

Expert Evaluation of a Spoken Dialogue System in a Clinical Operating Room

Juliana Miehle¹, Nadine Gerstenlauer¹, Daniel Ostler², Hubertus Feußner²,
Wolfgang Minker¹, Stefan Ultes³

¹Institute of Communications Engineering, Ulm University, Albert-Einstein-Allee 43, 89081 Ulm, Germany

²Minimally-invasive Interdisciplinary Therapeutical Intervention, Klinikum rechts der Isar,
Technical University of Munich, Trogerstr. 26, 81675 München, Germany

³Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK
juliana.miehle@uni-ulm.de, nadine.gerstenlauer@uni-ulm.de, daniel.ostler@tum.de,
hubertus.feussner@tum.de, wolfgang.minker@uni-ulm.de, su259@cam.ac.uk

Abstract

With the emergence of new technologies, the surgical working environment becomes increasingly complex and comprises many medical devices which have to be monitored and controlled. With the aim of improving productivity and reducing the workload for the operating staff, we have developed an Intelligent Digital Assistant for Clinical Operating Rooms (IDACO) which allows the surgeon to control the operating room using natural spoken language. As speech is the modality used by the surgeon to communicate with their staff, using it to control the technical devices does not pose an additional mental burden. Therefore, we claim that the surgical environment presents a potential field of application for Spoken Dialogue Systems. In this work, we present the design and implementation of IDACO as well as the evaluation in an experimental set-up by specialists in the field of minimally invasive surgery. Our expert evaluation yields promising results and allows to conclude that clinical operating rooms are indeed an expedient area of application for Spoken Dialogue Systems.

Keywords: Intelligent Assistance, Natural Spoken Language, Surgical Environment

1. Introduction

Finding new technological solutions in order to enhance the work in clinical operating rooms has been in the focus of research for many years. However, with the emergence of new technologies, the surgical working environment comprises nowadays many medical devices which have to be monitored and controlled and thus becomes increasingly complex. Therefore, new strategies are needed to keep the working environment manageable and to reduce the workload for the surgical team, thus allowing them to fully focus on the actual surgical procedure. In this context, the *Operating Room of the Future* is a keyword often used (Feußner, 2003). It describes the application of new technologies such as computer-enhanced systems to create an intelligent operating room (OR) that facilitates work and reduces the staff needed during a surgical intervention (Pransky, 2001). This reduces personnel cost and promises to lessen the rate of avoidable incidents caused by human error.

However, having computer systems that monitor the surgical devices is not sufficient. An intelligent OR also needs an intelligent human-computer interface as the surgeon as well as the whole surgery team must not be disturbed by the usage of complex computer applications and numerous devices. Therefore, the interface has to be designed so as to be simple and intuitive (Feußner et al., 2014). During a procedure, the surgeon needs his hands to operate on the patient and his eyes for being aware of what he is doing. Thus, any graphical and gesture-based systems are not well suited for this purpose. Being hands- and eyes-free, speech as an input and output modality seems to be a good choice. Moreover, it is the modality used by the surgeon to communicate with their staff. Thus, using speech to control the technical devices does not pose an additional mental burden. The surgeon can focus on the surgery and control the technical environment at the same time without taking care

of how to interact with the system. Therefore, we claim that the surgical environment presents a potential field of application for Spoken Dialogue Systems.

Until now, the use of voice interaction systems in clinical operating rooms has only been scarcely researched. Most research projects focus on visualisation and intraoperative imaging, telesurgery and robotic surgery as well as educational features (Bharathan et al., 2013). Moreover, there exist some projects about workflow modelling (Sutherland and van den Heuvel, 2006; Padoy, 2010; Katić et al., 2014; Kranzfelder et al., 2011; Kranzfelder et al., 2013; Kranzfelder et al., 2014; Agarwal et al., 2007). To the best of our knowledge, there exists only one speech-enabled operation assistant. HERMES built by Computer Motion Inc. (Roe and Wang, 2000) connects all devices in the OR over a central network. Then, the scrub nurse has the possibility to control all instruments and devices over a central touch screen instead of numerous control panels. Moreover, HERMES provides the surgeon a speech control over several devices. Feedback to the surgeon's request is shown on the endoscopic video screen and the system provides audio feedback if one of the controlled devices loses power or gets disconnected from the network. However, even though the use of voice commands was well received among surgeons and OR staff, HERMES implements an audio-visual interface and pure speech interaction is not supported.

Overall, the existing applications mostly concentrate on workflow modelling or single smart features rather than building up a complete intelligent interface. In contrast, our aim is to develop and evaluate an intelligent spoken language operation assistant offering several functionalities in order to provide the surgeon assistance in many different situations before and during an ongoing procedure. The work described in this paper builds upon and extends work published in (Miehle et al., 2017b; Miehle et al., 2017a; Gerstenlauer, 2017).

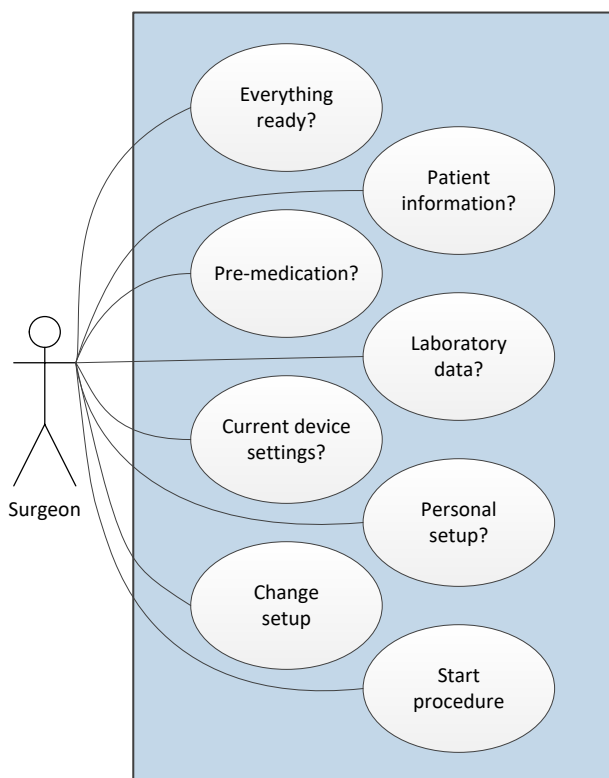


Figure 1: Before the operation starts, IDACO supports the surgeon during the team time-out and provides all necessary data.

The structure of the paper is as follows: In Section 2., the design and implementation of our system are described. The evaluation of IDACO is presented in Section 3., before concluding in Section 4.

2. Intelligent Digital Assistance for Clinical Operating Rooms

In order to increase productivity and reduce the workload for the operating staff, our system acts cooperatively and supports the surgeon autonomously during the surgery. IDACO escorts the surgery team throughout the entire procedure and provides assistance where necessary. The main functionalities of the presented speech-based assistant for clinical operating rooms have been identified during a requirement analysis with specialists in the field of minimally invasive surgery. They are described in the next section.

2.1. Functionalities

During the requirement analysis, we have identified two major parts where intelligent assistance is considered desirable by the medical specialists: the pre-procedural and the procedural part. For both parts, we determined several functionalities which should be taken on by IDACO.

Before the operation starts, the surgery team uses to carry out the team time-out in order to avoid any kind of confusion concerning the patient and the upcoming surgery. In the course of this procedure, the patient as well as the surgical team are identified and the surgery type is checked. IDACO supports the surgeon during the team time-out and provides data about the surgery type and the operating team

as well as the patient, including pre-diseases, medical treatment and laboratory data. The variety of information the user can request from the system in this mixed-initiative dialogue part is illustrated in Figure 1.

Afterwards, the surgeon starts with the operation and IDACO escorts the team throughout the entire surgery. The system tracks the usage of surgical instruments and material (e.g. trocars, different types of clips, suturing material) by constantly listening to the operating surgeon and compares the thereby observed course of the procedure to the predicted surgical workflow. If the usage differs, it reacts proactively and utters a warning via speech. Moreover, IDACO controls surgical devices automatically at the right time of the operation when the surgeon confirms the execution. For example, the presented system is able to start the insufflator, increase the gas insufflation, turn off and on the light and tilt the table. In order to operate the devices autonomously, the system saves the preferred device settings for each surgeon and transmits the parameters to the surgical devices (e.g. OR table, room light, insufflator, suction and irrigation unit). Moreover, it allows the surgeon to retrieve and change their pre-settings.

For unseen incidents, IDACO encompasses an emergency mode which includes a "silent mode" to prevent further distractions by the system. This mode may be activated by the surgeon at each arbitrary point of the operation. Having it activated, IDACO does not track the usage of surgery material until it is explicitly told to end this silent mode. Meanwhile, the surgeon can operate without being disturbed by the system. Furthermore, the system offers starting the emergency mode proactively in delicate situations.

2.2. Challenges

Enabling an intelligent operating assistance system to follow a surgery and control surgical devices automatically bears several challenges.

For keeping track of the procedure and automatically controlling surgical devices, the system needs to know when to perform which action on which device and when to stay in the background. Therefore, it has to be aware of the whole context of the surgery, i.e. the current point of the procedure and all past and future actions. This means that a reliable method for tracking the course of the surgery needs to be developed, thus allowing to detect unscheduled events. Moreover, it has to be clearly defined how the system is supposed to react in tenuous situations. For this purpose, standardized surgeries need to be described in detail, allowing the system to compare the actual course of the procedure to the schedule (Feußner and Wilhelm, 2016). Using this medical domain knowledge, exact models of the complex surgery structure need to be created which are then applied to the voice interaction system. Additionally, an interface needs to be designed and implemented which allows intercommunication between the voice interaction system and the surgical devices as well as the clinical information system. Moreover, with respect to patient safety, appropriate strategies need to be defined in order to maintain full control of the medical devices even if IDACO is allowed to perform some pre-defined actions during the surgery and control devices automatically.

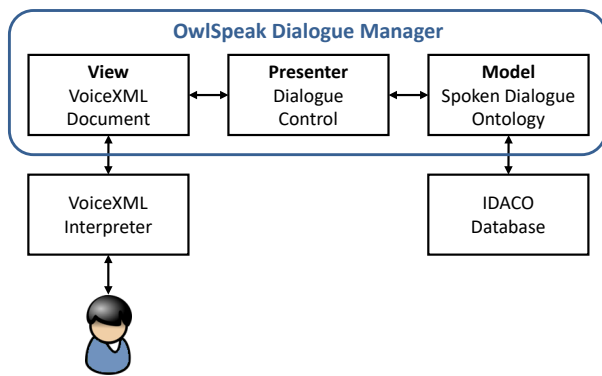


Figure 2: The overall architecture of IDACO comprising OwlSpeak, a VoiceXML interpreter and the IDACO Database.

2.3. Dialogue Modelling

For the dialogue modelling, we used the ontology-based Dialogue Management System OwlSpeak developed by Heinroth et al. (2010) and further extended by Ultes and Minker (2014). The overall architecture of our system can be seen in Figure 2. OwlSpeak is based on the model-view-presenter design pattern (Potel, 1996) which allows to separate data management (model), dialogue interface (view) and dialogue logic (presenter). The model is thereby implemented in the form of a Spoken Dialogue Ontology using the Web Ontology Language (OWL) (Antoniou and Van Harmelen, 2004) which contains the description of the dialogue and the current dialogue state. The view is represented by a VoiceXML document which is dynamically created at each dialogue turn and the presenter comprises the dialogue control logic. As OwlSpeak provides a new VoiceXML document at each turn, a VoiceXML interpreter by Voxeo¹ has been integrated. Moreover, OwlSpeak has been connected to the IDACO Database which acts as the interface between the Dialogue Manager and the Clinical Information System as well as the surgical devices. This database is described in detail in Section 2.4.

As a first prototype, we modelled the dialogue flow for a laparoscopic cholecystectomy. The dialogue is thereby divided into two parts: the pre-procedural and the procedural part. For the entire dialogue, the system utterances are concise and direct, meaning that the requested information is output very concretely and without any additional information that might be inappropriate.

During the pre-procedural part, IDACO supports the surgeon during the team time-out and provides data about the surgery type, the operating team and the patient, including pre-diseases, medical treatment and laboratory data. Moreover, the surgeon has the possibility to ask the system to adopt the devices in use to his personal pre-settings. He can also ask IDACO to read out the current device settings or his personal pre-settings stored in the IDACO Database and to change his personal set-up for a specific device. The user has the dialogue initiative and thus the full control of how to start the interaction. Hence, the surgeon can decide on which information should be provided and

which device settings should be changed or adopted. If he does not want to request any information, he can even skip the information-providing pre-procedural part and inform IDACO to start with the procedure immediately.

When the surgeon confirms to start with the surgery, the second part of the dialogue begins. In contrast to the first part which is very flexible and allows the user to control the dialogue, the procedural part follows an exact surgery schedule which has been modelled in the Spoken Dialogue Ontology. Keeping track of the surgery is thereby done by tracking the tool usage as described by Padoy (2010). Therefore, we introduced variables for all kinds of instruments which are used and all assistance actions which are performed during this specific surgery. The system listens to each of the surgeon's instructions and increments the variables after each user utterance corresponding to its specific purpose. The grammars that describe the range the system is able to understand are thereby kept flexible in order to avoid the need for command-based user inputs. In contrast, the operating surgeon does not need to concentrate on how to interact with IDACO but can just talk to the surgical staff and the system reacts to keywords used by the surgeon. The workflow and hence the current part of the operation are then derived from the history of used tools at any point of the surgical intervention. The observed course of the procedure is compared to the surgery schedule which has been modelled in the Spoken Dialogue Ontology. In case of a deviation from the regular course, the system reacts proactively and utters a warning. The surgeon can then correct the amount of used material or tell the system that the expected usage has to be adapted for the rest of the procedure.

Moreover, the surgery schedule which has been modelled in the Spoken Dialogue Ontology defines at which point of the operation IDACO should perform a control action on medical devices such as starting the insufflator, increasing the gas insufflation, turning off and on the light and tilting the table. Whenever the system recognises such a time for action, it asks the surgeon whether the action should be performed. The surgeon can then confirm the execution or tell IDACO to wait until he explicitly utters to do so. This allows the surgeon to maintain full control of the medical devices, thus ensuring the patient safety. For the emergency mode, we introduced an Agenda² without any system move and only one possible user move which is the user giving the command to deactivate this mode.

The pre-operational part of the system can be reused for various kinds of surgical interventions. Only the variables storing the device settings and the personal device set-up need to be adjusted to the specific operation type what can be easily done using the IDACO Database. The actual procedural part uses an exact model of the complex surgery structure. Hence, only features that do not refer to any surgery specific data like the emergency mode can be reused for different surgical interventions but not the implemented course of the procedure.

¹<https://evolution.voxeo.com/>

²Concept used by OwlSpeak to bundle several moves that belong to a specific dialogue turn (Ultes and Minker, 2014).

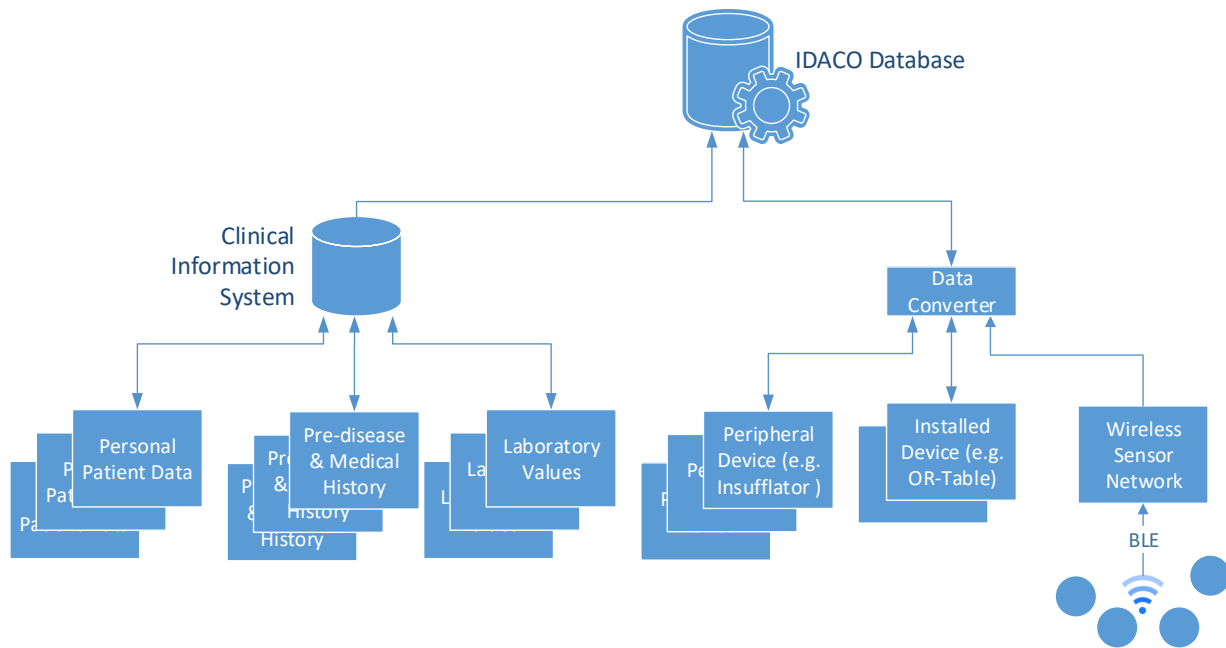


Figure 3: The schematic overview of the IDACO Database containing the mirrored patient- and intervention-specific information as well as the collected device and sensor data.

2.4. The IDACO Database

As an interface between the OwlSpeak Dialogue Manager and the Clinical Information System as well as the surgical devices, we have implemented a database which allows accessing necessary data and facilitates controlling of the surgical devices. A schematic overview of the IDACO Database is depicted in Figure 3.

The central point of the database is a table containing the intervention record from which all currently required data is concluded. Due to privacy issues, the system has no direct access to the Clinical Information System. However, the IDACO Database contains a mirrored image of all relevant information about the intervention. This covers patient data including personal data, pre-diseases, medical treatment and laboratory values, as well as data about the surgery type and the operating team. The table about the operating team contains all staff members who attend the current surgery, their name, their role and their experience. Moreover, as the system ought to store and adjust the surgeon's preferences, there is an additional database table containing the personal set-ups for each surgeon and each surgery type. One set-up contains all devices needed for the underlying operation as well as the corresponding parameters as pre-sets.

Furthermore, the IDACO Database provides an interface in order to control the surgical devices in the OR. Therefore, the database contains a list of all existing devices as well as the corresponding device parameters. After the confirmation by the surgeon, IDACO is able to set target values for the device parameters defined in the IDACO Database. Then, the data converter serves as interface to the specific device protocols of the peripheral and installed devices and changes the current parameters of the devices accord-

ingly. Thus, IDACO is able to automatically control several devices at the right time of the procedure. Moreover, IDACO derives the information about which persons are currently present in the operating room from a Bluetooth-Low-Energy (BLE) tracking system.

3. Evaluation

The system was implemented in an experimental set-up in order to get an expert evaluation from medical specialists. As a first prototype, we have modelled the dialogue flow for a laparoscopic cholecystectomy. According to Cuschieri (1999), this is the gold standard for the treatment of gallstones and a highly standardized surgical procedure which can be segmented into ten procedural tasks:

1. Insertion of a Veress needle
2. Creation of Pneumoperitoneum
3. Insertion of Ports
4. Initial Inspection and Exposure of the Triangle of Calot
5. Dissection of Cystic Pedicle
6. Operative Cholangiogram
7. Closure of Cystic Duct and Detachment of Gallbladder from Liver
8. Extraction of Gallbladder
9. Final Inspection
10. Closure of Port Wounds

Each procedural part comprises several steps which are directly linked with the usage of certain material and instruments. As these instruments and materials, which are necessary to perform each procedural task, are clearly defined, it is possible to predict the surgeon's utterances during each step. For our prototype, we used the knowledge about these utterances in order to define a dialogue corresponding to the surgical workflow.

For example, the first procedural task of the laparoscopic cholecystectomy, the insertion of a Veress needle, comprises four steps:

1. Incision with a scalpel
2. Application of the first Backhaus clip
3. Application of the second Backhaus clip
4. Insertion of the Veress needle

Hence, IDACO needs to track "scalpel", "Backhaus clip", "Backhaus clip" and "Veress needle" in order to complete the first part of the surgery and to move on to the second procedural task which is the creation of the Pneumoperitoneum. In order to begin with this part, the gas insufflation needs to be started. Therefore, IDACO asks the surgeon whether this action should be performed. The surgeon can then confirm the execution or tell IDACO to wait until he explicitly utters to do so. The resulting dialogue excerpt looks as follows:

SURGEON: Scalpel, please.

SURGEON: Backhaus clip.

SURGEON: Please give me another Backhaus clip.

SURGEON: Now the Veress needle.

IDACO: Shall I start the gas insufflation?

SURGEON: Yes, please.

IDACO: The gas insufflation has been started.

The complexity of the dialogue increases with the complexity of the surgery structure. In total, our dialogue comprises 126 Agendas and 96 Variables which are used to control 18 installed and peripheral devices and to read out all necessary data from the Clinical Information System.

The implemented prototype was then evaluated by specialists in the field of minimally invasive surgery where it received good feedback. The speech interface and the dialogue were perceived very positively. IDACO is designed not to annoy the surgeon and the operating staff with inappropriate behaviour and unnecessary system prompts during the surgery. Therefore, the system utterances are concise and direct. This communication style has been assessed well-suited for the underlying dialogue scenario in an OR. Moreover, IDACO stays in the background if the procedure goes as scheduled. However, the evaluation with the experienced physicians indicated that passive system behaviour makes the surgery team insecure. The operating staff prefers to get any kind of feedback on what the system recognises and understands. The team of surgical specialists therefore suggested to equip the scrub nurse with a tablet PC showing each user utterance the system receives.

In case of a misunderstanding by the voice interaction system, the nurse could then correct the recognised input. In doing so, the nurse has the possibility to observe the system behaviour and correct speech recognition errors manually.

Additionally, our evaluation pointed out that a close cooperation between system developers and medical scientists is inevitable in the design and implementation of intelligent systems for clinical operating rooms. Due to the highly specific field of application, it is hard for non-specialists to decide on optimal system behaviour.

4. Conclusion and Future Directions

We have presented the design, implementation and expert evaluation of a speech-based assistant for clinical operating rooms which supports the surgeon and their operating staff before and during a surgery. To the best of our knowledge, the presented system is the first intelligent spoken language operation assistant putting together several technologies in an OR and allowing to control the technical devices using speech. Enabling the operating surgeon to control devices inherent to the OR by himself as well as autonomously controlling some of the surgical devices reduces the workload for the surgical team as well as the amount of staff needed to assist during a surgical intervention and promises to lessen the rate of avoidable incidents caused by human error. Using IDACO, the surgeon can focus on the surgery and control the technical environment at the same time without taking care of how to interact with the system as speech is the modality used by the surgeon to communicate with their staff and IDACO just listens to the surgeon's utterances. Hence, it does not pose an additional mental burden on the surgical staff.

Our evaluation by specialists in the field of minimally invasive surgery showed that the system perceived very positive expert feedback. The only issue which remains open to debate is how the system should give feedback to the surgery team as, on the one hand, the surgeon should not be annoyed by unnecessary system prompts, but on the other hand, completely passive system behaviour makes the surgery team insecure. However, the speech interface, the dialogue and the communication style were assessed positively which leads us to the conclusion that we can confirm our claim that the surgical environment presents a field of application for Spoken Dialogue Systems.

In future work, the remaining issue of appropriate system feedback needs to be resolved. Afterwards, a broader evaluation needs to be done in order to get quantitative results regarding the performance. Moreover, a generic method for modelling the surgery control needs to be developed. For the presented system, we modelled exemplarily the procedure of a laparoscopic cholecystectomy. However, the multitude of existing surgical procedures makes it impossible to implement each one individually.

5. Bibliographical References

Agarwal, S., Joshi, A., Finin, T., Yesha, Y., and Ganous, T. (2007). A Pervasive Computing System for the Operating Room of the Future. *Mobile Networks and Applications*, 12(2-3):215-228.

- Antoniou, G. and Van Harmelen, F. (2004). Web ontology language: Owl. In *Handbook on ontologies*, pages 67–92. Springer.
- Bharathan, R., Aggarwal, R., and Darzi, A. (2013). Operating room of the future. *Best practice & research. Clinical obstetrics & gynaecology*, 27(3):311–322.
- Cuschieri, A. (1999). Laparoscopic cholecystectomy. *Journal of the Royal College of Surgeons of Edinburgh*, 44(3):187–192.
- Feußner, H. and Wilhelm, D. (2016). Minimalinvasive chirurgie und "robotic surgery": Chirurgie 4.0? *Der Chirurg*, 87(3):189–194.
- Feußner, H., Reiser, S. B., Bauer, M., Kranzfelder, M., Schirren, R., Kleeff, J., and Wilhelm, D. (2014). Technische und digitale weiterentwicklung in der laparoskopischen/offenen chirurgie. *Der Chirurg*, 85(3):178–185.
- Feußner, H. (2003). The operating room of the future: a view from europe. In *Seminars in Laparoscopic surgery*, volume 10, pages 149–156. Sage Publications Sage CA: Thousand Oaks, CA.
- Gerstenlauer, N. (2017). Controlling an intelligent operating room with spoken language.
- Heinroth, T., Denich, D., and Schmitt, A. (2010). OwlSpeak - Adaptive Spoken Dialogue within Intelligent Environments. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2010 8th IEEE International Conference on*, pages 666–671. IEEE.
- Katić, D., Wekerle, A.-L., Gärtner, F., Kenngott, H., Müller-Stich, B. P., Dillmann, R., and Speidel, S. (2014). Knowledge-Driven Formalization of Laparoscopic Surgeries for Rule-Based Intraoperative Context-Aware Assistance. In David Hutchison et al., editors, *Information Processing in Computer-Assisted Interventions*, volume 8498 of *Lecture Notes in Computer Science*, pages 158–167, Cham. Springer International Publishing.
- Kranzfelder, M., Schneider, A., Gillen, S., and Feußner, H. (2011). New technologies for information retrieval to achieve situational awareness and higher patient safety in the surgical operating room: the MRI institutional approach and review of the literature. *Surgical endoscopy*, 25(3):696–705.
- Kranzfelder, M., Schneider, A., Fiolka, A., Schwan, E., Gillen, S., Wilhelm, D., Schirren, R., Reiser, S., Jensen, B., and Feußner, H. (2013). Real-time instrument detection in minimally invasive surgery using radiofrequency identification technology. *The Journal of surgical research*, 185(2):704–710.
- Kranzfelder, M., Schneider, A., Fiolka, A., Koller, S., Reiser, S., Vogel, T., Wilhelm, D., and Feußner, H. (2014). Reliability of sensor-based real-time workflow recognition in laparoscopic cholecystectomy. *International journal of computer assisted radiology and surgery*, 9(6):941–948.
- Miehle, J., Gerstenlauer, N., Ostler, D., Feußner, H., Minker, W., and Ultes, S. (2017a). An intelligent digital assistant for clinical operating rooms. In *Proceedings of the 21st Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL)*, pages 164–165, August.
- Miehle, J., Ostler, D., Gerstenlauer, N., and Minker, W. (2017b). The next step: intelligent digital assistance for clinical operating rooms. *Innovative Surgical Sciences*, 2(3):159–161.
- Padoy, N. (2010). *Workflow and Activity Modeling for Monitoring Surgical Procedures*. Dissertation, Université Henri Poincaré, Nancy.
- Potel, M. (1996). Mvp: Model-view-presenter the taligent programming model for c++ and java. *Taligent Inc*, page 20.
- Pransky, J. (2001). An intelligent operating room of the future—an interview with the university of california los angeles medical center. *Industrial Robot: An International Journal*, 28(5):376–382.
- Roe, D. B. and Wang, Y. (2000). A voice-controlled network for universal control of devices in the OR. *Minimally invasive therapy & allied technologies : MITAT : official journal of the Society for Minimally Invasive Therapy*, 9(3-4):185–191.
- Sutherland, J. and van den Heuvel, W. J. (2006). Towards an intelligent hospital environment: adaptive workflow in the or of the future. In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*, volume 5, pages 100b–100b. IEEE.
- Ultes, S. and Minker, W. (2014). Managing adaptive spoken dialogue for intelligent environments. *Journal of Ambient Intelligence and Smart Environments*, 6(5):523–539.