What is the future of minimally invasive surgery in rhinology: marker-based virtual reality simulation with touch free surgeon's commands, 3D-surgical navigation with additional remote visualization in the operating room, or ...?

Ivica Klapan1*, Alen Duspara2, Zlatko Majhen3, Igor Benić4, Milan Kostelac5, Goranka Kubat6, Nedjeljka Berlengi7

1School of Medicine, Josip Juraj Strossmayer University in Osijek, Osijek, Croatia, EU.
2Klapan Medical Group University Polyclinic, Zagreb, Croatia, EU.
3Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia, EU.
4Avanza, Zagreb, Croatia, EU.
5Juraj Dobrila University, Pula, Croatia, EU.
6Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Zagreb, Croatia, EU.
7Sunce Polyclinic, Zagreb, Croatia, EU.

Abstract

The navigation-Leap Motion-virtual endoscopy and virtual surgery suggests that real and virtual objects definitely need to be integrated by use of real 'in the air' control with simulation of virtual activities that requires real-time visualization of 3D-virtual endoscopy motions, following the action of the surgeon that may be moving in the virtual reality area. It is achievable with simple hand gestures of the surgeon, which do not differ in any way from all other natural everyday hand or body movements. In this way, the surgeon can predict the course of the surgery. He 'travels' through the virtual space, 'plays' without additional body, head and/or hand gesture changes. Simply, solely with his gaze, the surgeon coordinates his right hand movements, thus enabling, with contact free commands, the course of their preference or needs through the world of 3D-model, virtual endoscopy or virtual surgery, without stopping the endoscopic procedure in process (the endoscope is held by the left hand in the real surgery field). With this, all assumptions that the surgeon/tele surgeon is positioned in this virtual world which they entirely control but which realistically does not exist are achieved.

Keywords: Virtual reality; Leap motion; Marker-based virtual reality simulation; Gesture control; Virtual endoscopy; Virtual surgery; Tele surgery; Telemedicine

Introduction

As one of the 3D-CAS [1] (computer assisted surgery) originators and pioneers in Tele-3D- NESS [2, 3] (navigation endoscopic sinus surgery), having implemented these novel technologies for more than two decades now, occasionally we found it impossible to get an ideal perception of the surrounding world of head anatomy during real operations (nose, sinuses, base of the skull), in spite of using the latest technologies (modern CAS/NESS [4] realistic simulations of 3D volume rendering of the real intra operative anatomy, VE/virtual endoscopy, VS/virtual surgery, etc). As we know, this must be feasible completely and throughout the operation time.

Therefore, we understood the need to define a new surgical approach which would entirely satisfy the demands of the surgeon to have the impression of the presence in the virtual world of the human head anatomy and to navigate and manipulate (VS) with the virtual anatomic non-existing environment [5]. Consequently, we improved the application of the touch screen, voice control and intelligence feedback in the OR (operating room; since 2013), with the application of the hardware and original SW (software) - solutions that enable the surgeon to touch free ('in the air') 3D simulation and motion-control of 2D and 3D-medical images, as well as VE and VS, in real time [Figure1]. For this, until now new and in many ways unique but very simple application, fast and realistic visualization of medical objects in the OR with virtual and realistic planning of the course of the future surgical procedure [6, 7] we did not feel the need to form a separate IT (information technology) control center that would control the aforementioned processes in the OR. We did not need additional IT-personnel, equipment or cables in the OR, or additionally designed plug-and-play consoles that are currently in the phase of testing in the world for similar needs [8]. We concluded, according to long lasting experience and multiple testing in the OR, it to be unnecessary and potentially dangerous, primarily because during continuous usage there is the need for constant changes in the position of the body, head and hands of the surgeon in relation to the endoscopic display of the operative/surgical field in the patient's head operative/surgical monitor-LM controller, which severely impacts the surgeon's perception, focus, orientation and movements in the real world in correlation to the virtual environment, and vice versa. Knowing this, let us to question the surgeon’s complete understanding.
of anatomical 3D or tele-3D-world, orientation in the surgical field, and above all, their continuous timely understanding and fusion of image content with 2D-world of medical MSCT (multi slice computer tomography) /MRI (magnetic resonance imaging) display.

After extensive testing of this new approach in the surgery of the head ‘with touch free surgeon’s commands’, we believe that any application of new techniques and/or technology in medicine, especially in the OR, has to be simple and safe, which we repeatedly stressed during development and application of CAS and tele-CAS surgery [1, 3, 7, 9] as well as minimally invasive surgery, based on marker-based virtual reality simulation (VS) with 3D- NESS and remote visualization in the OR (if necessary).

In this way, we tried to understand the new, visualized ‘world of anatomic relations within the patient’s head’ by creating an impression of virtual perception of the given position of all elements in a particular anatomic region of the head, which exists in the real world [5] and having an impression of the presence in VW (virtual world) [9], navigating and/or tele-navigating through it and manipulating with virtual objects [5, 10].

Materials and methods

Osirix MD

In a vast selection of available tools that Osirix [11] provides to the end-user, there is a subset that significantly simplifies working with ROI (region of interest) segments [12]. Visualization of the ROIs on 3D rendered DICOM (Digital Imaging and Communications in Medicine) slices is one of the essential functionalities in the OR, offering the surgeon better understanding of the patient anatomy and pathology observed. The best approach to get appropriate insight into the treated area is highlighting and naming the relevant segments in pathology observed. The best approach to get appropriate insight into the treated area is highlighting and naming the relevant segments in pathology observed.

In the aforementioned example in the surgery of inverted papilloma the radiologist should do segmentation and pre-opearations in collaboration with technicians specialized in creating effective 3D scenes to increase the efficiency of visualization. Freehand tools like pencil, brush, polygon and other similar shapes enable precise selection of segments, whereas more sophisticated 2D/3D grow region tool can automatically find edges of the selected tissue by analyzing surrounding pixel density and recognizing similarities [12].

LM-VE-3DVR-VC description

The use of Leap Motion sensor as an interface for camera positioning in 3DVE views, with additional integration of speech recognition as a VC solution, is now possible with our recently developed special plug-in application for Osiri X platform. In this way, using only one hand, different types of gestures for 3DVR enable navigation through virtual 3D space, adjusting viewing angle and camera position [12]. Finally, we minimized surgeon distraction while interacting, with natural movements, with the positioning system in the real-world objects [12].

There is a distinction between the interpretation of gestures in 3DVR and VE because of a different manner of movements while navigating throughout the virtual space. In 3DVR view, the camera is pivoting around a focus fixed on the observed object, which gives the impression of panoramic viewing (which was of utmost importance in the aforementioned example in the surgery of inverted papilloma of the maxillary sinus, which spread to the other Para nasal sinuses ending up in the base of the skull). Because the focus is on the fixed position, moving in the direction of focus is changing distance between the camera and the focus, which is presented as a zooming effect. Furthermore, in case of VE, the focus is revolving around the camera, whereas moving in the direction of the focus results in change in the position of the camera and the focus, translating the
position in the direction of the movement vector. This produces the effect of free first-person mobility through VE space mimicking movements and view from the endoscope.

Definition of the system setup, as well as all other important factors we took into account on deciding on the computer and LM controller position in the OR, we have described previously [12], such as: (a) position of LM controller in correlation to the surgeon, anthropometry and ergonomics; (b) position of the screen monitor (distance, height, angle), quality of display; (c) position of coworking personnel (nurses, assistants); (d) quality of the screen monitor and calibration of DICOM image display; and (e) computer, OS, cables.

We propose a set of gestures that provide the surgeon with the power of manipulating crop 3D cube selection [Figure 2]. enabling him to easily create cross section overview [12].

**Case Report**

Clinical details: a 59-year-old male patient had left sinus syndrome with lateralized uncinate process, opacified left maxillary sinus, and both anterior and posterior ipsilateral ethmoid sinuses visible on MSCT scan. However, the major reason for his visiting our Clinic was nasal blockage, left-sided cheek discomfort, and occasional tooth discomfort. More recently, he had a viral upper respiratory tract infection, which entailed a condition resembling a significant left-sided sinusitis. The patient also reported epistaxis. He described his health state as normal until 2 days before, when he had experienced nose bleed from the left nostril. It stopped after a few minutes of pressure upon it. He had one more bleed from the left nostril in the past 24 hours, which also stopped after a few minutes. Multiple MSCT acquisition was performed through para nasal sinuses with 2D reconstruction [Figure 3].

Anterior group (maxillary, anterior ethmoid and frontal air cells): the left maxillary sinus and anterior ethmoid air cells were completely opacified with a mass of soft tissue. There was a bony defect of the posterior medial wall of the sinus, suggesting penetration of the pathotissue into the nasal cavity. There was mild mucosal thickening in the right maxillary sinus with liquid content in the alveolar recess. The right frontal and right anterior ethmoid air cells were patent. Mucosal thickening caused narrowing of the maxillary ostia but they remained patent. The right ostiomeatal complex and fronto ethmoidal recess

![Figure 2](image1.png)

**Figure 2**: The set of our gestures that provide the surgeon with the power of manipulating crop 3D cube selection.

![Figure 3](image2.png)

**Figure 3**: Case report: on the left side, the maxillary sinus and the anterior/posterior ethmoid air cells were completely opacified with a mass of soft tissue, as well as the entire nasopharynx and nasal cavity.

remained patent. Posterior group/posterior ethmoid and sphenoid air cells: the right air cells were patent with non significant mucosal thickening evident. The sphenoid ethmoidal recess was patent. Posterior ethmoid air cells were completely opacified with a mass of soft tissue. There was no demineralization or erosion of the sinus walls.

Nasal cavity: the mass was evident within the left nasal cavity. No abnormality of the maxillary dental row was evident. The mastoid air cells were patent bilaterally. There was no aggressive osseous lesion and no abnormality of the orbits. Within limitations of the non-contrast study, the whole left nasopharynx was obstructed with a mass of soft tissue. Histo pathologic diagnosis: inverted papilloma of the nose, nasopharynx and left maxillary sinus.

Comment: stage II of disease (limited to the ethmoid sinuses and medial and superior portions of the maxillary sinuses [11]). Altogether, we considered the situation a good indication for functional endoscopic approach in the management of the tumor, with the extensive open procedure, as recommended previously (osteoplastic approach; association of inverted papillomas with malignancy).

Operation protocol: endoscopic revision of anterior and posterior ethmoids, maxillary sinus and nasopharynx. The navigational set-up put on and calibrated and sub millimeter accuracy achieved. Nasal endoscopy on the left: the middle turbinate was fixed to the lateral nasal wall by scars. The latter area was resected first together with a mass of the soft patho tissue, and the turbinate was gently medialized. There was edematous mucosa in the entire region. Through the basal lamella, the posterior ethmoid was inspected. Here the same mass of the patho-tissue was resected, and the rest of the mucosa appeared normal. Good communication passage was created. Then, following skull base, preparation was taken, anteriorly and after resection of terminal recess of the infundibulum and intra frontal cell, good passage was created to the frontal sinus proper. Finally, the maxillary sinus ostium was identified and enlarged at the expense of the posterior fontanelle, with additional evidence of a bony defect of the posterior medial wall of the sinus (as described previously), with a mass of the patho-tissue in the whole nasal cavity, naso pharynx and maxillary sinus. After functional endoscopic and extensive open approach, the cavity was filled with hemostyptic CMC-foam.

Discussion

The application of DICOM-image viewer in the OR [11], without any physical contact with diagnostic imaging testing, medical or IT apparatus, imposes a brand new standard in the development of this century’s surgery enabling the surgeon, in real time, to possess and control all medical patient data, as well as all digital imaging diagnostic testing, which cannot be said for the usual and until now most common ‘analysis’ of medical imaging contents [Figure 4].

With continuous usage, application and perfect control of 2D-MSCT and/or MRI medical images, as well as the navigation of 3D-models, VE or VS, simple no-touch ‘in the air’ motions/commands of the surgeon’s hand in the OR, we demonstrated the application of special plug-and-play devices for conducting this sophisticated surgery to be unnecessary [Figure 5]. Furthermore, we are of the opinion that, contrary to several ongoing projects in the world, which are still in their testing/experimental phase, the plug-and-play devices impose a burden for the surgeon and negatively contribute to their focusing on the actual surgical procedures. Seeing how they are designed as additional HD, which is placed on the side of the operating table [13], it disrupts the focus of the surgeon who should be fully focused on the surgical procedure.

Therefore, on planning and defining the setups of all the important elements in the development of our form of this sort of surgery, in which the surgeon becomes ‘the master and coordinator of time and space’ in the OR, we evaded the basic problems which are self-imposed in developing this surgery strategy. We propose a simpler and more intuitive way of activating positioning control which is based on waiting for the users’ hand to enter the central trigger area to activate the interaction with the interface, with touch free surgeon’s commands [12].

The main principle of our surgical philosophy is very simple. While navigating through the narrow pathways in3D/VE preope-ratively, we noticed that the camera could stray into the tissue. Following this, getting back on the track seemed hard, sometimes even impossible without going back to the starting position. We propose a solution with predefined paths that would allow the user (surgeon) to navigate through VE, which would reduce the interface control difficulty. These paths should allow branching and labeling with specific colors and names. The surgeon should have the option to look around and change camera orientation, while camera position cannot leave the predefined path. With the aim of reducing positioning time within VE, we also suggest using predefined points on paths (checkpoints) that can be used to relocate the camera. Relocation functions should be callable as VCs (voice command) to maintain the contactless interface.

After entering the trigger area, the surgeon gains control over positioning and maintains this control while remaining within the whole active sensor area. After losing control, it is necessary to re-enter the trigger area (in our case, it was the skull base), and we did so very often during the operation of our patient. Using this approach, the surgeon can use swipe-like gestures for positioning, which are more natural to humans than sign gestures. Releasing the controls in the current position should still be possible by using the sign gesture (closed fist). After closing the fist, it should not be necessary to keep the hand closed while exiting the sensor area.

In many cases, when entering the sensor active area, the surgeon would keep an open hand (active state), which mostly resulted in sudden change in the position and losing orientation in space. The reason for this is probably because the surgeon is concentrated on the surgery being performed, and after he makes a decision to use VE, he does not think about the hand position which should be used when entering the active area (closed fist, indicating for inactive state). Instead, he automatically tries to get fast response from the system by entering the active area. Another similar problem that happened frequently was entering the active area too close to the sensor, which led to reduced freedom of hand movement.

We also observed that the surgeon tried to use palm rotation to rotate around the focus, when practicing touch free commands during the work in the OR. With the addition of such control, the surgeon would have complete control over camera position and orientation. For easier navigation through 3DVR view, we suggest adding resistance to camera position changes. Introduction of this concept would mean better visualization of rendered images and user perception of space.

Since the cropping of the rendered object is a good approach of visualizing patient anatomy and inspecting the pathology, there is a need for integration of crop controls into contactless interface. We propose a set of our gestures that provide the surgeon with the power of manipulating crop 3D cube selection, enabling him to easily create cross section overview. Positioning of the cube can be achieved in the same manner the position of the 3D object is manipulated in
Figure 4: 'Standard' basic CT diagnostics has become an important aid in the diagnosis of chronic sinusitis, in terms of the follow-up and prognosis of the course and treatment outcome [3].

Figure 5: The use of Leap Motion sensor as an interface for camera positioning in 3D/VE views, with additional integration of speech recognition as a VC solution, is now possible with our recently developed special plug-in application for OsiriX platform.

3DVR combined with additional six gestures, each representing a single cube side. Starting position assumes the user has one or both hands over sensor, depending on whether the user wants to manipulate one or two sides at the same time. Each hand can use one of three gestures (one, two or three finger gesture) that correlate to one side of the cube. This tool is used for cropping 3D models to what remains within the tools 3D cube, which is manipulated and resized before. Everything positioned between the cube sides is left visible after cropping and the remaining regions outside become hidden.

With the application of this approach to the surgery, based on non-touch Osirix-LM system which enables control of the 'live' video imagery of the surgery field, 2D-MCT-MRI Diagnostic images, VE and previously determined computer VE-tracking, VS, and 3D differentially colored models of patient head, we simulated the real operation that cannot be explored by the existing methods, thus providing significantly better survey of the procedure [5, 7, 9]. We are able to create various complex anatomic structures of the nose, paranasal sinuses and skull base, which provide a highly precise diagnosis [9, 14], as well as the most sophisticated guidance through the operation itself [6] (preoperative planning, even before the operation in the OR [Figure 6]. This 'artificially created' world represents a true copy of real anatomy of the operative field, which defines with great accuracy comprehension of the uniqueness of the pathologic processes and their extension, and lastly how it could be successfully surgically solved. To be the 'active part' of this unreal anatomic world gives the opportunity to assess all options of performing future surgery, and represents a very powerful additional element in the world of surgery and/or tele surgery of the future.

Telesurgery, as remote visualization of the operative field, in any OR has to allow surgeons not only to transfer 3D computer models and surgical instrument movements with 2D images and 3D-model manipulations, but also to define the pathology, to produce an optimal path to the pathology, and to decide how to perform the real surgery [3, 7, 9]. Using tele-fly-through or tele-VE through 3D-models, both surgeons can preview all the characteristics of the region, and so predict and determine the next steps of the operation, as we have already shown so many times since 1998 [Figure 7].
Figure 6: Various preoperative 2D/3D diagnostics through the narrow pathways inside the patient’s head, when we got an ideal perception of the surrounding world of head anatomy prior to the operation (nose, sinuses, skull base); we considered the situation a good indication for functional endoscopic approach in the management of this tumor, with extensive open procedure (osteplastic approach/inverted papilloma).

Figure 7: Using tele-fly-through or tele-VE through 3D-models, both surgeons can preview all the characteristics of the region, and so predict and determine the next steps of the operation, as we have done so many times since October 1998) (taken with permission of Klapan Medical Group University Polyclinic, Zagreb, Croatia, EU).

Conclusion

- Using DICOM viewer, VE/VS per via m LM with touch free commands, animated images of the real surgery can be designed intra-operatively, very truly and precisely.

- Even more, this approach offers the possibility of preoperative planning (or the use of virtual simulator [15]) and has become a very important segment in surgical training and planning of each individual N-OsiriX-LM, robotic [16], or telesurgical intervention [2].

- The N-OsiriX-LM suggests that real and virtual objects definitely need to be integrated by use of real ‘in the air’ control with simulation of virtual activities that requires real-time visualization of 3D-VE motions [12, 7], following the action of the surgeon that may be moving in the VR area [12].

- It is achievable with simple hand gestures of the surgeon which do not differ in any way from all other natural everyday hand or body movements. In this way, the surgeon can predict the course of the surgery. He ‘travels’ in the virtual space, ‘plays’...
without additional body, head and/or hand gesture changes. Simply, solely with his gaze, the surgeon coordinates his right hand movements, which enables, with contact free commands, the course of their preference or needs through the world of 3D model, VE or VS, without stopping the endoscopic procedure in process (the endoscopy is held in the left hand in the real surgery field). With this, all assumptions that the surgeon/tele surgeon is positioned in this virtual world which they entirely control, but which realistically does not exist, are achieved.

- After just a second, the surgeon recognizes this unreal world and his brain activates the centers that control his further physical activities.

**Abbreviations**


**References**


15. Caversaccio M, Eichenberger A and Hauserl R. Virtual simulator as a training tool for endonasal surgery. Am J Rhinol. 2003; 17:283-290. [Crossref]