

딥러닝을 이용한 3차원 사람모델형상 변형⁺

(3D Human Shape Deformation using Deep Learning)

김 대 희¹⁾, 황 본 우²⁾, 이 승 욱³⁾, 곽 수 영^{4)*}

(DaeHee Kim, Bon-Woo Hwang, SeungWook Lee, and Sooyeong Kwak)

요 약 최근 가상현실 및 증강 현실 기술을 이용한 다양한 응용분야가 각광받으면서 빠르고 정확한 3차원 모델 생성이 요구되고 있다. 본 논문에서는 옷을 입은 3차원 사람 모델을 포인트 클라우드의 형상으로 변형하는 온-사이트 학습 (On-site learning) 기반 형상 변형 방법을 제안한다. 제안하는 알고리즘은 사전 학습과 온-사이트 학습 두 개의 파트로 구성되어 있으며, 각각의 학습은 인코더 네트워크, 템플릿 변형 네트워크, 디코더 네트워크로 구성된다. 딥러닝 네트워크 학습은 3차원 포인트 클라우드와 템플릿 정점 사이의 챔퍼 거리 (Chamfer distance)를 주요 손실 함수로 사용하는 비지도 학습을 적용한다. 입력된 포인트 클라우드 형태의 데이터에 대해 온-사이트 학습을 진행함으로써 추론의 결과물에 대한 높은 정확도를 얻을 수 있으며 이를 실험을 통해 제시한다.

핵심주제어: 딥러닝, 온-사이트 학습, 형상 변형, 3차원 사람 모델

Abstract Recently, rapid and accurate 3D models creation is required in various applications using virtual reality and augmented reality technology. In this paper, we propose an on-site learning based shape deformation method which transforms the clothed 3D human model into the shape of an input point cloud. The proposed algorithm consists of two main parts: one is pre-learning and the other is on-site learning. Each learning consists of encoder, template transformation and decoder network. The proposed network is learned by unsupervised method, which uses the Chamfer distance between the input point cloud form and the template vertices as the loss function. By performing on-site learning on the input point clouds during the inference process, the high accuracy of the inference results can be obtained and presented through experiments.

Keywords: Deep learning, On-site learning, Shape deformation, 3D human model

* Corresponding Author: sykwak@hanbat.ac.kr

+ 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2018년도 문화기술연구개발지원사업의 연구결과로 수행되었음(R20183 0391)

Manuscript received February 01, 2020 / revised March 24, 2020 / accepted March 30, 2020

1) CONTEC, 제1저자

2) 한국전자통신연구원, 제2저자

3) 한국전자통신연구원, 제3저자

4) 한밭대학교 전자제어공학과, 교신저자

1. 서 론

최근 3차원 광학 기술 및 컴퓨팅 성능의 발전으로 인해 3차원 복원 기술에 대한 수요가 높다. 광학 장치를 이용한 3차원의 외형 복원은 카메라 이미지를 이용한 수동 센서방식과 능동 센싱 방식으로 나누어진다. 수동 센서방식은 영

상의 해상도가 증가함에 따라 처리하는 데이터의 양이 급속히 늘어나므로 해상도에 한계가 있으며 조명의 영향을 많이 받는 문제가 있다. 능동 센서방식은 정확도는 높지만 가격이 비싸며 숙련된 엔지니어가 필요하다. 최근에는 광학 소자 기술의 발전으로 인해 낮은 가격으로 보급이 확산되고 있으며, 컴퓨팅 성능의 발전으로 카메라의 해상도 및 처리 능력이 향상되면서 인체의 3차원 복원에 관한 연구가 각광 받고 있다. 인체의 3차원 복원 연구는 광학 센서로부터 획득한 데이터를 정확한 3차원 데이터를 표현하기에는 노이즈 및 부족한 표면처리 등이 문제가 있어 원시 데이터에 대한 후처리 과정이 필수적이다. 일반적으로 세밀한 3차원 모델링은 그래픽스 툴을 이용하여 수작업을 통해 제작되는데 이는 작업시간이 많이 필요하며 숙련된 디자이너가 요구된다.

본 논문에서는 3차원 프린팅 및 애니메이션 등 다양한 응용에 적용 가능한 옷을 입고 있는 사람에 대한 3차원 형상을 다양한 자세로 변형하는 방법을 제안한다. 포인트 클라우드로 표현된 다양한 자세의 입력 형상과 텍스처를 표현한 템플릿이 주어졌을 경우 두 입력의 상관관계를 파악하여 템플릿의 형상을 옷을 입은 다양한 자세의 형상으로 변환하는 네트워크를 제안한다. 본 논문은 사람 형태에서 다양한 옷에 대하여 세밀한 형상 변형을 위해 전체 네트워크의 온-사이트 학습 (On-site learning)을 통해 더 높은 정확도를 얻고자 한다.

2. 관련연구

3차원 데이터를 표현하는 방식은 프리미티브 (Primitive), 프로젝션 (Projection), RGB-D, 다시점으로 표현하는 유클리디언 (Euclidean) 방식과 메쉬 (Mesh), 그래프, 포인트 클라우드 (Point cloud)로 표현하는 비 유클리디언 (Non euclidean)으로 나눌 수 있다.

형상 변형을 위해서는 데이터를 표현하는 방식에 따라 네트워크의 모델링 방식이 달라진다. 본 논문에서는 비 유클리디언 방식으로 표현된 3차원

데이터에 대한 형상 변형 네트워크를 다룬다.

최근 비 유클리디언 방식으로 표현된 3차원 데이터 표현을 다루기 위해 그래프 기반 네트워크 또는 포인트 클라우드 기반 네트워크에 대한 연구가 활발히 진행중에 있다.

메쉬 그래프 네트워크는 입력으로 그래프 구조와 각 노드 별 특징 정보를 받는다. 입력으로 받은 특징 정보와 그래프 내에서 나타나는 이웃 정보를 바탕으로 각 노드 별 벡터 형태를 출력 결과로 얻어낸다. 각 노드들은 그래프 상의 이웃 정보와 자기 자신의 정보를 이용하여 벡터를 만든다. 컨볼루션 뉴럴 네트워크 (Convolutional neural network)에서 인접한 벡터의 정보를 함께 사용하는 필터처럼 그래프 뉴럴 네트워크 (Graph neural network)는 인접한 노드들의 정보를 공유하는 구조이다. 이러한 벡터를 생성하는 층을 다중으로 쌓는다면 3차원 그래프 구조 및 메쉬 형태로 학습이 가능하다. 그래프 기반 네트워크는 인근 노드의 잡음 혹은 누락된 데이터에 대한 처리의 어려움을 극복하기 힘들다. 또한, 데이터의 정점이 많을수록 그래프에 대한 벡터를 생성하는 그래프 네트워크 층을 깊게 쌓아야 하는 문제로 인해 일반적으로는 저용량의 정점 데이터를 사용한다.

3차원 데이터 표현 방식의 그래프 기반 네트워크 연구 중 하나로 Wang et al. (2018)은 2차원 RGB영상을 삼각 메쉬 형태로 3차원 형상을 생성하는 네트워크를 제안하였다. 이는 입력 영상에서 추출된 전반적인 특징 벡터를 이용하여 그래프 기반 컨볼루션 신경망을 통해 타원체의 3차원 형상을 점진적으로 변형시키는 방법이다. 네트워크는 이미지 특징 추출 네트워크와 계층적 메쉬 변형 네트워크로 구성되어 있다. 이미지 특징 추출 네트워크는 입력 이미지에서 지각적 특징(Perceptual feature)을 추출하는 2차원 컨볼루션 뉴럴 네트워크이며, 추출된 특징은 타원체 메쉬를 3차원 모델로 점진적으로 변형시키기 위한 메쉬 변형 네트워크에 의해 활용된다. 계층적 메쉬 변형 네트워크는 그래프 기반 컨볼루션 뉴럴 네트워크를 사용하며, 3개의 메쉬 변형층과 메쉬 변형층 사이에 2개의 그래프 언풀링층(Graph unpooling layer)로 구성된다. 각각의 메쉬 변형층은 현재 메쉬 모델을 나타내는

입력 그래프를 이용하여 새로운 정점 위치를 특정하여 정점을 생성한다. 그래프 언폴딩층은 현재 메쉬의 정점 수를 늘려 세부적인 특징을 늘리면서 삼각형 메쉬 토폴로지를 유지한다. 또한, Chen et al. (2019)은 3차원 스캐너를 이용하여 다양한 의복을 착용한 사람 형태에서 인체 모양과 의복을 각각 추론하는 네트워크를 제안하였다. 3차원 메쉬의 주요 특징(관절, 손, 머리, 발 등)을 정렬하여 기하학적 영상에 매핑하였다. 3차원 메쉬에 사용된 데이터는 골격과 관절 각도를 다양하게 변형할 수 있는 인간 모델을 사용하였다(Loper et al., 2015). 신체에 대해 추론을 위해 Conditional GAN을 훈련하여 색상, 질감, 곡률, 위치 등을 포함한 특징 벡터를 추출한다. 신체와 기하학적 영상 사이의 벡터를 신체적 특징의 척도로 사용하고 이를 이용하여 의복을 제외한 신체 모양을 예측한다.

메쉬 기반 네트워크 및 포인트 클라우드 기반 네트워크는 모든 점에 대한 정보를 단일 벡터로 특징을 일반화하기 때문에 잡음 혹은 누락된 데이터에 강인하고 대용량 정점에도 사용할 수 있다. 기존의 3차원 딥러닝 네트워크는 사용되는 3D 데이터를 볼륨 표현을 이용하여 유클리디언 방식으로 변형하여 진행하였다. 하지만 유클리디언 방식으로 변형된 데이터는 불필요한 빈 공간이 발생하여 컴퓨터의 연산 처리량이 많아지고, 데이터 손실이 반드시 발생한다. 이러한 단점을 막고자 Qi et al. (2017)은 비 유클리디언 방식의 비정형 데이터 중 하나인 포인트 클라우드를 직접 사용한다. 포인트 클라우드 형식의 데이터를 사용하기 위해서는 두 가지의 성질을 만족해야 하며, 첫 번째로 순열 불변성이다. 순열 불변성이란 입력 벡터 요소의 순서와 상관없이 같은 출력을 생성하는 모델을 말한다. 포인트 클라우드 안에 n 개의 포인트가 있다면, 이러한 포인트 클라우드를 처리하는 모델은 $n!$ 개의 순열에 대해서 불변성을 지니는 것이다. 포인트 넷에서는 다층 퍼셉트론과 최대 풀링을 이용하여 포인트 클라우드의 각 포인트로부터 추출한 정보들을 합치는 과정을 수행하여 순열 불변성을 만족한다.

두 번째로 이동 불변성이다. 포인트 클라우드의 회전이나 따라 분류 결과를 변경하지 않아야 한다.

포인트 넷에서는 공간 변환 네트워크 (Jaderberg et al., 2015)모듈을 사용하여 3차원 객체의 골격에 해당하는 핵심 포인트를 학습하여 이동 불변성의 문제를 해결하였다. n 개의 포인트 클라우드를 입력으로 입력 및 특징 변환을 수행하고, 최대 풀링을 통해 모든 점에 대한 정보를 단일 벡터로 특징을 일반화 한다.

3차원 포인트 클라우드 기반의 네트워크의 연구 중 하나로 Groueix et al. (2018)는 다양한 자세를 갖는 입력 데이터와 3차원 템플릿 데이터를 인코더-디코더 형태의 형상 변형 네트워크를 이용하여 자세가 일치하는 변형된 템플릿 형상 생성을 목표로 한다. 3차원 인체 모델의 자세 일치성을 찾는 기존과는 다르게 (Allen et al., 2002; Allen et al., 2003; Zuffi et al., 2015) 인코더-디코더 형태의 딥러닝 네트워크인 3차원 형상 변형 네트워크를 이용하여 형상을 일치시키기 위한 연구다. 형상의 변형은 포인트 클라우드 기반 네트워크를 사용하여 입력 포인트 클라우드와 템플릿 포인트 클라우드의 전역 벡터를 추출하고 입력과 템플릿 사이의 챔퍼 거리를 최소화하는 손실함수를 이용하여 학습한다. 이때, 입력 포인트 클라우드의 형상과 템플릿 포인트 클라우드 형상의 차이가 큰 경우에는 형상이 심하게 왜곡될 수 있기 때문에 이러한 왜곡 방지를 위해 2개의 손실함수를 추가하여 사용하여 대략적인 형상의 변형을 학습하였다. 추론 과정에서 디코더 네트워크 모듈만을 재학습하여 입력의 형상에 최적화하여 형상을 변형하는 방식이다.

본 논문에서는 3D-CODED 방법의 왜곡에 대한 문제점을 온-사이트 학습방법을 통해 개선한 형상 변형 네트워크를 제안하고 3D-CODED의 변형 모델보다 개선된 성능을 보이는 결과를 제시한다.

3. 형상 변형 네트워크 설계

메쉬 그래프 기반 형태의 네트워크는 인접한 노드의 정보를 공유하는 구조에서 인근 노드의 잡음 혹은 누락된 데이터에 대한 어려움을 극복하기 힘든 단점이 있다. 적은 양의 데이터에서는 높

은 효율을 보이지만 데이터 크기가 커질수록 그래프에 대한 벡터를 생성하는 그래프 네트워크 층을 다중으로 쌓아야 하므로 학습에 있어서 비효율적이다. 따라서, 본 논문에서는 이러한 문제점을 극복하기 위해 포인트 클라우드 형식 데이터를 입력으로 사용하였다. 포인트 클라우드 형태의 입력 형상을 템플릿과 매칭시키기 위해서 기본적으로 인코더-디코더 형태의 네트워크를 이용하여 학습한다. 제안하는 형상 변형 네트워크는 인코더 네트워크, 템플릿 변형 네트워크, 디코더 네트워크 3가지로 구성되어 있으며 Fig. 1에 나타내었다. 인코더 네트워크에서는 입력 데이터의 전역 특징을 추출하고, 템플릿 변형 네트워크에서는 템플릿의 국부특징을 추출한다. 전역 특징과 국부 특징을 디코더 네트워크의 입력값으로 사용하여 최종적으로 변형된 형상을 얻게 된다.

3.1 인코더 네트워크

Fig. 2의 인코더 네트워크 모듈에서는 입력 객체의 $n \times 3$ 의 포인트 클라우드 형태를 입력 받는다. 각각의 점은 공유되는 3개의 은닉층 (Hidden layer)과 배치 정규화 (Batch normalization), 최대 풀링 (Max pooling)을 적용한 후, 정류 선형 유닛 (Rectified linear unit: ReLU)을 활성화 함수로 적용시켰으며, 이후 선형 층 (Linear layer)을 거쳐 n 개의 포인트 클라우드에 대한 1024 차원의 전역 특징벡터를 생성한다.

3.2 템플릿 변형 네트워크

Fig. 3의 템플릿 변형 네트워크 모듈은 템플릿 데이터 포인트 클라우드 형태 $n \times 3$ 을 입력 받으며, 3개의 은닉층과 배치 정규화, 쌍곡탄젠트 함수 (Hyperbolic tangent function)를 거쳐 템플릿의 포인트 클라우드를 변형하여 8 크기의 국부 특징 벡터를 생성한다.

3.3 디코더 네트워크

Fig. 4의 디코더 네트워크 모듈은 템플릿 변형 네트워크 모듈 출력의 템플릿 각 점의 국부 특징

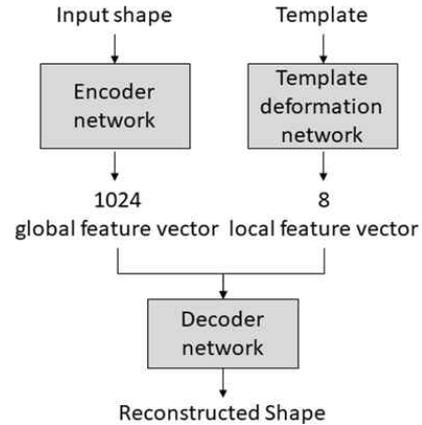


Fig. 1 Proposed Shape Deformation Network

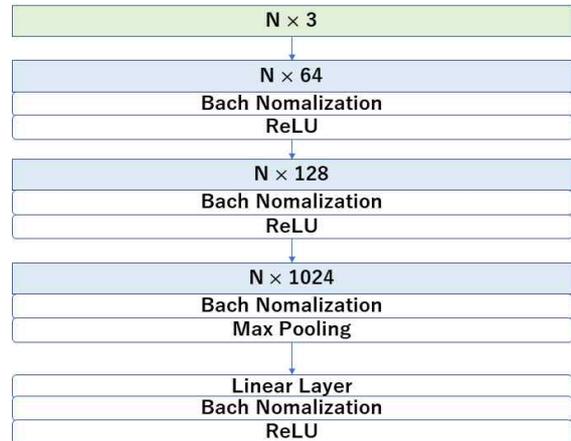


Fig. 2 Encoder Network Module

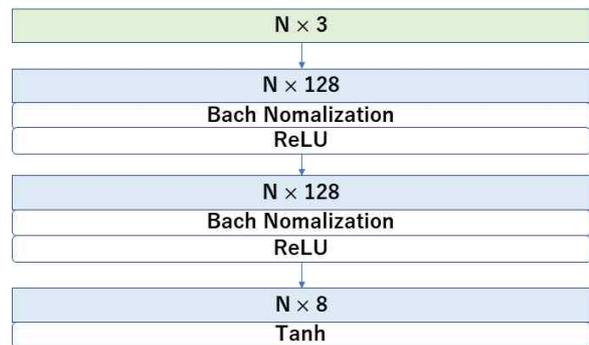


Fig. 3 Template Deformation Network Module

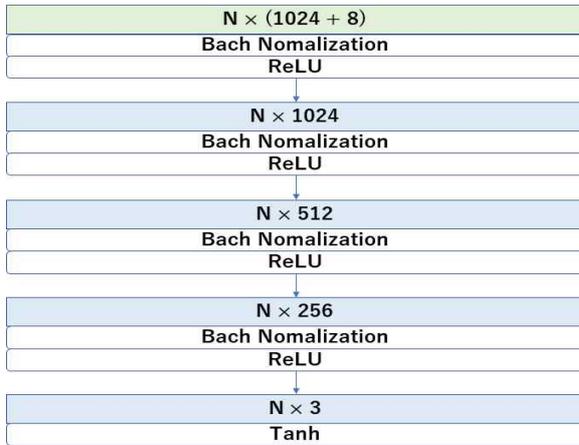


Fig. 4 Decoder Network Module

과 인코더 출력의 전체 입력 포인트 클라우드의 전역 특징이 합쳐져 디코더 네트워크 모듈의 입력으로 사용된다. 디코더 네트워크 모듈은 4개의 은닉층과 배치 정규화, 쌍곡선 함수로 구성된다. 제안된 인코더-디코더 형태의 네트워크는 중단 간 학습 방식으로 학습이 진행된다.

4. 온-사이트 학습

기존의 머신러닝에서 많이 사용하는 비지도 학습방법을 적용하게 되면 변형된 형상의 정확도가 떨어지는 문제점이 있다. 본 논문에서는 이러한 문제점을 개선하기 위해서 사전 학습된 모델로 유추한 정보를 초깃값으로 두고 지정된 형상 간의 변형을 전체 네트워크의 재학습을 통해서 세밀한 변형 결과를 얻는 온 사이트 학습을 제안한다.

4.1 사전 학습

사전 학습에서는 학습 데이터의 형상과 학습 템플릿 데이터의 형상 사이의 변형을 예측하고 형상 사이의 차이를 줄이는 방법을 학습한다. 여기서 사전 학습은 온-사이트 학습의 초깃값 역할을 한다. 사전 학습에 대한 손실 함수가 낮게 나올수록 온-사이트 학습의 초깃값이 정확하게 입력할 수 있으며 사전 학습에 사용되는 손

실 함수는 식 (1)과 같다.

$$L_{prelearning} = L^{Chamfer} + \lambda_{Lap} L^{Lap} + \lambda_{edge} L^{edge} + \lambda_{fnormal} L^{fnormal} \quad (1)$$

식 (1)에서 $L^{Chamfer}$ 는 입력 데이터와 네트워크로 재구성된 포인트 클라우드 사이의 차이를 얻기 위해 챔퍼 거리를 사용하였고 (Fan et al., 2017), 이를 최소화함으로써 입력 데이터의 형상에 근접한다. 입력 형상과 템플릿 형상 간의 차이가 적을 경우에는 챔퍼 거리만 사용해도 되지만 입력 포인트 클라우드의 형상과 템플릿 포인트 클라우드 형상의 차이가 큰 경우에는 챔퍼 거리 손실 함수만 학습에 적용할 경우 형상이 심하게 왜곡될 수 있다. 발생하는 왜곡을 방지하고자 $\lambda_{Lap} L^{Lap}$, $\lambda_{edge} L^{edge}$, $\lambda_{fnormal} L^{fnormal}$ 3개의 정규화 손실을 추가한다. L^{Lap} 은 로스라플라시안으로 템플릿과 형상 변형된 템플릿 메쉬에 대한 라플라시안 형태 변형 정도를 나타낸다. L^{edges} 는 로스 엣지로 재구성된 데이터와 템플릿에 대한 에지 길이의 변화 정도를 나타내며, $L^{fnormal}$ 는 재구성된 데이터와 템플릿의 메쉬 노멀 간의 코사인 유사도 (Cosine Similarity)를 의미한다. 이때 λ_{Lap} , λ_{edge} , $\lambda_{fnormal}$ 는 $L^{Chamfer}$ 값을 조절하는 가중치이다. 사전 학습에서는 이 4가지 성분을 포함한 손실 값을 최소화한다. 챔퍼 거리와 3개의 정규화 손실에 각각 가중치를 적용한다. 가중치는 손실 함수에 있어서 중요도에 따른 크기와 학습 진도에 따라 가중치를 감소하여 적용한다.

4.2 온-사이트 학습

추론 과정에서 새로운 입력 데이터에 대한 높은 정확도를 얻기 위해 온-사이트 학습을 진행한다. 온-사이트 학습에서는 사전 학습과 다르게 입력된 하나의 데이터에 대해서만 손실 함수를 갱신하기 때문에 사전 학습 모델의 손실 함수를 사전 학습보다 더욱 빠르게 최소화 할 수 있다. 온-사이트 학습에 사용되는 손실 함수는 빠른 학습을 위해 식 (2)와 같이 챔퍼 거리만 적용한다.

$$L_{on} = L^{Chamfer} \quad (2)$$

온-사이트 학습은 추론에 있어 반복 횟수만큼 시간이 소모된다는 단점이 있지만 이를 보완하기 위해 챔퍼 거리만 손실 함수로 적용한다. 사전 학습된 모델에는 손실 함수를 학습시킨 모델의 가중치에 이미 4가지의 손실이 적용되어 있어 챔퍼 거리만 사용한 온-사이트 학습만으로 템플릿 변형 왜곡이 크게 발생하지 않는다. 따라서 온-사이트 학습에는 전체에 손실 함수를 계산하지 않고 챔퍼 거리에 관련 손실 함수만을 계산함으로써 학습시간을 줄이면서 높은 정확도를 얻을 수 있다. 기존의 3D-CODED 방식은 디코더만 학습에 참여하여 사람 혹은 동물 형태의 다양한 자세에 대하여 사전 학습된 기존의 템플릿 데이터에서 입력 데이터의 대략적인 변화를 얻을 수 있다. 본 논문에서 제안하는 온-사이트 학습은 인코더, 템플릿 변형, 디코더 네트워크에 대하여 학습한다. 전체 네트워크에 대해 학습하는 이유는 정확도를 높이기 위함이다. 입력 데이터의 전역 벡터를 온-사이트 학습 횟수만큼 학습하여 템플릿의 포인트 클라우드가 이전 학습에서 미처 이동하지 못한 전역 벡터에 대하여 이동함으로써 해당 입력 데이터와 템플릿 데이터에 최적화되어 다양한 의상이나 체형에 대해서 더욱 정밀한 결과를 얻을 수 있다.

5. 실험 및 결과 분석

5.1 실험 환경 및 구성

본 논문에서 제안하는 온-사이트 학습 기반의 형상 변형에 대한 성능을 평가하기 위해 첫 번째로, 사전 학습만 사용한 추론 결과의 입력 데이터 간 챔퍼 거리와 사전 학습을 초기값으로 두고 온-사이트 방식을 적용한 추론 결과의 입력 데이터 간 챔퍼 거리를 비교하여 온-사이트 방식을 이용한 형상 변형의 정확성을 증명한다. 두 번째로, 기존의 연구인 3D-Coded의 디코더 학습 방식과 본 논문의 온-사이트 방식의 형상 변형에 대한 정확

성을 비교하여 제안하는 방식의 정확성을 증명한다. 본 논문의 실험은 Table 1과 같은 환경에서 구현하여 실험하였다.

Table 1 Computer Hardware and Software Specifications used in the Experiment

CPU	Intel Xeon(R) Gold 6140 2.30GHz * 2
Memory	32GB DDR PC4-21300 * 4
Graphic Card	NVIDIA TITAN Xp * 2
OS	Microsoft Windows 10
Tool	Python 3.7.3(Pytorch 1.0.0)

5.2 데이터 셋

본 실험에 사용된 데이터는 3차원 캐릭터 생성 툴인 DAZ Studio (Daz productions)를 이용하여 기반 캐릭터를 생성하였다. Fig. 5는 DAZ Studio 사이트에서 생성한 캐릭터 Paige 캐릭터를 나타낸 것이다. 3차원 캐릭터를 바탕으로 각각 직접 생성한 여성, 남성 각각 상의 50개, 하의 50개를 조합 (Fig. 6)하여 캐릭터에 착장하여 총 2,500개의 3차원 캐릭터를 생성하였다. 이때 사용한 생성한 데이터는 정점 2만개, 삼각형태의 면 4만개 내외로 PLY 파일 형식으로 생성하였다. 입력 데이터의 사전 정규화를 통해 포인트 넷에서의 공간 변환 네트워크(Qi et al., 2017) 모듈을 생략하였다.

5.3 사전 학습과 온-사이트 학습 분석

사전 학습을 초기값으로 설정하고 초기값의 전역 벡터를 온-사이트 학습 횟수만큼 학습하여 사전 학습과 온-사이트 학습의 챔퍼 거리를 비교하여 성능을 평가하였다. 이때 챔퍼 거리의 단위는 cm이다. Table 2는 각각의 데이터에 대해서 사전 학습과 온-사이트 학습 방식의 챔퍼 거리를 비교하여 나타내었다. 온-사이트 학습을 통해 7,000 epoch만큼 학습시켰을 경우 사전학습결과보다 챔퍼 거리가 평균적으로 0.84% 줄어드는 것을 알 수 있으며, 추론의 반복횟수가 늘어날수록 에러가 줄어드



Fig. 5 Pace Model in DAZ Studio



Fig. 6 50 Tops and Bottoms Paige Model

Table 2 Comparison of Chamfer Distances between Pre-learning and On-site Learning (cm)

	Pre-learning	On-site learning epoch=1,000	On-site learning epoch=3,000	On-site learning epoch=5,000	On-site learning epoch=7,000
Data 1	7.78	0.79	0.68	0.64	0.61
Data 2	3.30	0.82	0.72	0.68	0.66
Data 3	2.40	0.80	0.67	0.62	0.59
Data 4	1.98	0.69	0.61	0.58	0.57
Data 5	2.39	0.72	0.63	0.60	0.58
Data 6	6.61	1.52	1.07	0.82	0.79
Data 7	1.66	0.85	0.71	0.67	0.64
Data 8	5.47	1.49	1.21	1.10	1.08
Data 9	7.50	3.10	2.82	2.59	2.31
Data 10	2.41	0.82	0.65	0.60	0.58

는 것을 볼 수 있다. Fig. 7, 8, 9는 각각 데이터 1, 2, 3에 대한 온-사이트 학습 횟수에 따르는 형상 변형 결과 부위의 확대결과를 나타내었다. 사전학습과 온-사이트 학습의 챔퍼 거리 비교를 통해 템플릿의 포인트 클라우드가 이전 학습에서 미처 이동하지 못한 전역 벡터에 대하여 이동함으로써 해당 입력 데이터와 템플릿 데이터에 최적화되어 다양한 의상에 대해서 더욱 정밀한 결과를 얻을 수 있음을 보인다.

5.4 3D-Coded와 제안한 방법 비교 분석

Table 3은 각각의 데이터 1, 2, 3에 대해서 온-사이트 학습과 기존 방식의 디코더 학습 횟수에 따르

는 챔퍼 거리를 나타내었다. Table 3과 Fig. 7에서 제안한 방법이 기존의 방법과 비교하여 형태 변형에 있어서 시각적인 면에서나 챔퍼 거리를 통한 정량적인 비교에서 더 높은 정확도를 보이는 것을 확인할 수 있다.

6. 결론 및 향후 연구방향

본 논문에서는 템플릿을 입력받은 형상으로 변형하는 방법에 대해 다루었다. 변형의 정확도를 위해 포인트 클라우드 형태를 직접 입력하는 방식의 포인트 넷을 사용하였으며, 정확도를 높이기 위해 온-사이트 학습 기반의 형상 변형 방법을 제안했다. 결과

Table 3 Comparison of Chamfer Distances between 3D-Coded and Proposed Method (cm)

		Epoch=1,000	Epoch=3,000	Epoch=5,000	Epoch=7,000
Data 1	3D-Coded	2.07	1.66	1.50	1.38
	Proposed	0.79	0.68	0.64	0.61
Data 2	3D-Coded	1.62	1.39	1.29	1.23
	Proposed	0.82	0.72	0.68	0.66
Data 3	3D-Coded	1.41	1.26	1.17	1.12
	Proposed	0.80	0.67	0.62	0.59

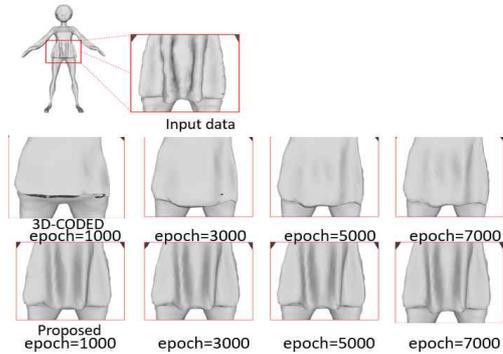


Fig. 7 Comparison Results between 3D-Coded and Proposed Method with Data 1

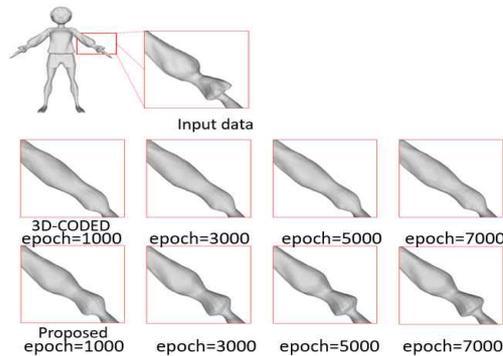


Fig. 8 Comparison Results between 3D-Coded and Proposed Method with Data 2

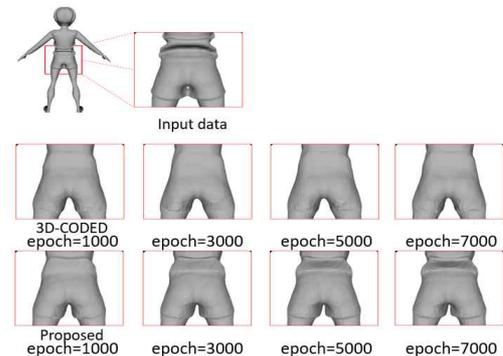


Fig. 9 Comparison Results between 3D-Coded and Proposed Method with Data 3

적으로 기존에 사용되던 3D-Coded 네트워크를 이용하여 형상 변형 한 결과보다 온-사이트 학습 기반의 예측 결과가 높은 정확도를 보여주었다. 향후 더 높은 정확도와 온-사이트 학습에서의 계산량 감소를 위해 각각의 영역별로 분할한 디코더 네트워크 적용을 추가하여 3차원 포인트 클라우드 형상의 각각의 부분별로 학습하여 온-사이트 학습을 시킨다면 이 방법은 현재보다 더욱 정확도가 높은 형상 변형 네트워크가 될 것으로 기대한다.

본 논문의 3차원 토폴로지를 토대로 리깅 및 텍스처를 설계한다면 프린팅, 애니메이션, 게임 및 영화 등 다양한 응용에 적합한 형상을 적용할 수 있을 것으로 기대된다.

References

- Allen, B., Curless, B., Popovic, Z. (2002). Articulated Body Deformation from Range Scan Data, *Transactions On Graphics*, 21(3), 612-619.
- Allen, B., Curless, B., Popovic, Z. (2003). The Space of Human Body Shapes: Reconstruction and Parameterization from Range Scans, *Transactions On Graphics*, 22(3), 587-594.
- Chen, X., Pang, A., Zhu, Y., Li, Y., Luo, X., Zhang, G., Wang, P., Zhang, Y., Li, S., Yu, J. (2019). Towards 3D Human Shape Recovery Under Clothing, *arXiv preprint arXiv:1904.02601*.
- Daz Productions, <https://www.daz3d.com/> (accessed Aug 10, 2019)
- Fan, H., Su, H., Guibas, L. (2017). A Point Set Generation Network for 3D Object

Reconstruction from a Single Image, *The IEEE Conference on Computer Vision and Pattern Recognition*, July. 21-26, Honolulu, USA, pp. 605-613.

Groueix, T., Fisher, M., Kim, Vladimir, G., Russell, Bryan, C., Aubry, M. (2018). 3D-CODED: 3D Correspondences by Deep Deformation, *The European Conference on Computer Vision*, Sep. 8-14, Munich, Germany, pp. 230-246.

Jaderberg, M., Simonyan, K., Zisserman, A. (2015). Spatial Transformer Networks, *The Advances in Neural Information Processing Systems*, Dec. 7-12, Montreal, Canada, pp. 2017-2025.

Loper, M., Mahmood, N., Romero, J., Moll, G., Black, MJ. (2015) SMPL: A Skinned Multi-person Linear Model, *The ACM Transactions On Graphics*, 34(6), 248.

Qi, CR., Su, H., Mo, K., Guibas, LJ. (2017). Pointnet: Deep Learning on Point Sets for 3D Classification and Segmentation, *The IEEE Conference on Computer Vision and Pattern Recognition*, July. 21-26, Honolulu, USA, pp. 652-660.

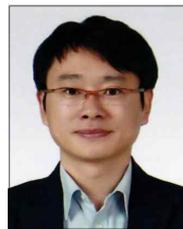
Wang, N., Zhang, Y., Li, Z., Fu, Y., Liu, W., Jiang, Yu-Gang. (2018). Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images, *The European Conference on Computer Vision*, Sep. 8-14, Munich, Germany, pp. 52-67.

Zuffi, S., Black, MJ. (2015). The Stitched Puppet: A Graphical Model of 3D Human Shape and Pose, *The IEEE Conference on Computer Vision and Pattern Recognition*, June. 7-12, Boston, USA, pp. 3537-3546.



김 대 희 (DaeHee Kim)

- 한밭대학교 제어계측공학과 공학학사
- 한밭대학교 제어계측공학과 공학석사
- (현재) (주)컨택 연구원
- 관심분야: 인공지능, 컴퓨터 비전 및 그래픽스, 형상 변형, 객체 변형



황 본 우 (Bon-Woo Hwang)

- 성균관대학교 전자공학과 학사
- 성균관대학교 전자공학과 석사
- 고려대학교 컴퓨터학과 박사
- 미국 카네기멜론대학교 로봇공학연구소 박사후과정
- (현재) 한국전자통신연구원 책임연구원
- 관심분야: 3D 프린팅, 3D 얼굴 복원, 3D 객체 변형, 컴퓨터 비전 및 그래픽스, 인공지능



이 승 욱 (SeungWook Lee)

- 홍익대학교 전과공학과 공학사
- 광주과학기술원 정보통신공학과 공학석사
- 한국전자통신연구원 콘텐츠 연구본부 책임연구원
- 관심분야: 딥러닝, CG/Vision, 표준화



곽 수 영 (Sooyeong Kwak)

- 정회원
- 연세대학교 컴퓨터과학과 공학박사
- 삼성전자 영상디스플레이사업부 책임 연구원
- (현재) 한밭대학교 전자제어공학과 교수
- 관심분야: 컴퓨터비전, 영상 분석, 딥러닝