

Context-aware Augmented Reality in laparoscopic surgery

Darko Katić^{a,*}, Anna-Laura Wekerle^{b,1}, Jochen Görtler^a, Patrick Spengler^a,
Sebastian Bodenstedt^a, Sebastian Röhl^a, Stefan Suwelack^a, Hannes Götz Kenngott^b,
Martin Wagner^b, Beat Peter Müller-Stich^b, Rüdiger Dillmann^a, Stefanie Speidel^a

^a Institute for Anthropomatics (IFA), Karlsruhe Institute of Technology (KIT), Germany

^b Department of General, Abdominal and Transplantation Surgery, University of Heidelberg, Germany

A B S T R A C T

Augmented Reality is a promising paradigm for intraoperative assistance. Yet, apart from technical issues, a major obstacle to its clinical application is the man-machine interaction. Visualization of unnecessary, obsolete or redundant information may cause confusion and distraction, reducing usefulness and acceptance of the assistance system.

We propose a system capable of automatically filtering available information based on recognized phases in the operating room. Our system offers a specific selection of available visualizations which suit the surgeon's needs best. The system was implemented for use in laparoscopic liver and gallbladder surgery and evaluated in phantom experiments in conjunction with expert interviews.

Keywords:

Augmented Reality

Context-awareness

Laparoscopy

Knowledge representation

1. Introduction

With advances in medical imaging and the increasing availability of intraoperative sensors, there is more and more information available to surgeons. High-resolution endoscopes, real-time tracking of instruments as well as instrument recognition via RFID and image processing generate large amounts of information. The downside of this development, however, is increased workload for surgeons. The ensuing information overflow can easily be distracting and tiresome. In order to integrate and make use of the additional information from different sources, assistance systems with new user interfaces are needed. Augmented Reality (AR) has emerged as a paradigm for this purpose. Yet, AR even with high accuracy registration and advanced visualization metaphors alone may prove insufficient to deal with huge quantities of data. Abundant visualizations pose a risk because they distract surgeons from the surgical field, especially since visualizations are commonly overlaid on top of areas of interest. Context-aware visualizations, on the other hand, provide visualizations just in time without handicapping the surgeon. Such systems have a certain understanding of the situation, enabling them to automatically derive the current information needs of the surgeon. This enables them to filter the huge amount of available information and selectively

display only what is currently relevant. For instance, it is only necessary to display vital structures like nerves or blood vessels when they are actually in danger of being harmed. Thus, the system automatically adapts itself to the needs of the surgeon without taking away focus from his actual task, the execution of the surgery.

Based on concepts introduced in [1,2], we develop a context-aware system for laparoscopic liver surgeries. The basic idea is to monitor the progress of the surgery with sensors, e.g. intraoperative data from medical devices in the operating room (OR), positional information from tracking systems or analyzed endoscope images. The current situation is then expressed in a computational model using Description Logics and enriched with background information from an ontology. This allows us to use interpretation algorithms to recognize the current phase of the operation. After that, it is possible to let the visualization adapt itself automatically in accordance to the current context. This is of special importance in laparoscopic liver surgery since the area of interest contains numerous vital structures which should not be harmed. Displaying all of them during the entire surgery would not be helpful. It would rather overwhelm the surgeon and make him disregard the visualizations after a while. Context-awareness, on the other hand, allows the system to selectively display vital structures just when needed. The visualizations, therefore, are not distracting and the surgeon is more likely to pay attention to them since they are a rather rare but important occurrence. Similarly, navigational information about the position of tumors and preoperatively collected image data, while highly important

* Corresponding author. Tel.: +49 72160848082.
E-mail address: katic@kit.edu (D. Katić).

¹ These authors contributed equally to this study.

for patient outcome, is only useful during specific parts of the surgery and should just be displayed in those cases.

We present the basic architecture of the system as well as the algorithms we developed, additionally to results from experiments on a phantom and in a realistic setup for laparoscopic cholecystectomy using a pig liver in a surgical simulator. Our main contributions are methods for ontology-based situation interpretation in the OR. We aim to show what context-awareness can add to AR systems in medical applications such as laparoscopy and dental implant surgery [3].

2. Context-aware and Augmented Reality systems in surgery

AR systems have long been in focus of research activities. The main issues revolve around finding ways to display visualizations and to track real-world objects. A categorization and review of current systems and research is presented by Nicolau et al. [4]. Endoscopic soft tissue surgery poses a special challenge since established optical and electromagnetic tracking devices only deal with rigid structures and are not readily able to track soft tissue deformations. To overcome these problems, a number of approaches have emerged. Intraoperative image acquisition efforts with ultrasound, MRI and CT have been undertaken as well as approaches using endoscopes as sensors to track artificial or natural landmarks, as reviewed by Baumhauer et al. [5] and Mirota et al. [6].

Those systems, in contrast to context-aware ones, are aimed at improving image and registration quality as well as accuracy, especially under soft tissue conditions. Yet to actually bring AR systems into the OR, improvements in accuracy and display quality are fundamental, but not sufficient. New interfaces and interaction techniques will still have to be developed. To handle the increased workload, information flow and to manage complex assistance functions, context-awareness is a very promising paradigm. Apart from the intraoperative information filtering and automatic selection of visualizations, context-awareness itself can also support intraoperative decision making by comparing current situations with previously recorded ones. In this way, context-aware systems can intraoperatively provide information about how other, possibly more experienced, surgeons have handled similar situations. Context-awareness also helps in robotic surgery by facilitating human-machine collaboration [7,8]. Furthermore, instrument changes needed in upcoming phases can be predicted to help staff members to prepare in time [9]. Estimations of the surgery's remaining duration can be used to plan anesthetizations. Postoperatively, semi-automatic creation of reports and quality assessment such as surgical skill evaluation [10,11] are possible.

Situation interpretation is a crucial part of context-aware systems since it allows a certain kind of understanding of the events in the OR. For this purpose two main algorithmic approaches have emerged. On the one hand, there are efforts to use machine learning techniques to interpret situations. Blum et al. extract several features from endoscopic images like gradients, histograms and scaled versions of the original image [12,13]. These feature vectors are used to represent surgeries. Interpretation is done by using statistical models, namely Hidden Markov Models (HMMs) and Dynamic Time Warping. Lallys et al. use HMMs and Support Vector Machines to recognize surgical phases [14]. A combination of Bayesian networks and HMMs is used in [15]. In a similar vein, Rosen et al. employ discrete Markov models to model minimally invasive surgeries [16].

On the other hand, there are efforts to use formal methods such as ontologies to model medical background knowledge in a machine readable fashion [17,18]. They can be directly

incorporated with knowledge representation techniques, which in our case are Description Logics in OWL (Web Ontology Language) [19,20]. Additionally, this approach improves safety since large parts of the hypothesis and the underlying assumptions are explicitly stated in a human readable fashion. This leads to easier inspection and verification.

In contrast to existing systems, we aim to provide context-aware AR assistance, with an emphasis on seamless integration in the surgeon's workflow. Adaption to the current conditions and processes in the OR, without direct input from the surgeon, is key to this aim. We follow a novel approach using generic knowledge representation techniques, as offered by OWL, to convey general information about abdominal surgeries. This allows us to adapt our system more easily to different types of interventions. In contrast to machine learning-based approaches, we need fewer training samples. Pieces of information known to medical experts do not need to be learned by using training samples. This is of special importance in the medical field, where expert knowledge, from literature and clinicians, is plentiful while training data is sparse and hard to come by. Compared with the existing work in formal description of surgeries, our emphasis lies in the real-time analysis of incoming data and the reuse of our ontology and interpretation algorithms in a Description Logic based framework. Our goal is to develop a system which can be easily adjusted for different laparoscopic surgeries. Furthermore we aim to realize the idea of an information filter by intraoperatively controlling visualizations to limit informational overflow and minimize the time spent configuring the system so that surgeons can fully focus on the intervention itself. The technical idea is to mix formal approaches with machine learning techniques to overcome limitations inherent to both paradigms. Compared to our work in dental implant surgery [3], the field of laparoscopy offers more diverse and complicated situations, as there are more instruments, anatomical structures and surgical skills involved. This calls for more background knowledge and an extended ontology.

3. Methods

Our methods focus on algorithms for situation interpretation since they are vital for the reliable display of context-aware visualizations. The situation interpretation is based on the surgical events currently happening in the OR and its aim is to recognize the surgical phase. For the purpose of this paper, we define a surgical event to be a vector with three components: the instrument, the surgical activity performed and the anatomical structures acted upon. For instance, we consider lifting the greater omentum with an atraumatic grasper or cutting the gastrocolic ligament with a ligasure to be single events. A surgical phase, as we define it, is a sequence of events which are logically coupled and lead to a predefined, immediate goal, such as mobilizing an organ or the resection of diseased tissue. Situation interpretation and phase recognition are used synonymously, since the result of the interpretation is the recognized phase. The technical process for providing situation interpretation services is based on knowledge-based techniques using subsymbolic input. It is divided into three individual steps. First, we collect sensor data and convert the measurements in a logic-based representation of the current scene using Description Logics. In the current experimental setup, the sensor data consists of distance measurements between the instrument and anatomical structures. The resulting description is then interpreted using a rule-based approach in order to recognize the current phase of the operation. In dependence of this result, an appropriate visualization is chosen and displayed in the final third step. These steps are detailed in the following.

3.1. Linking numerical measurements to ontological vocabulary

Handling of subsymbolic, numerical data within an ontological framework is of particular importance in the intraoperative use case since measurements from medical devices, tracking systems, etc. are usually not available in symbolic form. In order to close this semantic gap, we developed algorithms to convert subsymbolic data into logical predicates.

First, we calculate a fuzzy description of the real-valued measurement. In the case of distance measurements, we calculate the membership values to fuzzy sets like “near”, “medium” and “far”. They can be regarded as values of the linguistic variable “distance”. This step can be performed by using different approaches. We developed and compared three methods to learn the corresponding membership functions from data samples. In the second step, the measurements are converted into logical predicates. The membership values are used to find the most fitting element in the ontology, which is then assigned to the measurement. Finally, this information is used to form a computational model of the current situation, which is then interpreted to recognize the current phase of the surgery. These three steps are discussed in the following.

In order to find the fuzzy sets, we first collect labeled training samples. They consist of labels describing the values of the linguistic variable along with numerical values. To compute the fuzzy sets, we developed and evaluated three different methods: the evidence-based approach, the Bayesian approach with Gaussians and the Bayesian approach with Histograms.

3.1.1. Evidence-based approach

The evidence-based approach is motivated by Weisbrod [21]. Each training sample can be seen as proof that this particular value belongs to the corresponding fuzzy set. In this way, information is used positively, i.e. new observations enlarge the amount of real-valued measurements which belong to a certain fuzzy set. The membership functions for the fuzzy sets are represented as Gaussian mixtures, normalized to the interval [0, 1]. For each sample in the training set, a new Gaussian curve is added with a weight of 1 and a mean equal to the value of the measurement. In this way, the evidence for areas with a high density of observations is highly elevated while outliers contribute only little to their respective regions. Given a training set T of real-valued measurements m , assigned to the fuzzy set s , the membership value $\mu_s(x)$ for a measurement x is computed as:

$$\mu_s(x) = w \sum_{m \in T} e^{-(1/2)(x-m)/\sigma^2)}$$

where σ is the standard deviation and w a normalization factor with

$$w^{-1} = \max_x \sum_{m \in T} e^{-(1/2)(x-m)/\sigma^2)}$$

3.1.2. Bayesian approach with Gaussians or histograms

Both the other approaches are based on the idea that $P(c|x)$, i.e. the probability that predicate c holds, given a measurement x , can be used as a measure of how strongly the measurement x belongs to the fuzzy set corresponding to c . According to Bayes theorem, $P(c|x)$ can be computed as:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$P(x|c)$ and $P(x)$ can be estimated from training samples. We developed two different approaches to do this. In the Gaussian-based method we assume a normal distribution and estimate the corresponding parameters: mean and standard deviation. For the

histogram-based estimation we use histograms as an estimate of the actual distribution. The resulting values are normalized to [0, 1], as conforms with the definition of fuzzy sets.

3.2. Conversion to logical predicates

In order to assign new measurements to their logical predicate, we use a correspondence function. Given the current distance between two objects, we find the fuzzy set with the greatest membership value and use the corresponding ontological relation to denote the relationship between those objects. Given a set P of predicates and equally named fuzzy sets the corresponding predicate p for a measurement x is given by:

$$p = \operatorname{argmax}_{s \in P} \mu_s(x)$$

3.3. Situation interpretation

Our situation representation is grounded in the OWL standard of Description Logics. The idea in OWL is to distinguish between background knowledge, stored in the terminological box (TBox) and situational knowledge in the assertional box (ABox).

We developed a TBox representing knowledge from laparoscopic abdominal interventions. It consists of over 150 concepts, like “instrument”, “diaphragm” and “common-bile-duct” on several granularity levels and over 30 relations like “cut”, “grasp” or “bloodflow-between”. The system has extensive background knowledge about instruments, anatomical structures and surgical activities as well as relations between them. We use definitions of the domain and range of relations to denote the type of objects between which the relation is valid. Therefore, the system for instance knows that the relation “cut” only occurs between a “sharp-instrument” and an “organ”. Using the well-defined set of concepts and relations as a vocabulary, we describe the current situation in the ABox, using relations computed from numerical measurements. Due to restricted sensor data, the only relations computed on-line concern descriptions of distances (“near”, “medium” and “far”). The actual situation interpretation is performed on discretised sensor data using OWL-reasoning services on the ABox and TBox. The idea is to use handcrafted rules to describe phases and have them operate on an ABox as exemplified in Fig. 1. If the conditional term of the rule holds, a certain surgical phase is assumed. Thanks to the availability of reasoning services, the rules can be written in a general way, making use of powerful OWL features like qualified cardinality restrictions, classification and reflexive, irreflexive, symmetric, and anti-symmetric properties. For instance, risky situations, can be identified by looking for “sharp instruments”, “near”, “vitalstructures” in the current description of the situation in the ABox. The exact nature of the rules depends on the underlying type of surgery and on the phase to be recognized. Especially for standardized procedures they can be constructed from literature and with knowledge extraction techniques, e.g. interviews, directly from experts.

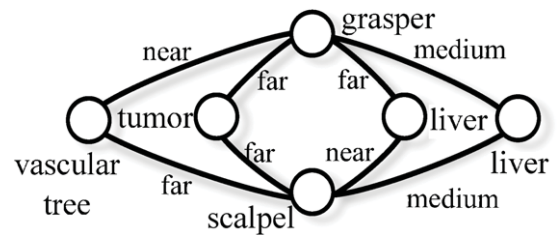


Fig. 1. ABox description of a scene.

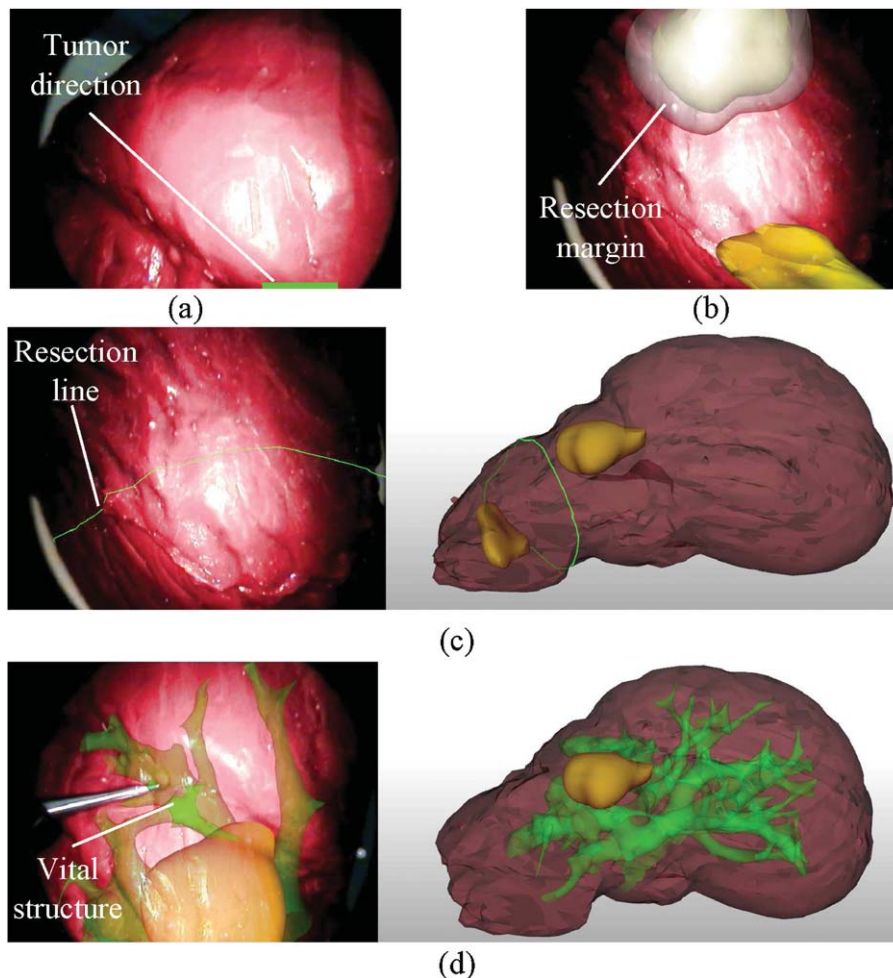


Fig. 2. Visualizations for (a) directing the surgeon to the tumor, (b) showing the resection margin and (c) the resection line and (d) displaying vital structures currently at risk.

3.4. Visualization

Depending on the recognized phase an appropriate visualization is chosen. For the alignment of the endoscope toward the target structure we developed an intuitive visualization, displaying a green rectangle at the edge of the screen, indicating the direction in which to move the endoscope to see the tumor. For the treatment of the tumor, its resection margin is displayed. Endangerment of vital structures is conveyed by displaying a virtual overlay of the structure with its opacity increasing proportionally to the proximity of the endangering instrument. Also, if the system assumes that the liver is about to be resected, it displays the preoperatively planned resection trajectory. All visualizations are shown in Fig. 2.

4. Evaluation

The evaluation is split into two parts. In the first part, we used a silicon phantom of a single liver to evaluate the basic functions of the system, giving a quantitative assessment of the phase recognition rate. In the second part, we used a pig liver, stored in a box trainer to perform a selected section of a cholecystectomy, using context-aware visualizations. The second evaluation offers qualitative results of the applicability to real-world scenarios in the OR. Both evaluations use slightly different sets of rules, which are detailed in their respective sections. The evaluation as a whole was conducted on a regular PC using Pellet [22] for logical reasoning and Protege [23] as an OWL-editing software.

4.1. Phantom evaluation

The phantom experiments evaluate the situation interpretation module and the visualizations. The experimental setup consisted of a silicon liver phantom, an endoscope and laparoscopic instruments. All entities were tracked using the Polaris optical tracking system. For the purposes of the evaluation, soft tissue aspects of the liver were not taken into consideration as all registration processes were done rigidly. The AR visualizations were superimposed on the endoscopic images and displayed using a standard monitor. The aim of this evaluation is to get quantitative data about the phase recognition rate by comparing the system's classification to a ground truth given by annotations. Furthermore, we aim to evaluate the plausibility of the learned fuzzy sets and get insight on the usability of the visualizations.

4.1.1. Evaluation of the ontology-based situation interpretation

To evaluate the phase recognition rate, two data sets were collected for training and evaluation purposes respectively. The training set consists merely of distance information and the corresponding classifications to “near”, “medium” and “far”. The evaluation set, on the other hand, contains semantic content like information about the instrument, the treated structure and the current surgical phase, in addition to the numerical values. For the evaluation, the training set is used to compute the fuzzy sets. The recognition rate is then only computed on the level of surgical phases. The transition to the symbolic domain is just a means to an

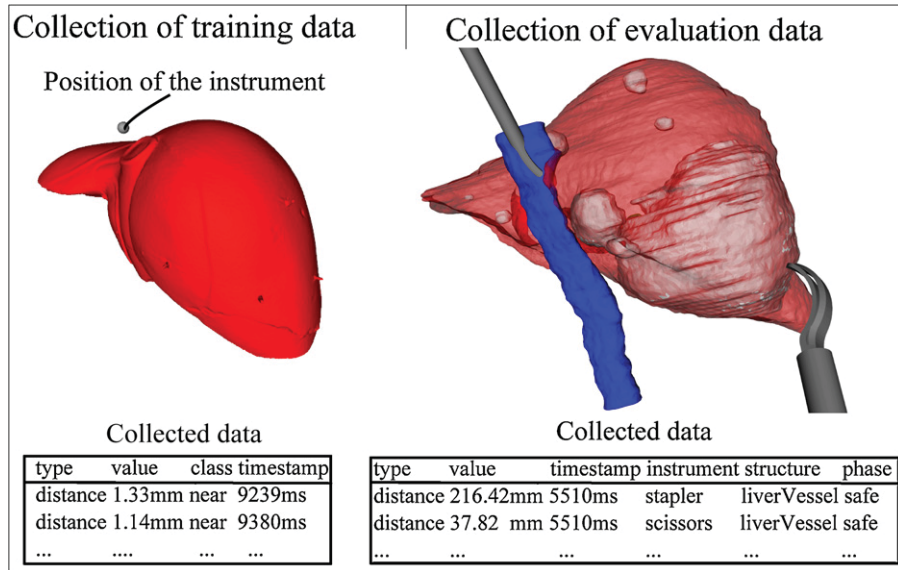


Fig. 3. Collection of samples.

end. Therefore, we treat the evaluation data set as a simulation of intraoperative sensors which are fed into the system in real time. We judge its performance based on the eventual result, the recognized surgical phase. The steps necessary to achieve this are shown below.

4.1.2. Situation interpretation

In this evaluation scenario, four phases are considered: “safe situations”, “risky situations”, “treatments of diseased structures” (e.g. tumors) and “search for target structures”. Phases can also overlap in time. The phase “search for target structures” is recognized when the target structure, i.e. the liver, is not within the endoscopes viewing frustum. This is calculated by knowing the viewing frustum of the endoscope, as given by our calibration process, and checking if the registered 3D model of the liver, as tracked by Polaris, is inside of it. As a binary, symbolic property, it can be directly handled in OWL, without any conversion from the subsymbolic to the symbolic domain. In this case the tumor is assumed to be immobile, or at least that its movements are negligible in the context of this assistance function. The performance of the system degrades gracefully with this approach. If the position of the tumor change only slightly the visualization will appear and disappear when the tumor is close to the viewing frustum (instead of exactly on the border), thus only slightly impeding the usefulness of the visualization.

The phase “treatment of diseased structures” is characterized by proximity of anatomical structures and appropriate instruments such as tumors and scalpels. If any sharp instrument is near a vital structure, e.g. a blood vessel, this structure is assumed to be endangered and a “risky situation” is detected. All other situations are considered to belong to the phase of “safe situations”. Due to the taxonomy, all the possible cases are not represented explicitly. For the recognition of “risky situations”, it is sufficient to look for “vital-structures” “near” “sharp-instruments”. This encompasses all cases with specific instruments, whether they are scissors or scalpels. An example of the resulting ABox representation, is shown in Fig. 1.

Collection of training data. In order to collect training data, a surgeon performed tasks common to liver and gallbladder surgeries on the phantom. The positions of the surgical instrument were recorded using the NDI Polaris tracking system. The corresponding annotation, i.e. the classification to “near”, “medium” and “far” was announced verbally by the surgeon and input into the system by the

click of a button in the GUI of the recording software by an assistant. All data samples collected from this moment on were considered to have the same label until the surgeon announced a new one and the assistant changed the settings accordingly. This method of on-line annotation leads to noisy data. There is a short delay between the announcement of the new label and the adjustment by the assistant. Also, there are cases where the instrument briefly leaves the intended distance range of the current label without the surgeon announcing it. Both effects lead to smeared boundaries between labels where similar values are recorded with different labels at different times. Therefore, the processing algorithms of this data must be able to deal with noisy data. On the other hand, the collection of data can be performed very quickly. In fact, the whole process to record the annotated samples took less than 3 min, once the experimental setup was ready. The 1244 resulting training samples, consisting of tuples with numerical measurements, a timestamp and the corresponding labels, were used for training purposes and contained no semantic information about instruments, anatomical structures or surgical phases, but merely (subsymbolic and symbolic) distance information. An excerpt of this process is shown in Fig. 3.

Collection of evaluation data. The evaluation set was collected by having the same surgeon perform typical tasks found in minimally-invasive liver surgeries using two optically tracked instruments. This set, in contrast to the training set, consists of tuples with numerical distances, annotated surgical phases, instruments, anatomical structures and timestamps. The annotation was done manually by a medical expert with our own software tool. It contains occurrences of all phases considered, spread over 14,605 evaluation samples. The virtual phantom consisted of a liver with several tumors and vital structures as shown in Fig. 3. Additionally, the whole process was partitioned in four aforementioned phases.

Evaluation process. In order to get insights on the effectiveness of our system, we used the data from the evaluation set, excluding the information about the current phase, as input to our situation interpretation system, which was trained using the training data. Using timestamps, the recorded values were played back, in real-time, and regarded as sensor results. At a rate of about 1 kHz, the currently recognized phase was compared to the annotated one. In this way, the system’s assumption about the current phase was constantly checked against the ground truth given by the annotation. The rate of phase recognitions is 2 Hz on average. This

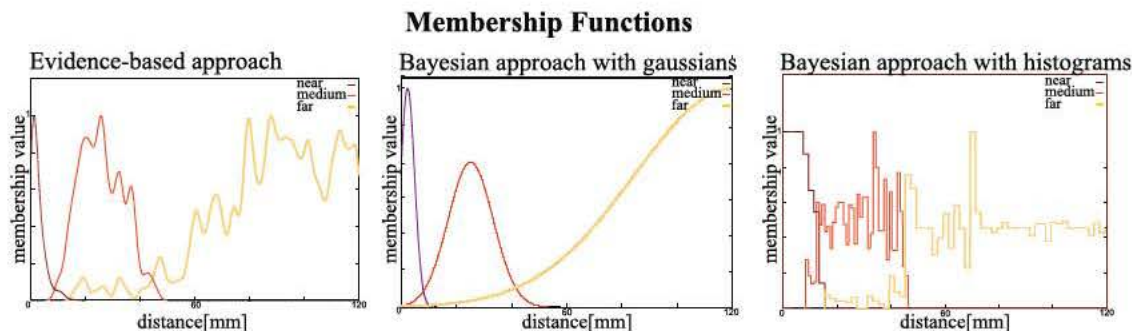


Fig. 4. Learned membership values.

allows us to consider real-time requirements of the system in the evaluation. If the system eventually recognizes the phase correctly, but takes too long to do so, the period of false recognitions will be accurately recorded at a high temporal resolution (of 1 kHz) until the systems assessment again matches the ground truth.

Plausibility of the fuzzy sets. For the evaluation, we collected 1244 training samples in two repetitions of the procedure and used all three approaches to learn the corresponding fuzzy sets. The results are shown in Fig. 4. As can be seen, all three offer plausible, yet differing results. The Bayesian approach with Gaussians yields to clean membership values. This is due to the fact that the shape of the function is highly restricted; the only degrees of freedom to be defined by the training data are the parameters for the Gaussians. On the other end of the spectrum lies the Bayesian approach with Histograms. Here, only few assumptions about the shape of the function are made. This allows for a greater flexibility but also increases the need for more data. Similarly, the evidence-based approach allows a high degree of flexibility, at the cost of higher requirements for training data. In comparison to the Bayesian approach with Histograms, the resulting function is smoother due to the underlying Gaussian curves.

Recognition rate. For the evaluation, the evaluation samples were played back to the system, which was to determine the current phase. The system's assertion was periodically compared to the annotated phase. The recognition rate was computed as the ratio of the number of times in which the phase was recognized correctly and the total amount of checks made. The phase "searching for target structures" was omitted from the evaluation, as it does not rely on the fuzzy sets. It is activated whenever the target structure is outside of the endoscope's viewing frustum and behaved, according to the participants of the study, just as expected. Overall a recognition rate of over 93% using the evidence-based approach, 94% using the Bayesian approach with Gaussians and 97% with histograms was reached.

4.1.3. Evaluation of the visualization

To evaluate the visualization, the system was presented to four medical experts in the same experimental setup as used in the phantom evaluation shown in Fig. 2. After a brief demonstration, they used the system on their own. Their reactions were recorded and they were interviewed afterwards. The feedback concerning the alignment of the endoscope toward the target structure was very favorable. The visualization is clear and concise without taking away focus from the clinician or disturbing the view on the patient. The display of the resection margin was also received well. It was suggested to add a better visualization of possible overlaps between vital structures and the resection margin. Particularly emphasized was the usefulness of vital structures being displayed only once they are actually endangered, since it lowers the informational load on the surgeon. As a possible improvement, it was suggested to augment the cutting trajectory with the corresponding CT-slices.

Overall, the visualizations in conjunctions with their context-aware management were considered useful in the proposed setting.

4.2. Laparoscopic cholecystectomy on a pig liver

In contrast to the phantom evaluation, the focus of this part emphasizes the medical merits of the system. We used a pig liver to perform a selected section of a cholecystectomy using context-aware visualization to assist the process. The goal was to visualize risk structures in the triangle of Calot such as the common hepatic duct and the common bile duct. After the experiment, we interviewed the surgeon.

Experimental setup. For the trial, we used the following experimental setup. A box trainer by Karl Storz (Tuttlingen) was fitted with a special underlayment of acrylic glass because the original, metallic one causes artifacts in CT scans. Furthermore, we installed a common earth electrode which was connected to the laparoscopic tower Fig. 5(a). As preparation of the liver, the hepatoduodenal ligament was dissected proximally to the duodenum. Contrast medium diluted with water (1:1) was injected in the common bile tract. Then the common bile duct was ligated and the liver was put over the earth electrode Fig. 5(b). The caudal part of the liver containing the gallbladder was exposed, just as in a real surgery. Subsequently the fundus of the gallbladder was sutured with the soft part of the cover. In doing so, a constant tension was created. Therefore, the organ was largely immobile throughout this procedure. After that, a CT scan was made of the liver in the box trainer Fig. 5(c). Afterwards the liver, the gallbladder with the cystic duct, the common bile duct and the common hepatic duct were segmented with MITK [24]. A Polaris tracker was affixed on the box trainer and registered using rigid registration methods.

Situation interpretation and assistance functions. For this evaluation, we defined a different set of phases and assistance functions. The main goal was to address the special challenges of the dissection of the cystic duct, namely to avoid injuries to vital structures. Firstly, the cystic duct may only be cut, if it was clipped beforehand. Otherwise, severe complications can occur. Secondly, the hepatic duct must not be cut under any circumstances.

The first phase to be recognized is therefore the endangerment of the hepatic duct by a sharp instrument. It is recognized by checking if the relation "near" holds between the respective instances in the ABox. For the cystic duct, we introduced a new Boolean attribute "clipped" which is set to true, if a clipping action was observed at this organ. We can differentiate between the phase "safe situations" when a sharp instrument is near a clipped cystic duct and "risky situations" when a sharp instrument is near an unclipped one. In the later case a warning is issued by displaying an AR overlay on the vital structure. For the recognition of clipping, we use a simple heuristic: a clipping action is assumed, if a suitable instrument is observed near a clipable anatomical structure. Since we do

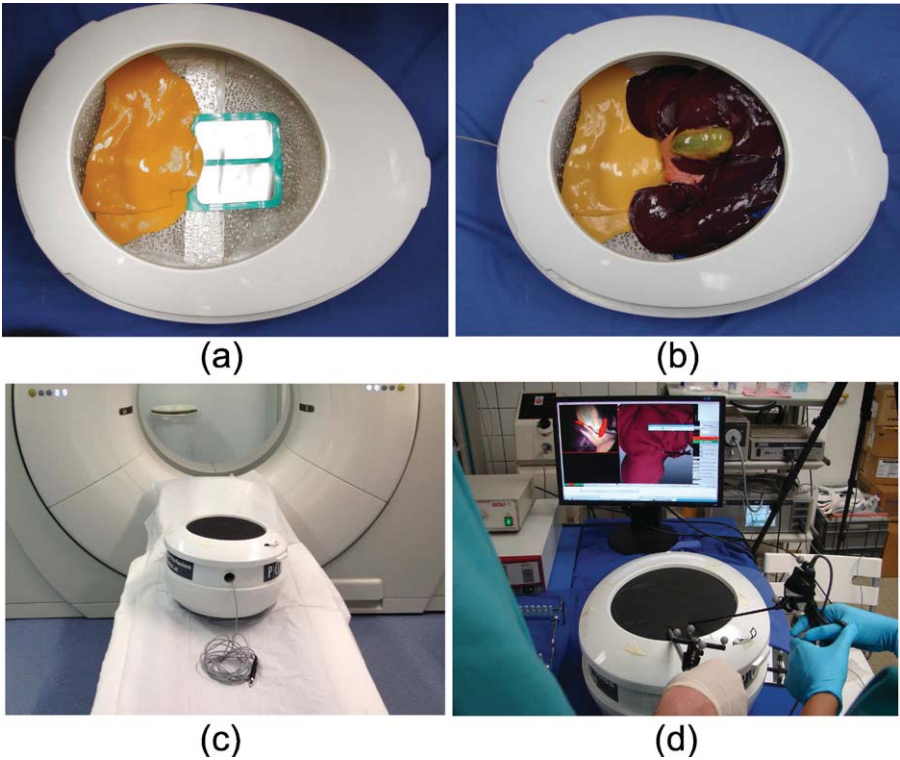


Fig. 5. (a) Underlayment and a common earth, (b) inserted pig liver, (c) CT scanning of the box trainer, and (d) execution of the experiment.

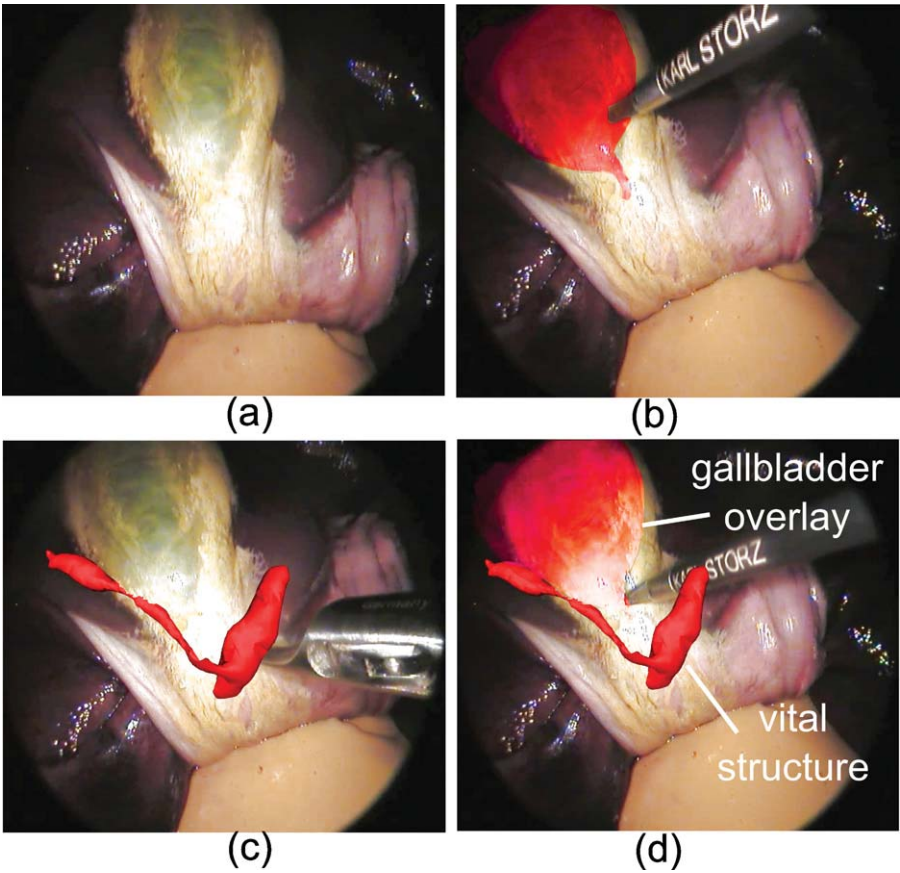


Fig. 6. (a) Unaugmented view on the liver, (b) display of the gallbladder, (c) display of risk structures, and (d) visualization of both structures.

not account for soft tissue deformation, all distance measurements rely on the positions defined by the preoperative CT scans and rigid registrations of the box trainer and the instruments. This is justified by the fixation of the organ using a suture, as described in the experimental setup. The visualizations are shown in Fig. 6. Note that a left and right hepatic duct with early separation are present and the cystic duct inserts in the right hepatic duct as a variant to normal anatomy.

Execution of the surgical procedure. An endoscope, tracked using Polaris, was inserted in the box trainer. With inserted scissors, we initially approached the triangle of Calot to find the location of the risk structures such as the common hepatic duct, the common bile duct and the structure of interest, the cystic duct. Considering the visualization, we dissected the cystic duct. By approaching the scissor to the unclipped cystic duct a warning visualization was displayed. Subsequently the scissors were exchanged with the clip. The clip was moved near to the common hepatic duct which depicts a potential risk situation so that the common hepatic duct was visualized. Then the clip was exchanged for the scissors. They were brought closer to the clipped cystic duct which is not considered risky and therefore was not visualized. The execution of the procedure is illustrated in Fig. 5(d).

Interview. According to the interview conducted after the trial, our system helped to dissect the triangle of Calot quickly and safely, without damaging the common hepatic duct or the common bile duct. The context-aware visualization was described as a very useful tool to support the surgeon by detecting life threatening situations during the operation. Also mentioned was the advantage of the explicit knowledge representation so that surgeons can understand and anticipate the system behavior and thus put more trust in it. The interview confirmed that the AR visualizations help with perception of depth, even in their simple form. Suggestions of photo-realistic rendering additions such as realistic shadows or lighting effects were dismissed, as the simple visualizations add less visual complexity to the scene. The placements of the virtual overlays were still regarded as satisfactory although they only relied on CT-Data and rigid tracking without compensation for soft tissue movements. The visualizations were considered as intuitive, with their meaning and intent easily comprehensible. Overall, it was confirmed that the considered section of the surgery can be performed faster and easier, compared to an execution without assistance. However, as an improvement, it was suggested to reduce the opaqueness of the visualization of risk structures because otherwise visualizations may cover the underlying anatomy completely and thus can complicate surgical evaluation. Also there was a short, yet noticeable delay in the display of the endoscopic images as the system first had to process and augment them. This was somewhat discomforting for the surgeons used to a faster, almost instant display of the images. But overall, the delay was still considered acceptable.

5. Discussion

The evaluation showed the feasibility and use of ontology-based situation interpretation for laparoscopy in a clinical scenario. With a phase recognition rate of over 93%, the recognition process showed satisfactory results in the phantom experiments. Although there is some information loss involved, the discretized computational model in the ABox leads to a much clearer representation of the situation in the OR and allows for the use of logical inference mechanism and discrete reasoning over real-valued data. Most false recognitions occurred for distances that fall into the boundary between two sets, especially between “near” and “medium”. In these cases, the correct predicate was not always chosen which ultimately led to false recognitions. This problem can be solved by

adding an additional set between “near” and “medium” to better capture the transition. Other false recognitions occurred when an appropriate instrument was near a tumor just by chance, without the surgeon actually doing any work on the diseased structure. The system then falsely assumed the structure was being worked on. Future work will include Case Based Reasoning to capture such variations. Due to the nature of the errors, the system is suitably accurate, in the confines of the experimental setup. As for the differences between the three approaches, it can be said that the Bayesian approach with Histograms and the evidence-based approach have the distinct advantage of being able to approximate any distribution. However, they also require more training samples to restrict the numerous degrees of freedom. The Bayesian approach with Gaussians makes assumptions about the underlying distribution, therefore limiting the degrees of freedom as well as its flexibility at the same time. The process to generate training samples can be done in a matter of minutes (less than 3 min in our evaluation) and only takes negligible amounts of disc space (37 kB for the 1244 samples used in the presented evaluation). Therefore, it is feasible to record such samples for each surgeon and customize the system individually.

The membership values are intuitive representations of their respective lingual concepts. As described in [21], the membership values can be regarded as belief values, denoting the degree to which classifications can be trusted. Based on the observed samples, high values for a given measurement x mark high confidence that x actually belongs to the fuzzy set at hand. Low values mean that there is no or little reason to believe so. It has to be noted that this is not equivalent to the statement that x does not belong to the set. It rather states that the system has not enough information to make the distinction. This behavior of the algorithm is desirable since it allows for safe classifications with a high degree of confidence. The system can also be implemented in a way that refuses to map measurements to predicates if the belief values are too low. This event occurs if the given value falls in a range which is not sufficiently represented in the training set.

Surgeons found the visualizations to be sufficiently clear as not to distract them from their work. Due to the context-awareness, visualizations only appear when needed and are inactive most of the time. This allows the surgeon to operate freely, while still receiving assistance when necessary.

In this work, we only considered phases which can be recognized using distance measurements. Future work will focus on expanding the system to use further, more sophisticated sensors. Integration of image processing algorithms to analyze endoscope images in order to recognize instruments or events like bleedings is planned, as well as the integration of the status of surgical devices, e.g. Thermoflator or Endomat. Also preoperative measurements from medical imaging devices are to be considered and integrated into the system. With the addition of new sensors, more phases can be recognized and further assistance can be granted to the surgeon. While the basic structure and the underlying algorithms will stay the same, the system’s usefulness and ability to understand and assist in more complex situations will grow. The goal is to use all information that is available, while relaying only a minimal, yet relevant and sufficient fraction of it to the surgeon.

In the case of the cholecystectomy using the pig liver, the organ was fixed and sewed to the top of the box trainer and therefore widely immobile. In doing so, we were able to largely avoid problems introduced by soft tissue movement. While such operating techniques are not always an option, the evaluation shows the possibility of obtaining enough situational information to provide context-aware assistance, such as approximate distances, even with the soft tissue deformation problem still being unsolved. For initial registrations of preoperative images from CT or MRI, we seek to employ our in-house developed methods, as described in

[25,26]. These types of algorithms will be used to generate input for the situation interpretation in various interventions, even when the soft tissue problem cannot be dealt with using specialized operating techniques. Furthermore, by using soft computing methods, namely fuzzy logics, to analyze the incoming data, we have achieved enough robustness to produce useable result even with distorted distance information provided by imperfect soft tissue tracking. In addition to registering techniques, methods described in [27] can be used to identify instruments and surgical activities via image analysis algorithms.

6. Conclusion

We presented a system which offers context-aware AR visualization for laparoscopic surgeries. By mapping numerical measurement to logical predicates, we created a computational model of the current scene, described as an ABox in OWL. The model was interpreted in real-time to recognize certain phases which occur during operations. The system was evaluated both in a phantom experiment and in a cholecystectomy using a pig liver in a surgical simulator, showing an overall recognition rate of over 93%. However, the training and evaluation data originated from the same person. It would be interesting to evaluate the differences between several surgeons to see whether customized fuzzy sets are necessary for individual surgeons or if it is possible to use a single representation for everyone. Also, the surgical phases considered in this study are rather simple and lend themselves to be expressed using rules. Further research will focus on more complex phases and extensions to our systems necessary to accommodate them.

By combining learning mechanisms and rule-based systems, the problem was solved with limited need for labeled training samples and simple rules. Due to machine learning techniques, the difficulty of extracting tacit knowledge from experts concerning the interpretation and classification of real-valued measurements was avoided.

Our context-aware system acts as an intelligent assistant with a certain kind of understanding for the situation and the medical background of the procedure. The surgeon should no longer be obliged to take care of the AR system but rather the system adapts its behavior to him. This high degree of integration into existing workflows allows surgeons to focus on the intervention itself and less on the management of the assistance system. We believe that this is an important step in raising acceptance for AR systems and ultimately bringing them into clinical application.

Acknowledgements

The present research was conducted within the setting of the “Research training group 1126: Intelligent Surgery – Development of new computer-based methods for the future workplace in surgery” and the “SFB TRR 125 Cognition-Guided Surgery” founded by the German Research Foundation. It is furthermore sponsored by the European Social Fund of the State Baden-Wuerttemberg. We would like to thank the Division of Medical and Biological Informatics at the German Cancer Research Center for providing virtual liver models.

References

- [1] Neumann B, Moeller R. On scene interpretation with description logics. *Image Vis Comput* 2008;26(1):82–101.

- [2] Navab N, Traub J, Sielhorst T, Feuerstein M, Bichlmeier C. Action- and workflow-driven augmented reality for computer-aided medical procedures. *IEEE Comput Graph Appl* 2007;27(5):10–4.
- [3] Katic D, Sudra G, Speidel S, Castrillon-Oberndorfer G, Eggers G, Dillmann R. Knowledge-based situation interpretation for context-aware augmented reality in dental implant surgery. *Med Imaging Augment Real Lecture Notes Comput Sci* 2010;6326:531–40.
- [4] Nicolau S, Soler L, Mutter D, Marescaux J. Augmented reality in laparoscopic surgical oncology. *Surg Oncol* 2011;20(3):189–201.
- [5] Baumhauer M, Feuerstein M, Meinzer HP, Rassweiler J. Navigation in endoscopic soft tissue surgery: perspectives and limitations. *J Endourol* 2008;22(4):751–66.
- [6] Mirota DJ, Ishii M, Hager GD. Vision-based navigation in image-guided interventions. *Annu Rev Biomed Eng* 2001;13:297–319.
- [7] Padoy N, Hager GD. Human-machine collaborative surgery using learned models. *IEEE Int Conf Robot Autom* 2012;528:5–529, 2.
- [8] Kragic D, Hager G. Task modeling and specification for modular sensory based human-machine cooperative systems. *Proc IEEE/RSJ Int Conf Intell Robots Syst* 2003;4:3192–7.
- [9] Yoshimitsu K, Miyawaki F, Vain J, Fukui Y, Hashimoto D, Masamune K. Surgical model based automatic instrument inserting control of scrub nurse robot for endoscopic and laparoscopic surgery. In: MICCAI workshop on modeling and monitoring of computer assisted interventions; 2009.
- [10] Lin HC, Shafraan I, Yuh D, Hager GD. Towards automatic skill evaluation: detection and segmentation of robot-assisted surgical motions. *Comput Aided Surg* 2006;11(5):220–30.
- [11] Reiley C, Lin H, Yuh D, Hager G. A review of methods for objective surgical skill evaluation. *Surg Endosc* 2010;25(2):356–66.
- [12] Blum T, Feußner H, Navab N. Modeling and segmentation of surgical workflow from laparoscopic video. *Med Image Comput Comput Assist Intervention* 2010;13(3):400–7.
- [13] Padoy N, Blum T, Ahmadi A, Feußner H, Berger MO, Navab N. Statistical modeling and recognition of surgical workflow. *Med Image Anal* 2010;16(3):632–41.
- [14] Lalys F, Riffaud L, Morandi X, Jannin P. Surgical phases detection from microscope videos by combining SVM and HMM. In: Proceedings of the international MICCAI conference on Medical computer vision: recognition techniques and applications in medical imaging. 2011. p. 54–62.
- [15] Bouarfa L, Jonkera PP, Dankelmana J. Discovery of high-level tasks in the operating room. *J Biomed Inform* 2010;44(3):455–62.
- [16] Rosen J, Brown J, Chang L, Sinanan M, Hannaford B. Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete markov model. *IEEE Trans Biomed Eng* 2006;53(3):399–413.
- [17] Jannin P, Morandi X. Surgical models for computer-assisted neurosurgery. *Neuroimage* 2007;3(37):783–91.
- [18] Neumuth T, Jannin P, Schlomberg J, Meixensberger J, Wiedemann P, Burgert O. Analysis of surgical intervention populations using generic surgical process models. *Int J Comput Assist Radiol Surg* 2010;6(1):59–71.
- [19] Baader F, Calvanese D, McGuinness DL, Nardi D, Patel-Schneider PF. The description logic handbook: theory implementation applications. Cambridge, England: Cambridge University Press; 2003.
- [20] Hitzler P, Krötzsch M, Parsia B, Patel-Schneider PF, Rudolph S, editors. OWL 2 Web Ontology Language Primer. See <http://www.w3.org/TR/owl2-primer/>. W3C Working Draft; 2009.
- [21] Weisbrod J. Unscharfes Schliessen. Heidelberg, Germany: Akademische Verlagsgesellschaft AKA GmbH; 1996.
- [22] Sirin E, Parsia B, Grau B, Kalyanpur A, Katz Y. Pellet: a practical OWL-DL reasoned, web semantics: science, services and agents on the world wide web. *J Web Semant* 2007;5(2):51–3.
- [23] Tudorache T, Vendetti J, Noy N. Web-Protégé: a lightweight owl ontology editor for the web. In: Fifth international workshop of OWL: experiences and directions. 2008.
- [24] Wolf I, Vetter M, Wegner I, Böttger T, Nolden M, Schöbinger M, et al. The medical imaging interaction toolkit. *Med Image Anal* 2005;9(6):594–604.
- [25] Röhl S, Bodenstedt S, Suwelack S, Kenngott H, Müller-Stich BP, Dillmann R, et al. Dense GPU-enhanced surface reconstruction from stereo endoscopic images for intraoperative registration. *Med Phys* 2012;39(3):1632–45.
- [26] Suwelack S, Roehl S, Dillmann R, Wekerle AL, Kenngott H, Müller-Stich BP, et al. Quadratic corotated finite elements for real-time soft tissue registration. In: MICCAI workshop: computational biomechanics for medicine; 2011:39–50.
- [27] Speidel S, Benzko J, Sudra G, Azad P, Müller-Stich BP, Gutt C, et al. Automatic classification of minimally invasive instruments based on endoscopic image sequences. In: Proceedings of SPIE medical imaging: visualization, image-guided procedures, and modeling. 2009. p. 7261.