ENPM673: Perception for Autonomous Robots
Homework 1

Project Report Submitted by:

Akanksha Patel - 116951029
Achal Vyas - 116869560
Sri Manika Makam - 116697415

1. Assume that you have a camera with a resolution of 5MP where the camera sensor is square shaped with a width of 14mm. It is also given that the focal length of the camera is 15mm.

i) Compute the Field of View of the camera in the horizontal and vertical direction:

Ans:

We know that the FoV (Field of View) depends on the focal length of the lens. The size of Field of View is governed by the size of the camera retina.

$$\o = tan^{-1}\ (\frac{d}{2f})$$

In the given equation, d is the width of the camera sensor and f is the focal length of the camera lens. Thus, in order to find FoV in the form of angle, we first found out the value of $\frac{d}{2f}$ and then took the tan inverse of the value in order to get the Field of View of the given camera.

$$\therefore \o = tan^{-1}\ (\frac{14}{30})$$

$$\therefore \o = tan^{-1}\ (0.4667)$$

$$\therefore \o = 25^{\circ}$$

Thus, the Field of View of the camera in the horizontal and vertical direction is 50°.

ii) Assuming you are detecting a square shaped object with width 5cm, placed at a distance of 20 meters from the camera, compute the minimum number of pixels that the object will occupy in the image:

To find the minimum number of pixels occupied by the mentioned object in the image frame, we first found the area of FoV of the camera, by using the similar triangle property.

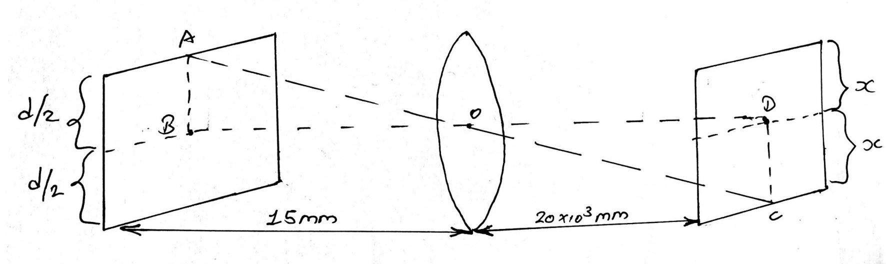

Figure 1: Image Formation

Using the similar triangle property in the triangles $\triangle ABO$ and$\triangle CDO$, we get that:

$$\frac{AB}{CD} = \frac{BO}{DO}$$

$$\therefore \frac{\frac{d}{2}}{x} = \frac{15}{20000}$$

$$\therefore \frac{\frac{14}{2}}{x} = \frac{15}{20000}$$

$$\therefore x = 9333\ mm$$

Let, number of pixel per mm of the FoV = n

Area of the object = A

Area of FoV = a

Minimum number of pixel occupied by the object = N

Now, as the area of the FoV = $(2x)^2$

$$\therefore a = 348,419,556\ mm^2$$

Now, in order to find the number of pixel per mm of the FOV, we divide area of FoV by resolution of the camera.

$$\therefore n = \frac{348419556\ mm}{5 * 10^6}$$

$$n = 0.014\ \frac{pixels}{mm}$$

Thus, the minimum number of the pixels occupied by the given object in the image is calculated by finding product of area of image and number of pixels per mm.

$$A = 50^2 = 2500\ mm^2$$

$$\therefore N = 2500 * 0.014$$

$$N = 35.8\ pixels$$

Thus the given square shaped object with width of 5cm will occupy 35.8 pixels.

2. Two files of 2D data points are provided in the form of CSV files. The data represents measurements of a projectile with different noise levels and is shown in figure 1. Assuming that the projectile follows the equation of a parabola

i) Find the best method to fit a curve to the given data for each case. You have to plot the data and your best fit curve for each case. Submit your code along with the instructions to run it.

Ans:

To find the best fit for the data, we tried two different algorithms, Ordinary Least Square (LS) and Random Sample Consensus (RANSAC). **For the first data points, Ordinary Least Square (LS) resulted in a better fit while for the second data points RANSAC performed better.** The final result of LS and RANSAC is shown in Figure 2 and Figure 3 respectively.

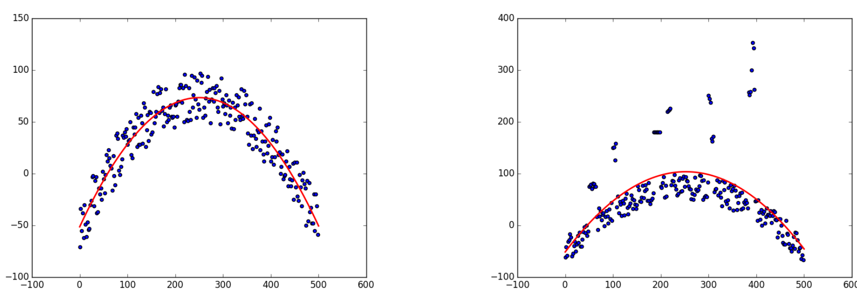

Figure 2: Curve fitting using Least Square (LS)

ii) Briefly explain all the steps of your solution and discuss why your choice of outlier rejection technique is best for that case.

Ans:

The two algorithms are discussed below:

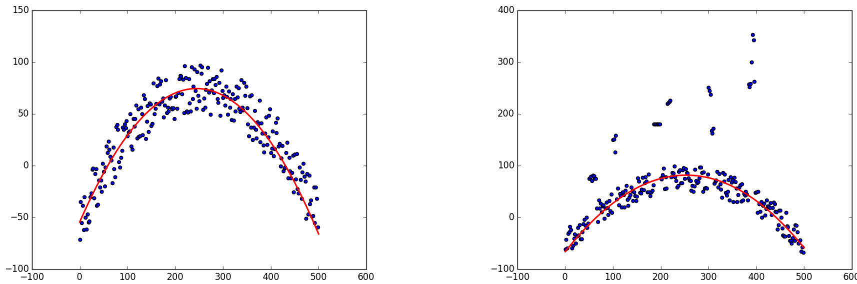

Figure 3: Curve fitting using RANSAC

i) Ordinary least squares: Ordinary Least Squares is a graph-fitting algorithm that finds a best-fitting curve for a given set of data points by minimizing the sum of the squares of the residual terms for each point. We know that the projectile follows the equation of parabola. Let the equation be given by $y = ax^2 + bx + c$. For a given set of data points $(x_1, y_1), \ldots (x_n, y_n)$, the problem can be reformulated in terms of a system of linear equations as follows:

$$Y = XB$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad X = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} \qquad B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

We want to minimize the error given by: $\sum_{i=1}^{n} (y_i - ax_i^2 - bx_i - c)$

$$\begin{aligned} E &= \|Y - XB\|^2 = (Y - XB)^T (Y - XB) \\ &= Y^T Y - 2(XB)^T Y + (XB)^T (XB) \end{aligned}$$

In order to find the minima, we differentiate E and equate it to 0.

$$\begin{gathered} \frac{dE}{dt} = 2X^T XB - 2X^T Y = 0 \\ X^T XB = X^T Y \\ B = (X^T X)^{-1} (X^T Y) \end{gathered}$$

Therefore, the equation of parabola corresponding to the best-fitting model is given by $B = (X^T X)^{-1} (X^T Y)$.

ii) RANSAC: Random Sample Consensus(RANSAC) is a curve fitting algorithm that is generally used when the data contains outliers. RANSAC estimates a model that fits the data by neglecting the effect of outliers in the estimation. Our implementation of RANSAC iteratively follows the following steps:

1. Randomly select three points from the data.
2. Generate a model (constants a, b, c in the equation of parabola) that fits the selected points.
3. For each remaining point, calculate the error. If error is greater than a specified threshold (th, specified by the user) mark it as an outlier, otherwise add the point to the set of inliers.
4. If the number of inliers $< d$ (specified by the user), discard the model. Go to step 1.
5. Update the model to best fit all the points in the inliers, $updated\_model$. (We have used our implementation of least squares for computing the $updated\_model$.)
6. Compute the error for all the points in inliers, $updated\_error$.
7. If $updated\_error <$ minimum error encountered till now, store $updated\_model$ as the best model

The steps above are repeated k times (specified by the user) and finally the model with the least error is chosen.

The value of threshold th was selected by eye inspection of the data. The value of minimum number of inliers for a model to be selected as a good fit for the data, d is 175 (70% of the data points as the number of outliers is low). And the number of iterations $k = 100$

Observations:

	LS	RANSAC
Data Set 1	12.3256	13.4443
Data Set 2	35.2158	30.9959

Table 1: Estimated error for different data sets

We observed that LS performed better on data points 1 and RANSAC performed better on data points 2. Table 1 summarizes the resultant errors estimates for LS and RANSAC. This was expected as LS results in a curve that minimizes the sum of the squares of the offsets. Therefore, in absence of outliers LS results in the best-fitting curve. However, because the error function in LS uses square of offsets, outliers have a significant affect on the resultant curve. This is what is noticed in Figure 2 where the curve is bulged upwards because of the outliers present in the data points. This scenario is handled by RANSAC as it neglects the outliers while computing the best-fitting model.

3. The concept of homography in Computer Vision is used to understand, explain and study visual perspective, and, specifically, the difference in appearance of two plane objects viewed from different points of view. This concept will be taught in more detail in the coming lectures. For now, you just need to know that given 4 corresponding points on the two different planes, the homography between them is computed using the following system of equations Ax = 0, where:

$$A = \begin{bmatrix} -x1 & -y1 & -1 & 0 & 0 & 0 & x1*xp1 & y1*xp1 & xp1 \\ 0 & 0 & 0 & -x1 & -y1 & -1 & x1*yp1 & y1*yp1 & yp1 \\ -x2 & -y2 & -1 & 0 & 0 & 0 & x2*xp2 & y2*xp2 & xp2 \\ 0 & 0 & 0 & -x2 & -y2 & -1 & x2*yp2 & y2*yp2 & yp2 \\ -x3 & -y3 & -1 & 0 & 0 & 0 & x3*xp3 & y3*xp3 & xp3 \\ 0 & 0 & 0 & -x3 & -y3 & -1 & x3*yp3 & y3*yp3 & yp3 \\ -x4 & -y4 & -1 & 0 & 0 & 0 & x4*xp4 & y4*xp4 & xp4 \\ 0 & 0 & 0 & -x4 & -y4 & -1 & x4*yp4 & y4*yp4 & yp4 \end{bmatrix}$$

$$X = \begin{bmatrix} H_{11} \\ H_{12} \\ H_{13} \\ H_{21} \\ H_{22} \\ H_{23} \\ H_{31} \\ H_{32} \\ H_{33} \end{bmatrix}$$

For the given point correspondences,

-	x	y	xp	yp
1	5	5	100	100
2	150	5	200	80
3	150	150	220	80
4	5	150	100	200

find the homography matrix

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix}$$

Ans:

We are given an homogeneous linear equation system given by the expression $Ax = 0$, where X is the vector of $N = 9$ unknowns and A is the matrix of M x N (8 x 9) coefficients.

The equations in $Ax = 0$ are linear in all unknowns and the right hand side is 0. Hence, the linear equation system allows the trivial solution x = 0. But, this is generally not the desired solution. To avoid this, we constrain the vector x to a fixed length, say $\|x\|^2 = 1$. The vector norm being $\|x\|^2 = \sum {x_i}^2$, we can say that this constraint is not linear, but quadratic in the elements of x.

Mathematically, the best way to solve the equation system $Ax = 0$, subject to the constraint mentioned above, is to perform Singular Value Decomposition (SVD) on the matrix A. SVD factors the matrix into a diagonal matrix D and two orthogonal matrices U and V, such that

$$A = UDV^T$$

U is a M x M (8 x 8) matrix with columns as orthogonal eigen vectors of AA^T.

$$U = \begin{bmatrix} 0.0118 & 0.0003 & -0.0516 & -0.4661 & -0.2603 & -0.0678 & 0.0108 & -0.8411 \\ 0.0118 & 0.0003 & -0.0872 & -0.4594 & -0.2491 & -0.0886 & 0.7655 & 0.3542 \\ 0.3587 & 0.6549 & 0.0135 & -0.4651 & 0.1701 & 0.2936 & -0.2784 & 0.1823 \\ 0.1435 & 0.2620 & -0.4454 & 0.1361 & -0.5008 & -0.5875 & -0.2731 & 0.1529 \\ 0.7750 & 0.0227 & 0.4085 & 0.2849 & 0.0320 & -0.2352 & 0.2627 & -0.1597 \\ 0.2818 & 0.0082 & -0.6922 & 0.3159 & 0.0114 & 0.5019 & 0.2466 & -0.1696 \\ 0.1846 & -0.3168 & 0.2485 & -0.0347 & -0.6983 & 0.4673 & -0.2524 & 0.1816 \\ 0.3693 & -0.6336 & -0.2889 & -0.3933 & 0.3189 & -0.1750 & -0.2614 & 0.1526 \end{bmatrix}$$

V is a N x N (9 x 9) matrix with columns as orthogonal eigen vectors of A^TA.

$$V = \begin{bmatrix} 0.0028 & 0.0031 & -0.2464 & -0.1586 & -0.1752 & 0.1767 & 0.9137 & -0.1203 & 0.0531 \\ 0.0024 & -0.0013 & -0.3770 & 0.1766 & 0.6895 & 0.5903 & -0.0529 & -0.0022 & -0.0049 \\ 0.0000 & 0.0000 & -0.0024 & -0.0037 & 0.0052 & 0.0075 & 0.0660 & 0.7860 & 0.6146 \\ 0.0011 & 0.0012 & 0.6612 & 0.3412 & 0.5017 & -0.2325 & 0.3721 & -0.0426 & 0.0177 \\ 0.0016 & -0.0029 & 0.5743 & -0.0710 & -0.3145 & 0.7499 & -0.0620 & 0.0046 & -0.0039 \\ 0.0000 & -0.0000 & 0.0058 & -0.0022 & 0.0029 & -0.0057 & -0.1225 & -0.6049 & 0.7868 \\ -0.6961 & -0.7180 & -0.0001 & -0.0038 & 0.0025 & -0.0002 & 0.0044 & -0.0006 & 0.0002 \\ -0.7180 & 0.6961 & 0.0016 & -0.0038 & 0.0025 & 0.0037 & -0.0006 & 0.0000 & -0.0000 \\ -0.0062 & 0.0000 & -0.1735 & 0.9067 & -0.3783 & 0.0622 & 0.0252 & -0.0025 & 0.0076 \end{bmatrix}$$

Eigen values ($\lambda_1 ... \lambda_r$) of AA^T and A^TA are equal. Here, we get 8 eigen values, that is $r = 8$. Then the singular values of A, σ_i is given by

$$\sigma_i = \sqrt{\lambda_i}$$

D is a M x N (8 x 9) matrix. First, a 8 x 8 diagonal matrix is formed with diagonal elements as $\sigma_1, \ldots, \sigma_r$. We then append a column of zeros so as to obtain a 8 x 9 matrix.

$$D = \begin{bmatrix} 60214.8954 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 31824.5207 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 260.8931 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 186.2193 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 145.6064 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 60.8809 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 3.8987 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.8102 & 0.0000 \end{bmatrix}$$

In the given problem, N is greater than M ($N > M$), that is the number of unknowns are greater than number of independent equations available to solve them. In this case, the equation system is under-constrained. So, we cannot find the exact solution but can find the solution which minimizes the sum of squared errors over all variables.

The least-squares solution is given by the column of V, which corresponds to the smallest eigen value of $A^T A$.

The eigen values of $A^T A$ are {3625833626.6986, 1012800114.9102, 68065.1927, 34677.6193, 21201.2335, 3706.4890, 15.2001, 0.6565, 0.0000}

We can see the smallest eigen value is 0. So, the solution x is given by the last column of V, which corresponds to the least eigen value.

$$x = \begin{bmatrix} 0.0531 \\ -0.0049 \\ 0.6146 \\ 0.0177 \\ -0.0039 \\ 0.7868 \\ 0.0002 \\ -0.0000 \\ 0.0076 \end{bmatrix}$$

Therefore, H is given by

$$H = \begin{bmatrix} 0.0531 & -0.0049 & 0.6146 \\ 0.0177 & -0.0039 & 0.7868 \\ 0.0002 & -0.0000 & 0.0076 \end{bmatrix}$$