Evaluation of haptic virtual reality user interfaces for medical marking on 3D models

Zhenxing Li, Maria Kiiveri, Jussi Rantala and Roope Raisamo

Faculty of Information Technology and Communication Sciences

Tampere University, Tampere, Finland

Abstract— Three-dimensional (3D) visualization has been widely used in computer-aided medical diagnosis and planning. To interact with 3D models, current user interfaces in medical systems mainly rely on the traditional 2D interaction techniques by employing a mouse and a 2D display. There are promising haptic virtual reality (VR) interfaces which can enable intuitive and realistic 3D interaction by using VR equipment and haptic devices. However, the practical usability of the haptic VR interfaces in this medical field remains unexplored. In this study, we propose two haptic VR interfaces, a vibrotactile VR interface and a kinesthetic VR interface, for medical diagnosis and planning on volumetric medical images. The vibrotactile VR interface used a head-mounted VR display as the visual output channel and a VR controller with vibrotactile feedback as the manipulation tool. Similarly, the kinesthetic VR interface used a head-mounted VR display as the visual output channel and a kinesthetic force-feedback device as the manipulation tool. We evaluated these two VR interfaces in an experiment involving medical marking on 3D models, by comparing them with the present state-of-the-art 2D interface as the baseline. The results showed that the kinesthetic VR interface performed the best in terms of marking accuracy, whereas the vibrotactile VR interface performed the best in terms of task completion time. Overall, the participants preferred to use the kinesthetic VR interface for the medical task.

Keywords— 3D visualization, Virtual reality, Haptic feedback, Force-feedback device, Medical marking.

1. Introduction

Three-dimensional (3D) visualization has been widely used in many professional fields, such as medicine (Preim and Bartz, 2007), architecture and industrial manufacturing (Bouchlaghem et al., 2005). In medicine, 3D visualizations of the human skeleton, organs and other anatomic structures are implemented based on radiological imaging, such as computed tomography (CT) and magnetic resonance imaging (MRI) scans (Sutton, 1993). 3D visualization technique offers numerous benefits. For example, volumetric medical images can enable medical students to better understand the spatial anatomy of body organs (Silén et al., 2008), improve the accuracy of medical diagnoses (Satava and Robb, 1997) and help surgeons plan and simulate surgical procedures (Gross, 1998).

Highlighting relevant points on volumetric medical images of CT and MRI scans is an important 3D manipulation task performed by medical practitioners in computer-aided medical diagnosis and planning. Medical practitioners manipulate the models (i.e., rotate, pan and zoom) and mark critical points for later inspection, measurement and analysis of skeletal relationships (Kula and Ghoneima, 2018), treatment planning (Harrell, 2007) and as a tool for discussing and developing treatment consensus (Reinschluessel et al., 2019). For example, during cephalometric tracing, medical practitioners select and mark a point on the skeleton model or surrounding soft tissue as a point of reference for operations related to positioning, measurement and orientation. The task difficulty depends on the marking locations and the structure complexity of the virtual models (Medellín-Castillo et al., 2016). The accuracy of the markers directly influences the results of the medical analyses, and thus, the overall quality of the medical services (Lindner et al., 2016).

Despite the recent advances in 3D visualization technology, the tools used to present and interact with these volumetric images have not changed in the field of medicine. A conventional 2D display is still the main visual channel to present volumetric data from CT and MRI scans, which provides the user with a fixed screen-based viewing perspective. Further, it is still a common practice to use a mouse with the rotate-pan-zoom technique to indirectly manipulate 3D models (Jankowski and Hachet, 2013). However, previous studies (Hinckley et al., 1997; Bowman et al., 2004) have argued that using the mouse-based interface for 3D manipulation is difficult. Some researchers have investigated the mouse-based rotation techniques to understand their issues for 3D manipulation (Bade et al., 2005). Other researchers have conducted comparative studies to

examine the usability of other user interfaces such as the tangible interface (Besançon et al., 2017) and the touchscreen-based interface (Yu et al., 2010). However, these interaction methods either did not exceed the performance of the mouse (Yu et al., 2010) or had a limited application area (Besançon et al., 2017).

Following the technical advances in virtual reality (VR), VR equipment (e.g., Oculus Rift (Oculus, 2020) and HTC Vive (VIVE, 2020)) has been developed. A combination of a VR headset and a handheld VR controller can provide the user with an intuitive and immersive interaction environment. In this VR interface, 3D models are presented to the user through the head-mounted display and the models can be manipulated by the user using the VR controller with six degrees of freedom (Oculus, 2020; VIVE, 2020). Compared with the traditional 2D interface, VR devices offer a flexible 3D view based on the position and orientation of the user's head and allow using 3D hand gestures to manipulate the objects. In addition, the VR controller can provide vibrotactile feedback to the user's hand and enable tactile interaction. Vibrotactile feedback as an augmentative sensory channel has many medical applications, such as, robot-assisted teleoperation (Peddamatham et al., 2008), minimally invasive surgery (Schoonmaker and Cao, 2006) and rehabilitation medicine (Shing et al., 2003). Because of the flexible viewing perspective and the natural hand-based input, the VR interface has been proposed to use in the field of medicine. For example, it has been employed to interact with skeleton and organ models for anatomy learning (Fahmi et al., 2019) and treatment planning (Reinschluessel et al., 2019). Multiple companies have employed it to develop software for medical diagnosis services (e.g., Surgical Theater, 2020; Adesante, 2020). However, the potentially beneficial vibrotactile feedback generated from the VR controller was not used in their interactive VR systems.

Further, force-feedback devices, such as the Geomagic Touch (3D systems, 2020) and the Novint Falcon (Novint, 2020), have been proposed as another beneficial interaction device for medical services (Ribeiro et al., 2016). These devices can support bidirectional kinesthetic exploration. The mechanical arm of the devices not only allows hand-based motions with six degrees of freedom for object manipulation but also transfers the generated kinesthetic feedback to the hand, to simulate the feeling of touch (Massie and Salisbury, 1994). Force-feedback devices have been used with a 2D display for, for example, anatomy education (Kinnison et al., 2009), surgery training (Steinberg et al., 2007; Webster et al., 2004) and medical analysis (Medellín-Castillo et al., 2016). The only study that has combined the force-feedback device with the VR headset, to the best of our knowledge, is the work by Saad et al. (2018). They have technically investigated the feasibility to connect these devices.

We combined a VR headset with a VR controller and a force-feedback device to create haptic VR interfaces which provide the user with a flexible viewing perspective, a natural hand-based input and haptic feedback simultaneously. These VR interfaces can enable intuitive and realistic 3D interaction, thus promising for the tasks involving 3D manipulation (Bowman et al., 2004) such as medical diagnosis and planning tasks. However, the usability of the haptic VR interfaces for these medical tasks, covering effectiveness, efficiency and satisfaction (Issa and Isaias, 2015), has not been explored. Furthermore, these two VR interfaces are based on similar interaction models but employ different interaction devices with different types of haptic feedback. Their difference in usability remains unclear in the context of medical diagnosis and planning. A comparison of two VR interfaces can help better understand the suitability of their interaction methods for 3D manipulation and reveal the effects of different types of haptic feedback for these high-standard medical tasks. More importantly, 2D interaction method using a mouse and a 2D display is still a powerful user interface and dominant in the field of medicine. A comparative study with the 2D interaction technique is necessary to explore the potential of the haptic VR interfaces to improve current medical diagnosis and planning work.

In the present study, we examined the two haptic VR interfaces, the kinesthetic VR interface using a force-feedback device and the vibrotactile VR interface using a VR controller, in an experiment involving medical marking on 3D models. To examine their practical usability, we compared two VR interfaces with the traditional 2D interface that uses a mouse and a 2D display as the baseline. In the experiment, because the structural complexity of the models and the marking locations can influence users' performance in the medical marking task, we employed three human anatomic structures as the experimental models with two different difficulty levels for the marking positions. To evaluate the three user interfaces, we collected both objective and subjective data. The objective data included task completion time and marking accuracy, and the subjective data included rating data for the perceived mental effort, hand fatigue, naturalness, immersiveness and user preference. The aim of the study was to answer the following questions in the context of medical marking:

- What are the differences between the kinesthetic VR interface and the vibrotactile VR interface, in terms of task completion time, marking accuracy and user experience? How do the marking locations affect users' performance with the two VR interfaces?

- What are the differences between the two VR interfaces and the traditional 2D interface? How do the marking locations affect users' performance with the traditional 2D interface?

This study makes the following contributions: We proposed two haptic VR interfaces to interact with volumetric medical images for computer-aided medical diagnosis and planning. The vibrotactile VR interface and the kinesthetic VR interface were evaluated based on a medical diagnosis and planning task on virtual models of the human skeleton and organ. The results revealed the strengths and weaknesses of two VR interfaces associated with current popular VR equipment and haptic device, which simultaneously provided empirical understanding for developing efficient and user-friendly interactive VR systems. In addition, through comparing with the 2D interaction technique, the better performances of two VR interfaces, in terms of marking accuracy (the kinesthetic VR interface) and task completion time (the vibrotactile VR interface), demonstrated their potential to replace the traditional 2D interface for these medical tasks.

The paper is organized as follows. Relevant previous studies are introduced, and then the prototype system and the experiment are described. The results are presented in detail, followed by the discussion of the main findings and the conclusion.

2. Background

2.1 2D and 3D visualization in the field of medicine

In current medical imaging systems, the most common information visualization is based on 2D slices. Slice-by-slice views support accurate exploration and diagnosis of medical imaging data (Tietjen et al., 2006). At the same time, 3D volume-rendering visualization has become a valuable technique in the diagnosis and planning phases. It helps medical staff in understanding 3D spatial relations and an overview of the model structure, as well as facilitates diagnostic analysis (Tietjen et al., 2006). For example, cephalometry analysis is an important tool in orthodontics (Kula and Ghoneima, 2018). Traditional cephalometry analysis on 2D slices suffers from visual distortion of skull structures and inaccurate marking locations, due to the overlap of skull structures in the 2D view (Lindner et al., 2016; Bholsithi et al., 2009). Some studies (Olszewski et al., 2007; Katsumata et al., 2005; Troulis et al., 2002) showed that cephalometry analysis on 3D models can improve the precision of the diagnoses. However, other studies (Van Vlijmen et al., 2010; Swennen and Schutyser, 2007) argued that compared to 2D-cephalometry, marking on the models during 3D-cephalometry analysis is difficult and time-consuming. These studies as examples indicated the importance of accurate medical marking on 3D models, but all were conducted using the traditional 2D interface based on a mouse and a 2D display.

In the present study, we contributed to this line of research by introducing two haptic VR interfaces to address the issues of medical marking on 3D models. Because the 3D models and the marking locations in medical diagnosis vary depending on the specific medical purpose, the experiment of this study involved marking tasks using several models of different parts of the human body with multiple selected marking positions. The aim was to examine the general usability of the VR interfaces for high-standard medical diagnosis and planning.

2.2 The dominant 2D interactive system

In traditional human-computer interaction systems, a 2D display is commonly used as the visual output channel, while the input may be provided with devices such as mice, trackballs (Imagine Media, 1996), joysticks and touch pads. Among the interaction devices, a mouse with a 2D display constitutes the most popular interactive system used in the field of medicine. To visually present 3D models using a 2D screen, a 3D projection technique is needed. By transforming and mapping 3D objects onto a 2D plane, the projection provides the visual effect of 3D models on the 2D screen as realistic as the human visual view in the physical world (Foley et al., 1995). However, the viewing angle of interacting with the models is often fixed, due to the nature of the 2D screen.

To manipulate 3D models, the rotate-pan-zoom technique is commonly used with mouse-based interfaces (Jankowski and Hachet, 2013). For example, to rotate the models, rotation techniques, such as Virtual Sphere (Chen et al., 1988) and ArcBall (Shoemake, 1992), are widely used. Both techniques adopt a virtual ball around the manipulating object and calculate the rotation axis and angle by utilizing the projection of the mouse location onto the sphere (Jankowski and Hachet, 2013). The panning operation is typically implemented by using the mouse to point at the object and dragging the mouse along the x-y plane while pressing a button of the mouse. The zooming function often adopts the method of rolling the mouse wheel with a discrete zooming step. The mouse-based user interface relies on these techniques to indirectly manipulate 3D models.

However, using 2D interaction devices such as the mouse to manipulate 3D models can be difficult (Bowman et al., 2004). For example, a previous study demonstrated that using a multidimensional interaction device could achieve more efficient interaction than using the mouse-based interface in a 3D manipulation task (Hinckley et al., 1997). Bade et al. (2005) have compared the existing mouse-based 3D rotation techniques and provided their design principles to address this issue. In addition, multiple studies have examined other interaction methods for 3D manipulation tasks. Yu et al. (2010) presented a touchscreen-based data exploration technique for 3D manipulation. The interaction method was easy to learn and use, but its task performance could not exceed the performance using the mouse-based interface. Besançon et al. (2017) compared a tangible interface and a touchscreen interface with a mouse-based interface. The results showed that using the tangible interface could lead to a shorter task completion time. However, it had a limited application area compared to the mouse. There are advanced 3D interactive systems implemented by using VR equipment and force-feedback devices. In this study, we experimentally compared these interactive VR systems with the mouse-based interface and investigated their strengths and weaknesses in medical marking.

2.3 The interactive VR system using a VR headset and a VR controller

VR equipment offers a new interactive experience with virtual objects in the field of medicine. Previous studies have used the CAVE environment for visualizing medical imaging data (Shen et al., 2008; Al-Khalifah et al., 2006). Currently the head-mounted VR display presents a more flexible and natural 3D view based on the position and orientation of the user's head. Many studies have employed the VR headset to visually present 2D slices (Wirth et al., 2018; King et al., 2016) and volumetric imaging data (Venson et al., 2017; Sousa et al., 2017; Randall et al., 2016) for radiologists. Other studies proposed to use the VR headset for the medical purposes such as treating chronic pain (Jones et al., 2016), anxiety disorders and phobias (Maples-Keller et al., 2017).

The VR controller allows 3D hand gestures to manipulate objects with optional vibrotactile feedback which can enable tactile interaction (Oculus, 2020; VIVE, 2020). Tactile interaction, as one branch of haptic interaction, concentrates on touch interaction that stimulates mechanoreceptors in the human skin (El Saddik et al., 2011). This technique has been widely used in many medical fields. For example, vibrotactile feedback was used to enrich interaction channels while using a surgical robot (Peddamatham et al., 2008), improve surgeons' performance in minimally invasive surgery (Schoonmaker and Cao, 2006) and help users in using a hand rehabilitation system (Shing et al., 2003).

Because of the natural interaction method, the VR interface using a VR headset and a VR controller has been used in an anatomy learning system (Fahmi et al., 2019). Reinschluessel et al. (2019) have proposed to use it to interact with 3D organ models for treatment planning. Some products such as Precision VR by Surgical Theater (2020) and SurgeryVision by Adesante (2020) have employed the VR interface to manipulate 3D models of human anatomic structures for medical analysis and other services. However, the vibrotactile feedback generated from the controller was missing in their interactive VR systems.

While the VR interface has been widely used, it remains unclear how well it compares with the traditional 2D interface using a mouse and a 2D display. It is likely that the benefit and cost are dependent on the context of the interaction such as the type of the task and the complexity of the model. For example, a previous study in the field of 3D geological modelling, comparing a VR controller with a mouse in manipulating an industrial software, showed that the controller was more difficult to use and caused more fatigue on the user's hand (Kim and Choi, 2019). In the present study, we evaluated the efficiency and user experience of this VR interface with vibrotactile feedback to interact with 3D objects in the context of medical marking.

2.4 The interactive VR system using a VR headset and a force-feedback device

Using force-feedback devices enables kinesthetic interaction which is another type of haptic interaction (El Saddik et al., 2011). The kinesthetic technique focuses on movement sensations originating in the muscles, tendons and joints (El Saddik et al., 2011). It enables bidirectional touch exploration closely as realistically as in the physical world and has been used for many medical practices, such as medical education (Kinnison et al., 2009), surgery operation simulation and training (Alaraj et al., 2015; Steinberg et al., 2007; Webster et al., 2004; Bielser and Gross, 2000), robot-assisted surgery (Okamura, 2004) and medical analysis (Medellín-Castillo et al., 2016).

For medical education, a kinesthetic simulator (Kinnison et al., 2009) has been demonstrated to be an engaging and efficient method for teaching students the human anatomy, which addresses current challenges in teaching using real human tissues

(e.g., anxiety and fear). Other studies have suggested that kinesthetic simulation is an efficient and repeatable method for surgery training without wasting real surgical samples. Kinesthetic simulators have previously been used, for instance, in dental (Steinberg et al., 2007) and eye (Webster et al., 2004) surgery simulations, basic cutting (Bielser and Gross, 2000) and aneurysm clipping simulation (Alaraj et al., 2015). In addition, the kinesthetic technique brings benefits in the context of robot-assisted surgery. For instance, a previous study (Okamura, 2004) has argued that offering kinesthetic feedback contributes significantly to the safe performance of the surgical procedures. In medical analysis, kinesthetic feedback was used to enhance medical marking in 2D, 2.5D and 3D computer-aided cephalometry analysis (Medellín-Castillo et al., 2016). The results indicated that haptic-enabled 3D cephalometry marking is better than the haptic-enabled 2D or 2.5D cephalometry marking, in terms of task accuracy.

Combining a force-feedback device with a VR headset as a new VR interface to interact with 3D models is feasible (Saad et al., 2018). In the present study, we evaluated two haptic VR interfaces (i.e., using either a VR controller or a force-feedback device) and examined the different performances of vibrotactile and kinesthetic feedback in medical marking. To demonstrate the practical usability, we compared these two interfaces with the traditional 2D interface.

3. Method

3.1 Design of the prototype system

The experimental prototype system included three user interfaces: vibrotactile VR interface (V), kinesthetic VR interface (K) and traditional 2D interface (T). We use the abbreviations (V, K, T) for the three interfaces in figures.

The vibrotactile VR interface employed a VR headset as the visual output channel and a VR controller as the manipulation tool with vibrotactile feedback. The VR headset provided realistic visual feedback with a flexible head movement-based viewing perspective. While pressing the trigger button of the controller with the index finger (see Fig. 1A), the user could manipulate the model by rotating (rotation along the x-, y- and z-axes), panning (displacement along the x- and y-axes) and zooming (displacement along the z-axis), following the movement of the whole arm (i.e., rotation was done by rotating the wrist, and panning and zooming were done by moving the arm along the x-, y- and z-axes respectively). In addition, the user could feel vibrotactile feedback from the controller while the cursor touched the model. The cursor was visible as a yellow sphere in the virtual environment with a 0.2 cm radius hanging on the tip of the controller. The controller was invisible to the user in the virtual environment. The duration of the haptic pulse for the vibrotactile feedback was set to 80 ms.

The kinesthetic VR interface used the same VR headset as the visual output channel, but a force-feedback device was used as the manipulation tool. Consistent with the vibrotactile VR interface, the user could rotate, pan and zoom the model with the
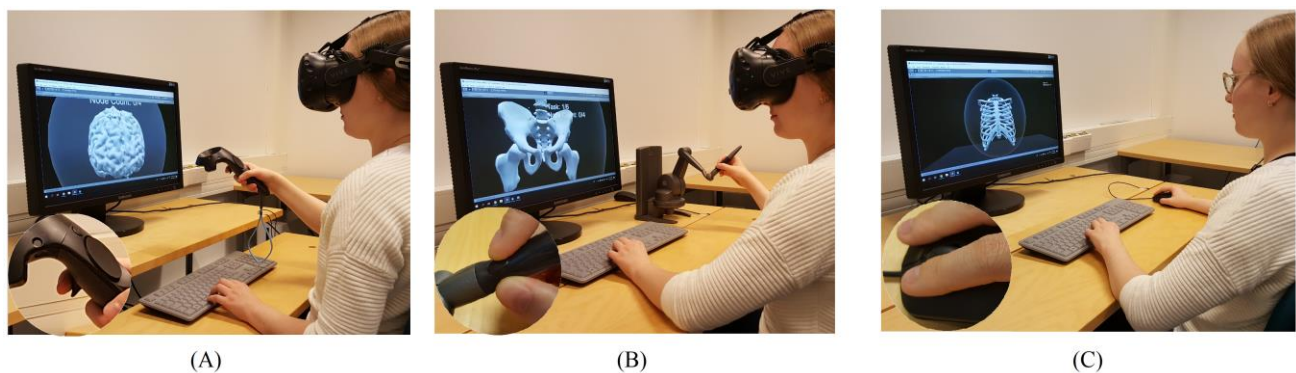

Fig. 1. (A) The vibrotactile VR interface using a Vive controller and a Vive VR headset. A close-up image of the trigger button of the controller is shown at the bottom left corner. The additional Samsung LCD 2D display (not visible to the user) shows the brain model. (B) The kinesthetic VR interface using a Geomagic Touch X force-feedback device and a Vive VR headset. A close-up image of the device button is shown at the bottom left corner. The additional 2D display shows the hipbone model. (C) The traditional 2D interface using a standard mouse with the Samsung LCD 2D display. A close-up image of the mouse buttons is shown at the bottom left corner. The screen shows the sternum model. With all three user interfaces, the space key on a standard keyboard was used to create the markers on the marking positions.

Table 1.

Specification of three user interfaces.

Interfaces	Equipment	Hand-based operation	Visual channel	Haptic feedback
Vibrotactile VR interface (V)	VR headset and controller	3D gesture along the x-, y- and z- axes	Head movement-based view	Vibrotactile
Kinesthetic VR interface (K)	VR headset and force-feedback device	3D gesture along the x-, y- and z- axes	Head movement-based view	Kinesthetic
Traditional 2D interface (T)	2D display and mouse	2D gesture along the x- and y- axes	Screen-based view	None

movement of the whole arm while pressing the button on the device with the index finger (see Fig. 1B). When the user touched the model with the cursor (visible as a yellow sphere with a 0.2 cm radius hanging on the tip of the device, and the device was invisible in virtual environment), he or she could feel kinesthetic feedback based on the surface properties of the model. The kinesthetic feedback involved the stiffness implemented by the linear spring law ($F = kx$), where k is the stiffness coefficient and x is the penetration depth of the interaction point, as well as the friction ($F' = \mu F_n$), where μ is the friction coefficient and $F_n = mg$, the normal force value.

With both haptic VR interfaces, the user could move the cursor to reach and touch any position on the medical model as the selected marking point for the experimental task.

The traditional 2D interface employed a standard 2D display as the visual output channel and a mouse as the manipulation tool. The 2D display also provided realistic visual feedback with a screen-based viewing angle. To rotate 3D models, the user could move the mouse along the x- and y-axes on the screen while holding down the left mouse button (see Fig. 1C). The rotation was based on the ArcBall technique with a selected rotation speed parameter of 3.5. Panning of the model (along the x- and y- axes) was implemented by dragging the model using the mouse while pressing down the right mouse button. Zooming of the model was done by rolling the mouse wheel with a selected zooming speed factor of 6. Marking on the model was based on the ray-casting technique (Poupyrev et al., 1998). This technique utilized the pointer position of the mouse on the 2D screen as the starting point of a virtual beam that traversed the virtual space along the viewing direction of the 2D screen. When the beam intersected with the 3D model, the intersecting point would be the selected marking point (Gallo et al., 2010). In addition, the mouse-based interface had no haptic feedback.

For all three user interfaces, the space key on a keyboard was used to create markers (visible as green nodes with a 0.2 cm radius) on the selected marking points. The three interfaces with the device buttons are shown in Fig. 1 and the specification of the user interfaces are listed in Table 1.

3.2 Experiment design

A within-subject experiment was conducted in a controlled laboratory setting with three experimental conditions: the vibrotactile VR interface, the kinesthetic VR interface and the traditional 2D interface. The experimental task was medical marking on 3D models, and there were three models involved in the experiment: the sternum, hipbone and brain (see Fig. 1). All models were inside a 30 cm diameter sphere. This ensured that the models were of a similar size. The selected stiffness coefficients (sternum and hipbone: 1.0; brain: 0.4) and friction coefficient (all: 0.7) of the models were set for kinesthetic feedback.

The task-related marking positions (visible as red dots with a 0.1 cm radius) were predefined on the models before the experiment. To mark these positions, the participants employed the interaction tools to manipulate the model for visually searching the marking positions (through the visual display) and then reach (using the VR controller or the force-feedback device) or point (using the mouse) at the positions on the model for determining the positions the participants wanted to mark. Markers were created on the positions by pressing the space key on the keyboard.

We varied the marking positions as an independent variable in the experiment. Because of the complex characteristic and structure of human anatomic models, we generally categorized the marking positions into two difficulty levels (easy and difficult). We placed the model away from the viewpoint (45 cm from the viewpoint to the centre of the model along z-axis)
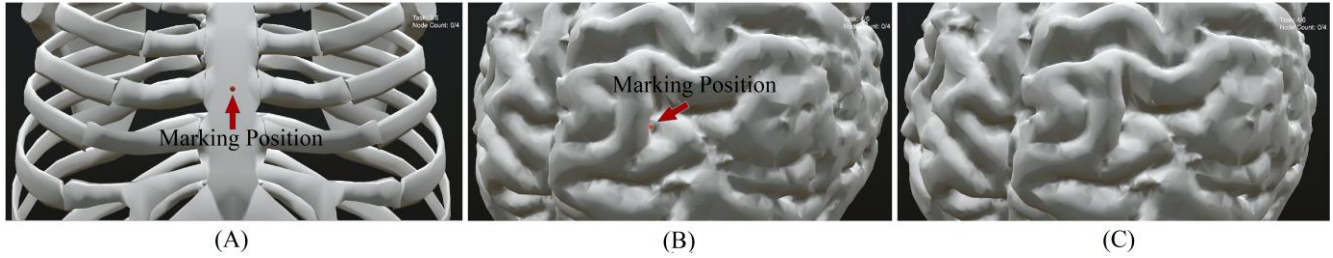
Fig. 2. (A) shows an example of the easy-to-mark task in which the marking position (red dot) was distinct on the model surface. (B) shows an example of the difficult-to-mark task in which the marking position was only visible from some viewing angles. (C) shows that the marking position could not be visible from some other viewing angles.

Table 2.
Statements in the questionnaire.

Statements	Description
S1	This user interface is mentally easy to use.
S2	This user interface does not make the hand tired.
S3	This user interface is natural to use.
S4	This user interface is immersive.

to make the model fully visible to the user and then determined the difficulty level of the marking position based on its visibility from the user's view. If the marking position could be visually blocked by other parts of the model surrounding it while rotating the model (excluding the situation of rotating to other faces of the model which made the marking surface completely invisible), we categorized the position as difficult. In this case, the user had to manipulate the model carefully and find an appropriate angle to mark the position. In contrast, if the marking position could not be visually covered by other parts of the model surrounding it while rotating the model, the user was able to spend less effort manipulating the model and marking the position and thus we categorized the position as easy (Fig. 2 shows examples of two marking difficulty levels). We predefined 24 possible marking positions on the surface of each model (12 for easy to mark and 12 for difficult to mark) for the experiment task.

The participants performed six trials with each interface: three trials were easy to mark and three trials were difficult to mark. Each trial involved only one 3D model. The order of the three models (i.e., the sternum, hipbone and brain) was randomly assigned for each difficulty level, and thus, the models were used twice, once for each difficulty level. In each trial, the participants needed to mark 4 positions randomly selected from 12 possible positions based on the marking difficulty level of the trial. Therefore, for each participant, there were a total of 18 trials (3 interfaces × 6 trials = 18 trials) with 72 marking positions (18 trials × 4 positions = 72 positions).

The prototype system recorded the task completion times and the positions of the participants' markers as objective data. In addition, a 7-point Likert scale questionnaire was used to record subjective data, including perceived mental effort, hand fatigue, naturalness and immersiveness (1: strongly disagree to 7: strongly agree). The questionnaire was inspired by the NASA Task Load Index (Hart and Staveland, 1988). The statements used in the questionnaire are shown in Table 2. Moreover, we used a post-test questionnaire in which the participants ranked the three user interfaces based on personal preference and gave comments.

3.3 Pilot study

We recruited four participants who had previous experience using the VR equipment and the force-feedback device to conduct a pilot study. The participants were asked to use the prototype system and comment which parameters were the most suitable.

- For the kinesthetic VR interface, we did not scale the movement (Fischer and Vance, 2003), that is, the movement of the hand holding the mechanical arm resulted in the same amount of movement of the cursor in the virtual space along the x-, y- and z-axes. This enabled more accurate control of the cursor (Li et al., 2020).
- For the traditional 2D interface, the parameter for the rotation (speed factor 3.5) was selected, so that users could quickly and accurately rotate the models. The zooming factor (6) was selected based on the distance between the viewpoint and the target model. With maximum zooming, the viewpoint could closely reach the model surface. By applying these parameters, users could smoothly use the mouse to manipulate the models for the medical task without any operational difficulty.
- For the vibrotactile VR interface, the duration of the vibrotactile feedback (80 ms) was selected. With this duration, the vibrotactile feedback could be easily perceived when the user touched the model, but it was not too disruptive to influence users' hand posture and stability.
- The sizes of the three experimental models within the sphere (30 cm diameter) were chosen so that they were suitable for the size of the workspace of the force-feedback device and the area of hand movement using the controller.
- The size of the interaction point (a sphere with a 0.2 cm radius) of two haptic interaction devices was chosen. It ensures that the interaction point was easy to control while simultaneously minimizing its influence on the marking accuracy. The size of the task-required marking positions (dots with a 0.1 cm radius) was selected so that it maintained a high-level requirement on the marking accuracy to examine the three user interfaces. At the same time, the size of the marking positions was big enough and thus clearly visible to the user.
- For the kinesthetic VR interface, stiffness coefficients (sternum and hipbone: 1.0 and brain: 0.4) and friction coefficient (all: 0.7) were chosen so that touching the virtual objects would resemble touching real objects as closely as possible.

The above parameter selections were made for the prototype system and the experiment, based on the users' performance and comments. We next introduce the experiment apparatus, participants and procedure.

4. Experiment

4.1 Apparatus and environment

The host computer in the experiment was an MSI GS63VR 7RF Stealth Pro laptop with an Intel i7-7700HQ processor, a GeForce GTX 1060 graphics card and a 16GB RAM. We used a Vive VR headset (VIVE, 2020) and a Samsung 245B Plus 24" LCD display as the visual displays in the experiment. A standard computer mouse, a Vive controller and a Geomagic Touch X force-feedback device (3D Systems, 2020) were utilized as the manipulation tools for the three user interfaces. A standard keyboard was used for the participants to mark the positions. The experimental setup and environment are shown in Fig. 1. The experimental system was developed on Unity (2020), along with SteamVR (Steam, 2020) and Geomagic OpenHaptics plugin (Unity haptic plugin, 2020).

4.2 Participants

Twenty-four participants were recruited from the local university community (9 women and 15 men). Their ages varied between 21 and 43 years (M = 28.33, SD = 7.65). No participants had previous experience with using a mouse and a 2D display to do medical marking or similar tasks. Seven participants had used a similar VR headset and a controller one or two times. Two participants had used a similar force-feedback device with a 2D display once, but they had no use experience with a VR headset. All participants reported that they were right-handed.

4.3 Experiment procedure

The participants were first introduced to the experimental task and the apparatus used. They signed an informed consent form and filled in the background information in the questionnaire before the experiment.

The order of the experimental conditions and the marking difficulty levels was counterbalanced among the participants. For each condition, the participants were introduced to the user interface and had up to 5 minutes to familiarize themselves with the use of the related devices. They sat comfortably on a chair and looked at the 2D display or wore the VR headset, and then used their dominant hand to hold the manipulation tool and their non-dominant hand to press the space key of keyboard to

mark the positions. For the two VR interfaces, head movement to change their viewing perspective was allowed. We did not provide any extra wrist-rest or arm-rest equipment for any of the interfaces.

The total experiment time for each participant was approximately one hour. The participants were informed that they needed to finish the experimental task as accurately and quickly as possible and the accuracy was the top priority. When each trial started, the model was placed at 45 cm away from the viewpoint along z-axis and its orientation was semi-random. We predefined a list of orientation values which could respectively make one of the six faces of the model to face the user (i.e., by considering the model as a cube) and the system randomly selected one of them as the starting orientation value for the model. For each marking position, the participants had only one chance to mark the position. In other words, they could not remove the marker and place it again. After the participants marked the four marking positions, the system automatically recorded the data and proceeded to the next trial. After finishing the six trials of one condition, the participants were asked to fill out the questionnaire, after which the next condition was executed. After completing all conditions, the participants commented the three interfaces and ranked them based on personal preference.

5. Results

We collected two types of objective data from the experiment: task completion times and the positions of the participants' markers. We used the task completion times to evaluate the interaction speed and calculated the error distances between the participants' markers and the marking positions required by the tasks to evaluate the marking accuracy. The Shapiro–Wilk Normality test showed that the data were not normally distributed (all $p < .001$). Thus, we analyzed the data using the 3×2 (interfaces × marking difficulty levels) aligned rank transform (ART) repeated-measures non-parametric ANOVA (Wobbrock et al., 2011) separately for the task completion times and the error distances. We used the Wilcoxon signed-rank test for post-hoc analysis of objective and subjective data and performed the Holm-modified Bonferroni correction (Holm, 1979) to control the family-wise type-1 error.

5.1 Task completion time

To evaluate the interaction speed of the three interfaces, we calculated the mean value of the task completion times of the six trials for each interface. Table 3 shows the p values of the ART ANOVA result for the task completion times. There were statistically significant main effects for both the interfaces and the marking difficulty levels. In addition, there was a statistically significant interaction effect between the interfaces and the marking difficulty levels.

Fig. 3 shows a boxplot of the task completion times with the three interfaces. The Wilcoxon signed-rank test showed that participants spent statistically significantly less time to complete the task while using the vibrotactile VR interface (M = 103.52, SD = 27.82) than while using the kinesthetic VR interface (M = 131.01, SD = 37.00; Z = −3.514, $p < .001$). The task completion time using the traditional 2D interface (M = 120.62, SD = 38.08) was not statistically significantly different from that using the vibrotactile VR interface (Z = −1.771, $p = .152$) and that using the kinesthetic VR interface (Z = −1.657, $p = .097$).

Not surprisingly, the main effect of the marking difficulty was statistically significant. Completing difficult marking trials (M = 173.44, SD = 50.76) took statistically significantly more time than completing easy marking trials (M = 63.33, SD = 12.49; Z = −4.286, $p < .001$). The more interesting part from the perspective of this study is the interaction effect between the interfaces and the marking difficulty levels.

Table 3.
Tests of within-subjects effects on task completion times.

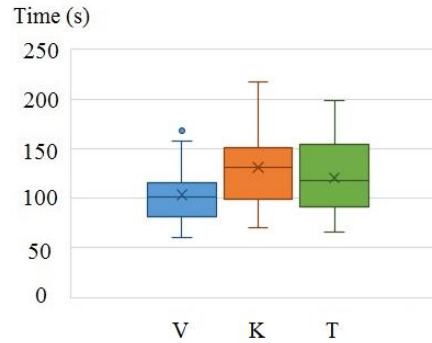| Sources	df	F-value	Sig.
Interfaces	2, 46	11.928	<0.001
Marking difficulty levels	1, 23	438.442	<0.001
Interfaces and marking difficulty levels	2, 46	8.963	0.001

Fig. 3. Overall task completion times for the three interfaces: the vibrotactile VR interface (V), the kinesthetic VR interface (K) and the traditional 2D interface (T). The cross mark shows the mean value and the line in the boxplot shows the median value (The same abbreviations and marks are used in the following figures).
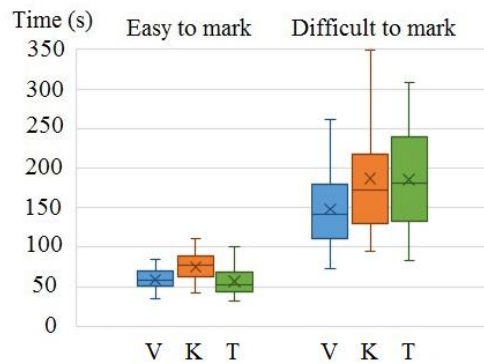

Fig 4. Task completion times using the three user interfaces, based on the marking difficulty levels.

Fig. 4 illustrates the task completion times of the three user interfaces based on the marking difficulty levels. According to the Wilcoxon signed-rank test, when the positions were easy to mark, the participants spent statistically significantly longer time using the kinesthetic VR interface (M = 74.83, SD = 17.77) than using the vibrotactile VR interface (M = 58.81, SD = 13.45; Z = −3.200, p = .002) and using the traditional 2D interface (M = 56.35, SD = 18.61; Z = −3.400, p = .003). There was no statistically significant difference in task completion times between the vibrotactile VR interface and the traditional 2D interface (Z = −0.800, p = .424). When the positions were difficult to mark, using the vibrotactile VR interface (M = 148.23, SD = 50.44) resulted in shorter task completion times than using the kinesthetic VR interface (M = 187.18, SD = 65.80; Z = −3.114, p = .006) and using the traditional 2D interface (M = 184.90, SD = 64.39; Z = −2.239, p = .05). There was no statistically significant difference in task completion times between the kinesthetic VR interface and the traditional 2D interface (Z = −0.514, p = .607).

5.2 Marking accuracy

To evaluate the marking accuracy of the three interfaces, we calculated the mean error distance for each trial consisting of four marking positions. Table 4 shows the results of the ART ANOVA for the distance data: the main effects of the interfaces and the marking difficulty levels, as well as their interaction effect, were statistically significant.

Fig. 5 shows a boxplot of the marking accuracy using the three interfaces. According to the Wilcoxon signed-rank test, the participants were statistically significantly more accurate in marking when using the kinesthetic VR interface (M = 0.13, SD = 0.04) than when using the vibrotactile VR interface (M = 0.35, SD = 0.20; Z = −4.286, p < .001) and the traditional 2D interface (M = 0.61, SD = 0.68; Z = −4.257, p < .001). There was no statistically significant difference in the marking accuracy between the vibrotactile VR interface and the traditional 2D interface (Z = −0.743, p =.458). For the main effect of the

Table 4.

Tests of within-subjects effects on distance data.

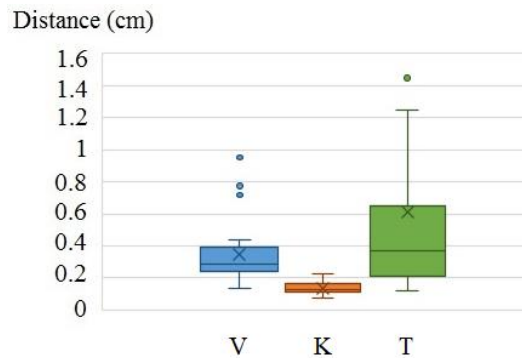| Sources	df	F-value	Sig.
Interfaces	2, 46	38.893	<0.001
Marking difficulty levels	1, 23	124.652	<0.001
Interfaces and marking difficulty levels	2, 46	39.728	<0.001

Fig. 5. Error distances between the marking positions and the participants' markers using the three interfaces.
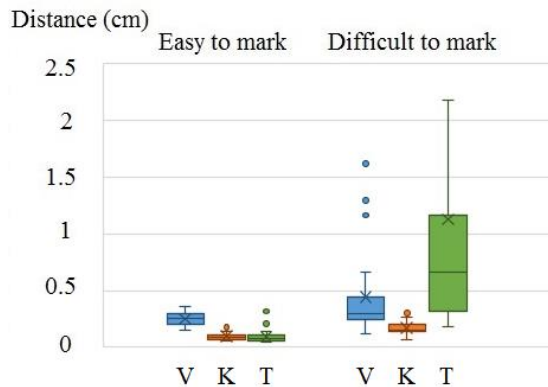

Fig. 6. Error distances between the marking positions and the participants' markers using the three interfaces, based on the marking difficulty levels.

marking difficulty, the participants were more accurate in marking when the positions were easy to mark (M = 0.15, SD = 0.04) than when the positions were difficult to mark (M = 0.58, SD = 0.46; Z = −4.286, p < .001).

Fig. 6 shows the marking accuracy using the three interfaces based on the marking difficulty levels. The Wilcoxon signed-rank test showed that when the positions were easy to mark, using the vibrotactile VR interface (M = 0.25, SD = 0.06) was statistically significantly less accurate than using the kinesthetic VR interface (M = 0.1, SD = 0.03; Z = −4.286, p < .001) and using the traditional 2D interface (M = 0.09, SD = 0.06; Z = −4.257, p < .001). There was no statistically significant difference between the kinesthetic VR interface and the traditional 2D interface (Z = −1.029, p = .304). When positions were difficult to mark, participants were statistically significantly more accurate in marking when using the kinesthetic VR interface (M = 0.17, SD = 0.06) than when using the vibrotactile VR interface (M = 0.44, SD = 0.38; Z = −4.286, p < .001) and using the traditional 2D interface (M = 1.13, SD = 1.36; Z = −4.257, p < .001). Using the vibrotactile VR interface was statistically significantly more accurate than using the traditional 2D interface (Z = −2.057, p = .04).

5.3 Subjective data

Fig. 7 shows the subjective data for perceived mental effort, hand tiredness, naturalness and immersiveness.

Mental effort. There were no statistically significant differences among the three interfaces in terms of perceived metal effort (the vibrotactile VR interface and the kinesthetic VR interface: $Z = -0.461$, $p = .645$; the kinesthetic VR interface and the traditional 2D interface: $Z = -1.685$, $p = .184$; the vibrotactile VR interface and the traditional 2D interface: $Z = -2.039$, $p = .123$).

Hand tiredness. The participants experienced statistically significantly more hand tiredness using the vibrotactile VR interface than using the kinesthetic VR interface ($Z = -2.913$, $p = .008$) and using the traditional 2D interface ($Z = -3.725$, $p < .001$). Furthermore, the participants experienced more hand tiredness using the kinesthetic VR interface than using the traditional 2D interface ($Z = -2.345$, $p = .019$).

Naturalness. The participants perceived the traditional 2D interface to be statistically significantly less natural than the kinesthetic VR interface ($Z = -2.721$, $p = .014$) and the vibrotactile VR interface ($Z = -2.994$, $p = .009$). There was no difference between the kinesthetic VR interface and the vibrotactile VR interface in terms of the naturalness ($Z = -0.661$, $p = .509$).

Immersiveness. The participants perceived the traditional 2D interface to be statistically significantly less immersive than the kinesthetic VR interface ($Z = -4.310$, $p < .001$) and the vibrotactile VR interface ($Z = -4.233$, $p < .001$). There was no difference between the kinesthetic VR interface and the vibrotactile VR interface ($Z = -0.966$, $p = .334$).

In addition, a post-test questionnaire was employed to collect data about user preference for the three interfaces. The participants could select multiple interfaces if they liked them equally. The results showed that the kinesthetic VR interface was voted the most preferred interface for the experimental task (19 votes), followed by the vibrotactile VR interface (6 votes) and the traditional 2D interface (5 votes). The participants also provided free form comments as follows.

P4 said, "*Controller excelled in rotation zoom and movement, but with no physical (kinesthetic) support, it felt very inaccurate for precision tasks.*" [preferred the kinesthetic VR interface and the traditional 2D interface].

P8 said, "*Finding the nodes with VR was way more efficient than 2D screen. The VR display felt natural and best suited for this task.*" [preferred the kinesthetic VR interface and the vibrotactile VR interface].

P22 said, "*Haptic (kinesthetic) feedback was more accurate in difficult positions. With the mouse it was almost impossible to landmark accurately. With the controller, the accuracy was also bad, especially towards the end. Haptic (kinesthetic) was much better in uneven/rough positions.*" [preferred the kinesthetic VR interface].

P24 said, "*The haptic (force-feedback) device gave the most realistic impression of the object, providing depth and 3D understanding*" [preferred the kinesthetic VR interface].

The above participants' comments as examples gave insights into their reason for preferring the interfaces. We next discuss the experiment results with the three user interfaces.
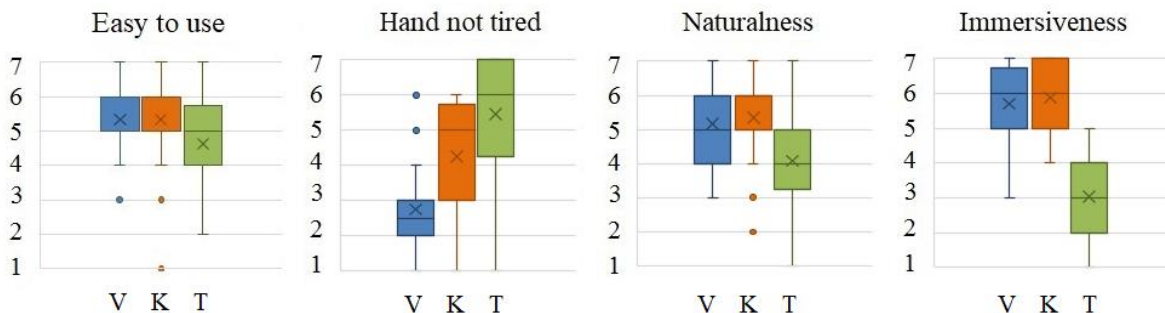

Fig. 7. Subjective results of the study.

6. Discussion

This study experimentally compared two haptic VR interfaces (vibrotactile and kinesthetic VR interfaces) with the traditional 2D interface. The interfaces varied in terms of the visual display as well as the manipulation tool. They were evaluated in the experiment of 3D medical marking task and the results showed that the different user interfaces influenced the task performance in terms of interaction speed and accuracy. In addition, the difference in the marking locations modulated the task performance. We now discuss the main findings in relation to the research questions.

6.1 Differences between the kinesthetic VR interface and the vibrotactile VR interface

The kinesthetic VR interface and the vibrotactile VR interface used the VR headset as the visual display and employed a force-feedback device and a handheld VR controller as the manipulation tool respectively. The results showed a speed-accuracy trade-off between the two interfaces. Generally, the vibrotactile VR interface was faster, but less accurate, than the kinesthetic VR interface (see Fig. 3 and Fig. 5).

Task completion time

Regardless of the marking difficulty, using the vibrotactile VR interface led to a shorter task completion time than using the kinesthetic VR interface (see Fig. 4). The two VR interfaces used the same VR headset as the visual display and allowed for head movement-based viewing perspective changes. The difference between two interfaces was in terms of the manipulation tool.

The vibrotactile VR interface using the controller allowed free hand-based input to reach the target and provided vibrotactile feedback as confirmation of contact with the 3D model. Further, natural 3D hand gestures allowed for efficient pan, rotation and zoom of the 3D models. All the marking tasks, especially the ones with high marking difficulty, required extensive use of these complex interactions to acquire the marking locations. In contrast, the mechanical arm of the force-feedback device and its length limited the flexibility of the gestures that can be used and the area that can be interacted with in the case of the kinesthetic VR interface (Massie and Salisbury, 1994). This often meant that the rotating, panning and zooming interactions were neither as intuitive nor as straightforward to perform. For example, when the user had to perform a large pan or rotation of the 3D model which was beyond the mechanical limits of the force-feedback device, the user had to break it down into multiple smaller pan movements or rotations. This may have contributed to the lower efficiency while using the kinesthetic VR interface.

Marking accuracy

The kinesthetic VR interface was more accurate in medical marking than the vibrotactile VR interface regardless of the marking difficulty (see Fig. 6). The kinesthetic and vibrotactile VR interfaces adopted a similar interaction mechanism, requiring the participants to reach and touch the model to determine the marking position. The different performance in marking accuracy is likely because of the different types of haptic feedback. Compared with the vibrotactile feedback which provided a simple vibration for touching, the bidirectional kinesthetic interaction allowed the user to understand the 3D structure of the model and the marking locations. This meant that when the surface near the marking location was uneven (e.g., the ridges and grooves on the brain surface), the user could accurately find the target location by relying on the kinesthetic feedback. Such feedback for the 3D structure was not available in the case of the vibrotactile VR interface.

Another important difference between the kinesthetic and vibrotactile VR interfaces was regarding the interaction boundary. The force-feedback device provided kinesthetic cues about the surface that was in contact. Once in contact with a surface, if the user applied further pressure inward, the device would produce a reaction force according to the material properties of the surface. It required a considerable amount of force for the kinesthetic interaction point to penetrate the surface of the 3D model. In the case of the vibrotactile VR interface, although vibrotactile feedback was provided when the user was in contact with the surface of the 3D model, any slight inward hand movement during the marking caused the controller to pass through the surface of the model leading to a depth error in the marking.

Further, the participants commonly performed mid-air hand gestures to touch the model for marking while using the VR controller. The stability of mid-air gestures is known to show considerable individual variability (Ángyán et al., 2007) and degrade with hand fatigue (Gates and Dingwell, 2011), which might have negatively affected the marking accuracy while using the vibrotactile VR interface. Although the participants also performed mid-air hand gestures while holding the arm of

the force-feedback device, the kinesthetic feedback could help maintain the hand stability during the touching process. This might further improve the marking accuracy while using the kinesthetic VR interface.

Previous studies have shown that kinesthetic feedback is a valuable complementary output modality in the fields of, for example, palpation simulation (Ribeiro et al., 2016) and robot-assisted surgery (Okamura, 2004). In the context of medical diagnosis and planning, Medellín-Castillo et al. (2016) have applied kinesthetic feedback for the 2D, 2.5D and 3D cephalometry marking and examined their differences. Our result showed that using kinesthetic feedback which allowed the user to realistically feel the model was better than using vibrotactile feedback that provided a simple tactile signal as confirmation of surface contact, in terms of marking accuracy. This finding directly demonstrated the significance of kinesthetic feedback for accurate medical marking while using hand gestures to touch the model to determine the marking position.

User experience

Based on the subjective data collected in the study (see Fig. 7), there were no statistically significant differences between the kinesthetic and vibrotactile VR interfaces, in terms of mental effort, naturalness and immersiveness. Both VR interfaces employed a VR headset, enabled changes in the head movement-based viewing perspective and allowed hand gestures to manipulate the models. It is not surprising that the participants rated both interfaces in a similar way. However, the participants reported more tiredness of the hand while using the vibrotactile VR interface compared with the kinesthetic VR interface.

The vibrotactile VR interface required the user to perform mid-air gestures, which is known to cause tiredness and heaviness of the hand especially when the interaction lasts for a long duration of time, a condition often referred as the gorilla arm effect (Hincapié-Ramos et al., 2014). The kinesthetic VR interface caused less fatigue, because the participants could rest their hand on the table while manipulating the models and on the models during the touching process.

6.2 Differences between the two VR interfaces and the traditional 2D interface

The traditional 2D interface that uses a 2D display and a mouse is still the most popular technique used to interact with volumetric images in the field of medicine. For the two haptic VR interfaces to be practically useful, they must perform well in comparison with the conventional 2D interface. The traditional 2D interface provided a fixed screen-based view, used a mouse-based rotate-pan-zoom technique to manipulate 3D models and used the ray-casting technique to determine the marking locations. In contrast, the two haptic VR interfaces employed a VR display for the visual output, which changed the scene perspective based on the user's head movement. Further, they relied on 3D hand gestures to manipulate the 3D models and determined the marking positions by reaching and touching the model.

Task completion time

Overall, there were no statistically significant differences between using the traditional 2D interface and using the two haptic VR interfaces, in terms of task completion time (see Fig. 3). However, a closer analysis based on the difficulty of the task revealed differences among the three interfaces. For the tasks with the low marking difficulty, the traditional 2D interface was faster than the kinesthetic VR interface. For the tasks with the high marking difficulty, the vibrotactile VR interface performed better than the traditional 2D interface and there was no difference between the kinesthetic VR interface and the traditional 2D interface (see Fig. 4).

The haptic VR interfaces and the traditional 2D interface have their own strengths and weaknesses for medical marking. On one hand, the marking method based on the ray-casting technique employed in the traditional 2D interface might lead to an improved interaction speed, due to the fast pointing capability of the mouse (along the x- and y- axes) on the 2D screen (Kim and Choi, 2019). The two VR interfaces required the user to reach and touch the model to determine the marking position. It is known that reaching an object in the 3D environment (along the x-, y- and z-axes) is difficult, and the speed of interaction is significantly influenced by the size and position of the target (Berthier et al., 1996). This especially provided an advantage for the traditional 2D interface when the marking targets were clearly presented on the surface of the 3D model without the need for any complex manipulation of the model.

On the other hand, the two haptic VR interfaces adopted hand-based 3D gestures which could implement rotating, panning and zooming simultaneously for manipulating the models, whereas the traditional mouse-based rotate-pan-zoom technique required the user to separately use three buttons for each manipulation function (i.e., the left mouse button for rotating, the

right mouse button for panning and the mouse wheel for zooming). These functions could not be performed simultaneously, which might have slowed down the interaction while using the traditional 2D interface. Further, the viewing perspective change following the user's head movement, provided by the VR headset, was more flexible than the fixed viewing angle presented by the standard 2D display. The participants could rotate the 3D model while also subtly changing the viewing angle with their head movement. This likely facilitated the visual search of marking positions. Therefore, it is likely that these benefits provided a distinct advantage to the VR interfaces, especially in the tasks involving the high marking difficulty.

Previous studies have explored to use the tangible interface (Besançon et al., 2017) and the touchscreen-based interface (Yu et al., 2010) for 3D manipulation tasks instead of using a mouse-based interface. We contributed to this research area by proposing two haptic VR interfaces and examining them with the traditional 2D interface. Through comparing with the traditional 2D interface in the medical marking task, the competitive performance of the two haptic VR interfaces in task completion time, especially the better performance using the vibrotactile VR interface in the difficult tasks, demonstrated their practical usability for 3D manipulation tasks.

Marking accuracy

Unlike the two haptic VR interfaces which required the user to reach the model to mark the positions, the traditional 2D interface used ray-casting to estimate the point of marking (Gallo et al., 2010). During the marking process, the participants commonly made displacement errors between the intended marking position and the actual marking position on the 2D screen while using the traditional 2D interface. When the marking difficulty was low, the surfaces of the marking positions were already parallel to the 2D screen or could be easily rotated parallel to the screen plane, and thus, a small displacement error on the 2D screen resulted in the same amount of displacement error of the user's marker on the 3D model. When the marking difficulty was high, the marking locations were easily visually blocked by other parts of the model and visible only from certain visual angles. The surfaces of the marking positions were difficult to rotate parallel to the screen plane while keeping the marking locations visible to the user. Thus, due to the angle that resulted between the marking surface and the screen plane, a small displacement error of the mouse pointer on the 2D screen would lead to a large displacement error of the user's marker on the 3D model, which negatively affected the marking accuracy. As expected, in comparison with the vibrotactile VR interface, using the traditional 2D interface resulted in an improved accuracy when the marking difficulty was low. In contrary, the traditional 2D interface had a lower marking accuracy when the marking difficulty was high (see Fig. 6).

Overall, the results suggested that the task performance of the traditional 2D interface was more affected by the marking positions in terms of marking accuracy. The results also indicated that, compared with the traditional 2D interface, using natural hand 3D gestures with haptic feedback to touch the model for medical marking could achieve more stable task performance regarding to the marking positions. Thus, it supported the potential of the two haptic VR interfaces, specifically the kinesthetic VR interface, for accurate medical marking.

User experience

Based on the subjective data (see Fig. 7), the two haptic VR interfaces were perceived to be similarly cognitively demanding and more natural and immersive compared with the traditional 2D interface. The VR interfaces adopted hand gestures to manipulate the 3D objects with haptic feedback and presented head movement-based viewing angles. These interaction experiences are similar to everyday interactions in the physical world, which might have made the participants feel easy to use and contributed to the improved perceived naturalness and immersiveness.

However, an advantage of the traditional 2D interface is that it caused the least hand fatigue. The mouse is a very familiar device to computer users, and users can easily rest their hands on the table while using the mouse. Using the vibrotactile VR interface led to the highest level of perceived hand fatigue, followed by the kinesthetic VR interface. In practical applications, extra arm-rest equipment may be useful to relieve the fatigue induced by the prolonged operation of the VR controller.

6.3 Limitations and future studies

This study has a few limitations. First, the experiment of the study involved medical marking on 3D models. For other scenarios in medical diagnosis and planning, such as highlighting a large area or manipulating multiple models, haptic VR interfaces may have different task performances. Further studies can employ the VR interfaces in different medical diagnosis and planning scenarios and explore their usability.

Second, we did not test the interfaces with medical professionals. It is possible that the results would have been different with a participant group who have previous experience in medical marking task using current medical marking systems with a mouse and a 2D display. They may have developed different strategies for interacting with complex 3D models following their years of usage experience. Thus, it is likely that the conventional 2D interface may have performed better if we had evaluated the system with medical professionals. However, for the sake of experimental design, it was a conscious choice to select participants for whom the experimental task and the interaction methods were novel. Future work can evaluate the effect of medical background on the results.

Third, for each medical marking trial, our participants performed a series of sub-tasks. For example, while searching and marking the target positions, the participants performed rotating, panning and zooming operations to manipulate the model. It is possible that the different interfaces we studied performed differently for each of the sub-tasks involved in terms of interaction speed. However, for the medical marking task, the participants often rotated, panned and zoomed the model simultaneously while using the controller and the force-feedback device. Future studies can investigate the performances of the VR interfaces in the sub-task level.

Fourth, this study included a short-term evaluation and the experiment task was new to the participants. It is likely that the task performance and user experience may change after more practice while using the three user interfaces, especially the haptic VR interfaces. The users are familiar with the traditional 2D interface but have only used the haptic VR interfaces for a short time. More training may further improve user performance while using the two VR interfaces. Thus, a long-term study can be conducted in the future to examine the learning effects of the use of the VR interfaces and the 2D interface.

Fifth, the two input manipulating tools used in the experiment for the haptic VR interfaces supported only single-point haptic interactions. Our everyday haptic interactions in the physical world commonly are multiple-point interactions. Multiple-point interaction method may largely improve haptic VR interfaces for manipulating 3D objects. To explore this, sophisticated and reliable haptic devices which support multi-point tactile and kinesthetic interaction are needed. This will require development of haptic technologies before such an experiment is feasible.

At last, we compared two haptic VR interfaces with a baseline of the traditional 2D interface. Such a comparison was necessary since, for any new interface to be practically useful and potentially be adopted by the medical practitioners, it is important for it to exceed the performance of the currently established methods. In addition, this comparative study contributes to the understanding of the strengths and weaknesses of the two VR interfaces. The VR headset with the head movement-based viewing perspective has contributed to the experimental task but was not fully explored in this study. To further explore the potential of the VR headset, a different comparative study is needed, which requires to have the type of displays as an independent variable and employ the same manipulation tool for the experimental interfaces. The mouse may not be an appropriate device for the VR environment and the VR controller is also specially designed for VR and not for a 2D display. One possible experimental setup to fully explore the effects of the VR headset in medical marking would be to use the force-feedback device as the manipulation tool with a 2D display and a VR headset respectively. This comparative study is different from our study but can be a follow-up study to continue from the basis of the present work. We hope our promising results will encourage more future studies in this research area.

7. Conclusion

Three-dimensional visualization has been widely used in computer-aided medical services such as highly accurate diagnosis and planning. The user interfaces for interacting with the 3D models are still largely based on the traditional 2D interaction method with a mouse and a 2D display. In this study, we proposed haptic VR interfaces to manipulate 3D models for the medical diagnosis and planning tasks and implemented a prototype system including two types of haptic VR interfaces (kinesthetic and vibrotactile VR interfaces). We conducted an experiment to compare the two VR interfaces with the traditional 2D interface for medical marking. The haptic VR interfaces showed promise in terms of interaction speed and accuracy. When the tasks involved complex 3D manipulation, using the vibrotactile VR interface led to the shortest task completion time and using the kinesthetic VR interface resulted with the best marking accuracy. The results demonstrated the potential of haptic VR interfaces to interact with volumetric medical images for medical diagnosis and planning. Our future work will investigate their usability in other medical diagnosis and planning scenarios and address how these haptic VR interfaces can be integrated as a part of medical workflow.

References

3D systems, 2020. Geomagic haptic device. https://www.3dsystems.com/ (accessed 15 September 2020)

Adesante, 2020. VR solution for Surgeons. https://www.surgeryvision.com/ (accessed 15 September 2020)

Alaraj, A., Luciano, C.J., Bailey, D.P., Elsenousi, A., Roitberg, B.Z., Bernardo, A., Banerjee, P.P., Charbel, F.T., 2015. Virtual Reality Cerebral Aneurysm Clipping Simulation with Real-Time Haptic Feedback. Operative Neurosurgery. 11(1), 52-58. https://doi.org/10.1227/NEU.0000000000000583

Al-Khalifah, A., McCrindle, R., Sharkey, P.M., Alexandrov, V.N., 2006. Using virtual reality for medical diagnosis, training and education. Int. J. Dis. Human Dev. 5 (2). 187-193. https://doi.org/10.1515/IJDHD.2006.5.2.187

Bade, R., Ritter, F., Preim, B., 2005. Usability Comparison of Mouse-Based Interaction Techniques for Predictable 3d Rotation. In: International Symposium on Smart Graphics. SG 2005, pp.138-150. https://doi.org/10.1007/11536482_12

Berthier, N.E., Clifton, R.K., Gullapalli, V., McCall, D.D., Robin, D.J., 1996. Visual information and object size in the control of Reaching. J. Mot. Behav. 28 (3), 187-197. https://doi.org/10.1080/00222895.1996.9941744

Besançon, L., Issartel, P., Ammi, M., Isenberg, T., 2017. Mouse, Tactile, and Tangible Input for 3D Manipulation. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI '17, pp.4727-4740. https://doi.org/10.1145/3025453.3025863

Bholsithi, W., Tharanon, W., Chintakanon, K., Komolpis, R., Sinthanayothin, C., 2009. 3D vs. 2D cephalometric analysis comparisons with repeated measurements from 20 Thai males and 20 Thai females. Biomed. Imaging Interv. J. 5(4), e21. https://doi.org/10.2349/biij.5.4.e21

Bielser, D., Gross, M.H., 2000. Interactive simulation of surgical cuts. In: Proceedings the Eighth Pacific Conference on Computer Graphics and Applications. https://doi.org/10.1109/PCCGA.2000.883933

Bouchlaghem, D., Shang, H., Whyte, J., Ganah, A., 2005. Visualisation in architecture, engineering and construction (AEC). J. Automation in Construction. 14 (3), 287-295. https://doi.org/10.1016/j.autcon.2004.08.012

Bowman, D.A., Kruijff, E., LaViola, Jr. J.J., Poupyrev, I., 2004. 3D User interfaces: theory and practice, 1st Ed. Addison-Wesley Professional Boston.

Chen, M., Mountford, S.J., Sellen, A., 1988. A study in interactive 3-D rotation using 2-D control devices. ACM SIGGRAPH Computer Graphics. 22 (4), pp. 121-129. https://doi.org/10.1145/54852.378497

El Saddik, A., Orozco, M., Eid, M., Cha, J., 2011. Haptics Technologies: Bringing Touch to Multimedia, in Avizzano, C.A., Azorín, J.M. (Eds.), Springer Series on Touch and Haptic Systems. Springer-Verlag Berlin Heidelberg.

Fahmi, F., Nainggolan, F., Siregar, B., 2019. 3D anatomy learning system using Virtual Reality and VR Controller. Journal of Physics: Conference Series 1235 012043. https://doi.org/10.1088/1742-6596/1235/1/012043

Fischer, A., Vance, J.M., 2003. PHANToM haptic device implemented in a projection screen virtual environment. In: Proceedings of the Workshop on Virtual Environments. EGVE '03, pp. 225-229. https://doi.org/10.1145/769953.769979

Foley, J.D., Andries, D., Feiner S.K., Hughes, J.F., 1995. Computer Graphics: Principles and Practice in C, 2nd Ed. Addison-Wesley Professional. Boston.

Gallo, L., Minutolo, A., De Pietro, G., 2010. A user interface for VR-ready 3D medical imaging by off-the-shelf input devices. Comput. Biol. Med. 40 (3), 350–358. https://doi.org/10.1016/j.compbiomed.2010.01.006

Gates, D.H., Dingwell, J.B., 2011. The effects of muscle fatigue and movement height on movement stability and variability. Exp. Brain Res. 209 (4), 525-536. https://doi.org/10.1007/s00221-011-2580-8

Gross, M.H., 1998. Computer graphics in medicine: from visualization to surgery simulation. ACM SIGGRAPH Computer Graphics. 32 (1), 53-56. https://doi.org/10.1145/279389.279462

Harrell, W.E., 2007. Three-dimensional diagnosis & treatment planning: The use of 3D facial imaging & ConeBeam CT (CBCT) in orthodontics & dentistry. Australasian Dental Practice. July/August 2007, 102-113.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Advances Psychol. 52, 139-183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hincapié-Ramos, J.D., Guo, X., Moghadasian, P., Irani, P., 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14, pp. 1063-1072. https://doi.org/10.1145/2556288.2557130

Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., Kassell, N., 1997. Usability analysis of 3D rotation techniques. In: Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology. UIST '97, pp.1-10. https://doi.org/10.1145/263407.263408

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Statist. 6 (2), 65-70. https://www.jstor.org/stable/4615733

Imagine Media, 1996. The Next Generation 1996 Lexicon A to Z: TrackBall. Next Generation. 15, pp. 42. https://archive.org/details/nextgen-issue-015

Issa T., Isaias P., 2015. Usability and Human Computer Interaction (HCI). In: Sustainable Design. Springer, London. https://doi.org/10.1007/978-1-4471-6753-2_2

Jankowski, J., Hachet, M., 2013. A survey of interaction techniques for interactive 3d environments. In: Eurographics Conference, EG 2013 - STARs. http://doi.org/10.2312/conf/EG2013/stars/065-093

Jones, T., Moore, T., Choo, J., 2016. The Impact of Virtual Reality on Chronic Pain. PLoS ONE, 11 (12), e0167523. http://doi.org/10.1371/journal.pone.0167523

Katsumata, A., Fujishita, M., Maeda, M., Ariji, Y., Ariji, E., Langlais, R.P., 2005. 3D-CT evaluation of facial asymmetry. J. Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod. 99 (2), 212-220. http://doi.org/10.1016/j.tripleo.2004.06.072

Kim, H., Choi, Y., 2019. Performance Comparison of User Interface Devices for Controlling Mining Software in Virtual Reality Environments. Appl. Sci. 9 (13), 2584. https://doi.org/10.3390/app9132584

King, F., Jayender, J., Bhagavatula, S.K., Shyn, P.B., Pieper, S., Kapur, T., Lasso, A., Fichtinger G., 2016. An Immersive Virtual Reality Environment for Diagnostic Imaging. J. Medical Robotics Research. 1 (1), 1640003. https://doi.org/10.1142/S2424905X16400031

Kinnison, T., Forrest, N.D., Frean, S.P., Baillie, S., 2009. Teaching bovine abdominal anatomy: Use of a haptic simulator. J. Anat. Sci. Educ. 2 (6), 280-285. https://doi.org/10.1002/ase.109

Kula, K., Ghoneima, A., 2018. Cephalometry in Orthodontics: 2D and 3D, 1st ed. Quintessence Publishing. USA.

Li, Z., Akkil, D., Raisamo, R., 2020. The Impact of Control-Display Gain in Kinesthetic Search. In: Nisky I., Hartcher-O'Brien J., Wiertlewski M., Smeets J. (eds) Haptics: Science, Technology, Applications. EuroHaptics 2020. Lecture Notes in Computer Science, vol 12272. Springer, Cham. https://doi.org/10.1007/978-3-030-58147-3_18

Lindner, C., Wang, C.W., Huang, C.T., Li, C.H., Chang, S.W., Cootes, T.F., 2016. Fully Automatic System for Accurate Localisation and Analysis of Cephalometric Landmarks in Lateral Cephalograms. Scientific Reports. 6, 33581. https://doi.org/10.1038/srep33581

Maples-Keller, J.L., Yasinski, C., Manjin, N., Rothbaum, B.O., 2017. Virtual reality-enhanced extinction of phobias and post-traumatic stress. Neurotherapeutics. 14 (3), 554-563. https://doi.org/10.1007/s13311-017-0534-y

Massie, T.H., Salisbury, J.K., 1994. The PHANToM haptic interface: a device for probing virtual objects. In: Proceedings of the ASME International Mechanical Engineering Congress and Exhibition, Dynamic Systems and Control, 55 (1), 295-301.

Medellín-Castillo, H.I., Govea-Valladares, E.H., Pérez-Guerrero, C.N., Gil-Valladares, J., Lim, T., Ritchie, J.M., 2016. The evaluation of a novel haptic-enabled virtual reality approach for computer-aided cephalometry. Comput. Methods Programs Biomed. 130, 46-53. https://doi.org/10.1016/j.cmpb.2016.03.014

Novint, 2020. Novint Falcon haptic device. https://www.vrs.org.uk/virtual-reality-gear/haptic/novint-falcon.html (accessed 15 September 2020)

Oculus, 2020. Oculus Rift VR. https://www.oculus.com/ (accessed 15 September 2020)

Okamura, A.M., 2004. Methods for haptic feedback in teleoperated robot-assisted surgery. Int. J. Industrial Robot. 31 (6), 499-508. https://doi.org/10.1108/01439910410566362

Olszewski, R., Zech, F., Cosnard, G., Nicolas, V., Macq, B., Reychler, H., 2007. Three-dimensional computed tomography cephalometric craniofacial analysis, experimental validation in vitro. Int. J. Oral Maxillofac. Surg. 36 (9), 828-833. https://doi.org/10.1016/j.ijom.2007.05.022

Peddamatham, S., Peine, W., Tan, H.Z., 2008. Assessment of Vibrotactile Feedback in a Needle-Insertion Task using a Surgical Robot. In: 2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. https://doi.org/10.1109/haptics.2008.4479920

Poupyrev, I., Ichikawa, T., Weghorst, S., Billinghurst, M., 1998. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. Computer Graphics Forum. 17 (3), 41-52. https://doi.org/10.1111/1467-8659.00252

Preim, B., Bartz, D., 2007. Visualization in Medicine: Theory, Algorithms, and Applications, 1st ed. Morgan Kaufmann. Burlington, USA.

Randall, D., Metherall, P., Bardhan, K.D., Spencer, P., Gillott, R., Noronha, R., Fenner, J.W., 2016. The Oculus Rift virtual colonoscopy: introducing a new technology and initial impressions. J. Biomedical Graphics and Computing. 6 (1), pp.34-42. https://doi.org/10.5430/jbgc.v6n1p34

Reinschluessel, A.V., Muender, T., Uslar, V.N., Weyhe, D., Schenk, A., Malaka, R., 2019. Tangible Organs: Introducing 3D Printed Organ Models with VR to Interact with Medical 3D Models. In: Proceedings of CHI Conference on Human Factors in Computing Systems Extended Abstracts. CHI'19 Extended Abstracts. https://doi.org/10.1145/3290607.3313029

Ribeiro, M.L., Lederman, H.M., Elias, S., Nunes, F.L.S., 2016. Techniques and Devices Used in Palpation Simulation with Haptic Feedback. J. ACM Computing Surveys. 49 (3), Article No. 48. https://doi.org/10.1145/2962723

Saad, E., Funnell, W.R.J., Kry, P.G., Ventura, N.M., 2018. A Virtual-Reality System for Interacting with Three-Dimensional Models Using a Haptic Device and a Head-Mounted Display. In: 2018 IEEE Life Sciences Conference. pp.191-194. https://doi.org/10.1109/LSC.2018.8572120

Satava, R.M., Robb, R.A., 1997. Virtual Endoscopy: Application of 3D Visualization to Medical Diagnosis. PRESENCE: Teleoperators and Virtual Environments. 6 (2), 179-197. https://doi.org/10.1162/pres.1997.6.2.179

Schoonmaker, R.E., Cao, C.GL., 2006. Vibrotactile force feedback system for minimally invasive surgical procedures, IEEE International Conference on Systems, Man and Cybernetics, vol. 3, pp. 2464-2469. https://doi.org/10.1109/ICSMC.2006.385233

Shen, R., Boulanger, P., Noga, M., 2008. MedVis: A Real-Time Immersive Visualization Environment for the Exploration of Medical Volumetric Data. In: 2008 Fifth International Conference BioMedical Visualization: Information Visualization in Medical and Biomedical Informatics. https://doi.org/10.1109/MediVis.2008.10

Shing, C.Y., Fung, C.P., Chuang, T.Y., Penn, I.W., Doong, J.L., 2003. The study of auditory and haptic signals in a virtual reality-based hand rehabilitation system. J. Robotica. 21 (2), 211-218. https://doi.org/10.1017/S0263574702004708

Shoemake, K., 1992. ARCBALL: A user interface for specifying three-dimensional orientation using a mouse. In: Proceedings of Graphics Interface '92. pp. 151-156. https://doi.org/10.20380/GI1992.18

Silén, C., Wirell, S., Kvist, J., Nylander, E., Smedby, O., 2008. Advanced 3D visualization in student-centred medical education. J. Med. Teach. 30 (5), e115-124. https://doi.org/10.1080/01421590801932228

Sousa, M., Mendes D., Paulo S.F., Matela N., Jorge J.A.P., Lopes D.S., 2017. VRRRRoom: Virtual Reality for Radiologists in the Reading Room. In: 35th Annual ACM Conference on Human Factors in Computing Systems. CHI '17, pp. 4057-4062. https://doi.org/10.1145/3025453.3025566

Steam, 2020. SteamVR. https://store.steampowered.com/app/250820/SteamVR/ (accessed 15 September 2020)

Steinberg, A.D., Bashook, P.G., Drummund, J., Ashrafi, S., Zefran, M., 2007. Assessment of faculty perception of content validity of PerioSim, a haptic–3D virtual reality dental training simulator. J. Dent. Educ. 71 (12), 1574-1582. https://doi.org/10.1002/j.0022-0337.2007.71.12.tb04434.x

Surgical Theater, 2020. VR surgical planning and navigation. https://www.surgicaltheater.net/ (accessed 15 September 2020)

Sutton, D., 1993. Textbook of Radiology and Imaging, 5th ed. Churchill Livingstone. London.

Swennen, G.R.J., Schutyser, F., 2007. Three-dimensional virtual approach to diagnosis and treatment planning of maxillo-facial deformity, in: Bell, W.H., Guerrero, C.A. (Eds), Distraction Osteogenesis of the Facial Skeleton, 1st ed. BC Decker Inc. Hamilton. pp.24

Tietjen, C., Meyer, B., Schlechtweg, S., Preim, B., Hertel, I., Strauß. G., 2006. Enhancing Slice-based Visualizations of Medical Volume Data. In: Conference EuroVis06: Joint Eurographics - IEEE VGTC Symposium on Visualization. pp. 123-130. https://doi.org/10.2312/VisSym/EuroVis06/123-130

Troulis, M.J., Everett, P., Seldin, E.B., Kikinis, R., Kaban, L.B., 2002. Development of a three-dimensional treatment planning system based on computed tomographic data. Int. J. Oral. Maxillofac. Surg. 31 (4), 349-357. https://doi.org/10.1054/ijom.2002.0278

Unity, 2020. Unity software. https://www.unity.com/ (accessed 15 September 2020)

Unity haptic plugin, 2020. Unity 5 Haptic Plugin for Geomagic OpenHaptics 3.3 (HLAPI/HDAPI). https://assetstore.unity.com/packages/essentials/tutorial-projects/unity-5-haptic-plugin-for-geomagic-openhaptics-3-3-hlapi-hdapi-34393 (accessed 15 September 2020)

Van Vlijmen, O.J., Maal, T., Bergé, S.J., Bronkhorst, E.M., Katsaros, C., Kuijpers-Jagtman, A.M., 2010. A comparison between 2D and 3D cephalometry on CBCT scans of human skulls. Int. J. Oral Maxillofac. Surg. 39 (2), 156-160. https://doi.org/10.1016/j.ijom.2009.11.017

Venson, J.E., Albiero Berni J.C., Edmilson da Silva Maia C., Marques da Silva A.M., Cordeiro d'Ornellas M., Maciel, A., 2017. A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality. Studies in Health Technology and Informatics, In: MEDINFO 2017: Precision Healthcare through Informatics, 245, 244-248. https://doi.org/10.3233/978-1-61499-830-3-244

VIVE, 2020. HTC VIVE VR. https://www.vive.com/ (accessed 15 September 2020)

Webster, R.W., Sassani, J., Shenk, R., Good, N., 2004. A haptic surgical simulator for the continuous curvilinear capsulorhexis procedure during cataract surgery. Studies in Health Technology and Informatics. In: Medicine Meets Virtual Reality 12, 98, 404-406. https://doi.org/10.3233/978-1-60750-942-4-404

Wirth, M., Gradl, S., Sembdner, J., Kuhrt, S., Eskofier, B.M., 2018. Evaluation of Interaction Techniques for a Virtual Reality Reading Room in Diagnostic Radiology. In: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. UIST '18, pp. 867-876. https://doi.org/10.1145/3242587.3242636

Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J., 2011. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '11, pp. 143-146. https://doi.org/10.1145/1978942.1978963

Yu, L., Svetachov, P., Isenberg, P., Everts M.H., Isenberg, T., 2010. FI3D: Direct-Touch Interaction for the Exploration of 3D Scientific Visualization Spaces. IEEE Transactions on Visualization and Computer Graphics. 16 (6), 1613-1622. https://doi.org/10.1109/TVCG.2010.157

Ángyán, L., Téczely, T., Ángyán, Z., 2007. Factors affecting postural stability of healthy young adults. Acta. Physiol. Hung. 94 (4), 289–299. https://doi.org/10.1556/APhysiol.94.2007.4.1
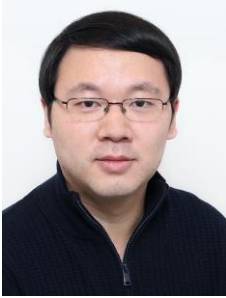
Zhenxing Li received his MSc degree in Radio Frequency Electronics from Tampere University of Technology and MSc degree in User Interface Software Development from University of Tampere, in 2009 and 2013 respectively. He is a researcher in Tampere Unit for Computer-Human Interaction (TAUCHI) Research Center at Tampere University and working toward PhD degree in Interactive Technology. His research interests include haptics, VR, gaze, multimodal interaction, mathematical modelling and algorithms.

Maria Kiiveri received her MSc degree in Computer Science from University of Tampere in 2019. She has worked in Tampere Unit for Computer-Human Interaction (TAUCHI) Research Center in Tampere University in a project that develops innovative technical solutions for medical and surgical use. She is currently employed in a company that provides software solutions for clinical laboratories.

Jussi Rantala received his MSc in Computer Science (2007) and PhD in Interactive Technology (2014) from the University of Tampere. He is a Senior Research Fellow at the Tampere Unit of Computer-Human Interaction (TAUCHI) at Tampere University. His research focuses on human-computer interaction, with special interest in haptics, olfaction, multisensory interfaces and XR.

Roope Raisamo is a professor of computer science in Tampere University, Faculty of Information Technology and Communication Sciences. He received his PhD degree in computer science from the University of Tampere in 1999. He is the head of TAUCHI Research Center (Tampere Unit for Computer-Human Interaction), leading Multimodal Interaction Research Group. His 25-year research experience in the field of human-technology interaction is focused on multimodal interaction, XR, haptics, gaze, gestures, interaction techniques, and software architectures.