

독립적 컬러채널을 이용한 얼굴검출 성능개선

이영복, 민현석, 노용만
한국정보통신대학교 영상 및 비디오시스템연구실
e-mail : yro@icu.ac.kr

Face detection enhancement using independent color channels

Young Bok Lee, Hyun Seok Min, Yong Man Ro
Image and Video systems Lab, Information and Communications University

요 약

본 논문은 기존의 질감기반 (texture) 얼굴검출 시스템에서 컬러 영상을 도입하여 성능개선의 중요한 부분인 얼굴 오검출율을 줄이는 방법을 제안한다. 얼굴 영상의 컬러 성분은 흑백 성분과 비교하여 낮은 공간 주파수 영역을 가지는 특징이 있다. 질감기반 얼굴검출에서 높은 대비 (contrast) 성분의 에지는 얼굴이 아닌 영역에서 얼굴로 오인할 수가 있다. 본 논문에서는 이런 오인을 감소하기 위해 독립적인 컬러 채널 성분들을 질감기반 얼굴 검출에 각각 이용하여 그 얻어진 결과들을 융합 (fusion) 하는 방법을 제안한다. 실험결과로 제안한 칼라 채널 융합 방법을 통해 얻은 얼굴 검출율은 기존 흑백 영상과 비슷하게 유지되며 오검출율을 현저히 줄이는 것을 보였다.

1. 서론

얼굴 검출은 얼굴 인식 시스템, 영상 감시 시스템, 생체 보안 시스템 및 다양한 분야에서 쓰이는 응용기술이다. 특히 생체 보안 산업의 증가로 인해 정확하고 효율성이 높은 얼굴 검출 기술이 요구되고 있다. 얼굴 검출 방법으로는 지식 기반 방법[1], 색상 기반 방법 [2], 특징 기반 방법 [3], 신경망 이용 방법[4] 등이 있다. 신경망을 이용한 얼굴검출은 정지 영상에서 우수한 성능을 보이지만 높은 해상도를 요구하며 검출 시간이 매우 느리다. 또한 색상 기반 방법은 주로 다른 방법과 결합되어 쓰이지만 조명에 큰 영향을 받는 단점이 있으며 얼굴과 같은 색이 있을 경우 검출의 정확성이 떨어지는 단점이 있다. 이미지내의 픽셀들의 정보에서 특징 (feature) 정보를 생성하여 쓰이는 특징 기반 방법이 있다. 최근에는 Haar-like 특징과 Integral 이미지를 적용하여 빠르고 정확한 얼굴검출을 수행하는 질감기반 (texture)방법이 연구되었다 [4].

질감기반 방법을 사용한 Viola [4]의 객체 검출 알고리즘은 매우 뛰어난 성능을 보인다. 하지만 빠른 검출 속도, 높은 얼굴 검출율 에도 불구하고 상당한 오검출율(False Positive rate)이 존재한다. 향상된 성능을 위해 얼굴 검출율 및 검출 속도에 집중하지만 이러한 오검출율을 줄이는 것 또한 성능개선에 있어 매우 중요하다.

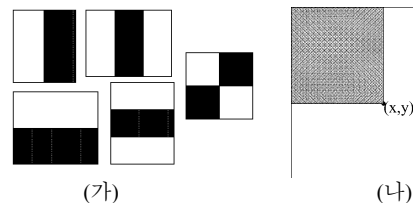
본 논문에서는 오검출율을 줄이기 위하여 오인식된 영역을 분석하고 독립적인 컬러 채널을 통해 얼굴 검출을 수행한 후 융합함으로써 얼굴 검출율 (True Positive rate)을 유지시키며 오검출율을 현저히

줄이는 방법을 제안한다.

논문의 구조는 다음과 같다. 2 장에서는 기존의 질감기반 방법을 사용한 Viola 의 얼굴검출에 대해 설명하고 3 장에서는 오인식된 영역을 분석한다. 또한 분석을 바탕으로 어떻게 독립적인 컬러 채널을 사용하고 융합을 하여 오검출율을 줄이는지에 대한 제안 방법이 설명된다. 4 장에서는 실험결과를 통해 제안된 방법이 성능개선에 있어서 우수성을 입증해 보인다. 마지막으로 5 장에서는 본 논문의 결론으로 제안된 방법에 대해 총 평가로 끝맺음을 한다.

2. 질감기반 (texture) 방법의 얼굴검출

Viola [5] 의 질감기반 (texture) 방법을 사용한 얼굴 검출 방법은 Papageoriou et al [6]에 제안된 간단하고 연산이 빠른 Haar-like 특징을 얼굴 검출을 위한 특징 집합으로 사용하였으며 Integral 이미지를 이용하여 빠르게 연산 하였다. 그림 1 에서는 Viola 가 사용한 Haar-like 특징 프로토타입 들을 볼 수 있다. 특징값 계산이 반복적으로 일어나기에 학습단계에서 상당한 시간이 소모된다. 이런 시간을 줄이기 위해 Integral 이미지가 사용되었는데 이는 영상처리에서 사용되는 영역 합 테이블 (Summed Area Table: SAT)과 유사하다.



(그림 1) 가) 5 개의 기본 특징 나) 인테그랄 이미지

그림 1 나에서 보듯이 (x,y) 의 위치는 다음과 같은 픽셀의 합을 포함하고 있다.

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y') \quad (1)$$

여기서 $i(x,y)$ 는 원본 사진이며 $ii(x,y)$ 는 인테그랄이미지이다. 이렇게 Haar-like 특징 추출 단계에서 생성된 얼굴 특징 계수를 이용하여 얼굴을 분류하기 위해서 Adaptive Boosting (AdaBoost) 알고리즘이 사용되었다. Adaboost의 기본 개념은 약한 분류기 (Weak classifier)를 선형적으로 결합하여 최종적으로 높은 검출 성능을 가진 강한 분류기 (Strong classifier)를 생성하는 것이다 [5]. AdaBoost 알고리즘은 초기 특징 집합에서 얼굴을 가장 잘 검출할 수 있는 특징의 선택과 분류기 학습이 동시에 이루어진다. 약한 분류기는 대부분의 초기 특징들을 제거하는데 사용되고 보다 강한 분류기는 낮은 오검출율에 도달하기 위해 사용된다. 오검출율을 최소화하기 위해 초기 AdaBoost 알고리즘의 문턱치값은 훈련 데이터에서 낮은 에러율을 가지고 설계된다. 일반적으로 보다 낮은 threshold는 높은 검출율과 높은 오검출율을 보인다.

이러한 특성으로 인해 Viola의 질감 기반 얼굴검출 방법은 높은 얼굴 검출율 일지라도 높은 오검출율을 수반한다[7]. 성능개선 부분에 있어 얼굴 검출율을 높이는 것 또한 중요하지만 오검출율을 줄이는 것도 성능개선에 있어 매우 중요한 사실이다. 이에 본 논문은 얼굴로 오인식된 영역을 분석하고 이를 바탕으로 얼굴 검출율을 적정선으로 유지하며 오검출율을 현저히 줄이는 방법을 제시한다.

3. 오인식 영역 분석 및 독립적 컬러 채널을 통한 얼굴검출

오인식된 영역은 검출 시스템이 얼굴이라고 판명된 비얼굴 영역이며 그림 2에서 확인할 수 있다.

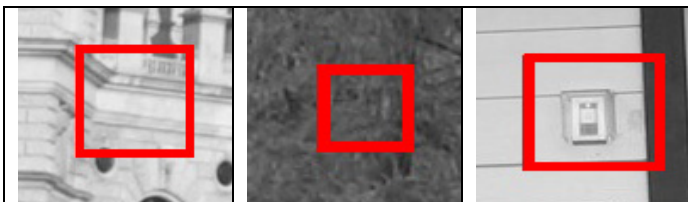


그림 (2) 오인식 영역들

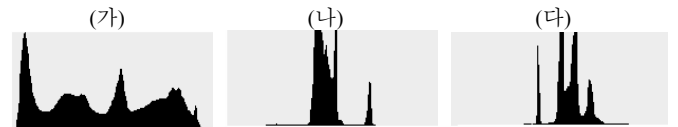
Viola의 알고리즘은 먼저 흑백으로 변환된 컬러영상이 얼굴 검출 시스템에 입력으로 들어간다. 이때 얼굴 검출 및 오검출이 이루어진다. Viola는 그림 1과 같은 특징을 쓰기 때문에 에지 성분이 뚜렷한 흑백영상의 특성과 더불어 오검출을 생성하게 된다. 이에 본 논문에서는 한 영상에 존재한 에지 성분을 다음 분산공식과 히스토그램을 이용하여 보여준다.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (2)$$

흑백 이미지의 성분을 비교하기 위하여 영상에서 처리가 간단하며 에지 성분이 적게 예상되는 Cb, Cr 컬러공간을 독립적으로 사용하여 비교해본다. 위의 분산 공식은 영상의 픽셀들을 분산 값으로 표현하기 위함이다. 분산 값이 높으면 높을수록 에지 성분이 높다는 것을 알 수 있다. 본 논문에서는 공식 DB인 MPEG-7 database의 몇 가지 이미지들을 예로 사용해 Cb나 Cr의 분산 및 히스토그램을 흑백영상과 비교해 에지 성분이 적은지 비교한다.

표 1. 각 채널의 분산값

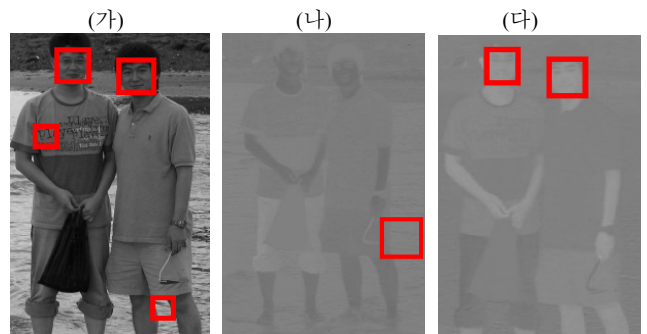
Filename	Y	Cb	Cr
al_00001.jpg	19.02	1.21	1.02
al_00004.jpg	13.58	0.16	0.10
al_00007.jpg	22.65	0.57	0.12
al_00009.jpg	16.60	0.43	0.75
al_000011.jpg	14.15	1.61	0.87



(그림 3) 각 채널의 히스토그램 가) Y 나) Cb 다) Cr

표 1과 그림 3에서 보듯이 Cb이나 Cr은 에지성분이 Y(흑백) 이미지보다 적다. 히스토그램에서도 보듯이 Y는 Cr이나 Cb에 비해 전체적으로 퍼져있음을 볼 수 있다.

위에서 언급된 바와 같이 Viola의 얼굴 검출 시스템에서의 오검출은 흑백영상 즉 Y(Luma) 이미지에 에지 성분이 높기 때문이고 이는 오인식 영역을 분석하여 증명하였다. 그러므로 본 논문에서는 이러한 사실을 바탕으로 독립적인 컬러채널(YCbCr)을 사용하여 얼굴 검출을 수행한다. 그림 4는 입력영상을 각 채널로 변환하여 얼굴검출을 수행한 결과의 예이다.

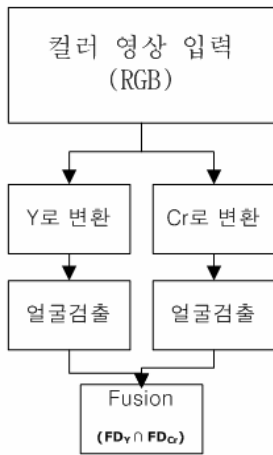


(그림 4) 가)Y 나)Cb 다)Cr로 얼굴검출 한 결과

그림 4에서 확인할 수 있듯이 Y에서는 얼굴검출이 정확히 이루어 졌으나 2개의 오검출을 보인다. Cb에서는 얼굴 검출이 없으며 오검출은 있다. Cr에서는

얼굴 검출이 있으나 오검출은 존재하지 않는다. 이는 Cr 에서는 위에서 증명되었듯이 에지 성분이 적기에 오검출이 없다. Cr (Chrominance red)은 또한 얼굴에 살색이 많이 포함 되어있는 특징을 부각시킨다[8]. 이로 인해 Cr 에서는 얼굴검출율이 높으면서도 오검출율을 낮은 성능을 보이게 된다. 이런 특성을 고려하여 본 논문에서는 독립적 컬러 채널인 Y 와 Cr 을 각각 얼굴검출을 수행한 후, 융합하는 방법을 제안한다. 여기서 Cb 는 전체적으로 얼굴검출율이 낮게 판명되어 사용되지 않았다.

제안된 방법의 순서도는 다음과 같다.



(그림 5) 제안된 방법을 사용한 얼굴 검출 시스템

입력 영상이 들어왔을 때 그림 5 처럼 우선 RGB 채널에서 Y, Cr 채널로 각각 얼굴 검출을 수행한다. 여기서 생성된 결과는 융합을 통해 성능을 개선 시킨다. 융합은 Y 에서 검출된 얼굴과 Cr 에서 검출된 얼굴이 일치 시에만 얼굴 검출이 되었다고 정의된다. 또한 Y 와 Cr 채널에서 생성된 오인식 영역들도 서로 일치할 경우 오검출율이 올라가게 된다. 하지만 정확히 일치 되어야만 하는 것은 아니라 영역이 서로 90% 이상 겹칠 때 같은 얼굴 영역 혹은 오인식 영역이라고 정의된다. 이렇게 제안된 방법을 통해 얻어진 성능개선은 실험 결과에서 입증된다.

4. 실험 결과

본 실험에서는 얼굴 검출에 사용되는 공식 DB 인 MPEG-7 Database 를 사용하였다. 총 669 장의 이미지에는 총 1161 개의 얼굴이 존재한다. 기존의 Viola 의 얼굴검출 방법과, 독립 채널인 Y 와 Cr 을 사용하여 융합한 제안된 방법과의 차이를 비교한다.

아래 표 2 에서는 TP rate (얼굴검출율) 와 FP rate (오검출율)를 통해 기존의 Y (흑백) 채널과 제안된 Cr 채널을 각각 통해 얻어진 얼굴검출 성능 결과이다. 또한 성능평가를 위한 TP rate 과 FP rate 는 다음과 같이 정의된다.

$$\text{True Positive Rate} = \frac{N_{TP}}{N_{True}} \quad (3)$$

$$\text{False Positive Rate} = \frac{N_{FP}}{(N_{TP} + N_{FP})} \quad (4)$$

여기서 N_{True} 는 전체 얼굴의 개수이고, $N_{Positive}$ 는 검출한 얼굴의 개수이다. 검출한 얼굴 중 실제 얼굴인 경우의 수는 N_{TP} 이고, 실제 얼굴이 아닌 경우의 수는 N_{FP} 이다.

다음은 제안된 방법으로 얼굴이 검출되었을 때의 성능을 기존 방법과 함께 비교한 것이다.

표 2. 각 채널의 얼굴검출 성과

	N_{True}	N_{TP}	N_{FP}	TP rate	FP rate
Y	1161	1120	165	96%	13%
Cr	1161	1073	81	92%	7%
$Y \cap Cr$	1161	1066	31	92%	3%

기존 방법은 Y (흑백) 일 때와 동일하다. 기존 방법은 96%의 높은 얼굴 검출 성능을 보였지만, 13%라는 오검출이 있다. Cr 로 변환된 입력으로 얼굴 검출 시, 92%라는 얼굴 검출율이 나왔으며 7%의 낮은 오검출율을 보였다. 그리고 융합을 하였을 때 92%의 얼굴 검출율과 3%라는 매우 낮은 오검출율이 나왔다.

기존 방법 Y 만 입력으로 쓰는 얼굴 검출 시스템은 높은 얼굴검출율을 보이지만 오검출율도 매우 높은 편이다. 본 논문에서 제안한 방법은 높은 얼굴검출율을 유지하면서 반면에 오검출율을 현저히 줄여 성능 개선에 있어 매우 중요한 역할을 하였다.

5. 결론

본 논문에서는 질감 기반 방법을 이용한 기존 얼굴 검출 시스템의 오인식 영역을 분석하고 이를 바탕으로 에지 성분이 강한 Y (흑백) 만을 사용하기보다, 에지 성분이 낮은 Cr 채널을 독립적으로 사용해 Y 와 Cr 을 각각 검출 시스템의 입력으로 사용하고 그 결과들을 조합시키는 방법을 제안하였다. 실험을 통해 제안된 방법의 우수성을 증명하였다. 이는 기존 방법 방 비교하였을 때 얼굴 검출율을 유지시키며 오검출율을 현저히 줄여지는 성능을 보였다.

참고문헌

[1] Ming-Hsuan Yang, Kriegman, D.J, and Ahuja N., "Detecting face in images : a survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol 24, Issue:1, pp34-58, 2002.
 [2] J. Yang and A. Waibel, "A real-time face tracker," *In Proc. 3rd Workshop on Appl. Of Computer Vision*, pp 142-147, 1996.
 [3] R. Brunelli and T. Poggio, "Face Recognition: features versus templates," *IEEE Transactions on PAMI*, Vol.15, No.10, pp1042-1052, 1993.
 [4] H.A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *In IEEE Patt. Anal. Mach. Intell*,

Vol.20, pp 22-38, 1998.

[5] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Computer Vision and Pattern Recognition 2001 (CVPR 2001)*, Vol.1, pp 511-518, Dec. 2001.

[6] M.Oren, C.Papageorgiou, P. Shinha, E.Osuna, and T. Poggio, "Pedestrian detection using wavelet templates," *In Computer Vision and Pattern Recognition*, pp 193-199, 1997.

[7] R. Lienhart, A. Kuranov, V. Pisarevsky, "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection," *DAGM'03 25th Pattern Recognition Symposium*, pp. 297-304, Sep. 2003.

[8] R. Hsu, M. Abdel-Mottaleb, A. Jain "Face Detection in Color Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, No.5 May 2002.