

단계형 구조와 SVDD를 이용한 실시간 얼굴 탐지 시스템

송지영⁰, 이한성, 임영희, 박대희
 고려대학교 컴퓨터 정보학과
 {songjiy⁰, mohan, yheem, dhpark}@korea.ac.kr

Real-time Face Detection System using Cascade structure and SVDD

Jiyoung Song⁰, Hansung Lee, Younghee Im, Daihee Park
 Dept. of Computer & Information Science, Korea University

요 약

본 논문에서는 점증적 분류 성능을 갖는 단계형(cascade) 분류기를 이용한 새로운 실시간 얼굴 탐지 시스템을 제안하고자 한다. 제안된 시스템의 첫 단계는 전처리 단계로써 매우 빠른 속도를 갖는 새로운 피부색 탐지기를 이용하여 탐색 공간을 대폭 축소하고, 두 번째 단계에서는 빠른 분류가 가능한 유사-하(Haar-like) 특징을 이용한 단계형 분류기를 배치하여 빠른 속도로 후보 얼굴을 검출한다. 마지막 단계에서는 탐지율을 높이기 위해 단일 클래스 SVM인 SVDD를 분류기로 사용하였으며, 실험을 통하여 제안된 시스템의 우수성을 보인다.

1. 서 론

최근 주목받고 있는 생체인식 분야 중 얼굴을 이용한 인식 방법은 사용자의 편의성 측면에서 가장 탁월한 시스템으로 각광받고 있다. 얼굴 인식의 성능을 높이고 실시간 얼굴 인식을 위해서는 빠르고 정확한 얼굴 탐지(face detection) 기술의 개발이 절실히 요구된다[1].

Zuo[2]는 컬러 정보를 이용한 점증적 분류 능력을 갖는 단계형 구조를 제안하여 매우 효율적인 얼굴 탐지를 가능하게 하였다. Zuo의 얼굴 탐지기는 1단계에 피부색 탐지기(skin-color detector)를, 2단계에 구조 기반(structure-based) 탐지기를, 3단계에서는 인공 신경망을 배치하여 초반부에서는 분류율은 다소 떨어지지만 빠른 속도를 낼 수 있는 분류기를 사용하고, 후반부로 갈수록 속도는 다소 느리지만 분류율이 뛰어난 분류기를 사용하는 전략을 사용하였다.

본 논문에서도 이러한 전략을 따르지만 Zuo의 구조와는 차별적으로, 새롭게 개선한 피부색 탐지기를 하나의 분류기가 아닌 전처리기로 따로 분리하여 좀 더 효율적인 얼굴 탐지를 가능하게 하였고, 2단계에서는 viola[3]가 제안한 인테그럴 이미지(Integral Image)와 유사-하(Haar-like) 특징을 이용한 매우 빠른 단계형(cascade) 구조의 분류기를 배치하여 고속의 탐지가 가능하게 하였다. 마지막 단계에서는 SVM(Support Vector Machine)의 한 종류인 SVDD(Support Vector Data Description)를 이용하여 우수한 탐지율을 갖도록 설계하였다.

Zuo의 피부색 탐지기는 주어진 입력 이미지로부터 모든 위치의 모든 크기의 부 윈도우(sub-window)에 대하여 피부색 탐지기를 통과시켜 얼굴 후보인지 아닌지를 결정한다. 그러나 본 논문에서는 피부색 탐지기를 탐색 공간의 축소 즉, 부 윈도우의 개수를 줄이는데 사용하여 고속의 탐지가 가능하도록 해준다. 또한 2단계에서는 정확한 탐지가 아니라 후보 얼굴의 검출이 목적이므로, 현재 가장 빠른 탐지기로 알려진 단계형 분류기를 사용하였다. 기존의 얼굴 탐지 모델을 위한 SVM의 경우, 얼굴/비얼굴로 분류하는 이진 SVM이 사용되었으나 본 논문에서는 단일 SVM인 SVDD를 사용한다. 이는 얼굴 탐지 문제를 비정상 상태 탐지(abnormality detection) 문제로 해석한 첫 번째 시도로써 그 의미가 매우 크다고 할 수 있다. 실험을 통하

여 제안된 방법론이 기존의 연구에 비하여 속도와 분류율면에서 모두 우수함을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 SVDD에 대하여 살펴보고, 3장에서는 새롭게 제안한 실시간 얼굴 탐지 시스템을 상세히 기술한다. 4장에서는 실험 및 결과를 서술하고, 5장에서는 결론 및 향후 연구 과제를 살펴본다.

2. SVDD(Support Vector Data Description)

단일 클래스 분류 문제(one-class classification problem)는 분류 대상이 되는 하나의 학습 클래스(target class)에 속한 데이터만을 이용하여 시스템의 학습(training)을 수행한다. 본 논문에서는 단일 클래스 분류 알고리즘 중 가장 대표적인 SVDD 방법[4][5]을 따르고자 한다.

d-차원 입력공간에 존재하는 N개의 데이터로 구성되는 학습 데이터의 집합 $D = \{x_i \mid i = 1, \dots, N\}$ 에 대해서 R^d 위에 정의되는 중심이 a이고 반경이 R인 원형체 B를 이용하여 학습 클래스의 존재 영역을 표현하고자 한다. 이 경우 각 학습 데이터 x_i 와 중심 a 사이의 거리가 R을 초과하는 경우에는 적절한 벌점(penalty)을 부과하는 전략을 사용하면 되며 다음과 같은 최적화 문제로 정의 된다.

$$\begin{aligned} \min L_0(R^2, a, \xi) &= R^2 + C \sum_{i=1}^N \xi_i \\ \text{s.t.} \quad \|x_i - a\|^2 &\leq R^2 + \xi_i, \xi_i \geq 0, \forall i. \end{aligned} \quad (1)$$

여기에서, ξ_i 는 i-번째 학습 데이터 x_i 가 원형체에서 벗어나는 벌점이며, C는 반지름과 벌점항의 상대적 중요성을 조정하는 상수(trade-off constant)이다. 라그랑제 함수(Lagrange function) L을 이용하여 위의 문제에 대응하는 쌍대 문제(dual problem)를 구한다. 쌍대 문제는 다음과 같은 QP(quadratic programming) 문제로 정리된다[5].

$$\min_{\alpha} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \langle x_i, x_j \rangle - \sum_{i=1}^N \alpha_i \langle x_i, x_i \rangle \quad (2)$$

$$s.t. \sum_{i=1}^N \alpha_i = 1, \alpha_i \in [0, C], \forall_i$$

학습이 끝난 후 적용 과정에서 결정함수는 다음과 같이 정의된다[4][5].

$$\begin{aligned} f(x) &= R^2 - \|x - a\|^2 \\ &= R^2 - (\langle x, x \rangle - 2 \sum_{i=1}^N \alpha_i \langle x_i, x \rangle \\ &\quad + \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \langle x_i, x_j \rangle) \geq 0 \end{aligned} \quad (3)$$

입력 공간 위에서 정의되는 원형체는 매우 간단한 형태의 영역만을 나타낼 수 있으므로 Mercer 커널 $k()$ 을 통하여 정의되는 고차원의 커널 특징 공간 F 위에서 정의되는 원형체를 사용하는 방향으로 확장될 수 있다. 가우시안 커널을 사용할 경우 최적화 문제 식(2)은 다음과 같이 보다 더 간단한 형태로 단순화될 수 있다[5].

$$\min_a \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j (x_i, x_j) \quad (4)$$

$$s.t. \sum_{i=1}^N \alpha_i = 1, \alpha_i \in [0, C], \forall_i$$

이 경우 커널 함수 도입후의 결정함수 식(3)은 아래의 식으로 대체된다.

$$\begin{aligned} f(x) &= R^2 - \|\phi(x) - a\|^2 \\ &= R^2 - (1 - 2 \sum_{i=1}^N \alpha_i k(x_i, x) \\ &\quad + \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j k(x_i, x_j)) \geq 0 \end{aligned} \quad (5)$$

3. 실시간 얼굴 탐지 시스템

3.1 시스템 구성도

본 논문에서는 그림 1과 같이 피부색 탐지기를 하나의 분류기가 아닌 전처리기로 따로 분리하여 좀 더 효율적인 얼굴 탐지를 가능하게 하였고, 2단계에서는 단계형(cascade) 구조의 분류기를 배치하여 고속의 탐지가 가능하게 하였다. 마지막 단계에서는 단일클래스 SVM인 SVDD를 배치하여 우수한 탐지율을 갖도록 설계하였다.

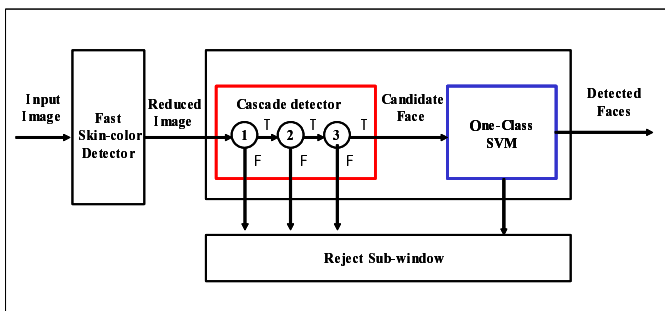


그림 1. 실시간 얼굴 탐지 시스템 구성도

3.2 전처리 역할의 피부색 탐지기

피부색 탐지기는 단독으로 얼굴 탐지에 사용하기에는 분류율 면에서 성능이 많이 떨어지지만, 다른 탐지기와 결합하여 사용하면 속도를 높일 수 있는 매우 적절한 방법이다. Zuo[2]의 피부색 탐지기는 주어진 입력 이미지로부터 모든 위치의 모든 크기의 부-윈도우(sub-window)에 대하여 피부색 탐지기를 통과시켜 얼굴 후보인지 아닌지를 결정한다. 그러나 본 논문에서는 모든 부-윈도우를 피부색 탐지기에 통과시키는 것이 아니라 입력 이미지의 크기를 줄이는 전처리기로 사용함으로써 고속의 탐지가 가능하도록 해준다. 즉, 본 논문에서 사용될 피부색 탐지기의 목적은 빠른 시간에 대략의 얼굴 후보 영역을 찾는 것이므로, 가장 계산량이 적고, 구현하기 쉬운 RGB 칼라 스페이스와 명시적 피부색 정의 모델을 이용하여 피부색 탐지기를 설계하였다. 이는 기존의 명시적 피부색 정의 모델의 계산 시간을 50% 이상 감소시킨 것으로 다음과 같이 유도된다.

$$\begin{aligned} (R, G, B) \text{ is classified as skin if :} \\ R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and} \\ R - G > 15 \text{ and } R > B \end{aligned} \quad (6)$$

그림 2는 식(6)을 이용하여 입력 이미지(a)에서 찾은 얼굴 후보 영역(b)이다.

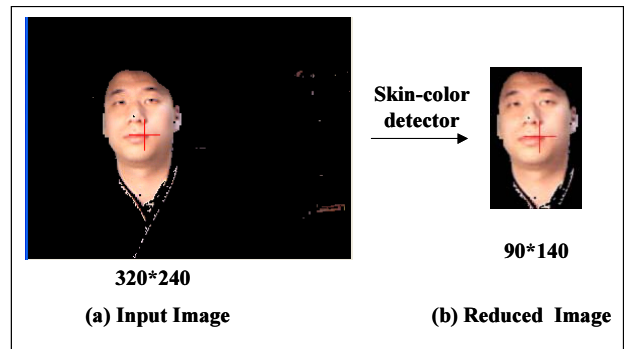


그림 2 피부색 탐지기를 통과후 축소된 이미지

3.3 고속의 단계형 분류기

2단계에서는 피부색 탐지기로부터 축소된 이미지를 입력으로 받아 후보 얼굴에 해당하는 부-윈도우를 SVDD의 입력으로 넘겨준다. 단계형 분류기 내부의 각각의 분류기(그림 1에서 ①②③으로 표시)는 분류율은 다소 떨어지지만 속도는 매우 빠른 약분류기(weak classifier)들이다. 이 분류기들은 비정상-오분류(false-negative)율은 크지만, 정상-오분류(false-positive)율은 거의 0에 가깝게 학습되므로 첫 번째 단계에서 입력으로 들어온 부-윈도우에 대해 매우 적은 계산량으로도 후보 얼굴(candidate face) 이미지인지 아닌지 결정할 수 있다. 즉, 1번 분류기에서 거절(reject)된 부 이미지는 다음 단계로 넘어가지 않고 바로 비 얼굴로 분류하고, 다시 새로운 부-윈도우를 분류하게 되므로 적은 양의 계산으로 비얼굴 부-윈도우를 걸러낼 수 있다[3].

3.4 분류율이 뛰어난 SVDD

탐지율을 높이고 비정상 오분류율을 높이기 위해서는 좋은 성능을 발휘하는 분류기가 필수적이다. Osuna[6]는 얼굴 탐지에 SVM을 도입하여 기존의 연구들에 비하여 탐지율이 뛰어난 모델을 제안하였다. 그러나 그림 3에서 보는 것처럼 정상 데이터

(얼굴)만 의미가 있고 비정상 데이터(비얼굴 즉, 배경)는 임의의 데이터를 사용할 경우 단일 클래스 분류기가 이진 클래스 분류기보다 유리하게 된다.

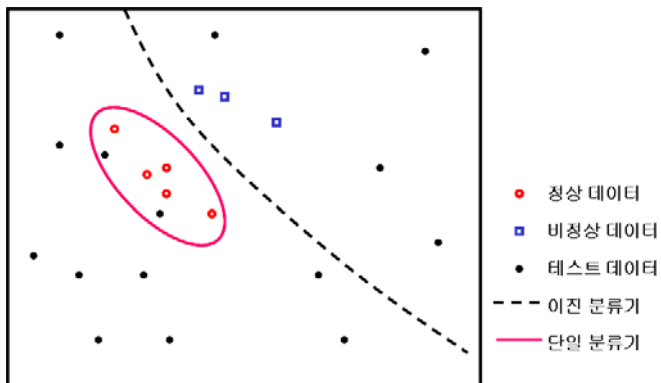


그림 3 이진 클래스 분류기와 단일 클래스 분류기의 경계 영역 비교

따라서 본 논문에서는 기존의 연구와는 차별적으로 단일 클래스 SVM인 SVDD를 도입하여 일반화 성능과 분류율이 뛰어난 뿐만 아니라 탐지 속도도 빠른 새로운 얼굴 탐지 시스템을 설계하였다. 단일 클래스 SVM은 다른 이름으로 비정상 상태 탐지라고도 불린다. 이 방법은 최근 들어 서포트 벡터 학습의 주요 응용 대상으로 주목받고 있다[4][5].

4. 실험결과

본 논문에서는 질 좋은 학습 데이터를 얻기 위하여 2004년부터 2005년 사이의 고려대학교 디지털 앨범에서 무작위 추출한 총 2000개의 얼굴 이미지를 이용하여 학습을 수행하였으며, 테스트 데이터는 인터넷상에서 취득한 이미지들을 사용하였다.

표 1의 결과에서 보듯이 제안된 시스템은 템플릿 매칭을 이용한 방법에 비하여 더 높은 탐지율을 보인다.

표 1. 얼굴 탐지율 비교

방법	얼굴 탐지율	정상 오분류율
템플릿 매칭	82.8 %	17.2 %
SVDD	99.3 %	0.7 %

웹카메라를 이용한 실시간 얼굴 탐지 실험에서는 기존의 연구결과[3](1초에 320×240 이미지 5장 정도를 처리)에 비하여 약 100% 향상된 초당 평균 10장 정도의 이미지를 처리할 수 있었다. 기존의 피부색 탐지기는 320×240 이미지에 대하여 부-원도우의 크기를 24×24부터 시작하여 25%씩 증가하면서 240×240까지의 부-원도우를 모든 위치에 대해 조사하는데, 이 경우 대략 180,000개 정도의 부-원도우가 나오게 된다. 비록 피부색 탐지기가 적은 양의 계산만으로 탐지를 할 수 있다고 하더라도, 전체 180,000개의 부 원도우를 다 비교하게 되면 상당히 많은 계산량이 필요하게 된다. 그러나 본 논문의 개선된 피부색 탐지기를 전처리기로 사용하게 되면 그림 2에서처럼 영역이 축소되어 부 원도의 개수가 10% 수준인 약 18,000개 정도로 줄어들게 돼 매우 빠른 탐지를 가능하다.

5. 결 론

본 논문에서는 점증적 분류 성능을 갖는 단계형 분류기를 이

용한 새로운 실시간 얼굴 탐지 시스템을 제안하였다. 제안된 시스템의 첫 번째 단계는 전처리 단계로써 매우 빠른 속도를 갖는 새로운 피부색 탐지기를 이용하여 탐색 공간을 대폭 축소한다. 새롭게 개선된 피부색 탐지기는 단순 연산수로 비교하면 기존의 피부색 탐지기 대비 54%(7/13)의 성능 향상을 가져 왔다. 두 번째 단계에서는 빠른 분류가 가능한 유사-하 특징을 이용한 단계형 분류기를 배치하여 빠른 속도로 비얼굴을 걸러 내고, 여기서 통과한 후보 얼굴들을 다음 단계의 입력으로 남겨 줌으로써 빠른 탐지가 가능하게 하였다. 마지막 단계에서 사용한 SVDD는 얼굴 탐지 분야를 이진 분류 문제가 아닌 비정상 상태 탐지 문제로 해석한 첫 번째 시도으로써 그 의미가 매우 크다고 할 수 있다. 실험을 통하여 제안된 시스템은 속도뿐만 아니라 분류율 면에서도 뛰어난 성능을 보였다. 향후 다각도 (multi-view) 얼굴 탐지 및 얼굴 추적(tracking)으로의 확장에 대한 연구가 요구된다.

참고 문헌

[1] M. H. Yang, D. Kriegman, N. Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 24, no. 1, pp. 34-58, 2002.
 [2] Zuo, F., With, P.H.N. de, "Fast Human Face Detection Using Successive Face Detectors with Incremental Detection Capability". SPIE Proc. of Image and Video Communications and Processing, IS&T/SPIE 15th Symp. on Electronic Imaging, pp. 831-841, 2003.
 [3] Paul Viola, Michael J. Jones, "Robust Real-time Object Detection", International Journal of Computer Vision, 2002.
 [4] David M.J. Tax and Robert P.W. Duin, "Uniform Object Generation for Optimizing One-class Classifiers", Journal of Machine Learning Research, Vol. 2, Issue 2, pp. 155-173, 2001.
 [5] 박주영, 임채환, "비정상 상태 탐지 문제를 위한 서포트 벡터 학습", 퍼지 및 지능시스템학회 논문지, vol. 13, no. 3, pp. 266-274, 2003.
 [6] E. Osuna, R. Freund, F. Girosi, "Training Support Vector Machines: an Application to Face Detection", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 130-136, 1997.