Carson Stevens - CSCI437 Homework 1

Sunday, September 8, 2019 12:08 PM

Hints:

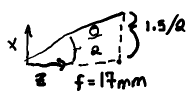
1. You will have to combine transformations; i.e., calculate the transformation from the camera to the world as VMHVMH .
2. The first two vertices of the cube, the ones with world coordinates (X,Y,Z) = (0,0,0) and (1,0,0), project to pixel locations (x,y) = (252, 240) and (301,255), rounded to the nearest pixel.

Turn in:

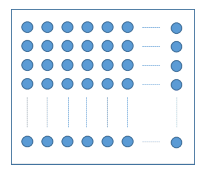
1. Your MATLAB program listing, with comments.
2. (In your PDF File with the rest of your answers:)
3. A description of your method of solution.
4. The image coordinates of the 8 projected points.
5. The 2D wireframe image
6. A plot of the 3D scene

1. (25 pts) We can treat the human fovea as a square sensor array of size 1.5 mm x 1.5 mm, containing about 337,000 cones (sensor elements) (See the figure below). Assume that the space between cones is equal to width of a cone, and that the focal length of the eye is 17 mm.

   1.1. What is the field of view (in degrees) of the human fovea?



$$FOV = 2\tan^{-1}\left(\frac{w/2}{f}\right) = 2\tan^{-1}\left(\frac{1.5/2}{17}\right) \approx \boxed{5.0522^\circ}$$



   1.2. Estimate the distance from Brown Hall to the top of South Table Mountain (you can find this using a map, or a webtool such as Microsoft Bing Maps, or Google Earth). What is the minimum size object that you can see with the naked eye on top of the mountain? Can you see a person on top of the mountain? Assume for simplicity that size of the image of the object must cover at least two receptors (cones).

- To find cones (N x N): $\sqrt{\sim 337{,}000} = 580 \text{ cones} \times 580 \text{ cones}$
- To find cone diameter: $C_d = \frac{1.5m}{580\underbrace{(2)}_{\text{space}}} = 0.0013mm$
- To find min size of cones to see: $d_s = 2(C_d)\cdot\underbrace{2}_{\text{space}} = 4(0.0013) = 0.0052mm$
- To find min size of object to be seen *(Distance from Brown to South Table ~ 1080m)* :

$$\frac{e_d}{d} = \frac{d_s}{f} \Rightarrow e_d = \frac{d(d_s)}{f} = \frac{1080(0.0000052)}{0.0017} \approx \boxed{0.330353m}$$

Since the minimum size we can see is ~ 0.330353m and a human is bigger, we should be able to see a human

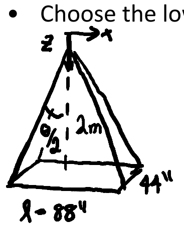
2. (25 pts) A pool-playing robot uses an overhead camera to determine the positions of the balls on the pool table. Assume that:

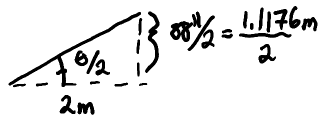
- We are using a standard billiard table of size 44" x 88".
- We need at least 100 square pixels per ball to reliably determine the identity of each ball.

- The center of the ball can be located to a precision of one pixel in the image.
- We need to locate the ball on the table to an accuracy of one cm.
- We are going to mount the camera on the ceiling, looking straight down. The distance from the camera to the table is 2 m.
- Billiard Ball Dimensions: 57mm diameter

Determine a configuration of the camera resolution and lens FOV that will meet these requirements. Assume that you can choose from the following parts:

- Lenses with field of view 30, 60, 90 degrees
- Cameras with resolutions of 256x256, 512x512, or 1024x1024 pixels
- Choose the lowest resolution that will meet the requirements.



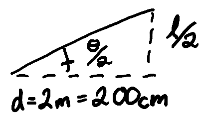


- To find FOV of camera:

$$\theta = 2\tan^{-1}\left(\frac{\frac{1.1176}{2}}{2}\right) \approx 58.3929°$$

*To cover the whole table, a FOV of **60°***

- To find NxN camera resolution:
  - To find length of image for camera:



$$\frac{l}{2} = d\tan\left(\frac{\theta}{2}\right)$$

$$l = 2\cdot 200\tan(30°)$$

$$l \approx 230.940\text{cm}$$

  - To find diameter of ball in pixels *(57mm => 5.7 cm)* :

$$l\cdot d_B = N\cdot 5.7 \qquad d_B = \frac{N\cdot 5.7}{l} = \frac{N\cdot 5.7}{230.940} \approx 0.024681\cdot N_{px}$$

  - To fine area of ball in image:

$$A = \pi r^2 = \pi\left(\frac{0.024681}{2}\cdot N\right)^2 \approx 4.7845\times10^{-4}\cdot N^2\ px^2$$

  - To account for ball constraint *(100px / Area ball)* :

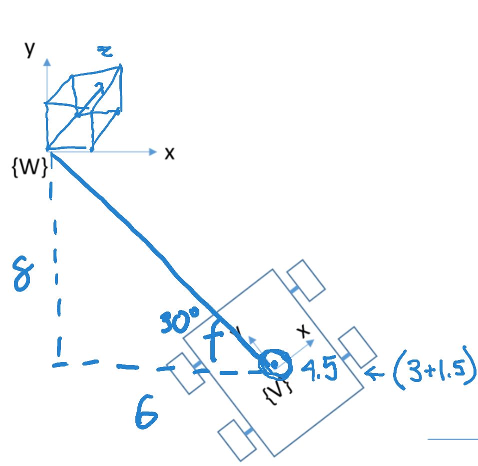
$$\frac{100px}{4.7845\times10^{-4}\cdot N^2} \rightarrow N = \sqrt{\frac{100}{4.7845\times10^{-4}}} = 457.1719_{px}$$

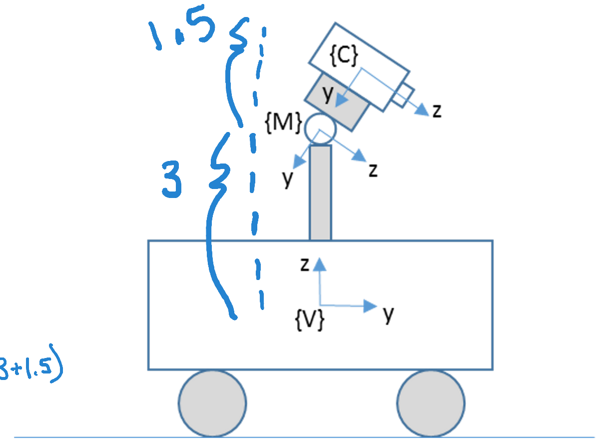
*To encompass the ~457.1719px, must use camera resolution of **512px x 512px***

3. (50 pts) A vehicle {V} is positioned at (6,-8,1) with respect to the world {W}. It is rotated by 30 degrees about the world Z axis, which points up. The figure below shows a top down view of the scene.

   A camera {C} is mounted on a rotational mount {M} on the vehicle, as shown in the figure below. The mount {M} is positioned directly above the vehicle origin at a distance = 3. It is tilted down by 30 degrees. The camera is rigidly attached to the mount. It is positioned directly above the mount origin at a distance = 1.5.

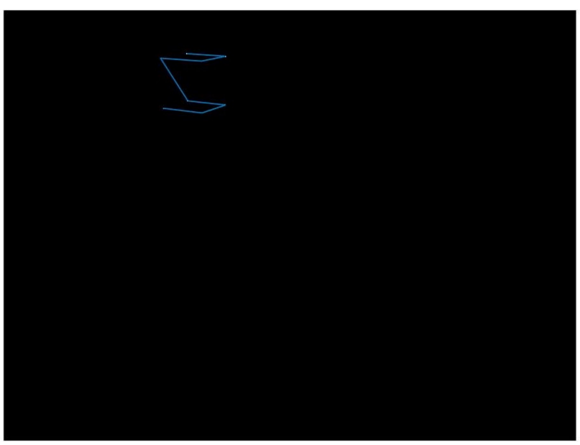
A cube has vertices in world coordinates: (0,0,0), (1,0,0), (1,1,0), (0,1,0), (0,0,1), (1,0,1), (1,1,1), (0,1,1). Using MATLAB, generate an image of a wireframe model of the cube as if were seen by the camera, similar to the figure below, and a 3D plot showing the poses of the camera, mount, vehicle, and cube. Assume a pinhole camera model, with focal length = 600 pixels, where the image size is 640 pixels wide by 480 pixels high.

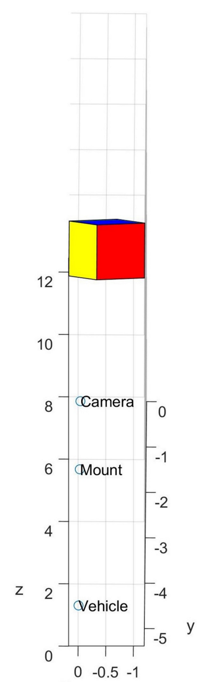




Wire frame model (3D => 2D)

3D => 3D





Method: I first tried to visualize it and then I started to make the Rotation Matrices and apply the translation to get the Homogeneous coordinates. From there, use that, K, and M to Find the projected points on the image or plot

Project Points 3D

X	Y	Z
-1.19615242270663	-4.09807621135332	11.6961524227066
-0.330127018922194	-3.84807621135332	11.2631397208144
0.169872981077806	-4.28108891324554	12.0131397208144
-0.696152422706633	-4.53108891324554	12.4461524227066
-1.19615242270663	-4.96410161513776	11.1961524227066
-0.330127018922194	-4.71410161513776	10.7631397208144
0.169872981077806	-5.14711431702998	11.5131397208144
-0.696152422706633	-5.39711431702998	11.9461524227066

Projected Points 2D

X	Y
110	179
115	222
106	248
102	206
54	176
57	22
52	249
49	205