THE UNIVERSITY OF BRITISH COLUMBIA

PRACTICE FINAL EXAM - 2013

COMPUTER SCIENCE 505

Image Understanding I: Image Analysis

TIME: 3 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. To answer a given question well, it is important to focus on the essential issues. Avoid simply stating facts you believe to be true in the hope of receiving partial credit. Be clear on how each statement you make is relevant to the question posed. MARKS WILL BE DEDUCTED FOR ANSWERS THAT INCLUDE IRRELEVANT DETAIL.
4. **If you find a question ambiguous then say so in your answer.**

   Do not be discouraged by the length of some questions. The required answers often are quite simple. The extra words are there to help you find the answers, not to create extra work (or anxiety).

   Despite everyone's best effort, it is possible that you might find the wording of a question ambiguous. In this case, 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. **Write legibly!** Your answers must be readable. Where appropriate, be sure to distinguish uppercase from lowercase symbols and Greek letters from standard alphanumeric characters.
6. **Good luck!**

THIS EXAMINATION CONSISTS OF 14 PAGES. CHECK TO ENSURE THAT THIS PAPER IS COMPLETE.

INSTRUCTOR'S NAME: Dr. R.J. Woodham
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 coefficients of a digital filter designed for differentiation should sum to zero.

(b) Ignoring image noise, all intensity profiles across edges in images of polyhedral scenes, where the objects are composed only of planar faces, are simple step functions.

(c) In edge detection, an operator that is optimal for detecting the presence of an edge will also be optimal for locating the position of that edge.

(d) The number of zero-crossings of the second derivative of a function $f(x)$ filtered by a Gaussian with parameter σ is a monotonically decreasing function of σ.

(e) If a problem is well-posed, in the sense of Hadamard, then there exists a numerically stable algorithm for computing the unique solution.

(f) Physical optics suggests that there are two kinds of reflectance, diffuse (i.e., Lambertian) and specular (i.e., mirror-like). Thus, in physics-based computer vision, it is sufficient to model reflectance as a linear combination of diffuse and specular components.

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

(h) The reflectance map, $R(p, q)$, generalizes the definition of the bidirectional reflectance distribution function (BRDF).

(i) In photometric stereo, one often uses paint (or other surface coating) to match reflectance properties between a calibration sphere and objects whose shape is to be analyzed. One could use a white paint or a darker (but otherwise identical) grey paint. For a given configuration of light sources, the identical measurements would be obtained in either case, provided only that the light sources were made appropriately brighter when using the darker grey paint. [Hint: Recall Assignment 5].

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

Question 2:

(a) Let $H = \{h_{ij}\}$ be the coefficients of an $n \times m$ digital filter, $i = 1, \ldots, n$, $j = 1, \ldots, m$. One can think of H as defining a 2-D array (or matrix). What does it mean for the 2-D filter, H, to be separable? Given that H is separable, what is the rank of H? [Hint: Recall that the rank of an $n \times m$ matrix is the number of linearly independent rows (or columns).]

(b) Some critics of Artificial Intelligence argue that humans are unlike machines because machines necessarily are "precise" in their perception. One example cited is the Mueller-Lyer illusion illustrated in Figure 1. By construction, both horizontal lines have equal length. To humans, as you hopefully agree, the lengths appear different. The critics argue that a machine is not subject to the illusion because it necessarily measures the correct length in each case.

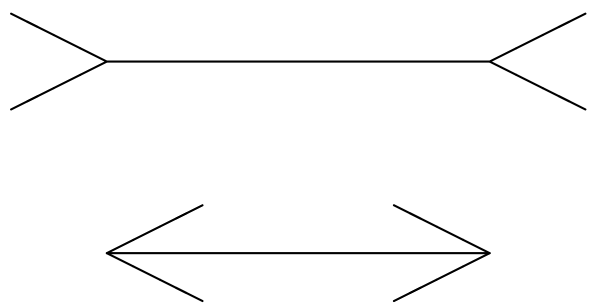

Figure 1: Mueller-Lyer Illusion

It is instructive to examine the Mueller-Lyer illusion in the context of edge detection. Based on your experience with how edge detection works at a corner, argue (one way or the other) whether a machine vision system is subject to the Mueller-Lyer illusion. [Note: It is useful to remember that edge detection is an ill-posed problem so that a smoothing (or regularizing) operator typically is required. How might this requirement influence the estimated lengths of the two horizontal lines?]

(c) Each of the initial NASA Landsat series of satellites carried a multispectral scanner (MSS) imaging system that recorded digital images of the earth's surface with a nominal ground resolution of 78m by 78m per pixel. The ground track of the satellite was not aligned with North-South. Thus, there was significant rotation between the actual coordinate axes of image acquisition and the natural orientation of coordinate axes typically used in mapping systems. In Canada, these Landsat MSS digital images routinely were geometrically corrected to Universal Transverse Mercator (UTM) map coordinates and resampled to a regular 50m by 50m square grid whose axes were aligned with North–South and East–West. From a image processing point of view, what would justify this oversampling? [Hint: Recall Assignment 2 and think of the ratio 78:50 as approximately $\sqrt{2} : 1$).]

(d) Woodham's 1990 formulation of multiple light source optical flow made use of three equations

$$\begin{array}{rcl} E_{1x}\,u \;+\; E_{1y}\,v \;+\; E_{1t} & = & 0 \\ E_{2x}\,u \;+\; E_{2y}\,v \;+\; E_{2t} & = & 0 \\ E_{3x}\,u \;+\; E_{3y}\,v \;+\; E_{3t} & = & 0 \end{array} \qquad (1)$$

How did Woodham get additional equations? What are E_{ix}, E_{iy} and E_{it}, $i = 1, 2, 3$? Under what circumstances are u and v the identical variables in each of the three equations? (i.e., Why are there no subscripts on u and v in Equations (1)?)

(e) Points in shadow cause difficulty in shape–from-shading and in the estimation of optical flow. Distinguish points in "self-shadow" from those in "cast-shadow."

(f) Stauffer & Grimson 2000 model the value of each pixel as a mixture of K Gaussian distributions. The value of K is limited by available memory and computational power. Stauffer & Grimson report using values of K in the range of 3–5. Why do they choose $K > 2$?

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.

Note: The actual final exam will have 5–6 questions to choose from in Part B.

Question 3:

We have considered three approaches to edge detection based on the work of Marr & Hildreth, Torre & Poggio and Canny. In each case, the approach was based on clearly stated design criteria. For the purposes of edge detection, what is meant by:

(a) detection

(b) localization in space

(c) localization in frequency

Canny included as one of his design criteria the elimination of multiple responses to a single edge.

(d) Why is this important in Canny edge detection?

(e) How do Marr & Hildreth handle multiple responses to a single edge? (Hint: Either explain how they handle multiple responses or give an argument for why the problem does not arise with their method).

Gaussian, or Gaussian-like, convolution occurs in much of the work on edge detection that we have studied. Concerning their work, Torre & Poggio state, “This result is the simplest and most rigorous proof that a Gaussian-like filter represents the correct operation to be performed before differentiation for edge detection.”

(f) Compare and contrast the way in which a Gaussian-like filter emerged as the correct operation in the Torre & Poggio approach with how the Gaussian filter emerged in the Marr & Hildreth approach and in the Canny approach.

Question 4:

This problem explores the technique of matched filtering. The goal is to establish that the correlation of an idealized pattern with the input image is the "optimal" matched filter in certain well-defined situations. Let $A = \{a_{ij}\}$ and $B = \{b_{ij}\}$ be two $n \times n$ digital images representing two (idealized) patterns. Suppose further that these patterns have been gray-level normalized so that

$$\sum_{ij} a^2{}_{ij} \ = \ \sum_{ij} b^2{}_{ij}$$

An $n \times n$ digital image $C = \{c_{ij}\}$ is obtained. C is assumed to be an image either of A or B, the problem is to decide which. To make the problem interesting, suppose each pixel of C is perturbed by additive, zero mean, Gaussian noise. That is, the pixels in C are not identically those of A or B but have added to them noise where the noise comes from a Gaussian distribution with zero mean. Further, suppose that the noise acts independently on each pixel in C. Recall that if x is a random variable from a Gaussian distribution with mean μ and standard deviation σ, then the probability density function of x is given by

$$p(x) \ = \ \frac{1}{\sqrt{2\pi}\sigma} \ \exp^{-\frac{(x \, - \, \mu)^2}{2\sigma^2}}$$

Thus, for each pixel in C, the probability of observing that pixel when the pattern is really A is given by

$$p\left(c_{ij}/a_{ij}\right) \ = \ k_1 \ \exp^{k_2(c_{ij} \, - \, a_{ij})^2} \quad \text{where} \quad k_1 \ = \ \frac{1}{\sqrt{2\pi}\sigma} \quad \text{and} \quad k_2 \ = \ -\frac{1}{2\sigma^2}$$

Since the noise is assumed to be independent for each pixel, the probability of observing the whole image C when the pattern is really A is just the product of the probabilities of each c_{ij} given a_{ij}. That is

$$p(C/A) \ = \ \prod_{ij} k_1 \ \exp^{k_2(c_{ij} \, - \, a_{ij})^2}$$

An identical argument gives

$$p(C/B) \ = \ \prod_{ij} k_1 \ \exp^{k_2(c_{ij} \, - \, b_{ij})^2}$$

It seems reasonable to decide that C is an image of A if and only if $P(C/A) \ > \ P(C/B)$.

(a) Prove that $P(C/A) \ > \ P(C/B)$ if and only if $\sum_{ij} c_{ij}a_{ij} \ > \ \sum_{ij} c_{ij}b_{ij}$. [Hint: Instead of comparing $P(C/A)$ and $P(C/B)$ directly as the products of exponentials, it is useful to take logarithms and compare the resulting sums. For $x \ > \ 0$ and $y \ > \ 0$, $x \ > \ y$ if and only if $\ln(x) \ > \ \ln(y)$].

Thus, in order to decide whether C is A or B it is sufficient to compare the correlation of A with C to the correlation of B with C. Said another way, A and B are their own best matched filters.

(b) Suppose C is an $n \times n$ digital image and A and B are $m \times m$ digital patterns with $n \gg m$. Given the result in part (a), describe, as concisely as you can, how you would go about finding instances of patterns A and B in the image C?

(c) What would be the effect on matched filtering, as implemented in part (b) above, if the patterns A and B were not gray-level normalized (i.e., $\sum_{ij} a^2{}_{ij} \neq \sum_{ij} b^2{}_{ij}$)?

(d) Is matched filtering, as implemented in part (b) above, likely to be robust with respect to differences in spatial scale between image and pattern? What about differences in rotation between image and pattern?

(e) Given the task of optical character recognition (OCR), how might you design a font to take best advantage of the results derived in this problem?

Question 5:

This problem demonstrates that, in certain circumstances, photometric stereo allows one to determine information about surface material as well as about surface orientation. As we have seen, the bidirectional reflectance distribution function (BRDF) of an ideal lossless Lambertian material is $f_r = 1/\pi$. In this context, lossless means that 100% of the surface irradiance is reflected. Suppose instead that only a fraction, ρ, $0 \leq \rho \leq 1$, of the irradiance is reflected. Then the BRDF of a Lambertian material is $f_r = \rho/\pi$. The fraction ρ is a reflectance factor that often is referred to as the "albedo."

We now show that three light source photometric stereo is sufficient to determine both the surface orientation and the reflectance factor, ρ, for Lambertian materials, even when ρ varies from point to point on the object surface. Suppose we obtain three images under three different light source conditions. Let

$$\begin{aligned} \mathbf{s}_1 &= [s_{11}, s_{12}, s_{13}]^{\mathrm{T}} \\ \mathbf{s}_2 &= [s_{21}, s_{22}, s_{23}]^{\mathrm{T}} \\ \mathbf{s}_3 &= [s_{31}, s_{32}, s_{33}]^{\mathrm{T}} \end{aligned} \tag{2}$$

be unit column vectors ($^{\mathrm{T}}$ denotes vector transpose) defining the directions of three distant, point sources of illumination. Define the 3×3 matrix, $\mathbf{S}$, by

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \tag{3}$$

Let $\mathbf{E} = [E_1, E_2, E_3]^{\mathrm{T}}$ be the column vector of irradiance values at an image point, (x, y), in each of the three images. Let $\mathbf{n} = [n_1, n_2, n_3]^{\mathrm{T}}$ be the column vector corresponding to a unit surface normal at (x, y) and let ρ be the associated reflectance factor.

(a) Express image irradiance, $E_i(x, y)$, as a function of ρ, $\mathbf{s}_i$ and $\mathbf{n}$ for $i = 1, 2, 3$. [Hint: Recall that the cosine of the angle between two vectors is the dot product divided by the product of the magnitudes. Also, since the absolute strength of the light sources is not specified, keep things simple by assuming the three light sources each have irradiance $E_0/\pi = 1$ where E_0 is the irradiance of the light source measured perpendicular to the beam of incident light].

(b) Combine the results of part (a) into a single matrix equation that expresses $\mathbf{E}$ as a function of ρ, $\mathbf{S}$ and $\mathbf{n}$.

(c) Indicate how the matrix equation of part (b) can be used to solve for ρ.

(d) Indicate how the matrix equation of part (b) and the result of part (c) can then be used to solve for $\mathbf{n}$.

(e) Under what conditions will the solution determined in parts (c) and (d) fail to exist?

Question 6:

In this problem, we explore edge detection at a corner. We use Gaussian smoothing and detect edges based on zero crossings of the second directional derivative along the gradient. To keep the problem tractable, we consider a single, isolated 2-D corner. Let $F(x, y)$ be the 2-D "image" defined by

$$F(x, y) \;=\; \begin{cases} 1 & x > 0 \;\; y > 0 \\ 0 & \text{otherwise} \end{cases}$$

$F(x, y)$ defines a "corner" at $x = 0$, $y = 0$, corresponding to the location where two (unit) step edges meet, one along the positive x axis and the other along the positive y axis. Thus, $F(x, y)$ can be written as

$$F(x, y) \;=\; f(x)\, f(y)$$

where

$$f(t) \;=\; \begin{cases} 1 & t > 0 \\ 0 & \text{otherwise} \end{cases}$$

First, we smooth the image, $F(x, y)$, with a 2-D Gaussian. The scale parameter, σ, of the Gaussian controls the degree of smoothing. But, notice that $F(x, y) = F(sx, sy)$, for all $s > 0$. This means that $F(x, y)$ is self-similar at all scales. Thus, it is sufficient to consider the case $\sigma = 1$. Let $G(x, y)$ be the 2-D Gaussian (with $\sigma = 1$) defined by

$$G(x, y) \;=\; \frac{1}{2\pi}\, \exp^{-\dfrac{x^2 + y^2}{2}}$$

$G(x, y)$ can be written as

$$G(x, y) \;=\; g(x)\, g(y)$$

where

$$g(t) \;=\; \frac{1}{\sqrt{2\pi}}\, \exp^{-\dfrac{t^2}{2}}$$

Note that $g(t)$ is the 1-D Gaussian (with $\sigma = 1$). Recall that convolution commutes so that

$$G(x, y) * F(x, y) \;=\; F(x, y) * G(x, y)$$

($*$ denotes convolution). Also, recall that 2-D and 1-D convolutions are defined, respectively, as

$$F(x, y) * G(x, y) \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x - u, y - v)\, G(u, v)\, du\, dv$$

and

$$f(x) * g(x) \;=\; \int_{-\infty}^{\infty} f(x - u)\, g(u)\, du$$

(a) Show that

$$G(x, y) * F(x, y) \;=\; \phi(x)\, \phi(y) \tag{4}$$

where

$$\phi(t) \;=\; \int_{-\infty}^{t} g(u)\, du$$

Let $\phi'(t)$ denote the derivative of $\phi(t)$. By the fundamental theorem of calculus,

$$\frac{d}{dx}\int_a^x g(u)\,du \;=\; g(x)$$

for any $x > a$ so that

$$\phi'(t) \;=\; g(t) \tag{5}$$

Recall that the second directional derivative along the gradient of a 2-D function, f, is given by

$$\frac{\partial^2}{\partial n^2} \;=\; \frac{{f_x}^2 f_{xx} \;+\; 2 f_x f_y f_{xy} \;+\; {f_y}^2 f_{yy}}{{f_x}^2 \;+\; {f_y}^2} \tag{6}$$

(subscripts denote partial differentiation).

(b) Is $\dfrac{\partial^2}{\partial n^2}$ a linear operator? [Note: No proof is required].

(c) Does $\dfrac{\partial^2}{\partial n^2}(G(x,y) * F(x,y)) \;=\; (\dfrac{\partial^2}{\partial n^2} G(x,y)) * F(x,y)$? [Note: No proof is required].

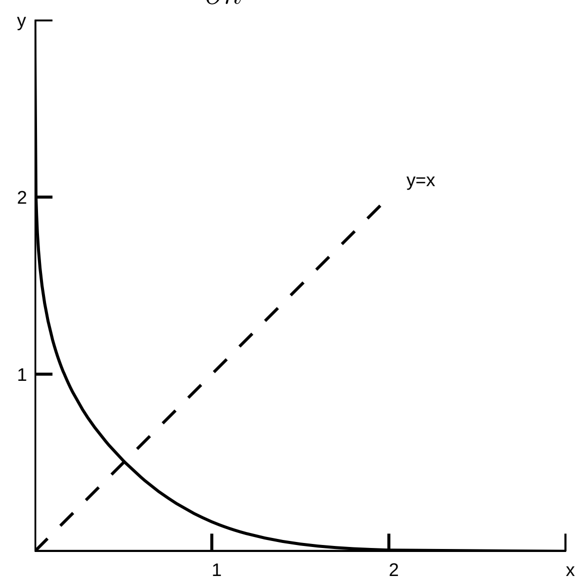

Figure 2: The 2-D corner and its bisector, the line $y = x$. Curved line shows detected edge points displaced near the corner at $x = 0$ and $y = 0$.

Edges are detected at zero crossings of $\frac{\partial^2}{\partial n^2}(G(x,y) * F(x,y))$. One could apply equation (6) to equation (4) to determine zeros directly. This was done to generate Figure 2. Note, however, that the problem is symmetric about the corner bisector (i.e., about the line $y = x$). It should be easy to convince yourself that, owing to symmetry, the zero crossing point corresponding to the corner at $x = 0$, $y = 0$, must lie along the line $y = x$. Also, owing to symmetry, the direction of the gradient is along the line $y = x$. Along the bisector, $G(x,y) * F(x,y) \;=\; \phi(x)\,\phi(x) \;=\; \phi^2(x)$ so that

$$\frac{\partial^2}{\partial n^2}(G(x,y) * F(x,y)) \;=\; \frac{d^2\phi^2(x)}{dx^2}$$

(d) Prove that $\frac{d^2\phi^2(x)}{dx^2} \;=\; 0$ if and only if

$$g(x) \;-\; x\,\phi(x) \;=\; 0 \tag{7}$$

[Hint: Use equation (5)].

Solving equation (7) numerically (to 10 decimal places) gives $x = 0.5060544690$, which we will say is approximately $x = 1/2$. Thus, the corner at point $(0, 0)$ is displaced to the point $(1/2, 1/2)$. If we now let σ be arbitrary, rather than $\sigma = 1$, then the self-similarity argument given above allows us to conclude that the corner at point $(0, 0)$ is displaced to the point $(\sigma/2, \sigma/2)$. Indeed, Figure 2 is self-similar at all scales and we can interpret scale on the x and y axes to be in units of σ.

Fredrik Bergholm of Sweden analyzed this and other special cases of 2-D image features. In particular, he also considered corners where the interior angle at the corner ranged from 0 to π radians, in addition to the $\pi/2$ case considered here.

(e) Use your intuition, rather than formal analysis, to argue whether the displacement at a corner with interior angle $\pi/4$ radians is going to be more (or less) than $(\sigma/2, \sigma/2)$.

Question 7:

In this problem, we develop a "direct" method to determine time to contact relative to a planar surface. The method is based on analysis of the 2-D motion field resulting from translational motion under perspective projection and the constant brightness assumption.

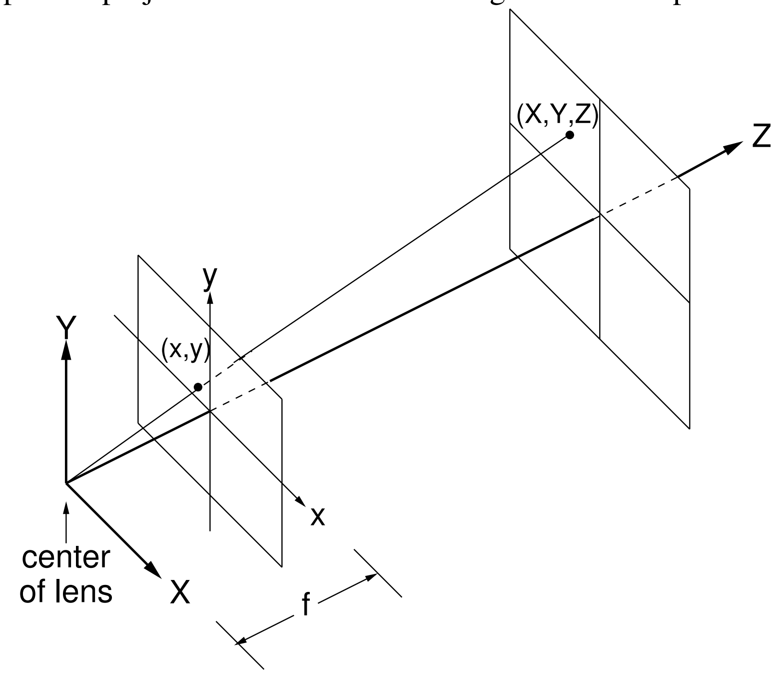

Figure 3: Perspective Projection

First, we deal with perspective projection. Figure 3 illustrates. 3-D point (X, Y, Z) projects to 2-D image point (x, y) where

$$x \;=\; f\,\frac{X}{Z} \;\text{ and }\; y \;=\; f\,\frac{Y}{Z} \tag{8}$$

Here, uppercase coordinates refer to quantities in the 3-D world and lowercase coordinates refer to quantities in the 2-D image plane. Let

$$[u, v] \;=\; \left[\frac{dx}{dt}, \frac{dy}{dt}\right] \qquad \text{and} \qquad [U, V, W] \;=\; \left[\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt}\right]$$

$[u, v]$ is the 2-D motion field. $[U, V, W]$ is the 3-D velocity of a point on the planar surface relative to the camera center of lens (COL) (which is opposite to the motion of the camera relative to the planar surface).

Differentiating the perspective projection Equations (8) with respect to time, we obtain

$$u \;=\; f\left(\frac{U}{Z} \;-\; \frac{X}{Z}\frac{W}{Z}\right) \;\text{ and }\; v \;=\; f\left(\frac{V}{Z} \;-\; \frac{Y}{Z}\frac{W}{Z}\right)$$

which can be re-written as

$$u \;=\; f\left(\frac{U}{Z} \;-\; \frac{x}{f}\frac{W}{Z}\right) \;\text{ and }\; v \;=\; f\left(\frac{V}{Z} \;-\; \frac{y}{f}\frac{W}{Z}\right)$$

or

$$u = \frac{1}{Z}(f\,U - x\,W) \text{ and } v = \frac{1}{Z}(f\,V - y\,W) \tag{9}$$

Consider the simple case of translational camera motion along the optical axis towards a planar surface oriented perpendicular to the optical axis.

(a) For this simple case, $U = V = 0$. Why?

Time to contact (TTC) is defined as the time remaining before the camera COL reaches the planar surface being viewed if the relative motion between the camera and the surface continues without change (i.e., if the relative motion continues at constant velocity).

Thus, time to contact, T, is the distance (to contact), Z, divided by the velocity of the camera COL, $-W$. That is,

$$T = -\frac{Z}{W} \tag{10}$$

Subsequent analysis will develop a method to solve for the inverse of TTC. Denote this inverse as C where

$$C = \frac{1}{T} = -\frac{W}{Z} \tag{11}$$

Suppose there is some object in the planar surface with (linear) size, S. Let the size of its image be s. Under perspective projection, we obtain

$$\frac{s}{f} = \frac{S}{Z}$$

which we re-write as

$$s\,Z = f\,S \tag{12}$$

Now let's differentiate Equation (12) with respect to time.

(b) The derivative of the right-hand-side (RHS) of Equation (12) is 0. Why?

If we apply the product rule for differentiation to the left-hand-side (LHS) of Equation (12) and combine this with the result of part (b) we obtain

$$s\,W + Z\,\frac{ds}{dt} = 0 \tag{13}$$

(c) Show that

$$T = s \bigg/ \frac{ds}{dt} \tag{14}$$

[Hint: Combine the results of Equations (10) and (13)

Equation (14) suggests the possibility of determining TTC based on image measurements alone. The challenge, of course, is to develop a method that is robust.

Now, consider image brightness, $E(x, y, t)$. As with Horn & Schunck 1981, apply the chain rule for differentiation to $E(x, y, t)$ to obtain

$$\frac{dE}{dt} \;=\; E_x\, u \;+\; E_y\, v \;+\; E_t$$

where subscripts denote partial differentiation. Suppose $\dfrac{dE}{dt} = 0$. This is the "constant brightness assumption."

(d) Explain, in simple English, what the "constant brightness assumption" means.

(e) Is the "constant brightness assumption" reasonable for the simple case of translational camera motion along the optical axis towards a planar surface oriented perpendicular to the optical axis? Briefly justify your answer.

With the constant brightness assumption, we obtain the (classic) optical flow constraint equation

$$E_x\, u \;+\; E_y\, v \;+\; E_t \;=\; 0 \tag{15}$$

Recalling the results in part (a) and Equations (9), we substitute $u = -x(W/Z)$ and $v = -y(W/Z)$ into Equation (15) to obtain

$$-\frac{W}{Z}(x\, E_x \;+\; y\, E_y) \;+\; E_t \;=\; 0$$

or

$$C\, G \;+\; E_t \;=\; 0$$

where C is the inverse of TTC, as in Equation (11), and G is the "radial gradient" $(x\, E_x \;+\; y\, E_y)$.

Finally, we formulate a least squares method to minimize

$$\sum\,(C\, G \;+\; E_t)^2 \tag{16}$$

Minimizing Equation (16) is straightforward. We differentiate with respect to C and set the result to zero to obtain

$$\sum(C\, G \;+\; E_t)\, G \;=\; 0 \tag{17}$$

or

$$C \sum G^2 \;=\; -\sum G\, E_t \tag{18}$$

so that C, the inverse of TTC, is

$$C \;=\; -\frac{\sum G\, E_t}{\sum G^2} \tag{19}$$

(f) We have not been specific about what region of the image to sum over in Equations (16)–(19). Over what region of the image should we sum? Briefly justify your answer.

The method derived here can be generalized to allow translational motion other than along the optical axis toward a planar surface at an arbitrary orientation to the optical axis. The mathematical details become more complicated (and thus beyond the scope of an examination question) but the essence of the method remains.

(g) For the simple case of translational camera motion along the optical axis towards a planar surface oriented perpendicular to the optical axis, is the estimation of TTC via Equation (19) likely to be robust? Briefly justify your answer.