

Kalman 필터를 이용한 비접촉식 응시점 추정 시스템에서의 빠른 머리 이동의 보정

Compensation for Fast Head Movements on Non-intrusive Eye Gaze Tracking System Using Kalman Filter

*김수찬, **유재하, ***남기창, ***김덕원
(Soochan Kim, Jaeha Yoo, Kichang Nam, Deokwon Kim)

Abstract - We propose an eye gaze tracking system under natural head movements. The system consists of one CCD camera and two front-surface mirrors. The mirrors rotate to follow head movements in order to keep the eye within the view of the camera. However, the mirror controller cannot guarantee the fast head movements, because the frame rate is generally 30Hz. To overcome this problem, we applied Kalman predictor to estimate next eye position from the current eye image. In the results, our system allows the subjects head to move 50cm horizontally and 40cm vertically, with the speed about 10cm/sec and 6cm/sec, respectively. And spatial gaze resolutions are about 4.5 degree and 4.5 degree, respectively, and the gaze estimation accuracy is 92% under natural head movements.

Key Words : gaze tracking, Kalman filter, non-intrusive, head movement

1. 장 서론

인간의 시각은 망막에 맺히는 상을 통해서 인식되는데, 막 중에서도 오목(fovea)이라는 좁은 영역만이 정확한 상을 볼 수 있는 곳으로 알려져 있다. 그러므로 안구의 시점으로 부터 사람의 의지나 의도, 행동, 또는 외부 자극에 대한 반응을 객관적으로 평가할 수 있는 유용한 정보를 얻을 수 있다[1]. 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 [2-5].

본 연구에서는 이상적인 응시점 추정 방법과 가장 근접한 영상을 이용한 비접촉식으로 자연스러운 머리 움직임이 가능한 시스템을 제안하고자 한다. 이를 위해 하나의 카메라와 각도 조절이 가능한 2개의 거울을 이용하였다. 얼굴에서 안구 영역만을 확대해 보기 위해 zoom in lens를 이용하기 때문에 시스템은 카메라의 view field 내에 안구 영상이 들어오도록 거울을 조절해야만 하나 빠른 머리 이동의 경우 조절을 제대로 못하는 경우가 발생한다. 이러한 문제를 해결하기 위해 Kalman predictor를 이용하였다. 안구 영상과 거울의 각도의 정보로부터 실제 응시점의 추출은 비선형 맵핑을 사용하였다.

저자 소개

- * 正 會 員: 韓京大學校 生物情報通信專門大學元 專任講師
韓京大學校 電子技術綜合研究所 研究員
- ** 正 會 員: 韓京大學校 電子工學科 助教授
韓京大學校 電子技術綜合研究所 研究員
- ***正 會 員: 延世醫大 醫學工學教室 研究講師
- ***正 會 員: 延世醫大 醫學工學教室 教授

2. 장 본론

2.1 절 시스템 구성

근적외선 CCD 카메라(TV ZOOM LENS M6Z, Hitachi, Japan)과 직경이 15cm인 2개의 거울을 이용하여 그림 1과 같이 시스템을 구성하였다. 시스템은 사용자 컴퓨터 모니터의 정면에 놓여지게 되고, 카메라는 Mirror 1과 Mirror 2로부터 반사되어온 영상을 얻게 된다. Mirror 1은 머리의 수직 이동을 보상하기 위해 y 축으로 회전할 수 있도록 하였고, Mirror 2는 수평 방향의 머리 이동을 보상하기 위해 z 축 방향으로 회전할 수 있도록 설계하였다.

주변 광원의 영향을 최소화하기 위해 시스템에 880nm에서 최고치를 갖는 근적외선 LED광원과 이 파장 대만 통과 시킬 수 있는 광학 band pass filter를 사용하였다 [2,3]. 그림 2a는 바깥과 LED가 켜졌을 때의 영상으로 각각 반사로 인해 동공의 중앙 부근에 밝게 보이는 glint가 형성된다. 그림 2b는 내부 광원이 켜졌을 때의 영상으로 적목 현상으로 인해 동공 부분만 밝게 되는 영상이다. 이러한 영상은 CCD camera로 얻어지는 형상에서 even과odd field 신호와 동기를 시켜 내부 및 외부 LED를 조절하여 얻었다.

2.2 절 Kalman filter를 이용한 머리 움직임 보정

머리의 움직임에 따라 2개의 거울은 CCD camera의 view field 내에 들어오도록 조절 되어 진다. 각각의 거울의 반사는 서로 독립적이므로 각각을 분리해서 컨트롤할 수 있다. 사용자의 머리가 왼쪽으로 움직이면 그림 1의 Mirror 2는 영상의 중앙과 동공의 중앙의 차이에 비례하여 시계 방향으로 회전시킨다. 사용자의 머리가 외로 움직이면 그림 1의 Mirror

1은 위쪽 방향으로 회전한다. 이러한 거울의 조절은 느린 머리 움직임의 경우 CCD camera의 view field가 안구 영상이 얻어지도록 조절이 가능하다. 그러나 빠른 머리의 움직임이 발생할 경우 작업은 실패할 때가 종종 있다. 왜냐하면, 영상의 sampling rate이 빠른 머리 움직임을 따라 갈만큼 충분히 빠르지 않기 때문이다. 이 문제를 풀기 위해, Kalman filter를 이용하여 현재의 안구 영상으로부터 다음 안구 영상을 추정하여 컨트롤의 성능을 향상 시켰다.

트래킹 알고리즘은 연속적인 영상 내에서 동공을 검사하는데 이용할 수 있다. 트래킹 되는 현재 상태(s_t)의 상태 $s_t = [x_t, y_t, v_{x,t}, v_{y,t}]^T$ 이고, 여기서 시간 t 에서의 동공의 중심은 각각 x_t 와 y_t 이고, 가속도는 $v_{x,t}$ 와 $v_{y,t}$ 는 이다. $t+1$ 에서의 상태 벡터는 $s_{t+1} = \Phi s_t + w_t$ 이고, 여기서 프로세스의 천이 행렬은 Φ 이고, zero mean과 Gaussian random process로서 상태의 perturbation을 모델링 한 것은 w_t 로 나타내었다. 프레임간의 움직임은 선형적이라고 가능하면, 상태 행렬 s_{t+1} 는 다음 방정식을 만족한다.

$$s_{t+1} = \Phi s_t, \quad \begin{bmatrix} x_{t+1} \\ y_{t+1} \\ v_{x,t+1} \\ v_{y,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ v_{x,t} \\ v_{y,t} \end{bmatrix}$$

측정 모델에 대한 filter는 $z_t = H s_t + \mu_t$ 이다.

H 는 상태 벡터와 측정 벡터간의 부잡음 상태에서의 관계이며 u_t 는 불확실성을 의미한다. 단순화시키기 위해 불확실성(u_t)을 0으로 하면,

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Kalman filter 알고리즘의 첫 단계는 다음 상태와 covariance를 추정하는 것이다.

$$s_{t+1}^- = \Phi s_t, \quad \Sigma_{t+1}^- = \Phi \Sigma_t \Phi^T + Q$$

여기서, Q 는 시스템의 error covariance이고, Σ_t 와 Σ_{t+1}^- 는 각각 현재 error covariance와 추정된 covariance이다. 다음 단계로 Kalman filter gain(K_{t+1})를 추정하고 posteriori state estimate s_{t+1} 을 얻는다.

$$K_{t+1} = \Sigma_{t+1}^- H^T (H \Sigma_{t+1}^- H^T + R)^{-1}, \quad s_{t+1} = s_{t+1}^- + K_{t+1} (z_{t+1} - H s_{t+1}^-)$$

마지막으로 posteriori error covariance를 갱신하고, 새로운 priori estimate를 예측한다.

$$\Sigma_{t+1} = (I - K_{t+1} H) \Sigma_{t+1}^-$$

위 과정을 반복한다. 시스템에 Kalman filter를 적용시키기 위해서 초기상태를 설정해야만 한다. 두 개의 frame 데이

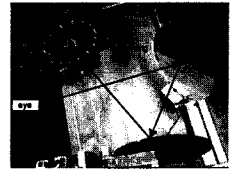
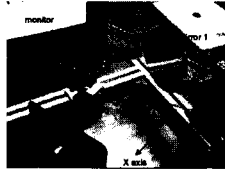


그림 1. 카메라 한대와 2개의 거울, 그리고 광원으로 구성된 안구 응시점 추정 시스템



그림 2. 두 개의 거울을 통해서 얻어진 이미지 (a) 외부 광원이 켜졌을 때 (b) 내부 광원이 켜졌을 때

터 정보로부터 초기 상태 벡터 s_0 를 얻어낼 수 있다.

$$s_0 = [x_{t+1}, y_{t+1}, x_{t+1} - x_t, y_{t+1} - y_t]^T$$

초기 error covariance matrix Σ_0 를 4x4 zeros matrix로 설정하자. 측정의 error covariance matrix(R)와 시스템의 error covariance matrix를 동공 운동에 기초하여 다음과 같이 설정하였고, 이를 stationary라고 가정하였다.

$$Q = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \quad R = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

2.3 절 신경망을 이용한 응시점 맵핑

컴퓨터 모니터 상에서 가로 6, 세로 4로 24개의 영역으로 나누었다. 각 응시점 마다 각기 다른 머리 위치로 하여 20개의 데이터 셋을 오른쪽 눈에서 모았다. 자유로운 머리 이동은 수평, 수직 각각 50cm와 40cm 정도로 제한하였다. 왜냐하면 모니터 앞의 머리 이동은 그 정도면 충분하기 때문이다.

응시점 추정을 위해 머리의 위치뿐만 아니라 동공과 관련된 보다 많은 정보가 필요하다. 응시점 추정에 활용한 동공 영상으로부터 얻어낸 정보들은 x , y , r , m_{mir} 와 m_{pan} 이다. x 와 y 는 동공과 glist사이의 x 축과 y 축에 대한 변위이다. m_{mir} 와 m_{pan} 은 Mirror 1과 Mirror 2의 회전 각도에 대한 정보로써 머리 이동에 대한 정보를 반영한다. 얻어진 5개의 특징점 중 표준 편차에서 많이 벗어나는 것은 제거하였다.

비선형 맵핑 함수로 Universal function approximator로써 개발된 generalized regression neural networks (GRNN)은 데이터 수가 많을 경우 노이즈에도 강한 특성을 보인다[5]. 따라서 본 연구에서는 맵핑 함수로써 GRNN을 사용하였고,

이웃 응시점 간의 오차를 줄이기 위해 아래 식과 같이 hierarchical GRNN (H-GRNN) classifier 를 이용하였다.

$$g = [\Delta x, \Delta y, r, m_{tilt}, m_{pan}]$$

$$1^{st} \text{ pass: } G_{point} = f_{GRNN}(g), \quad 2^{nd} \text{ pass: } Gaze_{point} = f_{HGRNN}^{(G_{point})}(g)$$

f_{HGRNN} : mapping function which is trained with the whole training data

$f_{HGRNN}^{(G_{point})}$: mapping function which is trained with the neighborhood data of G_{point}

H-GRNN은 단지 1차 맵핑 함수에서 얻어진 응시점 결과에서 그것의 이웃점들에 의해서 학습된 데이터만을 이용하여 만든 2차 맵핑 함수를 다시 수행하였다. Mapping function은 입력 인자인 x, y, m_{tilt} and m_{pan} 을 정규화한 이후에 활용하기 때문에 사용자 변경에 따른 재보정이 필요 없다.

3. 장 결과

3.1 절 머리 움직임 아래에서의 안구 운동 추적

시스템은 머리가 수평으로 50cm, 수직으로 40cm 내에서 자연스럽게 머리를 움직임을 보정 가능하였다. 시스템은 모니터 스크린 앞의 70~80cm 앞에서 앉은 상태에서 충분한 머리 움직임으로 보여 진다.

그림 3은 실제 동공의 위치와 Kalman filter로 예측한 위치를 보여주고 있다. 동공의 움직임이 느린 경우 두 가지 방법에는 거의 차이가 없었으나, 머리가 빠르게 움직일 때 예측된 동공 위치는 CCD 카메라의 광축 내에 동공을 유지 시키는데 충분히 유용하였다. 그림 3에서 보여지는 바와 같이 카메라의 시야를 벗어나기 전에 예측된 동공 위치를 이용하여 2개의 거울의 조절을 통하여 보다 빠른 머리 운동 하에서도 응시점 추정이 가능하였다.

3.2 절 머리 움직임 아래에서의 응시점 추정

표 1은 GRNN과 H-GRNN에 의한 응시점 추정 결과의 정확성을 보여 주고 있다. 이때 머리는 수평으로 50cm, 수직으로 40 cm 로 자연스럽게 움직이면서 응시점을 추정하였다. GRNN에 의한 응시점의 정확도는 85%였고, H-GRNN을 이용할 경우 92%였다. H-GRNN이 GRNN에 의한 것보다 높았다.

4. 장 결론 및 토의

대부분의 응시점 추정 시스템의 가장 큰 문제는 자연스러운 머리 움직임 아래에서 사용자의 응시점을 얻어내는데 있다. 본 연구에서는 단지 하나의 카메라와 2개의

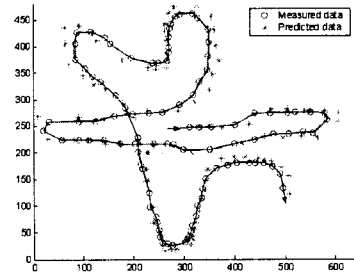


그림 3. 실제 동공의 위치와 예상된 동공의 위치의 이동 경로(o: 실제 측정된 데이터, *: Kalman filter를 이용하여 현재와 과거의 측정 결과를 바탕으로 추정된 데이터)

front-surface mirrors, 자체 광원을 이용하여 자연스러운 머리 움직임 하에서의 응시점 추정 시스템을 제안하였다. 머리 이동에 의한 안구 영상이 손실되는 것을 막기 위해 Kalman predictor를 이용하여 응시점 추정 성능을 향상시킬 수 있고, 이것은 효과적으로 머리 움직임으로 인한 안구 영상이 카메라 시각에서 벗어나는 것을 막을 수 있었다.

H-GRNN을 이용하여 응시점 분석 결과의 정확성을 GRNN보다 상당히 높일 수 있었다. 또한 머리와 얼굴이 카메라의 정면에서 약 ± 10 도 정도 벗어난 상태에서도 응시점 추정이 가능하였고, 입력 파라미터로 정규화된 값을 이용하므로 사용자마다 재 보정하는 번거로운 작업을 줄일 수도 있었다.

제안한 시스템은 휴먼-컴퓨터 인터페이스나 상호 작용, 진보된 그래픽 출력 장치 등과 같은 많은 분야에 적용할 수 있을 것으로 기대된다.

참 고 문 헌

- [1] Bekkering H. and Neggers S. F. W. (2002): 'Visual search is modulated by action intentions,' *Psychological Science*, 13, pp. 370-374
- [2] Ji Q. and Yang X. J. (2002): 'Real-time eye, gaze, and face pose tracking for monitoring driver vigilance', *Real-Time Imaging*, 8, pp. 357-377
- [3] Kim S.C., Sked M. and Ji Q. (2004): 'Non-intrusive eye gaze tracking under natural head movements', *Proc. 26th Int. EMBS, San Francisco, USA, 2004*, pp. 291
- [4] Ebisawa Y. (1998): 'Improved video-based eye-gaze detection method', *Instrumentation and Measurement, IEEE Transactions on*, vol. 47, pp. 948-955
- [5] Specht D. F. (1991): 'A General Regression Neural Network', *IEEE Transactions on Neural Networks*, 2, pp. 568-576