Real-time estimation of surgical procedure duration

Geudon, A.C.P.; Paalvast, M.; Meeuwsen, F.C.; Tax, D.M.J.; van Dijke, A.P.; Wauben, L.S.G.L.; Van Der Elst, M.; Dankelman, J.; van den Dobbelsteen, J.J.

DOI
10.1109/HealthCom.2015.7454464

Publication date
2015

Document Version
Final published version

Published in
HealthCom 2015 - 17th International Conference on E-Health Networking, Application and Services

Citation (APA)
Geudon, A. C. P., Paalvast, M., Meeuwsen, F. C., Tax, D. M. J., van Dijke, A. P., Wauben, L. S. G. L., Van Der Elst, M., Dankelman, J., & van den Dobbelsteen, J. J. (2015). Real-time estimation of surgical procedure duration. In *HealthCom 2015 - 17th International Conference on E-Health Networking, Application and Services: 1st International Workshop on IOR* (pp. 6-10). Article 7454464 Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/HealthCom.2015.7454464

Real-time estimation of surgical procedure duration

ACP Guédon, M Paalvast, FC Meeuwsen, DMJ Tax, AP van Dijke, LSGL Wauben, M van der Elst, J Dankelman, JJ van den Dobbelsteen

Abstract*— *Efficiency in the Operating Room (OR) is a topic of growing interest. Planning of care is a crucial element to ensure optimal use of the ORs. Currently, OR scheduling is considered as a complex task based on predictions of surgery duration. The latter are often based on average times, but turn out to be inaccurate in practice because of various factors (such as complexity, patient's characteristics, unexpected events, etc). The aim of this study is to develop a prediction system that estimates in real-time the remaining duration of a surgical procedure. The prediction system was based on monitoring the progress of a procedure by recording the activation of a single piece of equipment in the OR, the electrosurgical device. Support Vector Machines was then used as a classifier to predict the remaining surgical procedure duration and thereby the optimal timing to start preparing the next patient for surgery. The classifier was trained with data on the activation of the electrosurgical device during 55 laparoscopic cholecystectomies. The performance tests showed a mean error rate about 0.2, which means that about 80% of the procedures were classified correctly. The real-time prediction system is a promising tool to improve OR planning and decrease unnecessary patients' waiting times.

Keywords*— *operating room; planning; workflow; procedure duration; pattern recognition; support vector machines.

I. Introduction

The attention for process and cost efficiency in the Operating Room (OR), which is the largest cost and revenue center of hospitals, has grown in the past few years [1, 2]. Various studies have underlined the importance of adequate OR scheduling [3-6]. These studies show that factors such as the availability of personnel and equipment, elective and urgent surgeries, and preparation of patients prior to surgery influence the planning. Moreover, a reliable prediction of surgery duration for each individual case is needed to achieve an optimal OR planning [5, 6]. The many factors that have to be taken into account make OR planning complex and often inaccurate [5, 6]. Information on the progress of a surgery is crucial to constantly adjust the planning and keep it optimal throughout the day [5, 7]. In the hospital involved in this study, updates about the progress of a surgery are exchanged by means of phone calls between the OR team and the OR scheduler or by entering the OR to discuss it. Both ways are disrupting the surgical process and are therefore not desirable [8, 9].

Currently, an OR team member makes a phone call to the preoperative holding area when the next patient needs to be prepared for surgery. The timing of the phone call is watched by the nurse anesthetists and the surgeons, and is dependent on their personal views on the progress of surgery. This timing influences directly the journey of the patient in the hospital. A too optimistic view on the remaining surgery duration induces longer waiting times for the patients and overload of the preoperative holding area [10]. In contrary, if the patients are prepared too late, the OR will remain unnecessarily unused. An optimal timing for the start of the preparation of the next patient is desired in order to reduce waiting times, improve patients' experience and to streamline the patient flow to the preoperative holding area with the OR schedule. Monitoring the progress of procedures and providing automatically a reliable update of the estimation of the end time of surgery, and subsequently the timing for preparing the next patient, would therefore be valuable.

Previous studies investigated various monitoring methods for workflow mining in the OR [11-18]. They mainly monitored the usage of devices and instruments in order to recognize the surgical phase at each moment of the surgery. The two main methods for phase recognition were Dynamic Time Warping and Hidden Markov Models [15]. Dynamic Time Warping is a very reliable method for phase recognition but the main limitation is that data on whole procedures must be recorded before the method can be applied [12, 15]. Therefore, it is not suited for real-time predictions. Studies using Hidden Markov Models showed a high accuracy when data on the entire surgery was available [13, 15]. For this method, instrument and device usage was manually obtained from video data. Although it is technically feasible, the limitation is the necessity to develop sensors for monitoring automatically the numerous signals during surgery. For real time applications, this method is applicable but obtaining an accurate prediction is challenging [15]. For the sake of simplicity and practicality, we are aiming to limit the monitoring of the progress of the procedure to one single piece of equipment. A supervised classifier is used to predict for each point in time if the next patient can be called into the OR. In contrast to unsupervised Hidden Markov Models that model the phases of a procedure, a supervised classifier is not aiming at describing the phases, but focuses on predicting the timing for preparing the next patient as accurately as possible. Support Vector Machines (SVM) has been selected as the best performing classifiers among various others for the set of data presented in this paper.

The aim of this study is twofold: 1) to determine the optimal timing for the preparation of the next patient and 2) to develop a

This work was supported by DSW Zorgverzekeraar.

ACP Guédon, M Paalvast, FC Meeuwsen, LSGL Wauben, AP van Dijke, J Dankelman and JJ van den Dobbelsteen are with the department of Biomechanical Engineering, Delft University of Technology, Delft, the Netherlands (a.c.p.guedon@tudelft.nl).

M van der Elst is with the department of Surgery, Reinier de Graaf Groep, Delft, the Netherlands.

DMJ Tax is with the Pattern Recognition Laboratory, Delft University of Technology, Delft, the Netherlands.

real-time prediction system that estimates this optimal timing according to the progress of the procedure, by registering signals from a single piece of equipment in the OR.

II. Methods

A. *Timing for preparation of patients*

For this study, laparoscopic cholecystectomies were selected to test the prediction system. These procedures are performed frequently and are relatively standardized.

Data on surgical procedure duration, timing of the phone call to the preoperative holding area to start preparing the next patient and patient waiting times were gathered for laparoscopic cholecystectomies at a Dutch teaching hospital between 01/2013 and 10/2014. 157 procedures were selected, after removing the procedures that were missing data on procedure duration or timing of the phone call. Then, the optimal timing of the phone call was determined, taking additional information provided by the OR personnel and the protocols of the hospital into account.

B. *Prediction System*

1) Monitoring the activation of electrosurgical device

During laparoscopic cholecystectomy procedures, an electrosurgical device (Valleylab, Force FX or Valleylab, Force Triad) is used to remove the gallbladder from the liver. This device was selected as the piece of equipment to monitor as its activation matches certain stages of the procedure. A current sensor (Fig. 1) was developed to record the amount of current delivered to the electrosurgical device. The current sensor was placed between the power plug of the device and socket in the OR and logged the current used by the device approximately 10 times per second. This provided an accurate recording of the activation of the electrosurgical device during the procedure.

Fig. 1: Electrosurgical device and current sensor (encircled)

The data corresponding to laparoscopic cholecystectomies were filtered out by using the time of the first incision and the last suture registered in the hospital's digital planning system (Chipsoft EZIS). An example of raw data of a procedure is shown in Fig. 2.

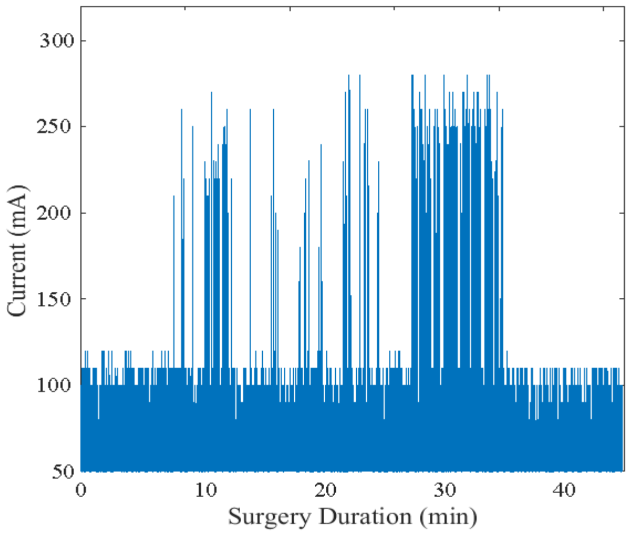

Fig. 2: Example of raw data obtained from the current sensor

2) Pattern recognition

The activation pattern of the electrosurgical device was used as input to recognize the progress of the surgery. Activations were defined when the measured value was larger than a selected threshold. From the activation patterns the following features were extracted each five minutes:

- First time the device was activated
- Last time the device was activated
- Number of activations
- Total duration of activation
- Number of activations with minimum intervals of 30 sec
- Binary indicating if the procedure was already finished

With these features, a prediction is made if, at a certain time point, the next patient should be called in for preparation or not. This means that there are two classes. One class tells that the surgery takes less than a certain time and hereby that it is time to start preparing the next patient. The other class tells the surgery takes a longer time and the preoperative holding can wait to prepare the next patient.

In this study, Support Vector Machines (SVM) was used as pattern recognition method. SVM is a binary classifier that attempts to linearly separate the time points in their distinct classes. After a SVM classifier is trained on the available example data (training data for which the end time of surgery is known), it can be applied to new and unseen data. The SVM classifier was trained to make a prediction based on the features extracted from the current sensor during 55 laparoscopic cholecystectomies using PRTools (statistical toolbox for MATLAB).

The features were normalized to make each feature weight the same in classification. Using forward feature selection, the above-mentioned features were determined to positively affect the performance of the classifier. Additionally, various features on patient and surgeon (patient age and BMI, surgeon ID, and surgeon in training ID) were tested, but did not increase the performance and were therefore abandoned.

3) Real-time prediction system

In order to use the prediction in real-time, the classifier must work with data of an on-going surgery. During the first 15 minutes after the first incision, the features are extracted from the data of the current sensor, they are scaled and placed in the classifier to obtain a prediction result. After the first 15 minutes, the classifier makes its first prediction on whether or not the next patient can be prepared. The system will keep making prediction every five minutes until the prediction that it is time to prepare the next patient or until 45 minutes have past. Therefore 6 classifiers had to be trained (15, 20, 25, 30, 35 and 45 minutes). The minimum of 15 minutes and the maximum of 45 minutes were set based on the availability of data to train the system. A prediction every five minutes was chosen to keep the amount of classifiers to train (and therefore the computation time) to a minimum and yet to provide enough accuracy about the timing.

The accuracy of the prediction system and the effect of the number of samples were tested by randomly splitting the data 25 times. 70 percent of the data was used to train the classifier and the remaining 30 percent of the data was used for testing the performance. The mean error and standard deviation were obtained for each of the classifiers. Finally, a simulation was performed where 17 procedures (30 percent of the data) are progressively processed by the classifier, until the prediction model determined it was time to start preparing the next patient, as it would do when tested in a real environment. The results of the simulation were compared to the ideal prediction, 25 minutes before the end of the procedures.

III. Results

A. Timing for preparation of patients

During the course of a procedure, the next patient is already prepared for surgery. That means that first, a phone call has to be made from the OR to the preoperative holding area to start the process. Then, the patient is transferred from the nursing department to the preoperative holding area. This should take a maximum of 15 minutes according to the hospital's protocols. Afterwards, the necessary preparations are performed before the patient can enter the OR. This should take a maximum of 10 minutes. 25 minutes are thus needed for the preparation of a patient. After discussions with the OR personnel, it was decided that 10 minutes extra buffer time was advisable. The time needed between the phone call and the patient entering the OR was set at 35 minutes. Consequently, these 35 minutes equals the preparation and waiting time for the patient.

For the 157 observed procedures, the mean preparation and waiting time for the patient was 47 minutes (SD: 17 min, median: 43 min). Plotting the preparation and waiting time on the y-axis and the procedure length of the previous surgery and the x axis (Fig.3) showed that the timing of the phone call was not optimal and that the majority of the patients had to wait for an unnecessarily long time.

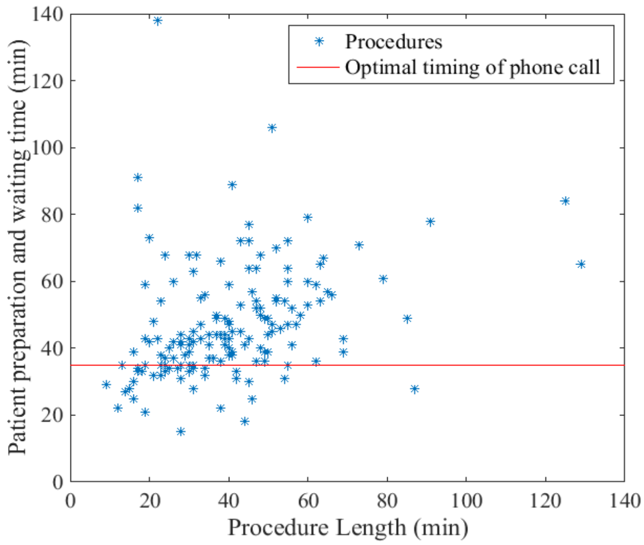

Fig 3. Timing of the phone call to prepare the next patient

B. Prediction system

After the last suture of a surgery, the patient needs to wake up, to be transferred to the recovery area and the OR needs to be cleaned, which take a minimum of 10 minutes. The optimal timing for the phone call to start preparing the next patient was therefore set at 25 minutes before the last suture of the surgery. The time that separates the two classes of the SVM classifier was consequently set at 25 minutes.

The accuracy of the prediction system for all classifiers is shown in Table 1 in terms of mean error rate and standard deviation. A prediction was considered as inaccurate when the classifier predicted the surgery to be in the wrong class (e.g. the classifier predicted a surgery to take less than a certain time, but it turned out to take a longer time).

The effect of the sample size on the mean error rate is shown in Fig. 4. The 45 minutes classifier is not displayed due to the low amount of samples available for one of its classes.

Table 1: Mean error rate and standard deviation of all classifiers

	Mean Error	SD
15 min classifier	0.25	0.10
20 min classifier	0.31	0.11
25 min classifier	0.22	0.08
30 min classifier	0.21	0.08
35 min classifier	0.16	0.07
40 min classifier	0.21	0.09
45 min classifier	0.14	0.09

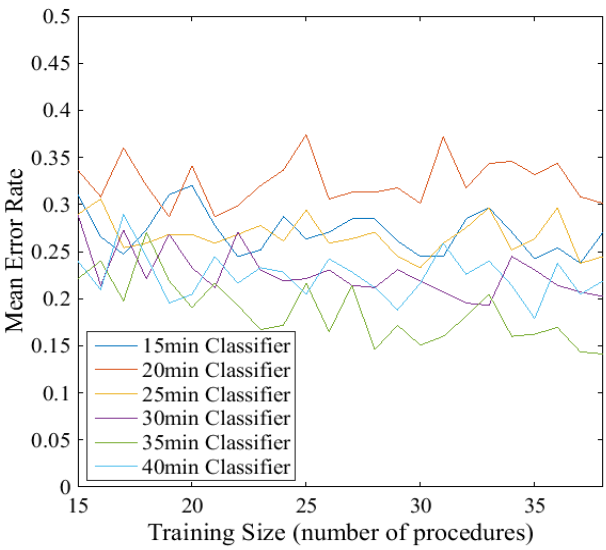

Fig. 4: Mean error rate with increasing training sample size

The mean error rate gives an indication of the performance of the prediction system in terms of right or wrong choice of class (i.e. shorter of longer than a certain duration). Additional information on the performance, in terms of proximity to the ideal time of prediction, is provided in Fig. 5, which shows the results of the simulation using 17 procedures. The x-axis represent the procedure duration and the y-axis the time at which the prediction model determined the next patient should be prepared. The red line represents the ideal time of the prediction, 25 minutes before the end of the procedure. Predictions beneath the red line were considered too early, while procedures above the line were considered too late. The closer a sample procedure is to the red line, the better the prediction. Predictions within a scope of ten minutes of the ideal line, shown in the figure between the two yellow lines, were considered as sufficiently close to the ideal prediction to be used in practice.

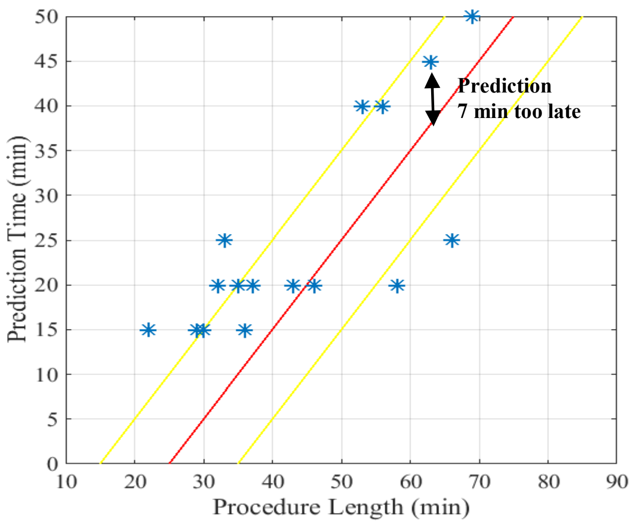

Fig. 5: Simulation using 17 procedures. The red line represents the ideal prediction time. The two yellow lines represent the range that is acceptable in practice

IV. Discussion

This paper presents a real-time system to predict the remaining duration of a surgical procedure, and hereby the optimal timing to start preparing the next patient. The system is based on monitoring the activation of a single piece of equipment in the OR, the electrosurgical device.

First, the optimal timing for preparation of the next patient was determined as 25 minutes before the last suture, according to historical data and discussions with OR personnel. Then, the prediction system was presented and the accuracy of the prediction was tested. Using the data of 55 laparoscopic cholecystectomies, an estimated error rate of 0.2 was found by splitting the available data in a set of 70% training data and 30% testing data. A slightly downward trend in error rate was observed when more samples were used for training the classifiers. The error rate is expected to decrease further if more data of procedures would be available.

Differences in the activation pattern of the electrosurgical device were observed between surgeons. Some of them activated the electrosurgical device before clipping the cystic duct, while others clipped first. This may have negatively affected the performance of the classifiers as the prediction system did not take into account which surgical method was applied. In the future, the classifiers could be trained for each surgeon to improve performance of the prediction system, keeping in mind that enough samples have to be available to train the classifiers. This leads us to a limitation of the proposed method; a certain minimum amount of training data have to be available first in order to properly train the classifiers.

This paper focused on laparoscopic cholecystectomies. However, the same prediction system could be used for other procedures. The classifiers would only have to be trained using data of the other types of procedures. Moreover, the prediction model used data on the activation of a single device in the OR. The same method could also be applied for activations of other devices by adding features from these devices to the existing list of features (assuming data about the use of these devices is available). The performance could hereby be improved further. Especially surgeries where equipment is used at a specific phase within the procedure would benefit from the extra features. This would allow the real-time prediction system to be widely usable to update the OR planning.

At this moment, parts of the prediction system are not automated yet, but it would be desirable in order to benefit at best from the application when used in an OR. The system should be started automatically at the first incision and the prediction for optimal timing for the phone call should be provided automatically to the appropriate OR personnel.

To conclude, the real-time prediction system presented in this paper is a promising tool to improve OR planning, decrease patients' waiting times and streamline the patient flow from the nursing department to the preoperative holding area and the OR.

REFERENCES

[1] B. Cardoen, E. Demeulemeester, and J. Beliën, "Operating room planning and scheduling: A literature review," *European Journal of Operational Research,* vol. 201, pp. 921-932, 2010.

[2] B. T. Denton, A. J. Miller, H. J. Balasubramanian, and T. R. Huschka, "Optimal Allocation of Surgery Blocks to Operating Rooms Under Uncertainty," *Operations Research,* vol. 58, pp. 802-816, Jul-Aug 2010.

[3] T. Archer and A. Macario, "The drive for operating room efficiency will increase quality of patient care," *Curr Opin Anaesthesiol,* vol. 19, pp. 171-6, Apr 2006.

[4] F. Dexter, R. H. Epstein, R. D. Traub, and Y. Xiao, "Making management decisions on the day of surgery based on operating room efficiency and patient waiting times," *Anesthesiology,* vol. 101, pp. 1444-53, Dec 2004.

[5] F. Dexter, R. H. Epstein, J. D. Lee, and J. Ledolter, "Automatic updating of times remaining in surgical cases using bayesian analysis of historical case duration data and "instant messaging" updates from anesthesia providers," *Anesth Analg,* vol. 108, pp. 929-40, Mar 2009.

[6] M. J. Eijkemans, M. van Houdenhoven, T. Nguyen, E. Boersma, E. W. Steyerberg, and G. Kazemier, "Predicting the unpredictable: a new prediction model for operating room times using individual characteristics and the surgeon's estimate," *Anesthesiology,* vol. 112, pp. 41-9, Jan 2010.

[7] I. B. Vermeulen, S. M. Bohte, S. G. Elkhuizen, H. Lameris, P. J. Bakker, and H. La Poutré, "Adaptive resource allocation for efficient patient scheduling," *Artificial intelligence in medicine,* vol. 46, pp. 67-80, 2009.

[8] A. N. Healey, N. Sevdalis, and C. A. Vincent, "Measuring intra-operative interference from distraction and interruption observed in the operating theatre," *Ergonomics,* vol. 49, pp. 589-604, Apr 15-May 15 2006.

[9] A. N. Healey, C. P. Primus, and M. Koutantji, "Quantifying distraction and interruption in urological surgery," *Qual Saf Health Care,* vol. 16, pp. 135-9, Apr 2007.

[10] L. S. G. L. Wauben, A. C. P. Guedon, D. F. De Korne, and J. J. van den Dobbelsteen, "Tracking surgical day care patients using RFID technology," *BMJ Innovations,* pp. 1-8, 2015.

[11] T. Blum, N. Padoy, H. Feußner, and N. Navab, "Workflow mining for visualization and analysis of surgeries," *International journal of computer assisted radiology and surgery,* vol. 3, pp. 379-386, 2008.

[12] T. Blum, N. Padoy, H. Feußner, and N. Navab, "Modeling and online recognition of surgical phases using hidden markov models," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, ed: Springer, 2008, pp. 627-635.

[13] N. Padoy, T. Blum, H. Feussner, M.-O. Berger, and N. Navab, "On-line Recognition of Surgical Activity for Monitoring in the Operating Room," in *AAAI*, 2008, pp. 1718-1724.

[14] S.-A. Ahmadi, T. Sielhorst, R. Stauder, M. Horn, H. Feussner, and N. Navab, "Recovery of surgical workflow without explicit models," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, ed: Springer, 2006, pp. 420-428.

[15] N. Padoy, T. Blum, S.-A. Ahmadi, H. Feussner, M.-O. Berger, and N. Navab, "Statistical modeling and recognition of surgical workflow," *Medical Image Analysis,* vol. 16, pp. 632-641, 2012.

[16] N. Padoy, T. Blum, I. Essa, H. Feussner, M.-O. Berger, and N. Navab, "A boosted segmentation method for surgical workflow analysis," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, ed: Springer, 2007, pp. 102-109.

[17] L. Bouarfa, P. P. Jonker, and J. Dankelman, "Discovery of high-level tasks in the operating room," *Journal of biomedical informatics,* vol. 44, pp. 455-462, 2011.

[18] T. Blum, H. Feußner, and N. Navab, "Modeling and segmentation of surgical workflow from laparoscopic video," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, ed: Springer, 2010, pp. 400-407.