Artificial Intelligence in Medicine 59 (2013) 185–196

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Image partitioning and illumination in image-based pose detection for teleoperated flexible endoscopes

Charreau S. Bell[a,*], Keith L. Obstein[b,a], Pietro Valdastri[a,b]

[a] *Department of Mechanical Engineering, Vanderbilt University, 2301 Vanderbilt Place, PMB 351826, Nashville, TN 37235-1826, USA*
[b] *Department of Gastroenterology, Hepatology and Nutrition, Vanderbilt University Medical Center, 1030C MRB IV, 2215 Garland Avenue, Nashville, TN 37232-0252, USA*

ARTICLE INFO

Article history:
Received 1 November 2012
Received in revised form
20 September 2013
Accepted 23 September 2013

Keywords:
Artificial neural networks
Pose estimation
Localization
Optical flow
Visual odometry
Closed-loop control
Teleoperation
Narrow band imaging
Flexible endoscopes
Colonoscopes
Gastrointestinal endoscopy

ABSTRACT

Objective: Colorectal cancer is one of the leading causes of cancer-related deaths in the world, although it can be effectively treated if detected early. Teleoperated flexible endoscopes are an emerging technology to ease patient apprehension about the procedure, and subsequently increase compliance. Essential to teleoperation is robust feedback reflecting the change in pose (i.e., position and orientation) of the tip of the endoscope. The goal of this study is to first describe a novel image-based tracking system for teleoperated flexible endoscopes, and subsequently determine its viability in a clinical setting. The proposed approach leverages artificial neural networks (ANNs) to learn the mapping that links the optical flow between two sequential images to the change in the pose of the camera. Secondly, the study investigates for the first time how narrow band illumination (NBI) – today available in commercial gastrointestinal endoscopes – can be applied to enhance feature extraction, and quantify the effect of NBI and white light illumination (WLI), as well as their color information, on the strength of features extracted from the endoscopic camera stream.
Methods and materials: In order to provide the best features for the neural networks to learn the change in pose based on the image stream, we investigated two different imaging modalities – WLI and NBI – and we applied two different spatial partitions – lumen-centered and grid-based – to create descriptors used as input to the ANNs. An experiment was performed to compare the error of these four variations, measured in root mean square error (RMSE) from ground truth given by a robotic arm, to that of a commercial state-of-the-art magnetic tracker. The viability of this technique for a clinical setting was then tested using the four ANN variations, a magnetic tracker, and a commercial colonoscope. The trial was performed by an expert endoscopist (>2000 lifetime procedures) on a colonoscopy training model with porcine blood, and the RMSE of the ANN output was calculated with respect to the magnetic tracker readings. Using the image stream obtained from the commercial endoscope, the strength of features extracted was evaluated.
Results: In the first experiment, the best ANNs resulted from grid-based partitioning under WLI (2.42 mm RMSE) for position, and from lumen-centered partitioning under NBI (1.69° RMSE) for rotation. By comparison, the performance of the tracker was 2.49 mm RMSE in position and 0.89° RMSE in rotation. The trial with the commercial endoscope indicated that lumen-centered partitioning was the best overall, while NBI outperformed WLI in terms of illumination modality. The performance of lumen-centered partitioning with NBI was 1.03 ± 0.8 mm RMSE in positional degrees of freedom (DOF), and $1.26 \pm 0.98°$ RMSE in rotational DOF, while with WLI, the performance was 1.56 ± 1.15 mm RMSE in positional DOF and $2.45 \pm 1.90°$ RMSE in rotational DOF. Finally, the features extracted under NBI were found to be twice as strong as those extracted under WLI, but no significance in feature strengths was observed between a grayscale version of the image, and the red, blue, and green color channels.
Conclusions: This work demonstrates that both WLI and NBI, combined with feature partitioning based on the anatomy of the colon, provide valid mechanisms for endoscopic camera pose estimation via image stream. Illumination provided by WLI and NBI produce ANNs with similar performance which are comparable to that of a state-of-the-art magnetic tracker. However, NBI produces features that are stronger than WLI, which enables more robust feature tracking, and better performance of the ANN in terms of accuracy. Thus, NBI with lumen-centered partitioning resulted the best approach among the different

* Corresponding author. Tel.: +1 615 3220688.
E-mail addresses: charreau.s.bell@vanderbilt.edu (C.S. Bell), keith.obstein@vanderbilt.edu (K.L. Obstein), pietro.valdastri@vanderbilt.edu (P. Valdastri).

http://dx.doi.org/10.1016/j.artmed.2013.09.002

variations tested for vision-based pose estimation. The proposed approach takes advantage of components already available in commercial gastrointestinal endoscopes to provide accurate feedback about the motion of the tip of the endoscope. This solution may serve as an enabling technology for closed-loop control of teleoperated flexible endoscopes.

1. Introduction

Each year, colorectal cancer claims the lives of more than 600,000 people worldwide, and is the fourth leading cause of cancer-related death in the world [1]. Colorectal cancer commonly progresses to malignancy in approximately 5–10 years. However, this type of cancer has the unique quality that if the tumor is detected at an early enough stage, the prognosis for survival is 90%, whereas if detected too late, it decreases to 5% [2]. This emphasizes the relevance of timely screening for population at risk (i.e., people over 50 years of age or having family history of colorectal cancer), even in the case that no symptoms are observed.

The most commonly used method for diagnostic and therapeutic assessment of colorectal cancer is through colonoscopy, an endoscopic procedure which requires the insertion of a 1.5-m long flexible tube through the anus. The endoscope provides illumination for visualization of the colon lumen, which enables detection and removal of polyps. Standard colonoscopy is performed under white light illumination (WLI). However, this approach can fail to reveal important information [3]. Even experienced endoscopists can miss up to 30% of all potential cancer lesions when using standard WLI [4].

In the last decade, narrow band imaging (NBI) has been introduced to improve diagnosis. NBI uses filters to narrow projected light to blue (415 nm) and green (540 nm) wavelengths to generate a colored image. Blue-green light enhances superficial mucosal capillaries and mucosal surface patterns; greater absorption of illuminating bands by hemoglobin causes the blood vessels to look darker. Despite recent literature demonstrating that NBI does not increase the diagnostic yield when compared to WLI [5], this imaging modality is today increasingly common in commercial colonoscopes (e.g., H180AL/I, Olympus, Japan).

Although a colonoscopy usually takes less than 30 minutes and is performed in outpatient surgery under sedation, patient compliance with recommended screening is low (i.e., 1 in 3 adults are not being screened [6]) due to the preparation required, fear of pain during the procedure, and perceived embarrassment. The main technological improvements in the field of flexible endoscopy aim to help patients to overcome these hindrances.

An approach for accomplishing this goal is through the development of increasingly flexible endoscopes, wireless capsule endoscopy (WCE), and virtual colonoscopy [2]. Complementary to these advances is the emergence of computer-assisted technologies to aid the doctor, whose purpose is to increase detection of malignancies and control over the intended trajectory of the endoscope. Robotics is playing an increasingly important role in this field with the development of fully- or semi-automated endoscopic systems [2,7–11]. This technological breakthrough has the potential to widen the implementation of colorectal cancer screening and surveillance programs to rural areas, to mobile camps, or to in-field military bases, and the physical presence of an expert endoscopist may no longer be required.

Real-time pose (i.e., position and orientation) detection of the tip of an automated flexible endoscope is crucial to achieving reliable and effective teleoperation. These devices operate in an intricate, complex environment and by definition are compliant; many variables exist that cannot be accounted for in a model, which severely limits the efficacy of open-loop control. Furthermore, medical procedures require a high degree of precision and accuracy; implementing real-time pose detection allows for calculated, controlled movements which enhance system stability [12]. In particular, the real-time estimated pose of the endoscope head can be used as feedback signal for a closed-loop control strategy, as represented in Fig. 1. This allows us to minimize the error between the intended pose (i.e., where the user wants the endoscope to move and orient the camera), and the reached pose (i.e., the measured pose of the endoscope tip).

In order to achieve real-time pose detection, magnetic tracking has emerged as a reliable method and there are several commercial manufacturers of 5 or 6 degree of freedom (DOF) trackers [13,14]. Magnetic trackers placed along the entire length of the colonoscope, such as in the commercially available ScopeGuide® (Olympus, Japan), have been used to provide the endoscopist visual feedback of the instrument pose with respect to a global coordinate frame [15]. Within bronchoscopy, the endoscopic camera stream has been used in conjunction with image registration and fluoroscopy for tracking of the endoscope [16–19].

However, magnetic trackers require additional space in the endoscope; this results in an increase in the size of the device, and a corresponding reduction in the flexibility of the endoscope. For endoscopes with extremely limited operating space, such as encephaloscopes, rhinoscopes, and bronchoscopes, minimization of the size of the endoscope is fundamental. Furthermore, commercial players in the field of gastrointestinal endoscopy are proposing platforms that are based on magnetic manipulation of the endoscopic device [20,21]. This promising approach is also being pursued by several research labs worldwide [22–26]. Magnetic trackers interfere with magnetic manipulation due to the presence of metallic components or because the localization principle itself is based on triangulation of electromagnetic fields.

Tracking of the endoscope head is even more crucial in soft body cavities (e.g., colon, small intestine), since image registration is not effective. Thus, a localization system which is independent of the technology platform to which it is applied, provides accurate pose estimation for real-time feedback, and neither creates unwanted disturbance to the system nor adds additional size to the endoscope will be beneficial for enabling closed-loop control of teleoperated flexible endoscopes.

1.1. Related work

The problem of real-time localization and steering of flexible endoscopes has a number of challenges [27,28]. Concerning the use of the image stream to steer the endoscope through the lumen, possible approaches include finding the darkest region of the image for lumen detection [29,30], identifying features such as the ring-like contours surrounding the lumen [31–33], and using highlights resulting from illumination [34]. Several works have used lumen center detection schemes to correct the current heading of the camera towards the lumen center in each control loop [8,35], providing a mechanism for automation. However, these solutions do not measure the change in pose of the endoscope, and thus cannot be used to implement closed-loop control (i.e., although the motors can be actuated towards the center of the lumen, there is no feedback as to whether the actuation was successful).