

# PointCloudXplore: Visual Analysis of 3D Gene Expression Data Using Physical Views and Parallel Coordinates

O. Rübél<sup>1</sup>, G.H. Weber<sup>2</sup>, S.V.E. Keränen<sup>3</sup>, C.C. Fowlkes<sup>4</sup>, C.L. Luengo Hendriks<sup>3</sup>, Lisa Simirenko<sup>3</sup>, N.Y. Shah<sup>2</sup>, M.B. Eisen<sup>3</sup>, M.D. Biggin<sup>3</sup>, H. Hagen<sup>1</sup>, D. Sudar<sup>3</sup>, J. Malik<sup>4</sup>, D.W. Knowles<sup>3</sup> and B. Hamann<sup>2</sup>

<sup>1</sup> International Research Training Group “Visualization of Large and Unstructured Data Sets,” University of Kaiserslautern, Germany

<sup>2</sup> Institute for Data Analysis and Visualization, University of California, Davis, CA, USA

<sup>3</sup> Life Sciences and Genomics Divisions, Lawrence Berkeley National Laboratory, CA, USA

<sup>4</sup> Computer Science Division, University of California, Berkeley, CA, USA

---

## Abstract

To allow a more rigorous understanding of animal gene regulatory networks, the Berkeley *Drosophila* Transcription Network Project (BDTNP) has developed a suite of methods that support quantitative, computational analysis of three-dimensional (3D) gene expression patterns with cellular resolution in early *Drosophila* embryos. Here we report the first components of a visualization tool, PointCloudXplore, that allows the relationships between different gene’s expression to be analyzed using the BDTNP’s datasets.

PointCloudXplore uses the established visualization techniques of multiple views, brushing, and linking to support the analysis of high-dimensional datasets that describe many genes’ expression. Each of the views in PointCloudXplore shows a different gene expression data property. Brushing is used to select and emphasize data associated with defined subsets of embryo cells within a view. Linking is used to show in additional views the expression data for a group of cells that have first been highlighted as a brush in a single view, allowing further data subset properties to be determined. In PointCloudXplore, physical views of the data are linked to parallel coordinates. Physical views show the spatial relationships between different genes’ expression patterns within the embryo. Parallel coordinates, on the other hand, show only some features of each gene’s expression, but allow simultaneous analysis of data for many more genes than would be possible in a physical view. We have developed several extensions to standard parallel coordinates to facilitate brushing the visualization of 3D gene expression data.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computing Methodologies/Computer Graphics]: Picture/Image Generation J.3 [Computer Applications/Life and Medical Sciences]: Biology and Genetics

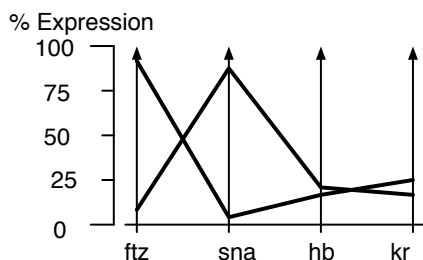
---

## 1. Introduction

Animal embryos comprise dynamic 3D arrays of cells that express gene products in intricate spatial and temporal patterns that determine the shape of the developing animal. Biologists have typically analyzed gene expression and morphology by visual inspection of photomicrographic images. Yet, to understand animal development, we need good methods to computationally describe gene expression data. To address this challenge, the BDTNP has developed image analysis methods to extract information about gene expression from imaging data, using early *Drosophila melanogaster* embryos as a model. Confocal image stacks of blastoderm stage *Drosophila* embryos are converted into matrices specifying the position of nuclei and the expression levels of select genes around each nucleus, (see Section 3). The resulting new datasets, termed PointClouds, promise to be an in-

valuable resource for studying development. Since available visualization tools are insufficient for comparing and analyzing 3D PointCloud data, we have developed PointCloudXplore as a tool to help biologists explore these datasets.

During embryogenesis complex regulatory networks are built up as transcription factors cross-regulate the expression of other transcription factors as well as enzymes, structural proteins, etc., guiding the development of animals [WKS\*03, Law92, SL05]. Since spatial regulation of gene expression directs animal morphogenesis, a major goal of the BDTNP is to decipher how spatial patterns of target gene expression are directed by the expression patterns of the transcription factors that regulate them. Because gene regulation depends on combinatorial inputs from many transcription factors, simultaneous analysis of many expression patterns is required. Therefore, we have included within



**Figure 1:** Using parallel coordinates to visualize imaginary expression levels of four genes

PointCloudXplore multiple visualization methods to allow specific relationships to be seen within highly complex expression data for many genes.

## 2. Previous Work

Generally, data can be displayed in multiple formats, or *views*, that each allow different relationships between the data components to be observed. The *linking* of multiple views is a well-established visualization method [BMMS91]. For example, it has been shown that linking abstract data displays, such as scatter plots, with physical data views, such as a 3D model of a catalytic converter, can improve data analysis significantly and provide insight into complex physical phenomena [KSH04, PKH04, DGH03]. Hauser et al. [HLD02] proposed integrating parallel coordinates with physical views for a better understanding of high-dimensional phenomena. Gresh et al. [GRW\*00] used parallel coordinates linked to physical views for visualizing biological data sets describing cardiac measurement and simulation experiments.

Parallel coordinates were proposed contemporaneously by Inselberg [Ins84] and Wegman [Weg90] and are a common information visualization technique for high-dimensional data sets. In a parallel coordinate view, a data set consists of a set of *samples*, which in our case are the cells in a *Drosophila* embryo. Each sample (cell) has a set of associated quantities, which in our case are the relative expression levels for multiple genes. Expression data for each gene corresponds to a dimension in the data set, with data for each gene being represented by one of a series of parallel vertical axes (see Figure 1). Each sample (cell) defines a *data line*, i.e., a zigzag line connecting adjacent parallel axes. The intersection point of the data line with each vertical axis corresponds to the value of the sample for the corresponding dimension (i.e., the relative expression level for the corresponding gene in that cell). Figure 1 illustrates parallel coordinates by using them to visualize imaginary expression levels for four genes (*fushi tarazu* (*ftz*), *snail* (*sna*), *hunchback* (*hb*), and *kruppel* (*kr*)) in two cells. In the example, one cell expresses *ftz* at 95%, *sna* at 5%, *hb* at 24% and *kr* at 26%; and the other cell expresses *ftz* at 10%, *sna* at 90%, *hb* at 25% and *kr* at 26% of maximum relative value.

Many extensions to standard parallel coordinates have been developed to make them more useful. Fua et al. [FWR99] proposed hierarchical parallel coordinates, including several techniques for visualization of selected subsets of the data. Distortion operations, such as dimensional zooming, for example, support a more detailed analysis of data subspaces. Color is widely used for improving parallel-coordinate views since dedicated line coloring eases following the course of data lines. Fua et al. [FWR99] and Novotny [Nov04] proposed the usage of color bands for visualization of brushes in parallel coordinate views. Wegman and Luor [WL97] proposed the usage of transparency and “over-plotting” translucent data points/lines. This method highlights dense areas while sparse areas fade away, thus revealing inherent data characteristics.

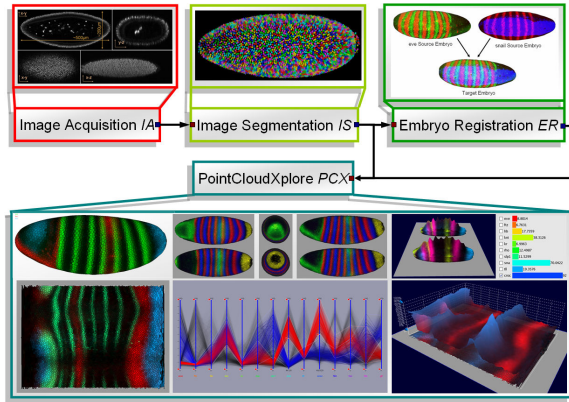
## 3. Gene Expression Data and Visualization Pipeline

A single PointCloud file contains the  $x,y,z$  location of each nucleus in one embryo and the relative concentration of gene products (mRNA or protein) associated to each nucleus [FLHK\*05]. These files are created in following manner (see Figure 2). Embryos are fixed, stained, and mounted, then imaged using a confocal microscope (Figure 2, *IA*). The obtained images are processed to detect all nuclei and measure the associated gene expression levels (Figure 2, *IS*). Embryos are typically labeled with one fluorophore to detect the nuclei, and with two others to detect two gene products. It is not practical to obtain the expression of more than a few genes in a single embryo, due to the limited number of different fluorophores that can be distinguished by the microscope, as well as the difficulty in adding these labels to the embryos. Because it is critical to compare the relationships between transcription factors and many of their target genes in a common co-ordinate framework, a set of PointClouds using both morphology and a common reference gene to determine correspondences (Figure 2, *ER*) are registered into Virtual PointClouds [FLHK\*05]. The resulting Virtual PointCloud contains averaged expression levels for many genes mapped on the nuclei of one of the embryos in the set. PointCloudXplore (Figure 2, *PCX*) can be used for visualization of both single-embryo PointClouds and Virtual PointClouds.

## 4. Visualization of Gene Expression in Physical Space

### 4.1. Multiple Embryo Views

We have developed several physical views (models) of the embryo to allow analysis of spatial gene expression patterns. In all these Embryo Views, each cell is represented by one 3D graphical object, positioned in space according to the physical position of the cell it represents. Gene expression values are visualized using color and, in the case of views using gene Expression Surfaces (Section 4.7), also by height. During the developmental stage that the BDTNP is



**Figure 2:** Gene Expression Data and Visualization Pipeline showing that PointCloudXplore is used to visualize data from single embryo PointClouds and Virtual PointClouds.

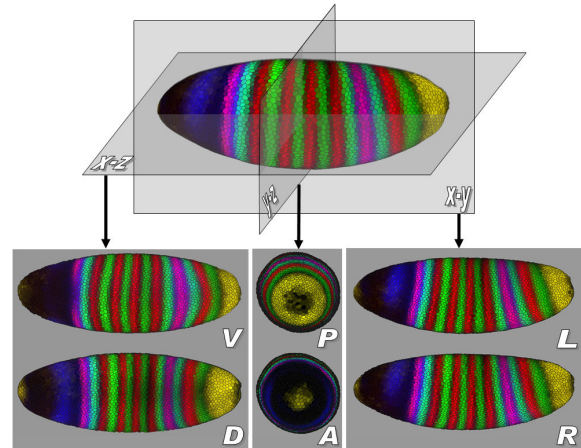
currently investigating — the blastoderm — all cells studied by the BDTNP are located on a roughly ellipsoid shell. Orthographic projection is used to display views from fixed angles that display expression and morphology along the three coordinate axes (Section 4.5). In another view, a cylindrical projection is used to map cells onto a plane to gain a global overview of the entire embryo (Section 4.6). Cells of interest can be selected in any of these views to create a so called *brush*, just by drawing on the surface of the embryo. The selected cells that comprise the brush are highlighted using color. The user can interact with all embryo views via interactive zooming, panning, and rotation.

#### 4.2. Representing Cells

In all our Embryo Views, each cell is represented by a polygon on the embryo surface, using the Eigencrust method [KSO04] for constructing an approximation of the Delauny triangulation of the surface of a PointCloud. The dual mesh of the triangulation is an approximation of the Voronoi diagram defined on the embryo surface [dB-vKOS00]. Each cell is represented by a Voronoi polygon with exactly one original data point in its center. The polygon sizes depend on the cells' distribution in the embryo, whereas the shape of the polygons has no direct meaning.

#### 4.3. Representing Expression with Color

Each polygon is colored according to expression values measured in the cell it represents. The color mapping is based on the HSV color model [FDFH95]. The basic color hue, H, of each gene is defined by the user. Gene expression values are mapped linearly to color brightness V. Saturation of color is always one, unless specified differently. For each gene a minimum and maximum value can be defined independently by the user. All expression values below the minimum are mapped to black, and all values above the maximum are



**Figure 3:** Orthographic Views showing expression patterns of four genes: *eve* (red), *ftz* (green), *Giant* (*gt*) (blue) and *fork head* (*fkh*) (yellow). The projections in the lower panels show the ventral (V), dorsal (D), anterior (A), posterior (P), left (L) and right (R) views.

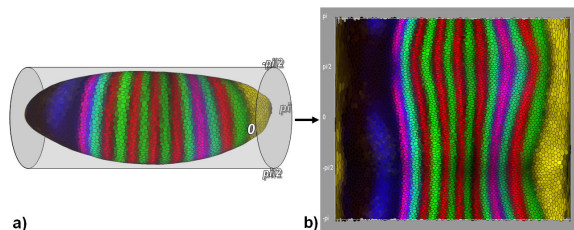
mapped to full intensity. Thus for a gene with a selected color of red, its expression will be displayed as bright red in cells in which expression is highest, dimmer red in cells in which expression is moderate, and black in cells in which the gene is not expressed.

#### 4.4. Defining Embryo Orientation and Position

To guarantee correct and accurate projection, the embryo must be aligned with the major coordinate system. Therefore, principal component analysis (PCA) is used to calculate the approximate degree and direction of the main directions — body axes — of the embryo. The PointCloud is then translated into the point of origin of the world coordinate system using the mean point of the PointCloud as reference point. Subsequently, the eigenvector of the largest eigenvalue is aligned with the first coordinate vector (1,0,0) via rotation of the PointCloud around the main coordinate axis. The eigenvectors of the second and third eigenvalues are then aligned with the second (0,1,0) and third (0,0,1) coordinate vectors respectively. In Figure 3, the upper embryo is oriented with anterior left, posterior right, ventral down, and dorsal up. The head of the fly will develop in the anterior, and the back of the fly on the dorsal side.

#### 4.5. Orthographic Views

After correction of position and orientation of the embryo, it is possible to use the three coordinate system planes as projection planes. To create the Orthographic Views, all cells are projected either onto the  $xy$ ,  $xz$  or  $yz$  plane using orthographic projection. Since the point of origin is in the center of the PointCloud, and since the main coordinate vectors are aligned with the PCA eigenvectors of the PointCloud, the



**Figure 4:** (a) shows a representation of a 3D embryo surrounded by a cylindrical projection. (b) shows an Unrolled View derived from the cylindrical projection. The genes whose expression patterns are displayed are the same as those in Figure 3.

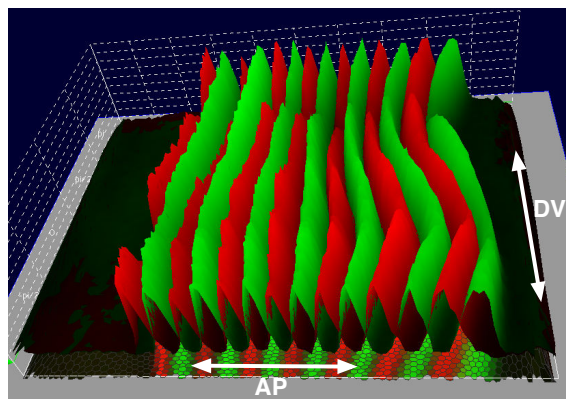
embryo is divided into two parts, depending on which side of the projection plane the cells are. The two sides of the embryo are separated, and in this way 2D views of the dorsal/ventral, anterior/posterior, and the left/right sides of the embryo are created (Figure 3). These Orthographic Views provide an overview of all cells while preserving the shape of the embryo as a frame of reference for a biologist. The curvature of the embryo leads to high densities of cells on the projection borders, but gives an impression of depth and shape. A general overview of all cells in an embryo can be obtained by switching between the three different Orthographic Views.

#### 4.6. Unrolled View

For a scientist who wants an instant overview of the whole blastoderm expression pattern, the Unrolled View uses a cylindrical projection to map all surface cells of the embryo onto a rectangular plane (see Figure 4). All cells are first projected onto a cylinder, which is unrolled in a plane. In this view, a complete overview of all cells is provided while the relative positions of cells on the anterior/posterior axis and around the embryo are preserved. Because of the ellipsoid shape of the embryo, cells in the anterior and posterior of the embryo are distorted by the projection in order to fill the rectangular plane. Shape and size of cells in the middle part of the embryo are less affected by distortion effects.

#### 4.7. Expression Surfaces

To allow a more quantitative analysis of gene expression data, Expression Surfaces can be defined above either the Orthographic or the Unrolled Views. Each Expression Surface displays data for one gene. The  $xy$  positions of Expression Surface points are determined by the positions of cells in the underlying views, whereas the height of an Expression Surface is determined by the expression values measured for the gene it represents. Spatial relationships between several genes' expression patterns can be viewed at once using multiple Expression Surfaces. For example, Figure 5 shows the quantitative relationship between *eve* (red) and *ftz* (green)



**Figure 5:** Gene expression surfaces for *eve* (red) and *ftz* (green) colored according to these two genes' expression values.

expression stripes, showing that these two genes' expression levels are spatially largely non-overlapping and change relative to one another along each body axis.

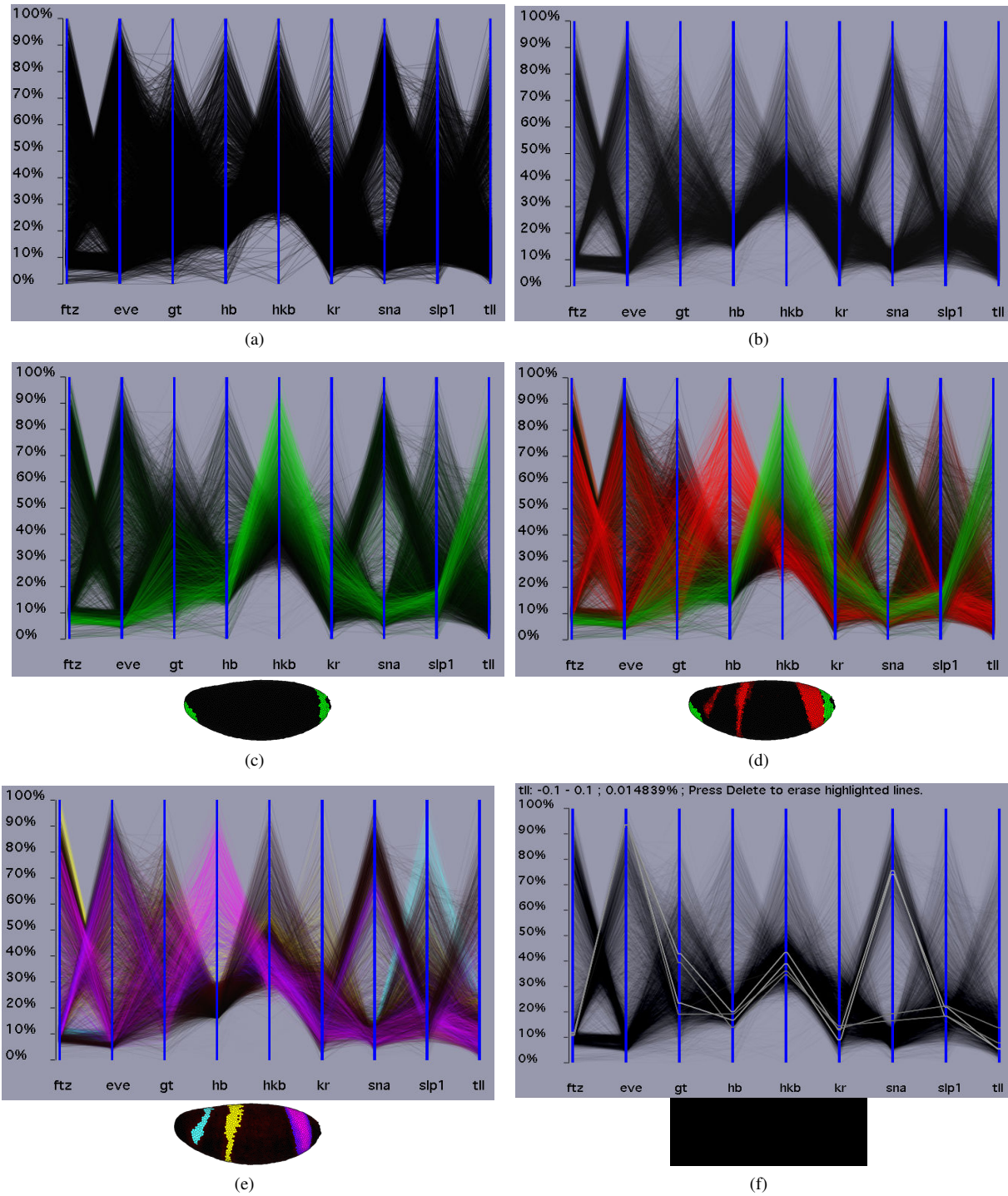
A variety of options can be used to improve the view, including different coloring strategies, or changing Expression Surface transparency or intensity values. An Expression Surface can, for example, be the same color as the gene in the 2D plot, but with an intensity that varies with gene expression level to allow comparison of expression patterns using both color and height in parallel (Figure 5). As in all other embryo views, the user can interact with the view via interactive zooming, panning and rotation.

### 5. Visualization in Gene Expression Space Using Parallel Coordinates

A limitation of all of the above Embryo Views is that when more than four or five genes whose expression overlaps are displayed at the same time, it is often not possible to distinguish each gene's expression. Therefore, we have adapted parallel coordinates to create several Parallel Coordinate Views, in which relationships between many genes' expression can be visualized (see Section 2 for an introduction to parallel coordinates). Further, we have linked Parallel Coordinate Views and Embryo Views, ensuring that all brushes defined in the Embryo Views can be displayed in the Parallel Coordinate Views and vice versa. In general, parallel coordinates suffer from several problems, such as occlusion and cluttering. To reduce these problems, several extensions to standard parallel coordinates have been developed, Sections 5.1 to 5.4.

#### 5.1. Varying Line Properties

Monochromatic data line drawings can lead to ambiguous interpretation of parallel coordinates when datasets contain as many samples as ours, since it is nearly impossible to trace



**Figure 6:** 2D Parallel Coordinate View showing expression data for nine genes. (a) A view in which data lines have maximum opacity. (b) A view in which data line transparency has been increased. (c) A view in which color from an Embryo View showing hkb expression (green) is shown. (d) A view in which colors from an Embryo view showing both hkb (green) and hb (orange) expression are shown. (e) A view showing two brushes drawn on an Embryo View to highlight distinguish between the three hb stripes of expression (light blue, yellow and pink). (f) A view showing line traces that highlight data associated with several cells.

a single data line through the graph (Figure 6(a)). Therefore, we have employed a number of line property management features to improve visualization.

By increasing line transparency, the user can improve the view by allowing only the denser clusters present in the data to be seen (compare Figure 6, (a) with (b)). The complex genetic regulatory networks guiding embryo development often cause the dimensions of gene expression data to be highly correlated. For example, in Figure 6(b), left two axes, it can be seen that in many cells in the embryo *eve* and *ftz* show an inverse correlation in gene expression because cells that express *eve* at high levels express *ftz* at low levels and vice versa. In most of the remaining cells, both genes are expressed at low levels, all consistent with the known expression patterns of these two genes.

To further improve visualization of the relationships between different gene's expression, particularly for genes displayed on non-adjacent parallel coordinate axes, we have developed four data line coloring schemes. In the first, colors defined in the Embryo Views can also be used in Parallel Coordinate Views. For example, in Figure 6(c) cells expressing *huckebein* (*hkb*) at high levels are shown in green in both views. In the Parallel Coordinate View, it can be seen that *hkb* expression is anticorrelated in most cells with all genes except *tailless* (*tll*), consistent with the fact that both *tll* and *hkb* are highly expressed at the posterior pole of the embryos, whereas the other genes shown are not. Addition of a second color to mark cells expressing *hb* at high levels (Figure 6(c), orange) allows additional relationships to be observed, such as the fact that cells that express both *eve* and *ftz* at low levels do not express *hb* at high levels but many do express *hkb* strongly.

Additional color options available in PointCloudXplore are: 1. A brush coloring scheme, in which a brush drawn on an embryo view can be used to highlight a set of cells in parallel coordinates with a user defined color (Figure 6e). 2. A cell position coloring scheme, in which the positions of cells along either the anterior/posterior axis or the dorsal/ventral circumference of the embryo are used for coloring data lines, providing a guide to the relative physical locations of cells. 3. A brush coloring scheme in which the distance to the average expression value of the gene displayed on the first parallel axis is displayed.

In all of these coloring schemes, cells of no interest are colored black. To further enhance visualization, the user can define an independent transparency value for black data lines. Increasing the transparency of these lines highlights the selected data lines, while the surrounding context is preserved (see Figures 6-7).

## 5.2. Line Trace Highlighting and Animation

The above data line transparency and coloring options improve the view, but it can still be hard to follow the course

of a single line through the graph. Line trace highlighting overcomes this problem. The user can select data lines by selecting a point, in other words a gene expression level, on a parallel coordinate axis. All data lines passing the region defined by this point are drawn in front of all other lines using a gray scale. In addition, the selected range and the percentage of highlighted lines are displayed (Figure 6(f)). Moving the cursor along a parallel coordinate axis reveals information about the data line density and line distribution along that axes.

Animation of line traces can help the user to follow the course of data lines as well. Line trace animation means that a user-defined subset of data lines is drawn in an animation process beginning from the first parallel axis to the last. To define a subset of data lines to be animated, the user needs to define both the number of lines to be traced and a data range on the first parallel axis. Data lines are then selected automatically, being roughly equidistant in the user-defined range. It is also possible to animate only lines selected by a user-defined brush.

## 5.3. Extending Parallel Coordinate View to 3D

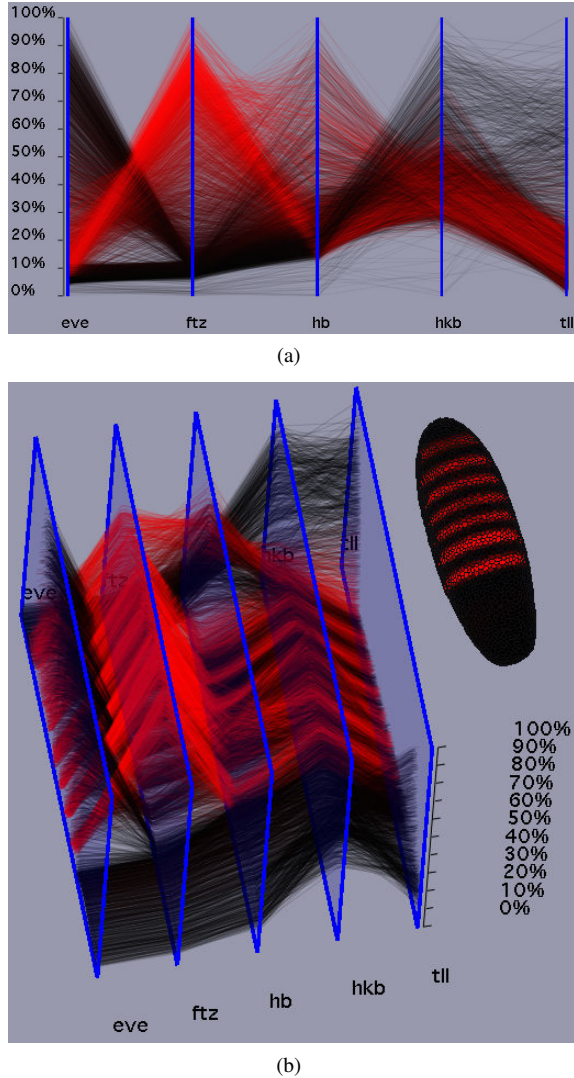
Information about the spatial relationships between different genes' expression patterns is essential for the analysis of regulatory networks. Information about relative cell positions along the two main axes of the embryo, anterior/posterior (AP) and dorsal/ventral (DV), can be derived from the Unrolled View described in Section 4.6. To display this information in Parallel Coordinate Views, the coordinate axes have been extruded into the third dimension (Figure 7b). Data lines are ordered from back-to-front according to cell positions along either the AP axis or the DV circumference. Along any given data line, the positional information is constant, such that data lines do not intersect each other in this third dimension. The 3D coordinate axes are drawn highly transparent with active z-buffering to prevent the addition of colors of overlapping parallel axes. This strategy guarantees a complete overview of the entire plot, with no details hidden. In addition, the outer frame of the coordinate axes are drawn with full opacity, which makes it easier to determine the position of the coordinate axis in 3D space.

By using this 3D visualization, spatial data dimensions are clearly separated from gene expression dimensions of the data and the basic character of the spatial gene expression patterns in one dimension is preserved. For example, if data lines are sorted according to the position of cells along the AP axis, then the stripe patterns of genes like *eve* or *ftz* are visible in the plot (Figure 7b). This 3D view also reveals what was not obvious in the 2D views shown in Figure 6 and Figure 7a: that cells expressing both *eve* and *ftz* at low levels are mainly at the anterior and posterior of the embryo and that a subset of these cells are in the principal areas where *hkb* and *tll* are highly expressed. Even if tens of additional

gene dimensions were to be added to the 3D view, these and doubtless other relationships could still be visualized.

#### 5.4. Brushing in Parallel Coordinate Views

The core concept of PointCloudXplore is to present data in multiple views that are linked together by brushing to select and highlight data associated with subsets of cells. In our Parallel Coordinates View, brushing can be performed by defining data ranges in each of the different data dimensions, including spatial dimension displayed in the 3D Parallel



**Figure 7:** 2D (a) and 3D (b) Parallel Coordinate Views comparing expression of five genes. The coloring of the expression levels of the *ftz* gene in an embryo view has been used to highlight *ftz* expressing cells in the parallel coordinates. In the 3D view, the anterior/posterior axis runs from top down, with anterior being at the bottom

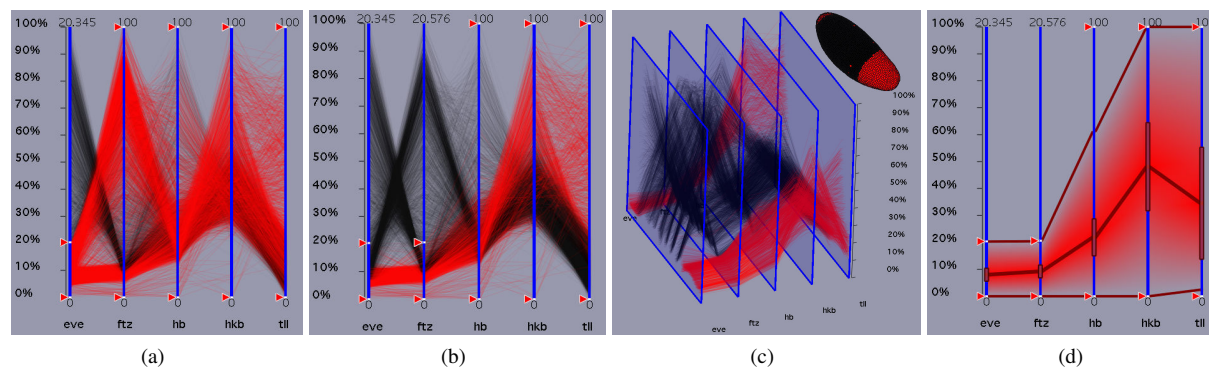
Coordinate View, using two sliders attached to each axis. For example, in Figure 8a cells expressing *eve* at low levels between two threshold markers are highlighted. By refining this brush to exclude cells expressing *ftz* at high levels (Figure 8b), it can be seen that cells that express *tll* at high levels do not express either *eve* or *ftz* at high levels. A 3D parallel coordinate view makes this clearer and shows the locations of these cells along the A/P axis (Figure 8c).

To facilitate brushing and to provide a general overview of basic brush properties, it is also possible to display brushes as just broad color bands (see Figure 8d). The width of each color band depends on the minimum and maximum data values selected by the brush. The color of the brush bands is defined by the user, whereas the transparency value of brush bands is defined automatically according to the distance to the average value of the brush in the different dimensions. The average gene expression values of the selected cells for each gene define the center of the brush band on each coordinate axis and are shown by maximum opacity, while transparency of the band is increased towards its edges. To avoid occlusion, the maximum opacity of brush bands is automatically lowered if several such brush bands are displayed, each highlighting different subsets of cells. The minimum-, maximum-, and average values of a brush can be displayed as an additional polyline, and the standard deviations of a brush can be displayed as rectangular boxes on the coordinate axes (Figure 8d).

Brushes define a subspace of the data across each parallel coordinate axis. To support exploration of the details in this subspace, each of the dimensions is scaled independently according to the minimum and maximum selected values. After scaling, the subspace within the brush fills the display area, and the details are revealed. This distortion technique is in general referred to as dimensional zooming. Both the use of brush bands and dimensional zooming were first proposed by Ying-Huey Fua et al. [FWR99]. We have extended this technique to 3D parallel coordinates. In PointCloudXplore, parallel coordinates can be displayed in three different viewers where one viewer always displays the standard 2D front view, a second one the 3D view, and a third one the 3D dimensional zoom. The three different views are displayed in a split view window. Via movement of the split bars, views can be resized or completely hidden if desired. As the overall context is always preserved in the other views, this strategy ensures that one does not get lost in details in some small subspace displayed in the dimensional zoom.

#### 6. Conclusions and Future Work

We have introduced PointCloudXplore and described a subset of its views and functionality. Dedicated physical views of the *Drosophila melanogaster* blastoderm (termed Embryo Views) allow the comparison and analysis of spatial gene expression patterns. Expression Surfaces provide an effective and intuitive way for quantitative analysis of gene



**Figure 8:** Defining brushes in Parallel Coordinate View. In (a) a brush is defined to exclude all cells expressing *eve* at more than 20%. In (b) this brush is further refined to also exclude all cells expressing *ftz* at greater than 20%. (c) shows a 3D Parallel Coordinate View of the brush defined in (b) in which location along the A/P axis of the embryo is shown in the 3D dimension. (d) Shows a broad color band display of the brush selected in (b), indicating the minimum, maximum, mean, and standard deviations for expression values for each gene.

expression data. To allow analysis of the relationships between genes directly in gene expression space, we have integrated parallel coordinates into the system (Parallel Coordinate Views). Several well-known extensions to standard parallel coordinates have been included. Parallel Coordinate Views have been extended to a 3D rendering, making it possible to present spatial and gene expression data dimensions in one plot, while both dimension types are visually separated and basic spatial properties of gene expression patterns are preserved. Dedicated line coloring schemes, line trace highlighting, and animation make it easy to follow the topology of data lines in parallel coordinates. All views are linked via the concept of brushing. PointCloudXplore makes interactive analysis of 3D expression data possible for the first time in this specific area of research.

Future work will concentrate on using additional information visualization techniques, such as scatter plots, for visualizing relationships in gene expression levels, and on integrating automatic data analysis tools into PointCloudXplore. Unsupervised clustering has been used previously to analyze microarray data and can also be applied to 3D gene expression data. Singular value decomposition (SVD) and other techniques have also been used to analyze gene expression and similar data. For analysis of 3D gene expression data, these existing techniques need to be modified, and new ones developed. Integration of such tools will further improve the utility of PointCloudXplore.

### Acknowledgments

This work was supported by an National Institutes of Health grant: GM70444; National Science Foundation awards: ACI 9624034 (CAREER Award), through the Large Scientific and Software Data Set Visualization (LSSDSV) program under contract ACI 9982251, a large Information Technology Research (ITR) grant; and by the LBNL Laboratory Di-

rected Research Development (LDRD) program; We thank the members of the Visualization and Computer Graphics Research Group at the Institute for Data Analysis and Visualization (IDAV) at the University of California, Davis; the members of the BDTNP at the Lawrence Berkeley National Laboratory (LBNL) and the members of the Visualization Group at LBNL.

### References

- [BMMS91] BUJA A., McDONALD J. A., MICHALAK J., STUETZLE W.: Interactive data visualization using focusing and linking. In *IEEE Visualization 1991* (1991), IEEE Computer Society Press, pp. 156–163.
- [dBvKOS00] DE BERG M., VAN KREVELD M., OVERMARS M., SCHWARZKOPF O.: *Computational Geometry*, second edition ed. Springer, Feb. 2000.
- [DGH03] DOLEISCH H., GASSER M., HAUSER H.: Interactive feature specification for focus+context visualization of complex simulation data. In *Data Visualization 2003 (Proceedings of the EUROGRAPHICS - IEEE TCVG Symposium on Visualization 2003)* (2003), Eurographics Association, pp. 239–248.
- [FDFH95] FOLEY J. D., DAM A. V., FEINER S. K., HUGHES J. F.: *Computer Graphics: Principle and Practice*, second edition in c ed. Addison-Wesley, July 1995.
- [FLHK\*05] FOWLKES C., LUENGO HENDRIKS C. L., KERÄNEN S. V. E., BIGGIN M. D., KNOWLES D. W., SUDAR D., MALIK J.: Registering drosophila embryos at cellular resolution to build a quantitative 3d map of gene expression patterns and morphology. In *CSB 2005 Workshop on BioImage Data Mining and Informatics* (Aug. 2005).
- [FWR99] FUA Y.-H., WARD M. O., RUNDENSTEINER E. A.: Hierarchical parallel coordinates for exploration of



- large datasets. In *IEEE Visualization 1999* (1999), IEEE Computer Society Press, pp. 43–50.
- [GRW\*00] GRESH D. L., ROGOWITZ B. E., WINSLOW R. L., SCOLLAN D. F., YUNG C. K.: WEAVE: A system for visually linking 3-d and statistical visualizations, applied to cardiac simulation and measurement data. In *IEEE Visualization 2000* (2000), IEEE Computer Society Press, pp. 489–492.
- [HLD02] HAUSER H., LEDERMANN F., DOLEISCH H.: Angular brushing of extended parallel coordinates. In *IEEE Symposium on Information Visualization (InfoVis'02)* (2002), IEEE Computer Society Press, pp. 127–130.
- [Ins84] INSELBERG A.: Parallel coordinates for multi-dimensional displays. In *Spatial Information Technologies for Remote Sensing Today and Tomorrow, The Ninth William T. Pecora Memorial Remote Sensing Symposium* (1984), IEEE Computer Society Press, pp. 312–324.
- [KSH04] KOSARA R., SAHLING G. N., HAUSER H.: Linking scientific and information visualization with interactive 3d scatterplots. In *Short Communication Papers Proceedings of the 12th International Conference in Central Europe on Computer Graphics, Visualization, and Computer Vision (WSCG)* (2004), pp. 133–140.
- [KSO04] KOLLURI R., SHEWCHUK J. R., O'BRIEN J. F.: Spectral surface reconstruction from noisy point clouds. In *Symposium on Geometry Processing* (July 2004), ACM Press, pp. 11–21.
- [Law92] LAWRENCE P. A.: *The Making of a Fly: The Genetics of Animal Design*. Blackwell Science, 1992.
- [Nov04] NOVOTNY M.: Visually effective information visualization of large data. In *Proceedings of Central European Seminar on Computer Graphics (CESCG)* (2004). Online at <http://www.cg.tuwien.ac.at/studentwork/CESCG/CESCG-2004/>.
- [PKH04] PIRINGER H., KOSARA R., HAUSER H.: Interactive focus+context visualization with linked 2d/3d scatterplots. In *Proceedings of the 2nd International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2004)* (July 2004).
- [SL05] STATHOPOULOS A., LEVINE M.: Genomic regulatory networks and animal development. *Developmental Cell* 9, 4 (2005), 449–462.
- [Weg90] WEGMAN E. J.: Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association* 85, 411 (Sept. 1990), 664–675.
- [WKS\*03] WEIGMANN K., KLAPPER R., STRASSER T., RICKERT C., TECHNAU G. M., JÄCKLE H., JANNING W., KLÄMBT C.: Flymove - a new way to look at development of drosophila. *Trends in Genetics* 19 (September 2003), 310–311. Online at <http://flymove.uni-muenster.de/>.
- [WL97] WEGMAN E. J., LUO Q.: High dimensional clustering using parallel coordinates and the grand tour. *Computing Science and Statistics* 28 (1997), 361–368.