Model-Based Segmentation of Hippocampal Subfields in Ultra-High Resolution In Vivo MRI

Koen Van Leemput¹,², Akram Bakkour¹,³, Thomas Benner¹, Graham Wiggins¹, Lawrence L. Wald¹,⁴, Jean Augustinack¹, Bradford C. Dickerson¹,³, Polina Golland², and Bruce Fischl¹,²,⁴

¹ Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, MGH, Harvard Medical School, USA
² Computer Science and Artificial Intelligence Laboratory, MIT, USA
³ Department of Neurology, MGH, Harvard Medical School, USA
⁴ Harvard-MIT Division of Health Sciences and Technology, MIT, USA

http://picsl.upenn.edu/caph08/papers/paper04.pdf
Motivation

✓ The hippocampus consists of multiple, interacting subregions (cf. talk by Susanne Mueller)

✓ Distinct hippocampal subregions:
  – are implicated in different memory subsystems
  – are differentially affected in different conditions and disease processes (normal aging, Alzheimer’s disease (AD), ...)

The ability to measure these subregions using in vivo neuroimaging is of great potential value:
  Basic neuroscience: insights into the function and structure of the hippocampus in the living human brain, and how it is affected in normal aging.
  Clinical research: sensitive, non-invasive biomarkers for early diagnosis and treatment evaluation in AD. Surrogate outcome markers in clinical treatment trials.
Recent developments in MR data acquisition technology are starting to yield images that show anatomical features of the hippocampal formation at an unprecedented level of detail.

- **Standard resolution**: $1 \times 1 \times 1 \ mm^3$
- **Ultra-high resolution**: $0.38 \times 0.38 \times 0.8 \ mm^3$

New opportunities for explicitly quantifying individual subregions, rather than their agglomerate, directly from *in vivo* MRI.

Analyzing large imaging studies of ultra-high resolution MRI scans requires automated computational techniques.
Model-based segmentation

✓ Derive computational models from manual delineations in a number of subjects

✓ Use those models to automatically segment MRI scans of new subjects
Bayesian modeling

“labeling model”

\[ p(L|\Phi_L) \]

\[ p(\Phi_L) \]

Label image \( L \)

“imaging model”

\[ p(Y|L, \Phi_Y) \]

\[ p(\Phi_Y) \]

MR image \( Y \)

Step 1

✓ Invent a generative model of image formation
  - A mathematical model of how an MRI image is formed
  - Depends on some model parameters

\[ \Phi = \{\Phi_Y, \Phi_L\} \]
Bayesian modeling

**Step 2**

- Use the generative model to obtain the most probable segmentation given an MRI image

\[
\hat{L} = \arg \max_L p(L|Y) \approx \arg \max_L \int_{\Phi} p(L|Y, \Phi)p(\Phi|Y)\,d\Phi
\]

- Involves **two** optimizations:
  - First estimate the optimal model parameters \( \hat{\Phi} \)
  - Then find the optimal segmentation based on those parameter estimates
Labeling model

\[ p(L | \Phi_L) \]
\[ p(\Phi_L) \]

Label image \( L \)

“labeling model”

“imaging model”

\[ p(Y | L, \Phi_Y) \]
\[ p(\Phi_Y) \]

MR image \( Y \)
Labeling model

The label image is modeled as being generated by deforming a probabilistic atlas, and sampling the voxel labels from the deformed atlas.

- Tetrahedral mesh atlas representation
- The parameters $\Phi_L$ are the location of the mesh nodes
- The prior $p(\Phi_L)$ penalizes deformations and is topology-preserving [Ashburner et al., 2000]
Labeling model

probability for white matter

probability for subiculum

... etc ...
Labeling model

deformed probability for white matter

deformed probability for subiculum

...etc...
Labeling model

... etc ...

Label image $L$
So where does the atlas come from?

- Manual delineations in N subjects are “augmented” with automated tissue classification in non-labeled voxels.
- The resulting label images are modeled as being generated by the deformable mesh-based labeling model described before.
- Bayesian inference: [Van Leemput, MICCAI 2006], [Van Leemput, TMI 2008 (submitted)]
  - For a given mesh representation, what are the most likely label probabilities in the mesh nodes? Group-wise registration
  - What is the most likely mesh representation?
    - Automatically penalizes overly complex models (cf. MDL)
    - Compact mesh representations
    - Avoids overfitting to the training data
Imaging model

```
\text{"labeling model"} \quad p(L|\Phi_L) \\
\quad p(\Phi_L) \\
Label \text{image } L
```

```
\text{"imaging model"} \quad p(Y|L, \Phi_Y) \\
\quad p(\Phi_Y) \\
MR \text{ image } Y
```
The intensity in each voxel is modeled as being drawn independently from a Gaussian distribution associated with its label.

- The model parameters $\Phi_Y$ are the means and variances of the Gaussians.
- We assume a uniform prior $p(\Phi_Y)$. 

"labeling model" $\downarrow$ \hspace{2cm} $\uparrow$

$p(L | \Phi_L)$  
$p(\Phi_L)$

Label image $L$

"imaging model" $\downarrow$ \hspace{2cm} $\uparrow$

$p(Y | L, \Phi_Y)$  
$p(\Phi_Y)$

MR image $Y$
Model parameter optimization

Given an MR image $Y$ that is to be segmented, what are the most probable model parameters $\Phi = \{\Phi_Y, \Phi_L\}$?

- Mean and variance of the Gaussian distribution associated with each neuroanatomical label
- Position of the nodes of the mesh (atlas warp)

Parameter optimization performed with a Generalized Expectation Maximization algorithm

- Repeatedly try to improve a lower bound to the objective function
- Each improvement automatically means an improvement in the objective function
Model parameter optimization

✓ Improve lower bound w.r.t. the atlas warp, keeping the Gaussian distribution parameters fixed at their current values
  – Calculate a statistical classification
  – Warp the atlas onto this classification (cf. [D’Agostino 2006, Pohl 2006])
  – Gradient descent

✓ Improve lower bound w.r.t. the Gaussian distribution parameters, keeping the atlas warp fixed at its current value
  – Calculate a statistical classification
  – Closed-form expressions for Gaussian distribution parameters

Statistical classification of each voxel depends on how well its intensity is explained by each label’s Gaussian distribution and by the prior probability for each label in the atlas in its current deformation
Model parameter optimization

- Improve lower bound w.r.t. the atlas warp, keeping the Gaussian distribution parameters fixed at their current values
  - Calculate a statistical classification
  - Warp the atlas onto this classification (cf. [D’Agostino 2006, Pohl 2006])
  - Gradient descent

- Improve lower bound w.r.t. the Gaussian distribution parameters, keeping the atlas warp fixed at its current value
  - Calculate a statistical classification
  - Closed-form expressions for Gaussian distribution parameters

Statistical classification of each voxel depends on how well its intensity is explained by each label’s Gaussian distribution and by the prior probability for each label in the atlas in its current deformation

Upon converge of the parameter estimation algorithm, the optimal labeling $\hat{L}$ is obtained by assigning each voxel to the most probable label
Experiments

✓ Ultra-high resolution (0.38 x 0.38 x 0.8 mm³) MRI data
  - 3T Siemens Trio with prototype custom-built 32-channel coil
  - Optimized MPRAGE sequence, 208 coronal slices
  - 5 acquisitions; motion-corrected and resampled to 0.38mm isotropic

✓ Manual segmentation of the right hippocampus in 5 subjects (2 young + 3 older cognitive normal)
  - Fimbria, CA1, CA2/3, CA4/DG, presubiculum, subiculum, hippocampal fissure + surrounding structures
  - Extremely time consuming: 1 -> 2 weeks *per hippocampus*

Validation of the automated segmentation algorithm using leave-one-out cross-validation
Qualitative results

Coronal

Axial

Sagittal

MRI data

Manual segmentation

Automated segmentation
Qualitative results

- Coronal
- Axial
- Sagittal

MRI data × manual segmentation × automated segmentation

- fimbria
- CA2 3
- CA1
- CA4 DG
- presubiculum
- subiculum
- hippocampal fissure
- hippocampus
- inf. lateral ventricle
- choroid plexus
Quantitative results: spatial overlap

- Dice coefficient for automated vs. manual segmentation
  \[ \frac{\text{volume of overlap}}{\text{average volume}} \]
- Excluded the “catch-all” label towards the tail of the hippocampus
Quantitative results: volume measurements

✓ Volume differences between manual vs. automated segmentations
✓ Linear regression between volume measurements with both methods (Pearson’s correlation coefficient)
Future (current) work

✔ More thorough validation
  - Repeated manual labelings of the same subjects
  - Human intra- and inter-rater variability provides context for validation

✔ Modified imaging model to take partial volume effect into account
  - Segment hippocampal subfields in standard resolution (1 x 1 x 1 mm$^3$) scans
  - Alzheimer’s Disease Neuroimaging Initiative (ADNI)
  - Open Access Series of Imaging Studies (OASIS) project
  - Hopefully with a reasonable reliability 😊

[Van Leemput, TMI 2003]
Support for this research was provided in part by:

- the NIH NCRR (P41-RR14075, R01 RR16594-01A1, NAC P41-RR13218, and the BIRN Morphometric Project BIRN002, U24 RR021382)
- the NIBIB (R01 EB001550, R01EB006758, NAMIC U54-EB005149)
- the NINDS (R01 NS052585-01, R01 NS051826)
- the MIND Institute.
- The Autism & Dyslexia Project funded by the Ellison Medical Foundation