 검출, 비강압식 캘리브레이션, 사용자 상호작용, 합성곱 신경망

keywords : gaze estimation, non-intrusive calibration, user interaction, convolutional neural network

* 과학기술연합대학원대학교 ICT전공

** 한국전자통신연구원 인공지능연구소

† 교신저자: 유장희(email: jhy@etri.re.kr)

접수일자: 2020.05.17. 심사완료: 2020.06.01.

게재확정: 2020.06.19.

1. 서 론

시선(gaze)은 관찰자가 눈을 통해 특정 사물을

보는 방향을 의미한다. 즉, 어떠한 사람의 시선을 알게 되면 그 사람이 무엇을 보고 있는지를 알 수 있다[1]. 사람은 일상생활 동안 끊임없이 시선을 움직이며, 시선 움직임에는 특정 사물에 대한 관심이 선행된다[2]. 따라서 시선 움직임을 분석하면 그 사람이 무엇에 관심이 있는지를 파악할 수 있다. 시선 움직임은 시선의 순간적 움직임(saccade)과 시선의 고정(fixation) 단계로 구분된다[3]. 시선의 순간적 움직임은 시선이 한 지점에서 다른 지점으로 빠르게 이동하는 것을 의미하며, 시선의 고정은 시선이 한 지점에 일정 시간 정지해 있는 것을 의미한다. 특히, 시선의 순간적 움직임은 특정 사물에 대한 관심의 결과로 나타나며, 시선의 고정은 시선의 순간적 움직임이 끝난 뒤에 나타나므로 그 지점에는 관심의 대상이 위치하게 된다. 최근 들어 이러한 시선 정보는 게임[4], 심리분석[5], 의료진단[6][7], 마케팅 등 매우 다양한 분야에 적용되고 있다.

이와 같이 시선 검출은 사람의 무의식적 관심을 인지할 수 있는 방법이지만 이를 개인 맞춤형 광고[8] 등 실세계 문제에 활용하기 위해서는 극복해야 할 많은 과제가 있다. 특히 현재 개발된 대부분의 시선 검출 방법은 적외선 조명과 같은 추가적인 장비를 사용하고 있으며[9][10], 이는 RGB 영상만을 다루는 일반 사용자의 PC 및 모바일 환경과 일치하지 않는다. 또한, 보다 정확한 시선의 검출을 위해서는 지루하고 복잡한 캘리브레이션(calibration) 과정이 선행되어야 한다. 캘리브레이션은 사용자들 간의 시선 특성을 반영하여, 시선 검출의 정확도를 높이기 위한 과정으로 시선 검출을 시작하기 전에 반드시 수행해야 하는 필수적 과정이다. 캘리브레이션을 위해서는 일반적으로 사용자의 참여가 필수적이며, 사용자는 정해진 방법에 따라 특정 위치를 응시하는 등 반복적 과정을 필요로 한다[9-12].

한편, 적외선 조명 기반의 다양한 시선 검출과 캘리브레이션 방법들이 제안되었다. 그림 1은 적외선 기반의 시선 검출 시스템에서 카메라와 적외선 조명의 배치 예를 보여주고 있다. 카메라는 모니터의 위 또는 아래에 위치하고 적외선 조명은 모니터의 각 모서리에 설치하거나 카메라 주변에 장착하여 시선을 검출하는 데 사용한다[9][10]. 다음으로 캘리브레이션을 위해서 모니터에 일정한 간격의 포인트가 순차적으로 나타나고, 이때 모니터에 나타난 포인트와 사용자가 해당 포인트를 바라보는 동안 추정된 시선 검출 결과를 비교하여 캘리브레이션을 수행하게 된다. 캘리브레이션 포인트의 개수는 1개[9]부터 9개[10], 24개[11] 그리고 29개[12]에 이르기까지 다양한 방법들이 사용되고 있다. 일반적으로 캘리브레이션 포인트의 개수가 많을수록 시선 검출의 정확도가 높아지지만, 캘리브레이션 과정에 걸리는 시간이 늘어나는 단점이 있다. 또한, 기존 캘리브레이션 방법들의 공통적인 문제는 얼굴의 포즈(pose)에 대한 의존성이 높아 캘리브레이션 이후에 발생하는 얼굴 포즈의 변화에 따라 시선 검출 성능이 크게 떨어질 수 있다는 것이다.

최근의 연구에서는 이러한 캘리브레이션 과정에서 발생하는 문제들을 극복하기 위하여, 캘리브레이션 과정을 없애거나 포인트 대신 사진을 이용하는 등의 비교적 자연스러운 과정을 통해 문제를 해결하는 방법들이 제안되고 있다. 대표적으로 다수의 카메라와 다수의 광원을 사용하여 캘리브레이션 없이 시선을 추정하는 방법[13][14]과 사용자가 사진을 보는 동안에 사진의 관심 영역 지도(saliency map)와 추정된 시선 분포 사이의 거리를 최소화하는 방법[15][16] 등이 제안되었다. 그러나 이러한 방법은 별도의 캘리브레이션 과정 없이 자연스러운 시선 검출이 가능하나 시선 정확도가 비교적 낮다는 단점이 있다.

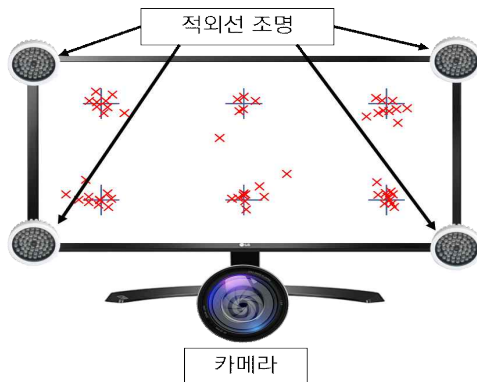


그림 1. 적외선 기반 시선 검출 시스템
Fig. 1. Gaze estimation system based on IR

본 논문에서는 별도의 조명이 없는 웹캠 사용 환경에서 사용자가 자연스럽게 컴퓨터를 사용하거나 콘텐츠와 상호작용하는 과정에서 캘리브레이션을 수행할 수 있는 비강압식(non-intrusive) 캘리브레이션 방법을 제안하였다. 이를 위해 정의된 상호작용이 발생한 위치에 사용자 시선의 고정(fixation)이 이루어지는 것으로 가정하였다. 즉, 상호작용 과정에서 사용자의 시선이 고정된 화면 좌표 데이터를 수집하고, 검출된 시선과 비교하여 캘리브레이션을 수행하게 된다. 그리고 시선 검출과 동시에 캘리브레이션을 수행하기 위해, 사용하는 시선 검출 CNN 모델[17]에 온라인 학습(online learning)[18] 방법을 적용하였다. 제안된 방법을 검증하기 위해 공개 데이터셋을 이용하여 시선 검출 모델을 훈련하였으며, 시험을 통하여 다양한 조합의 상호작용에 따른 시선 검출의 성능을 비교하였다. 또한, 제안된 방법을 사용했을 때 기존 방법보다 더 높은 정확도로 시선 검출이 가능하다는 것을 확인하였다.

본 논문의 구성은 2장에서 제안된 비강압식 캘리브레이션 방법에 대하여 기술하였으며, 3장에서는 2장에서 제시한 방법을 검증하기 위한 실험 및 결과를 보여주고, 4장에서 결론과 향후 연구 방향을 제시하였다.

2. 시선 검출을 위한 비강압식 캘리브레이션

본 논문에서 제안하는 시선 검출을 위한 캘리브레이션은 미리 학습된 시선 검출 모델의 정확도를 개선하기 위해 기존 모델을 보정(tuning)하는 과정이다. 그림 2는 비강압식 캘리브레이션 방법이 어떻게 시선 검출에 작용하는지를 나타내고 있다. 사용자 상호작용이 발생하는 순간에 수집된 데이터는 캘리브레이션을 위해 처리되고 시선 검출 모델에 입력되어 위치 좌표의 정확도를 보정하는 데 사용된다. 결과적으로 제안된 방법을 사용하면 상호작용이 많이 발생할수록 해당 사용자에게 맞추어 성능이 개선된다. 다음에서는 사용자 상호작용의 유형을 정의하고 이에 따라 캘리브레이션 데이터를 구성하는 방법과 이를 이용하여 시선 검출 모델을 훈련하고 캘리브레이션 하는 과정을 기술한다.

2.1 캘리브레이션을 위한 데이터 획득

사용자가 웹페이지를 사용하면서 발생하는 상호작용의 종류에는 여러가지를 생각할 수 있으나 본 논문에서는 마우스 클릭, 마우스 이동, 그리고 키보드 입력을 정의하였다. 캘리브레이션을 위해서는 정의된 모든 상호작용이 발생하는 즉시 시선 검출을 위해 얼굴 영상이 촬영된다. 시선 검출 모델을 보정하기 위한 사용자 시선의 위치는 정의된 상호작용이 발생한 화면상의 좌표와 일치하는 것으로 가정하였다.

마우스 커서 위치와 시선의 위치 사이의 관계를 밝히기 위한 여러 연구가 있었다. Liebling [19] 등은 마우스 클릭을 위해 처음 마우스를 이동할 때는 화면에서 커서와 시선 사이의 거리가 200px(pixel) 이상이지만 이동할수록 그 거리가 점점 가까워지고 마우스 클릭이 발생할 때는 완

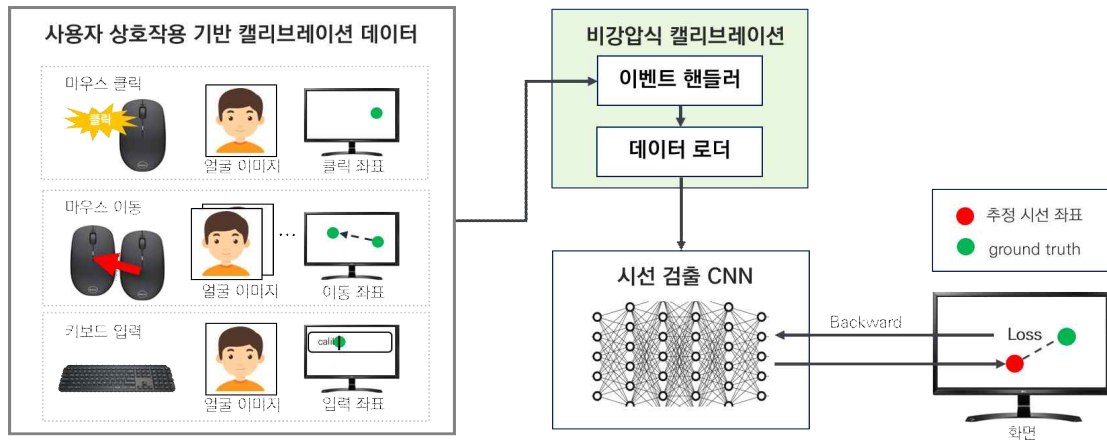


그림 2. 시선 검출을 위한 비강압식 캘리브레이션 방법
Fig. 2. Non-intrusive calibration method for gaze estimation

전히 일치함을 보였다. Huang[20] 등에 의하면 클릭을 위해 마우스를 이동할 때의 평균 거리가 77px, 클릭할 때의 평균 거리가 74px 임을 보였다. 이와 같이 화면에서 시선과 커서 사이의 거리는 수행되는 작업에 따라 다르지만, 마우스 이동보다는 클릭에서 더 짧은 것을 보여주고 있다. 따라서 마우스를 클릭할 때 수집되는 데이터는 모두 캘리브레이션에 사용하고, 마우스 이동의 경우에는 비교적 마우스 커서와 시선의 위치가 가까운 클릭 이전 200ms까지의 데이터를 캘리브레이션에 사용하였다.

키보드 입력의 경우에는 화면을 보고 입력하는 유형과 키보드를 보고 입력하는 유형으로 나뉜다. 키보드를 보고 입력하는 경우는 커서와 시선이 정확히 일치하지 않을 것이므로 캘리브레이션 데이터에서 제외하였다. Papoutsaki[21] 등에 따르면, 화면을 보고 입력하는 경우는 커서와 시선 사이의 거리가 평균 160px이고 키보드를 보고 입력하는 경우는 평균 352px로 키보드를 보고 입력했을 때의 평균 거리가 2배보다 더 크다. 따라서 키보드 입력 유형을 판단하기 위해, 문턱

값(threshold value) t 를 설정하고 이 값을 넘지 않는 경우에만 화면을 보고 입력한다고 판단하여 캘리브레이션 데이터로 사용하였다.

2.2 CNN 기반 시선 검출 모델

시선 검출을 위한 CNN 모델의 구조는 Zhang [17] 등이 제안한 모델에 기반하였으며, 그림 3에 제시 되었다. CNN 모델에 입력되는 영상은 검출된 얼굴 영역만 사용하며, AlexNet[22]의 합성곱층(convolutional layer)을 사용하여 특징 맵(feature map)을 생성한다. 이후 1×1 합성곱층을 통해 특징 맵에 대한 가중치 맵(weight map)을 만들고, 가중치 맵은 시선 검출을 위해 얼굴에서 중요한 영역에 높은 가중치를 주도록 훈련된다. 가중치가 적용된 특징 맵은 완전 연결층(fully connected layer)을 통해 화면에서의 시선 좌표로 변환된다.

훈련을 위해서는 얼굴 영상과 실제 시선 좌표(ground-truth)가 하나의 셋(set)으로 입력된다. 실제 시선 좌표와 추정된 좌표 사이의 오차는 식

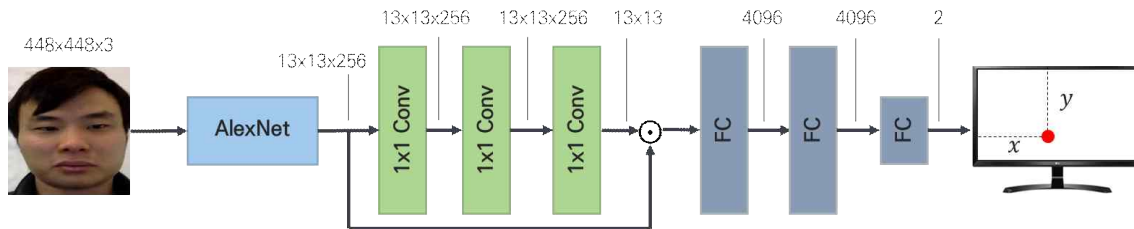


그림 3. 시선 검출을 위한 CNN 아키텍처
Fig 3. CNN architecture for gaze estimation

(1)과 같은 L 오류 함수를 사용하여 계산된다.

$$L = \| \hat{g} - g \|_1 \quad (1)$$

$$= |x' - x| + |y' - y|$$

식에서 $\hat{g}(x', y')$ 은 추정된 시선 좌표이며, $g(x, y)$ 는 훈련 데이터에서 주어진 실제 시선 좌표이다. 훈련 이후에 모델이 올바르게 훈련되었는지 확인하기 위해 테스트를 진행하며, 훈련에서와 동일한 L 오류 함수를 사용한다.

2.3 비강압식 캘리브레이션

앞서 기술했듯이 캘리브레이션은 훈련이 완료된 시선 검출 모델을 개별 사용자에게 적응시키기 위해 검출된 시선 좌표를 보정하는 과정이다. 즉, 시선 검출에 사용된 CNN 모델은 수집된 캘리브레이션 데이터를 이용하여 보정된다. 제안된 방법은 캘리브레이션 데이터의 수집이 충분히 진행된 후에 수행되는 것이 아닌, 사용자 상호작용이 발생하여 한 셋의 데이터가 수집되면 즉시 캘리브레이션에 사용된다. 이전 연구들에서는 캘리브레이션이 시선 검출 전에 미리 수행되었지만 본 연구에서 제안하는 방법은 시선 검출을 하면서 발생하는 상호작용과 함께 수행된다. 따라서 사용자 관점에서 매우 자연스럽게 편리한 방식이라 할 수 있다.

2.1절에서 기술한 그림 2와 같이, 정의된 사용자 상호작용이 발생하면 이벤트 핸들러(event handler)가 이를 감지하여 데이터 수집을 위한 신호를 데이터 로더(data loader)에 보낸다. 그리고, 데이터 로더는 얼굴 영상과 상호작용이 발생한 화면상의 위치인 좌표 정보를 획득하며, 이를 시선 검출 모델에서 사용할 수 있는 형태로 변환하여 전달한다. 전달된 데이터를 이용하여 시선 검출 모델은 실제 시선 좌표와 검출된 시선 좌표의 위치 차이를 식(1)의 L 오류 함수를 이용하여 계산하고, 오류를 줄이기 위해 학습을 통하여 모델을 최적화한다. 이렇게 사용자 상호작용 과정에서 발생하는 데이터를 이용하여, 온라인 학습으로 캘리브레이션이 진행되므로 시스템을 사용하면 할수록 해당 사용자에게 대한 시선 검출의 성능 향상을 기대할 수 있다.

3. 실험 및 결과 분석

3.1 실험 데이터 구성

시험을 위하여 WebGazer 공개 데이터셋[21]을 사용하였다. WebGazer 데이터셋은 51명으로부터 수집되었으며, 데이터셋의 구성은 웹 페이지를 탐색하는 동안의 얼굴 영상과 Tobii 상용 시선 추적기(eye tracker)[23]로 검출된 시선 좌표, 상호작용

표 1. 상호작용 시험을 위한 데이터셋의 구성
Table 1. Dataset for interaction test

사용자 상호작용	개수 (개)
마우스 클릭 (C)	1,419
마우스 이동 (M)	23,875
키보드 입력 (K)	69,225
전체	83,709

표 2. 비상호작용 시험을 위한 데이터셋의 구성
Table 2. Dataset for non-interaction test

데이터 분류	개수 (개)
시선 검출 모델 훈련	567,463
시선 검출 모델 테스트	141,866
전체	709,329

로그 기록이 포함된다. 데이터는 수집 시의 상호작용 발생 유무에 따라 상호작용 데이터와 비상호작용 데이터로 구성된다. 시험에 사용된 각각의 데이터셋 구성을 표 1과 표 2에 나타내었다. 상호작용 데이터는 시선 검출 모델의 캘리브레이션에 사용되었으며, 비상호작용 데이터는 시선검출 모델의 훈련 및 테스트에 사용되었다. 상호작용 데이터는 2장에서 정의한 상호작용의 종류에 따라 분류하였으며, 하나의 영상에 대해 두 개 이상의 상호작용이 중복될 수 있다. 비상호작용 데이터는 훈련 데이터와 테스트 데이터로 나누어지며, 구성 비율은 약 8:2 이다.

3.2 실험 절차 및 결과

CNN 모델을 훈련하기 위한 초기 설정에서 batch size는 128, epoch은 15로 하였다. 그리고 학습률(learning rate)은 0.01으로 설정하고, 학습률 감쇠(decay)는 사용되지 않았다. 캘리브레이션에 사용된 각 사용자 상호작용이 시선 검출 결과에 미치는 영향을 분석하기 위해 여러 상호작용의 조합으로 실험을 수행하였다. 상호작용의 조합은 표

1에서 정의된 것과 같이 3개의 상호작용에 대하여, C, C+M, C+K, C+M+K의 4가지로 구성하였다. 각 조합에서의 캘리브레이션은 독립적으로 수행함으로써, 서로 성능에 영향을 미치지 않도록 하였다. 또한, 각 조합에서 사용하는 데이터의 개수를 1,000개로 동일하게 하여, 데이터 개수에 따른 바이어스 가능성을 제거하였다.

표 3. 시선 검출 실험 결과
Table 3. Experimental results for gaze estimation

상호작용 조합	Hamming 거리 (px)
기준 연구 [24]	169.0
C	160.1
C+M	162.2
C+K	150.6
C+M+K	154.7

표 3은 정의된 사용자 상호작용의 조합을 통한 시선 검출 모델의 실험 결과를 보여주고 있다. 성능 비교를 위하여, 마우스 클릭 시에 획득한 데이터의 ridge regression에 기반한 기준 연구[24]의 결과를 함께 제시하였다. 성능 측정은 실제 시선 좌표와 검출된 시선 좌표의 거리 차이를 유클리드 거리(Euclidean distance)로 계산하였다. 제시된 실험 결과에서 모든 유형의 상호작용 데이터를 사용한 경우(C+M+K) 보다 마우스 클릭(C)과 키보드 입력(K) 데이터를 조합(C+K)한 경우가 가장 좋은 성능(4.1px 오류 감소)을 보여주었다. 또한, 마우스 클릭과 이동 데이터를 함께 사용한 경우(C+M)가 마우스 클릭만 사용한 경우(C)보다 오류가 2.1px 만큼 증가하였다. 이와 같은 결과를 종합해보면 마우스 이동 데이터(M)를 사용하는 경우는 오히려 성능이 감소되었지만, 기준 연구에서 제시된 성능보다는 좋은 결과를 보여주었다. 반면에 마우스 클릭(C)만 사용했을 때보다는 키보드 입력 데이터를 함께 사용한 경우(C+K)의 성능이 약

10px만큼 향상되었으며, 마우스 클릭과 마우스 이동을 함께 사용하고(C+M), 키보드 입력(K)을 추가한 경우(C+M+K)에도 성능이 7.5px 향상되었다. 이러한 결과를 통해 상호작용 기반의 캘리브레이션 과정에서 키보드 입력 데이터(K)를 함께 사용하면 시선 검출 성능에 긍정적인 영향을 준다는 것을 확인할 수 있었다. 결과적으로 실험에서 가장 높은 성능을 보인 조합은 마우스 클릭과 키보드 입력(C+K)을 모두 사용한 경우로 기존의 연구 방법과 비교하여 18.4px 향상 되었다.

4. 결 론

본 논문에서는 자연스러운 방식으로 캘리브레이션을 수행하기 위해 웹 페이지를 탐색하는 동안 발생하는 사용자 상호작용 과정에서 수집할 수 있는 데이터를 이용하여, 시선 검출 모델을 보정하는 비강압식 캘리브레이션 방법을 제안하였다. 제안된 방법을 검증하기 위해 다양한 상호작용 조합으로 실험을 진행한 결과 마우스 클릭과 키보드 입력(C+K)의 상호작용 조합이 가장 좋은 시선 검출 결과를 보였으며, 기존 방법 대비 10% 이상의 성능 향상을 보여주었다.

향후 연구에서는 보다 다양한 서비스의 개발을 위하여, 사용자의 시선이 고정되는 위치를 바탕으로 콘텐츠에 포함된 사물이 무엇인지를 인식하고 분류하는 연구가 필요할 것이다. 이를 통해 개개인이 보다 관심을 가질 수 있는 맞춤형 인터넷 광고를 노출하는 등 다양한 분야에 활용할 수 있을 것이다.

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 수행된 연구임. (2019-0-00330, 영유아/아동의 발달장애 조기선별을 위한 행동·반응 심리인지 AI 기술 개발)

참 고 문 헌

- [1] X. Wang, K. Liu, and X. Qian, "A Survey on Gaze Estimation", in Proceeding of the 10th International Conference on Intelligent Systems and Knowledge Engineering, pp.260-267, Taipei, Taiwan. Nov. 2015. DOI: <https://doi.org/10.1109/iske.2015.12>
- [2] K. Rayner, "Eye Movements in Reading and Information Processing: 20 Years of Research", Psychological Bulletin, vol.124, no.3, pp.372-422, Nov. 1998. DOI: <https://doi.org/10.1037/0033-2909.124.3.372>
- [3] G. Buscher, A. Dengel, and L. van Elst, "Eye Movements as Implicit Relevance Feedback", in Proceeding of the CHI'08 Extended Abstracts on Human Factors in Computing Systems, pp.2991-2996, Florence, Italy, Apr. 2008. DOI: <https://doi.org/10.1145/1358628.1358796>
- [4] L. E. Nacke, S. Stellmach, D. Sasse, and C. A. Lindley, "Gameplay Experience in a Gaze Interaction Game", arXiv preprint arXiv:1004.0259, 2010. URL: <https://arxiv.org/abs/1004.0259>
- [5] A. Navab, L. Gillespie Lynch, S. P. Johnson, M. Sigman, and T. Hutman, "Eye Tracking as a Measure of Responsiveness to Joint Attention in Infants at Risk for Autism", Infancy, vol.17, no.4, pp.416-431, Jul. 2012. DOI: <https://doi.org/10.1111/j.1532-7078.2011.00082.x>
- [6] M. Eizenman, H. Y. Lawrence, L. Grupp, and E. Eizenman, "A Naturalistic Visual Scanning Approach to Assess Selective Attention in Major Depressive Disorder", Psychiatry Research, vol.118, no.2, pp. 117-128, Jun. 2003. DOI: [https://doi.org/10.1016/s0165-1781\(03\)00068-4](https://doi.org/10.1016/s0165-1781(03)00068-4)
- [7] S. S. Deepika and G. Murugesan, "A Novel Approach for Human Computer Interface based on Eye Movements for

- Disabled People”, in Proceeding of the IEEE International Conference on Electrical, Computer and Communication Technologies, pp.1-3, Coimbatore, India, Mar. 2015. DOI: <https://doi.org/10.1109/icecct.2015.7226124>
- [8] J. P. Johnson, “Targeted Advertising and Advertising Avoidance”, The RAND Journal of Economics, vol.44, no.1, pp.128-144, Apr. 2013. DOI: <https://doi.org/10.2139/ssrn.2018938>
- [9] E. D. Guestrin and M. Eizenman, “Remote Point-of-Gaze Estimation Requiring a Single-point Calibration for Applications with Infants”, in Proceeding of the 2008 Symposium on Eye Tracking Research & Applications(ETRA), pp.267-274, Savannah, USA, Mar. 2008. DOI: <https://doi.org/10.1145/1344471.1344531>
- [10] Y. Durna and F. Ari, “Design of a Binocular Pupil and Gaze Point Detection System Utilizing High Definition Images”, Applied Sciences, vol.7, no.5, pp.1-16, May. 2017. DOI: <https://doi.org/10.3390/app7050498>
- [11] M. S. A. bin Suhaimi, K. Matsushita, M. Sasaki, and W. Njeri, “24-Gaze-Point Calibration Method for Improving the Precision of AC-EOG Gaze Estimation”, Sensors, vol.19, no.17, pp.1-13, Aug. 2019. DOI: <https://doi.org/10.3390/s19173650>
- [12] K. Harezlak, P. Kasprowski, and M. Stasch, “Towards Accurate Eye Tracker Calibration - Methods and Procedures”, Procedia Computer Science, vol.35, pp.1073-1081, Sep. 2014. DOI: <https://doi.org/10.1016/j.procs.2014.08.194>
- [13] T. Nagamatsu, J. Kamahara, and N. Tanaka, “Calibration-free Gaze Tracking using a Binocular 3D Eye Model”, in Proceeding of the CHI’09 Extended Abstracts on Human Factors in Computing Systems, pp.3613-3618, Boston, USA, Apr. 2009. DOI: <https://doi.org/10.1145/1520340.1520543>
- [14] S. W. Shih, Y. T. Wu, and J. Liu, “A Calibration-free Gaze Tracking Technique”, in Proceeding of International Conference on Pattern Recognition (ICPR), pp. 201-204; Barcelona, Spain, Sep. 2000. DOI: <https://doi.org/10.1109/icpr.2000.902895>
- [15] K. Wang, S. Wang, and Q. Ji, “Deep Eye Fixation Map Learning for Calibration-free Eye Gaze Tracking”, in Proceeding of the 9th Biennial ACM Symposium on Eye Tracking Research & Applications (ETRA), pp.47-55; Charleston, USA, Mar. 2016. DOI: <https://doi.org/10.1145/2857491.2857515>
- [16] Y. Sugano, Y. Matsushita, and Y. Sato, “Calibration-free Gaze Sensing using Saliency Maps”, in Proceeding of the International Conference on Computer Vision and Pattern Recognition (CVPR), pp.2667-2674; San Francisco, USA, Jun. 2010. DOI: <https://doi.org/10.1109/cvpr.2010.5539984>
- [17] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, “It’s Written All over Your Face: Full-face Appearance-based Gaze Estimation”, in Proceeding of the International Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp.51-60; Honolulu, USA, Jul. 2017. DOI: <https://doi.org/10.1109/cvprw.2017.284>
- [18] D. Sahoo, Q. Pham, J. Lu, and S. C. Hoi, “Online Deep Learning: Learning Deep Neural Networks on The Fly”, in Proceeding of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI), pp.2660-2666, Melbourne, Australia, Aug. 2017. DOI: <https://doi.org/10.24963/ijcai.2018/369>
- [19] D. J. Liebling and S. T. Dumais, “Gaze and Mouse Coordination in Everyday Work”, in Proceeding of the 2014 ACM International Joint Conference on

- Pervasive and Ubiquitous Computing: Adjunct Publication, pp.1141-1150, Seattle, USA, Sep. 2014. DOI: <https://doi.org/10.1145/2638728.2641692>
- [20] J. Huang, R. White, and G. Buscher, "User See, User Point: Gaze and Cursor Alignment in Web Search", in Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, pp.1341-1350, Austin, USA, May. 2012. DOI: <https://doi.org/10.1145/2207676.2208591>
- [21] A. Papoutsaki, A. Gokaslan, J. Tompkin, Y. He, and J. Huang, "The Eye of the Typer: a Benchmark and Analysis of Gaze Behavior during Typing", in Proceeding of the 2018 ACM Symposium on Eye Tracking Research & Applications (ETRA), pp.1-9; Warsaw, Poland, Jun. 2018. DOI: <https://doi.org/10.1145/3204493.3204552>
- [22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", in Proceeding of Advances in Neural Information Processing Systems (NIPS), pp.1097-1105; Lake Tahoe, USA, Dec. 2012. DOI: <https://dl.acm.org/doi/10.5555/2999134.2999257>
- [23] <https://www.tobiipro.com/ko/product-listing/tobii-pro-x3-120/>, 2020.
- [24] A. Papoutsaki, P. Sangkloy, J. Laskey, N. Daskalova, J. Huang, and J. Hays, "Webgazer: Scalable Webcam Eye Tracking using User Interactions", in Proceeding of the 25th International Joint Conference on Artificial Intelligence (IJCAD), pp.3839-3845, New York, USA, Jul. 2016. URL: <https://dl.acm.org/doi/10.5555/3061053.3061156>

저자 소개



이태균(Tae-Gyun Lee)

2018.02 한밭대학교 컴퓨터공학과 졸업
 2018.90-현재 : 과학기술연합대학원대학교
 ICT전공 석사과정
 <주관심분야> 인공지능, 딥러닝, 영상처리



유장희(Jang-Hee Yoo)

1988.02 한국외국어대학교 물리학과 졸업
 1990.02 한국외국어대학교 전산학과 석사
 2004.07 영국 University of Southampton
 전자 및 컴퓨터과학 박사
 1989.11-현재: 한국전자통신연구원
 인공지능연구소 책임연구원
 2005.09~현재: 한국저작권위원회 감정인,
 감정위원
 2007.03-현재: 과학기술연합대학원대학교
 ICT전공 전임교수
 2007.01~현재: 한국SW감정평가학회 이사
 2014.01~현재: 경찰청 과학수사자문위원
 2014.8~2015.8: University of Washington
 방문학자
 2018.3~2020.3: 국가지식재산위원회
 전문위원
 <주관심분야> 컴퓨터 비전, 인공지능, 생
 체인식, 휴먼모션분석, HCI 및 지능로봇 등