Statistical modeling and recognition of surgical workflow

Nicolas Padoy [a,*], Tobias Blum [b], Seyed-Ahmad Ahmadi [b], Hubertus Feussner [c], Marie-Odile Berger [d], Nassir Navab [b]

[a] Engineering Research Center for Computer-Integrated Surgical Systems and Technology, Johns Hopkins University, Baltimore, Maryland, USA
[b] Computer Aided Medical Procedures (CAMP), Technische Universitat Munchen, Germany
[c] Department of Surgery, Klinikum Rechts der Isar, Technische Universitat Munchen, Germany
[d] LORIA-INRIA Lorraine, Nancy, France

ARTICLE INFO

ABSTRACT

In this paper, we contribute to the development of context-aware operating rooms by introducing a novel approach to modeling and monitoring the workflow of surgical interventions. We first propose a new representation of interventions in terms of multidimensional time-series formed by synchronized signals acquired over time. We then introduce methods based on Dynamic Time Warping and Hidden Markov Models to analyze and process this data. This results in workflow models combining low-level signals with high-level information such as predefined phases, which can be used to detect actions and trigger an event. Two methods are presented to train these models, using either fully or partially labeled training surgeries. Results are given based on tool usage recordings from sixteen laparoscopic cholecystectomies performed by several surgeons.

1. Introduction

The operating room (OR) needs to be constantly adapted to the introduction of new technologies and surgical procedures. A key element within this process is the analysis of the workflow inside the OR (Cleary et al., 2005; Lemke et al., 2004). It has impact on patient safety, working conditions of surgical staff and even overall throughput of the hospital. While new technologies add much complexity to the daily routine of OR staff, they also facilitate the design of assistance systems that can relieve the surgical staff from performing simple but time-consuming tasks and assist them in the tedious ones.

In this paper, we focus on the design of a context-aware system that is able to recognize the surgical phase performed by the surgeon at each moment of the surgery. We believe that robust real-time action and workflow recognition systems will be a core component of the Operating Room of the Future. They will enable various applications, ranging from automation of simple tasks to detecting failures, suggesting modifications, documenting procedures and producing final reports.

Our contribution is threefold: In the introduction we identify the generic need for an automatic recognition system and introduce a signal based modeling of the surgical actions to achieve such automation. We then propose two statistical models constructed from generic signals from the OR: the annotated average surgery and the annotated Hidden Markov Model, for off-line and on-line recog-

nition of the surgical phases in a standard endoscopic surgery. The models are built based on a set of training surgeries where the phases have been labeled. We also propose a method only requiring partially labeled data. We finally demonstrate and evaluate the methods on the example of laparoscopic cholecystectomy using binary instrument usage information and illustrate their use in some potential applications: automatic generation of report sketches, automatic triggering of reminders for the surgical staff and automatic prediction of the remaining time of the surgery.

1.1. Motivation

In the last years, many experts have tried to predict which changes need to be made to the operating room in order to meet future requirements in terms of appearance, ergonomics and operability (Satava, 2003; Feussner, 2003; Berci et al., 2004). Aiming always at better patient treatment and higher hospital efficiency, issues and solutions have been further discussed in international workshops gathering clinicians, researchers and medical companies (Cleary et al., 2005; Lemke et al., 2004). The focus has been on surgical workflow optimization, system integration and standardization, in particular in new clinical fields like image guided surgery. Parallel to these studies, different testbeds have been established to experiment and report on the development of advanced surgery rooms (Sandberg et al., 2005; Agarwal et al., 2007; Mårvik et al., 2004).

A general consensus is that all systems present in the surgery room of the future should and will be fully integrated into a final

* Corresponding author.
  E-mail address: padoy@jhu.edu (N. Padoy).

system and network. For specific procedures, companies like Brain-Lab[1] and Medtronic[2] already provide fully integrated OR suites. In Meyer et al. (2007), a surgery room is presented where various contextual informations are presented on a unified display. Such ORs, where a multitude of signals and information are available within a unique computer system, offer great opportunities for the design of powerful context-aware systems which will be reactive to the environment.

Signals are already provided by anesthesia systems, medical and video images and digital patient files. In addition, more signals from advanced electronic surgical tools and navigation systems that are tracking patients, clinical staff, surgeon and equipments will provide an extensive and rich dataset in the future. Note that the use of RFID tags is being widely investigated for tracking of material and persons inside the hospital and the OR (Egan and Sandberg, 2007; Rogers et al., 2007). Another environment providing rich sensor information is robotic surgery. The availability of these signals on a daily basis will highly facilitate the analysis and recognition of all actions performed in the OR.

Context-awareness is not only profitable for assistance inside the OR. In many cases, OR delays come merely from poor synchronization between the workflows inside and outside the OR (Herfarth, 2003). Incorporating context-aware ORs inside a global hospital awareness system would therefore greatly improve the overall efficiency.

In this work, we present several methods to recognize the surgical phases of a surgery which has a well-defined workflow. They can either be applied on-line during the surgery for context-awareness, or off-line after the surgery for instance for documentation and/or report generation.

1.2. Related work

Close work comes from the robotic community, where awareness is required for robot control. Miyawaki et al. (2005), Yoshimitsu et al. (2007) target the development of a robotic scrub nurse that automatically provides the correct instrument to the surgeon. The approach is based on time-automata combined with a vision front-end. The conception of the model is however time-consuming as it is done by hand. A further limitation is that the prototype only works with a tiny set of instruments provided in a predefined order.

In Ko et al. (2007), a task model of cholecystectomy is designed for guidance of a laparoscopic robot which is controlling the camera pose. A viewing mode is assigned to each surgical stage and transition rules between the stages are manually defined based on the currently used tool that is detected using color markers. A surgical stage cannot be always uniquely recognized from the current surgical tool. Therefore some ambiguities cannot be distinguished with this deterministic approach.

Also demonstrated on cholecystectomy, we have proposed in previous work approaches based on Dynamic Time Warping for segmenting the surgical phases of the surgery using laparoscopic tool usage (Ahmadi et al., 2006; Padoy et al., 2007). In Padoy et al. (2008), we proposed an on-line approach using Hidden Markov Models constructed from data containing visual cues computed from the endoscopic video. In Blum et al. (2008), we addressed the automatic generation of human-understandable models of surgical phases. We propose in this paper a unified framework, address the case of data where the phases have only been labeled partially and present in more details the motivations and applications.

A statistical model, signals and a detection approach that relates the signals to the model are needed to recognize the surgical phase. Above references strive to provide a complete solution. Other existing work addresses one of these three aspects. In Jannin and Morandi (2007), a model based on Unified Markup Language (UML) is proposed in order to understand and optimize the usage of imaging modalities during a neurosurgical procedure. Neumuth et al. (2006) presents ontologies and tools to describe and record surgeries in a formal manner. The actions and interactions occurring in surgeries can be recorded manually by assistants using a software that helps generating standardized descriptions, which can be in turn used for an in-depth analysis of the workflow. The method has been validated on a large dataset of simulated data (Neumuth et al., 2009). This manual approach is complementary to ours in that it permits in-depth formal description and understanding of the workflow despite the fact that only few sensors are currently available in the OR, as required for monitoring.

Interesting signals for the analysis of surgical gestures are the positions of tools or the forces applied to them. They can be obtained indirectly using a tracking system or directly when a robot is used and thus provides the positioning information. Such signals have been mainly used for evaluating and comparing surgeons performing on a simulator (Rosen et al., 2006; Megali et al., 2006; Lin et al., 2006; Leong et al., 2007). For recognition of several actions in minimally invasive surgery, Lo et al. (2003), Speidel et al. (2008) use the endoscopic video. In James et al. (2007), the surgeon's attention is tracked with an eye-gaze system to detect the clipping of the cystic duct during a pig cholecystectomy.

Beyond surgeries, different work is also addressing context-awareness in clinical environment. For instance Xiao et al. (2005), Bhatia et al. (2007) use either vital signs available from the anesthesia systems or an external camera to automatically detect when patients are entering or leaving the surgery room.

Dynamic Time Warping (DTW) (Sakoe and Chiba, 1978) and Hidden Markov Model (HMM) (Rabiner, 1989) algorithms have emerged from the speech recognition community and have since been extensively used for classification in many domains. In the following, we do not use them for classification, but for constructing a statistical model of a surgical workflow, in which we intend to recognize the phases. This is permitted by the notion of an annotated model that we introduce.

1.3. Organization

The motivation for signal based workflow recognition was presented in Section 1.1, related work in Section 1.2. The remainder of this paper is as follows: the two core models are introduced in Section 2. Their use in segmentation and recognition is given in Section 3. Section 4 presents the medical application and an experimental comparison of their performance. Different applications of the off-line segmentation and on-line recognition as well as a discussion of the results are provided in Section 5. Finally, conclusions are given in Section 6.

2. Model construction

In this section, we explain how to construct two statistical models whose usage for phase recognition is described in Section 3: the annotated average surgery and the annotated Hidden Markov Model.

2.1. Assumptions and notations for signals and phases

We assume the signals from the operating room to be normalized and represented by a multidimensional time-series $\mathbb{O}$ where $\mathbb{O}_t \in [0, 1]^K$. In the example of laparoscopic cholecystectomy which is used in our experiments and detailed later in Section 4, the signals represent the instrument usage at a temporal resolution of one second and

$$\mathbb{O}_{t,k} = 1 \textbf{ if and only if} \text{ instrument } k \text{ is used at time } t. \tag{1}$$