Department of Computer Science
Digital Image Processing

N. Komodakis, Course instructor

Autumn 2022

Unsupervised Image Segmentation Algorithms

The objective of this work is the implementation of digital image segmentation and the evaluation of the parameters that determine their performance. By the term segmentation in the field of digital image processing we refer to the grouping of parts of the image into groups (classes) using some common feature, e.g. common color characteristics. The ultimate goal of this process is to exploit the segmented regions for the semantic analysis of the image content, at the level of objects, with the aim of highlighting correlations between them.

A good segmentation/partitioning algorithm should be able to group the data into as homogeneous subsets as possible. From the algorithms available in the literature, we will focus on those that possess the property of unsupervised learning.[1] In more detail in the present assignment we will be working with the unsupervised clustering and segmentation algorithms of (a) k-means[2], and, (b) Mean-Shift[3].

In this assignment you will be using RGB and depth data from the NYU Depth dataset. In the image 1, you will find an example showing the two types of data (image, depth) and the result of the segmentation with the Mean- Shift algorithm. The data can be found here[4].

a) Input Image b) Depth Map c) Segmentation Result

Figure 1: Input data and segmentation result using the Mean-Shift algorithm.

Part A: comprehension questions (30 credits)

Give short answers to the following questions about the characteristics of the algorithms. Rely on the course lectures and the useful links given to you in the lecture to understand the structure and operation of algorithms.

1. What are the advantages and disadvantages of the two algorithms?
2. What are the elements of the nearest neighbor algorithm that should be properly defined because they determine its performance, and what are those of the mean shift algorithm?

[1] https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning
[2] https://towardsdatascience.com/k-means-explained-10349949bd10 K-means explanation
[3] https://ml-explained.com/blog/mean-shift-explained
[4] https://drive.google.com/drive/folders/16H_SJ_roOOlI93cXw_96FRM-wPngrrL2?usp=share_link

Part B: segmentation algorithm implementation (40 credits)

In this task, to segment the set of images you will use (a) the MATLAB out-of-the-box implementation of the k-means algorithm, by calling the function *kmeans()*, and (b) a custom implementation of the mean shift algorithm, which can be found in the provided files in this footnote[5]. Specifically, you will segment each image by exploiting 3 data types: (a) color, (b) color + spatial coordinates, and (c) depth.

Specifically, for each image and each data type you will do the following steps. First, you will read your image (*imread()*), and load the array with the depth data(*load()*). Please don't forget to transform the image data to double precision with (*im2double()*).

Algorithm k-means (color, depth, color + spatial coordinates)

- partitioning for number clusters k = 3, 5, 7, and presenting the results in a common figure.
- for the set cluster k that gives the best result choose a different comparison metric and run the algorithm again. Specifically, check for 'cityblock', 'cosine', 'sqeuclidean' and present the results in a common figure.

Clarification: In the case of the 'cosine' metric, you should add +1 to its value each pixel. Can you think why such a thing is needed? (BONUS 5 credits)

Algorithm Mean-shift (color, color + spatial coordinates)

- partitioning for bandwidth (bw) = 0.2, 0.35, 0.6, and presenting the results in a common figure.
- for the bandwidth (bw) that gives the best result choose a different comparison metric and run the algorithm again. Specifically, check for 'cityblock' and display the results in a common figure.

To display the segmentation result you will use the function *labeloverlay()*. Useful examples of everything you'll need for the exercise can be found in the MATLAB documentation , so we encourage you to consult them!

Clarification: to implement the k-means algorithm, you will use the function *k-means()* and NOT the function *imsegkmeans()*.

Part C: analysis of results (30 credits)

Give short answers to the following questions.

1. Which type of data leads you to the best results and why?
2. Which algorithm gave you the best (visually) result? why do you think something like this happened?
3. Do the color characteristics of the objects present in the image and in relation to the background pattern play a role in the result of the segmentation?
4. Will the simultaneous use as input to the segmentation algorithms of several types of data (color, spatial coordinates, depth) improve the performance of the segmentation, or is it better not to increase the complexity of the data? Justify your answer

[5] https://drive.google.com/drive/folders/16H_SJ_roOOlI93cXw_96FRM-wPngrrL2?usp=share_link

BONUS(25 credits)

1. (5 credits) Can we transform the input data in some way to emphasize some of its features appropriately, so as to direct the segmentation algorithms appropriately towards the desired result?
   Hint: think about the things you implemented in the previous exercises.

2. (20 credits) Implement one of the ways you thought of. Read the image, transform it appropriately and apply one of the two segmentation algorithms and compare the result with the best result you got in the main part of the exercise.

Helpful Functions

- To display the depth data with (*imshow()*) , you should map the depth data to grayscale with (textitmat2gray()). Alternatively, to display the depth data combined with the original color image you can use the command:

  surf(depth data, image data, 'FaceColor', 'texturemap', 'EdgeColor', 'none')

Submission Instructions

- Your report should be submitted in *.pdf* file format and only through the *e-learn* platform . Otherwise it will not be graded. In both your report and your code you list your details.

- Your code should have comments in important sections as well as in each *figure* you should include the relevant title for what you are illustrating.

Problems and Questions

For any personal questions about the exercise, please contact the TAs, at kbacharidis@csd.uoc.gr, croditakis@csd.uoc.gr and for questions of a general nature in the course list.