

주성분분석을 통한 실시간 얼굴 검출 및 추적

박상용*, 이창우**, 이연철**, 김항준**
 *LG 전자㈜
 **경북대학교 컴퓨터공학과
 e-mail : baksang@lge.com

Real-Time Face Detection and Tracking Using PCA

Sang-Yong Park*, Chang-Woo Lee**, Yun-Chul Lee**, Hang-Jun Kim**
 *LG Electronics. Co. Ltd.
 **Dept. of Computer Engineering, Kyungpook National University

요 약

본 논문은 주성분분석(Principal Component Analysis, PCA)을 통하여 실시간 얼굴 검출 및 추적 방법을 제안한다. 제안된 방법은 얼굴 영역 검출과 추적의 두 단계로 구성되어 있다. 검출 단계에서는 피부 색깔 모델과 움직임 정보를 이용하여 얼굴 후보 영역들을 검출하고, 검출된 후보 영역들을 주성분 분석을 통하여 검증한다. 추적 단계에서는 검출된 얼굴들 중에서 현재 추적 중인 얼굴과 가장 유사한 얼굴을 찾아 전체 영상의 중심에 위치하도록 pan/tilt 위에 놓여진 카메라를 제어하여 추적한다. 제안된 방법은 잡음이 많은 배경 상황에서도 좋은 실험 결과를 보여준다.

1. 서론

사람 얼굴 영역을 검출하고 추적하는 것은 얼굴 인식, 얼굴 표정 인식, 입술 읽기(Lip Reading), 감시 시스템들의 필수적인 요소이다. 일반적으로 추적 시스템은 얼굴 검출과 추적의 두 단계로 구성된다. 실시간 카메라 영상에서 얼굴 영역을 검출하는 단계는 위에서 언급한 시스템들의 필수적인 요소이지만, 카메라 영상 내에 포함된 잡음(Noise)과 주변 환경의 변화 때문에 신뢰할 만한 결과를 얻는 것이 어렵다. 현재까지 얼굴 검출을 위하여 model-based, feature-based, sample-based 얼굴 검출방법 등이 제안되어왔다[1-5]. 그러나, 영상내의 조도 변화, 주변 환경의 변화, 영상 획득 단계에서 포함된 잡음 등으로 인해 많은 제약 조건들을 가지고 있다.

본 논문에서 실시간 Active 카메라 영상에서 얼굴 후보 영역들을 검출하고, 검출된 얼굴 후보 영역들을 주성분분석을 이용하여 검증하고 추적하는 시스템을 제안한다. 제안된 시스템은 그림 1 과 같이, 얼굴 검출 단계와 얼굴 추적 단계로 구성된다. 얼굴 검출 단계에서는 색깔 정보와 움직임 정보를 이용하여 얼굴 후보 영역들을 검출한다. 영상에 포함된 잡음으로 인해 색깔 정보만으로 얼굴을 검출하는 것이 어렵기 때문에 움직임 정보도 사용한다. 이렇게 검출된 후보

영역들은 모폴로지 연산과 연결성분 분석과정으로 이루어진 후처리 과정을 거친다. 그리고, 검출된 후보 영역들은 직교공간(Orthogonal Space)로 투영되고, 투영된 이미지를 학습 이미지와 비교하여 얼굴을 검증한다. 얼굴 추적 단계에서는, 검증된 얼굴들 중에서 추적중인 얼굴과 가장 유사한 얼굴을 찾아 카메라의 pan / tilt 를 제어함으로써 추적한다.

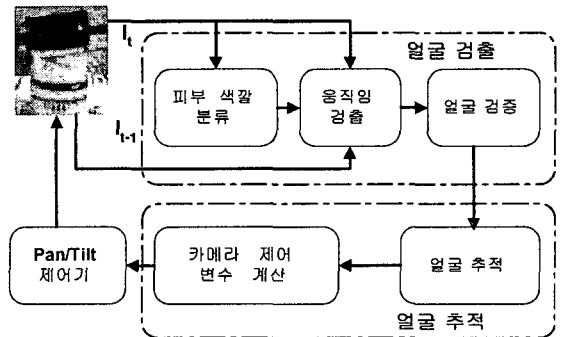


그림 1. 제안된 방법의 블록도

2. 얼굴 검출

얼굴의 검출은 얼굴 후보 영역 검출과 주성분분석을 통한 검증으로 이루어진다. 얼굴 후보 영역 검출을 위하여 실시간 카메라에서 캡처(Capture)한 영상은 일반적으로 많은 잡음이 포함되어 있기 때문에 피부 색깔 모델만으로 얼굴 영역을 검출하는 것이 어렵다. 본 논문에서는 얼굴 후보 영역을 검출하기 위해 피부 색깔 모델을 이용하여 피부 색깔 영역을 분할 하고, 분할 된 영역에 대하여 움직임 정보를 추가적으로 이용한다.

피부 색깔 모델은 색채 칼라 공간(Chromatic color space)에서 얼굴 피부 색깔의 Cr, Cg 성분이 2D-가우시안 모델(2D-Gaussian Model)을 따른다고 가정하고 근사화 시킨 모델이다. 사람 얼굴의 피부 색깔 분포 분석을 통한 Cr, Cg 성분의 평균과 공분산은 아래 수식 (1), (2)와 같이 계산한다[4, 5].

$$m = (\bar{r}, \bar{g}) \text{ where } \bar{r} = \frac{1}{N} \sum_{i=1}^N r_i, \text{ and } \bar{g} = \frac{1}{N} \sum_{i=1}^N g_i. \quad (1)$$

$$\Sigma = \begin{bmatrix} \sigma_{rr} & \sigma_{rg} \\ \sigma_{gr} & \sigma_{gg} \end{bmatrix} \quad (2)$$

피부 색깔을 정의하기 위하여 동양인 30 명을 대상으로 200 개 이미지의 피부 색깔을 추출하였다. 표 1 은 추출된 피부 색깔들로부터 일반적인 사람의 얼굴 피부 색채 분포를 구하기 위해 2D 가우시안 분포로 근사화한 파라미터들이다.

표 1. 2D 가우시안 파라미터들

Parameters	Values
μ_r	117.588
μ_g	79.064
σ_r^2	24.132
$\rho_{X,Y} \sigma_r \sigma_g$	-10.085
$\rho_{X,Y} \sigma_r \sigma_g$	-10.085
σ_g^2	8.748

얼굴 영역을 검출하기 위한 과정은 아래와 같다.

단계 1 : 연속적인 입력 영상 $I_{t-1}(x,y)$, $I_t(x,y)$ 을 캡처한다.

단계 2 : $I_t(x,y)$ 영상의 모든 픽셀을 피부 색깔과 비 피부 색깔로 분류한 이진화 영상을 생성한다.

단계 3 : 피부색깔 이진화 영상에 해당하는 $I_{t-1}(x,y)$ 와 $I_t(x,y)$ 의 차이영상(Frame difference)을 수식 (3)을 이용하여 생성한다.

단계 4 : 차이 영상을 모폴로지 연산과 연결성 성분 분석을 통하여, 영역(Region)화 한다.

단계 5 : 단계 4의 후보 영역에서 얼굴의 구조적 정보를 이용하여 얼굴 후보 영역을 검출한다.

$$M_t(x,y) = \begin{cases} 1, & I_t(x,y) \in S, \text{ \& } |I_t(x,y) - I_{t-1}(x,y)| > \theta \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

단계 5에서는 추출된 후보 영역에서 수평/수직 비율이 얼굴의 비율과 유사한 영역들을 얼굴 후보 영역으로 추정 검출한다.

검출된 얼굴 후보 영역들을 검증하고 추적하기 위하여 고유공간 상에서 각 영역의 가중치 벡터를 사용한다. $N \times N$ 크기의 이미지는 N^2 차원의 이미지 공간에서 한 점으로 표현할 수 있고, 이때 각각의 차원은 한 픽셀에 대응된다. 이러한 이미지를 심각한 정보의 손실이 없이 낮은 차원으로 표현하기 위하여 주성분분석 방법을 사용한다. M 개의 학습 이미지들로부터 M' 개의 고유얼굴($E_1, E_2, E_3, \dots, E_{M'}$)를 구하고, 특정 입력 이미지는 직교공간에서 M' 개의 고유얼굴들의 선형 조합(A linear combination)으로 표현될 수 있다[6, 7].

고유얼굴을 구성하는 방법은 다음과 같다. M 개의 학습이미지는 고유공간을 구성하기 위하여 열 벡터(Column vector) 인 $I_1, I_2, I_3, \dots, I_M$ 로 나타내고, 이들 입력 이미지의 평균 이미지인 $A = 1/M \sum_{m=1}^M I_m$ 를 구한다.

각 학습이미지와 평균 이미지와의 차이 벡터인 $\Phi_i = I_i - A$ 들로 이루어진 새로운 벡터집합 $Y = [\Phi_1 \Phi_2 \dots \Phi_M]$ 를 구한다. 얼굴 이미지의 분포를 최적으로 표현하는 M 개의 직교(Orthogornal) 벡터 u_k 를 구하기 위하여, 공분산 행렬(Covariance matrix)을 다음과 같이 구한다.

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = YY^T \quad (4)$$

C 행렬의 크기는 $N^2 \times N^2$ 으로 계산상의 효율성을 위하여 수식 (5)와 같이 C'을 사용한다.

$$C' = \frac{1}{M} \sum_{i=1}^M \Phi_i^T \Phi_i = Y^T Y \quad (5)$$

C' 행렬로부터 구한 M 개의 고유벡터(Eigenvector) v_k 와 고유값(Eigenvalue) λ_k 를 이용하여 u_k 를 다음과 같이 구한다.

$$u_k = \frac{Y \times v_k}{\sqrt{\lambda_k}} \quad k=1 \dots M \quad (6)$$

λ_k 의 값에 따라 u_k 중에서 중요한 M' 개를 선택하여 m 개의 학습 이미지에 대해서 고유벡터의 선형 조합, $W_m^{training} = [w_1, w_2, \dots, w_{M'}]$ 을 구한다.

$$w_k = u_k^T \Phi_i \quad k=1 \dots M' \quad (7)$$

얼굴 후보 영역의 검증을 위하여 k 개의 검출된 영역을 직교공간으로 투영하여 얻은 $w_k^{candidate}$ 값을 학습 이미지들의 얼굴 클러스터의 평균 \bar{W}_{face} 값(수식 (8))

과 얼굴이 아닌 클러스터의 평균 $\bar{W}_{nonface}$ (수식(9))를 수식 (10)과 같이 비교한다. 비교한 값이 학습 얼굴 이미지 클러스터에 가까우면 검증된 얼굴 영역으로 간주한다.

$$\bar{W}_{face} = \frac{1}{N_1} \sum_{i=1}^{N_1} W_i^{face} \quad (8)$$

$$\bar{W}_{nonface} = \frac{1}{N_2} \sum_{i=1}^{N_2} W_i^{nonface} \quad (9)$$

$$\min(\|W_k^{candidate} - \bar{W}_{face}\|, \|W_k^{candidate} - \bar{W}_{nonface}\|) \quad (10)$$

3. 얼굴 추적

얼굴의 추적을 위하여 이전 프레임에서 추적되고 있던 얼굴이미지의 W_{old} 값과 검출된 얼굴 영역들의 W_k 값 사이의 유클리드 거리(Euclidian distance)가 최소가 되는 얼굴 영역을 구한다.

$$obj = \arg \min_k \|W_{old} - W_k\| \quad (11)$$

추적할 얼굴 영역이 찾아지면, 전체 영상의 중심과 추적할 얼굴 영상의 중심간의 거리 벡터 $dist, (fa, scr)$ 를 이용하여 pan/tilt의 방향과 동작 시간을 계산하여 검출된 얼굴의 중심을 전체 영상의 중심에 위치하도록 제어한다.

$$dist, (fa, scr) = Face, (x, y) - Screen(h/2, w/2) \quad (12)$$

아래는 거리 벡터로부터 제어 파라미터를 결정하는 의사코드(Pseudo Code)이고, 표 2는 얼굴의 추적에 사용되는 pan/tilt의 제어 파라미터들이다.

제어 파라미터 계산을 위한 의사코드

```

Procedure Duration (x, y)
Begin
    Siga = None;
    distance =  $\sqrt{x^2 + y^2}$ ;
    IF distance >  $\theta_{close}$  then
        Siga = Close;
    ELSEIF distance >  $\theta_{far}$  then
        Siga = far;
    Return (Siga);
End Duration;
    
```

```

Procedure Orientation (x, y)
Begin
    Sigo = None;
    IF x >  $\theta_x$  then
        Add "RIGHT" to Sigo;
    ELSEIF x <  $-\theta_x$  then
        Add "LEFT" to Sigo;
    IF y >  $\theta_y$  then
        Add "UP" to Sigo;
    ELSEIF y <  $-\theta_y$  then
        Add "DOWN" to Sigo;
    Return (Sigo);
    
```

End Orientation;
표 2. Camera control parameters

Duration		Orientation (Pan/Tilt)
None	0 sec (0 degree)	Up, Down, Left, Right, Up-Left, Up-Right, Down-Left, Down-Right
Close	0.5 sec (3.1 degree)	
Far	1 sec (6.2 degree)	

4. 실험 결과

제안된 방법의 실험 환경은 조도의 변화가 심하고, 배경에 피부 색깔과 유사한 잡음이 많이 포함된 실내 환경을 대상으로 하였다. 본 시스템에서 사용한 pan/tilt는 삼성전자의 SPT-2410을 사용하였고, pan/tilt를 PC에서 제어가 되도록 8255 카드를 이용한 제어 회로를 제작하였다. pan 속도는 6.2°/sec이고, tilt 속도는 4.2°/sec이며, 사용된 PC는 PentiumII-400Mhz이다. 그림 2는 실험환경과 Window 98에서 VC++6.0으로 제작한 시스템 인터페이스 창을 보여준다.



(a) 실험환경 (b) 인터페이스 창

그림 2. 실험환경 및 인터페이스 창

얼굴 영역의 검출을 위해 연속된 카메라 영상에서 320x240의 칼라 영상을 초당 10 프레임씩 캡처하여 사용하였다. 학습 이미지는 10명의 각각 다른 각도에서의 얼굴이미지 5장과 팔과 손등 얼굴이 아닌 이미지 50장을 사용하였다. 그림 3은 사용된 얼굴 학습 이미지 중에서 일부를 보여준다.

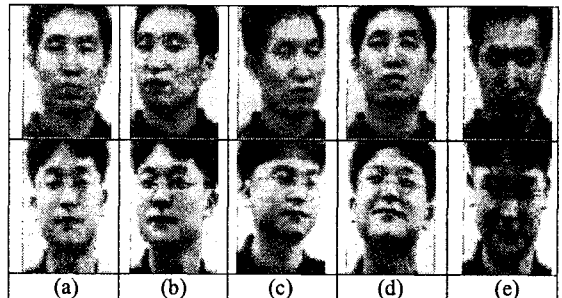


그림 3. 얼굴 학습 이미지

제안된 방법은 5분씩 20번의 실험을 하였으며, 표 3은 실험 중 캡처된 영상을 식 (13)와 (14)을 이용하여 분석한 결과를 나타낸다.

$$\text{Face verification rate} = \frac{\text{Number of correctly verified faces}}{\text{Number of images verified as true face}} \quad (13)$$

$$\text{Non-face verification rate} = \frac{\text{Number of correctly verified images as non_faces}}{\text{Number of images verified as non_face}} \quad (14)$$

표 3. 검증율 (Verification Rate)

Face	Non-Face	Total
96.5 %	92%	94.3%

그림 4 는 카메라의 입력으로부터 얼굴 영역을 검출하고, 검출된 영역을 추적하는 과정을 보여준다.

5. 결론

본 논문에서는 실시간 Active 카메라 영상에서 얼굴 영역을 검출하고, 검출된 얼굴 영역을 검증하여 추적하는 시스템을 제안하였다. 실시간 Active 카메라 영상에서 얼굴 영역의 검출을 위하여 칼라 정보와 움직임 정보를 이용하였고, 검출된 얼굴 영역을 주성분분석을 이용하여 검증하였다. 검증된 얼굴들 중 이전에 추적되고 있던 얼굴과 가장 유사한 얼굴을 전체 영상의 중심에 위치하도록 위치정보를 이용하여 실시간으로 Active 카메라를 제어함으로써, 검출된 얼굴을 추적하였다. 실험 결과 복잡한 배경 환경과 잡음

영상에서도 94.3%의 검증율을 보였다. 향후 과제로는 현재의 방법의 정확성 및 견고성을 높이고, 검출된 얼굴 영역을 이용하여 얼굴 판별 및 인식에 관하여 연구하고자 한다.

참고문헌

- [1] O. Ayinde, Y. Yang, "Region-based face detection", *Pattern Recognition*, in press
- [2] A. Yullie, P. Hallinan, D. Cohen, "Feature extraction from faces using deformable templates", *Int. J. Computer Vision*, Vol. 8, no. 2, (1992) pp.99-111
- [3] G. H. Rowley, S. Baluja, T. Kanade, "Neural network-based human face detection", *IEEE Trans. on Pattern Anal. Mach. Intell.*, Vol 20, no. 1, 1998, pp. 23-38
- [4] Jie Yang, Waibel, A., "A real-time face tracker", *Applications of Computer Vision, 1996. WACV '96, 1996*, pp. 142-147
- [5] S. H. Park, E. Y. Kim, S. W. Hwang, Y. C. Lee, H. J. Kim, "Face detection for security system on the internet", *Proceeding of IEEE Consumer Electronics, 2001, 2001*, pp. 276-277
- [6] M. Turk, A. Pentland, "Face Recognition Using Eigenfaces", *IEEE Proceedings CVPR '91, 1991*, pp. 586-591
- [7] Guan, A.X, Szu, H.H, "A local face statistics recognition methodology beyond ICA and/or PCA", *International Joint Conference on Neural Network, 1999*, pp.1016-1021



그림 4. 실험 결과