
주성분 분석을 이용한 기울어진 얼굴에서의 눈동자 검출*

최연석* · 문원호* · 차의영*

*부산대학교 컴퓨터공학과

Eye detection on Rotated face using Principal Component Analysis

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

*Dept. of Computer Engineering, Pusan National University

E-mail : yeonseokchoi@pusan.ac.kr

요 약

컴퓨터 비전을 이용한 눈동자 추적 기술은 Human-Computer Interface(HCI)의 중요성이 높아짐에 따라 많이 연구되고 있다. 본 논문에서는 HCI 장치를 위한 눈동자 검출 방법을 제안한다. 제안하는 방법은 기울어진 얼굴에서도 눈동자를 검출하기 위해 Principal Component Analysis(PCA) 방법을 이용하여 얼굴의 기울어진 정도를 검출하고 기울어진 정도를 이용하여 눈동자 영역의 위치를 계산한다. 최종적으로 눈 위치의 검출을 위해 눈동자 영역의 밝기 정보를 사용한다. FERET DB의 얼굴영상을 사용하여 실험한 결과, 기울어진 얼굴에서도 눈동자를 효과적으로 검출 할 수 있음을 확인하였다.

ABSTRACT

There are many applications that require robust and accurate eye tracking, such as human-computer interface(HCI). In this paper, a novel approach for eye tracking with a principal component analysis on rotated face. In the process of iris detection, intensity information is used. First, for select eye region using principal component analysis. Finally, for eye detection using eye region's intensity. The experimental results show good performance in detecting eye from FERET image include rotate face.

키워드

PCA, eye detecting, eye tracking, HCI

1. 서 론

새로운 환경을 위한 인간-컴퓨터 상호작용의 중요성이 높아짐에 따라 사람과 컴퓨터를 자연스럽게 연결하기 위한 인간-컴퓨터 인터페이스가 개발되고 있다. 이러한 인터페이스 중에서도 컴퓨터 비전 기술을 이용한 것들은 카메라 외에는 어떤 다른 특별한 장치도 필요치 않다는 사실은 큰 장점으로 작용한다. 또한 인간을 위한 인터페이스로써 시선 혹은 눈동자 추적은 인간의 관심이나 즉각적 반응, 흥미도 등을 표현할 수 있는 효과적인

도구로써 이용될 수 있다. 얼굴 인식, 표정 인식 등과 같은 얼굴 영상을 이용한 연구들은 일반적으로 입력 영상에 대한 정규화 과정이 필요하게 되는데 정규화는 주로 눈의 위치를 기준으로 이루어지며 그렇기 때문에 눈의 위치에 대한 정보는 필수적인 정보이다. 얼굴 영상에서 눈의 위치를 검출하기 위하여 많은 연구들이 진행되어 오고 있다. A.L. Yuille 등은 템플릿 정합 방법을 이용하여 눈 검출을 시도하였는데[1], 이 방법은 수행시간이 오래 걸리는 단점이 있다. A. Pentland 등은 Eigenspace 방법을 이용하여 눈 검출을 시

* 이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임.(No.2010-0028242)

도하였다[2]. 그러나 이 방법은 학습 데이터 영상에 큰 영향을 받는 단점이 있다. G.C. Feng 등은 variance projection function(VPF)을 이용하여 눈 검출을 시도하였다[3]. 이 방법은 눈이 작거나 감고 있는 영상에서는 검출하기 어려운 단점이 있다. 기존의 방법들은 자세, 표정, 조명 등의 변화에 잘 대처하지 못하거나 학습 데이터 영상에 민감한 문제점들이 있다.

따라서 본 논문에서는 먼저 AdaBoost를 이용하여 얼굴영역을 검출[4]한 후 얼굴에서 눈동자 영역을 찾기 위해 얼굴의 기울어진 정도를 주성분 분석을 통해 구하고 기울어짐과 얼굴의 형태학적 특징을 이용해 눈동자 영역을 찾는다. 그리고 눈동자의 명암도와 G.C. Feng 등이 제시한[3] VPF(Variance Projection Function)를 이용하여 눈동자 중심점을 찾는다.

II. 얼굴영역 검출 및 얼굴의 기울어짐 측정

눈동자 검출을 하기에 앞서 입력된 영상에서 얼굴영역을 먼저 검출해야 한다. 입력된 영상에서 얼굴 영역을 검출하기 위하여 본 논문에서는 AdaBoost 학습 알고리즘을 이용하며[4] 분류기는 Haar-like feature를 사용하였다[5].

두 번째로, 검출한 얼굴 영역에서 눈 영역을 포함한 후보 영역을 검출하기 위해 먼저 얼굴의 기울어짐을 측정한다. 얼굴의 기울어짐을 측정하기 위해 주성분 분석(Principal Component Analysis; PCA) 방법[6]을 사용한다. 주성분 분석 방법은 평균과 분산의 통계적 성질을 이용한 2차 통계적 기법으로 입력 자료에 대하여 공분산 행렬(covariance matrix)을 구하여 각 방향을 가리키는 정규화된 직교좌표 벡터를 찾는다. 공분산 행렬의 직교좌표 벡터중 장축에 해당하는 벡터의 기울기가 x축에 대한 얼굴의 기울기가 된다. 주성분 분석에 사용되는 데이터를 구하기 위해서는 먼저 얼굴 영역에서 기준이 되는 좌표를 추출해야 한다. AdaBoost 학습 알고리즘을 이용하여 얻어진 얼굴 영역에서 대략의 기준점은 눈동자 혹은 눈썹 영역으로 설정하며 눈동자 혹은 눈썹 영역을 검출하기 위해 먼저 얼굴의 형태학적인 특징을 이용하여 눈 영역을 설정한다.

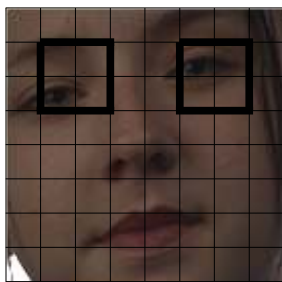


그림 1. 얼굴의 형태학적 특징

검출된 얼굴 영역에서 각각의 눈의 위치는 그림 1에서와 같이 얼굴 가로, 세로 길이의 8등분한 길이의 2배가 되며 얼굴 영역의 4등분한 영역중 1, 2사분면 가운데에 위치한다.

설정된 눈 영역에서 눈동자 및 눈썹을 검출하기 위해 평균 이진화를 적용한다. 이 때 눈동자 및 눈썹 픽셀의 명암도 값은 낮으므로 평균 이진화를 적용하여 얻은 임계치 값의 0.75배를 임계치 값으로 설정한다. 임계치보다 낮은 픽셀의 좌표는 주성분 분석의 데이터로 사용한다.



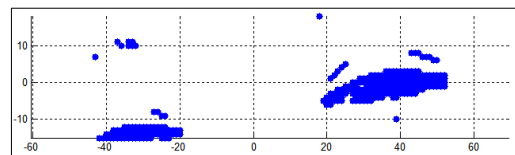
(a) 검출된 얼굴 영역 (b) 평균 이진화 영상

그림 2. 평균 이진화를 이용한 PCA 데이터 검출

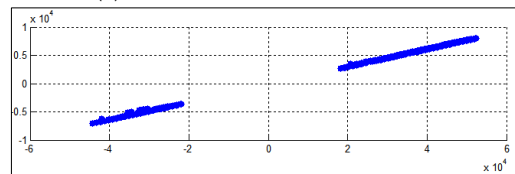
그림 2는 검출된 얼굴 영역(a)에서 평균 이진화를 적용(b)한 모습이다. 주성분 분석에 사용된 데이터는 그림 2의 (b)에서 검출된 픽셀의 x축 좌표 값, y축 좌표 값이 된다. 주성분 분석을 하기 위해 입력된 데이터의 공분산 행렬을 구한다. 공분산 행렬 C는 식 (2)와 같으며, X_i, Y_i 는 데이터 X, Y의 i번째 값, \bar{X}, \bar{Y} 는 데이터 X, Y의 평균 값을 나타낸다.

$$cov(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)} \quad (1)$$

$$C = \begin{pmatrix} cov(X, X) & cov(X, Y) \\ cov(Y, X) & cov(Y, Y) \end{pmatrix} \quad (2)$$



(a) 주성분 분석의 입력 데이터



(b) 공분산 행렬과 입력 데이터 곱한 결과

그림 3. 공분산 행렬 이용한 결과

주성분 분석의 입력 데이터의 공분산 행렬 $C(2 \times 2 \text{ matrix})$ 을 구한 후 입력 데이터 그림 3의 (a)($1 \times 2 \text{ vector}$)와 공분산 행렬 C 를 곱한 결과는 그림 3의 (b)와 같다. 주성분 분석의 입력 데이터는 그림 2의 (b)에서 검출된 좌표며 이 때, 공분산 행렬의 eigenvector의 기울기는 그림 3의 (b)에서 나타난 직선의 기울기가 된다. 다음 식 (3)에서 C 는 공분산 행렬이며 λ 는 eigenvalue, x 는 eigenvector가 된다.

$$Cx = \lambda x \quad (3)$$

eigenvalue는 식 (4)를 만족하는 λ 가 된다.

$$\det(C - \lambda I) = 0 \quad (4)$$

공분산 행렬 C 는 2차원 이므로 λ 의 값은 2개가 되며 이 때 큰 값에 해당하는 eigenvector의 기울기가 얼굴의 기울기가 된다. eigenvector $x = (v_x, v_y)$ 의 기울기는 식 (5)을 이용해서 구할 수 있다.

$$\theta = \tan^{-1}\left(\frac{v_y}{v_x}\right) \quad (5)$$

얼굴의 기울어짐 각이 구해지면 얼굴의 형태학적 특징을 이용하여 구한 눈동자 영역의 위치를 얼굴의 기울어짐 정도에 맞게 재조정 한다.

III. 눈동자 검출

각 각 검출된 왼쪽, 오른쪽 눈동자 영역에서 눈동자 중심의 최종 좌표는 눈동자 중심의 명암도 특징을 이용하여 구한다. 눈동자 중심의 명암도는 다른 영역의 명암도 보다 어두운 특징을 가지며 이것을 이용하기 위해 3×3 마스크 내 픽셀의 명암도 값의 합과 분산 값을 이용한다.

$$s(x, y) = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} I(i, j) \quad (6)$$

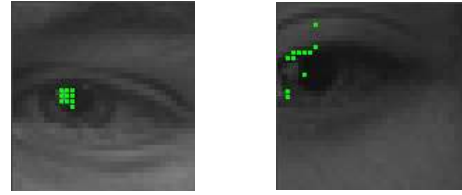
$$\overline{s(x, y)} = \frac{s(x, y)}{9} \quad (7)$$

$$v(x, y) = \sqrt{\sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} (I(i, j) - \overline{s(x, y)})^2} \quad (8)$$

$$p(x, y) = s(x, y) - v(x, y) \quad (9)$$

식 (6)에서 $s(x, y)$ 는 마스크 픽셀의 합, $\overline{s(x, y)}$ 는 $s(x, y)$ 의 평균값, $v(x, y)$ 는 마스크 픽셀에서의 분산의 값이 되며, $p(x, y)$ 의 값 중 작은 값 10개를 선택한다. 그림 4의 (a)에서와 같이 10개의 점이 한곳에 분포되어 있으면 이 영역의 중심점 좌표

가 눈동자의 중심 좌표가 되며 그림 4의 (b)와 같이 밀집되어 있지 않을 때에는 분산 프로젝션 함수를 이용하여 눈동자 중심 좌표를 구한다. 식 (10), 식 (11)은 IPF(Integral Projection Function)을 나타내는 식이며 $I(x, y)$ 는 (x, y) 위치에서 명암 값이고 $V(x)$ 는 영역 $[y_1, y_2]$ 의 명암도의 합이며 $H(y)$ 는 영역 $[x_1, x_2]$ 의 명암도의 합이다. $V_m(x)$, $H_m(y)$ 값은 어두운 영역에서는 값이 작게 나오며 밝은 영역에서는 값이 크게 나타나지만 영역의 밝기 합이 같을 경우 변화를 알아내는 것이 어렵다.



(a) 밀집되어 있는 경우

(b) 흩어져 있는 경우

그림 4. 값이 낮은 10개의 $p(x, y)$ 분포

이 함수를 보완하기 위해 VPF를 사용하며, VPF는 IPF를 사용하여 만들 수 있다. 수직 VPF를 $\sigma_v(x)^2$, 수평 VPF를 $\sigma_h^2(y)$ 로 정의하면 식 (14), 식 (15)과 같다.

$$V(x) = \int_{y_1}^{y_2} I(x, y) dy \quad (10)$$

$$H(y) = \int_{x_1}^{x_2} I(x, y) dx \quad (11)$$

$$V_m(x) = \frac{1}{y_2 - y_1} \int_{y_1}^{y_2} I(x, y) dy \quad (12)$$

$$H_m(y) = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} I(x, y) dx \quad (13)$$

$$\sigma_v^2(x) = \frac{1}{y_2 - y_1} \sum_{y_i=y_1}^{y_2} [I(x, y_i) - V_m(x)]^2 \quad (14)$$

$$\sigma_h^2(y) = \frac{1}{x_2 - x_1} \sum_{x_i=x_1}^{x_2} [I(x_i, y) - H_m(y)]^2 \quad (15)$$

최종적으로 눈동자 중심좌표 (x, y) 는 수직 VPF에서 가장 큰 peak 점의 x 좌표 수평 VPF에서 가장 큰 peak 점의 y 좌표가 된다.

IV. 실험 결과 및 분석

본 논문에서는 FERET 얼굴 데이터베이스를 이용하여 얼굴 영역에서 눈동자의 중심점 좌표를 찾는 실험을 하였다. 사용된 얼굴 영상은 정면 영

상을 사용하였으며 크기 256×384 영상을 사용하였다. 실험에 사용된 영상은 어두운 영상, 감은 눈 영상, 안경 쓴 영상이 포함되어 있다. 실험의 검증에 위해 실험 오차는 FERET DB의 ground truth 정보를 사용한다. ground truth 정보는 영상에서 왼쪽 눈, 오른쪽 눈에 대한 좌표를 말하며 이 좌표는 FERET에서 제공하는 좌표이다. 이 ground truth 정보를 이용하여 식 (16)을 구한다.

$$d_{eye} = \frac{\max(\|C_l - \tilde{C}_l\|, \|C_r - \tilde{C}_r\|)}{\|C_l - C_r\|} \quad (16)$$

여기서, C_l, C_r 은 왼쪽 눈, 오른쪽 눈 ground truth 이며, \tilde{C}_l, \tilde{C}_r 은 제안된 방법으로 검출한 왼쪽 눈, 오른쪽 눈의 좌표이다. $\|\cdot\|$ 은 Euclidean distance를 나타낸다. d_{eye} 값은 눈동자 좌표와 추출된 눈동자 좌표사이의 오차를 나타낸 값으로써 $\|C_l - C_r\|$ 은 눈 영역의 약 2배정도 값을 가진다. 따라서 $d_{eye} \leq 0.25$ 의 의미는 왼쪽 눈 또는 오른쪽 눈의 오차 중 큰 값은 눈 영역의 절반 안에 들어오는 것으로 간주할 수 있으며, $d_{eye} \leq 0.05$ 오차는 입력 영상 256×384 일 때, ground truth와의 오차가 2~3 픽셀 이내에 해당된다.

표 3. 제안한 방법으로 실험한 결과

d_{eye}	검출된 장 수(총 2408장)	검출률
0.05	1449	60.17%
0.10	2310	95.93%
0.15	2355	97.80%
0.25	2397	99.00%



+ 제안한 방법으로 검출된 좌표

그림 5. 눈동자 검출 결과

표 1은 제안한 방법으로 실행한 결과를 나타낸 표이며, 그림 5는 제안한 방법으로 눈동자를 검출한 결과 영상이다. 제안한 방법으로 실험한 결과 $d_{eye} \leq 0.10$ 일 때 95.93%의 검출률이 나왔으며 오류에 해당하는 4.07%의 영상은 대부분 눈을 감은 영상이다. 오차가 발생한 눈을 감은 영상의 ground truth 값은 그림 6에서와 같이 눈꺼풀을 나타내지만 제안한 방법으로 검출하였을 경우 그림 6에서와 같이 속눈썹 부분을 검출하기 때문에 d_{eye} 가 0.10보다 크게 나타났다.

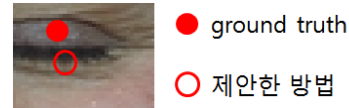


그림 6. 감은 눈 영상의 결과

V. 결 론

본 논문에서는 눈동자 영역을 검출하는 방법과 검출된 눈동자 영역에서 정확한 눈동자를 검출하는 방법을 제안하였다. 눈동자 영역을 검출하기 위해 주성분 분석 방법을 사용하였으며 눈동자를 검출하기 위해 명암도 값과 프로jection 함수를 사용하는 방법을 제안하였다. 제안한 방법은 다양한 얼굴 영상 혹은 기울어진 얼굴 영상에서 안정적인 성능을 보여주었으며 특히 주성분 분석을 이용하여 기울어진 정도를 측정함으로써 성능을 향상시킬 수 있었다.

하지만 감은 눈의 경우 ground truth를 어디로 하느냐에 따라 결과가 차이가 많이 발생하므로 감은 눈의 경우 새로운 검증 방법을 필요로 할 것이다.

참고문헌

- [1] A.L. Yuille, P.W. Hallinan, D.S. Cohen, "Feature extraction from faces using deformable templates," *Int'l J. Comput. Vision* 8(2), pp. 99-111, 1992.
- [2] A. Pentland, B. Moghaddam, T. Starner, "View-based and modular eigenspaces for face recognition," *Proc IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'94)*, pp.84-91, 1994.
- [3] G.C. Feng, P.C. Yuen, "Variance projection function and its application to eye detection for human face recognition," *Int'l J. Comput. Vision* 19, pp.899-906, 1998.
- [4] P. Viola, M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp.511-518, 2001.
- [5] R. Lienhart, J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection," *Proc. IEEE ICIP*, pp.900-903, 2002.
- [6] Lindsay I Smith, "A tutorial on Principal Component Analysis," www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf, 2002.