

## 저비용 3D 모델링을 위한 Instant-NGP 기반 파이프라인 설계 및 구현

김윤녕, 최주환

한국전자기술연구원 전북지역본부 IT응용연구센터

{kyn0217, netside}@keti.re.kr

Design and Implementation of Instant-NGP-Based Pipeline  
for Cost-Effective 3D Modeling

Youn Nyeong Kim, Ju Hwan Choi

Korea Electronics Technology Institute (KETI)

## 요약

최근 3D 모델링은 고가의 장비와 3D 설계 툴을 사용하여 직접 모델을 생성하는 추세지만, 일반 사용자가 이러한 장비를 갖추거나 툴을 다루기에는 비용적, 기술적 부담이 존재한다. 본 논문에서는 저비용으로 별도 장비 없이 일반 사용자 3D 모델을 구현할 수 있는 파이프라인을 제안한다. 제안된 파이프라인은 일반 스마트폰 카메라로 대상을 중심으로 원 모양으로 촬영하고 ffmpeg를 이용하여 프레임 추출한다. 추출된 프레임마다 COLMAP으로 카메라 파라미터를 생성한 뒤, Instant-NGP 학습을 통해 생성된 3D 모델을 구현한다. 마지막으로 3D 모델을 Mesh 변환 과정을 진행하여 3D 객체를 구현하였다. 제안한 파이프라인은 새로운 대상뿐 아닌 기존 시제품을 3D 모델로 만드는 것도 가능하며, 본 연구에서는 침수 시뮬레이션 모형을 대상으로 실험을 진행했다. 실험 결과 제안한 파이프라인을 통해 대상과 유사한 3D 모델을 얻었으며, Mesh 변환 과정의 한계와 향후 개선 방향을 제시한다.

## I. 서론

최근 3D 모델링은 제품 설계, 환경 분석, 시뮬레이션 등의 분야에서 중요한 역할을 하고 있다[1]. 특히 기후 변화로 인한 홍수 및 침수 재난의 발생 빈도가 높아지면서, 이를 효과적으로 예측 및 대응하기 위한 침수 재난 모델링 및 재난 대비 시뮬레이션의 필요성이 커지고 있다. 그러나 기존 3D 모델링은 고가의 센서나 LiDAR와 같은 촬영 장비를 사용하거나 3D 모델을 직접 설계툴을 사용하여 3D 모델을 직접 그려야 하는 한계가 존재한다.

본 논문에서는 재난 예방 제품을 제작하는 일반 사용자를 대상으로 별도의 장비를 사용하지 않고 일반 스마트폰 카메라로 촬영한 영상을 기반으로 오픈소스 툴인 COLMAP[2]과 Instant-NGP를 사용하여 모델을 만드는 저비용 3D 모델링 파이프라인을 제시한다. 이를 통해 누구나 별도의 장비 없이 제품 또는 환경에 대한 3D 모델을 생성하여 재난 대응 및 예방 기술 개발에 활용할 수 있음을 보이고자 한다.

## II. 본론

본 논문에서는 Instant-NGP와 COLMAP, ffmpeg를 사용하여 일반 스마트폰 카메라로 침수 실험 모형을 촬영한 영상을 가지고 3D 모델링을 진행하는 과정에 대해서 상세하게 설명한다. 그림 1은 3D 모델 구현 파이프라인의 흐름도를 나타낸다.

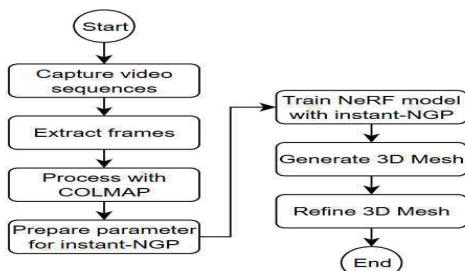


그림 1. 3D 모델 구현 파이프라인 흐름도

3D 모델 구현 파이프라인의 흐름은 모형을 만들 대상으로 잡은 대상을 일반 스마트폰 카메라로 촬영하는 것으로 시작한다. 촬영은 촬영 대상을 카메라 화면에 담을 수 있을 정도의 거리를 기준으로 한다. 촬영 방식은 대상을 기준으로 천천히 1초 내외로 정지하고 대상을 중심으로 하여 원을 그리듯이 이동하며 영상을 촬영한다. 촬영된 영상은 ffmpeg와 같은 오픈소스 영상 편집 툴을 사용해서 동영상상을 간단하게 모델을 학습하기 위한 1초당 2프레임부터 그 이상의 프레임으로 추출한다. ffmpeg는 촬영한 영상을 초당 프레임으로 추출하기 위해서 사용한 오픈소스 툴로 이와 유사한 기능을 하는 툴이 있다면 다른 툴을 사용해서 프레임을 추출해도 문제없이 진행할 수 있다. 본 실험에서는 안정적인 학습을 하기 위해서 초당 5프레임으로 추출을 진행했다. 추출한 프레임은 단순 이미지로만 구성되어 있어서 instant-NGP로 학습시키고 이로부터 나온 결과를 만들 수가 없다. 이를 해결하는 방법은 COLMAP을 사용하여 Instant-NGP에서 학습시킬 수 있도록 데이터를 생성하는 것이다.

COLMAP은 오픈소스 SfM(Structure-from-Motion) 및 MVS(Multi-View Stereo) 소프트웨어로 이미지로부터 카메라 포즈 계산, 카메라 내부 파라미터 산출, 포인트 클라우드 생성의 기능을 가진다. 이를 이용하면 프레임 이미지에 카메라 포즈, 초점 거리, 왜곡 계수 등의 학습에 필요한 데이터를 생성하여 프레임마다 정리한 파라미터 파일을 생성하여 Instant-NGP를 사용할 수 있게 해준다. COLMAP을 통해 추출한 프레임을 대상으로 이미지 간의 특징점을 찾기 위한 매칭 단계와 맵핑 단계를 거쳐서 프레임에 대한 파라미터를 정리한 transform.json 파일을 생성할 수 있다.

생성한 파일을 가지고 학습을 통해 NeRF(Neural Radiance Fields) 모델 중 하나인 Instant-NGP를 사용하여 3D 모델을 생성한다. 다른 모델이 많이 존재하지만, 그 모델 중 Instant-NGP를 선택한 이유는 적은 양의 데이터로도 NeRF 구현이 가능하며 학습 속도가 빠르다는 장점 때문에 저비용으로 누구나 사용 가능하도록 진행할 수 있어 채택하게 되었다[3].

다만, 생성된 json 파일을 Instant-NGP에서 바로 사용하려고 하면 오류가 발생하기 때문에 Instant-NGP의 요구 포맷으로 생성한 json 파일을 변경해

줘야 한다. 대표적인 변경점은 프레임마다 카메라 화각 정보가 존재하긴 하지만, 루트에 대표 화각이 추가되고 transform\_matrix\_start라는 변환 이전의 원본 행렬을 기록한 값이 추가된다. 그 외 Instant-NGP의 요구 포맷에 맞춰서 필드를 정리해야 한다.

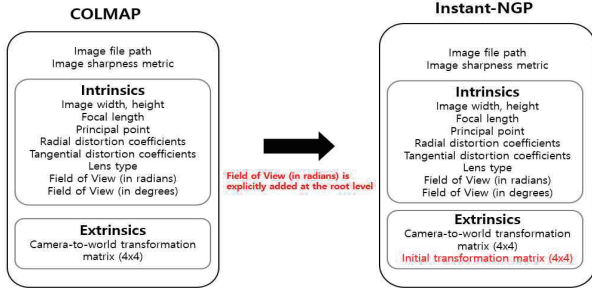


그림 2. Instant-NGP 학습을 위한 데이터 조정 과정

Instant-NGP 학습에 맞추도록 json 파일을 수정 완료하면 Instant-NGP에 학습시킬 프레임 이미지와 파라미터 json 파일을 가지고 학습을 시킬 수 있다. 실험에서는 학습 스텝 수는 모델 구조를 안정적으로 학습 값을 보여준 3만 정도의 수준으로 학습을 진행했다. 그림 3의 왼쪽 그림은 실제 촬영한 대상을 Instant-NGP를 통해서 구현한 것으로 촬영 모델을 실제와 비슷하게 구현된 것을 확인할 수 있었다.



그림 3. Instant-NGP를 통한 모델 구현

구현된 모델을 3D 모델로 만들기 위해 구현된 모델 외의 배경을 지워야 하는데 이는 Instant-NGP 안에서 공간을 특정 영역으로 축소하는 기능을 사용해서 진행한다. 해당 기능을 사용하게 되면 특정 지점을 기준으로부터 x, y, z 값을 축소하여 영역을 한정시키게 된다. 그림 3의 오른쪽 그림은 구현한 모델의 영역을 한정하여 촬영한 모델만 나오도록 진행한 것이다. 모델 근처의 영역만으로 한정시켰다면 그 뒤는 Mesh 파일로 결과를 변경시키는 것으로 Mesh 변환 과정도 Instant-NGP 안에서 지원하기 때문에 이를 가지고 모델을 Mesh 파일로 저장하면 된다.



그림 4. 3D Mesh 출력한 촬영 모델

저장되는 Mesh 파일의 형식은 obj 파일로 지정되며 색상 또한 obj 파일 안에 존재하기 때문에 blender와 같은 3D 그래픽 프로그램에서 확인할 수

있다. 그림 4는 Mesh로 출력한 모델을 blender에서 확인한 것을 보여준다.

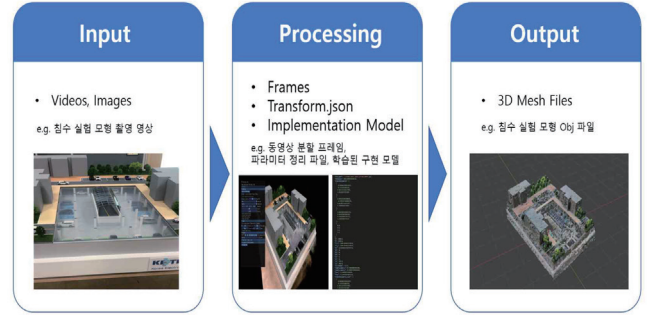


그림 5. Instant-NGP 기반 파이프라인 구현 요약

그림 5는 지금까지 진행한 Instant-NGP 기반 파이프라인에 따른 요약이며 Input부터 Output까지 진행한 것들을 정리한 것이며 이를 통해 모형을 촬영한 것부터 Obj 파일까지 변환하는 것을 확인할 수 있다.

### III. 결론

본 논문에서는 촬영부터 3D 모델링 과정까지의 전체 과정을 제시하고 이를 실제 실험을 통해 타당성을 검증하였다. 실험에서 침수 실험 모형을 대상으로 진행함으로써 복잡한 배수펌프부터 단순한 차수판, 플러드 백과 같은 단순한 제품도 따로 모델링하지 않고 3D 모델로 만들 수 있다는 것도 증명하였다. 더 나아가 시뮬레이션, 환경 분석, 디지털트윈 등에서도 3D 모델을 고가의 장비나 별도의 설계 없이 저비용으로 모델을 생성할 수 있는 것을 확인할 수 있었다. 다만, 학습을 통해 생성된 모델을 3D Mesh로 추출하는 과정에서 표면이 고르지 않게 나오는 부분과 빛이 투과되어 표현되는 부분이 뭉개져서 표현되는 문제가 존재했으며 이를 보정하는 부분에 있어 어려움이 존재한다는 것을 확인했다. 추후 연구에서는 이와 같은 문제를 해결하기 위해 NeRF 모델 변경을 통한 적용 및 Mesh 변환 과정에서 생기는 문제를 연구할 예정이고 Instant-NGP뿐 아닌 Gaussian splatting[4]과 같은 다른 3D 재구성 기술을 구현하여 3D 모델 생성까지의 성능을 비교할 예정이다.

### ACKNOWLEDGMENT

이 연구는 2023년도 행정안전부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구임(재난안전산업진흥시설조성지원, RS-2023-00239835)

### 참 고 문 헌

- [1] Zhou, Linglong, et al. "A comprehensive review of vision-based 3d reconstruction methods." Sensors 24.7 (2024): 2314.
- [2] Schonberger, Johannes L., and Jan-Michael Frahm. "Structure-from-motion revisited." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [3] Müller, Thomas, et al. "Instant neural graphics primitives with a multiresolution hash encoding." ACM transactions on graphics (TOG) 41.4 (2022): 1-15.
- [4] Xu, Zheng, et al. "A survey on surface reconstruction based on 3D Gaussian splatting." PeerJ Computer Science 11 (2025): e3034.