Motion Segmentation based on On-line Non-parametric Learning using RGB-D Data

Yuxiang Sun¹, Student Member, IEEE and Max Q.-H. Meng², Fellow, IEEE

Abstract—Motion segmentation is a fundamental technology in many robotic applications, such as mapping and navigation in dynamic environments. In this study, we propose a novel motion segmentation approach based on on-line non-parametric learning using RGB-D data. The proposed approach requires no prior information, such as hand-labelled initial segmentation. Foreground cues are derived from dense optical flow with the homography constraint. Visual and depth information of moving objects are learned on the fly to maintain a foreground model, which is incrementally updated during the iterations. We evaluate the approach using public sequences. The results demonstrate that our approach is able to effectively segment moving objects with a freely moving RGB-D camera.

I. INTRODUCTION

Motion segmentation from moving platforms has been studied over the past decades. It is a fundamental technology in robotic applications, such as mapping and navigation in dynamic environments [1]. Many methods have been proposed using various types of prior information and assumptions. For example, the shape or appearance information of moving objects can be learned from hand-labelled initial segmentation, and then the same pattern can be matched in the following frames to segment the moving objects [2]. In our method, we do not require such prior information. We use the geometric constraint to derive motion cues. The key assumption required in our method is that static objects dominate the scenes. In the following sections, we present the overview of our method and the experimental results tested using the public TUM RGB-D sequences [3].

II. METHOD OVERVIEW

We generally divide our approach into two stages: the learning stage that builds and updates a non-parametric foreground model, the inference stage that densely segments the foreground. Note that the two stages run simultaneously and our approach is on-line.

In the learning stage, we firstly use dense optical flow to calculate the dense pixel matchings between two consecutive RGB images. Then, we compute the homography between the two images with the pixel matchings and calculate geometric reprojection errors for each pixel using the homography. The reprojection errors serve as the motion cues for the foreground. We compute the foreground likelihood for each pixel using the reprojection errors. Finally, we use the RGB-D data from the points with higher foreground likelihood to build and update the foreground model. In the inference stage, we pixel-wisely compare the RGB-D data in the current frame with the built model. The foreground is densely determined through the comparison.

III. EXPERIMENTAL RESULTS AND DISCUSSION

Fig.1 shows sample experimental results. As we can see, two persons are moving in the scene and they are correctly segmented. This demonstrates the effectiveness of our method. However, there are some false positives in the segmentation results. For example, the books close to the left person in the left figure and the desk close to the right person in the right figure are misclassified as foreground. We think the reason for this is the close RGB-D values between the foreground and the background.

REFERENCES