

논문 2016-53-2-12

판별 얼굴 기술자 기반의 다중 해상도 분할 영역 히스토그램을 이용한 얼굴인식 방법

(Face Recognition Using Histograms of Multi-resolution Segments Based on Discriminant Face Descriptor)

이 장 윤*, 이 용 결*, 최 상 일**

(Jang-yoon Lee, Yonggeol Lee, and Sang-Il Choi[©])

요 약

본 논문에서는 얼굴 영상의 지역 정보를 효과적으로 활용하기 위해, 부분 영상에 대한 다중 해상도 히스토그램을 이용한 얼굴 인식 방법을 제안한다. 기존 DFD의 경우 단일 크기로 나누어진 부분 영역의 히스토그램을 통합하여 유사도를 비교하나, 이는 부분 가림이나 조명변이로 인해 변형된 영역이 단일 부분 영역 내에서 발생하지 않고 여러 개의 부분 영역에 걸쳐 발생할 수 있기 때문에, 지역 정보들의 특성을 활용하는 데에 효과적이지 못하다. 본 논문에서는 각각의 부분 영역에 대해 다중 해상도로 분할하여 여러 종류의 크기에 해당하는 부영역의 히스토그램을 사용함으로써, 인식 과정에서 지역 정보의 손실을 최소화하고자 하였다. YaleB, AR, CAS-PEAL-R1 데이터베이스에 대해 인식 실험을 수행한 결과, 제안한 방법이 여러 종류의 변이가 있는 경우에 인식 성능을 향상시키는 것을 확인 할 수 있었다.

Abstract

We propose a face recognition method using the histograms of multi-resolution segments in order to effectively utilize the local information of faces. Since the variations in faces can occur in various sizes, the DFD method, which uses the histograms from the sub-regions of the same size, is not effective for obtaining local information of faces. In this paper, we first divide an image into several sub-regions and extract the DFD(Discriminant Face Descriptor) from each sub-region. By dividing each sub-region into several segments with multi-resolution and extracting histograms for each segment, we reduce the loss of local information in the process of recognition. The experimental results for the Yale B, AR, CAS-PEAL-R1 databases show that the proposed method improves the recognition performance compared to the existing DFD based method.

Keywords : Face Recognition, Discriminant Face Descriptor, Multi-resolution Segmentation, Illumination Variation

I. 서 론

얼굴 인식은 특별한 동작이나 행위에 대한 요구 없이 사람을 인식하고 식별하는 기술로서, 출입 통제나 정보 단말기 인증 시스템 등의 보안 분야, 범죄자 검색, 멀티 미디어 검색 등 다양한 시스템에 사용될 수 있다. 최근

에는 스마트폰에 고성능 카메라가 지원되면서 얼굴 인식을 응용하는 다양한 프로그램들이 크게 증가하고 있다. 이러한 효용성으로 인해 그동안 다양한 특징 추출 방법 및 인식기가 개발되어 왔다.

얼굴 인식을 위해 얼굴 영상으로부터 특징을 추출하는 다양한 기술자(descriptor)들이 제안 되어 왔다. 기울기(gradient) 기반의 방법들 가운데, SIFT(Scale Invariant Feature Transform)^[1] 기술자는 특징점이 검출되면 특징점 부근의 관심 영역을 양자화 하여 히스토그램 기술자를 생성한다. SIFT는 크기, 방향, 회전 변이에 강인한 특징을 추출하지만 그만큼 연산 량이 많아진

* 학생회원, ** 정회원, 단국대학교 컴퓨터학과
(Department of Computer Science and Engineering,
Dankook University)

© Corresponding Author(E-mail: choisi@dankook.ac.kr)

Received ; August 20, 2015 Revised ; January 16, 2016

Accepted ; January 21, 2016

다는 단점이 있다. SIFT의 변형으로는 GLOH(Gradient Location and Orientation Histogram)^[2], SURF(Speeded Up Robust Feature)^[3] 등 다양한 방법들이 개발되었다. GLOH 기술자는 SIFT 기술자의 강건함과 변별력을 향상시키기 위하여 4×4 크기의 보조 영역 대신 log-polar location grid에서 총 272차원의 기술자를 생성한다. SURF 방법은 적분 영상(integral image)을 이용하여 SIFT보다 연산 속도를 향상시켰다.

텍스처(texture)기반의 지역 패턴(local pattern)을 이용하는 방법들도 제안되었는데, Ramin와 Woodfill이 제안한 CT(Census Transform)^[4]는 3×3 크기의 윈도우 내 중심값을 결정하기 위해 중심 픽셀의 주변부의 8개 픽셀과 평균값을 비교하여, 대소 관계에 따라 1 또는 0을 할당하고 이들 값을 일정한 방향으로 접합시킨다. 중심값은 1부터 $255(=2^8-1)$ 까지의 범위를 갖고, 픽셀의 주변 값들과의 상관관계를 이용하기 때문에 조명변이 환경에 대하여 강인한 성능을 보인다. Froba와 Ernst가 제안한 MCT(Modified Census Transform)^[5]은 중심값을 결정할 때, 윈도우 내 9개의 전체 픽셀과 평균값을 비교함으로써 1부터 $511(=2^9-1)$ 까지의 범위를 갖는 중심값을 추출한다. MCT는 CT와 달리 1개 비트만큼의 추가적인 비교 연산이 수행된다는 단점이 있지만, 풍부한 지역 정보를 추출함으로써 기존의 방법보다 높은 인식 성능을 나타내었다. LBP(Local Binary Pattern)^[6]는 인식 뿐만 아니라 검출, 추적 알고리즘 등에서도 널리 사용되고 있는 방법으로, 중심 화소 값과 이웃하는 화소 값과의 대소 관계를 통해 텍스처(texture) 정보를 추출한다. LBP는 영상에 변이가 발생하여도 화소 값의 크기 순서가 유지되는 한 좋은 성능을 보인다는 장점이 있지만, 고주파수 영역에서는 인접 픽셀마다 화소 값의 크기 순서가 바뀔 수 있어 고주파 잡음이 많이 발생하는 환경에서는 인식률이 떨어지는 경향이 있다^[7~10]. LBP의 변형(variant)으로서, uniform LBP^[11]는 특징점 정보의 대부분이 일부 LBP 코드에 집중되어 있다는 실험 결과를 근거로, 중심 픽셀을 기준으로 p 개의 주변 픽셀들을 고려하여 낮은 차원의 코드를 생성한다. CS-LBP(Center-Symmetric Local Binary Pattern)^[12] 방법은 중심 픽셀을 기준으로 대칭되는 픽셀만을 비교하여 LBP 코드를 생성한다. CS-LBP 코드는 8차원의 코드를 생성하기 때문에 기존 LBP 코드보다 차원은 낮지만 평활한 영역에서 가우시안 노이즈에 취약하다는 단점을 보인다. 이 외에도 메모리 사용과 매칭 속도의 개선을 위하여 두 픽셀의 밝기 값 차이로

이진 표현자를 생성하는 Brief(Binary robust independent elementary features)^[13] 기술자가 제안되었다.

한편, 지역 정보를 정해진 일정한 형식에 의해 추출하는 기존의 방법들과 달리 지역 패턴들로부터 판별력을 갖는 기술자를 학습하는 방법들이 제안되었다. DLBP(Discriminative Local Binary Pattern)^[14]는 LBP를 기반으로 영상 내 모든 5×5 크기의 윈도우 내 중심 값과 비교 가능한 모든 8개의 셋들로부터 Fisher criterion^[15]에 따라 가장 판별력이 높은 위치의 픽셀들을 선택한다. 판별 얼굴 기술자(Discriminant Face Descriptor, DFD)^[16]는 기존의 LBP 방법을 개선하여 같은 사람의 영상들에서는 특징 값들의 차이를 최소화하고 다른 사람의 영상들에서는 특징 값들의 차이를 최대화하는 방법을 이용하는 기술자(descriptor)이다. DFD는 일정한 크기의 부분 영역에 대해 판별 영상 필터 학습과 최적 이웃 샘플링을 통해 얼굴 인식을 위한 최적의 히스토그램을 추출한다. 하지만 얼굴 인식 과정에서 부분 영역에 대한 히스토그램 상호 교차 방법(intersection)^[17]을 사용할 경우, 본래 영상에 있는 얼굴의 지역적인 정보를 잃어 버려 조명변이 혹은 부분 가림에 있어서는 강인한 인식률을 보이지 못한다는 단점이 있다. 이는 얼굴 표정과 같이 윤곽선의 변형이 있는 영상을 인식하는데 강인하지만, 부분 가림이나 조명변이와 같은 넓은 영역에 걸친 변형이 있는 영상의 경우에는 취약하다. 즉 얼굴 영상의 일부분에 과도하게 음영이 발생하거나 타 물체로 가려지게 되면 해당 영역은 히스토그램 상호 교차 과정을 거치더라도 유효한 특징 값을 갖지 못하게 된다.

따라서 본 논문에서는 이러한 DFD의 단점을 보완하기 위하여 부분 영역을 여러 단계로 나누어, 단계별 히스토그램을 함께 사용하는 방법을 제안한다. 제안한 방법은 기존의 DFD와 달리, 여러 단계의 해상도로 다중 분할 히스토그램을 추출함으로써, 영상의 지역 정보의 손실을 최소화할 수 있다. 얼굴 인식 연구에서 많이 사용되는 YaleB 데이터베이스^[18], AR 데이터베이스^[19]와 CAS-PEAL-RI 데이터베이스^[20]에 대해 인식률을 평가해 본 결과, 제안한 방법을 통해 DFD의 성능이 향상됨을 확인할 수 있었다.

본 논문은 다음과 같이 구성되어 있다. II에서는 DFD에 대해 설명하고, 제안한 다중 해상도 분할 영역의 히스토그램 추출 과정을 설명한다. III에서는 다양한 얼굴 영상 데이터베이스에 대해 제안한 방법의 성능을 평가하고, 마지막으로 IV에서 결론을 맺는다.

표 1. DFD에 사용되는 변수 요약
Table 1. Summaries of the variables of DFD

Symbol	Definition
L	number of face classes
C_i	number of samples from the i th class
N	neighboring patches
dI_{ij}^p	PDM for j th image of i th class at position p
dm_{ij}^p	mean PDM for the i th class
dm^p	total mean PDM at position p

II. 본 론

DFD는 분별력(discriminant) 있는 기술자를 추출하기 위해 세 단계의 과정을 거친다. 첫째로 판별 영상 필터 학습을 거치고, 둘째로 최적 이웃 샘플링을 통해 소프트 샘플링을 하며, 셋째로 주요 패턴들을 통계적으로 구성한다^[14].

DFD의 히스토그램 특징을 이용한 얼굴 인식 기법은 얼굴 영상 내에서 임의로 구분시킨 부분 영역에서 PDM(Pixel Difference Matrix)^[21]을 추출하여 주요 패턴을 찾아내고 그 패턴들의 히스토그램들을 추출한다. 하지만 얼굴 인식을 위해 동일한 크기로 분할 된 영역에서 추출된 히스토그램들을 통합하여 비교를 수행하는 것은 실제 영상의 지역적인 특징을 인식과정에 충실히 반영하는데 어려움이 있다^[22]. 따라서 본 논문에서는 얼굴 인식에 지역 정보를 효과적으로 활용하기 위해, 각 부분 영역에서 PDM을 추출하여 히스토그램 특징을 추출하기에 앞서 해당 영역을 다중 해상도로 분할하여 개별 해상도로 분할 된 영역에서 각각 추출한 히스토그램 특징을 가중치 결합을 통해 함께 사용하는 방법을 제안한다.

1. DFD(Discriminant Face Descriptor) 학습

DFD 방법은 먼저 각 픽셀마다 픽셀 단위 간 차이 값을 나타내는 PDM이라는 행렬을 추출한다. DFD에서 사용되는 변수들의 요약을 표 1에 나타내었다. 영상 I 가 주어졌을 때, I 를 필터 f 로 필터링 한 영상을 $f(I)$ 라고 하면, PDM $dI^p \in R^{d_1 \times d_2}$ 로부터 만들어진 PDV(Pixel Difference Vector) $df(I)^p$ 는 다음과 같이 나

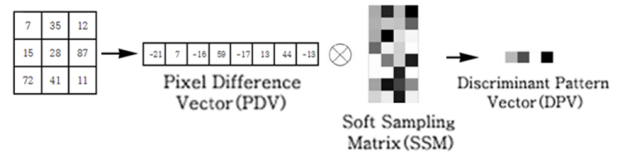


그림 1. SSM을 통한 DPV 생성 과정
Fig. 1. Generation of DPV using SSM.

타낼 수 있다^[21].

$$df(I)^p = [f(I)^{p_1} - f(I)^{p_0}, f(I)^{p_2} - f(I)^{p_0}, \dots, f(I)^{p_d} - f(I)^{p_0}] \quad (1)$$

여기서 p 는 필터의 위치를 나타내며 $p_i (i=1, \dots, d)$ 는 p 위치에 있는 필터 내부에서의 부 필터(sub-filter) 위치를 의미한다. DFD에서는 분별력 있는 필터를 설계하기 위해 Fisher criterion를 적용하는데 f 를 영상 필터 벡터 w 로 가정하면 $f(I)^p$ 는 $w^T I^p$ 로 표현할 수 있고, 결과적으로 분별력 있는 PDV는 아래와 같이 정의된 클래스 내 산란 행렬(within-class scatter matrix) S_w 과 클래스 간 산란 행렬(between-class scatter matrix) S_b 로부터 $S_b w = \lambda S_w w$ 를 만족시키는 w 에 의해 $df(I)^p = w^T dI^p$ 로 구할 수 있다.

$$S_w = \sum_{i=1}^L \sum_{j=1}^{C_i} \sum_{p=1}^N w^T (dI_{ij}^p - dm_i^p) (dI_{ij}^p - dm_i^p)^T w, \quad (2)$$

$$S_b = \sum_{i=1}^L \sum_{p=1}^N C_i w^T (dm_i^p - dm^p) (dm_i^p - dm^p)^T w \quad (3)$$

dI_{ij}^p 는 i 번째 클래스 내 j 번째 영상으로부터 추출된 PDM이며 dm_{ij}^p 는 i 번째 클래스의 평균 PDM 값, dm^p 는 모든 PDM의 평균 값을 나타낸다.

w 를 이용하여 구한 $df(I)^p (=w^T dI^p)$ 는 SSM(Soft Sampling Matrix) $v = [v_1, \dots, v_d]^T$ 에 투영되는데, SSM은 $df(I)^p$ 를 저차원의 최적 DPV(Discriminant Pattern Vector)를 구하기 위하여 이용된다^[21]. d_2 개의 이웃을 갖고 있는 각 PDM은 다음과 같이 정의된 S_w 와 S_b 로부터 Fisher criterion에 따라 같은 클래스 내의 공분산과 다른 클래스 간의 공분산의 비율을 최대로 하는 $w \in R^{d_1 \times d_1}$ 와 $v \in R^{d_2 \times d_2}$ 에 의해 판별력 있는 부분 공간(discriminant subspace)에 투영되어 $d_1' \times d_2'$ 차원의 DPV를 만든다(그림 1). 이렇게 만들어진 DPV들로부터 K-means 클러스터링^[23]을 통해 DP(Dominant Pattern)가 결정되고, 각 DP들에는 고유의 라벨 값이 부여된다.

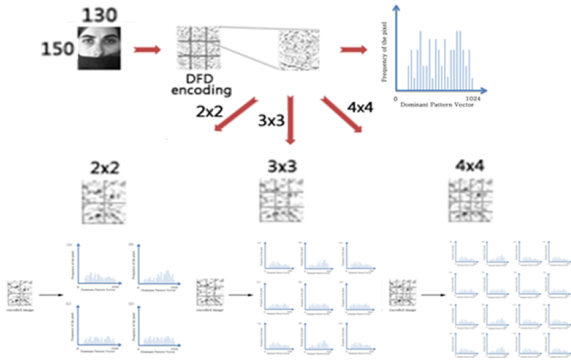


그림 2. 다중 해상도 분할 영역의 히스토그램 추출
Fig. 2. Histogram extraction of the multi-resolution segment.

SSM 을 통한 DPV 생성 과정은 그림 1에 나타나 있다.

2. 다중 해상도 분할 영역의 히스토그램 추출

주어진 영상에 대한 DFD는 영상의 부분 영역마다 개별 픽셀의 DPV를 구하여, 학습을 통해 만들어진 DP 들 중 가장 유사한 DP의 라벨 값으로 해당 픽셀 값을 인코딩(encoding)함으로써 생성된다. 얼굴 인식 과정은 각 부분 영역의 DFD 히스토그램을 모두 통합하여 영상 간의 히스토그램 상호 교차 방법을 통해 이루어진다. 그런데 이렇게 단순히 동일한 크기로 부분 영역의 히스토그램을 통합하여 비교를 하는 것은 지역 정보들을 특성을 활용하는 데에 효과적이지 못하다. 예를 들어, 부분 가림이나 조명변이로 인해 변형된 영역이 단일 부분 영역 내에서만 발생하지 않고 여러 개의 부분 영역에 걸쳐서 발생한 경우, 여러 개의 부분 영역의 히스토그램이 왜곡된 채로 교차과정에 사용될 수 있다. 반면, 부분 영역이 얼굴의 구성 요소(눈, 코, 입 등)보다 작을 경우에는 구성 요소들의 기하학적 특징들을 제대로 반영하지 못하게 된다. 따라서 본 논문에서는 각각의 부분 영역에 대해 다중 해상도로 분할하여 여러 종류의 크기에 해당하는 부영역의 히스토그램을 사용하여 지역 정보의 손실을 최소화하고자 하였다.

그림 2는 제안한 방법의 다중 해상도 분할 영역의 히스토그램 추출 과정을 보여준다 (본 논문에서는 얼굴 영상을 150×130 크기로 정렬하였다). 그림 2에서, 영상의 각 픽셀은 해당 부분 영역의 DP 라벨 값으로 인코딩 되는데(DFD 인코딩), 부분 영역의 히스토그램을 바로 추출하는 대신, 각 영역을 다중 해상도(2×2, 3×3, 4×4)로 다시 분할하여 단계별 영역 히스토그램을 추출한다. 이를 통해 보다 다양한 지역 정보를 활용할 수 있으며 세분화 정도는 그 분할치가 높으면 높

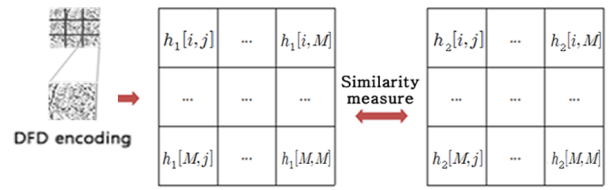


그림 3. 부분 영역의 다중 해상도 분할 영역 히스토그램 간 유사도 측정

Fig. 3. Comparison between the histograms of multi-resolution segment.

을수록 세밀한 지역 정보를 가질 수 있다.

3. 히스토그램 비교를 통한 영상의 유사도 측정

두 영상 I_1 와 I_2 의 유사도를 측정하기 위해 다중 해상도($M \times M$, $M=2,3,4$)로 분할한 영역에서 추출한 히스토그램들을 비교하였다. 그림 3에서와 같이, 한 개의 부분 영역에 대해, 단일 해상도로 분할했을 때 두 영상 I_1 와 I_2 에 대한 통합 히스토그램 $H_1^{M \times M}$ 와 $H_2^{M \times M}$ 는 각각 $\cup_{i=1, \dots, M} \cup_{j=1, \dots, M} h_1[i, j]$, $\cup_{i=1, \dots, M} \cup_{j=1, \dots, M} h_2[i, j]$ 이 된다. \cup 는 연결(concatenation) 연산자를 나타내며, 두 통합 히스토그램 $H_1^{M \times M}$ 와 $H_2^{M \times M}$ 의 유사도는 히스토그램 성분 간의 차이의 절대 값을 이용해 다음 식 (4)와 같이 측정하였다.

$$Sim_n(H_1^{M \times M}, H_2^{M \times M}) = \sum_{l=1}^m |H_1^{M \times M}(l) - H_2^{M \times M}(l)| \quad (4)$$

여기서 m 은 히스토그램의 전체 빈(bin)의 개수 이고, N 은 부분 영역의 개수이다. 그리고 $H_1^{M \times M}(l)$ 와 $H_2^{M \times M}(l)$ 은 $H_1^{M \times M}$ 와 $H_2^{M \times M}$ 의 l 번째 빈의 값을 나타낸다. 즉 $Sim(H_1^{M \times M}, H_2^{M \times M})$ 값이 작을수록 두 히스토그램은 유사하다고 볼 수 있다.

그러면, 단일 해상도 분할에 대한 두 영상 I_1 과 I_2 의 비교값 $COM(I_1, I_2)$ 은 다음과 같이 계산된다.

$$COM(I_1, I_2)^{M \times M} = \sum_{n=1}^N Sim_n(H_1^{M \times M}, H_2^{M \times M}) \quad (5)$$

다중 해상도를 사용할 경우, (본 논문에서는 세 종류의 해상도($M=2,3,4$)를 사용), 첫 번째 단계(level0)에서는 2×2 분할 크기로 구분하고, 두 번째 단계(level1)에서는 3×3, 세 번째 단계(level2)에서는 4×4 분할 크기로 구분하여 각 부영역마다 별도로 히스토그램을 구한 후, 이들 히스토그램들의 비교 측정치에 가중치 $(\alpha_1, \alpha_2, \alpha_3, \sum_{i=1}^3 \alpha_i = 1)$ 를 주어 최종적인 영상 간의 유사

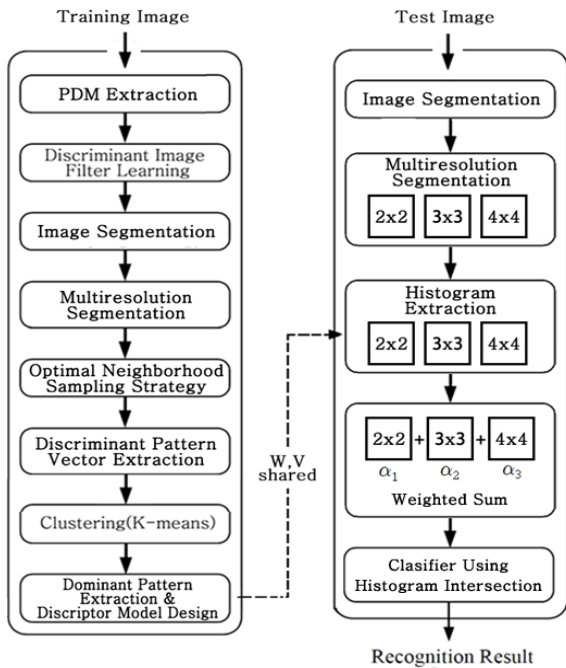


그림 4. 제안한 방법의 전체 흐름도
Fig. 4. Overall procedure of the proposed method.

도를 다음과 같이 측정하였다.

$$COM^{Final}(I_1, I_2) = \alpha_1 COM^{2 \times 2}(I_1, I_2) + \alpha_2 COM^{3 \times 3}(I_1, I_2) + \alpha_3 COM^{4 \times 4}(I_1, I_2) \quad (6)$$

제안한 방법의 전체 과정은 다음과 같이 요약할 수 있으며, 기존의 DFD를 이용한 얼굴 인식 과정과 제안한 방법의 전체 흐름을 그림 4에 나타내었다.

먼저 DP들을 학습하기 위해 주어진 영상의 부분 영역 별로 PDM 추출하고 판별 영상 필터 학습, 최적 이웃 샘플링 그리고 K-means 클러스터링 과정을 거친다. 학습 과정에서 구한 w, v 는 테스트 과정에서도 동일하게 사용된다. 부분 영역별로, 다중 해상도로 나누어진 DFD 히스토그램을 추출하여 해상도 별 통합 히스토그램($H^{M \times M}$)을 생성한다. 각각의 통합 히스토그램들을 비교한 측정치들($\sum_{n=1}^N Sim_n(H_1^{M \times M}, H_2^{M \times M})$)에 각각 다른 가중치를 주어 최종 비교 측정치들(COM^{Final})을 계산하여, 그 결과를 바탕으로 얼굴 인식률을 계산한다.

III. 실험

1. 실험 조건

얼굴 인식 연구에 많이 사용되는 YaleB 데이터베이스, AR 데이터베이스, CAS-PEAL-R1 데이터베이스



(a) AR 데이터베이스 예시 얼굴 영상



(b) CAS-PEAL-R1 데이터베이스 예시 얼굴 영상



(c) YaleB 데이터베이스 예시 얼굴 영상

그림 5. 여러 데이터베이스들의 영상 샘플들
Fig. 5. Samples of various databases.

들에 대하여 제안한 방법의 인식률을 평가하였다. 영상들은 데이터베이스의 종류에 따라 실제 환경에서 발생할 수 있는 조명, 포즈, 얼굴 표정, 부분 가림, 배경 등의 변이가 포함되어 있다(그림 5).

YaleB 데이터베이스는 10명의 정면 포즈 영상들을 포함하고 있으며, 각 사람당 45종류의 조명 조건에서 촬영된 영상들을 포함하고 있다. 10명에 대해 각각 3장씩, 총 30장의 영상을 동일하게 학습 영상과 갤러리 영상으로 사용하였으며, 나머지 420장은 테스트 영상으로 사용하였다.

AR 데이터베이스의 경우 118개의 클래스에 각 26장씩 총 3,068장으로 구성되어 있다. 영상에는 조명, 표정, sunglass나 scarf, 안경에 의한 부분 가림 등의 변이가 존재한다. 학습 영상으로는 각 클래스 당 sunglass나 scarf의 부분 가림이 없는 14장, 즉 총 1,652장이 사용되었으며, 부분 가림이 있는 나머지 1,416장의 영상에 대한 인식률을 측정하였다.

마지막으로 CAS-PEAL-R1 데이터베이스는 영상의 특성에 따라 두 종류의 데이터 영상을 구성하여 실험을 진행하였다. 먼저 첫 번째는 조명변이가 있는 Lighting 데이터 영상으로, 260개의 클래스에 각 9장씩 총 2,340장으로 구성되어 있다. 학습 영상으로는 각 클래스마다 3장씩 총 780장을 사용하였으며, 갤러리 영상으로는 각 클래스에서 1장씩을 사용하였다. 테스트 영상으로는 각 클래스에서 5장씩 총 1,300장을 사용하였다. 두 번째는 부분 가림이 있는 Accessory 데이터 영상으로, 436개의 클래스에 각 6장씩 총 2,616장으로 이루어져 있다. 학습 영상으로 각 클래스에서 2장씩 총 872장이 사용되었으며, 갤러리 영상으로 각 클래스에서 1장씩을, 테스

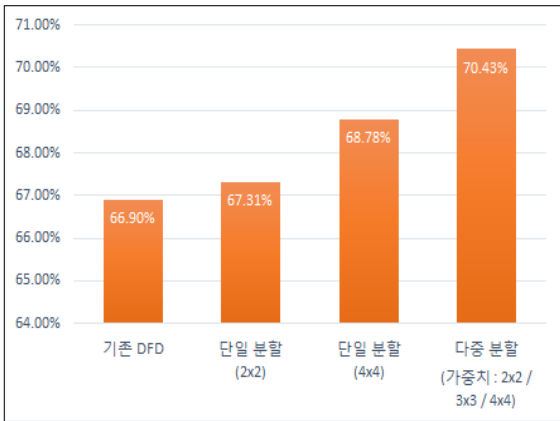


그림 6. 분할 세분화 정도에 따른 얼굴 인식률 (YaleB 데이터베이스)

Fig. 6. Face recognition rates for different sizes of segmentations. (YaleB database)

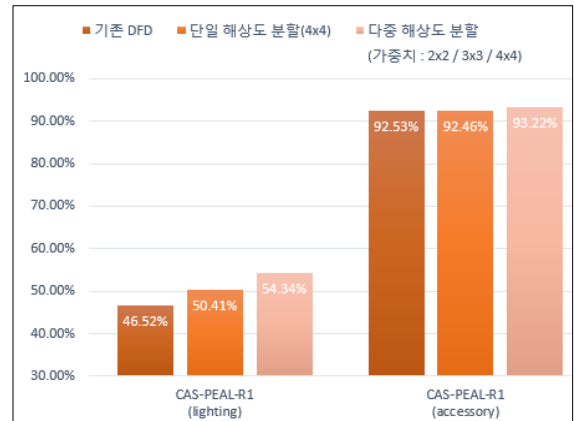


그림 8. 분할 세분화 정도에 따른 얼굴 인식률 (CAS-PEAL-R1 데이터베이스)

Fig. 8. Face recognition rates for different sizes of segmentations. (CAS-PEAL-R1 database)

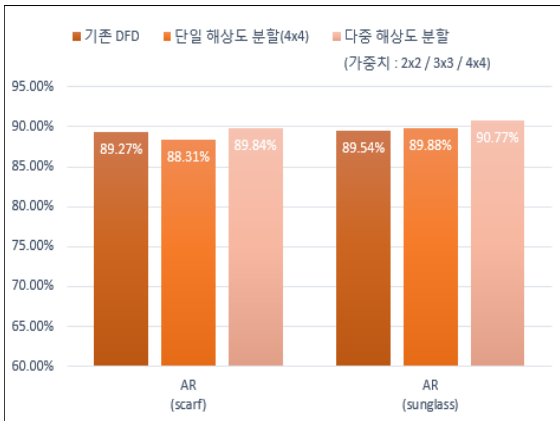


그림 7. 분할 세분화 정도에 따른 얼굴 인식률 (AR 데이터베이스)

Fig. 7. Face recognition rates for different sizes of segmentations. (AR database)

트 영상으로 각 클래스에서 3장씩 총 1308장을 사용하였다.

인식 실험에 사용된 모든 영상은 동일한 크기 (150×130)로 정렬(face alignment)하였으며, 정렬 기준으로는 수동으로 찾은 두 눈의 좌표를 이용하였다^[18,24]. 특징 추출에 앞서 정렬된 모든 얼굴 영상들은 DoG(Difference of Gaussian) 기반의 전처리(preprocessing) 방법^[25]이 적용되었다. 모든 영상에 대해 영상 필터의 크기는 7×7를 사용하였으며, PDM의 이웃(neighbor) 영역은 3픽셀로 정의하였다.

2. 얼굴 인식률 평가

먼저 YaleB 데이터베이스에 대해, 기존 DFD(H_{DFD})의 결과와 단일 해상도로 분할(2×2, 4×4)한 결과

($H_{DFD}^{2 \times 2}, H_{DFD}^{4 \times 4}$)를 비교하였다. 그림 6의 결과를 보면, 부분 영역에서 DFD 히스토그램을 바로 추출하여 사용하는 기존 방법에 비해, 부분 영역을 2×2, 4×4의 영역으로 나누어 히스토그램을 추출한 경우 각각 인식률이 소폭 상승하는 것을 확인할 수 있다. 4×4로 분할했을 때 2×2로 분할했을 때보다 인식률이 1.47% 더 높게 나왔는데 이는 조명변이가 심한 YaleB 데이터베이스의 경우, 투영 그림자(cast shadow)로 인해 고주파 성분의 디테일한 왜곡이 많이 발생하였기 때문인 것으로 보인다. 이 결과를 바탕으로, 최종 비교 측정치를 위한 해상도별 통합 히스토그램 비교치의 가중치는 다중 분할 고 해상도 일수록 높은 가중치를 부여하였다($\alpha_{2 \times 2} = 0.18$, $\alpha_{3 \times 3} = 0.32$, $\alpha_{4 \times 4} = 0.5$).

AR 데이터베이스에 대해서도 YaleB 데이터베이스에 서와같이 2×2, 3×3, 4×4의 다중 해상도를 사용하였다. 그림 7은 AR 데이터베이스에 대한 H_{DFD} 와 $H_{DFD}^{4 \times 4}$, 그리고 다중 해상도 분할 영역 히스토그램 (H_{DFD}^{Multi})의 인식률을 보여준다. 조명변이의 세기가 약하고 부분 가림의 특성이 있는 AR 데이터베이스의 경우 H_{DFD}^{Multi} 의 인식률이 H_{DFD} 에 비해 scarf와 sunglass의 경우 각각 0.57%, 1.23%의 인식률 상승을 확인할 수 있었다.

CAS-PEAL-R1 데이터베이스는 lighting과 accessory로 구분되어, 각각 조명변이와 부분가림의 특성을 지니고 있다. 그림 8은 CAS-PEAL-R1의 lighting과 accessory의 실험 결과를 보여준다. accessory의 경우 H_{DFD}^{Multi} 의 인식률이 H_{DFD} 에 비해 0.57%~1.23% 정도의 인식률 상승을 확인할 수 있으며, 조명변이가 있는 lighting의 경우에는 3.53%~7.82% 정도 인식률이

상승한 것을 확인할 수 있다.

그림 5~7의 실험 결과로부터, 제안한 방법은 부분 가림과 조명변이 모두에서 인식 성능을 향상시키는데 특히 조명변이에 더 효과적인 것을 확인할 수 있다. 이는 부분 가림의 경우, 영상의 sunglass나 scarf, accessory 등으로 인해 가려진 부분의 크기가 거의 동일하기 때문에 다중 해상도 적용의 효과가 크지 않고, 가려진 부분에서는 원래 얼굴의 정보가 완전히 다른 물체의 텍스처로 대체되었기 때문에 어떤 방식의 기술자를 사용하더라도 가려진 부분에 대해서는 의미가 사라지기 때문이다. 그럼에도 H_{DFD}^{Multi} 의 인식률이 H_{DFD} 보다 소폭 상승한 것은 가려지지 않은 영역에 대해 다중 해상도 분할이 다양한 크기의 얼굴 구성 요소(눈, 코, 입 등)의 특징을 잘 활용할 수 있었기 때문으로 보인다. 반면, 조명변이의 경우는 얼굴의 구성 요소들이 다양한 크기의 그림자를 생성하기 때문에 다양한 크기의 영역에서 추출한 히스토그램을 함께 사용하는 것이 보다 효과적으로 지역 정보를 추출할 수 있으므로 H_{DFD}^{Multi} 사용으로 인한 성능 개선이 두드러지게 나타났다.

IV. 결 론

본 논문에서는 다양한 변이에 대한 얼굴 인식 성능을 개선하기 위해 DFD 기반의 다중 해상도 분할 영역을 통한 얼굴인식 방법을 제안하였다. 동일한 크기로 나누어진 부분 영역에서 추출한 히스토그램을 통합하여 두 영상 간의 유사도를 비교하는 기존 DFD와 달리, 본 논문에서는 각각의 부분 영역에 대해 다중 해상도로 분할하여 여러 종류의 크기에 해당하는 분할 영역의 히스토그램을 사용하였다. 이를 통해 지역 정보를 보다 효과적으로 사용하여 얼굴 인식 성능을 개선할 수 있었다. 실험 결과, 조명변이 및 부분 가림이 있는 YaleB, AR, CAS-PEAL-R1 데이터베이스에 대하여 기존의 DFD을 이용한 얼굴인식 방법보다 약 0.57% ~ 7.82%의 얼굴 인식률이 향상하였음을 확인하였다.

REFERENCES

[1] D. Lowe, "Distinctive Image Features From Scale-invariant Keypoints," *International Journal of Computer Vision*, Vol. 60, no. 2, pp. 91-110, 2004.
[2] K. Mikolajczyk and C. Schmid, "A Performance

Evaluation of Local Descriptors," *IEEE Trans., Pattern Analysis and Machine Intelligence*, Vol. 27, no. 10, pp. 1615-1630, 2005.
[3] H. Bay, A. Ess, T. Tuytelaars and L.V. Gool, "Speeded Up Robust Features (SURF)," *Computer Vision and Image Understanding*, Vol. 110, no. 3, pp. 346 - 359, 2008.
[4] Z. Ramin and J. Woodfill. "Non-parametric Local Transforms for Computing Visual Correspondence," *Lecture Notes in Computer Science (European Conf. Computer Vision)*, Vol. 801, pp. 618-629, 1994.
[5] B. Fröba and A. Ernst, "Face Detection with the Modified Census Transform," in *Proc. IEEE Conf. Automatic Face and Gesture Recognition*, pp. 91-96, May 2004.
[6] T. Ahonen, A. Hadid and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 28, no. 12, pp. 2037 - 2041, 2006.
[7] Z. Cao, Q. Yin, X. Tang and J. Sun, "Face Recognition with Learning-Based Descriptor," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 2707-2714, June 2010.
[8] D. Maturana, D. Mery and A. Soto, "Face Recognition with Decision Tree-Based Local Binary Patterns," *Lecture Notes in Computer Science (Asian Conf. Computer Vision)*, Vol. 6495, pp. 618-629, 2010.
[9] D.J. Kim, S.H. Lee and M.K. Sohn, "A Study on Face Recognition Method Based on Binary Pattern Image Under Varying Lighting Condition," *Journal of the Institute of Electronics and Information Engineers*, Vol. 49, no. 2, pp. 61-74, 2012.
[10] S.J. Lee, D.H. Kim, Suryanto and S.J. Ko, "Improved Color-LBP Joint Histogram for Robust Object Tracking," *Journal of the Institute of Electronics and Information Engineers*, Vol. 47, no. 11, pp. 604-607, 2011.
[11] T. Ojala, M. Pietikänen and T. Mänpää, "Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, no. 7, pp. 971-987, 2002.
[12] M. Heikkila, M. Pietikainen and C. Schmid, "Description of Interest Regions with Local Binary Patterns," *Pattern Recognition*, Vol. 42, no. 3, pp. 425-436, 2009.
[13] M. Calonder, V. Lepetit, C. Strecha and P. Fua,

- “BRIEF: Binary Robust Independent Elementary Features,” Lecture Notes in Computer Science (European Conf. Computer Vision), Vol. 6314, pp. 778–792, 2010.
- [14] M. Daniel, Domingo Mery, and Alvaro Soto, “Learning Discriminative Local Binary Patterns for Face Recognition,” in Proc. IEEE Conf. Automatic Face and Gesture Recognition and Workshops, pp. 470–475, March 2011.
- [15] P.N. Belhumeur, J.P. Hespanha and D.J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, no. 7, pp. 711–720, 1997.
- [16] L. Zhen, M. Pietikainen and S. Li, “Learning Discriminant Face Descriptor,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 36, no. 2, pp. 289–302, 2014.
- [17] Y. Rubner, C. Tomasi and L.J. Guibas. “The Earth Mover’s Distance as a Metric for Image Retrieval,” International Journal of Computer Vision, Vol. 40, no. 2, pp. 99–121, 2000.
- [18] A.S. Georghiades, S. Athinodoros, P.N. Belhumeur and J. Kriegman, “From Few to Many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 23, no. 6, pp. 643–660, 2001.
- [19] A.M. Martinez and R. Benavente, The AR Face Database, CVC Technical Report #24, June 1998.
- [20] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang and D. Zhao, “The CAS-PEAL Large-scale Chinese Face Database and Baseline Evaluation,” IEEE Trans. Systems, Man, and Cybernetics, Part A : Systems and Humans, Vol. 38, no. 1, pp. 149–161, January 2008.
- [21] Z. Lei and S.Z. Li, “Learning Discriminant Face Descriptor for Face Recognition,” in Proc. Asian Conf. Computer Vision, pp. 748–759, November 2012.
- [22] L. Zhen, L. Shengcai, K. Jain and S.Z. Li, “Coupled Discriminant Analysis for Heterogeneous Face Recognition,” IEEE Trans. Information Forensics and Security, Vol. 7, no. 6, pp. 1707–1716, 2012.
- [23] J. MacQueen, “Some Methods for Classification and Analysis of Multivariate Observations,” in Proc. Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 281–297, 1967.
- [24] S.H. Kim, S.T. Chung, S.H. Jung and S.W. Jo, “Robust Eye Localization Using Multi-Scale Gabor Feature Vectors,” Journal of the Institute of Electronics and Information Engineers, Vol. 45, no. 1, pp. 25–36, 2008.
- [25] X. Tan and B. Triggs, “Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions,” IEEE Trans. Image Processing, Vol. 19, no. 6, pp. 1635–1650, 2010.

저 자 소 개



이 장 운(학생회원)
2013년 단국대학교 응용컴퓨터 공
학과 공학사
2015년 단국대학교 컴퓨터학과 공
학석사
<주관심분야 : 패턴인식, 머신러닝>



이 용 결(학생회원)
2012년 단국대학교 응용컴퓨터공
학과 공학사
2014년 단국대학교 컴퓨터학과 공
학석사
2015년~현재 단국대학교 컴퓨터
학과 박사 과정
<주관심분야 : 패턴인식, 컴퓨터비전>



최 상 일(정회원)
2005년 서강대학교 전자공학과
공학사
2010년 서울대학교 전기·컴퓨터
공학부 공학박사
2010년~2011년 Post Doctoral
Fellow, BK21 Information
Technology, Seoul National
University, Seoul, Korea
2011년 Post Doctoral Fellow, Computer Science
Department, University of Southern
California, CA, USA
2011년~현재 단국대학교 대학원 컴퓨터학과 조
교수
<주관심분야 : 패턴인식, 컴퓨터비전>