ENGN6528 Computer Vision – S2, 2020
Assignment-1

Objectives:

This is Assignment-1 of ENGN6528. The goal is to help you familiar with the concepts covered in the first 6 weeks of the lectures which include image formation, binary image analysis, edge/corner detection, image segmentation, clustering, and PCA.

Special Notes:

1. For this assignment, we will provide a tutorial but no lab sessions. You're required to complete it by yourself.
2. Your assignment will be marked based on the overall quality of your report. The report is to be uploaded to the Wattle site before the due time. The deadline for this assignment is 4th October 2020.
3. Your submission includes the report in PDF format as well as the code that generates the experimental result. Please submit a single zip file named <uid>-assignment-1.zip.
4. This assignment will contribute to 20% of the final marks.

Academic Integrity

You are expected to comply with the University Policy on Academic Integrity and Plagiarism. You are allowed to talk with / work with other students on lab and project assignments. You can share ideas but not code, you should submit your own work. Your course instructors reserve the right to determine an appropriate penalty based on the violation of academic integrity that occurs. Violations of the university policy can result in severe penalties.

Assignment-1 Tasks

Task-1: Image Filtering (9 marks)

1. Consider the filter f = [-1,2,-1] and the 1D image I = [0,1,2,4,5,10,3,3,3,2]. What is the result of f * I where * denotes the convolution? Pad the image with zeros at the boundary if necessary. [1 mark]

2. Given the following 5 x 5 image with the indicated grayscale intensities, what is the output of median filtering using a 3x3 window? Please make sure the output is the same size as the input by padding the input image with zeros. [2 marks]
   Hint: Median of a sorted array of length n, is the middle element, ie, the element at the position n/2 of this sorted array.

3	7	5	7	9
2	4	7	8	0
5	1	8	100	3
8	20	8	5	7
6	3	5	2	6

3. Bilateral filtering is a method to smooth an image while preserving edges. Briefly explain the main idea of bilateral filtering in your own words and how it preserves the edges in an image. [2 marks]
   Hint: Read the attached paper and/or corresponding section in Szleski's book.

4. Implement your own bilateral filter (*my_bilateral_filter()*) in your preferred programming language. Test it by varying the standard deviation of each kernel on the attached image and one of your face images. Report the results and the parameters that yield the (visually) best results. [4 marks]

Task-2: Hough Transform (8 + 2 marks)

1. Hough transform can be used to detect shapes in a given image. Please read the attached code for line detection using Hough transform. Modify the code to detect all the lines that have votes greater than a given threshold and test it on the given image by varying the threshold. [2 marks]

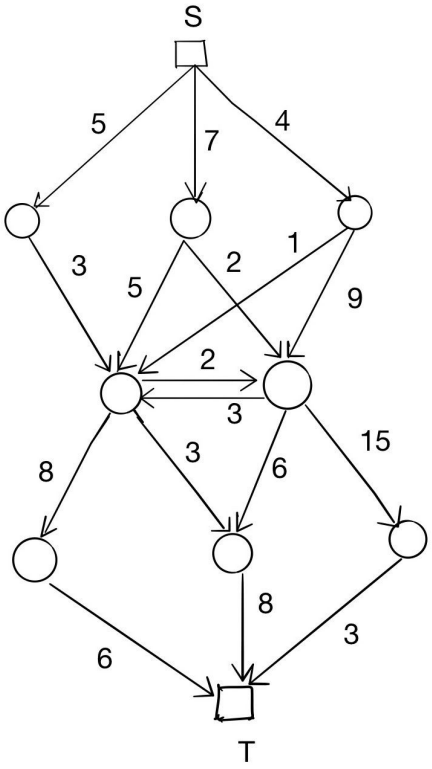
2. Briefly explain how one can use Hough transform to detect circles. You should provide necessary details such as the parametrization of a circle, the dimension of the accumulator, etc. [2 marks]

3. Please write your own *hough_circle()* function to detect circles and test it on the attached image and one of your frontal face photos by varying the threshold for number of votes [4 marks]. **Hint:** Take a photo without the eyeglasses.

4. [**Extra credit**] Tune the parameters in your code to detect only the eyeballs in your photo using your *hough_circle()* function. Report the final result along with the parameters that yield this result. [2 marks]

Task-3: MRFs (8 + 4 marks)

1. An st-graph is given below. Here, squares represent the source (S) and target (T) nodes, the circles represent the internal nodes, and the numbers next to each edge represent the edge weights. What is the min-cut partition of this graph and what is the cost of this partition? [2 marks] Briefly explain the algorithm you used to obtain the solution. [1 mark]



2. A multi-label pairwise MRF energy can be written in the following form:

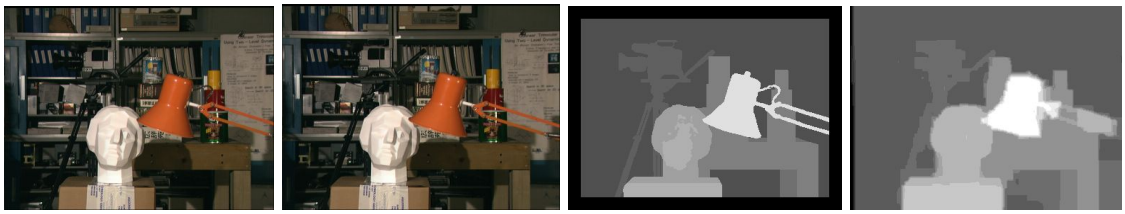
$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} E_i(x_i) + \sum_{(i,j) \in \mathcal{E}} E_{ij}(x_i, x_j)$$

Here, $\mathcal{V}$ is the set of pixels, $\mathcal{E}$ is the set of 4-connected edges, x_i is the label of pixel i, E_i is the unary potential and E_{ij} is the pairwise potential (regularization term) respectively.

As discussed in the lecture, a stereo problem can be formulated as an energy function of the form above where the labels are the disparity levels, unary potential is the color consistency between two pixels and the pairwise term has the following form:

$$E_{ij}(x_i, x_j) = w_{ij}(x_i - x_j)^2$$

where w_{ij} is the magnitude of the intensity difference between pixels *i* and *j*. The following is the result of minimizing the above energy function.



Left image Right image Ground truth disparity Our results

a. Currently the results are over-smoothed. Can you suggest some modifications to the above energy function to obtain better results? Explain why. [2 marks]
Hint: The quadratic regularization is too strong, think about a better pairwise term.

b. Please read the attached paper titled "Fast Approximate Energy Minimization using Graph Cuts". Briefly explain the expansion algorithm (including the pseudo code) using your own words. (there will be a penalty if your explanation exceeds 1 page) [3 marks]

c. [**Extra credit**] Consider the following five functions: quadratic (z^2), linear (z), truncated quadratic (min(z^2,k)), truncated linear (min(z,k)), and Cauchy function (k^2/2 log (1 + (z/k)^2)), where $z = |x_i - x_j|$ and $k > 0$. Which of these functions, if used as a pairwise term in the above energy function, can be optimized using the expansion algorithm? Justify your choices. [2 marks] Discuss why there is such a requirement for the expansion algorithm and how one could attempt to overcome that. [2 marks]

Hint: You can read the follow-up works on the expansion algorithm and its connection to st-graph cuts and submodular functions.

========== END of Assignment-1 ==========