Lab : LBG

Goal:

The goal is to study the a part of the Lloyd optimization algorithm that is known also as the k-means algorithm (which is also an unsupervised classification method). Note that the Lloyd optimization technique is used by the LBG algorithm described in the lecture.

Introduction:

The GNU/octave code presents 2 examples of unsupervised classifications by using the k-means method.

First a simple case is considered which synthetic data.

The second case achieves a colored image partition according to its hue and saturation components. The name of the main function script is "Main_kmeans.*m*", you can choose the name of the image to process, and the number of regions for the classification. The algorithm, from an initial random partition, iterates to converge, then its displays:

- the points cloud in the hue/saturation space, and the classes,
- the original image with the contours of the obtained regions.

To do:

First part (with synthetic data):

- Launch several times the algorithm with different random datas (the number of classes will stay equal to 2, but the level of noise and the "gap" between the 2 point clouds will be modified). Answer to these questions (and explain):
  - Does the k-means succeed to identify the 2 classes?
  - What does happen if the number of region is set as higher than 2?
  - A final class is characterized by its centroïd, that is appropriate here?
  - Does the final classification depend on the initialization?

Second part (with the image):

- Launch several times the algorithm with different number regions and images. Answer to these questions (and explain):
  - What is the criterion optimized by the k-means? Is this criterion OK to find the good number of regions?
  - A final region is characterized by its centroïd, that is appropriate here?
  - Is the optimization local or global?

General question: What are the new steps added by the LBG algorithm and why?

.