Adaptation of video game UVW mapping to 3D visualization of gene expression patterns

Peter D. Vize* and Victor E. Gerth
Departments of Biological Science and Computer Science, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N1N4, Canada

ABSTRACT

Analysis of gene expression patterns within an organism plays a critical role in associating genes with biological processes in both health and disease. During embryonic development the analysis and comparison of different gene expression patterns allows biologists to identify candidate genes that may regulate the formation of normal tissues and organs and to search for genes associated with congenital diseases. No two individual embryos, or organs, are exactly the same shape or size so comparing spatial gene expression in one embryo to that in another is difficult. We will present our efforts in comparing gene expression data collected using both volumetric and projection approaches. Volumetric data is highly accurate but difficult to process and compare. Projection methods use UV mapping to align texture maps to standardized spatial frameworks. This approach is less accurate but is very rapid and requires very little processing. We have built a database of over 180 3D models depicting gene expression patterns mapped onto the surface of spline based embryo models. Gene expression data in different models can easily be compared to determine common regions of activity. Visualization software, both Java and OpenGL optimized for viewing 3D gene expression data will also be demonstrated.

Keywords: UVW mapping, in situ hybridization, gene expression

1. INTRODUCTION

1.1 Why 3D analysis is important

Every cell in an animal contains exactly the same complement of genes within its chromosomes, with a few rare exceptions such as cells of the immune system. What makes one cell type different to another is the specific subset of the genes within its chromosomes that are active and copied into mRNA and protein. Mapping where each gene is expressed in space and time provides invaluable information on where and when it can contribute to the organism. For example protein-protein interactions play critical roles in both normal cell function and in disease. Many high throughput systems predict protein-protein interactions based on shape matching and in vitro/in vivo binding assays, yet if the two proteins are never present at the same place and time- they obviously cannot physically interact under normal circumstances. A second example is embryonic development, where temporal and spatial regulation of gene expression drives the formation of different cells and tissues. Knowing when and where genes are active is a powerful tool in the search for candidates that may regulate cell specification and also contribute to congenital disease.

pvize@ucalgary.ca; phone 403 220 8502; fax 403 289 9311

Multimedia Content Access: Algorithms and Systems, edited by Alan Hanjalic, Raimondo Schettini, Nicu Sebe,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6506, 65060H, © 2007 SPIE-IS&T · 0277-786X/07/$15

1.2 Generating and detecting 3D gene expression data

There are a number of methodologies for detecting where in an embryo, or tissue, a gene is actively producing a gene product such as an mRNA or a protein. The presence of an mRNA or protein in a cell indicates that the corresponding gene must be transcriptionally active, or has been within a few half lives of the gene product. When a gene is active, the gene is referred to as either expressed (gene expression) or transcribed (gene transcription). By far the greatest spatial resolution is obtained by analyzing the distribution of gene products in situ, or in their natural position or place. This can be done for example by staining a whole embryo for the presence of a gene product. mRNA distribution patterns can be determined by annealing a complimentary nucleic acid probe containing an detectable epitope to a whole embryo, then visualizing localization with some form of enzymatic assay (Harland, 1991).

The test system that we explore is genes active in the developing kidney. The embryos of the South African frog, *Xenopus laevis*, are used in these experiments as thousands of embryos can be obtained easily and cheaply, the later point being critical when one is staining thousands of gene patterns in tens of thousands of embryos. An example of this technique is illustrated in Figure1. In the first panel (Fig 1A) the frog embryo has been hybridized to an tagged RNA encoding a kidney ion transporter then the tagged RNA detected using an enzymatic assay that generates a colored precipitate. In the actual sample the embryo is a pale yellow color and the stain an intense dark purple (http://www.xenbase.org/index/kidney/NKCC2/32pfar.jpg). The second panel utilized a similar system but the enzymatic detection method generates an fluorescent product that cross links to the cells in which it is generated.

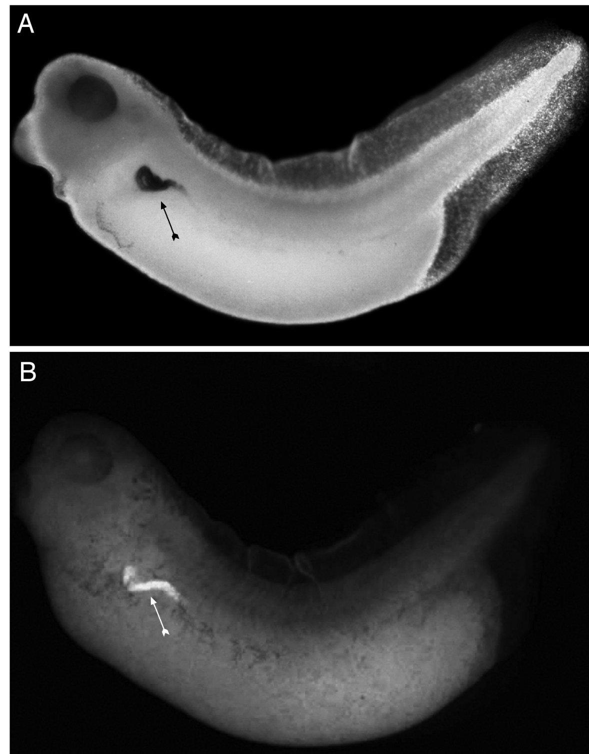

Figure 1. Expression of an ion transporter gene in frog tadpoles.
A. Colorimetric detection system generates an opaque purple colored stain in tissues actively transcribing the gene product. In this case the only tissue stained is a short segment of the kidney (black arrow). To view a color image of a similar embryo visit http://www.xenbase.org/index/kidney/NKCC2/32pfar.jpg B. Fluorescent detection system that generates a green fluorescent product in cells containing the gene product. The marked tissue (white arrow) is once again the developing kidney. To view a color image of a similarly stained sample, please see http://www.xenbase.org/methods/FISH/33ffar.jpg

While the first technique is an effective method to visualize where a gene is transcribed and it can be documented by a simple digital image it is not generally amenable to volumetric sampling, although this can be achieved by optical projection tomography (Sharpe et al., 2002; only available as a service) or by slicing up the stained sample, photographing each section, then building a 3D model (labor intensive). The second technique can be sampled either by digital imaging or by confocal optical sectioning. This technique uses a focused laser to activate fluorescence and overlapping optical sectioning to sample the entire embryo. Note that the two embryos, while very similar, do have differences in shape. While both have been stained to detect the transcriptional activity of the same gene, the shape of the stained structure is somewhat different between the two samples. The issue we face is nicely encapsulated by these two images: how can we compare expression patterns for different genes to each other, when the pattern of the same gene varies between samples?

One approach would be to simultaneously sample many genes within the same embryo. While colored stains like the purple one used in figure 1A cannot be accurately overlayed with other colored stains, fluorescently stained samples can. Different gene products can each be labeled with a different fluor, then the confocal can sequentially scan the embryo for each of the distinct excitation wavelengths (Kosman et al., 2004). This approach would allow multiple patterns to be compared with a common 3D space. However, at most 3 or 4 genes can realistically be sampled at one time, and performing such a combinatorial analysis of every gene product would be unrealistic.

To compare two different gene expression patterns we need to be able to compare results obtained in one sample to those obtained in another. Two different approaches to this problem are presented. In the first fluorescent volumetric sampling is used and in each sample a test gene is compared to a reference gene or genes. In the second approach colorimetric staining is used then this is mapped to a B-spline based reference system.

2. Comparison of 3D gene expression patterns

2.2 Volumetric data

In this example the expression of two different genes within an overlapping domain of the forming kidney was used. One gene is expressed throughout the forming kidney nephron and serves to mark the shape of the forming structure. This shape varies from embryo to embryo, and also over time. The detection of two different gene products was performed by labeling one gene product with a green fluor and one with a red fluor. A series of optical sections through the kidney region was then collected in each of the two independent samples using a confocal microscope and a 20x water immersion lens. The optical stacks from each data set (about 500 mb) were imported directly into Amira (http://www.amiravis.com) where the overlapping sections were segmented and used to build a surface mesh for each sample (Figure 2). For convenience one surface is colored red and the other green. If these two very different shapes are aligned within a common space their differences in topology result is very little overlap.

The 3D authoring package 3DStudioMax (Max) produced by Autodesk (http://www.autodesk.com/) has powerful object editing capabilities and is widely used in the video game authoring industry. We utilized this software to explore whether it could be used to edit the surfaces that depict the boundaries of gene expression domains illustrated in figure 2. The first step in this process was to generate an internal skeleton (Fig 2C). Each triangle of the surface mesh was then linked to the nearest point in the skeleton. The control points of the skeleton can then be used to stretch the linked mesh into an uncoiled, more two-dimensional shape (Fig 2D). This process was independently repeated for the two different gene expression domains then the two surfaces aligned using their skeletons. In panel 2F overlap between the red and green channels is shown in yellow. As the figure illustrates there are now very significant areas of overlap. The time line used to uncoil either one of the models can then be reversed to recoil the skeleton into its original shape, moving the now associated skeleton, and its associated surface mesh, into the same 3D shape. The resulting aligned 3D comparison clearly shows

that these two genes share a very large region of overlap in their expression domains.

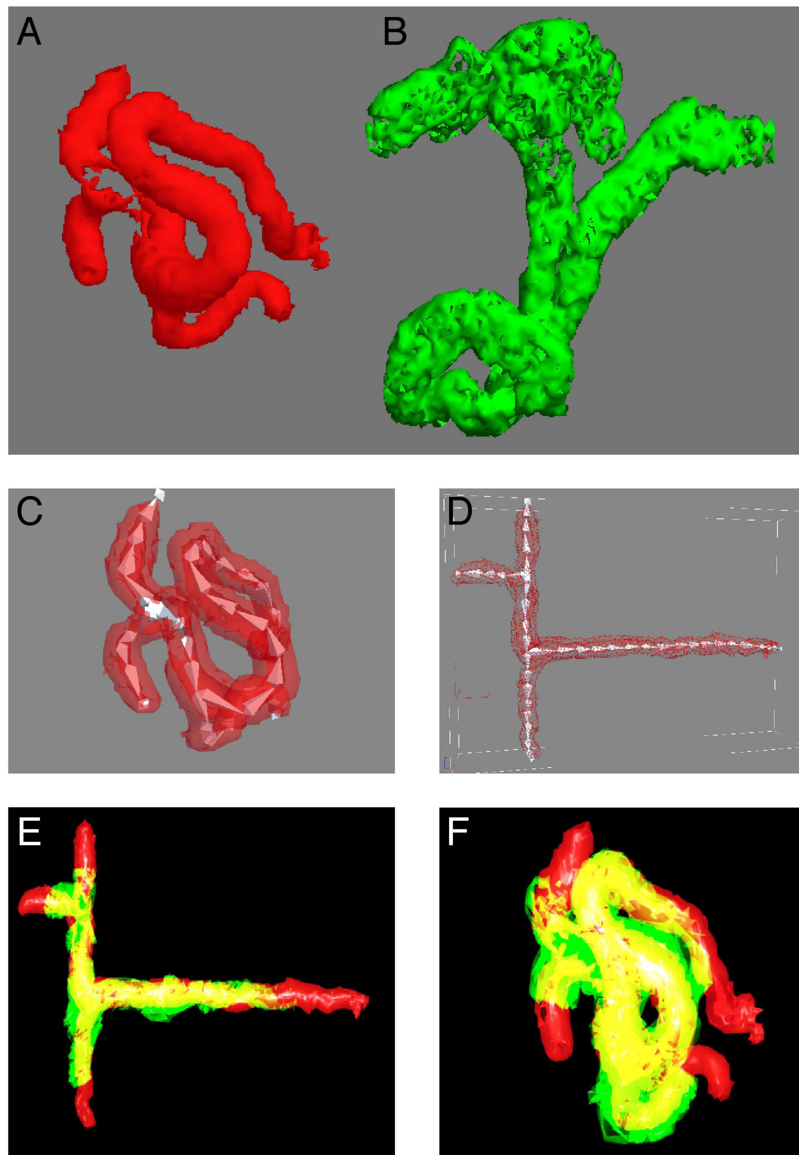

Figure 2. Max based alignment of two independent gene expression domains using skeletal deformation. A and B depict two independently sampled gene expression patterns from slightly different ages of development. C. Each model was provided with an internal skeleton. D. Skeletal control points were used to uncoil both samples. E. The two uncoiled samples were aligned. Additive color indicates areas of overlap as yellow. F. The timeline used to uncoil model 1 was reversed, recoiling both this sample and the now linked model 2. The area of overlap in 3D space appears yellow.

At the moment this alignment and editing process is performed manually, although we have made some progress at automated conversion of polygonal surfaces into B-spline networks. In examples where surface deformation is not required, for example when two samples are extremely similar in shape, we have written two different systems for manual alignment of gene expression data. The first of these is known as the Interactive 3D Gene Expression viewer (Gerth and Vize, 2005; http://www.xenbase.org/3DModels/release2/release2.html/). This is a customized Java based applet

based on the IDX3D software by Peter Walser (http://www.idx3d.ch/idx3d/idx3d.html). It allows users to load different models (in .3DS format) into different color spaces, independently align the models, then view overlap regions using additive color. A second system written in ObjectiveC that can both build surfaces directly from tiff stacks plus perform similar 3D alignments of multiple surfaces is under development and near completion.

While the above experiments show that manual alignment of very different samples is possible, it is not a simple process and the required sampling, processing, editing and manipulation makes pair-wise comparisons of thousands of gene expression patterns unrealistic. We therefore sought a simpler approach more amenable to rapid, if less accurate, comparisons.

2.3 UVW mapping

UV mapping links points in two dimensions (U and V- a third dimension, W, can also be used if needed) to other points in a distinct 3D space with X, Y and Z coordinates. To apply this technique to the mapping of gene expression patterns we use U and V coordinates within a digital color image to texture map the image to X, Y and Z coordinates within a B-spline framework representing an embryo. Databases of many thousands of digital images depict gene expression data such as that shown in figure 1 are already available in public databases (e.g. http://xenopus.nibb.ac.jp/). All that is required to convert this data from static 2D images to closely aligned 3D models is a large scale effort at UV mapping.

We first generated a series of B-spline frameworks that could be used as reference models and to which gene expression data could be mapped. Frog embryos of eight different developmental stages from pre-gastrula to tailbud tadpoles, were scanned on a confocal microscope. Once again Amira software was used to process confocal stacks and to build high-resolution threshold surfaces. Max was then used to convert polygonal edges to B-splines, and the B-splines in turn to produce Bezier surface patches and a 3D surface model.

Multiple images, for example images taken of a gene expression pattern from different angles, can be mapped onto the same model to generate a composite. Each image contains a lot of information in addition to the simple presence or absence of gene expression detected via a colorimetric staining protocol. For example in figure 1 the eye of the tadpole is obvious, as is the clear tail fin. These anatomical features aid in accurately mapping the image to the model. To aid in the mapping process, custom scripts were written in Max to load all of the images within a directory along with the closest aged model. Once this data is pre-loaded a software modeler can apply snap points and wrap one or more images as texture maps around the surface (Figure 3).

To visualize the UV mapped images we have been utilizing Shockwave, an OpenGL and Direct X tool that runs as a web browser plugin (http://sdc.shockwave.com/shockwave/download/). This system uses hardware acceleration to provide excellent performance and it is freely available for a variety of operating systems and browsers. Shockwave models can be directly exported from Max. If further model optimization is required, this can be achieved using the Adobe product Director (http://www.adobe.com/). To enhance the usefulness of this system we built a custom control bar to allow model magnification, rotation, and measurement, using Directors programming language, Lingo. For an example please see http://www.xenbase.org/3DModels/UVGeneMapping/40_e.html/

We have built over 180 models mapping gene expression data to B-spline based frameworks (VEG, Kaori Katsuyama, Atsushi Kitayama, Naoto Ueno and PDV, in prep.). These can be browsed and viewed at http://xenopus.nibb.ac.jp/

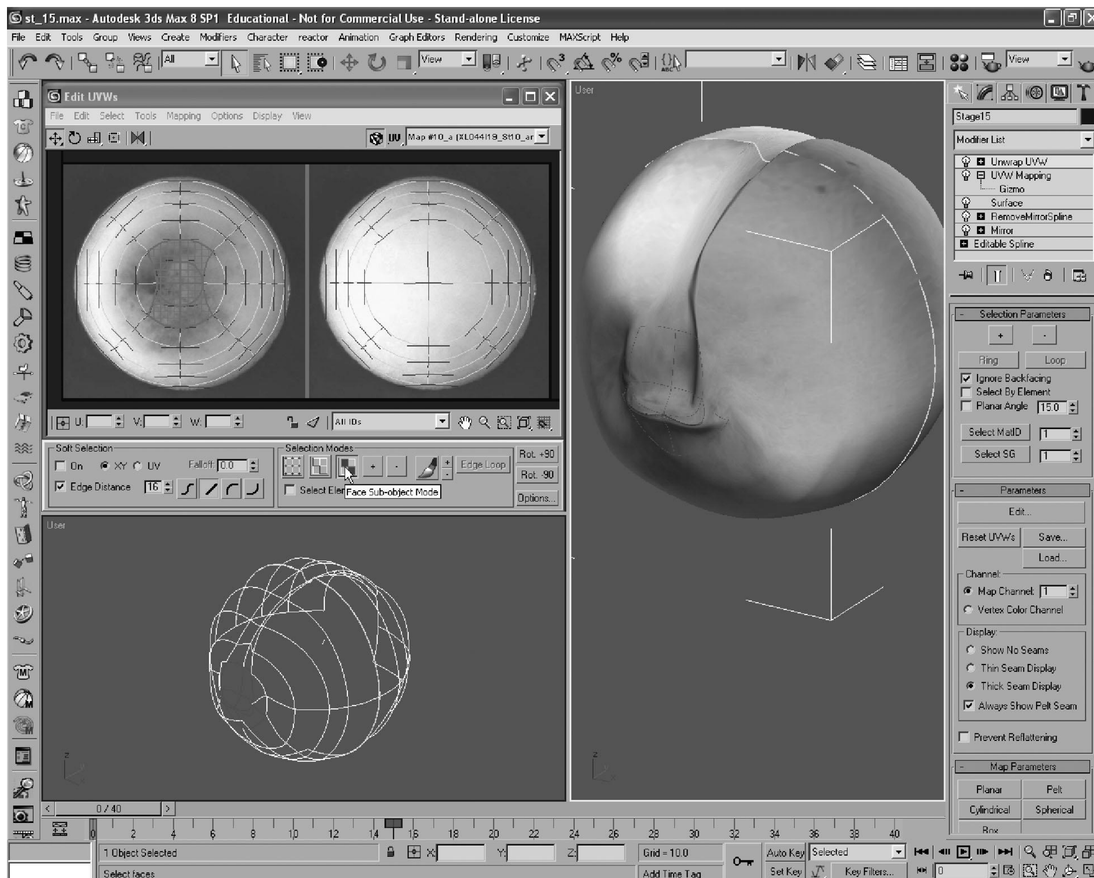

Figure 3. UV mapping.
Custom Max scripts load the surface plus any jpg images in a directory into the editing space. Reference points in the jpg image are then linked to the surface model.

In addition to simply building 3D models of gene expression data, the careful alignment of texture maps to the 3D model closely aligns all images mapped to a common model. While slight differences in view angle and sample to sample variation make pair-wise comparisons of photographs of very little usefulness, any two samples aligned by UV mapping to the same model can be compared to search for overlap. This is done by skinning the surface off the models using a common model dissection path to generate 2D representations of the 3D data (Figure 4). These can then be accurately aligned to compare patterns.

3. CONCLUSIONS

Two completely different approaches were used to explore whether video game authoring approaches to model generation and manipulation could be applied to the spatial analysis of gene expression data. The first system utilized very large data sets to build models, then extensive model manipulation used to map two independent patterns to each other in 3D space. While this approach works, it requires expensive and laborsome data collection and analysis. The second approach used was UV mapping. While this technique has a lower level of spatial resolution it also allows for sample to sample variation to be minimized and for pair-wise comparisons of gene expression data. UV mapping is

simple to perform and many thousands of gene expression data sets could be generated and compared. Given that many data sets are already available though image libraries and the simple hardware and software requirements of this system, we propose that it is a valid and powerful approach to the spatial analysis of gene expression.

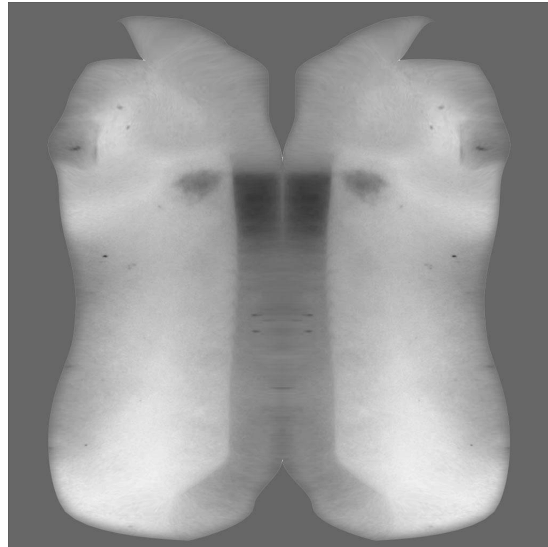

Figure 4. A 2D gene expression pattern map skinned from a UVW mapped 3D model. All skins removed in the same manner from the same surface can be accurately aligned and regions of overlap identified.

Acknowledgments

This work was supported by grants from the Alberta Heritage Foundation for Medical research, The Canadian Institutes for Health Research, and the Alberta Network for Proteomics Innovation.

References

Gerth, V.E. and Vize, P.D. (2005). A Java tool for dynamic web-based 3D visualization of anatomy and overlapping gene or protein expression patterns. Bioinformatics 21; 1278-1279

Harland RM. (1991) In situ hybridization: an improved whole-mount method for Xenopus embryos. Methods Cell Biol. 1991;36:685-695.

Kosman D, Mizutani CM, Lemons D, Cox WG, McGinnis W, Bier E. (2004) Multiplex detection of RNA expression in Drosophila embryos. Science 305: 846

Sharpe J, Ahlgren U, Perry P, Hill B, Ross A, Hecksher-Sorensen J, Baldock R, Davidson D. (2002) Optical projection tomography as a tool for 3D microscopy and gene expression studies. Science 296 :541-545.