EKF based Real Time Visual SLAM

Liang Xu
Electrical and Computer Enineering
virginia commonwealth university

***Abstract*—The past decades have been significant progress in robot navigation and Simultaneous Localization and Mapping(SLAM). SLAM is the computational problem of constructing a map of an unknown environment on the fly while simultaneously keeping track of a moving object with a sensor on it. It has been widely used in the robotic community. Different sensors can contribute their own knowledge with estimating the environment.**
Accelerator, IMU, GPS and many another sensors could be used for SLAM. However, many surprising reasons vision is an attractive choice for SLAM sensor: cameras are compact, accurate, informative, cheap and ubiquitous. Vision also is the primary navigation tool for human and animals. So this project will use the Extended Kalmen Filter(EKF) knowledge learned from the class and Computer Vision to build a real-time visual SLAM system.

I. Introduction

The real-time V-SLAM have two parts, one is the computer vision and another is the parameter estimation. The computer vision can give a better knowledge of the image get by the camera, and parameter estimation can help us to estimate the current position.
So, how to estimate the current point by just using the image? First, we are assuming the camera is moving, and it is a dynamic movement. This means the movement is a continuous function, and follow the Newton's Law. Second, there are Natural Visual Landmarks on each image. This marks could be some special points or corners in the image. There is the certain case, we can not get any feature points from the image, but we are not going to discuss that case. According to the first assumption, the object movement is a continuous function, so there are strong relation between the images. So we can use the position different between to image to estimate the position of the camera. This the whole idea of this project.

II. Feature Point Extraction

There are whole chapter and papers on the feature point extraction. So I am not going to discuss too much on this topic. Figure 1 is one of the examples of feature points, the red circle in the image are the feature points. Scale-invariant feature transform(SIFT) is the most common used feature points extraction technique, and the feature points are also called key points. SIFT is a robust algorithm to scale, transform and rotation. So the key-point find in the current frame has a high probability in the next frame also.
Feature matching also another topic. However, in this report, we only use this technique. Each of the key points will have its own descriptor, and the matting algorithm will search the descriptor for the target image. The expected movement of those key point is not too far from each other. So according to the sampling time, we can estimate the movement of the camera.

Fig. 1: Feature Point of an Image

III. Camera Position Transformation

Before estimate the camera position, we need to know how the 3D sense project to 2D image. The work flow is this: 3D sense project to 2D image using camera, estimate the position and movement using the key-point in 2D image, estimate the 3D position of the camera in the world coordinate, and then build a 3D map.
For the camera, there are intrinsic and extrinsic parameters Those parameters are fixed by the manufacturer. Figure 2 is the example of transform the world coordinate to the camera coordinate and also the image coordinate.

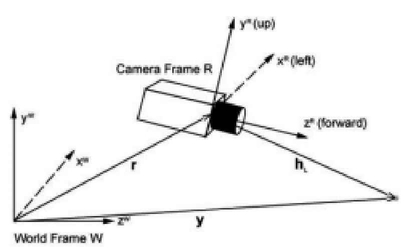

Fig. 2: Camera Parameter

IV. Camera Position Estimation

With the knowledge of the coordinate transform and feature points information, we can plug in the Extented Kalman Filter to estimate the position of camera in world coordinate. Figure 3 is a simulated image for camera position estimation

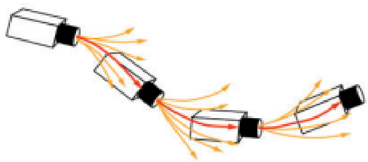

Fig. 3: Camera Parameter