

## 실시간 계층적 시선 식별

박호식, 박동희, 설중보\*, 손동주\*, 나상동\*\*, 배철수  
 관동대학교, (주)KT, 조선대학교  
 mediana@netsgo.com

### Real-Time Hierarchical Gaze Discrimination

Hosik Park, Donghee Park, Jeungbo Seol\*, Dongjoo Son\*, Sangdong Ra\*\*, Cheolsoo Bae  
 Kwandong University, Korea Telecom\*, Chosun University\*\*

#### 요 약

본 논문에서는 실시간 시선 식별 시스템을 위한 능동적 적외선 조명을 기반으로 한 컴퓨터 비전 시스템에 대하여 제안하고자 한다. 현존하는 대다수의 시선 식별 기술은 고정된 얼굴에서만 제대로 동작하고 개개인에 대한 교정 절차를 필요로 한다. 그러므로 본 논문에서는 교정 작업 없이 얼굴 움직임이 있는 경우에도 견실하고 정확하게 동작하는 시선 식별 시스템을 제안하고 있다. 제안된 시스템은 GRNN 을 이용한 동공 매개 변수로부터 스크린 좌표로 할당하는 시선 교정 방법을 사용하여 계층적인 구조를 구성되었다. GRNN 을 사용 함으로서, 시선 할당에 분석적 기능이나 얼굴 움직임에 대한 계산이 필요 없을 뿐 아니라 다른 개개인에 대하여 학습과정에서 할당 기능을 일반화할 수 있었다. 시선 예측 정확도를 개선하고자 계층적 식별을 제안함으로써 오분류율을 줄일 수 있었다. 공간 시선 해상도는 화면에서 1m 정도 떨어진 거리에서 수평으로 10cm, 수직으로 약 13cm 정도인, 즉 수평으로 8 도정도 수직으로 대략 5도 정도이었다. 실험 결과, 1차 시선 식별기보다 계층적 시선 식별기가 약 9% 정도 정확성이 향상되어 93%를 나타냄으로써 제안된 시스템의 유효성을 증명하였다.

#### I. 서 론

최근 컴퓨터 기술의 향상과 다양한 멀티미디어 장치들의 발달과 함께 컴퓨터를 편리하게 사용하려는 욕구 또한 높아져 휴먼-컴퓨터 인터페이스에 대한 관심이 크게 부각되고 있다. 키보드나 마우스와 같은 기존의 입력 장치는 사람의 감정이나 의도와 같은 감성 정보를 컴퓨터에 입력시키기 어렵고, 사용자의 직접적인 입력이 없으면 컴퓨터 스스로 어떤 일을 하거나 판단을 내릴 수 없다. 이에 따라 기존의 수동적인 입력 장치를 대신하는 제스처나 얼굴 표정, 음성, 시선 등을 이용한 멀티 모달(Multi-modal) 사용자 인터페이스에 대한 연구가 진행 중에 있다.[3] 그 중 마우스의 기능을 대신하여 사람의 시선 방향을 활용하는 방식은 컴퓨터가 사람의 의도를 인지한다는 측면에서 인간과 컴퓨터간의 새로운 의사 소통 방식의 하나로 간주되고 있다.[4]

일반적으로 시선 식별 시스템은 시선을 식별하기 위해 얼굴의 움직임과 눈동자의 움직임을 이용한다. 얼굴의 움직임만을 이용할 경우 얼굴의 위치에 따라 시선의 위치를 결정하게 되는데 미세한 시선 변화를 감지하기에는 어렵게 된다. 그리고, 눈동자의 움직임을 이용한 경우는 시선과 홍채, 동공의 기하학적 특성에 대한 관계에 근거하여 시선을 추정하게 된다. 동공과 각막 반사에 의한 글린트 사이의 공간적 위치 특성으로 시선을 식별하고, 동공의 위치와 형태, 왜곡 등의 정보로 시선을 실시간으로 식별한다. 이 때 신경망을 통하

여 눈과 명암에 대해 정규화 된 동공을 포함하는 작은 창을 추출하여 시선의 좌표를 결정하는데 사용한다. 현재까지 연구된 눈에 근거한 시선 추정 방식의 가장 큰 공통점은 각각의 글린트와 동공 사이의 상대적인 위치에 근거한다는 것이다.[4,5,6,7,8] 이러한 방식은 머리를 정지시킨 후, 기준점으로서 글린트를 사용하고, 동공의 중심으로부터 글린트 까지를 나타낸 벡터로 시선 방향으로 나타내었는데 얼굴의 작은 움직임에도 잘못된 결과를 얻을 수 있었다. 이러한 점은 인간과 컴퓨터 상호 작용에 있어 중요한 장애요소이다. 또한, 현존하는 눈과 시선 식별 시스템의 또 다른 문제는 각각의 사용자에 대하여 교정 처리를 필요로 한다는 것이다.

본 논문에서는 기존의 연구들이 가지고 있는 문제점들을 해결하고, 모니터 앞에 앉은 사용자의 자연스러운 얼굴 움직임에 의해 모니터상의 시선 위치를 파악할 수 있는 방법에 대하여 연구하였다. 이를 위하여 입력된 얼굴 영상으로부터 얼굴 영역 및 얼굴내의 눈과 동공의 위치를 추출하여, 얼굴의 전체적인 방향을 식별하였고, 추출된 동공과 글린트의 변화량, 동공의 타원 정도, 기술기 그리고 글린트의 좌표를 매개변수로 추출하였다. 추출된 매개변수를 가지고 GRNN (Generalized Regression Neural Network)을 계층적 구조를 가진 식별 시스템을 구성함으로써 시선 교정 없이 견실하고 정확하게 동작하는 개선된 시선 식별 시스템을 제안하고자 한다.