LECTURE 13: Clustering Based Medical Image Segmentation Methods

Dr. Ulas Bagci
HEC 221, Center for Research in Computer Vision (CRCV), University of Central Florida (UCF), Orlando, FL 32814.

bagci@ucf.edu or bagci@crcv.ucf.edu
Reminder (Project Selection)

- Lung Lobe Segmentation from CT Scans (Use LOLA11 Segmentation Challenge Data Set)
- Segmentation of Knee Images from MRI (Use SKI 2010 Data Set)
- Multimodal Brain Tumor Segmentation (Use BraTS Data Set)
- Automatically measure end-systolic and end-diastolic volumes in cardiac MRIs. (Use Kaggle Cardiac Data Set)
- Head-Neck Auto Segmentation Challenge (Use MICCAI 2015 Segmentation Challenge Data Set)
- CAD of Dementia from Structural MRI (Use MICCAI 2014 Segmentation Challenge Data Set)
- DTI Tractography Challenge (Use MICCAI 2014 Segmentation Challenge Data Set)
- EMPIRE 2010 - Pulmonary Image Registration Challenge ([http://empire10.isi.uu.nl/index.php](http://empire10.isi.uu.nl/index.php), (Links to an external site.) I have the team name and password for downloading the data set).
Reminder (Project Selection)

- MICCAI 2017 Challenges (end of march, sites will be fully functional)

<table>
<thead>
<tr>
<th>Full/Half</th>
<th>Challenge</th>
<th>Description</th>
<th>Team Members</th>
<th>Email(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>BRATS</td>
<td>BRATS 2017 - Multimodal Brain Tumor Segmentation Benchmark: &quot;Survival Prediction&quot;</td>
<td>Spyridon Bakas</td>
<td><a href="mailto:s.bakas@uphs.upenn.edu">s.bakas@uphs.upenn.edu</a></td>
</tr>
<tr>
<td>Half</td>
<td>CDMRI</td>
<td>Diffusion MRI data harmonisation</td>
<td>Francesco Grussu</td>
<td><a href="mailto:f.grussu@ucl.ac.uk">f.grussu@ucl.ac.uk</a></td>
</tr>
<tr>
<td>Half</td>
<td>WMH</td>
<td>WMH Segmentation Challenge</td>
<td>Hugo J. Kuijf</td>
<td><a href="mailto:h.kuijf@umcutrecht.nl">h.kuijf@umcutrecht.nl</a></td>
</tr>
<tr>
<td>Half</td>
<td>CoronARe</td>
<td>Coronary Artery Reconstruction (CoronARe)</td>
<td>Sercan Çimen, Mathias Unberath</td>
<td><a href="mailto:s.cimen@sheffield.ac.uk">s.cimen@sheffield.ac.uk</a>; <a href="mailto:mathias.unberath@fau.de">mathias.unberath@fau.de</a></td>
</tr>
<tr>
<td>Half</td>
<td>CPM</td>
<td>Computational Precision Medicine: Cancer Challenges in CADe, Radiomics, and Digital Pathology</td>
<td>Keyvan Farahani</td>
<td><a href="mailto:Farahani@nih.gov">Farahani@nih.gov</a></td>
</tr>
<tr>
<td>Full</td>
<td>EndoVis</td>
<td>Endoscopic Vision Challenge</td>
<td>Stefanie Speidel</td>
<td><a href="mailto:speidel@kt.edu">speidel@kt.edu</a></td>
</tr>
<tr>
<td>Full</td>
<td>ISLES</td>
<td>ISLES 2017 - Ischemic Stroke Lesion Segmentation</td>
<td>Mauricio Reyes, Roland Wiest, Stefan Wrinzeck</td>
<td><a href="mailto:mauricio.reyes@istb.unibe.ch">mauricio.reyes@istb.unibe.ch</a>, <a href="mailto:roland.wiest@insel.ch">roland.wiest@insel.ch</a>, <a href="mailto:sw742@cam.ac.uk">sw742@cam.ac.uk</a></td>
</tr>
<tr>
<td>Full</td>
<td>iSeg-2017</td>
<td>MICCAI Grand Challenge on 6-month Infant Brain MRI Segmentation</td>
<td>Dr. Li Wang</td>
<td><a href="mailto:li_wang@med.unc.edu">li_wang@med.unc.edu</a></td>
</tr>
<tr>
<td>Full</td>
<td>STACOM</td>
<td>M Challenge: Multi-Modality Whole Heart Segmentation (MM-WHS)</td>
<td>Xiaohai Zhuang</td>
<td><a href="mailto:zwh@fudan.edu.cn">zwh@fudan.edu.cn</a></td>
</tr>
<tr>
<td>Half</td>
<td>ACDC</td>
<td>Automatic Cardiac Delineation Challenge (AC-DC)</td>
<td>Pierre-Marc Jodoin</td>
<td><a href="mailto:pierre-marc.jodoin@usherbrooke.ca">pierre-marc.jodoin@usherbrooke.ca</a></td>
</tr>
<tr>
<td>Half</td>
<td>HRNTS</td>
<td>High resolution neuronal tissue segmentation challenge</td>
<td>Keerthi Ram</td>
<td><a href="mailto:keerthi@htic.iitm.ac.in">keerthi@htic.iitm.ac.in</a></td>
</tr>
<tr>
<td>Half</td>
<td>LITS</td>
<td>LITS - Liver tumor segmentation challenge</td>
<td>Patrick Christ</td>
<td><a href="mailto:patrick.christ@tum.de">patrick.christ@tum.de</a></td>
</tr>
<tr>
<td>Half</td>
<td>RETOUCH</td>
<td>RETOUCH-RETinal Oct fUid Challenge</td>
<td>Hrvoje Bogunovic</td>
<td><a href="mailto:hrvoje.bogunovic@meduniwien.ac.at">hrvoje.bogunovic@meduniwien.ac.at</a></td>
</tr>
<tr>
<td>Full</td>
<td>ENIGMA</td>
<td>ENIGMA Cerebellum</td>
<td>Jerry L. Prince</td>
<td><a href="mailto:prince@jhu.edu">prince@jhu.edu</a></td>
</tr>
</tbody>
</table>
Outline

• Clustering
  – K-means
  – FCM (fuzzy c-means)
  – SMC (simple membership based clustering)
  – AP (affinity propagation)
  – FLAB (fuzzy locally adaptive Bayesian)
  – Spectral Clustering Methods
What is Clustering?

• Organizing data into classes such that:
  – High intra-class similarity
  – Low inter-class similarity

• Finding the class labels and the number of classes directly from the data (as opposed to classification tasks)
What is a natural grouping?
What is a natural grouping?

Clustering is subjective

Simpson's Family  School Employees  Females  Males
What is similarity?
What is similarity?

Cluster by features

• Color
• Intensity
• Location
• Texture
• ....
Distance metrics

0.23

3

342.7
Motivation for Clustering in Medical Image Segmentation

- **Assumption:** The object of interest can be identified as a cluster in an appropriate feature space. Delineate the cluster to delineate the object.
• **Clustering algorithms** essentially perform the same function as **classifier methods** without the use of training data.

• Thus, they are termed **unsupervised methods**.

• In order to compensate for the lack of training data, clustering methods iterate between segmenting the image and characterizing the properties of each class.
Motivation in Biomedical Image Segmentation

Segmenting the image based on threshold implies underlying a posteriori probabilities for the different classes of which the histogram shows a non-normalized sum.
K-means Clustering
K-means Clustering
K-means Clustering
K-means Clustering
K-means Clustering
K-means Clustering
K-means Clustering
K-MEANS CLUSTERING

- Description

Given a set of observations \( (x_1, x_2, \ldots, x_n) \), where each observation is a \( d \)-dimensional real vector, \( k \)-means clustering aims to partition the \( n \) observations into \( k \) sets \( (k \leq n) \) \( S = \{S_1, S_2, \ldots, S_k\} \) so as to minimize the within-cluster sum of squares (WCSS):

\[
\arg \min_S \sum_{i=1}^{k} \sum_{x_j \in S_i} \|x_j - u_i\|^2
\]

where \( u_i \) is the mean of points in \( S_i \).

- Standard Algorithm

1) \( k \) initial "means" (in this case \( k=3 \)) are randomly selected from the data set.

2) \( k \) clusters are created by associating every observation with the nearest mean.

3) The centroid of each of the \( k \) clusters becomes the new means.

4) Steps 2 and 3 are repeated until convergence has been reached.
Recap: Core of Clustering

• Similarity metric
• Distance metric

Open-box Spectral clustering
Software for Medical images

Schultz and Kindlmann 2013, IEEE TVCG
Multispectral/Multimodal Image Segmentation

- The segmentation techniques based on integration of information from several images are called multispectral or multimodal.
Multispectral/Multimodal Image Segmentation

• The segmentation techniques based on integration of information from several images are called multispectral or multimodal.

• In multispectral images, each pixel is characterized by a set of features and the segmentation can be performed in multidimensional (multichannel) feature space using clustering algorithms.
Methods:

- $k$-Nearest Neighbor ($k$NN)
- C-Means ($CM$)
- Fuzzy C-Means ($FCM$)
- Simple Membership-based Classification ($SMC$)

Methods differ based on how clusters are defined, detected, and delineated.
**Rpl : Clustering - kNN**

**Training:**

For each of $L$ object classes, determine a true classification feature set $F_i$, $i = 1, 2, \ldots, L$.

**Classification:**

1. Choose and fix a value for $K$ (say $K = 7$).

2. For a given scene $S = (C, f)$ to be segmented, determine the feature vector $x_c$ associated with each voxel $c$. 
(3) To classify voxel \( c \) in \( C \), among all feature vectors in \( \bigcup_{i} F_{i} \), determine \( K \) vectors \( x_{c}^{1}, \ldots, x_{c}^{K} \) that are closest to \( x_{c} \).

(4) Classify \( c \) to that among the \( L \) classes which is represented maximally in \( x_{c}^{1}, \ldots, x_{c}^{K} \).
Applications in MR Brain Tissue Segmentation

(a) 
(b) 
(c) 
(d) 
(e) 
(f)
Rpl : Clustering - kNN

MRI T2
Original

MRI PD
Original

GM, WM, CSF
Segmented
**Rpl : Clustering - FCM**

It is an unsupervised method. No training needed.

- Requires number of object classes $L$ to be specified.
- Outputs fuzzy membership, at each voxel, in individual object classes.
- Also outputs class mean (centroid)
- Uses an optimization technique to determine how to optimally partition the feature space into $L$ classes.
$S = (C, f) : \text{a given scene}$

$u_{ck} : \text{membership of voxel } c \text{ in object class } k.$

$q : \text{a weighting exponent (} q = 2 \text{ often used).}$

$f(c) : \text{vector-valued scene intensity.}$

$\mu_k : \text{centroid (mean) of class } k.$

$\| \cdot \| : \text{any inner product norm (e.g., Euclidean norm)}$

$$J = \sum_{c \in C} \sum_{k=1}^{L} u_{ck}^q \| f(c) - \mu_k \|^2$$

Minimization of $J$ yields $u_{ck}$ and $\mu_k.$
Applications in T2-MRI Segmentation via K-means (K=9)
Quantitative Estimation of Volumetric Breast Tomosynthesis Images (Pertuz et al, Radiology 2015)
Quantitative Estimation of Volumetric Breast Tomosynthesis Images (Pertuz et al, Radiology 2015)

• Breast cancer is the most commonly diagnosed cancer in the United States and the second leading cause of death from cancer in women.
Quantitative Estimation of Volumetric Breast Tomosynthesis Images (Pertuz et al, Radiology 2015)

- Breast cancer is the most commonly diagnosed cancer in the United States and the second leading cause of death from cancer in women.
- Breast density is an independent risk factor for breast cancer
Quantitative Estimation of Volumetric Breast Tomosynthesis Images (Pertuz et al, Radiology 2015)

- Breast cancer is the most commonly diagnosed cancer in the United States and the second leading cause of death from cancer in women.
- Breast density is an independent risk factor for breast cancer.
- Objective and accurate methods for the estimation of density are needed to ensure reliable estimation and, ultimately, yield quantitative reproducible measures that are clinically useful.
Quantitative Estimation of Volumetric Breast Tomosynthesis Images (Pertuz et al, Radiology 2015)

• Breast cancer is the most commonly diagnosed cancer in the United States and the second leading cause of death from cancer in women.

• Breast density is an independent risk factor for breast cancer

• Objective and accurate methods for the estimation of density are needed to ensure reliable estimation and, ultimately, yield quantitative reproducible measures that are clinically useful.

• Volume-based quantitative density methods, the aim is to better estimate the amount of fibroglandular (i.e., dense) tissue with respect to the total volume of the breast.
Quantitative Estimation of Volumetric Breast Tomosynthesis Images (Pertuz et al, Radiology 2015)

Red: human annotator, yellow: computer algorithm
Let us dig into the theory a bit more…

- **Unsupervised learning**
  - Finding clusters
  - Dimensionality reduction (PCA, MDS, …)
  - Building topographic maps
  - Finding hidden causes or sources of the data
  - Modeling the data density
Let us dig into the theory a bit more...

- **Unsupervised learning**
  - Finding clusters
  - Dimensionality reduction (PCA, MDS,…)
  - Building topographic maps
  - Finding hidden causes or sources of the data
  - Modeling the data density

- **Uses of Unsupervised learning**
  - Classification
  - Data compression
  - Make other learning tasks easier
  - A theory of human learning and perception
  - ....
Mixtures of Gaussians

Goal: finding clusters in data.

To generate data from this model, assuming $K$ clusters:

- Pick cluster $k \in \{1, \ldots, K\}$ with probability $\pi_k$
- Generate data according to a Gaussian with mean $\mu_k$ and covariance $\Sigma_k$

\[
P(y) = \sum_{k=1}^{K} P(x = k)P(y|x = k)
\]
\[
= \sum_{k=1}^{K} \pi_k \mathcal{N}(y|\mu_k, \Sigma_k)
\]

x: input, y: output
Mixtures of Gaussians

**E-step:** Compute responsibilities for each data vec. $y^{(i)}$

$$r_{ki} \equiv P(x = k | y^{(i)}) = \frac{\pi_k \mathcal{N}(y^{(i)} | \mu_k, \Sigma_k)}{\sum_{\ell=1}^{K} \pi_{\ell} \mathcal{N}(y^{(i)} | \mu_{\ell}, \Sigma_{\ell})}$$

**M-step:** Estimate $\pi_k$, $\mu_k$ and $\Sigma_k$ using data weighted by the responsibilities.

The k-means algorithm for clustering is a special case of EM for mixture of Gaussians where $\Sigma_k = \lim_{\epsilon \to 0} \epsilon I$

You can download GMM train and test C/C++ code from my webpage:
http://www.cs.ucf.edu/~bagci/software.html

- Each tissue is represented by large number of Gaussians (to capture complex tissue spatial layout)

- Each tissue is represented by large number of Gaussians (to capture complex tissue spatial layout)
- **Intensity** is considered as a global parameter

- Each tissue is represented by a large number of Gaussians (to capture complex tissue spatial layout)
- **Intensity** is considered as a global parameter
- EM is utilized to learn parameter-tied CGMM.
  - Each Gaussian is linked to a single tissue and all the Gaussians related to the same tissue share the same intensity patterns.

- Each tissue is represented by large number of Gaussians (to capture complex tissue spatial layout)
- **Intensity** is considered as a global parameter
- EM is utilized to learn parameter-tied CGMM.
  - Each Gaussian is linked to a single tissue and all the Gaussians related to the same tissue share the same intensity patterns.
- Completely unsupervised, no alignment/atlas registration required.

*But note that general properties of T1 is used to include intensity info.*

(a) %9 noise, MRI from brain web, slice 95
(b) EM based algorithm
(c) CGMM algorithm results
Clustering methods for PET Images

• FCM (fuzzy c-means)
• FLAB (fuzzy locally adaptive Bayesian)
• K-means
• K-NN
• Spectral Clustering
• ….
Fuzzy Locally Adaptive Bayesian (FLAB) Approach for PET Segmentation (Hatt et al, TMI)

\[ P(X|Y) = \frac{P(X, Y)}{P(Y)} = \frac{P(Y|X)P(X)}{P(Y)} \]

- **Y**: observed image (noisy)
- **X**: (binary) segmented object
- **Bayesian setting**: given the image Y, what is the segmentation result X?
- **P (YIX)** is the likelihood of the observation Y conditionally with respect to the hidden ground-truth X
- **P(X)**: prior
- **P(X|Y)**: posterior distribution
Distribution of X (Prior)

- Two hard classes (Object + background) + finite number of fuzzy levels.
- Consider $Y_0$ and $Y_1$ distributions, with mean and variations

\[(\mu_0, \sigma_0^2) \quad (\mu_1, \sigma_1^2)\]

- Then, mean and standard deviation of each fuzzy levels are represented as

\[
\mu_{F_i} = \mu_0(1 - \epsilon_i) + \epsilon_i \mu_1
\]

\[
\sigma_{F_i}^2 = \sigma_0^2(1 - \epsilon_i)^2 + \epsilon_i^2 \sigma_1^2
\]

- Use EM algorithm to find unknowns.

*Epsilon 1/3, 2/3 used.*
Clustering methods for PET Images

- FCM (fuzzy c-means)
- FLAB (fuzzy locally adaptive Bayesian)
- K-means
- K-NN
- Spectral Clustering
- AP Clustering (affinity propagation)
  - State of the art for multi-focal uptake segmentation

Foster, Bagci, et al., 2014
Example: The 15 images with highest squared error under either affinity propagation or k-centers clustering are shown in the top row. The middle and bottom rows show the exemplars assigned by the two methods, and the boxes show which of the two methods performed better for that image, in terms of squared error. Affinity propagation found higher-quality exemplars.
**Affinity Propagation (AP) Based Clustering**

- Clusters data by locating exemplars (the data point that best describes a group of data) based on the similarities between pairs of points by passing “messages” between the data
  - Responsibility $r(i,k)$
    - How responsible is point $k$ for being $i$’s exemplar
  - Availability $a(i,k)$
    - How available is point $k$ to be point $i$’s exemplar
  - Messages are iteratively sent and compete until convergence
Motivation for AP in PET Images (multi-focal update)
Threshold Selection

Observed histogram of a PET image is the summation of the histogram of hidden objects

Objective for segmentation: Locate the optimal thresholding value(s) for best separation between these objects

\[
\begin{align*}
\rho(i, k) &\leftarrow s(i, k) - \max_{k' \neq k} \{ a(i, k') + s(i, k') \}, \\
a(i, k) &\leftarrow \min \left\{ 0, r(k, k) + \sum_{i' \neq i} \max \{ 0, r(i', k) \} \right\}.
\end{align*}
\]
Similarity Function Definition

\[ s(i,j) = -\left( | d_{ij}^x |^n + | d_{ij}^G |^m \right) \]

- **Intensity**
  - \( p_i \)
  - \( p_j \)

- **Intuitively:**
  - Probability difference - Distance includes information of the peaks and valleys between data points—needed for classification

- **Gradient information**
  - \( d_{ij}^G \)

- **Probability information**
  - \( d_{ij}^x \)

- **Weight parameters** learned for different image types and for small animal/human images
AP is an iterative process
Convergence of AP Segmentation
Comparison to Other Methods
Summary

- K-means
- FCM (fuzzy c-means)
- K-NN (k-nearest neighborhood)
- SMC (simple membership based clustering)
- AP (affinity propagation)
  - Useful for PET image segmentation (multi-focal uptake)
Slide Credits and References

• **Credits to:** Jayaram K. Udupa of Univ. of Penn., MIPG
• Bagci’s CV Course 2015 Fall.
• K.D. Toennies, Guide to Medical Image Analysis,
• Handbook of Medical Imaging, Vol. 2. SPIE Press.
• Handbook of Biomedical Imaging, Paragios, Duncan, Ayache.
• Seutens, P., Medical Imaging, Cambridge Press.
• Z. Ghahramani, U. Cambridge, UK.