様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　３ 年　 ６ 月 ２１ 日現在

機関番号：１７１０２
研究種目：若手研究
研究期間：2019～2020
課題番号：１９Ｋ２０２９７

研究課題名（和文）Unifying multiple RGB and depth cameras for real-time large-scale dynamic 3D modeling with unmanned micro aerial vehicles

研究課題名（英文）Unifying multiple RGB and depth cameras for real-time large-scale dynamic 3D modeling with unmanned micro aerial vehicles

研究代表者
ＴＨＯＭＡＳ　ＤＩＥＧＯ（Thomas, Diego）

九州大学・システム情報科学研究院・助教

研究者番号：１０８０４６５１

交付決定額（研究期間全体）：（直接経費）　　3,200,000円

研究成果の概要（和文）：３Dマッピングのために、計算速度を上げることなくデプス画像をロバストかつ正確に融合するための新しい方法を提案しました。また、動く人のようなダイナミックなシーンにも対応できる方法を提案しました。
デプス情報が利用できない状況を処理するために、単一のRGB画像からの3D形状推定の方法を提案しました。服を着た人間の詳細な形状を再構築するための新しいディープニューラルネットワークを提案しました。
RGB-Dデータをキャプチャできる民生用空中ドローンを搭載した新しい3Dスキャンシステムを提案しました。提案するシステムは、ポータブルバッテリーとRGB-Dカメラを搭載したミニコンピューターで構成されています。

研究成果の学術的意義や社会的意義
単一の画像から人体の緻密で詳細な3Dモデルを作成する最初の方法の1つを提案しました。 コンピュータビジョンのトップ国際会議で私たちの仕事を発表し、コードを一般に公開しました。 制御されていない環境での3Dシーン再構築の新しい可能性を開きます。
このシステムを消費者向けの空中ドローンに簡単に装備するためのソリューションと、それを制御するソフトウェアを紹介します。 実世界のデータをキャプチャし、システムを使用していくつかの最先端のRGB-DSLAM技術を評価しました。 空中3Dスキャンとマッピングの研究開発を後押しできるように、すべてのデータとコードをコミュニティで公開しました。

研究成果の概要（英文）：The goal of this project was to develop a real-time system for 3D reconstruction of dynamic scenes with unmanned micro aerial vehicles.
We proposed a new method for robust and accurate fusion of depth images without increasing computational speed. We also proposed a method that can handle dynamic scenes like a moving person. This was achieved by jointly optimizing non rigid motion and geometry. To handle situations when depth information is not available, we proposed a solution for 3D shape estimation from a single RGB image. We focused on the case of the human body and proposed a new deep neural network to reconstruct detailed shapes of humans wearing loose clothes from single RGB images.
We proposed a new 3D scanning system equipped on a consumer-grade aerial drone that can capture live sequences of RGB-D data. our proposed system consists of a minicomputer powered by a portable battery and an RGB-D camera. We shared material and code and captured real world data with our system.

研究分野：Computer Vision

キーワード：RGB-D SLAM　Aerial 3D capture　Dynamic scene　Aritificial intelligence　Deep neural network　3D human body

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

Real-time 3D reconstruction of large-scale dynamic scenes (i.e., scenes containing one or more moving objects to be modeled, possibly with shape deformation) remains an unsolved problem. This is a critical problem for systems in need of 3D information that adapt to quickly changing environments. For automatic monitoring of flooded areas, for example, it is vital to obtain large- scale 3D maps that are updated a few times per second. Without this, accurate prediction of dangerous zones cannot be made on time to identify safe escape routes. In this work, I will investigate fusion of multiple RGB and depth sensors mounted on multiple micro aerial vehicle for real-time 3D reconstruction of large-scale dynamic 3D scenes.

２．研究の目的

The objective of this research project was to develop a real-time system for 3D reconstruction of dynamic scenes with unmanned micro aerial vehicles (like commercial quadrotors). This is extremely difficult because consumer-grade 3D scanners (like RGB-D cameras) do not work well outside. As a consequence, multiple sensors must be equipped on the drone and fused together to capture 3D data of a large area. It is also desirable that the algorithm allows 3D reconstruction of non-rigid deformations.

３．研究の方法

To reconstruct non-static 3D scenes, deformations of the surfaces must be tracked while the 3D model is incrementally refined. The major challenges to overcome here are three-fold: (1) find a 3D representation that can model any shape with few memory and low processing consumption for scalability; (2) incrementally update the shape of the 3D model to integrate new measurements in real- time; (3) track non-rigid deformations and adapt the 3D model to these deformations. Findings unveiled here will be used to build real-time evolving 3D maps from unmanned micro aerial vehicles.

４．研究成果

First year of the project

Depth image fusion. In the first year of the project, we proposed a new method for robust and accurate fusion of depth images. We proposed a probabilistic optimization using variational message passing in a Bayesian network. Our formulation enabled us to fuse depth images robustly, accurately, and fast for high quality RGB-D keyframe creation, even if exact point correspondences are not always available. Our formulation also allows us to smoothly combine depth and color information without increasing computational speed. Our proposed framework achieved promising results for reconstructing accurate 3D models while using low

computational power and being robust against misalignment errors without post-processing. The results of this work were published at 3DV 2019.

3D reconstruction from single images. Depth information is not always available when reconstructing 3D models of outdoor scene. This is because of interferences from sunlight or distance between camera and object. Therefore, at some time only 2D color images can be. Recovering 3D data from a 2D image is ill-posed due to ambiguities but with the help of convolutional neural networks (CNN) and prior knowledge on the 3D scene it is possible to overcome such ambiguities and recover detailed 3D shapes from single images. We proposed a solution in the case of reconstructing human bodies from single images. We proposed a new volumetric model that is compact, dense, accurate, and yet well suited for CNN-based regression task. Our proposed network allows us to reconstruct detailed shapes of humans wearing loose clothes from single RGB images. The results of this work were published at CVPR 2020.

Second year of the project

Dynamic scene reconstruction. We studied the special case of the human body. The canonical Truncated Signed Distance Function (TSDF) grid equipped with a graph of warp nodes is popularly used for volumetric real-time non-rigid 3D shape reconstruction. Jointly optimizing the non-rigid warp field and the TSDF field in real-time, however, requires complicated implementation and significant engineering effort, deeming reproducibility hard to achieve. We propose a method for real-time human motion tracking and 3D body reconstruction that is cheap in memory consumption, handles fast motion and topological changes, while relatively simple to implement, and capable of producing 3D models with high accuracy. The results of this work were submitted for publication at ICIP 2021.

Drone-based 3D scanning. We designed and built the hardware to mount three RGB-D cameras on a consumer-grade quadrotor. Consumer-grade micro unmanned aerial vehicles (MUAV) have recently become popular and open new possibilities for automatic 3D reconstruction of large-scale outdoor scenes. However, it is extremely difficult to equip a MUAV with a 3D scanner because of strict constraints of space and weight. We proposed a new 3D scanning system equipped on a consumer-grade aerial drone that can capture live sequences of RGB-D data. our proposed system consists of a minicomputer powered by a portable battery and an RGB-D camera. We present a solution to easily equip this system onto a consumer-grade aerial drone (in a plug-and-play style), as well as the software to control it. We captured real world data and evaluated several state-of-the-art RGB-D SLAM techniques with our system. Our proposed method was submitted for publication at the international conference on robotics (IROS 2021) and MIRU 2021.

５．主な発表論文等

〔雑誌論文〕　計6件（うち査読付論文　6件／うち国際共著　6件／うちオープンアクセス　0件）

１．著者名 Thomas Diego、Sirazitdinova Ekaterina、Sugimoto Akihiro、Taniguchi Rin-ichiro	４．巻 0
２．論文標題 Revisiting Depth Image Fusion with Variational Message Passing	５．発行年 2019年
３．雑誌名 2019 International Conference on 3D Vision (3DV)	６．最初と最後の頁 328-337
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/3DV.2019.00044	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

１．著者名 Maxence Remy、Uchiyama Hideaki、Kawasaki Hiroshi、Thomas Diego、Nozick Vincent、Saito Hideo	４．巻 0
２．論文標題 Mobile Photometric Stereo with Keypoint-Based SLAM for Dense 3D Reconstruction	５．発行年 2019年
３．雑誌名 2019 International Conference on 3D Vision (3DV)	６．最初と最後の頁 574-582
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/3DV.2019.00069	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

１．著者名 Hayato Onizuka, Zehra Hayirci, Diego Thomas, Akihiro Sugimoto, Hideaki Uchiyama, Rin-ichiro Taniguchi	４．巻 0
２．論文標題 TetraTSDF: 3D Human Reconstruction From a Single Image With a Tetrahedral Outer Shell	５．発行年 2020年
３．雑誌名 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	６．最初と最後の頁 6011-6020
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

１．著者名 Sayo, Akihiko, Hayato Onizuka, Diego Thomas, Yuta Nakashima, Hiroshi Kawasaki, and Katsushi Ikeuchi	４．巻 0
２．論文標題 Human shape reconstruction with loose clothes from partially observed data by pose specific deformation	５．発行年 2019年
３．雑誌名 Pacific-Rim Symposium on Image and Video Technology (PSIVT)	６．最初と最後の頁 225-239
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1．著者名 Xue, Yu, Diego Thomas, Frederic Rayar, Hideaki Uchiyama, Rin-ichiro Taniguchi, and Boacai Yin	4．巻 0
2．論文標題 Blended-Keyframes for Mobile Mediated Reality Applications	5．発行年 2019年
3．雑誌名 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)	6．最初と最後の頁 211-216
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1．著者名 Sun, Cheng, Diego Thomas, and Hiroshi Kawasaki	4．巻 0
2．論文標題 Unsupervised 3D Human Pose Estimation in Multi-view-multi-pose Video	5．発行年 2021年
3．雑誌名 International Conference on Pattern Recognition (ICPR)	6．最初と最後の頁 5959-5964
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕　計3件（うち招待講演　1件／うち国際学会　2件）

1．発表者名 Maxence Remy, Hideo Saito, Hideaki Uchiyama, Hiroshi Kawasaki, Vincent Nozick, Diego Thomas
2．発表標題 Merging SLAM and photometric stereo for 3D reconstruction with a moving camera
3．学会等名 25th International Workshop on Frontiers of Computer Vision（国際学会）
4．発表年 2019年

1．発表者名 Hayato Onizuka, Diego Thomas, Zehra Hayirci, Akihiro Sugimoto, Hideaki Uchiyama, Rin-ichiro
2．発表標題 Regression of 3D human body shapes from a single image
3．学会等名 Machine Perception and Robotics 2019（国際学会）
4．発表年 2019年

１．発表者名 Diego Thomas
２．発表標題 3D human body reconstruction using RGB-D camera
３．学会等名 Asia Pacific Society for Computing and Information Technology 2019 Annual Meeting (APSCIT 2019 Annual Meeting)（招待講演）
４．発表年 2019年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
フランス	Gustave Eiffel University			
米国	Standford			