Hybrid Segmentation Framework for Tissue Images Containing Gene Expression Data

Musodiq Bello[1], Tao Ju[2], Joe Warren[2], James Carson[3], Wah Chiu[3], Christina Thaller[3], Gregor Eichele[4], and Ioannis A. Kakadiaris[1]

[1] Computational Biomedicine Lab, Dept. of Computer Science,
University of Houston, Houston TX, USA
[2] Dept. of Computer Science, Rice University, Houston TX, USA
[3] Verna and Marrs McLean Dept. of Biochemistry,
Baylor College of Medicine, Houston TX, USA
[4] Max Planck Institute of Experimental Endocrinology, Hannover, Germany

Abstract. Associating specific gene activity with functional locations in the brain results in a greater understanding of the role of the gene. To perform such an association for the over 20,000 genes in the mammalian genome, reliable automated methods that characterize the distribution of gene expression in relation to a standard anatomical model are required. In this work, we propose a new automatic method that results in the segmentation of gene expression images into distinct anatomical regions in which the expression can be quantified and compared with other images. Our method utilizes shape models from training images, texture differentiation at region boundaries, and features of anatomical landmarks, to deform a subdivision mesh-based atlas to fit gene expression images. The subdivision mesh provides a common coordinate system for internal brain data through which gene expression patterns can be compared across images. The automated large-scale annotation will help scientists interpret gene expression patterns at cellular resolution more efficiently.

1 Introduction

With mammalian genomes of over 20,000 genes [1] now sequenced, the next challenge facing the biomedical community is to determine the function of these genes. Knowledge of gene function is important for a better understanding of diseases and the development of potential new therapies. The mouse is a well-established model system for exploring gene function and disease mechanisms. Consequently, determining where genes are active in different mouse tissues can lead to a greater understanding of how gene products affect human diseases. Non-radioactive *in situ* hybridization (ISH) is a histological method that can be applied to reveal cellular-resolution gene expression in tissue sections [2]. This is an appropriate resolution for addressing the questions about the role of genes in cell identity, differentiation, and signaling. Robotic ISH enables the systematic acquisition of gene expression patterns in serially sectioned tissues [3]. By organizing a large collection of gene expression patterns into a digital atlas, ISH data can be used to make great advances in functional genomics as DNA sequence databases have done.

  A major step toward efficient characterization of gene expression patterns is the automatic segmentation of gene expression images into distinct anatomical regions and
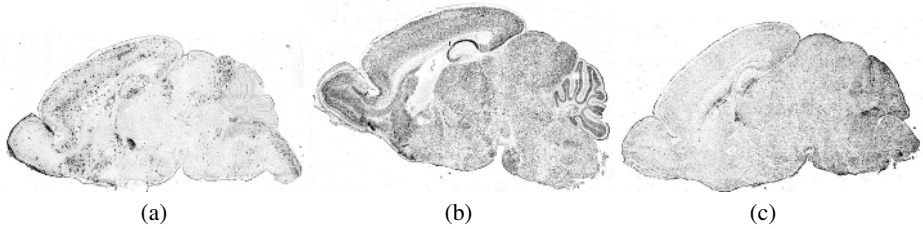
Fig. 1. Variation in the shape of the brain and expression pattern of (a) *Npy*, (b) *Cbfat2t1h*, and (c) *Neurog2* genes in mouse brain images

subregions. This is a challenging task, mainly because each gene is expressed differently from region to region leading to a substantial variation in the appearance of each image. There is also a natural variation in the shape of anatomical structures, compounded by the non-linear distortion introduced during sectioning of the brain. Moreover, there are many regions where no edges or intensity variation can be visually observed. Figure 1 depicts typical gene expression images.

To compare gene expression patterns across images, Ju *et al.* [4] constructed a deformable atlas based on subdivision surfaces which provide a common coordinate system when fitted to sagittal sections of the mouse brain. The 2D brain atlas is represented as a quadrilateral subdivision mesh, as shown in Fig. 2(a). Subdivision is a fractal-like process of generating a smooth geometry from a coarse shape [5]. Starting from an initial mesh, subdivision generates a sequence of refined meshes with increasing smoothness. In our application, the mesh is partitioned by a network of crease edges into sub-meshes, each modeling a particular anatomical region of the brain. The atlas was fitted to images using affine transformation to account for rotation and translation, and local deformation based on iterated least-squares to account for shape variation. However, the accuracy of the local fitting, and interior coordinate system resulting from the segmentation, is limited by its reliance on tissue boundary detection only. Thus, manual deformation of the internal regions of the atlas must still be performed.

In our previous work [6], we have extended that approach by identifying selected anatomical landmarks in expression images and using them to guide the fitting of internal regions of the mesh. Our method improved the general fitting of the internal regions,
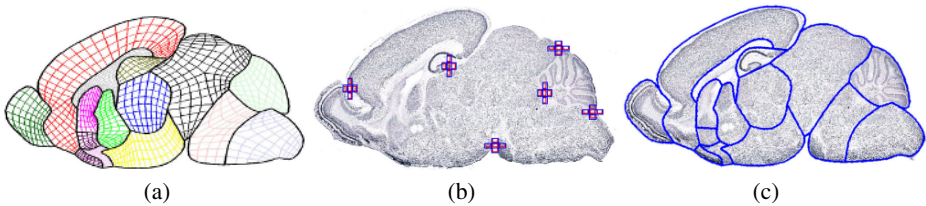

Fig. 2. (a) The atlas at subdivision level 2. (b) Feature-extracting templates overlaid on selected anatomical landmarks. (c) A typical gene expression image manually segmented into anatomical regions.

ensuring that specific landmarks were placed in appropriate regions. However, the region boundaries did not always match those drawn by neuroanatomists. In this paper, we propose a new hybrid segmentation framework that combines texture variation at region boundaries with textural features of specific landmarks to deform the subdivision atlas. In the rest of the paper, we explain the hybrid segmentation framework in detail in Section 2 and present results from using our algorithms in Section 3. Section 4 summarizes our work.

2   Hybrid Segmentation Framework

Our hybrid model is a triplet $\{S, B, L\}$ where S represents the shape of the subdivision mesh, B represents the appearance of the quads on the boundaries of anatomical regions, and L models the texture features of selected anatomical landmarks. The shape and boundary quad appearance are obtained for multiple mesh subdivision levels. Our framework consists of training and deployment stages. Training is performed on several mouse brain gene expression images which were previously fitted with a standard subdivision atlas by neuroanatomists (Fig. 2(c)). Deployment involves fitting the atlas to new gene expression images in order to segment them into anatomical regions.

2.1   Training

Shape: The shape term, S, defines the geometry and topology of the subdivision atlas (Fig. 2(a)) that will be fitted to each image. The geometry is a collection of the coordinates of the vertices of the mesh at a given subdivision level while the topology denotes the relationships between the vertices to form anatomical regions. The geometry is modeled as $\mathbf{x}^k = [x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]^T$ for a mesh at subdivision level k, where $[x_i, y_i]$ are the Euclidean coordinates of vertex i. For all N meshes in the training set, the mean shape is obtained as: $\bar{\mathbf{x}}^k = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}^k$. A training instance that is close to the mean was selected as a standard mesh. The shape is obtained for different subdivision levels of the mesh in a multi-resolution approach.

Boundary quad features: The second element, B, of our hybrid model captures information about the features at the anatomical region boundaries. It can be observed that the cell density pattern in the cerebellum is different from that of its neighbors in most images. A texture variation can similarly be observed along the boundaries of the cortex, septum, and thalamus. This slight variation in the texture patterns of anatomical regions is utilized to model the boundary quads at subdivision level k as $B^k = [B_1^k \ B_2^k \ ... \ B_s^k]$ for s selected segments in the mesh boundary. A boundary segment is a collection of adjacent crease edges and has quads from no more than two anatomical regions attached to it (Fig. 3). By separating the regional boundaries into segments, optimal features for each segment can be chosen, since no set of features will be equally optimal for all region boundaries. For each segment, $B_j = \{Q^j, F^j, p^j\}$ where Q^j is the set of all quads attached to segment j and distinguished by the side of the segment they belong to, F^j is the set of optimal features, and p^j is the set of classifier parameters to distinguish

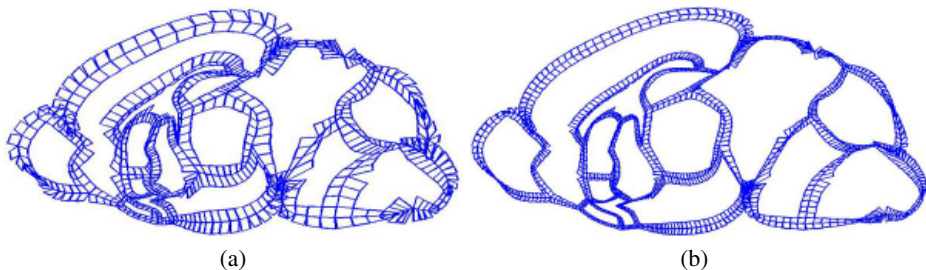(a)   (b)

Fig. 3. Boundary quads at subdivision level (a) 2, and (b) 3

between quads on either side of the boundary segment. For purposes of this research, we have selected the Support Vector Machine (SVM) [7] classifier due to its ability to obtain a good decision boundary even on sparse data sets. The optimal feature set, F^j, and classifier parameters, p^j, for each segment are obtained as follows:

Step 1: Extract features from quads. Each image was filtered with Laws' texture filters [8] to obtain a bank of 25 texture energy maps. The 1st, 2nd, and 3rd moments of the distribution of intensity values of all pixels in each quad on either side of the boundary were used as feature for the quad in each filtered image [4].

Step 2: Feature normalization. The feature values vary widely, necessitating normalization. A few feature normalization techniques were considered, including linear scaling to unit range, linear scaling to unit variance, transformation to uniform random variable, and rank normalization [9]. We obtained the best performance with rank normalization in which the feature values $f_1, f_2, \ldots, f_m$ are first ordered to obtain their ranks $(1 \ldots m)$, where m is the number of samples in the boundary segment. Each feature value is then replaced by $\tilde{f}_i = \frac{Rank(f_i)-1}{m-1}$.

Step 3: Optimal feature selection. The relevance of each feature f is computed using the Information Gain (IG) metric [10] after discretization using Fayyad and Irani's minimum description length algorithm [11]. The features were then sorted according to the relevance indices assigned by IG and each feature is included one-at-a-time in a feature set, F^j. The average error, E_c, of classifying with the feature set F^j is then obtained in a 10-fold cross-validation and the smallest set of features, F^j, with a stable and sufficiently low value of E_c is selected.

Step 4: Model parameter computation. For each segment j, a SVM classifier was trained to distinguish between quads on either side of the boundary segment based on the optimal features. Best performance was obtained by using the Radial Basis Function (RBF) kernel [12] with SVM. The optimal values for the kernel parameter and error penalty parameter are obtained by cross validation and used to compute SVM model parameters, p^j, for each segment as part of the hybrid model.

Anatomical landmark features: In addition to the region boundary quads, a few anatomical landmarks were modeled with respect to their texture features. Each landmark is associated with a vertex and they are used to guide the general position and

orientation of the mesh during fitting. Each landmark i is modeled as $L_i = \{v_i, F^i, p^i\}$, where v_i is the coordinates of the vertex that it is attached to at subdivision level 3, the highest level used in the model. The set of optimal features, F^i, that can be used to distinguish a landmark from its surrounding area and the set of classifier parameters, p^i, are computed as described above. To extract the features for the landmarks, a rectangular template was overlaid on the landmark in each of the texture maps and summary statistics for sub-windows in the template were used as features. Similar features were extracted from a 4-neighborhood (Fig. 2(b)) of the landmark to serve as non-landmark examples as described in our previous work [6].

2.2   Deployment

Given a new image, the model is fitted to the image by minimizing a quadratic energy function $E(\mathbf{x}^k)$ of the form: $E(\mathbf{x}^k) = E_f(\mathbf{x}^k) + E_d(\mathbf{x}^k)$ using a linear solver such as conjugate gradient. The energy term $E_f(\mathbf{x}^k)$ measures the fit of the mesh at subdivision level k to the image and $E_d(\mathbf{x}^k)$ measures the energy used in deforming the mesh. The fitting term $E_f(\mathbf{x}^k)$ is: $E_f(\mathbf{x}^k) = \alpha E_U(\mathbf{x}^k) + \beta E_B(\mathbf{x}^k) + \gamma E_L(\mathbf{x}^k)$, where $E_U(\mathbf{x}^k)$ is the fitting error of the outer boundary of the mesh to the outer boundary of the image, $E_B(\mathbf{x}^k)$ measures the fitting error of the regional boundaries resulting from the classification of the boundary quads, and $E_L(\mathbf{x}^k)$ measures the error of fit of the anatomical landmarks. The formulation of $E_U(\mathbf{x}^k)$ and the deformation energy term $E_d(\mathbf{x}^k)$ are the same as in [4]. The method for obtaining the other terms $E_L(\mathbf{x}^k)$ and $E_B(\mathbf{x}^k)$ is described below.

Step 1 - Shape Initialization: First, a global alignment of the reference shape to the image is performed. The image was segmented from the background using a flood-filling approach after a simple intensity threshold. Principal Component Analysis was applied to obtain the principal axes of the segmented image. The principal axes of the reference mesh are also obtained and an affine transformation of the mesh is performed to align the two pairs of axes [4].

Step 2 - Using the Landmarks for Fitting: Second, the fitting error, $E_L(\mathbf{x}^k)$, of the mesh to the landmarks is computed as $\sum_i (l_i - v_i)^2$, where v_i is the vertex of the mesh associated with landmark i detected at location l_i. Specifically, for each landmark, features are extracted for the pixels around the expected location and classified using the model parameters obtained at the training stage. For efficiency, classification is performed on every third pixel initially before conducting a pixel-by-pixel search around the area with the highest SVM decision values. This is possible because the decision values were found to monotonically increase towards the expected ground truth in all the images tested [6].

Step 3 - Using Boundary Quads for Fitting: Finally, for each segment j on the boundary at subdivision level 1, optimal features F^j are extracted for the quads Q^j in the segment and the model parameters p^j are used to classify each quad. There are four possibilities when two quads on opposite sides of a crease edge are classified with respect to the model. When the classification of both quads is in agreement with the model (Fig.
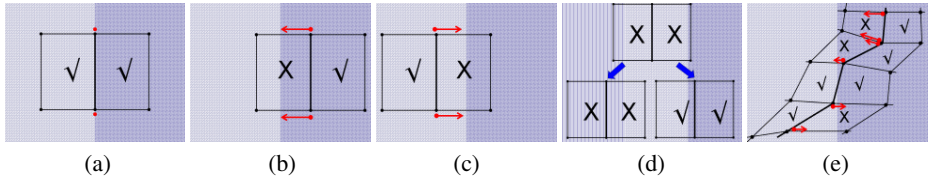
Fig. 4. Classifying opposite quads on a boundary crease edge (the background depicts two regions): (a) both are classified in accordance with the model, no displacement at the vertices. (b,c) Displacement of the boundary edge towards the misclassified quad. (d) Rare case of opposite quads being simultaneously incorrectly classified with respect to the model: position the quads on either side and select best match. (e) Various scenarios of quad classification and the resulting displacement at the vertices.

4(a)), the force exerted by the corresponding vertices is zero. When either of the quads is classified contrary to the model (Fig. 4(b,c)), the vertices exert a force pulling the boundary edge in the direction of the misclassified quad. The decision value returned by SVM gives an estimate of the confidence in the classification. In the event that both quads are incorrectly classified with respect to the model (this is very rare since the mesh boundaries are already quite close to the image boundaries after global fitting), the two quads are temporarily positioned on both sides of the segment and the position that results in correct classification is retained (Fig. 4(d)). If both positions still result in misclassified quads, the vertices are left unchanged. The various possibilities are illustrated in Fig. 4(e).

This process is performed iteratively until a specified ratio of the quads (e.g., 95%) are correctly classified, in which case the process is repeated at a finer subdivision level, up to level 3. With increasing subdivision level, the size of the quads decreases. This reduces the displacement of the vertices, resulting in a smooth fit.

3   Results and Discussion

Our experimental data are 2D images of sagitally sectioned (level 9) postnatal day 7 (P7) C57BL/6 mouse brains on which *in situ* hybridization has been performed to reveal the expression of a single gene. For computational efficiency, the images were scaled down by 25% from their original size of approximately 2000x3500 pixels. We trained our framework on 36 images manually fitted with subdivision meshes by neuroanatomists, and tested on 64 images. The weights for the terms $E_U(\mathbf{x}^k)$, $E_B(\mathbf{x}^k)$, and $E_L(\mathbf{x}^k)$ in the energy minimization equations were selected experimentally.

To quantify the quality of fit using our hybrid segmentation approach, we compared individual anatomical regions as delineated by our framework with those manually delineated by neuroanatomists. The number of overlapping pixels in both meshes was further normalized by the total number of pixels in the manually fitted mesh. In Fig. 5(a-e), the accuracy of fit of five major anatomical regions is compared for all 64 images. In regions where the distinction along the boundaries is pronounced such as the cortex, cerebellum and their adjoining regions, we expect to have higher
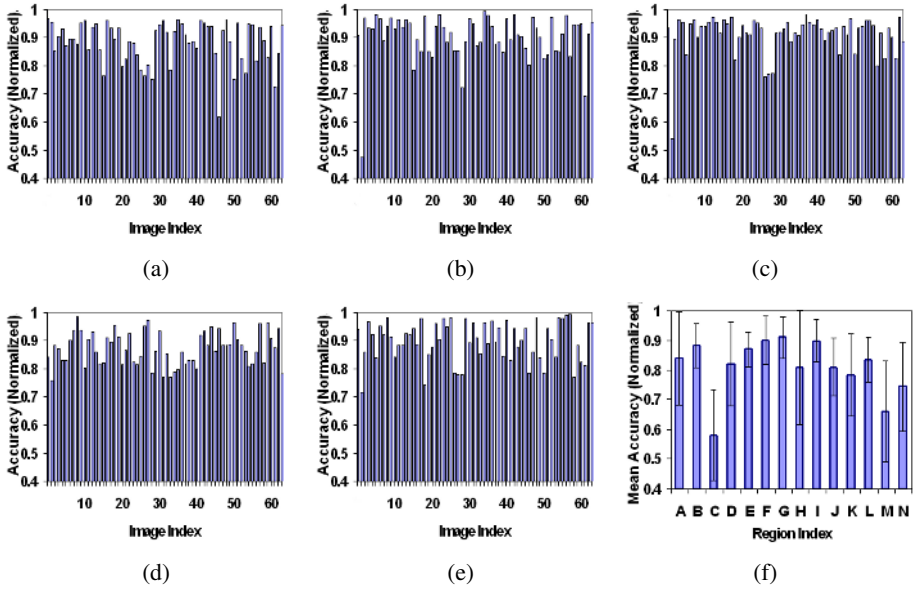
Fig. 5. Comparison of the accuracy of the fit in the (a) cortex, (b) medulla, (c) midbrain, (d) hypothalamus, and (e) pons in 64 expression images. (f) Mean accuracy of the fit in the 64 images across 14 regions using our hybrid segmentation framework. The error bars indicate the standard deviation. The letters indicate the following regions: A-cerebellum, B-cortex, C-basal forebrain, D-hippocampus, E-hypothalamus, F-medulla, G-midbrain, H-olfactory bulb, I-pons, J-septum, K-striatum, L-thalamus, M,N-ventral striatum.
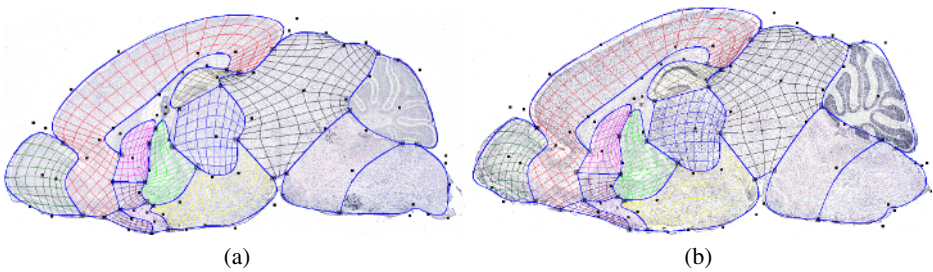

Fig. 6. The result of fitting the standard mesh on (a) *ChAT*, and (b) *BMAL1* gene expression images

accuracy in delineating the boundaries. Similarly, we expect lower accuracy in regions where this distinction is minimal, such as in the forebrain and the ventral striatum. This is confirmed by our results summarized for all 14 regions in Fig. 5(f). Note that the two sub regions of the ventral striatum are treated as two separate regions for purposes of comparison. Examples of fitting using our approach are illustrated in Fig. 6.

4   Conclusion

Due to its many advantages, subdivision surface modeling is becoming increasingly popular for geometric modeling and is also starting to appear in medical applications. For example, in gene expression images, subdivision surface modeling facilitates the comparison of expression patterns not only in regions, but also in subregions of the brain. The challenge is to fit a subdivision-based atlas to expression images accurately and automatically. Our approach combines the detection of selected anatomical landmarks with feature differentiation at regional boundaries using trained classifiers. Our results have been very encouraging.

References

1. Waterston, R., and the Mouse Genome Sequencing Consortium: Initial sequencing and comparative analysis of the mouse genome. Nature **420** (2002) 520–562
2. Albrecht, U., Lu, H.C., Revelli, J.P., Xu, X.C., Lotan, R., Eichele, G.: Studying Gene Expression on Tissue Sections Using In Situ Hybridization. In: Human Genome Methods. CRC Press, Boca Raton (1997) 93–119
3. Carson, J., Thaller, C., Eichele, G.: A transcriptome atlas of the mouse brain at cellular resolution. Curr Opin Neurobiol **12** (2002) 562–565
4. Ju, T., Warren, J., Eichele, G., Thaller, C., Chiu, W., Carson, J.: A geometric database for gene expression data. In: Eurographics Symposium on Geometry Processing, Aachen, Germany (2003) 166–176
5. Warren, J., Weimer, H.: Subdivision Methods for Geometric Design: A Constructive Approach. Morgan Kaufmann Publishers, San Francisco, CA (2002)
6. Kakadiaris, I.A., Bello, M., Arunachalam, S., Kang, W., Ju, T., Warren, J., Carson, J., Chiu, W., Thaller, C., Eichele, G.: Landmark-driven, atlas-based segmentation of mouse brain tissue images containing gene expression data. In: Proc. 7th International Conference on Medical Image Computing and Computer-Assisted Intervention, Rennes, France (2004) 192–199
7. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer-Verlag (2000)
8. Laws, K.: Textured Image Segmentation. PhD thesis, USC (1980)
9. Aksoy, S., Haralick, R.: Feature normalization and likelihood-based similarity measures for image retrieval. Pattern Recognition Letters **22** (2001) 563–582
10. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. Journal of Machine Learning Research **3** (2003) 1157–1182
11. Fayyad, U.M., Irani, K.B.: On the handling of continuous-valued attributes in decision tree generation. Machine Learning **8** (1992) 87–102
12. Schölkopf, B., Smola, A.: Learning with Kernels. MIT Press, Cambridge, MA (2002)