

논문 2014-09-32

픽셀 방향코드와 룩업테이블 분류기를 이용한 얼굴 검출

(Face Detection Using Pixel Direction Code and Look-Up
Table Classifier)

임길택, 강현우*, 한병길, 이종택

(Kil-Taek Lim, Hyunwoo Kang, Byung-Gil Han, Jong Taek Lee)

Abstract : Face detection is essential to the full automation of face image processing application system such as face recognition, facial expression recognition, age estimation and gender identification. It is found that local image features which includes Haar-like, LBP, and MCT and the Adaboost algorithm for classifier combination are very effective for real time face detection. In this paper, we present a face detection method using local pixel direction code(PDC) feature and lookup table classifiers. The proposed PDC feature is much more effective to detect the faces than the existing local binary structural features such as MCT and LBP. We found that our method's classification rate as well as detection rate under equal false positive rate are higher than conventional one.

Keywords : Face detection, AdaBoost, Local Direction Code, LBP, MCT

1. 서론

얼굴 검출 기술은 얼굴 인식, 표정 인식, 나이 및 성별 인식 등 다양한 얼굴 응용시스템의 완전한 자동화를 위해 반드시 필요하며, 보안시스템, 디지털카메라, 로봇, 자동차 등 다양한 카메라 응용 환경에서 활용되고 있다. 최근의 연구에서는 실시간 적용이 가능한 수준의 고성능 얼굴 검출 기술이 개발되는 등 기술의 활용도가 높아지고 있다.

얼굴 검출에는 신경망을 이용하는 방법, SVM(support vector machine)을 이용하는 방법, AdaBoost를 이용하는 방법 등이 많이 사용되고 있다. 초창기 얼굴 검출 연구의 대표적 연구로는 신경망과 통계적 얼굴모델을 이용하는 Rowley 등[1]과 Sung 등[2]의 연구가 있다. 이들의 방법은 검출 속도 측면에서 실시간 응용이 곤란하였으나 비교적 높은 검출률을 나타내었다. Heisele 등[3] 과 Le

등[4]은 SVM를 기반으로 하는 다단계 분류 방법을 적용하여 속도와 검출률을 개선한 얼굴 검출 시스템을 구현하였다. Viola 등[5]의 연구는 높은 검출률이면서 실시간 처리가 가능한 최초의 연구라고 할 수 있다. 실시간 얼굴 검출을 위해서 Viola와 Jones는 Haar-like 특징을 이용하는 약분류기들을 AdaBoost [6]로 결합한 캐스캐이드형 검출 방법을 제시하였다.

검출 대상 얼굴의 크기를 사전에 파악할 수 없는 경우의 얼굴 검출을 위해서 전체 입력 영상에 대한 영상 피라미드를 사용하는 것이 일반적이다. 영상 피라미드 내의 각 영상에 대해 전 영역에 걸쳐 탐색창을 띄워 스캐닝을 하면서 이진 얼굴 분류기를 적용한다. 예를 들어 VGA 영상(640×480)에서 20×20 크기의 얼굴을 찾겠다고 할 때 (640-20)×(480-20)개의 탐색창을 설정할 수 있어 분류횟수는 약 285,200이며, 영상 피라미드를 고려하면 수백만 회 이상의 분류를 하여야할 만큼 엄청난 탐색이 필요하다.

이와 같은 엄청난 계산량 때문에, 분류 능력은 뛰어나지만 속도는 느린 강력한 분류기는 실용적인 관점에서 전혀 쓸모가 없다. 따라서 AdaBoost를 이

*Corresponding Author(hwkang@etri.re.kr)

Received: 30 July 2014, Revised: 13 Aug. 2014,

Accepted: 25 Aug 2014.

K.-T Lim, H. Kang, B.-G. Han, J.T. Lee: ETRI

© IEMEK J. Embed. Sys. Appl. 2014 Oct.: 9(5) 261-268

ISSN : 1975-5066

http://dx.doi.org/10.14372/IEMEK.2014.9.5.261

용하여 고속 약분류기를 학습하고 선형 조합하는 Viola와 Jones의 방법이 가장 현실적이며 강력하다. Viola의 검출 프레임워크를 사용하는 Haar-like 특징 기반의 방법을 개선한 연구들이 있다[7-10]. Fröba 등[7]과 Zhang 등[8]은 Haar-like 특징보다 분류능력이 뛰어난 MCT (Modified Census Transform)[7]와 LBP (Local Binary Pattern)[8]라는 지역 이진 구조 패턴 특징을 사용함으로써 Viola의 방법에 비해 훨씬 작은 수의 약분류기만으로도 고속으로 얼굴과 비얼굴을 구분하였다. 이와 같은 지역 이진 구조 패턴 특징은 얼굴 검출뿐만 아니라 얼굴 인식에도 사용되는 등 얼굴 영상 분석 분야에서 가장 많이 사용되는 특징으로 알려져있다 [9].

본 논문에서는 고속 고정확도의 얼굴 검출기를 구현하기 위해, 얼굴 검출에 유용한 방향 코드 특징과 그 특징을 효율적으로 활용하는 룩업 테이블 분류기의 구현방법을 제안한다. 제안 하는 방법은 분류율, 기각률 대비 검출률, 메모리 요구사항 등의 관점에서 MCT와 LBP 특징을 이용하는 방법 보다 더 우수한 특성을 보여준다.

II. 특징과 얼굴 분류기

본 연구에서는 3×3 지역 윈도우내의 9개의 픽셀들이 나타내는 픽셀들의 구조적 특징을 얼굴 검출에 이용하는데, LBP, MCT, 그리고 본 논문에서 제안하는 픽셀 방향 코드 특징이다. 이들 특징을 입력으로 하는 룩업테이블 분류기들을 AdaBoost 방법을 이용하여 결합함으로써 최종 얼굴 검출기를 생성한다.

1. LBP(Local Binary Pattern) and MCT (Modified Census Transform)

LBP 특징은 3×3 커널을 통해 계산되는 픽셀들의 지역적인 구조 특성이다[8]. LBP 특징은 Census Transform[11]과 개념적으로 동일하다. 특정 픽셀에서의 LBP는 픽셀 p 와 그 주변의 8개의 픽셀들 ($p_n, n=0, \dots, 7$)의 밝기값을 비교하여 구해진다.

$$BP(p) = \sum_{n=0}^7 s(i(p_n) - i(p))2^n, \quad (1)$$

여기서 $i(p)$ 는 픽셀 p 의 밝기값이며, $s(a) = \begin{cases} 1, & \text{if } a \geq 0 \\ 0, & \text{if } a < 0 \end{cases}$ 이다.

110	130	150
120	135	130
125	125	130

(a)

0	0	1
0	-	0
0	0	0

(0010000)₂

(b)

0	1	1
0	1	1
0	0	1

(011110001)₂

(c)

그림 1. LBP와 MCT의 계산 (a) 3×3 지역 영상, (b) LBP, (c) MCT

Fig. 1 Calculation of LBP and MCT (a) 3×3 local image, (b) LBP, (c) MCT

MCT는 Fröba 등[7]이 Census Transform[11]을 개선한 것이다. 3×3 영역내에 존재하는 픽셀들의 평균 밝기와 9개의 각 픽셀들의 밝기값을 비교하여 이진 패턴으로 표현할 수 있다.

$$T(p) = \bigotimes_{q \in N(p)} C(i(p), i(q)), \quad (2)$$

여기서 $N(p) = N(p) \cup p$ 으로 $N(p)$ 에 픽셀 p 을 추가한 집합으로, $N(p)$ 는 픽셀 p 의 주변 픽셀들의 집합이다. $\bar{i}(p)$ 는 $N(p)$ 의 평균 밝기값, $C(a, b) = \begin{cases} 1, & \text{if } a \leq b \\ 0, & \text{if } a > b \end{cases}$ 그리고 \otimes 는 비트 연결 연산자이다. 즉, MCT는 LBP와 달리 영역내 평균값과 전체 9개의 픽셀 모두에 대해 크기 비교하여 그 논리값을 연결한 것이다.

LBP와 MCT 특징을 보다 쉽게 이해하기 위해 그림으로 표현하면 그림 1과 같다. 그림 1의 (a)는 3×3 윈도우내의 9개의 픽셀이며, (b)는 LBP, (c)는 MCT 값이다. LBP는 중심 픽셀과 인접하는 8개의 각 픽셀과의 크기 비교를 하고 그 결과 논리값을 연결한 값이 되고, MCT는 윈도우내 픽셀들의 평균값과 9개 각 픽셀과의 크기 비교에 따른 논리값으로 값이 결정된다.

2. PDC (Pixel Direction Code)

PDC는 3×3 영역의 중심 픽셀의 기술기벡터를 이용하여 계산되는 픽셀 방향코드로 다음의 과정을

통해 구한다. 수평 마스크(M_x)와 수직마스크(M_y)를 입력 영상(I)에 각각 컨볼루션 적용하여, 크기성분($G(\mathbf{p})$)와 방향성분($\theta(\mathbf{p})$)로 표현되는 기울기 벡터를 계산한다.

$$G_x(\mathbf{p}) = M_x * I(\mathbf{p}) \quad (3)$$

$$G_y(\mathbf{p}) = M_y * I(\mathbf{p}) \quad (4)$$

$$G(\mathbf{p}) = \sqrt{G_x(\mathbf{p})^2 + G_y(\mathbf{p})^2} \quad (5)$$

$$\theta(\mathbf{p}) = \tan^{-1} \left(\frac{G_y(\mathbf{p})}{G_x(\mathbf{p})} \right) \quad (6)$$

픽셀 \mathbf{p} 의 방향코드는, 0도에서 360도사이의 값을 가지는 방향성분을 구간화하여 얻어지는 인덱스 값과 관련이 있으며 식 7과 같이 표현된다.

$$K(\mathbf{p}) = \begin{cases} \lfloor \frac{\theta(\mathbf{p})}{\omega} \rfloor & \text{if } G(\mathbf{p}) > \tau \\ N & \text{otherwise} \end{cases} \quad (7)$$

영상의 밝기 값 변화가 어느 정도 두드러지게 존재할 경우($G(\mathbf{p}) > \tau$), 방향코드는 방향각도를 각도 해상도 ω 로 나눈 값으로 할당된다. 만약 기울기 벡터의 크기 성분이 임계값 τ 보다 작으면 방향코드 값으로 N 이 할당된다. 여기서 $N=360/\omega$ 이다. 예를 들어 ω 를 45로 할 경우, 크기 성분이 τ 보다 큰 픽셀은 0,1,...,7의 코드값 중 하나로 할당되며, 크기 성분이 τ 보다 작은 픽셀은 코드 8로 할당된다.

그림 2는 입력 얼굴 영상(a)과 영상으로 표현된 크기성분(b) 및 방향성분(c), 그리고 영상으로 표현된 방향코드(d), MCT(e), LBP(f)를 보여준다. 방향코드 영상에서 눈, 코, 입 등 주요 얼굴구성 요소를 이루는 픽셀들에서는 구분되는 코드가 할당되며, 배경에 해당하는 얼굴 면에서는 동일한 코드가 할당됨을 알 수 있다. LBP 영상에 비해 MCT 영상은 값의 동적 범위가 넓어 보다 부드럽게 보인다.

3. 간단한 고속 객체 분류기

본 논문에서는 PDC 특징을 룩업테이블(LUT; Look Up Table) 기반의 분류기에 적용한다. 이 LUT 분류기는 양성과 음성, 두 종류의 클래스를 구분하는 이진 분류기로 분류기의 입력 특징은 LUT 인덱스 값이다. LUT 분류기의 분류 규칙은 사전 학습에 의해 파악된 통계특성, 즉 입력 영상의 특징(LUT 인덱스)이 특정한 값으로 나타났을 때, 통계적으로 양성이 나타날 확률과 음성이 나타날 확률 중 어느 것이 높았는가에 따라 인식 결과 클

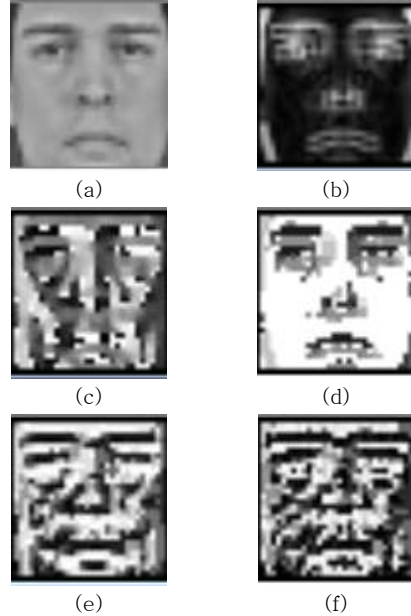


그림 2. (a) 입력 얼굴 영상, (b) 기울기벡터의 크기성분, (c) 기울기벡터의 방향성분, (d) 방향코드, (e) MCT, (f) LBP

Fig. 2 (a) Input face image, (b) magnitude of gradient vector, (c) angle of gradient vector, (d) direction code (e) MCT, (f) LBP

래스를 결정하는 것이다.

이를 수학적식으로 표현하면 식 8과 같다. 양성데이터에 대한 누적 빈도 테이블 L^{pos} , 음성데이터에 대한 누적 빈도 테이블 L^- 은 인덱스(특징)의 종류 ($N+1$)만큼의 빈도로 구성되어 있으며, 인덱스 k 로 접근한다. 분류기 h 의 클래스 결정은 식 8과 같이 $L^{pos}(k)$ 와 $L^-(k)$ 의 값 비교를 통해서 이루어진다.

$$h(\mathbf{x}, k) = \begin{cases} +1 & L^{pos}(k) > L^-(k) \\ -1 & \text{otherwise} \end{cases} \quad (8)$$

이와 같이 동작하는 LUT 분류기들은 AdaBoost 알고리즘을 사용하여 보다 효과적으로 학습되고 결합될 수 있다. AdaBoost 알고리즘 학습과정에서는 개별 약분류기 학습과 이들의 '결합'이 동시에 이루어진다. 학습 라운드마다 각각 다르게 부여되는 샘플 가중치는, 분류기들의 가중화된 결합뿐만 아니라 각 분류기의 학습에도 반영된다. 히스토그램 계산시 샘플의 가중치를 반영함으로써, LUT 분류기의 룩업테이블을 다음과 같이 구현한다.

$$D_i(k) = D(i) I(f(\cdot) = k) \quad (9)$$

$$L^-(k) = \sum_{\substack{i \in \\ i \in N_+}} D(i) I(f(\mathbf{x}) = k) \quad (10)$$

여기서 k 는 LUT의 인덱스이며, i 는 학습 샘플 데이터의 인덱스, N_{pos} 는 양성 샘플 데이터들의 인덱스 집합, N_- 는 음성 샘플들의 인덱스 집합이다. $I(\cdot)$ 는 인디케이트 함수로서 괄호내의 논리값이 참이면 1, 아니면 0의 값을 출력한다. $D(i)$ 는 인덱스 i 인 샘플 데이터의 가중값으로 AdaBoost 학습의 매 단계에서 샘플의 인식 결과에 따라 자동으로 변경된다.

4. AdaBoost를 이용한 LUT 약분류기 결합

AdaBoost는 분류율이 높지 않은 약분류기들을 순차적으로 학습한 후 이들을 선형 결합하여 높은 분류율을 가지는 강분류기를 구현하고자 할 때 사용되는 방법이다. AdaBoost 알고리즘을 보다 상세히 설명하면 다음과 같다.

AdaBoost 라운드: For $t=1, \dots, T$

1. 주어진 샘플 가중치를 반영하여 약분류기들을 학습하고, 오류가 가장 작은 약분류기를 선택
2. 약분류기 h_t 의 가중치 α_t 를 결정
3. 약분류기 가중치 α_t 와 인식 여부에 따라 샘플들의 가중치를 변경

AdaBoost 학습이 종료된 후 약분류기들의 선형 결합으로 생성되는 강분류기는 다음 식 11과 같이 표현된다.

$$C(\mathbf{x}) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(\mathbf{x}) \right), \quad (11)$$

여기서 \mathbf{x} 는 분류기의 입력 영상이며, $h_t(\mathbf{x})$ 는 t 번째 단계에서 선택된 이전 LUT 약분류기, α_t 는 그것의 가중치이다.

AdaBoost 라운드 1단계에서 약분류기의 선택기준은 다음의 식으로 표현되는 오류이다.

$$\epsilon_t = \sum_{i: h_t(\cdot) \neq y_i} D_t(i), \quad (12)$$

여기서 \mathbf{x}_i 는 인덱스가 i 인 샘플을 의미하며, y_i 는 그 샘플의 클래스이다. $D_t(i)$ 는 라운드 t 에서 \mathbf{x}_i 의

가중치이다. 부스팅 최초 라운드에서의 샘플 가중치 $D_1(i)$ 는 모두 동일하며, 모든 라운드에서 $\sum_i D_t(i) = 1$ 이 되도록 한다.

AdaBoost 라운드 2단계에서 약분류기의 가중치는 다음과 같이 표현된다.

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \quad (13)$$

AdaBoost 라운드 3단계에서 샘플들의 가중치는

$$D_{t+1}(i) = \frac{D_t(i) e^{-\alpha_t y_i h_t(\mathbf{x}_i)}}{Z_t} \quad (14)$$

와 같다. 여기서 Z_t 는 $\sum_i D_{t+1}(i) = 1$ 이 되도록 하기 위한 정규화항이다.

III. 실험 및 결과

1. 실험 데이터 및 방법

실험에 사용된 얼굴 영상은 학습 집합 3,200개, 테스트 집합 2,296개이며, 얼굴이 아닌 영상은 학습 집합 6,000개, 테스트 집합 14,727개이다. 얼굴 및 비얼굴 영상을 22×22의 명암도 영상으로 정규화하여 검출기의 입력으로 사용한다.

LUT 약분류기는 얼굴 영상을 이루는 전체 픽셀 위치에 존재할 수 있는데, 400개의 픽셀 포지션마다 구현될 수 있다. 상하좌우의 최외곽 픽셀들에 대해서는 3×3 마스크를 씌울수 없기 때문에 20×20의 픽셀들에 대해서만 약분류기를 구현한다. 400개의 LUT 약분류기들은 AdaBoost 알고리즘에 따라 학습되며 선택된다. 실험에서는 세 종류의 특징에 대해 AdaBoost 라운드를 30, 50, 70, 90, 110으로 하여 얼굴 검출기를 구현하였다. 특징별 약분류기의 수에 따른 검출률 비교 실험과 데이터셋별 검출률 비교 실험을 실시하였으며, 검출률과 오기각률에 실험도 실시하였다.

2. 실험 결과

AdaBoost 알고리즘에 따라 학습되고 선택된 각 특징별 LUT 약분류기들의 예를 그림 3에 나타내었다. 선택된 약분류기들 중 상위 20개의 약분류기들의 위치를 주황색 점으로 나타내었다. 세 특징 모두의 경우에서 얼굴의 특징적 구성요소인 눈, 코, 입 근방에서 약분류기들이 선택되었음을 보여준다. 특히 눈위치에 많은 약분류기가 분포하고 있어서 이

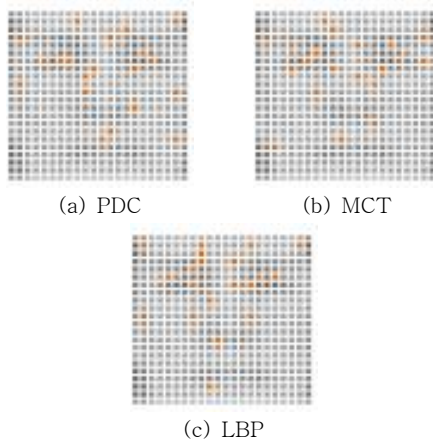


그림 3. 선택된 약분류기들의 위치
Fig. 3 Positions of weak classifiers

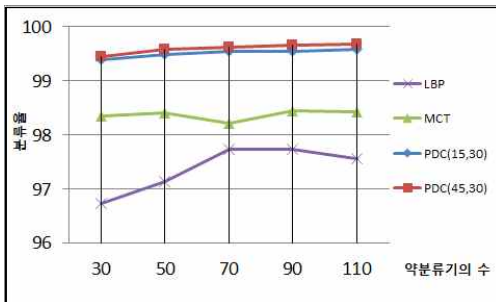


그림 4. 약분류기의 수에 따른 분류율 비교
Fig. 4 Comparison of classification rates with respect to the number of weak classifiers

위치가 얼굴이 아닌 것과 얼굴을 구분하는 중요한 특징 지점임을 알 수 있다.

선형결합되는 약분류기들의 수에 따른 각 특징별 얼굴 검출기의 분류 성능을 비교하기 위해, 약분류기들의 수를 30개에서 110개까지 변경하면서 분류율을 조사하였다. 테스트 데이터셋에 대해 식 (11)의 값이 0보다 크면 얼굴, 그렇지 않으면 비얼굴로 판단하는 것으로 분류율을 측정하였다. 그림 4는 약분류기의 수에 따른 특징별 분류율을 보여준다. 그림 4에서 알 수 있듯이 LBP 특징 얼굴 검출기가 가장 분류율이 낮았으며 제안한 PDC 특징 얼굴 검출기의 분류율이 가장 높았다. PDC 특징 분류기는 PDC(τ , γ)로 식 (7)의 파라미터 표기 방식에 따라 표현된다. 즉, PDC(45, 30)의 경우 방향코드 계산을 위한 각도 구간 해상도가 45도이며, 크기 성분은 30 이하인 픽셀의 방향코드는 8로 할당된

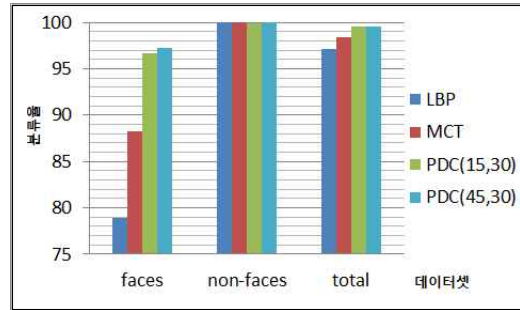


그림 5. 데이터셋에 대한 분류율
Fig. 5 Comparison of classification rates with respect to the data set

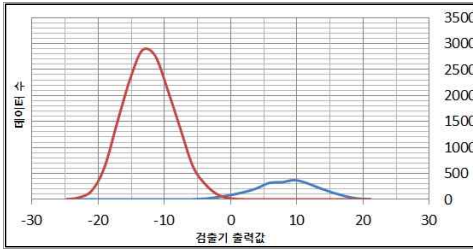
다.

특징별 LUT 약분류기의 크기는 테이블에 저장되는 정보의 양에 따라 결정된다. 50개의 LBP 특징으로 분류를 하는 LBP 검출기(50개의 LUT 약분류기 결합)는 $50 \times (256 + 2)$ 개의 숫자가 저장되어야 한다. 256은 LBP 특징 인덱스의 범위이고, 2는 약분류기의 픽셀 위치 정보이다. MCT 검출기는 $50 \times (511 + 2)$, PDC(15, 30) 검출기는 $50 \times (25 + 2)$, PDC(45,30) 분류기는 $50 \times (9 + 2)$ 개의 숫자가 저장된다. 제안한 PDC 검출기가 테이블의 크기가 가장 작다.

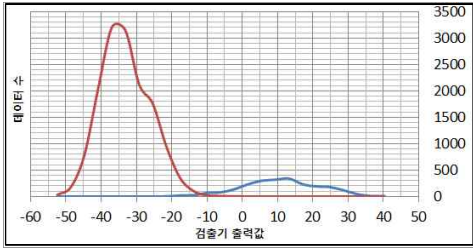
각 특징별 얼굴 검출기를 세분화된 데이터셋에 대해 적용하여 분류율을 조사해보았다. 얼굴 검출기의 약분류기의 수는 50개인 경우이다. 얼굴 데이터셋의 경우에서 특징별 분류율이 보다 더 뚜렷하게 차이가 남을 그림 5을 통해 알 수 있다. 비얼굴 데이터에 대한 분류율은 비슷하지만, 얼굴에 대해서는 제안하는 PDC 특징 분류기가 가장 우수한 것으로 나타났다.

얼굴 검출기의 최고 분류율을 살펴보기 위해 PDC 검출기와 MCT 검출기의 출력값 분포 특성을 그림 6에 나타내었다. 그림 6을 보면 검출기의 출력값이 0이하인 곳에서 대부분 분포하는 음성데이터와 0이상에서 대부분 분포하는 양성데이터를 구분하는 최적 임계값은 0보다 작은 값을 알 수 있다. 즉 0보다 작은 임계값을 식 (11)의 적용했을 때 최고 분류율을 얻을 수 있다.

그림 7에 최적 임계값을 적용했을 때의 특징별 최고 분류율을 나타내었다. 최적 임계값 적용전인 그림 5의 결과와 비교해보면, 세 종류의 특징별 얼굴 검출기의 성능 향상이 두드러지며, 제안하는 PDC 특징의 분류 성능이 여전히 가장 우수하다는 것을 알 수 있다. 그림 8에서 캡션 옆의 숫자는 해당 검출기의 최적 임계값이다.



(a) PDC



(b) MCT

그림 6. 얼굴 검출기의 출력값 분포

Fig. 6 Distributions of output values of face classifiers

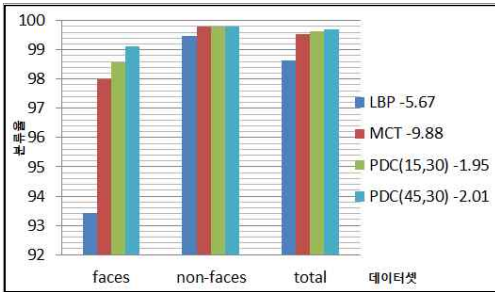


그림 7. 데이터셋 종류에 따른 최고 분류율

Fig. 7 Comparison of the best classification rates with respect to the data set

본 연구에서는 각 방법별 성능을 검출률과 오기각률 측면에서도 분석하였다. 오기각률은 비얼굴 영상 입력임에도 제대로 기각하지 못하고 얼굴 영상으로 잘못 분류한 것을 의미한다. PDC 특징은 구간 각도 해상도를 30과 60인 경우도 추가 하여 실험하였다. 임계치를 변경시키면서, 얼굴 분류기 출력값이 임계치 보다 크면 얼굴, 그렇지 않은 경우 배경으로 분류하여 두 가지 측정값을 얻었다. 그림 8을 살펴보면 동일한 오기각률에서의 검출률 측면에서, 제안하는 PDC 방법이 LBP 방법, MCT 방법 대체로 보다 우수함을 알 수 있다. PDC(60,30)의 경우 비교적 높은 오기각률에서는 MCT보다 검출률이 우

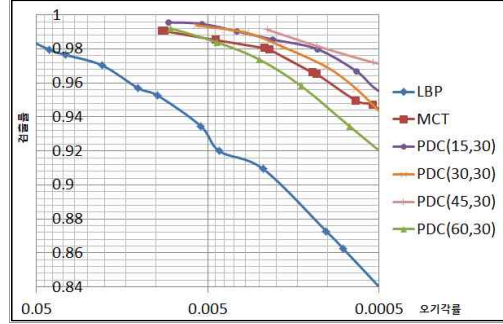


그림 8. 검출률 및 오기각률

Fig. 8 Detection rates(vertical axis) vs. false positive rates(horizontal axis) of LBP, MCT, and PDC classifiers

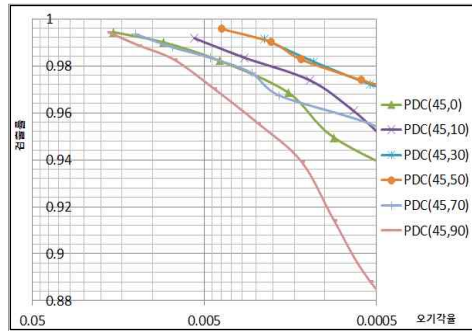


그림 9. PDC 검출률 및 오기각률

Fig. 9 Detection rates(vertical axis) vs. false positive rates(horizontal axis) of PDC classifiers

수하였지만, 오기각률을 낮게 했을 경우는 MCT보다 검출률이 좋지 않았다. 하지만 각도 해상도가 60인 경우 PDC 약분류기가 겨우 7개의 숫자 값을 저장하는데 비해, MCT의 경우는 511개의 숫자 값을 저장하는 것을 고려할 필요가 있다. PDC 방법간의 비교에서는 각도 구간 해상도를 45도로 하는 경우가 가장 우수하였다.

기울기 벡터의 크기성분 임계값 τ 에 따른 PDC 검출기의 검출률과 오기각률을 살펴보았다. 그림 9는 각 검출기의 검출률과 오기각률을 나타낸다. PDC(45, 30)과 PDC(45, 50)이 높은 검출률과 낮은 오기각률로 가장 성능이 우수함을 알 수 있다. 크기성분 임계값이 70인 경우는 임계값이 0인 경우와 비슷한 성능을 나타내며, 임계값이 90으로 큰 경우는 임계값이 0인 경우보다도 더 좋지 않은 성능을 나타내고 있다는 것을 주목할만하다.

IV. 결 론

본 논문에서는 기존의 Haar-like, MCT, LBP 특징 기반의 부스팅 방법을 개선한 얼굴 검출 방법을 제안하였다. 고속 고정확도의 얼굴 검출기를 구현하기 위해, 얼굴 검출에 유용한 방향 코드 특징과 그 특징을 효율적으로 활용하는 특업 테이블 분류기의 구현방법을 제안하였다.

제안하는 픽셀 방향코드 특징은 지역 기울기벡터의 크기성분과 방향성분을 조합함으로써, 기존의 LBP, MCT 등 지역 이진 구조 특징을 이용하는 경우 보다 얼굴과 비얼굴 데이터의 분별에 보다 더 유리하였다. 특업테이블 분류기는 특징값의 동적 범위에 해당하는 인덱스를 가지도록 설계하였으며 얼굴과 비얼굴의 분류에 효과적이었다. 방향 코드 특징을 입력으로 하는 특업테이블 분류기는 기존의 MCT, LBP 특징 방법에 비해 메모리 요구량이 훨씬 적은 장점도 있다. 또한 기존의 LBP, MCT 등 지역 특징을 이용하는 방법들보다 얼굴과 비얼굴 분류율이 높으며, 동일한 오기각률에서 검출률도 높았다.

References

[1] H. Rowley, S. Baluja, T. Kanade, "Neural network-based face detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, pp. 23-38, 1998.

[2] K. Sung T. Poggio, "Example-based learning for view-based human face detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, pp. 39-51. 1998.

[3] B. Heisele, T. Serre, S. Prentice, T. Poggio, "Hierarchical classification and feature reduction for fast face detection with wupport vector machines," Pattern Recognition, Vol. 36, No. 9, pp. 2007- 2017, 2003

[4] D. Le, S. Satoh, "A multi-stage approach to fast face detection," IEICE Trans. Inf. & Syst. Vol. E89-D, No. 7, pp. 2275-2285, 2006.

[5] P. Viola M. Jones, "Rapid object detetection using a boosted cascade of simple features," Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition,

pp. 511-518, 2001.

[6] Y. Freund, R. Schapire, "A Short Introduction to Boosting," Journal of Japanese Society for Artificial Intelligence, Vol. 14, pp. 771-780, 1999.

[7] B. Fröba, A. Ernst, "Face detection with the modified census transform," Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, pp. 91-96, 2004.

[8] L. Zhang, R. Chu, S. Xing, S. Liao, S. Li, "Face detection based on multi-block LBP representation," Lecture Notes in Computer Science, Vol. 4642, pp. 11-18, 2007.

[9] D. Huang, C. Shan, M. Ardebilian, L. Chen, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey," IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol. 41, No. 6, pp. 765-781 , 2011.

[10] B. Jun, D. Kim, "Robust facedetection using local gradient patterns and evidence accumulation," Pattern Recognition, Vol. 45, No. 9, pp. 3304-316, 2012

[11] T. Ojala, M. Pietikäinen, D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, Vol. 29, No. 1, pp. 51-59, 1996.

저 자 소 개

임길택



1993년 경북대학교 전자공학 공학사.

1995년 경북대학교 전자공학 공학석사.

1999년 경북대학교 전자공학 공학박사.

현재, 한국전자통신연구원 대경권연구센터 책임연구원.

관심분야: 컴퓨터비전, 패턴인식, 비디오보안
Email: ktl@etri.re.kr

강 현 우

2005년 경북대 컴퓨터 공학과 공학사.

2009년 경북대 전자전기 컴퓨터학부 공학석사.

현재, 한국전자통신연구원 대경권연구센터 선임연구원.

관심분야: 컴퓨터통신, 임베디드 소프트웨어, 컴퓨터 비전

Email: hwkang@etri.re.kr

한 병 길

2005년 경북대학교 전자 전기공학부 공학사.

2007년 경북대학교 전자 공학과 공학석사.

현재, 한국전자통신연구원 대경권연구센터 선임연구원.

관심분야: 영상처리, 로봇비전, 비디오보안

Email: kilyhan@etri.re.kr

이 종 략

2005년 한국과학기술원 (KAIST) 전기및전자공학과 공학사.

2007년 텍사스주립대 (University of Texas at Austin) 전자컴퓨터공학 공학석사.

2012년 텍사스주립대(University of Texas at Austin) 전자컴퓨터공학 공학박사.

현재, 한국전자통신연구원 대경권연구센터 연구원.

관심분야: 컴퓨터비전, 행동인식, 패턴인식

Email: jongtaeklee@etri.re.kr