Birthe Göbel*, Knut Möller

Challenging requirements and optical depth estimation techniques in laparoscopy

https://doi.org/10.1515/cdbme-2022-1175

Abstract: Minimally invasive surgery has many advantages and cannot be missed nowadays. It leads to faster recovery and less surgical trauma. In laparoscopic surgery years of training are required to optimize patient outcomes. To visually support the surgeon during stomach surgery a 3D-reconstruction of the whole organ shall be created prior and during the procedure. Results might be beneficial for various applications such as before-and-after documentation, navigation support and autonomous robotic surgery. The technical implementation of 3D-reconstruction requires depth estimation which is challenged due to the environmental conditions and surgical constraints that exist in the human body during minimally invasive surgery. This paper focuses on the requirements of 3D-reconstruction in laparoscopy, reveals current research challenges and proposes an evaluation framework for optical depth estimation techniques. Eight methods were included in the evaluation. Scores considering the requirements were established and assigned to each method. The methods Deformable Shape-from-Motion, Stereoscopy, Shape-from-Motion, Simultaneous Localization and Mapping, Structured Light and Light-Field Technology were shown to partially fulfill the requirements for laparoscopic 3D-reconstruction. Shape-from-Shading and Time-of-Flight need extensive modifications to be applicable. In conclusion it can be stated that currently no method exists to realize a real time high-resolution 3D-reconstruction of inner organs during laparoscopy.

Keywords: 3D-reconstruction, laparoscopy, endoscopy, optical depth estimation techniques

***Corresponding author: Birthe Göbel:** Erlanger Str. 38h, 95444 Bayreuth, Germany, e-mail: birthe.goebel@gmail.com
2nd Author: Knut Möller:
HFU, Villingen-Schwenningen, Germany

1 Introduction

So far research and industry could not develop a laparoscope with the ability to generate a 3D-reconstruction of the whole stomach and/or other organs. 3D organ models would come with multiple advantages. The surgeon could get an overview comparable to open surgery, the 3D-reconstruction could be used for documentation before and after an intervention and the 3D models might support autonomous robotic surgery [1].

This paper determines the prerequisites for 3D-reconstruction in laparoscopy and several optical depth estimation techniques with a strong intention to support researchers in improving laparoscopic surgery.

1.1 Depth estimation techniques used in research projects

A 3D point cloud containing the depth information of an object can be used to generate a 3D-reconstruction. A solution to create 3D points is using optical depth estimation techniques which are based upon image processing. To get the complete depth information, different camera perspectives are essential which is achieved with the help of camera motion. 3D-reconstruction of an object, which is also known as a 3D-model, is complete and can be visualized through packing each 3D point with the RGB values of the original images (Fig. 1).

Optical depth estimation techniques are also used in the smartphone industry for e.g. in face recognition [2] and autonomous driving in vehicle industry [3]. Compared to these fields the requirements to be fulfilled in minimally invasive surgery are a lot more sophisticated. This explains why image processing techniques used in smartphones or vehicles are not applicable in laparoscopes.

The following section presents optical depth estimation techniques used in research projects which emphasize the usability of laparoscopic surgery.

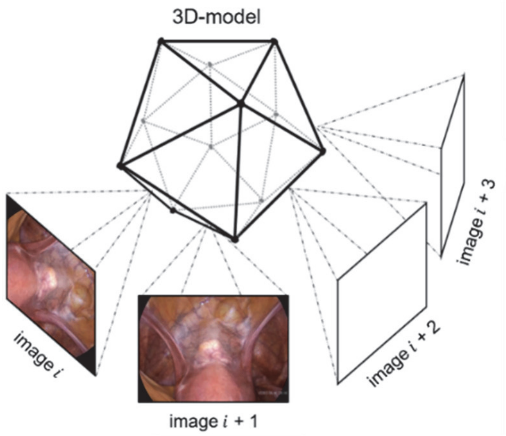

Figure 1: Basic visual of the generation of a 3D-model through image processing.

Stereoscopy is a depth estimation technique which is based on two cameras with a fixed baseline where the depth is calculated by triangulation with a common focal point [4].

Shape-from-X (SfX) is the overall description for algorithms which process mono camera images. One is called Shape-from-Motion (SfM) which uses camera motion to triangulate and determine depth [5]. Deformable Shape-from-Motion (DSfM) is similar to SfM but additionally uses templates to improve the reconstruction of surfaces especially if non-rigid [6]. Shape-from-Shading (SfS) uses pixel intensity to calculate depth [7].

Simultaneous Localization and Mapping (SLAM) is a framework for mono- and stereo camera system [8]. Structured Light (SL), for e.g. a projector, is used in parallel to a camera system and creates contrasts which are necessary to find corresponding points in image pairs [9]. Time-of-Flight (ToF) cameras measure the phase shift from the emitted light signal to the reflected signal which can be used to calculate object depth.

By using the Light-Field (LF) technique images are taken which capture light arrays from all directions separately due to an array of micro lenses between object and image sensor [10]. In one shot slightly different images arise. Comparable to stereoscopy, depth information can be calculated by triangulation. [11] uses LF to diagnose glaucoma.

1.2 Requirements to implement 3D-reconstruction in laparoscopy

Since the 1980s the topic of optical depth estimation has been captivating attention of medical researchers, especially in minimally invasive surgery which is exceedingly demanding. All of the following requirements should be considered when choosing a depth estimation method.

Requirement 1: The depth accuracy required in laparoscopic applications should be < 1 mm [12]. In laparoscopy the average object distance is about 10 – 200 mm. In comparison to smartphone or vehicle applications, the distances are 10 – 1000 times smaller.

Requirement 2: The target is a robust and dense 3D point cloud. For this a clear image quality and algorithms considering artefacts are needed. The conditions in the human body lead to some image quality issues which make it harder to precisely determine depth. During laparoscopic operations different issues can be observed for example [13]:

- tissue movements
- differences in the organ shape and surface texture
- under- and over-exposure because of illumination changes and organ topology
- blur when unsteady hand motion
- organ specularity and the presence of fluids
- smoke and bubbles
- light reflections
- no fixed landmarks
- no precise camera positioning

These issues have an influence on the corresponding point detection which is relevant for optical depth estimation [14]. Each image pixel works as an input for the algorithm and therefore should represent the true RGB value of an object. False pixel values can result in false 3D points and consequently false depth information. In some cases, the image quality is too bad to calculate a 3D point which can result in sparse 3D point cloud [15].

Requirement 3 + 4: Another significant requirement is real time estimation and a high resolution. A framerate of 20 fps implies that the sum of data acquisition and processing takes 50 ms per frame. Real time estimation requires 20 – 30 fps and for this most optical depth estimation techniques need GPU usage [14]. A high resolution about 1920 x 1080 Pixel (Full-HD) or even 3840 x 2160 Pixel (4K or UHD) should be feasible when using image sensors like OV5670 [16].

Requirement 5: Not only the software should fulfill certain requirements but also the hardware. Ideally the hardware set-up should not necessitate additional components compared to a standard system which is already launched into market. Additionally the endtip diameter should be as small as possible. This means for the hardware set-up only very small sized components should be used and the less the better. The market also claims a hardware set-up which is invulnerable to routine cleaning sessions.

2 Methods

This chapter explains how scores are retrieved to value each optical depth estimation technique regarding laparoscopy. Scoring is performed with reference to VDI 2221.

The score represents a usability measure of the shown methods in laparoscopic surgery based on a weighted fulfilling of the requirements from section 1.2. The higher the score the more promising the method. The table confirms that currently there is no ideal technique to be used in laparoscopy [14].
The score depends on how methods match the former declared requirements. The maximum score is 5 when all requirements are fulfilled. The different requirements were subjectively weighted as follows:

– Requirement 5: weight factor 1.5
– Requirement 2 + 4: weight factor 1.25
– Requirement 1 + 3: weight factor 0.5

Requirement 5: From a manufacturers perspective the hardware set-up is the most relevant demand because in minimally invasive surgery the puncture site should be as small as possible and laparoscopes should not become more complex than crucially required.

Requirement 2+4: Dense 3D point clouds and high-resolution image sensors lead to a 3D-model with a Full-HD or UHD resolution. These requirements are weighted as second most important.

Requirement 1+3: High precision and real time processing are less prioritized because these abilities count more for future applications like autonomous robotic surgery. As a first step the 3D-model can assist a surgeon in documenting the intervention. This application does not necessarily require the highest precision and real-time.

3 Results

Table 1 provides an overview of the optical depth estimation techniques, their advantages, their disadvantages and their score (range 1-5).

ToF cameras are the least useable as they require additional components which lead to increasing costs and an increasing complexity of the hardware set-up. Additionally, they suffer from low resolution chips and the predetermined chip position which means that chip-on-the tip cannot be realized. Although real time processing is possible and a dense 3D point cloud can be calculated, Requirements 4 + 5 do not hold. Thus, in total ToF cameras are assigned a score of 2.25.

SL requires additional components and it is built up next to a camera with a fixed baseline which can limit the range of object distances. E.g., when the baseline is very small, long distances cannot be captured. Although real time processing is possible and a dense 3D point cloud can be calculated, SL cannot fulfill Requirement 5 which leads to a score of 3.5. Due to the fixed baseline the final score is 3.

For LF additional parts are necessary (no Requirement 5). On the other hand real time processing as well as a dense 3D point cloud are possible, enough contrasts assumed [11], [17]. The score results in 3.5. There is no experience with LF in laparoscopy, leading to the final score of 3.

Stereoscopy, DSfM, SfM, SfS and SLAM do not need additional components and except DSfM, all of them can process in real time. DSfM gets a score of 4 (no Requirement 3). Stereoscopy and SLAM are limited by their fixed baseline and in some cases the 3D point cloud can be sparse. The score is 3.25.

SfM is valued with 3.25 because it is not robust for non-rigid objects (no Requirement 1 + 2). SfS leads in general to a sparse 3D point cloud and additionally the number of research projects with successful results is low. The score results in 2.

Table 1: Overview of optical depth estimation techniques, their advantages, disadvantages and score as an indication about the usability in laparoscopy [14].

Method	Advantage	Disadvantage	Score
Stereo	+ no additional components + real-time on GPU	- requires baseline - not always dense 3D point cloud	●●●○○
DSfM	+ no additional components + ideal for non-rigid objects	- no real time - not always dense 3D point cloud	●●●●○
SfM	+ no additional components + real time on GPU	- not robust for non-rigid objects	●●●○○
SfS	+ no additional components + real time	- sparse 3D point cloud	●●○○○
SLAM	+ no additional components + real time	- requires baseline - not always dense 3D point cloud	●●●○○
SL	+ dense 3D point cloud + real time	- additional components - requires baseline	●●●○○
ToF	+ dense 3D point cloud + real time	- additional components - low resolution - no chip-on-the tip	●●○○○

Method	Advantage	Disadvantage	Score
LF	+ dense 3D point cloud + real time	- additional components (array of micro lenses needed)	●●●●●

4 Discussion

This paper has presented eight optical depth estimation techniques and has invented an evaluation framework especially for laparoscopic applications. The framework suggests calculating a score (range 1-5) to value each depth measuring method. The score is grounded on five laparoscopic requirements which are weighted. The higher the score the better. The score bases on the author´s perspective and cannot be guaranteed as sufficient. Especially the weight factors are defined subjectively and changing these would have a strong influence on the evaluation results of the depth measuring methods. Moreover, this paper does not claim completeness regarding requirements and methods.

5 Conclusion

In conclusion, there is no ideal optical depth estimation method to create a high-resolution 3D-reconstruction in laparoscopy. The maximum score of five cannot be achieved. Recommendable depth measuring methods are DSfM, SfM, Stereoscopy, SL, LF and SLAM which have scores between three and four. As a future prospect the combination of different methods, high-performance computers and dye injections to increase contrasts for a better image processing will be considered.

Author Statement

Research funding: The authors state no funding involved. Conflict of interest: Authors state no conflict of interest. Informed consent: Informed consent has been obtained from all individuals included in this study. Ethical approval: The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration and has been approved by the authors' institutional review board or equivalent committee.

References

[1] Collins T, Bartoli A. 3D Reconstruction in Laparoscopy with Close-Range Photometric Stereo. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) 2012:634-642.

[2] Masiero A, Fissore F, Piragnolo M, Guarnieri A, Pirotti F, Vettore A. Initial Evaluation of 3D Reconstruction of close objects with smartphone stereo vision. International Society for Photogrammetry and Remote Sensing (ISPRS Technical Commission I) 2018:289-293.

[3] Iancu DT, Nan M, Ghiță AS, Florea AM. Vehicle depth estimation for autonomous driving. UPB Scientific Bulletin, Series C: Electrical Engineering and Computer Science. 2021:3-20.

[4] Wang C, Cheikh FA, Kaaniche M, Elle OJ. Liver surface reconstruction for image guided surgery. SPIE Medical Imaging 2018:105762H.

[5] Widya AR, Monno Y, Imahori K, Okutomi M, Suzuki S, Gotoda T, et al. 3D Reconstruction of Whole Stomach from Endoscope Video Using Structure-from-Motion. arXiv:1905.12988, 2019.

[6] Perriollat M, Hartley R, Bartoli A. Monocular Template-based Reconstruction of Inextensible Surfaces. Computer Vision-ACCV 2010:52-66.

[7] Collins T, Bartoli A. Towards Live Monocular 3D Laparoscopy using Shading and Specularity Information. Information Processing in Computer-Assisted Interventions 2012:11-21.

[8] Chen L, Tang W, John NW, Wan TR, Zhang JJ. SLAM-based dense surface reconstruction in monocular Minimally Invasive Surgery and its application to Augmented Reality. Computer Methods and Programs in Biomedicine 2018;158:135-146.

[9] Saeidi H, Opfermann JD, Kam M, Wei S, Leonard S, Hsieh MH, et al. Autonomous robotic laparoscopic surgery for intestinal anastomosis. Science Robotics 2022:1-13.

[10] Salem A, Ibrahem H, Kang HS. Light Field Reconstruction Using Residual Networks on Raw Images. Sensors 2022;22:1956.

[11] Schramm S, Dietzel A, Link D, Blum MC, Klee S. 3D retinal imaging and measurement using light field technology. Journal of Biomedical Optics 2021: 126002.

[12] Maier-Hein L, Groch A, Bartoli A, Bodenstedt S, Boissonnat G, Chang PL, et al. Comparative Validation of Single-Shot Optical Techniques for Laparoscopic 3-D Surface Reconstruction. IEEE Transactions on Medical Imaging. 2014:1913-1930.

[13] Ali S, Zhou F, Braden B, Bailey A, Yang S, Cheng G, et al. An objective comparison of detection and segmentation algorithms for artefacts in clinical endoscopy. Scientific Report 2020;10:2748.

[14] Maier-Hein L, Mountney P, Bartoli A, Elhawary H, Elson D, Groch A, et al. Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. Medical Image Analysis 2013:974-996.

[15] Allan M, Mcleod J, Wang C, Rosenthal JC, Hu Z, Gard N, et al. Stereo Correspondence and Reconstruction of Endoscopic Data Challenge. arXiv:2101.01133, 2021.

[16] OV5670. OMNIVISION.25.03.2022: https://www.ovt.com/sensor/ov5670/.

[17] Zhou S, Zhu T, Shi K, Li Y, Zheng W, Yong J. Review of light field technologies. Visual Computing for Industry, Biomedicine, and Art 2021:4:29.