
Kalman-Filer를 이용한 효과적인 실시간 시선검출

정유선* · 홍성수**

A Study on real time Gaze Discrimination Using Kalman Filter

You-sun Jeong* · Sung-soo Hong**

요 약

본 논문에서는 기존의 문제점인 얼굴 움직임이 있을 시 시선 식별이 어려운 점과 사용자에 따른 교정작업이 필요하다는 점을 해결하고자 새로운 시선 식별 시스템을 제안한다. Kalman필터를 사용하여 현재 머리의 위치정보를 이용하여 미래위치를 추정하였고 얼굴의 진위 여부를 판단하기 위해서 얼굴의 특징요소를 구조적 정보와 비교적 처리시간이 빠른 수평, 수직 히스토그램 분석법을 이용하여 얼굴의 요소를 검출한다. 그리고 적외선 조명기를 구성하여 밝은 동공효과를 얻어 동공을 실시간으로 검출, 추적하였고 동공-글린트 벡터를 추출한다.

ABSTRACT

In this paper, the movement faces the problem of the difficult points upon the gaze of the user that corrective action is needed to solve the identification system offers a new perspective. Using the Kalman filter using the position information of the current head position estimated the future.

In order to determine the authenticity of the face features of the face structural element information and the processing time is relatively fast horizontal and vertical histogram analysis method to detect the elements of the face. and people grow and infrared bright pupil effect obtained by constructing a real-time pupil detection, tracking and pupil - geulrinteu vectors are extracted.

키워드

Kalman, 베이지안, 신경망

1. 서론

기존의 시선 식별 시스템은 시선을 식별하기 위해 얼굴의 움직임과 눈동자의 움직임을 이용한다. 얼굴의 움직임만을 이용할 경우 얼굴의 위치에 따라 시선의 위치를 결정하게 되는데 미세한 시선 변화를 감지하지 못하는 단점을 지니고 있다. 현재까지 연구된 눈동자에 근거한 시선 추정 방식의 가장 큰 공통점은 각각의 글린트와 동공 사이의 상대적인 위치에 근거한

다는 것이다.[1-3] 눈과 시선 식별 시스템의 또 다른 문제는 각각의 사용자에 대하여 교정 작업이 필요하다는 점이다. 이러한 제한 사항을 극복하기 위하여, 안구의 모형[4]에 의한 새로운 시선 추적 기술을 제안하고, 2개의 저점만 교정이 필요하도록 하여 시선 교정 절차를 단순화 하였으나, 상대적으로 움직임이 적은 경우에 원활하게 동작하였고, 각각의 사용자에 대해 정확한 기하학적 안구 모형을 제시한다. 또한, 2대의 카메라와 눈과 눈 영상의 기하학적 특징[5]을 이용

* 조선대학교 컴퓨터공학과(ys1128f@naver.com)
접수일자 : 2010. 06. 10

** 교신저자 : 동강대학 컴퓨터정보과(hong34@hanmail.net)
심사(수정)일자 : 2010. 07. 01 게재확정일자 : 2010. 08. 05

하여 시선 교정 절차를 완전히 배제한다. 본 논문에서는 동공의 검출 후 객체 추적 알고리즘인 Kalman 필터 알고리즘에 의하여 연속되는 프레임에서 동공을 추적하였다. 1차 식별된 결과를 바르게 식별하였는지 검증하고자 인접 영역만을 가지고 재식별하도록 하여 실시간 시선 식별 인식률을 향상시켰다.

II. 베이지안 네트워크를 이용한 특징 추적

2.1. 베이지안 네트워크 이론

베이지안 네트워크는 하나의 데이터 집합으로부터 학습할 때 각 노드는 데이터 집합의 각 특징을 표현하고 각 가지(arc)는 특징들 간의 의존성을 표현하게 되며 이렇게 학습된 베이지안 네트워크를 기초로 목표 노드를 확률적으로 예측[5-7]한다. 각 노드의 조건 확률과 노드 간의 연결선들로 이루어진 베이지안 네트워크 구조가 주어지며, 각 노드의 확률을 데이터로부터 학습하게 되는데 식 (1)과 같이 베이지안 네트워크를 수식으로 표현할 수 있다.

$$p(x) = \prod_{i=1}^n p(x_i | \pi_i) \dots\dots\dots (1)$$

변수 A, B, C, D, E, F가 있고 각 변수들이 yes와 no 값을 갖는다고 할 때, 표 1-1의 조건확률표의 각 열은 이 두 변수 A와 B에 배정 가능한 값들을 나타내고, 각 행은 변수 D가 yes 또는 no를 가질 조건확률을 나타낸다.

표 1. 조건확률표
Table 1. Conditional Probability Table

	A(y), B(y)	A(y), ~B(n)	~A(n), B(y)	~A(n), ~B(n)
D(y)	0.4	0.1	0.8	0.2
~D(y)	0.6	0.9	0.2	0.8

베이지안 네트워크는 조건확률 계산식 식(2)와 같은 베이즈 이론을 이용한다.

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \dots\dots\dots (2)$$

주어진 학습 데이터 집합을 D라고 할 때, P(h|D)는 h의 사후 확률이고, 이를 구하기 위해서는 P(D|h), P(h), P(D)를 먼저 구해야 한다. 또한 최대 사후 확률은 식 (3)과 같다.

$$h_{MAP} \equiv \arg \max_{h \in H} P(D|h)P(h) \dots\dots\dots (3)$$

사전 가설의 확률이 없을 때 모두 같다고 가정하고 P(h) 값을 생략 한 후 최대 유사 확률은 식 (4)과 같이 표현한다.

$$h_{ML} \equiv \arg \max_{h \in H} P(D|h) \dots\dots\dots (4)$$

2.2. 얼굴, 눈동자 이용한 시선 식별법

동공과 홍채사이의 계조도 차이가 아주 작기 때문에 정확한 추출이 어렵고 정확도가 낮은 단점이 있다 [8,9]. 그러므로 적외선과 같은 특수 조명을 이용하여 동공의 움직임 추적함으로써 추출이 어려운 단점을 극복한다. 얼굴의 삼차원 움직임량을 구하는 연구는 이차원 카메라 영상에 투영된 얼굴 특징점의 움직임으로부터 얼굴의 삼차원 회전량 및 이동량을 추정하는 방법들을 사용한다.

2.3. 얼굴 특징 추적

Sirovich 와 Kirby[10]는 수염이 없는 얼굴영상에 대해 KL변환을 적용하여 얼굴을 표현하는 방법을 소개하였다. Hallinan[11,12]도 영상에서 눈을 검출하기 위해 템플릿 매칭을 사용하였다. 본 논문에서는 머리 영역 검출 후 적절한 상태가 판단되면 추출된 후보 영역이 얼굴인지의 여부를 판단하게 된다. 얼굴의 진위 여부를 판단하기 위해서 얼굴의 특징 요소인 양쪽 눈, 양쪽 콧구멍 및 입의 양끝을 사용하여, 구조적 정보와 비교적 처리 시간이 빠른 수평, 수직 히스토그램 분석법을 이용하여 얼굴의 요소를 검출하였다. 눈이 존재할 경우는 수평축과 수직축에서 히스토그램의 피크 위치가 나타나고 이 축이 만나는 두 점이 양 눈의 위치가 된다. 눈 위치 추출 시 안경 사용자의 경우, 안경 렌즈의 반사에 의해 눈동자 후화소가 많이 소실

될 수 있으므로, 눈 영역의 이진화시 p-tile방법을 사용하여 흑화소의 개수를 일정하게 유지함으로써 흑화소 소실 문제를 해결하였다. 콧구멍 역시 눈동자와 같은 방법으로 추출한다. 또한, 얼굴의 구조적 정보를 이용하여 입의 위치에 대한 존재 가능 범위를 설정한 후 이 영역에 대해 수직, 수평 히스토그램을 구한다. 입의 특성상 수직 방향 히스토그램은 일정 부분에만 존재하게 된다. 임계치 이상으로 갑작스럽게 변하는 부분을 입의 양 끝점 위치가 된다. 머리카락 영역은 얼굴 영역에 비해 이웃 화소간 밝기 성분이 크게 변하는 특징을 갖고 있고 식 (5)과 같은 분산 $V(x,y)$ 으로 표현 할 수 있다. 이 값이 정해진 문턱치 이상이면 머리카락 영역으로 표시한다.

$$V(x,y) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 |Y(x,y) - Y(x+i,y+j)| \dots (5)$$

머리카락 영역을 나타내는 영상 $f_H(x,y)$ 는 머리카락 영역은 255이고, 나머지는 0으로 표현된다. 기존 얼굴 방법으로 얻어진 $f_F(x,y)$ 과 머리카락 영상 $f_H(x,y)$ 을 식 (6)의 논리적 관계식으로 머리카락의 영향을 받지 않는 실질적인 얼굴 영역 영상 $f_{RF}(x,y)$ 을 얻을 수 있다.

$$f_{RF}(x,y) = f_F(x,y) \cap \overline{f_H(x,y)} \dots (6)$$

2.4. 얼굴 동작 인식을 위한 인공 신경망

본 연구는 지도학습법의 GRNN알고리즘으로 얼굴 인식을 연구한다. GRNN(generalized regression neural network)은 Specht에 의해 처음 사용된 것으로 관측된 데이터의 확률 밀도 함수에 기본[13]을 두고 있다. GRNN의 구성은 4개의 층으로 구성된다. 입력층(input layer)은 입력패턴을 각각의 숨김층(hidden layer)의 뉴런들에 배분하는 역할을 하며 두 번째 층으로 모두 연결된다. 신경망 중 가장 널리 쓰이는 것은 BP(back propagation)이므로 축색동기(axon)에 해당하는 뉴런(neuron)으로 연결되어진다. 뉴런에 입력된 입력신호들은 각각의 입력 신호 값과 관련된 가중치들에 의해 곱해진 후 합해지게 되고, 이 합은 비선형적인 활성화 함수에 넣어 출력값으로 나타나게 된

다. 이때 BP에서는 식(7)과 같은 시그모이드 (sigmoid)함수를 사용한다.

$$f(x) = \frac{1}{(1+e^{-x})} \dots (7)$$

계산된 출력값은 뉴런의 위치에 따라서 다음 층의 입력 데이터나 최종 출력값으로 사용된다. 인공신경망은 학습 데이터들을 사용하여 학습을 시키는데 인공신경망에서 계산된 값과 실제로 요구되는 측정값과의 오차의 자승으로부터 각 연결 가중치를 변화시킨다. 네트워크 진행 방향이 어떻게 될 것인가에 대한 이론적인 예측이 매우 힘들고, 초기 가중치와 숨김층에 있는 뉴런의 개수에 따라 많은 수의 가중치를 구해야할 경우에는 적절한 결과를 얻기 힘들다. 각각 숨김층 뉴런들이 가지고 있는 학습 데이터로부터 입력층 원소들을 빼어 제곱을 하거나 절대값으로 나타낸다. j번째 숨김층의 입력 함수는 식 (8)과 같다.

$$I_j = \sum_{i=1}^n |W_{ij} - X_{ij}| \text{ 또는 } I_j = \sum_{i=1}^n (W_{ij} - X_{ij})^2 \dots (8)$$

x_i 는 입력신호이고 w_{ij} 는 i 번째 입력층과 j 번째 숨김층 뉴런 사이의 가중치이며 n 은 입력층의 뉴런 수를 나타낸다. 이렇게 계산된 입력값은 식 (9)과 같은 지수함수의 형태를 가지는 비선형 활성화함수로 넘겨지게 된다.

$$f(I_j) = \exp\left(\frac{-I_j}{2\sigma^2}\right) \dots (9)$$

활성 함수에서 계산된 숨김층의 출력값은 합산층 (summation layer)으로 전달되며, 합산층에서는 식 (10)과 같은 적분을 실행하여 결과값을 계산하여 출력층(output layer)으로 보낸다.

$$Y(X) = \frac{\sum_{i=1}^n Y f(I_i)}{\sum_{i=1}^n f(I_i)} \dots (10)$$

Y'는 각 숨김층 뉴런이 가지는 출력값으로써 학습 데이터의 결과값에 해당한다. 위의 과정들을 정리하면

다음과 같다. GRNN은 각각의 학습 단계에서 네트워크에 입력하여 새로운 입력값에 대한 반응을 계산하는데 사용한다. 그리고 가우시안 커널(Gaussian Kernel)함수들을 각 학습 단계에 위치시킨다. 입력값으로 계산된 결과값은 학습 단계의 결과값 평균에서 가중치를 두어 계산되며, 가중치는 계산값과 학습 단계와 관련이 있다.역전파에 비해 GRNN이 가지는 큰 장점은 학습 데이터가 비교적 적게 요구된다는 점이다. 역전파와 비슷한 정확도를 얻기 위해 약 1%의 학습 데이터가 필요[13]하며, 학습 데이터가 많은 경우에는 비슷한 데이터들을 군집화시켜 사용할 수가 있다.

III. Kalman 필터 알고리즘 적용

3.1 Kalman 필터와 평균 이동 알고리즘을 이용한 동공추적

Kalman 필터를 기반으로 한 밝은 동공을 이용한 눈 추적기기를 이용하여 초기 프레임에서 동공의 위치를 확인하고 Kalman 필터에 의하여 밝은 동공을 추적하였다. 각각의 프레임에서의 동공의 움직임은 동공의 위치한 변화속도로 특성화 되어진다. 시점 t 에서 동공 중심의 화소의 위치를 (c_t, r_t) 라 하고 (u_t, v_t) 를 시점 t 에서 c 와 r 방향으로의 변화속도라 하면, 시점 t 에서 상태벡터를 $x_t = (c_t, r_t, u_t, v_t)t$ 로 나타낼 수 있다. 그리고 시스템을 식(11)과 같이 모델링 할 수 있다.

$$X_{t+1} = \Phi X_t + W_t \dots\dots\dots (11)$$

W_t 는 시스템의 불안정성을 나타낸다. 시점 t 에서 동공의 위치를 $Z_t = (\hat{c}_t, \hat{r}_t)$ 로 간주하면, Kalman 필터에 의한 측정모델을 식(12)와 같이 구할 수 있다.

$$z_t = Hx_t + v_t \dots\dots\dots (12)$$

식 (12)에서 v_t 는 불확정된 측정치이다. 밝은 동공 효과가 존재할 만한 예상 위치를 구하고 인접한 화소에서 문턱치값을 이용하여 추정한다. 상태 벡터 x_{t+1} 와 공분산 행렬 Σ_{t+1} 의 초기치를 얻고 예측을

위한 시스템 모델과 측정 모델을 갱신하게 된다.

Kalman필터로 동공 추적 시 눈이 감기거나 폐색에 의해 밝은 동공이 사라질 경우 Kalman필터에 의한 추적이 불가능해지는 단점이 있으므로, 평균 이동 알고리즘을 이용하여 눈 추적의 정확성을 높였다.

평균 이동 추적 알고리즘은 바타라치야 계수로 두 개의 명암 분포의 유사성을 측정하여 유사성에 따라 평균이동을 반복하여 모델과 가장 유사한 후보영역을 찾는다. Y지점에서 표적 밀도 q 를 위한 바타라치야 계수를 추정하고 표적 후보 밀도를 식(13)과 같이 구한다.

$$\hat{p}(y) \equiv p[\hat{p}(y), \hat{q}] - \sum_{u=1}^m \sqrt{p_u \hat{q}_u} \dots\dots\dots (13)$$

두 개의 분포에 대한 거리는 식(14)와 같다.

$$d(y) = \sqrt{1 - p[\hat{p}(y), \hat{q}]} \dots\dots\dots (14)$$

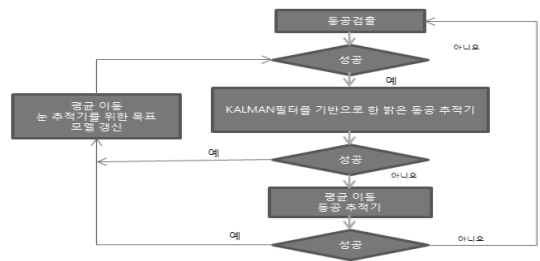


그림 1. 동공 추적 알고리즘
Fig. 1 Pupil tracking algorithm

눈과 눈이 아닌 영역에 대한 명암 분포에 대한 특성은 짝수와 홀수 필드 두 영상에 의해 확실히 특성지어진다. 어두운 동공과 밝은 동공 영상에 대응하는 각기 다른 특징의 2개의 확률 분포를 구하여 2차원 결합 히스토그램을 유도할 수 있다.

표 2. 동공 추적 비교

Table 2. Tracking statics Comparison for both pupil trackers

300프레임 영상		kalman필터 추적기		제한된 합성 추적기	
		인식된 프레임수	인식률	인식된 프레임수	인식률
왼쪽눈	(만는)2280프레임	2012	89%	2255	99%

	(감은눈)330프레임	0	0%	317	95%
	(감우어전눈)410프레임	0	0%	384	94%
오른쪽눈	(뜬눈)2125프레임	1945	92%	2117	99%
	(감은눈)330프레임	0	0%	307	93%
	(감우어전눈)545프레임	0	0%	499	92%
전체인식률		3357	66%	5879	98%

3.2 얼굴 식별 전처리

본 논문에서는 동공영상을 획득하기 위하여 Hutchinson의 방법에 근거하여 880nm파장의 적외선 LED를 두개의 고리 모양으로 구성된 능동적 적외선 조명기를 제작하여 40nm파장에서 32mW의 강도로 적외선을 조사하였다. 안경을 쓴 사용자의 경우 빛의 반사에 따른 오인식이 있을 수 있으므로 안경을 착용하지 않은 경우로 가정하였고 카메라와 사용자의 거리는 사용자가 컴퓨터 앞에 자연스럽게 앉아 있을 경우인 0.5~1m로 하였다. 광원을 교차 조사함에 따라 안쪽 고리의 적외선 LED들이 켜졌을 때는 동공에서 반사된 적외선이 검출되어 밝은 동공영상이 생성되었고, 바깥 쪽 고리의 적외선 LED들이 켜졌을 때는 동공에 반사된 적외선이 검출되지 않으므로 어두운 동공이 생성되었다. 짝수 필드 영상과 홀수 필드 영상을 만들어 같은 배경과 외부 조명을 사용하지만 짝수 필드의 동공이 홀수 필드보다 밝게 보였고, 배경과 외부 조명의 간섭을 제거하고자 짝수 필드 영상에서 홀수 필드 영상을 제거하여 차 영상을 만들었다. 차 영상에서는 대부분의 배경과 외부 조명의 간섭효과가 제거되었다.

3.3 매개변수 추출

시선 식별을 위하여 1차적인 매개변수만을 이용할 경우 움직임이 있는 경우에 대한 시선 식별이 잘 이루어지지 않는 경향이 있기에 본 논문에서는 얼굴 움직임에 대비한 얼굴 방향에 대한 매개변수를 추출하고자 하였다. 얼굴방향에 대한 동공 특성은 다음과 같다.· 얼굴이 전면으로부터 측면으로 회전함에 따라 동공 간 거리는 좁아진다.· 두 동공의 평균 명암 비율은 얼굴이 좌우 또는 상하로 회전하는 동안 1보다 커지거나 1보다 작아진다.· 두 동공의 모양은 얼굴이 좌우 또는 상하로 회전하면서 보다 타원형이 된다.· 동공의

크기는 얼굴이 좌우 또는 상하로 회전하면서 작아진다. 이는 동공으로부터 얼굴 방향을 추정하기 위한 기초로 사용하여, 좌우동공크기, 좌우동공간 거리 및 좌우 동공 모양과 같은 좌우 동공 명암을 이용하여 얼굴 방향을 분석하여 이와 같은 특징들이 일관성을 가질 수 있도록 이 매개변수들을 얼굴 전면에서의 측정값으로 나누어줌으로써 표준화하였다. 본 논문에 제안된 방법의 문제점은 동공이 눈꺼풀로 가려질 때 동공의 식별이 이루어지지 않을 수 있다는 것이다. 이러한 문제를 해결하기 위해서, 특정 시점에 대한 얼굴 방향을 그 전후 프레임들에 대한 얼굴 방향 측정치들을 포함한 평균치로 계산하여야 한다. 또한 머리의 움직임이 자연스럽다는 가정 하에, 전후 시점에 대해 측정된 얼굴 방향 등과 같은 경험적인 제약 조건이 필요하다.

3.4 시선교정

시선 식별을 위하여 전처리 과정에서 얻는 동공과 글린트를 가지고 동공-글린트 벡터를 구성하고, 동공-글린트 벡터를 모니터 상에 매핑하는 기능을 수행토록 하였다. 동공과 글린트에 관한매개변수가 충분히 주어진다면 시선에 따라 유일한 기능을 갖는 동공과 글린트 매개변수를 찾고자 GRNN 을 이용하여 교정을 일반화 하고자 하였다. GRNN에 사용할 입력 벡터는 식(15)와 같다.

$$g = \Delta x \Delta y r \theta g_x g_y \dots \dots \dots (15)$$

글린트-동공 벡터의 크기는 적외선 조명기에 내장된 카메라와 피사체의 거리와도 연관된다. 얼굴과 동공이 움직이면 얼굴 움직임과 동공 이동을 계산하여 그에 부합하는 매개 변수로 갱신함으로써, 머리 움직임에 따른 영향을 감소할 수 있다. 또한 매개변수들은 동공의 크기에 대해서 독립적이므로 시선 매핑 기능을 일반화 시킬 수 있어 다른 사용자를 위한 재계산 작업을 효과적으로 제거할 수 있었다.

IV. 실험 결과 및 고찰

4.1 얼굴 동작 분류 결과 분석

검출된 얼굴 특징점은 각 요소의 움직임을 파악하

게 되고 얼굴표정을 분석하는데 사용되었다. 총 5명의 피검자를 대상으로 각각 실내에 앉은 상태로 각성 혹은 부주의와 같은 상태와 하품, 잠듦과 같은 얼굴 눈꺼풀의 변화량에 따른 얼굴 시선 영역을 추정한다. 실험 환경은 적외선 카메라 시스템을 이용하여 초당 20 프레임 320*240 해상도로 얼굴 영상을 약 1분씩 총 5회 기록하여 실험한다. 이때 같은 동작으로 모든 단계 부주의, 하품, 잠듦에 대하여 상향식 분류에 사전 확률을 설정한다. 실험을 위해 눈 검출 영역, 머리 방향 추정, 머리 움직임추정, 얼굴추적 등을 조합하여 a,b,c,d,e,f 로 동작을 가정하고 각 동작으로 모든 단계에 부주의, 하품, 잠듦에 대해 사전 확률을 측정한다. Ia, Yw, Fa는 각각 부주의, 하품, 잠듦을 나타낸다.

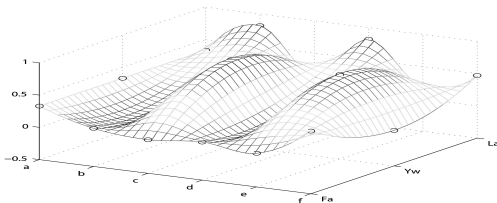


그림 2. 얼굴 동작 분류 결과
Fig. 2 The result of a face action grouping

그림 2-1에서 a는 Ia, Yw, Fa 값들이 0.5보다 작은 값인 0.33으로 동일하게 측정되었다. 이것은 각성 상태이거나 각성 상태일 경우 얼굴 동작의 변화가 없는 상태를 의미한다. b는 부주의의 상태에 대한 높은 사전 확률을 나타내었다. c는 Yw가 0.83으로 높게 측정되었으며 하품에 해당되고 d와 e는 분류 결과가 각각 부주의와 하품으로 측정된다. 즉, a,b,c,d,e는 피검자들의 여러 동작에 따른 얼굴 영역에서 시선 검출이 용이하다. 그러나 f는 부주의와 잠듦으로 사전 확률이 0.47로 동일하게 측정되었다.

모니터 상의 시선 위치를 파악할 수 있도록 적외선 조명기를 구성하여 밝은 동공 효과를 얻어 동공을 실시간으로 검출, 추적하였다.

4.2. 인접 영역에 대한 재식별 인식률 결과

학습에 참여하지 않은 사용자에 대한 실험 결과 전체적으로 84% 정도의 식별률을 나타내었다. 4.5번 영역과 인접한3.6 영역은 인접한 영역으로 오인식된 경

우가 많이 나타난다. 이는 학습에 참여한 경우보다 인식률이 현격하게 저하되는 것을 알 수 있으며 실제 응시 영역을 바르게 식별 못하고 인접한 것으로 식별한 것을 확인할 수 있다. 영역1을 바라보는 시선의 18%가 오인식되어 영역2로 나타나고 영역3을 바라보는 시선의 23%가 영역 4로 인식되었다. 결과적으로 오인식은 인접한 시선영역에서 주로 나타난다. 각각의 시선영역에 대한 인접 영역을 확인하고 인접한 영역에 대한 학습 데이터만을 이용하여 학습을 수행한 후, 인접한 영역만으로 재식별 하도록 하여 식별 결과를 검증한다. 재식별을 한 경우가 하지 않은 경우보다 9%정도 정확성이 향상되어 약 93% 정확도를 얻을 수 있었다. 특히, 인접영역 1과 2사이의 오인식률은 18%에서 8%로, 시선 영역 3과 4사이에 오인식률은 이전 24%에서 5%로 감소하였다. 다른 시선 영역에 의한 오인식 에러 역시 향상되었거나 동일한 성능을 나타내었다. GRNN을 통하여 시선 교정작업을 일반화함으로써 개개인에 대한 교정작업 이 필요없고, 얼굴의 움직임이 있는 경우에도 견실하게 동작하는 것이 큰 이점이라 할 것이다.

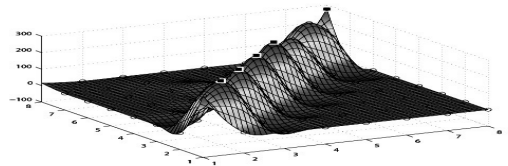


그림 3. 재식별기를 이용한 시선 식별 결과
Fig. 3 The gaze discrimination results using the gaze re-classifier

V. 결론

교정 작업 없이 얼굴 움직임이 있는 경우에 정확하게 동작하는 시선식별 시스템을 제안하기 위해서 모니터 좌표를 4*2로 양자화 하여 실험한 결과, 학습에 참여하지 않은 사용자가 얼굴의 움직임이 있을 경우에 2400개의 프레임 중 2075개의 프레임에 대하여 바르게 시선을 식별하여 약 84%의 정확성을 얻었으며, 바르게 인식되지 않은 결과를 분석한 결과 인접한 영역에 대하여 오인식이 있음을 알 수 있었다. 그러므로 오인식을 줄이고자 식별결과에 대하여 재 식별을 하

여 시선 영역을 검증토록 하였다. 인접영역의 시선을 재식별 하였을 경우 1과 2사이의 오 인식률은 18%에서 8%로, 3과 4사이에 오 인식률은 24%에서 5%로 감소하였다. 기존의 84% 정확성에 비해 약 9%정도 정확성이 향상되어 93%를 나타내었다. 제안된 시선 식별기는 자연스러운 얼굴 움직임이 있는 경우나 사용자에 따른 교정 작업이 없음에도 실시간 시선식별의 성능이 향상되었다. 향후에는 사용자의 안경착용 같은 제한사항을 극복하고 시선의 공간 해상도 높여서 시선 영역을 좀 더 세분화 하여 연구한다.

참고 문헌

- [1] J.Weng, D. L. Swets, "Face Recognition in Biometrics : Personal Identification in Networked Society", pp.67-86, Boston. MA KluwerAcademic,1999.
- [2] K. Sung, T.Poggio, "Example-based learning for view-based human face detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.20 Issue:1, pp.39-51, Jan. 1998.
- [3] L. Cohen, N. Sebe, A. Garg, M.A. Lew, T.S. Huang, "Facial expression recognition from video sequences", IEEE International Conference on Multimedia
- [4] T. Yamaguchi, M.Tominage, K. Mrakami, H. Koshimizu, "Regeneration of facial image eye-contacting with partner on TV Confernce environment", IEEE Iternational Conference on Systems, Man, and Cybernetics, vol.2, pp.1169-1174, 2000.
- [5] C. Collet, A. Finkel, R. Gherbi, "CapRe; a gaze tracking system in man-machine interaction", IEEE International Conference on intelligent Engineering Systems, pp.577-581, 1997.
- [6] P.Ekman and W.V. Friesen. "Emotion in the human face System."cambridge University Press, San Francisco, CA, second edition, 1982.
- [7] State of the art in facial Animation 22, acm siggraph '89 course notes,1989.
- [8] Isfan A. Essa and Alex P. Pentland. Facial expression recognition using visually extracted facial action parameters. In Int'l Work-shop on Automatic Face- and Gesture-Recognition '95,

pages 35-40,1995.

저자 소개



정유선(You-sun Jeong)

2005년 2월 조선대학교 대학원 컴퓨터공학과 (공학석사)

2009년 2월 조선대학교 대학원 컴퓨터공학과 박사수료

※ 주 관심 분야 : 컴퓨터통신, 컴퓨터그래픽스, voip, 영상처리



홍성수(Sung-soo Hong)

1996년 2월 조선대학교 산업공학과(공학석사)

2003년 2월 동신대학교 정보통신공학과(공학박사)

전 (주) 미래넷 대표이사

현 동강대학 컴퓨터정보과 부교수

동강대학 학술정보원 원장

KISTI 자문교수

※ 주 관심 분야 : 네트워크, 멀티미디어