Conditions for segmentation of 2D translations of 3D objects

Abstract

Various computer vision applications involve recovery and estimation of multiple motions from images of dynamic scenes. The exact nature of objects' motions and the camera parameters are often not known a priori and therefore, the most general motion model (the fundamental matrix) is applied. Although the estimation of a fundamental matrix and its use for motion segmentation are well understood, the conditions governing the feasibility of segmentation for different types of motions are yet to be discovered. In this paper, we study the feasibility of separating 2D translations of 3D objects in a dynamic scene. We show that successful segmentation of 2D translations depends on the magnitude of the translations, average distance between the camera and objects, focal length of the camera and level of noise. Extensive set of controlled experiments using both synthetic and real images were conducted to show the validity of the proposed constraints. In addition, we quantified the conditions for successful segmentation of 2D translations in terms of the magnitude of those translations, the average distance between the camera and objects in motions for a given camera. These results are of particular importance for practitioners designing solutions for computer vision problems.

Keywords — Motion segmentation, multibody structure-and-motion, fundamental matrix, robust estimation.