CS/RBE 549 Computer Vision
Fall 2022

Name:________________________________

Exam #2

This examination is open book and open notes. Please complete individually.
Partial credit will be given, but you must justify your work.

The examination will permit submission within 72 hours once it begins.Good luck.

Problem 1: /30
Problem 2: /30
Problem 3: /30
Problem 4: /10

Total: /100

1. **Iterative Optical Flow (30%).** To compute optical flow, we have learned an iterative method to update $u(x,y), v(x,y)$ at each iteration according to

$$\begin{bmatrix} u(x,y) \\ v(x,y) \end{bmatrix}^{\text{new}} = \begin{bmatrix} \lambda I_x^2 + 4 & \lambda I_x I_y \\ \lambda I_x I_y & \lambda I_y^2 + 4 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{n \in \text{neighbors}(x,y)} u^{\text{old}}(n) - \lambda I_x I_t \\ \sum_{n \in \text{neighbors}(x,y)} v^{\text{old}}(n) - \lambda I_y I_t \end{bmatrix}$$

Consider a local coordinate frame (x', y') where x' is aligned with the image gradient and y' is perpendicular to the image gradient. Likewise, $(u', v') = \left(\frac{dx'}{dt}, \frac{dy'}{dt}\right)$ are the image velocities in this frame. In this coordinate frame,

$$I_{x'} = \sqrt{I_x^2 + I_y^2} \text{ and } I_{y'} = 0$$

Show that the update equations

$$\begin{bmatrix} u'(x,y) \\ v'(x,y) \end{bmatrix}^{\text{new}} = \begin{bmatrix} \lambda I_{x'}^2 + 4 & \lambda I_{x'} I_{y'} \\ \lambda I_{x'} I_{y'} & \lambda I_{y'}^2 + 4 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{n \in \text{neighbors}(x,y)} u'^{\text{old}}(n) - \lambda I_{x'} I_t \\ \sum_{n \in \text{neighbors}(x,y)} v'^{\text{old}}(n) - \lambda I_{y'} I_t \end{bmatrix}$$

can reduce to

$$u'^{\text{new}}(x,y) = \bar{u}'^{\text{old}} - \frac{{I_{x'}}^2 \bar{u}'^{\text{old}} + I_{x'} I_t}{I_{x'}^2 + \frac{4}{\lambda}}$$

$$v'^{\text{new}}(x,y) = \bar{v}'^{\text{old}}$$

Please illustrate intermediate steps for partial credits.

2. **Object Representation (30%).** An object can be represented by its boundary $(x(s), y(s)), 0 \leq s \leq S$ where S is the length of the object's boundary and s is the distance along that boundary from some arbitrary starting point. We can combine x and y into a single complex function $x(s) = x(s) + jy(s)$. The Discrete Fourier Transform (DFT) of z is

$$Z(k) = \sum_{s=0}^{S-1} e^{-2\pi j \frac{ks}{S}} z(s), \qquad 0 \leq k \leq S-1$$

We can use the coefficients $Z(k)$ to represent the object boundary.

The limit on s is $S - 1$ because for a closed contour $z(s) = z(0)$.

The Inverse Discrete Fourier Transform is

$$z(s) = \frac{1}{S} \sum_{k=0}^{S-1} e^{+2\pi j \frac{ks}{S}} Z(k), \qquad 0 \leq k \leq S-1$$

a. (5%) Suppose that the object is translated by $(\Delta x, \Delta y)$, that is $z'(s) = z(s) + \Delta x + j\Delta y$. How is $Z'(k)$ as the DFT of z' related to $Z(k)$?

b. (10%) Suppose that the object is scaled by integer constant c, that is $z'(s) = cz(s)$. For simplicity, let's assume that $S' = S$. How is $Z'(k)$ as the DFT of z' related to $Z(k)$?

c. (10%) What object has $z(s) = \left[x_0 + R\cos\frac{2\pi s}{S}\right] + j\left[y_0 + R\sin\frac{2\pi s}{S}\right]$? Please sketch it.

d. (5%) What is $Z(k)$ corresponding to $z(s)$ from Part c? Hint: Most coefficients are 0.

3. **Stereo via Singular Value Decomposition (SVD) (30%):** Assume the usual stereo geometry, where the left and right cameras are offset by baseline $\vec{B}$ that is perpendicular to the common focal vector $\vec{F}$. Then the stereo imaging equations are

$$\vec{X}_L = \frac{|\vec{F}|^2}{\vec{F}\cdot\vec{X}^W}\left(\vec{X}^W + \frac{\vec{B}}{2}\right), \qquad \vec{X}_R = \frac{|\vec{F}|^2}{\vec{F}\cdot\vec{X}^W}\left(\vec{X}^W - \frac{\vec{B}}{2}\right)$$

In the presence of imaging errors and noise, these equations might not hold exactly. They can be approximated by

$$\vec{X}_L = \frac{|\vec{F}|^2}{\vec{F}\cdot\vec{X}^W}\left(\vec{X}^W + \frac{\vec{B}}{2}\right) \approx \vec{0}, \qquad \vec{X}_R = \frac{|\vec{F}|^2}{\vec{F}\cdot\vec{X}^W}\left(\vec{X}^W - \frac{\vec{B}}{2}\right) \approx \vec{0}$$

a. (15%) Show that these equations can be written as a 4x4 matrix operating on a column vector in homogeneous coordinates.

$$\begin{bmatrix} -f & 0 & x_L & -f\frac{b}{2} \\ 0 & -f & y_L & 0 \\ -f & 0 & x_R & f\frac{b}{2} \\ 0 & -f & y_R & 0 \end{bmatrix} \begin{bmatrix} x^W \\ y^W \\ z^W \\ 1 \end{bmatrix} \approx \vec{0}$$

Hint: combine the approximate imaging equations into a single matrix equation, multiply to eliminate the denominators, and simplify but not necessarily in that order.

b. (5%) The above equation can be written as $A\tilde{X}' \approx \vec{0}$. We can use SVD to find the singular vector $\tilde{X}'$that minimizes $|A\bar{X}'|^2$ subject to $|\bar{X}|^2 = 1$. Express world point $\vec{X}^W = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in terms of $\tilde{X}' = \begin{bmatrix} x' \\ y' \\ z' \\ w' \end{bmatrix}$.

c. (10%) When $y_L = y_R$, show that part a. gives $z^W = \frac{fb}{d}$, where d is the disparity.

4. **Vision Package (10%):** Robot SeeWee is so excited to have a new set of "eyes" – LiDAR (Light Detection and Ranging) sensor, which is a nice addition to the original Infra-Red sensor. SeeWee's thermal Infra-Red eyes are presumably sensitive to photons in the infra-red spectrum which are emitted uniformly in all directions from warm objects. Objects under Infra-Red eyes tend to lack distinct features because nearby parts of a single object have similar temperatures, SeeWee's LiDAR sensor emits a near-infrared laser and measures the reflection time from the object. It generates high-resolution maps of object surface structures.

   Please fill the table below by *briefly* discussing *one to three* challenges with LiDAR vision alone, Infra-Red vision alone, benefits of combination, and challenging situations that require additional vision capacities.

	Challenges with LiDAR Vision	Challenges with Thermal Infra-Red Vision	Benefits of Imaging the world with both LiDAR and Thermal Infra-Red Vision	What other Vision options SeeWee will benefit from.
#1				
#2				
#3				