Editors:
Larry Rosenblum and Simon Julier

Action- and Workflow-Driven Augmented Reality for Computer-Aided Medical Procedures

Nassir Navab, Joerg Traub, Tobias Sielhorst, Marco Feuerstein, and Christoph Bichlmeier

CAMP, Technische Universität München, Germany

In the history of computer science, the 1990s could be considered the point when researchers defined augmented reality (AR) and presented early ideas for solving practical problems in several application fields, including computer-aided diagnosis and surgery. At the earliest phases of AR's development, the scientific community focused mostly on establishing the fundamental technology as well as defining the new scientific field, its appropriate ontology, and its interaction with other established fields, such as virtual reality (VR), graphics, user interface design, and computer vision. By the start of the new millennium, basic tracking, visualization, and display technology had advanced enough to enable real-time augmentation of static scenes. However, the major promise of AR in allowing its users to function within their real environment could be advantageous only if it were seamlessly integrated into the workflow.

AR should augment users when needed; it shouldn't diminish users' efficiency during the rest of the workflow. If AR forces the user to leave the desired workflow and working environment entirely and immerse into a new environment for the sake of augmentation, it will lose its main advantage. In fact, users often need augmentation for only short periods of time; they can then continue working in their real environment for an extended time before needing another period of intelligent augmentation. For example, many surgical procedures could benefit from valuable in-situ visualization of preoperative planning at the beginning of the intervention. This visualization takes only a few minutes, but it could play an important role in the surgery's success. However, if this augmentation requires extensive online calibration, wearing cumbersome sensors and displays, or using large tracking equipment (reducing surgeons' working space), it might negatively affect other phases of the surgical procedure.

Therefore, one key to the success of a user interface including AR visualization is its ability to automatically recognize different phases of a workflow, which each require various levels of augmentation. It is also important for the AR system to be transparent to the user during the rest of the procedure—that is, it shouldn't complicate or disturb the rest of the workflow. These issues have greater importance when dealing with computer-aided surgery applications. In most of these applications, a surgeon needs augmentation for only quite brief periods, such as choosing the ports for a laparoscopic intervention or localizing the major arteries before starting a liver resection. These augmentations, however, can play an important role in the overall procedure's success.

During the past three years, we've tried to develop such integrated AR solutions in the context of minimally invasive surgery. We have therefore focused on four main issues:

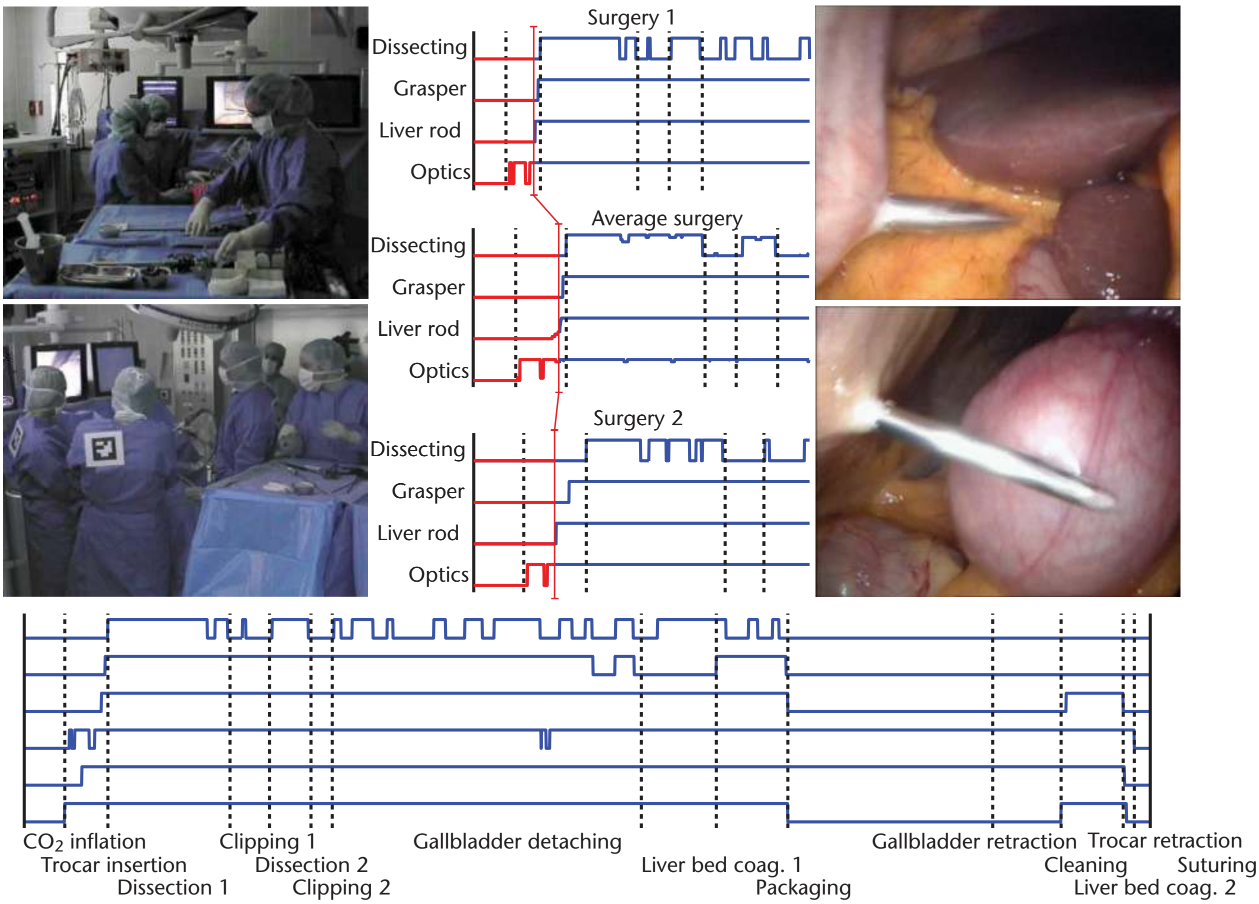
- recovery and monitoring of surgical workflow,
- integrating preoperative and intraoperative anatomic and functional data,
- improving visual perception in a mixed environment, and
- developing new user interaction paradigms for taking full advantage of the virtual data, while overlaid onto the real scene.

Each of these issues is the subject of many existing and future publications. Here, we provide a brief overview of our activities and current results in regard to each of these issues.

Recovery and monitoring of surgical workflow

For AR or VR to be integrated into a surgical procedure successfully, physicians must be able to automatically call on intelligent augmentation in the right format and at the right time during a surgical procedure. To achieve this objective, we need to recover and understand the workflow, model it appropriately, and then monitor it to develop action- and workflow-driven, adaptive, dynamic user interfaces. We accomplish this goal by recording as many relevant parameters as possible during a set of interventions of the same kind.[1] This can be done by acquiring the available signals within the surgery room, such as endoscopic camera images and electrocardiogram signals, and by using additional external sensors, such as external cameras, RFID, or magnetic or inertial sensors attached to tools or surgical staff. Each surgery is represented by a variable length matrix comprising all recorded signals (or representa-

 Published by the IEEE Computer Society



1 Recognition of actions and workflow phases is necessary for allowing an augmented reality system to provide the right information at the right moment and in the most suitable form.

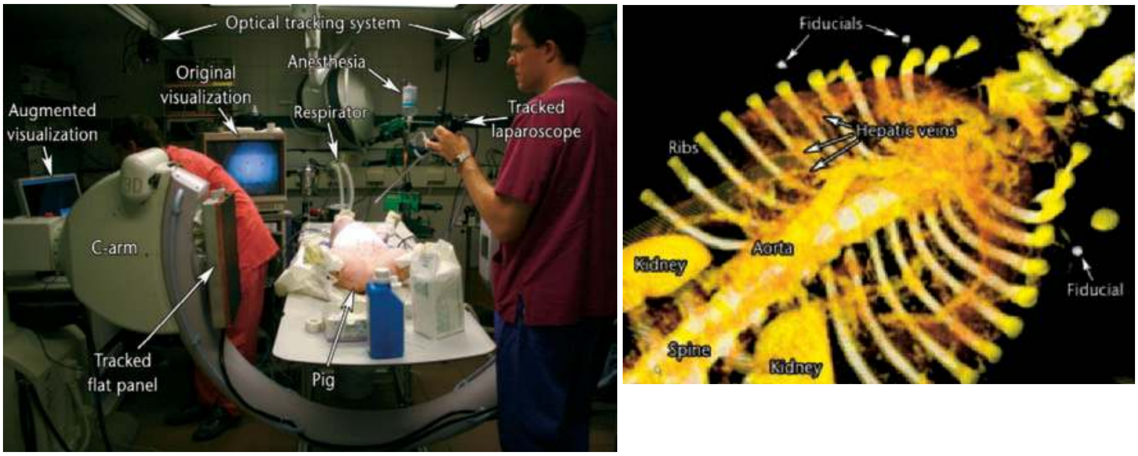
tive functions of them) over time (see Figure 1). If N is the number of observed parameters—for example, representing the use of different surgical tools and the specific parameters of electronic devices, or patient's health monitoring parameters such as blood pressure—and TS is the variable duration of a particular surgery S in units of time representing our sampling rate, each surgery is represented by an $N \times TS$ matrix of sensory data.

We study different methods for synchronizing multiple surgeries, segmenting surgeries into distinguished phases, and defining a virtual average surgery that represents one kind of surgery.[1] Signals acquired intraoperatively let us predict the current state of the surgery and trigger adapted visualization and user interfaces at each moment and for each task. We can augment the reality only if we know what it is and how it needs to be augmented. This information is different at each moment of the surgery and depends on surgical actions and the current state of important parameters in a surgical procedure, such as patient monitoring data, endoscopic videos, and active tools.

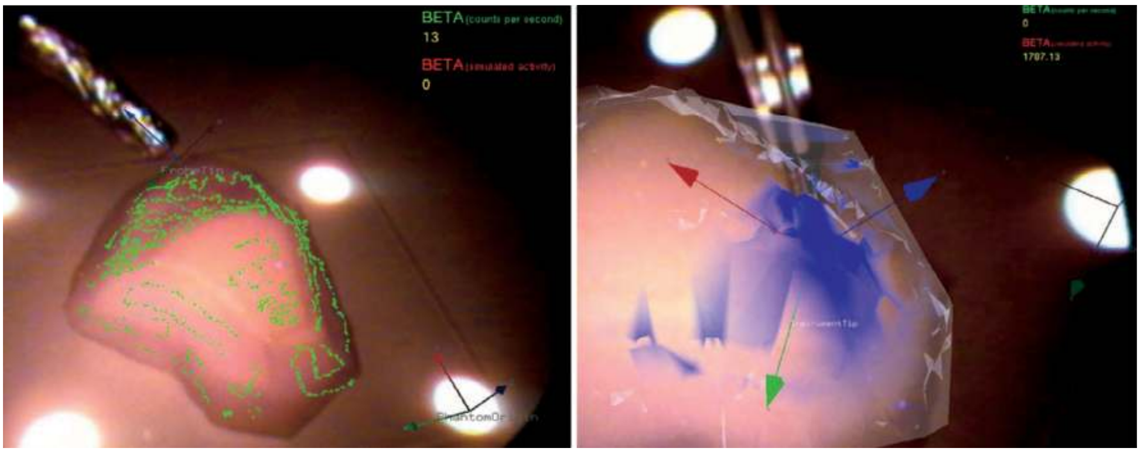
Integrating preoperative and intraoperative anatomic and functional data

Once we understand the medical workflow, the system must present the required information in its most intuitive and suitable form. Physicians often use preoperative images to plan the surgery. Intraoperatively, they need to update this planning, taking into account the possible modifications of a patient's anatomy, for example, due to the progress of a disease or preinterventional treatment such as chemotherapy or radiation therapy, the current filling of the stomach or bladder, or the body's rigid and deformable movements caused by such factors as the patient's positioning or breathing state.

The problem is almost solved for rigid anatomy. Commercially available systems exist for orthopedics as well as brain surgery, in which rigidity could be considered a reasonable assumption for several surgical procedures. Products register the preoperative and intraoperative data in this case. However, the problem is still open for deformable organs, such as in abdominal interventions. In such cases, even if deformable registration could provide some solutions, accuracy in the outcome is extremely hard to guarantee. An accurate registration requires detailed segmentation and physical modeling of organs, in terms of their deformability. This work is extremely complex in general and not always possible for pathological tissues caused by the disease's progress, which have different physical properties than the healthy tissue. A possible remedy is the use of intraoperative anatomical imaging or, even better, functional imaging. Although anatomical imaging enables visualization of a patient's anatomy, functional imaging focuses on visualization of pathologies, which are often the main target in surgical procedures.



2 Animal study. The surgeon moves the endoscope to a candidate position for port placement. The system presents a virtual fly-through into CT data corresponding to candidate port location.

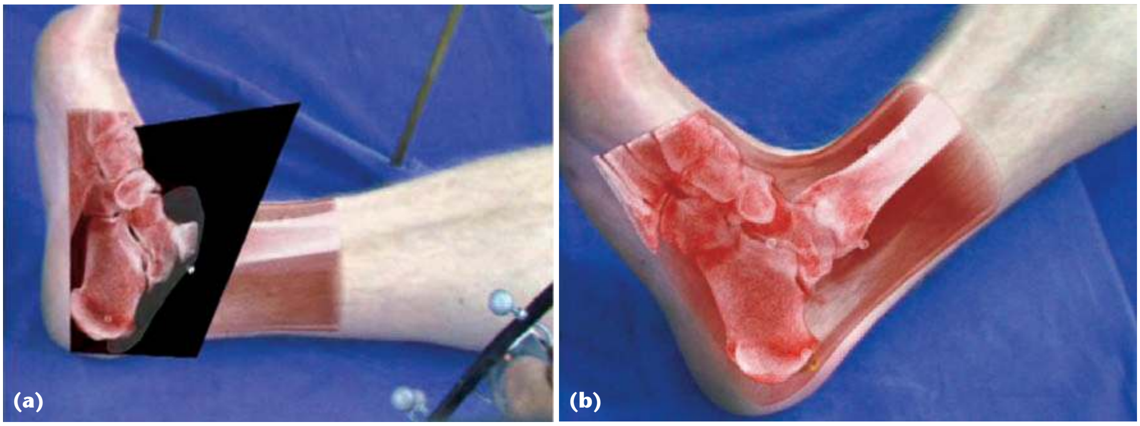


3 AR for intraoperative functional imaging: The surgeon moves the functional probe on the anatomy of interest. The system records the functional data synchronized with position information. Navigated functional imaging and AR visualization open new possibilities for open and minimally invasive surgery.

We believe that the future of surgery is in using both advanced surgical tools capable of functional imaging and advanced visualization systems to provide intuitive in-situ visualization of anatomical and functional data. In recent years, we've proposed several medical AR solutions. In particular, we've focused on minimally invasive surgery, in which the indirect view of the surgical site through an endoscopic camera makes the use of AR technology natural and intuitive.

One of our AR applications assisted the surgeon in port placement, which is one of the crucial steps of any keyhole surgery defining the access point to and working environment within the surgical volume.[2] In this application, through registering preoperative CT and intraoperative endoscopic camera views, we can superimpose the virtual CT data onto the patient's anatomy. In addition, we let the surgeons virtually fly through the potential port and observe the resulting surgical working space. Figure 2 shows a surgeon validating this solution through an animal study. Once port placement is complete, the patient's abdomen is inflated with carbon dioxide, resulting in extreme deformation of anatomy compared to the preoperatively acquired CT scan. Next, we proposed to track an intraoperative x-ray C-arm capable of cone-beam reconstruction in the same coordinates as the endoscopic camera. The C-arm rotates around the patient in exhalation, reconstructing the contrasted blood vessels. (For cone-beam reconstruction, the x-ray system requires hundreds of images taken from different angles around the anatomical target. This acquisition takes less than a minute. The reconstruction is valid, if the organ can be considered static. Therefore, neglecting the small motions due to heartbeat, the image acquisition is done during the exhalation.) As soon as the physician inserts the endoscope into the surgical area, the endoscope is augmented by the reconstructed blood vessels.[2] Gated by active breathing control to augment only during exhalation, this augmentation could enable the surgeons to update their surgical planning.

In our more recent project, a combination of tracked ultrasound and gamma probe guides the surgeon to malignant tumors (see our work in the proceedings of the 2007 Medical Image Computing and Computer-Assisted Intervention). This is one of the first medical AR applications to combine anatomical and functional imaging during surgery. Finally, after resectioning



4 Medical AR. EVI software system enables in-situ visualization of medical imaging data using the Ramp head-mounted display system. Thanks to a wireless mouse, the user can choose the visualization mode or move through different CT slices. (a) Volume rendering is combined with (b) slice view.

malignant tumors, physicians could use AR for intraoperative visualization of any residuals that a tracked beta-probe[3] detects (see Figure 3).

Improving visual perception in a mixed environment

One of the major issues in AR is correct depth perception in the presence of both virtual and real objects. In our laboratory, we developed a medical AR visualization package called EVI (easy visualization in-situ), using the Ramp system originally developed by Sauer and Khamene at Siemens Corporate Research. EVI was demonstrated at the German conference on orthopedics and trauma surgery in Berlin in October 2005. Figure 4 shows the CT image of a patient superimposed on direct views of his right foot seen through a stereo video see-through. Thanks to a wireless mouse, the user can select the desired visualization mode. (In this figure, both volume rendering and slice views are used.) The subject can displace his foot, which had small retroreflective markers attached to it, and the observer could also move around and manipulate the foot.

This real-time demonstration excited many surgeons. Using this visualization system, our surgical partners all believed that they could perform different surgical procedures when the target region is tracked or fixed. Therefore, we proposed that they do cadaver studies to confirm the system's usefulness. During a scheduled cadaver-based surgical training procedure, three surgeons tried using the video see-through system for intramedullary (IM) nail and pedicle screw placement procedures. Even when the surgeons found the system intuitive, they did not have the success they would have had using standard 2D navigation tools. This negative result was useful and motivated the surgeons and the engineers to find appropriate solutions.

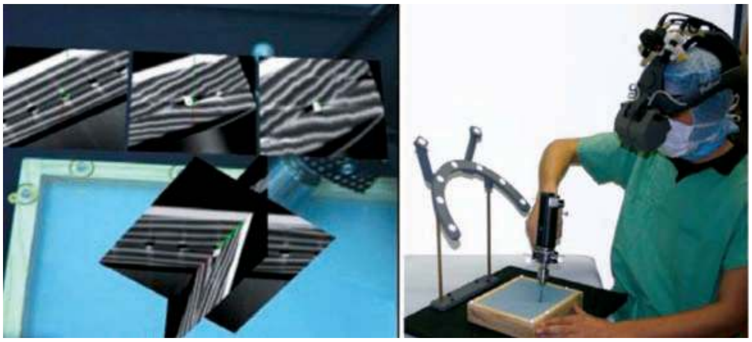
We therefore evaluated different visualization techniques through a series of studies with 20 surgeons. The results were published and presented at the 2006 Medical Image Computing and Computer-Assisted Intervention Conference. We jointly came to the conclusion that, for specific surgical tasks, hybrid user interfaces that appropriately combine orthogonal slice views and in-situ visualization could provide new optimal solutions.[4]

We continue developing and evaluating new visualization paradigms to appropriately integrate the virtual and real views, taking the objectives and actions of the surgeons into account. We therefore propose a set of workflow- and action-driven visualization techniques to improve surgeons' depth perception in a mixed environment (see Figure 5).

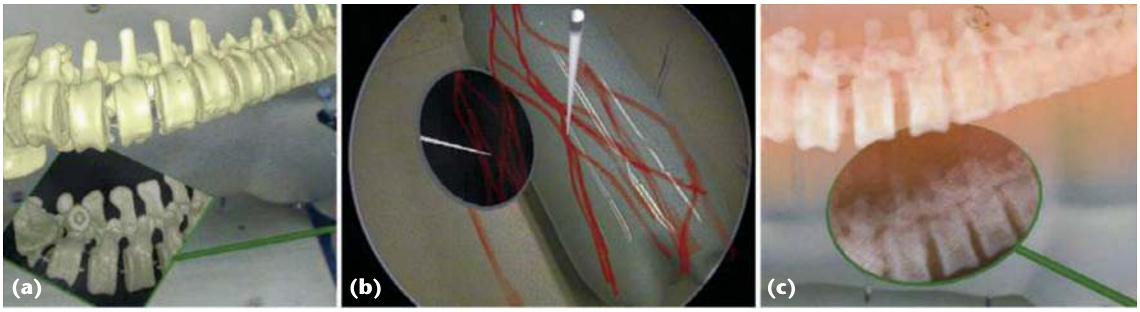
Developing new user-interaction paradigms for taking full advantage of AR visualization

3D virtual data proves its advantages only when the user interacts with the data. In traditional VR applications, this interaction is possible by moving the 3D object, changing the viewpoint in the virtual world, or letting the tracked user move around the virtual object. However, as soon as we are in an AR environment, the first two options are often not applicable. In medical AR applications, the user can move neither the patient undergoing surgery nor the virtual objects, which are aligned to the patient's anatomy. The surgeon's movement is also limited.

In the existing AR systems, the surgeons were therefore limited to the 3D perception obtained only through



5 A hybrid interface lets surgeons take advantage of both in-situ visualization and standard slice-based navigation user interface concepts. Three surgeons with different levels of expertise participated in a quantitative evaluation of the efficiency and accuracy of the proposed hybrid solution.



6 **The virtual mirror penetrating the 3D virtual space could reflect (a) surfaces or (c) rendered volumes, providing desired views of the 3D object from any viewpoint. In augmented laparoscopic surgery, the virtual mirror could provide additional views, solving the (b) 3D ambiguities of 2D projections. It also reflects the virtual models of tracked surgical instruments further improving hand-eye coordination.**

stereo displays. Manipulating the 3D object and observing it from all viewpoints was therefore impossible while preserving the real/virtual alignment. We achieved a breakthrough by proposing the use of an interactive virtual mirror that lets the user take full advantage of 3D data without losing the main advantages of in-situ visualization and real/virtual alignment.[5] Using such a virtual mirror is easy, since the concept is extremely familiar to the users. The virtual mirror is an interactive device, which, thanks to both visual perception and proprioception, allows its users to perceive well the 3D structures and study them without changing their viewpoint or moving the virtual objects (see Figure 6). We believe that this interactive virtual device will become an integrated part of many AR systems in the near future. In upcoming publications, we will present this concept in detail, propose and test different applications, and provide detailed user studies that demonstrate the high acceptance of such a virtual AR interaction device. ■

References

1. A. Ahmadi et al., "Recovery of Surgical Workflow without Explicit Models," *Proc. Medical Image Computing and Computer-Assisted Intervention* (MICCAI), LNCS 4190, Springer, 2006, pp. 420-428.
2. M. Feuerstein et al., "Intra-Operative Laparoscope Augmentation for Port Placement and Resection Planning in Minimally Invasive Liver Resection," to appear in *IEEE Trans. Medical Imaging*, 2007.
3. T. Wendler et al., "Navigated Three Dimensional Beta Probe for Optimal Cancer Resection," *Proc. Medical Image Computing and Computer-Assisted Intervention* (MICCAI), LNCS 4190, Springer, 2006, pp. 561-569.
4. J. Traub et al., "Hybrid Navigation Interface for Orthopedic and Trauma Surgery," *Proc. Medical Image Computing and Computer-Assisted Intervention* (MICCAI), LNCS 4190, Springer, 2006, pp. 373-380.
5. N. Navab, M. Feuerstein, and C. Bichlmeier, "Laparoscopic Virtual Mirror: New Interaction Paradigm for Monitor Based Augmented Reality," *Proc. Virtual Reality Conf.* (VR), IEEE Press, 2007, pp. 43-50.

Contact author Nassir Navab at navab@cs.tum.edu.

Contact the department editors at cga-vr@computer.org.