Medical Multimedia Information Systems (MMIS)

Klaus Schoeffmann, Bernd Münzer
Institute of Information Technology / Lakeside Labs
Klagenfurt University, Austria
{ks,bernd}@itec.aau.at

Michael Riegler, Pål Halvorsen
Simula Research and University of Oslo
Oslo, Norway
{michael,paalh}@simula.no

ABSTRACT

In hospitals all around the world, medical multimedia information systems have gained high importance over the last few years. One of the reasons is that an increasing number of interventions are performed in a minimally invasive way. These endoscopic inspections and surgeries are performed with a tiny camera – the endoscope – which produces a video signal that is used to control the intervention. Apart from the viewing purpose, the video signal is also used for automatic content analysis during the intervention as well as for post-surgical usage, such as communicating operation techniques, planning future interventions, and medical forensics. Another reason is video documentation, which is even enforced by law in some countries. The problem, however, is the sheer amount of unstructured medical videos that are added to the multimedia archive on a daily basis. Without proper management and a multimedia information system, the medical videos cannot be used efficiently for post-surgical scenarios. It is therefore already foreseeable that medical multimedia information systems will gain even more attraction in the next few years. In this tutorial we will introduce the audience to this challenging new field, describe the domain-specific characteristics and challenges of medical multimedia data, introduce related use cases, and talk about existing works – contributed by the medical imaging and robotics community, but also already partly from the multimedia community – as well as the many open issues and challenges that bear high research potential.

ACM Reference format:
Klaus Schoeffmann, Bernd Münzer and Michael Riegler, Pål Halvorsen. 2017. Medical Multimedia Information Systems (MMIS). In *Proceedings of MM'17, October 23–27, 2017, Mountain View, CA, USA.*, , 2 pages.
DOI: https://doi.org/10.1145/3123266.3130142

1 MOTIVATION

Over the last two decades, in addition to the personal and entertainment domain, we have observed the adoption of digital video recording and automatic content analysis in many professional domains (e.g., documenting construction sites, documenting and analyzing sports, surveillance of buildings/streets/traffic, etc.). One of these professional domains is medical science, where digital videos are recorded with increasing frequency to create exhaustive documentations of surgeries and inspections. This is especially true in the case of endoscopic interventions, as the video signal is inherently available and no additional equipment is required. Endoscopes are little cameras that are inserted into the patient's body in order to inspect body regions or to supervise and control surgeries in a minimally invasive way. Since recordings of these videos show exactly the same scene that was used on site by the surgeons, the videos contain highly valuable information for many additional purposes: explanations to patients, discussions with colleagues, training of new surgeons, medical case retrieval, search for similar pathologies in an archive of many patients (i.e., forensics), planning future surgeries, and even analysis of technical errors and adverse events that might have happened during the surgery.

The rapidly increasing amount of data that is collected in this way and the growing interest of health professionals necessitates the conception and development of special multimedia systems that are adapted to this domain. In general, a medical multimedia system is an interactive system that provides support for diagnostics, examination, surgery, reporting and teaching in a medical setting by combining all available information sources and putting them in the hands of medical professionals or patients [16]. An important point to add to this definition is that some medical information systems may be fully automatic, but should still be considered interactive, since a medical professional and/or patient must be in the loop to provide input, interpret and act on the results. A complete system has to fulfill several requirements, which include: (i) Efficient processing capabilities, which means that - depending on the actual use case - the system should either be able to process data in real-time such that the output can be used by medical experts during live procedures [11, 12] or - in the case of post-operative analysis scenarios - the processing is fast enough to keep up with the plethora of steadily produced data, (ii) a pipeline for the complete workflow from transferring medical expert knowledge into the system using annotation tools via an automatic analysis module through to a component for presentation of the systems' output [1] and (iii) the possibility to adapt and extend the system regarding different requirements such as different diseases [17].

Content analysis of medical videos is a very challenging area, since these videos contain a very special content [19]. For example, the actual content area is often circular (the frame is surrounded by a black border) and contains a considerable portion of irrelevant content, such as blurry, dark and noisy frames, as well as frames recorded outside of the patient [6]. More importantly, such videos contain highly self-similar content (similar colors, similar texture, similar motion) due to highly similar anatomical structures as well as similar actions and operations. Therefore it is very challenging to find discriminative content features [18] and sufficiently large and diverse training sets when using machine learning techniques [8, 14]. For some kind of medical videos also the motion is so slow and subtle that it can only be detected in high-resolution images [13].

MM'17, October 23–27, 2017, Mountain View, CA, USA.
 ISBN 978-1-4503-4906-2/17/10.
DOI: https://doi.org/10.1145/3123266.3130142

Over the last few years many researchers have contributed to solve specific problems in the challenging field of medical multimedia, with a clear focus on endoscopic image and video processing, as described in detail in a recent comprehensive survey authored by the applicants [5].

While the medical imaging community has contributed several methods for real-time detection of abnormalities/pathologies in medical images [2] and the robotics community has investigated computer vision techniques for real-time assistance in terms of augmented reality and robot-assisted surgery [4], the multimedia community has contributed methods for post-processing of endoscopic videos to enable efficient compression and storage, retrieval, similarity search, summarization, temporal segmentation and advanced interaction [3, 7, 15]. However, many problems still remain unsolved while at the same time people from the medical domain (researchers, surgeons, staff) demand for more effective and efficient tools to exploit the potential of ever-growing medical multimedia archives.

In this tutorial we want to provide a broad overview of the medical multimedia field, introduce the related applications, talk about existing methods and solutions, and prospect future challenges and issues that could and should be addressed by the multimedia community. We hope to raise awareness of the huge potential and growing significance of research in this field and as a result stimulate further research activities from the multimedia community.

2 COURSE OVERVIEW

As mentioned above, various issues and challenges arise when dealing with medical multimedia content. In this tutorial, we will give an overview of existing methods to approach named challenges and issues. This will include a detailed description about the data and its challenges, content analysis methods proposed for real-time detection as well as post-processing for content structuring, classification, and retrieval. Furthermore, we will talk about special interfaces that are required when interacting with medical multimedia data. The participants will learn best practices on how to process medical multimedia content and should be able to apply this knowledge in their own project upon completion of the tutorial.

The tutorial will roughly cover the following topics:

- Introduction and OVerview
- State-of-the-Art in Endoscopic Video Processing [5]
- Medical Videos & Datasets [9, 10]
- Diagnostic Decision Support [16, 17]
- Medical Multimedia Applications [8, 18, 20]
- Conclusions and Future Challenges

Target audience are researchers and practitioners working in at least one of the following fields: video content analysis, visual information retrieval, and human-computer interaction with multimedia data. Participants should have a basic understanding of multimedia content analysis and should be somewhat familiar with machine learning methods. We do not expect any medical background. One of our aims is to enthuse participants with this new application domain for multimedia research and stimulate further research in this area.

REFERENCES

[1] Zeno Albisser, Michael Riegler, Pål Halvorsen, Jiang Zhou, Carsten Griwodz, Ilangko Balasingham, and Cathal Gurrin. Expert driven semi-supervised elucidation tool for medical endoscopic videos. ACM, 03/2015 2015.

[2] M. Liedlgruber and A. Uhl. Computer-Aided Decision Support Systems for Endoscopy in the Gastrointestinal Tract: A Review. *Biomedical Engineering, IEEE Reviews in*, 4:73–88, 2011. 00042.

[3] Mathias Lux, Oge Marques, Klaus Schöffmann, Laszlo Böszörmenyi, and Georg Lajtai. A novel tool for summarization of arthroscopic videos. *Multimedia Tools and Applications*, 46(2-3):521–544, September 2009. 00000.

[4] G. P. Moustris, S. C. Hiridis, K. M. Deliparaschos, and K. M. Konstantinidis. Evolution of autonomous and semi-autonomous robotic surgical systems: a review of the literature. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 7(4):375–392, December 2011. 00050.

[5] Bernd Münzer, Klaus Schoeffmann, and Laszlo Böszörmenyi. Content-based processing and analysis of endoscopic images and videos: A survey. *Multimedia Tools and Applications*, pages 1–40.

[6] Bernd Munzer, Klaus Schoeffmann, and Laszlo Boszormenyi. Relevance segmentation of laparoscopic videos. In *Multimedia (ISM), 2013 IEEE International Symposium on*, pages 84–91. IEEE, 2013.

[7] B. Münzer, K. Schoeffmann, and L. Böszörmenyi. Domain-Specific Video Compression for Long-Term Archiving of Endoscopic Surgery Videos. In *2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 312–317, June 2016. 00000.

[8] Stefan Petscharnig and Klaus Schöffmann. Learning laparoscopic video shot classification for gynecological surgery. *Multimedia Tools and Applications*, pages 1–19, 2017.

[9] Konstantin Pogorelov, Kristin Ranheim Randel, Thomas de Lange, Sigrun Losada Eskeland, Carsten Griwodz, Dag Johansen, Concetto Spampinato, Mario Taschwer, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. Nerthus: A bowel preparation quality video dataset. In *Proc. of MMSYS*, 2017.

[10] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In *Proc. of MMSYS*, 2017.

[11] Konstantin Pogorelov, Michael Riegler, Pål Halvorsen, Thomas de Lange, Peter Thelin Schmidt, Sigrun L. Eskeland, Carsten Griwodz, and Dag Johansen. Gpu-accelerated real-time gastrointestinal diseases detection. IEEE, 07/2016 2016.

[12] Konstantin Pogorelov, Michael Riegler, Jonas Markussen, Håkon Kvale Stensland, Pål Halvorsen, Carsten Griwodz, Sigrun L. Eskeland, and Thomas de Lange. Efficient processing of videos in a multi auditory environment using device lending of gpus. pages 36:1–36:4. ACM, 05/2016 2016.

[13] Manfred Jürgen Primus, Klaus Schoeffmann, and Laszlo Böszörmenyi. Segmentation of recorded endoscopic videos by detecting significant motion changes. In *Content-Based Multimedia Indexing (CBMI), 2013 11th International Workshop on*, pages 223–228. IEEE, 2013.

[14] Manfred Jürgen Primus, Klaus Schoeffmann, and Laszlo Böszörmenyi. Instrument classification in laparoscopic videos. In *Content-Based Multimedia Indexing (CBMI), 2015 13th International Workshop on*, pages 1–6. IEEE, 2015.

[15] Manfred Jürgen Primus, K. Schoeffmann, and L. Böszörmenyi. Temporal segmentation of laparoscopic videos into surgical phases. In *2016 14th International Workshop on Content-Based Multimedia Indexing (CBMI)*, pages 1–6, June 2016. 00000.

[16] Michael Riegler, Mathias Lux, Carsten Griwodz, Concetto Spampinato, Thomas de Lange, Sigrun L. Eskeland, Konstantin Pogorelov, Wallapak Tavanapong, Peter T. Schmidt, Cathal Gurrin, Dag Johansen, Håvard Johansen, and Pål Halvorsen. Multimedia and medicine: Teammates for better disease detection and survival. In *Proceedings of the 2016 ACM on Multimedia Conference*, MM '16, pages 968–977, New York, NY, USA, 2016. ACM.

[17] Michael Riegler, Konstantin Pogorelov, Sigrun L. Eskeland, Peter T. Schmidt, Zeno Albisser, Dag Johansen, Carsten Griwodz, På l Halvorsen, and Thomas de Lange. From annotation to computer aided diagnosis: Detailed evaluation of a medical multimedia system. *Transaction on Multimedia*, 2017.

[18] Klaus Schoeffmann, Christian Beecks, Mathias Lux, Merih Seran Uysal, and Thomas Seidl. Content-based retrieval in videos from laparoscopic surgery. In *SPIE Medical Imaging*, pages 97861V–97861V. International Society for Optics and Photonics, 2016.

[19] Klaus Schoeffmann, Manfred Del Fabro, Tibor Szkaliczki, Laszlo Böszörmenyi, and Jörg Keckstein. Keyframe extraction in endoscopic video. *Multimedia Tools and Applications*, 74(24):11187–11206, 2015.

[20] Klaus Schoeffmann, Marco A. Hudelist, and Jochen Huber. Video interaction tools: A survey of recent work. *ACM Comput. Surv.*, 48(1):14:1–14:34, September 2015.