Ming Hsieh Department of Electrical Engineering

3D Face Modeling from One or More Images

USC Viterbi School of Engineering

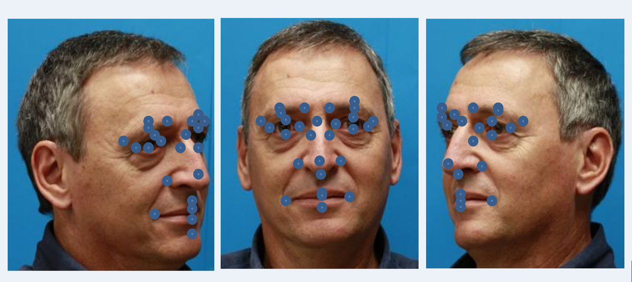
Jungyeon Kim, Jongmoo Choi, Gérard Medioni
Computer Vision Lab, Institute for Robotics and Intelligent Systems

1. Introduction

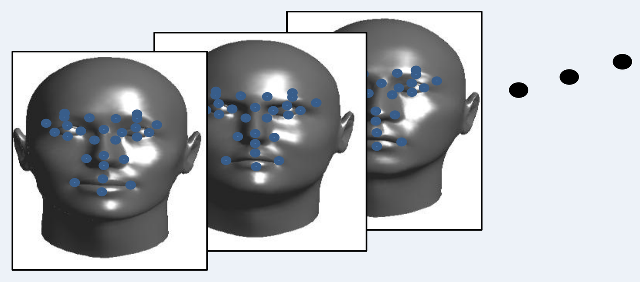
- **Why 3D face modeling ?**
  - Addresses most limitations of 2D face recognition
  - Pose, illumination, expression (PIE), aging, …
- **Goal: 3D face model from low-res. image(s)**
  - Current 3D face modeling techniques require high resolution images, videos, or many images
- **Our approach**
  - Morphable face model from a 3D face database
  - Pose estimation using affine camera model (1 or N images)
  - Update using view transfer and image features (if multiple images)

2. 3D head pose estimation from 1 or N images

• **2D image(s) with detected landmarks**



• **3D face models with pre-defined landmarks**



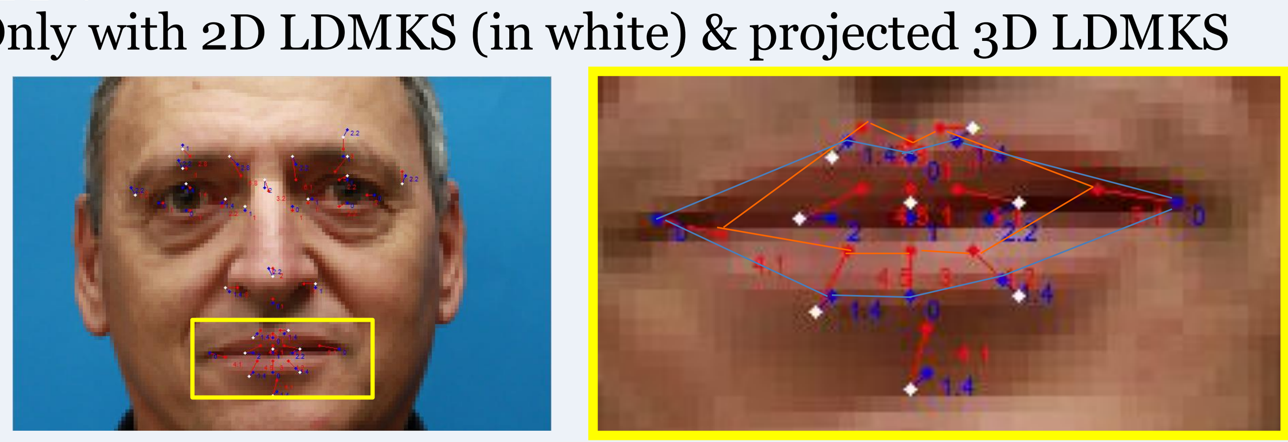
• If (iteration =1) Start from randomly selected 3D model
Else Start from the previous minimum distance error model

• **Pose estimation with a chosen 3D model**

$$\mathbf{p}_{2D} = \mathbf{A}\ \mathbf{P}_{3D} + \mathbf{t} = (\mathbf{s}\ \ \mathbf{R})\ \mathbf{P}_{3D} + \mathbf{t}$$

- $\mathbf{p}_{2D}$: 2D landmarks(LDMKS)
- $\mathbf{P}_{3D}$: 3D landmarks(LDMKS)
- $\mathbf{t}$: translation vector
- $\mathbf{A}$: affine matrix
- $\mathbf{s}$: scaling matrix
- $\mathbf{R}$: rotation matrix

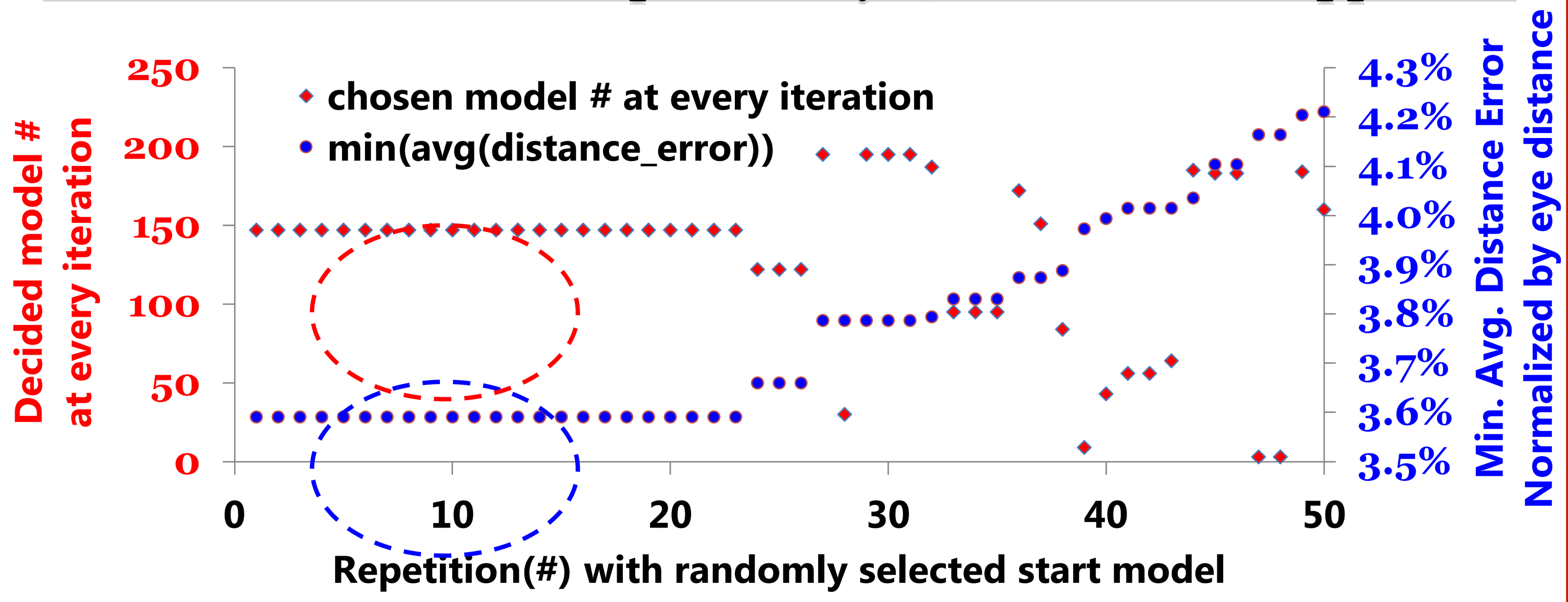
• **Local feature approach**
- **Purpose: Find min. avg. distance error model**
- **Apply "A & t" to 3D models**
- Only with 2D LDMKS (in white) & projected 3D LDMKS



(Blue: minimum distance case, Red: Maximum distance case)

• If (start model ≠minimum distance model) iteration +=iteration;
Else { iteration=1; save Affine & Local approach results}
• Repeat the whole sequence until convergence

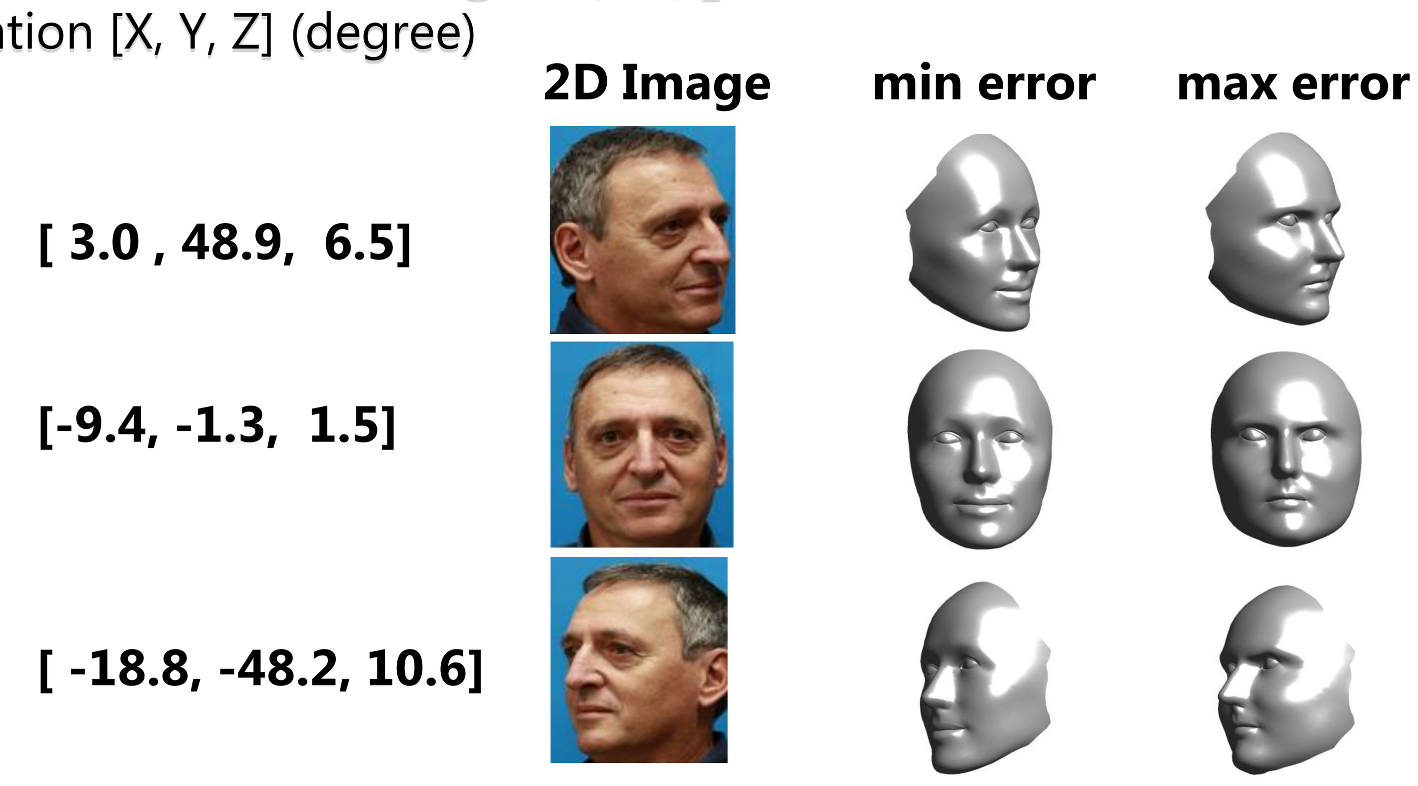
- **Min. distance error & repeatability @ Local feature approach**



(Sorted by the distance error with decided model#)

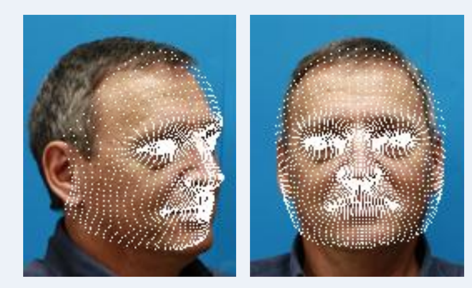
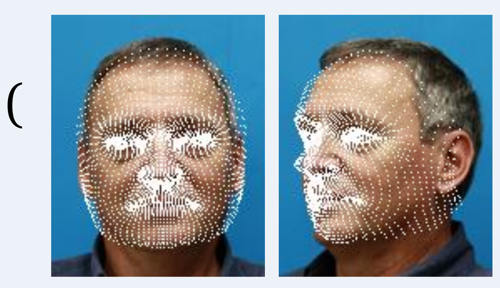
- **Pose estimation: (range: ±70°, pose error: < ±6°)**

Rotation [X, Y, Z] (degree)

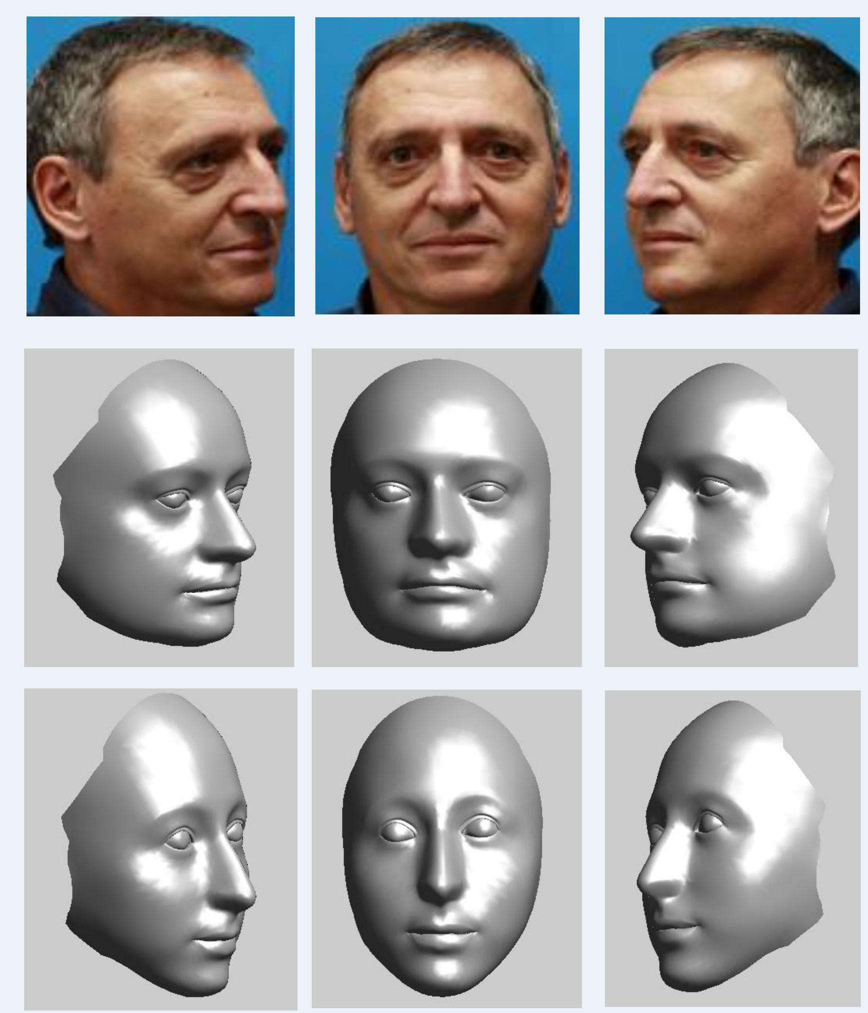


3. Global method using view transfer with N images

• **View transfer**
- Using the camera motion and 3D model, pixels in view (i) are transferred to the corresponding pixels in view (j).
- We find a set of 3D face models that provide consistency across views.
  - 2D image features (color values, features descriptors, …)
  - Example: the score is the sum of pixel correlation coefficients

correlation () + correlation ()

- **Example of chosen model**



4. 3D face modeling

• **Embedding/Mixing**
- Local feature approach: fast
- Global surface approach: slow and complex, but better results

• **Morphable 3D face modeling**
- For initial experiments, we use 1649 3D face models
- From selected models (e.g., 85 models), the final model is built by the maximum sum of correlation coefficients.

• **Texture mapping**

5. Future Work

- Performance evaluation with ground truth data
- Face expression normalization and neutralization
- Global method with more complex 2D features
- Improvement of deformable 3D face model
- Real application and performance for face recognition

Ming Hsieh Institute
Ming Hsieh Department of Electrical Engineering

jungyeon@usc.edu