論 文 の 内 容 の 要 旨

論文題目　Motion Estimation of Spherical Cameras and 3D Reconstruction Based on Sparse and Dense Pixel Flows
（スパース・デンスピクセルフローに基づく全天球カメラの運動推定と3 次元復元）

氏　　名　パトハック　サーサク

Cameras are ubiquitous sensors for robotics and can be used for measurement as well as visualization. One popular task involving cameras lies in motion estimation and 3D reconstruction of the surrounding environment. This is very useful for many applications such as for surveying disaster areas, digitization of structures, archeology, topographical mapping, etc. A common way to do this is by tracking the flow of image pixels in the view of the camera in order to back-calculate camera motion and to reconstruct their 3D positions. As compared to regular perspective cameras, spherical cameras are considerably advantageous. They can see in all spatial directions, they have no directional bias of information, and information almost never goes out of view. They can considerably broaden the scope of all such applications, especially in enclosed environments.

However, their projective properties are different from planar perspective cameras. This is because spherical images do not exist on a planar manifold like perspective images. Expressing them on a plane induces large distortions which change drastically as the camera moves. This can induce mistakes in tracking the flow of pixels. Since the consensus of many pixel flows is used for estimating camera motion, followed by 3D reconstruction, the presence of mistakes can cause large errors and considerably affect the accuracy and robustness of such systems. Furthermore, the large distortions also make it difficult to visualize spherical videos obtained from moving robots.

Hence, in this thesis, accurate and robust methods for camera motion estimation and 3D reconstruction using spherical cameras are proposed. This is done by using a combination of *sparse* and *dense* pixel flow information. *Sparse Pixel Flows* refers to the flow information of a sparse set of distinguishable pixels inside the image that can be estimated using feature detection and matching. This is possible between images captured over large camera displacements, but can have mistakes known as *outliers*. Meanwhile, *Dense Pixel Flows* refers to the flow information of every pixel inside the image obtained by dense optical flow techniques. This can only be estimated over small camera displacements, but is free of drastic outliers. This is because dense optical flow techniques employ various spatial smoothness constraints and suppress local outliers. In addition to a reliable consensus of camera motion, they also make it possible to reconstruct every pixel in 3D.

The advantages of estimating both sparse and dense pixel flows are combined in order to achieve robust and accurate estimation of camera motion and to obtain a dense 3D reconstruction of the surroundings from a sequence of spherical images. Sparse pixel flows are used initially to cover for large camera displacements and obtain a rough estimate of camera motion and a sparse 3D model. Following this, the estimate of camera motion is refined and a dense 3D model is obtained using the accurate consensus from dense pixel flows. In addition, dense pixel flows are also used to estimate camera motion in order to stabilize a spherical video. In each method, spherical camera geometry is taken into account, taking advantage of the complete spherical field of view while bypassing the high distortion.

In Chapter 1, the myriad applications that involve motion estimation and 3D reconstruction using cameras are introduced. It is explained how spherical cameras, that can see in all directions, can greatly increase the scope and effectiveness of all such applications. Following this, the challenges in applying spherical cameras to such applications are detailed. Specifically, it is explained how the distortion and difficulty of estimating pixel flows can reduce the robustness and accuracy of all such applications. Next, the aim of this thesis is detailed and the approach chosen to solve these issues, i.e. the combination of sparse and dense pixel flows, is explained.

In Chapter 2, the construction of spherical cameras using multiple wide-angle fisheye lenses is explained. This chapter also talks about the geometric basics that are required to understand parts of this thesis. Specifically, it covers the concepts of how environmental points are projected on a spherical camera and how spherical images are formed. Following this, the crucial constraints behind multi-view geometry i.e. the manners in which pixels flow within spherical images captured from a moving camera are explained. Finally, this chapter explains how these pixel flows can be estimated from image information.

In Chapter 3, various research related to the objective of this thesis is explained. The ideas behind each related research that inspired the concepts used in this thesis are also mentioned.

Next, the main content of this thesis, i.e. the research conducted towards achieving the objective mentioned in Chapter 1 is explained in a detailed manner, drawing from the theory explained in Chapter 2. Specifically, this research includes the following methods:

In Chapter 4, a method that uses sparse pixel flows, estimated from feature detection and matching, in order to estimate the trajectory of a moving spherical camera and sparsely reconstruct the environment is proposed. This method forms the initial rough estimation that is used to cover large camera displacements and initialize the refined estimation using dense pixel flows in the remaining chapters of the thesis. In addition, this research was also used for a practical application of spherical camera localization involving bridge inspection developed for the Cross-ministerial Strategic Innovation Promotion Program (SIP), Infrastructure Maintenance, Renovation, and Management. The results obtained during the field experiments are also presented in this chapter.

In Chapter 5, an approach that can decompose the dense pixel flow information between two images in order to estimate both their epipolar geometry and the surrounding 3D structure simultaneously is proposed. This is done via a novel rectification procedure that can completely bypass the distortion of spherical images, making it highly accurate and robust to spherical image distortions.

In Chapter 6, a technique that uses the previously estimated dense 3D reconstruction in order to estimate the positions and orientations of multiple spherical images in an image sequence is proposed. This is done by projecting the 3D reconstruction to create a `virtual image' at a given position and orientation and comparing it to the actual expected image. Thus, a motion estimation and 3D reconstruction pipeline based on dense pixel flows is set up, analogous to the method using sparse pixel flows in Chapter 4.

In Chapter 7, a fast method that uses dense pixel flows to stabilize the rotation of a freely moving spherical camera is proposed in order to provide a `rotationless virtual camera' for enhanced visualization of a spherical video. The method presented in this chapter uses a unique symmetry property of spherical cameras in order to ignore translation and avoids expensive 3D processing that is unnecessary for stabilization.

Finally, in Chapter 8, this thesis is summarized and potential future work is briefly explained.