Team

Members [team_list]

```
- Ben Robert Sturgis    03776460
- Kevin Qu              03730587
- Marco Busch           03779574
- Ryoto Kato            03767467
```

Tasks

- June 09
    - ☑ Project topic
    - ☑ Apply as a team via Google Form
    - ☐ Create proposal (Overleaf Latex)
    - ☐ Submit proposal
- June 16
    - ☐ weekly report
- June 23
    - ☐ weekly report
- June 30
    - ☐ weekly report
- July 7
    - ☐ weekly report
- July 14
    - ☐ weekly report
- July 21
    - ☐ weekly report
- **Friday July 28 23:59**
    - ☐ Final report
- **August 1**
    - ☐ Presentation
    - ☐ 5 mins videos

Proposals (Deadline: June 9)

- Overleaf

Introduction/Abstract

- Describe the problem you want to solve and why

- In the recent years, generative AI has changed our daily life and even our working style because of their potential and flexibility to various tasks. At the same time, our opportunities to interact with AI itself are increasing much more than ever and we are starting to trust AI gradually. In order to make AI more convincing to human, it is crucial to let them have human-like appearance and expressions, namely digital humans. The most difficulty is synthesizing convincing facial expression against human who are really sensitive to facial expression as we described with "the uncanny valley" reference. Especially, generating reliable facial expression in real-time requires us to parameterize them in low-dimensional space. One of parametric face models, so-called morphable model, could be obtained by principal component analysis on the aligned face templates on scans. Thies et al. used this model and develop to realize real-time facial reenactment of a monocular target video. The one of drawbacks of PCA model is that the target identity is only described by texture and shape of the face model. While people have different movement to express same expression, this type of techniques cannot synthesize personalized facial expression because they are synthesized average facial expressions by referring global expression change directions (global principal components). In this proposal, we are going to approximate distribution of target's facial expression (without calibration) by using nearest neighbor estimation techniques within parametric model to realize "personalized" and characteristic facial expression for more convincing digital human realization. While synthesis of personalized facial expression is implemented easily by calibration of the target's facial expression parameters and applying PCA on them, we have already known that statistical techniques can be applied on scanned human faces and it enables us to parameterize our complicated facial expression to some extent. By extending this knowledge, we aim to infer the distribution by referring data distribution in a parametric model and investigate the effectiveness of personalization (characterization) of facial expression compared to global parameterization techniques.

Technical Approach

- How do you propose to solve it

- Approach 1 (K-means clustering)

    - Apply K-means clustering on dataset
        - Obtain centroids (means of face vertices positions)
        - Obtain averaged $\mathbf{Eexp}$ per cluster
    - In preprocessing, we estimate Φ (Shape identity) and $\Beta$ (Material/Albedo) as Face2Face does.
    - By using estimated parameters, find a closest cluster and obtain its centroid and $\mathbf{Eexp}$
    - Use them as the principal component of the facial expressions
    - Construct energy function
    - Parameter Estimation as energy minimization
    - perform **Analysis-by-Synthesis**

- Approach 2 (Face skeletal structure descriptor)

    - By using detected facial landmarks, we obtain sparse and a low-dimensional representatives of target's face skeletal structure (facial mesh)
    - It can be used as input to get a descriptor of the identity.

- Since we have vertex correspondences between space face meshes within database, we can use some metric which are used for energy minimization in Non-rigid 3D reconstruction (Mesh deformation)
- As we measure deviation from rigidity by summing up deviation with respect to a fan, we can obtain "Face skeletal structure" difference between a target person and one from dataset $$ \mathbf{E_{fss}} = \sum_i\sum_{j\in\mathcal{N}(i)} ||(vi-vj)-(vi'-vj')||^2 \ \text{for ith vertex fan} $$
- Matching strategy
  1. Target vs one from parametric model and try to find a face model which has the highest similarity
  2. Target vs average face descriptor from clusters and find a cluster which has the highest similarity
- By matching, we will obtain the
- In preprocessing, we estimate Φ (Shape identity) and $\Beta$ (Material/Albedo) as Face2Face does.
- We derive a certain structure $\mathbf{Eexp}$ and its standard deviation by referring nearest neighboring face model $\mathbf{Eexp}$ and its standard deviation

- ,

Requirements

- Which datasets & libraries will you need ?
- Facial Landmark detection
- Dataset
-

Milestones

- Time estimate, tasks
  - PCA (if we do not have parametric model data base and only scanning data)
    - Align template to the every scanned face
    - PCA of the aligned face mesh
    - Obtain the parametric model
  - Face Landmark detection and data stream
  - Face template alignment
  - Face expression transfer
    - Second Proposal
      - Find personalized Facial expression deformation components
      - Personalized deformation coefficients
        - Goal: approximate plausible face deformation principal components for target face by referring neighboring face model in a certain feature space (Face identity itself or face skeletal structure)
  - Precomputing process: identify target RGB input Identity (Albedo and Identity) and Pose
  - Texture/Lighting
  - Rendering
-

- Team-Members
  - Group of 4

Reference

- @inproceedings{Mori2017TheUV, title={The Uncanny Valley: The Original Essay by Masahiro Mori-IEEE Spectrum}, author={Masahiro Mori and Karl F. Macdorman}, year={2017} }

Questions

- How can we obtain the dataset which is also used for the Face2Face
- We know that Face2Face used Facial Landmark detector but how can we obtain correspondence between face mesh template vertices and the detected landmark
-

Weekly report (Every Friday)

- Format: Google Docs
- Content
- about 1 page
  - What you have accomplished in the week
  - Which problems did you encounter
  - If possible, show some intermediate results
  - Outline plan for the next week

Final Report (Deadline: July 28)

- Content
  - 4 pages including figures and tables, excluding references
  - Should be structured like a paper
    - Introduction (Motivation)
    - Related Work (What has been done before)
    - Method (What we did)
    - Results (Quantitative & Qualitative)
    - Conclusion (What was achieved, what can be achieved in future)
- Deadline
  - Friday July 28 23:59

Topic: 3D Face reconstruction