

움직이는 관찰자의 두 눈 위치 검출 알고리즘

The Position Tracking Algorithm of Moving Viewer's Two-Eyes

허경무, 박영빈

(Kyung-Moo Huh and Young-Bin Park)

Abstract : Among the several types of 3D display methods, the autostereoscopic method has an advantage that we can enjoy a 3D image without any additional device, but the method has a disadvantage of a narrow viewing zone so that the moving viewer cannot see the 3D image continuously. This disadvantage can be overcome with the detection of viewer's positional movement by head tracking. In this paper, we suggest a method of detecting the position of the moving viewer's two eyes by using images obtained through a color CCD camera. The suggested method consists of the preprocessing process and the eye-detection process. Through the experiment of applying the suggested method, we were able to find the accurate two-eyes position for 78 images among 80 sample input images of 8 different men with the processing speed of 0.39 second/frame using a personal computer.

Keywords : autostereoscopic 3D display, eye-position detection, moving viewer, CCD camera

1. 서론

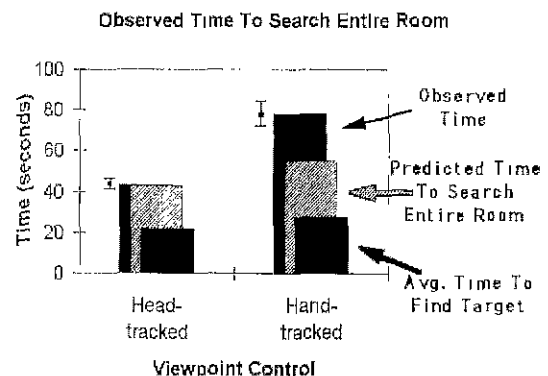
3차원 영상 매체 기술은 고도 정보화 사회에서 없어서는 안될 매우 중요한 원천 과학 기술의 하나로 이미 선진국들의 주요 연구 과제가 되어 개발이 활발히 진행되고 있다. 이 분야는 편광판이나 컬러 필터, 그리고 동기 신호가 주어지는 액정 셔터를 이용한 안경 방식의 입체 영상 방식과 렌티큘러(lenticular)와 액정을 이용한 무안경식의 입체 TV까지 수많은 결과물이 있고 또한 홀로그래픽 스크린, 홀로그래픽 비디오 등 입체 차원을 벗어난 3차원 영상 구현 방법도 개발되고 있다. 그러나 입체 방식은 시야각 확대, 다자 시청, 피로감, 깊이감 등 아직도 해결해야 될 문제가 산재해 있으며, 3차원 영상 구현의 경우는 신호추출, 정보량 처리, 실시간 신호처리, 표시장치, 기록장치 등 해결하기 쉽지 않은 수많은 문제들이 있다. 이와 같이 3차원 디스플레이 방법은 크게 나누어, 양쪽 눈에 서로 다른 화상을 보임으로써 입체 영상을 구현하는 방법과 공간상에 3차원 영상을 구현해 내는 방법의 2가지로 분류할 수 있는데, 그 중 가장 많이 쓰이는 방법이 양쪽 눈에 서로 다른 화상을 보이게 하되 시역(viewing-zone)에 위치한 인간의 눈에서는 완벽한 입체 영상이 되도록 하는 전자의 방법으로서 이의 대표적인 방법이 오토스테레오스코픽(autostereoscopic) 방법이다. 이러한 오토스테레오스코픽 방법은 고성능의 하드웨어나 어떠한 부가적인 장치 없이도 입체 영상을 즐길 수 있다는 장점을 가지고 있으나, 입체 영상을 즐길 수 있는 가시 영역이 매우 협소하고, 좌우가 바뀌는 영역이 존재하는 것 등이 가장 큰 문제점으로 알려져 있다. 즉 현재로서는 오토스테레오스코픽의 가장 큰 문제점은 일정하게 정해진 위치에서만 입체 영상을 감상할 수 있다는 것으로서, 움직이고 있는 사람은 입체 영상을 볼 수 없다는 점이다.

이러한 문제점을 보완하기 위한 방법으로 관찰자의 움직임을 측정해 내는 트래킹(tracking) 기술에 의한 입체 영상의 위치 이동방식에 대한 관심이 높아지고 있다. 즉, 움직이고 있는 사람이 계속 입체 영상을 볼 수 있게 하기 위해서는 움직임을 트래킹하여 움직임에 따른 입체 영상의 보상을 실시간적으로 수행시켜야 한다. 또한 홀로그램이나 가상현실 등의 응용에서 관찰자의 위치 이동에 따른 표시 영상의 변화도 구현할 수 있어 그 활용도가 매우 높다. 트래킹 방법으로는 크게 핸드 트래킹(hand-tracking)과 헤드 트래킹(head-tracking)의 두 방법을 사용할 수 있으며, 표 1은 이 두 방법간의 반응 속도를 비교한 결과이다. 그래프에서 볼 수 있듯이 물체를 관찰하고 예측하며 목표 물체를 찾는 데 걸리는 시간이 크기는 두 배까지 차이가 나는 것을 볼 수 있다. 물론 최근에 사용되는 장비를 이용할 경우에는 표 1의 결과보다 더 좋은 결과가 나올 수 있다.

헤드 트래킹을 구현하는 방법에는 여러 방법이 있을 수

표 1. Hand-tracking과 head-tracking의 수행 속도 비교.

Table 1 The comparison of execution time between hand-tracking method and head-tracking method.



접수일자 : 1999. 11. 30., 수정완료 : 2000. 4. 13.

허경무, 박영빈 · 단국대학교 공학과 전자공학전공

※ 본 연구는 한국과학재단 핵심전문연구(과제번호 981-0910-042 2) 지원으로 수행 되었음.

있었으나, 본 연구에서는 영상 입력 장치를 통해 들어오는 영상 정보를 분석, 비교하여 관찰자의 움직임을 추적하여 그에 상응하는 결과를 내놓는 방법을 선택하였다. 이러한 헤드 트래킹 방법으로는 Birchfield[1], Valentin[2], 그리고 Weng[3] 등이 제안한 여러 방법들이 제안되어 왔는데, 이들은 모두 인간의 얼굴을 검출하는 방법을 연구하였다. 특히 [1]에서는 얼굴을 타원형으로 모형화하여 타원 내부의 색상 히스토그램(color histogram)을 이용하여, 얼굴의 크기와 위치를 검출하는 방법을 제안하였다. 또한 Rowley는 [4]에서 동시에 여러 얼굴을 검출해 낼 수 있는 알고리즘을 제시한 바 있으나, 실시간 구현 가능성 및 위치 정밀도 면에서 상당히 뒤떨어지는 성능을 보이고 있다. 인간의 얼굴을 중점적으로 검출하는 알고리즘으로서, [5]에서는 얼굴 영역 검출과 얼굴의 특징점 마킹, 그리고 특징점 트래킹을 단계적으로 적용하여 얼굴을 트래킹하는 방법을 제안하였는데, 얼굴의 특징점 추출에 있어서는 STK (Shi-Tomasi-Kanade) 트래킹 알고리즘을 적용하여 얼굴 영역에 경계(landmark)를 나누어 눈, 입술의 끝, 콧구멍 등 얼굴의 특징점들을 추출하였다. 또한 [6]에서는 영상을 입력받아 웨이블릿(wavelet) 변환을 하고 영상을 저역 통과(low-pass)와 고역 통과(high-pass) 요소로 나눈 후, 에지의 히스토그램을 이용하여 얼굴의 위치를 찾은 다음 저역 통과와 고역 통과 영역에서의 얼굴 영역을 추출하였는데, 이 때 영상에서 얼굴 벡터(face vector)를 분류하기 위해 KL 변환과 FCM(Fuzzy C-Means) 알고리즘을 적용하였다. 그리고 [7]에서는 각 개인이 가지고 있는 얼굴의 특성을 비교함으로써 복잡한 배경 하에서도 얼굴을 트래킹할 수 있도록 하였는데, 얼굴 검출에 있어서 얼굴 영역 이외의 영역을 제거하기 위해 피부색에 민감하도록 RGB 모델 이미지를 YIQ 모델로 변환하고 차영상으로부터 얼굴 영역을 검출하였다. 그리고 이 밖에도 다수의 연구들이 관찰자의 얼굴을 중심으로 위치 추적을 다루고 있다[8][11]. 특히 인간의 얼굴을 배경 화면과 분리해 내는 방법에 대해서도 여러 가지 방법으로 연구되고 있는데, 다각형을 이용하여 얼굴의 기하학적 정보를 추출하여 얼굴을 추적하는 방법과 본 라인(bone-line)을 이용한 방법, 그리고 색을 이용한 방법과 텍스처를 이용한 방법 등 여러 가지 방법이 존재하고 있다. 지금까지의 대부분의 연구들은 워크스테이션을 1대에서 3대까지 이용하여 얼굴을 추적할 정도로 많은 연산량을 요구하여 왔다.

이와 같이 인간의 얼굴은 매우 복잡한 패턴을 갖고 있기 때문에 얼굴의 특징을 추출하고 규정하기는 상당히 어렵다. 그러나 눈은 얼굴 내에서 가장 뚜렷하고 단순한 특징을 가지고 있으므로 관찰자의 움직임을 추적하는데 가장 중요한 정보가 된다. 더구나 오토스테레오스코피 방식의 3차원 영상의 감상에 있어서는 관찰자의 두 눈의 위치가 직접적으로 관계되고 절대적인 역할을 하고 있으므로, 얼굴 자체를 검출하는 것보다는 두 눈의 위치를 직접 검출하는 것이 중요하고 필요하다. 즉 관찰자의 두 눈의 위치를 추적하는 것이 관찰자의 움직임에 따른 입체 영상의 실시간 보상을 위해 반드시 선행되어야 할 기술인 것이다. 그러나

위에서 언급한 대부분의 문헌들은 얼굴을 검출하여 추적하는 문제를 다루었다.

3차원 영상을 위해 눈의 위치를 검출하기 위한 방법으로 [12]에서는 관찰자의 눈 바로 위에 자기 위치 센서(magnetic position sensor)를 부착하여 위치를 검출하고 그에 따라 영상 이동 장치(image shifting optics)를 구동시키는 방법을 시도하였다. 그러나 이 방법은 그 시스템 구성이 복잡하고 한 쪽 눈의 위치만 검출함으로 인해 정확도가 떨어지는 단점이 있었다. 그리고 [13]에서는 관찰자가 카메라를 보고 있는 상태에서 관찰자의 눈을 검출하는 알고리즘을 제시하였는데, 흑백 카메라를 사용하여 흑백 영상으로 수행하였고 워크스테이션을 사용하여 구현하였으며, 눈 검출 성공률은 91.7%를 기록하였다. 또한 [14]에서는 컬러 영상으로부터 두 가지의 컬러 정보(피부색, 머리색)와 에지 정보를 사용하는 적분 투영법(integral projection method)을 적용하여 얼굴과 특징점을 검출하고, 얼굴내 특징점 영역의 위치는 Snake 모델(active counter model)을 이용하여 얼굴내 특징점들(눈, 눈썹, 입, 코)의 궤적(contour)을 구하는데 까지 수행하였는데, 그 처리시간은 한 프레임당 20초가 걸렸다. 그리고 [15]에서는 퍼지 패턴 매칭을 사용하여 이미지에서 얼굴과 유사한 영역을 검출하고, 얼굴의 특징점들이 그 안에 있는가를 체크하여 얼굴 후보 영역을 검사한 후, 얼굴의 특징점을 추출하였는데, 좌우 두 눈의 중심 좌표와 입의 중심 좌표를 구하였다. 이 실험 결과에서는 80장의 정지 영상을 실험하여 약 90%의 성능을 보였으며, 얼굴을 검출하는데 까지만 한 프레임당 2초가 걸렸다.

이 점에 착안하여 본 논문에서는 고성능의 하드웨어 장비가 아닌, 일반적으로 사용되는 개인용 컴퓨터에서도 인간의 두 눈을 찾아내고 관찰자의 움직임을 추적할 수 있도록 하는 알고리즘을 제안하고, 이를 적용한 실험 결과를 통해 관찰자의 두 눈 위치 검출 속도와 검출 성공률이 매우 향상되었음을 보인다. 본 실험에서는 관찰자가 카메라를 보고 있으며, 관찰자와 카메라와의 거리는 통상적인 관찰거리인 1.50m에서 1.85m의 범위 내이고, 관찰자는 검은 눈동자를 가진 동양인이라는 조건하에 수행하였으며, 이러한 조건들이 본 방법의 적용 한계이다.

II. 눈 검출 알고리즘

눈 검출 알고리즘은 크게 전처리(preprocessing) 과정과 두 눈의 중심의 픽셀 좌표를 얻는 눈 검출과정으로 나눌 수 있다.

1 전처리 과정

전처리 과정은 입력 영상으로부터 얼굴 영역의 위치를 결정하는 얼굴 영역 검출 과정과 조명 상태를 판단하여 알맞은 문턱(threshold) 값을 결정하는 과정, 그리고 얼굴 영역의 적색(red) 값의 평균 정규화 색상 계산 과정으로 나누어 수행하였다.

1.1 색상과 조명

색상 움직임을 이용한 얼굴 영역 추출 방법은 색상이 입력 영상에서 관심의 대상이 되는 물체의 영역을 추출해 내

는데 중요한 역할을 한다는 점에 착안하여 개발한 방법이다. 모든 색상은 R, G, B의 서로 다른 파장의 길이를 가지는 세 가지 성분의 조합과 빛의 휘도(brightness) 성분으로 표현되며, 일반적으로 비디오 카메라는 RGB 모델을 많이 사용하고 있으며 이 모델은 다른 컬러 모델로 쉽게 변환이 가능하다. 비디오 카메라나 프레임 그레버는 아날로그 신호를 각각의 화소에 대해 세 개씩의 레벨값[0-255]으로 변환한다. 즉, 한 화소는 각각 [0-255] 사이의 값을 가지는 R, G, B의 조합으로 표현되며, 이 모델은 하나의 색상뿐만 아니라 빛의 휘도까지 포함된 값이다. 따라서 이런 빛의 휘도 성분을 제외시킨 정규화 색상 모델을 사용한다. 만약 두 점 $Q_1=[r_1, g_1, b_1]$ 과 $Q_2=[r_2, g_2, b_2]$ 가 색상에 있어 (1)과 같은 관계를 가지고 있다면, 이 두 점은 같은 색상을 가지고 있으나 다른 휘도 값을 가지고 있다는 것을 의미한다.

$$\frac{r_1}{r_2} = \frac{g_1}{g_2} = \frac{b_1}{b_2} \quad (1)$$

색상을 이용하여 일련의 작업을 하려고 할 때 큰 문제로 나타나는 것이 조명의 영향이다. 조명의 변화에 색상은 영향을 받지 않으나 휘도는 크게 영향을 받으므로, 이러한 조명에 의한 영향을 제거시킬 수 있는 정규화 과정을 거칠 필요가 있다. 즉 (2)와 (3)의 정규화 과정을 거치게 되면 순수한 컬러 정보만을 지니게 된다. (3)에서 알 수 있듯이 r, g, b의 값은 비례적인 값을 나타낸다.

$$\begin{aligned} r &= \frac{R}{R+G+B} \\ g &= \frac{G}{R+G+B} \\ b &= \frac{B}{R+G+B} \\ r+g+b &= 1 \end{aligned} \quad (2)$$

$$(3)$$

본 연구에서는 후에 얼굴 영역의 적색(R) 값의 평균 정규화 색상을 구하는 부분에 이를 적용하였다.

1.2 얼굴영역 검출 과정

본 연구에서는 순수 칼라(true color)의 영상 이미지를 사용하였다. 순수 칼라인 경우 한 픽셀당 R, G, B의 3원색의 색상정보로 구성되어 있으며, R, G, B 3색의 각 레벨값이 작을수록 색은 검정색에 가까워진다. 여기에서 한 픽셀의 R 레벨값을 R_v , G 레벨값을 G_v , B 레벨값을 B_v 로 표기하였다.

본 얼굴 영역 검출 과정에서는 입력 영상에서 피부색과 가까운 부분만을 남기고 모두 제거한다. 즉 실험을 통해 얻어진 다음의 피부색 조건에 만족하는 부분만을 남기고 나머지는 제거한다.

<피부색 조건>

$$R_v > G_v, R_v > B_v, B_v > G_v$$

이렇게 처리된 이미지를 보면 얼굴 영역뿐만 아니라 다른 여러 부분들이 남게 되는데, 피부색 조건에 만족하는 픽셀이 가장 많이 그리고 넓게 분포되어 있는 부분을 얼굴 영역으로서 검출하고 검출된 영역의 상하 좌우의 얼굴영역 좌표를 구한다. 보다 구체적으로는 피부색 조건에 만족

하는 픽셀들이 가로, 세로 모두 크기가 64픽셀 이상이 되는 곳부터 시작하여 형성되는 사각형으로서 얼굴영역을 설정하였다. 그림 1은 검출된 얼굴 영역을 보여준다.

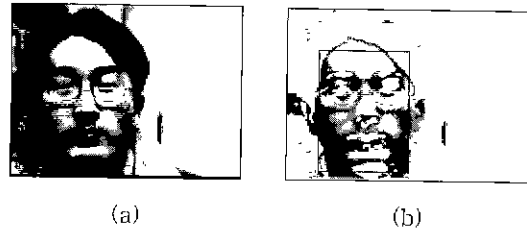


그림 1 얼굴 영역 검출 결과. (a) 입력 이미지, (b) 얼굴 영역 검출 결과.

Fig. 1. Detection result of face zone. (a) Input image, (b) Detection result of face zone.

1.3 Threshold 레벨 값 결정 과정

한 픽셀의 R, G, B 레벨 값은 색상 정보뿐만 아니라 빛의 휘도까지 포함하고 있다. 따라서 입력 이미지는 조명의 상태에 큰 영향을 받는다. 이 점을 보완하기 위하여 얼굴영역 검출 과정에서 얻은 이미지로부터 조명상태를 파악한다. 얼굴 영역 안의 피부색 픽셀들을 그레이 레벨로 변환하여 색상 성분을 없애고 조명의 휘도 정보를 찾아낸다.

$$Level = \frac{1}{i} \sum_{k=1}^i \frac{(R_{vk} + G_{vk} + B_{vk})}{3} \quad (4)$$

여기에서 Level은 얼굴 전체의 평균 Gray 레벨(0-255)을 나타내고, i는 얼굴영역 내부의 피부색 픽셀의 수를 의미하며, R_{vk}, G_{vk}, B_{vk} 는 얼굴 영역 내부의 피부색 픽셀들의 R, G, B 레벨값을 나타낸다.

1.4 얼굴 영역의 적색(R) 값의 평균 정규화 색상 계산

2.2절의 두 눈의 위치 결정 부분에서 입을 검출하는데 사용하기 위하여 우선 얼굴영역의 적색 값의 평균 정규화 색상을 구한다. 다음 (5)가 그 계산 과정이며, 여기에서 Q_{red} 가 얼굴 영역의 적색 값의 평균 정규화 색상이다.

$$Q_{red} = \frac{1}{i} \sum_{k=1}^i \frac{R_{vk}}{(R_{vk} + G_{vk} + B_{vk})} \quad (5)$$

2. 눈 검출 과정

눈 검출 과정은 전처리 과정에서 얻은 얼굴 영역 부분 안에서 관찰자의 눈을 찾기 위하여 우선적으로 눈과 비슷한 크기와 색상 정보를 가지고 있는 부분(유사 눈)을 추출해내는 유사 눈 추출 과정과, 유사 눈 추출 과정을 통하여 찾은 유사 눈들로부터 실제 두 눈의 조건과 일치하는 것을 찾음으로써 두 눈의 위치를 결정하는 과정으로 나누어 수행하였다.

2.1 유사 눈 추출 과정

전처리 과정에서 얻은 얼굴 영역 안의 영상이미지의 각 픽셀의 색상 정보를 모두 비교함으로써 눈과 유사한 부분을 검출해 낼 수 있다.

그림 2에서도 알 수 있듯이 눈동자는 주위의 색과는 구별되게 검은 색의 (R_v, G_v, B_v 값이 작은) 픽셀들로 구성되

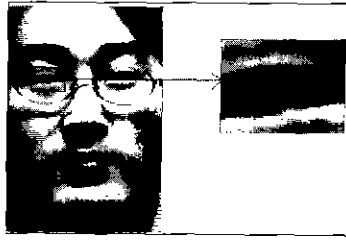


그림 2. 눈동자의 픽셀 구조.
Fig. 2. Structure of pupil.

이 있으므로, 먼저 눈동자의 기준이 될 수 있는 Rv, Gv, Bv 값의 범위를 정하였다. 이 부분에서 조명의 영향을 감소시키기 위해 전처리 과정에서 얻은 Level값에 대한 문턱 값 (Threshold Value, TV)을 결정한다. 그리고나서, 유사 눈의 조건의 범위에 드는 픽셀 들이 가로와 세로에 얼마나 연속적으로 분포되어 있는지를 분석함으로써 눈동자 부분 영역을 파악해 내도록 하는 방법이다. 표 2는 Level값에 대한 문턱 값으로서, 실험을 통해 구하였다.

표 2. Level 값에 대한 문턱 값.

Table 2. Threshold value about the level value.

Level 값	Threshold Value(TV)
90 이하	60
90 - 100	65
100 - 110	70
110 - 120	75
120 - 140	80
140 - 160	85
160 이상	90

이 때 기준이 되는 Rv, Gv, Bv 값의 범위와 눈동자의 가로, 세로 픽셀 수는 카메라와 관찰자의 거리에 따라 변화하게 된다. 본 연구에서는 카메라와 관찰자의 거리를 통상적인 관찰거리인 150-180(cm) 사이로 설정하여 기준 Rv, Gv, Bv 값의 범위와 유사 눈의 가로, 세로의 기준 픽셀 수의 범위를 설정하였다. 유사 눈의 연속된 가로 세로의 픽셀 수를 세는 과정에서, 유사 눈의 범위를 벗어나게 되면 일반적으로 Rv, Gv, Bv 레벨 값이 크게 변화하게 된다.

<유사 눈의 조건>

a) 유사 눈의 기준 픽셀의 Rv, Gv, Bv 값의 범위는 다음과 같이 설정하였다.

$$Rv < TV, Gv < TV, Bv < TV$$

b) 유사 눈의 가로 크기(W)와 세로 크기(L)의 기준 범위는 다음과 같이 설정하였다.

$$5 < W < 20 \quad (\text{픽셀})$$

$$3 < L < 15 \quad (\text{픽셀})$$

c) 유사 눈 내에서는 인접 픽셀 간의 Rv, Gv, Bv 레벨값의 차이가 다음과 같이 작아야 한다.

$$|Rv(x_{k+1}, y_k) - Rv(x_k, y_k)| < 15,$$

$$|Rv(x_k, y_{k+1}) - Rv(x_k, y_k)| < 15,$$

$$|Gv(x_{k+1}, y_k) - Gv(x_k, y_k)| < 20,$$

$$|Gv(x_k, y_{k+1}) - Gv(x_k, y_k)| < 20,$$

$$|Bv(x_{k+1}, y_k) - Bv(x_k, y_k)| < 20,$$

$$|Bv(x_k, y_{k+1}) - Bv(x_k, y_k)| < 20$$

(여기에서 x_k, y_k 는 각각 한 픽셀의 수평, 수직 좌표를 나타냄)

이상 3가지의 유사 눈의 조건을 만족하는 부분을 얼굴영역 안의 이미지를 분석하여 검출해낸다. 검출된 유사 눈들에 대하여 각각의 수평(X축) 및 수직(Y축) 중심좌표와 중심 픽셀 값을 구한다. 유사 눈들은 사각형의 모양으로 검출되므로 중심 좌표는 쉽게 계산되어진다. 그림 3은 입력 영상 전체를 분석하여 유사 눈의 조건에 부합되는 모든 부분을 흰색으로 변화시킨 것이다.

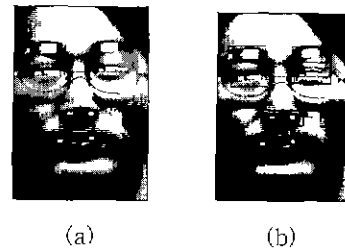


그림 3. 유사 눈 추출 결과 이미지. (a) 결과 이미지, (b) 유사 눈들의 그룹화.

Fig. 3. Extraction result image of the candidate eyes. (a) Result image, (b) Grouping of the candidate eyes.

이렇게 검출된 유사 눈들은 그림 3의 (a)에서 보이는 것과 같이 약 10-30개 정도이고, 그 분포를 보면 눈, 눈썹, 입, 그리고 어두운 부분 등 일정 영역에 집중되어 있음을 알 수 있다. 검출된 모든 유사 눈들에 대하여 눈의 위치를 판단하려 한다면 정확한 눈의 위치를 파악하기 힘들고 처리속도 또한 낮아지게 된다. 따라서 눈의 위치를 판단하기 전에 유사 눈들을 다시 한번 간소화 시켰다. 즉, 유사 눈들이 일정영역에 모여 있음을 감안하여, 두 눈 사이의 거리를 감안해 유사 눈들 간의 거리가 40 픽셀보다 작은 경우의 유사 눈들을 하나의 그룹으로 만들고, 한 그룹 당 하나의 유사 눈만을 추출하게 된다(간소화 단계) 두 눈 사이의 거리는 전처리 과정에서 얻은 얼굴 영역의 수평축(X축)의 크기에 따라서 달라진다.

그림 3의 (b)는 유사 눈들의 그룹을 일정 영역으로 묶은 것이다. 그림 3의 (b)를 보면 한 영역 안에 복수개의 유사 눈이 있음을 알 수 있다. 간소화 단계에서는 이 여러 개의 유사 눈들의 중심 픽셀의 Rv, Gv, Bv 값을 비교하여 그 중에서 가장 낮은 레벨 값(가장 검은 색상)을 갖는 유사 눈만을 남기고 나머지를 제거한다. 따라서 하나의 유사 눈 그룹에서 하나의 유사 눈만이 남게된다. 이 간소화 단계를 거친 후 남게 되는 유사 눈의 수는 10개 보다 작게 된다.

2.2 두 눈의 위치 검출 과정

두 눈의 위치 검출 과정은 간소화 단계를 거쳐 추출된

유사 눈들 중에서 두 눈의 조건과 일치하는 부분을 찾아내는 과정이다. 눈을 하나씩 나누어서 찾은 것보다는 사람의 두 눈의 조건을 이용하여 한꺼번에 찾아내는 것이 용이하므로, 간소화 단계를 거쳐 추출된 유사 눈들 중에서 두 눈의 조건과 일치하는 두개의 유사 눈을 찾아낸다. 두 눈의 조건은 다음과 같이 설정하였다.

<두 눈의 조건>

두 눈으로 판단하기 위한 조건은 다음의 a)-d)와 같다.

a) 두 눈 사이의 픽셀 거리(d)는 카메라와 관찰자와의 거리에 따라 변하게 되는데, 관찰자와 카메라 사이의 거리를 앞서 언급한 바와 같이 150-180(cm) 정도로 가정했을 때의 두 눈 사이의 픽셀 거리의 범위를 기준으로 설정하였다. 즉

$$40 \leq d \leq 80 \quad (6)$$

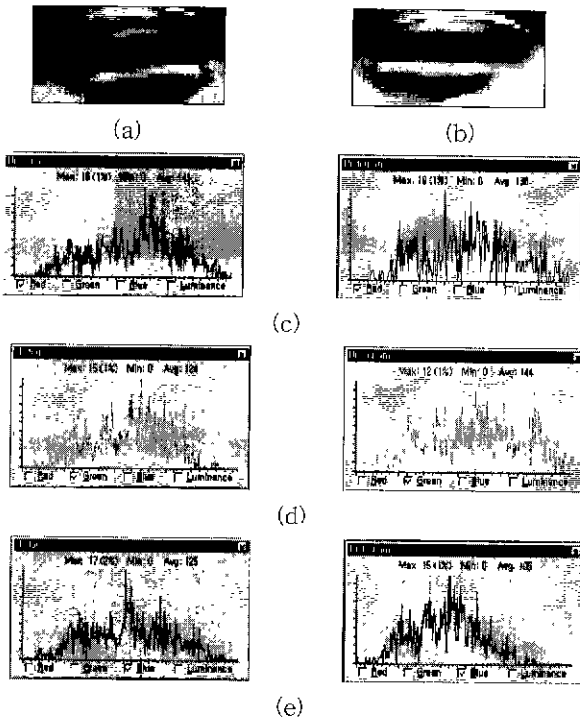


그림 4. 히스토그램 분석. (a) 오른쪽 눈 이미지, (b) 왼쪽 눈 이미지, (c) Red 값 히스토그램, (d) Green 값 히스토그램, (e) Blue 값 히스토그램.
 Fig. 4. Histogram analysis. (a) Image of right eye, (b) Image of left eye, (c) Red value histogram, (d) Green value histogram, (e) Blue value histogram.

이다. d 는 $d = \sqrt{|x_2 - x_1|^2 + |y_2 - y_1|^2}$ 의 식에 의해 계산되며, 여기에서 x_1, y_1 은 첫 번째 유사 눈의 수평(X), 수직(Y) 좌표를 나타내고, x_2, y_2 는 두 번째 유사 눈의 X, Y 좌표를 나타낸다. 이 거리 조건에 만족하는 두 유사 눈의 쌍들을 모두 구한다.

b) 얼굴 안에서 두 눈은 대칭이다. 그러므로 각 유사 눈을 중심으로 일정 영역을 나누어 히스토그램 분석을 통하여 두 유사 눈의 대칭 여부를 판단한다.

그림 4는 전형적인 왼쪽 눈과 오른쪽 눈의 R, G, B 3색

상의 히스토그램 곡선이다. 모두 유사한 곡선을 갖고 있다. a)의 거리 조건을 만족시키는 두 유사 눈의 쌍들 중에서 가장 대칭이라 판단되는 5개의 유사 눈의 쌍들을 선정한다. 여기서 대칭 정도의 판단은 R, G, B 각각의 색상 정보로써 판단한다. 즉 영역 안의 각각의 R, G, B 분포를 만들고 두 유사 눈 쌍의 각각의 R, G, B 분포의 차이 값을 계산하여, 모든 레벨(0-255)에 대하여 전부 합한 값이 가장 작은 것부터 5개의 유사 눈 쌍을 선정한다.

c) 눈동자는 얼굴 내에서 가장 낮은 색상을 갖는다. 즉 b)의 과정을 거쳐 나온 유사 눈의 쌍들로부터, 각 유사 눈의 중심 픽셀의 레벨값을 비교하여 그 값이 가장 작은 쌍부터 순위를 정한다.

d) 위의 a)-c)의 조건에 의해 순위가 결정된 유사 눈의 쌍으로부터, 정확한 두 눈의 쌍인가를 확인하기 위하여 입과의 위치관계를 확인한다. 일반적으로 사람의 얼굴은 그림 5와 같은 구조로 이루어져 있다. 그림 5에서 보이는 것처럼 사람의 두 눈 사이의 거리를 d 라 할 때 일반적으로 입은 두 눈의 중심(m)으로부터 약 $(4/3)d$ 만큼 아래에 위치하게 된다. 위의 a)-c)의 조건에 의해 판단된 두 유사 눈의 중심 좌표들로부터 두 유사 눈 사이의 중심점 m 과 입의 중심점 S 를 간단한 계산에 의하여 찾아 낼 수 있다. 이 S 를 중심으로 입 영역을 설정하여 입 영역 안에서의 적색(R) 값 평균 정규화 색상(Q_{redlip})을 구한다. 입 영역의 크기는 가로가 d 이고 세로가 15 픽셀이 되도록 설정하였다. 앞에서 정한 순위에 따라 Q_{redlip} 을 구하여 전처리 단계에서 구한 Q_{red} 와 비교한다. $Q_{red} < Q_{redlip}$ 의 조건이 성립되면 입이 S 점을 중심으로 위치해 있다고 판단하고, 이때의 유사 눈 쌍을 두 눈이라 결정한다. 그렇지 않은 경우는 다음 순위의 유사 눈 쌍을 대상으로 위의 과정을 계속 수행한다.

이상 위의 4가지 조건들을 이용하여 정확한 두 눈 좌표를 결정해낼 수 있다.

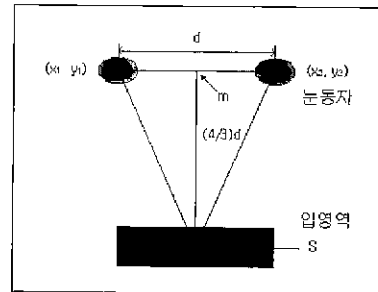


그림 5. 얼굴 구조.
 Fig. 5. Structure of face.

III. 실험 결과

본 실험에서는 다양한 관찰자를 대상으로 관찰자가 움직이는 환경 하에서 수행되었다. 입력되는 영상은 320*240의 순수 컬러이고, 다음의 조건하에서 실험을 수행하였다.

- 1) 관찰자는 카메라를 보고 있다
- 2) 관찰자와 카메라와의 거리 L 은 $1.50(m) < L < 1.85(m)$ 의 범위 내에 있으며, 이 안에서 관찰자가 움직인다.

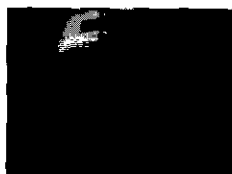
3) 관찰자는 검은 눈동자를 가진 동양인이다.

본 연구에서 사용한 방법은 관찰자의 변화와 거리의 변화에도 상당한 정도의 융통성을 보이며, 적응해 나갈 수 있었다. 모두 8명의 대상으로부터 80장의 입력 이미지를 얻어 실험을 실행한 결과 그중 78장의 이미지에서 두 눈의 위치를 정확하게 찾아 낼 수 있었다. 즉 97.5%까지 정확히 두 눈의 위치를 찾을 수 있었으며, 검출하는데 걸린 수행 시간은 평균적으로 약 0.39초가 소요되었다.

표 3에 기존의 방법과 제안한 방법의 성능을 비교하였으며, 여기에서 알 수 있듯이 제안한 방법이 PC를 사용하면 서도 수행 시간과 검출 성공률 면에서 훨씬 우수한 성능을 보이고 있음을 알 수 있다. 그림 6, 7, 8, 9는 각각 조명이 어두운 경우와 너무 밝은 경우, 얼굴이 많이 기울어진 경우와 배경이 피부색일 경우에 대해 실험한 결과로서, 이들 경우와 같이 매우 심한 환경의 변화에서도 두 눈 위치 검출을 성공적으로 하고 있어 제안한 방법이 환경에 상당히 견실(robust)함을 알 수 있다.

표 3. 기존의 방법과 제안한 방법의 성능 비교
Table 3. The performance comparison between the previous methods and the proposed method

	[12]의 방법	[13]의 방법	[14]의 방법	[15]의 방법	제안한 방법
특징	시스템 구성이 복잡	안정 작동시 30° 이상 얼굴 회전이 검출 곤란	얼굴 내 특징점들의 contour를 구함	퍼지 패턴 매칭법 사용	유사 눈 추출 후 두 눈의 조정을 이용해 검출
센서의 해상도	자기 위치 센서	흑박 카메라 640X480 픽셀	컬러 카메라 640X480 픽셀	컬러 카메라 80X 100~640X480 픽셀	컬러 카메라 320X240 픽셀
프로세서	PC 여러대	SUN Spark IPC Workstation	Indigo Workstation	Indigo ² Workstation	Pentium II PC 330MHz
성능	왼쪽 눈의 위치만 검출	두 눈 위치 검출	눈, 입, 코의 contour 검출	두 눈 위치 검출	두 눈 위치 검출
검출 소요시간			20 초	2 초	0.39 초
검출 성공률		91.7 %		90 %	97.5 %



(a) 어두운 조명

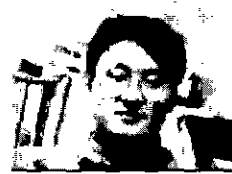


(b) 얼굴영역 검출결과



(c) 유사 눈 추출결과

그림 6. 눈 검출 과정 : 두 눈의 중심 좌표(125, 113)
Fig. 6 Eye detection process : center position of two eyes (125, 113)



(a) 밝은 조명



(b) 얼굴영역 검출결과



(c) 유사 눈 추출결과

그림 7. 눈 검출 과정 : 두 눈의 중심 좌표(172, 111).
Fig. 7. Eye detection process : center position of two eyes (172, 111).



(a) 기울어진 얼굴



(b) 얼굴영역 검출결과



(c) 유사 눈 추출결과

그림 8. 눈 검출 과정 : 두 눈의 중심 좌표 (220, 98).
Fig. 8. Eye detection process : center position of two eyes (220, 98).



(a) 피부색 배경



(b) 얼굴영역 검출결과



(c) 유사 눈 추출결과

그림 9. 눈 검출 과정 : 두 눈의 중심 좌표(210, 94).
Fig. 9 Eye detection process : center position of two eyes (210, 94).

IV. 결론

본 논문에서는 관찰자의 두 눈의 위치를 검출할 수 있도록 하는 알고리즘을 제안하고, 이의 적용 실험 결과를 제시하였다. 본 논문에서 제안한 방법을 통해 0.39초의 짧은 시간내에 약 97.5%까지 정확히 두 눈의 위치

를 찾을 수 있었다. 본 실험에서는 관찰자가 카메라를 보고 있으며, 관찰자와 카메라와의 거리는 통상적인 관찰거리인 1.50m에서 1.85m의 범위 내이고, 관찰자는 검은 눈동자를 가진 동양인이라는 조건하에 수행하였으며, 이러한 조건들이 본 방법의 적용 한계이다. 그리고 급격한 거리의 변화나 매우 심한 조명의 변화 하에서도 정확히 두 눈의 위치를 찾기 위해서는 더욱 많은 보완 연구가 필요하다고 판단된다.

참고문헌

[1] Stan Birchfield, "Elliptical head tracking using intensity gradients and color histograms," *IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, California, June, 1998.

[2] D. Valentin, H. Abdi, A. J. Otooole, and G.W Cottrel, "Connectionist models of face processing - A survey," *Pattern Recognition*, vol 27, no. 9, pp. 1209-1230, 1994.

[3] J. Weng, T. S. Huang, and N. Ahuja, *Motion and Structure from Image Sequences*, Springer- Verlag, Berlin, 1993.

[4] H. A. Rowley, S. Buluja, and T. Kanade, "Rotation invariant neural network-based face detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 38-44, June, 1998.

[5] K.-C. Jin, J.-H. Cho, "Automatic face feature tracking," *Proceedings of the 1999 IEEE Region 10 Conference TENCN'99*, vol. 1, pp. 68-71, Cheju, Korea, September, 1999.

[6]-C. Y. Yoon, J. H. Park, and Mignon Park, "Face recognition using wavelets and fuzzy C-Means clustering," *Proceedings of the 1999 IEEE Region 10 Conference TENCN'99*, vol. 1, pp. 80-83, Cheju, Korea, September, 1999.

[7] S.-J. Lee, S.-B. Jung, J.-W. Kwon, and S.-H. Hong, "Face detection and recognition using PCA," *Proceedings of the 1999 IEEE Region 10 Conference TENCN'99*, vol. 1, pp. 84-87, Cheju, Korea, September, 1999.

[8] Macro La Cascia. John Ishidoro, Stan Sclaroff, "Head tracking via robust registration in texture map images," *IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, California, June, 1998.

[9] Tomohiko Hattori, "Streographic image taking and output devices compatible with conventional infrastructure," *Proceedings of International Workshop on 3D Imaging Media Technology*, pp. 143-150, March 1996.

[10] C. Fermuller and Y. Aloimonos, "The hole of fixations in visual motion analysis," *Int'l J. Computer Vision*, 11:2, pp. 165-186, 1993.

[11] M. J. Swain and M. A. Strcker, "Promising directions in active vision," *Int'l J. Computer Vision*, 11:2, pp. 109-126, 1993.

[12] Hiroshi Imai, Masao Imai, Yukio Ogura, "Eye-position tracking stereoscopic projector using image shifting optics," *Proceedings of the Third International Display Workshops (IDW)*, pp. 449-452, Kobe, Japan, Nov., 1996.

[13] H.-S. Yoon, J. S., M. Wang, and B.-W. Min, "Human eye detection in a complex background," *The Fourth International Conference on Control, Automation, Robotics and Vision (ICARCV'96)*, vol. 1, pp. 380-384, Singapore, Dec., 1996.

[14] Taro Yokoyama, Haiyuan Wu and Masahiko Yachida, "Automatic detection of facial feature points and contours" *Proceedings of 5th IEEE International Workshop on Robot and Human Communication*, pp. 335-340, Tsukuba, Nov., 1996.

[15] Haiyuan Wu, Takeshi Fukumoto, Qian Chen, and Mashhiko Yachida, "Face detection and rotations estimation using color information." *Proc. of 5th IEEE International Workshop on Robot and Human Communication*, pp. 341-346, Tsukuba, Nov., 1996.



허경무

1979년 서울대학교 전자공학과 졸업
한국과학기술원 전기 및 전자공학과 석사(1981), 동대학 박사(1989). 현재 단국대학교 공학부 전자공학전공 부 교수. 주요 연구분야는 로봇트 제어, 3차원 디스플레이, 시스템 제어, 학습 제어, 디지털 제어기 설계 및 Simulation, Servo Motor 제어.



박영빈

1999년 단국대학교 전자공학과 졸업.
현재 동대학원 석사과정. 관심분야는 로봇 제어.