

계층적 분류기를 이용한 실시간 얼굴 검출 및 추적

김수희*, 양창호, 이배호
*전남대학교 컴퓨터정보통신공학과
e-mail:advance7@naver.com

Real-time face detection and tracking using hierarchical classifier

Su-Hui Kim*, Chang-Ho Yang, Bae-Ho Lee
*Dept. of Computer Engineering, Chonnam National University

요 약

본 논문은 계층적 분류기를 제안하여 실시간으로 얼굴 영역을 검출하고, PT(pan-tilt) 카메라를 통해 동적으로 얼굴을 추적할 수 있는 강인한 추적 알고리즘을 구현하고자 한다. 제안된 알고리즘은 분류기 학습, 실시간 얼굴 영역 검출, 추적의 세 단계로 구성된다. 분류기 학습은 AdaBoost 알고리즘을 이용하여, 독특한 얼굴 특징을 추출하는 계층적 분류기를 생성한다. 계층적 분류기는 높은 정확도를 가진 분류기들이 단계적으로 결합됨으로써 우수한 검출 성능으로 수행된다. 실시간 얼굴 영역 검출은 생성된 계층적 분류기를 통해, 빠르고 효율적으로 얼굴 영역을 찾아낸다. 추적은 PT 카메라를 통해 동적으로 검출 영역을 확장시키며, 이전 단계에서 추출된 얼굴 영역의 위치 정보를 이용하여 수행한다. 제안된 알고리즘은 계산의 효율성과 검출 성능을 동시에 증가시키며, 얼굴 검출 수행은 2초당 약 15프레임을 실시간으로 처리한다.

1. 서론

실시간 얼굴 검출과 추적 기술은 HCI(Human Computer Interaction), 영상 보안 및 감시 시스템, 얼굴 인식 시스템 등의 응용 분야에서 중추적인 역할을 한다. 그러나 얼굴 영상은 시점 변화, 기울기, 표정, 피부 색상, 머리 모양, 안경 등으로 인해 다양한 변화가 존재한다. 또한, 복잡한 배경, 조명의 영향, 영상의 질, 색상의 왜곡 등 외적인 요인들이 존재하기 때문에, 영상 내에서 얼굴 영역을 완벽하게 분리하고 추적하는데 현실적으로 한계가 있다[2].

최근의 얼굴 검출 기술은 시점 변화[3], 자세 및 조명 변화 등에 존재하는 다양성을 수용하도록 복잡한 알고리즘을 사용하며, 이는 시스템의 처리 속도에 큰 영향을 미친다. 따라서 검출 알고리즘의 처리 속도를 향상시키는 것은 실질적인 응용 시스템을 개발할 때 고려해야 할 중요한 문제이다[4].

본 논문은 계층적 분류기를 이용하여, 계산의 효

율성과 검출 성능을 동시에 만족시키는 강인한 추적 알고리즘을 제안한다. 제안된 알고리즘은 분류기 학습, 실시간 얼굴 영역 검출, 추적의 세 단계로 구성되며, 이에 대한 각 구성 모듈을 설명하고자 한다.

본 논문의 구성은 다음과 같다. 2장은 분류기를 학습하는 과정을 설명하고, 3장은 생성된 계층적 분류기를 통해, 실시간으로 얼굴 영역을 검출하는 방법을 설명한다. 4장은 추출된 얼굴 영역의 위치 정보를 이용하여, PT(pan-tilt) 카메라를 통해 얼굴을 추적하는 방법을 설명한다. 5장은 실험 결과를 분석하고, 6장은 결론을 제시한다.

2. 제안된 계층적 분류기 학습

2.1 사각형 마스크

본 연구는 고유한 얼굴 특징을 추출할 수 있는 사각형 마스크를 이용한 계층적 분류기를 제안한다. 사각형 마스크는 특정 영역에 대한 구조적인 정보를

제공하며, 얼굴 패턴을 분류하는데 결정적인 역할을 하는 특징을 추출한다.

사각형 마스크는 그림 1과 같이 8개의 사각형 형태로 구성되며, Haar 웨이브렛 함수와 유사한 구조를 가진다. 본 연구는 제안된 사각형 마스크를 이용하여, 입력 패턴에 대한 방대한 특징 집합을 생성한다[5]. 각 특징은 입력 패턴 내의 다양한 위치와 크기로 존재하게 된다.

사각형 마스크 값(RecF)의 계산은 식 (1)과 같이 정의되며, 특징 값의 계산에서 요구되는 방대한 연산량을 줄인다. 사각형 마스크 값(RecF)은 전체 사각형 영역의 픽셀 값(RecS_w)의 평균과 점정색 사각형 영역의 픽셀 값(RecS_b)의 평균에 대한 차 연산으로 구해진다.

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

식 (1)에서 m 은 RS_w 의 픽셀 수와 n 은 RS_b 의 픽셀 수이며, 사각형 영역인 $RecS_w$ 와 $RecS_b$ 는 SAT (Summed-Area Tables)를 이용하여 계산된 사각형 영역의 픽셀 값이다[6]. SAT는 입력 영상의 원점(0, 0)에서부터 특정한 좌표 영역(x, y)에 대해, 모든 픽셀 값을 더한 값을 가지는 테이블이다.

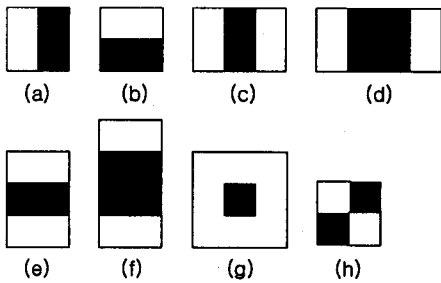


그림 1. 사각형 마스크

2.2 계층적 분류기 학습

본 논문은 AdaBoost 알고리즘을 이용하여 계층적 분류기를 학습한다. AdaBoost 알고리즘은 약한 분류기 (Weak classifier)를 결합함으로써, 최종적으로 높은 검출 성능을 가진 앙상블 분류기(Ensemble classifier)를 생성한다[1].

AdaBoost 알고리즘은 표 1과 같다. 분류기는 얼굴과 비얼굴 패턴으로 구성된 훈련 집합과 사각형 마스크 집합을 이용하여 학습된다. 입력 벡터 X 는 특징 집합이며, x_i 는 하나의 사각형 마스크이다. 출

력 Y 는 +1과 -1 값을 가지며, 각각 얼굴과 비얼굴을 나타낸다.

분류기 학습은 다음과 같이 진행된다. 약한 분류기 h_t 가 하나의 사각형 마스크 x_i 를 이용하여 훈련 패턴을 분류한다. 이때, 잘못 분류된 훈련 패턴은 가중치 $D_t(i)$ 를 증가시키고, 옳게 분류된 훈련 패턴은 가중치 $D_t(i)$ 를 감소시키는 과정을 반복한다. 따라서 분류기는 훈련 패턴의 변경된 가중치를 고려하여, 최소 에러율을 가지도록 하는 사각형 마스크를 선택한다. 이와 같이 학습된 분류기는 얼굴 패턴의 특징을 추출하는데 결정적인 역할을 하는 사각형 마스크를 선택한다.

약한 분류기 h_t 는 하나의 사각형 마스크 x_i 으로 분류하며, 최종적으로 생성된 앙상블 분류기 $H(x)$ 는 t 개의 약한 분류기가 결합된 형태이다. 즉, 앙상블 분류기는 t 개의 사각형 마스크가 선형적으로 결합된 형태이다. 앙상블 분류기는 약한 분류기의 수를 증가할수록 에러율이 영점에 지수적으로 근접한다[1]. 앙상블 분류기는 식 (2)와 같이 정의된다.

$$H(x) = \text{sign} \left(\sum_{i=1}^T \alpha_i h_i(x) \right) \quad (2)$$

제안된 사각형 마스크 기반 분류기는 높은 정확도

표 1. AdaBoost 알고리즘

<p>Given : $(x_1, y_1), \dots, (x_m, y_m)$ where $x_i \in X, y_i \in Y = \{-1, +1\}$</p> <p>Initialize $D_1(i) = 1/m$.</p> <p>For $t = 1, \dots, T$:</p> <ul style="list-style-type: none"> • Train weak learner using distribution D_t. • Get weak hypothesis $h_t: X \rightarrow \{-1, +1\}$ with error. • Choose $\alpha_t = \frac{1}{2} \ln \left(\frac{1-\epsilon_t}{\epsilon_t} \right)$. • Update: $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t} & \text{if } h_t(x_i) \neq y_i \end{cases}$ $= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$ <p>where Z_t is a normalization factor (chosen so that will be a distribution).</p> <p>Output the final hypothesis :</p> $H(x) = \text{sign} \left(\sum_{i=1}^T \alpha_i h_i(x) \right)$
--

를 가진 앙상블 분류기가 계층적으로 구성되며, 중요한 얼굴 패턴은 다음 레벨에 반복적으로 적용함으로써 우수한 검출 성능을 가진다.

3. 얼굴 영역 검출 알고리즘

제안된 알고리즘은 학습된 계층적 분류기를 통해, 빠르고 효율적으로 얼굴 영역을 찾아낸다. 검출 알고리즘은 전 처리, 실시간 얼굴 영역 검출의 두 단계로 구성된다.

3.1 전 처리

전 처리는 입력 영상에 대한 다중 해상도 피라미드를 구성함으로써, 다양한 위치와 크기의 얼굴 영역을 검출할 수 있다. 각 레벨의 영상은 5x5 가우시안 필터를 적용하여 잡음을 제거한 후, 20x20 크기로 일반화하여 입력 패턴을 생성한다. 입력 패턴은 대비 확장 알고리즘을 통해 색상을 보정한 후, 이에 대한 SAT를 생성하여 특징 값을 효율적으로 계산한다.

3.2 실시간 얼굴 영역 검출

실시간 얼굴 영역 검출은 사각형 마스크로 구성된 계층적 분류기를 이용하여, 전 처리 과정에서 생성된 입력 패턴을 분류한다. 계층적 분류기는 앙상블 분류기 $H(x)$ 가 계층적으로 구성된 형태이며, 앙상블 분류기 $H(x)$ 는 약한 분류기 h_i 의 선형적인 결합으로 구성된다. 즉, 앙상블 분류기는 사각형 마스크의 결합으로 구성되며, 얼굴 패턴의 특징을 추출하는데 결정적인 역할을 한다.

따라서, 제안된 분류기는 학습을 통해 설정된 크기와 위치의 사각형 마스크 값(RecF)을 계산한 후, 임계치를 만족하면 얼굴 패턴으로 분류한다. 최종적으로 분류기는 얼굴 패턴의 시퀀스를 출력한다.

4. 얼굴 추적 알고리즘

추적 알고리즘은 PT 카메라를 통해 동적으로 검출 영역을 확장시키며, 실시간으로 추출된 얼굴 영역의 위치 정보만을 이용함으로써 시스템의 처리 속도를 증가시킨다. 제안된 추적 알고리즘은 그림 2와 같으며, 분류기 학습, 실시간 얼굴 검출, 얼굴 추적의 세 단계로 구성된다.

분류기 학습 단계는 오프라인에서 진행되며, 훈련 집합과 특징 집합을 이용한다. 훈련 집합은 얼굴과 비얼굴 패턴으로 구성된 데이터베이스이며, 특징 집합은 8개의 사각형 특징을 스케일링(Scaling)과 쉬프팅(Shifting)하여 생성된 특징으로 구성된다. 학습 과정은 AdaBoost 알고리즘으로 진행되며, 얼굴과 비얼굴 패턴을 분류하는데 결정적인 역할을 하는 특징을 선택한다. 따라서, 앙상블 분류기는 특징들이 선형적으로 결합된 형태이며, 최종적으로 생성된 계층적 분류기는 앙상블 분류기가 단계적으로 결합된 형태이다.

얼굴 검출 단계는 생성된 계층적 분류기를 이용하여, 실시간으로 얼굴 영역을 찾아낸다. 계층적 분류기는 특징들을 이용하여, 입력 패턴에 대한 특징 값을 계산한다. 따라서, 입력 패턴은 계산된 특징 값이 임계값을 만족하면, 얼굴 영역으로 분류된다. 마지막

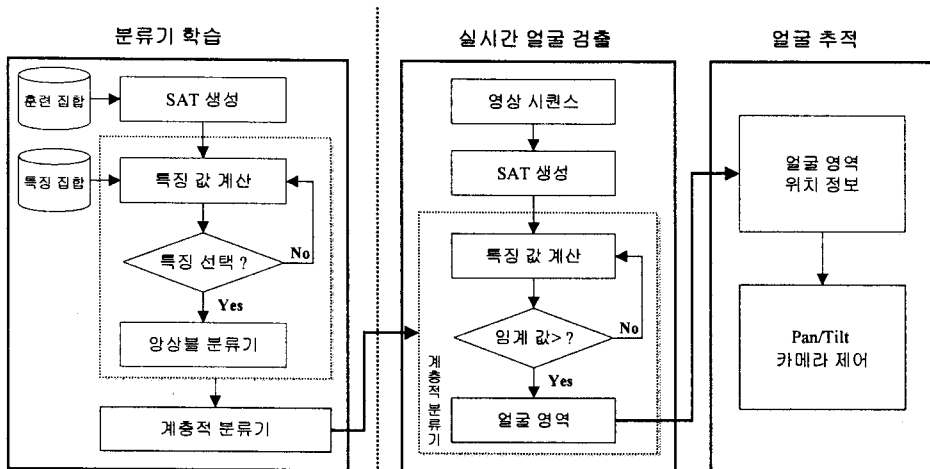


그림 2. 전체 시스템의 구성도

단계인 얼굴 추적 단계는 얼굴 영역의 위치 정보를 하여, 카메라의 Pan/Tilt를 제어한다.

5. 실험 결과

제안된 계층적 분류기는 MIT CBCL 데이터 집합으로 학습된다. 본 연구는 2,429개의 정면 얼굴 패턴과 4,548개의 비얼굴 패턴으로 훈련 집합을 구성한다. 훈련 패턴은 20x20 크기로 스케일링하여, 분류기 학습의 입력 패턴으로 사용한다.

특징 집합은 54,189개의 사각형 마스크로 구성된다. 이 특징 집합은 제안된 8개의 사각형 마스크를 이용하여, 20x20 크기의 영상 패턴에 대한 다양한 특징을 생성한다. 각 특징은 x방향과 y방향으로 각각 스케일링되며, 20x20 패턴 내의 다양한 위치에 존재하게 된다.

본 연구는 훈련 집합과 54,189개의 특징 집합으로 분류기를 학습한다. 학습된 분류기는 12개의 레벨과 120개의 특징으로 구성된다. 실시간 얼굴 검출 및 추적 결과는 그림 3과 같다. 제안된 검출 알고리즘은 얼굴 영역에서 발생하는 다양성을 수용할 수 있으며, 계산 효율적으로 실시간 얼굴 검출 및 추적을 이행한다. 이는 실험을 통해 입증된다. 그림 3에서

(a)와 (b)는 강한 조명과 어둠 속에서 얼굴을 검출한 결과이다. (b)와 (c)는 옆면 얼굴과 기울어짐이 발생한 얼굴을 검출한 결과이다. (c)와 (d)는 다양한 크기의 얼굴 영역을 검출한 결과이다. 검출된 얼굴 영역은 PT 카메라를 통해 실시간으로 추적된다.

제안된 사각형 마스크 기반 분류기는 320x240 크기의 입력 영상에 대해, 2초당 약 15프레임을 실시간으로 처리하며, 다양한 위치와 크기의 얼굴 영역을 빠르고 효율적으로 검출 및 추적한다.

6. 결론

본 논문은 PT 카메라를 통해 동적으로 얼굴을 추적할 수 있는 강인한 추적 알고리즘을 제안한다.

제안된 알고리즘은 AdaBoost 알고리즘을 이용하여, 독특한 얼굴 특징을 추출하는 계층적 분류기를 생성한다. 계층적 분류기는 높은 정확도를 가진 분류기들이 단계적으로 결합됨으로써 우수한 검출 성능을 가진다. 실시간 얼굴 영역 검출은 생성된 계층적 분류기를 통해, 빠르고 효율적으로 얼굴 영역을 찾아낸다. 추적은 PT 카메라를 통해 동적으로 검출 영역을 확장시킴으로써, 사각 지대를 최소화한다. 또한, 이전 단계에서 추출된 얼굴 영역의 위치 정보만을 이용함으로써, 계산량 감소로 시스템의 처리 속도를 증가시킨다. 제안된 알고리즘은 계산의 효율성과 검출 성능을 동시에 만족시키며, 얼굴 검출 수행은 2초당 약 15프레임을 실시간으로 처리한다.

참고문헌

- [1] Y. Freund, R.E. Schapire, "A Short Introduction to Boosting", *Journal of Japanese Society for Artificial Intelligence*, 14(5), pp. 771-780, 1999.
- [2] R. Chellappa, C.L. Wilson, S. Sirohey, "Human and machine recognition of faces: a survey", *Proceedings of the IEEE*, Vol. 83 Issue: 5, pp. 705-741, May 1995.
- [3] 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.
- [4] B. Heisele, T. Serre & S. Prentice, "Real-Time Face Detection".
- [5] 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.
- [6] F. Crow, "Summed-area tables for texture mapping", In *Proceedings of SIGGRAPH*, pp. 207-212, 1984.

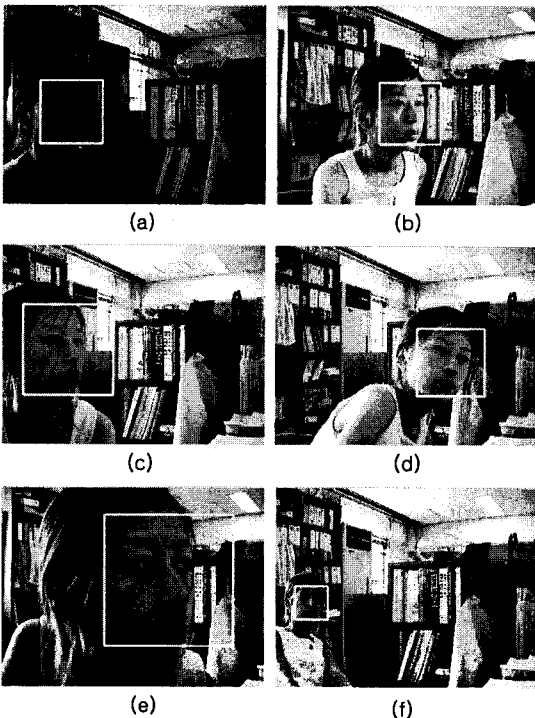


그림 3. 실시간 얼굴 검출 및 추적 결과