CIS 580, Machine Perception, Spring 2016
Homework 1
Due: 2016.02.8. 11:59AM

Instructions. This is an individual assignment. It contain written part and programming part. Written part should be in PDF form, programming part should contain all the source code and output files. Package all the files in zip form and submit it to Canvas.

1 Camera Model, Focal Length and Measurement

1. Camera sensor size. Pick a digital camera you have, for example your cell phone. What is the width (mm) and height (mm) of CCD of you camera? Include the web link of the reference.
2. Pixel size. What is the sensor pixel resolution? Compute the the size of a pixel in mm.
3. Focal length. What is the focal length of your phone camera? You can find this answer by looking up the specification of your camera. You can also compute it by finding the field of view (FOV) of your camera, and size of the CCD. Are they similar?

   Autofocus will change your focal length slightly, we can ignore this change for this problem.
4. Measurement. Capture an image of Levine building and measure the distance between the capturing location and the building. Calculate the height of the Levine building with detailed procedure.

   Hint: you can approximately measure distance by using Google map or counting the number of your steps.

2 Locate camera optical center by converging lines

The camera optical center is physically situated somewhere inside the camera, between the front of the camera lens and the image sensor. Locating camera optical center is hard but useful in certain applications such as panoramic photography. In this question, using a set of radiating lines, we will locate optical center relative to the front plane of the camera lens.

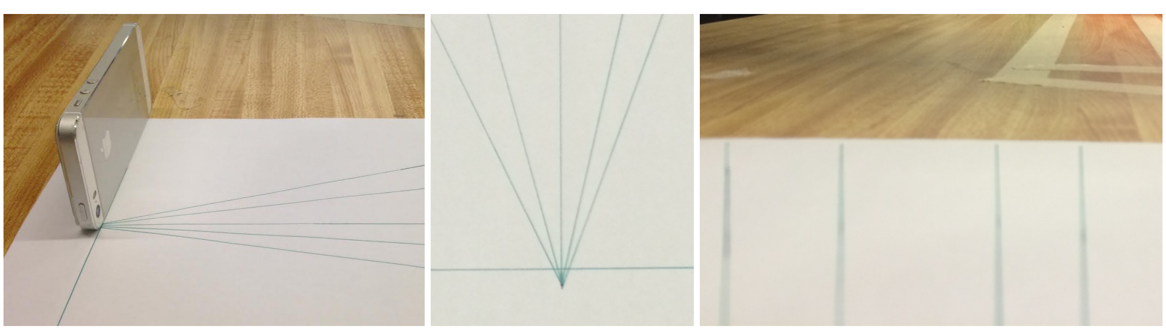

Figure 1: Locate camera optical center by converging lines. Place your cell phone camera (Left) and trace lines on the paper (Center) so the projection appears to be vertical (Right).

1. Place your cell phone camera vertically on a large piece of horizontal paper as shown in figure 1. Use a long ruler to draw a set of lines on the paper that appear perfectly vertical in the image. These lines should appear to be radiating out from a point which is the optical center.

   Mark the line position of the front plane of the camera lens. Trace the radiating line to find its intersection, and mark it as the optical center. Measure the distance, d_c, from optical center to the lens front.

   Alternative Method. Since your cell phone camera is very small, it might be difficult to carry out the method above precisely. Alternatively, you first use a long ruler to draw a set of lines radiating out from a fixed point. Given the lines, carefully place the camera so that each line would appear perfectly vertical when captured by the camera. Mark the line position of the front plane of the camera lens, and measure the distance from that line to the fixed converging point as d_c.

   Hint: Check Figure 2 for camera projection of this task.

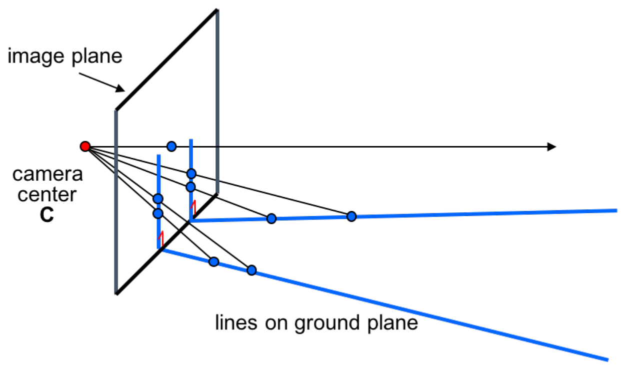

Figure 2: Camera Projection.

2. Given the lines and camera position in 2.1, if the camera is tilted forward, will the lines appear diverge (radiating out) or converge (radiating in) in the picture? Explain your answer intuitively, instead of experimentally.

3. We have another camera (with a different focal length), and place it so that its lens front lines up with the previous iphone. We took a picture of the same set of lines, and the resulting image is shown in Figure 3. The lines appear to be diverging in this example.

   What can we deduce about the location of the new camera's optical center? Is the optical center of the new camera roughly located at position **A** or **B** in Figure 3?

3 Estimate the height of objects

We can estimate the height of any object on the ground using three measurements in the image: 1) horizon line, 2) vanishing point in Z (perpendicular to the ground plane), and 3) a known object height on the ground. Using the method described in the lecture to complete this exercise.

1. Take a picture of Levine building. Include an object, or friend, with a known height in the picture. Make sure the bottom and top of the object (or your friend) is in the field of view. And try to keep the image plane vertical to the ground plane, so the vertical vanishing point remains at infinity.

2. Compute two vanishing points by intersecting parallel lines on the ground plane or on the building facade. Verify the parallel lines on the ground or facade converge to the same vanishing point in the image.

3. Compute and draw the horizon (a vanishing line) in the image.

4. Compute the vanishing point in the Z axis using vertical lines on the facade.

5. Compute the height of the front door of Levine building in `mm`.

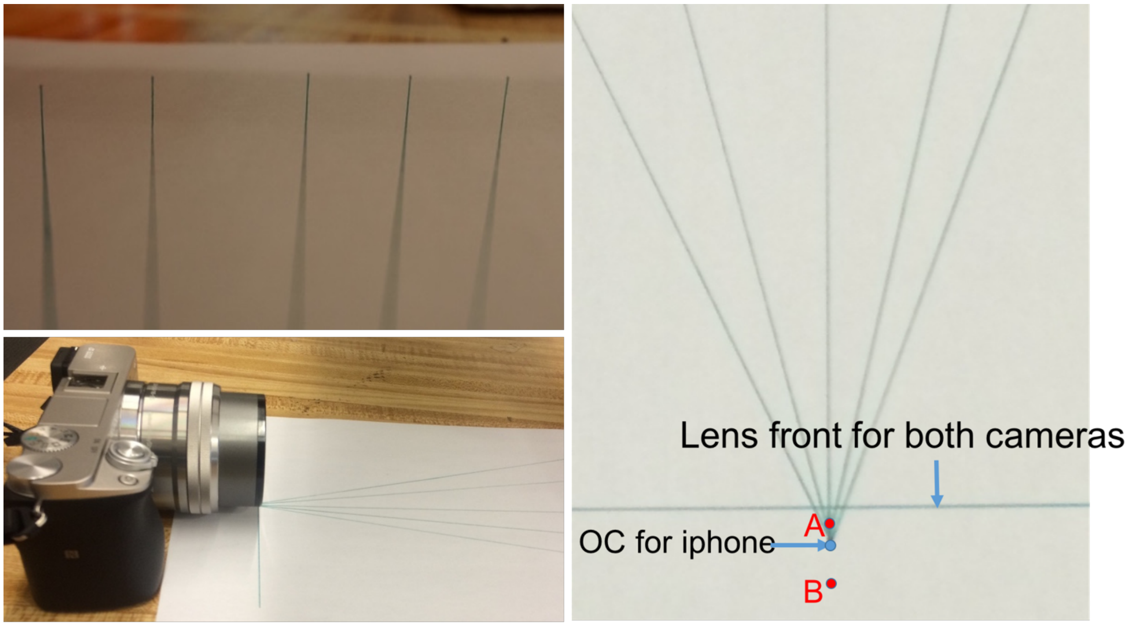

Figure 3: Locate camera optical center by converging lines.

6. Show how to measure the height of the cameraman using your or new image. Note that the image was taken such that the camera Z axis is parallel to the ground plane and the camera Y axis are aligned with the normal vector of the ground plane.

7. If we want you to be of the same height as the front door in picture (have same number of pixels in height), where should you stand? Specify the place in picture. To have a better visual result, you could also crop yourself out, adjust the scale and paste it back onto the image.

4 Dolly Zoom: depth compression due to focal length change

Dolly Zoom is used by filmmakers to create a sensation of vertigo, a "falling-away-from-oneself feeling". It keeps the size of an object of interests constant in the image, while making the foreground and background objects appear larger (or smaller) by adjusting focal length and moving the camera. In this problem, we will compute the camera focal length and position during a Dolly Zoom. We will render a Dolly Zoom video using a simple synthetic scene. Figure 5 shows its top view of the simple synthetic scene and the camera placement. One sample video of the Dolly Zoom is included on the course wiki page.

There are three objects in the scene, denoted as A (green cube), B (triangular pyramid) and C (blue cube). We will use the following setting:

- The image size is 1920×1080.
- The image plane is parallel to the xy plane. The horizontal direction is x axis. The vertical direction is y axis.
- The camera center, denoted as O_c, is moving on the z axis and oriented towards the positive z axis.
- In the first frame, the camera center is located at the origin, with focal length $f_o = 400$ (in pixel unit). Figure 6 (also see /code/start.png) shows the image captured for the first frame.

1. Let $\mathbf{p}=(p_x, p_y)$ denotes the position of each vertex in the image, and d its 3D depth (distance to the xy plane).

   The 3D position (X, Y, Z) for each visible vertex in Figure 6 can be computed as a function of (p_x, p_y), d and f_o. Write down the equation for this function.

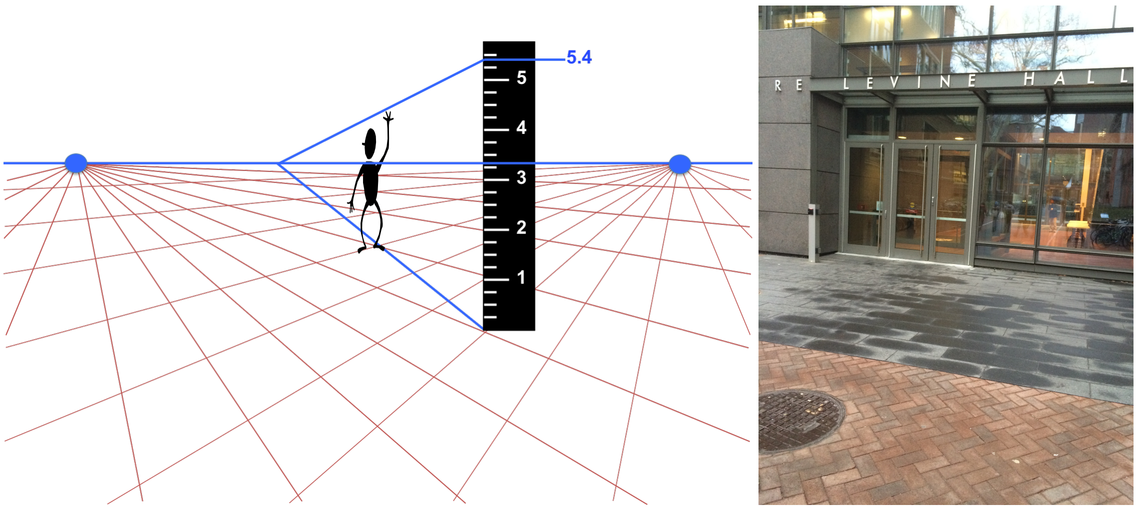

Figure 4: Single View Metrology: estimating the height of objects using a known reference object.

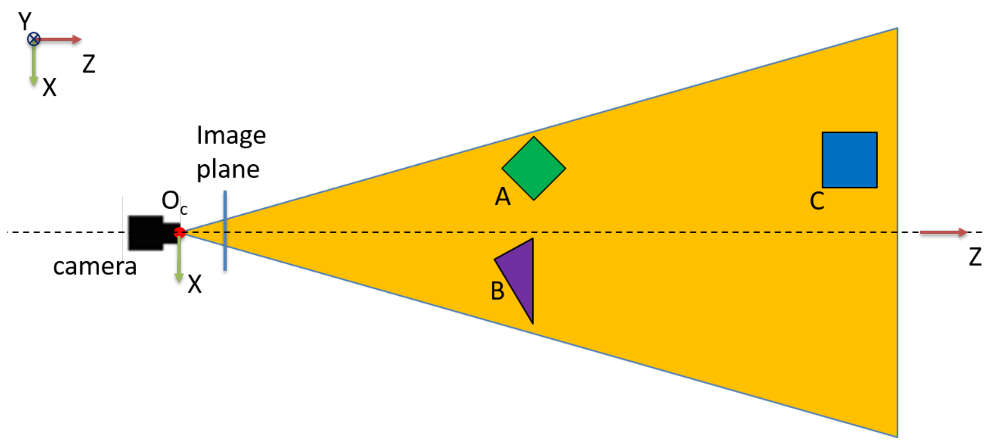

Figure 5: Top view of the synthetic scene.

Complete the Matlab function **unproject(p, d, f)**, and run **generate_points.m** to generate the 3D positions of the vertices.

The image positions (p_x, p_y) and 3D depth values d for the vertices are given in **start.mat** as A_p (n-by-2), A_d (n-by-1), B_p, B_d, C_p, C_d matrices.

The computed 3D positions need to be outputted to **point.mat**, containing $points\_A$ (n-by-3), $points\_B$ and $points\_C$ matrices.

2. As we move the camera backward along the z-axis, we want to magnify object C so it appears to be half of the height of object A, while keeping the object A constant in size.

   More specifically, we want $||A'_1A'_2|| = 2||C'_1C'_2||$ and $||A_1A_2|| = ||A'_1A'_2||$, where $C'_1C'_2$ and $A'_1A'_2$ are the object heights in current frame, and A_1A_2 is the height in the first frame.

   Compute the focal length f' and the position of camera center O'_c (on z axis) to achieve this goal.

3. To verify your 4.2 result, complete the Matlab function **project(p3d, f, pos)**, and use **Dolly_Zoom(f, pos)** to render the image with camera located at **pos**.

   More details can be found in **project.m** and **Dolly_Zoom.m** files.

   Include your synthetic image and image position of the vertices in the written part.

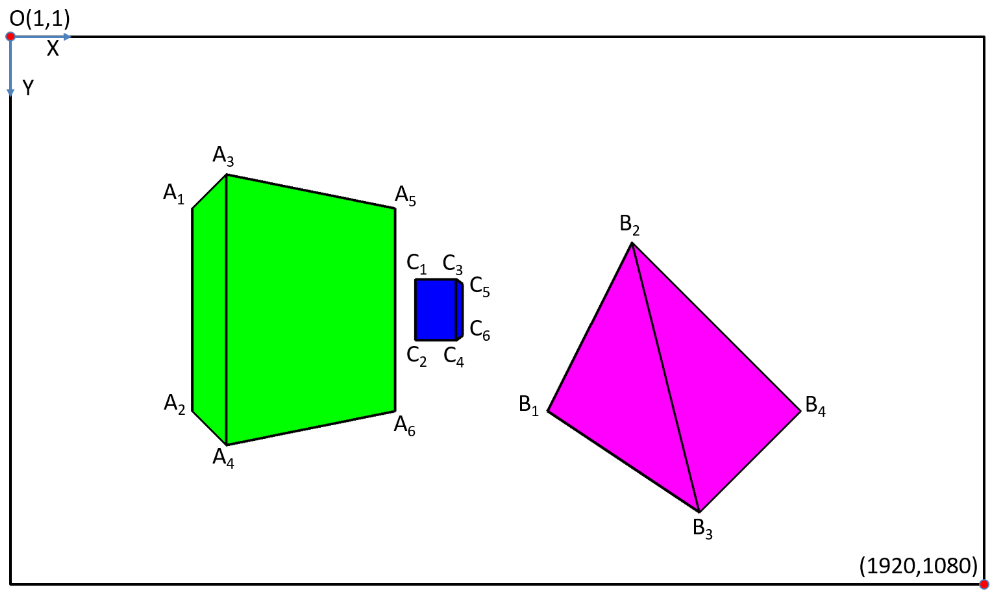

Figure 6: Camera image rendered for the synthetic scene: the first frame with pos=0, and f=400.

4. We keep moving the camera backward while maintaining the height of object A in image. Suppose the camera positions **pos** are given, compute the camera focal length we need to keep $||A_1A_2||$ same as the first frame.

   Complete the Matlab function of **compute_f(pos)** and use **generate_video.m** to render the video to show how images are changing during the Dolly Zoom. More details can be found in **compute_f.m** and **generate_video.m** files.

5. **Extra Credit** From the video, you can observe that at some period of time, object C was fully occluded by object A. Compute the range of camera center positions for this event.