Computers in Biology and Medicine 67 (2015) 83–94

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Fuzzy zoning for feature matching technique in 3D reconstruction of nasal endoscopic images

Surapong Rattanalappaiboon [a], Thongchai Bhongmakapat [b], Panrasee Ritthipravat [a,*]

[a] Department of Biomedical Engineering, Faculty of Engineering, Mahidol University, Nakorn Pathom, Thailand
[b] Department of Otolaryngology, Faculty of Medicine, Ramathibodi Hospital, Bangkok, Thailand

ARTICLE INFO

Article history:
Received 19 May 2015
Received in revised form
17 September 2015
Accepted 29 September 2015

Keywords:
3D reconstruction
Endoscope
Nasal cavity
Fuzzy logic
Structure from motion

ABSTRACT

3D reconstruction from nasal endoscopic images greatly supports an otolaryngologist in examining nasal passages, mucosa, polyps, sinuses, and nasopharyx. In general, structure from motion is a popular technique. It consists of four main steps; (1) camera calibration, (2) feature extraction, (3) feature matching, and (4) 3D reconstruction. Scale Invariant Feature Transform (SIFT) algorithm is normally used for both feature extraction and feature matching. However, SIFT algorithm relatively consumes computational time particularly in the feature matching process because each feature in an image of interest is compared with all features in the subsequent image in order to find the best matched pair. A fuzzy zoning approach is developed for confining feature matching area. Matching between two corresponding features from different images can be efficiently performed. With this approach, it can greatly reduce the matching time. The proposed technique is tested with endoscopic images created from phantoms and compared with the original SIFT technique in terms of the matching time and average errors of the reconstructed models. Finally, original SIFT and the proposed fuzzy-based technique are applied to 3D model reconstruction of real nasal cavity based on images taken from a rigid nasal endoscope. The results showed that the fuzzy-based approach was significantly faster than traditional SIFT technique and provided similar quality of the 3D models. It could be used for creating a nasal cavity taken by a rigid nasal endoscope.

1. Introduction

Three dimensional model reconstruction from multiple images takes an important role in various research areas, such as computer graphics, engineering design, and medical applications. The model can present an object of interest better than its 2D images. It is a very useful tool for medical education [1–3] in which medical students can thoroughly understand and visualize anatomical structure. The model may be printed out by a rapid prototyping process and used for surgical training, and prosthetic applications [4]. Basically, 3D reconstruction from multiple images can be categorized into 3 main approaches; (1) shape from silhouette [5], (2) shape from shading [6] and (3) structure from motion [7].

Shape from silhouette creates a 3D model of an object using its silhouettes taken from different views. This technique is simple and uses only geometrical information. The model can be reconstructed from intersection of visual cones that are generated by back projection of silhouette's object from each view. The intersection creates the "visual hull" [4]. Accuracy of the model depends on the number of acquired images. Even though there are enough input images, this technique cannot deal with concavities of the object.

Shape from shading recovers object's shape from gradual variation of shading in the object's image. Depth of the model is determined from a distance between the camera and the object surface. Surface reflectivity of the object and position of the light source must be known or predefined. This requirement is impractical in real applications. In addition, the technique is mathematically complicated. However, it had been used in endoscopic applications [8–11]. Shape from shading technique is suitable for an object that has uniform surface and textureless. For an object with glossy surface or textures, poor quality of the 3D model may be generated because different colors of textures or light reflection from glossy surface affect the object's shading.

Structure from motion uses camera movements and textures of an object to reconstruct the 3D model. This technique consists of four main steps. First, the camera is calibrated to get intrinsic and extrinsic parameters that are necessary for 3D reconstruction. Second, feature points on object's surface in each view are extracted. These points are used for representing the object. Third,

* Corresponding author. Tel.: +66 2 889 2138x63512.
E-mail address: panrasee.rit@mahidol.ac.th (P. Ritthipravat).

http://dx.doi.org/10.1016/j.compbiomed.2015.09.021
0010-4825/

feature points from different views are matched. Finally, triangulation method is applied to identify object location from the matched feature pairs. Camera parameters are used for estimating real dimension of the object. To obtain a good model, an object of interest should have sufficient feature points and these features should cover all areas of the object. Incomplete model may be reconstructed when feature points of a certain part are inadequately provided. Therefore, this technique is not suitable for a textureless object. Many researchers apply this technique to create 3D model of human organ such as gastrointestinal tract [12], face [13] and back [14].

In medical applications, objects of interest are relatively complicated. They may have both concave and convex areas. Shape from silhouette technique is thus rarely used. For shape from shading, many researchers had used this technique to reconstruct human's organ. However several assumptions are strongly assumed and may be invalid in reality. For example, the lambertian surface should not be assumed for face, bone, and some internal organs that contain glossy like surface, such as mucosal etc. Surface of human organ may possess various textures, such as pimple, scar or vascular patterns. These textures affect to the 3D reconstruction process. For structure from motion, it can be efficiently used with an object with textures. Strong assumption is not required. It is thus the most appropriate technique for medical applications. However, incomplete 3D model may be created when textures are not sufficiently generated.

In nasal endoscopic examination, rigid endoscope is basically used in investigating nasal obstruction, polyps, tumors, and bleeds. It is a routine daily practice for otolaryngologists. Interior surfaces and nasal passages can be viewed from a sequence of the images. 3D reconstruction of the nasal cavity from these images takes an important role in diagnosis and treatment in which irregularity can be obviously observed. The model can be used for demonstrating the narrowest or abnormal part for surgical correction of nasal cavity. Additionally, it can be used for estimating the degree of nasal obstruction and resistance. Presently, nasal endoscopy provides the distorted images because of fish eye lens. When the endoscope comes close to the object, the appearing size in monitor is bigger than that of the real object. With the 3D model, otolaryngologists can correctly estimate shape and size of the nasal cavity.

Two main approaches are generally studied in the endoscopic applications. They are shape from shading and structure from motion. For shape from shading, several researchers presumed position of the light source based on the structure of endoscope in which the camera is located next to the light source. For simplicity, the light source is normally assumed to be at the camera center [8–11]. Surface reflectivity is generally assumed to be the lambertian surface. This technique requires strong mathematical background to solve the image irradiance equation [8,9,11]. Nevertheless, shape from shading technique is not suitable for organs with a lot of textures.

For structure from motion, it can better handle with such organs and it does not require strong assumptions. Many researchers used this approach in 3D reconstruction of endoscopic images. In 2010, Sun et al. [15] studied 3D reconstruction of images acquired from a wireless capsule endoscope. Harris corner detection was used in feature extraction and Kanade–Lucus–Tomasi feature tracking was employed in feature matching. This system was tested with endoscopic images created from a phantom. The result showed that this approach could be applied to 3D model reconstruction of two endoscopic views. In the same year, Fan et al. [12] studied 3D reconstruction of real endoscopic images of gastrointestine tract that were captured by a wireless capsule endoscope and a traditional endoscope. SIFT was used in feature extraction and feature matching. The result showed that the proposed technique could generate 3D model of the gastrointestine tract. In 2011, Fan and Meng [16] applied Affine-scale invariant feature transform or ASIFT to 3D reconstruction of wireless capsule endoscopic images. ASIFT was used instead of SIFT because traditional SIFT technique was not suitable for unpredictable movement of wireless capsule endoscope. This technique was tested with real wireless capsule endoscopic images. The result showed that ASIFT provided a number of matching points more than traditional SIFT. In the same year, Du et al. [17] improved SIFT feature matching technique in 3D reconstruction of endoscopic images. The proposed algorithm was designed to reduce computational time of SIFT by narrowing a search space with a specified area. The number of matching candidate pairs could be decreased. This algorithm largely reduced computational time and increased the number of matched features for endoscopic images. However, fixed search space may be not appropriate for all features because each feature requires different size of search space based on depth of the feature.

Feature matching is the most important process in the structure from motion technique. The quality of the reconstructed model depends on the number of accurate matched pairs. Normally, Scale Invariant Feature Transform (SIFT) algorithm is used as the traditional technique for feature extraction and feature matching. SIFT is an algorithm that is widely used in object recognition, image stitching and 3D modeling. It can extract local features in an image that are invariant in scaling, illumination changing, rotation and translation. Though SIFT algorithm is a useful technique, its computational expense remains a major problem. Several researchers integrate fuzzy logic to SIFT technique in order to improve its performance. In 2008, Belacel et al. [18] applied the fuzzy set to SIFT feature. SIFT descriptor of every feature point in an image is transformed into fuzzy numbers. Fuzzy similarity measure is proposed and used in place of Euclidean distance. After testing with a single object image, the result showed that the proposed technique could be used for single object recognition with good performance. However, change of view angles could reduce the performance. In 2013, Shashikanth and Kulkarni [19] proposed a region based image similarity measure using fuzzy logic and SIFT. This technique required feature matching between a reference image and a test image. Feature point locations from both images were used as inputs of the fuzzy system in order to measure similarity of the features. The result showed that this proposed technique could remove mismatched feature points better than original SIFT.

From previous studies, however, SIFT algorithm and the proposed fuzzy based techniques still consumed large computational time particularly when it was applied to rigid endoscopic applications because each feature point from a certain frame must be compared with all features in its adjacent frame. Because view changes of the rigid endoscopic images are relatively small, locations of the features are gradually changed. It is not necessary to search in the entire image area.

This paper presents a fuzzy zoning approach to improve the feature matching process in 3D model reconstruction from a sequence of nasal endoscopic images. The proposed technique is combined with SIFT algorithm in order to reduce the computational time. Endoscopic images are created from phantoms. Performances between the proposed technique and SIFT algorithm are compared in terms of the matching time and errors of the reconstructed models. To show validity of the proposed technique, it is applied to two sequences of images taken from a rigid endoscope used in real nasal cavity examination.

2. Structure from motion

Structure from motion technique uses camera movement model and object textures in 3D reconstruction from different