

## Rotation Invariant Face Detection Using HOG and Polar Coordinate Transform

Kyung-Shik Jang\*

\*Professor, Division of Creative Software Eng., Dong-eui University, Busan, Korea

### [Abstract]

In this paper, a method for effectively detecting rotated face and rotation angle regardless of the rotation angle is proposed. Rotated face detection is a challenging task, due to the large variation in facial appearance. In the proposed polar coordinate transformation, the spatial information of the facial components is maintained regardless of the rotation angle, so there is no variation in facial appearance due to rotation. Accordingly, features such as HOG, which are used for frontal face detection without rotation but have rotation-sensitive characteristics, can be effectively used in detecting rotated face. Only the training data in the frontal face is needed. The HOG feature obtained from the polar coordinate transformed images is learned using SVM and rotated faces are detected. Experiments on 3600 rotated face images show a rotation angle detection rate of 97.94%. Furthermore, the positions and rotation angles of the rotated faces are accurately detected from images with a background including multiple rotated faces.

▶ **Key words:** Rotated face detection, Polar coordinate transform, Learning, HOG, SVM

### [요 약]

이 논문에서는 회전각도에 무관하게 회전 얼굴과 회전각도를 효과적으로 검출하는 방법이 제안되었다. 회전된 얼굴 검출은 회전에 따른 얼굴 외형의 큰 변화로 인해 어려운 분야이다. 제안한 극좌표계 변환 방법은 회전 각도와 무관하게 얼굴 구성요소들의 위치 정보가 유지되기 때문에 회전으로 인한 얼굴의 외형 변화가 없게 된다. 이에 따라 회전이 없는 정면 얼굴 검출에 사용되지만 회전에 민감한 특성을 갖는 HOG와 같은 특징들이 회전얼굴을 검출하는 과정에서 효과적으로 사용될 수 있다. 극좌표계 변환된 영상에서 얻은 HOG 특징을 SVM을 이용하여 학습하고 회전 얼굴을 검출하였다. 학습 데이터는 회전이 없는 정면 얼굴 영상만을 사용하였다. 3600개 회전 얼굴 영상에 대한 실험 결과 97.94 %의 회전각도 검출률을 얻었다. 또한, 다수의 회전 얼굴이 포함된 배경이 있는 영상들에서 회전 얼굴들의 위치와 회전 각도를 정확하게 검출하였다.

▶ **주제어:** 회전 얼굴 검출, 극좌표계 변환, 학습, HOG, SVM

• First Author: Kyung-Shik Jang, Corresponding Author: Kyung-Shik Jang  
\*Kyung-Shik Jang (ksjang@deu.ac.kr), Division of Creative Software Eng., Dong-eui University  
• Received: 2021. 10. 12, Revised: 2021. 11. 18, Accepted: 2021. 11. 18.

## I. Introduction

얼굴 인식은 컴퓨터 비전(Computer Vision), 패턴 인식 분야에서 가장 중요한 주제 중 하나이며, 이 과정에서 우선적으로 얼굴을 검출하는 것이 필요하다[1]. 얼굴 검출은 얼굴 영상에서 정보를 추출하는 컴퓨터 비전 시스템에서 매우 중요한 구성요소 역할을 한다. 생체인식, 감시 시스템 및 얼굴 표정 인식 등과 같은 실용적인 응용 분야에서 기본적으로 입력 영상에서 얼굴을 빠르고 정확하게 감지해야 하므로 많은 얼굴 검출 방법들이 활발히 연구되어 왔다.

Viola[2]는 AdaBoost 알고리즘으로 Haar-like 특징을 학습하여 얼굴을 검출하였다. Frowl[3]는 얼굴과 비얼굴 학습 이미지들로부터 계산된 MCT(Modified Census Transform) 히스토그램을 구한 후, 입력 영상에서 계산된 값과 비교하여 얼굴을 검출하였다. Rekha[4]는 HOG(Histogram of Oriented Gradient) 특징을 입력 영상에서 구하고 눈, 코, 입과 같은 각 얼굴 특징이 얼굴 형태에 기여하는 가중치를 계산하여 얼굴을 검출하였다.

그러나 이와 같은 방법들은 회전이 없는 정면 얼굴을 대상으로 좋은 결과를 보이고 있지만, 회전 얼굴을 검출하는 것은 회전으로 인한 얼굴의 외형 변화가 크기 때문에 여전히 어려운 분야이다[5]. 제약이 없는 환경에서 인물의 여러 가지 자세가 있을 수 있기 때문에 회전된 얼굴을 검출하는 것은 매우 중요하다.

Liu[6]은 회전된 얼굴을 검출하기 위하여 원형 형판(radial template)을 이용하는 방법을 제안하였다.  $360^\circ$ 를  $20^\circ$  단위로 18개 방향의 부 영역으로 나누고, 얼굴 중심을 기준으로 에지 성분을 방사 투영하여 원형 형판을 구성한다. 입력 영상에서 생성한 원형 형판과 미리 생성한 얼굴의 원형 형판을 비교하여 회전 얼굴을 검출하였다.

Rowly[7]은 회전된 얼굴을 검출하기 위하여 회전이 없는 정면 얼굴을 검출하는 기존의 시스템을 이용하는 방법을 제안하였다. 회전이 없는 정면 얼굴 영상을  $20^\circ$  씩 회전하여 18 방향의 얼굴영상을 생성하여 학습하고 회전된 얼굴을 검출하였다. Tsai[8]는 skin filter와 엔트로피(entropy)를 이용하여 검출한 얼굴 후보 영역에 PCA(Principal Component Analysis)를 적용하여 얼굴 회전각도를 계산하고 회전 얼굴을 검출하였다. Sujay[9]은 회전이 없는 정면 얼굴 영상을 15개 방향으로 회전시킨 영상들에서 LBP 특징을 얻고, SVM(Support Vector Machine)을 사용하여 얼굴을 인식하였다.

Shi[10]은 회전 얼굴을 검출하기 위하여 세 단계로 구성된 PCN(Progressive Calibration Networks)을 사용하였

다. PCN의 각 단계에서는 예측 얼굴 방향과 후보 얼굴의 회전 방향 차이를 조금씩 감소시킴으로써 비얼굴과 얼굴의 이진분류를 하고 회전 얼굴을 검출하였다. Shima[11]는 얼굴의 회전 방향을  $90^\circ$  단위의 4 가지로 분류하고 딥러닝(Deep Learning)을 이용하여 학습하였다. 영상의 특징을 추출하기 위하여 사전 훈련된 컨벌루션 신경망(convolutional neural network)을 사용하고 SVM으로 학습하여 분류하였다.

이 논문에서는 회전각도에 무관하게 회전 얼굴과 회전 각도를 효과적으로 검출하는 방법을 제안하였다. 회전 얼굴을 검출하기 위하여 눈, 입과 같은 얼굴 구성요소들의 상대적 위치 정보를 회전 각도에 무관하게 유지할 수 있는 극좌표계 변환 방법을 제안하였다. 극좌표계 변환 결과는 회전으로 인한 얼굴의 외형 변화가 전혀 발생하지 않는다. 따라서 회전이 없는 정면 얼굴 검출에 사용되지만, 회전에 민감한 특성을 갖는 HOG와 같은 특징들이 회전얼굴을 검출하는 과정에서 효과적으로 사용될 수 있다. 회전이 없는 정면 얼굴영상만을 학습 데이터로 사용하였으며, 극 좌표계 변환된 영상에서 HOG 특징을 얻고 SVM을 이용하여 학습하고 회전 얼굴과 얼굴 회전 각도를 검출하였다.

## II. Related Research

### 1. HOG(Histogram of Oriented Gradient)

HOG는 보행자를 검출하는 과정에서 효과적으로 사용되는 특징이다. 국부 영역에서 인접화소들 간의 밝기 차이를 기반으로 방향성과 기울기를 구하고 이들을 히스토그램으로 구성한 것으로 물체의 형상정보로써 사용된다[12]. HOG 특징은 간단하면서도 사람의 고유한 윤곽선 특징을 추출하기 유리하기 때문에 사람 및 얼굴 검출 분야 등에 많이 활용된다.

히스토그램으로 구성된 HOG 특징은 기하학적 변화에 강건하며, 조명의 영향에 민감하지 않은 특성을 가지고 있다. 그러나 국부 특징만을 사용하여 비교하기 때문에 회전 변환에는 매우 민감한 특성이 있다. HOG는 셀(cell)이라는 최소단위로 영상을 분할하고 국부적 특징을 추출한다. 일반적으로 각 셀은  $8 \times 8$  픽셀들로 구성되고, 한 블록(block)은  $2 \times 2$  셀들로 구성한다.

영상  $I(x, y)$ 의 각 화소에서 식 (1)을 이용하여  $x$  방향과  $y$  방향의 기울기  $g_x, g_y$  를 구한다. 기울기의 크기  $m(x, y)$ 는 식 (2)를 이용하여 구하고 기울기의 방향  $\theta(x, y)$ 는 식 (3)을 이용하여 구한다.

계산된 기울기의 크기  $m(x, y)$ 와 기울기 방향  $\theta(x, y)$ 를 이용하여 각 셀 내부의 기울기 방향에 대한 히스토그램을 작성한다. 기울기 방향 성분은  $0\sim 360^\circ$ 를  $0^\circ\sim 180^\circ$  기준으로 나타내며, 이를 양자화하여 매  $20^\circ$ 씩 9개의 bin으로 구성한다.

$$g_x(x, y) = I(x+1, y) - I(x-1, y) \quad (1)$$

$$g_y(x, y) = I(x, y+1) - I(x, y-1)$$

$$m(x, y) = \sqrt{g_x(x, y)^2 + g_y(x, y)^2} \quad (2)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right) \quad (3)$$

그림 1의 (a)와 같이 셀 내부의 기울기 방향 bin에 해당하는 기울기 크기를 누적시켜 각 셀에 대한 히스토그램을 구성한다.  $2 \times 2$  셀들로 구성되는 블록은 그림 1 (b)와 같이 블록 내부에 있는 셀들의 히스토그램을 하나의 벡터로 연결한다. 블록은 전체 영상에서 셀 하나의 크기만큼씩 슬라이딩 이동하며 생성한 각 블록의 히스토그램을 하나로 병합하여 최종적인 HOG 특징 벡터를 생성한다.

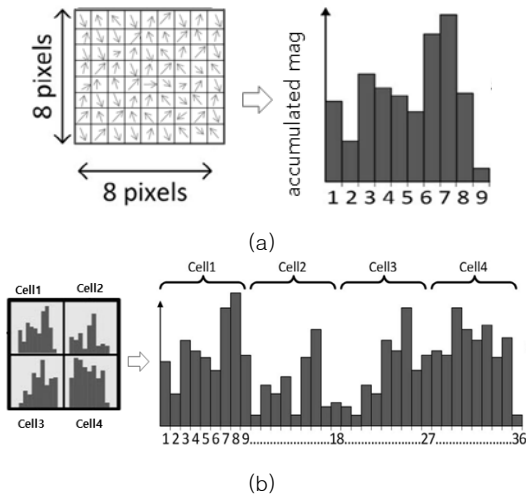


Fig. 1. HOG feature

## 2. SVM(Support Vector Machine)

SVM은 이진 패턴 분류문제를 해결하기 위하여 제안된 지도학습 모델이다. 학습 데이터를 사용하여 두 클래스를 구분하는 결정 경계면인 초평면(Optimal Separating Hyperplane)을 구한다. 초평면에서 가장 가까운 위치에 있는 데이터를 서포트 벡터라고 하며, 결정 경계면과 서포트 벡터와의 여백(margin)을 최대화시켜 두 클래스간의 구분을 최적화한다[13].

이진 분류인 경우 분류 초평면은 식 (4)와 같다.  $w$ 는 분류 초평면과 직교하는 가중치 벡터이고  $b$ 는 바이어스이며, 여백의 크기는  $\frac{2}{|w|}$ 이다. 결정 경계면부터 데이터까지의 거리인 여백이 클수록 좋은 분류기이다. SVM의 학습단계는 여백의 크기가 최대가 되도록 하는  $w$ 와  $b$ 를 얻는 것이다. 초기에 SVM은 선형분류 문제를 해결하기 위한 방법으로 많이 사용되었으며, 이후 특징 공간을 고차원 공간으로 변형시키는 커널을 사용하여 비선형 분류기로 확장되었다. 이 논문에서는 얼굴과 비얼굴 영상에서 얻어진 HOG 특징 벡터를 선형 SVM을 이용하여 학습하고 입력 영상에서 얼굴을 검출하였다.

$$w^T \cdot x_i + b = 0, \quad \forall i = 1, \dots, N \quad (4)$$

## III. The Proposed Scheme

### 1. Polar Coordinate Transform

얼굴뿐만 아니라 모든 물체는 회전에 따른 외형의 변화가 발생하며, 이러한 외형 변화는 회전된 물체의 검출을 어렵게 한다. 이 논문에서는 회전된 얼굴을 검출하기 위하여 입, 눈과 같은 얼굴 구성요소들의 상대적 위치 정보를 회전 각도에 무관하게 유지할 수 있는 극좌표계 변환 방법을 제안하였다. 제안한 극좌표계 변환 방법은 얼굴 요소들간의 공간 정보가 회전 각도와 무관하게 항상 유지되기 때문에 회전 효과가 제거된다.

$$x_1 = C_x + r \cos \theta \quad (5)$$

$$y_1 = C_y + r \sin \theta$$

$$X_1 = \theta - \theta_0$$

$$Y_1 = R - r$$

$$x_2 = C_x + r \cos(\theta + \pi) \quad (6)$$

$$y_2 = C_y + r \sin(\theta + \pi)$$

$$X_2 = \theta - \theta_0$$

$$Y_2 = R + (r - 1)$$

식 (5), (6)을 이용하여 직각 좌표계 영상  $I_R$ 을 극좌표계 영상  $I_P$ 로 변환한다.  $(C_x, C_y)$ 는 직각좌표계 영상의 중심점 좌표이다. 직각좌표계 영상에 내접하는 원의 반지름이  $R$  일 때,  $1 \leq r \leq R$  이다.  $\theta_0 \leq \theta < \theta_f$ 이며,  $\theta_0$ 는 변환 시작 각도,  $\theta_f$ 는 변환 종료 각도이다. 하나의  $(r, \theta)$ 값에

대해 직각좌표계 영상의 두 화소값  $I_R(x_1, y_1)$ ,  $I_R(x_2, y_2)$ 이 극 좌표계 영상의 두 화소값  $I_P(X_1, Y_1)$ ,  $I_P(X_2, Y_2)$ 로 각각 변환된다.  $I_R(x, y)$ 는 직각좌표계 영상의 좌표  $(x, y)$ 에서의 화소값이고  $I_P(X, Y)$ 는 극좌표계 영상의 좌표  $(X, Y)$ 에서의 화소값이다.

그림 2는 얼굴 영상의 중심을 기준으로  $0^\circ$ ,  $90^\circ$ ,  $-90^\circ$  회전한 영상들과 이들 각 영상을 극 좌표계 변환한 영상들이다.  $-270^\circ \leq \theta \leq 90^\circ$  각도 범위에서 극좌표계 변환하고 영상 폭을 축소하여  $126 \times 80$  크기로 정규화하였다. 극 좌표계 변환한 영상에서 얼굴에 해당하는 부분을 사각형으로 표시하였다.

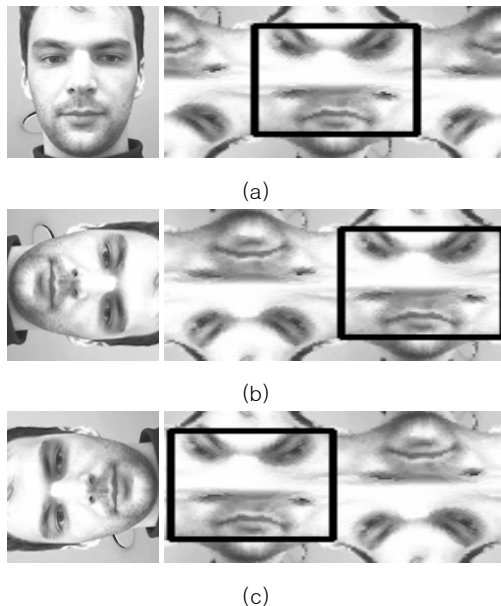


Fig. 2. Rotated faces and polar coordinate transform

그림 2의 극좌표계 변환된 영상에서 직각좌표계 영상의 회전각도에 따라 얼굴의 위치는 변화하지만 회전각도에 무관하게 얼굴 구성요소들 간의 공간 정보가 유지되어 동일한 얼굴 형태가 나타남을 알 수 있다. 극좌표계 변환된 영상에서 얼굴 부분만 비교하면 회전된 것인지 알 수 없기 때문에 회전 효과가 제거된다.

영상이  $81 \times 81$  크기인 경우, 반지름  $R = 40$  이며,  $-270^\circ \leq \theta \leq 90^\circ$  각도 범위에서 극좌표계 변환된 영상의 크기는  $360 \times 80$  이다. 극좌표계 변환된 영상의 폭 360은 극좌표계 변환 각도 범위  $360^\circ$ 에 해당되고 높이 80은  $2R$ 에 해당한다.

이 논문에서 회전 얼굴 검출을 위한 학습 데이터는 회전이 없는 정면 얼굴 영상만을 사용하였다. 그림 1의 (a)와 같은 회전하지 않은 정면 얼굴 영상들을  $81 \times 81$  크기로

정규화하고,  $-180^\circ \leq \theta \leq 0^\circ$  범위에서 극좌표계 영상으로 변환하여  $180 \times 80$  크기를 갖는 극 좌표계 영상을 생성한다. 영상 폭을 축소하여  $63 \times 80$  크기로 정규화된 학습 영상들을 생성하고 HOG 특징을 기반으로 SVM을 사용하여 학습하였다.

## 2. Rotated Face Detection

입력 영상을  $81 \times 81$  크기로 정규화하고  $-180^\circ \leq \theta \leq 360^\circ$  각도 범위로 극좌표계 영상으로 변환하여 회전된 얼굴을 검출하였다. 변환 각도 범위가 이보다 더 작으면 얼굴의 회전각도 크기에 따라 극좌표계 변환된 결과 영상에서 얼굴이 잘리는 경우가 발생한다. 극좌표계 변환된 영상은  $540 \times 80$  크기가 되며, 학습영상의 영상 폭을 줄인 비율에 따라  $189 \times 80$  크기로 정규화하고 회전 얼굴을 검출하였다.

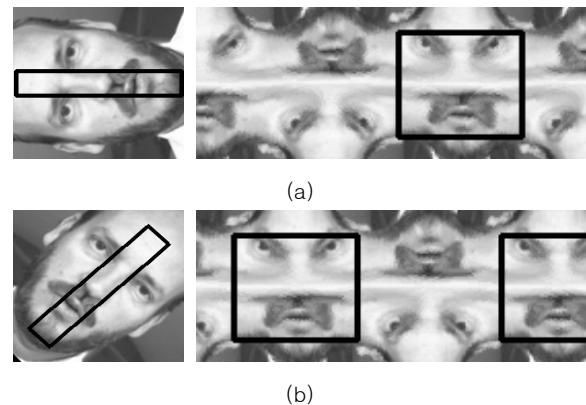


Fig. 3. Polar coordinate transform and detection result

그림 3의 (a)는  $-87^\circ$  회전한 왼쪽 입력 영상을 극 좌표계로 변환한 오른쪽 영상에서 검출한 얼굴 위치를 사각형으로 표시한 것이다. 검출한 얼굴 사각형의 중심 x 좌표가 얼굴의 회전각도이다. 왼쪽 입력 영상에 검출한 회전각도를 직사각형으로 표시하였다.  $-87^\circ$  회전한 것으로 정확하게 검출되었다.

그림 3의 (b)는  $48^\circ$  회전한 왼쪽 입력 영상을 극 좌표계로 변환한 오른쪽 영상에서 검출한 얼굴 사각형 위치를 표시한 것이다.  $46^\circ$ 와  $416^\circ$  두 곳에서 얼굴이 검출되었다.  $416^\circ$ 에서는 얼굴이 일부분만 존재함에도 불구하고 얼굴이 검출되었으며, 이때 검출 사각형은 얼굴 일부분만을 포함하는 크기의 사각형이 아니라  $46^\circ$ 에서의 검출 사각형과 같이 온전한 얼굴을 포함하는 크기의 사각형이 얻어진다. 이와 같이 여러 개의 얼굴이 검출되는 경우에는 얼굴의 대칭성을 이용하여 최종 검출 결과를 판정하였다. 일반적

으로 직각 좌표계 영상에서 얼굴은 좌우 대칭 형태이며, 극 좌표계 변환된 영상에서도 얼굴은 좌우 대칭이 성립한다.

극 좌표계 변환된 영상에서 얼굴이 검출된 각 사각형의 중심  $x$  좌표를 기준으로 검출 사각형들을 좌우 두 영역으로 이등분한다. 각 검출 사각형에서 오른쪽 영역의 영상을 좌우 대칭한 영상( $I_R$ )을 구한 후 왼쪽 영역의 영상( $I_L$ )과의 유사성을 측정한다. 그림 3의 (b)에서와 같이 4160에서 검출한 결과는 극 좌표계 변환된 영상에서 얼굴의 일부분만 존재하기 때문에 검출 사각형의 좌측  $x$  좌표  $x_l$ 를  $x_l = x_l - 360$  과 같이 보정한 사각형 위치에서의 영상을 사용하였다.

영상의 품질 측정에 사용되는 식 (7), (8)의 PSNR(Peak Signal-to-Noise Ratio)을 사용하여 검출한 각 얼굴 사각형에서 좌우 두 영역의 영상  $I_R$ 과  $I_L$ 의 유사성 측정을 하였으며, 가장 큰 PSNR 값을 갖는 검출 사각형을 최종 검출 얼굴로 판정하였다. 식 (7)에서  $W$ ,  $H$  는 검출 사각형을 이등분한 좌우 영역 영상의 폭과 높이이며, 식 (8)에서  $MAX_I$  는 영상의 최대값이다.

$$MSE = \frac{\sum_{y=1}^H \sum_{x=1}^W |I_L(x, y) - I_R(x, y)|^2}{W \times H} \quad (7)$$

$$PSNR = 10 \log_{10} \left( \frac{MAX_I^2}{MSE} \right) \quad (8)$$

#### IV. Experimental Result

BioID 얼굴 데이터베이스[14]에 있는 영상들 중에서 1000개의 얼굴 영상을 무작위로 선택하여 학습 영상으로 사용하였다. BioID 얼굴 데이터베이스는 모두 회전이 없는 23명의 정면 얼굴 영상들이며, 명암도 영상(gray image)들이다. 384 x 286의 해상도를 가진 1521개 영상들로 구성되었다. 정면 얼굴 검출에 사용되는 OpenCV의 AdaBoost 알고리즘을 이용하여 BioID 데이터베이스 영상에서 검출한 얼굴 영역을 81 x 81로 크기 정규화시킨 후, 극좌표계 변환한 영상들을 얼굴 학습 영상으로 사용하였다.

얼굴 검출 과정에서 비얼굴이 얼굴로 오검출된 영상과 인터넷상에서 얻은 영상을 포함하여 2000개의 영상을 비얼굴 학습 영상으로 사용하였다. 이들 얼굴 영상과 비얼굴 영상에서 HOG 특징을 얻고 SVM을 이용하여 학습하였다. 셀 크기를 5 x 5 로 설정하여 얻은 HOG 특징을 사용하였다.

Intel(R) Core i5 CPU 3.30GHz, 8G RAM 환경에서 실험하였다. OpenCV 4.1.2와 Dlib 19.22를 사용하였으며, Dlib 라이브러리에 구현된 HOG 특징을 이용한 얼굴 검출 알고리즘에 제안한 방법을 추가하여 실험하였다.

제안한 회전 얼굴 검출 방법의 성능을 평가하기 위하여 BioID 데이터베이스에서 학습에 사용하지 않은 영상 30개를 선택하여 실험 영상으로 사용하였다. 얼굴 중심을 기준으로 이들 영상을  $0^\circ \sim 357^\circ$  까지  $3^\circ$ 씩 회전시키고 얼굴 영역만 선택하여 총 3600개의 단일 얼굴 실험 영상을 구성하여 회전 얼굴의 회전 각도를 검출하였다. 표 1은 실험결과이다.

Table 1. Experiment result

error angle (degree)	detection number	detection rate
13	3589	99.69 %
10	3571	99.19 %
9	3526	97.94 %
8	3427	95.19 %
6	2993	83.14 %

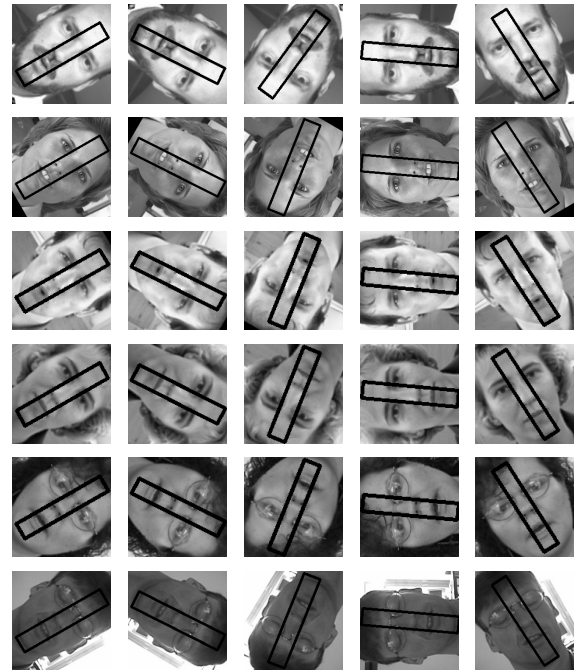


Fig. 4. Detection examples

검출 성공 허용각도를 변화하며 검출률을 비교하였다. 표 1은 입력 영상에서 검출한 얼굴의 회전 각도와 영상의 회전 각도 차이가  $13^\circ$ ,  $10^\circ$ ,  $9^\circ$ ,  $8^\circ$ ,  $6^\circ$  이내인 경우를 각각 검출 성공으로 판정한 결과를 나타낸 것이다.

$0^\circ$  회전한 영상에 있는 얼굴 자체가 조금 기울어진 것도 있기 때문에 검출한 얼굴 회전각도와 영상의 회전각도 차이가 약  $9^\circ$  이내인 경우를 회전각도 검출 성공으로 판정하

였으며, 97.94 %의 검출 성공률을 보인다. 검출 성공 허용 각도가 13°인 경우에는 99.69 %의 회전 각도 검출 성공률을 얻었으며, 이 경우 3600개 실험 영상 중에서 11개만을 검출하지 못하였다.

그림 4는 검출 결과의 예이다. 왼쪽 열부터 60°, 120°, 210°, 270°, 330° 회전한 얼굴 영상에서 얼굴의 회전 각도를 검출한 결과이다. 검출한 얼굴의 회전 각도를 직사각형으로 표시하였다. 일반적인 얼굴뿐만 아니라 안경을 착용한 얼굴과 조명이 밝거나 어두운 얼굴 등 다양한 얼굴 영상에서 회전된 얼굴의 회전 각도를 정확하게 검출하였다.

표 2는 HOG 특징을 얻는 과정에서 셀 크기를 다양하게 (2 x 2, 4 x 4, 5 x 5, 6 x 6, 8 x 8, 10 x 10, 12 x 12) 변화시켜 SVM으로 학습하고 셀 크기에 따른 회전 얼굴의 회전각도 검출 결과를 비교하였다. 셀 크기가 5 x 5 인 경우 가장 좋은 검출률을 얻었으며 이보다 셀 크기가 커지는 경우 점차 검출률이 낮아짐을 알 수 있다.

Table 2. Detection rate by cell size

cell size	detection number	detection rate
2 x 2	3415	94.86 %
4 x 4	3296	91.56 %
5 x 5	3526	97.94 %
6 x 6	3399	94.42 %
8 x 8	2539	70.53 %
10 x 10	2213	61.47 %
12 x 12	1409	39.14 %

배경이 있으며 다수의 회전 얼굴이 포함된 영상에서 제안한 방법을 사용하여 회전 얼굴들을 검출하였다. 그림 5는 MIT-CMU[15]에서 제공하는 회전 얼굴 검출을 위한 실험 영상에서 제안한 방법을 사용하여 회전 얼굴들을 검출한 결과이다.

단일 얼굴 영상에서 회전 얼굴을 검출하는 방법과 달리 이들 영상에서는 여러 얼굴들의 위치와 크기가 정해지지 않았기 때문에 얼굴들의 검출을 위하여 정사각형 형태의 슬라이딩 윈도우(sliding window)를 이용하여 1 화소(pixel)씩 이동하며 영상 전체를 탐색하였다. 각 위치에서 슬라이딩 윈도우 내의 영상을 극 좌표계 영상으로 변환하여 회전 얼굴과 얼굴의 회전 각도를 검출하였다. 영상에 있는 얼굴들의 크기가 다양하기 때문에 슬라이딩 윈도우 크기는 1.1 배씩 증가하였다.



Fig. 5. Multiple rotated faces detection results

인접한 위치에서 중복 검출되는 것을 방지하기 위하여 회전 얼굴이 검출된 슬라이딩 윈도우가 6개 이상 인접한 경우에만 인접한 슬라이딩 윈도우들의 사각형 좌표를 평균하여 하나의 검출 사각형 영역을 얻고 이를 회전 얼굴을 검출한 위치로 선택하였다. 그림 5에 표시한 정사각형들은 회전 얼굴이 검출된 슬라이딩 윈도우를 나타낸 것이다. 이와 같이 정사각형으로 표시된 얼굴 영역은 회전얼굴이 있는 위치만을 나타낸다. 6개 이상 인접한 슬라이딩 윈도우들 각 영역을 극 좌표계 변환한 영상들에서 검출한 얼굴 회전 각도는 동일한 값이 아닐 수 있다.

인접한 각 슬라이딩 윈도우의 극 좌표계 변환 영상에서 얼굴로 검출한 사각형들의 좌표를 특징으로 사용하여 군집화(clustering)를 수행하였다. 가장 많은 원소를 갖는 군집의 검출 사각형 중심  $x$  좌표를 해당 얼굴 위치에서 최종적인 얼굴의 회전 각도로 판정하였다. 그림 5에서 정사각형 내부의 직사각형은 검출한 얼굴 회전 각도를 표시한 것이다. 검출한 회전각도가 모든 얼굴들의 회전 각도와 유사함을 알 수 있다. 또한, 모두 서양인 얼굴들로 구성된 BioID 데이터베이스 영상을 학습 데이터로 사용했음에도 불구하고 그림 5의 네 번째 영상에서는 동양인들의 회전 얼굴들을 모두 정확하게 검출하였다.

그림 5에서 세 번째 영상의 아래쪽에 있는 얼굴 하나는 검출하지 못하였다. 검출된 모든 얼굴들은 정사각형으로 표시한 슬라이딩 윈도우가 얼굴 전체를 충분히 포함하지만 검출되지 않은 얼굴은 영상의 가장자리에 있기 때문에 슬라이딩 윈도우가 얼굴 전체를 충분히 포함하지 못하여 검출하지 못하였다.

회전에 의한 얼굴의 형태변화를 해결하기 위하여 기존 연구들은 정면 얼굴 영상을 지정된 개수의 각도로 회전하여 학습하는 추가적인 과정을 사용하기 때문에 학습시간이 증가할 뿐만 아니라 학습하지 않은 각도의 회전 얼굴은 검출하지 못하는 문제점이 발생할 수 있다. 제안한 방법의 극좌표계 변환 방법은 회전각도에 무관하게 동일한 얼굴 형태를 생성하기 때문에 회전에 의한 얼굴 형태변화가 제거되고 이로 인해 회전 얼굴을 매우 효과적으로 검출할 수 있었다.

## V. Conclusions

이 논문에서는 회전 각도에 무관하게 회전된 얼굴과 얼굴 회전 각도를 검출하는 방법을 제안하였다. 직각 좌표계 영상을 극좌표계 영상으로 변환하고 정면 얼굴 검출에 사용되는 HOG 특징을 이용하여 회전 얼굴을 검출하였다. 제안한 극좌표계 변환 방법은 얼굴 구성요소들의 공간 정보가 유지되기 때문에 임의의 각도로 회전이 발생하더라도 동일한 형태를 생성하기 때문에 회전 얼굴 검출에 매우 효과적이다. 또한, 극좌표계 변환을 사용함으로써 학습시 회전된 얼굴 영상이 필요하지 않고 정면 얼굴 영상만을 사용하여 학습을 수행하게 된다. 단일 얼굴 영상과 여러 회전 얼굴을 포함하는 배경이 있는 영상들을 사용한 실험결과 회전 얼굴을 검출하는 과정에서 제안한 방법이 매우 효과적임을 확인할 수 있었다.

이 논문에서 제안한 회전 얼굴 검출 방법은 회전에 의한 외형 변화 효과를 제거하기 때문에 정면 얼굴 검출에는 좋은 성능을 보이지만 회전에 민감한 특성을 갖는 기존의 정면 얼굴 검출 방법에 적용하여 회전얼굴을 검출하는 과정에서 효과적으로 사용될 수 있다. 또한, 회전 얼굴뿐만 아니라 회전된 일반 물체의 검출에도 효과적으로 사용될 수 있을 것으로 기대된다.

## REFERENCES

- [1] G. Singh and A. K. Goel, "Face Detection and Recognition System using Digital Image Processing," International Conference on Innovative Mechanisms for Industry Applications, pp. 348-352, Mar. 2020.
- [2] P. Viola and M. Jones, "Robust Real-Time Face Detection," IEEE International Conference Computer Vision, vol. 20, pp. 1254-1259, Jul. 2001.
- [3] B. Froba and A. Ernst, "Face detection with the modified census transform," IEEE Proc. of Automatic Face and Gesture Recognition, pp. 91-96, 2004.
- [4] Rekha N. and M. Z. Kurian, "Face Detection in Real Time Based on HOG," International Journal of Advanced Research in Computer Engineering & Technology, Vol. 3, No. 4, pp. 1345-1352, Apr. 2014.
- [5] Y. Chen and Y. X. W. Liu, "Multi-angle Face Detection with Step-by-Step Adjustment Networks," International Conference on Mechanical, Control and Computer Engineering, pp. 589-594, Sept. 2018.
- [6] H. Liu, S. Yan, X. Chen and W. Gao "Rotated face detection in color images using radial template (RT)," IEEE International Conference Multimedia and Expo, vol. 3, pp. 137-140, July 2003.
- [7] H. Rowley, S. Baluja, and T. Kanade, "Rotation Invariant Neural Network-Based Face Detection," IEEE International Conference on Computer Vision and Pattern Recognition, pp. 38-44, 1998.
- [8] Y. H. Tsai, Y. C. Lee, J. J. Ding, R. Y. Chang, and M. C. Hsu, "Robust in-plane and out-of-plane face detection algorithm using frontal face detector and symmetry extension," Image and Vision Computing, vol. 78, pp. 26-41, Oct. 2018.
- [9] S. N. Sujay, H. S. M. Reddy, and J. Ravi, "Face recognition using extended LBP features and multilevel SVM classifier," International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques, pp. 1-4, Dec. 2017.
- [10] X. Shi, S. Shan, M. Kan, S. Wu, and X. Chen, "Real-time rotation-invariant face detection with progressive calibration networks," IEEE Conference on Computer Vision and Pattern Recognition, pp. 2295-2303, Jun. 2018.

- [11] Y. Shima, Y. Nakashima, and M. Yasuda, "Detecting orientation of in-plain rotated face images based on category classification by deep learning," IEEE Region 10 Conference (TENCON), pp. 127-132, Nov. 2017.
- [12] Dalal and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 886-893, 2005.
- [13] E. Osuna, R. Freund, F. Girosi, "Training Support Vector Machine : An Application to Face Detection," Proceeding IEEE CVPR, pp 130-136, 1997.
- [14] O. Jesprsky, K. Kirchberg, R. Frischholz, "Robust Face Detection Using the Hausdorff Distance," Audio based Person Authentication, pp90-95, 2001.
- [15] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, pp. 23-38, Jan. 1998.

## Authors



Kyung-Shik Jang received the B.S., M.S. and Ph.D. degrees in Electronic Engineering from Yonsei University, Korea, in 1989, 1991 and 1996, respectively. Dr. Jang joined the faculty of the Division of Creative Software

Engineering at Dongeui University, Busan, Korea, in 1998. He is currently a Professor in the Division of Creative Software Engineering, Dongeui University. He is interested in image processing, pattern recognition and computer vision.