Joint Nonparametric Alignment for Spatial Gene Expression Pattern Analysis of Drosophila Imaginal Discs

Parvez Ahammad
EECS Dept., UC-Berkeley
Joint work with Cyrus Harmon,
Ann Hammonds, Shankar Sastry and Gerald Rubin

Some Basics Terms

- Drosophila Melanogaster
  - Fruitfly (very popular in gene expression studies)
- Nonparametric methods
  - Machine learning methods that use a set of examples as the learning primitives
- Spatial Gene Expression Patterns
  - Usually the micro array gene expression data is a time series with information about lots of genes but no spatial resolution
  - Patterns of spatial expression are considered to show the effect of a gene on a cell's evolution and behavior
- Imaginal Discs
  - Primordial tissues that will go on to become part of the exoskeletons of the Fruitfly
  - Separated in embryogenesis

So how do imaginal discs look like?

Motivation

- High throughput systems require automatic alignment and simple processing pipeline
  - Manual curation is expensive
- Align the imaginal disc shapes to facilitate:
  - Meaningful quantitative analysis of spatial gene expression patterns
  - Learning the underlying model of the imaginal disc shapes for use with model based techniques

Given

- A set of imaginal disc images of a specific class, with various stain patterns
  - Known class labels
  - Underlying shape model is unknown
  - Transformation parameters unknown

Difficulties

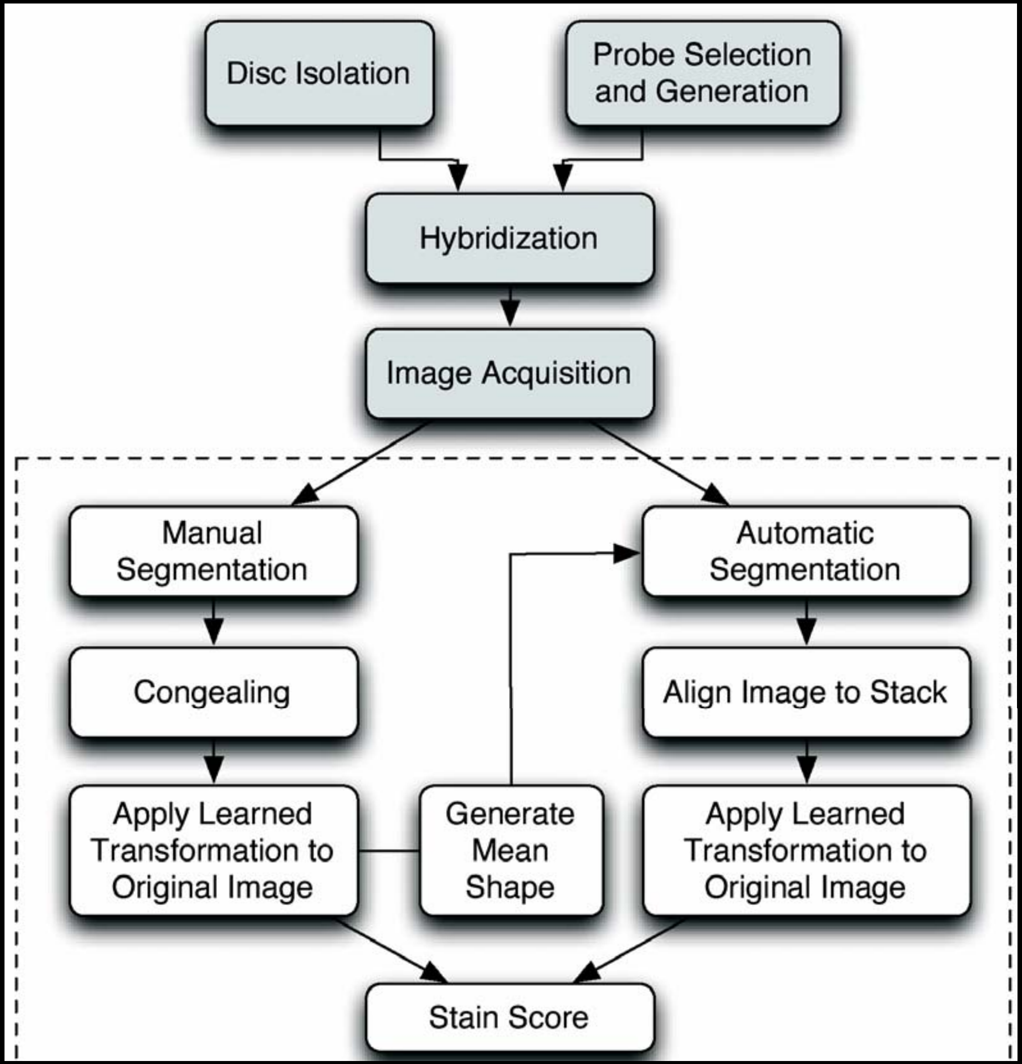
- A clean shape model of the tissue class is generally not available
- Manually selecting features on images is expensive and time consuming
- Images have a lot of clutter and noise

Revisit: Can you see the difficulties?

Prior Work

- Analysis of precise spatial gene expression patterns in Drosophila embryos through in situ hybridization[Berman et.al 2002, Tomancak et.al 2002]
  - Curation is manual
  - Requires annotation
- Automated embryo registration and stain classification [Peng and Myers (RECOMB 2004)] using Gaussian mixture models
  - Registration is very simplistic (major/minor axes)
- Shape learning and alignment is a well-studied problem
  - Miller et.al (CVPR 2000)
  - Charpiat et.al (ICIP2003)
  - Frey and Jojic (CVPR1999)
  - Tsai et.al(TMI 2003)

Data Flow in the Proposed Approach

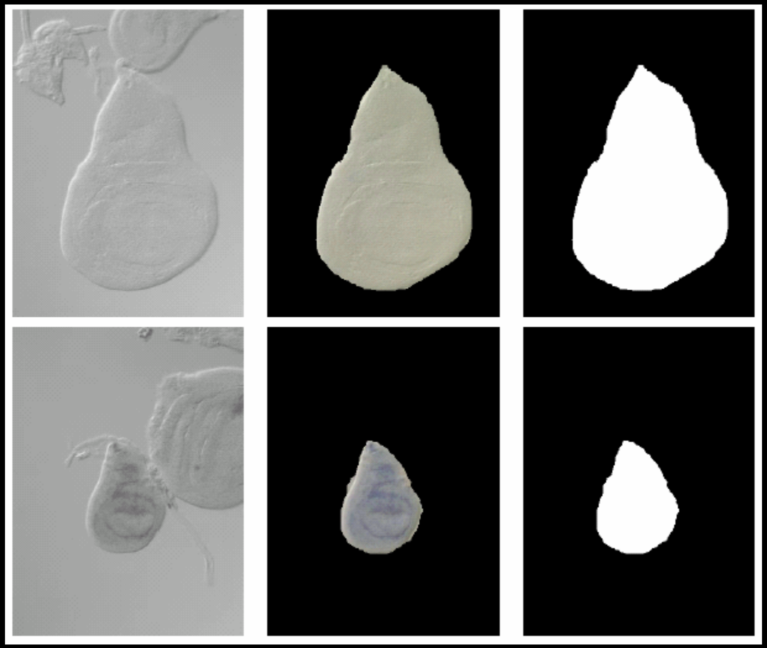


Highlights of the Approach

- Proposed approach makes no assumptions on the underlying anatomy or shape
  - Generalizable to other gene expression studies
- Removes the need for experts to mark points of interest
  - Amenable to large scale automation
  - Unsupervised
  - Registration performance improves asymptotically as number of samples (images) increases
- Learned transformations are semantically meaningful
- Augments model based registration techniques
  - Learned model can be used to bootstrap model based methods
- Extendable to 3D datasets and gray-scaled valued datasets

Segmentation

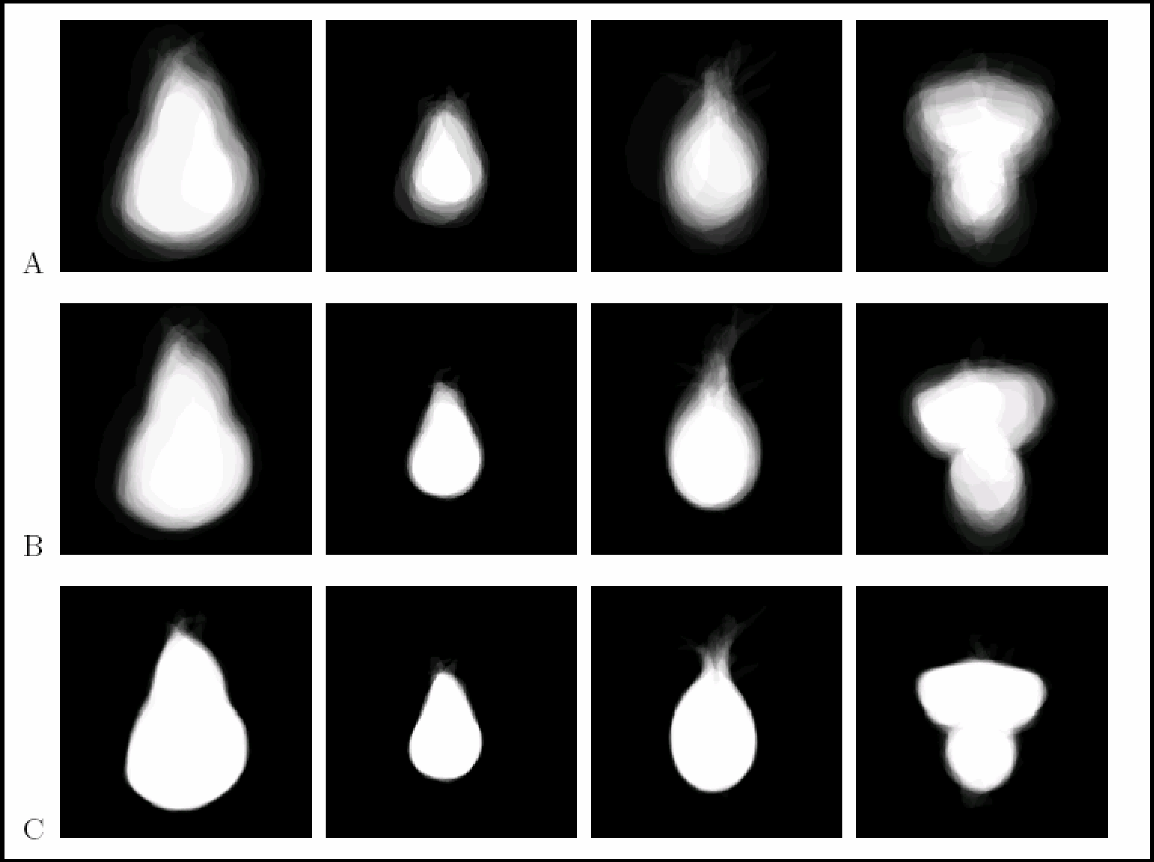
- Manually segment a few sample images
  - Learn shape model from these images
- Combine variance segmented images with learned shape model to automate segmentation process



Joint Nonparametric Alignment and Shape Learning

- Learn MAP estimate of the underlying shape model
- Entropy minimization algorithm: Congealing (ref: Miller et. al. – CVPR '00)
  - Coordinate descent method
  - Input: A set of binary shape masks of tissues of a given class
  - Output upon convergence: Aligned binary shapes + corresponding transformations
    - Over parameterization used: x-translation, y-translation, rotation, x-log-scale, y-log-scale, x-shear, y-shear

Mean Images from Shape Learning Stage

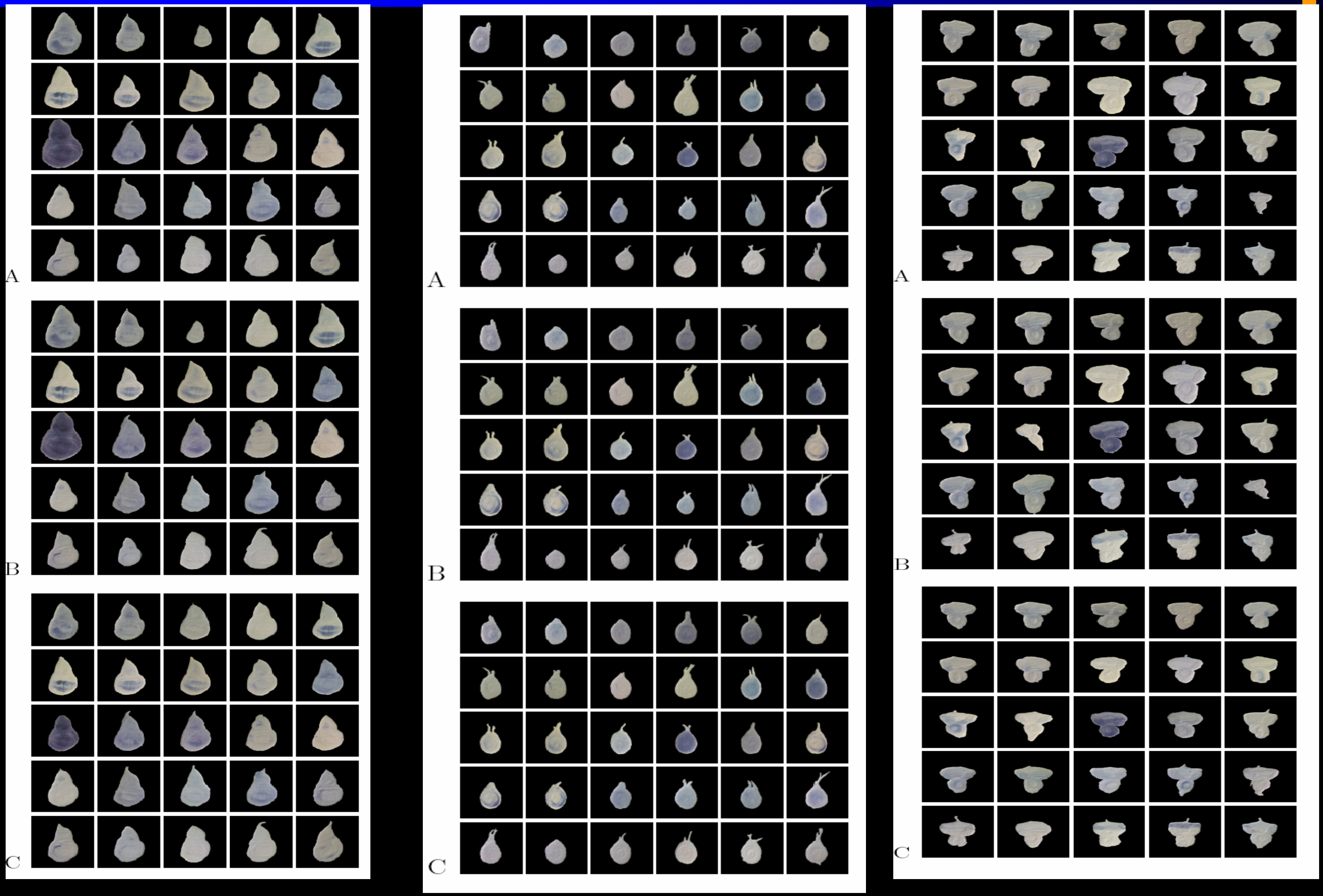


Wing, Haltere, Leg, Eye discs

A: Before Alignment; **B**: Alignment with 3 parameters;

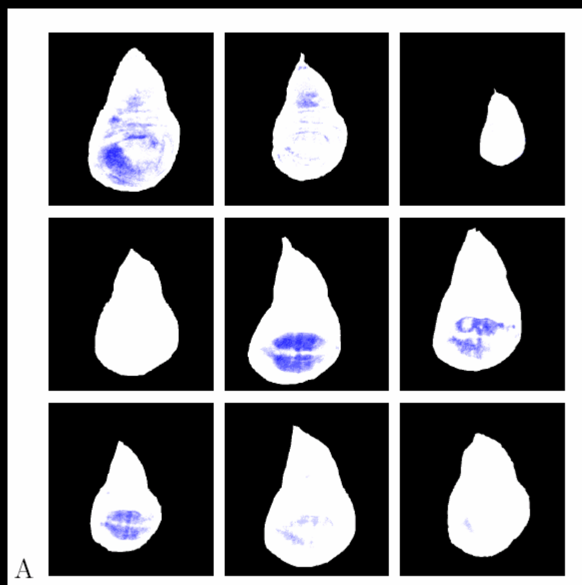
C: Alignment with 7 parameters

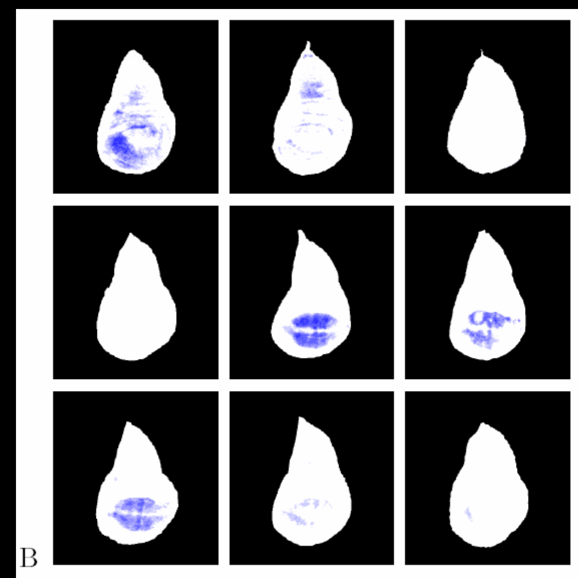
Alignment Results



Wing, Leg, Eye discs: **A**: Before Alignment; **B**: Alignment with 3 parameters; **C**: Alignment with 7 parameters

Stain Scored Patterns: Aligned Vs. Unaligned





Wing discs: **A**: Before Alignment; **B**: After Alignment with 7 parameters

References

- For Further Details:

  - **Parvez Ahammad**, Cyrus Harmon, Ann Hammonds, Shankar Sastry and Gerald Rubin, **'Joint Nonparametric Alignment for Analyzing Spatial Gene Expression Patterns of Drosophila Imaginal Discs'**, *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, June 20-25, 2005, San Diego, CA, USA.
  - Erik Miller, Nick Matsakis and Paul Viola, '**Learning from One Example Through Shared Densities on Transforms**', *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 464-471, 2000.

Thank you!

Questions?