Automatic Region Landmark Detection for Active Atlas

Yuncong Chen, David Kleinfeld and Yoav Freund

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

Abstract. Brightfield and fluorescent imaging is of central importance in mouse brain studies. Current technology allows a whole brain to be scanned in a matter of hours. At the same time, there is an increasing number of fluorescent markers that can be used to label neurons with specific phenotypes. As the number of available images mounts, the manual work required to analyze these images becomes a major bottleneck and automated tools that aggregate information in the images become necessary.
In this work we describe the first steps in a project whose goal is to build an "active" atlas, which combines annotation, registration and atlas building in a semi-automated iterative framework. This atlas explicitly models region landmarks, uses these models to detect landmarks from new images for registration; meanwhile, the models are updated based on newly registered images.
The part of the project described here is a semi-supervised learning system for identifying region landmarks. It detects and models regions that have distinct texture and are consistent across sections. It also models stable boundary segments. Experiments show that the detected landmarks correspond well with labellings by experienced anatomists. The results serve as a promising basis for the next stage of registration and atlas building.

Keywords: landmark detection, atlas building, mouse brain, gabor filter

1 Introduction

Brightfield and fluorescent imaging is of central importance in mouse brain studies. As automation is improving, a whole brain can be imaged in a matter of hours . At the same time, there is an increasing number of fluorescent markers that can be used to label neurons with specific phenotypes. As the number of available images mounts, the manual work required to analyze these images becomes a major bottleneck and tools that aggregates information in the images such as standardized atlases become necessary. A standardized atlas maps image series of different specimens onto a common 3D coordinate system. On the one hand, this allows the new specimen to be compared with annotations and cell type maps collected from previous specimens; on the other hand, new specimens

augment the atlas by introducing annotations for new structures or by providing additional cell type data demonstrated by new markers.

Much effort has been made to construct digital atlases for a variety of species. A popular atlas for mice is the Allen Reference Atlas. It consists of section images of a single brain, with major anatomical structures annotated and localized in the 3D reference space. While this atlas is useful as a visual guide for annotating new images by hand, there is no easy solution for aligning new image series to the atlas. Software for registration usually requires heavy human intervention, such as picking the landmarks or adjusting parameters that lack intuitive meaning.

We propose building an "active" atlas. In such an atlas, characteristics of annotated regions are modelled using machine learning. These models are then used to detect similar regions in new images. Detected regions serve as landmarks for both intra-specimen registration and specimen-atlas registration. Once the new images are aligned to the atlas, region models are updated based on characteristics of detected regions in the aligned new images. Essentially, the atlas "actively" adapts to registered specimens, and "actively" annotates new specimens for easier registration. In contrast to most existing work that either focuses on building an atlas by registering a single image series, or on aligning image series to an fixed atlas, the active atlas combines annotation, registration and atlas building into an iterative framework that is particularly suitable for processing a large number of specimens in an incremental fashion.

One crucial component in building an active atlas is identification of the landmarks. Until the region models are trained with sufficiently many examples, the system must obtain the annotation in other ways. Having human labellers do the annotation is time-consuming. In this paper, we propose an automatic approach.

Our particular study focuses on the brainstem. Unlike the cerebrum and the cerebellum, where neurons form high-contrast layered structures such as the cortex, the brainstem mostly contains nuclei, which are compact clusters of neurons. Neurons in each nucleus have similar size, shape and stain sensitivity, often also demonstrating uniform density and particular directionality. This gives each nuclei a distinctive texture, allowing them to be robustly identified when building the atlas. The fact that they are well localized also makes them perfect landmarks for registration. In addition to nuclei, fiber tracts are also recognizable structures in the brainstem. In particular, the directionality of a tract serves as an important feature of the landmark.

We therefore aim to model the landmarks based on texture. Section 3 describes how we represent textures using Gabor filters and superpixels. Section 4 explains how we identify potential landmark regions based on texture distinctiveness. Section 6 describes how we detect boundary segments based on regions in order to improve robustness. Section 7 describes how landmarks are matched between sections.

2 Related Work

Intensity based Registration Mostly intensity based

Point Landmark Detection
SIFT
Saliency and Objectness Detection
global rarity scheme
center-surround scheme
[YF: Is there work that uses notion of statistical significance in this context?]
Texture Representation
gabor filter
textons

3 Representing Texture using Histograms of Gabor Textons

Our algorithm starts with representing textures using responses of Gabor filters. Regions are detected based on similar textures. Regions with high contrast are selected.

Boundaries that are consistently part of region
We filter the image use Gabor filters.
rotation-invariant k-means clustering to form textons.
Over-segment into superpixels.
Describe texture using histogram of textons

4 Detecting Significant Region Using Center-Surround Contrast

[YF: Define the problem of finding regions+textures of high statistical significance.]

Brain anatomists usually label section images in terms of nuclei, which are dense groups of gray matter with the same phenotype. For example, Figure 1 shows the facial motor nucleus.

Given the texture descriptors obtained from Gabor textons, we define the saliency of a region in terms of statistical significance computed for texton distributions, as we will elaborate below.

[YC: We aim to find an algorithm that detects salient regions define saliency in terms of statistical significance. That is, salient regions should be those whose texton histogram is unlikely to be generated by those of the surrounds.]

First, region growing is performed on each superpixel. Starting from a region with one particular superpixel, the greedy procedure considers all superpixels that are neighbors of the current region, and iteratively adds the one whose texton distribution is most similar to the average distribution of the current region. Distance of texton distributions is computed using the Jenson-Shannon divergence, which is a symmetric measure.

Instead of using a pre-determined threshold on the distribution distance between the newly added superpixel and the current region average as the termination criteria, we over-grow the region (until 10% of the total area) while

recording the distances at every iteration, and eventually return the region when the distance is the largest in retrospect. In this way, a region automatically grows to the place where the interior-exterior contrast is the greatest.

We call the region that grows out of a seed superpixel the *expansion cluster* of that superpixel.

The significance of a region can be defined using a similar metric, that is the smallest distribution distance between any neighbor and the current region average. If the distance is the Kullback-Leibler divergence, then via Sanov's theorem, this value can be interpreted as the statistical significance of observing the interior texton histogram given that the closest neighbor's histogram is the true distribution.

5 Human Supervision

[YF: Describe how salient region detection improves efficiency of human labeling.]

6 Detecting Boundaries by Region Consensus

[YC: Major points for this section: (1) Boundaries are more robust landmarks than regions. (2) Why detect boundary segments? It is too crude to compute a single saliency value for a region. It is important to characterize saliency in different sides of the region, that is why we associate saliency value with boundary segments, rather than with entire regions. Also often regions are salient relative to each other, in this case using boundary is a more compact representation.]

Because region growing is not perfect, the expansion clusters of superpixels that belong to the same nucleus are often not identical. Figure x shows one such example. In most cases, boundaries are more robust landmarks than areas.

In addition to closed contours, partial boundaries are useful in cases when there is a clear boundary on one side of the nucleus where most clusters agree on, while on the other side no definitive boundary is present.

We aim to identify robust boundary segments that are supported by a large number of clusters. The strategy is to let the clusters vote for their boundaries using their respective significance scores. Superpixels receive high boundary votes if they are at the exterior of many salient clusters.

Instead of letting each superpixel as boundary separately, vote for them as a set.

Each region votes according to their saliency scores.

Each boundary is described by a tuple that consists of four elements:

1. x-y positions of every superpixel on the boundary
2. centroid of the expansion cluster inside the boundary
3. the average texton distribution of interior superpixels
4. a list of texton distributions of exterior superpixels (more precisely, the closest layer of superpixels on the outside of the boundary).

7 Matching Boundaries from Different Sections

distance = interior + exterior + shape + location

Distance between boundaries are defined as a weighted combination of:

1. Jenson-Shannon divergence between interior distributions

2. symmetric Hausdorff distances between the two sets of exterior distributions. That is, the maximum among the distances between each distribution and its closest distribution from the other set. Here the distance is the Jenson-Shannon divergence.

3. shape dissimilarity: total chi2-distances of shape context descriptors after correspondences are identified for superpixels on two boundaries using dynamic programming. (This is essentially the Shape Distance in Section 5.1 of Belongie's paper)

4. spatial distance: Euclidean distance between cluster centroids

Boundaries are detected from two sections and their pairwise distances are computed. A pair of boundaries are matched if they are the closest boundary of each other.

8 Experiments

8.1 comparison with human labelings

Shows the results of our algorithm is comparable to human labeling.

show results for RS141. (Do we need to do more than one stacks here? I think one stack already shows enough variations. A coronal stack would be good, but we don't have any now).

8.2 robustness of matching

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 Future Work

use detected landmarks for registration

Learn dictionary using deep neural networks, ICA, ...

References

1. Clarke, F., Ekeland, I.: Nonlinear oscillations and boundary-value problems for Hamiltonian systems. Arch. Rat. Mech. Anal. 78, 315–333 (1982)