

논문 2004-41SC-5-8

# 머리 움직임이 자유로운 안구 응시 추정 시스템

## (Eye Gaze Tracking System Under Natural Head Movements)

김 수 찬\*, Sked Matthew\*\*, Ji Qiang\*\*

(Soochan Kim, Matthew Sked, and Qiang Ji)

### 요 약

한 대의 카메라와 반사각의 조절이 가능한 2개의 거울, 그리고 별도의 적외선 광원을 이용하여 자유로운 머리 움직임이 가능한 안구 응시점 추정 시스템을 제안하였다. 거울의 회전 각도는 카메라의 광축(optical axis) 상에 안구가 올 수 있도록 공간 좌표계와 선형 방정식을 이용하여 계산하였다. 제안한 시스템은 수평 방향으로 90cm, 수직 방향으로 60cm 범위 내에서의 머리 움직임이 가능하였고, 응시점의 공간 해상도 각각 6°, 7°이며, 시간 해상도는 10~15 frames/sec이었다. Generalized regression neural networks(GRNN)을 기반으로 하여 2단계의 GRNN을 거치는 소위 hierarchical generalized regression neural networks(H-GRNN)을 이용하여 얻어진 인자를 모니터 좌표로 변환하였다. GRNN을 한번 사용하였을 경우 정확도가 85%이었으나 H-GRNN을 이용할 경우 약 9% 높은 94%의 정확도를 얻을 수 있었다. 그리고 입력 파라미터의 정규화를 통하여 재보정의 불편함을 제거했을 뿐만 아니라 약간의 얼굴 회전이 발생하였을 경우에도 동일한 성능을 보였다. 본 시스템은 공간 해상도는 크게 높지 않으나 자유로운 머리 움직임을 허용되므로 안정성과 피검자의 활동에 제약을 줄였다는 점에서 의의를 찾을 수 있다.

### Abstract

We proposed the eye gaze tracking system under natural head movements, which consists of one narrow-view field CCD camera, two mirrors which of reflective angles are controlled, and active infra-red illumination. The mirrors' angles were computed by geometric and linear algebra calculations to put the pupil images on the optical axis of the camera. Our system allowed the subjects head to move 90cm horizontally and 60cm vertically, and the spatial resolutions were about 6° and 7°, respectively. The frame rate for estimating gaze points was 10~15 frames/sec. As gaze mapping function, we used the hierarchical generalized regression neural networks (H-GRNN) based on the two-pass GRNN. The gaze accuracy showed 94% by H-GRNN improved 9% more than 85% of GRNN even though the head or face was a little rotated. Our system does not have a high spatial gaze resolution, but it allows natural head movements, robust and accurate gaze tracking. In addition there is no need to re-calibrate the system when subjects are changed.

**Keywords :** Eye movement, Gaze tracking, Non-invasive, Generalized Neural Network

### I. 서 론

인간의 시각은 망막에 맺히는 상을 통해서 인식되는데, 망막 중에서도 오목(fovea)이라는 좁은 영역만이 정

확한 상을 볼 수 있는 곳으로 알려져 있다. 그러므로 안구의 응시점으로부터 사람의 의지나 의도, 행동, 또는 외부 자극에 대한 반응을 객관적으로 평가할 수 있는 유용한 정보를 얻을 수 있다<sup>[1,2]</sup>. 특히, 응시점의 이동 경로(scan path)는 시각 인지(visual cognition)나 언어 인지 능력, 언어 학습 과정, 그리고 휴먼 컴퓨터 인터페이스(human computer interface, HCI)와 가상 현실(virtual reality) 등에 활용할 수 있는 주요한 인자 중에 하나이다<sup>[3-6]</sup>.

응시점을 얻는 방법으로 안구 주변에 전극을 부착하여 각막과 망막간의 전위를 측정하는 전기 안진기

\* 정회원, 한경대학교 생물정보통신대학원  
(Graduate School of Bio&Information Technology, Hankyong National University)

\*\* Rensselaer Polytechnic Institute  
(Electrical, Computer, and Systems Engineering Department, RPI)

※ 이 논문은 한국과학재단의 해외 Post-doc. 연수지원에 의하여 연구되었음.

접수일자: 2004년5월3일, 수정완료일: 2004년9월8일

(electrooculograph, EOG), 주변에 전자기장을 가하고 홍채 위에 직접 코일 전극을 얻겨 자기장 변화를 측정하는 sclera search coil(SSC), 적외선(Infrared)을 이용하여 홍채와 공막의 반사율과 굴절율을 이용하는 limbus tracker, 그리고 카메라로 직접 안구 영상을 측정하는 방법이 있으나 본 연구에서는 영상을 이용한 응시점 추정 시스템을 제안하고자 한다.

먼저, 영상을 이용한 방법의 장, 단점을 고찰해 보고자 한다. 영상을 이용한 방법에도 크게 2가지로 나누어 볼 수 있다. 첫째는 영상을 획득하기 위하여 카메라를 피검자에게 직, 간접적으로 부착하는 방식(contact method)과 일정 거리 떨어진 곳에 있는 카메라를 이용하여 특별한 추가적 장비의 부착을 요구하지 않는 비접촉식 방법(non-contact method)이 있다<sup>[7-14]</sup>. 첫 번째 방식은 응시점 추적에 핵심 정보를 담고 있는 안구 영상을 손쉽게 얻을 수 있고, 특히 카메라를 눈 가까이 부착함으로 고해상도의 안구 영상을 얻을 수 있어 오차가 1도 이하를 갖는 응시점 추정 시스템의 구현이 가능하다<sup>[12]</sup>. 하지만, 정확한 응시점 추정을 위해서는 안구의 절대적 위치를 알아야 하므로 공간 좌표를 얻기 위한 magnetic sensor와 같은 추가적인 장비나 공간 좌표계의 변화를 없애기 위해 머리의 움직임을 제한한다. 추가적 장비의 착용으로 인해 피검자의 시야를 방해하고, 착용이 번거롭기 때문에 연구나 임상을 위한 제한된 목적으로서는 유용하지만, 휴먼 컴퓨터 인터페이스나 일반인을 위한 시스템으로서는 이상적인 방법이라 볼 수 없다. 두 번째 방식인 비접촉식은 사용자의 시야나 활동에 전혀 방해를 주지 않기 때문에 현재까지 추구하는 응시점 추정 방법 중 가장 이상적인 방법에 가깝다. 그러나 어느 정도 떨어진 거리에서 안구 영상을 얻기 때문에 고글을 이용한 방법에 비하여 높은 해상도의 안구 영상을 얻기 어렵다. 광학 줌 렌즈(zoom lens)를 이용하여 안구 부분의 고해상도 영상을 얻는다 하더라도 카메라의 view field가 좁아지게 되어 조그마한 머리 움직임으로 인해 카메라는 안구 영상을 잡지 못하는 경우가 발생하게 된다. 이를 해결하고자 대부분의 연구에서는 wide-view field를 갖는 카메라와 이동이 가능한 narrow-view field(motorized camera)를 갖는 카메라를 함께 사용하여 이러한 문제를 해결하고자 하였다<sup>[13]</sup>. 하지만 2대 이상이 카메라를 이용할 경우 영상 획득보드(frame grabber)의 수가 증가되고, 카메라를 움직일 경우 카메라의 무게로 인해 카메라 제어를 위한 컨트롤 시스템이 복잡해지고, 시스템 구축을 위한 비용이 증가

되어 실용적인 응시점 추정 시스템 구현이 어렵다.

본 논문에서는 시스템을 단순화 하고, 시스템 구성비용을 고려하여 한 대의 카메라, 2개의 거울, 그리고 자체 적외선 광원(active infrared light source)을 이용하여 머리 이동을 자유롭게 보장하면서 동시에 응시점 추정이 가능한 시스템을 제안하고자 한다. 그리고 안구 영상에서 얻은 인자(parameter)로부터 응시점을 추정하는 함수로 generalized regression neural network (GRNN)을 이용하여 사용자가 바뀔 때 마다 재 보정을 해야 하는 불편함을 제거하고자 한다.

## II. 본 론

### 1. 시스템 구성

응시점 모니터링을 위한 시스템은 근적외선 대역까지 감지가 가능한 한 대의 카메라(TV ZOOM LENS M6Z, Hitachi, Japan)와 2개의 거울로 그림 1과 같이 구성하였다. 시스템은 컴퓨터 모니터 앞에 놓여지지만, 사용자 컴퓨터 작업을 하는 것에는 영향을 주지 않는다. 안구 영상은 Mirror 2와 Mirror 1을 거쳐 카메라로 입사된다. 각 거울의 직경은 15cm, 두께는 2mm로 거울의 반사각 조절이 가능하도록 모터(Futaba FPOS148 RC, Pontech, USA)에 부착하였다. Mirror 1은 수직 방향의 머리 움직임을 보정하기 위해 Y축 중심으로 회전하며, Mirror 2는 수평 방향의 머리 움직임을 보정하기 위해 Z축을 중심으로 회전한다. 모터는 RS232 포터를 통해 제어되고, 회전력(torque)은 3kg·cm이고, 거울의 회전 각도의 정밀도는 0.7°이다.

주변 조명에 의한 영향을 덜 받으면서 동공의 위치를 얻기 위해 자체 적외선 광원을 이용하였다. 광원으로 사용한 적외선 발광 다이오드의 주 파장대역은 880nm이다. 적외선 광원 이외에 다른 파장대역 즉 주변광의 영향을 줄이기 위해 880nm에서 대역폭이 10nm인 광학

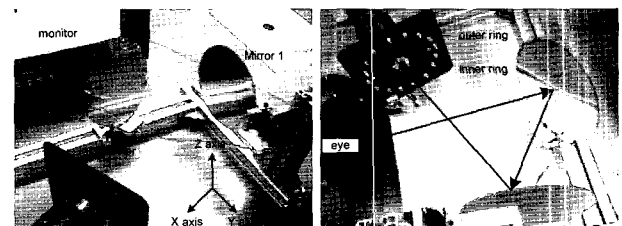


그림 1. 한 대의 카메라, 두 개의 거울, 그리고 자체 광원으로 구성된 안구 응시점 추정 시스템

Fig. 1. Eye gaze estimation system consisting of one camera, two mirrors, and active light source.

대역 통과 필터(band pass filter)를 사용하여 자체 광원에 의한 영상만 얻어지도록 하였다. 광원은 그림 1에서와 같이 안쪽 광원과 바깥쪽 광원으로 나뉘어 진다. 안쪽 광원은 카메라의 광축(optical axis)과 일치하도록 하여 적목 현상(red eye effect)을 유발시켰다. 그림 2(a)는 안쪽 광원이 켜졌을 때 얻어진 적목 현상이 있는 동공 영상(bright pupil image)이다. 바깥 광원이 켜지면, 적목 현상 없는 그림 2(b)와 같은 영상이 얻어진다. 각 광원의 on/off는 입력 영상 신호에서 필드 분리 신호와 동기를 맞추어 even field에는 적목 현상이 있는 영상, odd field에서는 적목 현상이 없는 영상을 얻었다. 적목 현상의 유무에 관계없이 얻어진 모든 영상에서는 모두 매우 밝게 반사되는 하얀 점을 볼 수 있는데, 이는 각막 반사에 의한 것으로 glint라 불리며, 응시점 추정에 유용한 정보를 제공한다<sup>[8],[9]</sup>.

영상 획득 보드로는 Winnov사의 Videum 1000 Plus를 사용하였고, 컴퓨터 환경은 Pentium IV(3GHz), 메모리 512MB에서 Windows XP 운영체제에서 Visual C++ 6.0으로 소프트웨어를 구현하였다.

2. 거울을 이용한 머리 움직임의 보정

머리 움직임에 따라 2개의 거울은 카메라의 광축에 동공의 이미지가 오도록 조절되어야 한다. 각 거울의 움직임은 서로 수직이므로 각각을 분리하여 각도를 계산할 수 있다.

먼저 수평 방향의 view field와 관련이 있는 Mirror 2의 조절에 대해서 알아보자. 그림 3(a)와 같이 카메라의 광축에 동공의 이미지를 움직이는 것과 사용자의 visual sight축이 Mirror 1과 2를 거친 후 카메라의 광축과 이루는 각도  $D$ 를 일정하게 유지하는 것과 동일한 문제이다. 그림 3(a)는  $Z$ 축을 중심으로 회전하는 Mirror 2의 회전 각도  $\Phi_2$ 를 계산하기 위해  $X-Y$  평면상으로 살펴 본 것이다. 각  $A, B, C$ , 그리고  $D$ 는 아래 식으로 나타낼 수 있다.

식 1로부터  $\Phi_2$ 와  $D$ 와의 관계를 식 2와 같이 나타낼 수 있고, 기존 각도  $D$ 에서 머리 이동에 의한 동공 위치의 변화된 각도  $D'$ 을 다시 각도  $D$ 로 유지하기 위해서는 Mirror 2의 이전 각도  $\Phi_{2old}$ 에서 식 (3)과 같이 각도  $(D-D')/2$  만큼의 각도 차이만큼 보정하면 된다.

수직 방향의 view field를 조절하는 Mirror 1의 각도의 계산은 Mirror 2에 비하여 간단하다. 왜냐하면 Mirror 2의 경우 Mirror 1의 각도도 함께 고려하여야

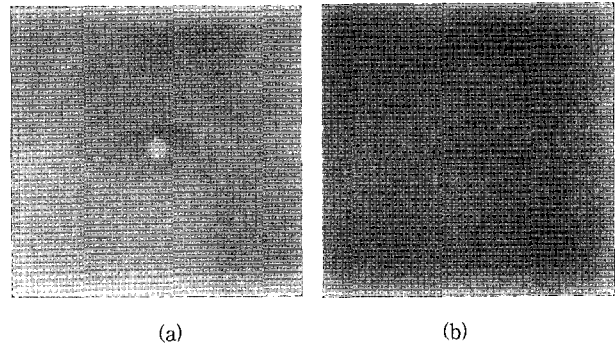


그림 2. 거울에 반사된 영상 (a) 적목 현상으로 인해 동공 영역이 밝게 얻어진 영상 (b) 적목 현상이 얻어진 영상  
 Fig. 2. The images reflected from two mirrors (a) the bright pupil image by red eye effect, (b) the dark pupil image.

하나 Mirror 1의 경우 Mirror 2의 반사를 고려하지 않아도 되기 때문이다.  $\Phi_1$ 은 Mirror 2로 반사된 동공의 각도와 Mirror 1과 카메라 광축과의 관계이다. 그림 3(b)는  $Y$ 축을 중심으로 회전하는 Mirror 1의 회전 각도  $\Phi_1$ 을  $X-Z$  평면상으로 살펴본 것이다. 식 (4)와 같이 Mirror 2에서 들어오는 동공 영상의 새로운 수직 입사각  $\Theta'_1$ 을 안다면  $\Phi_1$ 은 식 (4)와 같이 간단히 구해진다.

공간 좌표계의 절대 위치는 perspective camera 모델을 이용하여 계산하였고, 카메라 보정을 위한 외부 인자(extrinsic parameter)들은 checker board를 이용한 측정을 통하여 얻었다. 얻어진 영상으로부터 실제 공간상의  $\Theta_1$ 와  $\Theta_2$ 을 얻어 거울의 회전 각도  $\Phi_1$ 와  $\Phi_2$ 를 계산하였다.

$$A = \frac{\pi}{2} - \Phi_2 = \pi - B - C = \Theta_1 + \Theta_2$$

$$B = \frac{\pi}{2} - \Theta_2$$

$$C = \frac{\pi}{2} - \Theta_1$$

$$D = \pi - 2\Theta_1 - 2\Theta_2 = \pi - 2(\Theta_1 + \Theta_2) \tag{1}$$

$$\Phi_2 = \frac{\pi}{2} - \Theta_1 - \Theta_2 = \frac{1}{2} D \tag{2}$$

$$\Phi_{2new} = \Phi_{2old} + \frac{D - D'}{2} \tag{3}$$

$$\Phi_{1new} = \Phi_{1old} + \frac{(\Theta'_1 - \Theta_1)}{2} \tag{4}$$

3. 응시 방향 추정

응시점 추정에 기본 목표는 그림 4와 같이 컴퓨터 전

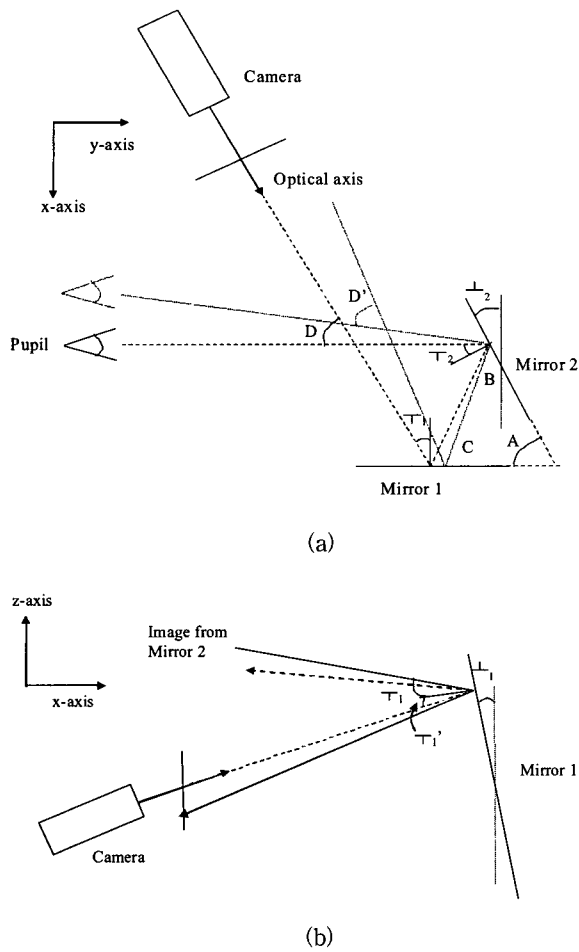


그림 3. 거울, 동공, 그리고 카메라의 광축 간의 관계 (a) Mirror 2의 반사각 계산 (b) Mirror 1의 반사각 계산

Fig. 3. The relationship among mirrors, pupil, and the optical axis of a camera (a) calculation of reflection angle for Mirror 2 (b) calculation of reflection angle for Mirror 1.

체 모니터를 총 24개 영역(6x4)으로 구분하여 이것에 대한 응시점 추정을 시도하였고, 각 번호는 응시 영역을 의미한다. 머리의 움직임을 추정하는 것과 안구의 응시점을 추정하는 것은 다른 문제이다. 응시점을 추정하기 위해서는 안구 및 머리의 위치뿐만 아니라 동공에 관한 보다 많은 정보가 필요하다. 본 연구에서는 안구 영상으로부터 응시점과 관련있는  $\Delta x$ ,  $\Delta y$ ,  $r$ ,  $g_x$ ,  $g_y$ ,  $\Phi_1$ , 그리고  $\Phi_2$ 의 정보를 얻었다. 여기서,  $\Delta x$ 와  $\Delta y$ 는 동공의 중심과 각막 반사로 인해 발생하는 glint 좌표에 대한 X와 Y축의 각각의 차이(displacement)이다.  $r$ 은 동공의 모양과 관련된 인자로 동공의 장축과 단축의 비이다. 동공은 이상적인 원이 아닌 타원에 가까우며 특히 카메라와 안구가 정면이 아닐 경우 타원으로 얻어진다. 이를 이용할 경우 사용자의 머리 회전이 발생하였

을 경우에도 응시점 추정이 가능할 수 있다<sup>[14]</sup>.  $g_x$ 와  $g_y$ 는 안구 영상에서의 glint의 좌표로써 평면상에 머리의 이동을 고려해 주기 위해서 사용하였다.  $\Phi_1$ 와  $\Phi_2$ 는 머리의 위치 정보를 얻기 위해 사용하였다.

안구 영상에서 동공 이미지의 정보를 얻어 내기 위해 적목 현상이 있는 영상에서 없는 영상을 뺀으로써 동공 부분의 안구 영상을 보다 손쉽게 얻을 수 있다. 하지만 광학 필터를 사용하더라도 주변광(형광등 혹은 자연광)에도 880nm 대역의 광이 존재함으로써 인해 잡음과 영상 자체의 노이즈가 여전히 존재한다. 이를 줄이기 위해 영상을 이진화한 후 morphological 방법을 이용하여 백색 잡음과 같은 노이즈를 제거하였다. 노이즈가 제거된 동공 영상을 ellipse fitting으로 모델링 한 후, 이것으로부터 동공의 중심과 동공의 장축과 단축을 얻어  $r$ 을 계산하였다<sup>[10,14]</sup>. Glint는 적목 현상이 없는 영상에서 가장 밝은 점의 집합들을 찾고, 그것들의 중심점( $g_x$ ,  $g_y$ )으로 설정하였다.

얻어진 7개의 인자로부터 응시점을 얻어낼 수 있는 함수로 universal function approximator로 잘 알려진 generalized regression neural network(GRNN)을 이용하였다. 왜냐하면, 학습 속도가 빠르고, 충분한 데이터가 주어진다면 노이즈에도 강한 특성을 보이기 때문이다<sup>[15]</sup>. GRNN의 경우 여러 입력에 대하여 연속적인 하나의 출력을 만드는데, 본 연구에서는 응시점의 가로와 세로 축 두 개의 위치이므로 각각에 해당되는 GRNN을 사용하였다. 그리고 인접한 이웃 응시점 간의 오차를 줄이기 위해 GRNN을 그림 5와 같이 2단계의 GRNN을 사용하였다(이하, hierarchical GRNN). 첫 단계에 있는 분류기(classifier)는 전체 학습 데이터(training data)를 이용하여 학습시켰고, 두 번째 단의 sub-classifier는 첫 단계에서 추정된 응시점과 그것과 인접한 응시점 영역의 학습 데이터만으로 학습하여 얻었다. 예를 들어, 첫 번째 응시점의 추정 결과가 그림 4에서의 영역 9라 가정하면, 자기 자신의 영역인 9와 이와 인접한 3, 8, 10, 15에 대한 학습 데이터를 이용하여 sub-classifier를 구성하였다. 9와 직접적으로 인접하지 않은 2, 4, 14, 그리고 16은 인접 영역에서 제외되었다. 왜냐하면 수평과 수직 모두 응시점이 잘못 추정되는 경우는 예비 실험 결과, 거의 없었기 때문이다. 이러한 sub-classifier는 각 응시 영역마다 존재하므로 총 24개가 얻어지며 첫 단의 분류기와 같이 가로축과 세로축에 대한 분류기로 각각 구성된다. sub-classifier의 선택은 첫 단계에서 추정된 응시점에 의해서 이루어진다.

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24

그림 4. 모니터 상의 응시점 위치  
Fig. 4. The eye gaze regions on the screen.

학습 데이터는 피검자가 머리를 고정한 상태에서 24개의 응시점을 그림 4에서의 1번째 영역부터 마지막까지 24번째 영역까지 차례로 응시할 때 얻었으며, 각 응시점마다 얻은 샘플의 수는 20개이다. 이후 머리를 조금 이동하여 새로운 머리 위치에서 앞과 동일한 방법으로 응시점에 대한 학습 데이터를 얻었다. 동일인에 대하여 이러한 과정을 머리의 움직임을 화면의 가운데, 좌측, 우측, 상측, 하측으로 이동하여 각각 얻었다. 총 12명에 대하여 오른쪽 안구 영상에 대하여 학습 데이터를 얻었다. 각 피검자로부터 얻은 데이터 중 동일한 응시점에 대하여 각 인자들의 평균과 표준편차를 이용하여 표준편차에서 3배 이상 벗어나는 데이터는 응시점의 부적절한 고정으로 인한 에러로 간주하여 제거하였다. 이러한 에러 값의 발생 원인으로서는 사람이 특정 곳을 늘 정확하게 응시할 수는 없기 때문이다.

GRNN의 학습을 위해 입력으로 넣기 전에 모든 데이터는 각 인자의 최대, 최소값을 기준으로 정규화 과정을 거친 후 입력으로 사용하였다. 이러한 정규화 과정은 사용자가 변경될 때마다 보정해야하는 불편함을 제거하기 위함이다.

### III. 실험

#### 1. 머리 움직임에 대한 안구 운동 추적

그림 6은 제안한 방법으로 머리가 상하, 좌우로 움직일 때 안구 영상을 계속적으로 카메라의 광축에 오도록 거울을 조절하여 머리 움직임을 보정해 주는 실험의 예이다. 거울에 반사되어서 입사된 영상은 직접 얻어진 영상에 비하여 상당히 어두운 결과를 보였다. 그러나 안구의 추적과 응시점 추정을 위한 인자의 추출은 가능하였다. 카메라와 거울의 초기 위치에 안구가 그림 6(a)와 같이 잡혔을 때 시스템은 정상적으로 구동하게 된다. 그림 6(b)와 같이 머리가 왼쪽으로 이동하게 되면

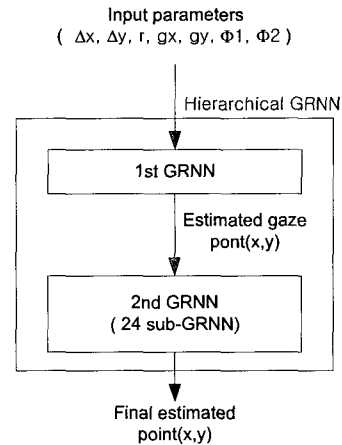


그림 5. 응시점 맵핑 함수로 사용한 H-GRNN의 구조도  
Fig. 5. H-GRNN structure used by gaze mapping function.

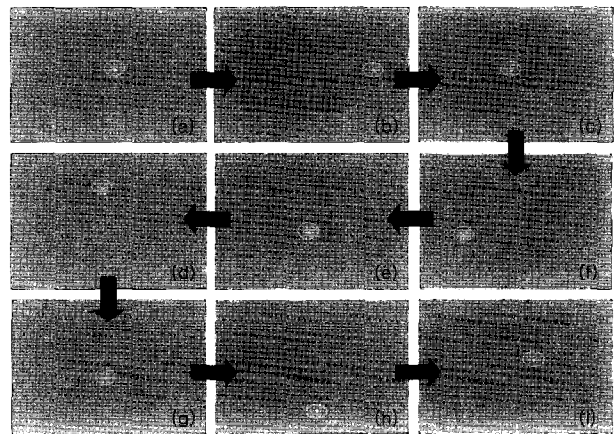


그림 6. 머리 움직임에 따라 안구를 추적하는 예  
Fig. 6. Eye tracking during head movements.

수평 조절 거울인 Mirror 2의 각도가 조정되어 그림 6(c)와 같이 안구가 카메라 중앙에 오도록 조절되었다. 머리가 그림 6(d)와 같이 위로 움직일 경우 수직 조절 거울인 Mirror 1이 조정되어 그림 6(e)와 같이 보정되었다. 그러나, 그림 6(i)의 경우 거울의 이동으로 머리의 이동을 보정한 직후의 영상임에도 불구하고 안구를 정확히 카메라 광축으로 이동시키지 못하여 동공이 영상의 중심에 오지 않고, 중심에서 약간 벗어나 있음을 볼 수 있다. 이것은 거울의 이동을 담당하는 모터의 회전 해상도가 정밀하지 못하기 때문으로 현재 해상도인 0.7°보다 정밀도가 높은 것을 이용하면 해결될 것이다.

머리의 빠른 이동이 발생할 경우 스텝 모터의 늦은 반응 속도로 인해 추적을 못하는 경우도 발생하는데, 이 역시 모터 드라이브 자체의 반응 시간이 적어도 5ms 이상의 시간 지연과 영상 획득 속도가 30Hz로 낮기 때문이다<sup>[16]</sup>.

표 1은 제안한 시스템으로 머리 이동을 보정할 수 있

표 1. 머리 움직임의 보정 능력

Table 1. Compensation capability of natural head movements.

Movement Direction	Left	Right	Up	Down
Length (cm)	36	58	30	30
Speed (deg/sec)	60	60	60	60

는 범위를 보여준다. 왼쪽보다 오른쪽이 허용 범위가 넓다. 이것은 그림 1에 보듯 광원의 부착으로 오른쪽의 경우 영상이 가려지기 때문이나 이것으로도 충분히 자유로운 머리 움직임을 보장한다.

거울에 반사되어 얻어진 그림 6과 같은 영상들이 거울을 사용하지 않고 직접적으로 얻어진 안구 영상보다 상당히 어렵다. 대부분 일반 거울의 가시광선 대역에서 반사율도 60~75%이며 적외선대의 반사율은 가시 광선 대보다 낮다. 예를 들어, 거울의 반사율이 70% 정도라고 가정했을 경우, 두 개의 거울을 거치게 되므로 실제적인 반사율은 50% 정도에 불과하다. 또한 반사되는 과정에서 일반 거울은 반사물질이 거울의 뒷면에 코팅된 back-surface 방식으로 빛이 유리 표면을 통과하여 반사가 이루어지므로 영상이 거울을 사용하지 않았을 경우보다는 흐려지게 되는 현상이 발생된다. 이러한 문제를 극복하기 위해서는 적외선 대역의 반사율이 우수하고 빛이 유리 표면을 통과하지 않고 표면에서 바로 반사가 되는 front-surface 거울을 사용할 것을 제안한다. 그러나 적외선대의 front-surface 거울은 일반 거울보다 수십 배 이상 비싼 편이므로 상품화를 고려한다면 신중히 고려해야 된다. 또 다른 방법으로는 광원으로 사용하는 적외선 다이오드의 수를 증가시키거나 전류를 증가시켜 보다 밝은 광원을 만드는 것이다. 그러나 광원을 밝게 할 경우 적외선으로 인해 안구 질환을 유발시킬 수 있으므로 국제 권고안(ANSI Z136.1) 등을 고려하여야 한다. 또 다른 한 가지 방법으로는 영상 획득 보드의 양자화 과정에서 dynamic range를 조절할 수 있는 것을 이용하는 것이다. 그러나 광원을 밝게 하거나 dynamic range를 조절할 수 있는 영상 획득 보드를 사용하여도 일반 거울을 이용할 때 발생하는 영상의 흐려짐은 향상시킬 수 없다.

## 2. 응시점 추정

표 2는 자연스러운 머리 움직임을 하면서 검사자가 레이저 포인터를 화면에 지적하였을 때, 피검자가 그곳을 응시하도록 하였고, 이때 제안한 시스템의 응시점

표 2. GRNN과 H-GRNN으로 얻어진 응시점 추정의 결과, GRNN (MEAN: 86±8%), H-GRNN (MEAN: 94±4%)

Table 2. Gaze classification results for the GRNN and the H-GRNN classifier. GRNN (mean: 86±8%), H-GRNN (mean: 94±4%).

[%]

		Horizontal					
		1	2	3	4	5	6
Vertical	1	GRNN 92	85	85	92	88	92
	H-GRNN	<b>96</b>	<b>87</b>	<b>94</b>	<b>96</b>	<b>96</b>	<b>100</b>
2	GRNN	72	88	87	89	89	78
	H-GRNN	<b>96</b>	<b>89</b>	<b>89</b>	<b>95</b>	<b>99</b>	<b>88</b>
3	GRNN	88	93	95	95	65	80
	H-GRNN	<b>94</b>	<b>93</b>	<b>95</b>	<b>95</b>	<b>93</b>	<b>92</b>
4	GRNN	70	85	93	92	90	74
	H-GRNN	<b>86</b>	<b>89</b>	<b>97</b>	<b>96</b>	<b>94</b>	<b>98</b>

추정 결과와 비교하였다. 각 응시점마다 얻어진 데이터의 수는 150개 이고, 10명을 대상으로 하였다. 머리 움직임은 제한한 시스템에서 제공하는 최대 범위까지 하지 않고, 수평 방향으로 40cm, 수직 방향으로 30cm 정도로 피검자가 모니터 앞에 앉은 상태에서 충분히 머리를 자유롭게 이동이 가능한 범위 내에서만 수행하였다. GRNN의 경우 평균 정확도는 85%였고, H-GRNN의 경우 94%였다. 예상과 같이 H-GRNN이 GRNN보다 정확도가 조금 더 높게 나타났다. 특히 응시점이 스크린 내부인 경우보다 바깥쪽에 해당되는 영역 3, 5, 7, 12, 13, 18, 19, 그리고 24(그림 4 참조)에서 정확도가 10% 내외로 두드러지게 향상되었다. 이것은 인접한 주변 영역만을 이용하여 학습된 sub-classifier가 해당 영역에 대해서는 보다 높은 공간 해상도를 가지고 있기 때문인 것으로 사료된다.

학습에 사용된 데이터 이외의 새로운 데이터 및 새로운 피검자를 시스템의 재 보정(re-calibration) 과정 없이 사용하였음에도 불구하고, 그 정확도가 90% 이상으로 높게 나온 것은 GRNN의 입력으로 데이터를 넣기 전에 수행한 정규화 과정의 효과로 볼 수 있다.

Narrow field view를 갖는 하나의 카메라만을 이용하므로 안구의 초기 위치를 찾는 것과 빠른 머리 운동으로 안구 추적을 성공적으로 수행하지 못하였을 때에는 사용자가 시스템의 view field내에 안구가 들어가도록 위치를 잡아야 한다.

모든 실험에 사용한 데이터는 오른쪽 안구와 가능한 정면을 바로 보면서 수행하였으나 머리를 약 ±10 정도 좌우로 회전한 상태에서도 정확도에 큰 차이 없이 응시

점 추정이 가능하였다. 하지만, 왼쪽 눈을 이용할 경우 오른쪽보다 정확하게 응시점 추정은 할 수 없었다. 그러므로 오른쪽 눈을 이용하여 학습된 시스템은 오른쪽 눈을 사용할 것을 권한다.

#### IV. 결 론

응시점 추정 시스템의 가장 큰 문제점 중에 하나가 머리의 움직임이 자연스럽게 하면서 사용자의 응시점을 정확하게 찾는 것이다. 본 연구에서 한 대의 카메라, 2개의 거울, 그리고 자체 광원을 사용하여 자연스러운 머리 움직임을 보장하는 응시점 추정 시스템을 개발하였다. 거울의 회전 각도는 공간 좌표계와 선형 수학을 이용하여 계산하였다. 제안한 시스템은 자연스러운 머리 움직임 하에서 응시점을 추정하였다. 또한 H-GRNN을 이용하여 GRNN보다 높은 정확도를 얻을 수 있었다.

제안한 방법은 사용자의 머리의 움직임에 제약을 주지 않는다는 점에서 기존의 방법과 큰 차이점이 있으며 입력 파라미터를 정규화하여 사용하였기 때문에 사용자가 바뀔 때 마다 보정할 필요가 없다. 제안한 방법은 물리적, 심리적 활동의 제약을 최소화하였기 때문에 측정으로 인한 바이어스가 없어야 되는 인지 과학이나 머리의 움직임을 고정하기 어려운 소아를 대상으로 하는 아동 심리학 연구에 활용 가치가 높을 것으로 기대된다.

#### 참 고 문 헌

- [1] H. Bekkering and S. F. W. Neggers, "Visual search is modulated by action intentions," *Psychological Science*, vol. 13, pp. 370-374, 2002.
- [2] P. Vuilleumier, "Perceived gaze direction in faces and spatial attention: a study in patients with parietal damage and unilateral neglect," *Neuropsychologia*, vol. 40, pp. 1013-26, 2002.
- [3] M. L. Phillips and A. S. David, "Visual scan paths are abnormal in deluded schizophrenics," *Neuropsychologia*, vol. 35, pp. 99-105, 1997.
- [4] G. Csibra, "Teleological and referential understanding of action in infancy," *Philos Trans R Soc Lond B Biol Sci*, vol. 358, pp. 447-58, 2003.
- [5] M. K. Tanenhaus and M. J. Spivey-Knowlton, "Eye-tracking," *Language and Cognitive Processes*, vol. 11, pp. 583-588, 1996.
- [6] C. L. Kleinke, "Gaze and eye contact: a research review," *Psychol Bull*, vol. 100, pp. 78-100, 1986.
- [7] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner, "Pupil detection and tracking using multiple light sources," *Image and Vision Computing*, vol. 18, pp. 331-335, 2000.
- [8] Y. Ebisawa, "Improved video-based eye-gaze detection method," *Instrumentation and Measurement, IEEE Transactions on*, vol. 47, pp. 948-955, 1998.
- [9] D. H. Yoo, J. H. Kim, D. H. Kim, and M. J. Chung, "A human-robot interface using vision-based eye gaze estimation system," presented at Intelligent Robots and System, 2002. IEEE/RSJ International Conference on, 2002.
- [10] Q. Ji and X. J. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real-Time Imaging*, vol. 8, pp. 357-377, 2002.
- [11] K. Ukai, S. Saida, and N. Ishikawa, "Use of infrared TV cameras built into head-mounted display to measure torsional eye movements," *Japanese Journal of Ophthalmology*, vol. 45, pp. 5-12, 2001.
- [12] 김수찬, 남기창, 이원상, 김덕원, "Development of Videoculograph for Vestibular Function Test", *대한전자공학회 논문지 SC편* 40(3):189-198, 2003.
- [13] B. Nourddin, A Non-Contact Video-oculograph For Tracking Gaze in a Human Computer Interface, Thesis(M.A.Sc.), University of British Columbia, 2003.
- [14] Q. Ji and Z. Zhu, "Non-intrusive Eye and Gaze Tracking for Natural Human Computer Interaction," *MMI-Interaktiv Journal*, vol. 6, 2003.
- [15] D. F. Specht, "A General Regression Neural Network," *IEEE Transactions on Neural Networks*, vol. 2, pp. 568-576, 1991.
- [16] Pontech, SV203 Servo Motor Controller Board User's Manual, [http://www.pontech.com/products-/sv200/sv203\\_v.1.20\\_webmanual.pdf](http://www.pontech.com/products-/sv200/sv203_v.1.20_webmanual.pdf)

저 자 소 개



김 수 찬(정회원)  
 1994년 인제대학교 의용공학과  
 졸업(학사)  
 1998년 연세대학교 생체공학 협동  
 과정 졸업(공학석사)  
 2003년 동대학원 졸업(공학박사).  
 2003년 3월~2003년 11월 연세대  
 의대 의학공학교실 연구강사.

2003년 12월~2004년 9월 Rensselaer Polytechnic  
 Institute 박사후 과정,  
 2004년 9월 현재 한경대학교 전임강사  
 <주관심분야: 생체계측 및 처리, Affective  
 computing 등>



Qiang Ji  
 1998년 University of Washington  
 졸업(공학박사).  
 1999년~2000년 University of  
 Nevada 조교수  
 2001년~현재 Rensselaer Polytechnic  
 Institute 조교수

<주관심분야: 컴퓨터 비전, 비간섭형 응시점 추정,  
 피로도 측정>



Matthew Sked  
 2004년 Rensselaer Polytechnic  
 Institute 석사 졸업  
 (공학 석사)