March 9, 2018

Shape Modeling in CellOrganizer

Robert F. Murphy
Ray & Stephanie Lane Professor of Computational Biology and
Professor of Biological Sciences, Biomedical Engineering and Machine Learning
External Senior Fellow, Freiburg Institute for Advanced Studies
Honorary Professor, Faculty of Biology, University of Freiburg, Germany

Carnegie Mellon University
Computational Biology Department

MMBioS
An NIH Biomedical Technology Research Center

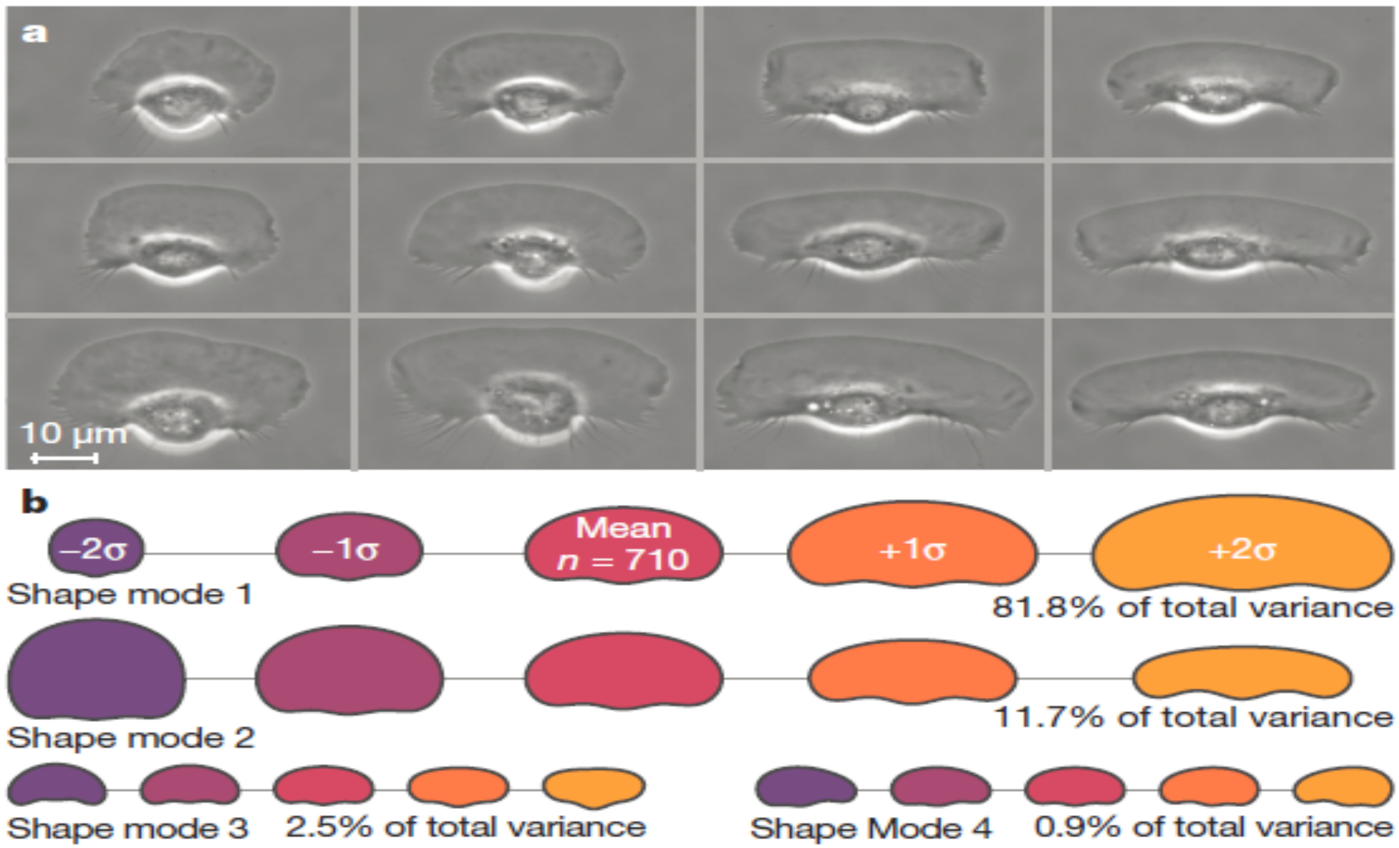
Methods for modeling cell shape

- Parametric
  - Outline
  - Ratio-metric relative to nuclear shape
- Nonparametric
  - Diffeomorphic
  - Autoencoder

Parametric shape outline models

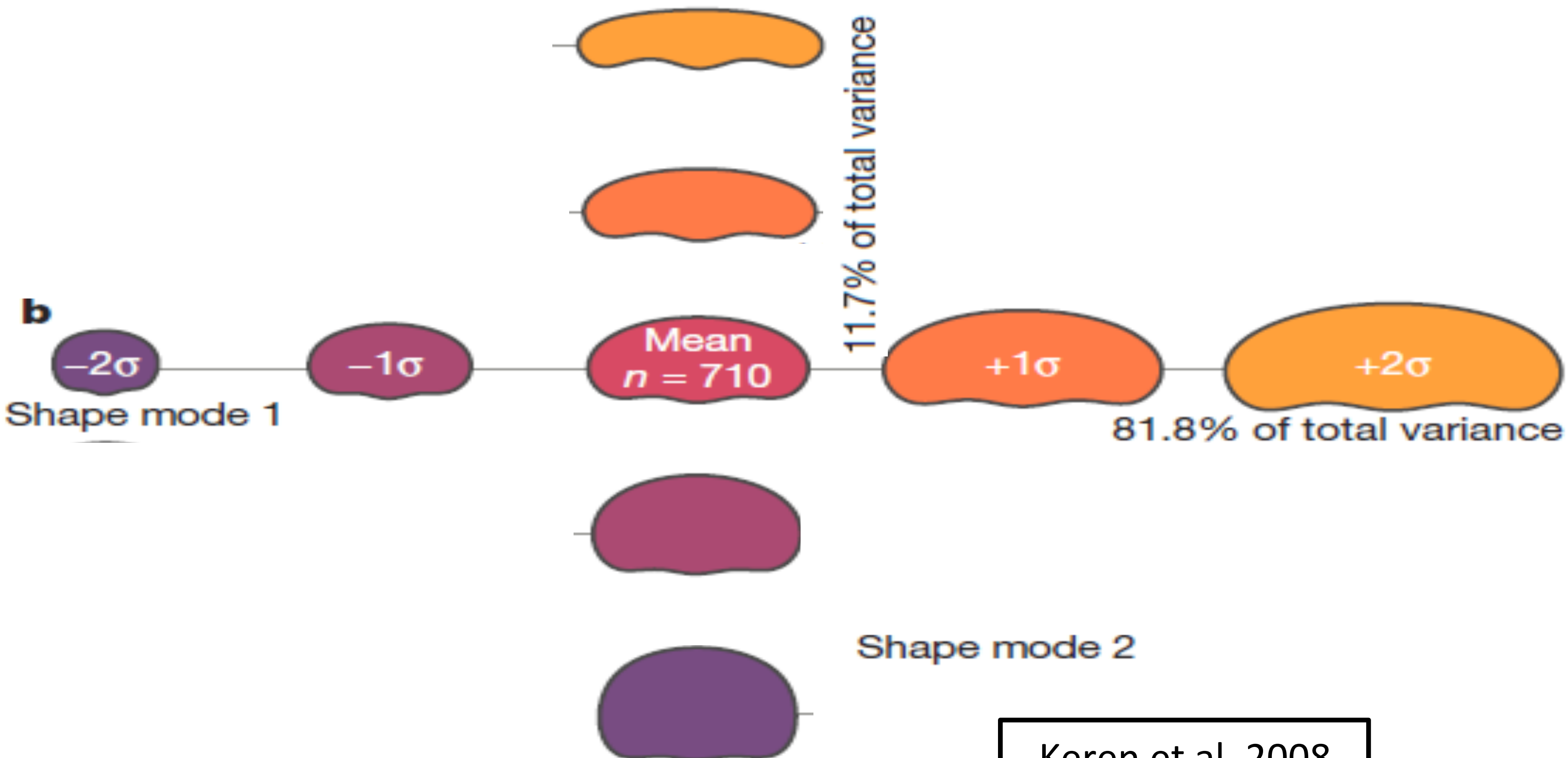
Images showing real shapes

Generative shape model



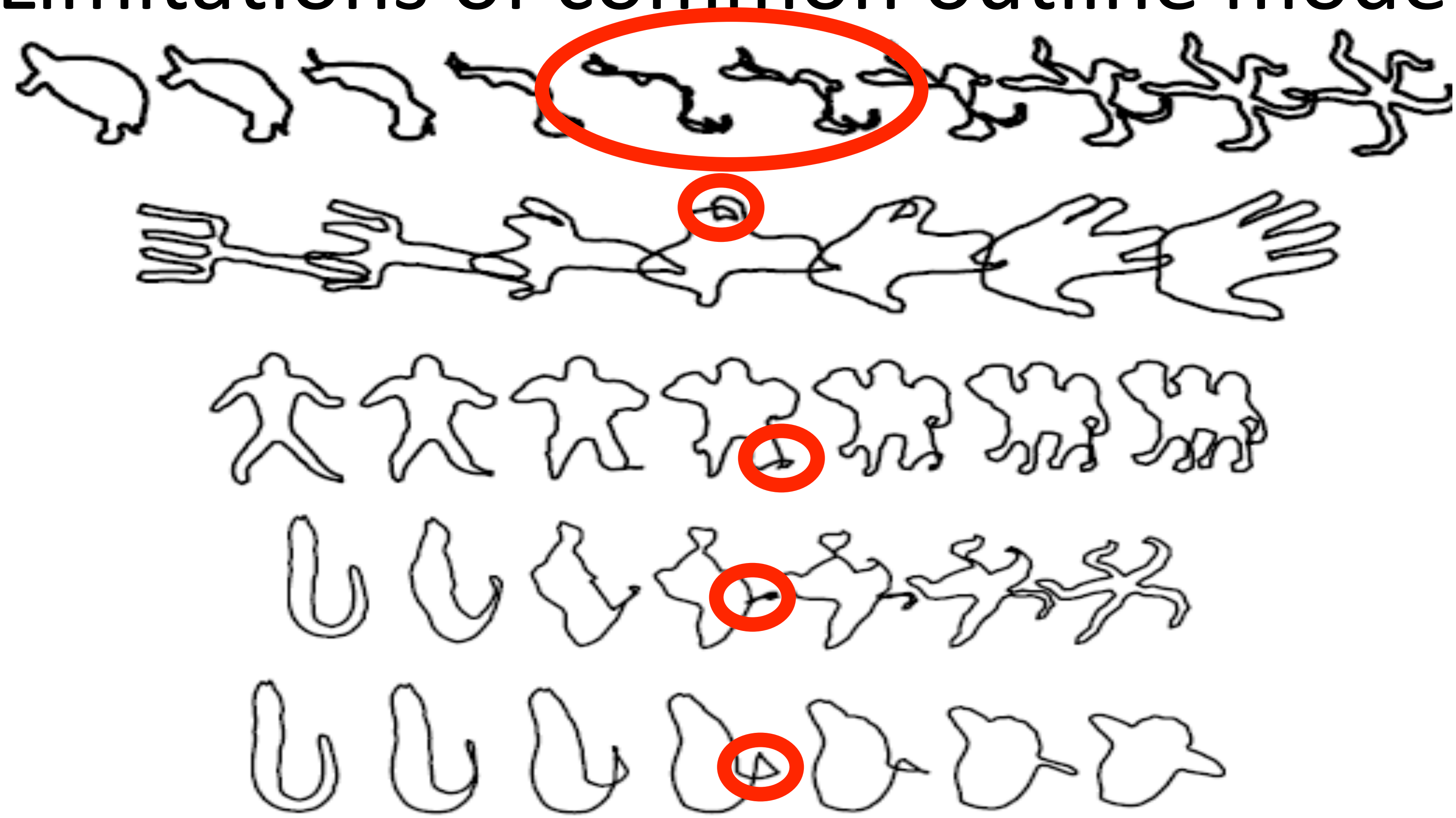
Keren et al. 2008

Shape space



Keren et al. 2008

Limitations of common outline model



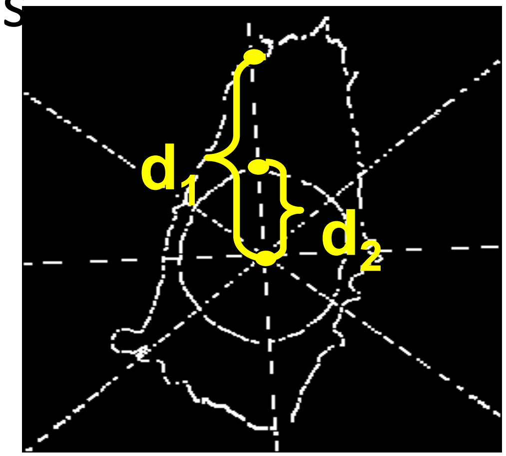
Srivastava et al. 2005

Cell shape: eigenshape ratio model

- Conditioned on nuclear shape:
  - Sample evenly around a circle to represent the shape by radius ratios
  - Parameterization

$$r = d_1 / d_2$$

$$\mathbf{r} \approx \bar{\mathbf{r}} + \sum_{i=1}^{k} \underline{b_i} \boldsymbol{\varphi}_i$$



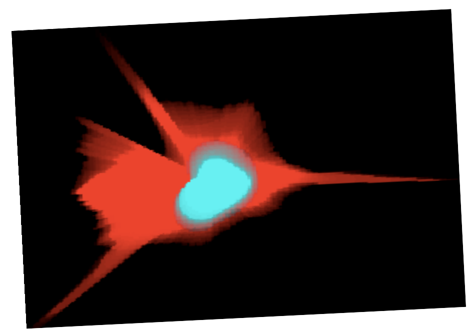
  - Keep 10 principal components (2D)
  - Keep 25 principal components (3D)

However…

- Cells don’t always satisfy assumptions of parametric models.

Segmented PC12 cell

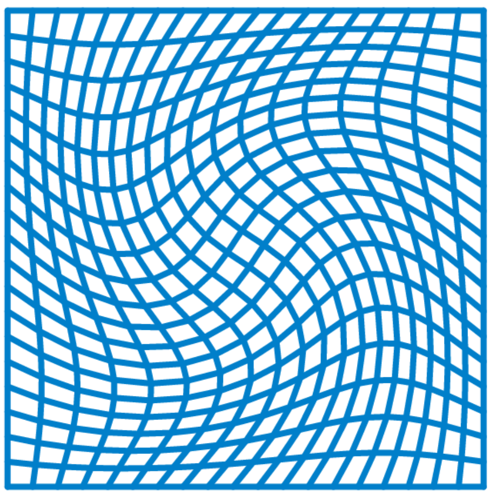
Star-polygon ratio model representation



LDDMM - Large Deformation Diffeomorphic Metric Mapping

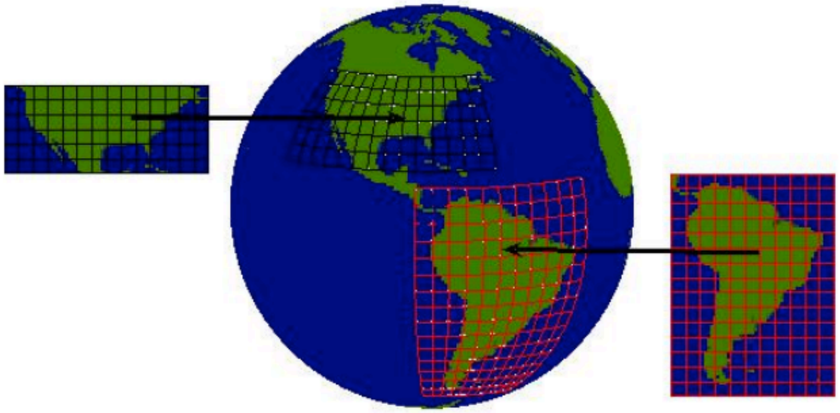
What is a diffeomorphism?

- Essentially a smooth and invertible mapping from one coordinate space to another



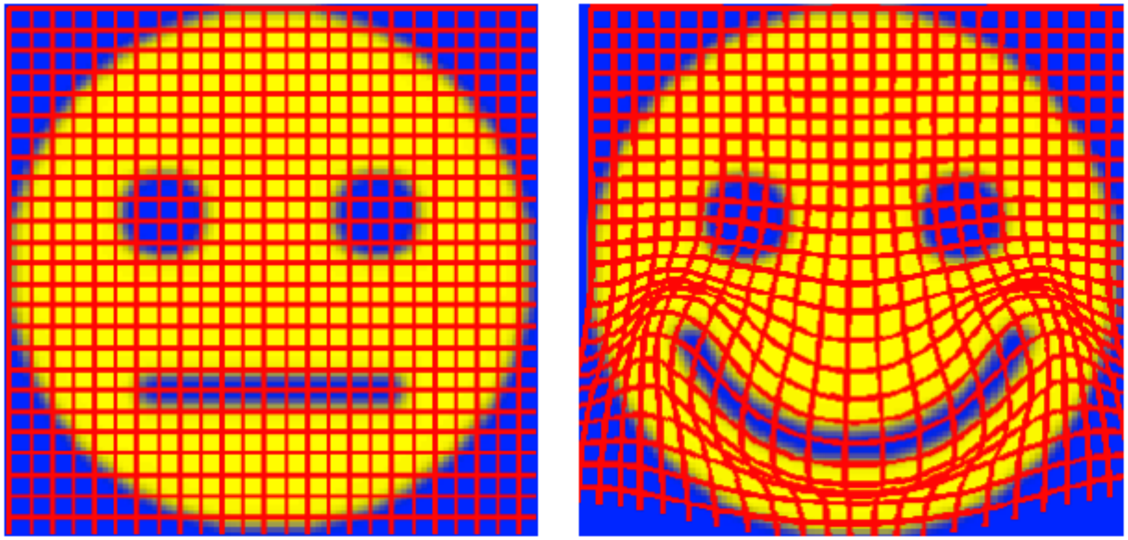
A diffeomorphic mapping from a regular rectangular grid.

https://en.wikipedia.org/wiki/Diffeomorphism



Diffeomorphic mappings of continents to a 2D projection of a globe

http://wwwx.cs.unc.edu/~mn/classes/comp875/doc/diffeomorphisms.pdf



A diffeomorphic mapping from one image to another.

http://wwwx.cs.unc.edu/~mn/classes/comp875/doc/diffeomorphisms.pdf

Nonparametric shape image-based models

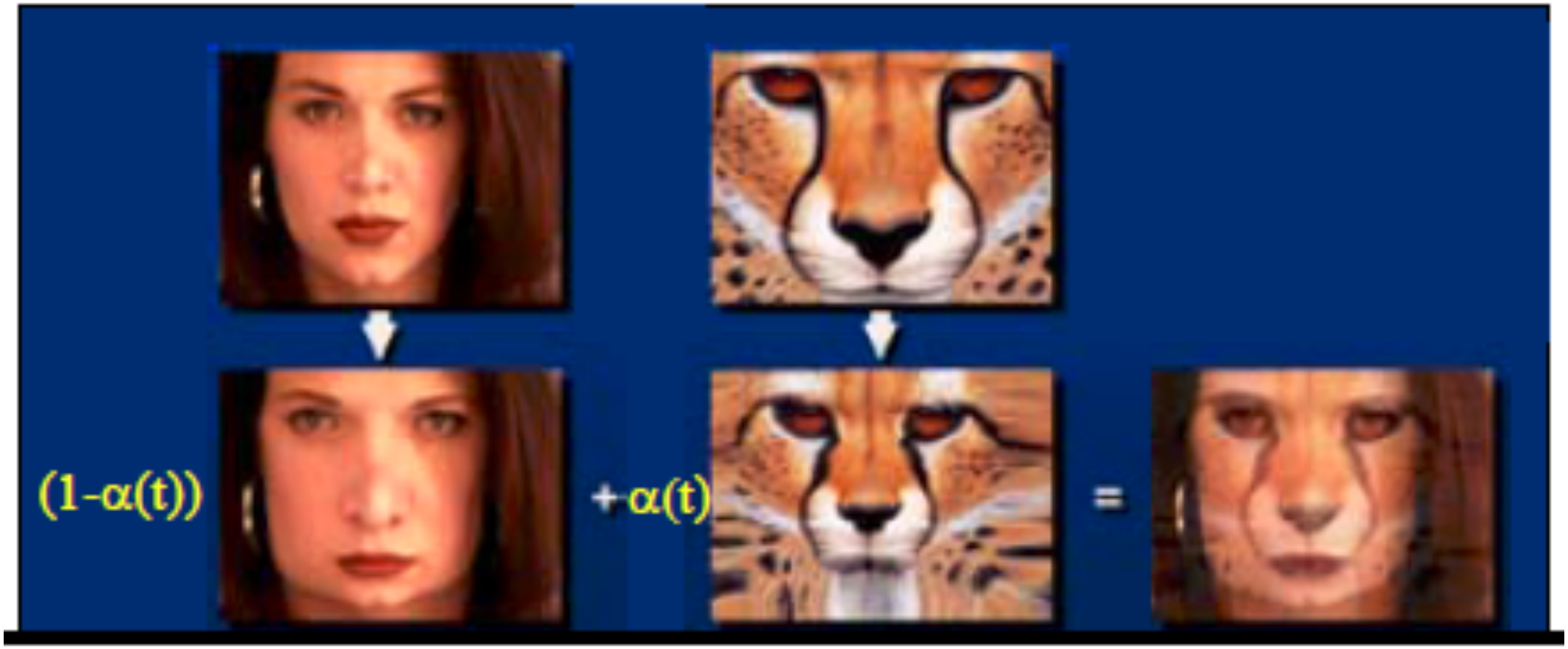
Real 2D nuclear shapes

Peng et al. 2009

Cannot just interpolate images as if they were vectors

http://alumni.media.mit.edu/~maov/classes/comp_photo_vision08f/

Morphing to interpolate images



http://alumni.media.mit.edu/~maov/classes/comp_photo_vision08f/

Distance between two shapes

Work so far

Distance

Shape A

Shape B

0

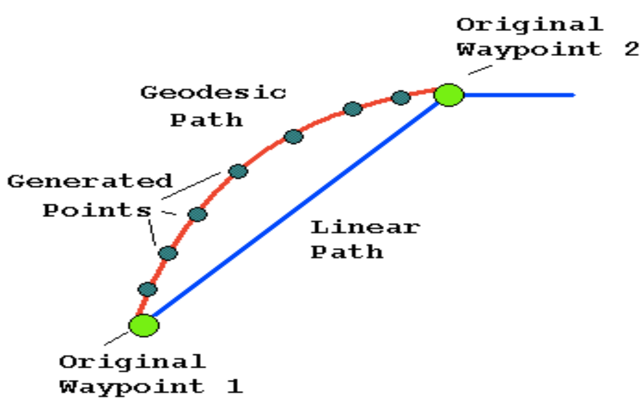
0.0195

Iterative reduction in difference between deformed shape A and B

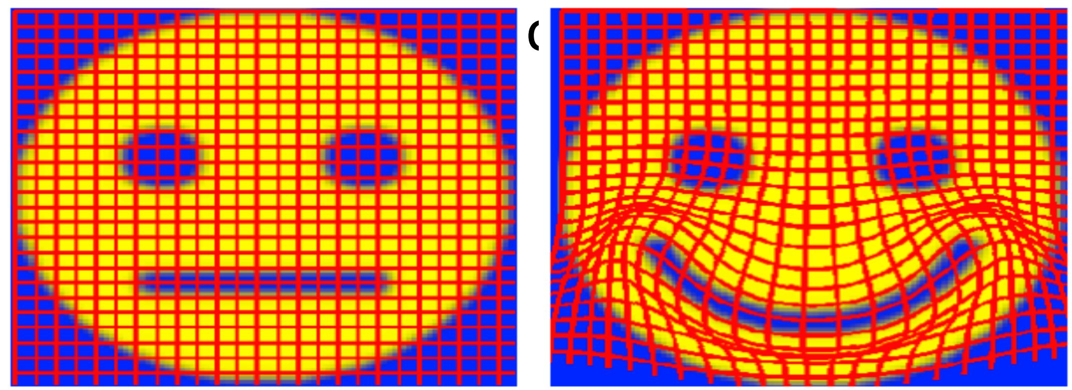
Peng et al.

LDDMM - Large Deformation Diffeomorphic Metric Mapping

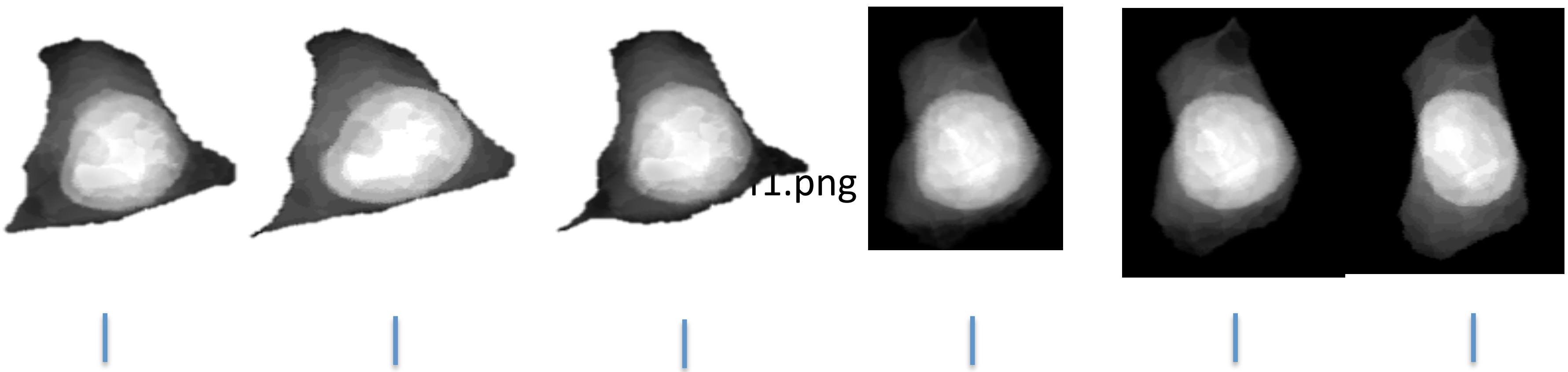
- Minimal energy transformation with respect to the gradient of the deformation field i.e. Geodesic distance
- Deformation field is a nonlinear manifold that contains the information of the image, including gradient, second order dat (



Shadel 1974



http://wwwx.cs.unc.edu/~mn/classes/comp875/doc/diffeomorphisms.pdf


1.png

Constructing a cell shape space

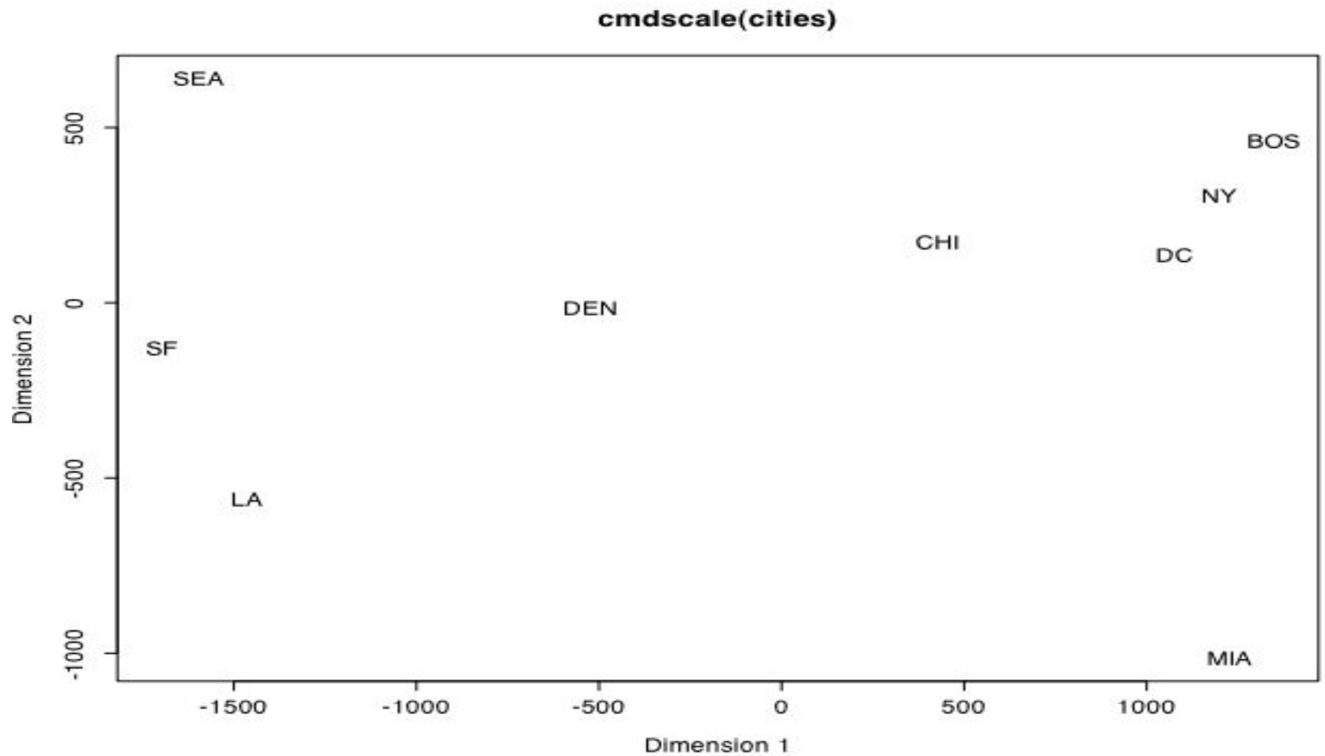
- Find distances of every cell to every other cell
- Try to find a “map” that puts each cell the correct distance from the others (i.e., puts cells with short distances near each other)

Distance matrix…

	BOS	CHI	DC	DEN	LA	MIA	NY	SEA	SF
BOS	0	963	429	1949	2979	1504	206	2976	3095
CHI	963	0	671	996	2054	1329	802	2013	2142
DC	429	671	0	1616	2631	1075	233	2684	2799
DEN	1949	996	1616	0	1059	2037	1771	1307	1235
LA	2979	2054	2631	1059	0	2687	2786	1131	379
MIA	1504	1329	1075	2037	2687	0	1308	3273	3053
NY	206	802	233	1771	2786	1308	0	2815	2934
SEA	2976	2013	2684	1307	1131	3273	2815	0	808
SF	3095	2142	2799	1235	379	3053	2934	808	0

http://personality-project.org/r/mds.html

… to coordinates

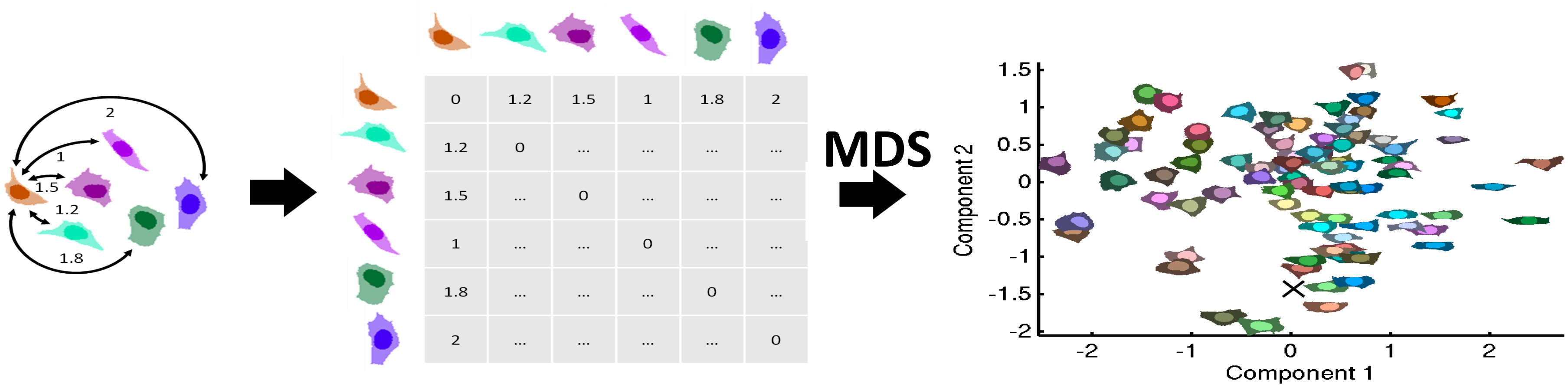


http://personality-project.org/r/mds.html

Shape space

Diffeomorphic Training

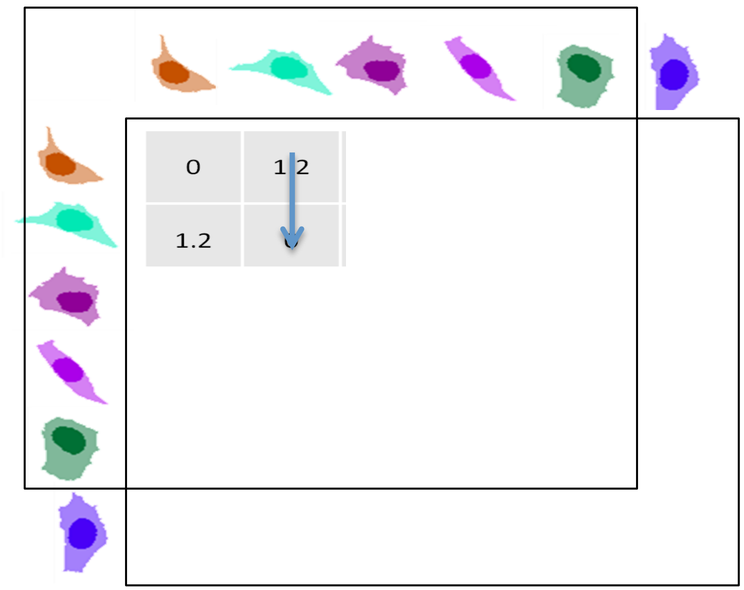
Shapes to Space



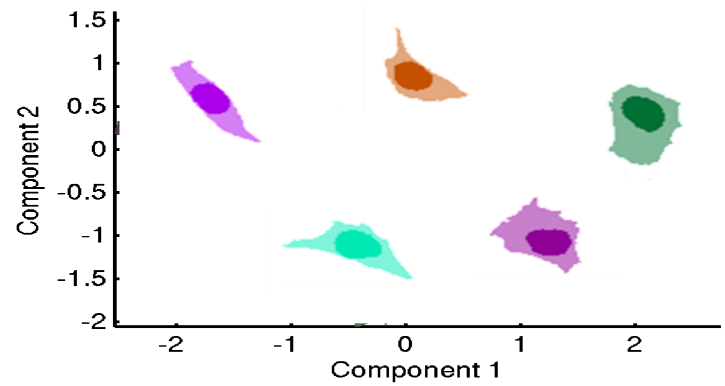
But this takes a lot of time

Partial Distance Matrix Learning

- Most complete shape space



MDS



Shape Space Embedding

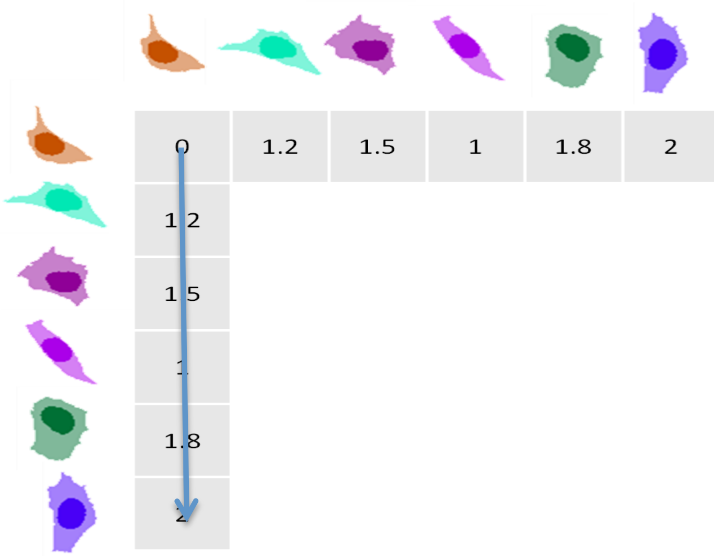
- This embedding method requires all pairs of distances
- Let's say we have 250 cells and it takes ~30 sec to register a pair
- $(30 \text{ sec} * 250^2)/2 \approx 10$ days. Way too long…..
- can we infer embedding with missing data?
- MDS with missing data

$$\{x_1, \ldots, x_n\} = \ argmin_{\{x_1, \ldots x_n\}} \sum_{i<j\leq n} w_{i,j}\left(d\left(x_i, x_j\right) - D_{i,j}\right)^2$$

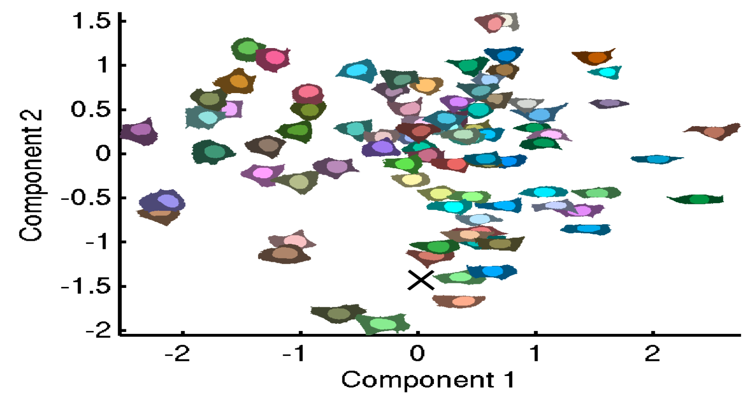
where $w_{i,j}$ is the weight of observation $D_{i,j}$

Partial Distance Matrix Learning

- Landmark MDS

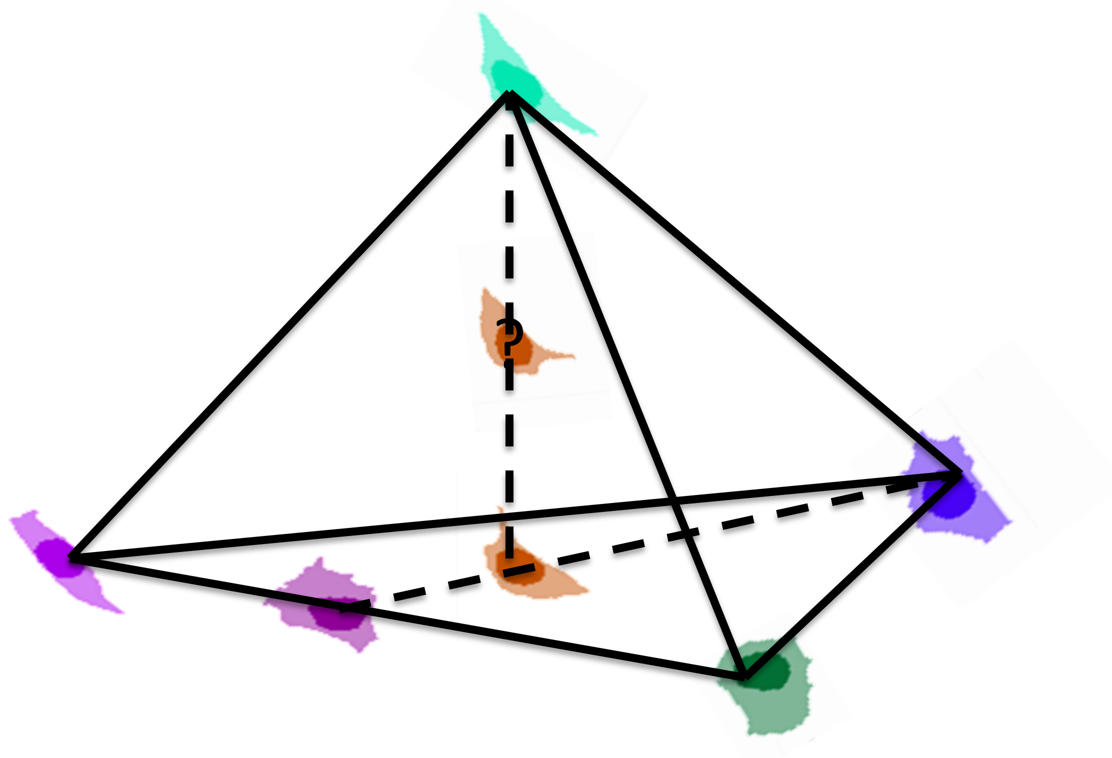


Approximate MDS



Diffeomorphic Synthesis

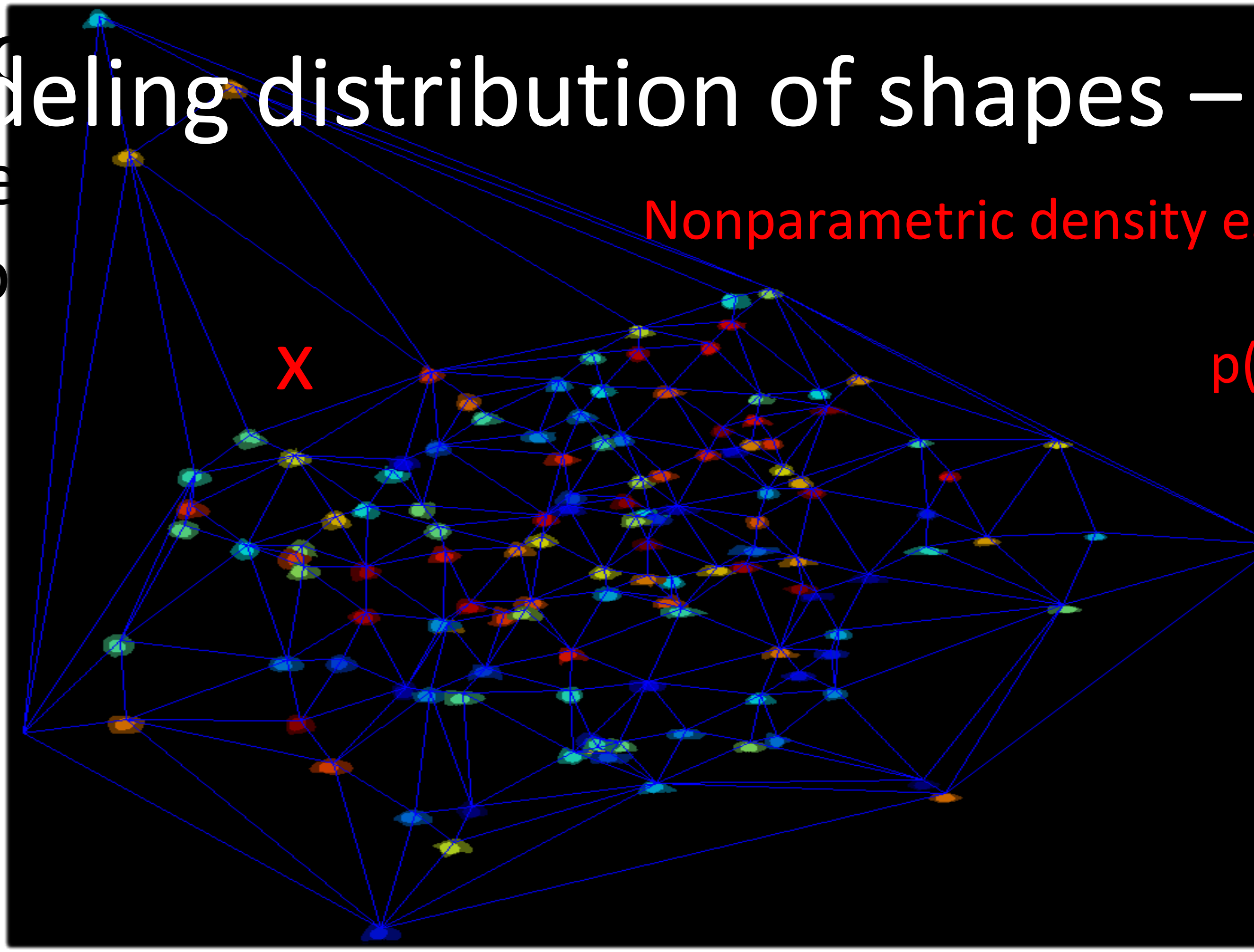
Space to Shapes



Synthesis strategy for new points

Modeling distribution of shapes –

- The
  pro

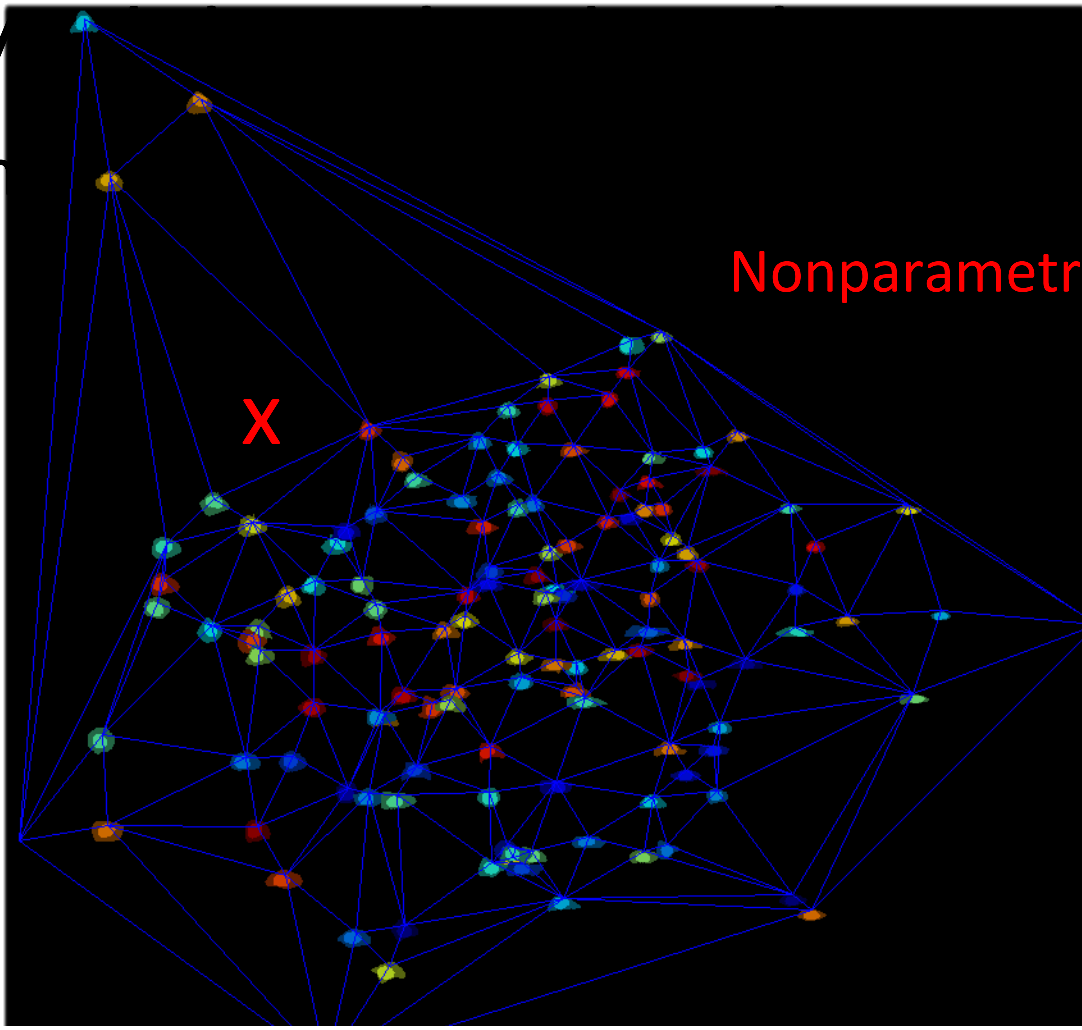


Nonparametric density estimation

$p(x) = 1/v_i n$

M   of shapes

- Th   robability density.

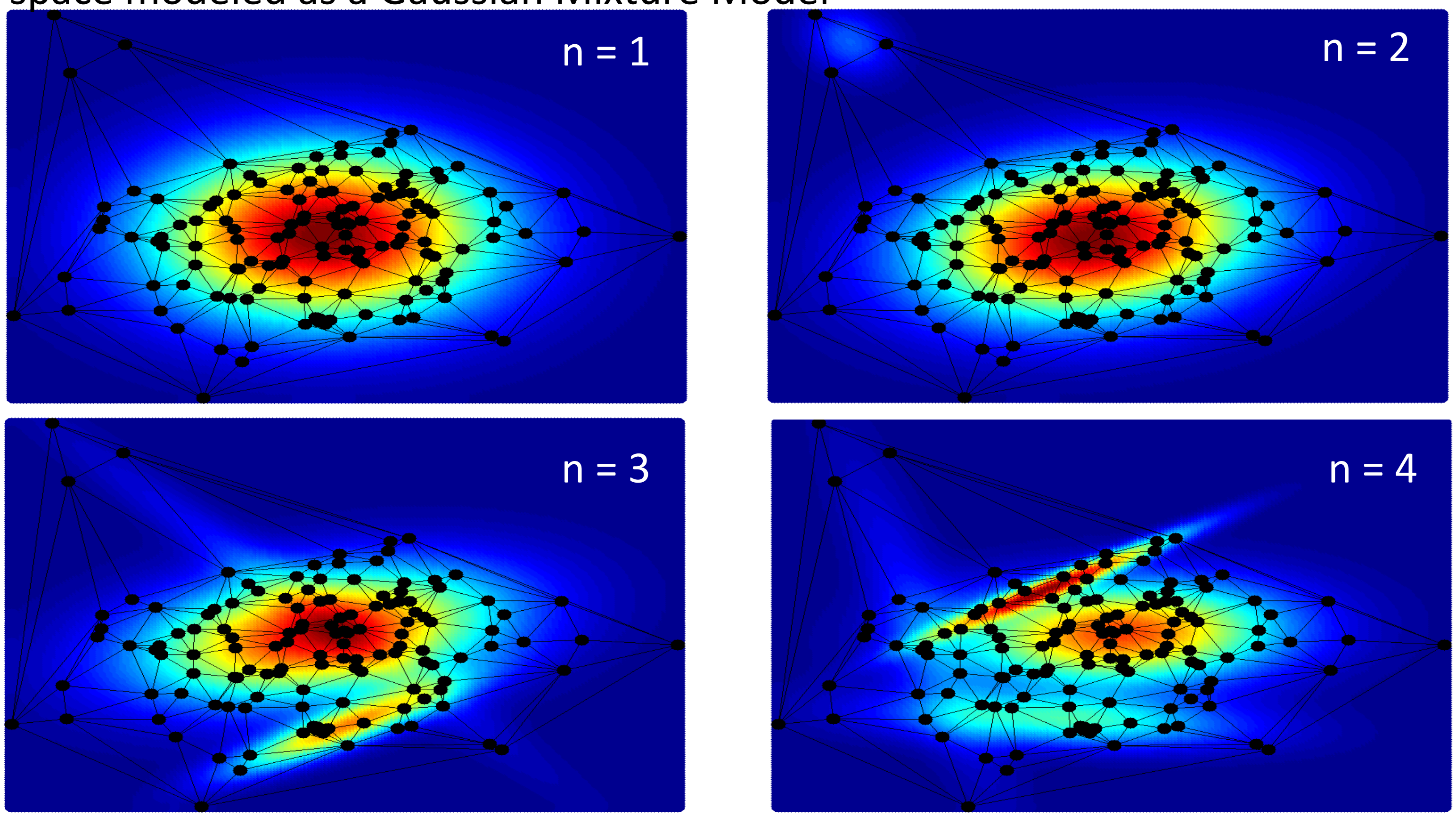


Nonparametric density estimation

$p(x) = 1/v_i n$

Modeling distribution of shapes – p(x)

Shape space modeled as a Gaussian Mixture Model

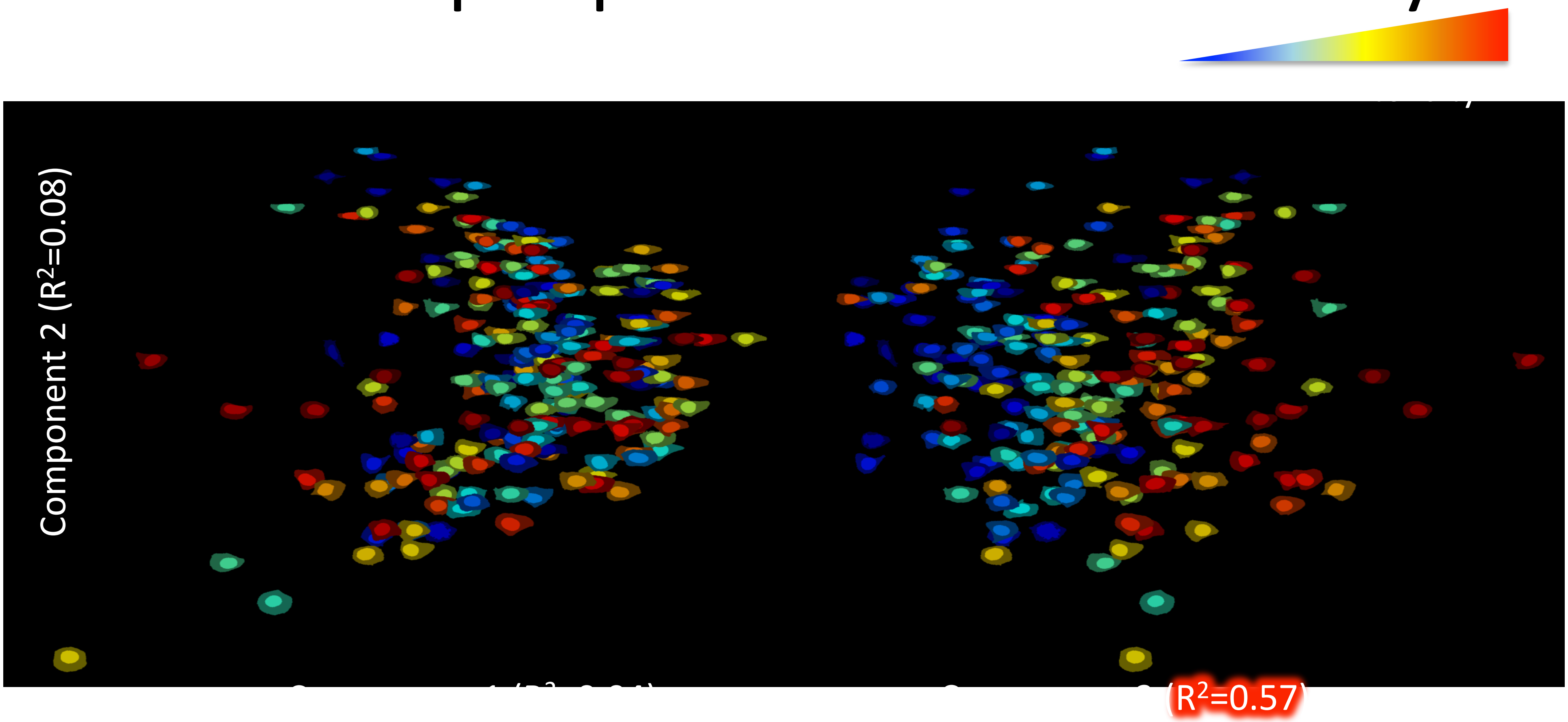


Diffeomorphic space

- New feature space
  - Positions in space correspond to a real image
  - Feature dimensions correspond to dimensions with highest eigenvalues

  - Can be treated exactly like a normal feature space

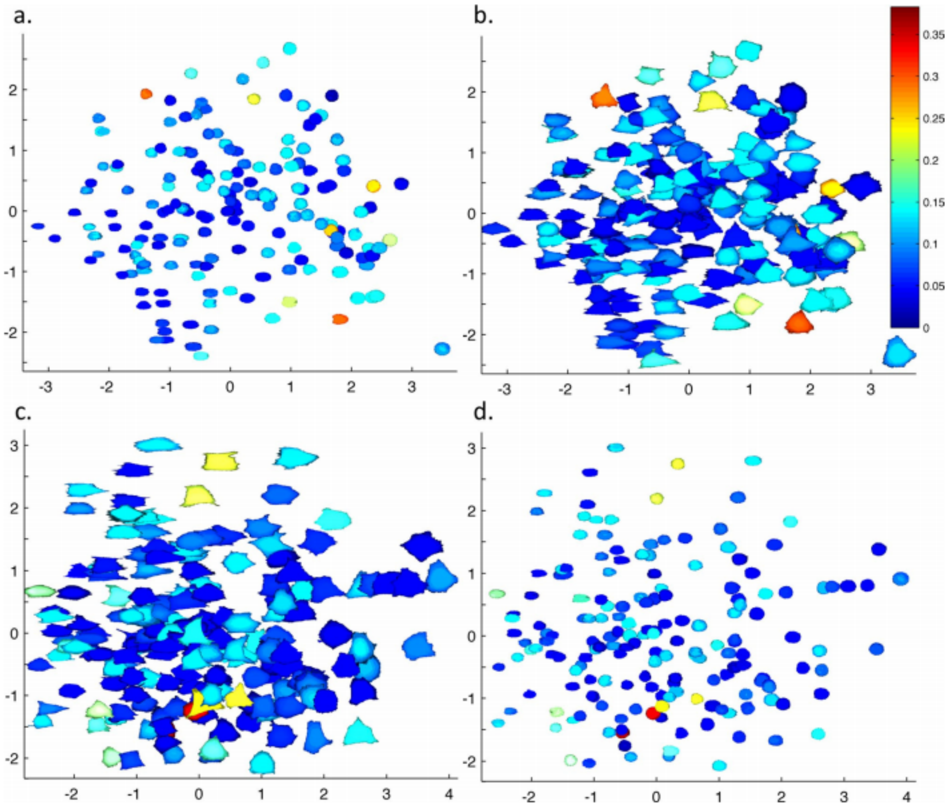
HeLa shape space with DNA intensity



Do cell and nuclear shape depend on each other?

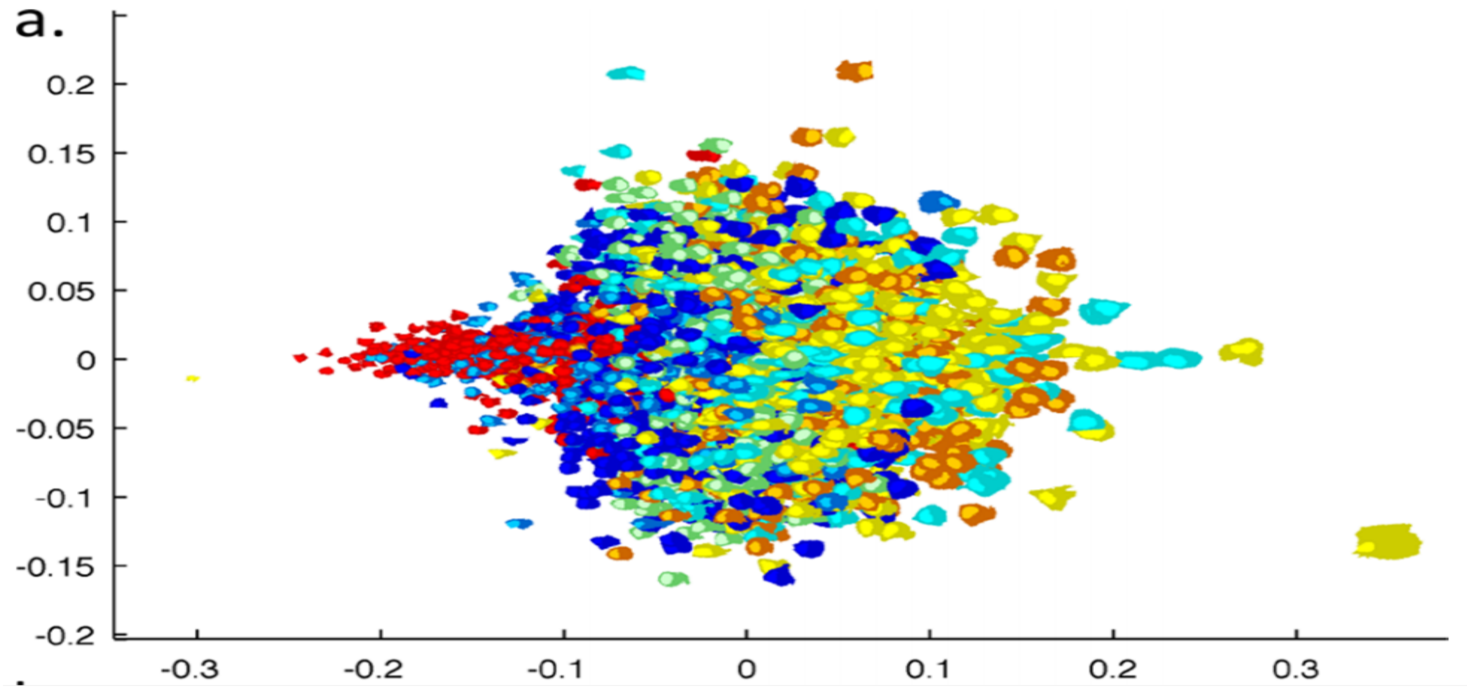
- Build shape spaces for
  - nuclear shapes only
  - cell shapes only
- For each nuclear shape, predict a position in cell shape space for it by interpolating at the same relative distances from the cell shapes of its neighbors in nuclear shape (and vice versa)
- Measure prediction error as the distance in the shape space between the predicted position and the actual position

Prediction of cell and nuclear shape dependency



Johnson et al. 2015

H1299 shape space, colored by protein label



Johnson et al. 2015