Introduction to Medical Image Segmentation

HST 582
Matthew Toews
(slides adapted from William Wells III)
Outline

• Terminology & Applications
• Probability Review
• Intensity-Based Classification
• Prior models
• Morphological Operators
Image Segmentation

Partitioning an image into regions defined by pixel intensity and geometry.

Input: Brain MRI
Output: Pixel Labels

- Grey Matter (GM)
- White Matter (WM)
- Cerebrospinal Fluid (CSF)
- Multiple Sclerosis Lesions
Anatomical Description Hierarchy

WHOLE
BRAIN/STRUCTURE

LOBES

SYSTEMS

CYTO- & MYELO-
ARCHITECTURE

Provided by D Kennedy
Stages of Anatomic Analysis

Original "General" Segmentation.
Subcortical Parc.
Cortical Parcellation
White Matter Parc.

Provided by D Kennedy
Medical Applications of Segmentation

• Image Guided Surgery
• Surgical Simulation
• Neuroscience Studies
• Therapy Evaluation
fMRI and Electro Corticography

Applications of Segmentation

- Image Guided Surgery
- Surgical Simulation

Personalized orthopedic surgery
Applications of Segmentation

• Neuroscience Studies
Statistical Map of Cortical Thinning: Aging

p < 10^{-6}

Thanks to Drs. Randy Buckner and David Salat for supplying this slide.

Provided by Bruce Fischl
Applications of Segmentation

- Therapy Evaluation
  - Multiple Sclerosis
  - Knee Cartilage Repair
Results: Segmentation of Femoral & Tibial Cartilage

MRI Image  Model-Based Segmentation  Manual Segmentation

Provided by T Kapur
Limitations of Manual Segmentation

- slow (up to 60 hours per scan)
- variable (up to 15% between experts)

[Warfield + 2000]
Automatic Segmentation

An automated segmentation method needs to reconcile

- Gray-level appearance of tissue
  - Characteristics of imaging modality
- Geometry of anatomy
Terminology: *Segmentation*

- **HST 582:**
  - Labeling images according to tissue type (e.g. White / Gray Matter)
  - Can include ‘object detection’, e.g. a brain region in an image

- **Graphics Community:**
  - Any process that turns images into models
Probability Review

• Discrete Random Variables (RV)
  – Probability Mass Functions (PMF)
• Continuous Random Variables
  – Cumulative Distribution Functions (CDF)
  – Probability Density Functions (PDF)
• Conditional Probability
• Bayes’ Rule
Discrete Random Variable

• Characterized by *Probability Mass Function* (PMF)
  – (sometimes called Distribution)
  – Maps values $x$ to their Probabilities $P(x)$

$$0 \leq P(x) \leq 1$$

$$\sum_{x} P(x) = 1$$
Continuous Random Variables

• Define Cumulative Distribution Function (CDF) on RV $x$

$$F_X(x) = P(X \leq x)$$

$$0 \leq F_X(x) \leq 1$$

• Non-Decreasing
• Sometimes called Distribution Function
Continuous Random Variables...

- Define *Probability Density Function* (PDF)

\[ p(x) = \frac{d}{dx} F_X(x) \]

- Easy to show, using Fundamental Theorem of Calculus:

\[ P(a \leq x \leq b) = \int_a^b p(x) \, dx \]
More on PDFs: $p(x)$

- Non Negative
- Integrates to One
- (Value can be Greater than One)
Conditional Probability

• Define Conditional Probability:

\[ P(X|Y) = \frac{P(X \& Y)}{P(Y)} \]
Bayes’ Rule (easy to show)

\[ P(A|B) = \frac{P(B|A)P(A)}{P(B)} \]

- Frequent Situation:
  - \( A \): State of the World
  - \( B \): Measurement, Observation
  - \( P(B|A) \): Measurement Model
  - \( P(A) \): A-Priori Model
Intensity-Based Segmentation

- Statistical Classification
  - ML
  - MAP, a-priori models
Segmentation

• Easy Segmentation
  – Tissue/Air (except bone in MR)
  – Bone in CT

• Feasible Segmentation
  – White Matter/Gray Matter
  – M.S. Lesions
Statistical Classification

- Probabilistic model of intensity as a function of (tissue) class
- Intensity data
- Prior model

[Duda, Hart 78]
Measurement Model

- Characterize sensor

\[ p(I|\text{Tissue class } J) \]

Tissue class conditional model of signal intensity

Mean intensity of tissue \( J \)

Intensity \( I \)
Example

\[ p(I|\text{gray matter}) \]

\[ p(I|\text{white matter}) \]
Maximum Likelihood Classification

• Measure intensity, $I_o$, and we want to know the tissue class
  
  $L(\text{TC}_j) = p(I_0 \mid \text{TC}_j)$

• Pick tissue class that maximizes $L$
• $L$ is not a probability
  – Called: Likelihood
Example - revisited

- white matter
- gray matter

threshold
Anatomical Knowledge

• *A priori* model
  – Before the measurement is considered

\[ P(TC_j) \]
MAP Classifier

- Choose TC to Maximize the *A Posteriori* probability

\[
P(TC \mid I_0) = \frac{p(I_0 \mid TC)P(TC)}{p(I_0)}
\]
Measurement Model

- Training data
  - Get an expert to label some of the voxels

- Optional: Use a parametric model
  - Assume functional form
    - Popular choice: Gaussian
Gaussian Density – 1D

• Why?
  – Central Limit Theorem
  – Makes math easy (when doing parameter estimation)

\[ G(\mu, \sigma, x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]
Choosing $\sigma$ and $\mu$

- Use training data: $\{I_1, I_2, \ldots, I_N\}$
- ML parameter estimation
  
  $$
  \mu = \frac{1}{N} \sum_i I_i \quad \sigma^2 = \frac{1}{N} \sum_i (I_i - \mu)^2
  $$

- MAP tissue classifier with Gaussian measurement model: choose tissue class to maximize:
  
  $$
  P(TC_j | I) = \frac{G(\mu_j, \sigma_j, I)P(TC_j)}{\ldots}
  $$
Gaussian Density – 2d Data

• Example

\[ I = \begin{cases} 
\text{proton density intensity} \\
\text{T2 weighted intensity}
\end{cases} \]

Mean Vector \( M \)
Covariance Matrix \( \Sigma \)

\[
G(M, \Sigma, X) = \frac{1}{N} \frac{1}{(2\pi)^{\frac{N}{2}} |\Sigma|^{\frac{N}{2}}} e^{-(I-M)^T \Sigma^{-1} (I-M)}
\]
2D Gaussian: Example

iso probability contour is ellipse
Multiple Sclerosis Example

- Dual echo MRI
  - 1 x 1 x 3 mm
  - Registered slice pairs
- Proton density image
  - Good: white/gray
  - Bad: gray/CSF
- T2-weighted image
  - Not so good: white/gray
  - Good: CSF/MS lesions
Dual Echo MRI Feature Space

- T2 Intensity
- PD Intensity

- CSF
- GM
- WM
- Severe lesions
- Air
• MS Lesions are “graded phenomenon” in MRI, and can be anywhere on the curve
Multiple Sclerosis

Provided by S Warfield
Background: Intensity Inhomogeneities in MRI

- MRI signal derived from RF signals...
- Intra Scan Inhomogeneities
  - “Shading” … from coil imperfections
  - interaction with tissue?
- Inter Scan Inhomogeneities
  - Auto Tune
  - Equipment Upgrades
EM-Segmentation

E-Step
Compute tissue posteriors using current intensity correction.

M-Step
Estimate intensity correction using residuals based on current posteriors.
Dual Echo Longitudinal Study

PDw

T2w

Provided by S Warfield
Tissue classification

No Intensity Correction

EM Segmentation
Generative Model

Image Intensities $I(x)$
(observable)

Tissue Labels $C(x)$
(unknown)

Image lattice, indexed
By pixel location $x$
Gaussian Mixture Model (GMM)

Likelihood Prior
(e.g. Gaussian density) (discrete)

\[ p(I) = \sum_{i} p(I \mid C_i) P(C_i) \]

Image Intensity Histogram

\[ p(I) \]

Intensity \( I \)
GMM Example

Two Tissue Types
Grey Matter (GM), White Matter (WM)

ML Parameter Estimation
- Expectation-Maximization (EM) Algorithm

\[ p(I) = p(I \mid GM)P(GM) + p(I \mid WM)P(WM) \]
GMM Segmentation

Maximum A-Posteriori (MAP) Tissue Estimation
- Identify most probable $C$ given intensity sample $I$

$$p(C_i | I) \propto p(I | C_i) P(C_i)$$

$$p(I) = p(I | GM) P(GM) + p(I | WM) P(WM)$$

Decision Boundary $I$

Intensity $I$
GMM Sampling

Generate new image histogram
- draw samples from GMM
Check modeling assumptions

\[ p(I) = \sum_{C_i} p(I \mid C_i) P(C_i) \]

Sampled Histogram

Intensity \( I \)
GMM Image Example

Visualize as RGB Colors

PD  T1w  T2w
GMM Image Example

\[ p(I) = \sum_{i} p(I \mid C_i) P(C_i) \]

Learn 4-tissue model

\[ p(C_i \mid I) \propto p(I \mid C_i) p(C_i) \]
GMM Image Example

\[ p(I) = \sum_i p(I \mid C_i) P(C_i) \]

MAP Tissue Class Estimation

\[ p(C_i \mid I) \propto p(I \mid C_i) p(C_i) \]
GMM Image Example

\[ p(I) = \sum_{i} p(I \mid C_i) P(C_i) \]

Generate new image from model

\[ p(C_i \mid I) \propto p(I \mid C_i) p(C_i) \]

No spatial component
Prior Models

- Markov Random Fields (MRF)
- Structurally-Conditioned Models
- Average Brain
Markov Random Fields (MRF)

Image Intensities $I(x)$

Tissue Labels $C(x)$

$\mathbf{x}$ Pixel Location

$N_x$ Set of neighbors
Markov Random Fields (MRF)

Image Intensities $I(x)$

Tissue Labels $C(x)$

Markov Assumption
Label at $x$ independent of all others, given $N_x$

Hammersley–Clifford Theorem
Probability is a Gibbs distribution over all neighborhood ‘cliques’ $Cl(X)$

Energy Function: Positive

$$p(C \mid I, X) \propto \frac{1}{Z} \exp \sum_{\xi \in Cl(X)} - E(\xi, I, C)$$
Markov Random Fields (MRF)

Penalize intensity/label inconsistencies

\[ E_1(x, I, C) \]

Penalize inconsistent neighbor labels, e.g. Potts model

\[ E_2(x, x', C) = \begin{cases} 1, & \text{if } C(x) = C(x') \\ 0, & \text{otherwise} \end{cases} \]

\[
p(C | I, X) \propto \exp \left[ - \sum_{x \in X} E_1(x, I, C) - \sum_{x' \in N_x} E_2(x, x', C) \right]
\]
Markov Random Fields (MRF)

Iterative Algorithms
• Monte-Carlo Markov Chain
  • Gibbs Sampling
  • Expectation-Maximization

Exact Algorithms
• Graph Cuts
Average Brain Models

- Construct a spatial prior model by averaging tissue distributions over a population [MNI].
P(white matter $| \ x$)
Generated Sample Images

GMM: No spatial component

MRF: Spatial neighborhood

Average Brain: GMM conditioned on location $x$

$$p(I \mid x) = \sum_i p(I \mid C_i) P(C_i \mid x)$$

Note conditional independence assumption:

$$p(I \mid C_i) = p(I \mid C_i, x)$$
Structurally-Conditioned Prior Models

• From (Kapur 1999)
  – Modeling Global Geometric Relationships between Structures
Modeling Global Geometric Relationships between Structures

- Relative Geometry Models
- Motivate Using Knee MRI
Segmented Knee MRI
Motivation

• Primary Structures
  – image well
  – easy to segment

• Secondary Structures
  – image poorly
  – relative to primary

Provided by T Kapur
Relative Geometric Prior Approach

• Select primary/secondary structures
• Measure geometric relation between primary and secondary structures from training data
• Given novel image
  – segment primary structures
  – use geometric relation as prior on secondary structure in EM-MF Segmentation
Segment Primary Structures: Femur, Tibia

Seed
Region Growing
Boundary Localization

Provided by T Kapur
Status

• Have Bone

• Want Cartilage
Measure Geometric Relationship between Primary and Secondary Structures

- Using primitives such as
  - distances between surfaces
  - local normals of primary structures
  - local curvature of primary structures
  - etc.
Measure Geometric Relationship between Primary and Secondary Structures

\[ \rho_s \equiv \text{distance to closest point on bone (femur)} \]

\[ n_s \equiv \text{normal to bone (femur) at closest point} \]

\[ P(x_s \in \text{Cartilage} \mid \text{Bone}) \]

\[ \approx \frac{P(\rho_s, n_s \mid x_s \in \text{Cartilage})P(\text{Cartilage})}{Z} \]
Estimate of \( P(\rho_s, n_s \mid x_s \in \text{Cartilage}) \)

\[ P(\rho_s, n_s \mid x_s \in \text{Fem. Cartilage}) \quad P(\rho_s, n_s \mid x_s \in \text{Tib. Cartilage}) \]

Provided by T Kapur
Results: Segmentation of Femoral & Tibial Cartilage

MRI Image

Model-Based Segmentation

Manual Segmentation

Provided by T Kapur
Morphological Operations

- Erosion
- Dilation
- Opening
- Closing

- [Haralick + 1989]
Morphological Operators...

• Ubiquitous simple tools. Useful for ad-hoc clean-up of results from Statistical Classification.
Dilation

- Binary (or Boolean) images
- Represent image by a set of coordinate vectors of pixels with value 1

\[ A \oplus B \equiv \{ c \mid c = a + b, \text{for some } a \in A, b \in B \} \]

Typical structure elements:

\[
\begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}
\]
Dilation

- Continuous analogy
- Makes structures \textit{fatter}

Erosion

• Erosion is dual of dilation
  – complement A
  – reflect B (negate coordinates)
  – dilate
  – complement result

\[ A \otimes B = \overline{A} \oplus \hat{B} \]

• Frequently, B is symmetric and then reflection can be ignored
Erosion

- Erosion by simple S.E.’s makes structures thinner
- Analog analogy:

![Diagram showing erosion and structure elements](Source: Wikipedia)
Opening

• Opening = Erode then Dilate

Break thin connections

Removes small junk
Closing

• Closing = Dilate then Erode
• Can attach objects that have become fragmented
Erosion and Dilation

• Common trick in brain isolation “de-scalping”
  – Erode “it”
    • to disconnect brain from head
  – Dilate “it”
    • But *only* mark pixels that were originally “brain”
Connectivity

• Define neighbor relation

- There are some inconsistencies that a 6-neighbor relation can fix
Connected Components

- Input: Boolean image objects
- Output: Unique label for each separate object
Finding Connected Components

- \( N = 1 \) (initialize label)
- Repeat until all pixels are labeled
  - Pick an unmarked \( l \) pixel
  - Label it, and all of its \( l \) neighbors,
    and all of their neighbors: \( N \)
  - \( N \leftarrow N + 1 \)

See Flood Fill Algorithm
Segmentation Quality Evaluation

Compare overlap between automatic and manual (ground truth) segmentation.

\[ DICE \equiv 2 \frac{|A \cap B|}{|A| + |B|} \]

Dice, Jaccard Measures
Topics Not Discussed

• Edge-base segmentation
  – Active contours, level sets

• Hierarchical representations
  – Coarse-to-fine modeling
Selected References

Further Reading


Feature-based Analysis

- Model image as a set of image patches
- 3D scale-invariant features (SIFT)

Feature-based Analysis

- Alignment/Registration: fast and robust

Feature-based Alignment (robust to perturbation)

Global Image Alignment (sensitive to perturbation)
Feature-based Analysis

- Model features over many images.

Infant brain development.

A Feature-based Developmental Model of the Infant Brain in Structural MRI.