

# 이질적 얼굴인식을 위한 심층 정준상관분석을 이용한 지역적 얼굴 특징 학습 방법

최여름<sup>†</sup>, 김형일<sup>\*\*</sup>, 노용만<sup>\*\*\*</sup>

## Local Feature Learning using Deep Canonical Correlation Analysis for Heterogeneous Face Recognition

Yeoreum Choi<sup>†</sup>, Hyung-Il Kim<sup>\*\*</sup>, Yong Man Ro<sup>\*\*\*</sup>

### ABSTRACT

Face recognition has received a great deal of attention for the wide range of applications in real-world scenario. In this scenario, mismatches (so called heterogeneity) in terms of resolution and illumination between gallery and test face images are inevitable due to the different capturing conditions. In order to deal with the mismatch problem, we propose a local feature learning method using deep canonical correlation analysis (DCCA) for heterogeneous face recognition. By the DCCA, we can effectively reduce the mismatch between the gallery and the test face images. Furthermore, the proposed local feature learned by the DCCA is able to enhance the discriminative power by using facial local structure information. Through the experiments on two different scenarios (i.e., matching near-infrared to visible face images and matching low-resolution to high-resolution face images), we could validate the effectiveness of the proposed method in terms of recognition accuracy using publicly available databases.

**Key words:** Heterogeneous Face Recognition, Deep Learning, Deep Canonical Correlation Analysis, Local Feature Learning

### 1. 서 론

최근 얼굴인식 기술이 실제 환경의 다양한 응용분야(예: 감시시스템, 범죄수사 등)에 채택됨에 따라 실용적 이슈에 관한 활발한 논의가 진행되고 있다[1], [2]. 보편적인 얼굴인식 방법에서는 같은 환경에서 촬영된 얼굴 영상에 대해 훈련집합과 검증집합을 분할하여 얼굴인식 성능을 검증하였다[3]. 하지만, 실제 환경에서는 시스템에 등록된 얼굴 영상과 실제

촬영되는 얼굴 영상 사이에 차이(예: 해상도 및 조명 조건의 차이)가 존재하게 되고, 이러한 불일치(mismatch)는 얼굴인식 성능을 매우 악화시킨다고 알려져 있다[6]. 이로 인해 서로 다른 양식(modality) 으로부터 얻은 얼굴 영상 간의 인식 방법인 이질적(heterogeneous) 얼굴인식 기술이 주목받고 있다[3-5]. 이 때, 이질적 얼굴인식에서는 갤러리(gallery) 얼굴 영상은 일반 카메라로 촬영된 고해상도의 얼굴 영상(visible (VIS) 영상)을 사용하게 되고, 테스트

\* Corresponding Author: Yong Man Ro, Address: (34141) KAIST 291, Daehak-ro, Yuseong-gu, Daejeon, Korea, TEL : +82-42-350-3494, FAX : +82-42-350-5494, E-mail : ymro@kaist.ac.kr

Receipt date : Feb. 1, 2016, Revision date : Mar. 2, 2016  
Approval date : Apr. 4, 2016

<sup>†</sup> School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST)  
(E-mail : cyr0703@kaist.ac.kr)

<sup>\*\*</sup> School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST)  
(E-mail : hyungil.kim@kaist.ac.kr)

<sup>\*\*\*</sup> School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST)

\* This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No.2015R1A2A2A01005724).

(probe) 얼굴 영상은 다른 양식으로부터 얻은 얼굴 영상(예: 감시시스템에 적합한 근적외선 영상 또는 범접 수사를 위한 스케치 영상)을 사용하게 된다.

예를 들면, 조도 변화가 발생하는 실외 CCTV (closed-circuit television) 촬영 환경 또는 조도가 매우 낮은 환경인 감시시스템 환경에서는 조도 변화에 강인하기 위해 근적외선(near-infrared (NIR)) 센서에 의해 얼굴 영상이 촬영되게 된다. 이 때, 기존에 등록된 VIS 얼굴 영상과 NIR 얼굴 영상 사이의 매칭을 통한 이질적 얼굴인식 기술이 요구된다[3]. 또한 대부분의 감시 카메라 환경에서는 저해상도의 얼굴 영상만을 가지므로 이러한 저해상도 얼굴 영상과 기존에 등록된 고해상도 얼굴 영상간의 매칭이 필요하다[5]. 이처럼 실제 환경에서의 이질적 얼굴인식은 다방면으로 유용하게 활용될 수 있다. 그러나 서로 다른 양식으로부터 얻은 얼굴 영상은 서로간의 큰 불일치로 인해 매우 도전적인 문제로 알려져 있다 [6].

Fig. 1은 서로 다른 양식으로부터 얻은 얼굴 영상의 화소 값들을 t-SNE[7]를 이용하여 2차원 공간에 시각화한 것으로, 실제로 매우 큰 불일치가 존재함을 확인할 수 있다. [8]에서는 이러한 불일치를 단순히 선형 부공간 학습(linear subspace learning)방법에 의해 해결하려고 하였지만 두 영상 사이의 불일치는 간단한 선형 모델로 해결될 수 없었다. 최근에는 다양한 커널 함수와 handcrafted 특징에 기반한 비선형 모델 학습방법[4]이 제안되고 있으나, 제한적인 비선

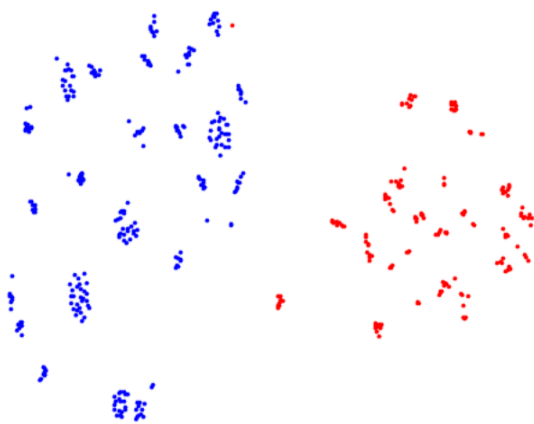


Fig. 1. Visualization of sample distribution from different modalities in feature space. Each dot represents a sample (Red : VIS, Blue: NIR).

형성과 분별력 문제는 여전히 해결하기 어려운 것으로 알려져 있다.

본 논문에서는 이처럼 서로 다른 양식으로부터 얻은 얼굴 영상 사이에 존재하는 매우 복잡한 비선형성 문제를 다루기 위해 심층 학습(deep learning) 기법 중 하나인 심층 정준상관분석(deep canonical correlation analysis, DCCA) 방법[9]을 채택한다. 이 때, DCCA 모델은 심층 신경망(deep neural network, DNN)에 의해 추출된 두 양식으로부터 얻은 얼굴 특징 사이의 상관도를 최대가 되도록 DNN을 학습하게 된다. 이러한 학습을 통해 서로 다른 양식으로부터 얻은 얼굴 영상 사이의 큰 차이를 줄일 수가 있게 된다. 추가적으로, 얼굴의 구조정보를 활용하고 DCCA의 분별력을 높이기 위한 지역적 얼굴 특징 학습 방법을 제안한다. 서로 다른 양식의 정보는 지역적인 부분에서 더 작은 차이를 보이게 된다[10]. 또한 지역적 얼굴 특징 학습으로 인해 같은 얼굴 영상에 대해 더 많은 정보를 얻을 수 있게 된다.

본 논문의 구성은 2절에서 제안 방법을 자세히 설명하고, 3절에서 실험 결과를 보이며, 마지막으로 4절에서 결론을 맺는다.

## 2. 제안하는 방법

본 논문에서는 서로 다른 양식으로부터 얻은 두 얼굴 영상 사이의 매우 복잡한 비선형 차이를 최소화하기 위해 DCCA 방법 기반 특징 학습 방법을 제안한다. 추가적으로, 얼굴의 구조정보를 활용한 분별력 있는 DCCA 특징을 학습하기 위한 지역적 특징 학습 방법이 제안한다. 2.1절에서는 전역적 얼굴 영상에 대한 심층 정준상관분석(DCCA) 기반 특징 학습 방법에 대해 설명하고, 2.2절에서는 지역적 특징 추출을 위한 지역 DCCA 구조를 설명한다. 마지막으로 2.3절에서는 DCCA 구조를 통해 얻은 특징 벡터를 이용해 분류하는 방법에 대해 설명한다.

### 2.1 DCCA 기반 전역적 이질적 얼굴 영상 특징 학습

서로 다른 양식으로부터 얻은 얼굴 영상에 대한 차이를 최소화하는 특징을 학습하기 위해서, 먼저 Fig. 2와 같이 각각의 양식으로부터 얻은 얼굴 영상에 대한 두 개의 DNN 구조를 도입한다. 이 때, 각 DNN의 출력들 사이의 상관도가 최대가 되도록

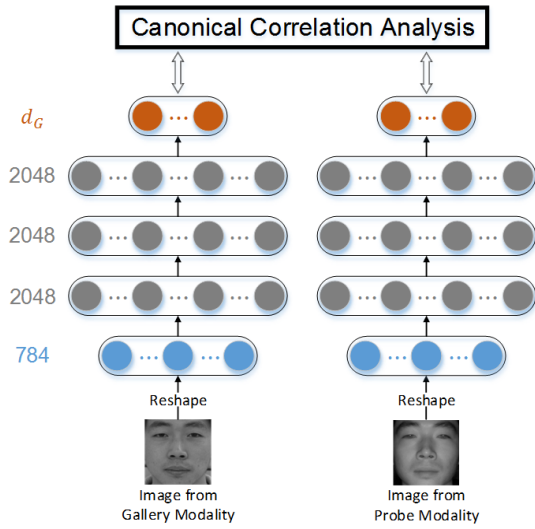


Fig. 2. Global DCCA structure for feature learning that maximizes correlation between two face images from different modalities (Left: DNN for an image from gallery modality, Right: DNN for an image from probe modality).

DCCA 목표 함수에 의해 DNN 구조를 학습하게 되면 두 얼굴 영상 사이의 특징간의 상관도가 높아지게 되고, 결국 얼굴 영상 사이의 차이가 상쇄되게 된다. [9]에서는 DCCA 목적 함수를 다음과 같이 정의하고 있다.

$$\begin{aligned} & \max_{P, Q} \frac{1}{N} \text{trace}(\mathbf{P}^T \mathbf{f}_1(\mathbf{X}_1) \mathbf{f}_2(\mathbf{X}_2)^T \mathbf{Q}) \\ & \text{s.t. } \mathbf{P}^T \left( \frac{1}{N} \mathbf{f}_1(\mathbf{X}_1) \mathbf{f}_1(\mathbf{X}_1)^T + r_x \mathbf{I} \right) \mathbf{P} = \mathbf{I} \\ & \mathbf{Q}^T \left( \frac{1}{N} \mathbf{f}_2(\mathbf{X}_2) \mathbf{f}_2(\mathbf{X}_2)^T + r_y \mathbf{I} \right) \mathbf{Q} = \mathbf{I} \\ & \mathbf{p}_i^T \mathbf{f}_1(\mathbf{X}_1) \mathbf{f}_2(\mathbf{X}_2)^T \mathbf{q}_j = 0 \text{ for } i \neq j, \end{aligned} \quad (1)$$

여기에서,  $\mathbf{X}_1$  과  $\mathbf{X}_2$  는 각각의 양식으로부터 얻은 얼굴 영상을 벡터화한 것이며, 함수  $\mathbf{f}_1$  과  $\mathbf{f}_2$  는 이렇게 벡터화 된 입력 얼굴 영상을 DNN의 출력인  $d_G$  차원의 특징 벡터로 맵핑시키게 된다. 그리고,  $\mathbf{P}$  와  $\mathbf{Q}$  는 DNN의 최종 출력 사이의 상관도를 최대화 하는 공간으로 투영시키는 행렬이며,  $N$  은 각 얼굴 영상의 개수이다. 또한  $r_x, r_y > 0$  은 입력 정보들의 공분산을 추정하기 위한 조정 매개변수이다. 결국, 식(1)을 만족하는 행렬  $\mathbf{P}$  와  $\mathbf{Q}$  를 찾아 첫 번째 양식으로부터 얻은 얼굴 영상에 대해서는  $\mathbf{P}$  에 의해 변환되고, 두 번째 양식으로부터 얻은 얼굴 영상에 대해서는  $\mathbf{Q}$  에

의해 변환함으로써 두 얼굴 영상 사이의 차이가 줄어드는 특징 추출이 가능하게 된다. 마지막으로, 높은 분별력을 제공하기 위해 각 얼굴 영상의 라벨 정보에 기반한 fine tuning을 back propagation 방법에 의해 수행하게 된다.

### 2.2 DCCA 기반 지역적 이질적 얼굴 영상 특징 학습

2.1절에서 얻은 특징 정보가 낮은 차원의 특징 벡터일 경우에는 분별력있는 정보를 제공하기에는 한계가 존재하게 된다. 하지만, 전역적 얼굴에 대해 매우 높은 차원의 상관관계를 갖는 특징을 학습하게 되면 DCCA 방법이 수렴이 매우 오래 걸리거나 수렴에 실패하게 된다.

따라서, 서로 다른 양식으로부터 얻은 얼굴 영상 사이에 대해 분별력있는 정보를 추가적으로 얻기 위해서 DCCA 기반 지역적 특징 학습 방법이 제안한다. 특히, 지역적으로 얼굴 영상을 보게 되면, 전역적으로 볼 때보다 상대적으로 서로 다른 양식간의 차이가 덜 복잡하다고 알려져 있다는 것[10]에 착안하여, 본 논문에서는 Fig. 3에서와 같이 얼굴을  $m$  개의 지역으로 나누어서 각 영역에 대한 DCCA 구조를 학습하게 된다. 지역적 DCCA 구조의 출력으로 나오는 특징 벡터는 각각  $d_L$  차원을 가지게 된다. 이 때, 전역적 DCCA 구조와 지역적 DCCA 구조의 균형을 맞추기 위해  $d_G$  와  $m \times d_L$  의 차원을 같게 하도록 특징을 추출하게 된다.

### 2.3 분류

분류단계에서는 첫 번째 양식이 갤러리 얼굴 영상으로 사용되고 두 번째 양식이 테스트 얼굴 영상으로 사용되었다. 검증집합에서 테스트 얼굴 영상에 해당하는 두 번째 양식으로부터 얻은 얼굴 영상을 벡터화하여  $\mathbf{X}_{Tst}$  을 얻는다. 이 얼굴 영상을 2.1절로부터 학습된 DCCA 구조에 입력하게 되면  $d_G$  차원의 특징 벡터가  $\mathbf{f}_2(\mathbf{X}_{Tst})$  로 나오게 된다. 이와 같은 방법으로 2.2절에서 학습된 DCCA 구조에도 적용하게 된다. 얼굴 영상  $\mathbf{X}_{Tst}$  에 대해서  $m$  개의 지역으로 나누게 되면,  $\mathbf{X}_{Tst}^{R_1}, \dots, \mathbf{X}_{Tst}^{R_m}$  의 지역적 얼굴 영상을 얻게 된다. 이를 DCCA 구조에 입력하면  $\mathbf{f}_2^1(\mathbf{X}_{Tst}^{R_1}), \dots, \mathbf{f}_2^m(\mathbf{X}_{Tst}^{R_m})$  를 얻게 되며, 하나의 특징 벡터는  $d_L$  의 차원을 가지게 된다. 최종적으로 전역적 특징 벡터와 지역적 특징 벡

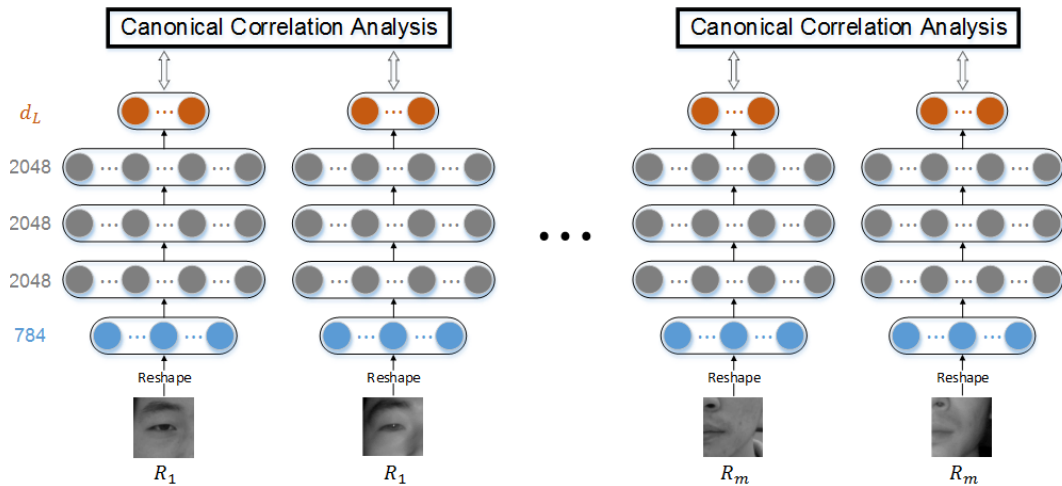


Fig. 3. Local DCCA structure for feature learning.

터를 결합하여  $d_G + m \times d_L$  차원의 특징 벡터

$\mathbf{P}_{Tst} = [\mathbf{f}_2^G(\mathbf{X}_{Tst}) \ \mathbf{f}_2^L(\mathbf{X}_{Tst}^{R_1}) \ \dots \ \mathbf{f}_2^L(\mathbf{X}_{Tst}^{R_m})]$ 를 얻게 된다.

얼굴인식을 위해서 갤러리 얼굴 영상 벡터  $\mathbf{X}_{Gal}$ 에 해당하는 특징 벡터를

$\mathbf{P}_{Gal} = [\mathbf{f}_1^G(\mathbf{X}_{Gal}) \ \mathbf{f}_1^L(\mathbf{X}_{Gal}^{R_1}) \ \dots \ \mathbf{f}_1^L(\mathbf{X}_{Gal}^{R_m})]$ 라 할 때, 전체 부류  $C$ 개의 갤러리 얼굴 영상 벡터  $\mathbf{X}_{Gal}^1, \dots, \mathbf{X}_{Gal}^C$ 에 해당하는 특징 벡터  $\mathbf{P}_{Gal}^1, \dots, \mathbf{P}_{Gal}^C$ 를 얻는다. 분류 단계에서는 유클리디언 거리(Euclidean distance)에 기반한 1-nearest neighbor 분류기를 사용하였다.

### 3. 실험 결과 및 분석

#### 3.1 DCCA 구조 및 실험 환경

DCCA 구조는 충분히 많은 집단(batch)을 통해 기울기를 구하는 최적화[9]를 통해 학습되었고, Fig. 2에서와 같이 DNN은 784개의 노드를 갖는 입력층, 2,048개 노드를 갖는 은닉층(hidden layer), 그리고 40개의 노드를 갖는 출력층으로 구성되어 있다. 지역적 특징 추출을 위한 DCCA 구조는 Fig. 3과 같이  $m$ 은 4로 총 4개의 지역으로 나누었다. 이보다 더 많은 지역으로 얼굴을 나누게 되면 중복된 정보를 가지는 큰 차원의 특징 벡터로 인해 오히려 성능에 저하가 있을 수 있다[11]. 이렇게 나누어진 4개의 지역에 대하여 각 DCCA 구조의 출력층은 10개의 노드를 갖는다. 결과적으로 최종 특징 벡터  $\mathbf{P}$ 는 80 차원을 가진다. 이 때, DCCA 학습을 위해 훈련 집합에서

같은 부류의 서로 다른 양식으로부터 얻은 얼굴 영상을 모든 경우에 대해 쌍을 이루어 구성했다.

#### 3.2 VIS-NIR 얼굴인식

VIS 얼굴 영상과 NIR 얼굴 영상간의 성능을 평가하기 위해 CASIA HFB 얼굴 데이터베이스[12]를 사용하였다. 이 데이터베이스는 202명의 부류로 구성되어 있고, 2,095장의 VIS 얼굴 영상 및 3,002장의 NIR 얼굴 영상을 포함하고 있다. 이 때, 검증에 위하여 [12]에서 제안하는 검증방법 1과 [3]에서 제안하는 검증방법 2의 두 가지 검증방법을 사용하였다. 첫 번째 검증방법은 훈련집합에 존재하는 인물이 검증에 사용되는 인물과 중복이 되며(closed-set), 두 번째 검증방법은 훈련집합에 존재하는 인물과 검증에 사용되는 인물 사이의 중복이 없다(open-set). 특히, CASIA HFB 데이터베이스는 VIS 얼굴 영상과 NIR 얼굴 영상이 서로 다른 환경에서 촬영된 점과 표정변화, 안경의 착용 등 때문에 매우 도전적인 데이터베이스로 알려져 있다[12].

Table 1에서는 [8]에서 제시하는 방법의 결과와 제안하는 방법에 의한 인식 결과를 다양하게 결합하여 보여주고 있다. 검증방법 2의 경우에는 테스트 얼굴 영상이 DCCA 구조를 학습할 시에 사용되지 않은 인물이므로 검증방법 1에 비해 상대적으로 낮은 성능을 보인다. 또한 검증방법 1의 경우에는 전역적 특징만 사용한 결과가 지역적 특징만 사용한 결과보다 높았는데, 이는 테스트 얼굴 영상이 이미 훈련집합에

Table 1. Face recognition rate for global and local DCCA on CASIA HFB database

	검증 방법 1	검증 방법 2
HCA[8]	3.75%	8.11%
전역적 DCCA( $d_G = 40$ )	92.00%	23.74%
지역적 DCCA( $d_L = 10, m = 4$ )	89.25%	30.22%
전역적 DCCA + 지역적 DCCA ( $d_G = 40, d_L = 10, m = 4$ )	99.75%	39.11%

존재하였던 인물이므로 얼굴 전체에 대한 정보를 사용하는 것이 더 효과적이기 때문이다. 이에 비해 검증방법 2의 경우에는 오히려 지역적 특징만 사용하였을 때 더 높은 결과를 보인다. 이는 테스트 얼굴 영상이 훈련집합에 존재하지 않았던 인물이므로 특징을 얻음에 있어서 특정 인물이 아닌 일반적인 얼굴 특징을 필요로 하게 되며, 이는 지역적인 정보를 통해 얻을 수 있기 때문이다. 결과적으로 전역적 특징만 사용하거나 지역적 특징만 사용하더라도 [8]에서 제시하는 방법에 비해 더 높은 성능을 보이는 것을 확인할 수 있으며, 두 특징을 결합하여 사용할 경우에는 그보다 더 높은 성능을 보인다.

### 3.3 고해상도와 저해상도 얼굴인식

고해상도 얼굴 영상과 저해상도 얼굴 영상간의 성능을 평가하기 위해 CMU Multi-PIE 데이터베이스 [13]를 사용하였다. 이 데이터베이스는 총 4개의 세션으로 337명의 부류로 구성되어 있는데, 이 중 실험에서는 조명 변화와 표정 변화를 모두 포함하는 정면 얼굴만을 사용하였다. 고해상도 얼굴 영상은 [14]에서 제안하는 얼굴의 특징점 추출 방법을 통해  $28 \times 28$  크기로 잘라서 사용되었으며, 저해상도 얼굴 영상은 고해상도 얼굴 영상을 저해상도로 먼저 축소한 뒤에 다시  $28 \times 28$  크기로 확대하여 사용하였다. 이때 사용된 저해상도 얼굴 영상의 크기는  $14 \times 14$ 와  $7 \times 7$ 이다. 본 실험에서는 훈련집합에 존재하는 인물과 검증에 사용되는 인물 사이의 중복이 존재하지 않는 검증방법만을 사용하였으며 처음 200명의 부류는 훈련집합, 나머지 137명의 부류는 검증집합으로 사용되었다. 훈련집합은 4,000장의 얼굴 영상을 포함하고 있으며, 갤러리 얼굴 영상과 테스트 얼굴 영상은 각각 137장, 2,603장을 포함한다.

Table 2. Face recognition rate for global and local DCCA on Multi-PIE database

	$14 \times 14$	$7 \times 7$
HCA[8]	62.66%	17.02%
전역적 DCCA( $d_G = 40$ )	64.89%	52.29%
지역적 DCCA( $d_L = 10, m = 4$ )	51.56%	49.56%
전역적 DCCA + 지역적 DCCA ( $d_G = 40, d_L = 10, m = 4$ )	66.50%	57.32%

Table 2에서는 제안하는 방법에 의한 인식 결과를 보이고 있다. 저해상도 얼굴 영상의 크기가  $14 \times 14$ 인 경우에, 전역적 DCCA는 [8]에서 제안하는 방법보다 더 높은 성능을 보이지만 지역적 DCCA는 보다 낮은 성능을 보인다. 그러나 두 특징을 결합하여 사용할 때에는 전역적 DCCA보다도 높은 성능을 보이는 것을 확인할 수 있다.  $7 \times 7$ 의 경우에는 전역적 특징만 사용하거나 지역적 특징만 사용하더라도 [8]에서 제시하는 방법보다 더 높은 성능을 보이며, 두 특징을 결합하여 사용할 때에는 더 높은 성능을 보이는 것을 확인할 수 있다. 이처럼 제안하는 방법은 더 낮은 저해상도 얼굴 영상에 대해서도 급격한 성능 저하 없이 높은 성능을 보이는 것을 알 수 있다.

### 3.4 특징 공간 분석

본 논문에서 제안하는 방법이 얼마나 효과적으로 서로 다른 양식 사이의 겹을 줄이는지 확인해보기 위해 DCCA 구조를 통과하기 전의 얼굴 영상과 DCCA 구조를 통과한 후의 특징 벡터를 t-SNE[7]를 통해 2차원 공간으로 투영하여 비교한다.

#### 3.4.1 VIS-NIR 특징 공간

Fig. 4는 검증방법 1에 대한 VIS 얼굴 영상과 NIR 얼굴 영상을 특징 공간에 투영한 결과로, 각각의 점은 특징 벡터 하나를 의미한다. 특징 벡터는 총 30개의 부류가 서로 다른 색으로 구분되어 있다. Fig. 4(a)에서 볼 수 있듯이 DCCA 구조를 통과하기 전에는 각각의 부류가 서로 구분되어 있지 않고 서로 겹쳐져 있거나 멀리 퍼져있는 것을 확인할 수 있다. 결과적으로 각각의 사람간의 구별이 불분명하게 된다. 그러나 Fig. 4(b)에서 DCCA 구조를 통과한 뒤의 특징 벡터를 보게 되면, 같은 부류별로 모임으로써 서로 간의 구별이 더 확실하게 됨을 볼 수 있다.

3.4.2 고해상도와 저해상도 특징 공간

Fig. 4(c), (d)는 고해상도 얼굴 영상과  $14 \times 14$  크기의 저해상도 얼굴 영상을 특징 공간에 투영한 결과이다. Fig. 4(c)는 DCCA 구조를 통과하기 전의 결과로 각각의 부류가 서로 뒤엉켜 있어 서로간의 구별이 힘든 것을 확인할 수 있다. 이에 비해 Fig. 4.(d)에서는 학습된 DCCA 구조를 통과한 특징 벡터들의 분별력이 크게 향상되어 같은 부류별로 모이는 현상을 확인할 수 있다. 그러나 그림 상에서 중앙 부근에는 같은 부류로부터 떨어져 있는 특징 벡터들을 볼 수 있다. 이는 해상도 차이로 인한 정보 손실로 인해 몇몇 저해상도 얼굴 영상에 대한 특징 벡터가 같은 부류로부터 떨어져 있는 것으로, 이로 인해 성능의 저하가 생기게 되었다.

4. 결 론

본 논문에서는 heterogeneous 얼굴인식을 위해 서로 다른 양식으로부터 얻은 얼굴 영상 사이의 차이를 줄이는 DCCA 구조를 학습하였다. 이 때, 추가적으로 분별력있는 특징을 학습하기 위해서 지역적 얼굴 영상에 대한 DCCA 구조를 학습하였으며, 최종적으로 전역적 그리고 지역적 DCCA 구조에 의한 특징이 얼굴인식을 위해 사용되었다. 실험을 통해서 VIS 얼굴 영상과 NIR 얼굴 영상 사이의 차이와 고해상도 얼굴 영상과 저해상도 얼굴 영상 사이의 차이가 특징 공간에서 봤을 때, 매우 감소되는 것을 확인할 수 있었다. 뿐만 아니라 인식 결과에 있어서도 제안하는 방법이 이질적 얼굴인식에 효과적임을 확인할 수 있었다.

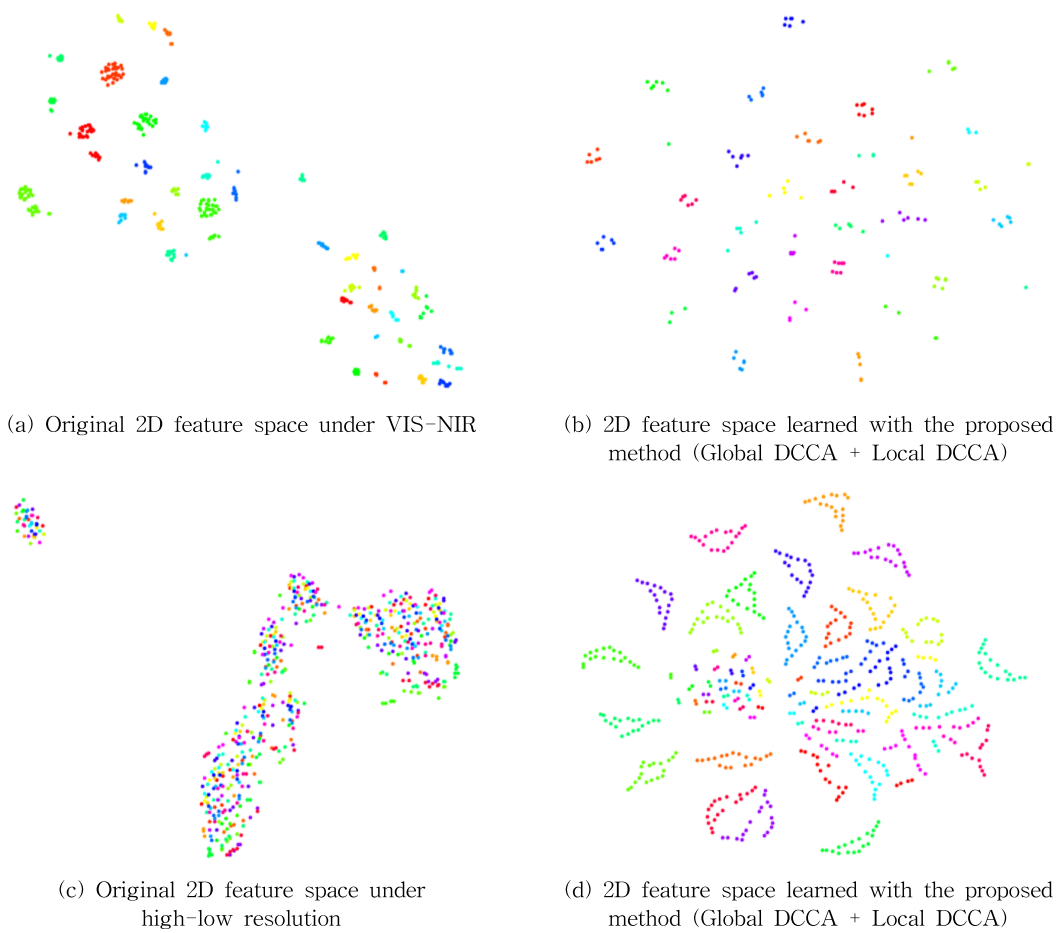


Fig. 4. Visualization of 2D feature spaces. Each dot represents a feature from 30 different classes.

## REFERENCE

- [1] Y.H. Kim and J.H. Kim, "Development of Real-Time Face Region Recognition System for City-Security CCTV," *Journal of Korea Multimedia Society*, Vol. 13, No. 4, pp. 504-511, 2010.
- [2] J. Pillai, V.M. Patel, R. Chellappa, and N.K. Ratha, "Towards a Practical Face Recognition System: Robust Registration and Illumination by Sparse Representation," *Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1838-1841, 2010.
- [3] B. Klare, and A.K. Jain, "Heterogeneous Face Recognition: Matching NIR to Visible Light Images," *Proceeding of International Conference on Pattern Recognition*, pp. 1513-1516, 2010.
- [4] B. Klare and A.K. Jain, "Heterogeneous Face Recognition Using Kernel Prototype Similarities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 6, pp. 1410-1422, 2013.
- [5] Z. Lei, S. Liao, A.K. Jain, and S.Z. Li, "Coupled Discriminant Analysis for Heterogeneous Face Recognition," *IEEE Transactions on Information Forensics and Security*, Vol. 7, No. 6, pp. 1707-1716, 2012.
- [6] P.J. Phillips, W.T. Scruggs, A.J. O'Toole, P.J. Flynn, K.W. Bowyer, C.L. Schott, *et al.*, *FRVT 2006 and ICE 2006 Large-scale Results*, Technical Report NISTIR 7408, NIST, 2007.
- [7] L.V.D. Maaten and G. Hinton, "Visualizing Data Using t-SNE," *Journal of Machine Learning Research*, Vol. 9, No. 85, pp. 2579-2605, 2008.
- [8] S.Z. Li, D. Yi, Z. Lei, and S. Liao, "The Casia NIR-VIS 2.0 Face Database," *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 348-353, 2013.
- [9] W. Wang, R. Arora, K. Livescu, and J. Bilmes, "On Deep Multi-view Representation Learning," *Proceeding of International Conference on Machine Learning*, pp. 1083-1092, 2015.
- [10] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of Dimensionality: High-dimensional Feature and Its Efficient Compression for Face Verification," *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3025-3032, 2013.
- [11] D. Donoho, *High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality*, AMS Math Challenges Lecture, 2000.
- [12] S.Z. Li, Z. Leii, and M. Ao, "The HFB Face Database for Heterogeneous Face Biometrics Research," *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1-8, 2009.
- [13] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie," *Journal of Image and Vision Computing*, Vol. 28, No. 5, pp. 807-813, 2010.
- [14] Z. Zhu, P. Luo, X. Wang, and X. Tang, "Deep Learning Identity-preserving Face Space," *Proceeding of IEEE International Conference on Computer Vision*, pp. 113-120, 2013.



최 여 름

2015년 KAIST 전기 및 전자공학과 공학사  
현재 KAIST 전기 및 전자공학부 석사과정  
관심분야: Pattern recognition and machine learning



노 용 만

1992년 KAIST 전기 및 전자공학과 공학박사  
1987년 Columbia University 연구원  
1996년 University of California, Berkeley 연구원



김 형 일

2013년 KAIST 전기 및 전자공학과 공학석사  
현재 KAIST 전기 및 전자공학부 박사과정  
관심분야: Image and video processing, Pattern recognition and machine learning, Face recognition

2006년 University of Toronto 초빙교수  
현재 KAIST 전기 및 전자공학과 정교수  
관심분야: Image and video processing, 3D video processing, Image and video feature extraction/recognition, Object detection/recognition, and medical image processing