

# 웨어러블 시스템에서 눈동자의 움직임에 이용한 인터페이스

권기문\*, 이정준\*, 박강령\*\*, 김재희\*  
연세대학교 전기전자공학과, 생체인식 연구센터\*  
상명대학교 미디어학부\*\*

## Eye Gaze Interface in Wearable System

Ki Moon Kwon\*, Jeong Jun Lee\*, Kang Ryoung Park\*\*, Jaihie Kim\*  
Department of Electrical and Electronic Engineering, Yonsei University\*  
Biometrics Engineering Research Center  
Division of Media Technology, SangMyung University\*\*  
E-mail : [kimoon@yonsei.ac.kr](mailto:kimoon@yonsei.ac.kr)\*

### Abstract

This paper suggests user interface method with wearable computer by means of detecting gaze under HMD, head mounted display, environment.

System is derived as follows; firstly, calibrate a camera in HMD, which determines geometrical relationship between monitor and captured image. Second, detect the center of pupil using ellipse fitting algorithm and represent a gazing position on the computer screen. If user blinks or stares at a certain position for a while, message is sent to wearable computer.

Experimental results show ellipse fitting is robust against glint effects, and detecting error was 6.5%, and 4.25% in vertical and horizontal direction, respectively.

### I. 서론

최근에 화두가 되고 있는 것이 유비쿼터스(Ubiquitous)란 말이다. 유비쿼터스란, 어원은 라틴어로 사용자가 시간과 장소에 구애 받지 않고 자유롭게 네트워크에 접속함을 의미한다. 이런 추세에 따라 기존의 컴퓨터 시스템에서 한 걸음 더 나아가서 이동 중에도 컴퓨터 통신을 지원하는 등의 미래 지향적인 컴퓨팅 환경이 연구되고 있다. 그 중 하나가 몸에 부착해서 사용할 수 있는 웨어러블 컴퓨팅 시스템이다. 웨어러블 컴

퓨팅 및 인터페이스 시스템은 현재 산업체 및 대학에서 연구를 시작하고 진행 중에 있다. 웨어러블 컴퓨팅 시스템에서는 사용자와 컴퓨터 간의 인터페이스를 위한 방법으로 음성이나 영상을 이용하는 방법들을 많이 사용된다.

음성에 의한 명령전달이 가장 자연스럽고 이상적인 방법이지만, 소음이 존재하는 환경에서는 사용하기 어려운 문제점이 있다. 이외에도 사용자의 손에 손 및 손가락 움직임을 추적할 수 있는 센싱 장비를 부착하여 이를 이용하여 인터페이스를 하는 연구가 있으나, 이 역시 양 손으로 다른 작업을 하고 있는 환경에서는 사용하기 어렵고 손가락이나 손의 움직임에 의한 명령 모드들을 형태를 외워야 한다는 어려움이 있다[1].

이러한 문제점들을 극복하기 위하여 소형카메라가 부착된 HMD(Head Mounted Display)환경에서 눈동자 움직임의 응시 위치추적을 통한 사용자와 웨어러블 컴퓨터간의 인터페이스 방법을 제안한다.

### II. 시스템 입력으로서의 눈의 특징

#### 2.1 일반적인 눈 움직임의 속도 특성

컴퓨터 모니터상에서 사람의 눈동자 움직임 속도는  $350^\circ/\text{sec} \sim 500^\circ/\text{sec}$  정도로 매우 빠르다[2]. 이는 21 인치 모니터의 좌측 상단부터 우측 하단을 쳐다 볼

때까지 걸리는 시간이 150ms 됨을 의미한다. 반면 기계적인 마우스 장치를 이용한 커서의 이동은 사용자의 손목 관절 운동에 의해 수행되므로, 눈동자의 움직임에 의한 마우스 커서 이동보다 느린 단점을 보인다[3].

## 2.2 마이더스 현상

사용자의 응시 추적을 통한 GUI(Graphic User Interface) 버튼을 선택하기 위해서, 사용자의 단순한 버튼 응시와 선택을 위한 응시를 구별하는 과정이 선행되어야 한다. 마이더스 현상(Midas Phenomenon)은 이를 구별하지 못하여 사용자가 바라보는 모든 버튼을 구동시키는 현상을 말한다[4]. 따라서 이런 두 종류의 응시를 구별하기 위해 본 논문에서는 선택하고자 하는 버튼을 일정시간이상 응시하거나, 눈의 깜박임을 선택을 위한 사용자 응시로 정의하였다.

## 2.3 눈동자의 Jittery 움직임

눈의 Jittery 움직임이란 눈동자가 불규칙적으로 진동하듯 움직이는 현상으로 일반적으로 사람들에게는 흔히 발생하는 현상이다[3]. 이런 눈의 Jittery 움직임으로 인해 사용자가 모니터상의 특정 위치를 응시하더라도, 정확한 응시위치를 추정할 수 없다. 그러므로 아주 작은 GUI 버튼을 정확히 응시하는 것은 어렵다.

## III. 동공의 중심 추출 및 응시 위치 추적

본 절에서는 카메라에 잡힌 눈동자 영상으로부터 사용자가 응시하는 위치를 추적하기 위한 방법에 대해서 설명한다. 본 논문에서는 사용자의 응시 위치를 파악하기 위하여 동공의 중심 위치를 추출하고 추출된 중심 위치로부터 모니터상의 응시위치를 파악하는 알고리즘을 사용하였다.

### 3.1 타원 매칭을 이용한 동공 중심 위치 추출

일반적으로 사람의 동공은 원이나 타원의 모양을 가지므로 눈 영상에서 타원 매칭(Ellipse fitting)방법을 이용하여 동공중심을 추출하였다. 우선 동공은 영상 내에서 동공 외 다른 부분보다 그레이값이 낮은 특징이 있으므로 이를 이용하여 동공의 대략적인 위치를 추정

할 수 있다. 흑화소의 대부분이 동공에 의하여 발생하므로 흑화소 영역의 무게 중심이 대략적인 동공의 위치라고 볼 수 있다. 이렇게 구한 동공의 위치로부터 탐색 영역을 설정하여 탐색 영역 내에서 좀더 정확한 동공의 위치를 추출한다. 그림 1 과 같은 실험영상에서는 동공과 홍채의 경계에서 그레이값의 차이가 30 이상으로 인하여 뚜렷하게 나타난다. 그러므로 동공의 외곽 경계 위에 있는 한 점을 일단 추출하고 이 점을 기준으로 전체 외곽 경계를 추출하는 방법을 사용하였다.

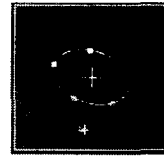


그림. 1 동공의 외곽경계와 동공 중심

이렇게 구해진 외곽 경계 중에서 5 개의 점을 균일한 간격으로 선택해서 이로부터 다음과 같은 타원의 방정식의 5 개의 계수를 구한다.

$$ax_i^2 + by_i^2 + cx_iy_i + dx_i + ey_i + 1 = 0, i = 1, 2, \dots, 5 \quad (1)$$

여기서,  $x_i$  와  $y_i$  는 가로축과 세로축의 좌표값

식 (1)에서 5 개의 파라미터를 구하기 위해서는 이론적으로 5 개의 선형 방정식이 필요하며, 본 논문에서 가우스-조르단(Gauss-Jordan) 방법을 사용하여 5 개의 계수를 구하였다[5]. 본 연구에서 눈동자 영상을 취득할 때에는 외부광의 영향을 없애고 눈동자 영상을 밝게 만들기 위하여 카메라 앞에 적외선 통과 필터를 부착하고, 적외선 조명을 사용하였다. 이때 적외선 조명에 의해 영상 내에 생기는 밝은 영역(그레이값이 220 이상)을 글린트(Glint)라고 한다.

그림 1 에서는 글린트를 제외한 순수한 동공의 외곽 경계점 5 개를 추출할 수 있다. 이와 같은 5 개의 점으로 5 개의 선형 방정식을 구하고 이로부터 가우스-조르단 방법과 식(1)을 이용하여 동공의 외곽경계를 모델링하는 타원의 방정식 및 타원의 중심 위치를 구할 수 있다.

### 3.2 응시 위치 추적

사용자의 응시위치 확인은 동공의 중심을 구하고 구한 중심에 대응하는 값을 모니터에 마우스 커서로 표시함으로써 가능해진다. 사용자가 HMD 를 통해서 모니터의 네 꼭지점과 중앙을 응시할 때 카메라에 획득되는 영상평면(Image Plane) 상에서는 위치 정보는 그림 2 와 같이 반드시 직사각형 모양이 아닐 수 있다. 이러한 이유로 본 논문에서는 아핀 변환(Affine Transform)을 이용한 영역별 매핑에 의한 방법으로 사용자의 응시 위치를 추적하였다.

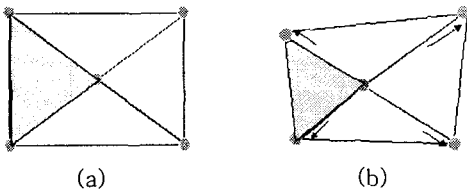


그림. 2 모니터의 네 꼭지점과 중앙(a)을 응시할 때의 동공중심의 영상평면에서의 위치 변화(b)

아핀 변환은 변환이전의 좌표와 변환 이후의 좌표 사이에 식(2)과 같은 관계식을 가진다[6].

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (2)$$

여기서,  $(x, y)^T$  변환전의 좌표

$(x', y')^T$  변환후의 좌표

아핀 변환은 점들의 이동(Translation), 회전(Rotation), 축소 및 확대(Scaling), 반사(Reflection)등을 수행할 수 있으며 아핀 변환의 계수들( $a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2$ )은 어떤 특성을 갖는 변환을 수행할지를 결정해 준다.

아핀 변환을 위한 Calibration 과정에서는 사용자가 모니터의 네 꼭지점과 중앙점을 쳐다볼 때 취득된 동공의 위치 정보들을 이용하였다.

#### IV. 명령전달 및 메뉴구동

본 논문에서 제안한 인터페이스 방법은 캘리브레이션(Calibration)과 종료모드로의 전환을 제외하고는 모두

눈동자에 의해 동작하도록 한다. 그러므로 사용자와 컴퓨터간의 명령전달과 메뉴구동을 하는 데 있어서 핸드 마우스와는 달리 새로운 정의가 필요하다. 명령전달 및 메뉴구동의 전체적인 흐름도는 그림 3 과 같다.

그림 3 에서 보듯이 명령을 전달하는데 있어서 오차 보정 모드라는 중간 단계를 거치게 된다. 응시 위치 모드에서 사용자가 모니터상의 한 지점을 일정시간이상 쳐다보게 되면 그림 4, 5 와 같이 이 영역의 화면이 확대되는 오차 보정 모드로 자동 전환된다. 즉, 현재 쳐다보는 화면 영상이 확대되어 있으면 화면내의 버튼 선택 역시 손쉬워지게 되는 것이다. 이러한 것은 HMD 에서 입력으로 받는 화면 해상도는 800×600 이고, 화면의 좌·우측 끝을 응시할 때 카메라로부터 파악되는 동공의 움직임 정도는 130pixel 이기 때문이다. 즉, 동공의 1 pixel 움직임이 화면상에는 약 6 pixel 움직임에 대응되므로, 이 이하의 세밀한 버튼 선택은 어렵게 되며 이러한 문제점을 해결하기 위하여 본 논문에서는 오차 보정 모드를 사용하여 세밀한 버튼 선택을 가능하게 하였다.

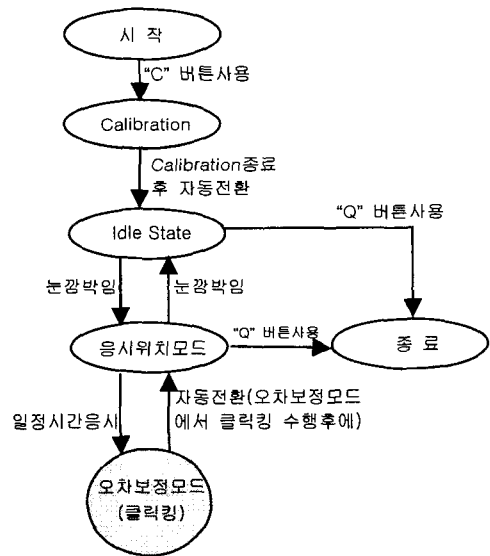


그림. 3 시스템의 모드 전환도

초기에 시스템을 구동하면 3. 2 에서 전술한 바와 같이 우선 캘리브레이션을 수행하도록 한다. 캘리브레이션이 종료되면 유희상태(Idle State)가 되는데, 이 상태는 눈동자를 움직이더라도 마우스 커서가 움직이지 않는 모드이다. 응시위치모드는 눈동자의 위치에 따라 마우스 커서를 이동시키는 모드이며 전술한 바

와 같이 특정 버튼 영역을 일정 시간이상 응시하게 되면 자동으로 오차 보정 모드로 전환된다. 동작 중에 Q 버튼을 누르게 되면 자동으로 시스템을 종료 시킬 수 있게 된다.

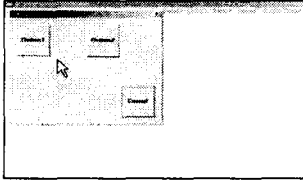


그림. 4 응시위치모드일 때 화면

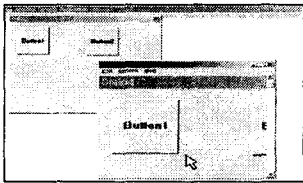


그림. 5 오차보정모드일 때 화면

### V. 실험 및 결과

본 논문에서는 타원 매칭 방법을 사용해서 동공의 중심을 찾는다. 이 방법은 글린트의 영향을 거의 받지 않으므로 강인한 방법임을 그림 6 으로부터 알 수 있었다.

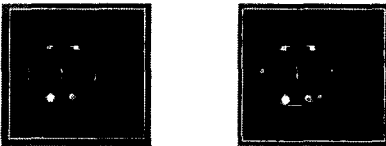


그림. 6 동공경계에 글린트가 존재할 때의 중심추출

표 1 은 응시위치를 매핑한 마우스 커서와 GUI 버튼 간의 거리 오차를 보여준다. 실험결과는 10 명의 사용자가 각각 모니터의 20 곳을 응시했을 때의 평균오차이다.

표.1 5 개점을 사용한 Affine 변환에 따른 응시위치오차

	응시위치오차	환경
X 축	34pixel	카메라영상크기 : 640×480 Virtual monitor : 15.4 inch at 70cm
Y 축	39pixel	

실험결과에 의하면 GUI 버튼의 크기가 너무 작게 되면 버튼 선택이 어렵다. 따라서 800×600 해상도 화면에서 70×80 크기로 버튼을 만들어서 사용자로 하여금

눈동자로서 마우스 커서를 움직이게 하고 명령을 전달할 수 있었다.

### VI. 결론

본 논문에서는 카메라를 통해 입력된 눈 영상을 이용하여, 사용자가 모니터 상에 쳐다 보고 있는 응시 위치 파악하고, 눈에 의한 명령 전달과 메뉴 구동하는 방법을 제안하였다. 먼저 캘리브레이션을 하고 눈 영상에서 타원 매칭 방법으로 동공 중심을 구한다. 그리고 나서 아핀 변환을 이용하여 사용자가 응시하는 위치에 따라 마우스 커서를 모니터에 디스플레이한다. 마지막으로 원하는 버튼으로 마우스 커서를 눈동자로 이동시킨 후 일정시간 동안 버튼 응시를 통해 클릭킹을 수행하게 된다. 실험 결과 GUI 버튼의 크기가 오차를 감안할 수 있는 70×80 의 크기로 선택하면, HMD(Head Mounted Display)가 부착된 웨어러블 컴퓨터에서 제안된 눈동자의 움직임을 이용한 인터페이스 방법이 실현될 수 있음을 볼 수 있다.

### Acknowledgement

본 연구는 한국과학재단 생체인식연구센터(BERC)의 지원으로 수행되었습니다.

### References

- [1] <http://www.sait.samsung.co.kr/sait/src/saitEnAward.jsp>
- [2] Ohno, T. "Quick Menu Selection Task with Eye Mark", Transactions of Information Processing Society of Japan, vol. 40, no. 2, pp.602-612, 1999.
- [3] Yamato, M., Monden, A., Matsumoto, K., Inoue, K. and Torii, K, "Quick Button Selection with Eye Gazing for General GUI Environments", International Conference on Software: Theory and Practice, August 2000.
- [4] Jacob, R.J.K. What You Look At is What You Get: Eye Movement-Based Interaction Techniques, Proc.CHI'1990.
- [5] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, Numerical Recipes in C, Cambridge University Press,1999.
- [6] Linda G. Shapiro and George C. Stockman, Computer Vision, Prentice Hall, 2001.