

실시간 얼굴 검출 시스템 설계 및 구현

조현섭*, 오명관**

*청운대학교 디지털방송공학과

**해전대학 디지털서비스과

e-mail:chohs@chungwoon.ac.kr, mkoh@hj.ac.kr

Design and Implementation of a Real-Time Face Detection System

Hyun-Seob Cho*, Myoung-Kwan Oh**

*Dept of Digital Broadcast Engineering Chungwoon University

**Dept of Digital service, Hyejeon College

요 약

본 논문에서는 적외선 조명을 이용한 밝은 동공 효과와 전형적인 외형을 기반으로 한 사물 인식 기술을 결합하여 외부 조명의 간섭으로 밝은 동공 효과가 나타나지 않는 경우에도 견실하게 눈을 검출하고 추적 할 수 있는 방법을 제안한다. 눈 검출과 추적을 위해 SVM과 평균 이동 추적방법을 사용하였고, 적외선 조명과 카메라를 포함한 영상 획득 장치를 구성하여 제안된 방법이 효율적으로 다양한 조명하에서 눈 검출과 추적을 할 수 있음을 보여 주었다.

1. 서 론

눈의 위치를 검출하고 추적하는 것은 인간과 컴퓨터 상호작용의 중요 분야 중 하나로 키보드나 마우스 대신 눈의 움직임으로 작동하는 컴퓨터 등 응용 범위가 넓어 최근 많은 연구가 진행되고 있다. 본 논문에서는 외관을 이용한 방법과 능동적 적외선 조명을 이용한 방법을 혼합하여 다양한 조명하에서 견실하게 실시간으로 눈을 추적하는 방법을 제안하고자 한다. 제안 된 방법은 두 방법의 장점을 결합시키고 단점을 보완한 것으로 밝은 동공 효과를 만들어 내기 위해 능동적 적외선을 이용하여 사용자의 얼굴을 조명하여 주었고 패턴 분류 기술인 SVM(Support Vector Machine) 과 객체 추적기술인 평균 이동 방법을 눈의 검출과 추적에 이용하였다.

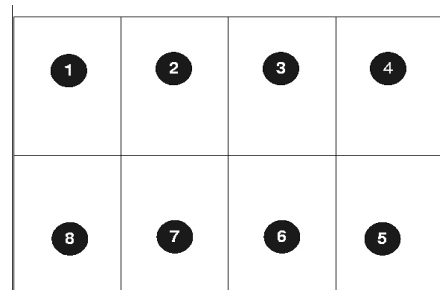
2. 영 상 획 득

영상 획득을 위하여 880nm 파장의 적외선 LED를 두 개의 고리모양으로 구성한 적외선 조명기를 제작하여 40nm 파장에서 32mW의 파워로 적외선을 조사하였다. 이 적외선 광원은 밝은 동공 영상과 어두운 동공 영상을 생성하였는데, 밝은 동공 영상은 안쪽 고리의 적외선 LED들이 켜졌을 때 생성되었으며, 어두운 영상은 바깥쪽 고리의 적외선 LED들이 켜졌을 때 생성된다. 적외선 조명 이외의 광원들로

부터 간섭을 최소화하여 영상의 질을 개선하기 위하여, 10nm 파장만을 통과시키는 광학 대역 필터를 사용하였고 필터를 사용함으로써 신호 대 잡음비를 20 이상 증가시키었다,

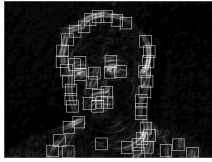
3. 동 공 검 출

교차된 입력 영상은 비디오 디코더에 의하여 잔상을 제거하여 그림2의 (a), (b)와 같은 짝수 필드 영상과 홀수 필드 영상을 만들어 내었다. 두 개의 영상은 같은 배경과 외부 조명을 사용하지만 짝수 필드의 동공이 홀수 필드보다 밝게 보였고, 배경과 외부조명의 간섭을 제거하고자 짝수 필드 영상에서 홀수 필드 영상을 제거하여 차영상을 만들었다. 차영상에서는 대부분의 배경과 외부 조명의 간섭효과가 제거되었다. 그림2 (c)에 차영상을 보여주고 있다.



[그림 2] 컴퓨터 스크린에 양자화 된 시선 영역

차영상은 동공 영역 외에도 많은 다른 영역을 나타낸다. 여기서 눈 영역을 찾아내기 위하여 우선적으로 동공이 있을 만한 모든 영역을 그림3과 같이 표시하였다.



[그림 3] 동공후보영역을 표시한 영상

일반적으로 동공은 타원 형태의 모양을 이루고 있으므로 타원 적합법[1]을 이용하여 각각의 후보영역을 추출한 후 형태와 크기가 다른 것을 후보 영역들 중에서 제거해 나간다. 즉, 영역의 크기가 크거나 좌우축 비율이 크다면 동공이 아니므로 후보 영역 중에서 제외하였다.



[그림 4] 형태나 크기 같은 기하학적 특성을 이용하여 일부 후보 영역을 제외한 영상

그림 4에서, 형태와 크기가 유사하여 동공과 구별하지 못한 영역이 일부 남아 있는 것을 볼 수 있다. 그러므로 본 논문에서는 SVM (Support Vector Machine) 분류기[2,3]를 이용하여 동공을 자동으로 구분하는 방법을 사용하였다.

3.1 SVM을 이용한 동공 검증

SVM은 두 범주를 갖는 객체들의 패턴 인식방법으로 Vapnik[4]에 의해 제안된 이후 그 성능을 인정받아 각광을 받고 있다. 표본으로부터 예측 학습을 다음과 같이 공식화 할 수 있다. 함수 f_α 의 집합이 다음과 같이 주어지고

$$\{f_\alpha : \alpha \in \Lambda\}, f_\alpha : R^N \rightarrow \{-1, +1\}$$

표본 집합 I 이 다음과 같다.

$$(x_i, y_i), K(x_i, y_i), K, (x_i, y_i), x_i \in R^N, y_i \in \{-1, +1\}$$

여기서 x_i 는 N 차원의 특징 벡터이고, y_i 는 -1 과 $+1$ 값의 범주를 나타내는데 확률 분포 $P(x, y)$ 로 각각 생성된다. 가장 적합한 함수 f_α^* 를 찾는 것이다.

$$R(\alpha) = \int |f_\alpha(x) - y| dP(x, y) \quad (1)$$

비선형 함수 g 를 이용하여 입력 데이터 X 를 고차

원 공간 Z 에 매핑함으로서 경계면 $D(X) = (w \cdot X + w_0)$ 을 SVM을 수행하여 분리한다. 최적의 경계면을 규정 짓고 최대 여백을 갖는 데이터 지점을 SV(Support Vector)라 한다. 학습 데이터는 최적의 경계면을 얻어야만 실효성을 갖는다. 동공 후보 지역을 얻은 후 어두운 동공 영상에서 경계면을 분리하였다. 보통, 분리된 영상에 눈이 포함되어 있다. 학습 전에 평활화와 정규화를 통하여 영상의 크기는 20×20 으로, 영상 데이터를 [5, 6]로 전처리 하였다. 눈 학습 영상은 양성 과 음성 두 개의 집합으로 구분되었다. 시선 방향이나 눈 개폐 여부, 안경 착용 여부에 관계없이 눈은 양성 영상 집합으로, 눈이 아닌 영상들이 음성 영상 집합으로 분류되었다. 그림 5와 6에 학습 집합에서의 눈영상과 눈이 아닌 영상을 보여주고 있다.



[그림 5] 양성 학습 집합의 눈 영상



[그림 6] 음성 학습 집합의 눈이 아닌 영상

3.2 분류되지 않은 데이터를 이용한 재학습

일반적으로 SVM은 분류된 학습 표본만을 사용하는 경우 정확성이 높지 않아 많은 분류되지 않은 데이터를 가지고 실험 한 후 잘못 분류된 데이터를 정확한 학습 집합에 넣어 재학습한다. 재학습을 위해서는 추가적인 시간이 필요하지만 정확성을 높일 수가 있다. 실험을 위하여 6명으로부터 눈 데이터 집합을 얻었다. 첫 번째 사람으로부터 얻은 데이터 집합을 눈 과 눈이 아닌 영상으로 수동으로 분류한 후 선형 SVM으로 학습하였다. 그리고 두 번째 사람으로 얻은 데이터 집합을 선형 SVM으로 자동 분류한 후, 잘못 분리된 데이터를 골라내어 정확한 학습 집합에 포함시켰다. 이와 같은 일련의 작업이 끝난 후, 향상된 학습 집합을 가지고 선형 SVM을 재 학습하였고, 다음 사람의 데이터 집합으로 가지고 위와 같은 작업을 반복하였다. 6명의 사람들의 모든 눈 후보 데이터 집합을 재학습을 마친 후 잘 선별된 학습 집합을 얻을 수 있었고 선형 SVM의 정확성을 높일 수 있었다.

3.3 SVM을 위한 학습 커널의 선택

위와 같은 단계를 마친 후, 558개의 양성 학습 벡터와 560개의 음성 학습 벡터를 얻을 수 있었다. 정확성을 높이기 위해서는 SVM을 위한 적합한 매개 변수를 선택하여야 한다. 표1에 학습에 참여하지 않은 사람으로 얻은 1757개의 눈 후보 영상을 매개변

수를 달리하여 세 가지 SVM 커널을 이용한 실험한 결과를 보여 주었다.

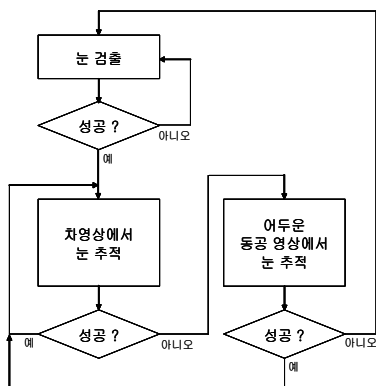
표로부터 최대 95.5037% 정확도를 얻을 수 있었다. 그러므로 SVM의 커널로서 시그마가 3인 가우시안 커널을 사용하였다.

[표 1] 다른 3가지 커널을 이용한 실험경과

커널 형태	Deg	Sigma	# Support Vectors	정확도
선형			376	0.914058
다항식	2		334	0.912351
	3		358	0.936255
	4		336	0.895845
Gaussian		1	1087	0.500285
		2	712	0.936255
		3	511	0.955037
		4	432	0.9465
		5	403	0.941377

4. 눈 추적 알고리즘

초기 프레임에서 동공을 검출하게 되면 동공을 다음 프레임에서 연속적으로 추적할 수 있다. 눈추적을 위한 기존의 방법[7]은 Kalman 필터를 이용하여 밝은 동공을 추적하는 것이었다. 그러나 이러한 방법은 얼굴 방향이나 외부 조명의 간섭으로 동공이 밝게 빛나지 않으면 동공은 차영상에서 제거되어 추적을 할 수 없게 된다. 본 논문에서는 차영상에서 밝은 동공을 찾지 못한 경우 바깥쪽 적외선 LED에 의한 독특한 패턴을 이용할 수 있다는 점에 착안하여 명암 분포 패턴을 이용한 방법과 밝은 동공 효과를 혼합한 눈 추적 방법을 제안한다. 그림 7에 제안된 추적 방법을 순서도로 나타내었다. 제안된 눈 추적 방법은 2단계로 이루어져 있다. 첫 번째 단계에서 밝은 동공 효과[8]를 이용한 Kalman 필터를 이용하고, 만약 추적에 실패할 경우, 두 번째로 어두운 동공 영상에서 평균이동알고리즘을 이용한 명암 분포에 근거하여 이전 프레임의 눈의 중심 근처에서 눈을 추적한다.



[그림 7] 눈 추적 알고리즘 순서도

4.1 평균 이동 눈 추적

4.1.1 추적 모델

눈표적의 특징 확률 분포를 표현한 m 과 회색조의 어두운 동공 영상을 끌어낸 1차원 히스토그램을 사용하였다. 히스토그램을 계산하기 전에, 눈 표적의 중심에서 먼 위치일수록 가중을 작게 할당하는 블록하고 단조 감소하는 커널 프로파일 k 를 선택하였다. $\{x_i\}_{i=1, \dots, n_h}$ 를 이전 프레임에서 y 의 중심인 눈 표적의 화소 위치라 하면, 눈 표적에 대한 명암의 분포 u 는 다음으로 구할 수 있다.

$$\hat{q}_u(y) = \frac{\sum_{i=1}^{n_h} k \left(\left\| \frac{y-x_i}{h} \right\|^2 \right) \delta[b(x_i)-u]}{\sum_{i=1}^{n_h} k \left(\left\| \frac{y-x_i}{h} \right\|^2 \right)} \quad (2)$$

여기서, $b(x_i)$ 는 히스토그램의 지표이고, h 는 커널 프로파일의 반지름, δ 는 크로네커 델타 함수이다. 표적 후보 분포 p 는 유사한 형태로 구축할 수 있었다.

4.1.2 알고리즘

이전 프레임에서 눈에 위치를 정한 후에, 이전 프레임에 발견된 눈을 근거로 하여 표적 눈 모형 q 를 구성한다. Kalman 필터를 이용하여 현재 프레임에서 눈의 위치 \hat{y}_0 를 예측하고, 다음의 알고리즘을 이용하여 초기 위치를 \hat{y}_0 로 설정하고 평균 이동을 반복하여 현재 프레임에서 눈 후보와 가장 유사한 곳을 찾는다.

1. 현재 프레임에서 표적의 위치를 \hat{y}_0 를 설정하고 분포 $\{\hat{p}_u(\hat{y}_0)\}_{u=1, \dots, m}$ 를 계산하고 값을 구한다.

$$\rho[\hat{p}(\hat{y}_0), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0) \hat{q}_u} \quad (3)$$

2. 다음에 의하여 가중치 $\{w_i\}_{i=1, \dots, n_h}$ 얻는다.

$$w_i = \sum_{u=1}^m \delta[b(x_i)-u] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \quad (4)$$

여기서 δ 는 크로네커 델타 함수이다

3. 평균 이동 함수를 이용하여 눈 표적의 새로운 위치를 얻는다.

$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} X_i w_i g \left(\left\| \frac{\hat{y}_0 - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n_h} w_i g \left(\left\| \frac{\hat{y}_0 - x_i}{h} \right\|^2 \right)} \quad (5)$$

여기서 $\{\hat{p}_u(\hat{y}_1)\}_{u=1, \dots, m}$ 를 갱신하고, 값을 구한다.

$$\rho[\hat{p}(\hat{y}_1, \hat{q})] = \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_1) \hat{q}_u} \quad (6)$$

4. $\rho[\hat{p}(\hat{y}_1, \hat{q})] < \rho[\hat{p}(\hat{y}_0, \hat{q})]$ 인 경우
 $\hat{y}_1 \leftarrow 0.5(\hat{y}_0 + \hat{y}_1)$ 를 실행한다.
5. 만약 $\|\hat{y}_1 - \hat{y}_0\| < \varepsilon$ 이면 중지하고
 다른 경우엔 , $\hat{y}_0 \leftarrow \hat{y}_1$ 로 설정하여 단계 1로 돌아간다.

현재 프레임에서 새로운 눈의 위치는 이전 눈 위치 근처를 전부 검색하는 상관관계를 이용한 방법에 비해 적은 반복 횟수로 찾을 수 가 있었다. 또한 계산들의 단순화하여, 상관관계에 비해 빠르게 동작하였다.

5. 결 론

본 논문에서는 다양한 조명과 다른 얼굴 방향에서 눈을 추적하는 새로운 방법을 제안하였다. 제안된 방법은 외관을 이용한 방법과 능동적 적외선을 이용한 밝은 동공 효과를 혼합하여 다양한 조명 조건하에서 견실하고 정확하게 눈을 추적하였다. 제안된 시스템은 강한 비적외선 조명과 통상 적인 적외선 조명하에서 기존의 밝은 동공 효과만을 이용한 방법보다 원활하고 정확하게 눈을 추적 할 수 있었다.

참 고 문 헌

- [1] BALUJA, S., AND POMERLEAU, D. "Non- intrusive gaze tracking using artificial neural networks." Technical Report CMU-CS-94- 102, Carnegie Mellon University. 2004.
- [2] COMANICIU, D., RAMESH, V., AND MEER, P. "Real-time tracking of non-rigid objects using mean shift." In *IEEE Conf on Comp. Vis. and Pat. Rec.* 2000.
- [3] CORTES, C., AND VAPNIK, V. "Support- vector networks." *Machine Learning* 20, 273- 297. 2005.
- [4] EBISAWA, Y., AND SATOH, S. "Effectiveness of pupil area detection technique using two light sources and image difference method." In *Proceedings of the 15th Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society*, 1268-1269. 2006.
- [5] FITZGIBBON, A. W., AND FISHER, R. "A buyers guide to conic fitting." In *Proc. 5th British Machine Vision Conference*, 513-522. 2007.
- [6] HARO, A., FLICKNER, M., AND ESSA, I. "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance." In *Proceedings IEEE CVPR 2000*. 2009.