Hierarchical Mode Analysis

Presented in Partial Fulfillment
of the Requirements for

CSC 471/578

in the Department of Computer Science

Ashlin Richardson

University of Victoria

April 18, 2013

Overview

Data driven clustering

Clustering problem: given N points in $\mathbb{R}^n$, assign each point a label. "Group data into statistically meaningful categories". ..Implementation[1][2] invited for inclusion in E.S.A. open source distro:

- **Data driven clustering:** find arbitrary shaped clusters, instead of generalized "spherical cows" $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$.
- **Point-cloud representation** of clusters (instead of, e.g., (μ, Σ) where $\Sigma = Cov(x_i, x_j)$).

Two steps:

- **Density Estimation**
- **Hill Climbing**

---

[1] A. Richardson *et al.*, "Unsupervised nonparametric classification of polarimetric sar data using the k-nearest neighbor graph," *proc. IEEE IGARSS*, 2010

[2] D. Goodenough *et al.*, "Mapping fire scars using radarsat-2 polarimetric sar data," *Can. J. Remote Sensing*, 2011

Data driven clustering: Step 1/2

Density Estimation e.g., by K-nearest Neighbor Density Estimate:

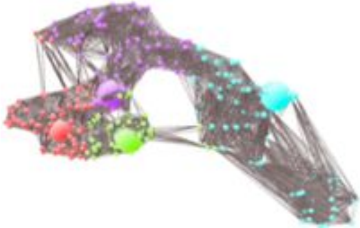
$$\rho(x) = \frac{1}{\frac{1}{K}\sum_{y \in N_x} d(x,y)}$$

- Evaluate this at each data point..
- K is the number of “nearest neighbors” to consider about a point
- N_x is the set of “K-nearest neighbours” about a point x

Data driven clustering: Step 2/2

Hill Climbing. Simple heuristic (avoids explicit gradient-estimation).

label[x] = **hillclimb**(x): //label the point x
  if($\rho(x) \geq \rho(N_x)$): //x highest-ρ in nbhd **(x is a "top")**
    return newLabel();
  else //"go up" **(find the centre)**
    return hillclimb$\left(\underset{x \in N_x}{\operatorname{argmax}}(\rho(N_x))\right)$;



Itty-bitty example..small number of points. **Peaks/tops** shown as large circles. Other points colored by label.

Results: Point-clouds good for compression. “Nearest Neighbor” ($\mu's$) bad..

Results: Point-clouds good for compression. "Nearest Neighbor" ($\mu's$) bad.. Also, Data-driven method more stable (w.r.t. parameters)..

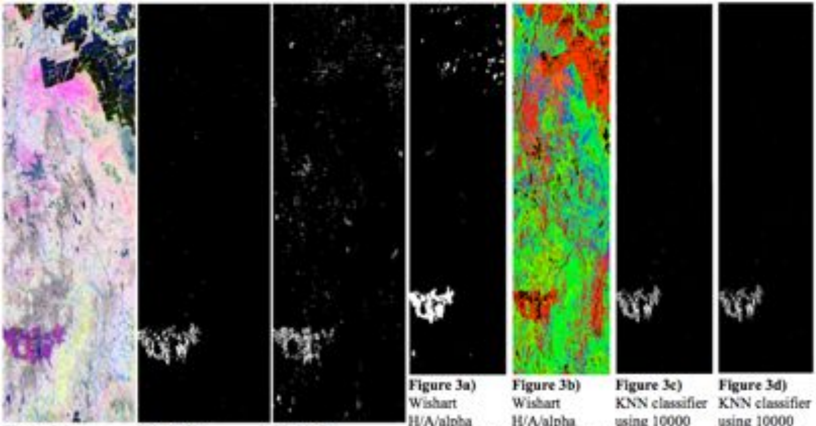


Figure 2a) Red=Entropy, Blue=Alpha, Green=Shannon Entropy.

Figure 2b) Lee algorithm – 200 initial segments – 35 final clusters.

Figure 2c) Lee algorithm – 200 initial segments – 40 final clusters.

Figure 3a) Wishart H/A/alpha classifier - 5x5 analysis window (fire scar class).

Figure 3b) Wishart H/A/alpha classifier – 1x1 analysis window (16 classes).

Figure 3c) KNN classifier using 10000 points, K=55, 22 clusters.

Figure 3d) KNN classifier using 10000 points, K=35, 51 clusters.

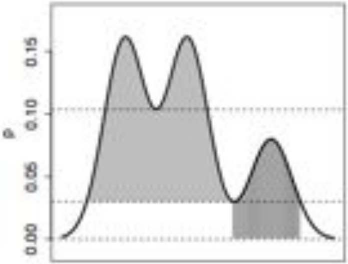
Hierarchical Mode Analysis: Data Driven Clustering in more general, Dendrographic (tree) Context

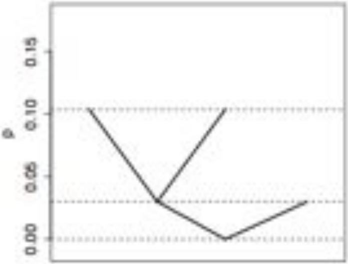
Review: categories (cf., labels) were representative of connected components of level sets L:

$$L(\lambda; \rho) = \{x | \rho(x) > \lambda\}$$

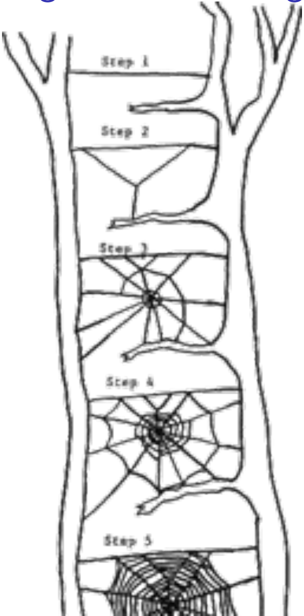
for some threshold λ ("one-level analysis").

Next: Hierarchical version! Cluster tree: (recall who's-your-daddy operator from last time)..Want russian-dolled (nesting) buckets.. **Coming up: Segmentation**





Hierarchical Mode Analysis: Data Driven Clustering in more general, Dendrographic (tree) Context

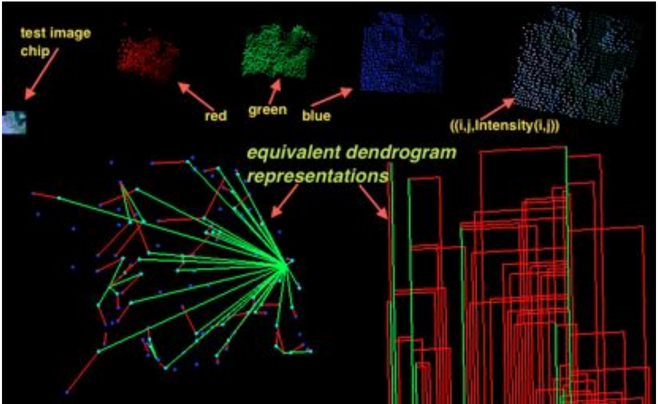


Segmentation Motivating Example: "Intensity Climbing"

In clustering procedure,

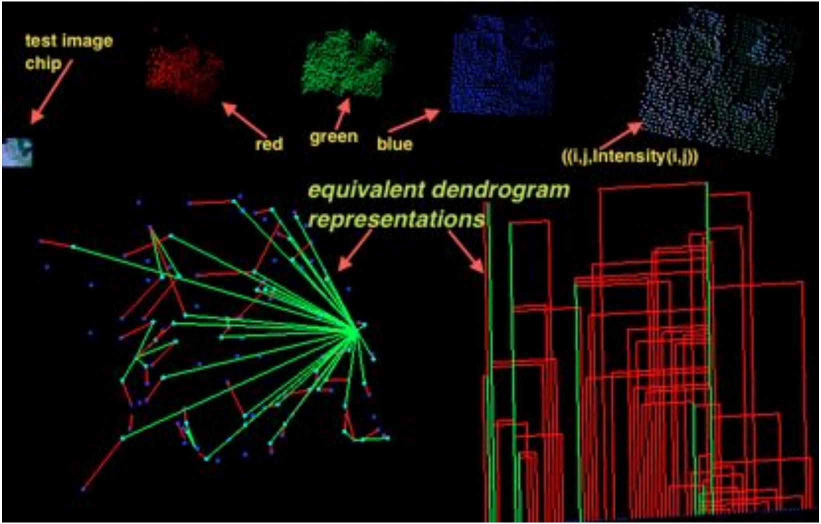
1 Use **Intensity** $I(x(i,j)) = \sum_k x(i,j)_k$ instead of $\rho(x)$!
2) Use a spatial neighbourhood. E.g., a square one: $N^{\text{Moore}}(x_{i,j}) = \{x_{i-1,j}, x_{i+1,j}, x_{j+1,i-1}, x_{j+1,i}, x_{j+1,i+1}, x_{j-1,i-1}, x_{j-1,i}, x_{j-1,i+1}\}$ instead of (spectral) K-Nearest Neighbourhood.

Result: map out hierarchical structure of peaks in image intensity surface (N.B., color information neglected..)



Segmentation Motivating Example: “Intensity Climbing”

Sample of “hyperspectral” segmentation result for image chip. The result is 5-dimensional (only three dimensions are visualized). Data is ALOS-PALSAR (spaceborne) polarimetric imaging (Japan Space Agency) over Keg river area, BC. The five channels are the following ”decomposition parameters”: cloude-pottier entropy, alpha, and anisotropy; shannon entropy; radar vegetation index.



Segmentation Motivating Example: "Intensity Climbing"

Q: how to generalize?

- How to incorporate color information?
- How to incorporate geometric (neighborhood information) simultaneously?

Segmentation via Hierarchical Mode Analysis

Q: how to generalize?

- Use simultaneous color/geometry information
- Use hierarchical formulation for scalar field segmentation (with underlying geometry: neighborhood structure).

Choice of neighbourhoods as K-Nearest-Neighbourhoods (KNN).. And choice of scalar field as KNN density estimate.. This algorithm reduces to the previous case! **Benefits of new version..**

- Arbitrary geometries (e.g., non-rectangular/higher-dimensional data, e.g., 6-d-toroidally sampled hyperspectral data..). N.b., issue of geometry estimation from sensor-point-clouds not addressed here..
- Being a little bit more vague (referring to ρ as a scalar field, rather than a density estimate).. more scope of applications! E.g., temperature measurements..

Segmentation via Hierarchical Mode Analysis: Formulation

Algorithm input: a set X of $N < \infty$ points $x \in \mathbb{R}^n$, $n \geq 1$

$$x = (x_1, x_2, \ldots, x_n)$$

, a scalar field $\rho(x) : X \to \mathbb{R}$,
and, a neighbourhood structure: $N_x(x \in X) : X \to 2^X$.

Output: labels with tree structure (associating tree-splits with scalar-field levels $\rho(x)$)

Note: treat color/geometry together via sensible choice of $\rho(x)$..

Segmentation via Hierarchical Mode Analysis: Formulation

Modes of scalar field

A regional maximum of the scalar field $\rho(x)$ ($\rho_{i,j} = \rho_{\pi(x)} = \rho(x)$ for conventional 2-d image) we shall refer to as a **Mode**, which we will **symbolize** by M, using this:

1. to refer to the **basin of attraction** associated with M
2. to refer to the associated **maximum** $\sup_{x \in M} C(x) = \sup C(M)$
3. as a label/name.

meta!

Segmentation via Hierarchical Mode Analysis: Formulation

Interfaces between modes

Suppose modes M_i, M_j share an interface:

$$\begin{aligned} I_{i,j} = I_{M_i,M_j} &= I_{M_j,M_i} \\ &= (N(M_i) \cap M_j) \cup (N(M_j) \cap M_i) \\ &= (\cup N(M_i) \cap M_j) \cup (N(M_j) \cap M_i) \\ &= \left\{ N(x \in M_i \cup M_j) \middle| N \cap M_j \cap M_i \neq \emptyset \right\} \end{aligned}$$

that is, the set of neighborhoods of points in modes M_i or M_j whose neighborhoods intersect with the other mode.

Segmentation via Hierarchical Mode Analysis: Formulation

Interfaces define the upper Level set inclusion structure (dendrogram)

For $\lambda > \sup\left(\rho(I_{M_i,M_j})\right)$, Modes M_i, M_j are no longer associated with the same connected component of the upper-level sets:

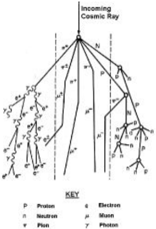
$$\{L(\lambda; \rho(x)) = \{x | \rho(x) \geq \lambda\}.$$

That is, the connected components of the upper-level sets associated with M_i, M_j are “joined” at level:

$$\lambda = \sup \rho(I_{i,j})$$

but are disjoint for

$$\lambda > \sup\left(\rho(I_{M_i,M_j})\right).$$



Segmentation via Hierarchical Mode Analysis: Formulation

Inclusion structure between neighboring modes

For a mode M we denote the neighboring modes (those modes that share interfaces with M) by

$$\mathbb{N}(M) = \left\{ M_i \middle| \cup N(M_i) \cap M \neq \emptyset \right\}.$$

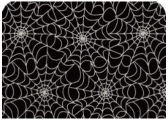
Then, for a node M,

$$S(M) = \underset{M_j \in \mathbb{N}(M)}{\arg\max} \left\{ \sup_{x \in I_{i,j}(M, M_j)} \rho(x) \right\}$$

refers to the (neighboring) component M_j which which M "joins" at level

$$\lambda_{S(M)} = \max_{M_j} \left\{ \sup_{x \in I_{i,j}(M, M_j)} \rho(x) \right\}.$$

Will refer to S(M) as the "high density neighbor mode" of M_i.



Segmentation via Hierarchical Mode Analysis: Formulation

Equivalence classes of connected components of level sets

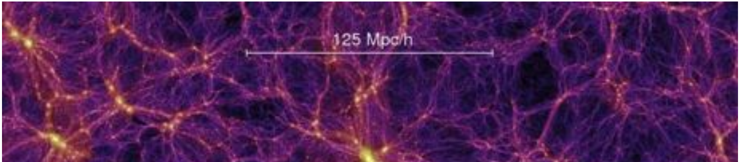
By $[M]_\lambda$ we denote the **equivalence class of connected components of upper-level sets** of $\rho(x)$ associated with M.

To illustrate.. In terms of decreasing heights λ, a node M "first" joins the same connected component as $S(M)$. Then, supposing M and $S(M)$ "join" the same connected component at

$$\lambda_{S(M)}$$

then

$$[M]_{\lambda_{S(M)}} \simeq_{[]_\lambda} [S(M)]_{\lambda_{S(M)}}.$$



Segmentation via Hierarchical Mode Analysis: Formulation

Inclusion trajectory diagram for a Mode (Segmentation Hierarchy) For M there is a sequence:

$$[M]_{\lambda_1}, [M]_{\lambda_2}, \ldots , [M]_{\lambda_n}$$

where from left to right the λ_i's are decreasing ($\lambda_i > \lambda_j$) for ($i < j$), and

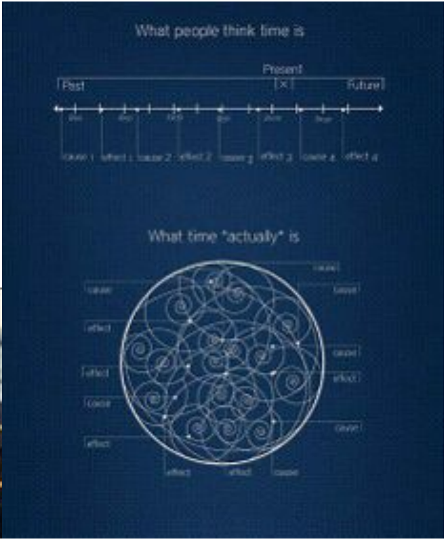
$$[M]_{\lambda_1} \subset [M]_{\lambda_2} \subset \ldots \subset [M]_{\lambda_n}$$

where the inclusion is proper (at each λ_i the connected component has addition(s)). Use notation

$$[M]_{\lambda_1} \to [M]_{\lambda_2} \to \ldots \to [M]_{\lambda_n}$$

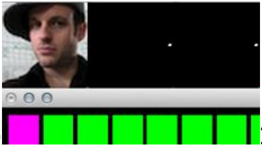
to emphasize connection with (discrete time) dynamical systems (bounded (spatio-temporal) (time)moment-(space)regions "included" into something greater (future accomodates past)?

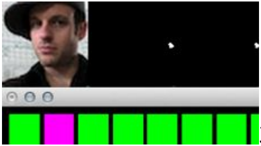

Segmentation via Hierarchical Mode Analysis: Formulation



Application: Smart Flood Fill

Select seed node (M) and successively apply “who’s your daddy operator” ($\rightarrow$)..

1) 

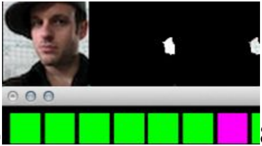
2) 

3) 

4) 

5) 

6) 

7) 

8) 

Note..Before taking color information into account, white background patch got added onto ear prematurely..

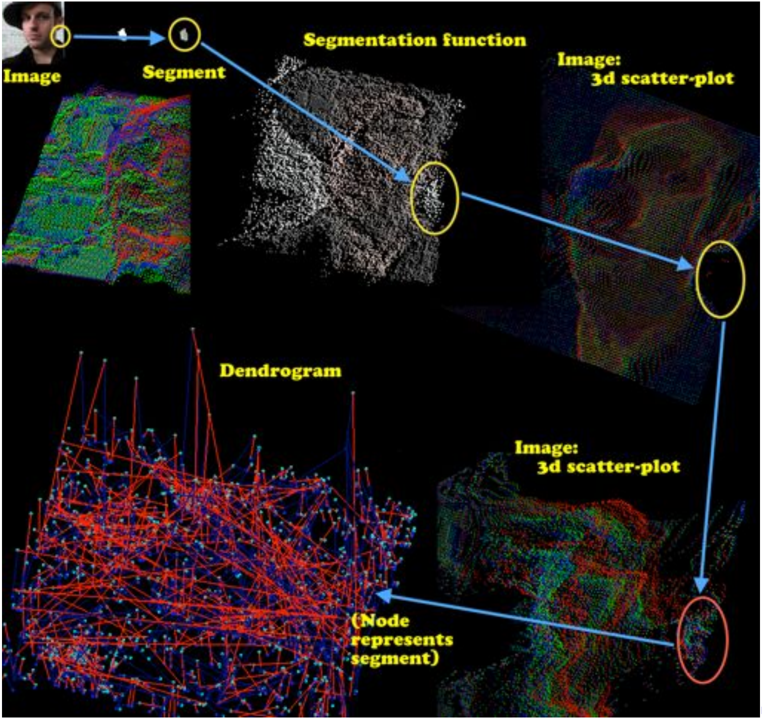


Image
Segment
Segmentation function
Image:
3d scatter-plot
Dendrogram
Image:
3d scatter-plot
(Node
represents
segment)

Acknowledgments

- Thanks to Prof. Bruce Gooch and the University of Victoria for financial support