Voice controller for Image Guided Surgery and personalized Interactive Visualization

Andoni BERISTAIN[abc], Alessandro DE MAURO[ab], Koen VAN DE WEYER[d], Dominique SEGERS[d]

[a]Department of eHealth and Biomedical Applications, Vicomtech-IK4
[b]Health Group, Bioengineering Area, Biodonostia Health Research Institute
[c]Computation Intelligence Group, Basque University (UPV/EHU)
[d]Barco N.V.

aberistain@vicomtech.org

Abstract. With the increasing availability of multimodal data, an important current topic is to provide surgeons with tools for the visual exploration of the patient data according to their own preferences and experience. In the operating theatre very rigid visualisation modalities are normally available and each of them is usually dedicated to one specific aspect of the operation workflow. A need of more personalised and adaptable solution is clearly identified in many types of surgery. This paper presents a system for a personalized visualization for each single intervention step. It permits to plan the visualization preferences and configure different monitors assigning several input sources. During the operation, natural user interfaces permit navigating among all the pre-set configurations. The natural interaction is achieved by means of a combination of speech and multi-touch user interface (UI).

1 Introduction

In the history of medicine, healthcare professionals have relied on their own senses to diagnose illness, monitor a patient's condition and perform invasive procedures. Nevertheless, during the last few decades, various three-dimensional (3D) medical imaging techniques, such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound, have become very effective tools for diagnosis, planning and intraoperative image modalities. Thanks to these technological advances it is now feasible to perform minimally invasive surgery instead of the traditional open surgery for many cases. The advantages of this kind of intervention are numerous:

higher productivity and efficiency, mortality rate reduction, reduced post-intervention issues and side effects, as well as a reduction in the hospitalization period.

The work presented in this paper has been carried out in the context of a European project called MEDIATE (Patient friendly medical intervention). The goal of the MEDIATE project (funded by ITEA2) is to increase productivity and effectiveness in healthcare and reduce patient risk by supporting healthcare professionals in the transition from invasive, open surgery to minimally invasive, image guided intervention and treatment (IGIT).

The world is facing several healthcare related trends: global economic growth and changes in demographics (aging world population), leading to surging healthcare costs. Combined with staffing shortages in personnel, there is a clear need for new methods to handle more patients within acceptable healthcare costs, while ensuring a high quality of care. Improved productivity and effectiveness in healthcare enabled by IGIT is expected to significantly lower healthcare costs due to shorter hospital stay and higher throughput.

IGIT involves medical diagnosis, planning and treatment of patients by minimally invasive placement of diagnostic and therapeutic devices such as catheters and stents inside the human body, enabled by medical image analysis and navigation methods. IGIT helps healthcare professionals to obtain a better clinical outcome of the treatment, a predictable procedure time, fewer complications, better service to the patient and lower morbidity & mortality rates. In addition, IGIT is considered to lead to a prolonged life, an improved quality of life and less discomfort for the patient.

From the practical perspective, one of the objectives of the project is to provide the surgeons with advanced medical interfaces for the interaction with the complex workflow that image guided surgery requires. Specifically, the targeted interventions are those related to cardiovascular, oncologic and orthopaedic areas.

In this context, the current paper describes the experimental use of natural user interfaces along the surgical workflow with the final target of optimizing the intraoperative visualization procedures in IGIT. A system has been developed which permits the surgeon to easily select what each of the video displays available in the operating room (OR) must show for each intervention step using simple voice commands.

This paper is structured as follows; first the related background is presented in section two. Then the developed system is described in sections three and four. Section three focuses on the functional description, while section four describes the implementation and technical aspects. The paper finishes with some conclusions and future work presentation.

2 Background

This section presents a review of User Interfaces (UI) which have been used in the medical domain, more specifically, Natural User Interfaces, as well as some related developments.

2.1 User Interface Requirements in the Operating Room (OR)

Not all the traditional user interfaces are suitable for the medical field. For example, keyboard and mouse are not well suited for surgical environments because: they are not appropriate for 3D interaction, they cannot be placed in the surgeon workspace and they are obstructive for the surgeons.

For intra-operative surgery assistance, there are some particular characteristics to be considered. First, sterilization of the operating room, including the entire environment and all the equipment. Then, no discharged air of computers or projectors is allowed. The possibility of electromagnetic interference (tracking systems vs. imaging devices) needs to be considered too, as well as avoiding clutter and maximizing ergonomics.

The UI design regarding usability aspects and in general for medical devices is a crucial aspect that has to be considered in any medical development [3].

A recently published international standard (ISO/IEC 62366: Application of Usability Engineering to Medical Devices) requires manufacturers of medical devices to follow a systematic usability process and a user centric design.

The documents is particular important because is setting the direction for the usability of medical user interfaces. In particular it describes a Usability Engineering Process which is an iterative process composed of nine stages:

1. Specify the application of the medical device.
2. Identify the device’s frequently used functions.
3. Identify hazards and hazardous situations related to usability.
4. Identify the device's primary operating functions.
5. Develop the usability specification.
6. Prepare the usability validation plan.
7. Design and implement the user interface.
8. Verify the user interface design.
9. Validate the usability of the medical device.

2.2 Natural User Interfaces (NUI)

Natural user interfaces (NUI) refers to a user interface that is effectively blended into the environment, or becomes part of it with successive learned interactions, to its users. Most computer interfaces use artificial control devices whose operation has to be learned and only take advantage of a little part of the humans’ interaction capabilities. The most common approach of that is mouse and display based UI, which according to O’Sullivan [1] would make users look as a hand with one finger, one eye and two ears, from the computer perspective.

NUI tries to go a step further by taking advantage of the innate expressive capabilities of human beings. In this case the goal is to make the user feel like a natural interaction. The learning curve for this type of interfaces is so short that the user feels to be instantly and continuously successful. From the technological point of view, all these devices allow the users to carry out relatively natural motions, movements or gestures. The conventional approaches use hand gestures and touch, speech and eye gaze to interact with the underlying system in a way that tries to emulate common everyday life interaction situations. For some further discussion on NUI, focused on multi-touch UI, the reader is referred to [2].

In the medical field, voice-activated control systems are being introduced into clinical practice, for minimally invasive surgery. The utility of the voice-activated control system in this OR has already been objectively evaluated [5] with the conclusion that improves communication efficiency among OR staff. Surgeons are afforded the most timely equipment adjustment possible while circulating nurses can concentrate on patient care instead of equipment adjustment during the course of the surgery. Luketich [6] also assesses the impact of this type of systems on operating room efficiency and user satisfaction concluding that physician and nurse acceptance of it was very high because of the smoother interruption-free environment.

SidneHD device control (by Stryker) is platform designed to network the OR, integrating surgical devices, which can be, controlled by simple voice commands. Many pieces of surgical equipment are outside the range of sterility for the surgeon and must be manipulated by a surgical staff while SidneHD enables all needed equipment to be directly under the surgeon's control. The staff has the freedom to use a centralized touch panel, a hand-held tablet or voice recognition to control a wide range of surgical equipment (such as tables, lights, video cameras). Some other current implementations of speech UI in clinics are presented in part III of [4].

Multi-touch devices are the current trend for many equipment in the OR because of the flexibility of controls, familiarity of the surgeons (e.g. smartphones) and because they are either simple to sterilize. For an incomplete chronology of multi-touch techniques, readers are referred to [7]. And for some advices on interface design the user is referred to [8].

2.3 Related systems

State of the art visualization systems for this type of procedure (i.e. Siemens Artis Zeego, GE Healthcare Valve Assist or Philips HeartNavigator) provides multiple view inputs simultaneously and over 200 layout configurations (each of the systems) on several screen combinations.

Both the surgeon needs of the customization of these complex configurations and the need of a fast and simple way to switch from one to another during the procedure are the main motivations to our research and development.

3 Methods

During a surgery intervention in the Operating Room, especially for delicate interventions, the surgeon needs to have at hand the most relevant visual information at every time, which is depicted on the available displays at the OR. In addition, the relevant information depends on the particular step of the intervention. Traditionally, the process of changing what is depicted on every display in the OR according to the particular step of the intervention was performed by the surgeon's assistant, following the surgeon's instructions. This can be a slow, tedious and error prone procedure because the surgeon needs to explain to his assistant what he wants to view and how for every display. The prototype system presented in this paper tries to simplify this process and make it quicker.

The development presented in this paper is used in two different stages, the pre-operative stage and the intra-operative stage. During the pre-operative stage the surgeon prepares the display setup using a multi-touch user interface (UI) and then, during the intra-operative stage he uses voice commands to navigate among the different display setups for each intervention step as presented in **Fig. 1**.

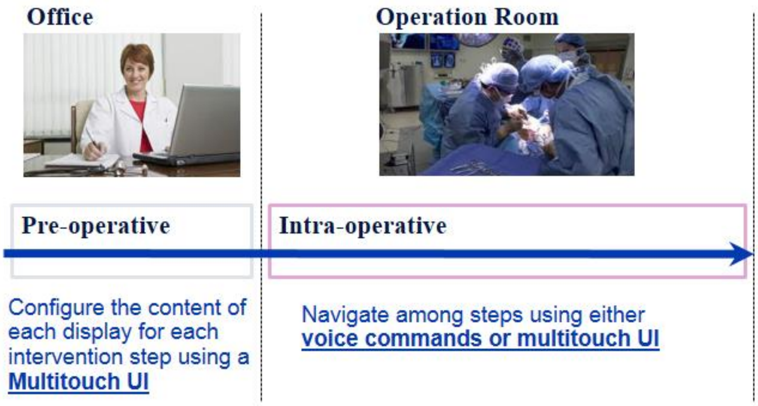

Fig. 1. Prototype usage in the different intervention stages

After consulting the surgeon requirements, expectations, ergonomics issues and environmental constraints, this set of considerations have been taken into account for the UI selection:

- The surgeon is using both hands for manoeuvring catheters and looking up at the displays for hand-eye coordination. UI devices should preferably be operable without the need to use the hands or at least be findable and operable without looking at them
- Procedures can be quite long (up to several hours) and tiring.
- The sterile drapes are bunched up and removed from the room after the case. In practice, small devices like remote controls etc. are prone to be discarded with the drapes. Low-cost disposable devices might be a solution here or RFID tags to detect valuables in the drapes

Users of medical devices are also users of consumer devices. Therefore, they have certain expectations from the way they interact with for instance mobile phones, tablet PCs, computers and gaming consoles. Creating a new paradigm only for the medical world is not going to work; the concepts that are introduced should align with the expectations arising from these much wider consumer domains. As a result of these considerations, the main objective has been to minimize the surgeons' strain during the interventions, simplifying the control and giving them a kind of UI which they are comfortable with (i.e. smartphones' and tablets' multi-touch UI and speech commands to their assistants).

During the preoperative phase surgeons can use a multi-touch UI application to set the surgical layout (**Fig. 2**) (which input video sources wants to view on each available display, for every step in the intervention). At a lower level, the application permits to create a layout which configures which video input sources need to be depicted on each display monitor available in the OR.

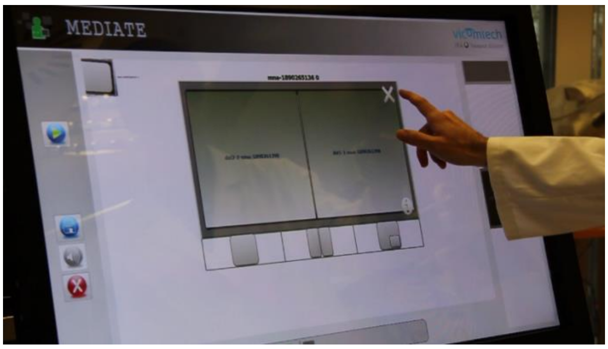

Fig. 2. Multi-touch application GUI

This system permits to depict up to two different video sources on a single display monitor using either a Picture and Picture (PaP) or Picture in Picture (PiP) arrangement (**Fig. 3**). Once the sequence has been created by the surgeon or his assistant, it can be stored, so that it can be used later during the intervention. For a given distribution, the sequence is created once, but then it can be used as many times as needed.

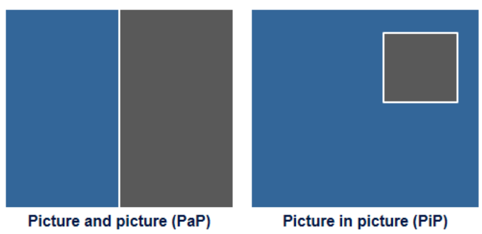

Fig. 3. PiP and PaP modalities

During the intervention, once the sequence has been loaded, the surgeon interacts with the sequence of previously defined display layouts using voice commands. Eventually, during the intervention, the surgeon may need a special arrangement in a particular case. In that situation, the assistant can still create an ad-hoc layout, using a multi-touch application similar to the one used during the pre-operative stage and following the surgeon's instructions.

The developed software application is relying on the state-of-the-art solution for operating room management API of Nexxis® (Barco). This is a fully IP-centric solution for image distribution in the operating room that has been integrated with an adhoc workflow management system and voice control.

3.1 Practical use case, Transcatheter aortic valve implantation (TAVI)

Transcatheter aortic valve implantation (TAVI) has been one of the practical intervention application cases in the Mediate project. And the prototype presented in this paper has been tested for that particular intervention workflow.

Aortic stenosis is the most common valvular lesion occurring among elderly patients and has become extremely frequent because of changing demographics in industrialised countries. Surgical risk after the age of 70 has increased. Transcatheter aortic valve implantation (TAVI) has emerged as a promising alternative to conventional aortic valve replacement for patients with severe, symptomatic aortic stenosis who are otherwise left untreated due to the perceived high risk of operative mortality.

This procedure consists on the insertion of a new artificial heart valve inside the old tight valve using a balloon catheter. The valve is composed of a metal frame (stent) and the outer lining (pericardium) of a cow's heart. The procedure is carried out under general anaesthetic and can be carried out using a transfemoral (through the femoral artery) or transapical (through the left side of the chest) access approach.

The procedure is done in the X-ray operating room using contrast dye and echocardiography to guide the valve into the correct position. The total duration of the procedure is usually around 1.5 hours.

Like the majority of surgeries, TAVI workflow can be parted in three main phases as reported in [11]. **Fig. 4** depicts a simplified version of this workflow.

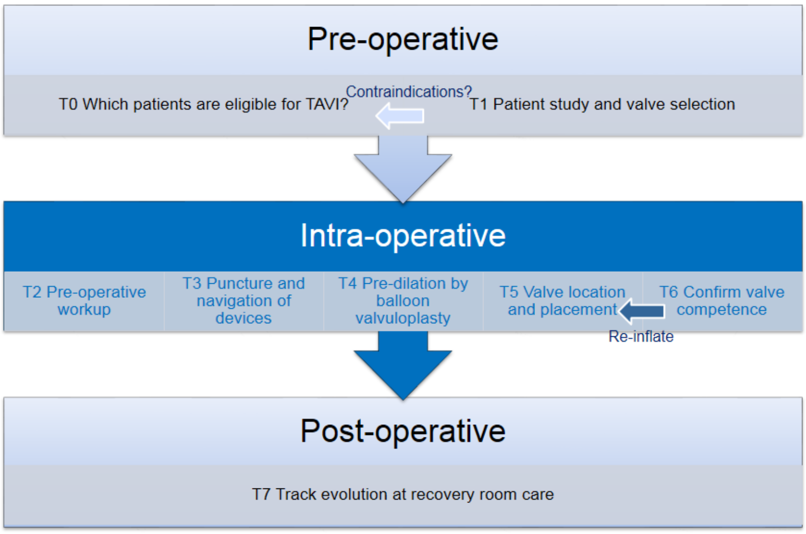

Fig. 4. TAVI workflow

3.2 Interaction with an intra-operative clinical decision support system (CDSS)

During the Mediate project the eXit research group at the University of Girona (Spain) developed a clinical decision support system (CDSS) for the TAVI intervention, which covered the pre-operative, intra-operative and post-operative stages.

The integration of all the systems involved in the intervention workflow was also one of the requirements from the medical perspective. Therefore, a preliminary integration of both prototypes was performed. This integration permits the GUI from the CDSS to be shown in one of the displays in the OR during an intervention. But more importantly, when the surgeon uses a voice command to navigate among the different steps in an intervention, the CDSS is notified of this change, so that it provides support for the current step. This combined prototype seems to be a very promising tool for the surgeons.

4 Materials

As it was presented before, the developed prototype system permits to personalize which video sources are depicted on every display in the operating for every phase in an intervention and then, to navigate among this set of configurations during the intervention using voice commands and a multi-touch application operated by an assistant when necessary.

The created software prototype seats on top of a system developed by BARCO called Nexxis OR Management Suite [9, 10]. The Nexxis Suite is composed of several hardware components: encoders, which convert the analog or digital video signal into network packages, decoder which do the opposite operation, a cable network, a network switch and a manager. The network packages travel from a source device (i.e. image production device such as an endoscopic camera) to the destination device (i.e. video display) through a conventional cabled network. Package delivery is controlled by a high-performance network switch, which is managed by a server computer. This manager computer also provides external access to the underlying hardware control through a REST API, which has been used to develop the high level interaction prototype presented in this paper. The architecture of the whole system is presented in **Fig. 5**.

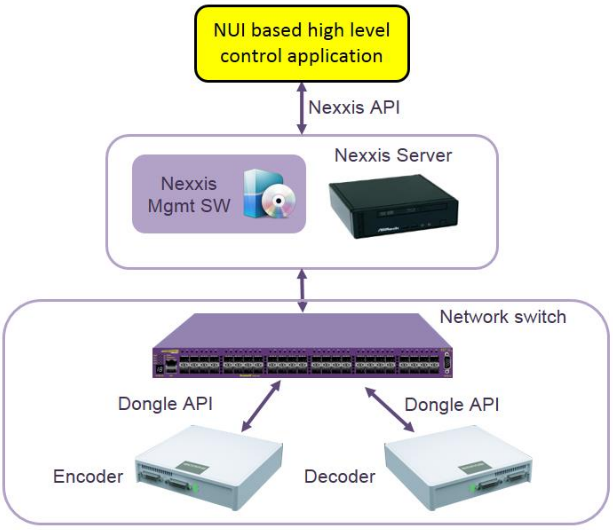

Fig. 5. Prototype system architecture

For the prototype itself C++ has been used as the primary development language. Digia's Qt libraries [12] have been used for the GUI implementation. More specifically, the Qt Quick modules have been used, which permit the implementation of highly interactive GUIs, adapted to multi-touch interaction.

In addition, Microsoft's Speech API has been used for speech recognition aspects, and wireless headsets and microphones have been used for the speech interaction.

5 Conclusions and future work

This paper has presented a prototype system for the simple yet efficient management of video sources and displays in the operating room using natural user interfaces. It combines the use of multi-touch and voice command based UI to support image guided surgical interventions in the operating room. This is part of a complex research analysis on the topic of surgeon interaction in the OR settings and how the traditional protocols are changing due to upcoming interfaces and paradigms.

The use case considered is the TAVI procedure but the principles could be extended to other types of surgery in which a particular stress on smart interaction with the visualisation is needed.

Future work is required to test and validate the prototype in a live surgery.

Acknowledgement

This work has been partially funded by the Mediate project n. 09039 under the ITEA2 (Information Technology for European Advancement), a EUREKA framework strategic cluster.

References

1. O'Sullivan, D. (2004) Physical Computing: Sensing and Controlling the Physical World with Computers. Thomson
2. NUI. (2009). Multi-Touch Technologies. Retrieved from http://nuigroup.com/log/nuigroup_book_1/.
3. Nielsen, J. (2007). How to Kill Patients through Bad Design.
4. Neustein, A. (2010). Advances in Speech Recognition. Springer US.
5. Salama, I. A., & Schwaitzberg, S. D. (2005). Utility of a voice-activated system in minimally invasive surgery. Journal of laparoendoscopic & advanced surgical techniques. Part A, 15(5), 443-6.
6. Luketich, J. D., Fernando, H. C., Buenaventura, P. O., Christie, N. A., Grondin, S. C., & Schauer, P. R. (2002). Results of a randomized trial of HERMES-assisted versus non-HERMES-assisted laparoscopic antireflux surgery. Surgical endoscopy, 16(9), 1264-6.
7. Buxton, B. (2011). Multi-Touch Systems that I Have Known and Loved. Retrieved from http://www.billbuxton.com/multitouchOverview.html.
8. Saffer, D. (2008). Designing Gestural Interfaces: Touchscreens and Interactive Devices (p. 272). O'Reilly Media.
9. NEXXIS: http://www.barco.com/en/Products-Solutions/Networked-solutions/Digital-operating-room/IP-imaging-and-audio-management-platform.aspx
10. NEXXIS White Paper: http://www.barco.com/~/media/Downloads/White%20papers/M-P/Whitepaper_Nexxis_TCO_synopsis%20pdf.PDF
11. El-Fakdi, A. & Meléndez, J. & Gamero, F. & Petite, R (2013). EXITCDSS: A framework for a workflow-based CBR for interventional CDSS, Application to Transcatheter Aortic Valve Implantation (TAVI), BIOSTEC 2013- HEALTHINF International Conference on Health Informatics.
12. Digia Qt software libraries. Content last accessed in February 2014. http://www.qt-project.org