Camera Movement during Telementoring and Laparoscopic Surgery: Challenges and Innovative Solutions

Andrius Budrionis, Gunnar Hartvigsen, Johan Gustav Bellika

Norwegian Centre for Integrated Care and Telemedicine,
University Hospital of North Norway, Tromsø, Norway

Abstract

Advances in surgical telementoring systems open new grounds for enhancing the remote supervision and training of the surgeons. Verbal communication links between the mentor and mentee are being upgraded with graphical annotation, medical image and 3D model overlay capacities. While the advancing setups are regulated as verbal communication systems, the potential patient safety risks introduced by the extra features are rarely evaluated. In this paper surgical camera movement is researched in the context of the advanced telementoring systems. The analysis of potential computational techniques to eliminate its impact on the quality of remote supervision is presented. Focusing on the mobile platforms as common remote side telementoring hardware, an architecture for distributed telementoring system is proposed that eliminates the resources bottleneck for highly computationally intensive live video processing tasks.

Keywords:
telementoring, telestration, laparoscopic, endo-scopic, annotation, camera movement, tracking, distributed architecture

Introduction

Medical services, delivered via an interactive audio-video communication channel without direct physician-patient interaction (telemedicine) have become a common research area [1]. The advances in Information Communication Technologies (ICT) provide a solid basis for supporting a variety of telemedical tools. The forecasted shortage of surgical workforce is an important motivator for the rapid spread of the online practices. Given the constantly growing population, the increasing demand of surgical workload and the time required to train a surgeon, we are facing a rising demand of surgical personnel with a limited supply [2]. The ongoing development of new surgical techniques also requires tools for disseminating knowledge and providing a medium for the practical training of surgeons. Telementoring is an important tool to facilitate collaboration among surgeons and provide better quality of healthcare service to the patient [3].

The paper explores remote supervision of novice surgeons via video conferencing (VC), while the expert is not available onsite (telementoring) [1]. VC in combination with live free-hand annotations of the video from laparoscopic camera (telestration) offers valuable benefits concerning improved clinical and educational outcomes, increased speed of procedures and reduced relocation costs of the expert [4]. Telementoring is a platform for surgical education, where the remote expert acts as a personal tutor for the surgeon performing the operation. ICT provide a medium for widening the scope of education from a typical mentor-mentee interaction onsite to an n-way mentoring-education process (multiple mentors, additional participants following the session for interactive training) [5]. The overall benefits are not limited to the potentially improved patient outcomes, but they also contribute to increasing the availability of the experts and improving surgical education by using live content based on actual cases instead of imitations.

Telementoring for minimally invasive (laparoscopic) procedures has attracted major focus lately. One of the main characteristics of laparoscopic surgeries is the importance of the visual component: the progress of the procedure in the OR is presented on a monitor in front of the surgeon, sharing this information with a remote mentor without losing much is highly feasible.

Camera movement is an important factor, potentially leading to the decreased accuracy in the actions of the surgeon. Moreover, if the telestration feature is introduced, it might result in even more challenges. Graphical guidance (annotations) are produced over a particular location, for example, a mentor draws a line marking the location of the incision. In the case of camera movement, the line will shift to a different location and result in inaccurate advice, which may then lead to critical failures. After analyzing a set of laparoscopic videos (N = 10 laparoscopic sigmoidectomy procedures), we came to the conclusion that camera movement is an integral part of the procedure that cannot be omitted [6]. The movement can be minimized when dealing with less complicated cases or when the assisting surgeon is highly experienced. However, even with the simplest case and the most experienced surgeon, it is not possible to avoid camera movement while telestrating on a live video stream. Moreover, the moving camera is not the only factor resulting in the repositioning of annotations. The movement and poor color distinction of intra-abdominal soft tissues also contributes to the complexity of the situation (Figure 1).

The use of telestration while mentoring is the first step towards surgical navigation systems: verbal interaction between the mentor and mentee becomes guidance by graphical artifacts, produced over laparoscopic video. It introduces additional risks for the patient safety. The intended influence on the flow of the surgery makes such systems medical devices, with regulations for development and application [7]. However, in practice they are often treated as communication modules, denying the potential patient safety threats to avoid the additional complexity in complying with the regulations.

The paper looks into the specifics of surgical telementoring. Although the actual impact of telestration is unknown, the earlier mentioned assumptions call for solution anchoring the annotations to their initial locations. A review of the potential solutions is presented in later sections followed by an architecture for a distributed telementoring system able to handle resource intensive computations on relatively low-powered hardware.

Materials and Methods

To maintain constant position of annotations, they need to be anchored to continuously identifiable features in the video. Two potential approaches for solving the camera movement problem (CMP) were identified in the literature: anatomical landmark tracking and 3D modeling for laparoscopic surgery. Additional method to avoid the problem - combination of laparoscopic video and still images for telestration was presented. A systematic literature search to support the approaches was performed in the PubMed, the Association of Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE) online research databases using a predefined set of keywords (vessel tracking, vessel identification, vessel detection, tissue tracking, track tissue deformations, 3d laparoscopic, 3d mentoring, 3d laparoscopic models) on the 22nd of November, 2014. Publication period between 01/01/2000 and 22/11/2014 was covered. Eight papers were included in the full text analysis.

The results formed a sufficient body of literature to support the initial claims. Challenges to adopt the techniques for the use on mobile devices were reported. To identify the feasible methods for solving the CMP in a mobile medium, guidelines to overcome the challenges were defined and summarized in a proposal for a distributed telementoring system architecture.

Results: Can Computational Techniques Solve the CMP?

In the previous section we mentioned the external factors (level of complexity, experience of a surgeon) that can improve the accuracy of mentoring advices presented in the form of annotations. However, the camera movement problem remains. Therefore, we suggest a set of computational solutions. To simplify the description, the area observed by the surgeon in the screen is defined and referred to as the operative field.

Combining Video and Still Images

The operative field is in continuous motion. Regardless the instable human-held camera, the surrounding environment (soft tissues) is also moving, either responding to the changing pressure inside the cavity or to the tools controlled by the surgeon (Figure 1). These factors prevent having stable landmarks to ensure the accurate positioning of telestrations. From the mentor's perspective, it may be difficult to provide an accurate graphical advice (annotation) under such conditions. To ensure the necessary accuracy, the best solution may be to pause for a moment, stabilize the movement of the internal tissues as much as possible (for example, keeping the pressure constant) and observe the scene. This representation makes it easier for the mentor to analyze the situation, as it becomes less ambiguous due to the minimized influence of changes. The stability of the operative field enables accurate telestration as the reference points maintain a constant position. If we look at this situation from another perspective, it is a shift from a video stream to a still image. Due to the minimized influence of the motion, the laparoscopic video, observed on a monitor by the local surgeon and the remote mentor, becomes a still image for the purpose of accurate telestration.

The schematic view of the proposed solution is depicted in Figure 1. The main idea is to add extra functionality to the video conferencing software that enables the video to be stopped at any moment, converts the frame to a still image that supports a discussion between the mentor and mentee (including the telestration capacity if it is necessary). After the discussion is completed by a corporate decision, the procedure, as well as the video-based mentoring session, is resumed until another critical moment, requiring a discussion to find the best solution, occurs.

Regardless the simplistic implementation, maintaining the accuracy in transition from the still image to the dynamic representation of the operative field may be challenging for the surgeon. An overlay of the scene used in the discussion over the actual video could offer an improvement, assuming no major changes occurred in the operating field after the screenshot was taken. However, to avoid the ambiguities between the representations used for mentoring and the surgery, employing the live video feed is discussed in later sections.

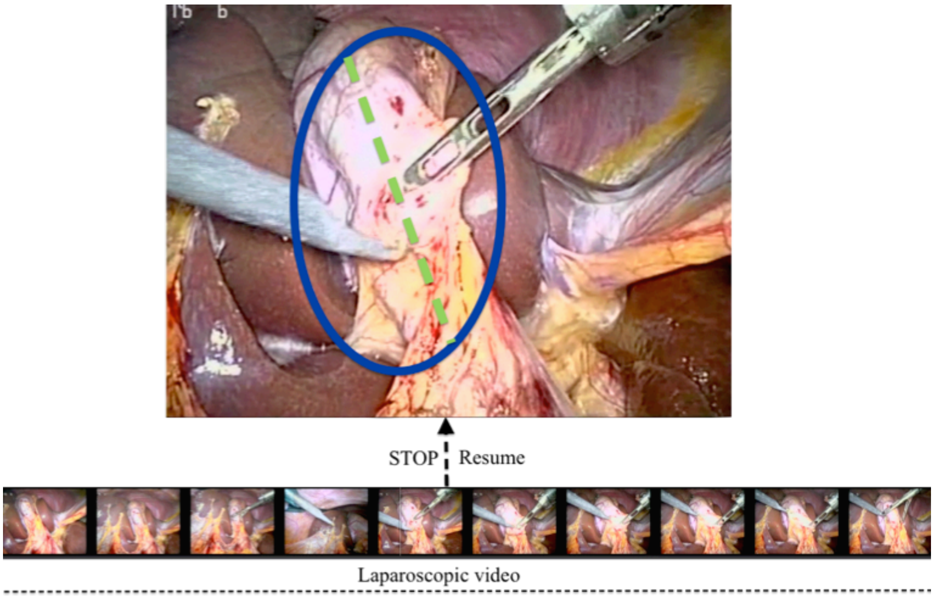

Figure 1- Combination of video and still images for telestration

Anatomical Landmark Tracking

To attach telestrations to anatomical landmarks, a set of stable reference points is required. The higher the number of references, the more reliable and robust the anchoring is. Research on blood vessel identification and tracking has revealed the growing potential of applications [8–10]. Computer vision techniques implement the idea of comparing the intensiveness of the pixels in the image with respect to the background. A blood vessel is identified when the result of the comparison exceeds a threshold value [9]. Although all the attempts have dealt with tracking blood vessels in still images, we propose extending this functionality to video content analysis. As the tracking of all full-length blood vessels (including forking and crossing) in the environment is not necessary in the telementoring case, the computations could be simplified and accelerated. The continuous detection of reference points attaches the annotations and stabilizes their position despite the camera movement. However, the approach is not representative due to the deformation of the soft tissues: the annotation may be stable, but reacting to the surface changes in the close vicinity is required (for example, inside the blue circle in Figure 1).

Tracking a high number of reference points on the surfaces appearing in the operative field extends the approach to the soft tissue deformation tacking. The core idea is to build an artificial representation of the deforming surfaces, observable in an operative field and represent their motion based on the tracked reference points. Attempts to use the captured reference points as anchors for annotations were also reported. Although the algorithms are computationally intensive, positive results in applying the techniques to process live video streams were presented [11,12].

The solution encounters a number of challenges to be dealt with. In some cases, the problem of identifying the landmarks becomes complicated due to limited visibility or reappearing artifacts (for instance, surgical tools). The homogenous structures in the operating field influence the detection process. For example, the more fat the tissue contains, the more complicated it is to find a blood vessel or another distinct object. Moreover, big camera movements call for reidentifying the same points, which were out of the operative field for a while. Fiducial markers could be of help as suggested by Mountney et al. to create a traceable landmark [13].

3D Models for Laparoscopic Surgery

Research dealing with the 3D modeling of the operative field encompasses the achievements of the tracking techniques. Su et al. presented an approach of combining Computed Tomography (CT) or magnetic resonance (MR) images and laparoscopic video. A markerless tracing system for real-time visualizations was investigated, enabling the modeling of the surfaces of the tissues in the operative field. The internal structure of a particular organ, based on CT or MR images, was designed and overlaid on a laparoscopic video, creating a three-dimensional representation of the operative field. The consolidation of two modalities (CT/MR images and live video) enables an augmented perception of the operative field. It allows internal and external observation before performing the incisions [14].

The described approach lays a fertile ground for solving the CMP. The CT/MR based model is a perfect artifact for attaching the annotations. Moreover, the internal model is aligned according to the reference points tracked on the surface of the tissue [14], making it responsive to the deformations.

The objective of this paper is to analyze the solutions to the CMP and adapt them for use in mobile environments. Although 3D modeling looks like the most promising approach, it has high demand for computational and representational power. The application of the discussed methods to mobile platforms would be feasible only after the complexity of computations is decreased or distributed among the components of the system. Therefore, we propose an architecture for a distributed telementoring system, capable of providing the required computational power.

A Distributed Telementoring System

The main drawback of tissue tracking and modeling approaches is the high demand for computational resources. Mobile deployment platforms increase this limitation and prevent straightforward development. However, the distribution of the computations among the components of the system looks promising. This section presents the architecture for a distributed telementoring system, aligned with the achievements of Zhou and Liu in developing an n-way visual content sharing

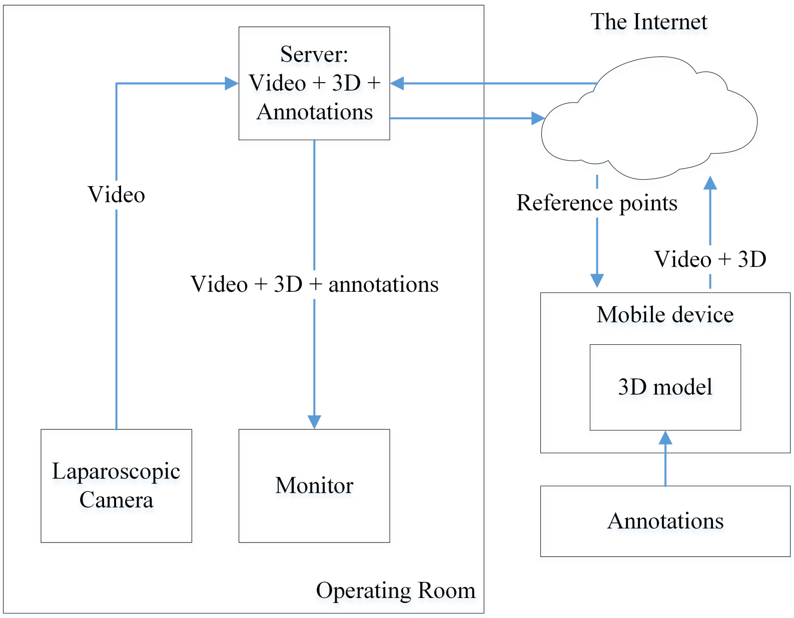

Figure 2- Architecture of a distributed telementoring system

and telestration technology, storing video streams and annotations separately and overlaying them when necessary [15]. The architectural design we present, models the approach discussed in section "3D Models for Laparoscopic Surgery". However, it could be easily adopted for the design and implementation using any of the approaches discussed earlier.

The system is divided into two parts (Figure 2): local (operating theatre) and remote (mentor). The local side is stationary and responsible for providing sufficient computational power for data processing. The remote side is mobile and acts as the representational medium to capture the reference points for annotations made by the mentor. The main procedures performed by the server on the stationary (local) side are:

1. Maintaining live audio-video conferencing between the operating theatre and the remote expert;
2. Overlaying 3D models of the tissue/organ on the laparoscopic video and aligning them according to the deforming surfaces observed in the laparoscopic video;
3. Recreating the annotations of the mentor received from the remote side in the form of reference points, combining the video, the 3D model and annotations;
4. Controlling the model on the mentor side, making it react to the tissue deformations.

The 3D model is developed before the procedure from medical imaging. It displays the internal anatomical structure and also functions as a coordinate system for telestration. The same model is used on both (local and remote) sides, to determine the accurate position of the annotations defined in reference points. Attaching annotations to the model eliminates the CMP, as the model maintains a constant position regardless of the instability of the operative field.

The main challenge is to accurately reflect the deformations of the operative field in the 3D model. It is a complex task requiring the computational resources delivered by the stationary platforms. As the system employs two instances of the 3D model, the mentor side is left with the representation of the results of the model deformation computed by the server side. Bandwidth consumption is minimized by only transferring the deformations of the model between the nodes and performing rendering locally.

Comparison to the alternative architectures

To highlight the advantages of the proposed architecture a comparison to other architectural approaches is provided bellow.

3D model on the client side

Utilizing a 3D model at the origin of annotations (the remote mentor side) seems like a straightforward solution anchoring the telestrations right after they were produced. Translation from 3D to 2D annotations would generate transferable artifacts for displaying them for the mentee in the OR. The same 2D representation can be displayed for the mentor, making sure both sides of the link follow exactly the same scene. Despite the technically feasible anchoring of telestrations, making the 3D model react to the soft tissue deformations seem to be too computationally intensive for the mobile device of the mentor. A non-responsive model in this case is as useful as a static grid placed over the dynamic video.

3D model on the server side

Placing the model on the server side provides the required computational and bandwidth resources. However, the annotations originate on the mentor side in 2D form, transforming them into 3D on the OR may not accurately communicate the message of the mentor. Moreover, mentor and mentee sides would be using different representations of the operative field, which may increase the ambiguities between the peers and result in patient safety threats.

3D model on both ends

Placing the 3D models on both ends was suggested in the previous section and visualized in Figure 2. Regardless a more complex distributed architecture, the proposal ensures the same representation of the operative field on both ends of the link and anchors the annotations eliminating their repositioning. Distribution of the computationally intensive tasks provides sufficient resources for enhancing telementoring experiences for the mobile users.

Discussion

The discussed advances of the communication system within the OR bring telementoring closer to the surgical navigation systems. Such solutions are common in the procedures, requiring extremely high precision, however often performed in less dynamic tissues (for instance, brain surgery). Mapping the surgical tools into the 3D model of the internal body structure enables the augmented perception and coordination of the moves based on the artifacts invisible for the naked eye. The overall representation (a responsive 3D model and tracked tools) provides a machine-readable documentation of the procedure. This vision not only introduces additional safety assurance procedures by adding the ability to warn the performing surgeon about the potential errors (cut of the blood vessel hiding inside the tissue), it also questions the utility of the human mentor.

Is it likely that an artificial intelligence system, trained on a high number of surgery cases would be able to provide a satisfactory advice for the surgeon? Does it make the machine "a trustworthy partner" in providing suggestions and making decisions, while the experience of the performing surgeon is exploited as a tool to complete physical actions?

Not only does it sound scary, but it also raises numerous ethical considerations. However, we cannot deny the efficiency of the artificial intelligence techniques in completing highly specific tasks, cost reduction and voiding of errors caused by the human factor. Regardless the futuristic scenario, the required technologies to support the workflow are available. The discussed ideas give an absolutely different meaning to telementoring, together with numerous computational challenges to cope with.

Conclusions

The paper looked into the attempts to integrate the computational image processing techniques into surgical telementoring to compensate for the instability of the laparoscopic camera. To the current date, the CMP has not been in a focus of the researchers, however, the importance of a reliable solution is fundamental in developing a comprehensive telementoring system, meeting the international regulatory requirements.

The reviewed publications point in one direction – 3D modeling of the operative field offers the biggest enhancements to telementoring, nevertheless it is associated with numerous challenges. Image guidance capabilities promise exceeding the benefits of onsite mentoring, making telementoring a preferred approach [16]. Such changes could have ground shaking consequences in the stagnated adoption of telementoring in surgery. The review and the proposed architecture emphasize the complexity of the solution to CMP and suggest the direction for further research and development.

Acknowledgements

This research was funded by a Helse Nord grant (ID 5614/HST1025-11) to the Norwegian Centre of Integrated Care and Telemedicine (NST), University Hospital of North Norway (UNN). The study would not have been possible without the support of the Mobile Medical Mentor (M3) project team at NST.
The authors want to thank prof. Rolv-Ole Lindsetmo at the Department of Gastrointestinal Surgery (UNN) for sharing his expertise on the topics discussed in the paper.

Address for correspondence

Andrius Budrionis, Norwegian Centre for Integrated Care and telemedicine, University Hospital of North Norway, Tromsø, Norway.
Andrius.Budrionis@telemed.no

References

[1] Society of American Gastrointestinal and Endoscopic Surgeons. Guidelines for the Surgical Practice of Telemedicine. 2010.

[2] Williams TE, Ellison EC. Population analysis predicts a future critical shortage of general surgeons. Surgery 2008;144:548–56. doi:10.1016/j.surg.2008.05.019.

[3] Augestad KM, Bellika JG, Budrionis A, Chomutare T, Lindsetmo R-O, Patel H, et al. Surgical telementoring in knowledge translation--clinical outcomes and educational benefits: a comprehensive review. Surg Innov 2013;20:273–81. doi:10.1177/1553350612465793.

[4] Budrionis A, Augestad KM, Patel HR, Bellika JG. An Evaluation Framework for Defining the Contributions of Telestration in Surgical Telementoring. Interact J Med Res 2013;2:e14. doi:10.2196/ijmr.2611.

[5] Budrionis A, Augestad K, Patel H, Bellika J. Towards Requirements for Telementoring Software. Scand Conf Health Inform 2012 2012.

[6] Budrionis A, Augestad KM, Bellika JG. Telestration in Mobile Telementoring. ETelemed 2013 Fifth Int. Conf. EHealth Telemed. Soc. Med., 2013, p. 307–9.

[7] European Commission. Guidelines on the Qualification and Classification of Stand Alone Software Used in Healthcare Within the Regulatory Framework of Medical Devices. 2012.

[8] Yin Y, Adel M, Bourennane S. Automatic Segmentation and Measurement of Vasculature in Retinal Fundus Images Using Probabilistic Formulation. Comput Math Methods Med 2013;2013:1–16. doi:10.1155/2013/260410.

[9] Jiang Y, Bainbridge-Smith A, Morris AB. Blood vessel tracking in retinal images. Proc. Image Vis. Comput., 2007.

[10] Kirbas C, Quek F. A review of vessel extraction techniques and algorithms. ACM Comput Surv 2004;36:81–121. doi:10.1145/1031120.1031121.

[11] Visentini-Scarzanella M, Merrifield R, Stoyanov D, Yang G-Z. Tracking of irregular graphical structures for tissue deformation recovery in minimally invasive surgery. Med Image Comput Comput-Assist Interv MICCAI Int Conf Med Image Comput Comput-Assist Interv 2010;13:261–8.

[12] Yip MC, Lowe DG, Salcudean SE, Rohling RN, Nguan CY. Real-Time Methods for Long-Term Tissue Feature Tracking in Endoscopic Scenes. In: Abolmaesumi P, Joskowicz L, Navab N, Jannin P, editors. Inf. Process. Comput.-Assist. Interv., vol. 7330, Berlin, Heidelberg: Springer Berlin Heidelberg; 2012, p. 33–43.

[13] Mountney P, Yang G-Z. Context specific descriptors for tracking deforming tissue. Med Image Anal 2012;16:550–61. doi:10.1016/j.media.2011.02.010.

[14] Su L-M, Vagvolgyi BP, Agarwal R, Reiley CE, Taylor RH, Hager GD. Augmented Reality During Robot-assisted Laparoscopic Partial Nephrectomy: Toward Real-Time 3D-CT to Stereoscopic Video Registration. Urology 2009;73:896–900. doi:10.1016/j.urology.2008.11.040.

[15] Zhou Q, Liu D. Interactive visual content sharing and telestration: A novel network multimedia service, IEEE; 2010, p. 1–6. doi:10.1109/ICIN.2010.5640915.

[16] Hollan J, Stornetta S. Beyond Being There. Proc. SIGCHI Conf. Hum. Factors Comput. Syst., New York, NY, USA: ACM; 1992, p. 119–25. doi:10.1145/142750.142769.