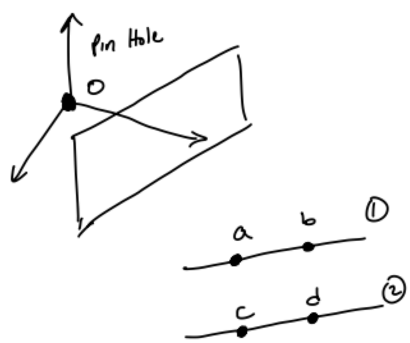
CSC I6716 Fall 2023 – Assignment 3
Allen Lau
EMPL (Last 4 Digits): 3838

1. (**Camera Models- 20 points**) Prove that the vector from the viewpoint of a pinhole camera to the vanishing point (in the image plane) of a set of 3D parallel lines is parallel to the direction of the parallel lines. Please show the steps of your proof.

   Given:

   The pin hole camera center is defined as point O and is defined as the origin. Two parallel lines, lines 1 and 2, that exist in the world. On line 1, there exists point a and b, and on line 2, there exists points c and d.



   Solution:

   The parametric representation of line 1 shown below, where $[X,Y,Z]^T$ is any point existing on the line, defined from $P_0 = P_1$:

$$\text{Parametric Equation of Line 1:} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} P_{1x} \\ P_{1y} \\ P_{1z} \end{bmatrix} + t \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

   As a result, points a and b on line 1 can be described by the following parametric equations:

$$\text{Parametric Equation of Point a on Line 1:} \begin{bmatrix} X_a \\ Y_a \\ Z_a \end{bmatrix} = \begin{bmatrix} P_{1x} \\ P_{1y} \\ P_{1z} \end{bmatrix} + t_a \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

$$\text{Parametric Equation of Point b on Line 1:} \begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} = \begin{bmatrix} P_{1x} \\ P_{1y} \\ P_{1z} \end{bmatrix} + t_a \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

Using the perspective projection equations, we can take points a and b defined in the world and project them onto the image plane defined at z = f away from the camera center, O:

$$\text{Projection of a to Image Plane: } x_a = f\,\frac{P_{1x}+t_aV_x}{P_{1z}+t_aV_z}\ , y_a = f\,\frac{P_{1y}+t_aV_y}{P_{1z}+t_aV_z}\ , z_a = f$$

$$\text{Projection of b to Image Plane: } x_b = f\,\frac{P_{1x}+t_bV_x}{P_{1z}+t_bV_z}\ , y_b = f\,\frac{P_{1y}+t_bV_y}{P_{1z}+t_bV_z}\ , z_b = f$$

Using the same procedure above, below are the parametric equations for line 2 and points c and d on line 2. Note that $P_0 = P_2$, however $V_1 = V_2 = V$ because the two lines are parallel.

$$\text{Parametric Equation of Line 2: } \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} P_{2x} \\ P_{2y} \\ P_{2z} \end{bmatrix} + t \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

$$\text{Parametric Equation of Point c on Line 2: } \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} P_{2x} \\ P_{2y} \\ P_{2z} \end{bmatrix} + t_c \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

$$\text{Parametric Equation of Point d on Line 2: } \begin{bmatrix} X_d \\ Y_d \\ Z_d \end{bmatrix} = \begin{bmatrix} P_{2x} \\ P_{2y} \\ P_{2z} \end{bmatrix} + t_d \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

$$\text{Projection of c to Image Plane: } x_c = f\,\frac{P_{2x}+t_cV_x}{P_{2z}+t_cV_z}\ , y_c = f\,\frac{P_{2y}+t_cV_y}{P_{2z}+t_cV_z}\ , z_c = f$$

$$\text{Projection of c to Image Plane: } x_d = f\,\frac{P_{2x}+t_dV_x}{P_{2z}+t_dV_z}\ , y_c = f\,\frac{P_{2y}+t_dV_y}{P_{2z}+t_dV_z}\ , z_d = f$$

We know that the perspective projection of parallel lines onto an image plane will converge to the vanishing point. This is simply the intersection of two lines, which can be derived in a general form as seen below:

$$line\ 1\colon a_1x + b_1y = c_1$$

$$line\ 2\colon a_2x + b_2y = c_2$$

$$Solving\ the\ system\ of\ equations\ for\ x\ and\ y\colon$$

$$x = \frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1}\ , y = \frac{c_1}{b_1} - \frac{a_1}{b_1}\left(\frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1}\right)$$

As a result, given two parallel lines each with two points defined, we can derive the vanishing point in terms of those 4 points using the following procedure:

1. After projecting points a and b from line 1 and c and d from line 2 onto the image plane, we can construct two equations of lines in general form, Ax + By = C, that exist on the image plane.
2. Solve for the intersection of the two resultant lines that exist on the image plane to find the vanishing point.

To solve for the equation of the projection of line 1 on the image plane, we compute the slope and write it in point-slope form. This is then converted to the general equation of the line:

Slope and point-slope form of line 1 projection:

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

$$= \frac{\cancel{f}\,\dfrac{P_{1y} + t_b V_y}{P_{1z} + t_b V_z} - \cancel{f}\,\dfrac{P_{1y} + t_a V_y}{P_{1z} + t_a V_z}}{\cancel{f}\,\dfrac{P_{1x} + t_b V_x}{P_{1z} + t_b V_z} - \cancel{f}\,\dfrac{P_{1x} + t_a V_x}{P_{1z} + t_a V_z}}$$

$$= \frac{\dfrac{(P_{1y} + t_b V_y)(P_{1z} + t_a V_z) - (P_{1y} + t_a V_y)(P_{1z} + t_b V_z)}{\cancel{(P_{1z} + t_b V_z)(P_{1z} + t_a V_z)}}}{\dfrac{(P_{1x} + t_b V_x)(P_{1z} + t_a V_z) - (P_{1x} + t_a V_x)(P_{1z} + t_b V_z)}{\cancel{(P_{1z} + t_b V_z)(P_{1z} + t_a V_z)}}}$$

$$= \frac{\cancel{P_{1y}P_{1z}} + t_a P_{1y} V_z + t_b P_{1z} V_y + \cancel{t_a t_b V_y V_z} - (\cancel{P_{1y}P_{1z}} + t_b P_{1y} V_z + t_a P_{1z} V_y + \cancel{t_a t_b V_y V_z})}{\cancel{P_{1x}P_{1z}} + t_a P_{1x} V_z + t_b P_{1z} V_x + \cancel{t_a t_b V_x V_z} - (\cancel{P_{1x}P_{1z}} + t_b P_{1x} V_z + t_a P_{1z} V_x + \cancel{t_a t_b V_x V_z})}$$

$$m = \frac{t_a P_{1y} V_z + t_b P_{1z} V_y - t_b P_{1y} V_z - t_a P_{1z} V_y}{t_a P_{1x} V_z + t_b P_{1z} V_x - t_b P_{1x} V_z - t_a P_{1z} V_x}$$

$$y - f\,\frac{P_{1y} + t_a V_y}{P_{1z} + t_a V_z} = \frac{t_a P_{1y} V_z + t_b P_{1z} V_y - t_b P_{1y} V_z - t_a P_{1z} V_y}{t_a P_{1x} V_z + t_b P_{1z} V_x - t_b P_{1x} V_z - t_a P_{1z} V_x}\left(x - f\,\frac{P_{1x} + t_a V_x}{P_{1z} + t_a V_z}\right)$$

General Equation of Line 1 Projection:

$$\frac{t_a P_{1y} V_z + t_b P_{1z} V_y - t_b P_{1y} V_z - t_a P_{1z} V_y}{t_a P_{1x} V_z + t_b P_{1z} V_x - t_b P_{1x} V_z - t_a P_{1z} V_x}\, x - y =$$

$$\frac{t_a P_{1y} V_z + t_b P_{1z} V_y - t_b P_{1y} V_z - t_a P_{1z} V_y}{t_a P_{1x} V_z + t_b P_{1z} V_x - t_b P_{1x} V_z - t_a P_{1z} V_x}\left(f\,\frac{P_{1x} + t_a V_x}{P_{1z} + t_a V_z} \right) - f\,\frac{P_{1y} + t_a V_y}{P_{1z} + t_a V_z}$$

Using the same procedure as above, we can write the general equation of line 2's projection:

$$\frac{t_c P_{2y} V_z + t_d P_{2z} V_y - t_d P_{2y} V_z - t_c P_{2z} V_y}{t_c P_{2x} V_z + t_d P_{2z} V_x - t_d P_{2x} V_z - t_c P_{2z} V_x}\, x - y =$$

$$\frac{t_c P_{2y} V_z + t_d P_{2z} V_y - t_d P_{2y} V_z - t_c P_{2z} V_y}{t_c P_{2x} V_z + t_d P_{2z} V_x - t_d P_{2x} V_z - t_c P_{2z} V_x}\left(f\,\frac{P_{2x} + t_c V_x}{P_{2z} + t_c V_z} \right) - f\,\frac{P_{2y} + t_c V_y}{P_{2z} + t_c V_z}$$

The parametric equation of the line from the pin hole camera (located at the origin) to the vanishing point can be written as the below:

$$\begin{bmatrix} x_0 \\ y_0 \\ f \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix}$$

Using the general intercept of two lines equation as described above, we can solve for the vanishing point (x_0, y_0) by plugging in the respective variables from the two general equations of line 1 and 2's projection on the image plane. Next, we can isolate the direction components of the resulting expression such that V_a is a function of only terms in the V_x direction, V_b is a function of only terms in the V_y direction, and V_z is a function of only terms in the V_z direction. This results in each direction vector only being multiplied by scalar values (the t and P values).

From this information, we know that the parametric equation of a line from the camera center to the vanishing point and the parallel lines existing in the world are scalar multiples of each other, with the same direction vector V; therefore, the vector from the pin hole to the vanishing point of a set of 3D parallel lines is parallel to the direction of those parallel lines.

2. (**Camera Models- 20 points**) Show that relation between any image point $(x_{im}, y_{im})^T$ of a plane (in the form of $(x_1,x_2,x_3)^T$ in projective space) and its corresponding point $(X_w, Y_w, Z_w)^T$ on the plane in 3D space can be represented by a 3×3 matrix. You should start from the general form of the camera model $(x_1,x_2,x_3)^T = M_{int}M_{ext}(X_w, Y_w, Z_w, 1)^T$, where $M = M_{int}M_{ext}$ is a 3×4 matrix, with the image center (o_x, o_y), the focal length f, the scaling factors(s_x and s_y), the rotation matrix R and the translation vector T all unknown. You should use the general form of the projective matrix (5 points), and the general form of a plane $n_x X_w + n_y Y_w + n_z Z_w = d$ (5 points), work on an integration (5 points), to form a 3×3 matrix between a 3D point on the plane and its 2D image projection (5 points).

Given:

General Form of the Projective Matrix:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = M_{int}M_{ext} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Where,

$$M_{int} = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}, \ f_x = \frac{f}{s_x} \ and \ f_y = \frac{f}{s_y}$$

$$M_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

Solution:

The product of $M_{int}M_{ext}$ is computed below:

$$M_{int}M_{ext} = \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

Substituting $M_{int}M_{ext}$ into the general projective matrix results in:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Distributing the world spatial components to their respective columns results in:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -f_x r_{11} X_w + o_x r_{31} X_w & -f_x r_{12} Y_w + o_x r_{32} Y_w & -f_x r_{13} Z_w + o_x r_{33} Z_w & -f_x T_x + o_x T_z \\ -f_y r_{21} X_w + o_y r_{31} X_w & -f_y r_{22} Y_w + o_y r_{32} Y_w & -f_y r_{23} Z_w + o_y r_{33} Z_w & -f_y T_y + o_y T_z \\ r_{31} X_w & r_{32} Y_w & r_{33} Z_w & T_z \end{bmatrix}$$

Given the equation of a plane $n_x X_w + n_y Y_w + n_z Z_w = d$, Z_w can be written as a function of X_w and Y_w:

$$Z_w = \frac{d - n_y Y_w - n_x X_w}{n_z}$$

Z_w is then substituted into the projective matrix equation:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$

$$\begin{bmatrix} -f_x r_{11} X_w + o_x r_{31} X_w & -f_x r_{12} Y_w + o_x r_{32} Y_w & -f_x r_{13} \frac{d - n_y Y_w - n_x X_w}{n_z} + o_x r_{33} \frac{d - n_y Y_w - n_x X_w}{n_z} & -f_x T_x + o_x T_z \\ -f_y r_{21} X_w + o_y r_{31} X_w & -f_y r_{22} Y_w + o_y r_{32} Y_w & -f_y r_{23} \frac{d - n_y Y_w - n_x X_w}{n_z} + o_y r_{33} \frac{d - n_y Y_w - n_x X_w}{n_z} & -f_y T_y + o_y T_z \\ r_{31} X_w & r_{32} Y_w & r_{33} \frac{d - n_y Y_w - n_x X_w}{n_z} & T_z \end{bmatrix}$$

Distributing the column of the M matrix corresponding to Z_w to the other three columns, simplifying the terms, and then pulling the world spatial components out of the matrix results in the finalized equation.

For the sake of brevity in the report, only computations for the first row of the 3x3 matrix is shown below:

Row 1, Z component:

$$-f_x r_{13} \frac{d - n_y Y_w - n_x X_w}{n_z} + o_x r_{33} \frac{d - n_y Y_w - n_x X_w}{n_z}$$

$$\frac{-f_x r_{13} d}{n_z} + \frac{f_x r_{13} n_x}{n_z} X_w + \frac{f_x r_{13} n_y}{n_z} Y_w + \frac{o_x r_{33} d}{n_z} - \frac{o_x r_{33} n_x}{n_z} X_w - \frac{o_x r_{33} n_y}{n_z} Y_w$$

Distributing each component to the other three columns:

Row 1, Column 1:

$$-f_x r_{11} X_w + \frac{f_x r_{13} n_x}{n_z} X_w + o_x r_{31} X_w - \frac{o_x r_{33} n_x}{n_z} X_w$$

$$-f_x \left(r_{11} - \frac{r_{13} n_x}{n_z} \right) X_w + o_x \left(r_{31} - \frac{r_{33} n_x}{n_z} \right) X_w$$

Row 1, Column 2:

$$-f_x r_{12} Y_w + \frac{f_x r_{13} n_y}{n_z} Y_w + o_x r_{32} Y_w - \frac{o_x r_{33} n_y}{n_z} Y_w$$

$$-f_x \left(r_{12} - \frac{r_{13} n_y}{n_z} \right) Y_w + o_x \left(r_{32} - \frac{r_{33} n_y}{n_z} \right) Y_w$$

Row 1, Column 3:

$$-f_x T_x - \frac{f_x r_{13} d}{n_z} + o_x T_z + \frac{o_x r_{33} d}{n_z}$$

$$-f_x \left(T_x + \frac{r_{13} d}{n_z} \right) + o_x \left(T_z + \frac{r_{33} d}{n_z} \right)$$

The relation between any image point $(x_{im}, y_{im})^T$ of a plane (in the form of $(x_1,x_2,x_3)^T$ in projective space) and its corresponding point $(X_w, Y_w, Z_w)^T$ on the plane in 3D space, represented by a 3×3 matrix, is shown below:

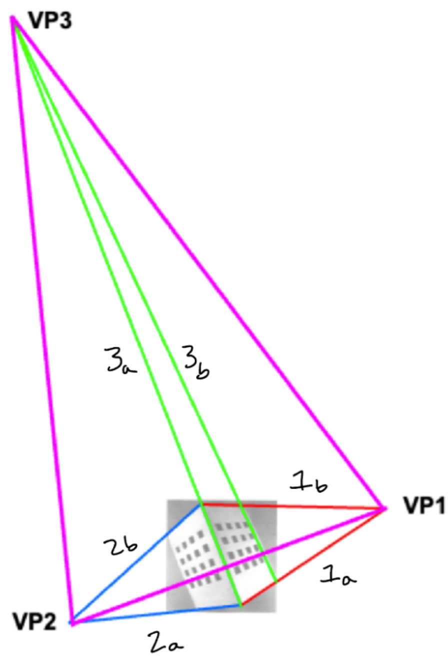
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$

$$\begin{bmatrix} -f_x \left(r_{11} - \frac{r_{13} n_x}{n_z} \right) + o_x \left(r_{31} - \frac{r_{33} n_x}{n_z} \right) & -f_x \left(r_{12} - \frac{r_{13} n_y}{n_z} \right) + o_x \left(r_{32} - \frac{r_{33} n_y}{n_z} \right) & -f_x \left(T_x + \frac{r_{13} d}{n_z} \right) + o_x \left(T_z + \frac{r_{33} d}{n_z} \right) \\ -f_y \left(r_{21} - \frac{r_{23} n_x}{n_z} \right) + o_y \left(r_{31} - \frac{r_{33} n_x}{n_z} \right) & -f_y \left(r_{22} - \frac{r_{23} n_y}{n_z} \right) + o_y \left(r_{32} - \frac{r_{33} n_y}{n_z} \right) & -f_y \left(T_y + \frac{r_{23} d}{n_z} \right) + o_y \left(T_z + \frac{r_{33} d}{n_z} \right) \\ r_{31} - \frac{r_{33} n_x}{n_z} & r_{32} - \frac{r_{33} n_y}{n_z} & T_z + \frac{r_{33} d}{n_z} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

3. (**Calibration- 20 points**) Prove the Orthocenter Theorem by geometric arguments: Let T be the triangle on the image plane defined by the three vanishing points of three mutually orthogonal sets of parallel lines in space. Then the image center is the orthocenter of the triangle T (i.e., the common intersection of the three altitudes.

(1) Basic proof: use the result of Question 1, assuming the aspect ratio of the camera is 1. Note that you are asked to prove the Orthocenter Theorem, not just the orthocenter of a triangle (7 points)

<u>Given:</u>



Camera center, O, is defined as the origin.

Lines 1a and 1b, exist in the world and are parallel. The perspective projection of these lines on the image plane converges to the vanishing point VP1.

Lines 2a and 2b, exist in the world and are parallel. The perspective projection of these lines on the image plane converges to the vanishing point VP2.

Lines 3a and 3b exist in the world and are parallel. The perspective projection of these lines on the image plane converges to the vanishing point VP3.

Lines 1a, 1b, 2a, 2b, 3a, and 3b form a set of mutually orthogonal sets of parallel lines. Their vanishing points, VP1, VP2, and VP3 form a triangle T.

Solution:

From question 1, it is proven that the vector from the pin hole camera to the vanishing point of a set of parallel lines is parallel to those parallel lines. As a result:

- Vector OVP1 is parallel to lines 1a and 1b
- Vector OVP2 is parallel to lines 2a and 2b
- Vector OVP3 is parallel to lines 3a and 3b

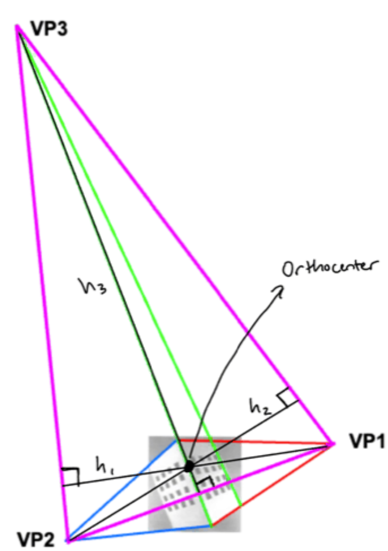
Since these vectors from the camera to the vanishing points are parallel to mutually orthogonal sets of parallel lines, this means that:

- OVP1 is orthogonal to OVP2 and OVP3
- OVP2 is orthogonal to OVP1 and OVP3
- OVP3 is orthogonal to OVP1 and OVP2

From the above information, we also know that the vectors from the camera center and the sides of triangle T are orthogonal:

- OVP1 is orthogonal to VP2VP3 (VP2VP3 = OVP2 – OVP3)
- OVP2 is orthogonal to VP1VP3 (VP1VP3 = OVP1 – OVP3)
- OVP3 is orthogonal to VP1VP2 (VP1VP3 = OVP1 – OVP3)

Defining the altitude for the triangle from each vanishing point to its corresponding triangle edge (where the vector from the vanishing point to the edge intersects the edge at 90 degrees) can be represented as variables h1, h2, and h3. The intersection of these three lines represents the orthocenter.



The image center, which is defined by the vector from the camera center to the image plane is, by definition, perpendicular to the image plane. This image center is located at the orthocenter because we know that the planes created from the camera center, image center, and each vanishing point would intersect at the orthocenter, where they would be orthogonal to each other.

(**2)** If you do not know the focal length of the camera, can you still find the image center using the Orthocenter Theorem? Explain why or why not (3 points). Can you also estimate the focal length after you find the image center? If yes, how, and if not, why (5 points)

Yes, we can still find the image center using the Orthocenter Theorem because this theorem is dependent on the geometric relationship between the vanishing points of a set of mutually orthogonal parallel lines and the triangles resulted from projecting those lines onto the image plane. This relationship is not dependent on focal length since we can still find this relationship and the orthocenter without the focal length being known.

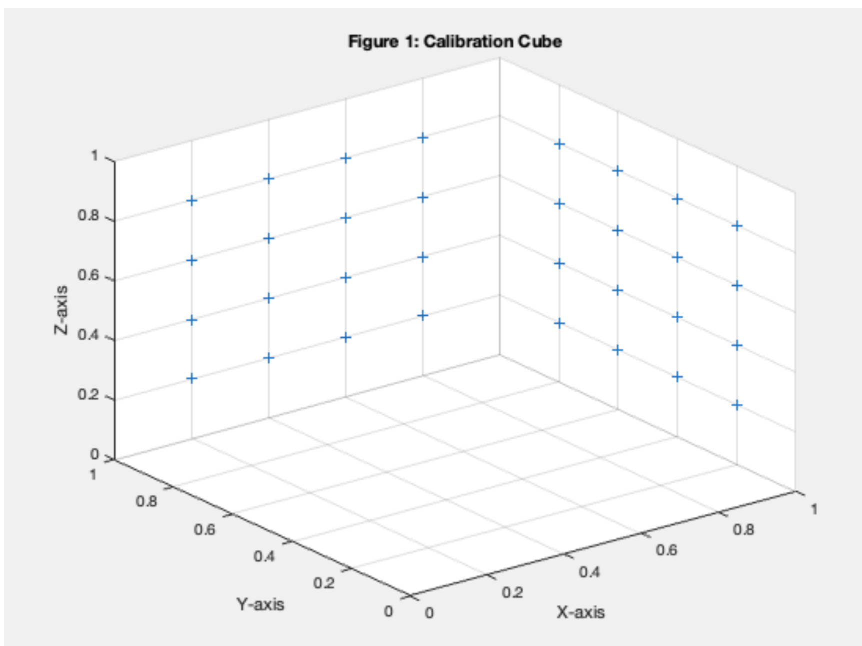
After estimating finding the image center, we should be able to estimate the focal length if given a calibration image like the example above. This allows us to have known world coordinates that can be paired with image coordinates to do a camera calibration process.

(3) If you do not know the aspect ratio and the focal length of the camera, can you still find the image center using the Orthocenter Theorem? Explain why or why not. (5 points)

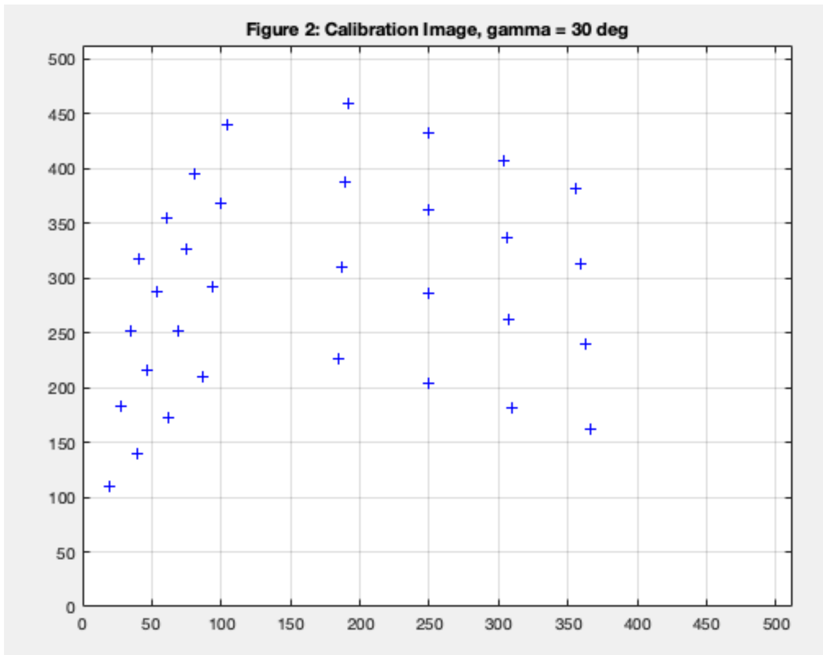
Without knowing the aspect ratio and focal length, we wouldn't be able to find the image center using the Orthocenter Theorem. This is because this theorem is dependent on the relationship between the relative geometric lengths and orientations of the lines and planes created from the vanishing points and orthocenter. If we do not know the aspect ratio, then these scales and orientations would not be fixed and information like the altitudes may not be found correctly.

4. **Calibration Programming Exercises (40 points)**: Implement the direct parameter calibration method to (1) learn how to use SVD to solve systems of linear equations; (2) understand the physical constraints of the camera parameters; and (3) understand important issues related to calibration, such as calibration pattern design, point localization accuracy and robustness of the algorithms. Since calibrating a real camera involves lots of work in calibration pattern design, image processing and error controls as well as solving the equations, we will use simulated data to understand the algorithms. As a by-product we will also learn how to generate 2D images from 3D models using a "virtual" pinhole camera.

- **A. Calibration pattern "design".** Generate data of a "virtual" 3D cube similar to the one shown in the lecture notes in camera calibration. For example, you can hypothesize a 1x1x1 m3 cube and pick up coordinates of 3-D points on one corner of each black square in your world coordinate system. Make sure that the number of your 3-D points is sufficient for the following calibration procedures. To show the correctness of your data, draw your cube (with the control points marked) using Matlab (or whatever language you are using). I have provided a piece of starting code in Matlab for you to use. **(5 points)**



- **B. "Virtual" camera and images.** Design a "virtual" camera with known intrinsic parameters including focal length f, image center (ox, oy) and pixel size (s_x, s_y). As an example, you can assume that the focal length is f = 16 mm, the image frame size is 512*512 (pixels) with an image center (o_x, o_y) = (256, 256), and the size of the image sensor inside your camera is 8.8 mm *6.6 mm (so the pixel size is (s_x, s_y) = (8.8/512, 6.6/512)). Capture an image of your "virtual" calibration cube with your virtual camera with a given pose (rotation R and translation T). For example, you can take the picture of the cube 4 meters away and with a tilt angle of 30 degree. Use three rotation angles alpha, beta, gamma to generate the rotation matrix R (refer to the lecture notes in camera model – please double check the equation since it might have typos in signs). You may need to try different poses to have a suitable image of your calibration target. **(5 points)**



- **C. Direction calibration method:** Estimate the intrinsic (f_x, f_y, aspect ratio a, image center (o_x, o_y)) and extrinsic (R, T and further alpha, beta, gamma) parameters. Use SVD to solve the homogeneous linear system and the least square problem, and to enforce the orthogonality constraint on the estimate of R.

  C(i). Use the accurately simulated data (both 3D world coordinates and 2D image coordinates) to the algorithms and compare the results with the "ground truth" data (which are given in step (a) and step (b)). Remember you are practicing a camera calibration, so you should pretend you know nothing about the camera parameters (i.e., you cannot use the ground truth data in your calibration process). However, in the direct calibration method, you could use the knowledge of the image center (in the homogeneous system to find extrinsic parameters) and the aspect ratio (in the Orthocenter theorem method to find image center). **(15 points)**

The below table describes the true camera parameters resulted from steps a and b:

True Camera Parameters	
Focal length (f)	0.016 m
Image Center (Ox, Oy)	(256 pixels, 256 pixels)
Pixel Size (Sx, Sy)	(0.0172 mm/pixel, 0.0129 mm/pixel)
Effective Focal Length (fx, fy)	(930.9091 pixels, 1241.2 pixels)
Aspect Ratio (asr)	0.750
Image Frame	512x512 pixels

The procedure for doing the direct parameter calibration is as follows:

1. Given a set of world calibration and their corresponding image points, build matrix A of the homogenous Av = 0 system.
2. Compute the SVD of A and taking the 8^{th} row of the V^T matrix, which represents the vector that contains the eigenvalue = 0. This vector of eight parameters represents v_bar with a given scale factor gamma.
3. Compute the magnitude of the scale factor gamma and aspect ratio.
4. With gamma and aspect ratio found, we can now compute the first two rows of the rotation matrix, T_x, and T_y, up to a common sign.
5. Find the sign, s, by assuming s is positive and computing $r_{11}X_w + r_{12}Y_w + r_{12}Z_w + T_x$. The image point x and the world point X should be opposite signs.
6. Compute the third row of the rotation matrix by taking the cross product of R1 and R2. Enforce orthogonality for the R matrix.
7. Compute T_z and F_x by constructing the matrix A and b given the equation $A*[T_z, F_x]^T = b$ using least squares method and SVD to find the inverse of A.
8. Compute F_y using F_x and the aspect ratio.

The below is a table of the estimated parameters, which are totally correct (no need to check them ☺):

Estimated Parameters; Image Center = (256, 256)	
Aspect Ratio (asr)	1.1709
Effective Focal Length (fx, fy)	(-0.3976 pixels, -0.3396 pixels)
Rotation Matrix (R)	$\begin{bmatrix} -0.4965 & 0.6273 & 0.5999 \\ 0.3147 & 0.7743 & -0.5490 \\ -0.8089 & -0.0838 & -0.5819 \end{bmatrix}$
Translation Matrix (T)	$\begin{bmatrix} 0 \\ -0.0243 \\ -0.0023 \end{bmatrix}$

C(ii). Study whether the unknown aspect ratio matters in estimating the image center **(5 points)**, and how the initial estimation of image center affects the estimating of the remaining parameters **(5 points)**, by experimental results. Give a solution to solve the problems if any **(5 points)**.

An unknown aspect ratio does not affect the estimation of the camera parameters for the direct parameter calibration method if we know the dimensions of the image frame. This method only requires a known image center and point pairs (calibration world coordinates paired with its x', y' coordinates). If the aspect ratio and the dimensions of the image frame are unknown, then computing the image center will not be possible since the Orthocenter Theorem requires the aspect ratio information.

The initial estimation of the image center should affect the accuracy of estimating the unknown parameters because it directly affects the setup of matrix A, which is one of the initial steps in calibration. In terms of its affect, the closer you are to the true image center, the more accurate the parameter estimation should be.

The below are experimental results exploring how the initial estimation of the image center affects the estimating of the parameters (once again, some amazing results! ☺):

Estimated Parameters; Image Center = (0, 0)	
Aspect Ratio (asr)	0.6180
Effective Focal Length (fx, fy)	(-0.2392 pixels, -0.3870 pixels)
Rotation Matrix (R)	$\begin{bmatrix} 0.5925 & 0.2099 & 0.7777 \\ -0.5343 & 0.8249 & 0.1844 \\ -0.6028 & -0.5248 & 0.6010 \end{bmatrix}$
Translation Matrix (T)	$\begin{bmatrix} 1.2217 \\ 0.7539 \\ 0 \end{bmatrix}$

Estimated Parameters; Image Center = (128,128)	
Aspect Ratio (asr)	1.0574
Effective Focal Length (fx, fy)	(0.4894 pixels, 0.4629 pixels)
Rotation Matrix (R)	$\begin{bmatrix} 0.5209 & 0.6123 & 0.5948 \\ 0.5027 & -0.7832 & 0.3660 \\ 0.6899 & 0.1083 & -0.7157 \end{bmatrix}$
Translation Matrix (T)	$\begin{bmatrix} -0.4362 \\ -0.2730 \\ -0.0009 \end{bmatrix}$

Estimated Parameters; Image Center = (256, 256)	
Aspect Ratio (asr)	1.1709
Effective Focal Length (fx, fy)	(-0.3976 pixels, -0.3396 pixels)
Rotation Matrix (R)	$\begin{bmatrix} -0.4965 & 0.6273 & 0.5999 \\ 0.3147 & 0.7743 & -0.5490 \\ -0.8089 & -0.0838 & -0.5819 \end{bmatrix}$
Translation Matrix (T)	$\begin{bmatrix} 0 \\ -0.0243 \\ -0.0023 \end{bmatrix}$