Joint Nonparametric Alignment for Analyzing Spatial Gene Expression Patterns in Drosophila Imaginal Discs

Parvez Ahammad,* Cyrus L. Harmon,† Ann Hammonds,† S. Shankar Sastry* and Gerald M. Rubin‡

Abstract

To compare spatial patterns of gene expression, one must analyze a large number of images as current methods are only able to measure a small number of genes at a time. Bringing images of corresponding tissues into alignment is a critical first step in making a meaningful comparative analysis of these spatial patterns. Significant image noise and variability in the shapes make it hard to pick a canonical shape model. In this paper, we address these problems by combining segmentation and unsupervised shape learning algorithms. We first segment images to acquire structures of interest, then jointly align the shapes of these acquired structures using an unsupervised nonparametric maximum likelihood algorithm along the lines of 'congealing' [12], while simultaneously learning the underlying shape model and associated transformations. The learned transformations are applied to corresponding images to bring them into alignment in one step. We demonstrate the results for images of various classes of Drosophila imaginal discs and discuss the methodology used for a quantitative analysis of spatial gene expression patterns.

1. Introduction

Microarray technologies have enabled researchers to measure levels of expression of large numbers of individual genes in a single experiment, thereby providing a greatly enhanced view of the genes that are active at a given time in specific tissues [10, 1]. These techniques generally require a large tissue sample and provide no information about the spatial patterns of gene expression. In situ hybridization of tissues with labeled probes for individual genes facilitates the measurement of precise spatial patterns of gene expression at high resolution. However, in situ hybridization can

*Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, USA.

†Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.

‡Howard Hughes Medical Institute and Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.
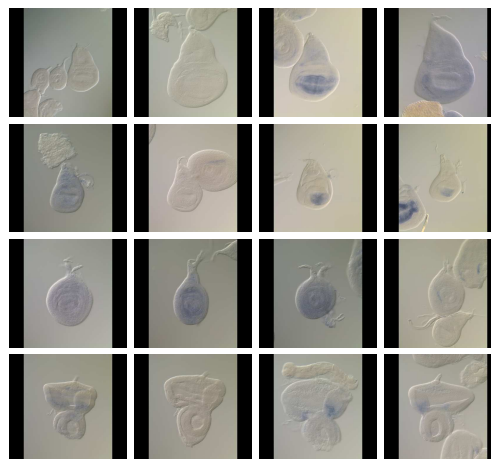
Figure 1. A typical set of in situ stained Drosophila imaginal disc images. Each row shows a different tissue class. First row: wing discs, second row: haltere discs, third row: leg discs, fourth row: eye discs. There is significant inter-class and intra-class variability both in the shapes, sizes and the stain patterns.

only be used to measure a limited number of genes at a time, usually one. It is desirable to be able to measure the spatial expression patterns of large numbers of genes and to be able to compare, cluster, and classify patterns of expression. Current experimental strategy entails setting up a high-throughput production system for the generation of large numbers of images which can then be processed by either human or computer. Many dipteran organisms, including the fruit fly *Drosophila melanogaster*, have a three-stage life-cycle in which the insect begins as an embryo, becomes a larva and metamorphoses into an adult, also known as an imago. The primordial tissues that will become the exoskeleton of the adult insect, or imago, called imaginal discs, are segregated from the larval tissues in embryogenesis and are maintained in sac-like structures that have a morphology similar to that of a flattened balloon[3].

1.1. Previous Work

Large-scale studies of patterns of gene expression in Drosophila have been performed using DNA microarrays both on whole organisms [2] and individual tissues such as imaginal discs. Klebes et al. compared differential gene expression in different imaginal discs and between imaginal discs and non-disc tissue [8]. Butler et al. manually dissected imaginal discs and were able to identify transcripts that were enriched in specific compartments of the wing discs [5]. Recent studies of precise spatial patterns of gene expression for large numbers of genes in developing Drosophila embryos through in situ hybridization [14, 4] require annotation, and suffer from the fact that the annotation of spatial expression pattern requires manual curation. Kumar et al. [9] applied machine vision techniques to low-resolution images of in situ stained embryos and developed an algorithm for searching a database of patterns of gene expression in the embryos. Peng and Myers [13] have performed automated embryo registration and stain classification by using Gaussian mixture models. Yet, most of the previous work makes some parametric assumptions on the shape of the tissue, and the registration techniques used are very simplistic (such as aligning the major and minor axes for embryo images [13]). Shape learning and alignment is a well studied problem in the computer vision community. In our work, we chose *congealing* [12] since we are interested in aligning the binary shape masks that we obtain after segmentation and *congealing* is readily applicable. For further discussion on relevant shape alignment algorithms, reader is referred to [6, 15, 7, 12].
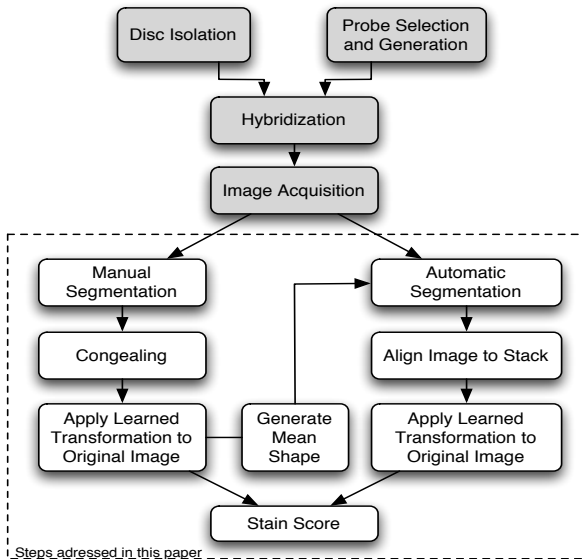

Figure 2. Data flow for our proposed approach.

1.2. Motivation and Problem Definition

The tissue structures in typical Drosophila imaginal discs have significant intra and inter-class variability in size, shape and stain patterns. Thus, shape models learned from one class cannot be used for another tissue class without making significant changes in the processing pipeline if one were to use model based alignment algorithms. Furthermore, there are far fewer identified and named morphological parts or regions in Drosophila imaginal discs than in the developing Drosophila embryo, for example. It is difficult to make any parametric or model based assumptions for tissue shapes given these limitations. Manual annotation and curation are extremely time-consuming and costly in high-throughput spatial gene expression analysis experiments. It is highly desirable to have a processing pipeline that can operate for various imaginal disc tissues such as wings, halteres, legs, eyes, etc., (shown in Figure 1) without significant re-structuring. In this paper, we propose a simple yet effective computational methodology (Figure 2) that addresses these demands of high-throughput systems for the analysis of spatial gene expression by combining segmentation and nonparametric alignment algorithms.

Given an input ensemble of noisy Drosophila imaginal disc images of a given tissue class, our goal is to learn the underlying shape model of the tissue nonparametrically while bringing the given images into alignment. This alignment greatly facilitates quantitative stain scoring analysis (Section 5) on the imaginal disc images.

2. Image Model

Imaginal disc has a morphology similar to that of a flattened balloon. Imaginal discs do have substantial depth to them, but we image a single plane from the discs and consider an idealized 2-dimensional representation of a disc.

We denote the set of input imaginal disc images of a given class as $\Phi \doteq \{I^i\}_{i=1}^N$ where N is the cardinality of the set. Each image $I^i(\cdot)$ can be represented as a map from the image $\mathbb{R}^3$ (in homogeneous coordinates) to the color space $C \subset \mathbb{R}^3$ with a small compact support $\Omega \subset \mathbb{R}^2$:

$$I^i(\boldsymbol{x}) :: \quad \boldsymbol{x} \in \Omega \ \mapsto \ \boldsymbol{c} = I^i(\boldsymbol{x}) \in C; \quad (1)$$

where $\boldsymbol{x} = [x, y, 1]^T \in \mathbb{R}^3$ (in homogeneous coordinates), $\boldsymbol{c} = [r, g, b]^T \in \mathbb{R}^3$ is a vector in the color space.

3. Extracting Tissue Shapes

Two salient features of our image dataset make the segmentation task relatively simple. First, Nomarski images of the discs [14] yield substantial highlights and lowlights at the periphery of the discs. The background of the images
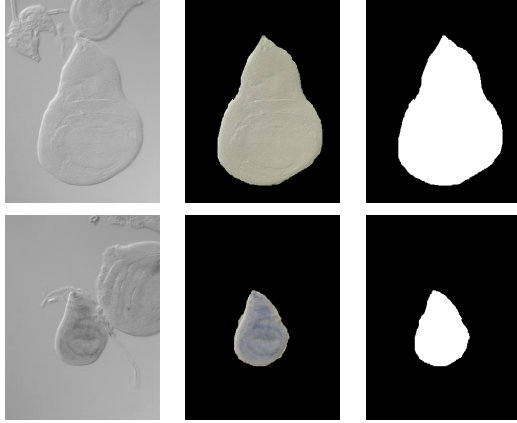
Figure 3. Example segmentation results for wing discs (*first row*) and haltere discs (*second row*) using the combined segmentation procedure. **Left Column**: Original image $I^i(\boldsymbol{x})$. **Middle Column**: Segmented tissue structure of interest $I^i_f(\boldsymbol{x})$. **Right Column**: Extracted binary shape $I^i_s(\boldsymbol{x})$.

Figure 4. Mean images from optimization process during *Congealing* for wing discs (*first column*), haltere discs (*second column*) leg discs (*third column*), eye discs (*fourth column*). **A**: Mean image of Φ_s *before* congealing. **B**: Mean image of Φ_s *after* congealing to convergence with only 3 parameters (t_x, t_y and θ). **C**: Mean image of Φ_s *after* congealing to convergence with 7 parameters (Equation 8).

is generally uniform and one can use the significant contrast generated at the edge of the discs to identify border regions. Second, compared to the background, the pixel intensities of the imaginal disc tissues have much more variability than background, even over a small window. Using these insights, we implemented a simple filter-and-threshold module for segmentation. It computes the *local variance* of the image in a small support region, estimates the bimodal distribution of variance in the filtered image and thresholds it appropriately to separate the disc region from the background.

The extracted shape image I^i_s is then calculated as:

$$I^i_s(\boldsymbol{x}) = \begin{cases} 1, & \text{if } var(I^i(\boldsymbol{x})) \geq \delta \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where $I^i_s(\boldsymbol{x})$ is a binary image of the extracted shape, $\boldsymbol{x} \in \Omega$ and δ is a threshold value where $\delta \in \mathbb{R}$. The segmented structure of interest (or the foreground) $I^i_f(\boldsymbol{x})$ can be computed by point-wise multiplication of $I^i(\boldsymbol{x})$ and $I^i_s(\boldsymbol{x})$:

$$I^i_f(\boldsymbol{x}) = I^i(\boldsymbol{x}) \cdot I^i_s(\boldsymbol{x}). \tag{3}$$

Sometimes this simple filter-and-threshold process results in unsatisfactory performance. This can be addressed by performing a template matching operation to identify the disc. Given the rough shape template of the disc, the template matching operation is quite simple but the problem is that there is no such clean shape template to begin with. We address this issue as follows: using manual segmentation on a set of disc images, we obtain relatively clean shapes of the tissue for each class. These relatively clean structures are then fed to the nonparametric shape learning algorithm to form a good canonical shape template.
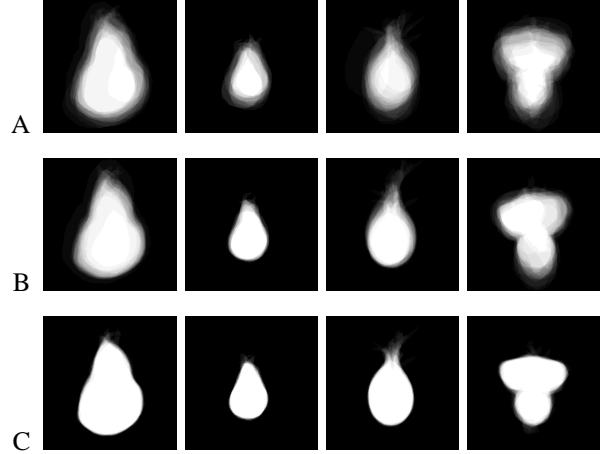
This learned shape template was used in conjunction with the simple filter-and-threshold algorithm to obtain better segmentation results automatically in cluttered images. The manually segmented shapes were also used as truth data for comparing the performance of our implementations of segmentation algorithms. Our current implementation gives satisfactory segmentation results (example: Figure 3).

4. Joint Nonparametric Shape Learning and Alignment

Once the shapes of the relevant disc tissues, $I^i_s(\boldsymbol{x})$ (Equation 2), are extracted in binary image format, we use this set of binary shapes to learn the canonical underlying shape model of the given class of disc tissues using a nonparametric learning algorithm called '*Congealing*' [12]. We denote the set of binary shape images as $\Phi_s \doteq \{I^i_s\}_{i=1}^{N}$ where N is the cardinality of the set. For a thorough discussion of this algorithm, the reader is referred to Miller et al. [12, 11].

Let us denote the latent binary shape of the given class of disc tissues as I_l. We model each shape image in Φ_s as I_l transformed through a geometric transformation. Given a class, the latent shape and the transformation are conditionally independent [11]. We assume that the transformations are affine and model the affine parameters as i.i.d. random variables. We shall assume that the transformation is a one-to-one and invertible mapping between I_l and I^i_s. We make the further assumption that the probability distribution of pixel values at each pixel location are *i.i.d.* Thus, for any
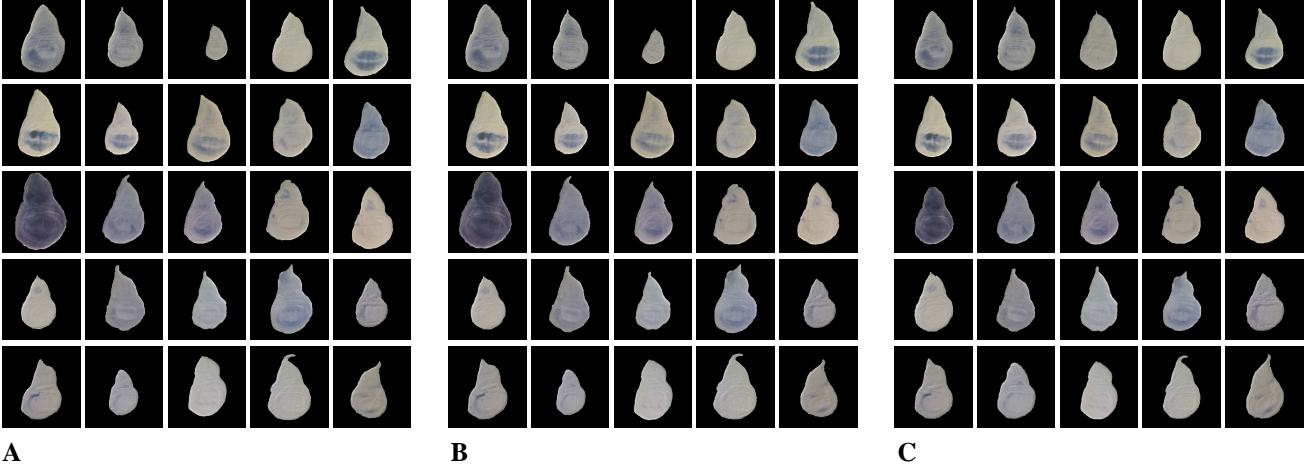
Figure 5. Alignment Results for wing discs. **A**: Segmented wing discs I_f^i *before* congealing. **B**: Segmented wing discs I_a^i *after* applying the transformations learned by congealing with only 3 parameters (t_x, t_y and θ). **C**: Segmented wing discs I_a^i *after* applying the transformations learned by congealing with 7 parameters (Equation 8). Note that the third image was taken with a 10x objective, while the rest were taken with a 20x objective. However, the algorithm was able to properly align the image as shown in **C**.

given pixel location $\boldsymbol{x} = \boldsymbol{x}^*$,

$$I_s^i(\boldsymbol{x}^*) = I_l(g_i^{-1}(\boldsymbol{x}^*)). \tag{4}$$

In our implementation, we parameterize the set of transformations g_i using the following component transformations: x-translation (t_x), y-translation (t_y), rotation (θ), x-log-scale (s_x), y-log-scale (s_y), x-shear (h_x), and y-shear (h_y). Clearly, this is an over-complete parameterization, but following the efficiency arguments presented by Miller [11] it allows faster convergence during the optimization routine. We experimented with different choices of parameterization, and we will show results based on the parameterization as shown in Equation 8.

Fixing the order of composition to ensure unique mapping (since the matrix multiplication is not commutative), this can be written as:

$$g = F(t_x, t_y, \theta, s_x, s_y, h_x, h_y) \tag{5}$$
$$g_i = F(\{v_j\}^i) \tag{6}$$
$$\{v_j\}^i = (t_x^i, t_y^i, \theta^i, s_x^i, s_y^i, h_x^i, h_y^i) \tag{7}$$

where $1 \leq i \leq N$, (N is the cardinality of the set Φ_s), $1 \leq j \leq K$, (K is the number of parameters chosen), and $v \in \mathbb{R}^{N \times K}$.

Writing g out explicitly, we get:

$$g = \begin{bmatrix} cos(\theta)e^{s_x} - sin(\theta)e^{s_y}h_y & cos(\theta)e^{s_x}h_x - sin(\theta)e^{s_y} & t_x \\ sin(\theta)e^{s_x} + cos(\theta)e^{s_y}h_y & sin(\theta)e^{s_x}h_x + cos(\theta)e^{s_y} & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

4.1. Shape Learning

The goal is to find the transformation g_i that converts a given I_s^i into the most likely form of I_l. Formulating this as the maximum likelihood estimation problem, and by searching through the affine parameter space as defined above, we want to find g_i that maximizes the probability $\hat{\Theta} = \arg\max_{g_i} P(g_i^{-1}|I_s^i)$. We can write $\hat{\Theta}$ as:

$$\hat{\Theta} = \arg\max_{g_i} P(I_l^i(I_s^i, g_i^{-1})). \tag{9}$$

Using our *i.i.d.* assumptions and Equations 4, 6 and 7,

$$\hat{\Theta} = \arg\max_v \prod_{\boldsymbol{x} \in \Omega} \prod_{v \in \mathbb{R}^{N \times K}} p_{\boldsymbol{x}}(I_s(g_i(\boldsymbol{x}))). \tag{10}$$

Taking log-probabilities, Equation 10 becomes:

$$\hat{\Theta} = \arg\max_v \sum_{\boldsymbol{x} \in \Omega} \sum_{v \in \mathbb{R}^{N \times K}} \log p_{\boldsymbol{x}}(I_s(g_i(\boldsymbol{x}))). \tag{11}$$

I_s^i is a binary image so the pixel stack at $\boldsymbol{x}$ would consist of 1's and 0's. $p_{\boldsymbol{x}}(I_s(g_i(\boldsymbol{x})))$ is the empirical probability of 1's in the pixel stack at $\boldsymbol{x}$ Since we model the transformations as the random variables causing $I_s^i(\boldsymbol{x})$ to vary from $I_l(\boldsymbol{x})$, we can see that these two will be the same when the randomness due to g_i is removed. We take entropy as our choice of estimator $\hat{\Theta}$ for randomness caused by g_i. This allows us to write:

$$\hat{\Theta} \approx \arg\min_v \sum_{\boldsymbol{x} \in \Omega} H(\boldsymbol{x}). \tag{12}$$

This approximation becomes equality when N is very large. Here, $H(\boldsymbol{x})$ is the Shannon entropy of the binary pixel
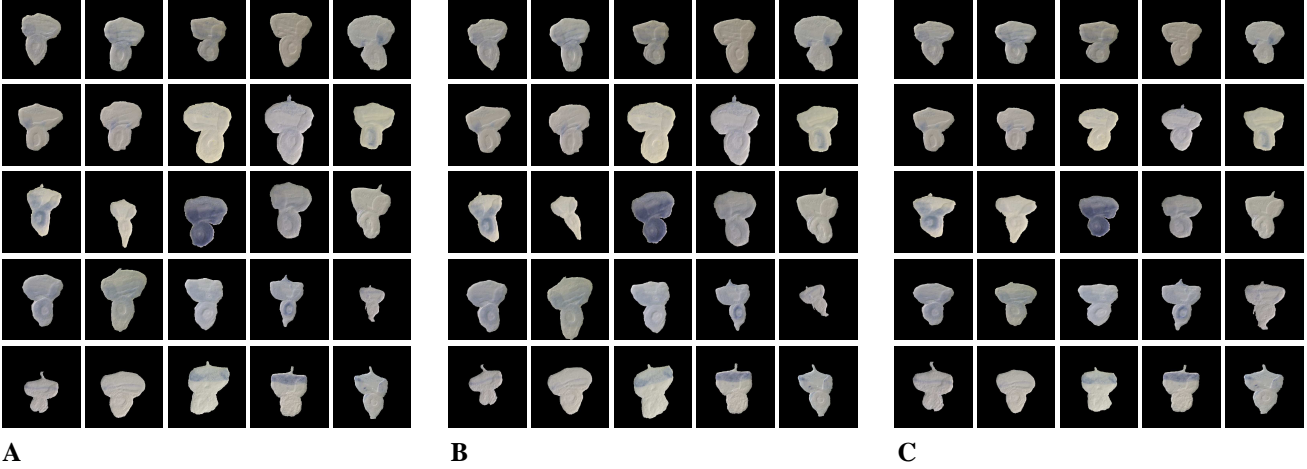
A   **B**   **C**

Figure 6. Alignment Results for eye discs. **A**: Segmented eye discs I_f^i *before* congealing. **B**: Segmented eye discs I_a^i *after* applying the transformations learned by congealing with only 3 parameters. **C**: Segmented eye discs I_a^i *after* applying the transformations learned by congealing with 7 parameters (Equation 8).

stack at $\boldsymbol{x}$. This ML estimation can be seen as solving an optimization problem with the objective function Ψ defined as

$$\Psi \doteq \sum_{\boldsymbol{x} \in \Omega} H(\boldsymbol{x}) + \sum_{i=1}^{N} |v^i| \qquad (13)$$

where v^i are the vectors of transformation parameters (Equation 7), and $|\cdot|$ is some norm (or penalty term or regularization term) on these vectors to keep the shape images from shrinking to zero or undergoing other extreme transformations. Ψ is called the penalized pixel-wise entropy [11].

The learning algorithm proceeds as follows:

1. Maintain a transform parameter vector v^i (Equation 7) for each shape image I_s^i. Each parameter vector will specify a transformation matrix $\hat{g}_i = F(v^i)$ according to Equation 8. Initialize all v^i to zero vectors. This has the effect of initializing all of the transformation matrices $\hat{g}_i$ to the identity matrix.

2. Compute the penalized pixel-wise entropy Ψ for the current set of images from Equation 13.

3. For each shape image I_s^i, repeat until convergence:

   (a) Calculate the numerical gradient $\bigtriangledown_{v^i} \Psi$ of Equation 13 with respect to the transformation parameters v_j^i's for the current image ($1 \leq j \leq K$).

   (b) Update v^i as: $v^i = v^i + \gamma \bigtriangledown_{v^i} \Psi$ (where the scaling factor $\gamma \in \mathbb{R}$).

   (c) Update γ (according to some reasonable update rule such as the Armijo rule.

At convergence of this optimization procedure, the set of shapes $\Phi_s = \{I_s^i\}_{i=1}^{N}$ are all aligned, and the associated transformations $\{g_i\}_{i=1}^{N}$ are computed. To visualize the entropy of the transformed image set for a class at each step of the optimization, one can construct an image (Figure 4) in which each pixel is the mean of its corresponding pixel stack.

4.2. Joint Alignment

We apply the $\{v_j\}^i$ learned from the congealing process to the extracted structures I_f^i to bring all the images into alignment in one step.

$$I_a^i(\boldsymbol{x}) = I_f^i(g_i(\boldsymbol{x})). \qquad (14)$$

where $1 \leq i \leq N$. We show our results in Figures 5 and 6.

5. Stain Scoring

Mass-isolated imaginal discs were placed in 96-well plates and stained with digoxigenin-labeled RNA or DNA complementary to genes of interest. Images were acquired using a light microscope equipped with Nomarski optics as described by Tomancak et al. [14]. The local presence of stain results in the appearance of blue in the image; darker blue suggests a greater local concentration of the gene of interest. However, there is substantial probe-to-probe variability and these intensities should not be relied on as an accurate quantitative measure of gene concentration. Nevertheless, the different intensity values can be used to suggest where local gene concentration is high.
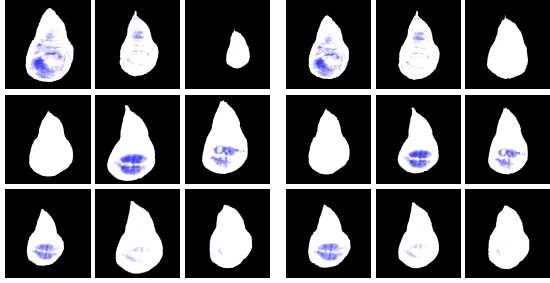
Figure 7. (This figure is better viewed in color) **Left**: Unaligned stain-scored wing disc images. **Right**: Aligned stain-scored wing disc images after congealing. Black pixels have been segmented as background, white indicates no stain and shades of blue indicate stained pixels calculated with the semi-quantitative stain scoring algorithm.

We developed a simple semi-quantitative metric that ranges from 0 to 5 where 0 indicates no expression of the gene of interest and 5 indicates high expression. Examples of segmented, stain-scored wing discs, both unadjusted and congealed, can be seen in Figure 7. Notice that the overall shape and size of the discs are more consistent in the congealed images and that a pixel-wise comparison of stain intensity of biologically similar patterns would appear more similar when comparing the congealed, stain-scored images than the unaligned stain-scored images.

6. Future Work

The proposed overall methodology shown in Figure 2 operates without making any assumptions about the underlying structure of a given tissue class. It is unsupervised and is highly amenable to large scale spatial expression analysis. It augments any model-based registration methods one may choose to apply by supplying the nonparametrically learned canonical structure model from the given ensemble of images for a given tissue class. We implemented and demonstrated the applicability of this methodology using Drosophila imaginal discs. To make our approach more general, we are currently investigating the possibility of incorporating more refined segmentation algorithms into our approach. We plan to perform detailed comparative analysis of spatial patterns of gene expression in aligned imaginal discs using pixel-wise comparisons. We also plan to extend this approach to three-dimensional datasets.

Acknowledgments

References

[1] A. A. Alizadeh et al. Distinct types of diffuse large b-cell lymphoma identified by gene expression profiling. *Nature*, 403:503–511, 2000.

[2] M. N. Arbeitman et al. *Science*, 297:2270–2275, 2002.

[3] M. Bate and A. M. Arias. *The development of Drosophila melanogaster*, volume II. 1993.

[4] B. Berman et al. Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the drosophila genome. *Proc Natl Acad Sci U S A*, 99(2):757–62, 2002.

[5] M. J. Butler et al. Discovery of genes with highly restricted expression patterns in the Drosophila wing disc using DNA oligonucleotide microarrays. *Development*, 130:659–670, 2003.

[6] G. Charpiat, O. Faugeras, and R. Keriven. Shape metrics, warping and statistics. In *Proceedings of the International Conference on Image Processing*, 2003.

[7] B. Frey and N. Jojic. Learning mixture models of images and inferring spatial transformations using the em algorithm. In *Proceedings of the IEEE Conference Computer Vision and Pattern Recognition*, 1999.

[8] A. Klebes, B. Biehs, F. Cifuentes, and T. Kornberg. Expression profiling of drosophila imaginal discs. *Genome Biol*, 3(8):RESEARCH0038, 2002.

[9] S. Kumar et al. BEST: a novel computational approach for comparing gene expression patterns from early stages of drosophila melanogaster development. *Genetics*, 162:2037–2047, 2002.

[10] R. Lipshutz, S. Fodor, T. Gingeras, and D. Lockhart. High density synthetic oligonucleotide arrays. *Nat Genet*, 21(1 Suppl):20–4, 1999.

[11] E. G. Miller. Learning from one example in machine vision by sharing probability densities. *PhD Dissertation*, Massachusetts Institute of Technology, 2002.

[12] E. G. Miller, N. Matsakis, and P. A. Viola. Learning from one example through shared densities on transforms. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2000.

[13] H. Peng and E. W. Myers. Comparing in situ mrna expression patterns of drosophila embryos. In *Proceedings of the Eighth Annual International Conference on Research in Computational Molecular Biology*, 2004.

[14] P. Tomancak et al. Systematic determination of patterns of gene expression during drosophila embryogenesis. *Genome Biology*, 3(12):1–14, 2002.

[15] A. Tsai et al. A shape-based approach to curve evolution for segmentation of medical imagery. *IEEE Transactions on Medical Imaging*, 22, No. 2:137–154, February 2003.