

스마트폰 사진 합성을 통한 3D 아바타 모델링

한제완, 이창규, 송인석, 남재우, 권기학, 문현준
세종대학교 소프트웨어융합대학 컴퓨터공학과
e-mail: hmoon@sejong.edu

3D Avatar Modeling through Composite Photograph for Smartphone Environment

Je-Wan Han, Chang-Gyu Lee, In-Seok Song, Jae-Woo Nam,
Gi-Hak Kwon, Hyeonjoon Moon
Department of Computer Science and Engineering , Sejong University

요 약

현대 사회의 발전으로 인해 사람들의 삶의 질이 향상됨에 따라 사람들은 다양한 방식으로 자신 및 자신의 개성을 표출하려는 시도를 한다. 특히 IT 기술의 발전은 가상현실 및 3D 기술의 성장을 이끌어냈다. 본 논문은 다가올 4차 산업혁명에 발맞추어 사용자의 개성을 표출할 실용적이고 개성 있는 3D 모델링 아이디어를 제안하고자 한다. 스마트폰 사진 촬영과 동시에 사용자가 선택한 다른 캐릭터 사진과의 합성 사진을 Convolutional Neural Network (CNN)과 Generative Adversarial Network (GAN) 기반 딥러닝 기술을 통해 생성한다. 생성된 이미지는 사용자의 모습과 합성의 대상이 되는 캐릭터의 모습을 동시에 담고 있다. 본 연구의 결과물로 생성된 합성 사진을 3D 프린터를 이용하여 자신만의 모습이 담긴 굿즈를 생산 혹은 이모티콘을 생성하는 등 다양한 실용적인 응용분야에 적용 가능하다.

1. 서론

정보화시대가 열리면서 다양한 방식을 통해 자신의 아바타를 만들려는 시도가 있었다. 특히 VR과 게임 산업에서 이러한 노력이 더욱 두드러진다. 게임 회사 닌텐도의 경우 이러한 아바타화에 대한 사람들의 열망을 깨닫고 2000년대부터 Mii라는 아바타 콘텐츠를 만들었다. 많은 선택 옵션을 주어 온라인상에서 사용할 자신의 3D 아바타를 꾸밀 수 있는 콘텐츠였다. 지금은 더욱 발전하여 많은 닌텐도 게임들과 연계되고 있다. 2018년 EA의 농구 게임인 NBA 2K18에서는 스마트폰 앱을 연동하여 자신의 얼굴을 그대로 게임 캐릭터로 만드는 콘텐츠를 제공했다. 이외에도 가상현실 게임을 비롯한 많은 게임에서 플레이어의 모습을 게임에 더 잘 투영하기 위한 기술들이 연구 중이다.

4차 산업사회에 가까워질수록 게임뿐이 아닌 더욱 많은 분야에서 3D 모델링 및 3D 아바타 기술은 중요해질 것이다. 따라서 더욱 보편적이고 범용적으로 사용할 수 있는 3D 아바타 기술 및 체계가 필요하다. 본 논문이 제안하고자 하는 것은 스마트폰을 활용해 한 장의 정면 사진만을 가지고 손쉽게 3D 모델링 및 3D 아바타, 캐릭터와 콘텐츠를 제공하는 것이다. 여기서 그저 얼굴만 가지고 아바타를 만드는 것은 아바타화 기술로서 효용성이 떨어지기 때문에, 합성 기술을 이용해 실제 인물이나 다른 캐릭터와의 얼굴 특징과 결합된 이미지를 창조해서 3D 모델링을 진행할 것이다. 또한 모델링에서 그치는 것이 아니라 사용자가 3D 프린터만 있으면 손쉽게 스마트폰을 연동해서 캐릭터 상품을 제작해낼 수 있는 유기적인 시스템을 제안한다.

2. 관련 연구

3D 아바타 모델링에 관한 연구는 이미지 한 장으로 얼굴을 3D 모델링 한다. 이 연구는 Aaron Jackson, a PhD student from the University of Nottingham에 의해 진행되어 완성되었다. 이 기술은 딥러닝을 이용하여 데이터 셋을 통해 얼굴의 특징들을 학습시킨다. Input image가 들어오면 학습된 알고리즘으로 image에 있는 얼굴의 특징들과 깊이를 검출한다. 3D 평면에 검출된 특징 값과 깊이에 따라 얼굴의 각 부위마다 z축이 다르게 설정된다. 이 연구에 쓰인 데이터 셋은 6만장이 넘는다.



그림 1 Mii 아바타 콘텐츠와 NBA 2K18 아바타



그림 3 VRN을 이용한 2D to 3D Modeling

3. 기술 개요

본 논문이 제안하는 시스템은 크게 총 4가지 기능으로 나뉘게 된다.

1. 인텔에서 제공하는 오픈 소스 컴퓨터 비전 라이브러리인 OpenCV를 활용한 머신러닝 기술을 활용한다. 스마트폰으로 촬영된 사진 및 기존에 존재하는 사진에서 얼굴을 검출한 후, 검출된 얼굴을 토대로 아바타화를 진행한다.
2. 사용자의 모습을 아바타 화하는 것뿐만 아니라 사용자의 개성이 있는 다양한 캐릭터의 모습을 유기적으로 합성한다. 합성에는 CNN 및 GAN을 기반으로 한 딥러닝 기술을 활용한다.
3. 합성된 사진을 바탕으로 3D 모델링을 진행한다. 이때 CNN을 통해 학습된 데이터를 바탕으로 얼굴 굴곡 및 특징에 대한 수치를 측정하여 3D화 한다.
4. 완성된 3D 모델을 안드로이드 웹 애플리케이션을 활용해 3D 프린터에 3D 아바타 모델 데이터를 전송하고 출력한다.

4. 얼굴 검출

아바타화를 위해 사진에서 얼굴을 찾고 얼굴을 추출해야 한다. 본 논문에서 사용하고자 하는 기술은 LBP 기반 검출기를 통한 얼굴 검출 기술이다. LBP 기반 얼굴 검출은 잠재적으로 검출 속도가 가장 빠르고 인공지능 학습 시간도 가장 빠르기 때문에 스마트폰 어플에서 적용하기에 알맞은 선택이다.

먼저, 얼굴 검출을 위해 LBP 검출기가 사람의 얼굴을 검출하도록 학습시킨다. 이후 얼굴을 검출하고 싶은 이미지를 입력받는다. 전처리 과정을 통해 얼굴을 검출할 수 있는 형태의 이미지로 변환하고, 학습된 검출기를 통해 얼굴 요소들의 특징을 찾아 얼굴의 위치를 탐색한다. 이어서 전처리 과정을 통해 합성에 필요한 얼굴 이미지만 수집한다.

5. 얼굴 합성을 통한 아바타화

수집된 얼굴 이미지를 세팅되어 있는 특정한 대상의 얼굴 특징과 합성한다. 이때 합성할 대상의 종류에 따라 각각 고유하게 생성된 학습데이터를 지닌다. 이러한 학습과 합성을 실행하는 기술로 적대적 생성신경망인 GAN을 활용한다. GAN은 생성자와 구분자라는 두 개의 신경망을 경쟁적으로 학습시켜 새로운 무언가를 만들어내는 생성 모델이다. GAN은 진짜같은 가짜를 만들어내는데 특화된 알고리즘이다. 특히 이미지 분야에서 뚜렷한 성과를 내고 있기 때문에 개성있는 아바타를 창조하는데 적합한 기술이다. 기본 골조는 생성자가 가짜 이미지를 만들어내면 구분자는 그것이 진짜인지 가짜인지 구분하고, 이 데이터를 토대로 생성자는 구분자가 최대한 구분할 수 없는 가짜 이미지를 다시 창조해내는 구조이다. 따라서 생성자는 진짜같은 가짜를 만들 수 있도록, 구분자는 진짜와 가짜를 더욱 잘 구별할 수 있게 서로 번갈아가면서 학습해나간다. 최종적으로는 생성자에 의해 진짜라고 해도 손색없는 가짜가 만들어지게 된다.

알고리즘 측면에서 살펴보자면 랜덤한 벡터 z 를 입력받아 하나의 위조품을 출력하는 생성자 $G(z)$ 와 위조품을 입력받아 그 위조품이 진짜인지 가짜인지 1과 0으로 판별해내는 구분자 $D(x)$ 가 존재한다. 구분자에 경우 들어오는 입력 값이 실제 데이터 분포에서 뽑아낸 샘플일 경우 1이 나와야 하고, 랜덤한 z 값에 의해 만들어진 생성자의 위조품일 경우 0이 나오게 학습이 진행된다. 반대로 생성자는 더욱 정밀하게 구분할 수 있게 진화되어가는 구분자를 속이기 위해 학습해나간다. 수식으로 표현하면 아래와 같다.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

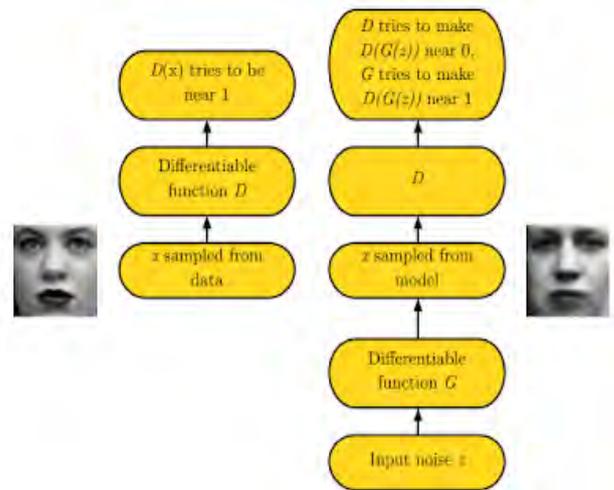


그림 2 GAN학습원리

추출한 얼굴 이미지와 합성하기 위해 각각의 대상마다 고유의 학습된 GAN 데이터 및 알고리즘이 필요하다. 특

유의 레이블을 지정해서 GAN을 통해 학습하고 이를 통해 완성된 생성기에 사용자의 검출된 얼굴 이미지를 입력 받아 개성있는 이미지를 만들어 낸다.

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학 사업의 연구결과로 수행되었음 (2015-0-00938)

6. 3D 모델링

개성있게 변형된 사용자의 얼굴 이미지를 3D 모델링한다. 본 연구에서 제안된 방법으로 CNN을 이용한 회귀 모델을 사용하고자 한다. 회귀 모델은 신경망을 통한 학습을 거듭하여 원하는 값을 예측하는 모델이다. 선형회귀는 최대한 실제 값의 특징들을 최대한 예측할 수 있는 하나의 선형방정식을 찾아내기 위한 학습을 시행한다.

사진에서 얼굴의 입체 데이터를 추출하기 위해 신경망을 학습한다. 학습된 신경망에 완성된 아바타 이미지를 입력한다. 선형회귀 모델을 통해 학습된 신경망은 이미지에서 얼굴의 입체적 데이터를 정밀한 예측을 통해 추출해낸다. 이를 기반으로 2D 기반의 아바타 이미지를 3D 모델로 변환한다. 더욱 효과적인 3D화를 위해 GAN 기술을 사용해보는 것도 시도해볼만 하다.

7. 결론 및 활용 방안

CNN 및 GAN 기술을 활용하여 사용자의 모습과 세팅되어 있는 특정한 캐릭터의 모습을 담고 있는 합성 사진을 생성해 낼 수 있다. 생성된 합성 사진의 활용에 대한 2가지 방안을 제시하고자 한다.

첫 번째는 3D 프린팅이다. 3D 프린팅 기술은 날로 발전하고 있으며 실물과 다른 것을 찾아보기 힘들 정도의 정교함을 자랑한다. 따라서 제시된 기술을 활용해 사용자만의 이미지가 담겨져 있는 굿즈를 제작하고 이를 활용하는 방안을 제시한다.

두 번째는 자신의 모습의 이모티콘이다. 이미 자신의 모습으로한 이모티콘 기능은 상용화가 되어 있는 상태이다. 하지만 제시된 기술로 생성된 합성 3D 이미지 모델을 활용한 이모티콘은 이와는 다르게 자신의 모습뿐만 아니라 다양한 캐릭터의 모습을 담고 있기 때문에 특수성을 보이며 자신만의 개성을 표출하는데 큰 역할을 할 수 있을 것으로 예상된다.

인공지능 및 딥러닝 기술의 발전에 따른 다양한 합성 사진이 생성되고 있다. 이렇게 생성된 합성 사진은 실물과 상당히 유사하며 이를 이용한 악성 행위들도 많이 행해지고 있는 것이 현실이다. 반면 제시된 기술 및 방안을 활용하여 생성된 이미지는 사용자의 개성을 표출, 굿즈라는 상품을 제작하는 등의 생산성 있는 긍정적인 효과를 창출해낸다. 이와 같은 합성 기술에 대한 긍정적인 면을 활용해 악성 행위를 근절하는 등의 연구를 시도하는 것도 유의미할 것이다.

참고문헌

- [1] Roger S. Pressman "Software Engineering A Practitiners' Approach" 3rd Ed. McGraw Hill
- [2] <https://github.com/AaronJackson/vrn>
- [3] <http://cvi-demos.cs.nott.ac.uk/vm/view.php?name=../59b1bd503517f>
- [4] <https://fosbytes.com/computer-scientists-finally-developed-3d-selfies/>
- [5] Generative Adversarial Nets - Ian Goodfellow et al. 2014 논문
- [6] NIPS 2016 Tutorial: Generative Adversarial Networks [7].<http://jaejunyoo.blogspot.com/2017/01/generative-adversarial-nets-1.html?m=1>
- [8] OpenCV 컴퓨터 비전 프로젝트, 저자 다니엘 렐리스 바지우 외 6명