

Boosted Random Ferns 를 이용한 회전 불변 얼굴 검출

김후현, *조동찬, *배종엽, **김희율
한양대학교

hhkim@vision.hanyang.ac.kr, *dccho@vision.hanyang.ac.kr,
*jybae@vision.hanyang.ac.kr, **wykim@hanyang.ac.kr

Rotation Invariant Face Detection with Boosted Random Ferns

Hoo Hyun Kim *Dong-Chan Cho *Jong Yeop Bae **Whoi-Yul Kim
Department of Electronics and Computer Engineering, Hanyang University

요 약

본 논문은 Boosted Random Ferns 기반의 회전 불변 얼굴 검출 방법을 제안한다. 기존 Random Ferns 의 경우 특징값을 추출할 때 임의로 선택한 두 픽셀의 밝기값 비교를 통하여 이진 특징값을 추출한다. 이 경우 해당 픽셀의 밝기값에 잡음이 포함되면 특징값이 부정확하게 추출되는 문제가 있다. 본 논문에서는 이러한 문제를 해결하기 위하여 임의로 두 블록을 선택하고 해당 블록내 밝기값의 평균을 비교하여 이진 특징값을 추출하였다. 또한 픽셀 위치를 임의로 선택하여 ferns 를 구성하였던 기존의 방법 대신 최고의 분류 성능을 가지는 fern 들을 이용하여 분류기를 구성하기 위해, AdaBoost 의 방법을 Random Ferns 에 맞게 변경하였다. Boosted Random Ferns 를 트리 구조의 cascade 노드에 방향과 각도에 따라 배치하여 연산 속도를 향상시키고 false-positive 를 줄이는 효과를 보였다. CMU Rotated Face Database 를 사용하여 평가하였을 때, 기존 Random Ferns 는 false-positive 의 수가 57 개 일 때 66%의 검출률을 보인 반면, Boosted Random Ferns 는 false-positive 의 수가 45 개 일 때 88%의 검출률을 보였다.

1. 서론

최근 영상처리 분야에서 얼굴검출은 개인식별, 얼굴 사진 정합, 군중감시와 얼굴 재구성과 같은 다양한 응용 분야에서 활용되고 있다. 얼굴을 검출하는 방법은 얼굴의 일반적인 요소를 검출하는 기법과 영역 분류에 의한 기법으로 나뉜다. 얼굴요소를 검출하는 방법은 눈, 코, 입, 입술을 검출한 후 기하학적인 특징에 의해 그룹화하여 독립된 얼굴을 검출 하는 기법으로, 각각의 얼굴요소들은 공간 필터, 이웃 영역과의 밝기 차, 피부색 모델 등을 이용하여 추출할 수 있다. 영역 분류에 의한 기법은 SVM, 신경망, AdaBoost 와 같은 분류기를 이용하여 국부 영역을 탐색하고 얼굴과 얼굴이 아닌 영역을 분류한다. 특히 Viola 가 제안한 얼굴 검출 방법은 haar-like 특징값과 약 분류기를 수집하여 강 분류기를 만드는 방법인 AdaBoost 를 사용하여 회전하지 않은 정면 얼굴에 대하여 90%이상의 검출률을 보이고 있다[1].

하지만 회전된 얼굴의 회전을 고려하지 않았기 때문에 회전된 얼굴을 포함하는 영상에서 얼굴 검출이 불가능하다. 이러한 한계를 극복하기 위해 회전 특성을 포함하는 특징벡터를 구성하는 방법으로 Kim [2]은 zernike 모멘트를 이용하였고, Christopher [3]는 특징 벡터를 각 방향에 따라 추출할 수 있도록 gradient, gaussian, gabor 필터를 구성하였다. 또한 기존 Viola 의 특징 추출 함수를 변형한 방법으로 Shaoyi [4]는 회전된 haar-like 패턴을 사용하여 각 회전에 대한 cascade 를 구성하고 회전된 얼굴을 검출하였고 Jones [5]방법은 Diagonal Filter 를 사용한다. 하지만 [3, 4, 5]의 방법들은 모두 검출률에 비해 false-positive 를 많이 포함한다.

본 논문에서는 얼굴의 방향에 따라 클래스를 구분하고 Random Ferns 를 이용하여 회전 불변의 얼굴 검출을 수행하는 방법을 제안한다. 또한 기존 Random Ferns 의 특징 검출 함수

는 두 픽셀의 비교를 통한 특징값을 검출하기 때문에 픽셀 잡음에 취약한 단점을 가지고 있지만, 본 논문은 블록 단위의 비교를 통해 잡음에 강인하게 구성한다. 그리고 기존 Random Ferns 는 임의로 fern 들을 추출하는 반면 제안하는 방법은 Boosted Random Ferns 를 통해 최적의 fern 들로 분류기를 구성한다.

2. Random Ferns

Random Ferns 는 다중 클래스 분류가 가능하고 영상내 밝기값 비교만을 이용하여 특징값을 추출하기 때문에 연산 속도가 빠르며 높은 검출률을 보이는 분류기이다[6].

H 개의 다중 클래스를 $c_i, i=1, \dots, H$ 로 정의하고, N 개의 특징 추출 함수를 $f_j, j=1, \dots, N$ 로 정의했을 때, 특징값이 속하는 최적의 클래스 \hat{c}_i 를 아래의 수식과 같이 구할 수 있다.

$$\hat{c}_i = \arg \max_{c_i} P(C = c_i | f_1, f_2, \dots, f_N). \quad (1)$$

Bayes 정의를 이용하여 식(1)을 다음과 같이 정의할 수 있다.

$$P(C = c_i | f_1, f_2, \dots, f_N) = \frac{P(f_1, f_2, \dots, f_N | C = c_i)P(C = c_i)}{P(f_1, f_2, \dots, f_N)} \quad (2)$$

여기서 $P(C)$ 와 $P(f_1, f_2, \dots, f_N)$ 를 정해진 확률값으로 가정하면, 다음과 같은 식(2)를 재 정의 할 수 있다.

$$\hat{c}_i = \arg \max_{c_i} P(f_1, f_2, \dots, f_N | C = c_i). \quad (3)$$

식(3)에서 $P(f_1, f_2, \dots, f_N | C = c_i)$ 를 구하기 위해 각 특징 추출 함수들을 독립적으로 생각한다면 식(4)와 같이 계산 할

수 있다.

$$P(f_1, f_2, \dots, f_N | C = c_i) = \prod_{j=1}^N P(f_j | C = c_i). \quad (4)$$

하지만 일반적으로 몇몇 특징 추출 함수들끼리 의존적인 특징을 보이기 때문에 이에 대한 고려가 필요하다. Random Ferns 는 특징 추출 함수들 사이에 상관 관계가 있다고 가정하고, 상관관계가 있는 특징 추출 함수들을 묶어 하나의 fern 이라고 지칭한다. Random Ferns 를 이용하면 수식(4)를 다음과 같이 수정할 수 있다.

$$P(f_1, f_2, \dots, f_N | C = c_i) = \prod_{k=1}^M P(F_k | C = c_i). \quad (5)$$

수식(5)에서 $F_k = \{f_{\sigma(k,1)}, f_{\sigma(k,2)}, \dots, f_{\sigma(k,S)}\}, k = 1, \dots, M$ 는 k 번째 ferns 을 나타내는데 내부에 S 개의 특징 추출 함수를 포함하고 있다. $\sigma(k, j)$ 는 범위가 $1, \dots, N$ 인 랜덤 순열 함수로 S 개의 특징 추출 함수가 임의로 선택된다. 총 M 개의 fern 의 결과를 이용하여 분류를 수행한다[6].

3. 제안한 방법

본 논문은 Random Ferns 을 이용하여 얼굴을 회전 각도에 따라 분류하여 회전 불변 얼굴 검출이 가능한 방법을 제안한다.

3.1. 특징 추출

기존의 Random Ferns 은 서로 다른 두 위치의 픽셀 밝기 차이에 따라 특징 추출 함수를 구성한다[5].

$$f_i = \begin{cases} 1 & \text{if } I(\mathbf{d}_{j,1}) < I(\mathbf{d}_{j,2}) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$I(\mathbf{d}_{j,1})$ 와 $I(\mathbf{d}_{j,2})$ 는 j 번째 특징 추출 함수에서 1 번 위치와 2 번 위치에서의 픽셀 밝기값을 나타낸다. 하지만 픽셀 단위의 특징 추출 함수를 사용할 경우, 해당 픽셀에 잡음이 발생하면 특징 값이 다르게 나타나는 문제가 있다. 따라서 본 논문에서는 잡음의 영향을 최소화 하기 위해 픽셀 단위가 아닌 블록 단위의 밝기 합의 차이를 비교하는 특징 추출 함수를 수식(7)와 같이 계산할 수 있다.

$$f_i = \begin{cases} 1 & \text{if } B(\mathbf{d}_{j,(1,w,h)}) < B(\mathbf{d}_{j,(2,w,h)}) \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$B(\mathbf{d}_{j,(1,w,h)})$ 와 $B(\mathbf{d}_{j,(2,w,h)})$ 는 j 번째 특징 추출 함수에서 길이가 w 이고 높이가 h 인 1 번과 2 번 위치의 블록 밝기 합을 나타낸다(그림 1).

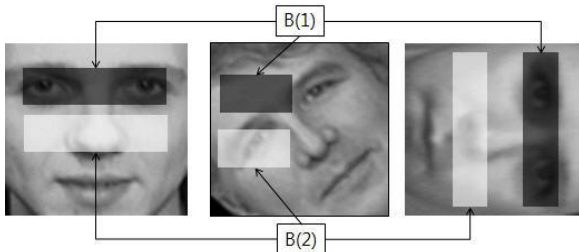


그림 1. 두 블록의(B(1), B(2)) 밝기 합의 차이를 이용한 특징 추출 함수 구성

3. 2. Boosted Random Ferns

기존 Random Ferns 의 경우 fern 을 구성할 때 임의로 선택한 두 픽셀 위치의 밝기값 차이를 이용하기 때문에 여러 개의 fern 들로 최종 Random ferns 를 구성하더라도 몇 개의 fern 들은 성능이 떨어질 수가 있다. 이러한 문제를 해결하기 위해 본 논문에서는 여러 개의 fern 들을 생성한 후 최상의 분류 성능을 가지는 fern 만을 선택하기 위해 AdaBoost[7]의 방법을 random ferns 에 맞게 변형(Boosted Random Ferns)하였으며, 클래스들마다 훈련 데이터의 가중치를 다르게 주어 데이터의 분류에 변별력을 주었다.

AdaBoost[7]는 부스팅과정에서 선택된 약 분류기 $h_t(x)$ 의 선형 합을 이용하여 강 분류기 $H(x)$ 를 수식(8)과 같이 만든다.

$$H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right) \quad (8)$$

수식(8)에서 α_t 는 선택된 약 분류기의 성능에 따른 가중치를 나타낸다. 본 논문에서는 수식(8)과 같이 강 분류기 $H(x)$ 를 구성하기 위해 약 분류기 $h_t(x)$ 로 다중 분류가 가능한 fern 을 이용하며, 부스팅과정에서 T 개의 $h_t(x)$ 를 생성한다.

부스팅과정에 앞서 약 분류기 모음인 fern 들을 미리 훈련시키기 위해, 수식(9)과 같은 방식으로 각 fern 에 대한 클래스 확률 분포를 생성한다.

$$P_{k,c_i} = \frac{N_{k,c_i} + N_r}{N_{c_i} + K \times N_r} \quad (9)$$

이 수식에서 N_{k,c_i} 은 fern 값 k 로 평가된 c_i 클래스의 훈련 데이터 수이고, N_{c_i} 는 c_i 클래스의 훈련 데이터 총 수이다. K 는 fern 이 이진화로 표현 할 수 있는 최대의 값으로 특징 함수의 개수에 따라 정해지며, N_r 은 확률값이 0 이 되는 것을 막기 위한 값으로 상수로 설정한다[6].

Boosted Random Ferns 는 먼저 n 개의 훈련 데이터 x_i 와 훈련 데이터의 클래스 레이블 ω_i 를 이용하여 훈련 집합 $(x_1, \omega_1) \dots (x_n, \omega_n)$ 을 생성한다. 그리고 훈련 데이터의 초기 가중치는 클래스에 따라 달리 구성이 되는데 수식(10)과 같이 훈련 데이터를 각 클래스에 따라 분류하고, 클래스에 포함된 데이터의 개수로 나누어 계산한다.

$$D_t^{\omega_i} = \frac{1}{N_{\omega_i}} \quad (10)$$

이 수식에서 $D_t^{\omega_i}$ 는 해당 데이터의 클래스 가중치를 나타내며, N_{ω_i} 는 각 클래스에 포함된 데이터의 수를 나타낸다.

Boosted Random Ferns 의 부스팅과정은 각 클래스의 데이터의 오류가 최저가 되는 fern 을 찾는 과정이다. 각 fern 에 훈련 데이터의 특징값을 추출하고, 특징값에 따른 확률이 가장 높은 클래스를 선택한 후, 훈련 데이터의 클래스와 같은지를 수식(11)과 같이 비교한다. 만약 훈련 데이터의 클래스가 다르면 해당 데이터의 가중치를 더해 fern 의 오류값을 측정한다. 최적의 fern 을 선택할 시에는 오류값이 가장 낮은 fern 을 선택하며 수식(12)과 같다.

$$y(h(x_i)) = \begin{cases} 1 & \text{if } L(h(x_i)) = \omega \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

$$h_t(x) = \arg \min_{h(x)} \left\{ \sum_{i=1}^n D_i^{\omega_i} \exp(-1 \times y(h(x_i))) \right\} \quad (12)$$

수식(11)에서 $L(h(x_i))$ 는 데이터를 약 분류기에 넣었을 때 가장 높은 확률로 출력되는 클래스이고, ω 는 해당 데이터의 클래스이다. 선택된 fern 은 수식(12)에서 계산된 오류값을 통해 fern 에 대한 가중치를 수식(13)와 같이 계산 할 수 있다. 이 가중치는 fern 의 성능을 나타내며, fern 의 오류값이 작을수록 높은 가중치를 부여한다.

$$\alpha_t^{\omega} = 1 / 2 \log((1 - e(h_t(x))) / e(h_t(x))) \quad (13)$$

α_t^{ω} 는 $h_t(x)$ 의 클래스에 따른 가중치이고, $e(h_t(x))$ 는 $y(h_t(x))$ 가 0 일 때, 해당 데이터의 $D_t^{\omega_i}$ 들의 합이다.

부스팅과정이 끝났다면, 다음 부스팅 과정을 위해 분류가 안된 훈련 데이터들의 가중치를 높여 주고, 분류가 된 훈련 데이터의 가중치를 낮게 설정 함으로써, 다음 부스팅과정에서 분류가 안된 데이터를 분류가 가능하게 설정 하는 과정(수식(14))이 필요하다.

$$D_{t+1}^{\omega_i} = D_t^{\omega_i} \times \exp(-\alpha_t \times y(h_t(x))) / Z_t^{\omega_i} \quad (14)$$

$$Z_t^{\omega_i} = \sum_{i=0}^n D_t^{\omega_i} \times \exp(-\alpha_t \times y(h_t(x_i))) \quad (15)$$

$Z_t^{\omega_i}$ 는 각 클래스의 데이터 가중치 합으로 이루어진 정규화 값이다. 그림 2 는 Boosted Random Ferns 과정을 도식화 하여 설명한다.

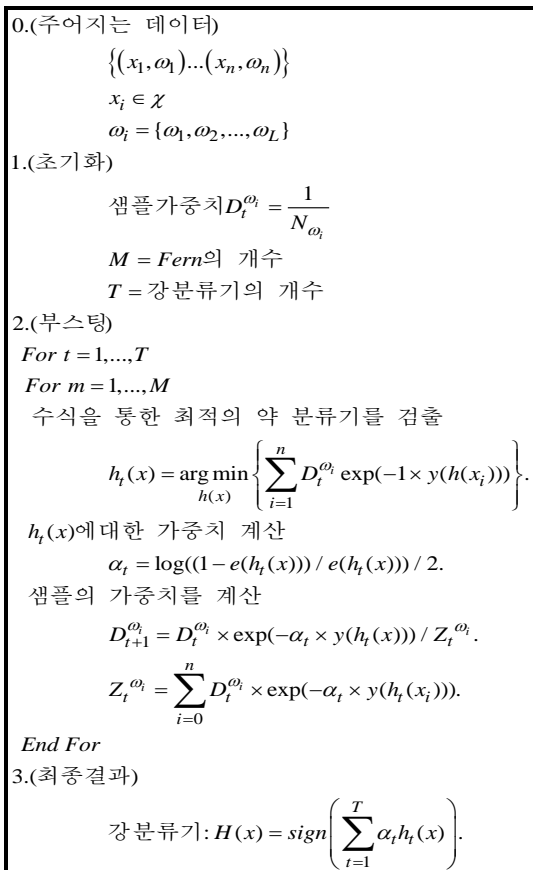


그림 2. Boosted Random Ferns

3. 3. 트리 구조의 cascade

선형적인 구조의 Boosted Random Ferns 를 이용하여 얼굴의 각도를 구성 한다면, 클래스의 개수가 각도(30, 60, 90, -30, -60, -90, B)에 따라 7개가 된다. 이는 클래스의 수가 많아 지기 때문에 false positive 가 많이 발생하게 되는 원인이 된다. 따라서 제안하는 방법은 그림 3 과 같이 트리 구조의 cascade 를 구성한다. Level1 는 얼굴의 회전된 방향을 구분하는 분류기 $H_1(x)$ 을 구성하고, Level2 는 얼굴의 각도를 구분하는 분류기 $H_2(x)H_4(x)$ 를 구성한다. 이 방법은 노드의 클래스가 4 개로 줄어 들게 되어 false positive 를 줄일 수 있게 된다. 또한 분류기가 증가하게 되어 연산 속도가 느려 지는 것 막기 위해 각 분류기에 배경 클래스를 포함시켜 걸러 줌으로써 연산 속도를 증가 시킨다.

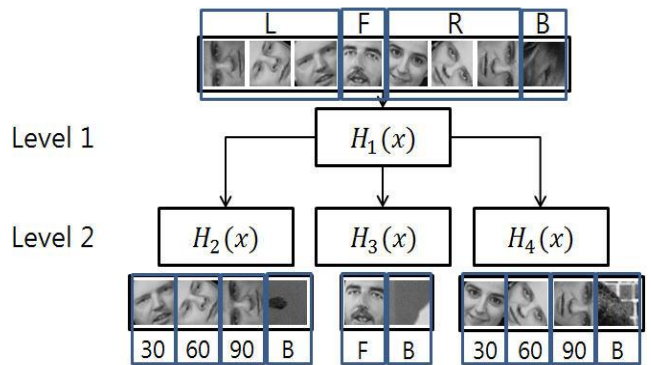


그림 3. 트리 구조의 Cascade

3. 3. 얼굴검출

얼굴검출은 각 노드에 Boosted Random Ferns 를 사용하여 생성된 분류기를 이용한다. 영상이 입력되면 최상의 노드에서 얼굴의 좌, 우, 정면, 비 얼굴을 검출하며, 노드에 결과에 따라 다음 Level 의 노드로 이동하거나, 얼굴검출을 멈추게 된다.

각 노드를 거쳐 리프노드에 도달하면, 최종 적으로 얼굴의 각도가 출력된다. 하지만 출력된 결과가 배경과 얼굴이 비슷한 확률을 가질 경우 어떤 클래스를 선택 할 것인지가 문제이다. 본 논문은 배경과 얼굴의 비율을 통해 클래스를 수식(12)과 같이 선택 한다.

$$\frac{H(C|x)}{H(B|x)} < \beta \quad (12)$$

이 수식에서 $H(C|x)$ 는 x 영상이 입력 될 때 C 클래스일 확률 이고, $H(N|x)$ 는 x 영상이 입력 될 때 B 배경 클래스일 확률이다. 두 확률의 비율이 β 이하일 때 얼굴로 판단하고 β 이상이면 배경으로 판단한다. β 는 0.90~0.93 사이에 가장 좋은 결과를 도출한다.

4. 실험 결과

본 논문은 CMU Rotated Face Database 을 이용하여 블록 기반 특징 함수의 성능과 Boosted Random Ferns 의 성능을 비교하고, 마지막으로 타 방법과의 얼굴 검출률을 비교한다.

● 특징 함수 비교

기존 Random Ferns 의 특징 추출 함수(픽셀)와 제안하는 방법(블록)의 성능을 평가 하기 위해, 가우시안 형태의 잡음을 테스트 영상에 추가하여 검출률을 비교한다.

가우시안 잡음 (sigma)	픽셀기반 검출률 (%)	블록기반 검출률 (%)
10	90	100
20	86	100
30	60	95

표 1. 가우시안 잡음 변화에 따른 특징 추출 함수 검출률

표 1 의 결과는 블록 기반 방법의 검출률이 픽셀 기반 검출률 보다 높다는 것을 알 수 있다. 이는 블록 기반 검출은 픽셀 변화에 민감하지 않기 때문에 검출률이 떨어지지 않는다(그림 4).

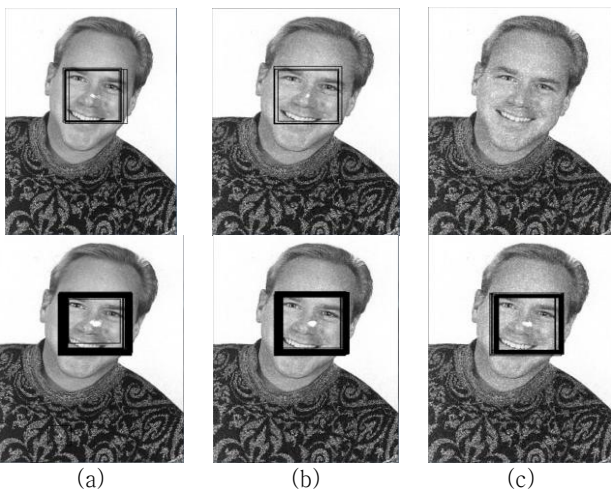


그림 4. 잡음에 따른 얼굴 검출(위: 픽셀 기반, 아래: 블록 기반) (a) sigma = 10 (b) sigma = 20 (c) sigma = 30

● Boosted Random Ferns 성능

표 2 은 Boosted Random Ferns 를 사용하여 최적의 ferns 를 구성 했을 때의 검출률을 기존 방법인 random 한 ferns 로 선택 했을 때와의 검출률을 비교 한다. 두 방법 모두 fern 의 개수는 10 개로 하고 false-positive 가 60 개 내인 상황에서 측정 한다.

ferns 구성 방법	검출률 (%)	false-positive
Boosted Random Ferns	88%	45
Original Random Ferns	66%	57

표 2. ferns 구성 방법에 따른 검출률 및 false-positive 의 수

Boosted Random Ferns 를 이용하면 최적의 fern 들을 추출 하여 분류기를 구성 할 수 있기 때문에 같은 개수의 fern 들을 구성 하였을 때, 기존 Random Ferns 보다 검출 성능이 우수하다.

● 얼굴 회전 검출률 비교

표 3 은 제안한 방법의 얼굴 회전 검출 성능을 평가하기 위해, [3, 4, 5]방법과 false-positive 및 검출률을 비교한다. [3, 4, 5]방법은 검출률은 높지만 false-positive 가 많이 발생 하

는데 비해, 제안하는 방법은 false-positive 가 타 방법보다 적은 것을 볼 수 있다.

얼굴 검출 방법	검출률 (%)	false-positive
[3]의 방법	93%	137
[4]의 방법	90.1%	221
[5]의 방법	89.7%	221
제안하는 방법	93%	109

표 3. 얼굴 검출 방법에 따른 검출률 및 false-positive

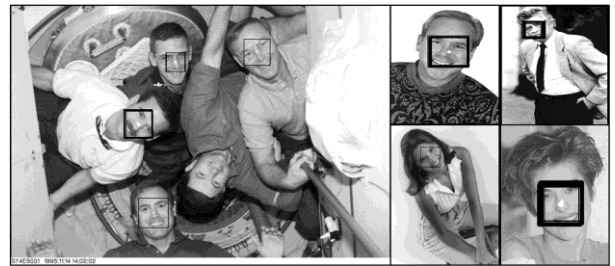


그림 5. 얼굴 검출 결과 (CMU Rotated Face Database)

5. 결론

본 논문에서는 Boosted Random Ferns 를 이용한 회전 불변의 얼굴 검출 알고리즘을 제안하였다. 기존 Random Ferns 의 픽셀 기반 특징 추출 함수를 블록 기반 특징 추출 함수로 변경하여 픽셀의 잡음에 강인한 특징값을 추출하였다. 또한 AdaBoost 방법을 변형한 Boosted Random Ferns 를 제안하여 검출 성능을 향상 시켰고, 트리구조의 cascade 를 구성하여 false-positive 를 줄일 수 있었다.

제안하는 방법을 평가하기 위해, CMU Rotated Face Database 를 사용하였다. 기존 Random Ferns 의 픽셀 기반 특징 추출 함수를 사용하였을 때, 잡음이 증가 할수록 검출률이 감소 하였다. 하지만 제안하는 블록 기반 특징 추출 함수는 잡음이 증가 하여도 검출률이 감소하지 않았다. 또한 기존 방법인 임의의 fern 을 선택 하였을 때 검출률이 66%인 반면, Boosted Random Ferns 를 사용하여 fern 를 선택 하였을 때 검출률이 88%로 기존 방법보다 검출 성능이 우수하다.

참고문헌

- [1] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. CVPR, 2001.
- [2] H.J Kim. zernike 모멘트를 이용한 얼굴 검출. 한국멀티미디어학회, 2006.
- [3] A. Christopher and L.Xiuxen. Rotation Invariant Face Detection using Spectral Histograms and Support vector Machines. ICIP, 2006.
- [4] D. Shaoyi and Z. Nanning. Rotated Haar-Like Features for Face Detection with In-Plane Rotation. VSMM, 2006.
- [5] M. Jones and P. Viola. Fast multi-view face detection. Technical report, 2003.
- [6] O. Mustafa and F. Pascal. Fast Keypoint Recognition in Ten Lines of Code. CVPR, 2007,
- [7] F. Yoav and S. Robert. A decision theoretic generalization of on line learning and an application to boosting. ICLT, 1995,