Introduction to Medical Image Segmentation

HST 582

Matthew Toews

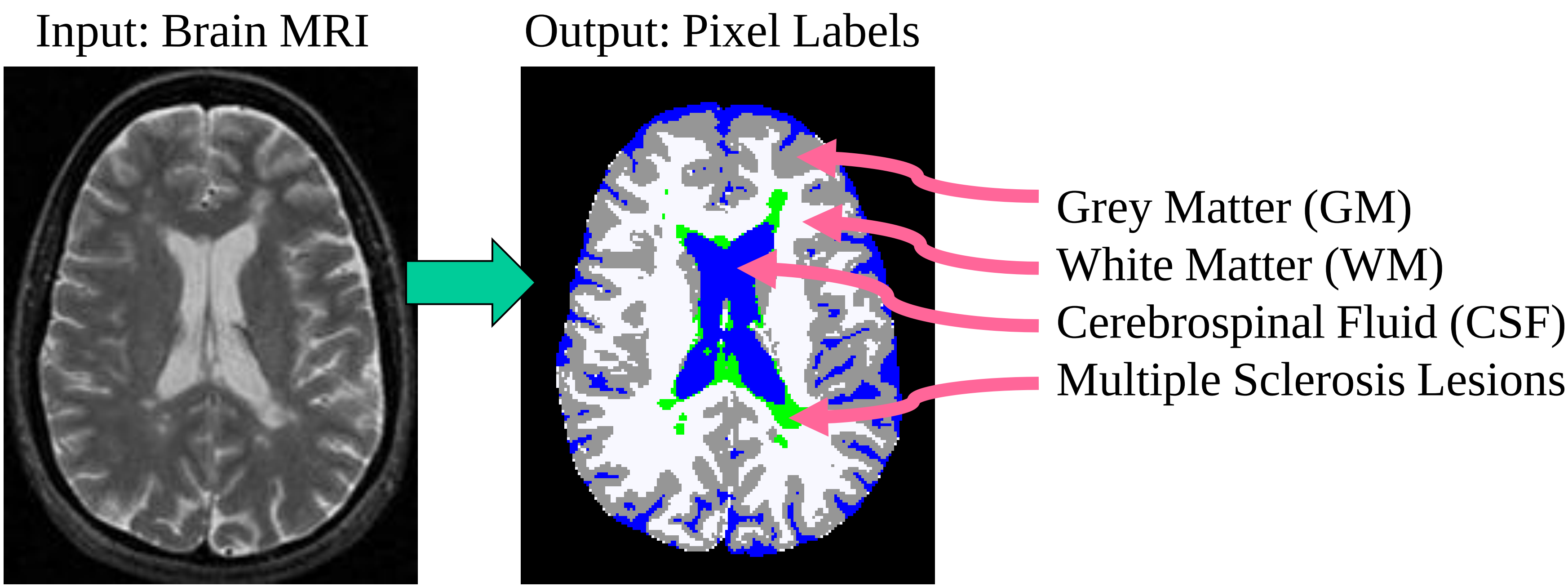
(slides adapted from William Wells III)

Outline

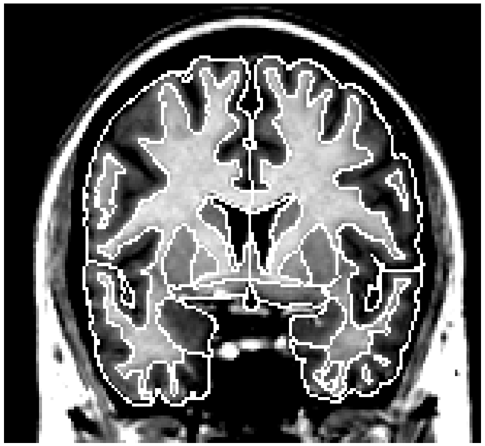
- Terminology & Applications
- Probability Review
- Intensity-Based Classification
- Prior models
- Morphological Operators

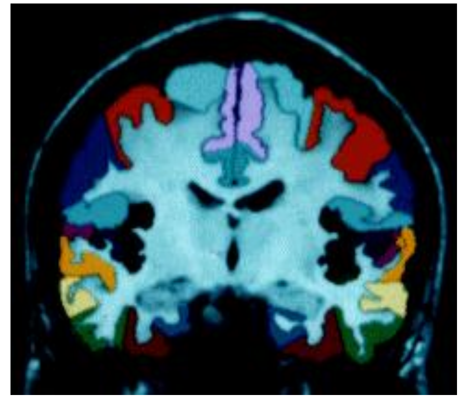
Image Segmentation

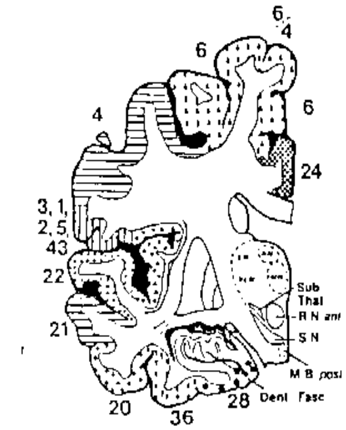
Partitioning an image into regions defined by pixel intensity and geometry.

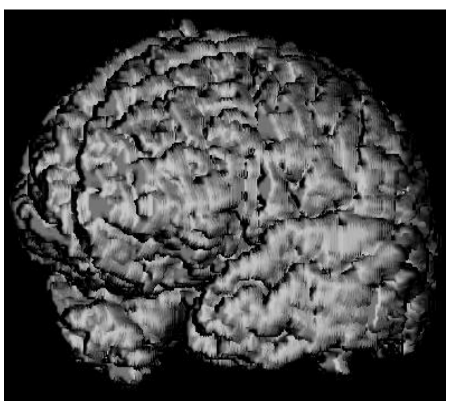


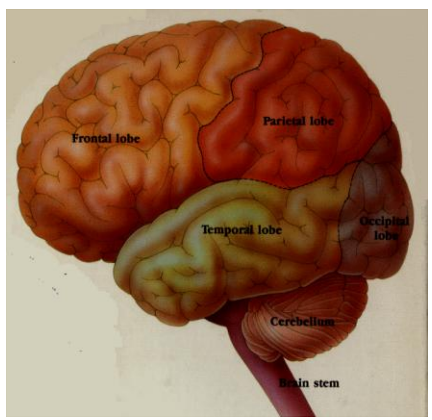
Anatomical Description Hierarchy

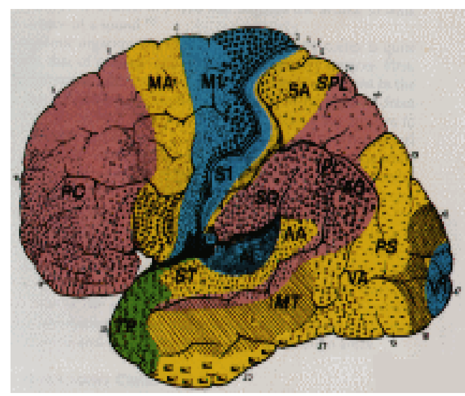


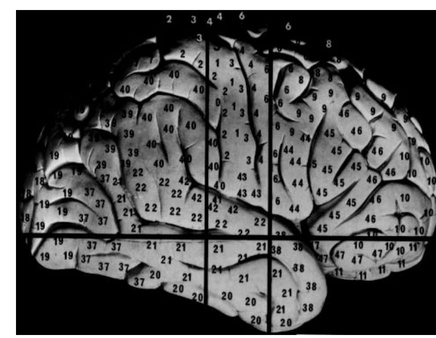












WHOLE BRAIN/STRUCTURE

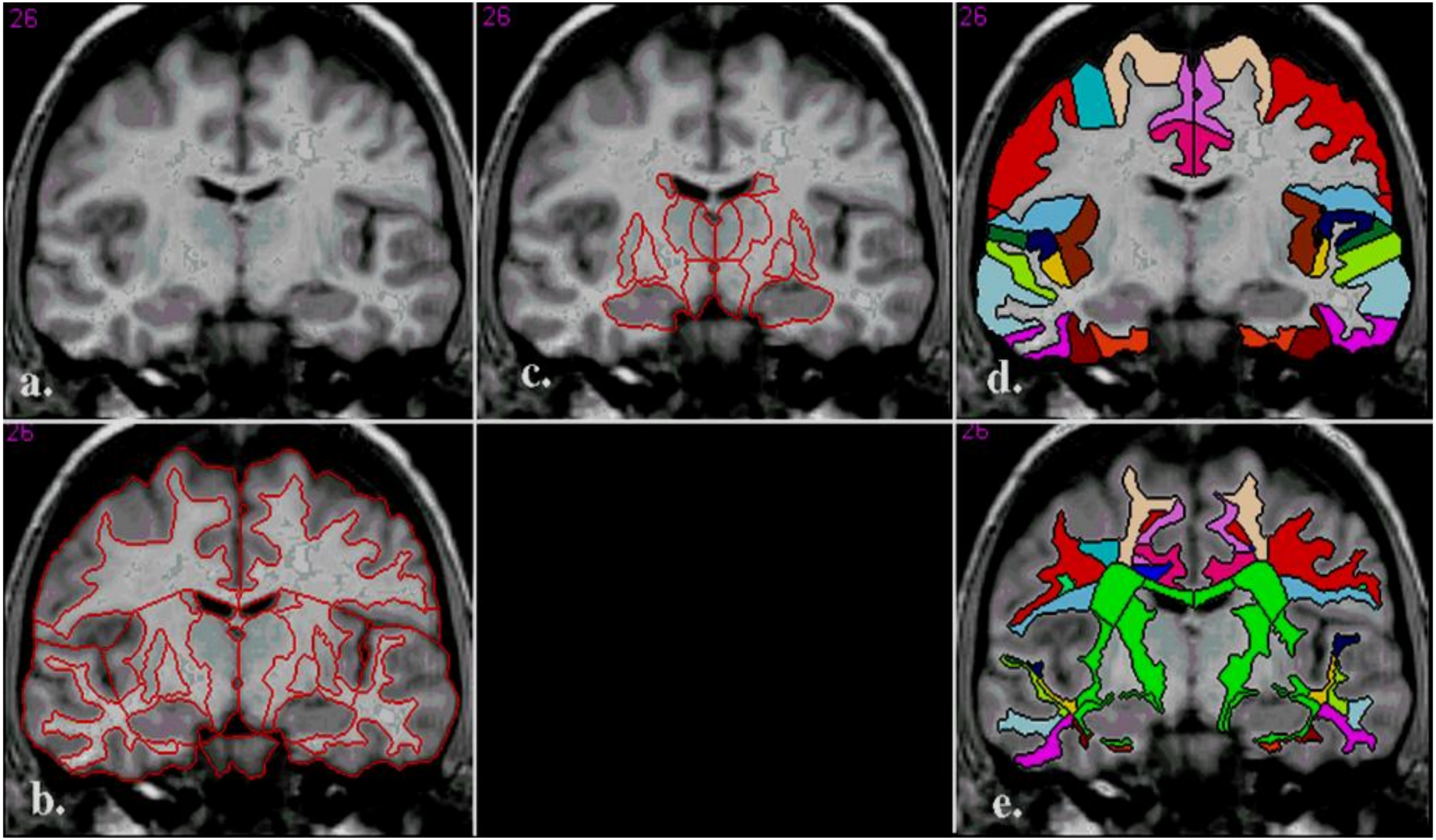
LOBES

SYSTEMS

CYTO- & MYELO-ARCHITECTURE

Provided by D Kennedy

Stages of Anatomic Analysis



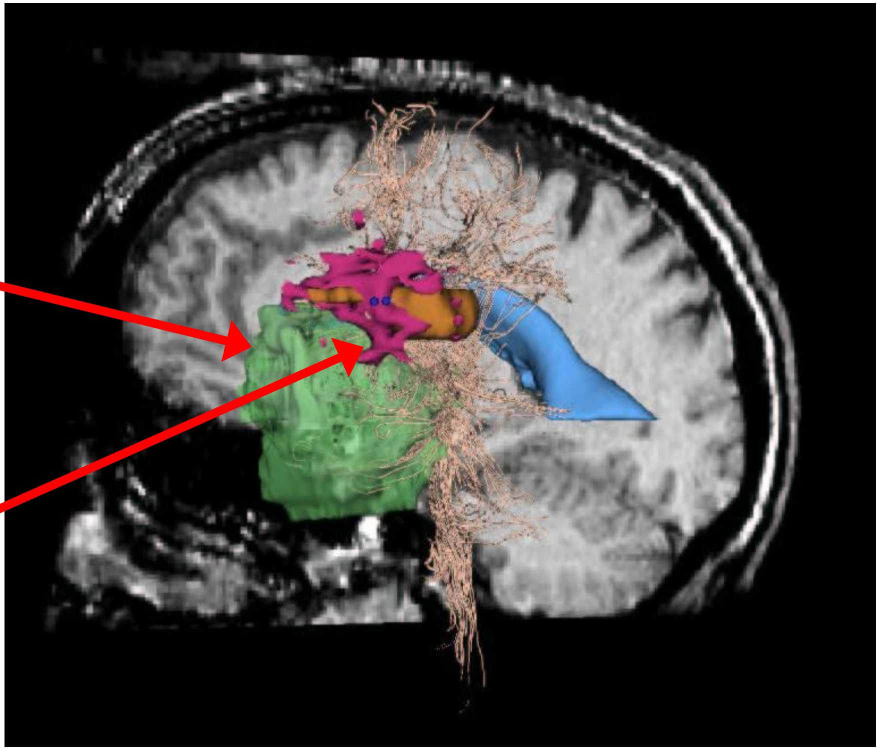
Provided by D Kennedy

Medical Applications of Segmentation

- Image Guided Surgery
- Surgical Simulation
- Neuroscience Studies
- Therapy Evaluation

fMRI and Electro Corticography

Tumor to be resected



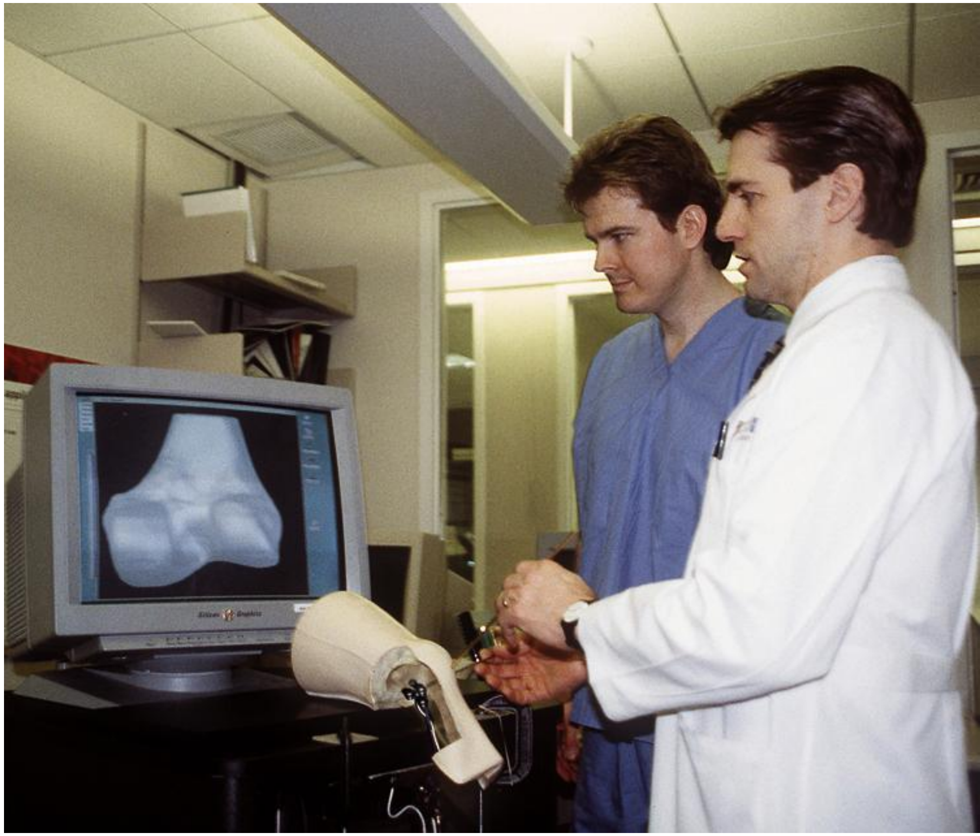
Functional regions to be preserved (e.g. language)

- **Quantitative comparison of functional MRI and direct electrocortical stimulation for functional mapping.** Larsen S, Kikinis R, Talos IF, Weinstein D, Wells W, Golby A. Int J Med Robot. 2007 Sep;3(3):262-70.

Applications of Segmentation

- Image Guided Surgery
- Surgical Simulation

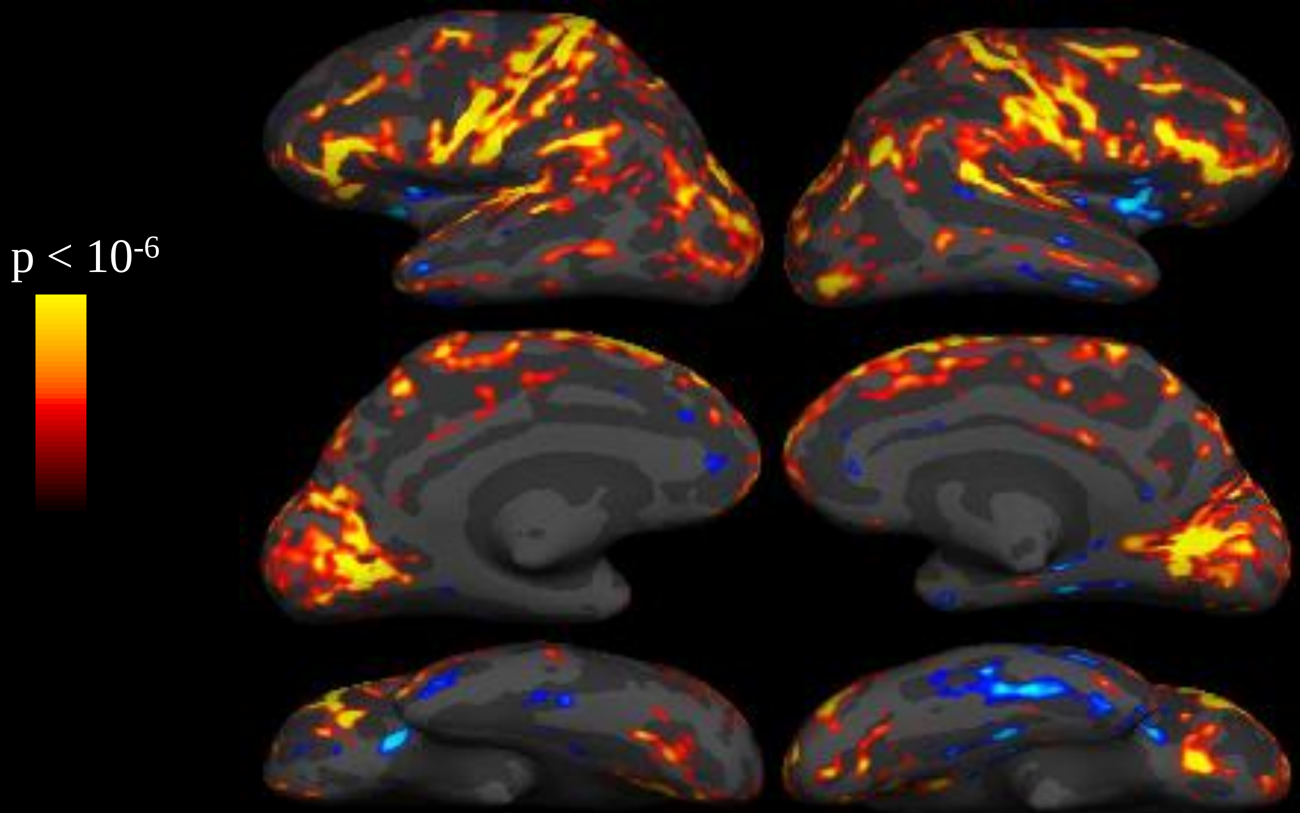
Personalized orthopedic surgery



Applications of Segmentation

- Neuroscience Studies

Statistical Map of Cortical Thinning: Aging



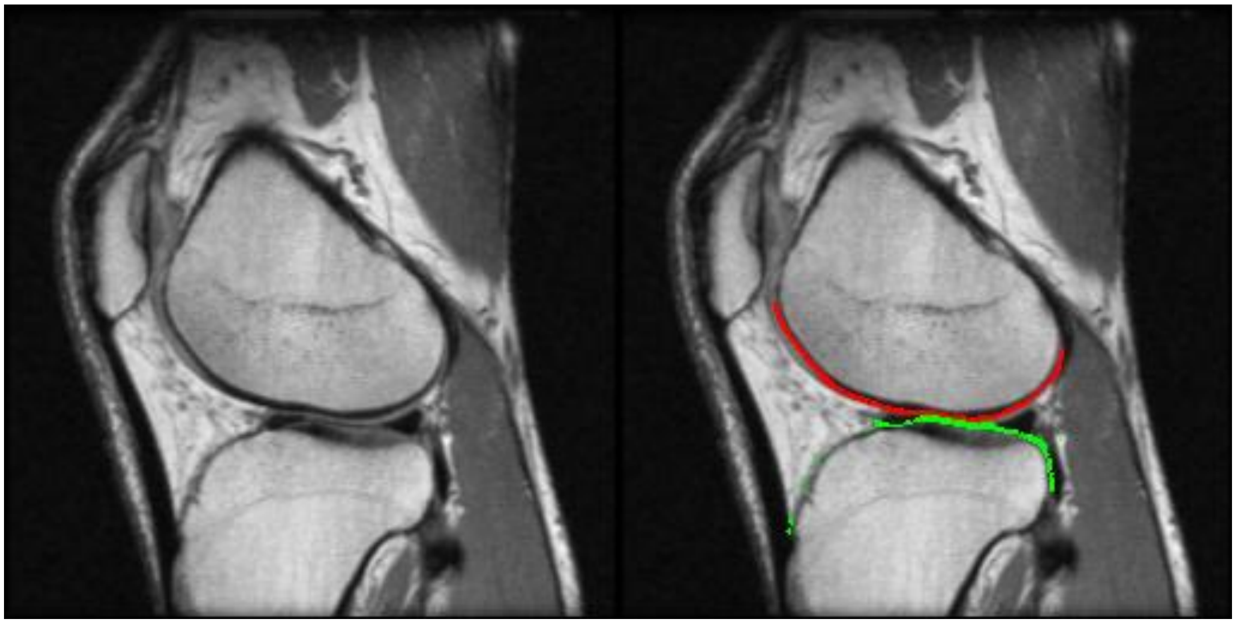
Thanks to Drs. Randy Buckner and David Salat for supplying this slide.

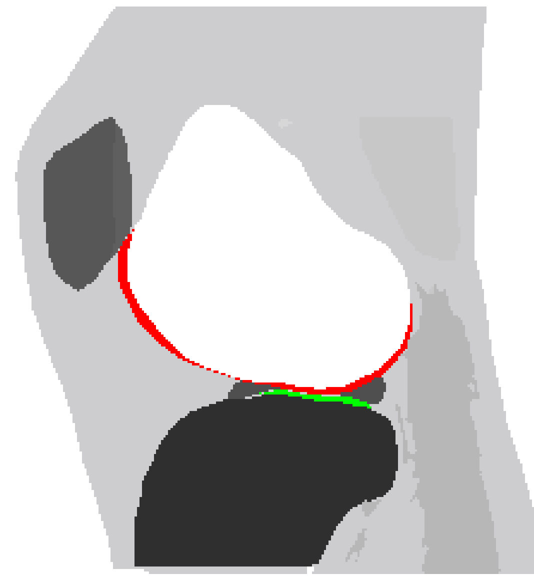
Provided by Bruce Fischl

Applications of Segmentation

- Therapy Evaluation
  - Multiple Sclerosis
  - Knee Cartilage Repair

Results: Segmentation of Femoral & Tibial Cartilage





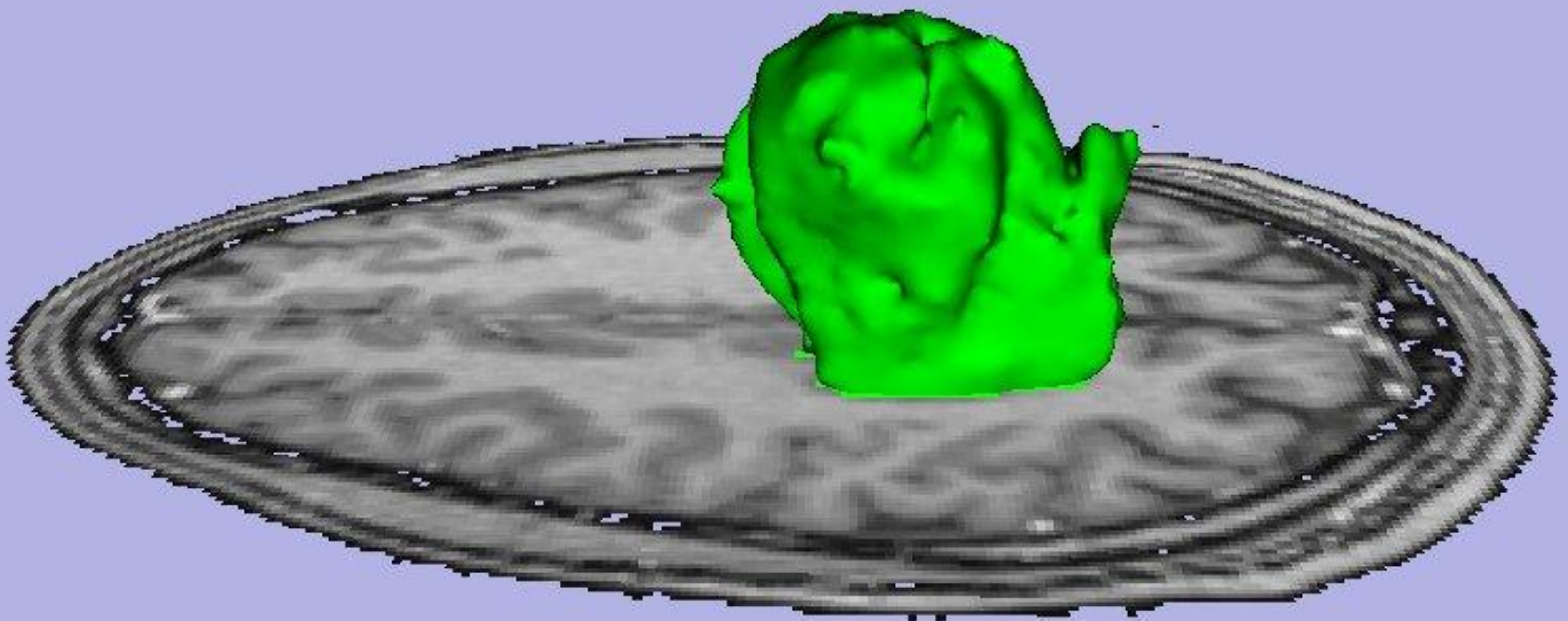
MRI Image

Model-Based Segmentation

Manual Segmentation

Provided by T Kapur

Manual Segmentation



Provided by D. Gering

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

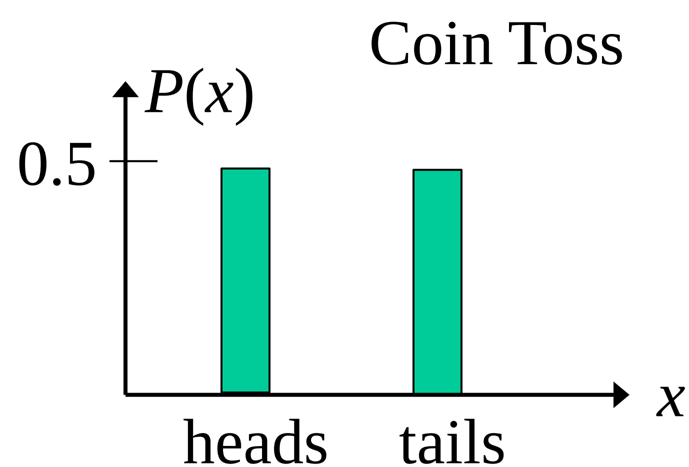
- Discrete Random Variables (RV)
  - Probability Mass Functions (PMF)
- Continuous Random Variables
  - Cumulative Distribution Functions (CDF)
  - Probability Density Functions (PDF)
- Conditional Probability
- Bayes' Rule

Discrete Random Variable

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

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

$$\sum_x P(x) = 1$$



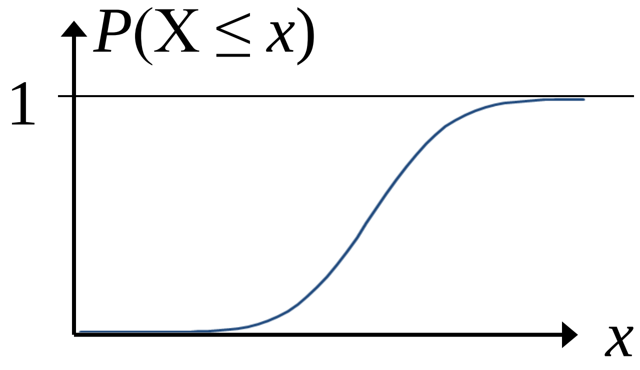
Continuous Random Variables

- 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*



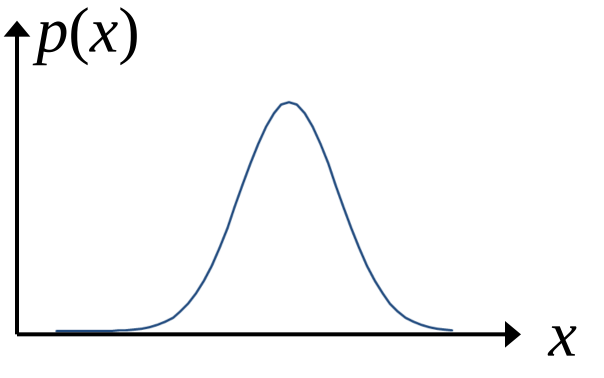
Continuous Random Variables...

- Define *Probability Density Function* (PDF)

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

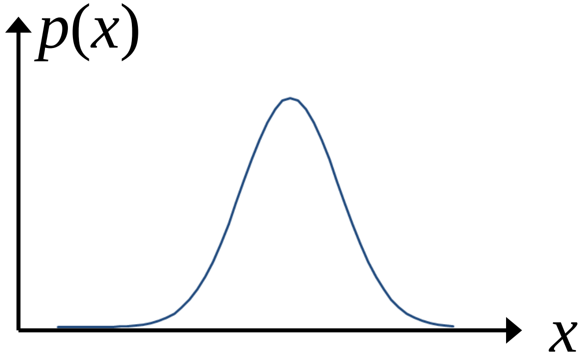
- Easy to show, using Fundamental Theorem of Calculus:

$$P(a \leq x \leq b) = \int_a^b p(x)dx$$



More on PDFs : $p(x)$

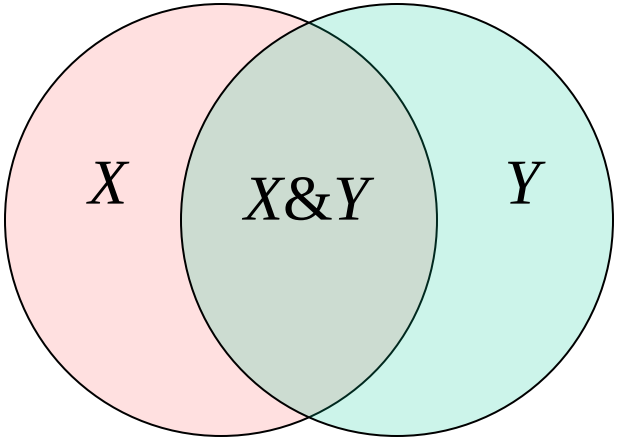
- Non Negative
- Integrates to One
- (Value can be Greater than One)



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

Classification of voxels

[Duda, Hart 78]

Measurement Model

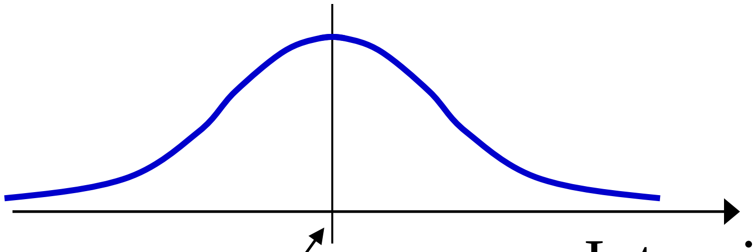
- Characterize sensor

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

intensity

probability density

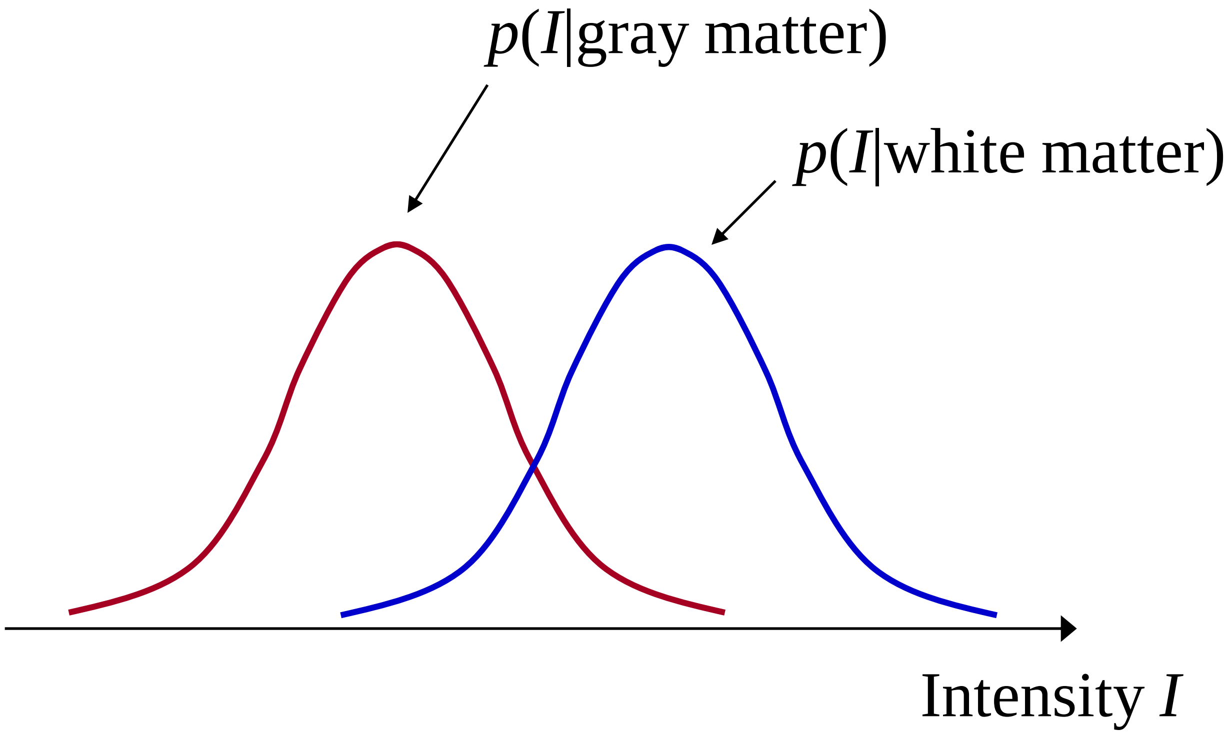
Tissue class conditional model of signal intensity



Intensity I

Mean intensity of tissue J

Example



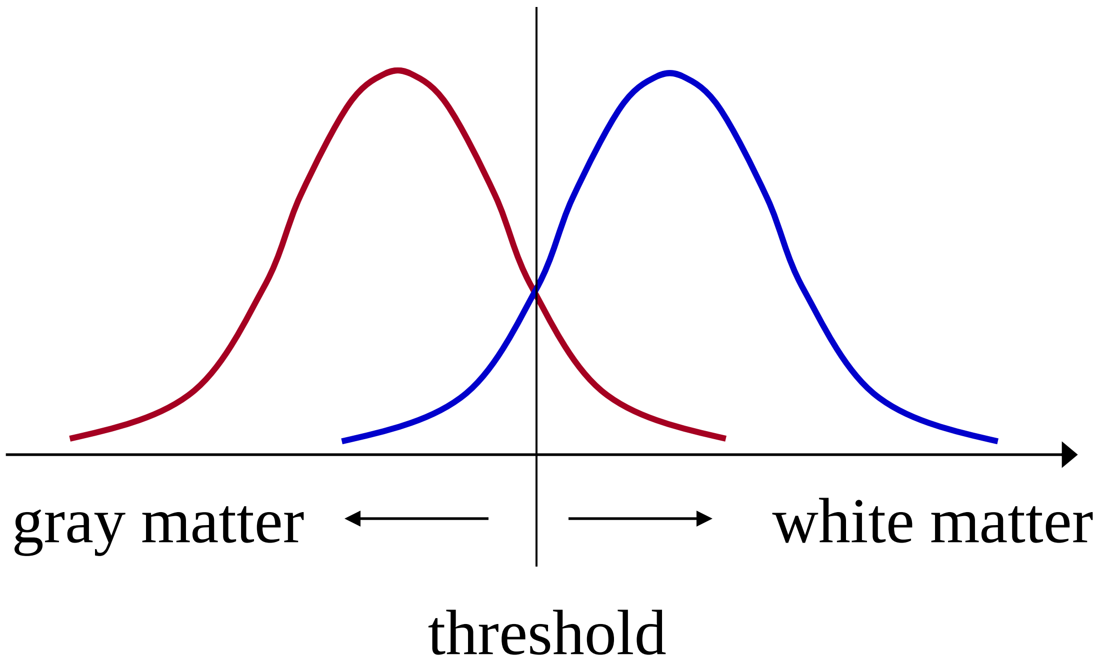
Maximum Likelihood Classification

- Measure intensity, I_o, and we want to know the tissue class

$$\boxed{L(TC_j) = p(I_0 \,|\, TC_j)}$$

- Pick tissue class that maximizes L
- L is not a probability
  – Called: Likelihood

Example - revisited



Anatomical Knowledge

- *A priori* model
  - Before the measurement is considered

$$\boxed{P(TC_j)}$$

MAP Classifier

- Choose TC to Maximize the *A Posteriori* probability

measurement model

prior

$$\boxed{P(TC \mid I_0) = \frac{p(I_0 \mid TC)P(TC)}{p(I_0)}}$$

posterior probability

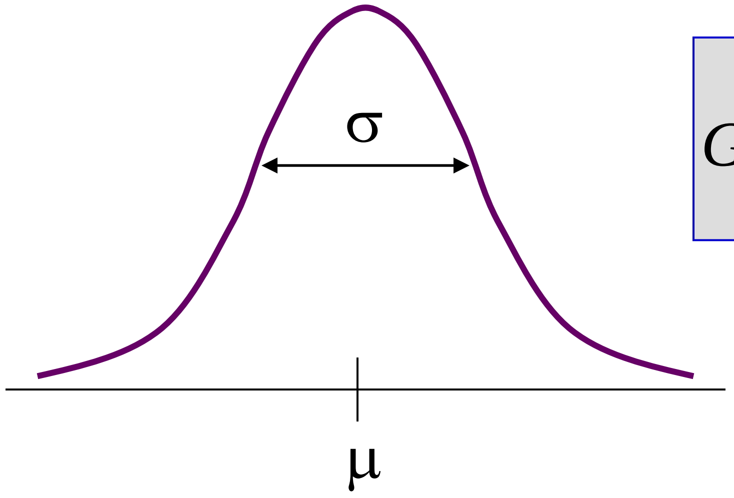
not important

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 σ and μ

- Use training data: $\left\{ I_1, I_2, \ldots, I_N \right\}$
- ML parameter estimation

$$\mu = \frac{1}{N} \sum_i I_i \qquad \sigma^2 = \frac{1}{N} \sum_i (I_i - \mu)^2$$

- MAP tissue classifier with Gaussian measurement model: choose tissue class to maximize:

$$P(TC_j \mid I) = \frac{G(\mu_j, \sigma_j, I) P(TC_j)}{\ldots}$$

Gaussian Density – 2d Data

- Example

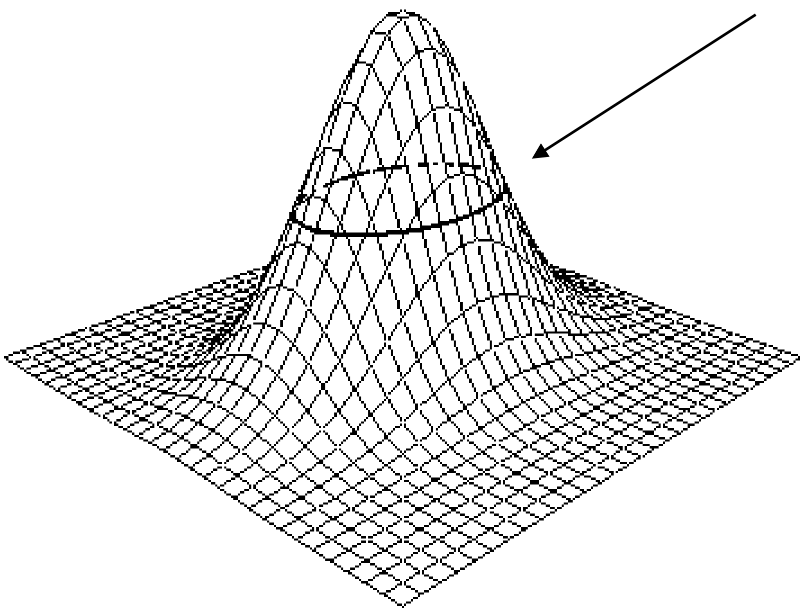
$$I = \begin{pmatrix} \text{proton density intensity} \\ \text{T2 weighted intensity} \end{pmatrix}$$

Mean Vector M

Covariance Matrix Σ

$$\boxed{G(M, \Sigma, X) = \frac{1}{(2\pi)^{\frac{N}{2}} |\Sigma|^{\frac{1}{2}}} e^{-(I-M)^T \Sigma^{-1} (I-M)}}$$

2D Gaussian: Example

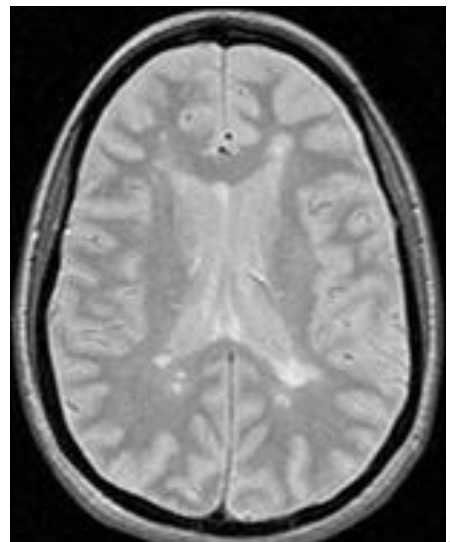


iso probability contour is ellipse

Multiple Sclerosis Example

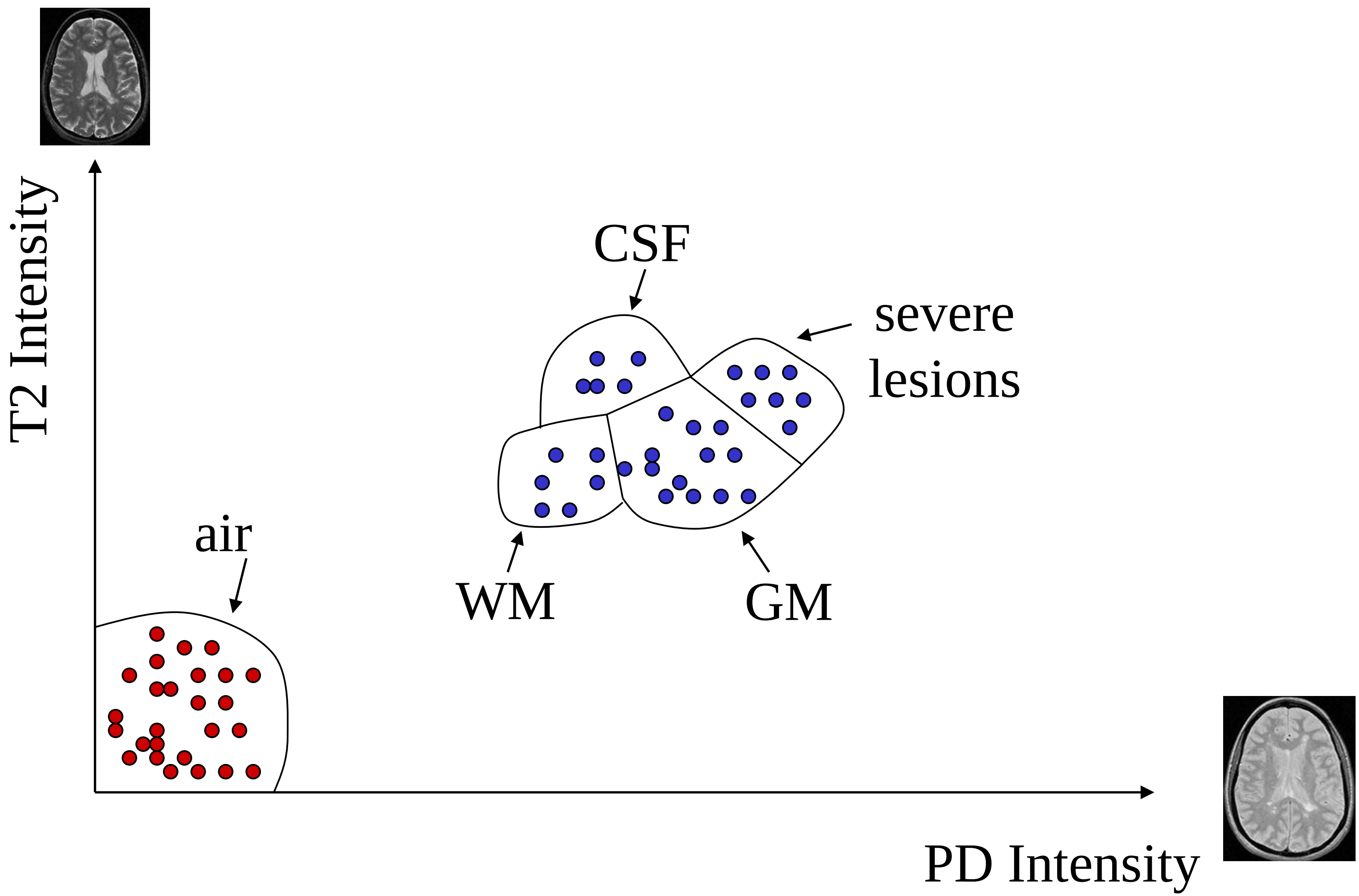
- Dual echo MRI
  - 1 x 1 x 3 mm
  - Registered slice pairs
- Proton density image
  - Good: white/gray
  - Bad: gray/CSF
- T2-weighted image
  - Not so good: white/gray
  - Good: CSF/MS lesions

PDw



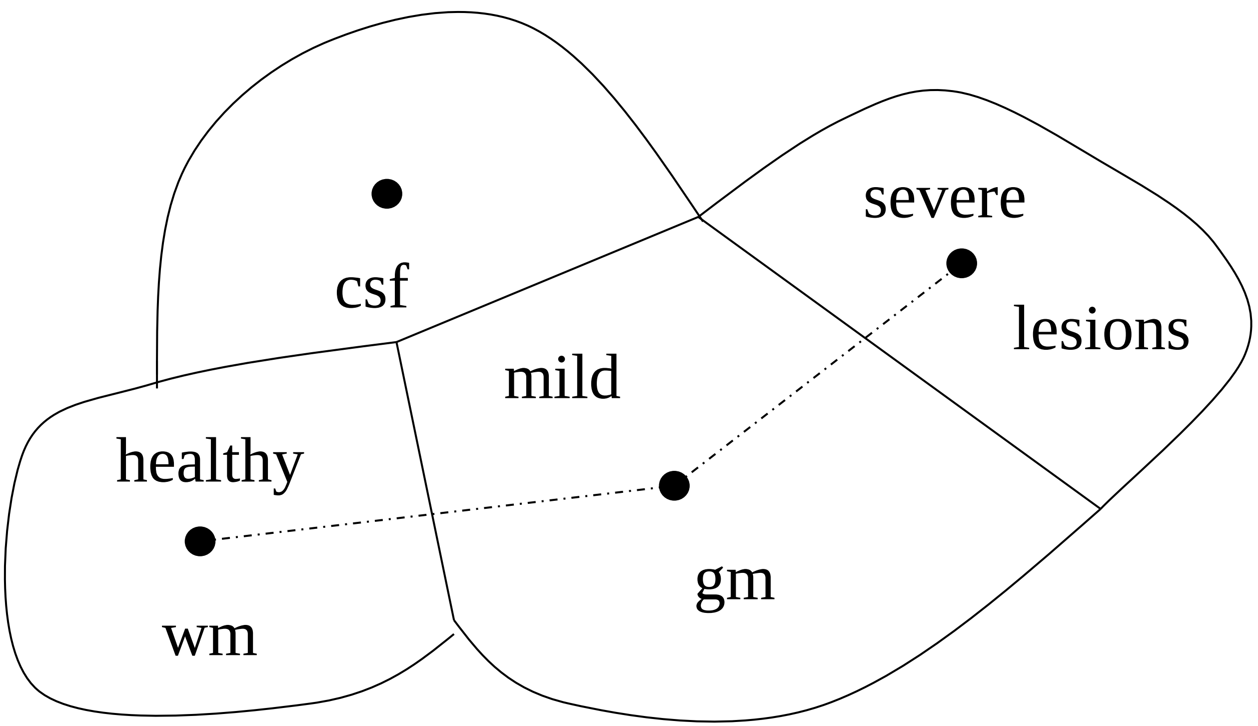
T2w

Dual Echo MRI Feature Space

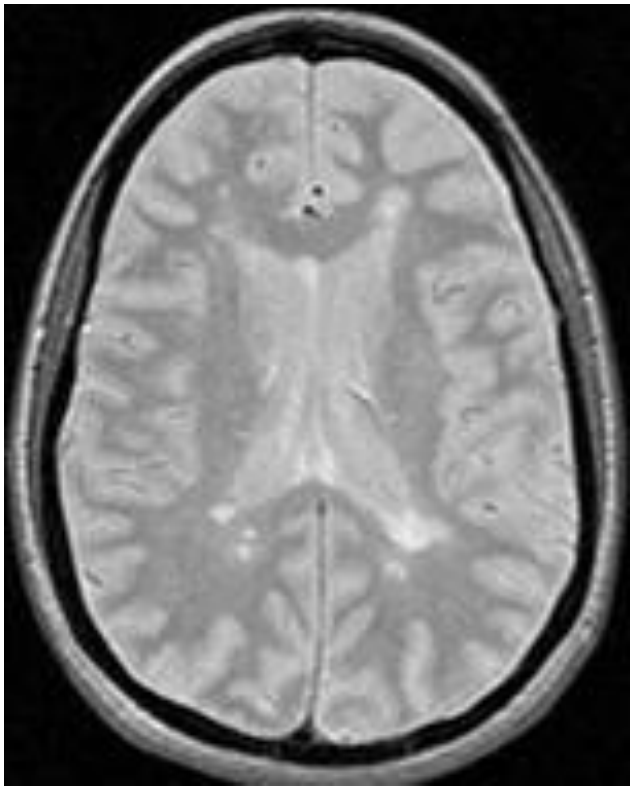


Detail

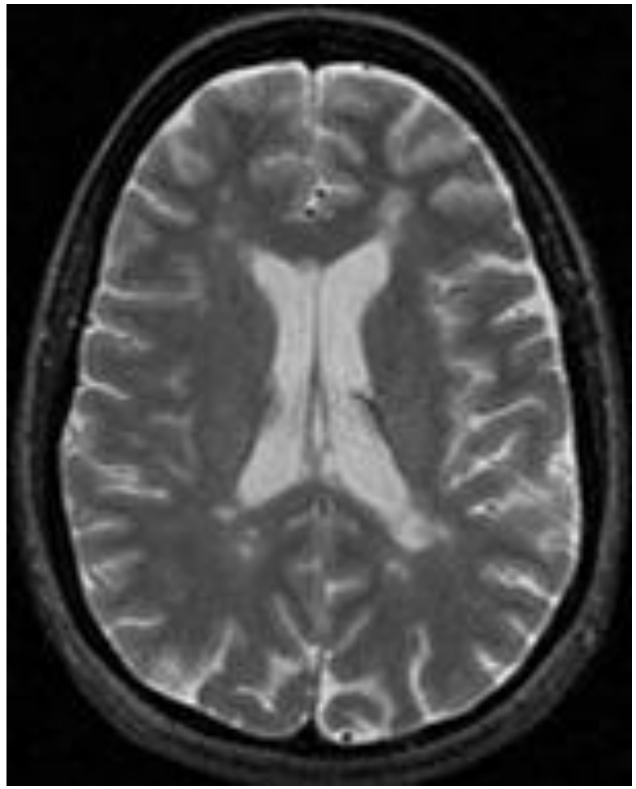
- MS Lesions are “graded phenomenon” in MRI, and can be anywhere on the curve



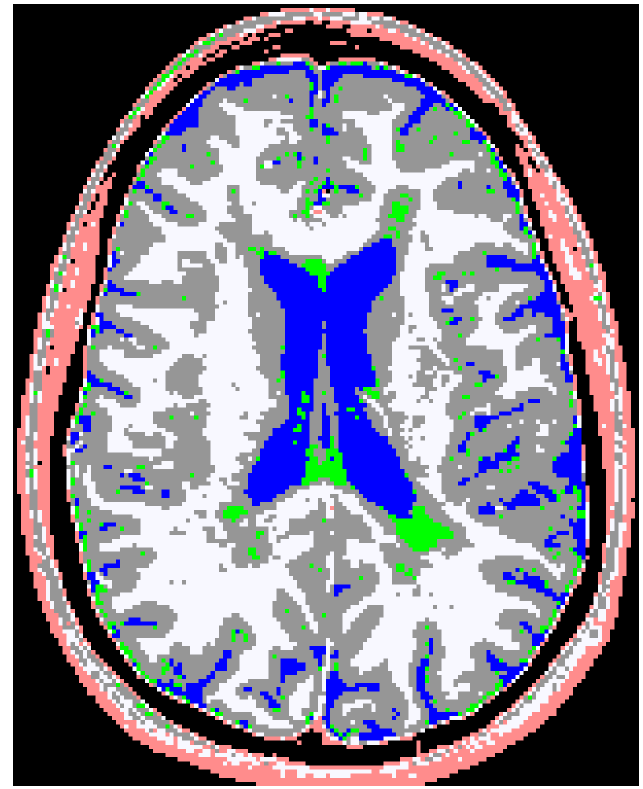
Multiple Sclerosis



PDw



T2w



Segmentation

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

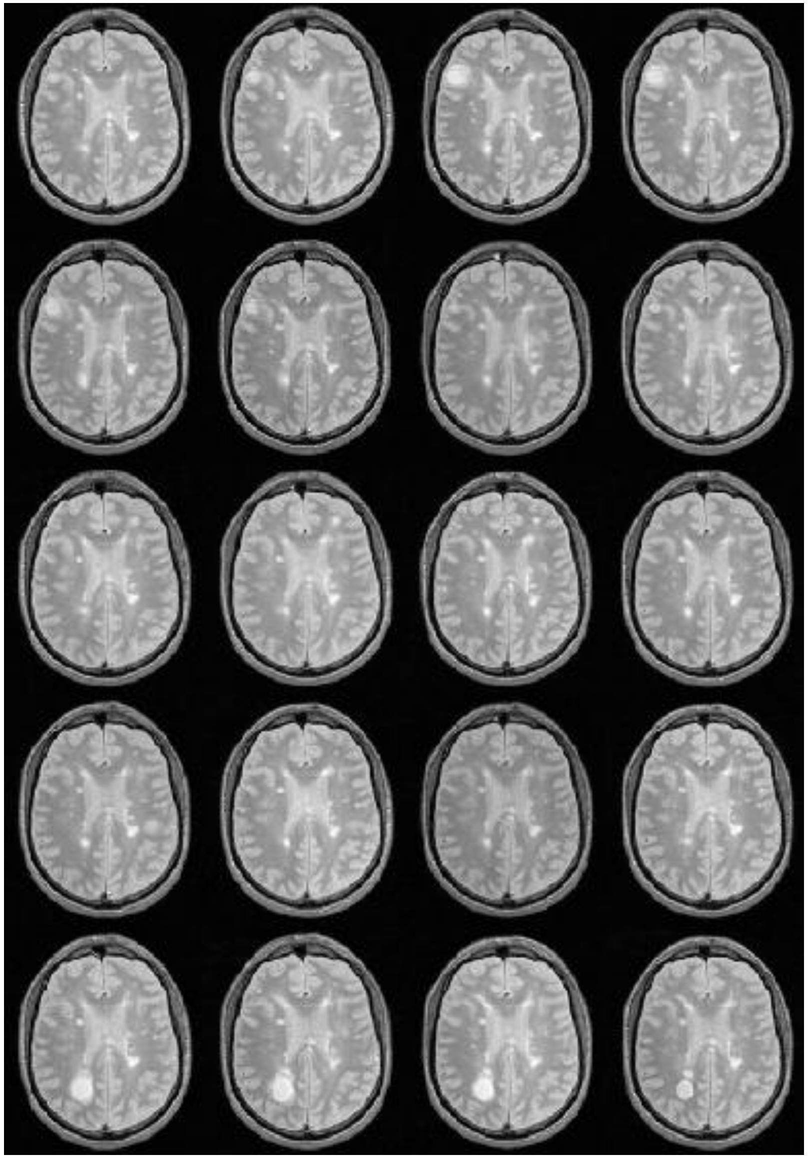
Compute tissue posteriors using current intensity correction.

Estimate intensity correction using residuals based on current posteriors.

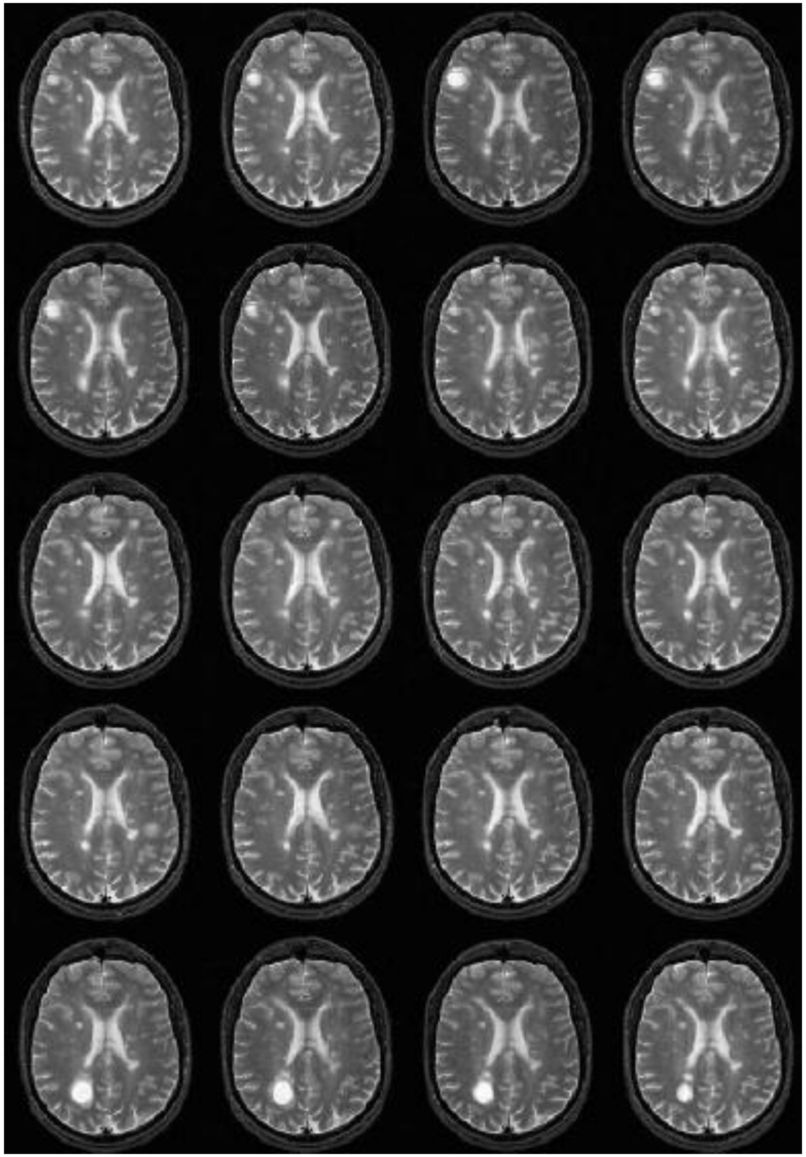
M-Step

Provided by T Kapur

Dual Echo Longitudinal Study



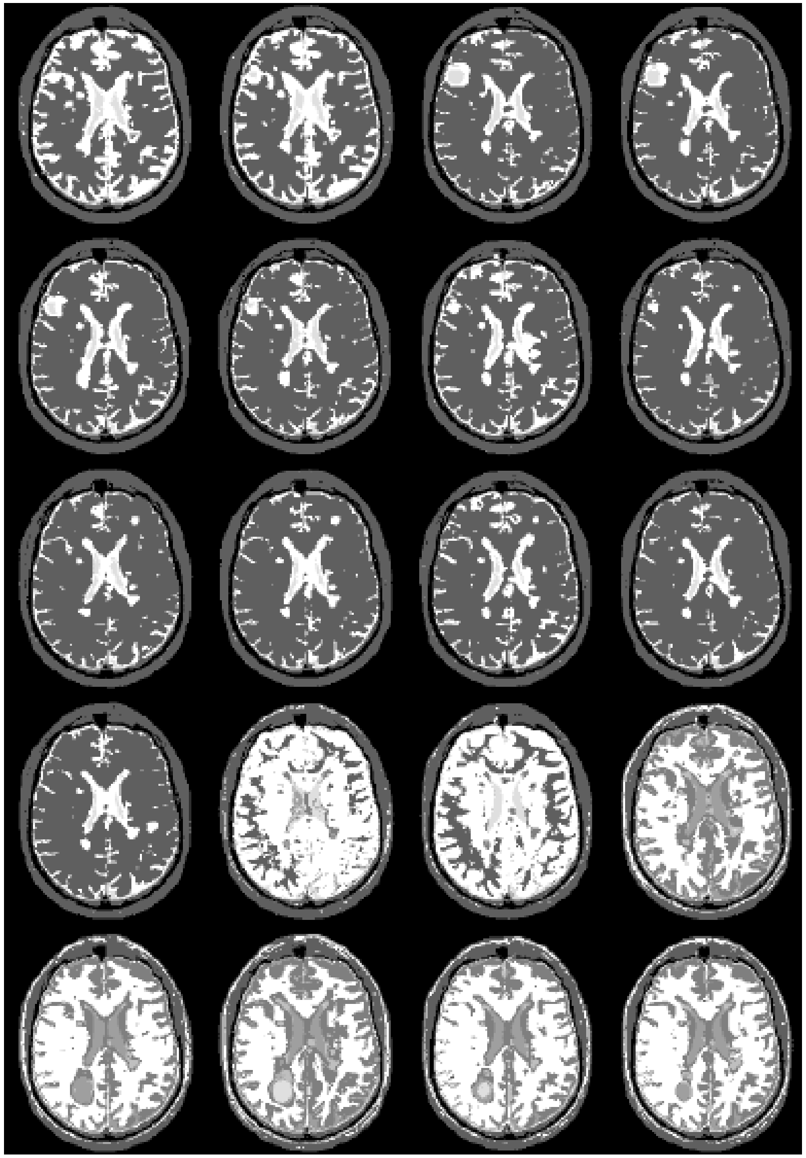
PDw



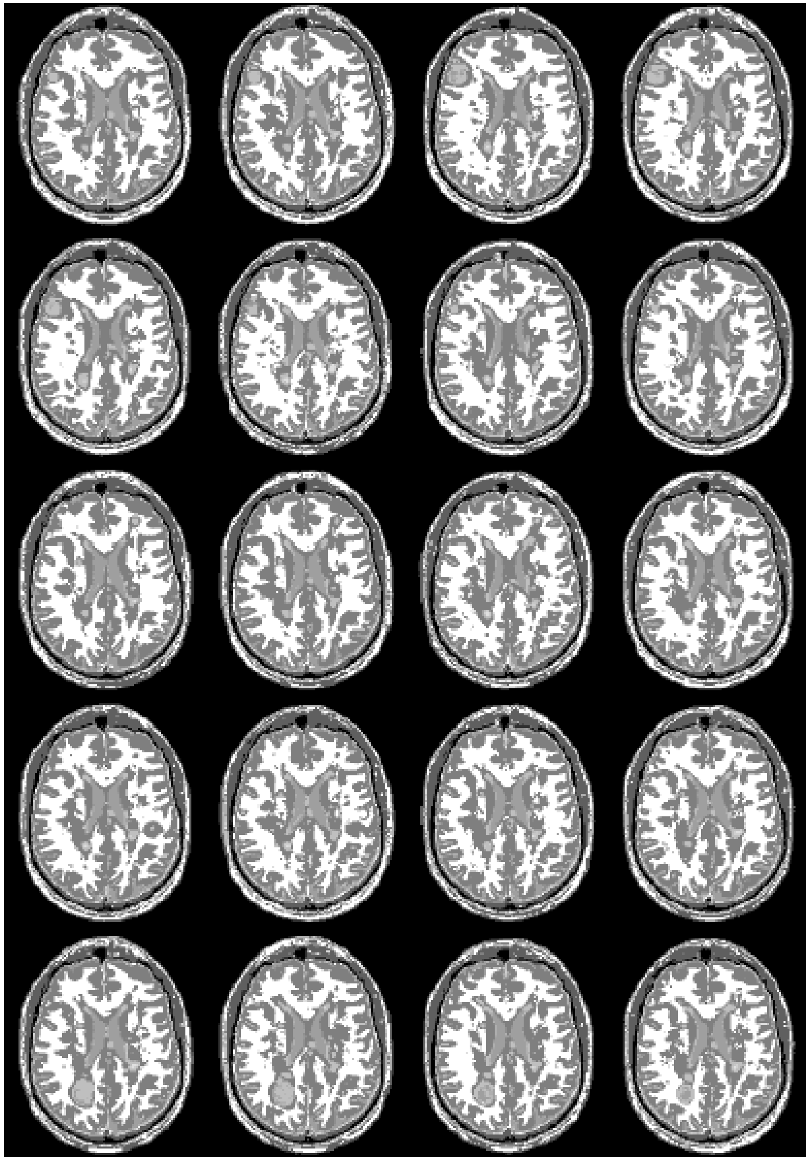
T2w

Provided by S Warfield

Tissue classification

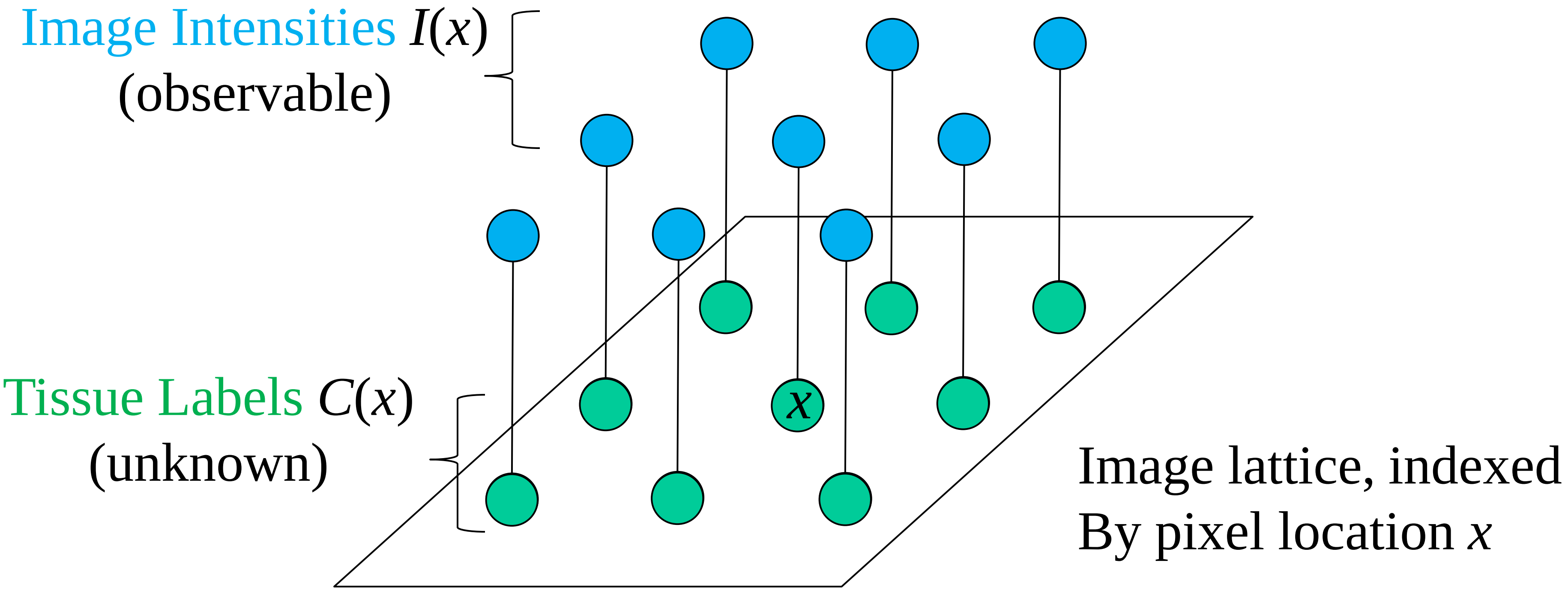


No Intensity Correction



EM Segmentation

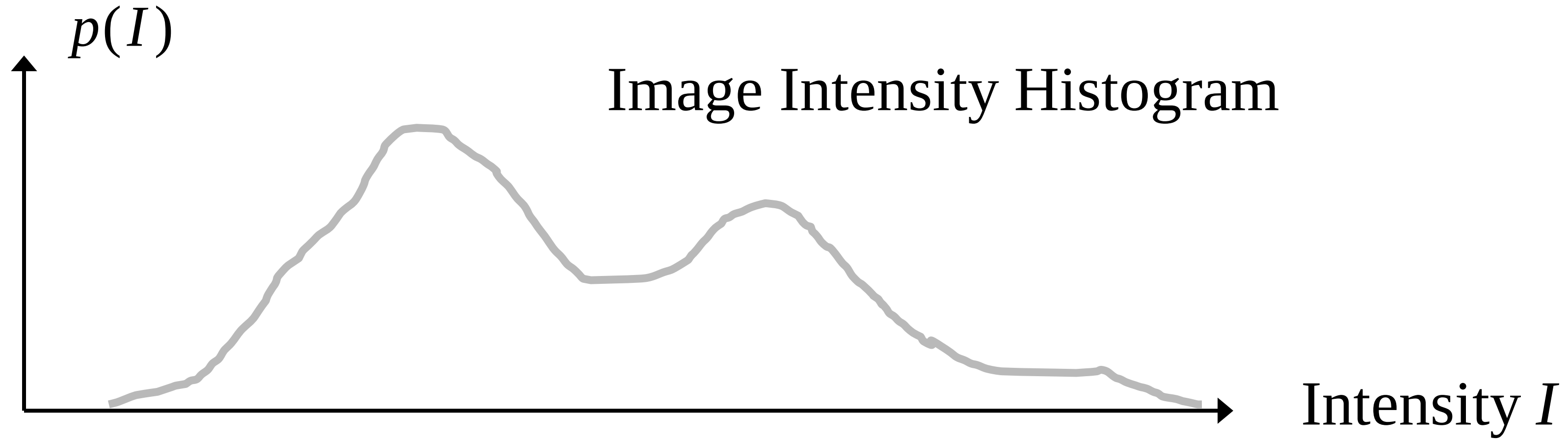
Generative Model



Gaussian Mixture Model (GMM)

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

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



GMM Example

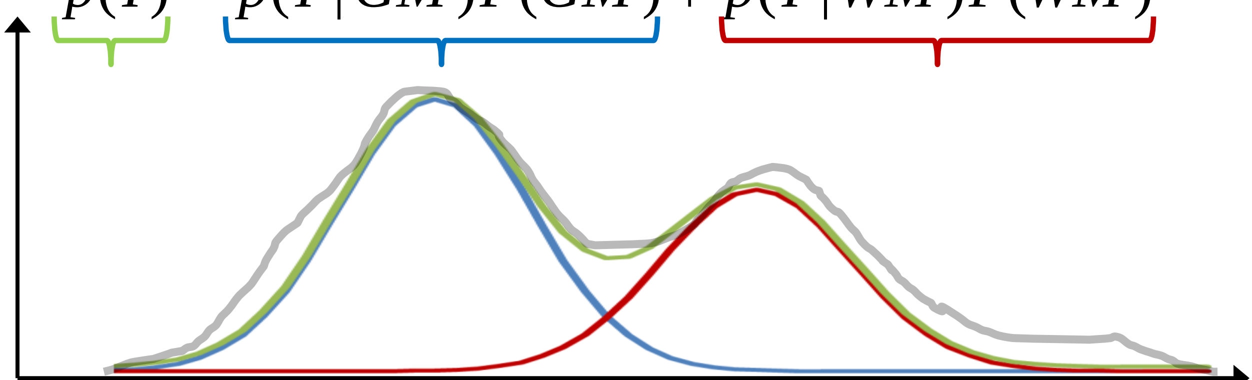
Two Tissue Types

Grey Matter (GM), White Matter (WM)

ML Parameter Estimation

- Expectation-Maximization (EM) Algorithm

$$p(I) = p(I \mid GM)P(GM) + p(I \mid WM)P(WM)$$



Intensity I

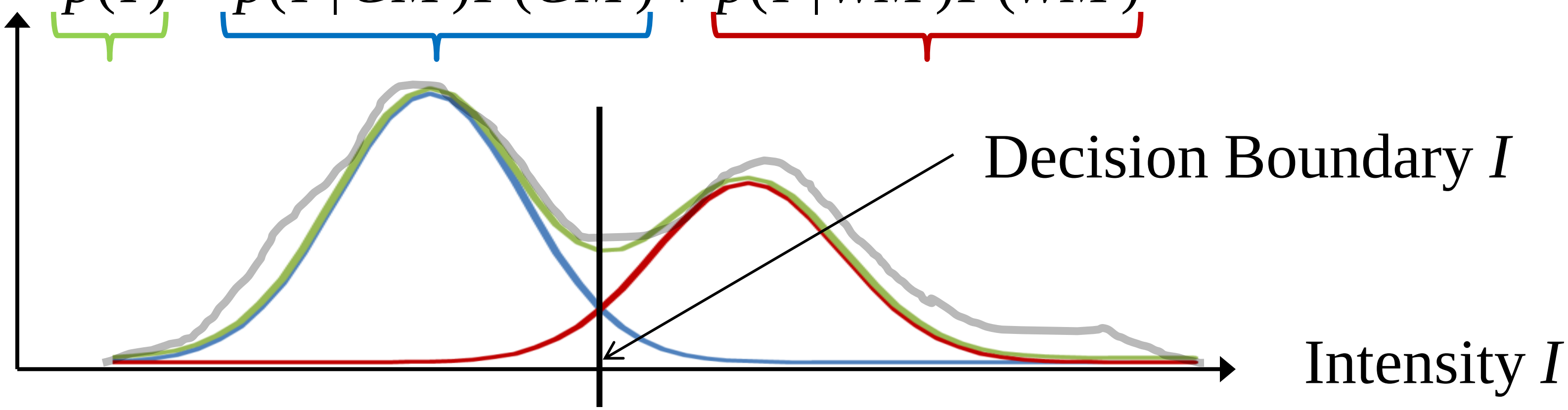
GMM Segmentation

Maximum A-Posteriori (MAP) Tissue Estimation

- Identify most probable C given intensity sample I

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

$$p(I) = p(I \mid GM) P(GM) + p(I \mid WM) P(WM)$$



GMM Sampling

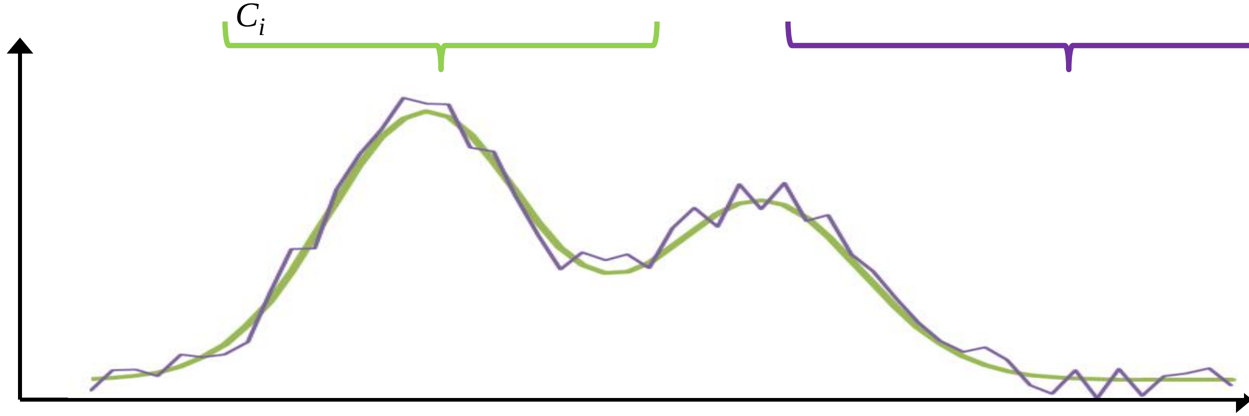
Generate new image histogram

- draw samples from GMM

Check modeling assumptions

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

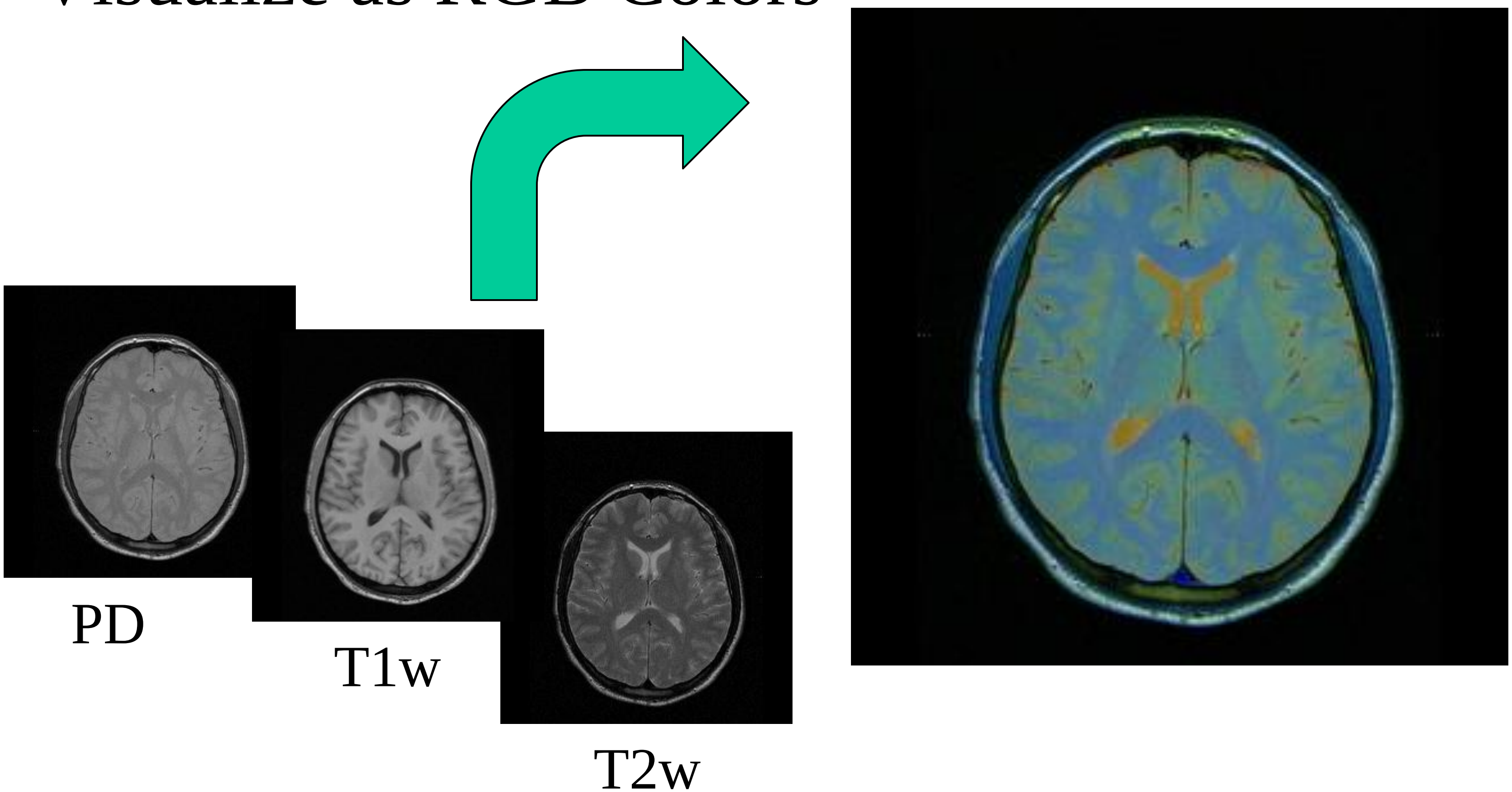
Sampled Histogram



Intensity I

GMM Image Example

Visualize as RGB Colors

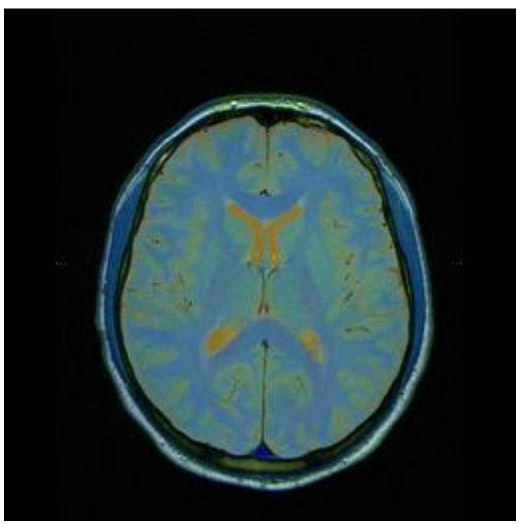


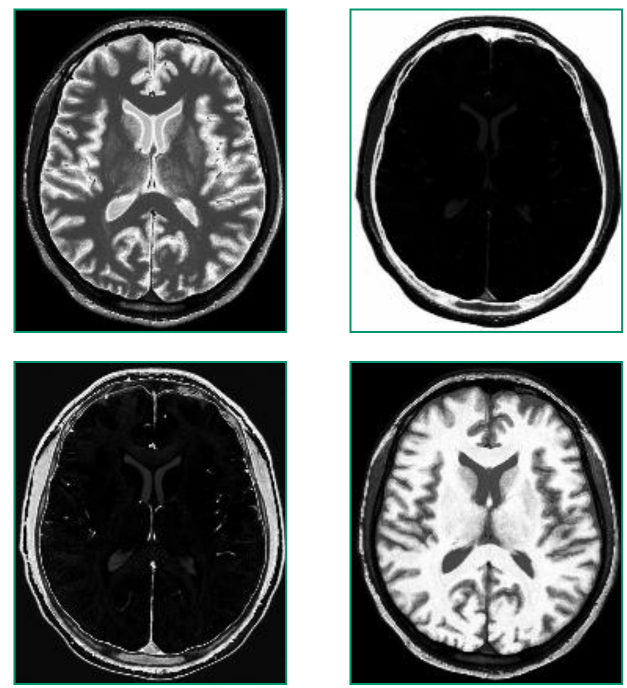
GMM Image Example

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

Learn 4-tissue model

$$p(C_i \mid I) \propto p(I \mid C_i) p(C_i)$$



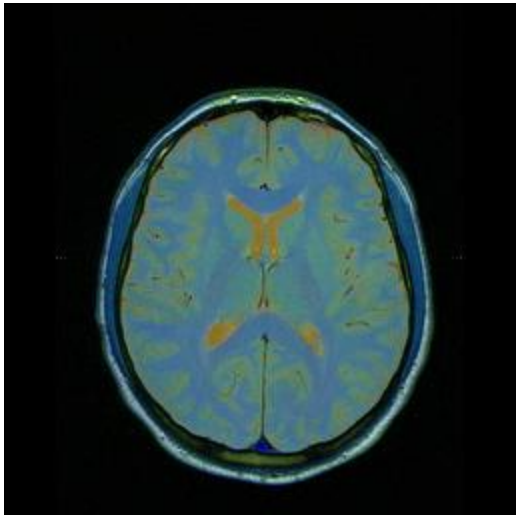


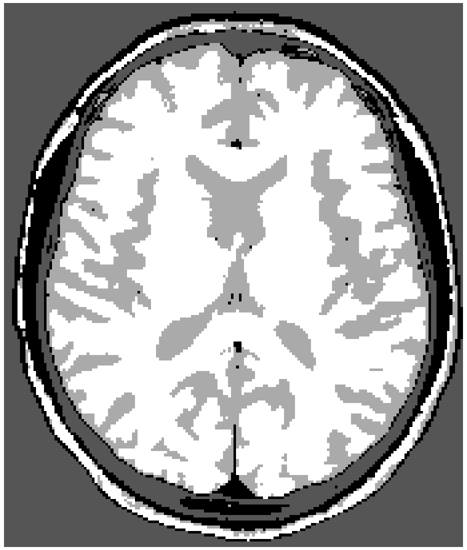
GMM Image Example

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

MAP Tissue Class Estimation

$$p(C_i \mid I) \propto p(I \mid C_i)\, p(C_i)$$



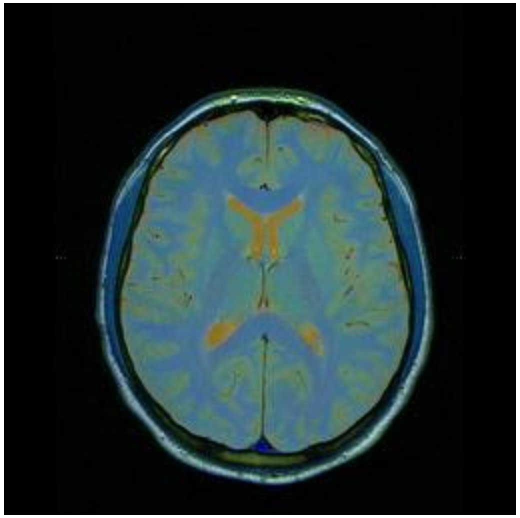


GMM Image Example

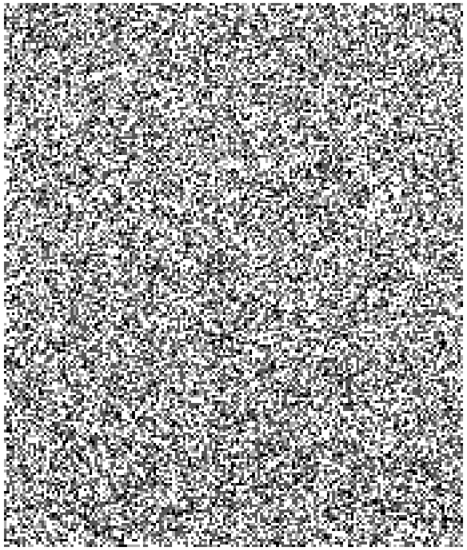
$$p(I) = \sum_i p(I \mid C_i) P(C_i)$$

Generate new image from model

$$p(C_i \mid I) \propto p(I \mid C_i) p(C_i)$$



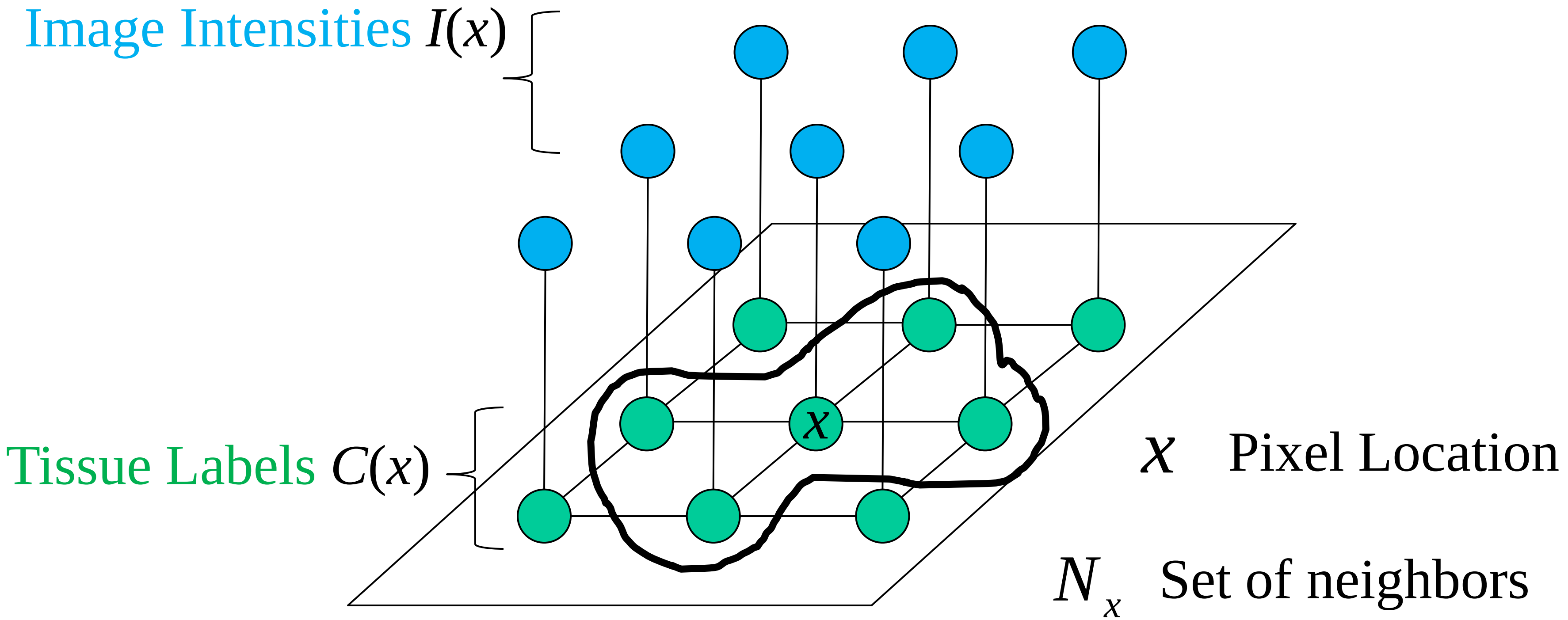
No spatial component



Prior Models

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

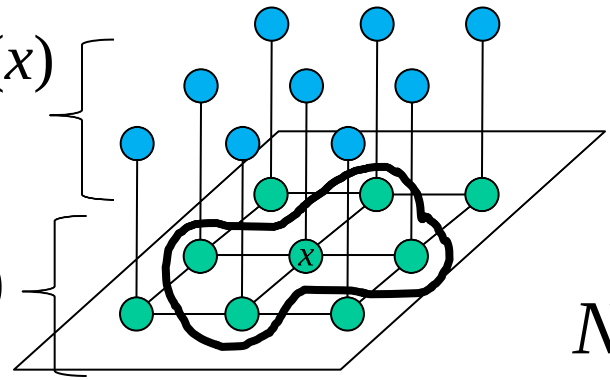
Markov Random Fields (MRF)



Markov Random Fields (MRF)

Image Intensities $I(x)$

Tissue Labels $C(x)$



x Pixel Location

N_x Set of neighbors

Markov Assumption

Label at x independent of all others, given N_x

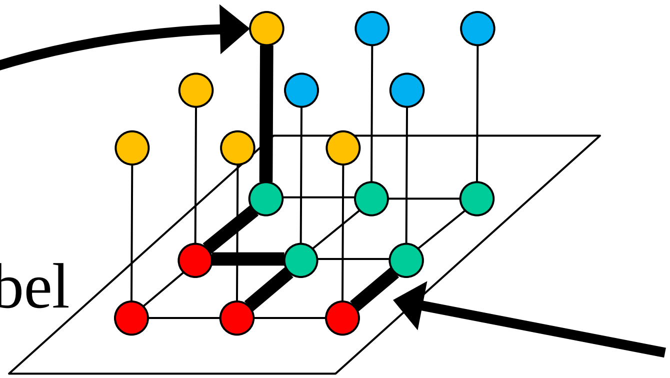
Hammersley–Clifford Theorem

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

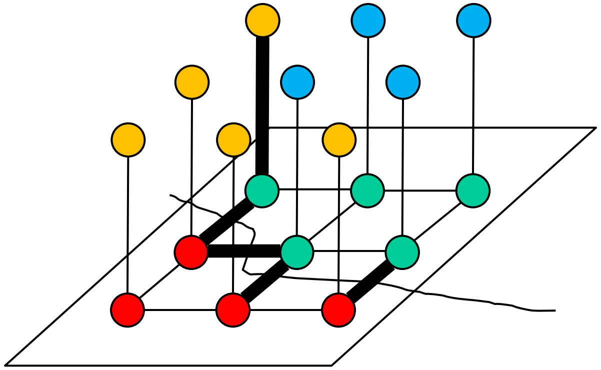
$E_1(x, I, C)$

Penalize inconsistent neighbor labels, e.g. Potts model

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

$$p(C \mid I, X) \propto \exp\left[-\sum_{x \in X} E_1(x, I, C) - \sum_{x' \in N_x} E_2(x, x', C) \right]$$

Markov Random Fields (MRF)



<u>Iterative Algorithms</u>

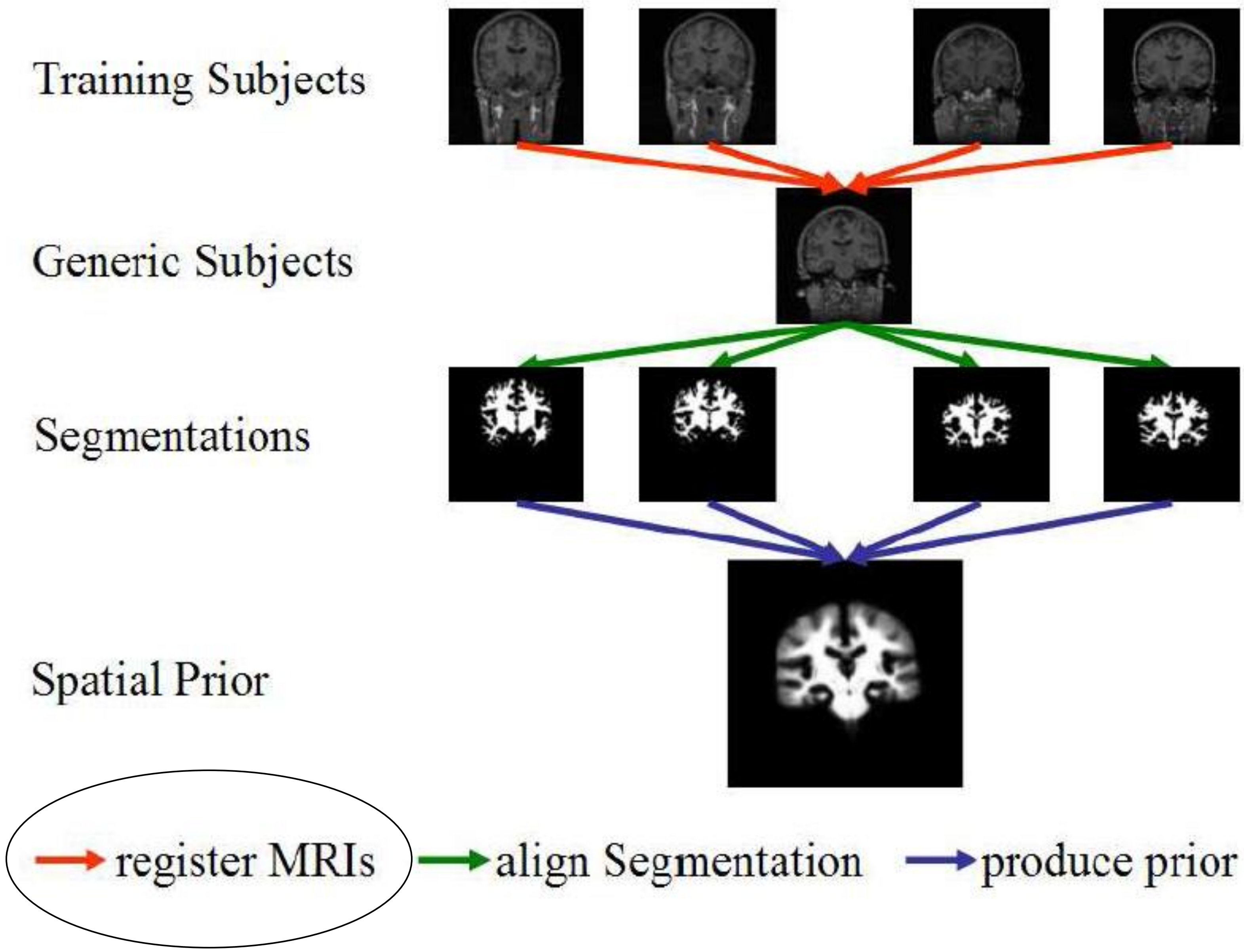
- •Monte-Carlo Markov Chain
  - •Gibbs Sampling
- •Expectation-Maximization

<u>Exact Algorithms</u>

- •Graph Cuts

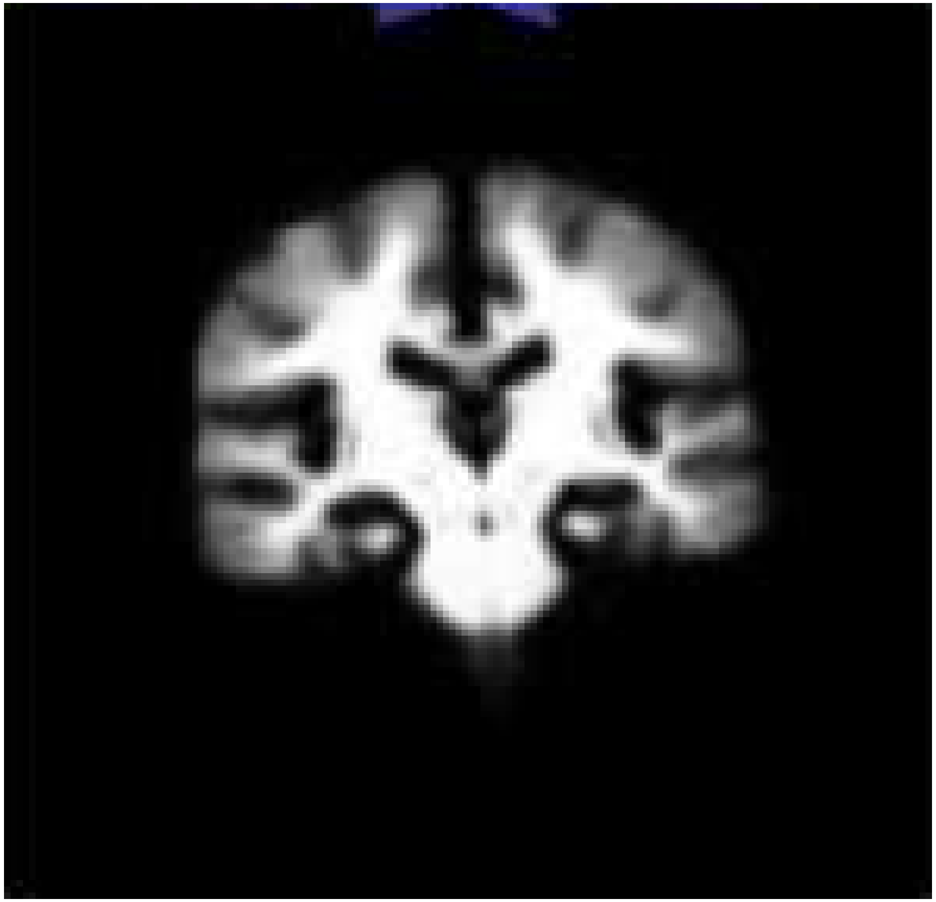
Average Brain Models

- Construct a spatial prior model by averaging tissue distributions over a population [MNI].



Provided by Kilian Pohl

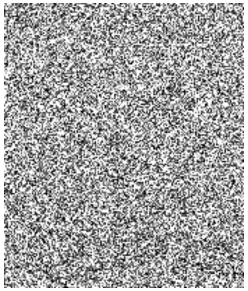
P(white matter | x)



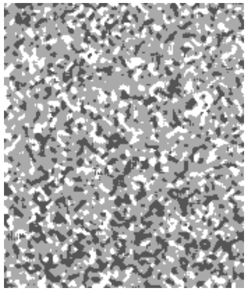
Provided by Kilian Pohl

Generated Sample Images

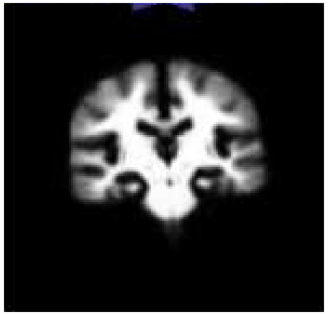
GMM: No spatial component



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:

$$p(I \mid C_i) = p(I \mid C_i, x)$$

Provided by Kilian Pohl

Structurally-Conditioned Prior Models

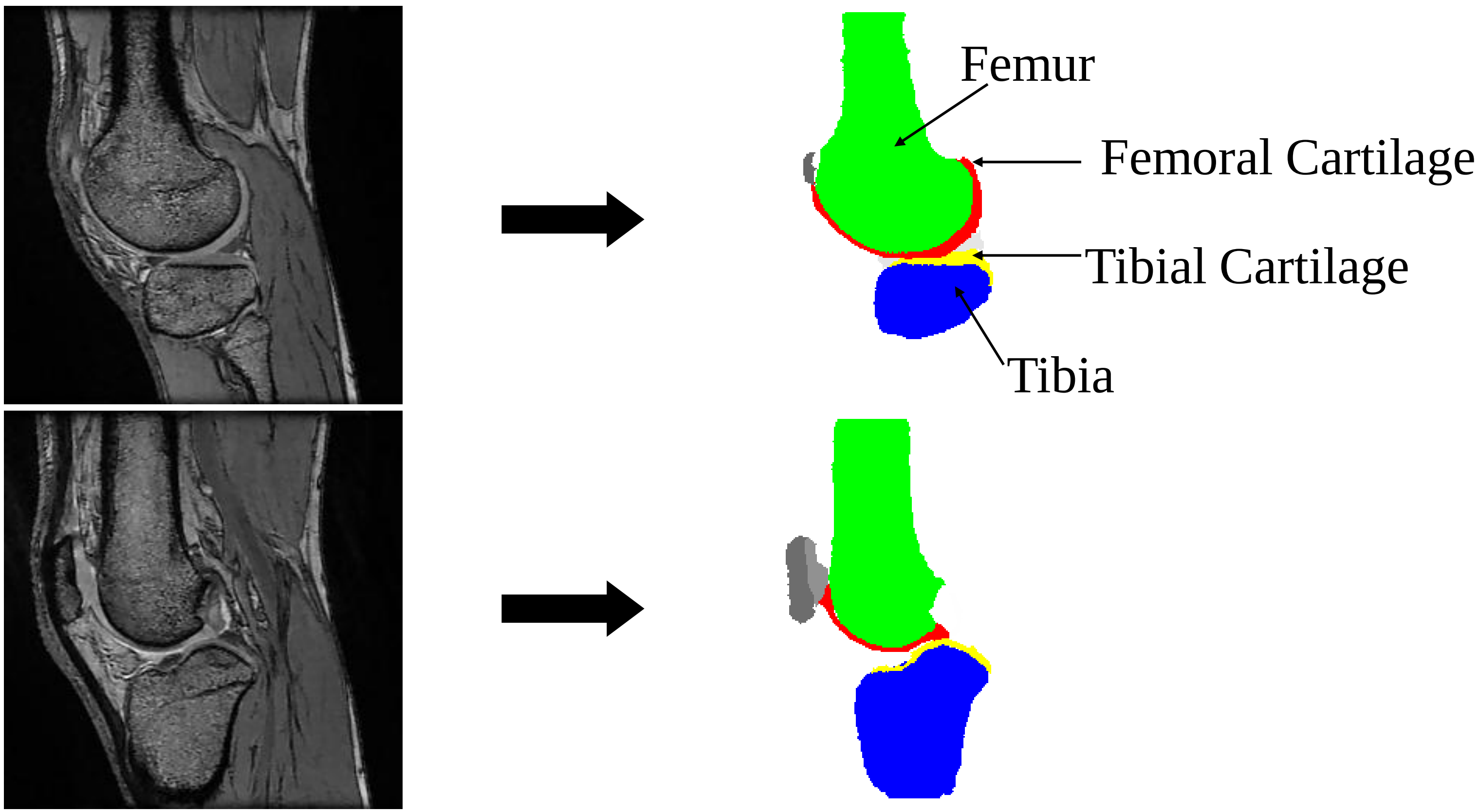
- From (Kapur 1999)
  - Modeling Global Geometric Relationships between Structures

Modeling Global Geometric Relationships between Structures

- Relative Geometry Models
- Motivate Using Knee MRI

MERL, SPL, MIT, CMU
Surgical Simulation
(Sarah Gibson, PI)

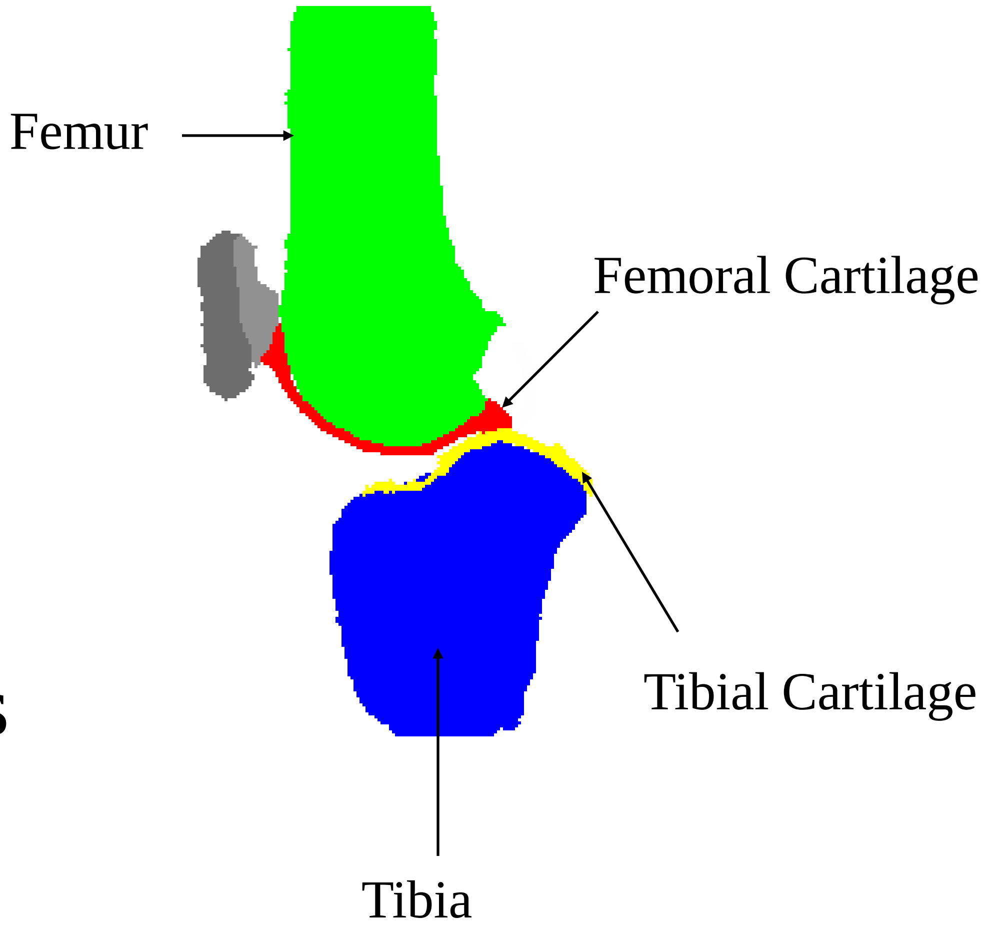
Segmented Knee MRI



Provided by T Kapur

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

Provided by T Kapur

Segment Primary Structures: Femur, Tibia

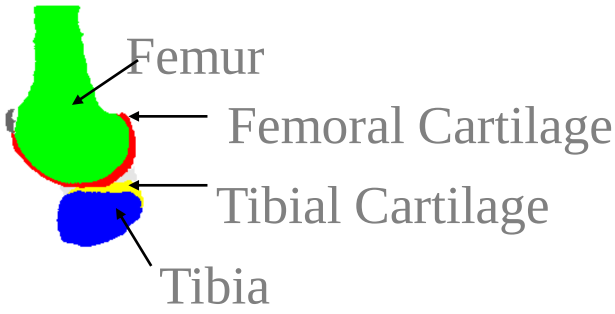
Seed Region Growing Boundary Localization

Provided by T Kapur

Status

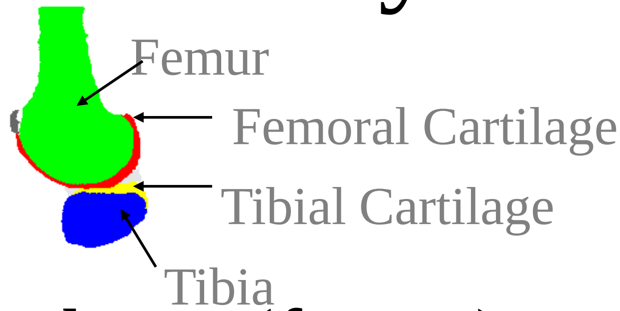
- Have Bone
- Want Cartilage

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



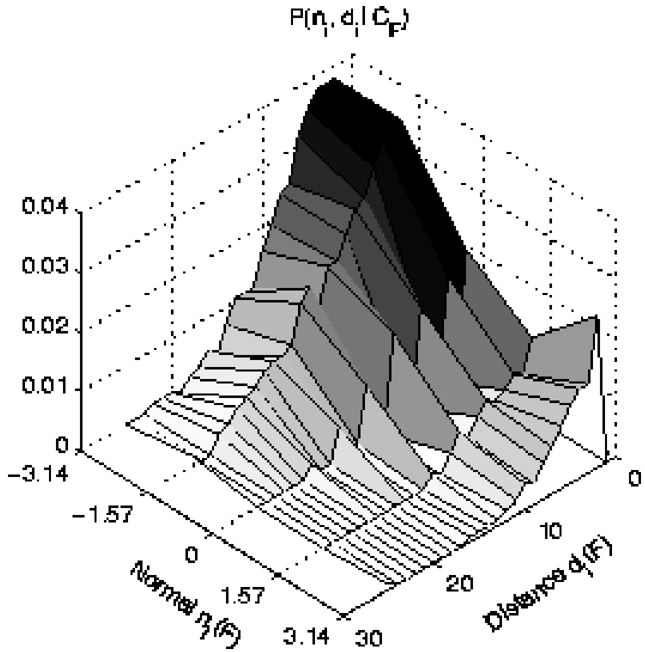
$$\rho_s \equiv \text{distance to closest point on bone (femur)}$$

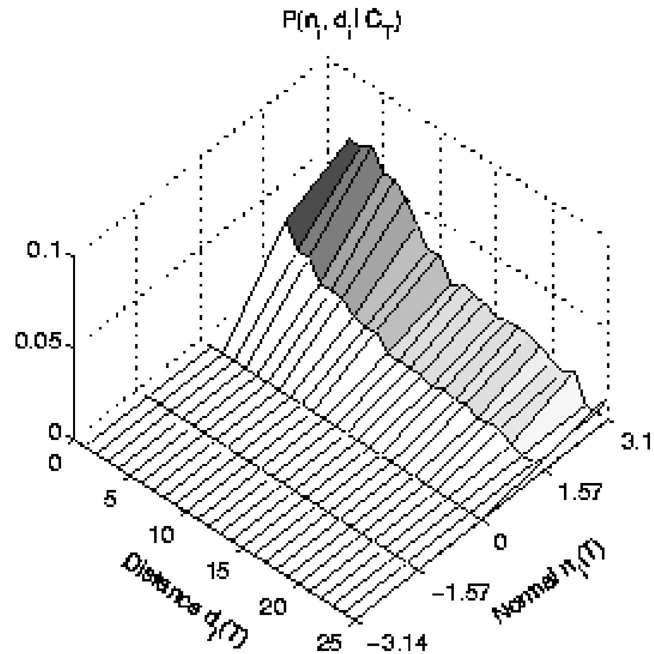
$$n_s \equiv \text{normal to bone (femur) at closest point}$$

$$P(x_s \in \text{Cartilage} \mid \text{Bone}) \cong \frac{P(\rho_s, n_s \mid x_s \in \text{Cartilage}) P(\text{Cartilage})}{Z}$$

Provided by T Kapur

Estimate of $P(\rho_s, n_s \mid x_s \in \text{Cartilage})$



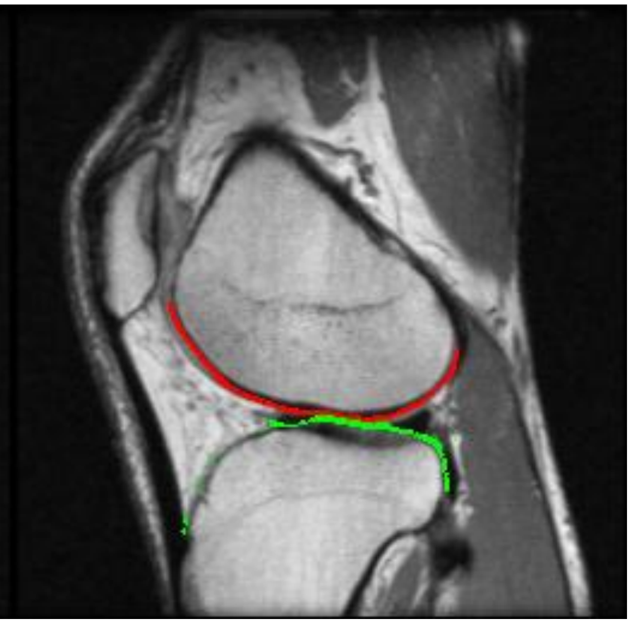


$P(\rho_s, n_s \mid x_s \in \text{Fem. Cartilage})$ $P(\rho_s, n_s \mid x_s \in \text{Tib. Cartilage})$

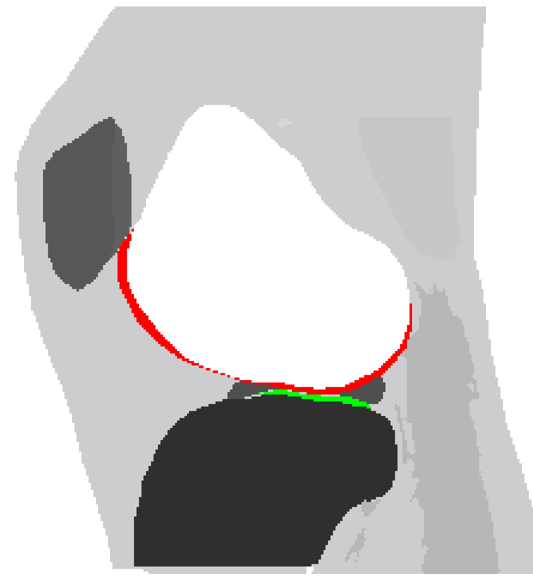
Provided by T Kapur

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

image

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

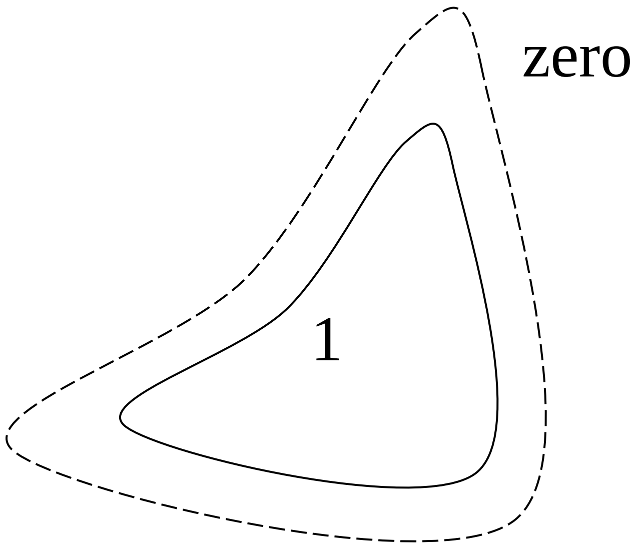
vector addition

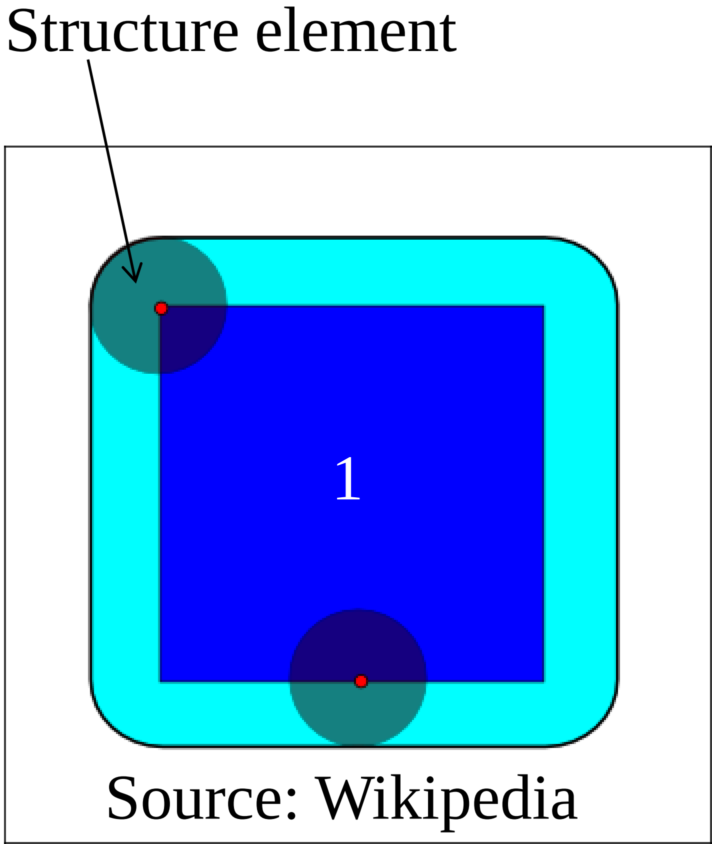
Typical structure elements:

$$\left\{ \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix} \qquad \begin{matrix} & 1 & \\ 1 & 1 & 1 \\ & 1 & \end{matrix} \right.$$

Dilation

- Continuous analogy
- Makes structures *fatter*





Source: Wikipedia

Erosion

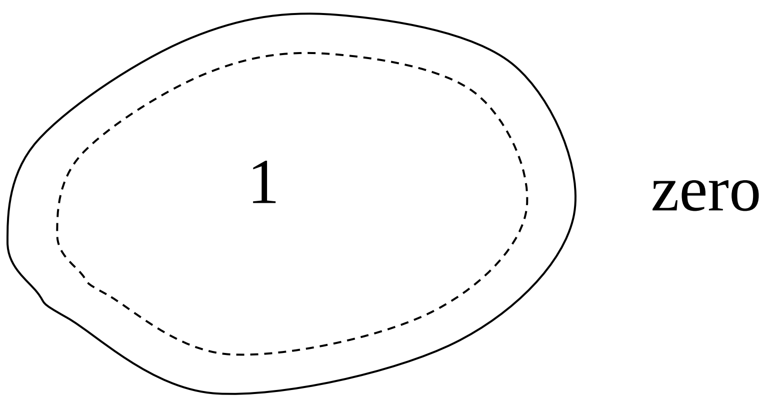
- Erosion is dual of dilation
  - complement A
  - reflect B (negate coordinates)
  - dilate
  - complement result

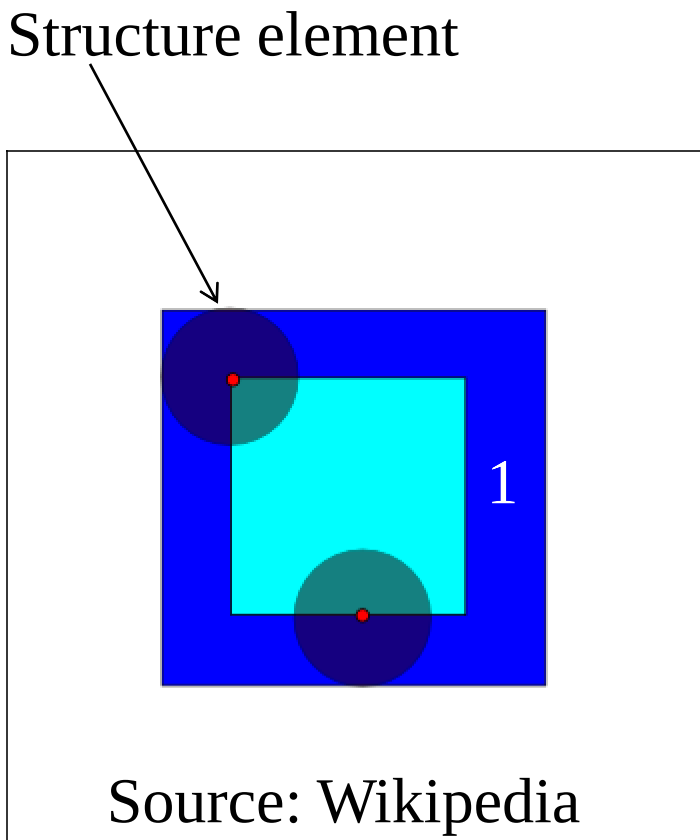
$$\boxed{A \otimes B = \overline{\overline{A} \oplus \hat{B}}}$$

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

Erosion

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

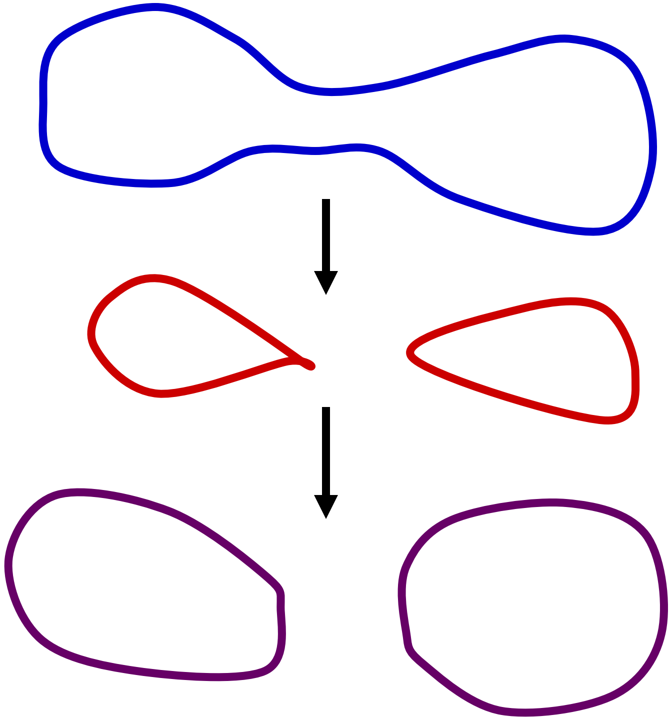




Source: Wikipedia

Opening

- Opening = Erode <u>then</u> Dilate



Break thin connections

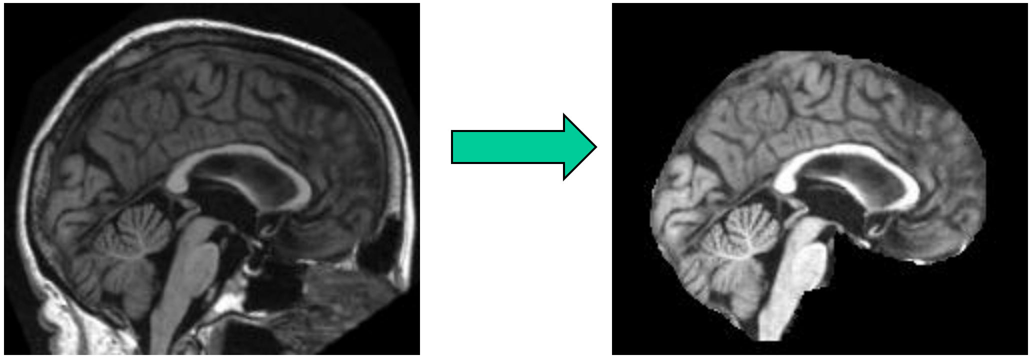
Removes small junk

Closing

- Closing = Dilate then Erode
- Can attach objects that have become fragmented

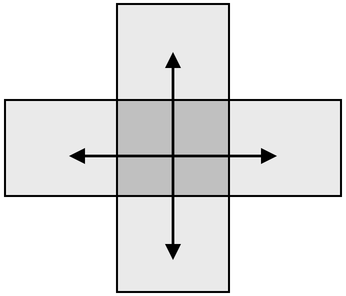
Erosion and Dilation

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

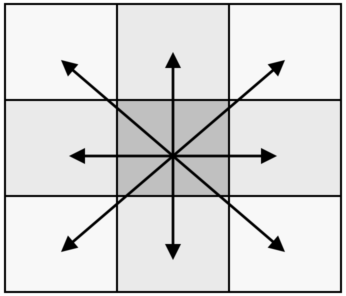


Connectivity

- Define neighbor relation



4-neighbor

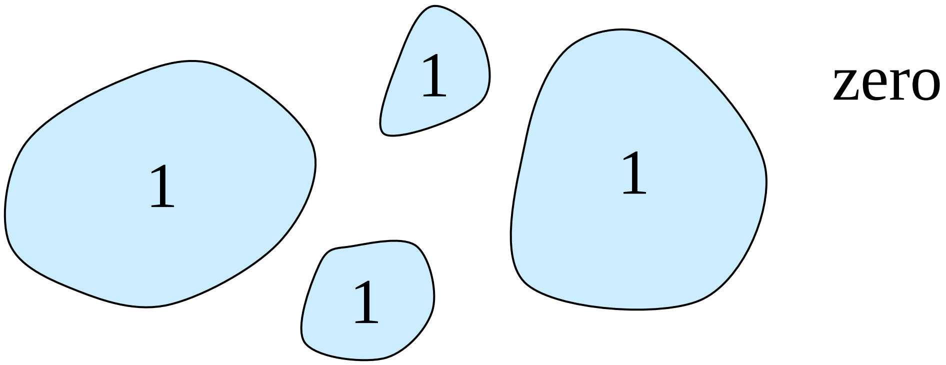


8-neighbor

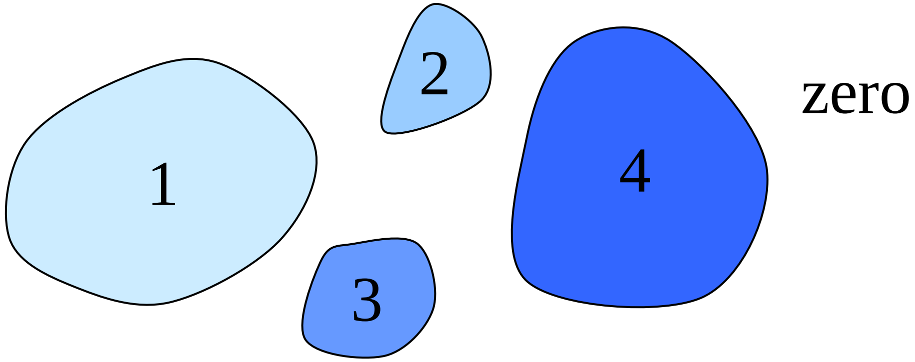
  - There are some inconsistencies that a 6-neighbor relation can fix

Connected Components

- Input: Boolean image objects



- Output: Unique label for each separate object



Finding Connected Components

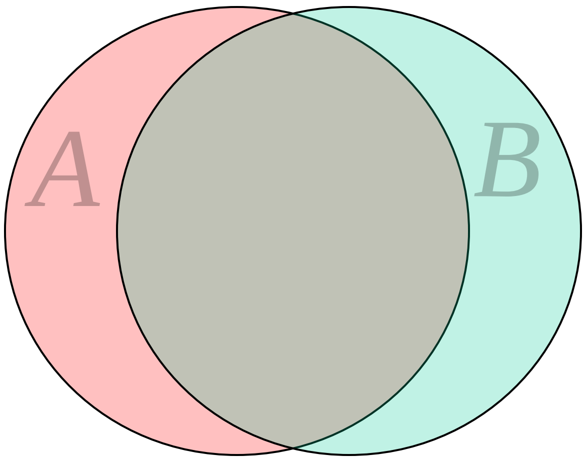
- N = 1 (initialize label)
- Repeat until all pixels are labeled
  - Pick an unmarked *1* pixel
  - Label it, and all of its *1* neighbors, and all of their neighbors: N
  - N ← N + 1

Recursive

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

- **[Duda and Hart1973] Duda, R. and Hart, P.1973. Pattern Classification and Scene Analysis. John Wiley and Sons.**
- **[Gonzales + 2001] R Gonzales and R Woods. Digital Image Processing, 2nd Ed. Prentice Hall 2001.**
- **[Haralick + 1989] R Haralick and S Steinberg. Image Analysis Using Mathematical Morphology. IEEE Transactions PAMI 1989.**
- **[Kapur 1999] T Kapur. Model Based Three Dimensional Medical Imaging Segmentation. PhD Thesis, MIT EECS, 1999.**
- **[Warfield + 2000] S Warfield, J Rexilius, M Kaus, F Jolesz, R Kikinis. Adaptive template moderated spatial varying statistical classification. Med. Image Analysis, 2000.**
- **[Wells + 1996] W Wells, E Grimson, R Kikinis, F Jolesz. Adaptive segmentation of MRI data. IEEE Trans. Med. Img. 15, 1996.**

Further Reading

Kass, Michael, Andrew Witkin, and Demetri Terzopoulos. "Snakes: Active contour models." International journal of computer vision 1.4 (1988): 321-331. Cited by 16117

Otsu, Nobuyuki. "A threshold selection method from gray-level histograms." Automatica 11.285-296 (1975): 23-27. Cited by 14825

Geman, Stuart, and Donald Geman. "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." Pattern Analysis and Machine Intelligence, IEEE Transactions on 6 (1984): 721-741. Cited by 15644

Shi, Jianbo, and Jitendra Malik. "Normalized cuts and image segmentation." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.8 (2000): 888-905. Cited by 8646

Boykov, Yuri, Olga Veksler, and Ramin Zabih. "Fast approximate energy minimization via graph cuts." Pattern Analysis and Machine Intelligence, IEEE Transactions on 23.11 (2001): 1222-1239. Cited by 4449

Zhang, Yongyue, Michael Brady, and Stephen Smith. "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm." Medical Imaging, IEEE Transactions on 20.1 (2001): 45-57. Cited by 2201

Feature-based Analysis

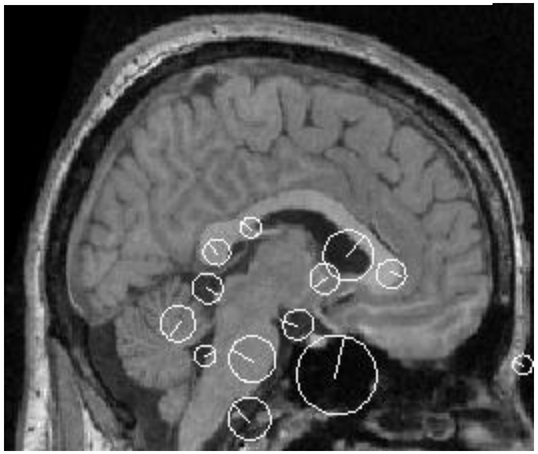
- Model image as a set of image patches
- 3D scale-invariant features (SIFT)

Feature-Based Morphometry: Discovering Group-related Anatomical Patterns.
Toews, Matthew and Wells III, William M., and Collins, D. Louis and Arbel, Tal.
NeuroImage, 2010, Vol 49 (3), 2010, pp. 2318-2327.

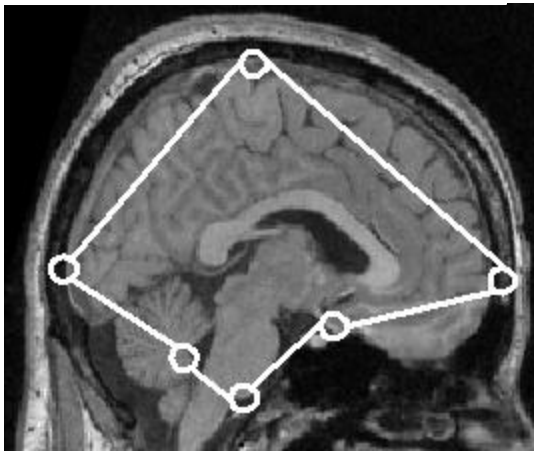
Feature-based Analysis

- Alignment/Registration: fast and robust

Feature-based Alignment
(robust to perturbation)



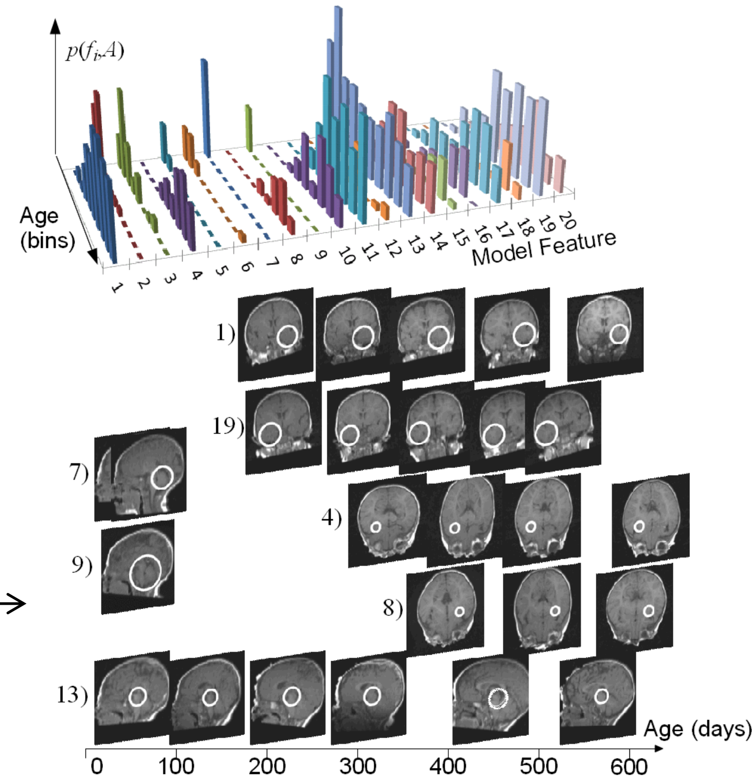
Global Image Alignment
(sensitive to perturbation)



Feature-based Analysis

- Model features over many images.

Infant brain development.



A Feature-based Developmental Model of the Infant Brain in Structural MRI.
M. Toews, W.M. Wells III, Lilla Zöllei.
Medical Image Computing and Computer Assisted Intervention, 2012.