

# 비강압적 홍채 인식을 위한 전 방향 카메라에서의 다각도 얼굴 검출

\*이현수, \*배광혁, \*김재희, \*\*박강령  
\*연세대학교 전기전자공학과, 생체인식연구센터  
\*\* 상명대학교 미디어학과, 생체인식 연구센터  
e-mail : alary@yonsei.ac.kr

Multi-views face detection in Omni-directional camera  
for non-intrusive iris recognition

\*Hyeon-Soo Lee, \*Kwanghyuk Bae, \* Jaihie Kim, \*\*Kang-Ryung Park  
\*Dept of Electric and Electronic Engineering , Yonsei University  
Biometrics Engineering Research Center,  
\*\* Division of Media Technology, SangMyung University

## Abstract

This paper describes a system of detecting multi-views faces and estimating their face poses in an omni-directional camera environment for non-intrusive iris recognition. The paper is divided into two parts; First, moving region is identified by using difference-image information. Then this region is analyzed with face-color information to find the face candidate region. Second part is applying PCA (Principal Component Analysis) to detect multi-view faces, to estimate face pose.

## I. 서론

홍채 인식 시스템은 홍채 영역에 존재하는 무늬 패턴(texture pattern)을 이용하여 개인을 식별하는 것으로 홍채 이미지를 획득하기 위한 입력 장치와 획득된 홍채 이미지를 이용해서 개인을 식별하는 인식 장치로 나눌 수 있다. 홍채 이미지를 획득하는 방식은 사용자의 협조가 필요한 수동식, 반자동식과 카메라가 사람의 얼굴과 눈의 위치를 파악해서 비강압적으로 홍채 영상을 취득하는 자동식으로 나눌 수 있다. 일반적인 field of view)와 협시야 장치(narrow field of view)의 두 부분으로 구성되어 있다. 광시야 장치 부분에서는 사용자들의 접근을 파악하고 사용자의 얼굴을 검출하

여 눈 위치 정보를 획득한다. 협시야 장치 부분에서는 광시야 장치 부분에서 찾아낸 홍채 인식 대상자의 눈 위치 정보를 바탕으로 홍채 인식이 가능한 눈 영상을 획득한다. 본 논문에서 사용된 시스템은 광시야 장치 부분에서 전 방향 렌즈(Omni-directional Lens)를 이용한 비강압적 홍채 인식 시스템을 사용한다[1].

본 논문에서는 전 방향 카메라 환경에서 사용자의 다각도의 얼굴을 검출한다. 전 방향 카메라 환경의 경우 정면 얼굴 영상뿐만 아니라 다각도의 얼굴 영상도 많기 때문에 이를 검출하는 과정이 필요하다. 논문의 전체적인 구성은 다음과 같다. 2장에서 본 시스템의 전체적인 소개를 하고, 3장에서는 얼굴 후보 영역을 찾는다. 그 다음으로 4장에서는 다각도의 얼굴 검출 과정을 설명하고, 마지막으로 5장에서 실험결과를 분석한다.

## II. 시스템 개요와 전 방향 카메라 영상

### 2.1 시스템 개요

전 방향 렌즈를 사용자의 눈높이에 위치시킨다(그림 3(a)). 전 방향 렌즈의 경우 360의 영상이 한 프레임에 들어오기 때문에 영상의 해상도가 제한된 부분을 영상으로 받는 일반 카메라보다 해상도가 낮다. 이를 보완하기 위해 사용자의 눈높이에 위치시킨다.

본 시스템의 전체적인 과정은 그림1과 같다. 전체적인 과정은 움직임 정보와 색깔 정보를 이용해서 얼굴

후보 영역을 추출하고, 주성분 분석 방법을 사용해서 얼굴 후보 영역에 대한 얼굴 검출과 얼굴 포즈를 예측한다. 각각의 단계는 새로운 움직임이 나타날 때마다 계속적으로 진행된다.

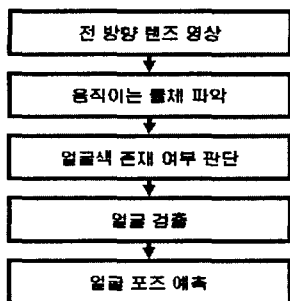


그림 1 시스템 개요

### 2.2 전 방향 카메라 영상

전 방향 렌즈에서 얻어진 영상은 그림 3(b)에서와 같다. 이 영상에서는 물체의 위치가 좌표계로 표현되기 때문에 연산이 복잡할 뿐만 아니라 영상의 왜곡이 발생할 수 있다. 그래서 영상을 그림 3(c)과 같이 수직 좌표계로 변환한 후에 다음 과정을 수행한다.



(a)

(b)



(c)

그림 3 전방향 렌즈 영상 : (a) 전 방향 렌즈 (b) 전 방향 렌즈 영상 (b) 펼쳐진 전 방향 렌즈 영상

수직 좌표계로 변환하는 과정에서 원 영상의 안쪽 경계와 바깥쪽 경계의 길이가 다르기 때문에 안쪽 경계의 길이를 바깥쪽 경계의 길이와 같게 펼치는 작업(Stretching)을 한다. 이 과정에서의 왜곡을 줄이기 위해 선형 보간법(Linear Interpolation)을 사용한다.

## III. 얼굴 후보 영역 검출

### 3.1 움직임을 이용한 후보 영역 추출

전 방향 렌즈를 사용하면 그림3(b) 영상과 같이 공간에서 360도의 전체 영역이 입력 영상으로 주어진다. 이 부분은 매우 넓기 때문에 얼굴 후보 영역을 찾기 위해 색깔 정보만을 이용하면 후보 영역이 너무 많이 발생해 이를 확인하는 과정에 많은 시간이 소요된다. 따라서 전체 영역 중에서 움직임이 있는 일부 영역만을 구별해 얼굴 후보 영역을 찾기 위한 탐색 영역을 축소시킨다.

움직임 정보를 얻기 위해서 가장 일반적으로 사용되는 누적 차영상(Absolute Accumulated Difference Picture)을 이용한다[2]. 이는 지속적인 움직임과 잡음을 구별하기 위해 차영상을 계속 누적시키는 방법이다. 이 과정에서 현재 프레임의 차영상과 지난 프레임의 차영상의 가중치를 다르게 하여 지난 차영상의 정보 비중을 현재 차영상보다 작게 한다. 이를 통해 지난 영상에서 발생한 잡음의 영향을 현재 누적 차영상에서는 줄일 수 있다. 그림 4(a)에 펼쳐진 영상에서 움직임 차를 이용해서 구한 누적 차영상이다.

$$AADP_n(x, y) = \alpha AADP_{n-1}(x, y) + \beta DP(x, y, n) \quad (1)$$

$$DP(x, y, n) = \begin{cases} 1 & \text{if } |F(x, y, n) - F(x, y, n-1)| > T \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$AADP_n$  : n번째 프레임에서의 누적 차영상

$DP$  : 차영상

$\alpha, \beta$  : 가중치

$F(x, y, n)$  : n번째 프레임의 (x,y) 지점에서의 밝기 값

### 3.2 컬러 정보를 이용한 후보 영역 추출

사람의 피부색은 다른 배경의 색과 다른 고유한 값을 가진다. 이런 특성을 이용해서 얼굴 영역을 구별할 수 있다. 얼굴 영역을 구별하기 위해 가장 많이 사용되는 좌표계는 HSV(Hue:색상, Saturation:채도, Value:밝기)이다. HSV좌표계에서 H와 S에서 사람 피부색의 값은 정규 분포(Normal Density)를 이룬다는 것이 증명되었다[3]. 이런 특징을 이용해서 움직임이 있는 영역에서만 HSV 좌표계로 변환해 얼굴 영역과 비 얼굴 영역을 구분한다. 그림 4(b)는 그림 4(a)에서 찾은 움직임 영역에 대해서만 HSV 좌표계를 이용해서 얼굴 후보 영역을 찾은 예이다. 검은 색으로 표시된 부분이 얼굴 후보 영역에 해당한다.

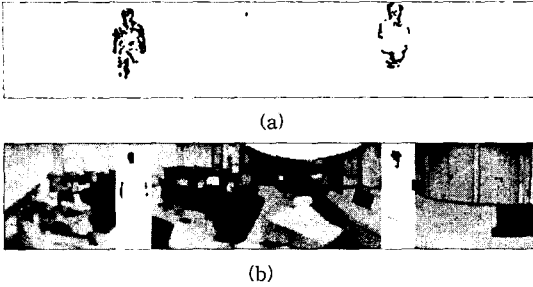


그림 4 얼굴 후보 영역 추출 : (a) 누적차영상  
(b) 움직임 영역에서 색깔 정보 이용

#### IV. 다각도 얼굴 검출

움직임과 컬러 정보를 이용해 얻은 얼굴 후보 영역에는 실제 얼굴 영역뿐만 아니라 비얼굴 영역도 포함되어 있다. 또한 전 방향 렌즈 환경에서 자연스럽게 움직이는 사용자의 경우 정면 얼굴(frontal face) 뿐 아니라 측면 얼굴(profile face)도 많이 나타나기 때문에 측면 얼굴을 검출하고 얼굴 포즈를 예측하는 과정이 필요하다.

##### 4.1 정면 얼굴 검출

주성분분석 방법을 이용한 얼굴 검출은 입력  $\mathbf{x}$ 를 얼굴 영역의 고유 공간에 투영시킨 분포값과 비얼굴 영역의 고유 공간에 투영시킨 분포값을 비교해서 결정한다 [4]. 얼굴 영역의 고유 공간은 얼굴 영상에서 학습하는 얼굴 벡터의 변위를 가장 크게 하는  $M$ 개의 고유 벡터를 선택해서  $M$ 차원의 고유 공간(eigen space),  $F$ 를 생성한다. 또한 선택되지 않은  $N - M$ 개의 고유 벡터를 이용해서 주성분과 직교하는 고유 공간  $\bar{F}$ 를 정의한다. 식(1)과 같이 입력  $\mathbf{x}$ 를 고유 공간에 투영시킨  $\mathbf{y}$ 의 고유 공간에 대한 얼굴 영역에 대한 분포값은 식(2)로 결정된다.

$$\mathbf{y} = \Phi^T (\mathbf{x} - \mathbf{m}_x) \tag{1}$$

$$P(\mathbf{y} | \Omega_1) = \frac{\exp\left(-\frac{1}{2} \sum_{i=1}^M \frac{y_i^2}{\lambda_i}\right)}{(2\pi)^{M/2} \prod_{i=1}^M \lambda_i^{1/2}} \left[ \frac{\exp\left(-\frac{1}{2} \sum_{i=M+1}^N \frac{y_i^2}{\lambda_i}\right)}{(2\pi)^{M/2} \prod_{i=M+1}^N \lambda_i^{1/2}} \right] \tag{2}$$

$$= P_F(\mathbf{y} | \Omega) P_{\bar{F}}(\mathbf{y} | \Omega)$$

##### 4.2 다각도 얼굴 검출

공간  $F$ 에서 포즈별 분포는  $-90$ 도에서부터  $90$ 도 각도까지 모두 연속적으로 분포되어 있는 것이 증명되었다 [4]. 고유 공간을 이용해서 얼굴 포즈를 결정하는 방법은 각 포즈별의 분리된 고유 공간을 생성하는 포즈별 접근 방법(view-based formulation)이다. 이 방법은 포즈별 부분 공간에서 각각의 부분 공간을 대표할 수 있는 고유 공간을 생성한다. 그리고 입력 벡터  $\mathbf{y}$ 에 대해서 각각의 고유 공간에 투영시켜서 가장 가까운 고유 공간의 얼굴 포즈로  $\mathbf{y}$ 의 포즈를 결정하는 것이다. 즉,  $N_c$ 개의 포즈가 존재하면 독립된  $N_c$ 개의 고유 공간이 생성한다. 그리고 입력  $\mathbf{x}$ 를 각각의 고유 공간에 투영시키고, 공간 안에서의 분포가 가장 큰 포즈를 선택한다(그림 5). 그러나 이 방법의 경우 모든 입력 영상에 대해서 모든 포즈 공간과의 연산을 해야 하기 때문에 계산량이 상당히 많다는 문제점이 있다.

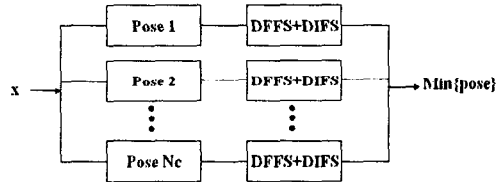


그림 5 얼굴 포즈 결정

##### 4.3 제안하는 포즈 결정 방법

얼굴 포즈별 부분 공간은 얼굴 고유 공간  $F$ 에서 연속적인 분포를 가지기 때문에 몇 개의 포즈별 부분 공간만을 이용하여 얼굴 포즈를 추측할 수 있다. 본 논문에서 고유 공간을 생성하는 포즈는  $p_1, p_2$ 이지만 이들의 고유 공간과의 거리를 비교하여 포즈  $p_1$ 과 포즈  $p_2$ 의 사이 포즈를 추정한다. 포즈  $p_1$ 의 고유 공간에서 포즈  $p_1$ 과 포즈  $p_2$ 의 사이 포즈의 분포는 비슷하다. 또한 포즈  $p_2$ 의 고유 공간에서 포즈  $p_2$ 과 포즈  $p_1$ 과 포즈  $p_2$ 의 사이 포즈의 분포는 비슷하다. 이를 이용하여 입력 영상의 포즈를 구분한다. 전체 과정은 다음과 같다.

1. 입력 벡터  $\mathbf{x}$ 에 대하여 포즈별 고유 공간  $\Psi_{p_1}, \Psi_{p_2}, \dots, \Psi_{p_w}$ 에 투영시킨다.

$$\mathbf{y}_{p_1} = \Psi_{p_1}^T \mathbf{x}$$

$$\mathbf{y}_{p_2} = \Psi_{p_2}^T \mathbf{x}$$

$$y_{p_n} = \Psi_{p_n}^T x$$

2. 각 고유 공간에 대해서 조건부 확률을 구해 얼굴 여부를 판단한다.

$$P(x | \Omega_{\psi_n}) > P(x | \Omega_E)$$

$$P(x | \Omega_{\psi_{n+1}}) > P(x | \Omega_E)$$

$$P(x | \Omega_{\psi_{p_n}}) > P(x | \Omega_E)$$

3. 2번 조건을 만족시키는 포즈가 한개 일 경우, 해당 하는 포즈로 결정한다.

4. 2번 조건을 만족시키는 포즈가 이웃하는 두개의 포즈일 경우 두 포즈의 고유 공간에서의 확률값의 비 ( $\alpha$ )를 통해 결정한다.

$$\frac{P(x | \Omega_{\phi_{pl}})}{P(x | \Omega_{\phi_{pm}})} = \alpha$$

### V. 실험 결과

전 방향 렌즈를 사람의 눈높이인 1.75m 정도에 위치시키고, 692×520의 영상을 Pentium IV 2.4GHz 환경에서 실험했다. 주성분분석 방법을 위한 얼굴 학습 영상은 얼굴 크기가 100×100이고, 한국인의 20대로 한정된 30명의 사람에서 각 포즈 당 60장씩 총 23400장의 영상을 획득했다. 각 영상은 왼쪽 90도에서 오른쪽 90도 까지 15도의 간격으로 영상을 획득했다. 학습 영상은 사람들마다 눈, 코, 입 등의 몇 개의 특징점이 포즈별로 일정한 위치에 위치하도록 하고 20×20으로 얼굴 영상을 축소해 사용했다. 포즈별 고유 공간은 정면 영상, 오른쪽 30도, 왼쪽 30도, 오른쪽 60도, 왼쪽 60도를 생성하였다. 표1은 기존의 방법과 제안한 방법의 영상 1장당 얼굴 검출과 포즈 판단에 소요된 시간이다. 표2은 기존의 방법을 적용시킨 얼굴 포즈 판단의 결과다. 표에서 기타는 얼굴 검출에 실패한 것이고, +는 왼쪽, -는 오른쪽을 나타낸다. 표3는 제안한 방법으로 예측한 포즈 판단의 결과이다. 포즈 판단의 정확도는 비슷하나 소요시간이 상당히 감소한 것을 살펴 볼 수 있다.

표 1. 영상 1장당 처리 시간 비교

	기존의 방법	제안한 방법
소요시간(ms)	4.218	2.75

표 2. 기존의 방법을 이용한 얼굴 포즈 판단

	정면	+15도	-15도	+30도	-30도	기타
정면	94	1	5	0	0	0
+15도	23.2	55	0.7	16	0	5
-15도	32	0	40	0.7	20.6	7.4
+30도	1.45	21	0.7	68.9	0	8
-30도	6	0	19	0	69.4	6

표 3. 제안한 방법을 이용한 얼굴 포즈 판단

	정면	+15도	-15도	+30도	-30도	기타
정면	80.2	19	0	0	0	1
+15도	20.3	50.7	0	21.8	0	7.1
-15도	20	0	60.3	0	11	8
+30도	2	13	0	70.3	0	14.3
-30도	3.7	14.2	0.5	0	73	9

### VI. 결론

본 논문에서는 비강압적 홍채 인식을 위해 전 방향 카메라를 사용한 광시야 장치 부분에서 사용자의 움직임을 파악하고 다각도의 얼굴을 검출하였다. 제안한 얼굴 포즈 예측 방법은 기존의 방법에 비해 유사한 정확도를 보이면서도 처리 시간을 감소시켰다. 그러나 사용자의 움직임이 겹칠 경우 사용자를 구분하지 못하는 문제점이 있다. 따라서 이를 해결하기 위한 추가 연구가 요구된다.

감사의 글

본 연구는 한국과학기술재단 생체인식 연구센터(BERC)의 지원으로 수행되었습니다.

### 참고문헌

[1] S.K. Nayac, "catadioptric omni-directional Camera", IEEE Computer Society conf. on CVPR, pp 482-488, 1999

[2] Ahmed Elgammal, Ramani Duraiswami, David Harwood, Larry S. Davis, "Background and Foreground Modeling Using Non-parametric Kernel Density Estimation for Visual Surveillance", Proceeding of The IEEE, VOL. 90, No.7, 2002.

[3] Ying wu, "Nonstationary Color Tracking for Vision-Based Human-Computer Interaction", IEEE Trans. Neural Network, VOL. 13, NO. 4, July, 2002

[4] Turk and A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, No.1, pp.71-86, 1991.