Stereo RGB-D indoor mapping with precise stream fusing strategy

Shengjun Tang [a,] *, Weixi Wang [a], Xiaoming Li [a] , Zhilu Yuan [a]

a Research Institute for Smart Cities & Shenzhen Key Laboratory of Spatial Information Smart Sensing and Services, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, PR China; shengjuntang@szu.edu.cn (S.T.), wangwx@szu.edu.cn (W.W.),lixming@szu.edu.cn

* Corresponding author

Keywords: Stereo camera, Camera tracking, Point Cloud, Indoor Mapping

Abstract:

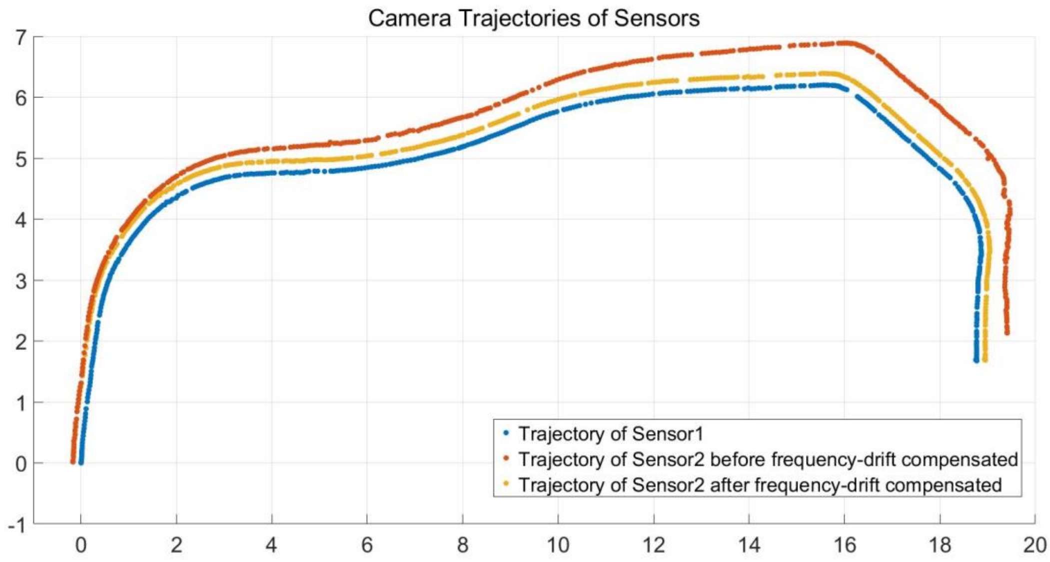
In order to achieve more robust pose tracking and mapping of visual SLAM, the robotics researcher has recently shown a growing interest in utilising multiple camera, which is able to provide more sufficient observations to fulfil the frame registration and map updating tasks. This implies that better pose tracking robustness can be achieved by extending monocular visual SLAM to utilise measurements from multiple cameras.[1] proposed a visual SLAM method using multiple RGB-D cameras, which integrate the observations from multi-camera for camera tracking. However, they ignored the time-drift between the frames obtained by different cameras, which may result at inaccurate positions of observation used for map updating. Besides, loop closure detection was not been implemented. [2] constructed a multiple RGB-D system with three Kinects V2 camera. This work mainly concentrated on the intrinsic and extrinsic calibration and verify the effectiveness of mapping using multiply RGB-D cameras.

Figure 1. Left: Stereo RGB-D mapping system setup, the two Kinects mounted on a Jetson Tx2, one facing upwards and the other one facing downwards. Middle: Camera views of two camera at the same time-stamp. Right: The built dense point cloud using the two Kinects in a SLAM process. The trajectory and key frame poses are also plotted in this map.

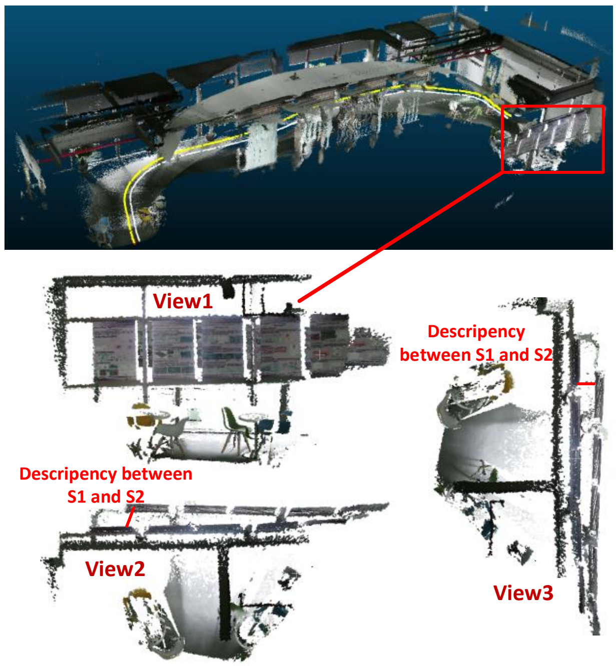
While in theory, a fix rigid transformation should be sufficient to register the frames with the same timestamp from two sensors. Since synchronising multiple Kinect sensors is impossible, there existed frequency-drift due to different topic publish rate of sensors. In this work, we present a frequency-drift compensated closed-form solution for multiply RGB-D SLAM. The intrinsic parameters for each sensor are obtained with a standard camera calibration process and the extrinsic orientation parameters achieved through a coarse-to-fine scheme that solves the initial exterior orientation parameters (EoPs) from sparse control markers and further refines the initial value by an iterative closest point (ICP) variant minimizing the distance between the RGB-D point clouds and the referenced laser point clouds. Then with the assumption of fix transformation between the frames with the same timestamp, we define one sensor as reference sensor and the other sensor as slave sensor and the slave frames can be mapped to the timeline of the references sensor. Rather than endow the camera pose of the nearest frame to the slave frames, we derive the accurate camera pose for slave frames in a spatially variant way. For each slave frame, we make a hypothesis that there exist a corresponding reference frame with the same timestamp and two adjacent frames can be found for each fictitious frame. A linear basis is imposed on the translation and rotation to recover the camera pose of the fictitious frame. A scale parameter is computed from the time interval between the fictitious frame and the adjacent frames. While trilinear interpolation is used to interpolate translation quantities, rotations have to be interpolated over the sphere to achieve constant-speed motion. This is achieved by the slerp operation. Therefore, the pose relations between the slave frame and the adjacent

29th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan | https://doi.org/10.5194/ica-abs-1-362-2019

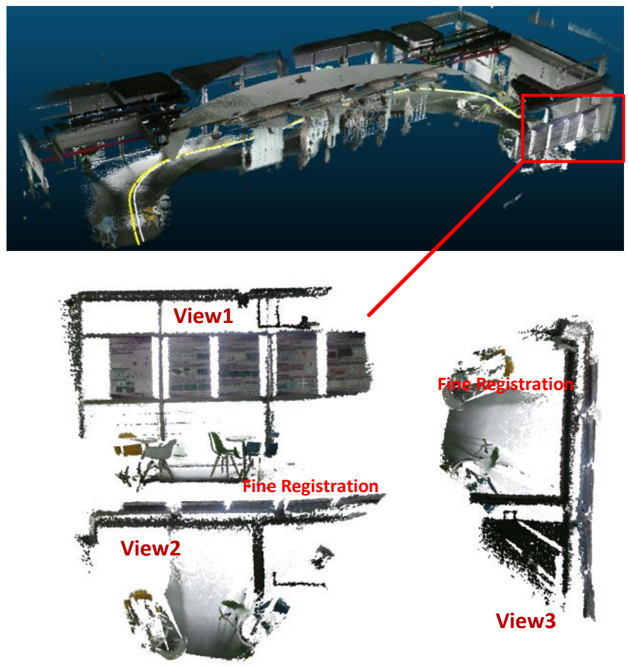
reference frame can be derived, which provided opportunity to use the more accuracy observations from multiple frames for better tracking and global optimization. Finally, the experiments in complex indoor scenarios demonstrate the efficiency of the proposed multiple RGB-D slam algorithm.



(a)Camera trajectories of sensors



(b)Dense point cloud obtained by the system before frequency-drift compensated



(c)Dense point cloud obtained by the system after frequency-drift compensated

Figure 2. Comparison of camera trajectories and dense point cloud between before and after frequency-drift compensated

[1] S. Yang, X. Yi, Z. Wang, Y. Wang, and X. Yang, "Visual SLAM using multiple RGB-D cameras," in *Robotics and Biomimetics (ROBIO), 2015 IEEE International Conference on*, 2015, pp. 1389-1395: IEEE.

[2] C. Chen *et al.*, "Calibrate Multiple Consumer RGB-D Cameras for Low-Cost and Efficient 3D Indoor Mapping," *Remote Sensing*, vol. 10, no. 2, p. 328, 2018.