MBioS

An NIH Biomedical Technology Research Center

Image-derived spatiotemporal models of subcellular organization, differentiation and perturbation

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

RAY AND STEPHANIE LANE
Center for Computational Biology
Carnegie Mellon

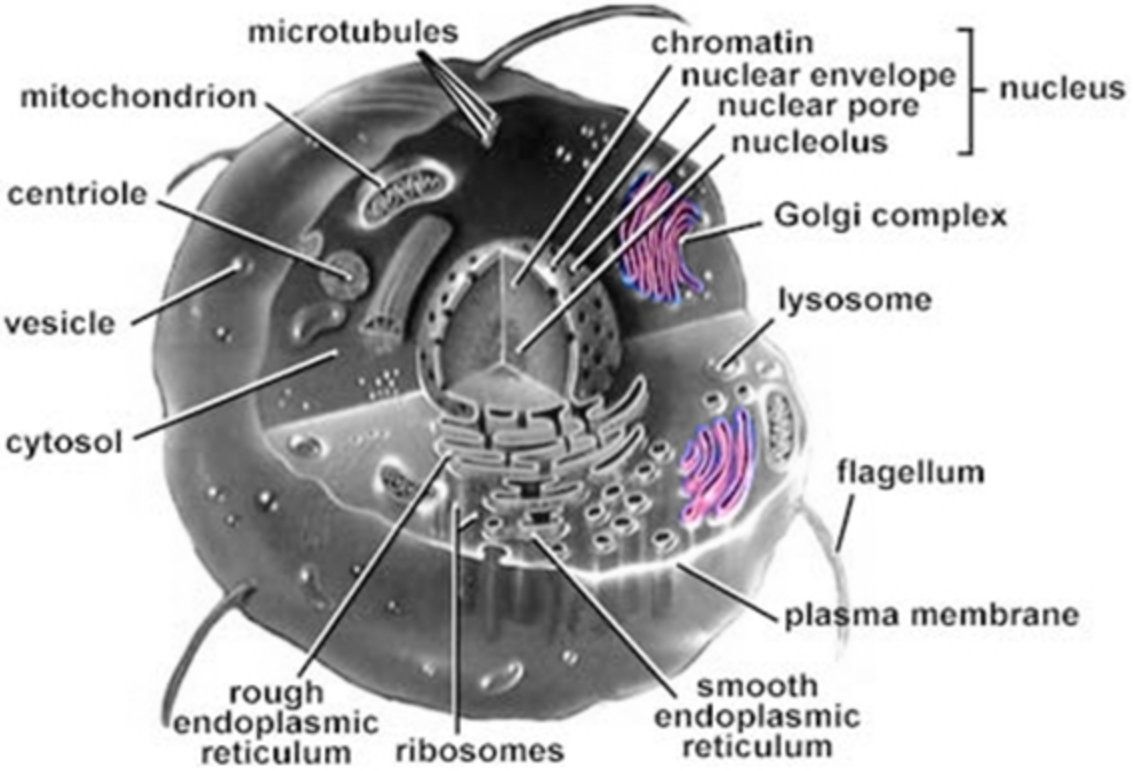
ALBERT-LUDWIGS-
UNIVERSITÄT FREIBURG

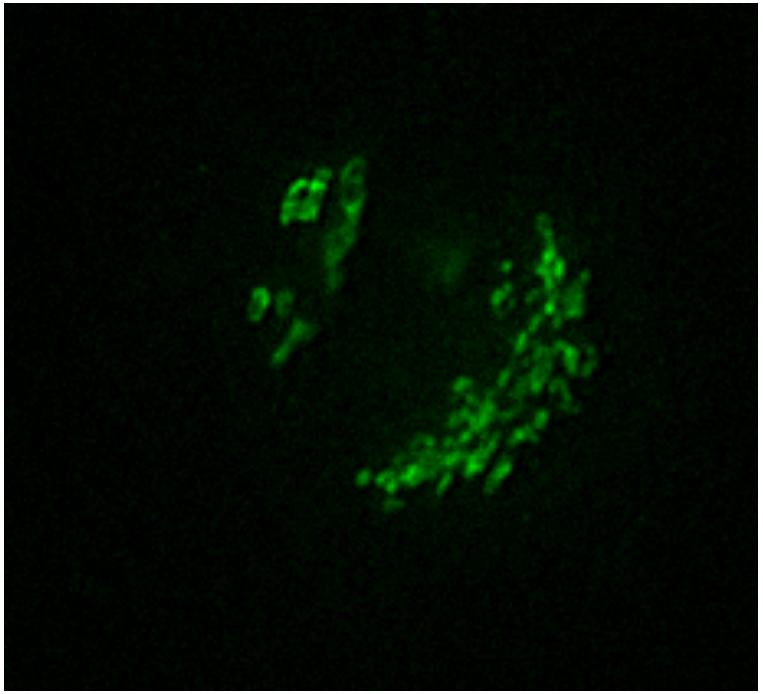
FRIAS
FREIBURG INSTITUTE
FOR ADVANCED STUDIES

Spatially-accurate cell simulations

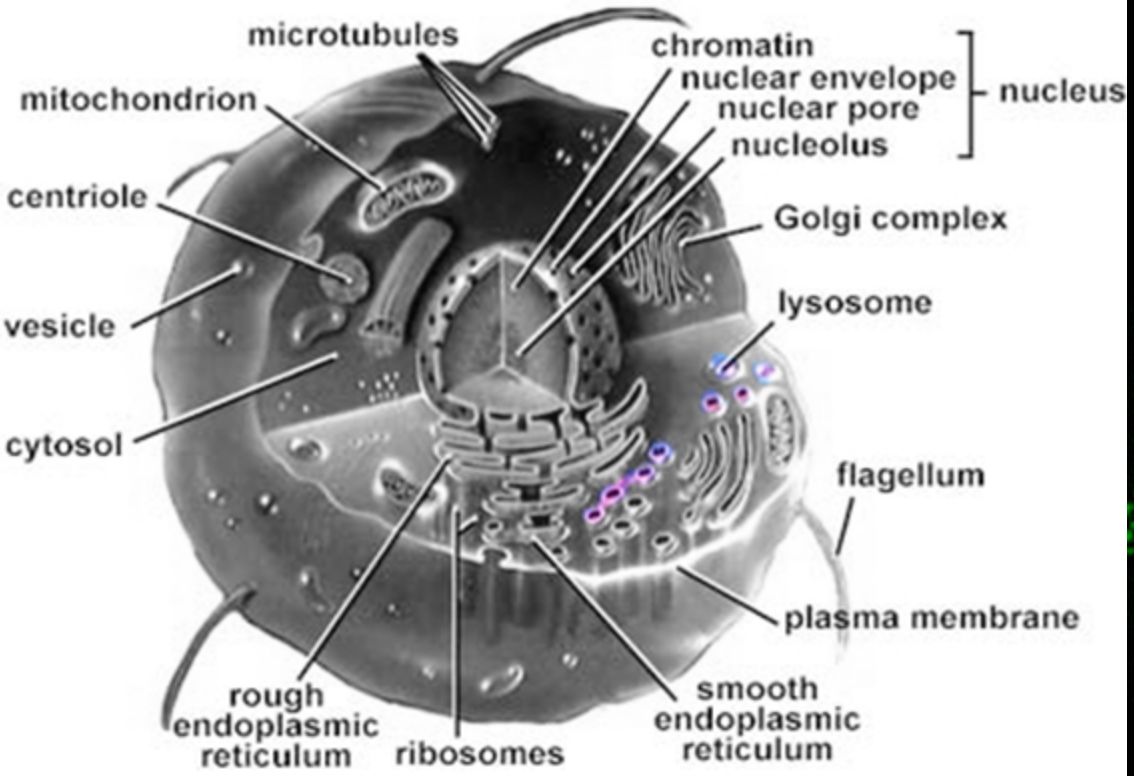
- Where do we get accurate information on the spatial distribution of proteins/ organelles in order to incorporate it into cell/tissue simulations?
- How do we learn and predict cell-type specific cell organization differences and how they are affected by perturbagens?

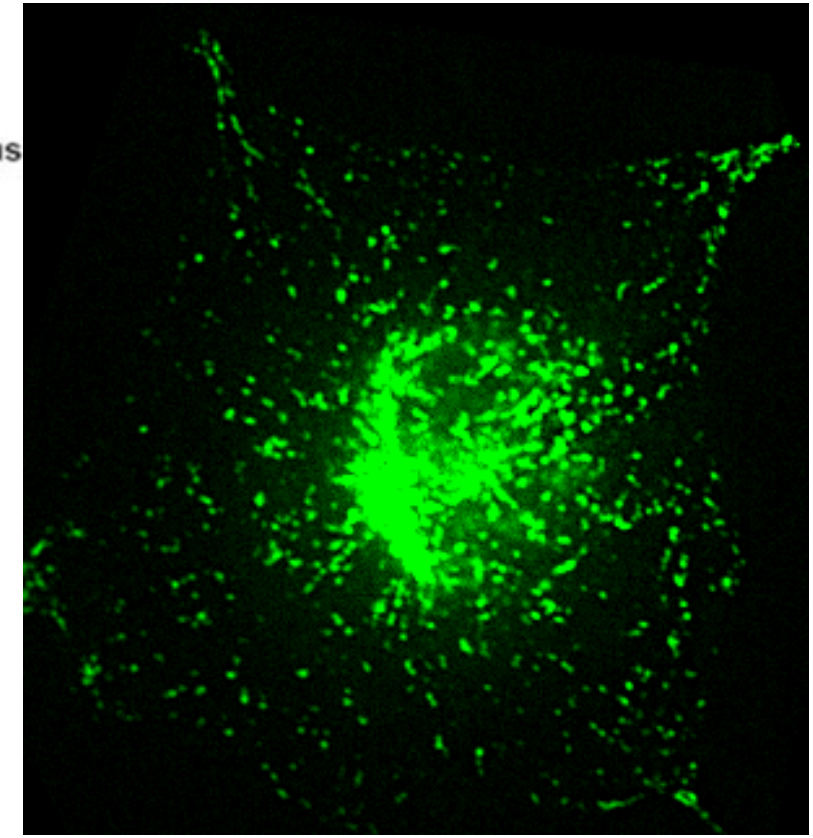
Subcellular Location



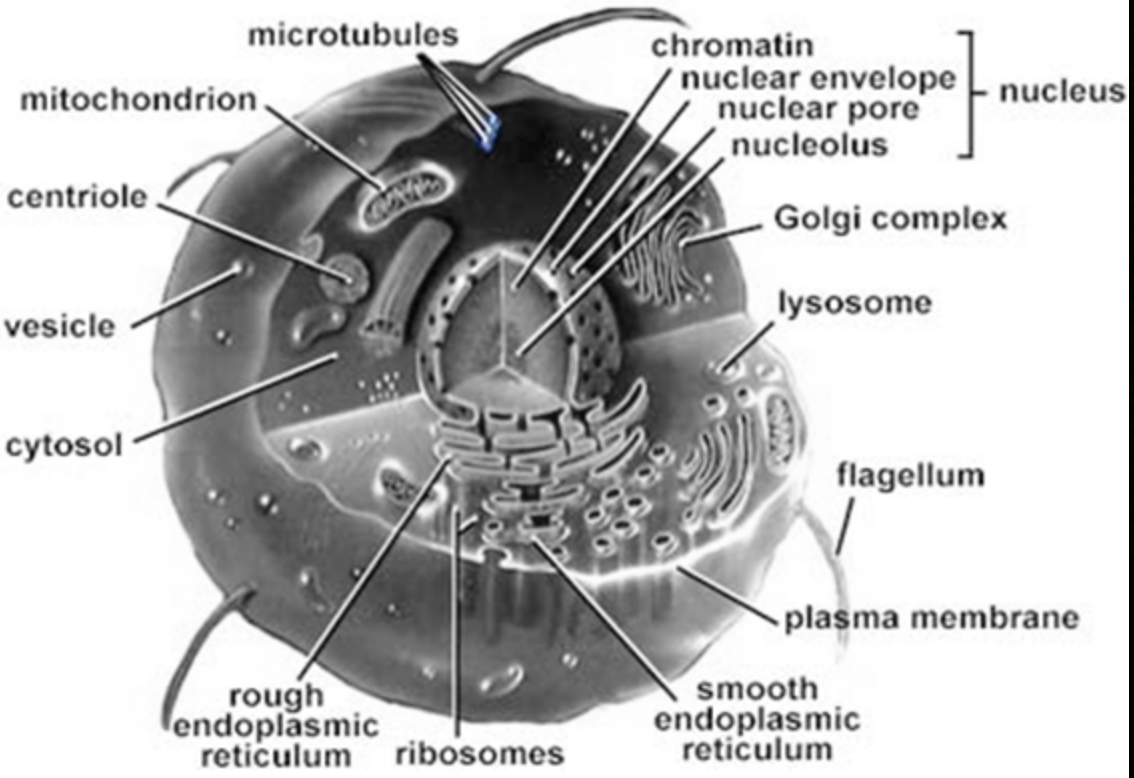


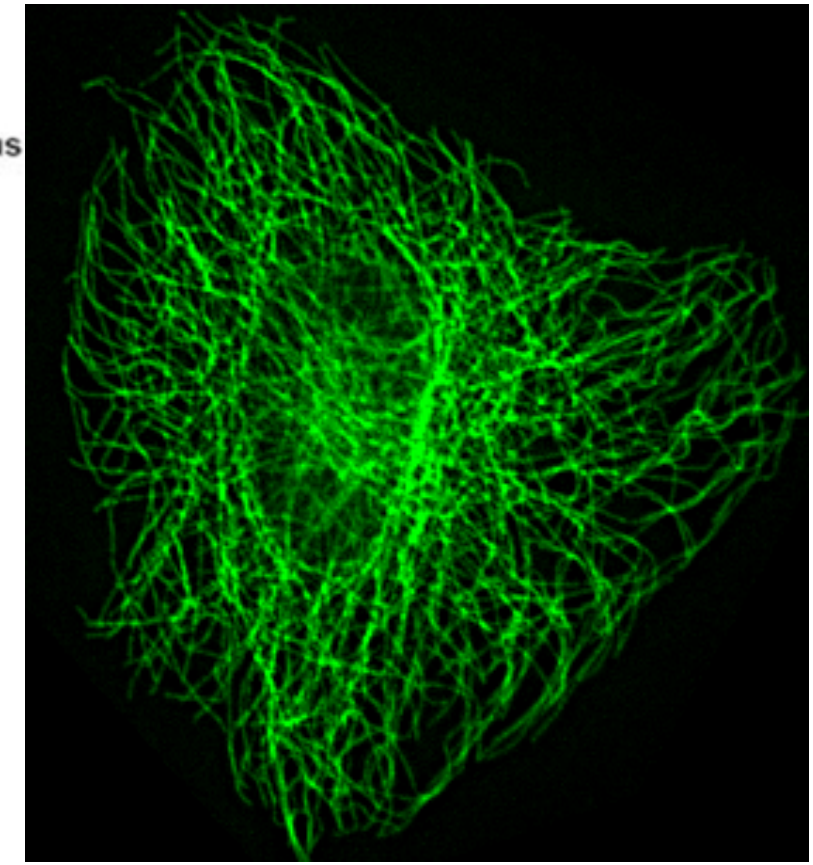
Subcellular Location



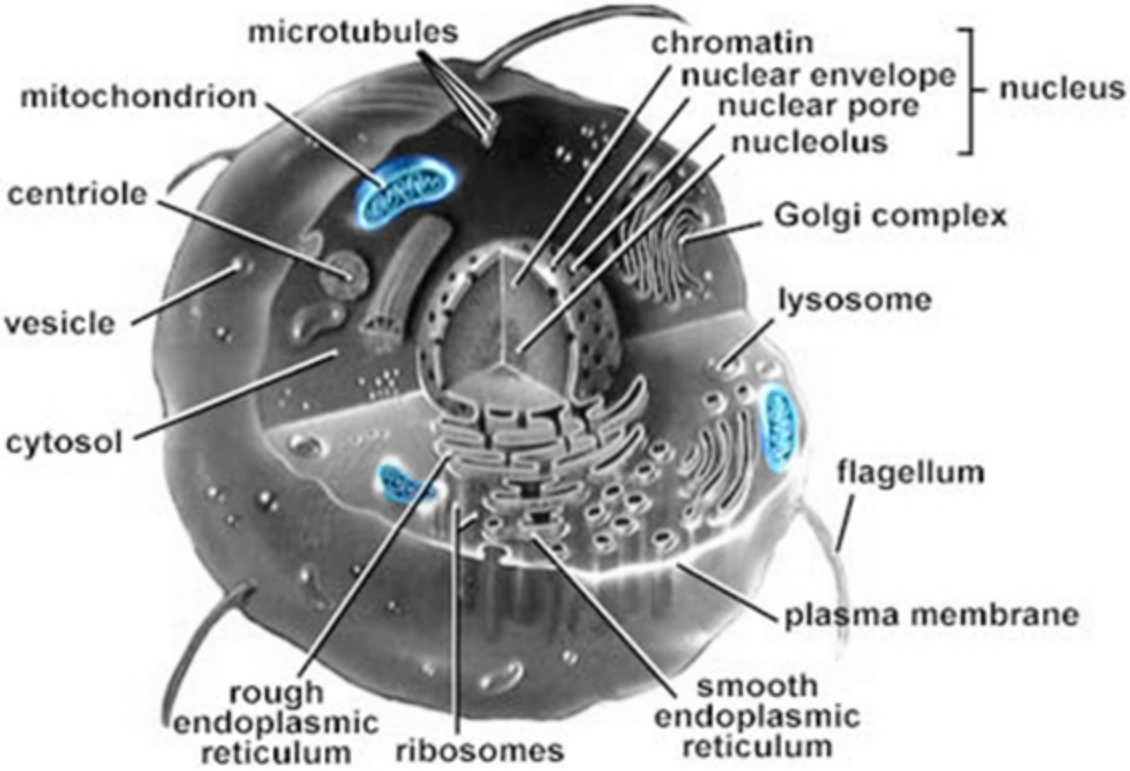


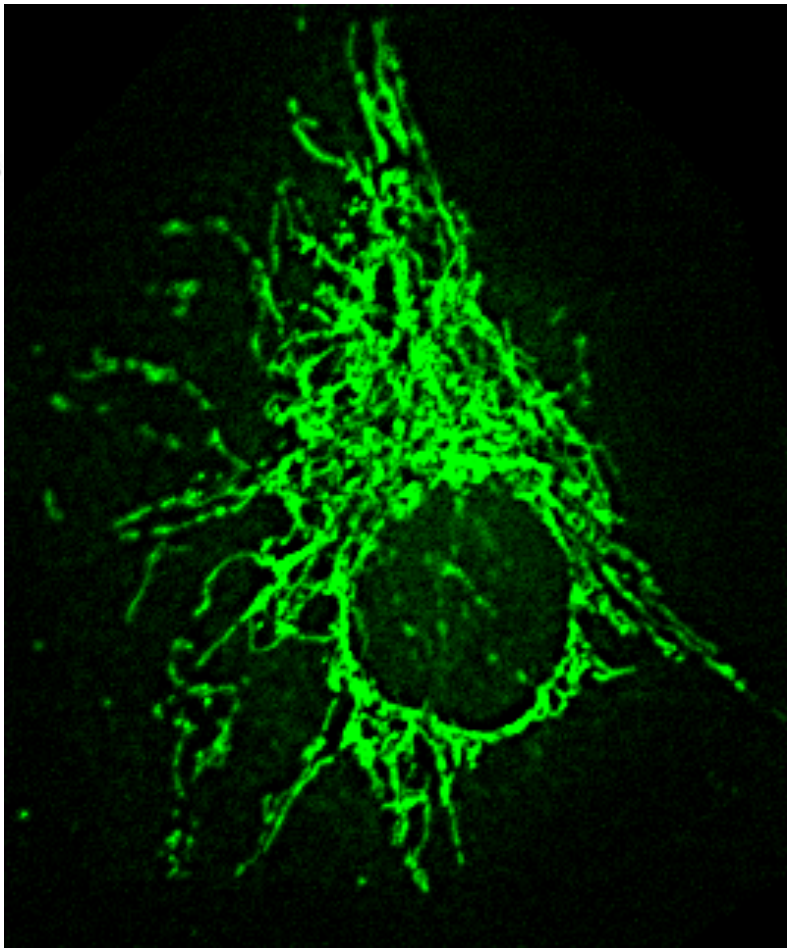
Subcellular Location





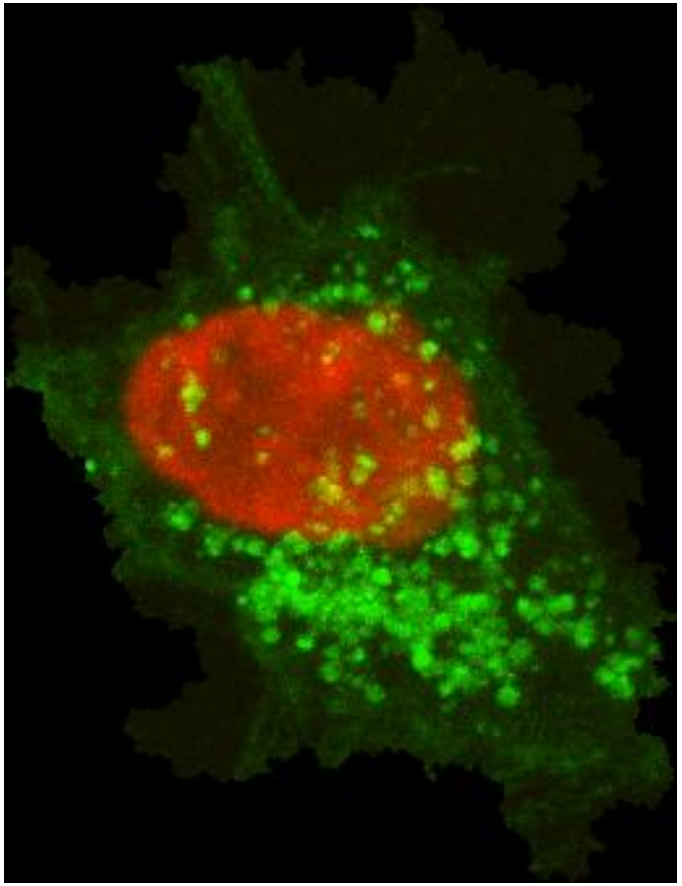
Subcellular Location





Traditional approach

- Take fluorescence microscope images of a tagged protein
- “Convert” them into words/GO terms to describe its subcellular location



Traditional approach

- Use visual inspection of images
- Assumes that
  - we know the “classes” that proteins should be placed in (e.g., GO terms)
  - people are good at assigning the terms after looking at one or more images

Supervised Machine Learning

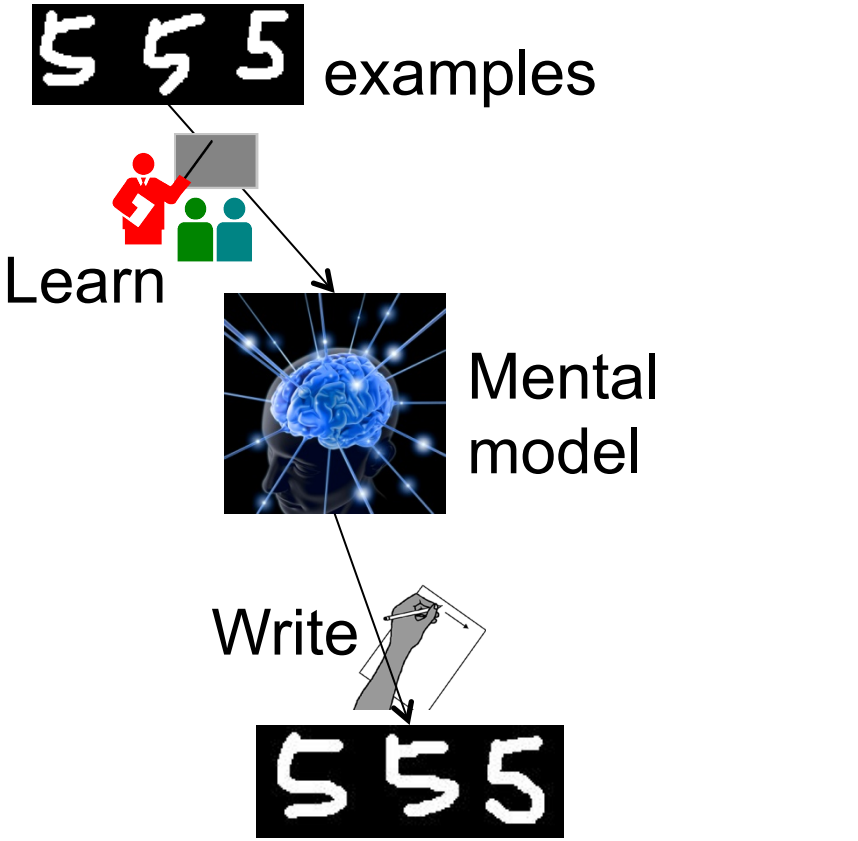
- Design features to describe subcellular patterns
- Use examples of images of proteins “known” to be in different subcellular patterns to train classifier

However…

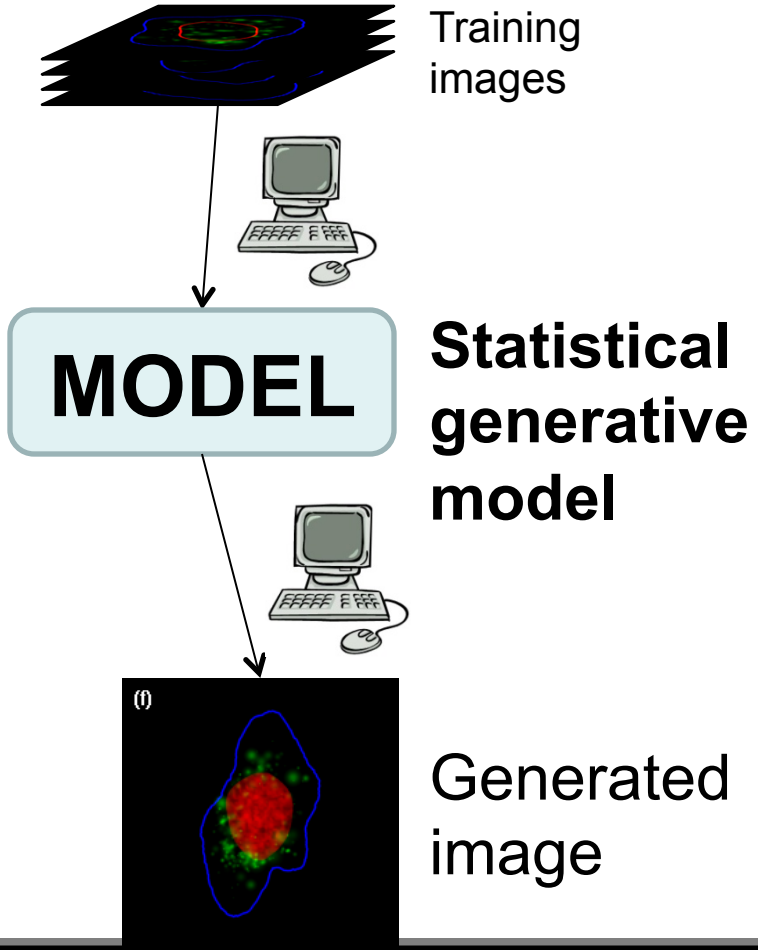
- Assigning words not sufficient
  - Knowing that apples and oranges can be distinguished by their color does not allow you to understand how either are formed
- Descriptive models may be wrong if features detect artifacts correlated with classes (e.g., background or illumination variations from different wells)

Alternative: Generative Modeling

- Human cognition



- Generative model

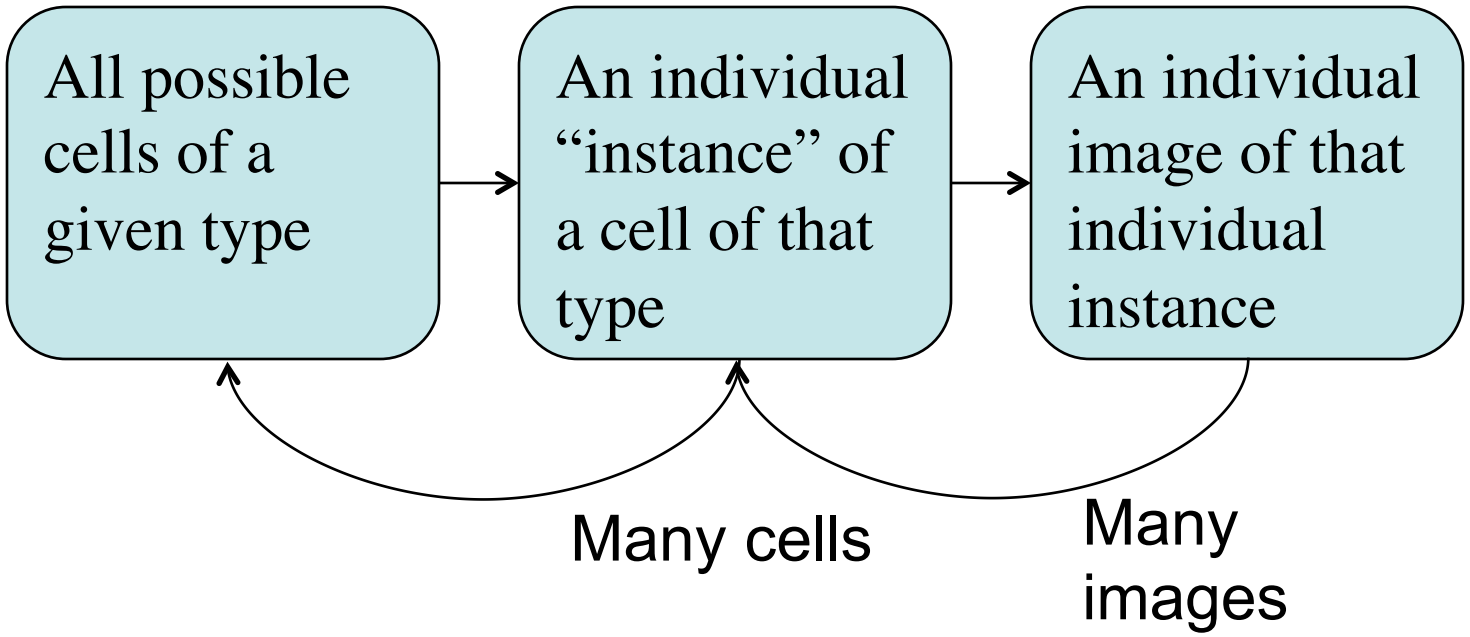


Descriptive vs. Generative Models

- Goal of descriptive models is to allow us to distinguish instances that we are given
  - Need “just enough” description
- Goal of generative models is to be able to create new instances
  - Need “complete” description

Generative models of images

- We seek the “underlying” model from which images are drawn



Classical inverse problem with statistical twist

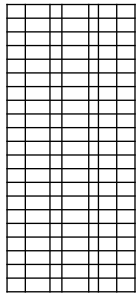
- Learn underlying reality observed via imaging
- Extensive work on image **reconstruction** to create a (higher resolution?) model of a conserved structure (e.g., nuclear pore, ribosome) by *removing noise and variation*
- Our goal is learning *statistical,* ***generative*** model of reality sampled via imaging by removing noise but *keeping* variation

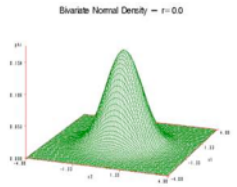
Making models generative

estimate generative parameters for each cell

build stat model

Many cell images





μ Σ

randomly sample new parameters

Synthetic cell images

http://CellOrganizer.org

CellOrganizer

Images ↔ Models

RAY AND STEPHANIE LANE Center for Computational Biology | Carnegie Mellon University | Center for Bioimage Informatics

Home
People
Publications
Downloads

May 17, 2013: Version 1.9.0 released!
New: Now allows synthesis of cell and nuclear shape instances for Hela cells using a diffeomorphic model.

The *CellOrganizer* project provides tools for

- learning generative models of cell organization directly from images
- storing and retrieving those models in XML files
- synthesizing cell images (or other representations) from one or more models

Model learning captures variation among cells in a collection of images. Images used for model learning and instances synthesized from models can be two- or three-dimensional static images or movies.

CellOrganizer can learn models of

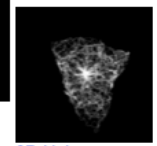
- cell shape
- nuclear shape
- chromatin texture
- vesicular organelle size, shape and position
- microtubule distribution.

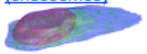
These models can be *conditional* upon each other. For example, for a given synthesized cell instance, organelle position is dependent upon the cell and nuclear shape of that instance.

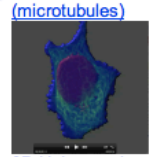
Cell types for which generative models for at least some organelles have been built include human HeLa cells, mouse NIH 3T3 cells, and Arabidopsis protoplasts. Planned projects include mouse T lymphocytes and rat PC12 cells.

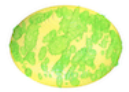
Synthesized Cell Images (click to view)


2D HeLa (endosomes)


3D HeLa (microtubules)


3D HeLa (mitochondria)

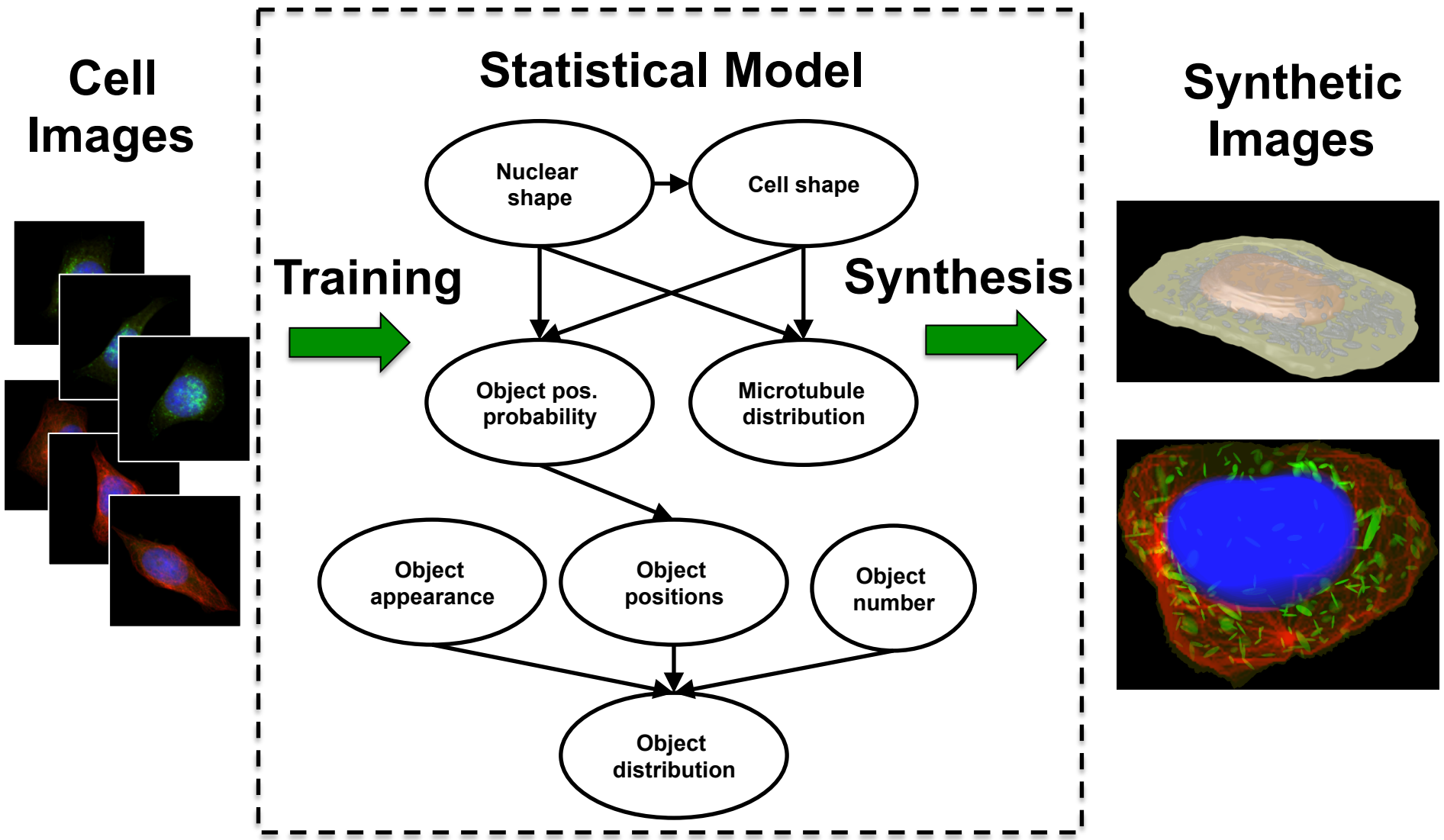

3D HeLa movie


3D protoplast (chloroplasts)

Support for *CellOrganizer* has been provided by grants GM075205 and GM090033 from the National Institute of General Medical Sciences, grants MCB1121919 and MCB1121793 from the U.S. National Science Foundation, by a Forschungspreis from the Alexander von Humboldt Foundation, and by the School of Life Sciences of the Freiburg Institute for Advanced Studies.

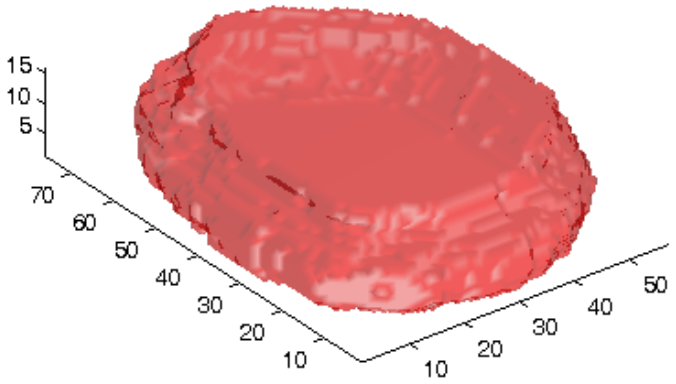
NSF | Alexander von Humboldt Stiftung/Foundation | Albert-Ludwigs-Universität Freiburg | FRIAS Freiburg Institute for Advanced Studies | Life Sciences - Lifenet

CellOrganizer

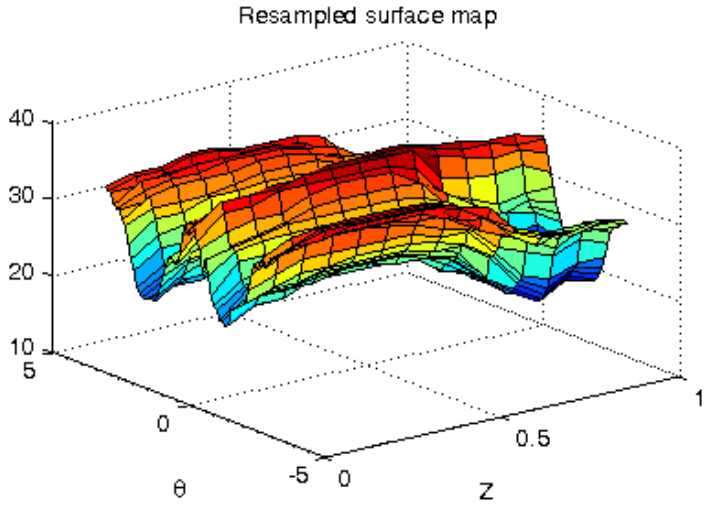


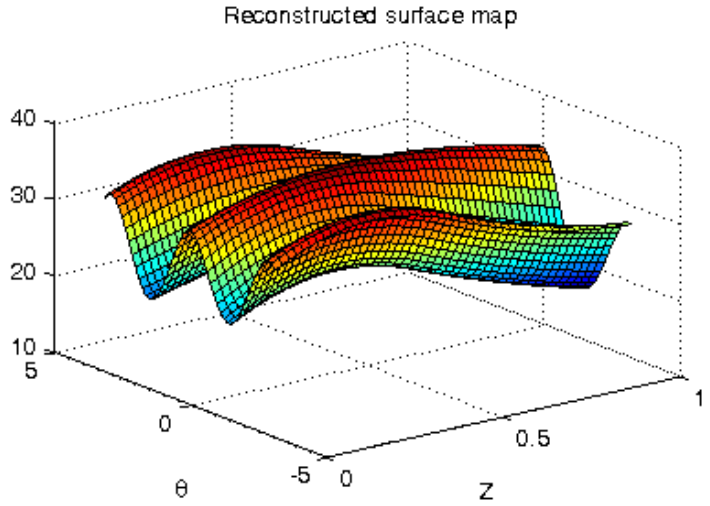
PARAMETRIC MODELS OF NUCLEAR SHAPE

3D Nuclear Shape – Cylindrical Spline Surface



Tao Peng

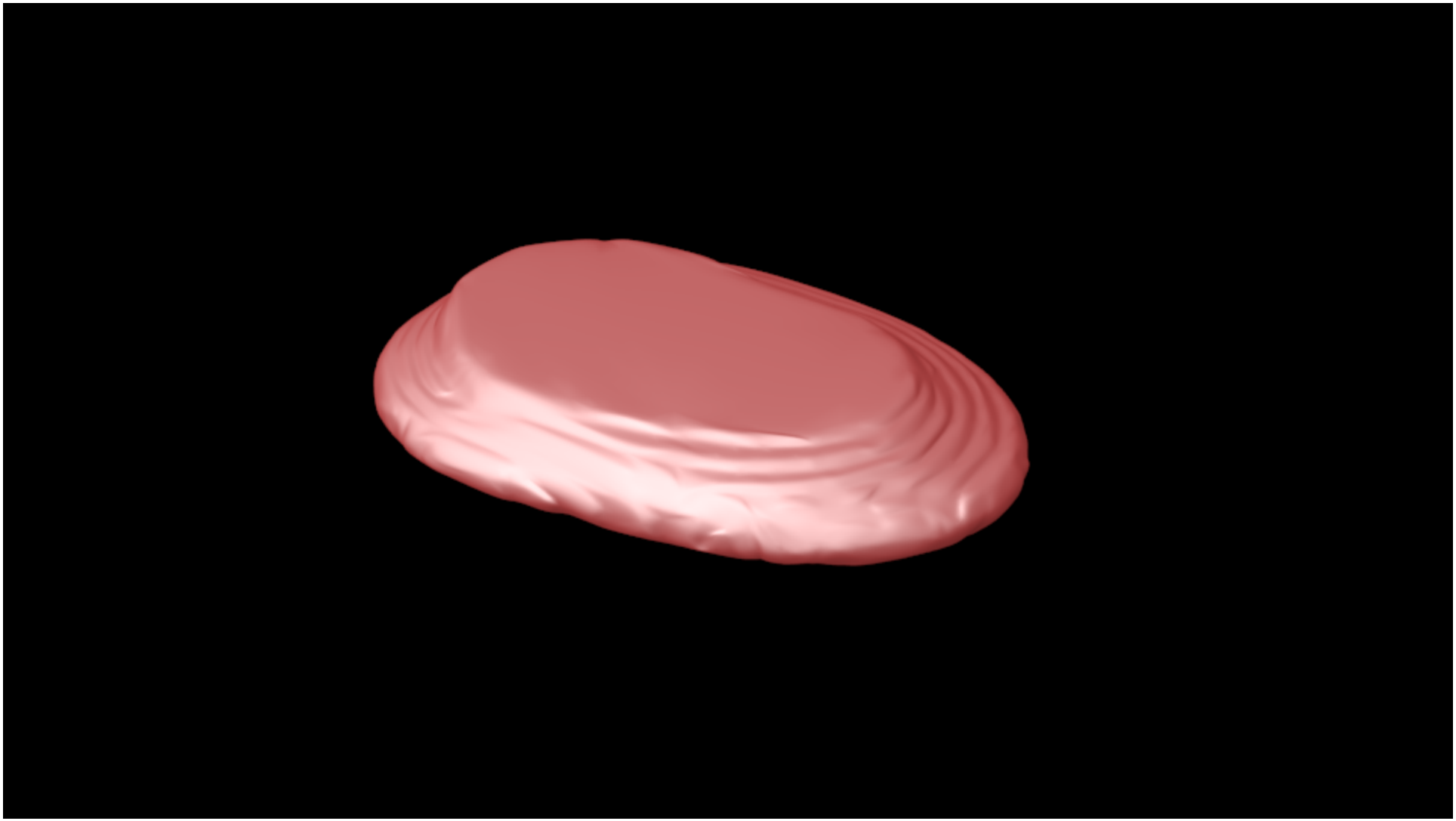




33 parameters (32 spline coefficients + height)

Statistical Models

- Determine the 33 parameters for many cells
- Learn appropriate statistical distribution (e.g., multivariate Gaussian)
- Sample from this distribution to synthesize new nuclei

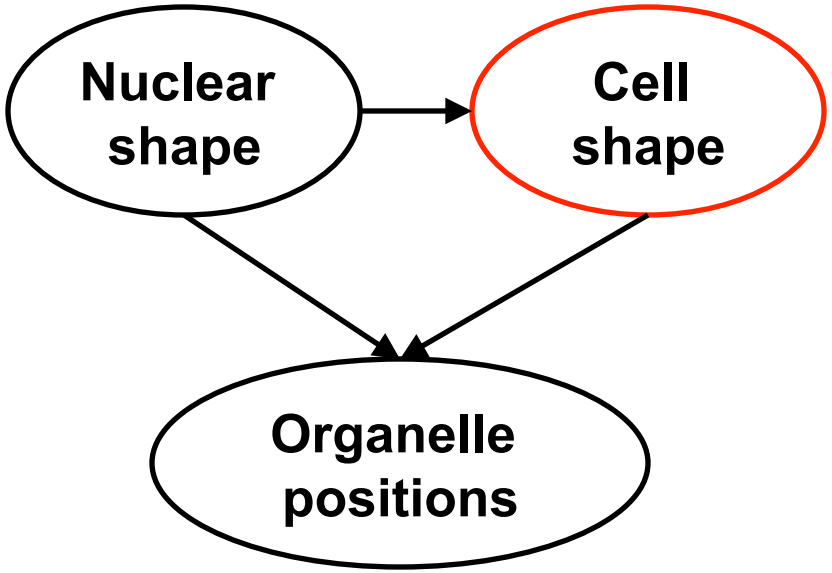


CELL SHAPE

Conditional models

- To ensure that the proper relationship exists between a synthetic nucleus and a synthetic cell surface, the models must be *conditional*
- Model for cell shape captures how cell shape *depends on* nuclear shape

MODEL:
CELL SHAPE

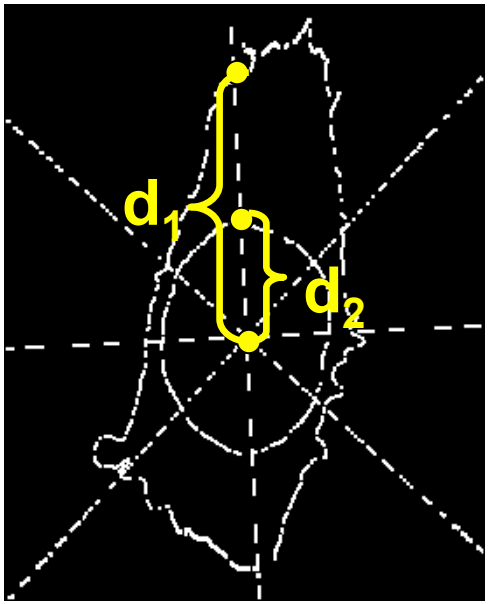


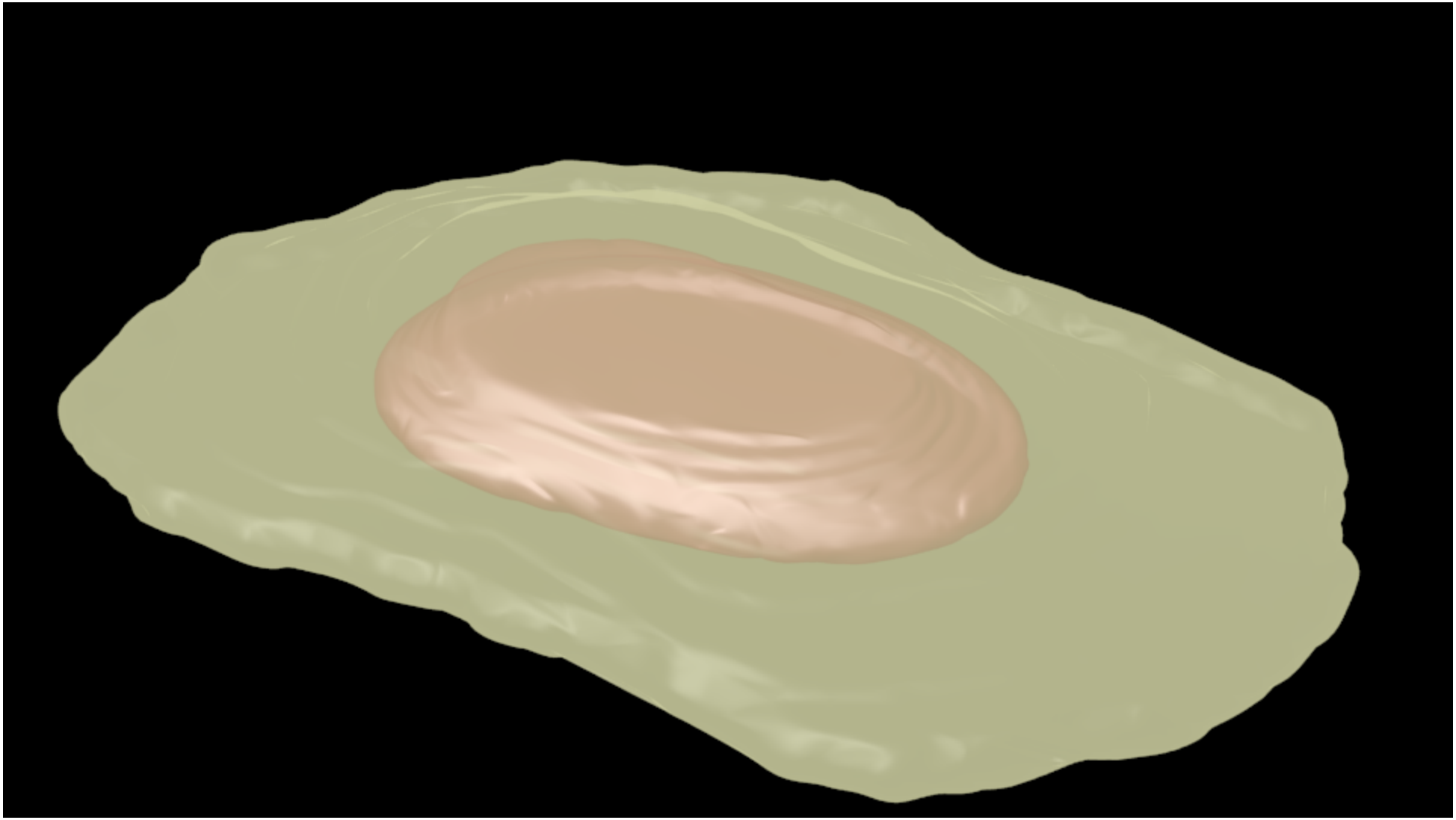
Cell shape: Ratio model

- Conditioned on nuclear shape: $r = d_1 / d_2$
  - Sample evenly around a circle to represent the shape by radius ratios (360 values)
  - Parameterization

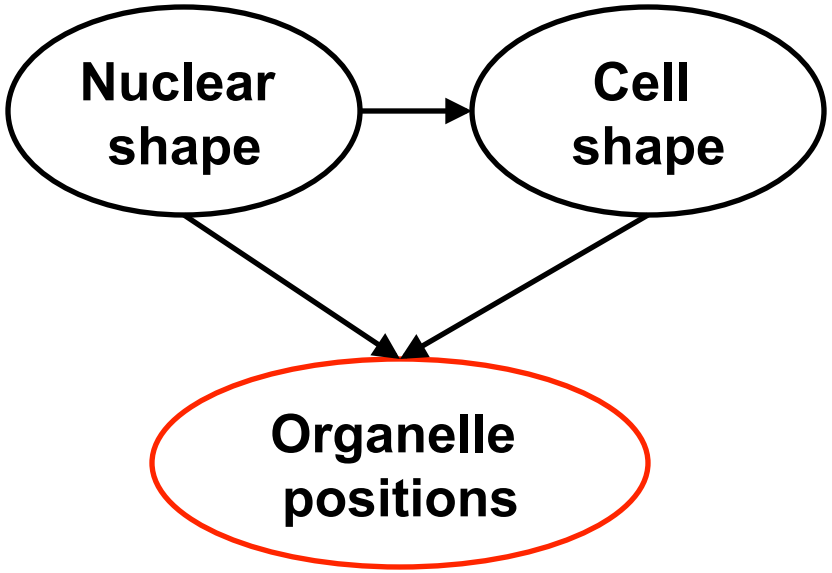
$$\mathbf{r} \approx \bar{\mathbf{r}} + \sum_{i=1}^{k} b_i \boldsymbol{\varphi}_i$$

    - Keep
    - 10 principal components for 2D
    - 25 principal components for 3D



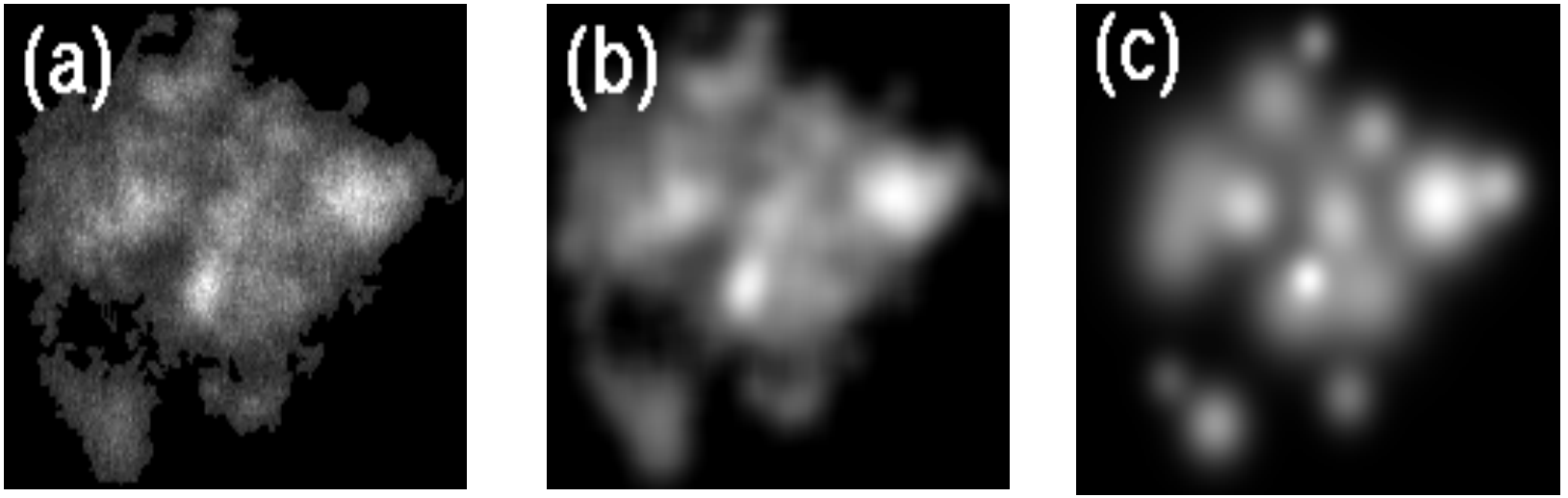


MODEL: VESICULAR ORGANELLES

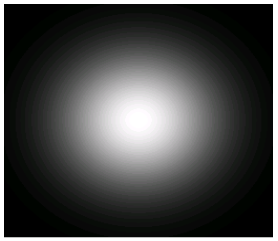


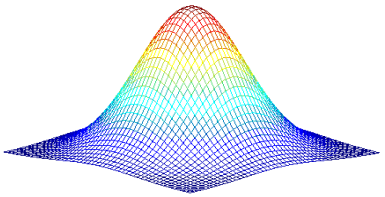
Modeling Vesicular Organelles

Original | Filtered | Fitted Gaussians



Gaussian object modelparameters





Number of objects in a cell N
Object size σ_x, σ_y, σ_z
Single object fluorescence F

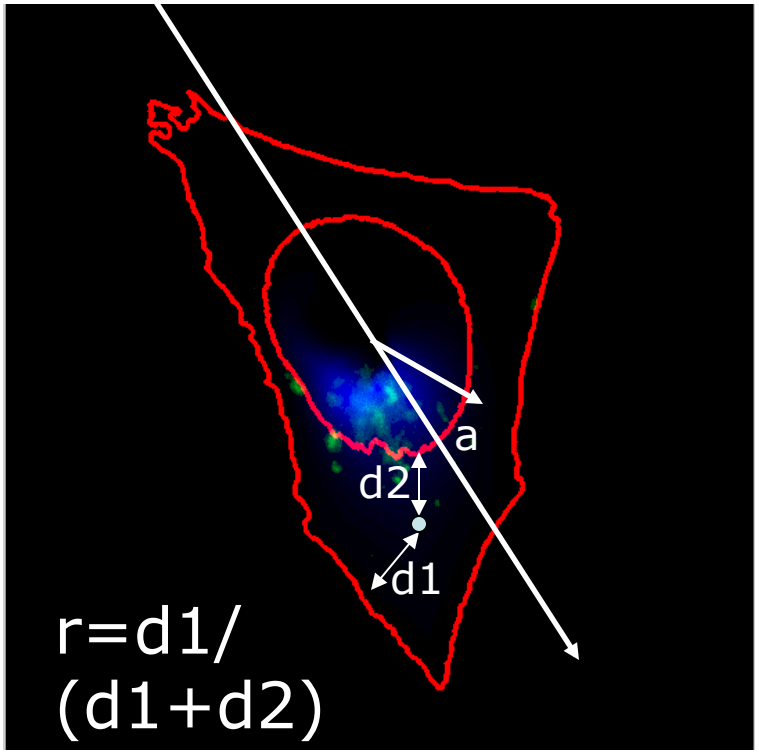
Protein Object Model: size, shape, intensity

- Size of each ellipsoid represented as distribution of length of major axis and conditional distributions of lengths of other axes relative to major axis
- Exponential distribution for intensity of each object

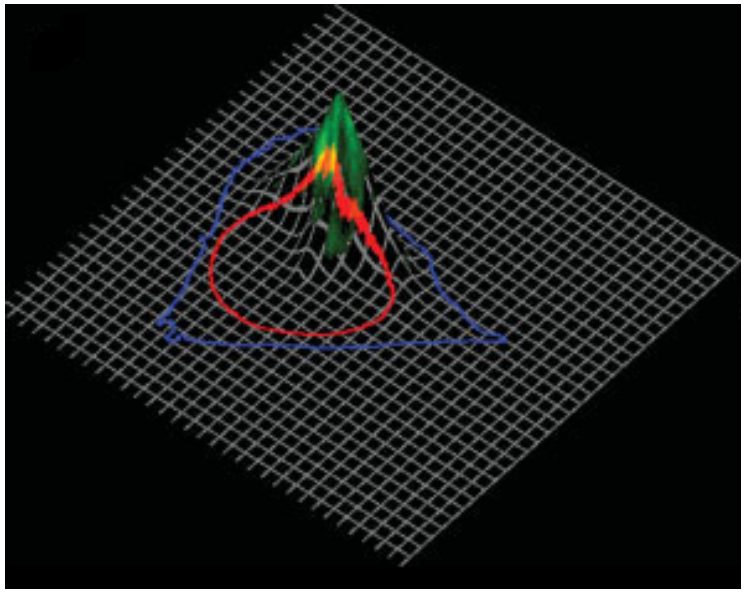
Protein Object Model: Position

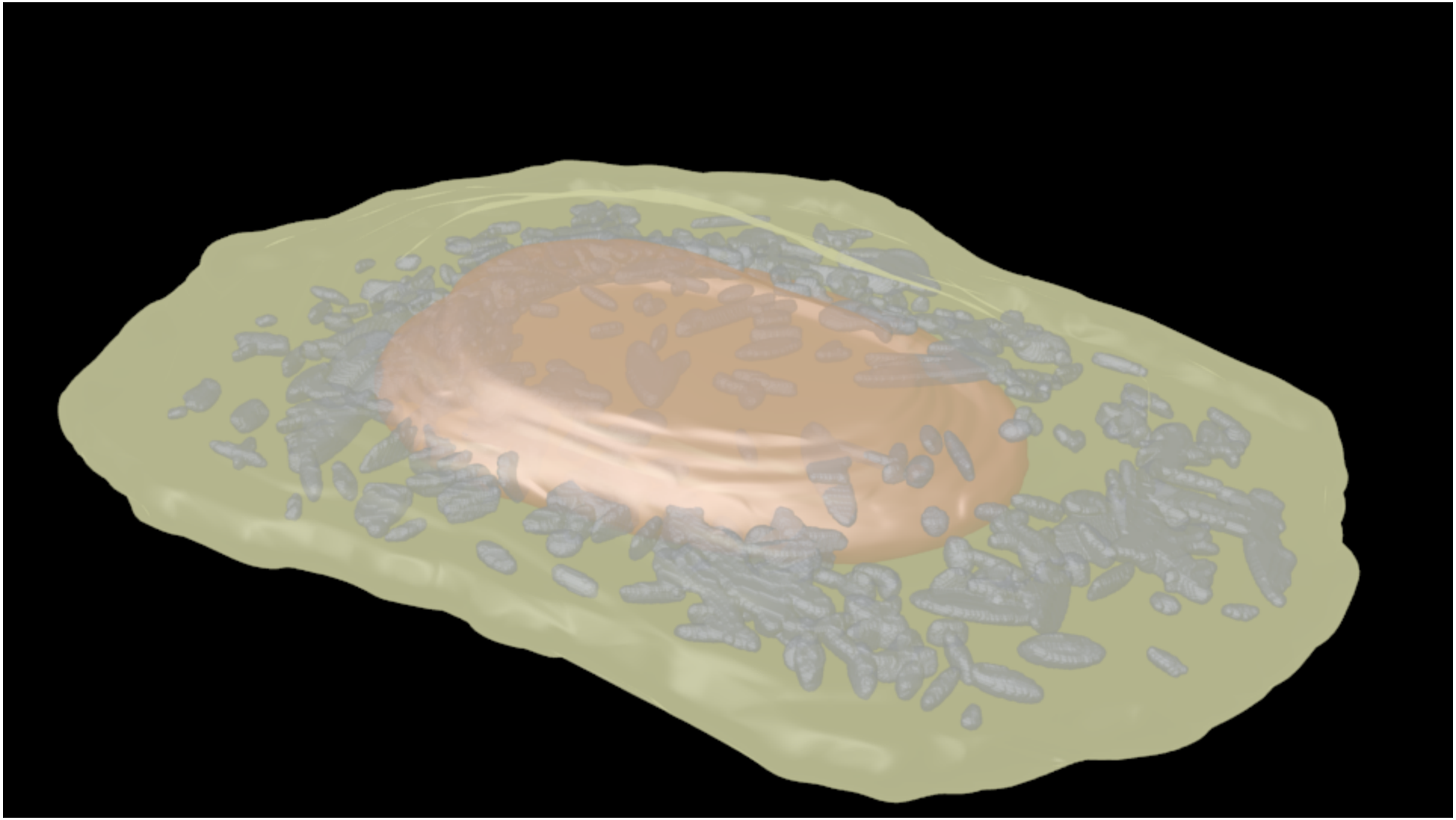
$$\text{Potential}: \quad P(r,a) \propto \frac{\tau(r,a)}{1+\tau(r,a)}$$

$$\text{where } \tau(r,a) = \exp\left(\underline{\beta_0} + \underline{\beta_1} r + \underline{\beta_2} r^2 + \underline{\beta_3} \sin a + \underline{\beta_4} \cos a\right)$$



Normalized potential map:
A 2-d density function





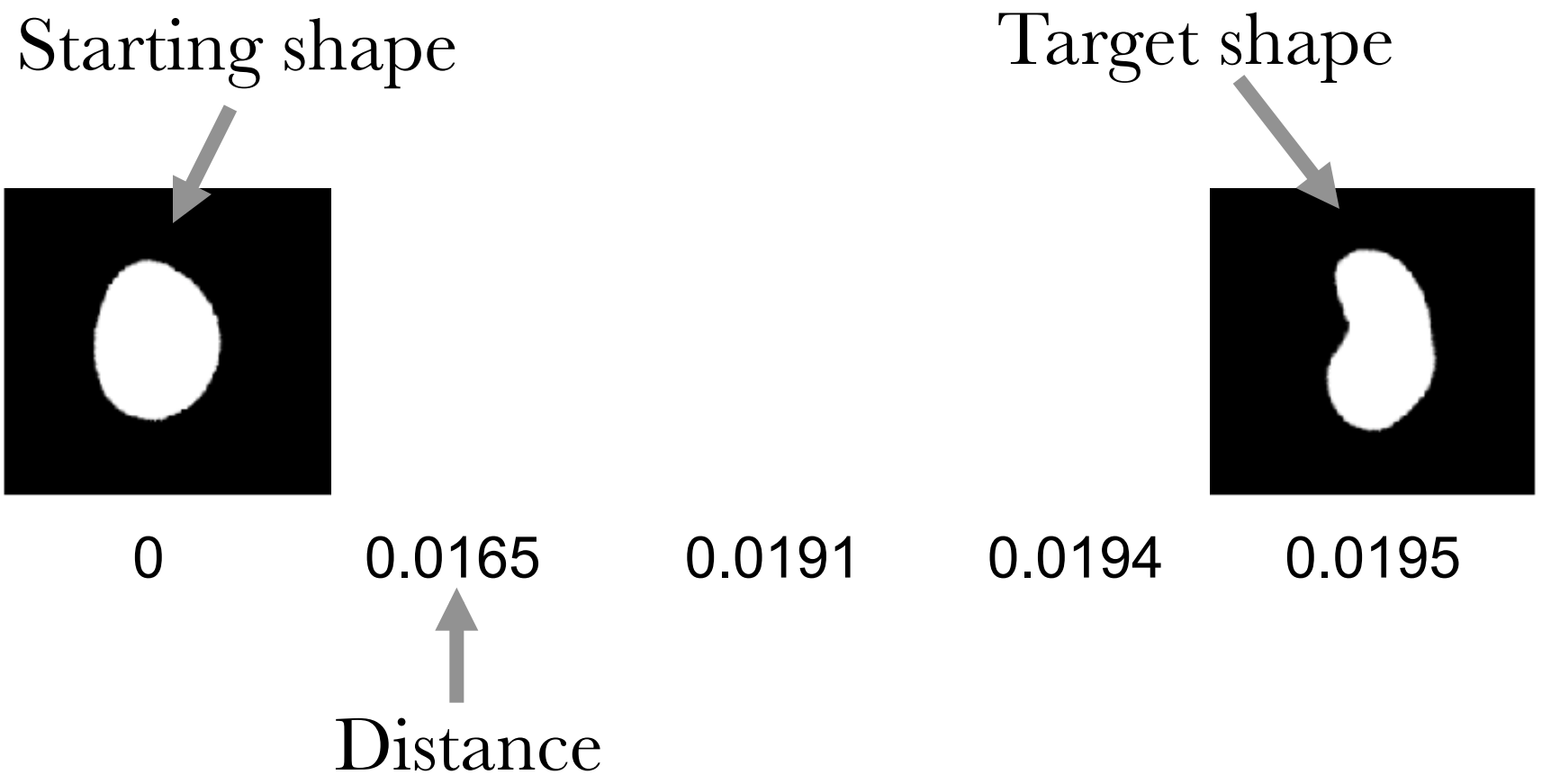
NON-PARAMETRIC MODELS

Diffeomorphic analysis of shape

- Sometimes cell or nuclear shapes are irregular
- Can use distance between shapes to characterize shape instead of parameters of model (based on work by Michael Miller and colleagues)

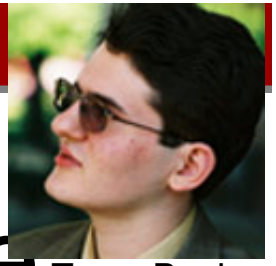
Gustavo Rohde

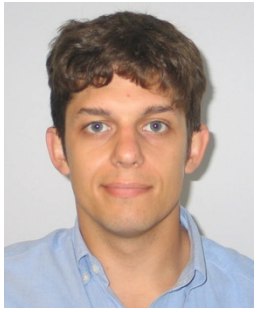
Morphing one shape into another



Constructing a shape space

- Once we have distances of every cell to every other cell, we can 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)


Taraz Buck


Greg Johnson

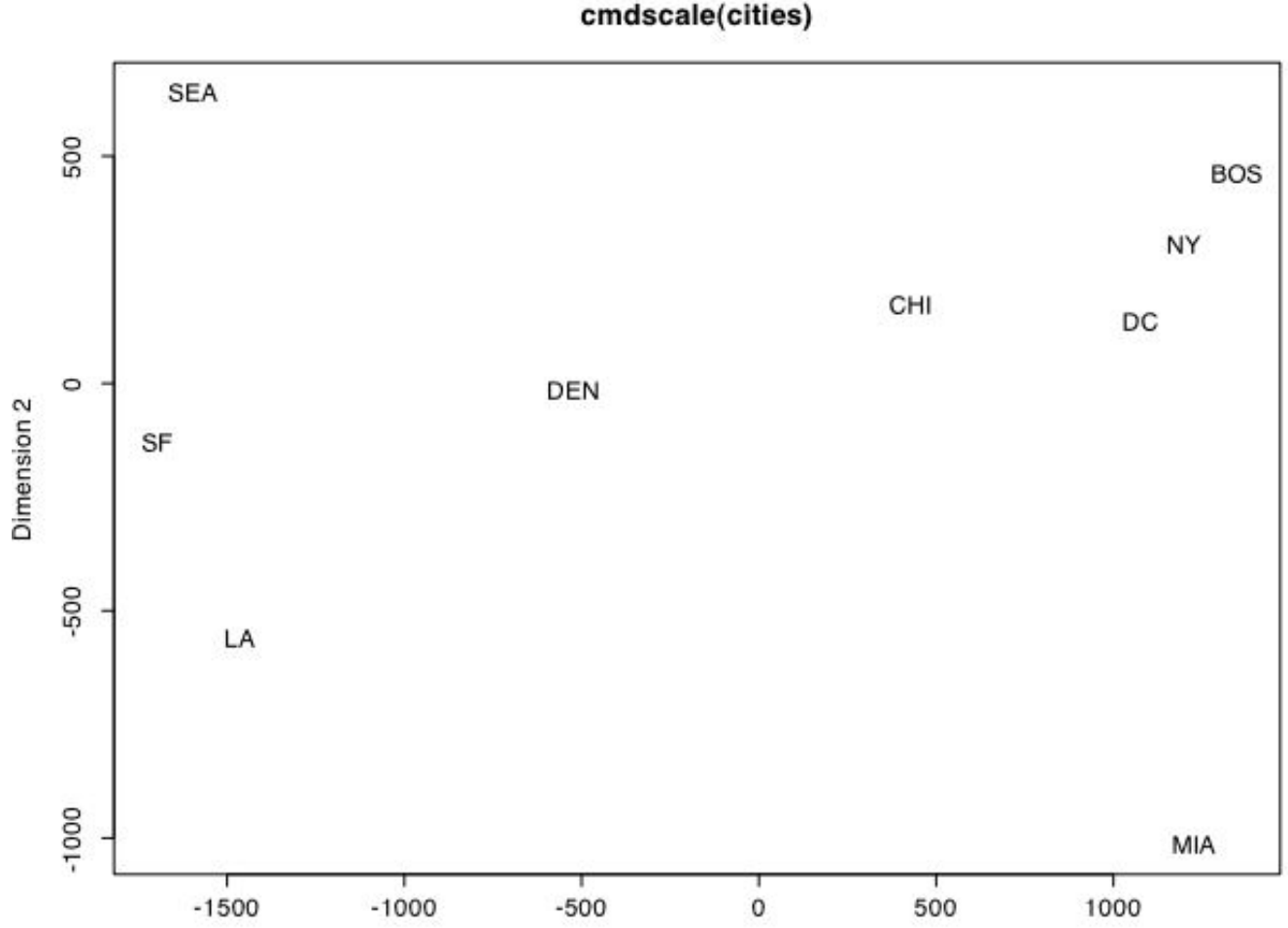

Devin Sullivan

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

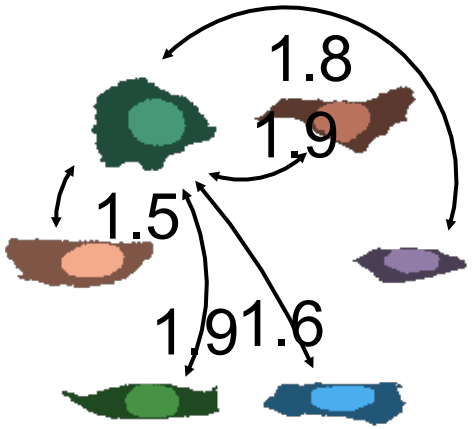
… to coordinates



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

Constructing a shape space

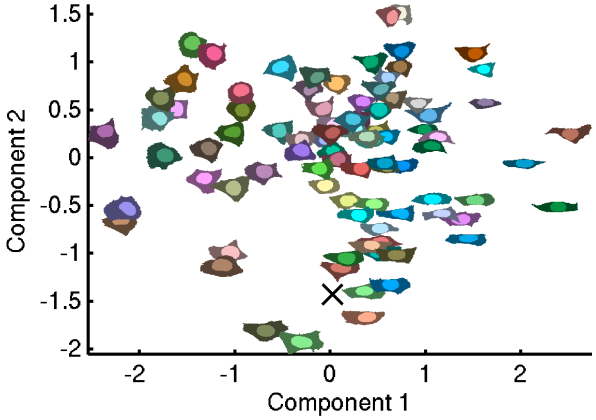
For a set of shapes, compute distances between pairs…



… to produce a distance matrix…

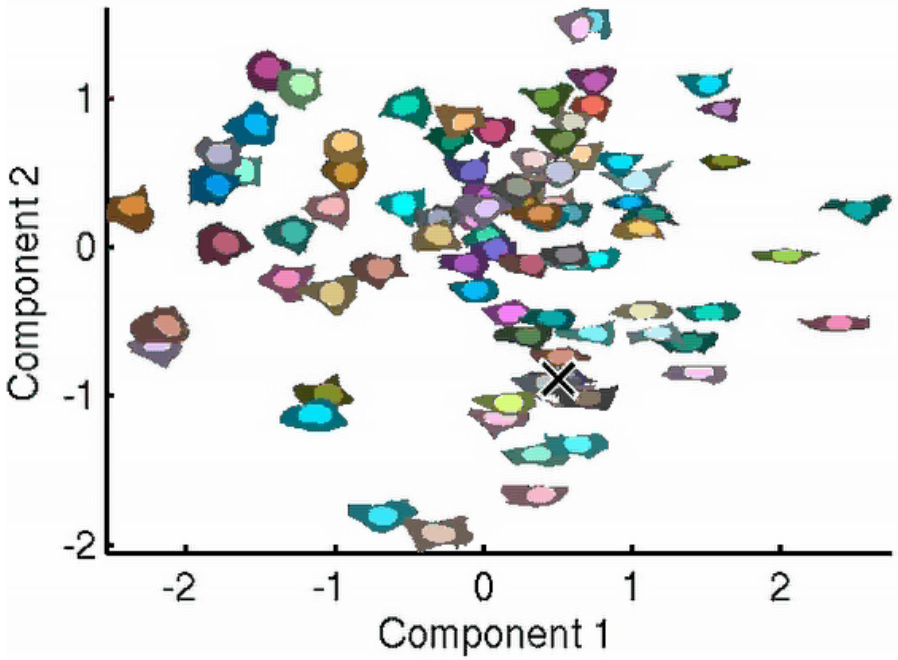
0	1.9	1.8	1.6	1.9	1.5
1.9	0				
1.8		0		…	
1.6			0		
1.9		…		0	
1.5					0

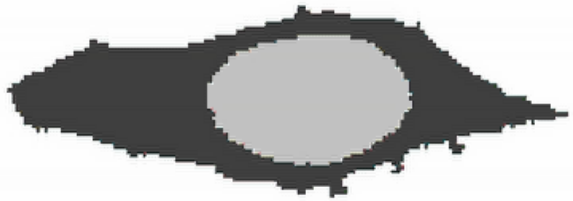
… and find coordinates for each shape that reproduce those distances



Shape space

Synthesized shape for current position in shape space





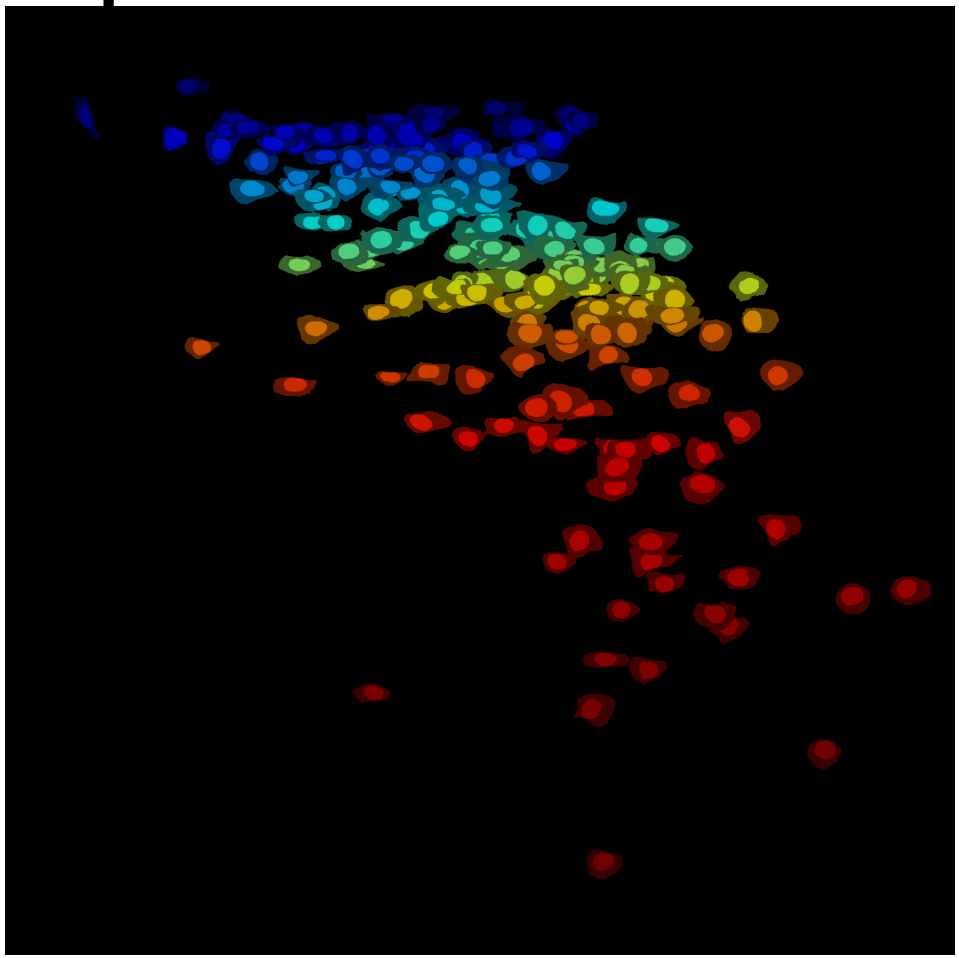
× = current shape's position

Training data shapes are shown at their respective locations in random

Static to dynamic

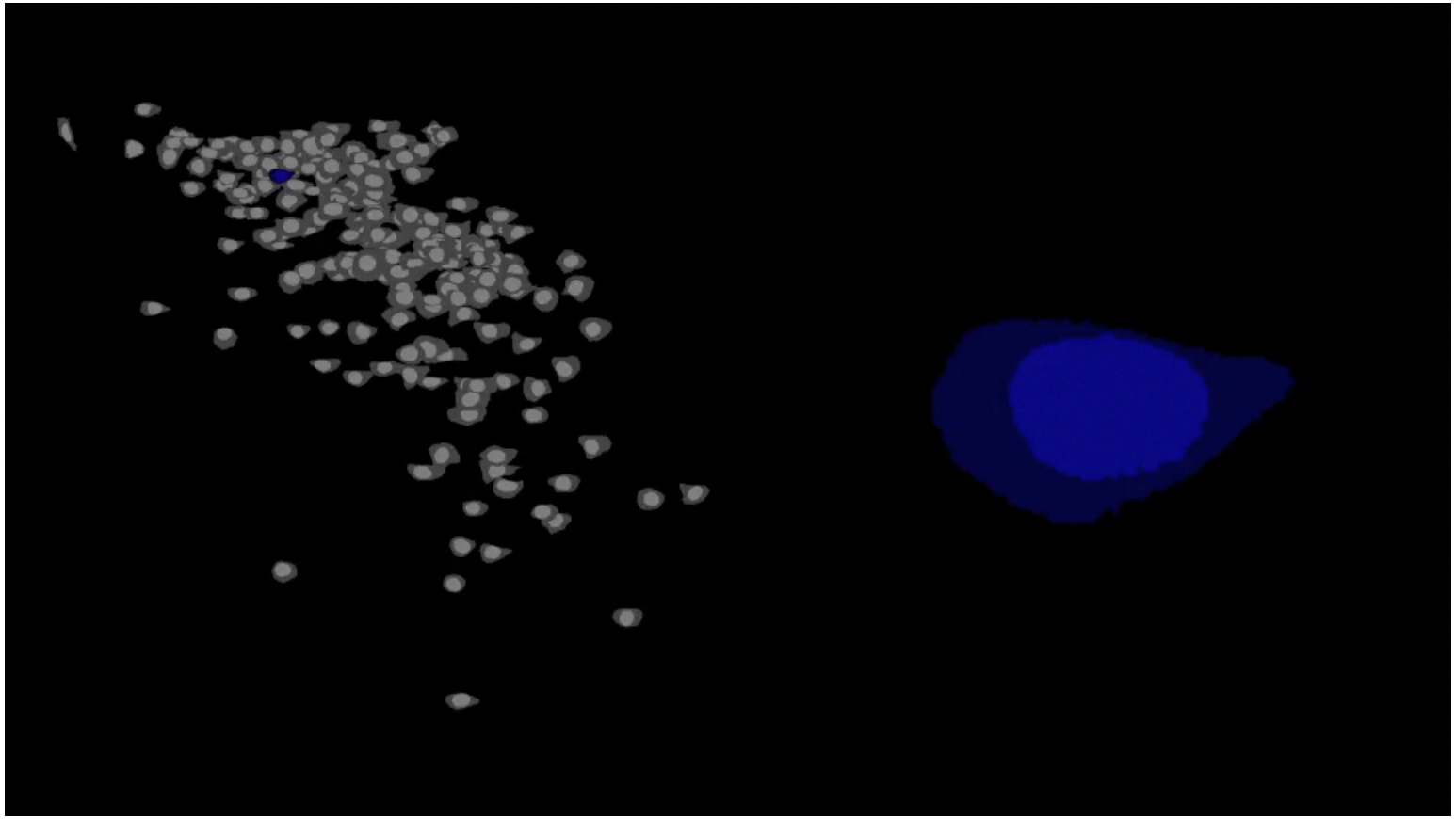
- Methods such as these allow a collection of static images of different cells to be converted into a model of the dynamics of cells
- Need to have a way to learn the rules governing “trajectories” in the model

Construct space to relate shape to DNA content



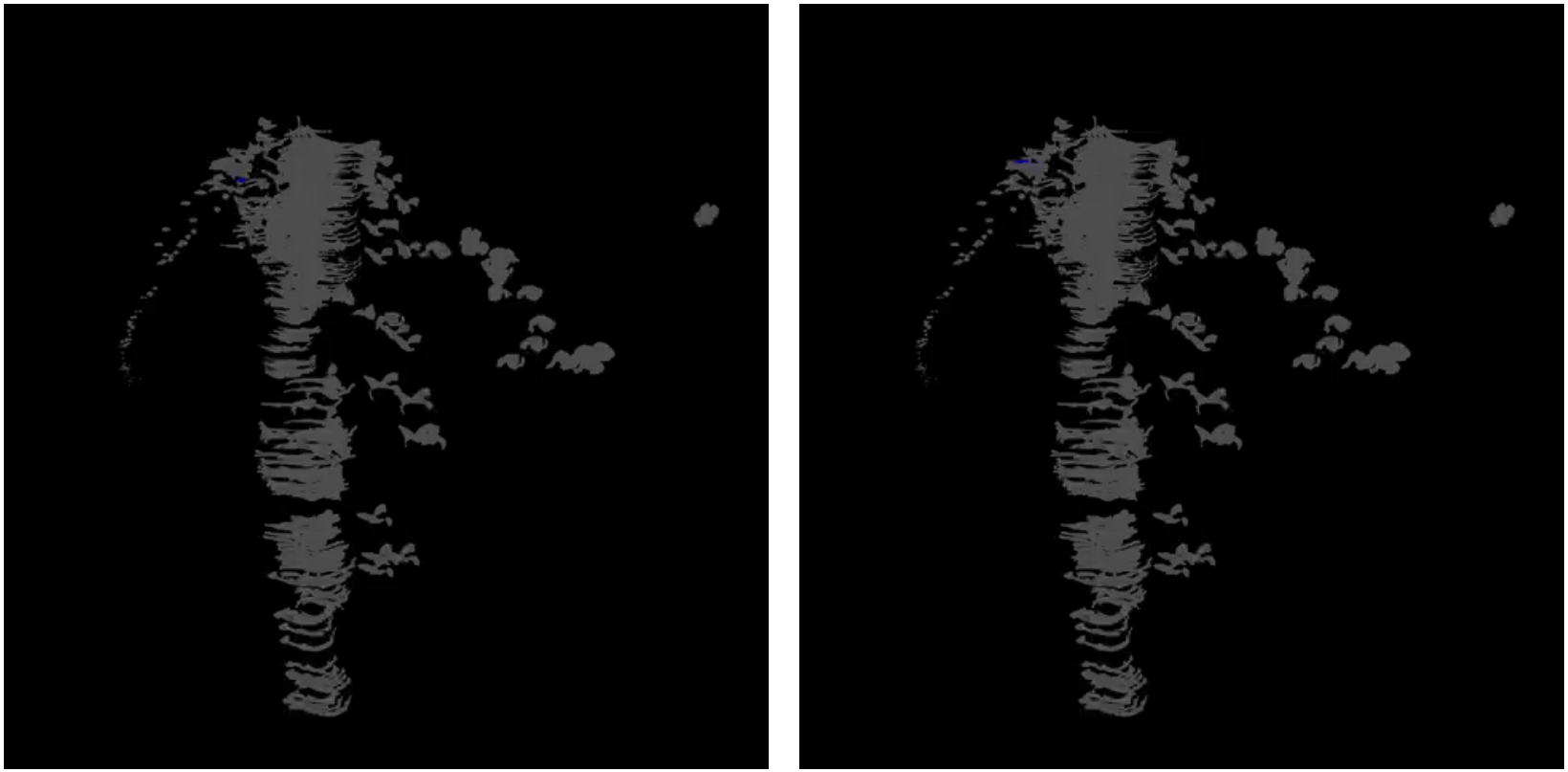
3D HeLa

Directed walk in DNA-shape space



3D HeLa

Actual trajectories in shape space



2D C2C12

MODELING SUBCELLULAR DISTRIBUTION CHANGES DURING CELL SIGNALING

Modeling non-organellar patterns and dynamics

- What are the spatiotemporal patterns of proteins involved in signaling during antigen presentation?

Taraz Buck

Baek Hwon Cho

Christoph Wülfing

Singleton et al. 2009

Analysis of T cell synapse patterns

- DIC and GFP fusion protein intensity

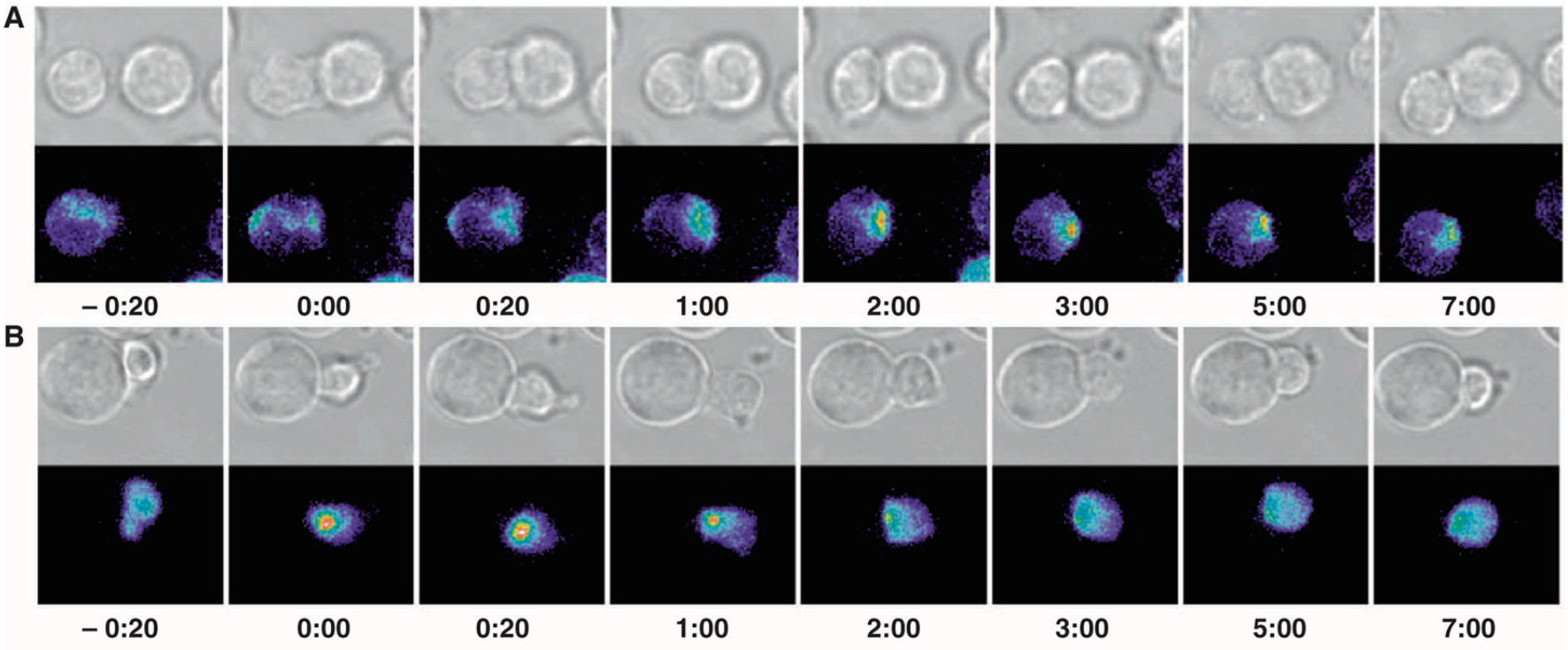
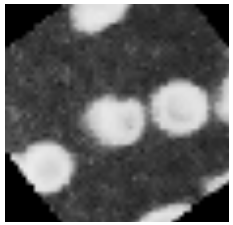
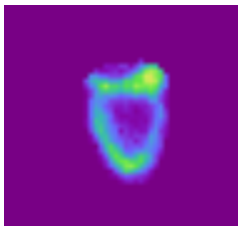
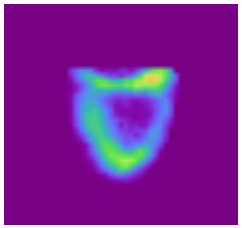
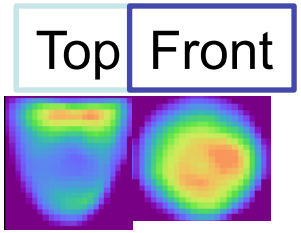


Image processing pipeline









Images of fields

Images of individual cells

Individual cells with uniform shape

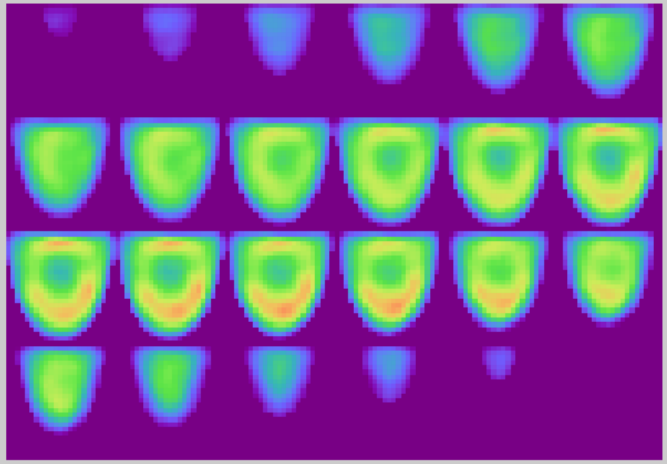
Protein distribution models

Segmentation

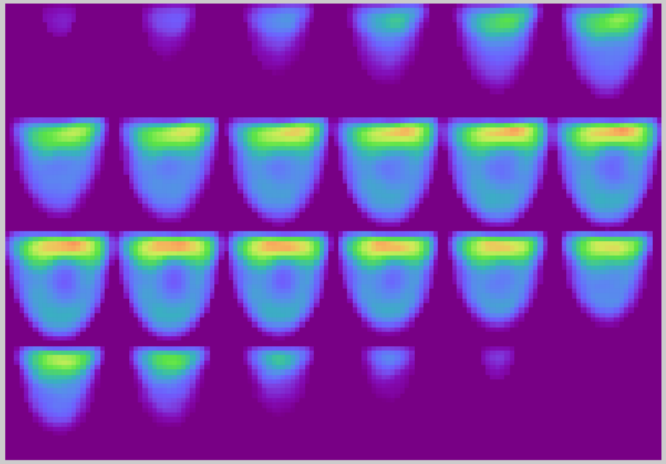
Rigid alignment and morphing to half-ellipsoid template

Parametric model construction and comparison

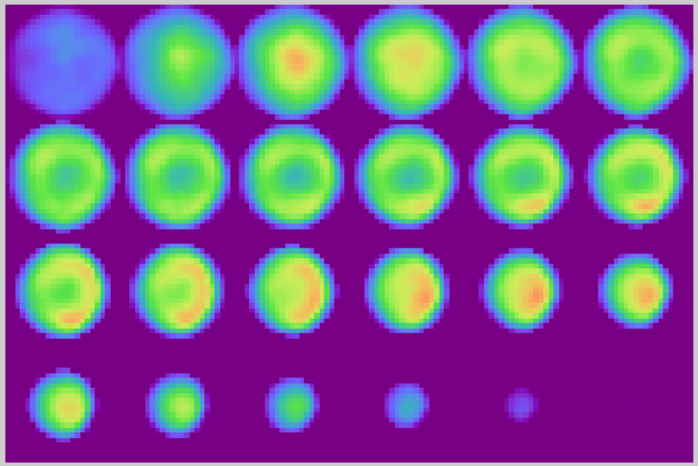
ARP3
relative time -2, 107 cells (bottom to top slices)



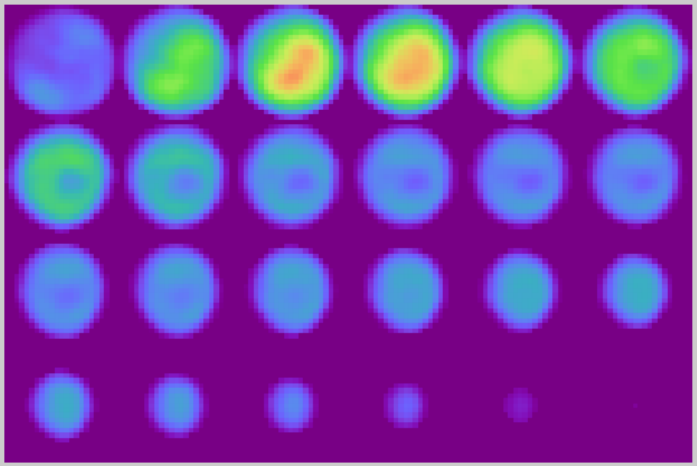
ARP3
relative time 0, 109 cells (bottom to top slices)



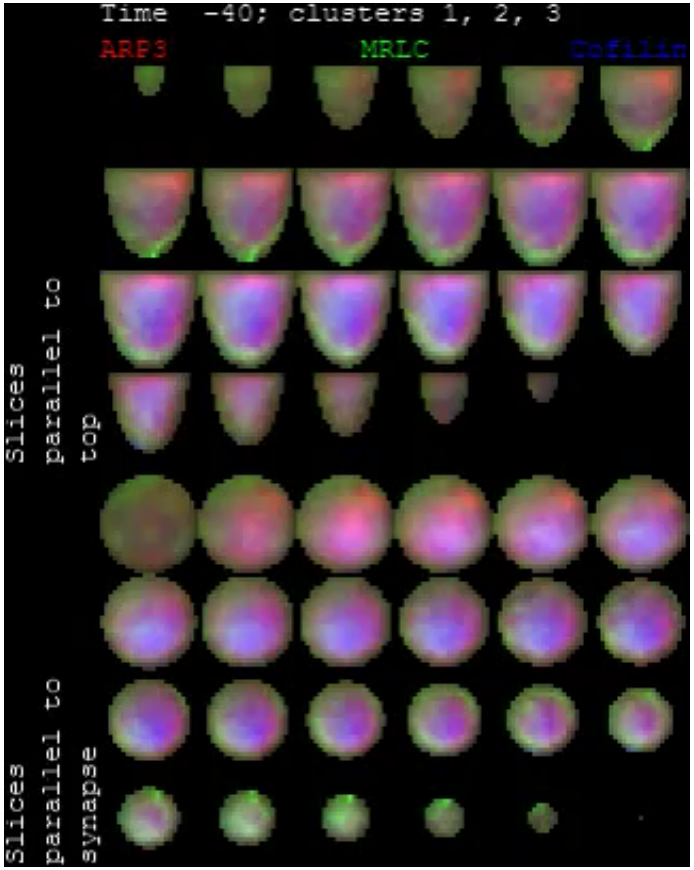
ARP3
relative time -2, 107 cells (synapse to end slices)



ARP3
relative time 0, 109 cells (synapse to end slices)

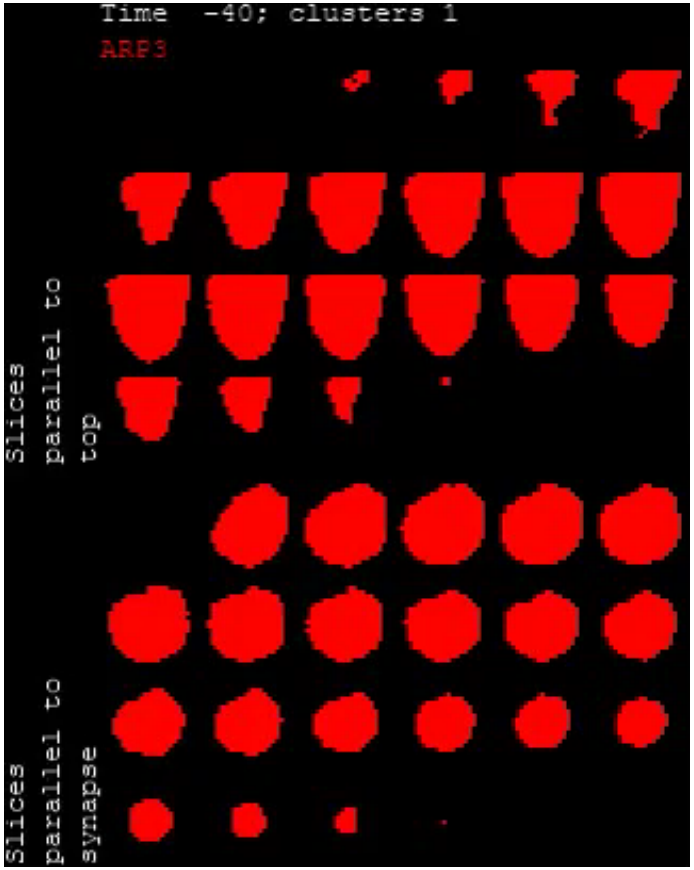


Redistribution Dynamics



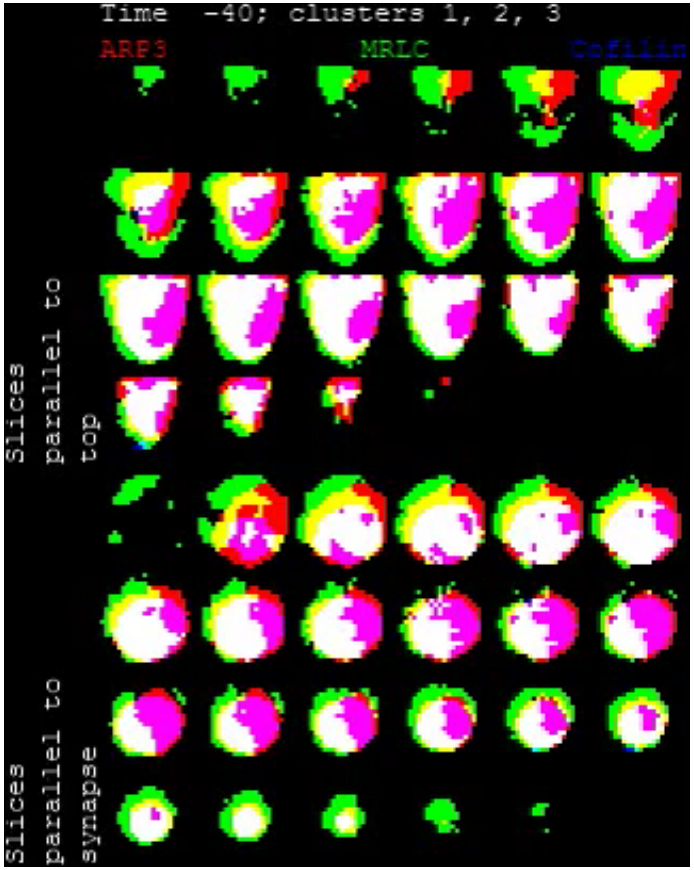
Note: intensity on log scale

Redistribution Dynamics



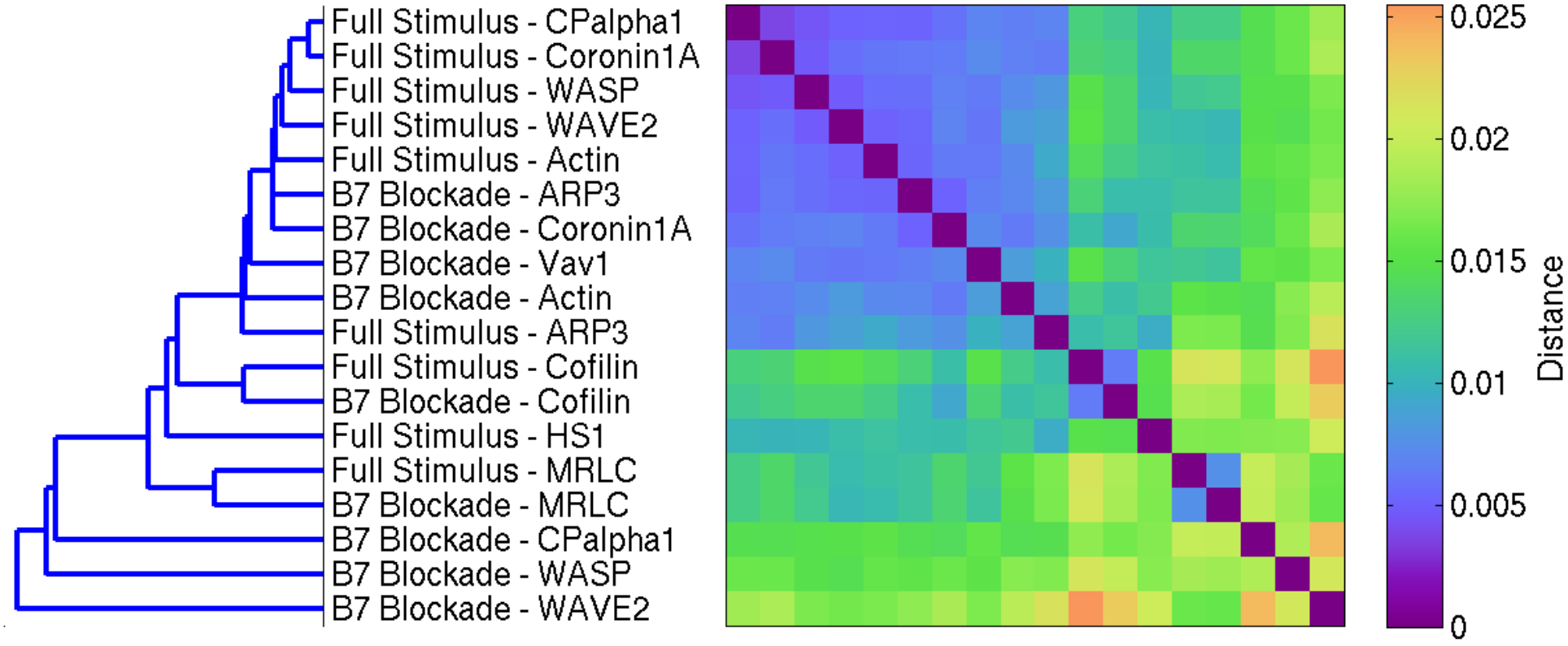
Thresholded intensity

Redistribution Dynamics



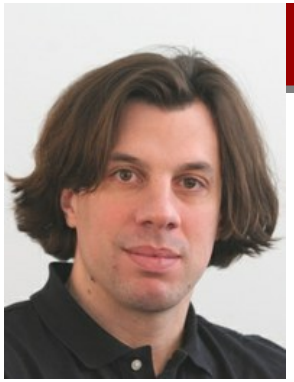
Thresholded intensity

Clustering of Spatiotemporal Patterns



Distance matrix for each average 4D map to each other

MODELING CHANGES DURING CELL DIFFERENTIATION

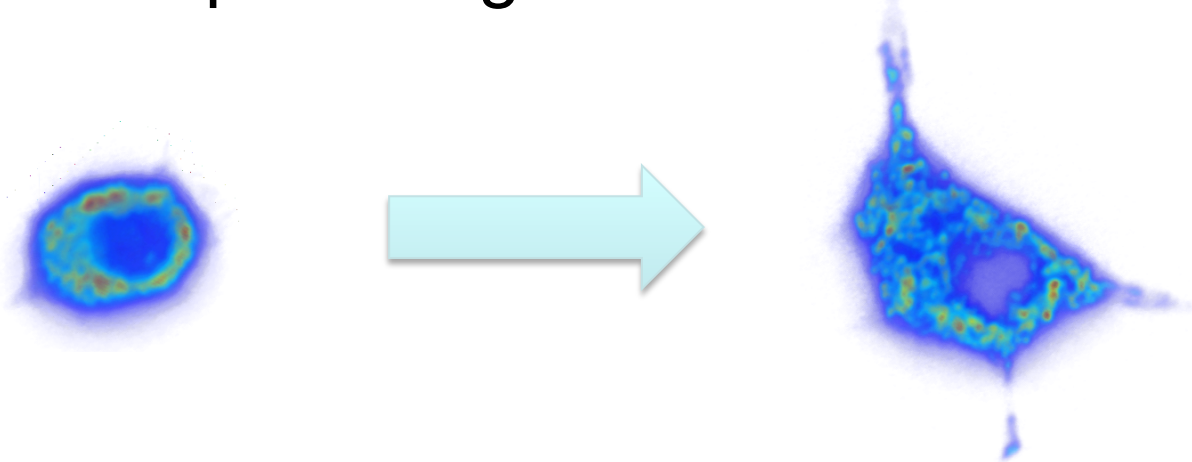


Hauke Busch

Model system

Melanie Boerries

- PC12 - Rat neuroendocrine tumor cells
  - Normal – very sensitive, divide
  - NGF – Stop dividing and start differentiating



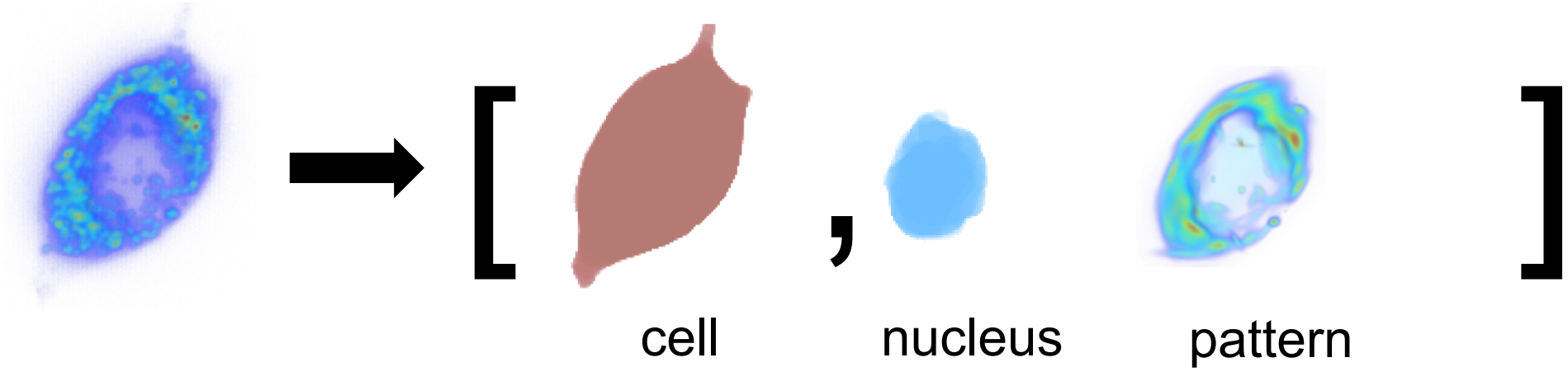
What is the phenotypic ordering of morphological changes that result in a differentiated cell?

Imaging

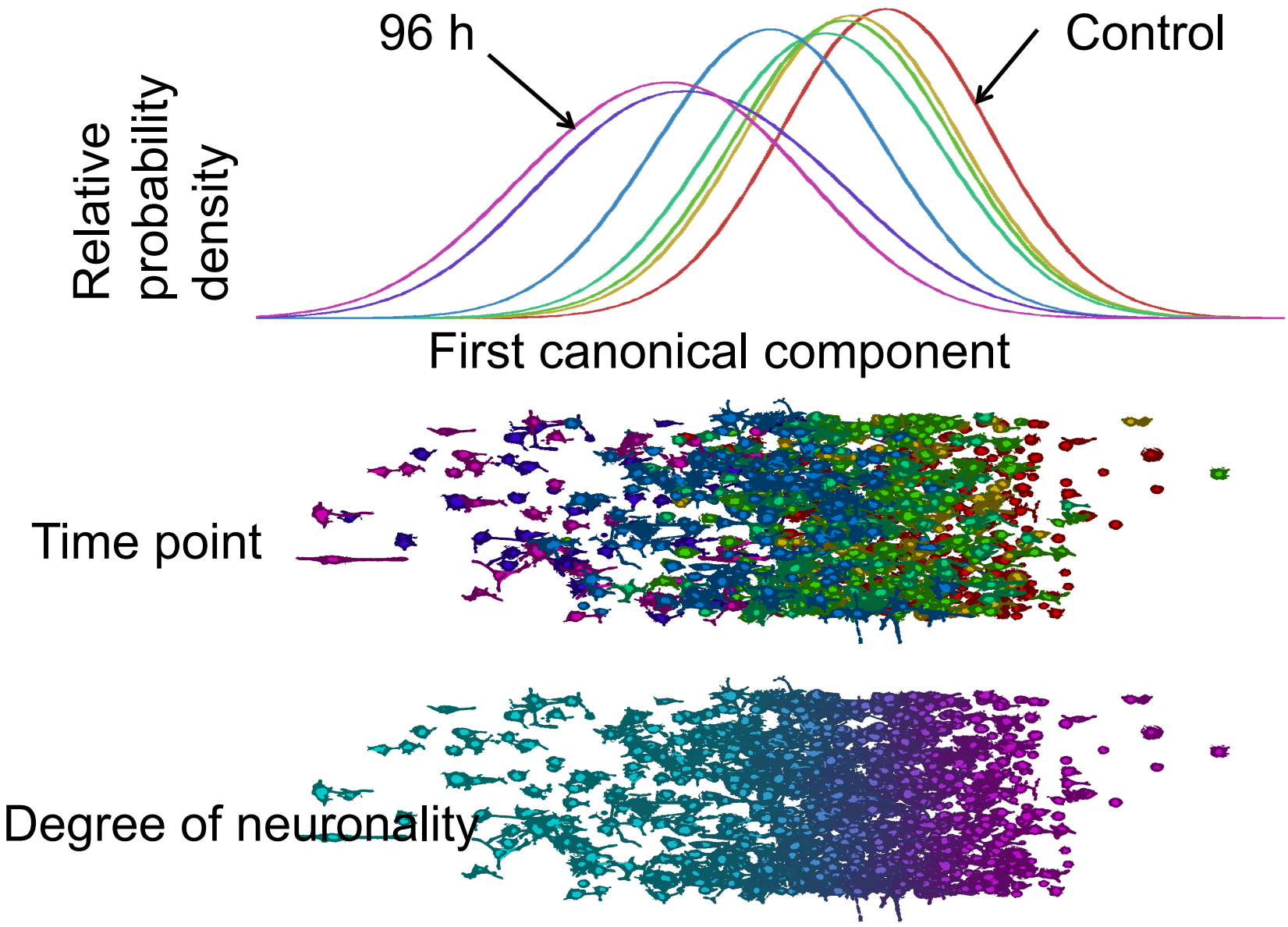
- Cells are cultured and exposed to NGF
- Cells are imaged at various times after adding NGF
  - 0, 12, 24, 36, 48, 72, 96 hours
  - Prior to imaging, stain with MitoTracker

Greg Johnson

Convert images to models



Learn one-dimensional embedding of parameter space (“extent of differentiation”)

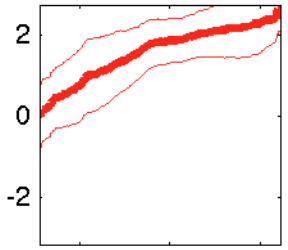

96 h
Control
Relative probability density
First canonical component
Time point
Degree of neuronality

Synthetic movie of cell and nuclear shape changes

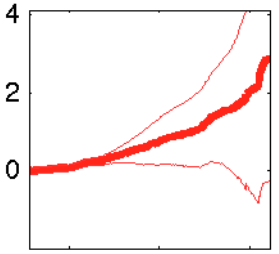
Identifying Parameter Changes

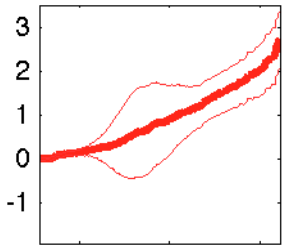
Cell Shape Parameters

Eccentricity



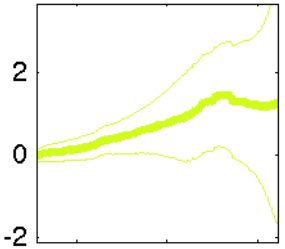
Skeleton Features



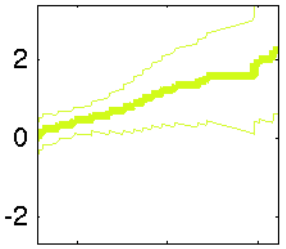


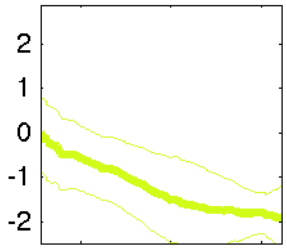
Nuclear Shape Parameters

Size



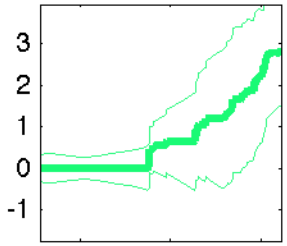
Skeleton Features



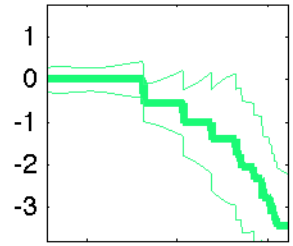


Mitochondria Position Parameters

Outer



Inner



Conclusions

- Tools beginning to be available to build image-derived generative models
  - Learn the underlying cell “model” from which individual cell images are drawn
- Useful for
  - Learning perturbation models – better than features!
  - Building spatially realistic cell simulations
  - Visualizing results from many noisy images

CellOrganizer

Project Leaders

Robert F. Murphy

Gustavo Rohde

Major Collaborators

Klaus Palme

Hauke Busch

Christoph Wülfing

Melanie Boerries

Jörn Dengjel

Ivo Sbalzarini

Current Team Members

Devin Sullivan

Gregory R. Johnson

Taraz Buck

Ivan Cao-Berg

Past Contributors

Ting Zhao

Tao Peng

Wei Wang

Aabid Sharif

Joshua Kangas

Jianwei Zhang

Alexander Dovzhenko

Rüdiger Trojok

Jieyue Li

Baek Hwan Cho

MMBioS

An NIH Biomedical Technology Research Center