

## 실시간 눈과 시선 위치 추적

조현설\*, 유인호\*\*, 김희숙\*\*\*

\*청운대학교, \*\*국립익산대학, \*\*\*아산정보기능대학

### Real Time Eye and Gaze Tracking

Hyun-Seob Cho\*, In-Ho Ryu\*\*, Hee-Sook Kim\*\*\*

Chungwoon University\*, Iksan National College\*\*, Asan Polytechnic College Academic\*\*\*

**Abstract** - 본 논문에서는 새로운 실시간 시선 추적 방식을 제안하고자한다. 기존의 시선추적 방식은 사용자가 머리를 조금만 움직여도 잘못된 결과를 얻을 수가 있었고 각각의 사용자에게 대하여 교정 과정을 수행할 필요가 있었다. 따라서 제안된 시선 추적 방법은 적외선 조명과 Generalized Regression Neural Networks(GRNN)를 이용함으로써 교정 과정 없이 머리의 움직임이 큰 경우에도 견실하고 정확한 시선 추적을 가능하도록 하였다. GRNN을 사용함으로써 매핑기능은 원활하게 할 수 있었고, 머리의 움직임은 시선 매핑 기능에 의해 적절하게 시선추적에 반영되어 얼굴의 움직임이 있는 경우에도 시선추적이 가능도록 하였고, 매핑 기능을 일반화함으로써 각각의 교정과정을 생략 할 수 있게 하여 학습에 참여하지 않은 다른 사용자도 시선 추적을 가능케 하였다. 실험결과 얼굴의 움직임이 있는 경우에는 평균 90%, 다른 사용자에게 대해서는 평균 85%의 시선 추적 결과를 나타내었다.

#### 1. 서 론

시선 위치 추적이란 현재 사용자가 응시하고 있는 위치를 자동으로 인식하는 것을 의미한다. 시선 위치 추적은 많은 응용분야를 가지는데, 그 대표적인 것이 양손을 사용하지 못하는 장애인들을 위한 컴퓨터 인터페이스 및 사용자의 시선 위치에 따른 기계 제어 등 많은 이점을 가지고 있어 최근 많은 연구가 진행되고 있다.

시선 위치를 추정하는 방법으로는 머리를 이용하는 접근 방식[9]과 눈을 이용한 접근 방식[1, 2, 3, 4, 5, 7, 8, 11], 그리고 머리와 눈을 모두 이용한 접근 방식으로 분류할 수 있다. 머리를 이용한 접근 방식은 머리의 위치에 따라 시선의 위치를 결정을 하게 되는데 미세한 시선 변화를 감지하기에는 어렵다고 생각되어진다. 그리고, 눈에 근거한 접근방식은 시선과 홍채, 동공의 기하학적 특성에 대한 관계에 근거하여 시선을 추정하게 된다. 동공과 각막 반사에 의한 글린트 사이의 공간적 위치 특성으로 시선위치를 결정하고, 홍채의 위치와 형태, 왜곡으로서 시선을 위치를 추적한다. 이 때 신경망을 통하여 눈과 명암에 대해 정규화된 동공을 포함하는 작은 창을 추출하여 시선의 좌표를 결정하는데 사용한다.

현재까지 연구된 눈에 근거한 시선 추정 방식의 가장 큰 공통점은 각각의 글린트와 동공 사이의 상대적인 위치에 근거한다는 것이다.[2,3,4,5,7,8] 이러한 방식은 머리를 정지시킨 후, 기준점으로서 글린트를 사용하고, 동공의 중심으로부터 글린트까지를 나타낸 벡터로 시선 방향으로 나타내었는데 머리의 작은 움직임에도 잘못된 결과를 얻을 수 있다. 이러한 점은 인간과 컴퓨터 상호 작용

에 있어 중요한 장애요소이다. 또한, 현존하는 눈과 시선 추적 시스템의 또 다른 문제는 각각의 사용자에게 대하여 교정 처리를 수행할 필요가 있다는 점이다

이러한 제한 사항을 극복하기 위하여, Ohno[8]는 안구의 모형에 의한 새로운 시선 추적 기술을 제안하였다. 2개의 지점만 교정이 필요하도록 하여 시선 교정 절차를 단순화하였으나, 상대적으로 움직임이 거의 없는 경우만 동작하였고, 각각의 사용자에게 대한 정확한 기하학적 안구 모형을 얻기에는 많은 어려움이 있었다. 또한, Koons[7]는 2대의 카메라와 눈과 눈영상의 기하학적 특징을 이용하여 시선 교정 절차를 완전히 배제하는 것에 대하여 연구하고 있다. 최근 다른 연구[3,12]는 다양한 조명하에서 눈 추적의 견실성을 개선하는데 초점을 맞추고 있다.

본 논문에서는, 시선 교정 없이 견실하고, 정확하고 무리 없이 동작 될 수 있는 개선된 시선 추적 방법을 제안하고자 한다.

#### 2. 눈 추적

눈을 추적함으로써 시선 추적은 할 수 있었고, 눈 추적을 위해서 동공을 먼저 추적하였다. 동공 추적을 위하여 880nm 파장의 적외선 LED를 두 개의 고리모양으로 구성한 적외선 조명기를 제작하여 40nm 파장에서 32mW의 파워로 적외선을 조사한다. 이 적외선 광원은 밝은 동공 영상과 어두운 동공 영상을 생성하였는데, 밝은 동공 영상은 안쪽 고리의 적외선 LED들이 켜졌을 때 생성되었으며, 어두운 영상은 바깥쪽 고리의 적외선 LED들이 켜졌을 때 생성된다. 적외선 조명 이외의 광원들로부터 간섭을 최소화하여 영상의 질을 개선하기 위하여, 10nm 파장만을 통과시키는 광학 대역 필터를 사용하였고 필터를 사용함으로써 신호 대 잡음비를 증가시켰다. 그림 1은 중심이 같은 2개의 적외선 LED 고리와 대역 필터로 이루어져 있는 적외선 조명기를 나타내었다.

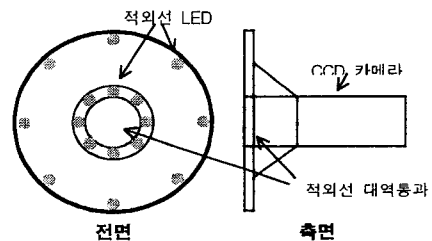


그림 1: 하드웨어 구성: 능동적 적외선 조명 카메라

동공 검출과 추적은 초기 프레임에서 검출된 동공 검출에 의하여 시작된다. 동공 검출은 동공의 명암과 SVM (Support Vector Machine)을 사용한 눈의 외관을 기준으로 하였으며, SVM을 사용함으로써 밝게 나타난 부분이 동공으로 오인식되지 않도록 하였다. 구체적인 동공 후보 지역은 어두운 동공 영상에서 밝은 동공 영상을 제한 차 영상에서 처음으로 검출되고, 동공 후보 지역들은 SVM에 의하여 가짜 동공 후보를 제거함으로써 유효하게 된다. 동공이 검출되어지면, Kalman 필터로 이전 프레임의 동공 위치를 가지고 현재의 프레임의 동공 위치를 예측할 수 있다.

### 3. 시선 결정 및 추적

제한된 시선 추정 알고리즘은 동공과 글린트 검출과 추적, 시선 교정, 그리고 시선 매핑으로 구성되어 있다. 본 논문에서는, 시선을 그림2에서와 같이 스크린(4x2)에 8개의 부분으로 양자화 하였다.

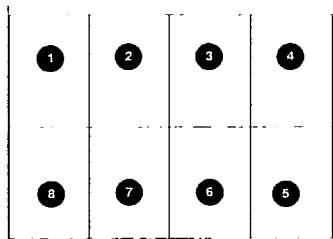
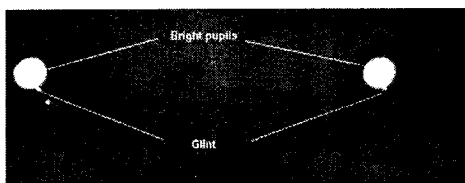


그림 2. 컴퓨터 스크린에 양자화 된 시선 영역

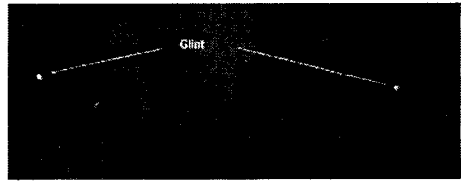
#### 3. 1 동공과 글린트 검출 및 추적

시선 추정은 동공과 글린트를 검출하고 추적함으로써 시작 된다. 시선 추정을 위해, 그림1과 같은 적외선 조명기를 사용하였고, 원하는 동공 효과를 만들기 위하여 2개의 LED 고리를 번갈아 키고 끄면서 비디오 디코더를 통하여 그림 3의 (a), (b)에 보여진 것과 같이 밝은 동공 효과와 어두운 동공 효과를 만들어 내었다.

글린트(작은 가장 선명한 반점)는 양쪽 영상 모두에서 나타났고, 밝은 동공 영상에서 동공의 검출과 추적을 할 수 있었다. 또한 매 프레임에서 동공의 위치는 중심 좌계에 의해서 특성 되어졌다. 글린트는 밝은 동공 영상에서 동공과 종종 같은 위치에 겹쳐지게 되므로 어두운 동공 영상에서 보다 쉽게 검출 될 수 있었다. 즉, 어두운 영상에서, 글린트는 눈 영상의 다른 부분 보다 밝게 되므로, 글린트 검출과 추적을 수월하게 한다. 동공 검출과 추적 방법은 어두운 영상에서 글린트를 검출하고 추적하는데도 사용될 수 있다. 그림4의 (c)는 발견된 글린트와 동공을 나타내었다.



(a) 밝은 동공 영상과 글린트

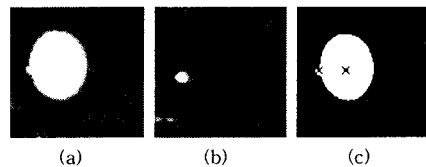


(b) 어두운 동공 영상과 글린트

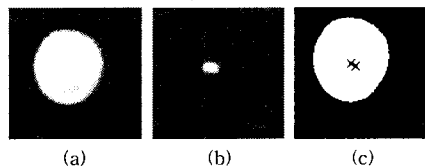
그림 3. 어둡고 밝은 동공 영상 및 글린트

#### 3. 2 시선 교정

검출된 동공과 글린트를 가지고 동공-글린트 벡터를 구성하고, 동공-글린트 벡터를 시선(스크린 좌표)에 매핑하는 매핑 기능을 수행토록 하였다. 그림 4는 글린트와 동공의 위치와 시선과의 연관성을 나타내었다. 매핑 기능은 교정을 통하여 결정되어지는데, 교정과정은 동공-글린트 벡터에 대응하는 스크린 좌표(시선)에 매핑하기 위한 매개변수를 결정하는 것을 의미한다. 시선 교정을 위해 전통적인 방법은 2가지 단점을 가지는데, 첫 번째는, 대부분의 매핑 기능이 원근 투영과 눈의 등근 표면으로 인한 적합하지 않을 수도 있는 다항식을 가지고 이루어진다는 것이고, 두 번째는 동공 중심과 글린트 위치에 대응하는 변위만을 시선 추정에 사용한다는 것이다. 이러한 점은 얼굴 위치에 따라 교정이 이루어지게 하고 또 다른 문제는 마지막 교정 후 머리가 아주 조금만 움직이게 되더라도 또 다시 교정을 필요로 한다는 것이다. 실제로, 현존하는 추적 방법으로는 사용자가 머리를 조금만 움직여도 잘못된 결과를 얻게 되므로 머리의 움직임을 시선 추정 과정에서 감안하도록 하여야 한다. 또 다른 문제는 사용자에게 따라 매핑 기능이 다르게 됨으로 각각의 사용자에게 대한 교정 과정이 수행되어야 한다는 것이다. 그러기에 본 논문에서는 이러한 문제점을 극복하기 위하여 신경망에 근거한 새로운 시선 교정방법을 제안하고자 한다.



좌측 주시



정면 주시

그림 4. 글린트와 밝은 동공 중심과의 공간적 연관성 (a) 밝은 동공 영상 (b) 글린트 영상 (c) 동공-글린트

#### 3.3 GRNN 을 통한 시선 교정

본 논문에서는 다양한 얼굴 자세와 다른 사람들의 시선을 추적하기 위해 신경망을 이용한 시선교정 방법을 제안한다. 동공과 글린트에 관한 매개 변수가 충분히 주어진다면 시선에 따라 유일한 기능을 갖는 동공과 글린트 매개변수가 있을 것이다. Specht[10]는 원형기준함수 네트워크(RBFN: Radial Basis Function Network)들과 확률 신경망 네트워크(PNN: Probabilistic Neural

Network)를 일반화 한 일반 회귀 신경망 네트워크 (GRNN: Generalized Regression Neural Network)를 제안하였다. GRNN은 확률 밀도 함수들의 추정에 근거한 기억 기반 피드 포워드 네트워크이다. GRNN의 특징은 빠른 학습 시간과, 비선형 기능의 모형을 만들 수 있다는 것과 충분한 데이터를 주어진 경우 복잡한 환경에서도 원활하게 동작한다는 것이다.

GRNN 위상은 입력 계층, 숨은 계층, 합산 계층, 그리고 출력 계층으로 이루어져 있다. 입력 계층은 1개의 출력 계층에 대하여 6개의 입력을 갖는다. 그리고 숨은 노드들의 수는 학습 샘플의 수와 같고, 합산 계층의 노드 수는 출력 노드 수에 1을 더한 것과 같다. 그림 5는 GRNN 구조를 보여주고 있다.

수직 수평 공간적 시선 분석에 의하여 출력 노드에서 시선 좌표 각각  $s_x$  와  $s_y$  를 나타내는 두 개의 동일한 GRNN을 구성한다. 입력 계층에 사용할 매개변수는 얼굴 거리와 방향을 따라 다양한 변화를 주어야 한다.

GRNN에 사용할 입력 벡터는 다음과 같다.

$$g = \{ \Delta x \ \Delta y \ r \ \theta \ g_x \ g_y \}$$

여기서  $\Delta x$  와  $\Delta y$  는 동공과 글린트 사이의 변화이고,  $r$  은 동공에 맞는 타원의 비율이다. 그리고  $\theta$  는 타원의 기울기이고,  $g_x$  와  $g_y$  는 글린트 영상의 좌표이다.

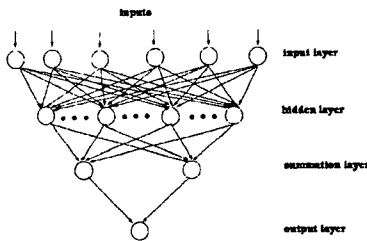


그림 5: 시선 교정을 위한 GRNN 구조

입력 매개변수는 다음에 근거하여 선택하였다.  $\Delta x$  와  $\Delta y$  는 글린트와 동공사이의 상대 이동을 위하여 사용한다. 글린트-동공 벡터의 크기는 적외선 조명기에 내장된 카메라와 피사체의 거리와도 연관된다.  $r$  은 얼굴 방향을 계산하기 위하여 사용한다. 얼굴이 정면을 향하고 있다면 비율은 1에 가까울 것이고, 얼굴이 좌우나 위아래로 움직인다면 비율은 1 보다 커지거나 작아진다. 기울기  $\theta$  는 얼굴 회전을 위하여 사용한다. ( $g_x$ ,  $g_y$ ) 는 평면 얼굴 변환을 위하여 사용되어진다. 이러한 매개변수들의 사용함으로써 얼굴과 동공이 움직이면 얼굴 움직임과 동공 이동을 계산하여 그에 부합하는 매개변수로 갱신할 것이다. 그럼으로 머리 움직임에 따른 영향을 감소 할 수 있다. 또한, 매개변수들은 동공의 크기에 대해서 독립적이므로 시선 매핑 기능은 일반화시킬 수 있어 다른 사용자를 위한 재계산 작업을 효과적으로 제거 할 수 있다.

입력 벡터는 GRNN에서 사용되기 전에 적합하게 정규화 되어야 된다. 정규화는 모두 입력 특징들이 같은 범위에 있게 하여 준다. 다른 머리 위치에 따른 막대한 학습 데이터가 GRNN의 학습으로 모아지고, 학습 데이터 수집되는 동안에, 사용자가 각 시선 영역에 시선을 고정하도록 요청하여, 각각의 시선 교정되는 동안 적합한 10 개의 매개변수 조합을 수집한다.

표1. 학습에 참여한 사용자의 얼굴 움직임 있는 100개의 시선 샘플에 대한 시선 추정 분류 결과

실제 주시 화면	예상 결과								성공률(%)
	1	2	3	4	5	6	7	8	
1	94	6	0	0	0	0	0	0	94
2	2	90	8	0	0	0	0	0	90
3	0	3	88	7	0	2	0	0	88
4	0	0	3	96	1	0	0	0	96
5	0	0	0	0	96	4	0	0	96
6	0	0	1	0	7	90	2	0	90
7	0	0	0	0	0	5	89	6	89
8	0	0	0	0	0	0	2	98	98

표2. 학습에 참여하지 않은 사용자의 60개의 시선 샘플에 대한 시선 추정 분류 결과

실제 주시 화면	예상 결과								성공률(%)
	1	2	3	4	5	6	7	8	
1	49	11	0	0	0	0	0	0	82
2	0	52	8	0	0	0	0	0	87
3	0	0	46	14	0	0	0	0	77
4	0	0	0	59	1	0	0	0	98
5	0	0	0	0	60	0	0	0	100
6	0	0	0	6	8	46	0	0	77
7	0	0	2	0	0	5	53	0	88
8	4	0	0	0	0	0	6	50	84

## N. 실험 결과 및 고찰

제한된 시선 추적기의 성능 실험을 실시하였다. 모니터 위에 카메라를 장착한 적외선 조명기를 설치하고 사용자가 1-1.5m 거리에서 약 15cm 정도 머리를 상하좌우로 움직이게 하거나, 좌우로  $\pm 20$ 도, 상하로  $\pm 15$ 도 정도 머리를 회전하도록 하여 자기 다른 얼굴 방향, 카메라의 거리변화, 다른 사용자 일 때의 성능을 각각 실험하였다. 표 1 과 2에 분류 결과들을 나타내었다. 제안된 시스템은 얼굴의 움직임 있는 경우에 대해 평균 90%, 다른 사용자에 대해서는 평균 85%의 결과를 나타내었다.

또한 실험결과 머리 움직임과 같은 외부 요인에 따라 수직 동공-글린트 벡터 측정이 수평에 비해 보다 민감하게 작용함으로써 동공의 수직 이동범위가 수평 이동 범위보다 작게 나타나는 것을 알 수 있었다. 이로 인해 수평 데이터에 비해 수직 데이터의 신호대 잡음비가 낮게 됨으로서 시선의 수직 분류를 수평 분류보다 작게 잡아 주게 되었다. 앞으로는 현재의 4x2 영역을 4x3, 그리고 5x4 영역으로 확장해 실험을 계속할 계획이다.

## V. 결론

본 논문에서는 시선 추적을 위한 새로운 방법을 제안하였다. 기존의 시선 추적 시스템과 비교하여 제안된 시스템은 교정 작업이 필요 없고 자연스러운 머리 움직임에도 정확하고 견실하게 시선을 추적할 수 있었다. GRNN을 이용한 새로운 시선 교정 방법을 사용함으로써 해석적 시선 매핑 기능을 사용할 필요가 없어졌고 머리의 움직임을 계산하여 매핑 할 수 있었다. 실험결과 움직임이 있는 얼굴에 대해 평균 90%, 다른 사용자에 대해서는 평균 85%의 시선 추적 결과를 나타내었다.

제한된 시선 추적기가 일부 기존의 시선추적기보다 정확하지는 않지만 머리의 움직임이 큰 경우에도 만족할만한 성능을 보여주고 교정작업이 필요 없다는 것이 중요한 이점이라고 사료된다.

## [참 고 문 헌]

- [1] S. Baluja and D. Pomerleau. Non-intrusive gaze tracking using artificial neural networks. Technical Report CMU-CS-94-102, Carnegie Mellon University, 1994.
- [2] Y. Ebisawa. Unconstrained pupil detection technique using two light sources and the image difference method. Visualization and Intelligent Design in Engineering, pages 7989, 1989.
- [3] Y. Ebisawa. Improved video-based eye-gaze detection method. IEEE Transactions on Instrumentation and Measurement, 47(2): 948955, 1998.
- [4] T. E. Hutchinson. Eye movement detection with improved calibration and speed. United States Patent [19], (4,950,069), 1988.
- [5] T. E. Hutchinson, K. White, J. R. Worthy, N. Martin, C. Kelly, R. Lisa, , and A. Frey. Human-computer interaction using eye-gaze input. IEEE Transaction on systems, man, and cybernetics, 19(6):15271533, 1989.
- [6] Q. Ji and X. Yang. Real time visual cues extraction for monitoring driver vigilance. in ICVS 2001: Second International Workshop on Computer Vision Systems, Vancouver, Canada, 2001.
- [7] D. Koons and M. Flickner. Ibm blue eyes project. <http://www.almaden.ibm.com/cs/blueyes>.
- [8] T. Ohno, N. Mukawa, and A. Yoshikawa. Freegaze: A gaze tracking system for everyday gaze interaction. Eye Tracking Research and Applications Symposium, 25-27 March, New Orleans, LA, USA, 2002.
- [9] R. Rae and H. Ritter. Recognition of human head orientation based on artificial neural networks. IEEE Transactions on Neural Networks, 9(2):257265, 1998.
- [10] D. F. Specht. A general regression neural network. IEEE Transactions on Neural Networks, 2:568576, 1991.
- [11] G. Yang and A. Waibel. A real-time face tracker. Workshop on Applications of Computer Vision, pages 142-147, 1996.
- [12] Z. Zhu, K. Fujimura, and Q. Ji. Real-time eye detection and tracking under various light conditions. Eye Tracking Research and Applications Symposium, 25-27 March, New Orleans, LA, USA, 2002.