

SVM 기반의 시선 인식 시스템의 구현

이 규 범[†] · 김 동 주[†] · 홍 광 석^{††}

요 약

시선 인식에 관한 연구는 현재 사용자가 응시하고 있는 위치를 파악하는 것으로 많은 응용 분야를 가지며 지속적으로 발전되어 왔다. 기존의 시선 인식에 관한 대부분의 연구는 적외선 LED 및 카메라, 고가의 헤드마운트드 장비 등을 이용하였기 때문에 범용 사용에 문제점을 가지고 있다. 이에 본 논문에서는 한 대의 PC용 웹 카메라를 사용한 SVM(Support Vector Machine) 기반의 시선 인식 시스템을 제안하고 구현한다. 제안한 시스템은 4방향과 9방향의 시선을 인식하기 위해 모니터를 가로 6, 세로 6, 총 36개의 시선 위치로 나누어 각각 9개, 4개씩 그룹핑 및 학습하여 사용자의 시선을 인식한다. 또한, 시선 인식의 성능을 높이기 위해 차영상 엔트로피를 이용한 영상 필터링 방법을 적용한다. 제안한 시스템의 성능을 평가하기 위하여 기존에 제시되었던 차영상 엔트로피 기반의 시선 인식 시스템, 눈동자 중심점과 눈의 끝점을 이용한 시선 인식 시스템, PCA 기반의 시선 인식 시스템을 구현하고 비교 실험을 수행하였다. 실험 결과 본 논문에서 제안한 SVM 기반의 시선 인식 시스템이 4방향은 94.42%, 9방향은 81.33%의 인식 성능을 보였으며, 차영상 엔트로피를 이용한 영상 필터링 방법을 적용하였을 경우에 4방향은 95.37%, 9방향은 82.25%의 성능을 보여 기존의 시선 인식 시스템보다 높은 성능을 나타내었다.

키워드 : 서포트 벡터 머신, 시선 인식, 차영상 엔트로피

An Implementation of Gaze Recognition System Based on SVM

Kue-Bum Lee[†] · Dong-Ju Kim[†] · Kwang-Seok Hong^{††}

ABSTRACT

The researches about gaze recognition which current user gazes and finds the location have increasingly developed to have many application. The gaze recognition of existence all about researches have got problems because of using equipment that Infrared(IR) LED, IR camera and head-mounted of high price. This study propose and implement the gaze recognition system based on SVM using a single PC Web camera. The proposed system that divide the gaze location of 36 per 9 and 4 to recognize gaze location of 4 direction and 9 direction recognize user's gaze. Also, the proposed system had apply on image filtering method using difference image entropy to improve performance of gaze recognition. The propose system was implements experiments on the comparison of proposed difference image entropy gaze recognition system, gaze recognition system using eye corner and eye's center and gaze recognition system based on PCA to evaluate performance of proposed system. The experimental results, recognition rate of 4 direction was 94.42% and 9 direction was 81.33% for the gaze recognition system based on proposed SVM. 4 direction was 95.37% and 9 direction was 82.25%, when image filtering method using difference image entropy implemented. The experimental results proved the high performance better than existed gaze recognition system.

Keywords : Support Vector Machine : SVM, Gaze Recognition, Difference Image Entropy

1. 서 론

시선 인식은 현재 사용자가 응시하는 위치를 시각적인 방법으로 파악하는 연구이다. 이러한 시선 인식 기술을 이용

한 인터페이스는 인간과 컴퓨터의 상호작용(Human Computer Interaction : HCI)에 있어서 매우 유용하게 사용될 수 있기 때문에 휴먼 인터페이스의 한 연구 분야로 많은 응용 분야를 가지며, 이에 대한 연구가 활발히 진행되고 있다.

대표적인 예로, 양 손을 사용하지 못하는 장애인을 위한 키보드 타이핑, 마우스의 커서 움직임 등을 사용자의 시선으로 대신 할 수 있고[1], 웹의 사용성 평가를 위한 정보 수집 방법으로 사용자의 지각 행위를 관찰하기 위해 시선 인식 기법이 이용되고 있다[2].

기존의 시선 인식 시스템은 얼굴의 움직임을 파악하여 시

※ 본 연구는 2008년도 중소기업기술혁신개발사업 (S1053744) 및 2009년도 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음(NIPA-2009-(C1090-0902-0046)).

† 준 회 원 : 성균관대학교 정보통신공학부 박사과정

†† 중 심 회 원 : 성균관대학교 정보통신공학부 교수

논문접수: 2009년 11월 30일

수정일: 1차 2009년 12월 24일

심사완료: 2010년 1월 14일

선을 인식하는 방법과, 눈동자의 움직임을 파악하여 시선을 인식하는 방법이 있다.

얼굴의 움직임을 이용한 시선 인식은 별도의 마커를 얼굴의 특정 위치에 부착하고, 이 마커의 위치를 파악하여 얼굴의 움직임을 추적함으로써 커서의 움직임을 제어하거나[3], 눈, 코, 입 등 얼굴의 특징점을 검출하고, 이 특징점들의 위치를 파악함으로써 사용자의 시선을 인식하는 방법이다[4]. 이러한 시스템들은 사용자의 시선을 인식하기 위해 별도의 마커를 얼굴의 특정 위치에 부착하여 얼굴의 움직임을 인식하였고, 마커 또는 얼굴의 특징점이 카메라의 범위를 벗어나면 시선 인식 성능이 저하되거나, 미세한 시선의 변화를 감지하지 못하는 단점을 지니고 있다.

눈동자의 움직임을 이용한 시선 인식은 동공과 각막 반사에 의한 글린트 사이의 공간적 위치 특성으로 동공을 식별하고, 동공의 위치와 형태, 왜곡 등의 정보로 시선을 인식하거나[5], 적외선 LED(Light Emittted Diode)와 2대의 웹 카메라를 이용하여 시선을 인식하는 방법이 있다[6]. 이 시스템들은 적외선 LED를 사용하여 동공을 검출하거나 웹 카메라를 머리에 장착하여 눈동자의 움직임을 추적하기 때문에 사용자에게 불편함을 초래할 수 있다는 단점이 있다.

따라서, 본 논문에서는 기존 시선 인식 시스템들의 단점을 극복하기 위해 SVM(Support Vector Machine) 기반의 시선 인식 시스템을 제안한다. 제안한 시스템은 4방향과 9방향의 시선 위치를 인식하기 위해 데스크 탑 모니터를 가로6, 세로6, 총 36개의 시선 위치로 나누어 각각 9개, 4개씩 그룹핑하여 사용자의 시선을 인식한다. 사용자의 시선을 인식하기 위한 첫 번째 단계로, Haar-like feature 기반의 AdaBoost 학습 알고리즘을 사용하여 입력된 영상으로부터 얼굴 영역을 검출하고, 검출된 얼굴 영역에서 눈 영역을 검출한다. 두 번째 단계로, 차영상 엔트로피를 이용한 영상 필터링 방법을 적용하여 검출된 눈 영역 영상을 시선 인식에 사용할 것인지 사용하지 않을 것인지 결정한다. 세 번째 단계로, 눈 영역 영상을 SVM 알고리즘을 이용하여 시선을 인식한다.

SVM은 최근 가장 널리 사용되는 분류 알고리즘들 중의 하나로 다른 알고리즘에 비해 높은 성능을 보이고, 지문 인식, 얼굴 인식, 문자 인식 등 많은 연구 분야 사용되고 있지만, 기존의 시선 인식 시스템에서는 거의 사용되지 않았거나, 적외선 조명 및 카메라 기반의 시선 인식 시스템에 사용되어져 왔다. 그 예로, Jie Xiang 등의 연구에서는 Tobii Eye Tracker를 사용하여 데이터를 획득하고, 이 데이터를 SVM 알고리즘에 적용하여 시선을 인식하였다[7]. Tobii Eye Tracker 시스템 역시 내장된 적외선 조명을 이용하여 동공과 각막에 의한 반사광을 추적함으로써 시선을 인식한다. 또한, Li Su-mei 등의 연구에서는 사용자의 시선 방향이 아닌 카메라의 응시 여부를 판단하는 연구를 수행하였다[8]. 따라서, 본 제안은 시선 인식을 위한 새로운 방법론을 제시할 수 있으리라 기대되며, 고가의 헤드 마운트드 장치나 적외선 LED, 상업용 Eye Tracker 등과 같은 별도의 장

치를 사용하지 않고, 일반적인 웹 카메라를 사용함으로써 사용자에게 편의성을 제공할 수 있다.

본 논문의 구성은 다음과 같다. 제 2 장에서는 시선 인식을 위한 눈 영역 검출 및 기존의 시선 인식 시스템에 대한 내용을 기술하고, 제 3 장에서는 본 논문에서 제안한 SVM 기반의 시선 인식 방법 및 차영상 엔트로피를 이용한 영상 필터링 방법에 대하여 설명한다. 제 4 장에서는 제안한 시스템에 대한 실험 및 결과를 통하여 성능을 평가하고, 마지막으로 제 5 장에서 결론을 맺는다.

2. 관련 연구

2.1 Haar-like 특징과 AdaBoost 학습 알고리즘

본 논문에서는 눈 영역 검출을 위한 특징으로 Haar-like 특징을 특징 집합으로 사용하였다. Haar-like 특징은 가장 간단하고 효율적인 방식으로 계산방법이 픽셀 자체를 이용한 연산이 아니라 특징값을 가지고 각 영역안에 있는 픽셀들의 값을 더하여 합을 구한 후, 그 값들에 가중치를 곱하여 합을 계산함으로써 특징값을 구할 수 있다[9].

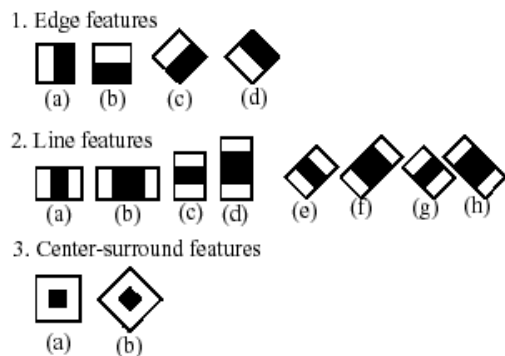
(그림 1)은 Haar-like feature의 프로토타입들로서 서로 다른 위치와 비율을 유지한 크기확장을 통해 윈도우에서 위치 할 수 있는 여러 형태로 변하기 때문에 많은 특징값을 얻을 수 있다[10, 11].

이러한 Haar-like 사각형의 픽셀 합을 빠르게 수행하기 위해서 인테그랄 이미지를 이용한다. 인테그랄 이미지는 좌표 $P(x, y)$ 의 왼쪽과 위쪽에 위치한 사각형 영역의 모든 픽셀들의 값의 합을 나타내고 식 (1)과 같다.

$$P(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

$P(x, y)$ 는 P영역의 이미지를 나타내고, $i(x', y')$ 는 원래의 입력 이미지를 나타낸다.

AdaBoost 학습 알고리즘은 실질적으로 눈 영역의 패턴을 분류할 수 있는 역할을 하는 강한 분류기(Strong classifier)의 생성을 목적으로 한다. 약한 분류기(Weak classifier)는 눈 영역 영상과 눈 영역을 포함하지 않은 배경 영상을 가장



(그림 1) Haar-like feature의 프로토타입

잘 분류하는 하나의 특징을 선택하는 역할을 하게 되며, 식 (2)와 같이 정의 할 수 있다[12, 13].

$$h_j(x) = \begin{cases} 1 & \text{if } f_j(x) < \theta_j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

식 (2)에서, $h_j(x)$ 는 약한 분류기, $f_j = w_j P(x, y)$ (w_j : 가중치)로서 특징값을 나타내고, θ_j 는 문턱치, x 는 sub-window 이다. 본 논문에서는 연산 속도를 향상시키기 위해 cascade 구조를 적용하였으며, 단계가 올라갈수록 각 Haar-like feature의 프로토타입들의 위치가 눈 영역의 세밀한 특징까지도 나타내는 장점을 가지고 있다[14].

2.2 차영상 엔트로피

엔트로피는 Shannon의 정보 이론에서 불확실성의 크기, 정보량의 기댓값을 뜻한다. 1948년 Shannon은 엔트로피를 식 (3)과 같이 정의하였다[15].

$$H = - \sum_{i=1}^m P_i \log \frac{1}{P_i} = - \sum_{i=1}^m P_i \log P_i \quad (3)$$

식 (3)에서, P_i 는 사건 i 의 발생 확률을 나타낸다.

차영상 엔트로피는 검출된 눈 영역 영상과 각각의 시선 위치별 참조 영상 또는 평균 영상과 계산된다. 차영상 엔트로피의 계산 방법은 다음과 같다. 입력된 영상으로부터 검출된 눈 영역 영상 I 와 각각의 시선 위치별 참조 영상 R 또는 평균 영상 A 와의 차영상을 구한다. 평균 영상 A 는 식 (4)와 같이 정의하고, 참조 영상 R 은 수식 (5)와 같이 정의 된다.

$$A = \frac{1}{M} \sum_{m=1}^M A_m(x, y) \quad (4)$$

식 (4)에서, A 는 평균 영상을 나타내고, M 은 평균 영상을 만들기 위해 사용된 영상의 수를 나타내며, $A_m(x, y)$ 는 m 번째 영상을 나타낸다.

$$R = T \times 0.7 + A \times 0.3 \quad (5)$$

식 (5)에서, T 는 시선 인식을 하기 전 칼리브레이션을 위해 사용자로부터 입력받은 눈 영역 영상이다. 참조 영상 R 은 실험 결과에 의해 T 에 0.7의 가중치를 부여하고, A 에 0.3의 가중치를 부여하여 생성한다. 차영상 D 는 식 (6)에 나타내었다.

$$D = I - R \quad \text{or} \quad D = I - A \quad (6)$$

차영상 D 는 입력영상과 각각의 시선 위치별 참조 영상 또는 평균 영상과의 픽셀 감산을 통해서 구해진다. 차영상 엔트로피는 식(3)을 적용하여, 식 (7)와 같이 정의하였다.

$$E = - \sum_{k=-255}^{+255} P_k \log \frac{1}{P_k} = - \sum_{k=-255}^{+255} P_k \log P_k \quad (7)$$

식 (7)에서, E 는 차영상 엔트로피 값을 나타내고, P_k 는 -255부터 +255까지의 차영상 히스토그램 빈도수의 확률을 의미하며, 식 (8)과 같이 정의된다.

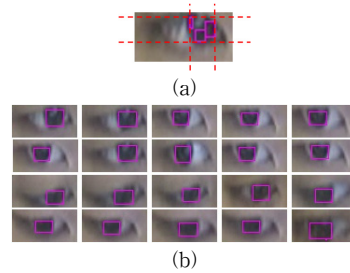
$$P_k = \frac{a_k}{G(T)} \quad (8)$$

a_k 는 k 번째 히스토그램의 빈도수를 가리키며, $G(T)$ 는 차영상에서 각 히스토그램 빈도수의 총 합을 의미한다.

2.3 눈동자 중심점 및 눈의 끝점 검출

눈동자의 중심점을 구하기 위해 눈 영역 영상을 이진화 하여 눈동자 영역을 검출한다. 이진 영상으로 변환할 때 고정된 임계값을 사용하지 않고, 영상 전체의 히스토그램을 이용하여 전체 픽셀 수의 10%가 되는 부분을 이진화 임계값으로 사용한다. 임계값을 설정하여 영상을 이진화 하였을 때 눈동자 부분이 정확하게 검출되지 않고, 여러 개의 작은 Blob으로 나뉘어 진다. 따라서 본 논문에서는 보다 정확한 눈동자 검출을 위해 (그림 2)와 같이 여러 개의 작은 Blob들을 하나로 묶어 눈동자 영역을 검출한다.

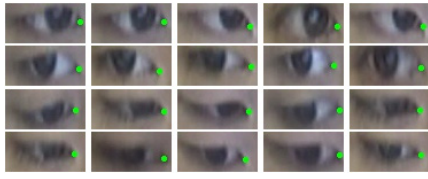
눈의 끝점을 검출하기 위해 눈 영역 영상을 두 부분으로 나누어 가버 필터를 적용한다[16]. 가버 필터는 방향(orientation)과 스케일(scale)에 따라 필터링 특성을 조절할 수 있다. (그림 3)은 가버 필터를 이용하여 눈의 끝점을 검출하는 과정을 나타내고, (그림 4)는 눈의 끝점을 검출한 예를 나타낸다.



(그림 2) (a)눈동자 영역 검출 방법과 (b)검출된 예



(그림 3) 가버 필터를 이용한 눈의 끝점 검출



(그림 4) 눈의 끝점을 검출한 예

2.4 주성분 분석(PCA)

주성분 분석(Principal Component Analysis : PCA)은 전체 학습 영상 데이터의 평균값과 표준편차, 차원 데이터들의 공분산을 이용한 2차 통계적 방법으로 데이터의 차원을 줄이기 위해 사용된다. PCA는 영상을 재구성하고 인식하는데 사용할 수 있는 고유벡터의 집합으로서, 영상 집합에 대해 통계적으로 매우 중요한 정보들을 추출할 수 있어 얼굴 인식 등에 많은 연구가 진행되고 있다[17, 18].

학습 영상들의 집합을 $X = [x_1, x_2, \dots, x_N]$ 이라 하고, 각각의 영상들로부터 평균영상벡터를 뺀 영상들의 집합을 $\Phi = [\Phi_1, \Phi_2, \dots, \Phi_N]$ 이라 할 때, 훈련 영상들의 공분산 행렬 C_T 는 식 (9)와 같다.

$$C_T = \frac{1}{N} \sum_{i=1}^N \Phi_i \Phi_i^T = \Phi \Phi^T \quad (9)$$

C_T 의 고유벡터 Ψ 와 고유값 Λ 는 식 (10)에 의해 구해진다.

$$(\Phi \Phi^T) \Psi = \Psi \Lambda \quad (10)$$

식 (10)의 양변에 Φ^T 를 곱하고, 행렬 $\Phi^T \Phi$ 의 고유값과 고유벡터를 Λ' 와 V 라 하면 식 (11)과 같이 된다.

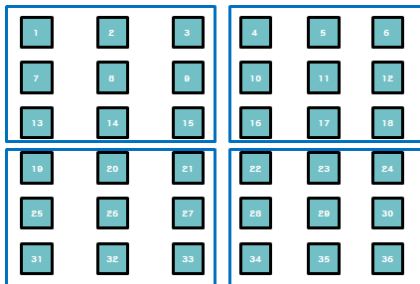
$$(\Phi^T \Phi) V = V \Lambda' \quad (11)$$

식 (11)의 양변에 Φ 를 곱하면 식 (12)와 같다.

$$\Phi (\Phi^T \Phi) V = (\Phi \Phi^T) (\Phi V) = (\Phi V) \Lambda' \quad (12)$$

고유값의 크기가 큰 순서대로 정렬하였을 때, p 개의 고유벡터로 구성된 행렬을 식 (13)에 의해 구할 수 있다.

$$W_{PCA} = [\Phi V_1, \Phi V_2, \dots, \Phi V_p] \quad (13)$$



(a) 4방향 인식

3. SVM 기반의 시선 인식 시스템

본 논문에서는 SVM 기반의 시선 인식 시스템을 제안한다. 제안한 시스템은 4방향과 9방향의 시선 위치를 인식하기 위해 데스크 탑 모니터를 가로6, 세로6, 총 36개의 시선 위치로 나누어 각각 9개, 4개씩 그룹핑하여 학습 시키고 사용자의 시선을 인식한다. 또한, 시선 인식의 성능을 높이기 위해 차영상 엔트로피를 이용한 영상 필터링 방법을 적용한다. (그림 5)는 시선 인식을 위한 인식 영역을 나타낸다.

3.1 SVM (Support Vector Machine)

SVM은 패턴인식 분야에서 이진분류 문제를 해결하는 분류기술로서, 학습 데이터를 비선형 매핑함수 $\Phi(x)$ 를 이용하여 입력 샘플 공간 X 로부터 선형 초평면이 만들어지는 고차원 특징 공간 F 로 매핑하고 인식 에러가 최소가 되게 하는 최적 초평면을 찾는다[19].

$$X: x = (x_1, \dots, x_n) \rightarrow F: \Phi(x) = (\Phi_1(x), \dots, \Phi_n(x)) \quad (14)$$

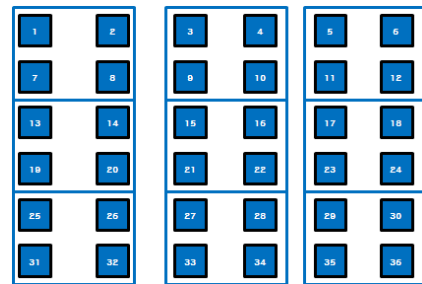
n 개의 학습데이터 중에서 i 번째 샘플 ($i = 1, \dots, n$) 인 x_i 의 클래스 레이블을 $c_i \in \{1, -1\}$ 라고 할 때, 최적 초평면에 대한 결정 함수의 식은 다음과 같다.

$$f(x) = \sum_{i=1}^n \alpha_i c_i K(x, x_i) + b, \quad K(x, x_i) = \Phi(x) \cdot \Phi(x_i) \quad (15)$$

식 (15)의 경우 α_i 는 x_i 가 초평면을 구성하는 샘플인 경우 0이 아닌값을 갖고, 그렇지 않을 경우에는 0의 값을 갖는 계수이다. 커널함수 $K(x, x_i)$ 는 비선형 매핑 함수의 내적 계산을 함수 형식으로 치환하여 계산을 간단하게 해주며, 선형(Linear) 커널, 다항식(Polynomial) 커널, 가우시안(Gaussian) 커널, 시그모이드(Sigmoid) 커널이 주로 사용된다.

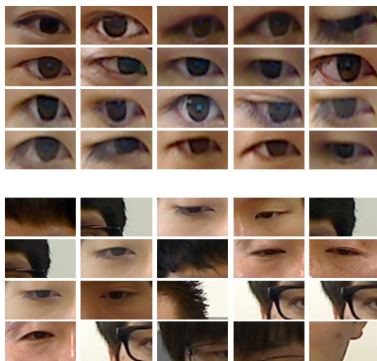
3.2 차영상 엔트로피를 이용한 영상 필터링

본 논문에서는 시선 인식 시스템의 성능을 높이기 위해 검출된 눈 영역 영상에 대해 차영상 엔트로피를 이용한 영상 필터링 방법을 적용하였다. 시선 위치별 평균 영상 또는



(b) 9방향 인식

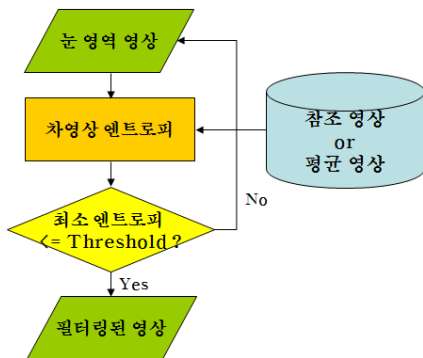
(그림 5) 시선 인식 시스템의 인식 영역



(그림 6) 마킹된 눈 영역과 검출된 눈 영역 영상 및 최소 엔트로피값



참조 영상과의 차영상 엔트로피를 구한 후, 최소 엔트로피 값을 가지는 영상이 실험을 통해 설정된 임계값보다 이하인 경우 시선 인식에 사용된다. (그림 6)은 눈 영역이 잘 검출된 영상과 눈 영역이 잘못 검출된 영상과 최소 엔트로피 값을 나타내고 (그림 7)은 차영상 엔트로피를 이용한 영상 필터링의 블록다이어그램을 나타낸다.



(그림 7) 차영상 엔트로피를 이용한 영상 필터링의 블록다이어그램

4. 실험 및 결과

본 논문에서 제안한 SVM 기반의 시선 인식 시스템의 성능을 평가하기 위하여 PC 화면을 가로 6, 세로 6 총 36개의 시선 위치로 나누어 4방향과 9방향에 대한 인식 실험을 하였다. 실험 영상은 36개의 시선 위치별 각각 10장씩 50명, 총 18000장의 영상을 사용하여 인식 성능을 평가하였다. 18000장의 영상 중 14400장의 영상을 이용하여 학습하였고, 학습에 참여하지 않은 3600장의 영상을 이용하여 인식 실험을 수행하였다. 또한, 차영상 엔트로피를 이용한 영상 필터링 방법을 적용하여 실험하였으며, 그 결과를 <표 1>, <표 2>, <표 3>, <표 4>에 나타내었다.

제안한 시스템의 성능을 평가하기 위해 기존에 제시되었던 차영상 엔트로피 기반의 시선 인식 시스템, 눈동자 중심점과 눈의 끝점 기반의 시선 인식 시스템, PCA 기반의 시선 인식 시스템을 구현하고 비교 실험을 수행하였다.

<표 1> SVM 기반의 시선 인식 실험 결과 (4방향), 3399 / 3600, 94.42%

입력 \ 출력	1	2	3	4
1	862 / 900	25 / 900	15 / 900	6 / 900
2	20 / 900	852 / 900	7 / 900	23 / 900
3	13 / 900	4 / 900	839 / 900	25 / 900
4	5 / 900	19 / 900	32 / 900	846 / 900

<표 2> 영상 필터링 적용 후, SVM 기반의 시선 인식실험 결과 (4방향), 3358 / 3521, 95.37%

입력 \ 출력	1	2	3	4
1	853 / 886	22 / 879	14 / 882	5 / 874
2	18 / 886	838 / 879	5 / 882	16 / 874
3	11 / 886	3 / 879	834 / 882	20 / 874
4	4 / 886	16 / 879	29 / 882	833 / 874

차영상 엔트로피 기반의 시선 인식 시스템에 대한 실험을 수행하기 위해 36개의 시선 위치별로 1명에 1장씩, 50장을 이용하여 평균 영상을 생성하고 실험하였고, 눈동자 중심점 및 눈의 끝점 기반의 시선 인식 시스템에 대한 실험을 수행하기 위해 모든 영상에 대해 눈동자 중심점과 눈의 끝점을 검출하여 그 좌표값을 이용하여 시선 인식 실험을 하였다. PCA 기반의 시선 인식 실험을 하기 위해 14400장을 이용하여 학습하고 나머지 3600장을 이용하여 인식 실험을 하였다.

차영상 엔트로피, 눈동자 중심점과 끝점, PCA 기반의 시선 인식 시스템 모두 영상 필터링 방법을 적용하여 실험을 수행하였다.

본 논문에서 제안한 SVM 기반의 시선 인식 시스템(방법 1)의 경우, 4방향은 94.42%, 9방향은 81.33%의 성능을 보였으며, 차영상 엔트로피를 이용한 영상 필터링 방법을 적용하였을 때, 4방향 시선 인식은 95.37%, 9방향 시선 인식은 82.25%의 성능을 보여 영상 필터링을 적용하였을 때 약 1%의 성능 향상을 보였다.

〈표 3〉 SVM 기반의 시선 인식 실험 결과 (9방향), 2928 / 3600, 81.33%

출력 \ 입력	1	2	3	4	5	6	7	8	9
1	360 / 400	26 / 400	2 / 400	22 / 400	5 / 400	6 / 400	4 / 400	4 / 400	2 / 400
2	14 / 400	325 / 400	13 / 400	16 / 400	14 / 400	10 / 400	6 / 400	4 / 400	3 / 400
3	5 / 400	18 / 400	367 / 400	3 / 400	7 / 400	17 / 400	3 / 400	5 / 400	4 / 400
4	8 / 400	9 / 400	2 / 400	323 / 400	28 / 400	11 / 400	17 / 400	11 / 400	2 / 400
5	4 / 400	7 / 400	4 / 400	24 / 400	291 / 400	19 / 400	10 / 400	15 / 400	9 / 400
6	2 / 400	6 / 400	7 / 400	3 / 400	22 / 400	308 / 400	6 / 400	12 / 400	11 / 400
7	4 / 400	2 / 400	1 / 400	6 / 400	9 / 400	4 / 400	331 / 400	29 / 400	3 / 400
8	2 / 400	4 / 400	1 / 400	3 / 400	18 / 400	7 / 400	19 / 400	282 / 400	25 / 400
9	1 / 400	3 / 400	3 / 400	0 / 400	6 / 400	18 / 400	4 / 400	38 / 400	341 / 400

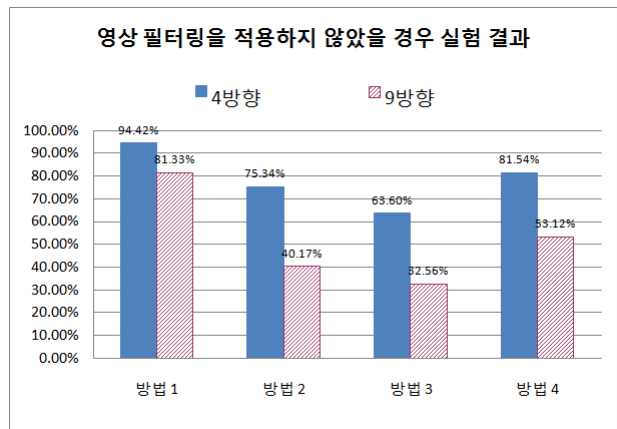
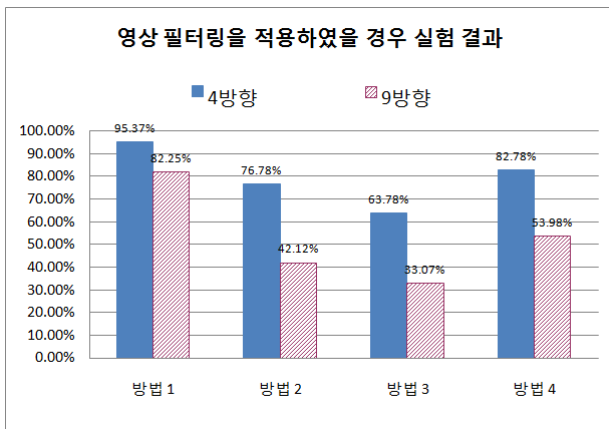
〈표 4〉 필터링 적용 후, SVM 기반의 시선 인식 실험 결과 (9방향), 2896 / 3521, 82.25%

출력 \ 입력	1	2	3	4	5	6	7	8	9
1	357 / 393	24 / 392	1 / 388	20 / 393	3 / 385	6 / 391	3 / 394	3 / 392	2 / 393
2	13 / 393	321 / 392	10 / 388	16 / 393	12 / 385	10 / 391	6 / 394	4 / 392	2 / 393
3	5 / 393	17 / 392	365 / 388	3 / 393	6 / 385	15 / 391	2 / 394	5 / 392	4 / 393
4	7 / 393	9 / 392	1 / 388	319 / 393	26 / 385	10 / 391	16 / 394	11 / 392	2 / 393
5	4 / 393	7 / 392	3 / 388	23 / 393	286 / 385	19 / 391	10 / 394	15 / 392	7 / 393
6	1 / 393	6 / 392	5 / 388	3 / 393	21 / 385	304 / 391	6 / 394	12 / 392	11 / 393
7	3 / 393	1 / 392	0 / 388	6 / 393	9 / 385	4 / 391	328 / 394	27 / 392	3 / 393
8	2 / 393	4 / 392	1 / 388	3 / 393	16 / 385	7 / 391	19 / 394	278 / 392	24 / 393
9	1 / 393	3 / 392	2 / 388	0 / 393	6 / 385	16 / 391	4 / 394	37 / 392	338 / 393

차영상 엔트로피 기반의 시선 인식 시스템(방법 2)의 경우, 4방향은 75.34%, 9방향은 40.17%의 성능을 나타내었고, 영상 필터링을 적용하였을 때, 4방향은 76.78%, 9방향은 42.12%의 성능을 보였으며, 눈동자 중심점과 눈의 끝점 기반의 시선 인식 시스템(방법 3)의 경우, 4방향은 63.60%, 9방향은 32.56%의 성능을 나타내었고, 영상 필터링을 적용하였을 때, 4방향 63.78%, 9방향 33.07%의 성능을 나타내었다. PCA 기반의 시선 인식 시스템(방법 4)의 경우, 4방향은 81.54%, 9방향은 53.12%의 성능을 보였으며, 영상 필터링을

적용하였을 때, 4방향은 82.78%, 9방향은 53.98%의 성능을 보였다. (그림 8)에 기존의 시선 인식 시스템과 SVM 기반의 시선 인식 시스템의 실험 결과를 그래프로 나타내었다.

실험 결과, 본 논문에서 제안한 SVM 기반의 시선 인식 시스템이 기존에 제시되었던 다른 시선 인식 시스템보다 더 높은 인식 성능을 나타내었을 뿐만 아니라, 다른 알고리즘 기반의 시선 인식 시스템들도 차영상 엔트로피를 이용한 영상 필터링 방법을 적용하였을 때, 인식 성능이 향상됨을 알 수 있다.



(그림 8) 기존의 시선 인식 시스템과 SVM 기반의 시선 인식 시스템의 실험 결과

〈표 5〉 시선 인식에 소요된 시간

	방법 1	방법 2	방법 3	방법 4
시간(s)	0.055	0.140	0.037	0.092

5. 결 론

본 논문에서는 SVM 기반의 시선 인식 시스템을 제안하였다. 또한, 시선 인식 성능의 향상을 위해 차영상 엔트로피를 이용한 영상 필터링 방법을 적용하였다. 제안한 시스템의 성능을 평가하기 위해 기존에 제시된 차영상 엔트로피, 눈동자의 중심점과 끝점, PCA 기반의 시선 인식 시스템을 구현하였고, 제안된 SVM 기반의 시선 인식 시스템과 비교 실험하였으며, 실험 결과 제안된 시스템이 4방향은 95.37%, 9방향은 82.25%의 성능을 보여 기존에 제시되었던 시선 인식 시스템들 보다 높은 성능을 나타내었다.

기존에 제시된 시선 인식 시스템들은 눈의 모양, 얼굴의 각도 및 기울임 정도에 따라 인식 성능이 크게 달라지는 반면, 본 논문에서 제안한 시스템은 각각 실험자의 개인별 특성에 큰 영향을 받지 않아 월등히 높은 성능을 나타내었다.

실험에 사용된 PC의 프로세서는 Core2 Duo 2.40GHz이며, <표 5>는 시선을 인식하는데 소요된 시간을 나타낸다. 제안한 시스템(방법 1)은 시선을 인식하는데 평균 약 0.055초의 시간이 소요되어, 약 0.14초가 소요된 차영상 엔트로피 기반(방법 2)의 시선 인식 시스템, 약 0.092초가 소요된 PCA 기반의 시선 인식 시스템(방법 4)보다 빠른 인식 결과 나타내어 인식 성능과 속도에서 더 높은 성능을 나타내었다. 반면, 약 0.037초가 소요된 눈동자 중심점과 끝점 기반의 시선 인식 시스템(방법 3)에 비해 다소 늦은 인식 결과를 나타내었다.

본 논문에서 제안한 시선 인식 시스템을 이용하여 기존의 입력 수단인 마우스 또는 키보드가 아닌 사용자의 시선을 이용한 객체의 선택 및 이동 등을 이용한 각종 게임에 적용될 수 있고, 음성, 제스처 등 인간의 오감을 활용한 멀티모달 휴먼인터페이스에 직접 적용될 수 있을 것으로 사료되며, 다방향 인식 및 인식 시간의 단축, 얼굴 움직임의 반영과 관련한 알고리즘 개발 및 연구가 향후 과제라 하겠다.

참 고 문 헌

[1] 김정규, 이창정, 박종승, “눈동자 검출 및 시선추적을 통한 간접 인터페이스 제어 시스템,” 한국정보기술학회논문지, 제6권, 제4호, pp.158-164, 2008.

[2] 김병주, 이진표, “웹 사용성 평가를 위한 통합평가모형 제안 및 도구 개발,” 한국디자인학회 2006 봄 학술발표대회 논문집, pp.74-75, 2006.

[3] Alexey Karpov, Andrey Ronzhin, Alexander Nechaev, Svetlana Chernakova, “MULTIMODAL SYSTEM FOR HANDS FREE PC CONTROL,” EUSIPCO2005, 2005. 9.

[4] K. Kinoshita, Y. Ma, S. Lao, M. Kawade, “A Fast and Robust 3D Head Pose and Gaze Estimation System” Proceedings of the 8th international conference on Multimodal interfaces, ACM, pp.137-138, 2006.

[5] T. E. Hutchinson, K. White, J.R. Worthy, N. Martin, C. Kelly, R. Lisa, A. Frey, “Human Computer interaction using eye-gaze input,” IEEE Transaction on systems,

[6] Tomasz Kocejko, Adam Bujnowski, Jerzy Wtorek, “Eye Mouse for Disabled,” Human System Interactions, 2008 Conference on 25-27 May 2008.

[7] Jie Xiang, Huixia Yan, Junjie Chen, “The Study of Decoding Mental State Based Eye Tracks Using SVM,” Proceedings of the IWISA 2009, pp.568-571, 21-22 November, 2009.

[8] Li Su-mei, Zhang Yan-xin, Chang Sheng-jiang and Shen Jin-yuan, “Gazing-detection of human eyes based on SVM,” Optoelectronics Letters, Vol.1, No.1, pp.65-68, 7, 2005.

[9] P.Viola and M.J. Jones, “Robust real-time object detection,” Technical Report Series, Compaq Camb ridge research Laboratory, CRL 2001/01/, Feb., 2001.

[10] 이훈진, 박현선, 김경수, 김희정, 정병희, 하명환, 김희율, “DCT/LDA를 이용한 얼굴 인식,” 대한 전자공학회 하계 학술대회 26권 제1호, pp.2024-2027, 7, 2003.

[11] 정성욱, “AdaBoost기반 실시간 얼굴 표정 인식을 위한 효과적인 사각 특징점 추출,” 한국과학기술원 석사학위 논문, 2004. 12. 14.

[12] 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.

[13] 민지홍, 이원찬, 홍기천, “Haar-like feature/HMM을 이용한 얼굴 검출 및 인증 시스템,” 2004년 한국정보과학회 가을 학술발표논문집 Vol.31, No.2, 2004.

[14] 박성훈, 이재호, 김희율, “Haar-like feature/LDA를 이용한 얼굴 인식,” 2004년 제16회 영상처리 및 이해에 관한 워크샵 발표 논문집, 2004년 1월.

[15] Shannon, C.E. “A Mathematical Theory of Communication,” The Bell Systems Technical Journal, 27, pp.379-423, 1948.

[16] Zhonglong Zheng, Jie Yang, Meng Wang, and Yonggang Wang, “A Novel Method for Eye Features Extraction,” LNCS Vol.3314, pp.1002-1007, 2005.

[17] L. Sirovich and M. Kirby, “Low dimensional Procedure for the Characterization of Human Faces,” J.Opt. Soc. Amer. 4, pp.519-524, 1987.

[18] Erik. Hjelm, “Face Detection : A Survey,” Computer Vision and Image Understanding, Vol.83, pp.236-274, 2001.

[19] C.-W. Hsu and C.-J. Lin, “A comparison of methods for multiclass support vector machines,” IEEE Trans. Neural Networks, Vol.13, No.2, pp.415-425, 2002.



이 규 범

e-mail : leo0608@skku.edu

2006년 서울보건대학 전산정보처리과 졸업
(학사)

2008년 성균관대학교 정보통신공학부(공학
석사)

2008년~현 재 성균관대학교 정보통신공학
부 박사과정

관심분야: 시선 인식, 멀티모달 인터페이스, HCI



홍 광 석

e-mail : kshong@skku.ac.kr

1985년 성균관대학교 전자공학과 졸업(학사)

1988년 성균관대학교 전자공학과(공학석사)

1992년 성균관대학교 전자공학과(공학박사)

1990년~1993년 서울보건대학 전산정보처리
과 전임강사

1993년~1995년 제주대학교 정보공학과 전임강사

1996년~현 재 성균관대학교 정보통신공학부 교수

관심분야: 오감 인식, 융합, 재현 및 HCI



김 동 주

e-mail : radioguy@skku.edu

1998년 충북대학교 전자공학과 졸업(학사)

2000년 충북대학교 전자공학과(공학석사)

2001년~현재 성균관대학교 정보통신공학부
박사과정

관심분야: 음성 및 영상 신호처리, 생체
인식, HCI