Name: Entry no:

COL7(6)83 End-Semester Exam

20 November, 2025

1. Answer each of the following questions with a brief justification (1-2 sentences).

   (a) Starting with a low-contrast image, histogram equalization is performed, then it is performed again. What happens the second time?

   Answer: Nothing happens (0.5 marks) because histogram is already uniform/flat after first histogram equalization (0.5 marks).

   (b) In colour JPEG compression, RGB values are converted to YCbCr as follows (approximately):

   $$Y = 0.3R + 0.6G + 0.1B, \qquad C_B = 128 + (B - Y)/1.8 \qquad C_R = 128 + (R - Y)/1.4.$$

   Which component(s) correspond to luminance and which to chromaticity?

   Answer: Y corresponds to luminance, C_B and C_R to chromaticity (0.5 marks) because Y is weighted sum of intensities in R, G, B, while C_B, C_R correspond to their differences (0.5 marks) (Alternative justification: if gray colour $R = G = B$ then Y is the same, C_B, C_R are zero.)

   (c) Find the Fourier transform of a tent function $t(x) = \max(1 - |x|, 0)$, using the fact that it is the convolution of a unit box function $b(x)$ with itself.

   Answer: Fourier transform $T = B^2$ (0.5 marks), and $B(u) = \mathrm{sinc}(u)$ so $T(u) = \mathrm{sinc}(u)^2$ (0.5 marks).

   (d) Consider a grayscale image corrupted by only pepper noise, i.e. some pixels are very dark. Suggest a way to eliminate the noise without affecting small/thin bright regions.

   Answer: Closing with small SE (1 mark).

   (e) When restoring a degraded image $g = h * f + \eta$, what will happen if we attempt to reduce the noise using Gaussian convolution, then apply an inverse filter to undo the blurring?

   Answer: In frequency domain $G = HF + N$, after Gaussian it becomes $H_g HF + H_g N$, after inverse filter $H_g F + H_g N/H$ (0.5 marks). So original image will become blurred by Gaussian (0.5 marks).

2. (a) Sketch the basis functions s_0 and s_1 for the discrete cosine transform (DCT) on a discrete signal of length 8. Label both x and y axes.

   Answer: s_0 is constant (0.5 marks) with value $1/\sqrt{8}$ (0.5 marks). s_1 is half-period of cosine (0.5 marks) decreasing from $2/\sqrt{8}$ to $-2/\sqrt{8}$ (0.5 marks). If x-axis does not go from 0 to 7, -0.5 marks.

   (b) For the Haar wavelet transform, sketch the scaling function ϕ_{11} and the wavelet function ψ_{20}. Label both x and y axes.

   Answer: ϕ_{11} is constant (0.5 marks) with value $\sqrt{2}$ on $[1/2, 1]$ (0.5 marks). ψ_{20} is $+1$ then -1 (vice versa is also allowed) (0.5 marks) on $[2, 3]$ (0.5 marks).

(c) Consider (i) a smoothly varying signal with no discontinuities, (ii) a piecewise constant signal with hard boundaries between flat regions. For each of these types of signals, explain whether the DCT or the Haar wavelet transform will give a better representation, i.e. able to approximate the signal well with a small number of coefficients.

3. (a) Consider a memoryless source with two symbols a and b having probabilities $p(a) = 0.2$ and $p(b) = 0.8$. A message of length 3 was encoded using arithmetic coding, and the result was 0.1. Decode the message, showing the decoding process.

   Answer: First step: range is $[0|0.2|1]$, message 0.1 is in first half so first symbol is a. Second step: range is $[0|0.04|0.2]$, so second symbol is b. Third step: $[0.04|0.072|0.2]$, third symbol is b. Decoded message = abb. (2 marks, cut marks for conceptual error but not for calculation error. Not giving detailed rubric because not sure what kind of partially correct answers can come; please discuss any such cases if they arise.)

   (b) Suppose I try to create a modified version of JPEG compression for colour images, in which I convert the image to the HSI colour space instead of YCbCr. I observe that the results are very poor, especially for red objects. Explain in detail why this occurs, describing: (i) which of the H, S, and I channels will be problematic, (ii) what will happen after block transform coding, and (iii) why the resulting coefficients will not compress well.

4. (a) Let f be an arbitrary grayscale image, and b be the 3×3 square structuring element. Can dilation twice, $(f \oplus b) \oplus b$, be computed using a single morphological operation with a (possibly different) S.E.? Similarly, is closing twice, $(f \bullet b) \bullet b$, equivalent to a single morphological operation? Justify your answer.

   Answer: Dilation twice = max over adjacent of adjacent pixels = dilation with 5×5 square SE (1 mark). Closing twice = closing once with 3×3 SE because closing is idempotent! (1 mark)

   (b) You are given a binary image A containing several non-overlapping white disks of various radii between 5 and 50 pixels. How can one use morphological operations to obtain an image containing only the disks of radius between 20 and 25 pixels?

   Answer: Let B_1 = disk SE of radius 20, then opening $A \circ B_1$ = image containing disks with $r \geq 20$. B_2 = disk SE of radius 26, $A \circ B_2$ = image containing disks with $r \geq 26$. $(A \circ B_2) - (A \circ B_1)$ is desired image. (1 mark for at least one opening with disk, 1 mark for difference of two openings.)

   (c) Suppose some of the disks in the input image A are overlapping. Suggest a simple method to find the centers and radii of all the disks (regardless of size).

   Answer: Compute distance transform of A. Locations of peaks are centers (0.5 marks) and values there are radii (0.5 marks).

5. (a) Given a binary image obtained from edge detection, design an efficient Hough transform algorithm for detecting all circles passing through a specified point (p_x, p_y).

   (If needed, you may assume a procedure is available for enumerating all pixels lying on an arbitrary line $ax + by + c = 0$.)

(b) Consider 3 adjacent pixels i, j, k in an image, with intensities $f(i) = f(j) < f(k)$, and suppose the intensity $f(j)$ is more similar to that of known foreground pixels than to background pixels. In graph cut segmentation, should the edge weight w_{ij} be greater than w_{jk} or less than it? What about the weights of edges joining pixel j with the foreground/background terminal nodes, $w_{j,\text{fg}}$ vs. $w_{j,\text{bg}}$? Justify your answer.

Answer: Edge weights are cost to cut edge, should be large for similar nodes. Should have $w_{ij} > w_{jk}$ because intensities of i, j are more similar (1 mark). Similarly should have $w_{j,\text{fg}} > w_{j,\text{bg}}$ because j is more similar to foreground. (1 mark)

6. (a) Give the formulas for the second-order shape moments of a region, $\mu_{20}, \mu_{11}, \mu_{02}$, and discuss how they vary under translation, rotation, and scaling of the region. Are they invariant, covariant, or neither with respect to each of the three transformations?

Answer: $\mu_{20} = \sum_{(x,y)\in R}(x - m_x)^2(y - m_y)^0$, similarly for μ_{11}, μ_{02} (1 mark). Invariant to translation because uses displacement relative to mean (0.5 marks), neither w.r.t. rotation e.g. 90 degree rotation will switch x and y (0.5 marks), covariant to scaling: scaling region by s scales moments by s^2 (0.5 marks).

(b) Recall that the SIFT descriptor uses a magnitude-weighted histogram of gradient orientations in a neighbourhood of the keypoint. Explain what this means.

(c) After normalization by the sum of the histogram bins, is the resulting SIFT descriptor invariant, covariant, or neither, with respect to (i) intensity offsets $f(x, y) \mapsto f(x, y) + c$, (ii) intensity scaling $f(x, y) \mapsto af(x, y)$?

Answer: Invariant to intensity offsets because only uses intensity gradients (0.5 marks). Invariant to intensity scaling because all magnitudes will be scaled by same amount (0.5 marks) which will be undone by normalization (0.5 marks).