Machine Vision - Assignment 2

Supankan

March 1, 2026

Question 1

We are given a dataset in `lines.csv` which represents three point scatters conforming to three lines. The first five rows reveal that the coordinates of the points for the three lines are organized into columns $(x_1, x_2, x_3, y_1, y_2, y_3)$.

(a) Total Least Squares on the First Line

To fit a line strictly using the data corresponding to the first line (the x_1 and y_1 columns), we perform Total Least Squares (TLS). This involves centralizing the coordinates by subtracting their mean and then performing Principal Component Analysis (PCA) or Singular Value Decomposition (SVD) on the centered data matrix. The singular vector associated with the smallest singular value provides the optimal normal vector (a, b) of the fitted line $ax + by + d = 0$.

By computing this on the first line's data, we obtain the following parameters for the line equation:

$$a = 0.7736$$
$$b = -0.6337$$
$$d = -3.7942$$

This defines the line:
$$0.7736x - 0.6337y - 3.7942 = 0$$

Or in the standard $y = mx + c$ format:

$$y = 1.2207x - 5.9872$$

(b) Fitting Three Lines using RANSAC

Now, treating all points collectively as a single dataset (by flattening all X columns and Y columns), we run the RANSAC (Random Sample Consensus) algorithm to iteratively find the three lines. The process is as follows:

1. Randomly sample two distinct points to form a line hypothesis.

2. Calculate the perpendicular distance of all available points to this line.

3. Count the number of inliers (points within a specified distance threshold).

4. Retain the line with the largest consensus set (most inliers).

5. Mask (remove) the winning inliers from the dataset.

6. Repeat the process on the remaining points to discover the next line.

After running this iterative process (with an inlier threshold of 0.2, and roughly 2000 iterations per line), and refining each consensus set with TLS, the resulting parameters for the three discovered lines are:

Line 1:

- Parameters: $a = -0.7284$, $b = 0.6851$, $d = -0.6595$

- Equation: $-0.7284x + 0.6851y - 0.6595 = 0 \implies y = 1.0631x + 0.9625$

- Inlier count: 43 points

Line 2:

- Parameters: $a = -0.3801$, $b = -0.9250$, $d = 2.0074$

- Equation: $-0.3801x - 0.9250y + 2.0074 = 0 \implies y = -0.4109x + 2.1703$

- Inlier count: 39 points

Line 3:

- Parameters: $a = 0.7894$, $b = -0.6139$, $d = -3.5935$

- Equation: $0.7894x - 0.6139y - 3.5935 = 0 \implies y = 1.2859x - 5.8537$

- Inlier count: 36 points

Note: As RANSAC is a randomized algorithm, the exact parameters and ordering of the extracted lines may vary slightly across different runs depending on the initialization and chosen threshold.

Question 2

We are tasked with finding the physical sizes of a pair of earrings shown in an image, given the following camera parameters:

- Focal length (f): 8 mm

- Distance to object (Z): 720 mm

- Pixel size: $2.2 \, \mu\text{m} = 0.0022$ mm

Assuming a standard pinhole camera model where the optical axis is perpendicular to the imaging plane (so the object is fronto-parallel), the magnification M is given by:

$$M = \frac{f}{Z}$$

The physical size of the object X relates to its size on the sensor x as:

$$X = x \cdot \frac{Z}{f}$$

Since the sensor size x can be expressed in terms of the number of pixels spanning the object and the pixel size:

$$X = (\text{pixels} \times \text{pixel\_size}) \cdot \frac{Z}{f}$$

By writing a simple Python script (using OpenCV) to analyze the image `earrings.jpg`, we threshold the image to find the bounding boxes of the two earrings. The script detects that both earrings span a bounding box of 382×400 pixels.

Applying the formula:

$$\text{Width} = (382 \times 0.0022 \text{ mm}) \times \frac{720 \text{ mm}}{8 \text{ mm}} = 0.8404 \text{ mm} \times 90 = 75.64 \text{ mm}$$

$$\text{Height} = (400 \times 0.0022 \text{ mm}) \times \frac{720 \text{ mm}}{8 \text{ mm}} = 0.8800 \text{ mm} \times 90 = 79.20 \text{ mm}$$

Thus, the physical dimensions of each earring are approximately **75.64 mm $\times$ 79.20 mm**.