

실시간 능동 비전 카메라를 이용한 시선 위치 추적 시스템

(Gaze Detection System using Real-time Active Vision Camera)

박 강 령^{*}
(Kang Ryoung Park)

요 약 이 논문에서는 컴퓨터 시각 인식 방법에 의해 모니터 상에 사용자가 쳐다보고 있는 시선 위치를 파악하기 위한 새롭고 실용적인 방법을 제안한다. 일반적으로 사용자가 모니터 상의 한 위치를 쳐다보기 위해서는 얼굴 및 눈동자를 동시에 움직이는 경향이 있다. 기존의 시선 위치 추적 시스템은 사용자의 얼굴 전체를 취득할 수 있는 단 하나의 광각 카메라 시스템을 주로 많이 이용하였다. 그러나 이러한 경우 영상의 해상도가 많이 떨어져서 사용자의 눈동자 움직임을 정확하게 추적하기 어려운 문제점이 있다. 그러므로 이 논문에서는 광각 카메라(얼굴의 움직임에 의한 시선 위치 추적용) 및 눈 영역을 확대하여 취득하는 협각 카메라(눈동자 움직임에 의한 시선 위치 추적용), 즉 이중 카메라를 이용하여 시선 위치 추적 시스템을 구현하였다. 얼굴의 움직임 시 전체적인 위치가 변화될 눈동자의 움직임을 정확히 추적하기 위해, 협각 카메라에는 광각 카메라로부터 추출된 눈 특징점의 위치를 기반으로 한 자동 초점 및 자동 상하/좌우 회전 기능이 포함되어 있으며, 눈 특징점을 보다 빠르고 정확하게 추출하기 위해 이중 적외선 조명을 사용하였다. 실험 결과, 본 논문에서는 실시간으로 동작하는 시선 위치 추적 시스템을 구현할 수 있었으며, 이때 얼굴 및 눈동자 움직임을 모두 고려하여 계산한 모니터상의 시선 위치 정확도는 약 3.44cm의 최소 자승 에러성능을 나타냈다.

키워드 : 시선 위치 추적, 이중 카메라, 자동 초점, 자동 상하/좌우 회전, 이중 적외선 조명

Abstract This paper presents a new and practical method based on computer vision for detecting the monitor position where the user is looking. In general, the user tends to move both his face and eyes in order to gaze at certain monitor position. Previous researches use only one wide view camera, which can capture a whole user's face. In such a case, the image resolution is too low and the fine movements of user's eye cannot be exactly detected. So, we implement the gaze detection system with dual camera systems(a wide and a narrow view camera). In order to locate the user's eye position accurately, the narrow view camera has the functionalities of auto focusing and auto panning/tilting based on the detected 3D facial feature positions from the wide view camera. In addition, we use dual IR-LED illuminators in order to detect facial features and especially eye features. As experimental results, we can implement the real-time gaze detection system and the gaze position accuracy between the computed positions and the real ones is about 3.44 cm of RMS error.

Key words : Gaze Detection, Dual Camera System, Auto Focusing, Auto Panning/Tilting, Dual IR-LED illuminators

1. 서 론

시선 위치 추적이란 현재 사용자가 응시하고 있는 위치를 파악하는 연구이다. 이 논문에서는 모니터 상에 사

용자가 쳐다보고 있는 위치를 파악하기 위한 새롭고 실용적인 방법을 제안한다. HCI(Human Computer Interaction)분야에서 사용자가 모니터 상에 쳐다보고 있는 위치는 여러 가지 의미에서 많은 중요한 정보를 갖는다. 또한 이러한 시선 위치 추적은 많은 응용 분야를 갖고 있는데 대표적인 예가 일반적인 컴퓨터 인터페이스 환경, 3차원 가상현실 시뮬레이터에서 사용자의 시선 위치에 따른 화면 조정, 양 손을 사용하지 못하는 장애인들을 위한 인터페이스 수단 및 양 손으로 화면상의

· 본 연구는 한국과학재단 지정 생체인식연구센터의 지원을 받아 이루어졌습니다.

^{*} 정 회 원 : 상명대 미디어학부 교수
parkgr@smu.ac.kr

논문접수 : 2003년 5월 7일
심사완료 : 2003년 8월 13일

버튼을 제어하고 있는 상황에서 제 3의 제어를 해야 하는 경우 등등이 있다.

기존의 연구들은 주로 얼굴의 2차원 및 3차원 움직임 량을 추정하거나[1-4], 얼굴의 움직임 만에 의한 시선 위치를 파악하거나[5-14], 눈동자만의 움직임에 의한 사용자의 시선 위치를 파악하는 연구가 주종을 이루었다[15-21]. 그러나 얼굴 및 눈동자 움직임을 함께 고려하여 시선 위치를 파악하는 연구는 거의 이루어지지 않았다.

Ohmura와 Ballard[7,8] 등의 연구에서는 초기에 얼굴 특징점의 3차원 거리(depth) 정보를 알고 있어야 하며, 시선 위치를 파악하기 위해서는 많은 처리 시간(약 1분 이상)이 소요되는 단점이 있다. Gee[9]와 Heinzmann[10] 등의 연구에서는 얼굴 좌표계에서의 시선 벡터의 방향을 계산하였을 뿐, 이로부터 모니터 상에 사용자 시선 위치 등을 구하지 않았다. 또한, 이들의 연구에서는 얼굴의 3차원 회전 및 이동이 동시에 발생하지 않는다고 가정했다. 이는 얼굴의 회전과 이동이 동시에 발생했을 때, 그들의 논문에서 사용하는 최소 자승 정합 알고리즘(least-square fitting algorithm)에서의 연산 복잡성과 처리 시간의 상승 등으로 3차원 움직임량을 정확하게 추정하기 어려웠기 때문이다. Rikert[11]의 연구에서는 학습 및 테스트 환경에서 얼굴 및 모니터 스크린사이의 거리가 변하지 않아야 한다는 가정이 있으며, 이러한 것은 실제 사용에 있어서 많은 불편함을 제공하게 된다. 기타 다른 연구들[12,13,16,19,20]에서는 사용자로 하여금 구분점이 부착된 안경을 착용하게 함으로써 얼굴 특징점을 추적하는 연구를 수행하였는데, 이처럼 별도의 안경을 착용해야 하므로 사용자에게 불편함을 제공하는 결과를 낳게 되었다. [5,6,14]의 연구에서는 눈동자의 움직임은 전혀 고려치 않고 단지 얼굴의 움직임 만에 의한 시선 위치 파악 연구를 수행하였다. 얼굴 및 눈동자의 움직임을 같이 고려하여 시선 위치를 파악한 연구[22]에서는 한 개의 광각 카메라를 이용하여 전체 얼굴을 촬영하는 방식을 취하였다. 그러나 이러한 경우, 눈동자 영상의 해상도(image resolution)가 많이 떨어지며 그 결과 사용자의 정교한 눈동자 움직임을 추적하기 어려운 문제점이 있다. 그러므로 이 논문에서는 광각 카메라(얼굴의 움직임에 의한 시선 위치 추적용) 및 눈 영역을 확대하여 취득하는 협각 카메라(눈동자 움직임에 의한 시선 위치 추적용), 즉 2개의 카메라를 이용하여 시선 위치 추적 시스템을 구현하였다.

얼굴의 움직임 시 전체적인 위치가 변화될 눈동자의 움직임을 정확히 추적하기 위해, 협각 카메라에는 광각 카메라 영상에서 추출된 얼굴 특징점의 위치를 기반으로 한 자동 초점(auto focusing) 및 자동 상하/좌우 회

전(auto panning/tilting) 기능이 포함되어 있다. 추가로, 본 논문에서는 얼굴내 눈 특징점을 보다 빠르고 정확하게 추출하기 위해 적외선 조명을 사용하였다. 이때 안경 사용자의 경우 안경 면에서 생기는 반사위치를 피하여 눈 영역을 정확하게 추출하기 위하여, 2개의 적외선 조명을 번갈아 키는 방법을 사용하였다. 실험 결과, 약 3.44cm의 시선 위치 추적 에러(최소 자승 에러) 성능을 얻을 수 있었다.

2. 얼굴 특징점 추적

2.1 시선 위치 추적용 이중(dual) 카메라 시스템

모니터 상에 사용자가 응시하고 있는 위치를 파악하기 위해서는 입력 영상에서 얼굴 특징점의 위치를 정확하게 추출할 수 있는 기술이 요구된다. 본 논문에서는 얼굴 특징점으로 양 눈동자의 중심 및 눈의 구석, 콧구멍 및 입의 양 끝점을 사용하였다. 얼굴 특징점의 위치를 찾기 위해 이 논문에서는 입력 영상으로부터 눈 영역을 먼저 추출한 후, 이를 기준으로 기타 다른 얼굴 특징점들의 위치를 실시간으로 추적하는 방법을 사용하였다[22].

일반적인 사무실 환경에서 사용자의 얼굴 뒤에는 복잡한 배경이 존재하며, 또한 외부 광(실내등 혹은 창문을 통해 들어오는 태양광) 등의 영향으로 얼굴 및 얼굴 특징점의 위치를 영상 처리 방법만으로는 실시간으로 추출하기 어렵다는 것이 일반적으로 알려져 있다. 그러므로 이 논문에서는 그림 1과 같이 적외선 조명 기반 카메라 시스템을 개발하고, 이를 이용하여 입력 영상으로부터 얼굴 특징점을 실시간으로 추출하였다. 전술한 바와 같이 입력 영상에서 눈 위치를 먼저 찾은 후 이를 기준으로 다른 얼굴 특징점의 위치를 찾는 방식을 사용했으며, 눈 위치를 추출하기 위하여 본 논문에서는 적외선 조명을 켜줄 때 눈의 각막(Corneal)에서 발생하는 반사위치(Specular Reflection)를 찾는 방법을 사용하였다[22].

그림 1을 보다 자세히 설명하면 다음과 같다. 적외선 조명(IR-LED(InfraRed-Light Emitting Diode))(1)은 눈에서의 반사 위치(Specular Reflection)를 생성하는데 사용된다. 이때 적외선조명으로는 사람 눈으로 감지하기 어려운 880nm의 조명을 사용함으로써, 동작 중 사용자의 눈부심 현상이 없도록 하였다. 카메라 렌즈 앞에 부착된 HPF(High Pass Filter)(2)는 근 적외선 영역(700nm)이상의 적외선 조명만을 통과시키기 때문에, 카메라 센서를 통한 영상 취득 시 700nm이하의 외부광(가시광선 대)이 영향을 주지 못하게 된다. 즉, 입력된 영상 내에는 카메라에 설치되어 있는 적외선 조명(1)만이 주로 영향을 주기 때문에, 외부광에 의한 발생하는 영상내의 그림자 등의 영향은 적어지게 된다. 그런데, 일반적인 카메라 CCD(Charge Coupled Device) 센서

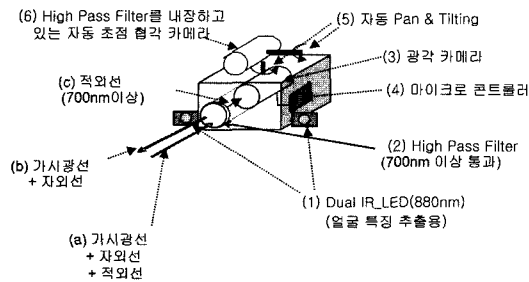


그림 1 시선 위치 추적용 이중 카메라 시스템

의 경우, 가시광선보다는 적외선 영역으로 갈수록 영상의 감도가 낮아지는 특성을 나타낸다. 본 논문에서는 사용자의 눈에 눈부심을 주지 않기 위하여 880nm의 적외선을 사용했으며, 이러한 경우 영상에서의 감도가 많이 낮아 얼굴 특징점을 추출하기 위해 충분한 밝기를 내기 위해서는 5 개의 적외선조명의 광을 모아 사용해야만 했다.

또한, 본 연구에서는 일반적인 격행 주사(Interlaced Scan) CCD 카메라를 사용하였으며 이와 같은 카메라 및 조명의 동작을 제어하기 위한 마이크로 컨트롤러(4)를 사용했다. 이를 이용하여 그림 2에서 나타나 있듯이 CCD의 출력 신호 중 Even 과 Odd 필드의 VD(Vertical Drive)신호에 맞추어 조명의 On/Off를 조정하게 하였다. 그림 2를 보다 자세히 설명하면 다음과 같다. 초기에 사용자가 시선 위치 추적 시스템을 시작하려는 순간, PC쪽에 구현되어 있는 프로그램으로부터 카메라 마이크모 RS-232C 인터페이스를 거쳐 시작 신호가 전달된다(①). 이 다음 frame부터, 카메라에 있는 마이크모는 그림 2에 나타나 있듯이 CCD의 매 Even 및 Odd Field의 시작 위치 VD신호에 맞추어 적외선 조명을 계속 On/Off 시키게 된다. 이처럼 적외선 조명을

On/Off시킨 영상이 입력되면, 이로부터 차영상을 구하여 눈에 발생하는 적외선 반사 위치를 구하게 된다(②)[22]. 이때 눈에 발생하는 반사위치(Specular Reflection)의 영상 그레이 값(Grey Value)은 다른 반사위치(뺨, 혹은 이마 등에 생기는 조명 반사)보다 높기 때문에, 차영상을 구하게 되면 눈에 발생하는 반사 위치를 쉽게 찾을 수 있게 된다[22].

그러나 때에 따라 화장을 진하게 한 여성의 경우, 피부에서의 조명 반사가 심해서 입력 영상 자체에서 눈에 의한 반사 위치와 다른 피부 등에서의 반사 위치가 거의 비슷한 그레이 레벨(255)을 나타내는 경우가 있다. 이러한 문제를 해결하기 위하여 본 논문에서는 추가로 프레임 그래버(Frame Grabber)의 디코더 설정값을 변경하는 방법을 이용하여 보다 정확한 눈의 반사위치를 찾는 방법을 취했다.

프레임 그래버는 카메라에서 출력되는 아날로그 신호(NTSC 혹은 PAL)를 입력받아 이를 디지털로 변환(Analog to Digital Conversion)함으로써 최종적으로 컴퓨터에서 처리할 수 있는 디지털 영상을 취득하는 기능을 수행하는 장비이다. 이러한 프레임 그래버 보드는 크게 2가지 기능을 수행하는 칩들로 구성되어 있는데 그 첫 번째 것이 아날로그 신호를 디지털로 변환하는 디코더(Decoder) 칩과 디코더로부터 나오는 디지털 신호를 컴퓨터 PCI로 전송하는 멀티미디어 브리지(Multimedia Bridge)가 있다. 그림 1에서와 같이 시선 위치 추적 시스템에는 협각 및 광각 두개의 카메라가 있으며, 각 카메라로부터 나오는 NTSC 신호를 입력받기 위해 본 연구에서는 2개의 디코더와 1개의 멀티미디어 브리지를 갖는 2채널 프레임 그래버(PCI인터페이스)를 구현하였다. 그런데 일반적으로 NTSC 신호는 넓은 신호 범위(0~2¹⁰-1)를 갖고 있고, 일반적인 디코더는 이

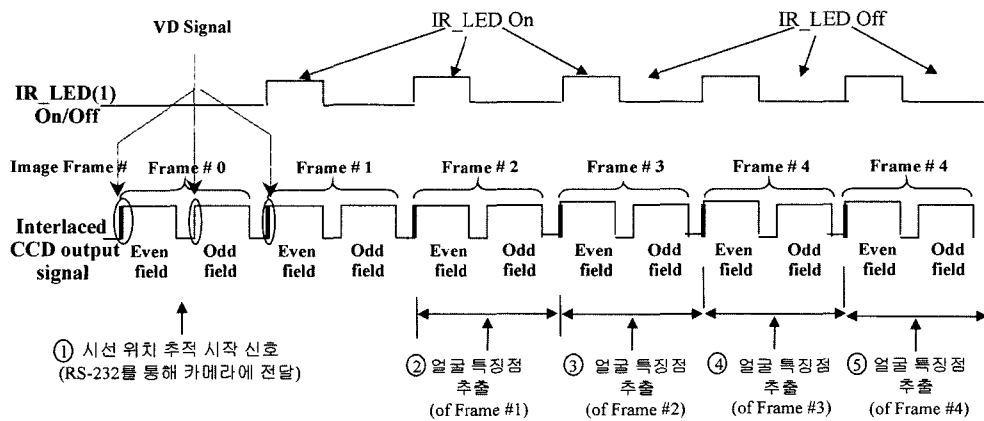


그림 2 얼굴 특징점 추출을 위한 적외선 조명 조정

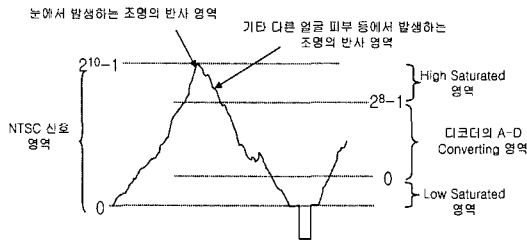


그림 3 NTSC 신호 영역 vs. 디코더의 A-D 변환 영역

에 비해 적은 범위의 아날로그/디지털 변환 능력(0 ~ 2^8-1)을 갖고 있기 때문에, 입력된 NTSC 신호는 이러한 디코더를 통해 다 표현되지 못하고 그림 3과 같이 잘려지는 부분이 발생하게 된다. 그림 3에서 high saturation된 NTSC 신호는 결과적으로 영상에서 $2^8-1(255)$ 의 그레이 레벨을 갖게 된다.

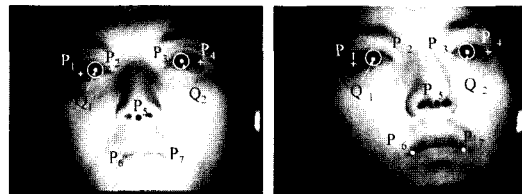
그러므로 피부 반사가 심한 사람의 경우, 그림 3에서 보는 바와 같이 눈에서 발생하는 조명의 반사 영역(Specular Reflection)과 얼굴의 피부 등에서 발생하는 다른 반사 영역들은 디지털 영상 상에서는 같은 $2^8-1(255)$ 의 레벨을 갖게 된다. 이러한 경우, 영상 처리 방법 만에 의해서는 눈에서 발생하는 반사 영역과 기타 다른 피부 등에서 발생하는 영역을 분리하기가 쉽지 않다. 그러나 그림 3과 같이 NTSC 아날로그 신호를 보았을 때는 눈에서 발생하는 반사 영역의 레벨이 다른 피부 등에서 발생하는 레벨보다 높기 때문에(본 연구에서는 NTSC 신호를 직접 오실로스코프(Oscilloscope) 장비로 입력받아 이러한 현상을 확인하였다), 본 논문에서는 디코더에서의 아날로그/디지털 변환 시 사용하는 설정값(Brightness 값)을 조정함으로써 눈에서 발생하는 반사 영역과 기타 다른 피부 등에서 발생하는 영역을 분리하였다. 즉, 디코더의 Brightness 값을 전체적으로 낮추면, 그림 3에서 디코더의 아날로그/디지털 변환 영역이 위로 이동(Shift)하게 되어 결과적으로 High Saturated 되는 영역은 없어지게 된다. 즉, 이전에는 둘 다 255로 표시되던 눈에서 발생하는 반사 영역과 다른 피부 등에서 발생하는 영역이 영상 자체에서 다른 값으로 표현됨으로써 궁극적으로는 두 영역을 쉽게 분리해 낼 수 있게 되는 것이다.

2.2 눈 및 기타 얼굴 특징점 추출

입력 영상으로부터 눈의 반사위치를 추출하게 되면, 이로부터 제한된 영역에서 정확한 눈 위치(눈동자의 중심 및 눈의 구석)를 추출하게 된다. 눈동자의 중심 위치를 추출하기 위해 이 논문에서는 원형 경계 검출 알고리즘(Circular Edge Detection)을 사용하였다. 눈동자의 중심 위치를 추출한 후, 본 논문에서는 눈구석 모양 템플릿(Eye Corner Shape Template)을 이용하여 눈의

양 구석 위치를 추출하였다[22]. 본 논문에서는 그림 1에서와 같이 HPF(2)를 사용하므로 입력 영상 내에 외부광에 의한 영향을 최소화 할 수 있었으며, 그 결과 입력 영상이나 템플릿 등을 조명에 대해 별도로 정규화할 필요가 없었다. 이와 같은 눈구석 모양 템플릿과 함께 본 논문에서는 SVM(Support Vector Machine)을 이용하여 정확한 눈의 구석 위치를 추출할 수 있었다[22].

이처럼 입력 영상에서 사용자의 눈 위치를 찾게 되면, 찾은 눈 위치에 대해 코 및 입의 상대적인 존재 가능 영역을 설정할 수 있게 된다. 이 논문에서는 설정된 존재 가능 영역 내에서, SVM을 이용하여 눈의 구석 위치를 찾았던 것과 유사한 방법으로 양 콧구멍 및 입의 끝점 위치를 찾았다. 입력 영상에서 얼굴 특징점의 위치가 추출되면, 그림 4에 나타난 것과 같이 총 9개의 특징점($P_1, P_2, P_3, P_4, P_5, P_6, P_7, Q_1, Q_2$)을 선정하게 되는데 이러한 특징점들은 얼굴의 3차원 움직임 량(3차원 회전 량 및 이동량) 및 눈동자 움직임량을 추정하기 위해 사용된다. 사용자가 모니터상의 한 지점을 응시하게 되면 그림 4의 (b)에서 나타나 있는 것처럼 9개의 특징점 위치가 ($P'_1, P'_2 \sim P'_7, Q'_1, Q'_2$)로써 변하게 된다. 여기서 ($Q_1, Q_2, P_1, P_2, P_3, P_4$)과 ($Q'_1, Q'_2, P'_1, P'_2, P'_3, P'_4$)는 5장에서 언급할 눈동자 움직임에 의한 시선 위치 추출을 위하여서도 추가로 사용된다.



(a) 사용자가 모니터 증상을 응시할 때 (b) 사용자가 모니터 상의 좌측상단지점을 응시할 때

그림 4 얼굴 및 눈동자 움직임에 의한 시선 위치 추출을 위해 사용되는 얼굴 특징점

3. 얼굴 특징점의 초기 3차원 위치 추정

3.1 모니터상의 시선 위치를 파악하기 위한 4단계

그림 5에 나타나 있는 것처럼, 얼굴 특징점을 추출한 후 모니터상의 시선 위치를 파악하기 위해서는 본 논문에서는 4 단계의 과정을 거치게 된다[5,6,22]. 첫 번째 단계에서는 사용자가 모니터 상에 미리 지정된 5 위치(모니터 좌측 상단, 우측 하단, 중앙, 중앙과 좌측 상단 사이의 지점, 중앙과 우측 하단 사이의 지점)를 쳐다봄으로써 얻어진 정보들을 바탕으로 모니터 좌표계 및 얼굴 좌표계에서의 초기 얼굴 특징점의 3차원 위치를 자

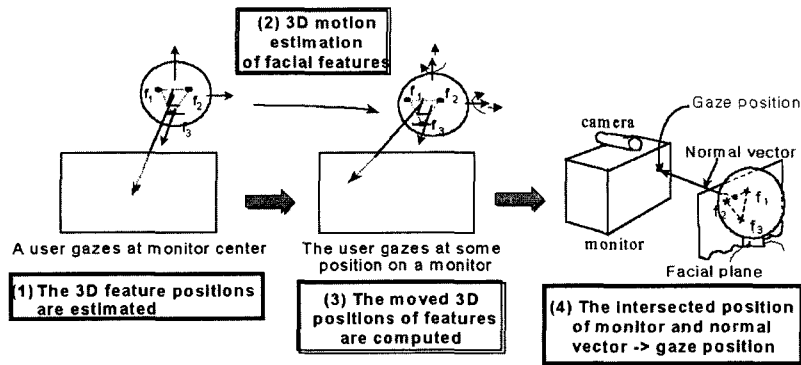


그림 5 모니터 상의 시선 위치를 파악하기 위한 4단계

동으로 구하게 된다. 본 논문의 실험 환경에서 사용자는 모니터상의 한 지점을 쳐다보고 이때 사용자의 얼굴 특징점은 모니터 위에 놓여져 있는 카메라로부터 관측되므로, 얼굴 좌표계, 모니터 좌표계, 카메라 좌표계의 3가지 좌표계를 동시에 고려해야 한다. 즉, 이러한 3가지 좌표계사이의 변환 관계를 정확히 알고 있어야만 얼굴 특징점의 3차원 위치를 구할 수 있게 된다.

먼저 모니터 좌표계에 대한 카메라 좌표계의 변환 관계를 파악하기 위해서는 모니터 중심 위치에 대한 카메라 광축 중심의 위치 변위, 모니터 중심에서의 법선 방향에 대한 카메라 광축의 기울어진 각도, 그리고 카메라 자체의 초점거리등을 알고 있어야한다. 또한 모니터 좌표계에 대한 얼굴 좌표계의 변환 관계를 알기 위해서는 모니터 중심 위치로부터 얼굴 좌표계 중심 위치까지의 Z거리를 알아야 한다. 이와 같이 모르는 미지의 변수들을 자동으로 구하기 위해 본 논문에서는 위의 3가지 좌표축을 결합하여 카메라 좌표계에서 관측되는 얼굴 특징점의 위치 변화로부터 위의 미지 변수들을 구할 수 있는 식을 유도하였다[5,6,22]. 그러나 이 식은 그 해를 산술적으로 구할 수 없는 비선형 방정식이므로, 본 논문에서는 초기에 사용자가 모니터상의 미리 알고 있는 5지점을 쳐다볼 때 카메라 좌표계에서 얻어진 얼굴 특징점의 위치 변화 값들과 이때 사용자가 쳐다보는 모니터상의 실제 위치 정보들을 대입한 반복적 변수 추정 방법(iterative parameter estimation)을 이용하여 앞에서 언급한 미지의 변수를 구하게 된다. 이러한 변수들로부터 모니터 좌표계 및 얼굴 좌표계에서의 초기 얼굴 특징점의 3차원 위치를 구할 수 있게 되었다[5,6,22].

이후 그림 5의 두 번째 및 세 번째 단계에서 사용자가 모니터상의 다른 한 지점을 쳐다볼 때 변화된 얼굴 특징점의 3차원 위치를 구하기 위하여, 본 논문에서는 얼굴 좌표계에서의 얼굴 특징점의 회전과 이동량을 3차원 움직임 추정 방법 중의 하나인 확장 칼만 필터

(Extended Kalman Filter)를 이용하여 구하고, 이로부터 아핀 변환(Affine Transformation)을 적용하여 변화된 얼굴 특징점의 3차원 위치를 산술적으로 구할 수 있게 된다. 그림 5의 4번째 단계에서는 변화된 얼굴 특징점의 위치를 다시 모니터 좌표계를 기준으로 계산하고, 이 얼굴 특징점들의 위치로부터 하나의 얼굴 평면과 법선을 구할 수 있으며, 이 법선과 모니터 평면이 만나는 위치가 얼굴의 움직임에 의한 시선 위치로 정하게 된다.

3.2 자동 초점/상하/좌우 회전되는 협각 카메라에 의한 눈동자 위치 추적

그림 5의 세 번째 단계에서 얼굴 특징점의 변화된 3차원 위치가 계산되면, 이들은 모니터 좌표계 및 첫 번째 단계에서 구한 카메라 매개변수 정보 등을 이용하여 카메라 좌표계를 기준으로 계산되어 질 수 있다. 이와 같은 정보를 바탕으로, 본 논문에서는 정확한 눈동자의 위치를 취득하기 위해 협각 카메라(Narrow View Camera)를 자동 상하/좌우 회전(auto panning/tilting)시킬 수 있게 된다(그림 1의 (5) 참조).

본 논문에서 사용하는 협각 카메라는 작은 화각(약 30~45 mm의 큰 초점거리)을 갖고 있으며, 이러한 특성 때문에 눈 영역을 확대하여 촬영할 수 있다. 또한 사용자가 모니터의 한 지점을 쳐다보기 위하여 얼굴을 심하게 회전하는 경우, 얼굴 영상을 취득하기 위한 광각 카메라(Wide View Camera)에서 한쪽 눈 영역이 아예 사라져 버리는 경우가 발생한다. 이러한 경우, 협각 카메라를 아무리 상하/좌우 회전(auto panning/tilting)한다고 하더라도, 광각 영상에서 사라진 한 쪽 눈 영역을 찾을 수 없게 된다. 그러므로 이러한 경우에 대비하기 위하여 본 논문에서는 사용자가 모니터 중앙을 쳐다보는 경우 양쪽 눈을 모두 추적하고, 이후 사용자가 얼굴을 심하게 회전하여 한쪽 눈을 추적할 수 없는 경우, 나머지 보이는 다른 쪽 눈만을 추적함으로써 눈동자 움직임에 의한 정확한 시선 위치를 파악하였다. 협각 카메라

의 상하 및 좌우 회전을 위해 420 pps(pulse per second)를 갖는 2개의 스텝핑 모터(steping motor)를 사용하였다.

또한, 일반적인 협각 카메라의 경우 작은 범위의 DOF(Depth of Field)를 갖는 특성이 있으므로, 입력 영상은 사용자의 Z 방향 움직임에 따라 쉽게 초점이 흐려지는 특성을 나타낸다. 일반적으로 DOF는 카메라의 아이리스(Iris) 크기가 작을수록, 피사체의 거리가 멀수록 커지는 특성을 나타낸다. 그러나 이 논문에서와 같이 사용자의 시선 위치를 파악하는 응용 분야에서는 모니터 앞에 약 50~70 cm의 다양한 Z거리 범위에서 사용자가 앉을 수 있기 때문에 임의로 피사체의 위치를 멀게 만들 수 없으며, 카메라의 아이리스 크기를 작게 만드는 것은 궁극적으로 카메라 CCD 센서로 들어오는 빛의 양을 줄이게 되어 입력 영상이 어두워지는 결과를 낳게 된다. 그러므로 본 논문에서는 단초점 렌즈를 이용해서는 다양한 Z 거리를 갖는 사용자의 눈에 대한 초점을 맞추기 어렵기 때문에, 직경 10mm의 아이리스 크기를 갖는 가변 초점(auto focusing) 렌즈 및 자동 초점용 모터(auto focusing motor, 420 pps)를 이용하여 보다 선명한 눈 영상을 취득하였다(그림 1의 (6)참조). 이와 같은 자동 상하/좌우 회전 및 초점 기능은 그림 1에서 나타나 있는 카메라내의 마이크로 콘트롤러(그림 1의 (4)참조)에 의해 제어된다.

협각 눈 영상의 초점을 맞추기 해서는 기본적으로 카메라를 기준으로 사용자 눈까지의 Z 거리를 알아야 한다. 본 논문에서는 그림 5의 단계 3에서 변화된 얼굴 특징점(특히 눈)의 3차원 위치 및 Z 거리를 구할 수 있으므로 이를 바탕으로 눈 영상의 자동 초점 작용을 하게 된다. 그러나 협각 카메라는 DOF(Depth of Field)가 평장히 작기 때문에 이와 같은 Z 거리만 가지고는 정확한 자동 초점이 이루어지기 어렵다. 그러므로 본 논문에서는 입력 영상내의 픽셀들 간의 변화 차이를 측정하는 단순한 초점 측정 알고리즘을 병행하여 사용하였다. 즉, 그림 5의 단계 3에서 계산된 얼굴 특징점(눈)의 3차원 위치 정보를 바탕으로 협각 카메라를 상하 및 좌우 회전을 하고 자동 초점을 수행함으로써 눈 영상을 취득한다. 이후, 취득된 영상은 PC로 전송된 후, 영상내의 픽셀들 간의 차이를 측정하는 초점 측정 알고리즘에 의해 영상의 초점 정도를 평가한다. 만일 영상의 초점 정도가 미리 정해놓은 임계치를 넘지 못하는 경우, 카메라 마이크로 콘트롤러를 통해 자동 초점 렌즈에 움직임 명령을 전달하게 된다. 여기서, 초점이 맞지 않는 눈 영상이 취득된 경우, 초점 렌즈를 어느 방향으로 움직여야 하는 지(앞 혹은 뒤) 결정하기 어려운 문제가 있다. 이를 결정하기 위하여 본 논문에서는 기타 추가적인 정보를 바탕

으로 한 여러 가지 휴리스틱한 방법(눈 영상의 밝기 정보 혹은 피라미드 기반 전역 렌즈 탐색)등을 이용하여 초점 렌즈의 움직임 방향을 결정한다.

이 단계에서 우리는 안경에 의해 발생하는 반사 영역(Specular Reflection)을 고려해야 한다. 그림 6과 같이, 일반적인 안경 면은 조명에 의한 여러 종류의 반사를 생성하여 이를 협각 카메라에 들어가게 하며, 이러한 경우 협각 카메라로 취득한 눈 영상은 그림 7의 (b)와 같이 큰 반사 영역으로 덮여서 눈 영역을 정확하게 검출하기 어렵게 된다. 일반적으로 안경 면은 사용자의 시력에 따라 다양한 형태를 나타낸다. 즉, 사용자의 시력이 약할수록 안경 면은 평평해지는 경향을 나타내며, 이러한 경우 안경에 의한 반사 영역의 크기가 증가되어 전체 눈 영상을 모두 덮는 경우가 종종 발생한다. 이러한 문제점을 해결하기 위하여 본 논문에서는 그림 1에서 나타낸 것처럼, 두개의 적외선 조명을 사용하여, 각 조명을 번갈아 켜므로써 이와 같은 안경에 의해 발생하는 큰 반사위치를 피하도록 하였다. 즉, 협각 카메라를 통해 눈 영상이 입력되면 본 논문에서는 2장에서 언급한 방법인 프레임 그래버 디코더 값을 낮춤으로써 눈 혹은 안경에서 반사되는 영역과 기타 얼굴 피부에서 반사되는 영역을 쉽게 분리할 수 있다. 이때 디코더 값을 낮춘 입력 영상 내에서 그레이 값이 일정 임계치 이상 되는 영역(눈 혹은 안경에 의한 반사위치)이 영상의 중심을 기준으로 일정 비율이상 존재하면, 안경에 의한 큰 반사 위치가 생긴 것으로 판단하여 반대쪽 조명을 키게 된다. 이러한 경우 그림 7의 (a)와 같이, 안경에 의한 반사 위치는 없어지게 된다.

그러나 때에 따라서는 양쪽 조명을 번갈아 켜도 양쪽 조명에 의한 안경의 반사 위치가 모두 크게 발생하여 눈 검출이 어려운 경우가 발생한다. 이러한 경우는 전술한 바와 같이 주로 평평한 안경을 착용하는 사람들일수록 자주 발생한다. 본 연구에서는 양쪽 조명에 의해 생기는 반사 영역의 크기 등을 파악하여 이러한 경우가

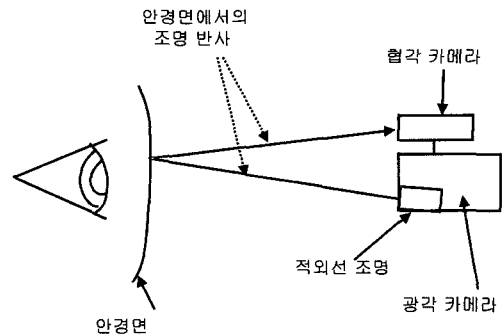
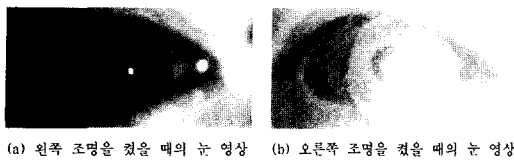


그림 6 안경 면에서의 조명 반사



(a) 왼쪽 조명을 켜었을 때의 눈 영상 (b) 오른쪽 조명을 켜었을 때의 눈 영상

그림 7 협각 카메라로부터 입력된 눈 영상에서의 조명 반사 영역

발생하면 사용자에게 음성 메시지(“고개를 조금 숙이세요” 혹은 “고개를 조금 드세요”)를 내보내어 사용자의 얼굴을 조금 숙이거나 들도록 유도하는 방법을 사용하였다. 이럼으로써 조명대비 사용자 안경면의 각도를 조정하게 함으로써 안경에 의한 반사는 피할 수 있게 된다.

초점이 맞은 눈 영상이 취득되면, 본 논문에서는 학습된 신경망(다층 퍼셉트론)을 이용하여 눈동자 움직임에 의한 모니터 상의 시선 위치를 파악하였다. 이후, 그림 5의 단계 4에서 구한 얼굴에 움직임에 의한 시선 위치와 눈동자 움직임에 의한 시선 위치는 그림 10과 같이 2차원 벡터 합성됨으로써 최종적인 모니터상의 시선 위치가 계산되어 진다. 이에 대한 보다 자세한 내용은 5장에서 거론된다. 그림 4의 첫 번째 단계에 대한 보다 자세한 방법은 [5]에서 참조할 수 있다.

4. 얼굴의 3차원 움직임 및 이동량 추정

이 장에서 설명할 내용은 3장에서 소개한 그림 5의 2번째 단계에 속한다. 기존에 많은 3차원 움직임 추정 방법들이 연구되어 왔는데, 예를 들어 확장 칼만 필터(Extended Kalman Filter)[1], 신경망[9] 그리고 아핀 투영 방법(affine projection method)[9,10] 등이 있다. 각 방법들은 많은 제한 요소와 문제점들을 가지고 있으며[22], 이러한 이유로 본 논문에서는 기존에 3차원 움직임 추정을 위해 가장 많이 사용되는 알고리즘인 확장 칼만 필터를 사용했다. 확장 칼만 필터와 아핀 변환을 이용하여 얼굴 특징점의 변화된 3차원 위치를 구하게 되었다[13,22].

5. 모니터 상의 시선 위치 파악

5.1 얼굴의 움직임에 의한 시선 위치 파악

이 장에서 설명하는 내용은 그림 5의 3, 4번째 단계에 해당된다. 앞서 3장에서 설명한대로 그림 5의 1단계에서 모니터 좌표계를 기준으로 파악된 초기 얼굴 특징점의 3차원 위치(그림 4의 $P_1 \sim P_7$)는 얼굴 좌표계를 기준으로 변환되어 질 수 있다. 이후, 이 특징점의 위치는 사용자가 모니터 상의 한 지점을 쳐다 볼 때 발생하는 회전 $[R]$ 및 이동 $[T]$ (4장에서 설명한 확장 칼만 필

터에 의해 추정되는) 행렬에 의해 변화되며, 이 행렬을 이용한 아핀 변환 계산에 의해 변화된 얼굴 특징점의 위치를 구할 수 있게 된다. 이처럼 사용자가 모니터 상의 한 지점을 쳐다볼 때의 얼굴 특징점의 3차원 위치를 계산하면, 이를 모니터 좌표계를 기준으로 다시 변환하고, 이 특징점들이 형성하는 얼굴 평면 및 평면의 법선이 모니터와 만나는 위치가 사용자의 시선위치가 된다.

5.2 눈동자의 움직임에 의한 시선 위치 파악

앞서 5.1에서 설명한 사용자의 시선 위치는 눈동자의 움직임은 고려하지 않고 얼굴의 움직임에 의해서만 모니터상의 시선위치를 파악한 것이다. 그러나 대부분의 경우, 사용자가 모니터 상의 한 지점을 쳐다볼 때에는 얼굴 및 눈동자의 움직임이 동시에 발생한다. 그러므로 본 논문에서는 보다 정확한 시선 위치를 파악하기 위하여 사용자의 눈동자 움직임을 파악하는 연구를 수행하였다.

사용자의 눈동자 움직임은 그림 9에 나타나 있듯이 입력 영상에서 추출된 특정 정보를 기준으로 파악된다. 그림 9는 앞서 설명한 자동 상하/좌우 회전 및 초점 기능을 가진 협각 카메라에 의해 취득된 눈 영상이다. 이 그림에서 추출된 눈동자 영역 및 구석 위치가 표시되어 있다. 여기서 눈동자 및 눈구석 영역을 추출하기 위하여 원형 경계 검출(circular edge detection) 및 SVM기반 눈구석 탐플릿 방법을 사용하였다. 이 방법은 2장에서 설명한 광각 카메라에서 눈 특징점을 추출하기 위해 사용했던 방식과 거의 유사하다. 전술한 바와 같이 사용자가 그의 얼굴을 심하게 회전하는 경우, 그의 한쪽 눈이 협각 카메라에서 보이지 않게 되는 경우가 종종 발생한다. 그러므로 본 논문에서는 그림 9의 (b)와 같이 사용자가 모니터 중앙을 쳐다 볼 때 양 눈을 모두 검출하고, 그림 9의 (a), (c)와 같이 사용자가 모니터의 한 지점을 쳐다보기 위해 얼굴을 심하게 회전하는 경우 보이게 되는 한쪽 눈(왼쪽 혹은 오른쪽 눈)만을 추적하게 된다.

일반적으로 사용자의 시선 위치에 따라 눈동자의 움직임량 및 형태 정보는 변하게 된다. 특히 눈동자의 중심위치(그림 4의 Q_1, Q_2 참조)와 왼쪽 혹은 오른쪽 눈의 구석 위치(그림 4의 $P_1, P_2, P_3, P_4, P'_1, P'_2, P'_3, P'_4$ 참조)사이의 거리는 사용자의 시선 위치에 따라 변화하게 된다. 본 논문에서는 그림 8과 같이 눈동자의 움직임량 및 시선 위치 사이의 관계를 학습하기 위하여 신경망(다층 퍼셉트론)을 사용하였다. 여기서 신경망을 위한 입력 값은 사용자가 모니터 중앙을 쳐다볼 때 취득된 눈 영상에서의 눈 중심과 구석 사이의 거리로 정규화된 다. 이와 같은 정규화가 필요한 이유는 본 논문에서는 줌 렌즈(zoom lens) 없이 초점 렌즈만을 이용하여 눈 영상을 취득하고 있기 때문에, 사용자가 모니터(카메라)

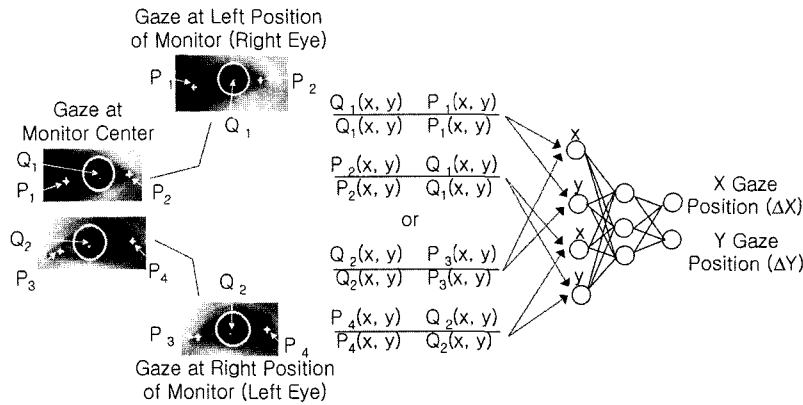


그림 8 눈동자 움직임에 의한 시선 위치 파악을 위한 특징점 및 신경망 구조

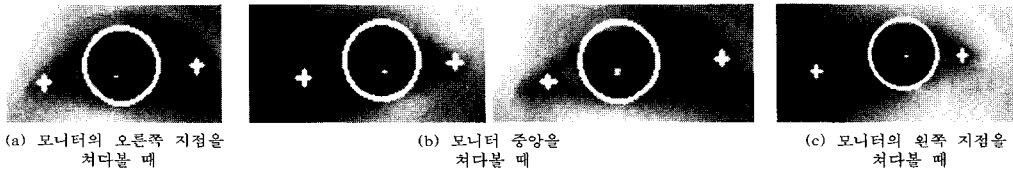


그림 9 자동 상하/좌우 회전 및 초점기능을 가진 협각 카메라를 이용한 눈 영상 취득

에 접근할수록 눈 영상이 확대되는 결과를 낳게 되기 때문이다. 이 논문에서 모니터 앞에 앉은 사용자의 Z거리는 50~70 cm 사이에 다양하기 때문에, 그림 8에서와 같은 Z 거리 정규화는 필요하다.

5.3 얼굴 및 눈동자 움직임에 의한 시선 위치 파악

이처럼 눈동자 움직임에 의한 시선 위치를 파악하게 되면, 5.1 및 5.2의 결과는 2차원 벡터 합성됨으로써 최종적으로 얼굴 및 눈동자 움직임에 의한 모니터 상의 시선위치를 그림 10과 같이 파악할 수 있게 된다. 즉, 5.1에서 설명한 대로 얼굴 움직임에 의한 시선위치를 파악한 후, 이를 기준으로 5.2와 같이 눈동자 움직임에 의한 시선 위치 변이량(ΔX, ΔY)을 파악함으로써 최종적인 모니터 상의 시선위치를 계산할 수 있게 되는 것이다.

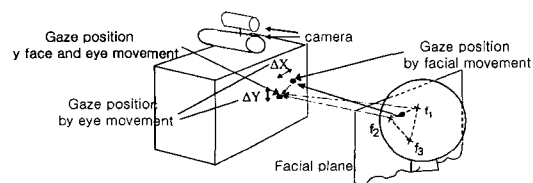


그림 10 얼굴 및 눈동자 움직임에 의한 모니터상의 시선 위치 파악

6. 실험 결과 및 고찰

6.1 얼굴 특징점 추출의 정확도

2장에서 설명한 바와 같이 본 논문에서는 광각 카메라에서 눈동자의 정확한 영역을 추출하기 위하여 원형 경계 검출 알고리즘(Circular Edge Detection)을 사용하였다. 적외선 조명을 이용하여 찾은 눈에서의 반사위치(Specular Reflection)를 기준으로 제한된 영역에서 눈동자 중심 위치를 추출하기 때문에, 정확한 위치를 추출하는 데에는 처리 시간이 거의 소요되지 않았다 (Pentium-II 550MHz PC에서 약 5~10 ms). 또한, 2장에서 설명한 눈구석 추출을 위하여 본 논문에서는 입력된 눈구석 영상을 30×30 픽셀 크기로 정규화하여 SVM의 입력으로 사용하였다. 입력 데이터로는 2000장의 연속된 영상을 취득하여(다양한 자세로 앉아 있는 20명의 사람×100장의 영상) 이로부터 8000개의 눈구석 영상 샘플(4개의 눈구석 영상 샘플/명×2000장의 영상)을 취득했으며, 이외에 1000장의 영상은 추가로 획득함으로써 SVM의 테스트를 위해 사용하였다. 실험 결과, 학습 데이터에 대한 분류 성능은 약 0.11%의 분류 에러(9/8000개)를 나타냈으며, 테스트 데이터 중에서는 약 0.2% (8/4000개)의 에러를 나타냈다. 또한, SVM을 사용했을 때의 분류 시간 역시 Pentium-II 550MHz 환경에서

13ms정도 소요됨을 알 수 있었다. 그러므로 실시간으로 얼굴 특징점 및 시선 위치를 파악하고자 하는 본 연구 분야에서는 충분히 사용 가능한 알고리즘임을 알 수 있었다.

이처럼 입력 영상에서 사용자의 눈 위치를 찾게 되면, 찾은 눈 위치에 대해 코 및 입의 상대적인 존재 가능 영역을 설정할 수 있게 된다. 이 논문에서는 설정된 존재 가능 영역 내에서, SVM을 이용하여 눈의 구석 위치를 찾았던 것과 유사한 방법으로 양 콧구멍 및 입의 끝점 위치를 찾았다. 얼굴 특징점 추출에 대한 실험 결과, 평균 1 픽셀(양 눈의 중심), 2 픽셀(양 눈의 구석 위치), 4 픽셀(양 콧구멍) 그리고 3 픽셀(입의 양 끝점)의 최소 자승 에러가 발생했음을 알 수 있었다. 본 논문에서의 입력 영상 크기는 640×480픽셀이며, 여기서 최소 자승 에러란 SVM으로 찾은 특징점의 위치와 사용자가 직접 눈으로 보고 찾은 위치 사이의 에러 값을 나타낸다. 본 논문에서는 총 3000장의 연속 영상(2000장의 학습 영상 및 1000장의 테스트 영상)을 이용하여 특징점 추출 성능을 테스트하였다[22].

6.2 얼굴 특징점의 초기 3차원 위치 추출의 정확도

그림 5의 첫 번째 단계에서 본 논문에서 사용한 방법에 의해 구한 초기 얼굴 특징점의 3차원 위치와 이 얼굴 특징점의 실제 위치(3D position tracker sensor[22]로 측정) 사이의 최소 자승 에러(RMS error)는 약 1.15cm(X축으로 0.64 cm, Y축으로 0.5cm, Z축으로 0.81 cm)의 정확도를 나타냈다. 실험을 위해서는 6.1절에서 얼굴 특징점 추출을 위해 사용했던 20명분의 데이터를 이용하였다.

6.3 얼굴 특징점의 3차원 움직임 량 추정 정확도

4장에서 설명한대로 본 논문에서는 얼굴의 3차원 움직임량 추정을 위하여 확장 칼만 필터를 사용했으며, 확

장 칼만 필터의 예측 정확도는 3차원 위치 추적 장비(polhemus sensor)[22]와 비교하였다. 실험 결과, 확장 칼만 필터와 3차원 위치 추적 장비와의 정확도 차이는 이동량 및 회전량에서 각각 1.4cm 및 2.98도의 에러 성능을 나타냈다.

6.4 모니터상의 시선 위치 추출 정확도

본 논문에서 제안하는 시선 위치 파악의 정확도는 표 1, 2와 같이 기존의 연구 결과[5,6,14,22]와 비교하였다. [5,6]의 연구에서는 본 논문에서와 같이 그림 5의 4단계를 거쳐 얼굴 특징점의 3차원 움직임 량 및 위치 등을 추정하고, 이로부터 계산되는 얼굴 평면의 법선이 모니터와 만나는 위치를 사용자의 시선 위치로 파악하는 방식을 사용하고 있다. 그러나 이 연구들에서는 사용자 눈동자 움직임을 전혀 고려하지 않았다. [19]의 연구는 [5,6]의 연구와 같이 얼굴 특징점의 3차원 위치 등을 구하지 않고, 선형 보간법, 신경망 등의 방법을 이용하여 카메라 영상에서 파악된 얼굴 특징점의 2차원 위치로부터 직접 모니터 상의 시선 위치를 매핑(mapping)하는 방법을 취하고 있다. 이 방법 역시 눈동자 움직임은 고려하지 않는다. [22]의 연구는 얼굴 및 눈동자 움직임을 모두 고려하여 모니터 상의 시선 위치를 구하였으나, 본 논문과 같이 두개의 카메라를 사용하지 않고 하나의 광각 카메라만을 이용하여 얼굴 및 눈동자 움직임을 추정하는 방식을 취했다.

본 논문에서 모니터상의 시선 위치 파악을 위한 실험 데이터로는 총 10명의 사용자가 19인치 모니터 상의 23개 지점을 쳐다볼 때 취득된 데이터를 사용하였다. 여기서 시선 위치 에러란 사용자가 모니터 상에 실제 쳐다보는 위치와 계산되어 나온 위치 사이의 최소 자승 에러를 나타낸다. 실험은 크게 두 가지로 나누어 사용자가 모니터상의 한 지점을 쳐다볼 때 얼굴의 움직임에 의해

표 1 얼굴만의 움직임만 존재하는 데이터에 대한 시선 위치 추출 정확도 (cm)

방법	선형 보간법 [19]	단일 신경망 [19]	다중 신경망 [19]	[2] 방법	[3] 방법	[22] 방법	본 논문의 방법
최소 자승 에러	5.1	4.23	4.48	5.35	5.21	3.40	2.98

표 2 얼굴 및 눈동자 움직임이 모두 존재하는 데이터에 대한 시선위치추출 정확도 (cm)

방법	선형 보간법 [19]	단일 신경망 [19]	다중 신경망 [19]	[2] 방법	[3] 방법	[22] 방법	본 논문의 방법
최소 자승 에러	11.8	11.32	8.87	7.45	6.29	4.8	3.44

서만 응시하는 경우(표 1)와 얼굴 및 눈동자 움직임을 같이 이용하여 자연스럽게 쳐다보는 경우(표 2)로 구분하여 수행하였다.

먼저 얼굴 움직임에 의해서만 쳐다보는 경우의 정확도는 표 1에 나타나 있는 것처럼 본 논문에서 제안한 방법이 기존의 많은 방법들에 비해 가장 우수함을 알 수 있다. 그리고 표 2와 같이 얼굴 및 눈동자 움직임이 같이 발생하는 실험 데이터에 대한 시선 위치 추출 정확도 역시 본 논문에서 제안하는 방식이 가장 우수함을 알 수 있다. 전반적으로는 얼굴의 움직임만이 존재하는 경우보다 얼굴 및 눈동자 움직임이 같이 존재하는 경우의 시선 위치 정확도가 다소 떨어짐을 알 수 있다.

두 번째 실험으로는 1280×1024 픽셀 해상도를 갖는 19인치 모니터에 수평 및 수직 150 픽셀 간격(2.8cm 간격)으로 반경 5 픽셀의 점들을 표시하고, 사용자가 이를 쳐다볼 때의 정확도를 측정하는 것이다. 이를 통해 총 560개의 테스트 샘플(10명분×56개의 시선 위치)을 취득하게 되었으며, 이러한 실험 조건은 Rikert의 연구[8]에서와 동일하다. 실험 결과 실제 응시 위치와 본 논문에서 제안한 방법에 의한 응시 위치 사이의 최소 자승 에러는 약 3.43cm인 것으로 나타났으며, 이러한 실험 결과는 Rikert의 방법에서의 정확도(약 5.08 cm의 최소 자승 에러)보다 우수함을 알 수 있다. 또한 Rikert는 사용자와 모니터사이의 Z거리는 학습 및 실제 사용 시 고정되었다고 가정하고 있으나, 본 논문에서 그러한 불편한 제약을 사용하고 있지 않다. 이의 일환으로 본 논문에서는 사용자의 Z거리를 변화시켜가면서(55, 60, 65cm) 시선 위치 정확도를 측정하였다. 실험 결과, 최종 자승 에러는 Z거리가 55cm일 때 3.38cm, Z거리가 60cm일 때 3.45cm 그리고 Z거리가 65cm일 때 3.49cm의 결과를 나타냈다. 전술한 바와 같이 본 논문에서의 방법은 사용자의 Z거리 변화에 거의 영향을 받지 않는다는 것을 보여주고 있다. 또한 Rikert의 방법은 처리 시간이 많이 소요되는 문제점(alphastation 333MHz에서 약 1 분)이 있었으나 본 논문의 방법은 이에 비해 상당히 빠르고 실시간으로 구현 가능한 결과를 나타내고 있다(Pentium-II 550MHz에서 약 1초 이내).

본 연구에서 개발한 시선 위치 추적 시스템의 재료비는 100달러 이하(협각 및 광각 카메라용 렌즈 가격 18달러, 자동 상하/좌우회전 및 초점 모터 가격 6 달러, 협각 및 광각용 CCD 센서 가격 20 달러, CCD 카메라를 위한 부가 칩들 8달러, 적외선 조명 가격 20달러, 카메라 케이스 및 적외선 필터 등등의 가격 20달러)이며, 이러한 경우 총 양산가는 300달러 이내(일반적으로 양산가는 재료비의 3배)가 되어 일반 소비자에게도 약 35만 원 대의 가격에 공급되어질 수 있다.

7. 결론

이 논문에서는 새로운 시선 위치 추적 방법을 제안하고 있다. 실험 결과에 나타나 있는 것처럼 시선 위치 정확도는 약 3.44cm인 것으로 알 수 있다. 이러한 시선 위치 추출 에러는 추가적인 얼굴 혹은 눈동자 움직임에 의한 시선 위치 이동(마우스 드래깅과 같은 효과)으로 보정되어질 수 있다.

참고 문헌

- [1] A. Azarbayejani. "Visually Controlled Graphics". IEEE Trans. PAMI, Vol. 15, No. 6, pp. 602-605, 1993.
- [2] J. Heinzmann et al., "Robust real-time face tracking and gesture recognition". Proc. of the IJCAI, Vol. 2, pp. 1525-1530, 1997.
- [3] T. BROIDA et al., "Recursive 3-D Motion Estimation from a Monocular Image Sequence". IEEE Trans. Aerospace and Electronic Systems, Vol. 26, No. 4, pp. 639-656, 1990.
- [4] T. Fukuhara et al., "3D-motion estimation of human head for model based image coding". IEE Proc., Vol. 140, No. 1, pp. 26-35, 1993.
- [5] K. R. Park et al., "Gaze Point Detection by Computing the 3D Positions and 3D Motions of Face," IEICE Trans. Inf. and Syst., Vol. E.83-D, No.4, pp.884-894, Apr 2000.
- [6] K. R. Park et al., "Gaze Detection by Estimating the Depth and 3D Motions of Facial Features in Monocular Images," IEICE Trans. Fundamentals, Vol. E.82-A, No. 10, pp. 2274-2284, Oct 1999.
- [7] K. OHMURA et al., "Pointing Operation Using Detection of Face Direction from a Single View". IEICE Trans. Inf. & Syst., Vol. J72-D-II, No.9, pp. 1441-1447, 1989.
- [8] P. Ballard et al., "Controlling a Computer via Facial Aspect". IEEE Trans. on SMC, Vol. 25, No. 4, pp. 669-677, 1995.
- [9] A. Gee et al., "Fast visual tracking by temporal consensus," Image and Vision Computing. Vol. 14, pp. 105-114, 1996.
- [10] J. Heinzmann et al., "3D Facial Pose and Gaze Point Estimation using a Robust Real-Time Tracking Paradigm". Proceedings of ICAFGR, pp. 142-147, 1998.
- [11] T. Rikert et al., "Gaze Estimation using Morphable Models". Proc. of ICAFGR, pp. 436-441, 1998.
- [12] Matsumoto-Y, et al., "An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement". Proc. the ICAFGR. pp. 499-504, 2000.
- [13] Newman-R et al., "Real-time stereo tracking for head pose and gaze estimation". Proceedings the 4th ICAFGR. pp. 122-8, 2000.

- [14] K. R. Park et al., "Intelligent Process Control via Gaze Detection Technology". EAAI, Vol. 13, No. 5, pp. 577-587, 2000.
- [15] A.Ali-A-L et al., "Man-machine interface through eyeball direction of gaze". Proc. of the South-eastern Symposium on System Theory, pp. 478-82, 1997.
- [16] A. TOMONO et al., "Eye Tracking Method Using an Image Pickup Apparatus". European Patent Specification-94101635, 1994.
- [17] Seika-Tenkai-Tokushuu-Go, ATR Journal, 1996.
- [18] Eyemark Recorder Model EMR-NC, NAC Image Technology Cooperation.
- [19] Porrill-J et al., "Robust and optimal use of information in stereo vision". Nature. vol. 397, no. 6714, pp. 63-6, Jan 1999.
- [20] Varchmin-AC et al., "Image based recognition of gaze direction using adaptive methods. Gesture and Sign Language in Human-Computer Interaction". Int. Gesture Workshop Proc. Berlin, Germany, pp. 245-57, 1998.
- [21] Betke-M et al., "Gaze detection via self-organizing gray-scale units". Proc. Int. Workshop on Recog., Analy., and Tracking of Faces and Gestures in Real-Time System. pp. 70-6, 1999.
- [22] K. R. Park et al., "Facial and Eye Gaze detection". LNCS, Vol.2525, pp. 368-376, 2002.



박 강 령

1994년 2월 연세대 전자공학과 졸업
 1996년 2월 연세대 전기컴퓨터공학과 석사.
 2000년 2월 연세대 전기컴퓨터공학과 박사.
 2000년 3월~2003년 2월 LG 전자 기술원, Digital Vision Group, 홍채인식팀 선임연구원.
 2003년 3월~ 상명대 미디어학부 전임강사. 관심분야는 컴퓨터 비전, 생체인식, 컴퓨터 그래픽스