

## Head Mounted Camera 환경에서 응시위치 추적

이철한, 이정준, 김제희  
연세대학교 전기·전자공학과  
전화 : 02-2123-4537 / E-mail:devices@hanmail.net

### Gaze Detection in Head Mounted Camera environment

Chul-Han Lee, Jeong Jun Lee, Jaihie Kim  
Dept. of Electrical and Electronic Engineering Yonsei University  
E-mail : devices@hanmail.net

#### Abstract

Gaze detection is to find out the position on a monitor screen where a user is looking at, using the computer vision processing. This System can help the handicapped to use a computer, substitute a touch screen which is expensive, and navigate the virtual reality. There are basically two main types of the study of gaze detection. The first is to find out the location by face movement, and the second is by eye movement. In the gaze detection by eye movement, we find out the position with special devices, or the method of image processing. In this paper, we detect not the iris but the pupil from the image captured by Head-Mounted Camera with infra-red light, and accurately locate the position where a user looking at by Affine Transform.

#### I. 서론

응시 위치 추적이란 카메라에 입력된 사용자의 얼굴이나 눈으로부터 사용자가 쳐다보고 있는 모니터상의 위치를 컴퓨터 비전(Computer Vision) 방법에 의해 파악해 내는 연구이다. 이런 응시 위치 추적은 장애인 단말기나, 가상현실에서 화면(view)의 조정, 터치 스크린(Touch Screen)의 대체 등에 응용 될 수 있다. 응시위치를 추적하는 연구는 크게 눈동자의 움직임을 이용한 응시위치 추적 연구와 얼굴 움직임에 의한 응시 위치 추적 연구가 있다. 눈동자에 의한 응시 위치 추적 연구로는

특수장비를 이용한 방법과 영상처리를 이용하는 방법이 있고, 얼굴의 움직임만을 이용하는 연구에는 얼굴의 삼차원 움직임을 구하는 방법, 시선자의 시선 벡터로 계산하는 방법 등이 있다.

본 논문에서는 적외선 LED가 부착된 Head Mounted Camera 환경에서 정확한 응시위치 추적을 위해 눈동자가 아닌 동공을 추적하고 이를 아핀 변환(Affine Transform)을 통해 좀더 정확한 응시위치 추적에 관해 논한다.

#### II. 동공의 중심 검출 및 추적

Head-Mounted 카메라 환경은 그림 1(a)과 같다. 헬멧 형태에 연결된 카메라는 얼굴 앞 10cm 정도에 달려 있고 이 카메라는 얼굴과 함께 움직여 얼굴움직임에 상관없이 그림 1(b)와 같이 고정된 영상을 얻을 수 있다.

눈동자에 의한 응시 위치 추적을 위해서는 정확한 눈동자의 움직임을 추적해야한다. 본 논문에 쓰인 장비는 카메라 주변에 적외선 LED 가 있어 홍채와 동공을 분리시키고 아래와 같이 두 단계를 걸쳐 정확한 동공의 위치를 추출한다.

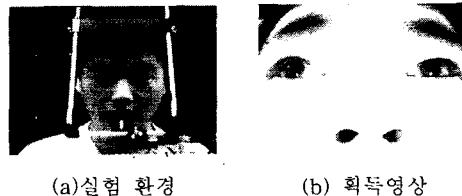


그림 1. 실험 환경과 획득영상

동공의 검출 및 추적의 전체적인 흐름도는 다음과 같다.

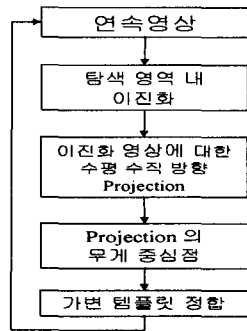


그림 2. 동공의 검출 및 추적의 전체적인 흐름도

### 2.1 무게중심을 이용한 동공의 초기 값 검출

(1) 초기 영상에서 동공의 초기 값 검출

Head-Mounted Camera 환경에서 들어온 영상은 그림 1(b)와 같이 항상 눈은 전체영상의 Y축으로는 중간에 위치한다. 이런 사전정보를 이용해 동공의 초기 값 검출을 위한 검색 범위를 줄일 수 있다. 탐색 영역 설정은 오류를 막기 위해 되도록 큰 탐색 영역을 설정하여야 한다. 본 논문에서는 실험을 통해 얻어진 결과를 바탕으로 전체영상의 중간 1/3을 탐색영역으로 설정하였다. 이 영상내에서 이진화(그림 3(a))를 하고 수직방향 투영(Projection)을 하면 그림 3(b)와 같은 결과를 얻을 수 있다. 이때 히스토그램이 2개의 피크(peak)를 가지면 상단의 피크를 눈썹이라 생각하고 제거하면 눈썹이 제거된다. 만약 피크가 하나만 나오면 영상을 제거하지 않고 그냥 사용한다.

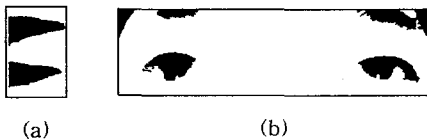


그림 3. 눈영역 이진화 영상(a) 및 수직방향 투영(b)

눈썹영역을 제거한 이진영상(그림 4(a))에서 수평방향 투영(그림 4(b))을 하고 수평 방향과 수직 방향 투영의 무게 중심점을 동공의 초기 좌표로 한다[2].

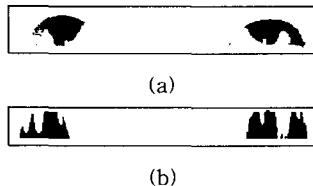


그림 4. 눈썹영역 제거 영상(a) 및 수직 방향 투영(b)

(2) 이후 영상에서 동공의 초기 값 좌표 검출

초기영상에서 구한 무게 중심점을 기준으로 일정 영역을 설정하여 그 영역에서만 수직 수평 히스토그램을 구하고 그 히스토그램의 무게중심점을 동공의 초기 좌표로 한다.

### 2.2 가변 템플릿(Deformable Template)을 이용한 동공의 검출 및 추적

가변 템플릿[3][4][5]은 실제 눈동자의 정확하게 정합되기 위해서 이동 방향과 크기를 변형시켜야 한다. 이동 방향과 크기를 결정하기 위해서는 윈도우 에너지를 중심으로 이동에너지와 크기변화 에너지를 정의하여야 한다. 먼저 윈도우 에너지는 식 (1)과 같이 정의된다.

$$f_w = \frac{\sum(\text{윈도우내의 각화서값})}{\text{윈도우내의 화소수}} - \frac{1}{2} \quad (1)$$

가변 템플릿 정합은 이진화된 영상을 사용하므로 화소 값들은 '0' 또는 '255'의 값을 갖는다. 그러므로, 윈도우 에너지의 값은 윈도우 내에 흑화소('0')의 개수가 백화소('255')보다 많은 경우 음(-)의 값을 갖게되고, 흑화소의 개수가 적을 경우 양(+)의 값을 갖게된다.

윈도우 에너지를 바탕으로 이동에너지는 식 (2)와 같이 정의된다.

$$f_{move_x} = \int_0^{2\pi} f_w(\theta) \cos \theta d\theta$$

$$f_{move_y} = \int_0^{2\pi} f_w(\theta) \sin \theta d\theta \quad (2)$$

템플릿은 눈동자와 정합되기 위해서 좌우(x축 방향) 또는 상하(y축 방향)로 움직여야한다. 그러므로 방향성분에 따라 에너지를 식 (2)와 같이 따로 정의하였다. x축 방향에너지는 템플릿의 모든 윈도우 힘의 cosine 성분으로 정의할 수 있고, y축 방향에너지는 템플릿의 모든 윈도우 힘의 sine 성분으로 정의할 수 있다. 이 방향 성분으로 템플릿을 움직임 방향을 결정할 수 있다. 크기 변화에너지는 식 (3)과 같이 정의된다.

$$f_{resize} = \int_0^{2\pi} f_w(\theta) d\theta \quad (3)$$

템플릿의 모든 윈도우의 에너지 합으로 정의한다. 에너지의 합이 양이면 템플릿은 수축하게 되고 음이면 팽창하게 된다.

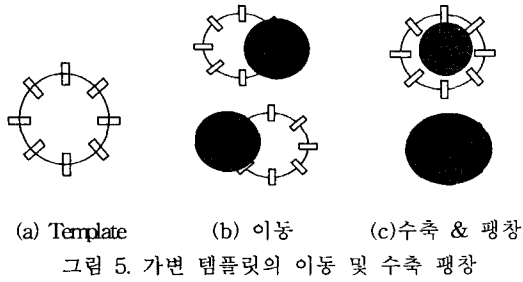


그림 5는 가변 템플릿의 변환을 보여주는 것이고 그림 6은 위에서 설명한 두 단계를 걸쳐 최종적으로 동공을 추출한 결과 영상이다.



그림 6. 동공 추출의 결과

### III. 응시 위치 추적

모니터를 동일하게 9등분하여 사용자가 9개점(그림 7 (a))을 본 결과 동공 좌표의 변화량이 모니터의 위치와는 다른 굴곡된 형태를 나타냄을 실험(그림 7 (b))을 통해서 확인했다.

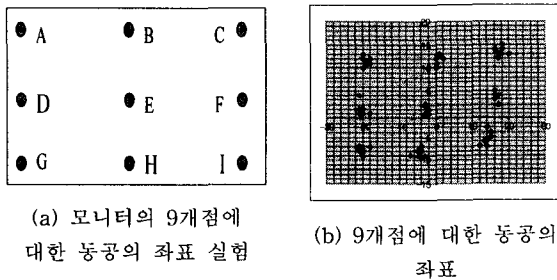


그림 7. 동공의 좌표 변화

본 연구에서는 동공의 움직임양 좌표를 모니터의 좌표계에 일치시켜 사용자의 응시위치를 정확하게 추적하기 위해 아핀 변환 방법을 사용하였다.

#### 3.1 Calibration을 이용한 아핀 변환 계수 값 결정

이차원 아핀 변환[6]은 다음과 같은 좌표변환 형태를 가진다.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{xx} & a_{xy} \\ a_{yx} & a_{yy} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_x \\ b_y \end{bmatrix} \quad (4)$$

원래 좌표(x,y)는 아핀 변환 계수를 통해 새로운 좌표 x',y'로 변환하는데  $a_{xx}$ ,  $a_{xy}$ ,  $a_{yx}$ ,  $a_{yy}$ ,  $b_x$ ,  $b_y$  는 변환 타입에 따라 정해지는 상수이다. 아핀 변환은 일반적으로 다음과 같은 특성이 있다. 평행의 직선은 평행한 직선으로 변환되며, 유한한 점들은 유한한 점들로 변화되는 선형적인 변환이다. 이동(translation), 회전(rotation), 확대 및 축소(scaling), 반사(reflection), 전단변환(shear)은 이차원 아핀 변환의 예이다. 어떤 이차원 변환도 위 변환들의 조합으로써 표현될 수 있다. 이 변환은 동공 움직임에 의한 응시위치 추적에 다음과 같이 이용될 수 있다. 앞에서 모니터의 9개점을 볼 때 동공의 좌표를 다시 표현하면 그림 8과 같은 형태를 가진다.

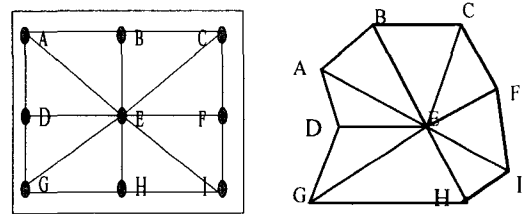


그림 8. 모니터의 9개점에 대한 동공의 좌표

동공의 좌표에 의해 굴곡된 8개의 삼각형을 각각 모니터의 직삼각형 좌표로 변환하기 위한 아핀 변환을 구하면, 이 아핀 변환을 통해 삼각형 내부의 점은 모니터의 한점으로 매칭 됨으로써 응시 위치를 추적 할 수 있다.

$$\begin{aligned} X &= a_{xx}x + a_{xy}y + b_x \\ Y &= a_{yx}x + a_{yy}y + b_y \end{aligned} \quad (5)$$

위 아핀 변환식(식 (5))에서 아핀 계수인  $a_{xx}$ ,  $a_{xy}$ ,  $a_{yx}$ ,  $a_{yy}$ ,  $b_x$ ,  $b_y$ 를 구하면 되는데 이 계수들을 구하기 위해 본 연구에서는 Calibration을 이용했다.

#### 3.2 Calibration 과정

사용자로 하여금 모니터 상의 9개점을 응시하게 하여 이때 구한 동공의 좌표를 아핀 변환의 계수를 결정하는 값으로 이용한다. Calibration의 오류를 미리 막기 위해 실험적으로 구한 각 9개점의 동공의 좌표범위를 구해 Calibration시 그 범위 안이 아니면 다시 그 점을 Calibration 시키고 그 범위 내부면 다음점을 Calibration

해 9개점에 대한 동공의 좌표를 구한다.

### 3.3 응시 위치 추적

Calibration 과정을 거치면 식 (6)의 아핀 계수를 구할 수 있다. 이 계수를 통해 새로운 동공 좌표(x,y)가 입력이 되면 모니터 상의 응시위치(X,Y)를 구할 수 있다.

$$X = a_{xx}x + a_{xy}y + b_x, Y = a_{yx}x + a_{yy}y + b_y \quad (6)$$

(X,Y : 응시 위치 좌표, x,y : 동공의 좌표,  $a_{xx}, a_{xy}, a_{yx}, a_{yy}, b_x, b_y$  : 아핀 계수)

## IV. 실험 및 결과

### 4.1 실험 환경

실험은 동공의 추적의 정확도와 시간, 응시 위치 추적의 정확도에 대해 시행하였다. 모니터와 사용자 사이에 일정한 거리(50cm)를 유지하고 얼굴의 움직임이 고정된 상태에서 모니터의 20영역을 응시하여 실험을 실시하였다. 실험에 쓰인 장비는 Petium III 500 MHz를 사용했다.

### 4.2 실험 결과 분석

동공의 정확도의 비교를 위해 동공의 중심점을 5사람이 증점으로 택한 점의 평균을 기준값으로 설정하였고, 동공추출의 결과에서 나온 중심치와의 거리를 오차로 정의하였고, 연속영상을 처리하기 위한 충분한 속도가 나오는지 또한 실험하였다. 실험결과(표 1) 연속 영상에서 응시위치 추적을 위한 정확도와 속도의 결과를 나타냈고 응시 위치 오차의 실험 결과는 표 2와 같다. 응시 위치의 오차는 Calibration의 정확성에 따라 많이 변함을 알 수 있었고, 얼굴고정의 가정이지만 사용자는 조금의 움직임이 있어 오차가 발생함을 알 수 있었다.

표 1 동공의 추적의 오차 및 속도

정확도	처리 속도	시스템 환경
X축 : 1.9 pixel Y축 : 1.8 pixel	0.006sec/frame	-영상크기 640 × 480 pixel -동공의 지름 25 pixel -Petium III 500 MHz 환경

표 2 응시 위치 오차

축	응시 위치오차	환경
X 축	1.1 inch	실험테이터 : 7명×20곳의 응시위치=140데이터 모니터 size : 19inch
Y 축	1.3 inch	

## V. 결론 및 추후연구

본 논문에서는 Head-Mounted 카메라 환경에서 응시위치 방법에 대해 논하였다. 연속영상에서 응시위치 추적을 위해서는 정확하고 빠른 동공의 위치를 추출하여야 한다. 이를 위해 본 논문에서는 무게중심으로 대강적인 눈의 위치를 알아내고 가변 템플릿을 이용하여 정확한 동공의 위치를 추적하는 두단계를 걸쳐 동공의 위치를 정확하게 추출하였고, Monitor의 좌우상하를 응시할 때 선형적이지 않은 눈동자의 움직임에 대한 응시위치를 추적하기 위해 9개점에 대해 Calibration 후 아핀 변환을 이용해 응시위치 추적함으로써 X축으로 1.1 inch Y축으로 1.3 inch 의 오차를 내었다.

향후에는 9개점을 Calibration 하는 불편함을 줄이는 방법에 대한 연구가 수행되어야 할 것이고 또한 눈동자에 의한 응시위치 추적만이 아니라 얼굴과 눈동자를 함께 고려한 응시위치 파악 연구도 수행되어야 할 것이다.

### 참고문헌

- [1] C. Colombo, A. Del Bimbo, S. De Magistris, "Human-Computer Interaction Based on Eye Movement Tracking" IEEE 1995.
- [2] 박현희, 이동재, 김재희 "눈 영상에서 동공 분할에 관한 연구"1999년 한국 통신 공학회 추계 종합 학술 발표회 논문집(하) Vol 20, No 2, pp1719 ~ 1722 1999년.
- [3] J. Y. Deng, Feipei Lai, "Region-Based Template Deformation and Masking for Eye-Feature Extraction and Edscription," Pattern Recognition, Vol. 30, No. 3, pp. 403-419, 1997.
- [4] X. Xie, R. Sudhakar, "On Improving Eye Feature Extraton using Deformable Templates," Pattern Recognition, Vol. 27, No. 6, pp. 791-799, 1994.
- [5] Yuille, A. L, P. W. and Cohen, D. S, "Detection facial features using deformable templates", International Journal of Computer Vision.
- [6] Donald Hearn, M. Pauline Baker "Computer Graphics C version" pp 183 ~ 215.