Final exam: CS 663, Digital Image Processing, 28$^{\text{th}}$ November

Questions

1. **Instructions (0 mark question):** There are 180 minutes for this exam (9:00 am to 12 noon). Answer all 9 questions, each of which carries 10 points. This exam is worth 20% of the final grade. Some formulae are listed in the beginning. Think carefully and write **concise** answers. At 12 noon, you must stop writing, scan your answers and submit the PDF on SAFE as well as moodle by 12:30 pm today. You must submit your screen recording link by 11:59 pm today and latest by 29th November 9 am.
   Some formulae
   (a) Gaussian pdf in 1D centered at μ and having standard deviation σ: $p(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/(2\sigma^2)}$.
   (b) 1D Fourier transform and inverse Fourier transform:
   $F(u) = \int_{-\infty}^{+\infty} f(x)e^{-j2\pi ux}dx, f(x) = \int_{-\infty}^{+\infty} F(u)e^{j2\pi ux}du$
   (c) 2D Fourier transform and inverse Fourier transform:
   $F(u,v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)e^{-j2\pi(ux+vy)}dxdy, f(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F(u,v)e^{j2\pi(ux+vy)}dudv$
   (d) Convolution theorem: $\mathcal{F}(f(x)*g(x))(u) = F(u)G(u); \mathcal{F}(f(x)g(x))(u) = F(u)*G(u)$ where $\mathcal{F}$ is the Fourier operator.
   (e) Fourier transform of $g(x-a)$ is $e^{-j2\pi ua}G(u)$. Fourier transform of $\frac{df^n(x)}{dx^n} = (j2\pi u)^n F(u)$ ($n>0$ is an integer).

2. Let $\delta(x,y)$ be a two-dimensional Dirac delta function which has an impulse of infinite height at the origin (0,0) and has the value zero everywhere else. What is the result of convolution of a given function f(x,y) with $\delta(x,y)$? What is the result of convolution of f(x,y) with $\delta(x-a,y-b)$? Using these results, derive the Fourier transform of the function $f(x,y) \triangleq A\sin(2\pi qx + 2\pi ry)\dfrac{\exp(-(x^2+y^2)/(2\sigma^2))}{2\pi\sigma^2}$. (Do not use any other method to determine this Fourier transform.) Given its Fourier transform, state a potential application of using $f(x,y)$ as a filter. [2+2+4+2=10 points]
   Answer: The convolution of f and δ gives back the same function f (proof not expected). The convolution of $f(x,y)$ with $\delta(x-a,y-b)$ gives back $f(x-a,y-b)$. (Proof not expected if the answer is right, otherwise proof is needed for partial credit if the answer were wrong. The proof is $h(u,v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(u-x-a,v-y-b)dxdy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x'-a,y'-b)\delta(u-x',v-y')dx'dy' = f(u-a,v-b)$.) To find the required Fourier transform, we use the variant of the convolution theorem. The Fourier transform of $g_1(x,y) \triangleq A\sin(2\pi qx + 2\pi ry)$ yields us $G_1(\mu,\nu) = (\delta(\mu-q,\nu-r) - \delta(\mu+q,\nu+r))/(2j)$ and the Fourier transform of $g_2(x,y) \triangleq \dfrac{\exp(-(x^2+y^2)/(2\sigma^2))}{2\pi\sigma^2}$ yields us $G_2(\mu,\nu) = \exp(-2\pi^2\sigma^2(\mu^2+\nu^2))$ using the formula provided in Q1. To see, this we know that $\mathcal{F}(e^{-ax^2})(\mu) = \sqrt{\pi/a}e^{-\pi^2\mu^2/a^2}$ from Q1. Putting in $a = 1/(2\sigma^2)$, we have $\mathcal{F}(e^{-x^2/(2\sigma^2)})(\mu) = \sqrt{\pi}\sqrt{2\sigma^2}e^{-2\pi^2\mu^2\sigma^2}$, and hence $\mathcal{F}(e^{-(x^2+y^2)/(2\sigma^2)})(\mu,\nu) = 2\pi\sigma^2 e^{-2\pi^2\sigma^2(\mu^2+\nu^2)}$ which further yields $\mathcal{F}(\dfrac{e^{-(x^2+y^2)/(2\sigma^2)}}{2\pi\sigma^2})(\mu,\nu) = e^{-2\pi^2\sigma^2(\mu^2+\nu^2)}$.
   The Fourier transform of their element-wise product is the convolution of G_1 and G_2. This is basically the convolution of a Gaussian with the sum of two impulses.

   Let us consider the case that the Gaussian had been $\dfrac{\exp(-(x^2+y^2)/(2\sigma^2))}{2\pi\sigma^2}$; we will change this later. Using the earlier result, we have
   $$F(\mu,\nu) = \frac{G_2(\mu-q,\nu-r) - G_2(\mu+q,\nu+r)}{2j}$$
   $$= \frac{\exp(-2\pi^2\sigma^2((\mu-q)^2+(\nu-r)^2)) - \exp(-2\pi^2\sigma^2((\mu+q)^2+(\nu+r)^2))}{2j}.$$ The application of this filter is

that when convolved with any image, it will weaken all frequencies in that image that are too far away from (q, r) or $(-q, -r)$ - how far depends upon the value of σ. A large σ in g_2 implies a small standard deviation in the μ, ν domain and vice versa.

But the Gaussian term is $\dfrac{\exp(-(x^2+y^2))}{2\pi\sigma^2}$. Its Fourier transform is equal to $\dfrac{1}{2\pi\sigma^2} \times \exp(-2\pi^2(\mu^2+\nu^2))$. In this case, the final answer is:

$F(\mu, \nu) = \dfrac{\exp(-2\pi^2((\mu-q)^2+(\nu-r)^2)) - \exp(-2\pi^2((\mu+q)^2+(\nu+r)^2))}{2j \times 2\pi\sigma^2}$.

Marking scheme: The Fourier transform of $f(x, y)$ must be derived using a method which involves convolution with the shifted delta function. Direct substitution is not allowed and is also considerably more complex. If the Gaussian standard deviation in the fourier transform is incorrect, deduct 1 point. If the shifting of μ, ν is not correct, then deduct 2 points. To get points for the application, the statement that there is weakening of frequencies far away from (q, r) or $(-q, -r)$ is a must. If the student assumed the Gaussian term in $f(x, y)$ to be $\dfrac{\exp(-(x^2+y^2)/(2\sigma^2))}{2\pi\sigma^2}$, the student should still get full points.

3. Given two signals $f(t)$ and $g(t)$, their cross-correlation is given as $h(s) = (f \star g)(s) = \int_{-\infty}^{+\infty} f^*(t)g(t+s)dt$ where x^* denotes the complex conjugate of x. This formula is slightly different from what we saw in class, because we were not considering complex numbers at that time. Your task is to prove that the Fourier transform of $h(t)$ is given by $F^*(\mu)G(\mu)$ where F, G stand for the Fourier transforms of f, g respectively. Use this result to argue whether or not cross-correlation is a commutative operation. [7+3=10 points]

   Answer: We have $\mathcal{F}(h(s))(\mu) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f^*(t)g(t+s)e^{-j2\pi\mu s}dsdt = \int_{-\infty}^{+\infty} f^*(t)\Big(\int_{-\infty}^{+\infty} g(t+s)e^{-j2\pi\mu s}ds\Big)dt$ $= \int_{-\infty}^{+\infty} f^*(t)G(\mu)e^{j2\pi\mu t}dt = \Big(\int_{-\infty}^{+\infty} f^*(t)e^{j2\pi\mu t}dt\Big)G(\mu) = \Big(\int_{-\infty}^{+\infty} f(t)e^{-j2\pi\mu t}dt\Big)^* G(\mu) = F^*(\mu)G(\mu)$. Note that we have used the Fourier shift theorem in one step for the Fourier transform of $f(t+s)$.

   This result shows that the Fourier transform of $(g \star f)$ is $G^*(\mu)F(\mu)$ which is unequal to the Fourier transform of $f \star g$ which is $F^*(\mu)G(\mu)$. This shows that cross-correlation is not commutative (unlike the convolution).

   Marking scheme: In the proof, correct substitution of Fourier transform for h fetches 1 point. Correct use of Fourier shift theorem for Fourier of $g(t+s)$ fetches 2.5 points, correct removal of $G(\mu)$ from out of the integral fetches 1 point, and correct final result with the complex conjugation step fetches the remaining 2.5 points. <u>Justification</u> that cross-correlation is not commutative is a must for the remaining 3 points (no points for only stating that it is not commutative.)

4. (a) Let (x_{1i}, y_{1i}) and (x_{2i}, y_{2i}) be a pair of physically corresponding points in two images J_1 and J_2. Let us suppose that the motion between the two images, and hence the motion between any pair of corresponding points between the two images, is given by the following two equations: $x_{2i} = a_1x_{1i}^2 + b_1y_{1i}^2 + c_1x_{1i}y_{1i} + d_1x_{1i} + e_1y_{1i} + f_1$ and $y_{2i} = a_2x_{1i}^2 + b_2y_{1i}^2 + c_2x_{1i}y_{1i} + d_2x_{1i} + e_2y_{1i} + f_2$. Here are unknown scalar constants. Write down a system of matrix-based equations to determine these coefficients from some N pairs of corresponding points in J1, J2. Write down the dimensions of each vector/matrix in this system clearly. What is the minimum number of point pairs required in order to get a unique solution to this system of equations, assuming that all point pairs are unique and that no three points in either image are collinear?

   (b) What is the advantage of reverse warping over forward warping? [(4+2)+4=10 points]

   Answer: For the first part, the matrix based equations are $\boldsymbol{P_2} = \boldsymbol{C}\boldsymbol{P_1}$:

$$\begin{bmatrix} x_{21} & x_{22} & ... & x_{2N} \\ y_{21} & y_{22} & ... & y_{2N} \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 \\ a_1 & b_1 & c_1 & d_1 & e_1 & f_1 \end{bmatrix} \begin{bmatrix} x_{11}^2 & x_{12}^2 & ... & x_{1N}^2 \\ y_{11}^2 & y_{12}^2 & ... & y_{1N}^2 \\ x_{11}y_{11} & x_{12}y_{12} & ... & x_{1N}y_{1N} \\ x_{11} & x_{12} & ... & x_{1N} \\ y_{11} & y_{12} & ... & y_{1N} \\ 1 & 1 & ... & 1 \end{bmatrix}, \tag{1}$$

   where $\boldsymbol{P_2}$ has size $2 \times N$, $\boldsymbol{C}$ has size 2×6 and $\boldsymbol{P_1}$ has size $6 \times N$. For a unqiue solution, we must have N to be at least 6.

For the second part, the advantage of reverse warping over forward warping is that: forward warping can leave holes in the image (eg: scaling to larger sizes) unlike the former; forward warping can also lead to multiple answers in a pixel if you downsample (necessitating awkward averaging operations) unlike the former.
Marking scheme: The matrix dimensions must be specified correctly, otherwise deduct 2 points. A transpose of the equation in my answer along with appropriate change in matrix dimensions is perfectly acceptable. 2 points for for the minimum number of points correctly stated (reason is not required). For the second part, there are two advantages - two points for each. Reasons for each advantage or definition of reverse/forward warping are not needed.

5. (a) Consider a 2D color (RGB) image with N pixels. Write the expression for the kernel density estimate of 5D vectors of the form $(x_i, y_i, R(x_i, y_i), G(x_i, y_i), B(x_i, y_i)$ assuming Gaussian kernels with standard deviation σ_s for the x,y coordinates and standard deviation σ_r for the R,G,B values.
   (b) Given a pixel (x_k, y_k) with intensity values $R(x_k, y_k), G(x_k, y_k), B(x_k, y_k)$, write the expression for its initial mean shift vector.
   (c) What would be the effect on image segmentation, if we ignored the x,y coordinates while building the kernel density estimate in mean shift and considered only the RGB values?
   (d) Mean shift involves doing gradient ascent on an appropriate probability density estimate. When does the ascent halt? [2.5 + 2.5 + 2.5 + 2.5 = 10 points]

Answers: For the first part, the expression is

$$p(x,y,R,G,B) = \frac{1}{N \times (2\pi)^{2.5}\sigma_s^2\sigma_r^3}\sum_{i=1}^N \exp\Big(-\frac{[(x-x_i)^2+(y-y_i)^2]}{2\sigma_s^2} - [\frac{(R-R(x_i,y_i))^2}{2\sigma_r^2} + \frac{(G-G(x_i,y_i))^2}{2\sigma_r^2} + \frac{(B-B(x_i,y_i))^2}{2\sigma_r^2}]\Big).$$

Second part: Define $\boldsymbol{v_k} \triangleq (x_k, y_k, R(x_k, y_k), G(x_k, y_k), B(x_k, y_k))$. The initial mean shift vector at the pixel (x_k, y_k) is given by:

$$\frac{\sum_{i=1}^N \boldsymbol{v_i} \exp\Big(-\frac{[(x-x_i)^2+(y-y_i)^2]}{2\sigma_s^2} - [\frac{(R-R(x_i,y_i))^2}{2\sigma_r^2} + \frac{(G-G(x_i,y_i))^2}{2\sigma_r^2} + \frac{(B-B(x_i,y_i))^2}{2\sigma_r^2}]\Big)}{\sum_{i=1}^N \exp\Big(-\frac{[(x-x_i)^2+(y-y_i)^2]}{2\sigma_s^2} - [\frac{(R-R(x_i,y_i))^2}{2\sigma_r^2} + \frac{(G-G(x_i,y_i))^2}{2\sigma_r^2} + \frac{(B-B(x_i,y_i))^2}{2\sigma_r^2}]\Big)} - \boldsymbol{v_k}.$$

Third part: If we ignored the (x, y) coordinates in mean shift, we would assign faraway pixels with similar color values to the same segment. If we consider (x, y) as well as (R, G, B) values together while building the density estimate, we would assign such pixels to different segments potentially which is more natural.
Last part: Mean shift halts when you reach a local mode (maximum) of the PDF. The step size of the ascent is automatically chosen because the mean shift vector is proportional to $\nabla p/p$. At a local maximum of p, we see that p is high and ∇p is zero, so the ascent automatically stops.
Marking scheme: For the first part, the constant factors should be correct, otherwise one point is deducted. The standard deviations in the Gaussian must be correctly placed, otherwise 1.5 points are lost. For the second part, you lose 1 point again if the factors involving the standard deviations are incorrect. In both parts, matrix-based expressions in the exponent are also fine. For the second part, the expression must produce a vector, otherwise all 2.5 points are lost. For the third part, no reasons for faraway pixels being assigned to the same segment are expected. For the fourth part, mentioning the mode is important, and you must also mention that at the mode, p is large and ∇p is small, so that the procedure stops.

6. Write down expressions or kernels for the following filters as direcred. You may leave your answer as an expression involving convolutions (wherever applicable), without solving the convolution itself:

   (a) Write down a 5 x 5 filter kernel which first replaces all pixel values by the average of the intensity values of their two left and two right neighbors, followed by replacing the resultant values by the difference between the values of their south and north neighbors.

Answer: $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ A 5 by 5 kernel is a must, otherwise, you will lose 1 point. *The first kernel* must add up to 1, otherwise you lose 1 point. *The second kernel must add up to zero otherwise you lose 1 point (no negative marks, though).*

(b) Write down a 5 x 5 filter kernel which which first replaces all pixel values by the average of the intensity values of all pixels in a surrounding 3 x 3 neighborhood but excluding the central pixel, followed by replacing the resultant values by their Laplacian.

Answer: $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -4 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1/8 & 1/8 & 1/8 & 0 \\ 0 & 1/8 & 0 & 1/8 & 0 \\ 0 & 1/8 & 1/8 & 1/8 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ The answer is reverse order is also fine as convolution is commutative. A 5 by 5 kernel is a must, otherwise, you will lose 1 point. The smoothing part of the kernel must add up to 1, otherwise you lose 1 point.

(c) Write down the expression for a Gaussian filter which smoothes more along the Y direction than the X direction. You need not write down a kernel.

Answer: $g_1(x,y) = \dfrac{1}{2\pi\sigma_x\sigma_y} \exp\Big(- x^2/(2\sigma_x^2) - y^2/(2\sigma_y^2)\Big)$ where $\sigma_y > \sigma_x$. You lose 1 point if the constant factor is incorrect and 2 points if the part on $\sigma_y > \sigma_x$ is missing.

(d) Write down the expression for a Gaussian filter which smoothes more along the edge at a pixel (x,y) than across the edge at that pixel.

Answer: The gradient direction is perpendicular to the edge at any location. The gradient vector at (x,y) is given as $\boldsymbol{g}(x,y) \triangleq \dfrac{I_x(x,y)}{\sqrt{I_x^2(x,y)+I_y^2(x,y)}}, \dfrac{I_y(x,y)}{\sqrt{I_x^2(x,y)+I_y^2(x,y)}}$ and the edge direction is given as $\boldsymbol{e}(x,y) \triangleq \dfrac{-I_y(x,y)}{\sqrt{I_x^2(x,y)+I_y^2(x,y)}}, \dfrac{I_x(x,y)}{\sqrt{I_x^2(x,y)+I_y^2(x,y)}}$. The required filter constructed around pixel (x,y) is given by

$g_2(x',y') = \dfrac{1}{2\pi\sigma_e\sigma_g} \exp\Big(- (\boldsymbol{v}\cdot\boldsymbol{e})^2/(2\sigma_e^2) - (\boldsymbol{v}\cdot\boldsymbol{g})^2/(2\sigma_g^2)\Big)$ where $\sigma_e \gg \sigma_g, \sigma_g \approx 0$ where $\boldsymbol{v}(x,y) \triangleq (x'-x, y'-y)$. You lose 1.5 points if the distinction between (x,y) and (x',y') is unclear. Alternatively, you can write an expression assuming $(x,y) = (0,0)$ as well. You lose 1 point if the constant factor is incorrect and 1 point if the part on $\sigma_e > \sigma_g$ is missing. If the edge vector and gradient vector are not carefully defined, you do not get any points.

(e) Which of these filters if any cannot be expressed as the result of convolution of the image with a single kernel? Why? [2+2+2+3+1=10 points]

Answer: The last filter cannot be expressed as the result of a convolution of the image with a single kernel. The filter is clearly not space-invariant and not linear globally as it depends on the signal gradient/edge. The reason must be correctly specified (any one reason is enough).

7. A medical technician is inspecting the quality of some gray-scale images acquired by a new type of sensor. The technician notices that all of the images have the following artifacts: (1) a few very high intensity and scattered dots in different places which have no medical significance, (2) change in the mean image intensity of the image, though the true mean value should be some known c if the image acquisition was good, (3) poor contrast, and (4) lack of image sharpness. As these artifacts hinder subsequent image inspection by a doctor, the technician wishes to correct all these problems, and then display the intensities within some range [a; b] in black color, while leaving other intensity values unchanged. Suggest a pipeline of steps (in the correct order) for the technician to remove/weaken the different artifacts. There is no need to describe each step in detail. For example, if you use bilateral filtering in one or more steps, you need not describe the filter from scratch, but just state that you use the bilateral filter in so and so step(s). [10 points]
Answer: The steps for enhancement are: (A) perform median filtering, preferably only in the close vicinity of the scattered dots to remove the spiky noise, (B) histogram equalization for removing poor contrast, (C) sharpening filter to boost higher frequencies and thus remove the blur, (D) Let $\bar{I}$ be the current average of the image I after step (C). Now we apply the operation $I(x,y) \leftarrow I(x,y) - \bar{I} + c$ to every pixel (x,y) to make the average value equal to c. Thereafter, (E) for the purposes of display, all intensities in $[a,b]$ are to

be set to 0 leaving the others unchanged.
Marking scheme: 2 points per step times 5 steps. The order of (B) and (C) may be interchanged. Apart from this, if the order is incorrect, deduct 1 point. Deduct 1 point if the student does not mention that the median filtering is performed in the close vicinity of the scattered dots. If you perform median filtering everywhere, you may subtly alter some important structural details of the rest of the image.

8. (a) Which properties of the image Laplacian render it useful for edge detection in image processing?
Answer: The Laplacian changes sign on opposite sides of an edge and is zero-valued on the edge. This property is very useful for edge detection. (Additionally, it is in principle rotationally invariant, but not in its discrete form – this part is not required for the answer.)

(b) Why does subtracting a quantity directly proportional to the image Laplacian at a pixel cause image sharpening? That is suppose you execute $J^{(t+1)}(x,y) = J^{(t)}(x,y) - \alpha[J^{(t)}_{xx}(x,y) + J^{(t)}_{yy}(x,y)]$, then why is $J^{(t+1)}$ a sharpened version of $J^{(t)})$? Answer this question from first principles based on properties of the Laplacian, do not just quote results pertaining to the heat equation.
Answer: The Laplacian changes sign on opposite sides of an edge and is zero-valued on the edge. For an edge that goes from lower to higher intensity, the Laplacian is positive on the side with lower intensity and negative on the side with higher intensity. When you deduct a quantity proportional to the Laplacian, the side with lower intensity acquires an even lower intensity, and the side with higher intensity acquires an even higher intensity. This causes edge sharpening. The edge itself does not change as the Laplacian is zero on the edge. For an edge from higher to lower intensity, the Laplacian again has negative values on the higher side and positive values on the lower side, and again you will get edge sharpening. (note: Merely quoting results pertaining to the heat equation or its reverse fetch no points. A student may explain that adding the Laplacian produces blur, and hence deduction will produce the opposite effect. This approach will fetch full points if the student has explained <u>why</u> adding the Laplacian produces blur.)

(c) Derive an expession for the Fourier transform of the Laplacian of an image $f(x,y)$. Assume that the image is of infinite extent and continuous valued. [3+4+3=10 points]
Let $F(\mu,\nu)$ be the Fourier transform of an image $f(x,y)$. Then using the differentiation theorem with $n=2$ (once for the double x-derivative, once for the double y-derivative), we have $\mathcal{F}(\nabla^2 f)(\mu,\nu) = (j2\pi\mu)^2 F(\mu,\nu) + (j2\pi\nu)^2 F(\mu,\nu) = -(4\pi^2[\mu^2+\nu^2])F(\mu,\nu)$.