

2021학년도 1학기 소프트웨어 종합설계

Parametric Control of Portrait Images

팀명 GrassC
이소정, 전영주, 정윤석
김선주 교수님, 전수빈 조교님

CONTENTS

Subject

- 연구주제 소개
- 연구 필요성
- 기존 연구와의 비교
- 결론

Method

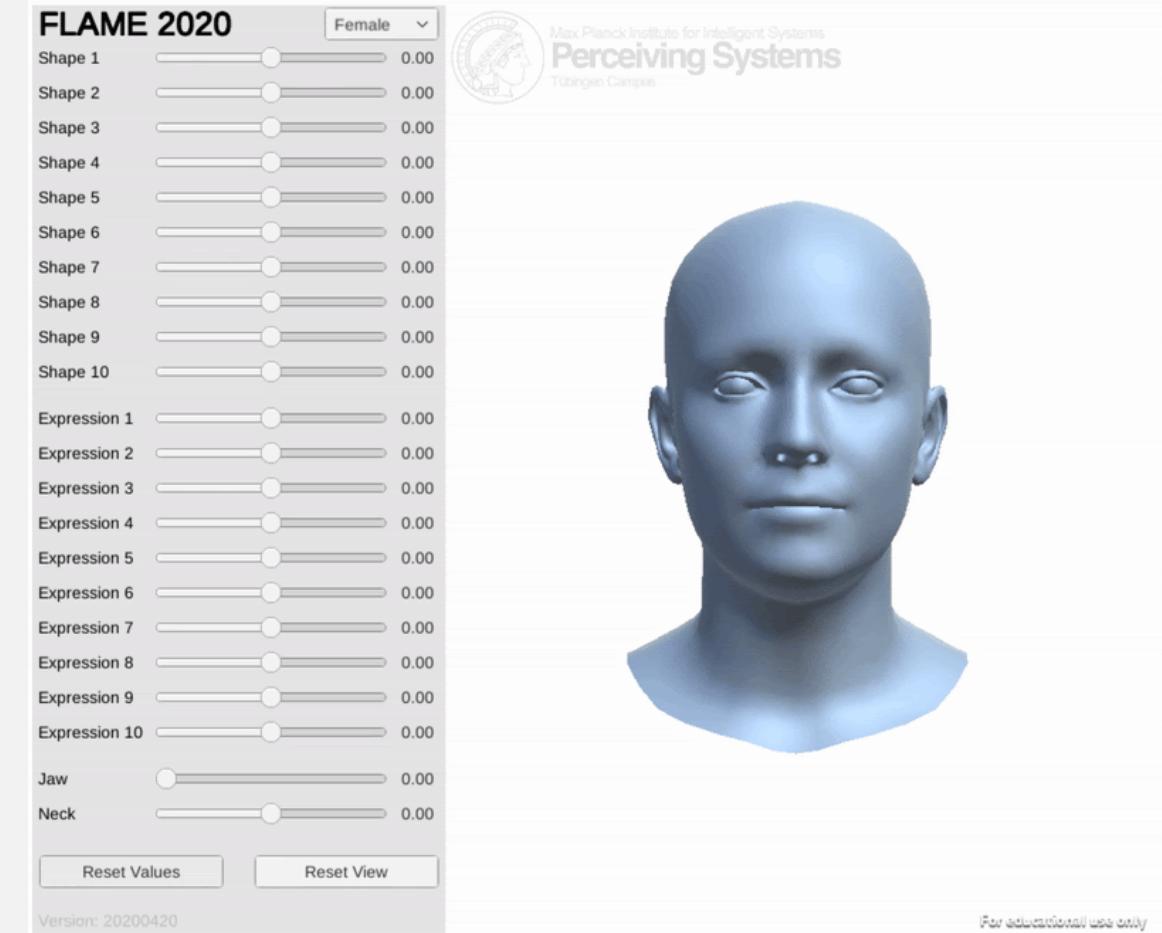
- 연구 방법
- 진행 상황

Result

- 연구 결과
- 개선점

Parametric Control of Portrait Images ▼

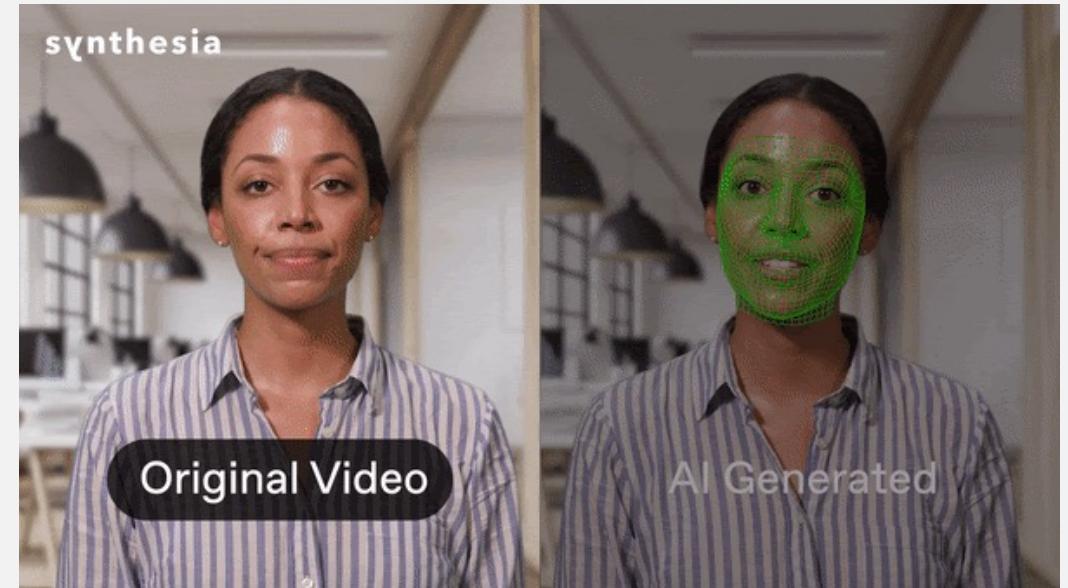
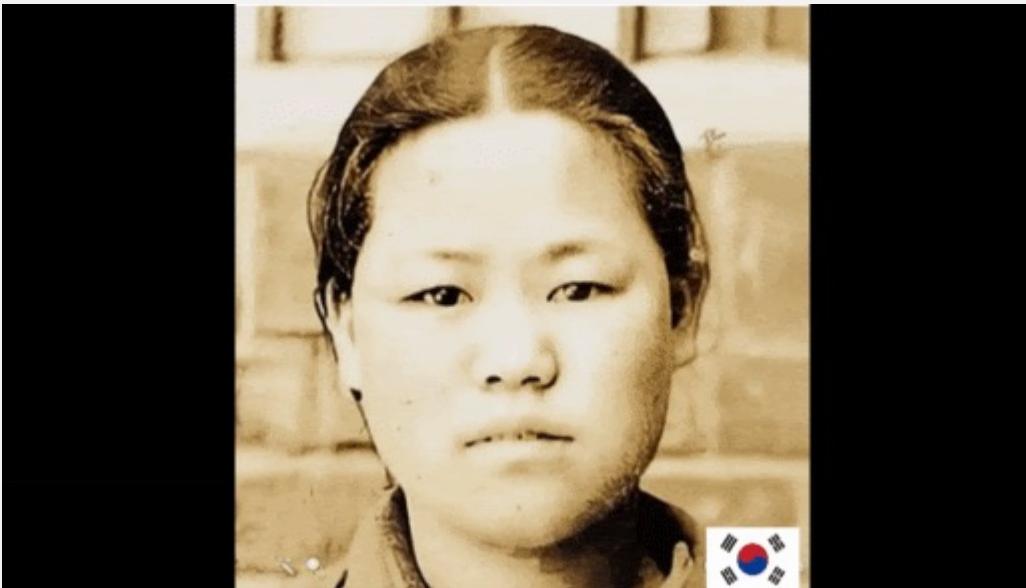
인풋으로 주어지는 사람의 single portrait 이미지의 pose와 expression을 parametric control 하는 방법 제안 및 인터페이스 설계



Subject
연구주제 소개

연구 필요성

1) Face editing의 수요가 기하급수적으로 증가하고 있으며, face reenactment 기술은 synthetic media의 핵심기술로 성장하고 있음.



2) Parametric Control의 필요성

- 사용 편의성을 위해 semantically meaningful parameterization을 통한 control이 가능하도록 제공 해주어야 한다.
- 그러나, 대부분의 기존 기술은 직관적이고 세밀한 control을 제공하고 있지 않거나 single isolated control parameter를 통한 대략적인 editing만이 가능하다.

PIE: Portrait Image Embedding for Semantic Control

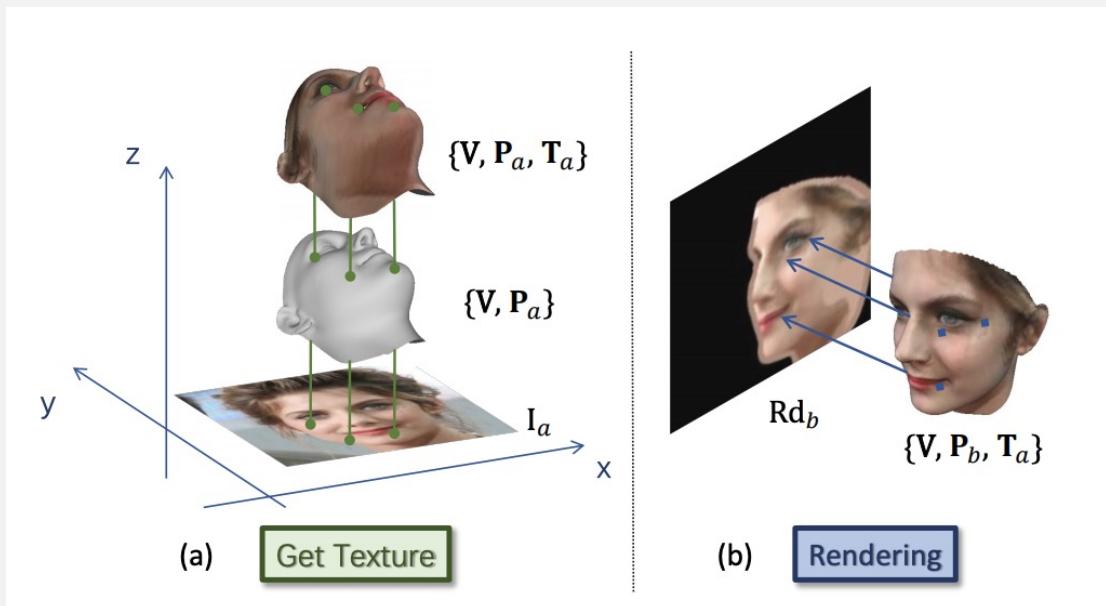
3) 얼굴 외 부분 합성의 필요성

- 얼굴 부분만을 가지고 3d model을 만드는 기법은 존재한다.
- 그러나, hair나 ear 등 얼굴 외의 부분까지 포함하여 3d model을 만들 경우, 사람의 얼굴 표현이 더 자연스럽다.
- 그러므로, 얼굴 외의 부분까지 포함하여야 상업적으로 더 유용하다.

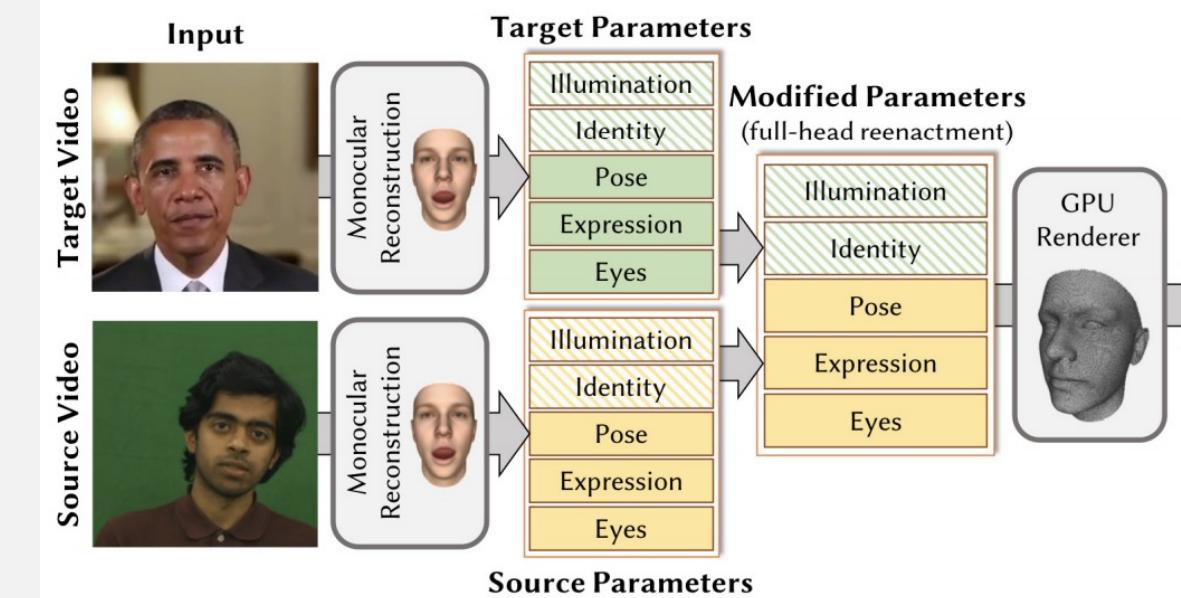
Subject
기존 연구와의 비교

기존 연구와의 차이점

Graphics pipeline



Rotate-and-Render: Unsupervised Photorealistic Face Rotation from Single-View Images (CVPR 2020)



Deep Video Portraits (Siggraph 2018)

기존 연구와의 차이점

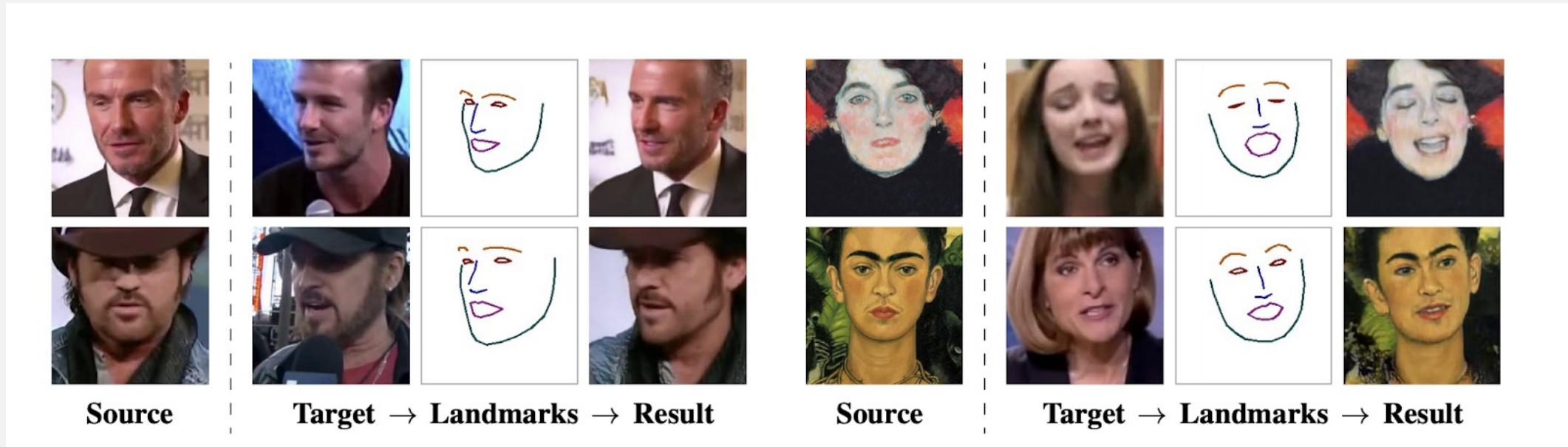
Graphics pipeline

- 장점 : 구축된 3d face 모델을 사용하기 때문에 expression에 대한 parametric control 하기에 용이
- 한계 : Face에 국한된 모델링/렌더링 → hair/ear 를 포함하는 portrait 이미지 생성에 어려움

Subject
기존 연구와의 비교

기존 연구와의 차이점

GAN 기법을 이용한 face reenactment



Few-Shot Adversarial Learning of Realistic
Neural Talking Head Models
(ICCV 2019)

기존 연구와의 차이점

GAN 기법을 이용한 face reenactment

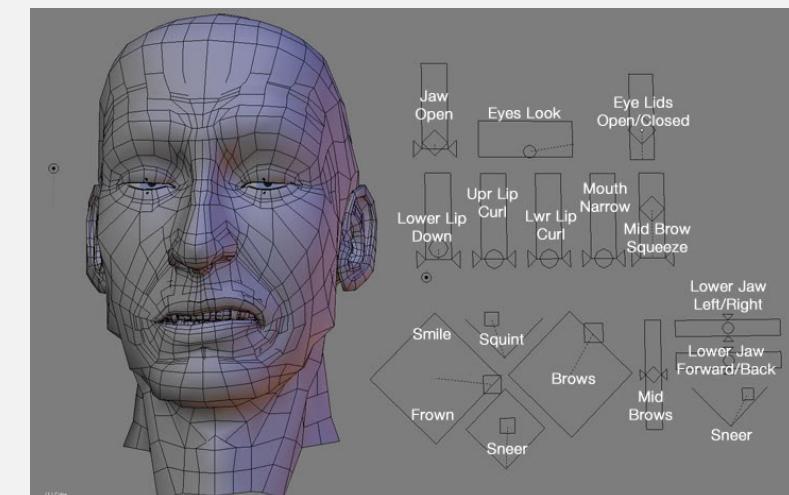
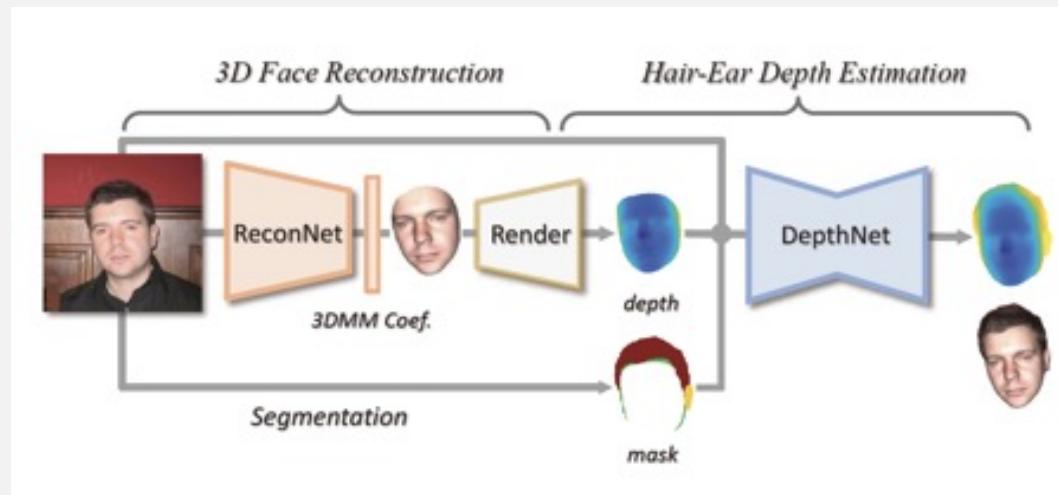
- 장점 : 3DMM같은 모델링 없이 더 넓은 영역(hair + ear)에 대한 사실적인 이미지 합성
- 한계 : 이미지 생성 과정에서 3d 모델을 생성하지 않고 GAN 기법을 사용하였기 때문에 3d 모델을 만들 때 geometric reasoning이 없다. 따라서 3D 모델을 만들 때의 parameter에 간섭을 할 수 없고 그에 따라 추가적인 parametric control이 불가능하다.

Subject
결론

연구 목표

- 얼굴 뿐만이 아니라 hair와 ear를 포함한 3D model 구축
- Pose와 expression을 parametric하게 control 가능

- 기존에 있는 모델인 3DMM을 사용하여 3D Face Reconstruction을 진행
- Sample 얼굴에서 expression basis를 추출하여 3d face에 적용
- Segmentation, Depth map 기법을 사용하여 hair, ear와 같은 face 이외의 부분 생성
- UI 설계 및 모델과 연동



Method
진행 과정

Face reconstruction 진행

- 기존에 있는 모델인 3DMM 중 Basel face model(BFM) 사용

Face reconstruction 진행

3DMM이란

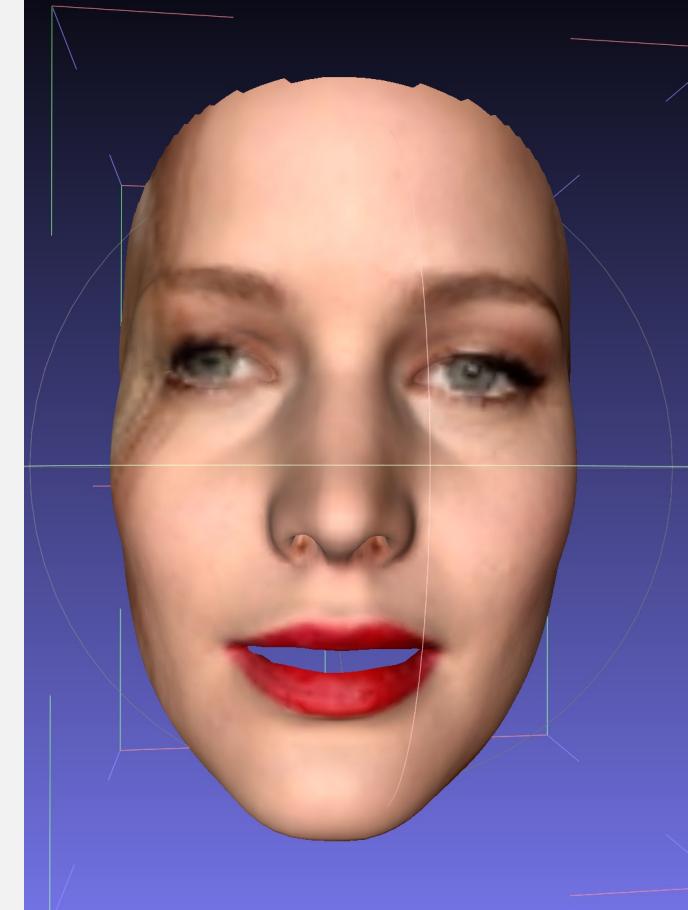
$$\mathbf{F} = \mathbf{F}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \bar{\mathbf{F}} + \mathbf{B}_{id}\boldsymbol{\alpha} + \mathbf{B}_{exp}\boldsymbol{\beta}$$

$$\mathbf{T} = \mathbf{T}(\boldsymbol{\delta}) = \bar{\mathbf{T}} + \mathbf{B}_t\boldsymbol{\delta}$$

- \mathbf{F} : Face Shape, \mathbf{T} : Texture
- $\bar{\mathbf{F}}$: Average Face Shape, $\bar{\mathbf{T}}$: Average Texture
- \mathbf{B}_{id} , \mathbf{B}_{exp} , \mathbf{B}_t : 각각 identity, expression, texture 의 PCA bases
- $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\delta}$: \mathbf{B}_{id} , \mathbf{B}_{exp} , \mathbf{B}_t 에 대응되는 coefficient vector
- $\bar{\mathbf{F}}, \bar{\mathbf{T}}, \mathbf{B}_{id}, \mathbf{B}_t$: BFM 모델 사용
- \mathbf{B}_{exp} : FaceWarehouse 사용

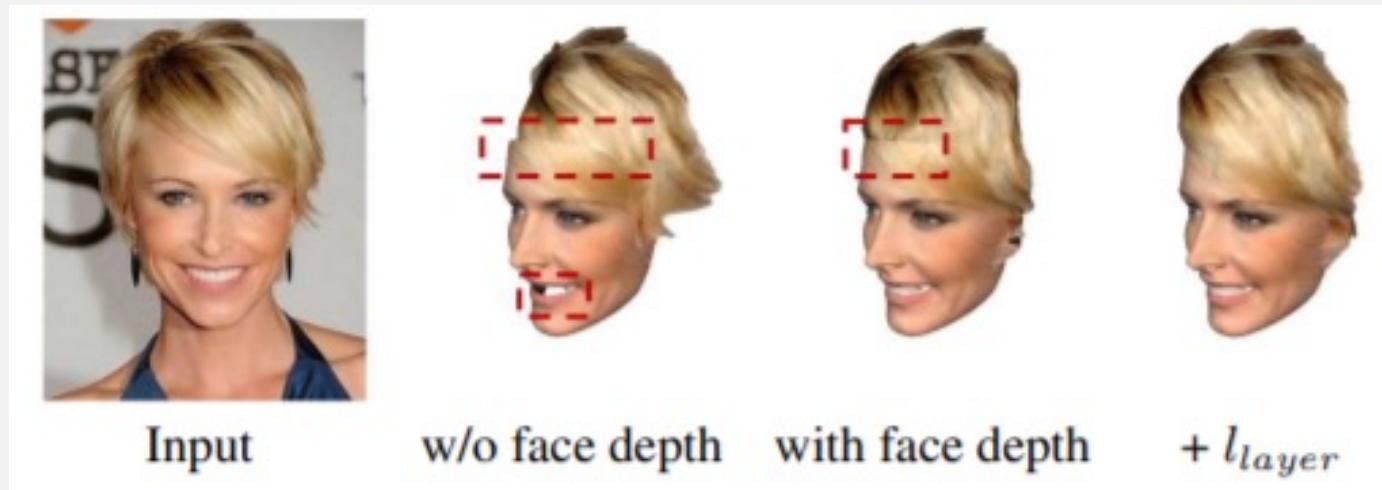
Method
진행 과정

Face reconstruction 진행 결과



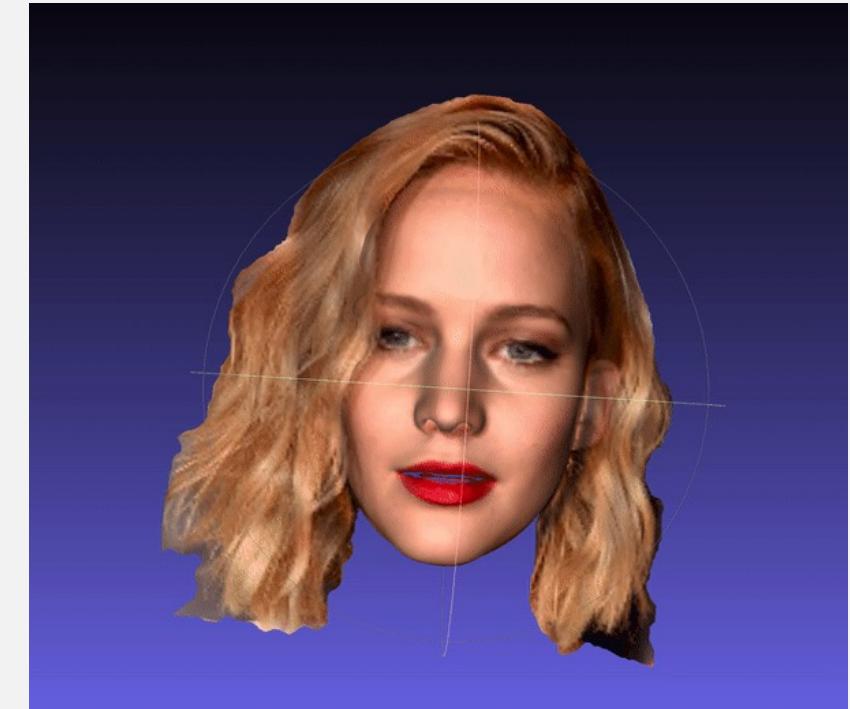
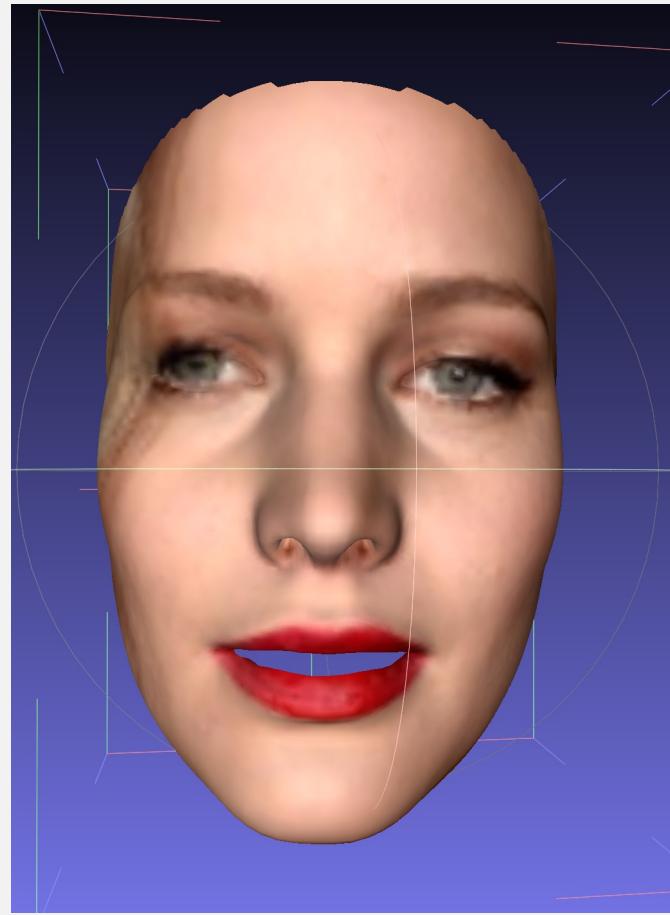
Segmentation과 Depth map 기법 적용

- 이미 존재하는 2D 이미지에 대한 segmentation을 사용
- Segmentation으로 hair와 ear 부분 mask 추출
- BFM으로 얻은 face 부분과 mask를 depth net을 이용하여 자연스럽게 합성



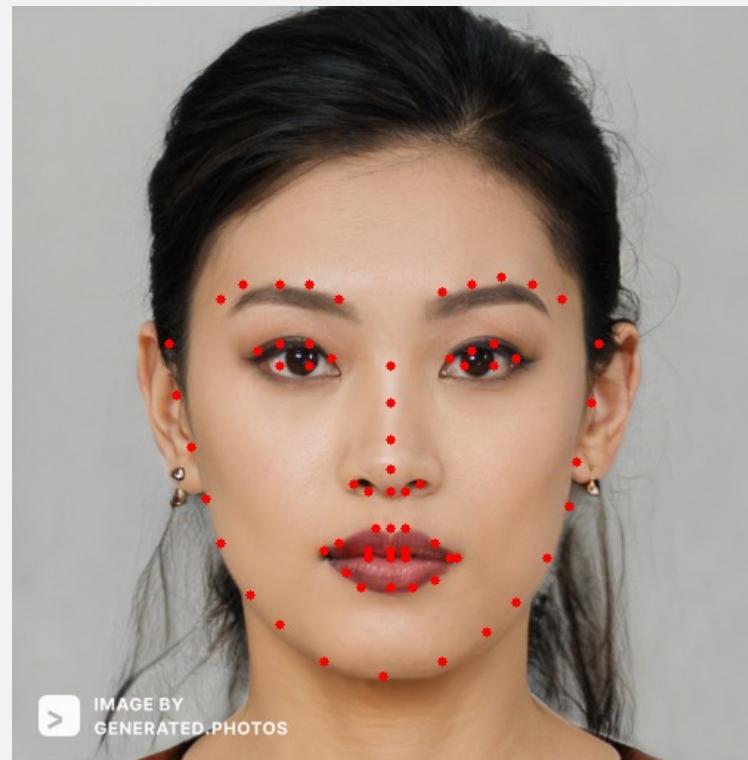
Method
진행 과정

Segmentation과 Depth map 기법 적용 결과

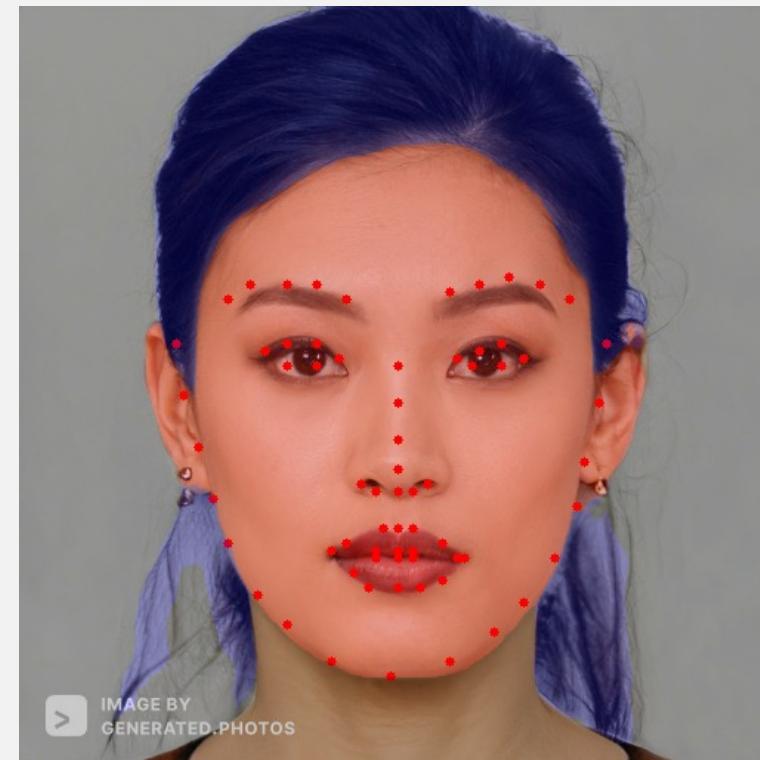


임의의 사진을 이용하기 위한 작업

1) Facial landmark 추출



2) Parsed segmentation 추출



Method

진행 과정

임의의 사진을 이용하기 위한 작업

2) Parsed segmentation 추출

- Segmentation 추출을 위해 *Face Parsing with RoI Tanh-Warping* 논문에서 언급된 RoI-tanh warping 기법을 활용한 face parsing이 이루어져야 함
- RoI-tanh warping 기법을 활용한 공식 코드가 없어서 앞서 설명한 대로 공식 리포에서 언급된 모델을 적용
- 기반 논문에서 언급한 공식 리포의 모델을 규격에 맞게 변형시켜 해결

Method
진행 과정

임의의 사진을 이용하여 생성한 3D model



Head Pose Control

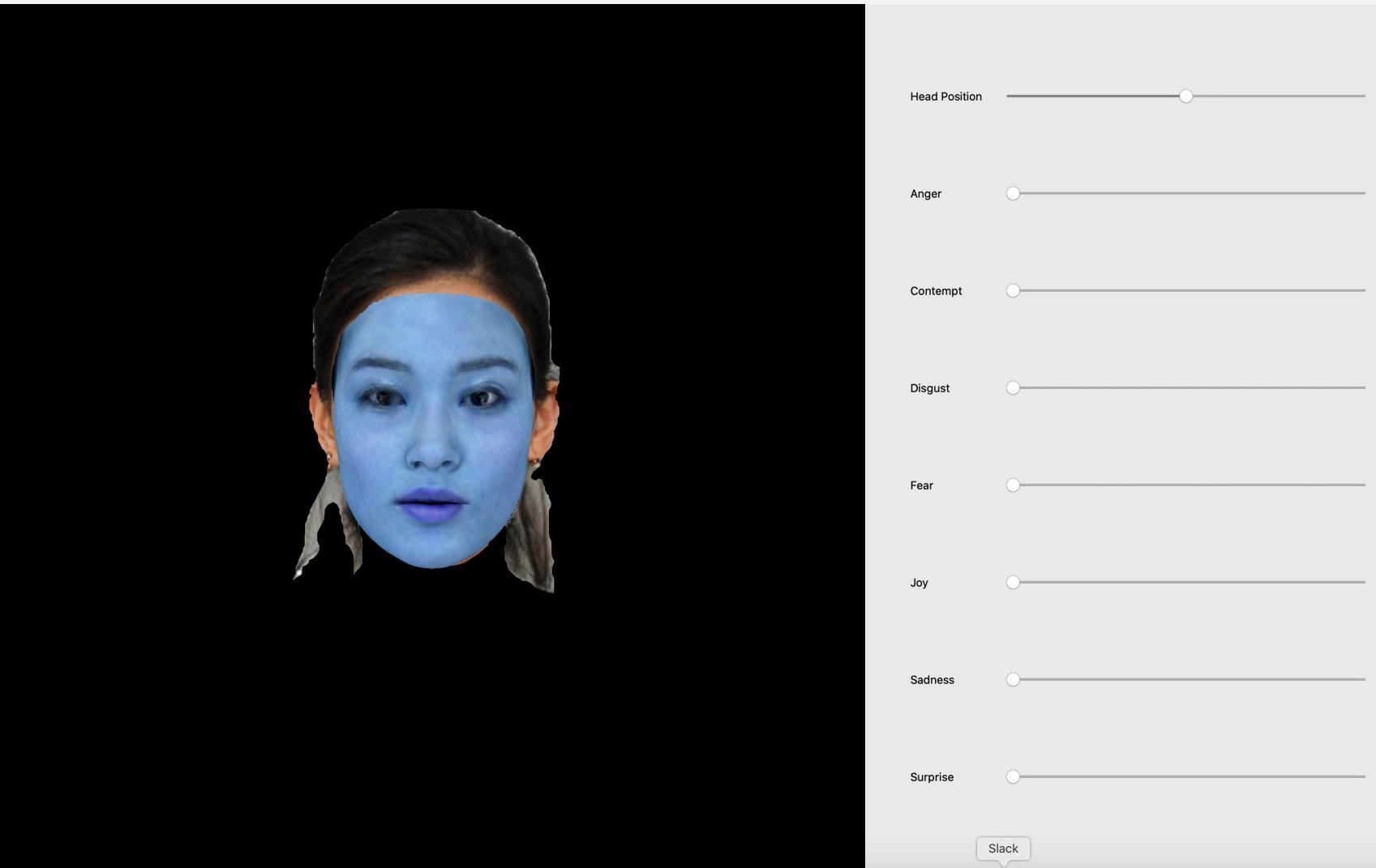
- 3D model을 2D 이미지로 변환하지 않고 UI 상에서 3D model의 head pose만 조정하면 되기 때문에 object 파일의 값 자체를 바꿀 필요가 없어서 UI 상의 camera를 rotation 함으로써 head pose control이 가능하게 함
- 추후에 2D 이미지로 다시 변환해야 하게 된다면 rotation에 따라 object 파일의 값도 변경할 예정

Expression Control

- 기존에는 input image의 expression PCA bases를 분석하여 각 표정과 관련된 유의미한 parameter를 추출하려고 하였으나 parameter와 expression 간의 관계를 직접 알아내기 어려웠다.
- Input image의 expression PCA bases를 분석하여 조정하는 방법 대신, 각 expression을 대표할 수 있는 image의 expression PCA bases를 추출하여 input image의 expression PCA bases에 적용시키는 방법을 사용하였다.
- 각 expression마다 10단계로 나누었으며, 단계별로 적용하고자 하는 값의 정도를 0~100%로 나누어 input image의 expression PCA bases에 적용하였다.

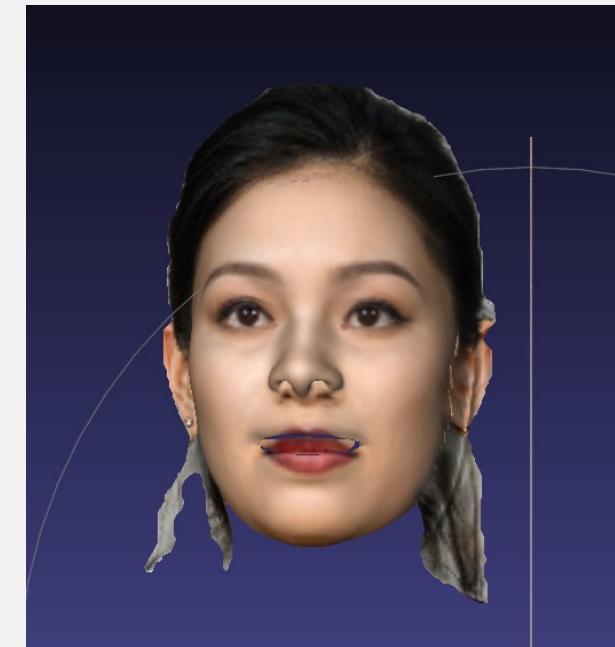
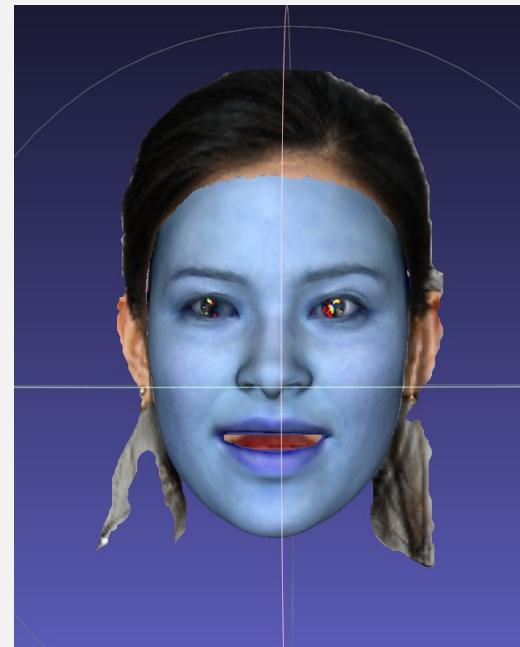
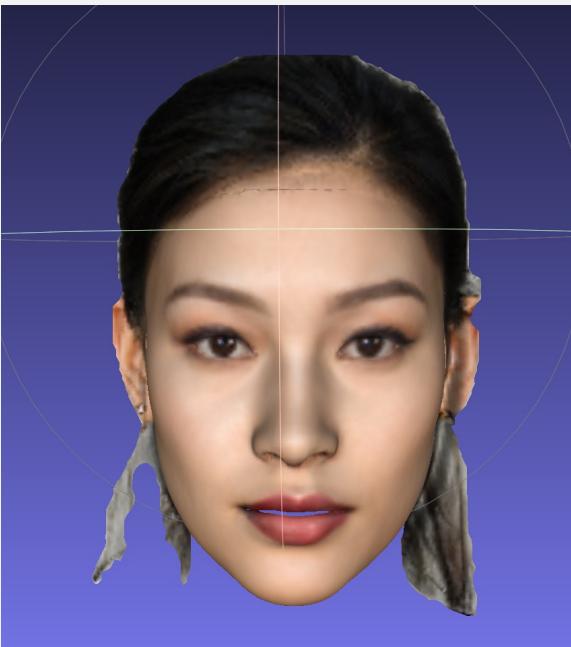
Result

연구 결과



Result 개선점

- Parametric control 이후에 3d model에 input image의 texture를 입히면 expression control이 눈에 잘 띠지 않는 현상을 발견하였다.



Reference

참고 문헌

<기반 논문>

- Sicheng Xu, Jiaolong Yang, Dong Chen, Fang Wen, Yu Deng, Yunde Jia, Xin Tong. 2020. Deep 3D Portrait from a Single Image. In Conference on Computer Vision and Pattern Recognition (CVPR).

<모델 사진>

- <https://generated.photos/face/neutral-asian-young-adult-female-with-long-black-hair-and-brown-eyes--5e6887b96d3b380006f1d845>

- Ayush Tewari, Mohamed Elgharib, Mallikarjun B R., Florian Bernard, Hans-Peter Seidel, Patrick Pérez, Michael Zollhöfer, Christian Theobalt. 20 Sep 2020. PIE: Portrait Image Embedding for Semantic Control. SIGGRAPH Asia 2020.
- Egor Zakharov, Aliaksandra Shysheya, Egor Burkov, Victor Lempitsky. 2019. Few-Shot Adversarial Learning of Realistic Neural Talking Head Models. In The IEEE International Conference on Computer Vision (ICCV).
- Hang Zhou, Jihao Liu, Ziwei Liu, Yu Liu, Xiaogang Wang. 2020. Rotate-and-Render: Unsupervised Photorealistic Face Rotation from Single-View Images. In Conference on Computer Vision and Pattern Recognition (CVPR).
- Hyeongwoo Kim, Pablo Garrido, Ayush Tewari, Weipeng Xu, Justus Thies, Matthias Nießner, Patrick Pérez, Christian Richardt, Michael Zollhöfe, Christian Theobalt. 2018. Deep Video Portraits. In the SIGGRAPH.
- Jinpeng Lin, Hao Yang, Dong Chen, Ming Zeng, Fang Wen, Lu Yuan. 2019. Face Parsing with Roll Tanh-Warping. In Conference on Computer Vision and Pattern Recognition (CVPR).

감사합니다