

눈 영역에 적합한 에지 추출과 밝기값 정보를 이용한 눈 검출*

문원호* · 최연석* · 김철기** · 차의영*

*부산대학교 컴퓨터공학과

**부산대학교 디자인학과

Detection of eye using optimal edge technique and intensity information

Won-ho Mun* · Yeon-seok Choi* · Cheol-ki Kim** · Eui-young Cha*

*Dept. of Computer Engineering, Pusan National University

**Dept. of Design, Pusan National University

E-mail : aquamariner@nate.com

요 약

사람의 눈동자는 얼굴 크기와 비교해 볼 때 상대적으로 일정한 거리를 가지고 있기 때문에 이미지가 정규화에 있어서 중요한 지표로 사용된다. 이 논문은 이러한 특징을 이용해 최적화된 세그멘테이션 방법을 사용하여 눈동자 검출의 새로운 접근방법을 소개한다. 눈 검출 방법은 세 가지 중요한 단계로 나눌 수 있다. (1)흑백 영상에서 눈 영역에 적합한 에지 추출 방법, (2)레이블링(labeling) 기법을 이용한 눈 영역 추출, (3)밝기값 정보를 이용한 눈동자 위치 검출. 실험 결과로는 다양한 조명 환경과 얼굴표정을 가진 2408장의 FERET 영상을 이용하여 98.9%의 검출 성능을 보였다.

ABSTRACT

The human eyes are important facial landmarks for image normalization due to their relatively constant interocular distance. This paper introduces a novel approach for the eye detection task using optimal segmentation method for eye representation. The method consists of three steps: (1)edge extraction method that can be used to accurately extract eye region from the gray-scale face image, (2)extraction of eye region using labeling method, (3)eye localization based on intensity information. Experimental results show that a correct eye detection rate of 98.9% can be achieved on 2408 FERET images with variations in lighting condition and facial expressions.

키워드

eye recognition, eye detection, edge extraction

1. 서 론

사람의 얼굴 이미지 분석과 인식은 지난 20년 동안 컴퓨터 비전과 패턴 인식 분야에서 매우 중요하게 연구되는 분야중 하나이다. 최근에는 인간-로봇의 상호작용과 관련한 지능형 로봇과 접목을 통해 얼굴검출[1], 얼굴인식[2], 얼굴표정인식[3]과 같은 거대한 잠재력이 있는 어플리케이션들이 요

구되고 있기 때문에 관심이 더욱더 집중되고 있다. 이러한 얼굴인식분야의 가장 필수적이지만 해결되어야 할 매우 중요한 문제는 바로 눈 검출이다. 눈은 사람의 얼굴에서 가장 두드러지고 안정적인 특징을 가지고 있기 때문이다. 때로는 눈 추출은 얼굴 찾기 알고리즘에서 가장 중요한 단계로 간주되기도 한다. 두 얼굴을 비교하기 위해서는 정렬과 크기의 정규화가 선행되어야 한다. 두 눈의 위치와 두 눈 사이의 거리는 상대적으로 일정한 비율을 가지기 때문에 눈 검출은 정규화 용도로도 사용된다[4]. 눈의 위치 측정은 얼굴의 다른 경계영역을 탐색하는데 용이하게 쓰여 지고

※ 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임.(No.2010-0028242)

있다. 눈은 사람의 감정을 반영하기 때문에 눈 트래킹과 함께 쓰여지면 얼굴 표정 분석, 휴먼-컴퓨터 상호작용, 혁신적인 사용자 인터페이스 분야에서 결정적인 진보를 가져올 수 있는 역할을 할 것으로 기대된다[3].

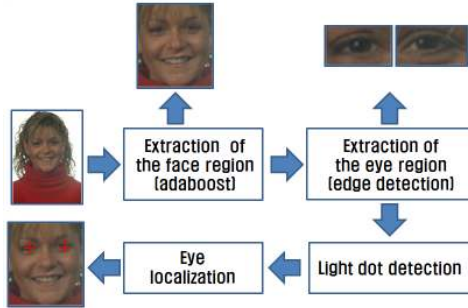


그림 1. 제안된 눈 검출 순서도

지금까지 많은 시간과 노력이 소비됐음에도 불구하고 아직도 자동 눈 검출에 있어서 그 복잡성에 기인하여 빛의 상태나 표정, 방해물(안경, 머리 카락) 등에 의해 각 알고리즘들의 성능은 의심할 여지없이 많은 영향을 받고 있다. 본 논문에서는 눈 영역에서 최적화된 에지 검출 기법과 밝기 정보를 이용하여 수행속도가 빠르고 좋은 성능을 가지는 새로운 눈 검출 방법을 제시하고자 한다. 그림 1은 본 논문에서 제안하는 눈 검출 방법의 순서도를 보여준다.

II. 얼굴 검출

Paul Viola와 Michael Jones에 의해 AdaBoost를 이용한 얼굴 검출기가 처음 개발되었고 Rainer Lienhart와 Jochen Maydt에 의해 확장된 얼굴 검출기에 대한 구현을 제공하였다[12]. Paul Viola와 Michael Jones가 개발한 얼굴 검출기를 보통 Viola-Jones 검출기라고 부른다. 그리고 이 검출기를 "Harr classifier"라고 지칭한다. 왜냐하면 이 검출기는 영상에서 특정 사각형 정보를 빼거나 더하는 과정을 통해 구해지는 Haar-like wavelet 또는 Haar feature를 사용하기 때문이다. 본 논문에서는 얼굴 검출을 위해 Viola와 Jones의 얼굴 검출기를 사용하였다. 본 장에서는 Viola와 Jones 검출기의 특징들을 간략하게 소개한다.

Viola와 Jones 분류기는 Haar-like 입력 특징들을 사용한다. 즉, 사각형 영상 영역의 합 또는 차에 임계값이 적용된다.

Viola와 Jones 분류기는 적분 영상(integral image)을 사용함으로써 사각형 영역 또는 45도 회전된 사각형 영역에서의 합계 연산을 매우 빠르게 수행한다. 적분 영상을 이용함으로써 Haar-like 입력 특징들을 빠르게 계산할 수 있다.

Viola와 Jones 분류기는 높은 검출률과 약한 거절 비율을 갖는 이진 분류 노드를 만들기 위하여 통계적 부스트를 사용한다.

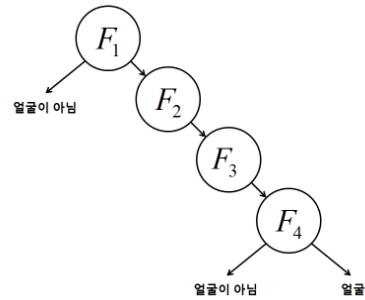


그림 2. Viola와 Jones 분류기의 거절 캐스케이드

Viola와 Jones 분류기는 그림 2과 같이 약한 분류기 노드를 이용하여 거절 캐스케이드를 구성한다. 첫 번째 분류기 집단은 많은 실수를 포함하지만 물체를 포함하고 있는 영역을 가장 잘 찾아내는 것으로 선택한다. 다음 분류기 집단은 약한 거절을 갖는 두 번째로 검출 성능이 좋은 것으로 선택하고, 이러한 방식을 반복한다. 실험 단계에서는 전체 캐스케이드를 다 통과한 경우에만 찾거나 하는 객체로 검출된다.

III. 눈 영역에 적합한 에지 추출

추출된 얼굴 영상에서 얼굴을 구성하는 각각의 구성요소(눈, 눈썹, 코, 입)들은 수평, 수직에지 성분을 많이 가지고 있다. 특히, 눈 영역은 다른 구성요소보다 더 많은 에지 정보를 가지고 있다.

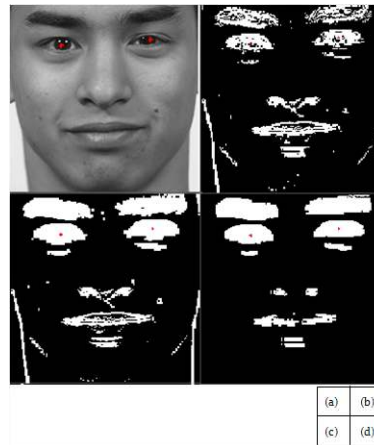


그림 3. 에지 추출 영상, (a) 원영상, (b) 제안한 에지추출 방법 결과 영상, (c) (b)영상을 팽창(dilation) 연산한 결과 영상, (d) 제안한 수직에지 제거 방법 결과 영상.

2차원 $Q \times Q$ 영상은 아래의 수식 1과 같이 정의할 수 있다.

$$\{f(x, y) | x = y = 0, 1, \dots, Q-1\} \quad (1)$$

$$\alpha = \frac{|I(k+1, y) - I(k, y)|}{I(k+1, y) + I(k, y)} \quad (2)$$

$$\beta = \frac{|I(x, k+1) - I(x, k)|}{I(x, k+1) + I(x, k)} \quad (3)$$

수식 2와 수식 3에서 $I(m, n)$ 는 원영상의 X축 좌표가 m 번째이고 Y축 좌표가 n 번째인 픽셀의 밝기값 정보를 의미한다. α, β 는 수평, 수직에지 성분을 검출하기 위한 지역적 적응형 경계값(threshold)을 나타낸다. 수식 2와 수식 3의 α, β 값을 기준으로 이진화된 두 영상을 AND 연산한 결과는 그림 3의 (b)와 같다. 본 논문에서는 실험 결과를 통해 최적화된 α, β 의 파라메타 수치를 각각 1.25, 1.175로 정하였다.

그림 3의 (c)영상은 얼굴 각 구성요소를 세그멘테이션 작업이 최적화 될 수 있도록 하기 위해 모폴로지 연산중의 하나인 팽창(dilation)을 수행한 결과 영상이다.

IV. Labeling을 통한 눈 영역 추출

그림 3의 (b)와 같이 지역적 적응형 경계값을 통해 이진화된 영상은 많은 에지 정보를 가지고 있기 때문에 세그멘테이션하기에는 적합하지 못하다. 특히 모폴로지의 팽창 연산의 결과로 눈과 눈썹의 영역이 명확하게 구분되지 못하는 문제점이 발생한다. 정확한 눈 영역을 세그멘테이션하기 위해서는 불필요한 수직에지 성분을 제거할 필요가 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 효과적인 에지 제거 방법을 제시한다.

$$\sigma_k = \frac{VPC_k}{HPC_k} > 1.5 \quad (k=0,1,2\dots) \quad (4)$$

수식 4에서 HPC (horizontal pixel connected)는 k 번째 픽셀 좌표에서 수평 방향으로 연속해서 연결되어 있는 총 픽셀 개수이고, VPC (vertical pixel connected)는 k 번째 픽셀 좌표에서 수직 방향으로 연속해서 연결되어 있는 총 픽셀 개수이다. 예를 들면, 그림 4에서 픽셀 P의 위치에서 $HPC = |AB|$ 이고 $VPC = |CD|$ 이다. σ 값은 수식 4를 통해 구할 수 있다. 분명한 것은 눈과 눈썹의 σ 값은 1보다 작다는 것을 알 수 있다. 본 논문에서 얼굴 경계영역(얼굴 윤곽)이나 방해요소(안경, 머리카락)를 제거하기 위해 최적화된 σ 의 값을 1.5로 정하였다. 그림 3의 (d)는 제안한 방법으로 수직에지 성분을 제거한 결과 영상을 보여준다.

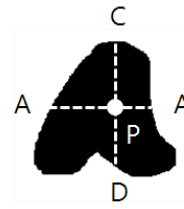


그림 4. HPC와 VPC 예제

제안된 방법으로 불필요한 수직에지 성분을 제거한 영상을 레이블링하게 되면 그림 5와 같이 여러 개의 눈 후보 영역이 세그멘테이션 된다. 레이블링 수행 중에서 후보대상에서 제거되는 조건은 다음과 같다.

- Noise size가 2보다 작을 때
- (완전히 연결되어 있지 않은 최대의 공백사이즈)
- Blob 영역에 포함된 유효 픽셀 수가 너무 작을 때, 20 pixels(256x384영상)
- Blob 영역의 가로크기(width)가 얼굴 영역의 1/3 이하의 크기일 때

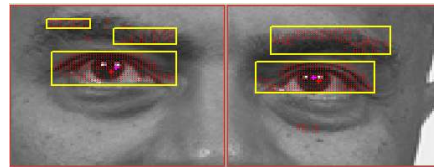


그림 5. 제안한 방법으로 레이블링한 결과영상

V. 밝기값 정보를 이용한 눈 영역 확인

본 논문에서는 정확한 눈 위치를 찾기 위하여 밝기값(intensity) 정보를 이용한다. 눈의 동공(iris)에는 빛이 반사되어 비춰지는 밝은 작은 원들이 존재한다. 이러한 작은 원은 상대적으로 동공과 비교해볼 때 매우 밝은 값을 가지고 있다. 그리고 어두운 영상에서는 동공의 빛 반사가 이루어지지 않는 대신 공막(sclera)이 상대적으로 매우 밝은 값을 가지고 있다. 이러한 특징을 이용하여 세그먼트의 가장 밝은 밝기값과 세그먼트의 평균 밝기값과의 대비, C (contrast)는 다음과 같이 구할 수 있다.

$$C_k = \frac{\max(x, y) \in \{L_k(x, y)\}}{AL_k} \quad (5)$$

$$C > \gamma \quad (6)$$

수식 5에서 L_k 는 k 번째 세그먼트의 해당 좌표의 밝기값 정보이고, $\max(x, y)$ 는 L_k 의 집합에 속하는 가장 밝은 밝기값이다. AL_k 는 k 번째 세그먼트의 평균 밝기값이다. 이러한 상대적인 대비 값이 눈 영역을 만족하기 위해서는 아래의 수식 6을 만족해야한다. 본 논문은 실험결과를 통해 최

적의 γ 상수의 값을 2.5로 제시한다.

VI. 실험 결과

눈 검출에 쓰여 지는 성능 평가 기준은 두 가지가 있다. 첫째는 검출 성공률, 둘째는 위치의 정확성을 평가하는 것이다. 검출 성공률은 정확하게 찾아진 데이터의 비율을 의미한다. 위치의 정확성은 수동으로 찾은 기준데이터(ground truth)와 자동으로 찾아진 눈의 위치의 차이를 의미한다. 본 논문에서는 Jesorsky et al.[13]이 제안한 수식 7과 같은 방법으로 실험결과를 평가하였다.

$$E = \frac{\max(D_L, D_R)}{D_{LR}} \quad (7)$$

E 는 눈 검출 오차율을 의미하고, D_L 은 왼쪽 눈 좌표의 오차율, D_R 은 오른쪽 눈 좌표의 오차율을 의미한다. D_{LR} 은 기준데이터의 왼쪽과 오른쪽 눈의 유클리디안 거리(Euclidean distance)를 의미한다. 이 평가 방법은 해상도와 무관하게 성능을 평가할 수 있다는 장점이 있다.

실험 결과는 FERET 데이터베이스를 이용하였고 0.25의 오차율(E)을 기준으로 98.9%의 높은 검출률을 보였다. FERET 데이터베이스 중 사람의 얼굴이 정면을 향해있는 사진의 개수는 총 2724장이다. 본 논문에서는 눈 검증 좌표(ground truth)가 제공되지 않은 316개의 영상을 제외한 2408장의 영상으로 실험을 하였다. 전체 영상은 안경을 쓴 영상이 260장, 눈감은 영상이 171장, 어두운 영상이 108장, 그 외는 일반 영상으로 분류되어진다. 그림 6은 실험 결과 예제 영상이다.

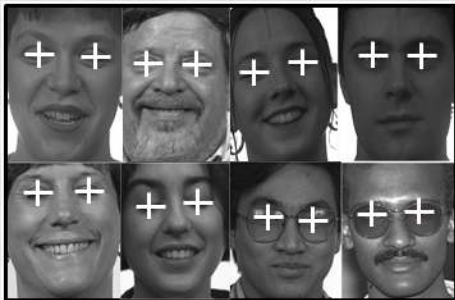


그림 6. 실험 결과 예제 영상

오검출된 영상을 분석한 결과, 전체 오검출 영상의 75% 이상이 안경을 쓴 영상에서 발생하였고, 나머지 25%는 어두운 영상, 눈 감은 영상, 일반 영상에서 오검출 되었다. 안경을 쓴 영상에서는 안경 테두리나 렌즈에서 반사되어지는 밝은 빛이 오인식의 원인이었고, 어두운 영상의 경우에는 에지 검출이 이루어지지 않아 후보 영역을 추출하지 못해 눈 검출 오인식의 원인이 되었다.

VII. 결 론

눈 검출은 얼굴 자동 인식 시스템에서 매우 중요한 단계로 쓰여 지고 있다. 그리고 눈 검출은 다양한 얼굴 표정, 조명환경, 눈 주위의 방해물(안경, 머리카락, 주름)과 같은 복잡성에 기인하여 다양한 영상에서 좋은 성능을 가지기 어렵다. 본 논문에서는 눈 영역에서 최적화된 에지 추출 기법과 밝기값 정보를 이용하여 새로운 눈 검출 방법을 제안하였다.

실험 결과, FERET 데이터베이스의 영상 2408장에서 98.9%의 높은 검출 성능을 보였고, 수행 속도는 Inter core2 Duo CPU E8400 환경에서 평균 62ms의 빠른 수행 속도를 보였다. 이와 같이 본 논문에서 제안한 눈 검출 방법은 빠르고 좋은 성능을 보여 주기 때문에 얼굴 표정 인식과, 얼굴 감지 시스템의 전처리 단계에서의 사용되기에 적합한 방법 중 하나로 제시될 것으로 보인다. 또한, 제안한 방법은 적외선을 이용한 눈 검출 시스템과 유사한 동공의 밝기 정보를 이용하지만, 적외선과 관련된 특정 장비가 필요가 없다는 장점을 가진다.

참고문헌

[1] M.-H.Yang, D.J.Kriegman, N.Ahuja, "Detecting faces in images: a survey," IEEE Trans. Pattern Anal. Mach. Intell. 24-1, pp.34-58, 2002.

[2] A.Samal, P.A.Iyengar, "Automatic recognition and analysis of human faces and facial expressions: a survey," Pattern Recognition 25-1, pp.65-77, 1992.

[3] B.Fasel, J.Luetttin, "Automatic facial expression analysis: a survey," Pattern Recognition 36, pp.259-275, 2003.

[4] J.Huang, H.Wechsler, "Visual routines for eye location using learning and evolution," IEEE Trans. Evol. Comput. 4-1, pp.73-82, 2000.

[5] D.J.Beymer, "Face recognition under varying pose," in: IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR'94), Seattle, Washington, USA, pp.756-61, 1994.

[6] O.Jesorsky, K.J.Kirchberg, R.W.Frischholz, "Robust Face detection using the Hausdorff distance," in: Proceedings of the Third International Conference on Audio- and Video-based Biometric Person Authentication, Halmstad, Sweden, pp.90-95, 2001.