CS GY 6643 - Computer Vision, Fall 2024

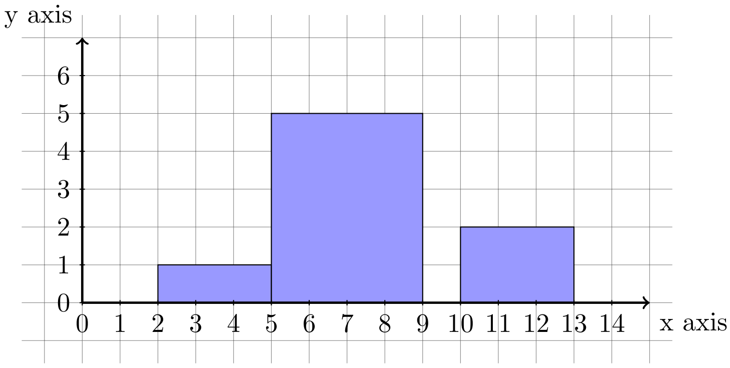
Homework 2

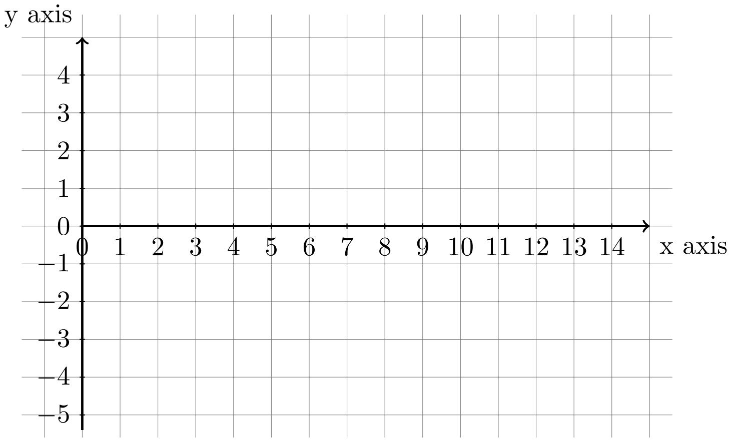
Due: 2024/10/18 11:59 PM

Note: Delays will incur 1 point deduction for every hour of delay in submissions(rounded down). Discussions are allowed on homework, but solutions must be written independently.

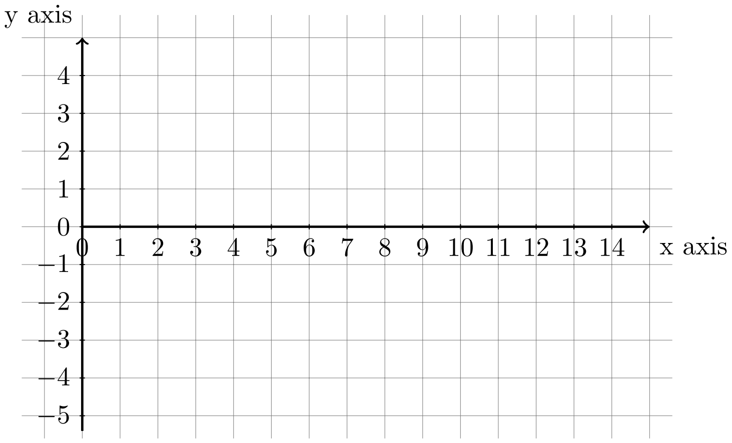
1 Edge Detection (5+5+(5+5))

(a) In the below 1-D image signal apply 1st order differentiation to find the edges and plot the resulting output and mark the edges, also mention the reasoning or heuristics behind your approach. Show the filter you used.

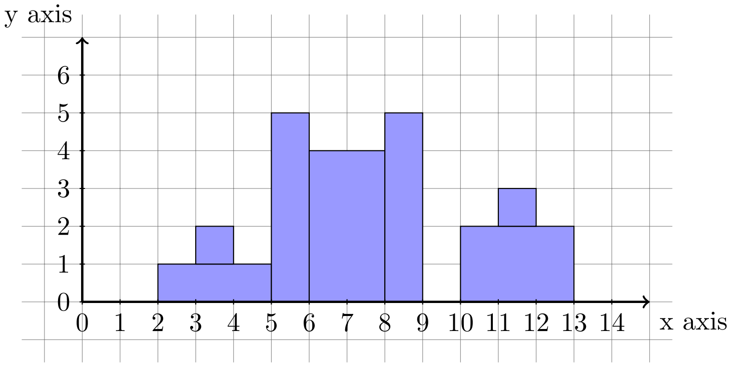


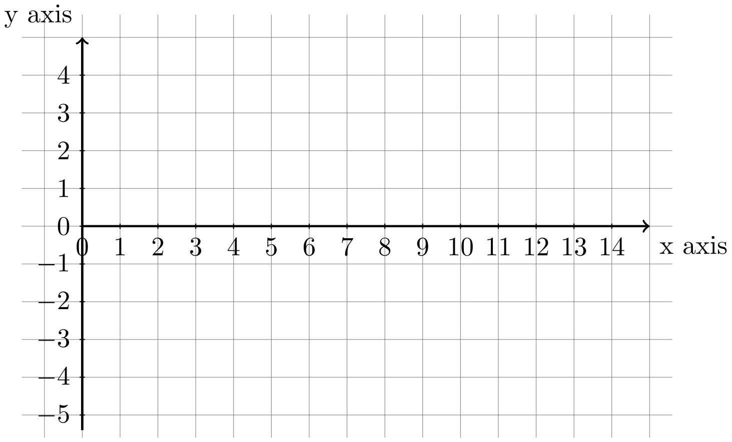


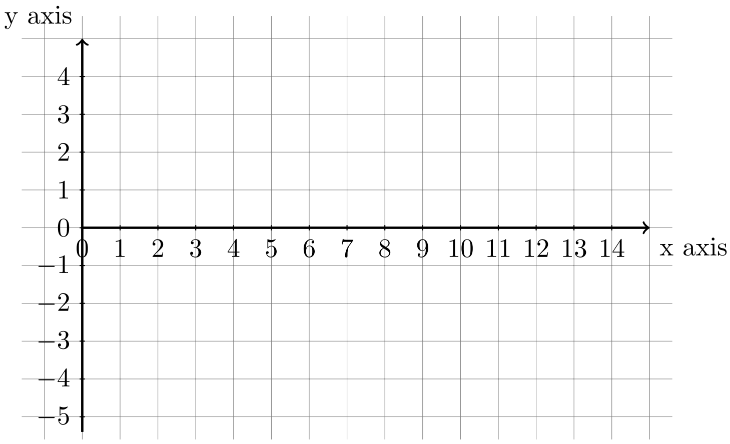
(b) On the same 1-D image signal apply 2nd order differentiation to find the edges and show the filter you have used. Plot the resulting output and mark the edges, also mention the reasoning or heuristics behind your approach. Can you reuse the previous filter from 1(a) to perform the 2nd order differentiation ?



(c) On the same 1-D image signal we have added some noise now apply methods from 1(a) and 1(b) and plot the outputs. Discuss your procedure how will you find the edges and critically examine which method is performing better.







2 Feature Descriptors and Fourier Space (6+8+(2+2+2))

(a) We all know **SIFT**(Scale-Invariant Feature Transform) feature descriptor, consider a new feature descriptor **ORB**(Oriented FAST and Rotated BRIEF) which is made up of two underlying algorithm one of it is **FAST**(Features from Accelerated Segment Test) and **BRIEF**(Binary Robust Independent Elementary Features). Now if ORB uses these two algorithm only without the modifications done in the "true" ORB algorithm, will ORB be scale and rotation invariant. Discuss why or why not it will be scale and/or rotation invariant.

(b) Prove convolution theorem in discrete 2D domain and demonstrate that for an image $I(x, y)$ and filter $h(x, y)$, convolution is equivalent to multiplication of $fft(I) \times fft(h)$.

(c) Derive the number of operations you need to convolve an image of dimension $N \times N$ and template of size $M \times M$ in spatial domain and frequency domain(fft). Now discuss how Image and template size(M and N) affects the computational complexity. Discuss various cases for the size of M and N sizes and judge which approach is better in each of the below scenarios

$$M << N$$

$$M < N$$

$$M = N$$

3 RANSAC(7+3+5+5)

Given a dataset with an outlier ratio O, you are using the RANSAC algorithm to fit a model. You want to ensure a 95% success rate for finding the correct model. Suppose the model requires $d + 1$ points for fitting.

(a) If you are working in a space with 4 dimensions(d), what is the minimum number of trials needed to guarantee a 95% success rate, given the outlier ratio O of 0.4 and that you are randomly selecting $d + 1$ points for each trial? Now find the number of trials if the outlier ratio changes to 0.2 and dimensions are increased to 7

(b) The RANSAC algorithm typically selects d+1 points to fit a model in a d-dimensional space. Explain why it is generally not advisable to use more than d+1 points when selecting correspondences for model fitting.

(c) Suppose you are given n unique points and need to select subsets of size $d + 1$ for RANSAC model fitting. What is the maximum number of possible combinations of d+1 points that can be drawn from these n unique points? (Hint: this is a combinatorial problem(n d+1))

(d) Based on this maximum number of combinations and the outlier ratio O, calculate the upper bound on the guaranteed success rate for the RANSAC algorithm.

4 Geometric Transformation Theory(6+(3+3+3)+5)

In image processing and computer vision, transformation operations like translation, rotation, and scaling can be applied to images. However, the order in which these operations are applied can significantly affect the result.

(a) Explain why the order of applying transformations, such as rotation followed by translation, may lead to different results compared to translation followed by rotation. Provide an example of a scenario where this difference is noticeable.

(b) When applying geometric transformations to images (such as resizing, rotating, or warping), interpolation is often required to compute pixel values at non-integer positions. Different interpolation methods, such as nearest-neighbor, bilinear, and bicubic interpolation, can handle this task differently. Explain the differences between nearest-neighbor, bilinear, and bicubic interpolation techniques in terms of their accuracy and computational cost.

(c) Which interpolation method is more likely to amplify noise when applied to a noisy image? Explain why certain interpolation methods tend to enhance noise more than others, and provide an example scenario where noise amplification could be problematic.

5 Geometric Transformation Working example(10+10)

You are given two sets of points in 2D:

$$SetA : A_1(1, 2), A_2(3, 4)$$
$$SetB : B_1(2, 1), B_2(4, 3)$$

The goal is to find the transformations that map Set A to Set B.

(a) Derive the proper rotation matrix that aligns Set A with Set B. Ensure your solution minimizes the squared error between the transformed points in Set A and their corresponding points in Set B.

(b) Derive the improper rotation matrix (which includes reflection) that also minimizes the squared error between the transformed points in Set A and Set B. What is the difference in interpretation between proper and improper rotation in this context?