

3D 변환을 활용한 NFT 기반 메타버스 패션 컨텐츠 플랫폼

김민호¹, 한수한², 박민규¹, 정동주³, 이병정²

¹서울시립대학교 전자전기컴퓨터공학부

²서울시립대학교 컴퓨터과학부

³(주)스마트 잭

alsgh8257@naver.com, hshhan1009@naver.com, pencake33@naver.com,
lostcode7@gmail.com, bjlee@uos.ac.kr

Towards a NFT-based Metaverse Fashion Contents Platform using 3D Conversion

Min-Ho Kim¹, Su-Han Han², Min-Gyu Park¹, Dong-Ju Jung³,
Byung-Jeong Lee²

¹Dept. of Electrical and Computer Engineering, University of Seoul

²Dept. of Computer Science and Engineering, University of Seoul

³Smart Jack Co., Ltd.

요 약

본 연구에서는 하나의 2D 이미지를 StyleGAN을 통해 다각도의 이미지를 생성하고, 그것을 다시 Kaolin으로 구현한 역그래픽 렌더러의 입력으로 받아 3D 오브젝트로 변환한다. 또한, 클레이튼 기반의 블록체인을 통해 NFT 기술을 통하여 3D 오브젝트를 NFT로 만들 수 있도록 한다. 최종적으로 2D 이미지를 메타버스에서 활용할 수 있는 3D 패션 아이템으로 변환하고 NFT를 발행하여 거래한다. 본 연구는 개인이 자유롭게 메타버스 콘텐츠를 제공하고 거래하여 메타버스 활성화에 기여할 것으로 기대한다.

1. 서론

메타버스에 대한 사회적 관심과 가상 디지털 콘텐츠 소비문화가 확산하면서, 사람들은 메타버스 세계에서 또 하나의 나를 만들고자 하는 추세다. 또한, NFT를 통해 나만의 특색있는 아이템을 소유하고자 하는 사람들이 늘어나고 있다. 이러한 추세를 반영하여 본 연구에서는 NFT를 활용하여 자신만의 캐릭터를 개성 있게 꾸밀 수 있는 플랫폼을 제공한다. 본 연구의 주요 목적은 자신의 2D 이미지 파일을 메타버스에서 사용할 수 있는 3D 패션 아이템으로 변환하여 나만의 패션 아이템을 만들 수 있도록 하는 것이다. 이렇게 만든 아바타 패션 아이템을 NFT로 만들어 소유자의 소유권을 증명하며, 또한 암호화폐를 통해 거래하는 환경을 구축한다.

2. 관련 연구

2.1 StyleGAN

PGGAN(Progressive Growing of GANs)은 생성기(generator)와 판별기(discriminator)의 해상도를 점진적으로 늘려 고해상도의 이미지를 생성하는 방

법이다[1]. 이를 바탕으로, 생성기의 각 계층에 스타일 정보를 입힘으로써 합성되는 이미지의 특징정보를 조절할 수 있는 StyleGAN이 제안되었다[2]. PGGAN은 입력 계층이 직접 합성곱(convolution)과 업샘플링(upsampling)을 거쳐 이미지로 변환된다. 이때, 입력 계층이 이미지의 특징정보와 비선형(non-linear)하게 매핑되어서 특징정보들이 서로 얽히게 된다. 따라서 입력의 작은 변화가 여러 특징에 큰 영향을 미치므로 출력 이미지의 특징정보를 조절하는 데 어려움이 있었다. 하지만 StyleGAN은 입력 계층이 8개의 완전연결계층(fully-connected layer)으로 구성된 매핑 네트워크(mapping network)를 거쳐 중간 벡터를 생성한 후, 이를 이미지의 스타일 정보로 사용하여 합성 네트워크에서 이미지로 변환된다. 매핑 네트워크를 통해 생성된 중간 벡터는 학습 데이터의 분포를 따를 필요가 없어 각 특징정보 간의 상관관계를 줄일 수 있다.

합성 네트워크의 각 단계에서 같은 중간 벡터를 이용하기 때문에 각 단계가 상관관계를 가질 수 있다. 이 상관관계를 줄이기 위해 모델은 두 개의 입

력 계층을 선택하여 합성 네트워크의 특정 지점부터 스타일에 사용하는 중간 벡터를 바꾸는 방식을 사용한다. 중간 벡터를 교체하는 지점을 매번 무작위로 선택함으로써 연속된 두 단계 사이의 상관관계를 줄일 수 있다.

2.2 Kaolin

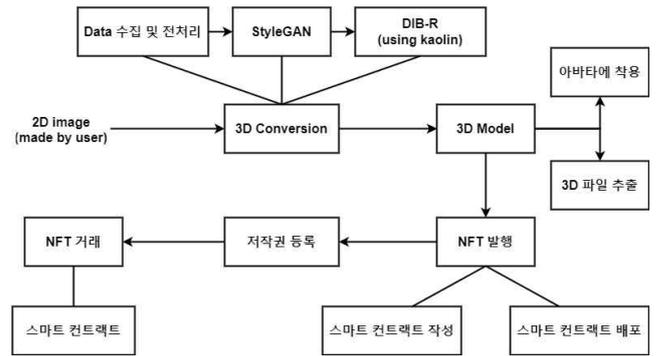
2차원 입력 영상과 그에 해당하는 깊이 정보를 이용하여 새로운 시점의 영상을 만들어내는 방식인 DIB-R이 제안되었다[3]. DIB-R은 이미지의 모든 픽셀에 대한 경사(gradient)를 계산할 수 있다. 이를 통하여 다양한 조명 모델을 통해 이미지의 정점(vertex) 위치와 색, 정규화된 값, 조명 방향 및 질감 정보에 대한 정확한 최적화를 수행한다. 새로운 피사체에 대한 3D 변환 학습은 오랜 시간이 걸리지만 한 번 학습이 완료된 후에는 100ms 이내에 3D 오브젝트를 출력할 수 있다.

NVIDIA는 DIB-R을 바탕으로 2D 영상을 3D 영상으로 변환할 수 있는 가속화 도구인 Kaolin을 개발했다[4]. 먼저 2D 이미지만을 사용하여 3D 메쉬(mesh)를 예측한다. RGBA 이미지를 RGB값인 I와 알파 값인 S로 나누어 I와 S에 대한 예측을 각각 진행한다. 이를 통해 S에 대한 예측인 실루엣 예측과 I에 대한 예측인 색상 예측에 대한 손실을 구분할 수 있다. 예측된 메쉬를 렌더링할 때는 실제 카메라 위치뿐만 아니라 임의의 각도에서 렌더링을 진행한다. 이를 통해 손실을 최소화할 수 있다.

2.3 NFT

본 연구에서는 3세대 블록체인인 클레이튼(Klaytn)을 사용하여 NFT(Non-Fungible Token) 거래를 구현한다. 클레이튼은 KIP-17로 발행되는 토큰을 이용해 거래하고 KIP-37를 통해 스마트 계약을 구현한다. KIP-17은 클레이튼의 NFT 표준안으로, 모든 아이템에 대한 고유성이 보장된다. KIP-37는 클레이튼의 스마트 계약에 대한 표준안으로, 서로 다른 코인 간 혼합 거래를 지원하고, 여러 아이템을 하나의 그룹으로 묶어 거래할 수 있는 묶음 토큰을 지원한다. 또한, 하나의 트랜잭션에서 여러 명에게 원하는 만큼의 아이템을 전송하는 다중 전송을 지원한다. 클레이튼은 가스 비용(gas fee)이 적고 처리량과 완결성, 확장성을 보장하기 때문에 개인 창작자들에게 적합하다.

3. 3D 변환을 활용한 NFT 기반 플랫폼



(그림 1) 개념 아키텍처

그림 1은 본 연구의 개념 아키텍처이다. 사용자는 사진, 그림 등의 2D 이미지 파일을 3D 변환을 통해 3D 오브젝트로 얻을 수 있다. 3D 변환은 StyleGAN과 Kaolin을 활용한 DIB-R을 통해 이루어진다. 변환된 3D 오브젝트는 3D 파일로 추출하여 본 플랫폼에서 제공하는 아바타에 착용할 수 있다. 이뿐만 아니라 다양한 메타버스 플랫폼에서 활용될 수 있도록 한다. 3D 오브젝트는 NFT로 발행되어 저작권 등록과 함께 스마트 계약 작성과 배포를 통한 NFT 거래 또한 가능하다.

3.1. 딥러닝 및 역그래픽을 활용한 3D 변환

3D 오브젝트를 생성하기 위해 다각도의 이미지 파일, 카메라 위치 정보, 분할 마스크(segmentation mask)가 필요하다. 본 연구는 다각도 이미지를 얻기 위해 StyleGAN을 사용한다. StyleGAN 합성 네트워크의 계층은 서로 다른 이미지 속성을 제어하는데, 그중 상위 계층을 통해 생성된 잠재 벡터들이 카메라의 시점을 조정한다[5]. 즉, 나머지 계층은 유지하고 상위 계층만 조절한다면 다각도의 이미지를 얻을 수 있게 된다. 만약 학습 데이터 세트에 충분한 다각도 이미지가 있다면 반대편을 포함한 다양한 이미지를 생성할 수 있다. 카메라 위치 정보를 얻기 위해선 객체의 키포인트에 주석(annotation)을 달아야 한다. 이때 위에서 얻은 잠재 벡터를 사용하면 주석다는 시간을 단축할 수 있다. n개의 다각도 이미지를 생성한다면 입력값이 달라도 사용된 잠재 벡터들은 모두 같기에 n개의 이미지에 대해서만 주석을 달아주면 된다. 주석을 단 후 SfM(Structure from Motion)을 사용하여 카메라 매개변수를 계산하면 카메라의 위치 정보를 얻을 수 있다. StyleGAN으로 생성된 이미지에는 우리가 목표로 하는 객체 이외에

불필요한 요소(배경, 다른 객체 등)들이 존재할 수 있고 이는 역그래픽 렌더러를 훈련하는 데에 방해요소가 된다. 그래서 MaskR-CNN으로 인스턴스 분할을 하여 불필요한 이미지를 분류한다.

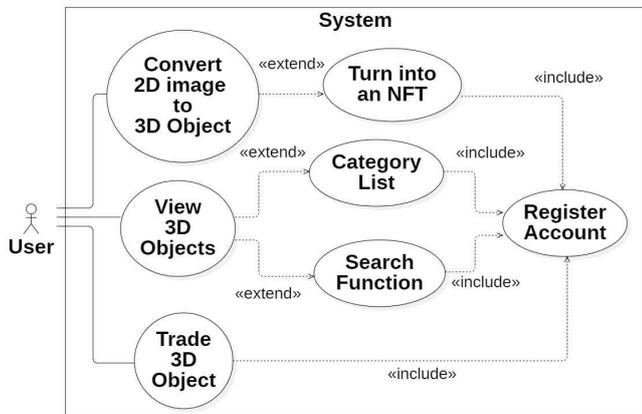
생성된 다각도 이미지, 카메라 위치 정보, 분할 마스크들은 역그래픽 렌더러(본 연구에선 DIB-R)의 입력이 된다. DIB-R에 사용되는 초기 3D 오브젝트는 구 형태에서 시작하여 학습할수록 우리가 원하는 3D 객체 형태를 띠게 된다. Kaolin을 통해 입력받은 카메라 위치에서 3D 객체를 바라봤을 때의 이미지와 마스크를 얻을 수 있고 이를 입력받은 실제 이미지, 마스크와 비교하여 손실을 계산한다. 이 손실을 생성된 3D 객체 형태를 통해 정규화한다. DIB-R은 손실을 줄일 수 있도록 학습을 진행하며 결과적으로 2D 이미지의 특징정보를 반영한 3D 오브젝트를 생성하게 된다.

3.2. NFT 발행 및 거래

NFT 발행 및 거래를 하기 위해 스마트 컨트랙트를 사용한다. 작품의 저작권에 관한 내용을 스마트 컨트랙트 전용 언어(Solidity 등)로 작성 후 컴파일하여 바이트코드(bytecode)를 생성한다. 이 바이트코드가 담긴 블록을 블록체인에 배포하여 NFT를 발행한다. NFT 거래는 NFT 발행 시 배포된 스마트 컨트랙트를 통해 수행된다. 스마트 컨트랙트에 작성된 거래 조건이 충족되면 계약이 체결되어 가상화폐가 블록체인 상으로 전송되고 거래된 NFT는 구매자 소유로 변환된다.

4. 애플리케이션 개발

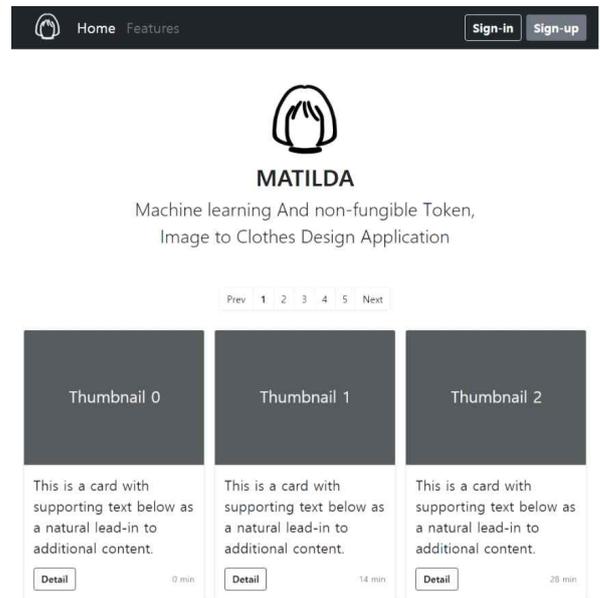
4.1 유스케이스도



(그림 2) 유스케이스도

그림 2는 본 연구에서 구현하는 애플리케이션의 유스케이스도이다. 애플리케이션은 사용자의 2D 이미지를 3D 오브젝트로 변환해 준다. 변환된 3D 오브젝트는 NFT로 웹사이트에 등록할 수 있다. 사용자는 웹사이트에 등록된 3D 오브젝트들을 볼 수 있다. 이때 카테고리 목록 또는 검색 기능을 이용해 원하는 3D 오브젝트를 확인할 수 있다. 마지막으로, 원하는 3D 오브젝트를 스마트 컨트랙트를 통해 거래할 수 있다.

4.2 UI 시나리오



(그림 3) 웹 페이지 UI

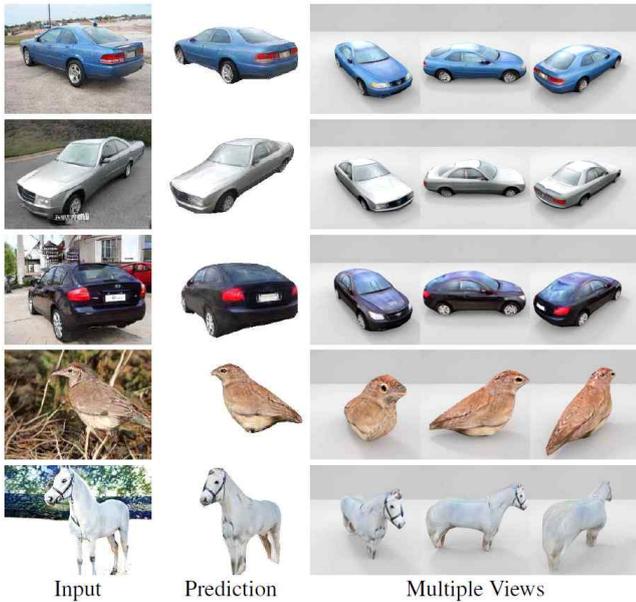
그림 3과 같이 웹사이트를 구축해 앱으로도 애플리케이션의 접근을 쉽게 한다. 시작 페이지에서는 다른 사람의 NFT를 개략적으로 볼 수 있게 구현되어 있고, 해당 카드를 클릭함으로써 NFT에 대한 자세한 정보를 확인할 수 있고, 원한다면 NFT를 구매할 수 있다. 회원가입과 로그인 기능을 통해 Sign-in이 되어 있다면 My Page를 통해 페이지에서 자신이 가지고 있는 NFT를 확인할 수 있다. 또한, 웹페이지에서 2D 이미지를 이용하여 3D 오브젝트를 만들어낼 수 있다. 마지막으로 변환을 통해 만들어진 3D NFT를 판매할 수 있도록 한다.

4.3 데이터 처리 순서

2D 데이터를 StyleGAN과 DIB-R 기술을 이용해 3D 오브젝트를 얻는다. 먼저, 구글링을 통해 학습과 테스트에 적합한 의류 자료집(clothing dataset)을

구한다. 자료집합을 StyleGAN으로 학습시키면 다양한 각도에서 바라본 2D 이미지 형태의 중간 출력물을 여러 개 만들어낼 수 있다. 중간 출력물을 DIB-R로 학습시키면 적절한 형태의 3D 오브젝트를 얻을 수 있다. 테스트 작업을 통해 원하는 결과물이 나올 때까지 학습시킨다. 학습이 끝난 StyleGAN과 DIB-R을 하나의 모듈로 묶는다.

4.4 구현



(그림 4) 3D 변환 예시

2D 이미지를 3D 오브젝트로 바꾸는 기술을 구현한다. 사용자는 자신이 가진 2D 이미지 파일을 어떤 패션 아이템으로 만들지 분류를 정하고 파일을 올린다. 애플리케이션은 그림 4와 같이 해당 이미지 파일을 StyleGAN과 DIB-R 기술을 이용하여 3D 오브젝트를 만든다. 3D 오브젝트를 아바타에 입혀 웹사이트에서 보여줄 수 있도록 한다. 또한, NFT를 거래하는 방법을 구현한다. 카카오 클레이튼 플랫폼에서 제공하는 블록체인을 기반으로 스마트 컨트랙트 기술을 이용해 다른 사람과 NFT를 거래하는 사이트를 구현한다.

5. 결론

본 연구에서는 StyleGAN과 Kaolin을 통해 2D 이미지를 3D 오브젝트로 변환하여 NFT로 발행하고 스마트 컨트랙트를 활용한 NFT 거래까지 구현한 애플리케이션을 제시하였다. 기존의 NFT 거래 플랫폼은 주로 2D 이미지나 영상 등에 집중되어 있었고,

메타버스 콘텐츠 또한 큰 규모의 회사에서 지엽적으로 제공하고 있었다. 하지만 본 연구를 통해 개인이 더욱 자유롭게 메타버스 콘텐츠를 제공하고 거래하여 메타버스 시장 활성화에 기여할 것으로 기대한다.

사사의 글

본 논문은 과학기술정보통신부 정보통신창의인재양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다(22_HF280).

참고 문헌

[1] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive Growing of GANs for Improved Quality, Stability, and Variation", arXiv:1710.10196v3, 2018.

[2] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks", arXiv:1812.04948v3, 2019.

[3] W. Chen, J. Gao, H. Ling, E.J. Smith, J. Lehtinen, A. Jacobson, and S. Fidler, "Learning to Predict 3D Objects with an Interpolation-based Differentiable Renderer", arXiv:1908.01210v2, 2019.

[4] K. M. Jatavallabhula, E. Smith, J. Lafleche, C. Tsang, A. Rozantsev, W. Chen, T. Xiang, R. Lebedian, and S. Fidler, "Kaolin: A Pytorch Library for Accelerating 3D Deep Learning Research", arXiv:1911.05063, 2019.

[5] Y. Zhang, W. Chen, H. Ling, J. Gao, Y. Zhang, A. Torralba, and S. Fidler. "Image GANs meet Differentiable Rendering for Inverse Graphics and Interpretable 3D Neural Rendering". arXiv:2010.09125v2, 2021.