

A User-Centered and Workflow-Aware Unified Display for the Operating Room

Ralf Stauder¹, Vasileios Belagiannis¹, Loren Schwarz¹, Ali Bigdelou¹,
Eric Söhngen², Slobodan Ilic¹, and Nassir Navab¹

¹ Computer Aided Medical Procedures, Technische Universität München, Germany
{stauder, belagian, schwarz, ali.bigdelou, slobodan.ilic, navab}@cs.tum.edu

² Klinikum rechts der Isar, Technische Universität München, Germany
soehngen@me.com

Abstract. In a modern operating room, usually a plethora of imaging devices is available. These come with at least one monitor each and additional, wall-mounted monitors are often available to review patient data and pre-operative imaging. Due to sterility constraints and limited space in the operating room (OR), most of these devices have to be organized around the operating table with some distance. This leads to the effect that the surgeon has to leave the sterile operation table in order to gain device or display access and usually relies on external help to manipulate the device parameters.

In this paper we propose an approach to this problem by employing a single, central and sterile draped display within the reach of the surgeon. This display receives live image data of all devices used in the OR, and pre-operatively acquired DICOM files. Depending on the current workflow phase of the operation, the surgeon has the possibility to switch among the sources with a gesture recognition system. Given the correct context, the surgeon can change selected parameters through other gestures, like adjusting the visible slice within a DICOM viewer. The system was designed in close partnership with our medical partner and evaluated through a usability study.

Keywords: User Interface, Surgical Workflow

1 Introduction

Modern surgery faces an ever increasing rise in the number of modalities and medical devices that are being used throughout the course of an operation. The merger of diagnostic and therapeutic technologies, often referred to as theragnostics will further promote this tendency in the future. The need for superior visualization and display techniques inside the operation theatre that provide visual information to both surgeons and assisting staff increases accordingly.

The growing complexity and the resulting wealth of intraoperatively available information also requires smart techniques, that filter relevant information for the surgical team and provide intuitive access to them. This includes closing the

usability gap introduced by the sterility requirements, that render direct access to the devices for the surgeons impossible and force them to rely on remote control through a non-sterile nurse. This could be approached by providing a sterile display to the surgeon with the possibility to pre-filter the shown information based on the current surgery phase.

1.1 Related Work

Much has been done in the field of surgical workflow detection and analysis in the last few years. To analyze recorded surgical workflows, [1] used a framework based on the concept of data warehousing. Detecting and visualizing the surgical workflow based on the knowledge, which instrument is in use at a given time was done by [2] using the methods of Dynamic Time Warping (DTW) and Hidden Markov Models (HMM). The required instrument usage vector can be collected by using marked instruments as in [3]. Based on HMMs and DTW, [4] uses this instrument usage vector to detect the currently active phase of an a priori known workflow online, while [5] detects the surgical workflow by directly evaluating the images acquired through a laparoscopic camera. On a coarser level of detail, [6] uses movement patterns of surgical staff to analyze the workflow, while [7] is working with gesture cues to detect intentions on a more detailed workflow level.

A popular trend in the field of human-computer interaction is the way towards enabling multiple users to collaborate on a single, shared display or in virtual collaboration. [8] introduced the concept of virtual collaboration to the office environment, while [9] focused on a shared display for multiple collocated users. [10] combines the shared display with smaller devices, so that every collaborator can use their own device.

Positioning and availability of imaging sources can have a significant influence on the result of a surgery. This was already shown in [11], in which the monitor showing the video of a laparoscopic operation was positioned in different ways around the surgeon. Another system using multiple data sources to show information dynamically on various displays in the operating room was implemented in [12]. Only a single, sterile draped, touch-enabled display was used in [13], attached directly to the intraoperative microscope to enable easy access to preoperative imaging data.

1.2 Contributions

In the present work we apply videostreaming of several intraoperatively used devices to a single, unified display. This display is visible to all members of the surgery staff without problems and provides convenient access to all relevant imaging data during the course of an operation.

This centralized and sterile draped display enables surgeons to directly control the display and potentially parameters of connected devices. While this can easily be done through the use of a touch-enabled interface on the display itself,

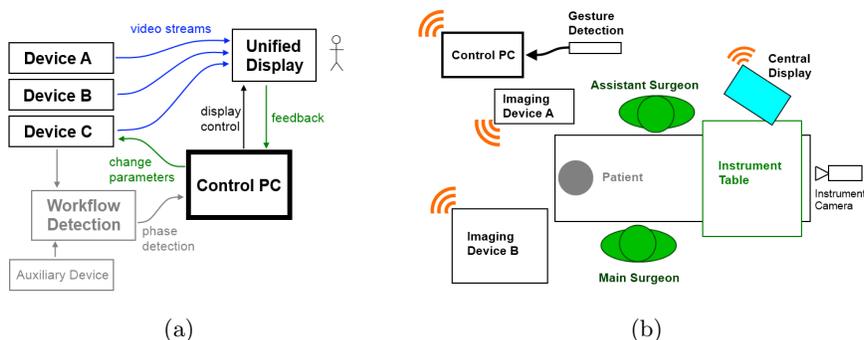


Fig. 1: A schematic view of our system. a) Interaction of the components. b) Layout of the examined OR.

we focus on manipulating selected parameters by employing a simple, touch-free gesture recognition approach.

The selection of the imaging source to be displayed is achieved by utilizing knowledge of the surgical workflow. Meaningful sources were predefined and connected to detected gestures for each workflow phase, allowing for a different image and interaction method in each phase. This workflow-oriented approach potentially eliminates the need for interaction with the system completely for the case that no complications occur during the surgery.

2 Method

All devices used in this system were directly or indirectly connected through a common network. As it can be seen in Figure 1, all control signals and video streams were sent through the network to the central display, while parameter changes and other feedback data could equally be sent back to the individual devices through the control PC. Prior to the surgery, the expected workflow phases are defined where the various image sources, data connections and recognized gestures are connected in a network on this control PC. These connections are depending on and mapped to every single workflow phase, as seen in Figure 4. This can be a simple, constant mapping to an image source, which will always be shown in one phase. Alternatively it can be up to several views that can be switched by a connected gesture in another phase. Furthermore the movement values of the gestures are mapped to the manipulation of parameters, e.g. the slice number of an image sequence viewer.

2.1 Gesture Based Input

The gesture recognition is based on the arm’s motion of the main surgeon. The body part detection relies on a Microsoft[®] Kinect[™] with its skeletal information, though any other gesture detection method could be introduced to the

system. A more sophisticated approach to gesture recognition with a KinectTM was done in [14], Time of Flight (ToF) cameras were used by [15], and inertial sensors in [16]. While we were using touchless gestures for this setup, previous experiments have shown that the capacitive touchscreens of most multi-touch enabled displays remain usable without any difficulty even under several layers of sterile sheeting and gloves. Thus a touch-based user interface directly on the draped display is also imaginable.

2.2 Workflow Detection

Figure 1a also depicts the possible use of a workflow detection module, automatically providing phase information to the control network. To this end several cameras are arranged to record the surgery from different viewpoints for further analysis, tracking and possible scene reconstruction. One camera is filming the operative field from the perspective of the surgical lighthouse while another camera targets the instrument table (see Figure 3a). These images could be used to identify and track the surgical instruments at any given time in a similar fashion to either [3] or [5]. This way the instrument usage vector (Figure 3b) can be created, which is necessary for the automatic workflow detection as described in [4]. An instrument usage vector consists of a binary value for each surgical instrument available, where the value of an instrument is 1 iff this instrument is in use at a specific time and 0 otherwise. A full surgery can be described by recording an usage vector for every time step, e.g. every second.

2.3 Emergencies

In the context of this prototype we did not handle inaccuracies, failures or emergencies within the procedure. Most surgical devices do not offer any special emergency handling, so they simply have to be moved away, which would also be true for our central display. Depending on the mount this can be done more quickly for a single display than for multiple devices. While in such cases the surgeon's talent for improvisation is most important, the workflow detection could identify an outlier [17] to allow the surgeon to freely switch between all available data sources.

3 Experiments and Results

For our experiments we recreated the situation of an operation room, set up the device network as described in section 3.1 and mimicked our target surgery according to a detailed script. After a series of recordings, we conducted a usability study with everyone using our prototype and asked a proficient expert for his opinion about the system.

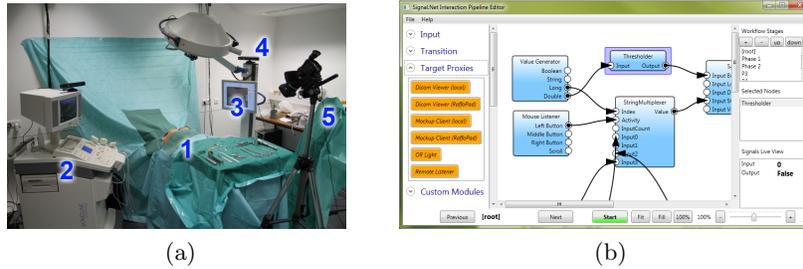


Fig. 2: a) The simulated OR of our experiment, based on the proposed method. 1 is the operative area with a dummy, 2 and 3 imaging devices used during the experiment, 4 the Kinect used for gesture recognition and 5 the camera recording the instrument table. b) The graphical network editor used to set up the connections of the different components, subject to the individual workflow phases.

3.1 System Setup

Our prototype consisted of four major components:

- A commercially available Slate PC on a stand, prepared with specialized software. This tablet was used as the central display of the surgery. The software running during the test included a DICOM viewer, which could be controlled through network commands, software to display the video streams received from the imaging sources and a customized application to switch between the different views based on received network commands. Additionally a screen recording software captured the visible screens throughout the procedure.
- An Ultrasound device with video output, connected to a regular PC with a video grabber card. The captured ultrasound video was streamed to the central display. (Number 2 in Figure 2a)
- A tracked freehand 3D SPECT device with an augmented video image. This video was streamed to the unified display with all associated additional data, usually displayed on the device’s own monitor. (Number 3 in Figure 2a)
- A control PC with customized software to detect gestures, define the workflow phases and the corresponding data network and send commands to the display through the network. The visual editor used to create this mapping can be seen in Figure 2b.

All components were connected through a shared IEEE802.11b/g wireless network. Four ceiling-mounted cameras were capturing the full scene from different viewpoints, one camera mounted on a tripod filmed the instrument table, another camera attached to the surgical light recorded the operative area. The simulated surgery was performed by two persons, imitating the main and assistant surgeon, working on a dummy. In the context of this experiment, we focused on the evaluation of the usability of the unified display, assuming the current workflow phase is known as described in [4].

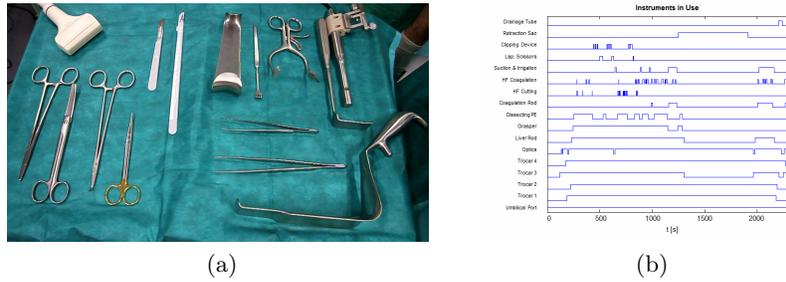


Fig. 3: a) The view of the camera filming the instrument table (Number 5 in Figure 2a). This information could be used to reconstruct the used instrument vector. b) An exemplary instrument usage vector.

#	Phase	Visible Imaging Source
1	“safe-surgery” check	DICOM images
2	Ultrasound check	Ultrasound image
3	3D SPECT image acquisition	3D SPECT video
4	Incision planning	
5	Incision, approaching tumor	DICOM images
6	Tumor removal	DICOM images
7	Marking of tumor margins	
8	Sentinel lymph node detection	DICOM images
9	Lymph node identification and localization	3D SPECT video
10	Confirmation for absence of radioactivity	3D SPECT video
11	Suturing	

Fig. 4: Short description of the surgical workflow used in our experiment.

3.2 Exemplary Workflow

The system was employed for assistance in a breast cancer case within a simulated operation room environment, in which the nonpalpable tumor in the breast and consequently a radiolabeled (sentinel) lymph node underwent surgical resection. The traditional surgical procedure was enhanced with the described multimodality visualization system and surgical workflow control. A short overview of the workflow is given in Figure 4.

For preparation, the surgeon completed a “safe-surgery” checklist, in which patient data, planned procedure and equipment were controlled. During this phase the tablet displayed preoperatively acquired DICOM images. In the real case, these images could include mammography, MRI and scintigraphy, while in the simulated case only a CT sequence was shown for sake of evaluation speed. Before initial incision, ultrasound of the breast was performed to confirm the tumor location, with the tablet display switching from DICOM to ultrasound. Using the freehand 3D SPECT device, the distribution of up to 24 hour priorly administered radioactivity was acquired with a navigated gamma probe to create

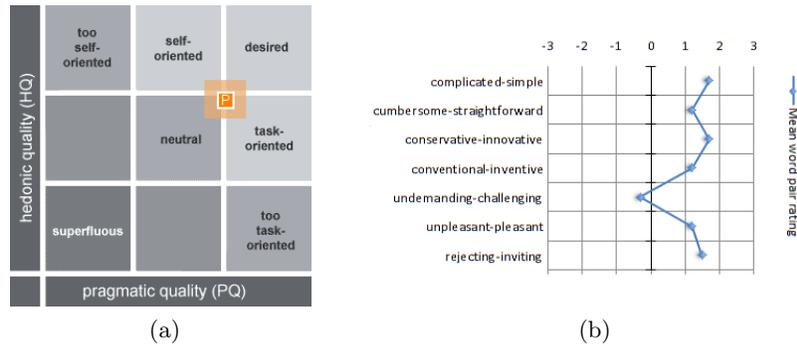


Fig. 5: a) Portfolio displaying the mean value (P) and confidence interval (transparent rectangle) of the usability test's results of our prototype. b) Selected word pairs of the questionnaire and their mean score.

a pre-precision imaging dataset. Image acquisition and probe guidance were displayed on the central tablet display. Based on the findings, the surgeon planned the incision and started the surgical procedure. While approaching to the tumor in the breast, the surgeon compared the intraoperative findings at the surgical site with the preoperative DICOM images, displayed on the tablet, to confirm depth and localisation.

After removal of the tumor, the margins of the resected tissue were marked with different colored threads for consequent histological sampling. Using the same incision (ADAM technique [18]), the surgeon proceeded into the axillary region for sentinel lymph node detection and excision. During these steps of the operation, frequent access to scintigraphy data and SPECT imaging was required while approaching to the site of the sentinel lymph node. The display therefore switched between preoperative and intraoperative imaging. After post-excision image acquisition with the SPECT system and confirmation of no remaining radioactive hot spots, tissue and skin layers were adapted with sutures.

Only a single image modality was shown on the central display at any given time, according to the traditional workflow. Usually switching between image sources also means actively moving a specific monitor or changing the viewer position to face another display and is therefore avoided as much as possible, so the advantages of using several modalities at the same time still have to be examined.

3.3 Qualitative Analysis

The user study involved a short AttrakDiff³ questionnaire of several word pairs with opposite meanings, for which every participant could choose on a 7-level Likert scale, which word of each pair best described the system usability. The

³ <http://www.attrakdiff.com>

results were grouped by the terms of pragmatic quality (PQ) and hedonic quality (HQ) later and plotted on a usability portfolio, as seen in Figure 5a. The participants were students of biomedical engineering, familiar with the situation and requirements of the OR, in the age group of 20-40 (N=4). The mean value of the results placed our system close to the *neutral* section on both axes, although the confidence rectangle indicates a trend towards the *desired* section. While the prototype suffered from noticeable network lag during the experiments, the potential to an optimal usability score is clearly visible.

Additionally to the user study, we asked Prof. Kazuhiro Hongo, Professor and Chairman, Dept. of Neurosurgery, Shinshu University, Japan, to share his evaluation of the proposed system with us. He already had experience with an operation-microscope mounted touch display, “in which a surgeon was able to access preoperative MRI imaging during the operation quite conveniently and independently right from the site of the operation.” With regards to our unified display solution he stated that the “integration of other modalities, that are frequently used or will be used intraoperatively, like ultrasound, navigation, intraoperative MRI [...] is the logical next step and could prove to be a valuable addition and make surgery more accurate and safe.” He also mentioned the “limited space around the surgical workspace that has to be shared with a number of medical devices also favours a ‘one-display solution’.”

4 Conclusion

We have presented the concept of a unified, workflow-driven display system in the OR, that not only combines the various imaging sources available during a surgery, but is also able to switch between them dynamically based on the current surgery phase. In addition, it enables the sterile surgeon to directly control specified device parameters.

Combining all imaging sources to a single display increased the ease of use and the ubiquity of those sources, while the room around the operating table is freed up, even with a large number of employed devices. Adjusting the shown image on the display to the current workflow phase assures the immediate access to the needed information at all times and reduces the need for interaction with the system. Apart from that the surgeon retains direct control of the system, its parameters and possibly the connected devices while satisfying all the necessary sterility requirements. Additional input methods, like different gesture recognition methods, are easily integrated into the system.

While it is clear that additional work will be required in the areas of surgical workflow detection, medical user interface design and efficient data transmission, our work can be seen as a general proof of concept for how the room for imaging guided surgeries can look in the near future.

5 Acknowledgements

This work was supported by the Deutsche Forschungsgesellschaft (DFG), project NA 620/11-1. Further thanks are due to Ash Okur for her valuable help with the experiments.

References

1. Neumuth, T., Mansmann, S., Scholl, M.H., Burgert, O.: Data Warehousing Technology for Surgical Workflow Analysis. *IEEE* (June 2008)
2. Padoy, N., Blum, T., Ahmadi, S.A., Feußner, H., Berger, M.O., Navab, N.: Statistical modeling and recognition of surgical workflow. *Medical image analysis* (December 2010) 1–22
3. Bouarfa, L., Akman, O., Schneider, A., Jonker, P.P., Dankelman, J.: In-vivo real-time tracking of surgical instruments in endoscopic video. *Minimally invasive therapy & allied technologies : MITAT : official journal of the Society for Minimally Invasive Therapy* **21**(3) (May 2011) 129–134
4. Blum, T., Padoy, N., Feußner, H., Navab, N.: Modeling and online recognition of surgical phases using Hidden Markov Models. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* **11**(Pt 2) (January 2008) 627–35
5. Blum, T., Feußner, H., Navab, N.: Modeling and segmentation of surgical workflow from laparoscopic video. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* **13**(Pt 3) (January 2010) 400–7
6. Nara, A., Izumi, K., Iseki, H., Suzuki, T., Nambu, K., Sakurai, Y.: Surgical workflow analysis based on staff's trajectory patterns. In: *M2CAI workshop, MICCAI, London*. (2009)
7. Lin, H.C., Hager, G.D.: User-Independent Models of Manipulation Using Video Contextual Cues. In: *M2CAI-Workshop*. (2009)
8. Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., Fuchs, H.: The office of the future. In: *Proceedings of the 25th annual conference on Computer graphics and interactive techniques - SIGGRAPH '98, New York, New York, USA, ACM Press* (July 1998) 179–188
9. Stewart, J., Bederson, B.B., Druin, A.: Single display groupware. In: *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99, New York, New York, USA, ACM Press* (May 1999) 286–293
10. Paek, T., Agrawala, M., Basu, S., Drucker, S., Kristjansson, T., Logan, R., Toyama, K., Wilson, A.: Toward universal mobile interaction for shared displays. In: *Proceedings of the 2004 ACM conference on Computer supported cooperative work - CSCW '04, New York, New York, USA, ACM Press* (November 2004) 266
11. Hanna, G.B., Shimi, S.M., Cuschieri, A.: Task performance in endoscopic surgery is influenced by location of the image display. *Annals of surgery* **227**(4) (April 1998) 481–4
12. Meyer, M.a., Levine, W.C., Egan, M.T., Cohen, B.J., Spitz, G., Garcia, P., Chueh, H., Sandberg, W.S.: A computerized perioperative data integration and display system. *International Journal of Computer Assisted Radiology and Surgery* **2**(3-4) (August 2007) 191–202

13. Soehngen, E., Rahmah, N.N., Kakizawa, Y., Horiuchi, T., Fujii, Y., Kiuchi, T., Hongo, K.: Operation-Microscope-Mounted Touch Display Tablet Computer for Intraoperative Imaging Visualization: Technical Note and Comparison with Other Modalities. *World Neurosurgery* (2011)
14. Gallo, L., Placitelli, A.P., Ciampi, M.: Controller-free exploration of medical image data: Experiencing the Kinect. In: 2011 24th International Symposium on Computer-Based Medical Systems (CBMS), IEEE (June 2011) 1–6
15. Soutschek, S., Penne, J., Hornegger, J., Kornhuber, J.: 3-D gesture-based scene navigation in medical imaging applications using Time-of-Flight cameras. In: 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE (June 2008) 1–6
16. Schwarz, L.A., Bigdelou, A., Navab, N.: Learning gestures for customizable human-computer interaction in the operating room. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* **14**(Pt 1) (January 2011) 129–36
17. Chuang, Y.C., Hsu, P.Y., Wang, M.T., Chen, S.C.: A Frequency-Based Algorithm for Workflow Outlier Mining. *Future Generation Information Technology* (2010) 191–207
18. Paepke, S., Schwarz-Boeger, U., Kiechle, M., Jacobs, V.R.: Axillary Dissection with Access Minimized (ADAM): a new technique for lymph node dissection in conservative surgery for breast cancer. *International journal of fertility and women's medicine* **48**(5) (2003) 232–7