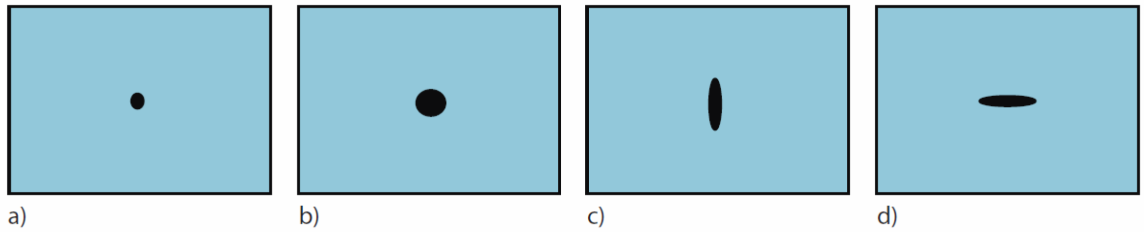
**Submission details**

- This assignment should be done in pairs (contact the TA if this is a problem).
- The topic of this assignment is The Pinhole Camera, Camera Calibration, Single View Metrology and Epipolar Geometry.
- The submission date is **4/25/2016**. Please pay attention to the late submission policy.
- Coding should be done in Matlab or Python. We prefer Python 2, but Python 3 is possible.
- You are not required to use any specific function or library, unless stated otherwise. If in doubt, please contact the TA via email or the Piazza website.
- For submission, package up your code as a zip file. Include your written answers as a pdf file named writeup.pdf. Include graphs and images in your writeup, and please label them so we know which figures go with which sub-problem.
- Send the final zip file to the TA. Add the course number to the subject of the email.
- If you have any questions about this assignment, please contact the TA stinger@tx.technion.ac.il.

**Task 1: The Pinhole Camera**

1. Consider a pinhole camera in which we change the geometry of the aperture as shown in the following image:




   Consider a basic scene situated in front of the pinhole camera. Illustrate the images that would be taken by the pinhole camera for each of the four aperture shapes. Explain what leads you to your answers. Note that this question is qualitative. It is not important to take into account the exact size of the aperture.

2. Explain the problems affecting an image taken by a pinhole camera when the diameter of the aperture is too small, and when the diameter of the aperture is too large.

**Task 2: Camera Calibration**

Here the goal is to compute the 3x4 camera matrix $P$ describing a pinhole camera, given the coordinates of 10 world points and their corresponding image projections. Then you will decompose $P$ into the intrinsic and extrinsic parameters. You should write a simple Python script that works through the stages below, printing out the important terms.

Use the provided ASCII files, world.txt and image.txt. The first file contains the $(X,Y,Z)$ values of 10 world points, and the second file contains the $(x,y)$ projections of those 10 points.

Add to your writeup significant values resulting from the stages below.

**1.** Find the 3x4 matrix $P$ that projects the world points $\underline{X}$ to the 10 image points $\underline{x}$. This should be done in the following steps:

- Since $P$ is a homogeneous matrix, the world and image points (which are 3-D and 2-D respectively), need to be converted into homogeneous points by concatenating a 1 to each of them (thus becoming 4-D and 3-D respectively).
- We now note that $\underline{x} \times P\underline{X} = 0$ , irrespective of the scale ambiguity. This allows us to setup a series of linear equations of the form:

$$\begin{bmatrix} \underline{\mathbf{0}}^T & -w_i\underline{\mathrm{X}}_i^T & y_i\underline{\mathrm{X}}_i^T \\ w_i\underline{\mathrm{X}}_i^T & \underline{\mathbf{0}}^T & -x_i\underline{\mathrm{X}}_i^T \\ -y_i\underline{\mathrm{X}}_i^T & x_i\underline{\mathrm{X}}_i^T & \underline{\mathbf{0}}^T \end{bmatrix} \begin{pmatrix} P_1^T \\ P_2^T \\ P_3^T \end{pmatrix} = \underline{\mathbf{0}}$$

  for each correspondence $\underline{\mathrm{x}}_i \leftrightarrow \underline{\mathrm{X}}_i$, where $\underline{\mathrm{x}}_i = (x_i, y_i, w_i)^T$ , $w_i$ being the homogeneous coordinate, and $P_j$ is the j[th] row of $P$. But since the 3[rd] row is a linear combination of the first two, we need only consider the first two rows for each correspondence. Thus, you should form a 20 by 12 matrix $A$, each of the 10 correspondences contributing two rows. This yields $A\underline{p} = \underline{\mathbf{0}}$ , $\underline{p}$ being the vector containing the entries of matrix $P$.
- To solve for $\underline{p}$ , we need to impose an extra constraint to avoid the trivial solution $\underline{p} = \underline{\mathbf{0}}$ . One simple one is to use $\|\underline{p}\|_2 = 1$ . This constraint is implicitly imposed when we compute the SVD of $A$. The value of $\underline{p}$ that minimizes $A\underline{p}$ subject to $\|\underline{p}\|_2 = 1$ is given by the eigenvector corresponding to the smallest singular value of $A$. To find this, compute the SVD of $A$, picking this eigenvector and reshaping it into a 3x4 matrix $P$.
- Verify your answer by re-projecting the points $\underline{\mathrm{X}}$ and checking that they are close to $\underline{\mathrm{x}}$ .

Now that we have $P$, we can compute the world coordinates of the projection center of the camera $\underline{C}$ . Note that $P\underline{C} = \underline{\mathbf{0}}$ , thus $\underline{C}$ lies in the null space of $P$, which can again be found using SVD. Compute the SVD of $P$ and pick the vector corresponding to this null-space. Finally, convert it back to homogeneous coordinates to yield the $(X,Y,Z)$ coordinates.

**Task 3: Single View Metrology**

1. A chess board of $1 \times 1$ meters is aligned such that its lines are parallel to the axes X and Z, and one of its corners is at the origin of the axes. Each of the corners between the board's checkers is located in the position $X_{ij}$ that fits a 2D integer index $(i, j)$, $0 \le i, j \le 7$ . The origin of the axes fits the index $(0,0)$: $X_{00} = (0,0,0)^T$ . An ideal camera with a focal length f is located at $X_c = (0,1,0)^T$ (one meter above the origin), and is directed at an angle of $\theta$ relative to the Z axis such that the camera's main axis is parallel to the board, and one of the axes of the image plane is parallel to the Y axis. Assume that the size of the focal plane and the image in it is unlimited.
   a. Calculate the locations of the corners of the board's checkers in the 3D world as a function of $(i, j)$ .

b. Calculate the locations of the corners of the board's checkers in the image plane as a function of $(i, j)$ and $\theta$.

c. Calculate the location of the vanishing line belonging to the plane of the chess board.

2. Properties of the orthographic projection.
   a. Suppose we have a line given by a point **p** and a direction vector **v** so that the equation for the line is $l(t) = \mathbf{p} + t\mathbf{v}$. A parallel line would be any line that has the same direction **v**. Show that under orthographic projection, parallel lines project to parallel lines.
   b. Given a line from **a** to **b**, and any point **c** along this line, show that the quantity

$$\frac{|\mathbf{c}-\mathbf{a}|}{|\mathbf{b}-\mathbf{c}|}$$

   is invariant under orthographic projection (Hint: **c** can be expressed as $\mathbf{a} + t(\mathbf{b}-\mathbf{a})$, where $0 \le t \le 1$).
   c. Why do the above properties not hold for perspective projection? Explain briefly.

3. Let the vectors **a**, **b**, **c**, **d** correspond to the corners of a square mapped under perspective projection, where the homogeneous coordinates of the vectors are

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 4 \\ 8 \\ 2 \end{bmatrix}, \ \mathbf{c} = \begin{bmatrix} 15 \\ 12 \\ 3 \end{bmatrix}, \ \mathbf{d} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$$

   Calculate the vanishing point for each pair of parallel lines. Express your answer in homogeneous coordinates (Hint: Try plotting the points on a plane).

4. <u>Bonus</u>: Prove that under perspective projection, all the vanishing points that are received from lines parallel to a certain plane (which are not perpendicular to the image plane) lie on straight line.

**<u>Task 4: Epipolar Geometry</u>**

1. Basic properties of epipolar geometry.
   a. What is the relation between the epipole and the epipolar line?
   b. How many epipoles exist in a set of two cameras? Explain.
   c. Given the fundamental matrix $F$ in a set of two cameras, express the epipoles in the set as a function of $F$ (Hint: think of a property that an epipole on the right camera has to follow with a certain point in the image of the left camera).
   d. In a set of two cameras, one of the cameras is located at the origin of the axes, its main axis is the $Z$ axis and the axes of its image plane are parallel to the axes $X$ and $Y$. The second camera is located by setting it where the first camera is and then translating it to a distance of $d$ in the direction of the $Y$ axis. Write the locations of the epipoles in this set. Explain.
   e. Consider the set in the previous item, but instead of translating the second camera it is only rotated with a rotation matrix $R$. Assume that both cameras have the unity matrix as their intrinsic matrix. Explain what happens to the epipoles, lines and matrices in this set.

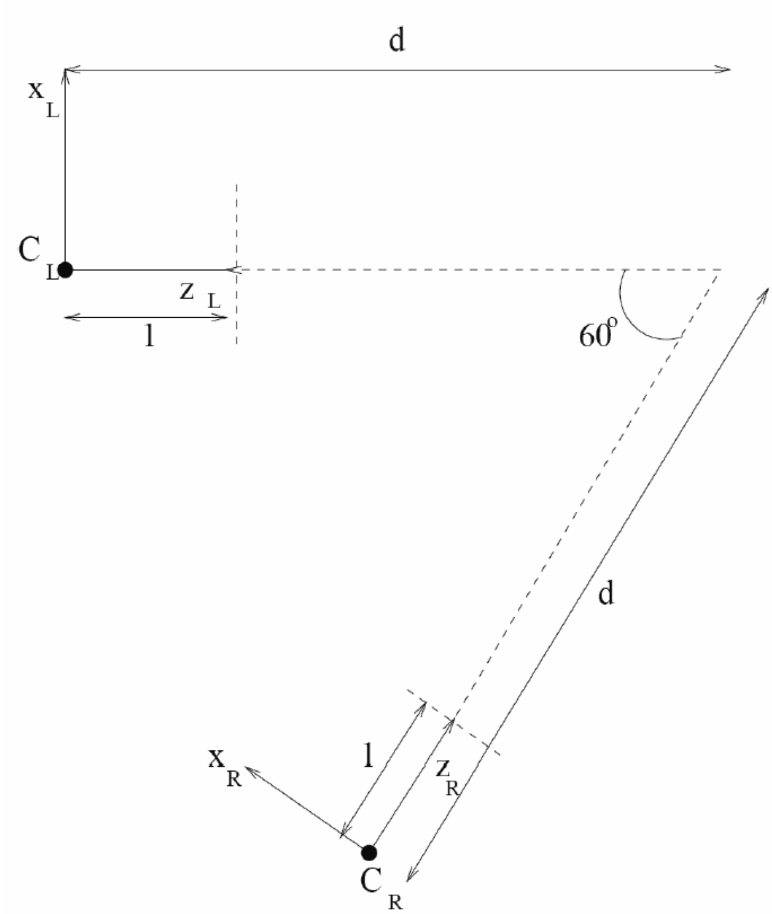
2. The figure below shows a pair of cameras, each with a focal length of a unity, whose principal axes meet at a point. The $Y$ axes of both cameras are parallel and point out of the page. Assume that the left camera (center $C_L$) lies at the origin of the world coordinate system.
   a. Write down the camera matrices for this configuration and verify that the fundamental matrix $F$ is:

$$\begin{pmatrix} 0 & -d/2 & 0 \\ -d/2 & 0 & -\sqrt{3}d/2 \\ 0 & \sqrt{3}d/2 & 0 \end{pmatrix}$$

   Hint: Take care when constructing $M'$, the projection matrix of the right camera. Do it by deriving $X'$ in terms of $X$, via a sequence of stages:

      i. a translation of the coordinate frame from $C$ to the intersection point;
      ii. a rotation of the coordinate frame about the intersection point and
      iii. another translation to $C'$.

   b. Compute the epipolar line in the right image corresponding to the homogeneous point $x = (1,1,1)^T$ in the left from $l = Fx$.
   c. Using Figure 2.1 determine where potential correspondences to the left image point $(x, y) = (0,0)$ can lie in the right image.
   d. Describe the rotation and translation that should be applied to the left camera that would make the epipolar lines in the two images horizontal.




3. For the given pair of chapel images:

   a. Find matching interest points using SIFT or Harris, similarly to what you did on HW3.
   b. Use the set of matched points you found to estimate the fundamental matrix $F$ automatically using RANSAC and the normalized 8-point algorithm.
      i. Indicate what test you used for deciding inlier vs. outlier.
      ii. Display F after normalizing to unit length.
      iii. Plot the outliers with green dots on top of the first image.
   c. Choose 7 sets of matching points that are well separated (can be randomly chosen). Plot the corresponding epipolar lines (in red) and the points (in green) on each image. Show the two images (with plotted points and lines) next to each other.

   Notes: Your epipolar lines should pass very near (e.g., within 1 pixel) your plotted points, but the solution might vary slightly from run to run. Consider your outlier criterion carefully.