Image Segmentation and Clustering Techniques

Statistical Computing midterm project

Ding-Chih Lin

National Tsing Hua University, Institute of Statistics

May 11, 2021

Agenda

Table of Contents

Image segmentation

Image Segmentation:

- Partition a digital image into multiple segments
- Simplify the representation of an image into something which is meaningful
- Assign a label to every pixel in an image such that pixels with the same label shared certain characteristics.

Practical applications

Practical applications:

- Tumor segmentation from medical images
- Object detection for autonomous vehicles or satellites
- Color quantization
- Optical character recognition (OCR)

Table of Contents

Method 1: K-means clustering

K-means Clustering:

- *Compactness*, cluster the objects which share lower distances
- Use *random partition method* to decide initial centroids
- Utilize naive K-means algorithm

Method 2: Spectral clustering

Spectral Clustering:

- *Connectedness*, cluster the objects which are connected to each other
- Build a graph $G(V, E)$ with corresponding weights (similarity) w_{ij}
- Identify the clusters of nodes based on the edges connecting them
- Utilize information from the eigenvalues of the W built from the graph

Method 3: Hidden Markov Model

Hidden Markov Model:

- Discrete-time doubly embedded stochastic process
- Process X being modeled is assumed to be a Markov process with unobserved states Y that depends on X
- Learn about X in order to observe or obtain Y
- Assume the emission probabilities are normally distributed

Simulation Study

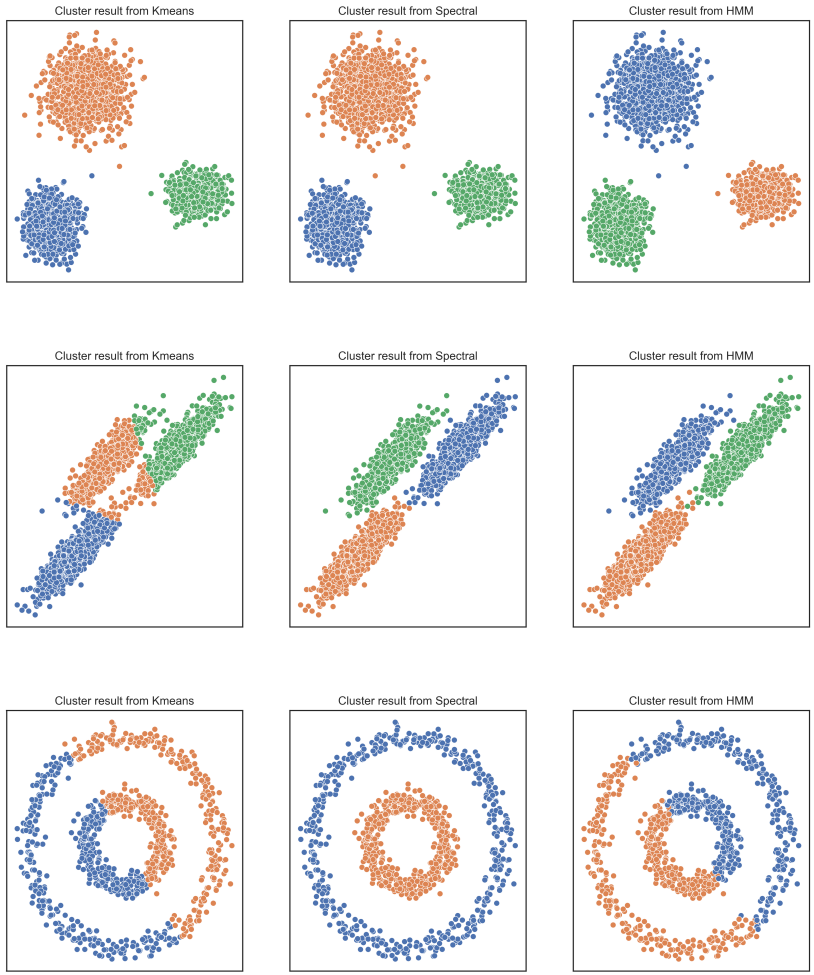


Figure: Clustering techniques in different scenarios

Table of Contents

Color Quantization

Color quantization reduces the number of distinct colors of an image while keeping the new image visually similar to the original one.

(a) Original image

(b) 256 code vectors

(c) 8 code vectors

Figure: Color quantization, *London Bridge, London, UK*

Color Quantization (Conti.)

This image consists of 1200x1200 pixels:

- Each pixel is represented as 3 bytes, which occupies 24 bits of storage
- 256 x 256 x 256 = 16.7 million possible combination colors
- It requires 4 MB of storage

Compressed image:

- with 64 code vectors requires 1.44 MB of storage
- with 8 code vectors only requires 540 kB of storage

Trade-off between image quality and storage requirements.

Brain Tumor Segmentation

Medical image segmentation is a task of automatically segmenting the targets of interest in a medical image.

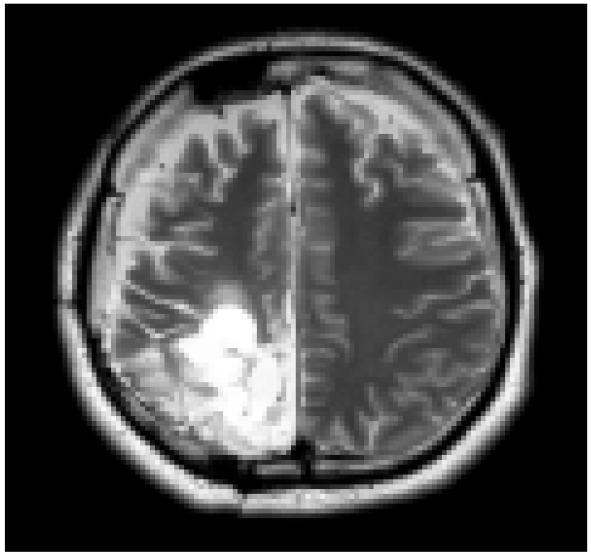
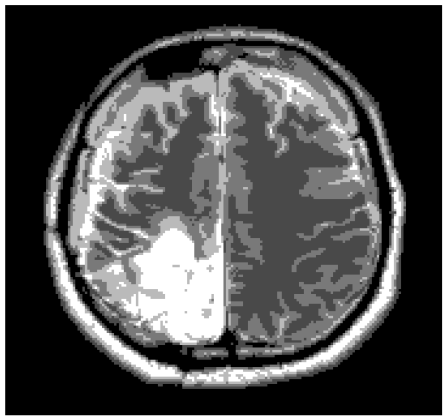


Figure: *tumor brain image* obtained from Stanford brain tumor center

Brain Tumor Segmentation

Spectral clustering outperforms the others in brain tumor segmentation.



(a) K-means clustering

(b) HMM clustering

(c) spectral clustering

Figure: Segmentation results