Partial Matching of Interpose 3D Facial Data for Face Recognition

IEEE BTAS 09

Boris Efraty

P. Perakis, G. Passalis, T. Theoharis, G. Toderici, I. A. Kakadiaris

University of Houston

Motivation & Challenges

- Motivation
  - Pose variations can result in missing data (e.g., half of the face)
    - Common in uncontrolled environments & uncooperative subjects

- Previous Approaches
  - Do not handle extensive missing data due to pose variation
  - Mostly rely on almost frontal scans
  - Do not perform recognition across scans of different poses

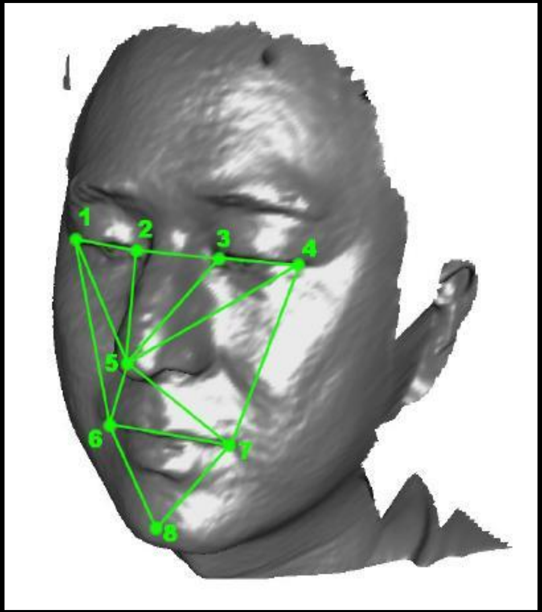
- Desired Attributes
  - Fully automatic
  - Interpose identification
  - Robust to missing data

Overview of Our Approach

- Model-based
  - 3D Facial Landmark Models (FLMs)
  - 3D Annotated Face Model (AFM), PAMI 2007
- Recognition process
  1. Preprocessing
     - Filter raw data
  2. 3D Landmark Detection & Pose Estimation
     - A novel landmark detector is used to estimate rough pose
  3. Registration
     - The facial data are registered to the AFM
  4. Deformable Model Fitting
  5. Wavelet Analysis

The Facial Landmark Models

- Eight anatomical landmarks
- Three Facial Landmark Models (FLMs)
  - FLM8 (Landmarks: 1,2,3,4,5,6,7,8): Visible on frontal datasets
  - FLM5R (Landmarks: 1,2,5,6,8): Visible on right facial datasets
  - FLM5L (Landmarks: 3,4,5,7,8): Visible on left facial datasets



1. Right eye outer corner
2. Right eye inner corner
3. Left eye inner corner
4. Left eye outer corner
5. Nose tip
6. Mouth right corner
7. Mouth left corner
8. Chin tip

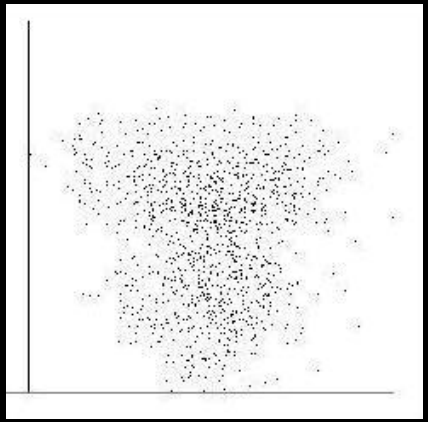
Construction of FLMs

- Statistical Mean Shape for each landmark set (FLM8, FLM5L, FLM5R)
  - Using a manually annotated training set of 150 frontal facial data sets with neutral expressions from FRGC v2, using Procrustes Analysis.

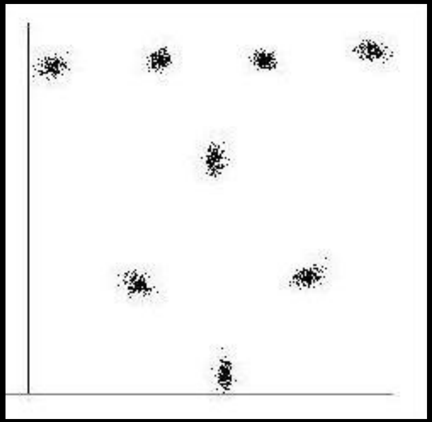
- Variations of each FLM
  - Landmark shape variations were computed by applying Principal Component Analysis (PCA) to the aligned training set.

Mean Landmark Shape (FLM8)

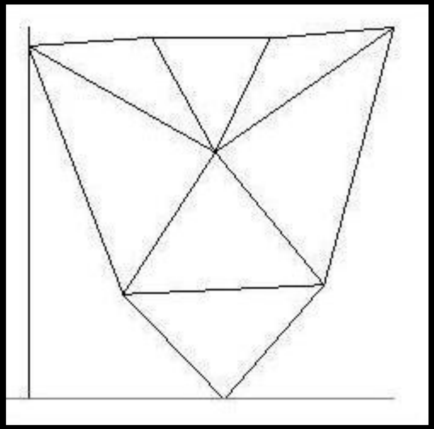
IEEE BTAS 09



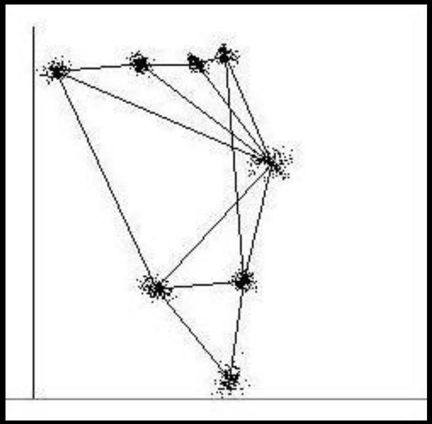
(a) Unaligned Landmarks



(b) Aligned Landmarks



(c) Mean Shape

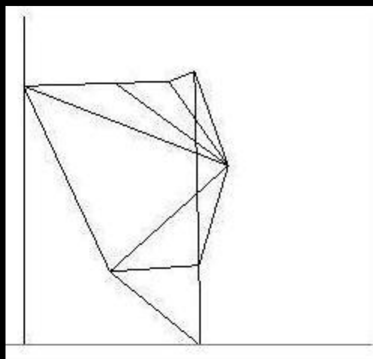


(d) Landmark Clouds & Mean Shape
(rotated by 60° around the y-axis)

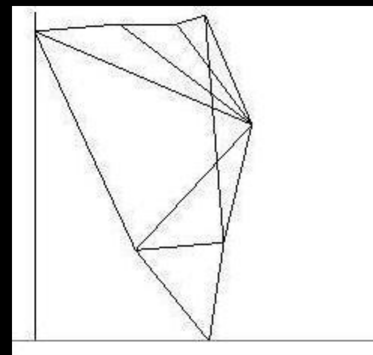
Landmark Shape Variations (FLM8)

IEEE BTAS 09

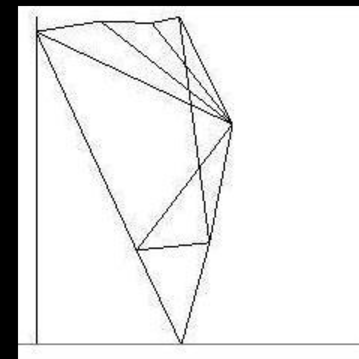
1st Mode of Variation (face shape: Round vs Oval)



$b_1 = -3\sqrt{\lambda_1}$

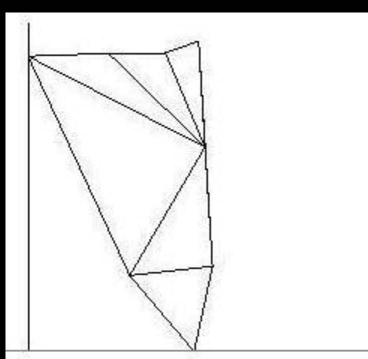


$b_1 = 0$

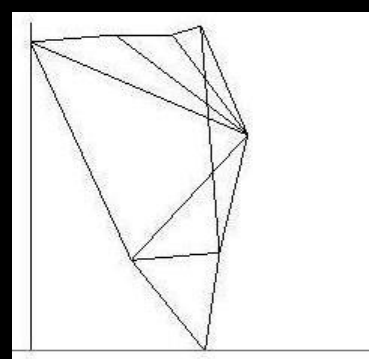


$b_1 = +3\sqrt{\lambda_1}$

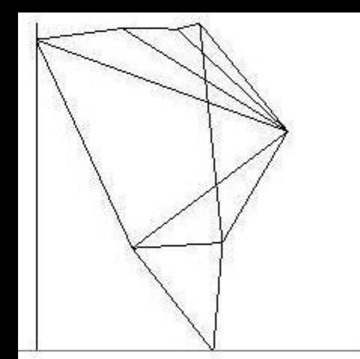
2nd Mode of Variation (nose shape: Flat vs Extruded)



$b_2 = -3\sqrt{\lambda_2}$

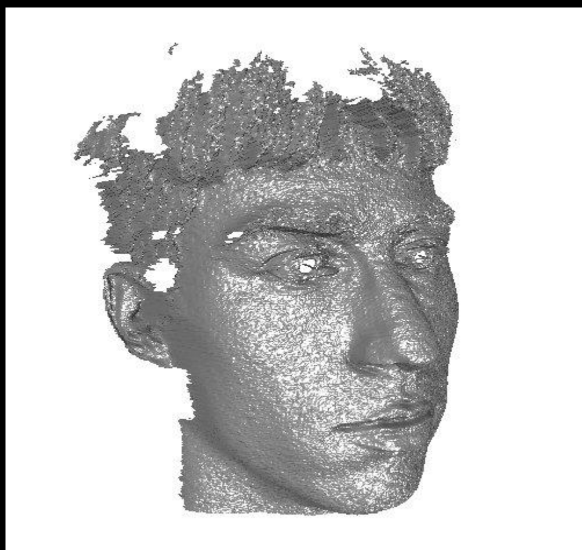


$b_2 = 0$



$b_2 = +3\sqrt{\lambda_2}$

1. Preprocessing



Scanner range-data

Polygonal data

Filters

- Median Cut
- Hole Filling
- Smoothing
- Subsampling

2. Landmark Detection & Pose Estimation

a. Extract candidate landmarks

b. Compute the rigid transformation that best aligns combinations of 5 and 8 landmarks with a Landmark Model (FLM5L, FLM5R, FLM8)

c. Discard combinations of candidate landmarks that are not consistent with the FLMs

d. Select the combination of landmarks with the minimum Procrustes distance from the FLMs

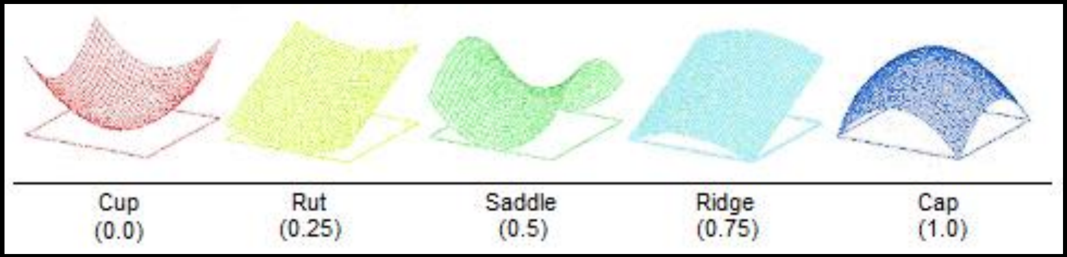
e. Label landmarks and estimate pose

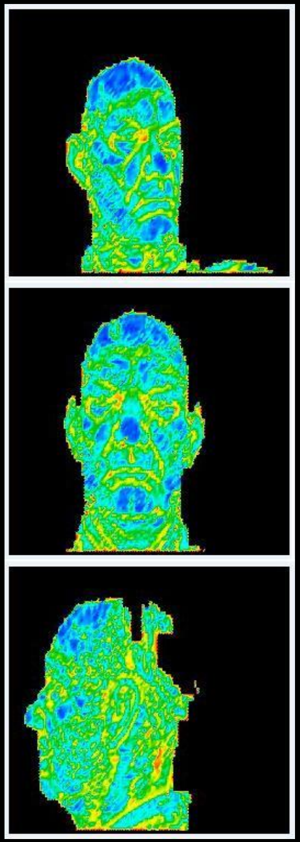
2.a Candidate Landmarks

- Shape Index represents the type of local curvature of a 3D object at a point **p**.

$$SI(p) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\frac{k_{max}(p) + k_{min}(p)}{k_{max}(p) - k_{min}(p)}$$

- To find candidate Landmarks:
  - Create Shape Index maps
  - Local minima (Cups) are candidate landmarks for eye and mouth corners
  - Local maxima (Caps) are candidate landmarks for nose and chin tips, further filtered by Extrusion map





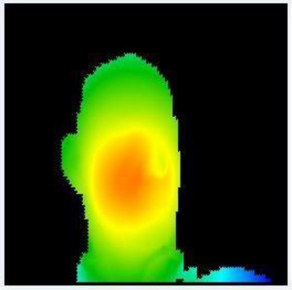
Shape Index maps of facial datasets

2.a Candidate Landmarks (2)

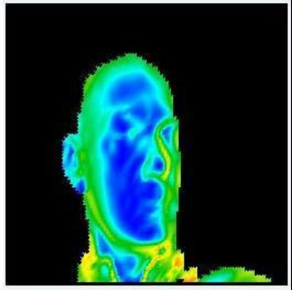
- Radial map
  - Represents the distance of a 3D point from the centroid (measure of radial vector)
- Tangent map
  - Represents the cosine of the angle between the normal vector at a 3D point and the radial vector from the centroid (tilt of normal vector)
- Extrusion map
  - Represents the product of the normalized values of the Radial and Tangent maps

Candidate Landmarks:

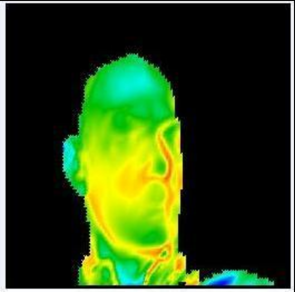
- Compute the Extrusion map
- Local maxima (most extruded points) that are also SI caps, are candidate landmarks for nose and chin tips



Radial map



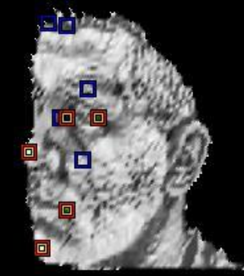
Tangent map



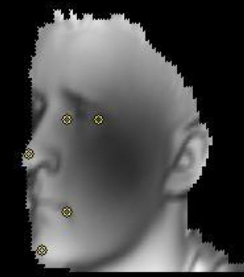
Extrusion map

2.b-d Resulting Landmarks

- Get combinations of 5 candidate landmarks, and create left and right sets of landmarks.
- Align left and right landmark sets to the FLMs.
- Filter out the landmark sets that do not conform with neither FLM5L nor FLM5R
- Fuse right and left consistent landmark sets in complete sets of 8 landmarks.
- Align complete landmark sets to the FLM8.
- The complete landmark sets that do not conform with FLM8 are discarded.
- The resulting landmark set is the one that has the minimum Procrustes distance to the corresponding model (FLM5R, FLM5L, FLM8).



Best Landmark sets (FLM5L, FLM5R), no consistent FLM8



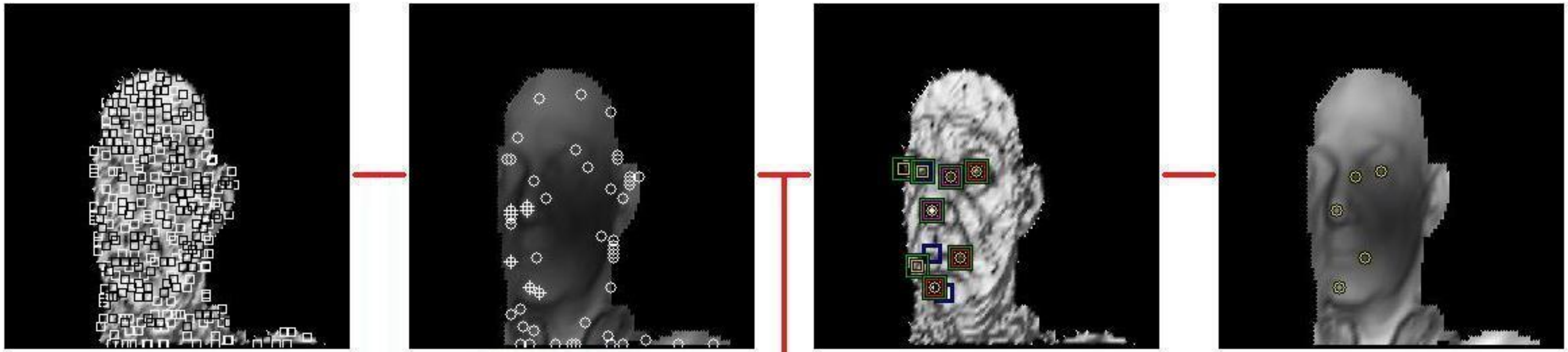
Resulting Landmark set (FLM5L)

Landmark Identification

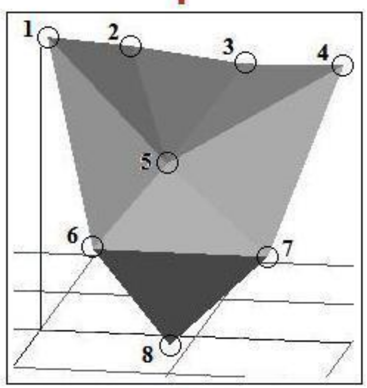
- A landmark set is considered as a plausible shape by checking the deformation parameters b to be within certain margins: $|b_i| \leq 3\sqrt{\lambda_i}$

- Fitting a candidate landmark set to the corresponding FLM
  - minimizing the Procrustes distance of a candidate landmark set from the corresponding FLM with a rigid transform

Resulting Landmarks (2)



(a) (b) (c) (d)



(e)

(a) Shape Index extrema
(Eye & Mouth corners)

(b) Extrusion map maxima
(Nose & Chin tips)

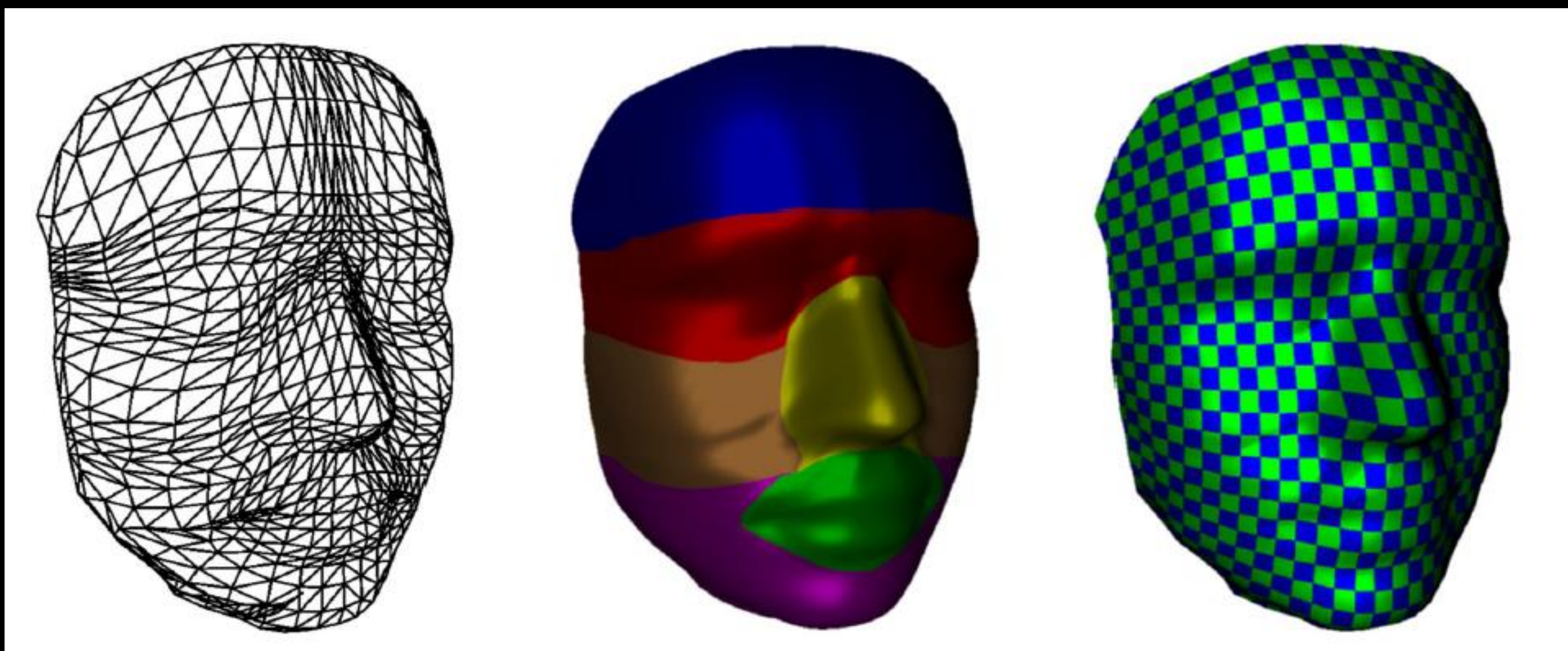
(c) Consistent Landmark Sets
(Left, Right, Frontal)

(d) Best Landmark Set

(e) Landmark Model

Annotated Face Model (AFM)

- 3D model of the human face
- Constructed only once
- Used for alignment, fitting and metadata generation



Polygonal Mesh | Annotated Areas | UV Parameterization

3. Registration

IEEE BTAS 09

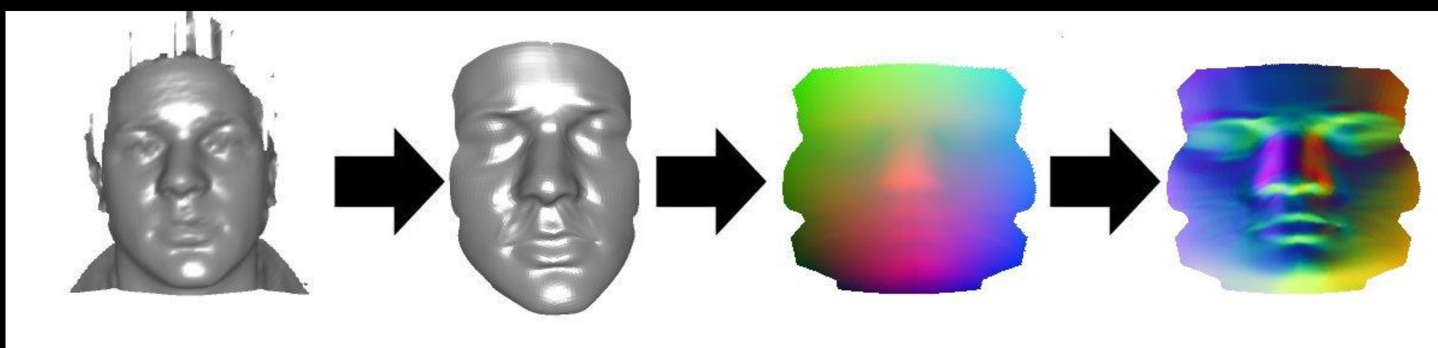
- Coarse alignment of facial data and the AFM using pose estimated from detected landmarks
- Tighter registration by using Simulated Annealing optimization on depth images
- Objective function (sum of z-buffer differences):

$$E = \sum_{i=1}^{R} \sum_{j=1}^{R} |D_{model}(i,j) - D_{data}(i,j)| \Rightarrow \omega = [t_x, t_y, t_z, \phi, \theta, \psi]$$

- This step fine-tunes the registration – it cannot alleviate errors caused by landmark detection failure
- For side scans only one half of the model's z-buffer is used in the objective function

4. Deformable Model Fitting

– For side scans symmetric fitting is used (to alleviate missing data)
– Fitting is performed by deforming the AFM



Raw Facial Data **Fitted AFM** **Geometry Image** **Normal Image**

– The AFM acquires the shape of the facial data
– The AFM is converted to a Geometry and Normal Image

Deformable Model Framework

BTAS 09

- Basic equation: $$\mathbf{M}_q \frac{d^2\vec{q}}{dt^2} + \mathbf{D}_q \frac{d\vec{q}}{dt} + \mathbf{K}_q \vec{q} = \vec{f}_q$$
- No motion: $\mathbf{M} = O$ and $\mathbf{D} = O$.
- External forces drive the deformation
- Internal forces resist the deformation
- The stiffness matrix determines the elastic properties:

$$\mathbf{K} = \mathbf{K}_{fo} + \mathbf{K}_{so} + \mathbf{K}_{sp}.$$

$$\Rightarrow$$

$$E_{fo} = \frac{1}{2}\kappa_{fo}\vec{q}^T \mathbf{K}_{fo}\vec{q}$$
$$E_{so} = \frac{1}{2}\kappa_{so}\vec{q}^T \mathbf{K}_{so}\vec{q}$$
$$E_{sp} = \frac{1}{2}\kappa_{sp}\vec{q}^T \mathbf{K}_{sp}\vec{q}$$

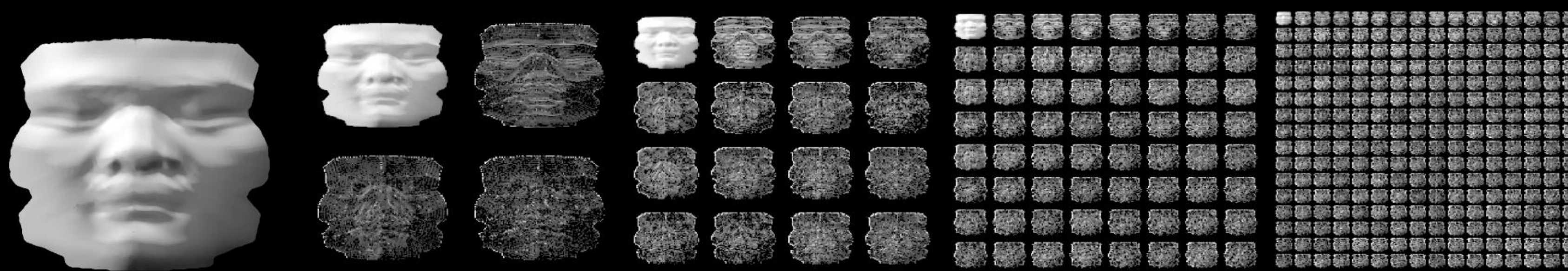
- Equations are solved by the subdivision-based Finite Element Method in an iterative way

5. Wavelet Analysis

- Walsh Wavelets Packet Decomposition of images
- 4 Haar Filters ($g^T\mathbf{x}g$, $g^T\mathbf{x}h$, $h^T\mathbf{x}g$, $h^T\mathbf{x}h$)

$$g = \frac{1}{\sqrt{2}}[1 \quad 1] \qquad h = \frac{1}{\sqrt{2}}[1 \quad -1]$$

- Level 4 decomposition: 256 (16x16) wavelet packets



Original Image | **1st Level Transform** | **2nd Level Transform** | **3rd Level Transform** | **4th Level Transform**

- Metadata represent the biometric signature
- Comparison: weighted L1 distance metric

Experimental Data

Datasets of common subjects from frontal scans from **FRGC v2** Database and side scans at 45° and 60° from **UND** Database

DB45LR: 119 subjects at 45°. Left scans as gallery, right as probe.

DB60LR: 88 subjects at 60°. Left scans as gallery, right as probe.

DB45F

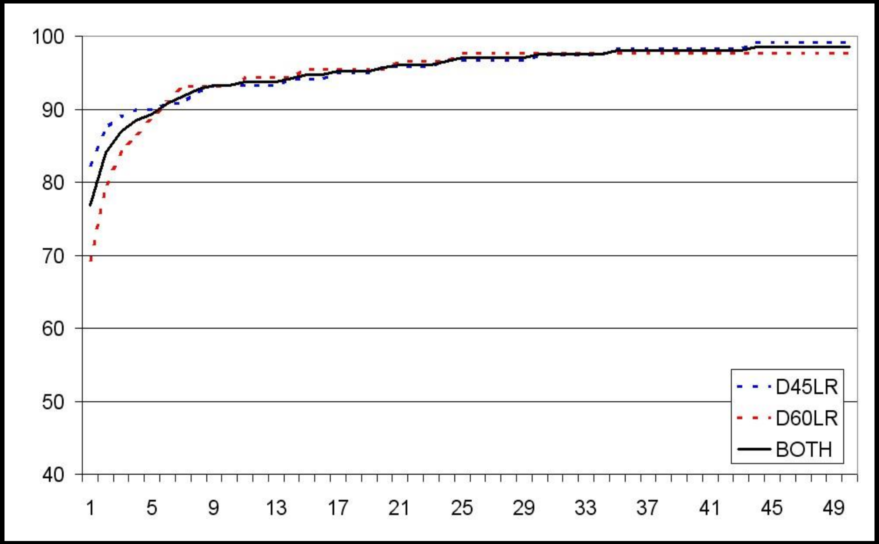
- Gallery: frontal scans of 466 subjects
- Probe: left and right scans at 45° of 39 subjects

DB60F

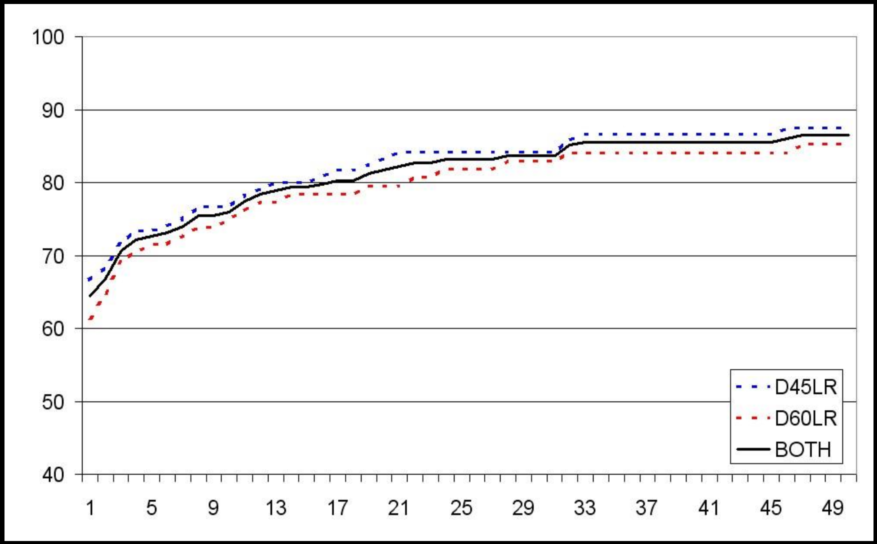
- Gallery: frontal scans of 466 subjects.
- Probe: left and right scans at 60° of 33 subjects.

Performance Evaluation

- CMC graphs for matching left side scans (gallery) with right scans (probe) using **DB45LR** and **DB60LR**



Automatic Landmarks



Manual Landmarks

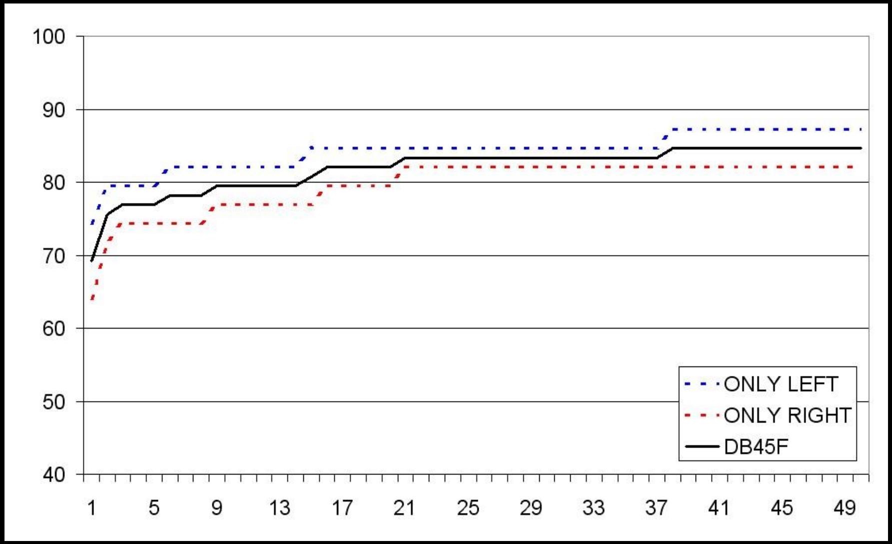
Rank-one rate:

DB45LR: 67% (automatic landmarks) and 82% (manual landmarks)
DB60LR: 60% (automatic landmarks) and 69% (manual landmarks)

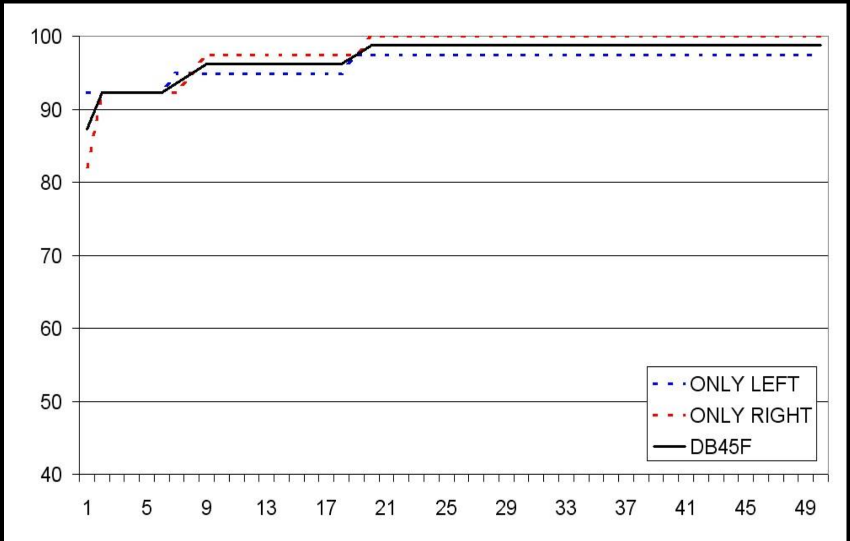
Performance Evaluation (2)

IEEE BTAS 09

- CMC graphs for matching frontal scans (gallery) with left and right scans (probe) using **DB45F**



Automatic Landmarks



Manual Landmarks

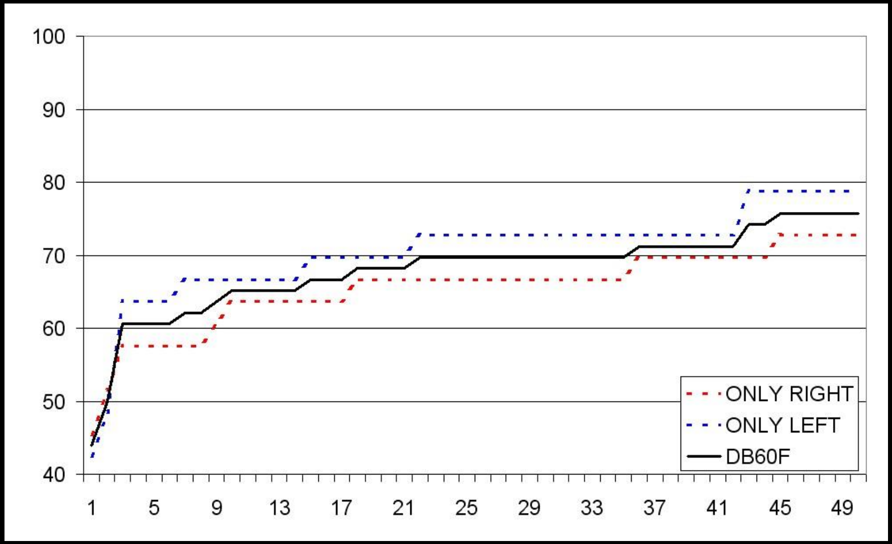
Rank-one rate:

DB45F: 69% (automatic landmarks) and 87% (manual landmarks)

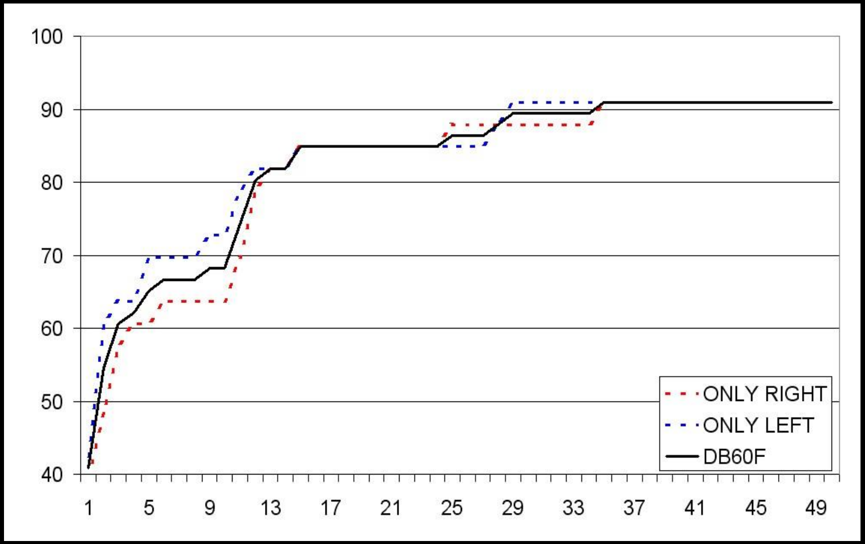
Performance Evaluation (3)

IEEE BTAS 09

- CMC graphs for matching frontal scans (gallery) with left and right scans (probe) using **DB60F**



Automatic Landmarks



Manual Landmarks

Rank-one rate:

DB60F: 44% (automatic landmarks) and 41% (manual landmarks) !

Performance Evaluation (4)

- Manually placed landmarks improve the recognition rate by approximately 10%
  - 10% is the approximate rate of total failures for the automatic landmark detector.
- The 60° side scans yield lower results than the 45°
  - This is due to the fact that 60° side scans have more missing data
- 60° left vs 60° right performs better than 60° left/right vs frontal
- Left scans generally performed better than right scans
  - Probably due to quality of scans

Conclusion

– Fully automatic
– Can handle up to 80° yaw rotations
– Can handle partial data (even if half of the face is missing).

Contact

Ioannis Kakadiaris
ioannisk@uh.edu