

효과적인 계단식 얼굴 검출을 위한 다중 특징 추출

소형준⁰, 남미영, 이필규
 인하대학교 컴퓨터 정보 공학과
 {mahaso⁰, rera}@im.inha.ac.kr, phrhee@inha.ac.kr

Multiple Feature Representation for Efficient Cascaded Face Detection

Hyoung Jun So⁰, Mi Young Nam, Phill Kyu Rhee
 Computer Science & Information Engineering, Inha University

요 약

본 논문은 복잡한 배경에서의 얼굴 검출에 있어서 다중 특징 추출 데이터로 학습한 계단식 분류기에 의한 방법을 제안한다. 얼굴 검출에서 얼굴의 패턴은 상당히 다양한 영상 표현으로 나타나기 때문에 하나의 특징 추출 방법은 사람의 얼굴을 모델링 하기에는 부족하다. 따라서 여기서는 얼굴의 전체적인 지역적인 특징을 나타내는 Sub-region과, 얼굴의 주파수 특성에 따라 좀 더 세밀하고 다양한 속성들을 나타내는 Haar 웨이블릿 변환을 이용하여 다중으로 특징을 추출하여 효과적인 모델링을 시도하였다. 특징을 추출한 얼굴과 비얼굴의 패턴(pattern)을 구분하기 위해서 패턴들의 통계적인 특성을 이용하여 각 추출방법에 맞게 학습된 Bayesian 분류기를 직렬로 연결하여 사용하였으며 비얼굴은 얼굴과 유사한 비얼굴(face-like nonface) 패턴들을 사용하여 모델링 하였다. 제안한 얼굴 검출 방식의 성능은 MIT-CMU 시험 영상들[1]을 이용하여 평가하였다. 그 결과 한 가지 특징 추출을 사용하는 것 보다 두 가지 특징 추출을 병행한 계단식 구성이 더 정확한 검출 결과를 나타내었다.

1. 서론

얼굴 검출은 얼굴 인식 시스템에서 얼굴을 인식하기 전에 이루어져야 하는 중요한 단계이다. 얼굴 검출은 여러 가지 방법으로 수행될 수 있다. 본 논문에서는 공통적으로 얼굴의 특정한 패턴을 학습 시키고, 새로 입력된 영상의 패턴과 학습된 통계적 특성을 비교하여 얼굴을 검출하는 외형 기반 모델을 사용하였다[2][3]. 얼굴 패턴을 sub-region과 Haar 웨이블릿 변환을 이용하여 모델링하고 Bayesian 분류기를 이용하여 단계적으로 얼굴의 검출을 시도한다. Sub-region은 얼굴 영상의 지역적인 특징을 나타낼 수 있다. 흑백 얼굴 영상에서의 명암 값들은 눈, 코, 볼 등과 같은 특정 부분에서 비슷한 값을 가지게 된다. 따라서 이런 영역들의 평균값을 이용하면 얼굴을 모델링 할 수 있다. 구체적인 내용은 2-1에서 설명하고 있다. Haar 웨이블릿 변환은 얼굴 영상의 다양한 주파수, 방향, 공간 특성을 나타낼 수 있다[4]. 그리고 주파수 영역으로 표현된 영상은 여러 주파수 성분 중에서 낮은 주파수 대역이 그 영상의 특성을 잘 나타내게 되므로

분별력이 높은 주파수 대역보다 크다. 이런 특성을 이용하여 영상을 나타내는 차원이 낮은 특징 벡터로 영상을 표현할 수 있다. 자세한 사항은 2-2에서 설명하고 있다. 위의 두 가지 특징 추출 방법의 효과적인 조합을 위해 두개의 분류기를 직렬로 연결하여 모든 검사영역의 패턴을 분류하게 된다. 기본적인 구조는 그림 1과 같다.

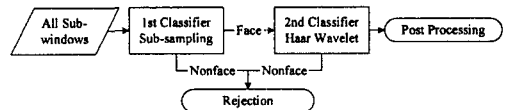


그림 1. 분류기의 계단식 구성

2절에서는 제안하는 특징 추출 방법에 대해서 설명하고, 3절에서는 사용된 Bayesian 분류기를 위한 통계학적 모델링 중 얼굴과 유사한 비얼굴의 모델링에 대해서 설명한다. 4절에서는 실험의 과정과 결과에 대해서 기술하고, 마지막으로 5절에서 결론을 맺는다.

2. 특징 추출

본 논문에서는 16×16 크기의 검출 영역을 기준으로 영상에 포함된 얼굴을 찾는다. 그러나 16×16의 검출 영역의 모든 점들을 그대로 이용해서 패턴으로 사용하면 복잡도가 증가하기 때문에 계산량이 많아지게 되고 상당히 많은 표본의 수가 필요하게 된다 [5]. 이러한 이유로 본 논문에서는 영상의 특징을 추출하여 차원을 감소 시켜서 사용한다[6]. 각 특징 추출 방법을 적용하기 전에 영상의 질을 개선하기 위해서 전처리 단계로 히스토그램 평활화(histogram equalization)를 적용한다.

2-1 Sub-region을 이용한 특징 추출

Sub-region을 이용하여 특징을 추출하면 얼굴 영상의 지역적인 특징을 나타낼 수 있다. 얼굴 영상에서 눈, 코, 입 등의 각 얼굴 요소들에 대한 점의 값들은 비슷하게 이루어져 있다. 따라서 이러한 특성을 이용하면 얼굴의 특징을 추출할 수 있다. 이러한 특성을 이용해서 16×16의 검출 영역에 대해서 54차원의 특징을 추출한다. 얼굴 영상을 2×2, 3×3, 4×4, 5×5로 줄여서 각 sub-region의 평균값들을 검사하는 영역의 특징 벡터로 사용한다(그림 2).

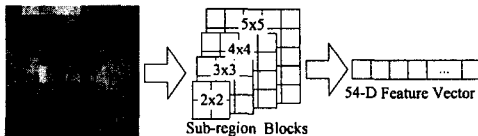


그림 2. 차원 특징 벡터의 추출

2-2 Haar 웨이블릿 변환을 이용한 특징 추출

Haar 웨이블릿 변환을 이용하면 얼굴 영상의 주파수 특성을 추출할 수 있기 때문에 얼굴이나 보행자를 검출하는데 효과적이다[7]. 주파수 영역에서의 영상의 특징은 고주파 성분보다 저주파 성분에서 크게 나타난다. 따라서 그림에서 보는 것처럼 16×16의 영상에서 3 단계의 Haar 웨이블릿 변환을 이용하여 특징을 추출하였다.(그림 3) 이는 검사 영역을 8×8의 영상으로 변환하여 Haar 웨이블릿 변환을 수행한 것과 같은 결과를 얻는다. 각 하위 대역은 위치에 따른 주파수 특성과 방향, 공간 특성 등을 갖는다[4]. 8×8 영상에서 LL3은 한 개의 실수 값이 되는데, 이는 전체 영상의 평균값을 가지게 된다. 이 변환을 통해서 64 차원의 특징 벡터를 얻을 수 있다. 이는 주어진 영역의 주파수 특성을 가지게 된다.

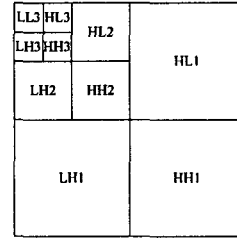


그림 3. Haar 웨이블릿 변환

3. 통계학적 모델링

본 논문에서 제안하는 방법으로 추출된 특징 벡터들은 Bayesian 분류기를 통해서 얼굴과 비얼굴로 분류한다. Bayesian 분류기는 학습 패턴들의 통계적인 특성을 분석하여 입력 패턴을 얼굴과 비얼굴로 분류하게 되므로 이를 위해서는 각 분류에 대한 통계학적 모델링이 필요하다[8]. 모델링 중 특히 비얼굴의 모델링이 어려운 문제인데 이는 얼굴이라는 대상을 제외한 모든 것이 비얼굴이기 때문에, 이 나머지 모든 것들의 특징들을 뽑아서 비얼굴의 모델을 만들기는 어려운 일이기 때문이다. 따라서 여기서는 얼굴 표본에 대한 Mahalanobis 거리를 기준으로 분류에 가장 영향을 줄 수 있는 얼굴과 유사한 비얼굴(face-like nonface)을 선택한 후 모델링 한다. 이런 비얼굴을 선택하는 방법은 아래의 과정으로 이루어진다.

1. m_f 개의 얼굴 표본 집합을 T_f 라 하고, T_f 의 공분산 행렬인 Σ_f 를 생성한다.
2. 모든 $f_i \in T_f$ 에 대해서 Σ_f 에 대한 Mahalanobis 거리인 d_{f_i} 를 계산한다.
3. θ 를 $\max\{d_{f_i} | 1 \leq i \leq m_f\}$ 로 설정한다.
4. m_n 개의 비얼굴 표본 집합을 T_n 라 하고, 각 표본, $n_j \in T_n$ 에 대해서 d_{n_j} 를 계산한다.
5. $d_{n_j} < \theta$ 인 n_j 를 얼굴과 유사한 비얼굴 표본 집합인 T_n 로 저장한다.
6. 비얼굴의 공분산 행렬을 Σ_n 라 하고, 이를 T_n 로부터 생성한다.

이런 방식으로 선택된 집합의 각 요소들은 얼굴과 일정한 Mahalanobis 거리, θ 안에 있는 얼굴과 유사한 비얼굴의 집합으로 구성되며 θ 는 실험적으로 정해야 하는 parameter이다

4. 실험 및 결과

Bayesian 분류기를 학습시키기 위해서 FERET 데이터베이스를 이용하였다. FERET 데이터베이스는 1개의

얼굴을 포함한 256x384 크기의 영상으로 구성되어 있다. 정면 얼굴을 두 눈 사이의 거리를 기준으로 동일한 크기로 축소되고, 똑바로 선 얼굴이 되도록 회전시켜서, 16x16의 크기로 정형화 하여 얼굴 표본 집합에 대한 영상으로 사용하였다 (그림 4).

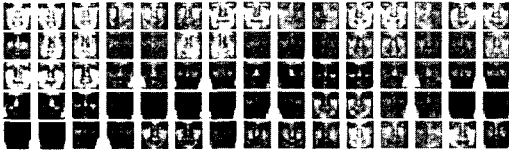


그림 4. 16x16의 얼굴 영상

비얼굴에 대한 영상은 얼굴을 전혀 포함하지 않은 자연 영상에서 8,297,936개의 추출하였으며 효과적인 학습을 위해 학습 데이터의 신뢰 가중치를 이용한 AdaBoost를 사용하였다[9]. 실험은 MIT-CMU 시험 영상들을 이용하였다. 검사 영역은 16x16 크기로 원본 영상의 왼쪽 위에서부터 수평과 수직 모두 1픽셀(pixel) 단위로 이동하면서 검사하게 된다. 검사 영역보다 큰 얼굴을 검출하기 위해서 원본 영상이 16x16 보다 작아질 때까지 1.2배씩 줄이면서 반복해서 검출하게 된다. 기울어진 정면 얼굴들을 검출하기 위해서 -20°에서 20°까지 5°단위로 회전시키면서 검출을 시도하였다. Bayesian 분류기가 이진화된 결과를 내는 것이 아니기 때문에 Haar 웨이블릿 변환을 사용한 두 번째 분류기의 τ 를 조정하면 아래 그림 5과 같은 ROC(Receiver Operating Characteristic) 곡선을 얻을 수 있다.

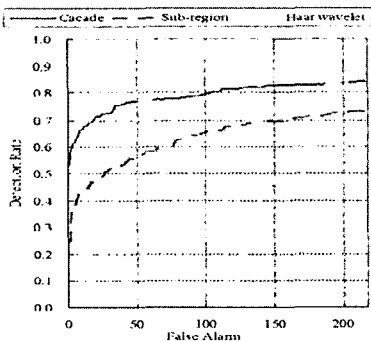


그림 5. ROC 곡선

그림 5에서 볼 수 있는 것과 같이 각각의 분류기를 사용한 것 보다 두 개의 분류기를 직렬로 사용한 것이 더 좋은 결과를 나타내었다. False alarm이 100일 때 79.7%의 검출률을 나타내었다. 검출 결과 영상의 일부를 그림 6에

나타내었다.



그림 6. 검출 결과 영상

5. 결론

본 논문에서는 얼굴 검출을 위해서 sub-region과 Haar 웨이블릿 변환을 이용한 두 가지 특징 추출 방법을 제안하였다. Bayesian 분류기를 사용하여 두 방법으로 추출된 특징 벡터들의 통계적인 특성을 이용해서 검출하였다. False alarm이 100일 때 79.7%의 검출률을 나타내었다.

참고문헌

- [1] H.A. Rowley and T. Kanade, "Neural Network-Based Face Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, January 1998.
- [2] Ming-Hsuan Yang, David J. Kriegman, and Narendra Ahuja, "Detecting Faces in Images: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, January 2002.
- [3] A. Yuille, P. Hallinan, and D. Cohen, "Feature Extraction from Faces using Deformable Templates," International Journal of Computer Vision, Vol. 8, No. 2, pp.99-111, 1992.
- [4] Henry Schneiderman and T.Kanade, "A Statistical Method for 3D Object Detection Applied to Faces and Cars," Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Vol.1, June 2000
- [5] Kah-Kay Sung and Tomaso Poggio, "Example-Based Learning for View-Based Human Face Detection," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, January 1998
- [6] Paul Viola and Michael Jones, "Rapid Object Detection using A Boosted Cascade of Simple Features," Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2001
- [7] Constantine P. Papageorgiou, Michael Oren, and Tomaso Poggio, "A General Framework for Object Detection," Proc. Conf. Computer Vision, pp. 555-562, Jan. 1998.
- [8] Chengjun Liu, "A Bayesian Discriminating Features Method for Face Detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No.6, JUNE 2003
- [9] R.E. Shapire, and Y.Singer, "Improving Boosting Algorithms using Confidence-rated Prediction." Machine Learning 37(30):pp. 297-336.
- [10] Richard O. Duda, Peter E. Hart, and David G. Stork, "Pattern Classification," Second Edition, A Wiley-Interscience Publication, 2000.