

차영상 엔트로피를 이용한 시선 인식 시스템의 구현

이 규 범[†] · 정 동 근^{††} · 홍 광 석^{†††}

요 약

본 논문에서는 차영상 엔트로피 기반의 시선 인식 시스템을 제안한다. 차영상 엔트로피는 현재 입력된 영상과 참조 영상 또는 시선의 위치별 평균 영상들로부터 차영상을 획득하고, -255부터 +255까지의 히스토그램 빈도수를 이용하여 계산한다. 차영상 엔트로피 기반의 시선 인식 방법은 2가지 방법이다. 1) 첫 번째 방법은 현재 입력된 영상과 시선 위치별 45개의 평균 영상들과의 차영상 엔트로피를 계산하여 현재 응시하고 있는 방향을 인식하고, 2) 두 번째 방법은 현재 입력된 영상과 45개의 참조 영상들과의 차영상 엔트로피를 계산하여 현재의 응시 방향을 인식한다. 참조 영상은 네 방향의 영상을 입력받아 시선 위치별 45개의 평균 영상을 이용하여 생성한다. 제안한 시스템의 성능을 평가하기 위해 PCA 기반의 시선 인식 시스템과 비교 실험을 하였고, 인식 방향은 좌상, 우상, 좌하, 우하 네 방향으로 하였으며, 45개의 참조 영상 또는 평균 영상에 대하여 인식 영역을 변경하여 실험하였다. 실험 결과 차영상 엔트로피는 97.00%, PCA는 95.50%의 인식률을 보여 차영상 엔트로피 기반의 시스템이 PCA 기반의 시스템보다 1.50% 더 높은 인식률을 나타내었다.

키워드 : 차영상 엔트로피, 시선인식, 주성분분석

An Implementation of Gaze Direction Recognition System using Difference Image Entropy

Kue-Bum Lee[†] · Dong-Keun Chung^{††} · Kwang-Seok Hong^{†††}

ABSTRACT

In this paper, we propose a Difference Image Entropy based gaze direction recognition system. The Difference Image Entropy is computed by histogram levels using the acquired difference image of current image and reference images or average images that have peak positions from -255 ~ +255 to prevent information omission. There are two methods about the Difference Image Entropy based gaze direction. 1) The first method is to compute the Difference Image Entropy between an input image and average images of 45 images in each location of gaze, and to recognize the directions of user's gaze. 2) The second method is to compute the Difference Image Entropy between an input image and each 45 reference images, and to recognize the directions of user's gaze. The reference image is created by average image of 45 images in each location of gaze after receiving images of 4 directions. In order to evaluate the performance of the proposed system, we conduct comparison experiment with PCA based gaze direction system. The directions of recognition left-top, right-top, left-bottom, right-bottom, and we make an experiment on that, as changing the part of recognition about 45 reference images or average image. The experimental result shows that the recognition rate of Difference Image Entropy is 97.00% and PCA is 95.50%, so the recognition rate of Difference Image Entropy based system is 1.50% higher than PCA based system.

Keywords : Difference Image Entropy, Gaze Direction Recognition, PCA

1. 서 론

유비쿼터스 시대의 도래와 더불어 정보 기술이 급속도로 발전하면서 PC나 PDA, 휴대폰 등과 같은 정보화기기들은 그래픽 사용자 인터페이스(Graphic User Interface : GUI) 뿐만 아니라 정보의 입·출력 방식을 다양한 매체를 이용하

여 기계와 사용자 간의 자연스러운 인터페이스에 대한 사용자들의 요구가 더욱 증가되고 있는 가운데[1], 인간의 감각 요소 중 가장 많은 정보를 수집하는 수단인 시각을 이용한 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.

기존의 시선 인식 시스템은 1) 얼굴의 움직임을 이용한 시선 인식 방법과 2) 눈동자의 움직임을 이용한 시선 인식 방법이 있다. 얼굴의 움직임을 이용한 시선 인식 방법으로 A. Ronzhin과 A. Karpov는 3개의 마커를 이용하여 영상 기

[†] 준 회 원 : 성균관대학교 정보통신공학부 박사과정
^{††} 정 회 원 : 을지대학교 의료산업학부 의료전선학 교수
^{†††} 종신회원 : 성균관대학교 정보통신공학부 교수
논문접수: 2008년 12월 10일
수정일: 1차 2009년 1월 30일
심사완료: 2009년 1월 31일

반의 얼굴 움직임을 인식하여 커서의 움직임을 제어하는 HANDS-FREE PC CONTROL 시스템을 제안하였다[2]. 제안된 시스템은 얼굴의 움직임을 인식하기 위해 별도의 마커를 사용하였으며, 미세한 시선의 변화를 감지하지 못하는 단점을 지니고 있다. 또한, 눈동자의 움직임을 이용한 시선 인식 방법으로 홍채나 동공의 기하학적 특성을 이용하여 시선을 인식하게 되며, T. E. Hutchinson 등은 동공과 각막 반사에 의한 글린트 사이의 공간적 위치 특성으로 동공을 식별하고, 동공의 위치와 형태, 왜곡 등의 정보로 시선을 실시간으로 인식하였다[3]. 이 시스템은 적외선 LED(Light Emitted Diode)를 사용하여 동공을 검출 하였으며, 동공의 중심으로부터 글린트까지의 벡터로 응시방향을 나타내었다. 또한, T. Kocejko, A. Bujnowski, J. Wtorek는 4개의 적외선 LED와 2대의 웹 카메라를 이용하여 한 대는 눈동자의 움직임을, 그리고 다른 한 대의 카메라는 사용자의 얼굴을 컴퓨터 모니터에 맞추기 위해 사용하였다[4]. 그러나 이 시스템은 웹 카메라를 머리에 장착하여 눈동자의 움직임을 인식하기 때문에 사용자에게 불편함을 초래할 수 있다는 단점이 있다.

따라서, 본 논문에서는 차영상 엔트로피(Difference Image Entropy)를 이용한 시선 인식 시스템을 제안한다. 눈동자의 시선을 인식하기 위해 Haar-like feature와 AdaBoost 학습 알고리즘을 사용하여 입력된 영상으로부터 얼굴 영역을 검출하고, 검출된 얼굴 영역에서 눈 영역을 검출하였다. 검출된 눈 영역 영상과 참조 영상 또는 시선의 위치별 평균 영상과의 차영상 엔트로피를 구한 후, 각각의 차영상 엔트로피 값을 이용하여 참조 영상 또는 평균 영상을 선택함으로써 눈동자의 시선을 인식하였다.

본 제안에서 기반 알고리즘으로 제시된 차영상 엔트로피의 특징은 목표 영상에 대한 히스토그램의 국부적 변화도에 민감하고 빨리, 에지 등의 특징을 잘 검출하는 장점이 있음에도 불구하고 기존의 시선 인식 시스템에서는 거의 사용되지 않았거나 검색 또는 엔트로피 연산자에 대한 연구 등으로 제한적인 범주 내에서 수행되어져 왔으므로 본 제안은 시선 인식을 위한 새로운 방법론을 제시할 수 있으리라 기대된다. 또한, 고가의 헤드 마운티드 장치나 적외선 LED 등과 같은 별도의 장치를 사용하지 않고, 일반적인 웹 카메라를 사용함으로써 사용자에게 편의성을 제공할 수 있다.

본 논문의 구성은 다음과 같다. 제 2 장에서는 본 연구의 관련 연구에 대해 기술하고, 제 3 장에서는 본 논문에서 제안한 차영상 엔트로피 기반의 시선 인식 방법에 대하여 설명하였다. 제 4 장에서는 제안한 시스템에 대한 실험 및 결과를 통하여 성능을 평가하였고, 마지막으로 제 5 장에서 결론을 맺는다.

2. 관련 연구

2.1 Haar-like feature와 AdaBoost 학습 알고리즘

본 논문에서는 눈 영역 검출을 위한 특징으로 Haar-like

feature를 특징 집합으로 사용하였다. Haar-like feature는 가장 간단하고 효율적인 방식으로써 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[5].

(그림 1)은 Haar-like feature의 프로토타입들로서 서로 다른 위치와 비율을 유지한 크기 확장을 통해 윈도우에서 위치 할 수 있는 여러 형태로 변하기 때문에 많은 특징값을 얻을 수 있다[6,7].

이러한 Haar-like 사각형의 픽셀 합의 계산을 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.

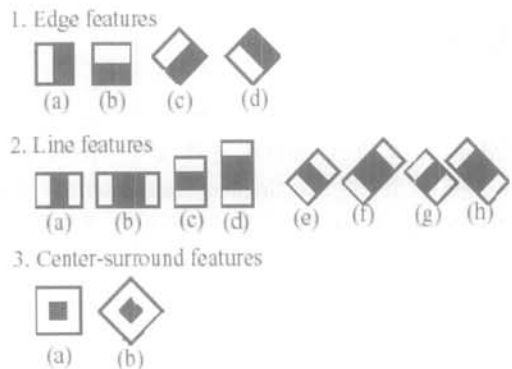
$$P(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y') \quad (1)$$

$P(x,y)$ 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 잘 분류하는 하나의 특징을 선택하는 역할을 하게 되며, 식 (2)와 같이 정의 할 수 있다[8,9].

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

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-window이다. 참고문헌 [9]의 경우, HMM을 이용한 얼굴 인증을 위해 Haar-like feature와 AdaBoost 알고리즘을 이용하여 얼굴을 검출하였으며, 본 논문에서는 차영상 엔트로피를 이용한 시선 인식을 위해 Haar-like feature와 AdaBoost 알고리즘을 이용하여 눈 영역을 검출하였다. 또한, 연산 속도를 향상시키기 위해 cascade 구조를 적용하였



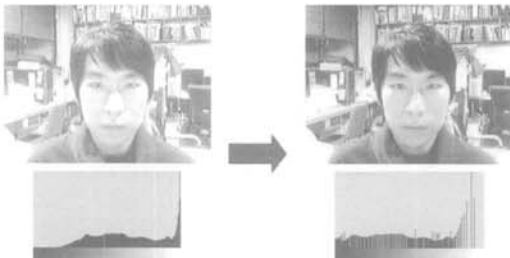
(그림 1) Haar-like feature의 프로토타입

으며, 단계가 올라갈수록 각 Haar-like feature의 프로토타입들의 위치가 눈 영역의 세밀한 특징까지도 나타내는 장점을 가지고 있다[10].

2.2 히스토그램 평활화와 ROI 선정

카메라로부터 입력된 영상에서 얼굴 영역과 눈 영역을 검출할 때, 빛이나 조명의 영향에 따른 검출 오류를 줄이기 위해 히스토그램 평활화 방법을 이용하여 전처리를 한다. 히스토그램 평활화를 이용하면 영상의 히스토그램이 펼쳐지고 영상의 대조를 높여준다. 획득한 얼굴 영상의 조명이 밝고 어두움에 따라 휘도치의 분포는 한곳에 모이는데, 이때 히스토그램 평활화를 사용하면 골고루 분포시킬 수 있다[11].

또한, 눈 영역을 검출하기 위한 시간 단축과 눈 유형이 아닌 다른 영역이 검출되는 오류를 줄이기 위해 입력 영상으로부터 검출된 얼굴 영역의 영상을 1/2로 나누어 윗부분을 ROI(Region of Interest)로 선정한다.



(그림 2) 히스토그램 평활화



(그림 3) ROI 선정 영상

2.3 주성분 분석과 k-NN

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[12,13].

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 (3)과 같다.

$$C_T = \frac{1}{N} \sum_{i=1}^N \Phi_i \Phi_i^T = \Phi \Phi^T \quad (3)$$

C_T 의 고유벡터 Ψ 와 고유값 Λ 는 식 (4)에 의해 구해진다.

$$(\Phi \Phi^T) \Psi = \Psi \Lambda \quad (4)$$

식 (4)의 양변에 Φ^T 를 곱하고, 행렬 $\Phi^T \Phi$ 의 고유값과 고유벡터를 Λ' 와 V 라 하면 식 (5)와 같이 된다.

$$(\Phi^T \Phi) V = V \Lambda' \quad (5)$$

식 (5)의 양변에 Φ 를 곱하면 식 (6)과 같다.

$$\Phi (\Phi^T \Phi) V = (\Phi \Phi^T) (\Phi V) = (\Phi V) \Lambda' \quad (6)$$

고유값의 크기가 큰 순서대로 정렬하였을 때, p 개의 고유벡터로 구성된 행렬을 식 (7)에 의해 구할 수 있다.

$$W_{PCA} = [\Phi V_1, \Phi V_2, \dots, \Phi V_p] \quad (7)$$

k-NN(k-Nearest Neighbor)은 nearest neighbor 방법 중에서 가장 일반적으로 사용되는 분류 방법으로 단 하나의 가장 가까운 이웃만으로 구하는 것이 아니고 k개의 가장 가까운 이웃 중에서 선출하여 임의의 샘플들을 분류하는 방법이다. k-NN은 분류하고자 하는 집단의 종류에 대해서는 알고 있지만, 각각의 샘플들에 대한 확률밀도 함수를 알지 못하는 경우에 사용된다. 따라서 각 샘플에 대한 확률적 파라미터들을 사전에 구하지 않고 샘플의 값 자체를 그대로 좌표에 표시하여, 가장 유사하거나 거리상으로 가장 가까운 집단으로 분류한다.

3. 차영상 엔트로피를 이용한 시선 인식 시스템

본 논문에서는 차영상 엔트로피를 이용한 시선 인식 시스템을 제안한다. 차영상 엔트로피는 현재 입력된 영상과 참조 영상 또는 시선의 위치별 평균 영상들로부터 차영상을 획득하고, -255부터 +255까지의 히스토그램 빈도수를 이용하여 계산한다. 또한, 시선 인식 시스템의 성능을 높이기 위하여 전체 차영상 엔트로피를 구하기 전, 각 히스토그램의 차영상 엔트로피에 히스토그램 거리 값을 가중치(Distance Weight)로 적용하고, k-NN의 k에 가중치를 적용한다.

3.1 차영상 엔트로피

엔트로피는 C. E. Shannon의 정보 이론에서 불확실성의 크기, 정보량의 기대치를 뜻한다. 1948년 Shannon은 엔트로피를 식 (8)과 같이 정의하였다[14].

$$H = \sum_{i=1}^m P_i \log \frac{1}{P_i} = - \sum_{i=1}^m P_i \log P_i \quad (8)$$

식 (8)에서, P_i 는 사건 i 의 발생 확률을 나타낸다.

본 논문에서는 검출된 눈 영역 영상과 각각의 시선 위치별 참조 영상 또는 평균 영상과의 차영상 엔트로피를 계산한다. (그림 4)는 차영상 엔트로피의 블록다이어그램을 나타낸다. 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 (9)와 같이 정의하고, 참조 영상 R 은 수식 (10)과 같이 정의한다.

$$A = \frac{1}{M} \sum_{m=1}^M A_m(x, y) \quad (9)$$

식 (9)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.

$$R = T \times 0.7 + A \times 0.3 \quad (10)$$

식 (10)에서, T 는 시선 인식을 하기 전 칼리브레이션을 위해 사용자로부터 입력받은 좌상, 우상, 좌하, 우하를 응시하는 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 (11)에 나타낸다.

$$D = I - R \quad \text{or} \quad D = I - A \quad (11)$$

차영상 D 는 입력영상과 각각의 시선 위치별 참조 영상 또는 평균 영상과의 픽셀 감산을 통하여 구한다. 차영상 엔트로피는 식(8)을 적용하여, 식 (12)와 같이 정의한다.

$$E = \sum_{k=-255}^{+255} P_k \log \frac{1}{P_k} = - \sum_{k=-255}^{+255} P_k \log P_k \quad (12)$$

식 (5)에서, E 는 차영상 엔트로피 값을 나타내고, P_k 는 -255부터 +255까지의 차영상 히스토그램 빈도수의 확률을 의미하며, 식 (13)과 같이 정의한다.

$$P_k = \frac{a_k}{G(T)} \quad (13)$$

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

3.2 히스토그램 거리값 가중치와 k 가중치

본 논문에서 제안한 시선 인식 시스템의 성능을 높이기 위하여 식 (14)와 같이 전체 차영상 엔트로피를 구하기 전, 각 히스토그램의 차영상 엔트로피에 히스토그램 거리 값 가중치(Distance Weight)로 적용한다.

$$E = \sum_{k=-255}^{+255} P_k \log \frac{1}{P_k} |DW| \quad (14)$$

$$= - \sum_{k=-255}^{+255} P_k \log P_k |DW|$$

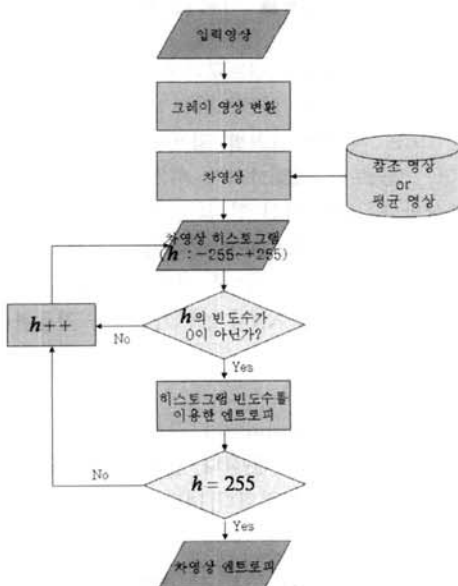
식 (14)에서, $|DW|$ 는 히스토그램 거리 값 가중치를 나타내며, $|k|$ 의 값과 같다.

또한, k-NN을 적용하여 차영상 엔트로피의 값이 작은 순서대로 k 개의 참조 영상 또는 평균 영상을 선택한 후, 그 중 차영상 엔트로피 값이 가장 작은 영상에 가장 높은 가중치를 부여하고, 차영상 엔트로피 값이 가장 큰 영상에 가장 낮은 가중치를 부여한다.

3.3 차영상 엔트로피를 이용한 시선 인식 시스템

본 논문에서 제안한 시선 인식을 위해 Haar-like feature 와 AdaBoost 검출 기법을 이용하여 입력된 영상으로부터 얼굴 영역을 검출한 후, 같은 방법을 이용하여 눈 영역을 검출한다. 카메라로부터 입력된 영상에서 빛이나 조명의 영향에 따른 검출 오류를 줄이기 위해 히스토그램 평활화 방법을 적용하고, 눈 영역을 검출하기 위한 탐색 시간을 줄이기 위해 검출된 얼굴 영상을 1/2로 나누어 윗부분을 ROI로 선정한다.

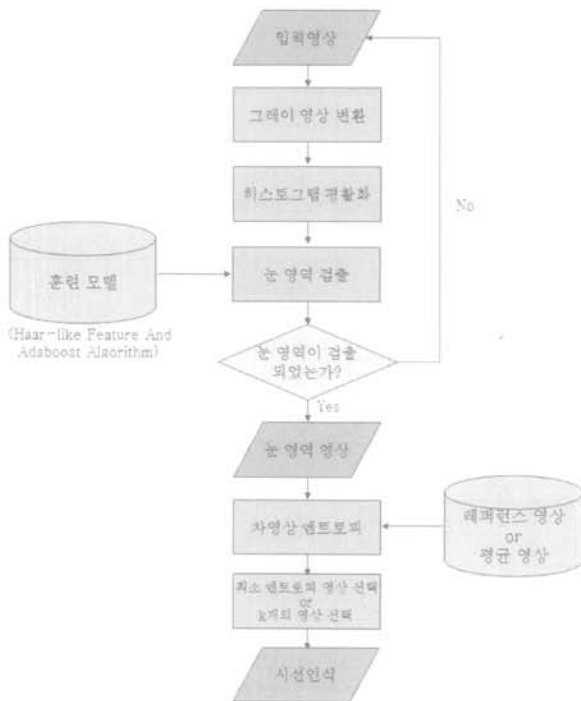
본 논문에서 제안한 차영상 엔트로피 기반의 시선 인식 방법은 두 가지 방법이다. 1) 현재 입력된 영상과 시선 위치별 45개의 평균 영상들과의 차영상 엔트로피를 구한 후 현재 응시하고 있는 방향을 인식하는 방법과, 2) 시선 인식을 하기 전 좌상, 우상, 좌하, 우하 네 방향의 영상을 입력받아 시선 위치별 45개의 평균 영상을 이용하여 참조 영상을 생성하고 입력된 영상과 45개의 참조 영상들과의 차영상 엔트로피를 구한 후 응시 방향을 인식하는 방법이다.



(그림 4) 차영상 엔트로피의 블록다이어그램

검출된 눈 영역 영상과 참조 영상 또는 평균 영상들과의 차영상 엔트로피를 구한 후, 각각의 차영상 엔트로피 값을 비교하여 최소 엔트로피 값 또는 k-NN을 이용하여 참조 영상 또는 평균영상을 선택하여 현재 응시하고 있는 방향을 인식한다. (그림 5)는 차영상 엔트로피를 이용한 시선 인식 시스템의 블록다이어그램을 나타낸다.

시선을 인식하기 위한 첫 번째 단계로, 입력된 영상을 그레이 스케일 영상으로 변환한 후, 영상 보정을 위해 히스토그램 평활화를 수행한다. 두 번째 단계로 Haar-like feature와 AdaBoost 알고리즘을 이용하여 얼굴 영역을 검출하고, 얼굴 영역 영상에 ROI를 적용하여 눈 영역을 검출한다. 세 번째 단계로 입력 영상으로부터 검출된 눈 영역 영상과 각각의 시선 위치별 참조 영상 또는 평균 영상과의 차영상 엔트로피를 계산한다. 눈동자의 시선을 인식하기 위한 마지막 단계로는, 입력 영상과 각각의 시선 위치별 참조 영상 또는 평균 영상들로부터 구해진 각각의 차영상 엔트로피 값을 비교하여 최소 엔트로피 값을 가지는 영상을 선택하거나,



(그림 5) 차영상 엔트로피를 이용한 시선 인식 시스템의 블록다이어그램

k-NN을 적용하여 k개의 참조 영상 또는 평균 영상들을 선택함으로써 눈동자가 응시하고 있는 방향을 인식한다.

4. 실험 및 결과

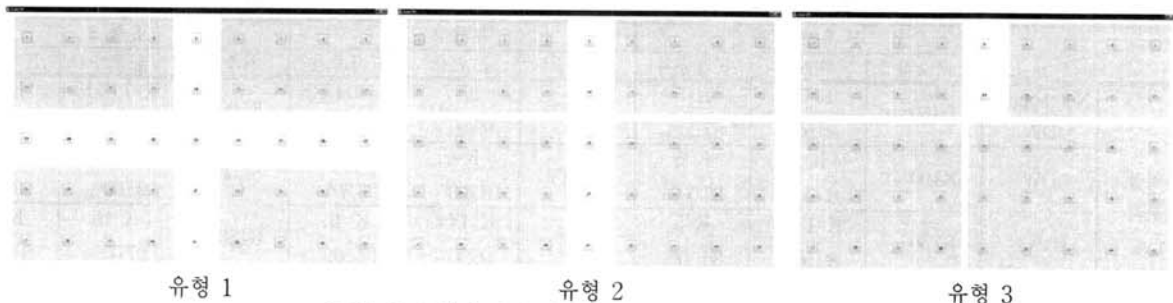
본 논문에서 제안한 차영상 엔트로피를 이용한 시선 인식 시스템의 성능을 평가하기 위하여 PC 화면을 9X5로 나누어 45개의 참조 영상 또는 평균 영상에 대한 인식 실험을 하였다. 인식 방향은 좌상, 우상, 좌하, 우하 네 방향으로 나누어 실험하였고, 실험 영상은 각 방향별로 150장씩 총 600장의 영상을 사용하여 인식 성능을 평가하였다. (그림 6)과 같이 45방향에 대한 인식 유형을 변경하여 실험하였고, k에 가중치를 적용하는 k-NN을 이용하여 실험하였다. 또한, 차영상 엔트로피를 구하기 전, 차영상의 각 히스토그램에 거리 값 가중치를 적용하여 인식 성능을 높일 수 있었고, PCA와 k-NN을 이용하여 비교 실험을 하였다.

평균 영상을 이용하여 차영상 엔트로피와 PCA를 각각 실험한 결과를 <표 1>에 나타내었으며, DW는 히스토그램 거리 값 가중치를 나타낸다.

실험 결과 차영상 엔트로피는 인식 유형 2에서 거리 값 가중치를 제공하여 적용하고 k는 11,13일 때 94.00%로 가장 높은 인식률을 보였으며, PCA는 인식 유형 3에서 k가 11,13일 때 92.50%로 가장 높은 인식률을 나타내어 차영상 엔트로피가 PCA 보다 2.00% 더 높은 인식률을 나타내었다. 또한, 45개의 평균 영상 중 3, 7, 39, 43의 4개의 영상만을 사용하여 인식 실험을 한 결과 차영상 엔트로피는 91.50%, PCA는 89.00%의 인식률을 보여 45개의 평균 영상을 사용하였을 때 더 높은 인식률을 보였다.

<표 2>는 참조 영상을 이용하여 차영상 엔트로피와 PCA를 각각 실험한 결과를 나타낸다. 실험 결과 차영상 엔트로피는 인식 유형 2에서 거리 값 가중치를 제공하여 적용하고 k는 13일 때 97.00%로 가장 높은 인식률을 보였으며, PCA는 인식 유형 3에서 k가 15일 때 95.50%로 가장 높은 인식률을 나타내어 차영상 엔트로피가 PCA 보다 1.50% 더 높은 인식률을 나타내었다. 또한, 45개의 참조 영상 중 3, 7, 39, 43의 4개의 영상만을 사용하여 인식 실험을 한 결과 차영상 엔트로피는 93.00%, PCA는 91.50%의 인식률을 보여 45개의 참조 영상을 사용하였을 때 더 높은 인식률을 보였다.

<표 3>과 <표 4>는 본 논문에서 제안한 차영상 엔트로



(그림 6) 45개의 시선 위치를 이용한 4방향 인식 유형

피와 PCA를 결합하여 실험한 결과를 나타낸다.

실험 결과 차영상 엔트로피만을 사용하였을 경우와 비슷한 인식률을 보였고, PCA만을 사용하였을 경우보다 더 좋은 인식률을 보였다.

차영상 엔트로피와 PCA를 각각 실험한 결과와 차영상

엔트로피와 PCA를 결합하여 실험한 결과 모두 칼리브레이션을 통하여 참조 영상을 생성하고, 이를 이용한 시선 인식 시스템의 성능이 평균 영상을 이용한 시스템보다 더 높은 성능을 나타내었다.

〈표 1〉 평균 영상을 이용하여 차영상 엔트로피와 PCA를 각각 실험한 결과

		유형 1			유형 2			유형 3		
		최소값	분포수	K가중치	최소값	분포수	K가중치	최소값	분포수	K가중치
차영상 엔트로피	Not applied DW	83.50%	K=11,13 88.50%	K=7 90.00%	84.50%	K=11,15 90.50%	K=15 90.50%	82.00%	K=13,15 91.00%	K=13 89.50%
	DW	84.00%	K=11,15 87.00%	K=10 87.50%	85.00%	K=13,15 92.50%	K=15 91.00%	80.50%	K=11,13 92.50%	K=15 89.00%
	(DW) ²	85.50%	K=13,15 89.50%	K=9,13 88.50%	84.50%	K=11,13 94.00%	K=13,15 91.50%	81.50%	K=11 93.00%	K=15 88.50%
	\sqrt{DW}	83.50%	K=13 88.00%	K=9,13 88.00%	83.00%	K=15 91.50%	K=11,15 90.00%	79.50%	K=13,15 91.00%	K=15 88.50%
PCA		79.50%	K=9 82.00%	K=11 81.00%	83.50%	K=15 90.50%	K=13,15 90.00%	83.00%	K=11,13 92.00%	K=13 91.00%

〈표 2〉 참조 영상을 이용하여 차영상 엔트로피와 PCA를 각각 실험한 결과

		유형 1			유형 2			유형 3		
		최소값	분포수	K가중치	최소값	분포수	K가중치	최소값	분포수	K가중치
차영상 엔트로피	Not applied DW	87.50%	K=13 90.50%	K=7 90.00%	86.50%	K=13,15 94.00%	K=15 94.50%	84.00%	K=13,15 94.00%	K=13 91.50%
	DW	85.00%	K=11,13,15 88.00%	K=9 89.50%	86.00%	K=15 94.50%	K=15 95.00%	82.50%	K=15 94.50%	K=15 91.00%
	(DW) ²	87.00%	K=15 90.00%	K=7,15 89.00%	87.00%	K=13 97.00%	K=15 96.00%	84.50%	K=11 96.50%	K=15 91.50%
	\sqrt{DW}	84.50%	K=11 88.50%	K=7,15 89.00%	85.00%	K=15 95.50%	K=15 96.00%	82.50%	K=15 95.50%	K=15 91.50%
PCA		81.00%	K=7 84.00%	K=9,11 84.00%	86.00%	K=15 93.00%	K=15 92.00%	86.00%	K=15 95.50%	K=13 93.00%

〈표 3〉 평균 영상을 이용하여 차영상 엔트로피와 PCA를 결합한 실험 결과

		유형 1			유형 2			유형 3		
		최소값	분포수	K가중치	최소값	분포수	K가중치	최소값	분포수	K가중치
차영상 엔트로피 + PCA	Not applied DW	83.50%	K=7,11 84.50%	K=5,7 85.50%	88.00%	K=7,10 92.00%	K=10,13 91.00%	81.00%	K=7,15 91.50%	K=11,13 90.00%
	DW	84.50%	K=7,10 85.50%	K=5 86.00%	87.50%	K=9,11 92.00%	K=15 91.50%	80.50%	K=13 92.00%	K=10,15 91.50%
	(DW) ²	84.50%	K=11,13 82.50%	K=5 84.50%	89.50%	K=11,13 94.50%	K=15 92.00%	82.50%	K=13,15 94.00%	K=13 91.50%
	\sqrt{DW}	83.00%	K=9 83.00%	K=7 86.50%	87.00%	K=13,15 93.00%	K=15 89.50%	80.00%	K=15 92.00%	K=13,15 89.00%

〈표 4〉 참조 영상을 이용하여 차영상 엔트로피와 PCA를 결합한 실험 결과

		유형 1			유형 2			유형 3		
		최소값	분포수	K가중치	최소값	분포수	K가중치	최소값	분포수	K가중치
차영상 엔트로피 + PCA	Not applied DW	89.50%	K=5,7 86.50%	K=5 87.50%	90.00%	K=7 95.00%	K=10,13,15 93.00%	84.00%	K=7,13,15 94.50%	K=13 93.00%
	DW	88.00%	K=7 86.50%	K=5 87.00%	89.00%	K=7 95.00%	K=15 93.50%	82.50%	K=13 95.00%	K=10 93.00%
	(DW) ²	87.50%	K=13,15 86.50%	K=5 86.50%	91.00%	K=13,15 96.50%	K=15 94.50%	84.50%	K=15 97.00%	K=13 94.00%
	\sqrt{DW}	88.50%	K=7 87.00%	K=5 87.50%	89.00%	K=15 95.50%	K=15 92.50%	82.50%	K=15 96.00%	K=13 93.00%

5. 결 론

본 논문에서는 차영상 엔트로피를 이용한 시선 인식 시스템을 제안하였다. 제안한 시스템의 성능을 평가하기 위해 기존에 제시된 PCA를 이용한 시선 인식 시스템을 구현하였고, 제안된 차영상 엔트로피 기반의 시선 인식 시스템과 비교 실험하였으며, 평균 영상을 이용한 시선 인식 방법과 참조 영상을 이용한 시선 인식 방법을 이용하여 실험하였다. 그 결과 참조 영상을 사용하였을 경우 차영상 엔트로피 기반의 시스템이 97.00%의 성능을 보여 95.50%의 성능을 보인 PCA 기반의 시선 인식 시스템보다 1.50% 더 높은 인식 성능을 보였다.

제안한 시스템은 응시 방향을 인식하는데 평균 약 0.14초의 시간이 소요된 PCA 보다 많은 평균 약 0.25초의 시간이 소요되었고, 이를 보완하기 위해 차영상 엔트로피와 PCA를 결합하여 실험한 결과 차영상 엔트로피만을 사용하였을 때와 비슷한 성능을 나타내어 차영상 엔트로피만을 사용하는 것이 가장 우수한 성능을 나타내었으며, 칼리브레이션 과정을 통하여 참조 영상을 생성하고, 이를 이용한 시선 인식 시스템이 평균 영상을 이용한 시선 인식 시스템보다 더 높은 인식 성능을 나타냄을 알 수 있었다.

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 단축과 관련한 알고리즘 개발 및 연구가 향후 과제라 하겠다.

감사의 글

본 연구는 2008년도 중소기업기술혁신개발사업 (S1053744) 및 2008년도 정보(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임. (R01-2008-000-10642-0)

참 고 문 헌

[1] 이현주, 석민수, "멀티미디어 환경에서 멀티모달 휴먼인터페이스 기술", 전자공학회지, 제22권 제4호, pp.407-417, 1995.
 [2] A. Karpov, A. Ronzhin, A. Nechaev and S. Chernakova, "Multimodal System for Hands Free PC Control," EUSIPCO 2005, 2005. <http://signal.ee.bilkent.edu.tr/defevent/papers/cr1105.pdf>
 [3] T. E. Hutchinson and K. White, J.R. Worthy, N. Martin, C. Kelly, R. Lisa, A. Frey, "Human Computer Interaction using Eye-Gaze Input," IEEE Transaction on systems, man, and cybernetics, Vol.19, pp.1527-1533, 1989.
 [4] T. Kocejko, A. Bujnowski, and J. Wtorek, "Eye Mouse for Disabled," Human System Interactions 2008, pp.199-

202, 2008.
 [5] P.Viola and M.J. Jones, "Robust real-time object detection," Technical Report Series, Compaq Cambridge research Laboratory, 2001.
 [6] 이훈진, 박현선, 김경수, 김희정, 정병희, 하명환, 김희율, "DCT/LDA를 이용한 얼굴 인식", 대한 전자공학회 하계 학술 대회 제26권 제1호, pp.2024-2027, 2003.
 [7] 정성욱, "AdaBoost기반 실시간 얼굴 표정 인식을 위한 효과적인 사각 특징점 추출", 한국과학기술원 석사학위 논문, 2004.
 [8] A. Kuranov, R. Lienhart and Vadim Pisarevsky, "An Empirical Analysis of Boosting Algorithms for Rapid Objects With an Extended Set of Haar-like Features," Intel Technical Report, 2002.
 [9] 민지홍, 이원찬, 홍기천, "Haar-like feature/HMM을 이용한 얼굴 검출 및 인증 시스템", 2004년 한국정보과학회 가을 학술발표논문집 제31권 제2호, pp.739-741, 2004.
 [10] 박성훈, 이재호, 김희율, "Haar-like feature/LDA를 이용한 얼굴 인식", 제16회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp.271-276, 2004.
 [11] 전종배, 하길람, 김동주, 홍광석, "이동 환경에서 히스토그램 평활화와 ROI를 이용한 치열영역 검출", 한국신호처리·시스템학회 2007 추계 학술 대회 논문집, pp.211-214, 2007.
 [12] L. Sirovich and M. Kirby, "Low dimensional Procedure for the Characterization of Human Faces," J.Opt. Soc. Amer. Vol.4, No.3, pp.519-524, 1987.
 [13] E. Hjelmas, "Face Detection : A Survey," Computer Vision and Image Understanding, Vol. 83, pp.236-274, 2001.
 [14] C. E. Shannon, "A Mathematical Theory of Communication," The Bell Systems Technical Journal, Vol.27, pp.379-423, 1948.



이 규 범

e-mail : leo0608@skku.edu
 2006년 서울보건대학 전산정보처리과(학사)
 2008년 성균관대학교 정보통신공학부(공학 석사)
 2008년~현 재 성균관대학교 정보통신공학부 박사과정

관심분야 : 시선 인식, 멀티모달 인터페이스, HCI



정 동 근

e-mail : tchung@eulji.ac.kr

1986년 성균관대학교 전자공학과(학사)
1988년 성균관대학교 전자공학과(공학석사)
1996년 성균관대학교 전자공학과(공학박사)
1990~현 재 을지대학교 의료산업학부
의료전산학 교수

관심분야 : 유비쿼터스 응용, 임베디드 시스템, 신호 및 제어



홍 광 석

e-mail : kshong@skku.ac.kr

1985년 성균관대학교 전자공학과(학사)
1988년 성균관대학교 전자공학과(공학석사)
1992년 성균관대학교 전자공학과(공학박사)
1990년~1993년 서울보건대학 전산정보처
리과 전임강사

1993년~1995년 제주대학교 정보공학과 전임강사
1996년~현 재 성균관대학교 정보통신공학부 교수
관심분야 : 오감 인식, 융합, 재현 및 HCI