Journal of Biomedical Informatics 72 (2017) 140–149

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

On the utility of 3D hand cursors to explore medical volume datasets with a touchless interface

Daniel Simões Lopes [a,*], Pedro Duarte de Figueiredo Parreira [b], Soraia Figueiredo Paulo [a], Vitor Nunes [c], Paulo Amaral Rego [d,e], Manuel Cassiano Neves [f], Pedro Silva Rodrigues [g], Joaquim Armando Jorge [a,b]

[a] INESC-ID Lisboa, IST Taguspark, Avenida Professor Cavaco Silva, 2744-016 Porto Salvo, Portugal
[b] Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
[c] Surgery Department, Hospital Prof. Doutor Fernando Fonseca, E.P.E., IC19, 2720-276 Amadora, Portugal
[d] Hip Surgery Unit, Orthopedic Surgery Department, Hospital Beatriz Ângelo, Av. Carlos Teixeira, 3, 2674-514 Loures, Portugal
[e] Department of Orthopaedic Surgery, Hospital da Luz, Avenida Lusíada, 100, 1500-650 Lisboa, Portugal
[f] Department of Pediatric & Adolescent Orthopaedic Surgery, Hospital CUF Descobertas, Rua Mário Botas, Parque das Nações, 1998-018 Lisboa, Portugal
[g] Oral Implantology Group, Clínica Universitária Egas Moniz, Rua D. João IV Nº 23ª, 2800-114 Almada, Portugal

ARTICLE INFO

Article history:
Received 17 December 2016
Revised 30 June 2017
Accepted 10 July 2017
Available online 15 July 2017

Keywords:
3D hand cursor
Bi-manual gestures
Medical volume data
Touchless interface
User study

ABSTRACT

Analyzing medical volume datasets requires interactive visualization so that users can extract anatomo-physiological information in real-time. Conventional volume rendering systems rely on 2D input devices, such as mice and keyboards, which are known to hamper 3D analysis as users often struggle to obtain the desired orientation that is only achieved after several attempts. In this paper, we address which 3D analysis tools are better performed with 3D hand cursors operating on a touchless interface comparatively to a 2D input devices running on a conventional WIMP interface. The main goals of this paper are to explore the capabilities of (simple) hand gestures to facilitate sterile manipulation of 3D medical data on a touchless interface, without resorting on wearables, and to evaluate the surgical feasibility of the proposed interface next to senior surgeons (N = 5) and interns (N = 2). To this end, we developed a touchless interface controlled via hand gestures and body postures to rapidly rotate and position medical volume images in three-dimensions, where each hand acts as an interactive 3D cursor. User studies were conducted with laypeople, while informal evaluation sessions were carried with senior surgeons, radiologists and professional biomedical engineers. Results demonstrate its usability as the proposed touchless interface improves spatial awareness and a more fluent interaction with the 3D volume than with traditional 2D input devices, as it requires lesser number of attempts to achieve the desired orientation by avoiding the composition of several cumulative rotations, which is typically necessary in WIMP interfaces. However, tasks requiring precision such as clipping plane visualization and tagging are best performed with mouse-based systems due to noise, incorrect gestures detection and problems in skeleton tracking that need to be addressed before tests in real medical environments might be performed.

1. Introduction

Exploring 3D medical images (i.e., volume) is a daily activity performed in a variety of healthcare scenarios such as diagnostics, surgical planning, anatomical 3D modeling, medical education and even patient communication. A particularly challenging scenario for 3D medical image exploration consists of surgical navigation, where there is a demand to maintain a sterile environment. Ultimately, the need to manipulate images forces the surgeon to move away from the surgical table and towards the image terminal, where they use a sterile towel over the mouse to interact with the graphical interface. Alternatively, surgeons often request other

Abbreviations: 2D, two-dimensional; 3D, three-dimensional; CT, computed tomography; DICOM, digital imaging and communications in medicine; GUI, graphical user interface; LAN, local wireless network; M, mean; MRI, magnetic resonance imaging; SD, standard deviation; USB, universal serial bus; VE, Voxel Explorer; Vol, Volview; WIMP, windows-icons-menus-pointers.

* Corresponding author at: INESC-ID Lisboa, IST Taguspark, Room 2N9.1, Avenida Professor Cavaco Silva, 2744-016 Porto Salvo, Portugal.

E-mail addresses: daniel.lopes@inesc-id.pt (D.S. Lopes), pedro.f.parreira@tecnico.ulisboa.pt (P.D.de F. Parreira), soraiafpaulo@inesc-id.pt (S.F. Paulo), vitor.m.nunes@hff.min-saude.pt (V. Nunes), prego@hospitaldaluz.pt (P.A. Rego), manuel.cassianoneves@jmellosaude.pt (M.C. Neves), pmsilvarodrigues@egasmoniz.edu.pt (P.S. Rodrigues), bjorgej@tecnico.ulisboa.pt (J.A. Jorge).

http://dx.doi.org/10.1016/j.jbi.2017.07.009
1532-0464/

personnel to manipulate images in their place. These ill-posed behaviors lead to delays, misinterpretations or miscommunications, eventually interfering with the surgeons' ability to interpret the medical images and, consequently, complications upon the operated patient [25].

With an increased reliance on imaging, there will be a greater need for aseptic navigation and due to the current improvements in real-time volume rendering, it becomes increasingly important to design more natural and aseptic prone interfaces, so that users drive a better understanding and visual awareness of the anatomo-physiological information contained within the volume data [22,26,23,7,15,31,16].

The vast majority of volumes rendering systems offer several functionalities operate upon 3D entities: volume data, clipping planes, cursors or tags. Probably the most commonly used functionalities are volume reorientation, transfer function definition and clipping plane visualization. Other software features such as measuring lengths, angles and adding tags are also available although less employed.

However, conventional volume rendering systems rely on traditional 2D input devices, such as mice and keyboards. It has been reported that exploring 3D content using such devices is often difficult, time consuming, promote potential non-sterility issues in surgical navigation, and are cumbersome as the user is forced to manipulate three dimensional objects using a bi-dimensional cursor. In other words, there is a gap between the degrees of freedom (dof) of the content (a tag has 3 dof while a volume has 6 dof) and the cursor (2 dof). Most noticeably, the gap stands out whenever a user intends to obtain specific orientations of the volume or a clipping plane using a mouse, usually requiring a long thread of sequenced rotations to obtain the desired point of view [30].

In order to address this gap, this paper addresses the problem of whether touchless interfaces operated with 3D hand cursors can improve 3D medical images exploration, when compared to more traditional WIMP interaction. The main goals of this paper are to explore the capabilities of (simple) hand gestures to facilitate sterile manipulation of 3D medical data on a touchless interface, without resorting on wearables, and to evaluate the surgical feasibility of the proposed interface next to senior surgeons (N = 5) and interns (N = 2). To this end, we developed a touchless interface that tracks body postures and hand gestures, and conducted usability tests to evaluate user response and performance with laypeople, senior surgeons and professional biomedical engineers.

2. Related work

Several papers have specifically explored new forms of interaction that aim to improve interactive visualization of medical volume. They usually combine Augmented/Virtual Reality technologies, stereoscopic displays with input controllers and interactive surfaces [3,17,29,4,20,14].

A keystone paper elaborating on bimanual interaction for 3D medical visualization relied on passive interface props [13]. Through physical manipulation of these props, the user was able to specify spatial relationships between the tangible object and the digital content. The props-based interface was applied in neurosurgical planning for 3D brain manipulation, clipping plane selection and trajectory selection. Results from informal evaluation sessions have shown that the interface facilitates a natural two-handed interaction, while providing haptic and tactile feedback for the user. However, this interface presented several limitations, namely, the user cannot easily express constrained motions, and it lacked a proper clutching mechanism to stop tracking the props as users relied on a foot pedal and buttons mounted on the props to clutch the props.

A distinguishable paper was presented by [19], which have attempted to segment a data volume without the loss of context or distortion of the volume. They considered a bimanual interaction approach based on the metaphor of cracking a volume open. Using dedicated hardware, users can grab the volume and separate distinguished parts of the volume, without any loss of data. The results were encouraging as users reported that the system was easy to learn and use. This type of interaction is relatively more difficult to translate to WIMP-based approaches, and it heavily depends on expensive dedicated hardware.

Less cumbersome devices has been developed for physicians to explore medical images using simple hand gestures. Small wearable devices such as [1,24] consist of small wrist bands equipped with EMG and accelerometers which interpret muscle activation, that can be mapped to specific interaction events. However, these devices allow for very few controllable gestures. Therefore, more demanding operations will mandatorily require a larger number of gestures to obtain the desired result, turning image exploration a more timely and cumbersome task. In addition, these wearable devices usually do not track hand position are require recurrent calibration, which seriously hampers a proper interaction with the 3D content.

A different approach, which discards any use of wearable or haptic devices, are the touchless interfaces [15,16,21,18]. For example, [27] relies on simple hand gestures captured by a depth camera to interact with volumes generated from medical images. In this case, a Leap Motion camera is placed underneath the hand to capture gestures, hand positions and orientations. While simpler and less obstructive, this setup offers a very small acquisition space and must be placed in the vicinity of the user's hands, thus, limiting the reachable space of the surgeon's arm and hand movements.

Smaller cameras have been suggested to make the touchless approach more suitable, such as the Soli project from Google [28] which uses a sonar to detect and identify very fine hand and finger movements, but still requires the use of dedicated hardware and must be placed at a short distance due to its limited acquisition space which also leads to less natural and more restricted gesture interactions.

On the contrary, the most familiar depth camera Kinect offers a much wider acquisition space. The Kinect has been used in several healthcare applications [9]. The company Gestsure [11] has taken advantage of the Kinect to create an easy to use system that allows surgeons to interact with their image viewing software using simple hand gestures to explore image slices or manipulate volumes during surgery. O'Hara et al. [25] also took into account not only the needs of the physicians during surgery, but also the socio-technical concerns at play when designing this type of technology, giving further insight into the trials and tribulations inside the surgical block. Unfortunately, these projects merely use the Kinect to convert 3D hand positions and gestures into interactive 2D cursors. Such approaches emulate standard mouse and keyboard controls and do not taking full advantage of depth cameras potential to capture three-dimensional hand information.

In this paper, we present a touchless interface for manipulating medical volume data using both hands as interactive 3D cursors. A comparative study between WIMP and touchless approaches was conducted with laypeople and the proposed system was evaluated by professionals that deal with medical images in a daily basis.

3. Touchless interaction system

A touchless interaction system, called Voxel Explorer, was developed to interact with 3D medical images without any physical contact or wearable device. Interaction with the application is done exclusively with a depth camera (Kinect One®) which

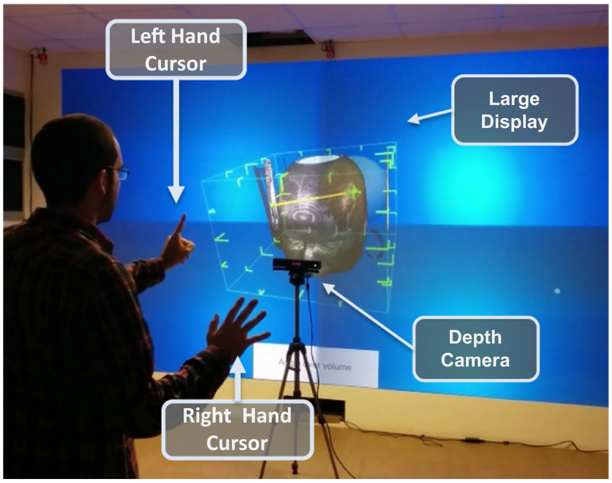

Fig. 1. Hardware setup of the touchless interaction system.

captures kinematic body data at approximately 30 frames per second. The camera is positioned in front of the user, at approximately 1.5 m, while the display sits behind the camera (Fig. 1). A dedicated PC with a USB 3.0 port running Windows 8 was used to connect the depth camera. A second PC was required to run the Voxel Explorer application and display the generated graphical content. The application was developed using Unity version 5.1.2f1 and Kinect SDK version 2. Data acquired from the depth camera, namely joint positions and hand gestures, is sent via a LAN to the computer running the application.

The hardware setup allows the user to move within the depth camera's acquisition frustum, hence, does not require the user to stay at a fixed location. In addition, different users may swap places, without loss of functionality, since the controls will always be adapted to the position and body size of the new user.

3.1. Volume data

Stacks of 2D medical images, or slices, are converted from DICOM (∗.dcm) to bitmap (∗.bmp), and the intensity value of each pixel is allocated into a 3D Texture to hold the color information.

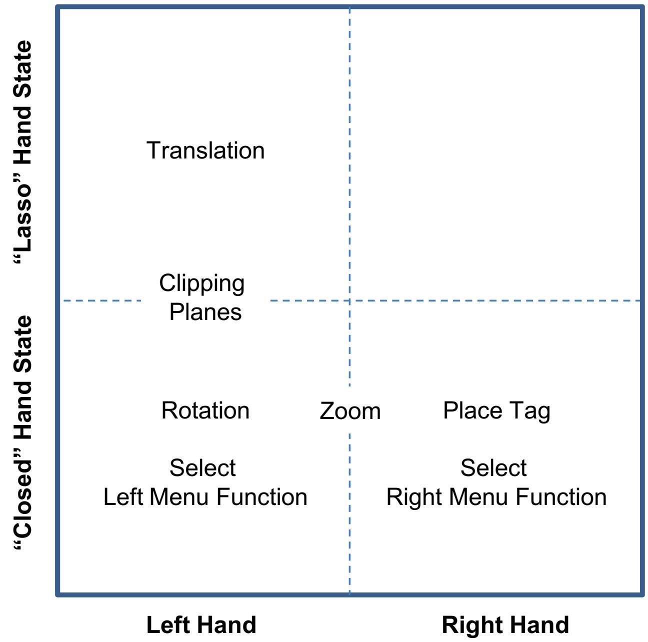

Fig. 3. User actions mapped according to hand maneuverability and hand gesture of the proposed touchless interaction system.

The texture dimensions correspond to the images original dimensions, $512 \times 512 \times N$, where N is the number of available slices. The 3D Texture is then applied as a material property of a parallelepiped geometric primitive that is properly scaled according to pixel dimensions, number of slices and slice spacing. The resulting volume is rendered with conventional ray marching algorithms, and in-house GLSL shaders were coded to account for depth and opacity properties.

3.2. Graphical user interface

The GUI is composed by a volume in the center of the screen (Fig. 2(a)), a display panel at the bottom of the screen to indicate which functionality is active (Fig. 2(b)), a directional box or cube with orientation lettering and colored axis at the lower left (Fig. 2(c)), and two lateral off-screen vertical panels with buttons

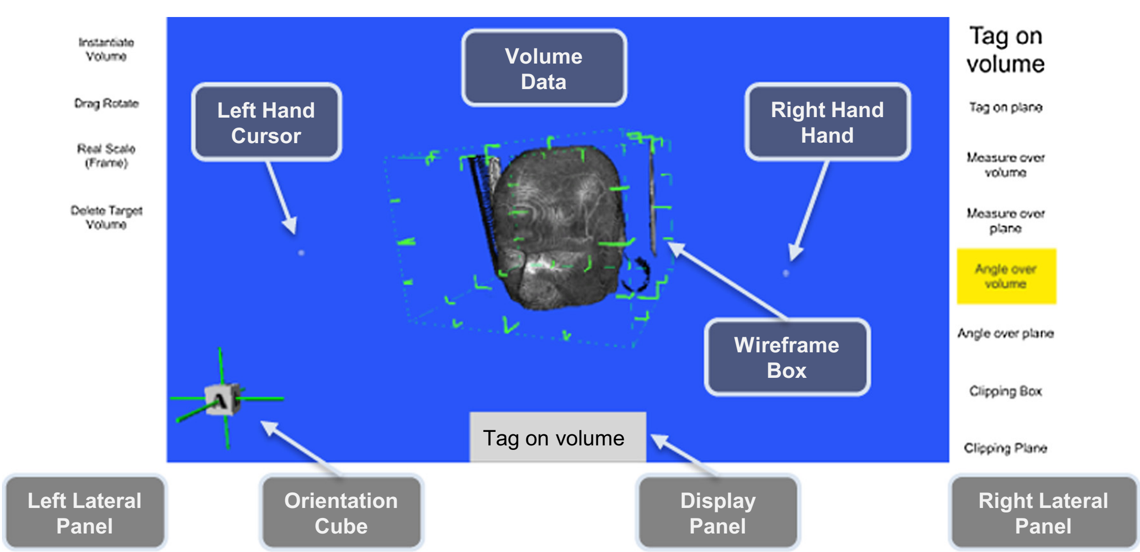

Fig. 2. GUI of the Voxel Explorer application (right lateral panel with highlighted function, in yellow). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

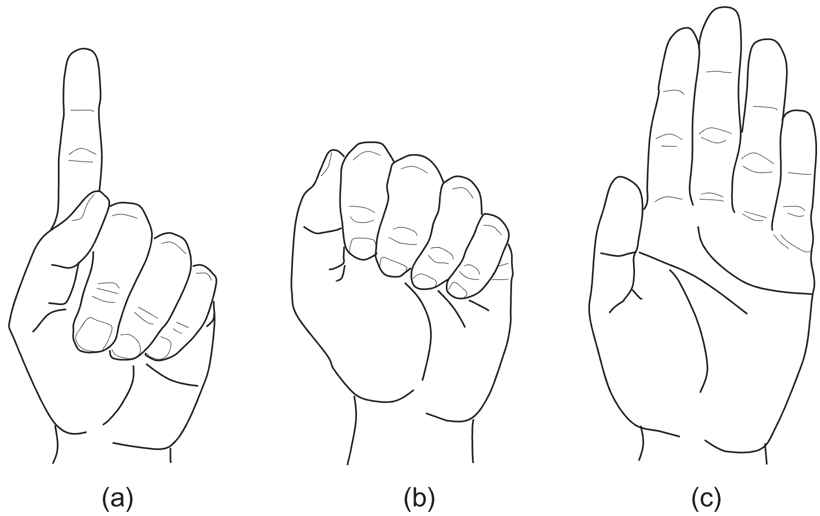

Fig. 4. Considered hand gestures for volume exploration: (a) lasso; (b) closed; and (c) open.

that appear once reached by a hand cursor (Fig. 2(d) and (e)). Note that the GUI is suited for hand and arm gesture interaction, as it consists of an minimalistic display of on screen buttons which only appear whenever the user intends to evoke a given functionality. In order to provide a relative sense of scale and perception of volume dimension, a wireframe box with equally spaced ticks was rendered on top of the volume. Two spherical cursors are also displayed to indicate both hand positions.

3.3. Interaction

Interaction is based on an asymmetrical bimanual paradigm [12] were the left hand is used primarily for positioning and reorienting 3D content (i.e., volume, clipping plane, tags), while the right hand is used mainly for selecting functions and activating/deactivating commands (Fig. 3). Note that, the interaction design follows the metaphor of a traditional game pad, where the digits of the left hand deal with directional input while the right hand digits trigger action buttons. Besides the 3D hand positions, a set of three simple hand gestures also play an essential role in interaction. To this end, the depth camera not only acquires the positions of both hands but also the following left or right hand gestures (Fig. 4: lasso (closed hand except for the index finger which points outwards; closed (hand forms a closed fist; and open (all fingers pointing outwards forming an open palm.

Interaction with the available 3D content and lateral panels is facilitated by segmenting the reachable space of the screen, as specialized areas can trigger specific events. Such segmentation avoids cursor ambiguity between the 3D content and lateral panels. To this end, an invisible window (i.e., set of vertical and horizontal boundaries) was attached to a point close to the user's upper part of the sternum or manubrium (also known as the *spineshoulder* label of the Kinect V2 Joint ID Map). This window is centered at this point and has a height determined by 1.5x the length between the position of the head joint and the center point, and a width determined by 1.5x the length between each shoulder. The invisible window also defines upper and lower strips, which can be used for context sensitive commands, namely for undo/redo or reset operations.

3.4. Content selection

There are two main content types to interact with: (i) lateral panels or menus with function buttons; and (ii) 3D content, which includes volume of medical images, clipping planes and tags.

Function button selection is performed by accessing the lateral panels and selecting the desired option. Whenever a hand is positioned outside the left or right boundaries of the window, yet within the horizontal boundaries, then side menus appear. The height of the hand then determines the hovered button, which appears highlighted in yellow. Selection is performed by closing the hand during 2 s, after which the button will be colored red. A list of the functionalities displayed at each lateral panel is presented in Table 1. The left panel contains several functions related to manipulating the display of the volume and the right panel contains several functions that are applied over the volume, such as measurements and clipping.

When the users' hands are positioned inside this window, the corresponding hand cursors appear on screen, and it is possible to select and manipulate the 3D content. If the volume is selected, then the wireframe box is rendered.

Table 1
Lateral panels and corresponding functions.

Left panel	Right panel
Instantiate volume Instantiates a new volume in the workspace	**Tag volume** Places a cursor over the cube surrounding the volume
Real scale (Frame) Instantiates a frame surrounding the volume, which allows seeing it in real scale	**Tag plane** Places a cursor over the clipping plane applied to the volume
Delete target volume Deletes the selected volume from the scene	**Measure over volume** Allows to measure a length over the cube surrounding the volume
	Measure over plane Allows to measure a length over the clipping plane applied to the volume
	Angle over volume Allows to measure an angle over the cube surrounding the volume
	Angle over plane Allows to measure an angle over the clipping plane applied to the volume
	Clipping box Allows to clip the volume along the canonical axis
	Clipping plane Allows to clip the volume along an arbitrary plane

3.5. Volume manipulation

By default, translation and rotation of the volume is available whenever no function is selected to avoid conflict with other functionalities. Once the left hand cursor hovers the 3D content, users can either translate ("lasso") or rotate ("closed") the volume or other objects. Deselecting is performed with the "open" gesture.

The amount of volume displacement or angular rotation is determined by two consecutive 3D positions of the left hand cursor. At every frame, a vector is computed from the previous position to the current hand position. This vector is then added to the position of the volume or its rotation is changed so that it faces the new point, according to which functionality is currently active. To reset the left hand position, the user performs an "open" gesture.

For more constrained manipulations, a directional box is placed at the lower left corner of the GUI. Besides providing orientation information of the volume according to the visible cube faces A – anterior, R – right, L – Left, P – posterior, U – upper, L – Lower), by interacting with the axes, it also serves as an interactive widget for axis constrained rotation and translation along the selected axis, allowing a finer control over the volume [6]. The active axis is displayed in red, while the remaining axes are displayed in green. For example, when the user grabs the z axis ("closed" left hand) the volume will only rotate and translate along that axis.

3.6. Clipping planes

The users have the ability to clip the volume along a plane perpendicular to each axis. When the option is selected, the user may choose which plane to move by using the regular translation methods when a plane is selected. When the cursor is over a plane, that plane is highlighted in green, indicating it has been selected, and then highlighted in red when moved. An arbitrary clipping plane can also be applied by using the same translation and rotation controls applied to the volume. Only the voxels of the volume above the plane are rendered. A normal vector is placed in the center of the plane to better distinguish the area above and under the plane. The normal vector is defined by the left hand position and the center of the plane. A clipping plane is then set after performing the "open" left hand gesture. Both forms of clipping are cumulative, which means that both canonical and free clipping can be applied to the volume simultaneously.

3.7. Tagging and measurements

Users have the ability to place tags to mark points of interest inside the volume, to measure distances or angles. All tags and lines are placed upon clipping planes and are permanently rendered "over" the volume. The position of the tag is determined by the intersection of a ray, which is casted from the left cursor's x and y position along the forward direction, and a clipping plane. If no collision is detected, a cross is rendered red inside the cursor. In case of collision, a green cross is rendered at the intersection point. To set a marker, the user closes the right hand during a couple of seconds. This avoids unwanted placement of markers if the user leaves the right hand open for too long.

Two tags are necessary to measure lengths between distinct points, and are rendered as a connecting line with the length value (in centimeters) displayed midway. Angles are determined in a similar fashion, utilizing three markers, with the measured angle (in degrees) appearing over the second marker.

4. Usability studies

To validate our system, two distinct user evaluations were performed in a controlled environment. Firstly, formal usability tests were conducted with laypeople which compared our touchless interaction system to WIMP interfaces. Secondly, senior surgeons and interns were invited to assess the medical benefits and limitations of our system and its adequacy to support surgical navigation.

4.1. User tests with laypeople

Fifteen unpaid participants (13 male and 2 female), with ages ranging from 20 to 24 years old (M = 22.6; SD = 1.3), were recruited for the comparative studies. Two participants reported to be left handed. One user reported to be experienced with motion controls, while the remainder reported only basic knowledge regarding this technology.

To evaluate users' ability to rotate volumes with each system, users were asked to achieve, both in Volview and Voxel Explorer, a similar orientation for a given volume rendering image. Users were also asked to apply clipping planes to reveal a specific structure within the volume, over which they would be asked to measure a distance, an angle and tag a specific and easily recognizable anatomical landmark. Clipping tasks were performed in three different platforms. Due to considerable differences between clipping with Volview and Voxel Explorer, a third software was used, 3DVol, due to its similarity in controls to Voxel Explorer, while still using a mouse.

A total of three volumes were considered for the 3D content manipulation tests (Fig. 5): a training volume represented by a CT scan of a human head (BRAINIX data set, Osirix, 2015) was used so that users had the opportunity to familiarize themselves with the application before the tests; and two task volumes of anonymized data represented by a CT scan of a human thorax and an MRI scan of the pelvic region, which were made available by Hospital Professor Doutor Fernando Fonseca, (E.P.E.). The user would test the systems up to 5 min to achieve habituation, although in average they took close to 3 min. To avoid bias of results caused by the order in which the tests were presented, the first system to be tested (Volview or Voxel Explorer; 3DVol

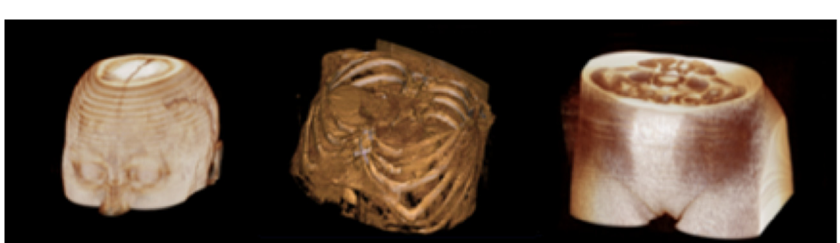

Fig. 5. Test volumes 0, 1, and 2.

or Voxel Explorer) was randomized for each user, as was the order of the test volume.

For each test (i.e., rotation, clipping plane, measurements and tagging), a total of five tasks were performed and users were asked to obtain a close result to a given volume rendering image. The time required for task completion was measured, along with the number of attempts a user needed to complete each task. To determine the number of attempts needed to obtain the desired result, a single attempt was counted at the beginning of each task, an a new attempt was added whenever: (1) the user would reset the orientation of the volume (except if this was performed at the very beginning of the task); (2) the user would move the volume randomly (e.g., mistakenly dragging the mouse or rotating their hands); (3) the user would lose control of the volumes' rotation (either by accident or frustration) and would have to try again.

4.2. User evaluation with surgeons

While the usability studies with laypeople served to evaluate the overall performance of the touchless interface and its ease of use by inexperienced people, it was still necessary to assess the interface's adequacy as a tool to be used by surgeons in their practices. Therefore, seven surgeons of different medical backgrounds were invited as specialists to test our touchless interface in the context of surgical navigation. Each professional was asked to analyze the benefits and limitations of Voxel Explorer as a image navigation tool, from a professional viewpoint (Fig. 6).

We evaluated the system with seven medical professionals, none of which were female. The ages of the invited surgeons ranged between 25 and 50 years old. Only one was left handed. Six were surgeons: one general surgery senior with 27 years of experience, two general surgery interns with 2 and 3 years of experience, one pediatric orthopedic senior with 40 years of experience, one orthopedic hip senior with 17 years of experience, and one orthopedic knee senior with 5 years of experience. Another participant was a medical dentist with 6 years of experience in oral surgery. None of the participants reported previous experience with spatial or touchless interfaces, but four of them were aware of the Wiimote gamepad and Kinect sensor. All reported that they analyze 2D medical images for pre-operative purposes on a routinely basis and, with the exception one participant, they also occasionally rely on 3D images during the pre-operative stage.

We wanted to assess whether the proposed touchless interface can be used in professional settings to overcome surgical practice issues related to image navigation and asepsis. The evaluation conducted with professional surgeons had a different configuration from the tests conducted with laypeople. These specialists were asked to test the Voxel Explorer system alone, after a thorough demonstration of all available functions and their application. To this end we employed medical images retrieved from anonymized CT scans compatible to the surgeons specialties (general surgery and hip orthopedics: middle aged female pelvis containing a small fracture and a metal hip prosthesis; elbow orthopedics: middle aged male elbow joint without any visible joint pathology; maxillofacial: middle aged male with brain tumor). Evaluation sessions comprised four stages: (i) introduction, (ii) free experimentation, (iii) questionnaire, and (iv) guided interview. We asked participants to stand through the test during approximately 30 min, with 5 to freely observe the volume data.

Fig. 6. One of the invited surgeons testing the Voxel Explorer system.

During free experimentation, users not only had the opportunity to familiarize themselves with the system but were asked to freely explore the images in order to characterize the anatomical content contained in the data set. The only restriction being they had to experiment with all available functions at least once. In average they took less than 3 min to achieve habituation. Afterwards, surgeons were asked to complete a questionnaire to raise the participants profile regarding their experience with interaction systems and image navigation, along with a questionnaire on the quality of their experience, mainly referencing their preferences between current surgical navigation practices and the presented touchless interface. The objective was to classify the level of difficulty felt in task performance and in the use of the available features. We assessed user preferences with a list of statements scored on a 6-point Likert Scale (6 indicate full agreement). Finally, we conducted a guided interview to capture their impressions about the system and its possible application in surgical practice, both professionally and in medical education. Data gathered also included transcripts of the interviews and observational notes taken during evaluation sessions. The interviews also answered to the questionnaire, inviting them to elaborate as much as possible on the subjects they found to be the most relevant from their experience. We also requested their thoughts and opinions regarding improvements to the system.

5. Results and discussion

In this section, we present the main observations made during the tests with laypeople and surgeons, as well as the difficulties felt during the evaluation and suggestions to improve the touchless interaction system. We also present the analysis made based on the results of the questionnaires and log files obtained during the tests.

5.1. Tests with laypeople

Regarding the usability tests with laypeople, when performing volume rotation users presented different task completion timings (Fig. 7). The average completion times for Test Volume 1 were shorter using Volview (M = 31.6 *s*, SD = 13.7 *s*) than when using Voxel Explorer (M = 39.8 *s*, SD = 31.5 *s*). The paired *t*-test ($p < 0.05$) reveals that the tasks performed with Voxel Explorer were substantially slower than Volview ($p = 0.04$). However, the results for Test Volume 2 were more similar between Volview (M = 28.2 *s*, SD = 11.2 *s*) and Voxel Explorer (M = 31.1 *s*, SD = 18.2 *s*), although with no statistically relevant differences ($p = 0.25$).

The completion time differences between both systems may result from a steeper learning curve regarding Voxel Explorer's interface, which is expected due to the lack of familiarity with motion based controls reported by the majority of the users. As for the performance differences between Test Volume 1 and Test Volume 2, we noticed that users seemed to gain dexterity during

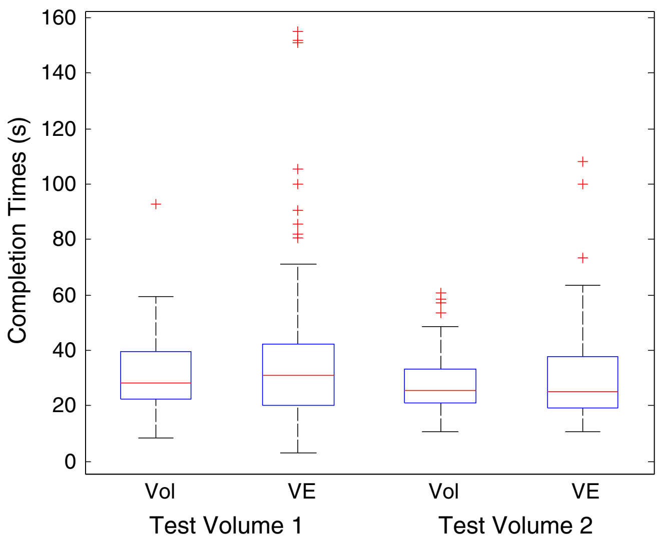

Fig. 7. Completion times of volume rotation using Volview (Vol) and Voxel Explorer (VE) regarding rotation of Test Volumes 1 and 2.

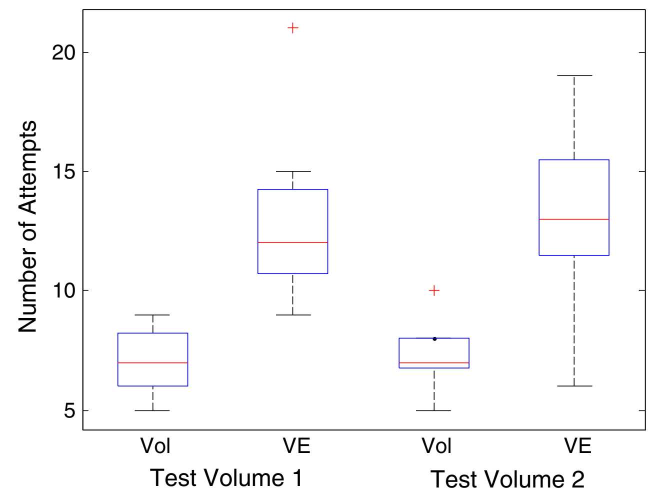

Fig. 9. Average number of attempts of volume rotation tasks using Volview (Vol) and Voxel Explorer (VE) performed upon Test Volumes 1 and 2.

the first battery of tests, which may have led to a better performance in Test Volume 2. However, there are no statistically relevant differences between completion times for the second set of tests.

As for the average completion time per attempt to obtain the desired orientation, Voxel Explorer performed better than traditional WIMP approach (Fig. 8). The first set of tests revealed that the average completion time per attempt when performing rotation with Voxel Explorer (M = 12.9 *s*, SD = 2.59 *s*) was significantly smaller ($p = 0.0017$) than Volview (M = 21.5 *s*, SD = 5.64 *s*). The second battery of tests also reveal that completion time per attempts in Voxel Explorer (M = 12.7 *s*, SD = 3.42 *s*) is significantly faster ($p = 0.0013$) than in Volview (M = 18.9 *s*, SD = 3.46 *s*).

Fig. 9 shows that the number of attempts required for volume manipulation is significantly higher (Test Volume 1: $p = 7 \times 10^{-4}$; Test Volume 2: $p = 0.0015$) when using Voxel Explorer (Test Volume 1: M = 13.0, SD = 3.54; Test Volume 2: M = 13.1, SD = 3.82) comparatively to Volview (Test Volume 1: M = 7.11, SD = 1.45; Test Volume 2: M = 7.22, SD = 1.39), which explains the higher average completion times using Voxel Explorer. This occurs mainly because of the lack of familiarity of a touchless interface, consequently resulting in a larger number of trial-and-error attempts to obtain the desired result.

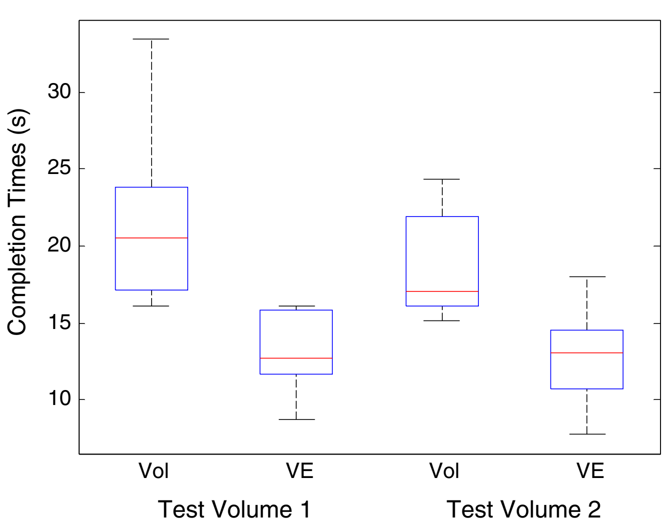

Fig. 8. Completion times per attempt of volume rotation tasks using Volview (Vol) and Voxel Explorer (VE) to perform rotation upon Test Volumes 1 and 2.

However, some users reached the desired rotation with a single attempt in Voxel Explorer. On the other hand, users who relied too heavily in rotating the volume using the fixed axis faced substantial difficulties to obtain the desired orientation. This is especially evident when observing task completion times, as the time required to obtain the desired results was substantially longer than those in the same test volume and the overall mean value. While useful for small adjustments, these constrained rotations are clearly not meant to be used as a main method for manipulating volumes, mainly due to the difficulty in selecting the desired axis with the Kinects' lack of precision.

Regarding the remaining exploration tasks, namely, application of clipping planes, distance and angle measurement, and tagging, users performed better using the traditional software on all tasks and both test volumes, while tasks performed with Voxel Explorer took considerably longer (Fig. 10). There was also no substantial improvement in performance between the first and second volumes in any of the tasks using Voxel Explorer ($p > 0.05$), which indicates that user experiences role in this case is either indifferent or insubstantial in the tests performed.

One major hurdle noted during the execution of these tasks, which presents the main reason for users' poor performance, was the Kinects lack of precision. While users were relatively quick to identify the structures of interest, they had some trouble selecting the desired slice or placing the markers in the correct position, thus increasing the time necessary to complete the task.

After performing the tasks, users were asked to answer two surveys. To obtain an overall opinion of user's preferences, the first survey consisted of a binary response scale screening questionnaire [8]. Users preferred the use Voxel Explorer for both rotation (60%) and translation (100%) of the volume over the conventional mouse controls. However, users preferred Volview for measuring (73.3%) and tagging (60%). Regarding the application of clipping planes, users preferred Voxel Explorer (66.7%) over 3DVol (33.3%). The second survey present a Likert-type scale for rating questionnaire responses related to more specific features as shown by (Tables 2) and (3). When asked about the level of difficulty felt when performing translation and rotation, users felt no significant difference between interfaces, as suggested by previous results, although they slightly preferred Voxel Explorer for rotating a volume. In addition, users reported that Voxel explorer provides a bet-

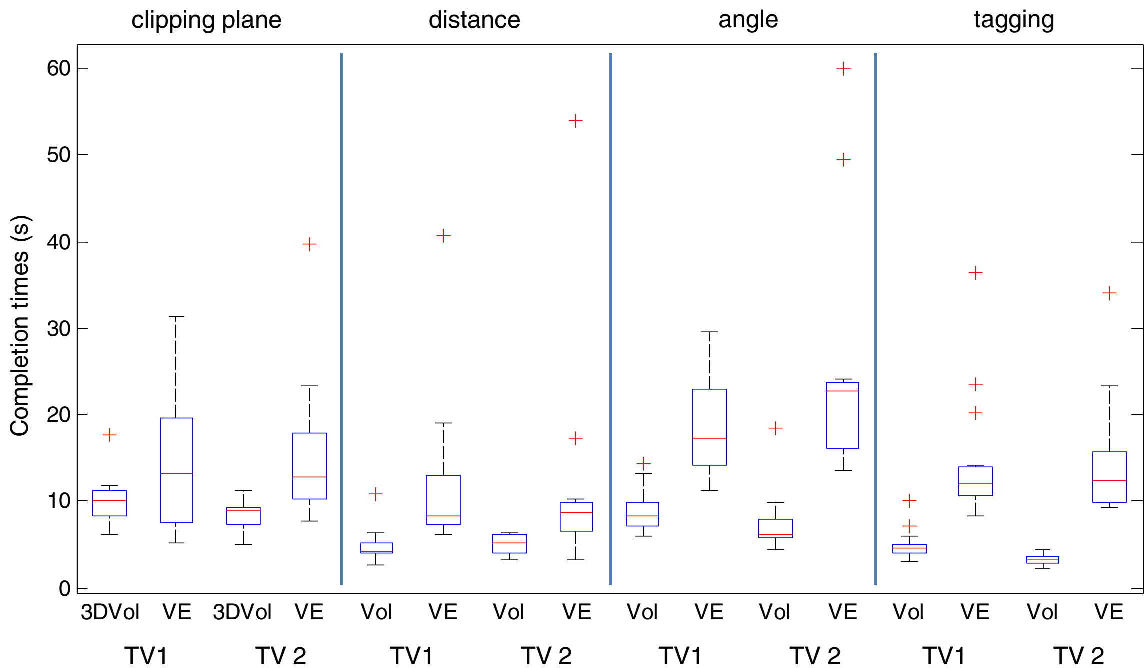

Fig. 10. Completion times of other exploration tasks applied upon Test Volume 1 (TV 1) and Test Volume 2 (TV 2) using Volview (Vol), 3DVol (3DV) and Voxel Explorer (VE) to apply clipping planes, measure distances and angles and apply tags.

Table 2
Participants preferences regarding different criteria for the evaluated systems: Median (inter-quartile range). Likert scale: 1 – totally disagree and 4 – totally agree.

	Volview	Voxel Explorer
In general, it was easy to rotate the volume in 3D	2(1)	3(1)
In general, it was easy to translate the volume in 3D	4(1)	4(0.5)
It was easy to identify how the control buttons worked[a]	2(0)	3(0)
The interface has a coherent layout[a]	3(1)	3(1)

[a] Indicates statistical significance.

Table 3
Participants opinion regarding fatigue and spatial awareness features of the touchless interface: Median (inter-quartile range). Likert scale: 1 – totally disagree and 4 – totally agree.

Operating with a touchless interface is more fatiguing than with a WIMP interface	3(0)
Applying clipping planes with a touchless interface enables a better spatial awareness of the structures inside the volume than with a WIMP interface	3(1)

ter spacial awareness of the structures inside the volume when compared to the slice views in 3DVol. Noticeably, users referred that the touchless interface imposes some physical strain or fatigue when performing mid-air tasks.

While the task results may not reveal an immediate benefit for the use of Voxel Explorer when compared to the traditional software, a deeper analysis of the data reveals that the strength of this new approach is in the ability to perform the trial and error procedure often associated with these tasks at a much faster rate. This means that even if the user fails to obtain the desired orientation of the volume in a certain attempt, resetting the volume and attempting again is an easy and quick task to accomplish, even with limited training and exposure to the controls. As the tasks proceeded, users required less time to obtain the desired results, which shows that the short time they had with the platform was enough to gain experience enough to improve the outcome.

The greatest limitations of the proposed touchless interface were the unfamiliarity with touchless interfaces and, more importantly, the imprecise body tracking of the depth camera [2,10,5]. Most body tracking and hand gesture recognition errors were introduced by the Kinect sensor, as it often fails to recognize hand position and gesture whenever (i) the user's hands are positioned too close to each other, (ii) if they are positioned near to the body and/or close to the medial axis of the body, (iii) if the hand is pointing directly to the camera (i.e., "lasso"), this gesture often is misinterpreted as being "closed". The Kinect imprecise tracking and recognition is probably the main reason why users performed worst when making measurements and using clipping planes. Users would often take longer to place makers on the indicated position. Similarly, when applying clipping planes, users would often lose a few more seconds trying to place the plane in the correct position.

5.2. Tests with surgeons

In terms of user experience, the responses to the questionnaire suggest that surgeons find the interaction design and graphical user interface to be adequate for manipulating 3D medical images. For instance, one of the participants referred that "having the possibility to clip the volume and visualize the data magnified by a large scale factor, without asking a coworker to manipulate the images, is fundamental for us surgeons". Although none of the participants reported arm fatigue during the session, all participants adverted that prolonged hand gesture manipulation may cause discomfort, suggesting the need for an improved "economy of gestures by keeping the elbows close to the trunk while gesticulating ones hands within smaller ranges of motion". Even so, users rapidly understood how to interact with the touchless interaction system as they easily manipulated the volumes, without having used the spatial interaction system before. Despite most of the participants being right-handed, none reported any issue

Table 4
Questionnaire results regarding user experience and preferences. R – rotation; CR – constrained rotation; T – translation; CB – clipping box; CP – clipping plane rotation and translation; D – distance measurement; A – angle measurement; Tg – tagging. Median (inter-quartile range). Likert scale: 1 – totally disagree and 6 – totally agree.

	R	CR	T	CB	CP	D	A	Tg
Usefulness	5(1)	5(1.75)	4(0.75)	5(0.75)	5(0.75)	5(0.75)	5(1.75)	5(1.5)
Easiness	5(1.75)	5(1.75)	5(1.75)	4(1.75)	4(1.5)	4(2.5)	3(1.75)	3(1.75)
Gesture	5(1)	5(1)	5(0.75)	4(1)	4(1.75)	4(1.5)	3(1)	4(1.5)
Recallability	5(1.5)	5(0.75)	6(1.75)	5(1.5)	5(1.5)	4(0.75)	4(1.5)	4(0.75)

towards using the left hand to perform all manipulation tasks. Interestingly, most surgeons reported that translation is not a very meaningful tool for 3D image manipulation and that they are more accustomed to interact with planar medical images (e.g., axial, sagittal and coronal planes) during surgery than with 3D medical images.

Table 4 reports the participants opinions regarding usefulness and performance of the interaction techniques employed. All participants considered the gestures and the visual feedback to be adequate (Table 4) and, in general, managed to achieve the desired 3D point of view with a median of 5 (IQR = 2.5) for rotation, median of 5 (IQR = 1.5) for constrained rotation and a median of 5 (IQR = 2.25) for translation. When asked whether it was easy to transition between rotation, translation and scale by using the simple set of hand gestures (Fig. 4), participants classified it with a median of 5 (IQR = 1).

Concerning the graphical user interface, the queried surgeons considered that the amount of buttons was sufficient for surgical navigation, with median of 5 (IQR = 0.75), as they mentioned that the existence of all the buttons did not force them to memorize all of the commands. They noticed the importance of having pop up lateral menus that are reachable and the clarity of the text associated to each functionality, hence considered that the lateral menus were adequate, with a of 5 (IQR = 0.75). They also found that the wireframe box that surrounds the volume data was informative with a median of 5 (IQR = 0.75). One of the senior surgeons mentioned that Voxel Explorer "is an extraordinarily useful tool to be used on a daily basis, but the interface needs to be tailored with the individual needs of each specialty. For instance, it would be useful to also have sub-windows displaying 2D slices".

In terms of viability, the surgeons demonstrated a good level of acceptance and willingness to use this technology in surgical practice as the proposed interface promotes an effective interactive manipulation of the volume. Moreover, all of the queried physicians consider such touchless interface as a powerful alternative to traditional WIMP controls to understand the patient's anatomical information, not only because of asepsis constraints but also due to the interaction space inside the operatory block is relatively limited due to the presence of several workers and equipment. Furthermore, our setup's low cost and portability were referred to as excellent features for surgical settings.

In their interviews, all surgeons noticed the importance magnifying the volume data which is an important advantage of the system compared to WIMP systems. According to them, the system provides a very large screen area that displays magnified anatomical information in a very responsive manner. Such is fundamental during surgical navigation, as surgeons constantly require to see small and detailed anatomical features contained within the images.

In terms of recallability, one surgeon stated that "our visual memory does not always match what we see, but using this system does help to clearly identify subject-specific anatomical details that are very important to remember during surgery". For instance, the general surgeons mentioned that "hepatic surgery is very image demanding in the sense that millimetric details condition the surgeon's judgment, forcing to constantly remember anatomical details and consult the images".

All surgeons referred Voxel Explorer as a desirable tool provided that the following limitations were overcome: (i) the lateral planes of the clipping box should be more accessible, (ii) the Kinect sensor does not recognize well the approaching and moving away hand movement which is used to translate the plane; (iii) tagging, distance and angle measurements lacked the required precision for surgical; (iv) the time required to activate and select a tool on the menus was too prolongated and needs to be reduced; (v) transitioning between translation-rotation-scaling operations could deconfigure the desired spatial setting; (vi) rotation could be improved by allowing access to the roll angle using both hands (as if steering a wheel) or, better yet, by accessing all six degrees of freedom of a single hand; and (vii) the rotation operation was slightly sensible to hand gestures (i.e., relatively small covered distances would produce slightly large rotations). Additionally, one of the participants pointed out it would be useful to resort on the right hand to apply fine tuned translations or rotations.

Finally, each surgeon highlighted the importance of the touchless interface for educational and training purposes, namely, to train surgical procedures to interns and for students endorsed in radiology anatomy courses. Such interactive visualizations of anatomical structures would enable students and interns to understand and explore anatomy more effectively; hence, students would be better prepared to deal with real life scenarios were such 3D medical images take part.

6. Conclusion and future work

This paper addresses the use of simple hand gestures as interactive 3D cursors to explore volumetric medical images. Hand gestures and body postures are directly mapped to 3D affine transformations of the volume, avoiding a physical interface. A comparative study between a conventional volume rendering system and a touchless interface was conducted among users with and without a medical image background. Results from the usability tests, where users were asked to rotate and position medical volumes and clipping planes using their hands as real 3D cursors without any wearable device, indicate that such a toucheless approach is easy to use, promotes a greater awareness of the image data, and provides a more efficient manipulation of volume data when compared to conventional WIMP applications. User's also considered the touchless interface easy to learn, even though they were not used to such spatial interaction. The use of a third dimension for a cursor, while unfamiliar at first, results in a much more direct manipulation of volume data and produces faster results when compared to traditional applications.

We conducted an evaluation with several surgeons of different specialities and years of experience, who routinely analyze medical images. All surgeons were positively impressed, suggested that the user experience and interaction methods were adequate, and highlighted many promising facets in the proposed touchless interface. Most importantly the evaluation results with professional surgeons are encouraging and bring about the possibility of adopting touchless interface as an intra-operative surgical navigation, surgical training or medical educational tool in everyday practice.

However, noisy measurements, incorrect gesture recognitions, and faulty skeleton tracking are still serious limitations that need to be addressed before this tool can be used in highly demanding scenarios, such as surgical navigation. In particular, the most prominent hurdle revealed during task performance was the lack of precision offered by the depth cameras. While users were relatively quick to identify the structures of interest, they had some trouble selecting the desired slice or placing the markers in the correct position, thus increasing the time necessary to complete the task. Another limitation of the usability studies performed was that the proposed the proposed system was only tested with 7 professionals. More participants would provide greater statistical significance. Even so, the results were validated from feedback collected from professionals and points towards new research paths regarding touchless surgical navigation systems.

As future work, it is necessary to improve the current state of the touchless interface so that professional users can perform spatial interactions while maintaining asepsis in sterile environments such as the operating room. This would facilitate interactive visualization of medical images without the need to wear any special equipment and aiding surgical planning and navigation tasks, hence, a greater involvement of the surgeons is required. Since the conducted user tests took place in closed, controlled laboratorial environment, the proposed interface needs to be tested in surgical environments to better determine how well it behaves in real situations, where possible tracking limitations may occur in scenarios crowded with staff and equipment may occlude the depth cameras performance.

Conflict of interest

None to report.

Acknowledgements

All authors are thankful for the financial support given by Portuguese Foundation for Science and Technology (FCT). In particular, the first author thanks for the postdoctoral grant SFRH/BPD/97449/2013. This work was also partially supported by national funds through FCT with reference UID/CEC/50021/2013 and IT-MEDEX PTDC/EEI-SII/6038/2014.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbi.2017.07.009.

References

[1] Adora – http://adora-med.com/#section_video.
[2] C. Amon, F. Fuhrmann, F. Graf, Evaluation of the spatial resolution accuracy of the face tracking system for kinect for windows v1 and v2, in: Proceedings of the 6th Congress of the Alps Adria Acoustics Association, 2014.
[3] H. Azari, I. Cheng, A. Basu, Stereo 3D mouse (S3D-Mouse): Measuring ground truth for medical data in a virtual 3D space, in: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, 2009, pp. 5744-5747. 10.1109/IEMBS.2009.5332603.
[4] D. Coffey, N. Malbraaten, T.B. Le, I. Borazjani, F. Sotiropoulos, A.G. Erdman, D.F. Keefe, Interactive slice WIM: navigating and interrogating volume data sets using a multisurface, multitouch VR interface, IEEE Trans. Visual. Comput. Graph. 18 (10) (2012) 1614–1626, http://dx.doi.org/10.1109/TVCG.2011.283.
[5] A. Corti, S. Giancola, G. Mainetti, R. Sala, A metrological characterization of the Kinect V2 time-of-flight camera, Robot. Auton. Syst. 75 (2016) 584–594, http://dx.doi.org/10.1016/j.robot.2015.09.024.
[6] F. Decle, tBox: A 3D transformation widget designed for touch-screens, Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, 2011, 3005–3008. http://doi.org/10.1145/1978942.1979387.
[7] S. Diepenbrock, J. Praßni, F. Lindemann, H. Bothe, T. Ropinski, Interactive visualization techniques for neurosurgery planning, Eurographics (2011) 13–16.
[8] S. Dolnicar, B. Grün, F. Leisch, Quick, simple and reliable: forced binary survey questions, Int. J. Market Res. 53 (2) (2011) 231–252.
[9] L. Gallo, A.P. Placitelli, M. Ciampi, Controller-free exploration of medical image data: experiencing the Kinect, in: Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on, 2011, pp. 1–6, 27–30 June 2011, doi: 10.1109/CBMS.2011.599913.
[10] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, L. Rochester, Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease, Gait Posture 39 (4) (2014) 1062–1068, http://dx.doi.org/10.1016/j.gaitpost.2014.01.008.
[11] Gestsure, 2015, Retrieved October 27, 2015, http://www.gestsure.com/.
[12] Y. Guiard, Yves Guiard Centre National de la Recherche Scientifique & Université de la Méditerranée CP 910, 163 avenue de Luminy 13288 Marseille cedex 09 – France, J. Motor Behav. 4 (1987) 1–23.
[13] Passive real-world interface props for neurosurgical visualization, Hinckley et al. CHI '94, pp. 452–458.
[14] P. Issartel, F. Guéniat, M. Ammi, Slicing techniques for handheld augmented reality, 3D User Interfaces (3DUI), IEEE Symposium on Minneapolis, MN 2014 (2014) 39–42, http://dx.doi.org/10.1109/3DUI.2014.6798839.
[15] M.G. Jacob, J.P. Wachs, R.A. Packer, Hand-gesture-based sterile interface for the operating room using contextual cues for the navigation of radiological images, J. Am. Med. Infor. Assoc.: JAMIA 20 (e1) (2013) e183–e186, http://dx.doi.org/10.1136/amiajnl-2012-001212.
[16] M.G. Jacob, J.P. Wachs, Context-based hand gesture recognition for the operating room, Pattern Recogn. Lett. 36 (2014) 196–203, ISSN 0167-8655, http://dx.doi.org/10.1016/j.patrec.2013.05.024.
[17] V. Konchada, B. Jackson, T. Le, Iman Borazjani, F. Sotiropoulos, D.F. Keefe, Supporting internal visualization of biomedical datasets via 3D rapid prototypes and sketch-based gestures, in: Symposium on Interactive 3D Graphics and Games (I3D '11). ACM, New York, NY, USA, 2011, pp. 214–214, doi: http://dx.doi.org/10.1145/1944745.1944794.
[18] Lachat, E., Macher, H., Mittet, M. -a., Landes, T., & Grussenmeyer, P. (2015). First Experiences With Kinect V2 Sensor for Close Range 3D Modelling, ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W4(February), pp. 93–100. http://doi.org/10.5194/isprsarchives-XL-5-W4-93-2015.
[19] B. Laha, D.a. Bowman, Poster: volume cracker: a bimanual 3D interaction technique for analysis of raw volumetric data. 3D User Interfaces (3DUI), IEEE Symposium on, C 147–148 (2013), http://dx.doi.org/10.1109/3DUI.2013.6550221.
[20] B. Laha, D.A. Bowman, J.D. Schiffbauer, Validation of the MR simulation approach for evaluating the effects of immersion on visual analysis of volume data, in: IEEE Transactions on Visualization and Computer Graphics, Proceedings of Virtual Reality, 2013, vol. 19, no. 4, pp. 529–538.
[21] M.R. Marner, R.T. Smith, J.a. Walsh, B.H. Thomas, Spatial Interfaces, 2014.
[22] K. Montgomery, M. Stephanides, S. Schendel, M. Ross (2005). User interface paradigms for patient-specific surgical planning: lessons learned over a decade of research, Computerized Medical Imaging and Graphics, vol. 29, no. 2–3, March–April 2005, pp. 203–222, ISSN 0895–6111, http://dx.doi.org/10.1016/j.compmedimag.2004.09.014.
[23] K. Mühler, C. Tietjen, F. Ritter, B. Preim, The medical exploration toolkit: an efficient support for visual computing in surgical planning and training, IEEE Trans. Visual. Comput. Graph. 16 (1) (2010) 133–146, http://dx.doi.org/10.1109/TVCG.2009.58.
[24] Myo Gesture Control Armband. Retrieved July 2016 from https://www.myo.com/.
[25] K. O'Hara, N. Dastur, T. Carrell, G. Gonzalez, A. Sellen, G. Penney, M. Rouncefield, Touchless interaction in surgery, Commun. ACM 57 (1) (2014) 70–77, http://dx.doi.org/10.1145/2541883.2541899.
[26] L.L.C. Polidais, Changing the Landscape. How Medical Imaging Has Transformed Health Care in the U.S., 2006, p. 28. http://doi.org/703.841.3200.
[27] B. Polley, Form and Function 3D: a Gesture-based User Interface by Brendan Polley. Retrieved October 27, 2015, https://vimeo.com/108086973.
[28] I. Saul, This New Google Project Is So Futuristic You Won't Be Able To Understand It. Retrieved October 27, 2015, http://aplus.com/, http://aplus.com/a/google-soli-project-future-is-here.
[29] P. Song, W.B. Goh, C.-W. Fu, Q. Meng, P.-A. Heng, WYSIWYF: Exploring and Annotating Volume Data with a Tangible Handheld Device ACM Conference on Human Factors in Computing Systems (CHI), 2011, pp. 1333–1342.
[30] R. Teather, Comparing 2D and 3D Direct Manipulation Interfaces Retrieved from http://www.cas.mcmaster.ca/~teather/pdfs/thesis.pdf, 2008.
[31] H.W. Yao, X.Y. Shi, M. Chen, R. Rong, Y.H. Liu, (2013); Improvement of standardization and accuracy of medical imaging and pathological diagnosis for better diagnosis and treatment of rectal cancer, Zhonghua Wei Chang Wai Ke Za Zhi 16 (6) (2013) 505–508. PMID: 23801198.