

“

나만의 캐릭터로 만드는 모션 캡처 3D 애니메이션

”

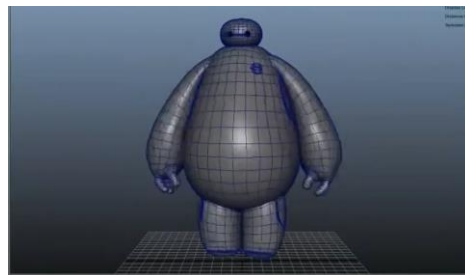
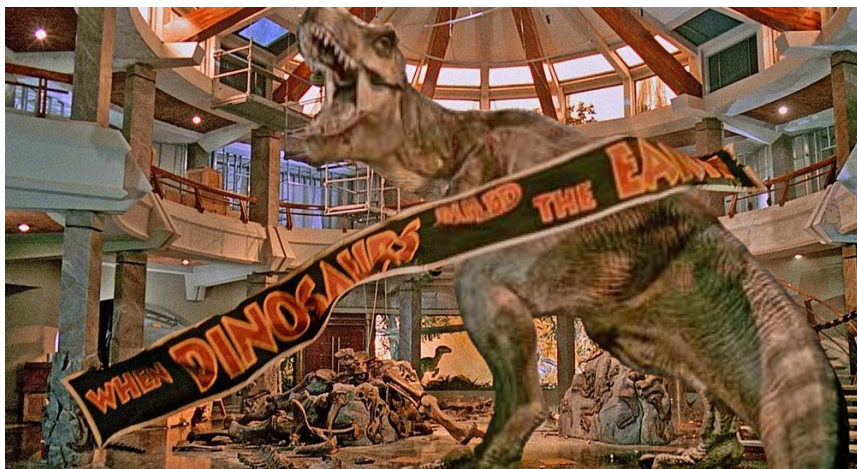
JurassicStudio

조준규, 서희원, 조대호, 최연석, 홍수정

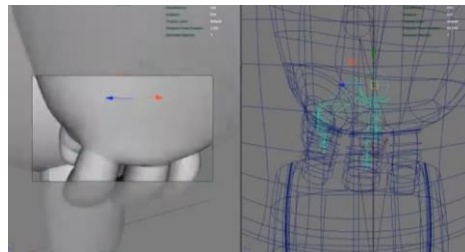
INDEX

- 01** 프로젝트 배경
- 02** 프로젝트 수행 절차 및 팀 구성
- 03** 프로젝트 각 파트 별 모델 및 수행 결과
- 04** 프로젝트 추후 개선 방안

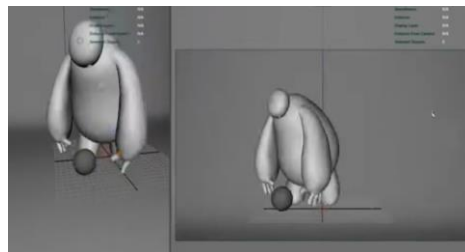
01 프로젝트 배경



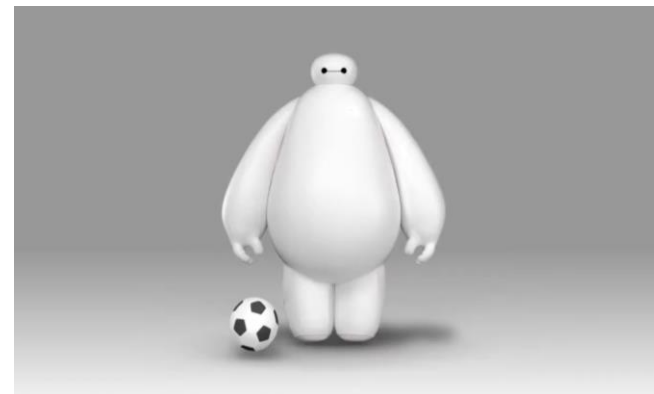
Modeling



Rigging



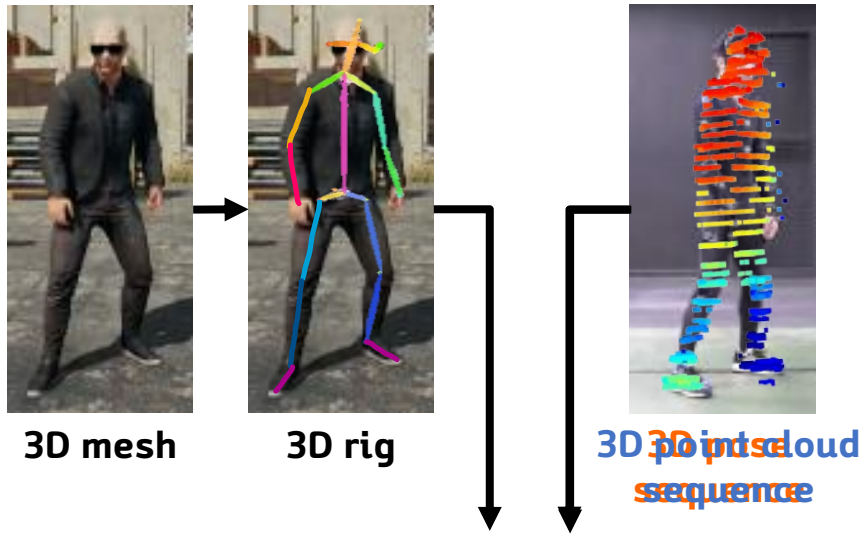
Animation



Rendering



01 프로젝트 배경

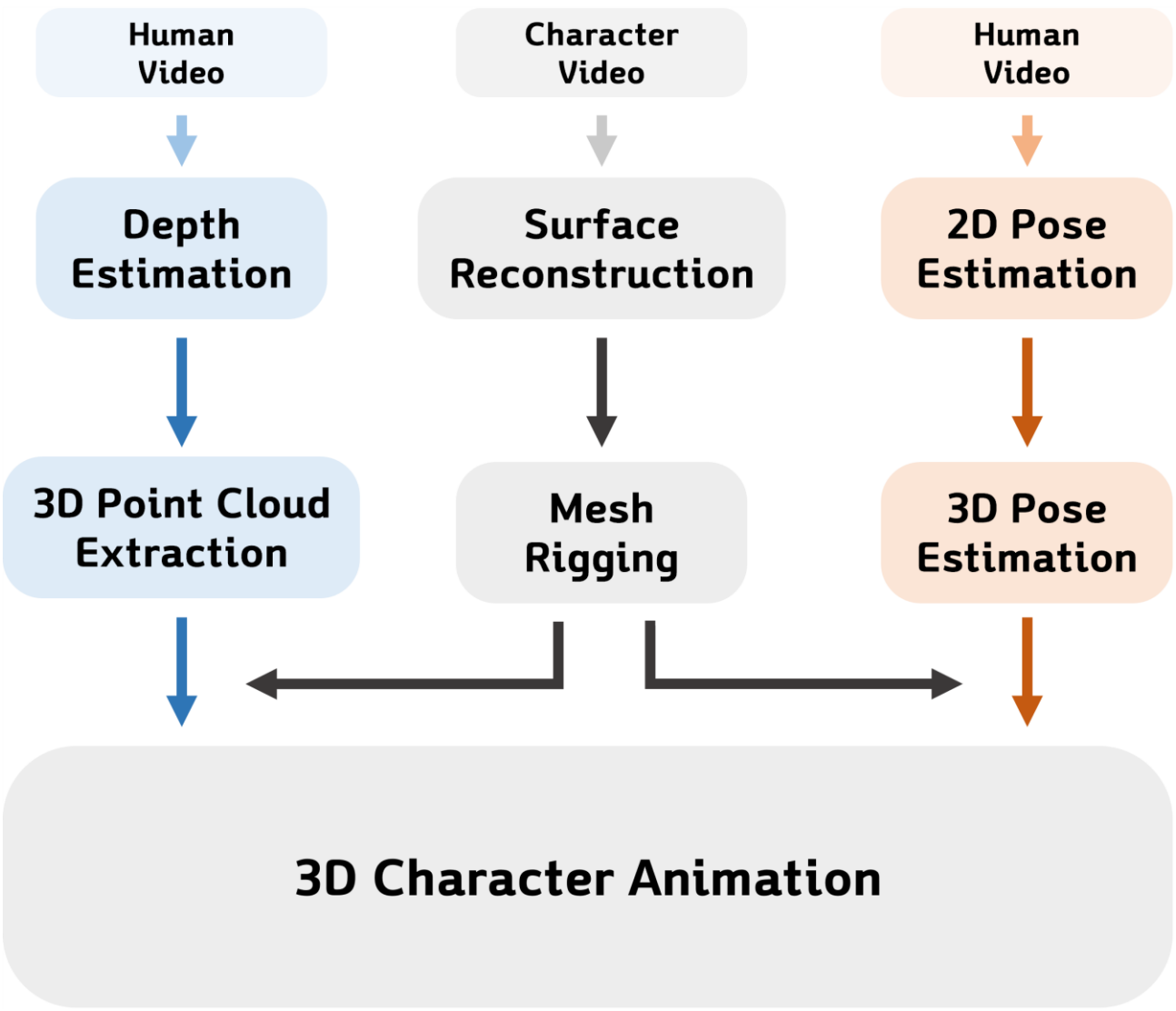


3D animation

- 나만의 캐릭터 동영상으로부터 3D mesh 및 rig 생성
- 나만의 움직임 동영상으로부터 3D point cloud sequence 생성
- 3D mesh와 3D point cloud sequence로 3D animation 생성

02 프로젝트 수행 절차 및 팀 구성

훈련생	역할
조준규 (팀 리더)	<ul style="list-style-type: none">캐릭터 영상으로부터 3D mesh 생성 (Neuralangelo)3D 캐릭터 mesh의 rig 예측(RigNet)
최연석 (팀원)	<ul style="list-style-type: none">캐릭터 영상으로부터 3D mesh 생성 (Neuralangelo)
조대호 (팀원)	<ul style="list-style-type: none">사람 동작 영상으로부터 depth 예측 (MiDaS)Depth map으로 3D point cloud 생성
홍수정 (팀원,서기)	<ul style="list-style-type: none">캐릭터 mesh와 3D point cloud로부터 3D animation 생성(Morig)
서희원 (팀원)	<ul style="list-style-type: none">캐릭터 mesh와 3D human pose로부터 3D animation 생성(MotionBERT)



03 3D mesh & rig extraction

- NeRF(**N**eural **R**adiance **F**ield)
 - CNN을 이용하여 다각도의 2D 이미지로부터 생성되는 3D 가상 공간
 - 동영상을 일정 간격의 frame으로 나누어 2D image set을 형성
 - 2D image set에서 feature를 extract & match 하여 3D로 재구성(Volume Rendering)
 - 재구성된 3D data를 활용해 다른 시점에서 관측될 2D 이미지를 추정하여 Rendering(View Synthesis)

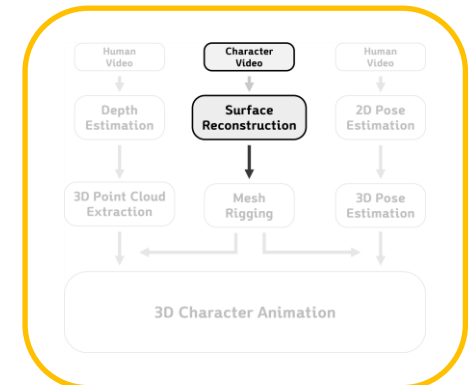
2D Images of
target object

NeRF model

Render 3D Radiance Field



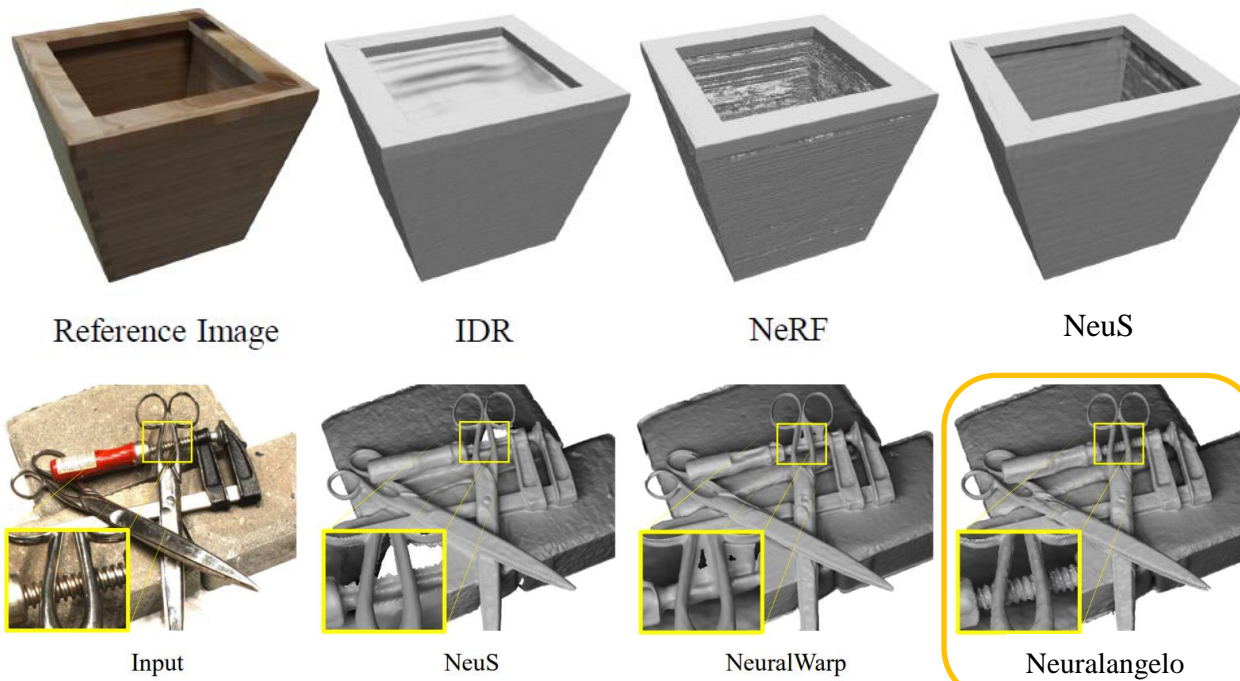
NeRF 3D 생성 과정



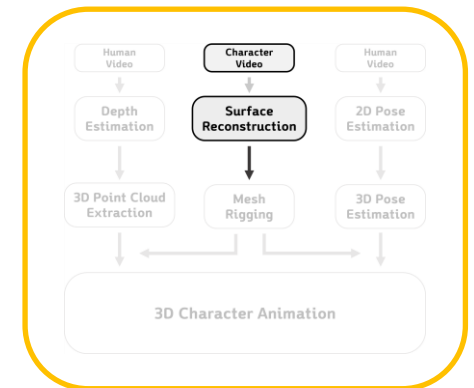
03 3D mesh & rig extraction

- **Neuralangelo(High-Fidelity Neural Surface Reconstruction)**

- NeRF를 surface reconstruction 관점에서 좋은 성능을 나타내도록 발전시킨 모델
- 기존 NeRF로 mesh를 생성하면 noise가 많이 생기는 점을 보완
- 물체 표면의 급격한 변화에도 잘 예측하여 실사와 같은 mesh 추출 가능

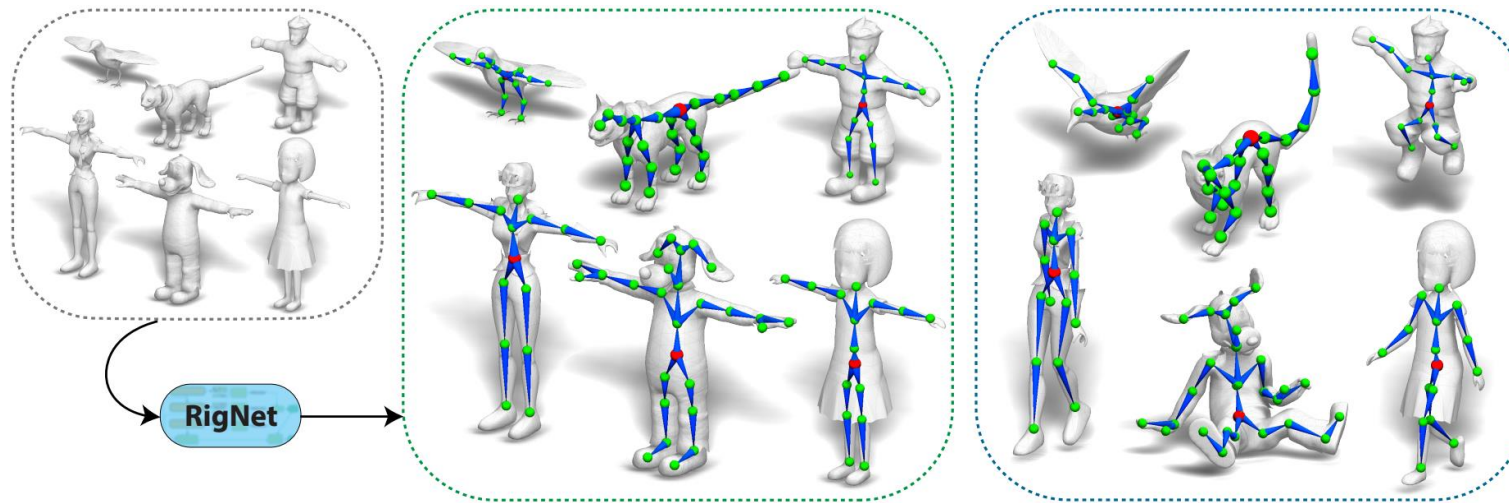


3D mesh 생성 모델 간의 비교

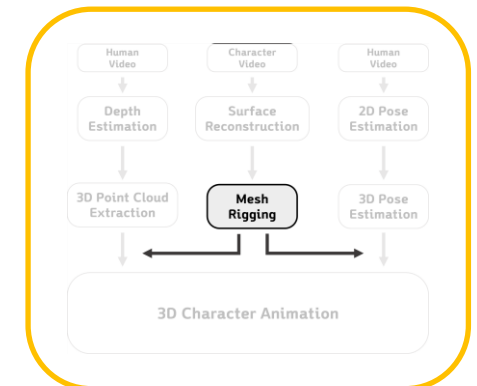


03 3D mesh & rig extraction

- **RigNet(Neural Rigging for Articulated Characters)**
 - 3D mesh로부터 공간 정보를 추출하여 rig 정보를 예측하는 모델
 - Mesh와 Rig 정보를 바탕으로 각 joint와 surface 사이의 관계를 학습
 - ModelResource dataset으로 pretrain
 - Neuralangelo에서 생성된 mesh의 각 꼭짓점을 simplify한 후 rig 정보 예측



RigNet input & output



03 3D mesh & rig extraction

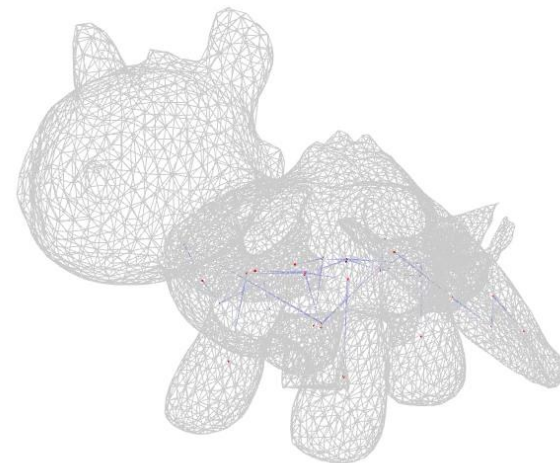
새로운 캐릭터 mesh 생성, rig 정보 예측



NeRF output mesh

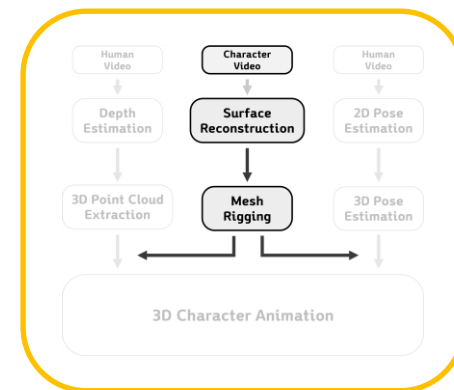


Neuralangelo output mesh



RigNet output rig

- NeRF output mesh 또한 선명했으나 noise가 많고 bounded 하지 않음
- 안과 밖의 경계가 명확하게 구분되지 않은 mesh로 인해 rig 정보 계산 불가
- Neuralangelo로 생성한 mesh로 rig 정보 예측

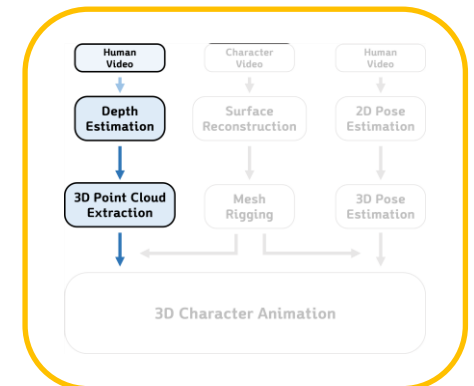


03 3D point cloud sequence extraction

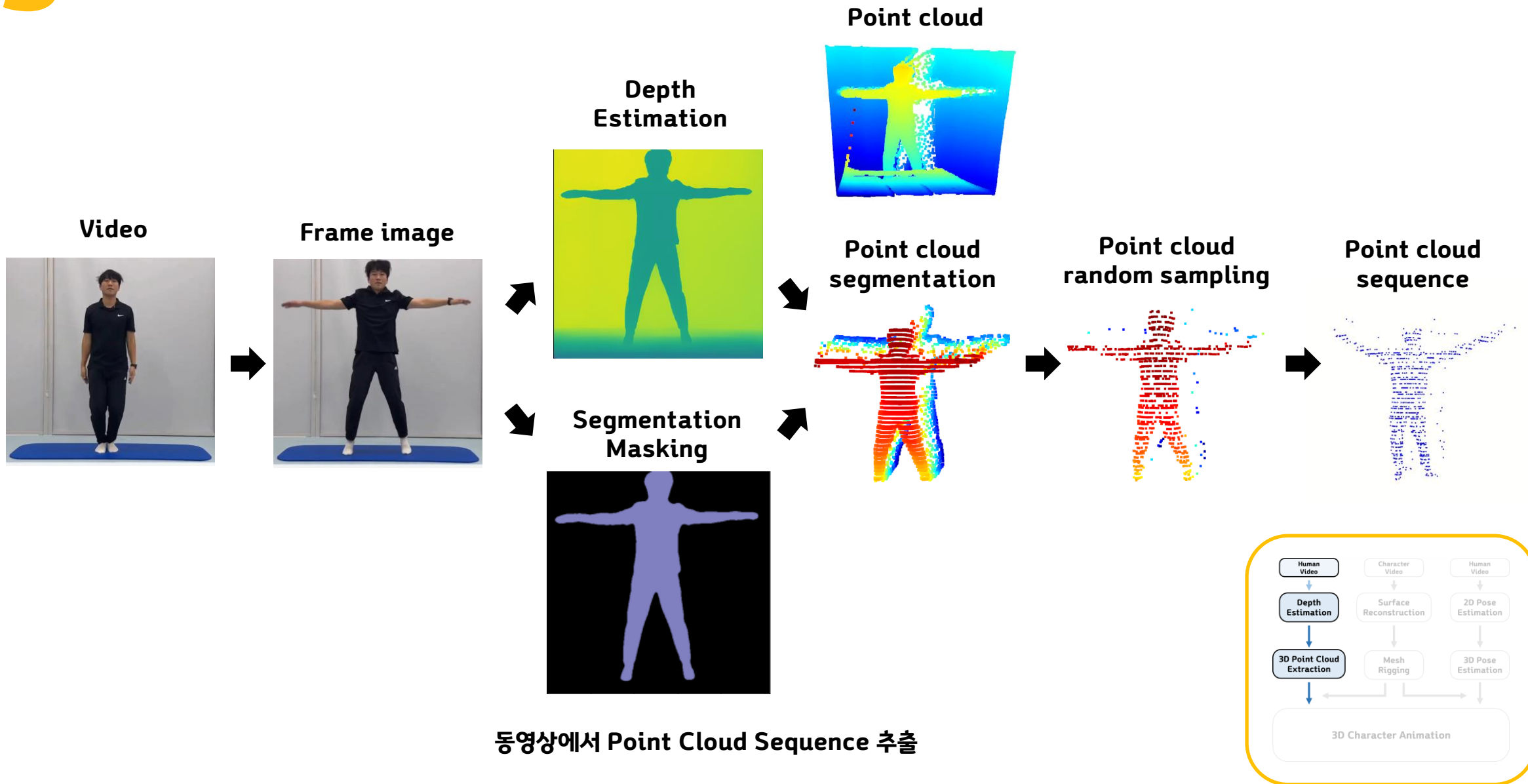
- MiDaS(Mixed Datasets for Adaptive Scaling)
 - 성능 향상을 목적으로 다양한 데이터 세트에서 학습된 monocular depth estimation 모델
 - 한 장의 image를 input으로 넣어주면 해당 scene에 대한 depth map 예측
 - 사람 동작 동영상으로부터 frame 추출 후 각 frame마다 depth map 예측
 - 이후 depth to point cloud 방식을 적용하여 3D point cloud sequence 추출



Depth map 예시 (밝은 노랑 - 근거리, 짙은 파랑 - 원거리)

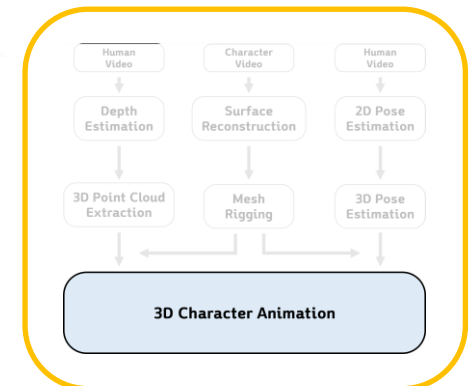
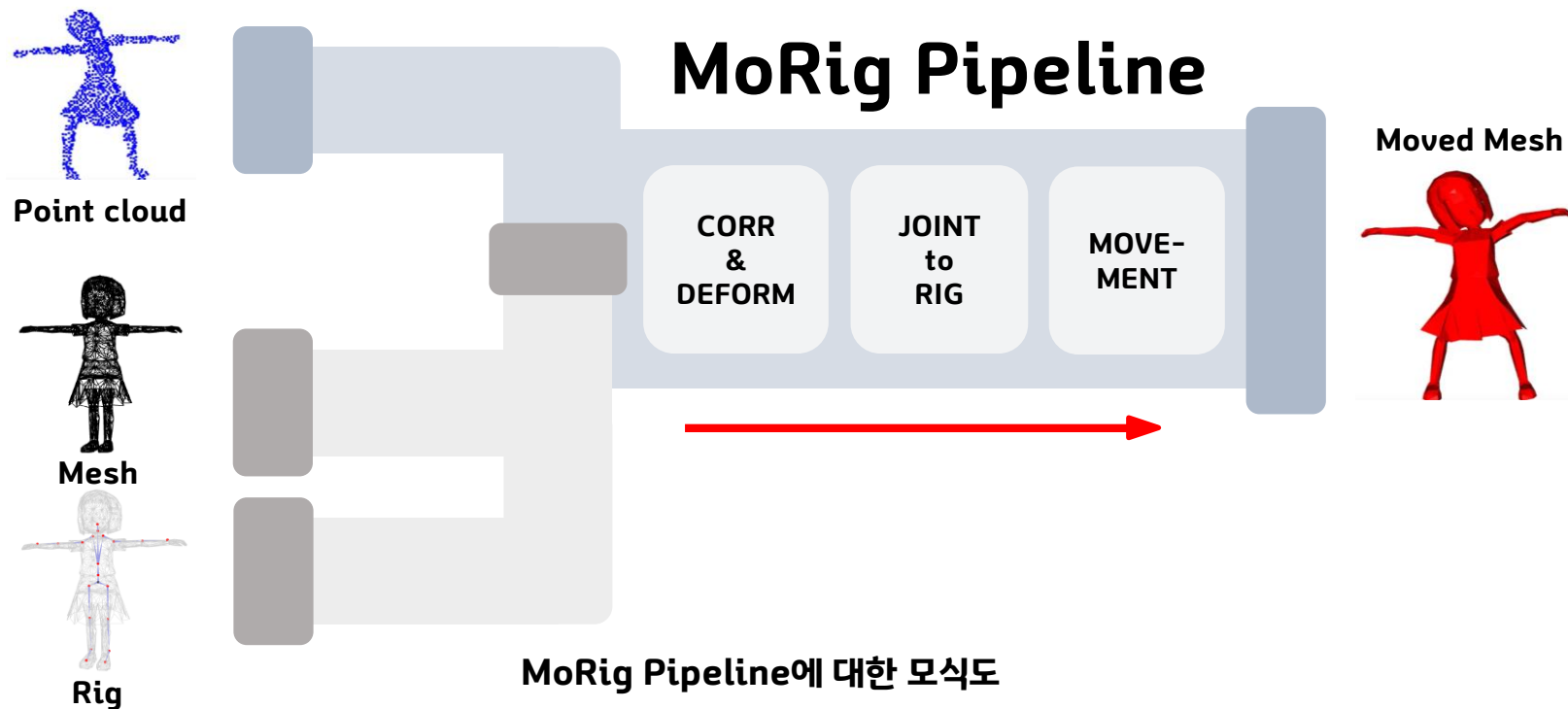


03 3D point cloud sequence extraction



03 3D character animation with 3D point cloud sequence(track 1)

- MoRig(Motion-aware Rigging of Character meshes from Point cloud)
 - Point cloud와 mesh로부터 계산한 rig 정보, geometric descriptor들을 input으로 받음
 - Mesh와 point cloud간이 대응 관계를 예측, 계산하고 관절과 그 대응관계들을 기반으로 mesh의 vertex 좌표를 예측하는 모델 * vertex : mesh의 색상, 벡터, 텍스처 등의 정보를 가진 위치

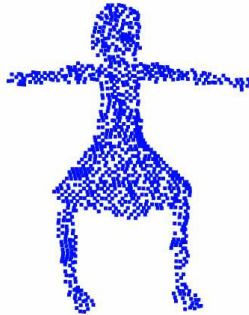


03 3D character animation with 3D point cloud sequence(track 1)

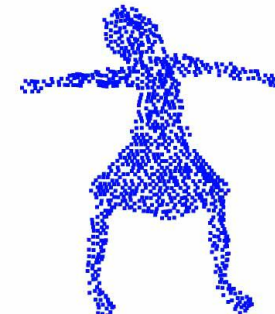
- MoRig를 사용하면 point cloud의 움직임을 mesh에 투영하게 됨
- 제공 Dataset에서 mesh가 다른 point cloud의 움직임을 따라하는 것을 확인

빨간색 : Mesh

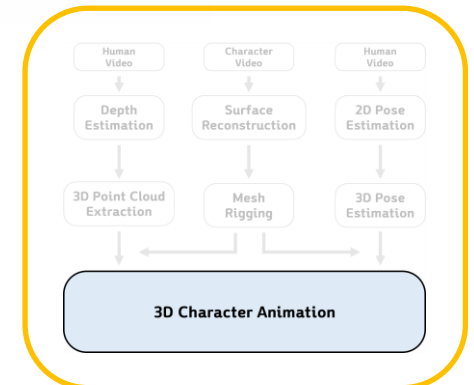
파란색 : Point cloud



MoRig 적용 전



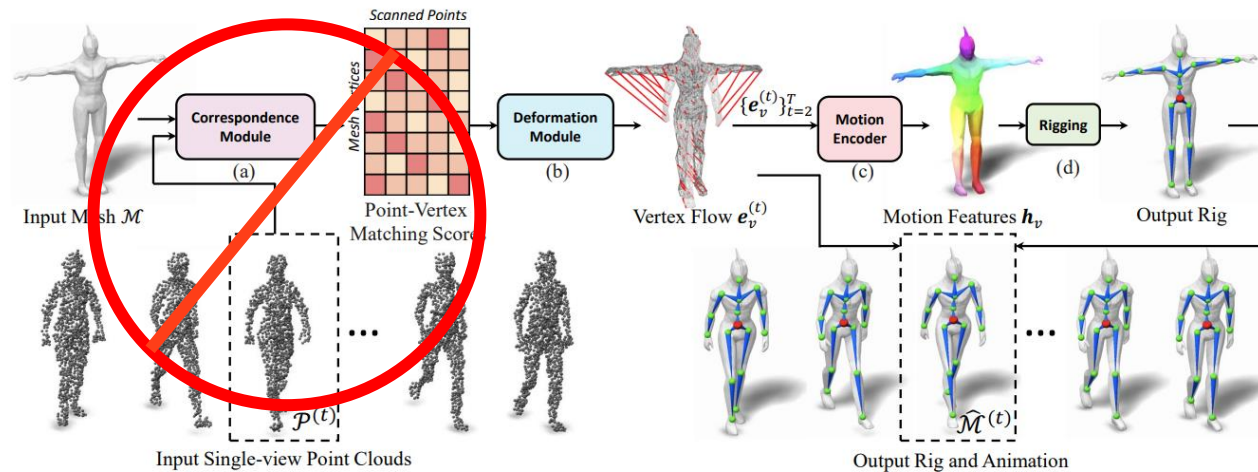
MoRig 적용 후



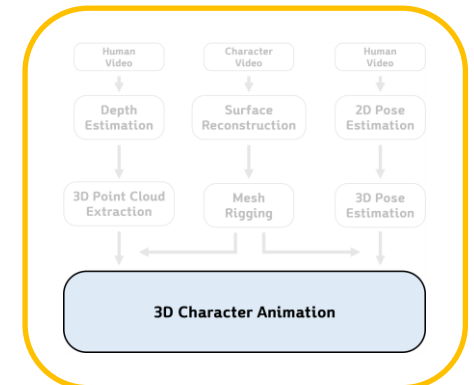
03 3D character animation with 3D point cloud sequence(track 1)

- 발생 이슈

- 새로 생성한 mesh와 point cloud 대응점을 찾기 위해서 correspondence를 구하는 과정에 누락된 source code가 있는 것을 확인
- 현재 제공 코드에서는 기존 데이터셋 외의 데이터에 적용할 수 없음

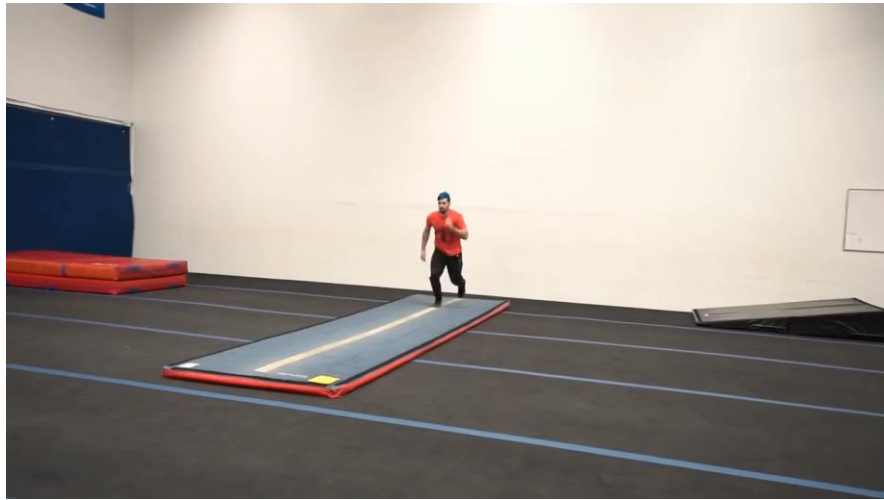


MoRig의 상세 진행 과정

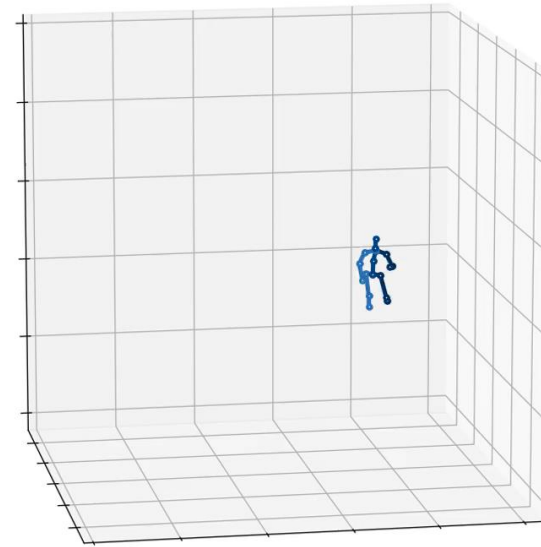


03 3D character animation with 3D human pose sequence(track 2)

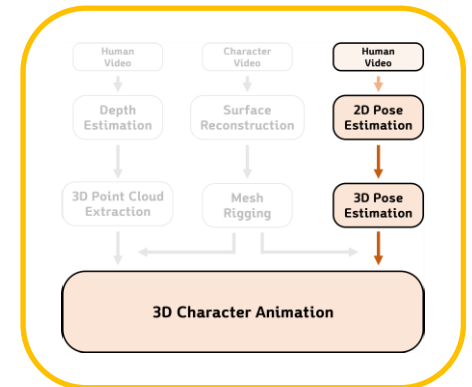
- **MotionBERT(A Unified Perspective on Learning Human Motion Representations)**
 - 동영상에서 추출된 2D pose sequence를 기반으로 3D pose sequence를 예측하는 모델
 - 동영상에서 AlphaPose 모델로부터 추출된 2D pose sequence를 사용
 - 2D pose 기반의 3D pose를 예측하여 motion sequence 생성
 - Rig 정보를 가진 캐릭터에 frame 별로 motion 정보를 입력하는 traditional rigging



원본 동영상



추출된 3D pose



03 3D character animation with 3D human pose sequence(track 2)

- 발생 이슈

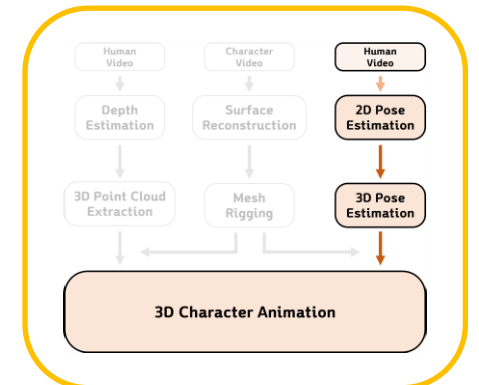
- 후처리 작업 혹은 다른 애니메이션 작업에 용이하도록 FBX로 export는 가능
- 하지만 추출된 3D pose는 x, y, z 형태의 직교 좌표계로 반환
- Euler angle 형태로 변환하는 과정에서 문제 발생 * Euler angle : x, y, z 각 축을 중심으로 회전한 세 개의 각도
- 현재로서는 정확한 결과가 나타나지 않음



원본 동영상

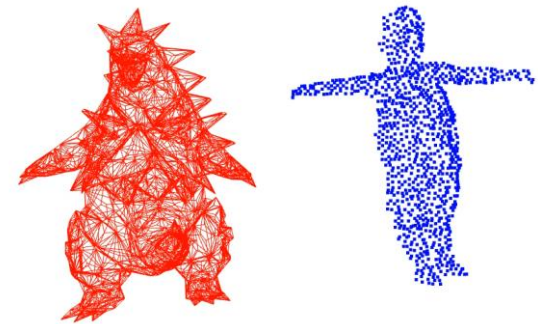


추출된 3D animation



04 프로젝트 추후 개선 방안

Part	가능 개선 방향
Mesh	<ul style="list-style-type: none">○ Input 영상에서 손떨림 등의 흔들림 보정
Point Cloud	<ul style="list-style-type: none">○ Image segmentation 대신 point cloud 군집화로 segmentation
RigNet	<ul style="list-style-type: none">○ Rigging 정확도 개선 필요○ 동물형 데이터셋을 활용한 finetuning
MoRig	<ul style="list-style-type: none">○ Correspondence 알고리즘 구현○ Correspondence 관련 모델의 finetuning
Pose	<ul style="list-style-type: none">○ x, y, z to Euler angle 변환 알고리즘 구현



사람처럼 움직이고 싶었는데...



놀랍게도 텀블링 중입니다



Reference

1. 프로젝트 배경 : (150) 배그 이모트가 만들어지는 과정 - 2 - YouTube
2. 프로젝트 배경 : (156) 디즈니 픽사 3D 애니메이션 빅히어로 베이맥스 스피드 모델링 애니메이션 렌더링 [CGusLab] - YouTube
3. NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction (arxiv.org)
4. Point cloud 원본 영상 : <https://www.youtube.com/watch?v=ISewSH6DsKU>
5. Omnidata paper: <https://arxiv.org/pdf/2110.04994.pdf>
6. MiDaS paper: <https://arxiv.org/pdf/1907.01341.pdf>

T H A N K Y O U

감사합니다

AIFTEL

모두의연구소