

AdaBoost 알고리즘을 이용한 실시간 얼굴 검출 및 추적

이우주[†], 김진철^{‡‡}, 이배호^{***}

요 약

본 논문은 AdaBoost(Adaptive Boosting)알고리즘을 이용한 실시간 얼굴 검출 및 추적에 대한 기법을 제안한다. 얼굴 검출은 8종류의 간단한 웨이블릿 특징 모형을 이용한다. 각각의 특징들은 20*20의 훈련 영상에서 다양한 크기와 위치로 배치되어 초기의 특징 집합을 구성한다. 초기의 특징 집합과 훈련 영상은 AdaBoost알고리즘의 입력으로 사용된다. AdaBoost알고리즘의 기본원리는 약한 분류기를 선형적으로 결합하여 최종적으로는 계층적 구조를 갖는 강한 분류기를 생성하는 것이다. 본 논문에서는 AdaBoost알고리즘에서 훈련 영상과 초기의 특징 집합 간에 이루어지는 반복적 계산량을 줄이기 위해 SAT(Summed-Area Table) 기법을 이용하였다. 얼굴 추적은 Pan-Tilt 카메라를 통해 동적으로 가시 영역을 확장해 가면서 검출된 영역의 위치와 크기 정보를 이용하여 실시간으로 이루어진다. 검출된 얼굴 영역의 중심을 전체 영상의 중심으로 이동하는 방법을 사용하였다. 실험결과 92.5%의 얼굴 검출율과 평균 12프레임의 얼굴 추적속도를 얻었다.

Real-Time Face Detection and Tracking Using the AdaBoost Algorithm

Wu-Ju Lee[†], Jin-Chul Kim^{‡‡}, Bae-Ho Lee^{***}

ABSTRACT

In this paper, we propose a real-time face detection and tracking algorithm using AdaBoost(Adaptive Boosting) algorithm. The proposed algorithm consists of two levels such as the face detection and the face tracking. First, the face detection used the eight-wavelet feature models which are very simple. Each feature model applied to variable size and position, and then create initial feature set. The intial feature set and the training images which were consisted of face images, non-face images used the AdaBoost algorithm. The basic principal of the AdaBoost algorithm is to create final strong classifier joining linearly weak classifiers. In the training of the AdaBoost algorithm, we propose SAT(Summed-Area Table) method. Face tracking becomes accomplished at real-time using the position information and the size information of detected face, and it is extended view region dynamically using the Pan-Tilt camera. We are setting to move center of the detected face to center of the image. The experiment results were amply satisfied with the computational efficiency and the detection rates. In real-time application using Pan-Tilt camera, the detecter runs at about 12 frames per second.

Key words: AdaBoost Algorithm(AdaBoost 알고리즘), Face Detection(얼굴 검출), Face Tracking(얼굴 추적), Feature(특징 모형), Pan-Tilt Camera(Pan-Tilt 카메라)

* 교신저자(Corresponding Author) : 이우주, 주소 : 광주광
역시 북구 용봉동 300번지(500-757), 전화 : 062)530-0719,
FAX : 062)530-1759, E-mail : 2000ljw@hanmail.net

접수일 : 2006년 3월 23일, 완료일 : 2006년 9월 8일

[†] 준희원, 전남대학교 대학원 컴퓨터정보통신공학과 박사
과정

^{‡‡} 준희원, 전남대학교 대학원 컴퓨터정보통신공학과 박사
과정

(E-mail: kjc1015@korea.com)

^{***} 정희원, 전남대학교 전자컴퓨터공학부 교수

(E-mail: bhlee@chonnam.ac.kr)

* This work was supported by the RRC-HECS, CNU under grant R12-1998-032-08005-0, and NURI-CEIA, CNU.

1. 서 론

얼굴 검출은 영상 감시 시스템, 원격 회의 시스템, HCI(Human Computer Interaction) 및 얼굴 인식 시스템 등 다양한 분야에 응용될 수 있는 기술이다. 그러나 얼굴 영상은 조명, 시점 변화, 표정, 머리 모양, 화장, 안경 등 다양성으로 인해 큰 변화가 존재한다. 또한 얼굴과 유사한 수많은 비얼굴 데이터가 존재하기 때문에, 복잡한 배경에서 얼굴 영역을 완벽하게 분리하는데 현실적으로 한계가 있다[1].

기존의 얼굴 검출 기술은 신경망 이용 방법[2], 색상 기반 방법[3], 가우시안 복합 모델 기반 방법[4], 특징 기반 방법[5-6] 등이 있다. 신경망 이용 방법은 정지 영상에서 얼굴을 찾는데 우수한 성능을 보이지만 자연 배경과 같은 무수히 많은 비얼굴 데이터를 학습시키는데 한계점을 지닌다. 색상 기반 방법은 단독적으로 사용되지 않고 주로 다른 방법과 결합된 형태로 이용된다. 가우시안 복합 모델 기반 방법은 특정 패턴에 대해서 비교적 높은 성능을 보이지만 시점의 변화와 같은 영상 전체의 변화에 대해서는 추출하기가 어렵다. 특징 기반 방법은 크기, 방향, 얼굴의 시점 변화 등에 상관없이 유용하게 적용할 수 있는 방법이지만, 눈, 코, 입과 같은 얼굴의 특징을 찾는 것이 쉽지 않다. 최근에는 간단한 웨이블릿 특징을 기반으로 Boosting 알고리즘을 이용하여 실시간으로 얼굴 영역을 검출하는 연구가 활발히 진행되고 있다[7-8].

본 논문은 계층적 분류기를 제안하여, 계산의 효율성과 검출 성능을 동시에 만족시키며, 얼굴 영역의 다양성을 수용할 수 있는 강인한 검출 및 추적 알고리즘을 구현하고자 한다. 제안한 알고리즘은 분류기 학습, 실시간 얼굴 영역 검출, 추적의 세 단계로 구성된다.

분류기 학습 단계에서는 고유한 얼굴 특징을 추출하는 사각형 특징 마스크로 구성된 계층적 분류기를 생성한다. 분류기는 3,000여개의 얼굴 영상과 5,600여개의 비얼굴 영상으로 구성된 훈련 영상 집합과 57,761개의 특징 집합으로 AdaBoost 알고리즘을 적용하여 학습된다. 이때의 특징 집합은 제안된 8개의 웨이블릿 특징 모형을 기반으로 하여 다양한 특징을 생성한다. 생성된 계층적 분류기는 높은 정확도를 가진 분류기들이 단계적으로 결합됨으로써 우수한 검

출 성능을 가진다. 실시간 얼굴 영역 검출은 생성된 계층적 분류기를 통해, 빠르고 효율적으로 얼굴 영역을 찾아낸다. 얼굴 추적은 Pan-Tilt 카메라를 통해 동적으로 가시 영역을 확장시키며, 이전 단계에서 추출된 얼굴 영역의 위치와 크기 정보를 이용하여 수행된다.

본 논문은 총 5장으로 구성된다. 1장은 서론으로 얼굴 검출 및 추적 시스템에 대한 전반적인 소개와 본 논문에서 제시하고자 하는 방법을 요약하고 있다. 2장에서는 8종류의 특징 모형을 기반으로 생성된 특징들을 AdaBoost 알고리즘을 통하여 학습하는 방법과 학습을 통하여 생성된 계층적 분류기를 이용하여 얼굴을 검출하는 과정을 설명한다. 3장에서는 Pan-Tilt 카메라를 이용하여 카메라의 동적인 시야를 확보하고 그것을 이용하여 얼굴을 추적하는 방법을 설명한다. 4장에서는 제안한 알고리즘에 대한 항목별 실험결과를 보이고 분석 및 평가한다. 마지막 5장은 본 논문의 결론으로 제안된 알고리즘에 대한 총체적인 평가로 끝맺음을 한다.

2. AdaBoost 알고리즘을 이용한 얼굴 검출

2.1 웨이블릿 특징 모형

웨이블릿 기반의 특징 계산은 픽셀 기반의 특징 계산보다 계산을 효율적으로 처리할 수 있다. 또한 간단한 구조에도 불구하고 특정 영역에 대한 에지나 라인 등의 정보를 잘 제공해 준다[9-7]. 특징 모형의 선택은 이후에 전체 얼굴 검출의 연산 속도에 결정적인 영향을 미치게 되기 때문에 매우 중요하다. 파울비올라는 4종류의 특징 모형을 사용하여 좋은 결과를 얻었다[7]. 본 논문에서는 파울비올라의 방법을 확장하여 [그림 1]과 같이 고유한 얼굴의 특징을 추출하는데 이용할 8종류의 웨이블릿을 제안하였다. [그림 1]의 (a), (b)는 에지의 특징, (c), (d), (f), (g)는 라인의 특징, (h)는 포인트의 특징, (e)는 대각선의 특징을 나타낸다. 각각의 특징 모형은 얼굴 요소에서 특징들을 만들어 낸다.

8종류의 간단한 웨이블릿 특징 모형은 20*20의 훈련영상에 맞추어 다양한 위치와 다양한 크기로 초기 특징 집합을 생성한다. 초기 특징 집합은 특징 모형에 의해 생성되어지고 얼굴 검출에 사용될 가능성이 있는 특징들을 전부 모아 놓은 것이다. 이때 8종류의

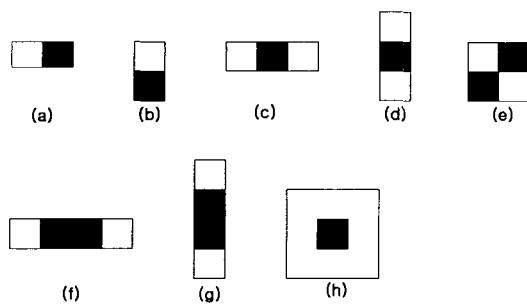


그림 1. 웨이블릿 특징 모형

특징 모형은 무수히 많은 특징들을 생성하게 된다. 후에 초기 특징 집합들은 다시 학습을 통하여 얼굴을 가장 잘 구별할 수 있는 특징들로 선택될 가능성이 있기 때문에 매우 중요하다. 제안된 8종류의 특징 모형을 사용하여 생성된 초기 특징 집합의 전체 수는 57,761개이다.

학습을 위하여 초기 특징 집합은 특징 값으로 변환되어진다. 그림에서와 같이 특징 모델을 이용하여 내부의 특징 값을 구하기 위해서는 내부의 모든 픽셀의 연산이 필요하다. 이러한 과정은 여러 가지 특징들과 훈련 영상 이미지의 연산에서 반복적으로 일어난다.

특징값 계산이 반복적으로 일어나기 때문에 학습 단계에서 많은 시간이 걸린다. 본 논문에서는 이와 같은 계산 시간을 줄이기 위해 학습하기 전에 훈련 영상의 집합을 특징값 누적 영상(Integral Image)으로 변환하였다. 훈련 영상을 특징값 누적 이미지로 변환하는 이유는 방대한 특징 집합과 계산에서 연산량을 줄이기 위해서다. 특징값 누적 영상을 생성하기 위해 [그림 2]와 같이 영역의 합 테이블(Summed Area Table)을 이용하였다[10]. SAT는 영상 내에 존

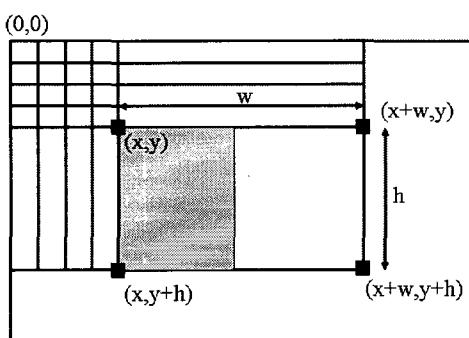


그림 2. 영역의 합 테이블(SAT : Summed Area Table)

재하는 특정한 좌표 영역(x, y)의 값은 입력 영상의 원점(0, 0)에서부터 좌표 영역(x, y)를 포함하는 사각형 영역 내에 존재하는 모든 픽셀 값을 더한 것이며, 식(1)과 같이 정의 된다.

$$S(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (1)$$

따라서, 하나의 사각형 특징값($RecS$)은 식(2)와 같이 정의된다. 이 값은 미리 계산된 SAT 테이블을 이용하여, 네 개의 SAT를 사용하여 효율적으로 계산된다.

$$RecS(x, y, w, h) = S(x+w, y+h) - S(x+w, y) - S(x, y+h) + S(x, y) \quad (2)$$

식(2)에서 x 와 y 는 가로 x , 세로 y 인 특정한 좌표 영역이며, w 와 h 는 너비 w , 높이 h 인 사각형 영역을 나타낸다. SAT는 [그림 2]에서 도식화된다.

선택되어진 전체 영역의 특징값($RecF$)은 식(2)를 사용하여 전체 사각형 영역을 포함하는 흰색 사각형 픽셀값($RecSw$)의 평균과 검정색 사각형 픽셀값($RecSb$)의 평균에 대한 차로 구하며, 식(3)와 같이 정의된다. 즉, 사각형 특징 마스크의 특징 값은 흰색과 검정색 영역의 각 평균을 차 연산한 결과 값으로 얻어진다.

$$RecF(x, y, w, h) = \frac{1}{m} \sum_m RecSw - \frac{1}{n} \sum_n RecSb \quad (3)$$

사각형 특징 마스크는 [그림 1]에서 언급한 두 개의 흰색과 검정색 사각형 영역이 오버랩된 형태이다. 식(3)에서 m 은 $RecSw$ 의 픽셀 수와 n 은 $RecSb$ 의 픽셀 수이다.

이와 같이 생성된 누적 영상과 특징값 계산 기법은 최종적으로 약한 분류기를 구성하는 함수의 계산 방법에 이용된다.

2.2 AdaBoost알고리즘

본 논문에서는 얼굴 검출에 필요한 특징을 추출하기 위해 AdaBoost(Adaptive Boosting)알고리즘을 사용하였다. AdaBoost알고리즘의 기본 개념은 약한 분류기(Weak classifier)를 선형적으로 결합하여 최종적으로 높은 검출 성능을 가진 강한 분류기(Strong classifier)를 생성하는 것이다[11-12]. AdaBoost알고리즘에 의해 생성된 강한 분류기는 [그림 3]과 같

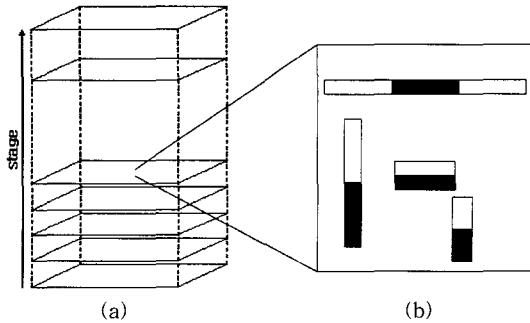


그림 3. 계층적 분류기: (a) 강한 분류기, (b) 약한 분류기

이 계층적인 체계를 이룬다. 이것은 기존의 다른 방법들이 하나의 복잡한 마스크 형태의 분류기를 이용한 것과 달리 간단한 마스크를 여러 개의 층으로 형성한 것이다. 기존의 분류기들은 실제 영상에서 얼굴을 검출하는데 복잡한 마스크와의 계산이 이루어지기 때문에 계산량이 많고 시간이 많이 소요되었다. 그러나 AdaBoost 알고리즘을 이용하여 생성된 계층적 분류기는 앞쪽 부분에 간단하면서도 얼굴을 가장 잘 검출하는 것을 배치하고 뒤쪽 부분에 잘못 검출된 것을 제거하는 형식으로 구성되어 있다. 이는 실시간 검출에 있어서 기존의 방법에 비하여 뛰어난 성능을 보인다. 이런 점을 고려하여 실시간으로 얼굴을 검출하고자 본 논문에서는 AdaBoost 알고리즘을 사용하였다.

초기 특징 집합과 얼굴, 비얼굴 훈련 영상을 이용하여 학습을 한다. AdaBoost 알고리즘은 초기 특징 집합에서 얼굴을 가장 잘 검출할 수 있는 특징의 선택과 분류기(classifier) 학습을 동시에 한다. 즉, AdaBoost 알고리즘 학습 결과로 얼굴 검출을 위한 특징과 분류기가 생성된다. 각각의 훈련 영상에 57761개의 초기 특징 집합이 결합된다. 이렇게 되면 기본 핵심의 수보다 더 많은 경우가 생긴다. 각 특징은 SAT방법에 의해 효율적으로 계산되지만 결합된 집합은 비교적 많은 시간이 걸린다. 효율적인 분류기를 생성하기 위해서는 적은 수의 특징으로도 얼굴을 검출할 수 있는 것을 찾아야한다. 이러한 목표에 도달하기 위해 약한 분류기는 얼굴과 비얼굴을 가장 잘 구별할 수 있는 하나의 특징만을 선택하도록 설계된다. [그림 4]와 같이 약한 분류기는 각 특징에 대하여 에러를 최소로 갖는 최적화 경계 분류 함수(optimal threshold classification function)를 결정한다. 약한 분류기는 식(4)와 같이 구성된다.

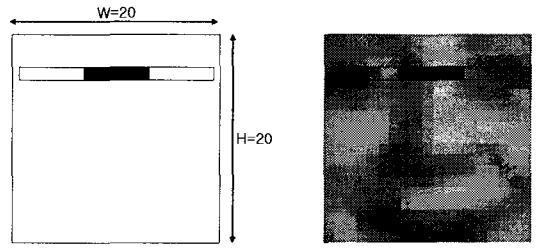


그림 4. 약한 분류기

$$h_j(x) = \begin{cases} 1 & \text{if } f_j(x) < \theta_j \\ -1 & \text{if } f_j(x) \geq \theta_j \end{cases} \quad (4)$$

그러나 실제 실험에서는 하나의 특징으로는 낮은 에러율에서 분류 작업을 실행할 수 없었다. 따라서 일정한 기준 에러율을 두고 거기에 맞추어 선택적으로 특징을 구하였다.

강한 분류기는 약한 분류기의 선형적 결합형태로 여러 개의 특징을 약한 분류기를 결합해서 실질적으로 얼굴의 패턴을 구별하는 역할을 한다. 본 논문에서 사용한 AdaBoost 알고리즘은 [그림 5]와 같다.

2.3 AdaBoost를 이용한 계층 분류기의 학습

여기에서는 계산 시간을 줄이면서도 성공률을 높

1. Given N examples $(x_1, y_1), \dots, (x_N, y_N)$ with $x \in R^k, y_i \in \{1, -1\}$
2. Initialize weights $w_1(i) = \frac{1}{m} \cdot \frac{1}{n}, i = 1, \dots, N$
m: number of positive, n: number of negative
3. Repeat for t=1, ..., T
 - (a) Train weak learner using weight w_t
 - (b) Repeat for j=1, ..., number of intial feature
Get weak classifier $h_t = X \rightarrow \{1, -1\}$
with error $\epsilon_{j1} = \sum_i w_t(i) |h_t(x_i) - y_i|$ (positive)
 $\epsilon_{j2} = \sum_i w_t(i) |h_t(x_i) - y_i|$ (negative)
 - choose the classifier using error (ϵ_t, h_t)
 - (c) weight update
 $\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right)$
 $w_{t+1}(i) = w_t(i) \cdot \begin{cases} e^{-\alpha_t} & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t} & \text{if } h_t(x_i) \neq y_i \end{cases}$
4. Output the final hypothesis(strong classifier)
 $H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$

그림 5. AdaBoost 알고리즘

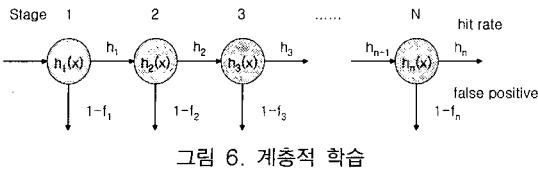


그림 6. 계층적 학습

이기 위해 계층적 분류기를 어떻게 구성하느냐에 대하여 설명한다. 분류기들은 얼굴을 검출하고 비얼굴은 제거하도록 구성된다. 간단한 분류기는 대부분의 초기 특징들을 제거하는데 사용되고 보다 복잡한 분류기는 낮은 (비 얼굴을 얼굴로 검출할 오류)false positive rate에 도달하기 위해 사용된다. 첫 번째 분류기의 positive result는 높은 검출률에 도달하기 위해 조절되는 두 번째 분류기의 평가를 변환한다. 두 번째 분류기의 positive result는 세 번째 분류기를 변환한다. 계층들은 AdaBoost를 이용한 분류기 학습에 의해 생성되고 그 후에 (얼굴을 비 얼굴로 검출 할 오류)false negative rate를 최소화하기 위해 문턱치를 조절한다. 초기 AdaBoost 알고리즘의 문턱치는 훈련데이터에서 낮은 에러율을 가지고 설계된다. 일반적으로 보다 낮은 문턱치는 높은 검출율과 높은 false positive rate를 낳는다.

3. Pan-Tilt 카메라를 이용한 얼굴 추적

얼굴 추적은 Pan-Tilt 카메라를 통해 동적으로 가시 영역을 확장해가며, 추출된 얼굴 영역의 위치와 크기 정보를 이용하여 실시간으로 추적한다. 추적 시, 검출된 얼굴 영역의 중심을 전체 영상의 중심에 두도록 설정하며, 따라서 검출된 영역의 중심 좌표와 전체 영상의 중심 좌표는 pan/tilt 카메라를 제어하는 파라미터로 이용된다.

전체 시스템은 [그림 7]과 같으며, 분류기 학습, 실시간 얼굴 영역 검출, 얼굴 추적의 세 단계로 구성된다.

분류기 학습 단계는 오프라인에서 진행되며, 훈련 영상 집합과 생성된 특징 집합을 이용한다. 훈련 집합은 얼굴과 비얼굴 패턴으로 구성된 데이터베이스이며, 특징 집합은 얼굴의 특징을 추출할 수 있는 사각형 특징 마스크들로 구성된다. 학습 과정은 Ada-Boost 알고리즘으로 진행되며, 분류기 학습의 궁극적인 목적은 얼굴과 비얼굴 패턴을 분류하는데 결정적인 역할을 하는 특징들을 선택하는 것이다. 학습을 통해 생성된 강분류기는 특징들이 선형적으로 결합

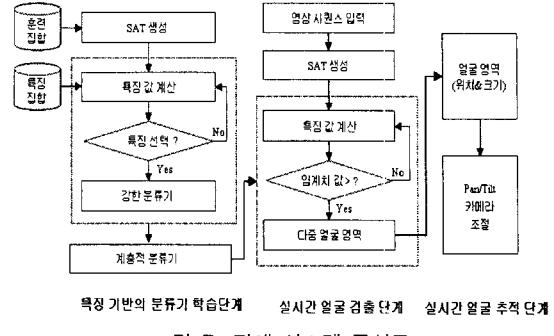


그림 7. 전체 시스템 구성도

된 형태이며, 최종적인 계층적 분류기는 강분류기가 단계적으로 결합된 형태이다.

실시간 얼굴 영역 검출 단계는 계층적 분류기를 이용하여, 실시간으로 얼굴 영역을 찾아낸다. 계층적 분류기는 사각형 특징 마스크들을 이용하여, 입력 패턴에 대한 특징 값을 계산한다. 따라서 입력 패턴은 계산된 특징 값이 임계 값을 만족하면, 얼굴 영역으로 분류된다.

얼굴 추적 단계는 이전 단계에서 검출된 얼굴 영역의 위치와 크기 정보를 이용하여, Pan-Tilt 카메라를 통해 실시간으로 수행된다.

Pan-Tilt 카메라를 이용한 얼굴 추적에는 프레임 간의 차영상, 얼굴의 움직이는 방향, 흐름 정보를 이용한다[13]. 그러나 카메라와 추적할 얼굴이 같이 움직인다면 알고리즘이 복잡해지고 이로 인하여 초당 10프레임 이상을 지원해야하는 실시간 얼굴 추적이 어렵게 된다. 또한 기존에 얼굴 추적 방법은 프레임 간의 차를 이용하여 움직임을 추적하는 방법이 사용되어 왔으나 대부분이 실시간을 고려하고 있지 않은 수학적 접근법을 사용하거나 알고리즘이 지나치게 복잡하여 실시간 구현에 용이하지 않는다[14]. 복잡한 얼굴 검출과 함께 얼굴 추적이 동시에 실시간으로 이루어지기 위해서는 보다 간단하면서도 실용적인 기법이 연구되어져야 한다.

얼굴 추적은 얼굴 검출율이 매우 높다는 것을 기본 체제로 하여 검출된 얼굴의 크기정보와 위치정보를 이용하여 추적한다.

이 방법은 몇 가지의 파라미터를 수식적으로 계산하기 때문에 추적 속도가 매우 빠르다. [그림 8]은 영상에서 얼굴을 추적하는 기본 방법을 나타낸 것이다.

검출된 얼굴의 크기 정보는 (w, h) 이며, 위치 정보는 (x, y) 이다. 이것을 이용하여 얼굴의 중심 좌표(x_c ,

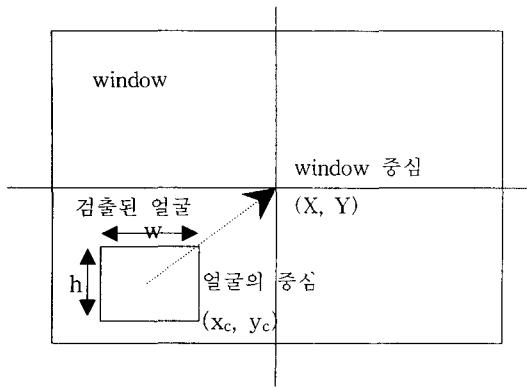


그림 8. 영상에 기반한 얼굴 추적

y_c)를 구하면 식(5)과 같다.

$$(x_c, y_c) = \left(\frac{x + (x + w)}{2}, \frac{y + (y + h)}{2} \right) \quad (5)$$

얼굴 추적은 얼굴의 중심을 영상의 중심(X, Y)에 위치하도록 하면서 이루어진다. 영상의 중심과 얼굴의 중심의 위치 정보를 이용하여 Pan-Tilt를 구동하기 위한 파라미터를 결정한다. 이때 얼굴의 움직임을 부드럽게 하기 위하여 얼굴의 중심과 영상의 중심의 거리에 따라 가우시안 분포 가중치를 적용한다. 가우시안 분포 가중치는 식(6)과 같다.

$$(x_m, y_m) = ((X - x_1) * Gx, (Y - y_1) * Gy) \quad (6)$$

이 방법은 얼굴 검출율이 높다는 기본 가정 하에 사용한 방법이다. 그러나 본 논문에서 사용한 AdaBoost 알고리즘을 이용한 평균 얼굴 검출율이 92.5% 이므로 순간적으로 얼굴 검출을 실패한 경우가 생기게 된다.

본 논문에서는 한 프레임에 다중의 얼굴이 검출되었을 경우 처음에 얼굴을 추적하기 위하여 영상의 중심과 각각의 얼굴의 중심의 거리에 따라 얼굴 추적의 우선순위를 두었다.

4. 실험 결과 및 분석

본 논문에서는 AdaBoost 알고리즘을 이용하여 얼굴 검출하기 위해 MIT CBCL 얼굴 데이터를 이용하였다. 원 MIT CBCL 훈련 데이터는 19*19 크기로 구성되어 있다. 본 논문에서는 얼굴 데이터들과 특징들 간에 계산의 효율성을 높이기 위해 20*20 크기로

정규화 하였다. 얼굴 데이터는 2901개를 사용하였고 비얼굴 데이터는 5600개를 사용하였다. 본 논문에서는 학습을 위한 파라미터를 몇 가지로 정하였다. 먼저 각 단계마다 최소한 얼굴 데이터를 얼굴로 검출할 확률(Min hit rate)을 0.995로 하였다. 비 얼굴을 얼굴로 검출할 최대 오류 확률(Max false positive)을 0.050로 하였다. 각 단계에 학습에서 최소한 얼굴 데이터를 얼굴로 검출할 확률이 0.995보다 작거나 비 얼굴을 얼굴로 검출할 최대 오류 확률이 0.050보다 커지면 다음 단계로 전이된다. 최종적인 학습 계층은 20계층으로 선택하였고 얼굴 검출율이 90% 이상이 될 때까지 학습하였다.

AdaBoost 알고리즘을 이용하여 실험한 결과 계층적 분류기의 계층은 20계층이 생성되었다. 특징 선택에 있어서는 초기 특징 집합 57761개 중 총 1630개가 얼굴 검출에 강인한 특징으로 선택되었다. 각 계층별 선택된 특징의 개수는 [그림 9]와 같다.

특징의 수는 전반부의 계층일 경우 수가 적게 나타나고 후반부로 갈수록 많아진다. 이것은 전반부의 경우 특징이 얼굴 패턴을 모두를 검출하기 위해서 대체적으로 크기가 큰 특징을 선택하기 때문이다. 반면에 후반부의 계층은 얼굴 패턴을 정교하게 검출하고 비얼굴 패턴을 확실하게 제거하기 위해서 크기가 작고 많은 수의 특징을 포함하게 된다.

계층별 얼굴 데이터를 얼굴로 검출할 확률(hit rate)은 [그림 10]과 같다. 계층별 얼굴 데이터를 얼굴로 검출할 확률은 각 단계를 통과하면서 나타나는 얼굴 데이터를 얼굴로 검출할 확률을 나타낸 것이다. 앞에 학습 파라미터에서 설명한 것과 같이 각 계층별 최소한 얼굴 데이터를 얼굴로 검출할 확률은 0.995이

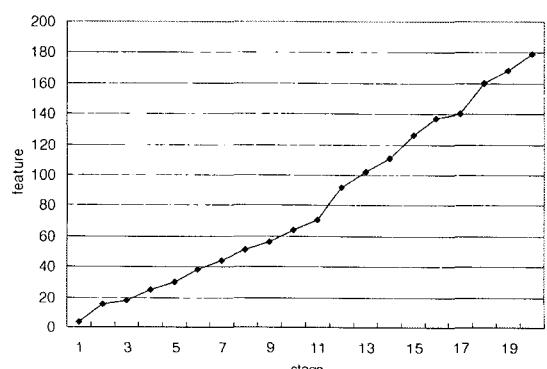


그림 9. 계층별 특징의 개수

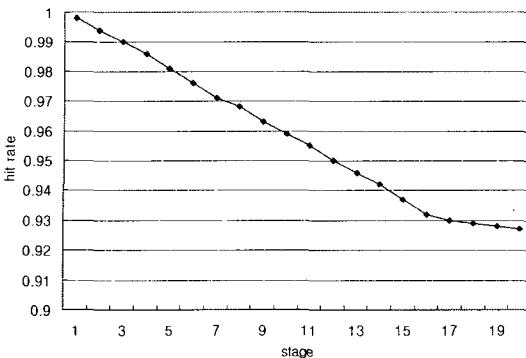


그림 10. 계층별 얼굴데이터를 얼굴로 검출할 확률

다. 이것을 바탕으로 각 계층을 통과할 때 0.995로 제한되어 있기 때문에 [그림 10]과 같이 감소하는 형태의 그래프가 나타난다.

실험분석결과 이와 같이 각 계층에서 비얼굴 패턴을 제거하기 위해 가중치를 조절하는 과정에서 장인하지 못한 얼굴 패턴이 일부 같이 제거되었기 때문에 감소 형태의 그래프가 나타난다.

계층별 얼굴 검출율은 [그림 11]과 같다. 최종 검출율은 92.5%를 나타내었다.

제안한 얼굴 검출 알고리즘의 성능 평가는 MIT CMU 테스트 집합을 이용하였으며, 총 130개의 영상을 이용하였다. 총 130개의 영상에는 507개의 얼굴이 포함되었다. 표 1은 제안된 알고리즘과 기존의 얼굴 검출 기법간의 검출율을 비교한 것이다. 각각의 기법에 따라 검출율을 계산하는 방식이 차이가 있어 대략적으로 평균 검출율을 기준으로 비교한 것이다. [표 1]에서와 같이 제안된 얼굴 검출 알고리즘은 Viola-Jones[7], Rowley[2], Sheiderman-Kanade[15]에 비해 우수함을 나타내었다.

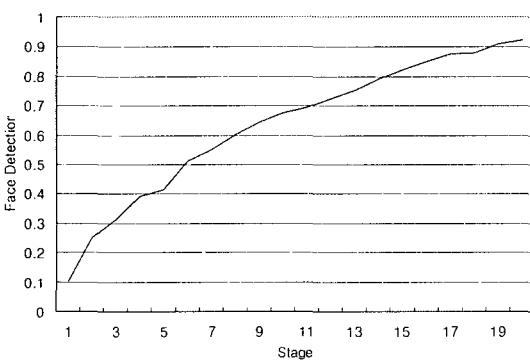


그림 11. 계층별 얼굴 검출율

표 1. 얼굴 검출 기법에 따른 검출율

얼굴 검출 기법	검출율
제안한 알고리즘	92.5%
Viola-Jones	92.0%
Rowley	90.1%
Sheiderman-Kanade	91.4%

표 2. 특징 모델 수에 따른 결과 비교

비교 항목 특징 모델 수	초기특징 집합수 (개)	선택된 특징수 (개)	계층수 (계층)	검출율 (%)
4개	32000	2217	30	92.0
5개	42000	2137	24	92.1
8개	57761	1630	20	92.5

특징 모델의 수에 따른 결과 비교는 [표 2]와 같다. 특징 모델의 수가 적을수록 초기 특징 집합수에 따른 선택된 특징수가 많고 계층수도 많다. 반면에 모델의 수가 많을수록 초기 특징 집합수에 따른 선택된 특징 수는 적고 계층수 역시 적다. 실험 결과 특징 모델의 수가 많을수록 얼굴의 복잡한 특징들을 잘 구별할 수 있고 목표 검출율에도 빠르게 도달하는 것으로 나타났다. 그러나 특징 모델의 수가 많을수록 학습에 필요한 시간은 증가하게 된다.

[표 3]은 이미지 Grab, 검출, 추적, 저장을 순차적으로 누적하여 처리한 평균 시간을 나타낸 것이다.

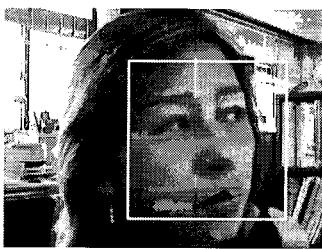
제안된 얼굴 검출 알고리즘은 다양한 크기의 얼굴과 다중 얼굴 검출이 가능하였다. 또한 다양한 조명 변화와 복잡한 환경, 얼굴 영역에 겹침 현상이 발생하였을 때도 장인함을 보였다. [그림 12]부터 [그림 16]은 다양한 조건하에서 실시간 얼굴검출 결과를 나타낸 것이다.

표 3. 실시간 처리

	이미지 Grab	얼굴 검출	얼굴 추적	이미지 저장
처리시간 (sec)	0.05	0.067	0.083	0.1
1초당 프레임수 (frame)	20	15	12	10

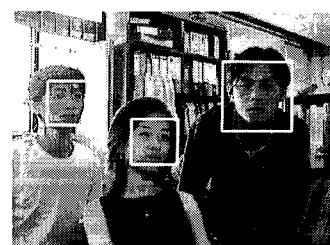


(a)



(b)

그림 12. 실시간 얼굴 검출 결과 : 다양한 얼굴 영역 크기: (a) 작은 크기, (b) 큰 크기

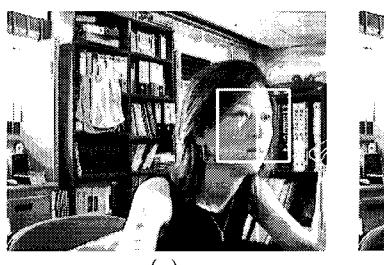


(a)

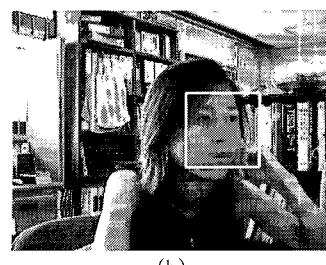


(b)

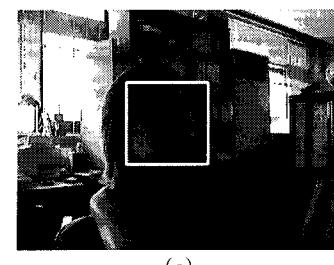
그림 13. 실시간 얼굴 검출 결과 : 다중 얼굴: (a) 세 사람, (b) 두 사람



(a)

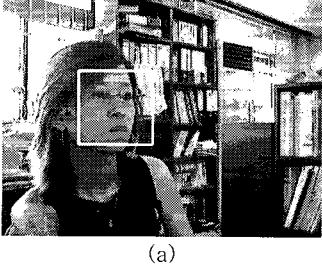


(b)

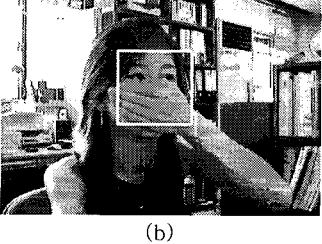


(c)

그림 14. 실시간 얼굴 검출 결과 : 다양한 조명 변화: (a) 밝은 조명, (b) 정상 조명, (c) 어두운 조명

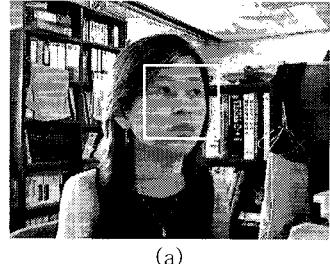


(a)

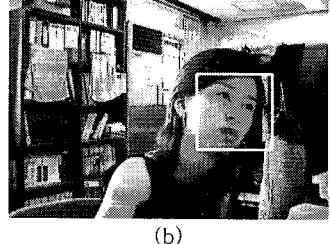


(b)

그림 15. 실시간 얼굴 검출 결과 : 얼굴 영역 겹침 발생: (a) 안경착용, (b) 얼굴영역 겹침



(a)



(b)

그림 16. 실시간 얼굴 검출 결과 : 얼굴의 기울어짐 발생: (a) 정상 모습, (b) 기울인 모습

5. 결 론

본 논문은 실시간 영상 감시를 위한 얼굴 검출 및 추적에 대한 알고리즘을 제안하였다. 얼굴 검출은 8종류의 간단한 웨이블릿 특징을 이용하였다. 각각의 특징들은 훈련 영상에 맞추어 다양한 초기 특징 집합을 구성하였다. 이렇게 만들어진 초기의 특징 집합과 얼굴과 배경으로 이루어진 훈련 영상은 AdaBoost알고리즘을 이용하여 얼굴과 비얼굴을 분류할 수 있는 분류기 생성 학습을 하였다. AdaBoost알고리즘을 통하여 계층적 분류기를 얻을 수 있었고 계층적 분류기는 실시간 얼굴 검출에 높은 성능을 보였다. 얼굴 추적은 Pan-Tilt 카메라를 통해 동적으로 가시 영역을 확장해가면서, 검출된 얼굴 영역의 위치와 크기 정보를 이용하여 실시간으로 이루어졌다. 얼굴 추적은 검출된 얼굴 영역의 중심을 전체 영상의 중심으로 이동하는 것을 기본 방법으로 설정하였다. 또한 얼굴 검출 실패에 따른 얼굴 추적 실패를 방지하기 위해 10프레임 간에 얼굴의 위치와 크기 정보를 저장하였다.

실험 결과 본 논문에서 제안한 AdaBoost 알고리즘을 이용한 얼굴 검출이 다른 얼굴 검출 알고리즘과의 성능 평가에서 우수한 성능을 나타내었다. 또한 특징 모델수를 4개나 5개를 사용했던 기존의 AdaBoost알고리즘보다 초기 특징 집합수에 따른 선택된 특징수도 적고 계층수도 적어 목표 검출율에도 빠르게 도달할 수 있었으며, 최종적으로 제안된 알고리즘은 평균 92.5%의 얼굴 검출율을 보였다. 제안된 얼굴 검출 알고리즘 다양한 조명 변화와 복잡한 환경, 얼굴 영역에 겹침 현상이 발생하였을 때도 장인함을 보였다. 얼굴 추적은 Pan-Tilt카메라를 이용하여 실험하였다. 실험결과 평균 12프레임의 속도로 얼굴 추적이 가능하였다.

제안된 알고리즘은 계산의 효율성과 검출 성능을 동시에 만족시키는 장인한 검출 및 추적 알고리즘이다. 알고리즘은 제안된 8종류의 특징 모형을 이용하여, 얼굴 패턴에 대한 중요한 구조적 특징을 성공적으로 추출한 것으로 평가된다. 따라서 얼굴 영상에서 발생하는 다양성을 수용할 수 있었으며, 기존의 특징 기반 얼굴 검출 방법보다 간단하면서도 효율적으로 얼굴을 검출을 수행할 수 있는 새로운 특징 추출 방법을 제공한다.

참 고 문 헌

- [1] R.O. Duda, P.E. Hart, and D.G. Strok, *Pattern Classification*, Second Edition by John Wiley & Sons, Inc, 2001.
- [2] H.A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *In IEEE Patt. Anal. Mach. Intell.*, Vol. 20, pp. 22-38, 1998.
- [3] J. Yang and A. Waibel, "A real-time face tracker," *In Proc. 3rd Workshop on Appl. of Computer Vision*, pp. 142-147, 1996.
- [4] K.-K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 20, Issue:1, pp. 39-51, Jan. 1998.
- [5] R. Brunelli and T. Poggio, "Face Recognition: features versus templates," *IEEE Transactions on PAMI*, Vol. 15, No. 10, pp. 1042-1052, 1993.
- [6] Y. Yagi, "Facial feature extraction from frontal face image," *Signal Processing Proceedings, 2000. WCCC-ICSP 2000. 5th International Conference on*, Vol. 2, pp. 1225-1232, 2000.
- [7] Paul Viola and M. Jones, "Robust real-time object detection," *International Conference on Computer Vision*, 2001.
- [8] S.Z. Li, L. Zhu, Z.Q. Zhang, and H.J. Zhang, "Statistical Learning of Multi-View Face Detection," *In Proc. 7th European Conference on Computer Vision*, Copenhagen, Denmark, May 2002.
- [9] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian detection using wavelet templates," *In Computer Vision and Pattern Recognition*, pp. 193-199, 1997.
- [10] F. Crow, "Summed-area tables for texture mapping," *In Proceedings of SIGGRAPH*, Vol. 18, No. 3, pp. 207-212, 1984.
- [11] Y. Freund and R.E. Schapire, "A Short Introduction to Boosting," *Journal of Japanese Society for Artificial Intelligence*, Vol. 14, No.

- 5, pp. 771-780, 1999.
- [12] Yoav Freund and Robert E. Schapire, "A decision theoretic generalization of on-line learning and an application to boosting," *In Computational Learning Theory: Eurocolt'95*, pp. 23-37, 1997.
- [13] Don Murray and Anup Basu, "Motion Tracking with an Active Camera," *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 5, 1994.
- [14] G. R. Bradski, "Computer Vision Face Tracking as a Component of a Perceptual User Interface," *IEEE Workshop on Application Computer Vision*, Princeton, pp. 214-219, 1998.
- [15] H. Schneiderman and T. Kandade, "A statistical method for 3D object detection applied to faces and cars," *In International Conference on Computer Vision*, 2000.

이 우 주



2002년 전남대학교 컴퓨터정보통신공학과 졸업(공학사)
 2004년 전남대학교 대학원 컴퓨터정보통신공학과 (공학석사)

2004년 ~ 현재 : 전남대학교 대학원 컴퓨터정보통신공학과 박사과정
 관심분야 : 영상처리, 신경회로망, 패턴인식, 얼굴인식, 인공지능



김 진 철

2000년 금오공과대학교 산업공학과 졸업(공학사)
 2003년 전남대학교 대학원 컴퓨터정보통신공학과 (공학석사)
 2003년 ~ 현재 전남대학교 대학원 컴퓨터정보통신공학과

박사과정

관심분야 : 영상처리, 신경회로망, 패턴인식, 인공지능, 얼굴인식



이 배 호

1978년 한양대학교 전자공학과 졸업(공학사)
 1980년 KAIST 전기 및 전자공학과 졸업(공학석사)
 1993년 University of Missouri, 전기 및 컴퓨터공학과 졸업(공학박사)

1980년 ~ 1983년 국방과학연구소 연구원
 1993년 ~ 현재 전남대학교 전자컴퓨터공학부 교수
 관심분야 : 멀티미디어 통신, 컴퓨터비전, 영상처리, 인공지능, 정보검색