

얼굴 및 눈 위치 추적을 통한 IPTV 화면 인터페이스 제어에 관한 연구

준회원 이원오*, 정회원 이의철**, 박강령*, 이희경***,
박민식***, 이한규***, 홍진우***

A Study on Controlling IPTV Interface Based on Tracking of Face and Eye Positions

Won Oh Lee* Associate Member, Eui Chul Lee**, Kang Ryoung Park*°, HeeKyung Lee***,
Minsik Park***, Han-Kyu Lee***, Jin Woo Hong*** Regular Members

요 약

최근 HCI 분야에서 사용자의 시선 추적을 통해 보다 편리한 입력 장치를 개발하려는 연구가 활발히 진행되고 있다. 하지만 기존의 시선 추적 방법들은 부가적인 사용자 착용형 장비를 필요로 하거나 원거리에서 작동되지 않는 문제 등으로 인해 IPTV 환경에서 적용하기 어려운 실정이다. 이에 본 연구에서는 사용자 착용없이 고정된 하나의 카메라를 이용하여 얼굴을 취득하고, 취득된 얼굴 영역 내에서 눈의 위치를 검출하여 IPTV의 화면 인터페이스를 제어할 수 있는 새로운 방법을 제안한다. 또한, Adaboost 방법으로 얼굴이나 눈이 성공적으로 검출되지 못했을 경우에도, 계층적 KLT (Kanade-Lucas-Tomasi) 특징 추적 방법을 통해 구해진 모션 벡터를 이용하여 화면 인터페이스를 제어할 수 있는 방법을 제안한다. 이처럼, 본 논문의 방법은 기존의 방법과는 달리 실제 IPTV의 시청 거리인 2m 정도의 원거리에서도 사용가능하며, 카메라 이외에 별도의 장치를 착용할 필요가 없으므로 편의성이 높고 얼굴 움직임의 제약이 없다는 장점이 있다.

실험결과, 입력되는 얼굴 영상을 초당 15프레임의 속도로 실시간 처리함을 확인할 수 있었으며, 기존 입력 장치의 역할을 충분히 대신할 수 있음을 알 수 있었다.

Key Words : IPTV, Human Computer Interaction, Gaze Tracking, Adaboost, KLT

ABSTRACT

Recently, many researches for making more comfortable input device based on gaze detection have been vigorously performed in human computer interaction. However, these previous researches are difficult to be used in IPTV environment because these methods need additional wearing devices or do not work at a distance. To overcome these problems, we propose a new way of controlling IPTV interface by using a detected face and eye positions in single static camera. And although face or eyes are not detected successfully by using Adaboost algorithm, we can control IPTV interface by using motion vectors calculated by pyramidal KLT (Kanade-Lucas-Tomasi) feature tracker. These are two novelties of our research compared to previous works.

* 본 연구는 지식경제부 산업원천기술개발사업의 일환으로 수행하였음. [2009-S-001-01(KI002070), IPTV용 Interactive 시점제어 기술 개발]

* 동국대학교 전자전기공학부(215p8@dongguk.edu, parkgr@dongguk.edu) (° : 교신저자), ** 국가수리과학연구소(eclee@nims.re.kr),

*** 한국전자통신연구원 (lhk@etri.re.kr, pms@etri.re.kr, hkl@etri.re.kr, jwhong@etri.re.kr)

논문번호 : KICS2009-12-639, 접수일자 : 2009년 12월 30일, 최종논문접수일자 : 2010년 5월 10일

This research has following advantages. Different from previous research, the proposed method can be used at a distance about 2m. Since the proposed method does not require a user to wear additional equipments, there is no limitation of face movement and it has high convenience.

Experimental results showed that the proposed method could be operated at real-time speed of 15 frames per second. We confirmed that the previous input device could be sufficiently replaced by the proposed method.

I. 서 론

사용자의 얼굴이나 눈의 움직임 정보를 분석하여 시선 위치를 추적하고 이를 통해 보다 자연스럽게 편리한 입력장치를 구현하려는 연구는 현재까지 활발히 진행되어 왔다^[1-7]. 시선 추적을 통한 입력장치의 장점은, 기존에 사용하던 마우스의 작동방법과 유사한 명령체계를 지니고 있어 사용자가 큰 어려움 없이 사용할 수 있다는 점이다. 또한 손을 사용하지 못하는 장애인에게 손동작을 필요로 하는 키보드, 마우스, 리모컨 등의 장치를 대신하여 컴퓨터 및 기타 기기들을 제어할 수 있는 기회를 제공한다. 는 장점이 있다.

사람의 시선 위치는 얼굴과 안구의 방향을 모두 고려하여 판단할 수 있다. 하지만 안구의 방향이 사람의 시선 위치를 결정하는 중요한 단서가 되므로, 기존의 시선 추적 방법에서는 주로 안구의 움직임을 추적하고, 얼굴 움직임에 의해 발생하는 오차를 보정하는 방식으로 연구가 진행되어 왔다. 기존의 시선 추적 방법들은 콘택트렌즈를 착용하거나, 피부전극(skin electrode)을 눈 주위에 부착하는 방법, 카메라 비전 시스템을 통한 영상 분석 방법들이 있었다^[8]. 콘택트렌즈를 이용한 방법은, 별도의 장비가 렌즈 면에 부착되어 있는 콘택트렌즈를 사용자의 각막 면에 부착되도록 착용해야 하고, 피부전극을 이용한 방법은 여러 개의 피부전극을 눈 주위에 부착해야 한다. 또한 카메라 비전 시스템을 이용한 방법은 얼굴 움직임에 영향을 받지 않고 고해상도의 눈 영상을 취득하기 위해 머리에 착용식 카메라를 장착하거나 회전 및 자동 초점 조절, 줌 기능을 가지는 고가의 카메라 장치를 사용해야 한다는 문제점이 있다. 결과적으로 세 가지 방법 모두 제작 단가 및 상용성 측면에서 단점을 가지고 있다^[1]. 이러한 문제점 때문에, IPTV 환경에서 기존 시선 추적 시스템을 적용하기에는 어려움이 있다.

일반적으로 컴퓨터에서는 사용자가 여러 가지 작업을 수행하는 경우가 많으므로 사용자 자세 및 얼굴 움직임에는 큰 변화가 없게 된다. 이에 비하여,

IPTV를 포함한 일반 TV 환경에서는 사용자가 특정 작업을 하는 경우보다는 편안한 자세로 앉아서 시청하는 경우가 많으므로 사용자의 자세 및 얼굴 움직임에 많은 자유도가 존재한다. 그러므로 기존의 컴퓨터 환경에서 개발된 알고리즘 보다, 얼굴 회전과 자세 변화에 강한 시선 추적 알고리즘이 필요하다. 본 연구에서는 이를 위해서, KLT 알고리즘을 사용하여 시청자의 얼굴이 회전하거나 포즈가 변하더라도 강인하게 얼굴의 움직임 벡터 및 얼굴 시선 위치를 추적함으로써, IPTV 인터페이스를 제어할 수 있는 연구를 수행하였다.

이를 위해 본 논문에서는 저가의 USB (Universal Serial Bus) 카메라를 이용해서 원거리에서 얼굴을 검출하고, 검출된 얼굴 영역 안에서 눈을 검출하여 그 위치 정보를 통해 IPTV의 화면 인터페이스를 조절하는 연구를 수행하였다. 또한 얼굴이나 눈이 성공적으로 검출되지 못했을 경우에도, 계층적 KLT(Kanade-Lucas-Tomasi) 특징 추적 알고리즘을 통해 구해진 모션 벡터를 이용하여 IPTV의 화면 인터페이스를 조작할 수 있는 연구를 수행하였다.

본 논문의 구성은 다음과 같다. II장에서는 본 논문에서 제안하는 알고리즘에 대해 설명하고, III장에서는 IPTV환경에서 제안하는 방법을 적용하여 수행한 실험에 대한 결과와 분석 내용을 포함하였다. IV장에서는 결론 및 추후 연구 방향을 제시한다.

II. 본 론

IPTV는 초고속 인터넷을 사용하여, 일반 텔레비전뿐만 아니라 인터넷 쇼핑, 동영상 콘텐츠, 오디오, 텍스트 데이터 등의 다양한 멀티미디어 서비스를 제공한다. IPTV는 방송 콘텐츠를 제공한다는 점에서는 일반 텔레비전과 차이가 없지만, 일반 텔레비전과는 달리 양방향성을 가진다는 특징이 있다. 이러한 양방향성 다양한 기능들 때문에, IPTV를 제어하는데 보편적으로 사용되는 리모컨(Remote Controller)은 그 크기가 점점 커지고 있으며, 그 복잡성도 증가하고 있다. 그러므로 시청자들은 IPTV

의 다양한 서비스를 이용하기가 힘들어지고, 보다 더 편리한 입력장치를 원하게 된다. 이러한 IPTV 입력장치가 갖추어야 할 요구조건들은 다음과 같다. 첫째, 시청자가 거부감을 느끼지 않는 환경에서, 편리하게 IPTV 인터페이스를 조작할 수 있어야 한다. 둘째, IPTV의 다양한 서비스를 이용하는데 불편함이 없도록, 복잡하지 않아야 한다. 이를 위해, 본 연구에서는 시청자가 자유로운 환경에서 보다 편리하게 IPTV 인터페이스를 조작할 수 있도록, 별도의 장비 착용없이 얼굴 및 눈 위치를 추적하여 IPTV 인터페이스를 제어할 수 있는 입력장치 알고리즘을 제안한다.

제안하는 알고리즘의 전체 흐름도는 다음 그림 1과 같다.

먼저 IPTV에 연결된 웹캠을 이용해서 얼굴 영상을 취득한다. 그 후 얼굴 패턴에 대한 사전 트레이닝 정보를 이용한 적응적 상승 기법 (Adaboost, Adaptive Boosting)을 이용하여 얼굴 영역을 검출한다⁹⁾. 검출된 얼굴 영역에서 미리 정해진 규칙에 의해 눈 후보 영역을 결정한 후, 사전에 눈 영역의 패턴에 대해 학습된 정보를 이용한 눈 검출을 위한 상승 분류기(Adaboost Classifier)를 통해 눈을 검출해 낸다. 검출된 눈의 위치 정보를 이용하여 IPTV 화면의 커서를 제어한다. 만약 얼굴이나 눈이 성공적으로 검출되지 못했을 경우, 계층적 KLT(Kanade-Lucas-Tomasi) 특징 추적 알고리즘을 통해 모션

벡터를 구하고, 이 정보를 이용하여 IPTV의 화면의 커서를 제어한다. 각 단계에 대한 자세한 내용은 2.1절에서 2.6절에 걸쳐 설명한다.

2.1 Adaboost기법을 이용한 얼굴 영역 검출

웹캠으로부터 얼굴 영상이 취득되면 그림 2와 같이 상승 분류기(Adaboost Classifier)를 이용하여 얼굴을 검출한다^{9,10)}. 상승 분류기란, 간단한 Haar wavelet 형태의 마스크를 사용한 통계적 모델로, 약한 분류기(Weak Classifier)를 선형적으로 결합하여 정확도가 높고 강한 분류기를 학습하는 방법이다¹⁰⁻¹²⁾. 이 때 분류기는 분류기의 배율을 조정하여 다양한 크기의 물체를 찾을 수 있다.

탐색 방법의 기본요소는 Haar 형태의 특징과 원하는 물체인지 아닌지를 구별하는 간단한 약한 분류기(Weak Classifier)이다. 각각의 특징은 템플릿(특징의 모양)으로 표현된다. Viola와 Jones⁹⁾의 연구에서는 8가지의 특징이 사용되었고, Lienhart¹¹⁻¹⁴⁾의 연구에서는 그림 2와 같이 총 14개의 특징이 사용되었다¹⁵⁾.

그림 2에서처럼, 특징들은 2개 혹은 3개의 검은 색과 흰색의 사각형으로 구성되어 있으며, 이는 수평, 수직 혹은 대각방향으로 기울어져 있다. 이러한 각각의 특징 템플릿이 얼굴 구성요소 (예, 양 눈 및 코 등, 왼쪽 눈과 눈꺼풀 등)를 나타낸다고 분수 있다.

실제의 분류기(Classifier)들은 이러한 수백 개의 특징들이 사용되기 때문에 직접적인 픽셀의 합과 같은 연산으로 처리속도가 매우 느리다. 하지만 Viola⁹⁾가 제안한 Integral 영상 방법을 사용하면 아주 빠르게 계산을 할 수 있다. 계산된 특징값은 보통 참과 거짓의 두 가지 값이나 세 가지 값을 나타낼 수 있는 아주 간단한 결정 트리 분류기 (Decision Tree Classifier)의 입력으로 사용되며, 이

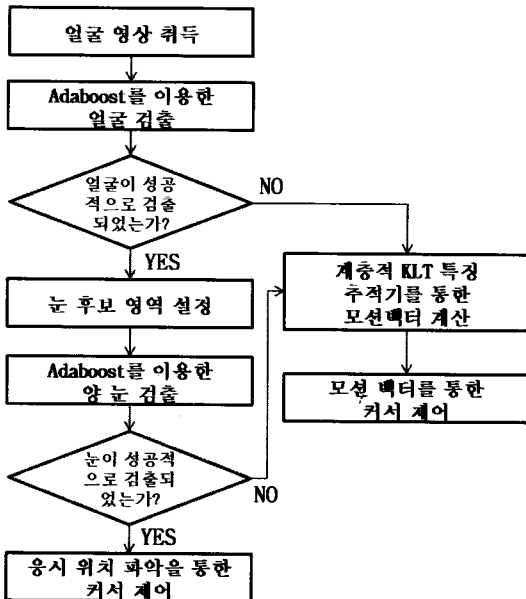


그림 1. 제안하는 알고리즘의 흐름도

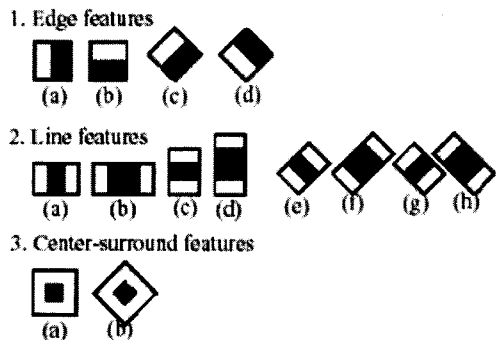


그림 2. 얼굴 검출을 위한 특징 템플릿들 [11-13]

때 각각의 분류기는 개별적으로는 원하는 물체를 찾을 수 없기 때문에 약한 분류기라고 불려진다.

다음 단계에서는 보다 복잡하면서도 강력한 분류기가 만들어진다. 이것은 Freund와 Shapire에 의해 소개된 방법이며 procedure called boosting 방법을 사용한 여러 개의 약한 분류기에서의 결과 값들을 합한 부호 값을 사용하게 된다^[16].

$$F = \text{sign}(c_1f_1 + c_2f_2 + \dots + c_nf_n) \quad (1)$$

식 1에서와 같이, 각각의 반복 단계에서 새로운 약한 분류기인 f_i 는 학습되고 누적된다. 학습 집합에서 분류기 f_i 의 에러 값이 작아질수록 계수 c_i 는 점점 커지게 된다. 그 후, 모든 학습 샘플들의 가중치가 갱신되고, 다음 반복 단계에서는 이미 만들어진 F 에 의해 잘못 분류된 샘플들의 역할이 강조되게 된다. 그것은 [16]의 연구에서 증명되었는데, 만약 f_i 가 임의로 추정된 것보다 조금이라도 크다면, F 는 임의의 큰 값을 취하게 된다. 만약 약한 분류기들의 개수가 충분히 크다면, (< 1)은 적중률이 되고, (> 0)는 해당 항목이 아니라도 해당 항목으로 분류될 확률이 포함되게 된다. 하지만, 실제로는 빠른 처리 시간을 위해서는 매우 많은 개수의 약한 분류기들 만큼이나 매우 많은 큰 학습 집합들이 요구된다^[15].

반면, Viola^[17]은 그림 3과 같이 계속적으로 복잡성을 증가시키는 부스트 분류기 F_k 를 만들었다. 그 후 간단한 분류기가 앞쪽에 있도록 직렬(Cascade) 형태로 배치를 하였다. 검출 단계에서, 그림 3과 같이 현재의 검색 윈도우는 검출하고자하는 물체를 판단하는 각각의 F_k 분류기에 의해 세부단계로 나누어진다. 이렇게 간단한 분류기를 앞에 둬으로써 검출에 필요한 연산시간을 줄일 수 있다. 여기서 각각의 단계들은 낮은 false-alarm rate보다는 높은 hit-rate를 지향한다. 즉, 얼굴이 아닌데 잘못 찾는 에러율을 줄이기보다는 얼굴인데 못 찾는 에러가

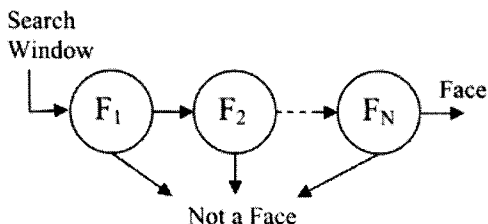


그림 3. 직렬 분류기 (Cascade Classifier)들에 의한 얼굴 검출 과정

없는 방향으로 학습한다. 이 때, 각 단계에서 이루고자 하는 hit-rate와 false-alarm rate를 미리 설정하고 각 단계에서의 정확도를 선택하면 좋은 검출 성능을 이룰 수 있다^[15].

본 연구에서는 이러한 방식으로 학습한 분류기를 이용하여, 입력 영상이 정면 얼굴인 경우에 99%의 정확도로 그림 4와 같이 얼굴 영역을 검출하였다.

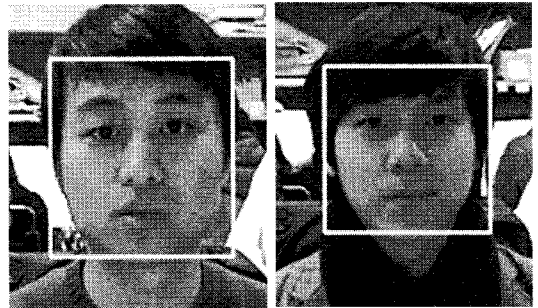


그림 4. Adaboost 얼굴 검출 기법을 이용한 얼굴 영역 검출 결과

2.2 눈 후보 영역 설정 및 눈 좌표 검출

검출된 얼굴 영역 내에서 그림 5와 같이 눈이 존재할 만한 영역을 얼굴의 비율을 이용하여 지정한다.

이 때 눈의 후보 영역은 검출된 얼굴영역을 수직 방향으로 4등분 하였을 때, 상위 두 번째 영역으로 실험적으로 결정하였다. 이는 32명분의 192장의 얼굴 영상을 분석한 결과, 눈이 모두 포함되는 영역을 기반으로 정한 것으로, 이렇게 눈 후보 영역을 결정함으로써, 처리 시간의 감소와 눈이 아닌 영역을 눈으로 잘못 검출하는 에러를 줄이는 효과를 볼 수 있다.

이 때 눈 후보 영역은 검출된 얼굴 영역에서 수직 방향으로만 탐색 범위를 좁혔을 뿐 수평 방향에 대해서는 탐색 범위를 좁히지 않았다. 이는 얼굴이

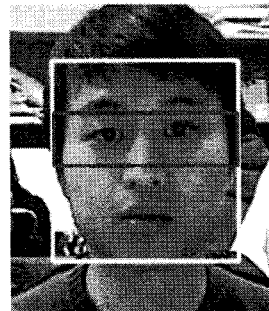


그림 5. 눈 후보 영역 지정

회전할 시 눈 영역이 얼굴 영역의 가장자리에 존재할 수 있기 때문이다.

2.3 눈동자 중심 검출

이 후, 눈 후보 영역 안에서 사전에 눈 영역만으로 학습된 정보를 이용한 상승 분류기 (Adaboost Classifier)를 통해 그림 6과 같이 눈 영역을 검출하였다. 검출된 눈 영역 안에서 Gonzalez와 Woods가 제안하는 적응적 임계치 추정 방법을 사용하여^[18] 이진화 연산을 수행하였다.

본 연구에서 사용한 적응적 임계치 추정 방법은 다음과 같다. 임계값 T의 초기 값을 임의로 추정하고 그 임계값 T로 먼저 이진화를 실시한다. 그 결과는 밝기 값이 T보다 크거나 같은 화소들로 구성된 G1 그룹과 밝기 값이 T보다 작은 화소들로 구성된 G2의 두 그룹으로 나누어진다. 영역 G1과 G2에 대해서 화소들의 평균 밝기 값 μ_1 과 μ_2 를 계산한다. 영역 G1과 G2의 평균값을 이용하여 새로운 임계값을 식(2)와 같이 계산한다.

$$T_{new} = \frac{1}{2}(\mu_1 + \mu_2) \quad (2)$$

연속적으로 T의 변화량이 미리 정의된 변화량 ΔT 보다 작을 때까지 이러한 과정을 반복한다^[19].

일반적으로 눈 동공 영역의 그레이 레벨이 주위 영역보다 낮다는 특성을 이용하여^[20], 이진화 후 동공을 분리해 내고, 눈 영역 안에 포함된 눈썹이나 기타 노이즈의 영향을 줄이기 위해서 성분 명명화 (component labeling) 기법을 적용한다. 최종적으로 분리된 동공 영역의 흑화소의 무게 중심을 구하여 동공의 정확한 중심을 구한다.

2.4 응시 위치 파악

응시 위치 파악을 위해서는 입력 영상에서 검출

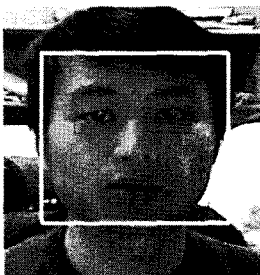


그림 6. 눈 영역 검출

된 눈 영역의 좌표(S_x, S_y)가 그림 7과 같이 IPTV 화면의 논리 좌표의 한 위치(M_x, M_y)로 사상 (Mapping)되어야 한다. 이 때, 그림 7의 (M_{x1}, M_{y1}), (M_{x2}, M_{y2}), (M_{x3}, M_{y3}), (M_{x4}, M_{y4})는 각각 IPTV 스크린의 좌측 상단, 우측 상단, 좌측 하단, 우측 하단 위치를 나타낸다.

그림 7에서 보는 바와 같이 초기에 IPTV의 중앙 지점을 쳐다볼 때 추출된 양 눈의 중심좌표 (S_x, S_y)를 스크린 상의 중앙 좌표 ($\frac{M_{x1} + M_{x2}}{2}, \frac{M_{y1} + M_{y3}}{2}$)로 대응시키고, 그 후 식(3)과 (4)에 의하여 스크린 상에 사상되는 커서의 위치를 계산한다.

$$M_x = \frac{(M_{x1} + M_{x2})}{2} + (C_x - S_x) \times \frac{(M_{x1} - M_{x2})}{\alpha} \quad (3)$$

$$M_y = \frac{(M_{y1} + M_{y3})}{2} + (C_y - S_y) \times \frac{(M_{y1} - M_{y3})}{\beta} \quad (4)$$

이 때, (C_x, C_y)는 이후 연속프레임에서 움직여진 양 눈의 중심 좌표를, (M_x, M_y)는 모니터 상에 사상되는 커서의 위치를 나타낸다. α 와 β 는 눈 영상에서 (C_x, C_y)의 변화에 따라 (M_x, M_y)의 변화량을 변화시켜서 커서 조절의 감도를 조정하기 위한 상수로써, 그 값이 작으면 (C_x, C_y)의 작은 변화에도 화면상의 커서를 많이 움직일 수 있다. 이런 경우 커서 이동이 민감하므로 화면의 여러 픽셀을 한꺼번에 건너뛰게 되므로 커서의 정교한 조정을 할 수 없어진다. 반대로 α 와 β 값이 커지면 정교한 조정이 가능한 대신, 커서를 움직이기 위해 많은 얼굴 움직임이 필요하므로 IPTV의 응용 프로그램에 따라 목적에 맞게 α 와 β 값을 적절하게 조절해야 한다. 본 연구에서는 시스템 사용 초기에 사용자로 하여금 선호도에 따라 α 와 β 를 설정할 수 있도록 하는

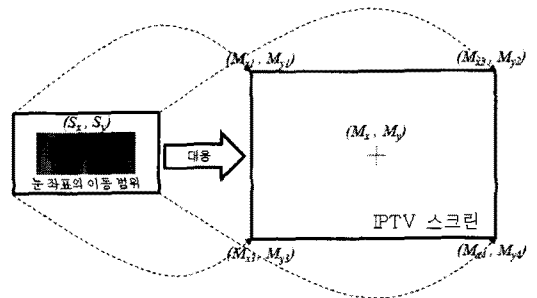


그림 7. 눈 좌표에서 IPTV 화면 좌표로의 사상 과정에 대한 개념도

기능을 제공한다.

또한, 본 연구에서는 카메라의 디지털 줌 기능을 사용하여 사용자의 작은 움직임으로도 커서가 많이 움직일 수 있도록 하여 사용자가 얼굴을 많이 움직여야 하는 불편함을 최소화 하였다. 또한 (C_x, C_y) 좌표가 영상의 이전 프레임과 대비하여 2픽셀 이하로 움직였을 때에는 이전 프레임에서 검출된 양 눈의 중심 위치로 (C_x, C_y) 를 대체함으로써, 커서의 떨림 보정 기능을 추가하였다.

2.5 계층적 KLT 특징 추적기 통한 모션벡터 계산

본 연구에서 얼굴 및 눈 검출을 위해 사용하는 Adaboost 기법은 정면 얼굴인 경우 검출하고자 하는 물체를 높은 정확도로 안정적으로 검출하지만, 측면 얼굴인 경우 그 정확도가 매우 떨어진다. 이러한 특성 때문에, IPTV 시청 환경에서, Adaboost 기법은 사용자의 자연스러운 얼굴 움직임을 검출을 지원해 주지 못한다. 즉, 사용자는 IPTV 인터페이스를 조작할 시 정면얼굴만을 사용해야 하는 불편함을 겪게 된다. 이러한 문제를 해결하기 위해 본 연구에서는 특징점 추적 알고리즘 중 계층적 KLT 특징 추적기(Pyramid Kanade-Lucas-Tomasi Feature Tracker)^[21]을 사용하여 Adaboost가 얼굴이나 눈을 검출하지 못했을 경우에도 자연스럽게 커서를 조정할 수 있도록 하였다.

KLT 특징 추적기는 평균 제곱오차(Mean Squared Error)를 기반으로 상관관계(Correlation)를 분석하고 이를 통해 특징점의 유사도를 측정하여 특징점의 위치 변화를 추적하는 방법이다. 이 때 특징점 윈도우의 크기가 크면, 움직임이 큰 변위를 얻을 수 있지만, 윈도우의 평균을 계산하는 연산들로 인해 작은 변위에 대한 추적 성능이 저하된다. 또한 그에 따른 연산량과 처리 속도도 증가하므로 윈도우 크기에 대한 적절한 선택이 필요하다. 이러한 문제를 극복하기 위하여 Bouguet^[22]는 KLT 특징 추적기의 계층식 구현을 제안하였고, 이를 통해, 변위가 큰 특징점과 세밀한 움직임을 갖는 특징점 모두에 강한 추적기가 개발되었다.

KLT 추적 알고리즘은 연속하는 두 프레임 사이의 특징점의 변위를 계산하여 특징점의 대응관계를 얻는다. 움직임 모델은 이미지 밝기 패턴의 움직임과 관련하여 식(5)과 같이 표현된다.

$$J(x) = I(x-d) - \eta(x) \quad (5)$$

이 때, 함수 J 는 특정 프레임에서의 밝기 값, x

는 픽셀의 위치를, d 는 변위 벡터를, η 는 형태의 변화로 인한 차이를, 함수 J 는 다음 프레임에서의 밝기 값을 각각 나타낸다. 특징점의 추적은 두 영상에 대응되는 해당 특징점 윈도우 W 의 평균 제곱오차 ϵ 를 식(6)과 같이 최소화하는 방향으로 결정된다.

$$\epsilon = \sum_{j \in W} (I(x_j-d) - J(x_j))^2 \quad (6)$$

밝기 함수를 테일러급수 확장에 의해 근사화할 수 있고, 이를 통해 변위 d 를 구할 수 있다. 이를 반복하여 ϵ 를 최소화하는 변위 d 를 찾아 준다.

그림 8은 계층적 KLT 특징 추적기를 이용하여 모션벡터를 찾아준 결과이다.

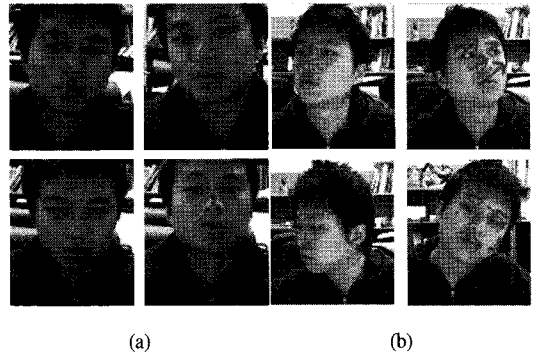


그림 8. 계층적 KLT 특징 추적기를 이용하여 모션벡터를 찾아준 결과 (a) 정면얼굴에서의 모션벡터 (b) 회전된 얼굴에서의 모션벡터

2.6 모션벡터를 통한 커서 제어

기존에 일반적으로 많이 사용하는 Adaboost face detector의 경우, 정면 얼굴 위주로 학습이 되어있어서 회전얼굴의 경우 검출율이 많이 떨어지는 문제점이 있다. 이를 해결하기 위해서는 정면 뿐 아니라 많은 회전얼굴을 포함하여 Adaboost 분류기를 다시 학습해야하지만 이는 시간 및 노력이 상당히 많이 요구되는 어려운 작업이다.

이러한 문제점을 해결하기 위하여 본 연구에서는, Adaboost 분류기를 이용하여 얼굴이나 눈을 성공적으로 검출하지 못했을 경우, 식 (7)과 (8)을 기반으로 KLT 특징 추적기를 이용하여 찾아준 모션벡터를 통해 커서를 제어한다.

$$M_x = M_{x_{prev}} + \sum_{i=0}^n \frac{d_{xi}}{n} \quad (7)$$

$$M_y = M_{yprev} + \sum_{i=0}^n \frac{\overrightarrow{d_{yi}}}{n} \quad (8)$$

이 때, (M_x, M_y) 는 모니터 상에 사상되는 좌표를, (M_{xprev}, M_{yprev}) 는 전 커서의 좌표를, $(\overrightarrow{d_{xi}}, \overrightarrow{d_{yi}})$ 는 KLT 특징 추적을 통해 구해진 모션벡터의 x, y 축 성분들을 각각 나타낸다.

제안하는 KLT 특징 추적 방법을 통해 시청자가 IPTV를 시청할 시, 얼굴의 회전이거나 자세에 제약받지 않는 자유로운 움직임이 가능하게 되었다.

III. 실험 결과 및 분석

본 연구는 아래와 같은 실험 환경에서 진행되었다. IPTV용 컴퓨터 하드웨어는 Intel(R) Core 2™ Duo CPU 2.4GHz, 3GB RAM를 포함하며, 카메라는 Logitech PN 960-00057을 이용하였다. IPTV용 모니터는 19인치 크기, 모니터 해상도는 1260 × 1024 픽셀이며, 사용자는 모니터와 80cm 이격된 거리에 위치하여 실험을 실시하였다. 본 연구에서 제안하는 프로그램은 Microsoft Visual Studio C++ 2008 개발환경에서 DirectX 9.0c SDK (Software Development Kit)의 DirectShow를 사용하여 구현하였다.

실험은 IPTV 화면에 보여지는 메뉴들의 크기를 고려하여 α 와 β 값을 (C_x, C_y) 와 (M_x, M_y) 의 움직임 비가 1:4가 되도록 설정하고 진행하였다.

3.1 객관적 성능 평가

본 연구에서는 화면에 무작위로(Randomly) 흩어져 있는 10개의 위치에, 사용자 얼굴 시선 위치에 따라 움직이는 마우스 커서를 순서대로 모두 위치 시키는데 걸리는 시간의 평균치를 총 25명의 실험 대상자로부터 측정하였다. 실험 결과는 그림 9와 같다. 실험 결과, 제안하는 방법은 초기에는 시간이 많이 걸리지만, 반복 사용에 따라 아이콘에 도달하는 시간이 감소하여 기계적인 마우스를 사용했을 경우와 거의 비슷해지는 것을 알 수 있다. 기계적인 마우스는 횡수의 변화에 따라 소요시간에 큰 변화가 없다. 이는 사용자 대부분이 기계적인 마우스에 익숙해져 있고, 제안하는 알고리즘에는 익숙하지 못하여 발생하는 결과로 판단된다.

제안하는 방법을 통해 얼굴을 포함하는 영상 한 프레임을 처리하여 IPTV 화면상의 시선 위치를 계

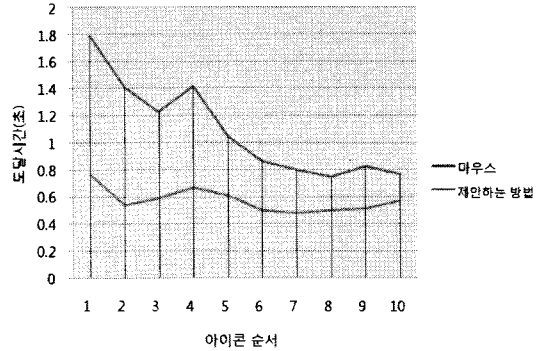


그림 9. 제안 하는 방법과 일반 마우스를 이용하여 무작위하게 위치한 아이콘에 도달하는데 걸리는 시간

산하는데 평균 64ms의 처리시간이 소요됨을 확인하였다. 이를 통해 제안하는 방법이 초당 15프레임의 영상을 취득하는 USB 카메라를 사용할 때 실시간으로 동작할 수 있음을 알 수 있었다.

3.2 주관적 성능 평가

주관적인 성능 평가에서는 총 25명의 실험 대상자에게 IPTV환경에서 기존의 기계적인 마우스를 사용하는 경우와 제안하는 방법 중 어느 것에 더 편의성 및 관심도를 느끼는 지에 대해 측정하였다^[23-25]. 편의성과 관심도에 관한 실험결과는 그림 10과 같다. 관심도에 관한 설문 조사를 보면 '제안하는 방법이 기계적인 마우스를 사용하는 것보다 관심도가 높다.'라는 의견이 많은 것을 알 수 있다. 반면에 편의성에 관한 조사에서는 기존의 마우스를 이용한 입력 방식이 제안하는 방법보다 편하다는 의견이 약간 많았다. 그 이유로는 '정확도가 마우스보다 떨어진다.', '손을 움직여 조작하는 것보다 얼굴을 움직여 조작하는 것이 더 불편하다.'를 꼽았다. 이와 같은 실험 결과, 편의성은 기계적인 마우스와 큰 차이가 나지 않는 반면, 관심도는 제안하는 방법이 기존 마우스에 비해 매우 높은 것으로 나타났다. 향후 연구를 통해 제안한 방법의 정확도를 일반 마우스 사용 수준으로 개선하여, IPTV에 적용할 경우 사용자의 관심도와 편의성이 모두 만족될 것으로 판단된다.

실험을 통해 제안하는 방법이 IPTV환경의 다양한 서비스 (인터넷 검색, Video-On-Demand 등)에서 사용할 수 있음을 확인하였고, 또한 TV시청 시 실내조명 상황에 관계없이 본 방법을 사용할 수 있음을 알 수 있었다.

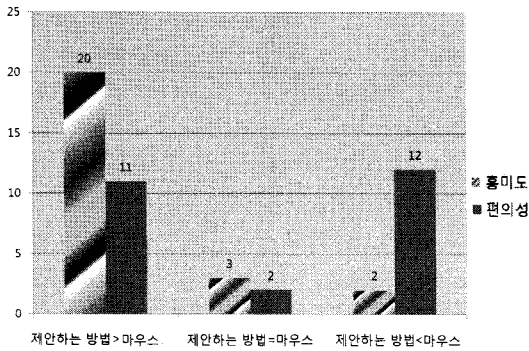


그림 10. 관심도와 편의성에 관한 설문 결과

IV. 결 론

본 논문에서는 입력영상에서 얼굴을 검출하고, 눈의 위치를 찾아 IPTV의 화면 인터페이스를 조작하거나, 얼굴이나 눈이 성공적으로 검출되지 않았을 시 계층적 KLT 특징 추적기를 통해 구해진 모션벡터를 이용하여 IPTV의 화면 인터페이스를 조작하는 연구를 수행하였다. 실험을 통해 제안하는 방법을 사용한다면 손이 불편한 사용자가 큰 불편을 겪지 않고 얼굴과 눈의 움직임만으로 IPTV를 조작할 수 있을 것이며, 일반 사용자도 별도의 장비 착용없이 거부감 없는 조작이 가능할 것으로 판단되었다.

본 연구 결과는 IPTV를 통한 인터넷 쇼핑이나 Video On Demand (VOD) 서비스 등 사용자와 TV사이의 인터랙션을 필요로 하는 쌍방향 서비스 분야에 유용하게 사용될 수 있을 것이다.

향후 보다 다양한 환경에서의 실험과 정확도 향상에 관한 연구가 진행돼야 할 것으로 예상된다.

참 고 문 헌

- [1] You Jin Ko and Kang Ryoung Park "Gaze Detection Based on Facial Features and Linear Interpolation on Mobile Devices," *Journal of Korea Multimedia Society* Vol. 12, No.8, August 2009
- [2] Jacob, R. and J. K., "Eye Movement-Based Human-Computer Interaction Techniques : Toward Non-Command Interfaces," *Advances in Human- Computer Interaction, H. R. Hartson and D. Hix, Editors., Ablex Publishing Co., pp.151-190, 1993.*
- [3] L. E. Sibert, and R. J. K. Jacob, "Evaluation of Eye Gaze Interaction," *Proceedings of the SIGCHI Conference*, pp.281-288, 2000.
- [4] Ohno, T. "Quick Menu Selection Task with Eye Mark," *Transactions of Information Processing Society of Japan*, Vol.40, No.2, pp.602-612, 1999.
- [5] Yamato, M., Monden, A., Matsumoto, K., Inoue, K. and Torii, K, "Quick Button Selection with Eye Gazing for General GUI Environments," *International Conference on Software: Theory and Practice*, August 2000.
- [6] Andrew T. Duchowski, Vinay Shiva shankariah, Tim Rawls, Anand K.Gramopadhye, Brian J. Melloy, and Barbara Kanki, "Binocular Eye Tracking in Virtual Reality for Inspection Training," *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, pp.89-96, 2000.
- [7] Jeong Jun Lee, Kang Ryoung Park, and Jaihie Kim, "Gaze Detection System under HMD Environment for User Interface," *Joint Conference of ICANN/ICONIP*, June 2003.
- [8] 이의철, 박강령, "HMD (Head Mounted Display)에서 시선 추적을 통한 3차원 게임 조작 방법 연구," *대한전자공학회 논문지*, 제 45권 2호 sp편, pp.49-64, 2008년 3월.
- [9] P. Viola, M. J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, Vol.57, No.2, pp.137-154, 2004.
- [10] Eui Chul Lee, You Jin Ko and Kang Ryoung Park, "Gaze Tracking Based on AAM and Multiple SVR on Mobile Device," *Optical Engineering*, Vol. 48, No. 7, pp.077002-1 - 077002-11, 2009.
- [11] Yoav Freund, Robert E. Schapire "A Short Introduction to Boosting", *Journal of Japanese Society for Artificial Intelligence*, pp.771-780, September 1999.
- [12] 이우주, 김진철, 이배호 "AdaBoost 알고리즘을 이용한 실시간 얼굴 검출 및 추적", *한국멀티미디어학회지*, 9권, 10호, pp.1266-1275,

2006년 10월.

[13] Rainer Lienhart and Jochen Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection," ICIP, 2002.

[14] Alexander Kuranov, Rainer Lienhart, and Vadim Pisarevsky, "An Empirical Analysis of Boosting Algorithms for Rapid Objects With an Extended Set of Haar-like Features," Intel Technical Report MRL-TR-July02-01, 2002.

[15] 박성효, 조달호, 박강령, "Adaboost를 이용한 모바일 환경에서의 홍채인식을 위한 눈 검출에 관한 연구", 대한전자공학회 논문지, 제 45권 CI편 제 4호, pp.1-11, 2008년 7월

[16] Byung Jun Kang, Kang Ryoung Park, "A Robust Eyelash Detection Based on Iris Focus Assessment", Pattern Recognition Letters, Vol. 28, Issue 13, 1 October 2007, pp.1630-1639

[17] Paul Viola and Michael J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE CVPR, 2001.

[18] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing Second Edition", 2002 by Prentice-Hall, Inc.

[19] 조재수, 강현수, 김홍수, 김성득 저 "멀티미디어 신호처리", 사이텍 미디어 p.41

[20] Dal-ho Cho, Kang Ryoung Park, Dae Woong Rhee, Yanggon Kim, and Jonghoon Yang, "Pupil and Iris Localization for Iris Recognition in Mobile Phones," Proceedings of SNPD, Las Vegas Nevada, USA, June 19-20, 2006.

[21] C. Tomasi and T. Kanade, "Detection and tracking of point features," Technical Report CMU-CS-91-132, Carnegie Mellon University, April 1991.

[22] J. Y. Bouguet, "Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm," Intel Corporation, 2003.

[23] 이의철, 박강령, "눈동자 시선 추적에 의한 3차원 1인칭 슈팅 게임", 한국정보처리학회 논문지, 제12-B권 제4호, pp.465~472, 2005년 8월

[24] 이의철, 박강령, "사용자 시선 추적을 통한 3차원 FPS 게임 화면 조정", 2006년 한국 차세대컴퓨팅학회 논문 발표회, 일산 KINTEX, 2006.11.16-17

[25] Eui Chul Lee, Kang Ryoung Park, Min Cheol Whang, Junseok Park, "Robust Gaze Tracking Method for Stereoscopic Virtual Reality System", Lecture Notes in Computer Science (HCI International 2007), Vol. 4552, pp.700-709, Beijing, China, July 22-27, 2007

[26] Eui Chul Lee, Kang Ryoung Park, Min Cheol Whang, Joa Sang Lim, "Vision-based Gaze Estimation Method For Stereoscopic Display Systems", 10th International Federation of Automatic Control (IFAC), Ritz-Carlton Hotel, Seoul, Korea, 4-6 Sep., 2007

이 원 오 (Lee Won Oh)

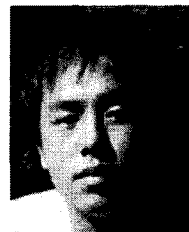
준회원



2009년 동국대학교 전자공학과 학사
2009년~현재 동국대학교 전자공학과 석사과정
<관심분야> 영상처리, 패턴인식, 컴퓨터 비전, HCI

이 의 철 (Lee Eui Chul)

정회원



2005년 상명대학교 소프트웨어 학부
2007년 상명대학교 컴퓨터과학과 석사
2010년 상명대학교 컴퓨터과학과 박사
2010년~현재 국가수리과학연구원 연구원
<관심분야> 영상처리, 패턴인식, 컴퓨터 비전, HCI

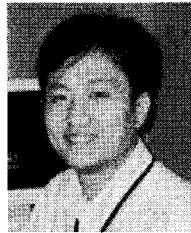
박 강 령 (Park Kang Ryoung) 정회원



1994년 연세대학교 전자공학과
학사
1996년 연세대학교 전자공학과
석사
2000년 연세대학교 전기·컴퓨터
공학과 박사
2000년~2003년 LG 전자 기

술원 Digital Vision Group 선임연구원
2003년~2008년 상명대학교 소프트웨어대학 디지털
미디어학부 조교수
2008년~현재 동국대학교 전자공학과 부교수
2003년~현재 생체인식연구센터 총괄과제책임자
<관심분야> Biometric 영상처리, 패턴인식, 컴퓨터
vision, 컴퓨터 그래픽스

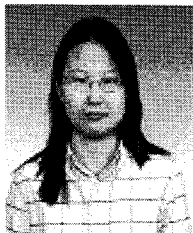
이 한 규 (Han-Kyu Lee) 정회원



1994년 경북대학교 전자공학
과 학사
1996년 경북대학교 전자공학
과 석사
2005년 한국정보통신대학교
공학부 박사과정 수료
1996년~현재 한국전자통신연
구원 방송미디어연구그룹

선임연구원
2005년~현재 한국전자통신연구원 방송미디어연구
그룹 융합서비스연구팀장
<관심분야> 영상신호처리 및 분석, 맞춤형방송 기
술 및 표준화

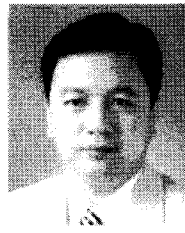
이 희 경 (HeeKyung Lee) 정회원



1999년 영남대학교 컴퓨터공학
과 학사
2002년 한국정보통신대학원대
학교 공학부 석사
2002년~현재 한국전자통신연
구원 방통융합미디어 연구부
선임연구원

<관심분야> MPEG-7, TV-Anytime, 디지털 방송,
맞춤형 방송, 콘텐츠 적응 변환, 시선추적

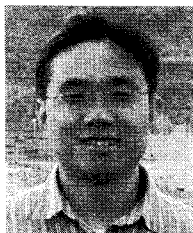
홍 진 우 (Jin Woo Hong) 정회원



1982년 광운대학교 응용전자공
학과 학사
1984년 광운대학교 전자공학과
석사
1993년 광운대학교 전자계산기
공학과 박사
현재 한국전자통신연구원 방송

통신융합미래기술연구부 부장
<관심분야> 통방융합 프레임워크 기술, 디지털방송
기술, 미디어 처리 기술, 디지털 콘텐츠 보호관리
기술

박 민 식 (Minsik Park) 정회원



1997년 광운대학교 전기공학과
학사
1999년 광주과학기술원 기전공
학과
2009년 충남대학교 컴퓨터 공
학 박사과정 수료
1999년~현재 한국전자통신연

구원 선임연구원
<관심분야> 디지털방송, 데이터방송, 맞춤형방송,
기계학습, 데이터 마이닝, HCI