

딥러닝을 이용한 영상기반 3차원 얼굴 복원 기법의 기술 동향

I. 서론

주어진 영상으로부터의 3차원 얼굴 복원은 얼굴 분석, 애니메이션 및 인식 등의 많은 컴퓨터비전 및 그래픽스 응용분야에서 중요한 역할을 하고 있으며, 특히 증강현실과 가상현실 분야의 핵심 기술로 촉망받고 있다. Blanz 와 Vetter가 3차원 얼굴 복원을 위해 3D morphable model(3DMM)^[1]의 사용을 제안한 이후로 3DMM을 기반으로 한 3차원 얼굴 복원은 가장 널리 쓰이는 방식이 되었다. 그로부터 지난 이십여 년 간의 연구는 3D 얼굴 모델의 평균 형상과 입력 영상간의 대응점을 찾고 비선형 최적화 함수를 통해 3DMM의 형상 및 표정 계수를 회귀 분석하는 방식이 대부분이었다. 그러나 이러한 방법들은 랜드마크의 정확도에 크게 의존하므로 다양한 조명, 자세 및 가려짐과 같은 극한의 조건하의 영상에 대한 3차원 얼굴 복원은 여전히 도전 과제로 남게 되었다.

CNN 기반의 딥러닝 방식이 물체 추적, 영상 분할 등의 여러 컴퓨터비전 분야에서 성공적으로 적용되어 온 것과 같이, 3차원 얼굴 복원 분야에서도 최근 CNN을 통해 주어진 입력영상으로부터 3차원 얼굴을 복원하는 방식이 개발되었다^[2-6]. CNN 기반의 방식은 여러 측면에서 3차원 얼굴 복원 능력을 향상시킨다. [35, 36]에서는 영상으로부터 보다 강인하게 랜드마크를 탐색하는 방법을 학습하고, 3DMM 계수를 계산한다. [7, 8, 9, 10, 11]은 단 하나의 네트워크를 통해 얼굴 랜드마크나 드러난 특징들을 사용하지 않고도 직접적으로 영상에서 3DMM의 형상 및 표정 계수를 학습하도록 하였다. 최근에 [12, 13]에서는 CNN을 통해 영상의 화소들을 3D 얼굴 구조에 직접 매핑 하는 방식을 제안하였다.

본 기고문은 딥러닝을 통한 3차원 얼굴 복원의 최근 동향을 소개한다. 기고문의 구성은 다음과 같다. 2장에서는 3DMM을 이용한 전통적



원 소 미
인하대학교 정보통신공학과



이 정 우
인하대학교 정보통신공학과



박 인 규
인하대학교 정보통신공학과



인 얼굴 형상 복원 기법을 소개하고 장단점을 살펴본다. 3장에서는 딥러닝을 이용한 최신 3차원 얼굴 복원 기법을 소개한다. 4장에서는 향후 연구 전망을 제시한다.

II. 3차원 얼굴 형상 복원을 위한 전통적 기법

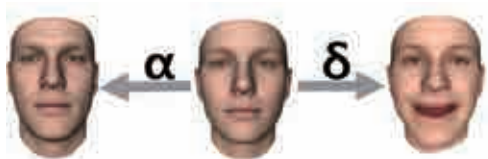
딥러닝 기법이 활발히 이용되기 이전의 전통적인 3차원 얼굴 복원 기법은 크게 3D morphable model(3DMM) 기반^[1,14-15,17-19]과 3D 메쉬 템플릿 기반^[21,27]의 두 가지 방식으로 나눌 수 있다.

1. 3D Morphable Model (3DMM) 기반 방식

Vetter와 Blantz는 3차원 얼굴을 표현하기 위해 주성분 분석(principal components analysis, PCA) 기반의 3D morphable face model (3DMM)을 고안했다. 이 방식에서는 개개인의 얼굴은 강한 형태의 유사성을 지니므로, 낮은 차원의 공간에서 사람의 다양한 얼굴 변화를 포착하는 통계적 형태의 모델이 얼굴의 복원을 위해 사용된다. 몇 가지 모델에서는 중립적 표정의 형상만을 포착^[1,8,14]하거나 또는 표정을 추가^[26,35]하기도 한다. 일반적으로 얼굴 모델은 아래와 같이 쌍 선형 표현으로 나타내 진다.

$$S_{geo}(\alpha, \delta) = \bar{S} + \sum_{i=1}^m \alpha_i \cdot S_{id}^{(i)} + \sum_{i=1}^n \delta_i \cdot S_{exp}^{(i)} \quad (1)$$

식 (1)에서 \bar{S} 는 평균 3차원 얼굴 형상, $S_{id}^{(i)}$ 는 어떠한 3차원 얼굴 데이터베이스를 PCA 분석하여 얻은 i 번째 얼굴 형상의 기저벡터 그리고 $S_{exp}^{(i)}$ 는 i 번째 얼굴 표정의 기저벡터를 나타낸다. m, n 은 형상과 표정 계수의 차원을 의미한다. 이 방식의 목표는 입력된 영상에 최대한 가깝게 3DMM을 변경할 수 있는 mn 개의 매개 변수 α, δ 를 추정해내는 것이다. <그림 1>은 3DMM의 형상 및 표정



<그림 1> 3D morphable face model (3DMM), 형상과 표정 매개 변수를 추정하여 입력 영상과 변형된 3차원 얼굴을 일치시킨다.

매개 변수의 역할을 보여주고 있다.

3DMM의 매개변수를 추정하기 위해 대부분의 기존의 방식은 입력 영상과 3차원 평균 모델간의 랜드마크나 국부적 특징의 대응점 간의 일치성을 확립하고자 한다. 이 방식에서는 비선형 최적화 함수를 풀기 위해 대응점간의 거리를 기준으로 식 (2)와 같은 에너지 함수를 설정하고 카메라 자세, 얼굴의 형상 및 표정 매개 변수를 반복적으로 추정한다.

$$E_{fit} = \sum_{k=1}^q \| sRF_{v,k} + T - U_k \| + \lambda_1 \sum_{i=1}^m \left(\frac{w_{id}^{(i)}}{v_{id}^{(i)}} \right) + \lambda_2 \sum_{i=1}^n \left(\frac{w_{exp}^{(i)}}{v_{exp}^{(i)}} \right) \quad (2)$$

여기서 첫 번째 항은 회귀 모델의 데이터 항이며, s 는 스케일, T 는 이동 벡터 그리고 R 은 회전 행렬을 나타낸다. $F_{v,k} \in R^3$ 는 3DMM의 k 번째 키 포인트의 좌표이며 U_k 는 그에 해당하는 2차원 영상의 랜드마크이다. λ_1, λ_2 는 정규화(regularization) 항의 가중치이다. $w_{id}^{(i)}$ 는 i 번째 형상 계수, $w_{exp}^{(i)}$ 는 i 번째 표정 계수이다. $v_{id}^{(i)}$ 와 $v_{exp}^{(i)}$ 는 SVD로 얻어진 특이값이다. 적절한 에너지 최소화 기법을 활용하여 카메라 자세, 형상과 표정 계수를 반복적으로 추정하여 최적의 값으로 수렴시킨다.

지난 십 수 년간 많은 연구자들은 다양한 입력 영상에 대해 얼굴 랜드마크와 지역 특징점을 정확하게 얻거나 3DMM 계수를 정확히 결정하기 위한 강인한 피팅 방법을 만들기 위해 노력하였다. 3DMM을 사용하는 가장 큰 장점은 해공간이 제한되어 문제를 단순화한다는 것이다. 그러나 PCA 기반의 방식은 개략적인 얼굴 형상의 표현은 가능하지만 피부 주름과 같은 얼굴의 세부적인 형상 표현은 제한된다. 게다가 랜드마크를 일치시키는 방식은 특징점 검출기의 정확도에 크게 의존하는 단점을 가진다.

2. 템플릿 기반 방식

템플릿 기반 방식^[21,27]은 입력 영상과 일치하도록 3차원 얼굴 메쉬 모델의 템플릿을 변형시켜서 문제를 해결하는 대안적 접근법이다. 주목할 만한 논문은 Kemelmacher 와 Basri의 논문^[27]이다. 이 논문에서는 참조 모델을 입력 영상과 결합하고, shade-from-shading (SfS) 기법을 사용하여 입력 영상에 일치하는 참조 모델의 표면 반사 특성을 조정한다. 저자는 이를 위



〈그림 2〉 3DMM 기반 방식과 템플릿 기반 방식의 비교.
(좌) 입력영상, (중) 3DMM 방식으로 모델링한 결과^[36],
(우) 템플릿 방식으로 모델링한 결과^[27]. 음영 정보를 결합함으로써
템플릿 기반의 방식은 얼굴의 세부묘사를 더 잘 할 수 있다.

해 Lambertian 반사 모델, 무한 광원, 그리고 참조 모델 간의 대략적인 정렬을 가정하였다. 얼굴 영상의 형성은 다음과 같이 나타낼 수 있다.

$$I(x, y) = \rho(x, y) \vec{Y}(\vec{n}(x, y)) \quad (3)$$

여기서 \vec{n} 는 무한대에 놓인 단일 점광원의 방향과 강도를 나타내는 3차원 벡터, $\vec{n}(x, y)$ 는 각 표면 점에서 표면 법선 벡터, $Y(\vec{n})$ 은 2차 구면조화함수(spherical harmonics function) 벡터이고, $\rho(x, y)$ 는 각 지점의 표면 반사율이다. 참조 모델을 사용하여 조명 및 자세를 추정할 수 있고, 이것들은 표면 반사율의 초기 추정치를 제공한다. 결과적으로 반사율 함수는 미지의 표면 법선의 함수가 된다. $z(x, y)$ 가 정점의 깊이일 때, $n_x = \frac{\partial z}{\partial x}, n_y = \frac{\partial z}{\partial y}$ 로 인해 표면 반사율 방정식은 편미분 방정식이 되고 그 깊이를 계산할 수 있게 된다. 이에 음영 정보를 결합함으로써 이 방식은 이론적으로는 얼굴 영상의 주름과 같은 세부 정보를 복구할 수 있다. 그러나 이 방식은 위에 언급했듯이 광원과 얼굴 반사율 등에 대해 강한 가정을 하였고 이러한 가정은 실제적인 적용 범위를 제한한다.

Hassner^[20]는 입력 영상과 템플릿 모델간의 외형과 깊이의 유사도를 최대화하는 방식을 제안하였다. 템플릿 기반 3차원 얼굴 복원 방식은 입력 영상의 세부 묘사를 복원하는 데에 보다 효과적이지만 템플릿 초기화에 크게 의

존한다. 〈그림 2〉에 3DMM 기반 방식과 템플릿 기반의 방식간의 차이를 나타내었다.

II. 딥러닝을 이용한 3차원 얼굴 형상 복원 기법

지난 수년간 컴퓨터비전 분야에서 CNN이 성공적으로 적용되면서 연구자들은 입력된 영상으로부터 3차원 얼굴을 복원하는 문제에도 CNN을 적용하기 시작했으며 여러 부분에서 비약적인 성능 향상이 이루어 졌다.

1. CNN 기반의 얼굴 특징점 정렬(alignment)

얼굴 특징점 정렬은 얼굴 영상을 정규화 할 수 있는 기준인 ‘눈꼬리’, ‘코 끝’ 그리고 ‘턱 중심’ 등과 같은 얼굴의 기준점들을 찾는 것을 목표로 한다. 이는 많은 얼굴 분석 작업들을 위한 필수적인 전처리 단계이다. 기존의 3DMM 기반의 방식은 대개 얼굴 정렬로부터 시작하여 얼굴 랜드마크와 대응되는 투영된 3차원 정점간의 오차를 최소화하는 식으로 3DMM 계수를 회귀분석하여 추정한다. 딥러닝이 출현하기 이전에는 점진적 회귀분석이 2차원 얼굴 정렬^[28-29]에서 최첨단기술로 쓰였다. 이러한 방법은 최근 LFPW^[30], Helen^[19] 및 300-W^[25] 등과 같은 자세 변화가 제어된 2차원 얼굴 정렬 문제를 통해 대략적으로 ‘해결’된 것으로 간주된다.

[16]에서 제안한 방식은 직렬로 연결된 CNN을 통해 얼굴 랜드마크의 위치를 회귀 분석하였다. 최근에 제안된 [23]의 방식은 발전된 랜드마크 지역화 구조와 잔차 블록(residual block)을 합성하여 고성능의 네트워크를 구축하였다. 이를 대용량의 2차원 얼굴 랜드마크 데이터 셋으로 훈련시킨 후 다른 모든 2차원 얼굴 랜드마크 데이터 셋과 비교하였다.

얼굴 정면에 대한 2차원 랜드마크에서의 큰 성공에도 불구하고 다양한 포즈를 가진 얼굴들은 스스로 가려지는 효과에 의해 일부 랜드마크가 보이지 않기에 여전히 도전적인 과제로 남아있게 된다. 최근 연구^[2-3]에서는 3차원 얼굴 랜드마크 정렬 방식을 제안하였으며 CNN을 사용하여 입력 영상에 반복적으로 3차원 변형 가능 모델을 적용



〈그림 3〉 기존의 얼굴 정렬 알고리즘에서는 다루 힘든 극단적인 자세에서의 [23]의 결과들. 빨간 선은 참값(ground truth), 흰색 선은 [23]의 예측값을 나타낸다. [23]의 예측값이 참값보다 정확하다.

〈표 1〉 다양한 기법의 정량적 성능비교. 300W 실험 데이터에서의 곡선 아래 영역(area-under-the-curve, AUC)과 실패율(0.08의 임계값으로 정규화)의 비교 결과로서 수치는 [37]에서 발췌함. AUG는 높은 값이 우수한 성능임.

알고리즘	평균오차(%)	AUC	실패율 (%)
ESR ^[28]	8.47	26.09	30.5
ERT ^[31]	8.41	27.01	28.83
LBF ^[32]	8.57	25.27	33.67
SDM+ ^[29]	5.83	36.27	13
MDM ^[33]	4.78	45.32	6.8
DAN ^[34]	4.3	47	2.57

하고 정점을 2차원 영상에 투영하여 랜드마크의 위치를 추정하였다. [23]에서는 3차원 얼굴 정렬을 위한 네트워크를 훈련시키고 새로이 만들어진 3차원 랜드마크가 표시된 약 230,000여개의 영상 데이터 셋에서 이를 평가하였다. 〈그림 3〉은 [23]의 얼굴 특징점 정렬 결과를 보인다. CNN 기반의 얼굴 정렬 방식은 어려운 자세와 가려짐에 대한 강한 이점을 나타내고 있다. 〈표 1〉은 서로 다른 얼굴 정렬 방식들의 성능을 보여주고 있다. CNN 기반의 방식이 자세나 가려짐에 의해 어려운 영상에서 더 나은 성능을 보이고 있음을 알 수 있다.

얼굴 특징점 정렬 이후, 3DMM과 영상간의 대응점이 설정될 수 있으며, 그에 따라 비선형 최적화 함수를 풀어서 얼굴 복원을 위한 형상과 표정 계수를 추정할 수 있게 된다.

2. CNN 기반의 3DMM 얼굴 형태와 얼굴 표정 계수 학습

최근의 연구들은 CNN을 통해 3DMM의 얼굴 형상과 표정 계수를 추정한다. [2, 3, 5]는 연속적으로 연결된



〈그림 4〉 랜드마크를 사용하지 않는 3차원 복원을 위한 순차 연결된 (cascade) 네트워크 구조^[6]. CoarseNet, FineNet과 렌더링 층을 결합하였다.



입력 영상 [6] [27] [4] [35] [6] [27] [4] [35]
 〈그림 5〉 입력 영상에 대한 서로 다른 네 가지([6][27][4][35]) 3차원 얼굴 복원 방식들의 결과를 두 개의 시점에서 표현한 결과의 비교. 실험결과는 [6]에서 발췌함.

CNN 구조를 사용하여 랜드마크의 위치와 3DMM 계수를 반복적으로 구하였다.

최근에는 랜드마크 기반의 방식 뿐 아니라 랜드마크를 사용하지 않는 딥러닝을 이용한 3차원 얼굴 복원 방식도 좋은 성능을 보이고 있다. [6]에서는 단일 네트워크를 사용하여 3DMM 계수를 추측한다. 이는 〈그림 4〉에서와 같이 네트워크의 이전 출력을 나타내는 피드백 채널을 일련의 특징지도로 추가하여 수행한다. 하나의 특징지도는 현재 정점의 RGB 채널을 x, y, z 좌표가 일치하는 평균 모델의 정점을 사용해 페인팅하여 계산된다. 다른 특징지도는 정점에서의 법선 벡터를 RGB값으로 사용하는 데에 사용된다. 특징지도는 얼굴 형태의 다양한 표현과 계수의 회귀분석을 돕는다. [8]에서는 형상 매개 변수의 탐지를 직접 학습하는 깊은 네트워크인 reset101을 사용하였다. 이러한 깊은 네트워크를 훈련시키기 위해 저자는 50만여 개의 영상을 사용하였다. [11]은 이 연구를 확장하여 랜드마크 탐색 없이 표정 계수들을 학습하도록 하였다. 이전에도 언급했듯이 3DMM을 통해서 세부정보를 포착하기가 어렵다. 따라서 3DMM의 계수를 추정한 후, [11]은 얼굴 기하 정보를 개선하기 위해 새로운 네트워크를 추가하고 두 네트워크를 연결하기 위한 새로운 렌더링 층을 만들었다. 〈그림 5〉는 서로 다른 방식들에 대한 비교



결과를 보이고 있다. <그림 5>에서 볼 수 있듯이 CNN 기반의 방식은 효율성과 정확성 측면에서 뛰어난 결과를 나타내고 있다.

3. CNN 기반의 영상으로부터 3차원 모델 직접 학습 기법

모델 기반 방식은 모델 형상 공간에 제약이 있는 제한된 형상을 생성한다. 널리 사용되는 Basel Face Model (BFM)^[1]은 무표정으로 제약된 조건에서 촬영한 200개의 피사체로 만들어졌으며, 표정 베이스는 주로 FaceWarehouse 데이터셋^[26]을 사용한다. 적은 트레이닝 세트는 얼굴 형상의 다양성에 한계를 가져온다. 한편 3차원 변형 가능 얼굴 모델은 평균 얼굴과 기초의 선형 조합을 통해 표현할 수 있다. 이때, 선형 행렬 연산으로 볼 수 있는 최적화 알고리즘을 사용하여 새로운 3차원 메쉬를 생성하는 계수를 추정한다. 그러므로 3DMM은 얼굴 모델을 입력 영상과 완벽하게 정렬하는 데에는 성능이 제한적이다.

이 문제를 해결하기 위해 [12]는 3차원 복셀 볼륨을 제안하고 192×192×200 해상도의 복셀 볼륨을 생성하는 volumetric regression network (VRN)을 사용하였다. 먼저 3DMM을 모두 3차원 볼륨으로 바꾸고 얼굴 복원 문제를 3차원 공간 분할 문제로 변환한다. 저자는 하나의 2차원 영상에서 3차원 얼굴 기하학의 체적 표현을 직접 회귀 분석하는 CNN인 VRN을 구축하였다. <그림 6>은 영상과 3차원 볼륨 구조로 훈련된 네트워크를 보이고 있다. 이 논문은 얼굴 복원을 위한 사전 작업으로서의 열



<그림 6> VRN의 네트워크 구조^[12]. 입력된 2차원 영상과 출력된 3차원 볼륨 구조를 보여줌.



<그림 7> VRN의 얼굴 복원 결과 ([12]에서 발췌함). 첫 번째 행은 입력된 2차원 영상이고 두 번째 행은 출력된 3차원 볼륨 구조이다.

굴 정렬을 필요로 하지 않는다. 이 방식은 3차원 얼굴 복원 문제에서 CNN의 가능성을 보여주는 개척자의 역할을 하고 있다. <그림 7>은 3차원 얼굴 복원의 결과를 보여주고 있으며, 그림에서 볼 수 있듯이 3차원 모델이 형상과 표정 모두 입력 영상과 잘 일치한다.

[13]은 완전한 3차원 얼굴 구조의 표현으로 UV 위치 지도를 제안하였다. UV 위치 지도는 UV 공간에서 3DMM의 정점 3차원 좌표를 기록하는 2차원 영상이다. 저자는 입력된 2차원 영상을 완전한 3차원 얼굴 구조로 회귀 분석하는 순차연결 네트워크를 구축하였다. 두 방법 모두 기존의 3DMM 및 맞춤 방식의 한계를 뛰어 넘는다. 딥러닝을 통해 화소를 직접 3차원 점에 매핑하므로 두 방식들은 모델 공간을 기반으로 한 방법이 아니다. 위의 이유로 3차원 모델이 2차원 영상에 거의 완벽하게 정렬된다^[24]. 또한 선형 3DMM 얼굴 모델의 한계를 다루며 기존의 선형 3DMM 보다 더 우수한 표현 능력을 갖는 비선형 모델을 학습하는 기법을 제안하였다.

IV. 향후 연구 전망

기존의 3차원 얼굴 복원에 비하여 CNN 기반의 방식은 다양한 자세와 같은 어려운 조건에서 영상으로부터의 3DMM 계수의 회귀 분석에서 더 뛰어난 성능을 보인다. 또한 딥러닝 네트워크는 기존의 선형 3DMM 모델의 한계를 넘어 입력 얼굴 영상의 화소로부터 직접 3차원 점의 위치를 추정해낼 수 있다. 그러나 CNN 네트워크는 많은 양의 데이터 셋을 통한 지도적 학습을 받아야 하며 입력 영상의 해상도가 상대적으로 낮게 고정된다는 단점이 있다.

향후의 연구는 지도 학습의 주석이 달린 데이터 셋의 부족에 대한 문제를 보완하기 위해 GAN (generative adversarial network)등을 이용한 비지도 학습에 초점이 맞춰질 것이다. 게다가 극단적인 조명, 심한 얼굴의 가려짐 등의 극한의 상황에서의 얼굴 복원 또한 해결되어야 할 문제로 남아있다. 최근 논문^[12-13]은 입력된 2차원 영상으로부터 직접적으로 3차원 구조를 학습하는 최근의 경향을 보여준다. 이는 3차원 변형 가능 얼굴 모델의 한계를 뛰어 넘는다. 그러나 본 방법은 여전히 출력된 3차



원 구조가 입력 영상으로부터의 세부 사항을 복구 할 수 없으며, 일반적으로 입력 해상도가 작은 크기로 고정되는 등의 단점을 가지고 있다. 이것은 미래 연구에서 해결해야 할 문제점이다.

참고 문헌

- [1] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," Proc. ACM SIGGRAPH, 1999
- [2] A. Jourabloo and X. Liu, "Large-pose face alignment via CNN-based dense 3D model fitting," Proc. IEEE CVPR, 2016.
- [3] X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li, "Face alignment across large poses: A 3D solution," Proc. IEEE CVPR, 2016.
- [4] E. Richardson, M. Sela, and R. Kimmel, "3D face reconstruction by learning from synthetic data," Proc. 3DV, 2016
- [5] F. Liu, D. Zeng, Q. Zhao, and X. Liu, "Joint face alignment and 3D face reconstruction," Proc. ECCV, 2016
- [6] E. Richardson, M. Sela, R. Or-EI, and R. Kimmel, "Learning detailed face reconstruction from a single image," Proc. IEEE CVPR, 2016.
- [7] P. Dou, S. K. Shah, and I. A. Kakadiaris, "End-to-end 3D face reconstruction with deep neural networks," Proc. IEEE CVPR, 2017.
- [8] A. T. Tuan, T. Hassner, I. Masi, and I. Medioni, "Regressing robust and discriminative 3D morphable models with a very deep neural network," Proc. IEEE CVPR, 2017
- [9] A. Jourabloo and X. Liu, "Pose-invariant 3D face alignment," Proc. IEEE ICCV, 2015
- [10] S. Laine, T. Karras, T. Aila, A. Herva, and J. Lehtinen, "Facial performance capture with deep neural networks," arXiv preprint arXiv: 1609.06536, 2016
- [11] F.-J. Chang, A. Tran, T. Hassner, I. Masi, R. Nevatia, and G. Medioni, "ExpNet: Landmark free, deep, 3D facial expressions," Proc. IEEE FGR, 2018
- [12] A. S. Jackson, A. Bulat, V. Argyriou, and G. Tzimiropoulos "Large pose 3D face reconstruction from a single image via direct volumetric CNN regression," Proc. IEEE CVPR, 2017
- [13] Y. Feng, F. Wu, X. Shao, Y. Wang, X. Zhou, "Joint 3D face reconstruction and dense alignment with position map regression Network," arXiv preprint arXiv:1803.07835, 2018
- [14] V. Blanz, K. Scherbaum, T. Vetter, and H. Seidel, "Exchanging faces in images," Computer Graphics Forum, 2004.
- [15] B. Chu, S. Romdhani, and L. Chen, "3D-aided face recognition robust to expression and pose variations," Proc. IEEE CVPR, 2014
- [16] Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," Proc. IEEE CVPR, 2013
- [17] S. Romdhani and T. Vetter, "Efficient, robust and accurate fitting of a 3D morphable model. Proc. IEEE ICCV, 2003
- [18] H. Tang, Y. Hu, Y. Fu, M. Hasegawa-Johnson, and T. S. Huang, "Real-time conversion from a single 2d face image to a 3D text-driven emotive audio-visual avatar," Proc. IEEE ICME, 2008
- [19] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang, "Interactive facial feature localization," Proc. ECCV, 2012
- [20] T. Hassner and R. Basri, "Example based 3D reconstruction from single 2D images," Proc. IEEE CVPR Workshops, 2006
- [21] T. Hassner, "Viewing real-world faces in 3D," Proc. IEEE ICCV, 2013.
- [22] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. "Deepface: Closing the gap to human-level performance in face verification," Proc. IEEE CVPR, 2014.
- [23] A. Bulat and G. Tzimiropoulos, "How far are we from solving the 2D&3D face alignment problem? (and a dataset of 230,000 3D facial landmarks)," Proc. IEEE CVPR, 2013
- [24] L. Tran and X. Liu, "Nonlinear 3D face morphable model," Proc. IEEE CVPR, 2018
- [25] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "A semi-automatic methodology for facial landmark annotation," Proc. IEEE CVPR, 2013.
- [26] C. Cao, Y. Weng, S. Zhou, Y. Tong, and K. Zhou, "FaceWarehouse: A 3D facial expression database for visual computing," IEEE TVCG, 2014.



[27] I. Kemelmacher-Shlizerman and R. Basri, "3D face reconstruction from a single image using a single reference face shape," IEEE TPAMI, 2011.

[28] X. Cao, Y. Wei, F. Wen, and J. Sun, "Face alignment by explicit shape regression," Proc. IEEE CVPR, 2012

[29] X. Xiong and F. DelaTorre, "Supervised descent method and its applications to face alignment," Proc. IEEE CVPR, 2013

[30] P. Belhumeur, D. Jacobs, D. Kriegman, and N. Kumar, "Localizing parts of faces using a consensus of exemplars," Proc. CVPR, 2011.

[31] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," Proc. IEEE CVPR, 2014.

[32] S. Ren, X. Cao, Y. Wei, and J. Sun, "Face alignment at 3000 fps via regressing local binary features," Proc. IEEE CVPR, 2014

[33] G. Trigeorgis, P. Snape, M. A. Nicolaou, E. Antonakos, and S. Zafeiriou, "Mnemonic descent method: A recurrent process applied for end-to-end face alignment," Proc IEEE CVPR, 2016

[34] M. Kowalski, J. Naruniec, and T. Trzcinski, "Deep alignment network: A convolutional neural network for robust face alignment," Proc. IEEE CVPR Workshops, 2017.

[35] Zhu, X., Lei, Z., Yan, J., Yi, D., Li, S.Z, "High fidelity pose and expression normalization for face recognition in the wild," Proc. IEEE CVPR, 2015.

[36] Grewe, C.M., Zachow, S, "Fully automated and highly accurate dense correspondence for facial surfaces," Proc. IEEE ECCV 2016.

[37] Y. Chen, C. Shen, H. Chen, X. S. Wei, L. Liu, J. Yang, "Adversarial Learning of Structure Aware Fully Convolutional Networks for Landmark Localization," arXiv preprint arXiv: 1711.00253, 2017.



원소미

- 2016년 8월 후남대학교 컴퓨터공학과 학사
- 2016년 9월~현재 인하대학교 정보통신공학과 석사과정

〈관심분야〉
컴퓨터 비전 및 그래픽스, deep learning



이정우

- 2018년 8월 인하대학교 정보통신공학과 학사
- 2018년 9월~현재 인하대학교 정보통신공학과 석사과정

〈관심분야〉
컴퓨터 비전 및 그래픽스, deep learning



박인규

- 1995년 2월 서울대학교 제어계측공학과 학사
- 1997년 2월 서울대학교 제어계측공학과 석사
- 2001년 8월 서울대학교 전기컴퓨터공학부 박사
- 2001년 9월~2004년 3월 삼성종합기술원 전문연구원
- 2007년 1월~2008년 2월 미국 MERL 방문연구원
- 2014년 9월~2015년 8월 MIT 미디어랩 방문부교수
- 2004년 3월~현재 인하대학교 정보통신공학과 교수

〈관심분야〉
컴퓨터비전 및 그래픽스(computational photography, light field, 영상 기반 얼굴/바디 모델링), GPGPU