CSC 47100-1EF (51286)

Computer Vision

Computer Science Department, The City College of New York

EXAM

April 25, 2022

Name: ____Ismail_Akram_____ **ID # (last 4 digits): ______8834_________**

Score: _________of 100

I. Multiple Choices (20 points)

(2x10 = 20 points – write only one of the labels (A, B, C, etc.) for each question from 1 to 10)

1. D
2. D
3. D
4. C
5. A
6. E
7. A
8. C
9. D
10. D

II Short Answers (80 points)

1. **Stereo Relation (20 points)**

Solving for R and T as functions of the extrinsic parameters of the two cameras.

(1) We are using Linear Algebra here.

a. Solving for R (in camera coordinates):

$$P_r = RP_l + T$$
$$P_r - T = RP_l$$
$$R = \frac{P_r - T}{P_l} \dots (a)$$

b. Solving for T (in camera coordinates):

$$P_r = RP_l + T$$
$$P_r - RP_l = T \dots (b)$$

Now we need to express this as functions of the extrinsic parameters of the two cameras. We sub in for P_r and P_l:

Using (a):

$$R = \frac{P_r - T}{P_l}$$
$$R = \frac{(R_r P + T_r) - T}{R_l P + T_l} \dots simualtaneous\ equation\ 1$$

Using (b):

$$T = P_r - RP_l$$
$$T = (R_r P + T_r) - R(R_l P + T_l) \dots simualtaneous\ equation\ 2$$

Combining the two equations will net us our R and T:
1 … I need to move to Question 2 but I basically needed to solve for R and T.

2. **Stereo Accuracy (20 points)**

1. $Z = \frac{fB}{d}$

   Solving for disparity $d = (fB)/Z$

Taking the partial derivative of 'Z' w.r.t 'd' gives us the error estimation equation:

$$\left|\frac{\partial Z}{\partial d}\right| = \frac{\partial\left(\frac{fB}{d}\right)}{\partial d}$$
$$= \frac{fB\ *\ \partial d^{-1}}{\partial d}$$
$$= fB\ * - d^{-2}\ * \frac{\partial d}{\partial d}$$
$$= \left|-\frac{fB}{d^2}\right|$$
$$= \frac{fB}{d^2}$$

$$So\ then: \partial Z\ = \frac{fB}{d^2}\ \partial d$$

But the disparity is dependent on the baseline and focal length.

To rectify this issue: replace 'd' in terms of independent variables 'f' 'B' and 'Z'.

$$\partial Z\ = \frac{fB}{d^2}\ \partial d$$
$$= \frac{fB}{\frac{fB^2}{Z}}\ \partial d$$
$$= \frac{Z^2}{fB}\ \partial d$$

We can then derive the equation for depth error as a function of dB as well by taking the partial derivative with respect to B.

$$\left|\frac{\partial Z}{\partial B}\right|$$
$$= \frac{\partial(fB)}{d}$$
$$= \frac{f}{d}\partial Z$$

$$= f/d\ \partial B$$

Substitute d: $\partial Z\ =\ f/(fB)/Z\ \partial B\ \ =\ Z/B\ \partial B$

2. Discussion

For $dB\ =\ 0$ and $dd\ !=\ 0$

We have:

Baseline length: as B increases, the depth error decreases (inverse relationship).

Focal length: as f increases, the depth error decreases (inverse relationship)

Depth: as depth increases, so will depth error (proportional relationship).

3. **Motion Basics (20 points)**

a.
Difference:
Image: only one rigid and relative motion between camera and scene object.
2D vector field of velocities induced by relative motion.
Can be viewed as a disparity map of the two frames captured at two camera locations consequently. (moving camera).

Optical flow: brightness of moving objects are constant (constantly equation).
Relation of apparent motion with spatial and temporal derivatives of the image brightness.
We also have the aperture problem: only the component of the motion field in the direction of the spatial image gradient can be resolved.

a. Obtaining shape from motion needs work. We're assuming rigidity of an object, so then we reconstruct it from its motion (or from the camera's motion).

If the camera is moving forward, the image flows contains the FOE point.

If object is stationary, then all the velocities of the images will meet at the FOE.
Moving objects will have different direction of image flow.
The depth z of a point can be obtained from its horizontal displacement in space and velocity of the observer.

Simple Example: (single object moving through a fixed background) using a fixed camera position.

We can find the FOE for that object as a single point where the object comes from.
Let's define velocity of the moving object as such:

$$u = \frac{dx}{dt}$$

$$v = \frac{dy}{dt}$$

$$w = \frac{dz}{dt}$$

We define perspective as such on the image plane:

$$x` = f * \frac{x}{z}$$

$$y` = f * \frac{y}{z}$$

For simplicity, let f = 1, focal length = 1.

Combining these equations:
using $[x_i, y_i, z_i]$ moving with a constant velocity in 3D space. The after time t we have:

$$x` = \frac{x_i + u * t}{z_i + w * t}$$

$$y` = \frac{y_i + v * t}{z_i + w * t}$$

So then we get motion by letting $t \rightarrow infinity$: so we can eliminate $[x_i, y_i, z_i]$ and get:

$$x` = \frac{u}{w}$$
$$y` = \frac{v}{w}$$

Finally, this is our FOE. It is a fixed point for movement with a constant velocity.

4. Vision System (20 points)

Video motion typically deals with video sequences. We can use Video Encoding and Compression in a MPEG 4 file to for example survey traffic:

So the main application is: surveillance of traffic speed.

Main components would be video surveillance, and cameras to triangulate speed.

Video surveillance simply monitors moving traffic. The important component is triangulating the speed by capturing various frames of distance of cars and measuring displacement. i.e. how quickly a vehicle moved from point a to b.

This would be implemented in a moving CCTV that can track the vehicles (so perhaps a mobile robot). Although a stationary camera with a wide FOV across the road will suffice.

We can also upgrade this system to check for drunk/reckless driving by using HCI using Human Gesture. We would need Deep Learning and Neural Networking (perhaps YOLO NN which is what I'm using for the project) (Artificial Intelligence would reduce the number of people needed to manually monitor) of various crash/reckless driving examples and feed it into our algorithm.

For safety, we would also need a industrial inspection (this can be done with a flagging system to alert a security guard) so as to not wrongly accuse drivers and have human monitoring. Most of the surveillance would be delegated to the camera itself.