Understanding Feature Matching, Essential Matrix, and Epipolar Geometry

February 11, 2025

1 Introduction

- Finding feature correspondences using SIFT and ORB
- Computing the fundamental and essential matrices
- Decomposing the essential matrix to find camera motion
- Visualizing epipolar lines

2 Feature Matching: Finding Correspondences

To estimate the relationship between two images, we first need to find corresponding points between them.

2.1 SIFT (Scale-Invariant Feature Transform)

SIFT is a feature detection algorithm that finds keypoints in an image based on differences in scale and orientation.

1. Convert the image to grayscale.
2. Generate a scale space by progressively blurring the image.
3. Use the Difference of Gaussians (DoG) to detect keypoints.
4. Assign an orientation to each keypoint.
5. Compute a descriptor by considering gradient orientations around the keypoint.
6. Match keypoints between images using the Euclidean distance of their descriptors.

The descriptor for a keypoint is represented as a 128-dimensional vector.

2.2 ORB (Oriented FAST and Rotated BRIEF)

ORB is a faster alternative to SIFT and is based on binary descriptors.

1. Detect keypoints using the FAST corner detector.
2. Assign an orientation using the intensity centroid method.
3. Compute binary descriptors using the BRIEF (Binary Robust Independent Elementary Features) method.
4. Match features using the Hamming distance instead of Euclidean distance.

3 Fundamental Matrix: Finding the Relationship Between Two Views

The fundamental matrix F represents the relationship between two images by encoding how points in one image correspond to points in the other.

3.1 Mathematical Definition

The fundamental matrix satisfies the equation:

$$\mathbf{x'}^T F \mathbf{x} = 0 \tag{1}$$

where:

- $\mathbf{x}$ is a point in the first image.
- $\mathbf{x}'$ is the corresponding point in the second image.
- F is a 3×3 matrix that encodes the relationship.

3.2 Computing the Fundamental Matrix

To compute F, we use the Eight-Point Algorithm:

1. Normalize the point coordinates.
2. Construct a matrix A using the correspondences:

$$A = \begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ x'_2 x_2 & x'_2 y_2 & x'_2 & y'_2 x_2 & y'_2 y_2 & y'_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{2}$$

3. Compute the singular value decomposition (SVD) of A, keeping only the smallest singular value.
4. Ensure F has rank 2 by forcing the smallest singular value to be zero.

4 Essential Matrix: Incorporating Camera Calibration

The essential matrix E is similar to the fundamental matrix but incorporates the camera's intrinsic parameters.

$$E = K^T F K \tag{3}$$

where K is the camera intrinsic matrix containing focal length and optical center.

4.1 Computing the Essential Matrix

To compute E:

1. Compute the fundamental matrix F.
2. Multiply by the intrinsic camera matrix K:

$$E = K^T F K \tag{4}$$

3. Ensure E has rank 2 by forcing one singular value to be zero.

5 Decomposing the Essential Matrix: Extracting Rotation and Translation

The essential matrix can be decomposed to extract the rotation and translation between two camera views.

5.1 Using Singular Value Decomposition (SVD)

We perform SVD on E:

$$E = U \Sigma V^T \tag{5}$$

where U and V are orthogonal matrices and Σ contains singular values.

5.2 Extracting Rotation and Translation

From SVD, the two possible rotation matrices are:

$$R_1 = UWV^T, \quad R_2 = UW^T V^T \tag{6}$$

where:

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

The translation vector is extracted as:

$$t = U[:, 3] \tag{8}$$

6 Epipolar Geometry: Visualizing the Correspondence

6.1 Epipolar Lines

Once we have F, we can draw epipolar lines, which are the possible locations where a point in one image appears in the other.

$$\text{epipolar line} = F\mathbf{x} \tag{9}$$

where $\mathbf{x}$ is a point in one image.

7 Conclusion

In this document, we covered:

- How to find corresponding points using SIFT and ORB
- How to compute the fundamental and essential matrices
- How to extract rotation and translation from the essential matrix
- How to visualize epipolar geometry

These concepts are fundamental for 3D reconstruction, visual odometry, and structure-from-motion tasks.