Biomedical Imaging & Analysis

Lecture 7, Part 2. Fall 2014

Image Segmentation (II)

*[**Text**: Ch: 10, Gonzalez and Woods, Digital Image Processing (3rd Edition) + Chapter 5, 8, & 9, Insight into Images, Terry Yoo, PhD]*

Prahlad G Menon, PhD

Assistant Professor

Sun Yat-sen University – Carnegie Mellon University (SYSU-CMU)

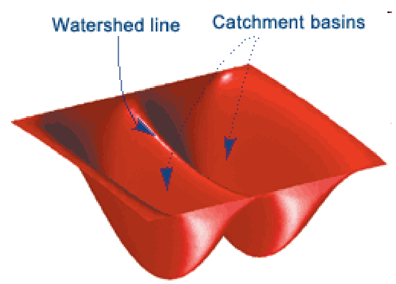
Joint Institute of Engineering

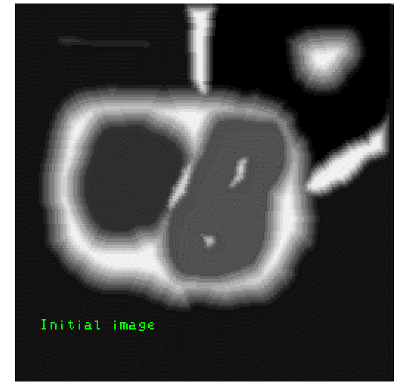
Watershed Segmentation

- Based on image morphology operations…

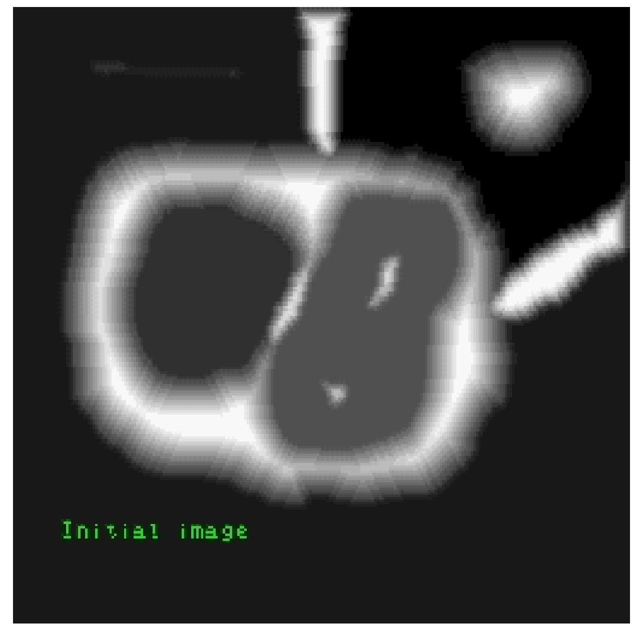
Watershed Segmentation

- Watershed segmentation is based on image morphology operations.





- Basic concepts
  - Watershed lines (divide lines)
  - Catchment basins



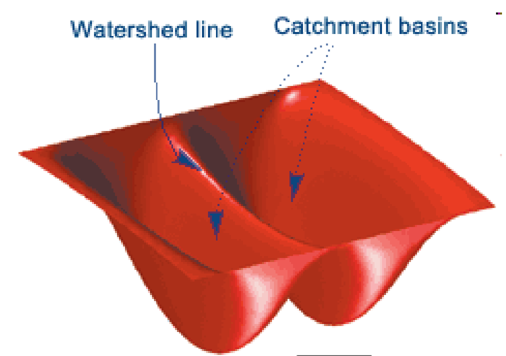
- Classification of points
  - Is a local minimum
  - Is not a local minimum
    - Belonging to a watershed line
    - Belonging to a catchment basin

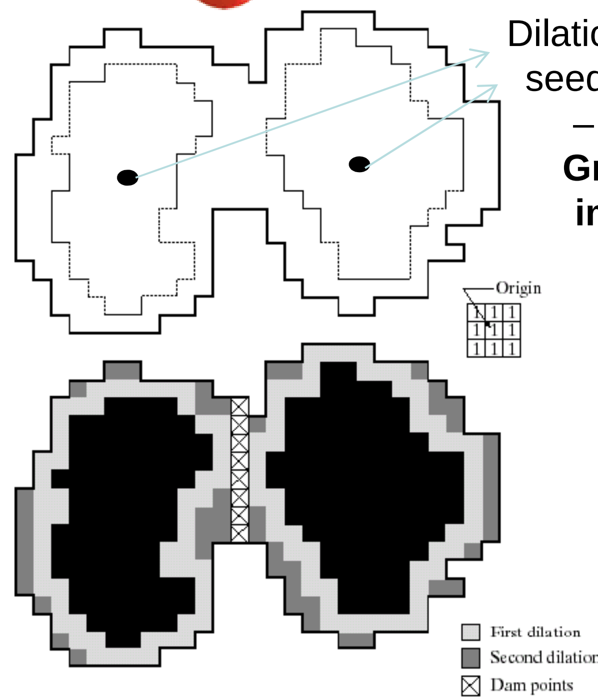
- Goal: to find the watershed lines

http://cmm.ensmp.fr/~beucher/wtshed.html

Watershed Segmentation

- The key step is to build the watershed lines which define the boundaries to be detected.
- Based on dilation starting from seed points (*region growing*), to create catchment basins which grow until they intersect.
- The intensity of each watershed line points is set to be higher than the maximum image intensity, to define a segmentation boundary.



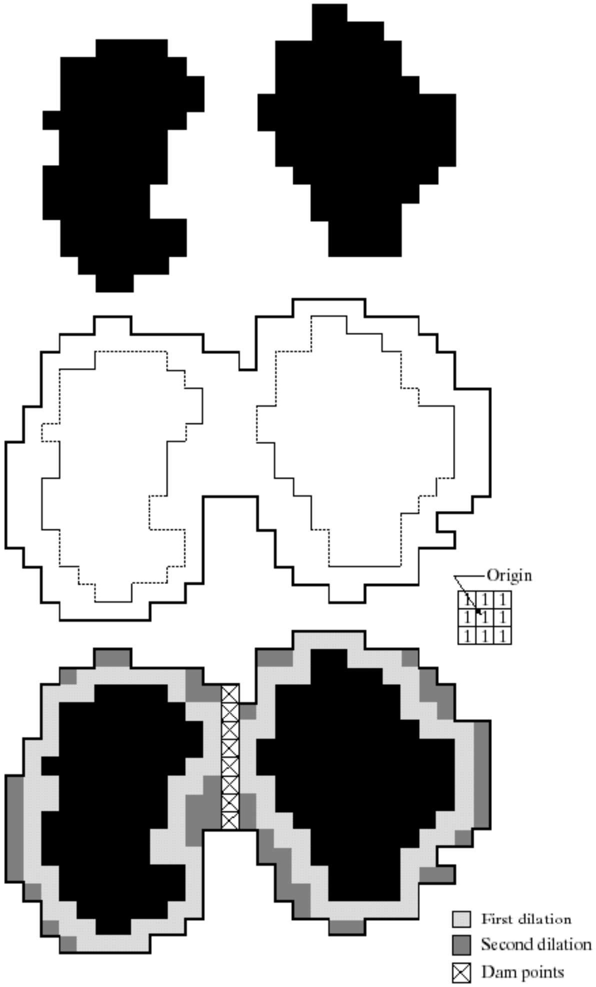


Dilation from seed points – start in **Gradient images!**

http://cmm.ensmp.fr/~beucher/wtshed.html

Watershed Segmentation

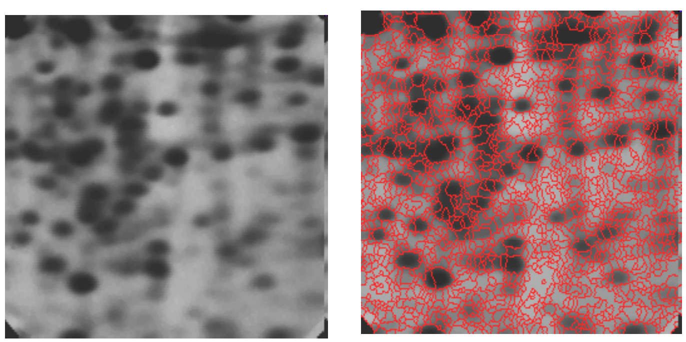
- The key step is to build the watershed lines.
- This is done based on dilation and intersection detection.
- The intensity of each watershed line points is set to be higher than the maximum image intensity.

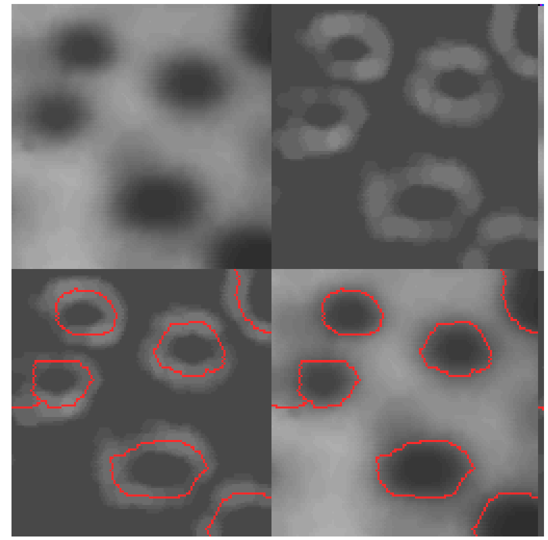


Watershed Segmentation (III)

- Watershed is usually applied to gradient images.
- Image noise often causes oversegmentation because of false local minima caused by noise.

- **One way to avoid too many lines!**
  - A criterion function built based on the relative heights of the walls separating the initial catchment basins to threshold out weak watershed lines.





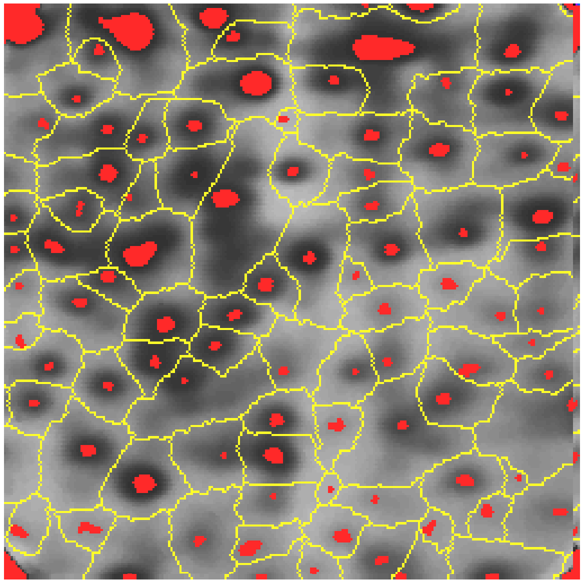
Watershed Segmentation

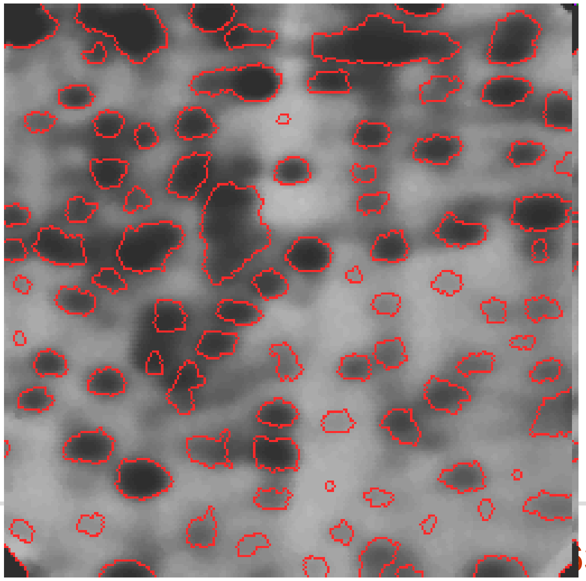
- Oversegmentation can be minimized by using markers.

- *<u>Construction of markers:</u>*

  - Low-pass smoothing of the original image

  - Identify connected local minimum regions as markers





Watershed Segmentation with **Markers**

Matlab implementation / tutorials:
http://www-rohan.sdsu.edu/doc/matlab/toolbox/images/morph16.html

http://www.mathworks.com/help/images/examples/marker-controlled-watershed-segmentation.html

http://cmm.ensmp.fr/~beucher/wtshed.html

Graph-Cut Segmentation

Basic Concepts of **Graphs**

- A graph G = (V, E) is an ordered pair that consists of a set of vertices and a set of edges that connect pairs of distinct vertices.
- A graph is weighted if a number is assigned to each edged.

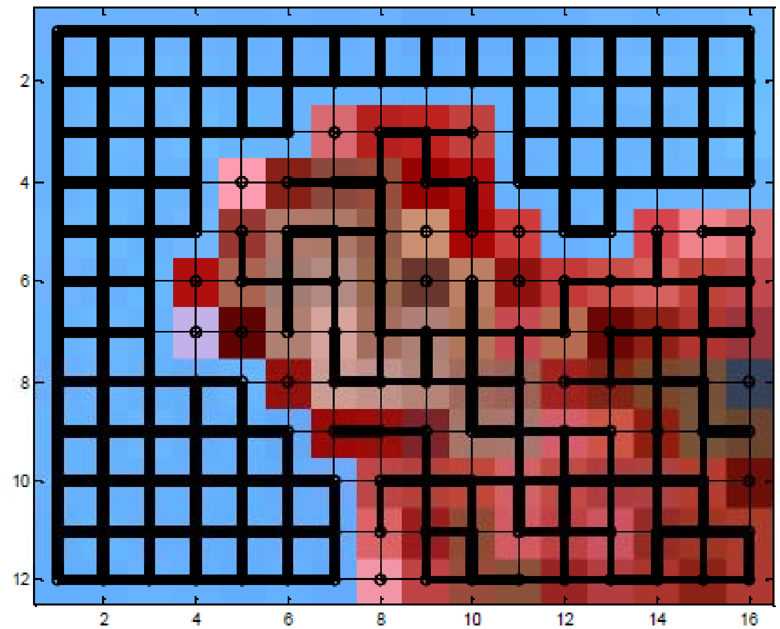
Applied to Segmentation:

- A **bipartite graph** is a graph whose vertices can be divided into two sets such that all edges connect a vertex in one set with a vertex in the other set.

Definition of Image Graph

Graph nodes: pixels

Edge weights: similarity between neighboring pixels



Figures are borrowed from
http://www.cs.yale.edu/homes/spielman/sgta/

$$w_{ij} = e^{-\frac{(I_i - I_j)^2}{\sigma^2}}$$

Matrices encountered

- Adjacency Matrix (*weighted*)

$$W = [w_{ij}]_{n\times n} \qquad w_{ij} = \begin{cases} 0 & i = j \\ e^{-\frac{(I_i - I_j)^2}{\sigma^2}} & i \neq j \end{cases}$$

- Degree Matrix

$$\mathrm{D} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}_{n\times n} \qquad \mathrm{d}_i = \sum_j w_{ij}$$

- Laplacian Matrix

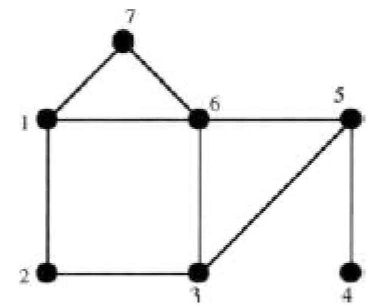
$$L = D - W \qquad l_{ij} = \begin{cases} d_i & i = j \\ -w_{ij} & i \neq j \end{cases}$$

Basic Adjacency Matrix & its Significance

$$a_{ij} = \begin{cases} 1, & \text{vertex } i \text{ is adjacent to vertex } j \\ 0, & \text{otherwise}\,. \end{cases}$$

The adjacency matrix for the undirected graph in Figure 4.2 is

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

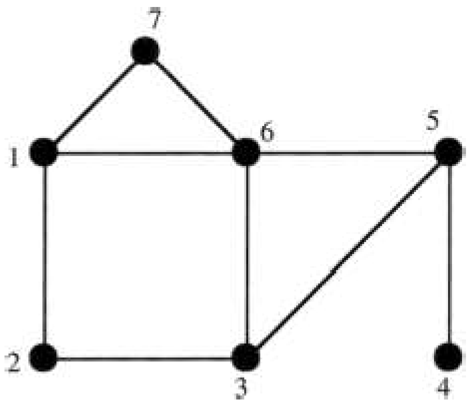


Applying the power method for symmetric matrices to A produces the estimates

$$\lambda_1 \approx 2.86081$$

$$\mathbf{v}_1 \approx \begin{bmatrix} 0.406691 & 0.290865 & 0.425420 & 0.134554 \\ 0.384933 & 0.541244 & 0.331352 \end{bmatrix}^T.$$

The Adjacency Matrix & its Significance



$$\mathbf{v}_1 \approx \begin{bmatrix} 0.406691 & 0.290865 & 0.425420 & 0.134554 \\ 0.384933 & 0.541244 & 0.331352 \end{bmatrix}^T.$$

Next, suppose the vertices in an undirected graph represent cities and an edge represents the existence of a direct traveling route between two cities. Geographers have shown that the entries in an eigenvector associated with the dominant eigenvalue of the adjacency matrix provide a measure of the accessibility of the cities (Straffin [12]). Thus, since the sixth entry in $\mathbf{v}_1$ is the largest, vertex 6 represents the most accessible city. Further, since the first, third, and fifth entries of $\mathbf{v}_1$ are roughly equal, the cities represented by vertices 1, 3, and 5 are nearly equal in terms of accessibility. Finally, as might have been expected, the city represented by vertex 4 is the least accessible.

Algorithm

- Given an image, set up a weighted graph $G = (V, E)$ and set the weight on the edge connecting two nodes to be a measure of the similarity between the two nodes.
- Solve $(D - W)x = \lambda Dx$ for the eigenvectors with the second smallest eigenvalue.
- Use the second smallest eigenvector to bipartition the graph.
- Decide if the current partition should be subdivided and recursively repartition the segmented parts if necessary.

Pair-wise similarity matrix W

Laplacian matrix D-W

Degree matrix D: $D(i,i) = \sum_j W_{i,j}$

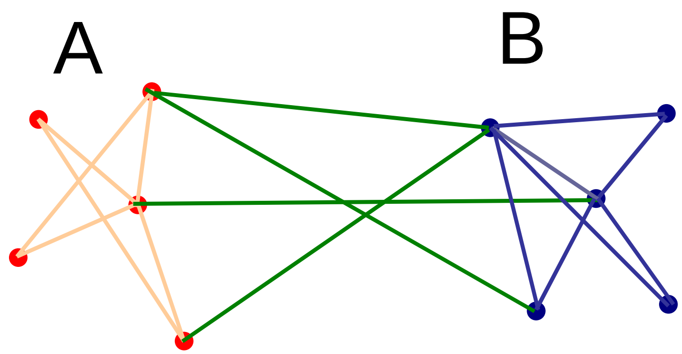
Basic Concept of **Graph Cuts** (I)

- A graph G = (V, E) can be partitioned into two disjoint sets A, B

$$A \bigcup B = V \qquad A \bigcap B = \varnothing$$

- Each vertex represents a <u>pixel</u> within the image.

- The weight of the edge connecting two vertices represents their <u>similarity</u>.

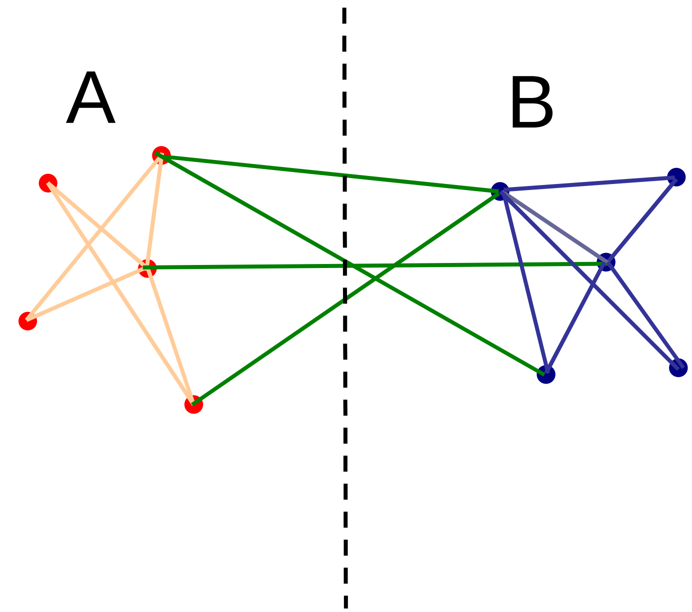


Shi & Malik, PAMI, 22:888-905, 2000

Basic Concept of **Graph Cuts** (II)

- A graph cut is defined as

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v)$$



- The goal is to find a partition that minimizes the cut.

- But there is a catch:
minimum cut favors small sets of isolated nodes.

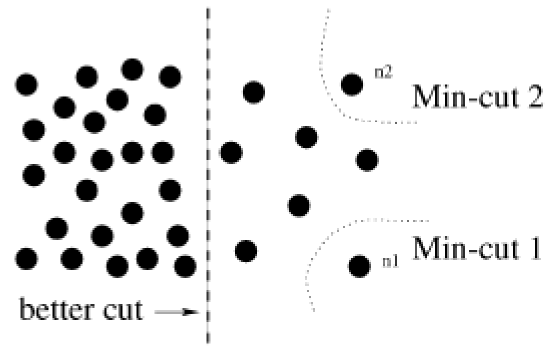


Fig. 1. A case where minimum cut gives a bad partition.

Formulation of Normalized Cut (I)

- Definition of normalized cut

$$Ncut(A,B)=\frac{cut(A,B)}{assoc(A,V)}+\frac{cut(A,B)}{assoc(B,V)}$$

$$\text{where } assoc(A,V)=\sum_{u\in A,t\in V} w(u,t)=assoc(A,A)+cut(A,B)$$

$$Ncut(A,B)=\frac{cut(A,B)}{assoc(A,A)+cut(A,B)}+\frac{cut(A,B)}{assoc(B,B)+cut(A,B)}$$

Formulation of Normalized Cut (II)

- Definition of normalized association

$$Nassoc(A,B) = \frac{assoc(A,A)}{assoc(A,V)} + \frac{assoc(B,B)}{assoc(B,V)}$$

$$= \frac{assoc(A,V) - cut(A,B)}{assoc(A,V)} + \frac{assoc(B,V) - cut(A,B)}{assoc(B,V)}$$

$$= 2 - \left(\frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \right)$$

$$\boxed{Nassoc(A,B) = 2 - Ncut(A,B)}$$

- Minimization of normalized cut is equivalent to maximization of normalized association.

Solution of Normalized Cut (I)

- Matrix formulation

$$Ncut(A,B) = \frac{\sum_{(x_i>0,x_j<0)} -w_{ij}x_i x_j}{\sum_{x_i>0} d_i} + \frac{\sum_{(x_i<0,x_j>0)} -w_{ij}x_i x_j}{\sum_{x_i<0} d_i}$$

$$d(i) = \sum_j w(i,j)$$

- Reformulate the problem into a matrix form

$$D = \begin{bmatrix} d_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & d_{N-1} & 0 \\ 0 & 0 & 0 & 0 & d_N \end{bmatrix} \text{ where } d_i = \sum_j w(i,j) \quad W(i,j) = w(i,j)$$

Solution of Normalized Cut (II)

- Solution of the normalized cut can be transformed into the minimization of the Rayleigh quotient.

$$min_x \ Ncut(x) = min_y \frac{y^T (D-W) y}{y^T Dy}$$

$$\text{where} \quad y = (1+x) - b(1-x), \qquad b = \frac{k}{1-k}, \quad k = \frac{\sum_{x_i > 0} d_i}{\sum d_i}, \quad d_i = \sum_j w(i,j)$$

$$s.t. \quad y_i \in \{1, -b\} \qquad y^T D1 = 0$$

- Exact solution of normalized cut is NP-complete.

Solution of Normalized Cut (III)

- A key relaxation is to allow y to take on continuous real values. Now y can be determined as the solution of the following generalized eigenvalue problem.

$$(D-W)\,y = \lambda D y$$

- After a further transformation, y is the solution of the following equation

$$D^{-\frac{1}{2}}(D-W)D^{-\frac{1}{2}}z = \lambda z \text{ where } y = D^{-\frac{1}{2}}z$$

$$\min_x Ncut(x) = \min_y \frac{y^T(D-W)y}{y^T D y}$$

$$\Updownarrow$$

$$(D-W)\,y = \lambda D y$$

$$\Updownarrow$$

$$D^{\frac{1}{2}}(D-W)D^{-\frac{1}{2}}z = \lambda z \qquad z = D^{\frac{1}{2}}y$$

Summary of the Solution Procedure

- Step 0: Compute *D, W*

- Step 1: Solve the eigenvector of the following equation

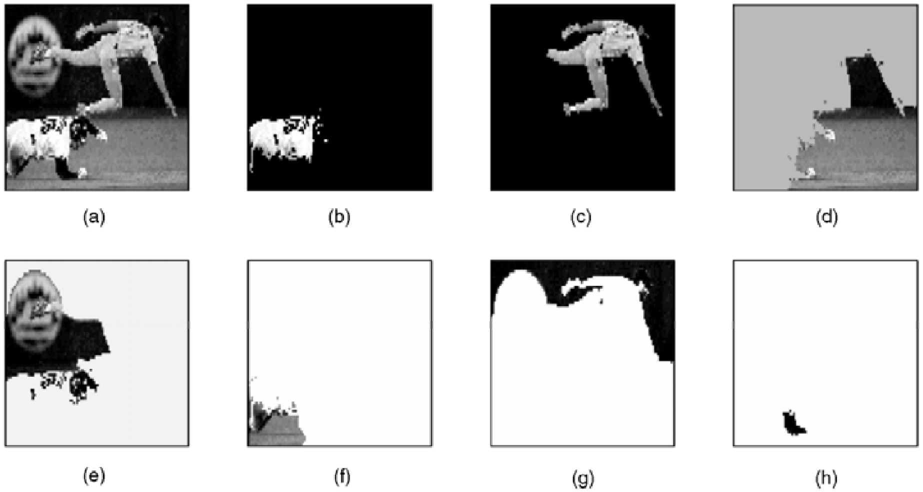
$$D^{-\frac{1}{2}}\left(D-W\right)D^{-\frac{1}{2}}z=\lambda z$$

- Step 2: Take the eigenvector corresponding to the second smallest eigenvalue and calculated

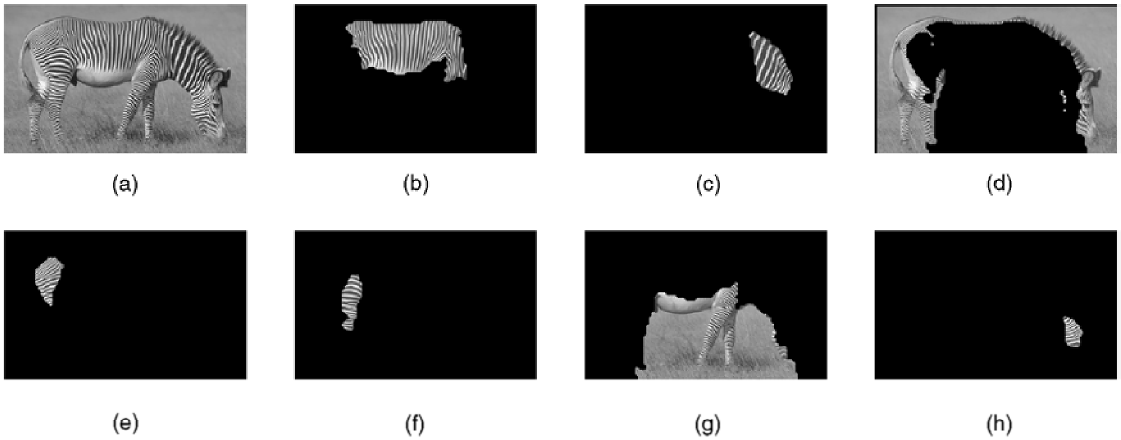
$$y=D^{-\frac{1}{2}}z$$

- Step 3: Partition y
  - By taking zero or the median as the splitting point
  - Search for the splitting point that minimizes *Ncut*

Results



(a) (b) (c) (d)
(e) (f) (g) (h)

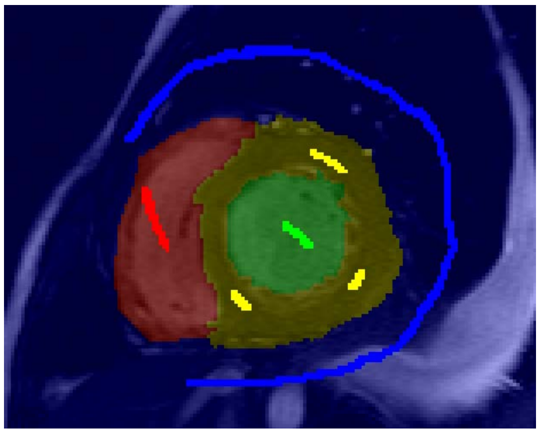


(a) (b) (c) (d)
(e) (f) (g) (h)

Random-Walk Segmentation

Random Walker - Concept

Given labeled voxels, for each voxel ask: What is the probability that a random walker starting from this voxel first reaches each set of labels?



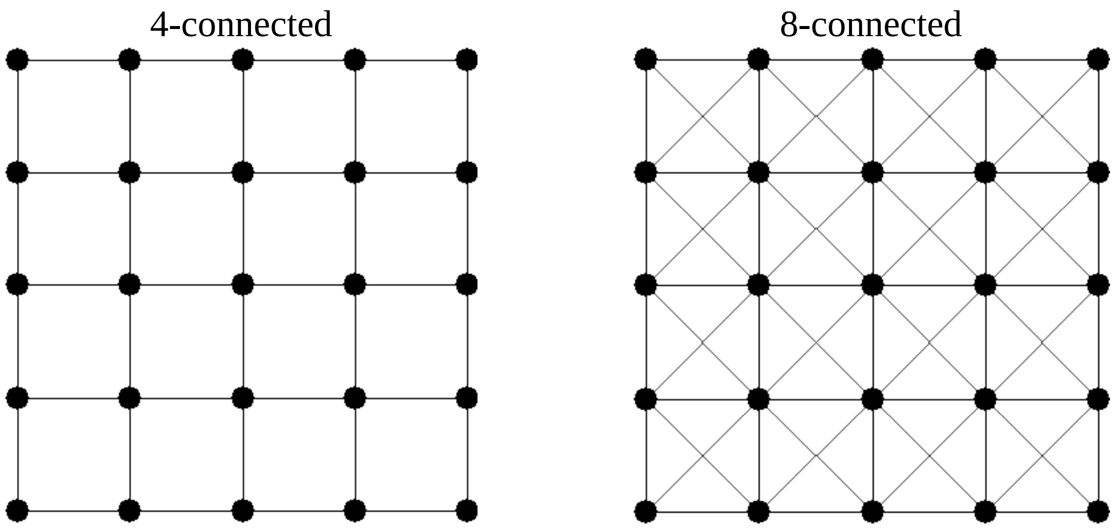
Algorithm summary:

1. Generate weights based on image intensities
2. Build Laplacian matrix
3. Solve system of equations for each label
4. Assign pixel (voxel) to label for which it has the highest probability

Random Walker - Theory

Parameters:

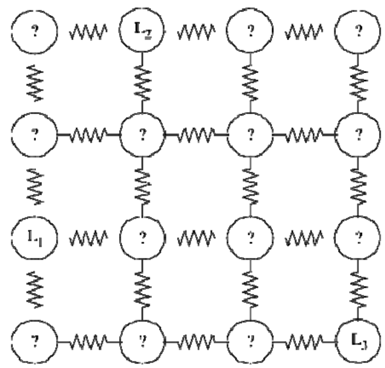
Only one parameter – In the function that maps intensity gradients to random walker biases/edge weights



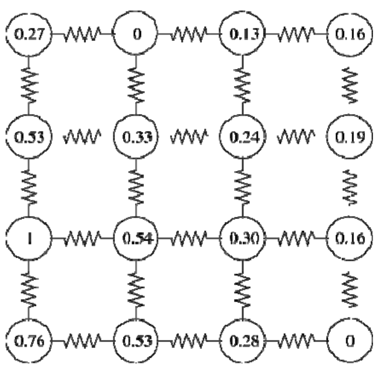
Random Walker - Theory

Situation exactly analogous to DC circuit in steady-state

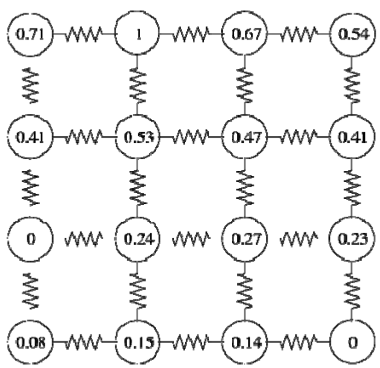
Initial labeling



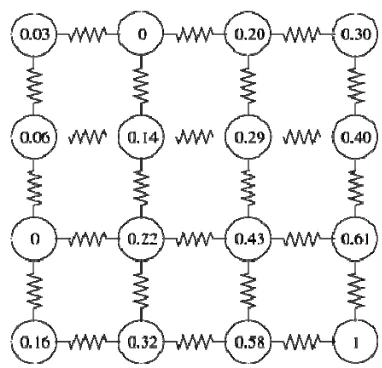
Label 1 prob.



Label 2 prob.



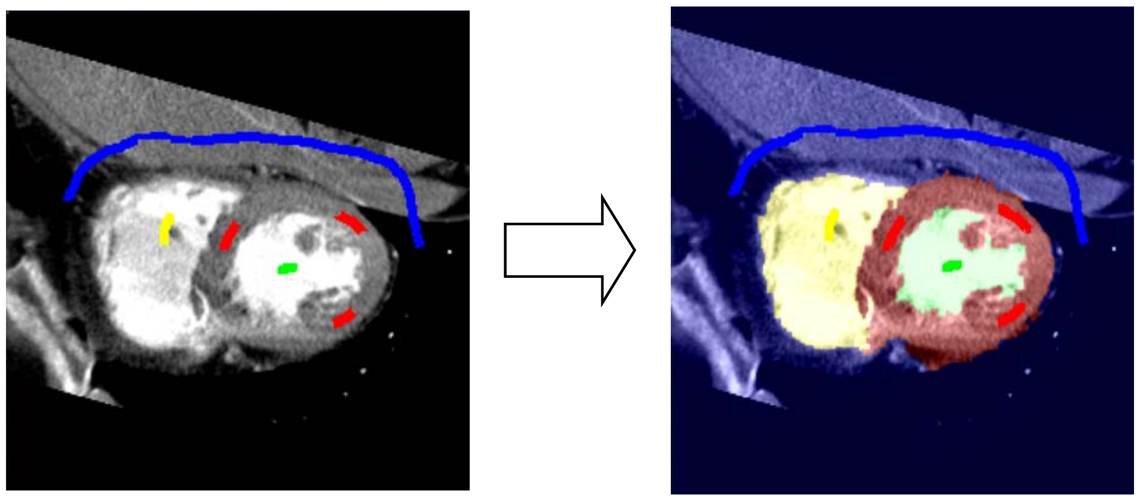
Label 3 prob.



Labels – Unit voltage sources or grounds
Weights – Branch conductances
Probabilities – Steady-state potentials

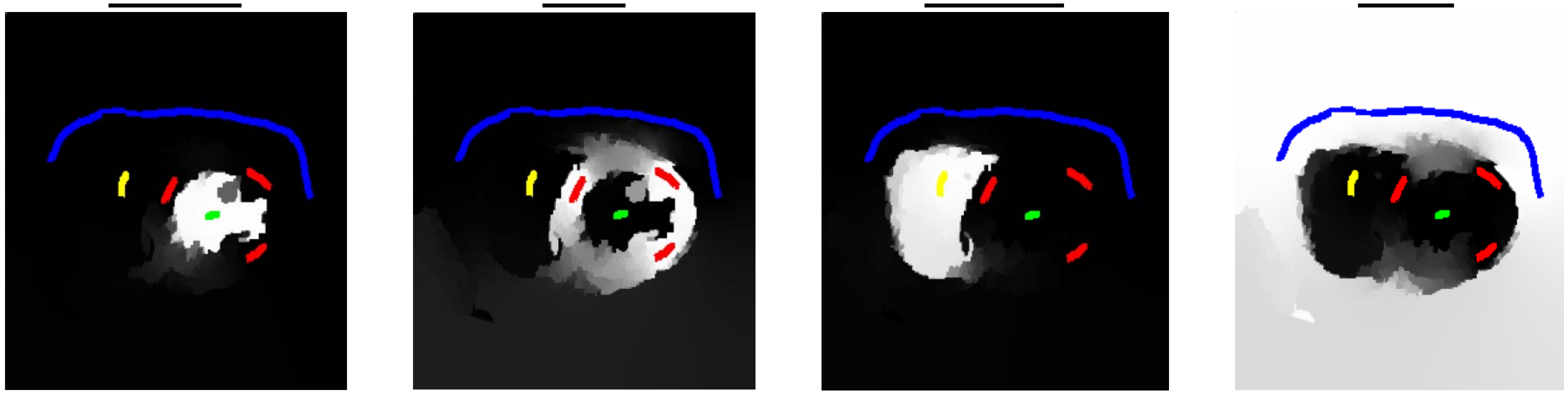
Random Walker – Concept (solved using circuit analogy)

Partially labeled image | Segmented image

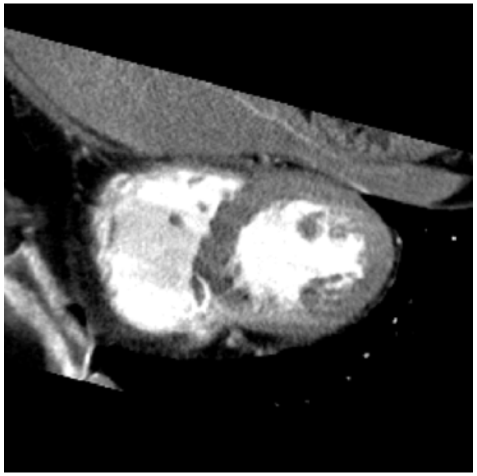


Probabilities

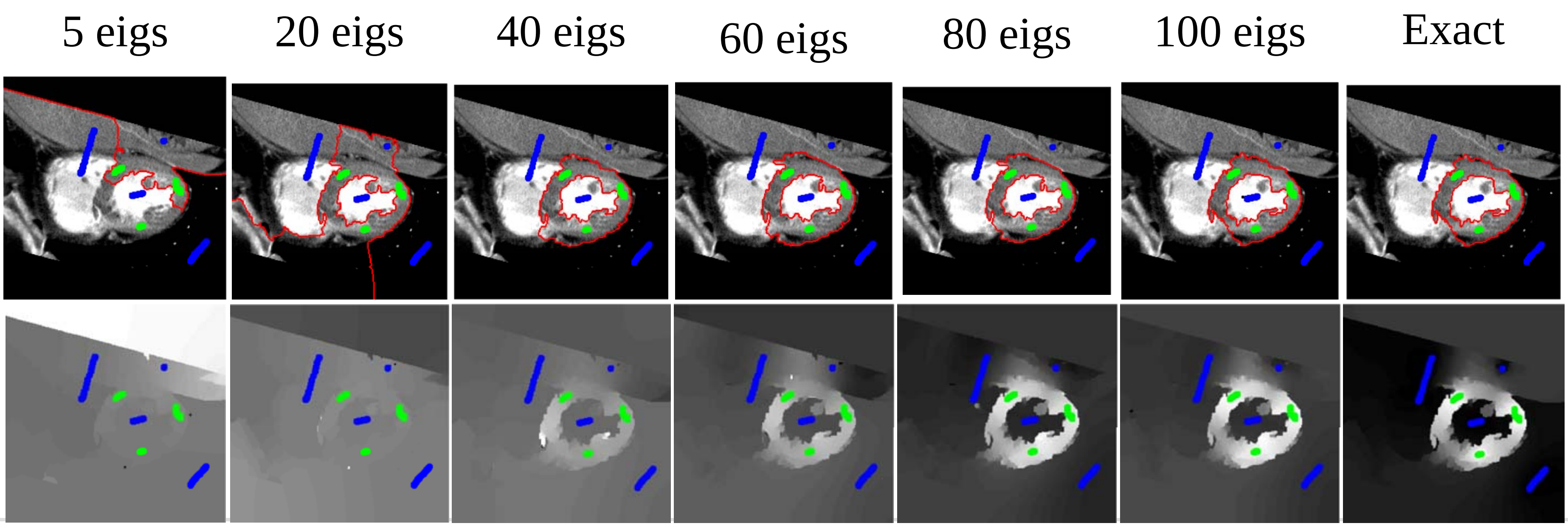
Green | Red | Yellow | Blue



Random Walker – Updated solution strategy since the 2006 version.

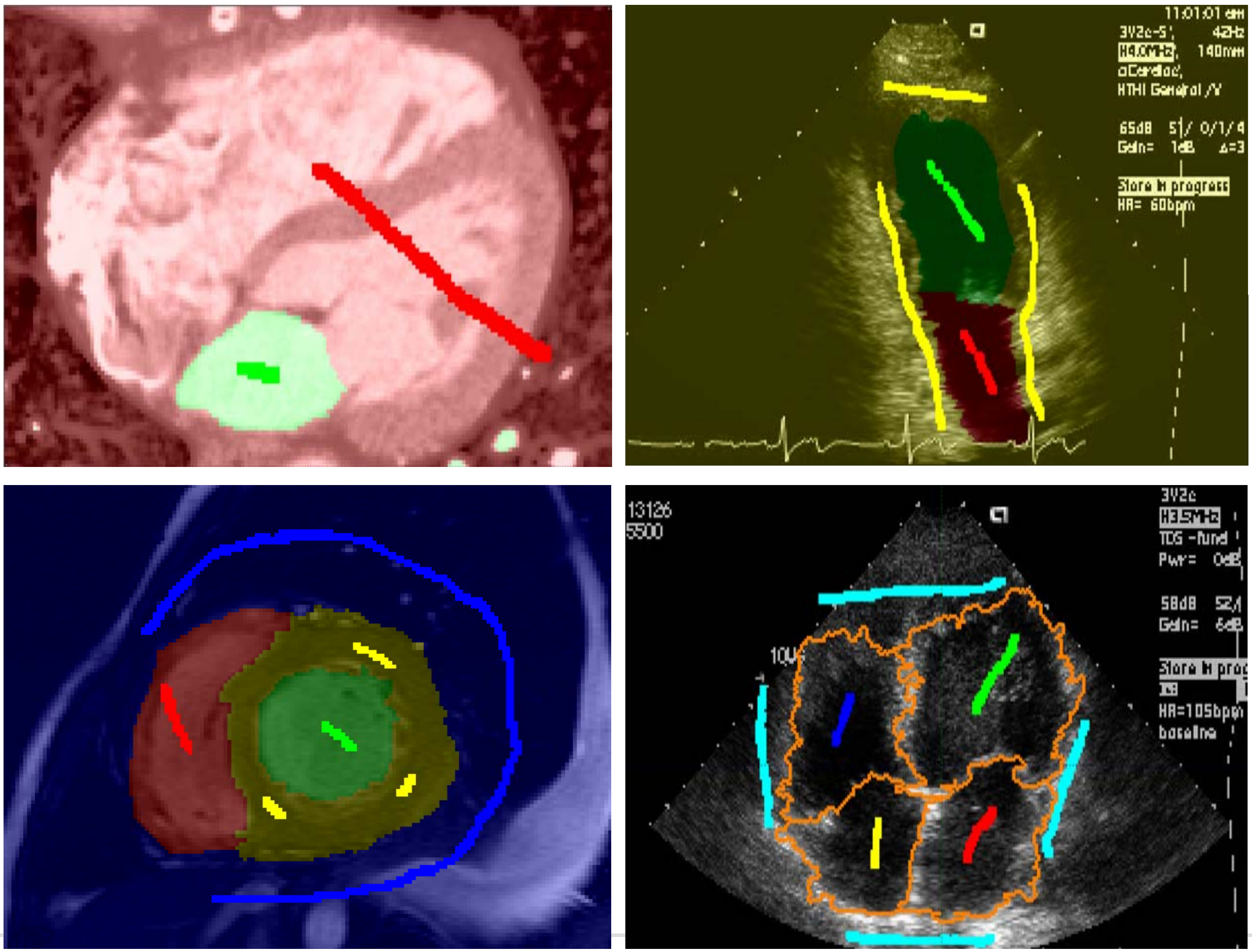


- Precompute eigenvectors of Laplacian
- Input seeds
- Instant result (approximation)



Random Walker - Results

Cardiac segmentation across modalities



Conclusion – More Information

Writings and code

Webpage:
http://cns.bu.edu/~lgrady

Random walkers paper:
http://cns.bu.edu/~lgrady/grady2006random.pdf

Random walkers MATLAB code:
http://cns.bu.edu/~lgrady/random_walker_matlab_code.zip

Random walker demo page:
http://cns.bu.edu/~lgrady/Random_Walker_Image_Segmentation.html

MATLAB toolbox for graph theoretic image processing at:
http://eslab.bu.edu/software/graphanalysis/

CVPR Short Course: Fundamentals linking discrete and continuous approaches to computer vision - A topological view
http://cns.bu.edu/~lgrady/Short_Course.html

Comparison of Graph-based Segmentation Methods

	Random Walk	Graph Cut	Normalized Cut
Objective function	$X^T LX$	$\sum_{(i,j)\in E} w_{ij} \mid x_i - x_j \mid$	$z^T D^{-1/2} LD^{-1/2} z$
Solution. (k=2)	Linear system	Maximal Flow	Eigenvector of the normalized Laplace
Solution. (k>2)	Linear system	Alpha-expansion (approximated)	Spectral clustering
Uniqueness	Unique	Not unique	Unique
Result bias	Shift related to seed placement	Small-cut bias	No bias

Image Segmentation Performance Evaluation

Evaluation Strategies

- Option 1: using manual segmentations as references.
- Option 2: relative rating by experts.
- Option 3: rating by consensus (Survey / Vote).
  - Consensus may not be accurate.

Segmentation Evaluation Criterion

- Different methods of comparison -

  - **Volumetric** overlap error

    Relative volume difference

  - **Surface** comparison error – distance metrics between shapes

    Average symmetric surface distance
    Root mean square surface distance
    Maximum symmetric surface distance

- **Popular scoring indices** -
  - **Dice** index
  - **Jaccard** Index
  - **P-values**: Comparing meshes in two populations
    - T-test : compute p-values on triangular meshes that is topologically equivalent to a sphere or ellipsoid (SPHARM method).
    - Useful code at: http://www.stat.wisc.edu/~mchung

Dice and **Jaccard** indices

- **In Matlab:**

$$Dice\ index = \frac{2|A \cap B|}{|A| + |B|}$$

(a)

$$Jaccard\ index = \frac{|A \cap B|}{|A \cup B|}$$

(b)

```
function [Jaccard,Dice,rfp,rfn]=sevaluate(m,o)
% gets label matrix for one tissue in segmented and ground truth
% and returns the similarity indices

m=m(:);                        % Gold standard segmentation / ground truth
o=o(:);                        % Segmentation to be tested for accuracy
common=sum(m & o);
union=sum(m | o);
cm=sum(m);                     % the number of voxels in m
co=sum(o);                     % the number of voxels in o

Jaccard=common/union;
Dice=(2*common)/(cm+co);
rfp=(co-common)/cm;            % false positive ratio
rfn=(cm-common)/cm;            % false negative ratio
```

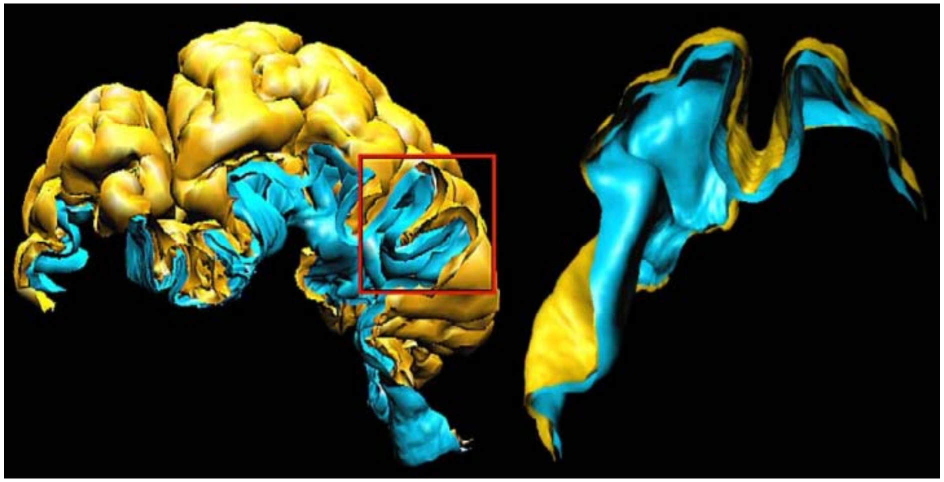
- **In ITK**, itk::MeshSimilarityCalculator filter is used to calculate the similarity of two labeled meshes: http://www.kitware.com/source/home/post/18

Comparing Segmentation Meshes

http://www.stat.wisc.edu/~mchung/papers/ni_heatkernel.pdf

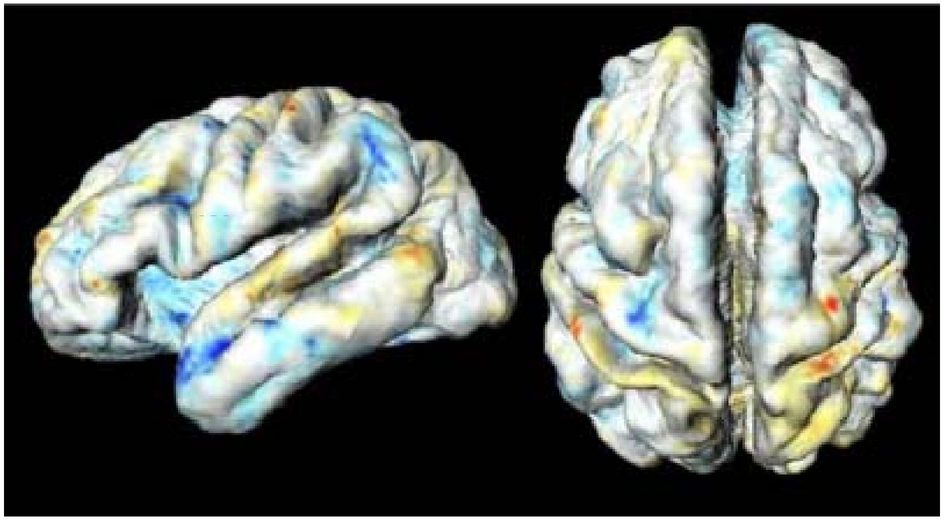
Type 1: Comparing two segmentations to get **cortical manifold regional thickness colormap comparison** based on distance metrics, spectral methods or spherical harmonic (SPHARM) decompositions, etc. i.e. methods of shape comparison.

SPHARM for example results in a mesh which starts an **ellipsoid located outside the brain and is shrunk** to obtain the inner and outer cortical surface with the same number of vertices and triangles in each segmented structure. Meshes can then be compared by **Euclidean distance**!



Type 2: Comparing cortical thickness colormaps in two groups of patients viz. autistic subjects and normal controls. For the *ith* group, let *ni* denote the number of subjects and *hi* denote the population mean thickness. Under stochastic model (1), we are interested in **testing if the thickness for the two groups are identical.**

P value maps projected onto the average normal population segmentation (Right).



M.K. Chung et al. / NeuroImage 25 (2005) 1256–1265

Reference

Heimann et al, *Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets, IEEE TRANSACTIONS ON MEDICAL IMAGING*, VOL. 28, NO. 8, pp. 1251-1265. 2009

Medical Imaging Computing and Computer-Assisted Intervention (MICCAI) http://mbi.dkfz-heidelberg.de/grand-challenge2007/index.html

- Liver segmentation: http://www.sliver07.org

- Caudate segmentation: http://www.cause07.org/

Introduction to ITK

- ITK: image segmentation and registration toolkit

- Started in 1999 through funding by the National Library of Medicine to support the Visible Human Project.

- Website: http://www.itk.org/

- ITK: insight toolkit

  - Open source software package for image registration and segmentation

Introduction to ITK (II)

- Language: 55% C++; 25% C; XML 11%; Other 9%

- Scale

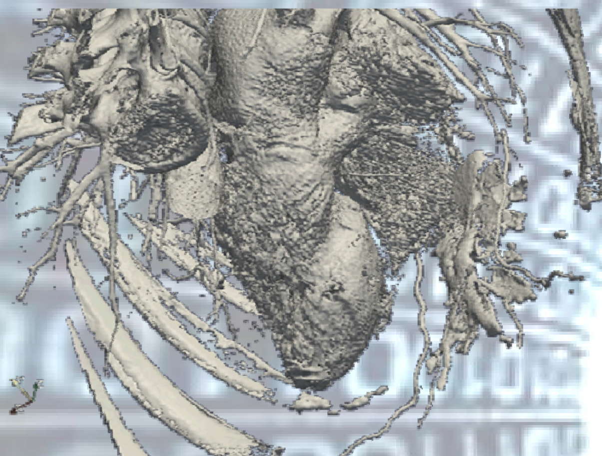
  - Approximately 2.2 million lines of code (as of Feb-27, 2012)
  - Initial cost: 718 person years, $39M (as of Feb-27, 2012)

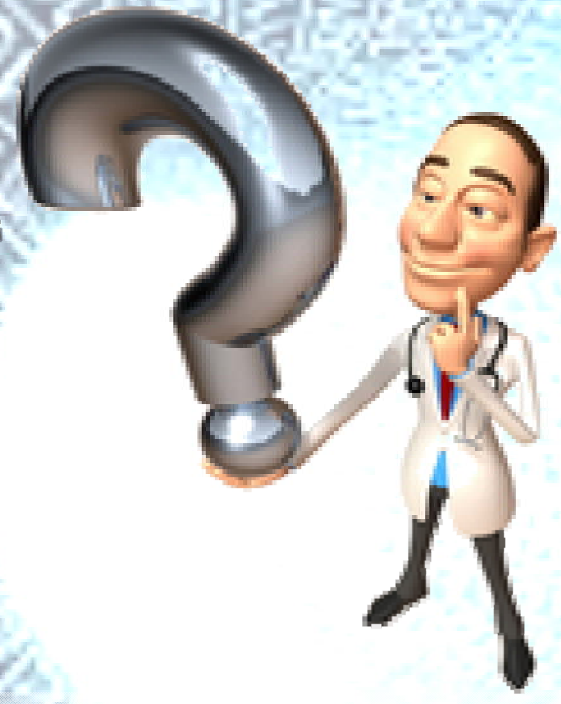
Introduction to MAT-ITK

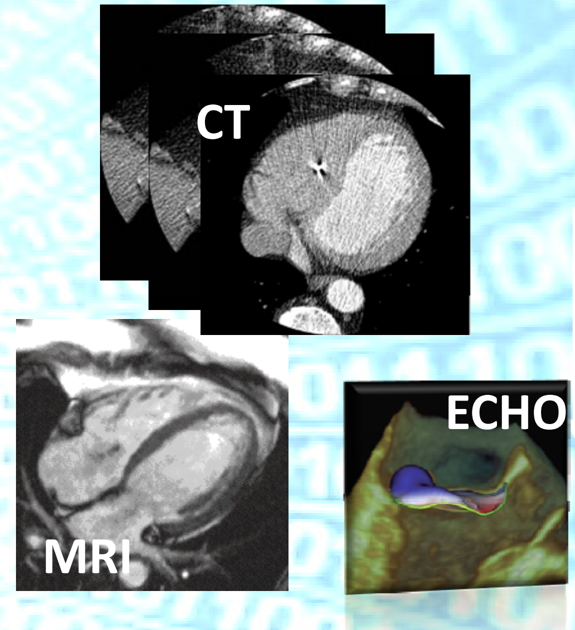
- Website: http://matitk.cs.sfu.ca/

Opcode	Method
SCC	ConfidenceConnectedSegmentation
SCSS	CellularSegmentationSegmentation(Debug)
SCT	ConnectedThresholdSegmentation
SFM	FastMarchSegmentation
SGAC	GeodesicActiveContourLevelSetSegmentation
SIC	IsolatedConnectedSegmentation
SLLS	LaplacianLevelSetLevelSetSegmentation
SNC	NeighbourhoodConnectedSegmentation
SOT	OtsuThresholdSegmentation
SSDLS	ShapeDetectionLevelSetFilter
SWS	WatershedSegmentation

BIA 2014







Prahlad G Menon, PhD

www.justcallharry.com

+1 412-259-3031

pgmenon@andrew.cmu.edu