

논문 2012-49-9-14

# 시선 응시 점 기반의 관심영역 확장을 통한 원 거리 얼굴 검출

## ( Far Distance Face Detection from The Interest Areas Expansion based on User Eye-tracking Information )

박 희 선\*, 홍 장 표\*, 김 상 열\*, 장 영 민\*, 김 철 수\*\*, 이 민 호\*\*\*

( Heesun Park, Jangpyo Hong, Sangyeol Kim, Young-Min Jang, Cheol-Su Kim, and Minho Lee )

### 요 약

영상처리 기법을 이용한 얼굴검출에 관한 많은 다양한 방법들이 제시되어 왔다. 일반적으로 가장 많이 쓰이는 얼굴 검출 방식은 Viola와 Jones이 제안한 Adaboost 방식이다. 이 방식은 Haar-like feature을 이용하여 얼굴영상을 선행 학습하고, 검출 성능은 학습된 DB에 의존한다. 이는 일정 거리 범위 안의 학습된 얼굴 크기에서는 얼굴 검출을 잘 수행하지만, 카메라에서 객체(얼굴)의 거리가 멀어지면 얼굴 크기가 작아져 기존에 학습한 Haar-like feature로 얼굴 검출을 하지 못하는 경우가 발생한다. 이에 본 논문에서는 생물학 기반의 선택적 주의집중 기반의 Haar-like feature 정보를 이용한 Adaboost 모델과 사용자의 시선 응시 점 정보를 이용하여, 사용자의 관심영역 확장을 통한 원거리 얼굴 검출 모델을 제안한다. 생물학적 기반의 선택적 주의 집중 모델인 돌출맵(Saliency map) 정보를 이용하여 입력 영상에 대하여 얼굴 후보 영역을 검출하고, 검출된 얼굴 후보 영역 중에서 선행 학습된 Haar-like feature 정보로 Adaboost 알고리즘을 이용하여 최종 얼굴 영상을 검출한다. 그리고 사용자의 시선 응시 점 정보는 관심영역을 선택 하는데 이용된다. 피 실험자가, 카메라로부터 멀리 거리 떨어져 얼굴의 크기가 얼굴검출이 힘들더라도 사용자 시선 응시 점 영역을 선행 보간법으로 확대하여 입력영상으로 재사용함으로써 얼굴 검출 성능을 높일 수 있다. 제안된 방법이 기존의 Adaboost 방법보다 얼굴 검출 성능과 수행시간 면에서 우수함을 실험을 통해 확인하였다.

### Abstract

Face detection methods using image processing have been proposed in many different ways. Generally, the most widely used method for face detection is an Adaboost that is proposed by Viola and Jones. This method uses Haar-like feature for image learning, and the detection performance depends on the learned images. It is well performed to detect face images within a certain distance range, but if the image is far away from the camera, face images become so small that may not detect them with the pre-learned Haar-like feature of the face image. In this paper, we propose the far distance face detection method that combine the Adaboost of Viola-Jones with a saliency map and user's attention information. Saliency Map is used to select the candidate face images in the input image, face images are finally detected among the candidated regions using the Adaboost with Haar-like feature learned in advance. And the user's eye-tracking information is used to select the interest regions. When a subject is so far away from the camera that it is difficult to detect the face image, we expand the small eye gaze spot region using linear interpolation method and reuse that as input image and can increase the face image detection performance. We confirmed the proposed model has better results than the conventional Adaboost in terms of face image detection performance and computational time.

**Keywords** : face detect, adaboost, eye tracker, selective attention, saliency map, pupil detect

\* 학생회원, \*\* 정회원, \*\*\* 정회원-교신저자, 경북대학교  
(Kyungpook National University)

※ 이 논문은 2011학년도 경북대학교 융·복합연구 지원프로그램의 지원을 받아 연구되었음.  
접수일자: 2011년12월14일, 수정완료일: 2012년7월13일

## I. 서 론

얼굴검출은 컴퓨터 비전 연구 및 응용에서 중요한 요소 기술로서 얼굴인식, 표정인식, 인간-컴퓨터 상호작용(Human computer interaction : HCI)등 여러 응용 분야에 사용된다. 각 응용분야에서 사람 얼굴의 존재여부를 판단하는 첫 번째 단계이기 때문에 얼굴 검출의 성능은 전체 시스템의 성능에 영향을 미치게 된다. 기존의 얼굴검출 기법으로는 skin color filter 기법<sup>[1~3]</sup>, template matching 기법, 학습 알고리즘 기반의 Adaboost<sup>[4]</sup> 알고리즘 등의 방법들이 연구되어왔다. 하지만 앞서 거론한 모델은, 일정 거리 내에 존재하는 제한된 크기의 얼굴영역 검출은 가능하지만, 원 거리에 있는 작은 얼굴영역은 잘 검출하지 못한다. 그뿐 아니라 얼굴영역이 아닌 영역을 얼굴영역으로 검출하는 오검출(false positive error)이 발생하기도 한다. 이에 본 논문에서는 원거리의 작은 화소(pixel) 수를 가지는 얼굴영역의 검출과, False positive의 오 검출률을 줄이기 위한 방법으로 생물학기반의 선택적 주의집중<sup>[5~8]</sup> 모델과 Haar-like feature 정보를 이용한 Adaboost<sup>[4]</sup>모델과 사용자의 시선 응시 점<sup>[9]</sup> 정보를 결합하여 사용자의 관심영역 내에 존재하는 원거리의 작은 크기의 얼굴영역을 검출할 수 있는 방법을 구현하였다. 기존의 얼굴 검출 방법으로 잘 알려진 Adaboost 모델의 성능 자체는 학습된 데이터에 매우 의존적이지만 얼굴 검출 면에서 본다면 우수한 성능을 나타내는 것으로 알려져 있다<sup>[4]</sup>. 하지만 학습된 데이터에 의존적이기 때문에 크기가 작은 얼굴 영상 또는 회전된 얼굴 영상 데이터가 학습되지 않으면 이에 대한 검출 성능이 현저히 떨어지게 된다. 이와 같은 한계점은 실 환경에 적용 시 많은 문제점을 야기할 수 있으므로 본 논문에서는 사용자의 시선 응시점 정보를 이용하여 사용자의 관심 영역 내에 존재하는 원거리의 작은 얼굴을 검출 가능한 모델을 제안한다. 사용자의 시선 응시 점 정보를 이용하면, 사용자가 보고 있는 영역, 즉 관심 영역 확인이 가능하고, 사용자의 관심 영역을 확대하여 얼굴 검출을 시행할 경우 원 거리에 있는 작은 크기의 얼굴영역을 검출할 수 있다. 또한, 생물학적 기반의 선택적 주의집중 방법을 도입하여 얼굴 후보 영역을 먼저 찾은 후, 그 후보 영역 내에서 Adaboost 방법으로 얼굴 영역을 최종적으로 검출하는 방법으로, 입력된 영상 전체에서 얼굴영역을 검출하지

않고, 선택 되어진 얼굴 후보 영역 안에서 후보 영역들이 얼굴인지 아닌지를 검출/판별하게 된다. 그러므로 얼굴 검출의 소요 시간을 줄일 수가 있다. 본 논문에서 수행한 거리에 따른 얼굴영역 검출 실험에서 원 거리의 작은 크기의 얼굴영역에 대하여 강건한 얼굴 검출 결과를 확인 할 수 있었다.

본 논문의 구성은 다음과 같다. II장에서 제안한 모델의 구현 방법을 설명하고, III장에서는 실제 실험한 데이터와 실험 결과를 분석을 하며, IV장에서 결론을 맺는다.

## II. 제안한 모델의 구현 방법

### 1. 시스템 구성

본 논문에서 사용하는 시선 응시 점 정보 검출 및 얼굴영역 검출을 위한 안경 타입의 플랫폼<sup>[9]</sup>은 그림 1과 같으며, 외부영상 획득을 통해 얼굴영역 검출에 사용되는 전방 주시 카메라, 사용자의 동공영역을 밝게 비추주는 두 개의 적외선 광원, 사용자의 시선 응시점 정보를 정확하게 획득하기 위한 반사경과 시선 추적 카메라, 그리고 카메라로부터 획득한 정보들을 처리하는 PC에 전송할 수 있는 인터페이스 보드박스로 구성되어 있다. 특히 반사경은 사용자의 시야 확보 편의를 제공하면서 시선 응시점 정보를 획득하는데 사용되었고, 시선 추적 카메라는 적외선 카메라 효과를 내기 위해 일반 USB Webcam에서 형광 필터를 제거하여 사용하였다.

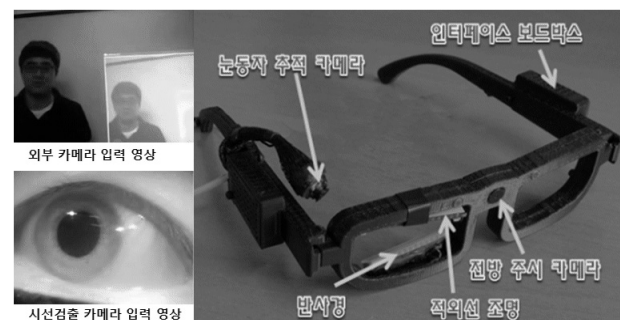


그림 1. 시선 응시 점 정보 기반 얼굴 검출 플랫폼.

Fig. 1. Face detection platform based on eye movement information.

### 2. 얼굴 검출 모델

본 논문에서 제안하는 얼굴 검출 방법은 크게 두 단

계로 구분된다. 첫 번째 단계에서는 사람의 시각 정보 처리 메커니즘을 모델링한 선택적 주의 집중 모델을<sup>[6~7]</sup> 이용하여 전체 입력 영상에서 얼굴 후보 영역을 선정한다. 전방 주시 카메라를 통해 들어오는 입력 영상의 밝기, 색상, 및 윤곽선과 같은 기초 시각 정보를 얻는다. 기초 시각 정보를 이용하여 생물학적 시각처리 메커니즘 모델을 통해 초기 시각 정보인 돌출맵(saliency map) 정보를 추출하여 주변보다 두드러진 영역에 대하여 우선적으로 선택적 주의집중을 할 수 있도록 한다. 다음으로는 앞서 선정된 얼굴 후보 영역에 대하여, Viola와 Jones 가 제안한 Adaboost<sup>[4]</sup> 알고리즘을 적용하여 실제 얼굴영역 여부를 판별한다.

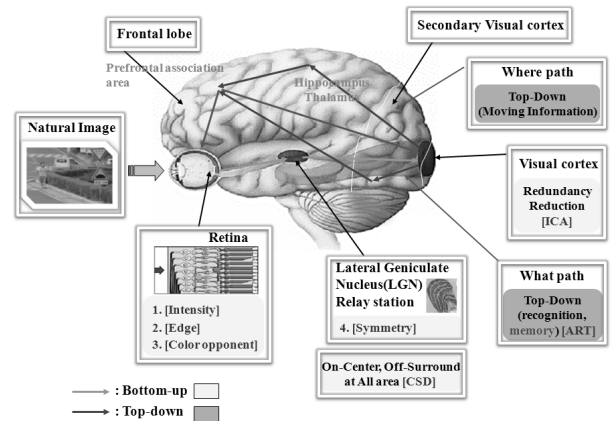


그림 2. 생물학적 시각 처리 경로  
Fig. 2. Biological visual pathway.

가. 생물학 기반의 얼굴지향 선택적 주의 집중 모델  
인간은 효율적인 시각 정보 처리를 위해 다양하고 복잡한 환경 인자 변화에 강건한 시각 처리 기능과 필요한 정보를 선별적으로 처리를 할 수 있도록 도와주는 선택적 주의 집중 메커니즘을 갖고 있다. 이러한, 인간 시각 시스템의 선택적 주의 집중 기능을 모방한 시각 시스템에 관련된 기존 연구로 Itti와 Koch는 Treisman의 심리학적 연구결과를 바탕으로 입력 영상에 대하여 밝기, 방향 및 보색 정보와 같은 특징 정보를 사용하여 주변보다 돌출적인 부분을 볼 수 있는 사람의 초기 시각 경로를 모방한 선택적 주의 집중 모델을 개발 하였으며<sup>[5, 10]</sup>, Ramstrom과 Christensen은 다중 해상도의 피라미드와 다중 특징을 사용하여 처리함으로써 돌출맵을 계산하였고<sup>[11]</sup>, Ziaoping Li는 돌출맵(Saliency Map : SM) 모델을 제안하였다<sup>[12]</sup>. 또한 Cynthia Breazeal, Brian Scassellati, 그리고 Gioriga Metta 등의 연구 그룹은 간소화된 선택적 주의 집중 모델을 이용하여 인간형 로봇 시스템에 적용시켜 보다 인간과 같이 동작할 수 있는 시스템을 제안하였다<sup>[13~15]</sup>. 그러나 이러한 여러 모델들은 인간의 시각 경로를 모방하여 시선 결정을 구현 하였지만, 극히 제한된 환경에서 동작이 가능하며, 객체지향 선택적 주의 집중을 고려하지 않은 모델들이다. 따라서 본 논문은 인간의 초기 시각 경로를 모방하여 입력된 영상에 대해 밝기, 색상, 윤곽선 정보 등의 초기 시각 정보를 이용하여 보다 인간과 유사하게 주변보다 두드러진 영역을 스스로 결정하는 기능을 갖는 상향식 과정(bottom-up)과 사용자의 선호도를 반영할 수 있는 하향식(top-down)모델이 고려된 실시간 객체 지

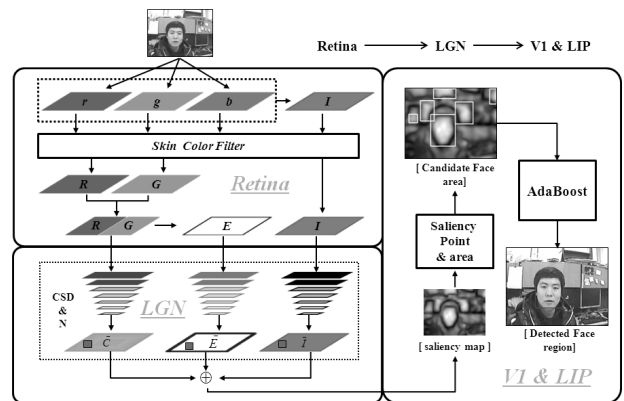


그림 3. 생물학적 기반 얼굴검출 선택적 주의 집중 모델  
Fig. 3. Biologically motivated selective attention model for face detection.

향 선택적 주의 집중 모델을 제안하며, 그 응용으로써 얼굴 후보 영역 검출에<sup>[16~17]</sup> 활용하였다.

그림 2는 인간의 생물학적인 시각 처리 메커니즘을 나타낸다.

그림 2에서 보는 바와 같이 상향식(bottom-up) 과정은 자료 주도적 처리과정을 말하는데, 이는 망막(retina)에서부터 시작하여 외측슬상체(lateral geniculate nucleus: LGN)를 거쳐 시각피질(visual cortex)에 이르게 된다. 이 과정에서 우리가 눈을 뜨고 대상을 바라보는 것과 같은 일차적 시각정보를 얻게 된다.<sup>[8]</sup>

그리고 하향식(top-down) 과정은 사전에 학습된 정보를 이용하는 개념 주도적 처리과정을 말하며, 그림 2에서 보는 바와 같이 고위 인지과정을 위한 두정엽(parietal lobe)으로 가는 ‘Where path’와 주변 사물의 인식 및 저장을 위한 측두엽(temporal lobe)으로 가는

‘What path’가 있다<sup>[8]</sup>. 본 논문에서는 얼굴지향 선택적 주의집중 모델에 초점을 맞추고, 하향식 과정의 메커니즘을 skin color filter의 정보로 활용하는 상·하향 통합 얼굴 지향 선택적 주의집중 모델을 제안한다.

그림 3은 본 논문에서 제안하는 인간의 생물학적인 시각 처리 메커니즘을 모델링한 상·하향식 시각 메커니즘을 보이고 있다.

제안된 시각 메커니즘의 얼굴 후보 영역 검출 모델에서는 망막에서 3가지 원추세포의 기능을 모델링하기 위해 먼저, 입력영상의 빨강색(r), 녹색(g) 그리고 파랑색(b) 정보를 추출하였고, 간상세포의 기능과 인간의 심리적 요소를 고려한 Treisman의 가설을 반영하기 위해 빨강색, 녹색, 파랑색의 평균값을 이용한 밝기정보를 식 (1)로 정의했다<sup>[10]</sup>.

$$I = \frac{r+g+b}{3} \quad (1)$$

그리고 사람의 얼굴영역의 선택적 주의집중을 하기 위하여 하향식 과정의 메커니즘인 skin color filter를 모델링하기 위하여, 입력 영상 r, g, b 값에 대한 skin color filter의 범위 값을 식 (2)와 같이 정의 하였다<sup>[18]</sup>.

$$\begin{aligned} r > 95, \quad g > 40, \quad b > 20 \quad \text{and} \\ \max\{r, g, b\} - \min\{r, g, b\} > 15 \quad \text{and} \\ |r - g| > 15 \quad \text{and} \quad r > g \quad \text{and} \quad r > b \end{aligned} \quad (2)$$

원추세포로부터 분리되고 skin color filter로 필터링된 빨강색, 녹색 및 파랑색 신호는 수평세포에 의하여 평활되어 R, G, B 신호로 나오고, 쌍극세포를 지나 아마크린세포의 역할로 인하여 Y 신호를 발생 하지만, 본 모델에서는 사람의 피부색과 상이한 B와 Y신호는 배제하고, 사람의 피부색과 유사한 R과 G 신호만을 식 (3)으로부터 구하여 활용하였다.

$$\begin{aligned} R &= \begin{cases} r - \frac{g+b}{2} & R > 0 \\ 0 & R \leq 0 \end{cases} \\ G &= \begin{cases} g - \frac{r+b}{2} & G > 0 \\ 0 & G \leq 0 \end{cases} \end{aligned} \quad (3)$$

또한 망막의 신경절세포의 윤곽정보 검출 기능을 모

방하기 위하여 Sobel필터<sup>[19]</sup>를 사용하여 R/G 색상대비의 윤곽정보를 검출하였다. 식 (4)는 제안 모델에서 사용된 Sobel 필터의 마스크(mask) 모양이다.

$$x_{mask} = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad y_{mask} = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad (4)$$

그리고 인간의 시각 처리 과정중 망막에서 추출한 기본 정보들에 대해 중앙 강화-주변 억제(on-center, off-surround) 그리고 중앙 억제-주변 강화(off-center, on-surround)기능이 신경절 세포에서 시각피질까지 일어나며, 이러한 기능을 모델링하기 위하여 중앙-주변 차(Center-Surround Difference: CSD) 알고리즘을 이용했다. 본 논문에서는 중앙-주변 차 알고리즘 구현 시 실시간 동작을 위하여 처리영상의 크기를 80×60 하였기 때문에 크기 0 ~ 크기 6의 계층을 갖는 가우시안 피라미드 영상을 사용하였으며, 각각의 특징맵(feature map)을 구성하는 각 계층의 피라미드 영상은 원 영상에서의 2n의 서브 샘플링(subsampling)에 의해서 만들어졌으며, 이러한 것은 망막의 비 균일한 분포를 반영한 것이다. 이 때 축소·확대는 이웃 값을 평균하여 구하는 방식을 사용하였고, 축소와 확대 시 왜곡을 부드럽게 하기 위해 전 처리 과정으로 가우시안 필터를 사용하였다. 본 논문에서 사용한 가우시안 필터는 식 (5)와 같고, 식 (5)의 필터를 영상과 컨볼루션(convolution)하면 부드러운 피라미드 영상을 얻을 수 있다.

$$g_{mask} = \begin{pmatrix} 2 & 13 & 2 \\ 13 & 40 & 13 \\ 2 & 13 & 2 \end{pmatrix} \quad (5)$$

가우시안 피라미드 CSD의 과정은 서브 샘플링을 통해 가우시안 피라미드 영상을 7개씩 만들고, 이를 다시 원시 영상의 크기로 확대한다. 이렇게 구해진 가우시안 피라미드 이미지를 이용하여 CSD처리를 한다. 식 (6)은 이 과정을 수식적으로 보여주고 있다. 식 (6)에서, c(center)는 피라미드 영상에서 확대·축소가 적게 되어 있는 선명한 영상을 의미하고, s(surround)는 피라미드 영상에서 확대·축소가 많은 흐린 영상을 의미하며, ‘⊖’ 연산자는 같은 위치의 화소 값 사이의 차이를 의미한다.

$$\begin{aligned}
 I(c, s) &= |I(c) \ominus I(s)| \\
 E(c, s) &= |E(c) \ominus E(s)| \\
 RG(c, s) &= |R(c) - G(s)| \ominus |G(c) - R(s)|
 \end{aligned}
 \tag{6}$$

여기에서는 I는 명암정보, E는 윤곽선 정보, RG는 평활화된 빨강색과 녹색정보의 색상대비를 의미한다. 또한 밝기 특징(I)과 윤곽선 특징(E), RG 색상 대비 특징에 대한 각각의 3레벨의 피라미드 영상에 대하여 식(6)과 아래의 식(7)을 이용하여 밝기(⊖), 윤곽선(⊖), 색상(⊖) 특징맵을 만들 수 있었다.

식(7)에서 “⊕”는 서로 다른 피라미드 위치간의 합연산을 의미하며, N은 정규화 과정을 의미한다.

$$\begin{aligned}
 \bar{I} &= \bigoplus_{c=2}^3 \bigoplus_{s=c+2}^{c+3} N(I(c, s)) \\
 \bar{E} &= \bigoplus_{c=2}^3 \bigoplus_{s=c+2}^{c+3} N(E(c, s)) \\
 \bar{C} &= \bigoplus_{c=2}^3 \bigoplus_{s=c+2}^{c+3} N(RG(c, s))
 \end{aligned}
 \tag{7}$$

그림 4는 특징 맵을 만들기 위한 과정을 보이고 있으며, 이렇게 만들어진 특징 맵의 단순 합으로 돌출 맵이 만들어지고, 이 돌출 맵으로부터 얼굴 후보 영역을 검출하였다.

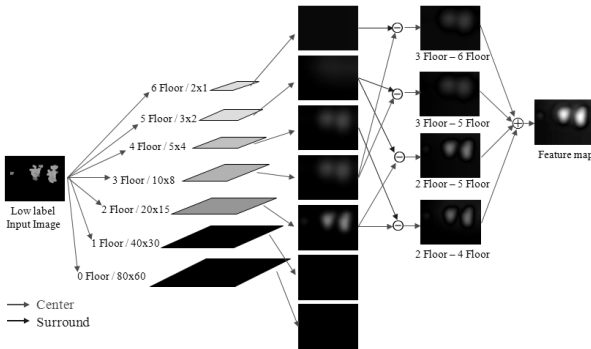


그림 4. 특징 맵의 처리 과정  
Fig. 4. The processing of feature map.

#### 나. Adaboost를 이용한 얼굴영역 판별

Haar-like feature<sup>[20-22]</sup>는 Viola가 얼굴영역 검출에 적용한 것으로 입력 영상의 명암도(Gray) 영상을 단순 합 이미지를 이용하여 특징 값을 표현한 것이다. 그리고 이를 AdaBoost 학습 알고리즘을 사용하여 약 분류기의 인식 자들을 합쳐서 강 분류기 인식자의 그룹으로

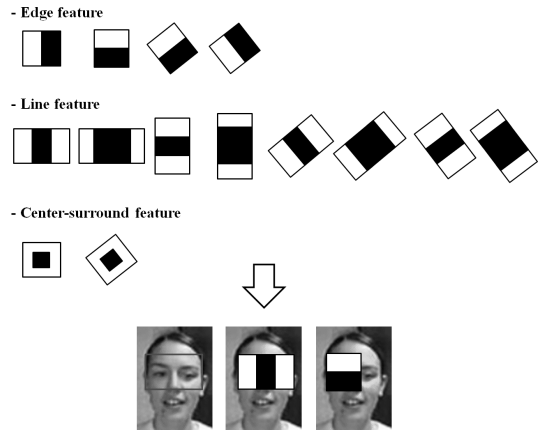


그림 5. Haar-like feature 프로토타입 세트  
Fig. 5. Haar-like feature prototype set.

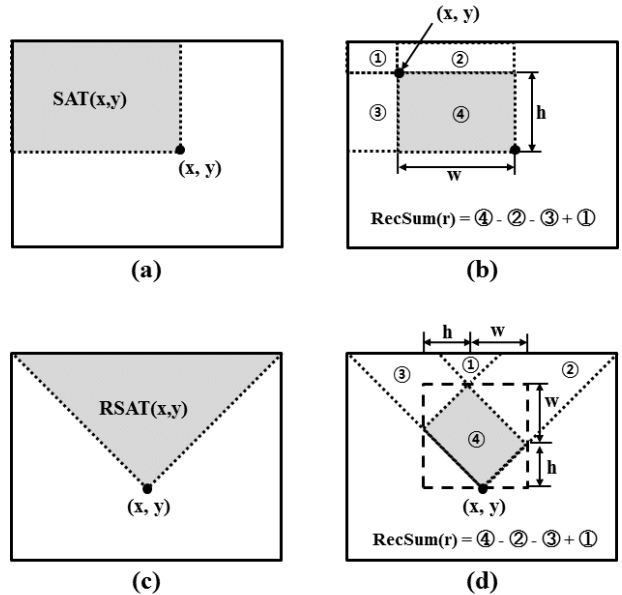


그림 6. Integral 영상, (a) SAT, (b) 사각형내의 화소 합계, (c) RSAT, (d) 회전 사각형내의 화소합계  
Fig. 6. Integral image, (a) SAT, (b) sum of the pixel within the rectangle, (c) RSAT, (d) rotated sum of the pixel within the rectangle.

만들었다. 이 때 생성된 특징 값은 얼굴의 특징을 잘 포함하는 장점을 가지며, 계산 방식이 단순 영역의 합연산(summed area table : SAT)으로 마스크 연산을 수행하므로 빠른 얼굴영역 검출이 가능하다.

Haar-like feature를 통해 얼굴 영역에 대하여 추출되는 값은 그림 5의 흰색 사각형 영역에 포함되는 화소의 합과 검은색 사각형 영역에 포함되는 화소 값의 합에 대한 차이로 나타난다. 그림 5는 얼굴 영역을 검출하기 위해 사용된 세 종류의 Haar-like feature 프로토

타입을 보이고 있다.

그림 6(a)의 SAT는 입력영상의 원점(0, 0)에서부터 좌표 영역(x, y)를 포함하는 사각형 영역의 모든 화소값의 합이며, 식 (8), 식 (9)와 같이 표현된다. 그림 6(b)에서 좌측 상단의 좌표가 (x, y)이고 너비와 높이가 각각 w와 h인 사각형 영역 r의 화소값의 합을 RecSum(r)로 표시한다. RecSum(r)의 값은 SAT를 이용하면 효과적으로 계산될 수 있다.

$$SAT(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (8)$$

$$SAT(x, y) = SAT(x, y-1) + SAT(x-1, y) + I(x, y) - SAT(x-1, y-1) \quad (9)$$

회전각도  $\alpha$ 가 45일 때 회전한 사각형 RSAT (rotated summed area table)는 사각형 영역의 모든 화소값의 합이며, 식 (10), 식 (11)과 같이 계산되고, RSAT를 이용하여 특정 위치에 존재하는 사각형 영역  $r = (x, y, w, h, 45)$ 의 합계 RecSum(r)은 그림 6 (c), (d)에서와 같이 식 (12)에 의해 계산된다.

$$RSAT(x, y) = \sum_{y' < y, y' \leq y - |x-x'|} I(x', y') \quad (10)$$

$$RSAT(x, y) = RSAT(x, y-1) + SAT(x-1, y) + I(x, y) - SAT(x-1, y-1) - RSAT(x, y-2) + I(x, y) + I(x, y-1) - RSAT(x, y-2) + I(x, y) + I(x, y-1) \quad (11)$$

$$RecSum(r) = RSAT(x-h+w, y+w+h-1) + RSAT(x, y-1) - RSAT(x-h, y+h-1) - RSAT(x+w, y+w-1) \quad (12)$$

Adaboost 알고리즘은 Freund와 Schapire<sup>[23-25]</sup>에 의해 소개된 알고리즘으로서 예러율은 약한 분류기의 수가 증가할수록 영점에 지수 적으로 근접함이 증명되었다. 즉, t개의 약한 분류기의 선형적인 결합을 통하여 최종적으로 높은 검출 성능을 가지는 강한 분류기를 생성하는 기법이다. 위에서 추출한 Haar-like는 하나의 약한 분류기가 되고 이는 학습데이터로부터 확률분포를 통해 확률이 높은 데이터를 바탕으로 단계적으로 접근한다. 단계적으로 접근하는 방식으로 Cascade 구조가

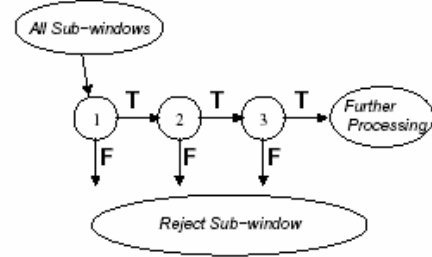


그림 7. Cascade 학습 접근법

Fig. 7. The approach of cascade for learning sequence.

쓰이는데, 이는 TLU(threshold logic unit)와 비슷하며 Cascade 구조는 상위에 높은 확률을 적용하고 순차적으로 낮은 확률을 갖는 Haar-like를 적용한다. 입력 영상에 각각의 Sub-windows가 적용되고 분류기를 거쳐 일정 단계를 거치는 도중 예러는 제거되고 단계별로 제거되어 그 속도 또한 증가한다. 그림 7에서 Cascade 구조로 학습하는 접근법을 보인다.

### 3. 사용자 안구운동 기반의 시선 응시 점 검출

#### 가. 시선 응시 점 정보 파악을 위한 동공영역 검출

시선 추적 카메라에는 반사경을 통해 비춰진 사용자의 눈 영역 전체가 입력영상으로 들어오기 때문에 사용자 시선 응시 점 정보 검출을 위하여 동공 영역만을 검출하는 영상 처리 기법을 사용 하였다. 동공영역을 검출하기 위해 불필요한 영역을 줄여나가는 순차적 방식을 사용하였는데, 눈 영역 검출과 동공 영역 검출로 나누어진다. 사용자의 눈 영역과 동공 검출을 할 때, Template matching<sup>[26]</sup> 기법을 사용하였으며, Template matching은 Template image를 기준으로, 특정 이미지에서 그와 가장 유사한 지점을 찾아주는 방법이다. Template matching은 크기변화와 회전에 약하다는 단점이 있지만, 본 논문에서 제안하는 모델에서는 안경타입의 플랫폼을 사용하기 때문에 시선 추적 카메라를 통해 들어오는 눈 영상은 크기 및 회전에 대한 변화가 거의 없으므로 이에 대한 영향을 배제시킬 수 있었다. 또한 Template matching을 이용해 불필요한 영역은 줄여나가므로 연산 속도나 정확도 면에서 뛰어난 효과를 보이고, 사용자 시선 응시 점 정보를 검출하기 위한 동공 영역 검출에서는 효율적이었다. 이러한 동공 검출 과정을 그림 8에 나타내었으며, 그 과정은 다음과 같다.

시선 추적 카메라를 통해 들어오는 입력 영상은 반사

경에 비추어진 사용자의 눈 영역이기에 눈 영역이 아닌 부분이 포함되어, 선행적으로 만들어 놓은 눈 영역에 대한 Template image를 이용하여, 입력영상에 대하여 Template matching을 수행하여 눈 영역을 우선적으로 검출 하였다. 여기에 사용자의 눈 영역을 검출하는데 사용되는 Template image는 특정 사용자의 눈 영상을 사용하지 않고, 여러 장의 눈 영상들을 겹친 영상을 사용함으로써 Template image의 사용 범용성을 높였다.

다음으로, 검출한 눈 영역 내에서 정확한 동공 영역의 검출을 위한 동공 Template matching을 실시한다. 본 논문에서 제안하는 안경 타입 플랫폼에서 시선 추적 카메라 입력 영상은 사용자 눈에 매우 근접하여 촬영하므로 사용자가 정면을 응시 할 때와 상/하/좌/우를 응시할 때 동공 모양/형태는 다르게 나타난다. 이는 사용자의 안구운동이 회전운동을 이용하여 사용자의 시선을 대상에 맞추기 때문이다. 이와 같은 안구운동의 특징 때문에 본 논문에서는 동공 검출의 신뢰도를 높이기 위해 총 11개의 동공 Template image를 사용하였다.

Template matching을 수행하여 입력 영상의 각 위치에 대한 상관계수 값을 얻은 후, 그 중 상관계수 값이 가장 높은 곳을 matching 지점으로 선택하게 된다. 제안한 방법에서는 11개의 동공 Template image로 각각 Template matching을 하기에 각 Template image에서의 상관계수 값 11개를 얻는다. 이렇게 구한 11개의 matching 지점들 중에서 상관계수의 값이 가장 큰 지점 하나만을 동공영역으로 결정하였다.

Template matching으로 검출된 사용자의 동공 영역에서 정확한 동공 영역만을 검출하기 위하여 검출된 동공 영역에 대하여 이진화, 라벨링, 타원 fitting 등의 영상처리 과정을 수행하였다. 이진화는 검출한 동공영역 내에서 동공 부분과 동공이 아닌 부분을 분리할 수 있도록 동적 이진화 방법을 수행하였다. 이는 적외선(IR; infrared)을 사용 하였기에 동공과 동공이 아닌 영역의 경계가 뚜렷하기에 간단한 동적 이진화 방법으로 구현할수 있었다. 이후 라벨링을 실시하게 되는데, 잡음이 낀 경우가 많으므로 특정 픽셀 이상일 경우에만 라벨링을 수행 하도록 설정하였다. 마지막으로 라벨링 된 동공 영역에 대하여 타원 fitting을 거쳐 타원의 중심 값을 동공의 중심 값으로 선정하여, 동공 좌표를 추출 하였다. 여기서 동공을 원이 아닌 타원으로 가정한 이유는 눈동자 추적 카메라와 눈 사이의 거리가 매우 짧으

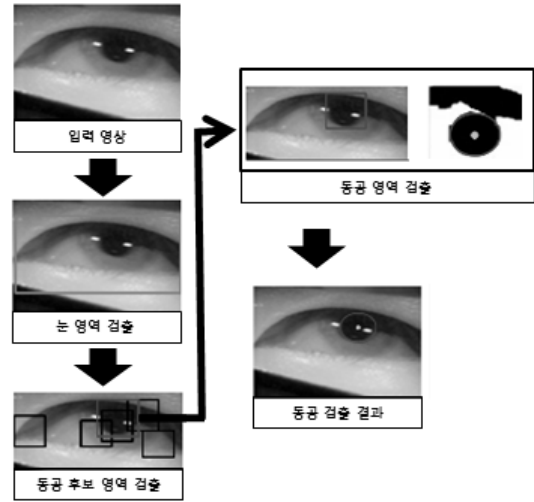


그림 8. 동공 영역 검출 과정  
Fig. 8. The sequence of pupil detection.

므로, 좌우를 응시할 경우 원보다는 타원에 가까워지기 때문이다.

나. Calibration 통한 외부 영역에서의 사용자 시선 응시 점 검출

시선 추적 카메라에서 검출된 사용자의 동공 중심좌표가 외부 영상의 어느 위치를 응시하는지 판단하기 위하여 동공 중심과 외부 영상 사이의 기하학적 위치 관계의 계산이 필요하다. 본 논문에서는 사용자의 시선(동공중심)과 외부 영상의 기하학적 위치관계를 계산하기 위하여 Calibration<sup>[27]</sup> 과정을 거친다. 이는 사용자의 시선 응시 점이 모니터 화면에 랜덤 순서로 발생시킨 9개의 점을 응시하는 것으로 계산되어 진다. 그림 9는 Calibration의 기하학적 변환(Geometric Transform) 방법으로 사용자 시선 응시 점을 2차원 평면모니터에 사상하는 방법을 보이고 있다. 그림 9에서 사용자 시선

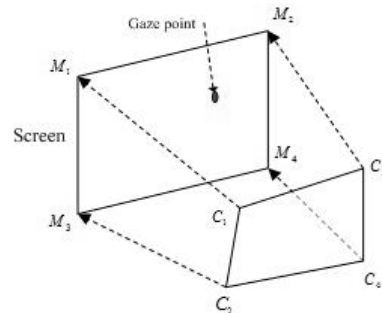


그림 9. 사용자 시선 응시 점 검출 방법  
Fig. 9. The gaze point detection method based on eye movement.

정보인 동공의 중심 좌표는 C이며, 이에 대응되는 외부 영상 스크린(screen)좌표는 M으로 표현하였다.

사용자의 시선 응시 점을 계산하기 위하여 식 (13)을 이용한다.

$$m_x = f(c_x, c_y), \quad m_y = g(c_x, c_y) \quad (13)$$

동공 중심을 보정하기 위한  $m_x$  와  $m_y$ 에 관한 이차 방정식은 식 (14)과 같다.

$$\begin{aligned} m_x &= a_0 c_x^2 + a_1 c_y^2 + a_2 c_x + a_3 c_y + a_4 c_x c_y + a_5 \\ m_y &= b_0 c_x^2 + b_1 c_y^2 + b_2 c_x + b_3 c_y + b_4 c_x c_y + b_5 \end{aligned} \quad (14)$$

이 때,  $a_0 \sim a_5$ 와  $b_0 \sim b_5$ 는 알 수 없는 계수이다.

본 모델은 9개의 Calibration 고정점을 사용한다. 12개의 미지수와 18개의 방정식으로 이루어져 있기에 주어진 방정식을 쉽게 풀 수 있으며, 계수들의 상관관계를 식 (15)과 같은 행렬의 형태로 나타낼 수 있다.

$$\begin{bmatrix} m_{x1} & m_{x2} & \cdots & m_{x9} \\ m_{y1} & m_{y2} & \cdots & m_{y9} \end{bmatrix} = \begin{bmatrix} c_{x1}^2 & c_{x2}^2 & \cdots & c_{x9}^2 \\ c_{y1}^2 & c_{y2}^2 & \cdots & c_{y9}^2 \\ c_{x1}c_{y1} & c_{x2}c_{y2} & \cdots & c_{x9}c_{y9} \\ c_{x1} & c_{x2} & \cdots & c_{x9} \\ c_{y1} & c_{y2} & \cdots & c_{y9} \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_0 & a_1 & \cdots & a_5 \\ b_0 & b_1 & \cdots & b_5 \end{bmatrix} \quad (15)$$

식 (15)의 행렬을 식 (16)과 같이 표현할 수 있다.

$$M = TC \quad (16)$$

따라서 Calibration 행렬 T는 식 (17)과 같이 표현된다.

$$T = MC^{-1} \quad (17)$$

행렬 T 값을 알면 외부 영상에 대한 시선 응시 점인 행렬 G를 결정 할 수 있다.

$$G = TW \quad (18)$$

이 때, W는 Calibration 좌표에 해당하는 입력 행렬로써 동공 중심 좌표 값이다. 그리고 식 (17)과 같이 역행렬을 구하는 과정은 많은 연산량을 필요로 하지만,

사용자 Calibration 한번만 계산하면 되므로, Calibration 이후 실시간으로 계산되는 시선 응시 점 추적의 수행시간에는 문제가 없다.

#### 4. 사용자 시선 응시 점 정보를 통한 관심 영역 기반 얼굴영역 검출 모델

전방 주시 카메라로부터 획득한 외부 영상의 위치 정보와 시선 추적 카메라로부터 획득한 동공의 좌표 값과의 맵핑 방법인 Calibration 과정을 통해 보정하면 외부 영상에 대한 사용자의 시선 응시 점 정보 기반의 관심 영역 정보를 얻을 수 있다. 본 논문에서 제안하는 모델은 전방주시 카메라를 통해 들어오는 영상 정보로부터 생물학 기반의 선택적 주의집중 알고리즘 및 Adaboost 모델을 이용하여 얼굴 영역 검출을 수행한다. 외부 영상에서 다수의 사람 얼굴 영역이 검출 될 수 있으며, 이 때 사용자의 시선 응시 점 정보를 기반으로 관심 있는 사람의 얼굴을 선택 할 수 있다. 또한 원 거리의 멀리 떨어져 얼굴 영상의 크기가 작아 검출이 되지 않는 얼굴 영상에 대해서는 사용자 시선 응시 점을 기준으로 50\*50 픽셀의 주변공간을 8배 확대하여 그 부분에 대한 얼굴 검출을 재 수행 한다. 선간 보간법에 의해 확대하므로 화질은 다소 떨어질 수 있으나, 기존의 Adaboost 방식으로는 검출할 수 없는 먼 거리의 작은 크기의 얼굴 검출도 가능하다.

### III. 실험 및 고찰

#### 1. 생물학 기반의 선택적 주의집중 기반의 얼굴 영역 검출 성능

인간의 시각 처리 메커니즘을 기반으로 본 논문에서 제안하는 생물학적 기반의 선택적 주의집중 모델을 이용하여 얼굴 후보 영역 검출 성능 실험을 수행 하였다. 그리고 기존의 skin color filter 기반 검출 방법과 성능을 비교하였다. 얼굴 후보 영역 검출을 위해서 사용하는 사람의 skin color filter 기반의 검출 모델은 단순한 배경을 가지는 입력영상에서는 얼굴 후보 영역이 잘 검출 될 수 있으나 복잡한 실세계 환경에서는 얼굴영역인데 얼굴 영역을 검출하지 못하거나, 얼굴이 아닌 영역을 얼굴 후보 영역으로 검출되는 오 검출률이 높게 나타난다. 이와 같은 문제점을 해결하고자 본 논문에서는 생물학적 기반의 선택적 주의집중 모델을 이용하여 얼



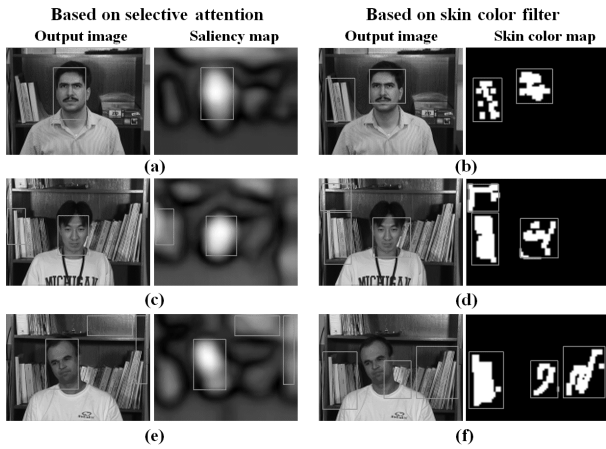


그림 10. 색상 필터 모델과 선택적 주의 집중 모델의 얼굴 후보 영역 선택 결과

Fig. 10. The candidate face detection result from skin color filter model and selective attention model

얼굴 후보영역을 결정하였다. 그림 10에 두 방법에 대한 얼굴 후보 영역 검출 실험 결과들을 나타내었다.

그림 10 (a), (c), (e)는 본 논문에서 제안하는 선택적 주의집중 기반의 얼굴 후보영역을 결정하는 모델의 결과이고, 그림 10 (b), (d), (f)는 skin color filter 기반의 얼굴 후보 영역을 결정하는 모델의 결과이다. 실험결과로부터 실제의 영상 잡음, 조명의 변화, 사람의 피부 색과 비슷한 색상에 대하여, skin color filter는 민감하게 동작하는 반면, 본 논문에서 제안하는 생물학 기반의 선택적 주의집중 모델은 강건한 얼굴 후보 영역의 검출 결과를 보여줌을 알 수 있다.

본 논문에서는 생물학적 기반의 선택적 주의집중 모델의 얼굴 후보 영역 검출 성능을 보이기 위하여 Georgia Tech face database[28]의 120개의 얼굴영역이 포함된 320×240 크기의 영상을 실험에 사용하였다. 표 1에서 Georgia Tech face DB를 이용하여, skin color filter 기반의 얼굴 후보 검출 성능과 제안하는 생물학적 기반의 선택적 주의집중 모델의 얼굴 후보영역 검출 성능을 나타내었다.

표 1로부터 본 논문에서 제안하는 모델의 얼굴 후보 영역 검출 성능이 색상필터 모델보다 우수함을 알 수 있으며, 두 모델에서 False Positive 에러가 높게 나타나는데 이는 조금 더 정확한 얼굴 후보 영역을 결정하기 위함이다. 본 실험에서 발생한 False Positive 에러는 다음 단에서 사용하는 분류기인 얼굴판별 함수에서 얼굴이 아닌 영역은 제거되기 때문에 문제가 되지 않는

표 1. 색상 필터 모델과 선택적 주의 집중 모델의 얼굴 후보 영역 검출 성능 비교

Table 1. The comparison of performance for face candidate regions detection between skin color filter model and selective attention model.

	True Positive	False Positive
Skin color filter	72.5%	75.6%
Selective attention	97.5%	70.1%

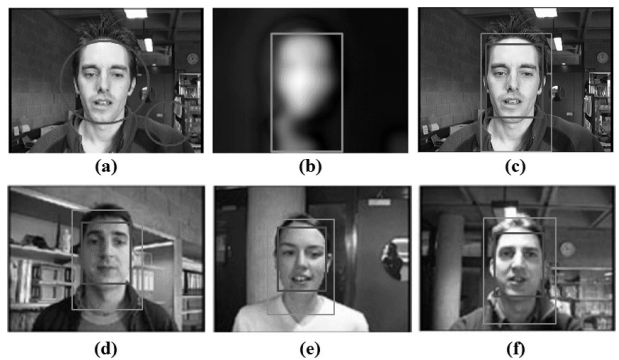


그림 11. 선택적 주의집중 기반의 Adaboost 모델을 이용한 얼굴검출 결과

Fig. 11. Face detection results using the Adaboost model based on selective attention.

는다.

앞서 언급한 생물학 기반의 선택적 주의집중 모델로 선택되어진 얼굴 후보 영역에 대하여 Haar-like feature 기반의 Adaboost 알고리즘을 통하여 최종 얼굴 영역만을 판단한다. 제안된 방법과 기존의 얼굴 검출 방법인 Adaboost 모델의 실험 결과를 그림 11에 나타내었다.

그림 11(a)는 단순히 Adaboost 알고리즘의 얼굴검출 실험 결과이며, 그림 11(b)는 생물학 기반의 선택적 주의 집중 모델로 구한 돌출맵 정보의 얼굴 후보 영역을 보이며, 그림 11(c)는 그림 11(b)의 얼굴 후보영역 내에서 결정되어진 얼굴영역의 결과를 보인다. 그림 11(a)에서는 얼굴 영역이 아닌 부분을 얼굴영역으로 판별하는 오 검출을 했는데 비하여 그림 11(c)에서는 얼굴 영역만을 검출하고 있다. 그림 11 (d)-(f)는 임의의 입력 영상에 대해 적용한 제안된 방법의 얼굴 검출 실험 결과를 나타낸다. 그림 11의 일반화 성능을 보이기 위하여 UCD database[29]의 530개의 얼굴영역이 포함된 360×288 크기의 영상을 사용하였다. UCD DB에 대해 기존의 Adaboost 알고리즘만으로 얼굴영역을 검출한

표 2. 기존의 adaboost 모델과 제안된 모델의 얼굴검출 성능 비교

Table 2. The comparison of face detection performance between the conventional Adaboost model and the proposed model.

	True Positive	False Positive
Adaboost	100%	8.4%
Proposed model	100%	3%

표 3. 기존의 Adaboost 모델과 제안된 모델의 얼굴검출 수행시간 비교

Table 3. The comparison of face detection processing time between the conventional Adaboost model and the proposed model.

	Saliency Map	Adaboost	Total
Adaboost	None	199.8 ms - 263.9 ms	206.4 ms - 270.8 ms
Proposed model	35.7 ms - 60.8 ms	7.75 ms - 240.1 ms	53.9 ms - 262.4 ms

결과와 제안한 모델에서의 얼굴 영역 검출 결과를 표 2에 나타내었다.

표 2에서 보이는 것처럼 두 모델 모두 True Positive 성능은 우수하지만, False Positive 에러와 같은 오 검출률면에서는 제안된 방법이 기존의 Adaboost 모델보다 낮게 나타남을 알 수 있다. 뿐만 아니라, 얼굴검출 수행 시간을 통해 기존의 Adaboost 모델과 제안한 모델의 성능을 비교하였으며, 그 결과를 표 3에 나타내었다. 표 3으로부터 제안된 모델의 얼굴 검출 수행시간이 기존의 Adaboost 모델을 이용한 경우보다 빠름을 알 수 있다.

### 2. Calibration 실험

전방 주시 카메라 통해 들어오는 외부 영상에 사용자의 시선 정보를 적용하기 위해서 전방 주시 카메라 영상과 시선 추적 카메라 영상간의 calibration 작업이 필요하다. 본 논문에서 제안하는 모델은 한번의 calibration 과정을 거친 후, 사용자의 시선 움직임에 의존적으로 만들어진 calibration table을 토대로 사용자의 시선 응시 점을 추적 할 수 있다. 사용자 시선 응시 점 추적의 정확도는 calibration 작업의 정확도에 의존하기 때문에 calibration 작업은 매우 중요한 요소이다. 본 실험

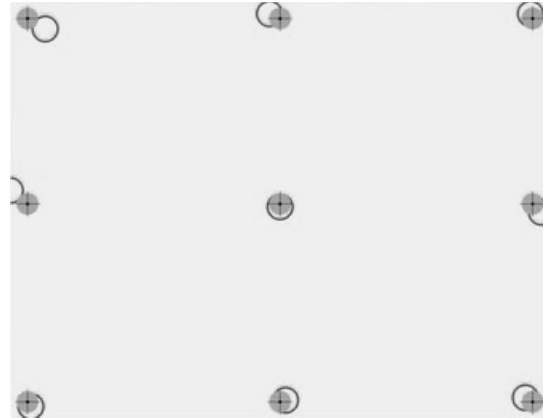


그림 12. Calibration 결과의 예  
Fig. 12. The example of calibration result.

험에서는 사용자가 모니터와 약 50cm 떨어진 거리에서 그림 12와 같이 9개의 점을 응시하는 calibration 과정을 수행하였다.

본 실험의 결과로는 그림 12에서 보이는 것처럼 모니터 상에 출력되는 점(주황색 짝 찬 동그라미)의 좌표와 사용자의 시선 응시 점(붉은색 빈 동그라미)간의 좌표 차이를 측정하여, calibration 작업의 정확도를 계산하였다. 보정 위치간의 차이를 측정하였다. Calibration 정확도 실험에 사용된 모니터는 해상도가 1600\*900 (가로\*세로)인 24인치를 사용하였으며, 50회의 반복 실험을 수행하여 얻은 실험 결과를 표 4와 그림 13에 나타내었다.

Calibration 실험 결과에서 모니터 상에 출력되는 기준 점의 좌표와 calibration 보정 작업 후 사용자 시선 응시 점의 좌표 사이의 평균 오차는 26.88 ±3.37의 오차를 가졌지만, 1600\*900 (가로\*세로)의 해상도를 감안하였을 때 신뢰할만한 수준의 calibration 작업이 이루어졌다고 할 수 있다.

표 4. 각 점에 대한 Calibration 오차 평균  
Table 4. The average of calibration errors for each mark.

Index	1	2	3	4	5
pixel difference	20.59 ±3.49	28.33 ±4.11	23.03 ±2.27	28.02 ±3.48	34.00 ±3.9
Index	6	7	8	9	total
pixel difference	29.66 ±3.11	22.40 ±2.99	36.13 ±4.46	19.76 ±2.53	26.88 ±3.37

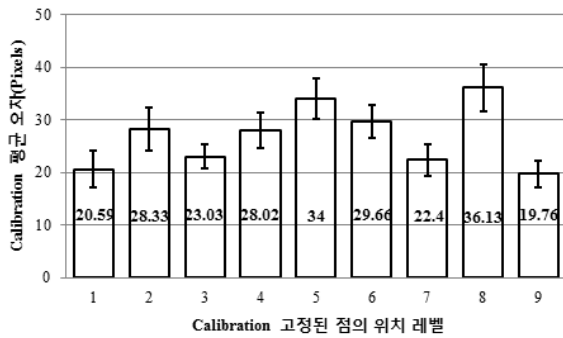


그림 13. 기준위치의 사용자 시선 응시 점 보정 오차  
Fig. 13. The average of calibration errors.

### 3. 사용자 시선 응시 점 기반의 정보선택 실험

본 논문에서 제안한 모델의 사용자 시선 응시 점을 이용한 정보선택의 정확도를 알아보기 위한 실험을 실시하였다. 그림 14와 같이 3m의 거리에 한 변이 30cm인 정사각형의 객체를 상하좌우로 4개를 배치하였다.

그 객체중 하나를 3초간 응시한 후, 그림 14와 같이 외부 화면에 표시된 시선 응시점이 정사각형 객체 안에 들어갔을 때에는 정보 선택이 성공한 것으로 간주하였고, 시선 응시점이 정사각형 객체 밖에 나타났을 때에는 시선 응시가 실패한 것으로 간주하였다.

총 400회 반복 실험하였으며, 그리고 사용자 시선 응시 점 추적 정보는 Calibration 작업에 의존하는 경향이 있으므로 Calibration 1회당 10번의 시선 응시 점 정보 선택 실험을 실시하였다. 또한 본 논문에서 제안하는 사용자 시선 응시 점 추적 모델의 검증에 위하여, tobii사에서 동공추적을 위해 상용화 제품으로 판매하고 있는 Glasses eye-tracker<sup>[30]</sup> 장비를 이용하여 성능 비교 실험을 실시하였다. 표 5에서도 볼 수 있듯이 본 논문에서 제안하는 모델의 정보 선택의 정확도는 87.25%로 tobii사의 Glasses eye-tracker의 정보선택 정확도는

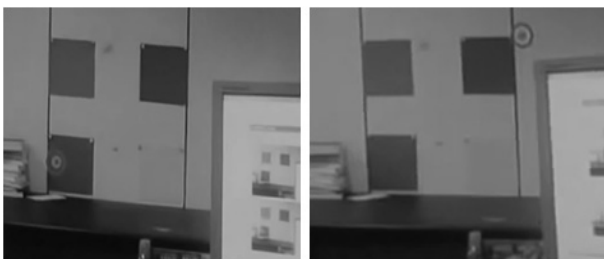


그림 14. 정보선택 실험 성공(왼쪽)과 실패(오른쪽)의 예  
Fig. 14. The example of success (left) and failure (right) of data selection accuracy test.

표 5. 정보선택의 정확도

Table 5. The accuracy of data selection test.

구분	tobii사 Glasses eye-tracker		제안 모델	
	성공	실패	성공	실패
횟수	366 / 400	34 / 400	349 / 400	51 / 400
비율(%)	91.5	8.5	87.25	12.75



그림 15. 얼굴 영역 정보 선택  
Fig. 15. Selective information of face area.

91.5%로 상용 제품과 유사한 결과(4.25% 차이)를 보였다. 앞서 언급한 사용자 시선 응시 점 기반의 정보선택 방법을 실영상에 적용시켜 피험자에게 관심 있는 얼굴을 선택시킨 결과를 그림 15에 나타내었다.

그림 15에서 초록색은 돌출맵 기반의 얼굴 후보 영역이며, Adaboost 알고리즘으로 판별된 얼굴 영역은 분홍색이다. 그리고 두 개의 얼굴 중 시선으로 정보 선택한 얼굴 영역은 붉은색의 사각형으로 표시하였다.

### 4. 시선 응시 점 기반의 관심 영역 확장을 통한 원거리 얼굴 검출 실험

본 논문에서 제안하는 사용자 시선 응시 점 기반의 관심 영역 확장을 통한 원거리의 작은 크기의 얼굴 영역 검출의 성능을 평가를 위하여 거리의 변화에 따른 얼굴영역 검출 실험을 수행하였다. 먼저, 피 실험자와 카메라 사이의 거리 변화에 따라 카메라를 통해 입력되는 얼굴 영상의 크기 변화 실험을 수행하였으며, 이를 그림 16에 나타내었다.

그림 16의 결과처럼 사용자와 피험자 간의 거리가 멀어질수록 전방 주시 카메라로 획득한 영상에서 피 실험자의 얼굴 영상이 차지하는 크기는 점점 작아지므로 얼굴영역 검출이 어려워 질것으로 예상되며, 이때, 사용자 시선 응시 점 정보를 기반으로 관심영역 확장을 통

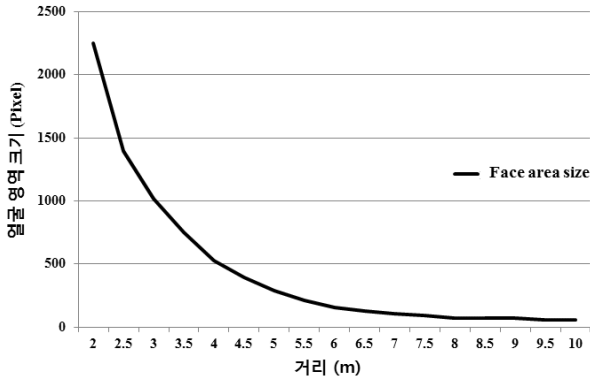


그림 16. 거리 변화에 따른 얼굴영역 크기 변화  
Fig. 16. The variation of face size according to the change of distance.



그림 17. 거리 및 얼굴크기 변화에 따른 시선 응시 점 기반의 얼굴검출  
Fig. 17. Face detection based on gaze spot information according to the change of distance and face size.

한 얼굴 영역 검출의 성능을 알아보기 위해 사용자와 피험자 간의 거리를 2m에서 시작하여 0.5m 간격으로 멀어지면서 얼굴영역이 검출되는지를 평가하는 실험을 수행 하였다.

그리고 본 실험에서는 각 거리 당 50개의 이미지를 저장하여 저장된 이미지에서 얼굴 크기에 따른 Haar-like feature 기반의 Adaboost 알고리즘의 얼굴 검출 성능을 평가 하였다. 본 실험의 얼굴영역 검출 성능을 그림 17과 표 6에 보인다. 거리가 가까울 때에는

표 6. 거리 및 얼굴크기 변화에 따른 얼굴검출 성능  
Table 6. The comparison of face detection performance according to the change of distance and face size.

거리(m)	2.0	2.5	3	3.5	4	4.5	5	5.5	6
얼굴크기 (가로*세로)	45*50	35*40	30*34	25*30	21*25	18*22	16*18	14*15	12*13
Adaboost (%)	100	100	100	100	100	60	16	0	0
제안모델 (%)	100	100	100	100	100	96	96	94	94

거리(m)	6.5	7	7.5	8	8.5	9	9.5	10
얼굴크기 (가로*세로)	11*12	10*11	9*10	8*9	8*9	8*9	7*8	7*8
Adaboost (%)	0	0	0	0	0	0	0	0
제안모델 (%)	90	84	94	96	92	50	46	0

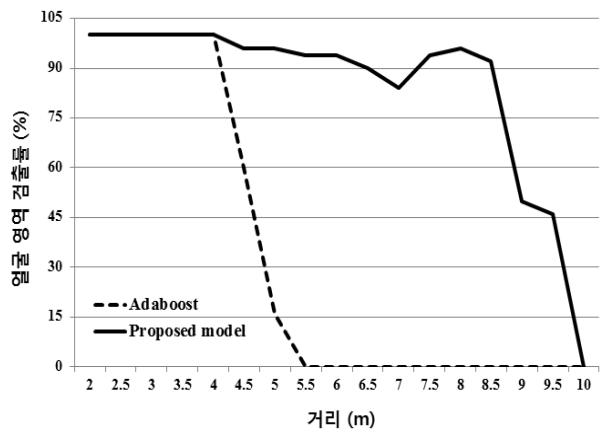


그림 18. 거리에 따른 얼굴검출 성능 비교  
Fig. 18. The comparison of face detection performance according to the distance.

기존의 Haar-like feature 기반의 Adaboost 알고리즘과 본 논문에서 제안한 모델 전부 잘 검출 되었다. 하지만 거리가 4.5m 이상 얼굴 영역의 크기가 18\*22 (가로\*세로)/(total pixel : 396) 보다 작아 졌을 경우 기존의 얼굴 검출 방법의 검출 성능 급격히 떨어져 원거리의 얼굴검출을 하기에는 어려운 점이 있다.

이에 반해 본 논문에서 제안한 모델의 경우 일정 거

리 이상으로 멀어지더라도 최대 9.5m까지 사용자 시선 응시 점 정보를 이용하여 관심 영역을 확대 한 후 얼굴 영역을 검출함으로써 원거리의 얼굴영역을 검출 할 수 있었다.

거리에 따른 얼굴 검출 실험 결과로 기존의 얼굴 검출 방법과 본 논문에서 제안하는 시선 응시 점 정보 기반의 관심 영역 확장을 통한 원거리 얼굴 검출성능을 그림 18에 그래프로 나타내었다.

#### IV. 결 론

본 논문에서는 기존의 얼굴검출 방법에 생물학 기반의 선택적 주의집중 모델과 사용자의 시선 응시 점 정보를 결합한 원거리 얼굴 검출 모델을 제안하였다. 또한, 안경 타입의 시선 추적 및 얼굴 검출을 위한 외부 카메라가 장착된 플랫폼을 제작하였는데, 전방 주시 카메라를 통해 획득한 영상은 선택적 주의 집중 모델을 적용하여 전체 영상에서 후보 얼굴 영상 영역을 선정하였으며, 후보 영역 중 Haar-like feature 기반의 Adaboost 알고리즘을 이용하여 최종 얼굴 영상을 검출함으로써 얼굴검출에 소요되는 연산시간을 줄였다.

그리고, 두 대의 카메라를 통해 획득되는 사용자의 시선 응시 점 정보와 외부영상의 좌표 정보의 맵핑 과정을 통해 사용자가 관심이 있는 영역을 찾을 수 있도록 하였고, 사용자의 시선 응시 점 정보는 카메라로부터 원거리에 떨어져 있는 작은 크기의 얼굴 영상도 응시점 주변을 선형 보간법으로 확대하여 입력영상으로 사용함으로써 최대 9.5m 떨어진 얼굴크기 7\*8 (가로\*세로)/(total pixel : 56)화소를 가지는 작은 얼굴 영역도 검출 할 수 있었다. 이러한 원거리 얼굴 검출 시스템은 사용자가 멀리 떨어져 있는 객체(얼굴)등의 정보를 알려고 할 때, 적응적으로 사용자의 관심영역을 파악하여 선행적으로 판별한 후, 사용자에게 알려주는 인지향상 시스템에 응용될 수 있다. 그리고 거리 변화뿐 아니라 얼굴의 기울임, 회전과 같은 다양한 왜곡에 대해서도 견실하게 얼굴 영역을 검출 할 수 있는 방법과 제안된 모델을 임베디드 시스템으로 구현하는 방법에 대한 연구를 진행하고 있다.

#### 참 고 문 헌

- [1] 정재찬, 신호철, 황대환, 조재일, “피부색 필터를 이용한 실시간 얼굴 검출기”, *대한전자공학회 하계 학술대회* 제 33권 1호, 253-254쪽, 2010년
- [2] 백영현, “피부색 모델 기반의 효과적인 얼굴 검출 연구”, *대한전자공학회 논문지-SP*, 45SP-6-5, 38-43쪽, 2008년
- [3] 황용원, 오상록, 유범재, 이지용, 박민용, 정문호, “컬러 정보를 이용한 신경 진동망 기반 얼굴 추적 방법”, *대한전자공학회 논문지-SC*, 48SC-2-7, 100-106쪽, 2011년
- [4] P. Viola, M. J. Jones, “Robust Real-Time Face Detection”, *International Journal of Computer Vision* 57(2) Kluwer Academic Publishers. Manufactured in The Netherlands, pp. 137-154, 2004.
- [5] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20(11), pp. 1254-1259, 1998.
- [6] J. Duncan, “Selective attention and the organization of visual information”, *Journal of Experimental Psychology: General*, 113, pp. 501-517, 1984.
- [7] S. Jeong, S.-W. Ban, and M. Lee, “Stereo Saliency Map Considering Affective Factors and Selective Motion Analysis in a Dynamic Environment”, *Neural Networks*, 21, pp. 1420-1430, 2008.
- [8] E. B. Goldstein, “Sensation & Perception, 4th edition,” *An international Thomson publishing company*, USA, 1995.
- [9] B. Hwang, C.-S. Kim, H.-M. Park, Y.-J. Lee, M.-Y. Kim, and M. Lee, “Development of Visualizing Earphone and Hearing Glasses for Human Augmented Cognition”, *LNCS 7063*, Springer, pp. 342-349, 2011.
- [10] A. M. Treisman, and G. Gelde, “A feature-integration theory of attention,” *Cognitive Psychology*, vol. 12, no. 1, pp. 97-136, 1980.
- [11] O. Ramström, and H. I. Christensen, “Visual Attention Using Game Theory,” *BMCV 2002*, LNCS 2525, pp. 462-471, 2002.
- [12] Z. Li, “A saliency map in primary visual cortex,” *Trends in Cognitive Sciences*, vol 6, no. 1, pp. 9-16, 2002.
- [13] B. Scassellati, “Investigating models of social development using a humanoid robot,” *in Proc.*

- IJCNN2003*, pp. 2704-2709, 2003.
- [14] C. Breazeal, A. Edsinger, P. Fitzpatrick, and B. Scasselati, "Active vision for social robots," *IEEE Trans. System, Man, and Cybernetics-Part A: Systems And Humans*, vol. 31, no. 5, 2001.
- [15] F. Orabona, G. Metta, and G. Sandini, "Object-based Visual Attentions: a Model for a Behaving Robot," *3rd International Workshop on Attention and Performance in Computational Vision*, 2005.
- [16] S.-W. Ban and M. Lee, "Biologically Motivated Visual Selective Attention for Face Localization", LNCS 3368 pp. 196-205, 2005.
- [17] W.-J. Won, Y.-M. Jang, S.-W. Ban, and M. Lee, "Biologically Motivated Face Selective Attention Model", LNCS 4984, pp. 953-962, 2008.
- [18] Kovač, J., Peer, P., Solina, F.: "Human skin colour clustering for face detection", *EUROCON 2*, pp. 144 - 48, 2003.
- [19] Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing*, 2ed edition, Prentice Hall, USA, 2002.
- [20] P. Viola and M.J. Jones, "Robust real-time object detection", *Technical Report Series, Compaq Cambridge research Laboratory*, 01, Feb. 2001.
- [21] 김석호, 김재민, 조성원, 이기성, 정선태, "Haar-like Feature 변형을 이용한 기울어진 얼굴 검출", *대한전자공학회 하계종합 학술대회* 31권 제 1호, 987-988쪽, 2008년
- [22] 배정민, 이영현, 송태엽, 구분화, 전승선, 고한석, "Edge와 Intensity 기반의 특징을 이용한 얼굴 검출", *대한전자공학회 하계종합 학술대회*, 967- 968 쪽, 2009년
- [23] Yoav Freund and Robert E. Schapire., "A decisiontheoretic generalization of on-line learning and an application to boosting." *Journal of Computer and System Sciences*, 55(1): pp. 119 - 139, 1997.
- [24] Robert E. Schapire and Yoram Singer. "Improved boosting algorithms using confidence-rated predictions", *In Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pp. 80 - 91, 1998.
- [25] Robert E. Schapire, Yoav Freund, Peter Bartlett, and Wee Sun Lee. "Boosting the margin: A new explanation for the effectiveness of voting methods", *In Machine Learning: Proceedings of the Fourteenth International Conference*, 1997.
- [26] R. Brunelli, "Template Matching Techniques in Computer Vision: Theory and Practice", Wiley, ISBN 978-0-470-51706-2, 2009.
- [27] Z. Ramdane Cherif, A. Nait-Ali, J. Motsch, M. Krebs, "An adaptive calibration of an infrared light device used for gaze tracking", *in Proc. of the IEEE Instrumentation and Measurement Technology Conference, Anchorage, AK*, pp. 1029 - 1033, 2002.
- [28] Georgia Tech face database, [http://www.anefian.com/research/face\\_reco.htm](http://www.anefian.com/research/face_reco.htm)
- [29] UCD Valid Database, <http://ee.ucd.ie/validdb/datasets.html>
- [30] Tobii glasses eye-tracker, <http://www.tobii.com/>

저 자 소 개



박 희 선(학생회원)  
2008년~현재 경북대학교  
전자공학부 재학  
<주관심분야 : 영상 신호처리>



장 영 민(학생회원)  
2006년 경일대학교 정보통신  
공학과 공학사  
2008년 경북대학교 전자전기  
컴퓨터공학부 공학석사.  
2011년 경북대학교 전자공학과  
박사수료  
<주관심분야 : 신경망, 시-청각 융합 정보처리,  
HRI, 뇌공학, 사람 의도 인식>



홍 장 표(학생회원)  
2008년~현재 경북대학교  
전자공학부 재학  
<주관심분야 : HCI, 인터페이스>



김 철 수(정회원)  
1989년 경북대학교 전자공학과  
공학사  
1991년 경북대학교 전자공학과  
공학석사  
1997년 경북대학교 전자공학과  
공학박사

1995년~1998년 김천대학교 전자통신과 전임강사  
1998년~2010년 경주대학교 컴퓨터정보과 부교수  
2011년~현재 경북대학교 IT대학 강사  
<주관심분야 : 신호처리, 3차원 디스플레이, BCI,  
HCI, 뇌공학>



김 상 열(학생회원)  
2012년 경북대학교 전자공학과  
공학사  
2012년~현재 LIG넥스원 연구소  
<주관심분야 : 신경망, 신호처리>



이 민 호(정회원)-교신저자  
1988년 경북대학교 전자공학과  
공학사  
1992년 한국과학기술원 전기 및  
전자공학과 공학석사  
1995년 한국과학기술원 전기 및  
전자공학과 공학박사

1995년~1998년 한국해양대학교 조교수  
1998년~현재 경북대학교 IT대학 전자공학부  
교수

<주관심분야 : 영상 및 음성 신호철, 생물학기반  
선택적 주의집중, 뇌-기계 상호작용, 뇌공학>