3D RECONSTURCTION OF SMALL-SCALE INTERIOR SURFACES

Katsiaryna Halavataya[1], Vasiliy Sadau[1]

[1]Faculty of Radiophysics and Computer Technologies, Belarusian State University, Belarus
katerina-golovataya@yandex.ru

3D reconstruction is a problem of creating a representation of real-world objects and structures as point cloud, mesh or polygonal 3D models based on regular observation methods. The most common form of input data for 3D reconstruction is a set of photographs of an object of study, usually taken manually with several points of view, by single or multiple cameras, in order to capture the object from different angles. Another relatively common form of input is a video sequence.

A photographic representation of an object can be thought of geometrically as a central projection of the scene onto the image plane. Epipolar camera optics and projective transformations provide a sufficient model for tracing projections of the same point across several images to a single point in three-dimensional space. These problems are traditionally solved using photogrammetry structure from motion (SFM) methods and involve a number of steps: calibration, keypoint detection and matching, bundle adjustment, sparse and dense point cloud reconstruction, mesh estimation and texturing [1].

However, there are several limitations of existing structure from motion methods that prevent their use in several important applications. One of them is the 3D reconstruction of medical images, allowing to enhance diagnostic capabilities and help with optimal treatment planning. Specifically, 3D reconstruction from video endoscopic images has the following distinctions:

- Object size. Traditional structure from motion photogrammetry assumes specific physical size of the objects and scenes. Near and mobile photogrammetry methods are used to reconstruct the objects with sizes ranging from 0.5m to 200m across any of the dimensions, aerophotographic and sattelite photogrammetry work with larger sizes, while microscopic image photogrammetry works with electron microscope scans for nanoscale objects with high magnification factor. Videoendoscopic research usually works with objects 0.1-5 cm of size.
- Optical system. Structure from motion methods are very sensitive to various geometric transformations imposed by optical system of a camera, specifically distortion and fish-eye effects. To mitigate these effects, camera calibration is mandatory to adjust to optical system irregularities around the edges of the image. However, videoendoscopic systems typically use ultra-wide angle cameras with fields of view up to 140°, which create heavy non-linear distortion effects around the entire image.
- Interior scene acquisition. It is assumed that objects described with photogrammetry methods can be isolated and exterior image acquisition can be used to observe the object independently of the background. In videoendoscopic research, an interior image acquisition is performed, most of the objects that need to be analyzed cannot be separated from the background, and the relevant part of a 3D model should be based on interior surface organ area.
- Dynamic nature of the environment. Usually, structure from motion analysis is performed on a static object that doesn't move, and its geometry is assumed to be constant across different camera positions. Surface areas of interior organs, however, may expand and contract with time and generally have a dynamic geometry within specific constraints.
- Dynamic nature of lighting and surface reflections. Classic photogrammetry methods assume static lighting and non-reflective surfaces of the object to correlate object points across images based on their absolute brightness. In videoendoscopic research, however, the light source is rigidly connected to the camera and moves along with it. Moreover, inner organs are usually covered in liquids, resulting in abdundant reflections and glares that fluctuate with the light source movement.

All of the described problems prevent the direct usage of traditional photogrammetry methods for endoscopic research 3D reconstruction and mandate the need to create adjustments to existing models [2, 3]. The paper proposes several solutions to these problems. For object size limitation, it's necessary to use more precise keypoint detection methods that select points of interest based on smaller features of the endoscopic image (like blood vessels and formation edges), as well as adjusted higher-resolution point cloud reconstruction techniques. Optical system specifics mean that a combination of ultra-wide projections must be calculated and used for distortion correction. Interior scene acquisition requires a reworked bundle adjustment process with point correspondence estimation across the depth axis of the image. Environment motion can be compensated with the introduction of correction matrix to the fundamental and essential matrix recaltulation algorithms during bundle adjustment. Dynamic nature of lighting can be mitigated by using a brightness-invariant colorspace transform, for example, 2-nd order component of principle component analysis, while glares can be compensated based on surrounding video sequence frames. Reconstruction based on adjusted methods allows for accurate and detailed 3D model acquisition for these types of images and video sequences.

[1] T. Luhman, S. Robson, S. Kyle, J. Boehm, Close-Range Photogrammetry and 3D Imaging, De Gruyter; 2nd ed. edition, 702 p. (2013).

[2] K. Halavataya, A. Kurachkin, Realtime computer-aided object detection in endoscopic screening, International conference Open Semantic Technologies for Intelligent Systems (OSTIS-2018), Minsk, BSUIR, pp. 247-250 (2018).

[3] J. C. McGlone, Manual of Photogrammetry, ASPRS; 6th ed. 1372 p. (2013).