

HCI를 위한 눈 응시방향 추적 시스템 개발

김도형, 유동현, 정명진
한국과학기술원 전기 및 전자공학과

Development of the eye-gaze tracking system for HCI

Do-hyoung Kim, Dong-hyun Yoo, and Myung-Jin Chung
Dept. of Electrical Engineering, KAIST

Abstract - 정보화 시대인 요즘 컴퓨터의 사용이 갈수록 증가하면서 사용의 편이를 높이기 위하여 사람과 컴퓨터간의 인터페이스에 대한 연구가 많아지고 있다. 새로운 형태의 인터페이스는 자연적인 사람의 행동 양식을 분석하여 그 정보로부터 사용자의 의도를 추출해내는 의도 파악(Intention reading)에 기반한 형태이다.

우리가 관심을 두고 있는 시스템은 사용자가 바라보고 있는 방향을 추적하여 그 방향을 이용하여 컴퓨터를 조작하는 시스템이다. 응시방향을 알아내기 위해 카메라를 이용하여 눈의 움직임을 관찰한다. 그리고 얼굴의 움직임이 있는 일반적인 경우 정확한 눈의 응시 방향을 결정하기 위해서 눈에 대한 정보만을 이용하지 않고 얼굴이 향하고 있는 방향을 추적하여 더 일반적이고 정확한 눈의 응시방향을 결정하고자 한다.

본 논문에서는 기존의 눈 응시 방향 추적 방법들에 대하여 소개하고 우리가 개발하려고 하는 시스템에 대해 언급하고자 한다.

1. 서 론

사람의 눈은 주위를 관측하기 위한 입력 기관이면서 출력 기관이라고 할 수 있는데, 이때 출력 기관이라는 의미는 우리가 무엇에 초점을 맞추고 있는지를 지시한다는 것이다. 이는 생리학적인 측면에서 살펴볼 수 있다. 인간은 살펴보고자 하는 세밀한 부분을 망막의 한 가운데에 놓고 의도하는 부분과 그 주위를 바라보게 되므로, 물리적인 기관이 표현하는 이러한 방향은 사람들이 의도하는 방향과 같다고 볼 수 있다.

이와 같은 사실을 이용하여 컴퓨터의 새로운 입력 장치를 개발하려는 시도가 있다. 컴퓨터의 개발 이후, 사용자와의 상호 통신의 발달은 컴퓨터로부터 사용자에게로 가는 의사통신에 대한 것을 중심이었다. 이것은 데이터들의 그래픽 표현, 윈도우 시스템 개발, 데이터를 표현하기 위한 사운드의 사용 등의 기능으로 표현되어 왔다. 반면에, 사용자에게서 컴퓨터로 향하는 통신은 여전히 손으로만 동작 될 수 있는 키보드, 조이스틱, 마우스에 한하고 있다. 그러므로, 사용자의 응시 방향을 추적하는 방법에 의해 알아낼 수 있는 사용자가 무엇을 바라보고 있는지에 대한 정보를 사용함으로써 사용자로부터 컴퓨터로 향하는 새로운 입력 장치를 개발할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 우리가 개발하려고 하는 눈 응시방향 추적 시스템에 대해 설명하고 3장에서는 기존의 눈 응시방향 추적 방법에 대해 소개하고 4장에서는 얼굴 방향 추정 방법에 대해 언급하고 5장에서 결론을 내고 끝맺는다.

2. 눈 응시방향 추적 시스템

우리가 개발하고자 하는 눈 응시방향 추적 시스템의 대략적인 구성은 그림1과 같다. 두 대의 카메라를 이용하는 한 하나는 눈 응시 부분만을 추적해 내는 망원렌즈를 이용한 카메라이고 다른 하나는 얼굴 전체를 넓은 각

도로 바라보는 카메라이다. 두 대의 카메라 모두 컴퓨터의 모니터 근처에서 사용자를 관찰하고 사용자는 어떠한 장치도 부착하지 않는 방법(Nonintrusive method)이다.

두 대의 카메라를 사용하는 이유는 눈의 응시 방향을 추적하기 위해서 가능한 눈을 크게 입력받을 필요가 있기 때문이다. 모니터 내에서의 눈동자의 움직임이 크지 않기 때문에 분해능을 높이기 위해 가능한 눈을 크게 입력받는다. 그런데 한 대의 카메라만으로 눈을 확대하여 바라보면서 추적하는 것은 어려운 일이다. 따라서 얼굴 전체를 관찰하는 카메라가 하나 더 필요하게 된다. 이 카메라는 얼굴을 관찰하면서 눈의 위치를 찾아서 눈 부분만을 바라보는 카메라에게 알려준다. 그리고 얼굴을 보는 카메라는 얼굴이 정면을 향하고 있지 않은 경우 얼굴이 바라보고 있는 방향을 고려하기 위한 보조 수단으로 사용된다.

3. 눈 응시방향 추적 기술

3.1 카메라를 이용한 기술들

3.1.1. 림버스(Limbus) 추적 방법

림버스(Limbus)는 눈에서 흰자위와 어두운 홍채사이의 경계를 나타내는 것이다. 이 경계는 일반적으로 눈의 밝은 부분(흰자위)과 어두운 부분(홍채, 동공)을 구분하는 것이므로, 영상 처리할 때 간단하게 이 경계를 쉽게 찾아내고 추적할 수 있다. 이 기술은 림버스(Limbus)의 모양이나 위치를 영상 처리과정을 통해 획득하고 이용한다. 그리고, 이 기술은 머리의 움직임에 영향을 받으므로 머리가 상당히 오랫동안 고정되어 있거나, 사용자의 머리에 고정되는 기구를 이용한다.[4]

하지만, 눈꺼풀에 의해 림버스(Limbus)의 위, 아래

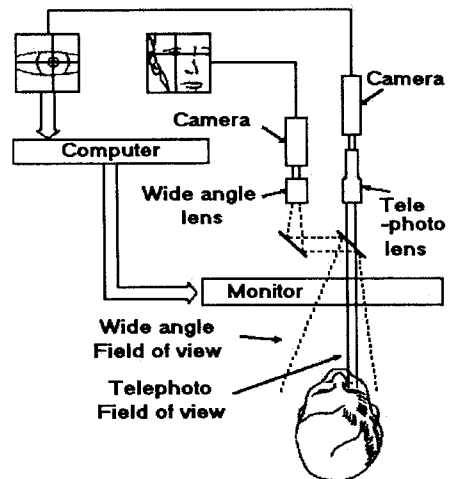


그림 1. 눈 응시방향 추적 시스템

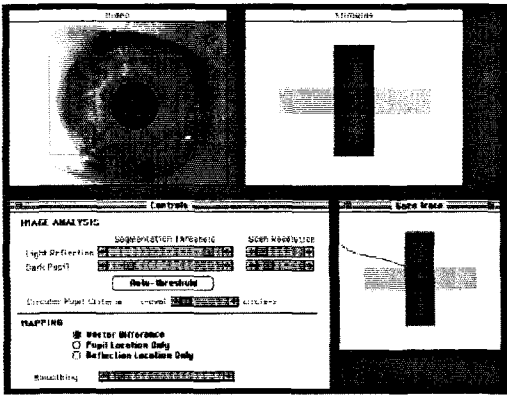


그림 2. 동공 추적 방법을 응용한 결과

부분이 대부분 가려지기 때문에 림버스(Limbus)를 이용한 방법은 오직 가로방향의 이동을 추적하는데 사용될 때 적절하다는 단점이 있다.

3.1.2. 동공 추적 방법

림버스(Limbus) 추적 방법과 비슷한 방법으로 동공과 홍채사이의 경계부분을 이용하여 눈의 응시방향을 추적하는 방법이다 이 방법도 완전히 머리의 움직임에 연관되어 동작하므로 머리에 고정할 기구부나 머리의 움직임이 오랫동안 정지할 필요가 있다.

림버스(Limbus) 추적 방법에 비하여 이 방법은 다음과 같은 장점들을 가지고 있다. 우선, 동공은 눈꺼풀에 의해 덜 가려지므로 세로방향의 이동에 대한 추적도가 높다. 그리고, 동공의 가장자리는 림버스(Limbus)보다 더 선명하게 나타나므로 더 정밀하게 움직임을 측정할 수 있다.[6][7]

하지만, 동공의 가장자리는 일반적으로 홍채와 동공의 색깔의 차이가 적어서 림버스(Limbus)보다 구분하기가 힘들므로 가장자리를 찾아내는 방법이 매우 어렵다.

그림2는 동공의 추적 방법과 3.1.3에서 소개할 발광 소자를 이용하여 각막에서 발생된 부분과의 관계에 의해 눈 응시 방향을 검출하고 추적하는 결과를 보여주고 있다.

3.1.3. 각막, 망막 반사의 상호관계에 의한 방법

사용자의 눈에 빛(적외선)을 쏘었을 때, 눈의 망막에 의해 빛이 반사되어 동공이 밝아지는 현상이 있는데, 이러한 현상을 동공이 밝아지는 현상(Bright-Eye)이라고 하며, 각막에서 반사되는 부분이 있는데 그 부분을 반짝거리는 부분(Glint)이라고 하는데, 이러한 두 부분의 위치 관계를 이용하면 눈의 응시 방향을 추적할 수 있다.

적외선을 눈에 쏘이면서 적외선에 민감한 카메라를 이용하여 이미지를 기록하면, 앞에서 언급한 동공이 밝아지는 현상(Bright-Eye)과 각막에 의해 반사되는 반짝거리는 부분(Glint)이 검출되고, 이때 눈이 상하좌우로 움직이면 밝아진 동공의 변화하는 중심점과 반짝거리는 부분(Glint) 사이의 상대적인 위치에 의해 응시 방향을 계산할 수 있다.[3]

그림2는 동공이 밝아지는 현상(Bright-Eye)에 의해 쉽게 검출할 수 있는 동공의 중심부분과 각막에 의한 반짝거리는 부분(Glint)과의 관계를 간단하게 표현한 것이다.

이러한 기술은 외부의 카메라를 이용할 경우에 눈의 이미지가 머리의 측면방향의 이동에 대해서는 초점이 맞지 않는다는 것과 추적할 수 있는 응시 방향의 범위가 제한적이라는 점이 있고, 기하학적인 교정을 수행하기 위한 복잡한 계산들이 필요하다는 것들이 문제점이 될 수 있다. 하지만, 이러한 문제점들은 좀더 향상된 소프트웨어의 교정 알고리즘이나 더 효과적으로 추적하는 카

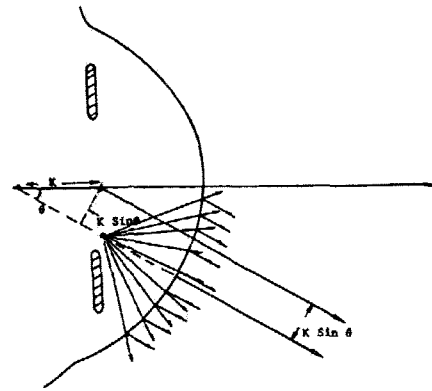


그림 3. 각막, 망막 반사의 상호관계

메라를 이용함으로써 해결할 수 있을 것이다.

3.2. 피부의 전기적 신호를 이용한 기술

EOG(Electro-oculography) 추적 방법은 눈이 회전함에 따라 정적인 전기적 영역이 존재한다는 사실을 감안한 것이다. 이러한 정전 영역의 발생에 의한 매우 작은 전위 차이들을 눈 주위의 피부에 전극들을 붙여서 기록함으로써 눈의 위치를 검출할 수 있다.[8]

이러한 기술의 문제점은 사용자에게 전극을 붙여야 한다는 점 때문에 매일 사용하는 목적으로는 적당하지 않다.

3.3. 콘택트 렌즈를 이용한 기술들

콘택트 렌즈를 사용하는 방법은 크게 다음과 같은 두 가지 방법이 있다.

첫째는 렌즈의 표면에 거울과 같이 반사율이 높은 것들을 한 개 이상 포함시켜서 빛(광선)의 반사를 이용하여 눈의 위치를 계산하는 방법이 있고, 두 번째는 렌즈에 매우 작은 감응 코일을 박아서 사용자의 머리 주위에 높은 주파수의 전기-자기장을 형성시킴으로써 렌즈의 정확한 위치를 기록할 수 있다.

이러한 기술도 매일 사용하기에는 높은 주파수의 전기-자기장이 건강에 미치는 영향이 아직 명확히 규명되지 않아 적합하지 않다. 또한 대부분의 사람들이 일상적인 의사소통을 하기 위해 콘택트 렌즈와 같은 것을 준비해야 된다는 점들은 그리 좋은 점이 되지 못한다.

4. 얼굴 방향 추정 기술

얼굴 방향 추정 기술은 사람이 어느 방향으로 얼굴을 향하고 있는지를 알아내는 것이다. 얼굴 인식 시스템에서 이러한 연구가 많이 있었다. 입력 영상의 얼굴과 데이터베이스의 얼굴을 비교하여 같은 얼굴을 찾을 때 입력 영상의 얼굴이 데이터베이스에 저장되어있는 영상과 다른 자세를 하고 있다면 같은 얼굴이라도 인식하지 못하는 경우도 발생한다. 따라서 강인한 얼굴 인식 시스템을 구현하기 위해서 입력 영상의 얼굴 방향을 추정하여 그 정보를 이용해 데이터베이스에 저장된 영상의 얼굴 방향과 일치하도록 보정을 해야 할 필요가 있다. 얼굴 인식 시스템을 연구하는 분야에서는 얼굴 방향 추정 기술을 보조 수단으로 사용하기 위해서 연구하였다. 그러나 최근 HCI(Human Computer Interaction) 분야에서 컴퓨터 인터페이스의 하나로 얼굴의 방향을 이용하는 연구가 진행되고 있다. 사람의 관심을 나타내는 자연적인 행동인 얼굴의 움직임 관찰하여 컴퓨터의 입력으로 이용하는 것이 목적이다.

기존의 연구는 사용자의 몸에 부착하는 특별한 장치의 유무에 따라 부착하는 방법과 그렇지 않은 방법으로 나

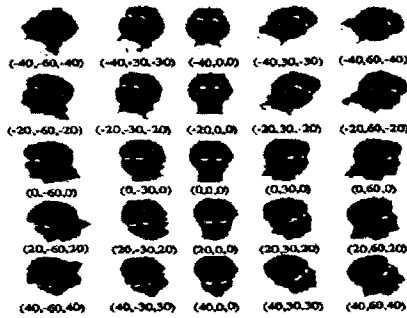


그림 4. 데이터베이스에 저장된 영상들

늘 수 있다. 그리고 이용하는 정보에 따라 영상 기반 방법과 모델 기반 방법으로 나눌 수 있다.

4.1 몸에 부착하는 장치의 유무에 따른 분류

4.1.1 장치를 부착하는 방법

이 방법은 사용자의 머리에 특정한 하드웨어 장치나 인공의 표식점을 부착하여 얼굴의 움직임을 측정하는 방법이다. 얼굴의 움직임을 측정하기 위한 장치로는 자이로스코프(Gyroscope)나 지자기 센서, 마그네틱 센서 등이 있다. 그러나, 이와 같이 머리에 장치를 부착하는 방식은 사용자에게 거부감을 느끼게 할 수 있다. 특히 장애인과 같이 정상적인 신체 조건이 아닌 경우에는 적합하지 않은 방법이다. 그래서 최근에는 장치를 부착하지 않는 방법에 대한 연구가 많아지고 있다.

4.1.2 장치를 부착하지 않는 방법

이 방법은 사용자에게 특정한 하드웨어 장치를 부착하지 않고 외부에서 머리의 움직임을 관찰하여 정보를 얻는 방법이다. 일반적으로 카메라를 이용하여 머리의 영상을 입력받아서 그로부터 정보를 얻어내는 방법이 많이 연구되었고 좀 더 유용한 정보를 습득하기 위해서 스테레오 카메라, 칼라 카메라, 거리 영상 등을 이용한 방법들이 연구되고 있다.

4.2 이용하는 정보에 따른 분류

4.2.1 영상을 기반으로 하는 방법

입력 영상과 데이터베이스에 저장되어 있는 영상들과의 상관 함수를 구하여 얼굴의 방향을 결정하는 방법이다. 이 때 얼굴의 데이터베이스를 구축하는 방법은 모든 가능한 자세와 조명 조건에 대해 다양한 영상을 저장하거나 여러 자세의 얼굴을 얻기 위해서 3D 머리 모델을 투영시킨다(그림4). (9) 그러면 다수의 영상들과 입력 영상을 비교하여 입력 영상에서의 얼굴이 어느 방향을 향하고 있는지를 찾아낸다. 그러나 이러한 방법을 이용하여 정확도를 높이기 위해서는 많은 영상을 데이터베이스에 저장해야 하므로 데이터베이스의 용량이 엄청나게 커진다. 그리고 데이터베이스에 저장되어 있지 않은 다른 사람 또는 다른 조명에 대해 적용하기 어렵다.

4.2.2 모델을 기반으로 하는 방법

사전에 알고 있는 머리에 대한 3차원 모델 정보를 이용하여 방향을 추정한다. 이 때 입력 영상에서 3차원 모델과 매칭하기 위한 3개 이상의 점을 구할 수 있어야 한다. 그림5는 얼굴에서 찾아내는 특징점들의 한 예이다.

R. Stiefelbogen 등은 눈, 코, 입의 특징점을 영상으로부터 추출하고 그 특징점을 추적하면서 얼굴의 방향을 구하는 실시간 시스템을 구현하였다. [10] 이들은 칼라 카메라로부터 입력받은 영상에서 통계적인 칼라 모델을 사용하여 얼굴을 찾아낸다. 그리고 얼굴 안에서 기하학적인 제한 조건을 만족하는 두 개의 어두운 영역을 찾는

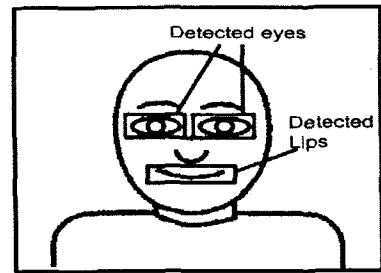


그림 5. 얼굴의 특징점들(눈, 코)

다. 이 때 기하학적인 정보는 얼굴과 눈에 대한 정보이다. 입술과 코를 찾는다. 그리고 특징점들을 계속해서 추적하면서 얼굴의 방향을 추정한다.

P. Ballard와 G. C. Stockman은 얼굴에 고정된 조명을 비추고 눈과 코에서의 반사를 이용하여 입력 영상에서 3개의 특징점을 추출하고 그 세 개의 특징점을 잇는 삼각형을 이용해서 얼굴의 방향을 결정하였다. [11] 이러한 방법의 단점은 영상에서 눈, 코, 입 등의 특징점을 정확히 찾는 과정이 중요하고 결과의 정확도에 밀접하게 연관된다.

5. 결 론

지금까지 사람의 눈 응시방향 추적 시스템에 대한 연구를 살펴보았다. 이러한 시스템은 손을 사용할 수 없는 장애인이나 노약자들이 컴퓨터를 조작할 수 있도록 할 수 있으며 일반인들도 기존의 입력 장치보다 편리하게 사용할 수 있게 될 것이다.

(참 고 문 헌)

- [1] Arne J. G., Theo E., "Eye controlled media: present and future state", Thesis for a Bachelor's Degree University of Copenhagen, 1995.
- [2] Karl F. A., "ViewPoint eye tracker", Arrington Research Report USA, 1997.
- [3] John M., Richard M., James L. P., "Remote measurement of eye direction allowing subject motion over one cubic foot of space", IEEE Trans. Biomed. Eng., vol. BME-21, pp. 309-317, 1974.
- [4] Baloh RW., Yue Q., Demer JL., "The linear vestibuloocular reflex in normal subjects and patients with vestibular and cerebellar lesions", J. Vestibular Research - Equilibrium & Orientation, 5: (5), 349-361, 1995.
- [5] Ivins JP., Porrill J., Frisby JP., "Deformable model of the human iris for measuring ocular torsion from video images", IEE Proceedings - Vision Image and Signal Processing, 145: (3), 213-220, 1998.
- [6] Ivins JP., Porrill J., "A deformable model of the human iris for measuring small three-dimensional eye movements", Machine Vision and Applications, 11: (1), 42-51, 1998.
- [7] Solomon D., Cohen B., "Stabilization of gaze during circular locomotion in light .1. Compensatory head and eye nystagmus in the running monkey", J. Neuro physiology, 67: (5), 1146-1157, May 1992.
- [8] A. Tukamoto, C. Lee, and S. Tsuji, "Detection and tracking of human face with synthesized templates", Proc. ACCV'93, pp.183-186, 1993.
- [9] R. Stiefelbogen, J. Yang, and A. Waibel, "A model-based gaze tracking system", Tech. report, Carnegie Mellon University.
- [10] P. Ballard, G. C. Stockman, "Controlling a computer via facial aspect", IEEE Transaction on system, man, and cybernetics, vol. 25, No. 4, 1995.