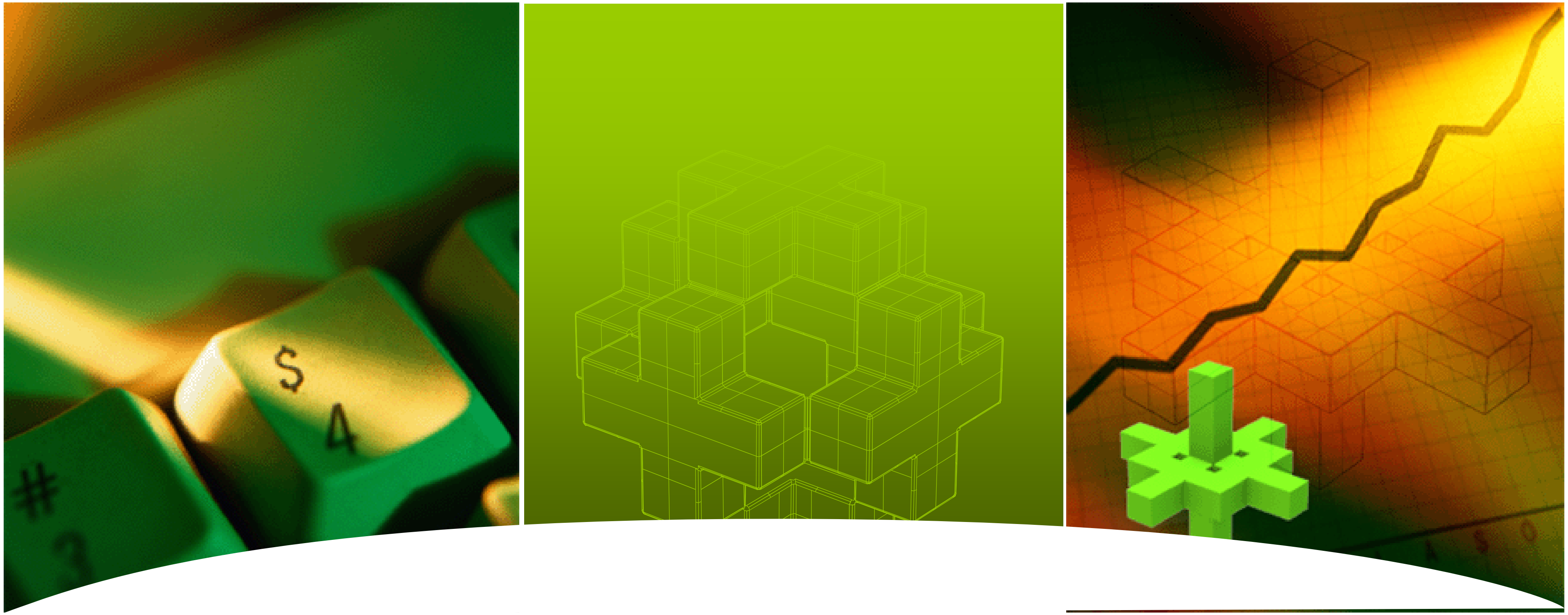


Lloyd Max Quantization

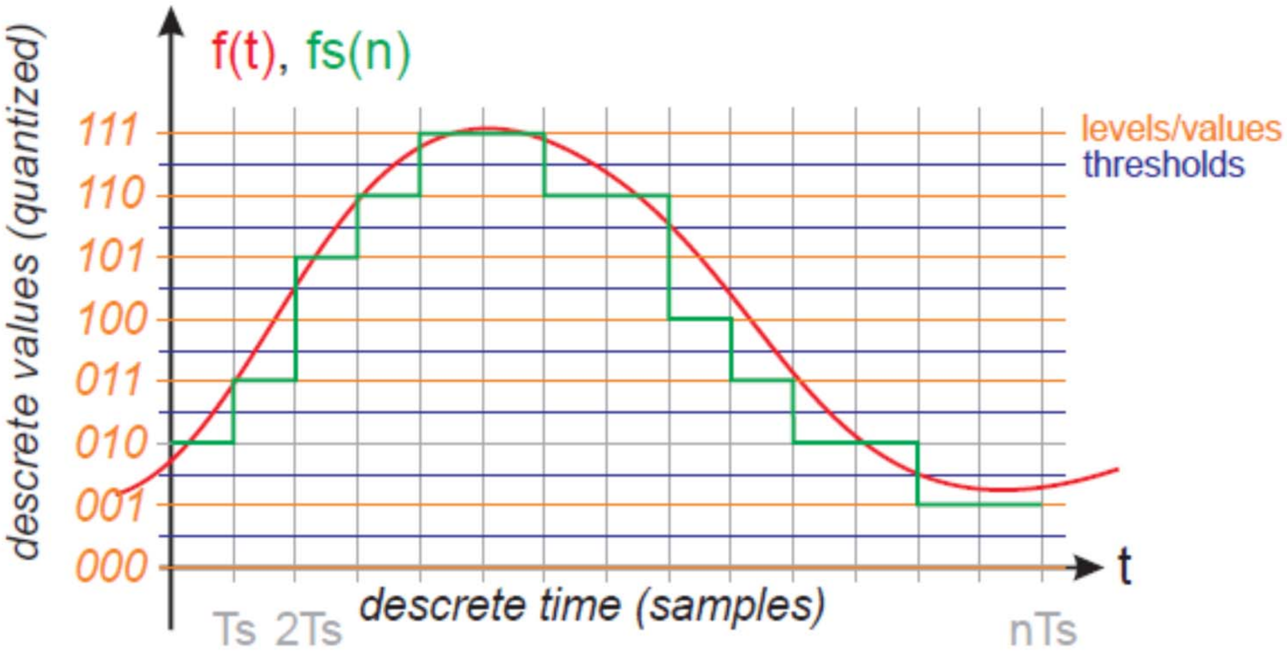
(K-means Clustering)

2015. 05. 27

Lecturer: Sang Hwa Lee

Quantization ?

❖ **Mapping continuous values to discrete levels**

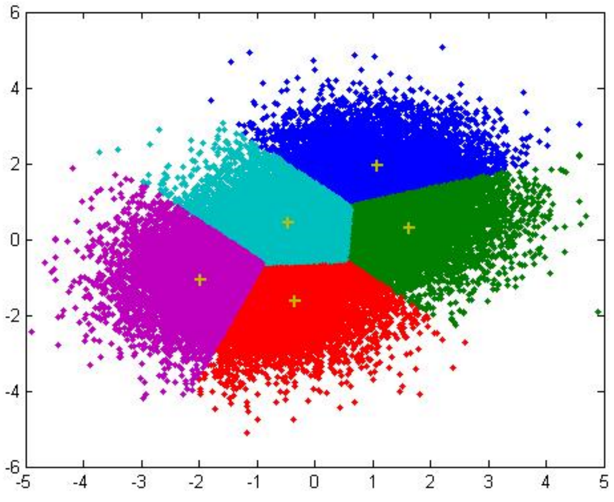


Quantization ?

- **Quantization parameters**
  - Step size, # of levels
  - Reconstruction value
- **Quantization is s kind of clustering**
  - Data coding and compression – quantization
    - Codewords
  - Pattern recognition – clustering
    - Representative Patterns
- **Scalar vs Vector quantization**
  - SQ: Partition of Single values
  - VQ: Partition of Vectors
- **Uniform vs non-uniform**

K-Means Clustering

- **Simplest unsupervised clustering of vectors**
  - Parameter K: number of clusters
- **Principle**
  - Euclidean distance based clustering
  - Cluster boundary: Equi-distance hyper-plane
    - Voronoi boundary



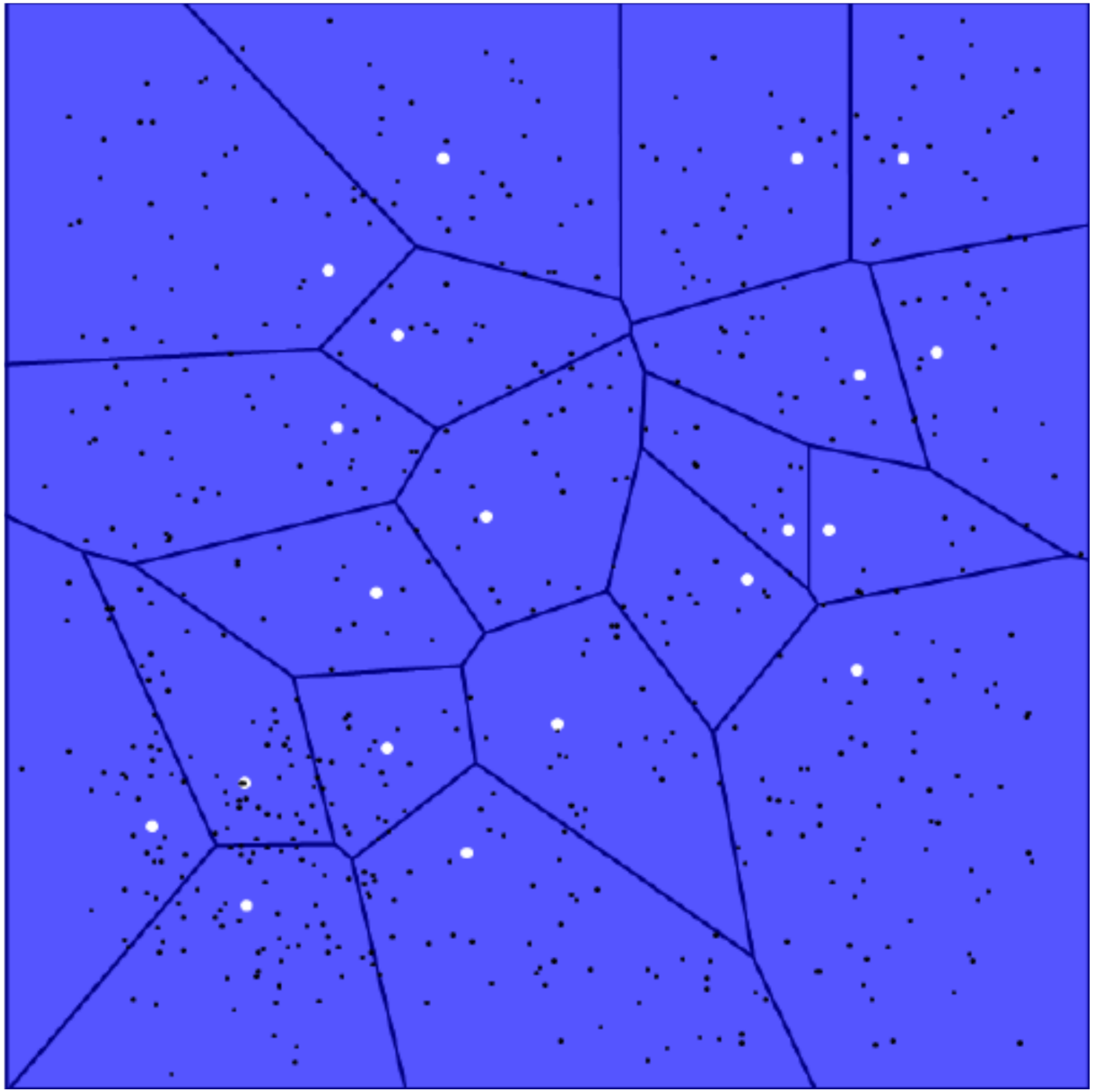
K-Means Clustering

❖ Basic Procedure

- **1. Select K initial centroid vectors**
  - Random, or simple rules
- **2. Find the closest vectors for every centroid using Euclidean distances**
  - Construct a tentative cluster for every centroid
- **3. Update the centroid of every cluster**
- **4. Repeat 2and 3 processes until stop conditions**
  - Little change of centroids
  - Maximum iteration number

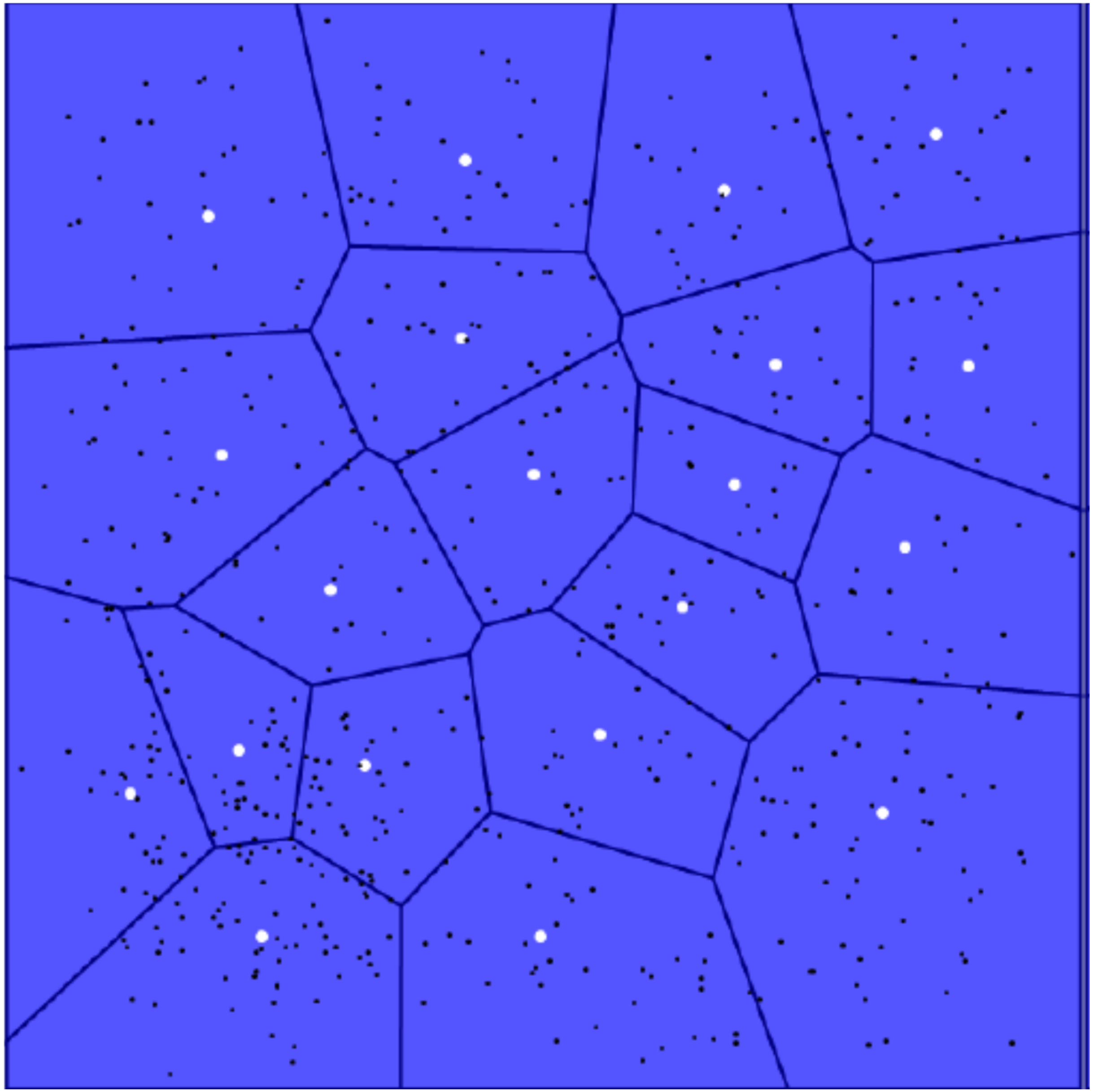
K-Means Clustering

1



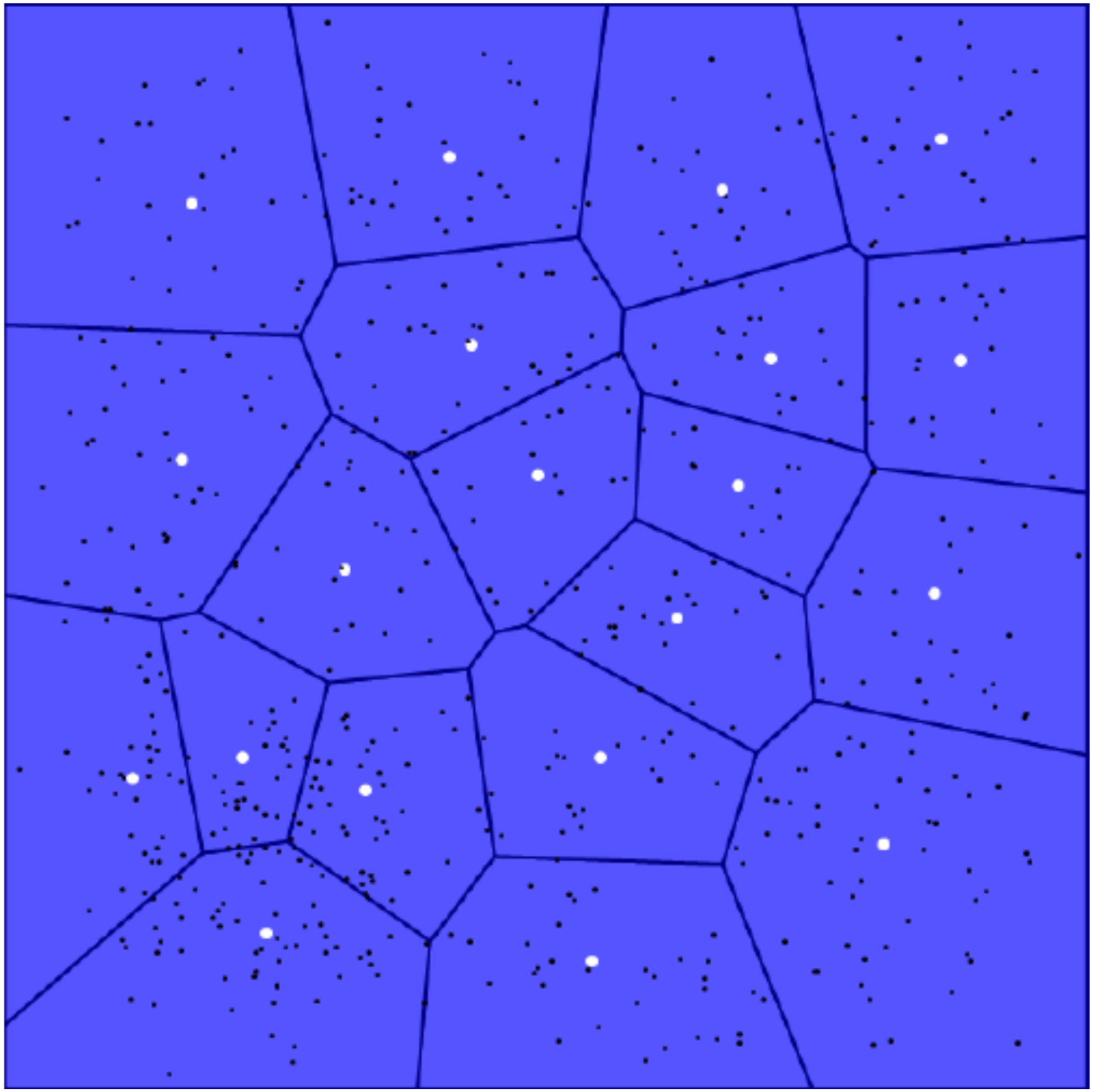
K-Means Clustering

2



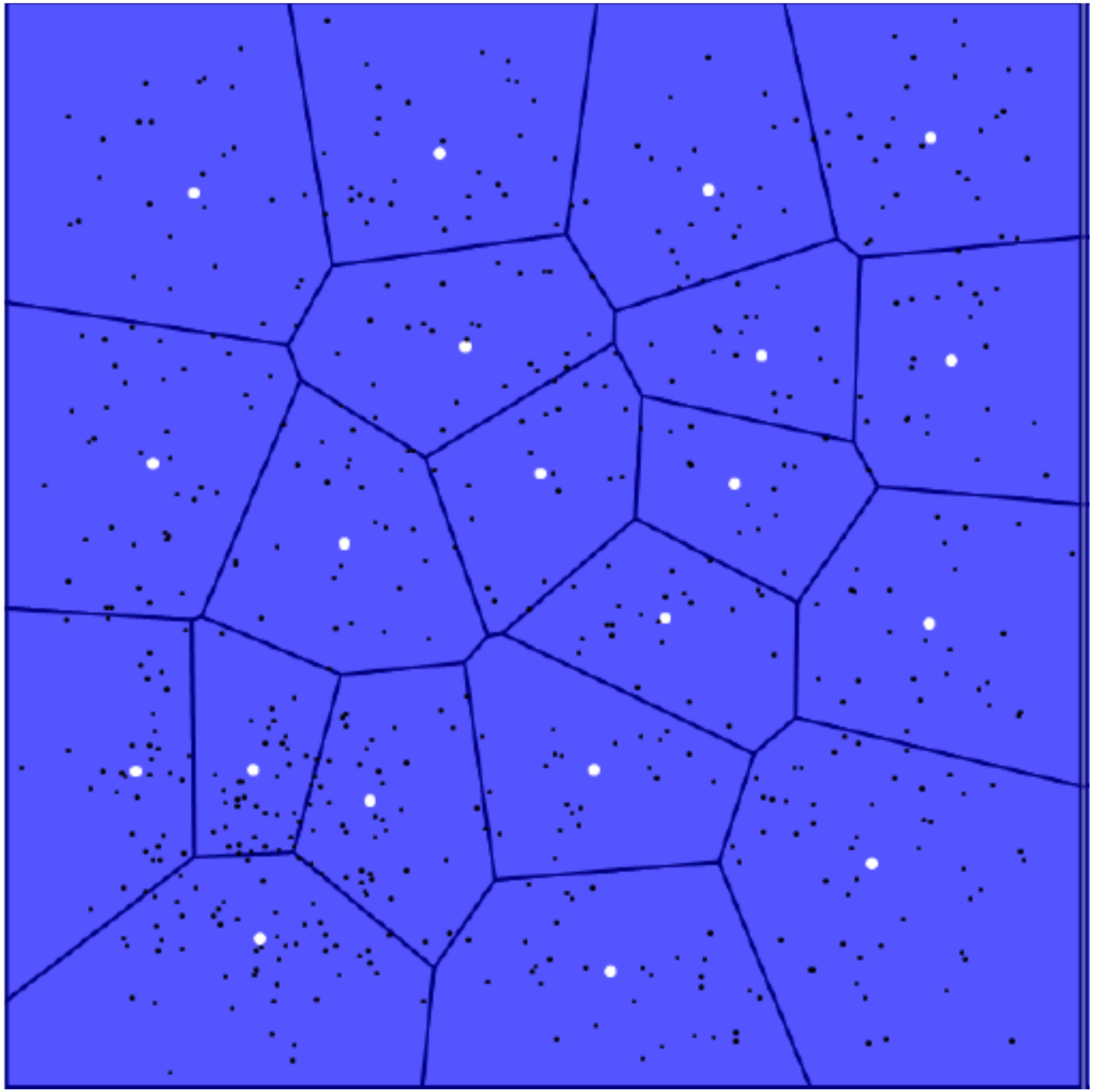
K-Means Clustering

3



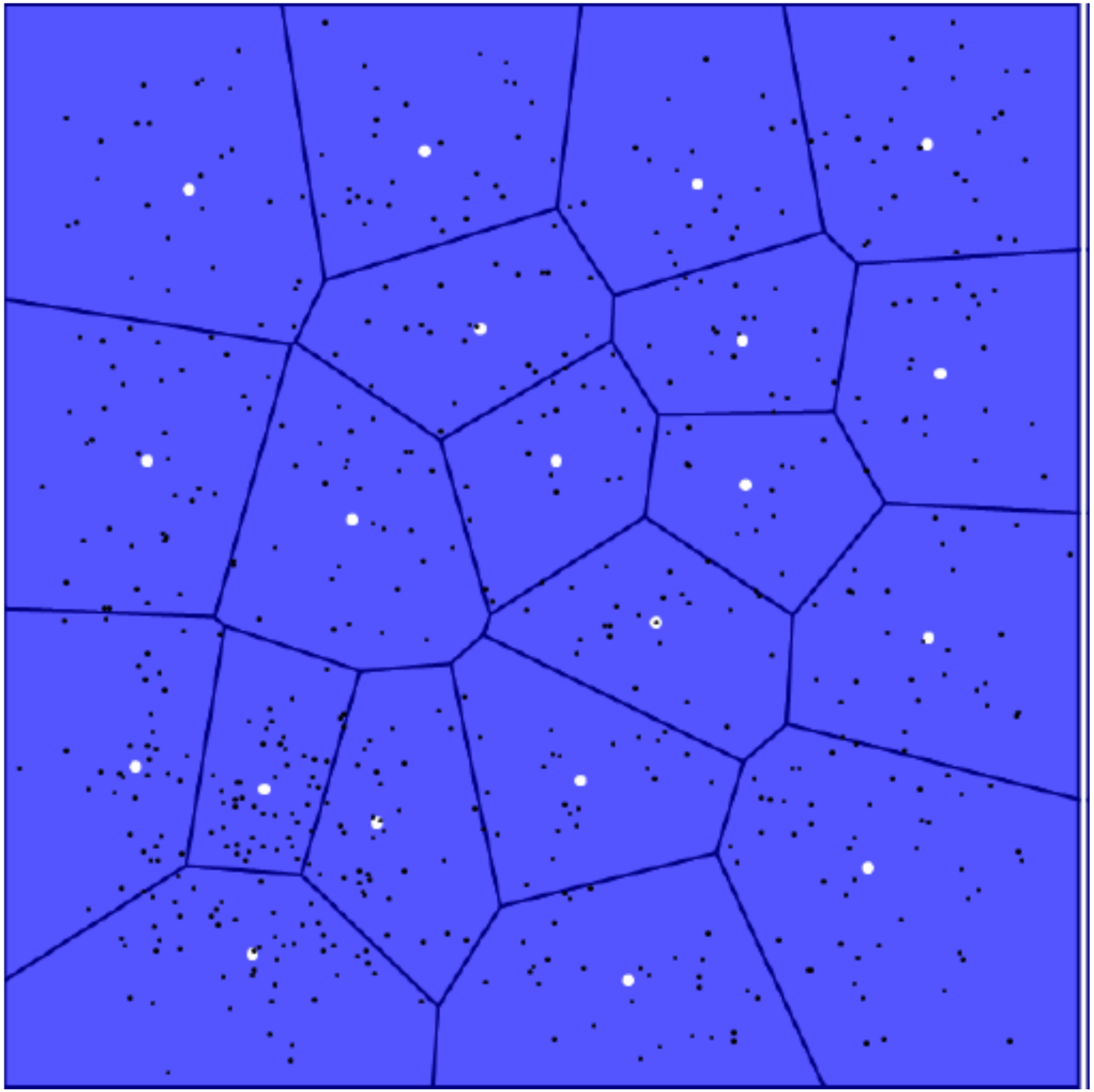
K-Means Clustering

4



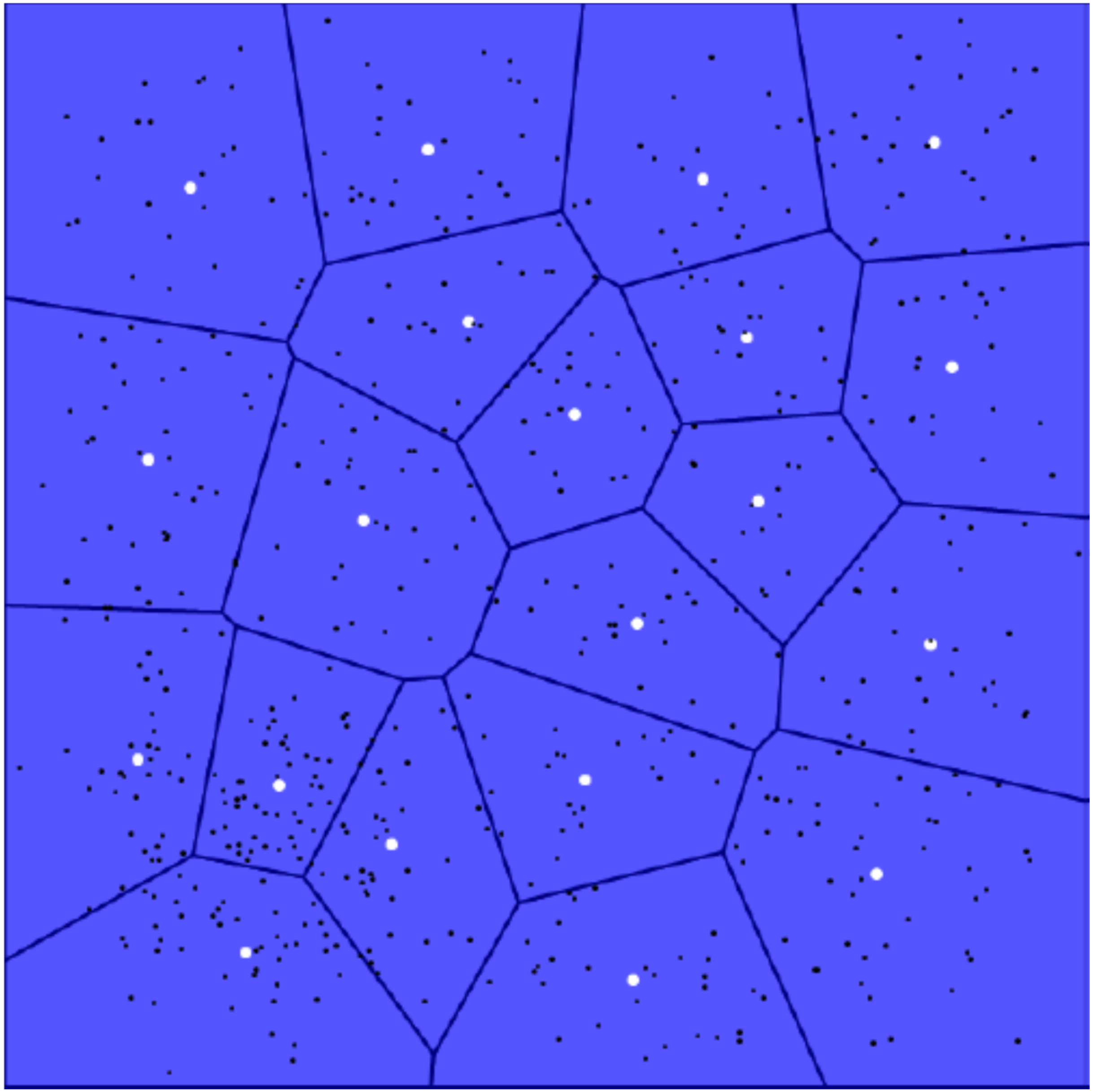
K-Means Clustering

5

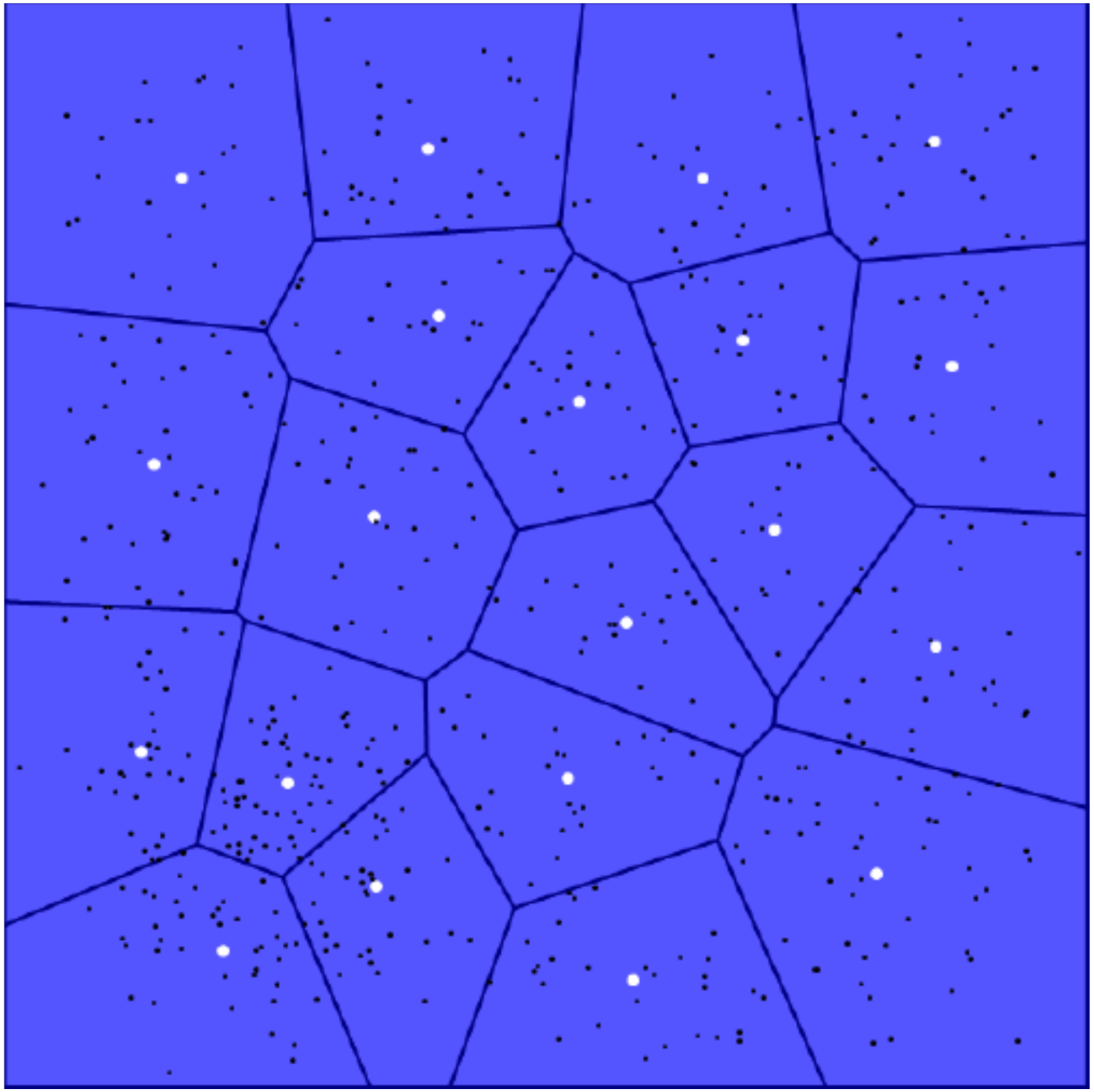


K-Means Clustering

6

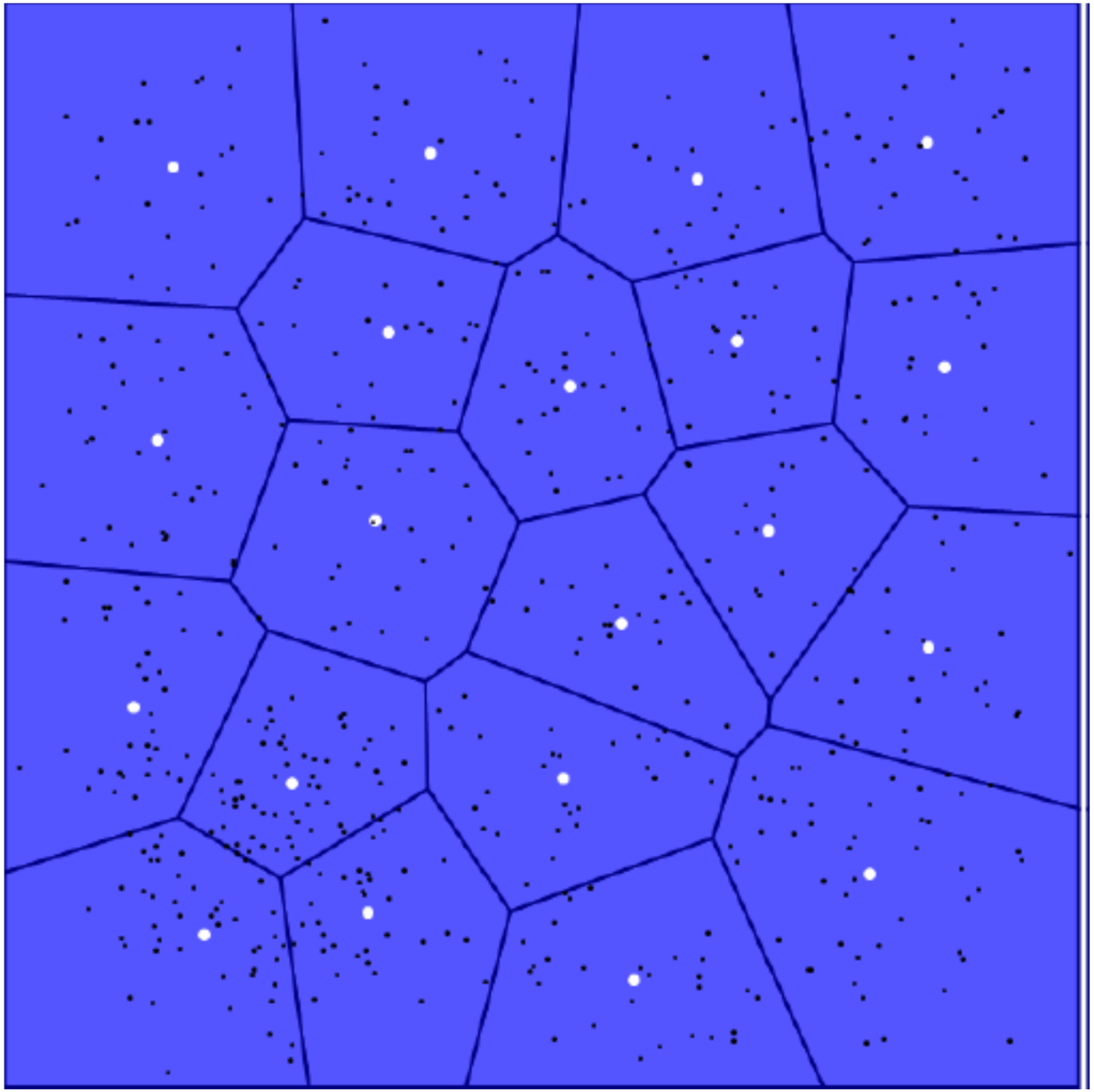


K-Means Clustering



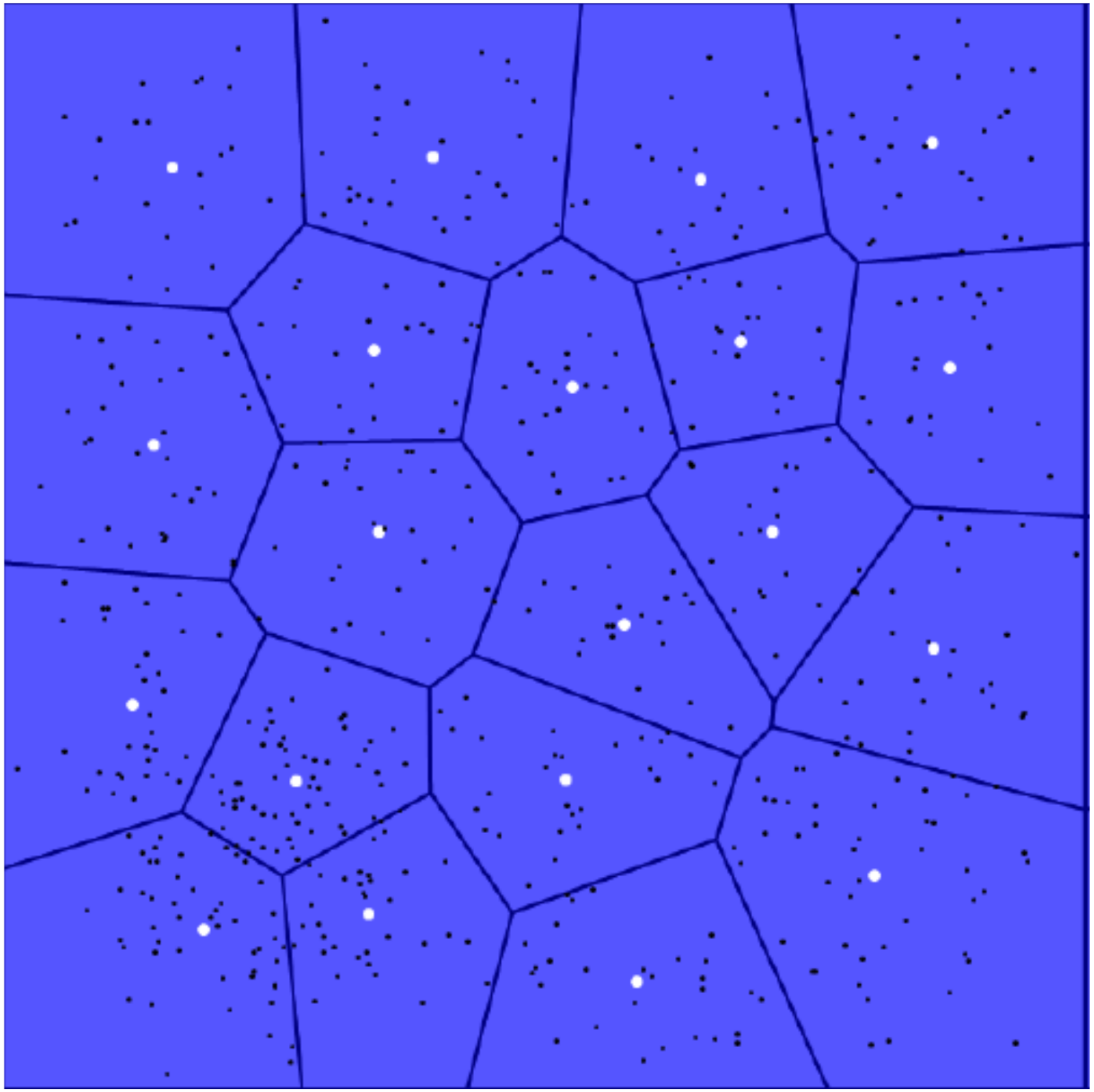
K-Means Clustering

8



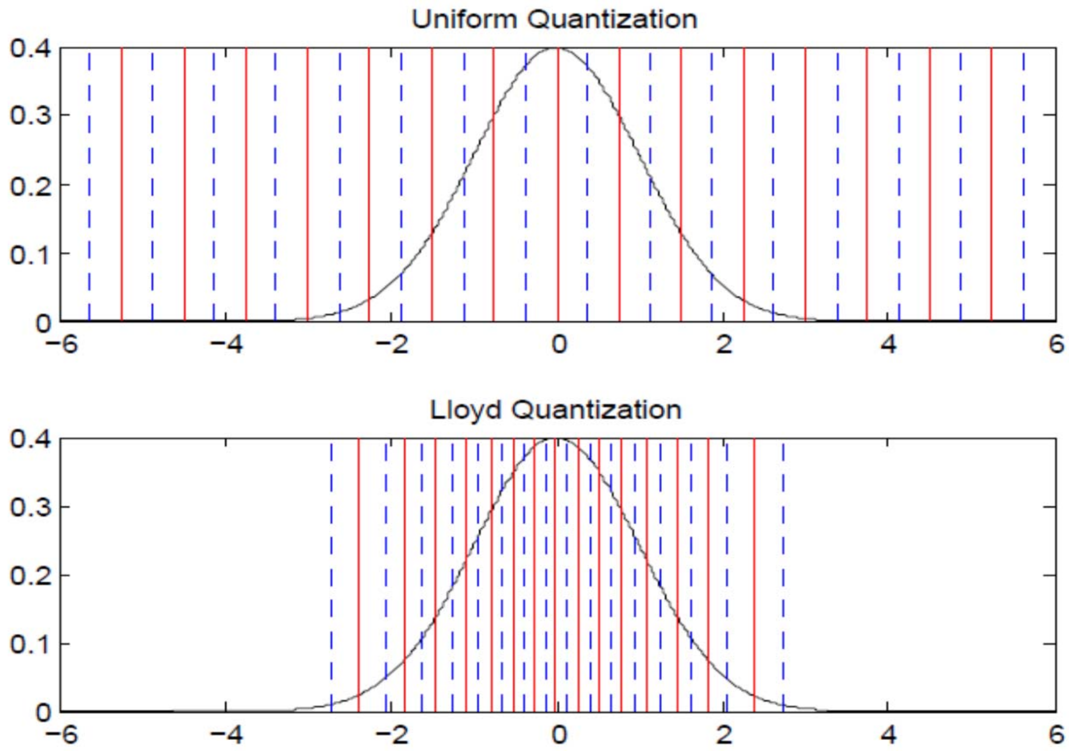
K-Means Clustering

9



Lloyd Max Quantization

- **K-means clustering**
  - Non-uniform quantization
- **Using probability of data samples**
  - Higher probabilities, smaller step sizes



Lloyd Max Quantization

❖ Mathematical formulation

Variables

- Bin-boundaries $\vec{L} = (L_1, L_2, \ldots, L_{n+1})$, with $\min \vec{x} = L_1 < L_2 < \cdots < L_n < L_{n+1} = 1 + \max \vec{x}$,
- and replacement values $\vec{p} = (p_1, p_2, \ldots, p_n)$.

Reconstruction (replacement) of x

$$q(x_i) = p_j, \text{ when } x_i \in [L_j, L_{j+1})$$

Energy minimization

$$E(\vec{L}, \vec{p}) = \sum_{j=1}^{n} \sum_{x_i \in [L_j, L_{j+1})} (x_i - p_j)^2$$

Lloyd Max Quantization

❖ Mathematical solution

- Optimal solution
- K-means clustering

$$\frac{\partial E}{\partial p_j} = \sum_{x_i \in [L_j, L_{j+1})} 2(x_i - p_j) = 0 \Leftrightarrow p_j = \frac{\sum_{x_i \in [L_j, L_{j+1})} x_i}{\#\{i \mid x_i \in [L_j, L_{j+1})\}} \quad (1)$$

$$\frac{\partial E}{\partial L_j} = 0 \quad \Leftrightarrow \quad L_j = \frac{1}{2}(p_{j-1} + p_j) \quad (2)$$

Lloyd Max Quantization

❖ Practical solution

- Iterative procedure
- Extend to vectors: K-means clustering

- Equations (1) are used to update the values $\vec{p}$:

$$p_j^{\text{new}} = ave \ \{x_i \mid x_i \in [L_j, L_{j+1})\}$$

- Equations (2) then yield new values for $\vec{L}$:

$$L_j^{\text{new}} = \frac{1}{2}(p_{j-1}^{\text{new}} + p_j^{\text{new}}) \ , \quad j = 2, \ldots, n$$

Lloyd Max Quantization

❖ Quantization Result (1)

- Image compression

Original: entropy=7.65

Quantized: entropy=4.56, PSNR=40.0 dB

Lloyd Max Quantization

❖Quantization Result (2)

Original: entropy=7.63

Quantized: entropy=3.76, PSNR=34.7 dB

Generalized Lloyd Max Quantization

❖ Mathematical formulation

- Using probability distribution
- Probability-weighted minimization

Energy minimization

$$MSE_m = \int_{t_{q,m}}^{t_{q,m+1}} (x - x_{q,m})^2 f_X \, dx$$

$$\frac{dMSE_m}{dx_{q,m}} = 0 \quad \Rightarrow \quad x_{q,m} = \frac{\int_{t_{q,m}}^{t_{q,m+1}} x f_X \, dx}{\int_{t_{q,m}}^{t_{q,m+1}} f_X \, dx}$$

$$\Updownarrow$$

$$t_{q,m} = \frac{x_{q,m+1} - x_{q,m}}{2}$$

Generalized Lloyd Max Quantization

❖ Quantization Result – 8 levels

Uniform quantizer

Lloyd Max quantizer

Generalized Lloyd Max Quantization

original

uniform

Lloyd

original

uniform

Lloyd

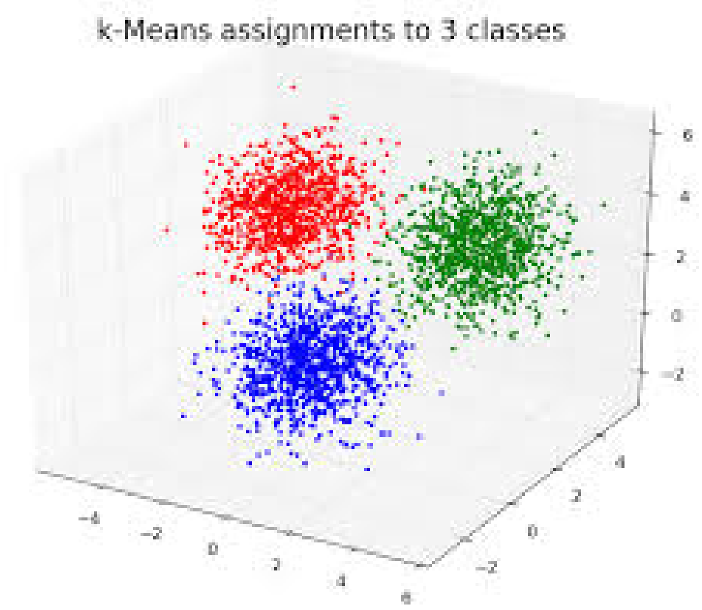
Comparison

❖ Unsupervised vector clustering

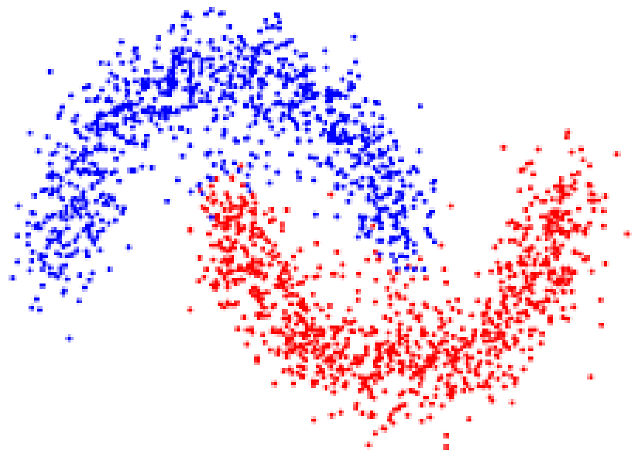
	Mean shift clustering	Lloyd Max clustering
Clustering Rule	• Mean shift convergence, • Boundary is not defined	• Equi-distance, • Voronoi boundry
Representative vector (clustering result)	• Convergent vector by mean shift process • Most probable vector	• Mean vector of cluster • Minimum distortion
# of clusters	• Not predefined, • Dependent on window size	• Predefined, • Dependent on rate-distortion
Classification	• Mean shift process, • Need all vectors of DB	• Nearest vector matching, • Need only mean vectors

Comparison

❖ **Classification results depend on the shapes of clusters.**

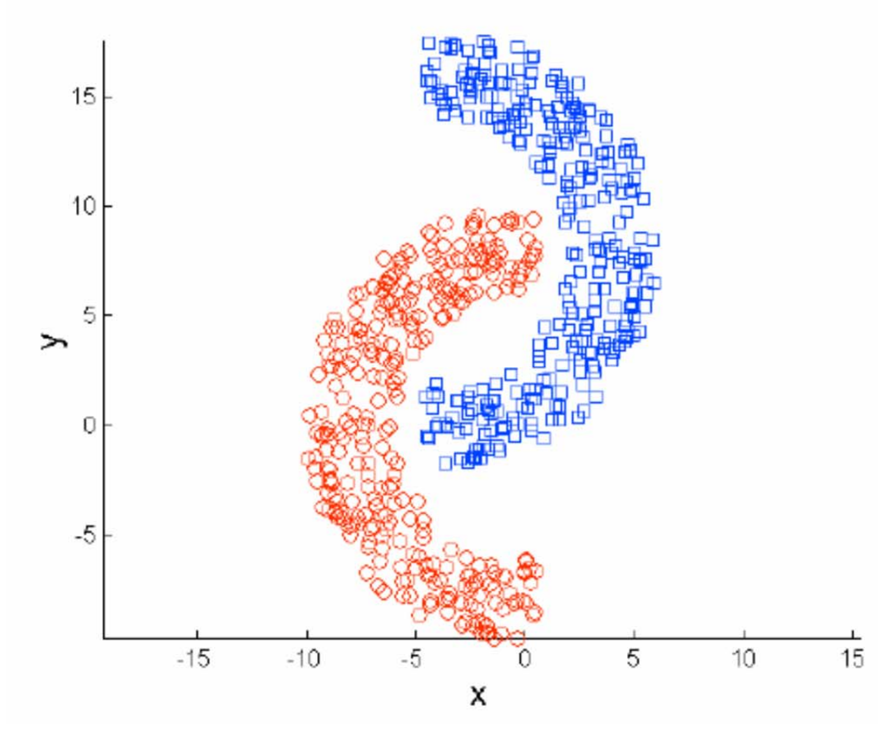


K-means clustering

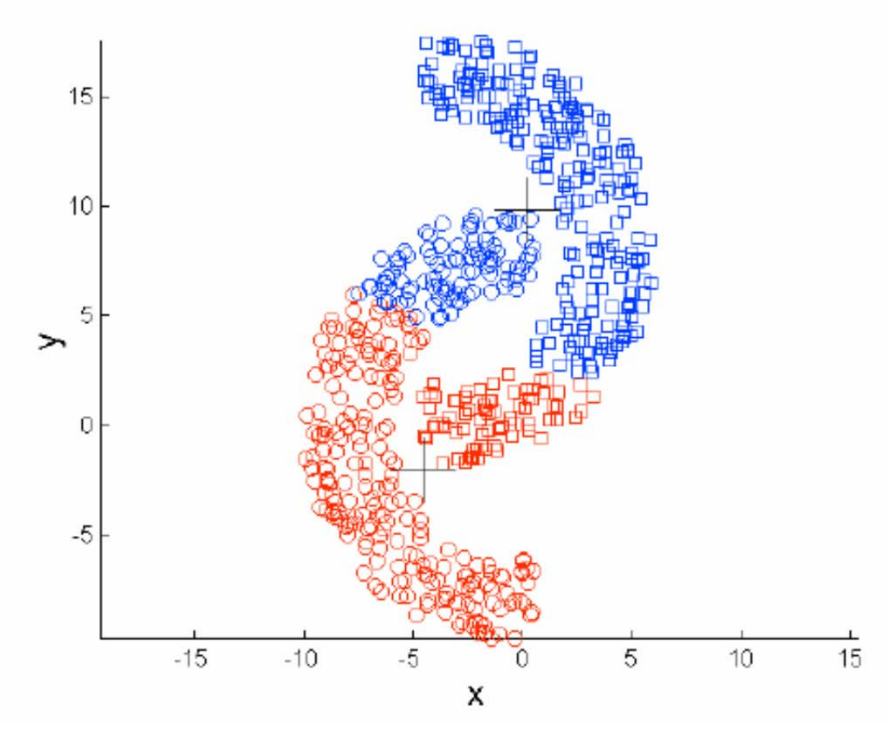


Suitable for mean shift clustering

Comparison



Mean shift clustering



K-means clustering