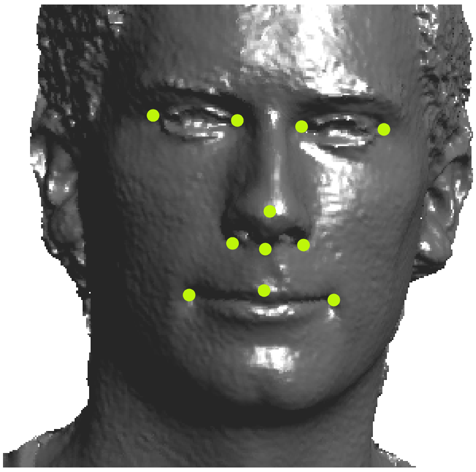
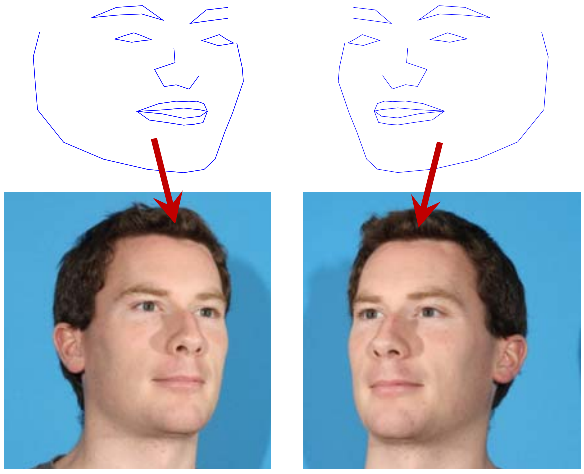
A quantitative assessment of 3D facial key point localization fitting 2D shape models to curvature information

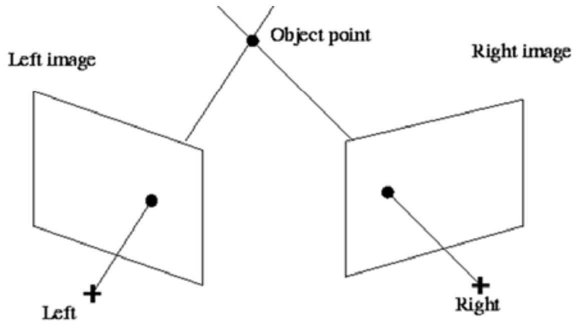
F. Sukno, T. Chowdhury, J. Waddington and P. Whelan

Dublin City University and
Royal College of Surgeons in Ireland

3D landmarks from 2D models



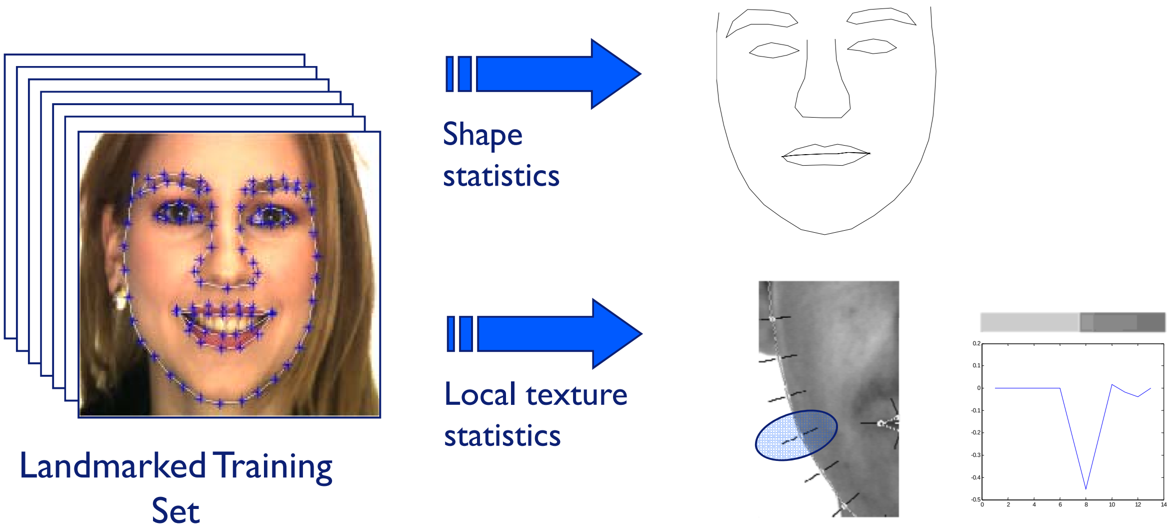




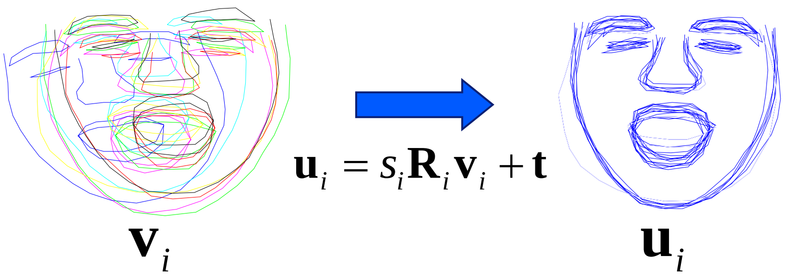
- 3D facial landmarks usually detected based on curvature information
- High quality texture is frequently available
- Potential for exploiting 2D methods for landmarks detection

Face Segmentation

- Based on Active Shape Models (ASMs)
  - Face outlines based on landmarks
  - Shape statistics to learn spatial relations
  - Texture statistics for image search



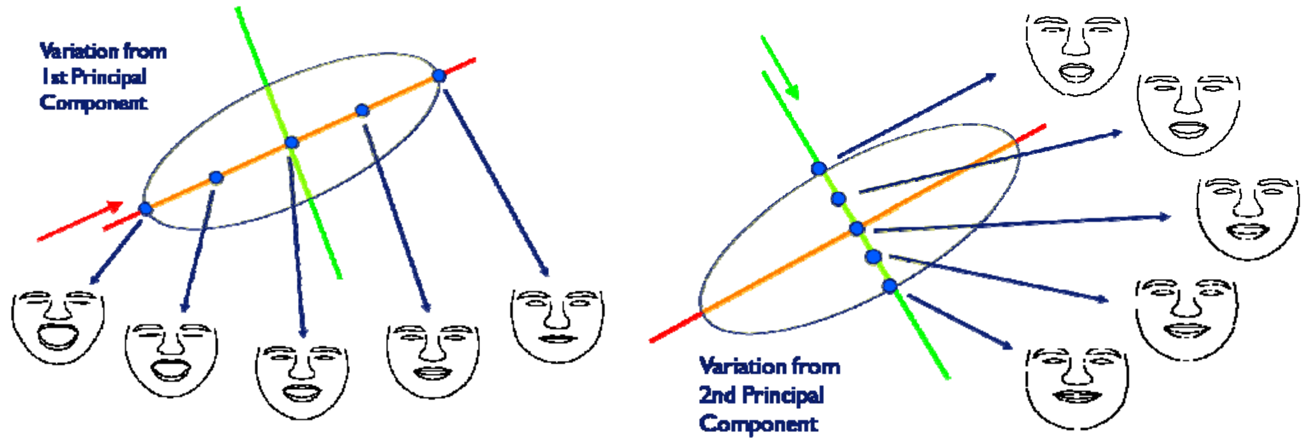
Shape statistics: Point Distribution Models



$$\mathbf{u}_i = s_i \mathbf{R}_i \mathbf{v}_i + \mathbf{t}$$

Image Coordinates

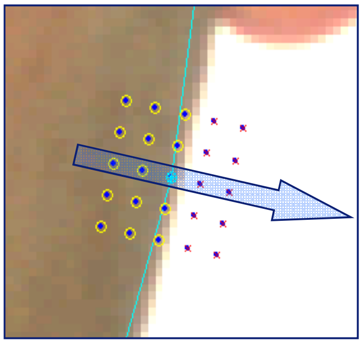
Model Coordinates

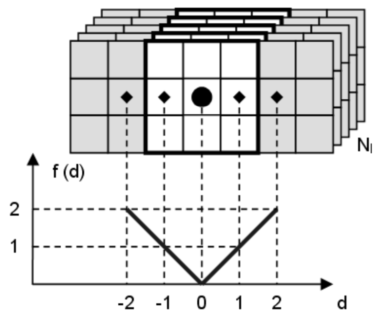


Invariant Optimal Features ASM

Sukno et al. (2007) IEEE Trans. Pattern Anal. Mach. Intell.

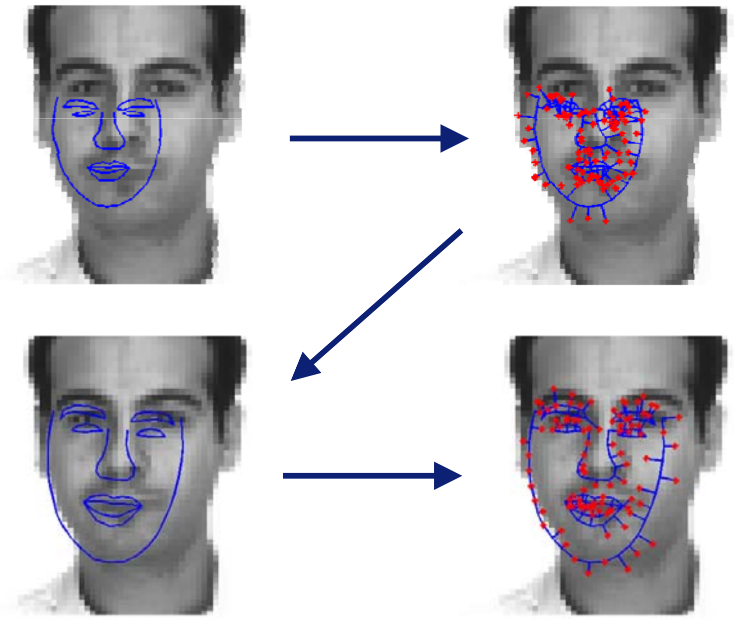
- Texture description based on differential invariants
  - Computed on a neighborhood of each landmark
- Non-linear classification
  - Multi-valued neurons
  - Separate classifier for each landmark
  - Allows for feature selection
- Unified profile shapes
  - Robust decisions
- Increased localization accuracy
  - 30% lower error than standard ASM

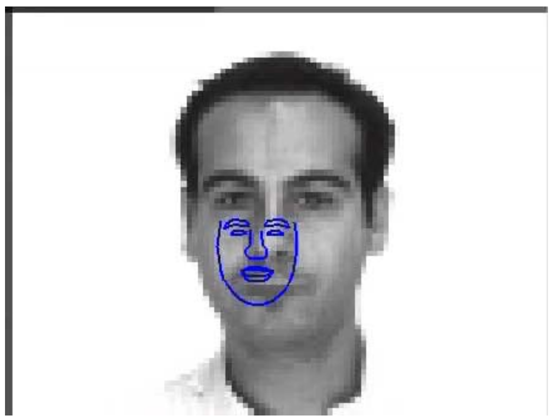




ASM: Model-to-image adaptation

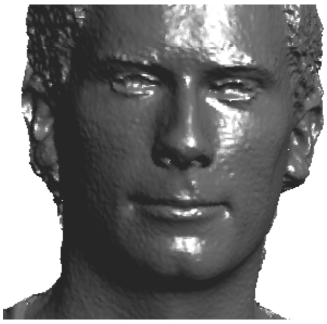
- Iterative process alternating
  - Local image search (best local displacement for each landmark)
  - PDM shape constraints (enforce global consistency of the shape)

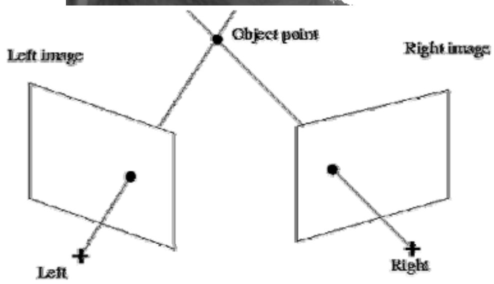




Example of model-to-image adaptation

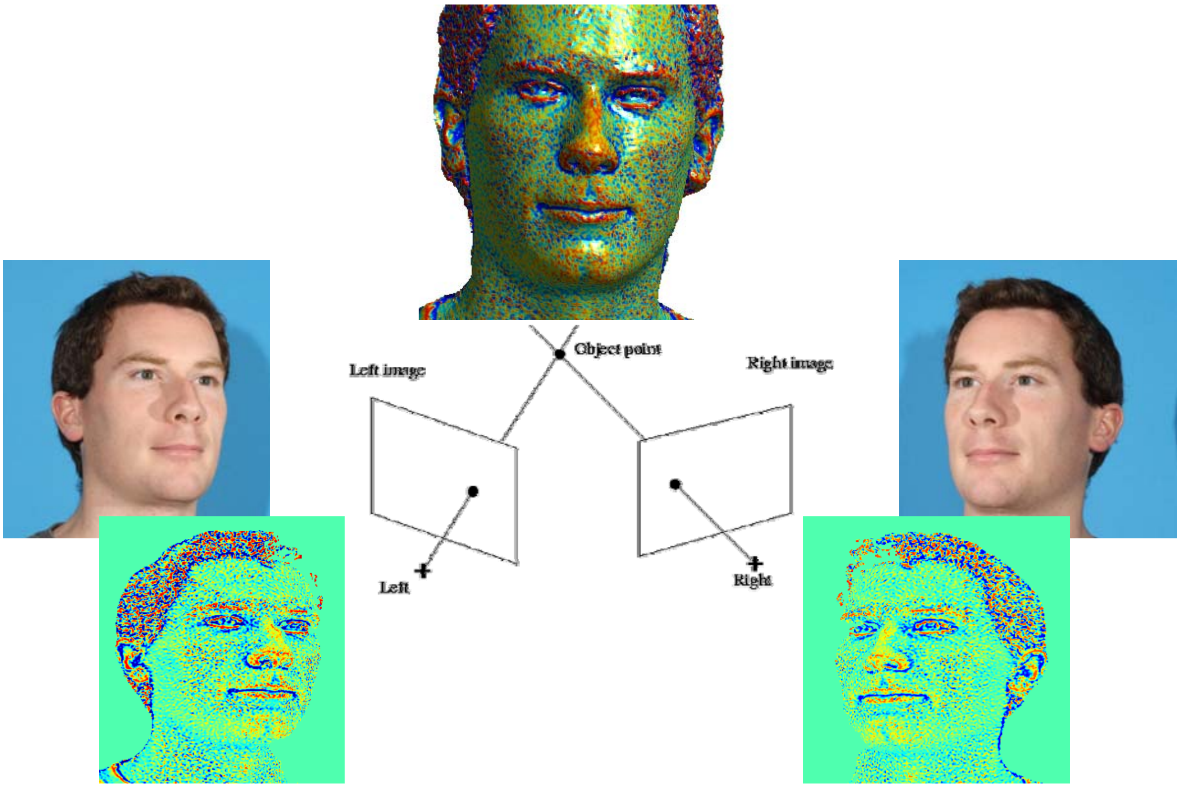
Mapping 3D features into 2D





- A mapping between the 2D views and the 3D mesh is available
  - Any features computed in 3D can be mapped back into the 2D views

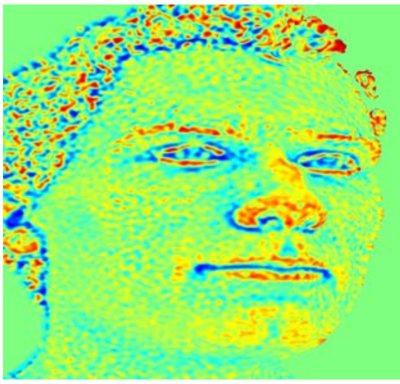
Comparing texture and curvature



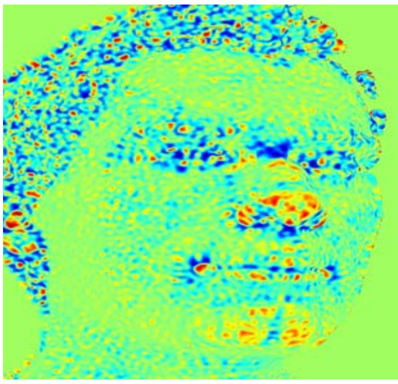
Curvature computation

Meyer et al. (2002) Int. Workshop on Visualization and Mathematics.

- Following the work by Meyer et al (2002)
  - Derivation of first and second order differential properties using averaging Voronoi cells + the curvature sign
  - Only mean curvature was used, as it was found considerably less noisy than Gaussian curvature



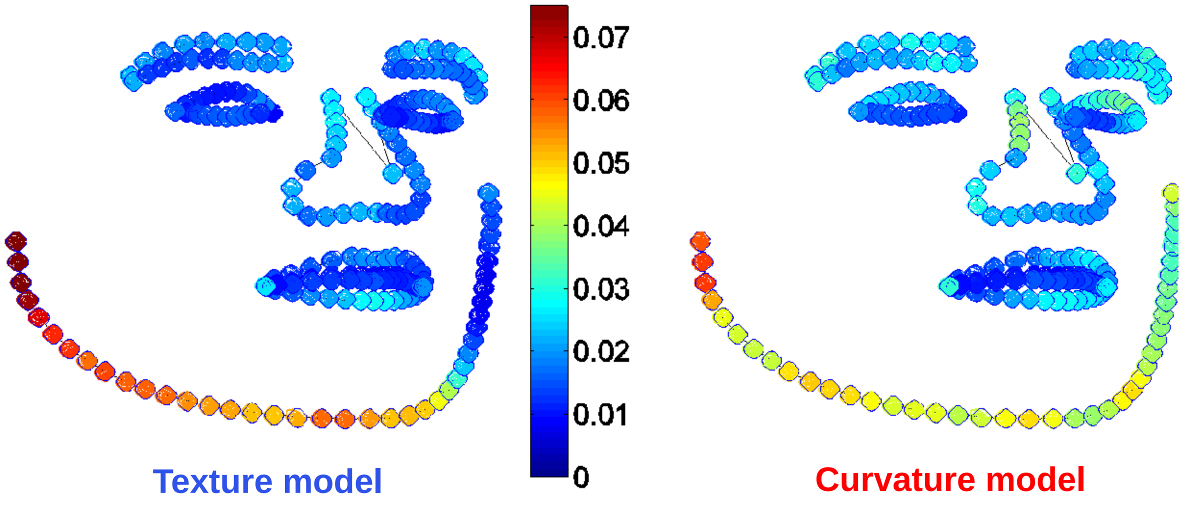
Mean curvature



Gaussian curvature

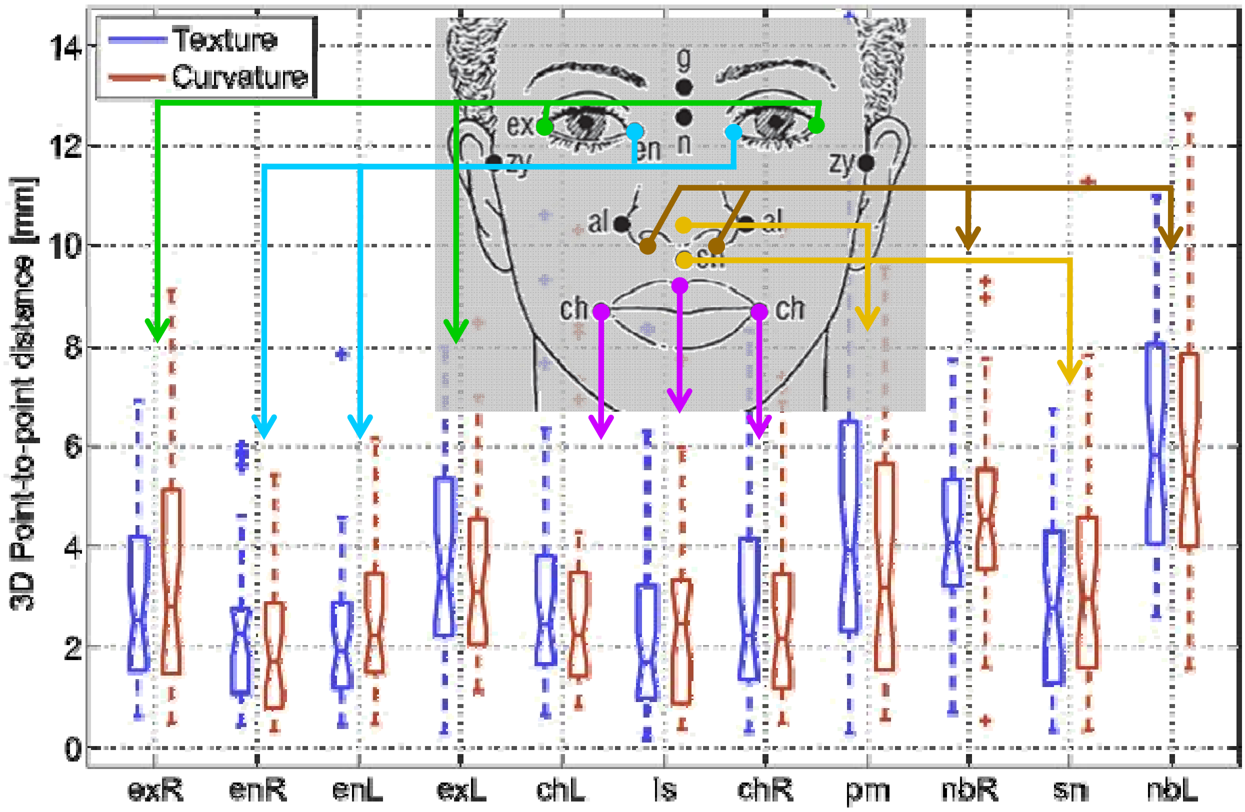
Model-to-image adaptation results

- Tests on a database of 34 facial scans from
  - 16 people, 61% neutral expression, 88% frontal pose
  - All experiments performed in 3-fold cross validation

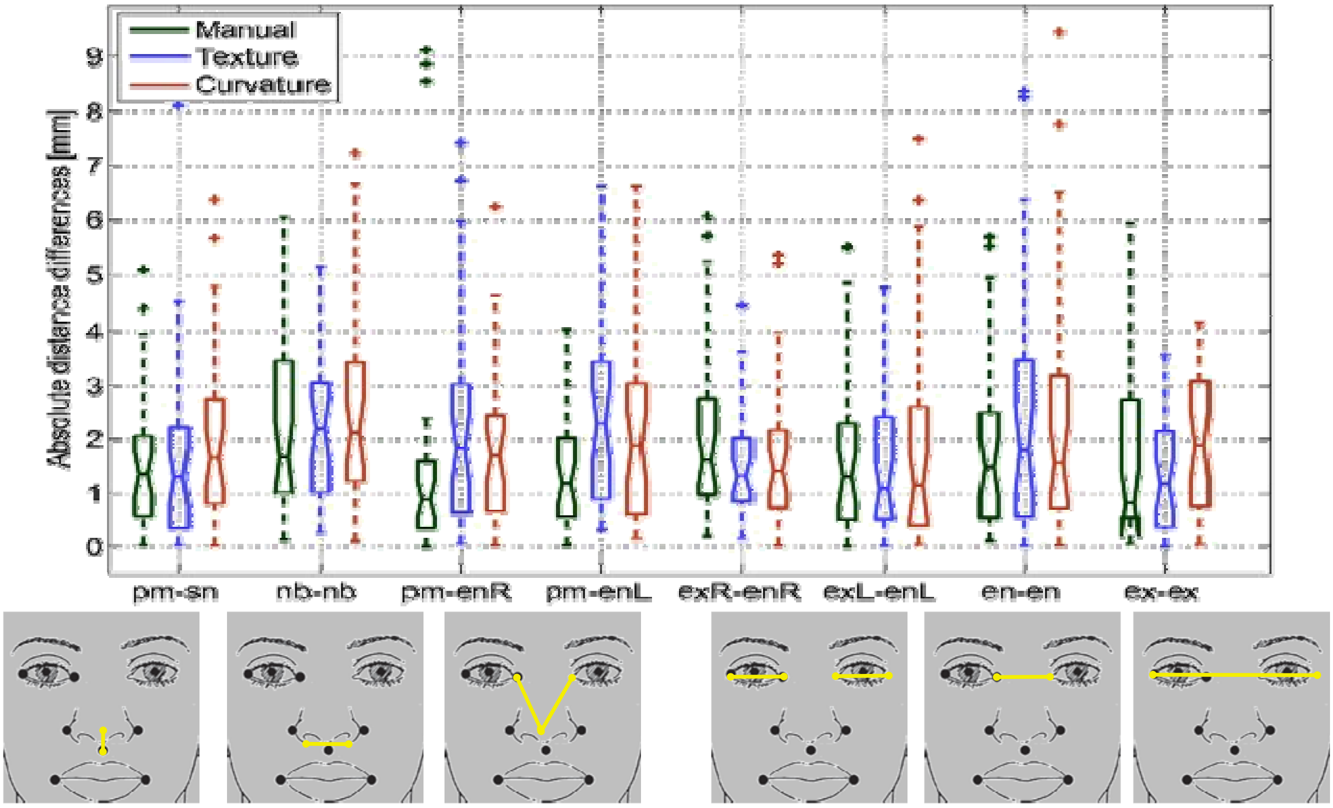


Point-to-curve errors normalized by the inter-ocular distance

3D Landmarks: distance to ground truth



Reproducibility of inter-landmark distances



Best reported averages on landmark localization errors [mm]

Method	ch	en	ex	ls	sn	nb / al	pm
IOF-ASM (texture)	3.2	2.3	3.3	2.4	2.9	5.1	4.9
IOF-ASM (curvature)	3.0	2.3	3.5	2.5	3.5	5.3	4.7
D'Hose et al. [4]	-	-	-	-	-	-	3.17
Lu and Jain [12]	6.1	8.05	9.9	-	-	-	6.1
Perakis et al. [16]	6.03	5.31	5.76	-	-	-	4.88
Segundo et al. [22]	-	3.52	-	-	-	5.34	1.87
Szeptycki et al. [25]	8.56	3.85	2.82	-	-	6.18	2.27
Yu et al. [29]	-	5.17	-	-	-	-	2.14
Zhao et al. [31]	3.93	3.21	4.27	2.72	-	4.47	2.68

- Results are encouraging for mouth and eyes corners and upper-lip centre / The nose tip is considerably less accurate
- The size of the database is comparatively small, hence the quantitative evaluation is only preliminary

Reported precision on manual measurements of inter-landmark distances [mm]

Method	en-en	en-ex	ex-ex	en-pm	nb-nb	al-al	pm-sn
Section III-C (different images)	3.0	1.8	2.0	2.2	2.8	-	1.8
Ainechi et al. (direct vs image) [1]	0.35	0.09	0.54	-	-	0.35	0.80
De Menezes et al. (direct vs image) [3]	-	-	0.62	-	-	-	0.28
Ghoddous et al. (direct vs image) [6]	5.0	-	0.6	-	-	-	2.6
Heike et al. (inter-observer) [8]	0.85	1.89	2.09	-	-	0.88	0.88
Wong et al. (direct vs image) [27]	1.0	-	0.5	-	-	0.8	0.7

- Reported values show great variability
  - Typically on very small datasets (N < 20)
  - Some studies use visible markers on the facial surface (as their aim is to compare direct measurements with image-based measurements)
  - The repeatability of manual measurements seems to be between 1mm and 2mm (for the most accurate landmarks)

Conclusions

- 2D models can be used to localize 3D facial landmarks even when no texture is available
  - Curvature-based information could be projected into synthesized 2D views
- The IOFASM based segmentation in a population of 34 facial scans
  - Showed state of the art results for mouth and eye points
    - Averages approximately in the range of 2 mm – 3 mm
  - Was considerably less accurate for nose points

The Face 3D project

Supported by wellcome trust

- Research into the analysis of three-dimensional facial dysmorphology
  - Relation with mental disorders of developmental origin
  - Application to surgical reconstruction
  - Need for highly accuracy landmark localization

Face3D RESEARCH CONSORTIUM