論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Stomach 3D Reconstruction from Monocular Gastroendoscopy Video
著者(和文)	アジ レシンドラ ウィヂャ
Author(English)	Aji Resindra Widya
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12080号, 授与年月日:2021年9月24日, 学位の種別:課程博士, 審査員:奥冨 正敏,蜂屋 弘之,塚越 秀行,原 精一郎,田中 正行
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12080号, Conferred date:2021/9/24, 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

（博士課程）
Doctoral Program

論 文 要 旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	Systems and Control Eng.	系 コース	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of (Engineering)
学生氏名： Student's Name	Aji Resindra Widya		指導教員（主）： Academic Supervisor(main)	奥富正敏
			指導教員（副）： Academic Supervisor(sub)	

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

Minimally invasive diagnosis and intervention is becoming the mainstream for giving a care to the patient, overtaking traditional surgical operations which can cause severe injury or trauma. Since minimally invasive diagnosis and intervention are able to be performed without the need of open surgery, it gives less pain experience for the patient which leads to less trauma and quicker recovery time.

In this dissertation, we are focusing on gastroendoscopy. Gastroendoscopy is one of the minimally invasive diagnosis and intervention procedures to perform a diagnosis of various lesions inside the patient's stomach using an endoscope. Even though gastroendoscopy has various advantages for the patients, it introduces some new challenges for the practitioners such as lack of depth perception and difficulty in knowing the exact pose of the endoscope. These conditions affect two fundamental issues in navigating the gastroendoscopy, i.e., navigating the endoscope and localizing any found lesions. On top of that, limited human vision and sense with only limited assistance (usually assisted by 2D endoscope video only) make gastroendoscopy has a steep learning curve for many surgeons to finally be able to do proper disease localization. Yet, disease or lesion localization is very important for deciding the next important surgical steps. In this dissertation, we are aiming to develop a pipeline for providing additional information such as textured whole 3D model of the stomach, depth, and pose information for practitioners to navigate the gastroendoscopy and to localize any found lesions. The organization of this dissertation is summarized as follows.

Chapter 1, *Introduction* - This chapter introduces the current state of minimally invasive diagnosis and intervention procedures, its advantages, and disadvantages. This chapter reviews the solutions that have been proposed to alleviate the challenges found endoscope system such lack of depth information, limited point-of-views, and instrument tracking problems. Those proposed solutions include radiology technique such as computed tomography (CT) scan and image-based view expansion such as simultaneous localization and tracking (SLAM). This chapter also points out why the previously proposed solutions are not enough to tackle the navigation and localization problem in gastroendoscopy.

Chapter 2, *Monocular Endoscope Dataset Creation* - This chapter briefly introduces and explains the dataset we use throughout this experiment, including the capturing and processing methods. Each chapter that follows will explain how the dataset is used for each proposed method.

Chapter 3, *Structure-from-Motion-Based Whole Stomach 3D Reconstruction* - This chapter explains our approach in order to identify the location of a gastric lesion such as early cancer and a peptic ulcer

within the stomach. This chapter addresses to reconstruct the color-textured 3D model of a whole stomach from a standard monocular endoscope video and localize any selected video frame to the 3D model in an offline manner. This chapter firstly examines how to enable structure-from-motion (SfM) to reconstruct the whole shape of a stomach from endoscope images, which is a challenging task due to the textureless nature of the stomach surface. Then, this chapter explains the investigation results of combined effect of indigo carmine (IC) dye and color channel selection on SfM to increase the number of extracted feature points. A plane fitting-based outlier removal algorithm to improve the 3D reconstruction quality is also introduced in this chapter.

Chapter 4, *Whole Stomach 3D Reconstruction Using Virtual Chromoendoscopy Images* – Realizing that in Chapter 3, the spraying of IC dye needs additional time which is not desirable for both patient and practitioners, an alternative way to achieve whole stomach 3D reconstruction without the need of the IC dye is introduced in this chapter. This chapter explains how to generate virtual IC-sprayed (VIC) images based on image-to-image style translation trained on unpaired real no-IC and IC-sprayed images. The effect of input and output color channel selection for generating the VIC images is also explored. This chapter shows that the reconstruction results by of using VIC images compared with the reconstruction results using real IC-sprayed images have close results.

Chapter 5, *Learning-based Depth Estimation for Monocular Endoscope Video* – This chapter reviews the importance of depth information in endoscopy and explains the challenges to obtain it. Following that, this chapter explains our approach to provide simultaneous acquisition of RGB and depth data which can boost the capability of the endoscopy for various potential diagnostic applications using a learning-based method. A novel data generation strategy for self-supervised training for depth estimation for monocular gastroendoscopy is proposed in this chapter. This chapter describes how the whole stomach 3D reconstruction pipeline using IC dye can be used to obtain dense reference depth data for training. This chapter also explores the use of CycleGAN to make the depth estimation network applicable to general endoscopic images without IC dye.

Chapter 6, *Learning-based Simultaneous Depth and Pose Estimation* – This chapter states that to effectively tackle the endoscope navigation and the lesion localization challenges, only providing depth information is not enough. In this chapter, a novel supervised approach to train depth and pose estimation networks to assist the endoscope navigation in the stomach is proposed. This chapter outlines a novel generalized photometric loss function to avoid the complicated process of finding proper weights for balancing the depth and the pose loss terms, which is required for existing direct depth and pose supervision approaches.

Chapter 7, *Conclusion* – This chapter revisits our key ideas, concludes the dissertation, and outlines possible future trends of the works.