Registration and Temporal Ordering of Images in Studies of Biological Development

Carmeline Dsilva[1], Bomyi Lim[1], Stanislav Shvartsman[1,2], and Ioannis Kevrekidis[1,3]

[1]Department of Chemical and Biological Engineering
[2]Lewis-Sigler Institute for Integrative Genomics
[3]Program in Applied and Computational Mathematics
Princeton University, Princeton, NJ

cdsilva@princeton.edu

15 July 2014

DOE CSGF

PRINCETON UNIVERSITY

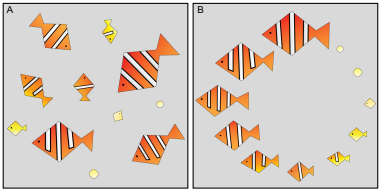
Reconstructing Dynamics from Snapshots

There are two types of data collection schemes

Longitudinal The developmental trajectory of *a single organism* is monitored over time

Cross-sectional Samples from *many organisms* are collected; each organism contributes only a single snapshot from the developmental process

We will look at cross-sectional imaging data.



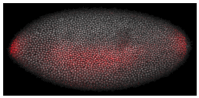
A caricature of ordering cross-sectional data.
(Left) Fish, each in a different stage of development. (Right) Fish, now registered and temporally ordered.
In this example, it is easy to order the data by hand.

Developmental Dynamics of *Drosophila* Embryogenesis

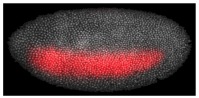
- *Drosophila melanogaster* (common fruit fly) is a common model organism in developmental biology
- We will look at *Drosophila* embryogenesis during the third hour of development

Images of *Drosophila* in the third hour of development

5 min



25 min

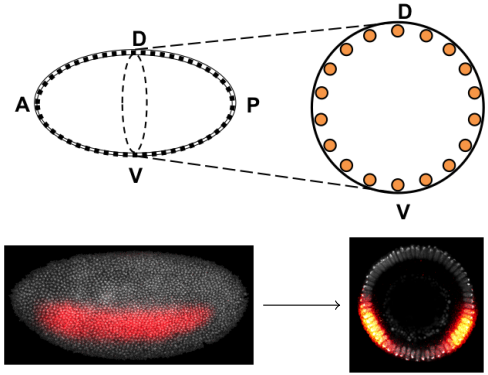


45 min

- Embryos are stained for important regulatory proteins (dpERK, shown in red)
- This staining cannot be done without first fixing the embryos

Drosophila Images

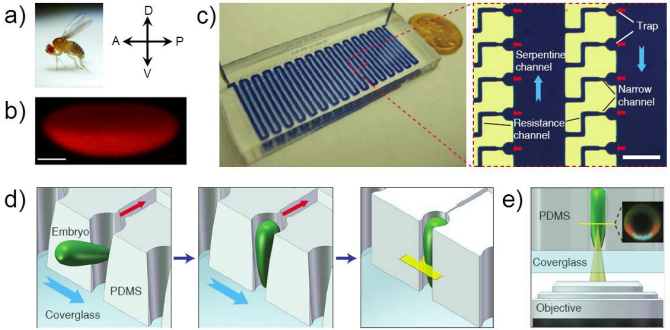
We will look at optical sections perpendicular to the long axis of the embryo

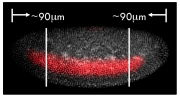


The stripes now appear as peaks around the circumference of the embryo

Data Collection

We use a microfluidic device [1] to easily obtain images of many embryos

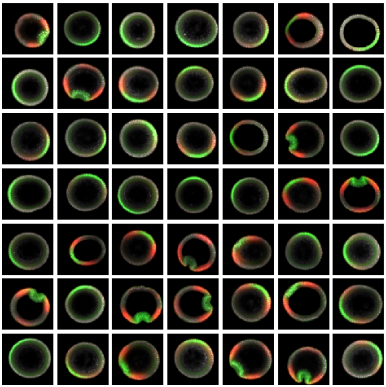




We obtain images at a fixed depth (90 μm) from the tip of the embryo.
Each embryo is in a different rotational orientation.

[1] Chung, K., Y. Kim, J. S. Kanodia, E. Gong, S. Y. Shvartsman, and H. Lu, *Nature Methods*, 2011

A Typical Data Set



- Each image is from a different embryo
- Each embryo is stained for two different proteins (Dorsal, in green, and dpERK, in red)
- Each embryo is at a different (and unknown) developmental time

Mathematical Methods for Registration and Ordering

We will show how we can *automatically* register and order the images

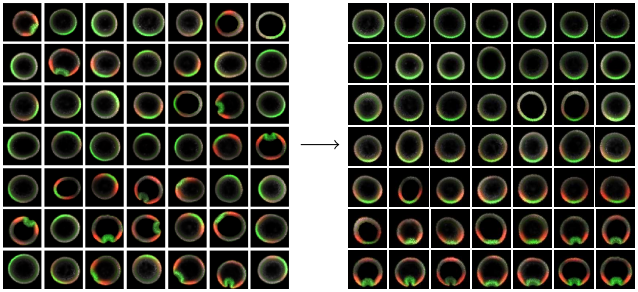


Image registration Angular synchronization [2] *(other methods: templates [3], feature extraction [4,5])*

Temporal ordering Diffusion maps [6] *(other methods: TSP [7], MST [8])*

Simultaneous registration and ordering Vector diffusion maps [9]

[2] Singer, A, *Applied and Computational Harmonic Analysis*, 2011

[3] Ahuja, S., I. G. Kevrekidis, and C. W. Rowley, *Journal of Nonlinear Science*, 2007

[4] Zhao, W., R. Chellappa, P. J. Phillips, and A. Rosenfeld, *Acm Computing Surveys (CSUR)*, 2003

[5] Schindler, G., F. Dellaert, and S. B. Kang, , 2007

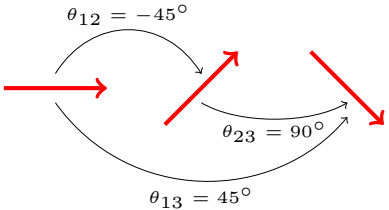
[6] Coifman, R. R. et al., *Proceedings of the National Academy of Sciences*, 2005

[7] Anavy, L. et al., *Development*, 2014

Image Registration: Angular Synchronization

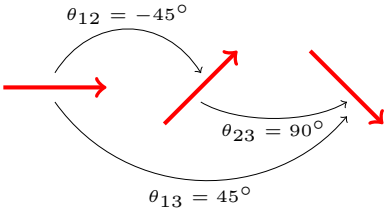
- We want to align or *register* the images, since each image is taken in a different rotation
- We do not know *a priori* what good image features are
- We also do not know what a good template function is

What we can do: compute pairwise alignments

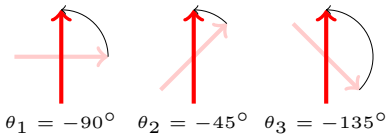


Angular Synchronization: The Assumptions

- Given data points $x_1, \ldots, x_m$, we can compute θ_{ij}, the angle needed to align x_i to x_j.
- We assume that each point x_i is some rotated copy of an underlying signal x_{true} (which we do not know).
- Let θ_i denote the angle to rotate x_{true} to x_i.
- Therefore, $\theta_{ij} \approx \theta_i - \theta_j$, i.e., rotating x_i to x_j is like rotating x_i to x_{true}, and then rotating x_{true} to x_j



Pairwise alignments θ_{ij} between set of vectors



Individual rotations θ_i for each vector

Angular Synchronization: The Algorithm

- We look at the matrix H which contains information about all pairwise alignments

$$H = \begin{bmatrix} e^{\theta_{11}} & e^{\theta_{12}} & \cdots & e^{\theta_{1m}} \\ e^{\theta_{21}} & e^{\theta_{22}} & \cdots & e^{\theta_{2m}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{\theta_{m1}} & e^{\theta_{m2}} & \cdots & e^{\theta_{mm}} \end{bmatrix}$$

- If our assumption holds ($\theta_{ij} \approx \theta_i - \theta_j$), then

$$H \approx \begin{bmatrix} e^{\theta_1-\theta_1} & e^{\theta_1-\theta_2} & \cdots & e^{\theta_1-\theta_m} \\ e^{\theta_2-\theta_1} & e^{\theta_2-\theta_2} & \cdots & e^{\theta_2-\theta_m} \\ \vdots & \vdots & \ddots & \vdots \\ e^{\theta_m-\theta_1} & e^{\theta_m-\theta_2} & \cdots & e^{\theta_m-\theta_m} \end{bmatrix} = \begin{bmatrix} e^{\theta_1} \\ e^{\theta_2} \\ \vdots \\ e^{\theta_m} \end{bmatrix} \begin{bmatrix} e^{-\theta_1} & e^{-\theta_2} & \cdots & e^{-\theta_m} \end{bmatrix}$$

and the top eigenvector of H gives us estimates of the optimal rotations

Temporal Ordering: Diffusion Maps

- We want to temporally order the (registered) images in time
- Our images are high-dimensional vectors
- We assume that our images lie on a one-dimensional (perhaps nonlinear) curve
- We also assume that this curve is parameterized by time

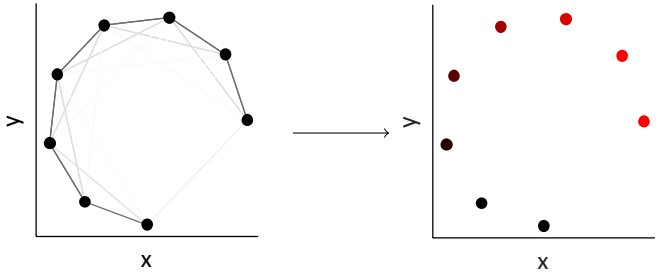


Illustration of two-dimensional data on a one-dimensional curve.

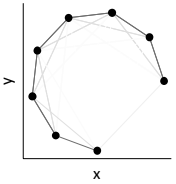
Data, now colored by the one-dimensional parameterization. We assume this parameterization is one-to-one with time.

Diffusion Maps Algorithm

- Given m data points $x_1, \ldots, x_m$, we want to find a coordinate transformation $y(x)$ that preserves local information: points that are close in the original space should also be close in the coordinate y.
- The first step is to construct the matrix $W \in \mathbb{R}^{m \times m}$, where W_{ij} is large if points x_i and x_j are "close."

$$W_{ij} = \exp\left(-\frac{d^2(x_i, x_j)}{\epsilon^2}\right)$$

$d(x_i, x_j)$ is the distance between x_i and x_j; ϵ is a characteristic scale



Two-dimensional data which lies on a one-dimensional curve. The edge strength is proportional to the kernel W_{ij}.

- To find this coordinate y, we want solve the following optimization problem [10]

$$\arg\min_y \sum_{ij} W_{ij}(y(x_i) - y(x_j))^2.$$

- The solution is given by the top (non-trivial) eigenvector of A, where $A = D^{-1}W$ and D is a diagonal matrix with $D_{ii} = \sum_{j=1}^{m} W_{ij}$.

[10] Belkin, M. and P. Niyogi, *Neural Computation*, 2003

Combined: Vector Diffusion Maps

Angular Synchronization

Calculate top eigenvector of H,

$$H_{ij} = e^{-i\theta_{ij}}$$

Diffusion maps

Calculate top eigenvectors of A,

$$A_{ij} = \frac{\exp\left(-\frac{d^2(x_i,x_j)}{\epsilon^2}\right)}{\sum_j \exp\left(-\frac{d^2(x_i,x_j)}{\epsilon^2}\right)}$$

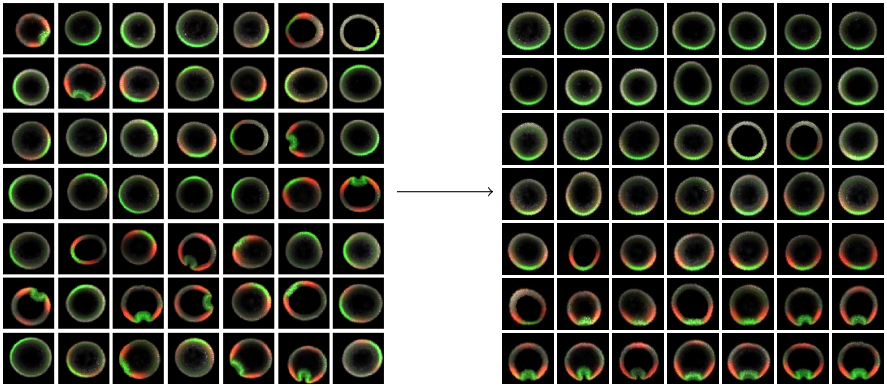
Vector Diffusion Maps (VDM) [11]

Calculate the top eigenvectors of S, where $S_{ij} = A_{ij}H_{ij}$

- The top eigenvector(s) of S give us the optimal alignments/rotations for our images
- Successive eigenvectors give us the embedding coordinates/temporal ordering for our images

[11] Singer, A. and H.-T. Wu, *Communications on Pure and Applied Mathematics*, 2012

Registration and Ordering of Images Using Vector Diffusion Maps

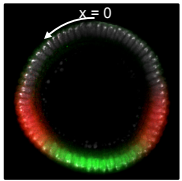


Images, unregistered and unordered

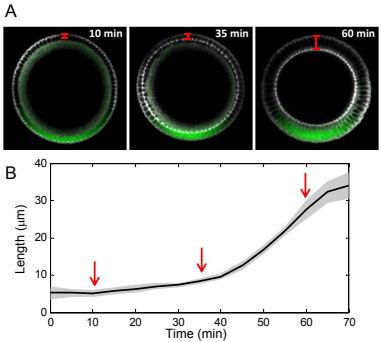
Images, registered and ordered using vector diffusion maps

Validating Registration and Temporal Ordering

For this particular system, both registration and temporal ordering can be done using prior knowledge



Registration: The green peak specifies the ventralmost point of the embryo



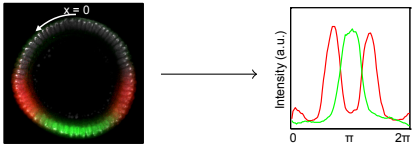
Ordering: Membranes grow in thickness as a function of time [12]

Note that the curve becomes flatter and noisier at the end of the developmental time window

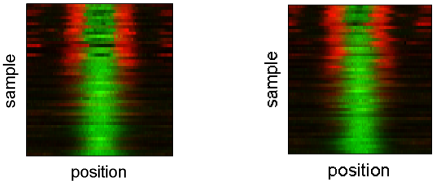
[12] Figard, L., H. Xu, H. G. Garcia, I. Golding, and A. M. Sokac, *Developmental cell*, 2013

Validation of Automatic Registration and Ordering

We compare concentration profiles around the circumference of the embryo



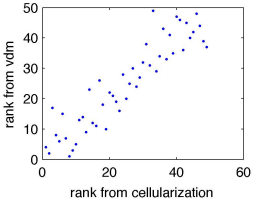
The results appear consistent



Registered using green peaks
Ordered using membranes

Registered and ordered using
vector diffusion maps

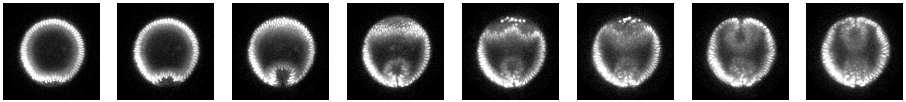
Rank correlation



$\rho = 0.9158$

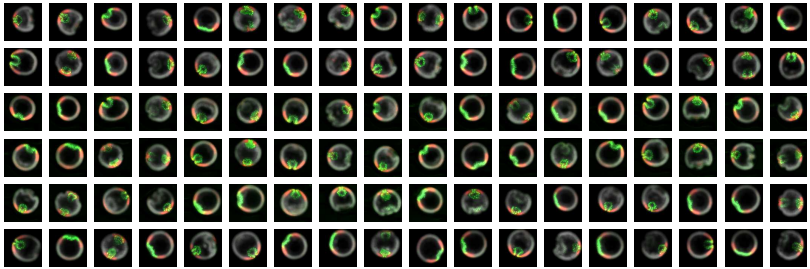
Registering and Ordering a More Complex Data Set

Following cellularization, there is gastrulation, when the embryos begin to morph and deform.



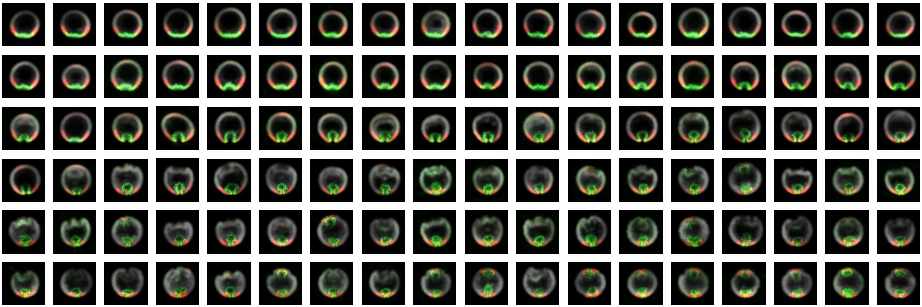
0 min | 2.5 min | 5 min | 7.5 min | 10 min | 12.5 min | 15 min | 17.5 min

During this time period, we have no good way to accurately measure the time of fixed images.

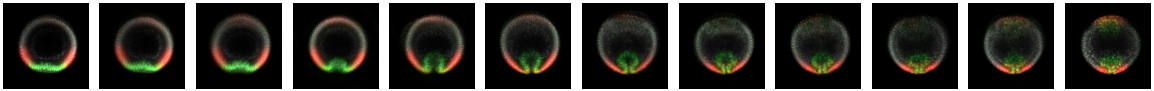


We want to register and temporally order this data.

Registered and Ordered using VDM



The data has now been registered and ordered using VDM



We can compute an *average* developmental trajectory.

Conclusions

- In many experiments, we are presented with cross-sectional imaging data that must be registered and temporally ordered
- Both of these tasks can be automated, using angular synchronization (for registration), diffusion maps (for temporal ordering), and vector diffusion maps (for simultaneous registration + ordering)
- In the first example, we have independent time and space markers, and we can validate our automatic methods
- We are now interested in matching fixed images with *live movies* that contain only one signal (e.g., nuclei) so that we can determine the rates of relevant developmental processes

Thank you to CSGF and Krell!

DOE CSGF