Inferring nuclear movements from fixed material

Charless Fowlkes
Jitendra Malik

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2006-142
http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-142.html

November 7, 2006

Inferring nuclear movements from fixed material

Charless C. Fowlkes, Jitendra Malik
Berkeley Drosophila Transcription Network Project
http://bdtnp.lbl.gov
fowlkes@eecs.berkeley.edu

Abstract

We describe a technique for inferring the typical movement of nuclei in Drosophila blastoderm using nuclear positions extracted from a large number of images of fixed embryos. Embryos are sorted into temporal cohorts and each cohort is represented by the average blastoderm shape and average density of nuclei along the blastoderm surface. To find cell movements, we formulate a cost function that measures how well a given placement of a set of "synthetic nuclei" respects the measured average density for the cohort. This function is optimized for each cohort in turn, initialized with the results of the previous time step. The result is a synthetic time series of changing nuclear locations which recapitulates average nuclear density and blastoderm shape seen under the microscope.

1 Introduction

Establishing a suitable coordinate system for analyzing morphology and spatial gene expression is a general problem in developmental biology because cells move. At times, these dynamics are obvious to the human eye, such as the rapid movements visible during gastrulation. To discern more subtle movements, developmental biologists typically track the locations of cells over time using live cell imaging. Identifying corresponding cells in successive frames of a live video is fairly straightforward, but this approach is often limited by poor signal to noise and high light attenuation along the optical axis of the microscope.

Imaging fixed material can overcome these limitations as it is generally possible to achieve higher quality fluorescent staining and, by manipulating the embedding media, to increase transparency of the tissue. However, estimating motions from fixed material poses it's own difficulties as it requires some notion of correspondence between cells in different animals. For tissues with fairly homogeneous structure, such as the Drosophila blastoderm, it is quite difficult (using morphology alone) to find corresponding cells across samples at the accuracy necessary to measure small motions.

In this report, we describe a computational approach to predicting nuclear/cellular movements which relies on having a large number of images of fixed material in order to estimate the average positions of all cells in the embryo or tissue under study.[1] The difficulties of correspondence are sidestepped by using average cell densities computed over hundreds of embryos. We suggest that when it is possible to obtain such data for cohorts at different time points, the strategy we have employed based on fixed material provides a powerful alternative to live cell analysis for estimating typical nuclear/cell movements.

2 Nuclear movement in Drosophila blastoderm

We employed a dataset of 1282 Drosophila embryos which were stained to fluorescently label nuclear DNA and imaged via two photon laser-scanning microscopy. The collection of embryos, spanning developmental stage 5, were grouped into 6 tightly spaced temporal cohorts based on visual examination of the extent of cell membrane invagination. The ~ 6000 nuclei visible in each volumetric image were localized and segmented using the techniques described in [1]. The result of this analysis is a list of estimated nuclear locations in 3D for each embryo which we refer to as a pointcloud.

[1] The biological implications of the movements we model in Drosophila are described elsewhere [2]

Although the nuclei in stage 5 blastoderm form a monolayer along the egg surface, we observed that the density of nuclear packing varies across this surface in a complex but consistent pattern which seems to foretell the much larger displacements that occur during gastrulation. In addition, the density patterns we observed change over the course of cellularization, with density increasing dorsally and decreasing ventrally and anteriorly (indicated by the colormap in Figure 1). Since nuclei do not divide during this time period one may conclude that these changes in density are due to movements of the nuclei.

From a collection of pointclouds, we would like to generate a synthetic time series that describes an average number of nuclei undergoing typical movements in wild-type embryos. The motion at a given time can be decomposed into two components. The first is a sliding motion of nuclei tangent to the blastoderm surface. The second is an inwards motion of nuclei which changes the overall blastoderm "shape". Our approach to recovering movements has a similar decomposition: we first set the shape of the blastoderm surface at a given time point and then place nuclei on this surface so as to replicate the observed density pattern.

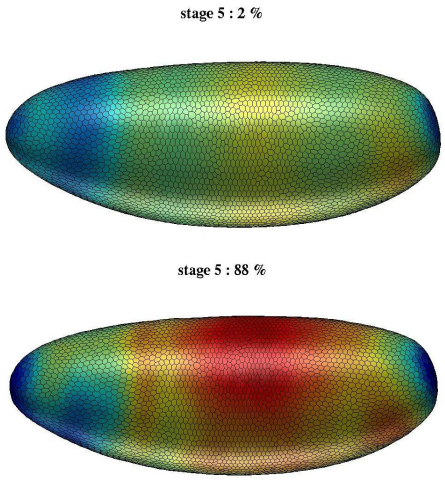

Figure 1: Blastoderm shape and nuclear density undergo subtle but significant changes over the course stage 5.

3 Parametrizing blastoderm surface shape

We coarsely align the embryos in a given cohort by a translating each embryo to place the center of mass at the origin and rotating to align the principal axis along the x coordinate. Embryos are also rotated around the central axis so that the dorsal mid-line (identified by eye) lies in the x-z plane.

We parametrize the surface of the blastoderm in spherical coordinates:

$$\begin{aligned} x &= r^t(\phi,\theta)\sin\phi\cos\theta \\ y &= r^t(\phi,\theta)\sin\phi\sin\theta \\ z &= r^t(\phi,\theta)\cos\phi \end{aligned}$$

Here r^t is the the average distance to the blastoderm surface for a given temporal cohort t which is computed by interpolating the pointcloud data to a fixed grid and averaging.

This parametrization serves to decouple the problems of surface shape and nuclear density. Given the (ϕ,θ) coordinate for each synthetic nucleus we simply plug it into our estimate of r^t, yielding 3D coordinates which automatically satisfy the shape constraint.

4 Inferring cell locations from density

Let $\mathbf{Y}_i^t = (\phi_i^t, \theta_i^t)$ be the location of the ith nucleus in spherical projection coordinates at time t. Given an estimate of the average density of nuclei in the spherical projection, D^t, specified at a set of locations $\mathbf{Z}_j$, we seek locations for the nuclei which minimizes:

$$C(\{\mathbf{Y}_i^t\}) = \frac{1}{2}\sum_j \left\| D^t(\mathbf{Z}_j) - \sum_i K_\sigma(\|\mathbf{Z}_j - \mathbf{Y}_i^t\|) \right\|^2$$

where the sum of kernels, $K_\sigma(\|\mathbf{Z}_j - \mathbf{Y}_i^t\|) = e^{-\|\mathbf{Z}_j - \mathbf{Y}_i^t\|^2/2\sigma^2}$ gives an estimate of the density at point $\mathbf{Z}_j$ as a function of the placement of the nuclei $\{\mathbf{Y}_i^t\}$ [4]. Estimating the density in this way has the advantage that the gradient of C is straightforward to express, e.g.

$$\frac{\partial C}{\partial \mathbf{Y}_i^t} = \frac{1}{\sigma^2}\sum_j (\mathbf{Z}_j - \mathbf{Y}_i^t) K_\sigma(\|\mathbf{Z}_j - \mathbf{Y}_i^t\|) \left(D^t(\mathbf{Z}_j) - \sum_i K_\sigma(\|\mathbf{Z}_j - \mathbf{Y}_i^t\|) \right)$$

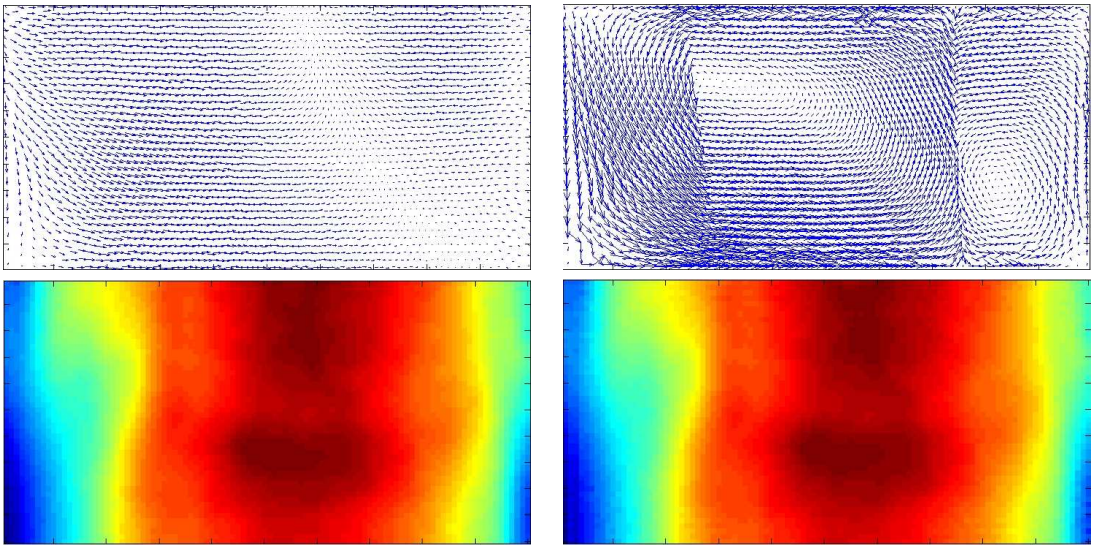

Figure 2: Two different flows which yield essentially identical final density patterns. The left column shows a slow, smooth motion and the resulting final density. The right shows the results of optimizing the un-regularized cost function with an extreme initialization of the motion and the resulting density pattern.

We find a set of nuclear locations that (locally) minimize C using conjugate gradient [3]. We initialize the optimization with a uniform triangular grid of 6078 points. Optimization typically converges within a few hundred line searches. Minimizers are clearly not unique since any permutation of the nuclei yields the same density pattern. However, this isn't a problem as the the nuclei are all identical for the purposes of this initialization step.

Given a placement of nuclei at timestep $t-1$ we would like to move these nuclei to new locations at time t which respect the updated density and shape. This is accomplished by local minimization starting from the initial condition $\{Y^{t-1}\}$ given by the previous time step.

As stated, this problem also does not have a unique solution since there are many possible motions which yield the same final pattern of density. Figure 2 shows an example of two such motions. To make the problem well formed, we add a regularization term to the cost function. This is also desirable from a biological perspective as the nuclear motions we observe in living embryos are slow and smooth. To enforce regularity in the resulting field of displacement vectors, $Y^t - Y^{t-1}$, we augment the cost function with two additional terms:

$$C(\{\mathbf{Y}_i^t\}) = \frac{1}{2}\sum_j \left\| D^t(\mathbf{Z}_j) - \sum_i K_\sigma(\|\mathbf{Z}_j - \mathbf{Y}_i^t\|) \right\|^2 + \frac{\alpha}{2}\sum_{i\neq j} \left\| \frac{(\mathbf{Y}_i^t - \mathbf{Y}_i^{t-1}) - (\mathbf{Y}_j^t - \mathbf{Y}_j^{t-1})}{\|\mathbf{Y}_i^{t-1} - \mathbf{Y}_j^{t-1}\|} \right\|^2 + \frac{\beta}{2}\sum_i \|\mathbf{Y}_i^t - \mathbf{Y}_i^{t-1}\|^2$$

The first regularization term, weighted by α, is a smoothness term which asks that neighboring cells (for which the denominator is small) have similar displacement vectors (numerator is small). This has a local smoothing affect as cells which are far away (large denominator) contribute very little to the penalty. The second term, weighted by β, enforces that the total motion be small.

In practice, we find that the results are quite robust to choice of the parameters α, β. For small time increments using the solution from the previous time step as initialization, the regularization terms are not even necessary. Since conjugate gradient performs a local search, starting from zero displacement assures the total displacement is small while the repulsion inherent in the density computation tends to yield smooth displacements. Qualitatively different solutions (as in Figure 2) only appear with very large perturbations of the initial conditions.

5 Using multiple coordinate charts

We have specified the quality of a proposed nuclear placement as a function of the nuclear coordinates in a spherical parametrization of the blastoderm surface. This has the short coming that near the poles, the density of nuclei in this co-

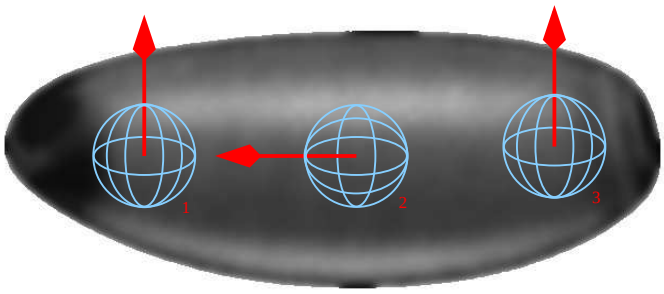

Figure 3: Covering the embryo with three different spherical charts. The orientations are chosen so that the coordinate system singularities do not overlap. The locations are chosen so that every part of the surface is nicely covered by some chart (i.e. chart 2 captures the middle 80% of the blastoderm while 1 and 3 capture the anterior and posterior poles respectively)

ordinate system goes to zero, as the "lines of longitude" converge to a point. This could be solved by a spatially varying σ parameter in order to give a smooth estimate of the density. However, the poles still pose the additional difficulties of how to specify the appropriate boundary conditions in estimating the densities and modeling the possibility of nuclei moving over the poles.

We overcome this problem by covering the blastoderm with three different spherical coordinate systems whose singularities occur at different locations on the surface (shown in Figure 3). For any point on the blastoderm surface, we can always choose a coordinate system to represent it which has no singularities nearby.

Let T^k be the transformation which takes coordinates for point on the surface of the blastoderm in coordinate system k to our standard coordinate system. Let D^k be the desired average density measured in the kth coordinate system. We augment our original cost function to be blend of cost functions from each coordinate system:

$$\begin{aligned} C(\{\mathbf{Y}_i\}) &= \sum_k C_k(T_k(\{\mathbf{Y}_i\})) \\ &= \frac{1}{2}\sum_{k,j} w_{k,j} \left\| D_k(\mathbf{Z}_{k,j}) - \sum_i K_\sigma(\|\mathbf{Z}_{k,j} - T_k(\mathbf{Y}_i)\|) \right\|^2 \end{aligned}$$

For clarity, we've dropped the time index and regularization terms. $w_{k,j}$ is a weight which specifies the extent to which the kth coordinate system should be used to represent the jth density measurement. In the vicinity of a coordinate system singularity we'd like this weight to be 0. In particular we let w be a smooth function of the ϕ coordinate of Z_k. For the standard coordinate system, w is a product of logistic functions centered at $\phi = \pi/10$ and $\phi = 9\pi/10$, covering the central 80% of the blastoderm. We use two additional spherical charts to cover the remaining 10% of each pole weighted by a single logistic function. Figure 4 shows the coordinate charts with their weightings.

Gradients of each C_k have the same form as those of the original cost function with the addition of appropriate spatial weighting. In order to minimize C, however, it is necessary to transform gradients computed in different coordinate charts back into the standard coordinate chart. In particular, the partial derivatives with respect to the standard coordinates are linearly related by the chain rule:

$$\begin{bmatrix} \frac{\partial C}{\partial \phi^2} \\ \frac{\partial C}{\partial \theta^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial \phi^1}{\partial \phi^2} & \frac{\partial \theta^1}{\partial \phi^2} \\ \frac{\partial \phi^1}{\partial \theta^2} & \frac{\partial \theta^1}{\partial \theta^2} \end{bmatrix} \begin{bmatrix} \frac{\partial C}{\partial \phi^1} \\ \frac{\partial C}{\partial \theta^1} \end{bmatrix}$$

For completeness, we give details of transforming gradient vectors from one coordinate system to another. To deal with the offsets between the centers of the three spherical coordinate systems, it is convenient to implement transforms by expanding the Jacobian in terms of intermediate Euclidean coordinates:

$$\begin{bmatrix} \frac{\partial \phi^1}{\partial \phi^2} & \frac{\partial \theta^1}{\partial \phi^2} \\ \frac{\partial \phi^1}{\partial \theta^2} & \frac{\partial \theta^1}{\partial \theta^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial \phi^1}{\partial x} & \frac{\partial \phi^1}{\partial y} & \frac{\partial \phi^1}{\partial z} \\ \frac{\partial \theta^1}{\partial x} & \frac{\partial \theta^1}{\partial y} & \frac{\partial \theta^1}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial \phi^2} & \frac{\partial x}{\partial \theta^2} \\ \frac{\partial y}{\partial \phi^2} & \frac{\partial y}{\partial \theta^2} \\ \frac{\partial z}{\partial \phi^2} & \frac{\partial z}{\partial \theta^2} \end{bmatrix}$$

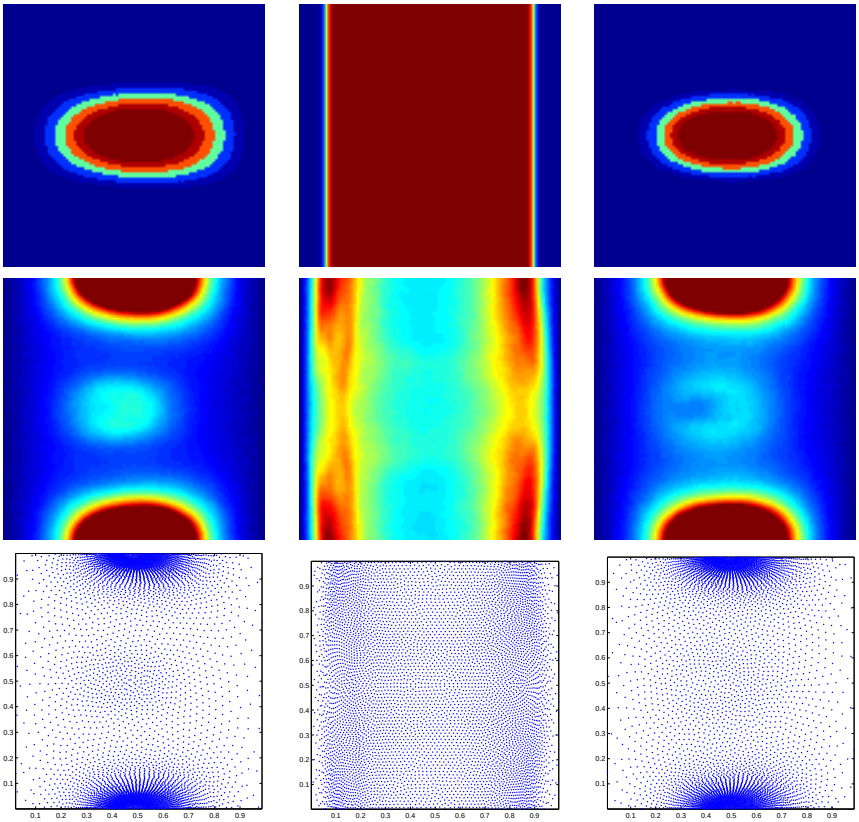

Figure 4: Each column corresponds to a coordinate chart in Figure 3. The top row shows the weighting $w_{1,2,3}$ used by the chart. The middle row shows the density $D_{1,2,3}$ of nuclei in the chart. Note the drop in the average density on the left and right edges of each chart ($\phi = 0, \pi$). The bottom row shows the synthesized nuclear locations $T_{1,2,3}(Y)$ which produce matching density maps.

Recall our parametric description of the embryo surface at time t in terms of spherical coordinates.

$$\begin{aligned} x &= r(\phi,\theta)\sin\phi\cos\theta \\ y &= r(\phi,\theta)\sin\phi\sin\theta \\ z &= r(\phi,\theta)\cos\phi \end{aligned}$$

Taking derivatives allows us to fill in the matrix, lifting tangent vectors in the spherical chart to the Euclidean embedding space:

$$\begin{aligned} \frac{\partial x}{\partial \phi} &= \frac{\partial r(\phi,\theta)}{\partial \phi}\sin\phi\cos\theta + r(\phi,\theta)\cos\phi\cos\theta \\ \frac{\partial y}{\partial \phi} &= \frac{\partial r(\phi,\theta)}{\partial \phi}\sin\phi\sin\theta + r(\phi,\theta)\cos\phi\sin\theta \\ \frac{\partial z}{\partial \phi} &= \frac{\partial r(\phi,\theta)}{\partial \phi}\cos\phi - r(\phi,\theta)\sin\phi \\ \frac{\partial x}{\partial \theta} &= \frac{\partial r(\phi,\theta)}{\partial \theta}\sin\phi\cos\theta - r(\phi,\theta)\sin\phi\sin\theta \\ \frac{\partial y}{\partial \theta} &= \frac{\partial r(\phi,\theta)}{\partial \theta}\sin\phi\sin\theta + r(\phi,\theta)\sin\phi\cos\theta \\ \frac{\partial z}{\partial \theta} &= \frac{\partial r(\phi,\theta)}{\partial \theta}\cos\phi \end{aligned}$$

This can then be inverted to yield the mapping from Euclidean back to spherical:

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{\cos\phi\cos\theta}{r(\phi,\theta)} \\ \frac{\partial \theta}{\partial x} &= \frac{\csc\phi\sin\theta}{r(\phi,\theta)} \\ \frac{\partial \phi}{\partial y} &= \frac{\cos\phi\sin\theta}{r(\phi,\theta)} \\ \frac{\partial \theta}{\partial y} &= \frac{\cos\theta\csc\phi}{r(\phi,\theta)} \\ \frac{\partial \phi}{\partial z} &= -\frac{\sin\phi}{r(\phi,\theta)} \\ \frac{\partial \theta}{\partial z} &= 0 \end{aligned}$$

6 Conclusion

Figures 5 and 6 summarize the results of our modeling effort. The synthetic nuclear density maps produced by this procedure agreed closely with maps measured from actual embryos at the same stages of development. Although density alone is a fairly weak constraint, the models requirement of a small, smooth movement resulted in a solution that was quite robust to perturbations of the constraints and initial conditions. Figure 6 shows the map of predicted nuclear movements between the early and late synthetic embryos. Qualitatively, the predicted movements matched those observed in the live data (see [2]), showing larger movements at the poles and dorsally than ventrally. Quantitatively, the movements were of a similar order as those we observed in living embryos.

Our work establishes a new strategy for measuring temporal changes in the locations of cells and gene expression patterns that uses fixed cell material and computational modeling. It also provides a coordinate framework for the blastoderm embryo which we hope will allow increasingly accurate spatio-temporal modeling of both morphogenesis and the transcriptional network that controls it.

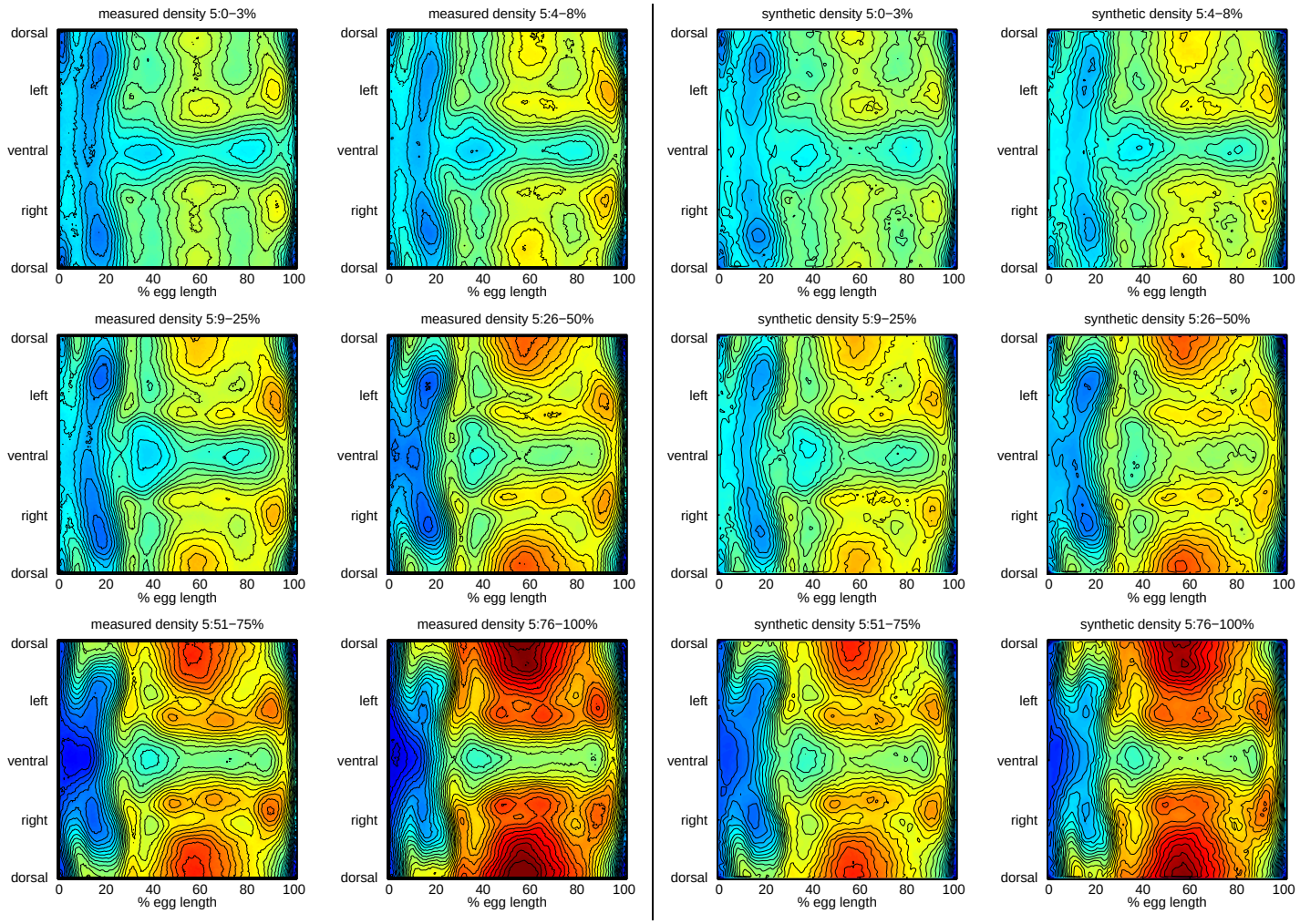

Figure 5: Nuclear density on the surface of the blastoderm displayed in cylindrical projection (anterior to the left). Left panel shows the average of densities measured in six temporal cohorts spanning the course of stage 5. Right panel shows the corresponding density pattern for the synthetic embryo undergoing estimated ”typical” nuclear movements.

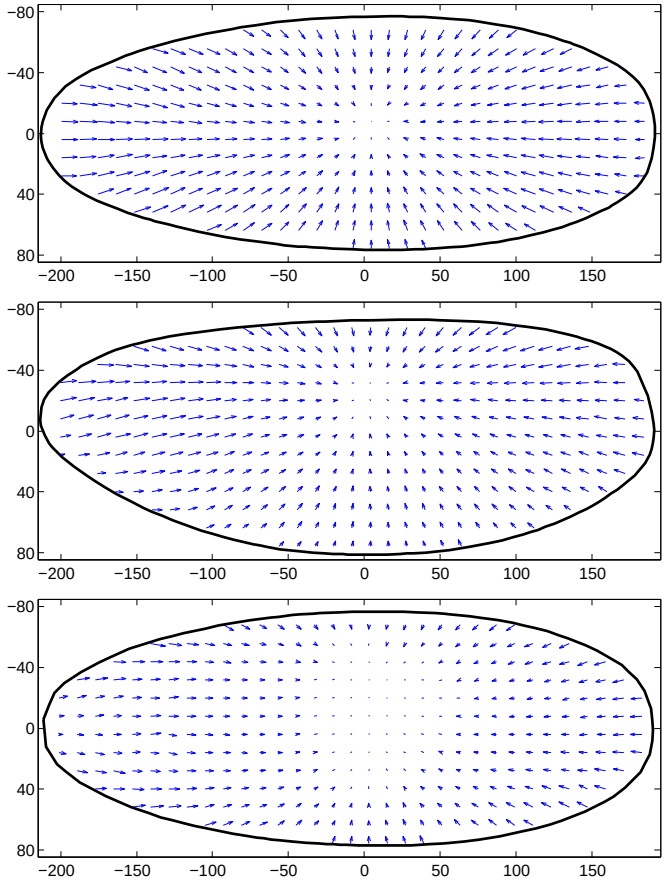

Figure 6: Predicted nuclear motions based on average shape and density changes.

References

[1] C. L. Luengo Hendriks, S. V. E. Keränen, C. C. Fowlkes, L. Simirenko, G. H. Weber, C. Henriquez, D. W. Kaszuba, B. Hamann, M. B. Eisen, J. Malik, D. Sudar, M. D. Biggin, and D. W. Knowles. 3d morphology and gene expression in the drosophila blastoderm at cellular resolution. *in press*, 2006.

[2] S. V. E. Keränen, C. C. Fowlkes, C. L. Luengo Hendriks, D. Sudar, D. W. Knowles, J. Malik, and M. D. Biggin. The dynamics of morphology and gene expression in the drosophila cellular blastoderm. *in press*, 2006.

[3] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P.Flannery. *Numerical Recipies in C, 2nd Edition*. Cambridge University Press, 1992.

[4] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapmann and Hall, 1986.