Robust Landmark Detection for Alignment of Mouse Brain Section Images

Yuncong Chen, David Kleinfeld, Martyn Goulding and Yoav Freund

Department of Computer Science and Engineering,
University of California, San Diego, La Jolla, CA 92122, USA

Abstract. Brightfield and fluorescent imaging of whole brain sections is a fundamental tool in anatomical studies on brain functions. As sectioning and imaging become more efficient, there is an increasing need to automate the post-processing of sections for alignment and three dimensional visualization. There is a further need to facilitate the development of a digital atlas, i.e. a brain-wide map annotated with cell type and tract tracing data, which serves as the underpinning for circuits analysis as well as for quantitative analysis of phenotypical variation. Currently, construction of such maps requires much human effort. In this work we describe the first steps in developing a semi-automated system to construct a mouse brainstem atlas that combines atlas-guided annotation, landmark-based registration and atlas generation in an iterative framework. In particulcar, we describe an unsupervised approach for identifying and matching region and boundary landmarks, based on modelling texture. Experiments show that the detected landmarks correspond well with human labellings. The results serve as a promising basis for the next stage of registration and atlas building.

Keywords: landmark detection, atlas building, mouse brain, registration, automated annotation

1   Introduction

Brightfield and fluorescent imaging is a fundamental tool in mouse brain studies. As sectioning and image acquisition become more efficient, there is a need to facilitate the development of a digital atlas, i.e. a brain-wide map annotated with anatomical structures as well as cell type, tract tracing and physiological data. Such map serves as the underpinning of brain circuit analysis and quantitative analysis of phenotypical variation.

An atlas is typically constructed by mapping image series of different studies to a common coordinate system, and aggregating the information. Images of different modalities or specimens stained for different targets can be superimposed. For example, the Allen Mouse Brain Atlas [8] computes the expression energy of different genes by collecting statistics from in-situ hybridization images of all the studies. In other cases, inter-specimen variation is interested. For example, the Waxholm Space atlas [7] provides an anatomical probability map that gives the probability of each voxel belonging to a particular anatomical structure. The probabilities are computed from a large number of co-registered specimens.

It is interesting to note that, on the one hand, if the specimen is labelled, annotated atlas facilitates landmark-based registration. On the other hand, atlas-guided annotation requires the specimen be brought into registration with the atlas first, which is often done by manually picking landmarks. In either case, time-consuming human labelling is needed. To alleviate this bottleneck, we propose to use machine learning to model the landmarks with appearance features. By training landmark detectors and storing them in the atlas, landmarks can be automatically identified in new images. These landmarks guide both intra-specimen and specimen-atlas registration. Landmark models are then updated based on registered new images.

This paper describes a first step towards building a three-dimensional histology atlas for mouse brainstem. The system will integrate atlas-guided annotation, landmark-based registration and atlas generation into an iterative framework, allowing the atlas to continue improving itself using new data. The particular aspect described in this paper is the automatic detection and modelling of brainstem landmarks.

Unlike the cerebrum and the cerebellum, where high-contrast cortex presents clear edges, the brainstem mostly contains nuclei, which are compact neuron clusters with homogeneous and distinct texture, and fiber tracts, which are recognizable for their striated texture. This motivates our algorithm to model texture, and detect landmarks including both compact regions and boundaries.

This paper is organized as the following: Section 2 describes how we represent textures using Gabor filters and superpixels. Section 3 and 4 explains how we identify potential landmark regions based on texture distinctiveness. Section 5 describes how we detect stable boundary segments to complement region landmarks. Section 6 describes how landmarks are matched between sections.

## 2   Representing Texture using Histograms of Gabor Textons

Images are first filtered using Gabor filters [6] with 9 orientations and 11 scales, resulting in 99 dimensional feature vectors. We reduce the data by quantizing these feature vectors into 100 clusters using K-Means. Clusters with close centroids are then merged, and the final cluster centroids are the textons. In our experiment, we have 14 textons.

We further reduce the data by describing texture at the level of superpixels. For each superpixel, texture is represented by the histogram of the textons it contains (Figure 1). We denote the texton histogram of the i'th superpixel by h_i. Superpixels are obtained using the SLIC algorithm [1].

Since different sections may be oriented at different angles, we need the features to be rotation-invariant. For this purpose, before doing K-Means, we compute the energy distribution over orientations for each Gabor feature vector, and shift all feature vectors such that the modes of the directional energy distributions are aligned. In this way, the texture is decoupled from directionality.

3   Evaluating Region Significance using Statistical Test

We define a region landmark as a connected set of superpixels. Before describing how to find significant regions, we need a scoring function that evaluates the significance
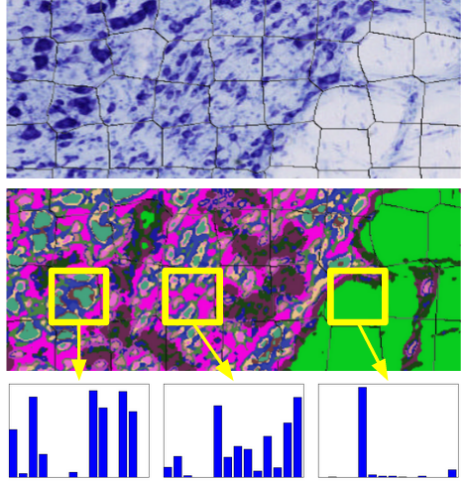
Fig. 1. Top: part of the original image with superpixels overlaid. Middle: texton map of the same area. Bottom: texton histograms of three superpixels with different textures.

of a region. We propose the significance score, $F(S)$, of a region S, to be a linear combination of three terms.

The first term is the distinctiveness of this region's texture, or in other words, how different its texture is from the surrounds. Since textures are represented as histograms, we formulate this problem as a statistical test on whether the region's average texton histogram and the texton histograms of the surrounding superpixels are sampled from the same distribution. We employ the chi-squared test for independence and the resulting p-value serves as a quantitative measure of the distinctiveness. The smaller the p-value, the more significant the region is. Since a region often has neighbours with different textures, the test is made between the region's histogram and that of every surrounding superpixel. To be conservative, the largest p-value among all tests is used. Denote the set of surrounding superpixels by $T(S)$, the first term is,

$$F^{cont}(S) = - \max_{j \in T(S)} pval(h_S, h_j)$$

where $pval(\cdot, \cdot)$ is the p-value of applying the chi-square test for independence to two histograms.

The second term is the texture homogeneity within the region, defined as the mean p-value of chi-square tests between all superpixels in the region and the region's average:

$$F^{coh}(S) = \frac{1}{|S|} \sum_{i \in S} pval(h_S, h_i)$$

The third term stems from the fact that most region landmarks have compact shapes. A common measure for the compactness of a closed curve is the isoperimetric quotient, defined as the enclosing area divided by the square of the circumference. We use the

number of surrounding superpixels as a proxy for the circumference, and the number of superpixels in the region as that for the area. Small

$$F^{comp}(S) = \frac{|T(S)|}{|S|^2}$$

The overall significance score $F(S) = F^{cont}(S) + F^{coh}(S) + F^{comp}(S)$.

4 Finding Significant Regions with Region Growing and Clustering

To find potential regions, we perform region growing for every superpixel. Starting with a singleton containing the seed superpixel, this greedy procedure considers all neighbours of the current set, and iteratively adds the one whose texture is closest to the current region's texture, measured by the χ^2 distance between texton histograms. The growing stops when the region reaches 10% of the total area. Eventually, the region at the point when the significance score is the largest is returned. Figure 2 shows examples of this procedure.
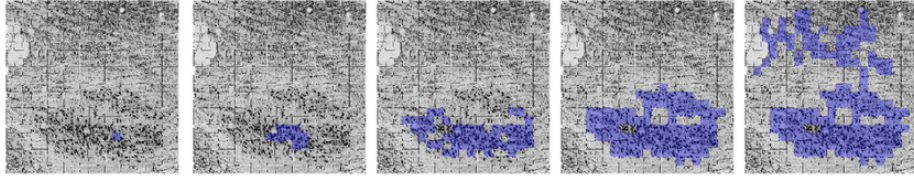

Fig. 2. Regions during growing, from left to right, at iteration 1, 10, 50, 96 (most significant), 150 (last iteration).

We call the set of superpixels that grows from a seed superpixel k the *region proposal* of the seed, denoted by S_k. If a region has distinct texture, the proposals of superpixels inside this region should be very similar, while proposals from within an inhomogeneous region are random (see Figure 3). In other words, the region proposals form clusters. The denser a cluster is, the more distinctive the region it represents. Using Jaccard index as pairwise distance, we group the region proposals using hierarchical clustering. Then within each group whose size is large enough, we select the proposal with the highest significance score to be the representative of that group. All representative proposals are ranked according to their significance scores.

5 Identify Robust Boundaries by Region Consensus

Sometimes a region gradually transitions into neighboring texture on one side, but has a clear boundary on the other side. The method described in the previous section may
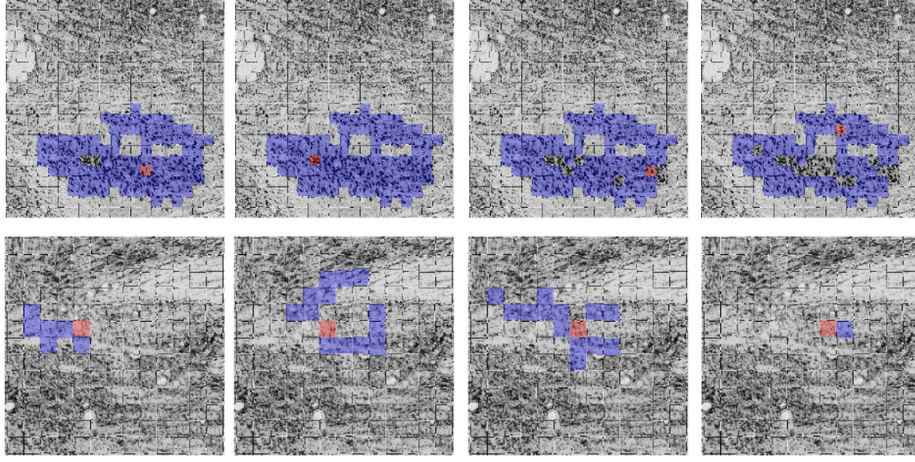
Fig. 3. The first row shows the region proposals of four different seeds (in red) in the facial motor nucleus. They are very similar. The second row shows the proposals of four neighboring seeds in an inhomogeneous region. They are very different.

not capture this inhomogeneous region, but the open boundary is nonetheless a perfect landmark. Figure 4 shows such an example. Notice how much the proposals from seeds in such region vary, but many of them still agree on the clear boundary. This motivates a consensus-based approach for evaluating boundary robustness, in which each region proposal votes for the segments on its boundary.
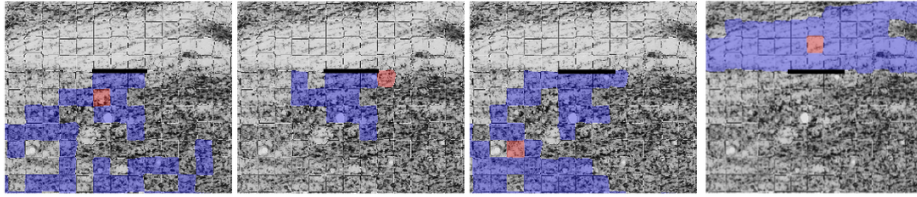
Fig. 4. The three region proposals in an inhomogeneous area are not consistent as a whole, but they all agree on a robust boundary segment (highlighted).

We represent a boundary segment between two superpixels by an ordered tuple (i, j), where i is the interior superpixel and j is the exterior superpixel. We also denote by δS_k, the set of segments on the boundary of a region proposal S_k. The vote a region proposal casts to a boundary segment depends on how contrasty the segment is, measured by the texture histogram distance between the region's average and the segment's exterior superpixel. We call the set of superpixels that vote for a segment the *supporter set* of the segment, and denote it by $R_{(i,j)} = \{k : (i, j) \in \delta S_k\}$. The total

score received by a segment (i, j) is then:

$$b_{(i,j)} = \sum_{k \in R(i,j)} \chi^2(h_{S_k}, h_j)$$

We discard segments whose vote is lower than a threshold. Figure 5 shows a vote map. Instead of modelling individual segments, we combine segments with similar supporter sets into groups, again using hierarchical clustering with Jaccard index between supporter sets as pairwise similarity. Each segment group represents a boundary. All boundaries are ranked according to the total vote received by its segments.
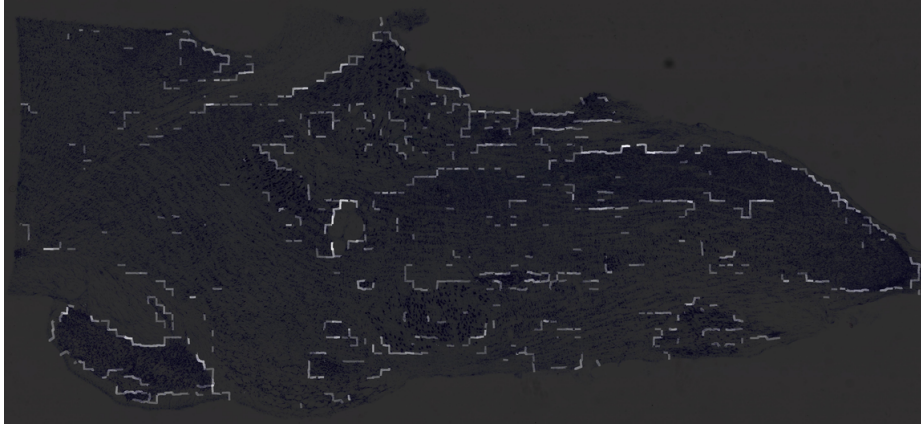
Fig. 5. The thresholded boundary vote map

6   Matching Landmarks from Different Sections

By representing region landmarks using closed boundaries and merging coinciding boundaries, we unify the two types of landmarks into a single set consisting of both open and closed boundaries. In order to use the landmarks for registration, correspondences must be made between them. In this section, we design a distance function for comparing two boundaries. This function is a weighted combination of the differences in four aspects, with weights chosen empirically:

$$D(\mathbf{B}_1, \mathbf{B}_2) = w^{int} D^{int}(\mathbf{B}_1, \mathbf{B}_2) + w^{shape} D^{shape}(\mathbf{B}_1, \mathbf{B}_2)$$
$$+ w^{ext} D^{ext}(\mathbf{B}_1, \mathbf{B}_2) + w^{loc} D^{loc}(\mathbf{B}_1, \mathbf{B}_2)$$

The first term is the difference of interior textures. For a region landmark S, the interior texture is simply the average texton histogram h_S. For a boundary B, the interior texture is the average texton histogram of the union of all segments' supporter

sets. Denote this union by $Q_B = \cup_{(i,j)\in B} R_{(i,j)}$, then $h_B^{int} = h_{Q_B}$. The difference is computed using the χ^2 distance:

$$D^{int}(\mathbf{B}_1, \mathbf{B}_2) = \chi^2(h_{B_1}^{int}, h_{B_2}^{int})$$

The second term measures the similarity of boundary shapes. We reduce the boundary to a point set consisting of midpoints of the segments. The shape distance between two point sets can be computed using shape context descriptors [2]. The shape context descriptors characterize the organization of other points around each point using a histogram. Two sets of points are matched by finding the minimum bipartite matching, where the edge weights are the chi-square distances between shape context descriptors. The Hungarian algorithm is used to find the minimum matching. The average cost of this matching, is used as the second term:

$$D^{shape}(\mathbf{B}_1, \mathbf{B}_2) = \frac{1}{|M|} \sum_{(i,j),(p,q)\in M(B_1,B_2)} \chi^2(c_{i,j}, c_{p,q})$$

where $M(B_1, B_2)$ is the minimum matching, $c_{i,j}$ and $c_{p,q}$ are the shape context descriptors of segment (i, j) and segment (p, q) respectively.

After the matching is made, we compute the third term, defined as the total distance between exterior textures of all matched segments:

$$D^{ext}(\mathbf{B}_1, \mathbf{B}_2) = \sum_{(i,j),(p,q)\in M(B_1,B_2)} \chi^2(h_j, h_q)$$

The fourth term measures the spatial distance. This is the thresholded Euclidean distance between the center of mass of the boundaries' midpoint sets:

$$D^{loc}(\mathbf{B}_1, \mathbf{B}_2) = \max(0, \|m_{B_1}, m_{B_2}\|_2 - l)$$

where m_{B_1} and m_{B_2} are the center of mass of boundaries B_1 and B_2, and l is a tolerance within which the position deviation is not penalized (set to 1 mm in our experiment).

Using this distance function, we compute the pairwise distances for two sets of landmarks detected from different sections. Two landmarks are matched if they are simultaneously the closest landmark to each another.

7 Human Supervision

This system reduces human supervision in annotating brain images. Without any prior annotation, the region landmarks detected by the system can serve as suggestions for significant regions that require annotation. A human labeller then inspects the suggested landmarks. If a landmark is judged to correspond to a real anatomical structure, the human labeller gives it a name and the model of this landmark is stored in the atlas. The human labeler can also manually correct landmarks that are misplaced or create new annotations. The stored models are compared to the landmarks detected on a new image. If some detected landmarks match the models in the atlas, annotation of these landmarks can be transferred to the new image automatically.

8 Experiments

8.1 comparison with human labelings

We test our algorithm on a series of section images of a mouse brainstem, stained with Cresyl Violet for Nissl substance in neurons. The images are scanned at 2 microns per pixel, showing individual neuronal cell bodies. An experienced anatomist is asked to label the same images for anatomical regions that are recognizable from the images .

Recall is more important than precision.

8.2 robustness of matching

Landmarks are detected for two sections that are three sections apart. We have two human labellers judge whether a matching is correct, partially correct, or incorrect.

Shows that matchings are robust to distortion and shape change. Also shows that our distance measure is a sensible one: each of the four terms is important. We show this by changing the term weightings, and then compare matching results.

9 Conclusion and Future Work

In this paper we described a system for detecting region and boundary landmarks from mouse brain images. It grows expansion regions from every superpixel. Based on these region proposals, significant regions such as nuclei and fiber tracts are identified using clustering. These regions are shown to correspond well with real anatomical structures. To complement region landmarks, confident boundary segments are also found using consensus voting. Landmarks from different sections can be matched using a distance function that is robust under distortion.

The next step of this project is to use matched landmarks to align image sections, as well as to align sections to the three-dimensional coordinate space defined by the atlas.

We will explore new representations of texture. A promising direction is to learn visual dictionary directly from the images using independent component analysis or deep neural networks.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. Pattern Analysis and Machine Intelligence, IEEE Transactions on 34(11), 2274–2282 (2012) 2
2. Belongie, S., Malik, J., Puzicha, J.: Shape context: A new descriptor for shape matching and object recognition. In: NIPS. vol. 2, p. 3 (2000) 6
3. Cheng, M.M., Zhang, G.X., Mitra, N.J., Huang, X., Hu, S.M.: Global contrast based salient region detection. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. pp. 409–416. IEEE (2011)
4. Clausi, D.A., Jernigan, M.E.: Designing gabor filters for optimal texture separability. Pattern Recognition 33(11), 1835–1849 (2000)

5. Gao, D., Vasconcelos, N.: Decision-theoretic saliency: computational principles, biological plausibility, and implications for neurophysiology and psychophysics. Neural computation 21(1), 239–271 (2009)
6. Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using gabor filters. In: Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on. pp. 14–19. IEEE (1990) 2
7. Lee, E.F., Jacobs, R.E., Dinov, I., Leow, A., Toga, A.W.: Standard atlas space for c57bl/6j neonatal mouse brain. Anatomy and embryology 210(4), 245–263 (2005) 1
8. Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al.: Genome-wide atlas of gene expression in the adult mouse brain. Nature 445(7124), 168–176 (2007) 1
9. Leung, T., Malik, J.: Representing and recognizing the visual appearance of materials using three-dimensional textons. International Journal of Computer Vision 43(1), 29–44 (2001)
10. Perazzi, F., Krahenbuhl, P., Pritch, Y., Hornung, A.: Saliency filters: Contrast based filtering for salient region detection. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. pp. 733–740. IEEE (2012)