

SVM을 이용한 정지영상에서의 얼굴 검출에 관한 연구

김정동, 김기완, 김익훈, 김중규
성균관 대학교 전기 전자과

A Study on Face Detection for Still Images using Support Vector Machine

JeongDong Kim, KiWan kim, IkHoon Kim, JungKyu Kim
School of Information & Communication Engineering, Sungkyunkwan University

요 약

본 논문에서는 SVM에 기반 하여 template matching 기법을 도입한 얼굴 검출 알고리즘을 제안한다. 또한 얼굴의 기하학적 특징에 기반 해 얼굴일 확률이 낮은 경우 비 얼굴로 인식하는 방법으로 SVM의 적용 횟수를 줄여 검출속도를 향상 시켰다. 그리고 single response criterion을 적용하여 얼굴을 중복검출하거나 얼굴 영역 내에서 얼굴을 찾는 경우를 제거하여 속도를 개선 하였다.

1. 서론

멀티미디어가 발달하고 컴퓨터 기술이 향상됨에 따라 이제는 단일 기계만을 사용하는 것뿐만 아니라 여러 기계들을 통합하여 인간의 편의를 향상시키기 위한 홈 네트워크 등의 연구가 증가 하고 있다. 이러한 연구에는 사람의 감정 등을 파악하기 위해서 얼굴 인지 및 표정 인지와 같은 많은 기술들을 필요로 한다. 이러한 기술들을 적용하기 위해서는 먼저 정확하게 얼굴을 검출하는 기술이 선행되어야만 한다.

그레이 영상에서 얼굴 검출에 관한 연구는 다음과 같이 4가지방향으로 진행되고 있다. 첫째 전형적인 얼굴을 구성하는 규칙들의 집합을 이용하는 방법, 둘째 얼굴 내에서 불변하는 특징 점들을 이용하는 방법, 셋째 및 가지 표준적 패턴을 이용하는 방법, 넷째 training 이미지를 통해 분류기를 train시키는 방법 등으로 나뉜다.

본 논문에서는 네 번째 방법 중 하나인 SVM에 기반 하여 template matching 기법 등을 도입한 얼굴 검출 알고리즘을 제안한다. 본 논문의 구성은 다음과 같다. 먼저, 2장에서는 얼굴 검출 방법 및 속도 개선에 대해서 설명하고 3장에서는 실험결과를 보인 후 4장에서는 결론 및 향후 연구 과제를 제시한다.

2. 얼굴 검출

2.1 훈련 영상

훈련 영상의 선택은 얼굴 검출에서 가장 중요한 부분 중 하나이다. 왜냐하면 훈련 영상이 얼마나 얼굴의 특징과 비 얼굴의 특징을 잘 나타내 주냐에 따라 얼굴 검출의 성능의 차이가 크게 발생하기 때문이다. 본 논문에서 사용된 얼굴 훈련 영상은 최근 얼굴검출 및 인식기 개발에 많이 사용되는 CMU DB, BioID DB, Slovenian DB 를 사용하였다. 서로 다른 데이터베이스들을 고루 사용함으로써 각 데이터베이스의 단점들을 서로 보완하고 얼굴 검출 시 변화에 좀 더 강한 성능을 가지게 된다. 비 얼굴 훈련 영상은 얼굴이 존재하지 않는 영상에서 임의로 획득하거나 차후 프로그램 구현된 후 over detection 되는 영상들로 구성 하였다. 본 논문에 사용된 훈련 영상은 검출 알고리즘에 적합하게 추출된 얼굴 영상 1100개와 비 얼굴 영상 2100개를 사용하고 있다.



그림 1. Positive image



그림 2. Negative image

2.2 특징 추출 과정

일반적으로 사람들의 얼굴은 약간의 개인차가 존재하나 기하학적으로 거의 유사한 형태를 가지고 있다. 그래서 본 논문에서는 효율적인 비교를 위하여 얼굴 훈련 영상들의 평균값들을 구하여 평균 얼굴 영상을 얻었다. 이 평균 얼굴 영상과 훈련 영상사이의 상관관계를 특징으로 추출하여 얼굴 영역을 검출하기 위한 특징으로 사용하였다.

2.2.1. Template matching

사람의 얼굴은 눈, 코, 입과 같은 구성요소들로 구성되어있기 때문에 이 정보들을 이용하여 평균 얼굴 영상과 눈, 코, 전체 얼굴로 구성된 template간의 관계로 특징 값을 구했다. template matching은 cross correlation 계수 값으로 계산한다.

$$R(T, D) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (T[i][j] - \mu_T)(D[i][j] - \mu_D)}{M * N * \sigma_T * \sigma_D}$$

(1) (T : template, D : data)

Template은 눈 부분(그림 3), 코 부분, 전체 영역 이렇게 세부분으로 분류하여 그 특징들을 얻었으며 더 향상된 검출을 위하여 sobel_edge영상을 구하여 위와 같은 과정으로 3가지 특징을 더 추출하였다.

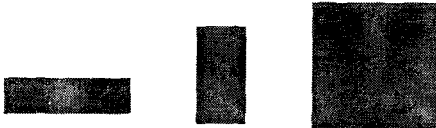


그림 3. Template(눈, 코, 얼굴 전체)

2.2.2. Horizontal, vertical amplitude projection

사람의 얼굴을 보면 눈 부분이 다른 부분에 비해 밝기 값이 어렵게 나오고 코 부분이나 볼 부분에서는 다른 부분에 비해서 밝기 값이 کم을 알 수 있다. 이러한 특징을 vertical, horizontal amplitude projection(그림 4)이 잘 보여주므로 이 값들을 이용하여 cross correlation 계수를 구해서 특징을 추출한다.



그림 4. Vertical, horizontal amplitude projection

2.3 SVM & kernel

SVM은 훈련 영상에서 얻은 특징 점들을 가장 효율적으로 분류하기 위한 방법이다. 효율적으로 분류하기 위해서 먼저 support vector라는 두 class의 경계점에 가까이 있는 특징 점들을 추출한 후 이 점들을 이용하여 두 class가 가장 큰 거리(margin)를 갖는 경계선(hyper plane)을 구하는 것이다. SVM을 간단히 식으로 나타내면 다음과 같다.

$$f(x, \alpha) = w \cdot \Phi(x) + b \quad (2)$$

여기서 α 는 경계선을 구할 때 각 support vector들이 미치는 영향에 비례하는 가중치이다. 그리고 w 는 경계 값의 기울기이고 b 는 threshold 값이다.

SVM값을 최적화 시키는 방법은 여러 가지가 있지만 본 논문에서는 Lagrange 방법을 사용하였다. α 값을 쉽게 구하기위해서 값의 범위에 대해서 제한을 두는데 본 논문에서는 Karush-Kuhn-Tucker(KKT)방법을 이용하여 제한을 두었다. 식(3)을 가장 크게 하는 α 값을 찾아서 w, b 값[식(6), (7)]을 구한다.(식(4), 식(5)은 α 값을 구할 때 제한 값들)

$$L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j K(x_i, x_j) \quad (3)$$

$$\alpha_i \geq 0 \quad i=1, l \quad (4)$$

$$\sum_{i=1}^l \alpha_i y_i = 0 \quad (5)$$

$$w = \sum_{i=1}^l \alpha_i y_i x_i \quad (6)$$

$$b = \frac{1}{N_{SV}} \left(\sum_{s=1}^{N_{SV}} \left(\frac{1}{y_s} - x_s^T w \right) \right) \quad (7)$$

위의 방법으로 경계 값을 구할시 특징들이 비선형

분포를 가지므로 효율적인 분류가 힘들다. 그래서 이러한 문제점을 극복하기 위해서 kernel 함수를 사용하는데 가장 많이 쓰이는 3가지 커널 함수 중 본 논문에서는 RBF 커널을 사용하였다. 커널 함수의 계수는 경험적으로 훈련 영상을 가장 효율적으로 나누는 기점을 기준으로 잡았다.

(λ : 커널함수 계수)

$$K(x, x_i) = e^{-1/2 * (\|x - x_i\|^2 / \lambda)} \quad (8)$$

2.4 속도 개선

2.4.1. 경계선을 구하는 support vector 개수 감소

얼굴 영역 검출 시 후보지역을 SVM으로 연산 시킬 때 support vector의 개수가 많을수록 많은 연산이 필요하게 되어 시간 소모가 증가하게 된다. 그러나 많은 support vector를 가지게 되면 성능 면에서는 missing할 확률이나 over detection이 감소하게 된다. 그러므로 속도와 성능간의 관계에서 성능 감소를 최소화하고 연산속도증가를 최적화하기 위해서 본 논문에서는 kernel 적용 시에 얼굴 훈련 영상과 비 얼굴 훈련 영상 간의 관계에 큰 영향을 미치지 않는 훈련 영상을 제거하여 속도 향상 및 적절한 성능을 유지시켰다.

2.4.2. 얼굴 특징을 이용한 연산 감소

입력영상의 모든 sub image에 대해 SVM을 적용할 경우 상당한 검출시간이 요구된다. 그러므로 얼굴일 확률이 희박한 sub image에 대해서는 SVM 연산을 생략하여 속도를 개선할 필요가 있다. 본 논문에서는 다음과 같은 방법으로 속도를 개선 시켰다.

2.4.2.1 눈과 미간의 밝기 차이 이용

얼굴에서 눈의 부분은 다른 지역에 비해 어두운 값을 가진다. 그러나 눈 사이와 코 부분까지의 영역은 눈에 비해 확연히 밝으므로 이 지역의 평균값이 양 눈의 평균값보다 크면 SVM을 적용 시키고 그렇지 않으면 생략하였다.



그림 5. Mask

2.4.2.2 Edge 차이 이용

얼굴 영역 중에서 눈 부분이 변화가 가장 크기 때문에 edge영상에서 보던 눈 영역이 가장 밝게 나온다. 그리고 볼 부분은 전체적으로 색변화가 크게 없기 때문에 edge영상에서 어둡게 나온다. 이것을 이용하여 앞에 경우와 동일하게 두 부분의 평균을 비교하여 만족시키지 않는 부분은 비 얼굴로 검출하였다. 단 두 영상에서 적용 부위의 크기는 훈련 영상에 적용 후 missing을 최소화하는 크기로 적용시켰다.

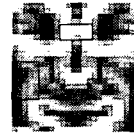


그림 6. Mask

2.4.3. Single response criterion

얼굴 검출 시에 실제 얼굴부근에서 여러 개의 검출결과를 보인다. 이러한 문제를 해결하고 검출속도 향상을 위해 본 논문에서는 single response criterion을 적용했다. 얼굴이라고 검출되는 지점이 나오게 되면 오른쪽 진행방향으로 7pixel 아래 방향으로 7pixel 범위에서 가장 얼굴일 확률이 높은 영역을 얼굴로 인식하고 검출된 영역 내에서는 얼굴이 발생할 확률이 없다고 보고 그 영역내의 값들은 SVM을 적용시키지 않으므로 속도를 향상시킨다.

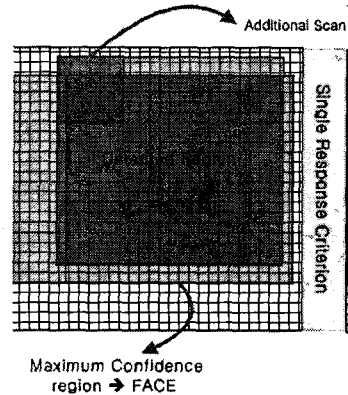


그림 7. Single response criterion



그림 8. 처리전



→ 처리후

3. 실험결과

본 논문의 실험 환경은 WINDOWS 2000 OS환경을 갖춘 CPU 1.6GHZ의 컴퓨터에서 Visual C++ 6.0로 구현하였다. 실험 영상은 얼굴 크기를 대략적으로 24*24로 고정하였고 정면 얼굴로만 훈련 시켰으므로 실험영상은 5%이내로 회전된 정면 얼굴로 제한하였다. 전체 250개 영상에서 600개의 얼굴을 가지고 실험하였고 결과는 $559/600 = 93.16\%$ 정도의 성능을 보였다.

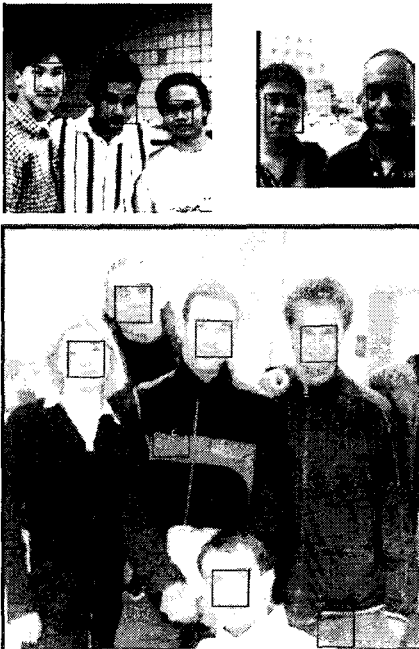


그림 9. 실험 결과

4. 결론 및 추후계획

본 논문에서는 SVM에 기반 하여 template matching 기법 등을 도입한 얼굴 검출 알고리즘을 제안한다. 그레이 영상에서 template matching등을 이용한 특징추출로 얼굴 검출의 효율을 개선하였고 얼굴 기하학적 특징 및 single response criterion을 적용하여 속도를 개선하였다. 제안한 알고리즘으로 얼굴 인식 등 얼굴 검출이 필요한 분야에 적용 가능할 것으로 기대된다. 향후 연구 과제로는 다차원에서의 얼굴 검출 및 동영상에서 실시간 검출과 다양한 각도에서의 검출에 대한 연구가 진행되어야 할 것이다.

[참고문헌]

- [1] C.Liu, "A Bayesian Discriminating Features Method for Face Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol25, no.6, pp.725~740, June2003
- [2] H.Ai, L.Liang, G.Xu, "Face Detection Based on Template Matching and Support Vector Machines", Proc. Int'l Conf. Image Processing, pp. 1006~1009, 2001 IEEE
- [3] R.Brunelli, T.Poggio, "Face Recognition: Features versus Templates", IEEE Trans. Pattern Analysis and Machine Intelligence, vol15, no.10, pp.1042~1052, October 1993
- [4] H.Ai, L.Ying, G.Xu, "A Subspace Approach to Face Detection with Support Vector Machines", Proc. Int'l Conf. Pattern Recognition, pp. 45~48, 2002 IEEE
- [5] H.A.Rowley, S.Baluja, T.Kanade, "Neural Network-Based Face Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.20, no.1, pp. 23~38, January 1998
- [6] E.Osuna, R.Freund, F.Girosi, "Training Support Vector Machines: an Application to Face Detection", Proc. CVPR'97, Puerto Rico, June, 1997
- [7] Vojislav Kecman. "Learning and Soft Computing", The MIT press, 2001