THE OHIO STATE UNIVERSITY

Applications of Elastic Functional and Shape Data Analysis

Sebastian Kurtek

Department of Statistics

The Ohio State University

Collaborators: Anuj Srivastava, Eric Klassen (Florida State University), Karthik Bharath (University of Nottingham), Ian Jermyn (Durham University), Hamid Laga (Murdoch University), Chafik Samir (University of Clermont), Shantanu Joshi (UCLA), Hassen Drira (Lille 1 University), Veera Baladandayuthapani (MD Anderson Cancer Center), Arvind Rao (University of Michigan), Qian Xie (Amazon), Gary Christensen (University of Iowa), Zhaohua Ding (Vanderbilt University), many more!

Quick Outline:

1. Functional Data
2. Shapes of Curves
3. Shapes of Surfaces

The Ohio State University

BAYESIAN REGISTRATION MODEL

Bayesian Model + Riemannian Geometry + Importance Sampling

(1) Bayesian Alignment Model:

- Allows comprehensive exploration of the variable space.
- Provides credible intervals of warping function estimates.
- Allows discovering multiple registration solutions through multimodal posteriors.

(2) Riemannian Geometry of Warping Group:

- Enables efficient Riemannian computation on the space of warping functions.
- Allows a geometric prior distribution on the space of warpings.

(3) Importance Sampling:

- Allows efficient sampling from the posterior distribution.

Kurtek, Electronic Journal of Statistics, 2017; Lu et al., Journal of Computational and Graphical Statistics, 2017

BAYESIAN REGISTRATION MODEL

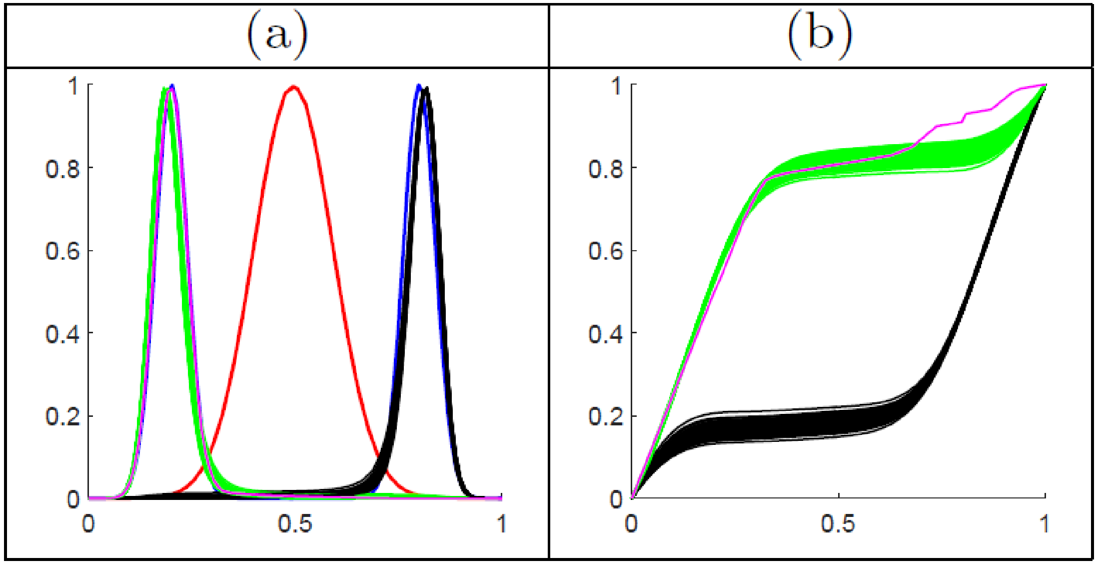
THE OHIO STATE UNIVERSITY

- Function discretization: $q([t]) = \{q(t_1), q(t_2), \ldots, q(t_N)\}$.
- Likelihood: $q_1[t] - q_2^*[t] \mid \psi \sim MVN(0_N, \Sigma_{N\times N}) \quad \Sigma_{N\times N} = 1/(2\kappa) I_{N\times N}$
- Prior on precision parameter: $\kappa \sim Gamma(1, 0.01)$
- Prior on warping functions: $\psi \sim TWN_A(1, \Sigma_\psi)$
- The prior is defined in the tangent space of the identity warping function for regularization, and wrapped onto the sphere.
- The prior is truncated to the positive orthant of the Hilbert sphere.
- We use Fourier-type basis in the tangent space. Other basis choices also available.
- Truncated Wrapped Normal:

$$\pi(\psi|1, K) \propto \exp\left(-0.5 \exp_1^{-1}(\psi) B^T K^{-1} B \exp_1^{-1}(\psi)\right) \mathbf{1}_{(\exp_1^{-1}(\psi)\in A)}$$

- K is diagonal with variances decaying as a function of the basis index.

SIMULATED EXAMPLE

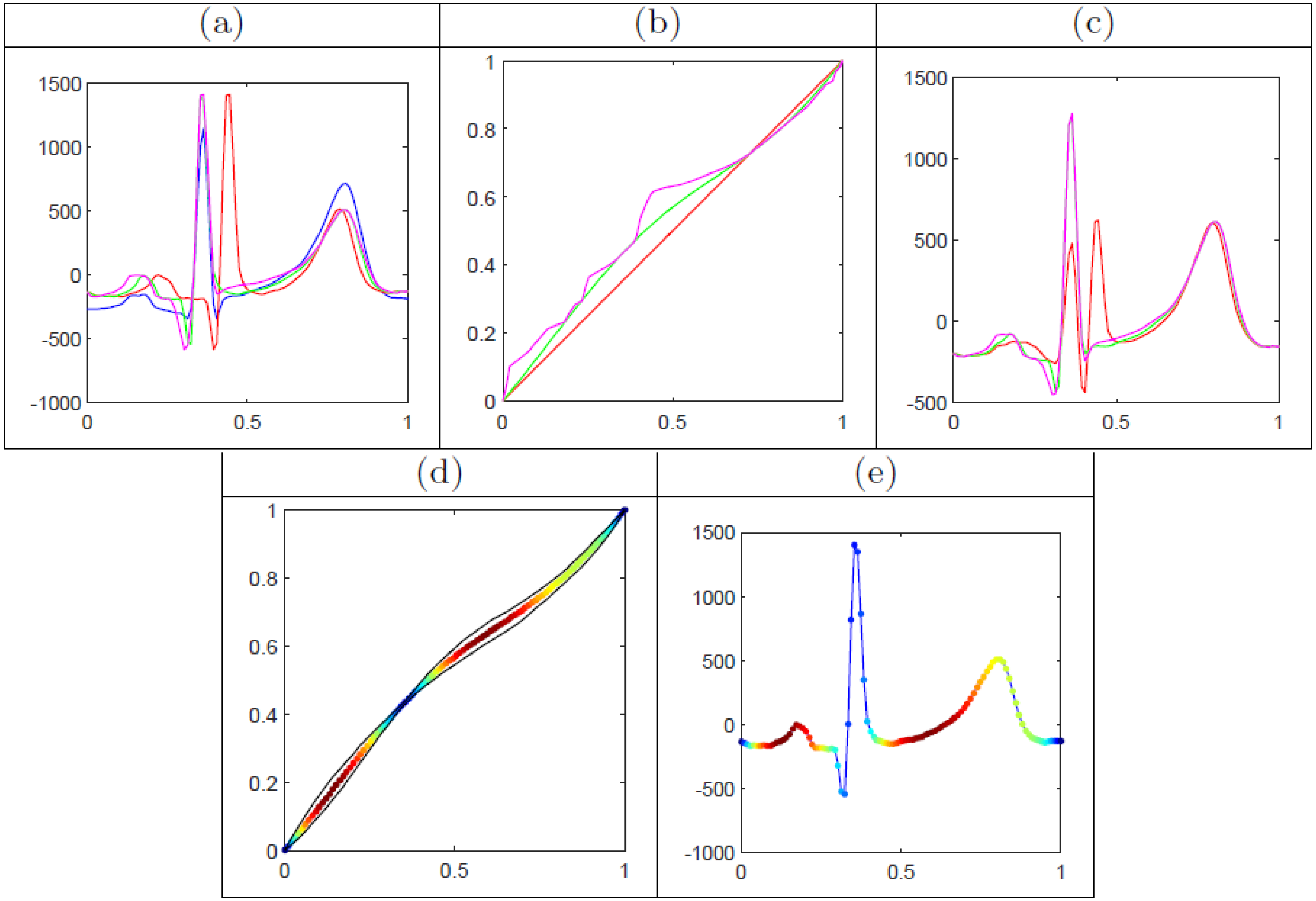


(a) Blue: f_1. Red: f_2. Magenta: f_2 after Dynamic Programming registration. Green and Black: f_2 after registration based on mean warping functions coming from the two posterior modes.

(b) Magenta: Dynamic Programming solution. Green and Black: mean warping functions corresponding to two posterior modes.

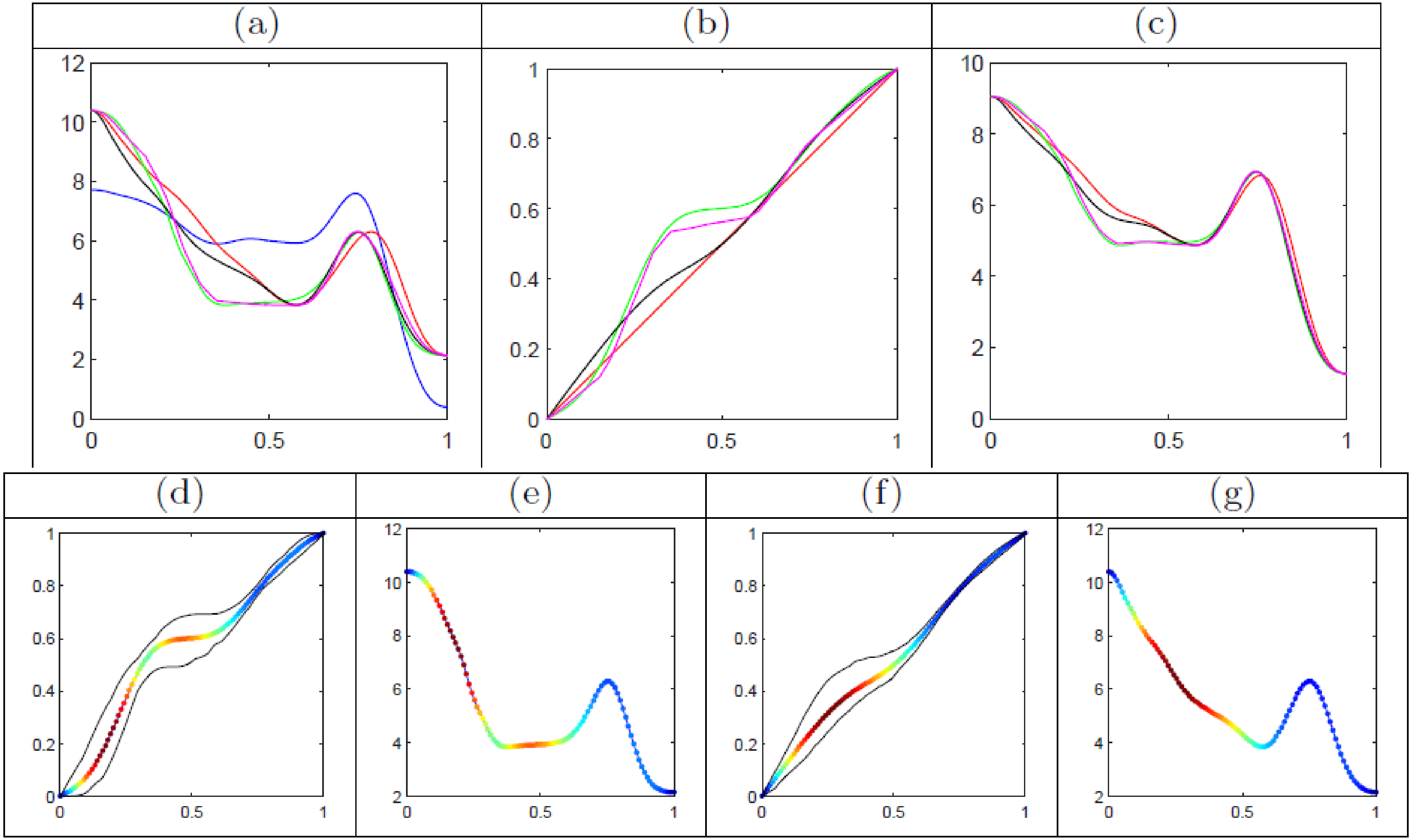
Note: Samples from each posterior mode are identified using intrinsic k-means clustering on the space of warping functions.

ECG PQRST CYCLES



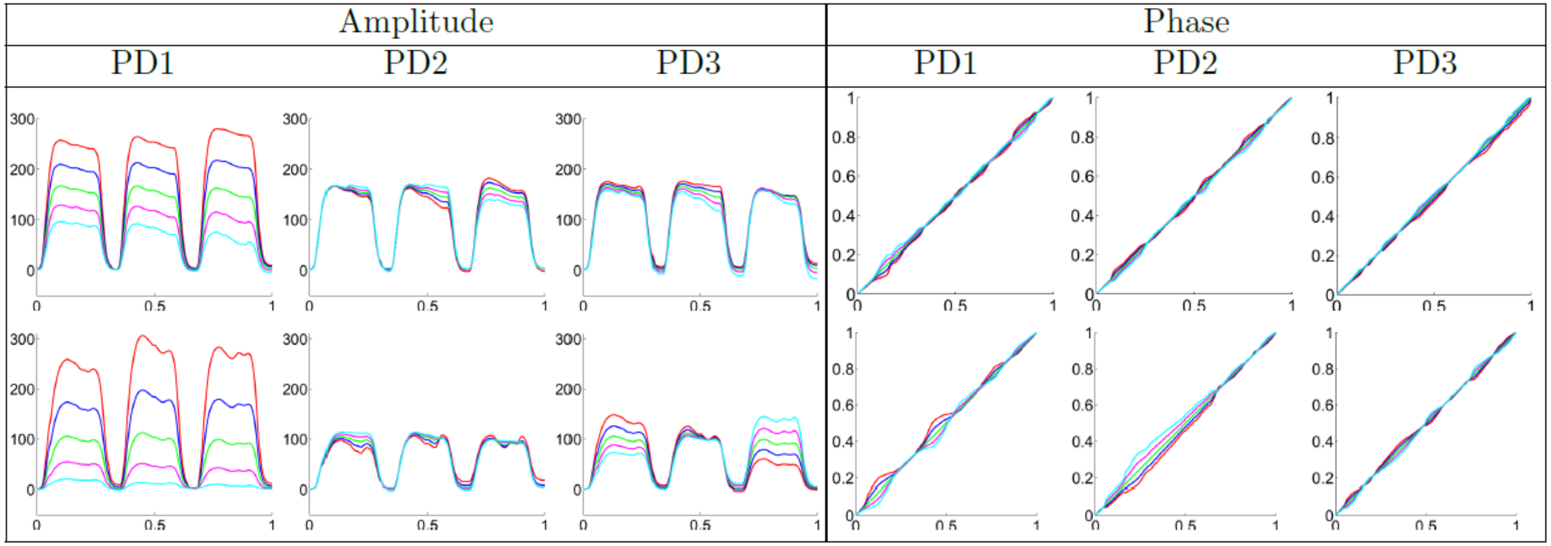
- The posterior is unimodal. There is very little uncertainty in the alignment of the large R peak.
- The posterior mean is much smoother than the Dynamic Programming solution and provides approximately the same alignment.

BERKELEY GROWTH DATA



- Two modes in the posterior reveal different patterns of growth dynamics.

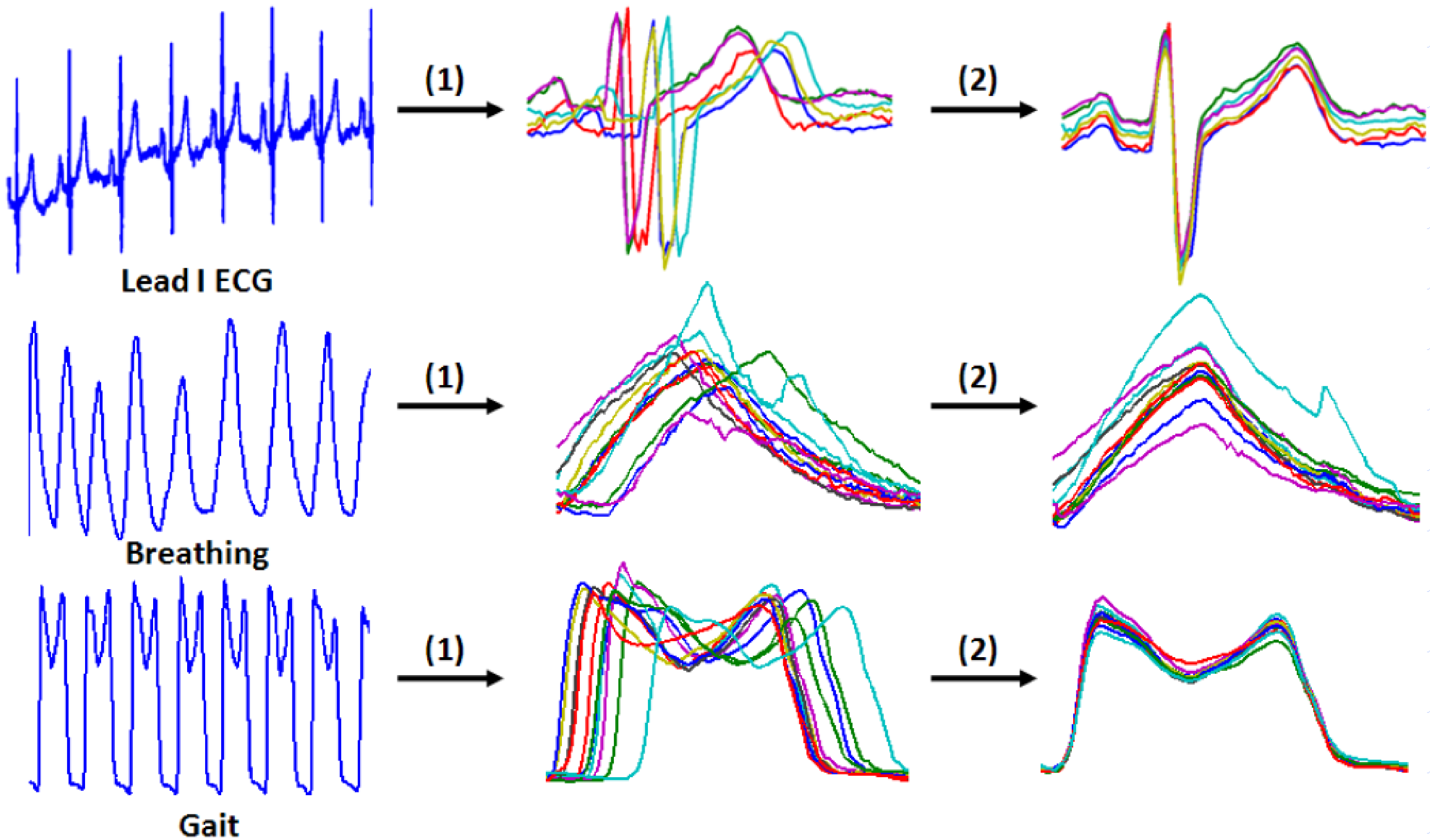
RHEUMATOID ARTHRITIS DIAGNOSIS



- Top: controls. Bottom: Rheumatoid Arthritis (RA) patients.
- Data: hand force functions for 50 subjects (20 controls, 30 RA patients).
- Functional diagnostic features: (1) first three amplitude and phase fPCA coefficients in control and RA groups, (2) norm of the estimated warping functions, (3) phase and amplitude distances to the control and RA means.
- Results: (1) Improved classification accuracy: 98% vs. 90% using classical features.
  (2) A sophisticated yet cost-effective, quick and easy diagnostic tool for determining hand impairment and function.

Samir et al., IEEE WACV, 2016

STATISTICAL ANALYSIS OF BIOSIGNALS



Kurtek et al., Journal of Applied Statistics, 2014

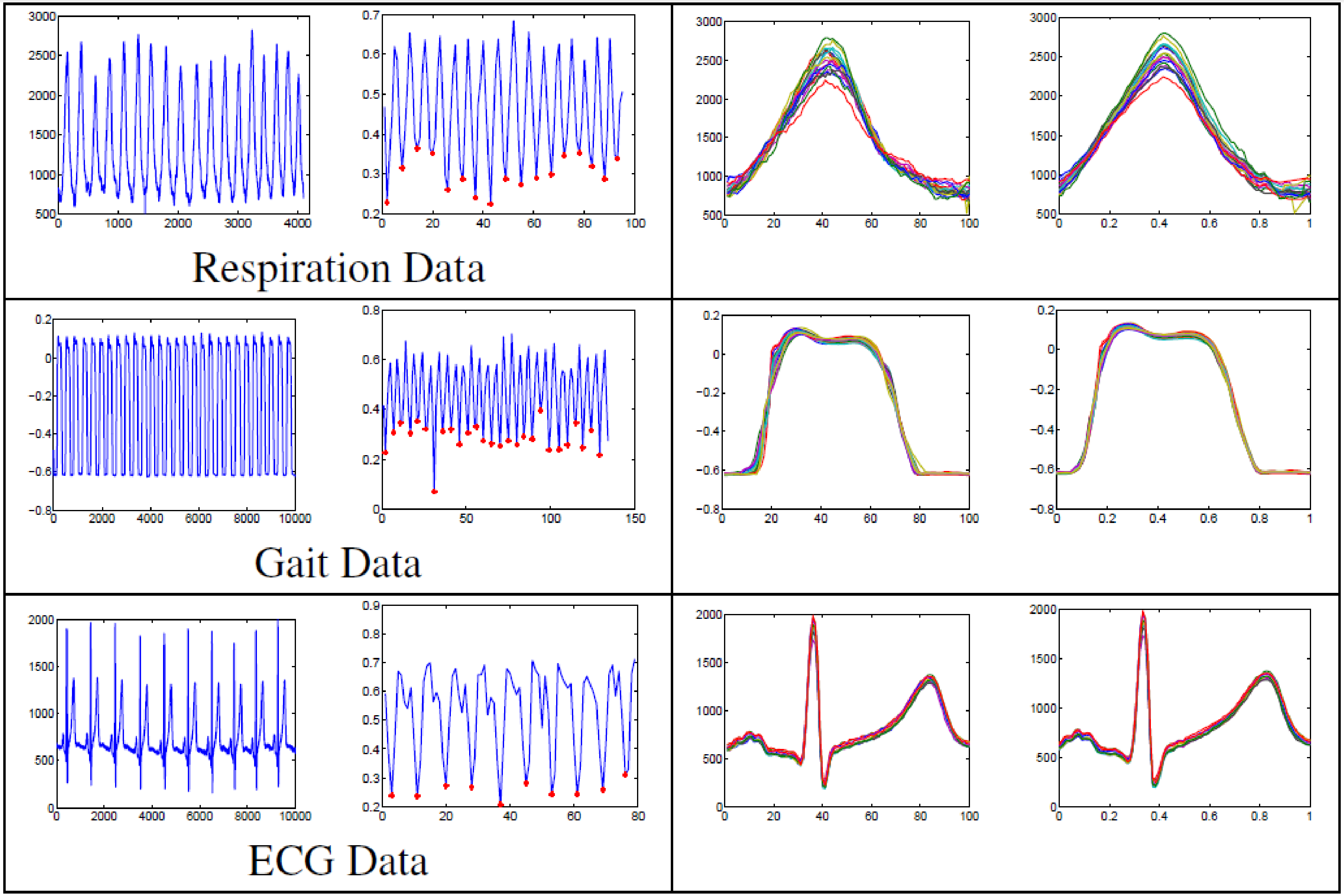
SEGMENTATION ALGORITHM

- Based on Fisher-Rao distance between warping functions. Robust to differing cycle scales.

Algorithm: Given a long cyclic signal and a cycle template.

1. Initialize with template at beginning of signal.
2. Compute optimal warping function by matching the template to overlapping signal.
3. Compute distance of optimal warping from identity warping.
4. Slide template along signal and repeat steps 2 and 3.
5. Resulting cost function is periodic. Extract minima and corresponding signal cycles.

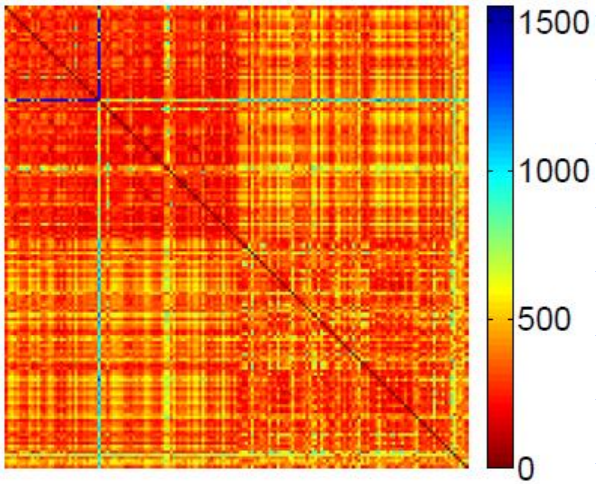
SEGMENTATION RESULTS

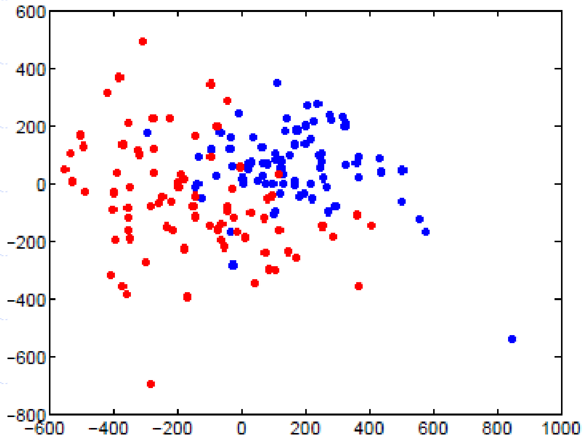


MYOCARDIAL INFARCTION DETECTION

The Ohio State University

Classifier Lead	Proposed Method SE(%)	SP(%)	AC(%)	$\mathbb{L}^2$ Distances SE(%)	SP(%)	AC(%)
I	71.25	86.25	78.75	58.75	56.25	57.5
II	73.75	75.00	74.38	70.00	53.75	61.88
III	72.50	91.25	81.88	55.00	67.50	61.25
aVR	77.50	90.00	83.75	63.75	63.75	63.75
aVL	77.50	92.50	85.00	52.50	65.00	58.75
aVF	61.25	88.75	75.00	61.25	58.75	60.00
V1	66.25	75.00	70.63	56.25	65.00	60.63
V2	68.75	91.25	80.00	58.75	66.25	62.50
V3	80.00	83.75	81.88	70.00	71.25	70.63
V4	76.25	83.75	80.00	68.75	62.50	65.63
V5	80.00	93.75	86.88	87.50	77.50	82.50
V6	78.75	82.50	80.63	72.50	61.25	66.88
VX	87.50	86.25	86.88	72.50	63.75	68.13
VY	70.00	83.75	76.88	55.00	60.00	57.50
VZ	78.75	81.25	80.00	66.25	57.50	61.88
Comb.	**83.75**	**96.25**	**90.00**	**72.50**	**62.50**	**67.50**

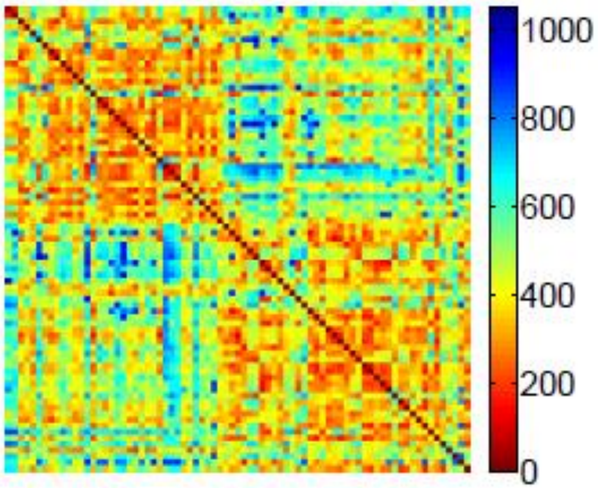


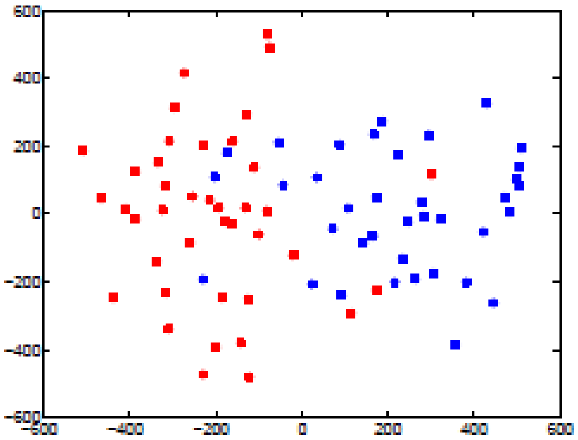


THE OHIO STATE UNIVERSITY

MYOCARDIAL INFARCTION LOCALIZATION

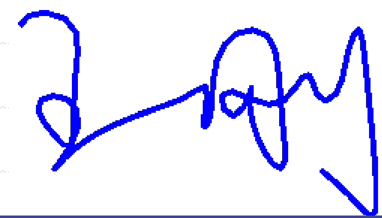
Classifier Lead	Proposed Method AC(%)	$\mathbb{L}^2$ Distances AC(%)
I	59.74	58.44
II	66.23	46.75
III	71.43	51.95
aVR	54.55	50.65
aVL	70.13	46.75
aVF	74.03	48.05
V1	71.43	67.53
V2	88.31	74.03
V3	90.91	76.62
V4	76.62	74.03
V5	53.25	55.84
V6	53.25	62.34
VX	71.43	54.45
VY	66.23	61.04
VZ	77.92	61.04
Comb.	**92.21**	**77.92**

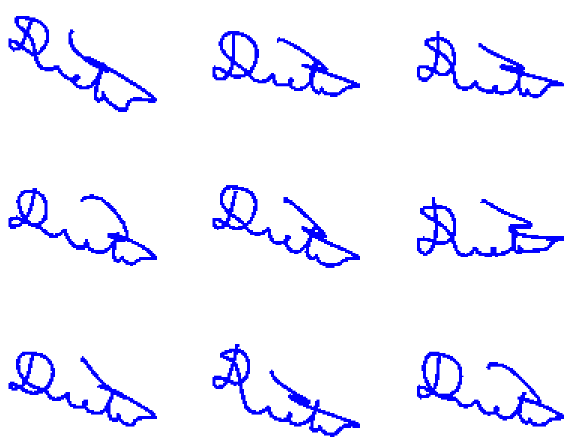




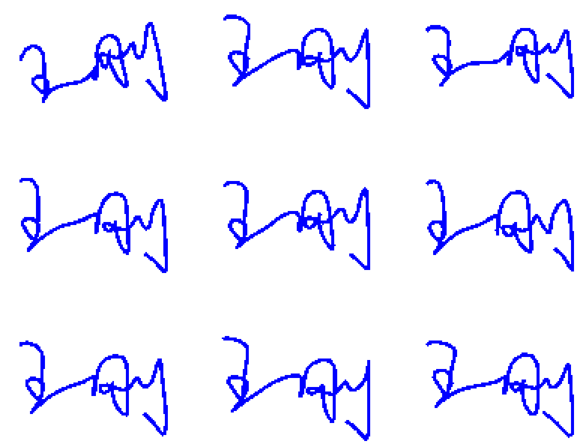
SIGNATURE STATISTICS

Mean Signature





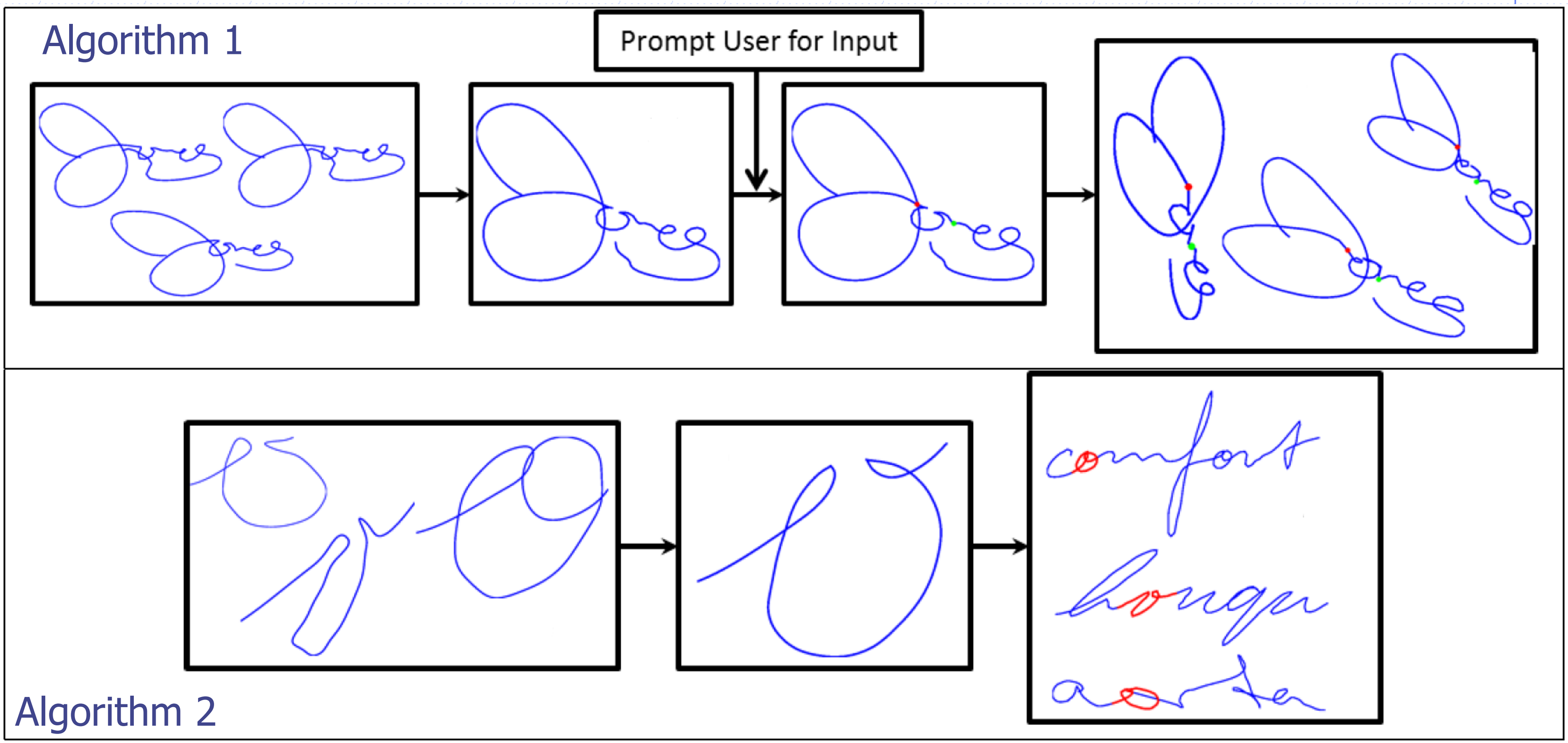
Random Signature Samples



Kurtek and Srivastava, ICPR, 2014

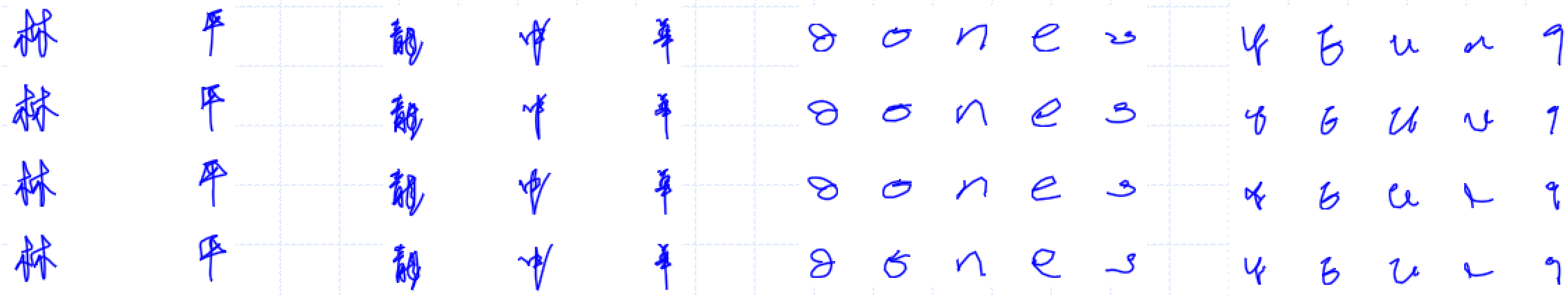
THE OHIO STATE UNIVERSITY

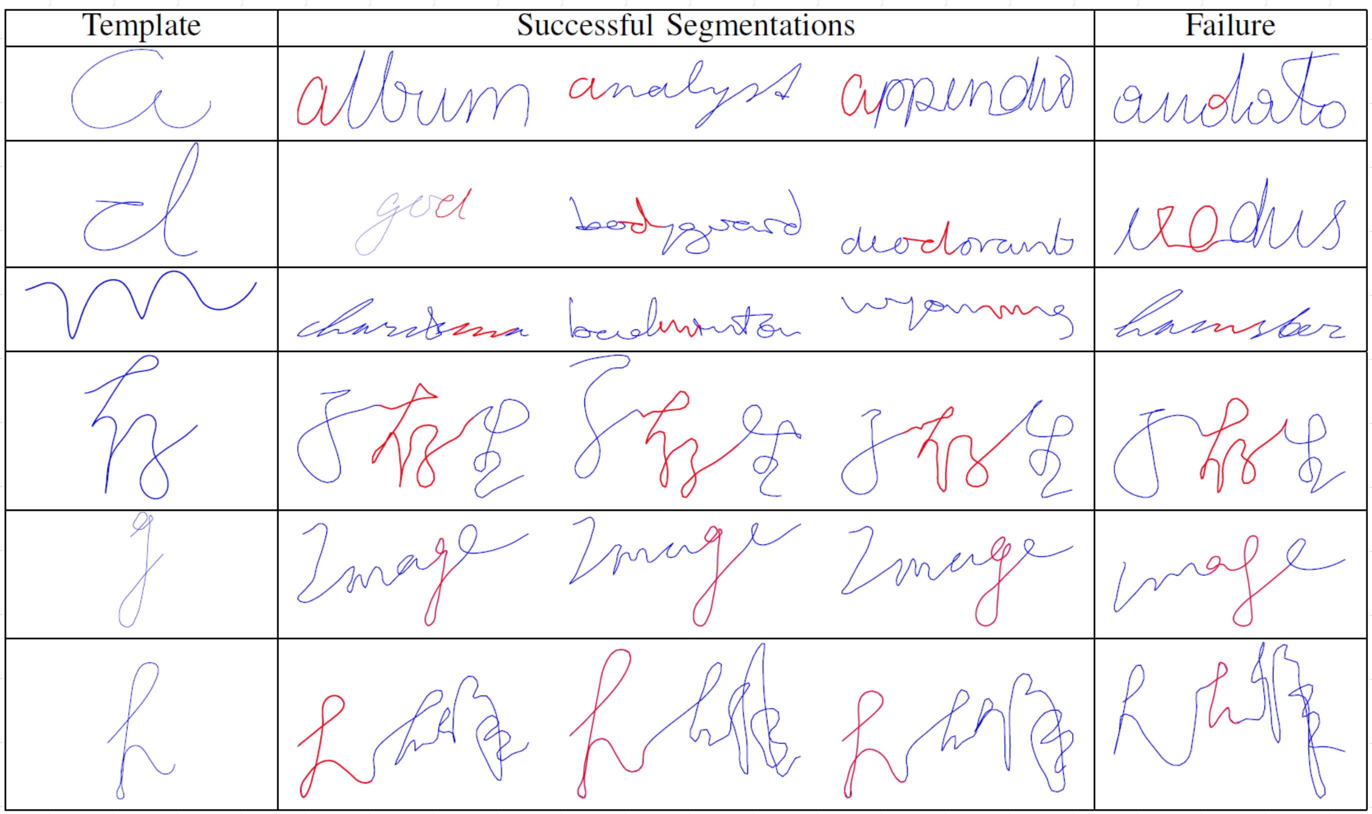
SIGNATURE SEGMENTATION



THE OHIO STATE UNIVERSITY

SIGNATURE SEGMENTATION RESULTS

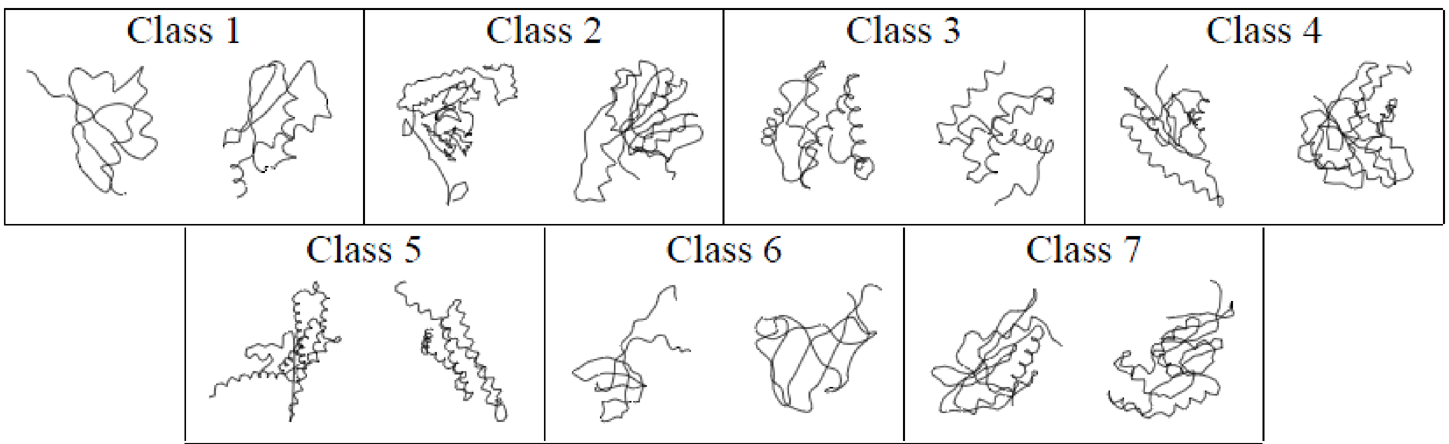


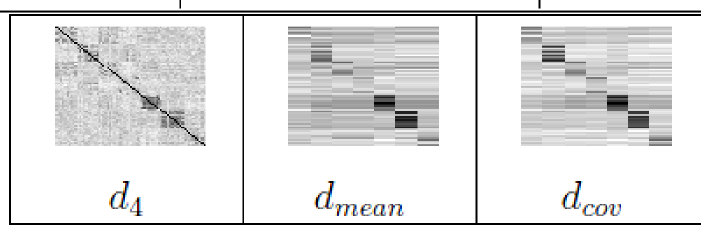


PROTEIN BACKBONE CLASSIFICATION USING TWN

1. Estimate the truncated Gaussian model in a leave-one-out manner.
2. Compute a covariance-adjusted distance (likelihood) to the mean.
3. Classify using nearest neighbor.
4. Compare results to all pair-wise distances (d_4), distance to the mean, and Procrustes distance.

Distance	d_4	d_{mean}	d_{cov}	Procrustes
Class. Performance	87.1%	68.6%	88.6%	77.5%



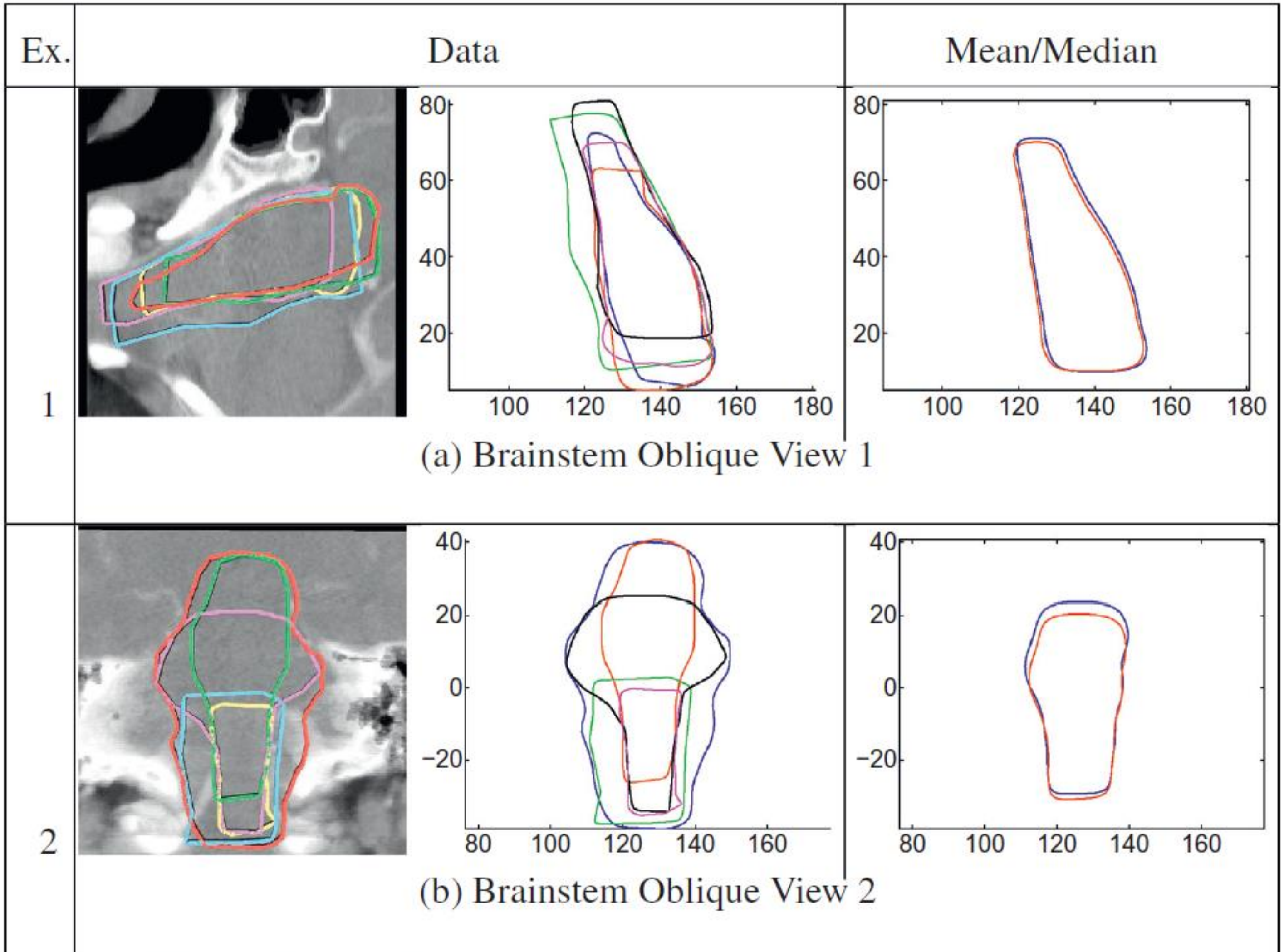


Kurtek et al., Journal of the American Statistical Association, 2012

THE OHIO STATE UNIVERSITY

MANUAL SEGMENTATIONS OF MEDICAL IMAGES

Citizen Science: non-experts helping scientists collect or analyze data.

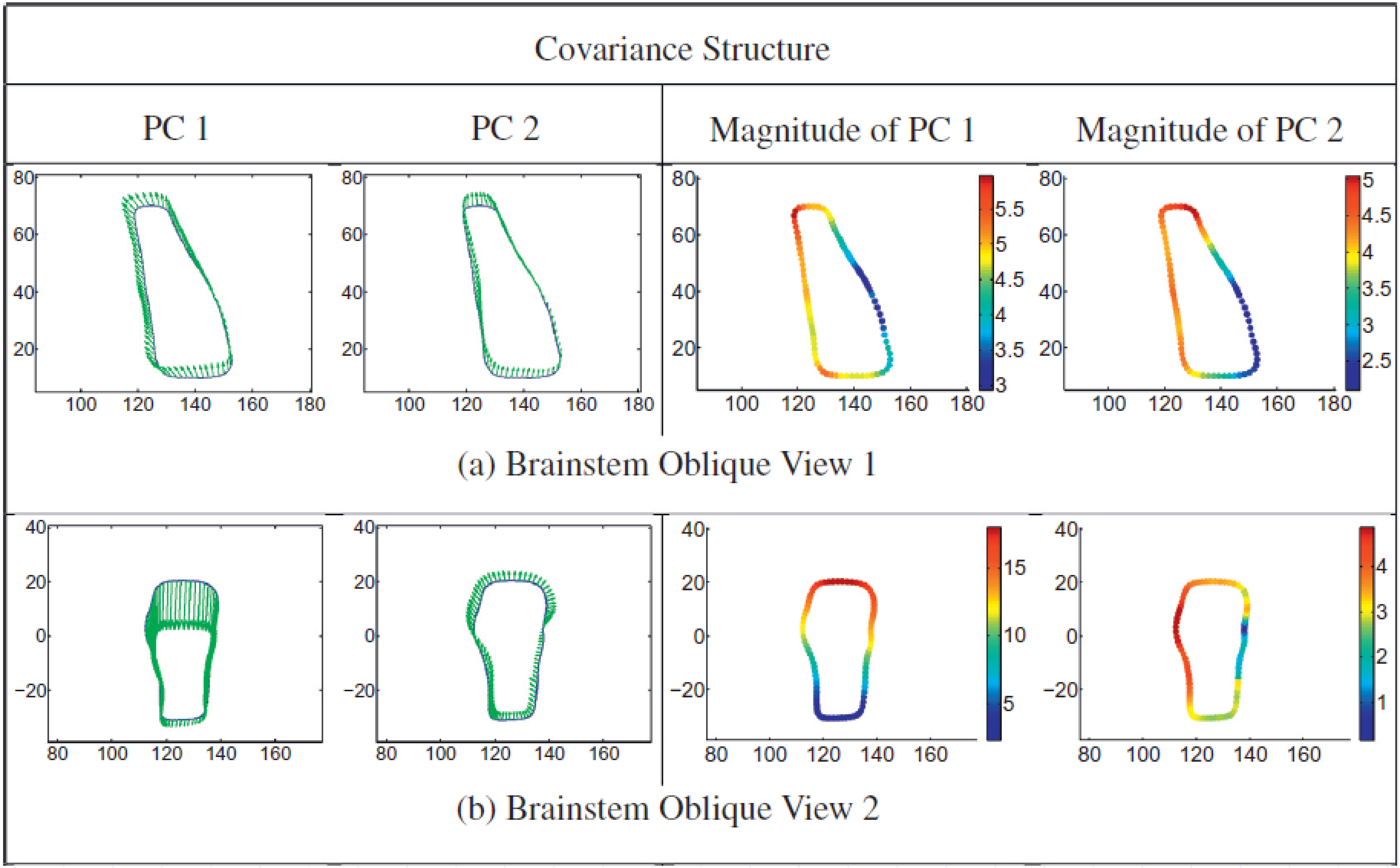


(a) Brainstem Oblique View 1

(b) Brainstem Oblique View 2

THE OHIO STATE UNIVERSITY

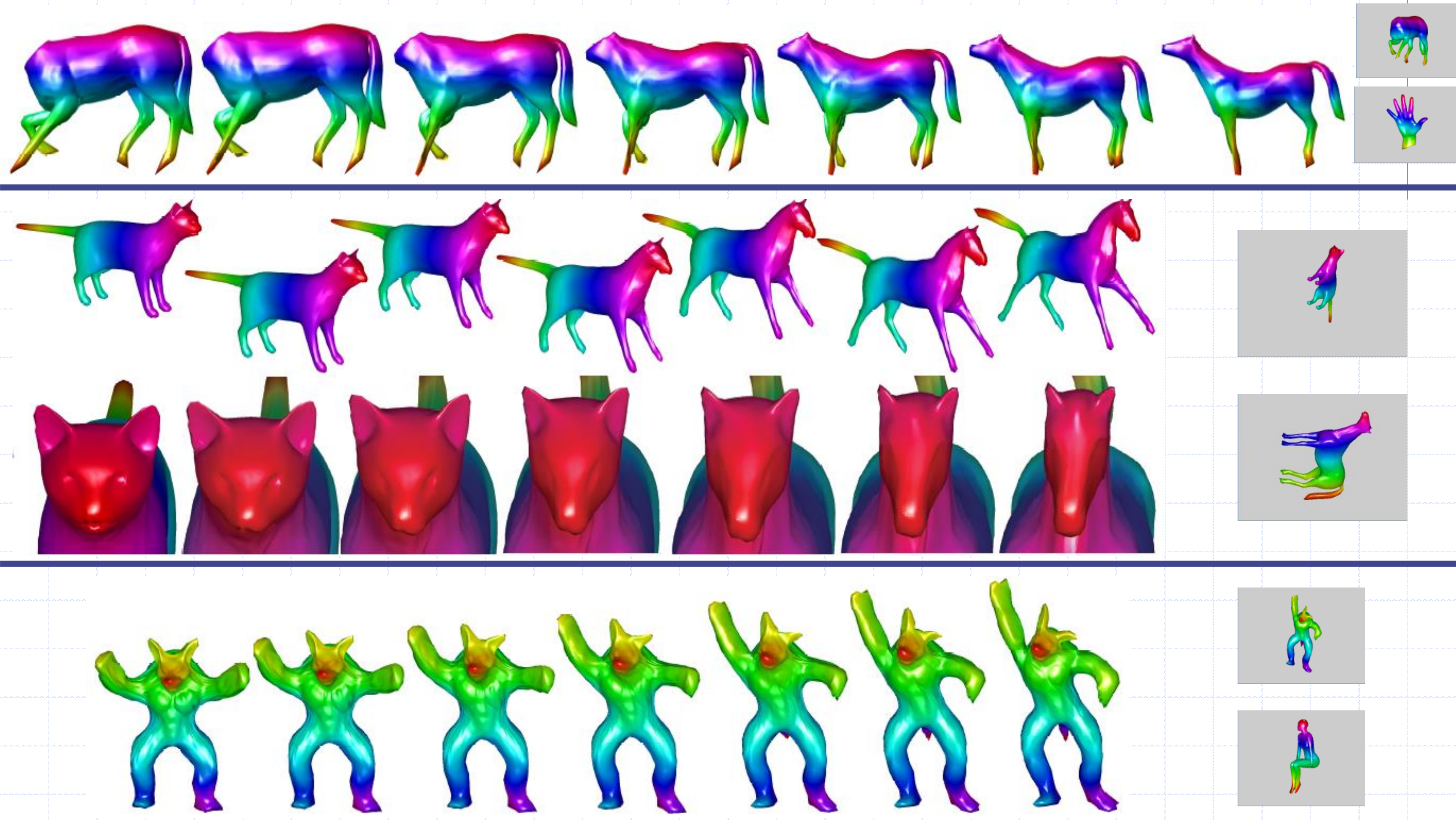
MANUAL SEGMENTATIONS OF MEDICAL IMAGES



(a) Brainstem Oblique View 1

(b) Brainstem Oblique View 2

GRAPHICS INSPIRED SHAPE GEODESICS



Kurtek et al., SIGGRAPH, 2012

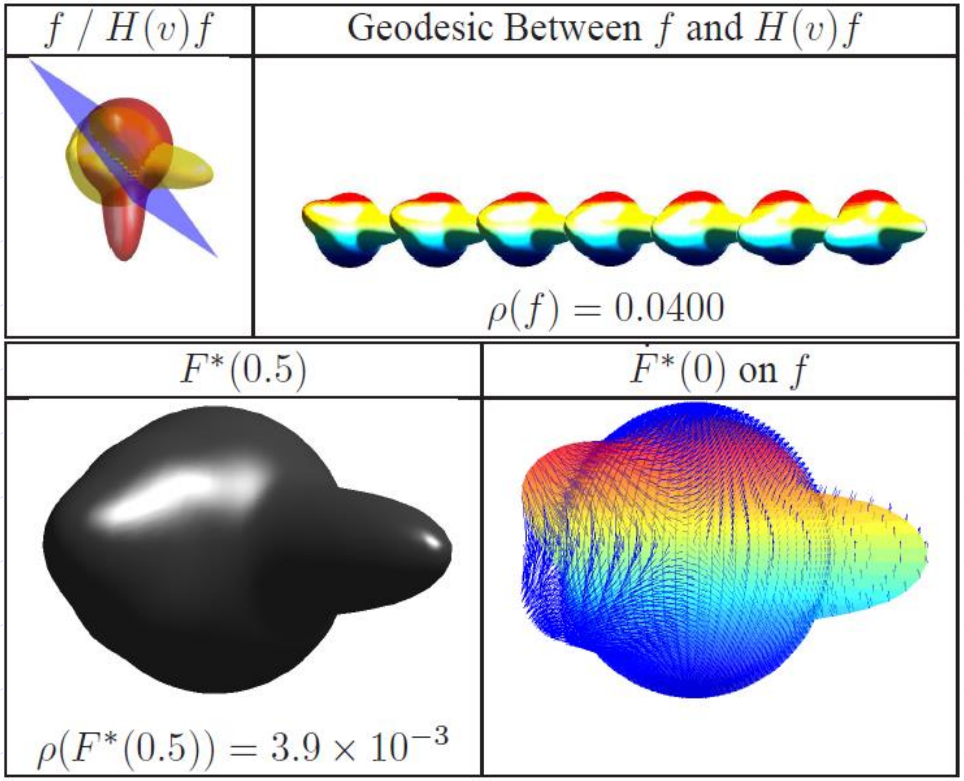
REFLECTION SYMMETRY ANALYSIS

Given a surface f and its reflection $\tilde{f}$.

Fact 1: Define the length of the geodesic path F* as a measure of symmetry of f (call this quantity ρ).

Fact 2: The halfway point along the geodesic, i.e. F*(0.5), is symmetric. If this geodesic path is unique, amongst all symmetric shapes, F*(0.5) is the nearest to f.

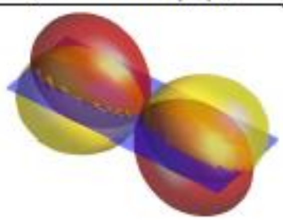
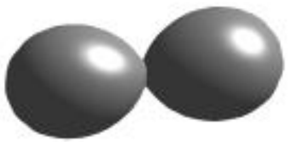
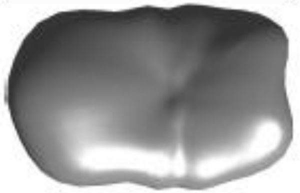
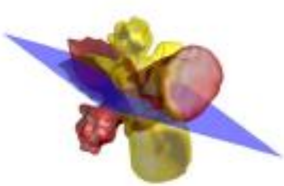
Fact 3: The velocity vector field $\dot{F}^*(0)$ provides a deformation (vector) field on f that deforms f into the nearest symmetric shape.



Kurtek and Srivastava, MMBIA, 2011

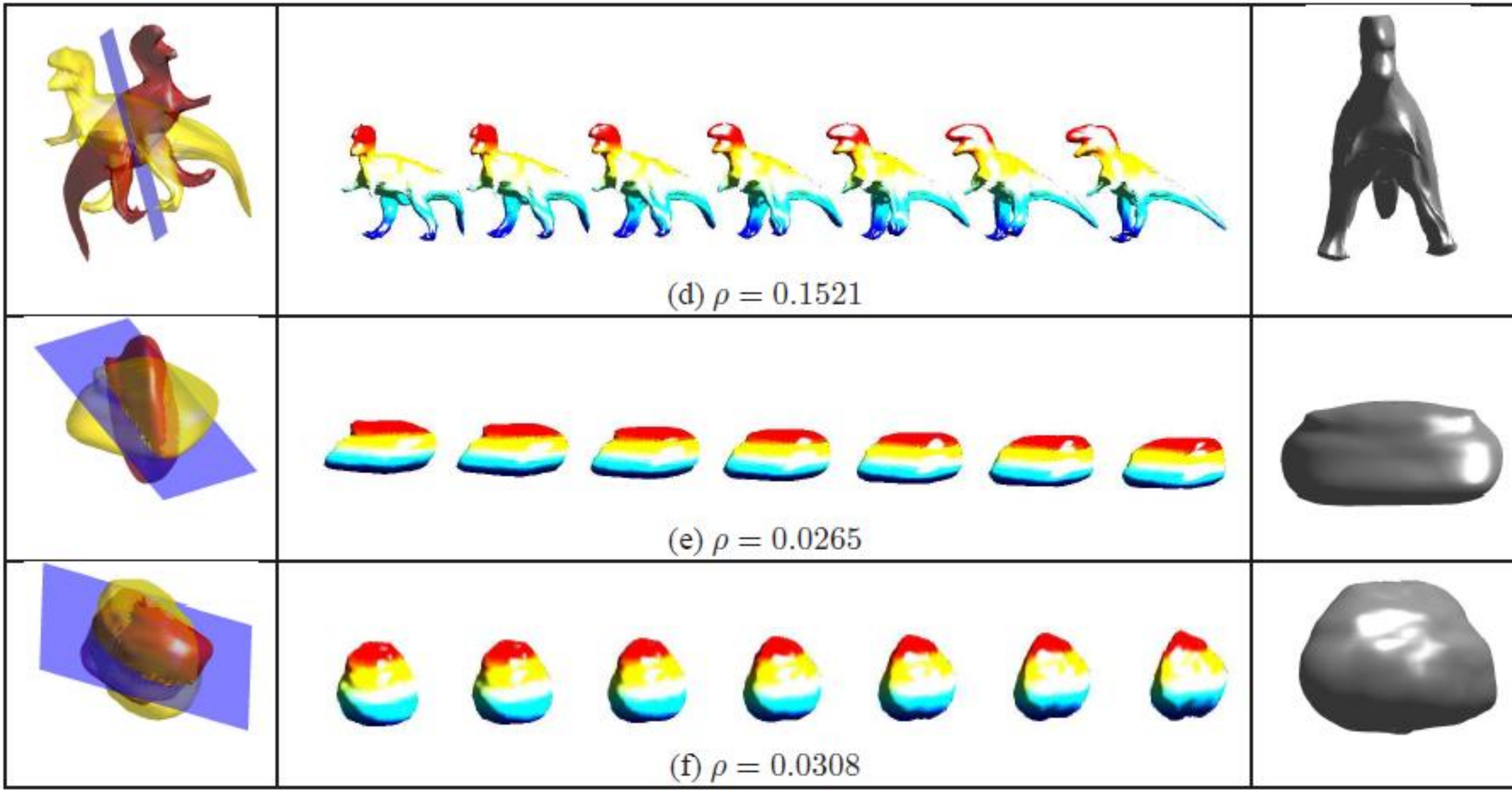
THE OHIO STATE UNIVERSITY

SYMMETRY ANALYSIS EXAMPLES

f and $H(v)f$	Geodesic Path	$F^*(0.5)$
	 (a) $\rho = 1.142 \times 10^{-8}$	
	 (b) $\rho = 0.0451$	
	 (c) $\rho = 0.0249$	

THE OHIO STATE UNIVERSITY

SYMMETRY ANALYSIS EXAMPLES

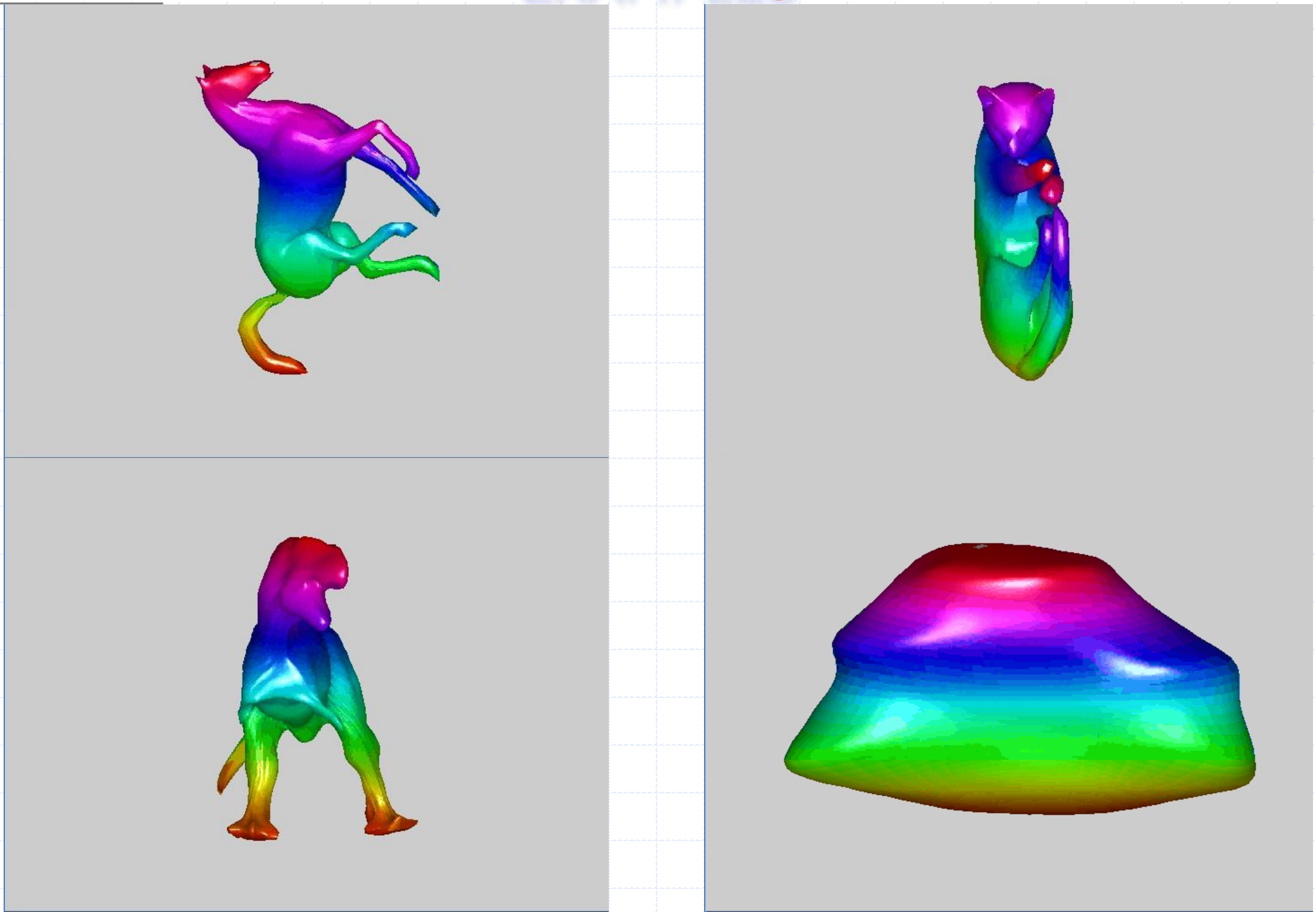


(d) $\rho = 0.1521$

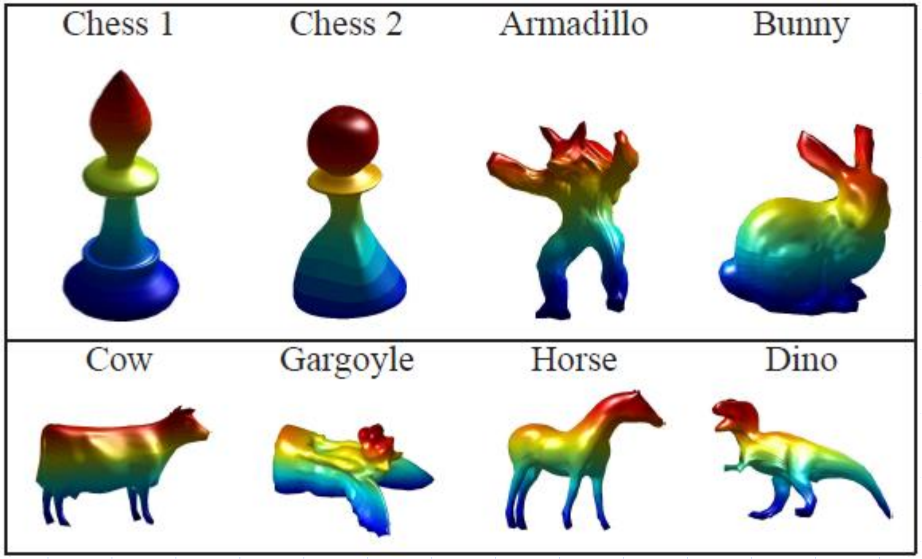
(e) $\rho = 0.0265$

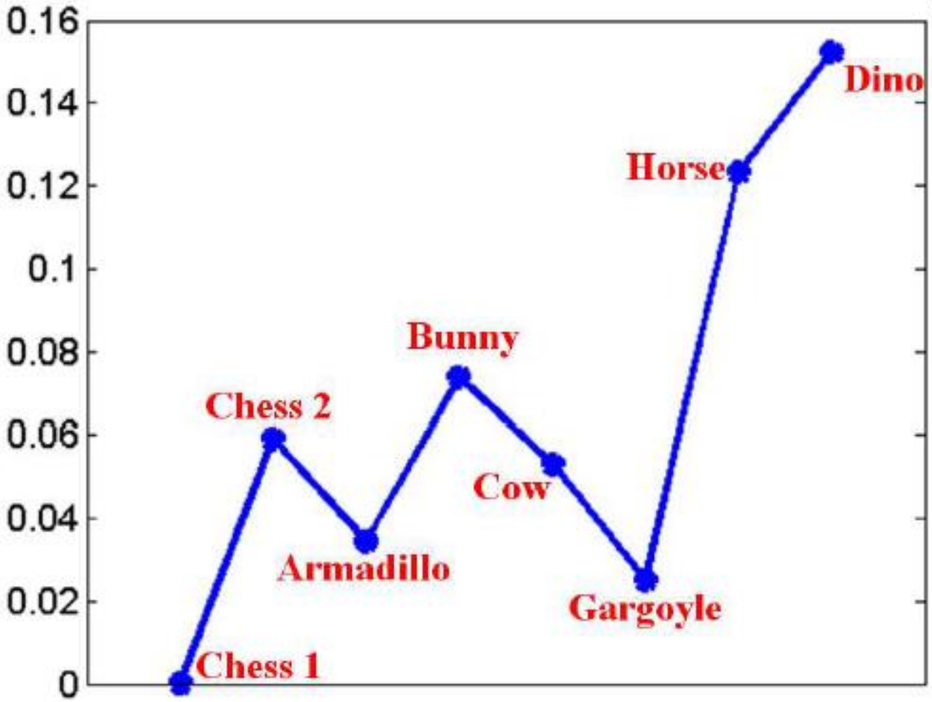
(f) $\rho = 0.0308$

SYMMETRY ANALYSIS: EXAMPLES



SYMMETRY ANALYSIS EXAMPLES





- The chess piece 1, armadillo and gargoyle are most symmetric.
- The bunny, horse and dino are least symmetric.

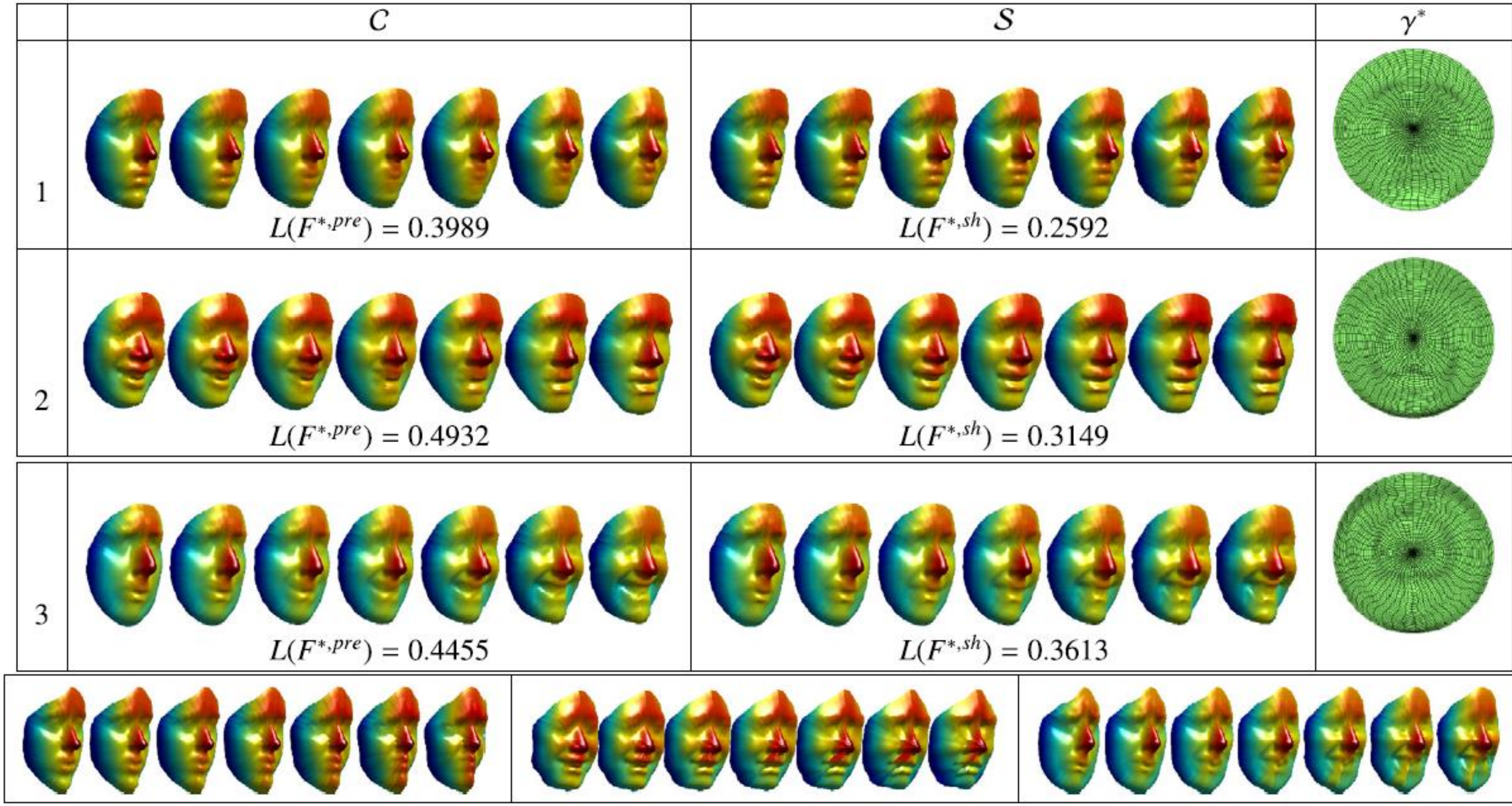
FACE SURFACES DATA DESCRIPTION

- Data description:
  - 66 cropped facial surfaces divided into six expression groups: anger, disgust, fear, happiness, surprise and sadness
  - 11 people per expression
- Each surface is parameterized using a disk domain with the origin fixed at the tip of the nose (automatically detected).
- We applied the elastic shape analysis framework via SRNFs to compute geodesics, averages, PCA and classification.
- As a form of comparison, we also applied the SRVF framework to the radial curves, which is a common approach in the literature.

Kurtek and Drira, Computers and Graphics, 2015

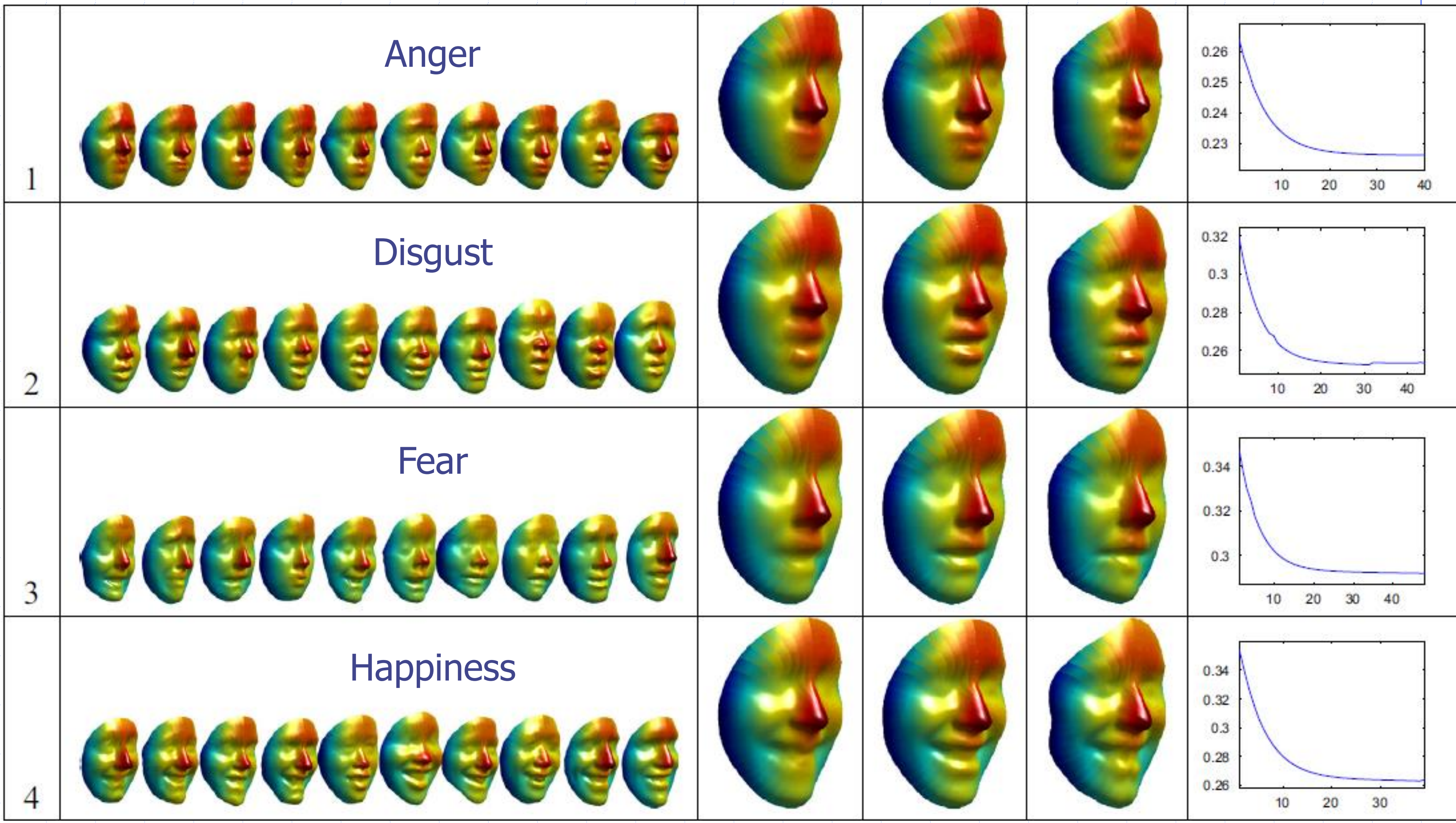
THE OHIO STATE UNIVERSITY

GEODESIC PATHS

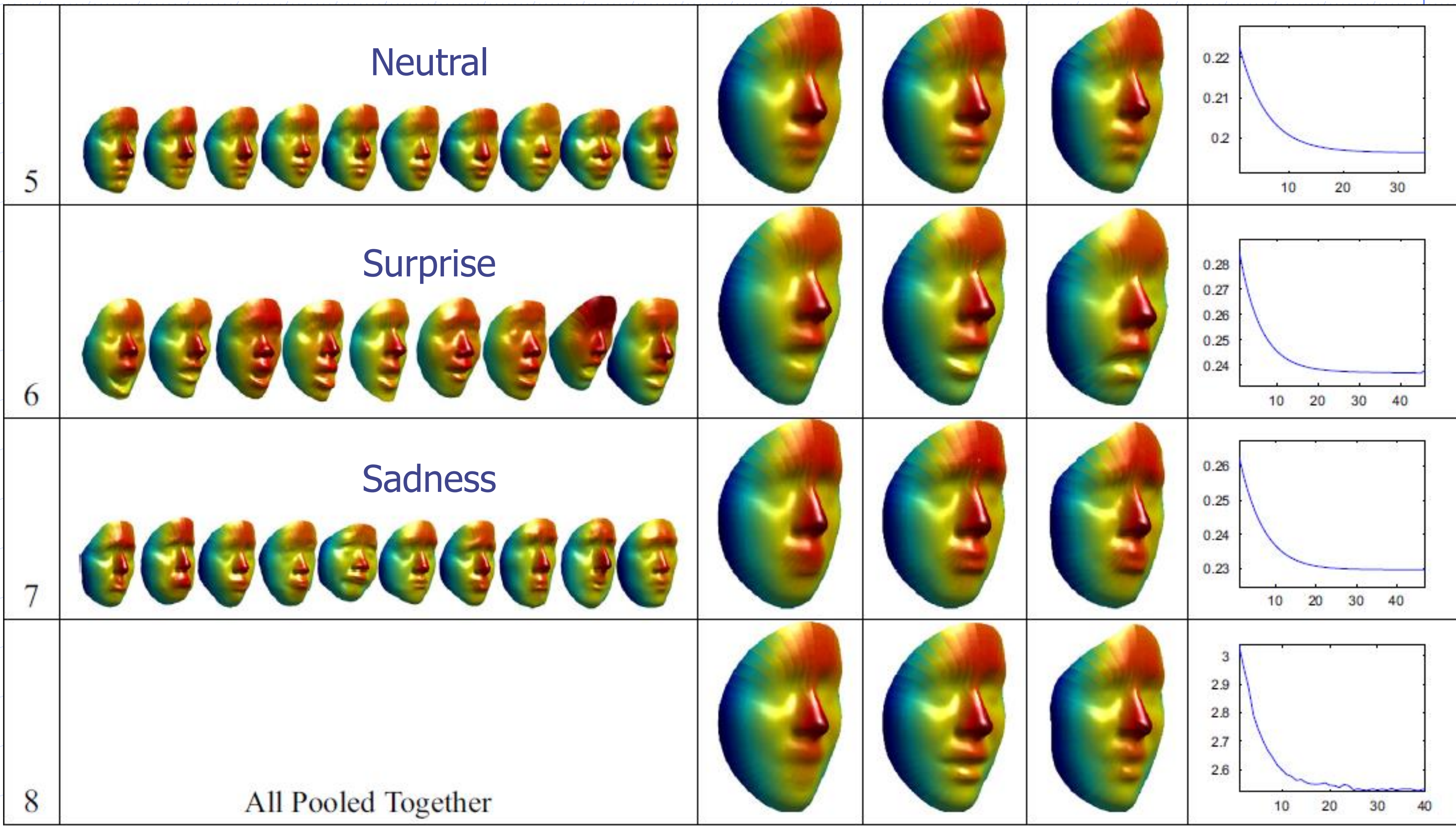


(1) neutral to anger, (2) happiness to disgust, (3) sadness to happiness

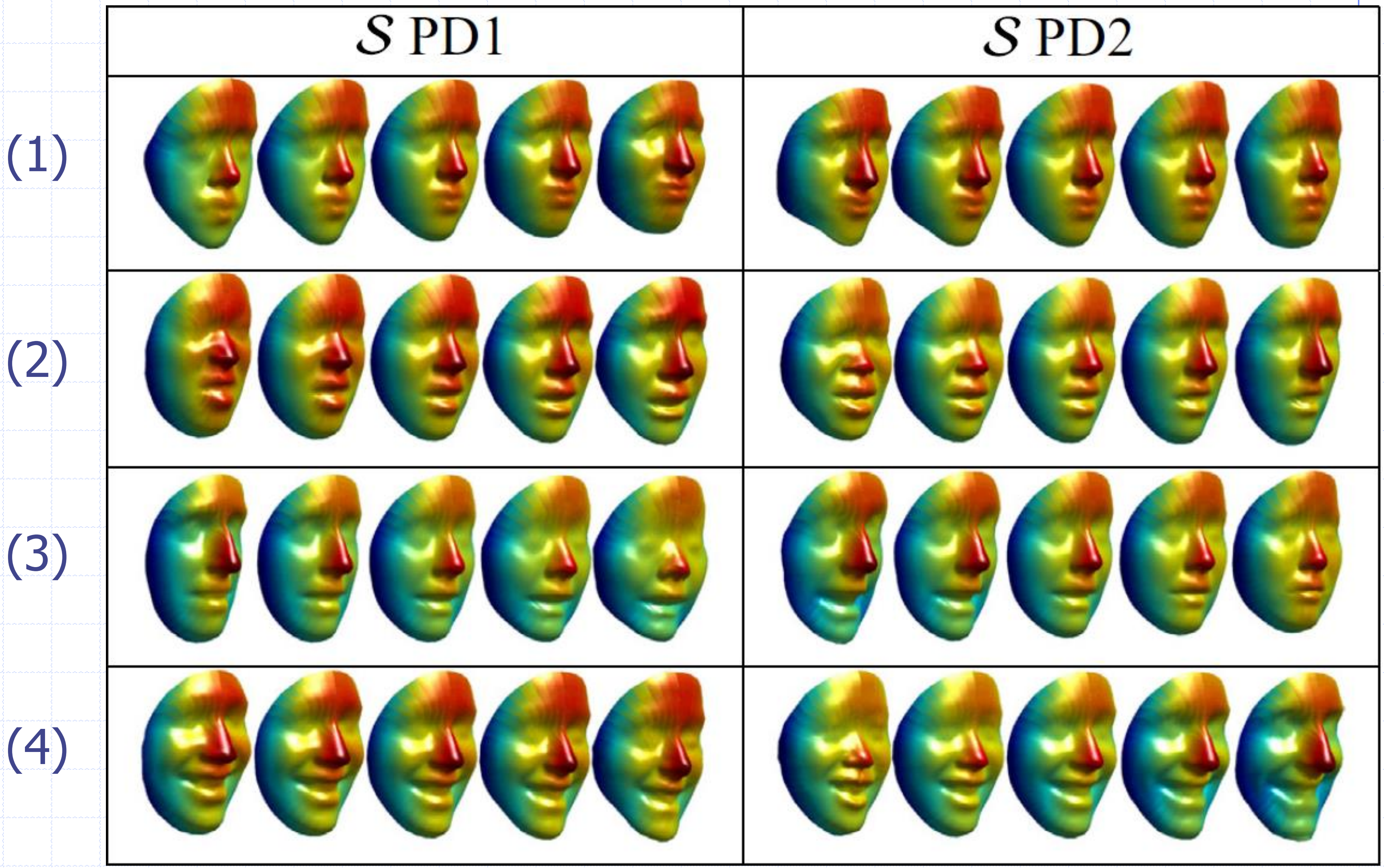
AVERAGING DIFFERENT EXPRESSIONS



AVERAGING DIFFERENT EXPRESSIONS

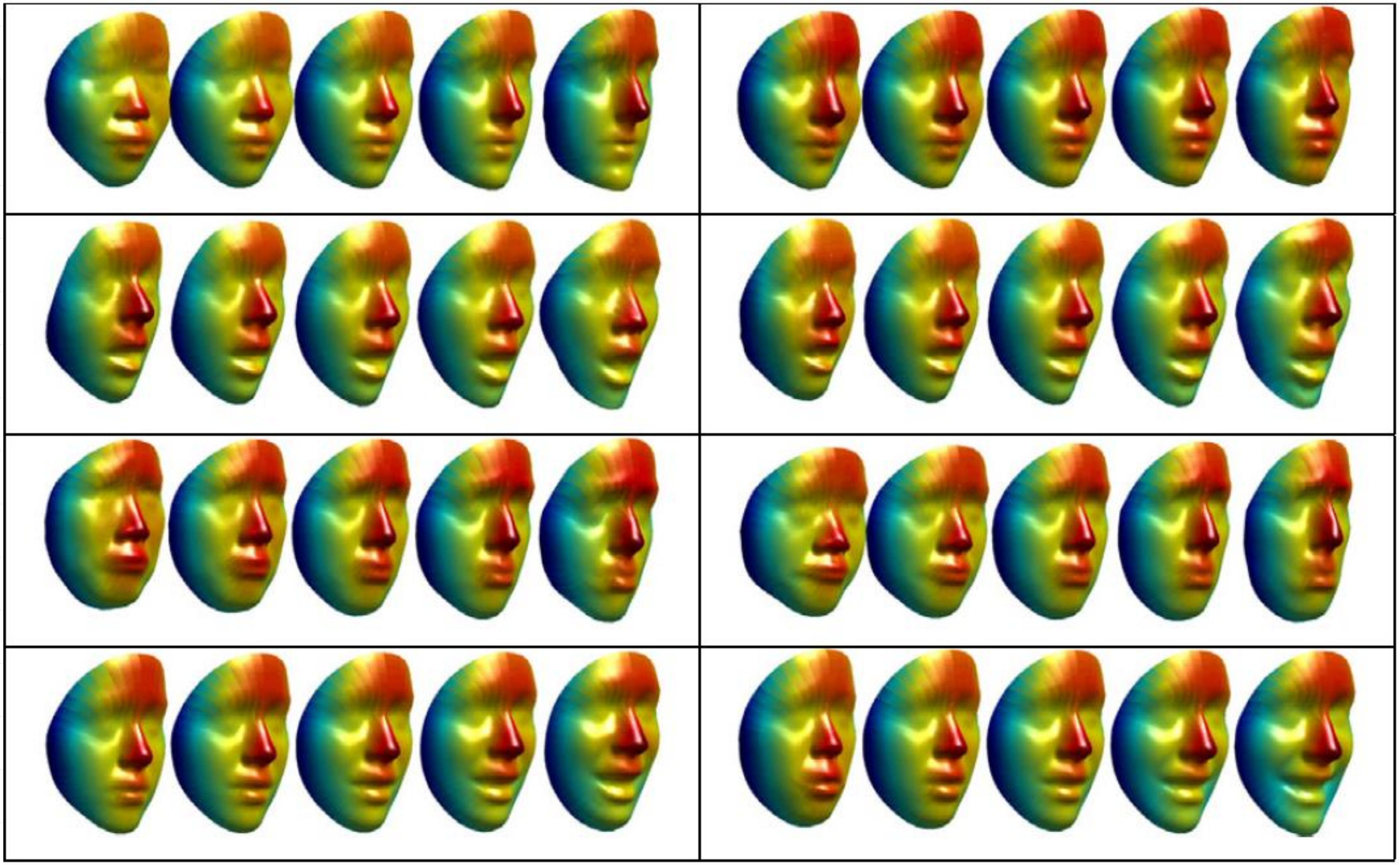


EXPRESSION PCA



EXPRESSION PCA

(5)
(6)
(7)
(8)



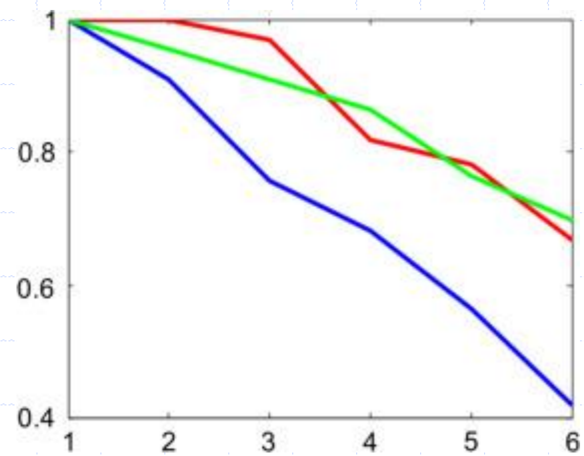
IDENTITY/EXPRESSION RECOGNITION

1. Distance-based expression recognition performance (leave out all expressions of the test subject).

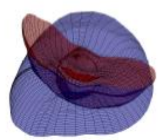
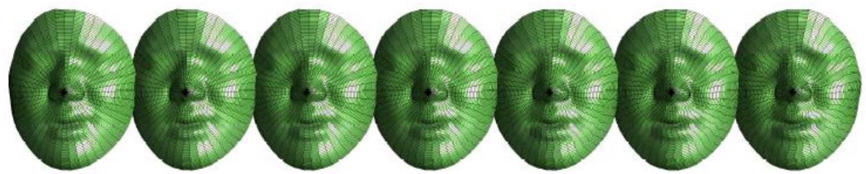
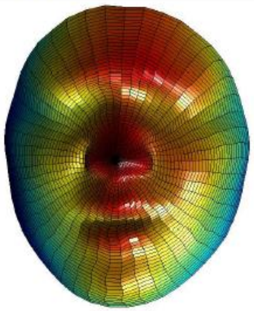
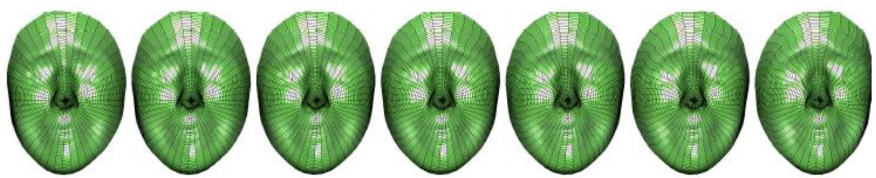
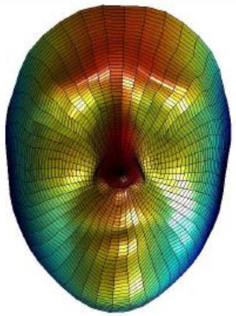
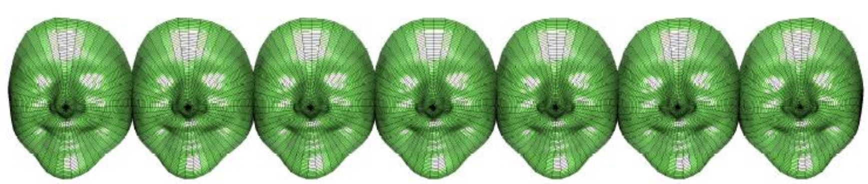
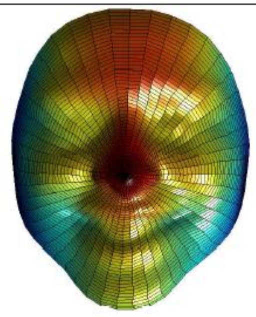
   (a) Elastic surface method: 74.24% accuracy
   (b) Non-elastic surface method (in pre-shape space): 62.12% accuracy
   (c) Elastic curve method: 68.18% accuracy

2. ROC curves for identity recognition on the same 66 subjects:

   Red: Elastic surface method
   Blue: Non-elastic surface method
   Green: Elastic curve method



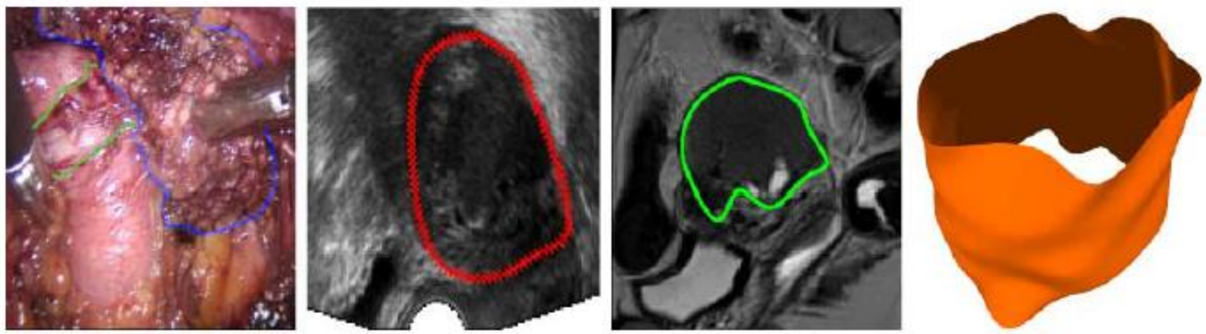
SYMMETRY ANALYSIS OF FACES

	(a) f and $\tilde{f}$	(b) Geodesic Path $F^{*,sh}$	(c) $F^{*,sh}(0.5)$
1		 $\rho(f) = 0.1626, \rho(F^{*,sh}(0.5)) = 0.0177$	
2		 $\rho(f) = 0.1953, \rho(F^{*,sh}(0.5)) = 0.0201$	
3		 $\rho(f) = 0.1041, \rho(F^{*,sh}(0.5)) = 0.0100$	

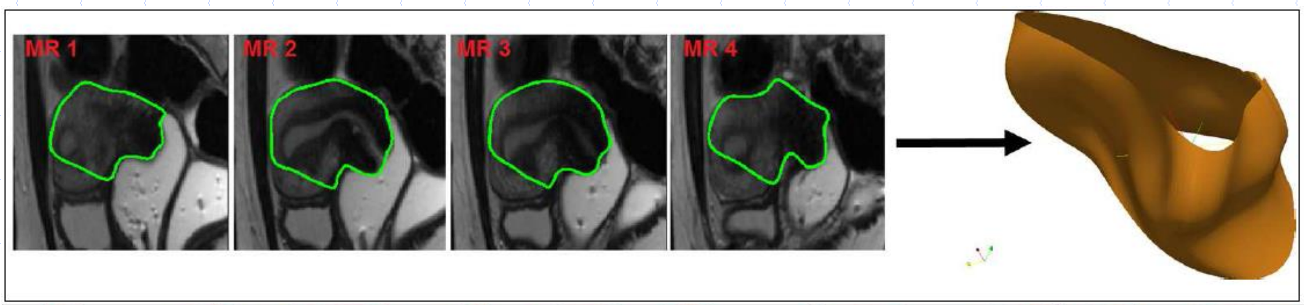
The Ohio State University

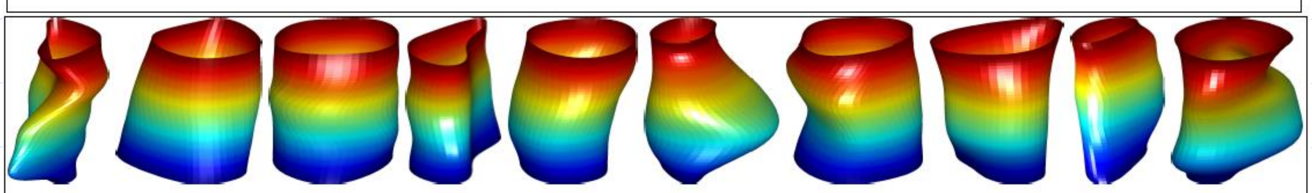
ENDOMETRIAL TISSUE SHAPES

- Data: 10 endometrial tissue surfaces manually segmented from Magnetic Resonance Images (MRIs) and corresponding curve from TransVaginal Ultrasound Images (TVUS)



Laparoscopic TVUS MRI





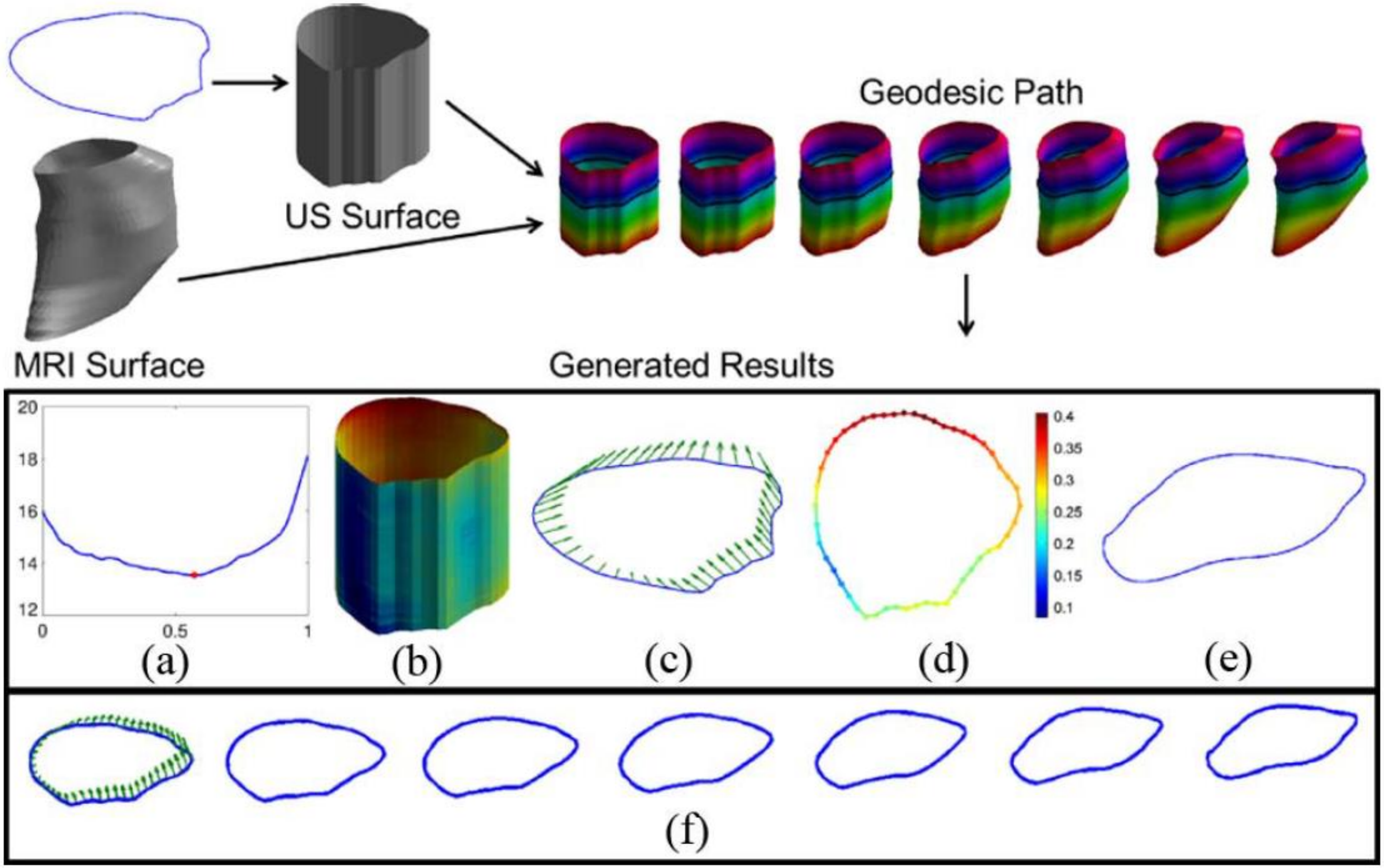
ENDOMETRIAL TISSUE SHAPES

THE OHIO STATE UNIVERSITY

- Goal 1:
  - Identify curve in MRI image that best corresponds to curve in TVUS image.
  - Compute optimal deformation between the two curves and quantify the elasticity of endometriosis using a measure of this deformation.
  - Extend this deformation from the optimal curve to the full MRI surface to help a physician in locating and assessing endometrium shape variability.
- Goal 2:
  - Define a shape model and generate random samples for validation purposes in endometriosis studies.
  - Simulate deformed endometrial tissue shapes as observed in Transvaginal Ultrasound (TVUS) images.

Samir et al., IEEE T Medical Imaging, 2014; Kurtek et al., Neurocomputing, 2016

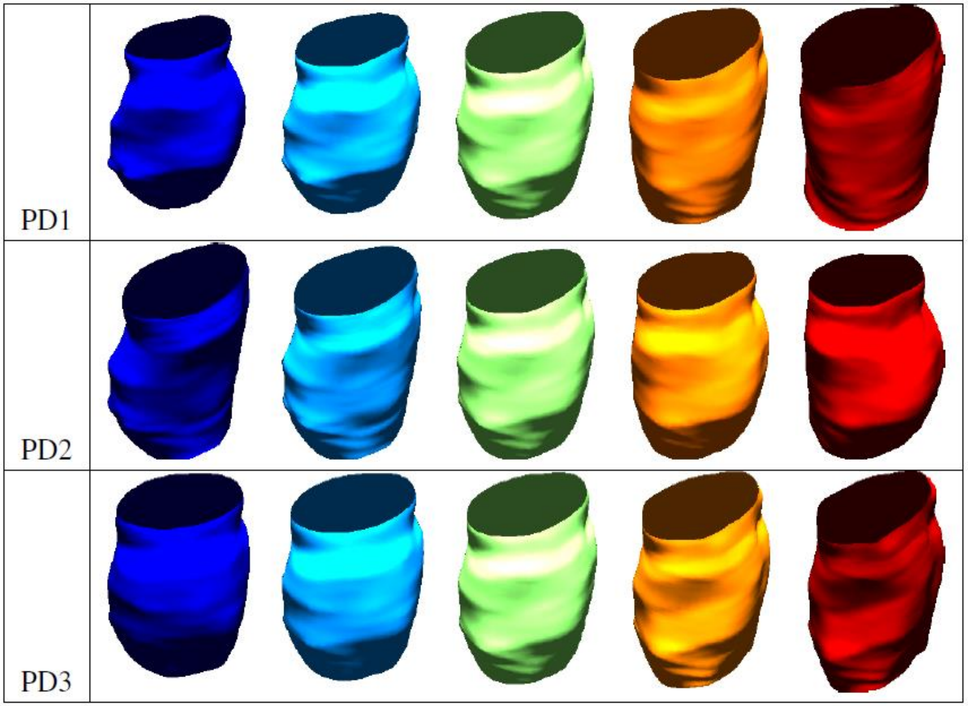
GOAL 1

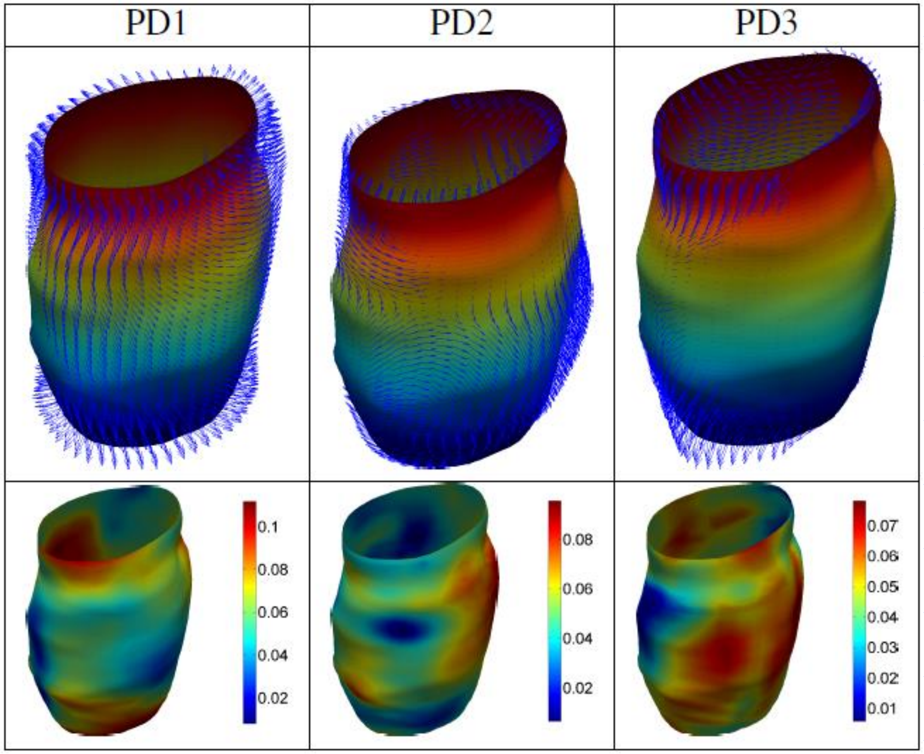

Geodesic Path
US Surface
MRI Surface
Generated Results
20
18
16
14
12
0
0.5
1
0.4
0.35
0.3
0.25
0.2
0.15
0.1
(a)
(b)
(c)
(d)
(e)
(f)

GOAL 2

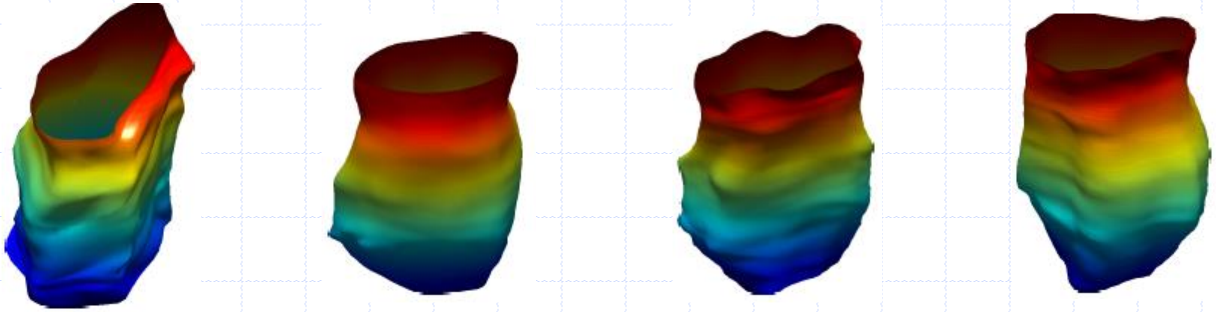
Shape model and random samples:

$$-1\sigma \rightarrow \mu \rightarrow +1\sigma$$



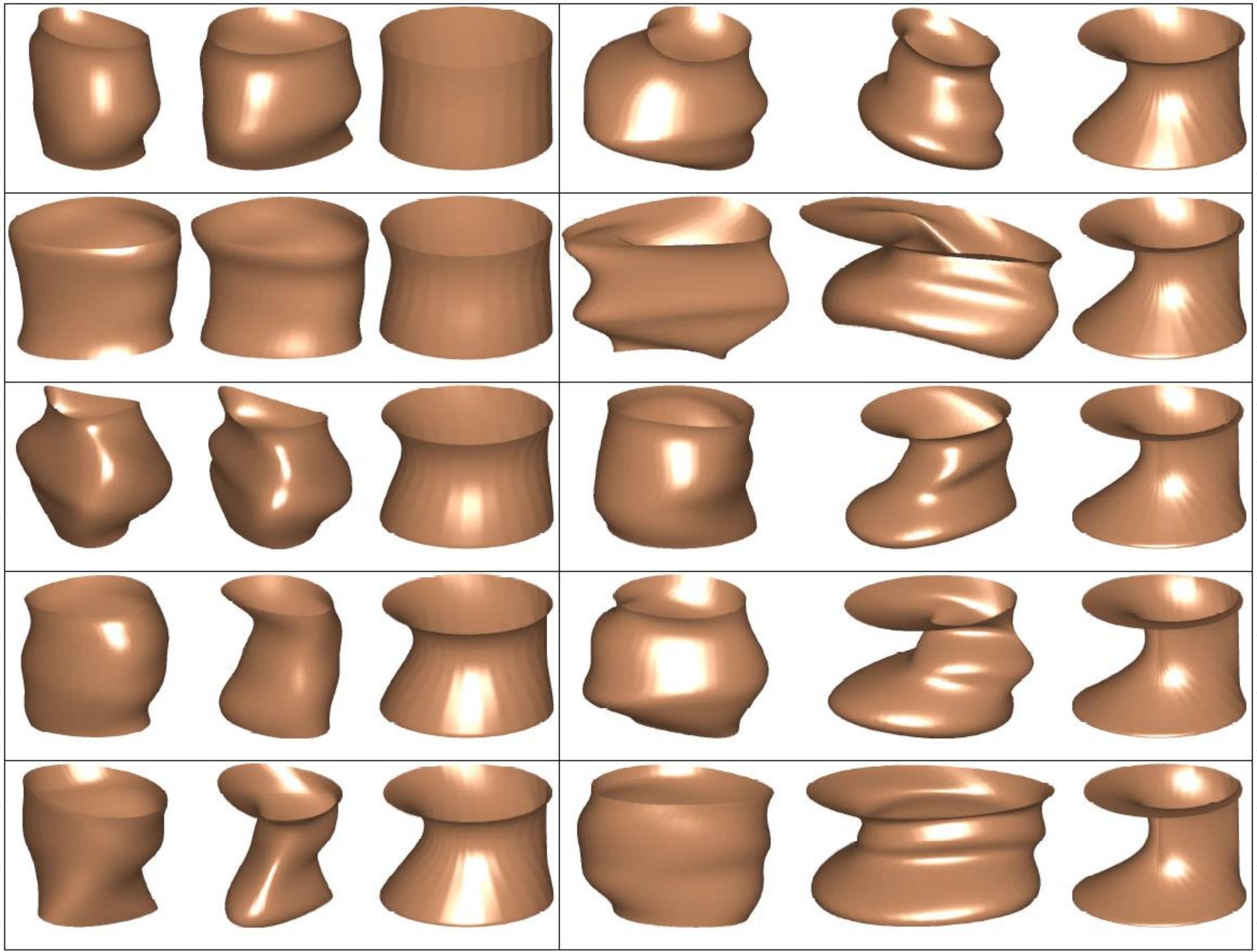


Random Samples



GOAL 2

Random samples with additional deformations simulating the transducer's pressure on the endometrial tissue.



The Ohio State University

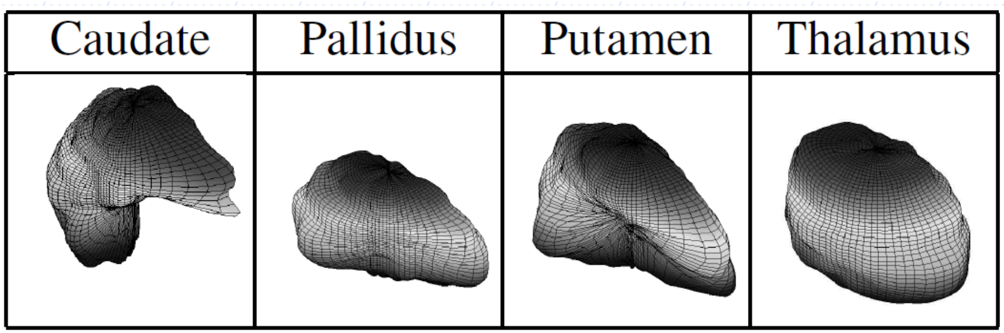
ELASTIC SHAPE ANALYSIS OF BRAIN STRUCTURES

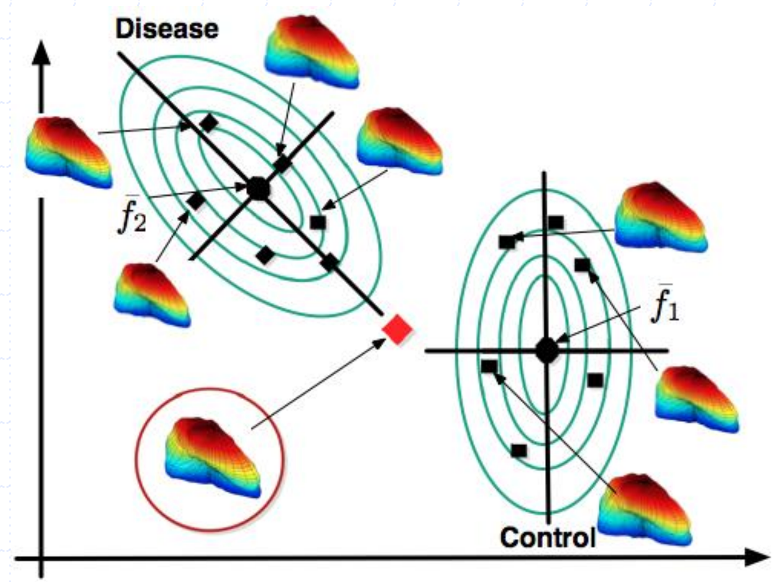
Relevance:

- Improvements in non-invasive medical imaging technology have enabled the study of biological variations of anatomical structures.
- Studying shapes of 3D anatomical structures in the brain is of particular interest.
- Many diseases are linked to altering the shapes of anatomical structures.
- In clinical studies it is typical to use clinical symptoms, such as behavioral tests, to classify disease types.
- In many neuro-degenerative diseases such as Alzheimer’s or Huntington’s, behavioral symptoms may not be present until full onset of the disease.
- The behavioral tests are subjective and qualitative.
- Shape analysis of anatomical structures provides an opportunity for early detection of diseases and a quantitative classification of disease types.

ADHD CLASSIFICATION

- Data: 34 subjects, 19 with diagnosed Attention Deficit Hyperactivity Disorder (ADHD) and 15 controls.
- We extracted four left and right subcortical structures from MRIs, which were chosen based on medical literature.
- We performed classification based on the shape differences of the subcortical structures.





Kurtek et al., IEEE T Medical Imaging, 2011; Xie et al., ECCV, 2014

ATTENTION DEFICIT HYPERACTIVITY DISORDER CLASSIFICATION

- Gaussian: estimate Gaussian model for each group in a leave-one-out manner and use likelihood ratio for classification
- NN: leave-one-out nearest neighbor distance-based classification
- SRF: square root function (previous idea for elastic shape analysis of surfaces)
- Harmonic: $\mathbb{L}^2$ distance between surfaces in pre-shape space
- ICP: iterative closest point algorithm
- SPHARM: spherical harmonics representation of surfaces

Method Structure	SRNF Gaussian	SRF Gaussian	SRF NN	Harmonic NN	ICP NN	SPHARM NN
L. Caudate	**67.7%**	-	41.2%	64.7%	32.4%	61.8%
L. Pallidus	**85.3%**	82.4%	76.5%	79.4%	67.7%	44.1%
L. Putamen	**94.1%**	88.2%	82.4%	70.6%	61.8%	50.0%
L. Thalamus	**67.7%**	-	58.8%	67.7%	35.5%	52.9%
R. Caudate	55.9%	-	50.0%	44.1%	50.0%	**70.6%**
R. Pallidus	**76.5%**	67.6%	61.8%	67.7%	55.9%	52.9%
R. Putamen	67.7%	**82.4%**	67.7%	55.9%	47.2%	55.9%
R. Thalamus	**67.7%**	-	58.8%	52.9%	64.7%	64.7%

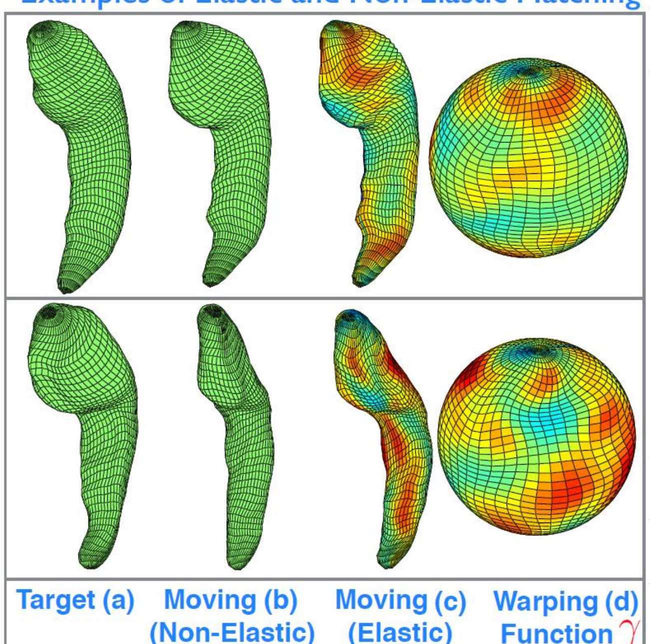
HIPPOCAMPAL SHAPE MORPHOMETRY IN ALZHEIMER'S DISEASE

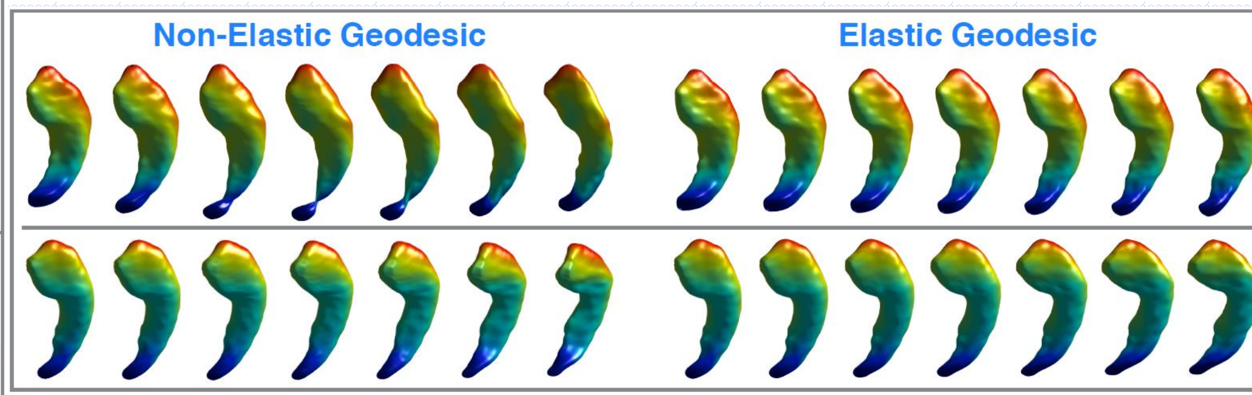
- Data: hippocampus surfaces extracted from MRIs of 120 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI).

- There are three groups in this data: Alzheimer's disease (AD), mild cognitive impairment (MCI) and normal controls (NC) with 40 subjects in each group.

Note: non-elastic refers to pre-shape and elastic refers to shape

Examples of Elastic and Non-Elastic Matching

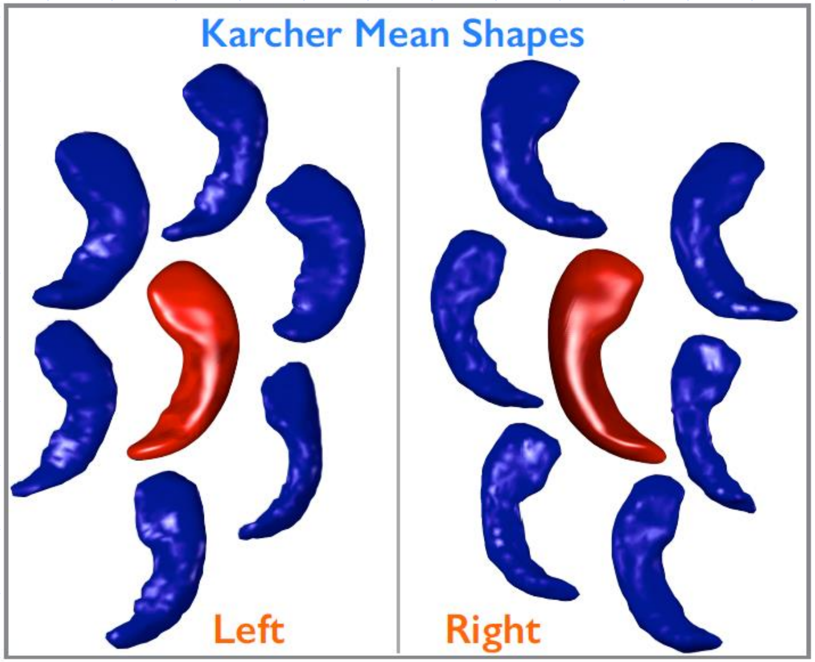




Joshi et al., DICTA, 2017

HIPPOCAMPAL SHAPE MORPHOMETRY IN ALZHEIMER'S DISEASE

The Ohio State University



Eigen-Shape variation for principal directions with significant disease effects

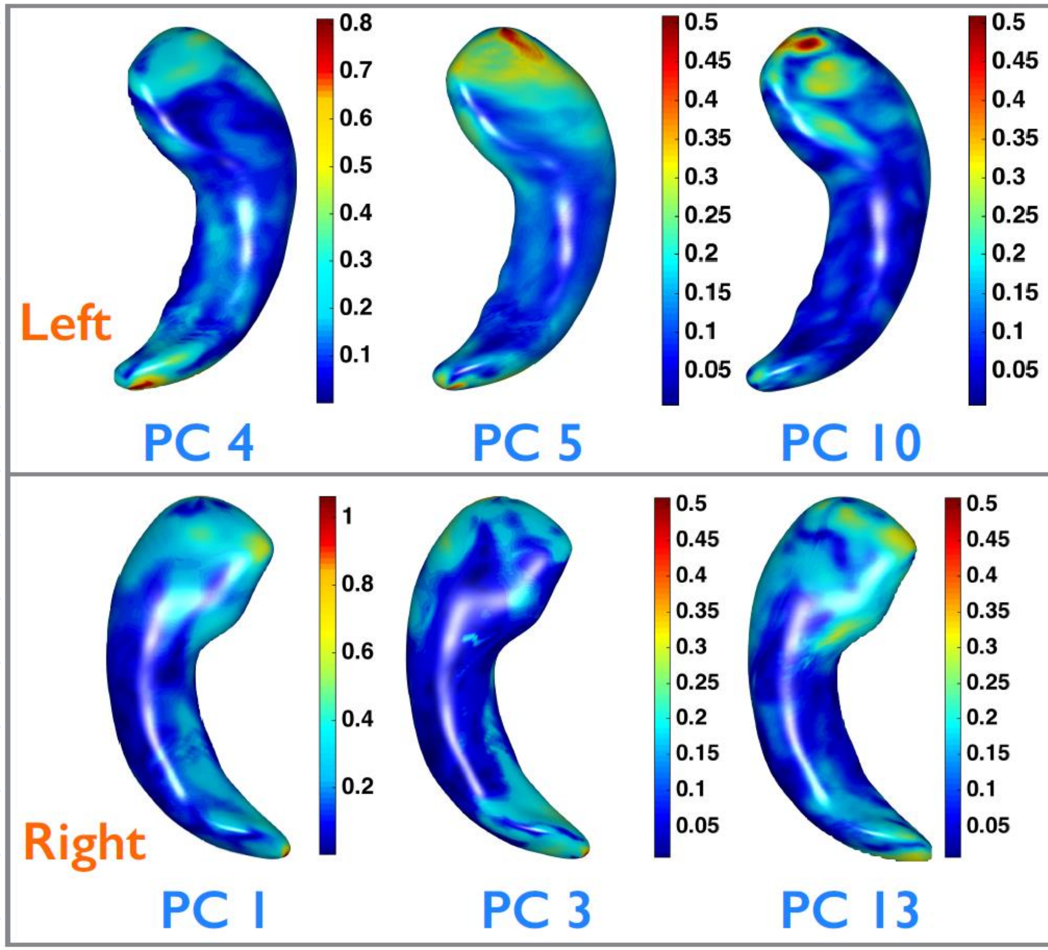
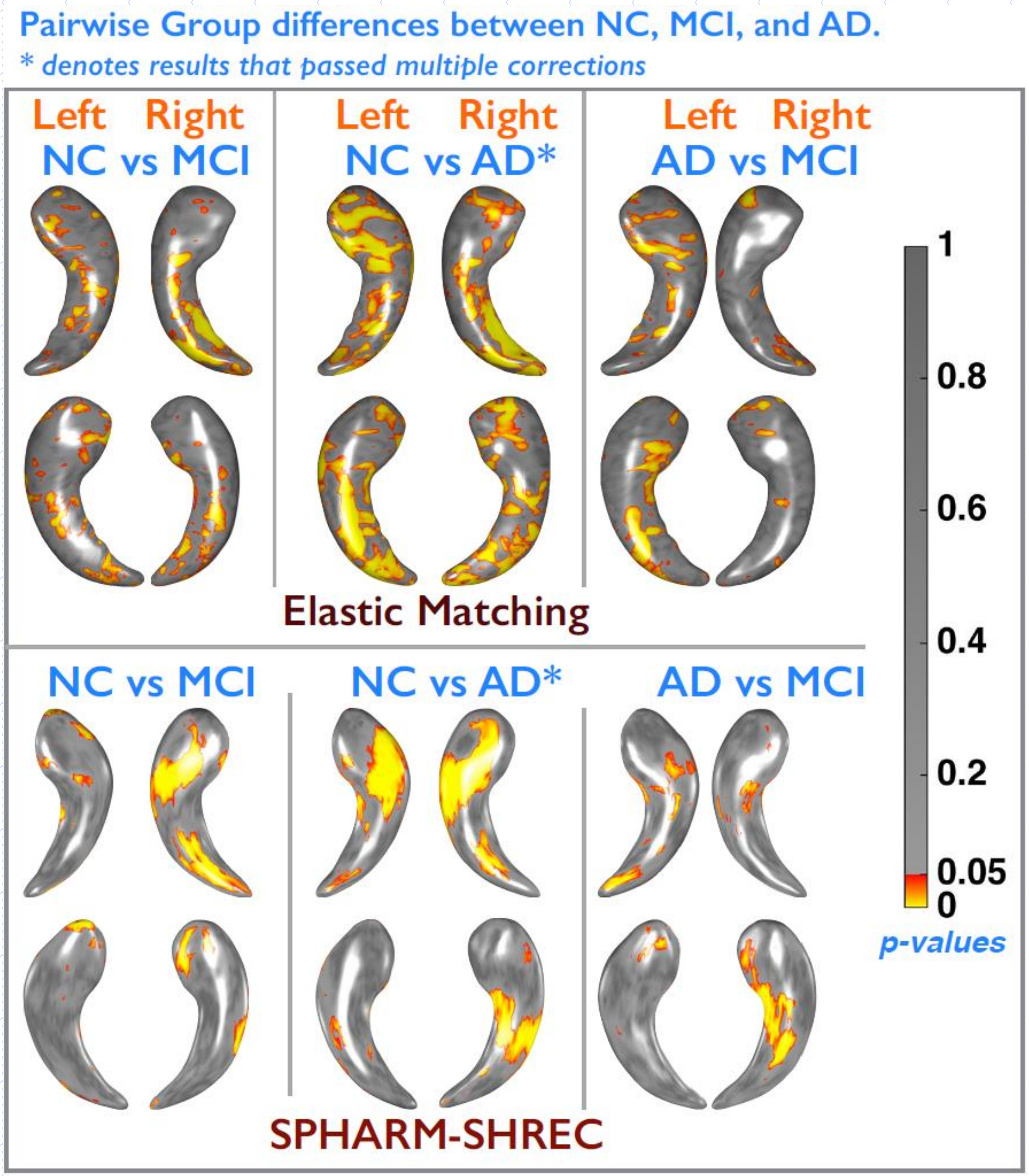


Table of Significant PCs

	AD vs NC			NC vs MCI	
Left	PC 4	PC 5	PC 10	PC 4	PC 10
	0.0007	0.0002	0.0002	0.0137	0.0006
Right	PC 1	PC 3	PC 13		
	0.0015	0.028	0.0126		

HIPPOCAMPAL SHAPE MORPHOMETRY IN ALZHEIMER'S DISEASE



THE OHIO STATE UNIVERSITY

THANK YOU!

I HOPE YOU ENJOYED THE WORKSHOP!

HAVE A SAFE TRIP HOME!