A hybrid 2D/3D user interface for radiological diagnosis

Veera Bhadra Harish Mandalika[1,4,5], Alexander I. Chernoglazov[1,4], Mark Billinghurst[2], Christoph Bartneck[1,5], Mike Hurrell[3], Niels de Ruiter[1,3,4,5], Anthony P. H. Butler[1,3,4], and Philip H. Butler[1,4]

[1] *University of Canterbury, Christchurch, New Zealand*
[2] *University of South Australia, Adelaide, Australia*
[3] *University of Otago, Christchurch, New Zealand*
[4] *MARS Bioimaging Ltd., Christchurch, New Zealand.*
[5] *HIT Lab NZ, Christchurch, New Zealand*

Abstract This paper presents a novel 2D/3D desktop virtual reality hybrid user interface for radiology that focuses on improving 3D manipulation required in some diagnostic tasks. An evaluation of our system revealed that our hybrid interface is more efficient for novice users, and more accurate for both novice and experienced users when compared to traditional 2D only interfaces. This is a significant finding because it indicates, as the techniques mature, that hybrid interfaces can provide significant benefit to image evaluation. Our hybrid system combines a zSpace stereoscopic display with 2D displays, and mouse and keyboard input. It allows the use of 2D and 3D components interchangeably, or simultaneously. The system was evaluated against a 2D only interface with a user study that involved performing a scoliosis diagnosis task. There were two user groups: medical students, and radiology residents. We found improvements in completion time for medical students, and in accuracy for both groups. In particular, the accuracy of medical students improved to match that of the residents.

Keywords 3D input · hybrid user interface · diagnostic radiology · medical visualization · user interface

1 Introduction

Imaging modalities such as X-ray computed tomography (CT), magnetic resonance imaging (MRI), ultrasound and nuclear imaging are the basis for diagnostic radiology. The quality of diagnosis depends on the radiologist's ability to identify features or anomalies within

Veera Bhadra Harish Mandalika
HIT Lab, University of Canterbury, New Zealand
E-mail: harish.mandalika@pg.canterbury.ac.nz

the data, and accurately measure and report the findings. Some diagnosis tasks involve the rotation of one or more of the three anatomical planes. These tasks are often difficult using conventional radiology software as they involve using a 2D input device, such as a mouse, to manipulate a 2D plane in 3D. Mouse input is precise for 2D manipulation tasks; however, previous research shows that using the mouse for 3D manipulation can be difficult [1–3]. A study also showed that radiologists were faster with a 6 degrees of freedom (DOF) device compared to the mouse, in reformatting a slice in 3D data [4]. It is challenging to design a system that improves 3D manipulation while maintaining the benefits of mouse interaction.

In this paper, we introduce a hybrid system combining 2D and 3D interface components for diagnostic radiology. The 3D component consists of 3D stylus input and a stereoscopic 3D display with head tracking, while the 2D component consists of mouse input and a 2D display (see Fig. 1). Our system displays anatomical objects in 3D along with a 2D slicing plane, on the zSpace stereoscopic display [5]. The anatomical object, as well as the 2D plane, can be manipulated in 3D using 3D stylus input. The plane is also displayed synchronously on the 2D display. This enables experienced users to utilize familiar 2D views to see slice information while using 3D stylus input for interaction. It also allows less experienced users to utilize the 3D display, to gain an overview of the anatomy, which in turn helps them identify interesting slices and explore the information within.

We also present an evaluation of our system for a scoliosis radiology diagnosis task comparing our hybrid interface to a standard 2D interface, and a 3D only interface. We chose two groups: radiology residents as our more experienced group, and medical students as our

Fig. 1 An image of the hybrid interface showing the 2D display at the top, followed by the zSpace display at the bottom. The image also shows mouse, keyboard and the 3D stylus input.

less experienced group. The results demonstrate that our hybrid system improves the performance of diagnosis by providing improved 3D manipulation compared to the 2D interface.

The next section discusses related work in 3D user interfaces, virtual reality (VR), 3D manipulation, stereoscopic 3D, 3D input and hybrid interfaces in the medical domain. This is followed by a formal description of our prototype hybrid system, how the system was evaluated, the results of the user study, a discussion of results, and a plan for future work.

2 Related Work

Researchers have long investigated virtual reality (VR) for viewing and interacting with medical volumetric datasets. An overview of VR in the medical domain can be found in [6], [7] and [8]. More recently, a VR user interface for viewing 3D medical images was proposed by Gallo et al.[9]. One of the key elements of using VR for

medical visualization is being able to support intuitive 3D interaction. This is a challenging topic in VR, with a lot of research conducted in 3D input techniques. An overview can be found in Hand's survey of 3D interaction techniques[10].

Performing 3D manipulation using a 2D mouse is a common task in fields such as computer-aided design (CAD), 3D modeling, surgery simulations, and desktop VR. Particularly, 3D rotation is commonly used for exploring medical data. The most commonly used 3D rotation techniques for mouse input include Shoemake's ARCBALL[11] and a virtual trackball surrounding the object[12]. A study comparing these techniques by Bade et al. [13] concluded that design principles were crucial for 3D rotation. The study also highlighted the need for intuitive interaction techniques.

As an alternative to using the mouse for 3D manipulation, 3D input devices were developed to offer more DOF for 3D manipulation. Some early 3D input devices, developed particularly for exploring volumetric medical data, include Hinckley's prop system [14] and the Cubic mouse [15]. More recently the Wiimote [16], and the Leap Motion controller[17] have been used to capture the natural gesture input. Furthermore, researchers focused on using contact free gesture input. For example, a controller-free tool for exploring medical images [18], and a touchless interface for image visualization in urological surgery [19].

Previous research suggests that 3D input devices outperform the mouse for the 3D manipulation tasks. A 1997 study by Hinkley et al., compared usability of 3D rotation techniques and found that 3D orientation input devices were faster than 2D input without sacrificing the accuracy[2]. Balakrishnan et al. modified a traditional mouse by adding two additional DOF forming the Rockin'Mouse, which was able to achieve 30% better performance for 3D interaction over mouse input [20]. A study from Dang et al. compared input devices for 3D exploration and found that a wand interface provided the best performance[21]. A study evaluating glove input in 2D and 3D for medical image analysis from Zudilova et al. showed that using more DOF improved performance and accuracy [22].

Contrary to the above studies, a study from Bérard et al.[23] found that 2D mouse (with 2 DOF) used with three orthographic views outperformed 3D input devices (with 3 DOF) used with a stereoscopic perspective view. A later modification to the task that involved adding pop-up depth cues within the stereoscopic perspective view, showed that the 3D input device outperformed the 2D mouse [24]. These studies led us to expect that the 3D stylus input with 6 DOF from zS-

pace should also provide similar benefits/improvements in 3D manipulation over 2D mice.

Most systems still strongly rely on some level of 2D user interaction. Hence, researchers started finding ways to combine 2D and 3D interaction to form hybrid user interfaces. Since VR hardware had limited resolution, early hybrid interfaces focused on providing means for completing high-resolution tasks in virtual environments[25]. The most common approach has been integrating handheld and touchscreen devices into virtual environments for 2D interaction [26–30]. Others have focused on providing hybrid input, such as the Virtual Tricoder [31], Pinch Glove based 2D/3D cursor approach [32], Pick and Drop approach [33] and other tangible user interfaces [34].

A summary of VR in health care can be found in [35], where they highlighted the need for more research focused on applications of VR in health care. A summary of VR in medicine can be found in [36], where they discuss the VR research in the medical domain, outlining major breakthroughs, and its impact in providing healthcare services. A more recent meta-analysis of VR research in medicine can be found in [37].

Some developments involved multi-display setups, for example, the hybrid interface for visual data exploration combining the 3D display (with glove input) and a tablet PC, which was proposed by Baumgärtner et al.[38]. Similarly, in the medical domain, another hybrid interface for manipulation of volumetric medical data for liver segmentation refinement was developed by Bornik et al[39]. This interface combined a tablet PC and an immersive VR environment to improve 3D manipulation. We extend their approach, our aim was to combine desktop VR with stylus input and traditional 2D displays to improve 3D manipulation for diagnostic radiology.

A 3D volume navigation tool for diagnostic radiology, that used a 3D mouse to provide both 2D as well as 3D input for exploring volumetric data, was developed by Teistler et al[40]. A tool for post-mortem CT visualization using a 3D gaming controller for interaction along with an active stereo screen was proposed by Teistler et al[41]. Another approach from Graves et al[4], used a fixed 6 DOF controller to reformat a slice in 3D data and found that radiologists were faster with their approach, compared to mouse input from the standard radiology workstation.

The literature shows that a number of researchers explored the benefits of hybrid interfaces for 3D object manipulation. However, as far as we are aware, there are very few medical interfaces that use hybrid input, and none that combine 2D and 3D screen viewing with stylus interaction in diagnostic radiology. Hence,

we present a hybrid user interface for diagnostic radiology and an evaluation against the existing 2D interface, comparing novice and experienced users.

3 Hybrid User Interface

This section describes the prototype hybrid interface we have designed. It starts by reviewing the hardware and software components, followed by an overview of interface design and walkthrough. Finally, we discuss various interaction tools.

3.1 Hardware Setup

The hardware setup consists of two parts: the zSpace desktop VR system, and a traditional 2D system. The zSpace consists of a 24 inch (1920x1080 pixels) passive stereo display with a refresh rate of 120 Hz. The zSpace screen also has embedded cameras for tracking. A pair of tracked circularly polarized glasses is used for viewing the stereo display. The zSpace also uses a 6 DOF 3D stylus, with a polling rate of 1000 Hz. The user's head and 3D stylus are tracked in 6 DOF by the cameras embedded in the display. The display is angled at 45° from the desk for convenient viewing (see Fig. 1).

The 2D component of our system consists of a 2D monitor, and keyboard and mouse input. The 2D monitor is placed perpendicular to the desk, above the zSpace display. The distance from the 2D screen to the user is approximately 30 cm. The keyboard is placed at the center below the zSpace display, with the mouse and 3D stylus, placed to the side. A single workstation PC (3.6GHz Intel Core i5-3470, NVidia Quadro K5200) is used to drive both systems.

3.2 Software Setup

The software system consists of a single application that drives the desktop VR system as well as the 2D system. The application is built on top of the MARS Vision framework[42–44]. MARS Vision contains DICOM dataset handling code based on DCMTK[45] for loading volumetric datasets, as well as the 2D desktop user interface, which is built using the Qt toolkit[46]. The desktop VR user interface uses a custom rendering engine built with OpenGL[47] and CUDA[48] APIs. The software was programmed in C++ with Visual Studio IDE for the Windows 7 operating system.

Once an anatomical dataset is loaded, it is visualized on both systems simultaneously. The three anatomical planes showing slice information are displayed on the

2D system along with other application controls, while an extracted 3D model and a 3D user interface are displayed on the VR system (see Fig. 1). Interaction with the data is possible with either system; however, the dataset loading and application controls are limited to the 2D desktop system. User interaction with the 2D system is performed using the 2D mouse while the VR system uses a 3D stylus. A keyboard is used for text input in both systems.

3.3 Interface Design
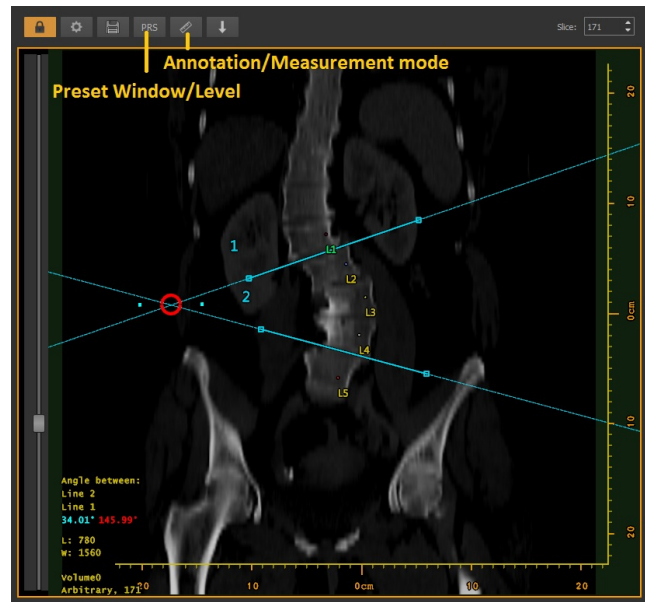
3.3.1 2D Interface

The 2D user interface was designed to closely resemble the standard radiology diagnosis tool, Inteleviewer PACS Viewer (Intelerad Medical Systems)[49]. This was done to minimize the differences for the users already familiar with traditional radiology software. The functionality, keyboard and mouse controls were also mapped in a similar fashion to that of Inteleviewer.
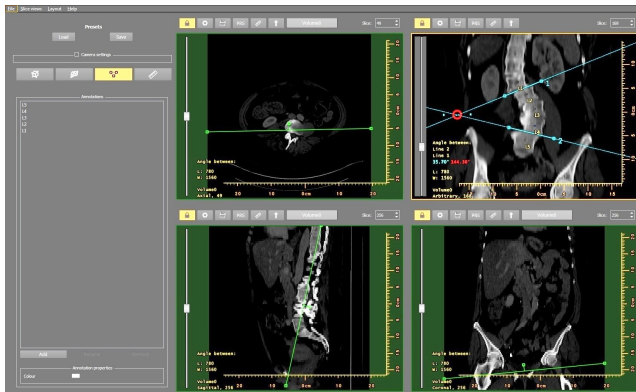
Fig. 2 2D interface showing the three anatomical planes, and the arbitrary plane (top right) in a 2x2 grid. The application controls tab lies to the left of the screen.

The 2D interface contains a fixed application controls tab on the left, followed by the three anatomical planes and an arbitrary plane in a two by two grid layout as shown in Fig. 2. The grid layout can be rearranged to the desired format. Each slice view also displays a scale, that is positioned to the right and bottom edges of the slice view. The user can interact with the 2D interface using mouse input alone, or in combination with a keyboard.

The mouse can be used to select and interact with the 2D slice views using the left mouse button. The user can hold the left mouse button and move the mouse to pan the slice, and use the scroll wheel for scrolling through slices. This can also be done by selecting the

Fig. 3 2D interface showing the slice view. The scroll bar to the left, slice control buttons including the PRS (preset window/level), ruler button to change interaction modes between annotation and measurement, slice spinner to the top right and the scale to the right and bottom edges showing length.

scroll bar at the left of each slice view with the left mouse button and dragging it or using the spinner at the top right corner of the slice view, see Fig. 3. The scroll wheel can be used while holding the 'Shift' key on the keyboard for changing the scale of the slice content.

The window and level can be changed by holding the middle or scroll mouse button and moving the mouse horizontally or vertically. Preset window and level settings can be directly selected using the 'PRS' button at the top of the slice view. The right mouse button is used to place annotations and measurements on the slice based on the interaction mode. Interaction modes can be changed using the application controls tab to the left (see Fig. 2), or the 'ruler' icon at the top of the slice view (see Fig. 3).

We also allow the possibility to remap functionality to any mouse buttons or the scroll wheel. This was done to allow experienced users, who had been using the mouse a certain way, the ability to remap the functionality to match their desired way. This would eliminate the need for experienced users to unlearn their preferred mouse interaction methods in favor of a fixed mapping in our 2D interface.

3.3.2 3D Interface

The 3D interface contains a 3D scene that shows a model, representing an iso-surface extracted from the volumetric CT data, with a strip of user interface but-

tons at the bottom as shown in Fig. 4. The scene is rendered in stereoscopic 3D relative to the user's tracked head position. The buttons are distinct 3D models created to represent various functions such as enabling the 2D plane, and changing between various interaction modes such as the 3D annotation mode and the 3D measurement mode (see Table 1). Buttons can be selected by touching them with the stylus. When selected, each button enlarges, its function is described underneath, and a translucent billboard with stylus controls is displayed for reference. While selected, a button can be triggered by pressing the front stylus button, see Fig. 5. The 3D stylus is the primary input for the 3D interface, while the keyboard is still used for text input.

Fig. 4 3D interface showing the zSpace display with stylus interaction. A model extracted form the CT dataset is shown along with the anatomical plane, and the 3D buttons at the bottom of the screen.
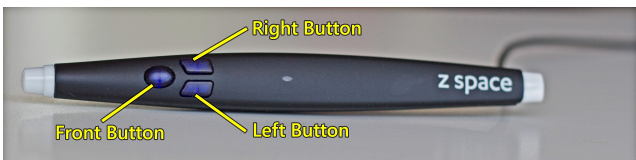
Fig. 5 zSpace 3D stylus showing its front, left and right buttons.

The stylus contains three buttons: front, left and right buttons, see Fig. 5. The front button is always used for picking and manipulating the model or the anatomical plane, as well as triggering all the 3D buttons. The functionality of the left and right buttons depends on the selected mode (see Table 1).

3.4 Interface Walkthrough

When a CT dataset is loaded, the 2D interface displays the three anatomical planes. The arbitrary plane is initially set to replicate the coronal plane.

Table 1 Functionality of stylus 'left'and 'right'buttons for various 3D interaction modes

Interaction Mode	Left Button	Right Button
Model Manipulation	Translate without rotating.	Increase or decrease the model scale.
Point Annotation	Delete the selected annotation.	Insert annotation at the end of the stylus.
Line Measurement	Delete the entire line connected by the selected point.	Insert an endpoint of the line to measure.
Angle Measurement	Clear selected lines.	Select lines to measure the angle between.

A 3D mesh is extracted in real-time from the CT dataset using the Marching Cubes algorithm[50]. The mesh extraction is performed on the GPU using CUDA. This is done each time it requires a change and not every frame.

However, the extraction time (15 milliseconds) is fast enough to be real-time. This mesh model is displayed on the 3D interface. Once the model is extracted, the rendering cost for one frame (for both eyes) is about 3 milliseconds. This results in an average frame rate of 330 frames per second (FPS) per eye.

The user can interact with the 2D interface by selecting and scrolling through any anatomical plane using the mouse. If the 3D plane is enabled in the 3D interface, its position and orientation on the object are synchronized to match the selected 2D plane. This concept of a synchronized slice being displayed in both 2D and 3D views is similar to the approach used by Teistler et al[40].

The part of the model in front of the plane is rendered translucent, where the level of transparency can be adjusted by the user. The model's front portion can also be made completely transparent, which is similar to the plane clipping technique used by Dai et al.[51]. The plane displays the CT densities from the CT dataset. This provides a rapid assessment of the arbitrary plane's position in the model. The user can synchronize the arbitrary plane orientation and position to any of the three orthogonal planes by selecting the orthogonal plane by left clicking on it while holding down the [control] key.

The user can rotate the arbitrary plane to a desired orientation using the mouse in the 2D interface. This can be done by holding down the [shift] key on the keyboard along with the left mouse button and moving the mouse. Horizontal movement rotates the plane around

the y-axis while vertical movement rotates the plane around the x-axis. The slice is always rotated about its own center point. This is similar to the interaction method used in Inteleviewer.

When the arbitrary slice is selected, its guidelines are displayed on the other orthogonal views. These guidelines can also be used to manipulate the position and orientation of the arbitrary slice. These actions are synchronized with the plane in the 3D interface.

Annotations are placed by right clicking on the slice, while the annotation mode is selected. Lines can be drawn by right clicking on the slice to start the line and dragging to the desired endpoint, while the line measurement mode is selected. The line can then be moved, or modified by moving its endpoints.

Angle measurement mode allows the selection of two lines to measure the angle. If only two lines are drawn, they are automatically selected and the angle between them is displayed. Annotations and measurements on the slice view, are also displayed in the 3D interface. Displaying 2D measurements in 3D is similar to the approach used by Preim et al[52].

In the 3D interface, the 3D model resides mostly in the negative parallax region. The stereoscopic screen being angled from the desk, combined with the head tracking, makes the model appear to float over the screen.

The 3D model can be manipulated by simply touching it with the stylus, and pressing the front stylus button to pick it up. Once picked up, the model is attached to the end of the stylus, and it follows the hand motion and orientation in a one to one fashion. The user can release the front button once the desired position and orientation is reached, see Fig. 6.

Similarly, the user is able to simply pick and move or rotate the arbitrary plane as desired, see Fig. 6. The plane maintains its relative position to the 3D model when the model is manipulated. Changes made to the plane orientation are always updated in the 2D interface. This can be seen in Fig. 1, where the 3D plane and the arbitrary plane (top right) are showing the same information.

The user can also perform other functions, such as annotations or measurements, in the 3D interface, as shown in Fig. 7. Annotations and measurements can be placed using the stylus, by first selecting the appropriate mode using the 3D mode buttons at the bottom of the 3D interface, see Fig. 7.

Annotations can be placed in by pressing the right stylus button. An annotation is placed at the end of the virtual stylus, represented by a sphere. Once placed, the keyboard is used to enter a name for it. The left stylus button can be used to delete an annotation.

Line measurements can be placed using the right stylus button to place the start and end points of the desired line, sequentially. The points are inserted at the end of the stylus each time the right stylus button is pressed. The endpoints are represented using spheres, with a 3D line between them. The line can be deleted, by touching one of its endpoints and pressing the left stylus button.

Angle can be measured using the right stylus button to select two lines for measuring the angle. The selected lines are displayed as white lines. The angle is displayed as white 3D text between the selected lines, once two lines have been selected. The line selection can be cleared using the left stylus button.

At any stage, the annotations or measurements can be picked and moved using the front stylus button. They maintain their relative position to the 3D model when the model is manipulated. They also scale with the model. All the stylus interaction modes and button mappings are shown in Table 1.

If the arbitrary plane is enabled while placing annotations or measurements, they will automatically snap to the slice plane, if they are within the distance of 1 cm from it. This makes it easy for the user to place measurements on a slice plane, using the 3D stylus input.
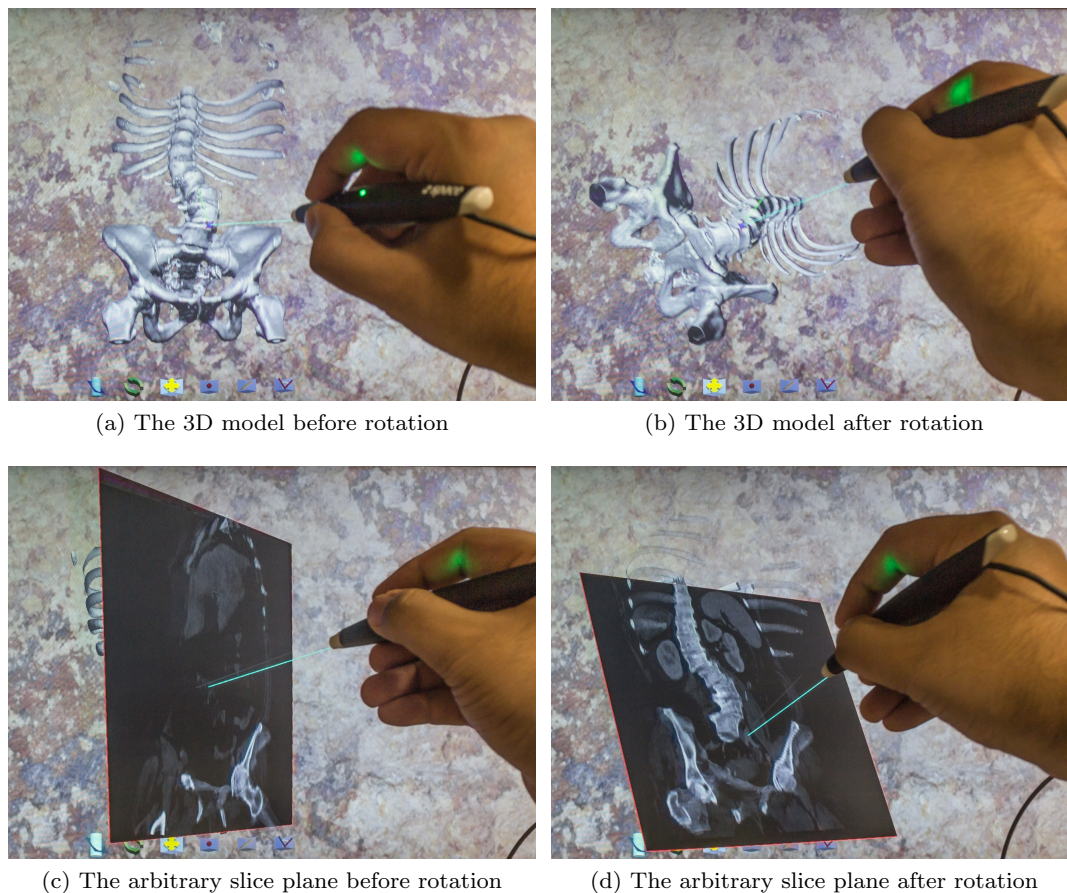
Annotations such as 3D points are shown in the 2D interface on the corresponding slice; however 3D measurements such as 3D lines are not shown in the 2D interface. The user can choose to interact with the 2D or 3D interface, or a combination of both in order to perform a diagnosis task as seen in Fig. 1.

3.5 Interaction Tools

Clinical information, such as the present symptoms and prior medical conditions, is used by the radiologist to rank the likelihood of various possible diagnoses when interpreting the scans.

Such information may also indicate the need to obtain diagnostic measurements, in which case the measurement task would typically comprise three steps: inspection, annotation (optional), and obtaining the measurement itself.

For our hybrid interface, the DICOM data loading and the 2D user interface, including the orthogonal views were part of the MARS Vision framework. We developed the arbitrary view, and the annotation and measurement tools, specifically for the purpose of testing our hybrid interface.

(a) The 3D model before rotation

(b) The 3D model after rotation

(c) The arbitrary slice plane before rotation

(d) The arbitrary slice plane after rotation

Fig. 6 An image of the 3D interface showing model and slice plane rotation using the 3D stylus input.

3.5.1 Inspection

The first step involves the visual inspection of slices from the CT dataset to find the region of interest (ROI) within a slice image. The user can use a set of tools that we refer to as the "2D slice manipulation tools" to perform the inspection. These tools include controls that not only allow the user to pan, scale or scroll through the slices, but also adjust the gray-scale window width and level to optimally display the slice. These tools also provide controls for rotating the slice planes. The inspection using our hybrid system additionally allows the user to use the 3D stylus to directly manipulate and re-orient a slice, to perform the inspection.

3.5.2 Annotation

Annotation tools are used to either annotate an ROI or an entire slice. Annotations contain text that provides additional information regarding the ROI or slice. Annotations can be used by radiologists to share findings. Annotations can also be used as a prerequisite for measurements. Typical annotation tools include point and line annotations as well as controls for annotating various ROIs within a slice or the entire slice. We only support point annotations in our system (see Fig. 8).
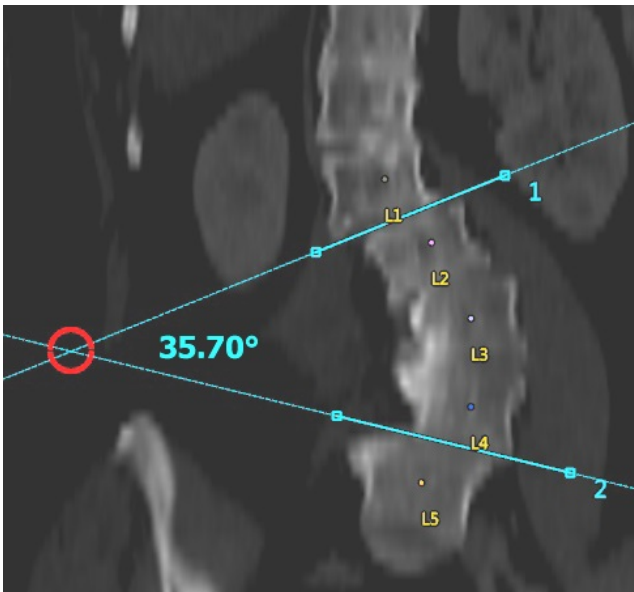
3.5.3 Measurement

Measurement tools are used to quantify various aspects of CT datasets. These measurements often include selecting a finite ROI on a slice using shapes such as rectangles, circles, or ellipses. The tool then provides min, max, and average values. The tool can also be used to measure the length, or area of certain parts of the slice. Measurements often form the basis for diagnosis reports. We have implemented most of the measurement tools for our 2D interface (see Fig. 8), along with line and angle measurement tools for our 3D interface (see Fig. 4).

3.6 Hybrid Interaction

Hybrid interaction can be performed in two ways, serial and parallel. In the serial approach, the tasks are
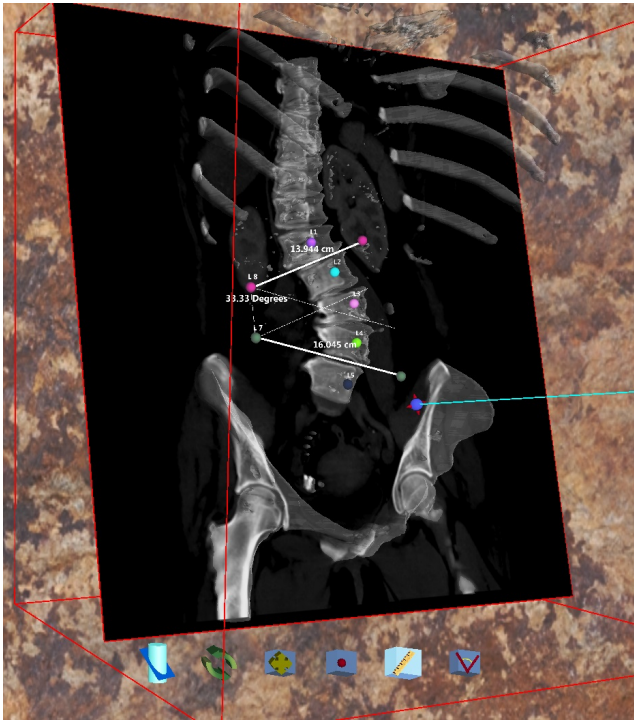
Fig. 7 An image of the 3D interface showing annotations and measurements tools used for a scoliosis angle measurement task.

Fig. 8 2D interface demonstrating the annotation and measurement tools used for a scoliosis angle measurement task on the arbitrary slice.

performed using the 2D and 3D components independently in a sequential manner. The diagnosis task can be divided into multiple steps, each of which can be performed using 2D or 3D components alone. Since both the components are synchronized with each other, the user can proceed with either interface for the next step.

In the parallel approach, both the 2D and 3D components can be used together at the same time to perform a single step in the diagnosis task. For example, a user can use the stylus to manipulate the arbitrary plane, while looking at the slice content on the 2D interface. The user can also choose to scroll through a slice using the mouse while looking at the 3D model to understand its relative position. Ambidextrous users can use the mouse and stylus together to manipulate the slice and the 3D model simultaneously.

Both approaches are supported in our system, however we only test the serial approach in our evaluation, as it requires considerably less training for users who have prior experience with the 2D component. Since the serial approach uses only mouse or stylus input at any given time, it can be easier and faster to learn the system.

We believe that the hybrid interaction method can have considerable advantages. The 2D component can be used for high precision tasks such as measuring line lengths, angles between lines, or marking 2D ROIs. Data inspection can also be performed in higher resolution since it is possible to view the slice information with a one to one mapping between the slice data and pixels. This is not ideal using the 3D component since slice information is textured on a 3D plane, which can appear distorted due to a non-optimum viewpoint when viewing it in the 3D view. Additionally, the circular polarization of the stereoscopic screen also reduces image brightness making it harder to view the slice information in the 3D view.

The 3D component can be used to gain a quick overview of the anatomy by manipulating the displayed anatomical object. It can also be used for tasks such as slice plane orientation, and 3D line and 3D ROI measurements. The speed/accuracy trade-off will depend heavily on the type of diagnosis task. The key factors would be the flow of actions in performing a particular diagnosis task, the level of synchronization and the user's ability to shift focus between the 2D and 3D components.

4 Evaluation

The goal of this evaluation was to determine the effectiveness of our hybrid interface for a specialized radiological diagnosis task, compared to a more traditional 2D interface. We wished to test the effect of providing an easier means for rotating the arbitrary plane, on task performance and accuracy, compared to using the 2D mouse. We achieved this by comparing the traditional 2D slice manipulation tools with our hybrid interaction

tools. We also chose a specialized task that required arbitrary slice rotation.

Since our hybrid interface is a combination of 2D and 3D components, we also included a 3D only interface to eliminate the possibility of the observed difference being due to the 3D component alone. We also implemented 3D measurement tools, specifically for the 3D only interface. So we evaluated three interface conditions: 2D, 3D, and hybrid.

4.1 Method

The key research question was to determine if our hybrid interface provides any improvements over the existing 2D interface in terms of task performance and accuracy. Furthermore, we wanted to study how the user's prior experience with the 2D interface influenced their task performance with our hybrid interface. Hence, the experiment was set up as a mixed *within/between* study in which the interface (2D, 3D, hybrid) was the within factor, and experience (student, resident) was the between factor. We obtained human ethics approval for our user study from University of Canterbury Human Ethics Committee (Ref# HEC 2016/35/LR-PS)[53].

4.1.1 Participants

For the evaluation, we chose two groups of participants: fourth-year medical students, and radiology residents (physicians training to become radiologists). The fourth-year medical students have the necessary medical background to perform the diagnosis tasks but have little to no experience using any diagnosis tools. Residents have at least one or two years experience using a 2D only interface; using it daily for various diagnosis tasks.

Since our hybrid interface uses a stereoscopic 3D display with 3D stylus interaction, we only chose participants with binocular or stereo vision, who had no hand related disabilities that would prevent them from using the stylus input. We had 31 participants for the evaluation. The student group consisted of 21 participants (11 male and 10 female) with an average age of 22, ranging from 21 to 27. The resident group consisted of 10 participants (6 male and 4 female) with an average age of 32, ranging from 28 to 38.

4.1.2 Task

Most diagnosis tasks are performed by exploring one or more of the three anatomical planes, namely sagittal, coronal, and transverse planes. The procedure for diagnostic radiology widely varies for each diagnosis task.

The conventional radiological workstation software allows exploration of 2D slices for diagnosis. Radiologists receive intense training on how to mentally visualize three-dimensional anatomical objects from these slices, and they require a lot of experience to do this task well[54].

Creating a 3D mental model helps them associate the location of each slice with its position inside the anatomical object. They find and examine the interesting slices, identify anomalies or features within these slices and report their findings. Although it is possible to render 3D anatomical objects from 2D slice data (for example volume rendering[55]), radiologists seldom rely on it. They find it easier and faster to diagnose using only the 2D slices.

We had two major criteria for selecting an evaluation task. Firstly, we wanted the task to have an element of 3D rotation for testing the effectiveness of our hybrid interface. Secondly, we wanted the task to be easy enough for medical students to perform, without much training.

After consultation with experienced radiologists, we chose to use a real diagnosis task, namely obtaining a scoliosis angle measurement task from abdominal CT scans. This task consisted of three simple steps: correcting the coronal slice orientation, annotating the spinal column, and measuring the scoliosis angle. Hence, the task could be performed by anyone with a basic knowledge of human anatomy.

We only chose abdominal CT scans where the patient's spinal column was not parallel to the coronal plane, as this was a very common scenario for such scans. Hence, the diagnosis task involved 3D manipulation of the coronal plane to correct its orientation. There were three similar, but different, datasets used for the experiment.

4.1.3 Process

Each participant first received a five-minute interface demonstration from an instructor followed by another five minute practice period. During the practice period, the participant was allowed to perform the diagnosis task with a practice dataset and was free to ask questions regarding the task or interface. Later, the participant performed the actual diagnosis task on a different data.

The scoliosis task consisted of three steps. The first step was is to adjust the orientation of the arbitrary 2D (coronal) slice to optimally display the spine. The second step was to examine the coronal slice and annotate the vertebrae. Finally, the user drew two lines between two sets of vertebral discs, and measured the scoliosis

angle between them. This procedure was repeated for each interface condition. The order of the conditions, and the use of datasets were both counterbalanced using a balanced latin square method [56].

The keyboard was used to label the annotations in all conditions. The 2D condition involved a 2D mouse, while the 3D condition used the 3D stylus. The hybrid condition was a mixture of both, utilizing the 3D stylus input for the first step, and mouse input for the rest of the task.

After each diagnosis task, participants were required to complete several questionnaires (see Section 4.1.4). At the end of the experiment, participants were asked to comment on their experience in a semi-structured interview.

To better control the environment of the task, the window/level settings to make the bones clearly visible within the slice were pre-set. This guaranteed the same initial conditions for the second step. In addition, we also chose task datasets that required the same set of vertebral discs to be measured for the scoliosis angle. This was communicated to the participants prior to starting each diagnosis task. Participants were also given reference cards showing a labelled diagram of a healthy spinal column, and illustrating the interface controls for all conditions.

4.1.4 Measures

Prior to the experiment, the scoliosis angle was measured by an expert radiologist using each of the three interface conditions. This was repeated three times with every interface, for each dataset. We observed only a very minor difference of about $0.5°$ over repeated angle measurements from the expert. An average of the measured angle was then recorded as the best solution for each dataset, to establish a baseline for comparison.

The effectiveness of our hybrid system can be quantified with task completion time and accuracy. For each task, a snapshot of the labeled vertebrae, the completion time, and the resulting scoliosis angle were recorded. The snapshot was used to verify the correct annotation of the spinal column.

The measured angle was later compared with the base line measurement provided by an expert, to establish an accuracy measure for each diagnosis task. The absolute difference in the angle was computed as the absolute error for comparison.

Additionally, since the 3D condition involved multiple mode changes for performing the diagnosis task, we also recorded the number of mode changes for each participant.

Another key measure is the usability of the system, which was measured by the System Usability Scale (SUS) questionnaire [57]. We also chose to measure physical and mental task load using the NASA TLX [58] questionnaires. We only used the physical and mental task load scales from the NASA TLX questionnaire for our study.

4.2 Results

The key question we wished to answer was whether our 2D+3D hybrid interface was more effective than the existing 2D interface, or 3D interface, in a radiological diagnosis task. To answer this question we conducted a two-way mixed ANOVA with two factors: interface condition (2D, 3D, hybrid) and experience (student, resident) on the resulting experimental data with completion time, absolute error, SUS score, Physical task load and Mental task load as the dependent variables. We will refer to our interface condition as an interface for the remainder of this section.

There were two significant outliers in the data, as assessed by inspection of a boxplot for values greater than 3 box-lengths from the edge of the box. The cause for these outliers was a hardware malfunction during the experiment. Therefore, both these outliers were removed from the analysis.

The data was normally distributed, as assessed by Shapiro-Wilks test of normality ($p > 0.05$). There was homogeneity of variance ($p > 0.05$) and covariances ($p > 0.05$), as assessed by Levenes test of homogeneity of variances and Boxs M test, respectively.

Mauchlys test of sphericity indicated that the assumption of sphericity was met for the two-way interaction for completion time ($\chi^2(2) = 1.238$, $p = 0.538$), absolute error ($\chi^2(2) = 1.324$, $p = 0.516$), physical task load ($\chi^2(2) = 0.262$, $p = 0.877$), and mental task load ($\chi^2(2) = 3.394$, $p = 0.183$), however it was violated for SUS score ($\chi^2(2) = 20.934$, $p < 0.0005$). The Greenhouse-Geisser correction was used for SUS score ($\epsilon < 0.75$)[59].

There was a statistically significant interaction between the interface and experience on completion time ($F(2, 54) = 37.835$, $p < 0.0005$, partial $\eta^2 = 0.584$), absolute error ($F(2, 54) = 4.416$, $p = 0.017$, partial $\eta^2 = 0.141$), SUS score ($F(1.288, 34.772) = 13.604$, $p < 0.0005$, partial $\eta^2 = 0.335$, $\epsilon = 0.644$), and physical task load ($F(2, 54) = 5.564$, $p = 0.006$, partial $\eta^2 = 0.171$), but the interaction was not statistically significant for mental task load ($F(2, 54) = 0.053$, $p = 0.948$, partial $\eta^2 = 0.002$).

For the mental task load, since the interaction was not statistically significant, we investigated main ef-

fects, but there was no statistically significant difference in mental task load between interfaces ($p = 0.089$) or groups ($p = 0.161$). For other variables, since statistically significant interaction implies that the impact of the interface is dependent on the group, we have not investigated main effects directly. Instead, we investigated the simple main effects of interface for each group individually and vice-versa, which are reported in two parts.

The first part consists of the effects of each *interface* on *experience groups* for the dependent variables. This is done by first splitting the data into two experience groups: students and residents, followed by running one-way repeated measures ANOVA on each group with the interface as the repeating factor. The mean (M), standard error (SE) and p-value are reported for each combination of interfaces.

The second part consists of effects of each *experience group* on the *interfaces* for the dependent variables. This is done by running independent samples t-test on the dependent variables for all interfaces with *experience* as the grouping variable with two defined groups: students and residents. The results are reported as mean ± standard error ($M \pm SE$), followed by the t-value and p-value for the mean difference. We also compute and report the effect size (d) for the t-test using Cohen's d [60].

4.2.1 Effects of interface and group on Completion Time

For the student group, the task completion time was statistically significantly higher using 2D compared to 3D ($M = 24.21$, $SE = 7.91\,seconds$, $p = 0.020$), 2D compared to hybrid ($M = 51.26$, $SE = 8.77\,seconds$, $p < 0.0005$), and 3D compared to hybrid ($M = 27.05$, $SE = 9.92\,seconds$, $p = 0.042$). For the residents, the task completion time was statistically significantly lower using 2D compared to 3D ($M = 113.3$, $SE = 16.27\,seconds$, $p < 0.0005$), and 2D compared to hybrid ($M = 44.7$, $SE = 11.67\,seconds$, $p = 0.012$). However, task completion time was statistically significantly higher in 3D compared to hybrid ($M = 68.6$, $SE = 15.68\ seconds$, $p = 0.005$).

The completion time of students, when compared to residents, was statistically significantly higher using the 2D interface ($26 \pm 10.15\,seconds$, $t(27) = 2.563$, $p = 0.016$, $d = 1.0$), but was statistically significantly lower using the 3D interface ($111.51 \pm 17.38\,seconds$, $t(27) = 6.415$, $p < 0.0005$, $d = 2.51$), and hybrid interface ($69.96 \pm 11.19\,seconds$, $t(27) = 6.255$, $p < 0.0005$, $d = 2.44$).
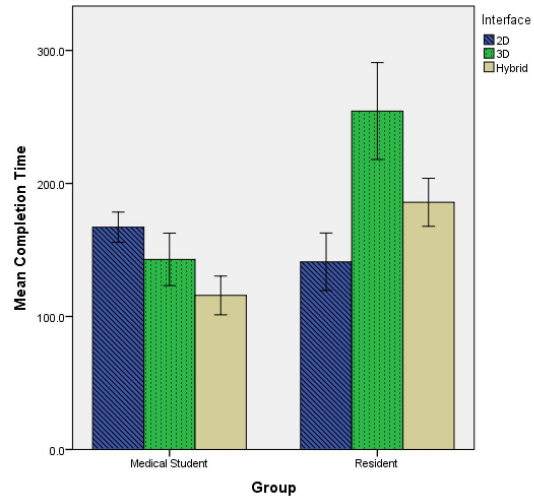
Fig. 9 Mean Completion Time

4.2.2 Effects of interface and group on Absolute Error

For the student group, the absolute error was statistically significantly higher using 2D compared to 3D ($M = 4.29$, $SE = 0.91°$, $p = 0.001$) and 2D compared to hybrid ($M = 4.789$, $SE = 0.86°$, $p < 0.0005$) but the absolute error was not statistically significantly different between using 3D and hybrid interfaces ($M = 0.502$, $SE = 0.743°$, $p = 1.0$). For the resident group, the absolute error was statistically significantly higher using 2D compared to hybrid ($M = 1.97$, $SE = 0.50°$, $p = 0.01$), but the absolute error was not statistically significantly different between 2D compared to 3D ($M = 0.91$, $SE = 0.47°$, $p = 0.257$) and 3D compared to hybrid ($M = 1.06$, $SE = 0.44°$, $p = 0.117$).
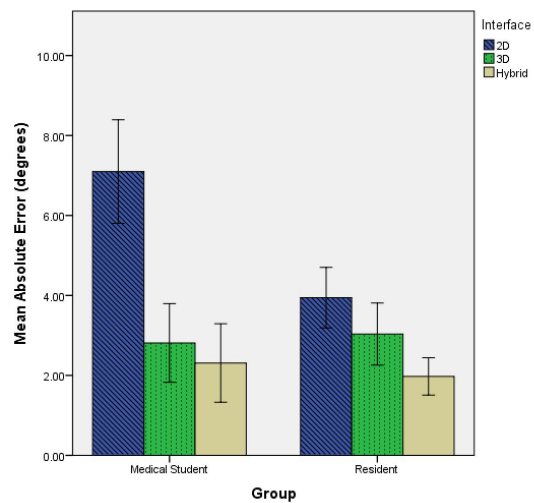
Fig. 10 Mean Absolute Error

The absolute error of students, when compared to residents, was statistically significantly higher using the 2D interface ($3.15 \pm 0.70°$, $t(25.7) = 4.494$, $p = 0.001$, $d = 1.38$). There was no statistically significant difference in absolute error between students and residents using the 3D interface ($-0.22 \pm 0.7°$, $t(27) = -0.320$, $p = 0.751$, $d = 0.12$), and the hybrid interface ($0.34 \pm 0.51°$, $t(23.901) = 0.657$, $p = 0.517$, $d = 0.2$).

4.2.3 Effects of interface and group on SUS score

For the student group, the SUS score was statistically significantly lower for 2D compared to 3D ($M = 14.08$, $SE = 3.87$, $p = 0.006$) and for 2D compared to hybrid ($M = 13.29$, $SE = 4.45$, $p = 0.024$) but the SUS score was not statistically significantly different between 3D and hybrid interfaces ($M = 0.79$, $SE = 1.47$, $p = 1.0$). For the resident group, the SUS score was statistically significantly higher for 2D compared to 3D ($M = 11.8$, $SE = 3.25$, $p = 0.016$) and for 2D compared to hybrid ($M = 6.8$, $SE = 1.79$, $p = 0.013$) but the SUS score was not statistically significantly different between 3D and hybrid interfaces ($M = 5.0$, $SE = 2.44$, $p = 0.213$).
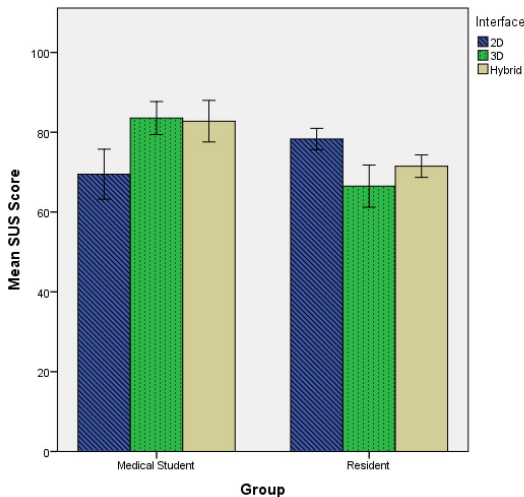
Fig. 11 Mean SUS Score : Higher score represents better usability.

The SUS score of students, when compared to residents, was statistically significantly lower for the 2D interface (8.83 ± 4.23, $t(27) = 2.085$, $p = 0.047$, $d = 0.82$), but was statistically significantly higher for the 3D interface (17.05 ± 3.2, $t(27) = 5.336$, $p < 0.0005$, $d = 2.08$), and hybrid interface (11.26 ± 3.56, $t(27) = 3.166$, $p = 0.004$, $d = 1.24$).

4.2.4 Effects of interface and group on Physical Task Load
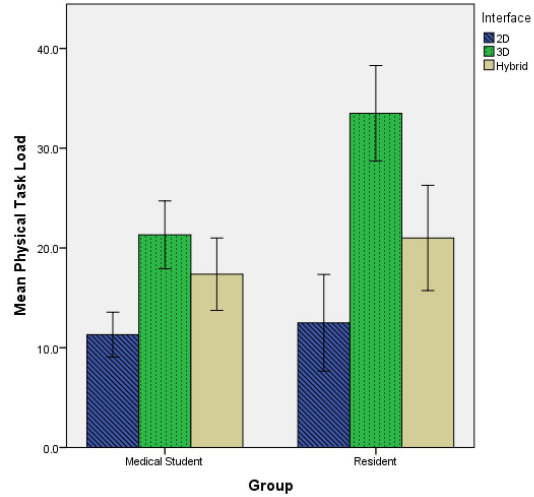
Fig. 12 Mean Physical Task Load

For the student group, the physical load was statistically significantly lower for the 2D interface compared to 3D ($M = 10.0$, $SE = 1.83$, $p < 0.0005$) and for 2D compared to hybrid ($M = 6.05$, $SE = 2.25$, $p = 0.045$) but the physical task load was not statistically significantly different between 3D and hybrid interfaces ($M = 3.947$, $SE = 2.08$, $p = 0.221$). For the residents group, the physical task load was statistically significantly lower for the 2D interface compared to 3D ($M = 21.0$, $SE = 3.06$, $p < 0.0005$) and for 2D compared to hybrid ($M = 8.5$, $SE = 2.59$, $p = 0.028$) but the physical task load was statistically significantly higher for the 3D interface compared to hybrid ($M = 12.5$, $SE = 2.5$, $p = 0.002$).

The physical task load for students, when compared to residents, was statistically significantly lower for the 3D interface (12.18 ± 2.71, $t(27) = 4.504$, $p < 0.0005$, $d = 1.76$). There was no statistically significant difference in physical task load between students and residents for the 2D (1.18 ± 2.13, $t(27) = 0.555$, $p = 0.583$, $d = 0.22$), and hybrid (3.63 ± 2.92, $t(27) = 1.244$, $p = 0.224$, $d = 0.49$) interfaces.

4.2.5 Mode changes with the 3D interface

We compared the average mode changes using the 3D interface between students($M = 4.58$, $SD = 1.07$) and residents($M = 7.5$, $SD = 3.28$) by running the independent samples t-test. The number of mode changes by residents was statistically significantly higher than

students (2.92 ± 0.81, $t(27) = 3.590$, $p = 0.001$, $d = 1.40$).

4.2.6 Age Correlations with 3D interface

We found a strong positive correlation between age and completion time using the 3D interface ($r(29) = 0.715$, $p < 0.0005$). We also found a strong negative correlation between age and SUS score for the 3D interface ($r(29) = -0.623$, $p < 0.0005$). However, there was no statistically significant correlation between age and task accuracy ($r(29) = -0.027$, $p = 0.889$).

4.2.7 Participant feedback from interview

All participants appreciated the availability of 3D model, as it offered a quick and complete view of the patient's anatomy. Most medical students reported feeling more confident while performing the diagnosis task in the 3D and hybrid conditions as the 3D model helped them orient the slice quickly. They also reported that 3D rotation of the arbitrary plane using the 2D mouse was difficult and that the guidelines on the other orthogonal planes were hard to follow and confusing.

The residents reported feeling more confident in the 2D and hybrid conditions, as they found it easy to use the 2D slice guidelines to verify the integrity of their slice orientation. All of them reported difficulty performing precise annotations and measurements with the stylus in the 3D condition. They suggested that some form of stabilization, or scaling, was essential for using the stylus.

Most participants felt that the mouse was more accurate for annotation and measurement steps. All participants felt that learning to use the stylus was not only very natural but straightforward and easy. They reported finding it difficult to annotate and measure using the 3D interface due to unsteady hands while using the stylus. A few participants mentioned that the limited virtual stylus length forced them to reach further, and thus made them experience more hand fatigue.

4.2.8 Observations

The stylus seemed very intuitive as no participant had any difficulty maneuvering the stylus. In the 3D condition, we observed that some participants found it difficult to perform precise tasks such as annotations and measurements using the stylus. To improve hand stability, most participants rested their elbow on the table, and some participants tried using their non-dominant hand to support their dominant hand.

All participants seemed to enjoy the stylus input and stereoscopic display. Near the end of their practice session, they often reloaded the dataset and explored the anatomical features of the 3D model using the stylus, in 3D and hybrid conditions. While in the 2D condition, they would proceed to start the evaluation task. Some participants even requested additional time at the end of their experiment, to further explore the 3D model.

In the 2D condition, medical students seemed confused by the slice guidelines on other orthogonal planes. They found it difficult to make precise 3D rotations using these guidelines. Medical students appeared more comfortable using the 3D interface compared to residents. The resident group seemed to have difficulty remembering the location and functions of 3D buttons in the 3D interface. They referred to the 3D control reference card a lot more frequently compared to medical students.

4.3 Discussion and Conclusion

Each participant only received five minutes of training followed by five minutes of practice time to familiarize themselves with the interface. Despite this, they were able to complete the task in less than five minutes with any interface. This showed that not much training was required to learn the interface and tools necessary for performing the diagnosis task.

The fastest condition for students was the hybrid interface followed by the 3D and 2D interfaces. They were also more accurate with the hybrid and 3D interfaces compared to 2D. The medical students' slow and inaccurate performance with the 2D interface could probably be attributed to their inexperience with 2D tools, and the difficulty of performing 3D rotation of the slice plane using the 2D mouse input.

Using the hybrid interface, students were able to achieve the same performance as the more experienced users. Since they had no prior experience with any interface, it shows that they were not only able to learn our hybrid interface quickly, but also to use it very efficiently. We believe this has major implications for improving diagnostic radiology training. While their accuracy using the 2D interface was low, using our hybrid interface, their accuracy improved significantly and was even comparable to that of the residents.

The fastest condition for residents was the 2D interface followed by the hybrid and 3D interfaces. However, much to our surprise, the residents were most accurate using the hybrid interface followed by 3D and 2D interfaces. Since the residents had been using the 2D interface daily for over a year, it is no surprise that they were

fastest with the 2D interface and significantly outperformed the students with it. Their expertise with the 2D interface made them faster with our hybrid interface as well, since annotation and measurement were still performed using the 2D tools.

The lower accuracy of residents using the 2D interface despite their prior experience could be attributed to the difficulty in orienting the arbitrary slice using the 2D mouse input. Such difficulty can result in a habit of considering measurements with potentially lower accuracy, as acceptable.

By providing an easier way to rotate the arbitrary slice in our hybrid interface, despite the habit of "rough measurements", we observe that the accuracy has increased. This is a significant finding. However, it should be noted that the cause could also be the novelty of our hybrid interface, leading to more attention from the residents during the diagnosis task. A future study looking at the use of our interface over multiple sessions could help identify the true cause of this effect.

We observed that residents found it difficult to cope with the different task workflow and relied more heavily on the interface reference cards. Our results show that using the 3D interface, on average, residents made significantly more mode changes than the medical students. This indicated that they found it difficult to follow the task workflow in the 3D condition. The residents did not have this problem in the hybrid condition, as it did not involve any 3D mode changes.

The 3D stylus had a one to one mapping of hand motion. While this was very natural for exploring the 3D model, precise movements were difficult. Unsteady hands made it even harder to perform precise interactions. This explains why participants experienced difficulties with annotation and measurements in the 3D condition. The stylus interaction also demanded higher physical hand movement compared to the 2D mouse. This explains the higher physical task load score for the 3D and hybrid interfaces.

We found a strong positive correlation between age and task completion time, with older participants taking longer to complete the task in the 3D condition. This behavior is similar to that observed by Zudilova et al [61] in their study, although their evaluation task was quite different. We also found a strong negative correlation between age and the 3D SUS score, with older users rating the 3D interface lower. However, their accuracy scores were not affected by age.

The resident group was the older user group, with age ranging from 28 to 38 years 4.1.1. Hence, we believe that the observed correlation is a result of their prior extensive experience with Inteleviewer's 2D interface.

It is hard to unlearn certain behaviors, that are similar in context but differ in interaction.

Feedback from the subjects in the unstructured interview, showed that most participants preferred the hybrid interface among the three conditions, since it combined the 3D model (showing the patient anatomy), the stylus input for 3D slice manipulation, the synchronized 2D view for verification, and the familiar and precise 2D tools (for annotations and measurements).

The medical students, in particular, appreciated having the 3D model, since it helped them better understand the patient's anatomy, compared to the 2D interface. This is likely the reason why they gave the lowest SUS score for the 2D interface and a relatively high SUS score for the 3D and hybrid interfaces despite the stylus precision issues. The residents gave the lowest SUS score to the 3D interface due to the change in task workflow, and stylus precision problems for the annotations and measurements. Their SUS score for the 2D interface was higher, but this was expected due to their familiarity with the system.

A baseline SUS score of 68 is often used to determine major usability issues with a user interface [62]. All the SUS scores we obtained were higher than this baseline score, within the margin of error. This shows that there were no major usability issues with any interface.

The 3D model was intended to give an overview of the anatomy for inexperienced users so that they could quickly and easily understand the position and orientation of a slice, relative to the entire scan. Although we used a mesh for visualizing the anatomy, other volume rendering techniques could be used to better represent soft tissues.

The stereoscopic 3D is a key component of our 3D interface. It was required to simulate the holographic 3D experience by rendering the 3D model in the negative parallax region. We made sure the scene contrast, eye separation, and scene depth were within the optimal range[63]. Since the maximum task completion time did not exceed 5 minutes, we believe that it is less likely that users experienced any ill effects from stereoscopic 3D. We did not get any feedback from the users, that would suggest otherwise.

4.3.1 Limitations

There were a limited number of radiology residents, hence the sample size for the resident group was relatively small. In the 3D condition, the lack of stylus stabilization, and the fixed virtual stylus length appeared to most negatively impacted task performance and contributed to increased fatigue.

For our prototype setup, a standard 2D screen was used for the 2D interface. We did not use a specialized 12-bit gray-scale radiology calibrated screen. Since our evaluation task only involved looking at bones, we believe that our standard 2D screen had sufficient contrast to perform the task. In the 2D interface, the users were required to hold the [shift] key in order to rotate the slice. This might have resulted in a minor disadvantage for the 2D condition.

In diagnostic radiology, the need for 3D measurements is rare compared to 2D measurements. Hence tasks that require 3D measurements are perceived as very specialized. Difficulty in performing 3D measurements can lead to simplified surrogate measurements being obtained. For example, abdominal aortic aneurysm maximum diameter measurements are performed horizontally in the transverse plane rather than representing the true maximum aneurysm diameter, introducing a small potential measurement error. Thus by making 3D manipulation and measurement easy, we can avoid the need for such surrogate measurements in diagnostic radiology.

The radiology diagnosis task used for our evaluation is a specialized case. We chose a diagnosis task that required some level of 3D manipulation to fairly evaluate our system, while still being easy enough for medical students. Only a limited number of diagnostic radiology tasks require 3D manipulation.

Fatigue from long diagnosis tasks can lead to mistakes in diagnosis. However, we believe that this might not be the case in our study, since the maximum task completion time for any of the interface conditions, did not exceed 5 minutes. Hence, the total task time for all three tasks was under 15 minutes. The effects of fatigue can be explored in a future study, by choosing a relatively complex diagnosis task and measuring the fatigue of subjects, using the full NASA TLX survey[58].

One of the issues mentioned was the mapping of buttons on the 3D stylus. While some participants were satisfied with the way stylus buttons were mapped, others preferred that the left and right buttons on the stylus to be mapped similar to a mouse, even though this was not ergonomic. The left button on the stylus was harder to reach with the index finger while holding the stylus. Despite this, participants preferred the most common functions to be mapped to the left stylus button, as this would be similar to the traditional mouse (where most functions are performed with the left mouse button).

4.3.2 Future Work

The 3D stylus precision can be improved by using some form of scaled manipulation such as PRISM[64] introduced by Frees et al. The speech input could be explored to improve the mode switching. Interaction with the 3D components could be improved by exploring other forms of 3D input such as freehand input.

Further studies can be run with different diagnostic radiology tasks, possibly with a larger pool of residents to observe differences in performance. Additional experimental data could be captured such as hand and head tracking to study user behavior in more detail. More subjective feedback could be gathered for future studies about the individual components within the hybrid interface. It would be interesting to explore effects (especially fatigue) while using the interface for a longer period of time.

Users would be expected to perform even better with additional training with the hybrid interface. It would be interesting to explore long-term learning effects.

4.3.3 Summary

We introduced a hybrid interface for improving 3D manipulation in radiological diagnosis. This interface combined the zSpace stereoscopic system with a 2D display, and mouse and keyboard input. We also presented an evaluation involving a user study diagnosing scoliosis for three conditions with two groups. The study results show that the hybrid interface allows users to achieve higher accuracy in tasks involving 3D rotation of anatomical planes, compared to the traditional 2D interface.

Users were able to learn all the interfaces (2D, 3D, and hybrid) after a five-minute training session and were later able to perform the scoliosis diagnosis task. Compared to the 2D interface, the novice users were able to perform the task faster and with a significant increase in accuracy using our hybrid interface. The experienced users were slightly slower using our hybrid interface, but their diagnosis was more accurate.

References

1. M. Chen, S. J. Mountford, and A. Sellen, "A study in interactive 3-D rotation using 2-D control devices," *ACM SIGGRAPH Computer Graphics*, vol. 22, no. 4, pp. 121–129, 1988.
2. K. Hinckley, J. Tullio, R. Pausch, D. Proffitt, and N. Kassell, "Usability analysis of 3D rotation techniques," in

Proceedings of the 10th annual ACM symposium on User interface software and technology, pp. 1–10, ACM, 1997.

3. D. Bowman, E. Kruijff, J. J. LaViola Jr, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice, CourseSmart eTextbook*. Addison-Wesley, 2004.

4. M. J. Graves, R. T. Black, and D. J. Lomas, "Constrained Surface Controllers for Three-dimensional Image Data Reformatting," *Radiology*, vol. 252, no. 1, pp. 218–224, 2009.

5. "zSpace zSpace, inc.." `https://zspace.com/`. Accessed: 2017-01-09.

6. T. Emerson, J. D. Prothero, and S. J. Weghorst, *Medicine and virtual reality: a guide to the literature (MedVR)*. Human Interface Technology Laboratory, 1994.

7. N. Ayache, "Medical computer vision, virtual reality and robotics," *Image and Vision Computing*, vol. 13, no. 4, pp. 295–313, 1995.

8. G. Székely and R. M. Satava, "Virtual reality in medicine," *BMJ: British Medical Journal*, vol. 319, no. 7220, p. 1305, 1999.

9. L. Gallo, A. Minutolo, and G. De Pietro, "A user interface for VR-ready 3D medical imaging by off-the-shelf input devices," *Computers in Biology and Medicine*, vol. 40, no. 3, pp. 350–358, 2010.

10. C. Hand, "A survey of 3D interaction techniques," in *Computer graphics forum*, vol. 16(5), pp. 269–281, Wiley Online Library, 1997.

11. K. Shoemake, "Arcball: a user interface for specifying three-dimensional orientation using a mouse," in *Graphics Interface*, vol. 92, pp. 151–156, 1992.

12. K. Henriksen, J. Sporring, and K. Hornbæk, "Virtual trackballs revisited," *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, no. 2, pp. 206–216, 2004.

13. R. Bade, F. Ritter, and B. Preim, "Usability comparison of mouse-based interaction techniques for predictable 3D rotation," in *International Symposium on Smart Graphics*, pp. 138–150, Springer, 2005.

14. K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell, "Passive real-world interface props for neurosurgical visualization," in *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, pp. 452–458, ACM, 1994.

15. B. Frohlich, J. Plate, J. Wind, G. Wesche, and M. Gobel, "Cubic-mouse-based interaction in virtual environments," *Computer Graphics and Applications, IEEE*, vol. 20, no. 4, pp. 12–15, 2000.

16. L. Gallo, G. De Pietro, and I. Marra, "3D interaction with volumetric medical data: experiencing the wiimote," in *Proceedings of the 1st international conference on Ambient media and systems*, p. 14, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.

17. S. Mauser and O. Burgert, "Touch-free, gesture-based control of medical devices and software based on the leap motion controller," *Stud Health Technol Inform*, vol. 196, pp. 265–270, 2014.

18. L. Gallo, A. P. Placitelli, and M. Ciampi, "Controller-free exploration of medical image data: Experiencing the Kinect," in *Computer-based medical systems (CBMS), 2011 24th international symposium on*, pp. 1–6, IEEE, 2011.

19. G. C. S. Ruppert, L. O. Reis, P. H. J. Amorim, T. F. de Moraes, and J. V. L. da Silva, "Touchless gesture user interface for interactive image visualization in urological surgery," *World journal of urology*, vol. 30, no. 5, pp. 687–691, 2012.

20. R. Balakrishnan, T. Baudel, G. Kurtenbach, and G. Fitzmaurice, "The Rockin'Mouse: integral 3D manipulation on a plane," in *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, pp. 311–318, ACM, 1997.

21. N. T. Dang, M. Tavanti, I. Rankin, and M. Cooper, "A comparison of different input devices for a 3D environment," *International Journal of Industrial Ergonomics*, vol. 39, no. 3, pp. 554–563, 2009.

22. E. V. Zudilova-Seinstra, P. J. de Koning, A. Suinesiaputra, B. W. van Schooten, R. J. van der Geest, J. H. Reiber, and P. M. Sloot, "Evaluation of 2D and 3D glove input applied to medical image analysis," *International journal of human-computer studies*, vol. 68, no. 6, pp. 355–369, 2010.

23. F. Bérard, J. Ip, M. Benovoy, D. El-Shimy, J. R. Blum, and J. R. Cooperstock, "Did "Minority Report" get it wrong? Superiority of the mouse over 3D input devices in a 3D placement task," in *IFIP Conference on Human-Computer Interaction*, pp. 400–414, Springer, 2009.

24. G. Wang, M. J. McGuffin, F. Bérard, and J. R. Cooperstock, "Pop-up depth views for improving 3D target acquisition," in *Proceedings of Graphics Interface 2011*, pp. 41–48, Canadian Human-Computer Communications Society, 2011.

25. S. Feiner and A. Shamash, "Hybrid user interfaces: Breeding virtually bigger interfaces for physically smaller computers," in *Proceedings of the 4th annual ACM symposium on User interface software and technology*, pp. 9–17, ACM, 1991.

26. G. W. Fitzmaurice, S. Zhai, and M. H. Chignell, "Virtual reality for palmtop computers," *ACM Transactions on Information Systems (TOIS)*, vol. 11, no. 3, pp. 197–218, 1993.

27. I. G. Angus and H. A. Sowizral, "Embedding the 2D interaction metaphor in a real 3D virtual environment," in *IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology*, pp. 282–293, International Society for Optics and Photonics, 1995.

28. M. Hachet, P. Guitton, and P. Reuter, "The CAT for efficient 2D and 3D interaction as an alternative to mouse adaptations," in *Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 225–112, ACM, 2003.

29. R. P. Darken and R. Durost, "Mixed-dimension interaction in virtual environments," in *Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 38–45, ACM, 2005.

30. J. Wang and R. W. Lindeman, "Object impersonation: Towards effective interaction in tablet-and HMD-based hybrid virtual environments," in *Virtual Reality (VR), 2015 IEEE*, pp. 111–118, IEEE, 2015.

31. M. Wloka, "Interacting with virtual reality," in *Virtual Prototyping*, pp. 199–212, Springer, 1995.

32. K. Coninx, F. Van Reeth, and E. Flerackers, "A hybrid 2D/3D user interface for immersive object modeling," in *Computer Graphics International, 1997. Proceedings*, pp. 47–55, IEEE, 1997.

33. J. Rekimoto, "Pick-and-drop: a direct manipulation technique for multiple computer environments," in *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pp. 31–39, ACM, 1997.

34. B. Ullmer and H. Ishii, "The metaDESK: models and prototypes for tangible user interfaces," in *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pp. 223–232, ACM, 1997.

35. G. Riva, "Virtual reality for health care: the status of research," *Cyberpsychology & Behavior*, vol. 5, no. 3, pp. 219–225, 2002.

36. T. N. Arvanitis, "Virtual reality in medicine," in *Handbook of Research on Informatics in Healthcare and Biomedicine*, pp. 59–67, IGI Global, 2006.

37. C. Pensieri and M. Pennacchini, "Virtual Reality in Medicine," in *Handbook on 3D3C Platforms*, pp. 353–401, Springer, 2016.

38. S. Baumgärtner, A. Ebert, M. Deller, and S. Agne, "2D meets 3D: a human-centered interface for visual data exploration," in *CHI'07 Extended Abstracts on Human Factors in Computing Systems*, pp. 2273–2278, ACM, 2007.

39. A. Bornik, R. Beichel, E. Kruijff, B. Reitinger, and D. Schmalstieg, "A hybrid user interface for manipulation of volumetric medical data," in *3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on*, pp. 29–36, IEEE, 2006.

40. M. Teistler, R. Breiman, T. Lison, O. Bott, D. Pretschner, A. Aziz, and W. Nowinski, "Simplifying the exploration of volumetric images: development of a 3D user interface for the radiologists workplace," *Journal of digital imaging*, vol. 21, no. 1, pp. 2–12, 2008.

41. M. Teistler, G. Ampanozi, W. Schweitzer, P. Flach, M. Thali, and L. Ebert, "Use of a low-cost three-dimensional gaming controller for forensic reconstruction of CT images," *Journal of Forensic Radiology and Imaging*, vol. 7, pp. 10–13, 2016.

42. R. Aamir, A. Chernoglazov, C. Bateman, A. Butler, P. Butler, N. Anderson, S. Bell, R. Panta, J. Healy, J. Mohr, *et al.*, "MARS spectral molecular imaging of lamb tissue: data collection and image analysis," *Journal of Instrumentation*, vol. 9, no. 02, p. P02005, 2014.

43. K. Rajendran, M. Walsh, N. De Ruiter, A. Chernoglazov, R. Panta, A. Butler, P. Butler, S. Bell, N. Anderson, T. Woodfield, *et al.*, "Reducing beam hardening effects and metal artefacts in spectral CT using Medipix3RX," *Journal of Instrumentation*, vol. 9, no. 03, p. P03015, 2014.

44. K. Rajendran, C. Löbker, B. S. Schon, C. J. Bateman, R. A. Younis, N. J. de Ruiter, A. I. Chernoglazov, M. Ramyar, G. J. Hooper, A. P. Butler, *et al.*, "Quantitative imaging of excised osteoarthritic cartilage using spectral CT," *European radiology*, vol. 27, no. 1, pp. 384–392, 2017.

45. "DCMTK OFFIS DICOM toolkit." `http://dicom.offis.de/dcmtk.php.en`. Accessed: 2017-01-12.

46. "Qt API Th Qt Company." `https://www.qt.io/ui/`. Accessed: 2017-01-12.

47. "OpenGL API Khronos Group." `https://www.opengl.org/`. Accessed: 2017-01-12.

48. "CUDA API NVIDIA Corporation." `https://developer.nvidia.com/cuda-toolkit`. Accessed: 2017-01-12.

49. "Inteleviewer Intelerad Medical Systems Incorporated." `http://www.intelerad.com/en/products/inteleviewer/`. Accessed: 2017-01-09.

50. W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," in *ACM Siggraph Computer Graphics*, vol. 21, pp. 163–169, ACM, 1987.

51. Y. Dai, J. Zheng, Y. Yang, D. Kuai, and X. Yang, "Volume-Rendering-Based Interactive 3D Measurement for Quantitative Analysis of 3D Medical Images," *Computational and mathematical methods in medicine*, vol. 2013, 2013.

52. B. Preim, C. Tietjen, W. Spindler, and H. O. Peitgen, "Integration of measurement tools in medical 3D visualizations," in *Proceedings of the conference on Visualization'02*, pp. 21–28, IEEE Computer Society, 2002.

53. "Human Ethics Committee university of canterbury." `http://www.canterbury.ac.nz/humanethics/`. Accessed: 2016-06-01.

54. W. E. Brant and C. A. Helms, *Fundamentals of diagnostic radiology*. Lippincott Williams & Wilkins, 2012.

55. R. A. Drebin, L. Carpenter, and P. Hanrahan, "Volume rendering," in *ACM Siggraph Computer Graphics*, vol. 22(4), pp. 65–74, ACM, 1988.

56. G. M. Cox and W. Cochran, *Experimental designs*. JSTOR, 1953.

57. J. Brooke *et al.*, "SUS-A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

58. S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," *Advances in psychology*, vol. 52, pp. 139–183, 1988.

59. S. E. Maxwell and H. D. Delaney, *Designing experiments and analyzing data: A model comparison perspective*, vol. 1. Psychology Press, 2004.

60. J. Hartung, G. Knapp, and B. K. Sinha, *Statistical meta-analysis with applications*, vol. 738. John Wiley & Sons, 2011.

61. E. Zudilova-Seinstra, B. van Schooten, A. Suinesiaputra, R. van der Geest, B. van Dijk, J. Reiber, and P. Sloot, "Exploring individual user differences in the 2D/3D interaction with medical image data," *Virtual Reality*, vol. 14, no. 2, pp. 105–118, 2010.

62. A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale," *Journal of usability studies*, vol. 4, no. 3, pp. 114–123, 2009.

63. R. E. Patterson, *Human factors of stereoscopic 3D displays*. Springer, 2015.

64. S. Frees and G. D. Kessler, "Precise and rapid interaction through scaled manipulation in immersive virtual environments," in *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pp. 99–106, IEEE, 2005.

65. "MARS Bioimaging Ltd." `http://www.marsbioimaging.com/mars/`. Accessed: 2017-01-09.