

웨어러블을 통한 시선추적 기법에 관한 연구

(A Study on Eye Tracking Techniques using Wearable Devices)

장재혁*, 정지우**, 박정훈***

(Jaehyuck Jang, Jiu Jung, Junghoon Park)

요약

시선추적 기술은 사회 다양한 분야에서 사용되고 있으며, 추적 기술의 정확성과 편리성 면에서 좋은 성능을 발휘하고 있다. 이는 터치를 하지 않고도 인터페이스(Interface)가 가능한 새로운 가능성을 보여주고 있고 할 수 있다. 이 기술은 루게릭 환자와 같이 신체의 한 부분, 한 부분이 서서히 마비되어 최종적으로는 눈을 움직일 수밖에 없는 상황의 환자에게 새로운 의사소통 방법이 될 수 있다. 이전의 루게릭 환자는 이 상황에 처하게 되면 가족 간의 의사소통은 물론, 죽음을 기다리는 것 외의 아무것도 할 수 없는 상황에 이를 수밖에 없었으나, 어렵지만 새로운 의사소통 방법인 눈을 이용한 인터페이스가 이들에게 도움이 될 수 있다. 현재로서도 시중에서는 시선추적 시스템과 전용 장비가 존재하나, 누워있는 이들에게는 사용법이 어렵거나, 개당 1200만원 이상의 높은 가격 등으로 인해 접근하기에 어려운 설정이다. 본 논문에서는 소수의 약자에게 도움을 줄 수 있음과 동시에 이 어려운 분들을 위해 저렴하게 사용할 수 있는 웨어러블 형태의 시선추적 장치를 제안하고 사회 전반의 미래 기술 성장 가능성을 극대화하기 위해 저렴하게 이용할 수 있는 시선추적 방법을 연구하여, 결론적으로 저사양/고성능의 웨어러블 디바이스 기반 시선추적 시스템을 설계하고 발전시키는 방법을 제안하고자 한다.

■ 중심어 : 안구추적 ; 시선추적 ; PCCR ; Glint ; Wearable

Abstract

The eye tracking technology is widespread all around the society, and is demonstrating great performances in both precision and convenience. Hereby we can glimpse new possibility of an interface's conduct without screen-touching. This technology can become a new way of conversation for those including but not limited to the patients suffering from Lou Gehrig's disease, who are paralyzed each part by part of the body and finally cannot help but only moving eyes. Formerly in that case, the patients were given nothing to do but waiting for the death, even being unable to communicate with their families. A new interface that harnesses eyes as a new means of communication, although it conveys great difficulty, can be helpful for them. There surely are some eye tracking systems and equipment for their exclusive uses on the market. Notwithstanding, several obstacles including the complexity of operation and their high prices of over 12 million won(\$9,300) are hindering universal supply to people and coverage for the patients. Therefore, this paper suggests wearable-type eye tracking device that can support minorities and vulnerable people and be occupied inexpensively and study eye tracking method in order to maximize the possibility of future development across the world, finally proposing the way of designing and developing a brought-down costed eye tracking system based on high-efficient wearable device.

■ keywords : eye tracking ; eye gaze tracking ; PCCR ; Glint ; Wearable

I. 서 론

영상처리 기술의 끊임없는 발달에도 불구하고, 하드웨어의 프로세서가 고사양이면 대부분

* 정회원, 아주대학교 전자공학과

** 비회원, 아주대학교 전자공학과

*** 정회원, 아주대학교 인공지능융합학과(교신저자)

접수일자 : 2023년 01월 09일

수정일자 : 1차 2023년 02월 01일, 2차 2023년 03월 03일

제재확정일 : 2023년 03월 14일

교신저자 : 박정훈 e-mail : stevejobs@ajou.ac.kr

좋은 성능을 발휘하는 반면 저사양의 소프트웨어/하드웨어로는 충분한 성능을 내기가 어려운 경우가 많다. 정보를 얻기 위해 끊임없이 움직이는 눈을 추적하는 기술은 그 사람이 가진 의도와 대상에 부여하는 중요도를 알 수 있게 해준다[1]. 따라서 이를 활용하면 AI가 사진을 이용하여 상황을 분석하고, 현장 파악을 위해 시각 정보를 선택적으로 취득하도록 학습시키는 데 필요한 데이터를 취득할 수 있다. 하지만, 정보를 받아들이는 눈을, 제어를 할 수 있는 컨트롤러로 사용하는 접근은 다소 생소한 접근일 수 있고, 사용자로 하여금 편한 손이나 말로써 할 수 있는 일을 왜 눈으로 하는지에 대한 의문이 있을 수 있다.

루게릭 병은 운동신경세포만 선택적으로 사멸하는 질환으로 증상은 서서히 진행되는 사지의 위약(weakness, 쇠약) 및 위축으로 시작하고, 병이 진행되면서 결국 호흡근 마비로 수년 내에 사망에 이르게 되는 치명적인 질환이다. 특히 혀 근육이 부분적으로 수축하여 식사할 때 사례가 들리거나 기침을 하고, 흡인성 폐렴(음식물 등이 식도로 넘어가지 않고 기도로 잘못 흡인되어 야기되는 폐렴) 등이 발생하기 쉬우며, 운동신경세포가 파괴되면서 뇌줄기, 목, 가슴, 허리영치분절이 담당하는 얼굴, 몸통, 사지에 위운동신경세포 손상 증상과 정후(위운동신경세포증후군)가 나타난다. 호흡곤란 또한 가로막과 갈비 사이 근육의 위약으로 발생할 수 있는데, 특히 가로막이 약하면 누워있을 때 복강 장기의 흉강 압박을 막지 못해 호흡곤란이 심해질 수 있다. 인지기능의 장애도 일부에서 동반된다고 알려져 있다[2].

혀까지 마비되어 말도 할 수 없고, 손까지 마비되어 손을 쓸 수 없는 루게릭 환자에게는 오직 눈만이 의사를 표현할 수 있는 방법이며, 의사와 간호사, 가족과는 노트에 글을 써서 의사를 전달하면 환자는 눈을 깜박인다든지, 눈을 응시하는 방법 등을 통해 최소한의 필요 대

화를 해왔다. 눈을 들여다 볼 수 있는 카메라와 얻어진 이미지 기반으로 시선을 추적하는 것이 가능하고, 이 시선추적 기술과 GUI(Graphic User Interface)는 팔다리를 사용함 없이 컴퓨터를 사용할 수 있게 해 줌으로써 루게릭 환자가 컴퓨터만 사용할 수 있는 지식만 있다면 컴퓨터를 자유자재로 사용할 수 있게 하여 디지털 디바이스를 조작하게 함으로써 디지털 접근권을 향상하고 행동 패턴과 사용 습관을 분석해 사용자 편의 및 자동화를 위한 데이터 구축에 큰 기여를 해 왔다[3]. 이 시선 추적 기술을 입력 장치에 활용할 경우 파킨슨·루게릭병 환자가 PC나 스마트폰을 잘 사용할 수 있도록 하여 다양한 방도로 사용하는 데 기여한다. 하지만 시선추적기술을 단순히 컴퓨터와 연결한다면, 루게릭 환자가 할 수 있는 일은 별로 없다. 단지 시선추적을 하여 좌표를 알려주는 것만으로는 아무것도 할 수 없기 때문이다. 현재 사용할 수 있는 방식은 의사와 간호사가 만든 알파벳 책을 눈 깜빡임을 활용하여 알파벳을 선택하거나 동의 여부를 묻는 수준의 의사소통 정도이다. 이러한 방식은 오랜 시간이 걸릴 뿐 아니라 다양한 표현을 전달하지 못하고, 무엇보다 환자 본인이 주체적으로 행동하지 못한다는 한계를 갖는다.

시선추적 기술이 보편화된다면 필요할 때만 과도한 태양광을 가려 주는 부분 선풍 기술 적용, 사물인터넷의 사용자 친화적 UX 설계 등 그동안 할 수 없었던 일들에 하나의 가능성으로 다가갈 수 있고, 여러 문제를 해결함과 동시에 눈을 새로운 인터페이스의 하나로 활용할 수 있을 것이다. 발전된 기술을 적용하여 환자나 장애인의 물리적 한계를 극복하고, 소통과 생활을 원활히 하는 데에 시선추적이 갖는 잠재력은 차례로 증명되고 있으며[4], 야구 batting과 온라인 학습 보조 시스템 등에서도 시선추적의 효용성이 지속해서 인정받으며 상용 기술에 대한 필요성은 점점 높아지는 추세

로 보이고 있다[5]. 최종적으로는 디지털 디바이스 사용 전반의 편리와 빅데이터 시대에 있어 무접점 인터페이스를 통한 자료 구축에 기여할 것으로 기대한다.

II. 본 론

마우스와 키보드는 대표적인 컴퓨터의 입력 장치로써, 많은 사람들에게 없어서는 안 될 입력인터페이스로 여겨져 왔다. 그 이유는 그 사용에 있어서 마우스의 섬세함과 키보드의 다양한 버튼 개수를 활용해 손을 사용하여 입력하는 방식이 보편적이고, 오랫동안 인터페이스의 대표를 누려왔기 때문이다. 이에 반해, 눈을 대체적인 입력 방식으로 사용하는 것은 그 자유성과 효용성에 있어서 키보드와 마우스의 비교 상대가 될 수 없다. 하지만, 루게릭 환자에게 있어서 마지막 남은 눈은 입력 도구인 동시에, 의사 표현의 도구인 제어기로 사용하는 것은 선택이 아니라 반드시 사용할 수밖에 없는 사용성으로 간주된다. 이 마우스와 키보드의 자유도를 눈을 이용한 입력 도구로 구현하는 것은 컴퓨터의 기본 기능도 잘 사용하지 못하는 환자에게는 시선으로만 입력 작업을 해결해야 하므로 몹시 가혹한 방법이다. 이를 극복하기 위해서는 컴퓨터 화면에 띄워지는 GUI가 편의적 기능을 사용할 수 있게 함과 동시에 직관적으로 사용자의 편리성을 최대화하는 방법으로 사용될 수 있다. 그리하여, 본 논문에서는 시선을 추적하는 방법과 시선이 모이는 점, 그리고, 그에 의해 동작을 정의할 수 있는 GUI의 연결 시스템으로, 저가에 만들 수 있지만 루게릭 환자에게 편리함을 주는 시선추적기반 대화 시스템을 제안하려 한다.

우선, 시선추적에는 여러 가지 방법이 존재한다. 피부의 변화를 기계적으로 감지하거나 눈의 움직임 자체를 포착할 수 있도록 코일이 삽입된 콘택트렌즈 등의 장비를 부착함으로써 기계적으로 정보를 획득하는 방법[6], 영상적으로

iris를 분석해내어 시선을 추적하는 방법[7]. sclera에서 반사된 적외선의 강도를 확인함으로써 눈의 위치를 파악하는 방법, 각막반사를 비롯한 눈의 특징을 이용하는 방법[8], 각막의 특수한 반사를 이용해 동공을 추적하는 PCCR(Pupil Center Corneal Reflection) 기법 등이 있다[9]. 안경형 웨어러블 장비로 동공을 이용해 시선을 추적하는 Pupil Labs의 eye tracker는 암동공 기반의 시스템을 갖추고 있으며, 사용자로 하여금 화면 상에 표시된 임의의 개수의 기준점을 바라보아 영상 정보와 마우스 좌표 간의 관계를 알아내는 캘리브레이션(calibration)과정을 거치게 한다. 또한, 거치형 장비를 사용하는 tobii의 eye tracker는 시선이 움직이는 속도를 분석해 부드러운 움직임과 단속적 운동을 분리하고, 최근 N개의 시선추적점 좌표 평균을 이용하는데, 이는 짧은 sampling 주기로 인해 안구의 사소한 움직임까지 포착하는 데에서 오는 noise를 보정하기 위함으로 보인다[10]. 이 장비는 45~95cm 거리에서 40도의 FoV(Field of View)를 지원하지만, 64bit Windows 10 이상의 운영체제와 6세대 이후 Intel Core, 최소 2GHz 8GB RAM pc를 요구하는 단점이 있다. 이에 비견하여 본 연구에서 제안하는 알고리즘은 라즈베리파이 4 보드(리눅스 기반의 1.5Hz, 2GB RAM) 사용 시, 비최적화 상태(Thonny프로그램)기반에서 작동되더라도 매끄럽게 구동할 수 있었음을 고려하면 두 시스템 사이의 사양 차이가 확인 가능하다. 여기에, 시선추적 기법 중 하나인 PCCR기법이 더해져, 계산량을 줄일 수 있었다고 볼 수 있다. PCCR(Pupil Center Cornea Reflection)은 명동공과 암동공이 차분된 영상을 활용하여 시선추적점을 찾는 방식으로, 영상을 보지 않고 기계적으로 차분하는 방식이 계산 결과가 의도한 바와는 다른 결과를 낳을 때가 많은 바에 비해, 조도와의 관계를 통해 들어오는 이미지를 다 계산할 필요 없이, 깨끗한 이미지만을

얻을 수 있다는 장점이 있다. 깨끗한 이미지를 얻지 못하는 경우에는 시선추적점을 계산하고 서도 결과물을 사용할 수 없는 경우에 처하게 된다. 이 모든 계산은 깨끗한 영상을 항상 얻는다는 가정상에서 이루어지는 경우가 대부분이기 때문에, 실험실이 아닌 실제 동작을 해내어야 하는 다양한 환경에서 역시 원하는 결과를 얻기 위해서는 반복해서 실험해야 하는 경우가 대부분이다. 좋은 이미지와 나쁜 이미지의 판단을 1초에 30프레임을 판단하기는 불가능에 가깝다고 할 수 있으므로, 다시 말해 조도와 이미지와의 관계는 끊으려야 끊을 수 없는 밀접한 관계로 이루어져 있다. PCCR의 경우에 깨끗한 이미지를 얻지 못하는 결과는 명동공이 안정적으로 검출되지 않아 가공된 프레임이 깔끔한 타원을 나타내지 않는 결과로 이어질 수 있어서 많은 주의가 요구된다. 안경과 같은 웨어러블形 디바이스를 이용하면 사용자의 눈에서 정보를 얻을 수 있도록 장치가 고안되어 이미지를 획득하므로 얼굴 전체나 주위 배경이 포함될 때보다 얼굴을 인식하는 알고리즘이 빠질 수 있어 훨씬 가벼운 알고리즘으로 동공 추적을 할 수 있게 할 수 있다[11]. 이렇게 하드웨어와 소프트웨어를 목적에 맞게 수많은 실험을 통해 프로토타입을 제작하였고, 다양한 시도들로 구현된 제안시스템은 다음과 같다.

1. 제안된 시스템

그림 1은 본 논문에서 제시하는 시선추적 시스템을 도식화한 것이다. PCCR을 적용하기 위해 bright pupil, dark pupil을 번갈아 촬영하고, 1초동안 30프레임 이상 들어올 수 있는 이미지 중 깨끗한 이미지상의 계산을 위해 프레임그랩(Frame grab)과 LED(Light Emitting Diode) 스위칭 순서를 주의하여 통제한다.

이렇게 얻은 두 장의 프레임을 차분하여 필터링을 통해 동공과 반사광의 이미지를 얻고,

각각의 좌표를 알고리즘에 의해 산출한다. 검출된 글린트(glint)를 바탕으로 동공(pupil) 탐색에 보정을 가해 사용자의 시선이 향하는 곳을 분석하고, 캘리브레이션을 통해 사용자의 각자 안구 크기에 따른 시선추적위치를 보정한다. 이처럼 시선에 따라 움직이는 마우스 좌표 끝점을 보완하면, 움직이는 좌표가 항상 시선이 움직이는 단점을 중심으로 하는 일정 반원 내에 있도록 보정하여 마우스 좌표를 계산, 화면에 표시할 수 있다. 사용자마다 안구의 지름이 다르므로 사용자별 시선추적 계산과정을 보정하기 위해 최초 16개 기준점을 이용해 캘리브레이션을 진행하고, 이후 이 정보를 토대로 시선추적 점이 안경의 위치 틀어짐으로 좌표오차가 발생했을 경우, 더 적은 캘리브레이션 포인트로 설정값을 다시 얻을 수 있다. 본 논문에서는, 최초 1회 동공으로 단정할 수 있는 타원을 엄격하게 검출한 후, 반복적으로 프로세스를 동작시켜 빠르고 정확하게 동공 좌표를 검출해 내는 것을 목표로 한다.

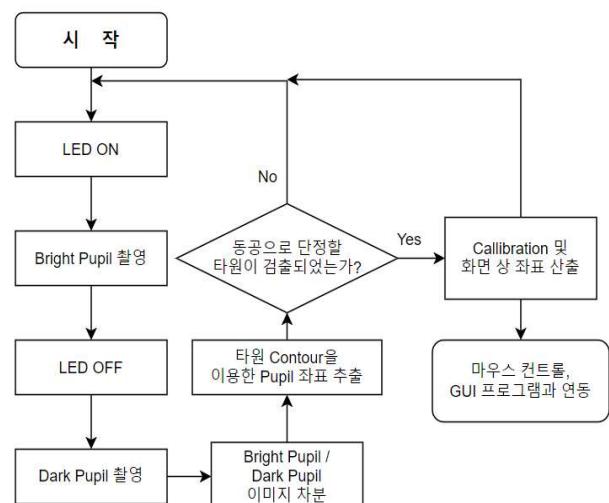


그림 1. 시선추적 시스템 흐름도

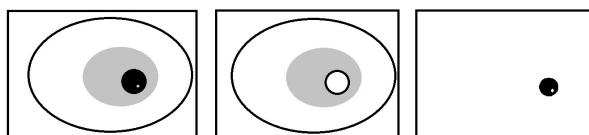


그림 2. 명암동공 및 glint 추출 이미지 모식도

가. 명암동공 및 glint 검출

그림 2는 bright pupil(가운데)과 dark pupil(왼쪽)을 차례로 촬영해 이미지를 차분(오른쪽)하는 과정을 나타낸 그림이다. 암동공 검출 시 보조 광원에 의해 발생한 glint가 함께 나타난 경우를 표시한 그림이며, 이후 차분된 이미지에서 동공과 glint를 소프트웨어적으로 추출한다. 명동공을 검출하기 위해서 카메라와 LED의 광축은 반드시 평행에 가깝게 할 필요가 있고, 이를 위해 안경형 장비에 이 둘의 평행을 반드시 확보해야 하는 형태가 요구된다. 명동공은 매우 제한적인 조건에서만 관찰되므로 조사각과 촬영각이 일치함을 섬세하게 확인하고 LED의 개수를 늘려 항상 깨끗한 이미지를 얻을 수 있는 환경을 확보해야 한다[12].

LED의 상태에 무관하게 랜덤하게 캡처된 이미지를 분석해 원하는 좌표를 알아내는 방법도 가능하지만 이 경우에는 많은 연산을 요구한다. 명동공과 암동공의 이미지를 차례로 획득하기 위해서는 프레임을 얻는 주체가 조도 확보를 위한 LED 회로 스위칭 상태를 반드시 알 필요가 있다.

또 하나의 방법은 회로가 ON 상태에 접어든 후 명동공이 검출되는 데 걸리는 시간과 반대로 암동공에 걸리는 시간을 미리 측정하는 것으로, 두 국소 시스템을 최초 실행 시 동기화하여 디지털 회로를 정확하게 통제하고 PC 카메라의 Vsync 신호를 이용하여 깨끗한 프레임을 선별해 이용한다. 또 하나의 다른 방법은 프레임그랩의 주체가 LED 회로를 통제하는 것이다. Arduino, Raspberry pi와 같은 하드웨어 저사양 보드를 이용하면 PC에서 구현된 프로그램으로 제작하드웨어와 통신하여 웹캠 촬

영 타이밍을 지시하거나 자체 연결된 IR 카메라로 적절한 순간에 이미지를 획득할 수 있다 [13].

(1) 회로 구성과 자체 발생 sync 기반 스위칭

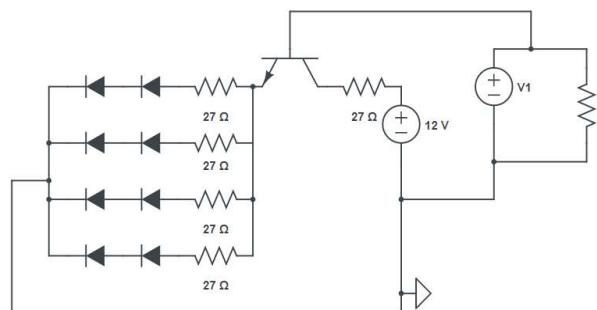


그림 3. 자체 sync발생 스위칭 회로도

그림3은 12V DC 전압원에 의한 power를 라즈베리파이 출력 전압(V1)으로 통제하는 회로를 간략하게 나타낸 것이다. 2SC1815 NPN Transistor를 활용하여, 구현한 회로로, 소형 보드를 사용하는 경우 3.3V GPIO Pin과 5V DC Pin이 있는데 전자는 직렬 연결된 다이오드에 2.6V의 충분한 전압을 인가하기 위한 V_{BE} 를 확보하기 어려우므로 GPIO 신호로 먼저 5V signal을 통제하고 이것을 High, 0V 상태 신호를 Low로 생각하여 설계한다. 기초 입력 High 신호가 높아질수록 독립 전류 소스로 동작하는 트랜지스터가 특정 전류(100mA)에 도달하는 시간이 줄어들지만, 부하 전압이 과도하게 증가해 소자 고장의 원인이 되므로 $V_B = 4.4V$ 일 때 load 전류가 충분한 안정성을 가짐을 확인하였고, 스위칭 후 다이오드가 충분히 켜지거나 꺼질 때, 즉 $2V_{LED}$ 가 각각 2.5V, 2.1V가 될 때까지 걸리는 시간은 각각 30ms, 20ms임을 또한 확인하였다.

(2) 하드웨어 구성

카메라 주위에 LED를 나란히 부착해 각막에

서 반사된 빛이 렌즈를 통해 평행하게 들어갈 수 있도록 구현하였고, 눈부심을 방지하고 반사된 빛을 카메라로 잘 관찰할 수 있도록 방사각이 76°인 870nm 파장의 고속 IRLED와 NOIR Camera를 이용하였다. 카메라 연결, sync 발생, 소프트웨어 처리를 위해 raspberry pi4B 제품을 채택하였으며, 별도의 보드에서 동작하기 때문에, PC에서 보다는 훨씬 부하가



그림 4. 다이오드의 광축과 평행한 IR 카메라

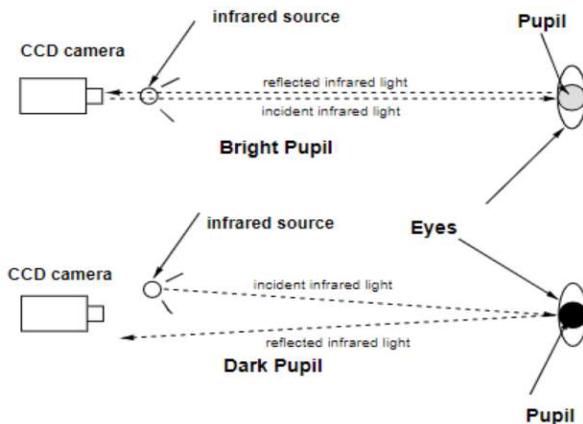


그림 5. 명동공과 암동공 효과의 원리[14]

작게 걸리는 알고리즘으로 동작한다는 것을 알 수 있다. 웨어러블 디바이스에 카메라와 LED를 연결하여, 명동공을 검출하기 위한 이 두 부품은 광축이 평행하게 배치하고 필요에 따라 주위에 glint 검출용 다이오드를 최소 11cm이상 떨어뜨려 카메라로 촬영한 이미지를 기반으로 동공 정보를 분석하고 여기서 획득한 좌표를 데이터로 활용하거나 마우스 조작에 사용할 수 있다. 본 연구와 같은 방식의 명동공

검출은 주변 환경의 밝기, LED의 조도, 카메라와 LED의 광축이 이루는 각도, 사용자의 눈과 광원의 거리 등 여러 요인에 의해 결정되므로, 각 부품의 SPEC을 잘 살펴보고 결정해야 한다. 이러한 조건이 맞지 않는 경우 대부분 홍채보다 동공이 어둡게 관측되어 암동공의 영상을 획득할 수 있고, 이때 각막에서 반사된 빛은 작은 점(glint)으로 나타난다[15].

(3) LED를 이용한 명암동공 및 glint 검출 조건

글린트를 검출하기 위해, 모든 실험은 내부 조명이 켜지지 않은 그늘진 환경에서 8개의 LED를 이용해 진행되었다. 명동공을 깨끗하게 검출하기 위해서는 충분한 광량이 필요하므로 광원이 밝고 주위 환경이 비교적 어두울수록 좋으며, 피사체에 직사광선이 닿을 경우 동공 주변의 색 대비가 혼란스러울 수 있고, 특히 안구의 특정 지점에서 집중적으로 반사광이 관찰되면 명동공, 암동공을 검출하는 것이 둘 다 어려워지는 경우가 있다. 기성 시선 추적 장비와 소프트웨어가 실내 사용을 권장하거나 강제하는 것은 이러한 이유에 기인하며, 제안하는 모델은 회로의 출력을 높이거나 주위 조도 분석을 통해 영상처리 알고리즘을 고도화하기 때문에 기존의 문제를 해결할 수 있다. 동공 주변에서 최대한 많은 빛이 반사되어야 명동공을 검출할 수 있으므로 복수의 다이오드를 사용했고, 방사각이 넓은 소자를 선택한 것은 출력된 빛이 한 지점에서 중첩되기 쉽게 하기 위함이다. 방사각이 26도, 54도인 LED를 이용해 같은 조건에서 실험을 진행했을 때는 5cm~50cm 범위의 안구-카메라 거리에서 뚜렷하거나 흐릿한 명동공을 관찰할 수 없어 보다 넓은 각도의 제품을 채택하였다. 조사각을 정확하게 맞추어 각각의 LED에서 나온 빛이 각막에서 반사되어 카메라에 곧이 흡수되게 할수록 다음의 세 가지 이점을 가진다. 첫째, 다이오드와 안구의 거리를 좁힐 수 있다. 거리가 멀어질수록 같은

방사각에서 더 넓은 조사 범위를 가지므로 파동을 중첩하기는 쉬우나 그만큼 단위 면적당 에너지는 감소할 수밖에 없다. 요구 거리가 짧아지면 하드웨어 구성의 어려움이 대폭 감소하고 무엇보다 안경형 디바이스의 특성 상 발생하는 돌림힘으로 인한 체감 무게 문제를 해결 할 수 있다. 둘째, 더 작은 방사각을 갖는 소자를 사용할 수 있다. 거리가 일정하여도 방사각이 작을수록 중앙의 빛이 강해진다.

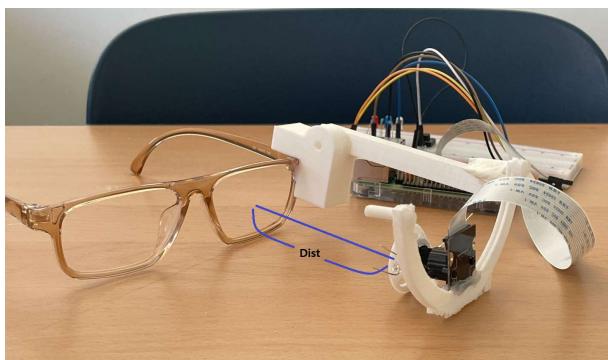


그림 6. 장비 구성의 예



그림 7. 명동공 검출의 예(안구-카메라 30CM)

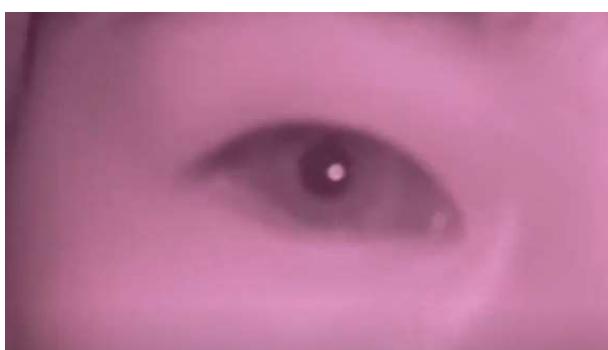


그림 8. Glint 검출의 예

셋째, 앞의 두 이점은 회로 출력이 더 높아진 경우와 같이 생각할 수 있다. 더 적은 수의 LED를 사용하거나, 고출력의 이점을 살려 장

비의 실외 착용 등 안정성, 활용성을 높일 수 있다.

안구-카메라 거리가 30cm일 때, 그럼 7과 같이 깔끔한 명동공이 관찰되었다. 25-27cm에서는 같은 범위에서 20%가량 다소 흐린 색상으로 관찰되었으며, 23cm 거리에서는 일부 범위에서만 명동공을 볼 수 있었다. LED가 20cm 이하로 다가갔을 때는 오직 암동공의 이미지만을 얻을 수 있었다. 세 명의 피실험자를 대상으로 2~3회씩 반복 측정했을 때 동일한 결과

표1: Iris 크기에 따른 glint 안정검출 요구 거리(단위 : cm)

Iris 지름	1.0	1.2	1.3
요구 거리	21	25	>25

과를 얻었으며, 실험실의 밝기를 크게 조절했을 때 검출 여부나 거리가 유의미한 변화를 보여 사용자보다는 사용 환경이 중요한 요인임을 확인하였다.

명동공에 비해 광량이 적거나 빛이 평행하게 반사되지 않을 때 발생하는 glint는 비교적 각도, 거리 조건에서 자유롭다. 본 연구에서는 명동공을 추출한 거리와 같은 25cm에서 주위에 고정대를 부착해 glint 검출용 추가 LED를 사용하는 시스템에 관심을 두어 실험하였고, 서로 다른 Iris 지름을 갖는 세 명의 피실험자를 대상으로 10회 이상씩 반복 실험하여 위와 같은 결과를 얻었다. 동공 검출 시와 마찬가지로 그늘진 실내 환경을 기준으로, 2개 이상의 LED가 있을 때 원하는 영상을 얻을 수 있었고 이때 크기와 화면상에서의 밝기는 조사각, 소자의 개수에 거의 무관하다. 밝은 자연광이 bright pupil 검출에 상당히 불리한 요소로 작용하기 때문에, 모든 실험은 실내에서 진행되었다. 안정적인 알고리즘을 위해 복수의 glint를 활용할 수 있는데, 이때 각 광원의 거리가 9cm일 때부터 두 점이 갈라져 보이기 시작하고 11cm 이상에서 완전히 분리된다. 60cm 거리에서 21inch 모니터를 바라볼 때 pupil과 iris이 보통 크기(약 12mm)인 경우, 내부에서 항상 glint가 검출되나 사용자에 따라 iris 내에

상이 맷하게 하기 위해서는 더 가까이에서 조사해야 하는 경우가 있을 수 있다. Pupil 크기는 실험자마다 다양하였으며 각각은 주위 밝기에 따라 지름이 2배 이상까지 변화하였으나 이것을 통제변인으로 지정하지 않아도 실제 작동을 하는 데에 장애 요소가 되지 않았다. 이때 테스터군은 전부 흑인이었다.

이러한 특징은 수직거리가 일정하여 반사광용 광원이 화면과 같은 평면상에 있을 때 수평거리에 독립적으로 나타나서 두 광원의 중심이 카메라로부터 최대 30cm 떨어져 있을 때까지 영상에서 유의미한 차이를 발견할 수 없었다.

나. 동공 추적 알고리즘

명동공이 검출될 때, 암동공에 비해 동공이 압도적으로 밝게 나타나 명도의 차이가 크다. 반면 iris, sclera 등의 주변 부위 및 사물은 오직 LED를 조명으로 사용한 정도만큼의 밝기 차밖에 보이지 않는다. 두 이미지를 차분하면 pupil을 제외한 대상은 매우 어둡게 나타나므로, 보정을 가한 후 threshold를 적용해 pupil 만이 나타난 상태에 가까운 이미지를 얻을 수 있다[15]. 정확한 시선 추적을 위해서는 카메라의 nodal point와 광원의 위치는 물론 cornea의 곡률을 비롯하여 고려해야 할 수많은 요소가 있으나 사용자의 안면에 대해 가깝게 고정된 장비를 사용한다면 이 중 대부분을 상수로 생각할 수 있고, 때문에 빛의 실제 반사 지점 등으로 계산되는 eye parameter를 고려할 필요가 없는 다중 카메라 시스템의 이점과 비슷한 효과를 기대할 수 있다[16]. 본 연구에서는 영상처리를 위해 python의 opencv 라이브러리를 활용하여 구현하였다. 이 이미지를 필터링하여 cv2.findContours 함수를 이용해 contour를 그리고 cv2.fitEllipse 함수로 ellipse 모양을 찾아낼 수 있다.

그림 9에서 사용한 함수 filter(img_new)는

openCV의 alpha, beta, gamma 필터링을 나타내며, 각 값은 이미지의 어두운 부분과 밝은 부분을 지정해 각각의 밝기 평균을 구하고, gamma 값을 조금씩 변화시키면서 gamma filtering한 이미지의 동일 부분 밝기 평균이 일정 값 이하일 때를 찾아서 설정한다.

이때 alpha, beta는 고정된 gamma값에 의해 다음과 같이 자동으로 정의된다. lookuptable은 gamma값 필터링을 위해 계산 결과를 저장하는 리스트이다.

```
img_d = 암동공 프레임  
img_b = 명동공 프레임  
img_new = (img_d - img_b)
```

```
img_new = filter(img_new)  
ellipse1, ellipse2, ellipse3, ... = getContour(img_new)  
  
if ellipse1 = is_good_ellipse(ellipse1)  
    then pupil = ellipse1  
if ellipse2 = is_good_ellipse(ellipse2)  
    then pupil = ellipse2  
...  
...
```

그림 9. 동공 추적 알고리즘 의사코드

```
if abs(area - area_cri) / area_cri <= 0.25  
and abs(ratio - ratio_cri) / ratio_cri <= 0.25  
and abs(ellipse_x - pre_x) < 80  
and abs(ellipse_y - pre_y) < 80  
then, return good
```

```
else, return bad
```

그림 10. ellipse 판별 의사코드. [17]의 접근법 참고[18]

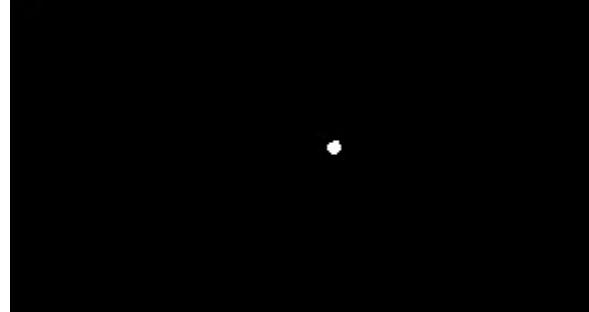


그림 11. 필터링된 Glint 이미지

```

black_max, white_min = mean(LUT(black,
lookuptable) , mean(LUT(white, lookuptable)
diff, sum = white_min - black_max, white
min + black_max

```

```

alpha = ceil(255/diff)*3
beta = (255-alpha*sum)/2

```

각각의 ellipse는 장/단축의 길이, 중심점, 회전 각 정보를 포함하므로 넓이와 위치를 이용해 동공에 해당하는 타원을 선별할 수 있다. Pupil 후보를 선별한 후 기존 획득한 결과와 위치나 크기가 비슷한 것을 새로운 frame의 pupil로 인식한다. 그림 10은 기존 동공 크기에서 추출한 표준크기 area_cri와 표준 장단반경비 ratio_cri, 그리고 장단반경의 각 길이를 이용해 동공 후보를 판별하는 코드이다. 정면에서 동공은 원에 가깝게 보이므로 최초 (1/1.3, 1.3) 반경 비 범위 내에서 동공을 탐색하고, 다양한 사용자 조건을 고려하여 범위를 (1/3, 3) 까지 늘린다. 이후에는 최신 몇 개의 장단경비 평균을 ratio_cri로 정하고 그것의 (1/1.5, 1.5)배의 ratio를 갖는 타원만을 동공 후보로 선별한다. area_cri 또한 마찬가지로 적용한다.

이러한 방식을 사용하면 최초 동공의 위치를 매우 까다롭게 요구해도, 검출 안정성을 높이면서도 반복적으로 동공을 찾지 못하는 불편함을 만들지 않을 수 있다. 특히 실내 환경에서 이 알고리즘을 적용했을 때 잘못된 타원을 동공으로 인식하는 frame은 거의 발생하지 않았다. 동공을 찾지 못했을 경우 눈을 감은 것으로 추정하고, ‘눈감음’ 신호를 발생시킨다. 캘리브레이션을 거친 좌표와 이러한 특수 입력을 소프트웨어를 통해 pc 조작에 활용할 수 있게 한다. ‘눈감음’ 신호는 HID 마우스의 물리 버튼 역할에 해당하며, GUI 등으로 그 기능을 클릭뿐 아니라 영역 설정, 스크롤 등으로 폭넓게 이용한다. 이렇게 얻은 좌표는 동공이 떨림에 따라 함께 흔들린다. 따라서 다음과 같은

방식으로 값을 일부 배제/보정하여 사용한다.

최근 10개 frame에서 검출한 동공 중심 좌표를 저장하여 직전 frame의 좌표에서 다음 좌표를 향하는 vector 9개를 생성한다. 내적값이 일정 값 미만일 때 terminal을 배제하여 나머지 값들에 대해 가중치를 적용한 평균 산출법을 이용한다. 내적을 사용하는 것은 시선의 흐름에 부자연스럽게 튕는 방향으로 오류가 발생하는 현상을 막기 위함이다.

하드웨어 자체의 흔들림, 틀어짐에 의한 오차를 배제하기 위해 글린트를 활용한다. 동공 또는 홍채(iris)위에서 반사광이 생기면 주변과

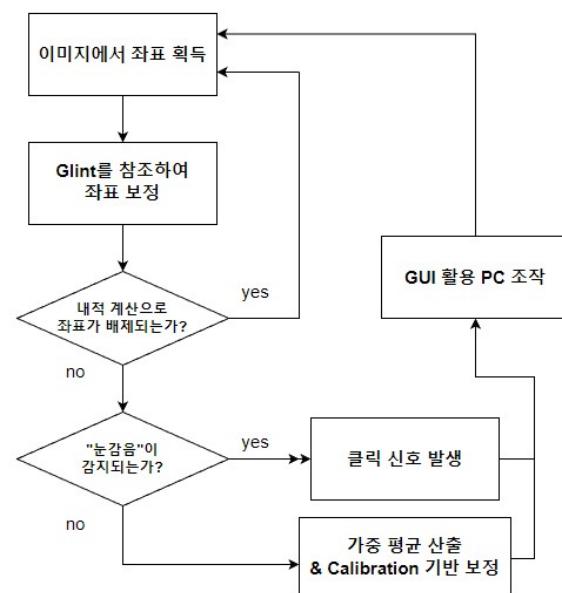


그림 12. 좌표 필터링 및 활용 순서도

밝기 대비가 매우 크고, 0~255의 밝기 범위에서 230 이상의 압도적인 값을 가지므로, 간단한 필터링만으로 그림11과 같은 결과를 얻을 수 있다. 이것은 고정된 점이므로 하나의 glint로 피사체의 평행이동을, 둘 이상으로 회전을 감지하고 바로잡을 수 있다. 또한, 캘리브레이션의 기준점이 많을수록 정확한 결과를 낼 수 있으므로, 간단한 모델은 다음과 같이 scalar field를 이중 생성하여 구현할 수 있다. 화면을 9개의 합동 직사각형으로 분할하고 16개의 꼭짓점을 기준으로 삼아, 화면에서 각 row의 네

점의 x좌표를 입력, 이 입력을 만들어낸 영상 내 동공 x좌표를 출력으로 생각한다. 입력과 출력을 반대로 놓는 것은 통상 화면의 해상도가 더 높아서 일대일대응 함수를 만드는 데 유리하기 때문이며, 이때 네 좌표 쌍을 대입하여 상용하는 삼차함수를 만들고, 다른 row에 대해서도 마찬가지로 각각 4개의 계수를 계산한다.

III. 결 론

본 논문에서는 안경형 장비에 카메라와 LED를 결합해 명암동공을 검출하고, 이를 통해 동공의 좌표를 계산해 사용자의 시선을 추적하는 방법을 다루었다. Raspberry pi 4를 이용하여 400*300 resolution의 차분되지 않은 이미지를, 보정 상수를 조정하여 제시한 알고리즘으로 처리했을 때 frame당 평균 150ms가 소모되고 같은 해상도의 두 이미지를 차분하는 데 평균 110ms가 소요되어 최종적으로 300ms이내에 동공 인식이 가능할 것으로 보인다. 모든 프로세스는 Thonny 환경에서 이루어졌으며, 제시된 소요 시간은 1천 개 이상의 샘플에서 계산된 것으로 99% 신뢰 범위 내에서 최대 172ms, 126ms 소요되어 안정적인 응답을 보였다. 또한 화면을 바라볼 때 동공 가동 범위 내 중심 좌표가 300픽셀 내에서 변화하며 캘리브레이션 테스트 시 사용자가 바라보는 아이콘을 특정할 수 있었으므로 실사용에 충분한 정밀도를 가졌다고 볼 수 있다. 실험적으로 장비를 구성한 가격은 안경테 1만 원, 지지대 3만 원, 라즈베리파이 20만 원 등을 합산하여 총 25만 원 미만으로 추가적인 pc의 구매가 불필요하다. 따라서 상용화될 경우, pc 불포함 35만 원가량에 판매되는 tobii 사의 eye tracker 등에 비해 월등한 가격 경쟁력을 가질 것으로 기대한다. 시선추적 장비는 질환으로 인해 사지가 자유롭지 않은 환자에게는 유일한 소통의 창구가 되며 사회 전반에서 통용될 때는 수많은 사용자의 데이터 축적을 통해 다양한 산업에서 활용될

수 있다. 이에 따라 저가의 시선추적 장치로서의 가능성을 보였고, 조금 더 섬세한 상용시스템에 상용하는 정확도를 위해서는 완전히 동일한 위치에서 촬영된 두 장의 명암동공 기반 차분 영상이 필요하고, 효과적인 이미지만을 전송하기 위한 IR LED의 최적 개수, 효과적인 명동공 추출 거리와의 관계 연구 및 사용자의 시선을 방해하지 않는 장비 구성을 위하여 15도가량 빗겨 led를 조사하고 안구를 촬영하는 방식에 대한 검토가 추가로 필요하다.

REFERENCES

- [1] Kiwon Yun, "Studying Relationships between Human Gaze, Description, and Computer Vision," *IEEE*, pp. 739-741, Portland, USA, 23-28 Jun. 2013.
- [2] 서울대학교병원 의학정보
<http://www.snuh.org/health/nMedInfo/nView.do?category=DIS&medid=AA000109>, (accessed Jul., 10, 2022).
- [3] 박정훈, 임강빈, "실시간시선추적기반스마트의료기 기고찰," *스마트미디어저널*, vol.10, no.1, pp. 9-13, 2021
- [4] Zhenhao Ji, "PCCR Based Wheelchair Control System," *IEEE Circuits and Systems Magazine*, vol.21, no.3, pp. 80-82, 13 Aug. 2021.
- [5] Satori Hachisuka, "Confirmation of the Significance of Facial Images in Online Learning Using Eye Gaze Tracking Measurement," *Springer*, pp. 595-599, 2021.
- [6] Andrew Duchowski, "Eye Tracking Techniques," *Springer*, pp. 1-32. 2007.
- [7] Kyung-Nam Kim, R. S. Ramakrishna, "Vision-Based Eye-Gaze Tracking for Human Computer Interface," *IEEE*, pp. 324-327, Tokyo, Japan, 12-15 Oct. 1999.
- [8] H.R. Chennamma, Xiaohui Yuan, "A SURVEY ON EYE-GAZE TRACKING TECHNIQUES," *IJCSE*, vol.4, no.5, pp. 388-391, Oct.-Nov. 2013.
- [9] CHI Jian-nan, "Key Techniques of Eye Gaze Tracking Based on Pupil Corneal Reflection," *IEEE*, pp. 136-137, Xiamen, China, 19-21 May. 2009.
- [10] The Tobii I-VT Fixation Filter(2012).
<http://www.vinis.co.kr> (accessed Jan., 27, 2023).
- [11] Ibrahim Shehi Shehu, "Remote Eye Gaze Tracking Research: A Comparative Evaluation on Past and Recent Progress," *MDPI*, pp.2-4, 8,

17-24, 2021.

- [12] Qiang Ji, Xiaojie Yang, "Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance," *Elsevier*, vol.8, no.5, pp. 359-364, Dec. 2002.
- [13] Aslam Forhad Symon, "Design and Development of a Smart Baby Monitoring System based on Raspberry Pi and Pi Camera," *IEEE*, 2017.
- [14] Carlos H. Morimoto, Marcio R.M. Mimica, "Eye gaze tracking techniques for interactive applications," *ELSEVIER*, pp. 118-119, Dhaka, Bangladesh, 28-30 Sep. 2005.
- [15] Zhiwei Zhu, Qiang Ji, "Robust real-time eye detection and tracking under variable lighting conditions and various face orientations," *Elsevier*, vol.98, no.1, pp. 129-132, Oct. 2004.
- [16] Elias Daniel Guestrin, "General Theory of Remote Gaze Estimation Using the Pupil Center and Corneal Reflections," *IEEE*, vol.53, no.6, pp. 1124-1131, Jun. 2006.
- [17] Dongheng Li, David Winfield, Derrick J. Parkhurst, "Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches," *IEEE*, pp. 3-6, San Diego, USA, 21-23 Sep. 2005.
- [18] Starburst at master git hub.
<https://github.com/Palpatineli/starburst/tree/master/starburst> (accessed May, 22, 2022).

저자소개



박정훈(정회원)

1996년 삼성전자 통신연구소 입사.
 2003년 연세대 전기전자공학과 석사.
 2015년 연세대 전기전자공학과 박사.
 ~2020년 삼성전자 AI센터
 2021년 아주대 인공지능융합학과 교수
 <주관심분야 : AI, 자율주행, 커넥티드카 보안 및, 비주얼 트래킹 등>



장재혁(정회원)

2022년 아주대학교 전자공학과 재학
 <주관심분야: 임베디드, 디스플레이 등>



정지우(비회원)

2022년 아주대학교 전자공학과 재학
 <주관심분야: 회로설계, 제어공학 등>