Project 5: Color Analysis and Image Segmentation

Each group should

- Submit one report per team containing procedures and results of this project by 11:59 pm, Feb. 29.
- Include a thoughtful, reflective paragraph from each of the team members summarizing what has been learned.
- Attach your well-commented code to the report.

Part A – Strawberry Color Analysis

In this assignment, you will analyze strawberry pixels and non-strawberry pixels in different color spaces and identify the best color component(s) to perform strawberry detection.

1. To build the strawberry-color model for strawberry detection, use all 20 images from the folder PartA, and manually draw a polygonal region of interest (ROI) around the strawberry to include as many strawberry pixels as possible in the ROI while leave non-strawberry pixels outside the ROI. Create a set of 20 mask images in which 1's indicate strawberry pixels and 0's indicate non-strawberry pixels. Label and save those mask images. Since the 20 images are of different sizes, you might want to resize images with high resolutions to keep all images at around the same size of the smallest image.
2. Examine color distributions for strawberry and non-strawberry pixels in RGB color space. For this, construct a histogram of all strawberry pixels from the images in folder PartA with the help of the mask images and compare it with the histogram of all non-strawberry pixels from the same set of images. You need to construct a total of 3 sets of histograms. Examine color distributions between strawberry and non-strawberry pixels. What do you observe?
3. Experiment with two other color spaces- normalized rgb and HSV (Here since hue is defined on a ring, it might be advantageous to represent strawberry hue in a continuous range, how can you do it?). Again, examine color distributions for the strawberry and non-strawberry pixels in those color spaces.
4. What is/are the best feature(s) for characterizing strawberry-color? Why?
5. Extra credit: Examine additional color spaces such as YCbCr, and Lab.

Part B – Strawberry Finding/Counting Using K-Means and EM.

In this assignment, you will perform image segmentation using both the K-Means and EM algorithms. You will also try to find and count the strawberries in the image based on the segmentation results.

1. Run K-Means and EM clustering algorithms using RGB values of a color image. Perform image segmentation on images in folder PartB. Display and discuss the segmentation results with different initializations and different values of K. Show the best result.
2. Based on the segmentation results in the last step, develop a strawberry finding and counting algorithm. Place a white bounding box around the detected strawberry.

3. Select the best feature (or features) from Part A and develop a strawberry counting algorithm. Compare the result with that in step 2.
   Note: You can use the MATLAB built-in functions for K-Means and EM implementation. You can earn extra credit if you write your own code.

Part C – K-Means and EM Clustering on Synthetic data

In this assignment, you will evaluate and compare the effectiveness of the two algorithms on clustering synthetic data.

1. Generate several (3-5) clusters of 2D data points by using *mvnrnd*(). Data in each cluster follow a multivariate Gaussian distribution. Vary the means and covariance matrices of the distributions as well as the number of points in each cluster.
2. Run K-Means on the synthetic data generated in Step1. Does it always yield the optimal configuration? Discuss the results.
3. Repeat Step 2 for EM clustering.
4. How do the two data clustering techniques compare?