ORIGINAL ARTICLE

Intra-operative surgical instrument usage detection on a multi-sensor table

Bernhard Glaser · Stefan Dänzer · Thomas Neumuth

Abstract

Purpose Automatic detection, classification and recording of operating room (OR) activities in real time during a surgical procedure requires a wide range of sensors to gather information on the activities of the surgeon and staff, the patient, and the OR equipment. The surgical instrument currently being used is an important parameter used to assess the instantaneous operating room status. An automated system was developed that detects unmodified surgical instruments in real time using a sensor-based table.

Methods A multi-sensor operating room table was designed featuring a 2D camera, digital scale, and infrared camera. Software was developed to detect and record the sequence of changes on the table during a surgical procedure. The detection rates were evaluated under laboratory conditions by recording the observed instrument usage in 27 functional endoscopic sinus surgeries (FESS).

Results The detection rate achieved using video-based detection alone was 84.9 %. The total detection rate achieved with the combined approach using both video- and weight-based information was 90.3 %.

Conclusion A multi-sensor table-based automated instrument tracking system was developed that provides a foundation for the intra-operative detection of surgical instruments without modifying the instruments in the surgical tray. This system was tested and found to satisfy clinical FESS requirements with a reasonable accuracy. This system may also be useful to improve patient safety, for example to prevent instruments being left in the patient.

B. Glaser (✉) · S. Dänzer · T. Neumuth
Innovation Center Computer Assisted Surgery,
University of Leipzig, Semmelweisstr. 14, 04103 Leipzig, Germany
e-mail: bernhard.glaser@medizin.uni-leipzig.de

Introduction

New strategies for intelligent automation in healthcare organizations are needed to save time and resources, accelerate throughput, enhance patient safety, and improve outcomes [1]. Considering that the operating room (OR) is the most expensive unit of a hospital [2], an optimization process may both save money and increase the quality of patient care [3]. Here, OR activities are addressed as a process to address their specific needs and possibilities. The generation [4], analysis [5], and application of models to the OR as a working environment play a key role in the OR of the future [6].

Early descriptions of surgical workflows were generated via human observation [7]. The goal is now to automate these observations and generate on-the-fly information of the current events. In addition to retrieving information on the functional state of peripheral devices and systems, a continuous, automated instrument surveillance system would provide essential information on the actual state of the surgical procedure [8].

Various approaches to the automatic detection and identification of surgical instruments have been developed. Most methods concentrate on laparoscopic surgery and exploit circumstances prevalent to this special case. The laparoscopic video images used for instrument tracking typically have the anatomic structures in better focus than the instrument itself.

In 1995, Uecker et al. [9] developed an image analysis and tracking algorithm to automate instrument localization and scope maneuvering for robot-assisted laparoscopic surgery.

Speidel et al. [10] presented an approach for classifying minimally invasive instruments using endoscopic images based on preliminary instrument tracking studies. The instruments were not modified with markers. This system segmented the instruments in the current laparoscopic image and recognized the instrument type based on three-dimensional models [11].

Voros et al. [12,13] tracked instruments in laparoscopic camera images by measuring their insertion point into the abdominal cavity. However, no instrument identification was performed.

Sznitman et al. [14] developed a unified instrument detection and tracking framework for retinal microsurgery.

Tonet et al. [15] localized endoscopic instruments in videos using a colored strip on the distal part of the instrument. Allen et al. [16] tracked the motion of standard laparoscopic instruments and their tips via video using standard FLS training boxes. Bouarfa et al. [17] presented a real-time multi-instrument tracker using compatible colored markers for in vivo use during surgery. Kranzfelder et al. [18] developed an automatic identification system that detects and registers individual instruments during their insertion into the trocar. The system was based on an optoelectronic object detection system using barcodes detected by a micro-endoscopic camera.

Outside the laparoscopic sector, instrument identification is possible by scanning an applied bar or matrix code [19] that manufacturers now generally equip on their instruments by default.

Several systems have been presented with the goal of retained clamp and surgical sponge prevention [20–22].

Neumuth and Meißner [23] presented an approach to automatic surgical instrument detection for workflow detection that is not restricted to laparoscopic surgery and focuses directly on the instrument table by marking the instruments with RFID tags and adding sensors to the situs and table.

The need to develop a system that detects intra-operative instrument usage across a comprehensive range of operations greatly limits previously proposed approaches.

Systems focused on a specific surgery (e.g., laparoscopic surgery) are not applicable to general operating setups.

Approaches based on existing bar or matrix codes require additional actions by the technical nurse during the operation, which is not applicable to real-life surgical conditions.

Meißner and Neumuth [24] presented promising results using RFID identification of the instruments during interventions; however, if additional marking of the instruments using either color codes or RFID tags is required, more challenges arise.

Egan and Sandberg [25] analyzed barcodes, Wi-Fi, and both active and passive RFID automatic identification technologies in the healthcare environment. They identified disadvantages for each type of tag, ranging from the detection range to battery issues. They concluded that the technology must truly fit the problem.

However, the instrument still requires modification, which raises sterilizability and durability issues for the markers or tags and the glue used to attach them. For very slim instruments, the application of a tag or code is often impossible.

Lemke and Berliner [26] stated that it is difficult to justify the cost of new technological and systemic advances in interventional procedures and redesigning healthcare infrastructure, such as ORs.

The addition of tags or markers requires the modification of all instruments to be tracked. Considering the number of instruments in a clinic and the instrument range of a manufacturer, which both easily exceed several thousand units, the cost effectiveness and feasibility of integrating strategies into the clinical workflow are doubtful.

The specified limitations demonstrate the need for an intra-operative surgical instrument identification system without modifications or limitations to laparoscopic surgery. Beeri and Einey [27] presented a rough concept, called VIS-ITS, for visually identifying instruments on trays and disposal surfaces before, during, and after surgery to prevent the retention of instruments and consumables. However, to the authors' knowledge, VISITS has never been elaborated or implemented.

Rattner and Park [28] suggested researching related fields to the one targeted when developing advanced operating room devices to reduce development efforts.

Following this guideline, a visual item verification system for fraud prevention in retail self-checkouts [29] and a smart tray solution [30] for robot rehabilitation systems [31] provide helpful basic concepts and inspirations to develop the multi-sensorial detection approach presented in this article.

In this article, we present a novel system design of a surgical instrument detection system without the need for modification of the instruments, which completely differs from all existing approaches and for the first time includes weight information for identification purposes.

Methods

System design

Structural system setup

The presented system combines multiple sensors arranged around an instrument table.

Figure 1 provides an overview of the setup and sensor equipment involved.

A 2D camera above the table views the entire scene. A Logitech Pro 9000 HD camera with a resolution of

1,600 × 1,200 pixels was used. An Optris PI 160 thermal imager with a 23° × 17° lens, 160 × 120-pixel resolution, 0.08 K temperature resolution (NETD) and system accuracy of ±2 °C was used as the supporting IR camera on the table side. The table itself consisted of a 79 × 59 cm tray mounted on a scale. A PCE BSH 10000 digital scale with a 10 kg weight range and a ±0.2 g accuracy was used. This scale features an RS232 interface that allows for access to the current weight using an intercalated RS232-to-USB converter.

All of the specified sensors were connected to the *Analyzer Central Unit*, a computer system with an Intel Core i7-2760QM Quad Core CPU at 2.4 GHz and 8 GB of RAM running on Windows 7.

The system running on the *Analyzer Central Unit* was divided into two modules: the *Builder* and *Analyzer*.

The *Builder* module allows a user to gather information on the different surgical instruments and to build an *instrument reference container*. To gather the necessary information, each instrument is weighed, and a number of reference images are obtained of the instrument in different positions. The number of images varies based on the shape and composition of each instrument from 2 for simple tools, such as a suction tube, to 12 for complex instruments, such as the scissor-like Blakesley tool, which can open to different angles and looks different on each side. When a new reference image is added, the *Builder* automatically performs several image data processing steps to accelerate the detection process. Therefore, a single image always generates a *reference image series,* where each image knows its algorithmic history. The user can add additional instrument information, such as the camera used, alternative instrument names, a description of the instrument, the manufacturer, a cataloged image, or other known instrument identifiers, such as barcodes. Figure 2 provides an overview of the underlying file structure for the *instrument reference container*.
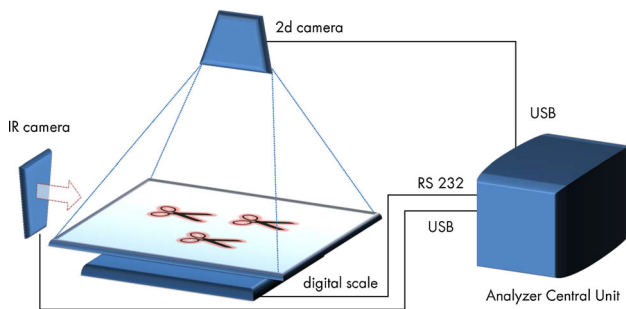

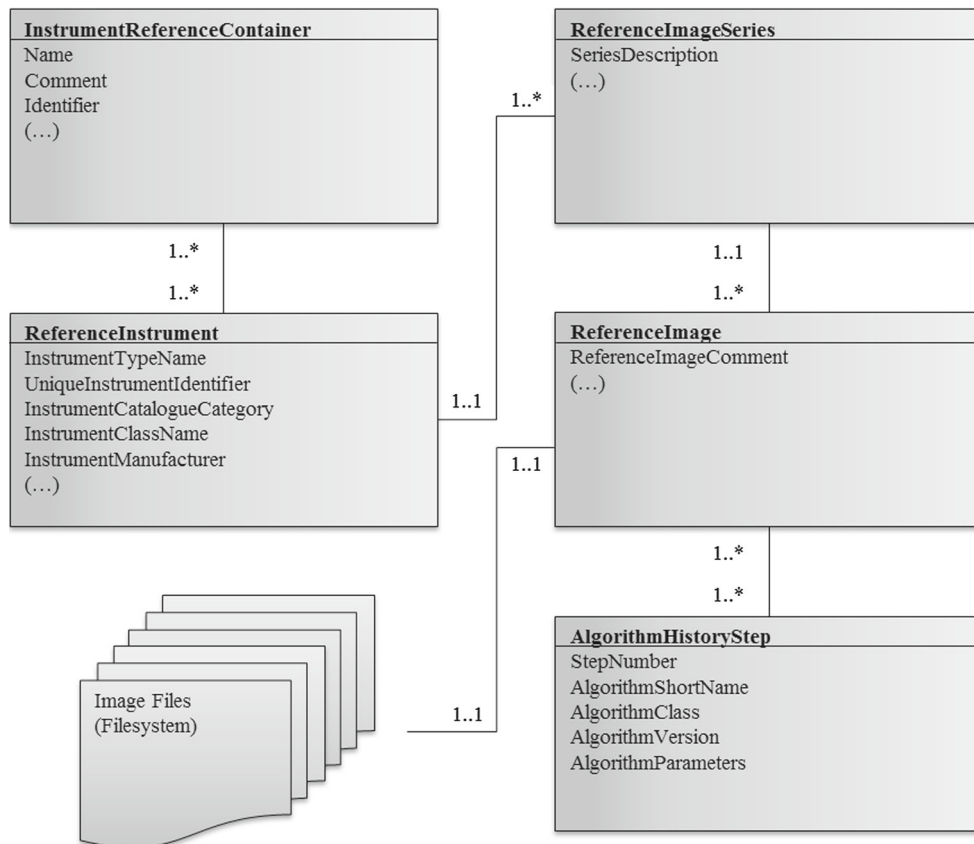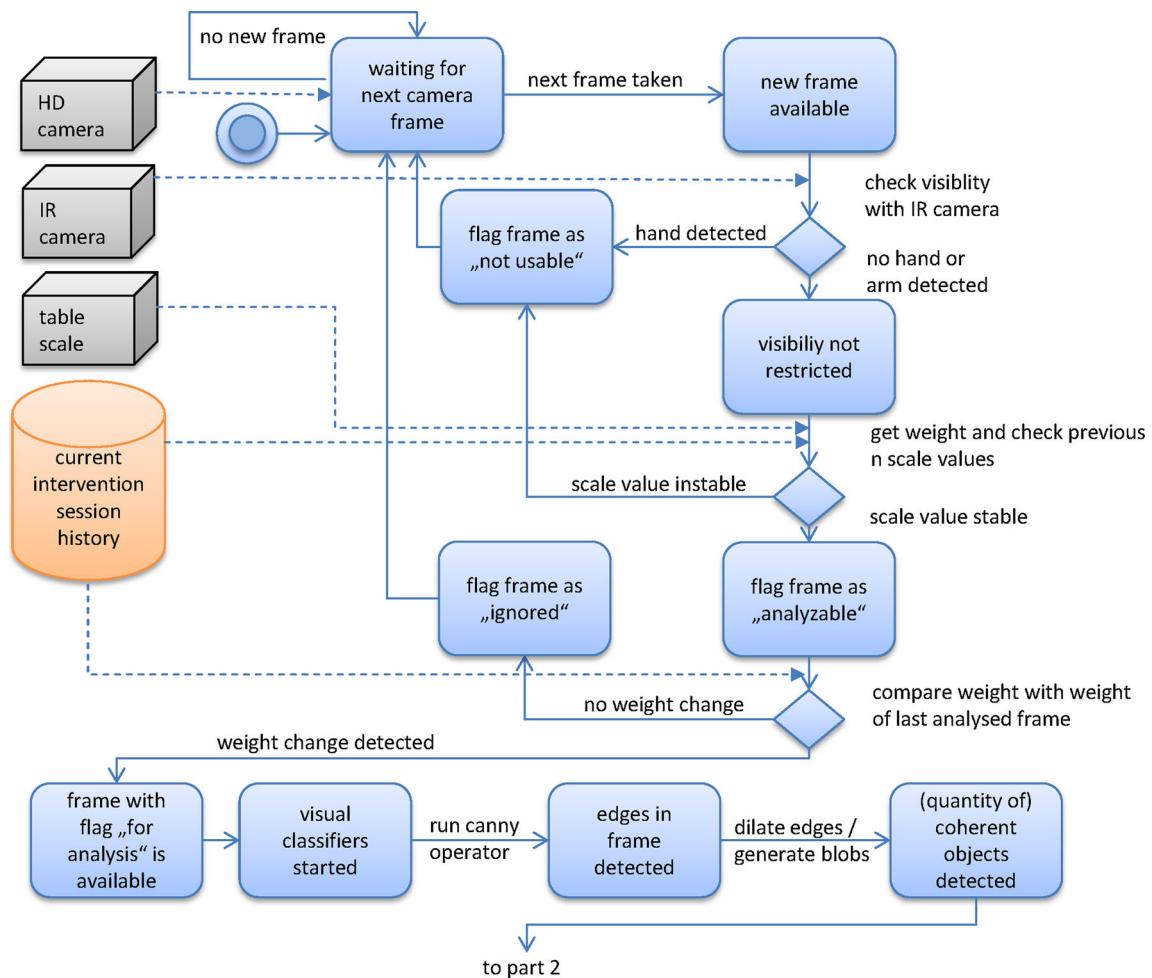**Fig. 1** Schematic of the proposed system

Fig. 2 Overview of the underlying data structure for the *instrument reference container* (excerpt)

HD camera

IR camera

table scale

current intervention session history

no new frame

waiting for next camera frame

next frame taken

new frame available

check visiblity with IR camera

flag frame as „not usable"

hand detected

no hand or arm detected

visibiliy not restricted

get weight and check previous n scale values

scale value instable

scale value stable

flag frame as „ignored"

flag frame as „analyzable"

compare weight with weight of last analysed frame

no weight change

weight change detected

frame with flag „for analysis" is available

visual classifiers started

run canny operator

edges in frame detected

dilate edges / generate blobs

(quantity of) coherent objects detected

to part 2

**Fig. 3** Overview of the detection procedure (part 1)

Dynamic system behavior

The second module, the *Analyzer*, references a previously generated *instrument reference container* upon start-up. This container typically holds information on all instruments in the surgical tray(s) for a single surgical intervention. Therefore, it summarizes all of the a priori information known to the system.

The *Analyzer* identifies objects on the instrument table as follows (Figs. 3, 9 provide a complete overview of the detection algorithm).

The 2D camera above the table takes a steady stream of snapshots called frames. Any frame with a visible hand or finger is discarded by analyzing the corresponding thermal images from the IR camera.

The underlying algorithm classifies the frames as "not usable" if the corresponding IR image exceeds a certain number of pixels above a temperature of 28 °C. Figure 4 presents an IR camera snapshot and temperature legend.

36°C
33°C
30°C
27°C
24°C
21°C
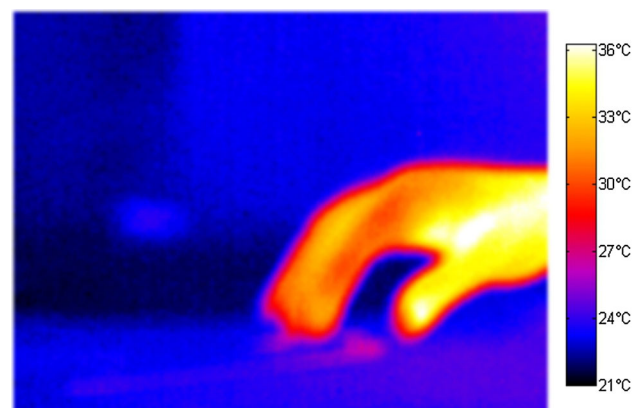
Fig. 4 Snapshot from the IR camera used for hand detection

Frames are also discarded if the weight was unstable at the moment of weighing, for which the last 4 weight values are compared to the tolerance levels of the scale. A weight sam-

pling rate of approximately 8 Hz for the scale also prevents the repeated visual analysis of frames when no noticeable weight change has occurred.

Figure 5 presents an extract of the digital scale values during object movements. The scale values stabilized in a reasonable time to deliver steady results within the accuracy range (±0.2 g).

Visual classification begins when a frame is flagged "for analysis". For this analysis, the frame undergoes edge detection using a Canny filter [32], and the edges are dilated to identify cohesive objects (Fig. 6). Very small blobs are considered noise and discarded. Rotation variance is removed from the procedure by performing a central axis transformation on each detected object (Fig. 7). The resultant rotated edge image is then compared to the a priori shape descriptions in the *instrument reference container,* and a confidence value is determined for each instrument in the reference container within the interval [0, 1].

Instrument candidates with a confidence value below a given threshold are discarded.

The following steps use the weight analysis to specify the results. Two different weight analyses are compared in this article:

The first analysis is called *PowerWeight* and uses the results from a prior visual analysis of the frame to restrict the (weight) search for instruments considered. This analysis is performed because searching for possible instrument combinations based solely on their total weight values without further information is a combinatorial optimization problem known as the knapsack problem. Due to the NP-complete nature of these problems, the computational complexity can easily become incalculable [33].
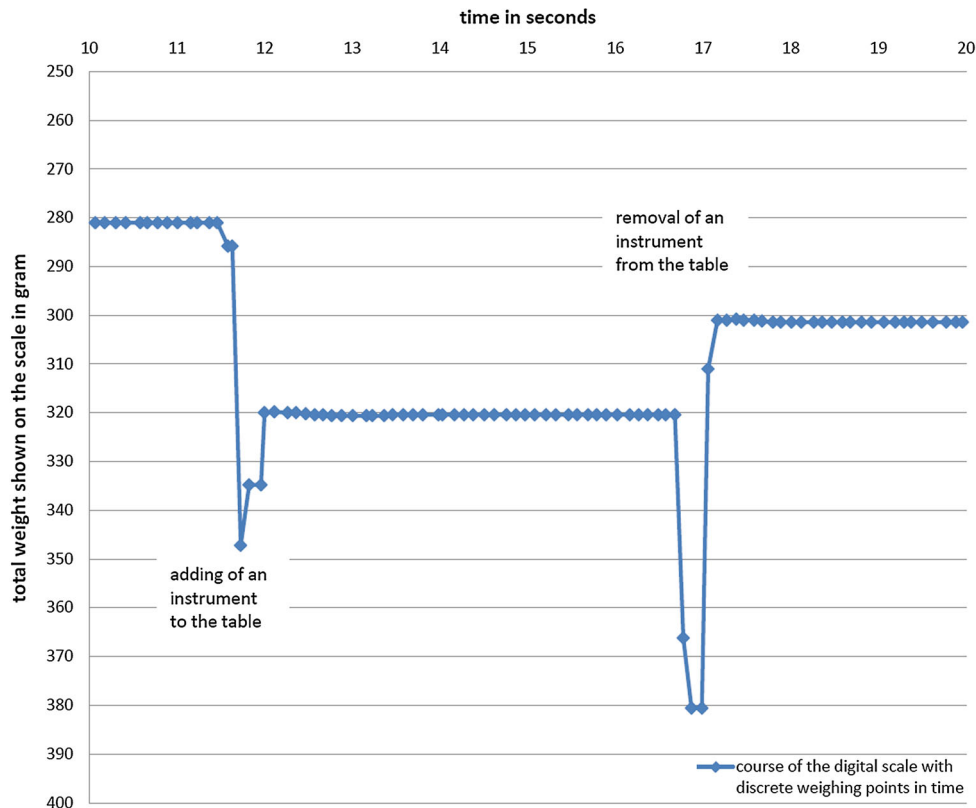

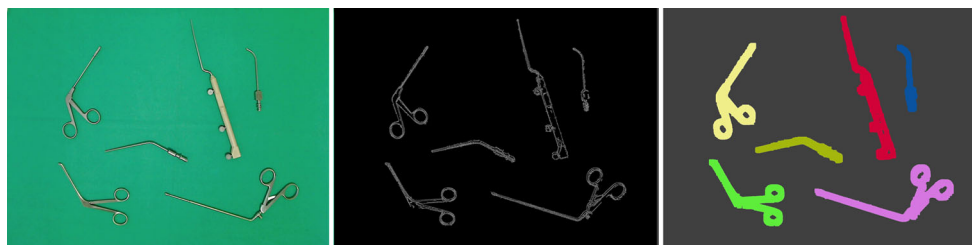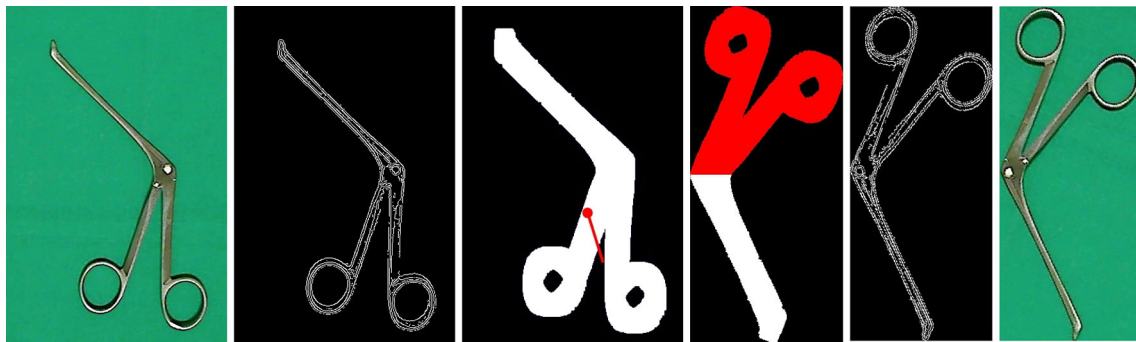**Fig. 5** Progression of digital scale values during object movements

Fig. 6 Edge detection and dilatation to identify coherent objects

Fig. 7 Edge detection and central axis transformation of a Blakesley instrument to remove rotational variance

Fig. 8 Overcoming the visual misclassification of a Blakesley instrument by ruling out a preferred visual candidate as a result of the follow-up PowerWeight analysis
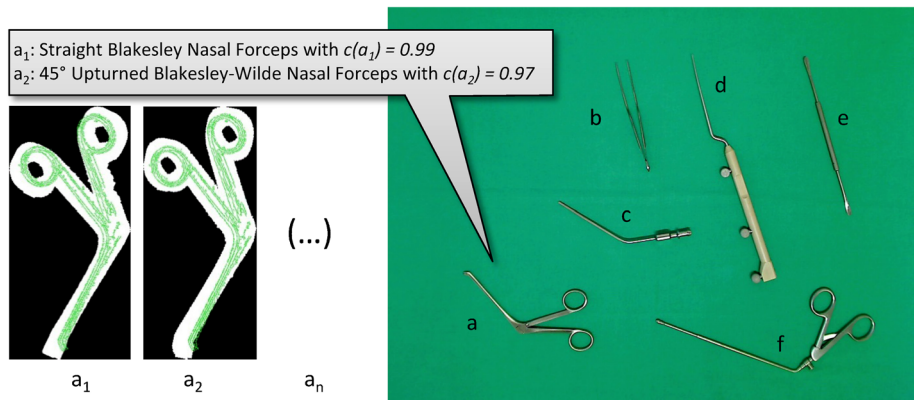

a_1: Straight Blakesley Nasal Forceps with $c(a_1) = 0.99$
a_2: 45° Upturned Blakesley-Wilde Nasal Forceps with $c(a_2) = 0.97$

Figure 8 shows an example frame, in which *instruments b* to *f* are visually detected with only one visual candidate, but the shown *instrument a* (45° upturned Blaskeley–Wilde Nasal Forceps) has multiple candidates and is misclassified as the very similar straight Blakesley nasal forceps when only relying on the best visual confidence value.

The *PowerWeight* analysis uses the multiple possible hits for *instrument a* as a base for the weight detection, by describing the table constellation as follows:

$$M_{\text{TableConstellation}} = (T(a_1) \veebar T(a_2) \veebar \ldots \veebar T(a_n))$$
$$\wedge T(b) \wedge T(c) \wedge T(d) \wedge T(e) \wedge T(f)$$

with $T(x_i)$: Instrument x_i is on table

For each possibility $M(a_i) := T(a_i) \wedge T(b) \wedge T(c) \wedge T(d) \wedge T(e) \wedge T(f)$; then, the total weight is calculated and compared to the weight value measured by the scale:

$$w(M(a_i)) := w_{ref}(a_i) + w_{ref}(b) + w_{ref}(c)$$
$$+ w_{ref}(d) + w_{\text{ref}}(e) + w_{ref}(f)$$

with $w_{\text{ref}}(x_i)$: reference weight of instrument x_i

If $|w(M(a_i)) - \omega| \le \varepsilon$ then $M(a_i)$ is a candidate else $M(a_i)$

is ruled out with ω : measured total weight from scale

For the shown example in Fig. 8 with a measured $\omega = 250.8\,\text{g}$ and a tolerance value $\varepsilon = 1.2\,\text{g}$, this leads to:

For a_1 : $|260.4 - 250.8\,\text{g}| > \varepsilon \Rightarrow M(a_1)$ is ruled out.

For a_2 : $|251.0 - 250.8\,\text{g}| \le \varepsilon \Rightarrow M(a_2)$ is a candidate.

$M(a_2)$ is now chosen as the detection result because a_1 with $c(a_1) > c(a_2)$ has been ruled out by weight information. As a consequence, the *PowerWeight* analysis delivers the correct result compared to the detection relying only on the visual analysis. However, in cases where no valid combination of weight values can be found or when unknown objects are detected on the table, the algorithm falls back to the video-only detection.

The detection procedure overview in Fig. 9 additionally illustrates the approach.

The second approach is called *PreFrameKnowledge* and uses historical knowledge of the current operation session to determine the instrument. This synthetic approach was added to investigate the influence of additional information (here, knowledge of the previous frame) not provided by reality.

Instead of using on the total table weight, the weight difference between the last and current table states is used in this weight analysis to determine the added (or removed) instrument.

Irrespective of the weight algorithm used, the final frame analysis yields a result by determining a combined confidence value for the individual analyses, which can be used to update the current operation session.
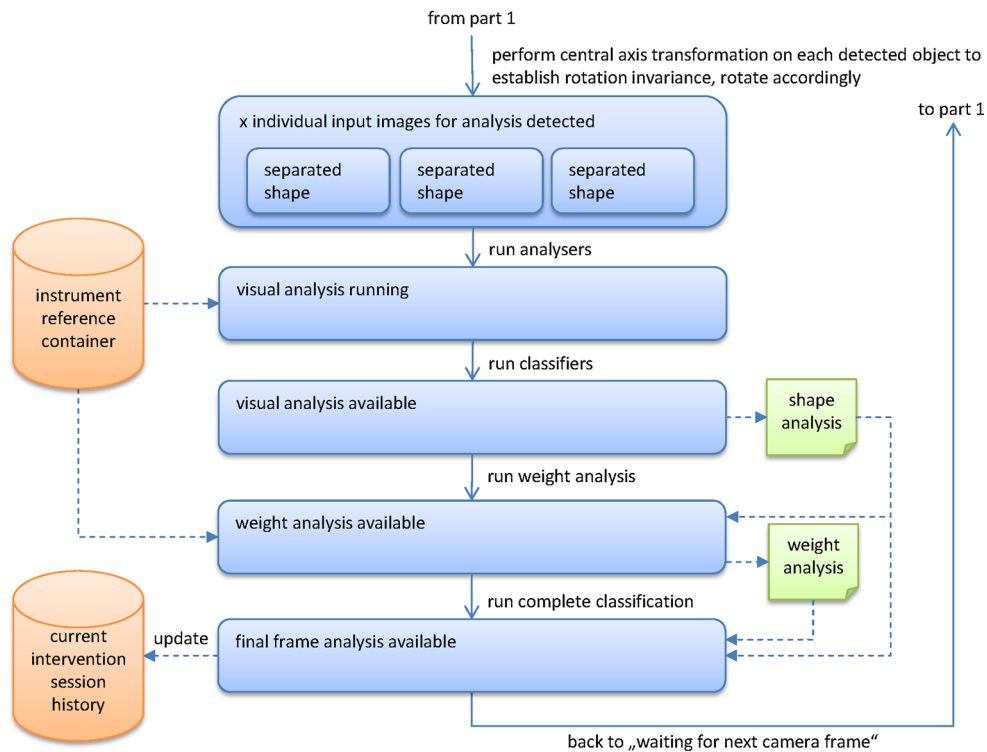
Fig. 9 Detection procedure overview (part 2)

Fig. 10 Comparison of real OR conditions (*left*) to the laboratory condition (*right*)

Evaluation study

Study design

A study to evaluate the recognition rates of the proposed system was conducted under laboratory conditions. Figure 10 compares the real OR and laboratory conditions.

A surgical intervention called functional endoscopic sinus surgery (FESS) was selected as the basis for the study based on its manageable time and instrument range.

For the chosen FESS operations, 27 surgical interventions were analyzed for movements involving the instrument table. The 27 patients (12 women, 15 men) were 19–72 years old with a mean age of 42 ± 15.6 years. Four skilled sur-

geons assisted by five technical nurses performed the interventions at the Acqua Klinik in Leipzig. The observations were restricted to the period from incision to suture with a total observed time of 13 h 4 min 40 s, a mean intervention period of 29 min 3 s, and a standard deviation of 11 min 5 s (minimum: 8 min 31 s, maximum: 58 min 21 s).

Identical surgical trays were used for each of the analyzed operations. The trays held a total of 74 items, including instruments and additional material. Special items not placed on the instrument table, such as cables and optics, and duplicate instruments were removed after consulting experts in the field (a senior technical nurse, two technical nurses, and a surgeon). An additional two items not in the tray but in use were added: a disposable scalpel and the pointing tool for
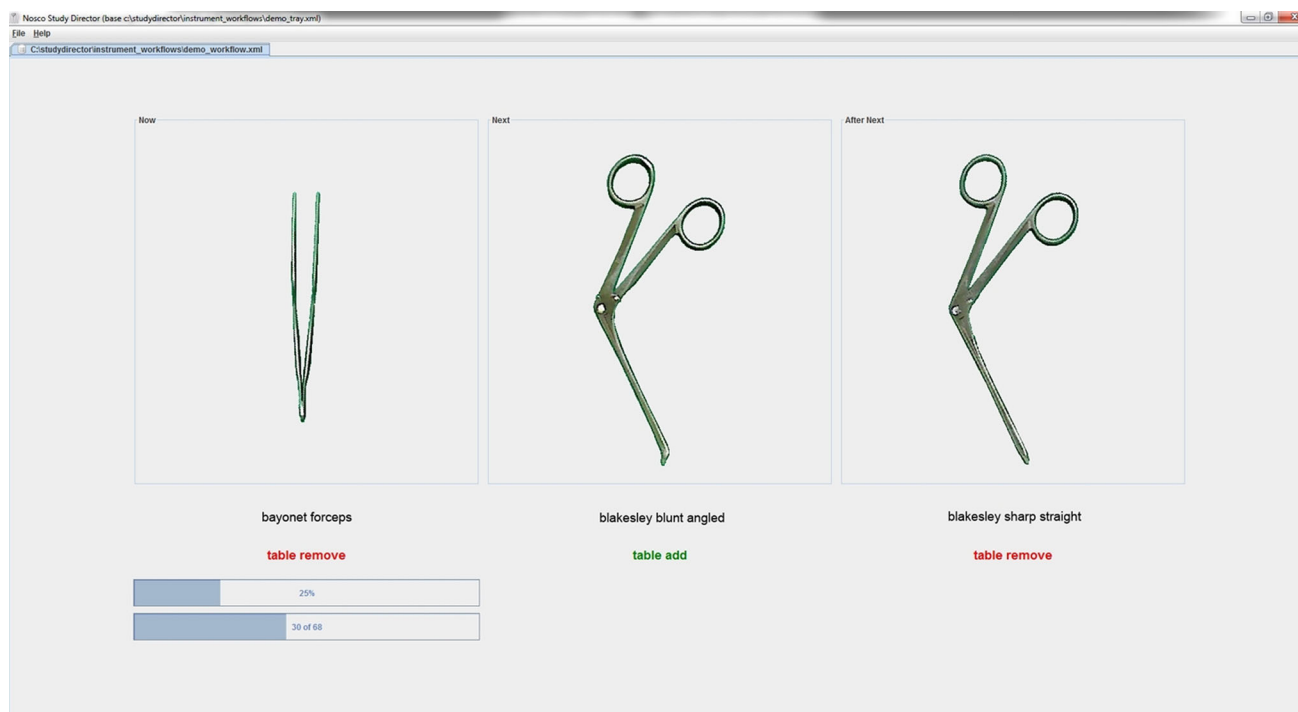
Fig. 11 Screenshot of the StudyDirector GUI

navigating operations. All items were captured by the *Builder* tool as described using a real surgical tray. In total, 49 different instruments were used to generate the corresponding *instrument reference container* using a total of 367 images, with an average of 7.49 images per instrument.

Two medical student observers recorded the instrument table activities using the ICCAS Workflow Editor [34] and generated an *instrument table workflow description model* for each intervention, which detailed when each instrument was moved to and from the instrument table. In total, 3,841 instrument movements were recorded in these models, with an average of 142.25 instrument movements per intervention.

The *instrument table workflow description models* were reduced to recreate the interventions.

The subject handled only the 10 most-used instruments and the pointer. These items accounted for 3,329 (86.7 %) of the total instrument movements, which reduces the average number of instrument movements per intervention to 123.3. The *instrument reference container* used for the detection algorithm still included the complete surgical tray.

To reduce the time requirements for the reenactments and the amount of accumulated data, the frequency of instrument table movements was normalized to 5 s, which reduced the total reenactment time for all interventions from 13 h 4 min 40 s to 4 h 27 min 25 s.

An application called *StudyDirector* was created to allow those without a technical nursing background to reenact the instrument table movements from the surgical procedures.

To this end, a display presents the next instrument to move to or from the table along with a progress time bar, which helps time the exact moment for the next movement. The next two pieces are also shown to mentally prepare the subject. Figure 11 presents a screenshot of the graphic user interface.

The data acquired from each reenacted surgical procedures were analyzed and compared to the known real situation. The detection results of the recreated instrument table movements were determined in three ways for each intervention: once using only visual detection, once combining visual detection and scale analysis, and once with this combination and including knowledge from the previous situation. A detection was considered a hit for a single instrument if it had the highest confidence value as determined by the specific algorithm. The results for each algorithmic approach were statistically analyzed. A Kruskal–Wallis rank sum test was performed to compare the three unmatched groups and test whether the samples originated from the same distribution [35].

Evaluation study results

The detection results from the reenacted instrument table movements were determined and presented three ways for all 27 workflows. Table 1 presents the detection results for each alternative.

The Kruskal–Wallis rank sum test yielded a p value below 0.001. Therefore, the null hypothesis was rejected, and the

Table 1 Detection results (%) for all reenacted instrument workflows

Instrument workflow	W01	W02	W03	W04	W05	W06	W07	W08	W09	W10	W11	W12	W13	W14
Detection rate video only (%)	83.4	83.8	90.5	86.7	81.2	89.7	77.2	84.8	87.0	80.8	83.0	86.0	91.9	85.7
Detection rate in % video and weight (%)	85.4	86.0	100.0	90.5	85.1	91.2	83.7	87.3	90.4	84.8	86.8	92.1	100.0	88.5
Detection rate in % preframe knowledge (%)	99.5	99.7	100.0	99.9	99.8	99.3	99.6	99.2	98.5	99.9	99.8	99.8	100.0	99.5

Instrument workflow	W15	W16	W17	W18	W19	W20	W21	W22	W23	W24	W25	W26	W27	AVG
Detection rate video only (%)	79.5	84.8	99.5	92.9	75.4	79.7	82.1	81.3	83.9	86.5	77.0	88.5	90.6	84.9
Detection rate video and weight (%)	98.8	89.0	96.7	100.0	82.7	86.9	86.5	82.9	87.7	91.0	84.4	100.0	100.0	90.3
Detection rate preframe knowledge (%)	100.0	98.9	99.6	100.0	98.5	99.9	98.6	99.2	99.8	99.6	99.5	100.0	100.0	99.6
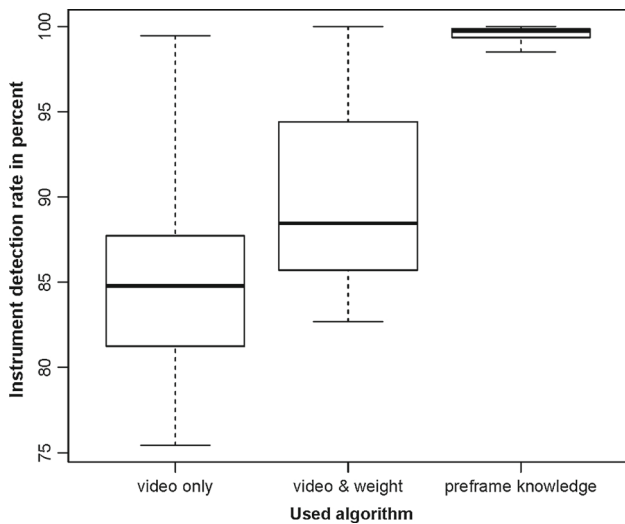
Fig. 12 Box-and-whisker diagram for the detection results

alternative hypothesis stating that the groups are not all from the same distribution is assumed.

Thus, the groups were subjected in pairs to a Mann–Whitney–Wilcoxon test to determine whether each pair is from the same population [36].

The comparison of "video only" to "video and weight" yielded a p value of 0.002.

The comparison of "video and weight" to "preframe knowledge" yielded a p value below 0.001.

The comparison of "video only" to "preframe knowledge" yielded a p value below 0.001.

Because all of the p values were below 0.01, the alternative hypothesis was assumed for each pairing, which leads to the assumption that the populations all differed.

Figure 12 presents a box-and-whisker diagram for the detection results.

The "video-only" detection rates had a mean of 84.9 %, a standard deviation of 5.4 %, a lower quartile at 81.2 %, and an upper quartile at 87.7 % (minimum: 75.4 %, maximum: 99.5 %). The "video & weight" detection rates had a mean of 90.3 %, a standard deviation of 6.0 %, a lower quartile at

85.7 %, and an upper quartile at 94.4 % (minimum: 82.7 %, maximum: 100.0 %). The "preframe knowledge" detection rates had a mean of 99.6 %, a standard deviation of 0.5 %, a lower quartile at 99.4 %, and an upper quartile at 99.9 % (minimum: 98.5 %, maximum: 100.0 %).

The detection rates using the "video and weight" algorithm are, with one exception, always higher than the "video-only" detection rates, and the "preframe knowledge" detection rates are always higher than the detection rates of the other two algorithms. Five of the workflows using the "video and weight" algorithm and six of the workflows using the "preframe knowledge" algorithm provided detection without error.

Discussion

We developed a system capable of autonomously identifying unmodified surgical instruments during an intervention. Instead, the presented system detects the instruments using a combination of sensors. A study was conducted to evaluate the recognition rates for this system under laboratory conditions and to compare different detection algorithms.

The detection rates when using the "preframe knowledge" algorithm had a mean value of 99.6 %; however, this detection rate cannot be transferred directly to an instrument table. Knowledge from the previous frame is not available in reality. However, these results indicate additional situational knowledge (in this case, the state of the table before its last modification) improves the detection rate. The detection rate using the "video and weight" algorithm can be directly transferred to the OR setting and yielded a mean value of 90.3 % with a minimum of 82.7 %. Therefore, combining shape and weight information with sensor data allows for the accurate identification of surgical instruments. With the exception of W17, the individual detection rates of all workflows were higher than the "video-only" detection rates, which had a mean of 84.9 %.

This 5.4 % increase in the detection rate indicates that the weight information contributes to the detection preci-

sion. The maximum detection rate improvement was for W15, which increased from 79.5 % for the "video-only" algorithm to 98.8 % with the addition of weight information. This 19.3 % increase demonstrates that the weight information can be decisive for the detection rate.

The algorithm used to visually detect the instruments has proven applicability to the given situation.

The *instrument reference container* of the system was created using the complete content of the surgical tray from the examined FESS interventions. The parallel use of a thermal camera to detect hands robustly discards frames with a visible hand and indicates table changes. When the subject is wearing surgical gloves, the measured hand temperature varies by up to 2 °C, which does not affect the hand detection algorithm.

The digital scale also yielded reasonable results. The RS232 interface delivered the scale values at a varying frequency (ca. 8 Hz).

Under real OR conditions, quicker instrument changes sometimes can keep the line of sight of the IR camera obstructed while more than one instrument change is happening. As a consequence, some instrument changes might happen unnoticed to the system. But since the video and weight variant does not rely on previously detected frames, the analysis detects that more than one change has occurred. With an additional inclusion of the scale movements, an indicator for instrument changes is also provided for phases when the line of sight of the IR camera is blocked.

Because the system only monitors the table, naturally, all circumventions of the instrument table considering instrument usage cannot be detected. Especially, the additional tray next to the instrument table must be monitored separately, as well as all additional devices used by the surgeon such as drills or surgical fraise.

As a central constraint, the visual algorithm cannot handle superposed objects, which are common for OR instrument tables. Any overlapping of an instrument will generate a blob containing two or more objects. Superimposed objects or overlaps can lead to a failing identification of these instru-

ments during the process, but the impact of missing sensor information can be compensated by follow-up systems [37]. Furthermore, the visual algorithm cannot handle flexible objects, such as consumables, e.g., swabs. These limitations require additional algorithms. However, both superposed and unknown objects have a high probability of being classified as unknown objects.

Because the visual detection algorithm depends only on the object edges, not the color information, it can be used under the steadily changing lighting situation of an OR. Strong shadows can negatively affect the detection results; however, the algorithm tolerances can be adjusted. As one might expect, a consistently lit instrument table is beneficial. Reflections can also affect the detection rates but were inconspicuous in this study.

The algorithm can consider instruments with movable parts, such as scissors, because it allows for multiple shapes for each instrument. Using precise instrument models, this method could even utilize cataloged instrument information, which would reduce the effort required to make instruments known to the system.

If additional instruments, which have not been made known to system previously, are added during surgery, which can happen in real OR situations, the system would most probably then classify these as unknown objects.

Some problematic cases for the shape- and weight-based approach were identified during this project. Figure 13(1) provides an example of two radically divergent instruments from an established instrument manufacturer with identical type numbers.

The pictured surgical retractors significantly differ in length (0.7 cm), width, height, and weight (2.0 g difference), whether this problem was a workmanship fault or is normal for the fabrication process has not been determined. Figure 13(2) provides an example of instruments that differ because of abrasion and maintenance modifications. The pictured tips of the tweezers vary by several millimeters in length because they are often shortened during maintenance to maintain their grip.
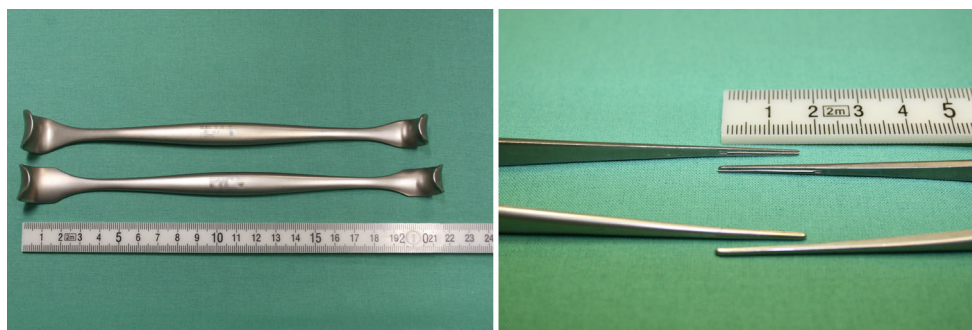
Fig. 13 Examples of differing instruments with identical part numbers. *Left* (1) length comparison of two surgical retractors, *Right* (2) length comparison of two tweezer tips

Such real-life circumstances will presumably negatively impact the detection rates, and the algorithm tolerance for shape and weight discrepancies must be adjusted if the instrument reference container was not initialized for the exact instruments being detected by the system.

Conclusions

To automatically detect the current OR situation, identifying the instrument currently being used during a surgical intervention is a critical parameter.

The proposed system uses combined sensors to detect the instruments and differs significantly from existing approaches. The evaluation study delivered detection rates of 90.3 % when combining video and weight information and even higher rates when information from previous frames was included. Compared to existing approaches for determining the instruments used during a surgical intervention, the proposed approach is not dependent on a special type of surgical intervention, such as laparoscopy. Furthermore, this technique circumvents the need to modify the instruments. Therefore, this method allows for the detection of instruments that cannot be modified because of their proportions or financial practicality, e.g., single-use instruments. The presented algorithm could also be enhanced to consider consumables. The proposed approach can also be combined with existing approaches using RFID or laparoscopic images in order to merge their individual advantages.

Detecting the instruments used during a surgical intervention contributes significantly to work that depends on the knowledge of the current OR situation, such as workflow-based assistance systems.

The presented system can also contribute to the statistical analysis of instrument usage rates and lead to their optimization because instruments that are randomly or never used on the trays could be identified. This system could also be used to resolve the problem of retained surgical instruments or consumables by monitoring objects given to the surgeon and adding a security count to those already performed by the persons responsible for the OR.

A follow-up project will transfer the demonstrated feasibility of this approach to a real OR instrument table and enhance the current algorithms to address some of the remaining restrictions.

References

1. Sutherland J, Van den Heuvel W (2006) Towards an intelligent hospital environment: adaptive workflow in the OR of the future. In: Proceedings of the 39th annual Hawaii international conference on system sciences, 2006, vol 5. HICSS'06, pp 100b–100b
2. Macario A, Vitez TS, Dunn B, McDonald T (1995) Where are the costs in perioperative care? Analysis of hospital costs and charges for inpatient surgical care. Anesthesiology 83(6):1138–1144
3. Archer T, Macario A (2006) The drive for operating room efficiency will increase quality of patient care. Curr Opin Anesthesiol 19(2):171–176
4. Neumuth D, Loebe F, Herre H, Neumuth T (2011) Modeling surgical processes: a four-level translational approach. Artif Intell Med 51(3):147–161
5. Neumuth T, Jannin P, Schlomberg J, Meixensberger J, Wiedemann P, Burgert O (2011) Analysis of surgical intervention populations using generic surgical process models. Int J Comput Assist Radiol Surg 6(1):59–71
6. Cleary K, Kinsella A, Mun SK (2005) OR 2020 workshop report: operating room of the future. Int Congr Ser 1281:832–838
7. McKenzie C, Ibbotson J, Cao C (2001) Hierarchical decomposition of laparoscopic surgery: a human factors approach to investigating the operating room environment. Minim Invasive Ther Allied Technol 10(3):121–127
8. Kranzfelder M, Schneider A, Gillen S, 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. Surg Endosc 25(3):696–705
9. Uecker DR, Wang YF, Lee C, Wang Y (1995) Laboratory investigation: automated instrument tracking in robotically assisted laparoscopic surgery. Comput Aided Surg 1(6):308–325
10. Speidel S, Delles M, Gutt C, Dillmann R (2006) Tracking of instruments in minimally invasive surgery for surgical skill analysis. In: Medical imaging and augmented reality. Springer, pp 148–155
11. Speidel S, Benzko J, Krappe S, Sudra G, Azad P, Peter B (2009) Automatic classification of minimally invasive instruments based on endoscopic image sequences. Proc SPIE 7261:72610A–1
12. Voros S, Long J-A, Cinquin P (2006) Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder. In: Medical image computing and computer-assisted intervention—MICCAI 2006. Springer, pp 535–542
13. Voros S, Long J-A, Cinquin P (2007) Automatic detection of instruments in laparoscopic images: a first step towards high-level command of robotic endoscopic holders. Int J Robot Res 26(11–12):1173–1190
14. Sznitman R, Richa R, Taylor RH, Jedynak B, Hager GD (2013) Unified detection and tracking of instruments during retinal microsurgery. IEEE Trans Pattern Anal Mach Intell 35(5):1263–1273

15. Tonet O, Thoranaghatte RU, Megali G, Dario P (2007) Tracking endoscopic instruments without a localizer: a shape-analysis-based approach. Comput Aided Surg 12(1):35–42

16. Allen BF, Kasper F, Nataneli G, Dutson E, Faloutsos P (2011) Visual tracking of laparoscopic instruments in standard training environments. Stud Health Technol Inf 163:11–17

17. Bouarfa L, Akman O, Schneider A, Jonker PP, Dankelman J (2012) In-vivo real-time tracking of surgical instruments in endoscopic video. Minim Invasive Ther Allied Technol 21:129–134

18. Kranzfelder M, Schneider A, Blahusch G, Schaaf H, Feussner H (2009) Feasibility of opto-electronic surgical instrument identification. Minim Invasive Ther Allied Technol 18(5):253–258

19. Shepherd JP, Brickley MR, Jones ML (1994) Automatic identification of surgical and orthodontic instruments. Ann R Coll Surg Engl 76(2 Suppl):59–62

20. Rogers A, Jones E, Oleynikov D (2007) Radio frequency identification (RFID) applied to surgical sponges. Surg Endosc 21(7):1235–1237

21. Rivera N, Mountain R, Assumpcao L, Williams AA, Cooper AB, Lewis DL, Benson RC, Miragliotta JA, Marohn M, Taylor RH (2008) ASSIST-automated system for surgical instrument and sponge tracking. In: IEEE international conference on RFID, pp 297–302

22. Kranzfelder M, Zywitza D, Jell T, Schneider A, Gillen S, Friess H, Feussner H (2012) Real-time monitoring for detection of retained surgical sponges and team motion in the surgical operation room using radio-frequency-identification (RFID) technology: a preclinical evaluation. J Surg Res 175(2):191–198

23. Neumuth T, Meissner C (2012) Online recognition of surgical instruments by information fusion. Int J Comput Assist Radiol Surg 7(2):297–304

24. Meißner C, Neumuth T (2012) RFID-based surgical instrument detection using Hidden Markov models. Biomed Tech 57:1

25. Egan MT, Sandberg WS (2007) Auto identification technology and its impact on patient safety in the operating room of the future. Surg Innov 14(1):41–50

26. Lemke HU, Berliner L (2011) Systems design and management of the digital operating room. Int J Comput Assist Radiol Surg 6(1):144–158

27. Beeri Y, Einey D (2003) VISITS—Visual surgical instrument tracking system, 2003. [Online]. http://www.cs.huji.ac.il/course/2004/postPC/projects/fall2003/visits.html. Accessed 22 Apr 2014

28. Rattner DW, Park AE (2003) Advanced devices for the operating room of the future. Semin Laparosc Surg 10:85–90

29. Bobbit R, Connell J, Haas N, Otto C, Pankanti S, Payne J (2011) Visual item verification for fraud prevention in retail self-checkout. In: IEEE workshop on applications of computer vision (WACV), pp 585–590

30. Volosyak I, Radchenko O, Pape A, Martens C, She H, Wendland E, Gräser A (2003) Smart tray for the support of a wheelchair mounted manipulator. In: Proceedings of the international conference on economic, engineering and manufacturing systems. ICEEMS 2003

31. Volosyak I, Ivlev O, Graser A (2005) Rehabilitation robot FRIEND II-the general concept and current implementation. In: 9th International conference on rehabilitation robotics, 2005. ICORR 2005, pp 540–544

32. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 6:679–698

33. Pisinger D (2005) Where are the hard knapsack problems? Comput Oper Res 32(9):2271–2284

34. Neumuth T, Durstewitz N, Fischer M, Strauß G, Dietz A, Meixensberger J, Jannin P, Cleary K, Lemke H. U, Burgert O (2006) Structured recording of intraoperative surgical workflows. In: Medical imaging, pp 61450A–61450A

35. Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. J Am Stat Assoc 47(260):583–621

36. Mann HB, Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat 18(1):50–60

37. Liebmann P, Meixensberger J, Wiedemann P, Neumuth T (2013) The impact of missing sensor information on surgical workflow management. Int J Comput Assist Radiol Surg 8(5):867–875