# Lecture 24
# Image Segmentation by Clustering
## ECEN 5283 Computer Vision

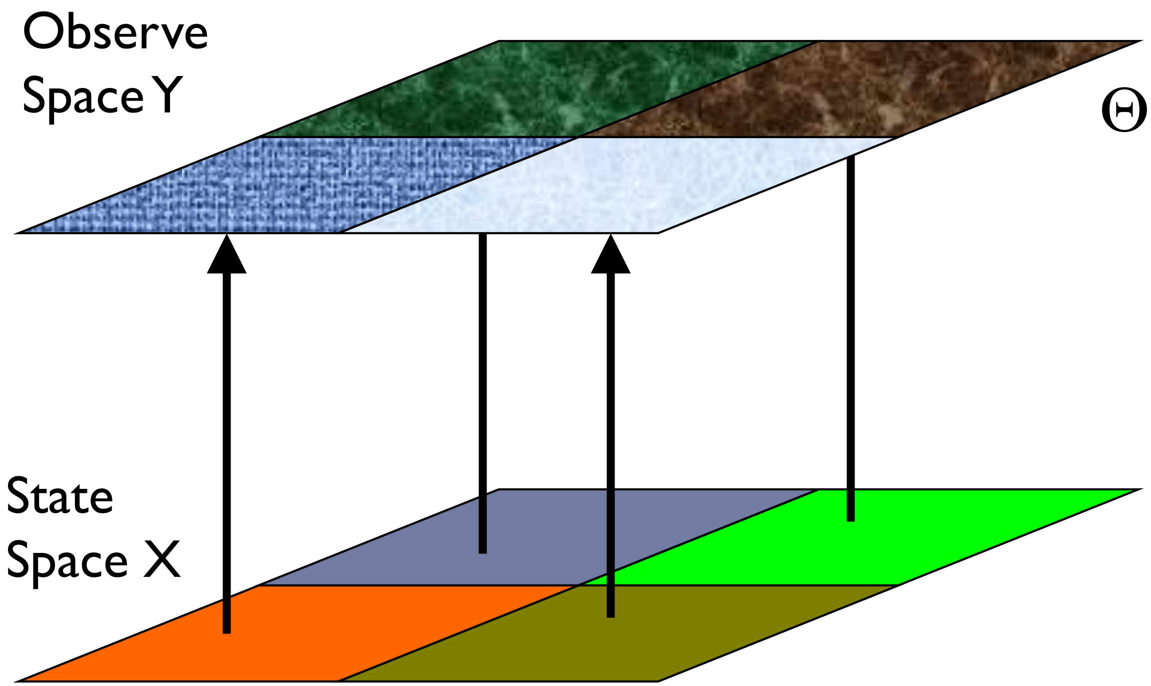
Dr. Guoliang Fan
School of Electrical and Computer Engineering
Oklahoma State University

# Goals

- To review the Expectation Maximization (EM) algorithm and compare EM with K-means.

- To introduce Project 4, image segmentation that involves texture analysis and clustering.

# Missing Data Problem




Parameter Space

$$\Theta = \{\alpha_i, \boldsymbol{\mu}_i, \Sigma_i \mid i = 1,2,3,4\}$$

$$p(x = i \mid y) \ \ \{i = 1, \ldots, k\}$$

(posterior probability)

$$\Theta^* = \arg\max_{\Theta} \log p(Y \mid \Theta) \quad \text{(log-likelihood of incomplete data)}$$

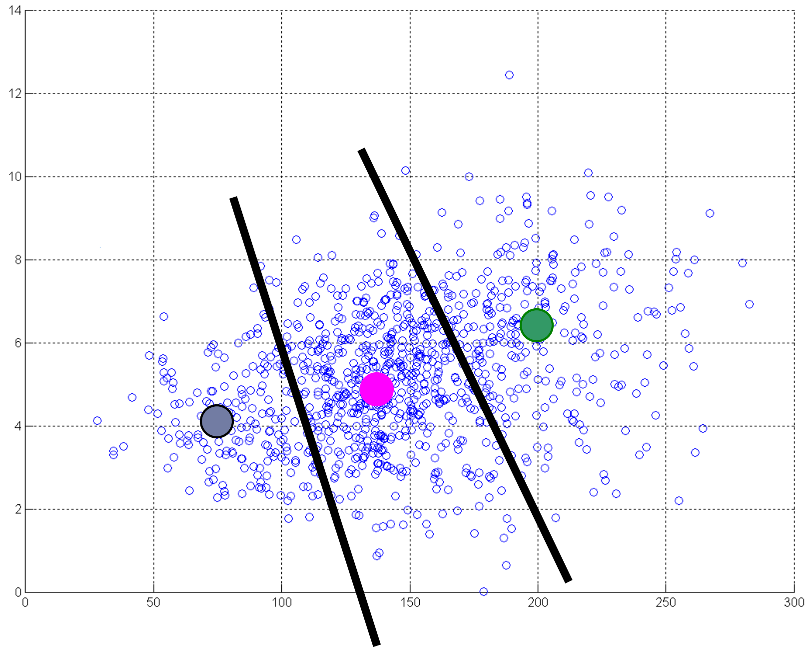
Can be computed but not easy to optimize

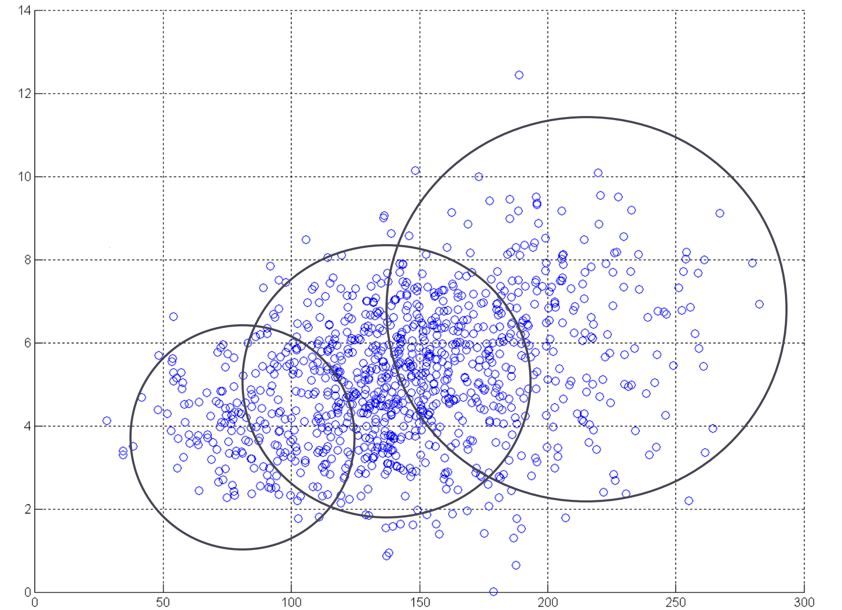
$$\Theta^* = \arg\max_{\Theta} \log p(X, Y \mid \Theta) \quad \text{(log-likelihood of complete data)}$$

Cannot be computed but can be optimized via a lower bound

# K-means vs. EM

$$\mathbf{Y} = \left(y_1, y_2, \ldots, y_N\right) \rightarrow \mathbf{X} = \left(x_1, x_2, \ldots, x_N\right)$$




$$x_i = \arg_{j\in\{1,\ldots,k\}} \min \left| y_i - \mathbf{c}_j \right|$$

$$x_i = \arg_{j\in\{1,\ldots,k\}} \max \, p(x_i = j \mid y_i, \Theta)$$

# K-mean vs. EM

- Both K-means and EM need to be initialized and involve two major steps during iteration.

| | K-mean | EM |
|---|---|---|
| **Initialization** | Initialize k means (cluster centers) $C^0 = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k\}$ | Initialize k Gaussian models that have equal weights. $\Theta^0 = \{\alpha_i, \mu_i, \Sigma_i \mid i = 1, \ldots, k\}$ |
| **Step 1.** | Assume the cluster centers are known, and classify each data sample to the closest cluster center. | Given the model parameters, estimate the missing data in terms of the posterior probability of each data sample. |
| **Step 2.** | Assume the allocation is known, and choose a new set of cluster centers. Each center is the mean of the points allocated to that cluster. | From the estimated missing data, to obtain the maximum likelihood estimate of the model parameters. |

# EM Algorithm: Initialization by K-means

- Randomly initialize k centers for the k-means multiple times.

$$C^0 = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k\}$$

- Run the k-means for each initialization until converge or with certain iteration number, and picks the best K-means result.

$$\Phi(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{k} \left\{ \sum_{x_j = i} \left| y_j - \mathbf{c}_i \right|^2 \right\}$$

- According to the class label of all samples $\{x_1, x_2, \ldots, x_N\}$, initialize the multivariate Gaussian models for the EM.

$$\alpha_i = \frac{\#(x_j = i \mid j = 1, \ldots, N)}{N}$$

$$\mu_i = \frac{\sum_{x_j = i} y_j}{\#(x_j = i \mid j = 1, \ldots, N)}$$

$$\Sigma_i = \frac{\sum_{x_j = i} (y_j - \mu_i)(y_j - \mu_i)^T}{\#(x_j = i \mid j = 1, \ldots, N)}$$

# EM Algorithm: Additional Issues

- Initialization:
  - Both EM and K-mean only converge to the local optimum.
  - Using K-mean is a practical way to initialize the EM algorithm.

- Iteration:
  - Iteration still the stop criteria is satisfied, e.g., no much change of the incomplete data log-likelihood

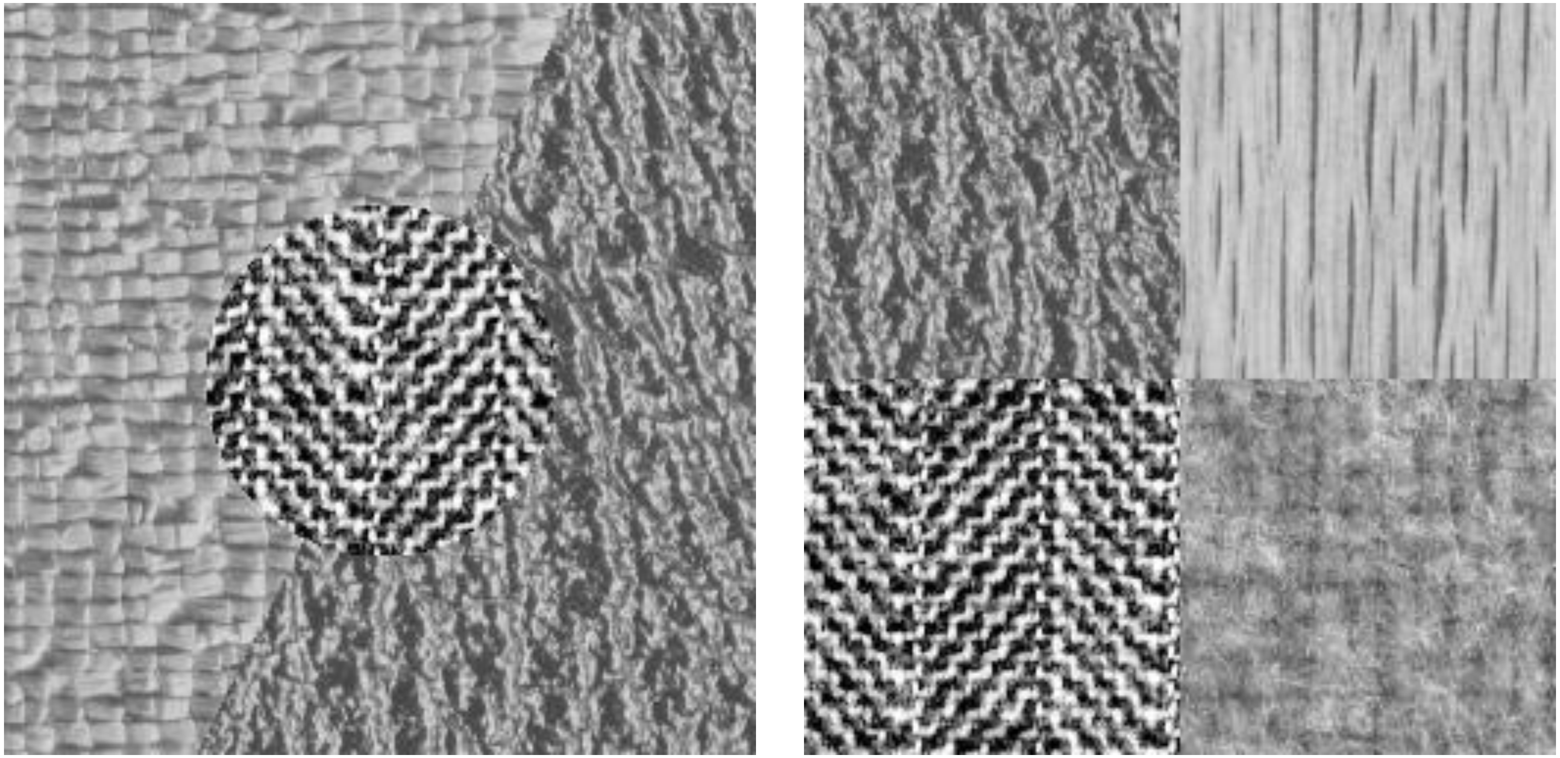
$$p(Y \mid \Theta^{(s+1)}) - p(Y \mid \Theta^{(s)}) < \Delta \ \rightarrow \log p(Y \mid \Theta^{(s+1)}) - \log p(Y \mid \Theta^{(s)}) < \Delta'$$

  - Or the iteration number can be fixed as a constant.

- Data classification
  - MAP classification $x_l = \arg_{m \in \{1,\ldots,g\}} \max \mathbf{I}(l,m)$

# Image Segmentation

- The goal of image segmentation is to segment an image into G homogeneous regions.



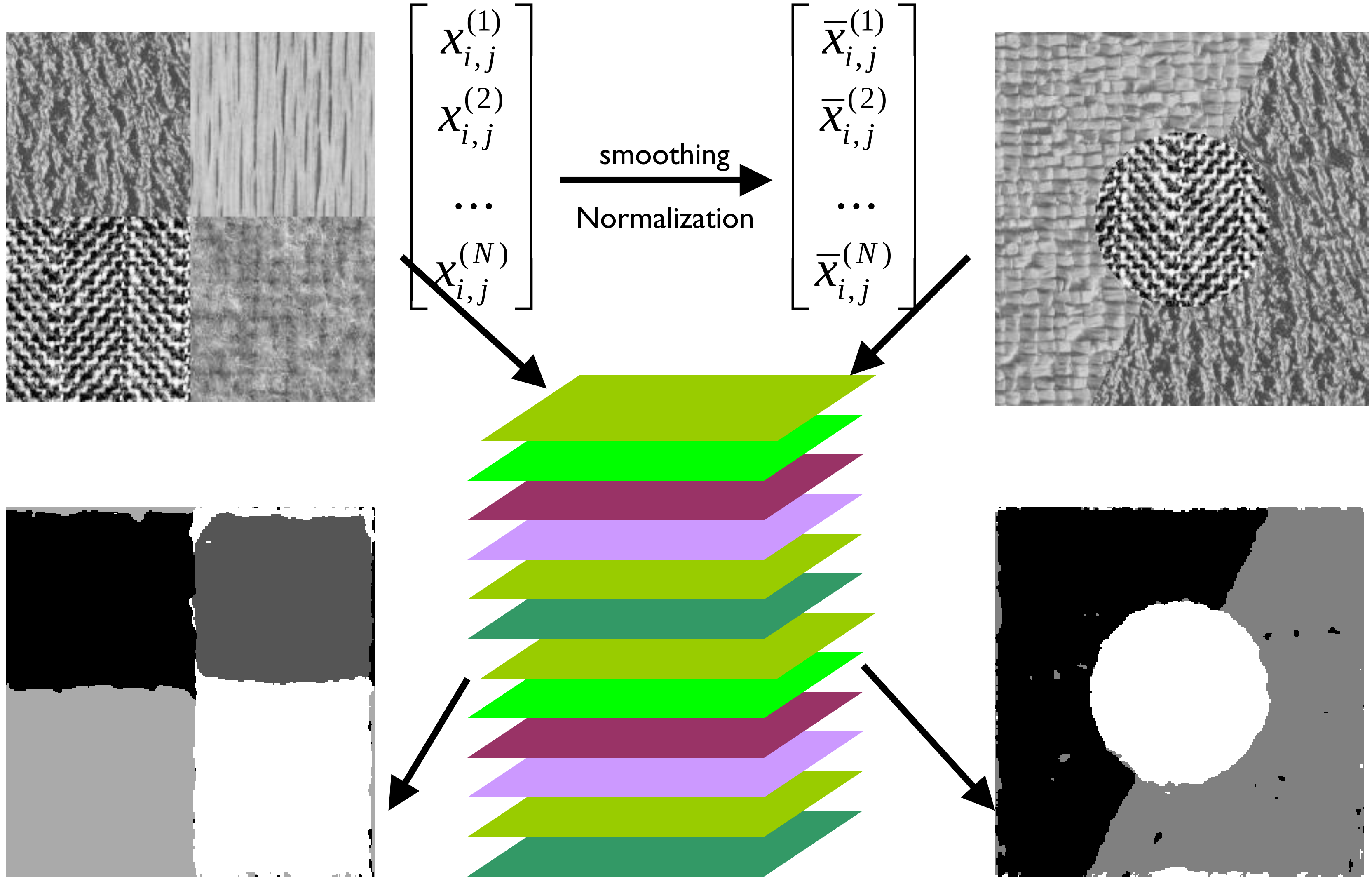
# Image Segmentation by Clustering

- Problem Formulation
  - An image has totally *N* pixels each of pixel is associated with a feature vector $Y = (y_l \mid l = 1,\ldots, N)$ , and the question is what is the class label for each pixel $X = (x_l \mid l = 1,\ldots, N)$?

- Assumption
  - It is assumed that feature vectors of G classes are defined by G multivariate Gaussians $\Theta = \{\alpha_m, \theta_m = (\mu_m, \Sigma_m) \mid m = 1,\ldots, G\}$ that defines the distribution of feature vectors and the ratio of each class.

- Solution:
  - Image segmentation becomes a clustering problem that can be solved by K-means or EM.

# Feature Extraction: Gabor Filtering



# Segmentation Evaluation

```
function Per=accuracy(Truth,Result,Num);
% Truth is the groud-truth map, Result is the obtained map, Num is the number class
% Per is the segmentation accuracy

[X Y]=size(Truth);

Z=zeros(256,256);

for i=1:X
    for j=1:Y
        p=Truth(i,j)+1;
        q=Result(i,j);
        Z(p,q)=Z(p,q)+1;
    end
end

T=sum(max(Z));
Per=T/X/Y;
```

Ground Truth

| 1 | 1 | 1 | 2 | 2 |
|---|---|---|---|---|
| 1 | 1 | 1 | 2 | 2 |
| 3 | 3 | 4 | 4 | 4 |
| 3 | 3 | 4 | 4 | 4 |
| 3 | 3 | 4 | 4 | 4 |

Clustering Result

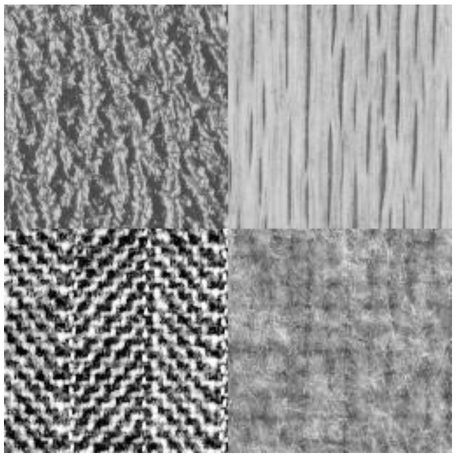
| 3 | 3 | 3 | 4 | 4 |
|---|---|---|---|---|
| 3 | 1 | 2 | 4 | 4 |
| 2 | 2 | 1 | 1 | 1 |
| 2 | 3 | 1 | 3 | 1 |
| 2 | 2 | 1 | 1 | 1 |

| | | | |
|---|---|---|---|
| 1 | 1 | 4 | 0 |
| 0 | 0 | 0 | 4 |
| 0 | 5 | 1 | 0 |
| 8 | 0 | 1 | 0 |

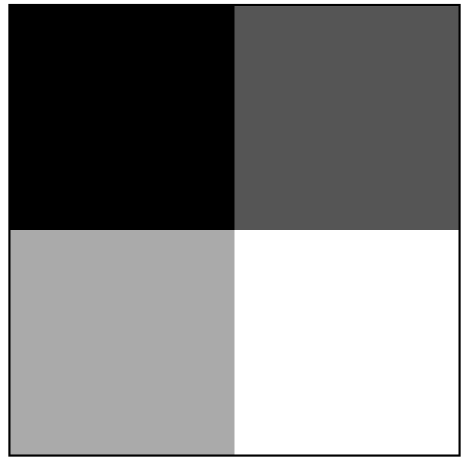
| 8 | 5 | 4 | 4 |
|---|---|---|---|

Sum=21

# Texture Segmentation using K-mean and EM Algorithms (Mosaic A)



Mosaic A



Ground truth segmentation

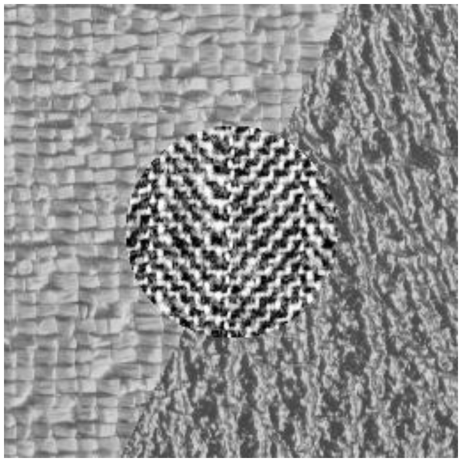
The 5-scales and 8-orientation Gabor filtering is used for texture feature extraction.

K-mean clustering result

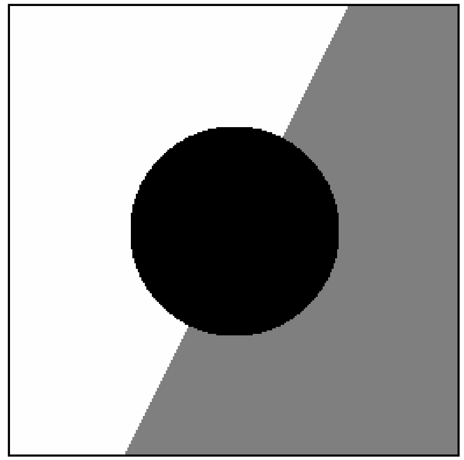
EM algorithm result

The dimension of feature vector is 5x8=40.

# Texture Segmentation using K-mean and EM Algorithms (Mosaic B)

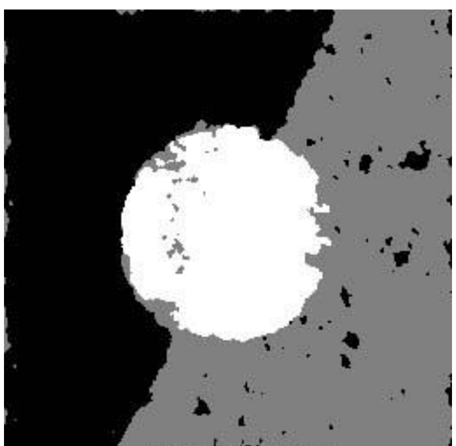


Mosaic B

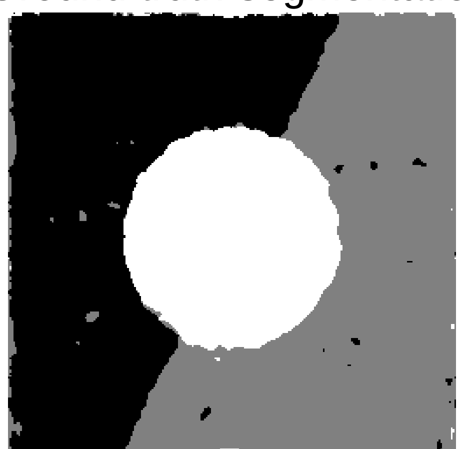


Ground truth segmentation

The 5-scales and 8-orientation Gabor filtering is used for texture feature extraction.



K-mean clustering result



EM algorithm result

The dimension of feature vector is 5x8=40.

# Requirements of Project 4.

- Implement two clustering algorithms for image segmentation
  - For K-means, try to use a random method to find a good initialization.
  - Discuss the effect of initialization to each of them.
  - Compare different initializations for EM-based image segmentation.

- Some plots to show for each clustering
  - Plot the incomplete data likelihood $p(Y|\Theta)$ vs. the iteration number.
  - Plot the segmentation accuracy vs. the iteration number

- Additional requirements
  - Create a video to show the segmentation map for each iteration.
  - Find a couple of color images and add RGB (3-channel colors) or XY (pixel coordinates) feature along with Gabor features for segmentation. Discuss your findings.