

論文 / 著書情報
Article / Book Information

題目(和文)	ズーム機能を持つ両眼三次元シーン再構築システム
Title(English)	An Intelligent Binocular Vision System with Zoom Lens for 3D Scene Reconstruction
著者(和文)	竇書洋
Author(English)	Shuyang Dou
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10547号, 授与年月日:2017年3月26日, 学位の種別:課程博士, 審査員:長橋 宏,熊澤 逸夫,小池 康晴,山口 雅浩,小尾 高史
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10547号, Conferred date:2017/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

専攻： 物理情報システム 専攻
Department of
学生氏名： 寶 書洋
Student's Name

申請学位(専攻分野)： 博士 (工学)
Academic Degree Requested Doctor of
指導教員(主)： 長橋 宏
Academic Advisor(main)
指導教員(副)：
Academic Advisor(sub)

要旨(英文 800 語程度)
Thesis Summary (approx.800 English Words)

3D scene reconstruction is a popular and promising research subject in Computer Vision (CV) field. It closely relates to many other research problems, like simultaneous localization and mapping, augmented reality, autonomous car, etc. By far, enormous efforts have been invested in this subject, and many sophisticated approaches have been proposed during the last decades. However, it is still considered to be one of the most challenging problems in CV area.

The multi-view reconstruction algorithm is a representative approach for reconstructing 3D scene. For such kind of solution, camera calibration is usually considered as a necessary step. The goal of camera calibration is mainly estimating intrinsic camera parameters. In many existing approaches, these parameters are either extracted from Exchangeable image file format (Exif) meta data or computed by some interpolation strategies. However, Exif meta data suffer from low accuracy, and interpolation method is time consuming if we want accurate calibration results.

In this work, we propose a new calibration method to estimate the focal length and principle point for zoom camera. Similar to the interpolation method, calibration results produced by an off-line algorithm at several positions are served as seeds data. Camera parameters are estimated by analyzing the change of corresponding points. Experimental results showed that our method could generate more accurate calibration results than the linear interpolation method.

One of the important problems when a zoom camera is used is how to control the zoom lens properly. Zooming out too much may suffer from obscure scene details, while zooming in too much may result in failed photos in 3D reconstruction due to insufficient correspondences, or planar degeneracy. Currently, most existing approaches need manual zoom control. However, it is hard and inefficient to decide the proper zoom position manually.

An automatic zoom control method is presented in this work. We define two terms, Pixel Per Unit (PPU) and Normalized PPU (NPPU), for measuring the zoom appropriateness. Roughly speaking, these terms measures the ratio between the size of an object in 3D world and its projection size on the image plane. The key for our proposed solution is to keep the PPU constant. By using our method, the size of an object's projection on the image plane will be close to unchanged, even if its distance to the camera varied. In order to compute the current PPU accurately, we employ a binocular vision system and carry out the sparse point cloud reconstruction. Simulation experiments showed that our computation of PPU were close to the ground truth.

We use two existing tools to estimate the 3D model of the scene. However, it is hard to perceive the scene with just this 3D shape model, because it only contains the geometric information about the scene. Proper textures need to be attached onto this shape model, in order to make the final scene model more realistic.

A new texture generation method is presented in this work. Our method can cope with both

image noise and out-of-focus blur. We optimally choose the most suitable view for texture extraction. This is done by minimizing an energy function defined on a Markov Random Field (MRF) which is constructed from the geometric mesh model. The energy term is defined so that views with large projection area and more feature points will be preferred. This term is actually a hybrid energy term. After extracting textures from selected views, we refine the texture seams among adjacent mesh faces by using a global correction algorithm, followed by a local Poisson image editing procedure.

A simple but efficient strategy for further improving the texture quality is also proposed. The key for this strategy is using additional large zoom photos only for texture mapping step. Most existing methods use the same image set for both 3D reconstruction and texture generation. Unlike these methods, we use additional images with much clearer scene details only for texture mapping. Though it is simple to just add these images for both 3D reconstruction and texture mapping. In order to suppress the computation cost, we suggest using them only for texture generation. Experimental results showed that by applying this strategy, texture quality could be improved a lot, while the total computation time only increased slightly.

By combining these algorithms together, we constructed a binocular vision system. In this system, two Pan-Tilt-Zoom (PTZ) network cameras are used for perceiving the scene, and a robot is used for movement. This system can dynamically estimate camera's focal length and principle point accurately. It can also automatically control the zoom lens properly. It has been shown that our system can easily reconstruct high quality 3D scene model.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note : Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).