Camera Pose Estimation

—

Shreehan Santosh Kate

National Institute of Technology,
Tiruchirapalli

Perspective-n-Point Algorithm

Perspective-n-Point or PnP algorithm deals with computing the calibration of the camera relative to the objects in the given image given a set of 3D coordinates in the world frame and corresponding 2D points on the image.

OpenCV provides a function cv2.solvePnP() that takes the following inputs:

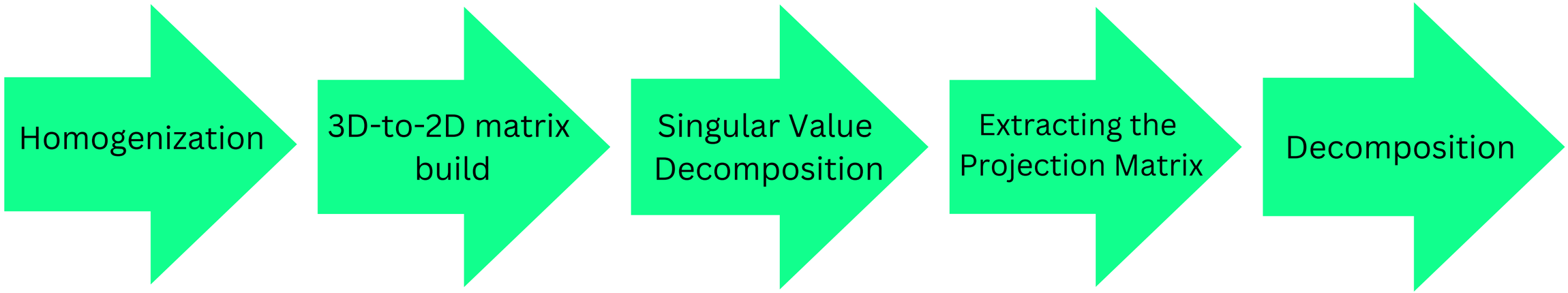
- Image Points: Set of 3D points in the World frame
- Object Points: Set of 2D pixel points in Image frame
- Camera Matrix: The Camera Model is computed beforehand with the help of already-known parameters such as camera focal lengths and principle points

Perspective-n-Point Algorithm

The Direct-Linear-Transformation (DLT) technique is incorporated in PnP algorithms to estimate the camera pose with the help of 3D and 2D coordinates.

The DLT method is based on principles of Projective Geometry and Linear Algebra

The following logic flow encapsulates the entire working of the DLT method:



Visual Inertial Odometry

Visual Inertial Odometry is used to estimate the pose and orientation of an unmanned vehicle by extracting datafrom visual and inertial sensors. Examples of visual and inertial sensors are cameras and accelerometers respectively

Visual Inertial Odometry incorporates visual and inertial measurements. Sensors are calibrated with the help of offset and relative poses. The camera’s intrinsic and extrinsic parameters are judged as well

Visual Inertial Odometry incorporates visual and inertial measurements. Sensors are calibrated with the help of offset and relative poses.
The camera’s intrinsic and extrinsic parameters are judged as well

VIO is widely used in robotics, Unmanned vehicles
and augmented reality applications

Simultaneous Locating and Mapping

Simultaneous locating and mapping (SLAM) algorithms are used to establish an unmanned vehicle’s position in a globally referenced environment with the help of visual sensors such as cameras.

Traditional SLAM methods rely on two features - Scale-Invariant-Feature Transform (SIFT) and Object Fast and Rotated BRIEF (ORB) for feature detection.

Optical flow algorithms estimate the apparent motion of objects in the scene by analyzing the displacement of image pixels between consecutive frames.

SLAM often incorporates a database that is used to access previously visited places by identifying certain elements in current/past frames.

Application of Deep Learning in SLAM involves feature extraction, depth estimation, dense mapping, semantic understanding among others.