

Camera grid calibration

Tim Lenertz

July 5, 2017

1 Introduction

This report describes the method used to calibrate the extrinsic camera parameters of the 3DLicorneA data sets. It is applicable for dense 2D data sets, where cameras are placed on a more or less regular grid on a plane parallel to the scene. The optical flow of tracked features, and aggregated values from the depth maps, are used to compute the camera positions and orientations. No calibration pattern needs to be present in the scene.

1.1 Requirements

To use the method, there are the following requirements:

- The camera intrinsic parameters are known. This is the camera matrix K and optionally the distortion coefficients. The method was used for the case where there is no distortion (Kinect v2), but with additional steps it can be used with distorted images.
- There is a depth map for all of some views. There is an extension to use the method without depth maps, by instead fixing the depth only for some feature points.
- The camera centers are arranged on an approximately regular 2D grid on a plane P . The distance between adjacent camera positions (in X and Y direction) is sufficiently small that feature points can be tracked using optical flow. For this disparities of features on adjacent views should be smaller than 15 pixels.
- It is assumed that camera centers lie always exactly on the plane. That is, they never move towards or back from the scene.
- The camera is facing approximately perpendicular to the plane P , towards the scene. There can be a small rotation \mathbf{R} of the camera relative to P . It is estimated from the images, as part of the calibration. The yaw, pitch and roll angles should be smaller than 5° . It is assumed that this angle remains constant for all views.
- There can be some missing images and depth maps in the data set.
- It is not necessary that the tracked feature points remain visible across the whole range of views. Calibration can be done on subsets of the camera positions, and then stitched together.

The method calculates one global rotation matrix \mathbf{R} , and for each view v , a 2D vector (x_v, y_v) of the camera position on P . From this it computes an extrinsic matrix \mathbf{Rt} for each view.

2 Method

Calibration is done in four steps: (1) Compute *image correspondences* using feature tracking. (2) Estimate the *camera rotation* \mathbf{R} . (3) Estimate *straight depths* of the tracked features, i.e. their distance to P . (4) Deduce camera positions.

First the algorithm selects several *feature points* on one on multiple *reference views*.

2.1 Image correspondences

2.1.1 Reference views

2.1.2 Choosing feature points

2.1.3 Optical flow tracking

2.1.4 Filtering features

2.2 Camera rotation

2.2.1 Optical flow slopes

2.2.2 View feature depths

2.3 Straight depths

2.3.1 Aggregating view feature depths

2.3.2 Depth from disparity

2.4 Camera positions

2.4.1 Relative camera positions

2.4.2 Stitching

2.4.3 Final camera extrinsics