THE UNIVERSITY OF BRITISH COLUMBIA
SESSIONAL EXAMINATIONS - DECEMBER 1996

COMPUTER SCIENCE 505
Image Understanding I: Image Analysis

TIME: 3.0 HOURS

SPECIAL INSTRUCTIONS:

1. This exam is closed book.

2. This exam consists of two parts.

   Part A is worth 40% of the total grade. Part A consists of two (2) questions worth twenty (20) marks each. Each question in Part A has multiple subparts. Within a question, each subpart has equal value. Answer all questions, and all subparts of each question, in Part A.

   Part B is worth 60% of the total grade. Part B consists of five (5) questions worth twenty (20) marks each. Answer any three (3) questions from Part B. Students who attempt more than three questions from Part B will be marked on their worst three answers, not on their best three answers.

3. You are allowed 3 hours to complete the examination.

   This time should be ample. Use some of your time to first read carefully all questions, and all subparts of all questions, before beginning to answer. MARKS WILL BE DEDUCTED FOR LONG-WINDED ANSWERS. You must try to be clear and concise. When you are asked to explain or interpret something, avoid simply stating facts you believe to be true. Instead, indicate how each statement is relevant to the question posed.

4. If you find a question ambiguous then say so in your answer.

   You might say, “I find this question ambiguous. If it means A then my answer is B. On the other hand, if it means C then my answer is D.” This is preferable to saying, “The answer is B or D or ... ”

5. Good luck!

   THIS EXAMINATION CONSISTS
   OF TEN (10) PAGES. CHECK TO
   ENSURE THAT THIS PAPER IS
   COMPLETE.

   INSTRUCTOR’S NAME: Dr. J.J. Little
   SECTION NUMBER: 101
Part A: Answer both Question 1 and Question 2.

Question 1:

For each of the following statements, indicate whether it is **TRUE** or **FALSE**. In each case, briefly justify your answer.

a) The Laplacian pyramid, in its standard implementation, decomposes an image into a series of bands, each of which occupies the same fraction of the full frequency range.

b) The value at a site in a Markov Random Field is independent of the values at other sites.

c) In edge detection, it is sufficient to model edges as simple step discontinuities in image brightness.

d) A Lambertian material has reflectance proportional to $\cos(i)$.

e) Let $f(x)$ be an even function with second derivative $d^2f(x)/dx^2$. Then $d^2f(x)/dx^2$ also is an even function.

f) Although the Fourier transform, $F(\omega)$, of a real function, $f(x)$, is in general complex, the imaginary part of $F(\omega)$ is zero if $f(x)$ is an even function.

g) When viewed from a fixed location, corresponding points in a sequence of images of a moving object will have exactly the same brightness values.

h) Any image operator that is linear can be implemented as a convolution.

i) The coefficients of a digital filter designed for interpolation should sum to one.

j) Let $f(x,y)$ be an image containing mostly vertical edges. Then its Fourier transform, $F(\omega_x, \omega_y)$, will contain values with non-zero magnitude mostly along the horizontal frequency axis.
Question 2:

a) A linear spatially invariant image processing operator can be described either by its transfer function or by its point spread function. What is a transfer function? What is a point spread function?

b) Let \((p_b, q_b)\) be the gradient corresponding to the direction to a single distant point source of illumination. Derive an equation, in terms of the gradient, \((p, q)\), for surface orientations that lie on the shadow terminator. [Hint: A surface orientation is on the shadow terminator if and only if it makes an angle of \(\pi/2\) radians with the direction to the light source. Said another way, \(\cos(i) = 0\) where \(i\) is the incident angle.]

c) What does it mean for a particular surface reconstruction algorithm (or method) to be “coordinate system dependent?” Give an example of an algorithm that is “coordinate system dependent” and explain what element(s) of the algorithm introduce this dependence.

d) Let \(f(t)\) and \(g(t)\) be two functions with Fourier transforms \(F(w)\) and \(G(w)\) respectively. Let \(h(t)\) be the product \(h(t) = f(t)g(t)\). Express \(H(w)\), the Fourier transform of \(h(t)\), in terms of \(F(w)\) and \(G(w)\).

e) What is the relationship between a reflectance map, \(R(p, q)\), and a bidirectional reflectance distribution function (BRDF)?
Part B:

Answer three (3) of the following five (5) questions. All subparts of a question are approximately equal in value. Students who attempt more than three questions from Part B will be marked on their worst three answers, not on their best three answers.

Question 3:

In image analysis, an "edge" is an abrupt change in image intensity. Typical edge detectors are designed to handle single, isolated step changes. Observations from real image data make clear that intensity profiles measured across edges have a variety of shapes and spatial scales. In this problem, we will reason about physical events in a scene domain consisting of simple polyhedra to learn how three non-step edge types can arise in practice.

In general, there are variety of scene properties that give rise to edges in an image.

   a) Identify four (4) distinct scene events that can give rise to abrupt changes in image intensity.

For the remainder of the problem, consider a restricted scene domain consisting only of polyhedra made of a given material. Assume that the reflectance map, $R(p, q)$, for the given material and illumination conditions is known. In this domain, all physical edges correspond to lines where two planes intersect. In the idealized model of this domain, we assume that the image irradiance equation is given by $E(x, y) = R(p, q)$ and that surface orientation is discontinuous at each edge. With these assumptions, all edges would indeed be step edges.

   b) Why, in the idealized model, are all edges step edges?

Of course, there are a variety of secondary factors that make the idealized model only a first approximation.

   c) What is inter-reflection? [Note: Inter-reflection is also called mutual illumination].

   d) With the aid of a diagram, describe a physical configuration that would give rise to a roof edge as shown in Figure ??.

![Figure 1: Roof Edge](image)

With real polyhedra, edges typically are slightly rounded so that the assumption of a discontinuity in surface orientation is not strictly true.
e) With the aid of diagrams, describe a physical configuration that would give rise to a highlight edge as shown in Figure ??.

Be sure to identify both the edge configuration in the scene and the locations of the associated gradients in the reflectance map, $R(p,q)$.

![Highlight Edge](image)

Figure 2: Highlight Edge

f) With the aid of diagrams, describe a physical configuration that would give rise to a lowlight edge as shown in Figure ??.

Be sure to identify both the edge configuration in the scene and the locations of the associated gradients in the reflectance map, $R(p,q)$.

![Lowlight Edge](image)

Figure 3: Lowlight Edge
Question 4:
This question will deal with binocular stereo. Assume cameras with parallel optical axes, perpective projection, where the optical centres of the cameras are at \((-b,0,0)\) for the left and \((b,0,0)\) for the right, in the standard left-handed coordinate system. The focal length of the cameras will be 1, for simplicity. We'll consider a variety of complicating factors in stereo computation.

a) Name two reasons why conjugate points (the projections of a surface point \((x,y,z)\) in the two sensed images) might have different measured brightnesses.

b) Consider a square planar patch (part of the viewed surface) on the surface of a viewed object. Under what conditions (position and orientation) will its projection be the same shape in both the left and right images?

c) Now, to understand the effects of viewing a patch from two different viewpoints we will study the *horizontal width* of the projection of the patch into the left and right images. Let's choose a simple orientation for the patch: it is vertical, but rotated around the \(y\) axis by the angle \(\theta\). The centre of the patch lies at \((-b,0,z)\) and its width is \(w\).

Describe the projected width of the patch, on a slice along the \(x\)-axis, in both left and right images.

d) We propose a new method for determining displacements. Now let's work in the image coordinate system. We will assume that the only difference between two rectangular patches, \(i_L\) in the left image and \(i_R\) in the right, is a displacement \(\delta x\) in the \(x\) direction. We will compute the Fourier Transform of the patches \(i_L\) and \(i_R\). Consider the left frame as the reference 2D coordinate system. Describe the Fourier transforms of the two patches.

e) Then describe the relation between the two Fourier Transforms. How can we recover the displacement \(\delta x\) between two patches simply by comparing their Fourier Transforms? [Hint: This method is called phase correlation.]

f) In the light of the result in part c, discuss whether the phase correlation method should work. Are its assumptions valid? [Hint: what does part c tell us about the frequencies of the projected patches?]

h) Does stereo work when the cameras are parallel and projection is orthographic?

h) Are there conditions (relations between the position and size of an object, and position of the camera) where orthographic projection is a good approximation of perspective projection?
**Question 5:**

The brightness constancy equation states that the brightness $E$ of an image remains constant with time:

$$ E(x, y, t) = E(x + x_d, y + y_d, t + t_d) $$ (1)

The first order Taylor series expansion about the point $(x, y, t)$ leads to the equation (with $t_d = 1$):

$$ E_x x_d + E_y y_d + E_t = 0 $$ (2)

In this problem we will assume that the image velocity $(x_d, y_d)$ is constant over the region of interest. We will denote the sampled values as $G(i, j, k)$, where $i, j$ are spatial values and $k$ is 0 or 1 for the two temporal images.

Because velocity is constant, we can solve the overconstrained system of equations of the above form. The solution is:

$$ (x_d^*, y_d^*) = \arg \min_{x_d, y_d} \sum_{i, j} (G_x x_d + G_y y_d + G_t)^2, $$ (3)

where the subscripts in $x$ and $y$ indicate differentiation. The $\arg \min$ function, with $x_d, y_d$ below, states that the solution is the $x_d, y_d$ that minimizes the following formula. We will ignore how to find the $x_d, y_d$.

Another way to determine displacements is by matching the two images, by minimizing the sum of squared differences between shifted versions of the two images:

$$ (x_d^*, y_d^*) = \arg \min_{x_d, y_d} \sum_{i, j} (G(i, j, 1) - G(i + x_d, j + y_d, 0))^2. $$ (4)

We will assume that the displacements are less than one pixel. Thus we need an interpolating function $F(i + x_d, j + y_d, k)$ to let us compare the sampled images between sample points.

Our goal is to show that there are equivalent formulations in both the optical flow and the SSD methods.

Let’s examine a 1-d case.

a) Write down an interpolant $F$, that linearly interpolates the $G$ values in space, in the $x$ dimension.

b) Replace $G$ with our linear interpolant $F$ and formulate the SSD equation, (4), in a 1-d case. Use a comparison between the point at $(i, 0)$ (image 1) with the interpolated point at $x_d$ in image 0.
c) We need to find derivatives, finite differences, that, substituted into equation (3) yield a similar formulation. There are three standard difference operators:

- forward: \( G(i + 1, k) - G(i, k) \)
- backward: \( G(i, k) - G(i - 1, k) \)
- central: \( (G(i + 1, k) - G(i - 1, k))/2 \)

Which one of the forward, backward or central difference operators should be used as the spatial derivative in (3)?

d) Which one should be used as the temporal derivative?

e) This indicates that the two formulations are equivalent. The formulation from brightness constancy is known to be “noisy”, but the SSD one is thought not to be noisy because it doesn’t involve differentiation. If they are equivalent, this means that the SSD method is also noisy.

Where does the noise arise in the brightness constancy formulation?

Where does the noise arise in the SSD formulation?

Explain the sources of the noise by analyzing the transfer functions of the relevant elements in the formulae.

f) Suggest a way to improve the SSD formulation so that it does not introduce noise. Is this practical?
Question 6:

Consider a fixed gradient, \((p_0, q_0)\), corresponding to the direction to a distant point light source. We wish to determine the locus in gradient space of surface orientations, \((p, q)\), that lie on the shadow line with respect to the direction to the given light source. (Points on the shadow line are orthogonal to the direction to the light source.)

(a) Write down the equation that \((p, q)\) must satisfy in order to lie on the shadow line with respect to the direction \((p_0, q_0)\).

[Hint: Consider the 3-D vector corresponding to the gradient and recall that two vectors are orthogonal iff their inner product is zero.]

(b) Give a geometric interpretation in terms of the gradient space to the constraint that follows when two gradients correspond to directions that are orthogonal in 3-D.

[Hint: Drawing a figure may help.]

Suppose in a polyhedral “blocksworld” scene, there is a trihedral vertex where three planes \(A\), \(B\) and \(C\) meet. If this is the vertex of a rectangular object, then the three planes are at right angles to each other (i.e., they are mutually orthogonal). Let \(G_A = (p_a, q_a)\), \(G_B = (p_b, q_b)\), and \(G_C = (p_c, q_c)\), be gradients corresponding to the planes \(A\), \(B\) and \(C\).

(c) If the vertex is known \emph{a priori} to be rectangular, is there enough information in the local image geometry to determine \(G_A\), \(G_B\) and \(G_C\)? (You don’t have to work out the equations but you should briefly justify your answer.)

We can also consider image brightness values. Suppose that the image irradiance equation is given by

\[
I(x, y) = R(p, q)
\]

where, to keep things simple, assume

\[
R(p, q) = 1 + p_0p + q_0q
\]

with \(p_0 = 0\) and \(q_0 = 1\). Suppose we observe intensities \(I_A\), \(I_B\) and \(I_C\) corresponding to planes \(A\), \(B\) and \(C\).

(d) Determine the gradients \(G_A\), \(G_B\) and \(G_C\) directly from the three intensity measurements \(I_A\), \(I_B\) and \(I_C\).

[Hint: This example is simple enough that you do not need to use the directions in the image of the three edges at the vertex.]

(e) Explain how, in general, one can check whether a given trihedral vertex is, in fact, rectangular.
**Question 7:**

In this problem, we explore the difference between Lambertian and specular surfaces for a moving light source and nearby viewer. In our standard left-handed coordinate system, assume that the viewer is positioned along the negative Z-axis a finite distance $d$, $d > 1$, from the XY-plane. That is, the viewer is located at point $(0, 0, -d)$. Further, suppose the viewer is viewing the (convex) hemispherical surface $z = -\sqrt{1 - (x^2 + y^2)}$, where $(x^2 + y^2) \leq 1$. As usual, let $p = \partial f(x, y)/\partial x$ and $q = \partial f(x, y)/\partial y$. Then, in this left-handed coordinate system, an outward pointing surface normal is $[p, q, -1]$.

Now, consider a single distant point source of illumination from direction $[s_1, s_2, s_3]$, where $s$ is a unit vector. If the surface is Lambertian with reflectance factor $\rho$, then it will have “normalized” intensity $\rho \cos(i)$, where the incident angle, $i$, is the angle between the surface normal and the light source direction. [Note: Normalized intensity here simply means that, for the purposes of the problem, we ignore camera parameters and the absolute value of the light source irradiance].

a) What is the surface normal of the brightest point on the surface? Recall that points on the unit sphere have coordinates identical to their surface normal direction. That is, if $(x, y, z)$ is a point on the sphere then the surface normal is $[x, y, z]$.

Now consider a hemispherical surface, of the same shape, that is perfectly specular. Recall that in this situation the incident ray, the emittance ray and the surface normal lie in a plane, and that $i = \epsilon$, where $i$ is the incident angle and where $\epsilon$ is the emittance angle.

b) Is the direction of the brightest point on the specular surface surface identical to that for the Lambertian surface? Is it closer to the viewer direction or farther away? It’s not necessary to perform the full calculation, but it should be possible to write the equation identifying the point on the sphere.

c) Does the location of the brightest point in part b) depend on $d$?

d) Consider the flat, planar specular surface $z = -1$ and the same illumination conditions, as above. Where is the brightest point on the surface?

e) Smoothly and at a constant rate, rotate the light source direction from $[s_1, s_2, s_3]$, through the viewer direction, $[0, 0, -1]$, to the direction $[-s_1, -s_2, s_3]$. Consider the velocity of the brightest point on the planar surface and on the hemispherical surface. Which is greater?

f) Consider the (concave) hemisphere, $z = \sqrt{1 - (x^2 + y^2)}$, where $(x^2 + y^2) \leq 1$. What is the relation between the velocity of the brightest point in this case and the cases described in part e)?