Ralf Hackner*, Thomas Eixelberger, and Thomas Wittenberg

Panorama mapping of the colon – a phantom study with robotic endoscopy

https://doi.org/10.1515/cdbme-2022-0029

Abstract: When documenting colonoscopic examinations, the challenge arises, that the exact location of the detected lesions (such as polyps) can usually only be determined with low precision. In order to improve the image-based documentation of such lesions with respect to their location and context, a system is currently investigated and developed, allowing the creation of a panoramic map of the interior of the colon's lumen during an endoscopic examination. To evaluate the proposed algorithms, a 3D-printed colon phantom with known geometry has been developed. Furthermore, to obtain image sequences with precise and known motion during a simulated examination for test purposes, a robot was used to perform the insertion and withdrawal of the endoscope with constant speed and a steady viewing angle. The thus obtained image data is used to generate an ideal panoramic map using our the proposed stitching algorithm. Hence, it is possible to obtain reference maps of the phantom, which can be used to evaluate the quality of the algorithm.

Keywords: Endoscopy, additive manufactured phantom, robot guided endoscopy, panorama mapping

1 Motivation

During diagnostic and interventional colonoscopy, the limited field of view of the image sensor into the colon's lumen remains a challenge. Hence, during the image-based documentation of the examination, interesting findings such as lesions (as e.g., adenomas or polyps) or landmarks (the cecum or flexures) are captured, but their anatomical context is usually missing in the pictures. Event though, the insertion depth of the colonoscope into the patient's colon – approximated from white ring markings on the endoscope sheath – are noted as reference for the depicted lesion, the context information of the lesion remains vague. Hence, to support gastroenterologists in their documentation of detected lesions, image panoramas or image maps can potentially be used to display their context in the surrounding anatomy. In the field of cystoscopy, systems and frameworks are already known [1, 2], which allow the computation an image map of the urinary bladder texture in real-time from monocular endoscopic video data. Besides mapping a lumen's internal texture, also a real-time visual feedback to the examiner can be achieved, proving information if all parts of an organ's tissue have been completely examined.

In contrast to the urinary bladder, which can be coarsely modeled as sphere, the anatomy of the colon is much more complex and variable. To achieve such a mapping of a dynamic colon from a monocular video endoscopy data stream, a set of complex image processing modules have combined, see Section 3.3. Nevertheless, on order to evaluate the results achievable by such an approach, studies on phantoms with known geometry and under controlled endoscopic movements are necessary. To examine the lumen of an additively manufactured colon phantom [3] (Sec. 3.1), in our previous work a flexible video-colonoscope was used, where the endoscope's tip was moving freely and partially uncontrolled during the manual withdrawal through the phantom's lumen. In the current work a rigid robot-controlled endoscope (Sec. 3.2) was used for evaluation purposes, eliminating involuntary jitter movements of the image sensor, and hence yielding an optimal mapping of the colon phantom.

2 Related Work

To qualitatively and quantitatively evaluate a 3D-reconstruction of the colon (and the polyps within) from (monocular) colonoscopic video data, some type of ground truth is need. Since the human colon has a variability of approximately fifty percent for all six degrees of freedom during the process of colonoscopy, exact information about its real geometry and extension is never known. Thus, alternative methods such as phantoms must be used. Nevertheless, all known and commercially available phantoms for the colon (as e.g., from Simbionix Corp. [4]) are usually based on soft silicon tubes [5, 6], transparent plastic [7], or on animal cadavers [8–10] and hence trying to mimic the human colon with respect of the deformation possibilities, and partially lack the vascularization of the tissue wall. A recent overview of all known phantoms used for the training of interventional endoscopic procedures has recently been provided by Finocchiaro et al. [11].

Recently AI-based research has been carried out with respect

***Corresponding author: Ralf Hackner,** Fraunhofer IIS, Erlangen, Germany, e-mail: ralf.hackner@iis.fraunhofer.de
Thomas Eixelberger, Fraunhofer IIS, Erlangen, Germany
Thomas Wittenberg, Fraunhofer IIS and FAU, Erlangen, Germany

to automated reconstruction of the colon's lumen. As currently no stereo-colonoscopes are commercially available, from the image analysis point of view, the challenge remains, how to make a 3D-reconstruction of the colon's lumen and possible polyps within, based on only monocular image data. To this end some sparse research can be found. On one hand, AI-extended SLAM (simultaneously locating and mapping) approaches have recently been applied and investigated [12–16], trying to predict the movement of the endoscope tip within the lumen while simultaneously reconstructing the (strongly deformable) surface around. On the other hand, different deep neural network architectures have been trained (mainly on synthetic endoscopic data) to approximate the corresponding depth map from any given colonoscopy view [17–19].

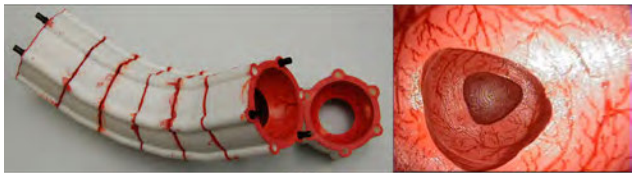

Fig. 1: Left:3D-printed modular colon phantom, image from [3]; right: endoscopic view into the phantom

3 Material and Methods

3.1 Additive-Manufactured Phantom

A modular, additive-manufactured, and stackable phantom of the colon was printed and manually colorized [3]. It has a known geometry and an inner luminal surface emulating the colon tissue, vascularization, and reflection similar to a real colon. All modules represent exactly one haustra, see Fig. 1, and can easily be combined in different order and rotation, thus yielding different flexures and curvature of a colon. One segment also includes polyps. As all modules are 3D-printed, a an exact geometric ground truth for the complete object is available. The precision of that ground truth is limited by the used layer thickness of $0.22mm$. The phantom used in this experiment has a length of about 28 cm with an inner diameter in the range of 3-5 cm. For this experiment only straight modules were used, in order to insert a rigid endoscope into the lumen.

3.2 Data Acquisition

To evaluate the proposed panoramic mapping approach of the colon (see Sec. 3.3), nine video sequences of the phantom (Sec. 3.1) and one sequence of the calibration chart, Fig. 2(H), were acquired. In contrast to our previous work [3] where two flexible video colonoscopes were used for image acquisition with a non-constant withdrawal-speed and a jittering endoscope tip, in the current experiments these strongly influential parameters were eliminated. Hence, instead a video colonoscope a rigid boroscope ($\oslash = 8.0$ mm, length = 30 mm, 0° frontal view, 67° viewing angle) (Karl Storz, Tuttlingen), Fig. 2(B), attached to an endoscope camera (Richard Wolf, Knittlingen), Fig. 2(C), yielded an unbiased direct view into the phantom's lumen, Fig. 2(A). To maintain a constant speed during the boroscope's insertion and withdrawal in an out of the phantom, a robot with seven axes (LBR iiwa, KuKa, Augsburg), Fig. 2(F) under manual control, Fig. 2(G) was applied. The endoscope's camera was attached to the robot's tip by a 3D-printed camera adapter, Fig. 2(E).

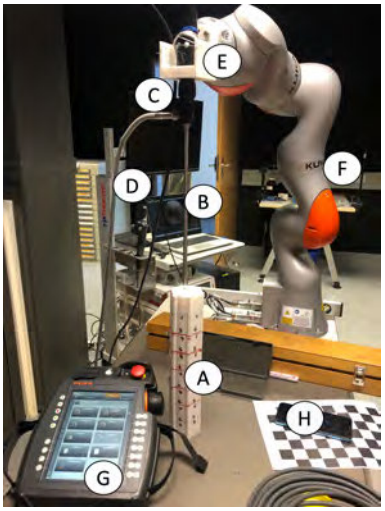

Fig. 2: Experimental setup: 3D-printed colon phantom (A); rigid endoscope (B); endoscope camera (C); endoscope tower with controller system, monitor, video grabber and laptop (D); additive manufactured camera adapter (E); 7-degree robot arm (F); manual robot control (G); calibration chart (H).

3.3 Image Processing

The workflow needed to transform an endoscopic (2D + t) image sequence of the colon into a corresponding 2D map consists of a set of image processing and transformation steps [3] depicted in Figure 3 and described below. The complete image processing chain runs in real-time with 19.5 frames per seconds (fps) and does not need any manual intervention.

3.3.1 Preprocessing

As not every frame of an endoscopic video stream has the quality and content be used as part of the colon panorama

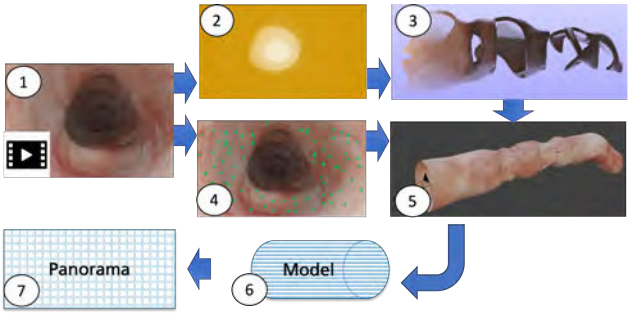

Fig. 3: Algorithm to obtain an image panorama (or map) of a colon lumen from a video sequence: colonoscopic video sequence (A), depth map per frame (B), longitudinal fused depth map (C), pairwise texture correspondences (D), 3D-approximation of the lumen (E), tube model (F), and final panorama (G).

map, a selection and prioritization of adequate frames must be made, while other frames have to be discarded. Hence, firstly, video frames with bad quality (e.g., blurred, or underexposed) are detected by a random forest based on traditional image features describing the image's structure (by the variance of Laplacian), color (32-bin HSV histogram), and edges (Canny edge detection). Secondly, video frames with foreign and irrelevant content (blood, stool, instruments, clips) are detected using a deep convolutional network, previously trained with over 7,000 frames from various public and private data collections [20]. Finally, to compensate the strong barrel distortion related to the endoscope's optics, previously acquired calibration information (see Fig. 2(H)) is used to undistort the remaining image frames. During the calibration process, the field-of-view of the sensor is obtained, needed later for the depth estimation.

3.3.2 Depth-map approximation

To estimate the 3D-geometry of the colon from the intensity and structures depicted a video frame $I(t)$ and its predecessor $I(t+1)$, a DCNN (a modified U-Net with a ResNet encoder) [17] is applied, trained on more than 16,000 image pairs. Each such image pair with known spatial relationship has been simulated from a digital colon model with real and artificial textures [21]. During the production state the DCNN receives a pair of endoscopic frames $(I(t), I(t+1))$ and returns an approximated depth map $D(t)$ and point cloud $P(t)$.

3.3.3 Point-cloud fusion

The hence computed point clouds are then concatenated and fused with each other in the 3D-space. To this end, image features are extracted using the SURF and AKAZE feature detectors. To compensate changes in illumination, an adaptive contrast correction is applied before. The extracted features are matched pairwise using a *fast library to approximate nearest neighbors* (FLANN). Based on the depth maps obtained from the DCNN, the 3D-coordinates of the features are computed. The final transformation for the registration step is estimated using a Levenberg-Marquardt approach. To reduce false registrations, the transformation is checked for plausibility.

3.3.4 Planar Mapping

The tubular object can be approximated by a sequence of cylinders. For the projection, the central axes of these cylinders are used. To obtain these cylinder axes, we determine centrally located reference point for each segment of the panorama and connect these points to a polygonal chain. Thus, a path is created which leads through the observed object. The reference point is currently defined as center of mass. A radial projection of the point cloud is then performed along these partial axes of the polygonal chain, yielding a planar map of the lumen.

4 Results

To evaluate the proposed panoramic mapping approach of the colon, a total of nine video sequences of the phantom (Sec. 3.1) and one sequence of the calibration chart (Fig. 2(H)), were acquired using the robotic endoscope setup described in Section 3.2. An example of an endoscopic image from the lumen of the phantom is depicted in Fig. 1 (right). All movements of the robot were made with constant speed and were manually controlled using a handheld controller. All image data was captured and stored and processed on a laptop-computer, while the image processing steps (Sec. 3.3) were computed on a standard desktop PC with a GPU (Nvidia RTX 2080 super). As result, textured point clouds of the colon phantom's lumen and the successive computed panoramic map were obtained. For image examples of the 3D-lumen and the corresponding map see Figures 4 and 5. The panorama depicts the mapping over three haustrae and includes two polyps.

5 Discussion

The resulting panoramic map, see Fig. 5, reflects the examined phantom in terms of its basic geometrical proportions and its surface texture. The two polyps placed in the model appear at the expected position in the panorama, Fig. 5(P). The unobserved (black) areas also appear at the expected positions, see

Fig. 4: 3D-approximation of the colon phantom's lumen.

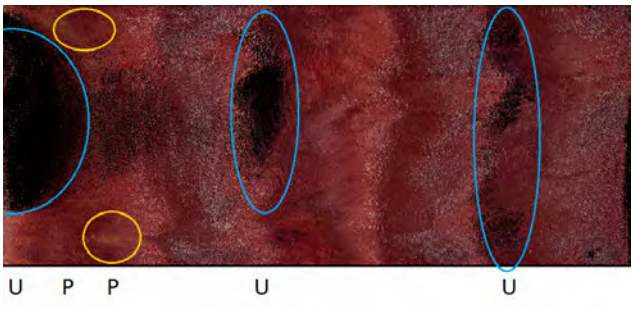

Fig. 5: Clipping from the extracted colon panorama: Left side: proximal end, right side: distal end of the phantom. P: polyps, U: unseen areas behind the haustrae

Fig. 5(U), but tend to be slightly smaller in the current projection than in reality. They are partially covered and occluded with material from visible areas that have been expanded in the projection process. Currently, the composed panorama map is still a back-projection from the fused point cloud, thus, the panorama image does not yet depict fine anatomical structures, such as vascularizations, and is still very grainy. As the goal of the research is a complete mapping of the colon's wall, radial errors in the 3D-reconstruction can be neglected.
However, the image resolution shall be improved in future versions of the back-projection module, by filling the gaps between the individual data points with image patches acquired from the original source images.

6 Conclusion

Even though the presented panorama reconstruction algorithms do not require the constant insertion and withdrawal motion related to the robotic guidance of the endoscope, the advantage of the constant robotic endoscope movement is, that the obtained image data can be used as a good reference to check and evaluate the quality of the proposed mapping algorithms for the colon.

Acknowledgment: This work was partially supported by the Bavarian Research Society for the project GastroMapper (AZ-1349-18).

Author Statement Authors state no conflict of interest.

References

[1] Bergen T (2017) Real-time endoscopic image stitching for cystoscopy. PhD Thesis, Univ. Koblenz-Landau.
[2] Hackner et al. Panoramic imaging assessment of different bladder phantoms – an evaluation study, Urology. 2021;156:e103–10.
[3] Hackner et al. 3D-reconstruction of the colon from monocular sequences – evaluation by 3D printed phantom data. Proc's Bildverarbeitung für die Medizin 2022, pp. 141-146.
[4] https://simbionix.com/simulators/gi-mentor/
[5] Reichel et al. Neuartiges Hands-on-Trainingsphantom für Proktologie & Koloskopie Z Gastroenterol 2016; 54 - KV096.
[6] Koch et al. Development and evaluation of an interventional training model for flexible endoscopy in Roux-En-Y anatomy Endoscopy 2021; 53(S 01): S159.
[7] King et al. Design and validation of a cost-effective physical endoscopic simulator for fundamentals of endoscopic surgery training. Surg Endosc 2016; 30, 4871–4879.
[8] Hochberger & Maiss. Currently available simulators: ex vivo models. Gastrointest Endosc Clin N Am. 2006 16(3):435-49.
[9] Chen et al. Training gastroenterology fellows to perform gastric polypectomy using a novel ex vivo model. World J Gastroenterol. 2011;17(41):4619-24.
[10] Ansell et al. (2013), The WIMAT colonoscopy suitcase model: a novel porcine polypectomy trainer. Colorectal Disease, 15: 217-223.
[11] Finocchiaro et al. Training simulators for gastrointestinal endoscopy: Current & future perspectives. Cancers 2021; 13(6): 1427.
[12] Grasa et al. Visual SLAM for handheld monocular endoscope. IEEE Trans. Medical Imaging. 2014, 33(1):135–46.
[13] Turan A non-rigid map fusion-based direct SLAM method for endoscopic capsule robots. Int. J. Intell Robotics & Applications. 2017;1(4):399–409.
[14] Ma et al. Real-time 3D reconstruction of colonoscopic surfaces for determining missing regions. MICCAI 2019, pp. 573-82.
[15] Chen et al. SLAM Endoscopy enhanced by adversarial depth prediction. ArXiv. 2019;abs/1907.00283.
[16] Xi et al. Endoscope localization and gastrointestinal feature map construction based on monocular SLAM technology. J. Infection & Public Health 2020;13(9):1314–21.
[17] Walluscheck et al. Partial 3D-reconstruction of the colon from monoscopic colonoscopy videos using shape-from-motion and deep learning. Curr. Dir. Biomed. Eng. 2021;7, 335-8.
[18] Liu et al. Self-supervised learning for dense depth estimation in monocular endoscopy, Proc's OR 2.0 Context-aware operating theaters, computer assisted robotic endoscopy, clinical image-based procedures & skin image analysis - pp. 128-38.
[19] Mahmood & Durr. Deep learning & conditional random fields-based depth estimation & topographical reconstruction from conventional endoscopy. Med Image Anal 2918; 48, 230-43.
[20] Kress et al. Automatic detection of foreign objects and contaminants in colonoscopic video data using deep leaning. Submitted to Curr. Dir. in Biomedical Engineering 8(2).
[21] Hackner et al. A geometric & textural model of the colon as ground truth for deep learning-based 3D-reconstruction. Proc's Bildverarbeitung für die Medizin 2021 - pp 298-303.