Title
An Active Texture-based Atlas for Automated Mapping of Structures and Markers Across Brains

Permalink
https://escholarship.org/uc/item/5zt9d974

Author
Chen, Yuncong

Publication Date
2018

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

An Active Texture-based Atlas for Automated Mapping of Structures and Markers Across Brains

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Computer Science

by

Yuncong Chen

Committee in charge:

Professor Yoav Freund, Chair
Professor David Kleinfeld, Co-Chair
Professor Sanjoy Dasgupta
Professor David Kriegman
Professor Zhuowen Tu

2018
The Dissertation of Yuncong Chen is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____________________________________

____________________________________

____________________________________

____________________________________

____________________________________  Co-Chair

____________________________________  Chair

University of California, San Diego

2018
TABLE OF CONTENTS

Signature Page ......................................................... iii
Table of Contents .................................................... iv
List of Figures ........................................................ vi
List of Tables ........................................................ xi
Acknowledgements .................................................... xii
Vita ....................................................................... xiv
Abstract of the Dissertation ........................................ xv

Chapter 1  Introduction .................................................. 1
  1.1 Outline ............................................................. 5

Chapter 2  Data Preprocessing ......................................... 8
  2.1 Preparation of Serial Sections ................................. 8
  2.2 3-D Reconstruction ............................................... 9
  2.3 Generating Tissue Masks ....................................... 11

Chapter 3  Human Annotation ......................................... 13

Chapter 4  Building Reference Atlas ............................... 18
  4.1 Reconstructing 3-D Structures ............................... 18
  4.2 Estimating Nominal Locations ............................... 19
  4.3 Estimating Average Shapes ................................... 20

Chapter 5  Training Texture Classifiers ........................... 25
  5.1 Learning from Example Patches ............................. 26
  5.2 Contrasting with Surrounding Textures .................... 26
  5.3 Encoding Texture with Convolutional Neural Network .... 27
  5.4 Classification with Binary Logistic Regression .......... 28

Chapter 6  Texture Classification for a New Brain .............. 30
  6.1 Generating Structure Probability Maps .................... 30
  6.2 Results ............................................................ 31

Chapter 7  Registration .................................................. 36
  7.1 Estimating Global Transform ............................... 37
  7.2 Estimating Structure-specific Transforms ................. 39
  7.3 Estimating Deformation Field ................................ 39
LIST OF FIGURES

Figure 1.1: Existing atlases. .................................................. 2
Figure 1.2: Appearance of the brainstem in raw histological images versus in an average template. ............................................. 4
Figure 1.3: Overall structure of an active atlas. ............................ 6
Figure 2.1: Reconstruction of serial sections of a mouse brain. ............ 9
Figure 2.2: Reconstructed graylevel volume. .................................. 11
Figure 2.3: Generation of tissue masks. ....................................... 12
Figure 3.1: Program for creating structure boundary annotation. ........... 14
Figure 3.2: Automatic generation of structure boundaries on intermediate sections (green) by interpolating hand-drawn boundaries on a sparse set of sections (red). .............................................. 15
Figure 3.3: Example expert-annotated section .................................. 15
Figure 4.1: Reconstruction of an annotation brain. ............................ 19
Figure 4.2: Estimation of standard locations. ................................... 20
Figure 4.3: Estimation of average shapes. ....................................... 22
Figure 4.4: The reference atlas. .................................................. 23
Figure 5.1: Selected image patches in an annotated section that are used to train the texture-based binary classifiers. .............................. 26
Figure 6.1: Texture-based detection of structures. .............................. 30
Figure 6.2: Classification performance. .......................................... 31
Figure 6.3: A demonstration of the advantage of using high-resolution texture. ......................................................... 32
Figure 6.4: Localization performance of classifiers trained using CNN texture features versus using mean gray level. .................................. 33
Figure 7.1: The registration process. .............................................. 36
Figure 7.2: Structure annotations derived from global registration and structure-specific registration for an example section. ................. 38
Figure 7.3: Deviation of the centroids of registered structures from expert annotation, normalized by the size of corresponding structure. ........... 41
Figure 7.4: Root-mean-squared error between registered probabilistic structures and expert annotations. ............................................ 41
Figure 7.5: Measures of registration confidence. ............................... 44
Figure 7.6: Comparison between texture-based method and graylevel-based method for registration. ................................................. 45
Figure 7.7: Comparison between texture-based method and graylevel-based method for atlas-based annotation. ..................................... 45
Figure 8.1: Twelve brains co-registered to atlas using global transforms only. . . . 49
Figure 9.1: Intensity normalization between fluorescent Neurotrace and bright-field thionin. 54
Figure 9.2: Visualization of the labeling motor and premotor inputs to the jaw muscle across all three planes. 55
Figure 9.3: Visualization of the labeling of different populations of premotor neurons in separate brains with overlapped density in the PCRt. 56
LIST OF TABLES

Table 3.1: List of annotated structures ........................................... 17
I would like to thank Professor Yoav Freund and Professor David Kleinfeld for providing academic guidance and financial supports throughout my PhD. I also thank Professor Thomas Bewley of UCSD’s Coordinate Robotics Lab for supporting and advising on my earlier works in robotics.

I appreciate Alex Tolpygo, Daniel Ferrente (Cold Spring Harbor Labs) and Lauren McElvain, Beth Friedman and Hannah Letchty (Kleinfeld Lab) for providing histological data of the highest quality. I also appreciate Professor Harvey Karten and Lauren McElvain for providing anatomical annotations with amazing patience.

I thank all team members in Kleinfeld lab for helpful discussions and moral support. I am also grateful to be accompanied by supportive friends Yinghao Chu, Dongjin Song, Duo Song and Lan Yu, who enlightened my life outside of the office. My deepest gratitude goes to my parents who never stop encouraging me and offering love in the most selfless way.

Chapter 1 through Chapter 10, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld, David; Freund, Yoav, Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI 2017. Lecture Notes in Computer Science, vol 10433, pp. 3-11, 2017, Springer.

Chapter 1 through Chapter 10, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The
dissertation author was the primary author of this material.
VITA

2010 Bachelor of Engineering, The Hong Kong University of Science and Technology, Hong Kong

2014 Master of Science, University of California, San Diego

2018 Doctor of Philosophy, University of California, San Diego

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

An Active Texture-based Atlas for Automated Mapping of Structures and Markers Across Brains

by

Yuncong Chen

Doctor of Philosophy in Computer Science

University of California, San Diego, 2018

Professor Yoav Freund, Chair
Professor David Kleinfeld, Co-Chair

An anatomical atlas constitutes a quantitative description of the structural organization of an organism. Nissl-stained serial brain sections reveal cellular texture (cytoarchitecture) and are the gold standard for defining structures. As high-throughput techniques advance and automated registration becomes commonplace, the role of an atlas grows beyond a static map towards a repository where a wide variety of experiment data acquired from different subjects can be integrated in a common template defined in a standard coordinate system. For histological data (0.5 μm resolution), however, the reliance of intensity-based
registration algorithms on downsampled pixel/voxel intensities (> 10 µm) and the resulting negligence of fine-scale textures means they often fail to accurately align structures whose boundaries are defined by cytoarchitecture rather than grayscale, such as many brainstem nuclei. The lack of a registration tool that strongly utilizes texture and a nucleus-level atlas to assist such registration has stifled comparison of results across experiments.

We demonstrate a data-driven active atlas system that automatically aligns brains based on cytoarchitectural landmarks. Our approach combines discriminative texture detectors based on convolutional neural network (CNN) features with a reference atlas that describes structure shapes as probabilistic volumes and their locations as Gaussian distributions. Histological serial sections are reconstructed in 3-D and converted to structure probability maps by classifiers trained to differentiate the texture inside versus outside each structure. Registration is achieved by maximization of the correlation between the probability maps and the reference atlas using a global affine transform followed by deformation interpolated from structure-specific rigid transforms. Initialized from annotation by expert neuroanatomists, the atlas is continuously refined after incorporation of new brains in a semi-supervised fashion.

Based on automated registration of twelve specimens, we developed an atlas for adult murine brainstem which defines 28 distinct structures in both hemispheres. The system’s utility in advancing brain circuitry study is demonstrated by the precise mapping of neuronal projections in cytoarchitecturally ill-defined regions across brains from different animals. Quantitative results showed that our approach produced accurate and confident registration and significantly reduces human labor.
Chapter 1

Introduction

**Histological serial sections** are among the most popular approaches to study the brain. The procedure usually involves cutting a brain into slices as thin as 10 microns, staining the slices with chemicals (such as thionin or Neurotrace) and imaging under microscopes (either in transmission or fluorescence) at sub-micron resolution. This allows the tissues to be visualized with cell-level details. In a histological image, one can clearly see regions with distinct textures (Figure 1.2a). The different textures are formed by cells of different shapes and in different organizations. The neuroscience term for these textures is **cytoarchitecture**. The study of cytoarchitecture dates back to the 1900s, when Korbinian Brodmann [14] used the texture to define structures in the brain. As each of these texturally homogenous structures tends to contain neurons that share similar projection pattern and histochemical properties, they are often regarded as nodes of the brain circuitry, also known as **nuclei**.

An **atlas** is a quantitative map of these structures. It describes a standard brain where different structures are defined in a standard coordinate space. An atlas often also provides reference images that are registered with the structure annotations. A user can then compare
Figure 1.1: Existing atlases. Left: One figure from the classical mouse brain atlas by Franklin and Paxinos. The sagittal part of this atlas was created by manually annotating Nissl-stained series of one brain. Right: Allen mouse brain atlas, common coordinate framework v3. This is a 3-D average intensity template derived from co-registration of the serial two-photon tomography images of more than 1,600 brains.

Her own images with the reference images to locate structures of interest. Traditionally, an atlas is built by manually annotating one brain. One example is the atlas by Franklin and Paxinos [48] (Figure 1.1a). Recently, due to the progress of high-throughput imaging techniques and automated registration algorithms, a large number of brains can be used to build an atlas. For example, the Allen mouse brain atlas [2] was constructed using more than 1,600 brains (Figure 1.1b), whose 3-D volumetric images acquired by serial two-photon tomography are co-registered to create an average intensity template.

The main computational issue this thesis aims to address is the registration between an atlas and an individual brain, or in other words, the establishment of a mapping between atlas coordinates and coordinates on the images of the serial sections. Successful registration simultaneously achieves two goals. On the one hand, by transforming the atlas to fit the subject, the segmentation or annotation defined in the atlas can be propagated to the images, allowing the interpretation of image data with anatomical context. This is known as atlas-
based annotation. On the other hand, by mapping a subject to the atlas, data on the subject can be placed in the standard space of the atlas. A large variety of experimental data from different brains (e.g. tract tracing, patterns of molecular marker, genetic expression) can therefore be integrated and analyzed under the same anatomical framework. In this sense, the role of an atlas has grown from a static visual guide to an ever-expanding spatial data repository. Developing a software system that facilitates the realization of this vision is the fundamental motivation for this project.

An active atlas is a software system for the automated registration of new brains to a reference atlas and the continuous refinement of the atlas. It consists of three components.

• A reference atlas that describes the nominal location and shape of structures defined by neuroanatomists in a standard coordinate system. New structures or new subdivisions of existing structures can be added.

• A cost function associated with variation around the nominal locations that describes biological plausibility of different variations. Variation of a particular structure is described by a spatial transform (possibly composed of multiple transforms).

• A measure of the goodness-of-fit between a brain and the transformed atlas.

With an active atlas, the registration of a new brain is automated by computing transforms that maximize the goodness-of-fit with the reference atlas while minimizing variation cost. Following a human verification step, the aligned new brain may be used to refine the reference atlas.

Traditionally, registration of histological data (0.35-0.5 µm per pixel) uses the same methods that were designed for registering magnetic resonance imaging (MRI) volumes which have far lower resolution (10-100 µm voxel size). These methods are based on
maximizing the similarity of voxel-based greylevel intensities, measured by normalized correlation coefficient or mutual information. Such measures only work in low resolution because as resolution increases the greylevel patterns can no longer be correlated across images. For histological data, this means that the images must be heavily downsampled before registration. The resulting smoothing away of the textures is particularly problematic for regions (such as the brainstem) that lack high-contrast macroscopic boundaries and must depend on high-resolution cytoarchitectural details to identify and align structures (Figure 1.2). Despite the scientific importance of brainstem, the lack of a registration tool that strongly utilizes texture and a nucleus-level atlas to assist such registration has stifled comparison of results across experiments.

In this thesis we describe an active atlas for the mouse brainstem based on texture detection (Figure 1.3). A supervised approach is adopted to create texture classifiers and to bootstrap a reference atlas. They are initialized from limited expert annotations and
iteratively refined through the alignment of a large number of unannotated brains. A set of structure probability maps generated by the classifiers for any new brain are compared with a transformed reference atlas to define the goodness-of-fit. This is accompanied by the variation cost that is adapted to the variability of previously registered brains to drive registration. Results on twelve brains showed accurate and confident registrations that allow human verification to take significantly less time than creating comparable annotations from scratch. We further demonstrate the scientific utility with an application in precise mapping of neuronal projections in cytoarchitecturally ill-defined regions across brains from different animals.

In this project we focus specifically on the brainstem. Brainstem is located in the posterior of the brain. It contains the main neural circuitry for sensorimotor processing, and plays a fundamental role in the regularization of orofacial behaviors such as breathing and whisking. Nonetheless, our approach is general and applicable to other parts of the mouse brain as well as to brains of other species.

1.1 Outline

We start in Chapter 2 with a description of the data and the procedure to reconstruct a new brain from serial sections. Then we explain each component of the active atlas system in sequence. Chapter 3 describes the manual labeling process. Based on expert labelings, Chapter 4 describes the process of building reference atlas and Chapter 5 describes training of the texture classifiers. Chapter 6 through 8 describes the process of incorporating a new brain. Chapter 6 addresses the detection and generation of structure probability maps. Chapter 7 addresses the registration with the reference atlas. Chapter 8 describes how the reference atlas is updated by the registration of this new brain. In Chapter 9 we demonstrate
Figure 1.3: Overall structure of an active atlas. (a,b) Inputs to the system are histological sections from a new brain, in these examples sagittal sections of a mouse brain with Nissl background staining, exemplified by thionin stained cells for brightfield data (panel a) and Neurotrace Blue stained cells for fluorescent data (panel b); the brain in panel b are also labeled with cell markers. (c) Texture classifiers. They are discriminative models responsible for localizing different structures from input brains. (d) Detection of texture generates a set of structure probability maps. (e) The reference atlas. It describes the shape and position of a set of anatomical structures. (f) Registration between atlas and input brains. (g) Successful registrations allow the soft annotation of any individual input brain, as well as the integration of data from all input brains. (h) Variations captured by the registrations can further refine the reference atlas and the soft annotations can be used to improve texture classifiers. (i) This semi-supervised learning loop is initialized by a small amount of expert-annotated brains.
the utility of the active atlas in mapping fluorescent markers. Finally, Chapter 10 reviews existing approaches for atlas construction and automatic annotation, which highlights the novelty of our use of texture for registration.

This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld, David; Freund, Yoav, Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI 2017. Lecture Notes in Computer Science, vol 10433, pp. 3-11, 2017, Springer.

This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Chapter 2

Data Preprocessing

2.1 Preparation of Serial Sections

Our main dataset are Nissl-stained sagittal serial sections of the brains of twelve P56 C57BL/6 male mice (Figure 2.1a). The brains were perfused, fixed and cryo-sectioned before being stained with thionin, which is a gold standard dye for defining brain structures based on cytoarchitecture.

The sections are prepared with a customized Cryo-Jane Tape-Transfer System [50, 1]. The system uses adhesive coated slides and adhesive tapes to capture sections from the tissue block instead of using an anti-roll brush as in conventional methods. This ensures the tissues are wrinkle-free and undistorted. These high-quality sections are crucial for the accurate 3-D reconstruction. A complete series from one brain contains around 400 sections, each 20 µm thick. The brainstem spans 250-300 sections.

Slides are scanned with Hamamatsu NanoZoomer at a planar resolution of 0.45 µm per pixel. The full-resolution image of each section has a dimension around 25,000 pixels \( \times \) 15,000 pixels. A 32-fold downsampled thumbnail is generated for every image (\( \approx 15 \).
Figure 2.1: Reconstruction of serial sections of a mouse brain. (a) Illustration of a mouse brain and its sagittal serial sections. Brainstem is located in the posterior part of the brain. (b) Aligned sections in the 3-D subject space.

µm per pixel). Since the thionin stain is largely monochrome, we converted the images to grayscale for any subsequent processing.

2.2 3-D Reconstruction

To reconstruct the 3-D anatomy of a brain, the serial sections need to be aligned. Because there is no noticable distortion to the tissues, we achieved good alignment using
in-plane rigid transforms. Using low-order transforms also means low-resolution thumbnails are sufficient. The small section spacing (20 µm) leads to high correlation between the appearance of consecutive sections, which facilitates registration.

A 2-D rigid transform is computed between every pair of consecutive sections using Elastix [36] by maximizing the mutual information of the distributions of pixel intensities [41]. The largest section in the series is selected as the anchor. For each remaining section, we compose the list of pairwise transforms between this section and the anchor section to derive a transform that directly registers it to the anchor section. Scaled-up versions of these transforms are then applied to the full-resolution images to generate co-aligned images that will be used for expert annotation.

We then define a 3-D coordinate space (coordinates are in µm) and place each section at the plane \( z = \text{sectionNumber} \times 20 \mu\text{m} \). This space is the subject space of this brain (Figure 2.1b). A mapping is thus established between 2-D coordinates of any image and 3-D locations in the subject domain. This mapping is used by many parts of the system, for example, to reconstruct detection maps in 3-D and to convert aligned atlas to 2-D contours on individual sections.

Figure 2.2 shows the reconstructed graylevel volume of a brain. It is formed by mapping the graylevel value of each image pixel into the subject space. One can dissect this volume in an arbitrary plane. From the virtual slices in the coronal and horizontal planes one can clearly identify structures such as the hippocampus. The smoothness of the structure boundaries demonstrates the high quality of reconstruction.
Figure 2.2: Reconstructed graylevel volume (voxel = 16 µm). Smoothness of structure boundaries on the coronal virtual sections demonstrate the quality of intra-series registration.

2.3 Generating Tissue Masks

Tissue masks allow texture classification to ignore the background, which often takes more than half of an image, saving computation time. For every section we manually specify an initial contour that loosely encircles the tissue and used morphological snake [3] to evolve the contour towards the tissue (Figure 2.3). After verifying that no debris is mistaken as tissue, a binary mask is generated for the section.

This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld, David; Freund, Yoav, Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI 2017. Lecture Notes in Computer Science, vol 10433, pp. 3-11,
**Figure 2.3**: Generation of tissue masks. (a) Section image with initial (red) and final (green) snake contours. (b) Resulting binary mask.

2017, Springer.

This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Chapter 3

Human Annotation

To bootstrap the iterative atlas building process, we used expert annotations of a small number of brains to create the first version of a reference model and texture classifiers. We randomly selected three brains out of the twelve. Two experienced neuroanatomists were asked to draw boundaries of major brainstem structures on full-resolution images of these three brains.

Annotation was performed using an in-house program designed to maximize efficiency and user-friendliness (Figure 3.1). The program loads the aligned full-resolution series. One can zoom in and out to observe cytoarchitectural textures at different magnifications and quickly browse through different sections. Structure boundaries are represented by closed polygons and their vertices were recorded.

Because some structures are easier to identify on planes of orientation other than sagittal, the program shows a reconstructed graylevel volume of the brain and allows one to navigate through it in any of the three orthogonal planes of orientation. Synchronization of
**Figure 3.1**: Program for creating structure boundary annotation. Main panel shows the original full-resolution section image. Side panels show virtual sections of the reconstructed graylevel volume in three orthogonal planes. All panels can be synchronized. Structure contours are shown on all panels.

All the panels allow one to simultaneously examine the three virtual planes passing through any 3-D location. Real-time reconstruction and resectioning of structures allow structure contours to be shown on all panels.

Annotation of large structures (e.g. spanning more than 40 sections) can be tedious if one was to annotate every single section. Since most structures change shape smoothly from one section to the next, we allow an annotator to annotate only one in every 5-10 sections and use linear interpolation to generate boundaries on intermediate sections (Figure 3.2). One can choose to refine an interpolated boundary, which still takes less time than drawing one from scratch.
Figure 3.2: Automatic generation of structure boundaries on intermediate sections (green) by interpolating hand-drawn boundaries on a sparse set of sections (red).

The whole process took around 30 hours. The results are 1300 contours of 28 structures (51 for both sides). 50% of all contours are created by interpolation and manually verified. The full list of structures is in Table 3.1. Figure 3.3 shows an example of the annotated sections.

Figure 3.3: Example expert-annotated section
This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld, David; Freund, Yoav, Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI 2017. Lecture Notes in Computer Science, vol 10433, pp. 3-11, 2017, Springer.

This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Table 3.1: List of annotated structures

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>3N</td>
<td>oculomotor nucleus</td>
</tr>
<tr>
<td>4N</td>
<td>trochlea nucleus</td>
</tr>
<tr>
<td>5N</td>
<td>trigeminal motor nucleus</td>
</tr>
<tr>
<td>6N</td>
<td>abducens nucleus</td>
</tr>
<tr>
<td>7N</td>
<td>facial motor nucleus</td>
</tr>
<tr>
<td>7n</td>
<td>facial nerve</td>
</tr>
<tr>
<td>10N</td>
<td>dorsal nucleus of vagus nerve</td>
</tr>
<tr>
<td>12N</td>
<td>hypoglossal nucleus</td>
</tr>
<tr>
<td>Amb</td>
<td>nucleus ambiguus</td>
</tr>
<tr>
<td>AP</td>
<td>area postrema</td>
</tr>
<tr>
<td>DC</td>
<td>dorsal cochlea nucleus</td>
</tr>
<tr>
<td>LRt</td>
<td>lateral reticular nucleus</td>
</tr>
<tr>
<td>LC</td>
<td>locus corelulus</td>
</tr>
<tr>
<td>IC</td>
<td>inferior colliculus</td>
</tr>
<tr>
<td>VCA</td>
<td>ventral cochlea nucleus, anterior</td>
</tr>
<tr>
<td>VCP</td>
<td>ventral cochlea nucleus, posterior</td>
</tr>
<tr>
<td>VLL</td>
<td>ventral lateral lemniscus</td>
</tr>
<tr>
<td>PBG</td>
<td>parabigeminal nucleus</td>
</tr>
<tr>
<td>Pn</td>
<td>pontine grey</td>
</tr>
<tr>
<td>R</td>
<td>red nucleus</td>
</tr>
<tr>
<td>RtTg</td>
<td>reticulotegmental nucleus</td>
</tr>
<tr>
<td>SC</td>
<td>superior colliculus</td>
</tr>
<tr>
<td>Sp5C</td>
<td>spinal-trigeminal nucleus, caudalis</td>
</tr>
<tr>
<td>Sp5I</td>
<td>spinal-trigeminal nucleus, interpolaris</td>
</tr>
<tr>
<td>Sp5O</td>
<td>spinal-trigeminal nucleus, oralis</td>
</tr>
<tr>
<td>SNR</td>
<td>substantia niagra, reticular</td>
</tr>
<tr>
<td>SNC</td>
<td>substantia niagra, compact</td>
</tr>
<tr>
<td>Tz</td>
<td>nucleus of trapezoidal body</td>
</tr>
</tbody>
</table>
Chapter 4

Building Reference Atlas

The reference atlas describes the nominal anatomy of a mouse brain in terms of the shape and centroid (center of mass) location of a set of cytoarchitecturally recognized structures defined in a standard coordinate system. To avoid the bias of any single brain, we used the three annotated brains to derive probabilistic summaries of these properties. Specifically, we compute for each structure, the average shape and the mean and covariance of centroid location.

In itself, the reference atlas is a quantified description of the organization of the mouse brain. With new brains, the reference atlas provides a template for the spatial normalization of the brains and can be used to create annotations on individual brains.

4.1 Reconstructing 3-D Structures

For every structure in each brain, we mapped the 2-D contours to the 3-D subject space (Figure 5.1), and linearly interpolated the contour series to create a smooth surface. By resampling the interior of the surface, a 3-D volume of this structure is reconstructed.
4.2 Estimating Nominal Locations

The brains are co-registered under a common coordinate space, which we call the atlas space. One brain is selected as the target and the other two brains are aligned to it. Alignment of two brains begins with aligning the mid-saggital planes, which were estimated by fitting a plane to the midway points of the centroids of structures that are in pairs. Under this constraint, we computed an affine transform that maximizes the correlation between the two brains (Figure 4.2a). The exact algorithm is a special case of that used for subject-atlas registration (Chapter 7).

Once all brains were co-registered in the atlas space, we compute for each structure a nominal location, which is defined as the location closest to the cross-brain mean centroid location of this structure, while satisfying the constraint that the nominal location of any singular structure is exactly on the mid-sagittal plane and those of any paired structure are symmetric about the mid-sagittal plane (Figure 4.2b)
Figure 4.2: Estimation of standard locations. (a) Illustration of three annotated brains (blue, green, red) brought in registration in atlas space by global affine transforms. (b) Top-down view of the atlas space. Each structure is represented by a color. Circles are instance centroids. Stars are the nominal centroids. Shaded plane is the common mid-sagittal plane. Note the symmetry of the nominal centroids of paired structure with respect to the mid-sagittal plane.

We also computed for each structure the covariance of the centroid location. It is used in subject-atlas registration to define the cost of deviation from the nominal centroid (see Equation 7.3).

Let $c_s \in \Omega_A$ denote the nominal location of structure $s$ and $\Sigma_s \in \mathbb{R}^3$ denote the covariance matrix. These are updated as new brains are incorporated (see Chapter 8).

### 4.3 Estimating Average Shapes

The average shape of each structure $s$ is a 3-D probabilistic volume, denoted by $M_s : \mathbb{R}^3 \rightarrow [0, 1]$, with the origin of the domain located at the center of (probability) mass of the volume.

The average shape of a structure is derived from a set of reconstructed instances
from both sides of all annotated brains. For structures in pairs, this set includes all the left-side instances and the mirror images of all right-side instances. For singular structures, this set additionally has the mirror image of each original instance. The set of instances were aligned using rigid transforms that maximize overlap and then resampled into volumetric representations in a common domain. We then derive the mean volume by computing the fraction of instances that contain each voxel. We did not further compute the variance because the number of instances (6 in most cases) is too few for finding meaningful variation modes. Figure 4.3 illustrates the results for facial motor nucleus. For a paired structure, the shape thus derived is for the left side, which is then mirrored to give the shape at the right side. For singular structures, a symmetric shape is derived by averaging with its own mirror image.

Unlike the nominal locations, the shapes are fixed after this initialization, and are not updated as new brains are incorporated. This is because our goal is to localize structures for the purpose of registration rather than to generate precise structure boundaries. Using fixed shapes are adequate for this purpose. Furthermore, the current number of training brains is not sufficient to learn generalizable shape models.

Figure 4.4 provides an illustration of the complete reference atlas.

This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld, David; Freund, Yoav, Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI 2017. Lecture Notes in Computer Science, vol 10433, pp. 3-11,
2017, Springer.

This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Figure 4.3: Estimation of average shapes. (a) Reconstructions of three annotated brains with facial motor nucleus in both sides highlighted. (b) All six instances of facial motor nucleus. (c) Instances aligned using rigid transforms. (d) Probabilistic average shape obtained by voxel averaging.
Figure 4.4: The reference atlas. (a) Illustration of reference atlas, showing the average shapes of all structures at respective nominal locations, and the brain outline (translucent). (b) Sagittal view of the atlas structures, labeled with acronyms. Shown for each structure is the $p=0.5$ iso-surface of the average shape.
Chapter 5

Training Texture Classifiers

In order to define a measure of goodness-of-fit on the basis of texture, we use texture classifiers to convert a brain into a set of structure probability maps. Each map conveys the probability that a particular structure exists at each location of the subject space. This represents the same type of anatomical information as the reference atlas, therefore they can be directly compared and, as we will describe in Chapter 7, the goodness-of-fit can be simply defined by the correlation.

Specifically, for each structure \( s \in S \), we define a classifier \( F_s \) that maps an image window to a value between 0 and 1. For each brain, we map the classifier output of each image window to a 3-D specimen domain location that is associated with the center of the image window. Each classifier \( F_s \) thus produces a structure probability map, defined on the subject space. Denote the probability map for structure \( s \) by \( D_s : \Omega_B \rightarrow [0, 1] \).

Note that the generation of structure probability maps is solely based on texture in images and is independent of any locational information. In this chapter we present the motivation for the design of the classification framework and describe how to train the classifiers using expert annotations.
5.1 Learning from Example Patches

The principle of cytoarchitectonics defines structures on the basis of texture. Textures of stained brain cells have rich delicate features that are negligible by untrained eyes and yet reveal properties of the tissue critical for structure differentiation. Neuroanatomists, through years of image inspection experience, have developed the visual acuteness to pick up these features and determine if they are indicative of particular structures. Similarly, we train classifiers using a large number of example images of various structures from many brains.

We used image patches as the basic unit of classification. Patch size is chosen to be $100 \mu m \times 100 \mu m$ ($224 \times 224$ pixels), which provides enough context to capture local cell organization and sufficient details to reveal the morphology of individual cells. Patches are collected based on a sliding window with a stride of $30 \mu m$ ($\approx 64$ pixels) that yields more than 50,000 patches per image.

5.2 Contrasting with Surrounding Textures

It is common for a neuroanatomist to characterize the texture of a structure by contrasting it with adjacent structures. For example, Sp5I are defined as having smaller neurons than Sp5O. Similarly, training a sensitive classifier requires not just images of the given structure, but also negative examples.

Specifically, we aim to differentiate the texture of each structure from that of its immediate surrounding. Since the reference model usually provides a strong estimate for the rough location of most structures, the goal of the classifiers is to provide fine-scale localization, which cannot be optimal if they are trained against a much more diverse background.
Training patches for a given structure are collected from all sections it is annotated on. A patch is labeled positive if three of the four corners are located inside a boundary of this structure. A negative patch must have three corners in a 500 µm-wide bordering zone of a boundary.

Figure 5.1: Selected image patches in an annotated section that are used to train the texture-based binary classifiers. Positive patches (green) are extracted from the interior of boundaries, and negative patches (red) are extracted from a 500 µm wide moat surrounding the landmark.

5.3 Encoding Texture with Convolutional Neural Network

Each image patch is encoded by a deep convolutional neural network (CNN) into a feature vector. A CNN encodes an input image through a series of convolution and pooling operations. Using back-propagation, a network with sufficiently many layers can learn a simple-to-complex hierarchy of latent visual elements from training images.

We used the Inception-BN network [?] (implemented by MXNet[15]), which is
a 27-layer CNN trained for multi-class classification on a subset of the natural image
dataset ImageNet [20] and achieved state-of-the-art performance. Despite being trained
on completely different types of images, we believe that the filters learned by the network
capture visual patterns equally applicable to histological textures. We used the 1024-
dimensional vector that feeds into the last fully-connected layer as feature. We also modified
the first layer to accommodate single-channel instead of RGB inputs.

5.4 Classification with Binary Logistic Regression

The feature vectors were used to train binary logistic regression classifiers (imple-
mented by Python scikit-learn). We choose to use logistic regression because it is fast
and performs comparably with more complex alternatives. A maximum of 15,000 positive
patches and 15,000 negative patches are sampled from all three annotated brains to train
each classifier.

This chapter, in part, have been published in the following proceeding. The dis-
sertation author was the primary author of this paper. “The Active Atlas: Combining
3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.;
Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld,
David; Freund, Yoav, Proceedings of Medical Image Computing and Computer Assisted
Intervention, MICCAI 2017. Lecture Notes in Computer Science, vol 10433, pp. 3-11,
2017, Springer.

This chapter, in part, is also currently being prepared for journal submission. Chen,
Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth;
Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author
was the primary author of this material.
Chapter 6

Texture Classification for a New Brain

6.1 Generating Structure Probability Maps

Given a new brain, we applied the full set of classifiers to a moving window on every section. The window has the same size as the training samples. The spacing is 30 µm (≈ 64 raw pixels). The output of each classifier for every window is mapped to the location in the specimen domain that corresponds to the window center, forming a sparse 3D probability map. This was then upsampled using cubic interpolation and discretized to create a dense probability map with 10 µm isothropic voxels. Figure 6.1 shows the results for three structures. The set of probability maps for all structures constitute a textural description of the input brain and serve as the basis for registration with the atlas.

One can regard each voxel as having a structure probability vector that represents the “soft” identity of this voxel. We argue that such vectors, as texture descriptors, are superior to the voxel intensity values, in terms of discriminating power.

We chose to construct the dense probability maps at a relatively low resolution (10 µm) so that all of them can be simultaneously loaded into the computer memory as
required by the registration algorithm. In practice, this resolution is sufficient for resolving all structures of interest. We also chose to use a large spacing for the moving window (30 \( \mu m \)) and then upsample the sparse maps to match the dense map resolution, because this reduces the number of patches whose CNN features need to be computed. The feature computation, even with GPU, is a time bottleneck. The typical extent of a section is 10 mm \( \times \) 5 mm. Consider only patches on the tissue mask, a spacing of 30 \( \mu m \) results in roughly 50,000 patches. These design choices allowed us to compute the probability maps for one section in about one minute.

### 6.2 Results

Since the ground truth is only available for three brains, we evaluate the classification performance using a leave-one-out strategy. For each structure, we evaluate three classifiers, each trained on two of the three brains and tested on the remaining brain. For each classifier, as many as 15,000 patches are randomly sampled from all sections of the training brains and 1,000 patches from the test brain.

The results are reported as receiver operator characteristic (ROC) curves. An ROC curve plots false positive rate against true positive rate as the threshold on the output of the classifier increases from 0 to 1 (Figure 6.2a). The area under the ROC curve (AUROC) measures the overall classification accuracy, which is 0.5 for chance and 1 for perfect accuracy.

We also compared with other texture representations including:

- Haralick features based on the greylevel co-occurrence matrix (GLCM) with 10 grayscale levels.
Figure 6.1: Texture-based detection of structures. (a) An unannotated set of Nissl-stained sections from a new brain. (b) One example patch that is passed through a CNN to be converted into a texture feature vector $X$. (c) Example of three of the 28 classifiers for each landmark that are applied to all texture feature vectors in the brain patches across the entire brain. This results in a probability map for each landmark, illustrated here for one section and throughout the brain. The cluster of probability mass indicated by each arrow corresponds to the relevant structure.

- Local binary patterns (LBP), using a radius of 3 pixels and 24 circularly symmetric neighbour set points.
- Average gray level of a 25µm window at patch center.

Figure 6.2b shows the results for each structure. CNN features achieve a median AUROC of 0.95 and have clear advantage over the alternatives.

The benefit of classifiers with strong texture differentiation is particularly evident in low-contrast areas. In Figure 6.3, the smoothed image does not have clear contrast
Figure 6.2: Classification performance. (a) ROC curves for the classifiers of different structures (for one particular combination of train/test brains). (b) Classification performance in terms of the area under ROC curve with different features. For most structures, CNN features (blue) achieved higher accuracy than LBP (orange), GLCM (green) and average gray levels (red).

between the three structures, while the probability maps generated by our classifiers precisely localized each structure. In Figure 6.4, compare to classifiers trained using average graylevel, our classifiers give significantly stronger score signal, which is critical for a precise and robust registration.

This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.;
**Figure 6.3**: A demonstration of the advantage of using high-resolution texture. Top left: high resolution image of a region that contains three cytoarchitecturally distinct structures. Top right: downsampled version lacks the fine-scale details needed to distinguish these structures. Bottom: probability maps for the three structures demonstrate clear differentiation and precise localization.


This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Figure 6.4: Localization performance of classifiers trained using CNN texture features versus using mean gray level. (a) Image of the forebrain and midbrain of a single section stained with thionin. The boxes contain two motor nuclei: the oculomotor nucleus (3N) and the hypoglossal nucleus (12N). (b,e) The area around each nucleus at higher magnification and a resolution of 0.5 μm. The nucleus appears distinct from its surrounding area. (c,f) The same areas as in panels b and e after downsampling to 25 μm of resolution. The semi-homogeneous intensity within a landmark is required by traditional alignment methods, but precision is sacrificed by the blurred boundaries. (d,g) Comparison of the structure probabilities obtained from the texture-based classifier versus the greylevel-based classifier along the lines across the nuclei in panels b to e. All values are normalized. Texture-detection results in a steeper plateau to yield a precise alignment.
Chapter 7

Registration

In Chapter 4 we described how to construct the reference atlas. In Chapter 6 we described how to convert a subject brain to structure probability maps. In this chapter we address the registration of the two.

We assume that the variation of a specimen from a nominal brain is a combination of extrinsic and intrinsic variations.

Extrinsic variations include (1) physical deformation of the tissues during the histology procedures, in particular, an overall scale change; (2) inconsistent angle of sectioning between brains which results in shearing; (3) bias introduced in the 3-D reconstruction, including the particular pose and cropping, which can be modeled by rotation and translation. All these effects can be captured by a global 3D affine transform.

Intrinsic variations are the biological particularities inherent to a specimen, such as the shape and position of a structure. In this work we only study the variability in structure position. We assume the variations of different structures are independent, and capture that of each with a separate rigid transform. The deformation of the space between structures is determined by the rigid transforms of nearby structures.
Let $G = [A|b]$ denote the global affine transform where $A \in \mathbb{R}^{3 \times 3}$ is a linear transform matrix that corresponds to scaling, rotation and shearing; $b \in \mathbb{R}^3$ is the translation component. Let $L_s = [R|t]$ denote the rigid transform specific to structure $s$ where $R \in \mathbb{R}^{3 \times 3}$ is a rotation matrix and $t \in \mathbb{R}^3$ is the translation component. Let $E : \Omega_A \to \Omega_S$ denote the deformation field derived by interpolating $\{L_s|s \in S\}$. The mapping from the specimen domain to the atlas domain is then formulated as the composition of the global transform $G$ and the deformation field $E$. Detailed definitions of these transforms are in Appendix A. Figure 7.1 illustrates the registration process.

### 7.1 Estimating Global Transform

The estimate for $G$ is the maximizer of the goodness-of-fit function $Q_{\text{global}}$, which is defined as the correlation between the probability maps of the specimen and the reference atlas over voxels of every structure:

$$
Q_{\text{global}}(G) = \sum_s \sum_{p \in \Omega_A^s} D_s(G(p)) \cdot M_s(p - c_s).
$$

(7.1)

where $D_s$ is the probability map for structure $s$, $M_s$ is the average shape, and $c_s$ is the nominal location.

This function is reminiscent of Normalized Correlation Coefficients (NCC), a commonly used metric in intensity-based registration methods. It is defined as

$$
NCC = \frac{\sum_{x \in \Omega_A} (A(x) - \bar{A}) \cdot (B(G(x)) - \bar{B})}{\sqrt{\sum_{x \in \Omega_A} (A(x) - \bar{A})^2} \cdot \sqrt{\sum_{x \in \Omega_A} (B(G(x)) - \bar{B})^2}}
$$

(7.2)

where $A$ is the fixed image and $B$ is the moving image, $\bar{A}$ and $\bar{B}$ are the average grey-values.
**Figure 7.1:** The registration process. (a) The atlas after global affine alignment to the probability maps for all landmarks. (b, c) Local alignment between individual landmarks in the new brain with those in the atlas in 3-dimensions (panel b) and for one section superimposed on the classifier scores (panel c). The thin colored mesh is the initial position and thick black mesh is the final position. Contours are cross-sections of $p = 0.5$ iso-surfaces of transformed nominal shapes. (d) Illustration of the final aligned result. Greyscale image volume is the reconstruction of the Nissl sections. Colored structures are the transformed reference atlas. (e) Contour lines from the aligned reference atlas overlaid on the section in Figure 6.1a.
The main difference between the two functions is the appearance descriptor used for each voxel. $Q^{global}$ replaces the normalized grey-values in NCC by the structure probability vectors which have greater discriminative power.

### 7.2 Estimating Structure-specific Transforms

In addition to maximizing the goodness-of-fit, the transform specific to each structure must be regularized to prevent unrealistically large deviation from the nominal location. The additional regularization term is defined using the centroid location covariance computed from all previously registered brains. Deviations for different structures in different directions are penalized differently.

Let $q = G(p)$ be the voxel location of $p$ after the global transform, and $\Omega_A^s$ be the subdomain that includes the structure $s$ as well as a 200 µm surrounding area. The inclusion of a surrounding area aims to increase the sensitivity of the function. For a given structure $s$, we seek a rigid transform that maximizes the function $Q^{local}$ defined as:

$$Q^{local}(L_s) = \sum_{p \in \Omega_A^s} D_s(L_s(q)) \cdot M_s(p - c_s) - \eta \mathbf{t}_s^\top \Sigma_s^{-1} \mathbf{t}_s. \quad (7.3)$$

where $\mathbf{t}_s$ is the translation component of $L_s$, $c_s$ is the nominal location and $\Sigma_s$ is the covariance matrix.

### 7.3 Estimating Deformation Field

The deformation for any space in between structures is determined by a weighted average of the transforms of the structures nearby. The influence of each structure transform
Figure 7.2: Structure annotations derived from global registration and structure-specific registration for an example section. Contours are the cross-sections of 0.5-level iso-surfaces of the aligned atlas structures. Grid lines represent the transformed result of a regular grid defined in atlas space. Left: results after global registration. Structures are placed reasonably close to the correct positions, but individual adjustment is still necessary. Grid lines exhibit an affine transformation. Right: results after structure-specific registration. Structure pose and locations are improved. Warped grid lines demonstrate the deformation field.

on a given location is inversely related to the distance between the structure centroid and the location.

Let $\mathbf{q} \in \Omega_A$ be a given atlas space location, and $L_s$ is the estimated transform for structure $s$. The deformation field $E$ is defined as:

$$E(\mathbf{q}) = \frac{1}{\sum_s w_s(\mathbf{q} - \mathbf{c}_s)} \cdot \sum_{s \in S} w_s(\mathbf{q} - \mathbf{c}_s) \cdot L_s(\mathbf{q})$$

(7.4)

where $w_s$ is a radial basis function $w_s(r) = \exp(-||r||/d_s)$. We set $d_s$ to be 70% of the size of the corresponding structure in the direction of $r$.

Figure 7.2 shows an example of the resulting structure annotations derived from global registration and structure-specific registration.
7.4 Optimization

Optimization of $Q^{global}$ uses gradient descent over each of the 12 parameters. For $Q^{local}$, the optimization is based on an Euler vector representation of rigid transform and gradient descent is performed on the logarithmic map of the Lie group SE(3). Details can be found in Appendix ??.

Gradient ascent finds a local maximum by iteratively updating the estimate according to direction of the gradient. A reasonable starting point is found using multi-resolution grid search over the three translation parameters. We used 5 test points for each parameter spanning an initial range of ± 500 µm, and in every successive iteration, re-center the search at the current best estimate with an exponentially reduced range, until the range gets below 30 µm.

Let $x^0$ be the estimate after grid search for a particular parameter $x$. At iteration $t$, let $T^t$ be the current estimate for all parameters and $g^t_x$ be the gradient of the objective function $Q$ with respect to $x$,

$$g^t_x = \frac{\partial}{\partial x} Q(T^t),$$ (7.5)

the parameter is updated according to

$$x^{t+1} \leftarrow x^t + \eta^t_x \cdot g^t_x,$$ (7.6)

Here $\eta^t_x$ is the learning rate specific to parameter $x$. Using a global learning rate for all parameters is not optimal when the magnitudes of different parameters are not of the same scale. Therefore we used Adagrad [24] to automatically rescale a base learning rate $\eta$ differently for each dimension according to the historical gradient magnitude on that
dimension:

\[ \eta_x' = \frac{\eta}{\sqrt{\sum_{\tau=1}^{t} \beta^x_{\tau}^2}}. \]

We set the base learning rate \( \eta = 10 \) for the three translation parameters and \( \eta = 0.1 \) for the other parameters.

To speed up gradient computation, a random set of 1,000 voxels are used to compute the gradient at each iteration. This is a 30-70% time saving compared to using all voxels of a structure. To verify that this does not bias the results, we arbitrarily selected a set of registrations and repeated each ten times with different random seeds. The standard deviation of the estimated translation is consistently within 2 \( \mu \)m.

### 7.5 Results

We evaluate the precision of structure localization in three ways.

#### 7.5.1 Compare with expert annotations

First, we registered the atlas to each of the annotated brains, without using the expert annotation in any way. We then compared every structure in the aligned atlas with the corresponding expert annotation which is reconstructed as a binary volume, in terms of the centroid distance and the per-voxel probability difference.

The distance between the centroid of every atlas structure and that of the expert annotation is normalized by the size of the structure in the direction of the deviation. Figure 7.3 plots the results for every structure. The median normalized deviation is 6.7%.

Figure 7.4 shows the results for the root-mean-squared (RMS) difference of the probabilities at every voxel between the registered structures and the expert annotations.
Figure 7.3: Deviation of the centroids of registered structures from expert annotation, normalized by the size of corresponding structure. Median = 6.7%.

Figure 7.4: Root-mean-squared error between registered probabilistic structures and expert annotations.

Even though the structure shapes are fixed, we achieved an average RMS difference of 0.23, suggesting a high degree of agreement.

7.5.2 Quantify expert corrections

Second, we aligned the atlas to the nine unannotated brains and asked two experts to review the results. The experts can manually adjust the placement of any structure by translating or rotating it in 3-D. We recorded the number of adjustments and the magnitude of each adjustment. Typically, less than 5 corrections need to be made on each brain (out of
51 structures, \(\approx10\%\) error rate), with a mean translation magnitude of 63.2 \(\mu m\). Of note, less than ten minutes was required for an expert to review and correct the annotations for an entire brain. This is approximately 200 times less than the 30 hours required for creating annotations from scratch.

### 7.5.3 Quantify significance

Thirdly, we developed a unsupervised metric to measure the significance of an estimated alignment (either global or structure-specific). Significance is measured in terms of a \(z\)-score, which measures the estimated maximum of the objective function relative to the mean and standard deviation of the values in its neighborhood. We used a neighborhood of \(\pm 100 \mu m\) in each translation direction. High \(z\)-score implies accurate fine-scale detection and successful optimization. (Figure 7.5a)

A complementary metric measures the width of the function maximum, characterized by the Hessian. From the eigenvalues and eigenvectors of the Hessian, we derived the most certain and the least certain directions of translation. We then computed for each direction a “margin”, defined as the distance from the maximum along the given direction that the \(z\)-score drops to 1. If the margin is large relative to the structure’s nominal size in the corresponding direction, then the sensitivity of the classifier is too low for precise localization.

The results for the registrations of 51 structures in 12 brains are plotted in Figure 7.5bc. For the global registration, the median \(z\)-score is 2.2. For the structure-specific registration, the median \(z\)-score is 1.41, and the median structure-size normalized margins in the least certain direction is 8.9\%. These results suggest that our registrations have a high level of confidence.

Based on these metrics we can identify a subset of high-confidence structures...
(z > 1.5) that are spatially spread out in the brainstem, which can have higher weights in the deformation field interpolation.

### 7.5.4 Comparison to intensity-based methods

To demonstrate the advantage of texture over low-resolution intensity for cross-brain registration, we aligned two manually annotated brains in 3-D using both methods and compare the results. For the intensity-based method, the brains are reconstructed as graylevel volumes (16 µm isotropic voxel). We first computed a global affine transform, and then computed a rigid transform for each structure by considering a 500 µm cubic region around each structure. Registration criteria is the normalized correlation. We then measure the overlap between the ground-truth annotations of the same structure in both brains, in terms of the Jaccard index. The Jaccard index between two binary masks $A$ and $B$ is defined as $\frac{|A \cap B|}{|A \cup B|}$ which ranges from 0 for completely disjoint to 1 for exact overlap. Figure 7.6 shows that for most structures our method gives a better alignment.

To make a more direct comparison in the context of atlas-based annotation, we aligned the annotated Nissl template of Allen CCF v3 (10µm voxel resolution) to one of our manually annotated brains by maximizing the mutual information of graylevel intensities of the Nissl images. Registration used deformable cubic B-splines. We then compared the structure boundaries derived from the registered template against the ground-truth annotation. As shown in Figure 7.7, the Allen template does not have enough details in the brainstem area to drive the registration, thus causing large errors. In contrast our method successfully registered all structures.

This chapter, in part, have been published in the following proceeding. The dis-
Figure 7.5: Measures of registration confidence. (a) Landscape of the objective function for a particular registration. Magnitude is normalized to yield z-scores. Significance metrics are the z-score of the estimated maximum, and the margin i.e. the distance from the maximum where z-score drops to 1. (b) z-scores of all structure-specific registrations. (c) normalized margin of all structure-specific registrations.
Figure 7.6: Comparison between texture-based method and graylevel-based method for registration. (a) Examples of registering two annotated brains using our method versus using an graylevel-based method. Images of one brain are shown in greyscale. Structures from the other brain after registration are shown in color overlay. Top row: texture-based registration accurately registers three structures that are at the center of respective images. Bottom row: graylevel-based registration often favors the alignment of high-contrast large structures at the expense of misaligning small nuclei, and cannot distinguish structures that are cytoarchitecturally different but have similar graylevels. (b) Plot of the resulting Jaccard index for each of 51 structures using our method against the intensity-based method. Our method is more accurate for most structures.

Figure 7.7: Comparison between texture-based method and greylevel-based method for atlas-based annotation. (a) Annotated Nissl template provided in Allen CCF v3. (b) An example Nissl section from our data, annotated using our method (blue) versus intensity-based registration of Allen’s template (red). Annotations given by our method achieved high consistency with the ground truth (green) throughout the brainstem, while the registered Allen’s template has large errors especially in the posterior part where high constraint landmarks are rare.

This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Chapter 8

Updating Atlas

8.1 Updating Reference Atlas

The structure-specific transforms, after verification and correction by experts, can be used to quantify population variability. In particular we used them to update the centroid location statistics in the reference atlas. Let $t_s$ be the translation component of the transform of structure $s$. The update to the mean centroid location is simply

$$c_s \leftarrow c_s + \frac{1}{n+1} t_s$$

where $n$ is the number of instances of this structure that have already been incorporated. Similarly the update to the covariance is

$$\Sigma_s \leftarrow \frac{n}{n+1} \Sigma_s + \frac{1}{n+1} t_s t_s^\top.$$ 

With the updated location statistics, we measure the variability of the position of structures across brains. This measure will consist of the natural biological variability as well
as any residual variability from errors in our annotation and our automated procedures. Thus the variability serves as an upper bound on biological variability as well as on our ability to gauge significance in the overlap of labels across brains. We quantified the variability in the location of the centroid of each landmark across twelve brains relative to the updated nominal locations. This provides a measure of the deviation of every landmark from the sample means. Our result showed that the sample-averaged standard deviation of the rostral-caudal variation is $\pm 90 \mu m$ (Figure 8.1). Importantly, there was no systematic increase in variability along the R-C axis, as might occur from poor brain-to-brain registration. Some landmark structures are non-isotropic in their variability. For example, the variability of spinal trigeminal nucleus caudalis (Sp5C) is predominantly along the medial-lateral axis, while that of the substantia nigra reticulata (SNR) is primarily along the dorsal-medial axis (Figure 8.1).

Since the transforms estimated do not include deformation of the structures, structure shapes in new brains are not captured. Therefore we do not update the average shapes defined in the atlas.

### 8.2 Updating Texture Classifiers

For brains that were not hand-annotated, the registered atlas structures also serve as “soft” annotations from which one can extract additional training examples to improve the texture classifiers. Although the training set from three annotated brains is more than sufficient to obtain good classifiers, because each brain is processed under a slightly different condition, data from a wider variety of brains improves robustness.

By re-sectioning the iso-surfaces of atlas structures that are registered to each brain, we can generate 2-D structure contours with specific probability levels on any section. This
Figure 8.1: Twelve brains co-registered to atlas using global transforms only, demonstrating the variation in positions of structures around respective nominal centroids. Different brains are represented by different colors. We show both the full brainstem and for three example structures. The plot shows the position variation in rostral-caudal direction for all structures. Shadow shows the sample-averaged standard deviation of ± 90 µm.

allowed us to collect additional positive patches from the interior of $p=0.8$ contours, and negative patches the 500 µm boundary zone exterior of $p=0.1$ contours. It is a future work to examine whether the classifiers trained using this new set improves the detection on more new brains.

This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld,
This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Chapter 9

Application to Neuronal Projection Mapping

Reconstruction of neuronal circuitry is a fundamental goal of modern neuroscience. Identification of the connections between structures is critical for understanding the mechanisms in which brain activities are coordinated to produce behaviors. Mapping the projections into our atlas allows the comparison and integration of results across experiments (e.g. different injection sites, different markers) within a common anatomical context.

Neural pathways can be visualized using immunohistochemical techniques in which neuron-infecting viruses that spread trans-synaptically are injected into a given site. Fluorescent antibodies (e.g. GFP) conjugate to the virus are then applied to brain sections to label neurons in the pathway. Finally, the sections are imaged using a fluorescence microscope and one can study the images to infer the projections originating from the injection site.

Because the fluorescent marker channel alone does not provide enough cytoarchitecture for texture detection, we also used a counter-stain called Neurotrace blue which, like the thionin, also stains the Nissl substances in neurons. In these images, the blue channels
are used for detection and alignment, while the other channels carry projection labelings.

9.1 Intensity Normalization Between Fluorescent and Bright-field Nissl Stains

Neurotrace stained sections are imaged in fluorescence (dark-field) while the thionin stained sections are imaged in transmission (bright-field). In addition, different chemicals yield slightly different intensity profiles. In order to apply the classifiers trained on thionin to Neurotrace, the intensities of the two stains must be related. The normalization of each patch into zero mean and unit standard deviation implicitly assumes a linear mapping between the two stains, which is not sufficient to eliminate the difference. Here we propose a histogram matching-based method to estimate a nonlinear mapping that results in greater similarity between the two stains.

We processed alternate sections of brains for thionin (Figure 9.1b) versus Neurotrace blue (Figure 9.1a). We converted thionin to grayscale and used the blue channel for Neurotrace. In each brain, sections of both stains are registered by mutual information maximization.

Registration ensures that adjacent sections have similar distribution of cells, which allows us to find a relationship between the intensities of the two stains by matching the histograms. Histogram matching works by mapping every percentile of the Neurotrace intensities to the same percentile of the thionin intensities. The mean of the mappings estimated from all adjacent pairs (Figure ??a.iii) is used to normalize all Neurotrace sections. The histogram and the appearance of the Neurotrace image converge to those of the thionin after this mapping (cf Figure 9.1b with e), allowing thionin-trained classifiers to
generate quality probability maps.

A more general approach to deal with novel non-nissl stains is to train new classifiers. As long as alternately processed brains are available, we can perform structure detection and registration based on only the thionin set. Once the atlas is registered, we can then derive structure boundaries on the sections of the new stain which can in turn be used to train stain-specific classifiers. This is a promising future direction.

9.2 Mapping neuronal projections

As a first example of the utility of automated alignment, we identify the 3-D spatial distribution of orofacial premotor neurons labeled with a retrograde viral tracer. Pseudorabies virus (PRV) that expresses GFP was injected into the jaw masseter muscle. The animal was sacrificed and perfused 86-hours after the injection; at this time all pre-motor neurons and some pre$^2$-motor are expected to be labeled. We observe GFP labeling of trigeminal motor (5N) neurons only on the ipsilateral side and extensive presynaptic populations throughout the brainstem and hypothalamus (Figure 9.2). Known premotor populations were labeled in diverse primary sensory nuclei, e.g., mesencephalic and spinal trigeminal nuclei, the nucleus of the solitary tract, the medial vestibular nucleus, the parvocellular, intermediate, gigantocellular, lateral paragigantocellular regions of the reticular formation, the pontine nucleus and the superior colliculus. This experiment replicates known connectivity, yet it further provides the first 3-D map of trigeminal premotor locations. Additional label in presumed pre$^2$-motor structures include the central amygdala, the zona incerta, the hypothalamus, and the periaqueductal grey.
Figure 9.1: Intensity normalization between fluorescent Neurotrace and bright-field thionin. (a) Example dark-field fluorescent Neurotrace image, showing the facial motor nucleus (red contour). (b) Same area in adjacent bright-field thionin section. (c,d) Normalization of the Neurotrace image by linear stretching resulted in poor performance of the thionin-trained classifier for facial motor nucleus. (e) Nonlinear mappings from Neurotrace to thionin estimated by histogram matching for different sections. Curves correspond to the section in this example (green), other sections (transparent) and the mean (black). (f,g) Nonlinear normalization resulted in much higher similarity to thionin images, and thus comparable detection performance.
Figure 9.2: Visualization of the labeling motor and premotor inputs to the jaw muscle across all three planes. The jaw region of trigeminal motor nucleus (5N) was injected with pseudorabies (PRV) that expressed GFP and visualized with a Neurotrace background stain. The PRV labeled cells were manually annotated and aligned with the reference atlas. Note the widespread, bihemispheric inputs and, critically, the absence of labeling from the contralateral motor nucleus (see sagittal panel).

We assessed the utility of our texture-based alignment for concatenating labeled neurons across multiple brains onto the same coordinate system. We injected G-protein deleted rabies (ΔG-RV) that co-expressed a fluorescent protein into the jaw region of the trigeminal motor nucleus (5N) (RFP; Figure 9.3i) and into the intrinsic protractor muscle region of the facial motor nucleus (7N) (GFP; Figure 9.3ii) of separate mice. The brains were processed and counter stained with Neurotrace blue. Two-channel fluorescent detection was used, with blue light for landmark detection and alignment to the reference atlas with our texture-based classifiers, and green or red light for detecting the viral label. The sagittal 3-D projection illustrates the dispersion and heterogeneity of these populations (Figure 9.3ii) and a close-up of the data reveals a subset of two populations with highly overlapped density in the reticular formations PcRT and IRt (Figure ??c.iii insert). The accurate alignment of fluorescent tracing data illustrates the power of texture-based classifiers, i.e., approximately 90 μm root-mean-square deviation (Figure ??) compared with an approximately 500 μm overlap (Figure ??c.iii insert). Thus texture-based registration provides a measure of
Figure 9.3: Visualization of the labeling of different populations of premotor neurons in separate brains with overlapped density in the PCRt. We labeled the premotor neurons of the jaw region of the trigeminal motor nucleus (5N) using ΔG-rabies-RFP (subpanel i) and the premotor neurons of the vibrissa region of the facial motor nucleus (7N) using ΔG-rabies-GFP (subpanel ii). Premotor neurons predominantly overlap in a border area of the reticular formations PCRt and IRt (subpanel iii). Red points are premotor neurons of 5N and green points are premotor neurons of the facial motor nucleus (7N). The insert shows a magnified view of the overlap of the two premotor populations.

This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld,

This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Chapter 10

Related Work

This project is related to several computational problems in medical and neuro image processing.

10.1 Brain Parcellation and Cytoarchitecture

Korbinian Brodmann’s classical 1909 atlas [14] proposed a brain mapping strategy based on subdividing Nissl-stained sections into areas with similar cellular and laminar structure. This approach studies cellular composition, also known as cytoarchitecture, demonstrated in stained sections as a mean to parse the brain. Nissl staining uses dyes, such as thionin, that stain "Nissl bodies" (rough endoplasmic reticulum in the RNA) which are abundant in neurons. It reveals details ranging from the macroscopic, such as the laminar pattern of the cerebral cortex or the interlocking nuclear patterns of the brainstem, to the microscopic, such as the distinctions between individual neurons and glia in subregions of the central nervous system. Nissl staining can be supplemented by immunohistochemistry and in situ hybridation techniques that label certain genes or proteins expressed by particular
cell groups as well as functional MRI that allows the correlation between function and structure [17]. However, unmatched for its comprehensiveness, Nissl staining remains the gold-standard approach for brain parcellation and anatomical atlas construction for a wide range of species [65].

10.2 Image Registration

Image registration (or image alignment) is a well-studied problem (see [44, 18] for an overview). The aim is to spatially transform a set of images, such that homologous locations in all images have maximal overlap. In the context of brain imaging, registration is usually between 2-D images of thin histological sections and between 3D volumes acquired by MRI or optical techniques.

Landmark-based registration methods minimize the distances between corresponding landmarks on each image, which can be fiducial markers or image feature points. Intensity-based methods maximizes a similarity metric between pixel/voxel-based intensity patterns of the whole images. Commonly used metrics include normalized correlation coefficient and mutual information. In addition, a regularization term is typically also added to encourage smoothness and penalize biologically implausible warps. Other methods [46, 31] used more complex features to capture textural signature.

The types of transform to estimate range from linear transforms such as rigid, affine, piece-wise affine[51], to deformable transforms [60, 35, 26] such as parametric models based on B-spline [55, 59] and thin-plate spline [12]), and diffeomorphic methods [8, 64, 9] that describe deformation field by partial differential equations).

Optimization usually uses gradient-based methods. To ensure faster convergence, a common practice is to use a multi-resolution approach by which a series of progressively
more flexible transforms are estimated using images at increasing resolutions, each stage using the result of the previous one as a starting point.

Our method combines landmark detection with correlation-based volume-to-volume registration, and estimates a deformation field composed of a global affine transform and a weighted combination of piecewise rigid transforms.

10.3 Atlas Construction

An atlas is an annotated map defined over the anatomical organization of an organism. A classical brain atlas takes the form of a series of 2D or 3D images, acquired from one or more specimens, annotated with information such as structure delineation, genetic expression, connectivity patterns and cell types.

10.3.1 Single-subject Atlas

Building a brain atlas is difficult because not only does it require exceptional experimental skills to produce high-quality brain sections, it also demands deep neuroanatomy knowledge as well as patience to perform accurate systematic annotation. As a result most traditional brain atlases are based on a very small number of specimens, typically one for each plane of section. The mouse brain atlas by Paxinos and Franklin [48] is based on alternate Nissl and AChE series of three brains cut in coronal, sagittal and horizontal planes respectively. Over 500 structures spanning the whole brain are delineated on images endowed with stereotaxic coordinates. The Allen Reference Atlas [21], initially published in 2005, was the first of a series of effort by the Allen Institute for Brain Science to create a "Common Coordinate Framework" (CCF) for the mouse brain. Based on a series of 528
aligned coronal Nissl sections from one specimen brain, they created a symmetric 3-D reference brain annotated with over 800 structures (2011).

10.3.2 Population-averaged Template

A small data set tends to introduce bias. As high-throughput processing and imaging technology advances, the data bottleneck has been largely eliminated. Increased computational power and faster automated registration algorithms allows modern atlasing projects to create average brains based on as many as thousands of specimens.

An average brain is usually in the form of an average shape, average intensity template with structure labels defined over the template. The labels can either be drawn on the template or on specimen images in which case the labels on the template can be probabilistic. Segmentation or annotation of user images is usually done by registering the images with the template, either manually or using an automated algorithm.

Construction of average brains is prevalent for low-resolution data such as magnetic resonance (MR) images [28, 43, 33, 10, 22, 32, 25, 49] where the voxel size is 10-100 µm. For histological data with 0.35-0.5 µm pixel resolution, intensity-based registration methods require downsampling or spatial smoothing before the averaging [54, 2, 42]. For example, Allen Institute’s CCF version 3 [2] is derived by deformably averaging the autofluorescence channel of serial two-photon (STP) tomography images of over 1,600 mouse brains. Compared to a 0.35 µm original planar resolution, the voxel size used for registration is only 10 µm.

The requirement for low-resolution data renders intensity-based registration unreliable in many applications. In particular, they cannot align brainstem structures accurately. In order to achieve precise localization, the similarity measure must be sensitive to small de-
viations from the correct location. The voxel-based measures are adequate at low resolution since the macroscopic appearance of different images are usually consistent, but they are not sufficient at sub-micron scale where the randomness of cell organization becomes evident and pixel/voxel-wise correlation across images is impossible. Settling with downsampled images, however, means many structures that are only differentiable from neighboring tissues by cytoarchitecture no longer have sufficient contrast. Such structures (nuclei) are particularly abundant in brainstem.

For construction of average templates, this problem is amplified by the low-pass nature of the averaging operation. While high-contrast structures that tend to be registered well will appear more clear in the average template than in the image of a single subject, structures that cannot be consistently aligned are smoothed away, making it impossible to robustly align the nearby areas. To complement average templates with cytoarchitecture information, atlas authors often also provide Nissl specimens that are registered to the templates [37]. This allows the user to manually warp their data to the template, but automated registration is still not adequate due to their reliance on low resolution.

Our approach tackles this issue by converting full-resolution textures into discriminative descriptors, which enables the use of correlation as similarity measure while utilizing high resolution information.

### 10.4 Automated Annotation

Various methods are proposed to automatically annotate images based on annotated atlases [27]. The classical methods deform an annotated template to fit images [67]. Multi-atlas label fusion [66, 23] uses an ensemble of templates representative of different variations to obtain candidate labels which are then consolidated.
Bayesian methods use the atlas as a spatial prior and train appearance models for different structures based on a variety of image features (mainly voxel intensities). Many works merge segmentation, registration and image artifact estimation into a single Maximum A Posteriori framework [62, 7, 52]. For generative models, EM is used to estimate the imaging and registration parameters treating the voxel labels as hidden variables. [61] used a combination of generative shape model and discriminative multi-class appearance model. Although our approach does not adopt a Bayesian formulation, we similarly employed a deep feature-based discriminative model for texture classification. This is combined with an anatomical model (the reference atlas) that separately models structure locations as normally distributed and shapes as fixed constraints. We emphasize that in our framework the texture classification is only a means to the end that is the registration.

10.5 Analysis of Histology Images

10.5.1 3D reconstruction of histology images

Unlike MRI volumes which inherently maintain anatomical integrity, histological sections are imaged after being removed from the tissue block. Handling and sectioning cause unpredictable distortion and a series of sections must be properly registered to reconstruct the 3-D anatomy. Most works apply in-plane nonlinear deformations [47, 6, 5, 34, 13]. In contrast, our system used adhesive tapes to fix the tissue before cutting to avoid deformation during transportation to slides, which minimizes warping and largely eliminates the need for deformable transforms in reconstruction.

To ensure the accuracy of overall shape, many approaches also involve registering the sections to an in vivo MRI template. For example, [69] interleaves a high-frequency
reconstruction of the section series and the low-frequency registration to MRI.

10.5.2 Texture recognition on histology images

Cellular texture is a defining characteristic of histological images. Automated classification of different regions based on texture is a common task. One example in the histopathological context is the segmentation and classification of different grades of cancer from healthy tissues. In images of brain sections the textural differences between structures are often more subtle.

Traditionally, a wide range of hand-crafted features are used ([57, 68], see [29] for a review). Generic texture features based on intensity statistics, include color histogram, intensity neighborhood [63], local binary patterns [45] and Haralick features derived from the grey-level co-occurrence matrix (GLCM) [30]. Features that aim to capture the patterns at different scales and frequencies include wavelet transform and the responses to filter banks such as Gabor filters [38]. Histology-specific features are also devised based on cell detection results, and various graph-theoretic statistics can be derived from a graph formed by connecting the cells [58]. The success of deep learning in pattern recognition also motivates the use of features extracted from deep neural networks, in particular convolutional neural network (CNN) [4, 16, 39].

One issue that complicates the processing of histology images is the intensity variation both within a single section and across sections due to uneven staining. Many intensity normalization methods are proposed, including linear normalization [53] and histogram matching [19] in a certain colorspace, such as lab or stain specific ones obtained by color deconvolution [40, 56]. Since both stains (thionin and Neurotrace blue) in our study are mostly monochrome, we applied histogram matching to the greylevel channels to correlate
the two stains.

This chapter, in part, have been published in the following proceeding. The dissertation author was the primary author of this paper. “The Active Atlas: Combining 3D Anatomical Models with Texture Detectors”, Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Karten, Harvey J.; Mitra, Partha P.; Kleinfeld, David; Freund, Yoav, Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI 2017. Lecture Notes in Computer Science, vol 10433, pp. 3-11, 2017, Springer.

This chapter, in part, is also currently being prepared for journal submission. Chen, Yuncong; McElvain, Lauren E.; Tolpygo, Alexander S.; Ferrante, Daniel; Friedman, Beth; Mitra, Partha P.; Karten, Harvey J.; Freund, Yoav; Kleinfeld, David. The dissertation author was the primary author of this material.
Chapter 11

Conclusion

In this thesis we developed a brain atlasing system that registers high-resolution serial sections of mouse brains based on cellular textures of different anatomical structures. We combined discriminative texture detectors with a reference model that probabilistically describes structure shapes and locations. Our algorithm encodes histological textures with a deep neural network, and trains classifiers to differentiate textures on the inside of a structure from the textures on its immediate surrounding area. We used correlation to measure the goodness-of-fit between the detection results and the transformed reference model. We perform an affine registration to correct for the global pose; and then perform structure-specific registrations to capture the variations unique to each structure in each particular brain. A deformation field is interpolated to enable the mapping of cell markers throughout the brainstem.

By strongly utilizing textural information, the system enables accurate and confident structure-level registration not possible by low-resolution greylevel similarity-based methods. The advantage is especially significant for regions that lack high-contrast boundaries. By "actively" identifying structures and perform structure-to-structure registration, it is also
more precise than methods that use structure-agnostic templates.

Initialized by minimal human annotation, a semi-supervised learning loop is realized by leveraging large amounts of unannotated data to update the texture detectors and the anatomical model. By reducing human effort to simply review of the registrations generated by the algorithm, which takes a fraction of the time required for manual annotation or manual registration, our system effectively eliminates one of the major bottlenecking tasks in the analysis of histological data.

One major limitation of our system is the inability to update structure shape models based on registered data. This requires the algorithm to estimate for each brain, not just the location and pose of each structure, but also its precise 3-D boundary. The “soft” annotations provided by the alignment of atlas structures are reasonable for structures whose shapes are relatively conserved across brains, but they are not adequate for structures whose variations are not captured by the three manually annotated training brains. One possible approach to allow the estimation of shape is to use more flexible transforms than rigid transform for structure-specific registrations, but methods that directly optimize boundaries such as active contours are likely to be more effective. The low number of training brains also limits our ability to learn generative models of structure shapes, which would provide not just the mean (like our average shapes), but also the variation, allowing Bayesian-style estimation of structure shapes from subjects.
Appendix A

Optimization of Registration Objective Functions

A.1 Optimization of Affine Transform

For a given structure $k$, write the spatial gradients of the fixed brain’s probability map as

$$ G_k(p) = \frac{\partial S_k(p)}{\partial p} \quad (A.1) $$

for $p \in \Omega_S$.

Let $g$ denote the 12-vector of the current estimate. Suppose $A(g)$ is the corresponding $3 \times 4$ matrix. Let $q$ denote $A(g)(p \ 1)$. Recall that the objective function is defined as,

$$ Q = \sum_{p \in \Omega_k} M_k(p - c_k) \cdot S_k(q) \quad (A.2) $$

$$ \frac{\partial Q}{\partial g} = \sum_{p \in \Omega_k} M_k(p - c_k) \cdot G_k(q) \cdot \frac{\partial A(g)p}{\partial g}, \quad (A.3) $$
where

\[
\frac{\partial A(g)}{\partial g} \mathbf{p} = \begin{pmatrix} \mathbf{p}_x & \mathbf{p}_y & \mathbf{p}_z & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{p}_x & \mathbf{p}_y & \mathbf{p}_z & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{p}_x & \mathbf{p}_y & \mathbf{p}_z & 1 \end{pmatrix}
\]

(A.4)

Gradient descent updates the parameters by

\[
g \leftarrow g + \eta \frac{\partial Q}{\partial g} \bigg|_{g=g}
\]

(A.5)

### A.2 Optimization of Rigid Transform

We perform the optimization of rigid transform on the manifold $SE(3)$ [11].

Let $v = \begin{pmatrix} t \\ \omega \end{pmatrix}$ be a 6-vector of coordinates in the Lie algebra $se(3)$, comprising of two separate 3-vectors: $\omega$, the vector that determine rotation, and $t$, which determines translation. The exponential map of $v$ is defined as

\[
e^v \equiv \begin{pmatrix} [\omega] \times & t \\ 0 & 1 \end{pmatrix}
\]

(A.6)

where the skew-symmetrix matrix operator $[\cdot] \times$ is defined as:

\[
\begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}
\]

(A.7)

Let $T$ denote the current estimate on the manifold $SE(3)$. Let $\varepsilon$ be a small increment in the Lie algebra coordinates. Let $e^\varepsilon$ be the corresponding Lie group exponentiation which
represents a small increment around $T$ in the manifold.

The gradient descent formula becomes

$$
\epsilon^* \leftarrow \eta \frac{\partial Q(e^T \epsilon)}{\partial \epsilon} \bigg|_{T=T}.
$$

(A.8)

$$
T \leftarrow e^{\epsilon^*} T.
$$

(A.9)

Let $q = Tp$. We obtain the Jacobian of the objective function $Q$ with respect to $\epsilon$:

$$
\frac{\partial Q}{\partial \epsilon} = \sum_{p \in \Omega_k} M_k(p - c_k) \cdot G_k(q) \cdot \frac{\partial e^\epsilon q}{\partial \epsilon}.
$$

(A.10)

where the last term is given by

$$
\frac{\partial e^\epsilon q}{\partial \epsilon} = \left( I_3 - [q]_\times \right)
$$

(A.11)

Lastly, the gradients of the regularization term $H = t_s^T \Sigma_s^{-1} t_s$ are given by,

$$
\nabla H = 2 \Sigma_s^{-1} t_s
$$

(A.12)
Bibliography


[10] Kanwal K Bhatia, Paul Aljabar, James P Boardman, Latha Srinivasan, Maria Murgasova, Serena J Counsell, Mary A Rutherford, Joseph V Hajnal, A David Edwards,


