

적응형 확대 기반의 시선 추적을 통한 정확도 및 정밀도 개선*

송현주¹, 조재민², 김보형³, 서진욱⁴

서울대학교 컴퓨터공학부^{1,2,4}, 분당서울대학교병원 영상의학과³

hjsong@hcil.snu.ac.kr, jmjo@hcil.snu.ac.kr, bhkim@snu.ac.kr, jseo@snu.ac.kr

Adaptive zoom-based gaze tracking for enhanced accuracy and precision

Hyunjoo Song¹, Jaemin Jo², Bohyoung Kim³, Jinwook Seo⁴

Dept. of Computer Science and Engineering, Seoul Nat'l Univ.^{1,2,4},

Dept. of Radiology, Seoul Nat'l Univ. Bundang Hospital³

요 약

카메라 기반의 원격 시선 추적 방식은 기기의 특성상 사용자의 움직임, 주변 조도 등에 의해 시선 추적의 정확성 측면에서 많은 영향을 받는다. 하지만 모든 변인을 통제하는 것은 힘들 뿐 아니라 사용자의 움직임을 제약하는 실험 환경은 시선 추적의 정확성은 높일 수 있으나 비침습적인 시선 추적 방식의 이점이 사라진다. 이 논문에서는 이를 극복하기 위하여 높은 정확도가 요구되는 관심 영역을 두 종류의 조건(상시, 조건부)에 의해 확대함으로써 시선 추적의 해상력을 높이고자 했고, 이를 기존 시선 추적 방식과 비교하였다.

1. 서 론

시선 추적 장비(gaze tracker)의 정확도(accuracy: 목표와의 거리 차이)와 정밀도(precision: 재현성, 일관성)는 시선 추적을 통한 상호작용 및 사용자의 시선 분포 분석 과정에서 매우 중요하다[1]. 이는 관심 영역(region of interest)의 크기가 작은 환경에서 더욱 강조되는데, 예를 들어 시선 추적기가 목표를 선택하기 위해 활용되는 경우, 목표의 크기보다 정확도가 높아야 안정적인 사용이 가능하다. 또한 문서를 읽을 때의 시선 패턴을 분석하는 실험의 경우 글자나 줄 단위의 분석을 위해서는 시선 추적 장비의 정확도와 정밀도가 일정 수준을 상회해야만 올바른 분석 결과를 도출할 수 있다.

하지만 카메라를 기반으로 한 원격 시선 추적 방식은 정확도와 정밀도가 제조사들이 보고하고 있는 사양보다 못 미치는 것으로 보고되고 있다[2]. 제조사들이 발표한 사양은 카메라를 기반으로 한 시선 추적 방식의 특성상 사용자의 머리 움직임이나 주변의 조도 등이 매우 이상적으로 통제된 상황에서 측정되었기 때문이다. 따라서 같은 성능을 기대하기 위해서는 턱받침 등을 사용하는 특수한 실험 환경 조성이 불가피하게 된다.

이에 따라 정확도와 정밀도를 높이기 위한 다양한 기법들이 연구되었다. 시선 추적을 시작하기 전에 진행되는 보정(calibration) 과정을 개선하는 형태로는 보정 지점(calibration point)의 수를 늘리거나 관심 영역이 몰려있는 곳에 보정 지점을 집중하여 배치하는 방안이 제시되었다. 그러나 이 경우 보정에 걸리는 시간이 늘어나거나, 보정이 되지 않는 영역에 대한

정확도가 낮아지는 단점이 있다. 정확도 개선을 위한 다른 접근 방법으로는 턱받침을 활용하여 사용자의 움직임을 제한하는 방안도 있으나, 비침습적인 원격 시선 추적 방식의 장점을 상쇄시킨다는 한계가 있다.

이 연구에서는 시선 추적 과정에서 사용자 보게 되는 화면상의 대상을 확대함으로써 카메라 기반의 시선 추적기가 가지고 있는 비침습적인 특성을 유지하면서 추가적인 보정 과정 없이 시선 추적 결과의 정확도를 향상시킬 수 있는 방안을 제시한다.

2. 시선 추적 대상의 적응형 확대 모델

시선 추적 장비의 정확도 향상을 위해 화면 위에 표시된 대상을 확대하고자 할 때 시선 추적기의 특성상 시야각(visual angle)을 키워야 하기에 화면과 사용자의 거리는 유지한 상태에서 대상의 크기를 늘리는 것이 필요하다. 하지만 시선 추적 장비의 동작 거리나 인식 가능한 각도의 한계 때문에 단순히 보다 큰 화면을 사용하는 접근 방법은 적용이 어렵다. 따라서 본 연구에서는 화면 크기는 유지한 상태에서 동적으로 화면상에 표시되는 대상의 크기를 렌즈 형태로 확대함으로써 정확도를 높이고자 했다(그림 1). 이를 위해 확대와 관련하여 고려해야 할 주요 인자들을 표 1과 같이 정리했다. 이들 중 일부는 복수의 산정 방안 중 한가지를 선택했으며 이번 연구에서 사용된 모델에서 선택된 것은 별표(*)로 표시했다.

해당 인자들 중 공간적 인자들의 값은 다음과 같은 기준으로 산정하도록 했다. 우선 확대를 해야 할 위치는 현재 사용자가 보고 있는 영역을 시선 추적 장비로부터 쉽게 얻을 수 있기에 이를 기준으로 했다. 다만 사람의 안구는 미세한 움직임(fixation jitter)이 지속되기 때문에 grab-and-hold(GHA) 알고리즘[3]을 도입하여 확대된

* This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. NRF-2014R1A2A2A03006998).



그림 1 적응형 대상 확대를 통한 시선 추적
 (A) 최초 측정된 시선 위치 (B) 사용자가 의도한 시선 위치 (C) 확대에 의해 보다 자세히 특정된 시선 위치 (D) 렌즈 형태로 확대된 영역

영역 크기 내에서 일어나는 시선 이동에 대해서는 최초의 위치를 유지하도록 했다. 확대 영역의 경우, 화면 전체를 확대하거나 관심 영역 주위만 한정하여 확대하는 방안이 있었다. 그러나 전체 영역의 확대는 주변 정보를 잃게 되는 단점이 있으며, GHA 알고리즘의 도입으로 인하여 최초 시선이 머물렀던 지점에서 이동을 할 때 추가적인 상호 작용 방안이 요구된다. 또한 확대가 일시적으로 이루어질 경우 원래 의도했던 목표 지정보다 더 가게 되는 문제(overshooting problem)가 발생기 쉽다는 단점이 있다. 이에 따라 최종 모델은 관심 영역만 확대하도록 했다.

확대 모델의 시간적 인자들 중 대상 확대의 지속 시간에 대해서는 항상 확대하는 방안과 필요에 의해 일시적으로 확대하는 방안 2가지를 고려했다. 하지만 기존 연구[4]에서의 비교 결과 안구의 단속성 운동(saccade) 도중에는 렌즈 형태의 확대 영역을 감추는 것이 작업의 수행 속도와 정확성 측면에서 더 뛰어났기에 최종 모델에서는 일시적인 확대를 선택하게 되었다. 아울러 이 경우 확대 렌즈의 표시 시점을 사용자가 수동으로 선택하거나 자동으로 띄워주는 2가지의 방안을 생각할 수 있었는데, 선행 연구[5] 결과에서 보고된 손과 시선 움직임의 차이를 감안하여 일정 시간(500ms) 이상 관심 영역에 머무를 경우 자동으로 렌즈가 표시되도록 모델을 설정했다.

마지막으로 확대 비율과 확대된 부분을 보여줄 보여줄 렌즈(그림 1D)의 크기는 기대 정확도를 기반으로 도출하였다. 이 중 확대 배율의 계산에 사용한 기대 정확도(e.g. 0.5°)는 3장에 설명된 예비 실험을 통해 구했으며, 실제 정확도(e.g. 1.0°)를 기대 정확도로 나눈으로써(e.g. 1.0° / 0.5° = 2) 확대 배율(e.g. 2배)을 계산할 수 있었다. 렌즈의 크기의 경우에도 실험을 통해 구한 실제 정확도를 감안하여 산출했는데, 이는 크기가 지나치게 클 경우 주변부를 과다하게 가릴 수 있고, 반대로 너무 작을 경우 시선 추적 장비의 오차에 의해

표 1 적응형 대상 확대 모델의 주요 인자

인자	비고
확대 위치	시선 위치 기반으로 산정
확대 대상	전체 화면 / 시선 주변의 관심영역(*)
확대 지속 시간	항상 / 일시적(*)
확대 시작 시점	수동 / 자동(*)
확대 배율	기대 정확도와 현재 정확도를 기반으로 산정
확대 렌즈 크기	중심와 시각(foveal vision)과 기대 정확도를 고려하여 산정

사용자가 원래 보고자 했던 영역을 렌즈 안에 담을 수 없기 때문이다. 이에 따라 최종 모델에서 렌즈의 크기는 중심와 시야각(foveal visual angle)과 실제 정확도를 합산한 각도를 기준으로 정하도록 했다.

3. 모델 평가를 위한 예비 실험 및 결과

이 연구에서 제안한 시선 추적 대상의 적응형 확대 모델에 대한 평가 실험을 진행하기에 앞서 주요 인자들의 값을 결정하기 위하여 30명의 참가자(남성 20명, 여성 10명, 평균 연령 25.4세)를 모집하여 2차례의 예비 실험을 진행했다. 실험에는 2대의 Tobii x60 시선 추적 장비가 사용되었으며, 2대의 1920 x 1080 픽셀 해상도의 24인치 모니터를 사용하여 동시에 두 명이 실험할 수 있는 실험 환경을 구성했다. 실험 참가자들은 모니터로부터 60cm 떨어진 곳에 턱받침 없이 앉아서 15개의 보정 지점을 사용하여 보정 과정을 거쳤다. 이후 실험 참가자들은 20군데의 무작위 지점에 차례로 하나씩 나타나는 38픽셀 크기(≈1°의 시야각)의 원을 응시했고, 그 결과 약 1.1°의 시선 추적 정확도를 얻을 수 있었다(SD=0.6°).

두 번째 예비 실험에서는 3명의 참가자(남성 2명, 여성 1명, 평균 연령 30.7세)를 모집하여, 앞선 예비 실험에서의 결과를 활용해서 확대 모델에 포함된 각 인자들의 값을 구하고자 했다. 이 실험에서도 첫 번째 실험에 사용된 것과 같은 환경(사용 장비, 모니터와의 거리, 조도 등)을 구성했으며, 실험에 앞서 9개의 보정 지점을 사용하여 보정 과정을 거쳤다. 이후 본 실험에서 참가자들은 화면 상에 나타난 1부터 9까지의 숫자를 오름차순으로 눈으로 따라가며 읽었다. 각 숫자들은 19픽셀 크기(≈0.5°의 시야각)로 표시되었고, 숫자들 중 일부는 몰려 있는 상태로 표시되었다(그림 2A). 이 실험에서는 항상 확대된 화면을 표시하는 조건과 일시적으로 표시하는 조건 2가지를 모두 사용했으며, 앞서 진행되었던 첫 번째 예비 실험결과로 얻은 실제 정확도를 0.5°로 개선하기 위하여 확대 배율을 2배로 설정했다. 그리고 렌즈의 크기는 중심와 시야각과 실제

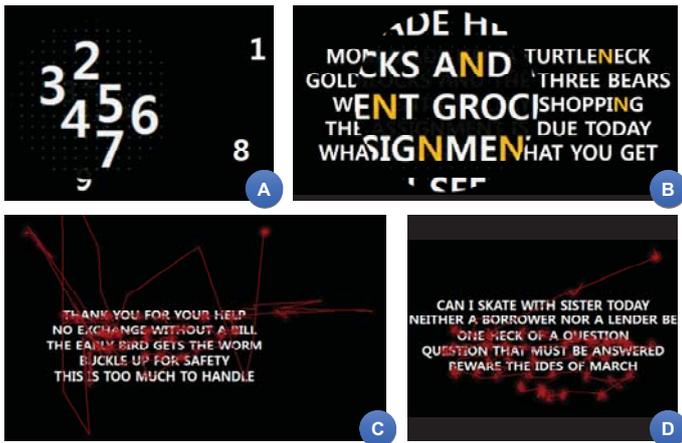


그림 2 세 가지의 수행 작업 조건(A, B, C)과 시선 추적 결과 (C, D)

정확도를 반영하여 97픽셀(≈2.55°의 시야각) 크기로 설정했다. 그 결과 아래와 같은 문제점을 찾을 수 있었다.

- 시선 추적 정확도가 낮을 경우 확대 렌즈가 잘못된 위치에 표시되어 사용자가 맥락을 잃는 경우가 나타났다.
- 일시적으로 확대가 되는 경우가 항상 확대되는 조건에 비해 공간적인 단서의 부족으로 겪는 어려움이 크게 나타났다.
- 확대 렌즈 주변의 테두리 선으로 인해 사용자들이 마우스 커서를 조작하는 것처럼 인위적으로 다른 곳을 바라봄으로써 위치를 옮기고자 했다.

이를 보완하기 위하여 이후 진행된 실험에서는 렌즈의 크기를 약 3배(285픽셀≈7.5°의 시야각)로 늘려 주변부에 대한 공간적 단서를 제공하고자 했고, 그와 동시에 렌즈를 반투명하게 표시하면서 테두리 선 대신 기존 연구[6]에서처럼 녹색 점으로 구성된 배경 무늬(그림 2A, 2B)를 도입하여 렌즈 영역을 구별할 수 있도록 했다.

4. 사용자 평가 실험 및 분석

앞서 진행된 예비 실험의 결과를 바탕으로 개선된 모델을 평가하기 위하여 추가로 6명의 참가자(남성 5명, 여성 1명, 평균 연령 27.5세)를 모집하여 두 번째 예비 실험에서 사용되었던 동일한 장비와 실험 환경을 사용한 사용자 실험을 진행했다. 실험은 3 (확대 조건: 없음, 일시적 확대, 항상 확대) x 3 (수행 작업: 숫자를 오름차순으로 찾기, 특정 알파벳의 등장 횟수 집계, 특정 알파벳의 위치 찾기) x 2 (대상의 크기: 0.5°, 1.0°)의 형태로 설계 되었다. 참가자들은 각 작업을 수행하기에 앞서 주어진 실험 설정으로 연습할 시간을

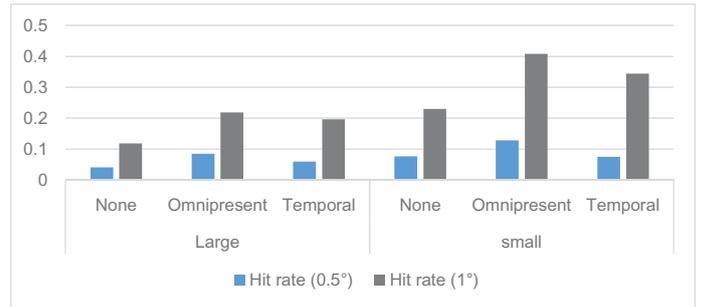


그림 3 관심 영역 적중률 (숫자 오름차순 찾기 작업)

충분히 가진 후 본 실험에 임하였다.

실험 결과는 관심 영역에 대한 적중률로 분석했는데, 그 결과 숫자를 오름차순으로 찾는 작업의 경우 그림 3에서 볼 수 있듯 종류에 관계 없이 확대가 적용되었을 때 적중률이 더 높아지는 것을 확인할 수 있었다. 하지만 일시적인 확대의 경우 예상과는 달리 항상 확대하는 조건보다 적중률이 낮은 것을 확인할 수 있었다.

5. 결론 및 향후 연구

이번 연구를 통해 카메라 시선 추적 장비의 정확도를 개선할 수 있는 가능성을 발견할 수 있었다. 향후 연구에서는 적응형 확대 모델의 각 인자들을 개선하고, 확장된 평가방안을 제안함으로써 제안된 방안이 다양한 작업들에 적용될 수 있도록 발전시키고자 한다.

참고 문헌

- [1] Adams, N., Witkowski, M., and Spence, R. The inspection of very large images by eye-gaze control. In *Proc. AVI 2008*, 111-118, 2008.
- [2] Nyström, M., Andersson, R., Holmqvist, K., and van de Wijer, J. The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods*, 45(1), 272-288, 2013.
- [3] Miniotas, D. and Špakov, O. An algorithm to counteract eye jitter in gaze-controlled interfaces. *Information Technology and Control*, 1(30), 65-58, 2004.
- [4] Ashmore, M., Duchowski, A.T., and Shoemaker, G. Efficient eye pointing with a fisheye lens. In *Proc. Graphics Interface 2005*, 203-210, 2005.
- [5] Pelz, J.B., Canosa, R., and Babcock, J. Extended tasks elicit complex eye movement patterns. In *Proc. ETRA 2000*, 37-43, 2000.
- [6] Kumar, M., Paepcke, A., and Winograd, T. EyePoint: Practical pointing and selection using gaze and keyboard. In *Proc. CHI 2007*, 421-430, 2007.