

# 지능형 공간에서 청각장애인의 시선 방향 검출

오 영 준<sup>†</sup> · 홍 광 진<sup>\*\*</sup> · 김 종 인<sup>†</sup> · 정 기 철<sup>\*\*\*</sup>

## 요 약

인간-컴퓨터 상호 작용은 인간공학과 정보기술을 융합하여 인간과 컴퓨터와의 상호 작용 기술을 연구하는 학문이고, 그 중에서도 지능형 공간은 정보화 사회에서 소외될 수 있는 장애인에게 더욱 효과적인 사용 환경을 제공할 수 있는 중요한 연구 분야이다. 장애인을 위한 지능형 공간에서의 정보 지원 방법은 장애 유형에 따라 달라지는데, 본 논문은 정보 지원 대상을 청각 장애인으로 한정한다. 청각 장애인에게 직접 접촉을 통해 정보 제공 위치를 인지시키는 방법을 제외하면, 시선이 향하고 있는 곳에 정보를 표시하는 것이 가장 효율적인 정보 제공 방법이기 때문에, 시선 방향 검출 방법은 필수적이다. 우리는 이처럼 청각 장애인에게 실내 생활 지원 서비스를 제공하기 위해 반드시 필요한 시선 방향 검출 방법을 제안한다. 제안된 방법은 다시점 영상(Multi-view Camera Image) 내에서 검출된 사용자 영역 정보를 이용하여, 시점 별 수평/수직 방향 시선각 후보를 생성하고 후보 간의 크기 비교를 통해 사용자의 시선 방향을 계산한다. 실험 결과에서, 제안된 방법은 높은 시선 방향 검출 성능을 보이며, 장애인을 위한 시나리오를 수행할 수 있는 가능성을 보였다.

키워드 : 장애인, 시선 방향 검출, 지능형 공간, 히스토그램, 칼라 영상 분할

## Detection of Gaze Direction for the Hearing-impaired in the Intelligent Space

Young-Joon Oh<sup>†</sup> · Kwangjin Hong<sup>\*\*</sup> · Jong-in Kim<sup>†</sup> · Keechul Jung<sup>\*\*\*</sup>

## ABSTRACT

The Human-Computer Interaction(HCI) is a study of the method for interaction between human and computers that merges the ergonomics and the information technology. The intelligent space, which is a part of the HCI, is an important area to provide effective user interface for the disabled, who are alienated from the information-oriented society. In the intelligent space for the disabled, the method supporting information depends on types of disability. In this paper, we only support the hearing-impaired. It is material to the gaze direction detection method because it is very efficient information provide method to present information on gazing direction point, except for the information provide location perception method through directly contact with the hearing-impaired. We proposed the gaze direction detection method must be necessary in order to provide the residence life application to the hearing-impaired like this. The proposed method detects the region of the user from multi-view camera images, generates candidates for directions of gaze for horizontal and vertical from each camera, and calculates the gaze direction of the user through the comparison with the size of each candidate. In experimental results, the proposed method showed high detection rate with gaze direction and foot sensing rate with user's position, and showed the performance possibility of the scenario for the disabled.

Keywords : The Disabled, Gaze Detection, Intelligent Space, Histogram, Color Segmentation

## 1. 서 론

인간-컴퓨터 상호 작용은 인간 공학과 정보 기술을 융합하는 상호 작용 기술을 연구하는 학문으로써, 인공지능, 영

상처리, 3차원 그래픽 투영 기술, 촉각 기술, 시스템 제어 기술, 인간 공학기술 등을 포괄하고 있다. Brooks가 제안한 지능형 공간(Smart Space)은 사용자와 실내 환경, 네트워크 통신 서비스 사이의 상호 연동이 가능한 기기와 응용서비스를 갖춘 소프트웨어에 의하여 구축되며, 실제 공간과 컴퓨터 사용공간을 융합하여 사용자의 실내 활동을 지원한다. 이러한, 지능형 공간 시스템은 유비쿼터스 컴퓨팅 환경과 사용자 간의 상호작용을 통하여, 사용자의 행동을 인식하며, 사용자에게 필요한 정보와 최적 서비스를 능동적으로 제공하는 동시에 시간이나 장소에 제약을 받지 않은 상태로 주

\* 본 연구는 숭실대학교 BK21 디지털영상산학공동사업단(2006303-2013.02)에서 지원을 받은 것임.  
† 준 회 원 : 숭실대학교 미디어학과 박사과정  
\*\* 준 회 원 : 숭실대학교 글로벌미디어학부 비전임강사  
\*\*\* 총신회원 : 숭실대학교 IT대학 글로벌미디어학부 교수(교신저자)  
논문접수 : 2011년 4월 29일  
수정일 : 1차 2011년 6월 8일, 2차 2011년 7월 1일  
심사완료 : 2011년 7월 7일

변 환경으로부터 생성되는 상황을 파악한다[1-2].

최근 교통사고, 약물사고, 산업재해, 각종 질병 등으로 인한 장애 발생률이 증가하고, 저출산 현상, 의학 기술의 발달과 함께 노인 인구의 비중이 증가하면서 많은 사회문제가 야기되고 있는데, 지능형 공간 시스템 연구는 이와 같은 문제 해결에 매우 중요한 역할을 할 수 있다. 그러나 현재까지의 지능형 공간 연구는 주로 일반인을 대상으로 진행되었기 때문에, 지원되는 다양한 기능들을 장애인에게 그대로 적용하기 어렵고, 장애 유형에 적합한 명령 입력 방법과 정보 제공 방법에 대한 추가적으로 필요하다.

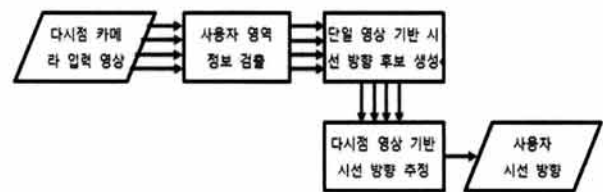
본 논문에서 우리는 장애 유형을 청각 장애인으로 한정하고, 장애인에게 보다 친화적인 지능형 공간을 제공하기 위한 다시점 영상(Multi-view Camera Image) 기반의 시선 방향 검출 방법을 제안한다. 일상 생활에서 청각 장애인이 정보 제공 여부를 인지하는 방법은 타 사용자와의 직접적인 접촉을 제외하면, 바라보고 있는 지점에 정보를 보여주는 것이 유일하기 때문에, 타 사용자의 도움을 받을 수 없는 지능형 공간과의 상호 작용을 위해서는 공간 내 사용자의 시선 방향 검출 방법이 필수적이다.

카메라 기반 시선 방향 검출 방법은 카메라 입력 영상을 이용하여 사용자의 얼굴이나 눈동자의 움직임을 검출하고, 시선 방향각을 추정함으로써, 사용자의 행동 및 의도를 분석하여 사용자에게 생활 지원 서비스를 제공할 수 있다[3]. 카메라 기반의 시선 방향 검출 연구로는 Ishii 등과 Hongo 등이 제안한 4 방향성 특징값(4-directional features) 기반 검출 방법[4-5]과 Wu와 Aghajan이 제안한 chamfer 매칭 기반 검출 방법[6], 공병용 등이 제안한 피부색 기반 검출 방법[7] 등이 있다. Ishii 등과 Hongo 등이 제안한 방법은 낮은 차원의 특징값을 이용하기 때문에 얼굴 영역 검출 시간이 빠르고 다양한 각도의 시선 방향에 대해서도 정확한 검출을 할 수 있다는 장점을 가지는 반면, 충분히 큰 시선 방향의 변화만 감지할 수 있다는 단점이 있다. Wu와 Aghajan이 제안한 방법과 공병용 등이 제안한 방법은 미리 정의된 시나리오를 바탕으로 사용자의 의도를 파악하기 때문에 시선 방향 검출과 사용자 행동 분석이 용이하지만 미리 정의된 시나리오 이외의 행동에 대해서는 반응하지 못한다는 단점이 있다. 또한, 시선 방향 검출 기술을 장애인 편의 시설 및 장치에 적용한 연구는 Y. Matsumotos 등[8]과 Mandel과 Frese[9], 정명진 등[10]이 제안한 시선 방향 검출 기반 휠체어 제어 기술, Spexard 등[11]이 제안한 인간 시선 방향-로봇 인터페이스 기술 등이 있다. 이들 연구는 모두 시선 방향 검출 기술을 통해 장애인에게 보다 편리한 환경을 제공할 수 있다는 장점이 있으나, 자유롭게 움직이는 실내 사용자에게 적용하기 어렵다는 단점이 있다.

본 논문에서 우리는 장애인 사용자에게 실내 생활 편의를 지원할 목적으로, 다시점 카메라 시스템(multi-view Camera System)[12,13]을 이용하여 실내 공간에서 자유롭게 움직이는 사용자의 시선 방향을 검출하는 방법을 제안한다. 다시점 카메라 시스템은 동일한 시간에 동일한 영역을 다양한

시점에서 촬영할 수 있는 시스템으로, 촬영 시점을 의미하는 하나의 프레임이 다수의 영상으로 구성되기 때문에 사용자 관심 영역(Region of Interest: ROI)에 대한 3차원 정보 추출에 주로 사용된다. 따라서 제안하는 방법은 기존의 2차원 평면에 국한된 시선 검출 방법을 3차원으로 확장하여 보다 효율적으로 공간을 활용할 수 있도록 한다.

(그림 1)에서 보는 것처럼, 우리가 제안하는 시선 방향 검출 방법은 사용자 영역 정보 검출 단계와 단일 영상 기반 시선 방향 후보 생성 단계, 다시점 영상 기반 시선 방향 추정 단계로 나눌 수 있다. 또한, (그림 1)의 화살표는 각 단계에서 출력되는 데이터의 흐름과 개수를 의미한다.



(그림 1) 프로그램 구성 및 데이터 흐름도

사용자 영역 정보 검출 단계는 입력 영상 안에서 시선 방향 검출을 위해 고려해야 할 데이터의 양을 줄이고 검출 속도를 향상시킬 수 있도록, 다시점 카메라 시스템을 통해 입력 받은 각각의 영상에 대해 사용자의 얼굴 영역과 위치를 검출한다. 단일 영상 기반 시선 방향 후보 생성 단계는 각각의 시점 별 영상에서 사용자의 좌우 및 상하 시선 방향각을 계산하여 카메라 수만큼의 후보를 생성하고, 다시점 영상 기반 시선 방향 추정 단계는 환경적인 요인에 의해 오류가 포함된 후보들을 제거하고 현재 프레임에 대한 사용자의 시선 방향을 결정한다.

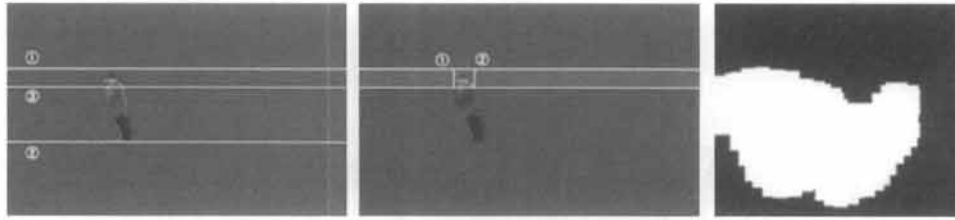
본 논문은 앞에서 간략하게 설명한 시선 검출 방법의 각 단계를 2~4장에서 상세하게 설명하고, 5장에서 실험과 그 결과를 보여주며, 6장에서 결론과 향후 연구 방향을 설명한다.

## 2. 사용자 영역 정보 검출

### 2.1 얼굴 영역 검출

제안하는 방법은 영상 내의 얼굴과 발 위치를 이용하여 사용자의 시선 방향을 계산한다. 우리가 제안하는 시선 방향 검출 방법은 크게 사용자 영역 정보 검출과 단일 영상 기반 시선 방향 후보 생성, 다시점 영상 기반 시선 방향 추정의 세 단계로 나눌 수 있고, 본 장에서는 영상 내의 사용자의 얼굴 영역과 발의 위치를 검출하는 사용자 영역 정보 검출 단계에 대해서 설명한다.

사용자의 얼굴 영역과 발 위치를 찾기 위해 먼저 영상에서 사용자(객체) 영역을 찾는 과정이 필요하다. 우리는 영상 내에서 객체를 추출하기 위해 차영상 방법을 사용한다. 입력 영상은 균일하지 못한 조명과 방향에 따른 그림자와 잡음을 포함하고 있기 때문에, 우리는 HSV 색상 공간을 이용



(a) 얼굴 후보 영역 선택 (b) 얼굴 영역 검출 (c) 이진화

(그림 2) 얼굴 영역 검출 과정

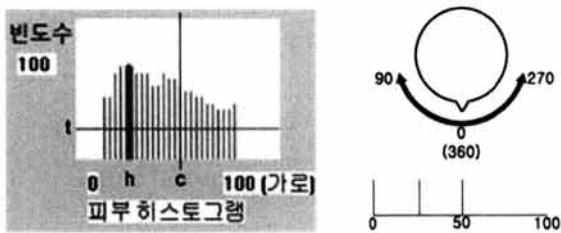
하여 차영상을 계산한다. 객체 추출 결과 영상을 이용하여 머리 끝(그림 2a)-①과 발 위치(그림 2(a)-②)를 찾고, 이 구간에서 얼굴 후보 영역을 결정한다(그림 2(a)-③). 선택된 후보 영역에서 얼굴의 왼쪽(그림 2(b)-①)과 오른쪽(그림 2(b)-②) 끝을 찾은 후, Y-Cb-Cr 색상 공간을 이용하여 피부색 영역만을 검출하고 모폴로지 연산을 통해 잡음과 홀을 제거하여 얼굴 영역을 검출한다(그림 2(c)). 최종적으로 검출된 영역의 크기를 100x100 픽셀 크기로 정규화한다.

### 3. 단일 영상 기반 시선 방향 후보 생성

단일 영상 기반 시선 방향 후보 생성 단계에서는 이전 단계에서 검출된 얼굴 영역과 발 위치를 이용하여 각각의 카메라 입력 영상에 대한 시선 방향 후보를 생성한다. 수평 방향 시선각은 사용자가 얼굴을 좌우로 움직이는 각도를 의미하며, X-Y 평면에서의 회전각을 측정하고, 수직 방향 시선각은 사용자가 얼굴을 상하로 움직이는 각도를 의미하며, Z축에 대한 회전각을 측정한다.

#### 3.1 수평 방향 시선각 추정

(그림 3(a))에서 보는 것처럼, 이진화된 얼굴 영역은 모든 픽셀을 x축에 대해 투영하고 그 누적된 값을 가로축 100, 세로축 100의 피부색 히스토그램으로 표현하였다. 검출되는 얼굴 영역은 카메라 위치에 따라 전면, 또는 후면으로 방향이 바뀌기 때문에, 본 논문에서는 히스토그램의 분포를 이용하여 얼굴의 방향을 판단하고, 각각의 방향에 대해 서로 다른 측정 방법을 적용하여 시선 방향을 추정한다.

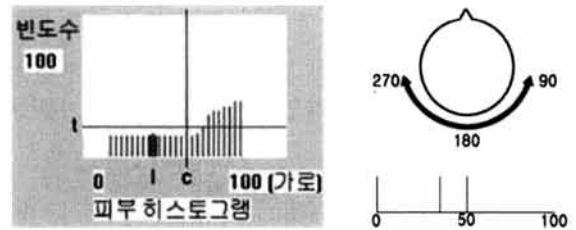


(a) 얼굴 전면 여부 판정 및 최고 빈도수 영역 탐색 (b) 히스토그램 구간에 대응하는 실제 얼굴 회전 범위

(그림 3) 얼굴 전면에 대한 수평 방향 시선각 계산

(그림 3)은 얼굴 전면에 대한 수평 방향 시선각을 계산하는 방법을 보여준다. 히스토그램의 중심선(c)을 기준으로 좌측과 우측 영역의 평균값이 모두 문턱치(t)보다 클 경우를 얼굴 전면이라고 판단하고, 가장 큰 도수를 가지는 영역(h)을 찾은 후, 비례식을 이용하여 해당 영역에 대한 각도를 계산한다. 이때, 최대값이 중심선의 왼쪽 영역에서 나타나면 사용자가 고개를 오른쪽으로 향하고 있는 상태이고, 반대 영역에서 나타나면 왼쪽을 향하고 있는 상태이다(그림 3(b)). 따라서 얼굴 전면 영상을 이용한 수평 방향 시선각( $D_h$ )은 아래의 식(1)에 의해 계산된다.

$$D_h = \begin{cases} \frac{9}{5}(50-h) & , \quad \text{if } h \leq 50 \\ \frac{9}{5}(250-h) & , \quad \text{otherwise} \end{cases} \quad (1)$$



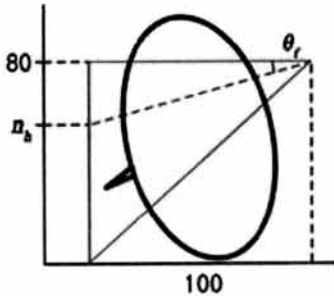
(a) 얼굴 후면 여부 판정 및 최저 빈도수 영역 탐색 (b) 히스토그램 구간에 대응하는 실제 얼굴 회전 범위

(그림 4) 얼굴 후면에 대한 수평 방향 시선각 계산

(그림 4)는 얼굴 후면에 대한 수평 방향 시선각을 계산하는 방법을 보여준다. (그림 4(a))에서 보는 것처럼, 우리는 히스토그램의 중심선(c)을 기준으로 좌측과 우측 영역의 평균값 중 하나라도 문턱치(t)보다 작을 경우를 얼굴 후면이라고 판단한다. 얼굴 후면이라고 판단될 경우에는 전면의 경우와 달리, 가장 작은 도수를 가지는 영역(l)을 찾고 비례식을 이용하여 해당 영역에 대한 각도를 계산한다. 얼굴 후면에 대한 수평 방향 시선각을 계산하는 수식은 다음과 같다.

$$D_h = \frac{9}{5}(150-l) \quad (2)$$

3.2 수직 방향 시선각 추정



(그림 5) 수직 방향 시선각 계산 과정

얼굴 영역은 모든 방향에 대해서 100x100 픽셀의 크기로 정규화되기 때문에, 지름 100, 높이 100의 원통이라고 가정할 수 있다. 또한, 4.1절에서 설명한 것처럼, 피부색 히스토그램에서 최대값을 가지는 영역(h)을 얼굴 전면의 가운데 영역으로, 최소값을 가지는 영역(l)을 얼굴 후면의 가운데 영역으로 가정하기 때문에, 얼굴 영역에서의 수직 방향 각도는 간단한 삼각 함수 연산에 의해 계산이 가능하다. (그림 5)에서 보는 것처럼, h의 도수 n\_h가 80일 때 사용자는 정면을 바라보고 있다라고 가정할 수 있고, 정면을 바라볼 때의 각도를 90로 정의하기 때문에, 입력 영상 내의 얼굴 영역이 전면일 경우, 수직 방향 각도를 계산하는 식은 다음과 같다.

$$D_v = (\tan^{-1}(\frac{n_h}{100})) + 51 \tag{3}$$

후면일 경우, l의 도수 n\_l이 10일 때 사용자가 정면을 바라보고 있다라고 가정하기 때문에 아래의 식(4)에 의해 계산된다.

$$D_v = 84 - (\tan^{-1}(\frac{n_l}{100})) \tag{4}$$

4. 다시점 영상 기반 시선 방향 추정

우리는 최종적으로 사용자의 시선방향을 결정하기 위해서, 3장에서 설명한 같은 시간에 촬영된 하나의 프레임을 구성하는 다양한 시점의 카메라 입력 영상에서 계산된 시선 방향 후보들의 평균값을 이용한다. 그러나, 단순히 후보들의 평균치를 이용할 경우, 특정 카메라에서의 프레임 입력 누락이나 계산 오류에 의해 잘못된 시선 방향을 얻을 수 있다. 이러한 문제를 해결하기 위해, 우리는 카메라 별 프레임 입력 누락 여부 확인, 후보 별 오류 포함 여부 확인, 시선 방향 각도 결정의 세 단계를 거쳐 지능형 공간 내 사용자의 시선 방향 각도를 결정한다.

본 논문에서 사용하는 시스템은 4대의 카메라를 통해 영상을 입력 받는다. 이론적으로 각각의 카메라는 초당 15프레임의 속도로 영상을 입력 받을 수 있으나, 데이터 전송 오류로 인한 프레임 누락이 발생할 수 있다. 따라서 우리는 동일

한 시간에 입력 받은 영상들의 인덱스 정보를 확인하여 프레임 누락 여부를 판정하고, 누락된 카메라에 대한 시선 방향 후보가 각도 결정에 영향을 주지 않도록 제외한다(식(5)).

$$drop_i = \begin{cases} 0 & , \text{if } Cam_i = null \quad (i : \text{id of camera}) \\ 1 & , \text{otherwise} \end{cases} \tag{5}$$

또한, 각각의 카메라 입력 영상은 영상 전처리 단계 중 차영상을 생성하는 과정에서 노이즈가 발생할 수 있다. 이 경우에 계산된 시선 방향 후보들은 오류를 포함하게 되고, 최종 시선 방향각 결정에 영향을 주기 때문에 각각의 후보들의 오류 포함 여부를 확인해야 한다. 오류 포함 여부 확인에 앞서 시선 방향 후보들 간의 좌표계를 일치 시켜주는 과정이 필요하다. 본 논문에서 사용한 4대의 카메라는 실험실 천정에 90° 간격으로 설치가 되어있지만, 각각의 카메라 입력 영상에서 계산된 수평 방향 시선각 후보들은 모두 사용자가 정면을 바라볼 때를 0°라고 가정하고 계산된 결과들이기 때문이다. 각도 보정을 위해 우리는 1번 카메라에서 계산된 수평 방향 시선각이 실제 시선 방향과 일치한다고 가정을 하고, 나머지 카메라의 각도를 아래와 같이 보정한다.

$$\begin{aligned} D_h(Cam_i) &= D_h(Cam_1) - 90(i-1) \\ D_h(Cam_i) &= 360 + D_h \quad , \text{if } D_h(Cam_i) < 0 \end{aligned} \tag{6}$$

오류가 포함된 후보(outlier)는 그렇지 않은 후보들과 비교적 멀리 떨어져서 나타나게 된다. 따라서, 우리는 후보들 간의 좌표계를 일치 시킨 후, 식(7)을 이용하여 최대/최소값을 갖는 후보를 제외한 나머지 후보들의 평균을 계산하고, 각각의 후보들과의 차를 비교함으로써 오류 포함 여부를 판단하여 outlier를 제거한다(식(8)).

$$Mean = \frac{\sum_{i=1}^n D_i - \min(D_1 : D_n) - \max(D_1 : D_n)}{\# \text{ of camera} - \# \text{ of drop}} \tag{7}$$

$$err_i = \begin{cases} 0 & , \text{if } |D_i - Mean| > \text{threshold} \\ 1 & , \text{otherwise} \end{cases} \tag{8}$$

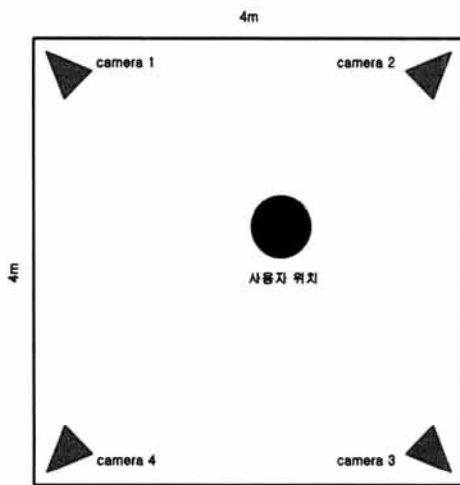
위의 과정을 통해, 프레임 누락 여부와 후보의 오류 포함 여부를 확인하면, 다음 식(9)와 같이 남은 후보들 간의 평균을 계산하여 시선 방향을 결정한다.

$$(\theta_h, \theta_v) = \frac{\sum_{i=1}^n Drop_i \times err_i \times (D_h - (Cam_i), D_v(Cam_i))}{\# \text{ of camera} - (\# \text{ of drop} + \# \text{ of err})} \tag{9}$$

5. 실험 및 결과

실험 환경은 1대의 PC와 4대의 카메라로 구성된다. PC는 Intel core2Quard 2.8Ghz CPU와 2GByte 메모리, 카메라 영

상 입력을 위한 IEEE 1394 캡처보드 2개로 구성된다. 카메라는 PointGrey Flea2와 3.5mm 광각 렌즈로 구성되어 있고, 각각의 카메라는 640×480 해상도의 영상을 15fps 속도로 입력 받는다. 카메라는 (그림 6)에서 보는 것처럼 바닥으로부터 2.5m 높이의 4m X 4m 천정 각 모서리에 90°간격으로 설치되어 있기 때문에, 지능형 공간 내부에 위치한 사용자의 수평/수직 방향 시선각과 위치를 측정하여 시선대상 물체 추정이 가능하다.



(그림 6) 지능형 공간 내의 카메라 배치

본 논문에서 우리는 실험을 통해 제안된 시선 방향 검출 방법 성능을 확인하기 위해, 계산 결과의 정확도를 측정하였고, 제안된 방법이 실제 장애인 사용자에게 적합한지 여부를 확인하기 위해, 지능형 공간에서 사용자가 응시하는 물체를 추정하는 시스템에 적용하고 그 결과를 분석하였다. 또한, 정확한 시선 방향 추정을 위해 다음의 두 가지를 한다.

- 첫째, 지능형 공간에서 사용자는 항상 서있는 상태이다.
- 둘째, 사용자는 귀와 목 뒷부분이 보이도록 머리카락을 정리한 상태이다.

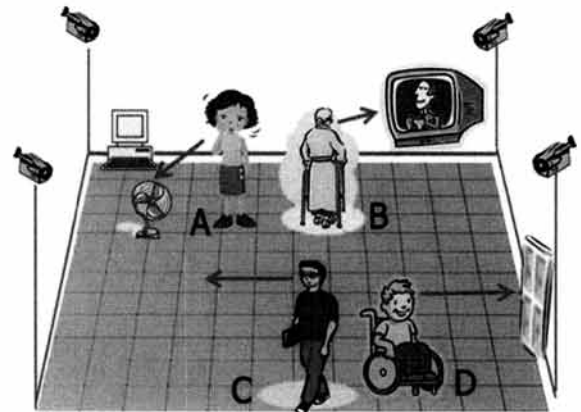
첫 번째 가정은 사용자의 서있는 자세를 얼굴 영역의 중심과 발바닥 위치를 연결하는 선으로 표현하기 위함이고, 두 번째 가정은 사용자의 얼굴이 향하고 있는 방향을 판단하기 쉽도록 하기 위해서이다.

5.1 실험 시나리오

(그림 7)에서 보는 것처럼, 우리는 지능형 공간 내에서 장애인이 생활할 때 장애 유형에 따라 시선 방향 검출이 필요한 상황을 아래의 4가지로 설정하고, 설정된 시나리오에 의해 실험 데이터를 수집하였다. 본 논문에서는 사용자의 시선 방향을 검출하고 그에 따른 지원 서비스를 제공하는 공간으로 가정하였기 때문에 장애 유형 중 시각 장애는 제외하고 시나리오를 설정하였다.

1. 사용자 A는 선풍기를 바라보고 있다.
2. 사용자 B는 TV를 바라보고 있다.
3. 사용자 C는 벽을 바라보고 있다.
4. 사용자 D는 문을 향해 이동하고 있다.

1, 2의 경우, 지능형 공간은 사용자가 바라보는 대상 물체를 추정하고, 사용자가 물체를 제어할 수 있도록 도와주고, 3의 경우 사용자가 바라보고 있는 곳에 사용자에게 필요한 정보를 표시해준다. 4의 경우, 사용자의 시선 방향에 위치하고 있는 문과 사용자 사이의 거리에 따라 일정 거리 이상 가까워질 경우 문을 열어 사용자가 지나갈 수 있도록 한다.

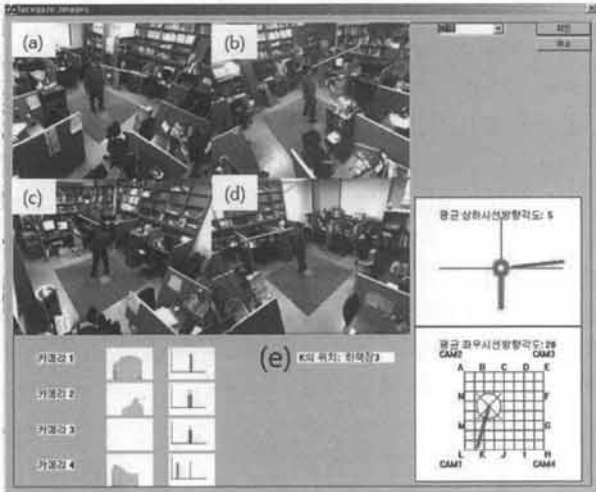


(그림 7) 지능형 공간 내 장애인의 시선 방향 검출 시나리오

5.2 시선 방향 검출 방법 성능 평가

제안된 방법의 성능을 평가하기 위해 실험 대상을 청각 장애인으로 설정하고 6.1절에서 설명한 시나리오에 따라 지능형 공간 안에서 위치를 이동하면서 미리 정한 28개의 대상 물체를 응시하도록 했다. 시선 대상 물체는 수평 방향에 대해 실험 공간을 14개 영역으로 균등 분할하고, 각 영역을 상하 2개 영역으로 구분한 뒤 각각의 영역에 포함되는 물체를 선택하였다. 총 7회에 걸쳐 약 9분 동안 실험을 수행하는 사이에 지능형 공간에 설치된 4대의 카메라로부터 입력 받은 9800장의 영상을 이용하여, 피부색 히스토그램을 생성하고, 수평 방향 시선각과 수직 방향 시선각을 추정하였다. 사용자의 위치가 고정되지 않은 환경에서 사용자의 공간 내 실제 시선각을 계산하는 것이 불가능하기 때문에, 추정 결과의 정확도는 지능형 공간 안에서 사용자가 실제로 바라보는 물체(그림 8(a-d))와 시스템이 추정한 시선 방향에 위치하는 물체(그림 8(e))의 일치 여부를 통해 계산하였다. 테스트를 위한 프로그램은 사용자의 위치와 수평 시선각을 이용하여 균등하게 분할된 14개 영역 중에서 1차적으로 시선 방향을 계산하고, 해당 영역에 대해 수직 방향 시선각을 적용하여 시선 방향에 위치한 물체를 추정하였다.

다음의 <표 1>과 <표 2>는 420장의 영상을 이용하여, 색상 공간의 종류에 따른 시선각의 정확도를 비교한 것이다. 아래의 결과에서 보는 것처럼, 단일 영상 시퀀스를 이용하는 경우와 다시점 영상 시퀀스를 이용하는 경우 모두에서,



(그림 8) 다시점 기반 시선 방향각 추정 프로그램

HSV 색상 공간이 RGB 색상 공간보다 높은 정확도를 보임을 확인 할 수 있었다.

<표 1> RGB-HSV 색상 공간 변환의 수평 방향 시선각 정확도(단위: %)

	단일 영상 시퀀스 이용				다시점 영상 시퀀스 이용
	Cam <sub>1</sub>	Cam <sub>2</sub>	Cam <sub>3</sub>	Cam <sub>4</sub>	
RGB 공간	89.5	93.0	81.9	89.5	88.5
HSV 공간	99.04	99.04	99.04	99.04	99.04

<표 2> RGB-HSV 색상 공간 변환의 수직 방향 시선각 정확도(단위: %)

	단일 영상 시퀀스 이용				다시점 영상 시퀀스 이용
	Cam <sub>1</sub>	Cam <sub>2</sub>	Cam <sub>3</sub>	Cam <sub>4</sub>	
RGB 공간	82.5	92.38	85.7	87.6	84.76
HSV 공간	95.2	100.0	95.2	99.04	99.04

다음의 <표 3>과 <표 4>는 단일 영상 시퀀스를 이용하여 사용자의 수평 방향 시선각과 수직 방향 시선각을 측정하고 그 정확도를 분석한 결과이다. 각각의 표는 단일 카메라에서 계산된 시선 방향각 후보들의 정확도와 4개 후보들을 평균하여 계산된 최종 결과의 정확도를 함께 보여준다. <표 3>과 <표 4>를 통해, 단일 영상을 이용하여 사용자의 시선각을 계산할 경우, 카메라의 시점과 사용자의 방향에 따라 정확도가 많은 차이를 보이고 있음을 알 수 있었다. <표 5>는 본 논문에서 제안한 방법을 이용하여 사용자의 시선각을 계산한 결과와 단일 영상을 이용하여 계산된 결과를 비교한 표이다. <표 5>에서 보는 것처럼, 본 논문에서 제안하는 다시점 영상 기반 사용자 시선각 추정 방법은 후

보 별 오류 검사와 데이터 보정 과정을 통해 카메라의 시점과 사용자의 방향에 영향을 적게 받기 때문에, 단일 영상 기반 추정 방법과 비교하여 보다 정확한 시선각 추정이 가능함을 확인하였다.

<표 3> 단일 영상 시퀀스를 이용한 수평 방향 시선각 정확도(단위: %)

실험 번호	획득 영상갯수	Cam <sub>1</sub>	Cam <sub>2</sub>	Cam <sub>3</sub>	Cam <sub>4</sub>	평균 정확도
1	704	96.6	94.3	93.2	93.8	94.5
2	1512	99.5	99.2	97.3	98.4	98.6
3	1508	97.3	98.4	97.6	98.9	98.0
4	1472	97.8	99.4	83.9	95.9	94.2
5	1492	100.0	98.9	97.5	98.9	98.8
6	1400	100.0	99.4	96.5	99.7	98.9
7	1712	92.9	99.76	90.65	99.76	95.8

<표 4> 단일 영상 시퀀스를 이용한 수직 방향 시선각 정확도(단위: %)

실험 번호	획득영상 영상갯수	Cam <sub>1</sub>	Cam <sub>2</sub>	Cam <sub>3</sub>	Cam <sub>4</sub>	평균 정확도
1	704	93.2	92.6	93.2	89.8	92.2
2	1512	100.0	99.5	99.5	100.0	99.75
3	1508	100.0	99.7	99.7	99.7	99.75
4	1472	100.0	99.7	89.0	95.3	96.0
5	1492	98.1	96.7	96.7	100.0	97.88
6	1400	92.9	97.1	96.3	99.1	96.35
7	1712	92.75	98.1	89.48	99.53	94.97

<표 5> 단일 영상과 다시점 영상 시퀀스를 이용한 시선각 정확도의 비교(단위: %)

실험 번호	수평 방향 시선		수직 방향 시선	
	단일 영상 이용	다시점 영상 이용	단일 영상 이용	다시점 영상 이용
1	94.5	94.9	92.2	99.4
2	98.6	98.9	99.75	99.7
3	98.0	97.6	99.75	98.4
4	94.2	99.1	96.0	99.7
5	98.8	99.1	97.88	98.1
6	98.9	99.7	96.35	99.1
7	95.8	99.3	94.97	99.76
평균	96.7	98.37	96.7	99.17

<표 6>은 기존의 시선 방향 검출 방법과 제안된 방법을 비교한 표이다. 각각의 방법의 성능은 측정 간격과 카메라 수, 정확도, 처리 속도의 4개 항목을 이용하여 비교하였다. <표 6>에서 보는 것처럼, 제안하는 방법은 다른 방법들과 달리 수직 방향의 시선각을 추정할 수 있으며, 시선 방향

측정 간격이 기존 방법들과 비교하여 조밀하고 정확도가 높기 때문에, 지능형 공간을 보다 효율적으로 활용할 수 있다. 시선각을 계산하기 위한 처리 시간은 초당 4프레임으로 실시간 정보 제공 기준인 15프레임에는 못 미치지만, 장애인들의 움직임 또한 조심스럽고 일반인과 비교하여 느리다는 점을 감안한다면, 지능형 공간 내 장애인들에게 정보를 제공하기에 충분하다고 할 수 있다.

〈표 6〉 기존 시선 방향 검출 시스템과의 비교

항목(단위)	제안하는 방법	W. Chan et al.[6]	공병용 등[7]
카메라 수(대)	4	4	1
정확도(%)	수평방향: 98.37 수직방향: 99.17	수평방향: 91.5 수직방향: -	수평방향: 76 수직방향: -
처리 속도(fps)	4	10-15	-

## 6. 결 론

본 논문에서는 지능형 공간에 있는 청각장애인 사용자에게 생활 지원 서비스를 제공하기 위한 사용자 시선 방향 검출 방법을 제안하였다. 제안하는 방법의 개발 목적은 지능형 공간 안에서 얼굴 시선 방향 검출을 통하여 시선 대상의 정보를 제공함으로써 장애인의 불편함을 개선하기 위함이다. 제안하는 방법은 다시점 카메라 입력 영상 내의 객체를 추출하고, 얼굴 영역 및 발 위치를 검출하여 수평/수직 방향 시선각을 계산하기 때문에 단일 카메라를 이용하는 기존 시스템과 비교하여 높은 정확도를 얻을 수 있었다. 향후, 얼굴 영역 검출 방법 개선을 통해 처리 속도를 향상시키고, 시선 대상 물체의 제어가 가능한 지능형 공간에 적용하여 장애인의 실내 활동 지원 서비스를 연구할 예정이다.

## 참 고 문 헌

[1] R. A. Brooks, "Intelligence without reason," Proceeding of the 9th International Joint Conference on Artificial Intelligence, pp.569-595, 1991.  
 [2] R. A. Brooks, D. Dang, J. D. Bonet, J. Kramer, J. Mellor, P. Pook, C. Stauffer, L. Stein, M. Torrance and M. Wessler, "Intelligent Room Project," Proceedings of the 2nd International Cognitive Technology, pp.271-279, 1997.  
 [3] 김종배, 김항준, "모델 기반의 SVM을 이용한 시선 방향 추정", 대한전기학회 2007 CICS 정보 및 제어 학술대회, pp.121-122, 2007.  
 [4] Y. Ishii, H. Hongo and K. Yamamoto, "Real-Time Face and Head Detection using Four Directional Features,"

Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp.403-408, 2004.  
 [5] H. Hongo, M. Yasumoto, Y. Niwa and K. Yamamoto, "Combination of Face Direction Estimation and Face Recognition Using Four-Directional Features," Proceeding of the 5th Asian Conference on Computer Vision, Vol.2, pp.737-742, 2002.  
 [6] C. Wu and H. Aghajan, "Head pose and trajectory recovery in uncalibrated camera networks - Region of interest tracking in smart home applications," Proceedings of the 2nd ACM/IEEE International Conference on Distributed Smart Cameras, pp.1 - 7, 2008.  
 [7] 공병용, 정도준, 김항준, "실내 환경에서 시선 방향을 고려한 사람 행동 인식", 대한전기학회 2007 CICS 정보 및 제어 학술대회, pp.119-120, 2007.  
 [8] Y. Matsumoto, T. Ogasawara and T. Ino, "Development of intelligent wheelchair system with face and gaze based interface," Proceeding of the 10th IEEE International Workshop on Robot and Interactive Communication, pp.262 - 267, 2001.  
 [9] C. Mandel and U. Frese, "Comparison of Wheelchair User Interfaces for the Paralyzed: Head-Joystick vs. Verbal Path Selection from an offered Route-Set," Proceedings of the 3rd European Conference on Mobile Robots, 2007.  
 [10] 정명진, 김재현, 이영진, 유동현, "척수 장애인의 HCI를 위한 시선방향 추적 시스템", 서비스 로봇 워크샵, pp.183-192, 2001.  
 [11] T. P. Spexard, M. Hanheide and G. Sagerer, "Human-oriented Interaction with an Anthropomorphic Robot," IEEE Transactions on Robotics, Vol.23, No.5, pp.852-862, 2007.  
 [12] 윤태현, "2차원 교차영역과 Visual hull을 이용한 실제-가상 객체 간 고속 3차원 충돌 감지 알고리즘", 숭실대학교 대학원 석사논문, 2010.  
 [13] Y. Oh, J. Kim, T. Yoon and K. Jung, "Face Detection From Four Captured Images Related To Intelligent Room For The Deaf," Proceeding of the 1st International Conference on Smart IT Applications, Vol.11, pp.1-5, 2009.

## 오 영 준



e-mail : yjooh@ssu.ac.kr  
 2001년 성공회대학교 정보통신학과(이학사)  
 2003년 숭실대학교 컴퓨터학과(공학석사)  
 2008년~현 재 숭실대학교 미디어학과  
 박사과정  
 2012년 숭실대학교 미디어학과(공학박사)

2004년~2008년 한국과학기술원 인간친화복지로봇시스템연구센터  
 전임연구원  
 관심분야: HCI, 컴퓨터 비전, 가상현실, 재활 공학, 수화 공학



### 홍 광 진

e-mail : hongmsz@gmail.com  
 2004년 숭실대학교 컴퓨터학부(공학사)  
 2006년 숭실대학교 미디어학과(공학석사)  
 2009년 숭실대학교 미디어학과(공학박사)  
 2003년~현 재 숭실대학교 글로벌미디어  
 학부 비전임강사

관심분야: HCI, 영상처리, Augmented Reality, Augmented  
 Game



### 김 종 인

e-mail : eyezzang@gmail.com  
 2005년 청운대학교 멀티미디어학과(이학사)  
 2008년 숭실대학교 미디어학과(공학석사)  
 2008년~현 재 숭실대학교 미디어학과  
 박사과정

관심분야: HCI, 컴퓨터 비전, 가상현실,  
 행동인식, 공간인식 인터페이스



### 정 기 철

e-mail : kcjung@ssu.ac.kr  
 1996년 경북대학교 컴퓨터학과(공학석사)  
 2000년 경북대학교 컴퓨터학과(공학박사)  
 1999년 Intelligent User Interface Group,  
 DFKI (The German Research  
 Center For Artificial Intelligence

GmbH), Germany 방문 연구원  
 1999년 Machine Understanding Division, Electro Technical  
 Laboratory, Japan 방문 연구원  
 2000년~2002년 PRIP Lab. Michigan State University, U.S.  
 박사후연구원  
 2003년~현 재 숭실대학교 IT대학 글로벌미디어학부 교수  
 관심분야: HCI, Interactive Contents, 영상 처리, 패턴 인식,  
 Augmented Reality, Mobile Vision