

신경망 기반 눈 영역 추정에 의한 실시간 얼굴 검출 기법

김 주 섭, 김 재 희
연세대학교 전기·전자공학과
전화 : 02-2123-4537 / Fax : 02-312-4584

Real-Time Face Detection by Estimating the Eye Region Using Neural Network

Jusub Kim and Jaihie Kim
Dept. of Electrical and Electronic Eng.
Yonsei University
E-mail : jskim@seraph.yonsei.ac.kr

Abstract

In this paper, we present a fast face detection algorithm by estimating the eye region using neural network. To implement a real time face detection system, it is necessary to reduce search space. We limit the search space just to a few pairs of eye candidates. For the selection of them, we first isolate possible eye regions in the fast and robust way by modified histogram equalization. The eye candidates are paired to form an eye pair and each of the eye pair is estimated how close it is to a true eye pair in two aspects : One is how similar the two eye candidates are in shape and the other is how close each of them is to a true eye image. A multi-layer perceptron neural network is used to find the eye candidate region's closeness to the true eye image. Just a few best candidates are then verified by eigenfaces. The experimental results show that this approach is fast and reliable. We achieved 94% detection rate with average 0.1 sec processing time in Pentium III PC in the experiment on 424 gray scale images from MIT, Yale, and Yonsei databases.

I. 서론

최근에 인터넷과 각종 정보망의 발달로 본격적인 정보화 시대로 돌입함에 따라 각종 생체인식기법을 이용한 정보 유출의 효과적인 차단에 관한 연구가 심도있게 진행되고 있다. 현존하는 여러 생체인식 방법 중 얼굴인식은 자신의 신체 일부를 인식장치에 직접 접촉시키지 않아도 되고 인식시 사용자의 특정행위를 요구하지 않는 비강제적인 방식으로 가장 자연스러운 생체인식 방법으로 주목받고 있다. 이러한 얼굴인식 시스템을 구현하기 위해서 주어진 영상에서 얼굴영역만을 분리해내는 일은 가장 중요하고도 필수적인 작업이다.

빠른 얼굴검출을 위해서는 탐색영역의 축소가 필수적이며 이를 위해서 기존의 논문에서 가장 많이 제안된 방법은 얼굴의 색상 정보를 이용하는 방법이다[1]. 하지만 이 방법은 조명에 민감하여 다양한 환경에 강인한 시스템을 구축하기 어려우며 대부분의 CCD카메라에서 사용하고 있는 계조도 영상(gray scale image)에서는 사용할 수 없는 문제가 있다. 이 밖에도 동영상일 경우에는 차영상을 이용하는 방법도 제안되었지만[2] 차영상만을 이용할 경우에 노이즈에 취약하고 한 명 이상의 얼굴은 검출하기 어려운 문제점이 있다. 이러한 문제를 해결하기 위하여 눈 영역 추정을 통한 얼굴검출 방법이 몇 번 제안되었으나[3][4][5] 눈 후보 영역 검출 방법에 단순한 이진화 또는 형태연산자

(Morphological Operator)를 사용하고 한 쌍의 눈인지 판단시 단순한 형태 정보 또는 가보웨이블렛(Gabor wavelet)을 사용함으로써 얼굴후보영역 추정의 정확성이 떨어지거나 시간이 오래 걸리는 단점이 있었다.

본 논문에서는 빠르면서 정확한 얼굴 후보 영역 추출을 위하여 변형된 히스토그램 평활화기법을 사용하여 눈 후보 영역을 분리하였으며 다층 신경망(multi-layer perceptron)을 이용하여 실제 눈일 가능성을 평가하였다. 고유얼굴기법을 이용하여 얼굴여부를 검증한 결과 제안한 방법이 빠르고 정확한 방법임을 보였다.

II. 탐색 영역 축소

2.1 눈 후보 영역 추출

얼굴의 주요 특징 중 눈 영역은 코나 입 영역보다 훨씬 조명이거나 크기 변화 등에 강인하게 검출할 수 있는 특징을 지닌다. 그것은 눈의 검은자위가 갖는 낮은 그레이 레벨 때문인데 이러한 성질로 인해 단순한 이진화 과정을 통해서도 눈 영역을 얼굴의 다른 부분과 분리해 낼 수 있다. 하지만 조명의 변화가 있을 경우 이진화 과정의 임계치가 매우 달라지는 문제가 발생한다. 이러한 문제를 해결하기 위해 Lam[4]등은 형태연산자(Morphological operator)를 사용하여 눈 영역을 분리해 냈으나 시간이 오래 걸리는 단점이 있었다.

본 논문에서는 변형된 히스토그램 평활화기법을 사용하여 빠른 시간내에 조명의 영향에 강인한 눈 후보 영역 추출 방법을 제안한다. 변형된 히스토그램 평활화기법을 사용하면 영상을 지역적으로 평활화하여 조명의 영향에 강인하게 지역적인 특징들을 찾아 낼 수 있다. 그런데 영상을 단순히 부분적으로 평활화할 경우 영상의 블록화 현상이 나타난다. 이러한 문제를 해결하기 위해 그림 2에서와 같이 부분 영역 R을 평활화할 때 그 주변 영역 N을 포함한 영역의 히스토그램 분포를 이용하여 평활화를 수행함으로써 블록화 현상을 제거하였다.

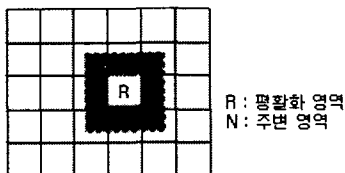


그림 2. 변형된 평활화 기법

변형된 히스토그램 평활화 기법을 적용하여 조명의

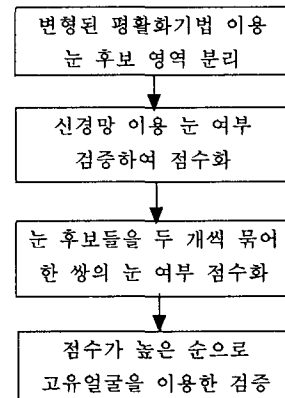


그림 1. 얼굴 검출 흐름도

영향을 최소화시킨 영상은 이진화 과정을 거친 뒤 영역 라벨링(labeling)을 수행한다. 라벨링 된 각 영역은 화소수, 가로/세로 비율등의 정보에 따라 눈의 검은자위가 될 수 없는 영역을 제거하며 아울러 눈과 눈썹이 붙은 경우 눈썹부분을 제거하는 등의 전처리 과정을 거치게 된다.

2.2 다층 신경망을 이용한 눈 후보 영역 검증

추출된 눈 후보 영역은 눈인지 여부를 검증하게 된다. 기존에 눈 영역 추정을 통하여 탐색 영역을 줄였던 유사한 방법을 제안한 논문에서는[3][4][5] 눈인지 여부를 검증하는 단계를 갖고 있지 않거나 단순한 원형 템플릿을 사용함으로써 많은 가짜 후보들을 가지고 다음 단계로 넘어가야 하는 문제를 가졌다. 이에 본 논문에서는 다층 신경망(multi-layer perceptron)을 이용하여 각 영역들의 눈 모양에의 근접성 정도를 점수로 매겨 사용하는 방법을 제안한다. 눈인지 아닌지 이분법으로 분류하지 않는 이유는 잘못하여 이 단계에서 눈 영역이 아니라고 판단내리는 경우 뒷 부분의 처리과정에 상관없이 얼굴 영역 추출에 실패하기 때문이다. 사용한 신경망은 그림 3과 같이 32-3-1의 구조를 갖는 다층 퍼셉트론으로서 눈일 경우 +1 눈이 아닐 경우 -1의 출력을 내도록 학습되었다. 학습영상은 MIT, Yale, Yonsei DB 에 있는 얼굴영상에서 얻어 졌으며 총 456 개의 눈영상과 눈이 아닌 영상을 가지고 역전파(Backpropagation) 학습법으로 학습시켰으며 학습영상에 가중치들이 과도하게 맞춰지는(Overfitting)것을 막기 위해 총 164개의 검증셋을 가지고 이른 정지(early stopping)방식으로 학습량을 조절하였다. 입력 신호의 차원을 32(8x4), 128(16x8), 512(32x16)로 변화시키고 은닉층의 뉴런수를 3, 7, 11, 15 로 변화 시켜

신경망 기반 눈 영역 추정에 의한 실시간 얼굴 검출 기법

가며 학습셋과 검증셋에 포함되지 않은 총 410개의 테스트 영상을 가지고 실험한 결과 은닉층의 뉴런수를 3개만 쓰고 입력벡터의 크기를 32차원으로 했을 때 84%의 검출률과 17%의 오검출률을 보이며 그 이상의 수를 갖는 뉴런과 입력벡터를 사용한 방법에 비해서 비슷하거나 오히려 나은 성능을 보여 최소 연결을 가지는 32-3-1의 구조를 선택하였다. 특히 은닉층의 뉴런수가 11개 이상이 되면 과도한 학습(Overfitting)의 영향으로 성능이 더 저하되었다.

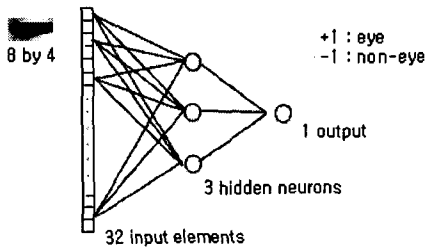


그림 3. 눈 후보 검증을 위한 다층 퍼셉트론 구조

2.3 얼굴 후보 영역 추출

눈 후보 영역이 추출되면 눈 후보 영역들은 두 영역이 한 쌍을 이루어 두 영역간의 유사도가 측정된다. 유사도도는 영상에 있는 한 얼굴의 두 눈은 그 형태나 모양이 비슷하다는 가정하에 화소수, 가로/세로 비, 계조도 평균값등에 의해 판단되며 여기에 앞 단계에서 구한 두 영역의 눈 모양 근접성 점수를 더하여 점수가 높은 순으로 얼굴 후보 영역의 검증 순서를 매기게 된다. 얼굴인지 여부의 검증의 대상이 되는 얼굴 후보 영역은 두 눈 영역의 위치를 기준으로 하여 일정 비율로 잘라낸 뒤 히스토그램 평활화하여 조명에 대한 정규화를 수행한 뒤 사용된다.

III. 얼굴 영역 검증

3.1 고유얼굴을 이용한 검증

얼굴 여부 검증에는 가장 널리 쓰이는 고유얼굴 기법을[6] 사용하였다. 대표적인 통계적 모델링 방법인 PCA(Principal Component Analysis)방법을 사용하여 구한 고유얼굴은 전체 영상공간상에서 얼굴만이 존재하는 낮은 차원의 영상공간을 구성하는 좌표계에 해당하며 원영상과 고유얼굴로 재합성된 영상간의 에너지 차이를 구하여 얼굴 여부를 검증하게 된다. 여기서 구



그림 4. 고유얼굴(eigenface)

한 에러는 잔여 에러(residual error)라 불리우며 후보 영상이 얼마나 얼굴모양에 적합한지를 나타내는 지표로 볼 수 있다. 에러가 적을수록 얼굴모양에 가깝다는 것을 의미하며 임계치 보다 작으면서 최소 잔여 에러를 갖는 영역을 최종적으로 얼굴영역이라고 판단 내리게 된다.

IV. 실험 결과

얼굴 검출 성능 평가를 위한 데이터베이스는 Yale, MIT, Yonsei 얼굴 데이터베이스에 있는 총 516 장의 영상으로 구성되었다. Yale 얼굴 데이터베이스는 흰색 바탕의 배경에 표정의 변화에 초점을 맞추어 15명의 165장의 영상으로 구성되었고 MIT 얼굴 데이터베이스는 비교적 단순한 배경하에 크기의 변화에 초점을 맞추어 16명의 144장 영상으로 구성되었으며 자체적으로 수집한 Yonsei 얼굴 데이터베이스는 매우 복잡한 배경하에 약간의($\pm 15^\circ$) 얼굴 돌아감(in-depth rotation)을 고려하여 15명의 207장의 영상으로 구성되었다. 약간의 in-depth rotation을 고려한 이유는 실제 실시간 정면 얼굴 검출시 완전한 정면이 아닌 경우가 많기 때문이다. 세 가지 얼굴 데이터베이스 모두 조명/인종/수염/안경 등의 다양한 변화를 포함하여 구성되었다.

전체 데이터베이스 516장의 영상중 대표영상 92장을 임의적으로 뽑아 얼굴 여부 검증을 위한 고유얼굴을 생성하였다. 검출 성능평가는 나머지 424장의 영상에 대해서 이루어졌다.

실험은 얼굴 후보 영역의 검증 순서를 정할 때 다층 신경망에 의한 눈 모양 근접성 정보를 포함하였을 때와 하지 않았을 때 후보 영역의 랭킹별 적응율을 구하여 얼굴 후보 영역 추정 정확성을 평가하는 방식으로 진행되었다. 실험 결과, 제안한 다층 신경망을 이용하여 눈 영역을 추정한 방식이 그림 5에 보이는 바와 같이 단순한 형태적 정보만을 이용한 기존의 방식[5]보다 최고 25%의 더 나은 적응률을 보였으며 80%의 적응률을 얻는데 필요한 랭킹의 차이는 약 4배 가까이 벌어졌다. 이것은 80%의 적응률을 얻기 위

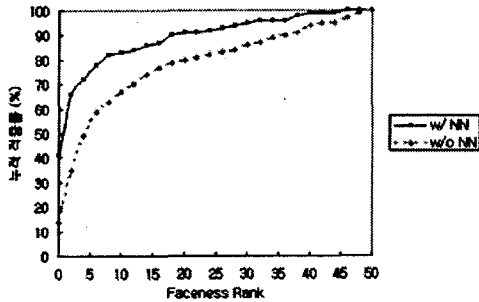


그림 5. 누적 적중률 vs. 얼굴후보영역 랭킹
해서 필요한 얼굴 후보 영역 수를 1/4로 줄였음을 의미한다. 전체 얼굴 검출률은 94%를 얻었으며 각 데이터베이스별 검출률은 표 1과 같다.

표 1) DB별 얼굴 검출률

	MIT	Yale	Yonsei	전체
검출률	83% (93/112)	96% (129/135)	99% (176/177)	94% (398/424)

나머지 6%의 에러중 눈 영역 검출에 실패하여 생긴 에러는 없었고 전부 같은 영상 내에 실제 얼굴 영역 보다 더 잔여 에러(residual error)가 적은 영역이 존재하여 발생하였다. 실제 실시간 시스템으로 구성하기 위하여 적중률 80%를 얻는 랭킹 5위 안에 드는 얼굴 후보영역에 대해서만 검증하도록 한 결과 320x240 크기의 영상을 기준으로 하였을 경우에 Pentium III 600MHz PC에서 초당 10 프레임의 영상을 처리할 수 있었다.

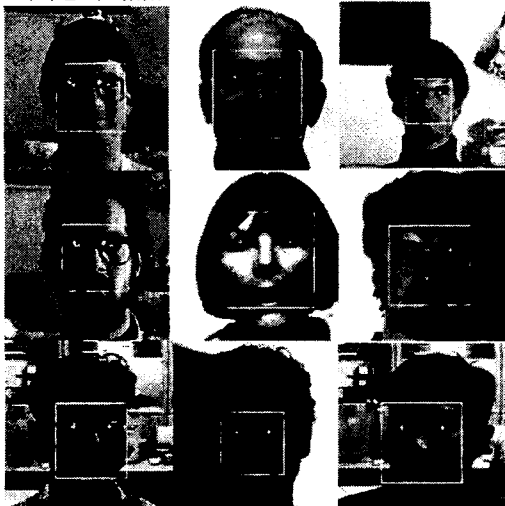


그림 6. 조명/표정/크기/포즈/안경/수염/배경의 변화를 포함한 얼굴 데이터베이스 상의 얼굴 검출 예

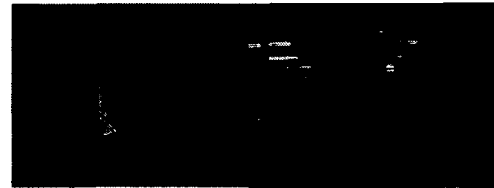


그림 7. 실시간 얼굴 검출 예

V. 결론

본 논문에서는 실시간 얼굴 검출을 위하여 다층 신경망 기반 눈 영역 추정을 이용한 빠른 얼굴 검출 기법을 제안하였다. 변형된 히스토그램 평활화 기법을 사용하여 조명에 강인하면서도 빠르게 눈 영역을 분리해 냈으며 각 눈 후보 영역마다 다층 신경망을 사용하여 눈 모양 근접성 정도를 점수화 함으로써 얼굴 후보 랭킹을 매겼을 때 적중률을 높였다. 고유얼굴 기법을 이용하여 검증한 결과 제안한 방법이 빠르면서 정확한 방법임을 보였다.

구현한 실시간 얼굴 검출 시스템은 향후 정확한 얼굴 영역 분리(segmentation)에 관한 문제와 좀 더 정밀한 얼굴 검증 기법에 관한 연구가 이루어진 후 실시간 정면 얼굴 검출을 요하는 응용분야에서 사용될 수 있을 것이다.

참고문헌

- [1] Haiyuan Wu, et al., "Detecting Human Face in Color Image", IEEE International Conf. on Systems, Man and Cybernetics, pp. 2232-2237, 1996.
- [2] 조성린, "얼굴인식을 위한 누적 차영상 기반의 얼굴 영역 검출에 관한 연구", 연세대 석사학위논문, 2001.
- [3] Jung-Il Choi, et al., "Face and eye location algorithms for visual user interface", IEEE First Workshop on Multimedia Signal Processing, pp. 239-244, 1997.
- [4] Kin-Man Lam, "A fast approach for detecting human faces in a complex background", ISCAS 98' Vol.4, pp. 85 -88. 1998.
- [5] Jong-Gook Ko, et al., "Facial Feature Tracking for Eye-Head Controlled Human Computer Interface", IEEE, TENCON 99', Vol 1, pp.72-75, 1999.
- [6] Matthew Turk and Alex Pentland, "Eigenfaces for Recognition", Journal of cognitive neuroscience, Vol 3, No. 1, 1991.