An Efficient 3D Scene Modeling Algorithm Based on a Pan-Tilt-Driven RGB-D Camera

Chi-Yi Tsai, Chih-Hung Huang, and Shang-Jhih Jhang

Abstract—This paper proposes an indoor scene three-dimensional (3D) reconstruction system using a pan-tilt-driven RGB-D camera. The proposed system is able to automatically reconstruct 3D indoor scenes on a fixed position. An efficient point cloud registration algorithm is proposed to align point clouds based on extrinsic parameters of the RGB-D camera from every presetted pan-tilt control points. Then, a local registration method is used to refine the alignment result. Experimental results verify the quality and efficiency of the proposed point cloud alignment algorithm.

I. INTRODUCTION

3D scene reconstruction is a fundamental function in many applications of robot vision, i.e. environment mapping [1], scene recognition [2] and augmented reality [3], etc. Modern point cloud alignment algorithms usually require a lot of computing power, and if the point cloud the greater the amount of data, the amount of computing will increase non-linearly. In this paper, we calibrate the transformation matrix of the motor control points in the offline state, and use the computed transformation matrix in the online state directly. This can be an efficient solution, because only need a transformation computation for aligning two point clouds.

II. METHODS AND RESULTS

The purpose of point cloud registration is to find the 3D transformation between two point clouds. The transformation can be defined by a rotation matrix $R$ and a translation vector $t$. To obtain both $R$ and $t$, there are two different approaches: local registration and global registration. Local registration requires two initial point clouds $P$ and $Q$ to be close enough, so that it can find a rigid body transformation of $P$ and $Q$. By contrast, global registration can start from two point clouds with any initial posture.

There are many existing methods that use global registration to find correspondences between point clouds or use geometric method to find the local correspondences. After finding the correspondences between $P$ and $Q$, the cost function $E_{\text{global}}$ is then defined as

$$E_{\text{global}} = \sum_{k} \|R_{p}^{k} t + \bar{q}^{k}_{f} - \bar{q}^{k}_{g}\|^{2},$$

where $p^{k}_{f}$ and $q^{k}_{f}$ are corresponding points. Global registration can be resolved using optimization, but its computation complexity is usually higher than the local registration. To speedup global registration, this paper proposes a 3D scene modeling method which combine motor control and camera calibration. The system performs camera calibration in offline, and obtains relative transformation matrices between fixed postures through camera calibration. These transformation matrices are then used in the online global registration.

The hardware devices used in the experiment includes a notebook computer, a RGB-D camera and a pan-tilt unit. We used ASUS Xtion Pro RGB-D camera and fixed the RGB-D camera to the D46-17 pan-tilt unit manufactured by FLIR. The pan-tilt unit was used to drive the RGB-D camera to capture point clouds. Figure 1 shows the experimental results, in which the pictures listed in upper row are the original point clouds. The lower row of Figure 1 presents the registration result of the proposed method. It is clear that an initial registration result can be obtained efficiently.

![Figure 1. Experiment results. The upper row shows two original point cloud datas, and the lower row is the experimental results of the proposed method.](image)

III. CONCLUSION

In this paper, we propose a method to align the scattered point clouds based on camera calibration. This method can quickly converge and is able to process in real time. In the future, we will be further extended this method with different motion platform to implement 3D map building function.

REFERENCES