

에러 보정 템플레이트를 이용한 모니터 응시점 추출

°박기범 최형일

송실대학교 컴퓨터학부

Extraction of gaze-points by circular error correction template

Ki-Bum Park, Hyung-II Choi

School of Computing Soongsil University

요 약

모니터의 흰 좌표를 응시하고 있을 때의 눈 영상을 카메라로 획득하고, 눈 영역 영상에 대한 눈동자의 좌표를 계산하여 사용자가 모니터의 어느 부분을 응시한 것인지를 알아내는 것이 본 논문의 주요 목적이다. 카메라로 획득하여 계산된 눈동자의 좌표는 2차원 평면상에서 계산된 것이므로, 응시 각도나 응시 거리등의 3차원 정보가 손실되어 응시점에 대한 보다 정확한 매핑을 하기 위해 눈동자 좌표의 보정이 필요하게 된다. 획득된 눈 영상의 눈동자 좌표 이외에 흰자위 비율 및 원형 템플레이트를 사용한 가중치 벡터를 더하여 개략적으로 얻어진 눈동자 좌표에 대한 응시 초점의 오차를 보정한다. 눈 영상에서의 보정된 눈동자 응시 초점 좌표를 모니터 평면 좌표로 매핑하여 모니터에 대한 사용자의 응시점을 찾아낸다. 이때, 계산된 모니터의 응시점이 명령 영역에 해당하면, 해당 명령을 실행시킴으로써 사용자는 모니터 화면을 응시하는 단순 동작만으로 컴퓨터에게 원하는 명령 실행을 요구할 수 있으므로 키보드나 마우스의 입력장치를 본 시스템으로 대체할 수 있다.

1. 서론

기존의 Eye tracking system에서는 작은 카메라가 달린 헬멧 장치를 머리에 쓰고 눈동자의 움직임을 카메라로 관찰하는 것이 대부분이었으며, 지진 압 적외선 발광장치를 눈에 쏘아 동공의 위치와 각막 반사광의 각도차기 학습 단계에서 calibration된 거리값들과 어느 만큼 다른지를 파악하여 눈동자의 위치를 계산하는 연구들이 있었다 [1, 2, 3]

본 시스템은 눈동자 추적을 위해 기존의 헬멧 장치나 적외선 발광장치를 눈에 쓰는 등의 사용자에게 부담을 줄 수 있는 요소들을 배제하였으며, 컴퓨터 비전 기술만을 이용하여 사용자의 모니터 응시점에 대한 명령을 실행시킴으로써 기존 시스템들과의 차별성을 두었다

획득한 눈 영상에 대한 칼라 정보를 이용하여 눈에 대한 레이블링과 MER을 구하여 눈 영역을 검출하였으며, MER의 범위에서 눈에 대한 프로젝션을 수행하여 눈 영상에 대한 눈동자 위치 좌표를 계산하였다. 또한, 이를 바탕으로 원형 템플레이트의 흰자위 비율을 이용하여 눈이 움직인 방향으로 가중치 벡터를 줌으로써 보다 정확한 눈동자의 응시좌표를 구한다. 이렇게 구해진 응시좌표를 모니터의 좌표로 인터폴레이션 매핑을 함으로써 모니터 평면에 대한 사용자의 응시점에 해당하는 명령을 수행한다.

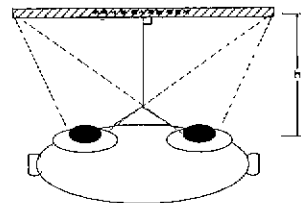
본문의 구성은 응시좌표 평면계에 대한 내용과 시스템의 구조도, 눈동자의 위치 계산 및 오차 보정 방향의 순서로 하였으며, 실험 및 결과의 향후 연구 방향에 대하여 논하겠

2. 응시 좌표 평면계의 정의 및 벡터표현

그림 1의 응시 좌표 평면계는 교와 수직적인 관계를 가지는 평면의 좌표표

현으로써, 머리를 고정시킨 상태에서 사물을 응시할 수 있는 모든 범위를 포함하는 평면 좌표계로 정의한다.

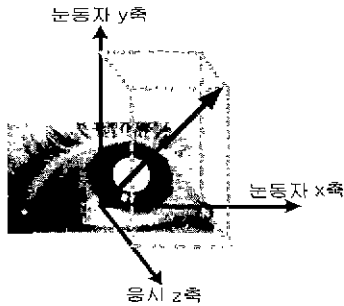
그러므로, 응시좌표 평면계의 좌표 범위는 좌측 최상단과 우측 최하단이 눈동자만을 이동하여 인기할 수 있는 최대한의 범위를 포함하며, 결국, 눈동자 기 응시하는 어느 한 점은 응시좌표 평면계의 한 점과 대응되고, 머리를 회전하면 그쪽 방향으로 응시좌표 평면계는 이동한다. 본 시스템은 사용자가 모니터를 응시하는 경만을 인정하므로 응시 좌표 평면계를 모니터 평면으로 제한하였으며, 응시 좌표 평면계와 눈의 거리는 모니터의 눈의 거리인 h 로 가정한다.



[그림 1] 응시 좌표 평면계

pupil(동공)은 cornea(각막)와 함께 실질적으로는 초의 움직임을 하므로, 벡터 축의 원점을 단구의 중심축으로 하고 눈동자가 움직이는 쪽을 x - y 축으로, 응시하는 쪽을 z 축으로 하였다. 그러므로, t 시점의 응시점은 x - y - z 축 상의 어느 한 점으로 대응되며, x - y - z 축의 방향(gaze direction)과 크기(눈-응

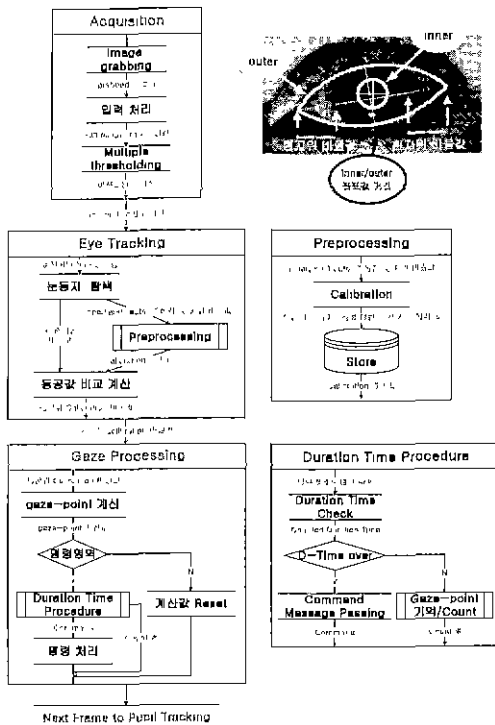
시점간 거리)로서 응시 가능한 모든 사물에 대하여 벡터 표현이 가능하다.



[그림 2] 응시점에 대한 벡터 표현

일반적으로, 연구가 움직일 수 있는 범위는 x-y축으로 제한적이지만, 응시하는 z축이 무한대의 가변적 성질을 가지므로, 가시범위 내의 모든 사물을 볼 수 있게 된다. 본 시스템에서 z축은 모니터외의 거리로 한정하여 일정거리를 유지하도록 하였다

3. 시스템 구조



[그림 3] 시스템 구성도

1) Acquisition

t 시점에 대한 한쌍 눈 영역의 영상을 카메라로 입력받아 메모리를 할당하고 처리를 하기 위해 샘플링하여 로드한다. 이때, 사용하는 모델 공간은 RGB 모델 공간으로서 (식 1)과 같이 (x, y)에서의 칼라 밝기값이 눈 영역에 해당하는 임계치(Th_{eye}) 이하인 것만을 받아들여 피부 영역을 제거한다

$$I_{eye}(x, y) \leq Th_{eye} \quad (식 1)$$

2) Preprocessing

모니터의 기준 응시점에 대한 눈동자의 (x, y) 위치와 눈 영역에 대한 MER을 이용하여 눈의 가로, 세로 크기 및 원자위의 비율값을 계산하고, 이때 모니터의 해당 기준 응시점 위치도 함께 저장한다.

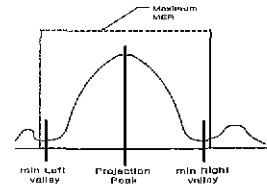
눈 요소에 대해 저장된 값들은 실행 단계에서 t 시점의 각 요소들과 비교하여 오차 보정 벡터를 결정하는 요소로 활용한다

3) Eye Tracking

영상을 스캔하면서 Th_{eye} 보다 작은 칼라값들에 대하여 8방향으로 연결이 되어 있으면, 동일한 번호로 할당하여 같은 영역으로 인식하고 연결이 끊어지면 할당할 번호를 증가시키며 입력된 눈 영상에 대하여 레이블링을 한다.

8방향으로 연결된 영역들에 대한 레이블링이 끝나면, 가장 크게 할당된 영역을 눈 영역의 MER로 받아들인다.

이와 같이 찾아진 눈의 영역에 대하여 Th_{eye} 에 해당하는 칼라값으로 X와 Y축으로 프로젝션을 시켜 눈동자의 크기 영역을 계산하고 MER 내에서 Peak를 중심으로 최소값을 가지는 왼쪽과 오른쪽 valley의 중심값을 눈동자의 중심좌표로 받아들인다.



[그림 4] MER 프로젝션

$$Central\ Coordinate(x, y) = \left[\frac{min\ Left\ valley + min\ Right\ valley}{2} \right]_{(x,y)} \quad (식 2)$$

눈동자의 중심좌표를 기준으로 calibration 피싱에서 학습된 눈동자에 대한 각 요소들을 비교하여 오차 보정 벡터를 계산한다

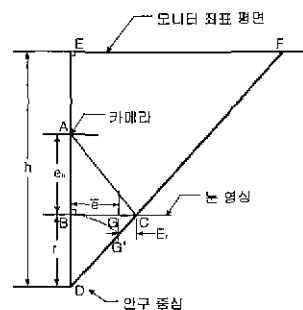
오차 보정 벡터 계산은 원형 템플릿의 원자위의 비율로써 눈동자의 이동 방향으로 기준치 벡터를 더하여 응시점에 대한 보정값을 얻는 것으로 5장에서 논한다

4) Gaze Processing

계산된 눈동자의 좌표값은 모니터 평면 좌표와 다르기 때문에 적당한 매핑 함수를 곱하여 모니터 좌표로 인터플레이션을 하는 것으로, 현재 t 시점의 눈 요소의 값들과 calibration의 비교 계산된 차이값만큼 보정하여 모니터의 응시점을 매핑한다.

모니터 좌표로 매핑한 계산된 응시점이 응시 지속 시간(gaze duration time)동안 명령영역에 머물러 있으면, 해당 명령을 수행하며, 만약 응시점이 명령영역에 있지 않다면 응시 지속 시간이 미달되었다면, 타이머 함수의 카운트를 증가시킨다

4. 보정된 응시점 매핑에 대한 수식적 표현



[그림 5] 응시점에 대한 도식

모니터의 초기 응시점 E를 보는 눈의 위치 B에서 G'으로 바뀌고, 응시점이 F로 변하였을 때 안구는 BG'을 이루는 호의 운동을 하며, 안구가 이루는 작은 $\angle BDG'$ 이 된다. 이러한 눈의 움직임을 카메라로 입력받으면, 카메라는 $\triangle ABG$ 의 영상을 찍게 되어, 보다 정확한 모니터의 응시점 F를 얻기 위해서는 선분 GC만큼의 보정이 필요하게 된다.

응시점 F를 보고 있을 때의 보정된 눈의 영상과 카메라 A를 잇는 $\triangle ABC$ 가 존재하면, 이것은 안구의 중심 D를 포함한 $\triangle DEF$ 와 대응된다

$$\triangle ABC \approx \triangle DEF$$

$$\frac{BC}{AB} = \frac{EF}{DE}, \quad \frac{\Delta e}{e_h} = \frac{\Delta d}{h} \quad (\text{식 3})$$

$$\Delta d = \frac{h}{e_h} \Delta e \quad (\text{식 4})$$

$$\bar{d} = K_h, \quad \bar{e} = K_h (\bar{e} + E_c) \quad (\text{식 5})$$

$$\left[\begin{array}{l} \Delta e = \bar{e} + E_c = \bar{e} \\ K_h = \frac{h}{e_h} \end{array} \right]$$

카메라로 입력된 영상에 대하여, 눈동자의 좌표 G인 \bar{e} 를 얻는다. 구해진 값에 이더 보정 벡터 E_c 를 더하여 $\triangle ABC$ 를 구한다. 이때, $\overline{BC} \parallel \overline{EF}$ 이며, $\angle ABC = \angle DEF = 90^\circ$ 이고, 눈의 특성상 $\angle CAB, \angle FDE$ 는 90° 보다 훨씬 작은 값이다

(식 3, 4, 5)에 의해, 응시점 F에 대한 영상의 \bar{e} 를 구할 수 있으며, 보정된 좌표를 대응되는 모니터의 좌표로 매핑하여 모니터 평면상의 응시점을 구할 수 있다.

5. 오차 보정 방안

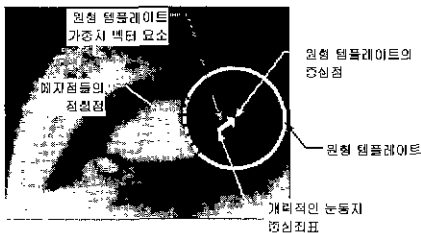
카메라로 입력받은 눈 영상에 대하여 발생하여 오차를 없애고, 보다 정확한 눈동자의 응시좌표를 얻기 위하여 응시하는 방향으로 가중치 벡터를 더해준다

가중치 벡터는 원자위 비율 및 원형 템플레이트의 중심점 계산을 이용하여 구하며, 응시방향이 커지면, 더해지는 가중치 벡터의 양도 커진다.

1) MER edge fitting으로 Circular template 중심점 벡터 보정

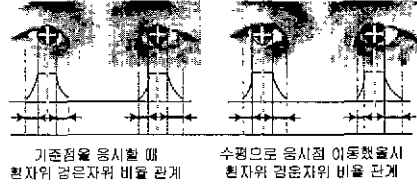
MER의 눈동자 부분에 대한 edge를 구하고 각 edge점들에 대하여 circular template을 눈동자 중심좌표로부터 상하좌우로 fitting시킨다. 원형 템플레이트의 반지름 길이는 calibration 과정에서 학습되어진다

edge점들과 fitting점의 정합이 Th 이상일 경우 fitting 작업을 마치고, 개략적 눈동자 중심좌표에서 위의 중심좌표까지의 거리를 가중치 벡터 요소로 받아들인다



[그림 6] 원형템플레이트를 이용한 보정

2) 눈동자 중심좌표의 수평축에 대한 원자위 비율

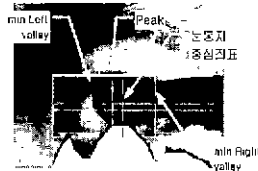


응시하는 쪽으로 김은자위의 비율은 변함이 없으나, 원자위의 비율은 작아진다. 눈동자 중심좌표 수평축의 움직임을 생각할 때 수평축상 원자위에 대한 화소수비 비율을 나타내며, 눈동자에 대하여 정규화한 곡선의 최대값을 가진 꼭지점이 눈 영역 안에서 어느 쪽으로 치우쳐 지는가를 계산하여 눈동자의 이동방향을 알아낸다. 눈 영역은 한정된 공간이므로, 중심축의 치우친 방향이 응시하는 방향이 된다.

(식 6)에서 W_{width} 는 원자위의 전체 크기를 나타내며, 원자위의 비율이 작아지는 방향으로 응시하고 있음을 나타낸다.

$$\left\{ \begin{array}{l} \frac{W_{left} < W_{right}}{W_{width}} = Left_Orientation \\ \frac{W_{left} > W_{right}}{W_{width}} = Right_Orientation \end{array} \right. \quad (\text{식 6})$$

6. 실험 및 결론



영상 크기 320x240
 MER 시작 좌표 : (59, 105)
 MER 끝 좌표 : (302, 185)
 minX (-) -> maxX : 117 ~ 215
 minY (-) -> maxY : 126 ~ 185
 눈동자 중심 좌표 : (166, 147)

입력영상에 대하여 RGB 칼라 모델 공간에서 칼라 히스토그램을 얻고, smoothing을 한 후, 눈동자에 대한 임계값을 구하여, 주어진 칼라 밝기값의 최초 mount와 전체영상에서 제일 많은 영역을 차지하는 skin에 대한 두 번째 mount의 사이에 있는 valley를 눈동자에 대한 Th_{eye} 값으로 인정한다. 이 과정에서 눈 영역을 제외한 skin을 별도의 sampling 작업없이 제거할 수 있었다.

구해진 임계값의 범위(0 ~ Th_{eye})에 존재하는 pixel들을 labeling하고 가장 큰 레이블 영역을 눈 영역의 MER로 하여, 전체 입력 영상의 탐색영역을 눈 영역으로 축소한 후, 임계값 범위에 해당하는 pixel들에 대하여 x, y축으로 projection하여 눈동자의 중심좌표를 구하였다. 향후에는 원형 템플레이트와 원자위 비율값을 이용하여 가중치 벡터를 구하고, 모니터 좌표로 매핑하는 연구가 집중적으로 필요하다.

7. 참고 문헌

- [1] Nancy R. Cleveland, R.N., B.S.N., "EYEGAZE HUMAN-COMPUTER INTERFACE FOR PEOPLE WITH DISABILITIES", 1994.
- [2] Arne John Glenstrup, Theo Engell-Nielsen, "Eye Controlled Media Present and Future State", 1995.
- [3] James E Chapman, M.D., Nancy R Cleveland, R.N., B.S.N., "THE EYEGAZE SYSTEM AND COGNITIVE RECOGNITION AND THERAPY", 1991.