Simultaneous Localization and Mapping, Outdoor Tracking

Simultaneous Localization and Mapping:

The simplest form of model-free tracking, which can be seen as a precursor to simultaneous localization and mapping (SLAM), is sometimes called visual odometry. In a nutshell, visual odometry means continuous 6DOF tracking of a camera pose relative to an arbitrary starting point. This approach originally comes from the field of mobile robotics. Visual odometry computes a 3D reconstruction of the environment, but uses it just to support the incremental tracking. A basic visual odometry pipeline encompasses the following steps:

1. Detect interest points in the first frame.

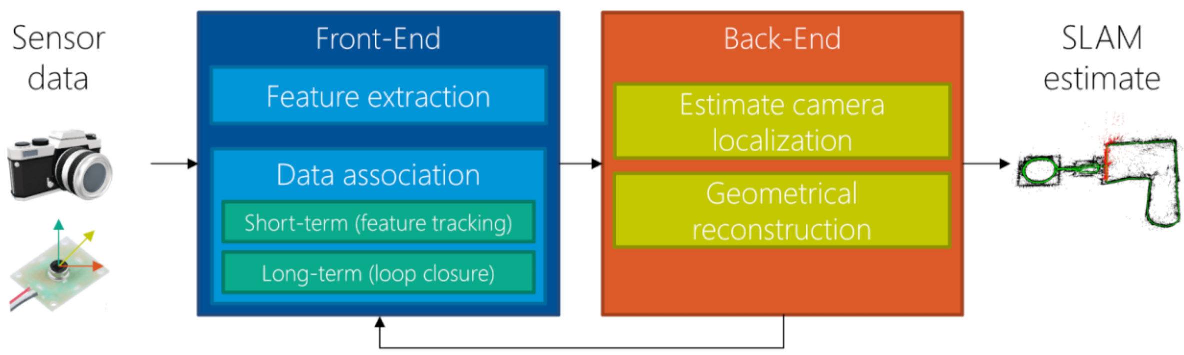
2. Track the interest point in 2D from the previous frame

3. Determine the essential matrix between the current and previous frames from the feature correspondences with a five-point algorithm

4. Recover the incremental camera pose from the essential matrix.

5. Since the essential matrix determines the translation part of the pose only up to scale, this scale must be estimated separately, so that it is consistent throughout the tracked image sequence. To achieve this aim, 3D point locations are triangulated from multiple 3D observations of the same image feature over time (see "Triangulation from More Than Two Cameras"). This approach is called structure from motion (SFM).

6. Proceed to the next frame.



Outdoor Tracking:

- Indoor environments are usually more predictable whereas the outdoor environments are limitless in terms of location and orientation.
- As stated earlier, GPS is a good tracking option when working outdoors.
- A differential GPS and a compass was used for position and orientation judgement. Latitudes and longitudes of several viewpoints were collected in a database along with the set of images captured at different times of the year with varying light conditions.
- Reference images were utilized for video tracking and matching was performed to discover these reference images for the outdoor AR system.
- A video image was examined with the reference images and a matching score was achieved. For the finest matching score, the 2D transformation was measured and the current camera position and orientation were deducted.
- This transformation was utilized to register the model on the video frame. The matching technique was based on Fourier Transformation to be robust against variation in lighting conditions hence it was limited to only 2D transformations like rotation and translation.
- This technique had a fixed number of computations therefore it was appropriate for real-time operation without using markers, it yet worked on 10Hz which is a low rate for real-time display.