수치해석

Final Term Project

이 상 화

Term Project

- Subject: Pattern recognition
  - Pattern => vector
  - Vector => coefficients(DFT) , sample values (RGB)
- Simulation approach
  - Clustering of vector data and classification
  - Vectors are randomly generated with different Gaussian distributions
  - K-means Clustering
  - Test for randomly generated vectors

Clustering of Vector Data

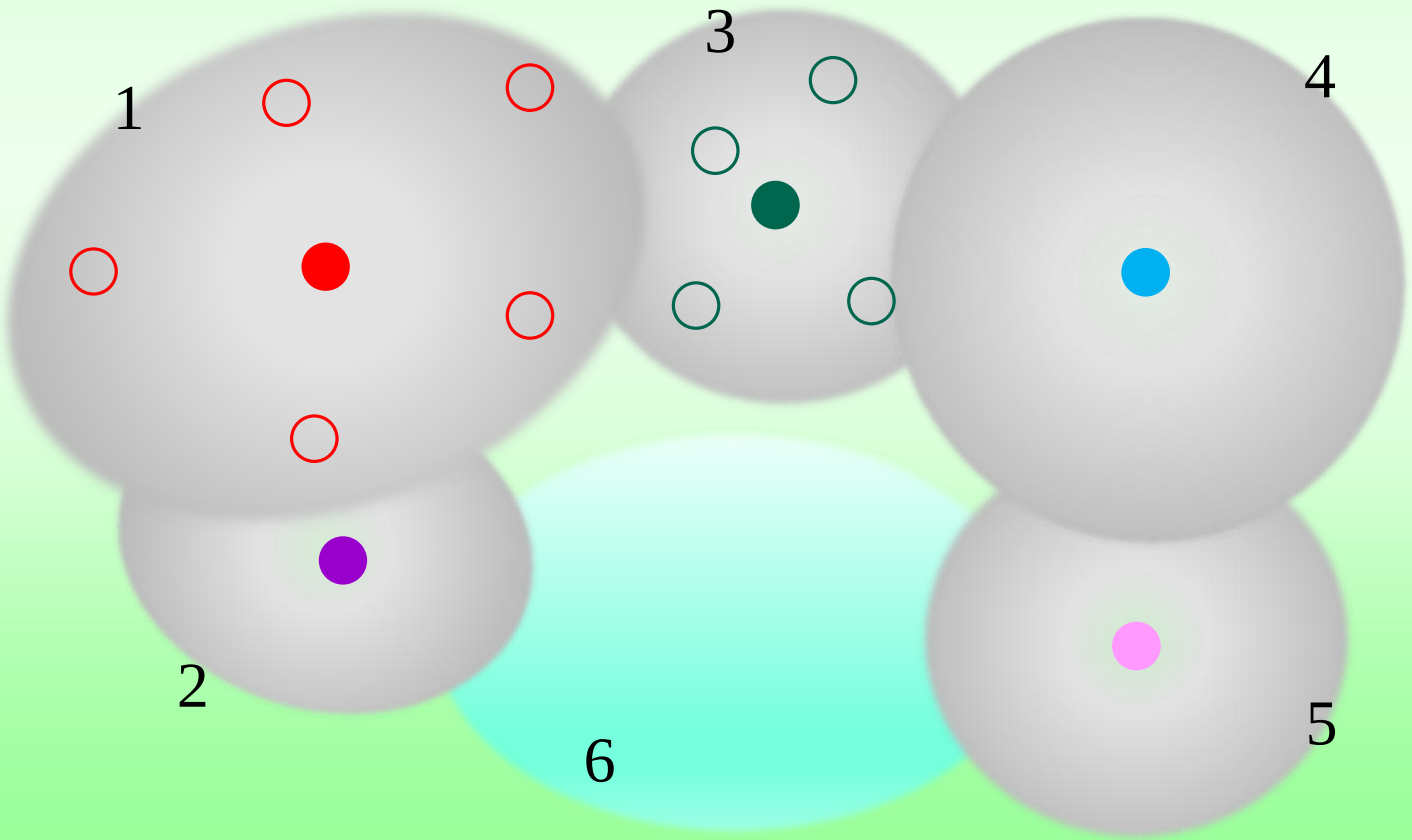
- Generate 5 classes (clusters) data using random number generation
  - 3D vector (X,Y,Z)
  - X, Y, Z are generated with different N(m, σ^2) respectively
    - Various values of m, σ^2 should be designed as you want.
    - EX: class#1: X = N(0, 1), Y= N(0, 1), Z= (1, 2)
  - 300 samples for each cluster
- Use K-means clustering for 3D vector data
  - Find the 5 mean vectors of clusters
  - Modeling the maximum distance for the clusters

Clustering of Vector Data

- Testing with new vectors
  - 3D vectors generated with the same distributions of 5 clusters
    - 100 samples for each class
    - Nearest neighbor with 5 mean vectors of clusters
  - 100 samples with different distributions
- Analysis of results
  - Recognition accuracy for each cluster
  - Analyze the results with respect to the Gaussian distributions that you used in generating the cluster samples
    - Distance of means, m
    - Size of variances, σ^2

Clustering of Vector Data

- m, σ^2 과 cluster 분포 특성



Submission

- Due: 2020. 12. 18 (Fri.), 23:00
  - Blackboard upload (pdf file upload)
  - Ppt: 16 pages maximum
- All HW Scores claim
  - 2020. 12. 21 ~2020. 12. 22 (20:00)
  - Email, message, mobile call